

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

PROPUESTA DE UNA METODOLOGÍA PARA LA FOCALIZACIÓN DEL SUBSIDIO ELÉCTRICO “TARIFA DE LA DIGNIDAD” A USUARIOS RESIDENCIALES DE LAS EMPRESAS ELÉCTRICAS DEL ECUADOR

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE
MAGISTER EN INGENIERÍA ELÉCTRICA EN DISTRIBUCIÓN**

ING. EDISON GABRIEL NOVOA GUAMÁN

DIRECTOR: DR. ING. GABRIEL BENJAMÍN SALAZAR YÉPEZ

Quito, diciembre 2020

AVAL

Certifico que el presente trabajo fue desarrollado por Edison Gabriel Novoa Guamán, bajo mi supervisión.

DR. ING. GABRIEL BENJAMÍN SALAZAR YÉPEZ
DIRECTOR DEL TRABAJO DE TITULACIÓN

DECLARACIÓN DE AUTORÍA

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

A través de la presente declaración dejo constancia de que la Escuela Politécnica Nacional podrá hacer uso del presente trabajo según los términos estipulados en la Ley, Reglamentos y Normas vigentes.

EDISON GABRIEL NOVOA GUAMÁN

DEDICATORIA

A mi amada esposa Eliana y a mi querido hijo Joaquín.

A todos mis familiares y amigos.

¡Los aprecio mucho!

AGRADECIMIENTO

A Dios por permitir que se haga en mí su voluntad todos los días de mi vida.

A mi mentor y director, Dr. Ing. Gabriel Salazar, a su esposa, Ing. Ivonne Pérez, y a toda su familia por su constante apoyo, guía y confianza.

A mi querida Escuela Politécnica Nacional por brindarme los conocimientos necesarios para poder ejercer, lo que más me gusta, la Ingeniería Eléctrica.

A mi esposa y a mi hijo por darme las fuerzas y motivación para vivir cada día.

A mis padres por siempre apoyarme en mis metas.

A mis familiares y amigos por tener siempre los mejores deseos y oraciones para mí.

¡Muchas gracias a todos!

ÍNDICE DE CONTENIDO

| | |
|---|------|
| AVAL | I |
| DECLARACIÓN DE AUTORÍA..... | II |
| DEDICATORIA..... | III |
| AGRADECIMIENTO..... | IV |
| ÍNDICE DE CONTENIDO..... | V |
| RESUMEN | VIII |
| ABSTRACT | IX |
| 1 INTRODUCCIÓN..... | 1 |
| 1.1 Objetivo General | 2 |
| 1.2 Objetivos Específicos | 2 |
| 1.3 Alcance | 3 |
| 2 MARCO TEÓRICO | 5 |
| 2.1 Fundamentos Sociales de los Subsidios..... | 5 |
| 2.1.1 La Sociedad | 5 |
| 2.1.2 La Pobreza | 5 |
| 2.1.3 El Subsidio | 6 |
| 2.1.4 Experiencias Internacionales de los Subsidios | 6 |
| 2.1.5 Focalización de Subsidios | 7 |
| 2.2 El Subsidio Tarifa de la Dignidad en el Ecuador | 8 |
| 2.2.1 Definición..... | 8 |
| 2.2.2 Procedimiento de Aplicación..... | 8 |
| 2.2.3 Resultados de la Aplicación del Subsidio..... | 10 |
| 2.3 Extracción del Conocimiento a partir de Datos | 12 |
| 2.3.1 Definición..... | 12 |
| 2.3.2 Selección, Limpieza y Emparejamiento de Datos | 13 |
| 2.4 Agrupamiento de Datos..... | 13 |

| | | |
|-------|---|----|
| 2.4.1 | Definición..... | 13 |
| 2.4.2 | Técnicas | 14 |
| 2.4.3 | Evaluación de Grupos | 15 |
| 2.4.4 | Ventajas y Desventajas | 15 |
| 2.4.5 | Aplicaciones | 16 |
| 2.5 | Herramientas Computacionales | 16 |
| 2.5.1 | Software ArcGIS..... | 16 |
| 2.5.2 | Software MATLAB | 17 |
| 3 | METODOLOGÍA | 18 |
| 3.1 | Recopilación y Procesamiento de Datos..... | 18 |
| 3.1.1 | Datos Técnicos..... | 18 |
| 3.1.2 | Datos Geográficos..... | 19 |
| 3.1.3 | Datos Socioeconómicos | 21 |
| 3.2 | Emparejamiento de Datos..... | 22 |
| 3.2.1 | Emparejamiento de Datos Técnicos y Geográficos..... | 22 |
| 3.2.2 | Emparejamiento de Datos Económicos y Geográficos | 23 |
| 3.2.3 | Emparejamiento de Datos Geográficos, Técnicos y Socioeconómicos | 25 |
| 3.3 | Agrupación de Datos..... | 28 |
| 3.3.1 | Definición de una Ecuación de Pobreza | 28 |
| 3.3.2 | Definición de la Técnica de Agrupación | 30 |
| 3.3.3 | Evaluación y Definición de Grupos | 32 |
| 3.3.4 | Criterio de asignación | 32 |
| 4 | APLICACIÓN Y RESULTADOS | 33 |
| 4.1 | Estructuración de la Base de Datos | 33 |
| 4.1.1 | Datos Técnicos..... | 33 |
| 4.1.2 | Datos Socioeconómicos | 35 |
| 4.1.3 | Consolidación de Datos Técnicos, Geográficos y Socioeconómicos | 41 |
| 4.2 | Agrupación de Datos..... | 43 |

| | | |
|-------|---|----|
| 4.2.1 | Agrupación de Usuarios según su Ubicación..... | 43 |
| 4.2.2 | Agrupación de Usuarios según su Consumo de Energía | 46 |
| 4.2.3 | Agrupación de Usuarios según la Valoración del Suelo | 47 |
| 4.2.4 | Agrupación de Usuarios según los Ingresos Económicos de la Zona | 48 |
| 4.2.5 | Agrupación de Usuarios según la Ecuación de Pobreza..... | 50 |
| 4.3 | Evaluación de Grupos | 51 |
| 4.4 | Asignación del Subsidio | 52 |
| 4.5 | Valoración Económica..... | 57 |
| 4.6 | Propuesta Metodológica..... | 60 |
| 5 | CONCLUSIONES | 62 |
| 6 | RECOMENDACIONES..... | 63 |
| 7 | REFERENCIAS BIBLIOGRÁFICAS | 64 |
| 8 | ANEXOS..... | 67 |

RESUMEN

El trabajo de titulación presenta una propuesta para la focalización del subsidio tarifa de la dignidad en los usuarios residenciales de las empresas de distribución de energía eléctrica del Ecuador. Para la focalización se utiliza la técnica de agrupación de datos considerando las características técnicas, geográficas y socioeconómicas de los usuarios.

El trabajo de titulación describe el proceso necesario para recopilar, depurar y procesar los datos a fin de que puedan ser utilizados para la focalización del subsidio. Luego, se detalla el proceso necesario para emparejar los datos y consolidarlos en una sola base de datos de usuarios residenciales. En una siguiente fase del trabajo de titulación, se define una ecuación de pobreza que considera, de manera ponderada, las características técnicas y socioeconómicas de los usuarios y que permite calcular su nivel de pobreza. Luego, se describe la forma de agrupar a los usuarios según su nivel de pobreza, así como la manera de establecer la mejor cantidad de grupos que se deben generar entre los usuarios. Finalmente, el trabajo de titulación describe la manera de definir qué grupos de usuarios y, en definitiva, qué usuarios deberían o no recibir el subsidio.

El trabajo de titulación se ha aplicado en los usuarios residenciales del Distrito Metropolitano de Quito y, para dichos usuarios, se presentan los resultados de la focalización y la estimación del ahorro que percibiría el Estado como resultado de incluir la metodología de focalización dentro de los criterios de asignación del subsidio tarifa de la dignidad.

PALABRAS CLAVE: Focalización, subsidio eléctrico, pobreza, procesamiento de datos, agrupación de datos.

ABSTRACT

This work includes a proposal for targeting the dignity tariff subsidy on residential users of electric power distribution companies in Ecuador. For targeting, the grouping data technique is used based on the technical, geographical, and socioeconomic characteristics of the users.

This work describes the process necessary to collect, refine, and process the data for targeting the subsidy. Then, this work shows the process to match the data and consolidate it into a single database of residential users. In a next phase, this work defines a poverty equation that considers, in a weighted way, the technical and socioeconomic characteristics of the users and that equation allows calculating the poverty level of the users. Then, the way to group users according to their poverty level is described, as well as the way to establish the best number of groups to be generated among users. Finally, the titling work describes how to define which user groups and, ultimately, which users should or should not receive the subsidy.

The titling work has been applied to the residential users of the Quito District Metropolitan and, for these users, the results of the targeting are showed. For the study case the estimated savings that the State would receive for including the targeting methodology are showed too.

KEYWORDS: Targeting, electricity subsidy, poverty, data processing, clustering data.

1 INTRODUCCIÓN

Actualmente, el subsidio tarifa de la dignidad, emitido mediante decreto ejecutivo No. 451-A vigente desde el 12 de julio de 2007, define que: aquellos usuarios residenciales cuyos consumos mensuales de energía sean inferiores a 110 kWh – mes en las empresas distribuidoras de la región Sierra y 130 kWh – mes en las de la región Costa/Oriente e Insular pagarán un costo preferencial de energía consumida igual a 0,04 US\$/kWh y un cargo por comercialización de 0,70 US\$ por abonado al mes, puesto que dichos usuarios representan a los quintiles 1 y 2 de ingresos según la Encuesta de Condiciones de Vida elaborada por el Instituto Nacional de Estadísticas y Cencos - INEC [1].

El subsidio tarifa de la dignidad se originó hace 13 años con el objetivo de reducir el valor de la planilla eléctrica de los usuarios de las Empresas Eléctricas con escasos recursos económicos; sin embargo, en la práctica, el subsidio ha terminado beneficiando a usuarios pobres y no pobres; y este, es justamente el problema, que el subsidio no está direccionado para beneficio exclusivo de usuarios pobres. El subsidio tarifa de la dignidad no se ha actualizado a lo largo de los años y su aplicación nunca ha dependido de las características socio económicas de los usuarios. Por tanto, es importante focalizar el subsidio tarifa de la dignidad puesto que permitiría reducir el dinero que destina el Estado para este beneficio.

La asignación del subsidio eléctrico tarifa de la dignidad ha tenido problemas desde su creación, desde el momento que se escogió como criterio de asignación condicional al consumo de energía de los quintiles más bajos de ingresos económicos del Ecuador. El problema es que no se consideró que dicho criterio condicional no es bidireccional y que, peor aún, no aplica necesariamente en el sentido inverso; es decir, el hecho de que un usuario consuma poca energía eléctrica no implica que sea pobre [2]. La literatura económica recomienda a los gobiernos la eliminación de los subsidios generalizados puesto que dan señales equívocas a los usuarios finales y distorsionan el mercado; sin embargo, la medida de eliminar algún subsidio tiene un costo político y social muy alto, razón por la cual, para disminuir los subsidios, se debe usar mecanismos de focalización basados en su adecuada asignación [3].

Si bien la Agencia de Regulación y Control de Electricidad (ARCONEL) ha tratado de focalizar la asignación del subsidio tarifa de la dignidad, utilizando el método de frecuencia de consumo estricto a 11 meses [1], este mecanismo sigue vinculado únicamente al consumo mensual de energía, parámetro que, por sí solo, no es un indicador contundente de la situación económica y social del consumidor.

Ante los evidentes inconvenientes existentes durante la aplicación del subsidio tarifa de la dignidad surgen varias preguntas como: ¿se podría reducir el error cometido al momento de asignar el subsidio? ¿es posible determinar geográficamente qué usuarios y qué zonas realmente requieren del subsidio eléctrico? Un ejemplo claro de que no se ha estudiado este tema, es el hecho de que no se puede saber con exactitud si el subsidio llega sólo a usuarios pobres, no se tienen cifras al respecto, nadie lo ha medido o ha investigado. Por tanto, al ser un trabajo especializado en el área de distribución de energía eléctrica, resulta un tema que merece estudiarse puesto que está estrictamente relacionado con el desarrollo comercial de las Empresas Eléctricas de distribución de energía eléctrica en el Ecuador [4].

El presente trabajo de titulación plantea el desarrollo una metodología general para la focalización del subsidio eléctrico tarifa de la dignidad que, en base al análisis de las características técnicas, geográficas y socioeconómicas de los usuarios, contribuya a resolver los problemas asociados a la inadecuada asignación del subsidio en las Empresas Eléctricas del Ecuador.

1.1 Objetivo General

Proponer una metodología general para la focalización del subsidio tarifa de la dignidad en las Empresas Eléctricas del Ecuador mediante el uso de la técnica de agrupamiento geográfico y el análisis de las características técnicas y socioeconómicas de sus usuarios residenciales para promover la reducción del gasto del Estado por concepto subsidio tarifa de la dignidad.

1.2 Objetivos Específicos

- Determinar las características básicas necesarias que deben poseer los datos de los usuarios residenciales del Ecuador, a fin de que puedan ser utilizados para la focalización del subsidio tarifa de la dignidad.
- Determinar la manera de preprocesar los datos de los usuarios residenciales del Ecuador, a fin de que puedan ser utilizados para la focalización del subsidio tarifa de la dignidad.

- Determinar la configuración de la técnica de agrupación para su aplicación en la focalización del subsidio eléctrico a usuarios residenciales de las Empresas Eléctricas del Ecuador.
- Aplicar la metodología general de focalización del subsidio eléctrico tarifa de la dignidad a los usuarios residenciales ubicados dentro del Distrito Metropolitano de Quito.
- Estimar el ahorro económico que podría percibir el Estado al aplicar la metodología de focalización del subsidio tarifa de la dignidad a los usuarios residenciales ubicados dentro del Distrito Metropolitano de Quito.

1.3 Alcance

El trabajo de titulación definirá una metodología general para la focalización del subsidio eléctrico tarifa de la dignidad dentro de la zona de concesión de cualquier Empresa Eléctrica del Ecuador. Para ello, se utilizará la técnica de agrupamiento geográfica, aplicada en los usuarios residenciales, con el objetivo de encontrar sectores de usuarios pobres y no pobres que serán definidos según sus características técnicas y socioeconómicas.

Inicialmente se recopilarán datos técnicos, económicos y geográficos de todos los usuarios residenciales pertenecientes a la Empresa Eléctrica definida como caso de estudio. Luego, se procesarán los datos a fin de que puedan ser manejables y compatibles entre las distintas herramientas informáticas: MATLAB, ARCGIS, Microsoft Office, entre otras. Finalmente, se armará una base de datos de usuarios residenciales que permita el emparejamiento de sus datos técnicos (consumo de energía), geográficos (ubicación georeferenciada) y socioeconómicos (relacionados a la pobreza).

Una vez que los datos de los usuarios residenciales estén listos para su uso se los filtrará y se los ingresará al software MATLAB, donde se programará un código que permita el uso de alguna de las técnicas de agrupación existentes y que permita encontrar grupos de usuarios geográficamente cercanos. Luego, se analizarán las características socioeconómicas de cada grupo resultante y se decidirá si estos pudieran ser considerados como pobres o no pobres. Dado que la técnica de agrupación requiere como variable de entrada el número de grupos a encontrar, se ejecutarán varias repeticiones del código de agrupación, cambiando el número de grupos, hasta llegar a encontrar la cantidad de grupos

que permita representar toda el área de estudio con sectores de usuarios de características socioeconómicas similares.

Como parte del trabajo de titulación se aplicará la focalización del subsidio tarifa de la dignidad a los usuarios de la Empresa Eléctrica Quito ubicados dentro del Distrito Metropolitano de Quito (DMQ) utilizando sus datos técnicos y todos los datos socioeconómicos que se pudieran obtener de los distintos actores como: el Municipio de Quito, la Agencia de Regulación y Control de Electricidad, Ministerio de Energía y Recursos Naturales no Renovables, Instituto Nacional de Estadística y Censos, entre otros [12] [13]. Al finalizar esta etapa se podrá estimar el ahorro económico que percibiría el Estado tras aplicar la metodología de focalización del subsidio tarifa de la dignidad en el DMQ.

Finalmente, con los resultados de las actividades anteriores, se definirán varios criterios técnicos conceptuales recomendados para la adecuada focalización del subsidio eléctrico y, con lo cual, se llegará a proponer una metodología general para focalizar del subsidio eléctrico tarifa de la dignidad en cualquier Empresa Eléctrica del Ecuador.

2 MARCO TEÓRICO

2.1 Fundamentos Sociales de los Subsidios

2.1.1 La Sociedad

La Real Academia de la Lengua define como sociedad al conjunto de personas que conviven bajo las mismas normas. En una sociedad las personas cumplen un rol específico, cooperan entre sí y tienen un objetivo común. En una sociedad las personas tienen deberes y derechos.

Para convivir en sociedad es primordial definir un esquema de justicia social que implica, según [5], la aplicación de al menos dos principios básicos:

- Toda persona tiene el mismo derecho a un esquema plenamente válido de libertades básicas que sea compatible con un esquema similar de libertades para todos.
- La igualdad social y económica debe satisfacer dos condiciones: la primera consiste en que los cargos jerárquicos dentro de la sociedad deben ser abiertos para todos en igualdad de oportunidades; la segunda establece que se debe buscar el mayor beneficio para los miembros menos aventajados de la sociedad.

Por tanto, el tratamiento preferente a los más débiles es una condición básica para la convivencia social.

2.1.2 La Pobreza

Una persona pobre es aquella que no tiene lo suficiente para vivir. Según [6] la pobreza está relacionada con la baja cantidad de ingresos económicos que recibe una persona y con las limitadas capacidades que tiene una persona para generar dichos ingresos económicos. A continuación, se describe la situación actual de la pobreza a nivel mundial y a nivel país.

La pobreza a nivel mundial, según el panorama general del banco mundial, tiene una tasa de pobreza extrema que disminuyó del 10,1% en el 2015 al 9,2% en el 2017, lo cual, representa una disminución de 689 millones de personas que vivían con menos de US\$ 1,90 al día. Para el año 2017, en las líneas de pobreza más altas, el 24,1% de la población mundial vivía con menos de US\$ 3,20 al día y el 43,6% con menos de US\$ 5,50 [7].

Según los resultados de la encuesta nacional de empleo, desempleo y subempleo realizada por el Instituto Nacional de Censos y Estadísticas (INEC), para diciembre de 2019, se considera a una persona pobre por ingresos si percibe un ingreso familiar per cápita menor a US\$ 84,82 mensuales y pobre extremo si percibe menos de US\$ 47,80. Así mismo, a diciembre de 2019, la pobreza a nivel nacional se ubicó en 25,0% y la pobreza extrema en 8,9% de la población. La ciudad de Quito tuvo un 8,3% de pobreza en su población, mientras que Guayaquil tuvo la mayor tasa de pobreza igual al 11,2% [8].

Las cifras de pobreza reflejan la necesidad de destinar los subsidios y recursos a quienes realmente lo necesitan.

2.1.3 El Subsidio

El subsidio es una ayuda, generalmente económica, que una persona o entidad recibe de un organismo oficial para satisfacer una determinada necesidad. En la práctica un subsidio corresponde a la diferencia entre el precio real de un bien o servicio y el precio efectivamente cobrado al consumidor por dicho bien o servicio. Un subsidio es una herramienta que permite al Estado redistribuir los recursos hacia los más necesitados a fin de que puedan disponer y hacer uso de determinados bienes o servicios [9].

Según [9] los subsidios deberían tener al menos tres características básicas:

1. Ser transitorios
2. Ser correctamente direccionados o focalizados
3. Ser eficientes

2.1.4 Experiencias Internacionales de los Subsidios

En Latinoamérica la mayoría de los países que subsidia el consumo de electricidad elige a los beneficiarios basándose únicamente en un límite de consumo de energía eléctrica mensual (kWh/mes). En Bolivia, por ejemplo, existe el subsidio Tarifa de la Dignidad que, desde el 2006, estipula reducciones tarifarias del 25% para los consumos inferiores a los 70 kWh/mes en el sector urbano e inferiores a los 30 kWh/mes en el sector rural. En Paraguay se beneficia a los usuarios con consumos menores a 150 kWh/mes considerando una tarifa igual al 25% de la tarifa normal para los consumos inferiores a 75 kWh/mes y una tarifa igual al 50% de la tarifa normal para los consumos superiores a 75 kWh/mes. Un esquema de subsidio particular aparece en Guatemala y Honduras donde el límite de consumo para acceder al subsidio, igual a 300 kWh/mes, es superior a los consumos medios de los usuarios en ambos países, lo cual ha provocado que más del 80% de los

usuarios accedan al subsidio en cada país [4]. En todos los casos mencionados el Estado financia los subsidios.

En Europa, uno de los ejemplos es el caso de España donde, según [10], aplica el bono social que consiste en una reducción del 25% del monto total de la factura eléctrica. El subsidio, en este caso, lo financian las empresas de generación de electricidad. Para aplicar al bono social se debe cumplir al menos uno de los siguientes requisitos:

- Tener una potencia contratada inferior a 3 kW.
- Tener 60 años o más de edad y ser pensionista del Sistema de Seguridad Social.
- Ser familia numerosa (que tiene 3 o 4 hijos).
- Formar parte de una unidad familiar que tenga a todos sus miembros en situación de desempleo.

Las experiencias internacionales descritas muestran implícitamente la necesidad de definir varios requisitos antes de entregar los subsidios a los usuarios finales.

2.1.5 Focalización de Subsidios

La focalización de recursos, en este caso los subsidios, es una herramienta social utilizada principalmente para el alivio de la pobreza. La focalización es una herramienta que tiene como objetivo destinar recursos a quienes más lo necesitan y solamente a ellos, los pobres [11].

El punto teórico a favor de la focalización es que: mientras más certero sea un subsidio en llegar a los pobres, menores serán el desperdicio y el costo para alcanzar el objetivo deseado de erradicar la pobreza. La focalización trata de mejorar la relación costo beneficio al momento de otorgar un beneficio. Si bien la focalización tiene algunos argumentos a favor, siempre se debe tener en cuenta los costos que genera, como: la manipulación de información, las distorsiones en los incentivos, el estigma social, las pérdidas administrativas, la invasión de la vida privada y la sustentabilidad política [12].

Por tanto, la focalización de recursos para la gente pobre debe realizarse con cierto grado de precisión, y no con tanta profundidad, puesto que se deben evitar dos errores comunes asociados a la focalización: el primero, sacar de la lista de gente pobre a personas que sí son pobres y, el segundo, incluir entre los pobres a los no pobres.

2.2 El Subsidio Tarifa de la Dignidad en el Ecuador

2.2.1 Definición

El subsidio tarifa de la dignidad, emitido mediante decreto ejecutivo No. 451-A vigente desde el 12 de julio de 2007, define que: aquellos usuarios residenciales cuyos consumos mensuales de energía sean inferiores a 110 kWh – mes en las empresas distribuidoras de la región Sierra y 130 kWh – mes en las de la región Costa/Oriente e Insular pagarán un costo preferencial de energía consumida igual a 0,04 US\$/kWh y un cargo por comercialización de 0,70 US\$ por abonado al mes, puesto que dichos usuarios representan a los quintiles 1 y 2 de ingresos según la Encuesta de Condiciones de Vida elaborada por el Instituto Nacional de Estadísticas y Cencos – INEC [1].

2.2.2 Procedimiento de Aplicación

El procedimiento de aplicación del subsidio tarifa de la dignidad, aprobado mediante Resolución No. ARCONEL-052/18, es el instrumento legal que habilita la otorgación de este beneficio. A continuación, se detallan los puntos más relevantes de dicho procedimiento.

- El artículo 15 de la Ley Orgánica del Servicio Público de Energía Eléctrica (LOSPEE), respecto de las atribuciones y deberes de la Agencia de Regulación y Control de Electricidad - ARCONEL, entre otras, dispone: “5. Realizar estudios y análisis técnicos, económicos y financieros para la elaboración de las regulaciones, pliegos tarifarios y acciones de control.”; así como, “6. Establecer los pliegos tarifarios para el servicio público de energía eléctrica y para el servicio de alumbrado público general.”.
- El artículo 59 de la LOSPEE determina: “Subsidios.- Si por circunstancias de carácter social o económico, el Estado hubiere otorgado o decidiera otorgar compensaciones, subsidios o rebajas directos y focalizados en el servicio público de energía eléctrica, a un determinado segmento de la población, mediante leyes, o políticas sectoriales, o si por intermedio de ARCONEL, aprobare o hubiere aprobado pliegos tarifarios que se ubiquen por debajo de los costos del servicio público de energía eléctrica, los valores que correspondan a estos subsidios, compensaciones o rebajas serán cubiertos por el Estado ecuatoriano, y constarán obligatoriamente en el Presupuesto General del Estado.”.
- La Disposición Transitoria Cuarta de la Ley Orgánica del Servicio Público de Energía Eléctrica dispone: “Aprobación de pliegos tarifarios.- A partir de la

aprobación de la presente ley, el pliego tarifario vigente aprobado por el CONELEC mantendrá su aplicación hasta la aprobación de un nuevo pliego tarifario sobre la base de lo establecido en esta ley y la regulación correspondiente. Los subsidios por Déficit Tarifario y Tarifa Dignidad mantendrán su vigencia en los términos y condiciones vigentes a la expedición de la presente ley, mientras no sean modificados o eliminados por el ARCONEL.”.

- Mediante Circular Nro. ARCONEL-DE-2015-002-OFC de 24 de marzo de 2015, la ARCONEL instruye a las Empresas Eléctricas de Distribución, se adopten las acciones que correspondan, para que, a efectos de la aplicación del Subsidio de la Tarifa de la Dignidad, se excluya a los “Consumo Cero” (0 kWh-mes), a partir de la facturación de abril de 2015.
- Finalmente, la Resolución No. ARCONEL 052/18, resuelve aprobar y disponer la modificación del Procedimiento para la Aplicación del Subsidio Otorgado por el Estado ecuatoriano mediante Decreto Ejecutivo No. 451-A “Subsidio de la Tarifa de la Dignidad” considerando la frecuencia de consumo estricta de 11 meses. A continuación, la Figura 2.1 muestra un ejemplo en el comportamiento de consumo mensual de energía eléctrica por usuario, a lo largo del año, con el que el usuario final podría continuar beneficiándose del subsidio según el criterio de frecuencia de consumo estricta a 11 meses.

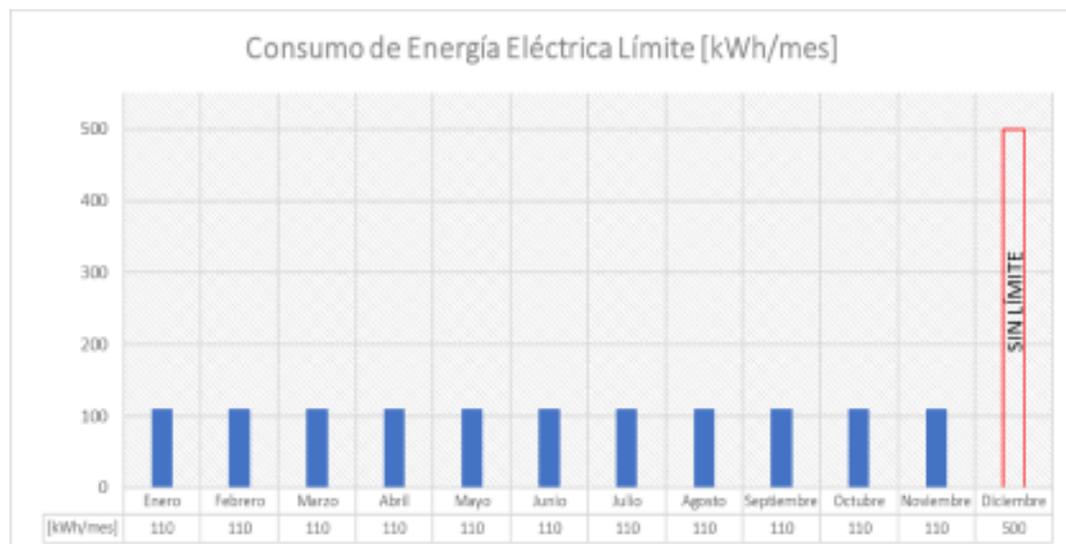


Figura 2.1. Ejemplo de frecuencia de consumo de energía estricta a 11 meses.

2.2.3 Resultados de la Aplicación del Subsidio

De acuerdo con los resultados oficialmente publicados por la ARCONEL [1], para el primer cuatrimestre del año 2019, a continuación, la Figura 2.2 y Figura 2.3 presentan la cantidad mensual de usuarios beneficiarios del subsidio tarifa de la dignidad en cada empresa eléctrica de distribución de energía y, así mismo, la Figura 2.4 y Figura 2.5 muestran los montos mensuales percibidos por cada empresa eléctrica de distribución de energía por concepto subsidio tarifa de la dignidad.

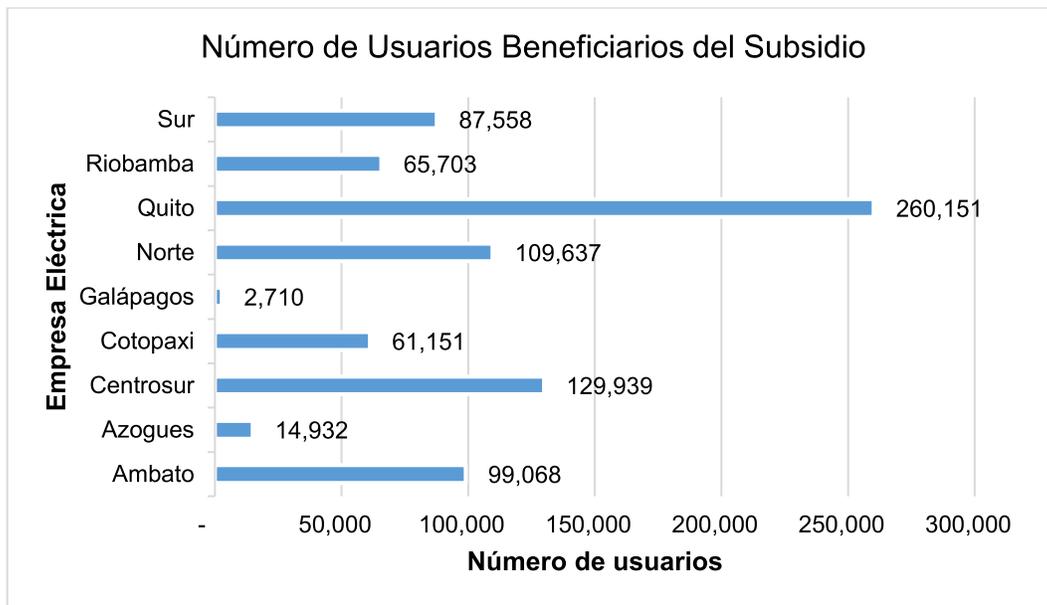


Figura 2.2. Número de usuarios beneficiarios del subsidio por empresa eléctrica de distribución fuera de CNEL EP.

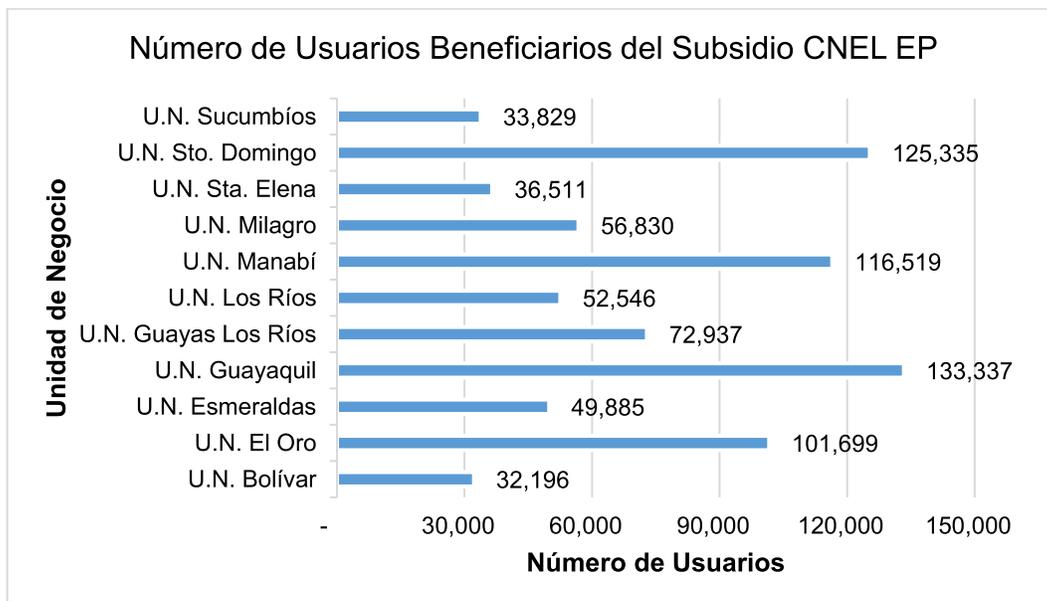


Figura 2.3. Número de usuarios beneficiarios del subsidio por empresa eléctrica de distribución en CNEL EP.

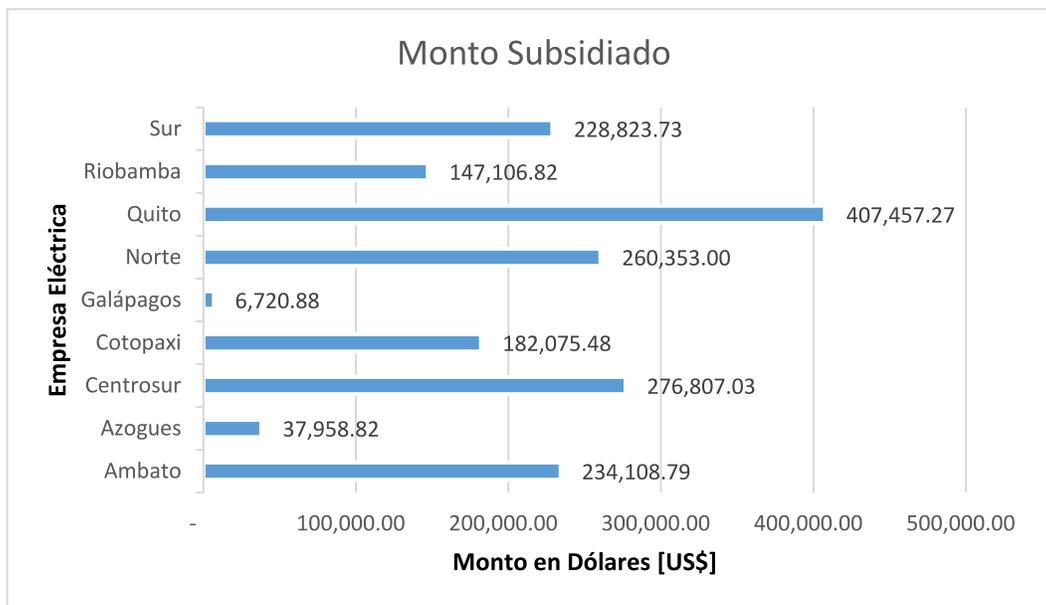


Figura 2.4. Montos percibidos por las empresas eléctricas, fuera de CNEL EP, debido al subsidio tarifa de la dignidad.

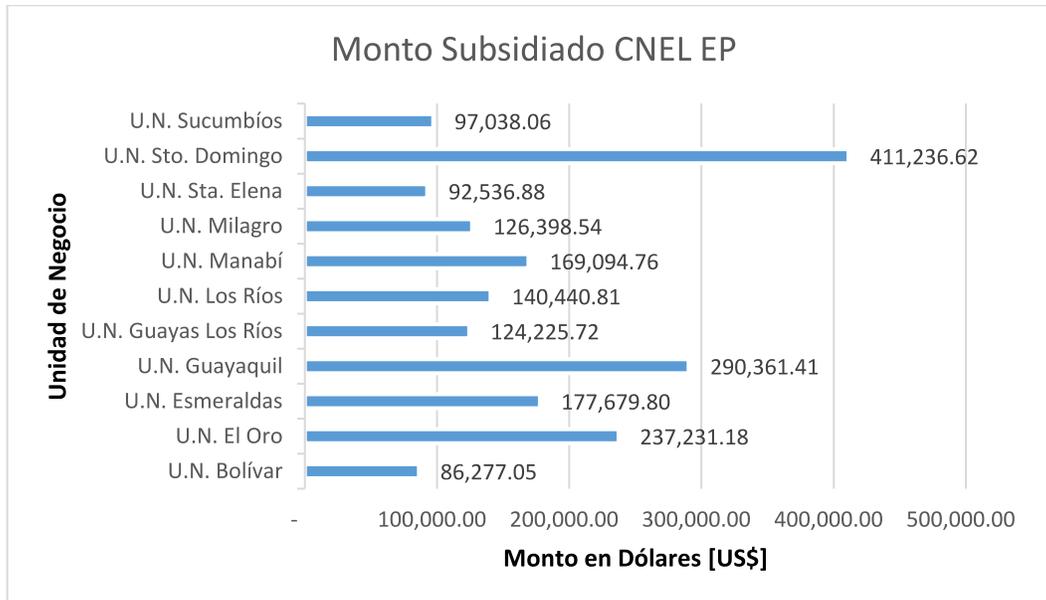


Figura 2.5. Montos percibidos por las unidades de negocio de CNEL EP debido al subsidio tarifa de la dignidad.

En términos generales, en promedio, existen cerca de 1'641.000,00 abonados mensuales que se benefician del subsidio tarifa de la dignidad en el Ecuador, esto equivale a 3'597.000,00 dólares que el Estado destina mensualmente a este subsidio; al año se subsidiaría 43'164.000,00 dólares.

2.3 Extracción del Conocimiento a partir de Datos

2.3.1 Definición

La extracción de conocimiento a partir de datos, en inglés *Knowledge Discovery in Databases* (KDD) es un proceso que utiliza la minería de datos para descubrir información significativa de grandes cantidades de datos. El objetivo del proceso KDD es permitir la toma de decisiones según la información extraída de los datos. A continuación, en la Figura 2.6 se transcribe el esquema del proceso KDD [13]:

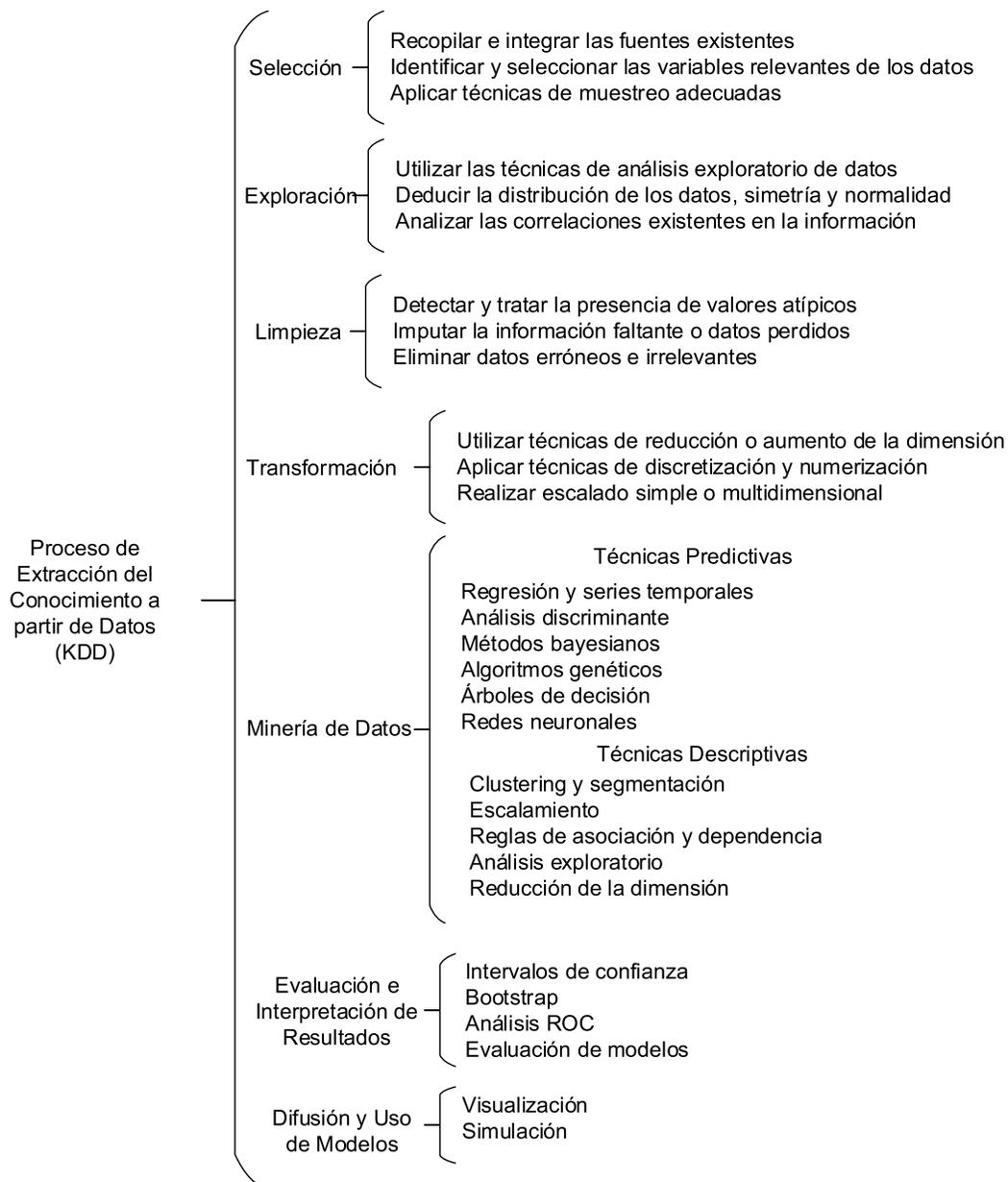


Figura 2.6. Esquema del proceso de extracción del conocimiento a partir de datos [13].

2.3.2 Selección, Limpieza y Emparejamiento de Datos

La fase de selección de datos está compuesta por actividades como la recopilación e integración de datos. En esta fase se determina qué fuentes de información son útiles para el trabajo y dónde se debe conseguir las.

Cuando ya se dispone de los datos lo primero que se debe hacer es identificar y seleccionar las variables o atributos más relevantes de los datos para el objeto de estudio. Para esta fase es común definir una muestra adecuada de datos para su análisis [13].

Una vez obtenidos los datos, la fase de limpieza de datos se compone de las siguientes actividades: primero, se define qué es un dato erróneo para el objetivo de análisis; luego, se identifican los datos erróneos dentro del conjunto de datos y, finalmente, se los corrige [14]. En esta fase, a fin de corregir los datos erróneos, se puede considerar dos opciones: eliminar el registro de datos erróneos o estimar un nuevo valor o atributo que reemplace al valor erróneo mediante alguna metodología matemática. Por ejemplo, para corregir un dato faltante o atípico es común calcular su nuevo valor mediante una interpolación lineal [15], [16].

Una vez que se ha conseguido y limpiado los datos, si el estudio lo requiere, se puede iniciar la fase de emparejamiento de datos la cual consiste en la unión de varios atributos de un dato en una sola base de datos. La unión o emparejamiento de datos se puede realizar solo si en las diferentes bases de datos existe un campo en común. Existen múltiples herramientas para el emparejamiento de datos como lo son Excel, Matlab, ArcGIS, entre otras.

2.4 Agrupamiento de Datos

2.4.1 Definición

La agrupación de datos o *clustering* es una herramienta de análisis de datos no supervisado que está enfocada en la búsqueda de grupos dentro de un conjunto de datos. A cada conjunto de datos se lo llama clúster y se caracteriza por poseer datos parecidos entre sí.

La agrupación de datos puede realizarse de diferentes maneras dependiendo del criterio de agrupación y definiendo de antemano el número de clústeres. Si el número de clústeres no es conocido la agrupación no tiene respuesta única para todos los casos y, más bien,

tiene un valor específico para cada problema. Por tanto, en la mayoría de los casos, la técnica de clustering requiere de humanos expertos en el dominio de su aplicación [17].

El procedimiento de agrupación más conocido es K-means, desarrollado por Steinhouse en 1956 y popularizado por MacQueen en 1967. Este método conceptualmente es simple y se define de la siguiente manera: si μ_1, \dots, μ_K son los K centros de los clústeres, K-means los encuentra minimizando

$$J(\mu_1, \dots, \mu_K) = \frac{1}{n} \sum_{k=1}^K \sum_{x_i \in G_k} \|x_i - \mu_k\| \quad (2.1)$$

Donde G_k es el clúster de elementos con centro más próximo a μ_k , es decir:

$$G_k = \left\{ x_i : \|x_i - \mu_k\| = \min_{1 \leq j \leq K} \|x_i - \mu_j\| \right\} \quad (2.2)$$

La limitación de este método es que no da buenos resultados si dentro de los datos existen datos atípicos o si las agrupaciones tienen grandes diferencias en su orientación, tamaño y forma [18]. Existen variaciones en la técnica de agrupación puesto que la forma de calcular la similitud $x_i - \mu_k$ puede realizarse de distintas maneras.

2.4.2 Técnicas

Se pueden mencionar tres grandes familias de técnicas de agrupación que son: clustering basado en centros, clustering jerárquico y clustering difuso [19]. Asimismo, existen múltiples algoritmos que permiten realizar la agrupación de datos, los más comunes se basan en:

- Particiones
- Jerarquías
- Densidades
- Rejillas
- Modelos
- Teoría de grafos
- Búsquedas combinatorias
- Técnicas Fuzzy
- Redes neuronales
- Datos secuenciales

2.4.3 Evaluación de Grupos

Para evaluar los resultados de la agrupación de datos existen dos tipos de indicadores, el primero, evalúa parámetros internos de los grupos de datos y, el segundo, evalúa parámetros externos de los grupos de datos.

El Índice de Davies-Bouldin es uno de los indicadores que permite evaluar los parámetros internos de los grupos de datos resultantes de la técnica de agrupación y considera que existe una mejor agrupación si la distancia media de un grupo de datos, respecto a su centro, dividida para la distancia existente entre los centros de los grupos es la menor. La definición de dicho índice es la siguiente

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left\{ \frac{\bar{d}_i + \bar{d}_j}{d_{ij}} \right\} \quad (2.3)$$

Donde DB es el índice de Davies-Bouldin, k es el número de grupos, \bar{d}_i es la distancia promedio de los datos al centro del grupo i, \bar{d}_j es la distancia promedio de los datos al centro del grupo j y d_{ij} es la distancia entre los centros de los grupos i y j. La agrupación óptima para diferentes números de grupos, según este criterio, es aquella con el menor valor de DB [20].

Otra forma de evaluar los resultados de la técnica de agrupación puede ser, por ejemplo, calculando la “pureza de los resultados” que mide la relación entre los datos de cada grupo con otros atributos externos que no fueron considerados para la agrupación.

2.4.4 Ventajas y Desventajas

A continuación, se enlistan algunas ventajas de la técnica de agrupación:

- Funciona con pocas o grandes cantidades de datos
- Puede manejar variables numéricas, binarias, ordinales, entre otras
- Como dato de entrada se requiere únicamente el número de clústeres

A continuación, se enlistan algunas desventajas o limitaciones de la técnica de agrupación:

- Los clústeres basados en distancias numéricas tienden a encontrar clústeres circulares
- Los clústeres basados en distancias numéricas pueden no dar resultados adecuados si los datos poseen valores atípicos

2.4.5 Aplicaciones

La técnica de agrupación puede utilizarse parcialmente o en su totalidad para los siguientes casos [21]:

- Banca y Comercio: Clasificación de clientes.
- Medicina y Farmacia: Clasificación de enfermedades y aplicación de tratamientos.
- Seguridad y detección de fraude: Reconocimiento facial.
- Recuperación de información: Agrupación de imágenes, texto, videos, sonidos.
- Astronomía: Identificación de nuevas galaxias.
- Geología, agricultura, minería y pesca: Identificación de áreas para distintos usos.
- Ciencias ambientales: Identificación de ecosistemas.
- Ciencias sociales: Planificación de ciudades.

Existe gran cantidad de aplicaciones más desarrolladas donde se ha utilizado la técnica de agrupación, por ejemplo, para el análisis de imágenes satelitales o de alta resolución, donde el objetivo ha sido encontrar objetos atípicos como un barco o un tornillo [18].

2.5 Herramientas Computacionales

2.5.1 Software ArcGIS

ArcGIS es un sistema que permite recopilar, organizar, administrar, analizar, compartir y distribuir información geográfica. A continuación, se enlista algunas de las actividades que se pueden realizar con dicho software.

- Crear, compartir y utilizar mapas inteligentes
- Compilar información geográfica
- Crear y administrar bases de datos geográficas
- Resolver problemas con el análisis espacial
- Crear aplicaciones basadas en mapas
- Dar a conocer y compartir información mediante la geografía y la visualización de mapas

ArcGIS permite la elaboración de mapas, basados en información geográfica, que sirven como herramienta para visualizar, explorar, analizar y actualizar datos espaciales. Los mapas, junto con su análisis espacial, sirven también como herramienta para buscar y comprender patrones y relaciones; realizar análisis y modelado a fin de resolver problemas específicos; visualizar y realizar un seguimiento geográfico del desarrollo de las ciudades para dar a conocer ideas, planes y diseños [22].

2.5.2 Software MATLAB

MATLAB es software, usado habitualmente en la Ingeniería, que sirve para el manejo y análisis de datos con un lenguaje de programación expresado directamente en la matemática de matrices y arreglos [23]. MATLAB posee algoritmos preprogramados que se pueden usar de manera sencilla y que permiten el desarrollo de aplicaciones prácticas sin la necesidad de ser un experto en programación. Por ejemplo, para el análisis de datos y el desarrollo de aplicaciones o estudios, MATLAB ofrece las siguientes herramientas:

- Machine learning
- Deep learning
- Visión artificial
- Procesamiento de señales
- Finanzas cuantitativas y gestión de riesgos
- Robótica
- Sistemas de control

3 METODOLOGÍA

En esta sección del Trabajo de Titulación se describe el proceso realizado para el desarrollo de una propuesta de metodología para la focalización del subsidio tarifa de la dignidad, basado en la agrupación de usuarios según sus características técnicas, geográficas y económicas, en las empresas eléctricas del Ecuador.

3.1 Recopilación y Procesamiento de Datos

En esta fase se ha revisado la información técnica, geográfica y socioeconómica de los usuarios residenciales del Ecuador que puede estar al acceso del personal de las empresas eléctricas. Una vez revisada la información disponible se ha determinado el uso de los siguientes datos:

- Datos técnicos: Consumo mensual de energía eléctrica por usuario en kWh.
- Datos geográficos: Ubicación georeferenciada de cada usuario, coordenadas (X,Y).
- Datos socioeconómicos: El valor básico referencial del metro cuadrado de suelo y los ingresos familiares mensuales promedio por sector.

3.1.1 Datos Técnicos

Los datos técnicos de los usuarios residenciales, desde el punto de vista eléctrico, deben comprender únicamente dos atributos: el código del cliente y el consumo de energía mensual. Estos datos pueden ser obtenidos de la base de datos del sistema comercial de cada empresa eléctrica y deberán ser filtrados considerando las siguientes condiciones:

- Se considera solo a usuarios con consumos inferiores a 110 kWh en la sierra o 130 kWh en la costa.
- Se considera solo a usuarios residenciales.

Dentro de la codificación de tarifas, según [24] y [25], correspondería a los usuarios con código BTCRSD01 (Residencial), BTCRSD02 (Residencial Temporal) y BTCRSD03 (Residencial para el programa de eficiencia energética en la cocción, PEC).

- No se han considerado usuarios, con bajos consumos, dentro de los grupos de consumo general-comercial, general-industrial y general-otros.

- Se retira a los usuarios con registros de consumo de energía igual a 0 kWh.
- Se retira a los usuarios con registros de consumo de energía negativos.
- Se considera a los usuarios independientemente de si pertenecen o no a la tercera edad. La pobreza puede estar presente en personas de cualquier edad.
- Se considera a todos los usuarios de tipo público o privado.
- Se retira a los usuarios con medición y conexión en medio voltaje puesto que dicho nivel de voltaje es utilizado principalmente para usuarios industriales.
- Se retira a los usuarios en los que el código de cliente posea valor cero o negativo.

Como resultado de esta fase, la Tabla 3.1 muestra un ejemplo de la estructura de los datos obtenidos:

Tabla 3.1 Estructura de los datos técnicos de los usuarios residenciales

| Código del Cliente | Consumo de Energía Mensual [kWh] |
|---------------------------|---|
| 851256 | 90 |
| ... | ... |
| ... | ... |
| ... | ... |
| 423552 | 75 |

Se debe evitar compilar información personal de los usuarios como: nombres, números de cédula, entre otros. El uso de grandes cantidades de datos podría complicar su manejo en los computadores. Por otro lado, la disponibilidad de datos personales podría servir para fines ajenos a la focalización del subsidio.

3.1.2 Datos Geográficos

Los datos geográficos de los usuarios residenciales del Ecuador se los obtienen de las geodatabases de las empresas eléctricas. Los datos geográficos deben estar compuestos por solo tres atributos: el código del cliente, la coordenada geográfica X y la coordenada geográfica Y. El sistema de coordenadas utilizado para ubicar a los usuarios eléctricos en el Ecuador es el sistema de coordenadas proyectado UTM WGS 1984 zona 17 Sur.

Los datos espaciales de los usuarios residenciales habitualmente estarán disponibles para su análisis mediante geodatabases, las cuales se pueden manipular y visualizar mediante

el software ArcGIS. Si la geodatabase presenta inconsistencias en sus datos se puede recuperarlos transcribiendo el siguiente código en la ventana de programación en Python, propia del software ArcGIS:

```
arcpy.RecoverFileGDB_management('C:\Users\Edison\Desktop\EEQ201801.gdb',  
'C:\Users\Edison\Desktop\BD RECUPERADAS', '201801new.gdb')
```

El código presentado requiere la definición de tres argumentos: el primero corresponde a la ubicación, en el computador, de la base de datos con inconsistencias; el segundo corresponde a la ubicación, en el computador, de la nueva base de datos que será recuperada; y, el último, corresponde al nombre que se le desea asignar a la nueva base de datos recuperada.

Una vez que los datos están disponibles para su uso se debe depurar los registros considerando los siguientes criterios:

- Se deben retirar los registros vacíos que no posean un valor en el código del cliente o en sus coordenadas geográficas. El código cliente y las coordenadas geográficas son valores mayores y distintos que cero.
- Se debe retirar registros duplicados.
- Se deben retirar los registros con valores tipo texto. El código cliente y las coordenadas geográficas son valores tipo número.
- Se deben crear nuevos registros para los casos en los que existe una misma ubicación para varios usuarios. Esta actividad puede realizarse utilizando Excel o Matlab y, en la siguiente etapa, facilitará el emparejamiento de datos.

Como resultado de esta fase, la Tabla 3.2 muestra un ejemplo de la estructura de los datos obtenidos:

Tabla 3.2 Estructura de los datos geográficos de los usuarios residenciales

| Código del Cliente | Coordenada X | Coordenada Y |
|---------------------------|---------------------|---------------------|
| 387320 | 780572.91 | 9981908.61 |
| ... | ... | ... |
| ... | ... | ... |
| ... | ... | ... |
| 1499703 | 776101.09 | 9977057.11 |

Otro dato geográfico necesario para la focalización del subsidio es la división político-administrativa del Ecuador, la cual contiene la codificación de provincias, cantones, zonas y sectores del Ecuador. Estos datos pueden encontrarse en la información cartográfica del Instituto Nacional de Estadísticas y Censos (INEC), en [26], y suelen presentarse en una base de datos de ArcGIS mediante un shapefile de datos tipo polígono. Los datos más útiles de la división político-administrativa del Ecuador para la focalización del subsidio se muestran en la Tabla 3.3.

Tabla 3.3 Estructura de los datos de la división político-administrativa del Ecuador

| Identificador Geográfico | Tipo de Dato | DPA_ZONA |
|--------------------------|--------------|-----------|
| 1627 | Polígono | 170150200 |
| ... | ... | ... |
| ... | ... | ... |
| ... | ... | ... |
| 2048 | Polígono | 170150272 |

3.1.3 Datos Socioeconómicos

Los datos socioeconómicos utilizados para este trabajo de titulación son la valoración económica del suelo y los ingresos económicos promedio de los núcleos familiares dentro de las zonas de concesión de las empresas eléctricas.

Los datos de valoración económica del suelo, por metro cuadrado geográficamente ubicado, se los pueden obtener de los Municipios en archivos shapefile de ArcGIS y deben poseer, al menos, los datos mostrados en la Tabla 3.4.

Tabla 3.4 Estructura de los datos de valoración económica del suelo

| Identificador geográfico | Tipo de dato | Valoración del suelo [US\$/m ²] |
|--------------------------|--------------|---|
| 250 | Polígono | 120 |
| ... | ... | ... |
| ... | ... | ... |
| ... | ... | ... |
| 2303 | Polígono | 500 |

Los datos de ingresos económicos de los núcleos familiares muestreados en el Ecuador se los puede conseguir de las encuestas realizadas por el Instituto Nacional de Estadísticas

y Censos (INEC). La encuesta de ingresos económicos por núcleos familiares ubicados por zonas y sectores está disponible en [27] mediante un archivo de datos en formato compatible con el software SPSS de IBM. Si bien la encuesta de ingresos y gastos realizada en el Ecuador posee mucha más información económica se deben depurar los datos a fin de estructurar solo aquellos mostrados en la Tabla 3.5.

Tabla 3.5 Estructura de los datos de ingreso económico por núcleo familiar

| Código Sector | Ingreso Económico Medio por Encuestado [US\$/mes] |
|---------------|---|
| 170150002006 | 977.11 |
| ... | ... |
| ... | ... |
| ... | ... |
| 170150104007 | 257.43 |

3.2 Emparejamiento de Datos

El emparejamiento de datos se lo puede realizar mediante cualquier algoritmo de emparejamiento y/o mediante cualquier software que posea esta función. Para el trabajo de titulación se ha seleccionado el uso del software Matlab, ArcGIS y Excel.

3.2.1 Emparejamiento de Datos Técnicos y Geográficos

Para emparejar los datos técnicos y económicos primero se cargan los datos técnicos en Matlab como una matriz de m códigos de cada cliente y sus n consumos de energía. Luego, se cargan los datos geográficos como una matriz de p códigos de cada cliente y q atributos, en este caso, las coordenadas X y Y. Una vez cargados los datos técnicos y geográficos se los empareja utilizando como atributo común los códigos de los clientes. A continuación, se transcribe un ejemplo de código en Matlab que permite este emparejamiento de datos.

```

clc
clear all

%dt es el archivo que contiene en la primera columna los códigos de los
%clientes y en la segunda columna los consumos mensuales en kWh
filename = 'dt.xlsx'; % dg es el archivo excel que contiene los datos
técnicos
dt = xlsread(filename);

%dg es el archivo que contiene en la primera columna los códigos de los
%clientes, en la segunda columna la coordenada X del cliente y en la
%tercera columna la coordenada Y del cliente

```

```

filename = 'dg.xlsx'; % dg es el archivo excel que contiene los datos
geográficos
dg = xlsread(filename);

[m,n]=size(dt); %m es el número de clientes y n la cantidad de atributos
de los datos técnicos
[p,q]=size(dt); %p es el número de clientes y q la cantidad de atributos
de los datos geográficos

cont=0; %cont es un contador que permite conocer la cantidad de datos
emparejados
df=0; %df es la matriz de datos finales resultado del emparejamiento

for i=1:1:m %lazo for para comparar valores en todo el rango de datos
técnicos
    for j=1:1:p %lazo for para comparar valores en todo el rango de datos
geográficos
        if dt(i,1)==dg(j,1) %si los códigos de los clientes son iguales
se emparejan los datos
            df(i,1)=dt(i,1); %se carga el código del cliente
            df(i,2)=dg(j,2); %se carga la coordenada X
            df(i,3)=dg(j,3); %se carga la coordenada Y
            df(i,4)=dt(i,2); %se carga el consumo en kWh
            cont=cont+1;
        end
    end
end

xlswrite('df.xlsx',df); % se exportan los resultados emparejados a Excel

```

El resultado de este primer emparejamiento de datos se observa en la Tabla 3.6.

Tabla 3.6 Estructura de los datos técnicos y geográficos emparejados

| Código del Cliente | Coordenada X | Coordenada Y | Consumo de Energía [kWh/mes] |
|---------------------------|---------------------|---------------------|-------------------------------------|
| 387320 | 780572.91 | 9981908.61 | 90 |
| ... | ... | ... | ... |
| ... | ... | ... | ... |
| ... | ... | ... | ... |
| 1499703 | 776101.09 | 9977057.11 | 75 |

3.2.2 Emparejamiento de Datos Económicos y Geográficos

Esta fase tiene como objetivo ubicar espacialmente los datos económicos de valoración del suelo y de ingresos económicos. Dado que los datos de valoración del suelo se los puede encontrar directamente en archivos geográficamente referenciados no hace falta emparejarlos. Sin embargo, en el caso de los datos de ingresos económicos por núcleo

familiar, es necesario primero inferir algunos datos para, luego, emparejarlos con su ubicación geográfica.

Antes de ubicar espacialmente los datos de ingresos económicos por núcleo familiar primero se han promediado las respuestas de los encuestados dentro de un mismo sector a fin de inferir que dicho resultado es el valor de ingresos por zona. Esta actividad se la puede realizar, por ejemplo, con el siguiente código.

```
%En esta sección se limpian las variables de matlab
clc
clear all
close all

%En esta sección se cargan los datos de ingresos obtenidos del INEC
filename = 'ingresosINEC.xlsx';
dinec = xlsread(filename);

%En esta sección se promedian los datos de ingresos por núcleo familiar
%El código funciona si los datos están ordenados tal como se los descarga
%de la página del INEC.

[m,n] = size(dinec); %se valora la cantidad de datos

cont = 1; %contador que permite conocer la cantidad de encuestados de un
mismo sector
j = 0; %variable que permite numerar los resultados
sum = dinec(1,2); %sumador que permite conocer la suma de ingresos de los
encuestados de un mismo sector
dpromsec = 0; %variable que contiene los resultados promedio de ingresos
por sector

for i = 2:1:m
    if dinec(i,1) == dinec(i-1,1) %compara si un encuestado pertenece al
mismo sector
        sum = sum+dinec(i,2); %si los encuestados pertenecen al mismo
sector se suman sus ingresos
        cont = cont+1; %se cuenta los encuestados que pertenecen al mismo
sector
    else
        j = j+1;
        dpromsec(j,1) = dinec(i-1,1); %se define el código de la zona en
los resultados
        dpromsec(j,2) = sum/cont; %se calcula y asigna el valor promedio
de ingresos a cada zona resultante
        sum = dinec(i,2);
        cont = 1;
    end
end

xlswrite('ingresossector.xlsx',dpromsec); % se exportan los resultados a
Excel
```

El cálculo resultado de los valores promedio de ingresos por zona tienen la estructura mostrada en la Tabla 3.7.

Tabla 3.7 Estructura de los datos de ingresos económicos promedio por zona

| Código Zona | Ingreso Económico Medio por Zona [US\$/mes] |
|--------------------|--|
| 170150002 | 750.11 |
| ... | ... |
| ... | ... |
| ... | ... |
| 170150104 | 320.43 |

Una vez obtenidos los resultados de ingresos promedio por zona se los puede emparejar con los datos geográficos de la división político-administrativa del Ecuador. Esta actividad se la puede realizar mediante la herramienta “Join” de ArcGIS utilizando como criterio de emparejamiento el código de la zona. Los resultados de este emparejamiento deben tener la estructura mostrada en la Tabla 3.8.

Tabla 3.8 Estructura de los datos de ingresos económicos promedio por zona

| Identificador Geográfico | Tipo de Dato | DPA_ZONA | Ingreso Económico Promedio por Zona [US\$/mes] |
|---------------------------------|---------------------|-----------------|---|
| 1627 | Polígono | 170150200 | 750.11 |
| ... | ... | ... | ... |
| ... | ... | ... | ... |
| ... | ... | ... | ... |
| 2048 | Polígono | 170150272 | 320.43 |

3.2.3 Emparejamiento de Datos Geográficos, Técnicos y Socioeconómicos

En esta sección se describe el proceso necesario para emparejar los datos geográficos (X,Y), técnicos (kWh) y socioeconómicos (Valor del suelo e ingresos promedio) en un solo archivo o base de datos.

Debido a que los datos provienen en distintos formatos y de distintas fuentes es necesario primero enlazar los datos técnicos de consumo con la ubicación de los usuarios residenciales; luego, es necesario enlazar los datos socioeconómicos con su ubicación

espacial para, finalmente, emparejar todos los datos en un solo archivos. Este último emparejamiento requiere del siguiente procedimiento:

1. Se debe generar un archivo tipo ráster con los datos de las áreas de valoración del suelo por metro cuadrado. Esta actividad se la puede realizar mediante la herramienta de conversión de datos de ArcGIS "*Polygon to Raster*" teniendo en cuenta que el valor de la rasterización es el valor del suelo por metro cuadrado.
2. Se debe generar un archivo tipo ráster con los datos de ingresos económicos promedio por zona. Esta actividad se la puede realizar mediante la herramienta de conversión de datos de ArcGIS "*Polygon to Raster*" teniendo en cuenta que el valor de la rasterización es el valor de ingresos económicos promedio por sector.
3. Una vez que los datos socioeconómicos se encuentran geográficamente ubicados, en un archivo tipo ráster, se los debe emparejar con los datos de ubicación de los usuarios residenciales y sus consumos; lo cual, se puede realizar mediante la herramienta de análisis espacial de ArcGIS "*Extract Values to Points*".

Los resultados del emparejamiento final de datos permiten obtener una base de datos con la estructura mostrada en la Tabla 3.9.

Tabla 3.9 Estructura de los datos de ubicación, consumo de energía, valor del suelo e ingresos económicos por usuario residencial

| Identificador de Usuario | Tipo de Dato | Código Cliente | Coordenada X | Coordenada Y | Consumo de Energía [kWh/mes] | Valor del Metro Cuadrado del Suelo por la Ubicación del Cliente [US\$/m ²] | Ingreso Económico Promedio por Cliente [US\$/mes] |
|--------------------------|--------------|----------------|---------------|---------------|------------------------------|--|---|
| Usuario 1 | Punto | 91931 | 766272.148125 | 10023235.6216 | 56 | 4.06 | 595.27 |
| Usuario 2 | Punto | 188454 | 747992.739375 | 10014971.7066 | 89 | 35 | 595.27 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| Usuario 100 | Punto | 212088 | 783591.05 | 9997044.49237 | 104 | 20 | 1545.73 |
| Usuario 101 | Punto | 870731 | 783483.529375 | 9999165.057 | 32 | 100 | 320.43 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| Usuario 1.000 | Punto | 334089 | 784199.51025 | 9996373.60837 | 55 | 140 | 1195.41 |

3.3 Agrupación de Datos

3.3.1 Definición de una Ecuación de Pobreza

Para la agrupación de los usuarios se definirá primero una propuesta de función de pobreza que considere, de manera ponderada, los criterios técnicos y socioeconómicos de los usuarios.

A continuación, se plantea la ecuación de pobreza en términos generales.

$$P(X) = w_1 \cdot X_1 + w_2 \cdot X_2 + w_3 \cdot X_3 + \dots + w_n \cdot X_n \quad (3.1)$$

P es la pobreza que depende de X datos o criterios normalizados, entre 0 y 1, ponderados con w pesos o costos; para la definición de la pobreza podrán existir n criterios, donde la suma de sus w_n pesos deberá ser siempre 1.

La normalización de datos, entre 0 y 1, de las variables X se realiza según la ecuación 3.2 y para su aplicación se debe conocer el valor máximo y mínimo ($X_{máx}$ y $X_{mín}$) de los datos originales.

$$X_{normalizada} = \frac{X_{original} - X_{mín}}{X_{máx} - X_{mín}} \quad (3.2)$$

La normalización de datos permite reducir el efecto de valores atípicos, extra tendenciales o demasiado grandes, volviéndolos adimensionales para el cálculo de la pobreza.

Una vez procesados los datos, considerando aquellos disponibles para este trabajo de titulación, la ecuación de pobreza puede definirse de la siguiente manera.

$$P(X_{criterios}) = w_{técnico} \cdot X_{técnico} + w_{suelo} \cdot X_{suelo} + w_{ingresos} \cdot X_{ingresos} \quad (3.3)$$

Donde:

$w_{técnico}$: Es la ponderación que se le asigna a los datos de consumo de energía en relación con la pobreza.

$X_{técnico}$: Son los datos normalizados de consumo de energía por usuario.

w_{suelo} : Es la ponderación que se le asigna a los datos de valoración del suelo en el que reside el usuario en relación con la pobreza.

X_{suelo} : Son los datos normalizados de valoración del suelo según la ubicación del usuario.

$w_{ingresos}$: Es la ponderación que se le asigna a los datos de ingresos económicos de los usuarios con relación a la pobreza.

$X_{ingresos}$: Son los datos normalizados de ingresos económicos por usuario.

Dado que los datos ya estarían normalizados se debe readecuar el cálculo de pobreza a fin de que tenga correspondencia con los resultados de pobreza esperados, 1 para usuarios pobres y 0 para usuarios no pobres. A continuación, se presenta una primera ecuación para el cálculo de pobreza por usuario propuesta para la focalización, la cual considera igual ponderación para cada criterio técnico, socioeconómico y geográfico.

$$P(X_{criterios}) = \frac{1}{3} * (1 - X_{técnico}) + \frac{1}{3} * (1 - X_{suelo}) + \frac{1}{3} * (1 - X_{ingresos}) \quad (3.4)$$

La ecuación anterior permite considerar que los usuarios con consumos de energía más bajos y que residen en sectores con valores de suelo e ingresos promedio más bajos son personas más pobres.

Si bien la Ecuación 3.4 es una muy buena aproximación para el cálculo de pobreza es necesario ajustar los pesos de los criterios considerando la calidad de los datos, la relación de los datos con la pobreza y los objetivos que tiene la focalización. La siguiente ecuación muestra el cálculo de pobreza propuesto en este trabajo de titulación.

$$P(X_{criterios}) = 0,1 * (1 - X_{técnico}) + 0,6 * (1 - X_{suelo}) + 0,3 * (1 - X_{ingresos}) \quad (3.5)$$

Para el criterio de consumo de energía se ha considerado un 10% de la ponderación total puesto que, a pesar de que los datos son buena calidad, no se tiene certeza de si un menor consumo de energía está asociado con la existencia de un usuario pobre. Para el criterio de valoración del suelo se ha considerado un 60% de la ponderación total puesto que la información es de calidad, tiene relación con la pobreza y cumple con el enfoque que se le quiso dar a la focalización. Finalmente, para el criterio de los ingresos económicos por núcleo familiar se le ha asignado solo un 30% de la ponderación total puesto que, si bien es un criterio importante para la focalización del subsidio, los datos que se obtuvieron del INEC, en varios casos, no presentan valores de ingresos registrados por los encuestados. En cualquier caso, los pesos o ponderación de cada criterio podrán ser ajustados por los organismos Estatales encargados de ejecutar la focalización.

3.3.2 Definición de la Técnica de Agrupación

Para definir la manera de resolver el problema de focalización, planteado en este trabajo de titulación, se han analizado tres principales aspectos:

- La naturaleza del problema
- La naturaleza de los datos
- La practicidad de la técnica de solución al problema

La naturaleza del problema: la focalización del subsidio eléctrico requiere necesariamente de alguna técnica que permita definir qué usuarios sí deben recibir el subsidio y qué usuarios no deben recibirlo; esto, solo se puede hacer mediante la técnica de agrupación de datos o mediante la técnica de clasificación de datos. La diferencia entre la técnica de agrupación de datos y la técnica de clasificación de datos es que en la técnica de agrupación de datos no se requiere conocer, como dato de entrada, si un usuario es pobre o no pobre; contrario a la técnica de clasificación de datos. Para el caso del Ecuador, dado que de entrada no se conoce qué usuarios son pobres o no pobres, la solución al problema de focalización demanda necesariamente del uso de la técnica de agrupación de datos.

La naturaleza de los datos y la practicidad de la técnica de solución al problema: para la focalización del subsidio eléctrico en los usuarios residenciales del Ecuador se han seleccionado variables técnicas y socioeconómicas que, por naturaleza, contienen valores solamente numéricos y que pueden ser tratados mediante arreglos de matrices o vectores. Dado que la técnica de agrupación de datos k-means puede trabajar con variables de datos inmersas en matrices y vectores y, dado que, además solo requiere de la definición del número de grupos que se desean obtener de la agrupación, se ha decidido, por practicidad, realizar la focalización con dicha técnica de agrupación de datos. Para la configuración de la técnica k-means en Matlab se ha mantenido, como medida de la similitud predeterminada, el valor de la distancia euclídeana cuadrada entre datos.

En todo caso, para el presente trabajo de titulación, no ha sido necesario realizar un análisis comparativo sobre el uso de los diferentes algoritmos de agrupación o sobre el uso de diferentes formas de calcular la similitud de los datos puesto que con cualquier algoritmo de agrupación y con cualquier forma de calcular la similitud de los datos siempre se puede llegar a obtener agrupaciones válidas de usuarios para la focalización del subsidio.

La técnica de agrupación definida y propuesta para la focalización del subsidio es *kmeans* y requiere, como datos de entrada, los valores resultantes de la ecuación de pobreza por cada usuario residencial y el número de grupos que se desean encontrar; la forma de definir la cantidad de grupos se estudiará en la sección 3.3.3.

La agrupación de usuarios junto con el cálculo del nivel de pobreza, definido en la sección anterior, se ha implementado en el software Matlab y se muestra en el código a continuación.

```
%En esta sección se borran las variables y datos previamente almacenados
%en Matlab
clc
clear all
close all

%En esta sección se importan los datos (depurados y emparejados) de la
%base de datos para la focalización
filename = 'Basededatosfocalización.xlsx';
ddff = xlsread(filename);

%En esta sección se definen los datos técnicos y socioeconómicos de cada
%usuario
dt = ddff(:,6); %Se definen los datos técnicos
daivas = ddff(:,7); %Se definen los datos de valoración del suelo
dingresos = ddff(:,8); %Se definen los datos de ingresos económicos

%En esta sección se normalizan los datos técnicos y socioeconómicos en un
%rango entre 0 y 1
dtn = normalize(dt, 'range'); %Se normaliza los datos técnicos
daivasn = normalize(daivas, 'range'); %Se normaliza los datos de
valoración del suelo
dingresosn = normalize(dingresos, 'range'); %Se normaliza los datos de
ingresos económicos

%En esta sección se calcula el nivel de pobreza por usuario para la
%focalización
p = 0.1*(1-dtn) + 0.6*(1-daivasn) + 0.3*(1-dingresosn);
%Valores de p cercanos a 1 indican que el usuario es un usuario pobre y
valores cercanos a 0 indicarían que el usuario es un usuario no pobre.

%En esta sección se agrupa a los usuarios según su nivel de pobreza
%considerando inicialmente, por ejemplo, 5 grupos.
k = 5;
[idx,C,sumd] = kmeans(p,k);

%En esta sección se agregan a la base de datos los índices de pobreza y
%de agrupación por cada usuario
ddff(:,9)=p;
ddff(:,10)=idx;

%En esta sección se exportan los resultados de la agrupación a Excel
xlswrite('grupopobrek13.xlsx',ddff);
```

3.3.3 Evaluación y Definición de Grupos

Para la evaluación y definición de grupos se ha decidido calcular el índice de Davies-Bouldin para distintas cantidades de grupos de usuarios, al menos entre 2 y 20 grupos, donde la cantidad de grupos recomendada para la focalización será aquella que minimice el índice Davies-Bouldin. Esta evaluación se puede realizar mediante la implementación del siguiente código en Matlab.

```
%Código para evaluar la mejor cantidad de grupos según el índice de
%Davies-Bouldin.

%Primero se define la variable de evaluación E y se utiliza la función
%predefinida evalcluster.

%evalcluster necesita como dato de entrada la variable en la que se desea
%realizar la agrupación, el índice que se desea evaluar en cada grupo y
%la cantidad de grupos que se desean evaluar, recomendado en al menos
%entre 2 y 20.

E = evalcluster(p, 'kmeans', 'DaviesBouldin', 'klist', [2:20]);
figure;
plot(E)

%El índice Davies-Bouldin calculado para cada agrupación se puede
%observar al graficar la variable E.
```

3.3.4 Criterio de asignación

Una vez que se ha logrado determinar la mejor cantidad de grupos de usuarios para la focalización se debe definir qué grupos de usuarios deben o no recibir el subsidio. Existen dos maneras de definir qué grupos de usuarios, que actualmente reciben el subsidio, deberían a futuro recibir este beneficio. La primera forma, la más restrictiva, se enfoca en asignar el subsidio a los usuarios con mayor nivel de pobreza los “más pobres” y, la segunda forma, la menos restrictiva, se enfoca en dejar de otorgar el subsidio a los grupos de usuarios con menor nivel de pobreza los “menos pobres”.

Dado que el subsidio surgió inicialmente con el objetivo de favorecer a los usuarios de los dos quintiles con ingresos económicos más bajos en el Ecuador, es decir el 40% de la población. Se propone, de manera similar, que para este trabajo de titulación se considere la focalización del subsidio para el 40% de usuarios “más pobres” según el cálculo de pobreza de la sección 3.3.1.

El porcentaje de focalización lo podrían definir los organismos del Estado de acuerdo con los resultados que se esperasen obtener de la focalización.

4 APLICACIÓN Y RESULTADOS

En esta sección se presentan los resultados de aplicar la metodología propuesta para la focalización del subsidio tarifa de la dignidad en los usuarios residenciales de la Empresa Eléctrica Quito (EEQ) dentro del Distrito Metropolitano de Quito (DMQ). Para el caso de estudio los datos técnicos obtenidos corresponden a las planillas de consumo y a la geodatabase de la EEQ para el mes de enero de 2018; los datos de la Encuesta Nacional de Ingresos y Vivienda y la división político-administrativa del Ecuador corresponden a los datos del año 2010 y, finalmente, los datos económicos de valoración del suelo en el DMQ corresponden a los datos del año 2020.

4.1 Estructuración de la Base de Datos

4.1.1 Datos Técnicos

La EEQ posee dentro de su zona de concesión, la provincia de Pichincha y parte de la provincia del Napo, cerca de 1.048.575 usuarios. A continuación, la Figura 4.1 y la Figura 4.2 muestran la ubicación geográfica de dichos usuarios en un mapa.



Figura 4.1. Usuarios de la Empresa Eléctrica Quito.

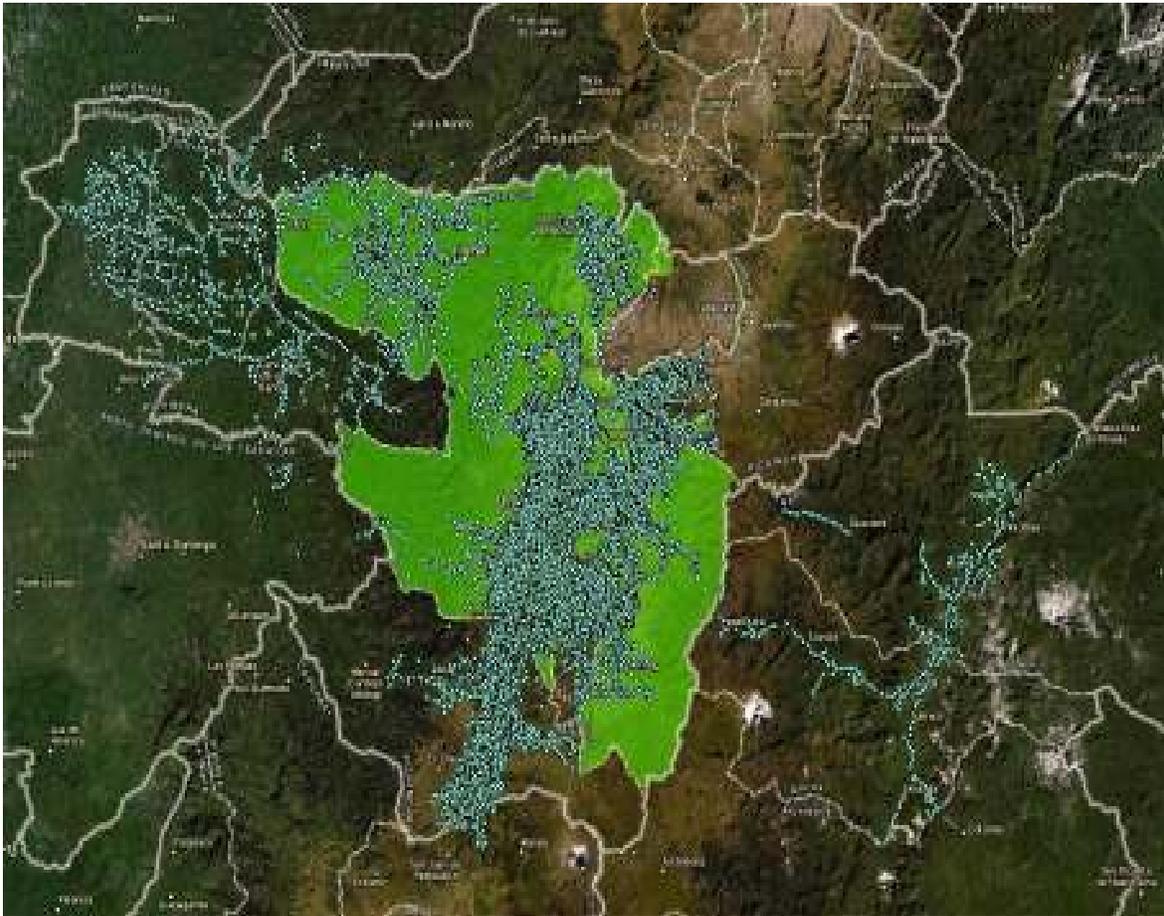


Figura 4.2. Usuarios de la Empresa Eléctrica Quito ampliado en el DMQ.

Los usuarios de la EEQ han sido depurados y emparejados según lo descrito en el capítulo 3 y se han seleccionado, para la focalización, aquellos ubicados dentro del DMQ. La cantidad de usuarios depurados que se han encontrado dentro del DMQ y que reciben el subsidio son cerca de 227.534 usuarios. A continuación, la Figura 4.3 muestra la ubicación de los usuarios depurados dentro del DMQ y su tabla de atributos muestra el resultado del emparejamiento de datos técnicos y geográficos.

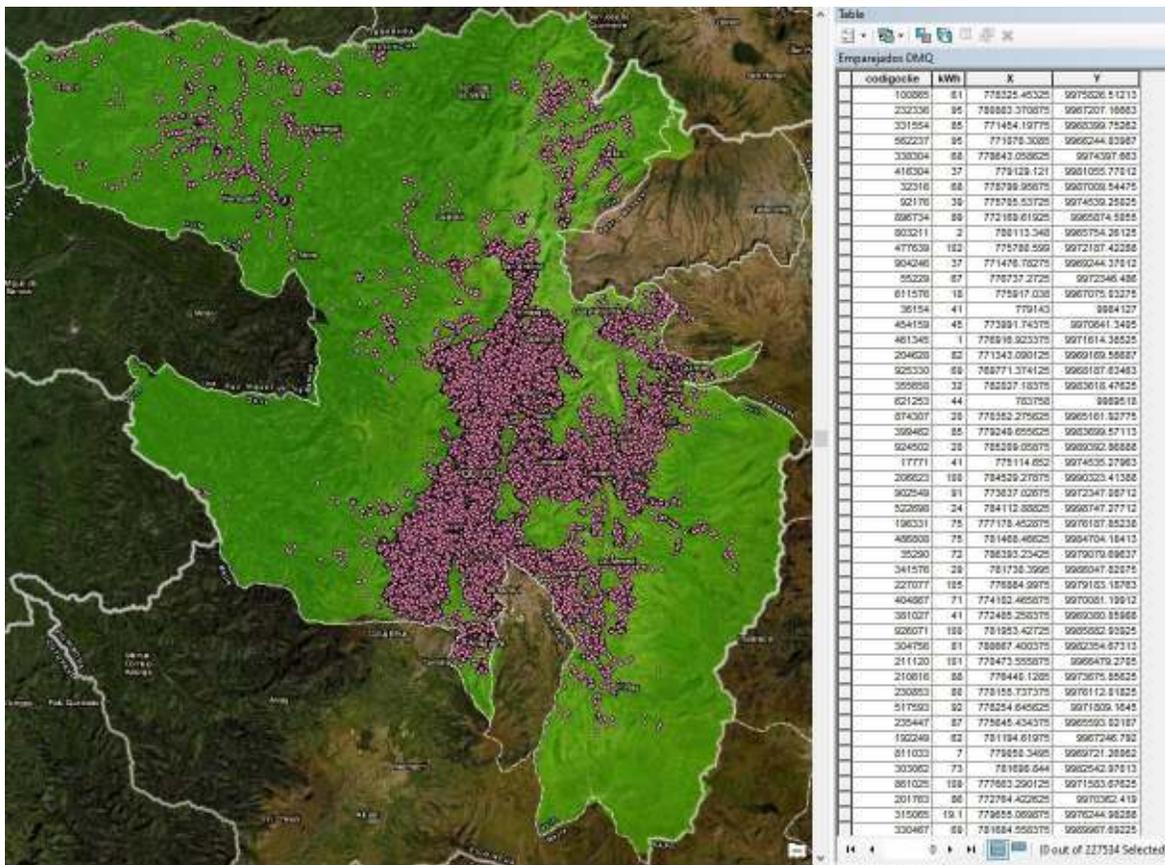


Figura 4.3. Usuarios del DMQ con su tabla de datos técnicos y geográficos emparejados.

4.1.2 Datos Socioeconómicos

4.1.2.1 Datos de Valoración del Suelo

El primer dato socioeconómico definido para la focalización del subsidio es el valor del suelo por metro cuadrado. Este dato se lo encuentra en el mapa de áreas de valoración o áreas de intervención valorativas (AIVAs) del DMQ, dicho mapa está disponible en los datos abiertos del DMQ en [28]. A continuación, la Figura 4.4 y la Figura 4.5 muestran un mapa térmico según la valoración del suelo en el DMQ, las gamas de color rojo representan los sectores donde el costo del suelo es más alto y las gamas de color verde representan los sectores donde el costo del suelo es el menor.

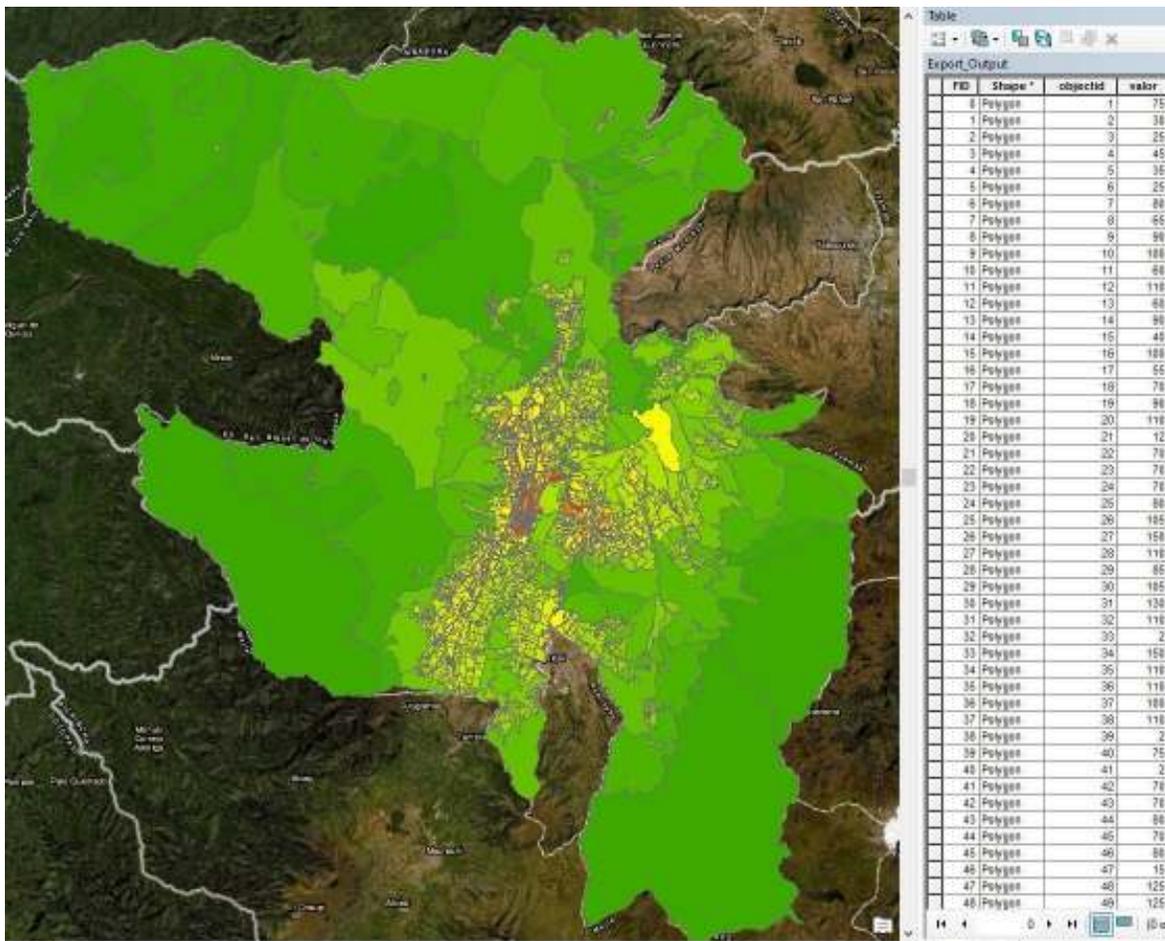


Figura 4.4. Mapa de valoración del suelo en el DMQ.

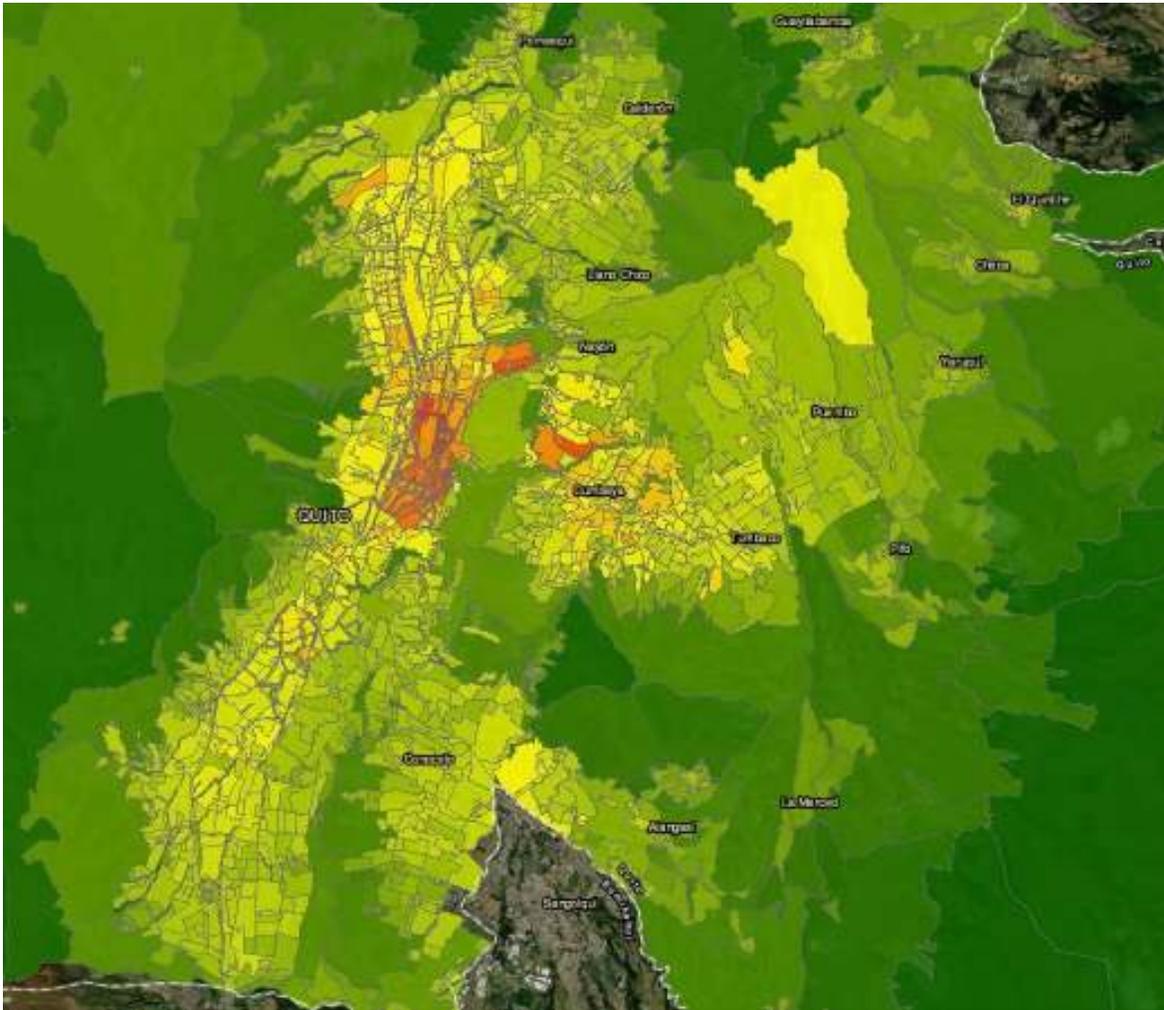


Figura 4.5. Mapa de valoración del suelo en el DMQ ampliado en la zona urbana.

4.1.2.2 Datos de Ingresos Económicos por Núcleo Familiar

Los datos de ingresos económicos obtenidos del INEC se pueden manipular mediante el uso del software SPSS de IBM. A continuación, la Figura 4.6 y la Figura 4.7 muestran la estructura de los datos que resultaron de la encuesta nacional de ingresos en el Ecuador.

05 ENIGHUR11_PERCEP TORNA_PARTEA.sav [Conjunto Datos] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda

| | Nombre | Tipo | Anchura | Decimales | Etiqueta | Valores | Perdidos | Columnas | Alineación | Medida | Rol |
|----|---------------|----------|---------|-----------|----------------------|-----------------|----------|----------|------------|---------|---------|
| 1 | Identif_2010 | Cadena | 12 | 0 | Identificador del... | Ninguno | Ninguno | 14 | Izquierda | Nominal | Entrada |
| 2 | Identif_hog | Cadena | 15 | 0 | Identificador del... | Ninguno | Ninguno | 17 | Izquierda | Nominal | Entrada |
| 3 | Identif_perna | Cadena | 17 | 0 | Identificador del... | Ninguno | Ninguno | 19 | Izquierda | Nominal | Entrada |
| 4 | Faxp_cen20... | Numérico | 17 | 13 | Factor de expa... | Ninguno | Ninguno | 17 | Derecha | Escala | Entrada |
| 5 | Regional | Numérico | 1 | 0 | Regional | {1, Norte} | Ninguno | 10 | Derecha | Nominal | Entrada |
| 6 | Área | Numérico | 1 | 0 | Área | {1, Urbana} | Ninguno | 6 | Derecha | Nominal | Entrada |
| 7 | Provincia | Numérico | 2 | 0 | Provincia | {1, Azuay} | Ninguno | 10 | Derecha | Escala | Entrada |
| 8 | Ciudad | Numérico | 6 | 0 | Ciudad | Ninguno | Ninguno | 8 | Derecha | Escala | Entrada |
| 9 | Ciudad_Auto | Numérico | 1 | 0 | Ciudad autopr... | {1, Quito} | Ninguno | 13 | Derecha | Escala | Entrada |
| 10 | Zona | Numérico | 3 | 0 | Zona | Ninguno | Ninguno | 6 | Derecha | Escala | Entrada |
| 11 | Sector | Numérico | 3 | 0 | Sector | Ninguno | Ninguno | 8 | Derecha | Escala | Entrada |
| 12 | Vivienda | Numérico | 2 | 0 | Vivienda | Ninguno | Ninguno | 10 | Derecha | Nominal | Entrada |
| 13 | Hogar | Numérico | 1 | 0 | Hogar | Ninguno | Ninguno | 7 | Derecha | Nominal | Entrada |
| 14 | Período | Numérico | 2 | 0 | Período de la M... | Ninguno | Ninguno | 9 | Derecha | Nominal | Entrada |
| 15 | Semestre | Numérico | 1 | 0 | Semestre | {1, 1} | Ninguno | 10 | Derecha | Escala | Entrada |
| 16 | Trimestre | Numérico | 1 | 0 | Trimestre | {1, 1} | Ninguno | 11 | Derecha | Escala | Entrada |
| 17 | Sem_Sec | Numérico | 2 | 0 | Semana secua... | {1, 1} | Ninguno | 10 | Derecha | Escala | Entrada |
| 18 | Domínio | Numérico | 2 | 0 | Domínio | {1, Rosta A... | Ninguno | 10 | Derecha | Escala | Entrada |
| 19 | Semana | Numérico | 1 | 0 | Semana | Ninguno | Ninguno | 8 | Derecha | Nominal | Entrada |
| 20 | Perna | Numérico | 8 | 2 | Perceptor No A... | Ninguno | Ninguno | 10 | Derecha | Escala | Entrada |
| 21 | Ocupa | Numérico | 8 | 2 | OCUPACION | {1,00, Ocup... | Ninguno | 10 | Derecha | Escala | Entrada |
| 22 | PERA00 | Numérico | 2 | 0 | PERA00. Nro | {1, 1401001} | Ninguno | 8 | Derecha | Escala | Entrada |
| 23 | PERA01 | Numérico | 7 | 0 | PERA01. Código | {1401001, 1... | Ninguno | 9 | Derecha | Escala | Entrada |
| 24 | PERA02 | Cadena | 32 | 0 | PERA02. Desc | Ninguno | Ninguno | 26 | Izquierda | Nominal | Entrada |
| 25 | PERA03 | Numérico | 9 | 2 | PERA03. Valor | {-1.00, 00 ...} | Ninguno | 11 | Derecha | Escala | Entrada |
| 26 | PERA04 | Numérico | 1 | 0 | PERA04. Frecu... | {1, Diario} | Ninguno | 8 | Derecha | Nominal | Entrada |
| 27 | PERA05 | Numérico | 9 | 2 | PERA05. Valor | {-1.00, 00 ...} | Ninguno | 11 | Derecha | Escala | Entrada |

Figura 4.6. Lista de variables resultado de la encuesta nacional de ingresos en el Ecuador.

05 ENIGHUR11_PERCEP TORNA_PARTEA.sav [Conjunto Datos] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda

| | Identif_2010 | Identif_hog | Identif_perna | Faxp_cen2010 | Regional | Área | Provincia | Ciudad | Ciudad_Auto | Zona | Sector | Vivienda | Hogar | Período | PERA00 |
|----|--------------|---------------|------------------|------------------|----------|------|-----------|--------|-------------|------|--------|----------|-------|---------|--------|
| 1 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 2 | 01010001003 | 0101000100311 | 0101000180021103 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 3 | 01010001003 | 0101000100311 | 0101000180021103 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 4 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 5 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 6 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 7 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 8 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 9 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 10 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 11 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 12 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 13 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 14 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 15 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 16 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 17 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 18 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 19 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 20 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 21 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 22 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 23 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 24 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 25 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 26 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 27 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 28 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 29 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 30 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 31 | 01010001002 | 0101000100211 | 0101000180021102 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 32 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 33 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 34 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 35 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 36 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 37 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 38 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |
| 39 | 01010001002 | 0101000100221 | 0101000180022101 | 05.3874666679432 | 4 | 1 | 1 | 1 | 10100 | 3 | 1 | 2 | 1 | 1 | 4 |

Figura 4.7. Lista de los primeros datos resultado de la encuesta nacional de ingresos en el Ecuador.



Figura 4.9. Mapa de la división político-administrativa del DMQ en zonas.

Finalmente, siguiendo el proceso descrito en el capítulo 3, se emparejan los datos de ingresos económicos promedio por sector con su ubicación en el mapa. A continuación, la Figura 4.10 muestra el resultado del emparejamiento de datos socioeconómicos y geográficos en un mapa con escala térmica sobre los ingresos económicos, las gamas de color rojo representan las zonas con mayores ingresos económicos y las gamas de color verde representan las zonas con menores ingresos.

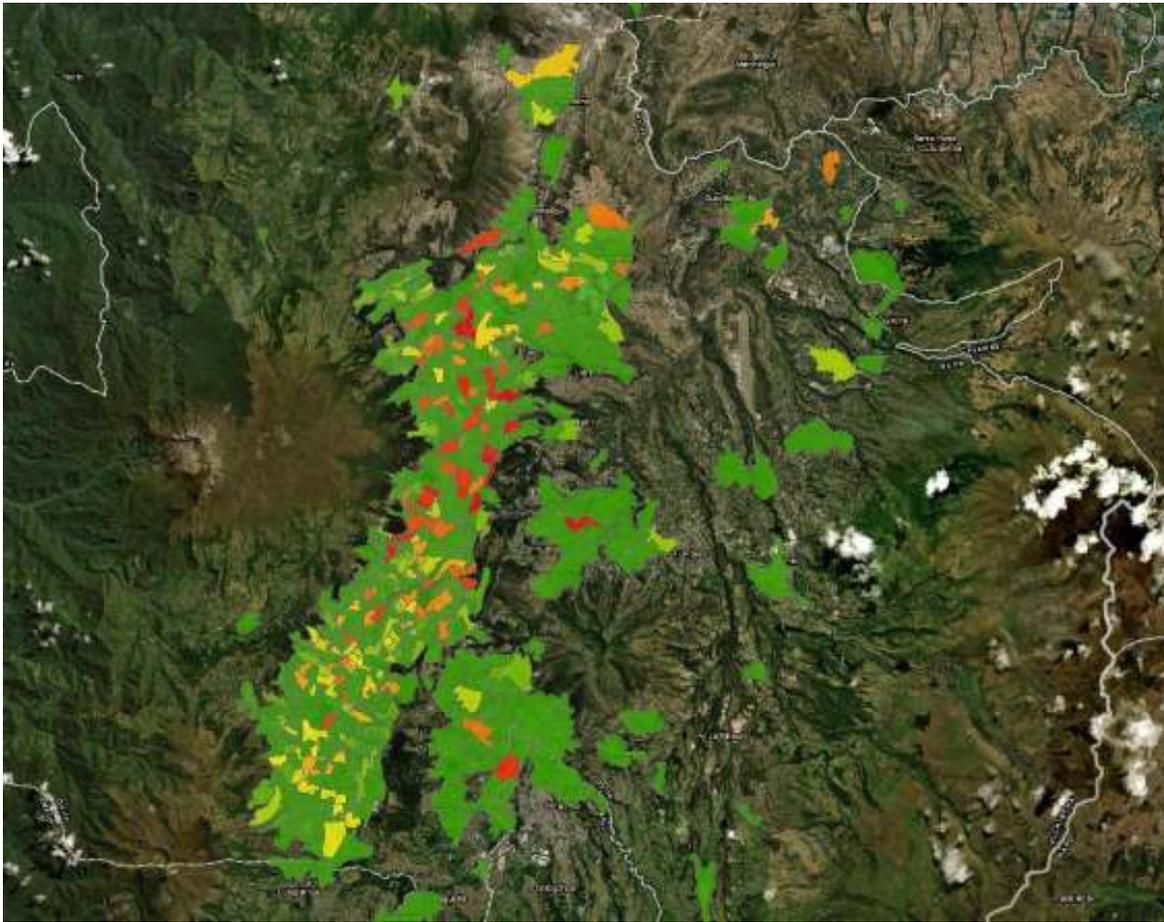


Figura 4.10. Mapa de la división político-administrativa del DMQ con escala térmica sobre los ingresos.

4.1.3 Consolidación de Datos Técnicos, Geográficos y Socioeconómicos

La fase de consolidación de datos está compuesta por el emparejamiento final de los datos técnicos, geográficos y socioeconómicos en una sola base de datos que contenga dicha información usuario por usuario. Una vez que se ha cumplido lo indicado en el capítulo 3 se logra obtener una base de datos que contiene la ubicación de cada usuario, la valoración del suelo donde está ubicado y los ingresos promedio que se perciben en su zona de ubicación. A continuación, se presentan los resultados de la consolidación de datos; la Figura 4.11 muestra, por ejemplo, como se puede representar a los usuarios según la valoración del suelo en el que residen y, la Figura 4.12 muestra, por ejemplo, como se pueden representar a los usuarios según sus ingresos promedio en la zona de residencia.



Figura 4.11. Mapa de los usuarios residenciales del DMQ con escala térmica sobre la valoración del suelo.

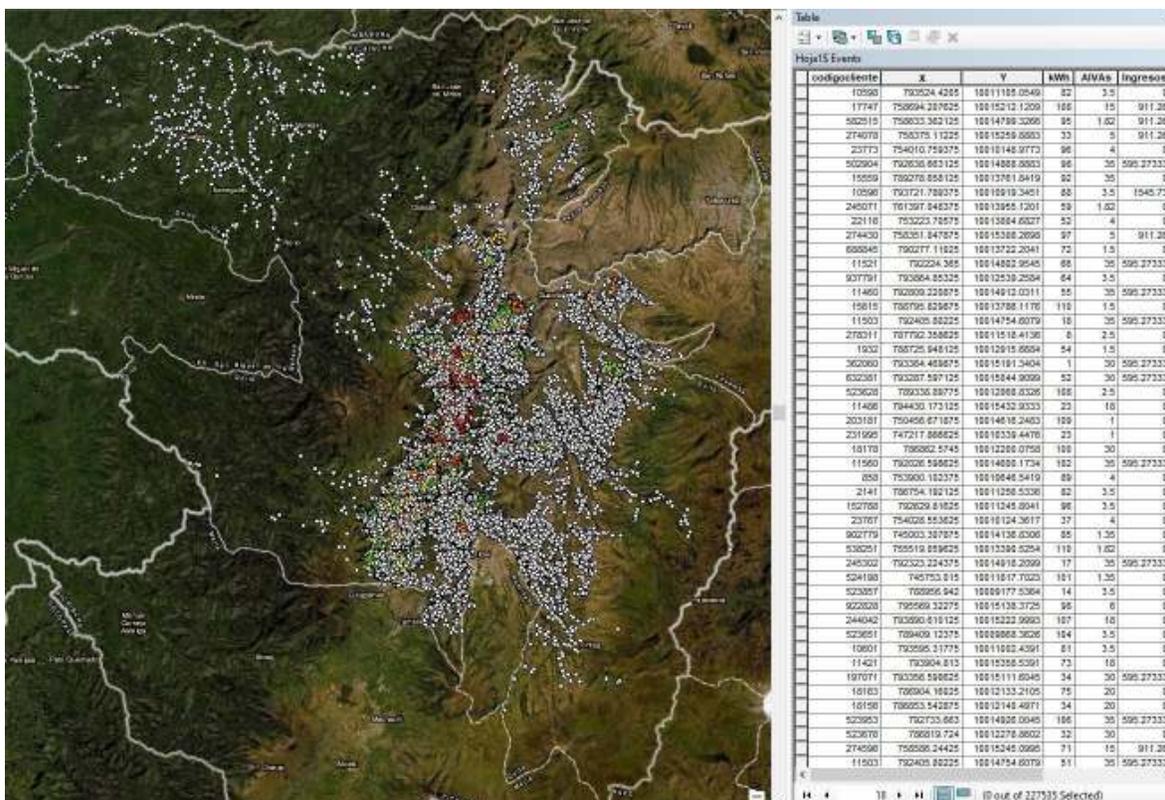


Figura 4.12. Mapa de los usuarios residenciales del DMQ con escala térmica sobre los ingresos por zona. (Los usuarios sin registro en color blanco).

4.2 Agrupación de Datos

En esta sección se presentan los resultados de la agrupación de usuarios basándose en distintos atributos de los datos.

4.2.1 Agrupación de Usuarios según su Ubicación

Una primera forma de agrupar los datos es de acuerdo con su ubicación espacial, se ha realizado esta forma de agrupación utilizando la ubicación de los usuarios residenciales del DMQ, que actualmente reciben el subsidio tarifa de la dignidad, utilizando la técnica *kmeans*. A continuación, desde la Figura 4.13 hasta la Figura 4.17 se muestran los resultados de la agrupación de usuarios para distintas cantidades de grupos.

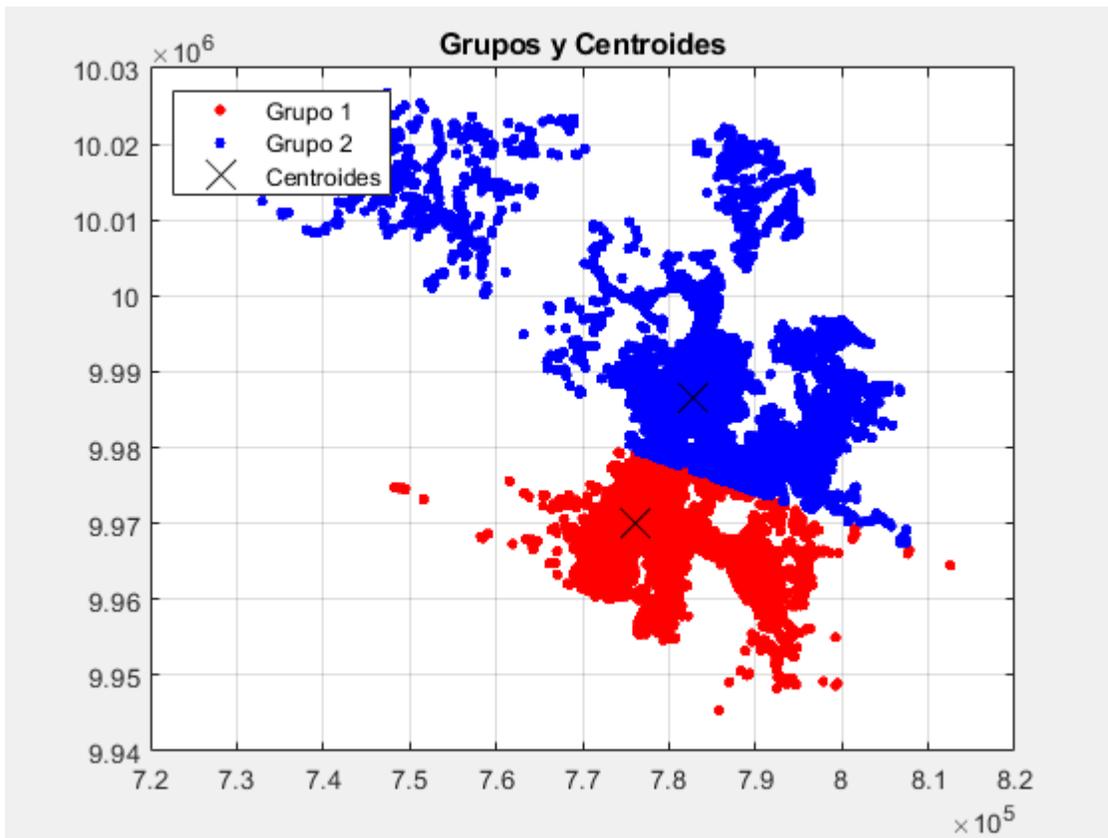


Figura 4.13. Agrupación espacial de los usuarios residenciales del DMQ en 2 grupos.

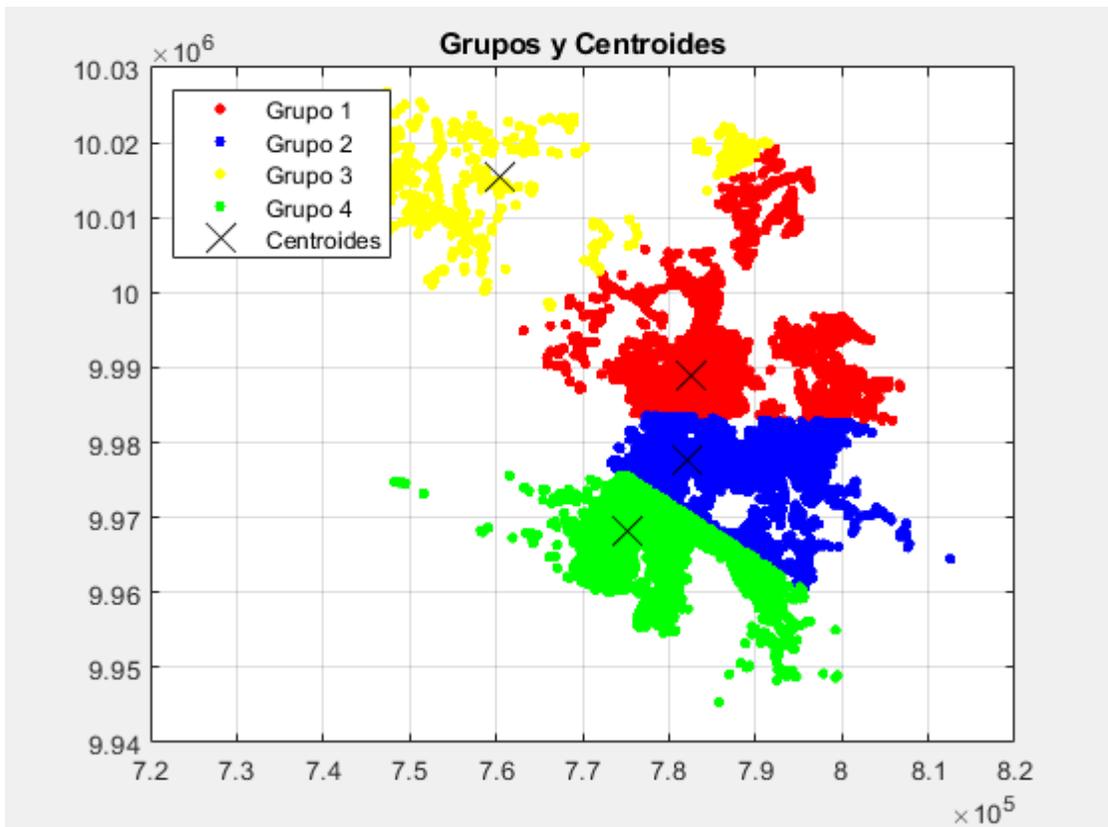


Figura 4.14. Agrupación espacial de los usuarios residenciales del DMQ en 4 grupos.

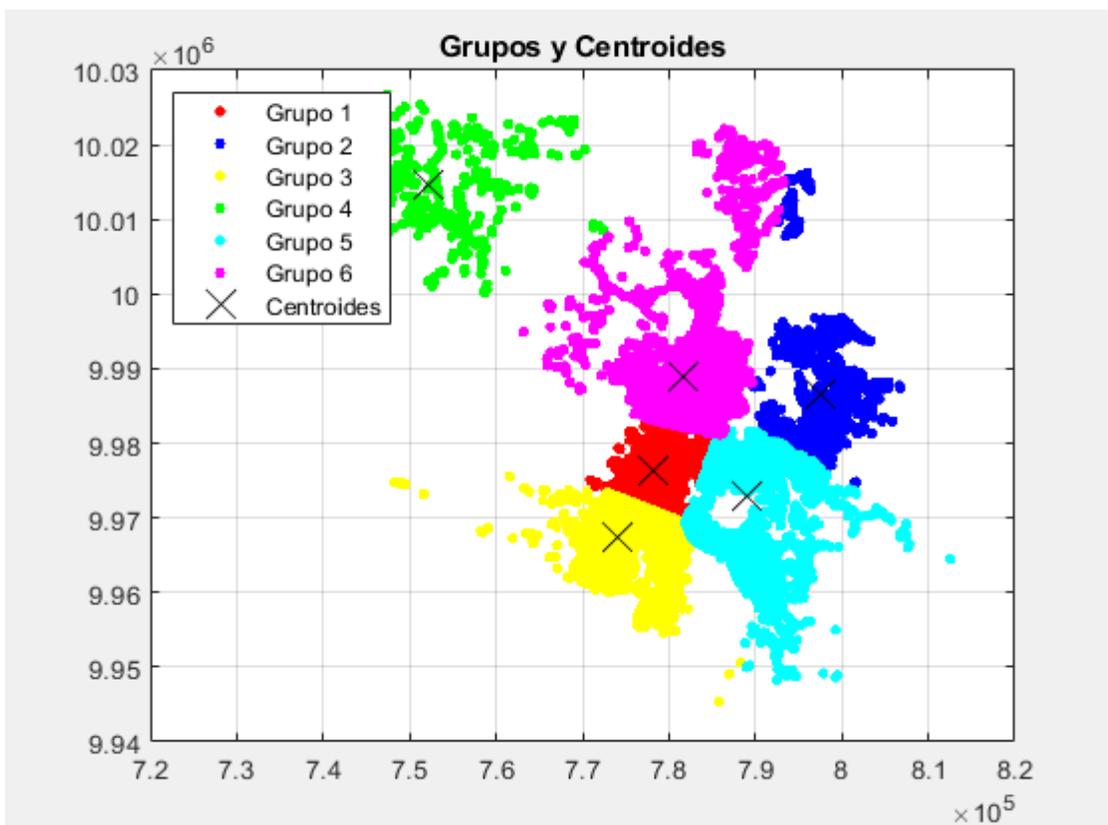


Figura 4.15. Agrupación espacial de los usuarios residenciales del DMQ en 6 grupos.

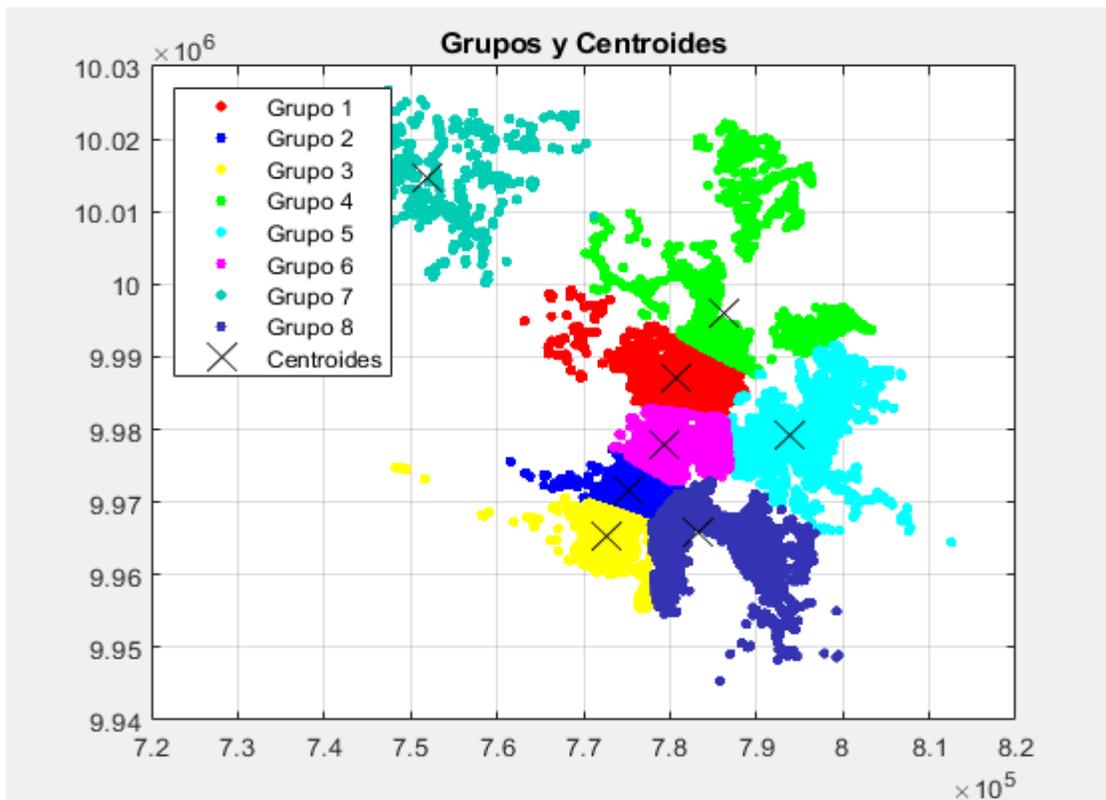


Figura 4.16. Agrupación espacial de los usuarios residenciales del DMQ en 8 grupos.

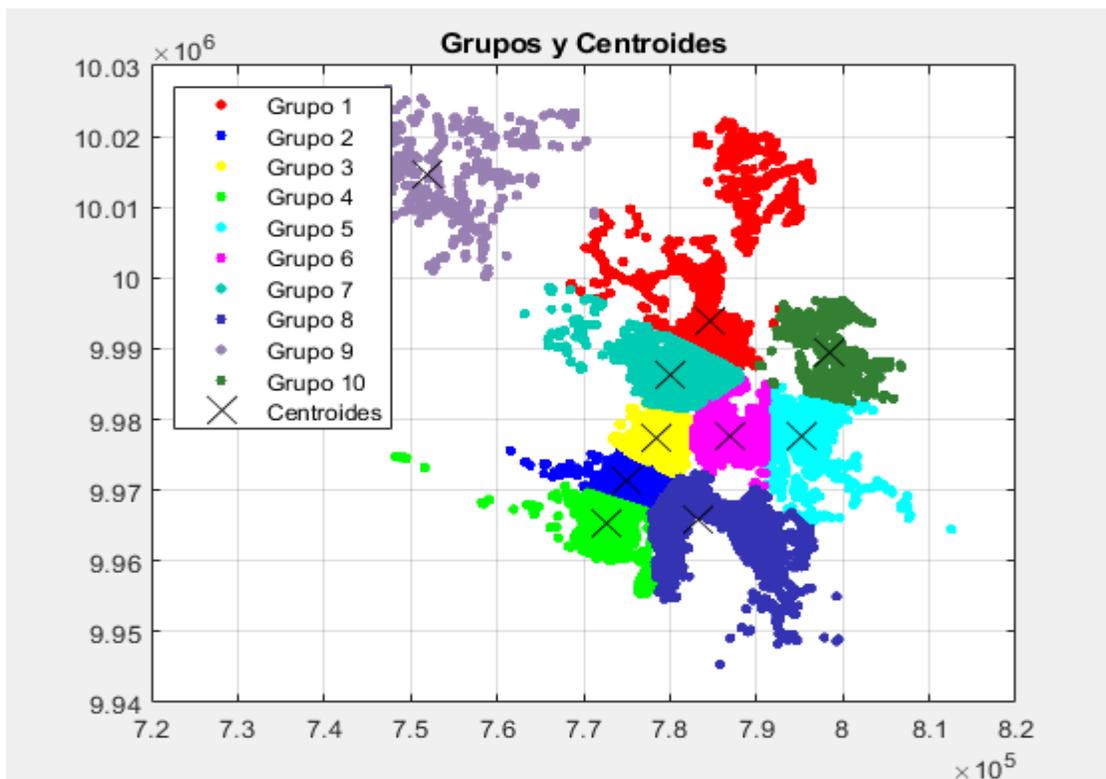


Figura 4.17. Agrupación espacial de los usuarios residenciales del DMQ en 10 grupos.

Los centroides resultantes de la agrupación se muestran en la Tabla 4.1.

Tabla 4.1 Centroides resultantes de la agrupación de usuarios

| Grupo | Centroide del grupo de consumo de energía [kWh/mes] |
|--------------|--|
| Grupo 1 | 9,43 |
| Grupo 2 | 36,58 |
| Grupo 3 | 59,43 |
| Grupo 4 | 79,78 |
| Grupo 5 | 100,03 |

Se puede observar que los 5 grupos de consumo resultantes están distribuidos en toda el área de estudio por lo que no se observa alguna relación entre los grupos de consumo de energía y los criterios socioeconómicos existentes en el DMQ.

4.2.3 Agrupación de Usuarios según la Valoración del Suelo

Una vez que se ha consolidado la base de datos técnicos, geográficos y socioeconómicos también se puede agrupar a los usuarios, que actualmente reciben el subsidio, de acuerdo con la valoración del suelo en el que residen. A continuación, la Figura 4.19 muestra, por ejemplo, el resultado de generar 5 grupos de usuarios en el DMQ según la valoración del suelo en el que residen. Los usuarios se han coloreado con escala térmica, las gamas de color verde representan a los usuarios que residen en predios de menor costo y las gamas de color rojo representan a los usuarios que residen en predios de mayor valor. Los centroides resultantes de la agrupación se muestran en la Tabla 4.2.

Tabla 4.2 Centroides resultantes de la agrupación de usuarios

| Grupo | Centroide del grupo de valoración del suelo [US\$/m²] |
|--------------|---|
| Grupo 1 | 88,81 |
| Grupo 2 | 201,12 |
| Grupo 3 | 447,96 |
| Grupo 4 | 928,23 |
| Grupo 5 | 2112,45 |

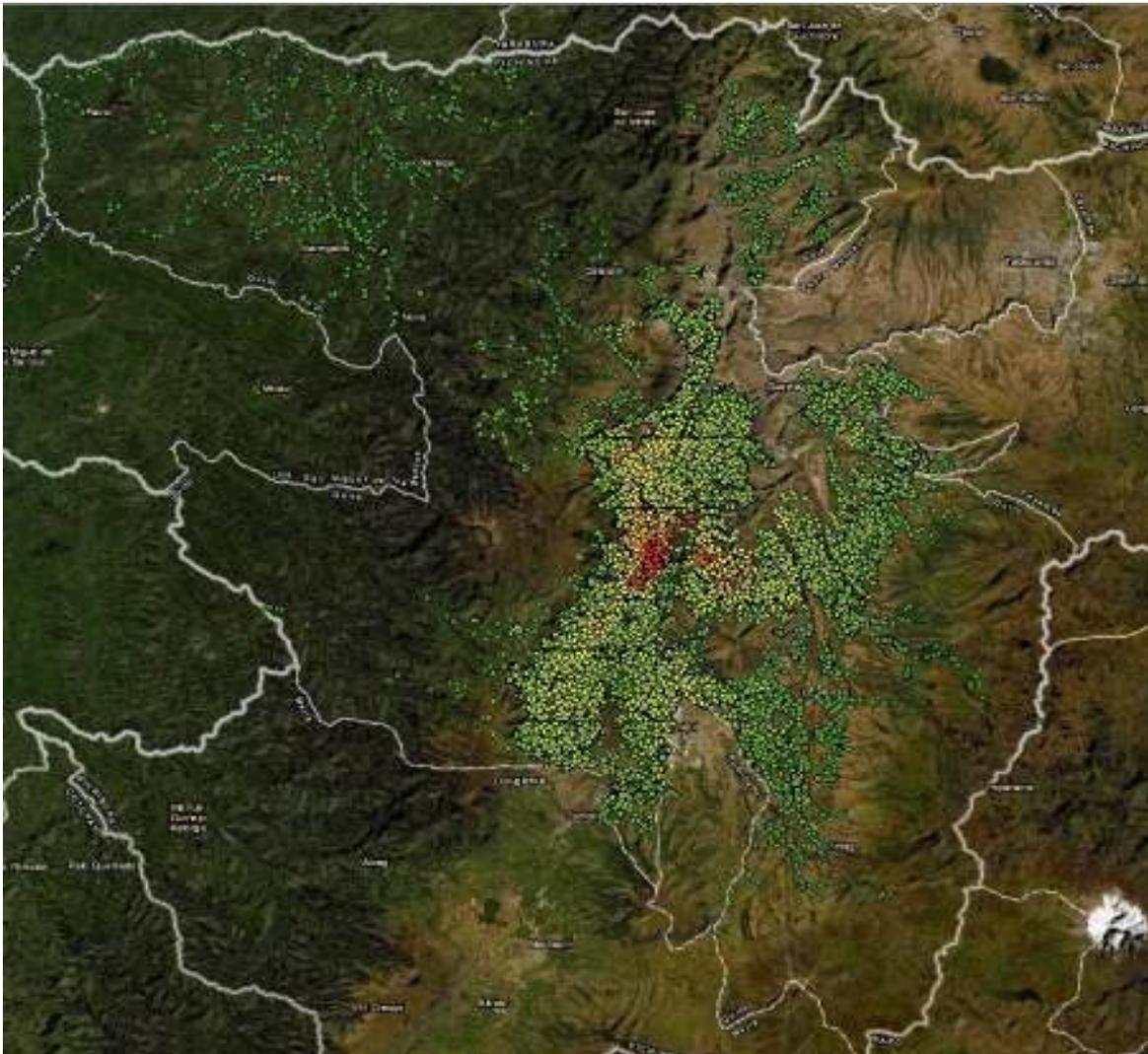


Figura 4.19. Agrupación de los usuarios residenciales del DMQ en 5 grupos según la valoración del suelo en el que residen.

Se observa, en este caso, que la agrupación de usuarios refleja una visible sectorización de usuarios en el DMQ según los valores del suelo.

4.2.4 Agrupación de Usuarios según los Ingresos Económicos de la Zona

También se pueden agrupar a los usuarios, que actualmente reciben el subsidio, de acuerdo con los ingresos económicos promedio que se perciben en la zona en la que residen. A continuación, la Figura 4.20 muestra, por ejemplo, el resultado de generar 5 grupos de usuarios en el DMQ según los ingresos económicos de la zona en la que residen. Los usuarios se han coloreado con escala térmica, las gamas de color verde representan a los usuarios que residen en predios de menor costo y las gamas de color rojo representan a los usuarios que residen en predios de mayor valor; en color blanco los usuarios que no registran valores en la encuesta de ingresos del INEC.

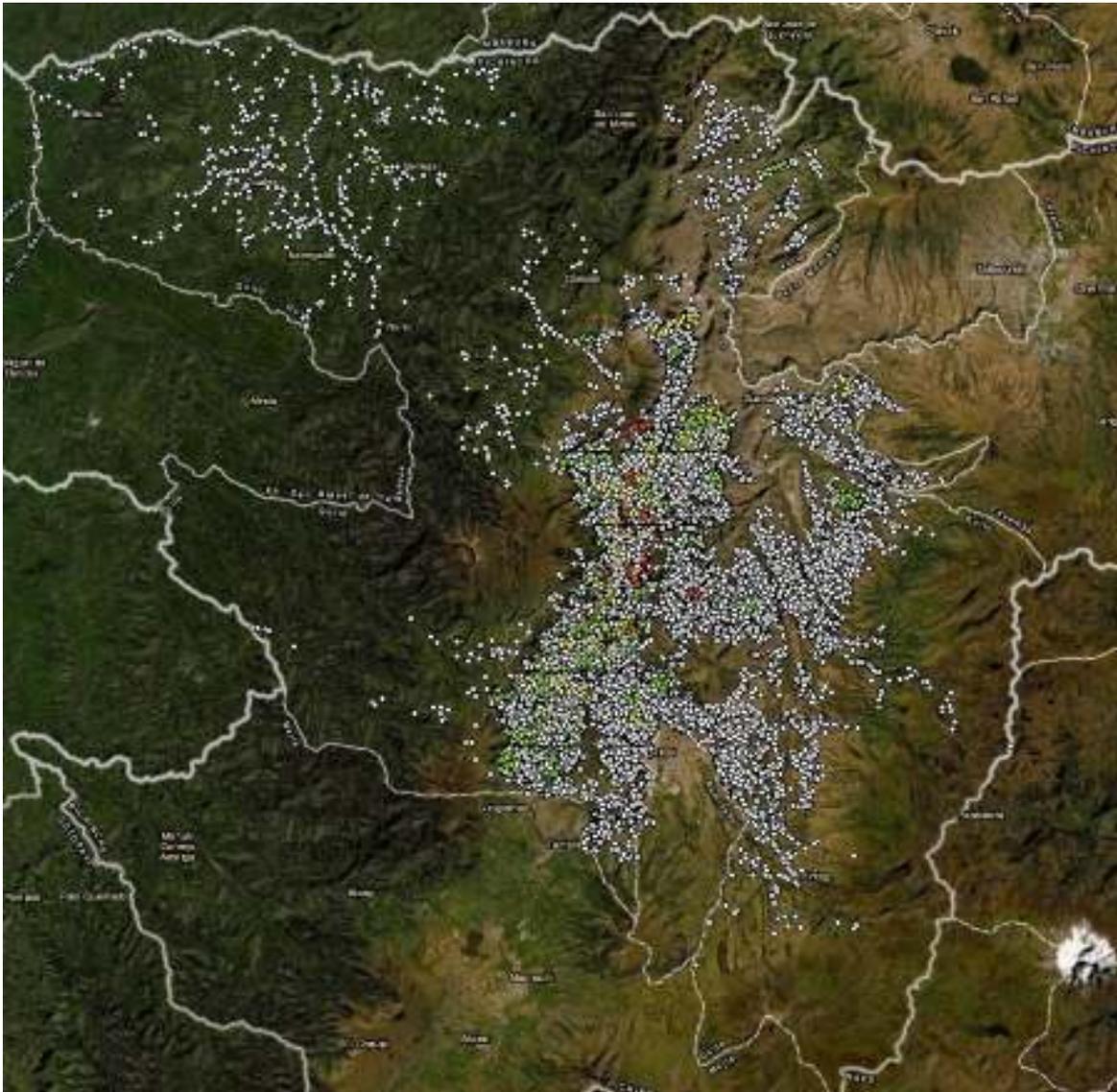


Figura 4.20. Agrupación de los usuarios residenciales del DMQ en 5 grupos según los ingresos económicos que se perciben en la zona en la que residen.

Los centroides resultantes de la agrupación se muestran en la Tabla 4.3.

Tabla 4.3 Centroides resultantes de la agrupación de usuarios

| Grupo | Centroide del grupo de ingresos económicos [US\$/mes] |
|--------------|--|
| Grupo 1 | 0,00 |
| Grupo 2 | 1.092,32 |
| Grupo 3 | 1.852,37 |
| Grupo 4 | 3.722,24 |
| Grupo 5 | 16.345,25 |

4.2.5 Agrupación de Usuarios según la Ecuación de Pobreza

De acuerdo con lo explicado en la sección 3.3. de este trabajo de titulación, para lograr considerar todas las características técnicas, socioeconómicas y geográficas de los usuarios, para la focalización del subsidio, se debe primero calcular el nivel de pobreza por usuario, 1 para usuarios pobres y 0 para usuarios no pobres; y, con dicho resultado, realizar la agrupación sobre los valores calculados de pobreza. A continuación, la Figura 4.21 muestra, por ejemplo, el resultado de formar 5 grupos de usuarios residenciales según su nivel de pobreza calculado.



Figura 4.21. Agrupación de los usuarios residenciales del DMQ en 5 grupos según su nivel de pobreza calculado.

Se observa que en el DMQ existen sectores de usuarios con altos niveles de pobreza y otros con bajos niveles de pobreza. Los centroides resultantes de la agrupación de usuarios se muestran en la Tabla 4.4.

Tabla 4.4 Centroides resultantes de la agrupación de usuarios

| Grupo | Centroide del grupo según su nivel de pobreza |
|--------------|--|
| Grupo 1 | 0,53 |
| Grupo 2 | 0,77 |
| Grupo 3 | 0,85 |
| Grupo 4 | 0,89 |
| Grupo 5 | 0,95 |

4.3 Evaluación de Grupos

Una vez que se ha definido que la agrupación de usuarios, según su nivel de pobreza, es la opción que permite considerar las características técnicas, geográficas y socioeconómicas de cada usuario se debe definir la mejor cantidad de grupos que se pueden generar en el universo de usuarios. Para esta actividad se ha realizado una evaluación utilizando el índice de Davies-Bouldin para distintas cantidades de grupos, al menos entre 2 y 20 grupos. A continuación, la Figura 4.22 muestra la variación del índice de Davies-Bouldin para los distintos números de grupos. Como resultado de la evaluación se puede observar que el índice va reduciendo de valor a medida que se forman más grupos y aparece un óptimo local al minimizarse el índice de Davies-Bouldin a 0,50 cuando se forman 13 grupos.

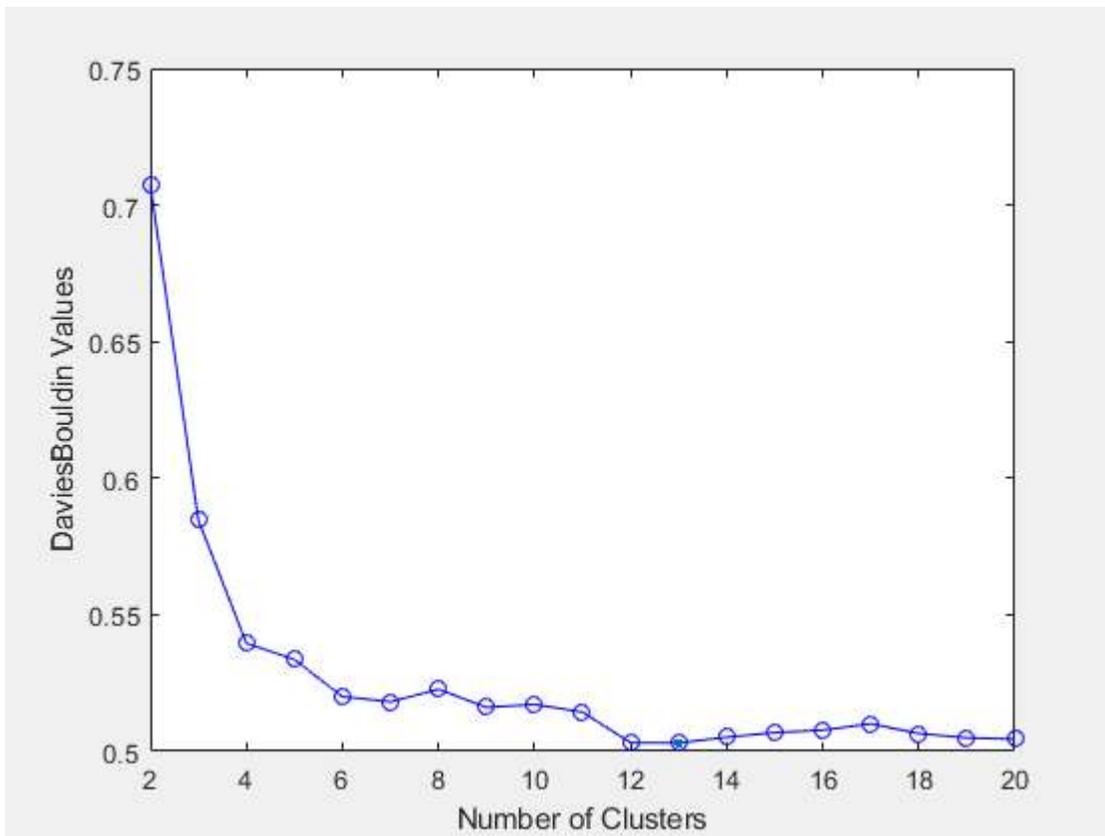


Figura 4.22. Índice de Davies-Bouldin para diferentes números de grupos.

4.4 Asignación del Subsidio

Una vez que se ha observado que, para la focalización del subsidio, la mejor forma de agrupar a los usuarios debe realizarse sobre el nivel de pobreza definido en este trabajo de titulación y, una vez que se ha determinado que la mejor cantidad de grupos para agrupar a los usuarios es 13 grupos, se ha vuelto a agrupar a los usuarios del DMQ en 13 grupos y los resultados se muestran en la Figura 4.23, en dicha figura cada color indica un diferente grupo de usuarios. Los centroides de cada grupo y la cantidad de usuarios por grupo se muestran en la Tabla 4.5.

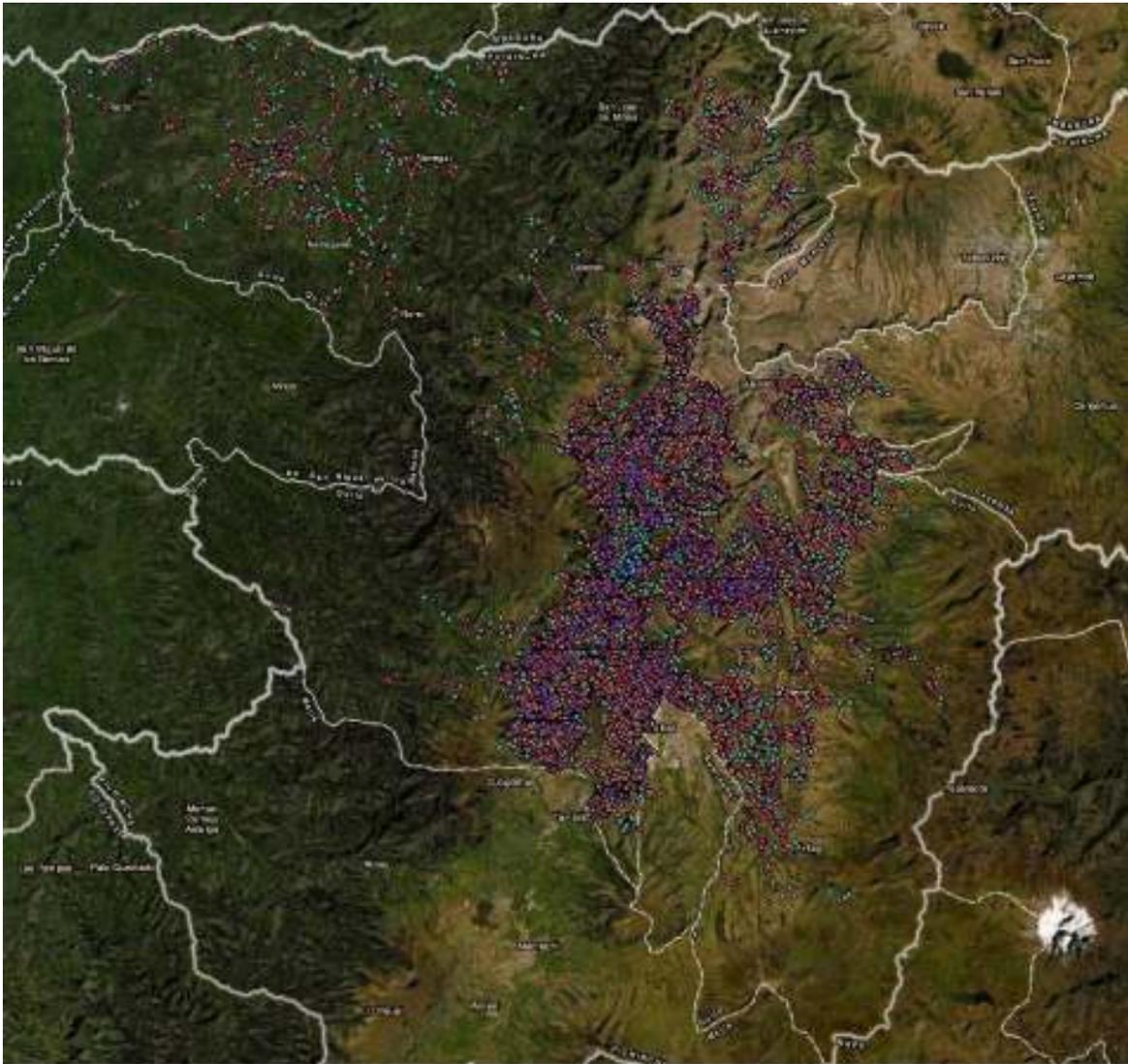


Figura 4.23. Agrupación de usuarios del DMQ, en 13 grupos, según su nivel de pobreza.

Tabla 4.5 Centroides y cantidad de usuarios por grupo

| Grupo | Centroide del grupo según su nivel de pobreza | Usuarios por Grupo |
|--------------|--|---------------------------|
| Grupo 1 | 0,39 | 405 |
| Grupo 2 | 0,57 | 1.188 |
| Grupo 3 | 0,70 | 2.297 |
| Grupo 4 | 0,76 | 4.071 |
| Grupo 5 | 0,80 | 9.201 |
| Grupo 6 | 0,84 | 16.912 |
| Grupo 7 | 0,86 | 27.409 |

| | | |
|----------|------|--------|
| Grupo 8 | 0,88 | 36.761 |
| Grupo 9 | 0,90 | 36.237 |
| Grupo 10 | 0,91 | 31.666 |
| Grupo 11 | 0,93 | 26.900 |
| Grupo 12 | 0,95 | 21.134 |
| Grupo 13 | 0,97 | 13.353 |

El último paso que queda por realizar para la focalización del subsidio es, según el criterio más restrictivo, filtrar el 40% de usuarios (91.014 usuarios) que se encuentren en los grupos con mayor nivel de pobreza. Según los resultados de la agrupación el 40% de usuarios más pobres se encuentran dentro de los grupos 10, 11, 12 y 13; dado que los usuarios de un grupo no se pueden separar para cumplir exactamente con el 40% de usuarios se ha incluido al grupo 10 con todos sus usuarios dentro de la lista de beneficiarios, dando un total de 93.053 usuarios que sí recibirían el subsidio. Para visualizar los resultados de la focalización se presenta, en la Figura 4.24, los usuarios que no recibirán el subsidio en color rojo y los que sí seguirán recibiendo el subsidio en color verde.

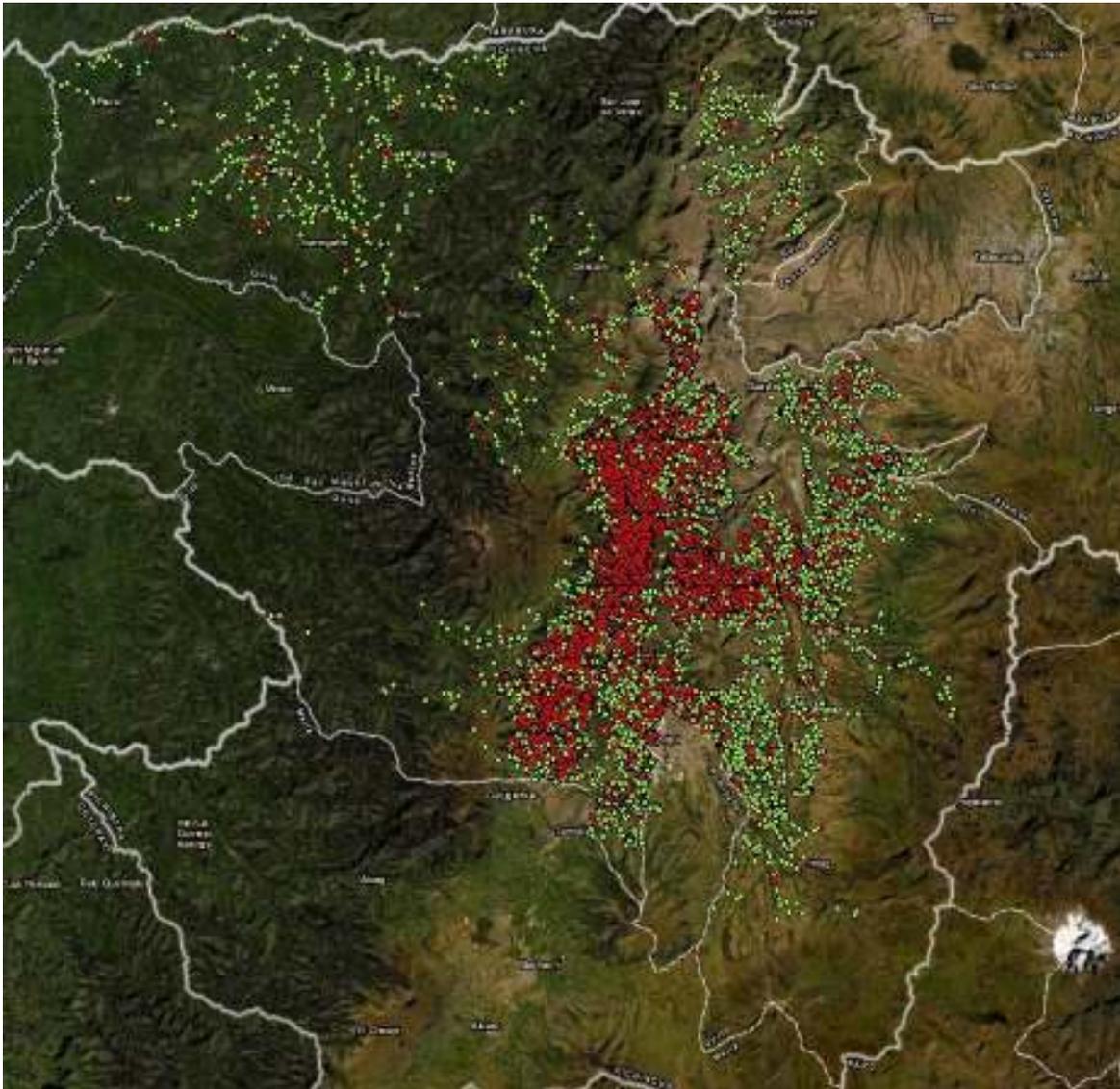


Figura 4.24. Mapa de focalización del subsidio en los usuarios residenciales del DMQ. Opción 1.

Si de otra manera se considera el criterio menos restrictivo, es decir, el de retirar de la lista de beneficiarios del subsidio al 40% de usuarios menos pobres, se obtiene que 129.290 usuarios continuarían recibiendo el subsidio. Para esta consideración la Figura 4.25 muestra con color verde a los usuarios que sí deberían recibir el subsidio y con color rojo a aquellos usuarios que no deberían recibir el subsidio.

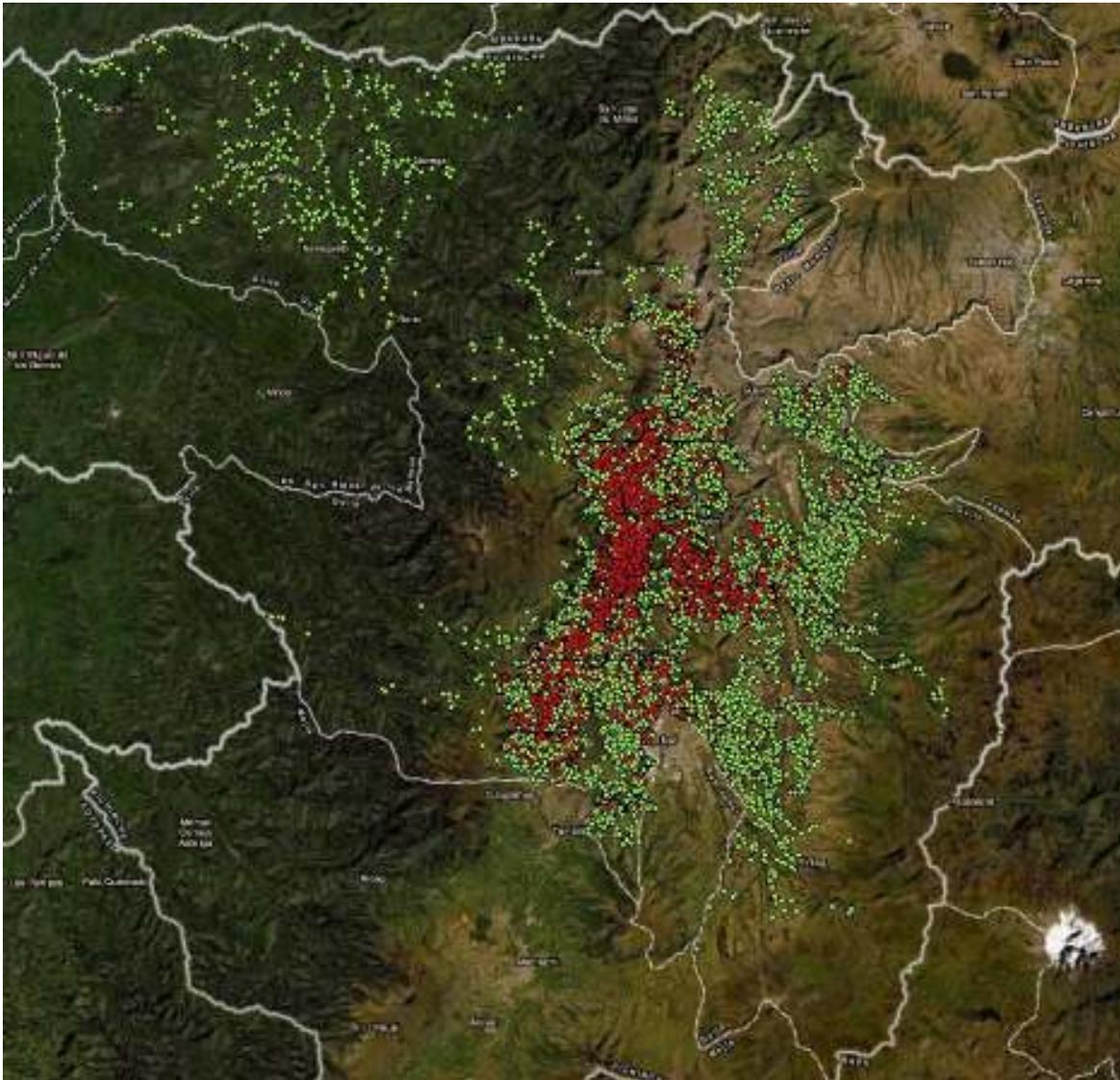


Figura 4.25. Mapa de focalización del subsidio en los usuarios residenciales del DMQ.
Opción 2.

Se verifica que, cualquiera que sea el criterio para definir el porcentaje de focalización del subsidio, la metodología de focalización propuesta permite definir qué usuarios sí deberían recibir el subsidio y qué usuarios no deberían recibirlo según sus características técnicas, geográficas y socioeconómicas.

En el Anexo 1 se presentan los resultados de la asignación del subsidio considerando el criterio más restrictivo y en el Anexo 2 se presentan los resultados de la asignación del subsidio considerando el criterio menos restrictivo; en ambos casos se presentan imágenes ampliadas para algunos sectores del DMQ

4.5 Valoración Económica

La valoración económica del monto que destina el Estado para el subsidio tarifa de la dignidad se lo realiza calculando la diferencia entre el pago que deberían realizar los usuarios residenciales considerando el pliego tarifario actual sin subsidio, disponible en [25], menos el monto que pagan debido a lo establecido en la definición del subsidio.

Sin la focalización del subsidio, para el caso de aplicación en el DMQ, existen 227.534 usuarios que reciben el subsidio tarifa de la dignidad. Para dichos usuarios se calcula la sumatoria de sus consumos de energía mensuales y luego se valora dicha energía total considerando el pliego tarifario al que se regirían y considerando la valoración definida en el subsidio, la diferencia entre estos dos últimos valores corresponde al valor subsidiado por el Estado. A continuación, la Tabla 4.6 muestra los resultados de dicha valoración.

Tabla 4.6 Valoración de la aplicación del subsidio tarifa del a dignidad en el DMQ. Sin la aplicación de la focalización.

| Valoración Económica | Usuarios Beneficiarios | Energía Total [kWh/mes] | Costo de la Energía [US\$/kWh] | Costo de Comercialización [US\$/mes] | Monto Facturado Total [US\$/mes] (Energía * Costo de la Energía + Usuarios*Costo de Comercialización) |
|---|------------------------|-------------------------|--------------------------------|--------------------------------------|--|
| Valoración del Subsidio según el Pliego Tarifario | 227.534 | 13'831.992,65 | 0,083 | 1,414 | 1'469.788,47 |
| Valoración del Subsidio según el Subsidio Tarifa de la Dignidad | 227.534 | 13'831.992,65 | 0,04 | 0,70 | 712.553,51 |
| Monto Subsidiado [US\$/mes] | | | | | 757.234,96 |

Nota: Los valores presentados corresponden al caso de estudio enero 2018.

A continuación, la Tabla 4.7 muestra los resultados de la valoración del subsidio considerando la metodología de focalización en la opción más restrictiva. Para este caso existirían 93.053 usuarios beneficiarios.

Tabla 4.7 Valoración de la aplicación del subsidio tarifa de la dignidad en el DMQ. Con la aplicación de la focalización. Opción 1.

| Valoración Económica | Usuarios Beneficiarios | Energía Total [kWh/mes] | Costo de la Energía [US\$/kWh] | Costo de Comercialización [US\$/mes] | Monto Facturado Total [US\$/mes] (Energía * Costo de la Energía + Usuarios*Costo de Comercialización) |
|---|------------------------|-------------------------|--------------------------------|--------------------------------------|--|
| Valoración del Subsidio según el Pliego Tarifario | 93.053 | 3'232.414,49 | 0,083 | 1,414 | 268.421,97 |
| Valoración del Subsidio según el Subsidio Tarifa de la Dignidad | 93.053 | 3'232.414,49 | 0,04 | 0,70 | 129.361,72 |
| Monto Subsidiado [US\$/mes] | | | | | 139.060,26 |

Nota: Los valores presentados corresponden al caso de estudio enero 2018.

A continuación, la Tabla 4.8 muestra los resultados de la valoración del subsidio considerando la metodología de focalización en la opción menos restrictiva. Para este caso existirían 129.290 usuarios beneficiarios.

Tabla 4.8 Valoración de la aplicación del subsidio tarifa de la dignidad en el DMQ. Con la aplicación de la focalización. Opción 2.

| Valoración Económica | Usuarios Beneficiarios | Energía Total [kWh/mes] | Costo de la Energía [US\$/kWh] | Costo de Comercialización [US\$/mes] | Monto Facturado Total [US\$/mes] (Energía * Costo de la Energía + Usuarios*Costo de Comercialización) |
|---|------------------------|-------------------------|--------------------------------|--------------------------------------|---|
| Valoración del Subsidio según el Pliego Tarifario | 129.290 | 5'768.781,97 | 0,083 | 1,414 | 661.624,96 |
| Valoración del Subsidio según el Subsidio Tarifa de la Dignidad | 129.290 | 5'768.781,97 | 0,04 | 0,70 | 321.254,28 |
| Monto Subsidiado [US\$/mes] | | | | | 340.370, 68 |

Nota: Los valores presentados corresponden al caso de estudio enero 2018.

Se observa que con el criterio más restrictivo de focalizar el subsidio al 40% de usuarios más pobres se logra un ahorro de aproximadamente el 72% del valor inicialmente subsidiado sin considerar la focalización. Por otro lado, con el criterio menos restrictivo de focalizar el subsidio al 60% de usuarios más pobres (equivalente a retirar el 40% de usuarios menos pobres) se logra un ahorro de aproximadamente el 65% del valor inicialmente subsidiado sin considerar la focalización.

4.6 Propuesta Metodológica

La propuesta metodológica para focalizar el subsidio tarifa de la dignidad considerando las características técnicas, geográficas y socioeconómicas de los usuarios residenciales de las empresas eléctricas del Ecuador, utilizando la técnica de agrupación de datos, se describe a continuación.

1. Recopilar los datos técnicos, geográficos y socioeconómicos de los usuarios residenciales de las Empresas Eléctricas del Ecuador.
2. Depurar los datos obtenidos y procesarlos por separado según los criterios de la sección 3.1.
3. Emparejar los datos para que cada usuario posea como atributos su ubicación, su consumo de energía y la valoración del suelo e ingresos económicos según el sitio en el que se ubica el usuario.
4. Definir una ecuación de pobreza que, para su cálculo, considere de manera ponderada los datos técnicos, geográficos y socioeconómicos de los usuarios.
5. Definir la ponderación o peso que se le dará a cada característica técnica o socioeconómica de los usuarios de acuerdo con la calidad de los datos y de acuerdo con la relación que tengan los datos con la pobreza.
6. Calcular el nivel de pobreza para cada uno de los usuarios residenciales según sus características técnicas, geográficas y socioeconómicas.
7. Determinar la mejor cantidad de grupos que se puede generar en la agrupación de usuarios residenciales minimizando, para ello, el índice de Davies-Bouldin.
8. Agrupar a los usuarios residenciales según su nivel de pobreza y en la cantidad de grupos que haya resultado del punto 7.
9. Ordenar a los grupos resultantes del punto 8 de acuerdo con el valor de su centroide de pobreza. Los grupos con un valor de centroide más cercano a 1 representan a los grupos de usuarios más pobres y los grupos con valor de centroide más cercano a 0 representan a los grupos de usuarios menos pobres.
10. Definir un porcentaje de focalización del subsidio.
11. Determinar la cantidad de grupos de usuarios y a los usuarios que, con mayor nivel de pobreza, ingresen dentro del porcentaje de focalización deseado.
12. En la base de datos del sistema comercial de cada Empresa Eléctrica se debe generar un atributo para cada usuario residencial que indique si es beneficiario o no del subsidio tarifa de la dignidad, dicho atributo se deberá llenar con los resultados de la focalización, según el código de cada cliente, para que así se lo

considere como beneficiario o no beneficiario del subsidio en la facturación de energía eléctrica.

A continuación, la Figura 4.26 muestra un diagrama de flujo que resume la metodología propuesta para la focalización del subsidio tarifa de la dignidad.

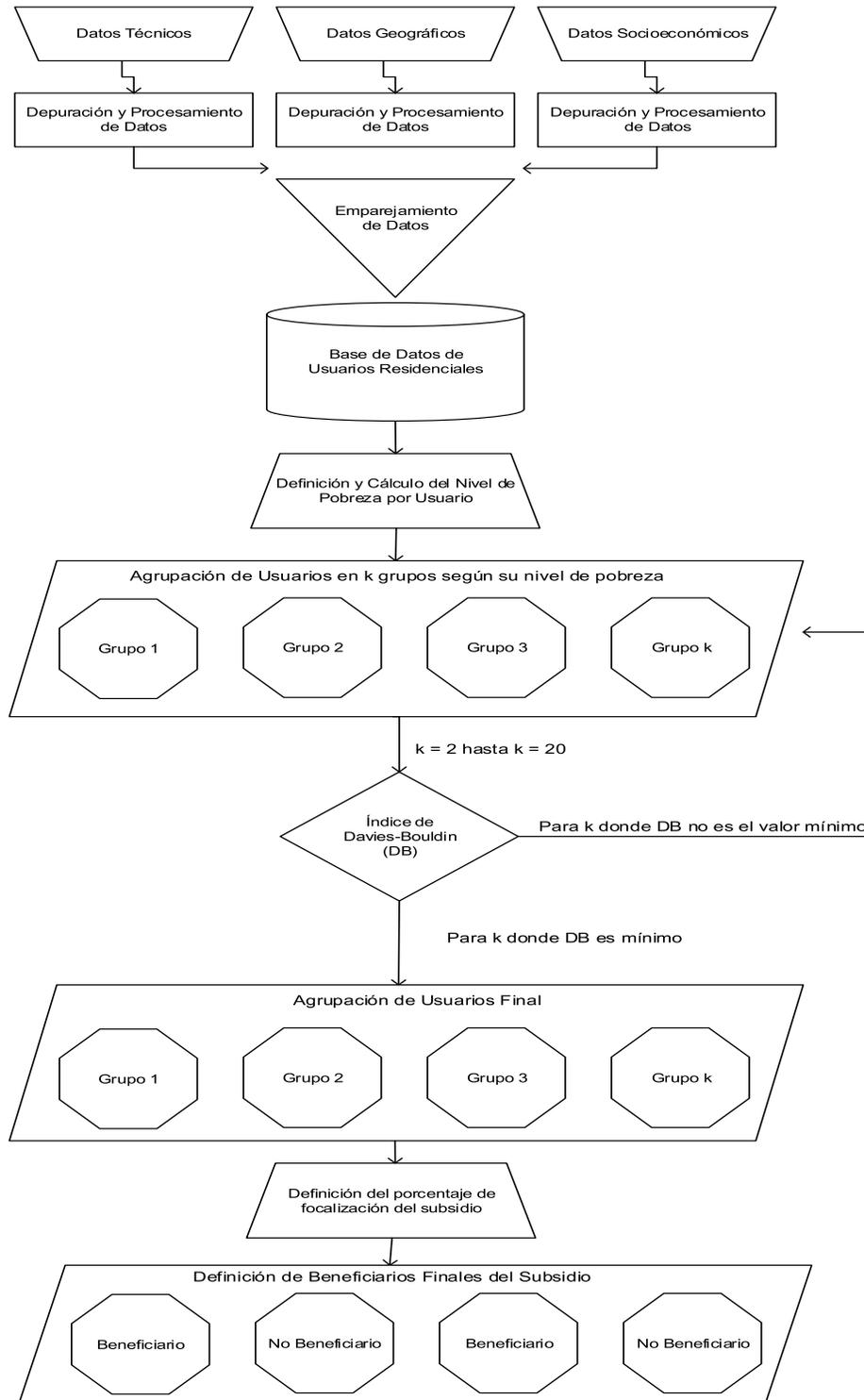


Figura 4.26. Diagrama de flujo para la focalización del subsidio tarifa de la dignidad.

5 CONCLUSIONES

- Para focalizar el subsidio tarifa de la dignidad es necesario estructurar una base de datos que permita relacionar la ubicación de los usuarios residenciales con sus características de consumo energético y con sus características socioeconómicas de valoración del suelo e ingresos económicos. Para la estructuración de la base de datos de usuarios se requiere el uso de Matlab, ArcGIS, SPSS y Excel puesto que los datos provienen en distintos tipos de archivos y, necesariamente, de distintos Organismos del Estado.
- La agrupación de usuarios, en base a su ubicación, no constituye, por sí sola, una forma de focalizar el subsidio tarifa de la dignidad debido a que los grupos y la cantidad de grupos que resultan de la agrupación no presentan una relación observable con la pobreza.
- Es posible focalizar el subsidio tarifa de la dignidad en los usuarios residenciales de las empresas eléctricas del Ecuador mediante el uso de la técnica de agrupación de datos y considerando las características técnicas, geográficas y socioeconómicas de los usuarios.
- La metodología de focalización propuesta en este trabajo de titulación sirve para focalizar el subsidio tarifa de la dignidad a cualquier porcentaje de usuarios. La metodología de focalización del subsidio propuesta permite identificar a quienes sí se debería otorgar el subsidio y a quienes no se debería hacerlo.
- El índice de Davies-Bouldin permite determinar la mejor cantidad de grupos de usuarios que se deben generar para focalizar el subsidio tarifa de la dignidad. Este índice permite encontrar la cantidad de grupos de usuarios en la que los usuarios de un grupo son más parecidos entre sí y a su vez son lo más diferentes que los usuarios de otros grupos.
- Para el caso de aplicación, correspondiente a la focalización del subsidio tarifa de la dignidad en el DMQ, con el criterio más restrictivo de focalizar el subsidio al 40% de usuarios más pobres se logra un ahorro del 72% del valor originalmente subsidiado sin considerar la focalización. Si, de otra manera, se considera el criterio menos restrictivo de focalizar el subsidio al 60% de usuarios más pobres se logra un ahorro del 65% del valor originalmente subsidiado sin considerar la focalización. El ahorro del Estado por concepto subsidio tarifa de la dignidad solo en el DMQ,

para el caso más restrictivo de focalización, sería de 618.174,70 dólares mensuales equivalentes a 7'418.086.40 dólares anuales.

6 RECOMENDACIONES

- La focalización del subsidio se debería realizar progresivamente, de manera anual, y con los mejores datos disponibles a la fecha de la focalización del subsidio. Esto permitirá que, luego de algunos años de focalización del subsidio, se elimine el subsidio tarifa de la dignidad por completo.
- Los recursos económicos que se logren ahorrar, como resultado de la focalización del subsidio tarifa de la dignidad, se deberían destinar para reforzar los programas sociales enfocados a la eliminación de la pobreza en el Ecuador.
- Los datos técnicos, geográficos y socioeconómicos deben definirse para un periodo estacional lo más parecido entre sí. Existe una gran oportunidad de focalizar el subsidio tan pronto se dispongan de los resultados del censo 2020.
- Para la ecuación de pobreza de la sección 3.3.1 se debería considerar la mayor cantidad de criterios socioeconómicos que estén relacionados con la pobreza de los usuarios puesto que esto permitirían una mejor focalización del subsidio. La ponderación de cada criterio socioeconómico adicional se podrá ajustar de acuerdo con la calidad de los datos y de acuerdo con la relación que tengan los datos con la pobreza.
- Para verificar los resultados de la focalización se debería realizar encuestas a los usuarios beneficiarios del subsidio que resulten de la focalización a fin de verificar su nivel real de pobreza.

7 REFERENCIAS BIBLIOGRÁFICAS

- [1] ARCONEL, «Agencia de Regulación y Control de Electricidad,» 08 10 2020. [En línea]. Available: <https://www.regulacioneolica.gob.ec/tarifa-dignidad/>. [Último acceso: 08 10 2020].
- [2] J. Franco y G. Aragón, «Consortio de investigación económico y social,» 20 Octubre 2020. [En línea]. Available: <http://www.cies.org.pe/sites/default/files/files/articulos/economiasociedad/02-franco-aragon.pdf>. [Último acceso: 20 Octubre 2020].
- [3] Bhagwati, Jagdish y Ramaswami, «Domestic Distortions, Tariffs, and the Theory of Optimum Subsidy,» *Journal of Political Economy*, vol. 71, nº 1, p. 7, 1963.
- [4] J. Franco y G. Aragón, «La focalización es relevante: propuesta de un esquema óptimo de subsidios al consumo eléctrico residencial e impactos sociales de su implementación,» Centro de Investigación de la Universidad del Pacífico, Lima, 2010.
- [5] J. Rawls, *A Theory of Justice*, Estados Unidos: Harvard University Press, 1971.
- [6] C. Verdú y M. Rodríguez, «Pobreza monetaria, exclusión educativa y privación material de los jóvenes,» *Revista de economía aplicada*, vol. 19, nº 56, pp. 59-88, 2011.
- [7] Banco Mundial, «Banco Mundial,» Grupo Banco Mundial, 08 10 2020. [En línea]. Available: <https://www.bancomundial.org/es/topic/poverty/overview>. [Último acceso: 08 10 2020].
- [8] Instituto Nacional de Estadísticas y Censos, «Indicadores de Pobreza y Desigualdad,» INEC, Quito, 2019.
- [9] L. Tobar, «Los subsidios en el Ecuador,» de *Análisis de Coyuntura de la Carrera de Economía de la Universidad Politécnica Salesiana Sede Cuenca*, Cuenca, 2020.

- [10] E. Jaio, J. Paredes y J. Sánchez, «EL BONO SOCIAL Y LAS COOPERATIVAS ENERGÉTICAS,» *REVESCO Universidad Complutense de Madrid*, vol. 1, nº 122, pp. 165-190, 2016.
- [11] J. Soto, «¿Universalización, participación o focalización?: la economía política de las estrategias para el alivio de la pobreza en México,» de *XXIII Congreso Internacional de la Asociación de Estudios Latinoamericanos*, Washington, D.C., 2001.
- [12] A. Sen, «The Political Economy of Targeting,» de *Annual Bank Conference on Development Economics, World Bank*, Washington, D.C., 1992.
- [13] C. Pérez y D. Santín, *Minería de Datos. Técnicas y Herramientas*, Madrid: Ediciones Paraninfo S.A., 2007.
- [14] O. Maimon y L. Rockach, *The Data Mining and Knowledge Discovery Handbook*, Israel: Springer, 2005.
- [15] C. Moler, *Numerical Computing with MATLAB*, SIAM, 2004.
- [16] D. Guevara, *PREPROCESAMIENTO DE LOS DATOS DE LAS UNIDADES DE*, Quito: Escuela Politécnica Nacional, 2019.
- [17] E. Gallardo, *Técnicas de clustering aplicadas a la resolución de problemas de optimización combinatoria con restricciones espaciales y temporales*, Leganés: Universidad Carlos III de Madrid, 2018.
- [18] J. González, *Métodos de clustering robustos*, Buenos Aires: Universidad de Buenos Aires, 2019.
- [19] L. Kaufman y P. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*, Edegem: John Wiley & Sons, 2009.
- [20] D. L. Davies y D. W. Bouldin, «A Cluster Separation Measure,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vols. %1 de %2PAMI-1, nº 2, pp. 224-227, 1979.

- [21] J. Riquelme, R. Ruiz y K. Gilbert, «Minería de Datos: Conceptos y Tendencias,» *Inteligencia Artificial: Revista Iberoamericana de Inteligencia Artificial*, vol. 10, nº 29, pp. 11-18, 2006.
- [22] ESRI, «ArcGIS Resources,» ESRI, 10 10 2020. [En línea]. Available: https://resources.arcgis.com/es/help/getting-started/articles/026n00000014000000.htm#ESRI_SECTION2_31E0F968AA8740CF8BD33F5CFAC50FAF. [Último acceso: 10 10 2020].
- [23] MATLAB, «MATLAB,» MathWorks, 10 10 2020. [En línea]. Available: <https://la.mathworks.com/products/matlab.html>. [Último acceso: 10 10 2020].
- [24] ARCONEL, «MANUAL PARA LA RECOPIACIÓN DE LA INFORMACIÓN DEL SECTOR ELÉCTRICO A TRAVÉS DEL SISTEMA SISDAT,» ARCONEL, 01 04 2018. [En línea]. Available: <https://www.regulacionelectrica.gob.ec/wp-content/uploads/downloads/2019/06/Anexo-5.-Manual-SISDAT-Participantes.pdf>. [Último acceso: 10 10 2020].
- [25] ARCONEL, «Pliego Tarifario para las Empresas Eléctricas de Distribución 2020,» ARCONEL, Quito, 2020.
- [26] Instituto Nacional de Estadísticas y Censos (INEC), «INEC - Información Cartográfica,» INEC, 01 12 2010. [En línea]. Available: <https://www.ecuadorencifras.gob.ec/category/cartografia-2/>. [Último acceso: 10 10 2020].
- [27] Instituto Nacional de Estadísticas y Censos (INEC), «Encuesta Nacional de Ingresos y Gastos de los Hogares Urbanos y Rurales – Bases de Datos,» INEC, 01 01 2011 - 2012. [En línea]. Available: <https://www.ecuadorencifras.gob.ec//encuesta-nacional-de-ingresos-y-gastos-de-los-hogares-urbanos-y-rurales-bases-de-datos/>. [Último acceso: 10 10 2020].
- [28] Gobierno Abierto Quito, «Gobierno Abierto Quito,» Municipio Metropolitano de Quito, 10 10 2020. [En línea]. Available: http://gobiernoabierto.quito.gob.ec/?page_id=1122. [Último acceso: 10 10 2020].

8 ANEXOS

Anexo 1. Presentación de los resultados de la focalización del subsidio tarifa de la dignidad. Se presentan gráficos ampliados de algunos sectores de Quito con los resultados de la focalización más restrictiva; en color verde se muestran los usuarios que sí recibirían el subsidio y en color rojo los que dejarían de recibirlo.



Figura 8.1. Mapa de focalización del subsidio. Zona Norte, sector Parque la Carolina.



Figura 8.2. Mapa de focalización del subsidio. Zona Centro, sector Centro Histórico.



Figura 8.3. Mapa de focalización del subsidio. Zona Sur, sector Las Cuadras (Quicentro Sur).

Anexo 2. Presentación de los resultados de la focalización del subsidio tarifa de la dignidad. Se presentan gráficos ampliados de algunos sectores de Quito con los resultados de la focalización más restrictiva; en color verde se muestran los usuarios que sí recibirían el subsidio y en color rojo los que dejarían de recibirlo.



Figura 8.4. Mapa de focalización del subsidio. Zona Norte, sector Parque la Carolina.



Figura 8.5. Mapa de focalización del subsidio. Zona Centro, sector Centro Histórico.



Figura 8.6. Mapa de focalización del subsidio. Zona Sur, sector Las Cuadras (Quicentro Sur).