

# **ESCUELA POLITÉCNICA NACIONAL**

**FACULTAD DE CIENCIAS**

**CONSTRUCCIÓN DE UN MODELO ESTADÍSTICO PARA  
CALCULAR EL RIESGO DE DETERIORO DE UNA CARTERA DE  
MICROCRÉDITOS Y PROPUESTA DE UN SISTEMA DE GESTIÓN  
PARA LA RECUPERACIÓN DE LA CARTERA EN UNA EMPRESA  
DE COBRANZAS**

**PROYECTO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE  
INGENIERO MATEMÁTICO**

**MARCO SANTIAGO JÁCOME JARA**

mjacome10@hotmail.com

**Director: Holger Efraín Benalcázar Paladines**

holger.benalcazar@epn.edu.ec

**DICIEMBRE 2014**



## DECLARACIÓN

Yo, Marco Santiago Jácome Jara, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

---

**Marco Santiago Jácome Jara**

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por Marco Santiago Jácome Jara, bajo mi supervisión.

---

**Holger Benalcázar Paladines**  
**DIRECTOR**

## **AGRADECIMIENTOS**

A la Escuela Politécnica Nacional, formadora desde siempre de profesionales con altísimo nivel para los mundos, comprometida con el desarrollo científico y productivo del Ecuador.

Al MSc. Holger Benalcázar, maestro y director de tesis, de quien recibí conocimiento para las aulas y fuera de ellas. Sus enseñanzas siguen vigentes, útiles y eficaces para enfrentar cada vez retos más grandes.

A la Facultad de Ciencias, a todos mis maestros y amigos; a los compañeros de trabajo; a mi querida familia, sobre todo a mis padres Marco y Luz: papá, me gustaría mucho tener varias de tus virtudes y ser más como quisieras que yo sea; mamá, tu apoyo y afecto siempre me han dado seguridad y bienestar para salir adelante; estoy seguro que eres la última persona que me fallaría. A mis hermanas, tía, prima, y en general a todo aquello que confabuló para llevar a feliz término este trabajo. Gracias totales!

## DEDICATORIA

Es obvio y necesario, dada mi naturaleza, que mis más grandes esfuerzos, mis más encarnizadas luchas, mis más sublimes conquistas, se dediquen a eso que es lo más grande en la vida de un hombre como yo: la mujer. Cito aquí un pequeño extracto que expresa de mejor forma lo que quiero decir:

*“Quien adora a Dios en la mujer no necesita ir a ningún templo.  
Quien no ama a la mujer, no sabe amar a Dios. Dios quiere lo que la mujer quiere.  
El hombre es mente que piensa, la mujer es intuición que inspira.  
Él: la fuerza, Ella: el consejo y la previsión; la fuerza vence, la sabiduría convence.  
Las gotas de rocío vivifican los pétalos de la marchita rosa y las lágrimas de la mujer resucitan las cualidades muertas en el corazón del hombre.  
Si eres hombre, debes divinizarte por la mujer; si eres Dios debes humanizarte por ella. Ella es el camino en la ida y en la vuelta.  
En las tribulaciones, el hombre se inclina a tomar por confidente a una mujer, y ésta, con su poder espiritual será el mejor consejero y enfermero; sus palabras son fuente de alivio, sus manos, manantial de salud, y de su encantadora sonrisa emana el valor.  
La mujer es el pensamiento más hermoso del Absoluto, que debe ser captada por la inteligencia, y no vista por los ojos.  
Dios es una palabra misteriosa y la mujer es su significado...”*

*Jorge Adoum*

En base a esto mi dedicatoria es a ustedes: Luz, Rosa, Guadalupe, Margarita, Steffani: mi hermosa madre, mi fuerte abuela paterna, mi inteligente abuela materma, la madre de mi hijo, mi gran amor del pasado.

## ÍNDICE DE CONTENIDO

LISTA DE FIGURAS .....	IX
LISTA DE TABLAS .....	X
LISTA DE ANEXOS .....	XIII
RESUMEN .....	XIV
ABSTRACT .....	XV
CAPÍTULO I .....	1
EL PROCESO DE COBRANZA DE UNA CARTERA DE MICROCRÉDITOS EN LA EMPRESA ABC .....	1
1.1 LAS MICROFINANZAS Y LAS MICROEMPRESAS .....	2
1.1.1 CONCEPTO DE MICROFINANZAS .....	3
1.1.2 IMPORTANCIA DE LA MICROEMPRESA .....	3
1.1.3 MARCO REGULATORIO Y LEGISLACIÓN RELACIONADA CON LA MICROEMPRESA.....	8
1.1.4 INDICADORES DE LAS MICROFINANZAS EN EL ECUADOR.....	9
1.2 PROCESO GENERAL DE COBRANZA EN ABC.....	10
1.2.1 EL SERVICIO DE COBRANZA.....	12
1.2.2 EL SISTEMA DE COBRANZA.....	14
1.2.3 LAS HERRAMIENTAS DE COBRANZA.....	14
1.3 SITUACIÓN ACTUAL DEL PROCESO.....	15
1.3.1 INGRESO Y DISPONIBILIDAD DE LA CARTERA A COBRAR .....	18
1.3.2 PLANIFICACIÓN DE LA COBRANZA EN ABC .....	19
1.3.3 FIDELIZACIÓN CON EL CEDENTE.....	21
1.3.4 GESTIÓN DE LA DEMANDA AL ÁREA DE SISTEMAS .....	22
1.3.5 PRODUCTIVIDAD Y CONTROL DE CALIDAD .....	22
1.3.6 GESTIÓN FINANCIERA ADMINISTRATIVA .....	23
1.4 SITUACIÓN ACTUAL DE LOS SUBPROCESOS .....	24
1.4.1 SUBPROCESO: GESTIÓN TELEFÓNICA .....	26
1.4.2 SUBPROCESO: GESTIÓN DOMICILIARIA O DE TERRENO .....	27

1.4.3	SUBPROCESO: GESTIÓN CON MEDIOS MASIVOS .....	28
1.5	PLANTEAMIENTO DEL PROBLEMA.....	29
1.5.1	FORMULACIÓN Y SISTEMATIZACIÓN DEL PROBLEMA .....	30
1.5.2	OBJETIVOS DE LA INVESTIGACIÓN .....	31
1.5.3	JUSTIFICACIÓN DEL PROYECTO .....	32
	CAPÍTULO II.....	34
	CONSTRUCCIÓN DE MODELOS ESTADÍSTICOS PARA CALCULAR EL RIESGO DE DETERIORO DE LA CARTERA MICROCRÉDITO EN ABC.....	34
2.1	MINERÍA DE DATOS .....	35
2.1.1	FILTRADO DE DATOS.....	35
2.1.2	SELECCIÓN DE VARIABLES .....	43
2.1.3	EXTRACCIÓN DEL CONOCIMIENTO.....	47
2.1.4	INTERPRETACIÓN Y EVALUACIÓN.....	48
2.2	CONSTRUCCIÓN DE MODELOS UTILIZANDO REGRESIÓN LOGÍSTICA CLÁSICA.....	49
2.2.1	SELECCIÓN DEL MEJOR MODELO .....	51
2.2.2	VALIDACIÓN DEL MODELO .....	52
2.2.3	VALIDACIÓN CON LA MUESTRA DE PRUEBA .....	55
2.2.4	PRESENTACIÓN DE RESULTADOS DEL MODELO LOGÍSTICO.....	56
2.3	MEJORAMIENTO DEL MODELO LOGÍSTICO UTILIZANDO ÁRBOLES DE DECISIÓN.....	57
2.3.1	CONSTRUCCIÓN DEL ÁRBOL.....	58
2.3.2	VALIDACIÓN SOBRE LA MUESTRA DE PRUEBA.....	60
2.4	LA DISTANCIA EN ESTADÍSTICA .....	61
2.4.1	APLICACIÓN DE LA REGRESIÓN BASADA EN DISTANCIAS PARA CALCULAR EL RIESGO DE CONVERTIRSE EN UN MAL CLIENTE DENTRO DE UNA EMPRESA DE COBRANZAS. ....	64
2.4.2	SISTEMAS DE CLASIFICACION BASADOS EN DISTANCIAS .....	65
2.5	ANALISIS COMPARATIVO DE MODELOS.....	71
2.5.1	CURVAS ROC.....	72

CAPÍTULO III .....	74
PROPUESTA DE UN NUEVO SISTEMA DE COBRANZA PARA LA CARTERA MICROCRÉDITO EN ABC .....	74
3.1 EVALUACIÓN DE FACTORES QUE AFECTAN LA RECUPERACIÓN DE LA CARTERA .....	74
3.1.1 FACTORES DE CONFUSIÓN .....	76
3.1.2 FACTORES DE INTERACCIÓN .....	77
3.1.3 VALORACIÓN DE LA CONFUSIÓN E INTERACCIÓN.....	77
3.2 MODELOS ESTADÍSTICOS BASADOS EN SCORE.....	87
3.2.1 CONSTRUCCIÓN DE LA TABLA DE DESEMPEÑO.....	88
3.2.2 CREACIÓN DE ESTRATEGIAS Y ACCIONES DE COBRANZA .....	92
3.2.3 CRITERIOS UTILIZADOS PARA SEGMENTAR LA CARTERA EN BASE A ESTRATEGIAS.....	93
3.2.4 CRITERIOS UTILIZADOS PARA DETERMINAR LAS ACCIONES DE COBRANZA PARA CADA ESTRATEGIA.....	96
3.2.5 COMPARACIÓN DE COSTOS ENTRE EL MODELO TRADICIONAL DE COBRANZA Y EL NUEVO SISTEMA DE GESTION BASADO EN SOCRE PARA LA EMPRESA ABC.....	99
CAPÍTULO IV .....	111
IMPLEMENTACIÓN DEL NUEVO SISTEMA DE GESTIÓN DE COBRANZA EN ABC.....	111
4.1 MONITOREO DEL MODELO ESTADÍSTICO .....	111
4.2 ACCIONES Y ACTIVIDADES NECESARIAS PARA LA IMPLEMENTACIÓN .....	122
4.3 INVENTARIO DE DOCUMENTACIÓN.....	127
CAPÍTULO V .....	130
CONCLUSIONES Y RECOMENDACIONES .....	130



## LISTA DE FIGURAS

Figura 1.1 - Estructura organizacional de la empresa de cobranza ABC .....	11
Figura 1.2 - Situación del proceso en la empresa ABC.....	16
Figura 1.3 - Flujograma del proceso general de cobranza en ABC.....	17
Figura 1.4 - Estructura funcional de la empresa de cobranza ABC .....	25
Figura 2.1 - Distribución de buenos y malos sobre la muestra de desarrollo.....	43
Figura 2.2 - Variables en la base de datos por su naturaleza.....	45
Figura 2.3 - Ejemplo de la no monotocidad del porcentaje de valores clasificados en función de K. ....	68
Figura 2.4 - Áreas bajo las curvas ROC de cada uno de los modelos .....	73
Figura 3.1 - Representación de los fenómenos de confusión e interacción.....	78
Figura 3.2 - Relación entre mora inicial, gestión telefónica del último trimestre y tipo de cliente .....	80
Figura 3.3 - Relación entre mora inicial y gestión domiciliaria último trimestre .....	84
Figura 3.4 - Esquema gráfico del diseño de estrategias basadas en modelos de score .....	92
Figura 3.5 - Segmentación de la variable sucursal por tipo de cliente utilizando árboles de decisión.....	95
Figura 4.1 - Cascada o matriz de rodamiento período t vs t+1.....	113
Figura 4.2 - Cálculo del cargo por edad de mora entre el período t y t-1.....	114
Figura 4.3 - Reporte de calidad de información y estabilidad de comportamientos .....	115
Figura 4.4 - Reporte de recaudo, cargo y deterioro por estrategia y mora .....	116
Figura 4.5 - Reporte de recaudo, cargo y deterioro por resultado de gestión.....	117
Figura 4.6 - Reporte de distribución de scorecard.....	118
Figura 4.7 - Reporte de mapa de estrategias.....	119
Figura 4.8 - Reporte cálculo de KS .....	119
Figura 4.9 - Reporte de comparación distribución buenos y malos. ....	120
Figura 4.10 - Reporte PSI .....	121
Figura 4.11 - Análisis FODA para la empresa ABC .....	123
Figura 4.12 - Diagrama causa efecto de la implementación del nuevo sistema en ABC..	124

## LISTA DE TABLAS

Tabla 2.1 - Cuadro de morosidad y efectividad en la cobranza por mes de desembolso ...	39
Tabla 2.2 - Distribución de buenos y malos clientes sobre la muestra de desarrollo.....	42
Tabla 2.3 - Variables recodificadas .....	46
Tabla 2.4 - Desarrollo de modelos utilizando regresión logística .....	50
Tabla 2.5 - Resultados de los modelos utilizando regresión logística.....	51
Tabla 2.6 - Resultados del modelo 17. ....	52
Tabla 2.7 - Distribución de buenos y malos, modelo 17 .....	53
Tabla 2.8 - Coeficiente Gini y Coeficiente K-S para el modelo 17.....	54
Tabla 2.9 - Coeficiente Gini y coeficiente K-S aplicado sobre la muestra de prueba.....	55
Tabla 2.10 - Distribución de buenos y malos sobre la muestra de prueba .....	55
Tabla 2.11 - Distribución de buenos y malos modelo final.....	56
Tabla 2.12 - Presentación de resultados modelo final con regresión logística clásica.....	56
Tabla 2.13 - Variables consideradas para el desarrollo del árbol de decisión.....	58
Tabla 2.14 - Distribución de buenos y malos con el desarrollo del árbol .....	59
Tabla 2.15 - Definición de nodos para buenos y malos .....	60
Tabla 2.16 - Distribución de buenos y malos con el desarrollo del árbol aplicado sobre la muestra de prueba.....	61
Tabla 2.17 - Variables del modelo logístico basado en distancias .....	64
Tabla 2.18 - Presentación de resultados de la regresión logística basada en distancias en la consola de R .....	65
Tabla 2.19 - Distribución de buenos y malos de la .....	65
Tabla 2.20 - Salida del paquete WEKA con la clasificación obtenida sobre la muestra de entrenamiento .....	69
Tabla 2.21 - Salida del paquete WEKA con la clasificación obtenida sobre la muestra de prueba .....	70
Tabla 2.22 - Análisis comparativo de modelos .....	71
Tabla 2.23 - Análisis comparativo de los modelos utilizando curvas ROC .....	72
Tabla 3.1 - OR entre mora inicial y tipo de cliente .....	79

Tabla 3.2 - OR entre mora inicial, gestión telefónica y tipo de cliente .....	80
Tabla 3.3 - Proporción de clientes sin gestión telefónica por edad de mora .....	81
Tabla 3.4 - Relación entre mora inicial y gestión telefónica último trimestre .....	82
Tabla 3.5 - ODD entre mora inicial y gestión telefónica último trimestre .....	82
Tabla 3.6 - Análisis de la interacción entre mora inicial y gestión telefónica último trimestre .....	83
Tabla 3.7 - Análisis de la interacción entre mora inicial y gestión domiciliaria último trimestre .....	83
Tabla 3.8 - Tabla de desempeño para el modelo de score de ABC .....	91
Tabla 3.9 - Estrategias por sucursales con alto o bajo riesgo .....	95
Tabla 3.10 - Estrategias por cantidad de cuotas pagadas con alto o bajo riesgo .....	96
Tabla 3.11 - Cuadro de asignación de acciones de cobranza por estrategia propuesto para la empresa de cobranza ABC .....	98
Tabla 3.12 - Estructura de costos de ABC.....	100
Tabla 3.13 - Tiempo promedio de gestión por estrategia de cobranza.....	101
Tabla 3.14 - Estructura de costos y factores de productividad para la cobranza telefónica en ABC .....	102
Tabla 3.15 - Costo por hora para un ejecutivo de cobranza telefónica en ABC .....	103
Tabla 3.16 - Costos y gastos por hora para un ejecutivo de cobranza telefónica en ABC.....	103
Tabla 3.17 - Cálculo del número de gestores telefónicos requeridos para ejecutar una estrategia de cobranza.....	105
Tabla 3.18 - Cálculo del número de gestores domiciliarios requeridos para ejecutar una estrategia de cobranza.....	106
Tabla 3.19 - Número de clientes por estrategia en el esquema tradicional de cobranza de ABC .....	107
Tabla 3.20 - Número de clientes por estrategia en el nuevo esquema de cobranza basado en score .....	107
Tabla 3.21 - Gestores telefónicos requeridos con el sistema tradicional de ABC.....	108
Tabla 3.22 - Gestores telefónicos requeridos con el sistema basado en score. ....	108
Tabla 3.23 - Gestores domiciliarios requeridos con el sistema tradicional de ABC .....	108
Tabla 3.24 - Gestores domiciliarios con el sistema de cobranza basado en score .....	109
Tabla 3.25 - Costo mensual de la gestión operativa con el sistema tradicional de ABC. ....	109

Tabla 3.26 - Costo mensual de la gestión operativa con el sistema basado en score. ....	109
Tabla 4.1 - Incidencias por criticidad de la categoría Tecnología. ....	125
Tabla 4.2 - Incidencias por criticidad de la categoría Operaciones. ....	126
Tabla 4.3 - Incidencias por criticidad de la categoría Planeación. ....	126
Tabla 4.4 - Incidencias por criticidad de la categoría Recursos Humanos. ....	126
Tabla 4.5 - Criticidad por categoría y factor empresa ABC. ....	127
Tabla 4.6 - Inventario de documentación. ....	128

## LISTA DE ANEXOS

ANEXO I.....	136
INDICADORES DE LA MICROEMPRESA EN ECUADOR .....	136
CARTERA BRUTA DE MICROEMPRESA .....	136
NÚMERO DE OPERACIONES DESTINADAS A LA MICROEMPRESA .....	136
TASAS DE INTERÉS.....	137
CARTERA VENCIDA SEGÚN TIPO DE CRÉDITO.....	138
ANEXO II .....	139
ANÁLISIS UNIVARIADO DE DATOS EMPRESA ABC .....	139
ANEXO III .....	161
REGLAS DE SELECCIÓN EN FORMATO SPSS UTILIZADAS PARA CONSTRUIR EL ÁRBOL SOBRE LA MUESTRA DE PRUEBA .....	161
ANEXO IV .....	166
CONSTRUCCIÓN DEL MODELO DE REGRESIÓN BASADO EN DISTANCIAS UTILIZANDO EL PAQUETE R.....	166
ANEXO V .....	169
LA DISTANCIA EN ESTADÍSTICA.....	169
EL MODELO DE REGRESIÓN BASADO EN DISTANCIAS.....	172
ANEXO VI.....	179
VALORACIÓN DE LA CONFUSIÓN E INTERACCIÓN PARA LAS VARIABLES INDEPENDIENTES DEL MODELO .....	179
ANEXO VII.....	195
COSTOS POR HORA PARA EL CASO DE UN GESTOR DE COBRANZA DOMICILIARIA EN LA EMPRESA ABC.....	195
ANEXO VIII .....	197
ANÁLISIS DE CRITICIDAD PARA LA IMPLEMENTACIÓN DEL NUEVO SISTEMA DE COBRANZA EN ABC .....	197

## RESUMEN

Esta tesis muestra el uso de modelos analíticos basados en score para la gestión de cobranza prejudicial en la empresa de cobranza ABC. Se basa en el cálculo de la probabilidad que un cliente deteriore su comportamiento de pago, utilizando para esto cuatro metodologías: i) regresión logística, ii) regresión logística mejorada con árboles de decisión, iii) regresión basada en distancias, iv) clasificación basada en distancias con el algoritmo K-NN, y en base a los resultados segmentar una cartera de microcréditos para definir acciones diferenciadas que ayuden a incrementar la recuperación a un menor costo y mejorando el servicio. El resultado de este trabajo muestra las ventajas que ofrece un nuevo sistema de cobranza con orientación al cliente sobre el esquema tradicional basado en la recuperación de saldos, donde lo que interesa es que las personas paguen lo que deben y no volver a tener otro tipo de relación comercial en el futuro con ellas. La cobranza orientada al cliente se refiere al uso del conocimiento sobre el desempeño y características del mismo, donde la probabilidad de deterioro del comportamiento de pago define las estrategias. Este modelo de gestión puede ayudar a los directivos de la empresa de cobranza a tomar decisiones rápidamente de forma automatizada, con un mayor conocimiento a priori de la operación y tipos de clientes a tratar: valoración cuantitativa del potencial de recuperación de un impago, aceleración del proceso de recuperación, reducción de costos en la gestión y mantenimiento de la relación comercial, porque los resultados de esta tesis presentan los problemas y el potencial para mejorar los niveles de recuperación de cartera con una adecuada administración de los recursos disponibles. En la parte final de este trabajo se propone algunas acciones para la implementación del nuevo sistema de cobranza, analizando las fortalezas de ABC y definiendo un esquema de reportes para control y seguimiento.

Palabras clave: Score, Cobranza Prejudicial.

## ABSTRACT

This thesis shows the use of analytical models based in score for the management of pre-judicial collection in the Collection Company ABC. It's based on the calculation of the probability that a client damages its behavior of payment, using four methodologies i) logistic regression, ii) logistic regression improved with trees of decision, iii) regression based on distances, iv) classification based on distances with the algorithm K-NN, and on the basis of the results, segment a microcredit portfolio to define differentiated actions that help to increase the recovery at less cost and improving the service. The result of this work shows the advantages that a new system of collection, oriented to the clients, offers over the traditional scheme based on the recovery of balances, where the goal is that people pay what they owe and then never have another type of commercial relation with them. The collection orientated to the client refers to the use of the knowledge of the performance and characteristics of the client, where the probability of deterioration of its behavior of payment defines the strategies. This model of management can help the executives of the collection company to take decisions fast and automatic, with better knowledge a priori of the operation and the types of clients they'll deal with: quantitative valuation of the potential of recovery of a non-payment, acceleration of the process of recovery, reduction of costs in the management and maintenance of the commercial relation, because the results of this thesis present the problems and the potential to improve the levels of recovery of portfolio with a suitable administration of the available resources. The final part of this work proposes some actions for the implementation of the new system of collection, analyzing ABC's strengths and defining a scheme of reports for control and follow-up.

Keywords: Score, Pre-judicial Collection.

## **CAPÍTULO I**

### **EL PROCESO DE COBRANZA DE UNA CARTERA DE MICROCRÉDITOS EN LA EMPRESA ABC**

Empiezo este trabajo presentando la evolución de las microfinanzas y la microempresa en Ecuador, cómo fueron sus inicios y las oportunidades que marcaron el desarrollo de esta clase de servicios financieros en el país, para citar luego descripciones y conceptos importantes respecto al tema que me ocupa. Al final describo el manejo que las instituciones bancarias hacen en la actualidad para mejorar sus operaciones, asegurando un control adecuado de los riesgos en un entorno de mercado cada vez más competitivo. En este sentido el desarrollo de estrategias y búsqueda de nuevos mecanismos de cobranza apunta a que las instituciones de microfinanzas opten por externalizar esta actividad, debido en parte a los altos costos de operación y también porque la especialización que existe hoy en las empresas dedicadas a este servicio ofrece buenos niveles de recuperación.

Una de estas empresas especializadas en cobranza para microfinanzas, misma que por razones de confidencialidad llamaré “Empresa de Cobranzas ABC”, operaba con un esquema tradicional de cobro, dando el mismo tratamiento a todos los clientes y basando su estrategia en la recuperación de saldos; esto originaba costos elevados en recursos humanos y técnicos, así como la pérdida de oportunidad para realizar una cobranza adecuada y rápida de la cartera, aumentando el riesgo de deterioro y la insatisfacción de los clientes. Para comprender mejor la situación de ABC, describo los procesos involucrados en la operación que realiza sobre una cartera de microcréditos: la relación comercial previa con los dueños de la cartera o Cedente, la infraestructura y servicios de cobranza que ofrece, su estructura organizacional, el intercambio de datos a través de los sistemas informáticos, los distintos tipos de interacción con los



clientes y los puntos de control sobre estas interacciones, el manejo y gestión de la información. A partir de esta descripción propondré en los siguientes capítulos un nuevo sistema de cobranza para mejorar la experiencia del cliente y la oferta en el mercado de ABC.

## **1.1 LAS MICROFINANZAS Y LAS MICROEMPRESAS**

El microcrédito, también denominado microfinanzas, se inició en el país en 1986 a través de un programa impulsado por el gobierno nacional, que consistió en la provisión de una variedad de servicios financieros tales como depósitos, préstamos y seguro a familias pobres que no tenían acceso a los recursos de las instituciones financieras formales. Los préstamos serían utilizados para invertir en microempresas, así como también en salud y educación, mejorar la vivienda o hacer frente a emergencias familiares.

En un principio los bancos fueron reticentes a operar estos microcréditos debido a:

1. Los montos crediticios eran muy pequeños.
2. Costos administrativos muy altos.
3. La imagen de los bancos se deterioraba.

Ante la actitud de la banca, el Estado intentó canalizar los recursos a través de ONGs, pero intereses políticos y económicos hicieron que estos programas fracasasen. Esta realidad fue el pretexto que el Banco Interamericano de Desarrollo (BID) utilizó para reformar las reglas de intermediación del microcrédito, exigiendo la utilización exclusivamente de la banca.

### **1.1.1 CONCEPTO DE MICROFINANZAS**

Según la definición de la Superintendencia de Bancos del Ecuador, el microcrédito es: “todo crédito concedido a un prestatario, sea persona natural o jurídica, o a un grupo de prestatarios con garantía solidaria, destinado a financiar actividades en pequeña escala, de producción, comercialización o servicios, cuya fuente principal de pago la constituye el producto de las ventas o ingresos generados por dichas actividades, adecuadamente verificados por la institución del sistema financiero prestamista”<sup>1</sup>. Además, el límite máximo para considerarlo microcrédito es de US\$ 20.000, el solicitante no necesita presentar estados financieros, sino que la institución crediticia deberá evaluar su capacidad de pago y la estabilidad de su fuente de recursos. Además establecerá el monto del crédito mediante la aplicación de tecnología crediticia específica en microfinanzas.

En la práctica la clasificación de los microcréditos se guía directamente por el monto del crédito sin analizar al destinatario del préstamo, con lo que se incluye en la cartera de microfinanzas una importante cantidad de créditos de consumo.<sup>2</sup>

### **1.1.2 IMPORTANCIA DE LA MICROEMPRESA**

El microcrédito no se limita a entregar servicios financieros solamente, sino que provee entrenamiento en el manejo del dinero y toca aspectos tales como liderazgo, confianza, autoestima, educación y manejo de microempresas que proveen de trabajo a un gran porcentaje de personas con ingresos medios y bajos en Ecuador. Estimaciones basadas en los datos del Censo de Población y Vivienda del 2001, indican que en ese año existían 1.027.299 microempresarios urbanos y 997.953 microempresarios rurales; cifras más actuales aseguran que

---

<sup>1</sup> Junta Bancaria, Resolución 457 del 2002.

<sup>2</sup> Son créditos de consumo los otorgados por las instituciones controladas a personas naturales que tengan por destino la adquisición de bienes de consumo o pago de servicios, que generalmente se amortizan en función de un sistema de cuotas periódicas y cuya fuente de pago es el ingreso neto mensual promedio del deudor, entendiéndose por éste el promedio de los ingresos brutos mensuales del núcleo familiar menos los gastos familiares estimados mensuales”. Fuente: Superintendencia de Bancos del Ecuador.

las actividades económicas que son consideradas microempresas lo desarrollan más del 50% de la población económicamente activa, es decir más de dos millones de personas, y solo en el sector urbano existen más de un millón de microempresas que aportan al producto interno bruto entre el 10% y 15%.

Para tener una visión más clara de la importancia de una microempresa y su impacto sobre la economía ecuatoriana, a continuación incluyo un resumen con las conclusiones más significativas del Estudio Nacional de Microempresas del Ecuador patrocinado por USAID<sup>3</sup>, mismo que fue desarrollado durante un período de 15 meses desde fines de octubre de 2003 hasta enero del 2005. El informe completo está publicado en la dirección electrónica:

[http://www.ruralfinance.org/fileadmin/templates/rflc/documents/1131636604206\\_Microenterprises\\_and\\_microfinance\\_in\\_Ecuador\\_ES\\_.pdf](http://www.ruralfinance.org/fileadmin/templates/rflc/documents/1131636604206_Microenterprises_and_microfinance_in_Ecuador_ES_.pdf)

### **Informalidad**

La mayoría de las microempresas opera en el sector informal. Aproximadamente un cuarto de las empresas tiene números de identificación de RUC (Registro Único de Contribuyentes) y un número igual tiene licencias municipales. Menos del 15% está inscrito en el sistema de seguridad social y solamente 20% llevan registros financieros formales.

### **Impacto de género de las microfinanzas**

Una característica de los microempresarios es el alto porcentaje de mujeres que dependen de una microempresa. Las mujeres comprenden solamente el 30.7% de la población económicamente activa pero constituyen el 46.7 % de los microempresarios. De las mujeres económicamente activas, 56.4% son

---

<sup>3</sup> United States Agency for International Development, institución creada por el gobierno de Estados Unidos en 1961 durante el gobierno de John F. Kennedy. Más información <http://www.usaid.gov>

microempresarias, gran parte cabezas de familia: solteras, separadas, viudas o divorciadas, demostrando que las microempresas constituyen una red de protección social importante para este sector de la población.

### **Concentración por sectores económicos**

Las microempresas están muy concentradas en el comercio, el 55.2 % de ellas están en este sector comparadas con el 25.7% en el sector de servicios y el 19.2% en producción. Dentro de cada uno de los sectores económicos hay una concentración importante de empresas en subsectores específicos, en el sector comercio por ejemplo el 60.2% de negocios están concentrados en comida, bebidas y vestuario. En el sector de servicios el 64.9% de microempresas están concentradas en cuatro subsectores: pequeños locales de comida y bebida (bares, restaurantes, y cafeterías), taxis, talleres de reparación automovilísticos y salones de belleza. En el sector de producción las microempresas están concentradas principalmente en ropa, mobiliario, artesanías y bebidas no alcohólicas, 56.2 % de las empresas en estos cuatro subsectores. Las mujeres microempresarias exhiben una concentración más grande en determinados tipos de negocios; por ejemplo en el sector de servicios, 85.0 % de las mujeres están en alimentos, bebidas y salones de belleza. En el sector de la producción 74.7% de las mujeres empresarias se encuentran en vestuario y bebidas no alcohólicas. En el sector comercial 67.9% de las mujeres están involucradas en la venta de alimentos, bebidas, vestuario y artículos misceláneos pequeños.

### **Impacto en el empleo**

Las microempresas son una fuente importante de empleo, sin embargo para la mayoría de microempresarios la empresa es una fuente de autoempleo. Cerca del 70% de microempresas ecuatorianas no emplean trabajadores o asistentes además del microempresario; además tienden a no crecer. En la gran mayoría de microempresas el empleo se genera solo cuando se forma y crece muy poco;

después, apenas el 10 por ciento de las microempresas han incrementado el nivel de empleo durante la vida del negocio.

### **Ventas e ingresos**

La mayoría de los microempresarios dependen del ingreso que ganan en su negocio: el 86.9% opera su negocio como una actividad a tiempo completo y el 94.3% dice que la empresa es su única fuente de ingresos. Para el 67.7 % de las familias la microempresa constituye la principal fuente de ganancias de la familia. La microempresa promedio tuvo ventas mensuales de \$778 que producen un ingreso familiar mensual neto de \$308, o un ingreso familiar anual neto de \$3,696. Las ventas y los ingresos constituyen la base para hacer crecer la empresa y para determinar la cantidad del crédito que un cliente puede razonablemente absorber y pagar. Basados en las ventas y ganancias, las microempresarias mujeres, por ejemplo, tienen una capacidad de pago de los préstamos más limitada que sus homólogos hombres, de lo que se puede concluir que las empresas cuyas propietarias son mujeres tienen ventas mucho más bajas y generan significativamente menos ganancias que las empresas de hombres. Similarmente las empresas de la Costa no parecen calificar para préstamos tan grandes como aquellas de la sierra y amazonía. Las diferencias entre empresas de varios subsectores en relación con su habilidad de generar ingresos son también relevantes, así por ejemplo las empresas en el sector de comercio tienen generalmente ventas más altas en general y generan más ganancias que las empresas de servicios o de producción.

### **Satisfacción y optimismo**

En general, los microempresarios están satisfechos con las ganancias que obtienen de sus empresas. La gran mayoría siente que sus ingresos están alrededor del "promedio", más de la mitad dicen que es bueno o muy bueno. Además casi el 70% dice que sus ingresos son mejores de los que podían recibir

en un empleo remunerado. Esta combinación del compromiso para trabajar independientemente y el deseo de ver crecer a la empresa son importantes para las instituciones de microfinanzas que consideran otorgar préstamos a microempresas. Los prestatarios que tienen fuertes sentimientos positivos sobre su empresa, que son optimistas respecto al futuro y que están comprometidos a ver que la empresa crezca, constituyen probablemente mejores sujetos de riesgo para el crédito que quienes montan una empresa debido a una falta de oportunidades, siendo pesimistas acerca del futuro y que preferirían tener empleo asalariado. El gran número de empresarios que dicen que les gustaría ver crecer sus empresas sugiere que las instituciones de microfinanzas tienen un mercado potencialmente grande para financiar las actividades empresariales crecientes.

### **Crecimiento y éxito**

Aunque los microempresarios expresan niveles altos de satisfacción y optimismo, muy pocos mostraron la clase de mejoras o crecimiento asociado con un exitoso o creciente negocio. Relativamente pocos microempresarios sentían que las ventas estaban aumentando, casi el 40% dijeron que las ventas fueron un poco o mucho menos que las del año anterior comparado con sólo el 15 % que dijeron que las ventas habían aumentado. Muy pocos han incrementado el número de empleados y otros tantos han hecho mejoras considerables a sus negocios.

Todos estos factores sugieren que la mayoría de microempresarios está operando en nichos de mercado muy limitados y altamente competitivos, por lo que su potencial para el crecimiento en las ventas y el ingreso es limitado. Esta aparente contradicción es considerada cuidadosamente por las instituciones de microfinanzas que buscan extender sus carteras, porque la ausencia de indicadores de crecimiento pudiera significar un sector estancado con una perspectiva pequeña de crecimiento en ausencia de otros cambios fundamentales en las microempresas y sus mercados.

### **Perfil del microempresario**

La mayoría de los microempresarios (55%) son adultos entre 31 y 50 años. Los jóvenes microempresarios entre 18 y 30 años alcanzan sólo el 20%. Por otra parte los mayores de 51 años representan al 25% de los microempresarios. Sobre las fuentes de financiamiento para iniciar la microempresa, la mayoría de microempresarios inicia su empresa gracias a ahorros personales (67,1%); la segunda fuente de financiamiento son los préstamos de familiares y amigos con un 12,6% en promedio. Muy pocos microempresarios indicaron no conocer alguna institución financiera, sólo el 8%. Sin embargo es muy bajo el nivel de financiamiento de la microempresa a través del sector formal de instituciones financieras. Sólo el 15,7 % de los microempresarios solicitó un préstamo en los últimos 12 meses con un 97% de éxito entre quienes lo solicitaron y obtuvieron, sin diferencias significativas entre hombres y mujeres o condición económica.

#### **1.1.3 MARCO REGULATORIO Y LEGISLACIÓN RELACIONADA CON LA MICROEMPRESA.**

Del apartado anterior se puede concluir que el empleo en la microempresa se relaciona por definición como empleo en el sector informal; las microempresas generalmente no cuentan con las licencias municipales requeridas, en la mayoría de veces los empleados de la microempresa no se encuentran anotados en los registros oficiales del ministerio de trabajo, ni registrados en la seguridad social; tampoco reportan sus niveles de ventas a la autoridad impositiva como el Servicio de Rentas Internas (SRI). Esto se debe en parte a que los costos asociados a registrarse (pago de impuesto, inscripciones, tasas, etc.) son muy altos y también porque el proceso es complicado y tedioso. Según el Banco Mundial, iniciar un negocio en Ecuador requiere cumplir con 14 trámites que incluyen desde un certificado de depósitos en el banco hasta registros en la Superintendencia de Compañías, SRI, Cámara de Comercio, Registro Mercantil, Seguridad Social, Ministerio de Trabajo y Municipio, entre otros. Para realizar

estos 14 trámites un empresario se demora en promedio 69 días hábiles y el costo asciende a US\$ 2.200.

En la actualidad la política económica y social apunta a una mayor participación del Estado en las actividades productivas de la vida nacional, buscando la generación de empleo y mejorando la redistribución de la riqueza a través de la contratación pública y manteniendo varios subsidios; sin embargo es imprescindible el establecimiento de políticas para impulsar el desarrollo del sector financiero y especialmente el sector microfinanciero, buscando no desmotivar la actividad emprendedora privada y orientando su desarrollo con políticas públicas que brinden confianza y seguridad a la inversión local y extranjera, esfuerzo que debe ser compartido por los entes de control. La proyección del Estado con la Superintendencia de Bancos y Seguros es fortalecer la relación entre las finanzas populares e instrumentar un marco de supervisión y control que propicie el crecimiento y desarrollo de las cooperativas de ahorro, entidades asociativas, de las cajas y bancos comunales y cajas de ahorro. La actual Constitución señala al sector popular como parte del sistema financiero y ordena la instrumentación de normas de control específicas que busquen su seguridad, estabilidad, transparencia y solidez.

#### **1.1.4 INDICADORES DE LAS MICROFINANZAS EN EL ECUADOR**

Información de los principales indicadores de microfinanzas, tales como la cartera bruta de la microempresa, el número de operaciones crediticias, las tasas de interés y la morosidad, junto con su evolución en el mercado ecuatoriano, se incluyen en el Anexo I.



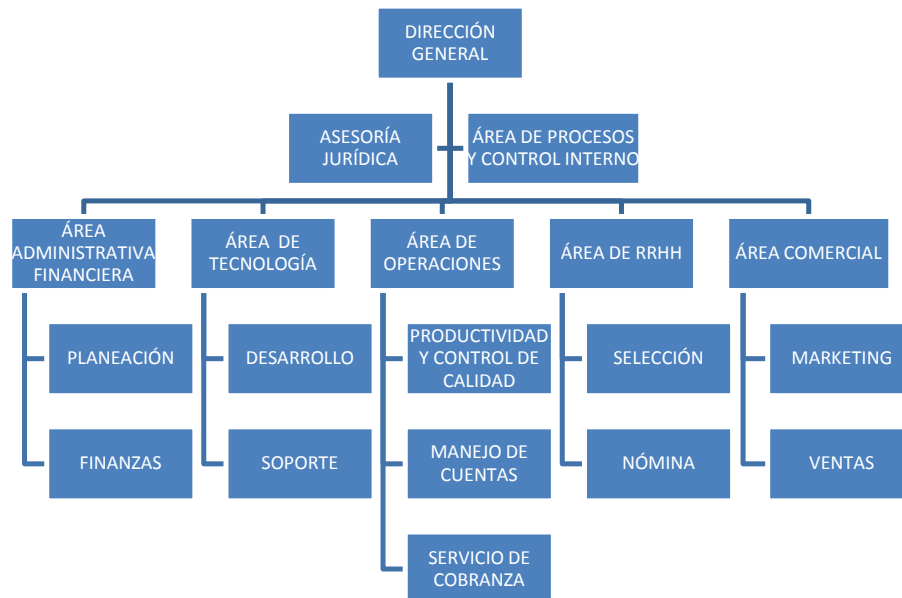
## **1.2 PROCESO GENERAL DE COBRANZA EN ABC**

ABC es una empresa de servicios en el área de Cobranza y recuperación de cartera; su objetivo final es maximizar los resultados operacionales como empresa, para lo cual define dos principios básicos que conforman el marco regulatorio de todo su accionar:

1. Consecución de los más altos estándares de productividad y satisfacción de su cartera de clientes.
2. Adecuada administración de los recursos propios.

A través del cumplimiento del primer principio que se conoce como la fidelización de los clientes a quien presta sus servicios, o Cedentes, se asegura y se incrementa su potencial de ingresos. Por su parte, el segundo principio junto con contribuir en forma directa sobre el primero, asegura niveles de gastos que resultan aceptables para los fines deseados. Ambos principios aplicados de la forma correcta procuran que ABC sea una entidad que tecnifica y profesionaliza el servicio de cobranza y recuperación de créditos.

ABC ha establecido de acuerdo a la norma ISO 9000:2000 la definición de proceso en la empresa y del área operativa en particular, como el “conjunto de actividades mutuamente relacionadas o que interactúan, las cuales transforman elementos de entrada en resultados”. Como estos resultados influyen sobre los procesos de otras áreas en la empresa, conviene tener claro de antemano la estructura organizacional (Figura 1.1) donde se desarrolla el mencionado proceso.



**Figura 1.1** - Estructura organizacional de la empresa de cobranza ABC

Más allá del concepto, es lógico y deseable en cualquier organización en materia del servicio de cobranza que esta integración exista, y para ello basta con ejemplificar algunas situaciones particulares que pude verificar en ABC:

- Si no existe venta del servicio, la estructura y herramientas no tienen sentido.
- La no disponibilidad y la no operatividad de un sistema informático hacen volver a procesos manuales de un alto costo y difícil aplicación.
- A pesar de efectuar la cobranza en todas sus fases y bajo cualquier esquema, si no hay herramientas de medición y evaluación los resultados no son medibles, no se aprecian, no existe comparación y no es posible la difusión a quienes contratan el servicio.
- Existiendo cobranza, si no se posee la estructura administrativa idónea para efectuar la facturación del servicio en forma clara y oportuna, se atenta contra la responsabilidad y seriedad del trabajo que efectúa ABC.

### **1.2.1 EL SERVICIO DE COBRANZA**

Las acciones básicas que constituyen la razón de ser en cualquier empresa dedicada al servicio de cobranza, ejecutadas en cada una de las fases que se distinguen en un proceso de recuperación de créditos afectos a morosidad, constituyen lo que se denomina el Servicio de Cobranza. Estas fases dependen de los acuerdos comerciales previos con el dueño de la cartera y son las siguientes:

#### **Cobranza Preventiva**

Está orientada a efectuar gestiones que permitan anticipar la morosidad de las deudas de acuerdo a un calendario mensual de vencimientos informado previamente por el Cedente y cuyo objetivo es afianzar la relación comercial, recordar el vencimiento y establecer un compromiso de pago. Adicionalmente permite entregar un servicio de postventa, detectando oportunamente los problemas que se traducen en objeciones de pago a futuro.

#### **Cobranza Prejudicial**

Esta fase del servicio está dirigida a recuperar deudas morosas mediante llamadas de cobranza telefónica, visitas domiciliarias y el uso de medios masivos. El objetivo de ABC es en primer término que el deudor se ponga al día en la o las cuotas morosas, logrando de este modo contener el período de morosidad evitando que aumente su deuda. La característica principal es que se trata de una fase amigable en donde se incentiva al cliente a normalizar su situación de morosidad, realizando acciones orientadas a:

- Localizar y contactar con el cliente deudor en forma directa.
- Incentivar al cliente a la normalización.
- Obtener el pago o comprometer el mismo en la fecha más temprana.

- Registrar cada acción y conformar la historia de gestiones del cliente.
- Validar los datos e informaciones sobre el cliente que están disponibles en la base de datos.

### **Cobranza Extrajudicial**

Normalmente la ejecución de acciones judiciales resulta de un alto costo para ABC y precisa de un tiempo no despreciable antes de conseguir resultados reales. Conforme con esa situación se establece este servicio como posterior a la gestión prejudicial, pero previo a la acción judicial. Su objetivo es la maximización del recupero mediante:

- Conseguir la mejor forma de pago, convenio o acuerdo de pago antes de iniciar el proceso judicial.
- Validar la información disponible de cada cliente, para seleccionar de la mejor forma aquellos clientes para los cuales resulta conveniente y recomendable desde la parte económica el accionar judicial.

El término de esta gestión está previamente definido con la empresa Cedente y depende en gran medida de sus políticas crediticias y comerciales.

### **Cobranza Judicial**

Este servicio se realiza para todos aquellos deudores que escapan a la acción extrajudicial y para los cuales el análisis de los anteCedentes así lo recomienda. En última instancia la determinación final de este proceder es materia de la empresa Cedente. El objetivo de esta fase es utilizar los elementos disponibles en la justicia para conseguir la recuperación de los capitales y créditos en mora. En términos de negocio, ABC procura siempre llegar a un acuerdo y negociación con los clientes, una vez más, en conformidad a las políticas de la empresa Cedente.

### **1.2.2 EL SISTEMA DE COBRANZA**

Los distintos tipos de interacción entre ABC y los clientes de la cartera del Cedente, desarrollados conforme a procedimientos operativos y administrativos conforman lo que se denomina Sistema de Cobranza y son las siguientes:

- Gestión telefónica
- Gestión domiciliaria o de terreno
- Gestión por medios masivos

La base conceptual del sistema de cobranza en ABC es administrar las diferentes acciones de cobro, de manera que se generen ya sea de forma manual o automática, nuevas acciones más eficaces dependiendo del resultado en cada interacción que se realice. Toda acción de cobranza y su resultado quedan registrados en la base de datos, conformando un registro histórico y una estructura de información que es utilizada en cualquier instancia en todas las fases del proceso; esto último facilita la obtención de estadísticas que permiten evaluar, analizar y predecir no sólo el comportamiento del cliente sino de toda la cartera a la cual pertenece. Cada forma particular de interacción se realiza conforme a una planificación, esto es, se especifica la acción en su forma y fondo para cada cliente y para cada cartera de clientes.

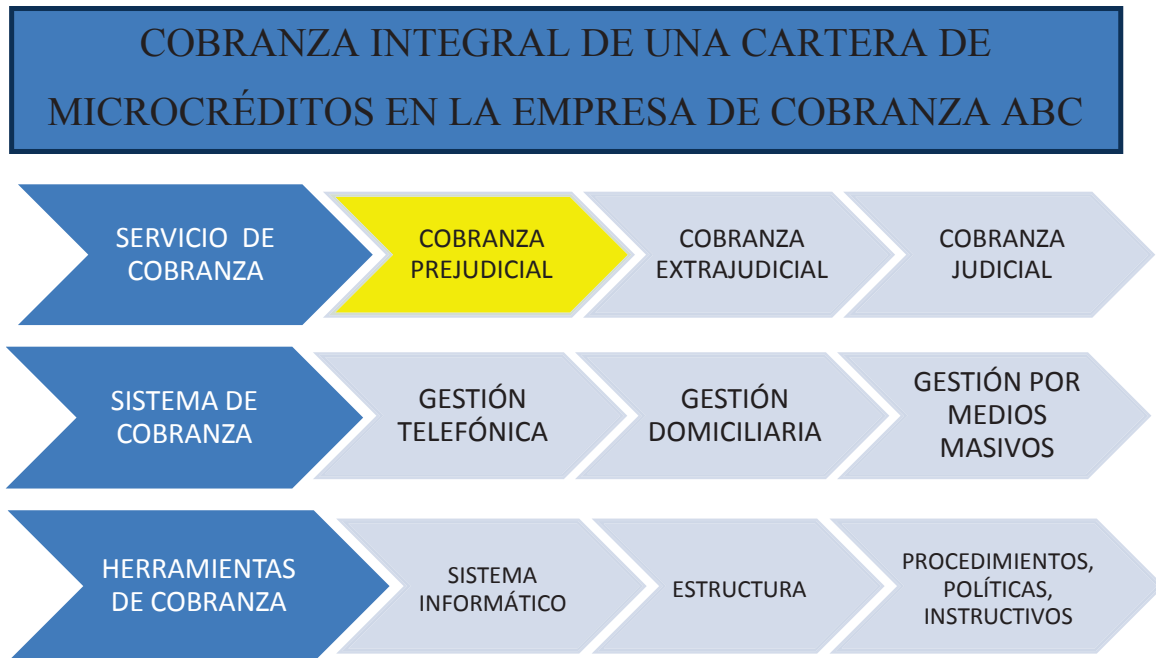
### **1.2.3 LAS HERRAMIENTAS DE COBRANZA**

Aun cuando parezca básico, la estructura de soporte de una empresa como ABC debe estar de acuerdo con los objetivos establecidos para ella y es precisamente aquí donde la experiencia propia de la empresa refuerza la idea de contar con una estructura suficiente, donde la función de cada uno de sus componentes esté orientada hacia el cumplimiento de estos objetivos, tomando en cuenta el hecho gravitante de contar con un sistema informático de cobranza propio porque es una herramienta clave para la consecución de los mismos. El adecuado uso de esta

herramienta y la interacción con los medios humanos, junto con los procedimientos, manuales y políticas que norman la operación, es lo que permite a la empresa conseguir el mejor nivel de satisfacción de su cartera de clientes y el cumplimiento de los indicadores clave que se han establecido para medir su desempeño.

### **1.3 SITUACIÓN ACTUAL DEL PROCESO.**

A través de los documentos realizados por el área de procesos de la empresa ABC, establecí que el foco de su accionar está orientado a ofrecer en todo momento lo que se llama Cobranza Integral, entendiendo por tal, los diferentes Servicios de Cobranza sobre cada una de las fases que se distinguen en un proceso de recuperación de créditos en cada una de las edades de mora. Para lograr este objetivo, ABC practica y desenvuelve los diferentes subprocesos operacionales y administrativos que conforman el denominado Sistema de Cobranza, sustentando los mismos en las llamadas Herramientas de Cobranza (Figura 1.2).



**Figura 2.2** - Situación del proceso en la empresa ABC

Establecí como proceso de estudio, a las actividades que se realizan en el área de operaciones durante la fase prejudicial de la cobranza a una cartera de microcréditos. Este proceso forma parte del macro proceso denominado Cobranza Integral y se relaciona con otros procesos de la empresa tales como:

Proceso proveedor de entradas:

Ingreso y disponibilidad de la cartera a cobrar.

Planificación operativa de la cobranza.

Fidelización con el Cedente.

Proceso que recibe las salidas:

Gestión financiera administrativa (facturación, rentabilidad, compensación).

Gestión de la demanda al área de sistemas.

Gestión de la productividad (productividad, control de calidad).

Cobranza extrajudicial y judicial.

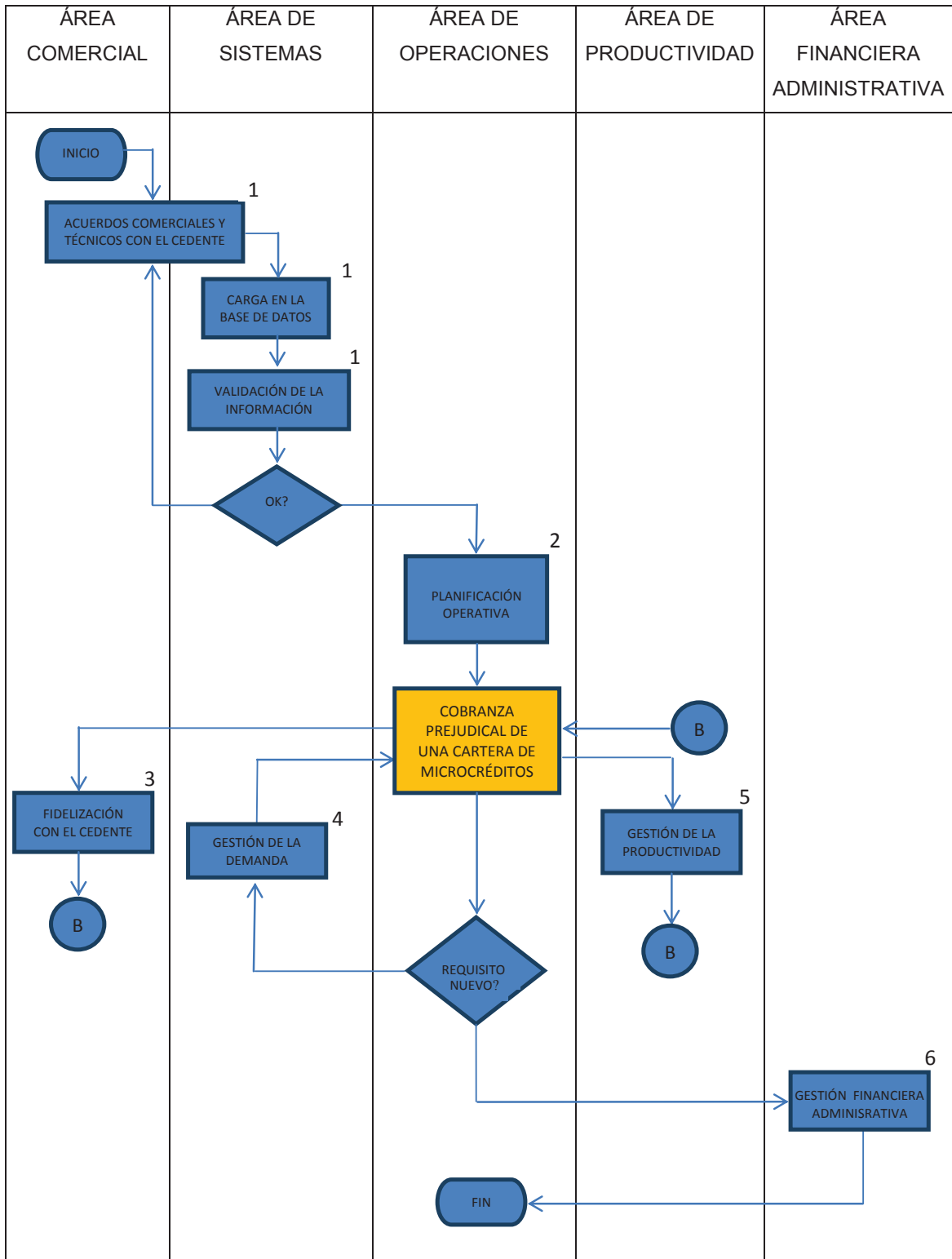


Figura 3.3 - Flujoograma del proceso general de cobranza en ABC



En la Figura 1.3 muestro paso por paso cómo interactúa entre sí toda la estructura y recursos, tanto humanos como tecnológicos en ABC, para ejecutar el sistema de cobranza y lograr finalmente los objetivos de la empresa.

### **1.3.1 INGRESO Y DISPONIBILIDAD DE LA CARTERA A COBRAR**

Se refiere a los aspectos más relevantes que ocurren desde que la carga o cartera de clientes de un Cedente llega a ABC, hasta que se encuentra en condiciones de ser accionada en cobranza a través de una o más de las gestiones que forman parte del Sistema de Cobranza. Involucra tanto al área de Informática como a la de Operaciones, la primera en la aplicación de todo el soporte tecnológico y la segunda principalmente en la determinación de la forma y oportunidad en que se requiere sea disponibilizada la información. En forma resumida este proceso contempla los grupos de actividades que se señalan a continuación:

#### **Acuerdos comerciales y técnicos con el Cedente.**

Esta fase preliminar que forma parte de las negociaciones previas con el Cedente es de la mayor importancia, ya que en ella se debe discutir y acordar todos aquellos aspectos y procedimientos técnicos y operacionales mediante los cuales se hará llegar la cartera a ABC y cómo ésta efectuará el paso de informaciones referidas a las gestiones y a las remesas<sup>4</sup> producto de los pagos que realicen los clientes en sus cajas.

#### **Carga y validación de la información de entrada.**

Con el fin de certificar la validez de la información recibida, el primer proceso es su validación conforme a patrones y procedimientos estándares.

---

<sup>4</sup> Dinero recaudado en cajas de la empresa de cobranza que debe ser transferido a las cuentas del Cedente.

Para el desarrollo de esta tarea los datos son cargados en tablas temporales desde las cuales es posible certificar su consistencia antes de formalizar su ingreso. Como resultado de esta fase de trabajo se puede contar con estadísticas preliminares sobre la cartera recibida, de forma que tanto el área de informática como el área de operaciones posean los antecedentes que le permitan planificar y montar las estrategias para llevar a cabo la mejor forma de cobranza.

### **1.3.2 PLANIFICACIÓN DE LA COBRANZA EN ABC**

Como paso previo a la fase operativa se requiere una coordinación con todas las áreas involucradas para formalizar en un documento los acuerdos del servicio que se entregará al Cedente, así como las definiciones comerciales y técnicas que servirán como base para la ejecución de las diferentes acciones en cada etapa de la gestión operativa, documento que es elaborado según las indicaciones entregadas al usuario responsable del área de operaciones. Este proceso contempla los siguientes aspectos:

#### **Distribución de la cartera**

Visto que una cartera de clientes normalmente posee una distribución geográfica que excede al de una ciudad o localidad, y que por lo tanto pertenece a diferentes sucursales o filiales del Cedente, se hace necesario parear esa sucursal Cedente con alguna sucursal de ABC que ofrezca una mejor atención desde el punto de vista del alcance geográfico. Al hablar de mejor atención, el objetivo está apuntado a la responsabilidad de ocuparse en forma directa por la calidad y oportunidad de gestión para los clientes asignados.

#### **Asignación de la cartera**

Se refiere al proceso de asignar los clientes para que sean gestionados ya sea en telefonía, en terreno, a través de métodos masivos o con la combinación de todos

estos mecanismos, y que se haya determinado en las estrategias de gestión. Consiste en establecer los “usuarios” que el sistema informático reconoce como ejecutivos de cobranza, a los cuales les serán asignados clientes para su gestión y responsabilidad.

### **Estrategia**

Es una agrupación de clientes que de acuerdo con criterios predefinidos, poseen la característica común de satisfacer esos criterios. Así por ejemplo, si se define una estrategia X que debe reunir a todos los clientes con una morosidad entre 31 y 60 días, se agrupa a todos los casos de la cartera que precisamente posean una morosidad como la indicada.

El propósito fundamental de crear estrategias es direccionar el proceso de gestión y de control hacia segmentos de carteras definidos. Los criterios más usados a la hora de establecer estrategias en ABC son los que se indican a continuación:

- Misma clasificación de la Superintendencia de Bancos.
- Morosidad en un rango de tiempo.
- Montos en un rango de morosidad.
- Clientes con juicios.
- Mezclas de los anteriores.

### **Prioridad**

La gestión presentada a los ejecutivos de cobranza de su cartera de clientes es en forma secuencial y por tanto es necesario parametrizar o entregar un criterio al cual obedece esa presentación; esto se conoce como prioridad y por tanto a partir de esta prioridad es que el sistema seleccionará el orden de presentación de los clientes.

Mediante algoritmos de cálculo internos del sistema se determina un valor de prioridad para cada cliente conforme al criterio parametrizado. Los criterios utilizados por ABC son:

- Monto adeudado (capital o sumatoria de cuotas morosas).
- Sector geográfico.

### **1.3.3 FIDELIZACIÓN CON EL CEDENTE**

Visto que una empresa como ABC subsiste sí y sólo si posee carteras para cobrar, es vital entonces que se logre una adecuada relación comercial con los dueños de la cartera o Cedentes, generalmente bancos, casas comerciales y empresas de servicios, entendiendo por fidelización a mantener la confianza de ese Cedente con ABC.

Para conseguir esta confianza no sólo basta el resultado de las gestiones, aspecto que de por sí es más que importante, sino que también se considere otros aspectos como los que se indican a continuación, cuya consecución lleva a la fidelización comentada:

- Adaptación a las exigencias del Cedente.
- Adecuada y oportuna transmisión de informaciones al Cedente.
- Atención y flexibilidad a los cambios.
- Planificación adecuada del trabajo y medición de resultados.

Lo anterior debe estar cuidadosamente en armonía con los intereses propios de ABC, lo que involucra para tales fines la optimización en el empleo de los recursos propios.

#### **1.3.4 GESTIÓN DE LA DEMANDA AL ÁREA DE SISTEMAS.**

El sistema informático de ABC se sustenta en una base de datos única que permite la interacción entre los diferentes módulos con funciones específicas y que provee esas facilidades hacia otros lugares geográficos en los cuales se requiere que ABC actúe con su servicio.

El área de informática está estructurada en dos grandes grupos de trabajo, el de Desarrollo cuya función es analizar, crear, actualizar e implementar soluciones informáticas de acuerdo a las necesidades. El segundo componente del área es el grupo de Redes y Comunicaciones, cuya función es la de proveer el funcionamiento óptimo de la línea de comunicaciones, servidores y dispositivos de hardware en general, tanto dentro de la compañía como con los componentes de enlace con el Cedente.

#### **1.3.5 PRODUCTIVIDAD Y CONTROL DE CALIDAD**

La gestión estratégica de la calidad y la productividad en ABC se realiza mediante el diseño, planeación, ejecución, control y mejoramiento de los procesos, en un marco de desarrollo adecuado para la empresa y su crecimiento. La meta es crear constancia en el propósito de mejorar el servicio, con el objetivo de llegar a ser competitivos y permanecer en el negocio, dejando de depender de la inspección para lograr la calidad. La comunicación de procedimientos y definición de los responsables del cumplimiento, es uno de los pilares básicos sobre los cuales se sustenta el funcionamiento adecuado de esta estructura y por tanto la consecución de los objetivos fijados para la empresa.

En lo que concierne a las metas, la actividad propia de la cobranza está siempre sujeta a conseguir resultados sobre el rendimiento y la producción en cobranza, de aquí entonces la necesidad de establecer correctamente estas metas y a la

vez enfocar el esfuerzo para que sean cumplidas a todo nivel, lo que implica la existencia de las instancias de control sobre las mismas.

### **1.3.6 GESTIÓN FINANCIERA ADMINISTRATIVA**

La gestión financiera administrativa de los recursos disponibles para el cumplimiento de los objetivos empresariales de ABC, permite que sus estrategias de crecimiento y desarrollo prevean los efectos y las necesidades en el tiempo, y con base en ésta se redefinan o se reorienten. Se realiza a través del cumplimiento de tres actividades:

#### **Control del gasto**

En el análisis del costo de operación, por una parte resulta fácil mezclar aquellos del tipo fijo con los variables, y por otra es también fácil no compatibilizar el gasto en el que incurren las diferentes áreas de forma unitaria en función de la productividad. Por esta razón se consideran como fundamentales este tipo de mediciones e indexaciones de forma que el resultado general como empresa no se vea afectado por un incremento irracional del gasto. Algunos de los ítems de control sobre los cuales ABC fija su atención son:

- Gastos imputables por cada cartera en gestión.
- Gastos por centros o unidades de operación al interior de la empresa.
- Índices de productividad por centros o unidades de operación.
- Índices de productividad por cada tipo de gestión dentro del sistema de cobranza: telefonía, domicilio, masivos.

#### **Control de Ingresos**

Para una empresa como ABC la única fuente de ingresos la constituye el ingreso de honorarios, sean éstos en la forma de un cargo adicional al cliente o bien en la

forma de cargos o comisiones al Cedente. Por tanto, es una actividad de mucha sensibilidad en donde la problemática no es la confianza comercial en el pago o imputación de esos cargos sino que está en la forma de identificarlos y relacionarlos cliente a cliente o bien cartera a cartera.

Es importante indicar que estas acciones están afectadas por situaciones normales dentro de la operación, como son las posibilidades de descuentos aplicados a estos cargos y las opciones de pagos tanto en cajas del Cedente como de la propia ABC según sea el caso. Por lo expuesto ABC considera funciones o unidades especializadas para atender directamente el proceso de facturación o de control detallado de ingresos.

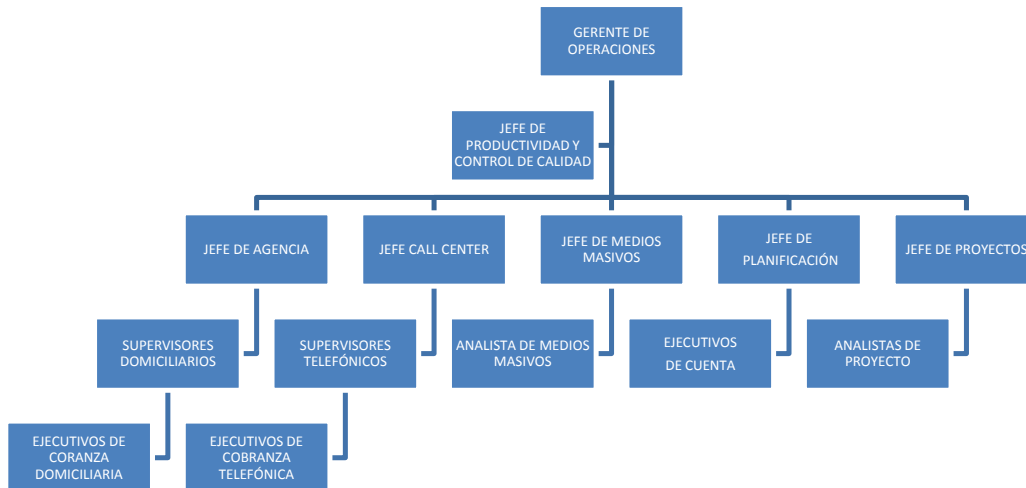
### **Cumplimiento de Presupuesto**

La naturaleza del negocio hace que toda la organización oriente sus esfuerzos hacia el cumplimiento del presupuesto de operación, lo cual requiere un estricto apego a la observación de las cifras presupuestadas tanto al nivel de ingresos como al nivel fijado como permisible para el gasto por cada área y por la empresa como un todo. De ahí entonces que la estructura debe contar con los elementos necesarios para formular y controlar las previsiones y proyecciones de esos índices.

## **1.4 SITUACIÓN ACTUAL DE LOS SUBPROCESOS**

El proceso operativo de cobranza de una cartera de microcréditos durante la fase prejudicial en la empresa ABC, está constituido por 3 subprocesos que se interrelacionan a través de su estructura funcional; la ejecución de este proceso es periódica mensual y los resultados son entregados a los procesos del área financiera administrativa y al departamento comercial del área de operaciones.

A continuación, junto con establecer las definiciones y explicaciones que contribuyan a entender los subprocesos, mostraré en la Figura 1.4 a modo de referencia la estructura funcional en ABC que se considera mínima para atender el flujo de tareas que se enmarca dentro del proceso interfuncional del servicio de cobranza para una cartera de microcréditos.



**Figura 4.4** - Estructura funcional de la empresa de cobranza ABC

De acuerdo con las características observadas en el organigrama funcional pude concluir que la división de funciones está muy bien delimitada, cada una de las personas conoce bien sus responsabilidades en el cumplimiento de objetivos y este conocimiento es impartido desde el primer día de labores por el personal más antiguo durante el período de inducción; sin embargo pude notar desconocimiento de las necesidades interdepartamentales y falta de comunicación para ser más efectiva la estrategia y planificación de la operación, siendo evidente que mucha de esta información no fluye adecuadamente en toda la estructura.



### **1.4.1 SUBPROCESO: GESTIÓN TELEFÓNICA**

#### **Objetivo**

Permitir de una manera masiva y rápida la localización y contacto con los clientes, contribuyendo a la mantención y mejoramiento del nivel de información en la base de datos que es fundamental dentro de la gestión de cualquier empresa de servicios, ofreciendo las mejores condiciones en términos de costo beneficio.

#### **Entradas, salidas y aspectos generales.**

Este subproceso tiene como entradas: la información que proporciona el sistema referente a números telefónicos del cliente y sus referencias, así como los datos de las operaciones vencidas del titular; el registro de gestiones previas y la calidad en el resultado de las mismas. Los productos que genera son: compromisos de pago, datos actualizados en el sistema y los reportes para el Jefe de Call Center.

Esta gestión está condicionada a la existencia de información sobre teléfonos del cliente, la calidad de los mismos y la disponibilidad del servicio externo y red de telefonía local. Al disponer de un centro grupal para esta gestión denominado Call Center, aparecen más claramente los beneficios económicos y de productividad, toda vez que es posible generar una economía de escala y una supervisión de procesos en forma más dedicada y específica.

Desde la perspectiva de la calidad de la gestión la telefonía resulta a veces menos intensa que un contacto personal con el cliente, ya que ofrece más facilidades para la evasión y/o falta de entendimiento de éste sobre sus obligaciones objeto de la cobranza. Esto es aún más notorio en moras avanzadas. Como contraparte, para las moras más recientes ese riesgo no es tan significativo.

## **1.4.2 SUBPROCESO: GESTIÓN DOMICILIARIA O DE TERRENO**

### **Objetivo**

Localizar y contactar de forma personal al cliente para incentivarlo a la normalización y solución de su morosidad, de acuerdo a la planificación de cobranza, al no disponer de información sobre teléfonos o no haber logrado éxito en la gestión telefónica (lo que se llama rebalse).

### **Entradas, salidas y aspectos generales**

Este subproceso tiene como entradas: la información que proporciona el sistema referente a direcciones domiciliarias del cliente y los datos de sus operaciones vencidas, mismas que son impresas en un formato en papel que se llama “ficha domiciliaria”; el registro de gestiones previas y la calidad en el resultado de las mismas. Los productos que genera son: dinero cobrado, convenios o acuerdo de pago, datos actualizados en las fichas domiciliarias y el sistema de cobranza, reportes para el Jefe de Agencia.

A diferencia de la gestión telefónica, la visita domiciliaria es una actividad que genera un mayor costo, esto fundamentalmente por la necesidad de contar con personal capacitado para atender una cartera significativamente menos voluminosa.

En forma similar a lo que ocurre en el caso de la telefonía, no existe una receta específica para determinar cuándo debe efectuarse esta gestión; la determinación sobre el efectuar o no efectuar una visita en terreno y cuándo hacerlo, formará parte de la planificación que debe estructurar el administrador, según la estrategia global para la cartera.

Existe cartera que estando en telefonía se trabaja también en domicilio, porque durante la gestión telefónica se determinó que perdió la característica de poseer algún teléfono vigente (teléfono fuera de servicio o no corresponde a clientes) o bien porque se excedió el plazo o número de intentos válidos para tratar de contactar telefónicamente al cliente y obtener un compromiso; a esto se le llama Gestión Doble.

Al hablar de Gestión Doble se piensa en efectuar simultáneamente tanto gestión telefónica como gestión domiciliaria. La realización de esta modalidad corresponde a una determinación operacional que debe considerar entre otros aspectos, las implicancias de costo y los acuerdos comerciales alcanzados con el Cedente.

### **1.4.3 SUBPROCESO: GESTIÓN CON MEDIOS MASIVOS**

#### **Objetivo**

El objetivo de esta gestión es contactar a grandes volúmenes de clientes a bajo costo, para lograr una vocación positiva de pago. Se utiliza distintos medios electrónicos e impresos masivos, tales como la mensajería de texto celular (SMS)<sup>5</sup>, correos electrónicos, cartas y sistemas automatizados de respuesta (IVR)<sup>6</sup>, mismos que a su vez pueden estar orientados para notificar la existencia de la mora, ofrecer condiciones de pago, entregar pines para efectuar pagos directos, ofertas de fidelización, entre otras.

---

<sup>5</sup> El servicio de mensajes cortos o SMS (Short Message Service) es un servicio disponible en los teléfonos móviles que permite el envío de mensajes cortos (también conocidos como mensajes de texto, o más coloquialmente, textos) entre teléfonos móviles.

<sup>6</sup> IVR son las siglas de Interactive Voice Response, que se traduce del inglés como respuesta de voz interactiva. Es un sistema automatizado de respuesta interactiva, orientado a entregar y/o capturar información a través del teléfono, permitiendo el acceso a servicios de información u otras operaciones.

**Entradas, salidas y aspectos generales.**

Este subproceso tiene como entradas: la información que proporciona el sistema referente a direcciones domiciliarias y teléfonos del cliente, así como los datos de sus operaciones vencidas. Los productos que genera son: llamadas de los clientes al Call Center, visitas a las oficinas, datos actualizados referentes a validez de teléfonos y direcciones, reportes para el Jefe de medios masivos.

El avance tecnológico y la integración de canales de comunicación telefónica y virtual a los sistemas de cobranza, permite utilizar medios alternativos a las tradicionales llamadas telefónicas y visitas domiciliarias para incentivar la normalización de las deudas vencidas de los clientes. Estos medios, si bien son más baratos, no sustituyen a la adecuada gestión por parte de un ejecutivo especializado, pero pueden ser tomados en cuenta como apoyo por el volumen de potenciales contactos que se pueden generar y la versatilidad para implementar campañas a diferentes segmentos. Ejemplos de estos medios modernos para la gestión de cobranza son las redes sociales.

**1.5 PLANTEAMIENTO DEL PROBLEMA**

El mercado de microcréditos en Ecuador se basa en el otorgamiento de préstamos de montos pequeños para desarrollar actividades generadoras de ingreso y autoempleo. La naturaleza masiva del negocio requiere una gran inversión en tecnología y capital humano para controlar todas las etapas del proceso de crédito. La gestión de cobranza no debe procurar únicamente la mayor recuperación de cartera sino también proveer información para mejorar la política de venta y los procesos de control, especialmente donde se presentan debilidades.

Una empresa de cobranzas tiene que innovar y mejorar continuamente su servicio, no solo por el cumplimiento de las metas de recupero impuestas por los Cedentes, sino también por las necesidades de información, complementación de servicios y disminución de costos que apuntan a mejorar su oferta en el mercado. Una forma para aumentar la recuperación de cartera, gastando menos y mejorando la relación comercial con los clientes es creando sistemas de gestión que utilizan soluciones analíticas, dentro de las cuales los modelos de score o tablas de puntuación son una de las herramientas más comunes.

La empresa de cobranzas ABC operaba con un esquema tradicional de cobro, dando el mismo tratamiento a todos los clientes y basando su estrategia en la recuperación de saldos; esto originaba costos elevados en recursos humanos y técnicos, pérdida de oportunidad de realizar una cobranza adecuada y rápida de la cartera, aumentando el riesgo de deterioro y la insatisfacción de sus clientes. Una gestión de cobranza sobre la base de un score estadístico permite focalizar los esfuerzos donde existe mayor riesgo y basa su estrategia en pronósticos y conocimiento de la cartera, obteniendo así mayor recuperación y mejores niveles de servicio. Para lograr este objetivo la estrategia debe adecuarse al tipo de segmento al que va dirigida y adicionalmente optimizar el uso de los recursos materiales y humanos que intervienen en el proceso.

### **1.5.1 FORMULACIÓN Y SISTEMATIZACIÓN DEL PROBLEMA**

- ¿Cómo se gestiona la cartera en una empresa de cobranzas tradicional?
- ¿La cobranza enfocada a la recuperación de saldos genera desgaste y favorece la deserción de clientes?
- ¿Cómo se puede aprovechar la información del cliente en el diseño de estrategias de recuperación?
- ¿Cómo se facilita la extracción de conocimiento contenido en grandes volúmenes de datos?

- ¿Qué aspectos podrían mejorar dentro de una empresa de cobranza si se utilizan soluciones basadas en pronósticos?
- ¿Cómo medir la eficacia de un modelo de recuperación de cartera?

## 1.5.2 OBJETIVOS DE LA INVESTIGACIÓN

### Objetivo General

Construir modelos estadísticos de score para calcular el riesgo de deterioro en los clientes de una cartera de microcréditos y plantear un sistema de recuperación basado en pronósticos para la empresa ABC, que ayude a mejorar la gestión que se realiza en la etapa prejudicial del servicio de cobranza.

### Objetivos específicos

- i. Describir los procesos involucrados en la gestión de recuperación de una cartera de microcréditos dentro de la empresa de cobranza ABC.
- ii. Mostrar una aplicación de los métodos más utilizados en la construcción de modelos de score, en el diseño de estrategias de cobranza dirigidas a clientes con similares patrones de riesgo, como son la regresión logística y árboles de decisión. Plantear la regresión logística basada en distancias y el clasificador KNN como herramientas alternativas al modelo logístico.
- iii. Validar el uso de un sistema de gestión de recuperación en base a los resultados obtenidos del modelo estadístico, que sea eficaz en el control de cartera vencida y considere la razón de negocio en la empresa de cobranzas ABC.
- iv. Proponer un nuevo sistema de gestión integral basado en mejorar la experiencia del cliente sobre las metas de recuperación, que mejore la oferta de mercado de la empresa de cobranza ABC.
- v. Mostrar las ventajas del nuevo sistema de gestión en cuanto a conocimiento del cliente, administración de la información y control del gasto operativo, para

- alinear los objetivos empresariales con los de la empresa Cedente en cuanto a riesgo, lealtad y post-venta.
- vi. Definir los reportes para seguimiento y control de la solución analítica, en base a los resultados que se obtengan en términos de recuperación y de la estabilidad del modelo estadístico a través del tiempo.
  - vii. Analizar la estructura requerida y definir las acciones necesarias en cada área de la empresa de cobranzas para la implementación del nuevo sistema de gestión.

### **1.5.3 JUSTIFICACIÓN DEL PROYECTO**

El modelo tradicional de recuperación tiene un enfoque extensivo, es decir, realizar la mayor cantidad de acciones de cobranza a todos los clientes de la cartera por igual; basa su gestión en la recuperación de dinero dando más importancia a la deuda que al cliente: lo que interesa es que las personas paguen lo que deben y no volver a tener otro tipo de relación comercial en el futuro. De esta manera clientes que deben poco pero con alto riesgo son ignorados y por el contrario gran cantidad de tiempo se aplica en clientes que deben mucho, pero con poco riesgo. Estas deficiencias que van en contra del enfoque moderno del servicio de cobranza, pueden ser vistas como una oportunidad de tener ventaja competitiva y mejorar el servicio.

Un modelo de score para recuperación de la cartera microcréditos con orientación al cliente, se refiere al uso del conocimiento sobre el desempeño y características del mismo para pronosticar el riesgo de deterioro en el futuro, permite tomar decisiones rápidamente de forma automatizada y en base a la información puede repercutir en un mayor conocimiento a priori de la operación y tipos de clientes a tratar; es un respaldo analítico a las políticas del proceso de venta y post-venta, valoración cuantitativa del potencial de recuperación de un impago, complemento a las estrategias de negocio, aceleración del proceso de recuperación, reducción de costos en la gestión de cobranza y mantenimiento de la relación comercial.

Lo que quiero es plantear una gestión de cobranza intensiva donde la probabilidad de recuperación defina las estrategias y que los esfuerzos sean dirigidos hacia segmentos con distinta probabilidad de recuperación, es decir, exista un trato diferenciado por cliente. Desde el punto de vista del dueño de cartera este modelo de gestión de recuperación contribuirá al cumplimiento de los objetivos comerciales de retención y mantenimiento de los clientes, ayudará a mejorar los niveles de riesgo, aumento del flujo de caja y control de provisiones. Desde el punto de vista de la empresa de cobranza mejorará los niveles de recuperación de cartera con una adecuada administración de los recursos disponibles y fortalecerá las relaciones comerciales con las empresas Cedentes, mejorando así su presencia en el mercado.



## **CAPÍTULO II**

### **CONSTRUCCIÓN DE MODELOS ESTADÍSTICOS PARA CALCULAR EL RIESGO DE DETERIORO DE LA CARTERA MICROCRÉDITO EN LA COBRANZA DE ABC**

En este capítulo seguí la técnica de minería de datos para la creación de modelos estadísticos de predicción, ajustarlos, validarlos y luego escoger el que mejor sirva para la construcción del nuevo sistema de cobranza de una cartera microcrédito en la empresa ABC.

Primero depuré la información para tener una base de datos completa y consistente, luego obtuve una estimación de la relación entre la variable dependiente o resultado y cada una de las variables independientes; como segundo paso pude crear múltiples modelos para estimar el riesgo de tener el efecto o desenlace establecido para un nuevo individuo con una cierta validez y precisión, quedándome al final con el de mejor predicción y que tuvo el menor número de variables. Para esto me valí de técnicas como la regresión logística clásica, árboles de decisión, regresión logística basada en distancias y el algoritmo de clasificación K-NN. En el tercer y último punto hice comparaciones entre los modelos obtenidos en el paso anterior y seleccioné el que tenía una buena predicción pero con mayor capacidad de explicación.

Como paso previo al análisis de datos fue necesario tener varias reuniones, por un lado con los gerentes de ABC para comprometer su apoyo en cada una de las etapas del proyecto, y por otro lado con las personas responsables de la administración de las bases de datos para estar seguro de contar con suficiente información histórica. A continuación en otra reunión más amplia, di a conocer al resto de la empresa los beneficios y alcance del trabajo que iba a emprender,

obteniendo de todos los involucrados su compromiso para la consecución de los objetivos planteados.

## **2.1 MINERÍA DE DATOS**

La minería de datos es un proceso analítico diseñado para explorar grandes volúmenes de información, con el objeto de descubrir patrones y modelos de comportamiento o relaciones entre diferentes variables. Desarrollar este proceso en ABC implicó cumplir con varias tareas que indico a continuación.

### **2.1.1 FILTRADO DE DATOS**

El formato de los datos contenidos en las bases no fue el idóneo, por ello filtré la información según las necesidades de la metodología escogida para desarrollar los modelos, depurando los datos duplicados, valores atípicos incorrectos y valores faltantes. También fue necesario un análisis univariado de cada una de las variables disponibles a través de técnicas estadísticas que me dieron un entendimiento básico de los datos, para identificar los valores atípicos que podían causar distorsiones en el estudio. La categorización de buen y mal cliente la expongo en detalle más adelante en el apartado correspondiente.

Este examen previo de los datos es un paso necesario que lleva tiempo y habitualmente se descuida, pero fue parte esencial de este trabajo, porque ayudó en la presentación formal de los resultados en ABC. Este análisis consta de manera detallada en el Anexo II.

Todos los datos que se ingresan a las bases provienen de dos fuentes: los usuarios de los diferentes módulos del sistema informático y del dueño de la cartera o Cedente.

**Información ingresada por los usuarios del sistema informático.**

El método de entrada de los datos que provienen de la interacción con los clientes en cobranza es a través de un sistema informático propio de ABC, mismo que dispone de módulos específicos diseñados para este propósito. Analicé 300 registros de gestiones de manera aleatoria para verificar la información registrada y pude concluir que aproximadamente un 3% de estas interacciones fueron ingresadas de forma incorrecta; por ejemplo en la gestión domiciliaria algunos ejecutivos de cobranza tenían la costumbre de dejar en blanco cuando la dirección no corresponde o no se encontró a ninguna persona en el domicilio, pero en el sistema se ingresa esta visita infructuosa como ausencia de gestión. Caso similar pasa con las llamadas telefónicas a los clientes.

**Información enviada por el dueño de la cartera o Cedente**

La base de datos también alberga información externa que previamente se acuerda y transfiere entre el dueño de la cartera o Cedente y ABC, desarrollando para esta tarea procesos informáticos automatizados. El Cedente de la cartera es una entidad bancaria que presta servicios financieros orientados al desarrollo del sector microempresarial; empezó sus operaciones en Ecuador en el año 1993 y actualmente tiene 20.000 clientes, con activos valorados en 18 millones de dólares. Como cualquier otro banco de este tipo busca su rentabilidad y crecimiento mediante la ampliación de cobertura e innovación de sus servicios, apoyándose en equipo humano y tecnología calificados.

Asumí que el proceso de entrada al sistema y gestión de la información por parte del Cedente estaba controlada, sobre todo aquella originada en la solicitud de crédito; sin embargo cuando analicé 300 casos de manera aleatoria pude encontrar por ejemplo que de todos los teléfonos ingresados al sistema un 7% son inválidos, es decir pasaron las validaciones del sistema del Cedente y cuando llegaron a cobranza no se los pudo utilizar. Esto puede conducir a decisiones erróneas, ocasionando pérdidas de tiempo, dinero y credibilidad, llevando a

costos importantes, como el sobreestimar el volumen de cobranza de los clientes, el exceso de gastos en los procesos de contacto con los clientes y pérdida de oportunidades de cobranza; sin embargo los datos que realmente necesitaba fueron filtrados y las fallas en la información de teléfonos o direcciones se consideraron como errores que afectan el proceso de gestión, más no a la construcción del modelo.

### **Selección de la ventana de muestreo**

La selección de la ventana de muestreo se realizó siguiendo los criterios de estabilidad, representatividad y madurez, es decir, la cartera seleccionada para la muestra debe ser estable en el sentido que una variable de interés en la población se mueva dentro de una franja no muy amplia, de forma que los datos de la muestra sean comparables con la actual población de clientes en mora y que la información disponible sea suficiente en tiempo para mostrar si el comportamiento de los clientes es bueno o malo.

Para medir la estabilidad de la muestra, definí a la morosidad como la característica que en primera instancia diferencia de manera más apropiada los buenos y malos clientes para ABC. Fue entonces necesario construir un indicador de mora.

**Indicador de mora.** Se calcula por mes de desembolso  $j$ , como la relación entre el número de clientes con un atraso máximo  $W$  y el total de clientes desembolsados en el mes  $j$ . El valor de  $W$  se estableció en uno, por una determinación directa del Cedente de la cartera, en relación de hasta qué punto la mora significa un riesgo significativo para sus intereses. De esta manera el indicador de mora queda definido por:

$$\text{Indicador de mora}_j = \frac{\text{Número de clientes del mes de desembolso } j, \text{ cuyo atraso mínimo es } \geq 1 \text{ día}}{\text{Número de clientes del mes de desembolso } j}$$

Por ejemplo, analizando para una fecha determinada que se conoce como cosecha, en el mes de junio del año 2010 se desembolsaron 809 microcréditos, de estos un 72.7% fueron pagados sin atraso hasta la fecha de vencimiento y el restante 27.3% ingresó en un tramo de morosidad con atraso de al menos un día, a una fecha determinada que se conoce como fecha de corte.

Como el indicador de mora va de la mano con la cobranza, fue necesario incluir información de la efectividad en la gestión de recuperación por cada tramo de morosidad, construyendo un indicador de efectividad en la cobranza.

**Efectividad en la cobranza.** Se calcula para el tramo  $r$ , como la relación entre el número de clientes que pagaron al menos una cuota dentro del período  $t$  y el total de clientes con mora en el tramo  $r$ .

$$\text{Efectividad cobranza}_r = \frac{\text{Clientes en el tramo } r \text{ que pagaron al menos una cuota dentro del período } t}{\text{Número de clientes vencidos en el tramo } r}$$

Siguiendo con el ejemplo de los 809 microcréditos desembolsados en el mes de junio del año 2010, 312 clientes pagaron la cuota mensual del crédito sin atraso; 232 clientes ingresaron a cobranza en el tramo de 1 a 30 días de atraso y al final del período de cobranza el 77% pagó al menos una cuota; 103 clientes ingresaron en el tramo de 31 a 60 días y el 52% pagó al menos una cuota; 63 clientes ingresaron en el tramo de 61 a 90 días y el 28% pagó; de la misma forma con el tramo de 91 a 120 días, donde ingresaron 40 clientes y pagó el 14%; 34 clientes al tramo de 121 a 150 días donde pagó el 12% y finalmente 25 clientes que ingresaron al tramo de 151 y más días donde pagó el 5% al menos una cuota.

Construí el cuadro de morosidad por mes de desembolso que se muestra en la Tabla 2.1, para mostrar cuándo la cartera se estabilizó y la efectividad de la cobranza en cada tramo de morosidad.

**Tabla 2.1-** Cuadro de morosidad y efectividad en la cobranza por mes de desembolso

Año de desembolso del crédito	Mes de desembolso del crédito	Total de clientes	Indicador de morosidad	Efectividad en la cobranza					
				Tramo de 1 a 30 días	Tramo de 31 a 60 días	Tramo de 61 a 90 días	Tramo de 91 a 120 días	Tramo de 121 a 150 días	Tramo de 151 días y más
2010	enero	105	19,0%	66%	50%	40%	33%	100%	-
	febrero	124	24,2%	53%	57%	33%	50%	0%	0%
	marzo	185	25,4%	53%	52%	30%	43%	25%	33%
	abril	601	34,8%	61%	48%	38%	19%	5%	5%
	mayo	891	29,1%	69%	57%	40%	17%	20%	5%
	junio	809	27,3%	77%	52%	28%	14%	12%	5%
	julio	761	24,2%	78%	50%	31%	17%	10%	6%
	agosto	917	27,3%	76%	52%	33%	19%	12%	8%
	septiembre	1051	26,8%	77%	54%	26%	14%	13%	7%
	octubre	1363	24,3%	78%	52%	30%	14%	11%	6%
	noviembre	1439	28,1%	77%	52%	31%	15%	9%	5%
	diciembre	1522	28,8%	74%	49%	32%	17%	8%	6%
2011	enero	1627	42,0%	64%	45%	27%	16%	10%	3%
	febrero	1516	41,3%	62%	40%	25%	15%	9%	4%
	marzo	1524	48,4%	58%	41%	28%	15%	10%	5%
	abril	1329	25,0%	57%	44%	26%	14%	11%	5%

Después de revisar la información de los créditos, supe que los préstamos iniciaron con plazos de 6 a 18 meses y después de un período de evaluación se redujo el plazo máximo a 15 meses. Actualmente el 70% de los microcréditos del Cedente son a 12 cuotas mensuales; esto depende usualmente del producto y del ambiente económico; por ejemplo, un microcrédito tenderá a madurar<sup>7</sup> mucho más rápido que un crédito de consumo o un crédito hipotecario, pero los tres se van a desempeñar de forma diferente en tiempos de inestabilidad económica.

A continuación investigué sobre actividades inusuales respecto a políticas de venta o aquellas emprendidas también por la empresa de cobranza; por ejemplo me enteré que no se implementaron cambios drásticos en las políticas de crédito pero sí existió, como es normal, controles de los procesos establecidos y se corrigieron errores a nivel operativo. La empresa de cobranza mantuvo siempre el esquema tradicional de recuperación basado en saldo capital y días mora.

<sup>7</sup> Un préstamo "madura" en la fecha final que debe ser pagado en su totalidad.

Finalmente pude seleccionar los meses para el desarrollo del modelo, que corresponden a los de microcréditos desembolsados entre junio y diciembre del 2010; la fecha de corte para extracción de la muestra se estableció en junio del 2011, lo que significa que la madurez de la ventana de muestreo fue entre 5 y 11 meses.

### **Definición de buenos y malos clientes para ABC**

Dar una definición adecuada de buenos clientes o de malos clientes requiere conocer a profundidad la cartera dentro de ABC, tanto aquella que está al día como la que ya presenta morosidad, así como también el proceso de cobranza que se realiza sobre ella. Esto último es particularmente importante cuando se trabaja con un esquema tradicional de recuperación, en donde la calificación de un cliente se hace en función del saldo capital y antigüedad en mora; por ejemplo uno con 30 días de atraso es considerado peor que uno con 5 días de atraso; un cliente con 500 usd de saldo en mora se considera peor que otro con 200 usd de saldo en mora.

La gestión de recupero moderna basada en CRM<sup>8</sup>, mejora estas conclusiones, indicando que puede ser apresurado calificar a un cliente como bueno o malo basándonos en su situación actual, sino que debe considerarse el comportamiento histórico, es decir, no es lo mismo un cliente con 30 días de mora que ha pagado puntualmente un 80% de su crédito que otro cliente con 5 días de mora que apenas empieza a pagar el préstamo; de forma similar tenemos el caso del cliente con 500 usd de saldo en mora pero que ya ha pagado 10 cuotas de su crédito en comparación con otro que no ha pagado ninguna cuota. La gestión basada en CRM considera además todas las acciones que se realicen para fidelizar, aumentar el margen por cliente y la gestión terminal.

---

<sup>8</sup> Customer Relationship Management es la administración basada en la relación con los clientes. CRM es un modelo de gestión de toda la organización, basada en la orientación al cliente.

Tomando en cuenta este enfoque moderno y utilizando el material que obtuve de las entrevistas con los responsables de la operación en ABC, pude formarme una idea clara de los atributos que debe tener un buen cliente y que nos permitió en conjunto definir el perfil de “buen cliente para ABC”. Para el desarrollo de este trabajo definimos como buen cliente, aquel que cumple todas las características mencionadas en dicho perfil y cualquier otro que no cumple con alguna de estas características se calificó como mal cliente para ABC.

1. Cliente al día o con mora menor a 15 días
2. Al menos 5 cuotas pagadas del último crédito vigente.
3. Máximo 12 gestiones telefónicas en los últimos dos meses, de las cuales al menos 1 fue de contacto con el cliente.
4. Clasificación de mora actual igual o menor a la mora máxima de los últimos tres meses.

Establecido el criterio para la calificación de buen y mal cliente, consideré a esta última como la variable dependiente del modelo, que es una variable dicotómica donde los buenos se presentarán como 0 (cero) y los malos con 1 (uno).

### **Obtención de la muestra**

Las expectativas de ABC respecto a los resultados de este trabajo me permitieron disponer de la información depositada en las bases de datos, manejando parámetros razonables de seguridad y confidencialidad<sup>9</sup>. La calidad de las bases y cumplimiento de las normas respecto a integridad expuestas anteriormente, evitó realizar procedimientos adicionales de captura de información incurriendo en nuevos costos. Estadísticamente se necesitaba para la construcción del modelo una muestra de aproximadamente 400 clientes, dado que se conoce el tamaño de la población (N=7862), asumiendo una distribución normal, y con un 95% de confiabilidad (Fórmula 2.1); sin embargo los recursos disponibles me permitieron

---

<sup>9</sup> Las políticas de seguridad en manejo de información de la empresa ABC, me autorizó utilizar el 50% de la base total de clientes en la ventana de muestreo.



utilizar una muestra mucho mayor, que luego de analizar la calidad de los datos quedó en 3700 clientes obtenidos aleatoriamente de la población de microcréditos en cobranza a nivel nacional.

$$n = \frac{Z^2 * p(1 - p)}{E^2} \quad (\text{Fórmula 2.1})$$

donde:

$n$  = tamaño de la muestra requerido, asumiendo que  $N$  es grande

$Z$  = nivel de confianza

$p$  = prevalencia esperada de ser mal cliente<sup>10</sup>

$E$  = precisión o error

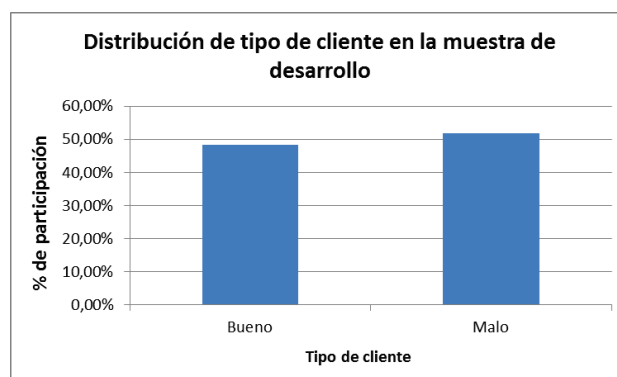
$$n = \frac{1,96^2 * 0,5(1 - 0,5)}{0,05^2} = 384,16 \sim 400$$

Es importante aclarar que la relación cliente:crédito es uno a uno, es decir un solo crédito por cada cliente, razón por la cual es indistinto hablar de cliente o crédito de ahora en adelante. Para la construcción del modelo se redujo la muestra original a 2931 clientes (80% \* 3700 clientes), de los cuales el 48,28% son buenos y el 51,72% son malos, tal como se muestra en la Tabla 2.2 y la Figura 2.1.

**Tabla 2.2** - Distribución de buenos y malos clientes sobre la muestra de desarrollo

Tipo de cliente	Número	Porcentaje
Bueno	1415	48,28%
Malo	1516	51,72%
Total	2931	100%

<sup>10</sup> Considerando el caso más desfavorable, es decir aquel que necesite el máximo tamaño de la muestra, tomé  $p=0,5$ .



**Figura 2.5** - Distribución de buenos y malos sobre la muestra de desarrollo

### Muestra de prueba

Como se puede obtener nuevos resultados utilizando el conocimiento extraído previamente sobre el desarrollo con la muestra del 80%, aproximadamente el 20% restante ( $20\% * 3700$  clientes) no se consideró en la construcción del modelo estadístico para al final poder validarlo, sabiendo con anticipación que podría ser necesario recalcular todas las estadísticas según se diera el caso. Fue necesario asegurarme que este 20% de clientes, seleccionados aleatoriamente de la muestra total, contenga una proporción de buenos y malos similar al de la muestra de desarrollo.

### 2.1.2 SELECCIÓN DE VARIABLES

Posterior al filtrado que realicé para normalizar la base de datos, fue necesario identificar aquellas variables que aportaban más información, agrupar las que aportaban poco y descartar las que no proporcionaron información relevante para el proyecto. Variables como por ejemplo *nacionalidad*, *existencia de verificación de la actividad laboral*, *profesión* y otras, no se tomaron en cuenta porque se encontraban escasamente pobladas o existían dudas acerca de su confiabilidad. Sin embargo fue importante investigar la posibilidad de crear nuevas características que resulten de cálculos basados en combinaciones de información ya existente y crear variables nuevas, como por ejemplo el número

de veces con mora mayor a 90 días que sirve para describir el comportamiento de pago de los clientes.

#### **Naturaleza de las variables del estudio.**

Existen variables que describen características propias del crédito y que fueron consideradas en su inicio para medir su riesgo, por ejemplo *saldo, cantidad de cuotas, valor cuota, edad, estado civil, nivel de educación, cargas familiares* y otras. Pero también existen variables propias de la gestión de cobranza, más ampliamente del proceso de mantenimiento de clientes que mencionaremos más adelante, que añaden valor a la cadena de la venta favoreciendo una recuperación eficiente y que determinan un perfil de buen o mal comportamiento de pago; ejemplo de estas variables son: la cantidad y tipo de teléfonos o direcciones del cliente, número de gestiones telefónicas y domiciliarias, cuántos contactos y compromisos de pago resultaron de la gestión. Finalmente también tomé en cuenta el comportamiento de la mora, por ejemplo la cantidad de cuotas pagadas, la conducta del atraso máximo y el número de ocasiones que el cliente ha tenido cuotas afectadas con morosidad.

Para la construcción de los modelos procedí a etiquetar las variables de acuerdo a su naturaleza en las diferentes etapas de la gestión de cobranza, obteniendo así los siguientes grupos:

**Variables de Inicio.** Son aquellas que se tomaron en cuenta para valorar el riesgo del cliente en el momento que ingresan a la gestión de recuperación. Ejemplo de estas variables son la *edad, cantidad de cargas, saldo promedio en el sistema financiero, cantidad de cuotas del crédito, nivel de educación.*

**Variables de comportamiento.** Son aquellas que resultan de la conducta de pago y que dan una idea del carácter del cliente y de la forma como éste hace honor a su deuda en el transcurso del tiempo. Ejemplo de estas variables son: *mora*

*máxima trimestral, mora máxima anual, mora anterior, contador de moras mayores a 15 días, cuotas pagadas, cuotas impagas.*

**VARIABLES DE GESTIÓN DE COBRANZA.** Son propias de las acciones que se realizan en la empresa de cobro para incentivar el pago de las obligaciones pendientes. Ejemplos de estas variables son la *cantidad de teléfonos vigentes, cantidad de llamadas telefónicas, cantidad de visitas domiciliarias, cantidad de compromisos de pago.*

Como ilustración incluyo la Figura 2.2 para esquematisar el papel de estos tres grupos de variables en la construcción de modelos.



**Figura 2.2** - Variables en la base de datos por su naturaleza

### **Recodificación de variables**

En el caso de las variables cuantitativas, comparé a través de una prueba t-Student los dos grupos que se establecieron por las categorías de la variable dependiente *calificación*, evaluando si había diferencia en las medias sobre cada una de estas variables independientes; pero esta metodología no permite calcular medidas de riesgo y por tanto tampoco intervalos de confianza, así que fue conveniente recurrir a una de las prácticas habituales en el procesamiento de los

datos como es la categorización de variables<sup>11</sup>, misma que además facilita la interpretación de los coeficientes de un modelo estadístico. Por este motivo procedí a dicotomizar las variables independientes cuantitativas, ayudándome de criterios empíricos que los obtuve de las entrevistas con los responsables de la operación en ABC, tales como tener o no gestión domiciliaria el último mes, existencia o no de teléfonos válidos, antes o después de pagar la mitad del crédito, si el cliente tiene la primera cuota en mora u otra; o también donde los análisis cuantitativos sugerían un punto de corte que facilite la interpretación de resultados, por ejemplo con el cálculo de los cuartiles.

**Tabla 2.3** - Variables recodificadas

VARIABLES		RECODIFICACIÓN		
TIPO DE VARIABLE	NATURALEZA	ORIGINAL	CATEGORIZADA	DICOTOMIZADA
<b>Nominal</b>	INICIALIZACIÓN	7	3	11
	COMPORTAMIENTO	1	1	15
	GESTIÓN	5	5	15
<b>Nominal Total</b>		<b>13</b>	<b>9</b>	<b>41</b>
<b>Ordinal</b>	INICIALIZACIÓN		7	
	COMPORTAMIENTO		8	
	GESTIÓN		4	
<b>Ordinal Total</b>			<b>19</b>	
<b>Continua</b>	INICIALIZACIÓN	8		
	COMPORTAMIENTO	15		
	GESTIÓN	11		
<b>Continua Total</b>		<b>34</b>		
<b>Grand Total</b>		<b>47</b>	<b>28</b>	<b>41</b>

En la Tabla 2.3 se puede observar como originalmente empecé el estudio con 47 variables, 13 cualitativas y 34 cuantitativas. De las primeras 13, 7 corresponden a variables de inicio, 1 a comportamiento y 5 a gestión de cobranza; de manera similar de las 34 variables cuantitativas, 8 son de inicio, 15 de comportamiento y 11 de gestión. Obtuve adicionalmente 28 variables cualitativas a partir de las variables originales y también 41 variables dicotómicas, terminando con 116 variables en total para los análisis.

<sup>11</sup> Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power of the cutpoint

En todo este proceso utilicé el paquete estadístico SPSS 15 con el objetivo de hacer posible el uso de los datos a cualquier técnica estadística. Con el fin de mejorar la disposición de los datos almacenados en la base para la realización del proyecto, creé un libro de códigos en el que se detallan los nombres de todas las variables utilizadas, su tipo de dato, rango de valores y significado dependiendo del tipo de variable. Aquellas que son categóricas se identifican como tales, de manera que la categoría de referencia “con menor riesgo” es la que se codifica como “cero”. El resultado se presenta en el Anexo II, Tabla A2.1.

Culminar con éxito esta actividad demandó una serie de reuniones con los encargados de la parte operativa, tecnológica y comercial de la empresa ABC, para conocer muy bien el tema de las cobranzas en general.

### **2.1.3 EXTRACCIÓN DEL CONOCIMIENTO**

En esta etapa definí los atributos más importantes que tienen las variables en función del poder discriminante que tiene cada una de ellos, de acuerdo a la definición de buenos y malos clientes, sobre la muestra seleccionada; obtuve un modelo de conocimiento que representa patrones de comportamiento observados en los valores de las variables consideradas o relaciones de asociación entre éstas.

Utilicé las variables recodificadas dicotómicas tomándolas de una en una para hacer valoraciones bivariantes y tener una primera estimación de la medida de asociación; luego pude clasificar estas variables según el valor de esta medida de asociación y de su significación estadística, aun reconociendo que pudieran tratarse de estimaciones sesgadas o poco informativas. La forma de evaluar esta asociación entre la variable independiente y dependiente la realicé con la construcción de tablas de contingencia en el paquete SPSS, donde analicé el valor del estadístico Chi-Cuadrado para contrastar la hipótesis de no asociación  $H_0$ , versus la hipótesis de relación entre las dos variables  $H_a$ , y puesto que se

trata de variables dicotómicas también calcular las medidas de riesgo. Los resultados se presentan en el Anexo II, Tabla A2.2.

### **Construcción de modelos**

Aquí utilicé regresión logística clásica sobre características de la cartera microcrédito en cobranza para calcular el riesgo de convertirse en un mal cliente. Con el fin de mejorar el análisis utilicé también métodos no paramétricos que se plantearon como complemento y alternativa al método logístico y que sirven para comparar y evaluar los resultados:

- Árboles de decisión
- Regresión Logística Basada en distancias
- Algoritmo K-NN (K-Nearest Neighbour)

#### **2.1.4 INTERPRETACION Y EVALUACION**

Una vez obtenido el modelo procedí con su validación, comprobando que las conclusiones que arroja son válidas y satisfactorias. El objetivo final fue asegurarme de tener el mejor modelo posible y que tenga capacidad de predicción, esto significa que por una parte necesitaba medir la capacidad de discriminación entre buenos y malos clientes que el modelo realiza y por otra parte conocer su estabilidad, es decir, comparar la capacidad de discriminación sobre una nueva muestra contra aquella que este modelo tenía cuando fue construido.

El modelo obtenido asigna un puntaje o nivel de riesgo a cada cliente y para rechazarlo su puntaje se compara con un punto de corte, de este modo el testeado de un modelo abarca tanto su estructura como el ajuste de su punto de corte. Un buen modelo es aquel donde el porcentaje de malos clientes que el modelo rechaza es cercano a 100% y el porcentaje de buenos clientes que el modelo rechaza es cercano a 0%.

## 2.2 CONSTRUCCIÓN DE MODELOS UTILIZANDO REGRESIÓN LOGÍSTICA CLÁSICA

En esta parte del trabajo mi objetivo fue la predicción del tipo de cliente para ABC, relacionado con el atributo “malo” de la variable *calificación* en la base de datos, encontrando una fórmula de cálculo para este resultado. Generé distintos modelos con regresión logística, considerando diferentes selecciones de variables de inicio, comportamiento y gestión para lograr una buena discriminación de la variable *calificación*. De cada modelización obtuve distintas ecuaciones, algunas de ellas con una marcada diferencia en cuanto al poder de predicción para clientes buenos y malos. A continuación en la Tabla 2.4 presento de forma resumida el desarrollo de los diferentes modelos y en la Tabla 2.5 los resultados obtenidos en cada uno de ellos.





**Tabla 2.5** - Resultados de los modelos utilizando regresión logística

MODELO	NÚMERO DE VARIABLES DE ENTRADA			MÉTODO MANUAL			AUTOMÁTICO HACIA ADELANTE			AUTOMÁTICO HACIA ATRÁS								
	INICIO	GESTIÓN	COMPORTAMIENTO	RESULTADOS DEL MODELO			RESULTADOS DEL MODELO			NÚMERO DE VARIABLES DE SALIDA								
				% CLASIFICADOS	Cox & Snell R Square	Nagelkerke R Square	% CLASIFICADOS	Cox & Snell R Square	Nagelkerke R Square	INICIO	GESTIÓN	COMPORTAMIENTO	% CLASIFICADOS	Cox & Snell R Square	Nagelkerke R Square	INICIO	GESTIÓN	COMPORTAMIENTO
1		22		92,6%	0,637	0,849	92,2%	0,632	0,843	0	9	5	92,4%	0,636	0,848	0	13	8
2		15		92,3%	0,633	0,845	91,5%	0,63	0,841	0	5	8	92,1%	0,632	0,844	0	6	13
3	11			77,9%	0,351	0,468	77,7%	0,35	0,467	6	0	0	77,4%	0,351	0,468	7	0	0
4	11			77,4%	0,371	0,495	77,4%	0,369	0,492	6	0	0	77,6%	0,371	0,495	8	0	0
5		17		85,7%	0,445	0,594	85,5%	0,444	0,593	0	12	0	85,7%	0,445	0,593	0	13	0
6		15		85,0%	0,41	0,546	85,0%	0,407	0,543	0	9	0	85,0%	0,409	0,546	0	11	0
7			15	89,4%	0,578	0,771	89,4%	0,573	0,764	0	0	6	89,3%	0,578	0,771	0	0	11
8			15	88,7%	0,586	0,782	88,4%	0,584	0,78	0	0	10	88,7%	0,586	0,782	0	0	13
9			14	88,7%	0,586	0,782	88,4%	0,584	0,78	0	0	10	88,7%	0,586	0,782	0	0	13
10	11	17		88,8%	0,559	0,746	88,6%	0,557	0,743	6	12	0	88,9%	0,559	0,745	7	14	0
11	11	15		87,6%	0,54	0,72	87,5%	0,537	0,717	5	8	0	87,4%	0,538	0,718	6	10	0
12	12		15	90,7%	0,6	0,8	90,3%	0,594	0,792	2	0	6	90,3%	0,598	0,798	3	0	10
13	12		15	90,3%	0,609	0,812	90,2%	0,606	0,809	3	0	10	90,5%	0,608	0,811	5	0	13
14	12	17	15	93,3%	0,652	0,869	93,1%	0,648	0,864	2	9	7	93,2%	0,651	0,868	2	9	7
15	12	17	14	93,4%	0,652	0,869	93,3%	0,649	0,865	3	6	8	93,4%	0,65	0,867	4	6	10
16	13	14	14	94,5%	0,671	0,896	93,2%	0,651	0,865	10	21	3	94,4%	0,668	0,891	15	36	7
17	12	15	15	93,4%	0,652	0,869	93,3%	0,649	0,865	3	6	8	93,4%	0,65	0,867	4	6	10
18	8	11	14	93,1%	0,651	0,868	93,3%	0,649	0,865	3	6	8	93,1%	0,649	0,866	4	6	9
19	14	14	15	95,6%	0,687	0,916	93,9%	0,664	0,886	36	8	3	95,1%	0,679	0,906	36	26	8
20	13	15	16	94,0%	0,666	0,889	93,8%	0,664	0,885	27	6	8	94,0%	0,664	0,886	27	6	10
21	3	5	6	93,4%	0,646	0,862												

### 2.2.1 SELECCIÓN DEL MEJOR MODELO

Los criterios para la selección del modelo más adecuado, tomando en cuenta la noción de negocio para ABC, fueron los siguientes:

1. El mejor valor de los mayores al 90% de bien clasificados.
2. La menor cantidad de variables posible, máximo 20.
3. Que tenga los tres grupos de variables: de inicio, de comportamiento y de gestión.

Los modelos que cumplen estos criterios son el 12, 13, 14, 15, 17 y 18. De estos los que mejor clasificación ofrecen fueron el 15 y el 17, ambos con el mismo número de variables, pero al final me decidí por el 17 con el método de selección automática hacia adelante, porque considera un mayor número de variables dicotómicas, mismas que son más fáciles de interpretar y mostraron ser mejores al momento de estimar la proporción de varianza explicada, calculando los

coeficientes de determinación de Cox-Snell y el de Nagelkerke. A continuación incluyo los resultados de este modelo en la Tabla 2.6.

**Tabla 2.6 - Resultados del modelo 17.**

		Variables en la ecuación						I.C. 95% para EXP(B)	
		B	E.T.	Wald	gl	Sig.	Exp(B)	Inferior	Superior
Paso 17a	estadoCiv_rec(1)	1,394	,180	59,688	1	,000	4,030	2,830	5,739
	tipoViv_rec(1)	1,332	,182	53,805	1	,000	3,790	2,655	5,411
	cuota_rec(1)	-1,077	,315	11,691	1	,001	,341	,184	,631
	contador1_rec(1)	3,929	,785	25,021	1	,000	50,847	10,907	237,042
	contador30_rec(1)	1,159	,263	19,393	1	,000	3,186	1,902	5,337
	contador60_rec(1)	3,198	,717	19,889	1	,000	24,477	6,004	99,791
	contador15_rec(1)	-2,815	,734	14,704	1	,000	,060	,014	,253
	Porcentaje_rec(1)	2,989	,233	164,381	1	,000	19,864	12,579	31,370
	moraTri_rec(1)	4,176	1,352	9,541	1	,002	65,123	4,601	921,709
	moraAno_rec(1)	-3,470	1,359	6,523	1	,011	,031	,002	,446
	telefonosPar_rec(1)	1,835	,211	75,824	1	,000	6,266	4,146	9,470
	vectorMor_rec(1)	,765	,373	4,217	1	,040	2,150	1,035	4,463
	gestionesDom_ant_rec(1)	-1,529	,282	29,327	1	,000	,217	,125	,377
	gestionesTel_tri_rec(1)	2,487	,555	20,093	1	,000	12,027	4,054	35,683
	direccionesLab_rec(1)	-1,014	,406	6,250	1	,012	,363	,164	,803
	telefonosLab_rec(1)	2,113	,186	129,049	1	,000	8,275	5,747	11,916
	gestionHist_rec(1)	-1,214	,524	5,367	1	,021	,297	,106	,830
	Constante	-3,906	,453	74,390	1	,000	,020		

En la tabla anterior se puede apreciar que las significancias de cada uno de los coeficientes es menor al 5%, por lo que acepté que no existen coeficientes estadísticamente nulos con un 95% de confianza.

## 2.2.2 VALIDACIÓN DEL MODELO

La validación fue realizada a través de dos técnicas: tablas de clasificación y pruebas de eficiencia.

### Tabla de Clasificación

Una de las formas de evaluar el ajuste del modelo es comprobar cómo clasifica a los individuos de la muestra en comparación con lo real. Para construir una tabla de clasificación o matriz de confusión consideré un punto de corte igual a 0,5 y a partir de este valor realicé las clasificaciones de buenos y malos; de esta forma el

modelo predice correctamente el 93,3% (Tabla 2.7), la proporción de variabilidad explicada es 65% según el R cuadrado de Cox y Snell y 87% según el R cuadrado de Nagelkerke.

**Tabla 2.7** - Distribución de buenos y malos, modelo 17

**Tabla de clasificación<sup>a</sup>**

Observado			Pronosticado		
			califica		Porcentaje correcto
			BUENO	MALO	
Paso 17	califica	BUENO	1331	84	94,1
		MALO	111	1405	92,7
		Porcentaje global			93,3

a. El valor de corte es ,500

## Pruebas de eficiencia

### Coefficiente Gini

Mide la eficacia del modelo, comparando el porcentaje de los buenos contra el porcentaje de los malos clientes para los mismos puntajes. Si el porcentaje de malos se traza contra el porcentaje de buenos para una serie de bandas de puntajes, el resultado es una curva. El coeficiente Gini se calcula como sigue:

$$\text{Gini} = 1 - \sum_{i=1}^n ((b_i - b_{i-1})(g_i + g_{i-1})), \quad b_0 = g_0 = 0$$

donde:

$n$  = tamaño de la muestra

$g_i$  = % acumulado de buenos en un puntaje dado

$g_{i-1}$  = % acumulado de buenos en el puntaje anterior a  $g_i$

$b_i$  = % acumulado de malos en un puntaje dado

$b_{i-1}$  = % acumulado de los malos en el puntaje anterior a  $b_i$

En la práctica  $i$  representa el percentil  $i$ , es decir se divide a toda la población en percentiles y para cada percentil se calcula el % acumulado de buenos y malos clientes, tomando en cuenta que el modelo “ideal”, es decir, el que predice exactamente los buenos y malos clientes, tendría un índice de Gini igual a 1. El resultado se muestra en la Tabla 2.8.

### Coeficiente K-S

Es la medida de la máxima diferencia absoluta entre las distribuciones acumuladas de buenos y malos. El índice de Gini es únicamente una medida global de calidad del modelo, mientras que en el coeficiente K-S, aparte de medir la calidad de ajuste, identifica el valor del score para el cual se maximiza este coeficiente. Al final del capítulo dos muestro la utilidad de este procedimiento utilizando curvas ROC. El resultado se muestra en la Tabla 2.8.

**Tabla 2.8** - Coeficiente Gini y Coeficiente K-S para el modelo 17

RANGO SCORE	% de buenos		% de malos		K-S	Gini
	% buenos	acumulado	% de malos	acumulado		
1 a 10	24,66%	24,66%	0,20%	0,20%	24,47%	0,05%
11 a 20	23,18%	47,84%	0,13%	0,33%	47,51%	0,08%
21 a 50	16,18%	64,03%	0,59%	0,92%	63,10%	0,12%
51 a 170	17,53%	81,55%	1,78%	2,70%	78,85%	0,42%
171 a 500	12,86%	94,42%	4,75%	7,45%	<b>86,96%</b>	0,84%
501 a 540	0,42%	94,84%	0,92%	8,38%	86,46%	0,02%
541 a 940	4,45%	99,29%	16,49%	24,87%	74,43%	0,78%
941 a 1000	0,71%	100,00%	75,13%	100,00%	0,00%	0,65%
Total general	100,00%		100,00%			
						<b>97,05%</b>

De estudios anteriores realizados en el mercado por la empresa Lisim<sup>12</sup>, supe que de su experiencia un coeficiente Gini debe estar en un valor cercano al 70% para un modelo de cobranza, dependiendo de la calidad y cantidad de información

<sup>12</sup> Compañía internacional de consultoría, experta en el diseño de soluciones de scoring.

disponible; de la misma forma el K-S debe ser mayor al 40%, por lo que concluí que los resultados obtenidos para el modelo de ABC eran consistentes.

### 2.2.3 VALIDACIÓN CON LA MUESTRA DE PRUEBA

Utilizando los mismos criterios que apliqué sobre la muestra de desarrollo, pude encontrar variables que siendo seleccionadas en la construcción del modelo, no resultaron ser significativas a partir de las pruebas de hipótesis correspondientes sobre la muestra de prueba y considerarlas ocasionaría un error de estimación y de interpretación del resultado por lo que fueron eliminadas.

Como medidas de calidad de ajuste sobre la muestra de prueba, analicé el coeficiente Gini y el coeficiente K-S que se muestran en la Tabla 2.9, y la clasificación de buenos y malos clientes en la Tabla 2.10.

**Tabla 2.9** - Coeficiente Gini y coeficiente K-S aplicado sobre la muestra de prueba

RANGO SCORE	% de buenos		% de malos		K-S	Gini
	% buenos	acumulado	% de malos	acumulado		
1 a 20	15,87%	15,87%	6,09%	6,09%	9,79%	0,97%
21 a 50	25,40%	41,27%	7,83%	13,91%	27,36%	3,53%
51 a 60	11,51%	52,78%	2,61%	16,52%	36,26%	1,20%
61 a 80	17,86%	70,63%	2,17%	18,70%	51,94%	0,85%
81 a 110	5,95%	76,59%	1,74%	20,43%	56,15%	0,23%
111 a 210	13,89%	90,48%	7,83%	28,26%	<b>62,22%</b>	1,33%
211 a 589	2,38%	92,86%	17,83%	46,09%	46,77%	0,61%
591 a 910	2,78%	95,63%	20,87%	66,96%	28,68%	1,07%
911 a 1000	4,37%	100,00%	33,04%	100,00%	0,00%	2,35%
Total general						<b>87,85%</b>

**Tabla 2.10** - Distribución de buenos y malos sobre la muestra de prueba

Tabla de clasificación				
Observado	Pronosticado			
	califica		Porcentaje correcto	
	BUENO	MALO		
califica	BUENO	233	19	92,5%
	MALO	92	138	60,0%
Porcentaje global				77,0%

Se puede observar que tanto el coeficiente K-S como el coeficiente Gini son menores a los obtenidos con la muestra de construcción y de la misma forma la discriminación del modelo ya que el porcentaje de bien clasificados se reduce del 93,3% al 77%.

## 2.2.4 PRESENTACIÓN DE RESULTADOS DEL MODELO LOGÍSTICO

Con todos los resultados previos volví a la fase de desarrollo para llegar al modelo final número 21, como se muestra a continuación en la Tabla 2.11 junto con el porcentaje de buenos y malos clientes en la Tabla 2.12.

**Tabla 2.11** - Distribución de buenos y malos modelo final

Observado	Pronosticado		
	califica		Porcentaje correcto
	BUENO	MALO	
Paso 1 califica BUENO	1336	79	94,4
MALO	113	1403	92,5
Porcentaje global			93,4

**Tabla 2.12** - Presentación de resultados modelo final utilizando regresión logística clásica

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Paso 1 <sup>a</sup> moraTrn(1)	,932	,220	17,949	1	,000	2,540	1,650	3,909
Percent(1)	2,981	,232	165,337	1	,000	19,710	12,512	31,047
estadoC(1)	1,374	,179	58,980	1	,000	3,952	2,783	5,611
tipoViv_(1)	1,336	,180	55,305	1	,000	3,803	2,674	5,408
gestionH(1)	-1,155	,523	4,885	1	,027	,315	,113	,877
cuota_re(1)	-1,033	,310	11,130	1	,001	,356	,194	,653
telefonoLa(1)	2,022	,182	123,481	1	,000	7,557	5,290	10,796
telefonoPa(1)	1,794	,208	74,275	1	,000	6,015	4,000	9,045
contado1(1)	3,861	,717	29,041	1	,000	47,525	11,669	193,561
contad15(1)	-2,342	,682	11,792	1	,001	,096	,025	,366
contad30(1)	1,330	,256	26,924	1	,000	3,780	2,287	6,246
contad60(1)	2,475	,561	19,502	1	,000	11,886	3,962	35,657
gestDom(1)	-1,532	,281	29,694	1	,000	,216	,125	,375
gestiTel(1)	2,449	,553	19,628	1	,000	11,577	3,918	34,208
Constante	-3,970	,448	78,571	1	,000	,019		

Almacené los valores de probabilidad pronosticados por el modelo de regresión logística para cada individuo de la muestra, en una nueva variable continua con el nombre pre\_1, que representa la probabilidad de convertirse en mal cliente obtenida con el modelo de regresión logística.

## **2.3 MEJORAMIENTO DEL MODELO LOGÍSTICO UTILIZANDO ÁRBOLES DE DECISIÓN.**

Aquí mi objetivo fue mejorar la clasificación obtenida con el modelo logístico para los clientes en mora, sobre los que se realiza una gestión de recuperación en ABC conforme si representan o no un riesgo de deteriorar su comportamiento de pago. Para esto utilicé la técnica de clasificación denominada Árboles de Decisión usando el paquete Answer Tree 3.0 como herramienta de análisis y las variables del modelo logístico final junto con las probabilidades obtenidas con el mismo.

Un árbol de decisión es un método no paramétrico de segmentación donde se divide repetidamente los datos en grupos mutuamente excluyentes, las ramas representan conjuntos de decisiones y cada decisión genera reglas sucesivas para continuar la clasificación (partición), formando grupos homogéneos respecto a la variable que se desea discriminar. Las particiones se hacen en forma recursiva hasta que se alcanza un criterio de parada y el árbol resultante se usa para clasificar nuevos datos. El objetivo es particionar la respuesta en grupos homogéneos y a la vez mantener el árbol razonablemente pequeño.

Para dividir los datos se requiere un criterio de particionamiento que determinará la medida de impureza; esta última establece el grado de homogeneidad entre los grupos.

### **Método Chaid**

Como todas las variables son categóricas, excepto la probabilidad obtenida con el modelo de regresión logística que es continua, decidí escoger el método chaid porque se puede trabajar con variables en cualquier nivel de medida, además de ser el algoritmo más utilizado en la creación de árboles. Dada una variable predictora, este método fusiona aquellas categorías consideradas estadísticamente homogéneas y deja las categorías heterogéneas inalteradas.



Seguidamente, de todas las variables predictoras potenciales elige la que tenga mayor valor del coeficiente para formar la primera rama del árbol, esto es, si la variable criterio es continua se utiliza la prueba F y si la variable criterio es categórica se utiliza la prueba Chi-cuadrado.

### 2.3.1 CONSTRUCCIÓN DEL ÁRBOL

Primero definí los parámetros para tener la amplitud necesaria en el desarrollo guiado del árbol, de la siguiente manera:

- Árbol sin validación
- Número de niveles: 6.
- Nodo parental: 5.
- Nodo filial: 1.

La variable criterio es calificación, y el resto son variables predictoras que salieron del modelo logístico. Analicé el valor del estadístico Chi-cuadrado (Tabla 2.13) para tener una idea de la relación entre las variables independientes y el resultado, criterio que me sirvió como guía para la segmentación del árbol.

**Tabla 2.13** - Variables consideradas para el desarrollo del árbol de decisión

variable	medida	definición	categoría expuesta	Chi-cuadrado
PRE_1	cuantitativa	probabilidad obtenida del modelo logístico	N/A	2.309,52
contador30_rec	categórica dicotómica	Tener al menos 1 mora entre 31 y 60 días en los últimos 12 meses	si	1.640,84
contador1_rec	categórica dicotómica	Siempre con mora menor o igual a 15 días en los últimos 12 meses	no	1.558,77
moraTri_rec	categórica dicotómica	Mora mayor a 30 días en el último trimestre	si	1.192,67
telefonosLab_rec	categórica dicotómica	Tener al menos 1 teléfono laboral vigente	no	877,64
contador60_rec	categórica dicotómica	Tener al menos 1 mora entre 61 y 90 días en los últimos 12 meses	si	875,38
contador15_rec	categórica dicotómica	Tener al menos 1 mora entre 16 y 30 días en los últimos 12 meses	si	761,15
cuota_rec	categórica dicotómica	Mayor a 47 usd	si	662,38
Porcentaje_rec	categórica dicotómica	Porcentaje de cuotas por pagar mayor a 67%	si	594,04
telefonosPar_rec	categórica dicotómica	Tener al menos 1 teléfono particular vigente	no	396,25
estadoCiv_rec	categórica dicotómica	Ser casado, viudo, divorciado o en unión libre	soltero	250,43
tipoViv_rec	categórica dicotómica	Tener casa propia, hipotecada, herencia o anticresis	arrendada	243,15
gestionesDom_ant_rec	categórica dicotómica	Tener al menos 1 gestión domiciliaria el mes anterior	no	126,98
gestionesTel_tri_rec	categórica dicotómica	Tener al menos 1 gestión telefónica el mes anterior	no	8,54
gestionHist_rec	categórica dicotómica	Haber tenido solo gestión de cobranza telefónica	otros tipos de gestión	4,43

### Desarrollo del nodo raíz

Antes de desarrollar el nodo raíz pude observar que en el mismo la proporción de clientes malos es del 51.72%. Observando la vista <Riesgos> del paquete Answer Tree, la estimación del riesgo a priori, es decir el porcentaje de mal clasificados antes de hacer cualquier desarrollo en el árbol, es del 48,28%. También fue muy importante definir con anticipación los criterios de parada para determinar si un nodo es terminal o no lo es; esta definición por su importancia la tomé en conjunto con los responsables de la operación en ABC, porque su conocimiento del negocio orientó bien el grado de sensibilidad que debía tener el árbol.

### Criterios de parada para nodos terminales

1. Diferencia entre el porcentaje de buenos y malos clientes del 100%.
2. Si el nodo contiene hasta el 10% de los datos de la muestra y la diferencia entre el porcentaje de buenos y malos es superior al 90%.
3. Que en la siguiente rama se mejore el rendimiento de la clasificación de malos clientes, al menos un 20%.

Al final pude obtener los resultados mostrados en la Tabla 2.14 para la variable dependiente *calificación*.

**Tabla 2.14** - Distribución de buenos y malos con el desarrollo del árbol

Matriz de clasificación errónea				
		Categoría real		
		BUENO	MALO	Total
Categoría estimada	BUENO	1348	85	1433
	MALO	67	1431	1498
	Total	1415	1516	2931
Estadísticos de riesgo				
Estimación de riesgo		0,0518594		
ET de la estimación de riesgo		0,00409583		

El riesgo de realizar una clasificación errónea con el modelo final es del 5.2%, hay 85 clientes pronosticados como buenos que realmente son malos y 67 clientes malos clasificados como buenos.

### Reglas de clasificación

Las reglas indican las sucesivas selecciones que se han realizado para crear un nodo determinado. Por ejemplo, el nodo 16 está constituido por aquellos que tienen probabilidad del modelo logístico entre 0.537 y 0.930, con ninguna gestión telefónica en los últimos tres meses. Información como esta me permitió comprender el modelo de clasificación y establecer los nodos que definen a los buenos y malos clientes (Tabla 2.15).

**Tabla 2.15** - Definición de nodos buenos y malos

NODOS DE BUENOS	9	1	22	34	19	29	31	25	18	39	35
NODOS DE MALOS	8	7	6	16	28	38	36	32			

Las reglas de selección que se generaron en formato SPSS se presentan en el Anexo III y fueron utilizadas en una ventana de sintaxis para clasificar nuevos casos. Con los valores calculados obtuve una nueva variable continua prob\_001, que es la probabilidad de convertirse en mal cliente, mejorando la probabilidad obtenida con el modelo de regresión logística utilizando Árboles de Decisión.

### 2.3.2 VALIDACIÓN SOBRE LA MUESTRA DE PRUEBA

Aplicando las reglas de clasificación del árbol a la muestra de prueba obtuve los resultados mostrados en la Tabla 2.16, donde pude notar más claramente la ganancia de utilizar este método combinado con regresión logística; el porcentaje de bien clasificados pasa del 77% al 84%, es decir hay una ganancia de 7 puntos porcentuales.

**Tabla 2.16** - Distribución de buenos y malos con el desarrollo del árbol aplicado sobre la muestra de prueba.

Tabla de clasificación					
Observado		Pronosticado			
		califica		Porcentaje correcto	
		BUENO	MALO		
Paso 1	califica	BUENO	207	45	82,1%
		MALO	31	199	86,5%
Porcentaje global					84,2%

## 2.4 LA DISTANCIA EN ESTADÍSTICA

Desde su principio, la estadística moderna ha dependido de la teoría de probabilidad, del análisis, de la teoría de la medida y del álgebra; la metodología estadística no podría avanzar sin los recursos que proporcionan estas áreas de la matemática. También desde los principios la geometría, y especialmente las propiedades topológicas derivadas del concepto de distancia, han desempeñado un papel importante en estadística aunque su incorporación como elemento de trabajo es más reciente. (Ver Anexo V)

### Distancias

Una distancia  $\bar{d}$  sobre un conjunto (finito o no)  $\Omega$  es una aplicación que a cada par de individuos  $(\omega_i, \omega_j) \in \Omega \times \Omega$ , le hace corresponder un número real  $\bar{d}(\omega_i, \omega_j) = \bar{d}_{ij}$ , que cumple con las siguientes propiedades básicas:

**P.1.**  $\bar{d}_{ij} \geq 0$

**P.2.**  $\bar{d}_{ii} = 0$

**P.3.**  $\bar{d}_{ij} = \bar{d}_{ji}$

Cuando, además, se cumple la desigualdad triangular

**P.4.**  $\bar{d}_{ij} \leq \bar{d}_{ik} + \bar{d}_{kj}$ ,

se dice que la distancia es métrica.

### Similaridades en general

En muchas aplicaciones es conveniente trabajar con similaridades, concepto dual al de distancias. Una similaridad  $s$  en un conjunto  $\Omega$ , es una aplicación que asigna a cada par  $(\omega_i, \omega_j) \in \Omega \times \Omega$ , un número real  $s_{ij} = s(i, j)$  que cumple:

1.  $0 \leq s_{ij} \leq s_{ii} = 1.$
2.  $s_{ij} = s_{ji}$

Cuando  $\Omega$  es un conjunto finito, entonces la matriz se denomina matriz de similaridades sobre  $\Omega$ .

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix}$$

Es inmediato pasar de similaridad a distancia y recíprocamente. Las dos transformaciones básicas son:

$$\begin{aligned} \bar{d}_{ij} &= 1 - s_{ij} \\ \bar{d}_{ij} &= \sqrt{1 - s_{ij}} \end{aligned}$$

En general una matriz de similaridades puede tener en su diagonal elementos  $s_{ii} \neq 1$ . La transformación que nos permite pasar de similaridad a distancia es:

$$\bar{d}_{ij} = \sqrt{s_{ii} + s_{jj} - 2s_{ij}}$$

Si las variables son mixtas, contínuas, binarias o cualitativas, es entonces adecuado utilizar la distancia de Gower,  $d_{ij}^2 = 1 - s_{ij}$ , siendo

$$s_{ij} = \left( \sum_{h=1}^{p_1} (1 - |x_{ih} - x_{jh}| G_h) + a + \alpha \right) / (p_1 + (p_2 - d) + p_3)$$

Una similaridad, donde  $p_1$  es el número de variables cuantitativas;  $a$  y  $d$  corresponden al número de coincidencias y no coincidencias para las  $p_2$  variables binarias, respectivamente; y  $\alpha$  es el número de coincidencias para las  $p_3$  variables cuantitativas.  $G_h$  es el rango de la  $h$ -ésima variable cuantitativa.

### **Aplicación de las distancias en estadística<sup>13</sup>**

Muchos métodos de estadística y análisis de datos utilizan el concepto geométrico de distancia entre individuos, entre poblaciones y de un individuo a una población. Esto es especialmente cierto en técnicas de representación de datos (análisis de correspondencias, análisis de coordenadas principales, análisis de proximidades), donde la distancia entendida como medida de diferenciación entre objetos constituye la base fundamental de la presentación de los resultados. Las distancias aparecen también en muchos otros aspectos de la estadística como son el contraste de hipótesis, la regresión, el análisis discriminante, etc.

### **Predicción con variables continuas, categóricas y mixtas**

El modelo basado en distancias se reduce al modelo clásico de regresión cuando la distancia utilizada es la euclídea y las variables son continuas. La situación cambia si las variables son mixtas, esto es, una mezcla de continuas, binarias y categóricas; entonces la distancia no es euclídea y una buena elección consiste en emplear la distancia de Gower. Existen muchos ejemplos que prueban que utilizando el método de regresión con esta distancia se puede obtener mejores predicciones que con el método clásico.<sup>14</sup>

---

<sup>13</sup> Para información más detallada del concepto de distancia y sus aplicaciones, referirse al documento “Enfoque basado en Distancias de algunos Métodos Estadísticos Multivariantes”, Josep Fortiana, October 30, 2001.

<sup>14</sup> Ver “Métodos Multivariantes basados en Distancias”, C.M. Cuadras, Enero 2007.

### 2.4.1 APLICACIÓN DE LA REGRESIÓN BASADA EN DISTANCIAS PARA CALCULAR EL RIESGO DE CONVERTIRSE EN UN MAL CLIENTE DENTRO DE LA EMPRESA DE COBRANZAS ABC.

Como aplicación desarrollé un modelo de regresión basado en la disimilaridad de Gower para calcular el riesgo de convertirse en un mal cliente dentro de la gestión de cobranza en la empresa ABC, utilizando para esto las funciones del paquete estadístico R. Los datos consisten en el conjunto de variables originales que se obtuvieron del análisis univariado: 34 continuas y 13 cualitativas que hacen de regresoras para la variable dependiente *calificacion* (Tabla 2.17). El número de observaciones en la muestra de desarrollo es 2931 y los resultados obtenidos se aplicaron luego sobre la muestra de prueba con 482 casos para validar el modelo.

**Tabla 2.17** - Variables del modelo logístico basado en distancias

Tipo de variable	Naturaleza	Número
Nominal	Inicio	7
	Comportamiento	1
	Gestión de cobranza	5
Total de variables nominales		13
Continua	Inicio	8
	Comportamiento	15
	Gestión de cobranza	11
Total de variables continuas		34
<b>Total general</b>		<b>47</b>

#### Presentación de resultados

A continuación presento en la Tabla 2.18 la salida del paquete R con el modelo que resultó de aplicar la regresión logística basada en distancias a la muestra de prueba, también el porcentaje de bien clasificados (Tabla 2.19) que en este caso resultó ser 88%. La construcción del modelo se presenta en el Anexo IV.

**Tabla 2.18** - Presentación de resultados de la regresión logística basada en distancias en la consola de R

```

R Console
Call:
glm(formula = Y ~ Xr, family = binomial, data = tesis[, -1])

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.2813 -0.3909  0.0186  0.2226  2.7673

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.82657    0.08409   9.830 < 2e-16 ***
Xr1          24.48536    0.97279  25.170 < 2e-16 ***
Xr2         -16.42474    1.10542 -14.858 < 2e-16 ***
Xr3           3.71060    0.94352   3.933 8.40e-05 ***
Xr4          -6.92023    0.94170  -7.349 2.00e-13 ***
Xr5          -7.34058    1.19612  -6.137 8.41e-10 ***
Xr6          -7.95772    1.23585  -6.439 1.20e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 4059.7  on 2930  degrees of freedom
Residual deviance: 1593.9  on 2924  degrees of freedom
AIC: 1607.9

```

**Tabla 2.19** - Distribución de buenos y malos de la regresión basada en distancias sobre la muestra de prueba

Tabla de clasificación				
Observado		Pronosticado		
		califica		Porcentaje correcto
		BUENO	MALO	
califica	BUENO	228	24	90,5%
	MALO	34	196	85,2%
Porcentaje global				88,0%

## 2.4.2 SISTEMAS DE CLASIFICACION BASADOS EN DISTANCIAS

En la parte final de este capítulo construí un nuevo modelo denominado clasificador, que utiliza la distancia como criterio de comparación principal para



desarrollar el concepto de aprendizaje sobre un conjunto de datos de entrenamiento, estableciendo la conexión entre las variables que son dadas y las que se van a predecir, y que al final se aplica sobre otro conjunto de prueba para estimar el error de generalización.

Encontré en la herramienta WEKA<sup>15</sup> una interfaz gráfica adecuada para facilitar el análisis, tanto en el desarrollo del aprendizaje como en la verificación de lo aprendido utilizando las pruebas de validación, tales como muestra sustituta, validación cruzada y partición porcentual.<sup>16</sup>

### **Concepto de aprendizaje**

Una de las tareas más desafiantes en la ciencia de la computación es construir máquinas o programas de computadoras que sean capaces de aprender, con lo que se abre una amplia gama de nuevas aplicaciones, además de entender las capacidades y limitaciones humanas de aprendizaje. El aprendizaje no solo se encarga de obtener el conocimiento, sino también de la forma en que este se representa.

Un programa de computadora se dice que aprende la experiencia  $E$  con respecto a una clase de tareas  $T$  y medida de desempeño  $D$ , si su desempeño en las tareas en  $T$ , medidas con  $D$ , mejoran con la experiencia  $E$ . (Mitchell, 1997)<sup>17</sup>

Existe una gran cantidad de técnicas para el aprendizaje de Sistemas de Clasificación:  $k$  vecinos más cercanos, discriminadores bayesianos, árboles de clasificación, sistemas basados en reglas, redes neuronales, máquinas de soporte vectorial.

---

<sup>15</sup> Waikato Environment for Knowledge Analysis es una plataforma de software para aprendizaje automático y minería de datos escrito en Java y desarrollado en la Universidad de Waikato. Weka es un software libre.

<sup>16</sup> Weka proporciona tres modos de prueba para un clasificador: aplicar sobre un nuevo fichero de datos (fichero sustituto), realizar  $k$  particiones de la muestra y tomar  $k-1$  de ellas como entrenamiento y la otra  $k$  como validación (validación cruzada) y definir un porcentaje de la muestra como validación (partición porcentual).

<sup>17</sup> Mitchell, T. Machine Learning. McGraw-Hill, New York, 1997.

### **El método K Nearest Neighbours (KNN)**

Es un método de clasificación no paramétrico que estima el valor de la función de densidad  $F(x/C_j)$  o directamente la probabilidad a posteriori de que un elemento  $x$  pertenezca a la clase  $C_j$  a partir de la información proporcionada por el conjunto de prototipos.

Suponiendo un espacio de representación bidimensional y una serie de prototipos de una misma clase representados en él, dado un patrón cualquiera  $X$ , si se considera los  $k$  prototipos más próximos a  $X$ , estos estarán localizados en un círculo centrado en  $X$ . El área del círculo que encierra un número fijo de puntos,  $k$ , es menor en regiones densamente pobladas que en regiones donde los puntos están más dispersos. En espacios multidimensionales el círculo se convierte en una hiperesfera.<sup>18</sup>

### **Entrenamiento de la máquina con el paquete WEKA**

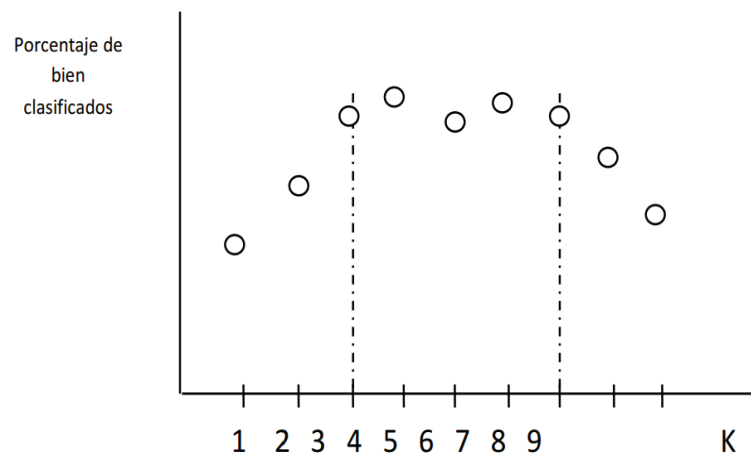
Antes de clasificar a los clientes de la cartera microcrédito sobre los que se ha realizado gestión de cobranza en la empresa ABC, conforme si representan o no un riesgo de deteriorar su comportamiento de pago, hubo que entrenar la máquina. Para esto utilicé el algoritmo IBK, que es una implementación del método KNN en WEKA, que consiste en almacenar en un vector al conjunto de entrenamiento junto con la clase asociada a cada muestra de este conjunto. A continuación, y con motivo del aprendizaje del algoritmo, se calcula la distancia euclídea de cada muestra de entrenamiento, a todas las demás que se tiene almacenadas en el vector mencionado anteriormente y de las que se conoce la clase a la que corresponden, quedándose con las  $K$  muestras más cercanas y clasificando la nueva muestra de entrenamiento en la clase más frecuente a la que pertenecen los  $K$  vecinos obtenidos anteriormente. A continuación se realiza

---

<sup>18</sup> Técnicas supervisadas II: Aproximación no paramétrica. Nikos Drakos, Computer Based Learning Unit, University of Leeds. [http://iie.fing.edu.uy/ense/assign/recpat/material/tema3\\_00-01/node33.html](http://iie.fing.edu.uy/ense/assign/recpat/material/tema3_00-01/node33.html)

el mismo proceso con los datos de validación y al final se calcula el porcentaje de clasificación sobre los casos de este conjunto (desconocidos en la tarea de aprendizaje) para conocer su poder de generalización.<sup>19</sup>

Es importante indicar que la determinación del valor  $k$ , constata empíricamente<sup>20</sup> que el porcentaje de casos bien clasificados es no monótono<sup>21</sup> con respecto de  $k$ , (Figura 2.3) siendo una buena elección valores de  $k$  comprendidos entre 3 y 7.



**Figura 2.3** - Ejemplo de la no monotocidad del porcentaje de valores clasificados en función de  $K$ .

Los resultados del entrenamiento se muestran a continuación en la Tabla 2.20, seleccionando la opción IBK con  $k=3$ , es decir, profundidad 3.

<sup>19</sup> Algoritmos de aprendizaje: KNN & KMEANS, Cristina García Cambroner e Irene Gómez Moreno, Universidad Carlos III de Madrid, 2008-2009.

<sup>20</sup> Aplicación de Predictores Conformales a Señales de Fusión, tesis de maestría en ingeniería informática para la industria, Norma Verónica Ramírez Pérez, Universidad Complutense de Madrid, 2010-2011.

<sup>21</sup> En matemáticas, una función entre conjuntos ordenados se dice monótona (o isotona) si conserva el orden dado.

**Tabla 2.20** - Salida del paquete WEKA con la clasificación obtenida sobre la muestra de entrenamiento

```

=== Summary ===
Correctly Classified Instances      2709      92.4258 %
Incorrectly Classified Instances    222      7.5742 %
Kappa statistic                    0.8486
Mean absolute error                 0.1056
Root mean squared error            0.2282
Relative absolute error            21.1393 %
Root relative squared error        45.6661 %
Total Number of Instances          2931

=== Detailed Accuracy By Class ===
      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.898   0.048   0.952   0.898   0.925   0.983   si
      0.952   0.102   0.897   0.952   0.924   0.983   no
Weighted Avg. 0.924  0.074  0.926  0.924  0.924 0.983

=== Confusion Matrix ===
  a  b  <-- classified as
1362 154 |  a = si
  68 1347 |  b = no

```

Se puede observar que el algoritmo IB3 (método KNN en WEKA con K=3) fue capaz de clasificar el 92.4% de clientes correctamente; la proporción que fue clasificada como mala, perteneciendo verdaderamente a la clase de los malos clientes, es del 89.8% y para el caso de los buenos que sí pertenecen a esa categoría es 95.2%. La proporción que fue clasificada como mala, perteneciendo a los buenos clientes, es 4.8% y para el caso de los buenos que en realidad pertenecen a los malos clientes, la tasa es 10.2%.

### Validación del modelo

La segunda tarea para diseñar el clasificador es realizar el mismo proceso con los datos de validación. Se calcula el porcentaje de clasificación sobre los ejemplos

de este conjunto (desconocidos en la tarea de aprendizaje) para conocer su poder de generalización. Realicé la validación de lo aprendido, aplicando el modelo de entrenamiento sobre la muestra de prueba, obteniendo los resultados que se presentan en la Tabla 2.21.

**Tabla 2.21** - Salida del paquete WEKA con la clasificación obtenida sobre la muestra de prueba

```

=== Summary ===
Correctly Classified Instances      437      90.6639 %
Incorrectly Classified Instances     45      9.3361 %
Kappa statistic                     0.8126
Mean absolute error                  0.1253
Root mean squared error              0.2729
Relative absolute error              25.0123 %
Root relative squared error          54.4544 %
Total Number of Instances           482

=== Detailed Accuracy By Class ===
      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.887   0.075   0.915   0.887   0.901   0.951   si
      0.925   0.113   0.9     0.925   0.912   0.951   no
Weighted Avg.  0.907   0.095   0.907   0.907   0.907   0.951

=== Confusion Matrix ===
  a  b  <-- classified as
204 26 | a = si
 19 233 | b = no

```

La proporción de clientes bien clasificados en la muestra de prueba, apenas si se reduce, pasando del 92.42% al 90.66%; la tasa de verdaderos positivos (TP Rate) para los malos clientes y de falsos positivos (FP Rate) para los buenos, no difieren mucho en relación a los obtenidos con la muestra de entrenamiento; sin embargo, la mayor variación se da en los verdaderos positivos para el caso de los buenos clientes y de falsos positivos para los malos, situación relacionada con un incremento de costos, al asignar recursos en cobranza a cartera de buenos clientes que fueron mal clasificados por el modelo.

## 2.5 ANÁLISIS COMPARATIVO DE MODELOS.

Para finalizar esta parte comparé la regresión logística clásica, la regresión logística mejorada con árboles de decisión, la regresión logística basada en distancias y el algoritmo de clasificación K-NN, sobre la base de analizar el resultado de sus clasificaciones sobre la misma muestra. Utilicé como criterio de selección del modelo las probabilidades de mala clasificación: la de buenos riesgos, la de malos riesgos y la global. Fue adecuado buscar un equilibrio entre las tres probabilidades, porque que a diferencia de un score de crédito minimizar la clasificación de malos riesgos no es lo único que importa, ya que minimizar la clasificación de buenos riesgos tiene un impacto en la cobranza por los altos costos que genera.

**Tabla 2.22** - Análisis comparativo de modelos

	Probabilidades estimadas de mala clasificación		
	Buenos riesgos	Malos riesgos	Global
Modelos			
Regresión Logística Clásica	<b>0,075</b>	0,400	0,230
Regresión Logística mejorada con árboles de decisión	0,179	0,135	0,158
Regresión logística basada en distancias	0,095	0,148	0,120
Clasificador con algoritmo K-NN	0,113	<b>0,075</b>	<b>0,093</b>

Se puede observar en la Tabla 2.22 que la metodología con menores probabilidades de mala clasificación, tanto para los malos riesgos como para la probabilidad global, se obtiene con el clasificador K-NN con valores de 0,075 y 0,093 respectivamente; le sigue la regresión logística mejorada con árboles de decisión con probabilidad de 0,135 para los malos riesgos. Centrándome en riesgo global, la siguiente técnica con menor probabilidad es la regresión logística basada en distancias con un valor de 0,12.

### 2.5.1 CURVAS ROC<sup>22</sup>

Para comprobar cómo se distribuyen los diferentes individuos en función de los valores pronosticados con cada uno de los modelos, utilicé curvas ROC para un punto de corte de 0,5. Los valores de clasificación pronosticados para cada individuo de la muestra de prueba se contrastan con la variable de estado que es *calificación*, misma que tiene los valores reales observados.

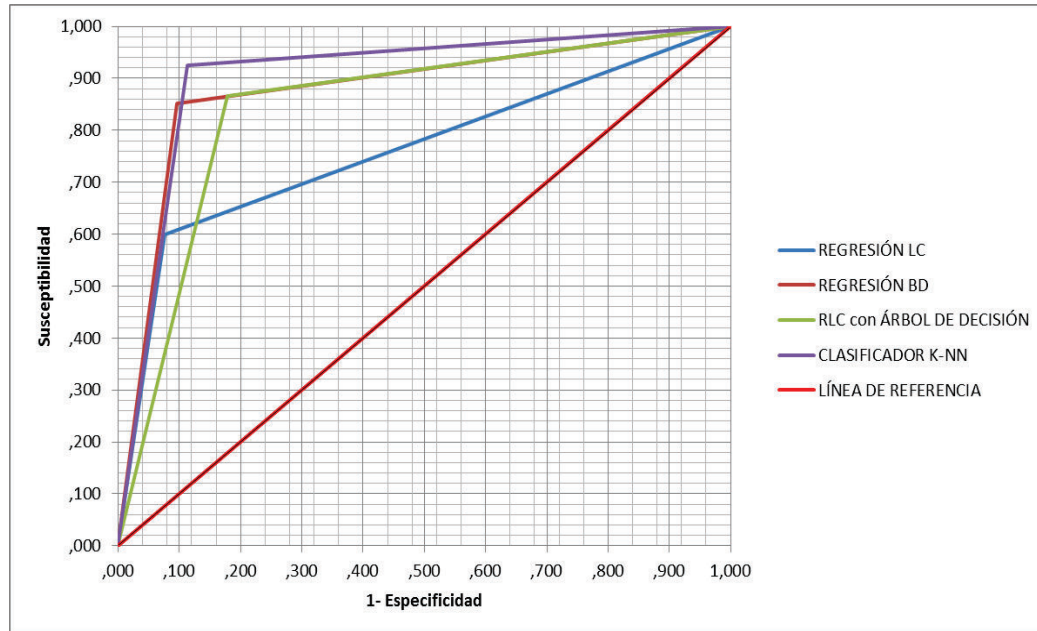
El gráfico ROC apuntando hacia arriba y hacia la izquierda en cada caso, es evidencia de una buena calidad de los modelos y el área bajo la curva indica el riesgo de deterioro del comportamiento de pago que un cliente seleccionado aleatoriamente del grupo de malos tiene respecto a un cliente elegido al azar del grupo de buenos (Figura 2.4). Los resultados son estadísticamente significativos para todos los modelos como se muestra en la Tabla 2.23.

**Tabla 2.23** - Análisis comparativo de los modelos utilizando curvas ROC

Área bajo la curva ROC					
Modelo	Área	Error tip.a	Sig. asintóticab	asintótico al 95%	
				Límite inferior	Límite superior
Regresión Logística Clásica	,762	,023	,000	,718	,807
Regresión Logística con árbol de decisión	,843	,019	,000	,806	,881
Regresión Logística basada en distancias	,878	,017	,000	,845	,912
Clasificador K-NN	,906	,015	,000	,875	,936

El modelo de mayor área bajo la curva ROC es el clasificador K-NN con un valor de 0,906 lo que significa que un cliente seleccionado aleatoriamente del grupo de malos tendrá el 91% de las veces un riesgo mayor de deteriorar su comportamiento de pago que un cliente elegido al azar del grupo de buenos clientes. Le sigue el modelo de regresión logística basado en distancias con un riesgo del 88%.

<sup>22</sup> ROC: Receiver Operating Characteristic o Característica Operativa del Receptor. La curva ROC fue desarrollada inicialmente por ingenieros para la estimación de errores en la transmisión de mensajes y se ha aplicado posteriormente en áreas como la medicina y la estadística.



**Figura 2.4** - Áreas bajo las curvas ROC de cada uno de los modelos

Con estos resultados pude concluir que la mejor metodología es el clasificador K-NN porque tiene la mayor precisión en la predicción. A continuación la regresión logística mejorada con árboles de decisión, porque a pesar de tener similares valores de precisión a la regresión logística basada en distancias, la primera es mejor en cuanto a malos riesgos. Finalmente la regresión logística clásica con la mayor probabilidad de mal clasificados y la peor en cuanto a malos riesgos.

Finalmente como el alcance del trabajo no se centra en la clasificación sino en entender además cómo se comporta la cartera, en conjunto con las gerencias de ABC seleccionamos el método de regresión logística mejorada con árboles de decisión como la mejor opción por la capacidad de explicación de esta metodología, misma que es tan importante o más que la precisión de la predicción. En el siguiente capítulo presento una aplicación del modelo escogido anteriormente, para la construcción del nuevo sistema de cobranza en ABC, y muestro la importancia de tener resultados simples y comprensibles, evitando la dificultad en la interpretabilidad de los modelos.



## CAPÍTULO III

### PROPUESTA DE UN NUEVO SISTEMA DE COBRANZA PARA LA CARTERA MICROCRÉDITO EN ABC

La parte más sensible de este trabajo fue mostrar a las gerencias de ABC el efecto que tienen las acciones de cobranza que se realizan sobre la cartera, explicando de manera sencilla y objetiva la realidad de la cobranza en la empresa, por ejemplo si llamar por teléfono a un cliente es estrictamente necesario, o dicho de mejor manera, ¿cuándo es oportuna la llamada telefónica?, ¿cuántas veces se debe llamar a los clientes?, ¿a dónde llamar, al domicilio o al trabajo?, ¿cuándo dejar de gestionar la cobranza y declarar una cartera como pérdida?; conocer si alguna acción adicional mejora el efecto de la cobranza que se lleva a cabo, es decir, a qué clientes y en qué momento se debe aplicar una combinación de gestiones que maximicen la recuperación?, o incluso, a qué clientes no conviene gestionar por teléfono o en terreno, sino de una forma masiva no personalizada?. Las respuestas a estos cuestionamientos se obtuvieron de los análisis que realicé utilizando regresión logística y los resultados sirvieron como base para definir las estrategias del nuevo sistema de gestión de cobranza para la cartera microcrédito en ABC.

#### 3.1 EVALUACIÓN DE FACTORES QUE AFECTAN LA RECUPERACIÓN DE LA CARTERA.

El objetivo es determinar factores que evaluados simultáneamente, muestren tener algún tipo de relación con la variable dependiente *calificación*, a fin de conocer si se modifica el efecto de estas variables independientes sobre el resultado. En este análisis no existe una variable independiente principal cuyo efecto se desea evaluar, sino varias, mismas que fueron seleccionadas tras

entrevistas con diferentes funcionarios de ABC expertos en el tema, que mostraron gran interés en conocer la naturaleza de estas relaciones.

El análisis con regresión logística es una buena herramienta para controlar estos posibles factores, y que al final me permitirá ratificar o no las hipótesis que se manejan habitualmente en el argot de las cobranzas, tales como las siguientes:

- Las mujeres jóvenes son mejores pagadoras que los hombres.
- Los adultos mayores tienen mejor comportamiento de pago que los jóvenes.
- Los clientes pagan mejor sus primeras cuotas.
- Hacer llamadas telefónicas solo disminuye el riesgo en mora temprana de convertirse en un mal cliente.
- Hacer llamadas y visitas aumenta la recuperación y mejora el comportamiento de pago.
- No tener gestión personalizada de cobro aumenta el riesgo de los clientes que no han pagado al menos la mitad de su crédito.
- No tener compromisos de pago previo aumenta el riesgo de los clientes que se gestionaron telefónicamente.
- Un cliente con atraso es un mal cliente y se debe utilizar todas las herramientas de cobranza disponibles para regularizar su situación de mora.

Cuando se evalúa la relación entre una variable independiente o exposición y una variable dependiente o respuesta, los resultados obtenidos pueden deberse a dos tipos de factores que pueden enmascarar o ejercer algún efecto sobre la relación encontrada, que no corresponden al objeto de estudio y que de no ser considerados pudieran llevar a conclusiones equivocadas. Se trata de variables predictoras que en conjunto se llaman “de control”, pero que juegan un papel diferente y tienen que ver con los fenómenos de confusión e interacción.

### 3.1.1 FACTORES DE CONFUSIÓN

Los llamados factores o variables de confusión, o confundentes, son variables externas a la relación que se evalúa, cronológicamente anteriores a la exposición y cuya presencia produce sesgos en la relación entre la variable dependiente e independientes y que se deben eliminar o controlar a través de ajustes estadísticos. El efecto que puede tener la presencia de esta variable extraña en el resultado puede ser diverso: en ocasiones determina la existencia de asociaciones positivas, las que en la práctica no son reales (error de tipo I o  $\alpha$ ) o bien, pueden esconder asociaciones reales, las que en presencia de esta variable extraña, quedan enmascaradas (error de tipo II o  $\beta$ ).

La confusión se manifiesta a través de una distorsión del efecto observado respecto al efecto real, debido a una desigual distribución de una variable extraña en los grupos estudiados<sup>23</sup>.

#### **Análisis estratificado.**

Una manera de poder superar la heterogeneidad en la distribución de una variable extraña en el estudio de la asociación, es el análisis considerando los diferentes estratos en los que esta variable está categorizada, calculando las medidas de riesgo<sup>24</sup> de acuerdo con los estratos y comparando el resultado con el análisis no estratificado o crudo.

---

<sup>23</sup> El análisis discriminante requiere que la variable por la que se estratifica o ajusta sea de tipo categórico y que el tamaño muestral sea suficientemente grande como para que no existan atributos con pocos representantes.

<sup>24</sup> La relación matemática que se establece entre la distribución del número de casos de una variable dependiente sobre los estratos de otra variable independiente, ambas dicotómicas, permite el cálculo de varias expresiones sin dimensión, tales como odds ratio, riesgo relativo y riesgo atribuible.

Con el análisis estratificado puede intuirse que hay confusión por una tercera variable cuando la razón de odds<sup>25</sup> (OR) de los estratos son parecidas y a su vez diferentes a la OR global o cruda; en cambio si se utiliza regresión logística se detecta cuando la OR que evalúa la fuerza de asociación entre la variable independiente y la variable dependiente, cambia de forma importante cuando se introduce en la ecuación de la regresión una tercera variable.

### 3.1.2 FACTORES DE INTERACCIÓN

Las variables de interacción o modificadoras de efecto, cambian la intensidad o el sentido de la relación entre el factor de análisis y la variable dependiente, descubrirlas es un objetivo del estudio. Con análisis estratificado se detecta que puede haber interacción si la OR en cada estrato son muy diferentes, siendo la OR global o cruda un promedio de ambas.

La interacción requiere introducir en la ecuación de regresión logística un término multiplicativo, compuesto por las dos variables independientes que se presuponen interactúan en su efecto sobre la variable dependiente, y una vez incluido, ver si su coeficiente de regresión logística ( $\beta$ ) es estadísticamente significativo.

### 3.1.3 VALORACIÓN DE LA CONFUSIÓN E INTERACCIÓN

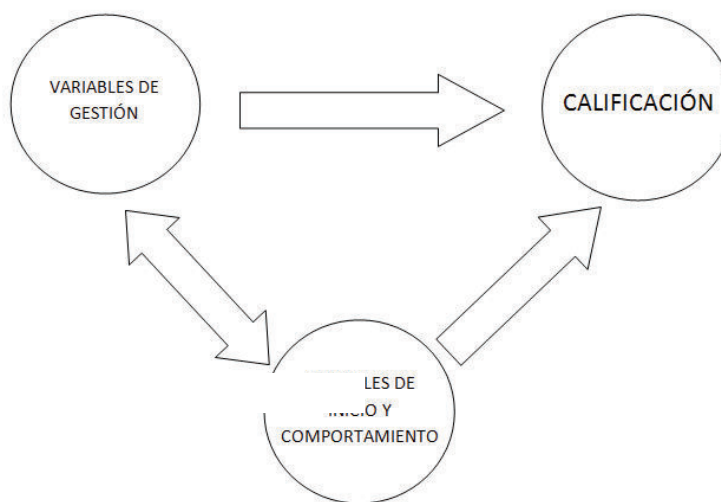
A continuación presento los resultados del análisis que realicé para determinar si existe relación entre las variables independientes, presumiblemente predictoras del efecto tipo de cliente (Figura 3.1) y mostrar la potencialidad que tiene la regresión logística en ofrecer medidas de asociación OR ajustadas o no confundidas. Utilicé las variables cualitativas dicotómicas en todo el análisis, porque si bien en un principio puede parecer que se pierde información, se gana en eficiencia a la hora del análisis y sobre todo claridad en la interpretación. Es

---

<sup>25</sup> La OR (Odds Ratio o razón de odds) expresa cuántas veces más riesgo tienen los expuestos a un resultado o efecto en relación con los no expuestos. En el caso de regresión logística se puede obtener directamente la OR a partir de los coeficientes de regresión:  $OR = e^{\beta}$ .

importante indicar que tanto la confusión e interacción pueden estar presentes, por esta razón evalué por separado la presencia de cada una de ellas en el programa SPSS, seleccionando el método manual de introducción de variables.

Para exponer la metodología empleada en la valoración de los fenómenos de confusión e interacción, a través del estudio de las variables reconocidas como factores de riesgo de tener mal comportamiento de pago en la empresa de cobranza ABC, voy a examinar el caso de las variables de comportamiento de pago y gestiones de cobranza, incluyendo los análisis para el resto de variables en el Anexo VI.



**Figura 3.1** – Representación de los fenómenos de confusión e interacción

### **Valoración de la confusión e interacción de los factores comportamiento de pago y gestiones de cobranza.**

Analicé las posibles relaciones entre las variables de comportamiento y las variables de cobranza, para detectar confusión o interacción entre ellas que modifiquen el desenlace de convertirse en un mal cliente para ABC, y las conclusiones que arroje este análisis junto con los que se adjuntan en el Anexo VI

para el resto de variables, servirán para construir el nuevo sistema de gestión de cobro. Empecé con el estudio de la influencia de la morosidad actual sobre la variable *calificación*, para comprobar si todos los casos que se han ubicado en un rango de mora sobre los 30 días, o mora tardía, se vuelven definitivamente malos (Tabla 3.1).

**Tabla 3.1** - OR entre mora inicial y tipo de cliente

Risk Estimate			
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for califica (BUENO / MALO)	19.881	15.693	25.186
For cohort moralni_rec = entre 0 y 30	2.201	2.073	2.337
For cohort moralni_rec = mayor a 30	.111	.090	.136
N of Valid Cases	2931		

Las tablas de contingencia muestran correlación entre las dos variables, en el sentido que tener una mora mayor a 30 días multiplica por 20 el riesgo de convertirse en mal cliente. Sin embargo de manera natural sale la pregunta: ¿la gestión de cobranza puede modificar este comportamiento? y de ser así ¿en cuánto cambia realizar estas acciones el riesgo de convertirse en un mal cliente?

Para responder estas cuestiones analicé primero las variables de cobranza individuales que representan la cantidad de gestiones, la respuesta que se obtuvo en la interacción con el cliente y el tipo de gestión utilizada, teniendo en cuenta que el servicio de ABC se compone de tres tipos de gestión: masiva no personalizada, gestión telefónica y gestión domiciliaria (ver capítulo uno). Todas estas acciones de cobro tienen como finalidad el recupero de cartera, sin embargo el alcance de este estudio pretende mostrar que existe la necesidad de crear un modelo exitoso en la recuperación, pero que además sirva para influir positivamente en el comportamiento de pago, en el sentido de mantener o mejorar la relación comercial con los clientes y tener los mejores resultados al menor costo posible.

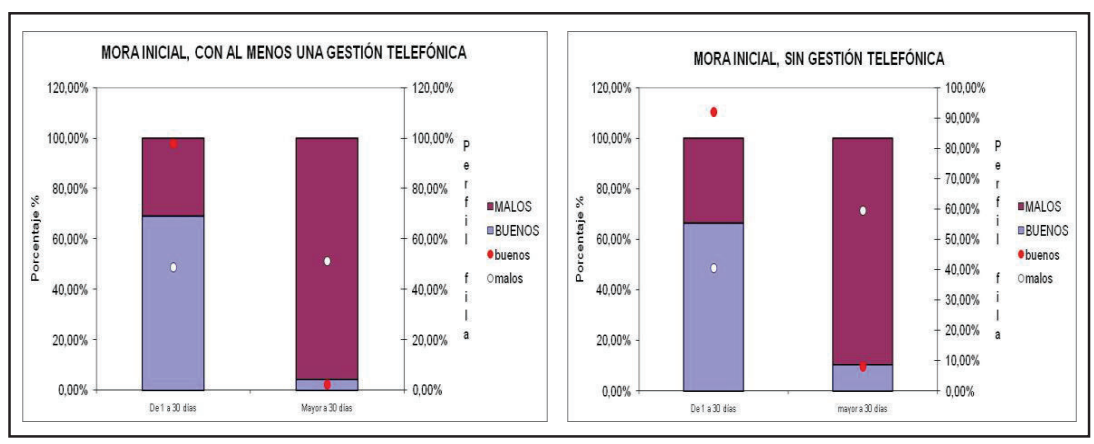
Empecé con el análisis de la gestión telefónica en el último trimestre, que agrupa por una parte a los clientes con al menos 1 gestión telefónica –categoría “0”– y por otro a los clientes sin gestión telefónica –categoría “1”-. Como se muestra en el Anexo II, Tabla A2.2, existe relación entre las gestiones telefónicas del último trimestre y la variable calificación, donde el contraste Chi cuadrado es significativo ( $p=0,003$ ) y la OR 1.3, traducida en que los clientes sin gestión telefónica comparados con aquellos que tienen al menos 1 gestión, incrementan su riesgo de convertirse en mal cliente un 30%. Sin embargo cuando se introducen simultáneamente la morosidad y la gestión telefónica, las medidas de asociación son distintas como se muestra en la Tabla 3.2 y en la Figura 3.2.

**Tabla 3.2** - OR entre mora inicial, gestión telefónica del último trimestre y tipo de cliente

**Variables in the Equation**

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>						
moralni_rec(1)	2.989	.121	610.187	1	.000	19.856
gestionesTel_tri_rec(1)	.014	.098	.020	1	.887	1.014
Constant	-.730	.084	74.626	1	.000	.482

a. Variable(s) entered on step 1: moralni\_rec, gestionesTel\_tri\_rec.



**Figura 3.2** - Relación entre mora inicial, gestión telefónica del último trimestre y tipo de cliente

El resultado obtenido es interesante, porque si bien la OR para la morosidad actual se mantiene sobre su estimación y el coeficiente de regresión es estadísticamente significativo, en el caso de las gestiones telefónicas en el último trimestre se reduce en proporciones importantes (casi un 20%) y el coeficiente de regresión no es estadísticamente significativo. No se puede afirmar que la gestión telefónica confunda el efecto sobre el tipo de cliente, sin embargo para completar el análisis exploré una posible interacción entre ellas y su acción sobre la variable calificación.

Comprobé primero si existe asociación entre las gestiones telefónicas y la morosidad, mediante el análisis con tablas de contingencia mostrado en la Tabla 3.3.

**Tabla 3.3** - Proporción de clientes sin gestión telefónica por edad de mora

		moralni_rec		Total
		entre 0 y 30	mayor a 30	
gestionesTel_tri_rec	al menos 1	Count 580 % within gestionesTel_tri_rec 74.6%	197 25.4%	777 100.0%
	ninguna	Count 1390 % within gestionesTel_tri_rec 64.5%	764 35.5%	2154 100.0%
Total		Count 1970 % within gestionesTel_tri_rec 67.2%	961 32.8%	2931 100.0%

Existe una desproporción entre los clientes sin gestión telefónica con menos de 30 días de mora (64.5%) y los que tienen más de 30 días de morosidad sin gestión telefónica (35.5%), es decir, la mayoría de los clientes sin gestión tienen morosidad menor a 30 días y la mayoría de los clientes con al menos una gestión también tienen morosidad menor a 30 días, siendo estas diferencias estadísticamente significativas según el contraste Chi cuadrado como muestro en la Tabla 3.4 y la Tabla 3.5.



**Tabla 3.4** - Relación entre mora inicial y gestión telefónica último trimestre

Chi-Square Tests					
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	26.511 <sup>a</sup>	1	.000		
Continuity Correction <sup>b</sup>	26.054	1	.000		
Likelihood Ratio	27.315	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	26.502	1	.000		
N of Valid Cases	2931				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 254,76.  
b. Computed only for a 2x2 table

**Tabla 3.5** - ODD entre mora inicial y gestión telefónica último trimestre

Risk Estimate			
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for <i>gestionesTel_tri_rec</i> (al menos 1 / ninguna)	1.618	1.346	1.945
For cohort <i>moralni_rec</i> = entre 0 y 30	1.157	1.099	1.218
For cohort <i>moralni_rec</i> = mayor a 30	.715	.626	.817
N of Valid Cases	2931		

A la vista de estos resultados no pude descartar que ambas covariables generen modificación de efecto o interacción sobre la variable calificación; continuando con el estudio, la salida de la regresión con el término de interacción multiplicativo *gestionesTel\_tri\_rec\*moralni\_rec* (Tabla 3.6), muestra que éste es estadísticamente significativo e introduce nueva información a la OR de la variable morosidad, multiplicando por 2 este valor hasta llegar a 47, pero a costa de que dicha estimación se hace más imprecisa traducida en un mayor error estándar de los coeficientes de regresión y un más amplio intervalo de confianza para su OR.

**Tabla 3.6** - Análisis de la interacción entre mora inicial y gestión telefónica último trimestre

Variables in the Equation								
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup> gestionesTel_tri_rec(1)	.122	.106	1.317	1	.251	1.130	.917	1.392
moralni_rec(1)	3.846	.353	118.786	1	.000	46.796	23.434	93.446
gestionesTel_tri_rec(1) by moralni_rec(1)	-1.029	.376	7.484	1	.006	.357	.171	.747
Constant	-.807	.090	80.512	1	.000	.446		

a. Variable(s) entered on step 1: gestionesTel\_tri\_rec, moralni\_rec, gestionesTel\_tri\_rec \* moralni\_rec.

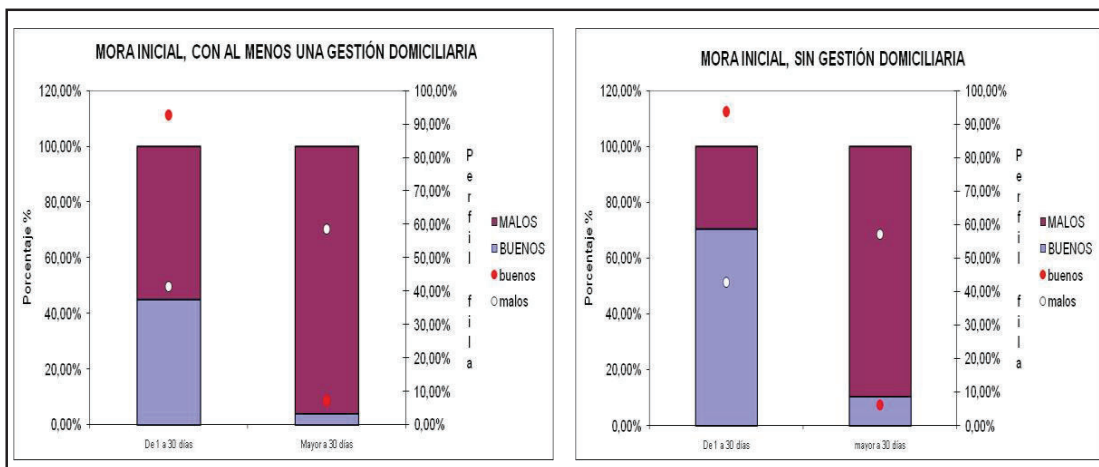
Respecto a la gestión telefónica, su OR no cambia significativamente y se hace no significativa. Se puede concluir entonces que el valor de 47 es el riesgo de ser mal cliente cuando no se hacen gestiones telefónicas y la mora inicial es mayor a 30 días.

Realizando el mismo análisis con la variable de las gestiones domiciliarias en el último trimestre, el efecto es diferente como se muestra la Tabla 3.7 y la Figura 3.3. A más de no ser estadísticamente significativo, el término de interacción tampoco introduce nueva información en la estimación de la OR para las variables de gestión domiciliaria y morosidad, mismas que se mantienen, o incluso para el caso de la morosidad, hace que las estimaciones sean más imprecisas.

**Tabla 3.7** - Análisis de la interacción entre mora inicial y gestión domiciliaria último trimestre

Variables in the Equation								
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup> moralni_rec	2.894	.363	63.463	1	.000	18.071	8.866	36.834
gestionesDom_tri_rec(1)	-1.071	.137	61.465	1	.000	.343	.262	.448
gestionesDom_tri_rec(1) by moralni_rec	.097	.386	.063	1	.802	1.102	.517	2.345
Constant	.197	.126	2.443	1	.118	1.217		

a. Variable(s) entered on step 1: moralni\_rec, gestionesDom\_tri\_rec, gestionesDom\_tri\_rec \* moralni\_rec.



**Figura 3.3** - Relación entre mora inicial y gestión domiciliaria último trimestre

Siguiendo la misma metodología realicé el estudio con el resto de variables, para descubrir si existen cambios importantes en las medidas de correlación y de riesgo cuando se introducen de forma simultánea otras variables que incrementan o ajustan dichos valores (Anexo VI). Finalmente cabe indicar que para realizar con éxito este análisis fue muy importante el conocimiento en materia de cobranzas, mismo que lo obtuve de las entrevistas con los diferentes responsables de la operación en ABC y que me ayudaron bastante en este cometido.

## Conclusiones

1. Tener gestión telefónica el mes anterior no guarda relación con ser mal cliente para ABC: que un cliente no tenga llamadas el mes anterior no aumenta su riesgo de convertirse en mal cliente.
2. Tener al menos una gestión domiciliaria el mes anterior se relaciona con el tipo de cliente para ABC: los clientes que no tuvieron gestión domiciliaria el mes anterior redujeron su riesgo de ser mal cliente. Esta conclusión hay que saberla interpretar puesto que en la empresa de cobranza se realiza gestión domiciliaria a los clientes que ya empezaron un proceso de deterioro en su

comportamiento de pago, o también a aquellos que se considera importantes para mantener la relación comercial con los Cedentes.

3. Tener al menos una gestión telefónica en los últimos tres meses se relaciona con la variable *calificación*: entre los malos clientes hay 47 veces más probabilidad de no tener gestión telefónica en el último trimestre que de haber tenido al menos una, es decir que los clientes de mora temprana que no tuvieron gestión telefónica en el último trimestre multiplicaron su riesgo por 47 de convertirse en malos en relación a los que sí la tuvieron.
4. Tener al menos una gestión domiciliaria en los últimos tres meses se relaciona con el tipo de cliente para ABC: los clientes que no tuvieron gestión domiciliaria en el último trimestre disminuyeron su riesgo de deteriorar su comportamiento de pago, en el mismo sentido explicado anteriormente para el caso de la variable gestión domiciliaria en el mes anterior. A la luz de estos resultados se debe reorientar la gestión domiciliaria, que es muy costosa, a la cartera que ya muestra problemas serios de recuperación.
5. Tener un contacto con el titular durante la gestión telefónica en el mes anterior y en el último trimestre, no influye en el tipo de cliente; tampoco el hecho de haber obtenido un compromiso de pago.
6. Tener direcciones y teléfonos se relaciona con la variable *calificación*: entre los malos clientes hay 5 veces más probabilidad que no tengan direcciones y teléfonos, a que tengan al menos una dirección o teléfono válido, es decir se multiplica por 5 el riesgo de convertirse en mal cliente para ABC en relación a los que sí tienen teléfonos y direcciones; entre los malos clientes existe 12 veces más probabilidad que no tengan un número de teléfono laboral que de tenerlo, es decir que se multiplica por 12 el riesgo de convertirse en malo en relación al que sí tiene teléfono laboral; entre los malos clientes hay 6 veces más probabilidad de no tener un número de teléfono particular que de tener al menos uno, es decir que no tener números de teléfono particular multiplica el

riesgo por 6; entre los malos clientes existe 6 veces más probabilidad de no tener al menos una dirección laboral que de tenerla, es decir, no tener al menos una dirección laboral multiplica por 6 el riesgo de deteriorar su comportamiento de pago en relación al que sí tiene dirección laboral. De todo esto se concluye lo importante que debe ser para ABC tener una base de teléfonos y direcciones correcta y actualizada.

7. El riesgo de pago de la cartera microcrédito en cobranza se incrementa rápidamente, incluso desde el primer día de atraso. Después del día 15 de morosidad el seguimiento debe ser permanente, de forma que el cliente perciba que recibe un servicio integral y de calidad.
8. Si el cliente paga con atraso una sola cuota puede ser síntoma que a futuro incrementa su riesgo, a pesar que se encuentre al día. En este caso puede ser importante una gestión antes de la fecha de vencimiento, utilizando una comunicación adecuada para que el cliente no reaccione de forma negativa y por el contrario lo perciba como un beneficio, formando un lazo emocional con el Cedente de la cartera.
9. El seguimiento debe continuar durante toda la madurez del microcrédito, porque el cliente puede llegar a “cansarse” de pagar cuota tras cuota, y al final sin motivo aparente dejar este comportamiento oportuno, porque percibe que dejó de recibir el acompañamiento que tuvo al inicio. Aquí la gestión debe orientarse no solo a la recuperación de las últimas cuotas, sino también a formar un promotor del producto, contribuyendo de esta manera el crecimiento del negocio del Cedente.
10. Que exista muchos clientes que no han pagado las primeras cuotas podría indicar falencias en el proceso de crédito. La gestión de cobranza debe retroalimentar al Cedente en esta situación, para que revise sus políticas y haga un control minucioso caso por caso, de manera que se corrijan errores y que su negocio crezca sanamente.

## 3.2 MODELOS ESTADÍSTICOS BASADOS EN SCORE

El score es una herramienta de decisión que ordena jerárquicamente una población por la opción que ocurra un cierto evento. Para esto utilicé una tabla denominada scorecard, construida en base a los resultados del modelo logístico mejorado con árboles de decisión y ajustando el valor respectivo del score por medio de la fórmula:

$$\text{Score} = p * 1000$$

donde  $p$  es la probabilidad de convertirse en mal cliente para ABC, de manera que aquel con menor riesgo de ser malo tenga el valor de 1 y el cliente con más riesgo tenga el valor de 999.

La construcción de un modelo basado en score obedece a dos posibles objetivos:

### 1. Predicción de los malos clientes en cobranza

Además del poder de discriminación del modelo estadístico, se debe estudiar su estabilidad en el transcurso del tiempo, así como el proceso continuo y automático de recolección de la información que alimenta al modelo y permite monitorearlo. En el capítulo cuatro voy a mostrar la metodología para realizar este seguimiento.

### 2. Definición de las estrategias de cobranza

Para validar la aplicabilidad del modelo estadístico construido, se debe generar reglas que permitan proponer acciones para minimizar la probabilidad de convertirse en mal cliente. El objetivo es identificar las cualidades y beneficios que se obtendrá implementando las estrategias basadas en el modelo de score construido para la empresa ABC.

Como paso previo, seleccioné los casos que son aptos para la aplicación o uso del modelo de score obtenido en el capítulo dos, sabiendo que pueden existir clientes de cartera microcrédito que a su vez tienen otros productos vigentes con el mismo Cedente, tales como créditos de consumo, hipotecarios o comerciales. Estos casos se excluyeron porque no formaron parte del desarrollo.

### **3.2.1 CONSTRUCCIÓN DE LA TABLA DE DESEMPEÑO**

De acuerdo a la teoría de gestión basada en CRM, el control de la morosidad debe estar enfocado al mantenimiento del cliente, particularmente a una adecuada gestión de cobro que lo motive a estar al día en los pagos e influya en su comportamiento para prevenir el deterioro progresivo de la cartera. Con este fin construí la Tabla 3.8 calculando el score sobre la muestra de prueba<sup>26</sup> y distribuyendo los clientes por el rango de sus días de atraso, de manera que me permitan realizar una interpretación adecuada del desempeño de pago asociado a cada uno de los ellos.

#### **Rangos del score**

Los rangos se establecieron en función de la acumulación del 10% de casos en la muestra de prueba, ordenando previamente los valores de score desde los puntajes más bajos a los más altos. Cuando el score disminuye, la proporción de clientes malos es menor, por lo que valores de score más altos indican un mayor riesgo.

#### **Clasificación de comportamiento**

A continuación analicé el comportamiento de pago utilizando la variable mora máxima alcanzada entre los 5 y 11 meses que duró la ventana de muestreo,

---

<sup>26</sup> Se generan las estrategias sobre la muestra de prueba, ya que se considera a esta como el escenario más adecuado al no formar parte de la muestra de construcción del modelo, evaluando así su capacidad de ser aplicada en clientes de otras muestras proCedentes de la misma población.

clasificando en una de las cinco posibles categorías a cada cliente de la muestra de prueba sobre el que se realiza gestión de cobranza en ABC:

1. Buenos/Mora máxima de 1 a 15 días
2. Mora máxima de 15 a 30 días.
3. Mora máxima de 30 a 60 días.
4. Mora máxima de 60 a 90 días.
5. Mora máxima mayor a 90 días.

### **Porcentaje Acumulado**

La cantidad de clientes que se acumulan en cada rango del score, expresada en términos porcentuales, establece un punto de corte que define un segmento de cartera sobre el que se aplica una estrategia de cobranza particular y diferente, en base al servicio acordado previamente con el dueño de la cartera; por ejemplo, una estrategia con corte en el valor 900 implica que el 40,25% del total de los clientes serán retirados de gestión o tramitados a través del área legal de ABC. Si se considera retirar de gestión aquellos clientes que tengan 980 puntos en adelante, se está identificando el 55% de los clientes malos con mora de 90 días en adelante, donde además solo el 6,9% de los clientes son buenos.

### **Distribución del intervalo**

Representa el porcentaje de créditos en una categoría de comportamiento con un score determinado, comparados con otras categorías en la tabla de desempeño. Por ejemplo 48 cuentas caen en el rango entre 494 y 500; de estos el 16.67% registró una mora de 30 días o más y solo el 2% registraron morosidad de 60 días en adelante.



**Odds a 30 días**

Representa la relación entre la cantidad de clientes buenos o con atraso hasta 15 días, por cada cliente que ya ha empezado su proceso de deterioro fijado en el umbral de 30 días de atraso. Por ejemplo en el segmento con bajo riesgo, score de 1 a 430 puntos, existen 45 clientes buenos por cada cliente malo y en el rango de 494 a 500 puntos existe 4.63 buenos por cada malo; esto quiere decir que hasta un punto de corte de 500 puntos existen 14.33 clientes buenos por cada cliente malo. Esta medida ayuda a definir los puntos de corte para cada estrategia, aumentando o disminuyendo la intensidad conforme se incrementa el riesgo.

**Tabla 3.8** - Tabla de desempeño para el modelo de score de ABC

Rango del Score	Total		Buenos/Mora 1 a 15 días		Mora de 15 a 30 días		Mora de 30 a 60 días		Mora de 60 a 90 días		Mora de 90 a 120 o peor		Otis a 30 días a 30 días acumulado	Otis a 30 días a 30 días acumulado	Homologación Resultado Decisión		
	Distribución intervalo	Porcentaje acumulado	Distribución intervalo	Porcentaje acumulado	Distribución intervalo	Porcentaje acumulado	Distribución intervalo	Porcentaje acumulado	Distribución intervalo	Porcentaje acumulado	Distribución intervalo	Porcentaje acumulado					
1	48	100,0%	45	93,75%	100,0%	1	2,08%	100,0%	1	2,08%	100,0%	0	0,00%	100,0%	45,00	RIESGO BAJO DE DETERIORO	
430	48	90,04%	45	93,75%	83,64%	1	2,08%	99,47%	0	0,00%	98,88%	0	0,00%	97,6%	45,00	RIESGO BAJO DE DETERIORO	
490	48	80,08%	44	91,67%	67,27%	2	4,17%	98,93%	0	0,00%	98,88%	0	0,00%	97,6%	33,50	ZONA GRIS	
492	48	70,12%	44	91,67%	51,27%	3	6,25%	97,86%	0	0,00%	98,88%	0	0,00%	97,6%	25,43	ZONA GRIS	
494	48	60,17%	37	77,08%	35,27%	8	16,67%	96,26%	1	2,08%	98,88%	1	2,08%	97,6%	14,33	ZONA GRIS	
500	48	50,21%	28	58,33%	21,82%	17	35,42%	91,98%	4	8,33%	97,75%	4	8,33%	95,2%	7,59	RIESGO MEDIO DE DETERIORO	
900	48	40,25%	13	27,08%	11,64%	34	70,83%	82,89%	10	20,83%	93,26%	9	18,75%	85,71%	3,88	RIESGO ALTO DE DETERIORO	
997	48	30,29%	9	18,75%	6,91%	37	77,08%	64,71%	20	41,67%	82,02%	6	12,50%	55,17%	2,57	RIESGO ALTO DE DETERIORO	
998	48	20,33%	0	0,00%	3,64%	45	93,75%	44,92%	26	54,17%	59,55%	7	14,58%	34,48%	1,79	RIESGO ALTO DE DETERIORO	
998	50	10,37%	10	20,00%	3,64%	39	78,00%	20,86%	27	54,00%	30,34%	3	6,00%	10,34%	1,47	RIESGO ALTO DE DETERIORO	
Totales	482		275	57,05%		187	38,80%		89	18,46%		29	6,02%		1,47	38,80%	1,47

Rango del score

Porcentaje acumulado

Distribución del intervalo

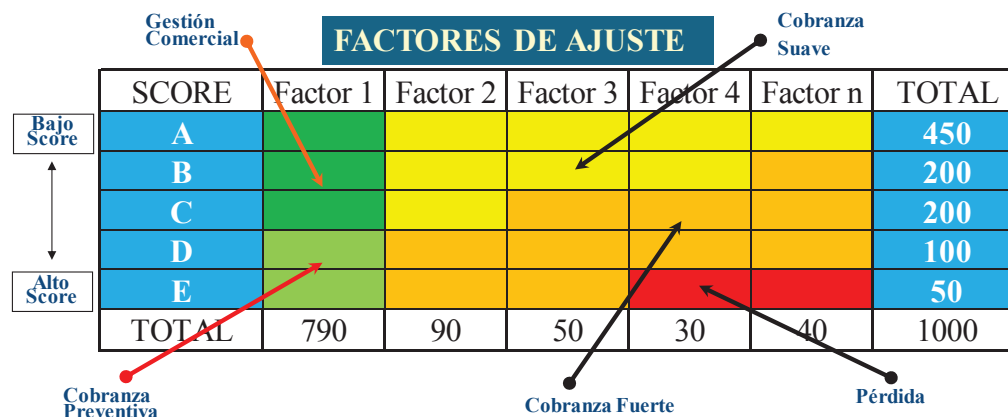
Clasificación de comportamiento



### 3.2.2 CREACIÓN DE ESTRATEGIAS Y ACCIONES DE COBRANZA

La creación de estrategias y acciones de cobranza nació de una reunión donde participaron representantes del área de call center, área domiciliaria, área comercial y tecnología de la información.

Una vez segmentada la población de acuerdo al pronóstico de ser mal cliente para ABC, el siguiente paso fue el diseño de estrategias de cobro diferenciadas por cliente, basado en una gestión particular sobre los menos riesgosos y aquellos con mayor probabilidad de deterioro, manteniendo un equilibrio entre la eficiencia del servicio y la buena relación comercial con el Cedente referida al cumplimiento de las metas de recuperación (Figura 3.4).



**Figura 3.4** - Esquema gráfico del diseño de estrategias basadas en modelos de score

La necesidad de contar con una adecuada planificación del trabajo basado en estrategias, encuentra su origen en dos aspectos principales que expongo a continuación:

**Conocimiento de la cartera.**

Orientar los esfuerzos hacia el total de los clientes, especialmente cuando éstos ya poseen gestiones históricas, es poco óptimo; lo mejor es actuar conforme sea el resultado de esa historia, lo que incluye el conocimiento de las características propias de la cartera y las exigencias planteadas por el Cedente a ABC. Entre éstas figuran:

- El tipo de metas planteadas
- La información que se necesita obtener.
- Las políticas de negociación de los saldos deudores.
- Los protocolos de servicio.
- Los productos comerciales y especiales.
- Las políticas de reprogramación de deuda.

**Viabilidad en términos económicos y de productividad.**

Los recursos disponibles para la gestión siempre serán escasos, lo que exige un mejor uso de aquellos que están disponibles. El aumento de gestores u operadores de cobranza no significa un aumento de productividad, pero sí implica un incremento del gasto. Para visualizar esto basta pensar que en el límite, con un operador para cada crédito u operación en mora, no se asegura la recuperación o normalización del mismo.

**3.2.3 CRITERIOS UTILIZADOS PARA SEGMENTAR LA CARTERA EN BASE A ESTRATEGIAS.**

El modelo de score es la base para construir el nuevo sistema de gestión de cobranza en ABC, sin embargo su correcta implementación requiere una combinación de técnica y experiencia para adaptar los procesos de gestión y las estrategias a la realidad del momento. En este sentido es necesario que de forma

complementaria se consideren otros criterios recogidos de las buenas prácticas, estableciendo criterios de premio o castigo que ajusten el resultado del score para dar o no mayor profundidad y atención a la cobranza, mismos que muestro a continuación por orden de prioridad.

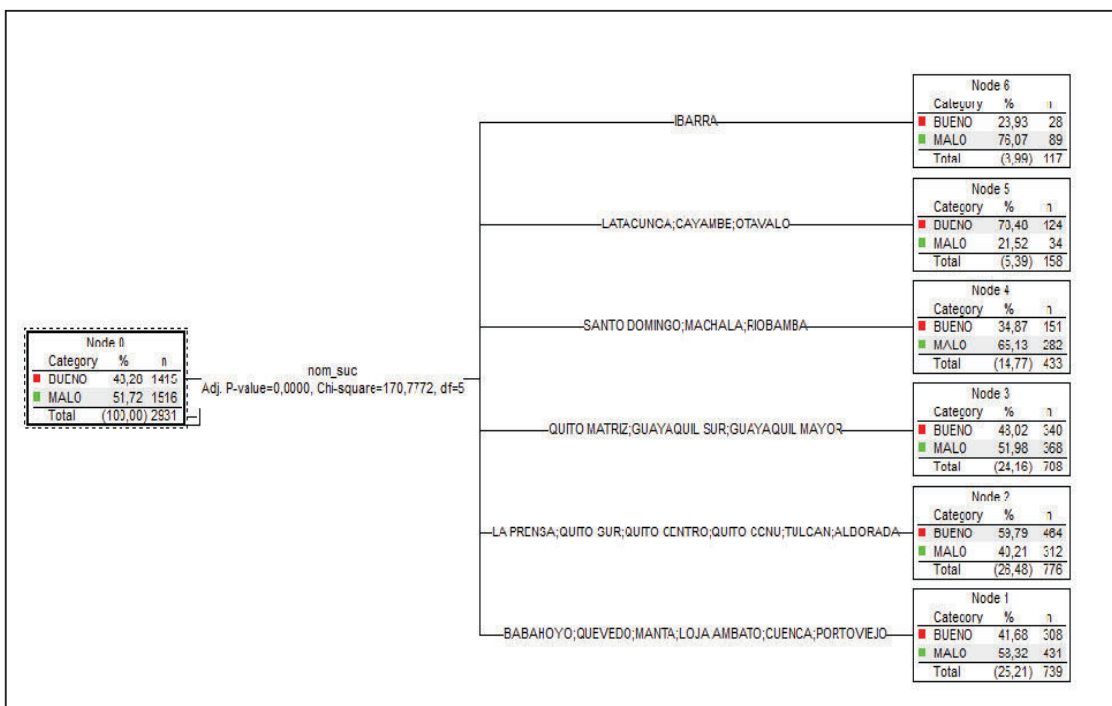
### **Valor del score**

Mientras más bajo sea el score la gestión es más comercial e informativa, para afianzar la relación comercial con el cliente; cuando el valor del score sube también debe incrementarse la intensidad de la gestión, hasta llegar a un punto en el cual se decida dejar esta cartera a cargo del área legal, para que realice una gestión extrajudicial y judicial.

### **Sucursal**

Del análisis bivariante que mostré en el capítulo dos, concluí que las agencias de Cayambe y Otavalo tienen porcentajes distintos de clientes buenos y malos, a pesar de ser cercanas geográficamente; caso similar se da entre las agencias Quito Matriz y Quito Norte. Por el contrario las sucursales de Manta y Portoviejo presentan similar porcentaje de buenos clientes y están ubicadas relativamente cerca una de la otra, aunque son poco pobladas. Entre las sucursales más grandes en cuanto a número de clientes, si comparamos Guayaquil Sur con Guayaquil Norte, existe una diferencia de 15 puntos en el porcentaje de clientes malos, lo que puede indicar un manejo inadecuado de cartera en la sucursal Guayaquil Sur. En las sucursales pequeñas existe un comportamiento similar entre Cayambe, La Prensa y Tulcán, donde predominan los clientes buenos, caso contrario sucede en Babahoyo, Ibarra, Machala, Manta y Portoviejo donde predominan los clientes malos.

Utilizando árboles de decisión para segmentar el tipo de cliente con la variable *sucursal*, obtuve 6 grupos de la manera como se muestra en el Gráfico 3.5:



**Figura 3.5** - Segmentación de la variable sucursal por tipo de cliente utilizando árboles de decisión.

En conjunto con ABC decidimos considerar la variable sucursal en la definición de estrategias, creando dos grupos como se muestra en la Tabla 3.9.

**Tabla 3.9** - Estrategias por sucursales con alto o bajo riesgo

Sucursal	Valoración	Ajuste del score
IBARRA, SANTO DOMINGO, MACHALA, RIOBAMBA, BABAHOYO, QUEVEDO, MANTA, LOJA, AMBATO, CUENCA, PORTOVIEJO, QUITO CENTRO, QUITO MAYOR, GUAYAQUIL SUR, GUAYAQUIL CENTRO	Sucursal de alto riesgo	Castigo
LATACUNGA, CAYAMBE, OTAVALO, LA PRENSA, QUITO SUR, QUITO NORTE, TULCÁN, GUAYAQUIL NORTE.	Sucursal de riesgo bajo	Premio

### **Cantidad de cuotas pagadas**

Si el cliente tiene cero, una o dos cuotas pagadas, se castiga con una gestión de cobranza más intensa (Tabla 3.10).

**Tabla 3.10** - Estrategias por cantidad de cuotas pagadas con alto o bajo riesgo

<b>Cantidad de cuotas pagadas</b>	<b>Valoración</b>	<b>Ajuste del score</b>
Cero, una o dos cuotas pagadas	Segmento de alto riesgo	Castigo
Al menos dos cuotas pagadas	Segmento de riesgo bajo	Premio

### **3.2.4 CRITERIOS UTILIZADOS PARA DETERMINAR LAS ACCIONES DE COBRANZA PARA CADA ESTRATEGIA.**

Para decidir qué estrategia se aplica a cada tipo de cliente, utilicé criterios relacionados con el servicio ofertado, los acuerdos comerciales y el impacto económico que se genera en la operación. Esta tarea la desarrollamos en conjunto con el área de operaciones de ABC, sin embargo volví a enfatizar que entre todas las áreas involucradas se debía levantar un procedimiento para la correcta ejecución de las estrategias y su monitoreo, tema que abordaré en el siguiente capítulo, completando de esta manera la base del nuevo sistema de gestión de cobranza para ABC (Tabla 3.11).

#### **De la urgencia en el inicio de la gestión**

La gestión telefónica normalmente es más rápida que una del tipo domiciliario. Si este factor es decisivo en el servicio, lo recomendable es iniciar primero la gestión telefónica antes que la domiciliaria.

**De los costos e incidencia económica del tipo de gestión.**

Una llamada telefónica normalmente es más barata que una visita, entre otras cosas porque un ejecutivo telefónico es capaz de gestionar una mayor cantidad de clientes en un menor tiempo y con mejores niveles de servicio, debido principalmente al uso de tecnología especializada que aumenta la productividad de los ejecutivos telefónicos, con el consecuente ahorro en costos.

**De la profundidad de la gestión**

Es posible que a través de las visitas en terreno se consiga una mayor intensidad de gestión, en la medida que se establezca un contacto directo con el titular del crédito. Si ese es el objetivo a perseguir, lo recomendable será optar por este tipo de asignación.

**De los aspectos comerciales coordinados con el Cedente.**

Si en los acuerdos comerciales con el Cedente se acordó un determinado tipo de gestión, la asignación debe obedecer a ese criterio y no a otro. Sólo se excluirán los clientes para los cuales su información o historial de gestiones no permitan su ejecución.

**De la antigüedad de la cartera.**

En una cartera muy antigua es mejor iniciar la gestión con llamadas telefónicas, para confirmar direcciones del domicilio o trabajo y solo entonces proceder con la gestión de terreno; cuando se trata de una cartera muy nueva es posible asumir que los datos son confiables y bastará sólo una pequeña cantidad de llamadas o gestión con procesos masivos para conseguir la normalización de la morosidad. En otras instancias donde se requiera mayor profundidad de la gestión, la combinación de llamadas y visitas es lo más apropiado.



**Tabla 3.11** - Cuadro de asignación de acciones de cobranza por estrategia propuesto para la empresa de cobranza ABC

Score	Sucursal	Cuotas Pagadas	Tramo de mora	Número de estrategia	Descripción	Acción relacionada	Observaciones
Alto	Indiferente	Indiferente	1 a 15 días	1	Clientes de Score Alto, alta probabilidad de deterioro, Sucursal Indiferente, Cuotas Pagadas Indiferente, entre 1 y 15 días de mora	Gestión Telefónica personal, llamada de cobro con guión 2. Visita del ejecutivo de crédito. Envío de mensajes de correo electrónico y mensajes de texto.	Llamada de cobro en tono suave. La gestión en terreno no debe ser de cobro, debe ser de seguimiento. Envío de mensajes de correo electrónico.
Bajo-Zona Gris	Castigo	Indiferente	1 a 15 días	2	Clientes con Score Bajo y Zona Gris, Castigo por Sucursal con alta probabilidad de deterioro, Cuotas Pagadas Indiferente, entre 1 y 15 días de mora	Gestión Telefónica personal, llamada de cobro con guión 1. Visita del ejecutivo de crédito. Envío de mensajes de correo electrónico.	Llamada de cobro en tono suave. La gestión en terreno no debe ser de cobro, debe ser de seguimiento.
Bajo-Zona Gris	Premio	Indiferente	1 a 15 días	3	Clientes con Score Bajo y Zona Gris, Premio por Sucursal con baja probabilidad de deterioro, Cuotas Pagadas Indiferente, entre 1 y 15 días de mora	Gestión telefónica personal. Llamada de tipo Comercial, no de cobro.	Es una gestión de seguimiento no de cobro
Alto	Indiferente	Indiferente	16 a 30 días	4	Clientes con Score Alto, alta probabilidad de fraude, Sucursal Indiferente, Cuotas Pagadas Indiferente, entre 16 y 30 días de mora	Gestión telefónica personal con guión 3. Visita por parte del ejecutivo de cobranza. Envío de mensajes de texto y correos electrónicos. Entrega de notificación de cobro 1	Llamada de cobro en tono medio. Entregar una notificación de cobro en la visita domiciliaria.
Bajo-Zona Gris	Indiferente	Castigo	16 a 30 días	5	Clientes con Score Bajo y Zona Gris, Sucursal Indiferente, Castigo por Cuotas Pagadas con alta probabilidad de deterioro, entre 16 y 30 días de mora	Gestión telefónica personal con guión 3. Visita por parte del ejecutivo de cobranza. Envío de mensajes de texto y correos electrónicos. Entrega de notificación de cobro 1.	Llamada de cobro en tono medio. Entregar una notificación de cobro en la visita domiciliaria.
Bajo-Zona Gris	Indiferente	Premio	16 a 30 días	6	Clientes con Score Bajo y Zona Gris, Sucursal Indiferente, Premio por Cuotas Pagadas con baja probabilidad de deterioro, entre 16 y 30 días de mora	Gestión telefónica personal con guión 3. Visita por parte del ejecutivo de cobranza para los clientes que no sean localizados.	Llamada de cobro en tono medio. Entregar una notificación de cobro en la visita domiciliaria.
Alto	Indiferente	Indiferente	30 a 60	7	Clientes de Score Alto, alta probabilidad de deterioro, Sucursal Indiferente, Cuotas Pagadas Indiferente, entre 30 y 60 días de mora	Gestión telefónica con guión 4 y visita domiciliaria acompañada de un coordinador, uso de guión 1. Investigación de datos para los clientes localizados. Envío de comunicación escrita 2 con un ejecutivo de cobranza, directamente al titular, anunciando reporte a buró de crédito. Envío de correos electrónicos.	Llamada en tono fuerte.
Bajo-Zona Gris	Indiferente	Indiferente	30 a 60	8	Clientes de Score Bajo y Zona Gris, Sucursal Indiferente, Cuotas Pagadas Indiferente, entre 30 y 60 días de mora	Gestión telefónica con guión 4 y envío de notificación escrita por correo al titular anunciando reporte a buró de crédito. Envío de correos electrónicos.	Llamada de cobro en tono fuerte. Visitar los clientes de acuerdo a los mayores montos y los menores scores
Indiferente	Indiferente	Indiferente	61 a 90 días	9	Clientes entre 61 y 90 días de mora de cualquier score, Sucursal Indiferente, Cuotas Pagadas Indiferente	Gestión telefónica con guión 5 y envío de notificación escrita por correo al titular. Reporte a buró de crédito. Visita al domicilio y trabajo, amenaza de embargo	Llamada en tono fuerte. Carta anunciando procesos judiciales. Inventariar activos para preocupar al cliente por el embargo.
Indiferente	Indiferente	Indiferente	91 a 120 días	10	Clientes entre 91 y 120 días de mora. Cualquier Score, Sucursal Indiferente, Cuotas Pagadas Indiferente	Visita al domicilio y al trabajo, uso de guión 2. Posible entrega de cartera a terceros. Paso al área legal con remuneración por valor recuperado.	Cobro Jurídico. Cartera de recuperación mínima.
Indiferente	Indiferente	Indiferente	Mayor a 120 días	11	Clientes mayor a 120 días de mora. Cualquier Score, Sucursal Indiferente, Cuotas Pagadas Indiferente	Evaluar el castigo, gestión terminal o venta de cartera	Negociar el mejor precio con el comprador de la cartera

### **3.2.5 COMPARACIÓN DE COSTOS ENTRE EL MODELO TRADICIONAL DE COBRANZA Y EL NUEVO SISTEMA DE GESTIÓN BASADO EN SCORE PARA LA EMPRESA ABC.**

En esta sección analicé el impacto económico que tiene la implementación del modelo propuesto en los costos operativos de la gestión de cobranza prejudicial, comparando los dos escenarios establecidos con el uso del modelo actual o tradicional en ABC y del nuevo sistema de gestión basado en score. Con este fin tuve varias reuniones con la gerencia financiera administrativa, quien me proporcionó los datos necesarios para realizar esta tarea.

Como en toda empresa de servicios, la base de la cadena productiva en ABC es el recurso humano, en este caso los ejecutivos de cobranza, quienes generan el margen a la empresa en cada interacción con los clientes a los que se realiza una gestión de cobro. Como mencioné en el capítulo uno, el estudio se centra en las actividades que se realizan en el área de operaciones durante la fase prejudicial de la cobranza, por este motivo los procesos relacionados con la selección y capacitación no se consideran para el análisis, así como tampoco otras competencias diferenciadoras para generar valor a los dueños de la cartera o Cedentes, tales como el uso de plataformas emergentes o software libre, desarrollo de competencias técnicas, entre otras.

#### **Estructura de costos para la gestión de cobranza en ABC.**

La unidad de la estructura de costos en una empresa de cobranzas es el gestor o ejecutivo de cobranza, quien realiza la gestión a los clientes de las empresas que solicitan este servicio, constituyéndose en la materia prima del negocio; significa el 64% de la estructura de costos de ABC (Tabla 3.12) y en función de la cantidad de gestores se dimensiona la plataforma tecnológica y la infraestructura.

**Tabla 3.12** - Estructura de costos de ABC

	Porcentaje
Personal	64%
Tecnología	9%
Red	11%
Facilidades	16%

La metodología empleada por ABC para el cálculo del costo de la gestión operativa de cobranza, pondera el costo por concepto de operadores<sup>27</sup> y gasto administrativo<sup>28</sup>, para luego contrastar con los ingresos comerciales (gasto/ingreso) y determinar desde el punto de vista de los intereses propios de la empresa los márgenes aceptables para la cartera. A diferencia del esquema tradicional que solo considera el número de clientes para calcular la cantidad de gestores necesarios, como opción de mejora propuse la utilización del tiempo medio de gestión por cada interacción con el cliente en la cobranza telefónica y domiciliaria (Tabla 3.13), porque de esta forma se toma en cuenta no solo el volumen de la cartera, sino también el protocolo de servicio establecido para cada tipo de gestión y los factores que determinan la estrategia, tales como la edad de mora, el riesgo y la oportunidad comercial.

---

<sup>27</sup> Considera el sueldo de los gestores, más la remuneración variable y los beneficios de ley.

<sup>28</sup> Considera los costos del servicio telefónico, servicios básicos, movilización, cafetería y limpieza.

**Tabla 3.13** - Tiempo promedio de gestión por estrategia de cobranza

Estrategia	Tiempo medio de gestión	Duración (segundos)	Distribución del tiempo medio de gestión
Mora menor a 30 días	Presentación y saludo por parte del gestor telefónico	25	15,43%
	Información del crédito y negociación de la deuda.	92	56,79%
	Despedida y tramitación de la gestión en el sistema.	45	27,78%
	<b>Total</b>	162	<b>100,00%</b>
Mora mayor a 30 días	Presentación y saludo por parte del gestor telefónico	25	8,28%
	Información del crédito y negociación de la deuda.	177	58,61%
	Despedida y tramitación de la gestión en el sistema.	100	33,11%
	<b>Total</b>	302	<b>100,00%</b>
Riesgo alto	Presentación y saludo por parte del gestor telefónico	25	5,66%
	Información del crédito y negociación de la deuda.	317	71,72%
	Despedida y tramitación de la gestión en el sistema.	100	22,62%
	<b>Total</b>	442	<b>100,00%</b>
Oportunidad comercial	Presentación y saludo por parte del gestor telefónico	35	17,50%
	Información del crédito y negociación de la deuda.	120	60,00%
	Despedida y tramitación de la gestión en el sistema.	45	22,50%
	<b>Total</b>	200	<b>100,00%</b>

Utilizando la metodología planteada del tiempo medio de gestión, realizamos junto con ABC el cálculo del costo por hora en el caso de un gestor telefónico, definiendo previamente con los expertos de la cobranza: gerente de operaciones y gerente general de la empresa, el tamaño y los participantes de la infraestructura necesaria; por ejemplo cuántos gestores o cobradores, la definición de salarios e incentivos, el personal de supervisión y apoyo, los sistemas de gestión de la información y la infraestructura tecnológica que

permiten mantener los niveles de calidad en el servicio, como se muestra en la Tabla 3.14.

**Tabla 3.14** - Estructura de costos y factores de productividad para la cobranza telefónica en ABC

<b>Estructura de costos Telefonía</b>	<b>VALOR</b>
<b>Cobranza operativa prejudicial</b>	
<b>Salario bruto anual por categoría</b>	
<b>Personal de operación</b>	
Gestor	3.900
<b>Personal de estructura de centro</b>	
Coordinador	6.600
Formador	5.400
Monitor	5.400
Administrativo	10.800
Supervisor	14.400
Responsable de operaciones	24.000
<b>Otros factores que inciden en el Costo</b>	
<b>Costo empresa por empleado</b>	37,15%
<b>Días y horarios especiales</b>	
Plus por hora nocturna	25,0%
Plus por hora suplementaria	50,0%
Plus por hora extraordinaria	100,0%
<b>Estructura de Centro</b>	
Número de gestores por Supervisor	20
Número de gestores por Formador	50
Número de gestores por Monitor	50
Número de gestores por Administrativo	100
Número de gestores por Jefatura	100
Número de gestores por Responsable de operaciones	200
<b>Factores de productividad</b>	
Horas nominales por año	1.920
% de tiempo de descanso a la hora	2,4%
% de tiempo de formación a la hora	2,0%
% de ausentismo medio	5,0%
Tasa de ociosidad estructural	0,0%
Tiempo medio de operación objetivo	162
Tasa de efectividad en cobranza telefónica	64,8%

La estructura de costos que se muestra en la Tabla 3.15, desgrega el costo por hora de un ejecutivo de cobranza telefónica en la fase prejudicial, considerando aspectos como la estructura de la empresa y los factores de productividad, estos últimos relacionados con la cantidad de llamadas, la cantidad de llamadas con contacto y la cantidad de contactos que al final se convirtieron en pagos.

**Tabla 3.15** - Costo por hora para un ejecutivo de cobranza telefónica en ABC

<b>Estructura de costos Telefonía Cobranza operativa prejudicial</b>	<b>VALOR</b>
<b>Productividad</b>	
Minutos productivos por hora	54
Segundos productivos por hora	3.262
Llamadas gestionadas a la hora	20
Cobros por hora	13
<b>Costos</b>	
<b>Costo/Año por categoría</b>	
Ejecutivo de cobranza	5.349
Coordinador	9.052
Formador	7.406
Monitor	7.406
Administrativo	14.812
Supervisor	19.750
Responsable de operaciones	32.916
<b>Costo/Hora del Ejecutivo de cobranza</b>	
Ejecutivo de cobranza	2,7859
Parte propocional de Coordinadores	0,2357
Parte propocional de Formadores	0,0771
Parte propocional de Monitores	0,0771
Parte propocional de Administrativos	0,0771
Parte propocional de Supervisores	0,1029
Parte propocional de Responsable de operaciones	0,0857
Parte propocional de otras categorías	0,0000
<b>Total Costo/hora Ejecutivo de cobranza telefónica</b>	<b>3,44</b>

Finalmente el cálculo del costo total por hora para un gestor de cobranza telefónica en ABC, se obtiene añadiendo los costos directos y otros gastos por hora, tal como se muestra en la Tabla 3.16.

**Tabla 3.16** - Costos y gastos por hora para un ejecutivo de cobranza telefónica en ABC

<b>Estructura de costos Telefonía Cobranza operativa prejudicial</b>	<b>VALOR</b>
<b>Otros costos y gastos</b>	
Alquiler de local	0,03
Computador	0,11
Servicios de agua y luz	0,06
Seguros	0,07
Comunicaciones (teléfono, red)	0,35
Suministros, varios	0,02
<b>Total costos directos por hora</b>	<b>4,08</b>
Gastos generales	0,06
Gastos en investigación y desarrollo	0,54
Márketing	0,16
<b>Total costos y gastos operativos por hora</b>	<b>4,84</b>
Depreciación	0,27
Incobrables	0,08
Gastos financieros	0,08
<b>Total costos y gastos antes de impuestos por hora</b>	<b>5,27</b>
Costo/Segundo productivo del Ejecutivo de cobranza telefónica	0,00162
Costo de la llamada	0,26184

Para la gestión de cobranza en terreno la estructura de costos disgrega el valor por cada visita donde se ha contactado personalmente al titular del crédito. El análisis completo para este caso del costo por hora de un gestor de campo consta en el Anexo VII.

**Comparación del costo de la gestión operativa de cobranza entre el sistema tradicional y el nuevo basado en score.**

Para cuantificar el impacto económico de la implementación del nuevo sistema de gestión de cobranzas, junto con ABC establecimos el uso de un esquema basado en la determinación del esfuerzo por cobrador, calculando el número de ejecutivos de cobranza en telefonía y terreno que se necesitan para gestionar un segmento particular de la cartera con una estrategia definida.

En atención a la naturaleza dinámica del negocio, en especial por aspectos de tecnología y de servicio a los clientes, consideré primero varias premisas para elaborar este análisis, de manera que los resultados puedan ser comparables:

1. Toda la infraestructura y plataforma tecnológica está montada, de manera que no hacen falta nuevas inversiones en activos fijos, tales como software, hardware, equipos de aire acondicionado, generadores de energía, muebles y enceres, inmuebles.
2. Los gastos generales que incluyen el alquiler el local donde opera ABC, el servicio de telefonía fija y celular, los seguros para los equipos, los servicios públicos, los suministros de oficina y las tasas municipales, se mantienen en los mismos valores.
3. Los sueldos de los ejecutivos de cobranza, personal administrativo y personal de soporte no cambian.
4. No se contrata personal adicional para los servicios de mensajería y limpieza.
5. La efectividad en la cobranza no se ve afectada durante la primera etapa de implementación del nuevo sistema de gestión.

### Definición del número de cobradores telefónicos por estrategia.

Analizamos las cifras y valores históricos referidos a la cantidad de clientes por ejecutivo de cobranza, tiempo promedio de cada llamada y rendimiento de cada operador por día, de la manera como se plantea en el ejemplo de la Tabla 3.17. Los pasos a seguir para cumplir esta actividad son los siguientes:

1. Determinar una estrategia según el tipo de cartera.
2. Definir el tiempo promedio de cada llamada telefónica, de acuerdo a la experiencia propia de ABC y a los acuerdos comerciales con el Cedente.
3. Definir un número de llamadas mínimas por cada cliente.
4. Revisar la duración de la jornada laboral por día, semana o mes.
5. Calcular la cantidad de ejecutivos telefónicos.

**Tabla 3.17** - Cálculo del número de gestores telefónicos requeridos para ejecutar una estrategia de cobranza

DEFINICIÓN DEL NÚMERO DE COBRANZADORES TELEFÓNICOS ESTRATEGIA XX		MES xx
Concepto	Cálculo	Indicador
Tiempo promedio de llamada		<b>2,7 minutos</b>
Total de minutos productivos día	54 * 7 horas diarias	<b>381 minutos</b>
Cantidad de llamadas posibles día	381 / 2,7	<b>141 llamadas día</b>
Cantidad de llamadas posibles mes	141 * 22 días	<b>3104 llamadas mes</b>
<hr/>		
Total de clientes en el rango X		<b>3500</b>
Llamadas por cliente requeridas		<b>2,5 llamadas</b>
Total llamadas requeridas	3500 * 2,5	<b>8750 llamadas</b>
Número mínimo de gestores telefónicos para gestionar los clientes del rango X	8750 / 3104	<b>3 gestores</b>



### Definición del número de cobradores de terreno por cada estrategia.

En general corresponde analizar cifras y valores históricos de acuerdo a la experiencia propia de ABC y los acuerdos comerciales con el Cedente, tales como rendimiento de los gestores en cuanto a número de visitas por día, la cantidad de clientes sin teléfonos que requieren gestión de terreno, la cartera que estando en telefonía necesita apoyo de la cobranza en campo. De forma similar al caso de los ejecutivos telefónicos, construí el siguiente análisis para el cálculo de los gestores requeridos para realizar las visitas (Tabla 3.18).

**Tabla 3.18** - Cálculo del número de gestores domiciliarios requeridos para ejecutar una estrategia de cobranza

DEFINICIÓN DEL NÚMERO DE COBRANZADORES EN TERRENO		
ESTRATEGIA XX		MES xx
CONCEPTO	CÁLCULO	INDICADOR
Cantidad de visitas día por cobrador		<b>25 visitas</b>
Cantidad de visitas posibles mes	25 * 22 días laborales	<b>550 visitas</b>
_____	_____	_____
Cartera Inicial para terreno		<b>240 clientes</b>
Estimado de rebalse	3500 * 10%	<b>350 clientes</b>
Apoyo de terreno a telefonía	3500 * 5%	<b>175 clientes</b>
Cartera total en terreno	240 + 350 + 175	<b>765 clientes</b>
_____	_____	_____
Visitas por cliente requeridas		<b>1,5 visitas</b>
Total visitas requeridas	765 * 1,5	<b>1148 visitas</b>
Número mínimo de gestores domiciliarios para gestionar los clientes de la estrategia XX	1148 / 550	<b>2 gestores</b>

Siguiendo la metodología descrita calculamos con ABC el número de ejecutivos de cobranza requeridos por cada estrategia, basándonos en los clientes de la

muestra de prueba que ya estuvieron segmentados, tanto para el esquema tradicional (Tabla 3.19) como para el nuevo sistema basado en score (Tabla 3.20)

**Tabla 3.19** - Número de clientes por estrategia en el esquema tradicional de cobranza de ABC

SISTEMA TRADICIONAL DE COBRANZA EN ABC							
Número de estrategia	Edad de mora	Rango monto *	Con direcciones y teléfonos	Solo con direcciones	Solo con teléfonos	Sin direcciones ni teléfonos	Total general
1	Entre 1 y	monto alto	92	34			126
2	15 días	monto bajo	124	54	2		180
3	Entre 16 y	monto alto	28	3	1		32
4	30 días	monto bajo	51	12	1	1	65
5	Entre 31 y	monto alto	15				15
6	60 días	monto bajo	36	3			39
7	Entre 61 y	monto alto	2				2
8	90 días	monto bajo	14	1	1		16
9	Mayor a 90	monto alto	2				2
10	días	monto bajo	5				5
Total general			369	107	5	1	482

\* monto alto: montos mayores o iguales al promedio, monto bajo: saldos menores al promedio

**Tabla 3.20** - Número de clientes por estrategia en el nuevo esquema de cobranza basado en score.

SISTEMA DE COBRANZA BASADO EN SCORE									
Número de estrategia	Score	Sucursal	Cuotas Pagadas	Tramo de mora	Con direcciones y teléfonos	Solo con direcciones	Solo con teléfonos	Sin direcciones ni teléfonos	Total general
1	Alto	Indiferente	Indiferente	1 a 15 días	98	58	1	0	157
2	Bajo-Zona Gris	Castigo	Indiferente	1 a 15 días	76	14	1	0	91
3	Bajo-Zona Gris	Premio	Indiferente	1 a 15 días	42	16	0	0	58
4	Alto	Indiferente	Indiferente	16 a 30 días	33	10	1	0	44
5	Bajo-Zona Gris	Indiferente	Castigo	16 a 30 días	1	0	0	0	1
6	Bajo-Zona Gris	Indiferente	Premio	16 a 30 días	45	5	2	0	52
7	Alto	Indiferente	Indiferente	30 a 60	25	3	0	0	28
8	Bajo-Zona Gris	Indiferente	Indiferente	30 a 60	26	0	0	0	26
9	Indiferente	Indiferente	Indiferente	61 a 90 días	16	1	1	0	18
10	Indiferente	Indiferente	Indiferente	91 a 120 días	7	0	0	0	7
11	Indiferente	Indiferente	Indiferente	días	0	0	0	0	0
Total general					369	107	5	1	482

**Tabla 3.21** - Gestores telefónicos requeridos con el sistema tradicional de ABC.

SISTEMA TRADICIONAL DE COBRANZA EN ABC						
Estrategia	Clientes a ser gestionados en telefonía	Tiempo medio de llamada	Cantidad llamadas posibles mes cobrador	Llamadas por cliente requeridas	Total llamadas requeridas	Cantidad de gestores telefónicos
1	92	162	3.104	3	276	0,09
2	126	162	3.104	2,5	315	0,10
3	29	162	3.104	4,5	130,5	0,04
4	52	162	3.104	4	208	0,07
5	15	302	1.665	6,5	97,5	0,06
6	36	302	1.665	6	216	0,13
7	2	442	1.138	8,5	17	0,01
8	15	442	1.138	8	120	0,11
9	2	442	1.138	12	24	0,02
10	5	442	1.138	10	50	0,04
Total general	374	3020	1.836	6,50	2.431	0,67

**Tabla 3.22** - Gestores telefónicos requeridos con el sistema de cobranza basado en score.

SISTEMA DE COBRANZA BASADO EN SCORE						
Número de estrategia	Clientes a ser gestionados en telefonía	Tiempo medio de llamada	Cantidad llamadas posibles mes cobrador	Llamadas por cliente requeridas	Total llamadas requeridas	Cantidad de gestores telefónicos
1	99	162	3.104	3	297	0,10
2	77	162	3.104	2,5	192,5	0,06
3	42	200	2.515	2,5	105	0,04
4	34	162	3.104	4	136	0,04
5	1	302	1.665	3	3	0,00
6	47	200	2.515	3	141	0,06
7	25	302	1.665	6	150	0,09
8	26	302	1.665	5,5	143	0,09
9	17	302	1.665	6,5	110,5	0,07
10	7	442	1.138	7	49	0,04
11	0	0	-	0	0	-
Total general	375	2536	2.405	3,91	1.466	0,59

**Tabla 3.23** - Gestores domiciliarios requeridos con el sistema tradicional de ABC

SISTEMA TRADICIONAL DE COBRANZA EN ABC					
Estrategia	Clientes a ser gestionados en campo	Cantidad llamadas posibles mes cobrador	Visitas por cliente requeridas	Total visitas requeridas	Cantidad de gestores domiciliarios
1	126	550	2	252	0,46
2	72,6	550	1,5	108,9	0,20
3	31	550	3	93	0,17
4	37,5	550	2,5	93,75	0,17
5	15	550	4	60	0,11
6	39	550	3,5	136,5	0,25
7	2	550	4,5	9	0,02
8	15	550	4	60	0,11
9	2	550	5,5	11	0,02
10	5	550	5	25	0,05
Total general	345,1	5500	3,55	849,15	1,54

**Tabla 3.24** - Gestores domiciliarios con el sistema de cobranza basado en score

SISTEMA DE COBRANZA BASADO EN SCORE					
Número de estrategia	Clientes a ser gestionados en campo	Cantidad llamadas posibles mes cobrador	Visitas por cliente requeridas	Total visitas requeridas	Cantidad de gestores domiciliarios
1	156	550	1,5	234	0,43
2	90	550	1,5	135	0,25
3	0	550	1,5	0	-
4	43	550	2,5	107,5	0,20
5	1	550	1,5	1,5	0,00
6	50	550	1,5	75	0,14
7	28	550	3,5	98	0,18
8	26	550	2,5	65	0,12
9	17	550	4,5	76,5	0,14
10	7	550	5,5	38,5	0,07
11	0	0	0	0	-
Total general	418	5500	2,36	831	1,51

**Tabla 3.25** - Costo mensual de la gestión operativa con el sistema tradicional de ABC.

SISTEMA TRADICIONAL DE COBRANZA EN ABC							
Estrategia	Cantidad de gestores telefónicos	Cantidad de gestores domiciliarios	Costo hora por gestor telefónico	Costo hora por gestor de campo	Total costo mes gestión telefónica	Total costo mes gestión de campo	Total costo Gestión operativa
1	0,08	0,46	\$ 5,27	\$ 6,47	\$ 68,00	\$ 177,87	\$ 245,87
2	0,09	0,20	\$ 5,27	\$ 6,47	\$ 77,61	\$ 76,86	\$ 154,48
3	0,04	0,17	\$ 5,27	\$ 6,47	\$ 32,15	\$ 65,64	\$ 97,79
4	0,06	0,17	\$ 5,27	\$ 6,47	\$ 51,25	\$ 66,17	\$ 117,42
5	0,05	0,11	\$ 5,27	\$ 6,47	\$ 44,78	\$ 42,35	\$ 87,13
6	0,12	0,25	\$ 5,27	\$ 6,47	\$ 99,21	\$ 96,34	\$ 195,56
7	0,01	0,02	\$ 5,27	\$ 6,47	\$ 11,43	\$ 6,35	\$ 17,78
8	0,10	0,11	\$ 5,27	\$ 6,47	\$ 80,67	\$ 42,35	\$ 123,02
9	0,02	0,02	\$ 5,27	\$ 6,47	\$ 16,13	\$ 7,76	\$ 23,90
10	0,04	0,05	\$ 5,27	\$ 6,47	\$ 33,61	\$ 17,65	\$ 51,26
Total general	0,61	1,54	\$ 5,27	\$ 6,47	\$ 514,86	\$ 599,35	\$ 1.114,21

**Tabla 3.26** - Costo mensual de la gestión operativa con el sistema basado en score.

SISTEMA DE COBRANZA BASADO EN SCORE							
Número de estrategia	Cantidad de gestores telefónicos	Cantidad de gestores domiciliarios	Costo hora por gestor telefónico	Costo hora por gestor de campo	Total costo mes gestión telefónica	Total costo mes gestión de campo	Total costo Gestión operativa
1	0,09	0,43	\$ 5,27	\$ 6,47	73,18	165,16	238,34
2	0,06	0,25	\$ 5,27	\$ 6,47	47,43	95,29	142,72
3	0,04	-	\$ 5,27	\$ 6,47	31,94	-	31,94
4	0,04	0,20	\$ 5,27	\$ 6,47	33,51	75,88	109,38
5	0,00	0,00	\$ 5,27	\$ 6,47	1,38	1,06	2,44
6	0,05	0,14	\$ 5,27	\$ 6,47	42,89	52,94	95,83
7	0,08	0,18	\$ 5,27	\$ 6,47	68,90	69,17	138,07
8	0,08	0,12	\$ 5,27	\$ 6,47	65,68	45,88	111,56
9	0,06	0,14	\$ 5,27	\$ 6,47	50,75	54,00	104,75
10	0,04	0,07	\$ 5,27	\$ 6,47	32,94	27,17	60,11
11	-	-	-	-	-	-	-
Total general	0,53	1,51	4,79	6,47	448,60	586,53	1.035,13

## **Conclusiones**

El nuevo sistema de gestión de cobranza basado en score utiliza menos gestores telefónicos (Tablas 3.21 y 3.22) que el sistema de cobranza tradicional de ABC, lo que traducido en costos representa un 13% de ahorro con el nuevo sistema.

El nuevo sistema de gestión de cobranza basado en score utiliza también menos gestores domiciliarios (Tablas 3.23 y 3.24) que el sistema de cobranza tradicional, resultado que en costos representa un 2.13% de ahorro con el nuevo sistema.

En términos del costo total de la cobranza operativa: gestión telefónica más gestión domiciliaria, el nuevo sistema basado en score es 7.4% más barato que el tradicional (Tablas 3.25 y 3.26), ahorro que se da principalmente por la reducción de gestores en el área de telefonía.

## **CAPÍTULO IV**

### **IMPLEMENTACIÓN DEL NUEVO SISTEMA DE GESTIÓN DE COBRANZA EN ABC**

En ese capítulo desarrollé el plan de implementación del nuevo sistema de gestión de cobranza basado en score, con las respectivas mediciones que cuantifiquen el esfuerzo y resultado logrado, para reaccionar oportunamente a los cambios que podrían afectar su desempeño. Lo primero fue determinar si el modelo estadístico trabaja según se diseñó en la etapa de desarrollo, estableciendo los indicadores básicos que valoran su eficiencia en el tiempo y los reportes que permiten monitorear los eventos significativos para entender con mayor precisión qué está sucediendo, cómo está sucediendo y por qué podría estar sucediendo. Finalmente para cerrar el proyecto, dado que el sistema de gestión funciona a cabalidad en la medida que su comunicación y control sean los requeridos, determiné la línea de acción para una planificación de trabajo sólida y oportuna, el manejo de información y conocimiento, la descripción de las actividades y los métodos de trabajo con los colaboradores y otros participantes, de forma que en revisiones futuras se pueda adoptar las medidas de conservación o corrección que aseguren la consecución de los objetivos de ABC.

#### **4.1 MONITOREO DEL MODELO ESTADÍSTICO**

Primero realicé un análisis de las necesidades de información para valorar la eficiencia del modelo y luego junto con ABC consideramos tres tipos de reportes, dependiendo del objeto y la etapa en la que se encuentre esta evaluación, en medida de los indicadores que se necesita medir y con qué prioridad. La generación de estos reportes requiere archivos con las variables consideradas en el desarrollo del modelo, así como el score calculado y la implementación de las estrategias.

**Reportes de estabilidad y comportamiento.** Sirven para verificar la estabilidad de las variables y que los comportamientos esperados se sigan manteniendo. Para esto se cruza la información de la variable contra el indicador de buenos y malos; en el caso de las variables del score se realizan análisis de frecuencias con el fin de garantizar la información requerida por el modelo.

**Reportes de perfil de clientes.** El objetivo de estos reportes es hacer seguimiento a los indicadores generales relacionados con el modelo, tales como su poder de discriminación y la adaptación en el sistema de cobranza para la creación de estrategias de gestión diferenciadas por segmento de riesgo.

**Reportes Back Testing.** Los reportes de Back Testing tienen la finalidad de observar la estabilidad del score y de la población, comparando su comportamiento al momento del estudio y en un período posterior.

#### **4.1.1 TIEMPOS Y FRECUENCIAS INVOLUCRADAS EN LOS REPORTE**

Los reportes tienen una frecuencia mensual, salvo que dentro de un reporte específico se mencione una frecuencia diferente. Por este motivo establecí notaciones de  $t$  en el período de referencia que se esté analizando y  $t+1$  ó  $t-1$  en caso de que una variable o una medida deba ser calculada en el mes siguiente o el mes anterior.

#### **4.1.2 GENERACIÓN DE INDICADORES Y REPORTE**

Dentro de los reportes se incluyen una serie de indicadores que sirven de medición de aquellas variables analizadas como dimensiones del reporte, incluyendo las notaciones de tiempo en la que se genera esta medida. La población sujeta de reporte y análisis en un determinado período, es aquella que se presenta en cobranza dentro de ABC como cartera microcrédito, que no tenga otros productos vigentes en el mismo Cedente. Los indicadores utilizados como medida son los siguientes:

### Indicador de buenos y malos

Relaciona la cantidad de clientes buenos o malos respecto al total. Permite medir la evolución de la cartera de un mes al otro, en el sentido de recuperación o deterioro del riesgo de los clientes en cobranza.

$$\text{Indicador de buenos/malos}_t = \frac{\text{número de clientes buenos/malos}_t}{\text{total de clientes}_t}$$

La cascada o matriz de rodamiento que muestro en la Figura 4.1, sirve de ejemplo para presentar la definición del indicador de buenos y malos. Un cliente que en el período t+1 haya cancelado su saldo total o se haya puesto al día, independientemente de la mora que presente en el período t, es definido como bueno en el período t. Los demás clientes existentes en el período t son considerados como malos.

Edad de mora período t+1	Edad de mora período t						
	0	1-15	16-30	31-60	61-90	91-120	>120
0	<b>BUENOS</b>						
Cancelados (pago total del saldo del crédito)							
1-15	<b>MALOS</b>						
16-30							
31-60							
61-90							
91-120							
>120							

**Figura 4.1** - Cascada o matriz de rodamiento período t vs t+1

### Indicador de recaudo

Representa el pago total de los valores vencidos y por vencer en un período t, más los recargos que haya generado la mora, respecto al saldo total en cobranza dentro del período t. Un indicador de recaudo del 100% es equivalente a que el cliente siga al día en caso de estar al día y a que se ponga al día en caso de estar



en mora. Un indicador de recaudo del 100% no representa cancelación total del saldo del crédito.

$$\text{Indicador de recaudo}_t = \frac{\text{saldo liberado}_t}{\text{saldo total}_t}$$

### Indicador de cargo

El valor del cargo representa la diferencia en provisiones<sup>29</sup> entre el período t y t-1, entendiéndose que en el escenario de cero recuperaciones esta diferencia será máxima (Figura 4.2).

Clasificación	Provisión del período (t-1)	Provisión del período (t)	Valor de cargo
A	Saldo en A * Porcentaje de provisión en A	Saldo en A * Porcentaje de provisión en A	Diferencia de provisiones, períodos (t) y (t-1)
B	Saldo en B * Porcentaje de provisión en B	Saldo en B * Porcentaje de provisión en B	Diferencia de provisiones, períodos (t) y (t-1)
C	Saldo en C * Porcentaje de provisión en C	Saldo en C * Porcentaje de provisión en C	Diferencia de provisiones, períodos (t) y (t-1)
D	Saldo en D * Porcentaje de provisión en D	Saldo en D * Porcentaje de provisión en D	Diferencia de provisiones, períodos (t) y (t-1)
E	Saldo en E * Porcentaje de provisión en E	Saldo en E * Porcentaje de provisión en E	Diferencia de provisiones, períodos (t) y (t-1)
Provisión total	Sumatoria de las provisiones en cada clasificación, período t-1	Sumatorio de las provisiones en cada clasificación, período t	Diferencia de provisiones, períodos (t) y (t-1)

**Figura 4.2** - Cálculo del cargo por edad de mora entre el período t y t-1

Este indicador muestra en porcentaje, la cantidad de cargo máximo que se redujo producto del pago de clientes, en el período t. Si existe un valor de cargo negativo se llama liberación de provisiones.

<sup>29</sup> Para la gestión del riesgo de crédito suele utilizarse el concepto de pérdida esperada de una operación, que es la esperanza matemática de un posible quebranto y generalmente se calcula como el producto de las probabilidades de incumplimiento, tamaño de la deuda y las pérdidas en caso de incumplimiento. Las entidades financieras en Ecuador están obligadas por la Superintendencia de Bancos y Seguros a mantener reservas o provisiones en porcentajes mínimos y máximos, para cubrir riesgos de incobrabilidad o pérdida del valor de los activos de riesgo, mismas que se hacen con cargo al estado de pérdidas y ganancias de dichas instituciones.

$$\text{Indicador de carga}_t = 1 - \frac{\text{carga real}_t}{\text{carga máximo}_t}$$

### Reporte de calidad de información y estabilidad de comportamientos

Un aspecto importante es que la funcionalidad de los modelos basados en score depende, entre otras cosas, de la información requerida o de relevancia, por tal razón es importante periódicamente revisar la calidad de información, porque sobre ella se soporta el modelo y su correcto funcionamiento. La validación de la calidad de la información contempla dos aspectos: que sea suficiente y completa, es decir que la totalidad de la población la contenga, y el segundo es la coherencia de los datos.

El objetivo de este reporte es observar mensualmente el comportamiento de las variables del score y adicionales que interesan a abc, versus el indicador de buenos y malos (Figura 4.3).

DISEÑO TÉCNICO		(Prioridad 1)				
Nombre del Reporte	REPORTE DE CALIDAD DE INFORMACIÓN Y ESTABILIDAD DE COMPORTAMIENTOS					
Frecuencia del Reporte	Mensual					
Variables que intervienen	Variables análisis e indicador de buenos y malos.					
POST-REPORTE						
Conclusiones						
Acciones						
DISEÑO REPORTE						
INDICADOR DE BUENOS Y MALOS PERÍODO t.						
VARIABLE PERÍODO t	buenos	% buenos	malos	% malos	TOTAL	% TOTAL
Categoría 1	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna
Categoría 2	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna
TOTAL	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna

Figura 4.3 – Reporte de calidad de información y estabilidad de comportamientos

### Reporte de recaudo, cargo y deterioro por estrategia y edad de mora

Este reporte tiene como finalidad conocer el deterioro presentado en cada una de las estrategias, para observar mejoras en el comportamiento y eficiencia en la gestión (Figuras 4.4 y 4.5).

FICHA TÉCNICA		(Prioridad 1)										
Nombre del Reporte	RECAUDO, CARGO Y DETERIORO POR ESTRATEGIA Y MORA											
Frecuencia del Reporte	Mensual											
Variables que intervienen	Estrategias, tramos de mora, indicador de buenos y malos, número de clientes, recaudo y cargo.											
POST-REPORTE												
Conclusiones												
Acciones												
DISEÑO DEL REPORTE												
		INDICADOR BUENOS Y MALOS (Deterioro t a t+1)							RECAUDO t		PROVISIÓN t	
EDAD DE MORA (t)	ESTRATEGIA (t)	BUENOS	% BUENOS	MALOS	% MALOS	% MALOS ANTERIOR (t-1)	TOTAL	TOTAL ANTERIOR	% RECAUDO (t)	% RECAUDO ANTERIOR (t-1)	VALOR CARGO (t)	VALOR CARGO ANTERIOR (t-1)
0	1	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	
0	2	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	
SUBTOTAL MORA		# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR	recaudo/ deuda	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR	valor cargo	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR
1-15	3	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	
16-30	4	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	
31-60	5	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	
61-90	6	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	
.....	.....	# Clientes	buenos/ total	# Clientes	malos/ total		# Clientes		recaudo/ deuda		valor cargo	

Figura 4.4 – Reporte de recaudo, cargo y deterioro por estrategia y mora

FICHA TÉCNICA		(Prioridad 1)									
Nombre del Reporte		RECAUDO, CARGO Y DETERIORO POR RESULTADO DE GESTIÓN									
Frecuencia del Reporte		Mensual									
Variables que intervienen		Estrategia, resultado de gestión, tramos de mora, indicador de buenos y malos, clientes, recaudo y cargo.									
POST-REPORTE											
Conclusiones											
Acciones											
DISEÑO DEL REPORTE											
TOTAL CLIENTES Y DETERIORO								RECAUDO (t)		PROVISIÓN (t)	
ESTRATEGIA (t)	TIPO DE GESTIÓN (t)	RESULTADO GESTIÓN (t)	TOTAL	% TOTAL	% TOTAL ANTERIOR	% MALOS	% MALOS ANTERIOR (t-1)	% RECAUDO	% RECAUDO ANTERIOR (t-1)	VALOR CARGO	VALOR CARGO ANTERIOR (t-1)
1	telefónica	promesa de pago	# Clientes	clientes/subtotal estrategia	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR	malos/total	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR	recaudo/deuda	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR	valor cargo	EXTRAER DATOS DEL MISMO REPORTE MES ANTERIOR
1	telefónica	sin contacto	# Clientes	clientes/subtotal estrategia		malos/total		recaudo/deuda		valor cargo	
1	telefónica	sin gestión	# Clientes	clientes/subtotal estrategia		malos/total		recaudo/deuda		valor cargo	
SUBTOTAL DE ESTRATEGIA			# Clientes	estrategia/total clientes		malos/total		recaudo/deuda		valor cargo	

Figura 4.5 – Reporte de recaudo, cargo y deterioro por resultado de gestión

### Reportes para evaluar el score

El reporte de mayor relevancia para evaluar el modelo basado en score es la distribución por rangos, esta distribución debe ser similar a la que se tenía en el estudio. Se calculan las diferencias en las proporciones de clientes esperados para cada rango y para ello se relaciona el porcentaje de malos por rango con el porcentaje de malos para toda la población (Figura 4.6).

DISEÑO TÉCNICO		(Prioridad 2)					
Nombre del Reporte		REPORTE DISTRIBUCIÓN SCORECARD					
Frecuencia del Reporte		Mensual					
Variables que intervienen		Score, indicador de buenos y malos, número de clientes, recaudo					
POST-REPORTE							
Conclusiones							
Acciones							
DISEÑO REPORTE							
INDICADOR DE BUENOS Y MALOS t							
RANGO SCORE (t)	buenos	%fila	malos	% malos	TOTAL	% TOTAL	RECAUDO (t)
J	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
I	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
H	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
G	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
F	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
E	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
D	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
C	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
B	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
A	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total
TOTAL	# Clientes	%fila	# Clientes	%fila	# Clientes	% columna	recaudo/total

Figura 4.6 – Reporte distribución de scorecard

### Reporte mapa de estrategias

El presente reporte sirve para evaluar la discriminación que está logrando el score por banda y por edad de mora, se utiliza por tanto para controlar el comportamiento de los segmentos de cada estrategia (Figura 4.7).

### Reporte cálculo de KS

Permite ver la diferenciación del score partiendo de porcentajes acumulados, muestra la distancia entre buenos y malos que idealmente debe ser superior al 40% (Figura 4.8).

DISEÑO TÉCNICO		(Prioridad 1)							
Nombre del Reporte		REPORTE MAPA DE ESTRATEGIAS							
Frecuencia del Reporte		Mensual.							
Variables que intervienen		Score, indicador de buenos y malos, recaudo.							
POST-REPORTE									
Conclusiones									
Acciones									
DISEÑO REPORTE									
EDAD DE MORA (t)									
RANGO DE SCORE (t)	1-30		31-60		>60		TOTAL		
A	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	
B	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	
C	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	
TOTAL	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	# Clientes	recaudo	

Figura 4.7 - Reporte de mapa de estrategias

DISEÑO TÉCNICO		(Prioridad 2)							
Nombre del Reporte		REPORTE CÁLCULO DE KS							
Frecuencia del Reporte		Mensual.							
Variables que intervienen		Score, indicador de buenos y malos, clientes							
POST-REPORTE									
Conclusiones									
Acciones									
DISEÑO REPORTE									
INDICADOR DE BUENOS Y MALOS (t)									
RANGO SCORE (t)	buenos	buenos acumulad	% buenos	malos	malos acumulad	% malos	TOTAL	total acumulad	% TOTAL
J	# Clientes	No. buenos(J)	%columna	# Clientes	No. malos(J)	%columna	# Clientes	# Clientes Totales(J)	%columna
I	# Clientes	No. buenos(J+I)	%columna	# Clientes	No. malos(J+I)	%columna	# Clientes	# Clientes Totales (J+I)	%columna
H	# Clientes	No. buenos(J+H)	%columna	# Clientes	No. malos (J+H)	%columna	# Clientes	# Clientes Totales (J+H)	%columna
G	# Clientes	No. buenos(J+H+G)	%columna	# Clientes	No. malos (J+H+G)	%columna	# Clientes	# Clientes Totales (J+H+G)	%columna
F	# Clientes	No. buenos(J+H...+F)	%columna	# Clientes	No. malos (J+H...+F)	%columna	# Clientes	# Clientes Totales (J+H...+F)	%columna
E	# Clientes	No. buenos(J+H...+E)	%columna	# Clientes	No. malos (J+H...+E)	%columna	# Clientes	# Clientes Totales (J+H...+E)	%columna
D	# Clientes	No. buenos(J+H...+D)	%columna	# Clientes	No. malos (J+H...+D)	%columna	# Clientes	# Clientes Totales (J+H...+D)	%columna
C	# Clientes	No. buenos(J+H...+C)	%columna	# Clientes	No. malos (J+H...+C)	%columna	# Clientes	# Clientes Totales (J+H...+C)	%columna
B	# Clientes	No. buenos(J+H...+B)	%columna	# Clientes	No. malos (J+H...+B)	%columna	# Clientes	# Clientes Totales (J+H...+B)	%columna
A	# Clientes	No. buenos(J+H...+A)	100%	# Clientes	No. malos (J+H...+A)	100%	# Clientes	# Clientes Totales (J+H...+A)	100%

Figura 4.8 - Reporte cálculo de KS

### Reporte comparación de distribución del score

La comparación permite observar cómo era la participación de buenos y malos por score al momento del estudio y cómo se comporta en un período escogido (Figura 4.9).

DISEÑO TÉCNICO		(Prioridad 2)						
Nombre del Reporte	REPORTE COMPARACIÓN DISTRIBUCIÓN BUENOS Y MALOS							
Frecuencia del Reporte	Mensual.							
Variables que intervienen	Rangos score, buenos y malos por banda score en estudio, buenos y malos por banda score al periodo futuro escogido.							
POST-REPORTE								
Conclusiones								
Acciones								
DISEÑO REPORTE								
RANGO SCORE	Buenos por score en la fecha de cálculo (PO)	% Buenos Observado	Buenos por score (t)	% Buenos Columna	Malos por score en la fecha de cálculo (PO)	% Malos Observado	Malos por score (t)	% Malos Columna
J	EXTRAER DE LOS DATOS OBTENIDOS EN LA FECHA DE CÁLCULO DEL MODELO(PO)	% columna	# Clientes	% colum	EXTRAER DE LOS DATOS OBTENIDOS EN LA FECHA DE CÁLCULO DEL MODELO(PO)	% colum	# Clientes	% columna
I		% columna	# Clientes	% colum		% colum	# Clientes	% columna
H		% columna	# Clientes	% colum		% colum	# Clientes	% columna
G		% columna	# Clientes	% colum		% colum	# Clientes	% columna
F		% columna	# Clientes	% colum		% colum	# Clientes	% columna
E		% columna	# Clientes	% colum		% colum	# Clientes	% columna
D		% columna	# Clientes	% colum		% colum	# Clientes	% columna
C		% columna	# Clientes	% colum		% colum	# Clientes	% columna
B		% columna	# Clientes	% colum		% colum	# Clientes	% columna
A		% columna	# Clientes	% colum		% colum	# Clientes	% columna
TOTAL		100.0%	# Clientes	100.0%		100.0%	# Clientes	100.0%

Figura 4.9 – Reporte comparación distribución buenos y malos.

### Reporte back testing – psi (population stability index)

Este indicador se utiliza para evaluar el cambio en la población en el transcurso del tiempo (Figura 4.10), midiendo la separación entre las distribuciones del score del período de desarrollo (E) y el período de validación (O); si existen en el

transcurso del tiempo cambios en la población, estos pueden resultar en cambios en el score. El psi se calcula con la siguiente fórmula:

$$PSI = \frac{\sum \left[ (O - E) * \ln\left(\frac{O}{E}\right) \right]}{100}$$

donde si el PSI es inferior a 0.1 representa que el modelo presenta estabilidad.

<b>DISEÑO TÉCNICO</b>		(Prioridad 2)				
<b>Nombre del Reporte</b>		REPORTE PSI				
<b>Frecuencia del Reporte</b>		Mensual.				
<b>Variables que intervienen</b>		Score, % malos por banda score en estudio, % malos por banda score un periodo atrás (equivalente a tiempo de maduración)				
<b>POST-REPORTE</b>						
<b>Conclusiones</b>						
<b>Acciones</b>						
<b>DISEÑO REPORTE</b>						
<b>RANGO SCORE</b>	<b>% Malos por score en la fecha de cálculo PO (Esperado=E)</b>	<b>% Malos por score (t) (observado=O)</b>	<b>O-E</b>	<b>ln (O/ E)</b>	<b>sum [(O-E)*ln(O/ E)]</b>	<b>PSI</b>
L	Observar documento estudio	Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
K		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
J		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
I		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
H		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
G		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
F		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
E		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
D		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
C		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
B		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
A		Extraer del reporte de distribución de la Scorecard	O-E por fila	ln (O/ E) por fila	[(O-E)*ln(O/ E)]	
<b>TOTAL</b>						<b>sum [(O-E)*ln(O/ E)]</b>

Figura 4.10 – Reporte PSI.



## **4.2 ACCIONES Y ACTIVIDADES NECESARIAS PARA LA IMPLEMENTACIÓN**

El tener un objetivo de trabajo, una estructura y una base informática para desarrollar las diferentes gestiones y procesos asociados a la cobranza, no bastan al momento de implementar un sistema basado en estrategias, porque implica trabajar con una nueva filosofía orientada al cliente y para eso es necesario ir construyendo una cultura dentro de la empresa. Voy a plantear las acciones básicas necesarias que se deben tomar en cuenta para la implementación de este nuevo esquema de cobro, para maximizar el uso de todos los medios y herramientas que están disponibles. Sin embargo, antes de ejecutar cada una de estas acciones, fue necesario definir y establecer cuáles serán las consideraciones de corte estratégico que están envueltas en todo este proceso de implementación.

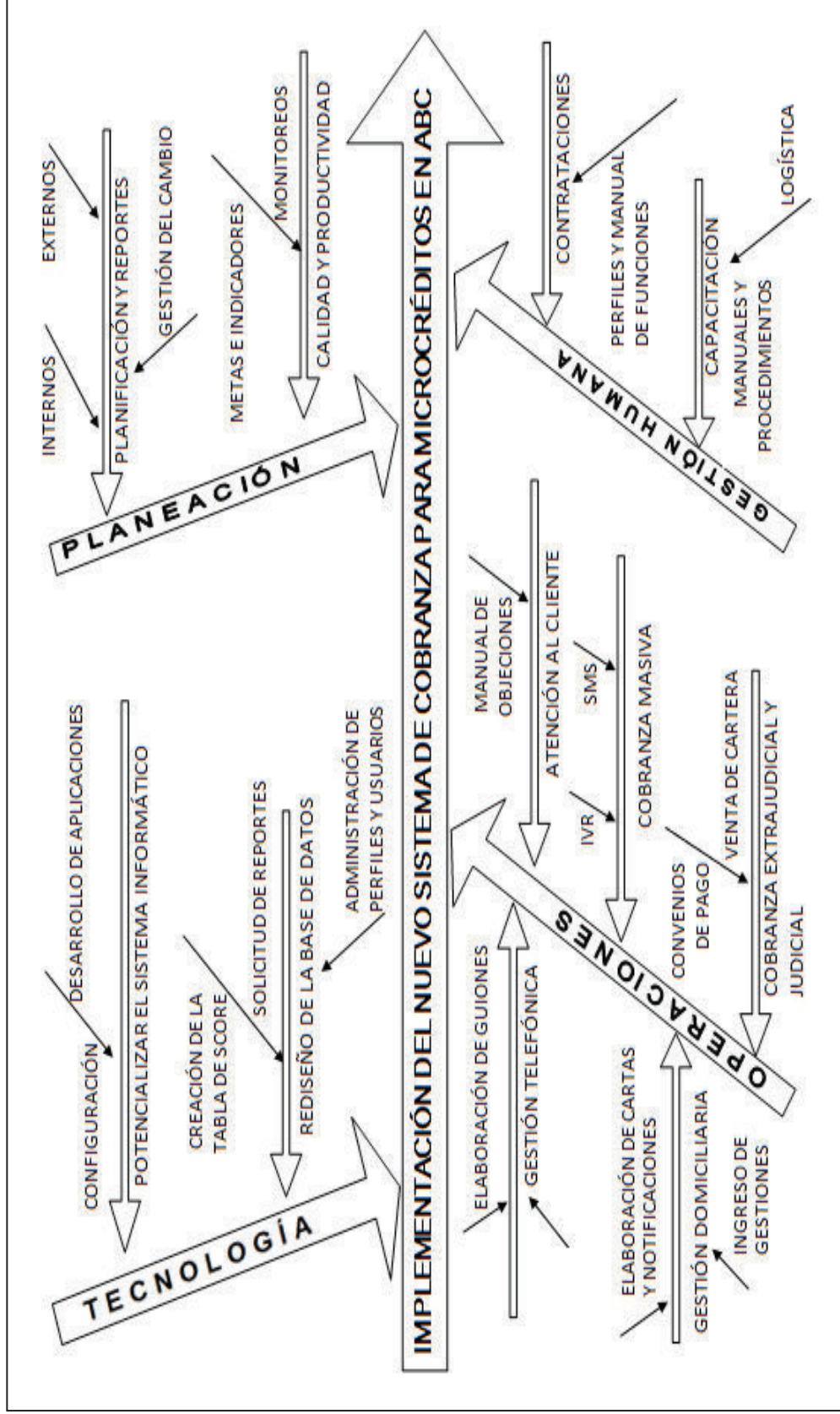
Visto de otra forma, la cobranza como toda actividad táctica requiere de una planeación y organización previa que garantice al máximo el resultado satisfactorio; esto involucra el diseño oportuno de todo un sistema que garantice que las cosas se harán de la mejor manera y por supuesto que estarán a cargo de la gente capacitada, de manera que queden definidos con la debida oportunidad los aspectos que tienen relación con el cómo se organizará el trabajo de cada uno de los participantes al interior de ABC.

Para empezar utilicé un análisis FODA (Figura 4.11) con el objetivo de determinar el desempeño actual y potencial de la organización ante una situación crítica como es la implementación de un nuevo sistema de gestión de cobranza basado en pronósticos y luego para obtener un diagnóstico preciso, construí un diagrama causa efecto (Figura 4.12) que me ayudó a graficar de mejor manera las relaciones entre las actividades que se deben ejecutar para llevar a buen término la implementación del nuevo sistema.

<b>FORTALEZAS</b>	<ul style="list-style-type: none"> <li>• Servicio de cobranza integral, desde la gestión preventiva hasta judicial y terminal.</li> <li>• Atención personalizada y permanente con el cliente.</li> <li>• Cobertura nacional con sucursales en todas las regiones.</li> <li>• Sólidos conocimientos en el área de cobranzas por la presencia de la operación en otros países.</li> <li>• Presencia de Cedentes con importancia en el mercado</li> <li>• Sistemas informáticos propios desarrollados a la medida.</li> </ul>
<b>OPORTUNIDADES</b>	<ul style="list-style-type: none"> <li>• Mercado Creciente para los microcréditos.</li> <li>• Políticas de Estado favorables para los microempresarios.</li> <li>• Poca especialización en la cobranza para el sector microempresarial.</li> <li>• Gestión de cobro judicial y extrajudicial no tecnificado.</li> <li>• Tendencia creciente en el medio para el fortalecimiento de las áreas de inteligencia de negocios y experiencia del consumidor.</li> </ul>
<b>DEBILIDADES</b>	<ul style="list-style-type: none"> <li>• Cultura empresarial orientada a la recuperación de capitales y no en el mantenimiento de clientes</li> <li>• No existen políticas de calidad en el servicio</li> <li>• Limitada tecnología e integración de telecomunicaciones</li> <li>• Deficiente gestión en la captación de nuevos Cedentes</li> <li>• Procedimientos desactualizados o inexistentes</li> </ul>
<b>AMENAZAS</b>	<ul style="list-style-type: none"> <li>• Instituciones financieras indiferentes a contratar el servicio</li> <li>• Nuevos competidores con mayor inversión</li> <li>• Regularizaciones del Estado para los honorarios por cobranza</li> <li>• Normativas estatales al régimen laboral actual.</li> </ul>

**Figura 4.11** - Análisis FODA para la empresa ABC

**DIAGRAMA DE ISHIKAWA (CAUSA - EFECTO) IMPLEMENTACIÓN DEL NUEVO SISTEMA DE COBRANZA PARA LA EMPRESA ABC**



**Figura 4.12** - Diagrama causa efecto para la implementación del nuevo sistema de cobranza en ABC.

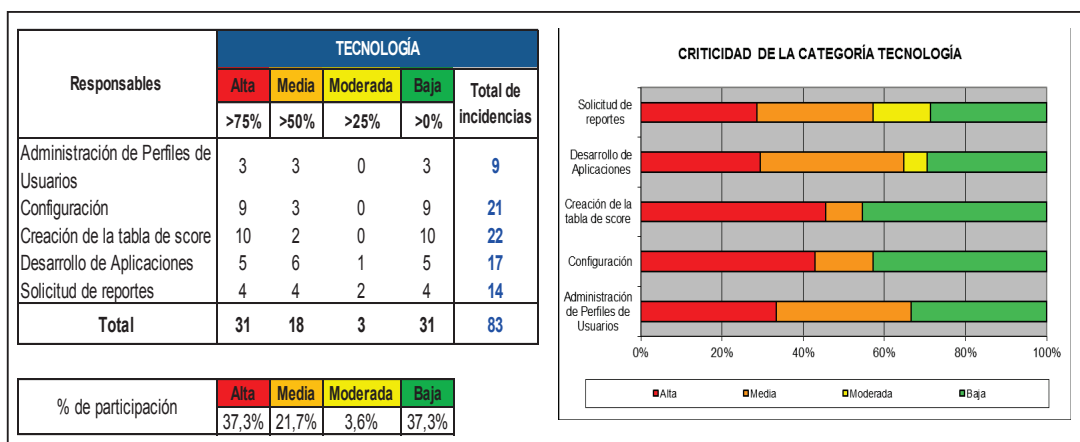
Para obtener la información necesaria involucré nuevamente a los principales responsables de la empresa, primero haciendo entrevistas individuales a cada uno de los expertos y luego analizando en grupo toda la información para hacer una lluvia de ideas. En el equipo me acompañaron las siguientes personas:

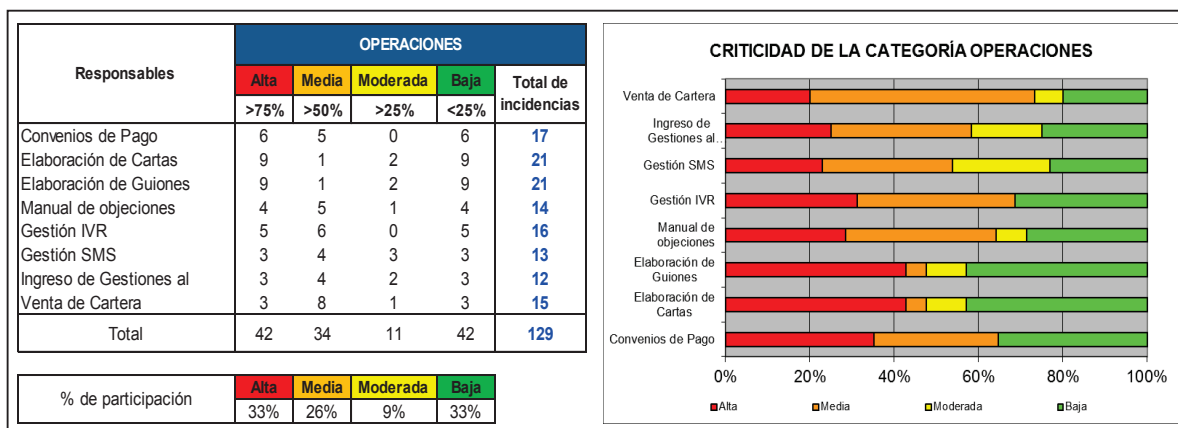
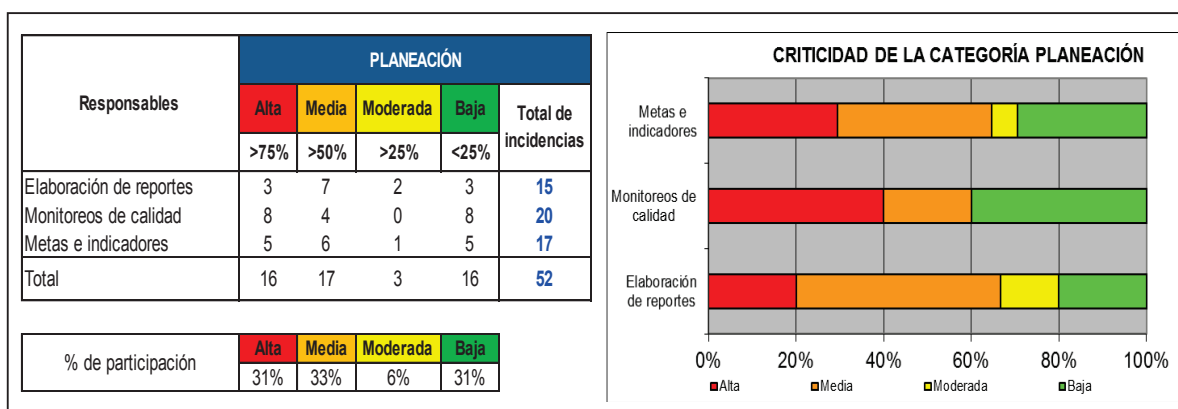
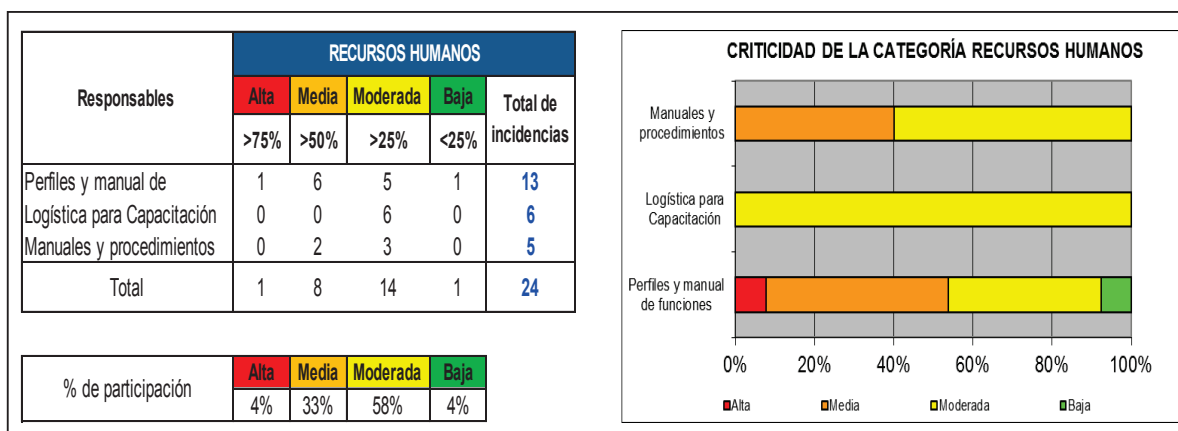
- Gerente de Operaciones
- Gerente de Sistemas
- Gerente Administrativo Financiero
- Gerente de Recursos Humanos

De la Figura 4.11 y 4.12 pudimos identificar 19 causas dentro de cuatro categorías que influyen sobre la correcta implementación del nuevo sistema: uso de la tecnología, buena planeación, correcta operación y gestión de RRHH.

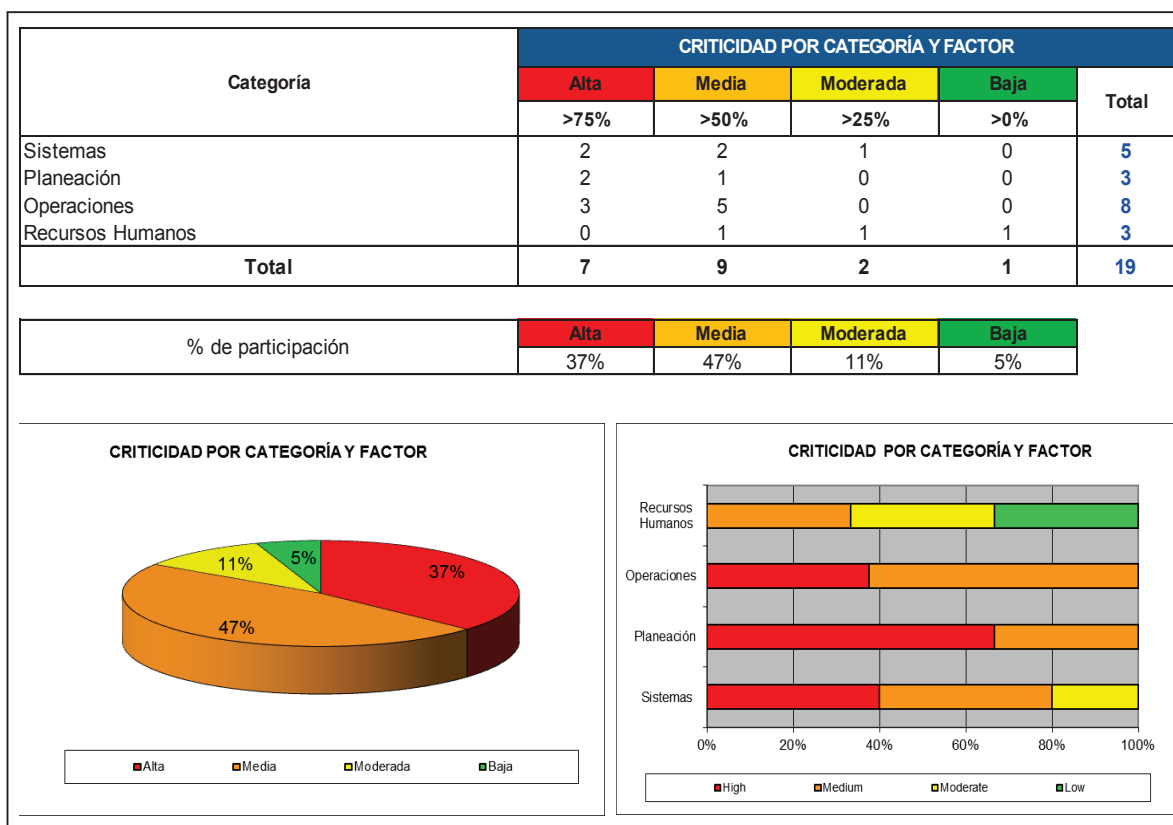
Finalmente, utilizando una matriz determinamos los atributos más críticos en cada una de estas categorías, ponderando la calificación promedio de todos los entrevistados sobre el impacto de cada uno de estos valores en la experiencia del cliente (factor 1 = 40%), la rentabilidad (factor 2 = 30%) y la comunicación entre las áreas (factor 3 = 30%), tal como se muestra a continuación en las siguientes tablas y en el Anexo VIII:

**Tabla 4.1** - Incidencias por criticidad de la categoría Tecnología.



**Tabla 4.2** - Incidencias por criticidad de la categoría Operaciones.**Tabla 4.3** - Incidencias por criticidad de la categoría Planeación.**Tabla 4.4** - Incidencias por criticidad de la categoría Recursos Humanos.

**Tabla 4.5** - Criticidad por categoría y factor.



### 4.3 INVENTARIO DE DOCUMENTACIÓN

La gestión de la información y el conocimiento se relaciona con la comprensión de la situación del sistema de cobranza, tanto antiguo como nuevo, y la organización de los datos disponibles para describirla. El manejo de la información significa seleccionar aquella que se necesita, quién la tiene, cómo la podemos obtener y guardar, y encontrar el método más adecuado para su uso<sup>30</sup>, teniendo en cuenta que ABC debe establecer los mecanismos que faciliten la transformación de la

<sup>30</sup> Project Management Information Systems: Guidelines for Planning, Implementing and Managing a DME Project Information System (2004), CARE International

información en conocimiento y una comunicación adecuada y amplia a todos los involucrados.

Documentar correctamente el plan de implementación del sistema de cobranza es importante, porque permite identificar más fácilmente los aspectos y características que forman parte de este proceso y le proporcionan identidad, de manera que los usuarios y responsables del mismo puedan reconocer las ventajas y desventajas, características y funcionalidades, así como los beneficios que implica el desarrollo del proyecto. A continuación incluyo en la Tabla 4.4 los documentos necesarios para respaldar la implementación del nuevo sistema de cobranza.

**Tabla 4.6** - Inventario de documentación.

CATEGORÍA	NOMBRE DEL DOCUMENTO	TIPO	EXISTE?
Tecnología	Creación de usuarios en el sistema informático	Procedimiento	SI
	Gestión de la demanda	Procedimiento	NO
	Solicitud de reportes	Procedimiento	NO
	Administración de perfiles de usuarios	Procedimiento	NO
Operaciones	Planificación de la gestión de cobranza	Procedimiento	SI
	Manejo de objeciones	Manual	NO
	Ingreso de gestiones al sistema informático	Manual	NO
	Venta de cartera	Política	NO
	Gestión judicial y extrajudicial	Política	SI
	Administración de convenios de pago	Procedimiento	NO
	Exoneraciones de gastos de cobranza	Política	NO
	Atención de quejas y reclamos	Procedimiento	NO
Recursos Humanos	Actas de reunión	Documento	NO
	Definición de perfiles de cargos	Documento	SI
	Funciones de cargos	Documento	SI
	Cálculo de comisiones	Política	NO
	Capacitación de personal	Procedimiento	NO
Planeación	Evaluación del personal	Procedimiento	SI
	Gestión de la información	Procedimiento	NO
	Elaboración de procedimientos y control de cambios	Procedimiento	SI
	Monitoreos de calidad	Procedimiento	NO

De los resultados mostrados concluí que los aspectos más críticos en el área de tecnología, son la creación de la tabla de puntuación con el 50% y la configuración del sistema de cobranza con el 45% de criticidad; en el área de operaciones la elaboraciones de guiones y cartas de notificación con el 45%, seguido por la implementación del esquema de convenios de pago con un 40%; en el área de planeación el proceso de control de calidad con el 40%, seguido por la definición de metas e indicadores con un 30% de criticidad; finalmente el área de recursos humanos con la elaboración de los perfiles y manual de funciones con un 10% de criticidad.

A pesar de ser una empresa robusta y con presencia en el mercado, ABC todavía debe afinar sus procesos para alinearse con las últimas tendencias en materia de cobranza, orientando su oferta sobre la calidad en el servicio a los clientes. La cultura organizacional que ha marcado su desarrollo, si bien ha sido motivo de grandes logros y reconocimiento, también es una piedra de tope a la hora de realizar los cambios requeridos. La implementación del nuevo sistema de cobranza basado en score obliga a realizar una serie de cambios e implementaciones en las diferentes áreas, mismos que evaluados por sus principales responsables en base al impacto sobre los clientes, los costos y la empresa en si, muestran una alta criticidad de todas las áreas (sobre el 30%), excepto en la parte de recursos humanos donde el impacto es menor (4% de alta criticidad), debido a que los procesos de selección, contratación y mantenimiento del personal se basan en las definiciones del área de operaciones.



## CAPÍTULO V

### CONCLUSIONES Y RECOMENDACIONES

El microcrédito en Ecuador continúa sobre la línea de apoyo a las personas que no disponen de acceso a crédito por la banca tradicional; sin embargo su desarrollo ha sido marcado por la política y la influencia de la banca. El Estado debe regular el mercado de los microcréditos a través de políticas económicas que fomenten la producción y el emprendimiento, mejorando el servicio y los productos de la banca estatal y promoviendo una adecuada relación con la banca privada, misma que con su tecnología e infraestructura juega un papel clave en el crecimiento y penetración de este producto financiero. Estas acciones serán beneficiosas siempre y cuando no se pierda la personalización y el trabajo directo y permanente con los microempresarios.

El uso de modelos analíticos para evaluar el riesgo asociado a cada cliente ayuda a bajar los costos de operación y mejorar el servicio, a través de la segmentación de carteras, definición de canales de interacción con los clientes, desarrollo de tecnología, procesos automatizados y estándar, medición y control de la productividad. La cobranza moderna se enmarca en el concepto de servicio, promoviendo la construcción de relaciones sólidas entre los dueños de la cartera y sus clientes, y también del cumplimiento de indicadores relacionados con la experiencia del consumidor<sup>31</sup> siendo al mismo tiempo eficaces en la recuperación oportuna de saldos al menor costo posible.

1. La banca opta por la tercerización total o parcial de sus departamentos de cobranza basada en el desarrollo técnico de empresas especializadas, que como nunca antes, extraen conocimiento de las carteras con un nivel de

---

<sup>31</sup> Algunos indicadores son el medidor de deserción voluntaria o involuntaria (CHURN), venta cruzada (CS), calidad percibida, predisposición a la recomendación (NPS).

profundidad que permite desarrollar sistemas de información muy completos para el diseño de estrategias de recuperación, mismos que se comparten de forma ágil y oportuna a todos los involucrados del proceso.

ABC debe ser muy cuidadosa en las estrategias que generan los ingresos a la compañía, así como en el control del gasto. Desde el punto de vista de los ingresos es ventajoso que tenga un cliente grande para llegar a una especialización y exclusividad; sin embargo por temas de riesgo esta estrategia no es la más adecuada, ya que el impacto financiero de quedarse sin este cliente podría provocar incluso el cierre de la empresa. El área comercial de ABC, además de generar y fortalecer buenas relaciones con el portafolio de clientes y captar otros negocios, debe propiciar la creación e innovación de nuevos servicios para fortalecer la presencia en el mercado.

Un adecuado y estricto manejo de la cantidad de mano de obra directa juega un papel decisivo, no solo con el uso de tecnologías que favorecen el incremento de llamadas o visitas, sino también por un seguimiento de las métricas de productividad por gestor, para cada una de las estrategias definidas dentro de la cartera. La ejecución de estos procesos de control se debe encargar a un área especializada, ya que de no hacerlo se convierte en un factor de riesgo que puede afectar la calidad y hasta la propia estabilidad de la empresa, por tanto hay que evaluar las ventajas de disponer de estos recursos.

2. La metodología estadística que más se utiliza en la construcción de modelos predictivos de cobranzas es la regresión logística clásica. Los resultados que se obtienen de esta manera son buenos, sin embargo se pueden mejorar y contrastar con el uso de otras técnicas, tales como árboles de decisión, regresión logística basada en distancias o métodos de aprendizaje basados en distancias. Con los resultados de los modelos construidos, pude concluir que la mejor metodología es el clasificador K-NN, porque tiene la mayor precisión

en la predicción (91%); a continuación la regresión logística mejorada con árboles de decisión, porque a pesar de tener similares valores de precisión con la regresión logística basada en distancias (84,2% y 88% respectivamente), la primera es mejor en cuanto a malos riesgos lo que repercute en el costo (13,5% y 14,8% respectivamente). Finalmente la regresión logística clásica con la menor probabilidad de bien clasificados y la peor en cuanto a malos riesgos (77% y 40% respectivamente). Como el alcance del trabajo no se centró en la clasificación, sino en entender cómo se comporta la cartera, elegimos como la mejor opción el método de regresión logística mejorada con árboles de decisión, principalmente por la capacidad de explicación.

3. Para validar el uso del sistema de gestión de recuperación en base a los resultados obtenidos del modelo estadístico, construimos con ABC estrategias diferenciadas por tipo de cliente con el fin de optimizar los recursos y dirigir los esfuerzos en la medida que se requiera. El nuevo sistema utiliza menos gestores telefónicos que el sistema de cobranza tradicional de ABC, lo que en términos monetarios representa un 13% de ahorro; también utiliza menos gestores domiciliarios, con un aporte en la disminución de costos del 2,13%. En términos del costo total de la cobranza operativa: gestión telefónica más gestión domiciliaria, el nuevo sistema basado en score es 7,4% más barato que el tradicional, ahorro que se da principalmente por la reducción de gestores en el área de telefonía. Finalmente, es importante indicar que el nuevo sistema de cobranza de ABC requiere no solo saber la naturaleza y comportamiento de pago que tienen los clientes, sino también la estacionalidad, los eventos externos influyentes como un aumento del ingreso por pago de sobresueldos o beneficios, feriados locales y nacionales; así se determina la mejor combinación de acciones motivadoras de pago.
4. El nuevo sistema de cobranza de ABC mejora la experiencia del cliente sin descuidar las metas de recuperación; de los resultados obtenidos se muestra

que desde el día 15 de morosidad, la cobranza debe ser permanente e incrementar de forma gradual su intensidad, utilizando una comunicación adecuada para formar un lazo emocional con el microempresario, misma que debe seguir durante toda la madurez del microcrédito hasta convertirlo en un promotor del producto. Si un cliente no tiene llamadas telefónicas, no tiene por qué deteriorar su comportamiento de pago inmediato, sin embargo el no tener al menos una gestión telefónica en los últimos tres meses incrementa considerablemente el riesgo de convertirse en malo. Se debe reorientar la gestión domiciliaria, ya que las visitas son efectivas siempre y cuando se dirijan hacia los clientes con más riesgo, incluso por el alto costo de las mismas. En cuanto al contacto directo con el titular se concluye que no afecta en el comportamiento de pago, pero sí afecta el no tener información de teléfonos y direcciones vigentes, multiplicando por 5 el riesgo de convertirse en mal cliente; también se muestra que entre los malos clientes existe 12 veces más probabilidad que no tengan un número de teléfono laboral y 6 veces de no tener un número de teléfono particular; finalmente en cuanto a las direcciones, no tener al menos una dirección laboral multiplica por 5 el riesgo de deteriorar el comportamiento de pago. Con todos estos resultados se recomienda que ABC tenga procedimientos robustos de investigación y rescate de datos, para tener una base de teléfonos y direcciones actualizada; se recomienda también que la construcción de estrategias se oriente a contener la morosidad en los primeros días, por lo que intensificar la cobranza preventiva y de mora temprana debe ser la línea de acción que tome ABC, no solo porque los clientes incrementan el riesgo conforme aumentan sus días de atraso, sino porque en una cartera sana se evita una gestión desgastante en los rangos de mora elevado.

5. Uno de los requisitos del nuevo sistema de cobranza de ABC es el mayor conocimiento del cliente, que va más allá de la obtención de una fórmula que discrimine los clientes más riesgosos para la construcción de estrategias; involucra un profundo conocimiento del negocio, los acuerdos comerciales que

se hayan negociado previamente, las competencias y habilidades para gestionar una cartera, la infraestructura necesaria, nuevas tecnologías, la coyuntura del crédito en el país, las regulaciones y restricciones estatales y las mejores prácticas en el medio. En el caso de ABC por ejemplo, las políticas de cobranza que utiliza son muy rígidas sobre los clientes impagos con más de 90 días de atraso (6% del total de la cartera), ejecutando estrategias agresivas como la gestión judicial, a pesar que tengan una probabilidad muy baja de recuperación. Una estrategia con corte en el valor 997 implicaría que el 30,29% del total de los clientes sean retirados de gestión para ser tramitados por el área legal, pero si se considera solo aquellos que tengan mora mayor a 90 días, se identifica apenas el 3% del total de la cartera, de los cuales el 55% son malos clientes. Con este conocimiento ABC puede mantener en gestión prejudicial a un 3% del total de la cartera, o lo que es lo mismo una reducción del 50% de los casos que maneja el área legal actualmente, con el ahorro correspondiente por los altos costos de la gestión judicial. De manera adicional ABC puede identificar clientes que necesitan una gestión menos intensa, y que sí la tienen con el esquema tradicional; por ejemplo 20 de cada 100 clientes ya no recibirán una cobranza desgastante y agresiva, sino se gestionarán de manera comercial, propiciando la fidelización; 30 de cada 100 clientes que en el esquema tradicional se visitan en el domicilio para presionar el cobro, con el nuevo sistema se gestionan solo telefónicamente, ya que muestran una tendencia a mejorar su comportamiento de pago en el tiempo.

6. En el esquema tradicional de ABC, los reportes para seguimiento y control se basan en el saldo recuperado respecto al saldo total en cobranza. Este indicador no considera la razón del negocio de manera completa, porque incluso están en contra de la misión y los valores empresariales de ABC, que citan en muchas partes la calidad en el servicio y la construcción de relaciones sólidas entre las empresas Cedentes de cartera y sus clientes. El nuevo sistema basado en score necesita mayor acompañamiento en virtud de las múltiples ventajas que representa para la operación, ABC debe designar la

persona o equipo responsable del mantenimiento del score, pero sobre todo del control de calidad en la información que se necesita para generar los resultados. El cálculo de indicadores debe llegar a niveles del ejecutivo de cobranza, ya que es muy importante medir los resultados de cada estrategia en cuanto a recaudaciones, clientes pagados, cuotas y provisiones, con el más alto nivel de detalle para identificar y corregir las desviaciones.

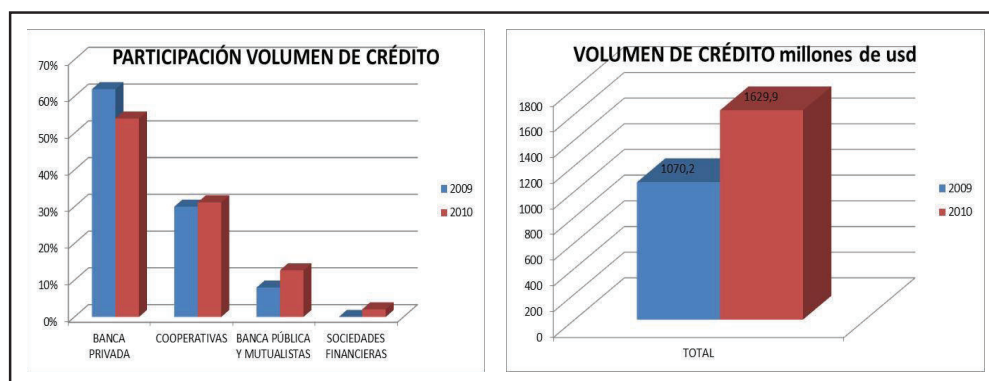
7. La estructura requerida para la implementación del nuevo sistema de gestión, requiere primero un cambio en la cultura organizacional de ABC, para pasar del esquema de volumen de gestiones y recuperación de saldos, al de segmentación y experiencia del cliente. Todas las áreas de ABC deben redefinir sus procesos, y este cambio no está exento de costos y riesgos que la dirección debe asumir y defender con miras al futuro y estabilidad de la empresa. La incorporación de personal capacitado en el análisis y construcción de modelos estadísticos podría ser muy recomendado para que ABC obtenga todos los beneficios del modelo analítico, no solo para el mantenimiento sino también para la presentación de resultados y diseños de los planes de mejora; en este sentido el apoyo del área de tecnología es muy importante para proveer de los sistemas informáticos necesarios, tanto en la administración de la información como también en el control de la productividad. El área de recursos humanos debe trabajar en actividades que promuevan los valores de este cambio, coordinando con la parte de operaciones la mejor estrategia para que la gestión diaria no se vea afectada. Finalmente el área financiera debe realizar un seguimiento de cerca a todos los hitos de la implementación del nuevo sistema de cobranza, para retroalimentar acerca del comportamiento de los ingresos y gastos, que pueden obligar a realizar correcciones o ratificar las acciones emprendidas, siempre con miras de hacer más rentable y eficiente el negocio de ABC, con alternativas innovadoras y técnicas, mejorando las condiciones de sus colaboradores.

## ANEXO I

### INDICADORES DE LA MICROEMPRESA EN ECUADOR

#### CARTERA BRUTA DE MICROEMPRESA

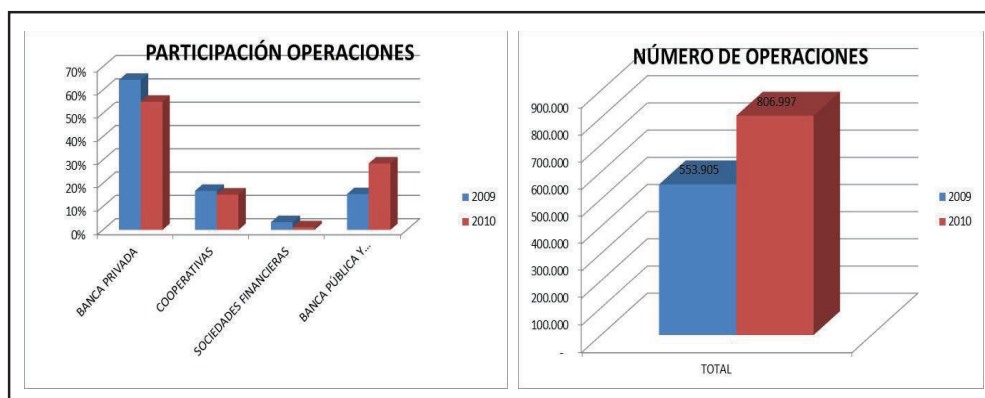
La cartera bruta de microempresa (Figura A1.1) a septiembre de 2010 fue de US\$2.025 millones, (3% del PIB), que comparado con el mismo período de 2009 (US\$1.560 millones), representa un crecimiento del 30%, y equivale al 15% de la cartera total del sistema financiero (US\$13.541 millones a septiembre de 2010). La banca privada otorgó US\$915 millones, equivalente al 45% del total de la cartera bruta de microempresa; las Cooperativas aportaron con US\$719 millones (36%); la Banca Pública otorgó US\$348 millones (17%).



**Figura A1.1 - Participación y volumen de Crédito Microempresa**

#### NÚMERO DE OPERACIONES CREDITICIAS DESTINADAS A LA MICROEMPRESA

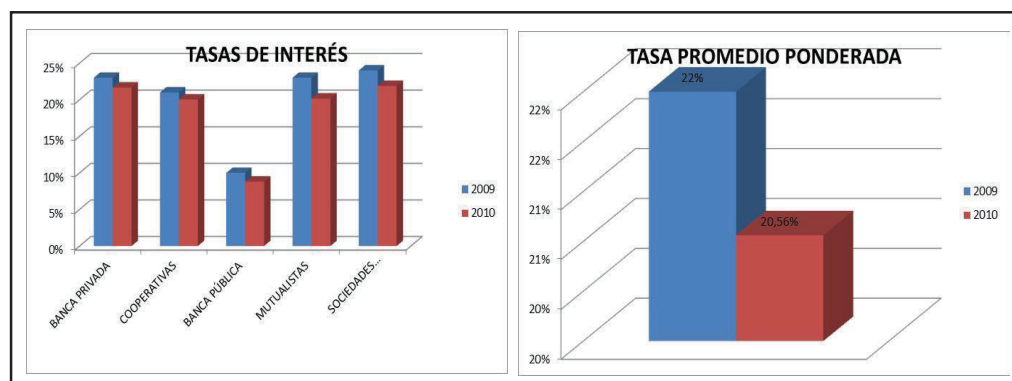
Conforme al volumen de crédito (Figura A1.2), el número total de operaciones registradas en el sistema micro financiero que mantienen saldo a septiembre de 2010, fue de 806.997 operaciones, mientras que a septiembre de 2009 fueron de 553.905, lo que representa un crecimiento del 45,7%.



**Figura A1.2** - Participación y número de operaciones Microempresa

## TASAS DE INTERÉS

Las tasas de interés (Figura A1.3) para el sector micro financiero conllevan “costos adicionales”, ya que implican mayor grado de especialización, atención personalizada en el campo y contar con tecnología específica para microfinanzas, razones por las cuales se diferencian de las tasas de interés para otros tipos de financiamiento como son: comercial, consumo y vivienda; sin embargo, la política gubernamental atada a una necesidad de brindar mayor acceso al financiamiento, vía transparencia y sensibilidad del sector privado, ha logrado ajustar las tasas de interés.

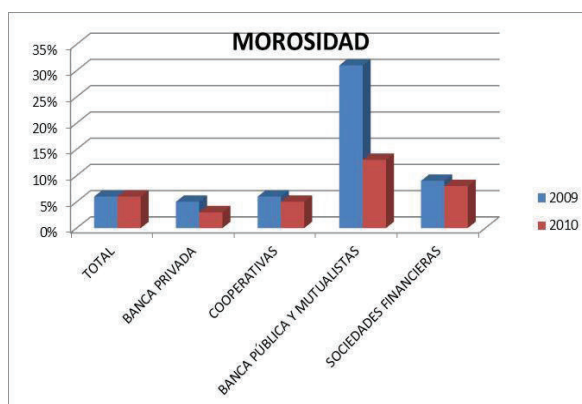


**Figura A1.3** - Tasa de interés sector microfinanciero



## CARTERA VENCIDA SEGÚN TIPO DE CRÉDITO

En lo que respecta al nivel de morosidad (Figura A1.4) de la industria micro financiera, los actores y operadores, desde su origen han establecido de cierta forma una estrategia para precautelar y cuidar el desarrollo del mercado microfinanciero, elementos como la especialización y la tecnología crediticia adquirida por los instituciones que hacen microfinanzas, el apoyo de organismos internacionales, la capacitación y modernización en la entidad de control, han contribuido a mantener un nivel de morosidad adecuado, inferior al 6%, pero precisa recalcar que el hecho fundamental que contribuye a mantener dichos niveles son las características propias del cliente micro financiero y su imperiosa necesidad de mantener viva la posibilidad de crecer y desarrollar su emprendimiento, a pesar de aquello, las cifras ya empiezan a ser preocupantes.

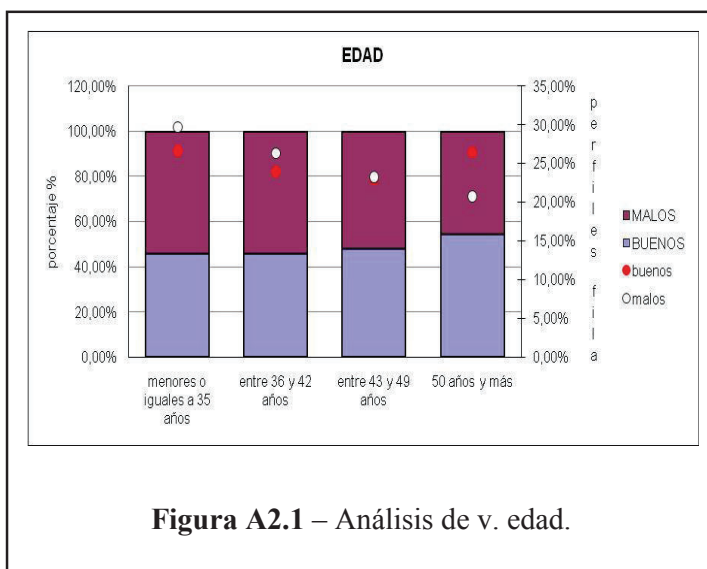


**Figura A1.4** - Morosidad del sector micro financiero

## ANEXO II

## ANÁLISIS UNIVARIADO DE DATOS EMPRESA ABC

## Variable Edad



Los clientes se distribuyen de forma similar en todas las categorías, aunque aquella menor o igual a 35 años destaca ligeramente y la 50 años o más es la que contiene el menor número de clientes.

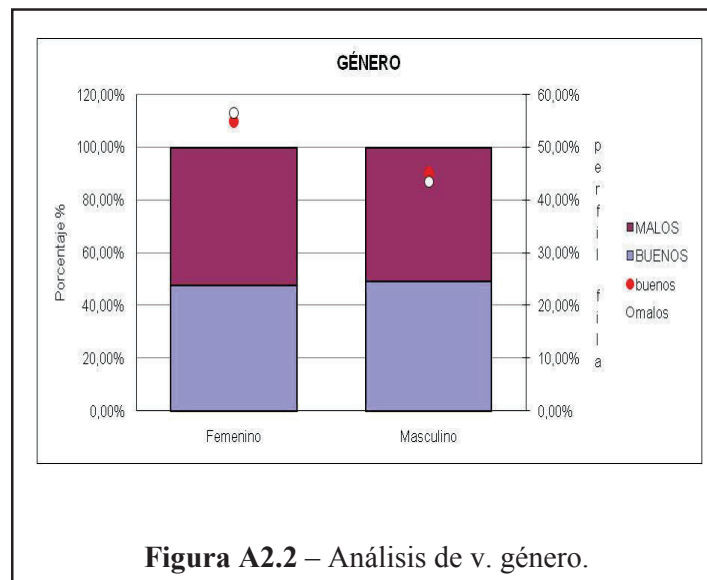
Dentro de la categoría menor a 35 años prevalecen los clientes

con mal comportamiento de pago mientras que en aquella mayor a 50 años predominan los clientes buenos. Se puede concluir de este análisis que los clientes tienden a ser mejores pagadores conforme tienen más edad. Todas las categorías son similares por cuanto no existen diferencias significativas entre los clientes buenos y malos.

## Variable Género

Predomina ligeramente la categoría de género femenino en la muestra, donde existe una mayor cantidad de clientes con mal comportamiento de pago. En la categoría género masculino existe una leve prevalencia de clientes con buen comportamiento de pago sin embargo no es suficiente como para asegurar que esta variable aporte con suficiente poder discriminatorio al modelo. Este hecho se

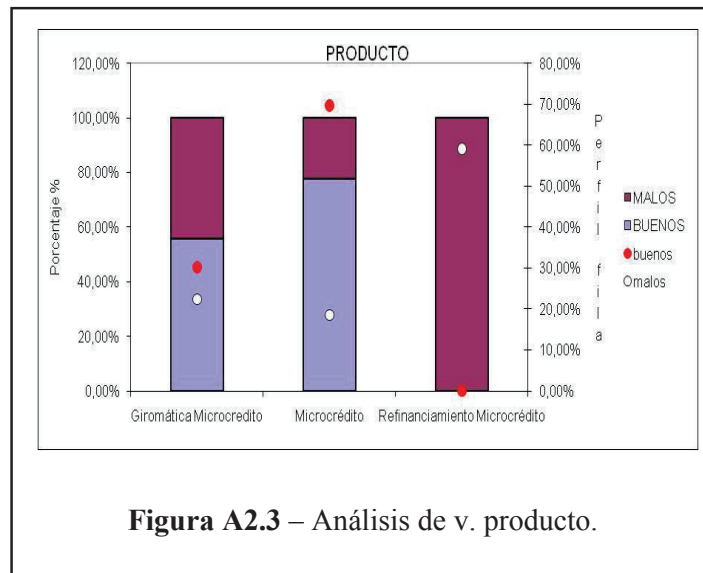
ratifica con el valor de significación de estadístico Chi Cuadrado, 0.392 que es bastante mayor al nivel de significación bilateral de 0,05 y por lo tanto se acepta la independencia de la variable género y tipo de cliente.



### Variable Producto

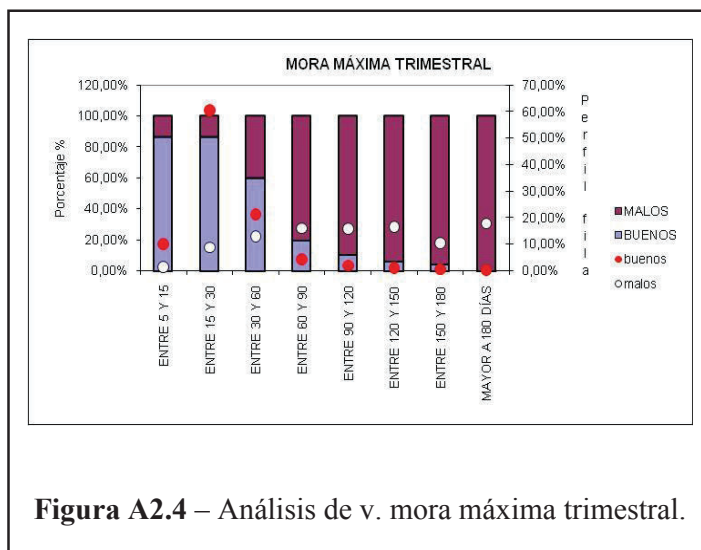
Existen más clientes con producto microcrédito normal que aquellos con giromática o refinanciamiento.

En los clientes con giromática y normales, especialmente en esta última, predominan los clientes con buen comportamiento de pago y en la categoría refinanciamiento predominan los clientes malos.



**Figura A2.3 – Análisis de v. producto.**

**Variable Mora Máxima Trimestral**



**Figura A2.4 – Análisis de v. mora máxima trimestral.**

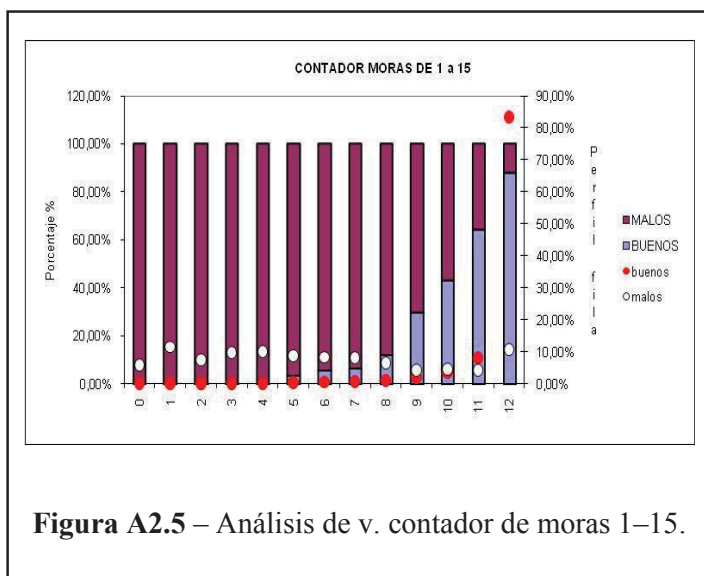
Se observa un comportamiento similar de esta variable respecto de la mora máxima anual medida desde el último pago. En las categorías con mora menor a 60 días predominan los clientes buenos mientras que en el resto de categorías predominan los clientes

con mal comportamiento de pago siendo notoria la diferencia respecto de los clientes buenos en todas las categorías mayores a 60 días de mora. Las categorías de 0 a 15 y de 15 a 30 días se comportan de manera similar en cuanto a la distribución de clientes buenos pero para el resto de categorías como es natural la proporción de clientes buenos disminuye conforme aumenta la mora; así por ejemplo la categoría de 30 a 60 días tiene 40,24% más de clientes buenos

que la siguiente categoría de 60 a 90 y 49,50% respecto de la categoría de 90 a 120 días de mora. Se observa que del 15% de clientes del rango de mora menor a 6 días que deterioraron su comportamiento de pago, un 53% continuará este deterioro; de aquí un 94% terminará siendo mal cliente lo que implica que la gestión de recuperación debe centralizar sus esfuerzos en contener la mora temprana hasta los 60 días porque a partir de este valor la cartera se hace completamente mala.

Es notorio el hecho que los clientes deterioran su comportamiento de pago drásticamente cuando superan el umbral de los 5 días de mora, existiendo 38% más de clientes malos en el rango de 6 a 30 días que en el rango de mora anterior. Los clientes que alguna vez tuvieron mora entre 6 y 30 días pero que actualmente tienen una mora inferior o están al día tienen tendencia a tener mejor comportamiento de pago que aquellos que empeoraron su calificación de mora. Del estadístico Chi Cuadrado se desprende la no independencia entre estas variables.

#### Variable Contador de Moras de 1 a 15 días



La categoría más poblada es aquella que contiene los clientes con 12 ocasiones en mora de 1 a 15 días y las categorías con 0 hasta 8 ocasiones son las menos pobladas. En las categorías con 0 ocasiones hasta 10 ocasiones en mora menor a 15 días predominan los clientes malos mientras que en

aquellas con 11 o 12 ocasiones predominan los clientes con buen comportamiento de pago. La categoría de 12 ocasiones tiene 23,54% más de clientes buenos que la categoría con 11 ocasiones, 44,65% más que la categoría con 10 ocasiones y 57,88% más que la categoría con 9 ocasiones.

### Variable Desembolso

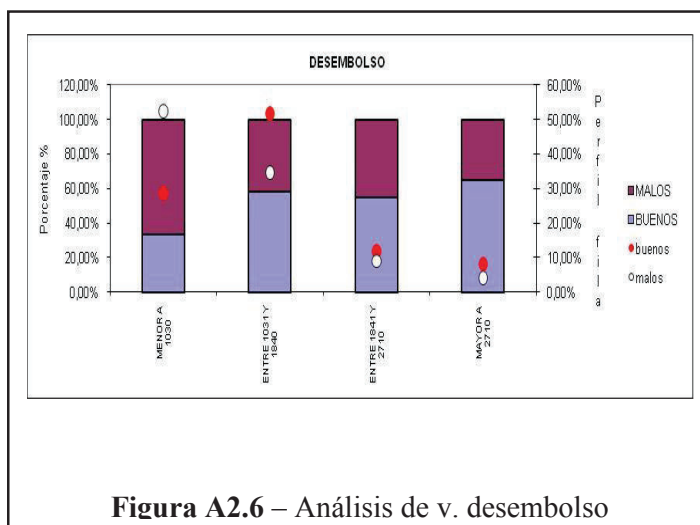


Figura A2.6 – Análisis de v. desembolso

Los clientes con desembolsos menores a 1030 usd son en general malos clientes. Las categorías de 1031 usd en adelante presentan mejor comportamiento de pago sobre todo aquellos con desembolsos mayores a 2710 usd que tienen un alto porcentaje de buenos

clientes. Los porcentajes de referencia más representativos pertenecen a los clientes con desembolsos menores a 1030 usd y mayor a 2710.

La categoría más poblada corresponde al rango de 1031 a 1840 usd de valor desembolsado y la menos poblada es aquella con montos mayores a 2710. En todas las categorías predominan los clientes con buen comportamiento de pago pero donde se aprecia con mayor claridad es en la categoría con montos entre 1031 y 1840 usd. Las tres categorías se comportan de forma similar aunque se puede apreciar que los clientes con mayores valores desembolsados tienden a ser mejores clientes.

## Variable Saldo

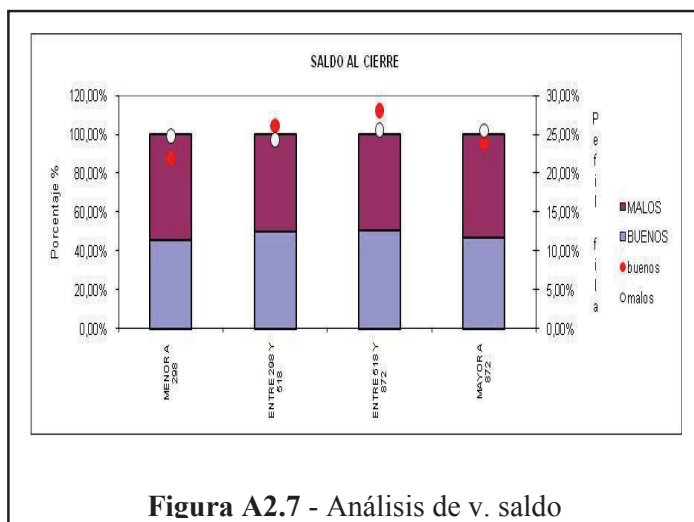


Figura A2.7 - Análisis de v. saldo

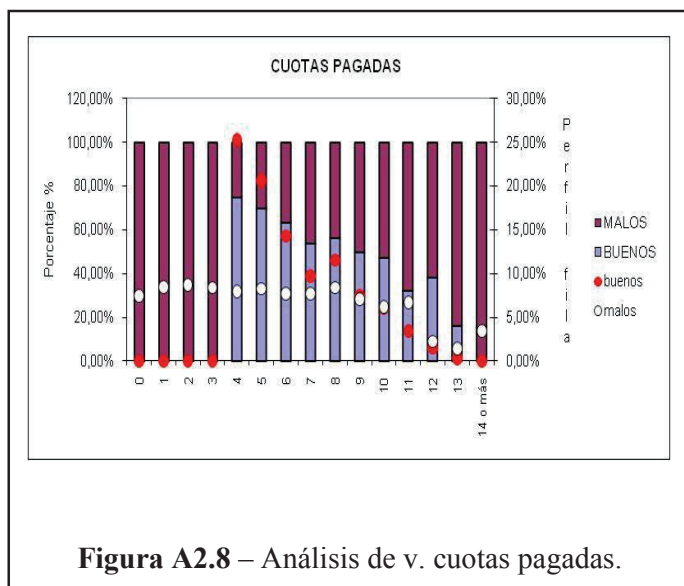
Los clientes con saldos al cierre entre 519 y 872 usd son los que predominan. En esta categoría y en aquella con saldos al cierre entre 299 y 510 predominan los clientes con buen comportamiento de pago mientras que en la categoría con saldos al cierre mayores a 872 usd

predominan los clientes malos. Las tres categorías son similares respecto a la distribución de clientes buenos y malos lo que sugiere que esta variable tiene poco poder discriminatorio y no entre en el modelo.

## Variable Porcentaje de Cuotas por Pagar

La categoría más representativa es aquella con los clientes que les falta entre el 52% y 67% de las cuotas de su crédito. La menos poblada corresponde aquella con los clientes que tienen pendiente menos del 31% de las cuotas de su crédito. Las categorías con menos del 52% de cuotas pendientes por pagar de su crédito se comportan de forma similar respecto de la distribución de buenos y malos clientes. En la categoría con más del 67% de cuotas por pagar no existe mucha participación de clientes con buen comportamiento, lo que indica que los clientes malos tienen problema desde un inicio con el pago de sus cuotas.

## Variable Cuotas Pagadas



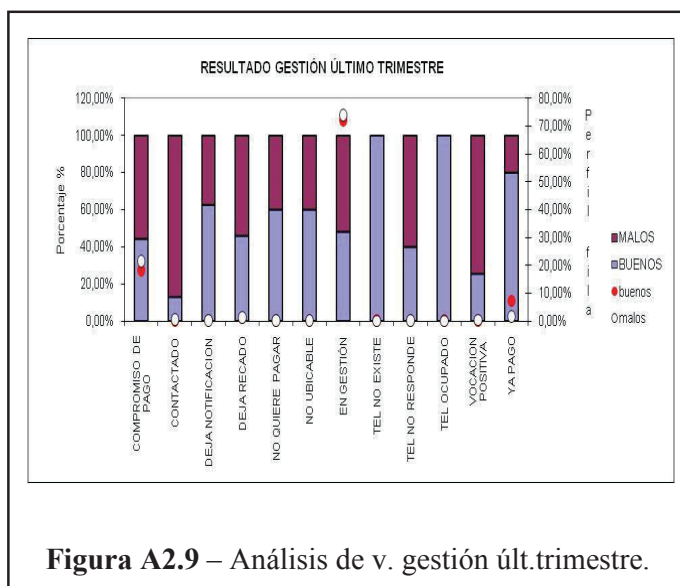
Los clientes que predominan son aquellos que han pagado 4 cuotas seguidos de cerca por los que han pagado 5 cuotas. Las categorías menos pobladas son aquellas con 12 o más cuotas pagadas y con menos de 4 cuotas pagadas. Los clientes que han pagado entre 4 y 5 cuotas se comportan de forma similar

en cuanto a la distribución de clientes buenos y malos, de la misma forma aquellos que han pagado entre 6 y 10 cuotas y entre 11 y 13 cuotas. Sin embargo de estas similitudes existe la tendencia que conforme aumenta el número de cuotas pagadas disminuye el porcentaje de clientes buenos. Así la categoría con 4 cuotas pagadas tiene un 4,72% más de clientes buenos que aquella de 5 cuotas pagadas y 11,42% respecto de la categoría de 6 cuotas pagadas.

Este análisis es importante a la hora de sugerir cambios a la política de crédito del banco puesto que el comportamiento de pago de los clientes cambia conforme el número de cuotas que vaya pagando siendo notorio que aquellos que aquellos que han pagado hasta 3 cuotas se consideran malos posiblemente porque pagan las cuotas con dinero del propio préstamo y por otro lado el hecho que el porcentaje de clientes buenos disminuya conforme aumenta el número de cuotas pagadas puede significar cansancio de cliente por el pago de muchas cuotas o porque él piense que ya ha pagado mucho. Estos resultados también arrojan conclusiones importantes a la hora de establecer políticas de conservación de clientes.



## Variable Respuesta de la Gestión Último Trimestre



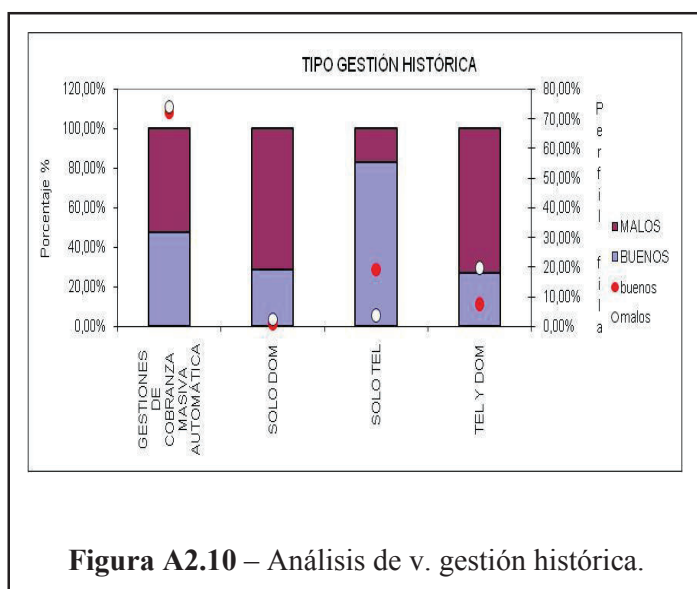
En el último trimestre se observa que predominan los clientes sin gestión personalizada, vía telefónica o domiciliaria. Le siguen los clientes con compromiso de pago. En ambas categorías predominan los clientes con mal comportamiento de pago. De forma similar al comportamiento mensual,

aquellas categorías que representan algún contacto, ya sea directo o indirecto con el cliente, presentan un mayor porcentaje de clientes buenos; por ejemplo existe 27,5% más de clientes buenos en la categoría contactados que en aquellos donde no se ha podido ubicar al cliente telefónicamente en los últimos tres meses.

En vista que los resultados de la gestión en el mes anterior y último trimestre son bastante similares, y existen categorías muy poco pobladas, se decidió agrupar estas gestiones en categorías que reúnan características similares y dejar solo los resultados del mes anterior. Por ejemplo se unió las categorías "Teléfono no existe", "Teléfono ocupado" y "Teléfono no responde" en una sola categoría "No contactados"; las categorías "Ya pagó" y "Contactados" se unieron en una sola categoría "Contactados"; finalmente las categorías "Vocación Positiva de Pago", "Deja notificación" y "Deja recado" se unieron en una sola categoría "En seguimiento"; el resto se dejó de la forma original. La categoría de gestión masiva de cobranza tiene un 4% más de clientes buenos que los compromisos de pago, la mayor parte incumplidos. El estadístico chi cuadrado tiene un nivel de significación menor a 0,05 pero existe un 66,7% de celdas con valor esperado

menor a 5 por lo que no se puede concluir la dependencia entre la variable tipo de cliente y gestión último trimestre.

### Variable Tipo de Gestión Histórica



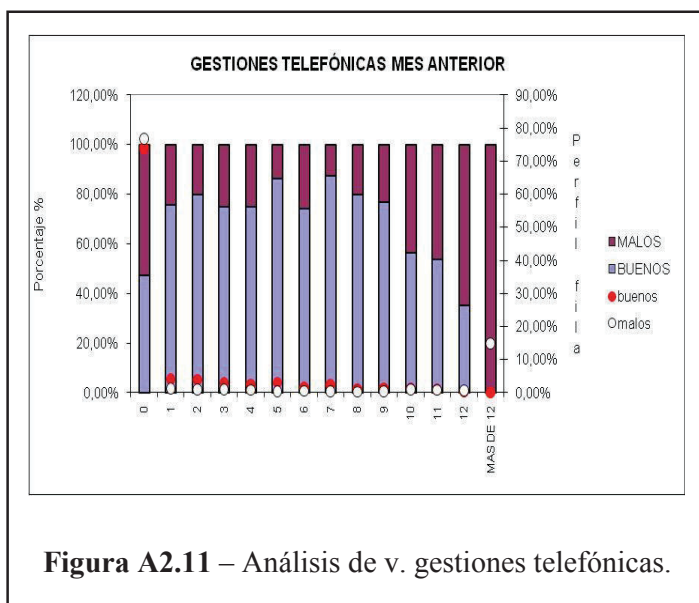
La categoría predominante es aquella que contiene los clientes a los que se les ha hecho gestión de cobranza masiva automática y la menos poblada corresponde a los clientes a los que solo se les ha hecho visitas domiciliarias. En todas las categorías, excepto en aquella de los clientes con gestión

telefónica, predominan los clientes con mal comportamiento de pago. La categoría solo gestión telefónica tiene 35,26% más de clientes buenos que la categoría gestión masiva de cobranza, 56,22% más que la categoría gestión telefónica y domiciliaria y 54,57% más que la categoría solo gestión domiciliaria. Estos resultados sugieren la importancia y la efectividad de la gestión telefónica en el recupero de cartera y ponen en duda la efectividad de la gestión doble y gestión domiciliaria. El establecimiento de la mejor estrategia de recupero, ya sea vía telefónica, vía visita domiciliaria o las dos se analizarán desde un punto de vista financiero más adelante. La categoría solo gestión telefónica tiene un 55% más de clientes buenos que la categoría Solo Gestión domiciliaria y 56% más de clientes buenos que la categoría Gestión Telefónica y Domiciliaria. Respecto de la categoría Gestión Masiva, esta tiene un 35% menos de clientes buenos que aquella Gestión Telefónica. El estadístico chi cuadrado muestra la no independencia entre estas dos variables.

Del análisis de esta variable se concluye la importancia que tiene la gestión telefónica en la recuperación de cartera y por consiguiente el adecuado manejo de la información del cliente acerca de sus direcciones y teléfonos principalmente.

### Variable Número de Gestiones Telefónicas en el Mes Anterior

La categoría más poblada corresponde aquellos clientes con ninguna gestión telefónica en el mes anterior. Aquí predominan levemente los clientes con mal comportamiento de pago. En el resto de categorías predomina escasamente el



**Figura A2.11** – Análisis de v. gestiones telefónicas.

buen comportamiento de pago, aunque contienen muy pocos individuos dan una idea que mejora el comportamiento de pago conforme se realiza al menos una gestión telefónica. El segmento que agrupa los clientes con más de 12 gestiones telefónicas en el mes anterior agrupa una cantidad importante de

clientes que en su mayoría tienen mal comportamiento de pago. Si analizamos el comportamiento de las categorías observamos un comportamiento similar entre aquellos que tienen 0 gestiones telefónicas y más de 10, en todas estas categorías existen más clientes malos que buenos. Por el contrario en aquellos clientes ubicados en las categorías con al menos 1 gestión telefónica hasta 9 gestiones, predominan los clientes con buen comportamiento de pago. Esta situación es particularmente importante a la hora de estructurar el nuevo modelo de gestión, para indicar el número adecuado de gestiones telefónicas a un grupo de clientes en particular.

### **Variable Número de Gestiones Domiciliarias Último Trimestre**

Lo que más hay son clientes con 0 gestiones domiciliarias; el resto de categorías son muy poco pobladas. En los clientes con 0 gestiones predominan los buenos mientras que en el resto de categorías predominan los clientes con mal comportamiento de pago. Las categorías con 0 y de 1 a 4 gestiones domiciliarias tienen comportamiento parecidos; sin embargo si analizamos los clientes con 0 gestiones y aquellos con 5 a 8 gestiones estos últimos tienen 43,8% más de clientes malos. Algo similar ocurre si comparamos la categoría 0 gestiones con las de 9 a 12 gestiones y más de 12 gestiones domiciliarias en el último trimestre.

### **Variable Vector de Moras**

El Vector de Moras se basa en el comportamiento de pago en los últimos 3 meses sin incluir el mes actual. Aquí A=mora de 0 a 5 días; B=mora de 6 a 30 días; C=mora de 31 a 60 días; D=mora de 61 a 90 días; E=mora mayor a 90 días.

Se puede observar que los mejores clientes son aquellos que se encuentran al día en el último mes y con una sola mora menor a 30 días hace dos y tres meses; es decir clientes con vectores de mora A/A/A, B/A/A, A/B/A. Los peores clientes son aquellos que en último trimestres presentan al menos 1 mora mayor a 30 días. Se va a simplificar estos resultados agrupando los rangos más representativos en cuanto a número de clientes se refiere y que tengan porcentajes de referencia considerables.

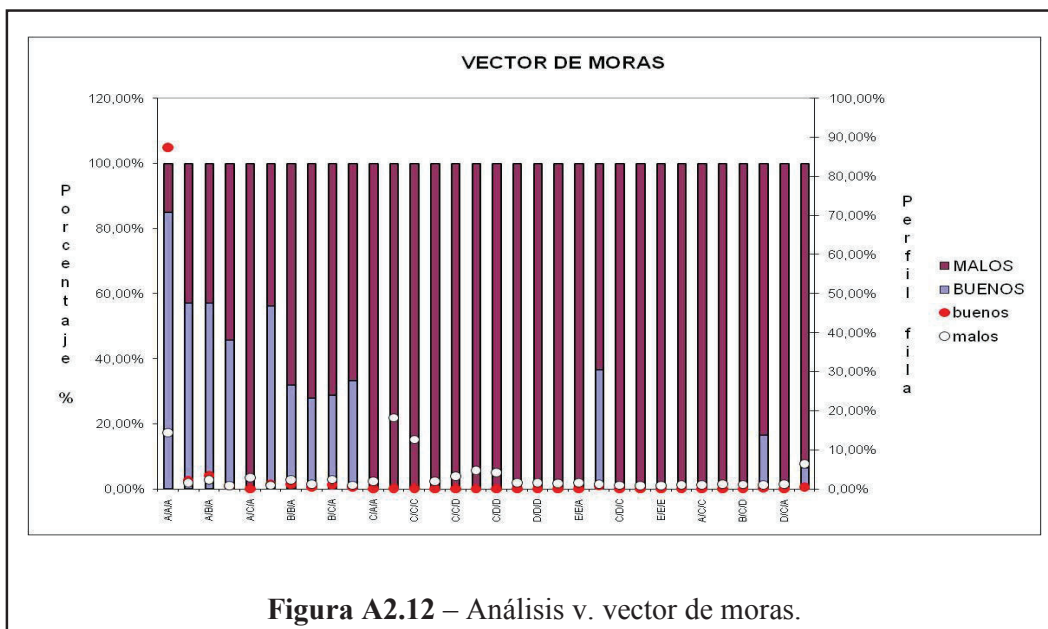
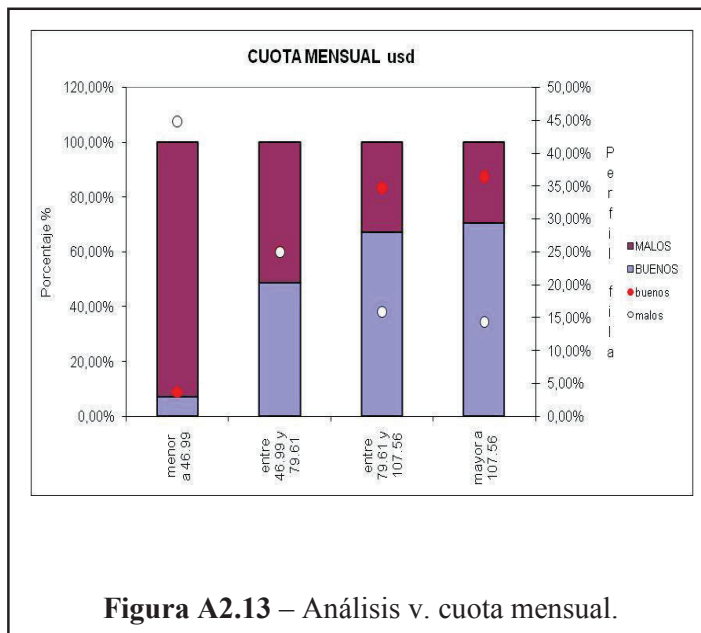


Figura A2.12 – Análisis v. vector de moras.

Lo que más hay son clientes al día en los últimos tres meses A/A/A; el resto de categorías son poco pobladas a excepción de aquellas con vectores de mora CCC y CCA posiblemente porque agrupan un segmento especial. En los clientes con vector de moras AAA predominan aquellos con buen comportamiento de pago, en las categorías AAB y BBA predominan muy levemente este comportamiento de los buenos; en el resto de categorías predominan los clientes con mal comportamiento de pago. Las categorías AAB y ABA se comportan de forma muy similar, de igual manera las categorías BBA, BBB, BCA, BCC, ABB con un porcentaje similar de clientes buenos.

La categoría AAA tiene 52,87% más de clientes buenos que la categoría BBA y 27,81% más de clientes buenos que la categoría AAB. En general aquellas categorías con más de 2 veces consecutivas con moras mayores a C tienen muy mal comportamiento de pago.

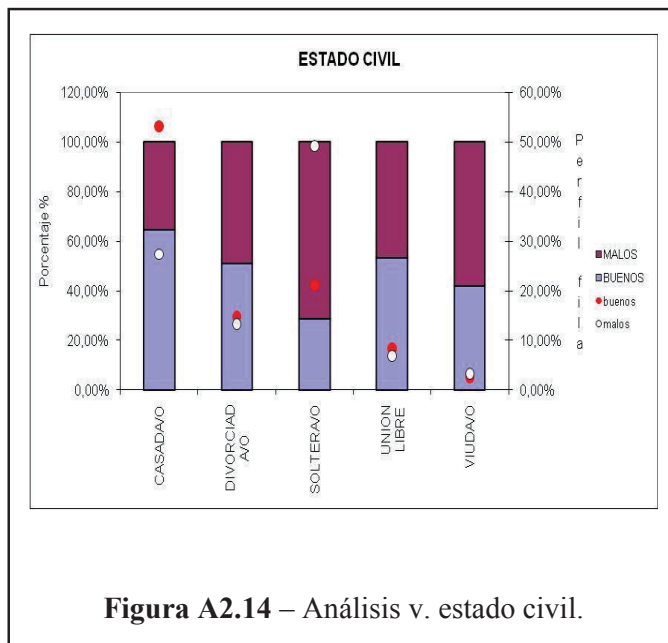
## Variable Cuota Mensual



Existen más clientes que tienen que pagar cuotas mensuales superiores a 79 usd. De estos, la distribución de clientes en cada categoría: "entre 79 usd y 107 usd" y "mayor a 107 usd", es muy parecida y también se comportan de forma similar en cuanto a la distribución de clientes buenos y malos en cada una de ellas. Aquí los

buenos pagadores son mayoría. Por otra parte en los grupos de cuotas inferiores a 79 usd predominan los clientes con mal comportamiento de pago, siendo más visible este comportamiento en los clientes con cuotas menores a 46 usd donde la mayoría son malos pagadores. En general existe una tendencia a mejorar el comportamiento de pago conforme aumenta el valor de la cuota mensual; así por ejemplo la categoría "mayor a 107 usd" tiene 63,19% más de clientes buenos que la de clientes con cuotas menores a 46 usd y 21,87% más que la de cuotas entre 46 usd y 79 usd.

## Variable Estado Civil

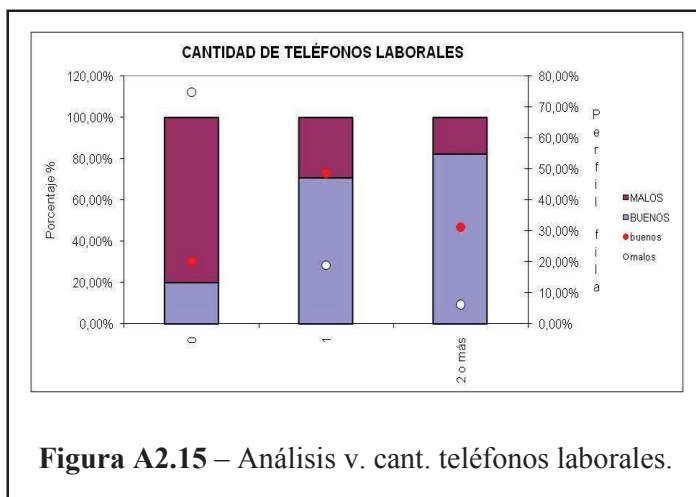


Predominan los clientes solteros y casados, el resto de categorías son poco pobladas. En la categoría Casado predominan claramente los buenos pagadores mientras que los solteros tienen a tener muy mal comportamiento de pago. En el resto de categorías predominan muy levemente los clientes buenos. La categoría divorciado presenta

un porcentaje similar de buenos y malos clientes; los clientes en Unión Libre mejoran levemente su comportamiento de pago respecto a los solteros y los viudos son en su mayoría malos pagadores. Analizando los grupos de casados y solteros, los primeros tienen 35,8% más de clientes buenos. Se observa que las categorías de soltero, divorciado y viudo se comportan cada una de forma diferente y por tanto no conviene unir categorías.

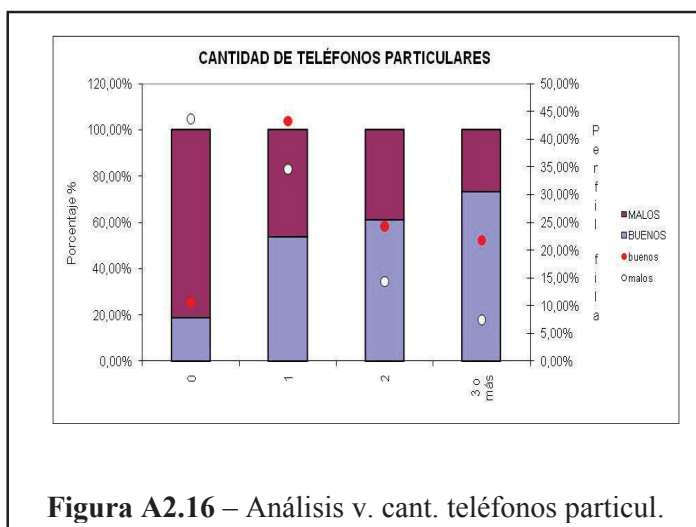
Las categorías casado y soltero se comportan de forma diferente, los primeros tienen un 35% más de clientes buenos que los segundos. Comparando la categoría casado y divorciado, la diferencia entre los porcentajes de clientes buenos disminuye. Los casados tienen 13% más de clientes buenos que los divorciados y estos a su vez un 22% más de clientes buenos que los solteros. Del estadístico Chi Cuadrado se concluye la no independencia entre estas variables.

### Variable Cantidad de Teléfonos Laborales Vigentes



La categoría más poblada es la de aquellos clientes sin teléfono laboral vigente, es decir, no se puede obtener un contacto con el titular cuando se gestiona al número laboral registrado. La mayor parte de clientes buenos tienen al menos un teléfono vigente mientras que aquellos con mal comportamiento de pago no tienen números telefónicos para comunicarse con ellos al sitio de trabajo.

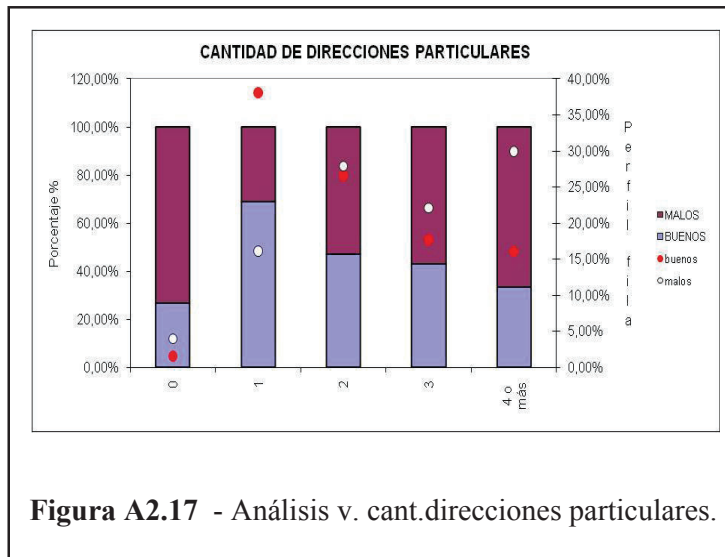
### Variable Cantidad de Teléfonos Particulares Vigentes



Lo que más existe son clientes que tienen al menos 1 teléfono particular para comunicarse telefónicamente con ellos; aquellos que tienen buen comportamiento de pago tienen al menos 1 teléfono particular vigente mientras que los calificados como malos no se los puede contactar telefónicamente a sus domicilios por cuanto no tienen teléfonos particulares vigentes. Observamos en la gráfica que mientras más teléfonos particulares vigentes tengan los clientes, mejor calificados están.



### Variable Cantidad de Direcciones Particulares Vigentes



Lo que más existe son clientes con al menos 1 dirección particular vigente mientras que la categoría menos poblada corresponde a los que no se puede visitar en el domicilio por falta de una dirección válida. Llama la atención sin embargo que exista una gran

cantidad de clientes con dos o más direcciones particulares vigentes; además del gráfico podemos concluir que en aquellos sin direcciones o con más de una dirección predominan los clientes con mal comportamiento de pago, lo que da a entender problemas de contactabilidad en la gestión de terreno.

Tabla A2.1 - Libro de códigos de las variables utilizadas

VARIABLE	DESCRIPCIÓN	MEDIDA	BASE DE DATOS
sucursal	Nombre de la sucursal	Nominal	0=Alborada 1=Ambato 2=Babahoyo 3=Cayambe 4=Cuenca 5=Guayaquil Mayor 6= Guayaquil Sur 7=Ibarra 8=La Prensa 9=Latacunga 10=Loja 11=Machala 12=Manta 13=Otavalo 14=Portoviejo 15=Quevedo 16=Quito CCNU 17=Quito Centro 18= Quito Matriz 19=Quito Sur 20=Riobamba, 21= Santo Domingo 22=Tulcán
producto	Nombre del producto	Nominal	0=Giromática 1=Microcrédito 2=Refinanciamiento Microcrédito 3=Refinanciamiento Microcrédito + Micro
calificación	Calificación del cliente	Nominal	0=Bueno 1=Malo
genero	Género del cliente	Nominal	0=Masculino 1=Femenino
estadoCiv	Estado civil del cliente	Nominal	0=Casado 1=Soltero 2=Divorciado 3=Unión Libre 4=Viudo
nivelEdu	Nivel de educación del cliente	Nominal	0=Ninguna 1=Especial 2=Primaria 3=Secundaria 4=Superior
tipoViv	Tipo de vivienda del cliente	Nominal	0=Anticresis 1=Arrendada 2=Herencia 3=Hipoteca 4=Propia
respuestaTel_ant	Respuesta a la gestión telefónica mes anterior	Nominal	0=Compromiso de pago, 1=Contactado, 2=Deja Notificación, 3=Deja Recado, 4=No quiere Pagar, 5=No ubicable, 6=Gestión de cobranza no personalizada 7=Tel No Existe, 8=Tel No Responde, 9=Tel Ocupado 10=Vocación Positiva de Pago, 11=Ya pago

VARIABLE	DESCRIPCIÓN	MEDIDA	ATRIBUTOS EN BASE DE DATOS
respuestaTel_tri	Respuesta a la gestión telefónica último trimestre	Nominal	0=Compromiso de pago, 1=Contactado, 2=Deja Notificación, 3=Deja Recado, 4=No quiere Pagar, 5=No ubicable, 6=Gestión de cobranza no personalizada 7=Tel No Existe, 8=Tel No Responde, 9=Tel Ocupado 10=Vocación Positiva de Pago, 11=Ya pago
tipoTel	Tipo de teléfonos vigentes	Nominal	0=Sin telefono 1=Telefonos laborales 2=Telefonos Particulares 3=Solo telefonos referencias 4=Telefonos laborales y particulares 5=Telefonos laborales y referenciales 6=Telefonos particulares, laborales y referenciales 7=Telefonos particulares y referenciales vigentes
tipoDir	Tipo de direcciones vigentes	Nominal	0=Solo direcciones particulares 1=Direcciones laborales y particulares vigentes 2=Sin direcciones 3=Solo direcciones laborales
resultadoGes	Resultado de la gestión de cobranza en el mes anterior	Nominal	0=Contactado 1=Compromiso de Pago 2=Seguimiento 3=No ubicable 4=En gestión
gestionHist	Tipo de gestión histórica	Nominal	0=Sin gestión 1=Solo gestión telefónica 2=Solo gestión domiciliaria 3=Gestión telefónica y domiciliaria juntas
vectorMor	Vector de moras	Nominal	Combinación de moras $i = X/X$ donde $X \in \{A, B, C, D, E\}$ e $i = \{0, 1, 2, \dots\}$ . A=mora de 0 a 15 días; B=mora de 16 a 30 días; C=mora de 31 a 60 días; D=mora de 61 a 90 días; E=mora mayor a 90 días
combinacionTelDir	Combinación de cantidad de teléfonos y direcciones	Nominal	0=Con telefonos y direcciones vigentes 1=Sin telefonos y direcciones vigentes 2=Solo direcciones vigentes 3=Solo telefonos vigentes

VARIABLE	DESCRIPCIÓN	MEDIDA	ATRIBUTOS EN BASE DE DATOS
ran_moraIni	rango de mora máxima del mes anterior	Nominal	0=mora de 1 a 15 días 1=mora de 16 a 30 días 2=mora de 31 a 60 días 3=mora de 61 a 90 días 4=mora de 91 a 120 días 5=mora de 121 a 150 días 6=mora de 151 a 180 días
ran_moraAnt	rango de mora máxima del mes anterior	Nominal	0=mora de 1 a 15 días 1=mora de 16 a 30 días 2=mora de 31 a 60 días 3=mora de 61 a 90 días 4=mora de 91 a 120 días 5=mora de 121 a 150 días 6=mora de 151 a 180 días
ran_moraAno	rango de mora máxima del último año	Nominal	0=mora de 1 a 15 días 1=mora de 16 a 30 días 2=mora de 31 a 60 días 3=mora de 61 a 90 días 4=mora de 91 a 120 días 5=mora de 121 a 150 días 6=mora de 151 a 180 días
ran_moraSem	rango de mora máxima del último semestre	Nominal	0=mora de 1 a 15 días 1=mora de 16 a 30 días 2=mora de 31 a 60 días 3=mora de 61 a 90 días 4=mora de 91 a 120 días 5=mora de 121 a 150 días 6=mora de 151 a 180 días
ran_moraTri	rango de mora máxima del último trimestre	Nominal	0=mora de 1 a 15 días 1=mora de 16 a 30 días 2=mora de 31 a 60 días 3=mora de 61 a 90 días 4=mora de 91 a 120 días 5=mora de 121 a 150 días 6=mora de 151 a 180 días
ran_moraPro	rango de mora promedio del último año	Nominal	0=mora de 1 a 15 días 1=mora de 16 a 30 días 2=mora de 31 a 60 días 3=mora de 61 a 90 días 4=mora de 91 a 120 días 5=mora de 121 a 150 días 6=mora de 151 a 180 días

VARIABLE	DESCRIPCIÓN	MEDIDA	ATRIBUTOS EN BASE DE DATOS
saldo	saldo deuda	Cuantitativa	Mínimo=27 Máximo=2500
cuotaAnt	Cuota más antigua en mora	Cuantitativa	Mínimo=1 Máximo=19
diasIni	Días en mora máximo inicial	Cuantitativa	Mínimo=3 Máximo=180
cantidadTel	Cantidad de Teléfonos	Cuantitativa	Mínimo=0 Máximo=15
cantidadDir	Cantidad de Direcciones	Cuantitativa	Mínimo=0 Máximo=10
telefonosLab	Cantidad de teléfonos laborales vigentes	Cuantitativa	Mínimo=0 Máximo=5
telefonosPar	Cantidad de teléfonos particulares vigentes	Cuantitativa	Mínimo=0 Máximo=5
telefonosRef	Cantidad de teléfonos de referencias vigentes	Cuantitativa	Mínimo=0 Máximo=5
direccionesLab	Cantidad de Direcciones Laborales vigentes	Cuantitativa	Mínimo=0 Máximo=5
direccionesPar	Cantidad de Direcciones Particulares vigentes	Cuantitativa	Mínimo=0 Máximo=8
cuotasTot	Cuotas totales del crédito	Cuantitativa	Mínimo=6 Máximo=18
cuota	Valor cuota mensual	Cuantitativa	Mínimo=25 Máximo=250
edad	Edad del cliente	Cuantitativa	Mínimo=18 Máximo=65
saldoSis	Saldo promedio en el sistema financiero	Cuantitativa	Mínimo=20 Máximo=2000
cuotaSis	Cuota promedio en el sistema financiero	Cuantitativa	Mínimo=20 Máximo=280
cargasFam	Número de cargas familiares	Cuantitativa	Mínimo=1 Máximo=5
diasAnt	Días en mora máximo mes anterior	Cuantitativa	Mínimo=3 Máximo=180
diasAno	Días en mora máximo anula	Cuantitativa	Mínimo=0 Máximo=180
diasSem	Días en mora máximo semestre	Cuantitativa	Mínimo=0 Máximo=180
diasTri	Días en mora máximo trimestre	Cuantitativa	Mínimo=0 Máximo=180
diasPro	Días en mora promedio anual	Cuantitativa	Mínimo=0 Máximo=180
contador1	Contador de moras entre 1 y 15 días	Cuantitativa	Mínimo=0 Máximo=12
contador15	Contador de moras entre 16 y 30 días	Cuantitativa	Mínimo=0 Máximo=5
contador30	Contador de moras entre 31 y 60 días	Cuantitativa	Mínimo=0 Máximo=12
contador60	Contador de moras entre 61 y 90 días	Cuantitativa	Mínimo=0 Máximo=4
contador90	Contador de moras mayor a 90 días	Cuantitativa	Mínimo=0 Máximo=1
desembolso	Valor del desembolso	Cuantitativa	Mínimo=200 Máximo=3800
porcentajePag	Porcentaje de cuotas por pagar	Cuantitativa	Mínimo=5% Máximo=100%
cuotasPag	Cuotas pagadas	Cuantitativa	Mínimo=0 Máximo=18
cuotasPen	Cuotas pendientes	Cuantitativa	Mínimo=1 Máximo=18
gestionTel_ant	Cantidad de gestiones telefónicas mes anterior	Cuantitativa	Mínimo=0 Máximo=12
gestionDom_ant	Cantidad de gestiones domiciliarias mes anterior	Cuantitativa	Mínimo=0 Máximo=2
gestionTel_tri	Cantidad de gestiones telefónicas último trimestre	Cuantitativa	Mínimo=0 Máximo=16
gestionDom_tri	Cantidad de gestiones domiciliarias último trimestre	Cuantitativa	Mínimo=0 Máximo=3

NOMBRE DE VARIABLE EN BASE DE DATOS	MEDIDA	DESCRIPCIÓN	REDISTRIBUCIÓN DE ATRIBUTOS		ATRIBUTOS EN ANÁLISIS
			ANTERIOR	NUEVO	
saldoSis_rec	Nominal	rango saldo sistema			0= menor a 324 1= 324 y más
cuotaSis_rec	Nominal	rango cuota sistema			0= menor a 53 1= 53 y más
saldo_rec	Nominal	rango saldo cierre			0=menor a 208 1= 208 y más
Porcentaje_rec	Nominal	rango % por pagar			0=menor a 67% 1= 67% y más
cuota_rec	Nominal	rango cuota mensual			0= menor a 47 1= 47 y más
telefonosLab_rec	Nominal	rango cantidad teléfonos laborales vig			0= al menos 1 1= ninguno
telefonosPar_rec	Nominal	rango cantidad teléfonos particulares vig			0= al menos 1 1= ninguno
direccionesLab_rec	Nominal	rango cantidad direcciones laborales vig			0= menor a 1 1= 1 y más
direccionesPar_rec	Nominal	rango cantidad direcciones particulares vig			0= menor a 1 1= 1 y más
cuotasTot_rec	Nominal	rango total cuotas			0= hasta 12 1= mayor a 12
cargasFam_rec	Nominal	rango cargas familiares			0=máximo 1 1= mayor a 1
contador1_rec	Nominal	rango contador moras 1 a 15			0= 12 1= menor a 12
contador15_rec	Nominal	rango contador moras 16 a 30			0= ninguna 1= al menos 1
contador30_rec	Nominal	rango contador moras 31 a 60			0= ninguna 1= al menos 1
contador60_rec	Nominal	rango contador moras 61 a 90			0= ninguna 1= al menos 1
contador90_rec	Nominal	rango contador moras mayor a 90			0= ninguna 1= al menos 1
cuotasPag_rec	Nominal	rango cuotas pagadas			0= al menos 6 1= menor a 6
cuotasPen_rec	Nominal	rango cuotas pendientes			0= máximo 6 1= mayor a 6
gestionesTel_ant_rec	Nominal	rango gestiones telefónicas mes anterior			0= al menos 1 1= ninguno
gestionesDom_ant_rec	Nominal	rango gestiones domiciliarias mes anterior			0= al menos 1 1= ninguno
gestionesTel_tri_rec	Nominal	rango gestiones telefónicas último trimestre			0= al menos 1 1= ninguno
gestionesDom_tri_rec	Nominal	rango gestiones domiciliarias último trimestre			0= al menos 1 1= ninguno

**Tabla A2.2** – Valor p asociado al estadístico chi-cuadrado para medir la asociación entre las variables independientes categorizadas y la variable respuesta calificación

VARIABLE INDEPENDIENTE CATEGÓRICA	Valor p asociado	OR
cargasFam_rec	0,627	0,955
nivelEdu_rec	0,59	1,052
respuestaTel_ant_rec	0,583	0,954
resultadoGes_rec	0,573	0,952
saldoSis_rec	0,516	0,953
genero	0,378	1,068
respuestaTel_tri_rec	0,319	1,087
gestionesTel_ant_rec	0,08	1,162
cuotasPag_rec	0,076	1,14
gestionHist_rec	0,035	1,196
saldo_rec	0,004	0,726
gestionesTel_tri_rec	0,003	1,277
edad_rec	0,003	1,248
contador60_rec	0	318,227
contador90_rec	0	249,013
contador30_rec	0	89,223
contador1_rec	0	42,045
vectora_rec	0	40,894
moraAno_rec	0	21,11
moraSem_rec	0	21,048
moraTri_rec	0	21,045
moralni_rec	0	19,881
moraPro_rec	0	19,159
Porcentaje_rec	0	14,54
moraAnt_rec	0	12,975
telefonosLab_rec	0	11,811
cuotasTot_rec	0	10,633
contador15_rec	0	10,247
TelefonosPar_rec	0	6,472
tipoTel_rec	0	5,837
combinacionTelDir_rec	0	5,178
direccionesLab_rec	0	4,627
tipoViv_rec	0	3,451
estadoCiv_rec	0	3,067
cuotasPen_rec	0	2,959
direccionesPar_rec	0	2,609
tipoDir_rec	0	0,383
desembolso_rec	0	0,361
gestionesDom_tri_rec	0	0,335
gestionesDom_ant_rec	0	0,252
cuotaSis_rec	0	0,236
cuota_rec	0	0,047

## ANEXO III

**REGLAS DE SELECCIÓN EN FORMATO SPSS UTILIZADAS PARA  
CONSTRUIR EL ÁRBOL SOBRE LA MUESTRA DE PRUEBA.**

**\* Node 1.**

```
DO IF (PRE_1 LE 0.00526777802972247).
  COMPUTE nod_001 = 1.
  COMPUTE pre_001 = 0.
  COMPUTE prb_001 = 1.000000.
```

```
END IF.
```

```
EXECUTE.
```

**\* Node 9.**

```
DO IF (SYSMIS(PRE_1) OR (VALUE(PRE_1) GT 0.00526777802972247 AND
VALUE(PRE_1) LE 0.0400706991209444)) AND
(MISSING(GESTIONESDOM_ANT_REC) OR GESTIONESDOM_ANT_REC NE 0).
  COMPUTE nod_001 = 9.
  COMPUTE pre_001 = 0.
  COMPUTE prb_001 = 0.986014.
```

```
END IF.
```

```
EXECUTE.
```

**\* Node 17.**

```
DO IF (SYSMIS(PRE_1) OR (VALUE(PRE_1) GT 0.00526777802972247 AND
VALUE(PRE_1) LE 0.0400706991209444)) AND (GESTIONESDOM_ANT_REC EQ
0) AND (telefonosPar_rec EQ 1).
  COMPUTE nod_001 = 17.
  COMPUTE pre_001 = 1.
  COMPUTE prb_001 = 1.000000.
```

```
END IF.
```

```
EXECUTE.
```

**\* Node 18.**

```
DO IF (SYSMIS(PRE_1) OR (VALUE(PRE_1) GT 0.00526777802972247 AND
VALUE(PRE_1) LE 0.0400706991209444)) AND (GESTIONESDOM_ANT_REC EQ
0) AND (MISSING(telefonosPar_rec) OR telefonosPar_rec NE 1).
  COMPUTE nod_001 = 18.
  COMPUTE pre_001 = 0.
  COMPUTE prb_001 = 0.857143.
```

```
END IF.
```

```
EXECUTE.
```

**\* Node 19.**

```
DO IF (PRE_1 GT 0.0400706991209444 AND PRE_1 LE 0.160156301655103) AND
(telefonosPar_rec EQ 1) AND (MISSING(gestionesTel_tri_rec) OR gestionesTel_tri_rec
NE 1).
```



```

        COMPUTE nod_001 = 19.
        COMPUTE pre_001 = 0.
        COMPUTE prb_001 = 0.906977.
    END IF.
EXECUTE.
* Node 20.
DO IF (PRE_1 GT 0.0400706991209444 AND PRE_1 LE 0.160156301655103) AND
(telefonosPar_rec EQ 1) AND (gestionesTel_tri_rec EQ 1).
    COMPUTE nod_001 = 20.
    COMPUTE pre_001 = 1.
    COMPUTE prb_001 = 1.000000.
END IF.
EXECUTE.
* Node 29.
DO IF (PRE_1 GT 0.0400706991209444 AND PRE_1 LE 0.160156301655103) AND
(MISSING(telefonosPar_rec) OR telefonosPar_rec NE 1) AND (gestionHist_rec EQ 0)
AND (MISSING(cuota_rec) OR cuota_rec NE 0).
    COMPUTE nod_001 = 29.
    COMPUTE pre_001 = 0.
    COMPUTE prb_001 = 0.902439.
END IF.
EXECUTE.
* Node 30.
DO IF (PRE_1 GT 0.0400706991209444 AND PRE_1 LE 0.160156301655103) AND
(MISSING(telefonosPar_rec) OR telefonosPar_rec NE 1) AND (gestionHist_rec EQ 0)
AND (cuota_rec EQ 0).
    COMPUTE nod_001 = 30.
    COMPUTE pre_001 = 1.
    COMPUTE prb_001 = 0.800000.
END IF.
EXECUTE.
* Node 22.
DO IF (PRE_1 GT 0.0400706991209444 AND PRE_1 LE 0.160156301655103) AND
(MISSING(telefonosPar_rec) OR telefonosPar_rec NE 1) AND
(MISSING(gestionHist_rec) OR gestionHist_rec NE 0).
    COMPUTE nod_001 = 22.
    COMPUTE pre_001 = 0.
    COMPUTE prb_001 = 0.970149.
END IF.
EXECUTE.
* Node 31.
DO IF (PRE_1 GT 0.160156301655103 AND PRE_1 LE 0.537344683973939) AND
(MISSING(gestionesDom_ant_rec) OR gestionesDom_ant_rec NE 0) AND
(telefonosLab_rec EQ 1) AND (MISSING(gestionesTel_tri_rec) OR gestionesTel_tri_rec
NE 1).
    COMPUTE nod_001 = 31.

```

```

        COMPUTE pre_001 = 0.
        COMPUTE prb_001 = 0.785714.
    END IF.
EXECUTE.
* Node 32.
DO IF (PRE_1 GT 0.160156301655103 AND PRE_1 LE 0.537344683973939) AND
(MISSING(gestionesDom_ant_rec) OR gestionesDom_ant_rec NE 0) AND
(telefonosLab_rec EQ 1) AND (gestionesTel_tri_rec EQ 1).
    COMPUTE nod_001 = 32.
    COMPUTE pre_001 = 1.
    COMPUTE prb_001 = 0.589744.
END IF.
EXECUTE.
* Node 39.
DO IF (PRE_1 GT 0.160156301655103 AND PRE_1 LE 0.537344683973939) AND
(MISSING(gestionesDom_ant_rec) OR gestionesDom_ant_rec NE 0) AND
(MISSING(telefonosLab_rec) OR telefonosLab_rec NE 1) AND (telefonosPar_rec EQ 1)
AND (MISSING(gestionHist_rec) OR gestionHist_rec NE 1).
    COMPUTE nod_001 = 39.
    COMPUTE pre_001 = 0.
    COMPUTE prb_001 = 0.833333.
END IF.
EXECUTE.
* Node 40.
DO IF (PRE_1 GT 0.160156301655103 AND PRE_1 LE 0.537344683973939) AND
(MISSING(gestionesDom_ant_rec) OR gestionesDom_ant_rec NE 0) AND
(MISSING(telefonosLab_rec) OR telefonosLab_rec NE 1) AND (telefonosPar_rec EQ 1)
AND (gestionHist_rec EQ 1).
    COMPUTE nod_001 = 40.
    COMPUTE pre_001 = 1.
    COMPUTE prb_001 = 1.000000.
END IF.
EXECUTE.
* Node 34.
DO IF (PRE_1 GT 0.160156301655103 AND PRE_1 LE 0.537344683973939) AND
(MISSING(gestionesDom_ant_rec) OR gestionesDom_ant_rec NE 0) AND
(MISSING(telefonosLab_rec) OR telefonosLab_rec NE 1)
AND (MISSING(telefonosPar_rec) OR telefonosPar_rec NE 1).
    COMPUTE nod_001 = 34.
    COMPUTE pre_001 = 0.
    COMPUTE prb_001 = 0.837838.
END IF.
EXECUTE.

```

**\* Node 25.**

DO IF (PRE\_1 GT 0.160156301655103 AND PRE\_1 LE 0.537344683973939) AND  
(gestionesDom\_ant\_rec EQ 0) AND (MISSING(contador1\_rec) OR contador1\_rec NE  
1).

    COMPUTE nod\_001 = 25.  
    COMPUTE pre\_001 = 0.  
    COMPUTE prb\_001 = 0.718750.

END IF.

EXECUTE.

**\* Node 35.**

DO IF (PRE\_1 GT 0.160156301655103 AND PRE\_1 LE 0.537344683973939) AND  
(gestionesDom\_ant\_rec EQ 0) AND (contador1\_rec EQ 1) AND (tipoViv\_rec EQ 0).

    COMPUTE nod\_001 = 35.  
    COMPUTE pre\_001 = 0.  
    COMPUTE prb\_001 = 0.600000.

END IF.

EXECUTE.

**\* Node 36.**

DO IF (PRE\_1 GT 0.160156301655103 AND PRE\_1 LE 0.537344683973939) AND  
(gestionesDom\_ant\_rec EQ 0) AND (contador1\_rec EQ 1)  
AND (MISSING(tipoViv\_rec) OR tipoViv\_rec NE 0).

    COMPUTE nod\_001 = 36.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 0.933333.

END IF.

EXECUTE.

**\* Node 37.**

DO IF (PRE\_1 GT 0.537344683973939 AND PRE\_1 LE 0.930639602650388) AND  
(gestionesTel\_tri\_rec EQ 0) AND (telefonosPar\_rec EQ 1)  
AND (MISSING(telefonosLab\_rec) OR telefonosLab\_rec NE 0).

    COMPUTE nod\_001 = 37.  
    COMPUTE pre\_001 = 0.  
    COMPUTE prb\_001 = 0.600000.

END IF.

EXECUTE.

**\* Node 38.**

DO IF (PRE\_1 GT 0.537344683973939 AND PRE\_1 LE 0.930639602650388) AND  
(gestionesTel\_tri\_rec EQ 0) AND (telefonosPar\_rec EQ 1) AND (telefonosLab\_rec EQ  
0).

    COMPUTE nod\_001 = 38.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 0.777778.

END IF.

EXECUTE.

**\* Node 28.**

DO IF (PRE\_1 GT 0.537344683973939 AND PRE\_1 LE 0.930639602650388) AND  
(gestionesTel\_tri\_rec EQ 0) AND (MISSING(telefonosPar\_rec) OR telefonosPar\_rec NE  
1).

    COMPUTE nod\_001 = 28.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 0.861538.

END IF.  
EXECUTE.

**\* Node 16.**

DO IF (PRE\_1 GT 0.537344683973939 AND PRE\_1 LE 0.930639602650388) AND  
(MISSING(gestionesTel\_tri\_rec) OR gestionesTel\_tri\_rec NE 0).

    COMPUTE nod\_001 = 16.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 0.866667.

END IF.  
EXECUTE.

**\* Node 6.**

DO IF (PRE\_1 GT 0.930639602650388 AND PRE\_1 LE 0.990926670371744).

    COMPUTE nod\_001 = 6.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 0.958904.

END IF.  
EXECUTE.

**\* Node 7.**

DO IF (PRE\_1 GT 0.990926670371744 AND PRE\_1 LE 0.998835245661788).

    COMPUTE nod\_001 = 7.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 0.993197.

END IF.  
EXECUTE.

**\* Node 8.**

DO IF (PRE\_1 GT 0.998835245661788).

    COMPUTE nod\_001 = 8.  
    COMPUTE pre\_001 = 1.  
    COMPUTE prb\_001 = 1.000000.

END IF.  
EXECUTE.

## ANEXO IV

### CONSTRUCCIÓN DEL MODELO DE REGRESIÓN BASADO EN DISTANCIAS UTILIZANDO EL PAQUETE R.

#### Captura de los datos en R

```
desarrolloABC <- read.table("desarrolloABC.txt", header=T, sep=" ")
dim(desarrolloABC)
n <- dim(desarrolloABC)[1]           # número de observaciones
p <- dim(desarrolloABC)[2] - 1       # número de variables regresoras
colnames(desarrolloABC)
```

#### Corrección del tipo de dato

```
desarrolloABC [,1]=as.factor(desarrolloABC [,1])
desarrolloABC [,36:48] <- lapply(desarrolloABC [,36:48],as.factor)
summary(desarrolloABC)
```

#### Cálculo de la disimilaridad

```
library(cluster)
Y <- desarrolloABC [,1]
Delta <- daisy(desarrolloABC [,-1],type=list(ordratio=
c("o1","o2","o3","o4","o5","o6","o7","o8","o9","o10","o11","o12"),asymm=c("b1","c2
5")))
class(Delta)
```

#### Análisis de coordenadas principales

```
mds <- cmdscale(Delta^(1/2), k = n-1, eig = TRUE)
names(mds)
round(mds$points[,1],4) # primera coordenada principal
```

Representación de las observaciones de las variables regresoras en dimensión 2 con las 2 coordenadas principales.

```
plot(mds$points[,1],mds$points[,2],
     main="Representación en dimensión reducida dim=2",
     xlab="Coordenada principal 1", ylab="Coordenada principal 2",
     pch=19, col="red") # falta identificar los puntos
abline(h=0)
abline(v=0)
```

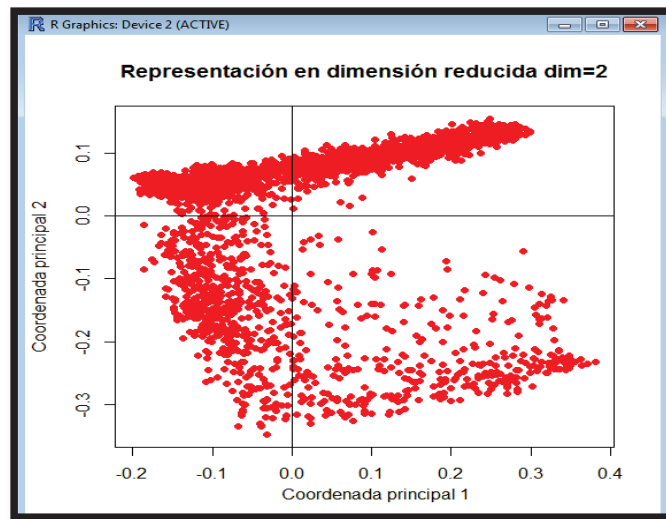


Figura A6.1 - Representación de las observaciones de las variables regresoras

La función `cmdscale` da a notar la presencia de algún valor propio negativo porque utilizamos un número excesivo de coordenadas principales. Se fija el rango de la matriz B y calcula la matriz X que contiene todas las coordenadas principales para ese rango.

```
m <- sum(mds$eig > 1.0e-15)
mds <- cmdscale(Delta^(1/2), k = m, eig = TRUE)
X <- mds$points
```

### Selección de coordenadas principales como regresoras

Para seleccionar las coordenadas principales como regresoras seleccionamos las de mayor valor propio. Presentamos los 100 primeros valores en concreto, sin embargo en el siguiente apartado se aprecia de mejor forma esta decisión.

```
ValoresPropios[1:100]
Regresión con las coordenadas seleccionadas
Xr <- X[,1:100]
rdb <- glm(Y ~Xr, data= desarrolloABC [, -1], family=binomial)
summary(rdb)
```

Se elimina del modelo las coordenadas donde el p-valor correspondiente no es significativo de modo que, en este caso, se puede definir el modelo de regresión definitivo con las coordenadas principales 1,3,5,6,7 y 9.

```
Xr <- X[,c(1,3,5,6,7,9)]
rdb <-glm(Y ~Xr, data= desarrolloABC[, -1], family=binomial)
summary(rdb)
```

### Predicción de nuevos valores sobre la muestra de prueba.

```
pruebaABC <- read.table("pruebaABC.txt", header = T, sep=" ")
dim(prueba3)
n <- dim(pruebaABC)[1] # número de observaciones
p <- dim(pruebaABC)[2] - 1 # número de variables regresoras
colnames(pruebaABC)
resultadoF <- rbind(desarrolloABC, pruebaABC)
d <- as.matrix(daisy(resultadoF[, -1], type=list(ordratio=
c("o1", "o2", "o3", "o4", "o5", "o6", "o7", "o8", "o9", "o10", "o11", "o12"),
asymm=c("b1", "c25")))[1:2931, 2932:3413])
b <- diag(X %*% t(X))
x.new <- (1/2)*diag(ValoresPropios[1:2676]^(-1)) %*% t(X) %*% (b-d)
y.pred <- t(c(1, x.new[c(1,3,5,6,7,9)])) %*% rdb$coef
y.pred
```

## ANEXO V

### LA DISTANCIA EN ESTADÍSTICA

#### Distancias para variables cuantitativas

Si  $\Omega$  es un conjunto finito,  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ , las distancias  $\bar{d}_{ij}$  se expresan mediante la matriz simétrica  $\Delta$ , llamada matriz de distancias sobre  $\Omega$ :

$$\Delta = \begin{pmatrix} \bar{d}_{11} & \bar{d}_{12} & \cdots & \bar{d}_{1n} \\ \bar{d}_{21} & \bar{d}_{22} & \cdots & \bar{d}_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \bar{d}_{n1} & \bar{d}_{n2} & \cdots & \bar{d}_{nn} \end{pmatrix} \quad \bar{d}_{ii} = 0, \quad \bar{d}_{ij} = \bar{d}_{ji}$$

Se llama preordenación de  $\Omega$  asociada a  $\Delta$ , a la ordenación de menor a mayor de los  $m = n \times (n - 1)/2$  pares de distancias no nulas:

$$\bar{d}_{i_1 j_1} \leq \bar{d}_{i_2 j_2} \leq \cdots \leq \bar{d}_{i_m j_m},$$

es decir, la ordenación de los pares  $(\omega_i, \omega_j)$  de  $\Omega$ , de acuerdo con su proximidad.

Una matriz de distancias  $\Delta$  puede ser transformada de diversos modos. Por ejemplo:

$$\bar{d}_{ij}^* = \begin{cases} 0 & i = j \\ \bar{d}_{ij} + c & i \neq j \end{cases} \quad (t1)$$

La transformación (t1), que consiste en sumar una constante fuera de la diagonal de  $\Delta$ , se llama **aditiva**. Otra transformación es:

$$\bar{d}_{ij}^2 = \begin{cases} 0 & i = j \\ \bar{d}_{ij}^2 + c & i \neq j \end{cases} \quad (t2)$$



que afecta al cuadrado de la distancia y la que llamaremos **q-aditiva**. Las transformaciones (t1) y (t2) son útiles para conseguir que la nueva distancia cumpla propiedades que la distancia original no posee, pero conservando la preordenación, es decir, las relaciones de proximidad entre los individuos.

Se supone ahora que cada individuo de  $\Omega$  puede ser representado por un punto  $x = (x_1, x_2, \dots, x_p) \in \mathbb{R}^p$ . Algunas distancias especialmente interesantes entre dos puntos  $x, y \in \mathbb{R}^p$ , son:

a) la distancia Euclídea,

$$d_E(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (d1)$$

b) la distancia "ciudad"

$$d_1(x, y) = \sum_{i=1}^p |x_i - y_i|, \quad (d2)$$

c) la distancia "valor absoluto"

$$d_A(x, y) = \sqrt{\sum_{i=1}^p |x_i - y_i|} \quad (d3)$$

Cuando  $\Omega$  se puede asimilar a una población normal multivariante  $N_p(\mu, \Sigma)$ , con  $\Sigma$  no singular, la distancia estadística (al cuadrado) más apropiada es

$$d_M^2(x, y) = (x - y)' \Sigma^{-1} (x - y) \quad (d4)$$

llamada distancia de Mahalanobis. Naturalmente, esta distancia puede ser definida en poblaciones  $(\mu, \Sigma)$ , es decir, con vector de medias  $\mu$  y matriz de covariancias  $\Sigma$ , sin necesidad de asumir normalidad.

### Similaridades y distancias con variables binarias

Se supone que existen  $p$  variables binarias  $X_1, X_2, \dots, X_p$ , donde cada  $X_i$  toma los valores 0 ó 1 según la presencia de una cierta característica. Son conocidos los siguientes coeficientes de similaridad entre cada par de individuos  $\omega_i, \omega_j$  :

$$s_{ij} = \frac{a + d}{p} \quad (\text{Sokal - Michener}) \quad (s1)$$

$$s_{ij} = \frac{a}{a + b + c} \quad (\text{Jaccard}) \quad (s2)$$

siendo  $a, b, c, d$  las frecuencias de  $(1,1), (1,0), (0,1)$  y  $(0,0)$ , respectivamente.

Nótese que  $p = a + b + c + d$ . Estas similaridades pueden ser transformadas en distancias.

### Similaridad con variables mixtas

Si las variables son mixtas, contínuas, binarias o cualitativas, es entonces adecuado utilizar la distancia de Gower,  $d_{ij}^2 = 1 - s_{ij}$ , siendo

$$s_{ij} = \left( \sum_{h=1}^{p_1} (1 - |x_{ih} - x_{jh}|G_h) + a + \alpha \right) / (p_1 + (p_2 - d) + p_3) \quad (s3)$$

una similaridad, donde  $p_1$  es el número de variables cuantitativas,  $a$  y  $d$  corresponden al número de coincidencias y no coincidencias para las  $p_2$  variables binarias, respectivamente, y  $\alpha$  es el número de coincidencias para las  $p_3$  variables cuantitativas.  $G_h$  es el rango de la  $h$ -ésima variable cuantitativa. Este coeficiente

admite la posibilidad de tratar datos faltantes y se reduce al coeficiente de Jaccard cuando  $p_1 = p_3 = 0$

### **EL MODELO DE REGRESIÓN BASADO EN DISTANCIAS DB (distance-based).**

Se considera ahora el modelo lineal que relaciona una variable respuesta con diversas variables explicativas:

$$y = X\beta + e. \quad (m1)$$

Aquí  $y(n \times 1)$  es un vector (conocido) con  $n$  observaciones de una variable respuesta cuantitativa  $Y$ , la matriz  $X(n \times m)$  es conocida de  $\text{rang}(X) = m$ ,  $\beta(m \times 1)$  es un vector (desconocido) de parámetros y  $e = (e_1, \dots, e_n)'$  es un vector aleatorio tal que

$$\begin{aligned} E(e_i) &= 0 \quad \text{var}(e_i) = \sigma^2 \quad i = 1, \dots, n \\ E(e_i e_j) &= 0 \quad i \neq j \end{aligned}$$

En muchas ocasiones se cumple que  $e$  es  $N_n(0, \sigma^2 I_n)$  y entonces se dice que  $y = X\beta + e$  es un modelo lineal normal. Se tiene que, la estimación LS (mínimos cuadrados) de  $\beta = (\beta_1, \dots, \beta_m)'$  es

$$\hat{\beta} = (X'X)^{-1} X'y \quad (m2)$$

la suma de cuadrados residual es

$$R_0^2 = (y - X\hat{\beta})'(y - X\hat{\beta}) = \hat{e}'\hat{e} \quad (m3)$$

y una estimación insesgada de  $\sigma^2$  es

$$\hat{\sigma}^2 = \frac{R_0^2}{n - m}$$

A continuación vamos a interpretar el modelo lineal desde la perspectiva de las distancias. Indiquemos por  $x_1, \dots, x_n$  las filas de  $\mathbf{X}$ . La distancia Euclídea cuadrática entre cada par de individuos  $\omega_i, \omega_j$  es

$$d_E^2(\omega_i, \omega_j) = (x_i - x_j)(x_i - x_j)' \quad (d5)$$

Sobre la matriz de distancias Euclídeas

$$D = (d_E(\omega_i, \omega_j))$$

se construye  $\mathbf{B} = \mathbf{H}\mathbf{A}\mathbf{H}$  y hallamos la descomposición espectral

$$\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}' = \tilde{\mathbf{X}}\tilde{\mathbf{X}}' \quad (m4)$$

Resulta entonces que las distancias entre las filas  $\tilde{x}_1, \dots, \tilde{x}_n$ , de  $\tilde{\mathbf{X}}$  reproducen exactamente la distancia (d5). Podemos, en consecuencia, reescribir el modelo (m1) como

$$y = \tilde{\mathbf{X}}\gamma + e \quad (m5)$$

donde  $\beta \rightarrow \gamma$  es una reparametrización.

Lo que importa aquí es notar que (m5) ha sido construido utilizando solamente distancias (en este caso la distancia Euclídea) y se le llama modelo DB (distance-based).

El vector proyector de  $y$  del modelo (m1) es

$$\hat{y} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y \quad (m6)$$

Análogamente se puede encontrar  $\hat{\tilde{y}}$  para el modelo (m6). Pero ambos modelos son esencialmente el mismo, puesto que

$$\hat{y} = \hat{\tilde{y}} \quad (m7)$$

Como se puede ver enseguida, el modelo DB presenta ventajas con otras distancias. La distancia Euclídea ( $d_5$ ) ha sido calculada utilizando las filas de la matriz de diseño  $\mathbf{X}$  que es conocida. Vamos a introducir el modelo de regresión DB (distance based) en general.

Supongamos que tenemos  $p$  variables  $W_1, W_2, \dots, W_p$  observables, de tipo contínuo, binario o categórico, o incluso los tres tipos a la vez, en cuyo caso diremos que los datos son mixtos. Sea  $d(\omega_i, \omega_j)$  una distancia adecuada entre pares  $\omega_i, \omega_j$  de individuos. Si los datos son binarios  $d(\omega_i, \omega_j)$  se puede basar en (s1) ó (s2), y si son mixtos en el coeficiente de similitud de Gower (s3). Supongamos que la distancia tiene la propiedad de ser Euclídea. A partir de  $d(\omega_i, \omega_j)$  se puede obtener la matriz  $n \times n$  de distancias  $\mathbf{\Delta}$ , y aplicando la descomposición espectral  $\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}' = \mathbf{X}\mathbf{X}'$ , obtendremos la matriz  $\mathbf{X}$ , de coordenadas principales que reproducen las distancias originales. El modelo DB que proponemos es entonces

$$y = \beta_0 \mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (\text{m8})$$

donde  $\mathbf{1}$  es el vector de unos, mientras que  $\mathbf{X}$ ,  $\boldsymbol{\beta}$  y  $\mathbf{e}$  tienen el mismo significado que en el modelo (m1). Observemos que, como  $\mathbf{B}\mathbf{1} = \mathbf{0}$ , tanto  $\mathbf{1}$  como las columnas  $X_1, \dots, X_m$  de  $\mathbf{X}$ , son vectores propios de  $\mathbf{B}$ .

Podemos también escribir

$$y = \beta_0 \mathbf{1} + \sum_{i=1}^m \beta_1 X_i + \mathbf{e} \quad (\text{m9})$$

donde  $m = \text{rang}(\mathbf{B})$  y  $X_1, \dots, X_m$ , juegan el papel de **variables predictoras**.

Las propiedades básicas del modelo DB son:

1. Las estimaciones de los parámetros de regresión son

$$\hat{\beta}_0 = \bar{y} = \mathbf{y}'\mathbf{1} / n \quad \hat{\beta}_i = \mathbf{y}'\mathbf{X}_i / \lambda_i$$

2. El vector predictor o proyección ortogonal de  $y$  es

$$\hat{y} = \bar{y}\mathbf{1} + \mathbf{X}\mathbf{\Lambda}^{-1}\mathbf{X}'\mathbf{y}$$

3. El coeficiente de correlación simple  $r_i = r(y, X_i)$  es

$$r_i^2 = (\mathbf{y}'\mathbf{X}_i) / n s_y^2 \lambda_i$$

donde  $s_y^2$  es la variancia muestral de  $y$ .

4. El coeficiente de correlación múltiple (al cuadrado)  $R^2$  entre  $y$  y  $X_1, \dots, X_m$  es

$$R^2 = \mathbf{y}'\mathbf{X}\mathbf{\Lambda}^{-1}\mathbf{X}'\mathbf{y} / n s_y^2 = \sum_{i=1}^m r_i^2$$

donde  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_m)$  contiene los valores propios de  $\mathbf{B}$ .

### Regresión DB en dimensión reducida

El modelo DB (m8-m9), es el **modelo completo**. A veces, sin embargo,  $m = \text{rango}(\mathbf{B})$  crece con  $n$  (incluso puede darse el caso en que  $m = n - 1$ ).

Entonces, el número de variables  $X_1, \dots, X_m$  (las columnas de  $\mathbf{X}$ ) puede resultar excesivo y de esa manera encontrarnos con un coeficiente de determinación  $R^2$  arbitrariamente próximo a 1. Para evitar este problema, es conveniente partir  $\mathbf{X}$  en dos partes

$$\mathbf{X} = (\mathbf{X}_{(k)}, \mathbf{Z})$$

donde  $\mathbf{X}_{(k)} = (X_1, \dots, X_k)$  contiene k columnas adecuadas de  $\mathbf{X}$ , y la matriz  $\mathbf{Z}$  contiene las restantes columnas.

Definimos de este modo el modelo DB en dimensión k, que puede ser expresado de dos maneras equivalentes:

$$y = \beta_0 \mathbf{1} + \mathbf{X}_{(k)} \beta_{(k)} + e_k,$$

$$y = \beta_0 \mathbf{1} + \sum_{i=1}^k X_i \beta_i + e_k.$$

Como valor de k, se puede tomar k= número inicial de variables observables explicativas.

Una buena selección de las columnas  $X_1, \dots, X_k$  de  $\mathbf{X}$  consiste en escogerlas por orden de correlación con  $\mathbf{y}$ , es decir,

$$r(\mathbf{y}, X_1) > r(\mathbf{y}, X_2) > \dots > r(\mathbf{y}, X_k).$$

Otra selección obvia consiste en ordenarlas de acuerdo con la variabilidad explicada por las variables predictoras (columnas de  $\mathbf{X}$ ):  $\lambda_1 > \dots > \lambda_k$ , es decir, seleccionar los k primeros ejes principales. Pero si resultara que la variable  $X_{k+1}$  tiene una correlación  $r_{k+1} = r(\mathbf{y}, X_{k+1})$  relativamente alta, podríamos haber perdido una variable predictiva importante.

### **Predicción DB sobre un nuevo individuo**

Se supone ahora que, sobre las variables (mixtas) explicativas, se ha obtenido la observación  $w_{n+1} = (w_1, \dots, w_p)$  sobre un nuevo individuo  $w_{n+1}$ . Entonces debería ser posible calcular las distancias

$$d_i^2 = d^2(\omega_i, \omega_{n+1}) \quad i = 1, \dots, n$$

entre  $w_{n+1}$  y los otros individuos cuyas observaciones conocemos para la variable respuesta  $Y$ . Se quiere evaluar

$$y_{n+1} = Y(w_{n+1}),$$

es decir, la predicción de  $Y$  sobre  $w_{n+1}$ .

Se puede obtener esta predicción mediante la fórmula de añadir un punto (f1), hallando las coordenadas del nuevo individuo a partir de las distancias. Estas coordenadas, utilizando el modelo completo, son

$$x = (x_1, \dots, x_m)' = \frac{1}{2} \Lambda^{-1} X'(b - d).$$

La predicción según el modelo (m8) es

$$\hat{y}_{n+1} = \bar{y} + \sum_{i=1}^m \hat{\beta}_i X_i = \bar{y} + x' \hat{\beta}$$

Substituyendo, se obtiene

$$\hat{y}_{n+1} = \bar{y} + x' \Lambda^{-1} X' y \quad (f2)$$

Si consideramos ahora el modelo DB en dimensión  $k$ , y hacemos las particiones

$$x = \begin{pmatrix} x_{(k)} \\ z \end{pmatrix} \quad X = (X_{(k)}, Z) \quad \Lambda = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \Lambda_{m-k} \end{pmatrix}$$

donde  $x_{(k)} = (x_1, \dots, x_k)'$  son las  $k$  coordenadas relativas a las  $k$  dimensiones predictivas, y la diagonal de  $\Lambda_k$  contiene los valores propios, se obtiene

$$\hat{y}_{n+1}(k) = \bar{y} + x'_{(k)} \Lambda_k^{-1} X'_{(k)} y + z' \Lambda_{m-k}^{-1} Z' y.$$



Si ahora se tiene en cuenta que  $Z'y \approx 0$  (ya que  $Z$  contiene las variables menos correlacionadas con  $Y$ ), se obtiene finalmente:

$$\hat{y}_{n+1}(k) = \bar{y} + x'_{(k)} \Lambda_k^{-1} X'_k y$$

### **Predicción con variables continuas, categóricas y mixtas**

El modelo DB se reduce al modelo clásico de regresión cuando la distancia utilizada es Euclídea (d1) y las variables son continuas. Se tiene que<sup>32</sup>:

- Para variables cuantitativas y distancia Euclídea, la fórmula de predicción (f2) brinda los mismos resultados.
- Para  $p$  variables cualitativas  $W_1, W_2, \dots, W_p$ , donde  $W_r$  tiene  $q_r$  estados ( $1 \leq r \leq p$ ), podemos tomar como distancia

$$d_{ij}^2 = 2(p - m_{ij}) \quad (d6)$$

donde  $m_{ij}$  es el número de coincidencias entre los individuos  $i$  y  $j$ . En ese caso, el modelo DB con la distancia (d6) vuelve a dar los mismos resultados que el modelo de regresión clásica, es decir las predicciones son las mismas. Naturalmente los resultados son diferentes si consideramos otras distancias.

- La situación cambia si las variables son mixtas, esto es, una mezcla de continuas, binarias y categóricas. Entonces la distancia no es Euclídea en el sentido de antes. Una buena elección consiste en emplear la distancia de Gower (s3). Existen muchos ejemplos que prueban que utilizando el método de regresión DB con esta distancia podemos obtener mejores predicciones que con el método clásico.

---

<sup>32</sup> Cuadras y Arenas (1990)

## ANEXO VI

### VALORACIÓN DE LA CONFUSIÓN E INTERACCIÓN PARA LAS VARIABLES INDEPENDIENTES DEL MODELO.

#### Factor género y factor edad

Se refiere a explorar el papel del género sobre el tipo de cliente, teniendo en cuenta la posible confusión que puede efectuar sobre esta asociación la variable edad, una dicotómica que establece dos perfiles de acuerdo a la mediana: “menores a 42 años” y “mayores o iguales a 42 años”, de forma que los clientes jóvenes podrían tener mayor riesgo de convertirse en impagos, y a su vez suponiendo una posible asociación con la variable género, de forma que tal vez los clientes hombres tengan menos preocupación por pagar correctamente sus obligaciones.

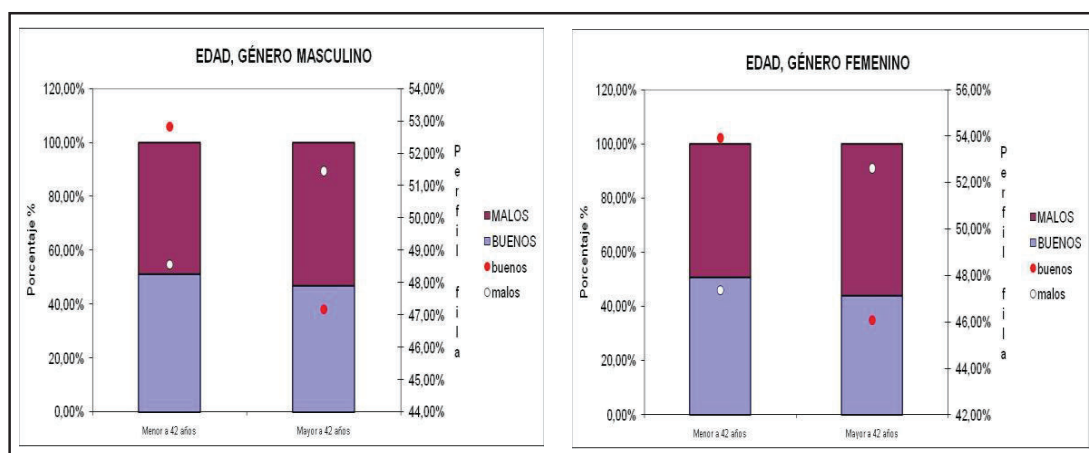
Investigué primero la relación principal entre el género y el tipo de cliente a través de un análisis bivariante simple, utilizando una tabla de contingencia 2x2 (Tabla A2.2) donde se muestra que existe una asociación estadísticamente significativa con el test Chi cuadrado ( $X^2 = 0.78$  con valor  $p = 0.378$ ) y una medida de fuerza de asociación  $OR(\text{masculino/femenino}) = 1.068$  (IC95% 0.923 a 1.236) que no identifica al género como un factor de riesgo de convertirse en mal cliente, es decir que entre los malos clientes no hay mayor probabilidad de ser mujer que de ser hombre. OR

A continuación realizando el mismo análisis, pero estratificando por la edad que podía ser un factor de confusión, obtuve el resultado mostrado en la Tabla A6.1.

**Tabla A6.1 - OR entre edad, género y tipo de cliente**

edad_rec		Valor	Intervalo de confianza al 95%	
			Inferior	Superior
42 y más	Razón de las ventajas para gen_cli (Masculino / Femenino)	1,020	,831	1,253
	Para la cohorte califica = BUENO	1,010	,914	1,117
	Para la cohorte califica = MALO	,990	,891	1,099
	N de casos válidos	1482		
menor a 42	Razón de las ventajas para gen_cli (Masculino / Femenino)	1,119	,908	1,377
	Para la cohorte califica = BUENO	1,063	,949	1,190
	Para la cohorte califica = MALO	,950	,864	1,045
	N de casos válidos	1449		

Se observa que hay dos estratos establecidos por las categorías de la variable edad y en cada estrato se muestran los valores para la distribución 2x2 de género y tipo de cliente, además de los valores de asociación OR con sus intervalos de confianza. Esta relación también se muestra en la Figura A6.1.



**Figura A6.1 - Relación entre género, edad y tipo de cliente**

Como en este caso ambas OR son parecidas (1.020 y 1.119) y prácticamente iguales a la OR de género sin estratificar por edad (1.068), no se puede concluir

de forma categórica la ausencia de confusión o interacción de la edad con la variable principal género. Para mejorar este análisis, suponiendo antes que no existe confusión, utilicé regresión logística para comprobar la existencia de interacción de la variable edad, introduciéndola simultáneamente con el género y además el término multiplicativo *edad\_rec\*genero\_cod*. El paquete SPSS arrojó los siguientes resultados:

**Tabla A6.2** - Análisis de la interacción entre género, edad y tipo de cliente

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	genero_cod(1)	.061	.102	.359	1	.549	1.063
	edad_rec(1)	-.222	.111	3.969	1	.046	.801
	edad_rec(1) by genero_cod(1)	.009	.149	.004	1	.953	1.009
	Constant	.136	.076	3.192	1	.074	1.146

a. Variable(s) entered on step 1: genero\_cod, edad\_rec, edad\_rec \* genero\_cod .

Como se puede ver en la Tabla A6.2, el término de interacción no es significativo ( $p=0.953$ ), por lo que puedo descartar que exista modificación de efecto. En este caso tiene sentido explorar en segundo lugar si hay confusión, sacando de las covariables el término de interacción para dejar el género y la edad como se muestra en la Tabla A6.3.

**Tabla A6.3** - Análisis de la confusión entre género, edad y tipo de cliente

		Variables in the Equation						95% C.I. for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 <sup>a</sup>	genero_cod(1)	.065	.075	.769	1	.381	1.068	.922	1.235
	edad_rec(1)	-.217	.074	8.565	1	.003	.805	.696	.931
	Constant	.134	.066	4.172	1	.041	1.143		

a. Variable(s) entered on step 1: genero\_cod, edad\_rec.

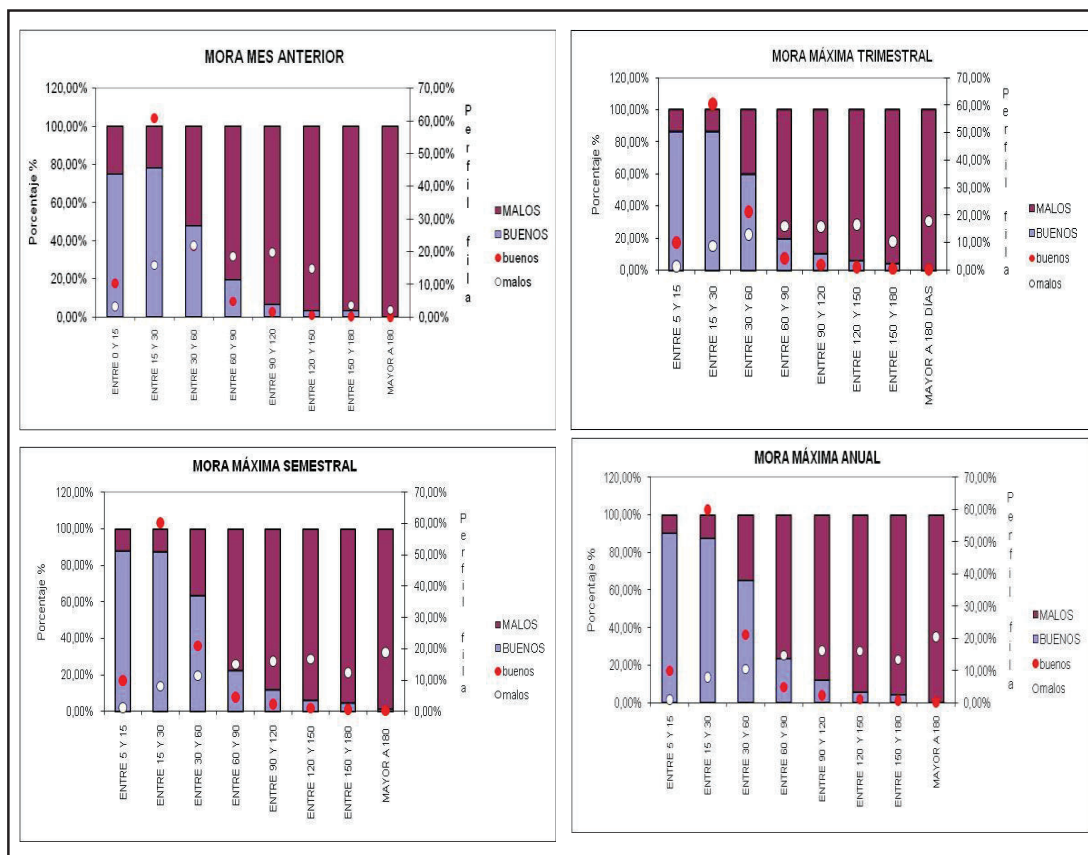
El coeficiente de regresión de la variable género es no significativo ( $p=0.381$ ) y los intervalos de confianza de la OR contienen el valor nulo (entre 0.922 y 1.235).

Esto indica a la vez que la variable edad es una predictora del tipo de cliente, puesto que la asocia a ella una OR - ajustada por la variable género - de 0.805 y con un coeficiente de regresión estadísticamente significativo. Se puede decir entonces que tener más de 42 años reduce el riesgo de ser mal pagador en un 20%.

### **Factor morosidad y factor edad**

Analizando las variables de morosidad: mora del mes anterior, mora máxima en el último trimestre, mora máxima en el último semestre y mora máxima en el año, de acuerdo a los gráficos de distribución para el tipo de cliente (Figura A6.2), se observa un comportamiento muy similar entre todas las variables que representan el atraso máximo en el pago de las cuotas mensuales, por esta razón solo desarrollé el análisis para el caso de la mora en el mes anterior.

La mora anterior muestra buen comportamiento de pago hasta los 30 días, luego, al pasar a un rango de mora de hasta 60 días, el comportamiento de pago se deteriora y finalmente cuando el cliente sobrepasa este umbral, el comportamiento se hace notoriamente malo. Por otra parte pude ver que los clientes con buen comportamiento de pago se concentran en los rangos de hasta 30 días, mientras que los malos lo hacen en el resto de rangos de morosidad.



**Figura A6.2 - Relación entre morosidad y tipo de cliente**

Con estas observaciones pude concluir que existen tres grupos bien definidos: de 1 hasta 30 días, de 31 a 60 días y el de 61 días de mora en adelante, sin embargo conviene analizar si colapsando esta variable ordinal en una categórica dicotómica se facilita el análisis y la interpretación de la asociación entre la morosidad y el tipo de cliente. Para esto analicé con regresión logística si existen diferencias significativas entre las medidas de asociación en ambos casos, ya que el uso de tablas de contingencia no proporciona mucha información cuando se trata de variables con más de dos categorías.

### Morosidad del mes anterior (tomando dos categorías: 0 y 30 días; mayor a 30 días)

Como puede apreciarse en la Tabla A6.4, el coeficiente de regresión de la variable morosidad del mes anterior es estadísticamente significativo ( $p < 0.05$ ) y su OR es 12.97, que significa que los clientes con más de 30 días de mora multiplican por 13 su riesgo de convertirse en malos pagadores.

**Tabla A6.4** - Relación entre morosidad del mes anterior y tipo de cliente

		Variables in the Equation					95% C.I. for EXP(B)		
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 <sup>a</sup>	moraAnt_rec(1)	2.563	.090	817.579	1	.000	12.975	10.884	15.467
	Constant	-1.209	.063	372.665	1	.000	.298		

a. Variable(s) entered on step 1: moraAnt\_rec.

### Morosidad del mes anterior (tomando tres categorías: entre 0 y 30 días; entre 30 y 60 días y mayor a 60 días)

Al tratarse de una variable policotómica introducida con tres categorías, el programa SPSS la convierte automáticamente en dos variables dicotómicas dummies, para poder calcular la OR de cada categoría frente a otra que sirve de referencia, como se muestra a continuación en la Tabla A6.5.

**Tabla A6.5** - Creación de variables dummy morosidad del mes anterior

		Categorical Variables Codings		
		Frequency	Parameter coding	
			(1)	(2)
ran_mor_ant_r1	entre 0 y 30	1440	.000	.000
	entre 31 y 60	530	1.000	.000
	mayor a 60	961	.000	1.000

Se crean dos variables nuevas tomando a la morosidad entre 0 y 30 días como categoría de referencia, porque es la que tiene una codificación más baja en la variable original; la segunda es una variable dicotómica en la que el valor “1” es “entre 30 y 60 días” y finalmente una tercera que es una dicotómica en la que el valor “1” es “mayor a 60 días”.

En la Tabla A6.6 muestro un modelo de regresión con estas dos variables nuevas que llevan información desagregada de la antigua morosidad; de hecho la variable original no tiene interpretación en la ecuación, está solo para indicar que de ella se han generado las dos variables dummies.

**Tabla A6.6** - Relación entre morosidad (dummy) y tipo de cliente

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	ran_mor_ant_r1			798.670	2	.000	
	ran_mor_ant_r1(1)	1.583	.108	213.556	1	.000	4.871
	ran_mor_ant_r1(2)	3.479	.127	747.922	1	.000	32.425
	Constant	-1.209	.063	372.665	1	.000	.298

a. Variable(s) entered on step 1: ran\_mor\_ant\_r1.

Pude concluir que, comparado con tener “entre 0 y 30 días de mora”, tener “entre 30 y 60 días de mora” multiplica por 4,87 el riesgo de convertirse en mal cliente y tener “más de 60 días de mora” multiplica por 32.42 las probabilidades de tener un mal comportamiento de pago; adicionalmente para estas dos variables se alcanza la significación estadística.

Finalmente analicé el efecto que produce la variable edad en la relación morosidad y tipo de cliente, para determinar si las personas mayores tienden a regularizar mejor su situación de morosidad comparado con las personas más jóvenes. Figura A6.3.



**Tabla A6.7** - Análisis de la interacción entre morosidad del mes anterior, edad y tipo de cliente

		Variables in the Equation					95% C.I. for EXP(B)		
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 <sup>a</sup>	moraAnt_rec(1)	2.574	.126	415.562	1	.000	13.117	10.241	16.799
	edadRec(1)	.281	.126	4.997	1	.025	1.324	1.035	1.694
	edadRec(1) by moraAnt_rec(1)	-.008	.180	.002	1	.962	.992	.697	1.411
	Constant	-1.351	.091	220.415	1	.000	.259		

a. Variable(s) entered on step 1: moraAnt\_rec, edadRec, edadRec \* moraAnt\_rec.

Como se observa en la Tabla A6.7, el término de interacción no es significativo ( $p=0.962$ ), por lo que descarté que exista modificación de efecto. En este caso tuvo sentido explorar si hay confusión, eliminando de las covariables el término de interacción y dejando las variables morosidad y edad como muestro en la Tabla A6.8.

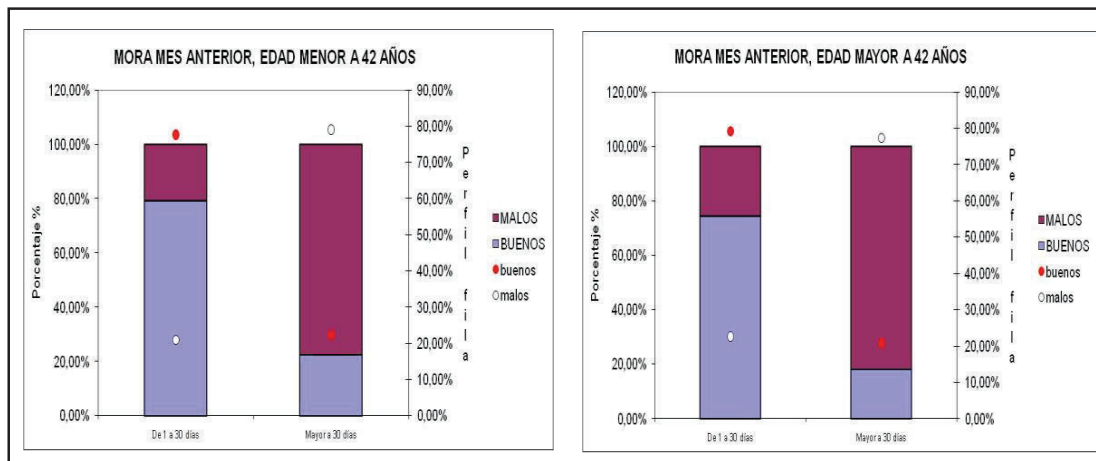
**Tabla A6.8** - Análisis de la confusión entre morosidad del mes anterior, edad y tipo de cliente

		Variables in the Equation					95% C.I. for EXP(B)		
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 <sup>a</sup>	moraAnt_rec(1)	2.570	.090	815.917	1	.000	13.062	10.950	15.581
	edadRec(1)	.277	.090	9.466	1	.002	1.319	1.106	1.573
	Constant	-1.349	.078	295.466	1	.000	.260		

a. Variable(s) entered on step 1: moraAnt\_rec, edadRec.

El coeficiente de regresión de la variable morosidad es estadísticamente significativo y su OR es 13.064, ajustada por la variable edad; la decisión de obtener un estimador diferente y menos sesgado de la asociación entre morosidad y tipo de cliente, depende si al introducir la edad cambia la OR cruda de forma importante, que para este estudio lo establecí en al menos un 10%. En este caso el cambio en la OR de la morosidad que se produjo al introducir la variable edad, no fue suficiente (pasó de 12.975 a 13.064 y esto es apenas un 1% aproximadamente) y a pesar que su coeficiente de regresión ( $B= 2.570$ ) es

significativo, concluí que la edad no modifica el efecto de la morosidad sobre la variable calificación.



**Figura A6.3-** Relación entre morosidad del mes anterior, edad y tipo de cliente

### Factor porcentaje de cuotas por pagar y factor morosidad

A continuación examiné lo que ocurre con otra variable que está reconocida como un factor de riesgo de tener mal comportamiento de pago, que es el porcentaje de cuotas por pagar, una categórica que distingue dos perfiles de clientes definidos por la mediana: “67% de cuotas pagadas o más” y “menos de 67% de cuotas pagadas”

Analizando el porcentaje de cuotas por pagar utilizando regresión logística (Tabla A6.9), pude observar que la OR es 14.55 y que el coeficiente de regresión es estadísticamente significativo; esto significa que los clientes que no han pagado un 67% de las cuotas totales de su crédito, tienen 15 veces más riesgo de ser malos que aquellos que sí lo tienen pagado.

**Tabla A6.9** - OR entre porcentaje de cuotas por pagar y tipo de cliente

		Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 <sup>a</sup>	porcen_rec(1)	2.677	.130	423.369	1	.000	14.545	11.271	18.770
	Constant	-.467	.044	112.518	1	.000	.627		

a. Variable(s) entered on step 1: porcen\_rec.

La salida de regresión introduciendo las dos covariables, porcentaje pagado y morosidad, se presenta en la Tabla A6.10.

**Tabla A6.10** - OR entre porcentaje de cuotas por pagar, mora mes anterior y tipo de cliente

		Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 <sup>a</sup>	morAnt_rec(1)	2.871	.108	701.760	1	.000	17.662	14.282	21.843
	porcen_rec(1)	3.095	.152	416.130	1	.000	22.083	16.403	29.730
	Constant	-2.011	.087	528.447	1	.000	.134		

a. Variable(s) entered on step 1: ran\_mor\_ant, porcen\_rec.

En este caso se incrementa el valor de la OR de morosidad, pasando de 12.98 a 17.66, es decir un 30%, con un coeficiente de regresión significativo ( $p=0.00$ ) y los intervalos de confianza de la OR sin contener el valor nulo, manteniendo su precisión porque el rango no se hace más grande al introducir la nueva variable. Descubrí que a la vez la variable porcentaje pagado es una predictora del tipo de cliente, puesto que la asocia una OR ajustada por la morosidad de 22.083 y con un coeficiente de regresión estadísticamente significativo.

Concluí entonces que la verdadera OR de morosidad para el desenlace predicho tipo de cliente es 17.7 y este es el riesgo que tienen los clientes con más de 30 días de atraso frente a los que no lo tienen de convertirse en “mal cliente”, sin tomar en cuenta el porcentaje de cuotas que les falta pagar. Pero también crece la OR de esta última variable sobre su estimación cruda en proporción importante

(57%) y permanece con coeficiente de regresión estadísticamente significativo con la única desventaja que aumenta el intervalo de confianza, que se traduce en estimaciones más imprecisas.

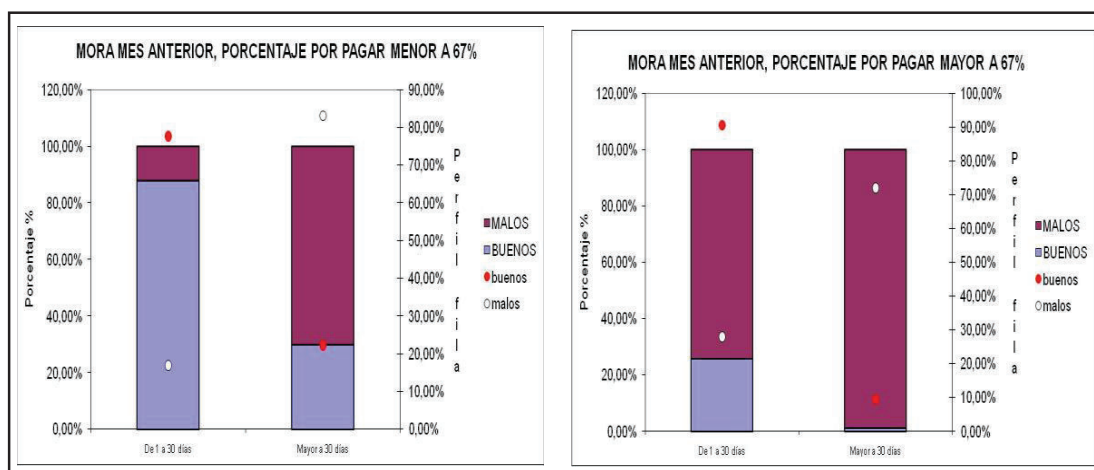
De cualquier manera las nuevas medidas de asociación son mejores que las originales, concluyendo entonces que la variable porcentaje por pagar confunde la relación entre la morosidad y el tipo de cliente. Cuando esto ocurre es necesario explorar no sola la asociación entre las variables predictoras, sino una posible interacción entre ellas y su acción sobre el efecto de convertirse en mal cliente.

Analiqué primero si existe correlación entre la morosidad del mes anterior y el porcentaje de cuotas por pagar a través de un análisis simple con tablas de contingencia, como muestro en la Tabla A6.11 y la Figura A6.4.

**Tabla A6.11** - relación entre porcentaje de cuotas por pagar y mora mes anterior

**ran\_mor\_ant \* porcen\_rec Crosstabulation**

			porcen_rec		Total
			menor a 67%	67% y mayor	
moraAnt_rec(1)	entre 0 y 30	Count	1184	256	1440
		% within ran_mor_ant	82.2%	17.8%	100.0%
	mayor a 30	Count	998	493	1491
		% within ran_mor_ant	66.9%	33.1%	100.0%
Total	Count	2182	749	2931	
	% within ran_mor_ant	74.4%	25.6%	100.0%	



**Figura A6.4** - Relación entre morosidad del mes anterior, porcentaje de cuotas por pagar y tipo de cliente

Existe una clara desproporción entre los clientes que ya han pagado al menos el 67% de las cuotas del crédito, independientemente de su comportamiento de pago. A continuación muestro la salida de la regresión logística con el término de interacción `moraAnt_rec*porcen_rec` (Tabla A6.12), para analizar si existe modificación de efecto sobre la relación entre el porcentaje de cuotas por pagar y el tipo de cliente por añadir la morosidad.

**Tabla A6.12** - Análisis de la interacción entre morosidad del mes anterior, porcentaje de cuotas por pagar y tipo de cliente

		Variables in the Equation					95% C.I. for EXP(B)		
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1 <sup>a</sup>	<code>porcen_rec(1)</code>	3.030	.168	325.379	1	.000	20.700	14.893	28.771
	<code>rmoraAnt_rec(1)</code>	2.842	.113	632.270	1	.000	17.155	13.746	21.409
	<code>porcen_rec(1) by rmoraAnt_rec(1)</code>	.361	.422	.732	1	.392	1.435	.628	3.279
	Constant	-1.993	.089	496.420	1	.000	.136		

a. Variable(s) entered on step 1: `porcen_rec`, `ran_mor_ant`, `porcen_rec * ran_mor_ant`.

A más de no ser estadísticamente significativo ( $p=0,392$ ), el término de interacción disminuye información en la estimación de las OR de morosidad y el porcentaje

de cuotas pagadas, mismas que se reducen a 17.1 y 20.7 respectivamente; además estas estimaciones se hacen con más imprecisión traducido en mayores errores estándar (S.E.) de los coeficientes de regresión y unos más amplios intervalos de confianza. Con estos resultados el término de interacción fue removido de la ecuación.

En la Tabla A6.13 muestro en qué porcentaje cambian las OR de las variables de cobranza cuando se introducen otras variables de comportamiento.

**Tabla A6.13** - Variación de las OR en las variables de cobranza

VARIABLES ORIGINALES	VARIABLES INTRODUCIDAS			
	<i>moraTri_rec</i>	<i>vector moras</i>	<i>contador moras 30</i>	<i>porcentaje</i>
combinacionTelDir_rec	42%	63%	22%	6%
direccionesLab_rec	-64%	-66%	103%	21%
direccionesPar_rec	-28%	-38%	58%	-3%
gestionesDom_ant_rec	-10%	-3%	-11%	-22%
gestionesDom_tri_rec	-18%	-4%	3%	-17%
gestionesTel_ant_rec	-23%	-27%	23%	-8%
gestionesTel_tri_rec	-32%	-34%	21%	-8%
gestionHist_rec	-21%	-33%	-9%	-31%
respuestaTel_ant_rec	-9%	-16%	15%	-14%
respuestaTel_tri_rec	-25%	-25%	18%	-10%
resultadoGes_rec	-6%	-12%	16%	-13%
telefonosLab_rec	-22%	-30%	26%	27%
telefonosPar_rec	23%	42%	32%	9%
tipoDir_rec	39%	61%	-37%	3%
tipoTel_rec	15%	32%	35%	10%

### Factores de comportamiento sobre el tipo de cliente

Realicé el análisis entre las variables de comportamiento empezando con la mora máxima del último trimestre, misma que está relacionada con la variable calificación como fue mostrado en el apartado anterior; pero a su vez quiero confirmar que podría también estar asociada con la morosidad inicial, de forma que tal vez los clientes que actualmente están en un rango de mora elevado,

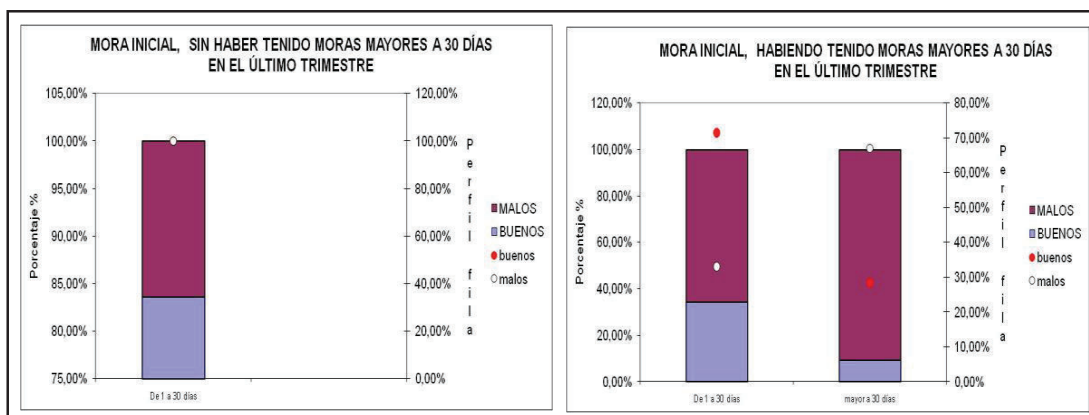
estuvieron en la misma situación hace tres meses y las gestiones de cobranza no han tenido el efecto deseado sobre la cartera para normalizar la morosidad (Figura A6.5). La salida de la regresión introduciendo las dos covariables, mora inicial y mora máxima trimestral, es como se presenta en la Tabla A6.14.

**Tabla A6.14** - Relación entre mora inicial, mora trimestral y tipo de cliente

Variables in the Equation		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 <sup>a</sup>	moralni_rec(1)	1.627	.138	139.181	1	.000	5.086	3.882	6.665
	moraTri_rec(1)	2.275	.111	420.306	1	.000	9.726	7.825	12.088
	Constant	-1.632	.075	478.657	1	.000	.196		

a. Variable(s) entered on step 1: moralni\_rec, moraTri\_rec.

Las OR se redujeron considerablemente, de 20 a 5 para la morosidad inicial y de 21 a 10 para la trimestral, es decir 75% y 52% respectivamente; los intervalos de confianza en ambos casos también se hicieron más pequeños lo que se traduce en estimaciones más precisas. Pude concluir que el tener mora inicial mayor a 30 días multiplica el riesgo por 5 de deteriorar el comportamiento de pago, mientras que tener mora máxima trimestral mayor a 30 días multiplica por 10 el riesgo de convertirse en mal cliente.



**Figura A6.5** - Relación entre mora inicial, mora trimestral y tipo de cliente

Si se incluye la variable morosidad del mes anterior en lugar de la mora máxima trimestral, los resultados son similares: la OR para la mora actual se reduce de 20 a 6 y para la del mes anterior pasa de 13 a 5, es decir un 70% y 62% respectivamente; los intervalos de confianza en ambos casos también se redujeron. Pude concluir entonces que el tener mora actual mayor a 30 días multiplica el riesgo por 6 de deteriorar el comportamiento de pago, mientras que tener una morosidad mayor a 30 días el mes anterior, multiplica por 5 el riesgo de convertirse en mal cliente.

Incluir la variable mora máxima semestral en lugar de la morosidad del mes anterior, arroja resultados muy similares a los obtenidos anteriormente y no vi razón para mayores análisis con otras variables de comportamiento de la morosidad. Los resultados se resumen en la Tabla A6.15, donde muestro el cambio porcentual en las OR de las variables originales, cuando además se introduce en el análisis otra variable presumiblemente confundente o modificadora de efecto.

La variable que representa la morosidad inicial se relaciona con la variable *calificación*, en el sentido que entre los malos clientes existe mayor probabilidad que tengan un atraso mayor a 30 días a que no lo tengan, es decir que esta morosidad mayor a 30 días multiplica por 20 el riesgo de convertirse en un mal cliente para ABC. Resultados parecidos se obtiene con las variables de mora trimestral, semestral o anual; sin embargo la naturaleza del negocio de la cobranza exige la aplicación de acciones rápidas, que no dejen deteriorar la cartera vencida con el pasar del tiempo y por el contrario contribuyan a su normalización y estabilización.



**Tabla A6.15 - Variación de las OR en las variables de comportamiento**

VARIABLES ORIGINALES	VARIABLES INTRODUCIDAS							
	contador1_rec	contador15_rec	contador30_rec	cuotasPag_rec	cuotasPen_rec	moraTri_rec	Porcentaje_rec	vectorMor_rec
contador1_rec	0%	1880%	-83%	3%	-26%	-67%	0%	-82%
contador15_rec	-100%	0%	-31%	-60%	-19%	34%	-37%	-27%
contador30_rec	-74%	824%	0%	-23%	-54%	133%	-21%	-69%
cuotasPag_rec	48%	-68%	6%	0%	-64%	132%	-75%	59%
cuotasPen_rec	-16%	-7%	-45%	404%	0%	38%	-62%	-11%
moralni_rec	0%	0%	0%	0%	0%	-74%	23%	-84%
moraAnt_rec	0%	0%	0%	0%	0%	-97%	32%	-79%
moraTri_rec	-87%	10%	-77%	56%	17%	0%	43%	-87%
moraSem_rec	0%	0%	0%	0%	0%	0%	46%	-92%
moraAño_rec	0%	0%	0%	0%	0%	0%	47%	-86%
Porcentaje_rec	-4%	-57%	-31%	0%	0%	65%	0%	10%
vectorMor_rec	-83%	0%	-83%	9%	-24%	-56%	6%	0%

## ANEXO VII

**COSTOS POR HORA PARA EL CASO DE UN GESTOR DE  
COBRANZA DOMICILIARIA EN LA EMPRESA ABC.**

**Tabla A7.1 - Estructura de costos y factores de productividad para la cobranza domiciliaria en ABC**

<b>Estructura de costos visitas en terreno</b>	<b>VALOR</b>	
<b>Cobranza operativa prejudicial</b>		
<b>Salario bruto anual por categoría</b>		
<b>Personal de operación</b>		
Gestor	3.900	
<b>Personal de estructura de centro</b>		
Coordinador	6.600	
Formador	5.400	
Monitor	5.400	
Administrativo	10.800	
Supervisor	14.400	
Responsable de operaciones	24.000	
<b>Otros factores que inciden en el Costo</b>		
<b>Costo empresa por empleado</b>	37,15%	
<b>Días y horarios especiales</b>		
Plus por hora nocturna	25,0%	
Plus por hora suplementaria	50,0%	
Plus por hora extraordinaria	100,0%	
<b>Estructura de Centro</b>		
Número de gestores por Supervisor	15	
Número de gestores por Formador	50	
Número de gestores por Monitor	50	
Número de gestores por Administrativo	100	
Número de gestores por Jefatura	100	
Número de gestores por Responsable de operaciones	200	
<b>Factores de productividad</b>		
Horas nominales por año		1.680
% de tiempo de descanso a la hora	9,1%	
% de tiempo de formación a la hora	12,0%	
% de ausentismo medio	15,0%	
% de movilización	33,0%	
Tiempo medio de operación objetivo		302
Tasa de efectividad en cobranza de terreno	44,2%	

**Tabla A7.2 - Costo por hora para un ejecutivo de cobranza domiciliaria en ABC**

	VALOR
<b>Estructura de costos visitas en terreno</b>	
<b>Cobranza operativa prejudicial</b>	
<b>Productividad</b>	
Minutos productivos por hora	19
Segundos productivos por hora	1.114
Visitas gestionadas a la hora	4
Cobros por hora	2
<b>Costos</b>	
Costo/Año por categoría	
Ejecutivo de cobranza	5.349
Coordinador	9.052
Formador	7.406
Monitor	7.406
Administrativo	14.812
Supervisor	19.750
Responsable de operaciones	32.916
Costo/Hora del Ejecutivo de cobranza	
Ejecutivo de cobranza	2,7859
Parte propocional de Coordinadores	0,3143
Parte propocional de Formadores	0,0771
Parte propocional de Monitores	0,0771
Parte propocional de Administrativos	0,0771
Parte propocional de Supervisores	0,1029
Parte propocional de Responsable de operaciones	0,0857
Parte propocional de otras categorias	0,0000
<b>Total Costo/hora Ejecutivo de cobranza telefónica</b>	<b>3,52</b>

**Tabla A7.3 - Costos y gastos por hora para un ejecutivo de cobranza domiciliaria en ABC**

<b>Estructura de costos visitas en terreno</b>	
<b>Cobranza operativa prejudicial</b>	
Alquiler de local	0,03
Comunicaciones(teléfono, red)	0,19
Servicios de agua y luz	0,06
Seguros	0,41
Movilización	0,61
Suministros, varios	0,16
<b>Total costos directos por hora</b>	<b>4,98</b>
Gastos generales	0,06
Gastos en investigación y desarrollo	0,84
Márketing	0,16
<b>Total costos y gastos operativos por hora</b>	<b>6,04</b>
Depreciación	0,27
Incobrables	0,08
Gastos financieros	0,08
<b>Total costos y gastos antes de impuestos por hora</b>	<b>6,47</b>
Costo/Segundo productivo del Ejecutivo de cobranza en terreno	0,00581
Costo de la visita	1,75373

## ANEXO VIII

## ANÁLISIS DE CRITICIDAD PARA LA IMPLEMENTACIÓN DEL NUEVO SISTEMA DE COBRANZA EN ABC

**Tabla A8.1 - Matriz de criticidad área de operaciones**

			ÁREA DE OPERACIONES					
AGRUPACIÓN	CATEGORÍAS	CAUSA	IMPACTO EN CLIENTES		IMPACTO EN COSTOS / INGRESOS/ RENTABILIDAD		IMPACTO EN ÁREAS	
Correcta operación	Servicio al cliente	Manual de objeciones	80%	Alta	80%	Alta	50%	Media
Correcta operación	Gestión judicial y extrajudicial	Convenios de pago	70%	Media	75%	Media	50%	Media
Correcta operación	Gestión judicial y extrajudicial	Venta de cartera	50%	Media	75%	Media	75%	Media
Correcta operación	Cobranza Masiva	Gestión IVR	50%	Media	75%	Media	50%	Media
Correcta operación	Cobranza Masiva	Gestión SMS	20%	Baja	75%	Media	30%	Mediana
Correcta operación	Gestión Domiciliaria	Elaboración de cartas	80%	Alta	40%	Mediana	40%	Mediana
Correcta operación	Gestión Domiciliaria y Gestión telefónica	Ingreso de gestiones al sistema	20%	Baja	75%	Media	30%	Mediana
Correcta operación	Gestión Telefónica	Elaboración de guiones	80%	Alta	40%	Mediana	40%	Mediana
Buena planeación	Planificación y reportes	Metas e indicadores	75%	Media	70%	Media	75%	Media
Buena planeación	Planificación y reportes	Elaboración de reportes	75%	Media	70%	Media	75%	Media
Buena planeación	Planificación y reportes	Monitoreos de calidad	80%	Alta	80%	Alta	80%	Alta
Gestión de talento humano	Capacitación	Logística para Capacitación	25%	Mediana	30%	Mediana	20%	Baja
Gestión de talento humano	Capacitación	Manuales y procedimientos	25%	Mediana	20%	Media	75%	Media
Gestión de talento humano	Contrataciones	Perfiles y manual de funciones	25%	Mediana	40%	Mediana	40%	Mediana
Uso de tecnología	Potencializar el sistema informático	Configuración	80%	Alta	80%	Alta	80%	Alta
Uso de tecnología	Potencializar el sistema informático	Desarrollo de Aplicaciones	70%	Media	75%	Media	80%	Alta
Uso de tecnología	Rediseño de la base de datos	Administración de Perfiles de Usuarios	10%	Baja	80%	Alta	10%	Baja
Uso de tecnología	Rediseño de la base de datos	Creación de la tabla de score	90%	Alta	90%	Alta	90%	Alta
Uso de tecnología	Rediseño de la base de datos	Solicitud de reportes	50%	Media	75%	Media	90%	Alta

Tabla A8.2 - Matriz de criticidad área de tecnología

			ÁREA DE TECNOLOGÍA					
AGRUPACIÓN	CATEGORÍAS	CAUSA	IMPACTO EN CLIENTES		IMPACTO EN COSTOS / INGRESOS/ RENTABILIDAD		IMPACTO EN ÁREAS	
Correcta operación	Servicio al cliente	Manual de objeciones	1%	Baja	100%	Alta	50%	Media
Correcta operación	Gestión judicial y extrajudicial	Convenios de pago	100%	Alta	100%	Alta	80%	Alta
Correcta operación	Gestión judicial y extrajudicial	Venta de cartera	100%	Alta	100%	Alta	60%	Media
Correcta operación	Cobranza Masiva	Gestión IVR	1%	Baja	100%	Alta	50%	Media
Correcta operación	Cobranza Masiva	Gestión SMS	1%	Baja	100%	Alta	30%	Mediana
Correcta operación	Gestión Domiciliaria	Elaboración de cartas	100%	Alta	100%	Alta	100%	Alta
Correcta operación	Gestión Domiciliaria y Gestión telefónica	Ingreso de gestiones al sistema	1%	Baja	74%	Media	1%	Baja
Correcta operación	Gestión Telefónica	Elaboración de guiones	100%	Alta	100%	Alta	100%	Alta
Buena planeación	Planificación y reportes	Metas e indicadores	100%	Alta	100%	Alta	74%	Media
Buena planeación	Planificación y reportes	Elaboración de reportes	100%	Alta	100%	Alta	74%	Media
Buena planeación	Planificación y reportes	Monitoreos de calidad	100%	Alta	100%	Alta	74%	Media
Gestión de talento humano	Capacitación	Logística para Capacitación	30%	Mediana	30%	Mediana	25%	Mediana
Gestión de talento humano	Capacitación	Manuales y procedimientos	30%	Baja	20%	Baja	25%	Baja
Gestión de talento humano	Contrataciones	Perfiles y manual de funciones	25%	Mediana	60%	Media	50%	Media
Uso de tecnología	Potencializar el sistema informático	Configuración	100%	Alta	100%	Alta	100%	Alta
Uso de tecnología	Potencializar el sistema informático	Desarrollo de Aplicaciones	75%	Media	75%	Media	85%	Alta
Uso de tecnología	Rediseño de la base de datos	Administración de Perfiles de Usuarios	10%	Baja	55%	Media	75%	Media
Uso de tecnología	Rediseño de la base de datos	Creación de la tabla de score	100%	Alta	100%	Alta	100%	Alta
Uso de tecnología	Rediseño de la base de datos	Solicitud de reportes	15%	Baja	15%	Baja	60%	Media

Tabla A8.3 - Matriz de criticidad área de recursos humanos

			ÁREA DE RECURSOS HUMANOS					
AGRUPACIÓN	CATEGORÍAS	CAUSA	IMPACTO EN SUBSCRIPTORES		IMPACTO EN COSTOS / INGRESOS/ RENTABILIDAD		IMPACTO EN ÁREAS	
Correcta operación	Servicio al cliente	Manual de objeciones	0%	Baja	50%	Media	40%	Mediana
Correcta operación	Gestión judicial y extrajudicial	Convenios de pago	100%	Alta	50%	Media	20%	Baja
Correcta operación	Gestión judicial y extrajudicial	Venta de cartera	100%	Alta	60%	Media	60%	Media
Correcta operación	Cobranza Masiva	Gestión IVR	100%	Alta	60%	Media	50%	Media
Correcta operación	Cobranza Masiva	Gestión SMS	100%	Alta	40%	Mediana	70%	Media
Correcta operación	Gestión Domiciliaria	Elaboración de cartas	100%	Alta	60%	Media	80%	Alta
Correcta operación	Gestión Domiciliaria y Gestión telefónica	Ingreso de gestiones al sistema	100%	Alta	40%	Mediana	70%	Media
Correcta operación	Gestión Telefónica	Elaboración de guiones	100%	Alta	60%	Media	80%	Alta
Buena planeación	Planificación y reportes	Metas e indicadores	80%	Alta	40%	Mediana	50%	Media
Buena planeación	Planificación y reportes	Elaboración de reportes	58%	Media	52%	Media	45%	Mediana
Buena planeación	Planificación y reportes	Monitoreos de calidad	80%	Alta	60%	Media	50%	Media
Gestión de talento humano	Capacitación	Logística para Capacitación	20%	Baja	20%	Baja	20%	Baja
Gestión de talento humano	Capacitación	Manuales y procedimientos	19%	Baja	46%	Mediana	29%	Mediana
Gestión de talento humano	Contrataciones	Perfiles y manual de funciones	60%	Media	40%	Mediana	50%	Media
Uso de tecnología	Potencializar el sistema informático	Configuración	50%	Media	50%	Media	60%	Media
Uso de tecnología	Potencializar el sistema informático	Desarrollo de Aplicaciones	30%	Mediana	50%	Media	60%	Media
Uso de tecnología	Rediseño de la base de datos	Administración de Perfiles de Usuarios	0%	Baja	0%	Baja	0%	Baja
Uso de tecnología	Rediseño de la base de datos	Creación de la tabla de score	100%	Alta	100%	Alta	70%	Media
Uso de tecnología	Rediseño de la base de datos	Solicitud de reportes	30%	Mediana	30%	Mediana	80%	Alta

Tabla A8.4 - Matriz de criticidad área de planeación

			ÁREA DE PLANEACIÓN					
AGRUPACIÓN	CATEGORÍAS	CAUSA	IMPACTO EN CLIENTES		IMPACTO EN COSTOS / INGRESOS/ RENTABILIDAD		IMPACTO EN ÁREAS	
Correcta operación	Servicio al cliente	Manual de objeciones	70%	Media	90%	Alta	70%	Media
Correcta operación	Gestión judicial y extrajudicial	Convenios de pago	90%	Alta	80%	Alta	60%	Media
Correcta operación	Gestión judicial y extrajudicial	Venta de cartera	75%	Media	50%	Media	40%	Mediana
Correcta operación	Cobranza Masiva	Gestión IVR	100%	Alta	80%	Alta	80%	Alta
Correcta operación	Cobranza Masiva	Gestión SMS	70%	Media	80%	Alta	70%	Media
Correcta operación	Gestión Domiciliaria	Elaboración de cartas	100%	Alta	80%	Alta	80%	Alta
Correcta operación	Gestión Domiciliaria y Gestión telefónica	Ingreso de gestiones al sistema	80%	Alta	90%	Alta	60%	Media
Correcta operación	Gestión Telefónica	Elaboración de guiones	100%	Alta	85%	Alta	80%	Alta
Buena planeación	Planificación y reportes	Metas e indicadores	95%	Alta	100%	Alta	70%	Media
Buena planeación	Planificación y reportes	Elaboración de reportes	80%	Alta	40%	Mediana	50%	Media
Buena planeación	Planificación y reportes	Monitoreos de calidad	90%	Alta	90%	Alta	70%	Media
Gestión de talento humano	Capacitación	Logística para Capacitación	20%	Baja	20%	Baja	25%	Mediana
Gestión de talento humano	Capacitación	Manuales y procedimientos	20%	Baja	20%	Baja	20%	Baja
Gestión de talento humano	Contrataciones	Perfiles y manual de funciones	70%	Media	70%	Media	95%	Alta
Uso de tecnología	Potencializar el sistema informático	Configuración	90%	Alta	95%	Alta	90%	Alta
Uso de tecnología	Potencializar el sistema informático	Desarrollo de Aplicaciones	90%	Alta	85%	Alta	90%	Alta
Uso de tecnología	Rediseño de la base de datos	Administración de Perfiles de Usuarios	70%	Media	80%	Alta	80%	Alta
Uso de tecnología	Rediseño de la base de datos	Creación de la tabla de score	90%	Alta	80%	Alta	70%	Media
Uso de tecnología	Rediseño de la base de datos	Solicitud de reportes	70%	Media	85%	Alta	90%	Alta

Tabla A5.5 - Matriz de criticidad del impacto total según expertos

AGRUPACIÓN	CATEGORÍAS	CAUSA	IMPACTO TOTAL SEGÚN EXPERTOS							
			IMPACTO EN CLIENTES		IMPACTO EN COSTOS / INGRESOS / RENTABILIDAD		IMPACTO EN ÁREAS		IMPACTO TOTAL	
Correcta operación	Servicio al cliente	Manual de objeciones	38%	Mediana	80%	Alta	53%	Media	55%	Media
Correcta operación	Gestión judicial y extrajudicial	Convenios de pago	90%	Alta	76%	Alta	53%	Media	75%	Alta
Correcta operación	Gestión judicial y extrajudicial	Venta de cartera	81%	Alta	71%	Media	59%	Media	72%	Media
Correcta operación	Cobranza Masiva	Gestión IVR	63%	Media	79%	Alta	58%	Media	66%	Media
Correcta operación	Cobranza Masiva	Gestión SMS	48%	Mediana	74%	Media	50%	Media	56%	Media
Correcta operación	Gestión Domiciliaria	Elaboración de cartas	95%	Alta	70%	Media	75%	Alta	82%	Alta
Correcta operación	Gestión Domiciliaria y Gestión telefónica	Ingreso de gestiones al sistema	50%	Media	70%	Media	40%	Mediana	53%	Media
Correcta operación	Gestión Telefónica	Elaboración de guiones	95%	Alta	71%	Media	75%	Alta	82%	Alta
Buena planeación	Planificación y reportes	Metas e indicadores	88%	Alta	78%	Alta	67%	Media	78%	Alta
Buena planeación	Planificación y reportes	Elaboración de reportes	78%	Alta	65%	Media	61%	Media	69%	Media
Buena planeación	Planificación y reportes	Monitoreos de calidad	88%	Alta	83%	Alta	69%	Media	80%	Alta
Gestión de talento humano	Capacitación	Logística para Capacitación	24%	Baja	25%	Mediana	23%	Baja	24%	Baja
Gestión de talento humano	Capacitación	Manuales y procedimientos	24%	Baja	26%	Mediana	37%	Mediana	29%	Mediana
Gestión de talento humano	Contrataciones	Perfiles y manual de funciones	45%	Mediana	53%	Media	59%	Media	51%	Media
Uso de tecnología	Potencializar el sistema informático	Configuración	80%	Alta	81%	Alta	83%	Alta	81%	Alta
Uso de tecnología	Potencializar el sistema informático	Desarrollo de Aplicaciones	66%	Media	71%	Media	79%	Alta	72%	Media
Uso de tecnología	Rediseño de la base de datos	Administración de Perfiles de Usuarios	23%	Baja	54%	Media	41%	Mediana	38%	Mediana
Uso de tecnología	Rediseño de la base de datos	Creación de la tabla de score	95%	Alta	93%	Alta	83%	Alta	91%	Alta
Uso de tecnología	Rediseño de la base de datos	Solicitud de reportes	41%	Mediana	51%	Media	80%	Alta	56%	Media



## REFERENCIAS BIBLIOGRAFICAS

- Rodríguez Torres Luis; Microcrédito, La Industria de La Pobreza; Avanzar; Quito; 2006.
- Ferron Aranaz Magdalena; SPSS para Windows, Programación y análisis estadístico, McGraw-Hill, España , 2008.
- Hill Charles; Jones Gareth; Administración Estratégica; McGraw- Hill Bogota; 1997.
- Superintendencia de Bancos; Volumen de Crédito; Quito; 2008, 2009, 2010.
- Brad Cleveland, Call Center Operations Management, Call Center Press, Anápolis USA, 2004 .
- Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors [commentary]. J Natl Cancer Inst 1994;86: 829-835.
- Maxwell S, Delaney H. Bivariate median splits and spurious statistical significance. Psychological Bull 1993;113:181-190.

### Papers

- Josep Fortiana, Enfoque Basado en Distancias de Algunos Métodos Estadísticos Multivariantes, Octubre 30 2001.
- C.M. Cuadras, Métodos Multivariantes Basados en Distancias, Enero 2007.
- Francesc Carmona, Métodos Multivariantes Basados en distancias con R, Departament d'Estadística, Marzo 5 2007.
- Jorge Rodríguez, Edwar Rojas, Clasificación de datos usando el método k-nn, Diciembre 03 2008.