



ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

APLICACIONES DE NUEVAS METODOLOGÍAS PARA EL MONITOREO MULTIVARIADO DEL RENDIMIENTO ESTUDIANTIL UTILIZANDO GRÁFICOS DE CONTROL Y SISTEMAS UMBRAL

**TESIS PREVIA A LA OBTENCIÓN DEL GRADO DE MAGISTER EN
ESTADÍSTICA APLICADA**

GUIDO WASHINGTON SALTOS SEGURA

guidosaltos@gmail.com

DIRECTOR:

ING. MIGUEL ALFONSO FLORES SÁNCHEZ. MSC.

miguel.flores@epn.edu.ec

CODIRECTOR:

MAT. LUIS ALCIDES HORNA HUARACA. PHD.

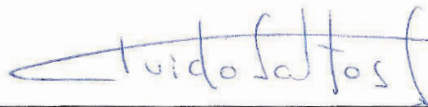
luis.horna@epn.edu.ec

Quito, Abril 2016

DECLARACIÓN

Yo, Guido Washington Saltos Segura, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada 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.



Guido Washington Saltos Segura

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Guido Washington Saltos Segura, bajo mi supervisión.

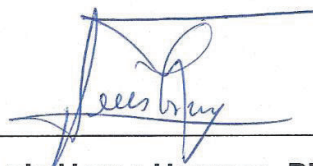


Ing. Miguel Flores Sánchez, MSC.

DIRECTOR

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Guido Washington Saltos Segura, bajo mi supervisión.

A handwritten signature in blue ink, appearing to read "Luis Horna Huaraca", is written over a horizontal line.

Mat. Luis Horna Huaraca, PHD.

CODIRECTOR

AGRADECIMIENTOS

Agradezco a Miguel Flores por su apoyo, asesoría y confianza en este proyecto

A Juan Carlos García por su comprensión y asesoría

A Alejandro Galvis por su apoyo y oportunos comentarios

Y a todos aquellos que de alguna u otra manera ayudaron a sacar adelante este proyecto.

Gracias!!!!

DEDICATORIA

Dedico este trabajo a mis seres queridos: mi Madre, mis dos hijos: Jael y Adrián, mi hermana Mara, los cuales estuvieron cerca en la distancia, o cerca en el corazón, pero cuya presencia fue siempre innegable en mi vida...

RESUMEN

El presente trabajo constituye una aplicación de varios conceptos en la búsqueda de nuevas metodologías para el monitoreo multivariado del rendimiento estudiantil.

Entre los conceptos aplicados se tiene básicamente el de profundidad de datos, así como el uso de una de las cartas de control no paramétricas de las desarrolladas por Regina Liu: La carta r .

Se parte de información proporcionada por una universidad y de la misma se escogen dos variables en determinados período de tiempo en dos contextos distintos:

Primero: notas promedio de varias materias relacionadas con la Estadística y porcentaje de alumnos aprobados de las mismas. Se analizaron datos que van desde el 2008 hasta el 2015 (16 semestres).

Segundo: notas promedio de todas las materias de un grupo de estudiantes que cursaron las mismas materias y porcentaje de aprobación de aquellas. Esto se analizó desde el 2008 hasta que culminaron sus estudios regulares en el 2011.

Con la información proporcionada y aplicando la metodología expuesta se logra reproducir en el campo de la educación lo obtenido por Regina Liu en otro campo: Análisis de Seguridad en la aviación.

Se obtienen gráficos bivariados que sirven para clasificación así como cartas de control r donde se monitorea el desempeño en los dos contextos explicados antes: tanto para notas promedio de materias - porcentaje de alumnos aprobados, como para notas promedio de estudiantes - porcentaje de materias aprobadas.

Finalmente, una parte muy importante del trabajo es que, la metodología utiliza un conjunto de referencia obtenido de los resultados dados, en lugar de los estándares ideales que sólo son característicos de datos atípicos o "outliers". Este conjunto de referencia sirve para calibrar la carta de control y así poder monitorear los datos subsecuentes.

Palabras clave: Profundidad de datos, Cartas de control no paramétricas, Educación.

ABSTRACT

This work is an application of several concepts looking for new methodologies for the multivariate students monitoring.

Among the concepts applied there are principally the idea of data depth and the use of one of the quality control charts developed by Regina Liu: The r – Chart.

Information provided by a university is used, and from it, two variables are chosen using two distinct contexts in determined time periods:

First: average scores of several subjects related with the field of Statistics and percentage of approved students in the same subjects. Data of 16 semesters were analyzed since 2008 until 2015.

Second: average scores of subjects from the whole student's group who studied the same subjects; and the percentage of approved subjects by the same students.

With the information provided and applying the above methodology it is possible to reproduce in the field of education what was obtained by Regina Liu in another field.

Bivariate Graphics are obtained which are used, among other things, for classification and also r charts where performance is monitored using the two contexts exposed above: both subjects average scores-percentage of approved students; and students average scores - approved subjects percentage.

Finally, one important part of the work is that its methodology use a reference set obtained of the results, instead of using ideal standards which only are characteristic of atypical data or outliers. This reference set is used for calibration of the control chart enabling the monitoring of subsequential data.

Keywords: Data depth, Nonparametric control charts, Education.

INDICE DE CONTENIDO

Resumen	vii
Abstract.....	viii
Lista de figuras	xii
Lista de tablas	xvi
1. Introducción.....	1
2. Marco teórico.....	4
2.1. Profundidad de datos	4
2.1.1. Aspectos generales.	4
2.1.2. Propiedades de profundidades.....	5
2.1.3 funciones de profundidad	6
2.2 Gráficos de control para procesos multivariados.....	8
2.2.1 Aspectos generales.	8
2.2.2 Estadísticos derivados de la profundidad de datos.....	9
2.2.3 Cartas de control basadas en profundidad de datos.....	10
2.2.4 Simulación de cartas de control	15
2.2.5 Sistema umbral.....	18
3. Metodología.....	21
3.1 Datos a usarse	21
3.1.1 Limitaciones	21
3.1.2 Datos representativos	23
3.1.3 Sistema umbral.....	24
3.1.4 Splicación de la carta r	25
4. Resultados y discusiones	27
4.1 Materias	27
4.1.1 Estadística descriptiva de notas por materia	27
4.1.2 Comportamiento de las notas promedio a lo largo del tiempo.....	28
4.1.3 Estadística descriptiva del porcentaje de aprobación de alumnos.....	29
4.1.4 Comportamiento del porcentaje de aprobación de alumnos a lo largo del tiempo	29
4.1.5 Clúster de desempeño medio.....	30
4.1.6 Comparación de la materia en particular con el clúster medio	31

4.1.7 Carta r	32
4.1.8 Clasificación del desempeño de las 7 materias	33
4.2 Estudiantes	37
4.2.1 Estadística descriptiva de notas por estudiante	37
4.2.2 Comportamiento de las notas promedio a lo largo del tiempo.....	38
4.2.3 Clúster de desempeño medio.....	39
4.2.4 Comparación del desempeño del estudiante en particular con el clúster medio .	40
4.2.5 Carta r del estudiante.....	40
4.2.6 Clasificación del desempeño de los 15 estudiantes.....	41
4.3 Carta r. Puntos justo en la envolvente convexa	45
5. Conclusiones y recomendaciones.....	47
5.1 Conclusiones.....	47
5.2 Recomendaciones	49
Referencias	50
Anexos.....	52
Anexo 1. Código r de cartas: r, s y q (simulación)	52
Anexo 2. código r de materias: notas promedio-porcentaje de aprobacion de alumnos...	57
Anexo 3. Codigo r de estudiantes: notas promedio-porcentaje de aprobacion de materias	64
Anexo 4. Código r de funciones.	71
Anexo 5. Código r de objeto: npqed.....	73
Anexo 6. Codigo r de objeto: npqcs	77
Anexo 7. Codigo r de objeto: npqcs.r.r.....	87
Anexo 8. Estadística descriptiva de notas promedio de materia.....	92
Anexo 9. Estadísticos descriptivos de notas de materias.....	94
Anexo 10. Estadística descriptiva de porcentaje de aprobacion de alumnos por materia.	95
Anexo 11. Estadísticos descriptivos de porcentaje de aprobación de alumnos.	97
Anexo 12. Comportamiento en el tiempo de las notas promedio por materia.....	98
Anexo 13. Comportamiento en el tiempo del porcentaje de aprobación de alumnos por materia	100
Anexo 14. Materias. Gráficos bivariados de notas - porcentaje de aprobación de alumnos	102
Anexo 15. Materias. Cartas de control r	104
Anexo 16. Estadística descriptiva de notas promedio de diversos estudiantes.....	106

Anexo 17. Estadísticos descriptivos de notas promedio de varios estudiantes.....	110
Anexo 18. Estudiantes. Estadística descriptiva de porcentaje de materias aprobadas...	112
Anexo 19. Estudiantes. Estadísticos descriptivos de porcentaje de materias aprobadas.	115
Anexo 20. Comportamiento en el tiempo de las notas promedio de los estudiantes.....	117
Anexo 21. Comportamiento en el tiempo del porcentaje de materias aprobadas.....	121
Anexo 22. Estudiantes. Gráficos bivariados de notas – porcentaje de aprobación de materias.....	125
Anexo 23. Estudiantes. Cartas de control r.....	133
Anexo 24. Datos originales.....	137

LISTA DE FIGURAS

Figura 1. Carta r.....	16
Figura 2. Carta Q, n=4.....	16
Figura 3. Carta Q, n=10.....	17
Figura 4. Carta S.....	17
Figura 5. Carta S*.....	18
Figura 6. Gráfico bivariado de la materia: AES7.....	25
Figura 7. Carta r de AES7.....	26
Figura 8. Estadística Descriptiva de AES2.....	27
Figura 9. Comportamiento de Notas Promedio por semestre de AES2.....	28
Figura 10. Estadística Descriptiva del Porcentaje de Aprobación de Alumnos en AES2.....	29
Figura 11. Porcentaje de Aprobación de Alumnos por semestre. AES2.....	30
Figura 12. Clúster de Referencia. Todas las materias.....	31
Figura 13. Gráfico Bivariado de AES2.....	32
Figura 14. Carta r. AES2.....	33
Figura 15. Clúster esperado con la clasificación de desempeños de las demás materias.....	33
Figura 16. Estadística Descriptiva de Notas promedio del Estudiante 7091.....	38
Figura 17. Comportamiento de las notas promedio del estudiante 7091 a lo largo del tiempo.....	39
Figura 18. Clúster de referencia o esperado de los 15 estudiantes.....	39
Figura 19. Comparación del desempeño del estudiante 7091 con el clúster medio.....	40
Figura 20. Carta r del Estudiante 7091.....	41
Figura 21. Clúster esperado con la clasificación de desempeños de los demás estudiantes.....	41
Figura 22. Estudiante 6054. Clúster Esperado y Puntos sobre la envolvente convexa.....	45
Figura 23. Carta r del Estudiante 6054.....	46
Figura A.8. 1 AES1.....	92
Figura A.8. 2 AES2.....	92
Figura A.8. 3 AES3.....	92
Figura A.8. 4 AES4.....	92
Figura A.8. 5 AES5.....	93
Figura A.8. 6 AES6.....	93
Figura A.8. 7 AES7.....	93
Figura A.10. 1 PAES1.....	95
Figura A.10. 2 PAES2.....	95
Figura A.10. 3 PAES3.....	95
Figura A.10. 4 PAES4.....	95
Figura A.10. 5 PAES5.....	96
Figura A.10. 6 PAES6.....	96
Figura A.10. 7 PAES7.....	96
Figura A.12. 1 SM-AES2.....	98
Figura A.12. 2 SM-AES1.....	98
Figura A.12. 3 SM-AES4.....	98

Figura A.12. 4 SM-AES3	98
Figura A.12. 6 SM-AES6	99
Figura A.12. 5 SM-AES5	99
Figura A.12. 7 SM-AES7	99
Figura A.13. 1 SP-AES1	100
Figura A.13. 2 SP-AES2	100
Figura A.13. 3 SP-AES3	100
Figura A.13. 4 SP-AES4	100
Figura A.13. 5 SP-AES6	101
Figura A.13. 6 SP-AES5	101
Figura A.13. 7 SP-AES7	101
Figura A.14. 1 B-AES2	102
Figura A.14. 2 B-AES1	102
Figura A.14. 3 B-AES4	102
Figura A.14. 4 B-AES3	102
Figura A.14. 5 B-AES5	103
Figura A.14. 6 B-AES6	103
Figura A.14. 7 B-AES7	103
Figura A.15 1 r-AES2.....	104
Figura A.15 2 r-AES1.....	104
Figura A.15 3 r-AES4.....	104
Figura A.15 4 r-AES3.....	104
Figura A.15 5 r-AES6.....	105
Figura A.15 6 r-AES5.....	105
Figura A.15 7 r-AES7.....	105
Figura A.16. 1 MEDA-6066.....	106
Figura A.16. 2 MEDA-6054.....	106
Figura A.16. 3 MEDA-6202.....	106
Figura A.16. 4 MEDA-6166.....	106
Figura A.16. 5 MEDA-6450.....	107
Figura A.16. 6 MEDA-6515.....	107
Figura A.16. 7 MEDA-7059.....	107
Figura A.16. 8 MEDA-6481.....	107
Figura A.16. 9 MEDA-7077.....	108
Figura A.16. 10 MEDA-7060.....	108
Figura A.16. 11 MEDA-7080.....	108
Figura A.16. 12 MEDA-7091.....	108
Figura A.16. 13 MEDA-8427.....	109
Figura A.16. 14MEDA-8426.....	109
Figura A.16. 15 MEDA-8509.....	109
Figura A.18. 1 PEDA-6202.....	112
Figura A.18. 2 PEDA-6166.....	112
Figura A.18. 3 PEDA-6450.....	112
Figura A.18. 4 PEDA-6515.....	112
Figura A.18. 5 PEDA-7059.....	113

Figura A.18. 6 PEDA-6841	113
Figura A.18. 7 PEDA-7060	113
Figura A.18. 8 PEDA-7077	113
Figura A.18. 9 PEDA-7091	114
Figura A.18. 10 PEDA-7080	114
Figura A.18. 11 PEDA-8426	114
Figura A.18. 12 PEDA-8509	114
Figura A.20. 1 MSERIE-6166	117
Figura A.20. 2 MSERIE-6054	117
Figura A.20. 3 MSERIE-6202	117
Figura A.20. 4 MSERIE-6066	117
Figura A.20. 5 MSERIE-6450	118
Figura A.20. 6 MSERIE-6515	118
Figura A.20. 7 MSERIE 6841	118
Figura A.20. 8 MSERIE 7059	118
Figura A.20. 9 MSERIE 7060	119
Figura A.20. 10 MSERIE 7077	119
Figura A.20. 11 MSERIE-7080	119
Figura A.20. 12 MSERIE-7091	119
Figura A.20. 13 MSERIE-8426	120
Figura A.20. 14 MSERIE-8427	120
Figura A.20. 15 MSERIE-8509	120
Figura A.21. 1 PSERIE-6054	121
Figura A.21. 2 PSERIE-6066	121
Figura A.21. 3 PSERIE-6166	121
Figura A.21. 4 PSERIE-6202	121
Figura A.21. 5 PSERIE-6515	122
Figura A.21. 6 PSERIE-6450	122
Figura A.21. 7 PSERIE-7059	122
Figura A.21. 8 PSERIE-6841	122
Figura A.21. 9 PSERIE-7077	123
Figura A.21. 10 PSERIE-7060	123
Figura A.21. 11 PSERIE-7080	123
Figura A.21. 12 PSERIE-7091	123
Figura A.21. 14 PSERIE-8509	124
Figura A.21. 13 PSERIE-8426	124
Figura A.22. 1 B6054	125
Figura A.22. 2 B6066	125
Figura A.22. 3 B6166	126
Figura A.22. 4 B6202	126
Figura A.22. 5 B6450	127
Figura A.22. 6 B6515	127
Figura A.22. 7 B6841	128
Figura A.22. 8 B7059	128
Figura A.22. 9 B7060	129

Figura A.22. 10 B7077	129
Figura A.22. 11 B7080	130
Figura A.22. 12 B7091	130
Figura A.22. 13 B8426	131
Figura A.22. 14 B8427	131
Figura A.22. 15 B8509	132
Figura A.23. 1 C6066	133
Figura A.23. 2 C6054	133
Figura A.23. 3 C6202	133
Figura A.23. 4 C6166	133
Figura A.23. 5 C6450	134
Figura A.23. 6 C6515	134
Figura A.23. 7 C7059	134
Figura A.23. 8 C6841	134
Figura A.23. 9 C7077	135
Figura A.23. 10 C7060	135
Figura A.23. 11 C7091	135
Figura A.23. 12 C7080	135
Figura A.23. 14 C8427	136
Figura A.23. 13 C8426	136
Figura A.23. 15 C8509	136

LISTA DE TABLAS

Tabla 1. Materias	22
Tabla 2. Siglas de Estudiantes	22
Tabla 3. Comparación Total de Materias	34
Tabla 4. Comparación de Materias	35
Tabla 5. Plan de Mejoramiento de Materias.....	36
Tabla 6. Plan de Mejora de Materias para los grupos de Advertencia	37
Tabla 7. Comparación Total de Estudiantes.....	42
Tabla 8. Comparación de 3 Estudiantes	43
Tabla 9. Plan de Mejoramiento de Estudiantes	44
Tabla 10. Plan de Mejora de Estudiantes para los grupos de Advertencia.....	44
Tabla 24. 1 Materias	137
Tabla 24. 2 Estudiantes.....	141

1. INTRODUCCIÓN

En la actualidad y como se lo ha hecho tradicionalmente, el rendimiento estudiantil universitario se lo mide por las notas académicas que obtiene un estudiante; vale mencionar que a las notas se les ha considerado un reflejo de los logros en los diferentes componentes del aprendizaje que incluye: aspectos personales, académicos y sociales. (Garbanzo, 2013).

No obstante, desde hace algunos años hasta ahora se han producido varios trabajos con enfoques diferentes, por ejemplo:

Montes y Lerner¹ (2010-2011) conceptúan el rendimiento académico como una función de 5 dimensiones, por decirlo así. Usan la herramienta de Análisis Factorial para definir variables latentes o no observables. Las dimensiones que consideran en su estudio son: dimensión académica, dimensión económica, dimensión familiar, dimensión personal, dimensión institucional. En su estudio, de alguna manera explican el rendimiento académico como el resultado de la interacción de los distintos elementos involucrados en cada una de las dimensiones mencionadas antes. El producto final sirve para de alguna manera explicar lo logrado por los individuos (estudiantes) durante el período de estudio.

Otro tipo de trabajos como el de Barahona (2014), han tratado de determinar las variables predictoras del rendimiento académico –en esta etapa se usó regresión múltiple- y posteriormente mediante regresión logística se determinaron ciertas variables que incidían en la probabilidad de mejorar el rendimiento académico del estudiante. Con este tipo de trabajos, ya se plantea la posibilidad de predecir o no si un estudiante aprobará o no.

Dentro de los modelos de predicción, tenemos por ejemplo el trabajo de Montero, Villalobos y Valverde (2007), con el uso de regresión multinivel para predecir calificaciones finales en cursos de carrera de determinadas muestras

¹ Montes I., Lerner J. (2010-2011). Rendimiento Académico de los Estudiantes de Pregrado de la Universidad EAFIT. Perspectiva cuantitativa. <http://www.eafit.edu.co>.

estratificadas. En este trabajo las variables independientes que se consideraron obedecían a dimensiones: institucionales, sociodemográficas, psicosociales y pedagógicas y la variable dependiente: la nota obtenida.

En otro tipo de trabajos se ha usado un enfoque econométrico, Ferreyra (2007) mediante el uso de cuantiles de regresión censurados y modelos Tobit estudia la influencia de: las características personales, entorno familiar, escuela secundaria y la etapa inicial de la vida universitaria en el rendimiento académico de distintos tipos de estudiantes.

Como lo dicen Montes y Lerner en su trabajo (2010-2011) el rendimiento académico es complejo en su definición y forma de abordarlo, además se modifica acorde al objetivo del estudio y el enfoque y puede ser amplio o limitado, analizarse cualitativa o cuantitativamente o ambas. Por esa razón se ha decidido enfocar este trabajo desde una perspectiva distinta: El Control de Calidad.

Los trabajos hechos antes permiten comprender los factores que influyen en el rendimiento estudiantil o dependiendo del enfoque, de alguna manera -lo que se diría en Control de Calidad- nos podrían dar una “alerta temprana” acerca del rendimiento predicho de un estudiante. El presente trabajo pretende mostrar el monitoreo del rendimiento estudiantil a lo largo del tiempo, y como se estila en control de calidad, también se apreciarán señales tempranas que indique que el resultado del rendimiento está saliendo de los límites aceptables de control, permitiendo tomar las acciones correctivas a tiempo.

El presente trabajo muestra el monitoreo de dos variables (bivariado). Se usará este enfoque para poder aplicar el concepto de Profundidad de Datos Simplicial (Enfoque no paramétrico), ya que este trabajo es una aplicación de la metodología expuesta por Cheng, Liu y Luxhoj en su publicación: “Monitoring multivariate aviation safety data by data depth: control charts and thresholds systems”, (2000) en el cual se usan principalmente dos variables. Además, posteriores estudios han mostrado que la profundidad Simplicial es muy complicada de trabajar con muchas dimensiones. (Casco, 2011).

Los datos a usarse provienen de una universidad la cual, para poder

proporcionarlos ha solicitado expresamente que no se mencione su origen. Se trabajará con notas y estudiantes de varias materias del área de estadística.

Dentro de la metodología usada, se selecciona un grupo de desempeño promedio y es con respecto a ese grupo que se comparan los desempeños individuales.

Al delimitar un desempeño medio esperado, se contará con especificaciones realistas de lo que pueden lograr los estudiantes medios. Es decir no se compararán los resultados con valores ideales como son notas máximas y perfectas, pues en la práctica no observamos nunca registros perfectos sino como observaciones de tipo “outlier” o atípicas.

2. MARCO TEÓRICO

2.1. PROFUNDIDAD DE DATOS

2.1.1. ASPECTOS GENERALES.

El Concepto de Profundidad, desarrollado primeramente por J. W. Tukey y posteriormente popularizado con el concepto de profundidad simplicial por Regina Liu, (Liu, 1988), ha demostrado ser una cantera muy fértil originando aplicaciones multivariante diversas como son: a) el Bagplot (originalmente desarrollado por Tukey como diagrama de caja o boxplot y luego generalizado por Rousseeuw, Ruts y el mismo Tukey como bagplot), b) Bondad de Ajuste, c) DD-plots, d) Análisis Discriminante, e) Análisis por Conglomerados, f) Regresión, g) Medidas de Riesgo, h) Control de Calidad, (Liu, 1995). Es en este último campo en el cual se centra el presente trabajo.

En el análisis multivariante, el término profundidad se refiere al grado de centralidad de un punto con respecto a una nube de puntos o una distribución de probabilidad. Si tenemos una muestra de datos univariantes, el punto más profundo constituye una medida de tendencia central de la misma y a medida que se consideran puntos más alejados de ese “centro” su profundidad disminuye. Un ejemplo de dato más profundo con datos univariantes sería la mediana, en donde la profundidad de los demás datos se podría establecer a partir del cuantil que representaría cada punto respecto de la distribución de probabilidad empírica o también por el orden del estadístico ordenado más cercano, (Casco, 2011).

Si tenemos una nube de puntos multivariante, se puede tomar como punto más profundo a una generalización multivariante de la mediana. Así la profundidad disminuirá a lo largo de cada uno de los rayos que surgen del punto más profundo.

Partiendo del concepto de profundidad y de una distribución de probabilidad, es posible definir un orden en el espacio euclídeo multidimensional. Este orden permite extender las técnicas de análisis de datos univariantes basadas en el orden natural al campo multivariante, (Casco y otros, 2011).

2.1.2. PROPIEDADES DE PROFUNDIDADES.

Según Dyckerhoff ,(Dyckerhoff, 2004) y expuestas por Cascos (Cascos, 2011) la definición y propiedades de la función de profundidad son las siguientes:

Definición 1:

Si se fija una distribución P en R^d , una función de profundidad es una función acotada $D_p : R^d \rightarrow R$ que asigna a cada punto de R^d , su grado de centralidad respecto de P .

Propiedades:

D1: Invarianza afín:

Para cualquier matriz no singular $A \in R^{d \times d}$ y $b \in R^d$, se tiene $D(Ax + b; P_{AX+b}) = D(x; P_X)$ donde P_X denota la probabilidad inducida por el vector aleatorio X

D2: Anulación en el infinito:

$$\lim_{\|x\| \rightarrow \infty} D(x; P) = 0$$

D3: Semicontinuidad superior:

El conjunto de puntos $\{x \in R^d : D(x; P) \geq \alpha\}$ cuya profundidad es mínimo α es cerrado para cualquier α ,

D4: Monotonía respecto del punto más profundo:

Si θ es un punto de máxima profundidad, entonces $D(x; P) \leq D(\theta + \lambda(x - \theta); P)$ para cualquier $0 \leq \lambda \leq 1$

D4': Cuasiconcavidad

Una función de profundidad puede satisfacer una propiedad más restrictiva que D4:

$$D(\lambda x + (1 - \lambda)y; P) \geq \min\{D(x; P), D(y; P)\} \text{ para cualquier } x, y \in R^d \text{ y } 0 \leq \lambda \leq 1.$$

Esta propiedad establece que los conjuntos de puntos cuya profundidad es al menos un cierto valor son convexos.

Simetría angular:

Una distribución tiene un centro claro cuando satisface algún concepto de simetría respecto del centro. Sea P una distribución de probabilidad, ésta será angularmente simétrica en torno a θ si para una observación X de la misma, los vectores aleatorios siguientes siguen la misma distribución: $(X - \theta) / \|X - \theta\|$ y $-(X - \theta) / \|X - \theta\|$.

2.1.3 FUNCIONES DE PROFUNDIDAD

Existen muchas funciones de profundidad y entre las más conocidas se tiene:

MP1. Profundidad Simplicial (Liu, 1990)

MP2. Profundidad de Mahalanobis (Mahalanobis 1936),(Cascos, 2011)

MP3. Profundidad Semiespacial (Half-Space) o de Tukey (Hodges, 1955, Tukey, 1975), (Cascos, 2011)

MP4. Profundidad de Verosimilitud (Likelihood Depth), (Fraiman y Meloche, 1996), (Liu, 1999)

MP5. Profundidad de las Proyecciones (Random Projection Depth) (Zuo y Serfling, 2000) (Cascos, 2011)

MP6. Profundidad de la Mayoría (Majority Depth) (Singh, 1991), (Cascos,2011)

MP7. Profundidad de peladura de la cáscara convexa (Convex hull pelling depth), (Barnett, 1976), (Liu, 1999).

MP8. Profundidad de Oja (Oja Depth) (Oja, 1983), (Liu, 1999),

MP9. Profundidad del Zonoide, (Koshevoy y Mosler, 1997), (Cascos, 2011)

MP10. Profundidad asociada a la Mediana.(Vardi y Zhang, 2000), (Cascos, 2011)

MP11. Profundidad por Bandas (López-Pintado y Romo, 2009), (Cascos 2011).

Las cinco primeras funciones de profundidad (MP1-MP5) están incorporadas en el paquete `fda.usc` de R.

Para este trabajo, el cual es una aplicación de lo expuesto por Liu (Liu, 2000) en su artículo: "Monitoring multivariate aviation safety data by data depth: control charts

and thresohlds systems” se ha escogido la Definición de Profundidad Simplicial, por lo que se analizará con más detalle:

MP1. Profundidad Simplicial:

Liu (Liu, 1990) definió la profundidad Simplicial como la probabilidad de que un punto esté en el simplex² cuyos vértices son d+1 observaciones independientes de una distribución P,

$$SD(x; P) = \Pr\{x \in co\{X_1, X_2, \dots, X_{d+1}\}\} \quad (1)$$

En donde: X_1, X_2, \dots, X_{d+1} son las observaciones independientes de P y “co” representa la envolvente convexa.

La versión muestral de SD , se obtiene reemplazando P por su estimado muestral P_n (Distribución empírica de la muestra), o calculando la fracción de los simples aleatorios de la muestra que contienen al punto x .

Si SD es la porción de simples que contienen al punto x dentro de todos aquellos cuyos vértices son d+1 observaciones distintas de la muestra, entonces:

$$SD = \left(\binom{n}{d+1} \right)^{-1} \sum_{1 \leq i_1 < \dots < i_{d+1} \leq n} I(x \in co\{X_{i_1}, X_{i_2}, \dots, X_{i_{d+1}}\}) \quad (2)$$

En donde $I(A) = 1$ si el evento A ocurre y será cero en caso contrario. La función

$I(A)$ se llama función indicadora. Por otra parte, La expresión: $\binom{n}{d+1}$

corresponde al coeficiente binomial.

Según Liu (Liu, 1990) Para saber si un punto X en R^d está o no dentro del simplex $co\{X_1, \dots, X_{d+1}\}$, es necesario resolver el siguiente sistema de ecuaciones:

$$X = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_{d+1} X_{d+1} \quad ; \quad \alpha_1 + \alpha_2 + \dots + \alpha_{d+1} = 1 \quad (2.a)$$

² Un simplex o n-simplex, o hipertetraedro es la generalización de una región tetrahedral de espacio a n dimensiones. Es el análogo de n dimensiones de un triángulo. Es la envoltura convexa de un conjunto de (n+1) puntos independientes afines en un espacio euclídeo de dimensión n o mayor. Así un 0-simplex es un punto; un 1-simplex es un segmento; un 2-simplex es un triángulo; un 3-simplex es un tetraedro, etc.

Entonces para un simplex no degenerado, el sistema (2.a) de $d + 1$ ecuaciones con $d + 1$ incógnitas: $\alpha_1, \alpha_2, \dots, \alpha_{d+1}$ tiene una solución única, y X estará dentro del simplex si y sólo si $\alpha_1, \alpha_2, \dots, \alpha_{d+1}$ son todas positivas. (Ver Liu, 1990).

Según Cascos, la Profundidad Simplicial cumple las propiedades D1-D3. (Cascos, 2011)

Acorde al citado autor, la profundidad Simplicial no es muy usada ya que las regiones inducidas por ella no son convexas y no es sencillo trabajar con ellas. No obstante, vale mencionar que el trabajo de Regina Liu fue el que impulsó todo el trabajo posterior de profundidad de datos y se usará su definición de profundidad para reproducir unos ejemplos de uno de sus artículos (Liu, 1995) en una parte posterior de este trabajo, así como para realizar ciertos análisis en dos dimensiones.

2.2 GRÁFICOS DE CONTROL PARA PROCESOS MULTIVARIADOS

2.2.1 ASPECTOS GENERALES.

Los gráficos o cartas de control son herramientas muy usadas en la industria en general. Con límites de control adecuadamente seleccionados una carta de control puede detectar un desplazamiento de una distribución de calidad “buena” a una “mala”.

Las cartas de control más usadas en la práctica son: La carta X , La carta \bar{X} (o de Shewart) y la carta de Suma Acumulada o *CUSUM*. Estas cartas, no obstante que son útiles en la práctica, tienen las siguientes limitantes:

Son útiles cuando la observación X es univariante y se debe asumir normalidad de X . En la práctica se tienen varias características de calidad simultáneamente y están correlacionadas y por otra parte no siempre son normales.

Por lo expuesto antes es necesario disponer de cartas de control que monitoreen medidas multivariadas de manera directa.

Las carta de Control que se usará es una de las propuestas por Regina Liu (Liu, 1995), la carta r . La idea principal bajo sus cartas de control es reducir cada

medida multivariante a un índice univariado, o mejor dicho a su ordenamiento relativo del centro hacia afuera inducido por profundidad de datos. Este planteo es completamente no paramétrico, y así las cartas resultantes son válidas sin suposiciones paramétricas en el modelo del proceso. Estas cartas permiten detectar simultáneamente el cambio en localización (desplazamiento de la media) y el incremento de escala (cambio en variabilidad) en un proceso. Liu desarrolló las cartas r , Q , y S , las cuales pueden ser vistas como generalizaciones multivariadas basadas en profundidad de datos de las cartas univariadas X , \bar{X} , y $CUSUM$ respectivamente.

2.2.2 ESTADÍSTICOS DERIVADOS DE LA PROFUNDIDAD DE DATOS

Si k , ($k \geq 1$) son las características de cada producto usadas para determinar la calidad del mismo. El proceso estará bajo control si las mediciones siguen una distribución de calidad determinada (requerida por el usuario o ingenieros de diseño).

Sea G la distribución k -dimensional, y Y_1, \dots, Y_m ; m observaciones aleatorias de G . La mencionada muestra es una muestra de referencia en el contexto de control de calidad y se considera como medidas de los productos producidos en un proceso bajo control. Sean X_1, X_2, \dots las nuevas observaciones provenientes del proceso de manufactura. Se asume que las $X'_{i,s}$ siguen una distribución F . Ahora bien, basado en las observaciones $X'_{i,s}$, se quiere determinar si la calidad del producto se ha deteriorado o si el proceso está fuera de control. Esto ocurriría si las $X'_{i,s}$ no siguen la distribución $G(\bullet)$ en cierto sentido. Por tanto, hace falta comparar F con G .

Se usará la definición de Profundidad Simplicial para los cálculos demostrativos de las cartas de control r , S y Q , (véase la sección 2.2.4).

La profundidad de datos induce un ordenamiento del centro hacia afuera de los puntos de la muestra si se calcula la profundidad de todos los puntos y se los compara.

Si se ordenan todas las profundidades $D_G(Y'_i)$ en orden ascendente y se usa $Y_{[j]}$ para denominar al punto de la muestra asociado con el j -ésimo valor de profundidad

más pequeño, entonces $Y_{[1]}, Y_{[2]}, \dots, Y_{[m]}$ son los estadísticos de orden de los Y_i 's, siendo $Y_{[m]}$ el punto más central. Mientras más pequeño el índice (o rango) de un punto, más externo será ese punto respecto de la distribución subyacente $G(\cdot)$.

Sea $Y \sim G$ para indicar que la variable aleatoria Y sigue la distribución G , luego Liu (Liu, 1995) define: el rango (Rank)

$$r_G(y) = P\{D_G(Y) \leq D_G(y) | Y \sim G\} \quad (3)$$

Y para el caso en que G es desconocida y sólo se conoce la muestra $\{Y_1, \dots, Y_m\}$, define:

$$r_{G_m}(y) = \#\{Y_j | D_{G_m}(Y_j) \leq D_{G_m}(y), j = 1, \dots, m\} / m \quad (4)$$

Si $F_n(\cdot)$ representa la distribución empírica de la muestra $\{X_1, \dots, X_n\}$, entonces en el mismo trabajo, Liu define, (Liu, 1995).

Índice Q:

$$Q(G, F) = P\{D_G(Y) \leq D_G(X) | Y \sim G, X \sim F\} = E_F[r_G(X)] \quad (5)$$

$$Q(G, F_n) = \frac{1}{n} \sum_{i=1}^n r_G(X_i) \quad (6)$$

y

$$Q(G_m, F_n) = \frac{1}{n} \sum_{i=1}^n r_{G_m}(X_i) \quad (7)$$

2.2.3 CARTAS DE CONTROL BASADAS EN PROFUNDIDAD DE DATOS

CC1. CARTA r

Según Liu (Liu, 1995) la carta en mención se construye de la siguiente manera:

Se calcula $\{r_G(X_1), r_G(X_2), \dots\}$ o $r_{G_m}(X_1), r_{G_m}(X_2), \dots$ si sólo está disponible Y_1, \dots, Y_m pero no G , usando las fórmulas (3) o (4). (En el presente trabajo se usará la expresión (4)).

La carta r es el gráfico de $r_G(X_i)$'s o $r_{G_m}(X_i)$'s versus el tiempo i , con la línea central

$CL = 0,5$ y el límite de control α . El proceso es declarado fuera de control si $r_G(\bullet)$ cae bajo α . Este último valor es la tasa de falsa alarma, la cual se pone cercana a cero, de manera que la carta r tiene un límite inferior $LCL = \alpha$ pero no límite superior UCL .

$r_{G_m}(X)$ acorde a la expresión (4) indica cuan externo es X con respecto a la nube de datos Y_i 's. Un valor muy pequeño de $r_{G_m}(X)$ significa que sólo una muy pequeña proporción de Y_i 's son más externos que X . Así X está en la capa externa y no es conforme con la mayoría de la parte central del conjunto de datos "buenos" o de referencia.

Si se asume que $X \sim F$, entonces un valor pequeño de $r_{G_m}(\bullet)$ sugiere una posible desviación desde G a F .

Puesto que $r_{G_m}(\bullet)$ se define acorde a la profundidad de datos, la posible desviación puede ser un desplazamiento en "centro" y/o un incremento en escala.

Por tanto, según Liu (Liu, 1995) la carta r con $LCL = \alpha$ correspondería a una prueba de hipótesis de nivel α , siendo sus hipótesis respectivas las siguientes:

$$H_0 : F = G$$

$$H_a : \text{Hay un desplazamiento de localización y/o un incremento de escala desde } G \text{ a } F \quad (8)$$

En la carta de control r , H_0 es rechazada cuando una observación cae bajo α

A continuación se muestra una proposición establecida por Liu en su trabajo de 1995 (Liu, 1995), la cual permite justificar la escogencia de los límites mencionados para CL y $LCL = \alpha$.

Proposición 1:

Se asume que $F = G$ y que $X \sim F$, además $U[0,1]$ es una distribución uniforme en el intervalo $[0,1]$. Entonces si $D_G(X)$ tiene una distribución continua, se cumple:

- a. $r_G(X) \sim U[0,1]$ y
- b. A medida que $m \rightarrow \infty, r_{G_m}(X) \rightarrow^L U[0,1]$ a lo largo de casi todas las secuencias $\{Y_1, \dots, Y_m\}$ siempre que $D_{G_m}(\bullet)$ converja a $D_G(\bullet)$ uniformemente cuando $m \rightarrow \infty$

En el mismo trabajo, Liu menciona que la convergencia uniforme de $D_{G_m}(\bullet)$ se mantiene para la profundidad Simplicial si G es absolutamente continua.

Ahora bien, bajo la hipótesis $H_0 : F = G$, la Proposición 1 implica que el valor esperado de $r_G(X)$ es 0,5 y que el de $r_{G_m}(X)$ es 0,5 casi con seguridad para todas las secuencias $\{Y_1, \dots, Y_m\}$ para un m grande. Esto justificaría el escoger $CL = 0.5$ para la carta r .

Cuando $r_G(X)$ o $r_{G_m}(X)$ es mucho más pequeño que 0,5; habría duda para H_0 y evidencia para aceptar H_a mostrando un posible deterioro en calidad. Por otra parte cuando $r_G(X)$ o $r_{G_m}(X)$ es mayor a 0,5; hay indicios de una disminución en escala con quizás un despreciable desplazamiento de localización. Esto es visto como una mejora en calidad dada una ganancia en la precisión, y así el proceso no se vería como fuera de control. Luego, hay sólo una LCL en la carta r . La distribución uniforme de $r_G(X)$ o $r_{G_m}(X)$ implica que LCL debería ser α .

Aunque la carta r no dispone de la línea UCL para hacer de CL la línea central de la región bajo control, la línea CL sirve como punto de referencia que permite observar si aparece o se está gestando un patrón o tendencia en una secuencia de muestras.

³ \rightarrow^L Significa convergencia en ley

CC2. CARTA Q

La carta Q se basa en la representación de promedios de subconjuntos de $r_G(X_i)_s$ o $r_{G_m}(X_i)_s$

Es decir se asume que cada subconjunto tiene tamaño n . La idea detrás de la carta Q es similar a la de carta univariada \bar{X} . Si X_1, X_2, \dots son univariadas y G es normal, la carta \bar{X} dibuja los promedios de subconjuntos consecutivos de las X_i 's.

Una característica de la carta \bar{X} es que puede prevenir una falsa alarma cuando el proceso en realidad está en control pero algún punto de la muestra individual cayó fuera de los límites de control debido sólo a fluctuaciones aleatorias. Esta es una ventaja de la carta \bar{X} sobre la carta X .

Entonces siguiendo la notación de las expresiones (6) y (7), los promedios de las $r_G(X_i)_s$ o $r_{G_m}(X_i)_s$ está dado por $Q(G, F_n^j)$ y $Q(G_m, F_n^j)$, aquí F_n^j es la distribución empírica de las X_i 's en el j -ésimo subconjunto, $j = 1, 2, \dots$. Así la carta Q dibuja los siguientes conjuntos de puntos: $\{Q(G, F_n^1), Q(G, F_n^2), \dots\}$ o $\{Q(G_m, F_n^1), Q(G_m, F_n^2), \dots\}$ si solo se dispone de Y_1, \dots, Y_m

Los límites de control de esta carta fueron hallados por Liu en su trabajo ya citado de 1995 y son los siguientes:

Para $n \geq 5$

$$CL = 0,5 ; LCL = (0,5 - Z_\alpha (12n)^{-1/2}) \text{ para dibujar } \{Q(G, F_n^j)\}'s \quad (9)$$

$$CL = 0,5 ; LCL = 0,5 - Z_\alpha \sqrt{\frac{1}{12} [(1/m) + (1/n)]} \text{ para dibujar } \{Q(G_m, F_n^j)\}'s \quad (10)$$

Para $n < 5$

$$CL = 0,5 ; LCL = (n! \alpha)^{1/n} / n \quad (11)$$

CC3. CARTA S

La carta S se inspira en la carta univariada CUSUM. Esta dibuja $\sum_{i=1}^n (X_i - \mu)$ lo cual refleja el patrón de la desviación total del valor esperado. Esta carta es más efectiva que las cartas X o \bar{X} en detectar pequeños cambios del proceso.

Según Liu, (Liu, 1995) en una carta S se dibujan los puntos correspondientes a las expresiones siguientes:

$$S_n(G) = \sum_{i=1}^n [r_G(X_i) - 1/2] \quad (12)$$

o

$$S_n(G_m) = \sum_{i=1}^n [r_{G_m}(X_i) - 1/2] \quad (13)$$

Los Límites de Control son los siguientes:

$$LCL = -(Z_\alpha(n/12)^{1/2}) \text{ carta S basada en } S_n(G) \quad (14)$$

$$LCL = -\{Z_\alpha \sqrt{n^2[(1/m) + (1/n)]/12} \text{ carta S basada en } S_n(G_m) \quad (15)$$

$$CL = 0 \text{ en ambos casos} \quad (16)$$

Los límites de control LCL en ambos casos constituyen una curva en vez de una recta; si n es grande la carta S fácilmente excedería el tamaño del papel, por lo que se estandariza de la siguiente manera:

Se dibuja:

$$S_n^*(G) = S_n(G) / \sqrt{n/12} \quad \text{o} \quad (17)$$

$$S_n^*(G_m) = S_n(G_m) / \sqrt{n^2[(1/m) + (1/n)]/12} \text{ para } n = 1, 2, \dots \quad (18)$$

Límites:

$$CL = 0 \quad (19)$$

$$LCL = -Z_\alpha \quad (20)$$

2.2.4 SIMULACION DE CARTAS DE CONTROL

Se reproducirán los ejemplos de simulación presentados en el artículo de Regina Liu : “Control Charts for Multivariate Processes”, (Liu, 1995). En esta simulación se calcula la Profundidad Simplicial: SD y a partir de ella se construyen todas las cartas expuestas a continuación. Se usa para ello el paquete en R: fda.usc el cual permite el cálculo de las profundidades MP1-MP5 descritas en 2.1.3

Carta r:

Se usa un conjunto de datos bivariado. La simulación se efectúa usando el lenguaje R. (Veáse el código respectivo en el Anexo 1

Sea $G \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$. Se generarán 540 puntos de muestra a partir de G ,

llamando a los primeros 500 como Y_1, \dots, Y_{500} y a los últimos 40 como X_1, \dots, X_{40} . A

continuación se generan 40 puntos de la distribución: $N\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}\right)$ y se llaman

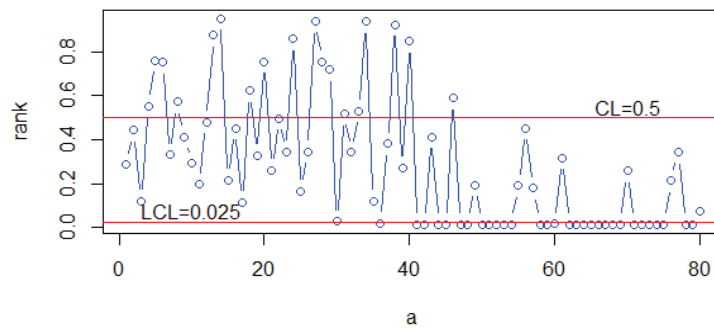
a estos puntos X_{41}, \dots, X_{80}

Procedimiento:

1. Se calcula la profundidad de los primeros 500 puntos bivariados de G Y_1, \dots, Y_{500} con respecto a ellos mismos.
2. Se juntan los últimos 40 puntos de G : X_1, \dots, X_{40} con los otros 40 puntos de la otra distribución X_{41}, \dots, X_{80} . Notar que ahora están en secuencia 40 puntos de G con 40 puntos de otra distribución con diferente vector de medias así como diferente matriz de varianzas-covarianzas.
3. Se calcula la profundidad de los 80 puntos últimos con respecto a los 500 puntos de G .
4. Se calculan los rangos de los 80 puntos analizados acorde a la expresión:
(4)
5. Se grafican en la carta r .

Las distribuciones escogidas son normales por facilidad, pero en realidad la normalidad no es un requisito de las cartas en estudio. Notar que existe un claro desplazamiento de la media y un incremento de escala en la distribución de los últimos 40 puntos. Es de esperarse que todas las cartas detecten este cambio.

Figura 1. Carta r



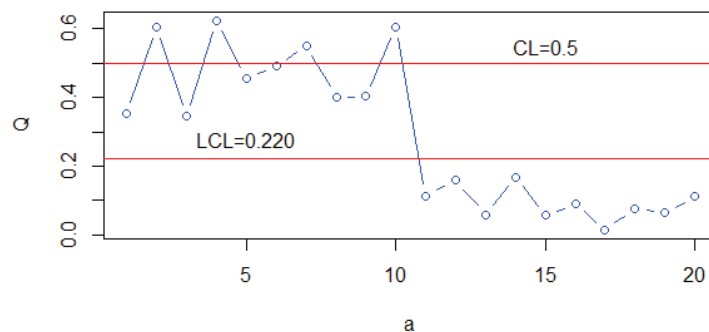
En la Figura 1 se muestran los límites de control (Líneas rojas), y se aprecia que a partir del dato 40 la carta r detecta el desplazamiento de la media y/o incremento de escala de una manera muy clara.

Carta Q ; n=4

Esta carta se muestra en la Figura 2 indicada más abajo. (Véase el código en el Anexo 1)

En la Figura 2, $CL = 0,5$ y $LCL = 0,220$; este último valor se ha obtenido usando la expresión (11).

Figura 2. Carta Q, n=4



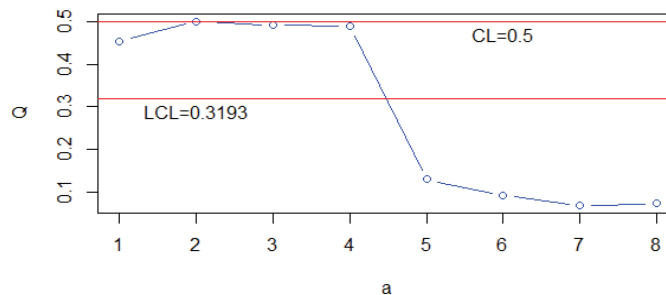
En la Figura 2, se aprecia que la carta detecta un cambio a partir del décimo primer punto. Considérese que Q es un promedio de cada 4 puntos, por tanto los primeros 40 puntos corresponderán a $40/4=10$ puntos de la carta mostrada en la figura 2. Nótese también que el promediado de los puntos ha eliminado las fluctuaciones aleatorias que aparecen en la Figura 1 (Carta r).

La carta Q con $n=10$ se muestra en la Figura 3. (Véase el código en el Anexo 1)

En la Figura 3, $CL = 0,5$ y $LCL = 0,3193$, este último valor se ha obtenido usando la expresión (10).

La carta detecta un cambio a partir del quinto punto. Nótese que Q es un promedio de cada 10 puntos, por lo cual los primeros 40 puntos corresponderán, a los 4 primeros de la carta Q correspondiente.

Figura 3. Carta Q, n=10

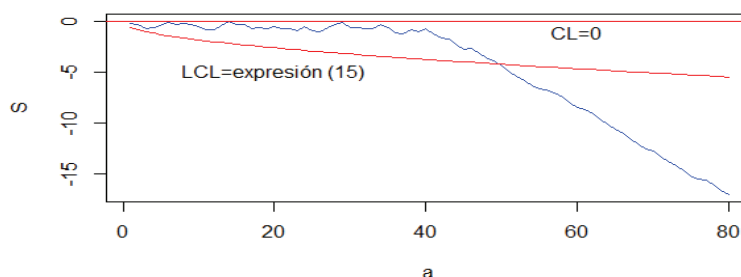


Carta S:

Se muestra en la Figura 4. Se han usado las expresiones (13), (15) y (16)

(Véase el código en el Anexo 1).

Figura 4. Carta S



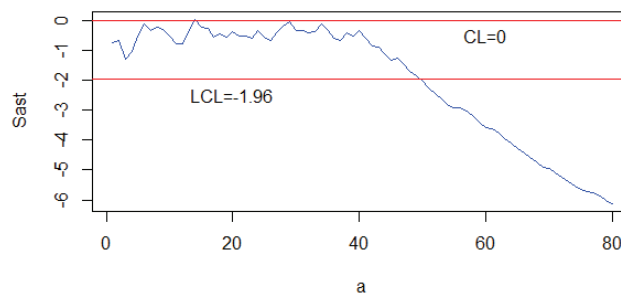
Carta S*.

(Véase el código en el Anexo 1)

Esta carta usa las expresiones (18), (19) y (20). $LCL = -1,96$ y $CL = 0$

Se muestra en la Figura 5

Figura 5. Carta S*



2.2.5 SISTEMA UMBRAL

SU1. Datos Representativos

Los datos a usar son una mezcla de datos de desempeño “normal” en sentido de ocurrencia típica, más un grupo de datos mejores que el desempeño normal y otro grupo de desempeño peor que el grupo de referencia.

Es necesario por tanto separarlos y tomar como referencia al núcleo típico.

Al delimitar un desempeño medio esperado, se contará con especificaciones realistas de lo que pueden lograr los estudiantes medios en la actualidad.

SU2. Clúster Medio:

Para seleccionar el conjunto de datos esperado, se seguirá el principio del Clúster Medio mostrado en el trabajo de Cheng, Liu y Luxhog (2000): “Monitoring multivariate aviation safety data by data depth: control charts and thresholds systems”, el cual establece que el escoger un clúster medio hace posible identificar en el total, desempeños que sean extraordinarios o fuera de lo ordinario. En el

mismo artículo se establece que el conjunto de referencia debería ser tal que el clúster medio sea lo suficientemente grande para ser representativo, de manera que se obtenga un umbral normal accesible, pero no tan grande de manera que el sistema umbral total sea suficientemente discriminante.

Liu define al clúster medio como el clúster geométrico central para dos o más dimensiones, (la mediana puede ser vista como el punto más central para datos de dos dimensiones, como es el caso que se trata en este trabajo y el clúster geométrico incluye a la mediana).

Puesto que el concepto de profundidad (la simplicial en este trabajo) induce un ordenamiento de adentro hacia afuera para el conjunto de datos, se usará este concepto para escoger el clúster medio o central.

Liu en el mismo artículo usa un procedimiento de separación y filtrado para configurar el clúster medio:

Procedimiento de Separación y Filtrado:

- a) Se calcula y ordenan los valores SD (profundidad Simplicial) de todas los conjuntos de individuos involucrados usando dos variables (Véase la sección 3.1.2 para la aplicación particular de este trabajo)
- b) Se selecciona el 80% más profundo o más central de los datos y se forma una región convexa englobando a estas medidas.
- c) Se considera al 80% de datos más profundos así seleccionados como el conjunto de referencia y a su región convexa como el umbral para identificar desempeños esperados.
- d) El 20% restante corresponderá al grupo de datos no representativos estructurado así: aproximadamente 10% como de atención requerida y 5% de cada uno de los extremos: inusualmente buenos o inusualmente malos. (El escoger 5% corresponde a la práctica usual de usar un p-valor del 5% en la prueba de cualquier hipótesis).

Según Liu (1999), hay dos razones para escoger al 80% como clúster medio. Primera: 80% es suficientemente grande para incluir “casi todos” los desempeños, de manera que el clúster medio es lo suficientemente amplio para ser visto como obtenible o de desempeño representativo. Esta escogencia obliga a ver el 20%

restante de las observaciones como “no esperadas” o “no representativas”, y así como Segunda: La fracción 20% es pequeña y aun suficientemente significativa para discriminar el desempeño no representativo.

SU3. Clasificación del desempeño

Para efectos de clasificación del desempeño, se considerarán los siguientes niveles:

- a) Esperado: Rango normal de una medida, es decir pertenece al conjunto de referencia.
- b) Advertencia: Indica que una medida es peor que el rango normal, o sea que no pertenece al conjunto de referencia.
- c) De Preocupación: Indica que una medida es significativamente peor que el rango normal.
- d) Informativa: Indica que una medida es mejor que el rango normal o sea mejor que el conjunto de referencia.

En el presente trabajo, se utilizarán dos variables para ser analizadas, por ejemplo Notas promedio en varias materias a lo largo del tiempo con porcentaje de materias aprobadas. (Véase la metodología).

Por lo expresado arriba se adoptarán los siguientes criterios para definir los niveles umbral:

- a) Esperado: Las dos variables son comparables con el clúster medio (conjunto de referencia)
- b) Advertencia: Los puntos se alejan un poco del clúster medio y generalmente una de las dos variables no es consistente con el conjunto de referencia.
- c) De Preocupación: Los puntos se alejan más del clúster medio y generalmente las dos variables son peores que las del conjunto de referencia
- d) Informativa: Las dos variables son mejores que el conjunto de referencia y están en la parte superior derecha más alejados del clúster medio.

3. METODOLOGÍA

Esta sección está estructurada de la siguiente manera:

- Primeramente, se describen los datos a usarse,
- A continuación se indica el procedimiento de separación del conjunto de referencia o “normal” para ser comparado con los datos individuales.
- Luego se explica lo que es el sistema umbral
- Finalmente se muestra el uso de la carta r.

Para todo lo anterior se ha seguido el siguiente algoritmo, el cual se ha implementado en el lenguaje R. (Se han usado además los paquetes qcr y fda.usc) (Flores, 2014) y (Febrero,2015).

Los pasos seguidos para la elaboración del Código R siguen la siguiente secuencia:

- 1) Carga de los Datos
- 2) Cálculo de las Profundidades
- 3) Obtención del Clúster de Referencia o Esperado
- 4) Para cada Materia o Estudiante se determinó:
 - Estadísticos Descriptivos (EDA)
 - Gráfico de Comportamiento en el tiempo (Serie)
 - Gráfico Bivariado para Clasificación
 - Monitoreo mediante carta r
- El código en R aparece en los Anexos 2 al

3.1 DATOS A USARSE

3.1.1 LIMITACIONES

Los datos que se usan en este trabajo provienen de una universidad, la cual ha solicitado que no se cite la fuente, debido a razones de sigilo educativo; no obstante se puede decir que son notas y otras variables como por ejemplo Número de Materias aprobadas, registro de asistencia, etc., de estudiantes que han cursado varias materias a lo largo de su carrera. Se dispone de información en diferentes materias a partir del año 2008 hasta el 2015. Las siglas de las materias son ficticias y son las siguientes:

Tabla 1. Materias

N°	Sigla
1	AES1
2	AES2
3	AES3
4	AES4
5	AES5
6	AES6
7	AES7

De cada materia se estudiarán dos variables: Nota Promedio semestral por un lapso de 16 semestres y % de Aprobación de Alumnos semestral.

También se estudiarán las notas promedio semestrales y % de Aprobación de materias semestral de 15 estudiantes los cuales se han escogido de la siguiente manera: Todos son de la cohorte del 2008 y concluyeron la malla respectiva en el 2011. Todos ellos cursaron materias de estadística. Las variables % de materias aprobadas así como porcentaje de alumnos aprobados, se han escogido pues constituyen variables de interés en otros trabajos realizados. Es importante mencionar que aunque estas variables están correlacionadas con la variable: notas, también se podrían usar variables no correlacionadas entre sí.

Los estudiantes se identifican con las siguientes siglas ficticias:

Tabla 2. Siglas de Estudiantes

N°	Identificación
1	8426
2	8427
3	8509
4	6054
5	6066
6	6166
7	6202

8	6450
9	6515
10	6841
11	7059
12	7060
13	7077
14	7080
15	7091

3.1.2 DATOS REPRESENTATIVOS

Los datos a usar son una mezcla de datos de desempeño “normal”, en el sentido de ocurrencia típica, más un grupo de datos o muy buenos o muy malos. Es necesario separarlos y tomar como referencia al núcleo típico.

Al delimitar un desempeño medio esperado, se contará con especificaciones realistas de lo que pueden lograr los estudiantes medios.

Procedimiento de separación del núcleo típico:

Este procedimiento se aplicará tanto al grupo de las 7 materias como el grupo de los 15 estudiantes. A continuación la explicación con las 7 materias:

- a) Se calculan y ordenan los valores de SD (Profundidad Simplicial) de todas las de materias (7 en total)
- b) Se selecciona el 80% más profundo o más central y se forma una región convexa a partir de estas medidas.
- c) El conjunto de referencia o representativo estará constituido por el 80% central seleccionado de esta manera. Cada Materia se comparará así con este conjunto de referencia para monitorear el desempeño.

El 80% es usado por Liu (1999) en su trabajo original y se escoge por las siguientes razones:

80% es suficientemente grande para incluir “casi todos” los desempeños, así que el clúster mediano es suficientemente amplio para ser visto como desempeño obtenible o representativo.

El 20% restante será considerado como no esperado o no representativo. La fracción 20% es pequeña y aún significativa para discriminar el desempeño no representativo.

Por lo anterior se tiene lo siguiente:

El 20% se distribuye así: Aproximadamente 5% en la dirección de ser superior y 5% de ser inferior al desempeño representativo. (El escoger 5% corresponde a la práctica usual de usar 5% =p-valor en probar cualquier hipótesis.

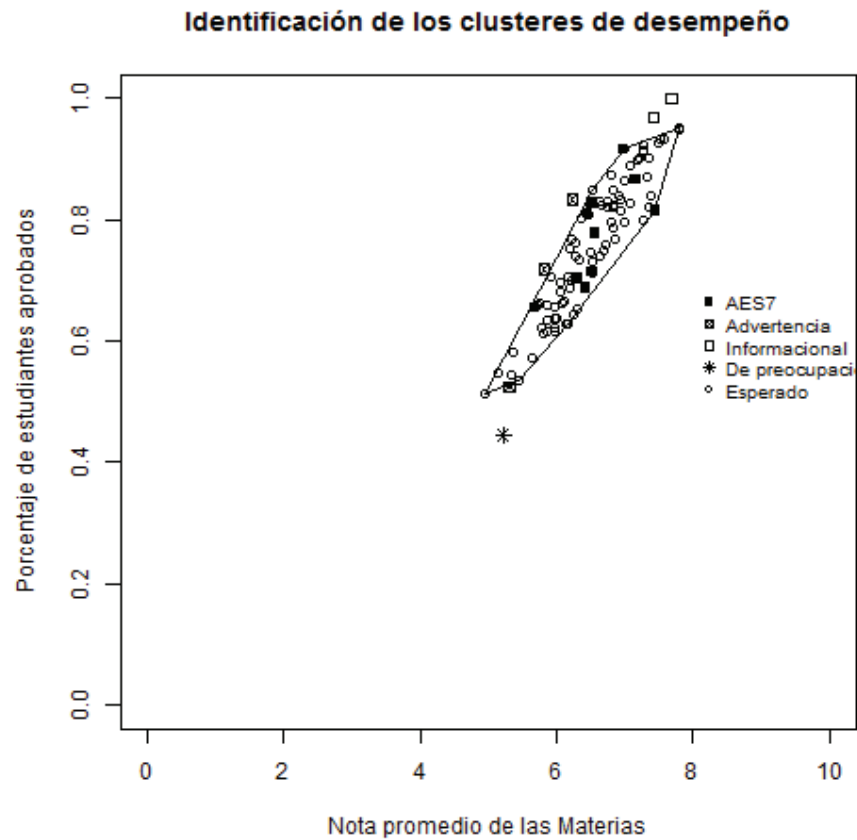
Este procedimiento provee un desglose de 80% de desempeños como esperados, 10% como atención requerida y 5% de cada extremo, inusualmente buenos o inusualmente malos.

3.1.3 SISTEMA UMBRAL

Se adoptarán los siguientes criterios para los niveles umbral:

- a) Esperado: Las dos variables son comparables con el clúster medio (conjunto de referencia). Véase por ejemplo la Figura 6: círculos blancos dentro de la envolvente convexa
- b) Advertencia: Los puntos se alejan un poco del clúster medio y generalmente una de las dos variables no es consistente con el conjunto de referencia. Véase por ejemplo la Figura 6: Cuadrados con una x fuera de la envolvente convexa.
- c) De Preocupación: Los puntos se alejan más del clúster medio y generalmente las dos variables son peores que las del conjunto de referencia. Véase por ejemplo la figura 6: Asteriscos fuera de la envolvente convexa
- d) Informativa: Las dos variables son mejores que el conjunto de referencia y están en la parte superior derecha más alejados del clúster medio. Véase por ejemplo la figura 6: cuadrados blancos fuera de la envolvente convexa y en la parte superior de la misma.

Figura 6. Gráfico bivariado de la materia: AES7



3.1.4 APLICACIÓN DE LA CARTA r

Se aplicará la carta r para monitorear las variables Nota promedio y % de Aprobación de estudiantes para cada materia mencionada (Tabla 1), así como nota promedio y % de materias aprobadas para cada estudiante (Tabla 2).

Los valores de las líneas de control en la carta r son las siguientes:

$$LCL = \alpha = 0,05$$

$$CL = 0,5$$

Sea G la distribución de los datos del clúster medio o esperado

Sea F la distribución monitoreada

Se probarán las siguientes hipótesis:

$$H_0: G = F$$

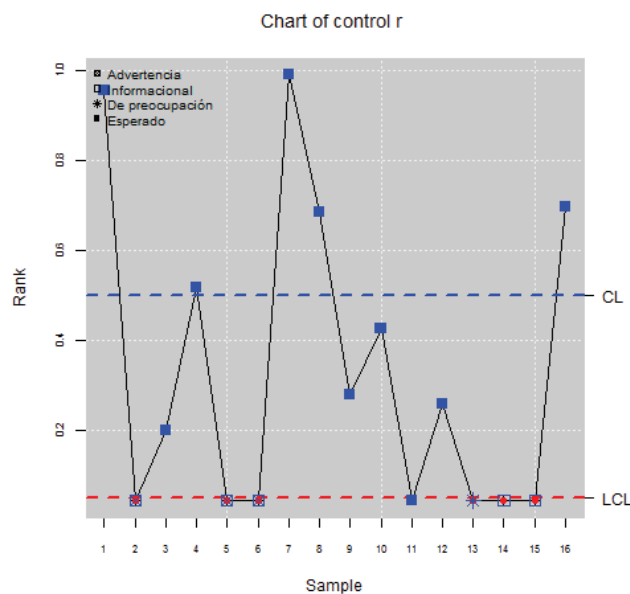
$H_a : G \neq F$ (Hay un desplazamiento de localización y/o un incremento de escala desde G a F)

Si $r < \alpha = 0,05$ pueden darse las siguientes situaciones: Los puntos corresponden a las condiciones: De Preocupación, De Advertencia; Informativa.

Como ejemplo se muestra en la Figura 7 la carta r de la misma materia analizada en la sección anterior:

Se pueden apreciar los seis puntos fuera de la envolvente convexa que caen bajo la línea LCL. Estos son: el 2,5,6,13,14,15. Hay un punto esperado: el 11 que aparece bajo la línea LCL, pero es porque se encuentra en la frontera de la envolvente convexa. (Véase la sección 4.3 para mayor detalle). En todo caso, se mantiene la información mostrada de la figura 6: Puntos 2,5 y 6 son puntos de advertencia; el 13 es de preocupación y los 14 y 15 son informativos. Nótese que en la carta r se conserva la secuencia en la que aparecieron los puntos en el tiempo. (Cuando los puntos caen bajo la línea LCL, se está aceptando que los puntos que lo hacen no pertenecen a la misma población que los del clúster medio, es decir ha habido un desplazamiento de localización y/o un incremento de escala desde G a F).

Figura 7. Carta r de AES7



4. RESULTADOS Y DISCUSIONES

Se usan los datos de una materia y de un estudiante para el análisis de esta sección. El resto de resultados aparecen en los Anexos 8-23

Primeramente se analiza la estructura de las variables involucradas y se comprueba si hay normalidad o no en los datos. Posteriormente se muestra la estructura en el tiempo de las variables involucradas.

A continuación se analiza el clúster de desempeño medio o de referencia que sirve para compararlo con las variables individuales: tanto de materias como de estudiantes. Finalmente se analizan las cartas r obtenidas.

4.1 MATERIAS

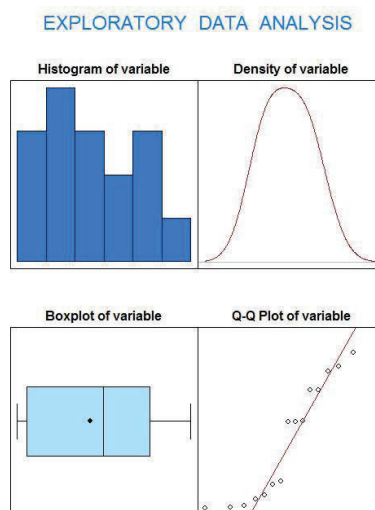
A continuación se discute una materia tanto por notas promedio, como porcentaje de aprobación de estudiantes. Todos los gráficos obtenidos de todas las materias se encuentran en los Anexos: 8 - 15

4.1.1 ESTADÍSTICA DESCRIPTIVA DE NOTAS POR MATERIA

Materia AES2:

Se muestra un resumen de Estadística Descriptiva de las Notas Promedio de la Materia mencionada, de esta manera se puede ver si la variable sigue o no una distribución normal:

Figura 8. Estadística Descriptiva de AES2



```
[1] "variable"--AES2
Size (n)  Missing  Minimum  1st Qu   Mean   Median  TrMean  3rd Qu   Max.   Stdev.
 16.000    0.000    5.963   6.077   6.831  6.980   6.831   7.524   8.012  0.734
   Var.  SE Mean  I.Q.R.   Range Kurtosis Skewness SW p-val
 0.539   0.184   1.447   2.049  -1.687   0.109  0.046
```

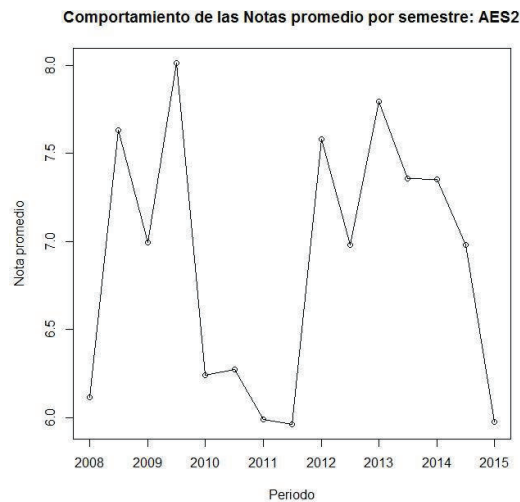
Vale la pena mencionar que la nota de aprobación en la Universidad de donde provienen los datos es de 6.0, ello explicaría porque los valores centrales como la media, la mediana, están alrededor de ese valor.

Por otra parte, el estadístico de Shapiro-Wilk (SW) para probar normalidad, muestra un p-valor =0,046. (El nivel de significancia usado en esta prueba es de $\alpha = 0,05$), indicando que los datos no se ajustan a una distribución normal.

4.1.2 COMPORTAMIENTO DE LAS NOTAS PROMEDIO A LO LARGO DEL TIEMPO

La Figura 9 muestra el comportamiento de las notas promedio de AES2 a lo largo del tiempo.

Figura 9. Comportamiento de Notas Promedio por semestre de AES2

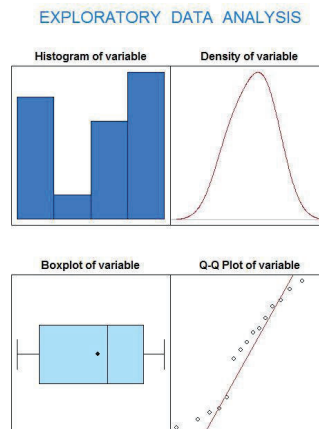


En la Figura 9 se aprecia que las notas tienen sendos bajones en los semestres del 2010 al 2011 volviendo a subir en el 2012 con un posterior bajón en el 2015.

4.1.3 ESTADÍSTICA DESCRIPTIVA DEL PORCENTAJE DE APROBACIÓN DE ALUMNOS

Se muestra a continuación un resumen de Estadística Descriptiva del Porcentaje de Aprobación de Alumnos de AES2:

Figura 10. Estadística Descriptiva del Porcentaje de Aprobación de Alumnos en AES2



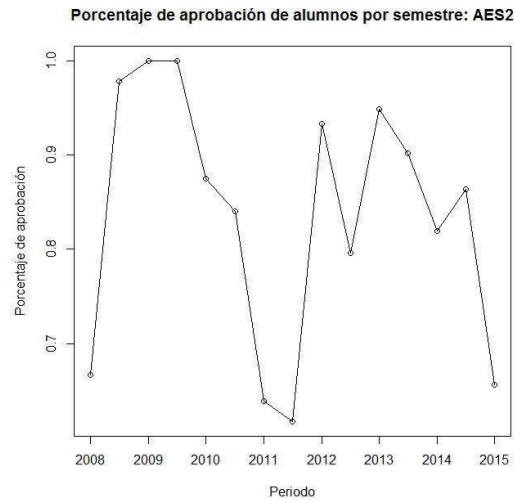
```
[1] "variable"- P-AES2
Size (n)  Missing  Minimum  1st Qu   Mean   Median  TrMean  3rd Qu   Max.   Stdev.
  16.000    0.000    0.617   0.674   0.827  0.852   0.827   0.945   1.000  0.134
  Var.  SE Mean  I.Q.R.   Range  Kurtosis  Skewness  SW p-val
  0.018  0.034   0.271   0.383  -1.523   -0.254   0.126
```

Del análisis descriptivo se ve que la media está alrededor del 82% y los datos siguen una distribución normal. El análisis se lo hace para ver con que tipo de variable se trabajará en la carta r: normal o no.

4.1.4 COMPORTAMIENTO DEL PORCENTAJE DE APROBACIÓN DE ALUMNOS A LO LARGO DEL TIEMPO

A continuación se muestra el comportamiento del Porcentaje de Aprobación de los alumnos a lo largo del tiempo de la materia AES2:

Figura 11. Porcentaje de Aprobación de Alumnos por semestre. AES2



La figura anterior muestra un comportamiento similar al de la Figura 9 en cuanto a los bajones ocurridos alrededor del 2011 y 2015.

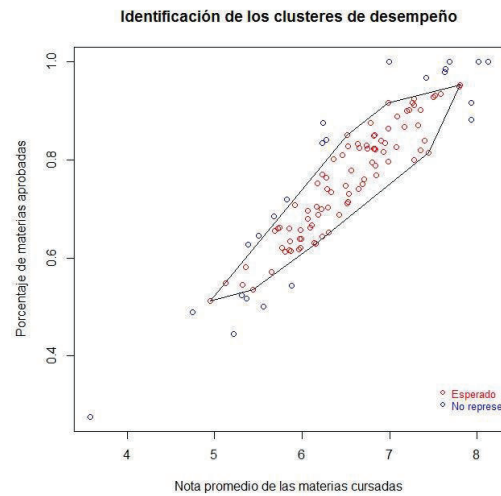
4.1.5 CLÚSTER DE DESEMPEÑO MEDIO

A continuación se muestra en la Figura 12, el Clúster de Desempeño esperado de todas las 7 materias consideradas. En la figura se puede apreciar la envolvente convexa que limita el 80% de los datos más profundos, tal como se explicó en la sección 3.2.1 Datos Representativos.

Se debe recordar que este conjunto constituye el de referencia. Cada materia se compara con este conjunto para obtener los datos esperados como los no representativos.

En la misma figura se aprecian algunos puntos fuera de la envolvente convexa que vienen a ser puntos no representativos

Figura 12. Clúster de Referencia. Todas las materias

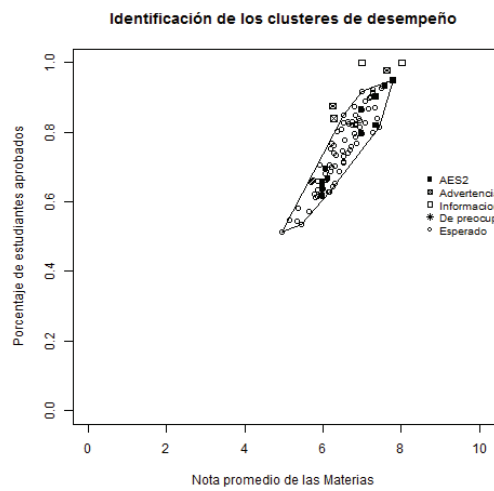


4.1.6 COMPARACIÓN DE LA MATERIA EN PARTICULAR CON EL CLÚSTER MEDIO

A continuación se comparan los datos de AES2 con los del clúster medio. (Nótese que se realiza una clasificación del desempeño).

En la Figura 13 se aprecia que hay 5 puntos fuera de la envolvente convexa, siendo tres de advertencia y dos de ellos informativos (éstos últimos con cuadros en blanco). Los de advertencia nos indican que se debe prestar atención al desempeño de las materias y los informativos muestran que ha ocurrido un desempeño superior al de referencia. Esta información se puede complementar con la que proporciona la carta r mostrada en la Figura 14.

Figura 13. Gráfico Bivariado de AES2..



4.1.7 CARTA r

En general la carta r (Figura 14) indica la secuencia obtenida del desempeño de la materia AES2 en el tiempo. Luego de los 5 puntos bajo la línea LCL se aprecia un comportamiento con tramos bajo CL y otros encima de la misma.

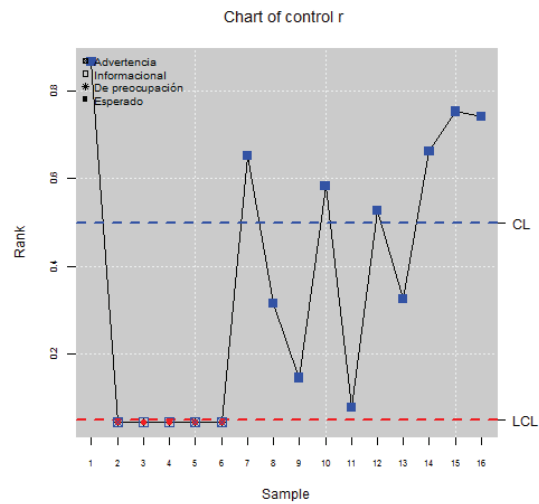
En la misma Figura se puede ver que hay 5 puntos (del 2 al 6) que están por debajo de la línea LCL. Estos puntos corresponden a los 5 puntos mostrados por el gráfico bivariado de la Figura 13.

En la Figura 14 se tiene que: los puntos 2, 5 y 6 corresponden a desempeños de advertencia y el 3 y 4 en cambio a desempeños mejores que el de referencia o informacionales.

Nótese que en el eje horizontal de la carta r se puede asumir que las unidades correspondientes son de tiempo, en este caso semestres, por tanto, los puntos extraordinarios ocurrieron en el segundo semestre del 2008 hasta el segundo semestre del 2010. Por ejemplo el evento del punto 5 ocurrió en el primer semestre del 2010.

Vale la pena aclarar que el paquete usado (qcr) para hacer las cartas de control r viene con la leyenda "sample" en el eje horizontal. (La modificación del mismo no es asequible, por ello se lo ha dejado igual).

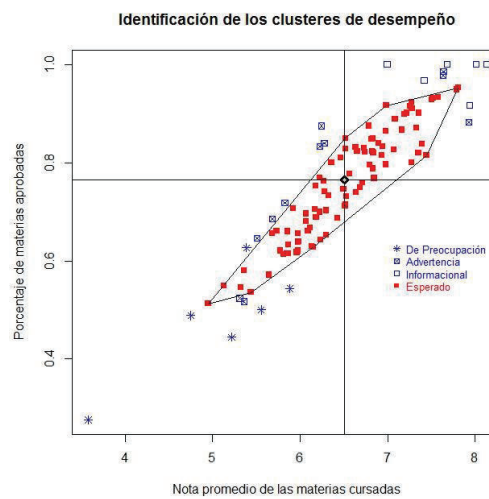
Figura 14. Carta r. AES2



4.1.8 CLASIFICACIÓN DEL DESEMPEÑO DE LAS 7 MATERIAS

La Figura 15 muestra el Clúster esperado o de referencia así como la clasificación del desempeño del total de las materias al ser comparadas con aquel.

Figura 15. Clúster esperado con la clasificación de desempeños de las demás materias.



4.1.9 COMPARACIÓN TOTAL DE MATERIAS

La tabla 3, muestra las medias y coeficiente de variación de las materias estudiadas por cada grupo de desempeño tanto para notas como para porcentaje de aprobación, en donde se puede apreciar que la categoría *Informacional* es la que tiene valores más altos tanto en nota como en porcentaje de aprobación; le siguen en ese orden las categorías de *Esperado*, *Advertencia* y *De preocupación*.

Respecto del %AP se observa que el valor más alto corresponde a la categoría *Informacional*, le sigue la categoría *Advertencia*, luego la categoría *Esperado* y finalmente la *De preocupación*. Con respecto a la homogeneidad de los datos, se aprecia que la categoría más homogénea es la *Informacional* pues su valores de CV%, tanto en nota como %AP son los más bajos: 5,6% y 3,5% respectivamente. Se aprecia también que los menos homogéneos son los grupos de *Advertencia* y *De preocupación* con valores de CV% más altos.

Tabla 3. Comparación Total de Materias

Categoría	Estadístico	Nota	%AP
<i>Informacional</i>	Media	7,694	98,1
	C.V. %	5,6	3,5
<i>Esperado</i>	Media	6,515	75,6
	C.V. %	9,5	14,7
<i>Advertencia</i>	Media	6,329	77,1
	C.V. %	15,2	21,4
<i>De preocupación</i>	Media	5,062	48,0
	C.V. %	16,2	24,4

4.1.10 EJEMPLO DE COMPARACIÓN DE MATERIAS

La tabla 4 muestra los valores medios de las variables Nota y % de Aprobación con sus respectivos coeficientes de variación. Se ha considerado para el ejemplo las siguientes materias: AES3; AES7 y AES2 en las categorías: *Informacional*, *Esperado*, *Advertencia* y *De preocupación*. Además se cuenta con la frecuencia de ocurrencia de las materias en las respectivas categorías.

En la tabla se aprecia que de las tres materias con respecto a la categoría de referencia (*Esperado*) la materia con mejor desempeño en cuanto a Nota fue la AES2 (6.7), aunque en porcentaje de aprobación (78%) está empatada prácticamente con AES7. En la categoría *Informacional*, en cambio la nota de

mayor valor es la de AES7 (7,5), pero en cambio el porcentaje más alto de aprobación está en AES2 (100%). En AES3 no se observa la categoría *Informacional*, y es en ésta donde se presentan los valores más bajos tanto en las categorías de *Advertencia* (5,68) y *De preocupación* (4,94); lo propio ocurre con el porcentaje de aprobación: *Advertencia* (68,5%) y *De preocupación* (45,2%).

En conclusión, el desempeño de la materia AES3 fue inferior al de las materias AES7 y AES2, siendo la mejor AES2. Véase en 4.1.11 lo que se podría hacer para mejorar el desempeño.

Tabla 4. Comparación de Materias

Categoría	Estadístico	AES3			AES7			AES2		
		Nota	%AP	n	Nota	%AP	n	Nota	%AP	n
<i>Informacional</i>	Media				7,548	98,4	2	7,503	100	2
	C.V.%				2,5	2,3		9,6	0	
<i>Esperado</i>	Media	6,589	75,7	11	6,606	77,7	10	6,741	77,6	11
	C.V.%	10,6	17,6		7,4	10,9		10,8	16,1	
<i>Advertencia</i>	Media	5,682	68,5	1	5,787	69,2	3	6,716	89,8	3
	C.V.%				7,9	22,5		11,8	8,0	
<i>De preocupación</i>	Media	4,942	45,2	4	5,217	44,4	1			
	C.V.%	20,8	26,5							

4.1.11 DEFINICIÓN DE METAS PARA IMPLEMENTAR UN PLAN DE MEJORAMIENTO DEL RENDIMIENTO DE MATERIAS

La tabla 5, propone metas que deberían alcanzarse para mejorar el desempeño de las materias a nivel general: para pasar de la categoría *De preocupación* a *Esperado*, la nota promedio del grupo *De preocupación* debería pasar de un máximo de 4,75 (este valor corresponde al dato más profundo de la categoría *De preocupación*, es decir el dato que más se acerca al grupo categoría *Esperado*) a un mínimo de 6,50 (este valor corresponde al dato menos profundo de los peores de la categoría *Esperado*, es decir el dato que más se acerca al grupo categoría *De preocupación*) y finalmente el porcentaje de aprobación promedio debería pasar de un máximo de 49 a un mínimo de 75 (ambos en la misma connotación que la descrita para las notas promedio).

Para pasar de la categoría de *Esperado* a *Informacional*, la nota promedio del grupo *Esperado* debería pasar de un máximo de 6,56 (este valor corresponde al dato menos profundo de los mejores de la categoría *Esperado*, es decir el dato que más se acerca al grupo categoría *Informacional*) a un mínimo de 7,93 (este valor corresponde al dato más profundo de la categoría *Informacional*, es decir el dato que más se acerca al grupo categoría *Esperado*). Y finalmente, el porcentaje de aprobación debería pasar igualmente de 78 a 92 (ambos en la misma connotación que la descrita para las notas promedio).

Tabla 5. Plan de Mejoramiento de Materias

	De preocupación a Esperado		De Esperado a Informacional	
	Nota	% AP	Nota	% AP
Nota de Mejora	4,751163	48,83721	6,562963	77,77778
	6,497015	74,62687	7,933333	91,66667

Con base a la información de la tabla 6, se plantean metas de mejoramiento para la categoría de Advertencia acorde a si las variables son mayor o igual que la mediana o sin son menor o igual que la mediana de la variable respectiva (Nota y %AP). Para esto se han considerado cuatro subgrupos. El primer subgrupo: $\geq y \geq$, deberán pasar de una Nota de 7,63 a 7,93 pasando así de la categoría de *Advertencia* directamente a la *Informacional*, en cuanto al %AP estaría bien pues su valor es superior al mínimo de Informacional que es 91,67 (tabla 5). Del segundo grupo: $\leq y \geq$, deberán pasar la Nota de 6,28 a 7,93 pasando así de *Advertencia* a *Informacional*, en cuanto a %AP deberá pasar de 84,0 a 91,67 (tabla 5). Los del cuarto grupo: $\leq y \leq$ deberían pasar del grupo de *Advertencia* al *Esperado* mejorando tanto la Nota: de 5,50 a 6,497 como el %AP de 65,0 a 74,6 (tabla 5).

La forma práctica de plasmar las mejoras perseguidas iría ligada a actividades departamentales como son por ejemplo tutorías sistemáticas a los grupos cuyo desempeño se busca mejorar, tal vez ajustar las metodologías que usan los profesores de manera que los más exitosos compartan sus métodos con los demás, etc.

Tabla 6. Plan de Mejora de Materias para los grupos de Advertencia

Med Nota	Med %AP	Nota	%AP	n	De Advertencia a:
≥	≥	7,63	97,8	3	<i>Informacional:</i> Nota necesaria= 7,93 // %AP = ok
≤	≥	6,28	84,0	3	<i>Informacional:</i> Nota necesaria = 7,93 // %AP = 91,67
≥	≤	No hay	No hay	0	
≤	≤	5,50	65,0	5	<i>Esperado:</i> Nota necesaria = 6,497 // %AP neces. = 74,6

4.2 ESTUDIANTES

A continuación se discuten los datos de un estudiante, tanto para notas promedio, como porcentaje de aprobación de materias. Todos los gráficos obtenidos de todos los estudiantes se encuentran en los Anexos 16 – 23.

Se analiza la estructura de las variables estudiando la estadística descriptiva para ver si son normales o no. A continuación se observa el comportamiento de las variables a lo largo del tiempo.

Seguidamente se muestra el Clúster de desempeño medio o de referencia comparándolo con el desempeño del estudiante en particular.

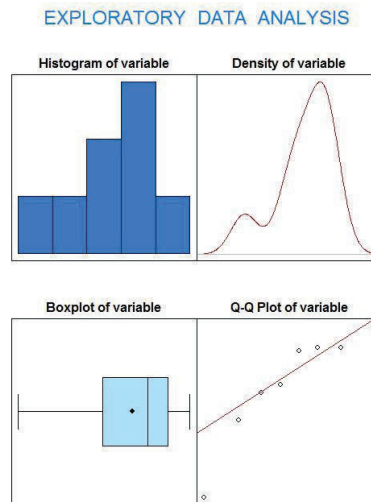
Finalmente se muestra la carta r del estudiante en particular.

4.2.1 ESTADÍSTICA DESCRIPTIVA DE NOTAS POR ESTUDIANTE ESTUDIANTE 7091

Se muestra un resumen de Estadística Descriptiva de las Notas Promedio del Estudiante mencionado:

Se puede apreciar que su media de notas es 7,62 y según la prueba de normalidad de Shapiro-Wilk los datos son normales pues p-valor=0,248.

Figura 16 . Estadística Descriptiva de Notas promedio del Estudiante 7091



[1] "variable"-ESTUDIANTE 7091

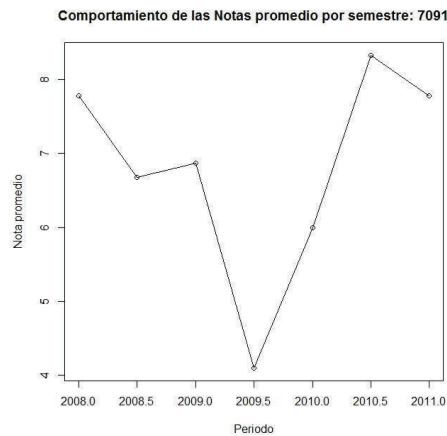
Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.414	6.712	7.620	7.397	7.620	8.125	10.000	1.140
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	1.299	0.403	1.413	3.586	-0.366	0.897	0.248		

4.2.2 COMPORTAMIENTO DE LAS NOTAS PROMEDIO A LO LARGO DEL TIEMPO ESTUDIANTE 7091

En la Figura 17 se aprecia el comportamiento de las notas promedio del mencionado estudiante a lo largo del tiempo.

Se puede apreciar que tuvo un promedio bajo 6 en el segundo semestre del 2009.

Figura 17. Comportamiento de las notas promedio del estudiante 7091 a lo largo del tiempo.

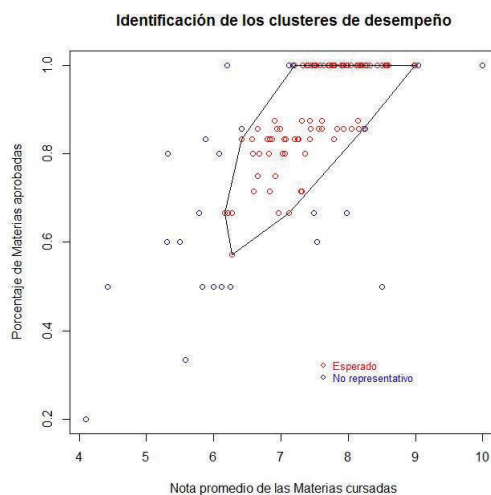


4.2.3 CLÚSTER DE DESEMPEÑO MEDIO

En la Figura 18 se muestra el clúster esperado o de referencia que se obtiene luego de considerar el 80% de datos más profundo de todas las notas de los 15 estudiantes.

Se puede apreciar la presencia de puntos esperados y puntos no representativos.

Figura 18. Clúster de referencia o esperado de los 15 estudiantes

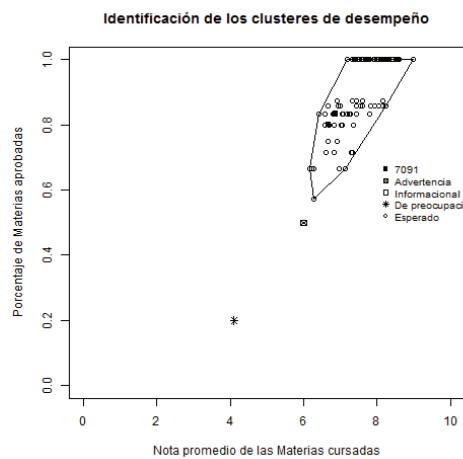


4.2.4 COMPARACIÓN DEL DESEMPEÑO DEL ESTUDIANTE EN PARTICULAR CON EL CLÚSTER MEDIO

La Figura 19 muestra la comparación del Estudiante 7091 con el clúster medio en cuanto al desempeño de las materias cursadas.

Se puede apreciar que hay dos puntos fuera de la envolvente convexa: uno de preocupación y uno de advertencia. Lo mismo se puede apreciar en la Figura 18 de la carta r del mencionado estudiante.

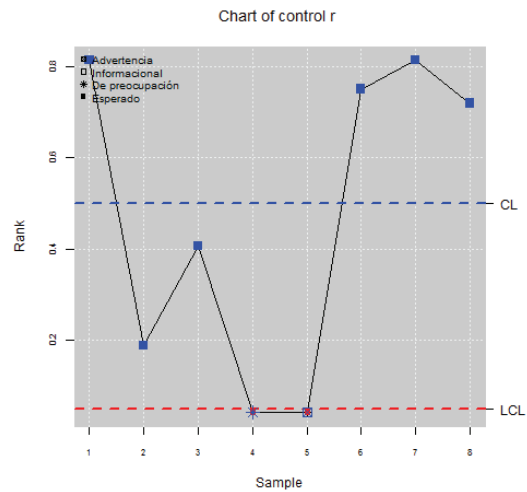
Figura 19. Comparación del desempeño del estudiante 7091 con el clúster medio.



4.2.5 CARTA r DEL ESTUDIANTE

En la Figura 20 se muestra la aparición de dos puntos bajo la línea LCL: uno en el segundo semestre del 2009 y el otro en el primer semestre del 2010, lo cual concuerda con lo observado en el gráfico bivariado de la Figura 19. Además se ve claramente el tipo de puntos que son: el primero es de preocupación y el segundo de advertencia. Nótese que con la carta r se puede ver la secuencia de ocurrencia de los eventos en el tiempo: cual ocurrió primero, de que tipo y cuando, a diferencia del gráfico bivariado que sólo muestra la existencia de los dos eventos extraordinarios fuera de la envolvente convexa.

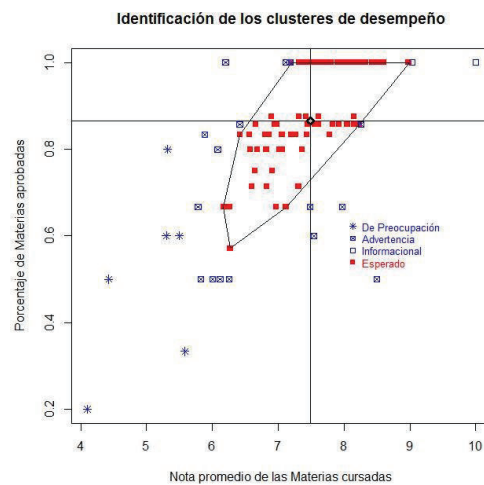
Figura 20. Carta r del Estudiante 7091



4.2.6 CLASIFICACIÓN DEL DESEMPEÑO DE LOS 15 ESTUDIANTES

La Figura 21 muestra el Clúster Esperado o de Referencia encerrado por la envolvente convexa del 80% de datos más profundos y además los otros puntos acorde a la clasificación que proporciona el gráfico bivariado.

Figura 21. Clúster esperado con la clasificación de desempeños de los demás estudiantes.



4.2.7 COMPARACIÓN TOTAL DE ESTUDIANTES

La tabla 7 muestra las medias y coeficiente de variación de los estudiantes considerados por cada categoría de desempeño tanto para notas como para porcentaje de aprobación, en donde se puede apreciar que la categoría *Informacional* es la que tiene valores más altos tanto en nota como en porcentaje de aprobación; le siguen en ese orden las categorías *Esperado*, *Advertencia* y *De preocupación*. En cuanto a homogeneidad el más homogéneo es el grupo *Informacional* y los menos homogéneos resultaron ser los de *Advertencia* y *De preocupación*, pues sus valores de CV% son los más elevados.

Tabla 7. Comparación Total de Estudiantes

Categoría	Estadístico	Nota	%AP
<i>Informacional</i>	Media	9,519	100
	C.V.%	7,1	0
<i>Esperado</i>	Media	7,582	90,5
	C.V.%	8,6	12,4
<i>Advertencia</i>	Media	6,788	71,5
	C.V.%	13,7	27,0
<i>De preocupación</i>	Media	5,038	50,6
	C.V.%	12,3	42,1

4.2.8 EJEMPLO DE COMPARACIÓN DE ESTUDIANTES

La tabla 8 muestra los valores medios de las variables Nota y % de Aprobación con sus respectivos coeficientes de variación. Se ha considerado para el ejemplo los siguientes estudiantes: 6515; 7091 y 7060 en las categorías: *Informacional*, *Esperado*, *Advertencia* y *De preocupación*. Además se cuenta con la frecuencia de ocurrencia de los estudiantes en las respectivas categorías.

De los tres el de mejor desempeño fue el 7060 el cual aparece en la categoría *Informacional* (1 vez) con Nota = 10 y %AP = 100%, además nunca apareció en la categoría *De preocupación*, pero sí en la de *Advertencia* (2 veces) con Nota de 6,8 y %AP= 93%. El de peor desempeño fue el estudiante 6515 el cual nunca apareció en la categoría *informacional* pero si en las categorías *Esperado* (2 veces), de

Advertencia (4 veces) y de preocupación (2 veces). Sus notas y %AP se pueden apreciar en la mencionada tabla.

Tabla 8. Comparación de 3 Estudiantes

Categoría	Estadístico	6515			7091			7060		
		Nota	%AP	n	Nota	%AP	n	Nota	%AP	n
Informativa	Media							10	100	1
	C.V.%									
Esperado	Media	6,68	84,52	2	7,522	93,9	6	7,473	88,1	5
	C.V.%	5,6	2,0		8,3	10,1		8,7	13,7	
Advertencia	Media	5,893	70,0	4	6	50,0	1	6,797	92,86	2
	C.V.	2,2	21,7					8	10,9	
De preocupación	Media	4,962	55,0	2	4,1	20,0	1			
	C.V.	15,3	12,9							

4.2.9 DEFINICIÓN DE METAS PARA IMPLEMENTAR UN PLAN DE MEJORAMIENTO DEL RENDIMIENTO DE ESTUDIANTES

La tabla 9, propone metas que deberían alcanzarse para mejorar el desempeño de los estudiantes a nivel general: para pasar de la categoría *De preocupación* a *Esperado*, la nota promedio del grupo *De preocupación* debería pasar de un máximo de 5,5 (este valor corresponde al dato más profundo de la categoría *De preocupación*, es decir el dato que más se acerca al grupo categoría *Esperado*) a un mínimo de 7,44 (este valor corresponde al dato menos profundo de los peores de la categoría *Esperado*, es decir el dato que más se acerca al grupo categoría *De preocupación*) y finalmente el porcentaje de aprobación promedio debería pasar de un máximo de 0,6 a un mínimo de 0,86 (ambos en la misma connotación que la descrita para las notas promedio).

Para pasar de la categoría de *Esperado* a *Informativa*, la nota promedio del grupo *Esperado* debería pasar de un máximo de 7,98 (este valor corresponde al dato menos profundo de los mejores de la categoría *Esperado*, es decir el dato que más se acerca al grupo categoría *Informativa*) a un mínimo de 9,04 (este valor corresponde al dato más profundo de la categoría *Informativa*, es decir el dato que más se acerca al grupo categoría *Esperado*). Y finalmente, el porcentaje de aprobación debería mantenerse en un 100%.

Tabla 9. Plan de Mejoramiento de Estudiantes

Nota de Mejora	De preocupación a Esperado		De Esperado a Informativa	
	Nota	% AP	Nota	% AP
	5,5	60	7,98	100
7.442857	85,71429	9,0375	100	

Con base a la información de la tabla 10, se plantean metas de mejoramiento para la categoría de Advertencia acorde a si las variables son mayor o igual que la mediana o si son menor o igual que la mediana de la variable respectiva (Nota y %AP). Para esto se han considerado cuatro subgrupos. El primer subgrupo ($\geq y \geq$) no se tienen observaciones mientras que en el segundo grupo: $\leq y \geq$, los estudiantes deberán pasar de una Nota de 7,18 a 9,04 pasando así de *Advertencia* a *Informativa*, en cuanto a %AP está bien pues tiene un valor de 100 (tabla 9). Los del cuarto grupo: $\leq y \leq$ deberían pasar del grupo de *Advertencia* al *Esperado* mejorando el %AP: de 66,7 a 85,7 (tabla 9)

Tabla 10. Plan de Mejora de Estudiantes para los grupos de Advertencia

Med Nota	Med %AP	Nota	%AP	n	De Advertencia a:
\geq	\geq	No hay	No hay	0	
\leq	\geq	7,18	100	3	<i>Informativa</i> : Nota necesaria = 9,04 // %AP = ok
\geq	\leq	8,257	85,7	4	<i>Informativa</i> : Nota necesaria = 9,04 // %AP=100
\leq	\leq	7,483	66,7	9	<i>Esperado</i> : Nota necesaria = ok // %AP neces. = 85,7

Asimismo las estrategias prácticas para conseguir implementar estas mejoras en notas podrían ser tutorías dirigidas a los grupos de interés, difundir técnicas de estudio adecuadas a los estudiantes, etc.

4.3 CARTA r. PUNTOS JUSTO EN LA ENVOLVENTE CONVEXA

Pueden ocurrir situaciones como las mostradas en las figura 22 y 23. En la primera se ve que hay dos puntos de desempeño esperado (cuadrados negros) que coinciden con la frontera o envolvente convexa (parte inferior de la envolvente convexa).

La figura 23 muestra la carta r de dicho estudiante donde aparecen dos puntos de desempeño esperado (cuadrados azules) bajo la línea LCL. El porqué de esta situación se debe al hecho explicado en el párrafo anterior, es decir hay dos puntos esperados ubicados justo en la envolvente convexa.

No obstante, la carta r indica este hecho, con lo cual ya se sabe cómo proceder en este caso: ignorar esas alertas.

Figura 22. Estudiante 6054. Clúster Esperado y Puntos sobre la envolvente convexa

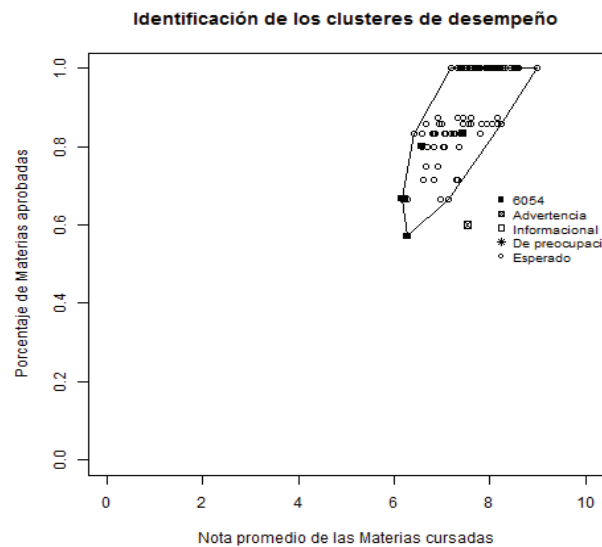
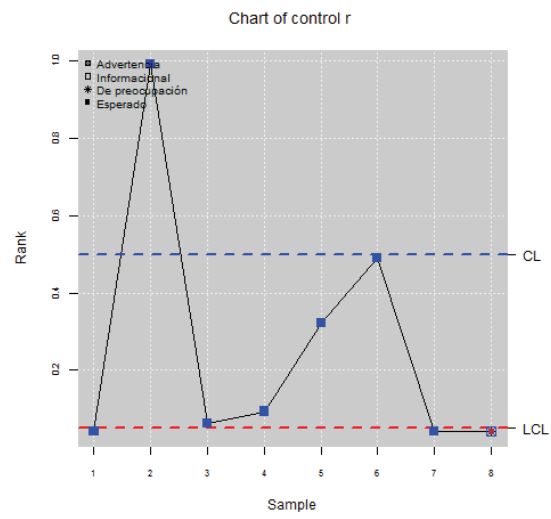


Figura 23. Carta r del Estudiante 6054



5. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- El concepto de profundidad permite transformar una observación multivariante a un índice univariante, el cual es susceptible de monitorearse en una carta de control.
- Acorde a lo investigado, la profundidad Simplicial cumple las propiedades D1-D3 y no es muy usada ya que las regiones inducidas por ella no son convexas. No obstante para dos dimensiones se pudo trabajar con la misma aprovechando el trabajo de Regina Liu.
- La Carta r es una generalización multivariada, basada en el concepto de profundidad de datos, de la carta univariada X .
- La Carta r es superior al gráfico bivariado para mostrar información, pues por un lado la carta r permite graficar el comportamiento del proceso para dos o más dimensiones, en tanto el gráfico bivariado sólo para dos dimensiones.
- En la carta r , se preserva la secuencia de ocurrencia de los eventos en el tiempo a diferencia de la gráfica bivariada que no lo hace.
- A diferencia de las cartas tradicionales univariadas, la carta r tiene sólo una línea de control: LCL , que sirve para separar los eventos esperados de los extraordinarios (buenos y malos).
- La carta r sirve para monitorear tanto datos normales como aquellos que no sigan tal distribución.
- El gráfico bivariado está limitado a mostrar información sólo de dos variables, en tanto la carta r puede hacerlo para observaciones multivariadas.
- El gráfico bivariado sirve para mostrar la clasificación de desempeño acorde a los criterios establecidos pero no muestra información secuencial en el tiempo.
- La carta Q es una generalización multivariada, basada en el concepto de profundidad de datos de la carta univariada \bar{X} , en este trabajo no se utilizó pero se citó su existencia.

- La carta S es una generalización multivariada, basada en el concepto de profundidad de datos, de la carta univariada $CUSUM$, en este trabajo no se utilizó pero se citó su existencia.
- En cuanto a materias y a porcentaje de alumnos aprobados, el análisis mostró que los datos de las materias seguían casi en su totalidad una distribución normal (excepto AES2).
- Respecto a materias y a porcentaje de alumnos aprobados, las cartas r muestran el desenvolvimiento de las materias en el tiempo y se ve claramente la presencia de altos y bajos en el desempeño. Los puntos extraordinarios brindan oportunidades de realizar una realimentación a los entes responsables del desarrollo de las asignaturas (profesores, coordinadores, etc.): aquellos que están mal para que mejoren y aquellos que están bien para que continúen así.
- Respecto a materias y a porcentaje de alumnos aprobados, el análisis mostró que los datos de las materias seguían en su totalidad una distribución normal.
- Respecto a notas promedio de los 15 estudiantes, el análisis mostró que los datos de notas promedio seguían casi en su totalidad distribuciones normales (Excepto: 6450 y 6066)
- Las cartas r muestran el desenvolvimiento de las notas promedio en el tiempo y se ve claramente la presencia de altos y bajos en el desempeño. Los puntos extraordinarios brindan oportunidades de realizar una realimentación a estudiantes: aquellos que están mal para que mejoren y aquellos que están bien para que continúen así.
- Con respecto a los 15 estudiantes en particular, el análisis mostró que el porcentaje de materias aprobadas seguía en su mayoría una distribución no normal (Véase el Anexo 19)
- Es posible delimitar fronteras de paso entre las categorías, lo cual permitiría establecer estrategias de mejora para intentar pasar de un grupo de desempeño inferior a un grupo de desempeño superior.

5.2 RECOMENDACIONES

- Si se trabaja con más de dos dimensiones, se recomienda trabajar con alguna de las otras definiciones de profundidad expuestas en el trabajo.
- Se podría implementar un sistema institucional que utilice la metodología propuesta para monitorear a estudiantes y materias.
- La implementación de un sistema de notas que monitoree el desempeño tanto de materias como de individuos, se lo puede realizar y entre sus características se usaría a R como motor ya que en varios paquetes existentes hay implementaciones de varias funciones de profundidad que incluyen la simplicial además de otras
- El enfoque de control de calidad usado permitiría realizar realimentaciones oportunas tanto a nivel de materias como a nivel individual de los estudiantes
- La universidad de donde provienen los datos considera este trabajo como un proyecto piloto. Para su aplicación a futuro deberían considerarse más variables, con lo cual la profundidad simplicial ya no se usaría, sino alguna otra.

REFERENCIAS

Barahona P., Aliaga V. Variables predictoras del rendimiento académico de los alumnos de primer año de las carreras de Humanidades de la Universidad de Atacama, Chile. *Rev. Int. Investig. Cienc. Soc.* Vol. 9(2): 207-220.

Cascos I.,López A.,Romo J. (2011). Data Depth in Multivariate Statistics. *Boletín de Estadística e Investigación Operativa*, Vol.27(3):151-174.

Cheng A., Liu R., Luxhoj J. (2000) Monitoring multivariate aviation safety data by data depth: control charts and thresholds systems. *IIE Transactions* 32: 861-872

Cheng, A., Liu, R. and Luxhoj (1997), Thresholds for safety inspection measurements based on control charts, *International Journal of Reliability, Quality and Safety Engineering*. Vol. 4,(2): 205-226.

Febrero M., Oviedo M. Galeano P.,Nieto A., García-Portugués E. (2015). Functional Data Analysis and Utilities for Statistical Computing, (fda.usc) V.1.2.2. CRAN Repository.

Flores M. , Naya S., Fernández R. (2014). Quality Control and reliability, (qcr). V01-18. CRAN Repository.

Ferreya M.(2007). Determinantes del Desempeño Universitario: Efectos Heterogéneos en un Modelo Censurado. Tesis de Maestría. Maestría en Economía. Universidad Nacional de La Plata.

Garbanzo G. (2013). Factores asociados al rendimiento académico en estudiantes universitarios desde el nivel socioeconómico: Un estudio en la Universidad de Costa Rica. *Revista Electrónica Educare On-line version ISSN 1409-4258 Educare* Vol.17(3).

Liu R. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*. Vol 18 (1); 405-414.

Liu R. (1995). Control Charts for Multivariate Processes. *American Statistical Association*. Vol 90(432);1380-138.

Liu R., Parelius J., Singh K. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference. *The Annals of Statistics*. Vol 27(3), 783-858.

Loaiza, R. (2013). Gráficos de Control no paramétricos basados en R-estadísticos con aplicación al caso multivariante. Proyecto de Fin de Carrera. Madrid España. <http://earchivo.uc3m.es/bitstream/handle/10016/18161/proyectoRubenLoaizaA.pdf?sequence=2>

Montero E., Villalobos J., Valverde A. (2007). Factores Institucionales, Pedagógicos, Psicosociales y Sociodemográficos asociados al rendimiento académico en la Universidad de Costa Rica: Un análisis Multinivel. *Revista Electrónica de Investigación y Evaluación Educativa*. Vol.13(2):215-234.

Montgomery D. (2009). *Introduction to Statistical Quality Control*. John Wiley & Sons, Inc. 6th Edition.

ANEXOS

ANEXO 1. CÓDIGO R DE CARTAS: r, S y Q (SIMULACIÓN)

```
## Simulación de 540 puntos normales bivariados (G)
```

```
library("mvtnorm")
```

```
set.seed(12345)
```

```
mu<-c(0,0)
```

```
Sigma<- matrix(c(1,0,0,1),2,2)
```

```
u <- c(2,2)
```

```
S <- matrix(c(4,0,0,4),2,2)
```

```
G <- rmvnorm(540, mean = mu, sigma = Sigma)
```

```
G
```

```
## Simulación de 40 puntos normales bivariados (x)
```

```
x<- rmvnorm(40,mean=u,sigma = S)
```

```
x
```

```
##Separación de Puntos
```

```
H <- G[1:500,]
```

```
H[1:5,]
```

```
I <- G[501:540,]
```

```
I[1:5,]
```

```
#Unión de las matrices I,x
```

```
J <- rbind(I,x)
```

```
J[1:5,]
```

```
## Cálculo de profundidad de SD G (y)
```

```

library(fda.usc)

Sdepth <- mdepth.SD(H,H)

Sdepth

class(Sdepth)

segundadepth <- mdepth.SD(J,H)

segundadepth

names(Sdepth)

class(Sdepth)

m <- unlist(Sdepth)

n <- unlist(segundadepth)

length(n)

length(m)

#####

#####

## Cálculo de rank (Carta r)

lista = NULL

for (i in 1:length(n)){

  r = sum(n[i]>=m)

  lista = c(lista, r)

}

lista

rank <- lista/500

rank[1:10]

length(rank)

a <- seq(1:80)

```

```

plot(a,rank,type="b",col="blue",main="Figura 1: Carta r")
abline(h=0.5,col="red")
abline(h=0.025,col="red")
text(70,0.55,"CL=0.5")
text(10,0.08,"LCL=0.025")

```

```

#####
#####

```

##Cálculo de Q(4) (Carta Q, n=4)

```

n <- rank
Q=NULL
for(i in seq(4,80,4)){
  Q=c(Q,mean(n[(i-3):i]))
}
Q
a<-seq(1:20)
plot(a,Q,type="b",col="blue",main="Figura 2: Carta Q , n=4")
abline(h=0.5,col="red")
abline(h=0.22,col="red")
text(15,0.55,"CL=0.5")
text(5,0.28,"LCL=0.220")

```

```

#####
#####

```

##Cálculo de Q(10) (Carta Q, n=10)

```

n <- rank

```

```

Q=NULL
for(i in seq(10,80,10)){
  Q=c(Q,mean(n[(i-9):i]))
}
Q
a<-seq(1:8)
plot(a,Q,type="b",col="blue",main="Figura 3: Carta Q, n=10")
abline(h=0.5,col="red")
abline(h=0.3193,col="red")
text(6,0.47,"CL=0.5")
text(2,0.29,"LCL=0.3193")
#####
#####

##Carta S
n <- rank
S=NULL
LCL=NULL
for (i in seq(1,80,1)){
  S=c(S,sum(n[1:i]-0.5))
  LCL=c(LCL,-1.96*sqrt((i^2)*((1/500)+(1/i))/12))
}
S
LCL
a <- seq(1:80)
plot(a,S,type="l",col="blue",main="Figura 4: Carta:S")
abline(h=0,col="red")

```

```

lines(a,LCL,col="red")
text(60,-1,"CL=0")
text(20,-5,"LCL=expresión (15)")

#####
#####

##Carta S* (Sast)

S
n <-S
Sast =NULL
for (i in seq(1,80,1)){
  Sast=c(Sast,(n[i]/sqrt((i^2)*((1/500)+(1/i))/12)))
}

Sast
a <- seq(1:80)

plot(a,Sast,type="l",col="blue",main="Figura 5:Carta S*")
abline(h=0,col="red")
abline(h=-1.96,col="red")
text(60,-0.5,"CL=0")
text(20,-2.5,"LCL=-1.96")

```


ANEXO 2.CODIGO R DE MATERIAS: NOTAS PROMEDIO-PORCENTAJE DE APROBACION DE ALUMNOS

```
#####
#####
```

CARGAR DATOS

```
#####
#####
```

```
#rm(list=ls())
```

```
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION - Funciones/funciones.r")
```

```
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION - Funciones/npqcd.r")
```

```
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION -Funciones/npqcs.r")
```

```
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION - Funciones/npqcs.r.r")
```

```
setwd("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION --Notas Materias")
```

```
clusters <- read.table("TodasNotas-Estadistica.txt", header=T,dec = ",",)
```

```
#####
#####
```

ESTADÍSTICA DESCRIPTIVA

```
#####
#####
```

```
library(TSA)
```

```
library(PASWR)
```

```
Materia <- unique(clusters$MATERIA)
```

```
inicio <- 2008
```

```
fin <- 2015
```

```
frecuencia <- 2
```

```
##### NOTAS
```

```

titulo <- "Comportamiento de las Notas promedio por semestre:"
ylab <- "Nota promedio"

for(id in 1:length(Materia)){
  variable <- clusters[clusters$MATERIA == Materia[id,],$Nota
  result <- ADE(Materia[id],variable,inicio,fin,frecuencia,titulo,ylab,type="M")
  print(result)
}

```

PORCENTAJE

```

titulo <- "Porcentaje de aprobación de alumnos por semestre:"
ylab <- "Porcentaje de aprobación"

for(id in 1:length(Materia)){
  variable <- clusters[clusters$MATERIA == Materia[id,],$X._AR
  result <- ADE(Materia[id],variable,inicio,fin,frecuencia,titulo,ylab,type="P")
  print(result)
}

```

```
#####
```

```
#####
```

PROFUNDIDAD

```
#####
```

```
#####
```

```
library(qcr)
```

```
H<-as.matrix(clusters[,c(5,6)])
```

```
data.npqcd <- npqcd(H,H)
```

```

str(data.npqcd)

png("Carta de control todos los datos.png")

res.npqcs <- npqcs.r(data.npqcd,method = "Liu", alpha=0.05,plot = T)

str(res.npqcs)

dev.off()

#identificar el 80% de los datos más profundos donde se considera como muestra de referencia

# a todos los datos de la muestra.

depth <- res.npqcs$depth.data

i <- order(depth,decreasing = T)

#depth.or <- depth[i]

n <- length(depth)

orden <- 1:n

clusters$orden[i] <- orden

n.central <- floor(n*0.8)

referencia <- i[1:n.central]

estado <- vector(length = n)

estado[referencia] <- "R"

estado[-referencia] <- "NR"

clusters$estado <- estado

# Resto

Nreferencia <- i[(n.central+1):n]

m <- length(Nreferencia)

m.bajo <- ceiling(n*0.05)

m.alto <- ceiling(n*0.05)

m.medio <- m - m.bajo-m.alto #(ceiling(n*0.10)

```

```

Grupo <- vector(length = n)

#Bajo

clusters$orden[Nreferencia]

cond.bajo <- (clusters$Nota[Nreferencia] < median(clusters$Nota)) &
  (clusters$X._AR[Nreferencia] < median(clusters$X._AR))
t <- sum(cond.bajo)

#clusters$orden[Nreferencia][cond.bajo]

#Nreferencia[cond.bajo]

m.bajo <- min(t,m.bajo)

ref.bajo <- Nreferencia[cond.bajo][(t-m.bajo+1):t]
Grupo[ref.bajo] <- "De preocupacion"

cond.alto <- (clusters$Nota[Nreferencia] > median(clusters$Nota)) &
  (clusters$X._AR[Nreferencia] > median(clusters$X._AR))
t <- sum(cond.alto)
m.alto <- min(t,m.alto)
ref.alto <- Nreferencia[cond.alto][(t-m.alto+1):t]
Grupo[ref.alto] <- "Informativa"
Grupo[referencia] <- "Esperado"
Grupo[Grupo==F] <- "Advertencia"

# variable que identifica a que grupo pertenece
clusters$Grupo <- Grupo

```

```
#####
#####

# UMBRAL

#####
#####

medianas <- c(median(clusters$Nota),median(clusters$X._AR))

win.graph()

plot(clusters$Nota[i],clusters$X._AR[i],

      main="Identificación de los clusteres de desempeño",

      xlab ="Nota promedio de las materias cursadas",

      ylab="Porcentaje de materias aprobadas",

      col=c(rep("red",n.central),rep("blue",n-n.central)))

legend(7,0.5,c("Esperado","No representativo"),

      text.col=c("red","blue"), bty = "n",pch = 1,

      title.col = "Black",col =c("red","blue"),cex = 0.8)#,fill=c("red","blue"))

#Envolvente convexa

X <- clusters[referencia,c(5,6)]

hpts <- chull(X)

hpts <- c(hpts, hpts[1])

lines(X[hpts, ])

savePlot(paste("Umbral",sep=""),type = "jpg")

dev.off()

#####
#####

# CLASIFICACIÓN

#####
#####
```

variable que identifica a que grupo pertenece

```
punto <- vector(length = n)
punto [Grupo=="Esperado"] <- 15
punto [Grupo=="Informativa"] <- 0
punto [Grupo=="Advertencia"] <- 7
punto [Grupo=="De preocupacion"] <- 8
```

```
clusters$punto <- punto
```

#View(clusters)

```
win.graph()
plot(clusters$Nota[i],clusters$X._AR[i],pch=punto[i],
main="Identificación de los clusteres de desempeño",
xlab ="Nota promedio de las materias cursadas",
ylab="Porcentaje de materias aprobadas",col=c(rep("red",n.central),rep("blue",n-n.central)))
abline(v=medianas[1])
abline(h=medianas[2])
points(x=medianas[1], y=medianas[2],lwd=3,pch=23,col="black",bg="grey")
lines(X[hpts, ])
legend(7,0.65,c("De Preocupación","Advertencia","Informativa","Esperado"),
text.col=c("blue","blue","blue","red"), bty = "n",pch = c(8,7,0,15)
,title.col = "Black",col =c("blue","blue","blue","red"),cex = 0.8)
savePlot("Clasificación",type = "jpg")
dev.off()
```

```
#####
#####
```

CARTAS DE CONTROL r

```
#####
#####
```

```
#####Materia
```

```
Materia <- unique(clusters$MATERIA)
```

```
H<-as.matrix(clusters[referencia,c(5,6)])
```

```
main="Identificación de los clusteres de desempeño"
```

```
xlab ="Nota promedio de las Materias"
```

```
ylab="Porcentaje de estudiantes aprobados"
```

```
umbral <- X[hpts, ]
```

```
for (id in 1:length(Materia)){
```

```
  muestra <- clusters$MATERIA == Materia[id]
```

```
  J<-as.matrix(clusters[muestra,c(5,6)])
```

```
  puntos <- clusters[(clusters$MATERIA == Materia[id]),10]
```

```
  monitoreo(Materia[id], H, muestra,puntos,J,main,xlab,ylab,referencia,umbral)
```

```
}
```

ANEXO 3. CODIGO R DE ESTUDIANTES: NOTAS PROMEDIO-PORCENTAJE DE APROBACION DE MATERIAS

```
#####
#####
# CARGAR DATOS
#####
#####
#rm(list=ls())

source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION -Funciones/funciones.r")
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION - Funciones/npqcd.r")
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION - Funciones/npqcs.r")
source("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION -Funciones/npqcs.r.r")

setwd("C:/Users/Guido/Dropbox/Guido/ALCANCE INFORMACION - Notas de Estudiantes")

clusters <- read.table("Calificaciones totales alumnos.txt", header=T,dec = ",")[-1]

#####
#####
# ESTADÍSTICA DESCRIPTIVA
#####
#####

library(TSA)

library(PASWR)

Matricula <- unique(clusters$MATRICULA)

inicio <- 2008

fin <- 2011

frecuencia <- 2

##### NOTAS
```



```
#####
#####

# PROFUNDIDAD

#####
#####

library(qcr)

H<-as.matrix(clusters[,c(5,6)])

data.npqcd <- npqcd(H,H)

png("Carta de control todos los datos.png")

  res.npqcs <- npqcs.r(data.npqcd,method = "Liu", alpha=0.05,plot = T)

dev.off()

#identificar el 80% de los datos más profundos donde se considera como muestra de referencia

# a todos los datos de la muestra.

depth <- res.npqcs$depth.data

i <- order(depth,decreasing = T)

#depth.or <- depth[i]

n <- length(depth)

orden <- 1:n

clusters$orden[i] <- orden

n.central <- floor(n*0.8)

referencia <- i[1:n.central]

estado <- vector(length = n)

estado[referencia] <- "R"

estado[-referencia] <- "NR"

clusters$estado <- estado
```

Resto

```

Nreferencia <- i[(n.central+1):n]
m <- length(Nreferencia)
m.bajo <- ceiling(n*0.05)
m.alto <- ceiling(n*0.05)
m.medio <- m - m.bajo-m.alto #(ceiling(n*0.10)

```

```

Grupo <- vector(length = n)

```

#Bajo

```

clusters$orden[Nreferencia]

cond.bajo <- (clusters$Nota[Nreferencia] < median(clusters$Nota)) &
  (clusters$X._AR[Nreferencia] < median(clusters$X._AR))
t <- sum(cond.bajo)
#clusters$orden[Nreferencia][cond.bajo]
#Nreferencia[cond.bajo]
m.bajo <- min(t,m.bajo)

```

```

ref.bajo <- Nreferencia[cond.bajo][(t-m.bajo+1):t]
Grupo[ref.bajo] <- "De preocupacion"

```

```

cond.alto <- (clusters$Nota[Nreferencia] > median(clusters$Nota)) &
  (clusters$X._AR[Nreferencia] > median(clusters$X._AR))
t <- sum(cond.alto)
m.alto <- min(t,m.alto)

```

```

ref.alto <- Nreferencia[cond.alto][(t-m.alto+1):t]

Grupo[ref.alto] <- "Informativa"

Grupo[referencia] <- "Esperado"
Grupo[Grupo=="F"] <- "Advertencia"

# variable que identifica a que grupo pertenece

clusters$Grupo <- Grupo

#####
#####

# UMBRAL

#####
#####

medianas <- c(median(clusters$Nota),median(clusters$X._AR))

win.graph()

plot(clusters$Nota[i],clusters$X._AR[i],

      main="Identificación de los clusteres de desempeño",

      xlab="Nota promedio de las Materias cursadas",

      ylab="Porcentaje de Materias aprobadas",

      col=c(rep("red",n.central),rep("blue",n-n.central)))

legend(7,0.5,c("Esperado", "No representativo"),

      text.col=c("red", "blue"), bty = "n",pch = 1,

      title.col = "Black",col =c("red", "blue"),cex = 0.8)#,fill=c("red", "blue"))

#Envolvente convexa

X <- clusters[referencia,c(5,6)]

```

```

hpts <- chull(X)
hpts <- c(hpts, hpts[1])
lines(X[hpts, ])
savePlot("Umbral",type = "jpg")
dev.off()

#####
#####

# CLASIFICACIÓN

#####
#####

# variable que identifica a que grupo pertenece

punto <- vector(length = n)
punto [Grupo=="Esperado"] <- 15
punto [Grupo=="Informativa"] <- 0
punto [Grupo=="Advertencia"] <- 7
punto [Grupo=="De preocupacion"] <- 8

clusters$punto <- punto

#View(clusters)

win.graph()
plot(clusters$Nota[i],clusters$X._AR[i],pch=punto[i],
      main="Identificación de los clusteres de desempeño",
      xlab ="Nota promedio de las Materias cursadas",
      ylab="Porcentaje de Materias aprobadas",col=c(rep("red",n.central),rep("blue",n-n.central)))

abline(v=medianas[1])
abline(h=medianas[2])

points(x=medianas[1], y=medianas[2],lwd=3,pch=23,col="black",bg="grey")

```

```

lines(X[hpts, ])

legend(8,0.65,c("De Preocupación", "Advertencia", "Informativa", "Esperado"),
      text.col=c("blue", "blue", "blue", "red"), bty = "n",pch = c(8,7,0,15)
      ,title.col = "Black",col =c("blue", "blue", "blue", "red"),cex = 0.8)

savePlot("Clasificacion",type = "jpg")

dev.off()

#####
#####

# CARTAS DE CONTROL r

#####
#####

#####Matricula

Matricula <- unique(clusters$MATRICULA)

H<-as.matrix(clusters[referencia,c(5,6)])

main="Identificación de los clusteres de desempeño"

xlab ="Nota promedio de las Materias cursadas"

ylab="Porcentaje de Materias aprobadas"

umbral <- X[hpts, ]

for (id in 1:length(Matricula)){

  muestra <- clusters$MATRICULA == Matricula[id]

  J<-as.matrix(clusters[muestra,c(5,6)])

  puntos <- clusters[(clusters$MATRICULA == Matricula[id]),11]

  monitoreo(Matricula[id], H, muestra,puntos,J,main,xlab,ylab,referencia,umbral)

}

```

ANEXO 4. CÓDIGO R DE FUNCIONES.

```
#####  
#####
```

```
# ADE
```

```
#####  
#####
```

```
ADE <- function(Materia, variable, inicio, fin, frecuencia,titulo,ylab,type){  
  variable.ts <- ts(data = variable,start = inicio,end = fin,frequency = frecuencia)  
  
  win.graph()  
  
  res <- EDA(variable)  
  
  savePlot(paste(type,"EDA-",Materia,sep=""),type = "jpg")  
  
  dev.off()  
  
  win.graph()  
  
  plot(variable.ts,  
        main=paste(titulo,Materia),  
        xlab ="Periodo",  
        ylab=ylab,type="o")  
  
  savePlot(paste(type,"Serie-",Materia,sep=""),type = "jpg")  
  
  dev.off()  
  
  return(res)  
}
```

```
#####  
#####
```

```
# MONITOREO
```

```
#####
#####
#ID=Materia[id]

monitoreo <- function(ID, H, muestra,puntos,J,main,xlab,ylab,referencia,umbral){

  data.npqcd <- npqcd(J,H)

  png(paste("C",ID,".png",sep=""))

  res.npqcs <- npqcs.r(data.npqcd,method = "Liu", alpha=0.05,plot = T)

  points(1:length(res.npqcs$statistics),res.npqcs$statistics,pch=puntos,cex=1.5,col="blue")

  legend("topleft",c("Advertencia","Informacional","De preocupación","Esperado"),

        bty = "n",pch = c(7,0,8,15)

        ,title.col = "blue",cex = 0.8)

  dev.off()

  png(paste("B",ID,".png",sep=""))

  plot(clusters[referencia,5],clusters[referencia,6],

        main=main,

        xlab =xlab,

        ylab=ylab,ylim = c(0,1),xlim = c(0,10))

  lines(umbral)

  points(clusters[muestra,5],clusters[muestra,6],pch=puntos,cex=1)

  legend(8,0.7,c(as.character(ID),"Advertencia","Informacional","De preocupación","Esperado"),

        bty = "n",pch = c(15,7,0,8,1)

        ,title.col = "Black",cex = 0.8)

  dev.off()

}
```


ANEXO 5. CÓDIGO R DE OBJETO: npqcd

```

#-----#
#
#          #
#    QUALITY CONTROL AND RELIABILITY IN R          #
#          #
# An R package for statistical in-line quality control.          #
#          #
# Written by: Miguel A. Flores          #
#    Student Master of Statistical Techniques          #
#    University of The Corunia, SPAIN          #
#    ma.flores@outlook.com          #
#          #
#-----#
#-----#
# Main function to create a 'npqcd' object
#-----#

##' Non Parametric Quality Control Data

##'

##' Create an object of class 'npqcd' to perform statistical quality control.

##' This object be used to plot Non Parametric Multivariate Control Charts.

##'

##' @aliases npqcd

##' @param x a matrix or data-frame or array where it should contain data. Dimension has to
be the same as that of the observations.

##' @param G The x as a matrix, data frame or list. If it is a matrix or data frame, then each
row is viewed as one multivariate observation.

##' @param data.name a string that specifies the title displayed on the plots.

```

```

## If not provided is taken from the name of the object x.
## @export
## @examples
##
## library(qcr)
##
## set.seed(356)
## data <- matrix(rnorm(999), nc = 3)
## x <- rexp(999,0.5)
## x <- matrix(x,ncol=3)
## data.npqcd <- npqcd(data,x)
## str(data.npqcd)

```

```

npqcd <- function(x, G = NULL, data.name = NULL)
#.....
{
  if (!is.matrix(x) & !is.data.frame(x) & !is.array(x))
    stop("object must be a matrix or data.frame or array")

  if (is.null(G)) G <- x

  if (!is.matrix(G) & !is.data.frame(G) & !is.list(G))
    stop("object must be a matrix or data.frame or list")

  if (class(x) == "matrix" || class(x) == "data.frame") {
    p <- ncol(x) # quality characteristics
    m <- nrow(x) # number of samples or observations

```

```
names <- colnames(x)
x <- array(data.matrix(x),c(m,p,1))
colnames(x) <- names
}

if(is.list(G))
{
  m=length(G)
  n=length(G[[1]])
  for(i in 1:m){
    if(length(G[,i])!=n)
      stop("When using a list, each element must be a matrix of the same length.")
  }
}

if (is.data.frame(G)) as.matrix(G)

if (is.matrix(G)){
  n=dim(G)
  m=dim(x)[1:2]
  if(sum(n!=m)==2)
    stop("The object x must be of the same dimension than observations.")
}

if (is.null(data.name))
  data.name <- "DATA"
```

```
result <- list(x = x, G = G)

attr(result, "data.name") <- data.name
attr(result, "type.data") <- "Multivariate"

oldClass(result) <- c("npqcd", "list")

return(result)

} # npqcd
#.....
```

ANEXO 6. CODIGO R DE OBJETO: npqcs

```

#-----#
#
#          #
#    QUALITY CONTROL AND RELIABILITY IN R          #
#          #
# An R package for statistical in-line quality control.      #
#          #
# Written by: Miguel A. Flores                               #
#    Student Master of Statistical Techniques                #
#    University of The Corunia, SPAIN                       #
#    miguel.flores@epn.edu.ec                              #
#          #
#-----#
#-----#
# Main function to create a 'npqcs' object
#-----#
##' Create an object of class 'npqcs' to perform statistical quality control.
##' This function is used to compute statistics required for to plot Non Parametric Multivariate
Control Charts
##'
##' @aliases npqcs summary.npqcs print.npqcs
##' @param x Object npqcd (Non Parametric Multivariate Quality Control Data)
##' @param method Character string which determines the depth function used.
##' method can be "Tukey" (the default), "Liu", "Mahalanobis", "RP" Random Project or "LD"
Likelihood depth.
##' @param ... arguments passed to or from methods.

```

```
##' @export
```

```
npqcs <- function(x, method = c("Tukey", "Liu", "Mahalanobis", "RP", "LD"), ...)
```

```
#.....
```

```
{
```

```
if(is.null(x) || !inherits(x, "npqcd"))
```

```
  stop("x must be an objects of class (or extending) 'npqcd'")
```

```
method <- match.arg(method)
```

```
#npqcd <- data.npqcd
```

```
npqcd <- x
```

```
x <- npqcd[[1]]
```

```
G <- npqcd[[2]]
```

```
depth.data.x <- matrix(,nrow=dim(x)[1],ncol=dim(x)[3])
```

```
for (i in 1:dim(x)[3]){
```

```
  depth.data.x[,i] <- f.depth(x[,i], G, method)
```

```
}
```

```
depth.data.G <- f.depth(G, G, method)
```

```
rank.depth.x <- vector (length = length(depth.data.x))
```

```
depth.data.x <- as.vector(depth.data.x)
```

```
for(i in 1:length(depth.data.x))
```

```
{
```

```

rank.depth.x[i]<-(sum(depth.data.G <= depth.data.x[i]))/length(depth.data.G)
}

rank.depth.x <- matrix(rank.depth.x, nrow=dim(x)[1], ncol=dim(x)[3])
depth.data.x <- matrix(depth.data.x, nrow=dim(x)[1], ncol=dim(x)[3])

result <- list (npqcd = npqcd, depth.data = depth.data.x, rank.depth = rank.depth.x)

oldClass(result) <- c("npqcs")

return(result)
} # npqcs
#.....

##' @export
print.npqcs <- function(x, ...) str(x,1)
#.....

##' @export
summary.npqcs <- function(object, ...)
#.....
{
type <- object$type
r <- object$statistics
cat("\nSummary of group statistics:\n")
print(summary(r))
cat("\nNumber of quality characteristics: ", dim(object$npqcd[[1]])[2])
cat("\nNumber of samples or observations: ", dim(object$npqcd[[1]])[1])

```

```

cat("\nNumber of observations or sample size: ", dim(object$npqcd[[1]])[3])

limits <- object$limits

if (!is.null(limits))
{ cat("\nControl limits:", "\n")
  print(limits)
}

if (length(object$violations)== 0){
  cat("\nNumber beyond limits: 0", "\n")
}
else {cat("\nBeyond limits of control:", "\n")
  print(object$statistics[object$violations])
}

invisible()

#.....
} # summary.npqcs

f.depth <- function(x, G = NULL, method = c("Tukey", "Liu", "Mahalanobis", "RP", "LD")){

if(is.null(G)) G <- x

if (!is.matrix(G)) stop("object must be a matrix")

method <- match.arg(method)

```



```

depth.data <- switch(method,"Tukey" = {mdepth.TD(x,G)$dep},
                    "Liu" = {mdepth.SD(x,G)$dep},
                    "Mahalanobis" = {mdepth.MhD(x,G)$dep},
                    "RP" = {mdepth.RP(x,G)$dep},
                    "LD" = {mdepth.LD(x,G)$dep})

return (depth.data)
}

# npqcs.add function
#-----
##' npqcs.add Add a matrix, data.frame or array object with a npqcs object
##'
##' This function is used to join two objects of type matrix, data.frame or array and npqcs.
##'
##' @param x Object type npqcs
##' @export
##'

npqcs.add <- function(x, ...){
  UseMethod("npqcs.add")
}

##' @rdname npqcs.add
##' @method npqcs.add default
##' @param value Object type data.frame, matrix or array
##' @param ... arguments to be passed to or from methods.
##' @export

```

```
npqcs.add.default <- function(x, value, ...){

  if (!inherits(x, "npqcs"))
    stop("object must be npqcs")

  if (!is.matrix(value) & !is.data.frame(value) & !is.array(value))
    stop("object must be a matrix, data.frame or array")

  data <- x$npqcd$x
  limits <- x$limits
  data.name = x$data.name
  method = x$method
  type <- x$type
  G <- x$npqcd$G

  if (class(value) == "matrix" || class(value) == "data.frame") {
    p <- ncol(value) # quality characteristics
    m <- nrow(value) # number of samples or observations
    names <- colnames(value)
    value <- array(data.matrix(value),c(m,p,1))
    colnames(value) <- names
  }

  n1 <- dim(data)[3]
  n2 <- dim(value)[3]
```

```

m1 <- dim(data)[1]
m2 <- dim(value)[1]
m <- m1 + m2

k1 <- dim(data)[2]
k2 <- dim(value)[2]

if (n1 == n2) stop("The samples must be of the same dimension")
if (k1 == k2) stop("The samples must be of the same quality characteristics")

xx <- array(dim = c(m,k1,n1))
for (i in 1:n1 ){
  xx[,i] <- cbind(data[,i],value[,i])
}

z.npqcd <- npqcd(x = xx, G = G, data.name = data.name)

z.npqcs <- switch(type,
  "r" = npqcs.r.npqcd(x = z.npqcd, limits = limits, method = method),
  "Q" = npqcs.Q.npqcd(x = z.npqcd, limits = limits, method = method),
  "S" = npqcs.S.npqcd(x = z.npqcd, limits = limits, method = method),
  NULL)
result <- z.npqcs
}

#-----
# npstate.control

```

```

#-----
##' This function is used for the state of the process is under control.
##' Were removed from the sample observations that violate the rules of a process under
control
##' @aliases npstate.control
##' @param x Object npqcd (Quality Control Statistical Non Parametric)
##' @param control a logical value indicating whether the initial sample comes from a process
under control.
##' @export
##' @examples
##'
##' ##
##' ## Continuous data
##' ##
##' library(qcr)
##' set.seed(356)
##' mu<-c(0,0)
##' Sigma<- matrix(c(1,0,0,1),nrow = 2,ncol = 2)
##' mu <- c(2,2)
##' S <- matrix(c(4,0,0,4),nrow = 2,ncol = 2)
##' G <- rmvnorm(540, mean = mu, sigma = Sigma)
##' x<- rmvnorm(40,mean=mu,sigma = S)
##' x <- rbind(G[501:540,],x)
##' M <- G[1:500,]
##' data.npqcd <- npqcd(x,M)
##' str(data.npqcd)
##' res.npqcs <- npqcs.r(data.npqcd,method = "Liu", alpha=0.025)
##' str(res.npqcs)

```

```

## summary(res.npqcs)

## plot(res.npqcs)

## new.npqcd <- npstate.control(x = res.npqcs)

## res.npqcs <- npqcs.r(new.npqcd)

## summary(res.npqcs)

## plot(res.npqcs)

npstate.control <- function(x, control = FALSE)

#.....

{

if (length(x$violations)>0){

  ii<-x$violations

  n <- dim(x$npqcd[[1]])[3]

  m <- dim(x$npqcd[[1]])[1]

  k <- dim(x$npqcd[[1]])[2]

  xx <- array(dim = c(m-length(ii),k,n))

  for (i in 1:n){

    xx[,i] <- x$npqcd[[1]][,i][-ii,]

  }

  if (control == TRUE){

    G <- x$npqcd[[2]]

  }else{

    G <- x$npqcd[[2]][-ii,]

  }

  result <- npqcd(x = xx, G, data.name = x$data.name)

```

```
} else {  
  cat("The process is under control")  
}
```

```
oldClass(result) <- c("npqcd", "list")
```

```
invisible(result)
```

```
} #npsate.control
```

```
#.....
```

ANEXO 7. CODIGO R DE OBJETO: npqcs.r.r

```

#-----#
#
#          #
#          QUALITY CONTROL STATISTICS IN R          #
#          #
# An R package for statistical in-line quality control.      #
#          #
# Written by: Miguel A. Flores          #
#          Student Master of Statistical Techniques          #
#          University of The Corunia, SPAIN          #
#          ma.flores@outlook.com          #
#          #
#-----#
#-----#
# r chart
#-----#
##' Function to plot r chart
##'
##' This function is used to compute statistics required by the r chart.
##'
##' @param x an R object (used to select the method). See details.
##' @param ... arguments passed to or from methods.
##' @export
##' @references Regina Liu (1995)
##' @examples
##'

```

```

### ## Continuous data

### ##

### library(qcr)

### set.seed(356)

### mu<-c(0,0)

### Sigma<- matrix(c(1,0,0,1),nrow = 2,ncol = 2)

### mu <- c(2,2)

### S <- matrix(c(4,0,0,4),nrow = 2,ncol = 2)

### G <- rmvnorm(540, mean = mu, sigma = Sigma)

### x<- rmvnorm(40,mean=mu,sigma = S)

### x <- rbind(G[501:540,],x)

### M <- G[1:500,]

### data.npqcd <- npqcd(x,M)

### str(data.npqcd)

### res.npqcs <- npqcs.r(data.npqcd,method = "Liu", alpha=0.025)

### str(res.npqcs)

### summary(res.npqcs)

### plot(res.npqcs,title =" r Control Chart")

npqcs.r <- function(x, ...) {
  UseMethod("npqcs.r")
}

### @rdname npqcs.r

### @method npqcs.r default

### @inheritParams npqcd

### @param limits a two-values vector specifying the control limits lower and central.

```



```

##' @param method Character string which determines the depth function used.

##' method can be "Tukey" (the default), "Liu", "Mahalanobis", "RP" Random Project or "LD"
Likelihood depth.

##' @param alpha it is the the significance level (0.01 for default)

##' @param plot a logical value indicating should be plotted.

##' @export

npqcs.r.default <- function(x, G, data.name = NULL, limits = NULL,
                           method = c("Tukey","Liu","Mahalanobis","RP","LD"), alpha = 0.01, plot = FALSE,
                           ...)
#.....
{

method <- match.arg(method)

obj <- npqcd(x = x , G = G, data.name = data.name)

result <- npqcs.r.npqcd(x = obj, data.name = data.name, limits = NULL, method = method, alpha =
alpha,
                       plot = plot, ...)

return(result)
} # npqcs.r.default
#.....

##' @rdname npqcs.r

##' @method npqcs.r npqcd

##' @export

```

```
##'
```

```
# x <-data.npqcd
```

```
# method="Tukey"
```

```
# alpha=0.5
```

```
npqcs.r.npqcd <- function(x, data.name, limits = NULL, method =
c("Tukey","Liu","Mahalanobis","RP","LD"), alpha = 0.01, plot = FALSE, ...)
```

```
#.....
```

```
{
```

```
  if(is.null(x) || !inherits(x, "npqcd"))
```

```
    stop("x must be an objects of class (or extending) 'npqcd'")
```

```
  if(dim(x[[1]])[3]!=1)
```

```
    stop("The dimension of object x must be valid for control chart r")
```

```
  method <- match.arg(method)
```

```
  if (is.null(limits)){
```

```
    central<-0.5
```

```
    limits <- c(lcl = alpha, cl = central)
```

```
  }else{
```

```
    limits <- limits
```

```
  }
```

```
  npqcs <- npqcs(x, method)
```

```
  statistics <- npqcs$rank.depth
```

```
  depth.data <- npqcs$depth.data
```

```
violations <- which(statistics < alpha)
```

```
data.name <- attr(x, "data.name")
```

```
result <- list(npqcd = x, type = "r", depth.data = depth.data, statistics = statistics, alpha = alpha,  
             limits = limits, data.name = data.name, method = method,  
             violations = violations)
```

```
oldClass(result) <- c("npqcs.r", "npqcs")
```

```
if(plot) plot(result, ...)
```

```
return(result)
```

```
#.....
```

```
} # npqcs.r.npqcd
```

```
#.....
```

ANEXO 8. ESTADÍSTICA DESCRIPTIVA DE NOTAS PROMEDIO DE MATERIA.

Figura A.8. 1 AES1

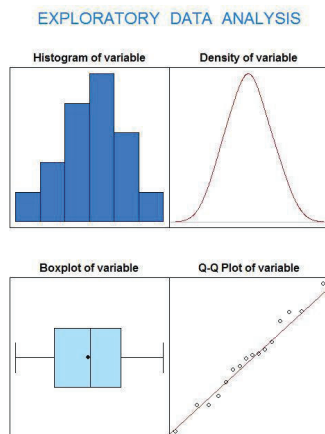


Figura A.8. 2 AES2

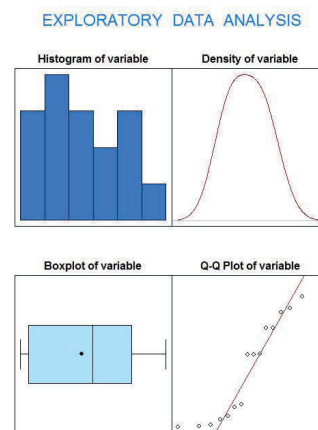


Figura A.8. 3 AES3

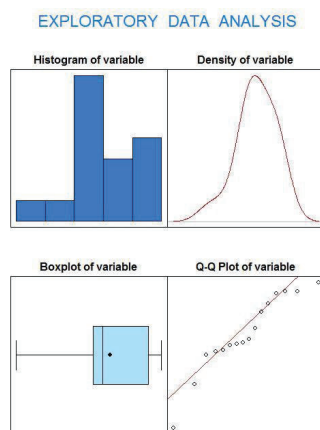


Figura A.8. 4 AES4

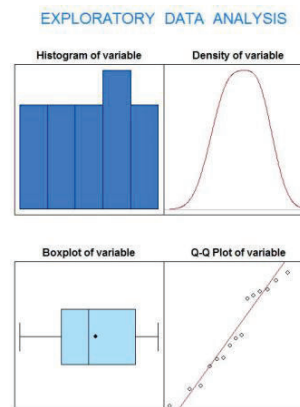


Figura A.8. 5 AES5

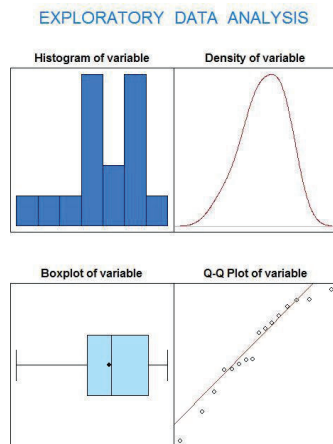


Figura A.8. 6 AES6

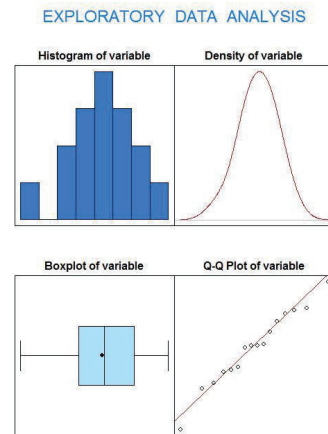
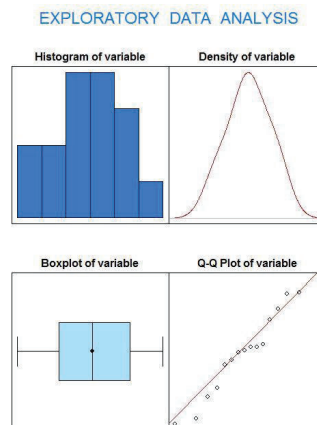


Figura A.8. 7 AES7



ANEXO 9. ESTADÍSTICOS DESCRIPTIVOS DE NOTAS DE MATERIAS.

MATERIAS-NOTAS

[1] "variable"-AES1

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	4.950	5.560	6.089	6.121	6.089	6.604	7.275	0.624
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.389	0.156	1.044	2.325	-0.897	0.056	0.978			

[1] "variable"-AES2

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	5.963	6.077	6.831	6.980	6.831	7.524	8.012	0.734
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.539	0.184	1.447	2.049	-1.687	0.109	0.046			

[1] "variable"-AES3

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	3.579	5.655	6.121	5.930	6.121	7.152	7.517	1.038
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
1.077	0.260	1.497	3.938	-0.050	-0.648	0.172			

[1] "variable"-AES4

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	5.129	5.769	6.305	6.200	6.305	6.922	7.278	0.700
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.490	0.175	1.153	2.149	-1.530	-0.115	0.250			

[1] "variable"-AES5

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	5.323	6.639	7.047	7.078	7.047	7.763	8.129	0.801
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.642	0.200	1.124	2.806	-0.812	-0.493	0.436			

[1] "variable"-AES6

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	5.771	6.287	6.500	6.515	6.500	6.782	7.086	0.329
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.108	0.082	0.495	1.315	-0.480	-0.296	0.942			

[1] "variable"-AES7

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	5.217	5.925	6.483	6.491	6.483	7.118	7.679	0.735
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.540	0.184	1.193	2.462	-1.073	-0.098	0.629			

ANEXO 10. ESTADÍSTICA DESCRIPTIVA DE PORCENTAJE DE APROBACION DE ALUMNOS POR MATERIA.

Figura A.10. 1 PAES1

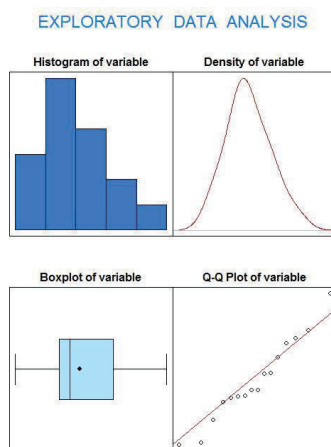


Figura A.10. 2 PAES2

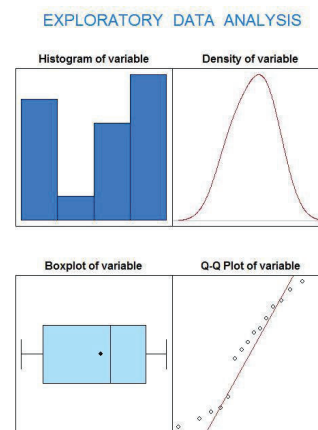


Figura A.10. 3 PAES3

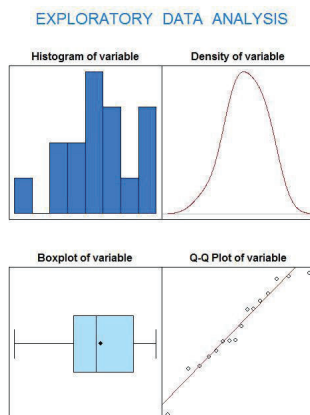


Figura A.10. 4 PAES4

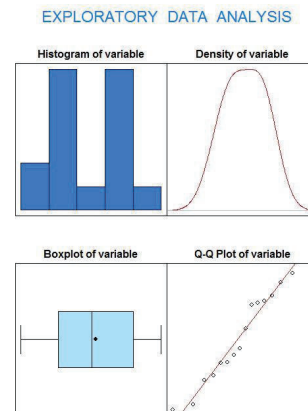


Figura A.10. 5 PAES5

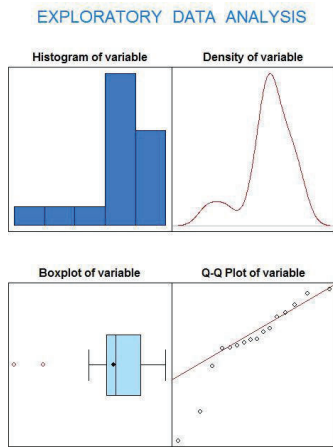


Figura A.10. 6 PAES6

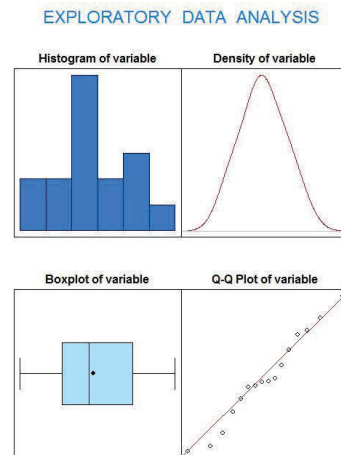
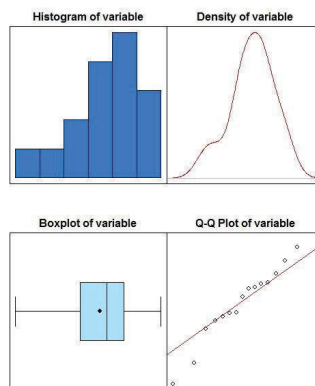


Figura A.10. 7 PAES7



ANEXO 11. ESTADÍSTICOS DESCRIPTIVOS DE PORCENTAJE DE APROBACIÓN DE ALUMNOS.

MATERIAS-PORCENTAJE DE APROBACIÓN DE ALUMNOS

[1] "variable"- P-AES1

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.512	0.631	0.687	0.661	0.687	0.778	0.924	0.111
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.012	0.028	0.147	0.412	-0.628	0.319	0.746		

[1] "variable"- P-AES2

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.617	0.674	0.827	0.852	0.827	0.945	1.000	0.134
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.018	0.034	0.271	0.383	-1.523	-0.254	0.126		

[1] "variable"- P-AES3

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.276	0.550	0.676	0.653	0.676	0.825	0.931	0.181
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.033	0.045	0.275	0.655	-0.726	-0.329	0.608		

[1] "variable"- P-AES4

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.535	0.635	0.735	0.726	0.735	0.836	0.910	0.123
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.015	0.031	0.201	0.375	-1.480	-0.101	0.318		

[1] "variable"- P-AES5

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.545	0.822	0.844	0.849	0.844	0.925	1.000	0.119
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.014	0.030	0.103	0.455	0.516	-1.020	0.054		

[1] "variable"- P-AES6

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.621	0.694	0.747	0.740	0.747	0.815	0.889	0.077
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.006	0.019	0.121	0.268	-1.015	0.053	0.846		

[1] "variable"- P-AES7

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
16.000	0.000	0.444	0.691	0.766	0.794	0.766	0.858	1.000	0.148
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.022	0.037	0.167	0.556	-0.464	-0.457	0.727		

ANEXO 12. COMPORTAMIENTO EN EL TIEMPO DE LAS NOTAS PROMEDIO POR MATERIA

Figura A.12. 1 SM-AES1

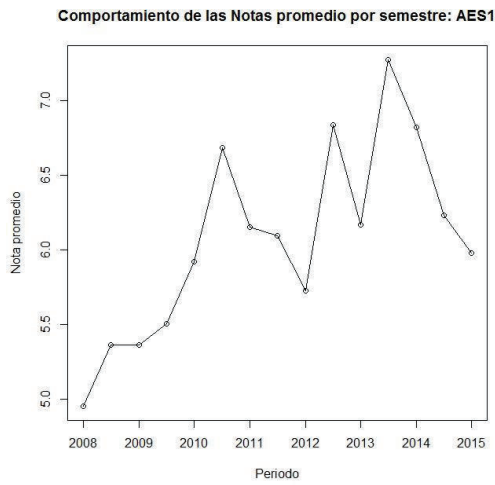


Figura A.12. 2 SM-AES2

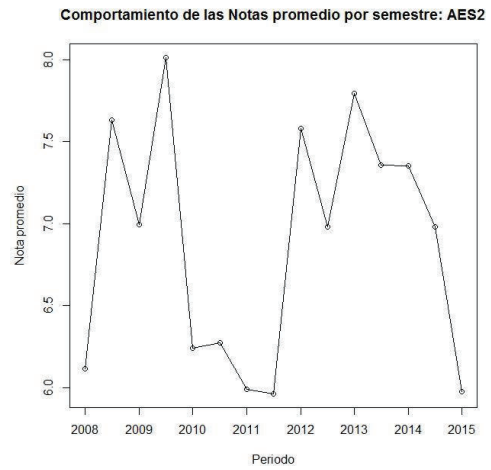


Figura A.12. 3 SM-AES3

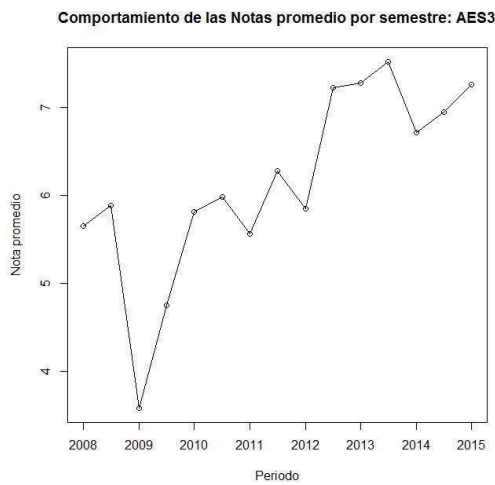


Figura A.12. 4 SM-AES4

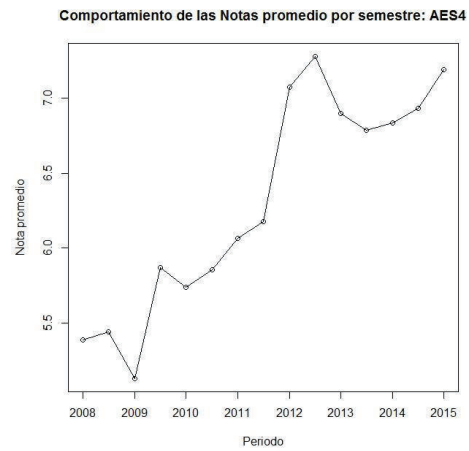


Figura A.12. 5 SM-AES5

Comportamiento de las Notas promedio por semestre: AES5

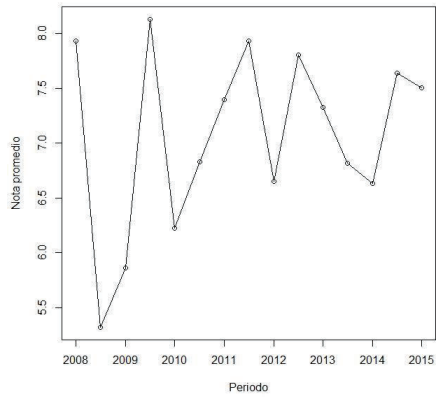


Figura A.12. 6 SM-AES6

Comportamiento de las notas promedio por semestre: AES6

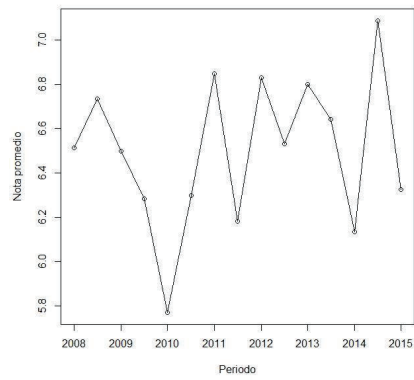
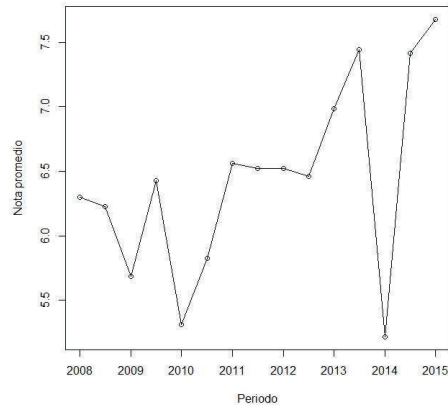


Figura A.12. 7 SM-AES7

Comportamiento de las Notas promedio por semestre: AES7



ANEXO 13. COMPORTAMIENTO EN EL TIEMPO DEL PORCENTAJE DE APROBACIÓN DE ALUMNOS POR MATERIA

Figura A.13. 1 SP-AES1

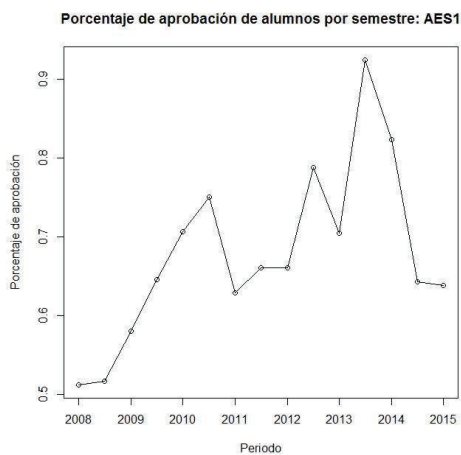


Figura A.13. 2 SP-AES2

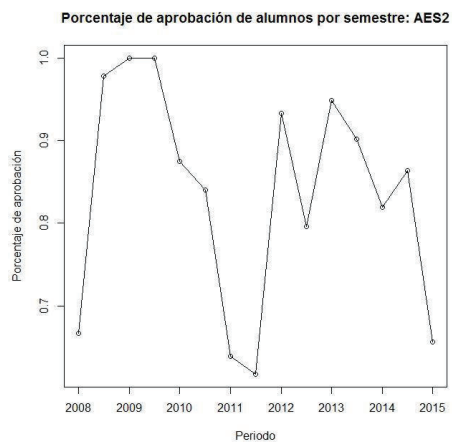


Figura A.13. 3 SP-AES3

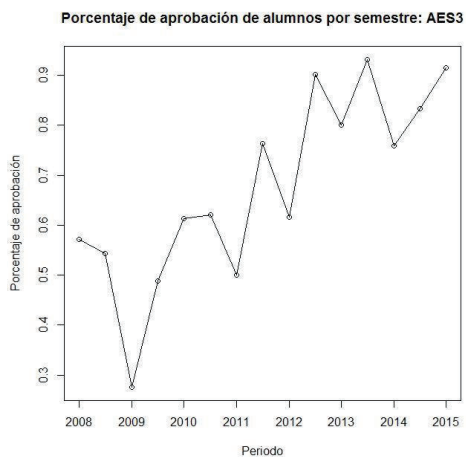


Figura A.13. 4 SP-AES4

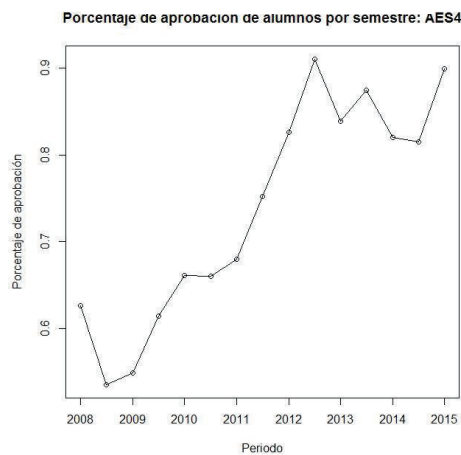


Figura A.13. 5 SP-AES5

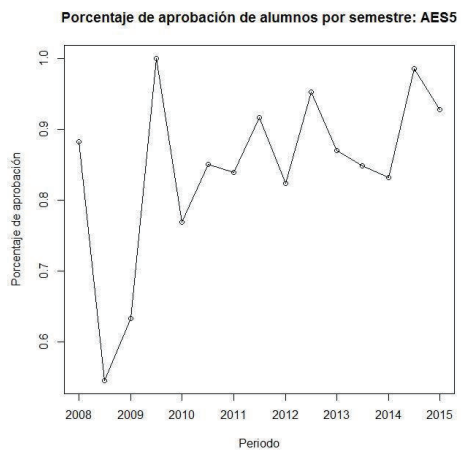


Figura A.13. 6 SP-AES6

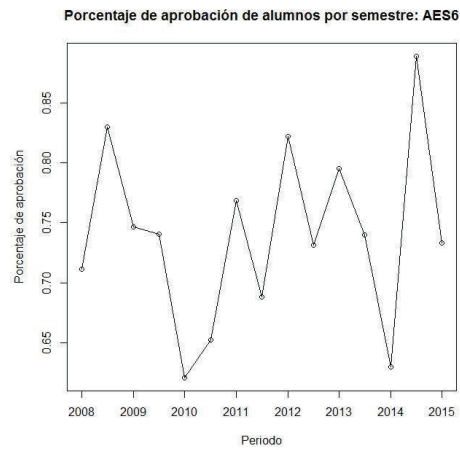
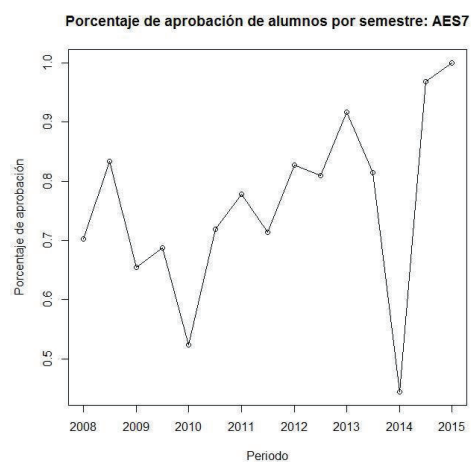


Figura A.13. 7 SP-AES7



ANEXO 14. MATERIAS. GRÁFICOS BIVARIADOS DE NOTAS - PORCENTAJE DE APROBACIÓN DE ALUMNOS

Figura A.14. 1 B-AES1

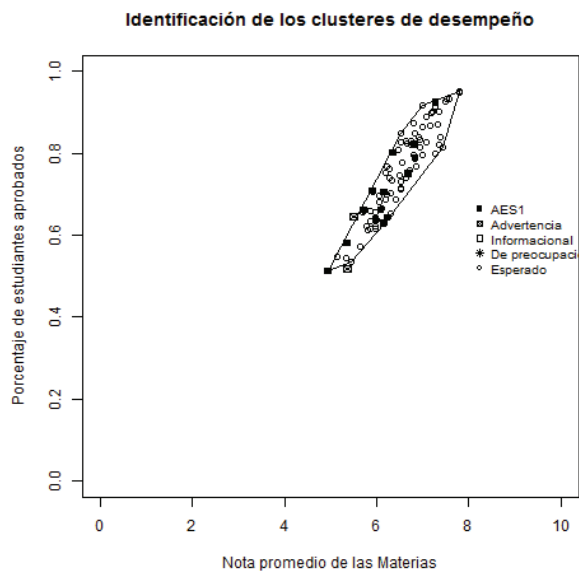


Figura A.14. 2 B-AES2

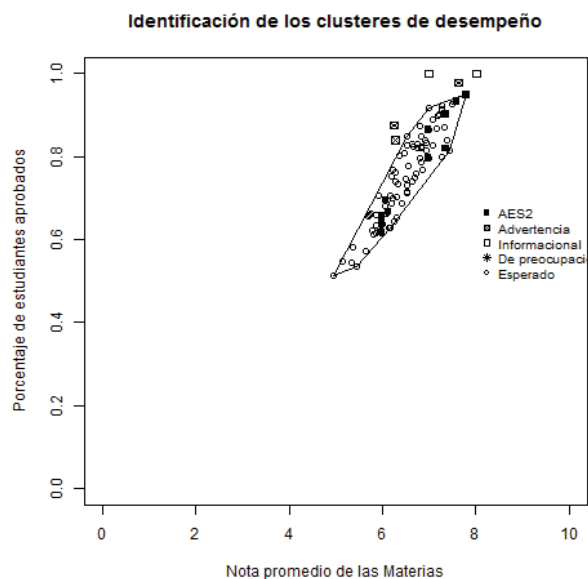


Figura A.14. 3 B-AES3

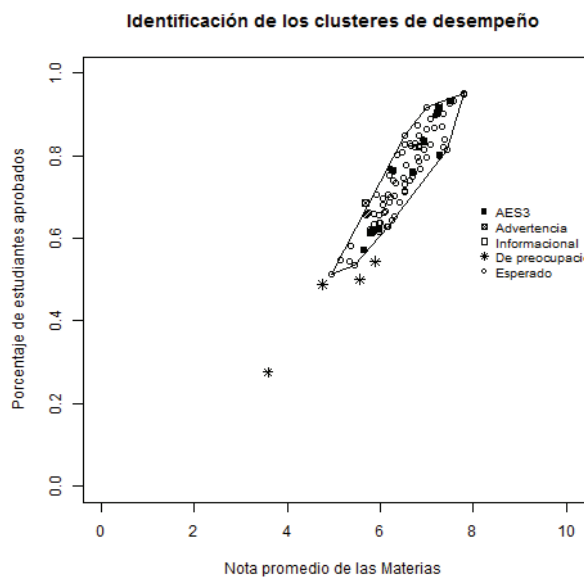


Figura A.14. 4 B-AES4

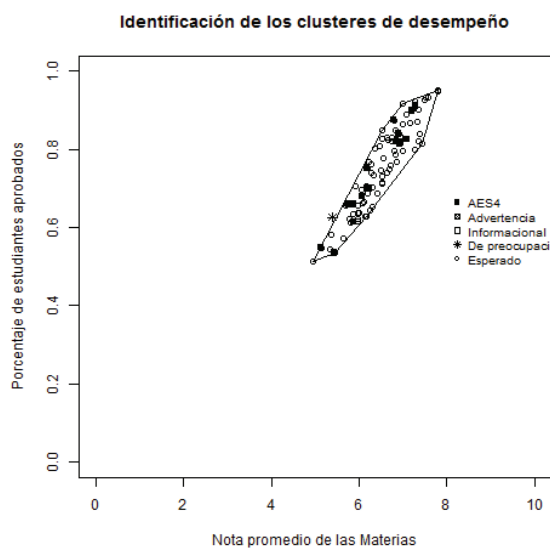


Figura A.14. 5 B-AES5

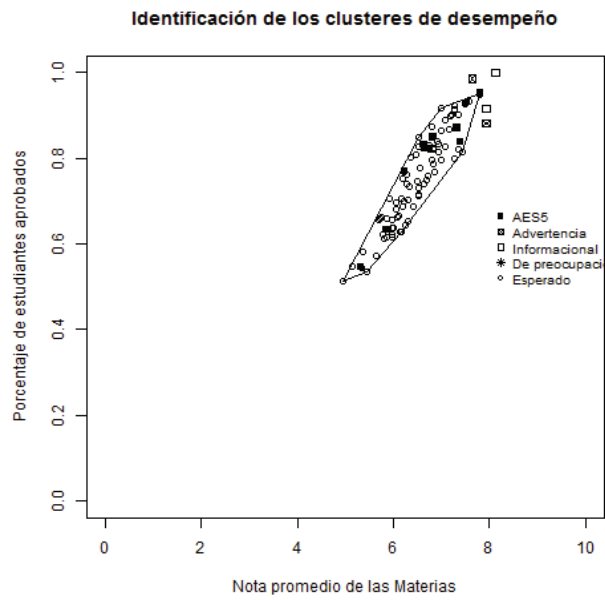


Figura A.14. 6 B-AES6

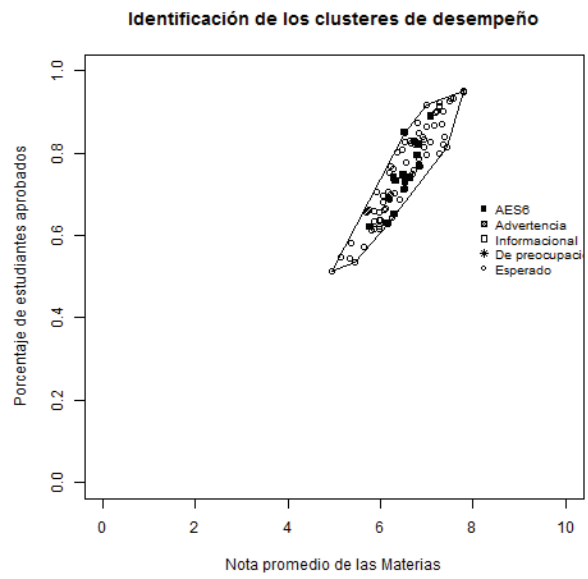
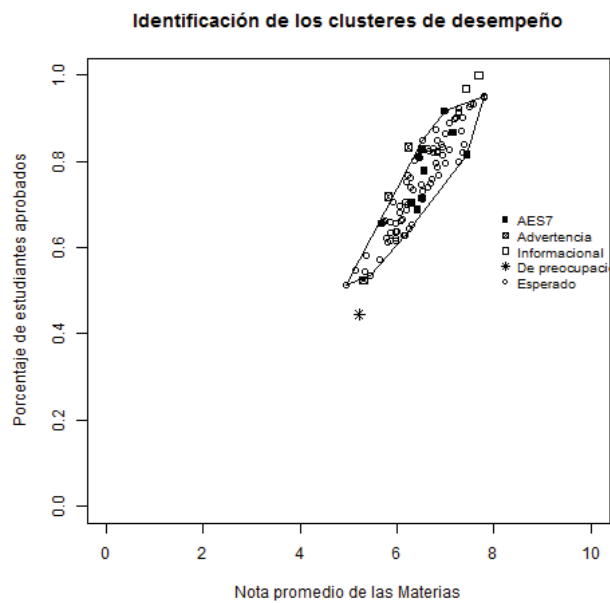


Figura A.14. 7 B-AES7



ANEXO 15. MATERIAS. CARTAS DE CONTROL r

Figura A.15 1 r-AES1

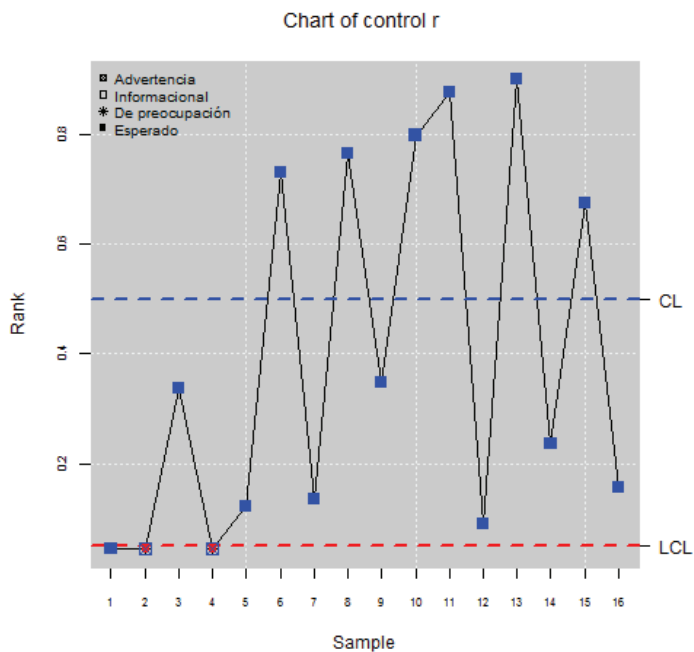


Figura A.15 2 r-AES2

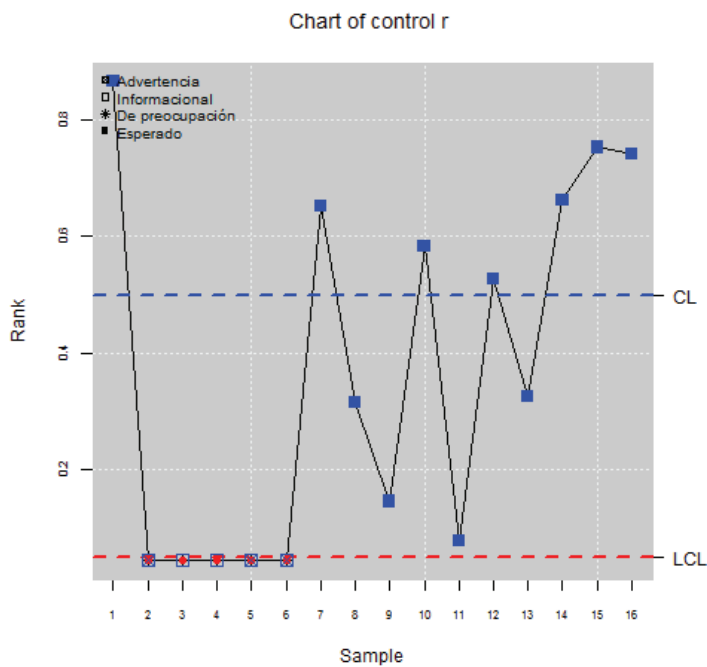


Figura A.15 3 r-AES3

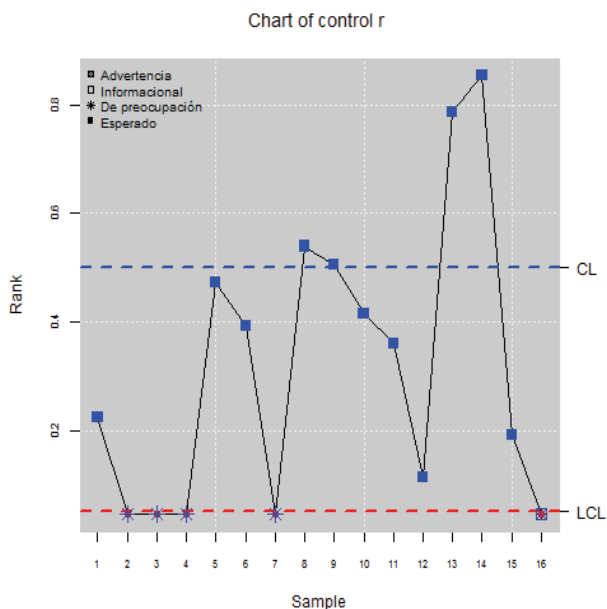


Figura A.15 4 r-AES4

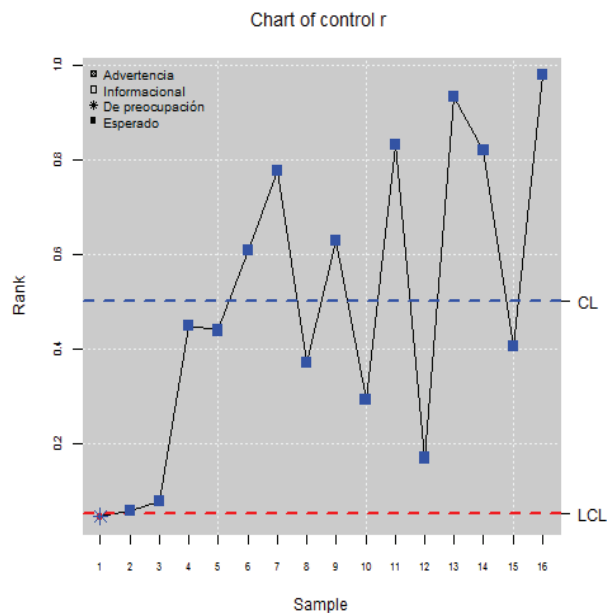


Figura A.15 5 r-AES5

Chart of control r

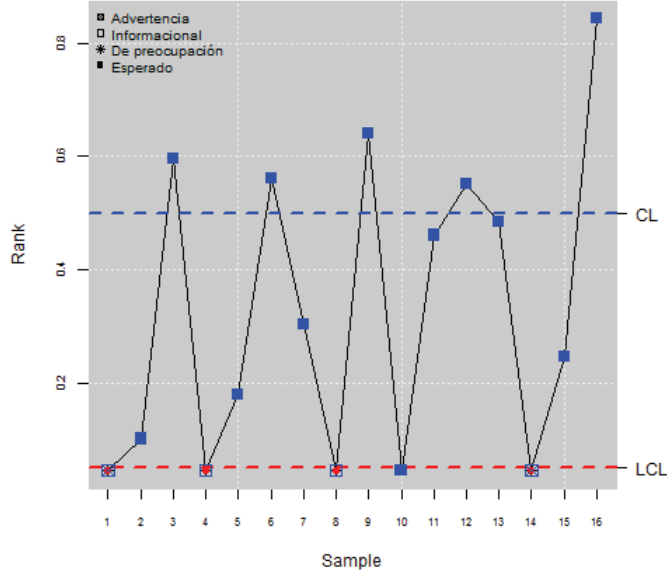


Figura A.15 6 r-AES6

Chart of control r

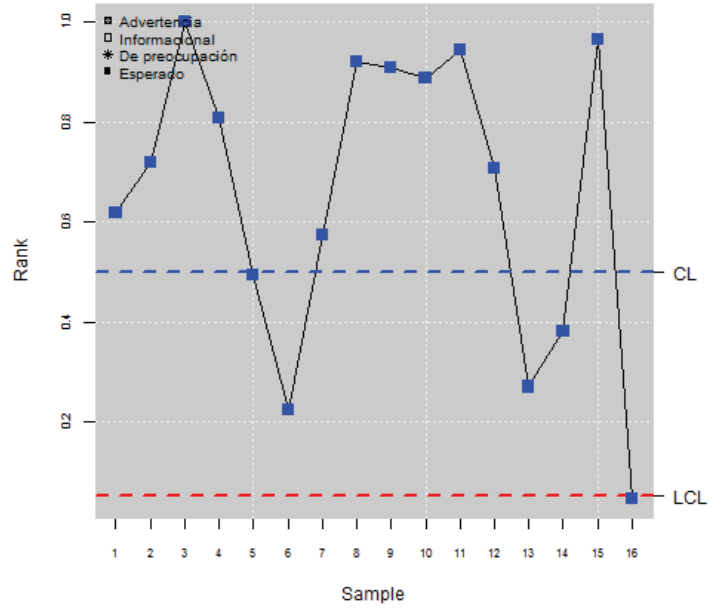
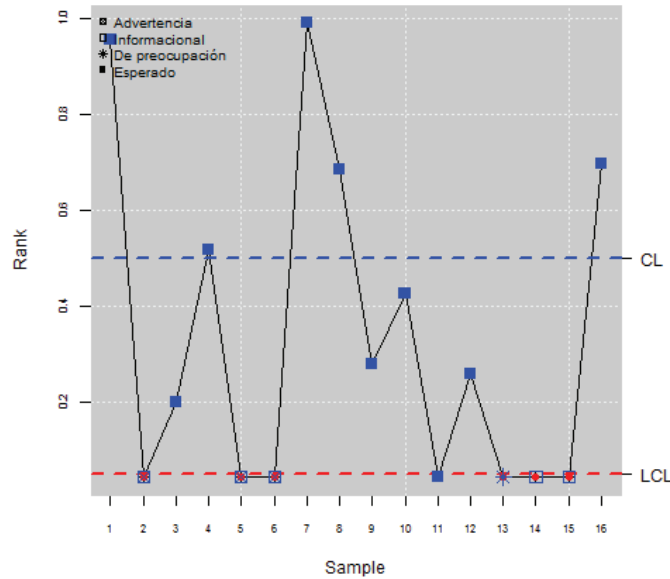


Figura A.15 7 r-AES7

Chart of control r



ANEXO 16. ESTADISTICA DESCRIPTIVA DE NOTAS PROMEDIO DE DIVERSOS ESTUDIANTES

Figura A.16. 1 MEDA-6054

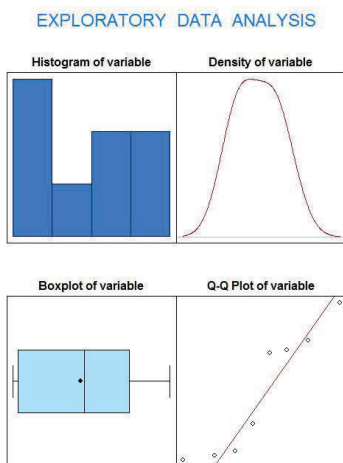


Figura A.16. 2 MEDA-6066

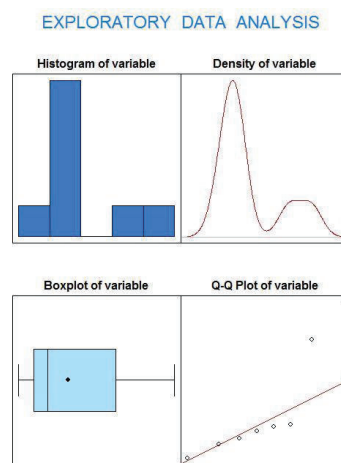


Figura A.16. 3 MEDA-6166

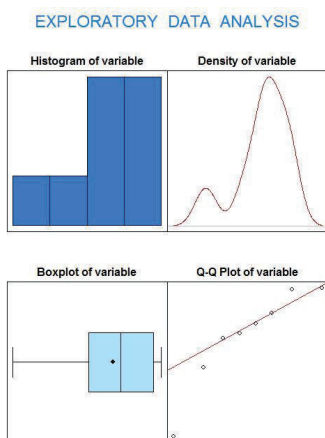


Figura A.16. 4 MEDA-6202

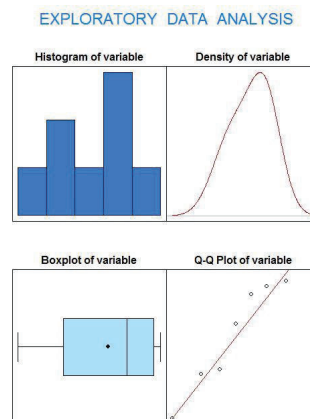


Figura A.16. 5 MEDA-6450

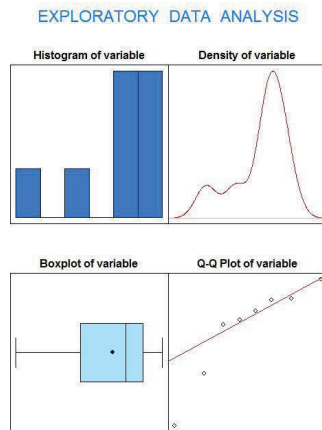


Figura A.16. 6 MEDA-6515

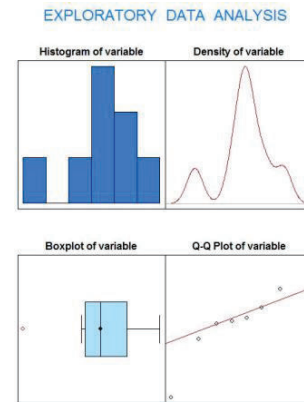


Figura A.16. 7 MEDA-6481

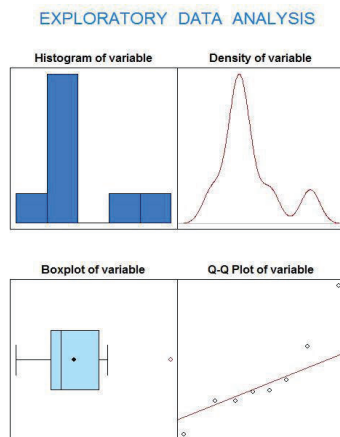


Figura A.16. 8 MEDA-7059

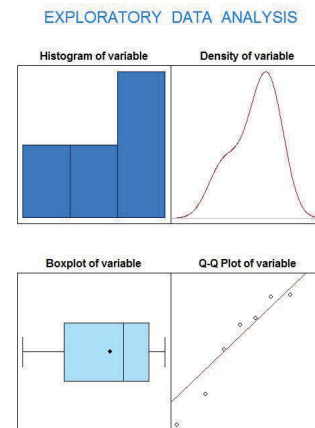


Figura A.16. 9 MEDA-7060

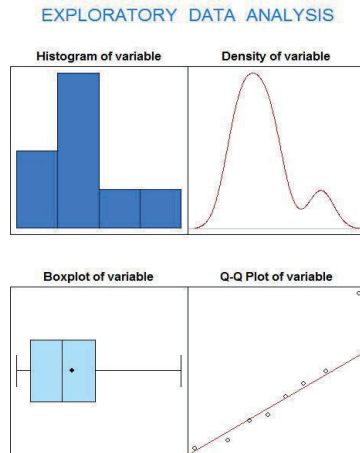


Figura A.16. 10 MEDA-7077

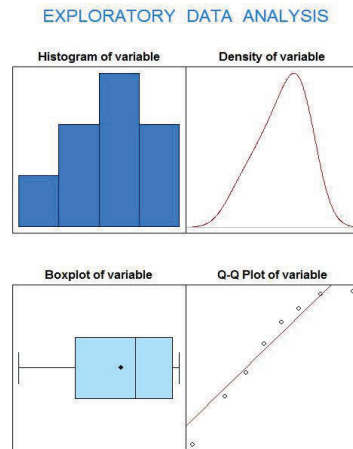


Figura A.16. 11 MEDA-7080

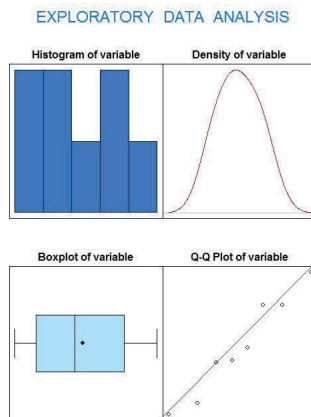


Figura A.16. 12 MEDA-7091

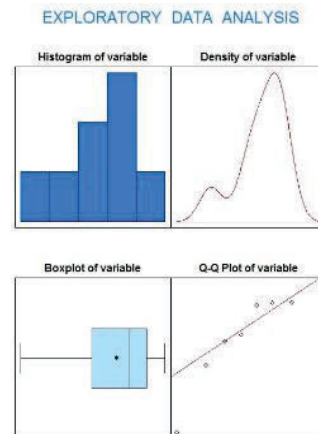


Figura A.16. 13 MEDA-8426

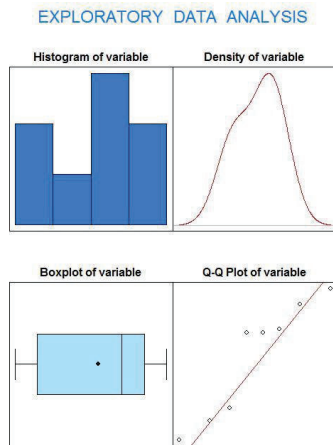


Figura A.16. 14 MEDA-8427

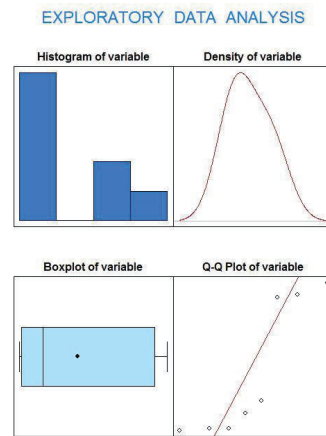
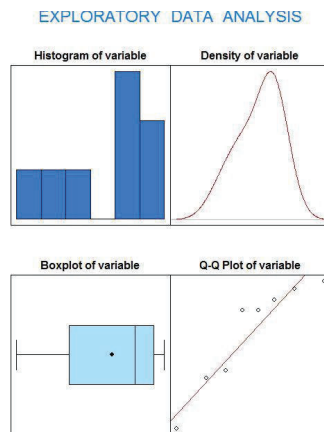


Figura A.16. 15 MEDA-8509



ANEXO 17. ESTADÍSTICOS DESCRIPTIVOS DE NOTAS PROMEDIO DE VARIOS ESTUDIANTES

ESTUDIANTES-PROMEDIO DE NOTAS

[1] "variable"-ESTUDIANTE 6054

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	7.300	7.401	7.675	7.783	7.675	7.886	7.986	0.259
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.067	0.092	0.485	0.686	-1.820	-0.289	0.199		

[1] "variable"-ESTUDIANTE 6066

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	8.429	8.438	8.669	8.525	8.669	8.984	9.037	0.277
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.077	0.098	0.546	0.608	-1.996	0.383	0.010		

[1] "variable"-ESTUDIANTE 6166

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	7.714	7.907	8.065	8.150	8.065	8.218	8.257	0.193
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.037	0.068	0.311	0.543	-1.344	-0.609	0.155		

[1] "variable"-ESTUDIANTE 6202

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.167	6.230	6.949	6.990	6.949	7.513	7.980	0.717
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.513	0.253	1.283	1.813	-1.938	0.095	0.142		

[1] "variable"-ESTUDIANTE 6450

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.912	7.075	7.445	7.225	7.445	7.948	8.575	0.595
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.354	0.210	0.873	1.663	-0.969	0.928	0.016		

[1] "variable"-ESTUDIANTE 6515

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	5.300	6.992	7.525	7.693	7.525	8.430	8.600	1.070
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	1.144	0.378	1.438	3.300	-0.394	-0.903	0.186		

[1] "variable"-ESTUDIANTE 6841

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	5.583	6.217	6.846	7.106	6.846	7.481	7.580	0.743
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.553	0.263	1.264	1.997	-1.584	-0.458	0.189		

[1] "variable"-ESTUDIANTE 7059

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	5.320	6.467	7.040	7.280	7.040	7.579	7.929	0.851
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.724	0.301	1.112	2.609	-0.634	-0.920	0.117		

[1] "variable"-ESTUDIANTE 7060

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	4.425	5.571	5.857	5.854	5.857	6.333	6.943	0.730
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.532	0.258	0.762	2.518	-0.482	-0.488	0.580		

[1] "variable"-ESTUDIANTE 7077

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.117	6.652	7.006	6.810	7.006	7.386	8.500	0.718
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.516	0.254	0.734	2.383	-0.321	0.897	0.161		

[1] "variable"-ESTUDIANTE 7080

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	7.020	7.428	7.886	8.017	7.886	8.271	8.433	0.495
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.245	0.175	0.843	1.413	-1.336	-0.572	0.364		

[1] "variable"-ESTUDIANTE 7091

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.414	6.712	7.620	7.397	7.620	8.125	10.000	1.140
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	1.299	0.403	1.413	3.586	-0.366	0.897	0.248		

[1] "variable"-ESTUDIANTE 8426

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.567	7.129	7.581	7.723	7.581	8.088	8.157	0.562
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.316	0.199	0.959	1.590	-1.297	-0.543	0.415		

[1] "variable"-ESTUDIANTE 8427

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	6.833	7.089	7.653	7.560	7.653	8.150	8.550	0.599
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.358	0.212	1.061	1.717	-1.593	0.045	0.700		

[1] "variable"-ESTUDIANTE 8509

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	4.100	6.170	6.904	7.283	6.904	7.780	8.325	1.359
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	1.846	0.480	1.610	4.225	-0.499	-0.901	0.163		

ANEXO 18. ESTUDIANTES. ESTADÍSTICA DESCRIPTIVA DE PORCENTAJE DE MATERIAS APROBADAS.

Figura A.18. 1 PEDA-6166

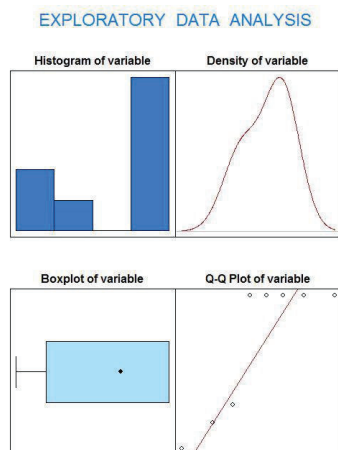


Figura A.18. 2 PEDA-6202

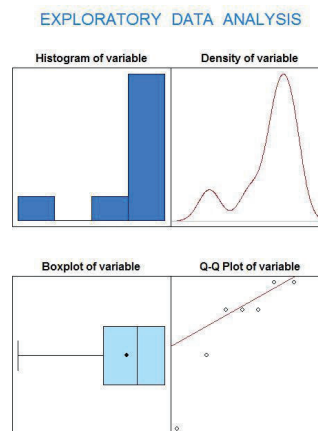


Figura A.18. 3 PEDA-6450

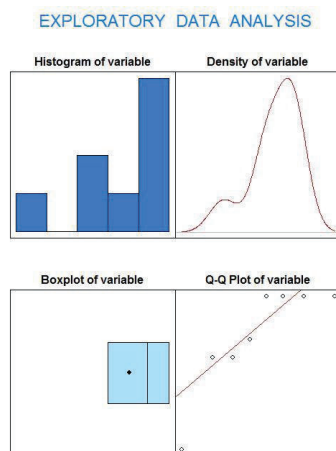


Figura A.18. 4 PEDA-6515

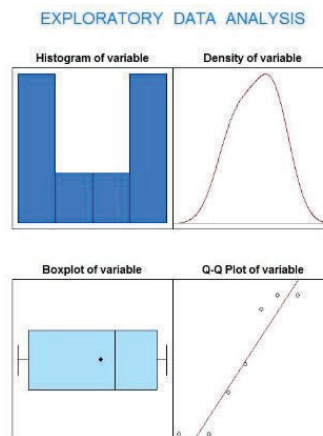


Figura A.18. 5 PEDA-6841

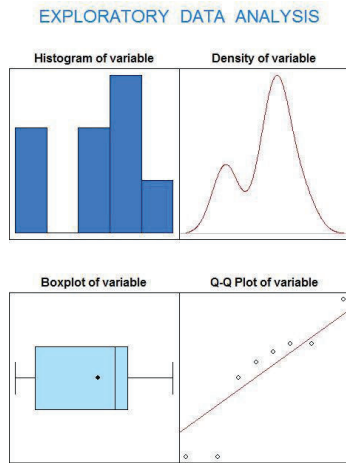


Figura A.18. 6 PEDA-7059

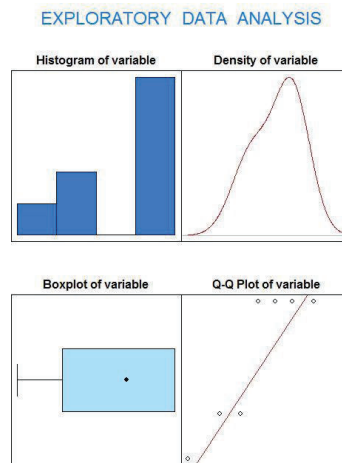


Figura A.18. 7 PEDA-7060

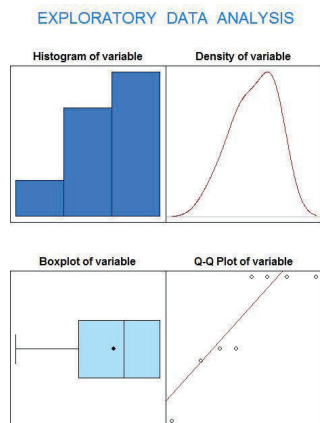


Figura A.18. 8 PEDA-7077

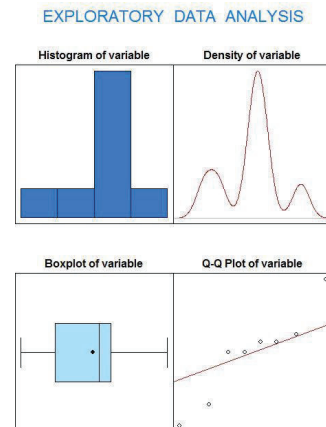


Figura A.18. 9 PEDA-7080

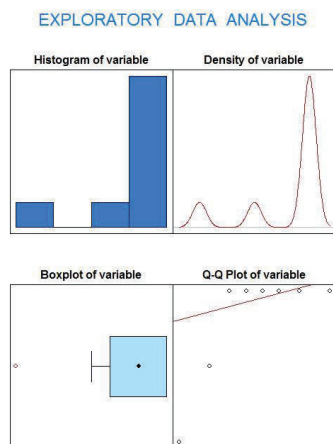


Figura A.18. 10 PEDA-7091

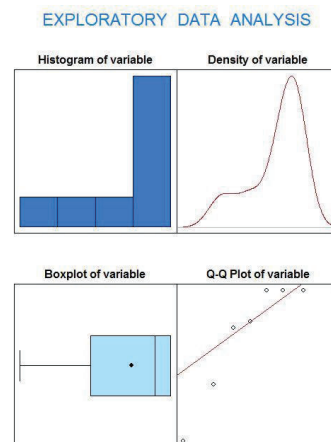


Figura A.18. 11 PEDA-8426

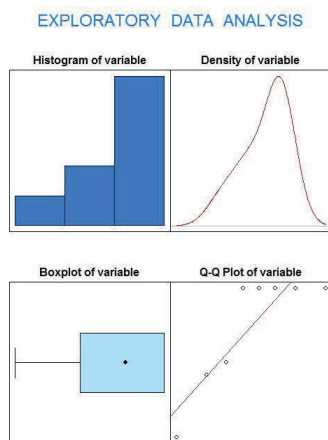
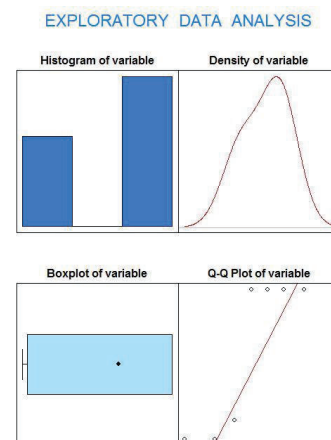


Figura A.18. 12 PEDA-8509



ANEXO 19. ESTUDIANTES. ESTADÍSTICOS DESCRIPTIVOS DE PORCENTAJE DE MATERIAS APROBADAS.

ESTUDIANTES - PORCENTAJE DE MATERIAS APROBADAS

[1] "x" - 8426

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.714	0.839	0.926	1.000	0.926	1.000	1.000	0.111
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.012	0.039	0.161	0.286	-1.123	-0.785	0.005		

[1] "x" - 8509

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.857	0.862	0.949	1.000	0.949	1.000	1.000	0.071
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.005	0.025	0.138	0.143	-1.998	-0.437	0.001		

[1] "x" - 6054

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.571	0.617	0.767	0.733	0.767	0.958	1.000	0.169
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.029	0.060	0.341	0.429	-1.726	0.292	0.240		

[1] "x" - 6066

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.667	0.762	0.840	0.833	0.840	0.958	1.000	0.114
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.013	0.040	0.196	0.333	-1.313	0.194	0.300		

[1] "x" - 6166

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.600	0.679	0.873	1.000	0.873	1.000	1.000	0.178
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.032	0.063	0.321	0.400	-1.858	-0.498	0.003		

[1] "x" - 6202

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.333	0.719	0.828	0.875	0.828	1.000	1.000	0.229
	Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val	
	0.052	0.081	0.281	0.667	-0.110	-1.148	0.013		

[1] "x" - 6450

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.500	0.800	0.870	0.929	0.870	1.000	1.000	0.175
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.031	0.062	0.200	0.500	-0.354	-0.987	0.017			

[1] "x" - 6515

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.500	0.525	0.699	0.733	0.699	0.833	0.857	0.152
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.023	0.054	0.308	0.357	-1.916	-0.244	0.091			

[1] "x" - 6841

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.500	0.562	0.762	0.817	0.762	0.857	1.000	0.177
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.031	0.063	0.295	0.500	-1.356	-0.464	0.150			

[1] "x" - 7059

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.800	0.857	0.939	1.000	0.939	1.000	1.000	0.086
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.007	0.030	0.143	0.200	-1.755	-0.536	0.003			

[1] "x" - 7060

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.714	0.839	0.908	0.929	0.908	1.000	1.000	0.108
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.012	0.038	0.161	0.286	-1.425	-0.469	0.041			

[1] "x" - 7080

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.667	0.875	0.938	1.000	0.938	1.000	1.000	0.124
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.015	0.044	0.125	0.333	-0.045	-1.280	0.000			

[1] "x" - 7091

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.200	0.575	0.792	0.917	0.792	1.000	1.000	0.295
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.087	0.104	0.425	0.800	-0.744	-0.950	0.014			

[1] "x" - 7077

Size (n)	Missing	Minimum	1st Qu	Mean	Median	TrMean	3rd Qu	Max.	Stdev.
8.000	0.000	0.667	0.744	0.830	0.845	0.830	0.871	1.000	0.102
Var.	SE Mean	I.Q.R.	Range	Kurtosis	Skewness	SW	p-val		
0.010	0.036	0.127	0.333	-0.983	-0.092	0.418			

ANEXO 20. COMPORTAMIENTO EN EL TIEMPO DE LAS NOTAS PROMEDIO DE LOS ESTUDIANTES

Figura A.20. 1 MSERIE-6166

Comportamiento del porcentaje de aprobación de materias por semestre

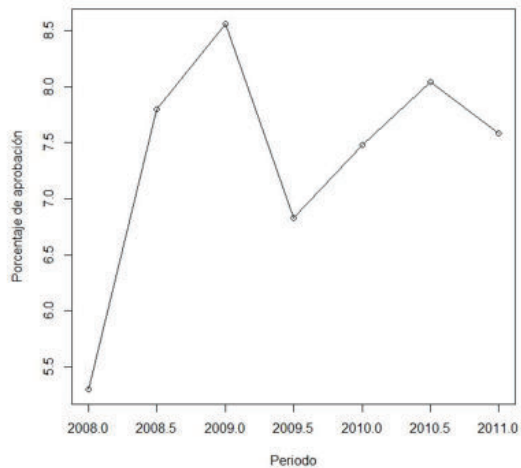


Figura A.20. 2 MSERIE-6054

Comportamiento del porcentaje de aprobación de materias por semestre

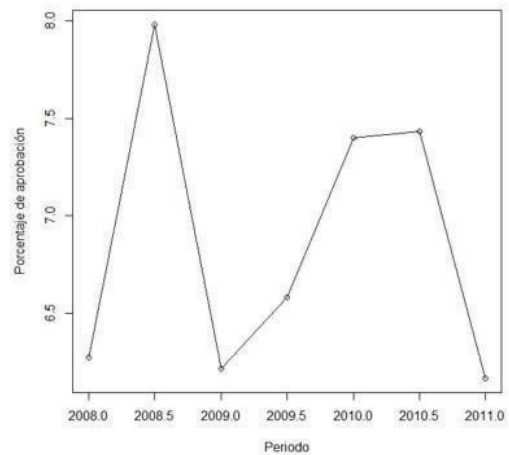


Figura A.20. 3 MSERIE-6202

Comportamiento del porcentaje de aprobación de materias por semestre

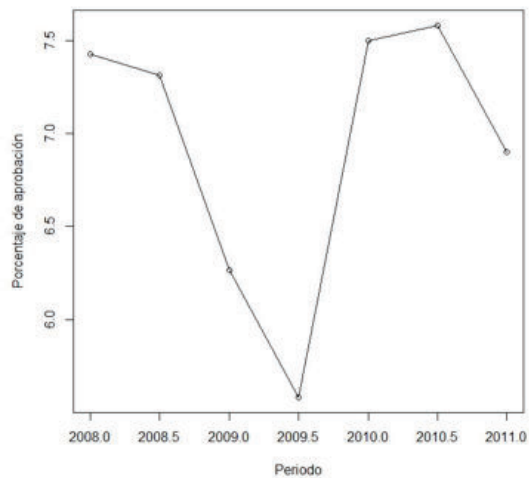


Figura A.20. 4 MSERIE-6066

Comportamiento del porcentaje de aprobación de materias por semestre

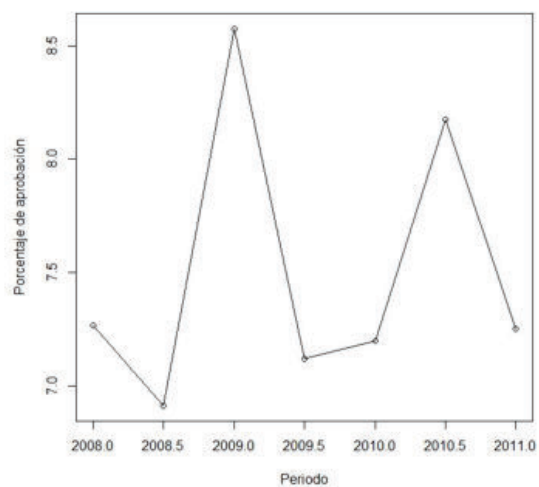


Figura A.20. 5 MSERIE-6450

Importamiento del porcentaje de aprobación de materias por semestre

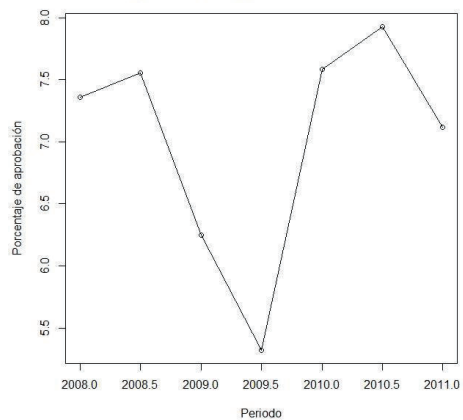


Figura A.20. 6 MSERIE-6515

Importamiento del porcentaje de aprobación de materias por semestre

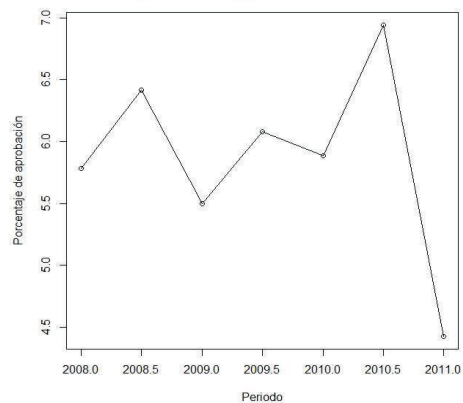


Figura A.20. 7 MSERIE 6841

Importamiento del porcentaje de aprobación de materias por semestre

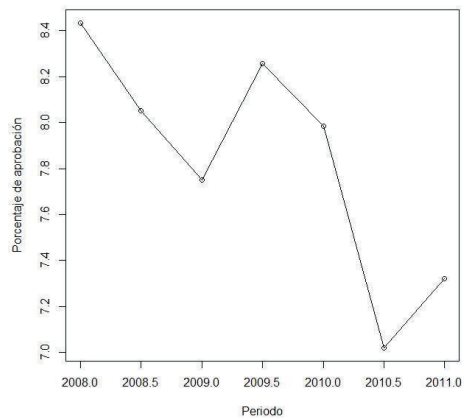


Figura A.20. 8 MSERIE 7059

Importamiento del porcentaje de aprobación de materias por semestre

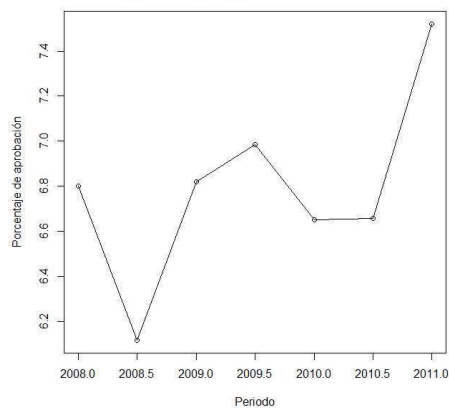


Figura A.20. 9 MSERIE 7060

Importamiento del porcentaje de aprobación de materias por semestre

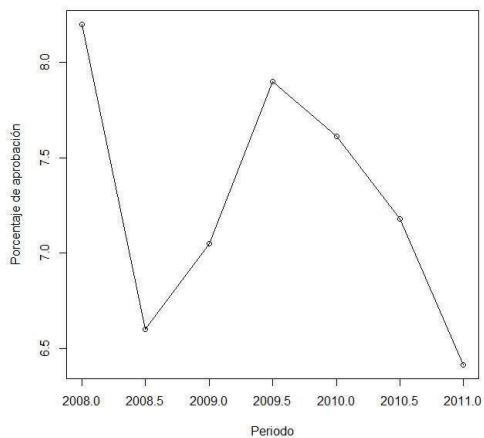


Figura A.20. 10 MSERIE 7077

Importamiento del porcentaje de aprobación de materias por semestre

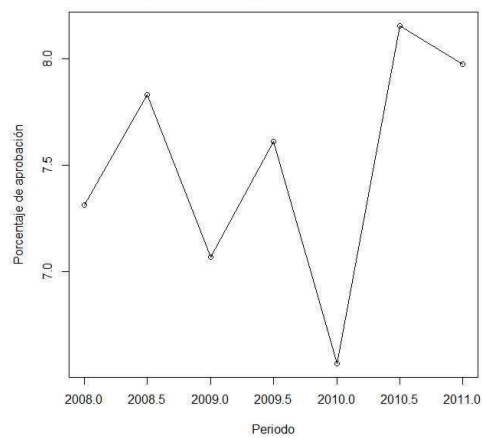


Figura A.20. 11 MSERIE-7080

Importamiento del porcentaje de aprobación de materias por semestre

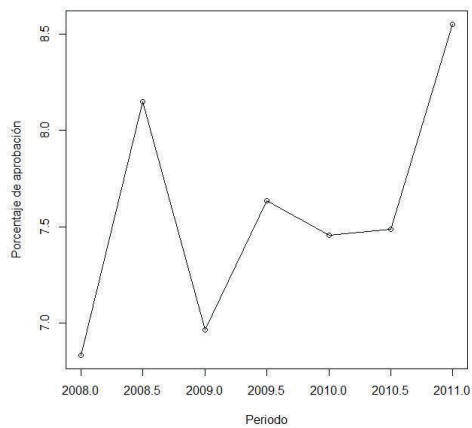


Figura A.20. 12 MSERIE-7091

Importamiento del porcentaje de aprobación de materias por semestre

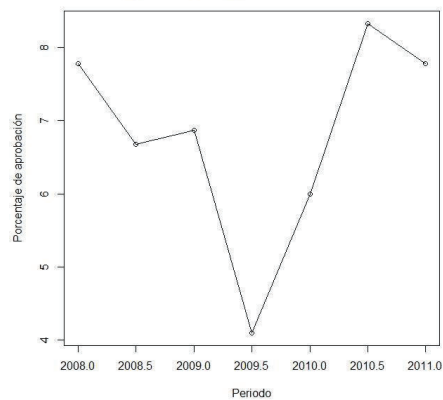


Figura A.20. 13 MSERIE-8426

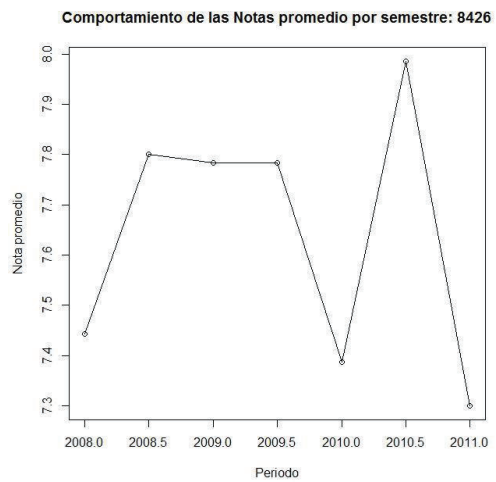


Figura A.20. 14 MSERIE-8427

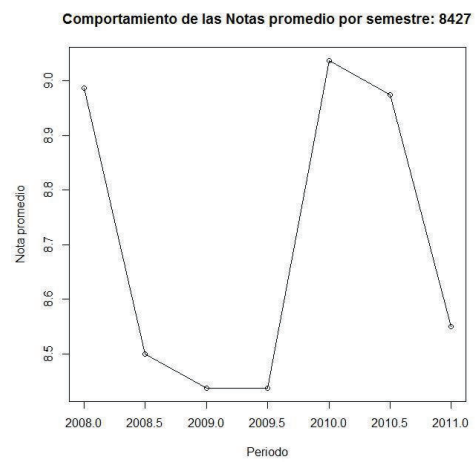
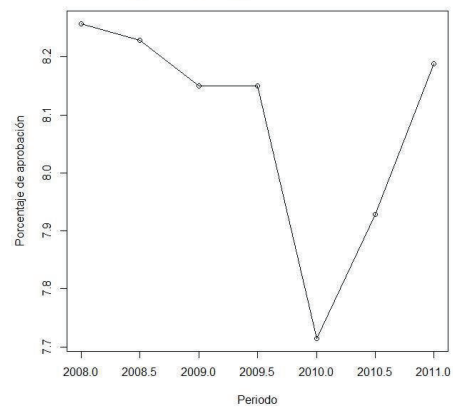


Figura A.20. 15 MSERIE-8509

Comportamiento del porcentaje de aprobación de materias por semestre



ANEXO 21. COMPORTAMIENTO EN EL TIEMPO DEL PORCENTAJE DE MATERIAS APROBADAS

Figura A.21. 1 PSERIE-6054

Comportamiento del porcentaje de aprobación de materias por semestre

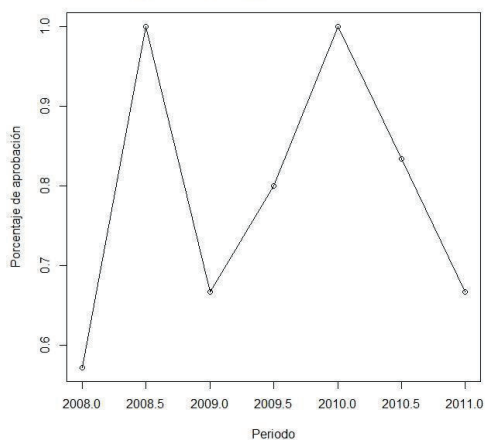


Figura A.21. 2 PSERIE-6066

Comportamiento del porcentaje de aprobación de materias por semestre

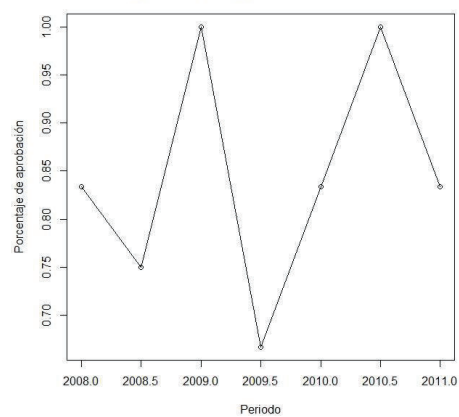


Figura A.21. 3 PSERIE-6166

Comportamiento del porcentaje de aprobación de materias por semestre

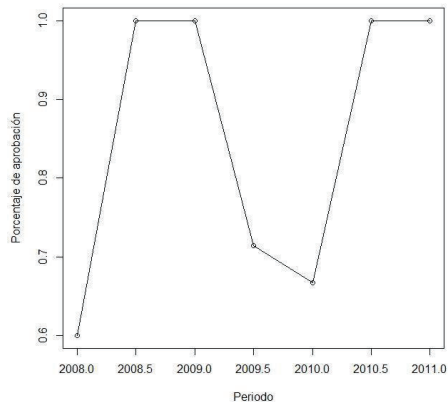


Figura A.21. 4 PSERIE-6202

Comportamiento del porcentaje de aprobación de materias por semestre

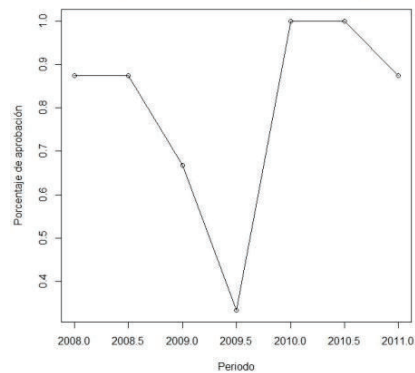


Figura A.21. 5 PSERIE-6450

Importamiento del porcentaje de aprobación de materias por semestre

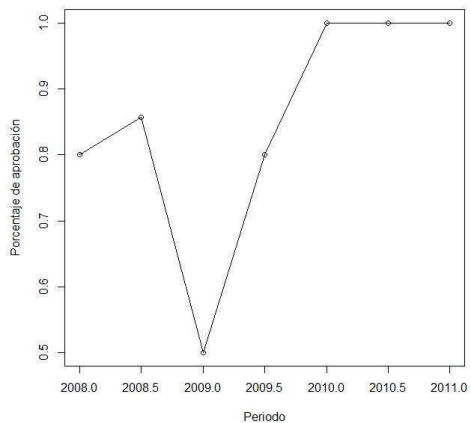


Figura A.21. 6 PSERIE-6515

Importamiento del porcentaje de aprobación de materias por semestre

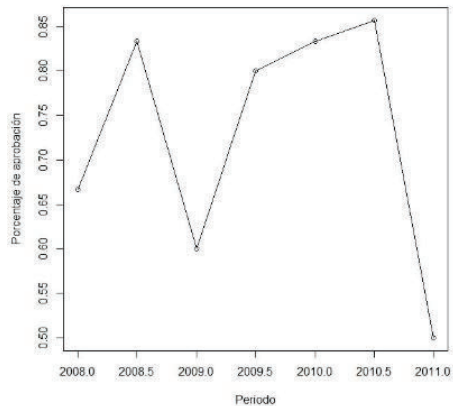


Figura A.21. 7 PSERIE-6841

Importamiento del porcentaje de aprobación de materias por semestre

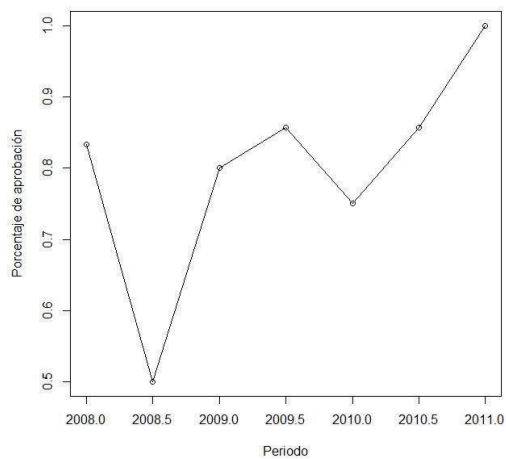


Figura A.21. 8 PSERIE-7059

Importamiento del porcentaje de aprobación de materias por semestre

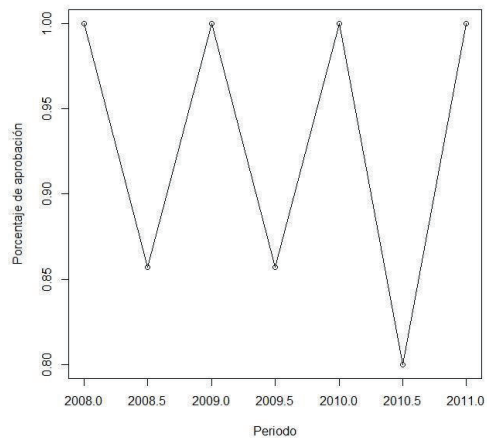


Figura A.21. 9 PSERIE-7060

Importamiento del porcentaje de aprobación de materias por semestre

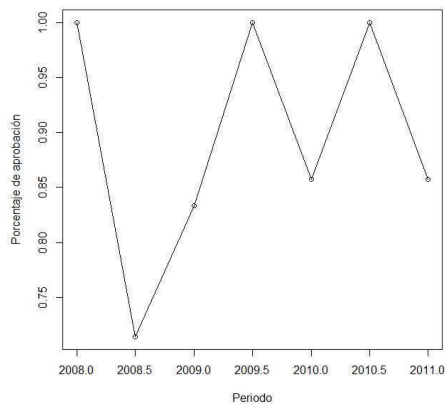


Figura A.21. 10 PSERIE-7077

Importamiento del porcentaje de aprobación de materias por semestre

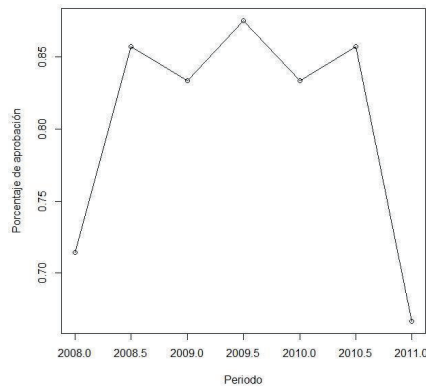


Figura A.21. 11 PSERIE-7080

Importamiento del porcentaje de aprobación de materias por semestre

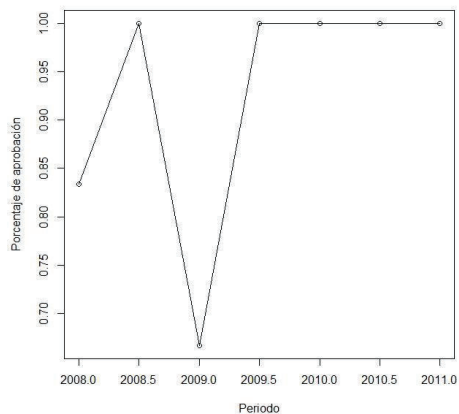


Figura A.21. 12 PSERIE-7091

Importamiento del porcentaje de aprobación de materias por semestre

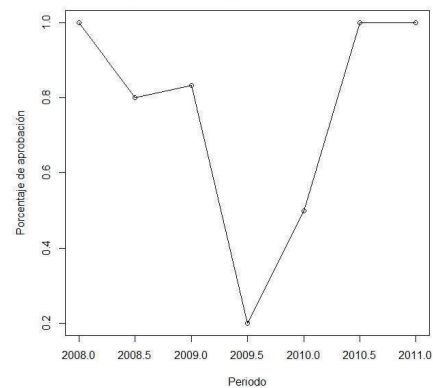
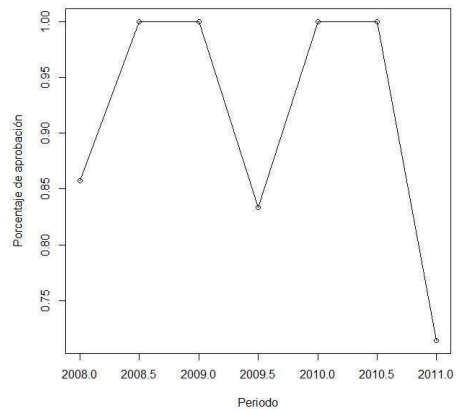
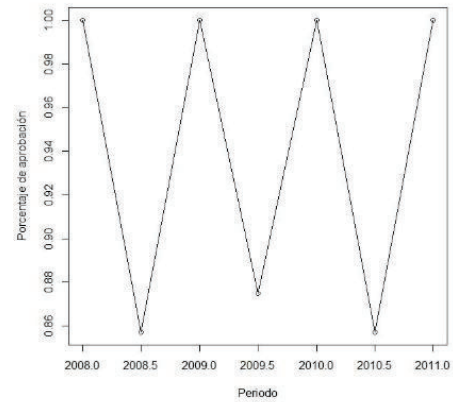


Figura A.21. 13 PSERIE-8426

Comportamiento del porcentaje de aprobación de materias por semestre

**Figura A.21. 14 PSERIE-8509**

Comportamiento del porcentaje de aprobación de materias por semestre



ANEXO 22. ESTUDIANTES. GRÁFICOS BIVARIADOS DE NOTAS – PORCENTAJE DE APROBACIÓN DE MATERIAS.

Figura A.22. 1 B6054

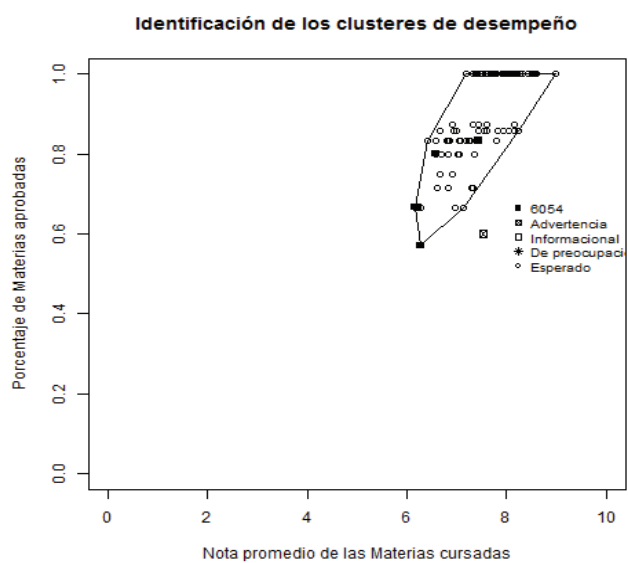


Figura A.22. 2 B6066

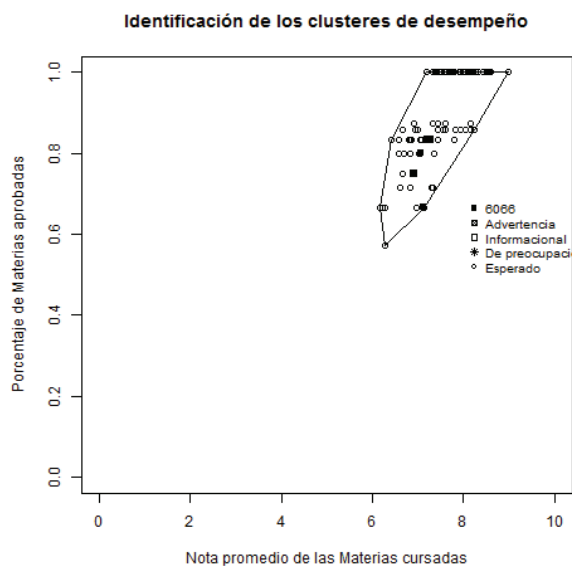


Figura A.22. 3 B6166

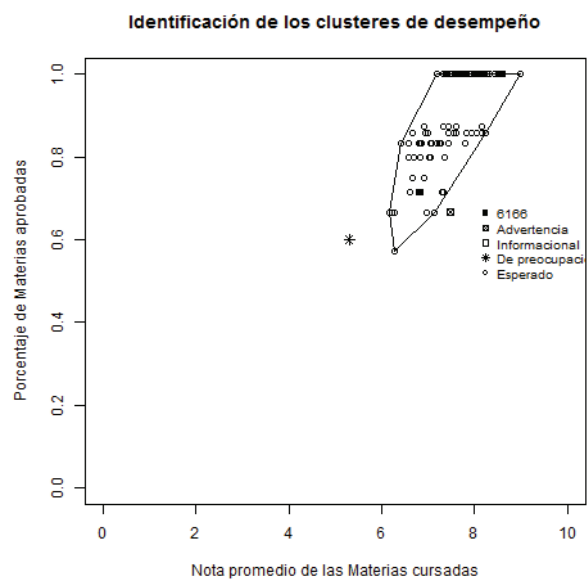


Figura A.22. 4 B6202

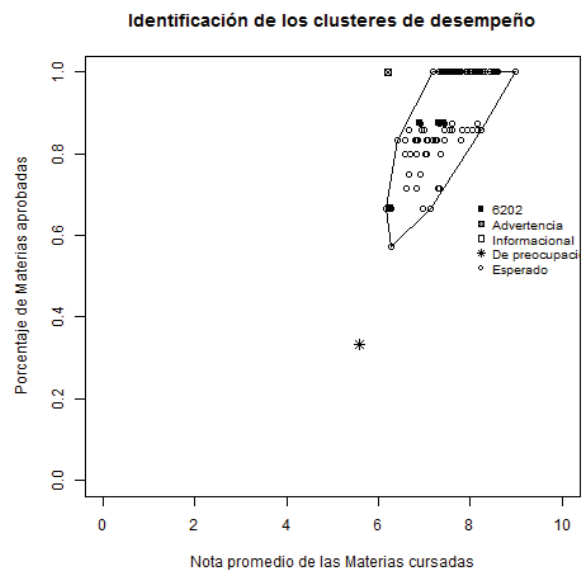


Figura A.22. 5 B6450

Identificación de los clusters de desempeño

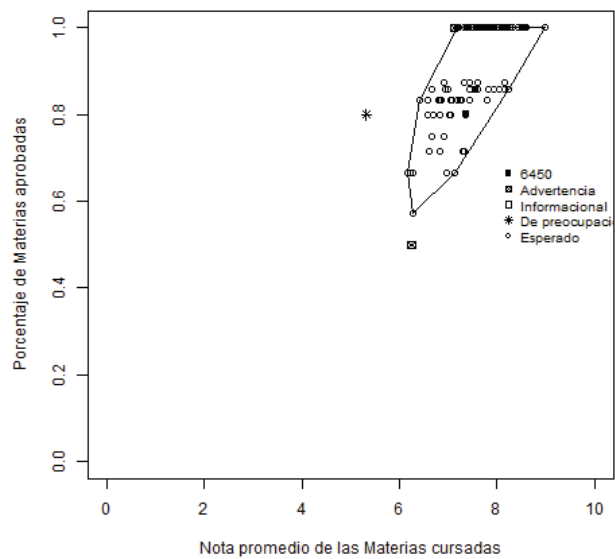


Figura A.22. 6 B6515

Identificación de los clusters de desempeño

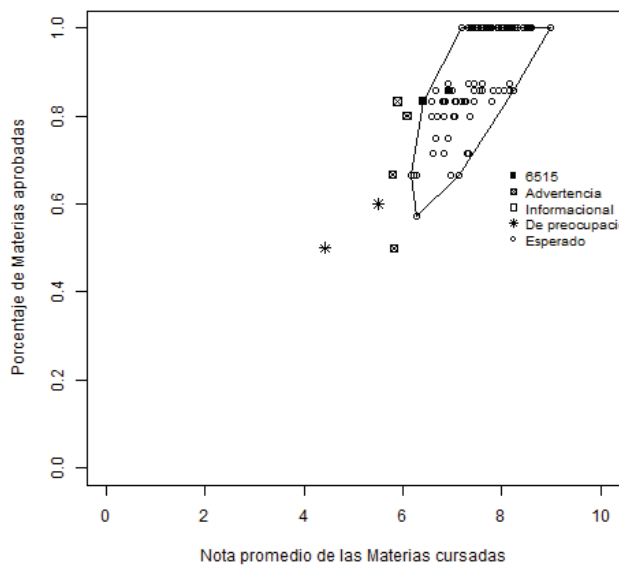


Figura A.22. 7 B6841

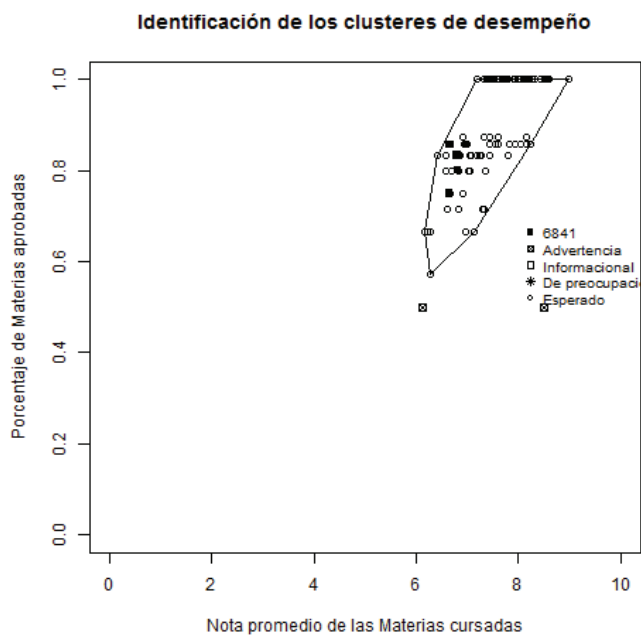


Figura A.22. 8 B7059

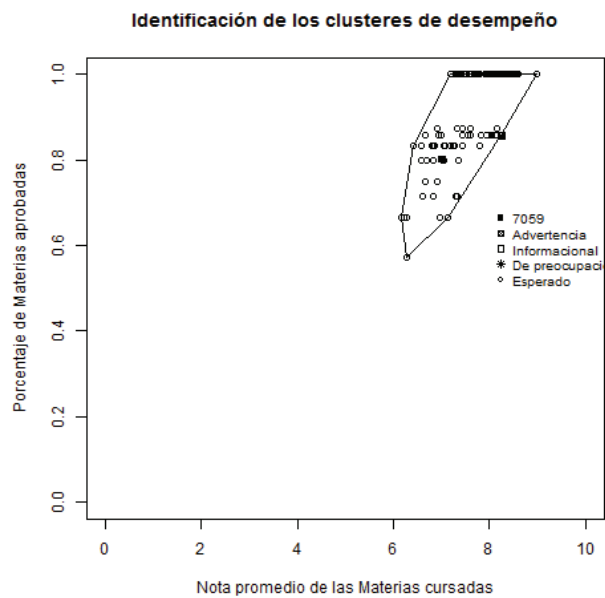


Figura A.22. 9 B7060

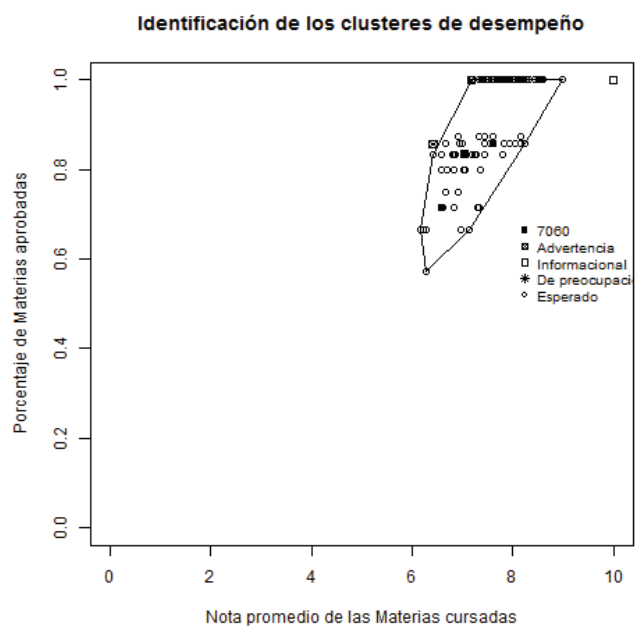


Figura A.22. 10 B7077

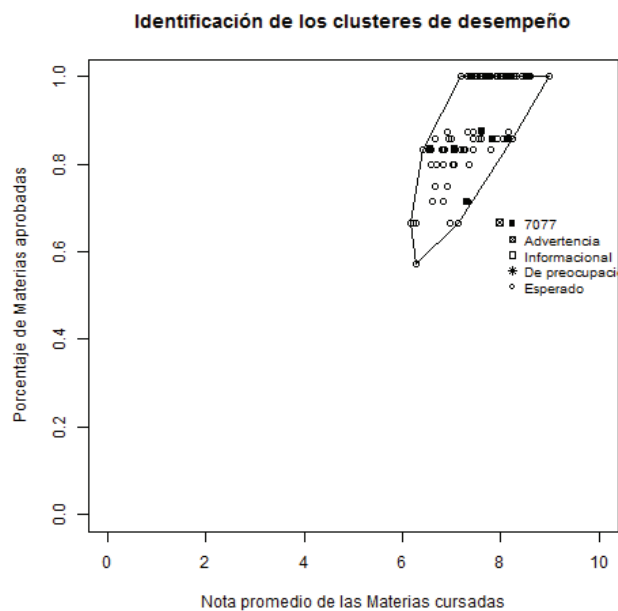


Figura A.22. 11 B7080

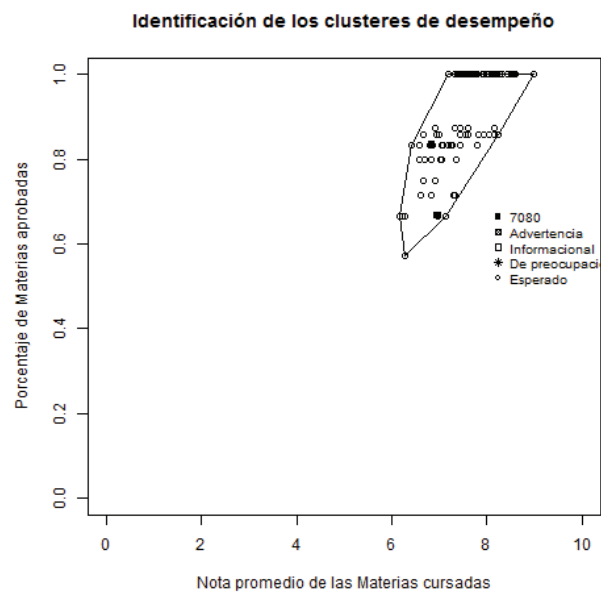


Figura A.22. 12 B7091

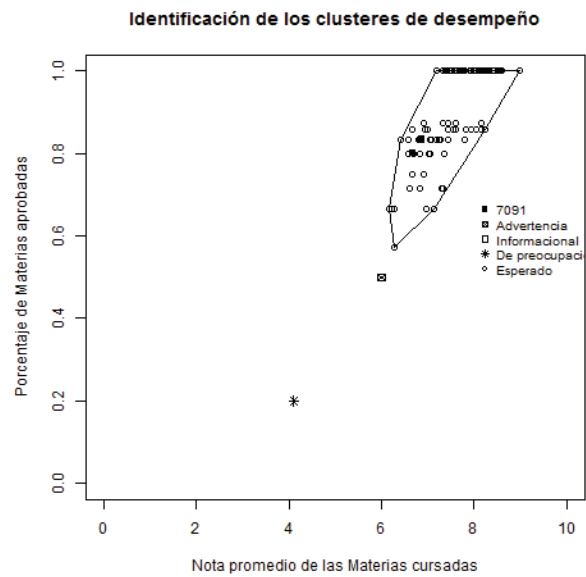


Figura A.22. 13 B8426

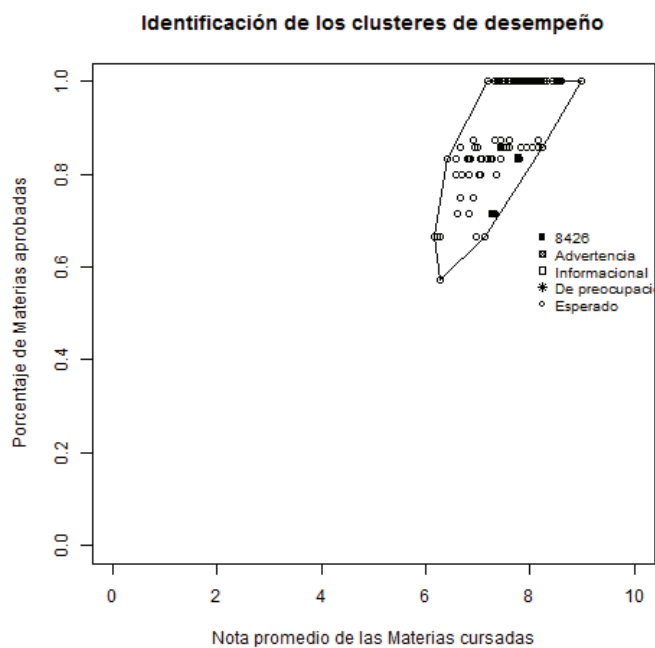


Figura A.22. 14 B8427

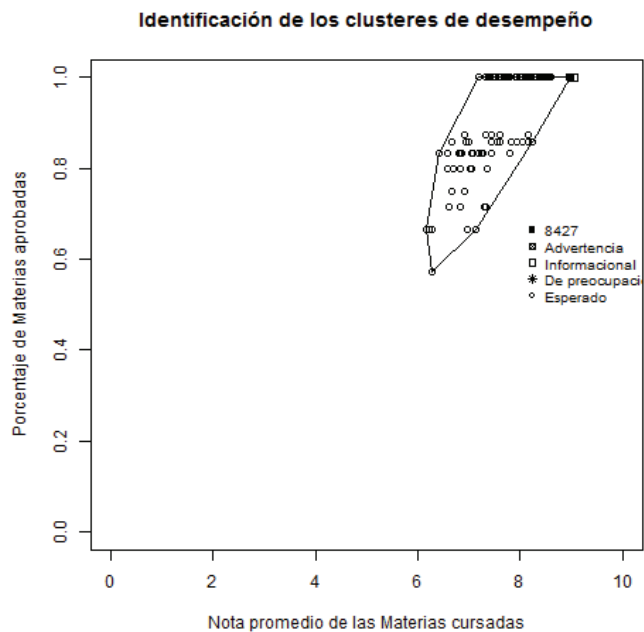
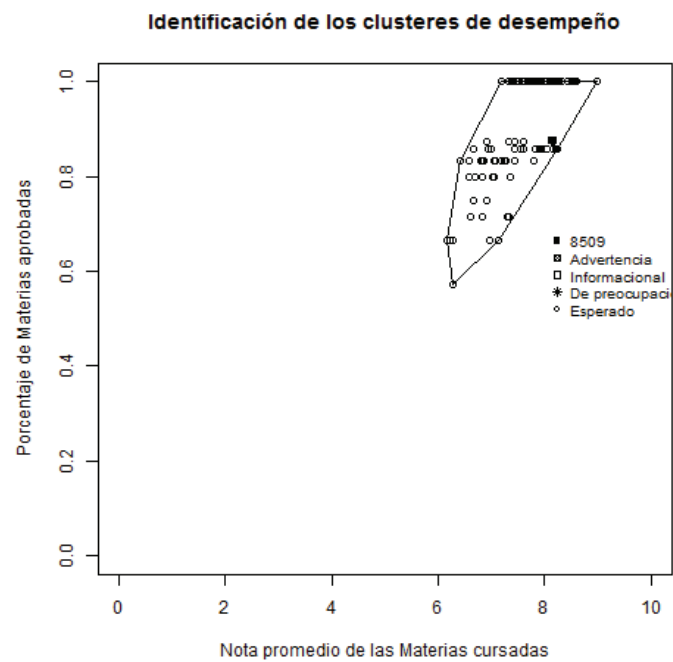


Figura A.22. 15 B8509



ANEXO 23. ESTUDIANTES. CARTAS DE CONTROL r

Figura A.23. 1 C6054

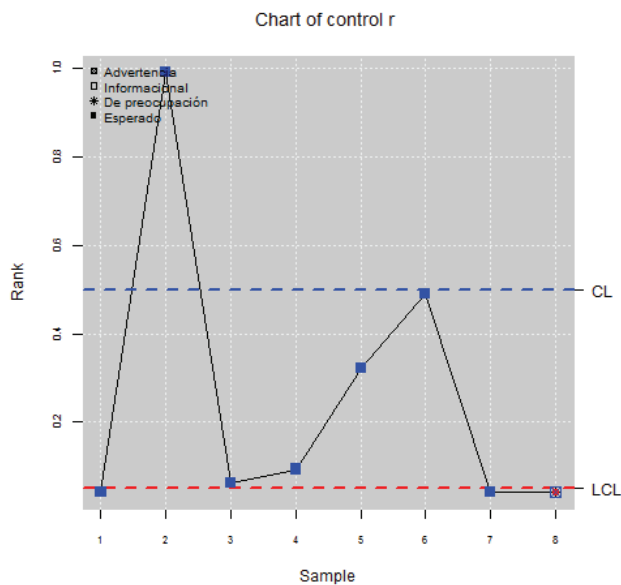


Figura A.23. 2 C6066

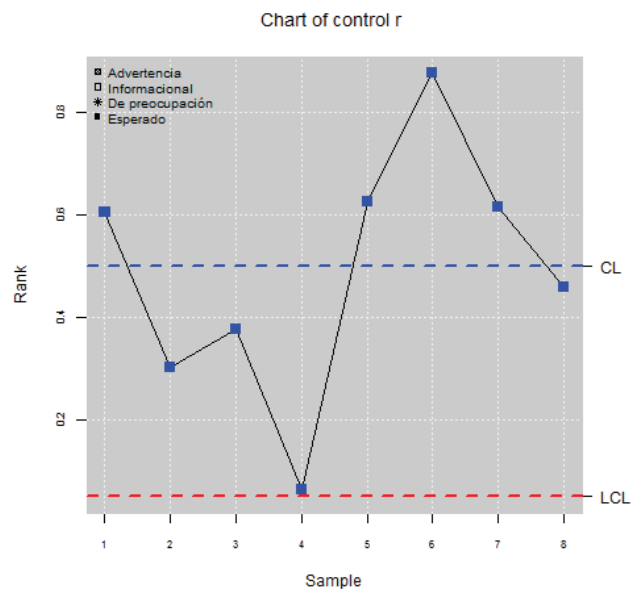


Figura A.23. 3 C6202

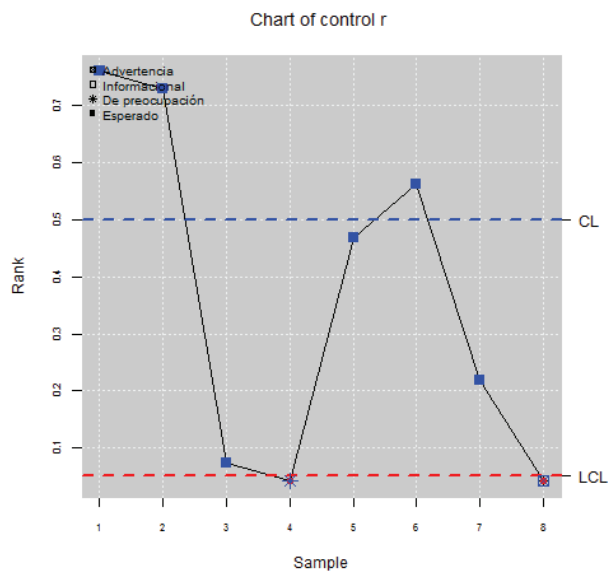


Figura A.23. 4 C6166

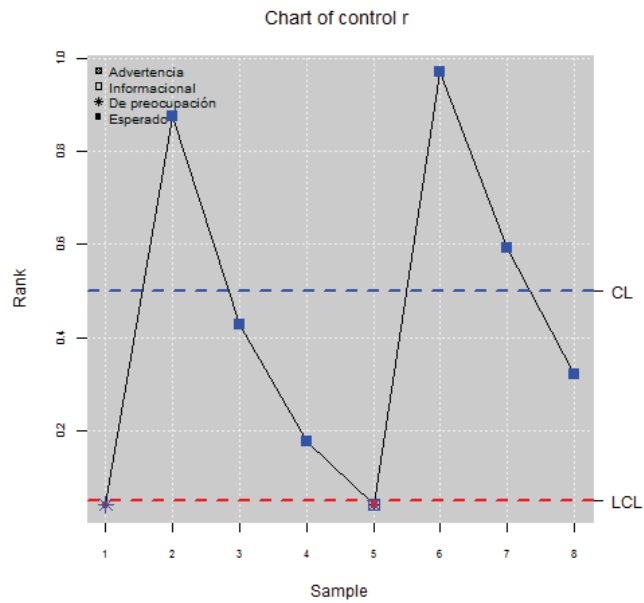


Figura A.23. 5 C6450

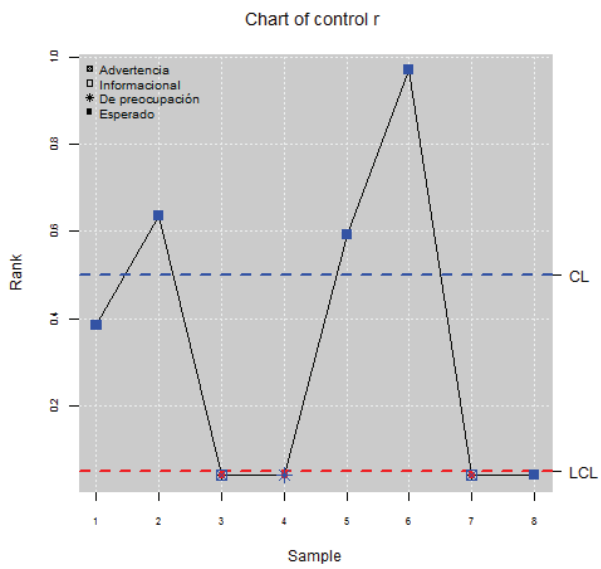


Figura A.23. 6 C6515

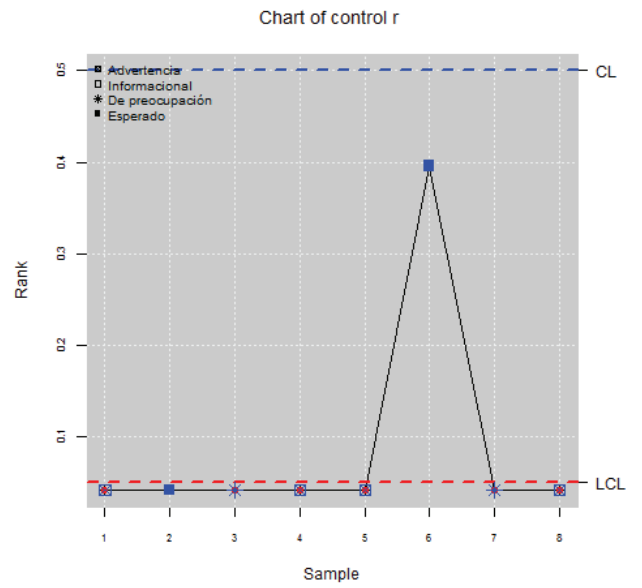


Figura A.23. 7 C6841

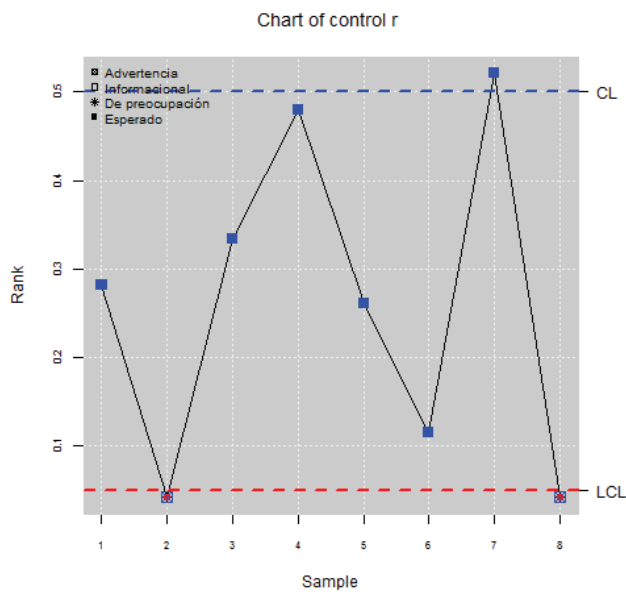


Figura A.23. 8 C7059

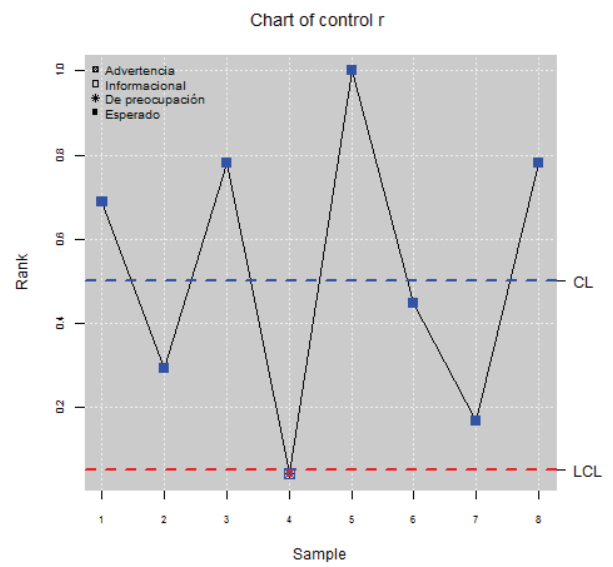


Figura A.23. 9 C7060

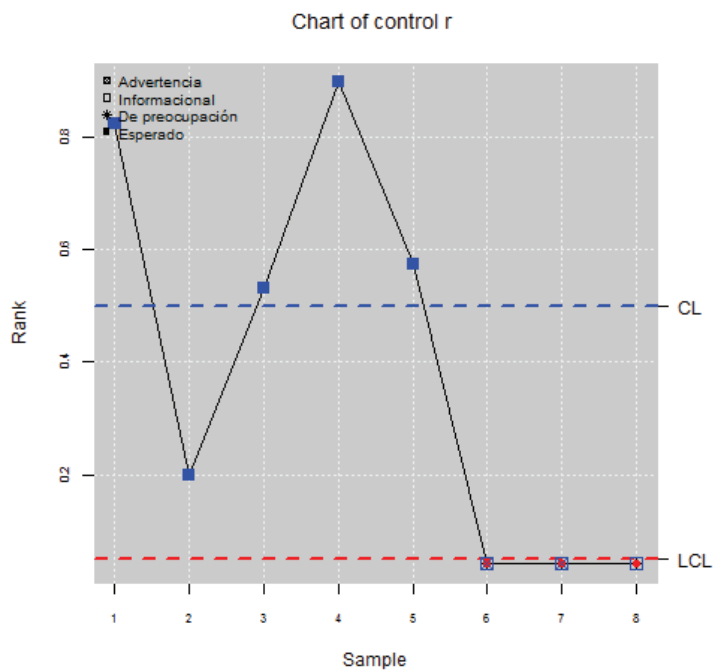


Figura A.23. 10 C7077

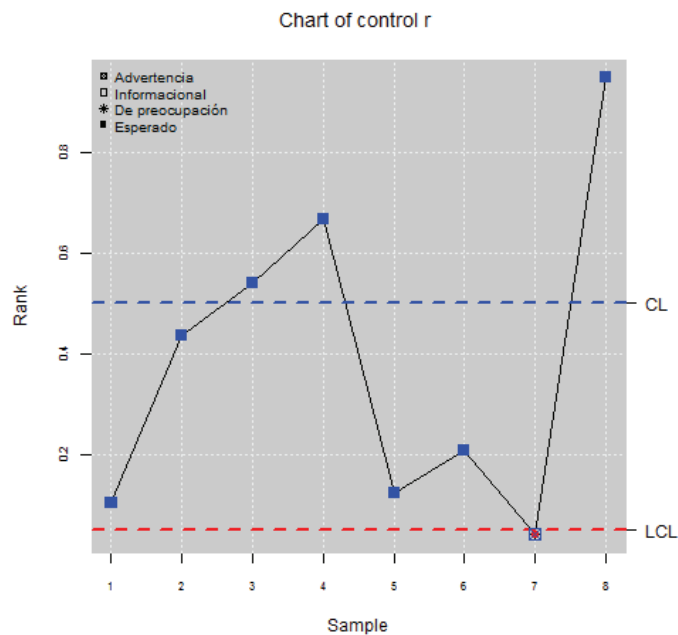


Figura A.23. 11 C7080

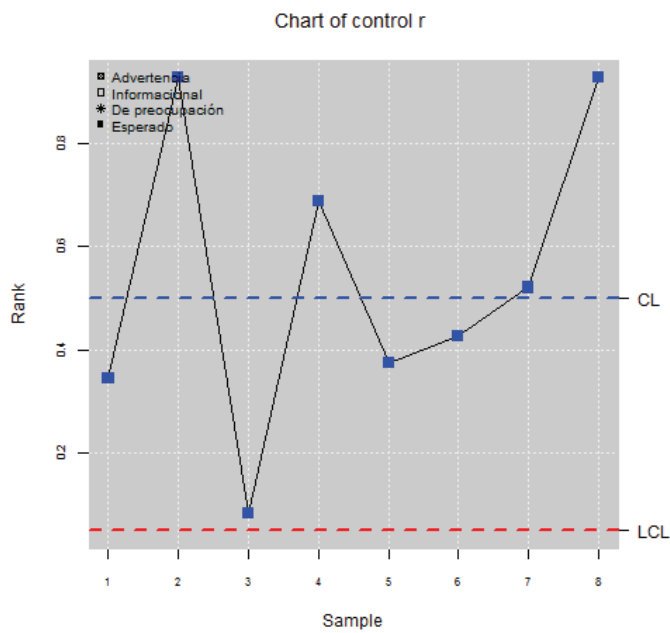


Figura A.23. 12 C7091

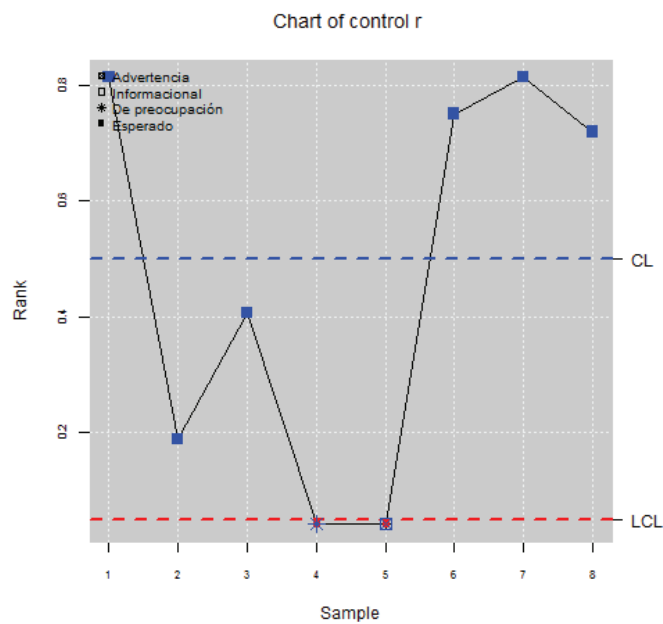


Figura A.23. 13 C8426

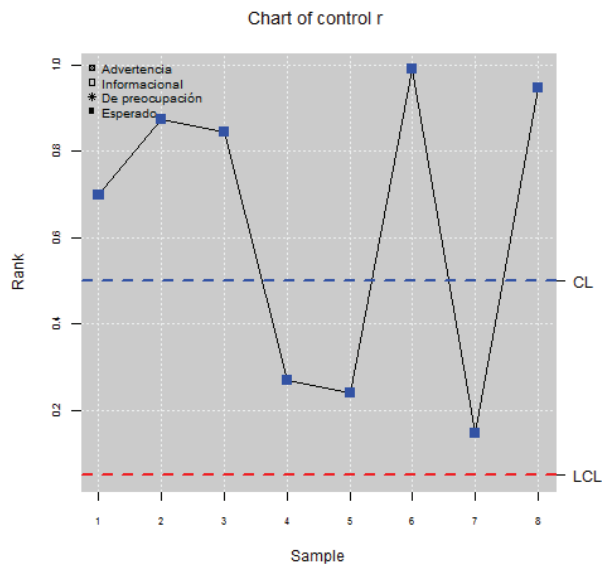


Figura A.23. 14 C8427

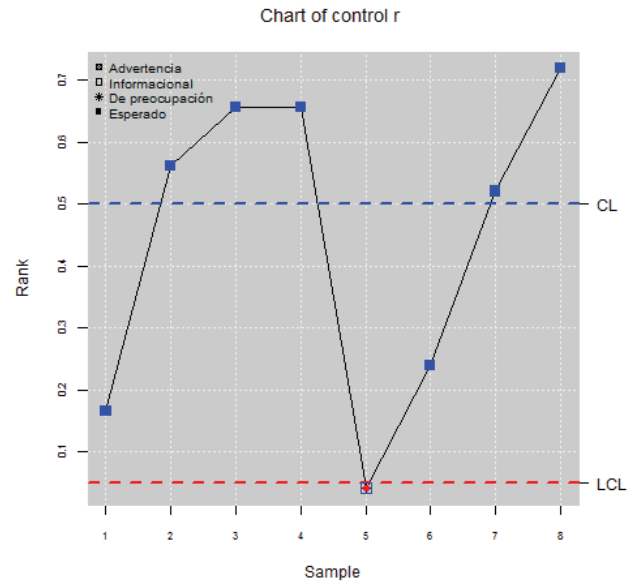
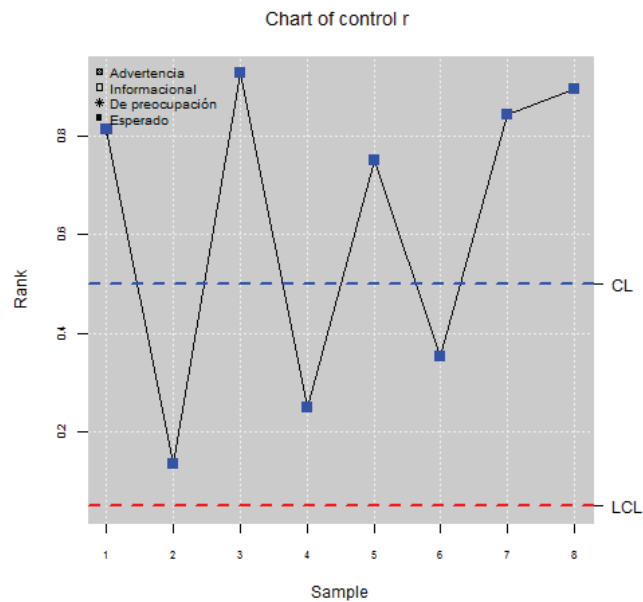


Figura A.23. 15 C8509



ANEXO 24. DATOS ORIGINALES.

Tabla 24. 1 MATERIAS

MATERIA	i	Año	Periodo	Nota	%_AR
AES1	1	2008	1	4,95041667	0,5125
AES1	2	2008	2	5,36149425	0,51724138
AES1	3	2009	1	5,35874126	0,58041958
AES1	4	2009	2	5,50425532	0,64539007
AES1	5	2010	1	5,92030075	0,70676692
AES1	6	2010	2	6,68653846	0,75
AES1	7	2011	1	6,15142857	0,62857143
AES1	8	2011	2	6,09152542	0,66101695
AES1	9	2012	1	5,72559524	0,66071429
AES1	10	2012	2	6,83484849	0,78787879
AES1	11	2013	1	6,16818182	0,70454546
AES1	12	2013	2	7,275	0,92391304
AES1	13	2014	1	6,82195122	0,82317073
AES1	14	2014	2	6,23285714	0,64285714
AES1	15	2015	1	5,97612903	0,63870968
AES1	16	2015	2	6,35787037	0,80092593
AES2	1	2008	1	6,11428571	0,66666667
AES2	2	2008	2	7,63043478	0,97826087
AES2	3	2009	1	6,99333333	1
AES2	4	2009	2	8,0125	1
AES2	5	2010	1	6,24166667	0,875
AES2	6	2010	2	6,276	0,84
AES2	7	2011	1	5,98888889	0,63888889
AES2	8	2011	2	5,96296296	0,61728395
AES2	9	2012	1	7,58	0,93333333
AES2	10	2012	2	6,97962963	0,7962963
AES2	11	2013	1	7,79230769	0,94871795
AES2	12	2013	2	7,35492958	0,90140845

AES2	13	2014	1	7,35245902	0,81967213
AES2	14	2014	2	6,97961165	0,86407767
AES2	15	2015	1	5,9765625	0,65625
AES2	16	2015	2	6,06470588	0,69607843
AES3	1	2008	1	5,64642857	0,57142857
AES3	2	2008	2	5,8826087	0,54347826
AES3	3	2009	1	3,57931035	0,27586207
AES3	4	2009	2	4,75116279	0,48837209
AES3	5	2010	1	5,80967742	0,61290323
AES3	6	2010	2	5,97758621	0,62068966
AES3	7	2011	1	5,55625	0,5
AES3	8	2011	2	6,27118644	0,76271186
AES3	9	2012	1	5,84871795	0,61538462
AES3	10	2012	2	7,22098765	0,90123457
AES3	11	2013	1	7,276	0,8
AES3	12	2013	2	7,51666667	0,93137255
AES3	13	2014	1	6,70886076	0,75949367
AES3	14	2014	2	6,94385965	0,83333333
AES3	15	2015	1	7,25915493	0,91549296
AES3	16	2015	2	5,68150685	0,68493151
AES4	1	2008	1	5,38674699	0,62650602
AES4	2	2008	2	5,43858268	0,53543307
AES4	3	2009	1	5,12941177	0,54901961
AES4	4	2009	2	5,86885246	0,6147541
AES4	5	2010	1	5,74032258	0,66129032
AES4	6	2010	2	5,856	0,66
AES4	7	2011	1	6,06557377	0,68032787
AES4	8	2011	2	6,1785124	0,75206612
AES4	9	2012	1	7,07318841	0,82608696
AES4	10	2012	2	7,27835821	0,91044776
AES4	11	2013	1	6,89732143	0,83928571
AES4	12	2013	2	6,7859375	0,875

AES4	13	2014	1	6,83652695	0,82035928
AES4	14	2014	2	6,93057325	0,81528662
AES4	15	2015	1	7,19208633	0,89928058
AES4	16	2015	2	6,22180451	0,69924812
AES5	1	2008	1	7,92941177	0,88235294
AES5	2	2008	2	5,32272727	0,54545455
AES5	3	2009	1	5,86333333	0,63333333
AES5	4	2009	2	8,12857143	1
AES5	5	2010	1	6,22692308	0,76923077
AES5	6	2010	2	6,83	0,85
AES5	7	2011	1	7,39354839	0,83870968
AES5	8	2011	2	7,93333333	0,91666667
AES5	9	2012	1	6,65294118	0,82352941
AES5	10	2012	2	7,80476191	0,95238095
AES5	11	2013	1	7,32592593	0,87037037
AES5	12	2013	2	6,81818182	0,84848485
AES5	13	2014	1	6,63473684	0,83157895
AES5	14	2014	2	7,63768116	0,98550725
AES5	15	2015	1	7,50361446	0,92771084
AES5	16	2015	2	6,74109589	0,82191781
AES6	1	2008	1	6,51443299	0,71134021
AES6	2	2008	2	6,73295455	0,82954546
AES6	3	2009	1	6,49701493	0,74626866
AES6	4	2009	2	6,28271605	0,74074074
AES6	5	2010	1	5,77126437	0,62068966
AES6	6	2010	2	6,29891304	0,65217391
AES6	7	2011	1	6,84631579	0,76842105
AES6	8	2011	2	6,18348624	0,68807339
AES6	9	2012	1	6,82970297	0,82178218
AES6	10	2012	2	6,5311828	0,7311828
AES6	11	2013	1	6,79871795	0,7948718
AES6	12	2013	2	6,642	0,74

AES6	13	2014	1	6,13555556	0,62962963
AES6	14	2014	2	7,08592593	0,88888889
AES6	15	2015	1	6,32666667	0,73333333
AES6	16	2015	2	6,516	0,85
AES7	1	2008	1	6,2972973	0,7027027
AES7	2	2008	2	6,225	0,83333333
AES7	3	2009	1	5,68275862	0,65517241
AES7	4	2009	2	6,425	0,6875
AES7	5	2010	1	5,30952381	0,52380952
AES7	6	2010	2	5,825	0,71875
AES7	7	2011	1	6,56296296	0,77777778
AES7	8	2011	2	6,52142857	0,71428571
AES7	9	2012	1	6,52068966	0,82758621
AES7	10	2012	2	6,46190476	0,80952381
AES7	11	2013	1	6,98333333	0,91666667
AES7	12	2013	2	7,44444444	0,81481482
AES7	13	2014	1	5,21666667	0,44444444
AES7	14	2014	2	7,41612903	0,96774194
AES7	15	2015	1	7,67894737	1
AES7	16	2015	2	7,16333333	0,86666667

Tabla 24. 2 ESTUDIANTES

Estudiante	MATRICULA	i	Año	Periodo	Nota	%_AR	N
1	8426	1	2008	1	7,44286	0,85714	7
1	8426	2	2008	2	7,8	1	7
1	8426	3	2009	1	7,78333	1	6
1	8426	4	2009	2	7,78333	0,83333	6
1	8426	5	2010	1	7,3875	1	8
1	8426	6	2010	2	7,98571	1	7
1	8426	7	2011	1	7,3	0,71429	7
1	8426	8	2011	2	7,91429	1	7
2	8427	1	2008	1	8,9875	1	8
2	8427	2	2008	2	8,5	1	7
2	8427	3	2009	1	8,4375	1	8
2	8427	4	2009	2	8,4375	1	8
2	8427	5	2010	1	9,0375	1	8
2	8427	6	2010	2	8,975	1	8
2	8427	7	2011	1	8,55	1	8
2	8427	8	2011	2	8,42857	1	7
3	8509	1	2008	1	8,25714	1	7
3	8509	2	2008	2	8,22857	0,85714	7
3	8509	3	2009	1	8,15	1	8
3	8509	4	2009	2	8,15	0,875	8
3	8509	5	2010	1	7,71429	1	7
3	8509	6	2010	2	7,92857	0,85714	7
3	8509	7	2011	1	8,1875	1	8
3	8509	8	2011	2	7,9	1	8
4	6054	1	2008	1	6,27143	0,57143	7
4	6054	2	2008	2	7,98	1	3
4	6054	3	2009	1	6,21667	0,66667	6
4	6054	4	2009	2	6,58	0,8	2
4	6054	5	2010	1	7,4	1	6
4	6054	6	2010	2	7,43333	0,83333	6

4	6054	7	2011	1	6,16667	0,66667	6
4	6054	8	2011	2	7,54	0,6	5
5	6066	1	2008	1	7,26667	0,83333	6
5	6066	2	2008	2	6,9125	0,75	8
5	6066	3	2009	1	8,575	1	4
5	6066	4	2009	2	7,12	0,66667	6
5	6066	5	2010	1	7,2	0,83333	6
5	6066	6	2010	2	8,175	1	4
5	6066	7	2011	1	7,25	0,83333	6
5	6066	8	2011	2	7,06	0,8	5
6	6166	1	2008	1	5,3	0,6	5
6	6166	2	2008	2	7,8	1	6
6	6166	3	2009	1	8,56	1	5
6	6166	4	2009	2	6,82857	0,71429	7
6	6166	5	2010	1	7,48333	0,66667	6
6	6166	6	2010	2	8,04	1	5
6	6166	7	2011	1	7,58571	1	7
6	6166	8	2011	2	8,6	1	2
7	6202	1	2008	1	7,425	0,875	8
7	6202	2	2008	2	7,3125	0,875	8
7	6202	3	2009	1	6,26667	0,66667	6
7	6202	4	2009	2	5,58333	0,33333	6
7	6202	5	2010	1	7,5	1	5
7	6202	6	2010	2	7,58	1	5
7	6202	7	2011	1	6,9	0,875	8
7	6202	8	2011	2	6,2	1	1
8	6450	1	2008	1	7,36	0,8	5
8	6450	2	2008	2	7,55714	0,85714	7
8	6450	3	2009	1	6,25	0,5	4
8	6450	4	2009	2	5,32	0,8	5
8	6450	5	2010	1	7,58571	1	7
8	6450	6	2010	2	7,92857	1	7

8	6450	7	2011	1	7,11667	1	6
8	6450	8	2011	2	7,2	1	3
9	6515	1	2008	1	5,78333	0,66667	6
9	6515	2	2008	2	6,41667	0,83333	6
9	6515	3	2009	1	5,5	0,6	5
9	6515	4	2009	2	6,08	0,8	5
9	6515	5	2010	1	5,88333	0,83333	6
9	6515	6	2010	2	6,94286	0,85714	7
9	6515	7	2011	1	4,425	0,5	4
9	6515	8	2011	2	5,825	0,5	4
10	6841	1	2008	1	6,8	0,83333	6
10	6841	2	2008	2	6,11667	0,5	6
10	6841	3	2009	1	6,82	0,8	5
10	6841	4	2009	2	6,98571	0,85714	7
10	6841	5	2010	1	6,65	0,75	4
10	6841	6	2010	2	6,65714	0,85714	7
10	6841	7	2011	1	7,52	1	5
10	6841	8	2011	2	8,5	0,5	2
11	7059	1	2008	1	8,43333	1	6
11	7059	2	2008	2	8,05	0,85714	7
11	7059	3	2009	1	7,75	1	6
11	7059	4	2009	2	8,25714	0,85714	7
11	7059	5	2010	1	7,98333	1	6
11	7059	6	2010	2	7,02	0,8	5
11	7059	7	2011	1	7,32	1	5
11	7059	8	2011	2	8,275	1	4
12	7060	1	2008	1	8,2	1	7
12	7060	2	2008	2	6,6	0,71429	7
12	7060	3	2009	1	7,05	0,83333	6
12	7060	4	2009	2	7,9	1	8
12	7060	5	2010	1	7,61429	0,85714	7
12	7060	6	2010	2	7,18	1	5

12	7060	7	2011	1	6,41429	0,85714	7
12	7060	8	2011	2	10	1	1
13	7077	1	2008	1	7,31429	0,71429	7
13	7077	2	2008	2	7,83333	0,85714	9
13	7077	3	2009	1	7,06667	0,83333	6
13	7077	4	2009	2	7,6125	0,875	8
13	7077	5	2010	1	6,56667	0,83333	6
13	7077	6	2010	2	8,15714	0,85714	7
13	7077	7	2011	1	7,975	0,66667	6
13	7077	8	2011	2	8,125	1	4
14	7080	1	2008	1	6,83333	0,83333	6
14	7080	2	2008	2	8,15	1	7
14	7080	3	2009	1	6,96667	0,66667	6
14	7080	4	2009	2	7,63333	1	6
14	7080	5	2010	1	7,45714	1	7
14	7080	6	2010	2	7,4875	1	8
14	7080	7	2011	1	8,55	1	6
14	7080	8	2011	2	8,15	1	2
15	7091	1	2008	1	7,78	1	5
15	7091	2	2008	2	6,68	0,8	5
15	7091	3	2009	1	6,86667	0,83333	6
15	7091	4	2009	2	4,1	0,2	5
15	7091	5	2010	1	6	0,5	4
15	7091	6	2010	2	8,325	1	4
15	7091	7	2011	1	7,78	1	5
15	7091	8	2011	2	7,7	1	7