

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA**

**OBTENCIÓN DE CURVAS DE CARGA PARA LOS DIFERENTES  
TIPOS DE USUARIOS RESIDENCIAL, COMERCIAL E INDUSTRIAL  
A TRAVÉS DEL ALGORITMO DE AGRUPAMIENTO K-MEANS.**

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO  
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN  
ELECTRICIDAD**

**WLADIMIR HUMBERTO MAFLA QUEZADA**

**wladimir.mafla@epn.edu.ec**

**DIRECTOR: MSc. PATRICIA ELIZABETH OTERO VALLADARES**

**patricia.otero@epn.edu.ec**

**DMQ, Octubre 2022**

## CERTIFICACIONES

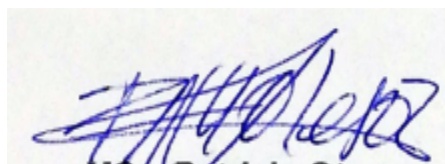
Yo, WLADIMIR HUMBERTO MAFLA QUEZADA declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.



---

**WLADIMIR MAFLA**

Certifico que el presente trabajo de integración curricular fue desarrollado por WLADIMIR HUMBERTO MAFLA QUEZADA, bajo mi supervisión.



---

**MSc. Patricia Otero**  
**DIRECTOR**

## **DECLARACIÓN DE AUTORÍA**

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el (los) producto(s) resultante(s) del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

WLADIMIR HUMBERTO MAFLA QUEZADA

MSc. PATRICIA ELEIZABETH OTERO VALLADARES

## **DEDICATORIA**

### **A DIOS:**

Primero dedico este trabajo a Dios ya que gracias a él se me ha regalado sabiduría, perseverancia y el conocimiento día a día. Y porque no dedicárselo si me ha brindado la oportunidad de llegar a culminar mis estudios con este proyecto tan importante para mi futuro.

### **A MIS PADRES:**

Se lo dedico a mis padres por su apoyo incondicional que me guio a donde estoy ahora, a ellos por enseñarme a superar los retos para alcanzar mis metas. Por ser las personas que me han motivado y por convertirme en el hombre que soy.

### **A MIS HERMANOS:**

A mis hermanos porque son personas que han influido en mi vida, siendo un apoyo incondicional y confiando en mí.

### **A MIS MAESTROS:**

Les dedico este trabajo como símbolo a la gratitud, respeto y admiración que tengo al labor diario que realizan, por compartir su conocimiento con cada uno de nosotros, por la atención y paciencia que nos han brindado a lo largo de la carrera.

### **A MIS COMPANEROS:**

Se los dedico a mis amigos, las personas que caminaron este trayecto conmigo con los que compartimos la experiencia de aprender no solo conocimientos, sino también muchos otros valores y desafíos.

## **AGRADECIMIENTOS**

### **A MI DIRECTORA DE PROYECTO:**

Expreso mi más profundo agradecimiento a la MSc. Patricia Otero por su gran apoyo y motivación en la realización de este proyecto, por haberme transmitido los conocimientos y ayuda necesarios para la realización de este proyecto.

### **A MIS PADRES:**

Agradezco a mis padres por haberme apoyado a estudiar en esta prestigiosa universidad, por todo lo que han hecho por mí a lo largo de toda la carrera y mi vida

### **A LA UNIVERSIDAD:**

Tengo un profundo agradecimiento con la universidad y todos los docentes y personal, por permitirme vivir esta experiencia.

### **A MIS MAESTROS:**

Mis más profundos agradecimientos a todos los Ingenieros de la ESCUELA POLITECNICA NACIONAL por darnos el impulso y no permitirnos desfallecer sobre las adversidades, en especial mención a mi tutor PhD. Fabian Pérez quien siempre estuvo pendiente de mi progreso y me guio a lo largo de mi carrera.

## ÍNDICE DE CONTENIDO

CERTIFICACIONES.....	I
DECLARACIÓN DE AUTORÍA.....	II
DEDICATORIA.....	III
AGRADECIMIENTOS .....	IV
ÍNDICE DE CONTENIDO.....	V
RESUMEN .....	VII
ABSTRACT .....	VIII
1 INTRODUCCIÓN.....	1
1.1 Objetivo general .....	2
1.2 Objetivos específicos .....	2
1.3 Alcance .....	2
1.4 Marco teórico .....	3
1.4.1 Sistemas de distribución .....	3
1.4.2 Definiciones básicas .....	3
1.4.3 Curvas de carga.....	4
1.4.3.1. Curvas de carga diaria.....	5
1.4.3.2. Curvas de carga Anual .....	6
1.4.4 Categorías tarifarias y niveles de voltaje .....	6
1.4.4.1 Categorías tarifarias .....	6
1.4.5 Minería de datos .....	7
1.4.6 Términos relacionados a los métodos de agrupamiento.....	9
1.4.7. Herramientas para la agrupación de patrones de carga .....	9
1.4.8. Algoritmo k-means.....	10
1.4.9. Software Matlab .....	11
1.4.9.1 K-means en Matlab .....	11
2 METODOLOGÍA.....	12
2.1 Normalización de los datos de potencia.....	13
2.3 Algoritmo k-means .....	15
2.4 Utilización del algoritmo k-means.....	17
2.5 Índices MIA Y CDI.....	21
2.4.1. Índice MIA.....	21

2.4.2. Índice CDI .....	21
3 RESULTADOS, CONCLUSIONES Y RECOMENDACIONES.....	22
3.1 Resultados .....	22
3.1.1 Segmentación de datos .....	22
3.1.2 Aplicación del Algoritmo K-means .....	25
3.1.3 Segunda Aplicación del Algoritmo k-means y obtención de las curvas de características de carga.....	29
3.1.4 Tabla de índices CDI y MIA .....	35
3.2 Conclusiones.....	35
3.3 Recomendaciones.....	37
4 REFERENCIAS BIBLIOGRÁFICAS .....	38
5 ANEXOS.....	40

## RESUMEN

El presente trabajo de integración curricular aborda la obtención de curvas de carga representativas mediante la utilización del algoritmo de agrupamiento k-means. Este algoritmo se aplica a las medidas tomadas de los diferentes tipos de consumidores, con la finalidad de definir las curvas de carga diarias y comprobar la adecuación de estas, mediante el cálculo de los indicadores CDI (Clustering Dispersión Indicator) e MIA (Mean Index Adequacy).

Los datos son obtenidos de las mediciones realizadas por los equipos de calidad instalados en las cabeceras de los alimentadores correspondientes a los usuarios conectados en medio voltaje y de los transformadores de distribución. Estos datos son proporcionados por la empresa CNEL-EP-Unidad de Negocios Azogues.

Las curvas de carga representativas se realizarán a través de dos etapas principales. La primera comprende procesar la información para segmentarla por tipos de clientes y a su vez en una subclasificación por días y tiempo entre la toma de medidas. La segunda etapa es el desarrollo de minería de datos en específico del método de agrupamiento k-means, dicho método hace uso de la técnica de cluster con el fin de obtener las curvas de carga representativas para los diferentes tipos de usuarios residencial, comercial e industrial. Esto es realizado a través del algoritmo de agrupamiento k-means implementado en el software Matlab, el cual procesa la información segmentada y da como resultado las curvas de carga representativas.

Adicionalmente, se realiza la obtención de los indicadores CDI y MIA con el fin de comprobar la adecuación del método para la obtención de las curvas de carga representativas.

**PALABRAS CLAVE:** curvas representativas de carga, k-means, cluster, CDI, MIA.



## ABSTRACT

This curricular integration work deals with obtaining representative load curves by using the k-means clustering algorithm. This algorithm is applied to the measurements taken from the different types of consumers, in order to define the daily load curves and check their adequacy, by calculating the CDI (Clustering Dispersion Indicator) and MIA (Mean Index Adequacy indicators).

The data is obtained from the measurements carried out by the quality equipment installed at the headends of the feeders corresponding to the users connected in medium voltage and from the distribution transformers. These data are provided by the company CNEL-EP- Unidad de Negocios Azogues.

Representative load curves will be made through two main stages. The first involves processing the information to segment it by types of customers and in turn in a subclassification by days and time between taking measures. The second stage is the development of specific data mining of the k-means clustering method, said method makes use of the cluster technique in order to obtain the representative load curves for the different types of residential, commercial and industrial users. This is done through the k-means clustering algorithm implemented in Matlab software, which processes the segmented information and results in representative load curves.

Additionally, the CDI and MIA indicators are obtained in order to verify the adequacy of the method for obtaining representative load curves.

**KEYWORDS:** representative load curves, k-means, cluster, CDI, MIA.

# 1 INTRODUCCIÓN

El presente trabajo de integración curricular se centrará en la obtención de curvas de carga representativas mediante la utilización del algoritmo de agrupamiento k-means. Este algoritmo se aplicará a las medidas tomadas de los diferentes tipos de consumidores, con la finalidad de definir las curvas de carga diarias, así como el cálculo de los indicadores CDI (Clustering Dispersion Indicator) e MIA (Mean Index Adequacy).

El análisis de los datos obtenidos de las curvas de carga diaria permite: identificar diversos tipos de factores, el comportamiento de los clientes u usuarios, así como cargas que resulten en variaciones en la demanda eléctrica. Todos estos datos pueden ser utilizados para desarrollar la planificación eléctrica a corto, mediano y largo plazo [1] [2].

El agrupamiento clásico de k-means es un algoritmo de clusterización, en el que se agrupa un conjunto de datos  $x^m$  ( $m = 1, \dots, M$ ) en  $k = 1, \dots, K$  grupos basándose en sus características mediante un proceso iterativo. El algoritmo comienza con una primera suposición para los centroides del cluster  $c^k$ , para  $k = 1, \dots, K$  generalmente son seleccionados al azar del conjunto de datos. Los centroides clasifican los datos siendo asignados al centroide si la distancia es mínima entre el dato y el centroide del grupo. Los centroides clasifican al conjunto de datos en grupos (siguiendo la norma Euclidiana), la ubicación de los centroides es actualizada tomando la nueva posición del promedio de los elementos del grupo perteneciente. Este procedimiento se reitera hasta llegar a la estabilización de los centroides. El número óptimo de clusters no se conoce y la calidad depende del valor que especifica el usuario [3], [4].

Se tomarán las mediciones obtenidas por los equipos de calidad instalados en las cabeceras de los alimentadores que corresponden a los usuarios que se conectan a medio voltaje y de los transformadores de distribución. Esta información es proporcionada por la empresa CNEL-EP-Unidad de Negocios Azogues. Antes de que se ingresen los datos al Algoritmo estos deben ser tratados o procesados para segmentarlos por tipo de clientes, así también en una subclasificación en: días y a su vez por el tiempo entre la toma de medidas.

Con la utilización del método de agrupamiento k-means, siendo implementado a través de un algoritmo en el software Matlab procesara la información previamente segmentada, dando como resultado las curvas de carga representativas y los respectivos indicadores para los diferentes tipos de usuarios.

Los índices de adecuación MIA y CDI son calculados una vez aplicada la técnica de agrupamiento k-means a los datos, para poder evaluar el rendimiento de este, conociendo

que valores menores de los índices se traduce a una mayor calidad de desempeño. El indicador MIA presenta que tan compactos son los grupos y a su vez CDI presenta que tan separados estos se encuentran.

## **1.1 Objetivo general**

Utilizar la técnica de agrupamiento k-means para la obtención de curvas de carga representativas para los diferentes tipos de consumidores residencial, comercial e industrial.

## **1.2 Objetivos específicos**

1. Realizar una investigación bibliográfica acerca de: curvas de carga, el algoritmo de agrupamiento k-means y como se lo utiliza para la obtención de curvas de carga.
2. Segmentar la información y elaborar una base de datos correspondiente a cada tipo de usuario.
3. Desarrollar el algoritmo k-means utilizando Matlab para obtener las curvas de carga representativas de los usuarios tipo residencial, comercial e industrial; así como los respectivos indicadores de adecuación.
4. Analizar los resultados obtenidos, emitir conclusiones y recomendaciones.

## **1.3 Alcance**

El trabajo de Integración comienza con la investigación bibliográfica del algoritmo K-means, y su utilización en la obtención de curvas de carga características para los diferentes tipos de consumidores: residencial, comercial e industrial.

Una vez obtenida la base teórica se puede segmentar y elaborar las respectivas bases de datos para cada tipo de consumidor y empieza la estructuración del algoritmo, el cual será implementado con la ayuda del software Matlab. El programa implementado en Matlab permitirá obtener y desplegar las respectivas curvas representativas para cada tipo de consumidor, así como sus respectivos indicadores MIA y CDI.

Con las curvas e indicadores obtenidos se realizará una evaluación de la técnica de agrupamiento utilizada, así también como su adecuación como herramienta para la obtención de curvas de carga representativas, permitiendo emitir las respectivas conclusiones y recomendaciones.

## 1.4 Marco teórico

A continuación, se presenta información teórica en la que está sustentada el proyecto, esta información es el soporte teórico para entender la obtención de las curvas características de carga para los diferentes tipos de usuarios: Industrial, Comercial, Residencial.

### 1.4.1 Sistemas de distribución

El sistema de distribución es caracterizado por ser el componente menos glamuroso en el sistema eléctrico. En la actualidad la operación de los componentes de generación y transmisión presentan muchos desafíos, las centrales de generación se hicieron cada vez más grandes; líneas de transmisión entrecruzadas. Por otro lado, los sistemas de distribución continúan entregando energía al medidor del usuario final, muchos de estos sistemas se encuentran sobre diseñados [1].

Un sistema eléctrico se compone de un sistema de generación, transmisión, subtransmisión y distribución. La tabla 1.4.1 muestra los diferentes niveles de un sistema típico a los que se encuentran estos componentes, así a su vez estadísticas de cada uno.

**Tabla 1.4.1** Niveles de voltaje de un sistema eléctrico [5]

Etapa del Sistema	Voltaje [kV]
Transmisión	500-230-138
Subtransmisión	138-69
Primario o Alimentador	34.5-22-13,8
Secundario o Acometida	0,12-0,24

Los sistemas de distribución generalmente empiezan con la subestación de distribución la cual se encuentra alimentada por una o más líneas de subtransmisión. En algunos casos la subestación de distribución se alimenta de una línea de transmisión de alto voltaje directamente, sin haber un sistema de subtransmisión; esto varía según la empresa. Cada subestación de distribución brindara el servicio a uno o más alimentadores primarios [2].

### 1.4.2 Definiciones básicas

A continuación, se enlistan algunas definiciones básicas referentes a la demanda:

- Demanda: “La demanda de una instalación o sistema es la carga en las terminales receptoras promediada sobre un intervalo de tiempo especificado” [1] [6]. La demanda es expresada en kilovatios [kW], kilovares [kVAR], kilo voltamperios [kVA], kiloamperios

[kA] o amperios [A]. En una definición más simple la demanda es la cantidad de potencia que se utiliza en cualquier momento por parte de un consumidor y la cual es variable en el tiempo.

- Intervalo de demanda: Es el periodo de tiempo en el que se promedia la carga. Este periodo por lo general toma valores de 15 min, 30 min, 1 h o incluso más.
- Máxima demanda: “La máxima demanda de una instalación o sistema es la mayor de todas demandas que han ocurrido durante el período de tiempo especificado” [1] [6]. La máxima demanda también es expresada debido al intervalo de tiempo utilizado para medirlo, por ejemplo, puede ser el máximo de las demandas diarias, mensuales o anuales.
- Demanda diversificada: Es la máxima suma de las aportaciones individuales de las demandas en un intervalo de tiempo específico.
- Densidad de carga: Es la relación entre la carga instalada y el área que ocupa el proyecto.
- Carga instalada: Es la capacidad de un usuario para consumir energía eléctrica, es expresada generalmente en kVA, MVA, kW o MW.
- Capacidad instalada: También conocida como capacidad nominal del sistema, es la suma de los transformadores, generadores conectados a líneas que suministran la potencia a las cargas.
- Factor de carga: Es la relación entre la carga promedio y la potencia máxima, determinadas durante un periodo de tiempo [1] [6].
- Factor de diversidad: Se lo calcula en base a la relación de la sumatoria de las demandas máximas individuales en las distintas subdivisiones del sistema ante la máxima demanda de todo el sistema [1] [6].

Factor de coincidencia: Se obtiene entre la relación entre las máximas demandas individuales de las cargas y el máximo valor de demanda del grupo de cargas [1] [6].

### **1.4.3 Curvas de carga**

La curva característica de carga consiste en plasmar la forma del consumo eléctrico de un cliente durante un intervalo de tiempo, se gráfica en base a puntos dados en un plano cartesiano, donde el eje x es representado por los intervalos de tiempo de manera ascendente y el eje y a su vez por el consumo eléctrico.

### 1.4.3.1. Curvas de carga diaria

Son curvas graficadas para el día pico de un periodo estadístico seleccionado. Se encuentran formadas por los picos obtenidos en pequeños intervalos de una hora para cada hora en el día. Estas curvas permiten tener una idea de las características de las cargas conectadas en el sistema sean estas residenciales, comerciales o industriales.

El análisis de estas curvas conduce a resultados similares a las curvas anuales, pero con mayores detalles sobre su forma de variación en el intervalo de tiempo. Estos análisis concluyen o permiten obtener tendencias predominantes de la carga, selección óptima de los equipos de transformación.

A continuación, en la figura 1.4.1 se presentan formas de curva de carga diaria típicas para diferentes tipos de consumidores comercial, residencial, industrial y alumbrado público.



Figura 1.4.1. Curvas de Carga diaria típicas [7]

### 1.4.3.2. Curvas de carga Anual

Por lo general estas curvas son graficadas para 4 años del periodo estadístico y muestra como incrementa la carga durante este periodo de tiempo, a través de su análisis se puede llegar a la obtención de la tasa de crecimiento de la demanda.

A las curvas de carga anual se las conforma de los valores de demanda pico durante cada mes, lo que permite observar los crecimientos y variaciones que se presentan en los picos de cada mes o año. El análisis de los resultados de esta curva generalmente guía a obtener conclusiones sobre el comportamiento del sistema y los factores que causan variaciones o afectaciones.

A continuación, en la figura 1.4.2 se presenta una curva de carga anual, donde los intervalos de tiempo son meses, y el consumo eléctrico en miles de KW.

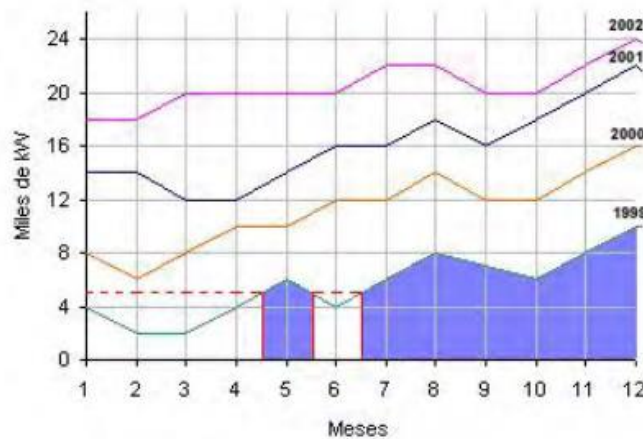


Figura 1.4.2. Curva de Carga anual [7]

### 1.4.4 Categorías tarifarias y niveles de voltaje

Dentro del sistema de servicio público de energía eléctrica se tiene en consideración dos categorías que son por características del consumidor: general y residencial; y, a su vez por su punto de entrega donde tenemos alto, medio y bajo voltaje.

#### 1.4.4.1 Categorías tarifarias

La categorización pasa a ser determinada y es responsabilidad de las empresas distribuidoras, a través de una evaluación de las características de la carga y el uso declarado por el consumidor de la energía. Así la empresa distribuidora aplica el tipo de tarifa correspondiente de acuerdo con el pliego tarifario [8].

Se consideran dentro de estas categorías:

Residenciales:

- a) Destinado de forma exclusiva al uso doméstico.
- b) Se incluye consumidores de bajos recursos económicos o bajo consumo que cuentan con una pequeña actividad artesanal o económica integrada a su vivienda.

Comerciales:

- a) Establecimientos comerciales (Privados o Públicos).
- b) Organismos de asistencia social.
- c) Organismos de beneficio público.
- d) Organismo Oficiales (sector público).
- e) Escenarios deportivos.
- f) Establecimientos destinados a la enseñanza o predicación de un culto religioso (capillas, iglesias, lugares de oración).
- g) Y demás establecimientos que no se encuentren considerados en la categoría Residencial.

Industriales:

- a) Todos los establecimientos u organismos encargados de operaciones de materiales para la obtención, transporte u transformación de productos naturales.
- b) Establecimientos o locales destinados a la elaboración de productos por cualquier proceso industrial y sus oficinas administrativas.
- c) Instalaciones de bombeo de agua incluido oficinas administrativas [8].

#### **1.4.5 Minería de datos**

Se la puede definir como un proceso de análisis y extracción de modelos, con la finalidad de encontrar patrones o tendencias. La minería de datos se la considera una de las etapas de un proceso mayor conocido como extracción de conocimiento esto a través de datos obtenidos de grandes bancos de información, donde se utiliza algunas técnicas como: inteligencia artificial, aprendizaje automático, gestión de la información y otras técnicas informáticas [7]. Para este trabajo se ha utilizado la minería de datos para obtener curvas de carga representativas de los diferentes tipos de usuarios residencial, comercial e industrial.

En la minería de datos tenemos las denominadas tareas las cuales se clasifican en predictivas o descriptivas.



Las predictivas se encuentran enfocadas en problemas donde como su nombre lo indica se necesita predecir para uno o más ejemplos, uno o más valores. Así esta correspondencia entre ejemplos y valores, definen algunas tareas predictivas como:

- Discriminación o Clasificación
- Clasificación Suave
- Estimación de la probabilidad de clasificación
- Categorización
- Priorización o preferencias
- Regresión

Las descriptivas por otro lado no se encargan de predecir datos nuevos, sino que describen los ya existentes. Estas definirían algunas tareas como:

- Clustering
- Factorizaciones y correlaciones
- Reglas de asociación
- Dependencias funcionales
- Detección de valores e instancias anómalas

Como se presentó con anterioridad cada una de estas tareas necesitan de algoritmos, métodos o técnicas para su resolución. Por lo cual en la siguiente Tabla 1.4.2. se presentará algunas técnicas.

**Tabla 1.4.2.** Técnicas de minería de datos [7]

Nombre	Predictivo		Descriptivo		
	Clasificación	Regresión	Clustering	Reglas de asociación	Correlaciones / Factorizaciones
Redes neuronales	X	X	X		
Arboles de decisión ID3, C4.5, C5.0	X				
Arboles de decisión CART	X	X			
Redes de Kohonen			X		

Regresión lineal y logarítmica		X			X
Regresión logística	X			X	
K-means			X		
Apriori				X	
Naive Bayes	X				

#### 1.4.6 Términos relacionados a los métodos de agrupamiento

- *Cluster*: Es un término de origen Ingles, el cual se lo traduce como conjunto o conglomerado [10].
- *Clustering*: Es categorizar la información del consumidor para generar segmentos relevantes, que compartan características similares. El clustering hace uso de un modelo matemático para descubrir grupos similares en función de encontrar hasta las pequeñas variaciones entre los usuarios de cada grupo [10].

#### 1.4.7. Herramientas para la agrupación de patrones de carga

Las técnicas de clustering parten de un conjunto de datos o información previamente recopilada, a continuación, se comparan diferentes herramientas para la agrupación de patrones de carga.

*Distancia Euclidiana*: la información es caracterizada por un vector que marca su posición, en este trabajo esta posición viene dada por un valor de potencia que se encuentra por unidad y a su vez en el dominio del tiempo el cual corresponde a intervalos de 1 hora por cada día. Así el algoritmo selecciona vectores de posición aleatorios para los centroides este número en algunos casos los escoge el usuario y en otros el mismo algoritmo los determina, una vez determinados, se mide la distancia entre los miembros del grupo y su respectivo centroide esta distancia es expresada como se presenta en la ecuación 1.1 [9].

$$d(X) = \sqrt{\frac{1}{2M} \sum_{m=1}^M d^2(x^m, X)}$$

**Ecuación 1.4.1.** Distancia Euclidiana

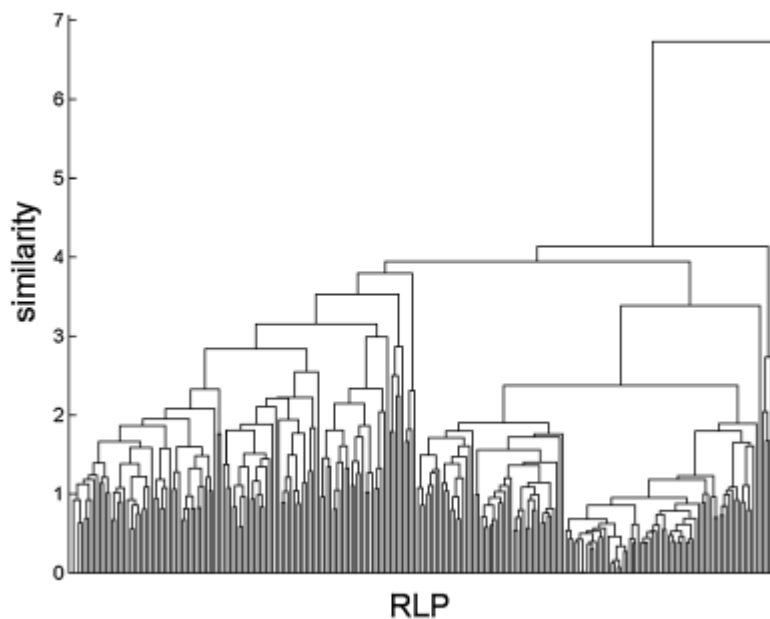
Donde:

M es el número de miembros pertenecientes al grupo X.

X un grupo de datos.

x uno de los elementos del grupo.

*Agrupación Jerárquica:* este proceso comienza con un número de cluster igual al número de datos obtenidos o parte de la base de datos, entonces se construye una primera matriz utilizando el criterio de la norma euclidiana. Posterior a esto los datos son agrupados en cluster binarios los cuales usan una matriz de similitud en base a un criterio de vinculación. Este proceso se repite de forma iterativa fusionando grupos en otros más grandes hasta llegar a un solo grupo, este proceso se basa en formar una estructura de árbol binario en donde la raíz es conformada por todos los datos de la base de datos y en su parte superior el proceso iterativo de la conformación de clusters. Para la conformación de estos conglomerados se tienen dos criterios los cuales son la vinculación única y la completa. Así de esta forma cada una utiliza el criterio de la distancia entre los puntos más cercanos y los puntos más alejados respectivamente; por lo que la vinculación única puede formar pocos grupos y a su vez la vinculación completa demasiados [10].



**Figura 1.4.3.** Representación del Árbol Jerárquico [4]

#### **1.4.8. Algoritmo k-means**

El algoritmo k-means toma un enfoque incremental, donde se agrega de forma dinámica un centro de cluster consecutivamente, llevando a una búsqueda determinística de N (donde N es el tamaño de un conjunto de datos) [12].

Es generalmente el método más usado y conocido. Es un método de agrupamiento no supervisado, donde se agrupa en base a patrones y a través de aprendizaje automático; el algoritmo es inicializado con un número de clusters [13].

El agrupamiento clásico de k-means es un algoritmo de clusterización, en el que se agrupa un conjunto de datos  $x^m$  ( $m = 1, \dots, M$ ) en  $k = 1, \dots, K$  grupos basándose en sus características mediante un proceso iterativo. El algoritmo comienza con una primera suposición para los centroides del cluster  $c^k$ , para  $k = 1, \dots, K$  generalmente son seleccionados al azar del conjunto de datos. Los centroides clasifican los datos siendo asignados al centroide si la distancia es mínima entre el dato y el centroide del grupo. Los centroides clasifican al conjunto de datos en grupos (siguiendo la norma Euclidiana), la ubicación de los centroides es actualizada tomando la nueva posición del promedio de los elementos del grupo perteneciente. Este procedimiento se reitera hasta llegar a la estabilización de los centroides. El número óptimo de clusters no se conoce y la calidad depende del valor que especifica el usuario [3], [4].

#### **1.4.9. Software Matlab**

En el presente trabajo se utilizará este software para obtener las curvas de carga de diferentes tipos de usuarios, con la información de la base de datos obtenidos.

MATLAB combina un entorno hecho para procesos iterativos y matemáticos, y los integra a diseños con un lenguaje de programación directo que expresa las operaciones lógicas y matemáticas [14].

En el presente trabajo hacemos uso de una de sus herramientas el cual es el Live editor que sirve para la creación de scripts que combinan código y salidas ejecutables.

##### **1.4.9.1 K-means en Matlab**

Los siguientes comandos sirven para desarrollar el algoritmo, el comando k-means, realiza el agrupamiento para segmentar los datos ingresados en k grupos devolviendo la ubicación de los centros de cada cluster, a su vez también permite obtener la distancia entre puntos y centroide; o la suma de todos los puntos del grupo a su respectivo centroide [14].

Sintaxis:

```
idx = kmeans(X,k)
```

```
idx = kmeans(X,k,Name,Value)
```

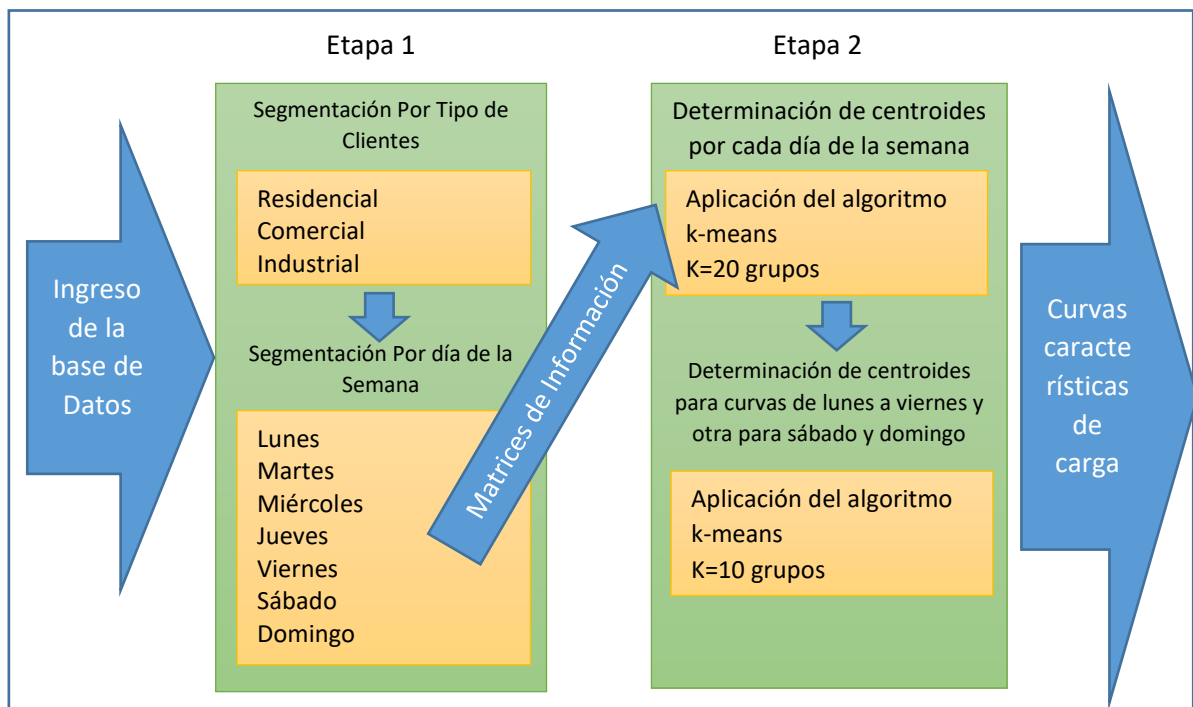
```
[idx,C] = kmeans(__)
```

```
[idx,C,sumd] = kmeans(__)
```

## 2 METODOLOGÍA PARA LA OBTENCIÓN DE CURVAS CARACTERÍSTICAS

El presente capítulo abarca el proceso para la obtención de curvas características de carga según los diferentes tipos de usuarios: residencial, comercial e industrial para la ciudad de Azogues. Se utilizará dos etapas, la primera comprende la segmentación de la base de datos por cliente y a su vez por días de la semana; la segunda emplea el algoritmo k-means para determinar puntos representativos de esta información y así determinar una curva característica para cada tipo de usuario.

Este algoritmo ha sido implementado con la ayuda de una herramienta computacional, Matlab, que es un sistema de cómputo numérico que tiene un lenguaje de programación propio (lenguaje M). Gracias a la facilidad de programación, así como las herramientas que hacen parte del software se pueden generar tablas, matrices y gráficos, lo cual lo convierte en una gran herramienta para interactuar con las grandes bases de datos proporcionadas por las empresas eléctricas. Además, se nos permite interactuar directamente con Excel, para el manejo de la información. La figura 2.1. ilustra el proceso a seguir de manera general.



**Figura 2.1.** Estructura de la metodología para la obtención de curvas características de carga

Como se muestra en la Figura 2.1., las flechas indican el flujo de datos, así como los procesos por los que pasará la información para llegar al resultado que serían las curvas

características de carga para los diferentes tipos de usuarios una para los días de lunes a viernes y otra para el fin de semana.

## 2.1 Normalización de los datos de potencia

Debido a la gran cantidad de datos que fueron proporcionados, estos deben ser segmentados y a su vez procesados ya que no pueden ser utilizados sin un proceso previo donde se escoge la información relevante para cumplir los objetivos.

Muchas veces los parámetros eléctricos optan por llevar a sus valores numéricos, a un valor relativo, ya sea en por unidad [p.u] o en por ciento [%]. Y esto es resultado de referenciar valores absolutos de dichos parámetros a sus correspondientes valores base [12].

Así expresar estos valores de esta forma presenta ventajas como:

- Tener una información relativa de la magnitud y así poder llevar a cabo una comparación pertinente de esta; por ejemplo, una pérdida de 2MW en una línea de transmisión es un dato sin relevancia sino es comparado con la cantidad de potencia activa de la línea.
- Se puede realizar una mejor comparación entre los valores de los equipos o elementos que conforman un sistema eléctrico.

Entonces para determinar este valor relativo en por unidad se utiliza la siguiente expresión matemática.

$$\text{Valor en [pu]} = \frac{\text{Valor Medido [unidad correspondiente]}}{\text{Valor Base Elegido [unidad correspondiente]}}$$

### **Ecuación 2.1** Ecuación para cambio de base a por unidad

Para el presente proyecto se utilizarán valores de potencias activas en el sistema por unidad tomando como valor base elegido a la máxima potencia del día.

Para conocer los hábitos de consumo se realizan estudios de caracterización de cargas, cuyo análisis resulta en un pequeño grupo de curvas de carga. Un estudio de caracterización de carga comienza con el estudio de las mediciones de la curva de carga.

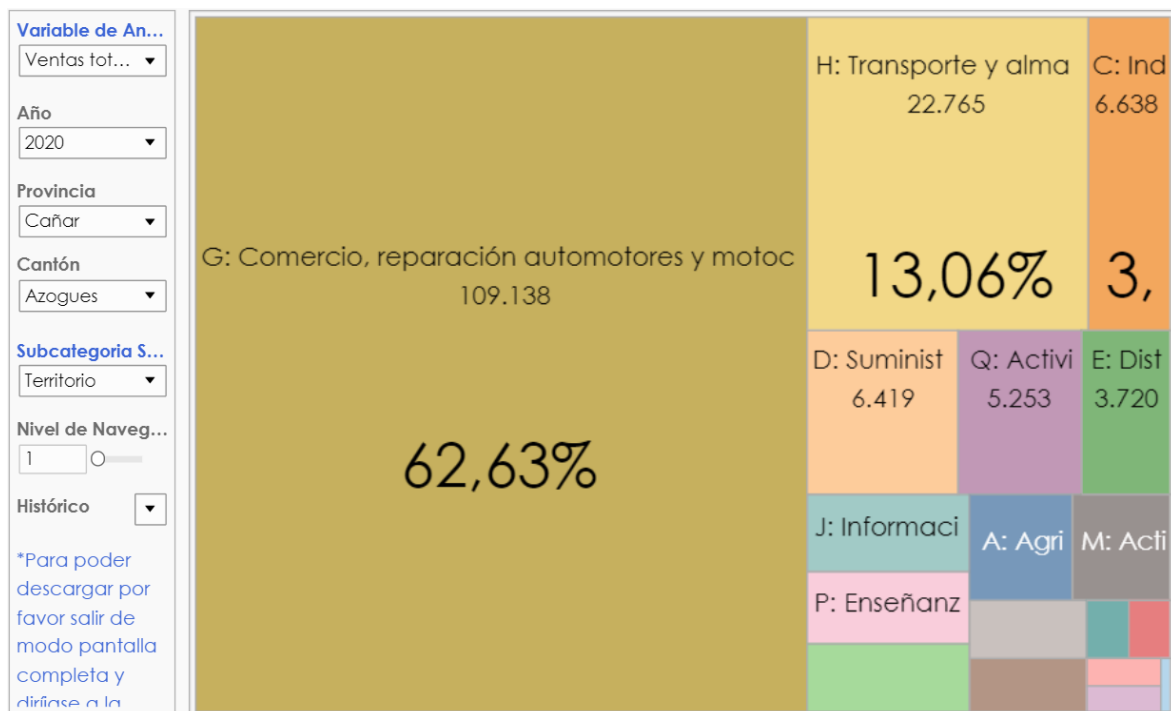
### 2.1.1. Periodo de Tiempo

Un grupo de medidas para la conformación de curvas de carga depende del tiempo en el que estas fueron efectuadas así obtenemos curvas de carga diarias, mensual o anual. Para el presente proyecto se tiene una toma de medidas por cerca de dos meses todos los días en un intervalo de 1 hora, lo cual nos resulta útil para conformar pares de potencia en por unidad y su respectiva hora del día.

## 2.2 Tipo de Clientes

La base datos que utiliza este proyecto es perteneciente a la ciudad de Azogues, ubicada en la provincia de Cañar. En la base de datos se tiene una especificación del tipo de cliente al que se realizó la recopilación de datos subdividiéndolos en 3 categorías que son industriales, comerciales o industriales.

Dentro de la ciudad se tiene diferentes tipos de negocios, los cuales se presentan en la figura 2.2, obtenida de la base de datos del INEC, específicamente en su sección Ecuador en cifras.



**Figura 2.2.** Segmentación de la Industria y el comercio de la ciudad de Azogues [16]

En la figura 2.3 se presenta como las industrias manufactureras basan sus actividades.



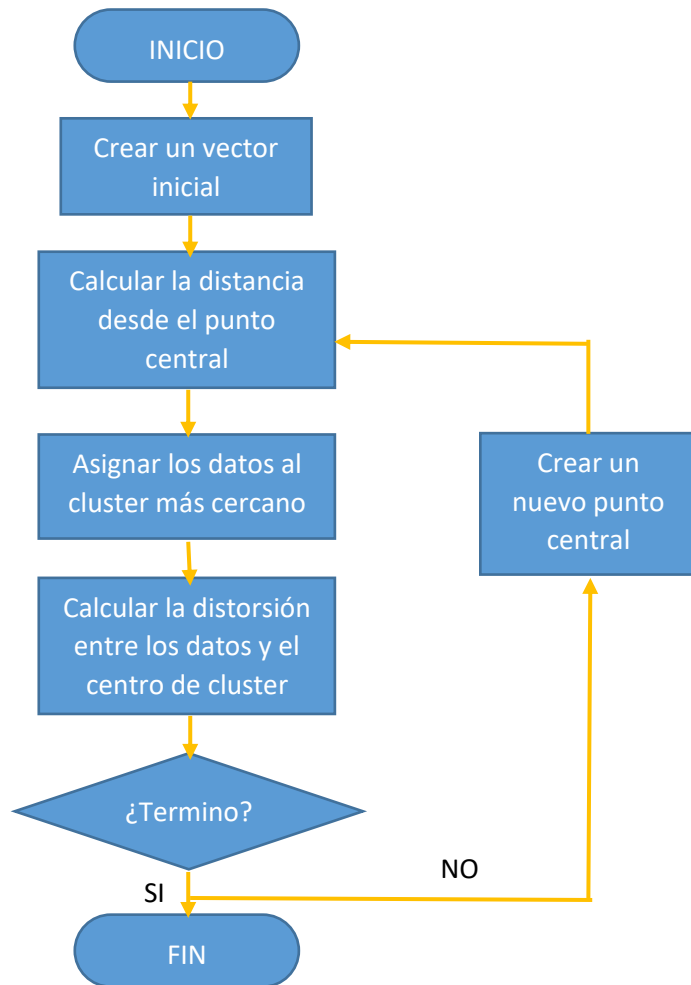
**Figura 2.3.** Segmentación de la Industria manufacturera en la ciudad de Azogues [16]

## 2.3 Algoritmo k-means

Se utiliza como herramienta el algoritmo k-means, para determinar centroides dentro de la gran variedad de datos obtenidos, estos centroides pueden considerarse puntos de carga representativos para sus respectivos grupos. Obteniendo un grupo de puntos que representan el consumo general y promedio de cada grupo de clientes sea este residencial, comercial e industrial.

El algoritmo de k-means fue propuesto en 1957 por Cox y en 1958 por Fisher, pero no es hasta el año de 1967 y 1975 que fue desarrollado por MacQueen y Hartigan respectivamente. Es un método usado comúnmente para la agrupación de datos, agrupándolos en “k” grupos, el cual es previamente seleccionado [16]. A continuación, se presenta un diagrama de flujo representando los pasos del algoritmo.





**Figura 2.4.** Diagrama de Flujos del Algoritmo K-means

Los pasos descritos a continuación, se extraen del diagrama de flujo anterior:

- ❖ Paso 1: Se selecciona ubicaciones arbitrarias para los primeros centros de los cluster, este número depende de la selección del usuario.
- ❖ Paso 2: Se mide la distancia euclidiana entre los diferentes puntos y cada uno de los centroides ubicados, y se busca la distancia mínima.
- ❖ Paso 3: El centroide más cercano al punto acogerá a este dato en su respectivo cluster o grupo, dividiendo el conjunto de datos en “k” grupos o clusters.
- ❖ Paso 4: Se renueva la ubicación de los centroides y se repite el proceso de agrupación.
- ❖ Paso 5: Se obtiene la suma total de las distancias entre los datos del grupo y su respectivo centroide.

- ❖ Paso 6: Si el valor de la suma es igual a los valores obtenidos previamente en la anterior iteración, el algoritmo se detiene caso contrario sigue dando nuevas ubicaciones a los centroides.

La presente exposición de pasos indica el funcionamiento del algoritmo, ahora se procederá a mostrar cómo fue implementado dentro de nuestro trabajo para la obtención de las curvas características de carga para los diferentes tipos de usuarios.

## 2.4 Utilización del algoritmo k-means

Una vez que se segmentó la información obtenida de la base de datos proporcionada se obtiene las respectivas matrices de los datos segmentados, se dividió los datos en respectivas matrices la primera fue por el tipo de cliente como se muestra a continuación en la Figura 2.5. esto se realizó con un lazo donde se comparan las descripciones por tipo de cliente y ordena los datos en 3 grupos: industrial, comercial, residencial.

```
for i = 1:n
    a=table2array(basedatos(i,12));
    b='Industrial';
    c='Residencial';
    d='Comercial';
    e='NoExiste';
    f='Otros';
    ind=strcmp(a,b);
    res=strcmp(a,c);
    com=strcmp(a,d);
    otr=strcmp(a,e);
    noex=strcmp(a,f);
    if ind==1
        industrial(j,:)=basedatos(i,:);
        j=j+1;
    elseif res==1
        residencial(k,:)=basedatos(i,:);
        k=k+1;
    elseif com==1
        comercial(l,:)=basedatos(i,:);
        l=l+1;
    elseif otr==1
        otro(m,:)=basedatos(i,:);
        m=m+1;
    else
        noexiste(n2,:)=basedatos(i,:);
        n2=n2+1;
    end
end
```

**Figura 2.5.** Lazo para la división de los datos según el tipo de cliente

Con la matriz para cada tipo de cliente, se procedió a dividir los datos en sub matrices dependiendo el día en la que fue tomada la división, esto para los 7 días de la semana obteniendo sub matrices por tipo de cliente y a su vez por día de la semana; en la presente Figura 2.6. se muestra el lazo que permite esta subdivisión.

```

for i = 1:n
    a=table2array(basedatos(i,10));
    if a==1
        ind_lunes(j,:)=basedatos(i,:);
        j=j+1;
    elseif a==2
        ind_martes(k,:)=basedatos(i,:);
        k=k+1;
    elseif a==3
        ind_miercoles(l,:)=basedatos(i,:);
        l=l+1;
    elseif a==4
        ind_jueves(m,:)=basedatos(i,:);
        m=m+1;
    elseif a==5
        ind_viernes(n2,:)=basedatos(i,:);
        n2=n2+1;
    elseif a==6
        ind_sabado(n3,:)=basedatos(i,:);
        n3=n3+1;
    else
        ind_domingo(n4,:)=basedatos(i,:);
        n4=n4+1;
    end
end

```

**Figura 2.6.** Lazo para la subdivisión de los datos según el día que fueron tomados los datos

Los datos de estas matrices son graficados en diagrama de puntos donde en el eje x tenemos la hora del día en la que fueron tomadas, estas horas se encuentran representadas en números enteros del 0 al 23; y a su vez en el eje “y” es representado por la potencia consumida en el sistema por unidad. En la presente Figura 2.7 se presenta la sección del código donde es seleccionado la información que se necesita para proceder a una gráfica de puntos conforme a la subdivisión por días de la semana.

```

%%POR DIAS
%%Lunes
ind_lunes_cluster(:,1)=ind_lunes(:,7);
ind_lunes_cluster(:,2)=ind_lunes(:,9);

```

**Figura 2.7.** Selección de la información necesaria para el ingreso al algoritmo k-means

Una vez obtenida ya una matriz con datos segmentados que se pueden ingresar al algoritmo k-means, dentro de Matlab se tiene ya un comando que permite realizar la agrupación para esto ingresan las sub matrices por día y se selecciona un numero de grupos en la que se dividirá todos estos, para la primera división se utilizó un numero de 20 grupos o 20 centroides, así el comando calcula estos centroides optimizados y entrega su ubicación. En la Figura 2.8 se presenta la aplicación del algoritmo para 20 centroides, esta sección también permite obtener la gráfica de puntos con los respectivos centroides ubicados.

```

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_lunes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'Linewidth',3)
title 'Datos Lunes';
legend('Cluster Centroid')
curv_ind_lunes=C;

```

**Figura 2.8.** aplicación del algoritmo k-means

Una vez obtenidos estos centros para cada día de la semana, sus ubicaciones representaran al grupo de medidas tomadas en ese día, y se agrupan estos puntos en 2 grupos nuevos, un grupo será conformado por los centroides encontrados en los días de lunes a viernes y otro grupo para los días sábado y domingo. Como se muestra en la Figura 2.9. se obtiene una matriz representativa, conformada por las submatrices de cada día.

```

%Dias_de_Lunes_a_Viernes
curv_ind_lunvie=[curv_ind_lunes;curv_ind_martes;curv_ind_miercoles;
    |curv_ind_jueves;curv_ind_viernes];

```

**Figura 2.9.** Combinación de las submatrices

Estas ubicaciones de los diferentes centroides localizados son ingresadas en una nueva matriz y a su vez ingresan por segunda vez al algoritmo k-means, con un nuevo número de clusters, para este caso se escogió 10 y estos 10 puntos permiten obtener una curva característica de carga que represente a todos los usuarios de los diferentes tipos sean estos residencial, comercial e industrial. En la Figura 2.10 se presenta la segunda aplicación del algoritmo k-means, con 10 puntos representativos o centroides.

```

% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility
Y2=curv_ind_lunvie;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...
    'Replicates',10,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes a Viernes';
legend('Cluster Centroid')
curv_ind_lunesaviernes=C;

```

**Figura 2.10.** Segunda aplicación del algoritmo k-means, con 10 centroides

Para la gráfica, el algoritmo despliega una matriz con las ubicaciones de los centroides, pero estos datos vienen desordenados según el orden en que fueron encontrados, para una apreciación de la curva característica, esta matriz debe ser ordenada por lo que se aplicó un lazo iterativo que ordene estos datos, para poder ser graficados y obtener la curva característica de carga, este proceso se aplicó para los tres tipos de usuarios. En la figura 2.11 se presenta el lazo iterativo que permite ordenar los centroides para su posterior grafica.

```

%% Curva de Carga Lunes a Viernes
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvalunvie(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)
        curvalunvie(2,:)=C(i,:);
    elseif arreglo(3,1)==C(i,1)
        curvalunvie(3,:)=C(i,:);
    elseif arreglo(4,1)==C(i,1)
        curvalunvie(4,:)=C(i,:);
    elseif arreglo(5,1)==C(i,1)
        curvalunvie(5,:)=C(i,:);
    elseif arreglo(6,1)==C(i,1)
        curvalunvie(6,:)=C(i,:);
    elseif arreglo(7,1)==C(i,1)
        curvalunvie(7,:)=C(i,:);
    elseif arreglo(8,1)==C(i,1)
        curvalunvie(8,:)=C(i,:);
    elseif arreglo(9,1)==C(i,1)
        curvalunvie(9,:)=C(i,:);
    elseif arreglo(10,1)==C(i,1)
        curvalunvie(10,:)=C(i,:);
    end
end
figure
plot(curvalunvie(:,1),curvalunvie(:,2))
title 'Curva de Carga para los días de Lunes a Viernes';

```

**Figura 2.11.** Lazo para ordenar la ubicación de los centroides

## 2.5 Índices MIA Y CDI

Los índices mean index adequacy y the clustering dispersion indicator permiten verificar el rendimiento de la técnica usada para el agrupamiento, valores pequeños de los índices indican una mayor calidad en el desempeño de la técnica [18].

### 2.4.1. Índice MIA

MIA (índice medio de adecuación), representa un valor dependiente de la cantidad de lo compacto que es cada grupo, es decir si los datos que pertenecen a ese cluster están muy juntos, el índice MIA es bajo [17] [18]. La ecuación 2.2. es la expresión matemática que se utiliza para el cálculo de este índice.

$$MIA = \sqrt{\frac{1}{K} \sum_{k=1}^K \sum_r d^2(r_k, C_k)}$$

**Ecuación 2.2.** Ecuación para el cálculo del factor MIA

Donde:

$C_k$ : son el conjunto de centroides.

$r_k$ : conjunto de miembros del cluster.

$k$ : numero de centroides

### 2.4.2. Índice CDI

CDI (indicador de dispersión de conglomerados), este valor depende de la distancia que existe entre los miembros del mismo grupo y adicional a esto incluye las distancias entre los diagramas representativos de cada cluster. Debido a esto el CDI mide lo compacto que esta los grupos y la cantidad por la cual difiere de los otros. Una buena solución al agrupamiento es relativa aún bajo valor del CDI [17] [18] . La ecuación 2.3. es la expresión matemática que se utiliza para el cálculo de este índice.

$$CDI = \frac{1}{d(C)} \sqrt{\frac{1}{K} \sum_{k=1}^K d^2(R_k)}$$

**Ecuación 2.3.** Ecuación para el cálculo del factor CDI

Donde:

C: son el conjunto de centroides.

$R_k$ : conjunto de miembros del cluster.

### 3 RESULTADOS, CONCLUSIONES Y RECOMENDACIONES

En el presente capítulo se muestran las matrices, tablas, gráficas de puntos y las respectivas curvas de carga obtenidas de la información proporcionada de la base de datos, para los diferentes tipos de usuarios.

#### 3.1 Resultados

En el presente apartado se presentarán las curvas características de carga obtenidas para los diferentes tipos de usuarios. Así como los factores MIA y CDI calculados.

##### 3.1.1 Segmentación de datos

Después de ingresar la matriz o tabla original dada por la empresa CNEL-EP-Unidad de Negocios Azogues, dada en una base de datos con formato Excel como se muestra a continuación en la Figura 3.1.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Medidor	Fec	Hora	Promedio P	Promedio Q	FechaSola	HoraSola	PmaxPorDia	PuPotProm	DiaSem	ClaveGCon	GrupodeConsum	Parr	ClaveParroquia
2	400096	11/6/2018	3:00:00 PM	3181.935	915.2416667	Tuesday		15	4260.81	0.746791103	2	30 Industrial	SAN MIGUEL	22
3	400096	11/6/2018	4:00:00 PM	3092.081667	880.815	Tuesday		16	4260.81	0.725702781	2	30 Industrial	SAN MIGUEL	22
4	400096	11/6/2018	5:00:00 PM	2969.395	825.0633333	Tuesday		17	4260.81	0.696908569	2	30 Industrial	SAN MIGUEL	22
5	400096	11/6/2018	6:00:00 PM	2994.276667	827.1266667	Tuesday		18	4260.81	0.702748225	2	30 Industrial	SAN MIGUEL	22
6	400096	11/6/2018	7:00:00 PM	2844.781667	764.7516667	Tuesday		19	4260.81	0.667662174	2	30 Industrial	SAN MIGUEL	22
7	400096	11/6/2018	8:00:00 PM	2844.718333	777.455	Tuesday		20	4260.81	0.66764731	2	30 Industrial	SAN MIGUEL	22
8	400096	11/6/2018	9:00:00 PM	2760.631667	727.685	Tuesday		21	4260.81	0.647912408	2	30 Industrial	SAN MIGUEL	22
9	400096	11/6/2018	10:00:00 PM	2780.085	693.4166667	Tuesday		22	4260.81	0.65247805	2	30 Industrial	SAN MIGUEL	22
10	400096	11/6/2018	11:00:00 PM	2736.586667	663.4633333	Tuesday		23	4260.81	0.642269115	2	30 Industrial	SAN MIGUEL	22
11	400096	11/7/2018	12:00:00 AM	2676.896667	504.2966667	Wednesday		0	4174.11	0.641309565	3	30 Industrial	SAN MIGUEL	22
12	400096	11/7/2018	1:00:00 AM	2554.196667	436.41	Wednesday		1	4174.11	0.611914077	3	30 Industrial	SAN MIGUEL	22
13	400096	11/7/2018	2:00:00 AM	2577.236667	440.91	Wednesday		2	4174.11	0.617433816	3	30 Industrial	SAN MIGUEL	22
14	400096	11/7/2018	3:00:00 AM	2596.155	438.4316667	Wednesday		3	4174.11	0.62196612	3	30 Industrial	SAN MIGUEL	22
15	400096	11/7/2018	4:00:00 AM	2594.54	439.3833333	Wednesday		4	4174.11	0.621579211	3	30 Industrial	SAN MIGUEL	22
16	400096	11/7/2018	5:00:00 AM	2568.195	447.1716667	Wednesday		5	4174.11	0.615267686	3	30 Industrial	SAN MIGUEL	22
17	400096	11/7/2018	6:00:00 AM	2587.271667	503.6966667	Wednesday		6	4174.11	0.619837922	3	30 Industrial	SAN MIGUEL	22
18	400096	11/7/2018	7:00:00 AM	2584.371667	492.9416667	Wednesday		7	4174.11	0.619143163	3	30 Industrial	SAN MIGUEL	22
19	400096	11/7/2018	8:00:00 AM	2586.781667	465.4283333	Wednesday		8	4174.11	0.619720531	3	30 Industrial	SAN MIGUEL	22
20	400096	11/7/2018	9:00:00 AM	2559.793333	451.76	Wednesday		9	4174.11	0.613254881	3	30 Industrial	SAN MIGUEL	22
21	400096	11/7/2018	10:00:00 AM	2522.116667	442.495	Wednesday		10	4174.11	0.604228606	3	30 Industrial	SAN MIGUEL	22
22	400096	11/7/2018	11:00:00 AM	2658.533333	501.3933333	Wednesday		11	4174.11	0.636910224	3	30 Industrial	SAN MIGUEL	22
23	400096	11/7/2018	12:00:00 PM	3027.101667	738.5966667	Wednesday		12	4174.11	0.725208887	3	30 Industrial	SAN MIGUEL	22

Figura 3.1. Porción de la matriz obtenida de la base datos proporcionada.

Esta matriz es ingresada dentro del software Matlab y se procede con su respectiva segmentación por tipo de clientes, como se puede apreciar en la figura anterior esta base de datos permite conocer qué tipo de cliente es y esta información se encuentra localizada en la columna L. Gracias a esto se compara y se divide esta matriz en 3 sub matrices una por cada tipo de consumidor como se muestra en las siguientes figuras 3.2, 3.3 y 3.4, los

datos han sido divididos por tipo de cliente, creando 3 grupos los que son Industriales, Comerciales y Residenciales.

industrial														
6950x14 table														
	1 Var1	2 Var2	3 Var3	4 Var4	5 Var5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10	11 Var11	12 Var12	13 Var13	14 Var14
1	400096	06-Nov-2018	0.6250	3.1819e+03	915.2417	'06-11-2018'	15	3.1819e+03	1	2	30	'Industrial'	'SAN MIGU...	22
2	400096	06-Nov-2018	0.6667	3.0921e+03	880.8150	'06-11-2018'	16	3.1819e+03	0.9718	2	30	'Industrial'	'SAN MIGU...	22
3	400096	06-Nov-2018	0.7083	2.9694e+03	825.0633	'06-11-2018'	17	3.1819e+03	0.9332	2	30	'Industrial'	'SAN MIGU...	22
4	400096	06-Nov-2018	0.7500	2.9943e+03	827.1267	'06-11-2018'	18	3.1819e+03	0.9410	2	30	'Industrial'	'SAN MIGU...	22
5	400096	06-Nov-2018	0.7917	2.8448e+03	764.7517	'06-11-2018'	19	3.1819e+03	0.8940	2	30	'Industrial'	'SAN MIGU...	22
6	400096	06-Nov-2018	0.8333	2.8447e+03	777.4550	'06-11-2018'	20	3.1819e+03	0.8940	2	30	'Industrial'	'SAN MIGU...	22
7	400096	06-Nov-2018	0.8750	2.7606e+03	727.6850	'06-11-2018'	21	3.1819e+03	0.8676	2	30	'Industrial'	'SAN MIGU...	22
8	400096	06-Nov-2018	0.9167	2.7801e+03	693.4167	'06-11-2018'	22	3.1819e+03	0.8737	2	30	'Industrial'	'SAN MIGU...	22
9	400096	06-Nov-2018	0.9583	2.7366e+03	663.4633	'06-11-2018'	23	3.1819e+03	0.8600	2	30	'Industrial'	'SAN MIGU...	22
10	400096	07-Nov-2018	0	2.6769e+03	504.2967	'07-11-2018'	0	3.1375e+03	0.8532	3	30	'Industrial'	'SAN MIGU...	22
11	400096	07-Nov-2018	0.0417	2.5542e+03	436.4100	'07-11-2018'	1	3.1375e+03	0.8141	3	30	'Industrial'	'SAN MIGU...	22
12	400096	07-Nov-2018	0.0833	2.5772e+03	440.9100	'07-11-2018'	2	3.1375e+03	0.8214	3	30	'Industrial'	'SAN MIGU...	22
13	400096	07-Nov-2018	0.1250	2.5962e+03	438.4317	'07-11-2018'	3	3.1375e+03	0.8275	3	30	'Industrial'	'SAN MIGU...	22
14	400096	07-Nov-2018	0.1667	2.5945e+03	439.3833	'07-11-2018'	4	3.1375e+03	0.8269	3	30	'Industrial'	'SAN MIGU...	22
15	400096	07-Nov-2018	0.2083	2.5682e+03	447.1717	'07-11-2018'	5	3.1375e+03	0.8185	3	30	'Industrial'	'SAN MIGU...	22
16	400096	07-Nov-2018	0.2500	2.5873e+03	503.6967	'07-11-2018'	6	3.1375e+03	0.8246	3	30	'Industrial'	'SAN MIGU...	22
17	400096	07-Nov-2018	0.2917	2.5844e+03	492.9417	'07-11-2018'	7	3.1375e+03	0.8237	3	30	'Industrial'	'SAN MIGU...	22
18	400096	07-Nov-2018	0.3333	2.5868e+03	465.4283	'07-11-2018'	8	3.1375e+03	0.8245	3	30	'Industrial'	'SAN MIGU...	22
19	400096	07-Nov-2018	0.3750	2.5598e+03	451.7600	'07-11-2018'	9	3.1375e+03	0.8159	3	30	'Industrial'	'SAN MIGU...	22
20	400096	07-Nov-2018	0.4167	2.5221e+03	442.4950	'07-11-2018'	10	3.1375e+03	0.8039	3	30	'Industrial'	'SAN MIGU...	22
21	400096	07-Nov-2018	0.4583	2.6585e+03	501.3933	'07-11-2018'	11	3.1375e+03	0.8473	3	30	'Industrial'	'SAN MIGU...	22
22	400096	07-Nov-2018	0.5000	3.0271e+03	738.5967	'07-11-2018'	12	3.1375e+03	0.9648	3	30	'Industrial'	'SAN MIGU...	22
23	400096	07-Nov-2018	0.5417	2.9624e+03	699.3800	'07-11-2018'	13	3.1375e+03	0.9442	3	30	'Industrial'	'SAN MIGU...	22
24	400096	07-Nov-2018	0.5833	3.1358e+03	816.6417	'07-11-2018'	14	3.1375e+03	0.9994	3	30	'Industrial'	'SAN MIGU...	22
25	400096	07-Nov-2018	0.6250	3.0845e+03	763.7550	'07-11-2018'	15	3.1375e+03	0.9831	3	30	'Industrial'	'SAN MIGU...	22

Figura 3.2. Porción de la matriz obtenida para los tipos de usuario: Industrial

comercial														
11160x14 table														
	1 Var1	2 Var2	3 Var3	4 Var4	5 Var5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10	11 Var11	12 Var12	13 Var13	14 Var14
1	101034	09-Jan-2018	NaN	1.2147e+03	1.0314e+03	'09-01-2018'	13	2.2019e+03	0.5517	2	20	'Comercial'	'BORRERO'	11
2	101034	09-Jan-2018	NaN	1.4062e+03	1.2101e+03	'09-01-2018'	14	2.2019e+03	0.6386	2	20	'Comercial'	'BORRERO'	11
3	101034	09-Jan-2018	NaN	1.3676e+03	1.2796e+03	'09-01-2018'	15	2.2019e+03	0.6211	2	20	'Comercial'	'BORRERO'	11
4	101034	09-Jan-2018	NaN	2.2019e+03	1.6772e+03	'09-01-2018'	16	2.2019e+03	1	2	20	'Comercial'	'BORRERO'	11
5	101034	09-Jan-2018	NaN	1.4782e+03	1.2204e+03	'09-01-2018'	17	2.2019e+03	0.6714	2	20	'Comercial'	'BORRERO'	11
6	101034	09-Jan-2018	NaN	1.6264e+03	1.2156e+03	'09-01-2018'	18	2.2019e+03	0.7386	2	20	'Comercial'	'BORRERO'	11
7	101034	09-Jan-2018	NaN	1.7367e+03	1.2130e+03	'09-01-2018'	19	2.2019e+03	0.7887	2	20	'Comercial'	'BORRERO'	11
8	101034	09-Jan-2018	NaN	1.6545e+03	1.1690e+03	'09-01-2018'	20	2.2019e+03	0.7514	2	20	'Comercial'	'BORRERO'	11
9	101034	09-Jan-2018	NaN	1.0950e+03	1.0495e+03	'09-01-2018'	21	2.2019e+03	0.4973	2	20	'Comercial'	'BORRERO'	11
10	101034	09-Jan-2018	NaN	823.3860	920.8590	'09-01-2018'	22	2.2019e+03	0.3740	2	20	'Comercial'	'BORRERO'	11
11	101034	09-Jan-2018	NaN	840.2470	900.6180	'09-01-2018'	23	2.2019e+03	0.3816	2	20	'Comercial'	'BORRERO'	11
12	101034	10-Jan-2018	NaN	674.6020	687.7110	'10-01-2018'	0	1.7882e+03	0.3773	3	20	'Comercial'	'BORRERO'	11
13	101034	10-Jan-2018	NaN	619.1290	601.9160	'10-01-2018'	1	1.7882e+03	0.3462	3	20	'Comercial'	'BORRERO'	11
14	101034	10-Jan-2018	NaN	680.3060	772	'10-01-2018'	2	1.7882e+03	0.3804	3	20	'Comercial'	'BORRERO'	11
15	101034	10-Jan-2018	NaN	714.0480	747.9590	'10-01-2018'	3	1.7882e+03	0.3993	3	20	'Comercial'	'BORRERO'	11
16	101034	10-Jan-2018	NaN	680.6140	672.2250	'10-01-2018'	4	1.7882e+03	0.3806	3	20	'Comercial'	'BORRERO'	11
17	101034	10-Jan-2018	NaN	659.7660	804.7510	'10-01-2018'	5	1.7882e+03	0.3690	3	20	'Comercial'	'BORRERO'	11
18	101034	10-Jan-2018	NaN	833.5110	956.5480	'10-01-2018'	6	1.7882e+03	0.4661	3	20	'Comercial'	'BORRERO'	11
19	101034	10-Jan-2018	NaN	809.6150	840.4270	'10-01-2018'	7	1.7882e+03	0.4528	3	20	'Comercial'	'BORRERO'	11
20	101034	10-Jan-2018	NaN	1.0141e+03	1.1429e+03	'10-01-2018'	8	1.7882e+03	0.5671	3	20	'Comercial'	'BORRERO'	11
21	101034	10-Jan-2018	NaN	1.4096e+03	1.2942e+03	'10-01-2018'	9	1.7882e+03	0.7883	3	20	'Comercial'	'BORRERO'	11
22	101034	10-Jan-2018	NaN	1.3643e+03	1.2445e+03	'10-01-2018'	10	1.7882e+03	0.7630	3	20	'Comercial'	'BORRERO'	11
23	101034	10-Jan-2018	NaN	1.4706e+03	1.2468e+03	'10-01-2018'	11	1.7882e+03	0.8224	3	20	'Comercial'	'BORRERO'	11
24	101034	10-Jan-2018	NaN	1.7111e+03	1.5391e+03	'10-01-2018'	12	1.7882e+03	0.9569	3	20	'Comercial'	'BORRERO'	11
25	101034	10-Jan-2018	NaN	1.6358e+03	1.4348e+03	'10-01-2018'	13	1.7882e+03	0.9148	3	20	'Comercial'	'BORRERO'	11

Figura 3.3. Porción de la matriz obtenida para los tipos de usuario: Comercial



	1 Var1	2 Var2	3 Var3	4 Var4	5 Var5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10	11 Var11	12 Var12	13 Var13	14 Var14
1	300786	18-Jan-2018	0.4167	4.8516e+03	3.5284e+03	'18-01-2018'	10	6.3871e+03	0.7596	4	10 'Residencial'	'AZOGUES'	11	
2	300786	18-Jan-2018	0.4583	2.8195e+03	1.2839e+03	'18-01-2018'	11	6.3871e+03	0.4414	4	10 'Residencial'	'AZOGUES'	11	
3	300786	18-Jan-2018	0.5000	4.9202e+03	3.5676e+03	'18-01-2018'	12	6.3871e+03	0.7703	4	10 'Residencial'	'AZOGUES'	11	
4	300786	18-Jan-2018	0.5417	3.4533e+03	3.9727e+03	'18-01-2018'	13	6.3871e+03	0.5407	4	10 'Residencial'	'AZOGUES'	11	
5	300786	18-Jan-2018	0.5833	4.3582e+03	3.5676e+03	'18-01-2018'	14	6.3871e+03	0.6824	4	10 'Residencial'	'AZOGUES'	11	
6	300786	18-Jan-2018	0.6250	5.6716e+03	3.4369e+03	'18-01-2018'	15	6.3871e+03	0.8880	4	10 'Residencial'	'AZOGUES'	11	
7	300786	18-Jan-2018	0.6667	6.3871e+03	3.7996e+03	'18-01-2018'	16	6.3871e+03	1	4	10 'Residencial'	'AZOGUES'	11	
8	300786	18-Jan-2018	0.7083	2.6528e+03	2.4503e+03	'18-01-2018'	17	6.3871e+03	0.4153	4	10 'Residencial'	'AZOGUES'	11	
9	300786	18-Jan-2018	0.7500	2.1236e+03	98.0100	'18-01-2018'	18	6.3871e+03	0.3325	4	10 'Residencial'	'AZOGUES'	11	
10	300786	18-Jan-2018	0.7917	2.4405e+03	360.6840	'18-01-2018'	19	6.3871e+03	0.3821	4	10 'Residencial'	'AZOGUES'	11	
11	300786	18-Jan-2018	0.8333	2.9991e+03	271.1650	'18-01-2018'	20	6.3871e+03	0.4696	4	10 'Residencial'	'AZOGUES'	11	
12	300786	18-Jan-2018	0.8750	1.8818e+03	0	'18-01-2018'	21	6.3871e+03	0.2946	4	10 'Residencial'	'AZOGUES'	11	
13	300786	18-Jan-2018	0.9167	1.8818e+03	0	'18-01-2018'	22	6.3871e+03	0.2946	4	10 'Residencial'	'AZOGUES'	11	
14	300786	18-Jan-2018	0.9583	1.8786e+03	0	'18-01-2018'	23	6.3871e+03	0.2941	4	10 'Residencial'	'AZOGUES'	11	
15	300786	19-Jan-2018	0	1.8622e+03	0	'19-01-2018'	0	7.9618e+03	0.2339	5	10 'Residencial'	'AZOGUES'	11	
16	300786	19-Jan-2018	0.0417	1.8459e+03	0	'19-01-2018'	1	7.9618e+03	0.2318	5	10 'Residencial'	'AZOGUES'	11	
17	300786	19-Jan-2018	0.0833	1.8393e+03	0	'19-01-2018'	2	7.9618e+03	0.2310	5	10 'Residencial'	'AZOGUES'	11	
18	300786	19-Jan-2018	0.1250	1.8426e+03	0	'19-01-2018'	3	7.9618e+03	0.2314	5	10 'Residencial'	'AZOGUES'	11	
19	300786	19-Jan-2018	0.1667	1.8426e+03	0	'19-01-2018'	4	7.9618e+03	0.2314	5	10 'Residencial'	'AZOGUES'	11	
20	300786	19-Jan-2018	0.2083	1.8361e+03	0	'19-01-2018'	5	7.9618e+03	0.2306	5	10 'Residencial'	'AZOGUES'	11	
21	300786	19-Jan-2018	0.2500	1.8263e+03	0	'19-01-2018'	6	7.9618e+03	0.2294	5	10 'Residencial'	'AZOGUES'	11	
22	300786	19-Jan-2018	0.2917	1.8524e+03	0	'19-01-2018'	7	7.9618e+03	0.2327	5	10 'Residencial'	'AZOGUES'	11	
23	300786	19-Jan-2018	0.3333	5.1227e+03	2.7443e+03	'19-01-2018'	8	7.9618e+03	0.6434	5	10 'Residencial'	'AZOGUES'	11	
24	300786	19-Jan-2018	0.3750	7.9618e+03	4.3844e+03	'19-01-2018'	9	7.9618e+03	1	5	10 'Residencial'	'AZOGUES'	11	
25	300786	19-Jan-2018	0.4167	6.7693e+03	3.6134e+03	'19-01-2018'	10	7.9618e+03	0.8502	5	10 'Residencial'	'AZOGUES'	11	

**Figura 3.4.** Porción de la matriz obtenida para los tipos de usuario: Residencial

A continuación, se realizó otra subdivisión de estas matrices obteniendo 7 sub matrices, una por cada día de la semana, y a su vez segmentando o eliminando la información que no es necesaria para el ingreso al comando de k-means; para esto solo se necesita la hora del día en el que fue tomada la medida y la potencia que fue consumida en su respectiva base por unidad. Las figuras 3.5, 3.6 y 3.7 representan la subdivisión por días de la semana y la información necesaria para el ingreso al algoritmo k-means.

	1 Var7	2 Var2
1	0	0.8346
2	1	0.8276
3	2	0.9000
4	3	0.8079
5	4	0.8424
6	5	0.7872
7	6	0.7908
8	7	0.8271
9	8	0.8396
10	9	0.8515
11	10	0.9197
12	11	0.8959
13	12	0.9457
14	13	0.9151
15	14	0.9502
16	15	0.9573
17	16	1
18	17	0.9679
19	18	0.9695
20	19	0.9351
21	20	0.9107
22	21	0.9241
23	22	0.8768
24	23	0.8671
25	0	0.4471

**Figura 3.5.** Porción de la matriz obtenida para los tipos de usuario: Residencial, lunes.

ind_lunes_cluster		
888x2 table		
	1 Var7	2 Var2
1	0	0.8346
2	1	0.8276
3	2	0.9000
4	3	0.8079
5	4	0.8424
6	5	0.7872
7	6	0.7908
8	7	0.8271
9	8	0.8396
10	9	0.8515
11	10	0.9197
12	11	0.8959
13	12	0.9457
14	13	0.9151
15	14	0.9502
16	15	0.9573
17	16	1
18	17	0.9679
19	18	0.9695
20	19	0.9351
21	20	0.9107
22	21	0.9241
23	22	0.8768
24	23	0.8671
25	0	0.4471

**Figura 3.6.** Porción de la matriz obtenida para los tipos de usuario: Industrial, lunes.

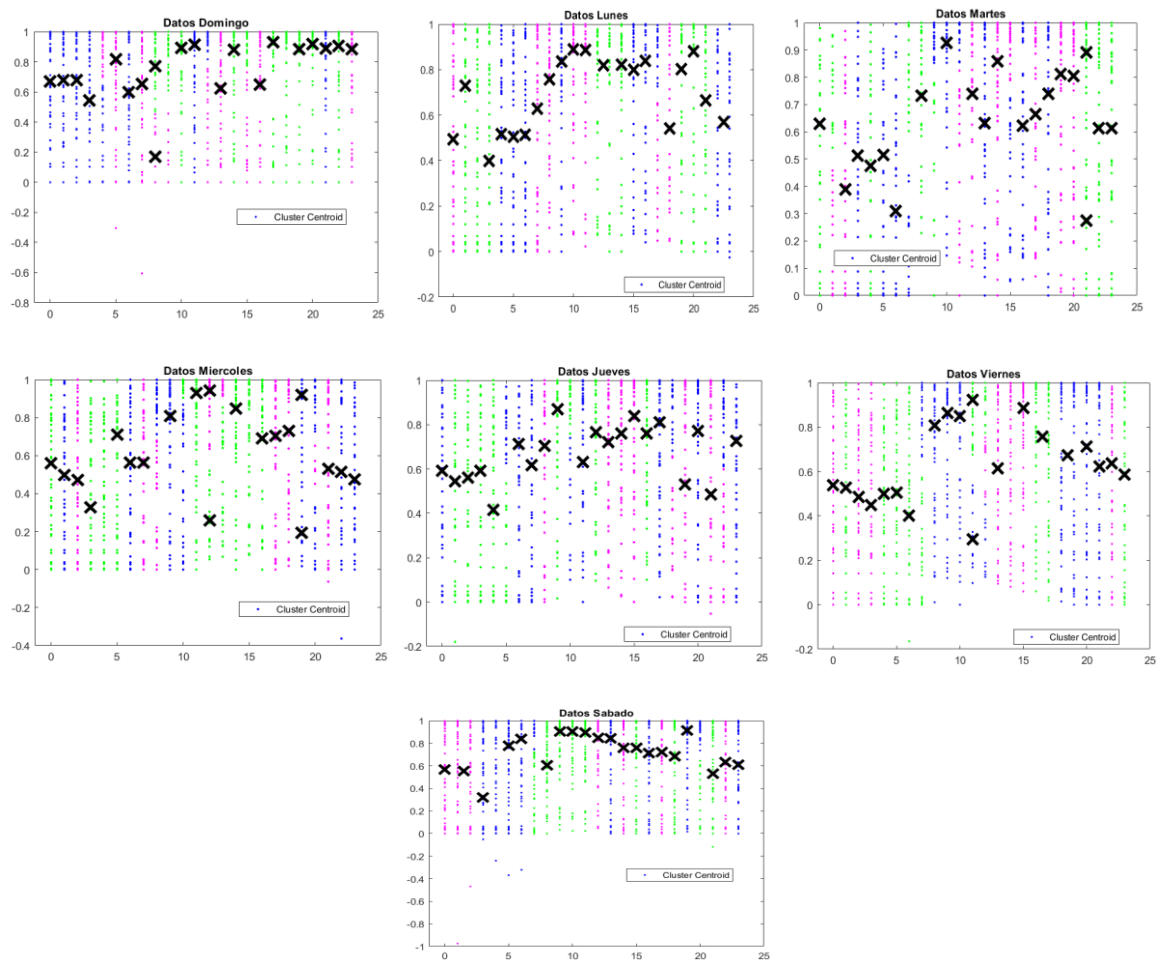
com_lunes_cluster		
1448x2 table		
	1 Var7	2 Var2
1	0	0.8346
2	1	0.8276
3	2	0.9000
4	3	0.8079
5	4	0.8424
6	5	0.7872
7	6	0.7908
8	7	0.8271
9	8	0.8396
10	9	0.8515
11	10	0.9197
12	11	0.8959
13	12	0.9457
14	13	0.9151
15	14	0.9502
16	15	0.9573
17	16	1
18	17	0.9679
19	18	0.9695
20	19	0.9351
21	20	0.9107
22	21	0.9241
23	22	0.8768
24	23	0.8671
25	0	0.4471

**Figura 3.7.** Porción de la matriz obtenida para los tipos de usuario: Comercial, lunes.

### 3.1.2 Aplicación del Algoritmo K-means

Una vez segmentado en estas sub matrices por día de la semana están ingresan al comando de k-means. Donde especificamos el número de centroides que queremos obtener en este primer caso hemos optado por un valor de 20 centroides. A continuación, se muestran las gráficas.

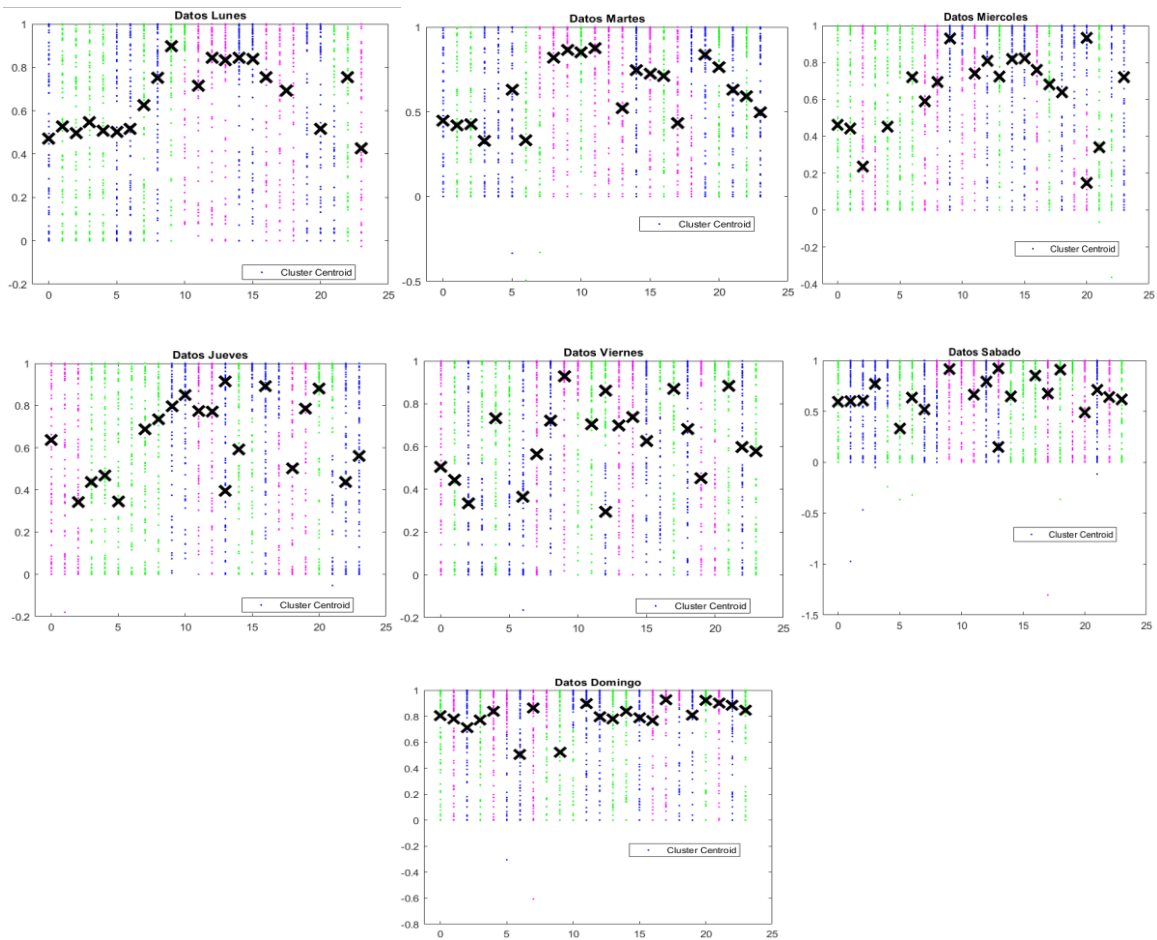
Clientes de tipo Industrial:



**Figura 3.8.** Graficas de puntos con 20 centroides por cada día de la semana, usuarios tipo: Industrial

Se observa que los centroides encontrados por el algoritmo no tienen una forma definida y para cada día se tiene una distribución diferente, esto representa el consumo representativo en cada día de la semana, sin embargo, de lunes a viernes se tiene una pequeña similitud, caso contrario con los días sábado y domingo; los cuales tienen una distribución muy diferente al resto. Se puede observar que a excepción del domingo la curva empieza a disminuir a partir de alrededor de las 8 de la noche, esto indica que dentro del sector industrial se maneja de mayor parte industrias que operan de manera laboral es decir que prenden y apagan sus equipos en los horarios establecidos, el domingo puede ser una excepción debido a que la industria representativa de este día es una que opera de manera continua es decir con sus equipos prendidos en todo momento. Los picos más altos se encuentran alrededor de las 10 de la mañana, que representa a los encendidos de la maquinaria y empiezo de producción.

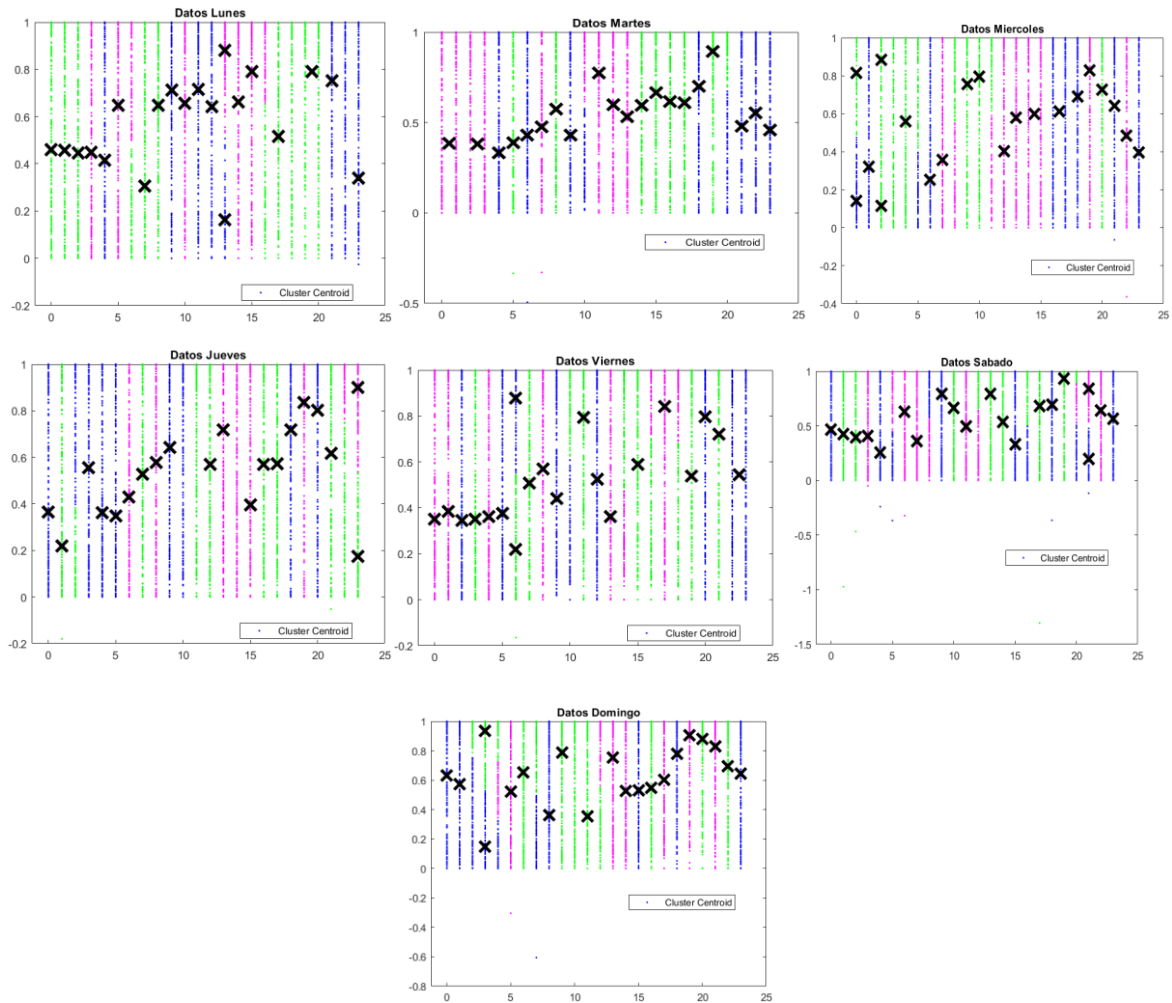
Clientes de tipo Comercial:



**Figura 3.9.** Graficas de puntos con 20 centroides por cada día de la semana, usuarios tipo: Comercial

Por las gráficas obtenidas, se puede apreciar los diferentes comportamientos del sector comercial, siendo la tendencia de consumo un sector que trabaja en jornadas laborales, es decir trabajan en un horario donde su consumo empieza a subir alrededor de las 7 de la mañana y a su vez disminuye alrededor de las 8 de la noche, siendo este tipo de consumidor el más representativo, sin embargo también se puede apreciar ubicaciones de centroides atípicas, las cuales pueden ser representadas por negocios que trabajen en el horario nocturno, este tipo de consumo se evidencia en el consumo del día domingo donde la curva viene a ser representada por este tipo de consumidor.

Clientes de tipo Residencial:



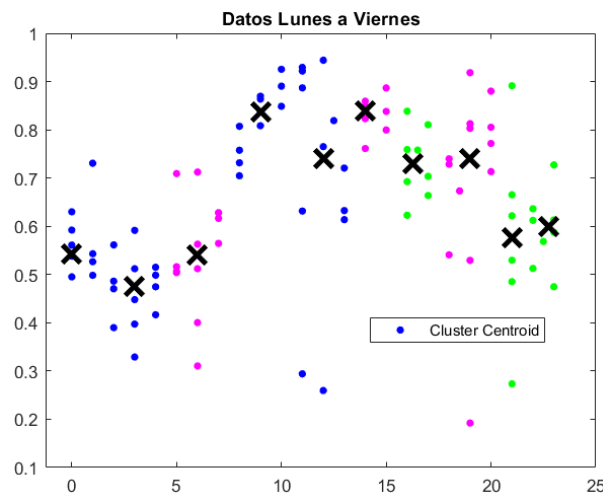
**Figura 3.10.** Graficas de puntos con 20 centroides por cada día de la semana, usuarios tipo: Residencial

Conforme a la ubicación de los centroides obtenidas por el algoritmo, se puede observar que el tipo de consumo de un usuario residencial es muy parecido en todos los días de la semana, empezando su horario de consumo alrededor de las 6 de la mañana y bajando desde las 20 horas, siendo los mayores picos alrededor de las 19 horas, siendo este el horario donde regresan del trabajo a sus casas. Debido a consumos diferentes no se tiene una caída de consumo en el horario de oficina es decir que la curva representativa viene dada por el consumo promedio de una familia donde algunos miembros se van a trabajar, pero otros se quedan dentro del hogar.

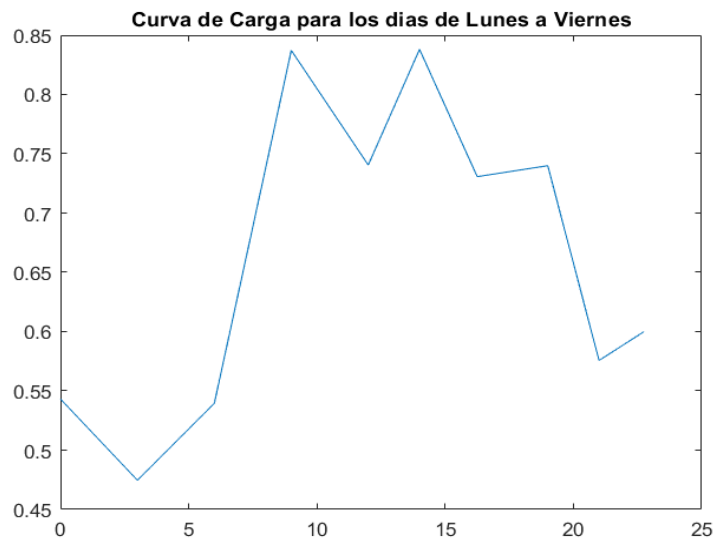
### 3.1.3 Segunda Aplicación del Algoritmo k-means y obtención de las curvas de características de carga.

Una vez se tiene la ubicación de estos 20 centroides se conforma una nueva matriz, con esta información y se procede a llevar a cabo una nueva aplicación del k-means separando los datos por una gráfica de lunes a viernes y otra para el fin de semana, para esta nueva aplicación se seleccionó un numero de centroides menor puesto que se tiene menos puntos, el numero seleccionado fue de 10 centroides; y estos representan a los puntos más representativos de cada día.

Usuarios tipo: Industrial.

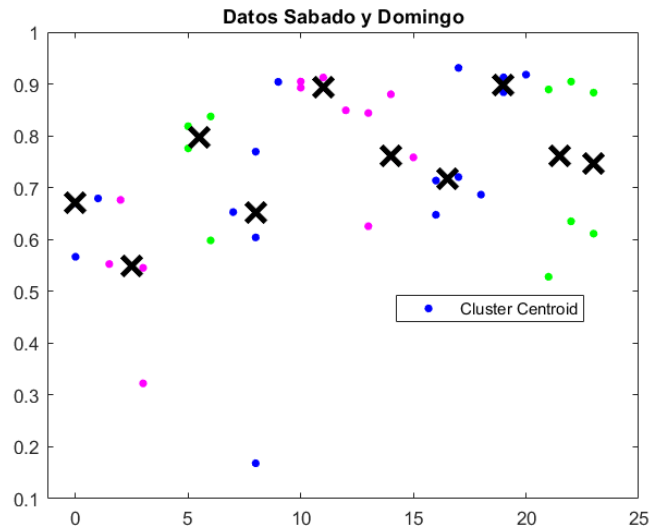


**Figura 3.11.** Graficas de puntos con 10 centroides usuarios tipo: Industrial, lunes a viernes

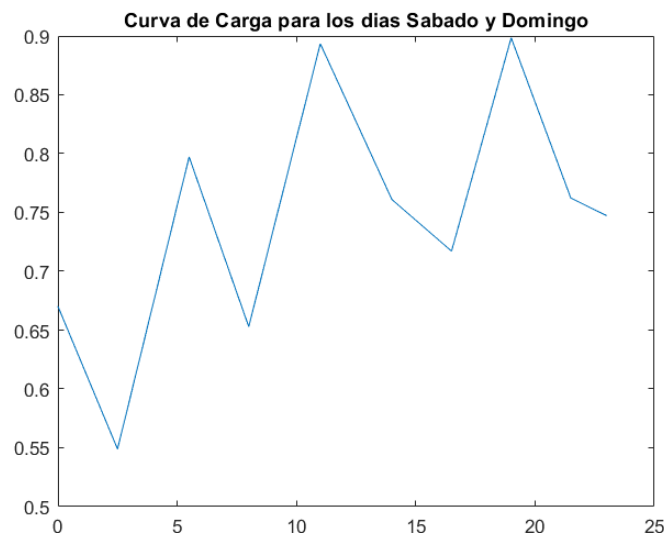


**Figura 3.12.** Curva característica de carga para usuarios tipo: Industrial, lunes a viernes

La curva característica de carga nos muestra el comportamiento representativo de los consumidores tipo industrial, en los días de lunes a viernes. Se puede observar que la industria representativa del sector industrial maneja horarios laborales es decir que apaga o detiene su consumo a partir de las 17 horas y vuelve a iniciar alrededor de las 7 horas, con un pico producido por el encendido o puesta en marcha de la maquinaria. Otro pico se observa en la tarde, esto puede deberse a que disminuyen su producción en la hora del almuerzo.



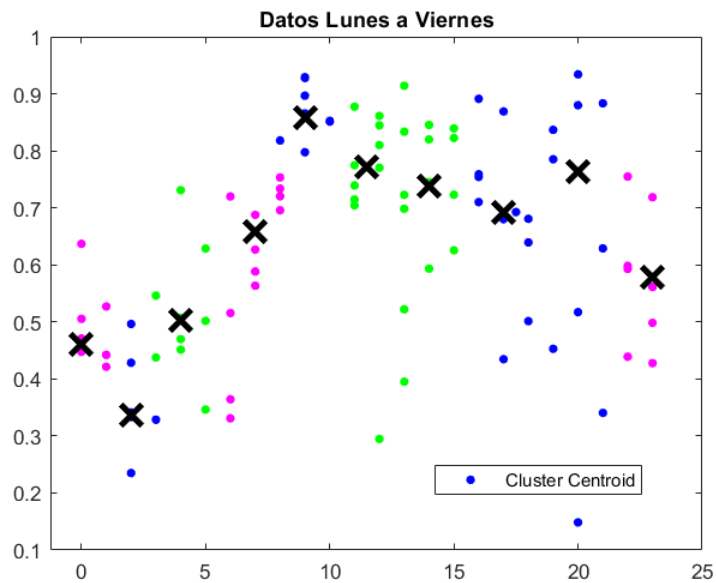
**Figura 3.13.** Graficas de puntos con 10 centroides usuarios tipo: Industrial, sábado y domingo



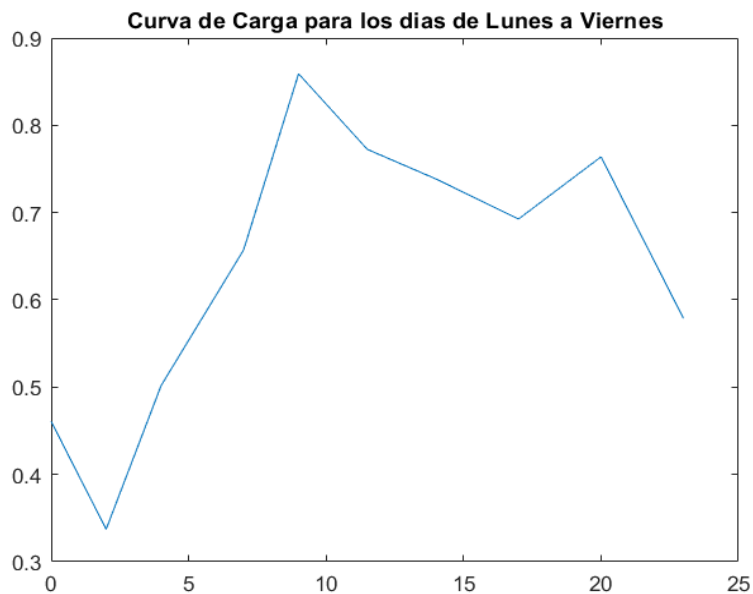
**Figura 3.14.** Curva característica de carga para usuarios tipo: Industrial, sábado y domingo

Para la curva característica de los días sábado y domingo, tenemos una forma muy diferente, esto se debe a que los fines de semana en especial como se mencionó, el domingo es representado por una industria que tiene un consumo permanente, lo que hizo que la curva de sábado y domingo tenga esta forma donde se evidencian picos, pero el consumo es casi permanente.

Usuarios tipo: Comercial.



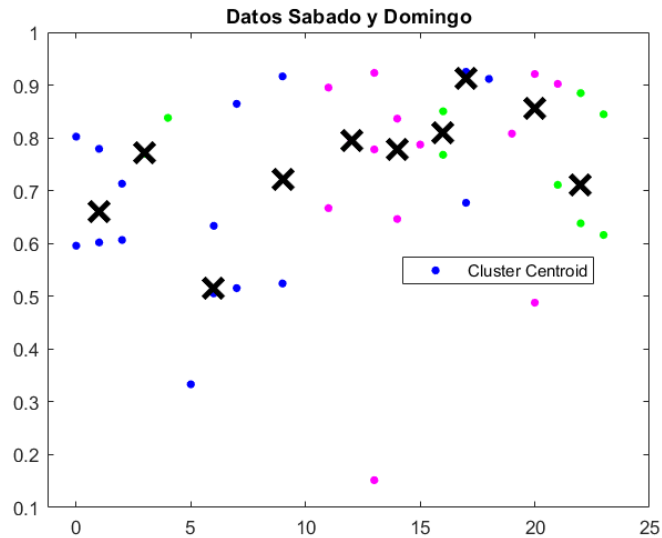
**Figura 3.15.** Graficas de puntos con 10 centroides usuarios tipo: Comercial, lunes a viernes



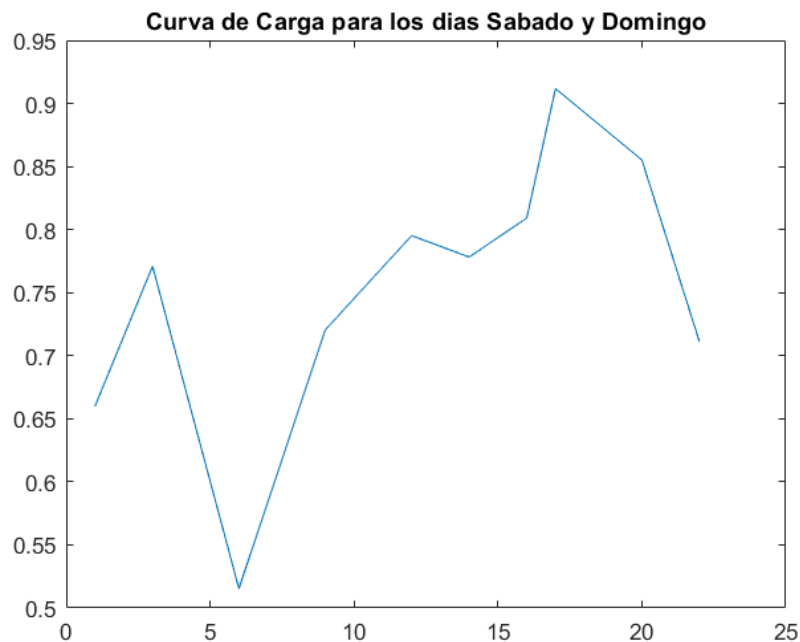
**Figura 3.16.** Curva característica de carga para usuarios tipo: Comercial, lunes a viernes



En base a la curva obtenida para los usuarios de tipo comercial, se observa que el consumo en los días de lunes a viernes es representado por un consumidor que tiene un horario laboral, comenzando desde alrededor de las 3 de la mañana que comienza a subir el consumo hasta llegar a un pico máximo que se da a las 10 de la mañana. El consumo disminuye hasta la tarde y finalmente decae, por cierre de negocios u oficinas.



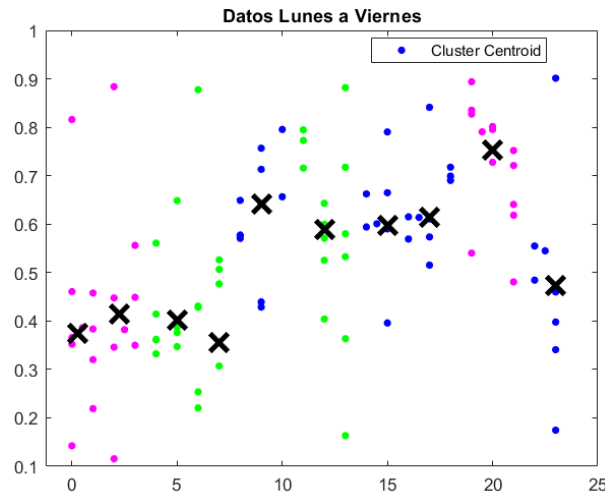
**Figura 3.17.** Graficas de puntos con 10 centroides usuarios tipo: Comercial, sábado y domingo



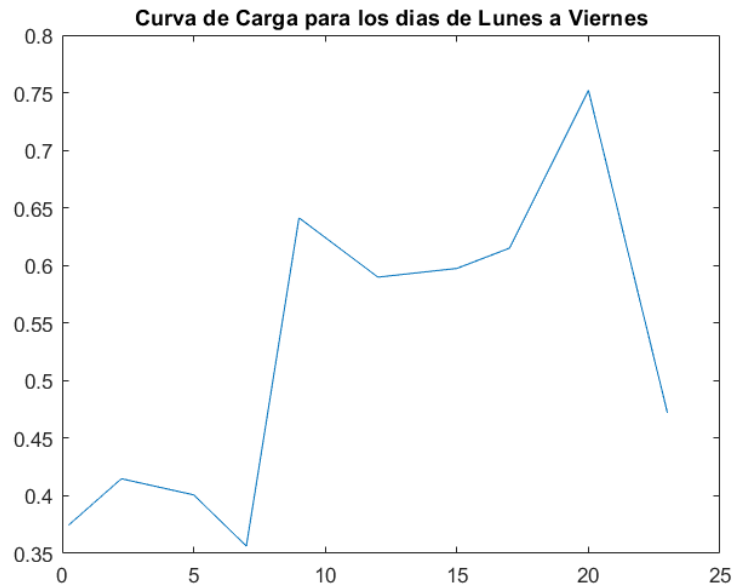
**Figura 3.18.** Curva característica de carga para usuarios tipo: Comercial, sábado y domingo

Para la curva representativa de los días sábado y domingo, se puede observar un pico inusual en la madrugada esto se puede deber debido al tipo de negocios que maneja la ciudad de Azogues, pero el resto de la curva es muy parecida a la presentada de los días lunes a viernes, con la diferencia que el pico más alto se encuentra en horas de la tarde, esto puede deberse a que el consumidor referencial de esta curva pueden ser locales o restaurantes que tienen su mayor clientela en horas de la tarde, pero así a su vez alrededor de las 8 de la noche tienen el cierre.

Usuarios tipo: Residencial.

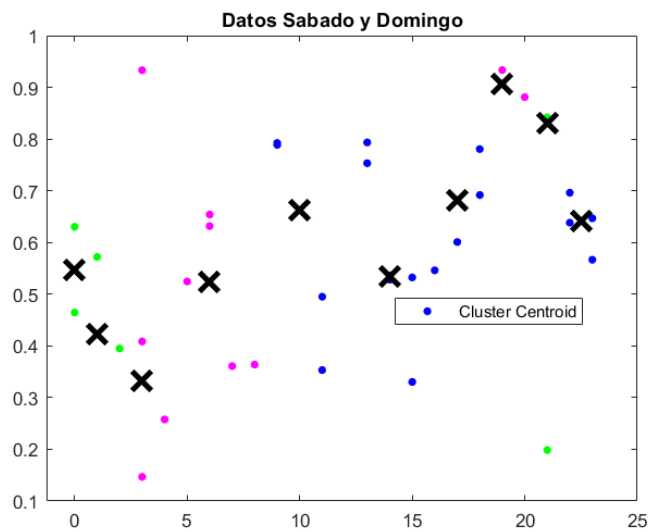


**Figura 3.19.** Graficas de puntos con 10 centroides usuarios tipo: Residencial, lunes a viernes

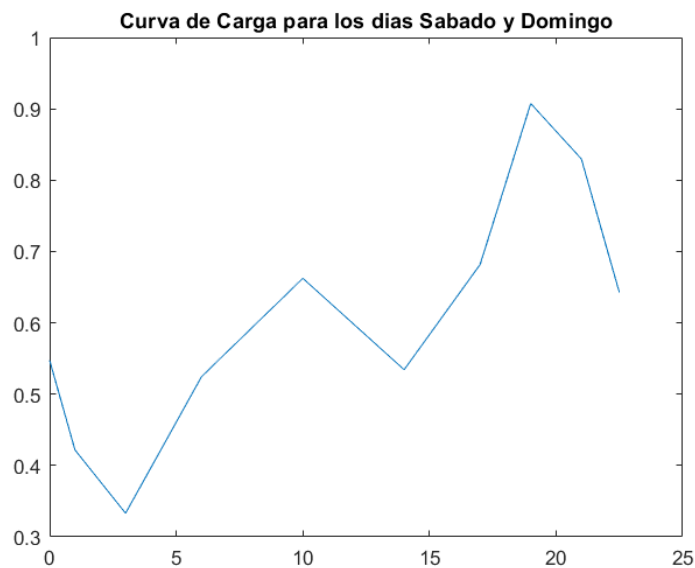


**Figura 3.20.** Curva característica de carga para usuarios tipo: Residencial, lunes a viernes

La curva representativa para usuarios de tipo residencial es muy parecida a la presentada en el marco teórico, esto se debe a que el consumo en este tipo de usuarios es muy parecido, donde inicia a las 7 de la mañana hasta las 9 que llega a su pico, se evidencia un consumo casi uniforme hasta alrededor de las 17 horas, esto se produce ya que el tipo de usuario representativo de esta curva se basa en un hogar donde algunos miembros salen a trabajar, pero otros se quedan y mantienen un consumo. El pico representa el horario que regresan de sus trabajos e incrementa el consumo hasta alrededor de las 20 horas y disminuye hasta la mañana siguiente.



**Figura 3.21.** Graficas de puntos con 10 centroides usuarios tipo: Residencial, sábado y domingo



**Figura 3.22.** Curva característica de carga para usuarios tipo: Residencial, sábado y domingo

La curva representativa de los días sábado y domingo muestra un comportamiento del usuario donde el primer pico representa el despertar de los hogares, pero se puede concluir por su forma que los usuarios representativos de estos días no permanecen en casa, y en la tarde que es su retorno el consumo se dispara hasta el pico de consumo más alto el que se da alrededor de las 19 horas, para finalmente decaer en horas de la noche y madrugada.

### 3.1.4 Tabla de índices CDI y MIA

**Tabla 3.1.** Índices MIA y CDI a los diferentes tipos de usuarios

	MIA Lunes-Viernes	CDI Lunes-Viernes	MIA Sábado-Domingo	CDI Sábado-Domingo
Industrial	2.46	1.09	1.51	0.61
Comercial	2.51	1.01	1.55	0.78
Residencial	2.52	0.84	1.56	0.52

Una vez obtenidas las curvas características, se procede a determinar los índices MIA y CDI con las ecuaciones presentadas anteriormente, en donde se obtienen los resultados presentados en la tabla 3.1, gracias a estos índices podemos observar cómo se está realizando el método de agrupación.

## 3.2 Conclusiones

Al utilizar el método k-means dentro del extenso grupo de datos obtenidos a través de mediciones correspondientes en lugares estratégicos, se puede observar que se ha realizado una agrupación en base a la cantidad de datos ubicados en la gráfica de puntos, donde se ah graficado potencia en base a la hora que fue tomada la medida, para esto se puede observar que la mejor manera de obtener una curva característica fue con la utilización del algoritmo k-means, pues al realizarse dos veces esta discrimino datos mal tomados como los negativos así como priorizo puntos estratégicos, para poder graficar las curvas características de carga.

En base a las curvas obtenidas se tiene una perspectiva del consumo realizado por los diferentes tipos de usuarios, se puede determinar horas pico, donde se dispara el consumo de estos usuarios, así como las horas donde mantiene un consumo constante y a su vez las horas donde registra consumos mínimos.

El correcto uso de la información obtenida de la base datos proporcionada es importante para conseguir la meta de este proyecto, el uso del sistema por unidad fue importante ya

que ayudó a poner en un mismo grado a toda la información obtenida para ser tratada de igual manera, esto sabiendo que no todos los usuarios en especial los de tipo industrial trabajan a los mismos niveles de voltaje, y tienen diferencia en sus consumos. Con la ayuda del software Matlab la información fue distribuida en matrices generales por tipos de usuario y estas a su vez fueron subdivididas por días de la semana, esto ayuda a manejar de forma más dinámica la información recopilada de la base de datos.

Las formas que tienen las curvas son representativas es decir son curvas que muestran un consumo general de los tipos de usuarios, a su vez como se sabe el consumo en la parte industrial es muy variado dependiendo de la actividad que esta realiza, por esto en base a la información obtenida en Ecuador en cifras una aplicación útil perteneciente a la entidad INEC, se sabe que la mayor parte de la actividad industrial en la ciudad no tiene un uso continuo es decir que hay horas donde se apagan las maquinas o suspenden producción, siendo activadas o prendidas de nuevo a horas de la mañana.

Los valores obtenidos de los índices MIA y CDI nos indican que hay mayor dispersión en los datos de lunes a viernes, en comparación con los de sábado y domingo. La mejor agrupación se la puede determinar por los valores obtenidos en el índice CDI y esta se encuentra en la categoría residencial debido a que la mayoría de los usuarios tienen el mismo índice de consumo, a su vez la que tuvo mayores dificultades es el sector industrial y comercial ya que los consumos son variados y no necesariamente comparten el mismo comportamiento de consumo.

Matlab ofrece convertirse en una herramienta muy útil, esto se debe a sus funciones preestablecidas, así como el permitir agregar nuevas. Su facilidad para manejar extensos grupos de datos en base a tablas o matrices facilita su manejo, así como el control de datos, su trabajo conjunto a Excel permite fortalecer el manejo de bases de datos, pudiendo compartir datos en ambos programas.

### **3.3 Recomendaciones**

Una vez propuesto y utilizado el método k-means para la obtención de curvas características de carga, se sugiere realizar comparaciones con otros métodos de clusterización, que tienen diferentes formas de agrupar los datos, para así poder determinar el mejor método o el que más se ajuste a las necesidades de la empresa.

Habiendo visto los resultados se sugiere el uso de k-means para cada día de la semana y a su vez en la gráfica general, ya que se obtuvieron mejores curvas representativas, para cada tipo de usuario comercial, industrial y residencial.

Con la presente obtención de curvas características de carga se puede tener una idea de la forma de consumo de energía de la ciudad de la que se ha tomado las mediciones, así tener una perspectiva para la determinación de las horas pico, donde se necesitara una mayor generación y prever el aumento de las instituciones, residentes, así como negocios.

La utilización y aplicación del algoritmo k-means en Matlab, será de mucha utilidad en futuros proyectos, debido a que su utilización es versátil y puede aplicarse en bases de datos más extensas, la comunicación entre Excel y Matlab es muy útil debido a que muchas bases de datos se encuentran realizadas en Excel. Las funciones utilizadas pueden ayudar a comprender mejor el funcionamiento de la minería de datos, por lo que es recomendable visualizar las herramientas que ofrece Matlab.

Con la obtención de las curvas características de carga, se puede llevar a cabo una comparación con la generación de energía y llegar a una planificación óptima, con el fin de minimizar costos de generación. Lo que puede ayudar en un mejor estudio de tarifas, para los diferentes tipos de consumidores.

## 4 REFERENCIAS BIBLIOGRÁFICAS

- [1] T. Gonen, *Electric Power Distribution Engineering*, New York: CRC Press, 2014.
- [2] W. H. Kersting, *Distribution System Modeling and Analysis*, Las Cruces, New Mexico: CRC Press, 2012.
- [3] P. B. S. Balasubramanian, "Characterising electricity demand through load curve clustering: A case of Karnataka electricity system in India," *Elsevier Ltd.*, p. 14, 2021.
- [4] F. P. Gianfranco Chicco, "Comparisons Among Clustering Techniques for Electricity Customer Classification," *IEEE Transactions on Power Systems*, vol. XXI, no. 2, p. 9, 2006.
- [5] H. L. Willis, "Power Distribution Planning Reference Book," in *Second Edition, Revised and Expanded*, North Carolina, U.S.A, Marcel Dekker, Inc., 2002, p. 11.
- [6] A. A. S. D. o. E. T. G. 3. G. of, *Transmission and Distribution*, VA: ASA C42.35, 1957.
- [7] S. R. Castano, *Redes de Distribucion de Energia*, Bogota: Manizales, Tercera Edicion , 2004.
- [8] A. d. R. y. C. d. E. y. R. N. n. Renovables, "Pliego Tarifario del Servivio Publico de Energia Electrica," Quito, 2021.
- [9] J. R. C. F. Jose Hernandez Orallo, *Introduccion a la Minería de Datos*, Valencia: Pearson.
- [10] R. N. P. P. M. S. a. C. T. G. Chicco, "Customer characterization options for improving the tariff offer," *IEEE Trans. Power Syst.*, vol. 18, no. 1, 2003, pp. 381-387.
- [11] M. R. Anderberg, *Cluster Analysis for Applications*, New York: Academic, 1973.
- [12] N. V. J. J. V. Aristidis Likas, "Reconocimiento de patrones," in *El algoritmo de agrupamiento global k -means*, ELSEVIER, 2003, pp. 451-461.
- [13] C. P. Sinaga and M.-S. Yang, "Algoritmo de agrupamiento de K-Means no supervisado," *IEEE*, 2006.
- [14] Mathworks, "Matlab," [Online]. Available: <https://www.mathworks.com/help/matlab/index.html>. [Accessed Febrero 2022].
- [15] O. Elgered, *Electric Energy Systems Theory*, McGraw Hill, 1982.
- [16] INEC, "Ecuador en Cifras," 2020. [Online]. Available: <https://public.tableau.com/app/profile/instituto.nacional.de.estad.stica.y.censos.inec>.

/viz/VisualizadordeEstadsticasesEmpresariales2020/Dportada. [Accessed 10 02 2022].

- [17] M. Amelot, "Programacion de Excel: Macros y lenguaje VBA," VBA excel 2016, Barcelona: ENI, 2016.
- [18] R. N. P. P. M. S. a. C. T. G. Chicco, "Customer characterization options for improving the tariff offer," in *Power Systems*, IEEE Transactions on, vol. 18, 2003, p. 381–387.
- [19] T. C. U. A. a. T. R. Ian Dent, "An Approach for Assessing Clustering of Households by Electricity Usage," University of Nottingham, Nottingham.
- [20] L. Carvajal, Metodología de la Investigación Científica. Curso general y aplicado, 28 ed., Santiago de Cali: U.S.C., 2006, p. 139.
- [21] H. L. Willis, Powe Distribution Planning Reference Book, New York: Marcel Dekker, 2002.
- [22] "Tnil," TOTVS, [Online]. Available: <https://blog.tail.digital/es/clusterizacion-que-es-y-por-que-es-tan-importante-para-el-marketing/#:~:text=La%20clusterizaci%C3%B3n%20es%20la%20categorizaci%C3%B3n,de%20personas%20con%20caracter%C3%ADsticas%20similares..> [Accessed 7 Diciembre 2021].
- [23] A. Carranza, "Crehana," 21 Enero 2021. [Online]. Available: <https://www.crehana.com/ec/blog/negocios/clusterizacion/>. [Accessed 7 Diciembre 2021].



## 5 ANEXOS

### ANEXO I

Código del Algoritmo en Matlab, en su herramienta Live Script

```
% PROYECTO DE INTEGRACION CURRICULAR

% Nombre: Wladimir Mafla
% Gr:7
% Determinacion de Curvas Carga e Indicadores

%% LIMPIEZA DE PANTALLA
clc;
clear all;
close all;
format long;

%% TITULO
fprintf('
                ESCUELA POLITECNICA NACIONAL \n');
fprintf('
                FACULTAD DE INGENIERIA ELECTRICA Y ELECTRONICA \n');
fprintf('
                TRABAJO DE INTEGRACION CURRICULAR \n');
fprintf('
                CURVAS DE CARGA E INDICADORES \n');
fprintf('NOMBRE:\n');
fprintf('MAFLA QUEZADA WLADIMIR\n');

%% TIPOS DE CLIENTES
fprintf('\nTIPOS DE CLIENTES\n');
fprintf(' 1.Clientes de tipo Industrial\n');

%% INGRESO DE DATOS
basedatos= readtable('BDEEA.xlsx','sheet','Base de
Datos','Range','A2:N54858');
n1=size(basedatos);
n=n1(1,1);

%%DIVISION POR TIPO DE CLIENTE
j=1;
k=1;
l=1;
m=1;
n2=1;

for i = 1:n
    a=table2array(basedatos(i,12));
    b='Industrial';
    c='Residencial';
```

```

d='Comercial';
e='NoExiste';
f='Otros';
ind=strcmp(a,b);
res=strcmp(a,c);
com=strcmp(a,d);
otr=strcmp(a,e);
noex=strcmp(a,f);
    if ind==1
        industrial(j,:)=basedatos(i,:);
        j=j+1;
    elseif res==1
        residencial(k,:)=basedatos(i,:);
        k=k+1;
    elseif com==1
        comercial(l,:)=basedatos(i,:);
        l=l+1;
    elseif otr==1
        otro(m,:)=basedatos(i,:);
        m=m+1;
    else
        noexiste(n2,:)=basedatos(i,:);
        n2=n2+1;
    end
end

%%Usuarios Industriales
n1=size(industrial);
n=n1(1,1);

j=1;
k=1;
l=1;
m=1;
n2=1;
n3=1;
n4=1;

for i = 1:n
    a=table2array(basedatos(i,10));
    if a==1
        ind_lunes(j,:)=basedatos(i,:);
        j=j+1;
    elseif a==2
        ind_martes(k,:)=basedatos(i,:);
        k=k+1;
    elseif a==3
        ind_miercoles(l,:)=basedatos(i,:);
        l=l+1;
    elseif a==4

```

```

        ind_jueves(m,:)=basedatos(i,:);
        m=m+1;
    elseif a==5
        ind_viernes(n2,:)=basedatos(i,:);
        n2=n2+1;
    elseif a==6
        ind_sabado(n3,:)=basedatos(i,:);
        n3=n3+1;
    else
        ind_domingo(n4,:)=basedatos(i,:);
        n4=n4+1;
    end
end

%%POR DIAS
%%Lunes
ind_lunes_cluster(:,1)=ind_lunes(:,7);
ind_lunes_cluster(:,2)=ind_lunes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_lunes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes';
legend('Cluster Centroid')
curv_ind_lunes=C;
%%Martes
ind_martes_cluster(:,1)=ind_martes(:,7);
ind_martes_cluster(:,2)=ind_martes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_martes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Martes';
legend('Cluster Centroid')

```

```

curv_ind_martes=C;
%%Miercoles
ind_miercoles_cluster(:,1)=ind_miercoles(:,7);
ind_miercoles_cluster(:,2)=ind_miercoles(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_miercoles_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Miercoles';
legend('Cluster Centroid')
curv_ind_miercoles=C;
%%Jueves
ind_jueves_cluster(:,1)=ind_jueves(:,7);
ind_jueves_cluster(:,2)=ind_jueves(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_jueves_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Jueves';
legend('Cluster Centroid')
curv_ind_jueves=C;
%%Viernes
ind_viernes_cluster(:,1)=ind_viernes(:,7);
ind_viernes_cluster(:,2)=ind_viernes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_viernes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on

```

```

plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Viernes';
legend('Cluster Centroid')
curv_ind_viernes=C;
%%Sabado
ind_sabado_cluster(:,1)=ind_sabado(:,7);
ind_sabado_cluster(:,2)=ind_sabado(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_sabado_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Sabado';
legend('Cluster Centroid')
curv_ind_sabado=C;
%%Domingo
ind_domingo_cluster(:,1)=ind_domingo(:,7);
ind_domingo_cluster(:,2)=ind_domingo(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(ind_domingo_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Domingo';
legend('Cluster Centroid')
curv_ind_domingo=C;

%Dias_de_Lunes_a_Viernes
curv_ind_lunvie=[curv_ind_lunes;curv_ind_martes;curv_ind_miercoles;curv_ind_
jueves;curv_ind_viernes];
% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility
Y2=curv_ind_lunvie;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...

```

```

'Replicates',10,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes a Viernes';
legend('Cluster Centroid')
curv_ind_lunesaviernes=C;
%% Curva de Carga Lunes a Viernes
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvalunvie(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)
        curvalunvie(2,:)=C(i,:);
    elseif arreglo(3,1)==C(i,1)
        curvalunvie(3,:)=C(i,:);
    elseif arreglo(4,1)==C(i,1)
        curvalunvie(4,:)=C(i,:);
    elseif arreglo(5,1)==C(i,1)
        curvalunvie(5,:)=C(i,:);
    elseif arreglo(6,1)==C(i,1)
        curvalunvie(6,:)=C(i,:);
    elseif arreglo(7,1)==C(i,1)
        curvalunvie(7,:)=C(i,:);
    elseif arreglo(8,1)==C(i,1)
        curvalunvie(8,:)=C(i,:);
    elseif arreglo(9,1)==C(i,1)
        curvalunvie(9,:)=C(i,:);
    elseif arreglo(10,1)==C(i,1)
        curvalunvie(10,:)=C(i,:);
    end
end
figure
plot(curvalunvie(:,1),curvalunvie(:,2))
title 'Curva de Carga para los días de Lunes a Viernes';
%%Indice MIA y CDI
ini1=sqrt(((C(10,1)-C(9,1))^2)+((C(10,2)-C(9,2))^2));
ini2=sqrt(((C(9,1)-C(8,1))^2)+((C(9,2)-C(8,2))^2));
ini3=sqrt(((C(8,1)-C(7,1))^2)+((C(8,2)-C(7,2))^2));
ini4=sqrt(((curvalunvie(7,1)-curvalunvie(6,1))^2)+((curvalunvie(7,2)-
curvalunvie(6,2))^2));
ini5=sqrt(((C(6,1)-C(5,1))^2)+((C(6,2)-C(5,2))^2));
ini6=sqrt(((C(5,1)-C(4,1))^2)+((C(5,2)-C(4,2))^2));
ini7=sqrt(((C(4,1)-C(3,1))^2)+((C(4,2)-C(3,2))^2));
ini8=sqrt(((C(3,1)-C(2,1))^2)+((C(3,2)-C(2,2))^2));
ini9=sqrt(((C(2,1)-C(1,1))^2)+((C(2,2)-C(1,2))^2));
d2c=ini4;
sumdt=sum(sumd);
Ind_mia_lunvie=sqrt((sumdt)/10);
Ind_cdi_lunvie=(Ind_mia_lunvie/d2c);

```

```

fprintf('El indice MIA es de: %u',Ind_mia_lunvie);
fprintf('El indice CDI es de: %u',Ind_cdi_lunvie);

%Dias_Sabado_y_Domingo
curv_ind_sabdom=[curv_ind_sabado;curv_ind_domingo];
% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility
Y2=curv_ind_sabdom;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Sabado y Domingo';
legend('Cluster Centroid')
curv_ind_sabadoydomingo=C;

%% Curva de Carga Sabado y Domingo
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvasabdom(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)
        curvasabdom(2,:)=C(i,:);
    elseif arreglo(3,1)==C(i,1)
        curvasabdom(3,:)=C(i,:);
    elseif arreglo(4,1)==C(i,1)
        curvasabdom(4,:)=C(i,:);
    elseif arreglo(5,1)==C(i,1)
        curvasabdom(5,:)=C(i,:);
    elseif arreglo(6,1)==C(i,1)
        curvasabdom(6,:)=C(i,:);
    elseif arreglo(7,1)==C(i,1)
        curvasabdom(7,:)=C(i,:);
    elseif arreglo(8,1)==C(i,1)
        curvasabdom(8,:)=C(i,:);
    elseif arreglo(9,1)==C(i,1)
        curvasabdom(9,:)=C(i,:);
    elseif arreglo(10,1)==C(i,1)
        curvasabdom(10,:)=C(i,:);
    end
end
figure
plot(curvasabdom(:,1),curvasabdom(:,2))
title 'Curva de Carga para los dias Sabado y Domingo';

%%Indice MIA y CDI

```

```

ini1=sqrt(((C(10,1)-C(9,1))^2)+((C(10,2)-C(9,2))^2));
ini2=sqrt(((C(9,1)-C(8,1))^2)+((C(9,2)-C(8,2))^2));
ini3=sqrt(((C(8,1)-C(7,1))^2)+((C(8,2)-C(7,2))^2));
ini4=sqrt(((curvasabdom(7,1)-curvasabdom(6,1))^2)+((curvasabdom(7,2)-
curvasabdom(6,2))^2));
ini5=sqrt(((C(6,1)-C(5,1))^2)+((C(6,2)-C(5,2))^2));
ini6=sqrt(((C(5,1)-C(4,1))^2)+((C(5,2)-C(4,2))^2));
ini7=sqrt(((C(4,1)-C(3,1))^2)+((C(4,2)-C(3,2))^2));
ini8=sqrt(((C(3,1)-C(2,1))^2)+((C(3,2)-C(2,2))^2));
ini9=sqrt(((C(2,1)-C(1,1))^2)+((C(2,2)-C(1,2))^2));
d2c=ini4;
sumdt=sum(sumd);
Ind_mia_lunvie=sqrt((sumdt)/10);
Ind_cdi_lunvie=(Ind_mia_lunvie/d2c);
fprintf('El indice MIA es de: %u',Ind_mia_lunvie);
fprintf('El indice CDI es de: %u',Ind_cdi_lunvie);

%%Usuarios Comerciales
fprintf('2.Clientes de tipo Comercial\n');
n1=size(comercial);
n=n1(1,1);

j=1;
k=1;
l=1;
m=1;
n2=1;
n3=1;
n4=1;

for i = 1:n
    a=table2array(basedatos(i,10));
    if a==1
        com_lunes(j,:)=basedatos(i,:);
        j=j+1;
    elseif a==2
        com_martes(k,:)=basedatos(i,:);
        k=k+1;
    elseif a==3
        com_miercoles(l,:)=basedatos(i,:);
        l=l+1;
    elseif a==4
        com_jueves(m,:)=basedatos(i,:);
        m=m+1;
    elseif a==5
        com_viernes(n2,:)=basedatos(i,:);
        n2=n2+1;
    elseif a==6
        com_sabado(n3,:)=basedatos(i,:);
        n3=n3+1;

```



```

        else
            com_domingo(n4,:)=basedatos(i,:);
            n4=n4+1;
        end
    end
end

%%POR DIAS
%%Lunes
com_lunes_cluster(:,1)=com_lunes(:,7);
com_lunes_cluster(:,2)=com_lunes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(com_lunes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes';
legend('Cluster Centroid')
curv_com_lunes=C;
%%Martes
com_martes_cluster(:,1)=com_martes(:,7);
com_martes_cluster(:,2)=com_martes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(com_martes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Martes';
legend('Cluster Centroid')
curv_com_martes=C;
%%Miercoles
com_miercoles_cluster(:,1)=com_miercoles(:,7);
com_miercoles_cluster(:,2)=com_miercoles(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility

```

```

Y2=table2array(com_miercoles_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Miercoles';
legend('Cluster Centroid')
curv_com_miercoles=C;
%%Jueves
com_jueves_cluster(:,1)=com_jueves(:,7);
com_jueves_cluster(:,2)=com_jueves(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(com_jueves_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Jueves';
legend('Cluster Centroid')
curv_com_jueves=C;
%%Viernes
com_viernes_cluster(:,1)=com_viernes(:,7);
com_viernes_cluster(:,2)=com_viernes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(com_viernes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Viernes';
legend('Cluster Centroid')
curv_com_viernes=C;
%%Sabado
com_sabado_cluster(:,1)=com_sabado(:,7);
com_sabado_cluster(:,2)=com_sabado(:,9);

```

```

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(com_sabado_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Sabado';
legend('Cluster Centroid')
curv_com_sabado=C;
%%Domingo
com_domingo_cluster(:,1)=com_domingo(:,7);
com_domingo_cluster(:,2)=com_domingo(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(com_domingo_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Domingo';
legend('Cluster Centroid')
curv_com_domingo=C;
%D1A2Y3E4L5I6N--T7E--A8M90--G1R2A3C4I5A6S-P102R-T102D30***
%Dias_de_Lunes_a_Viernes
curv_com_lunvie=[curv_com_lunes;curv_com_martes;curv_com_miercoles;curv_com_
jueves;curv_com_viernes];
% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility
Y2=curv_com_lunvie;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes a Viernes';
legend('Cluster Centroid')
curv_com_lunesaviernes=C;

```

```

%% Curva de Carga Lunes a Viernes
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvalunvie(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)
        curvalunvie(2,:)=C(i,:);
    elseif arreglo(3,1)==C(i,1)
        curvalunvie(3,:)=C(i,:);
    elseif arreglo(4,1)==C(i,1)
        curvalunvie(4,:)=C(i,:);
    elseif arreglo(5,1)==C(i,1)
        curvalunvie(5,:)=C(i,:);
    elseif arreglo(6,1)==C(i,1)
        curvalunvie(6,:)=C(i,:);
    elseif arreglo(7,1)==C(i,1)
        curvalunvie(7,:)=C(i,:);
    elseif arreglo(8,1)==C(i,1)
        curvalunvie(8,:)=C(i,:);
    elseif arreglo(9,1)==C(i,1)
        curvalunvie(9,:)=C(i,:);
    elseif arreglo(10,1)==C(i,1)
        curvalunvie(10,:)=C(i,:);
    end
end
figure
plot(curvalunvie(:,1),curvalunvie(:,2))
title 'Curva de Carga para los dias de Lunes a Viernes';

%%Indice MIA y CDI
ini1=sqrt(((C(10,1)-C(9,1))^2)+((C(10,2)-C(9,2))^2));
ini2=sqrt(((C(9,1)-C(8,1))^2)+((C(9,2)-C(8,2))^2));
ini3=sqrt(((C(8,1)-C(7,1))^2)+((C(8,2)-C(7,2))^2));
ini4=sqrt(((curvalunvie(7,1)-curvalunvie(6,1))^2)+((curvalunvie(7,2)-
curvalunvie(6,2))^2));
ini5=sqrt(((C(6,1)-C(5,1))^2)+((C(6,2)-C(5,2))^2));
ini6=sqrt(((C(5,1)-C(4,1))^2)+((C(5,2)-C(4,2))^2));
ini7=sqrt(((C(4,1)-C(3,1))^2)+((C(4,2)-C(3,2))^2));
ini8=sqrt(((C(3,1)-C(2,1))^2)+((C(3,2)-C(2,2))^2));
ini9=sqrt(((C(2,1)-C(1,1))^2)+((C(2,2)-C(1,2))^2));
d2c=ini4;
sumdt=sum(sumd);
Ind_mia_lunvie=sqrt((sumdt)/10);
Ind_cdi_lunvie=(Ind_mia_lunvie/d2c);
fprintf('El indice MIA es de: %u',Ind_mia_lunvie);
fprintf('El indice CDI es de: %u',Ind_cdi_lunvie);

%Dias_Sabado_y_Domingo
curv_com_sabdom=[curv_com_sabado;curv_com_domingo];
% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility

```

```

Y2=curv_com_sabdom;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Sabado y Domingo';
legend('Cluster Centroid')
curv_com_sabadoydomingo=C;

%% Curva de Carga Sabado y Domingo
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvasabdom(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)
        curvasabdom(2,:)=C(i,:);
    elseif arreglo(3,1)==C(i,1)
        curvasabdom(3,:)=C(i,:);
    elseif arreglo(4,1)==C(i,1)
        curvasabdom(4,:)=C(i,:);
    elseif arreglo(5,1)==C(i,1)
        curvasabdom(5,:)=C(i,:);
    elseif arreglo(6,1)==C(i,1)
        curvasabdom(6,:)=C(i,:);
    elseif arreglo(7,1)==C(i,1)
        curvasabdom(7,:)=C(i,:);
    elseif arreglo(8,1)==C(i,1)
        curvasabdom(8,:)=C(i,:);
    elseif arreglo(9,1)==C(i,1)
        curvasabdom(9,:)=C(i,:);
    elseif arreglo(10,1)==C(i,1)
        curvasabdom(10,:)=C(i,:);
    end
end
figure
plot(curvasabdom(:,1),curvasabdom(:,2))
title 'Curva de Carga para los dias Sabado y Domingo';

%%Indice MIA y CDI
ini1=sqrt(((C(10,1)-C(9,1))^2)+((C(10,2)-C(9,2))^2));
ini2=sqrt(((C(9,1)-C(8,1))^2)+((C(9,2)-C(8,2))^2));
ini3=sqrt(((C(8,1)-C(7,1))^2)+((C(8,2)-C(7,2))^2));
ini4=sqrt(((curvasabdom(7,1)-curvasabdom(6,1))^2)+((curvasabdom(7,2)-
curvasabdom(6,2))^2));
ini5=sqrt(((C(6,1)-C(5,1))^2)+((C(6,2)-C(5,2))^2));
ini6=sqrt(((C(5,1)-C(4,1))^2)+((C(5,2)-C(4,2))^2));
ini7=sqrt(((C(4,1)-C(3,1))^2)+((C(4,2)-C(3,2))^2));

```

```

ini8=sqrt(((C(3,1)-C(2,1))^2)+((C(3,2)-C(2,2))^2));
ini9=sqrt(((C(2,1)-C(1,1))^2)+((C(2,2)-C(1,2))^2));
d2c=ini4;
sumdt=sum(sumd);
Ind_mia_lunvie=sqrt((sumdt)/10);
Ind_cdi_lunvie=(Ind_mia_lunvie/d2c);
fprintf('El indice MIA es de: %u',Ind_mia_lunvie);
fprintf('El indice CDI es de: %u',Ind_cdi_lunvie);

```

```

%%Usuarios Residenciales

```

```

fprintf('3.Clientes de tipo Residencial\n');
n1=size(residencial);
n=n1(1,1);

```

```

j=1;
k=1;
l=1;
m=1;
n2=1;
n3=1;
n4=1;

```

```

for i = 1:n
    a=table2array(basedatos(i,10));
    if a==1
        res_lunes(j,:)=basedatos(i,:);
        j=j+1;
    elseif a==2
        res_martes(k,:)=basedatos(i,:);
        k=k+1;
    elseif a==3
        res_miercoles(l,:)=basedatos(i,:);
        l=l+1;
    elseif a==4
        res_jueves(m,:)=basedatos(i,:);
        m=m+1;
    elseif a==5
        res_viernes(n2,:)=basedatos(i,:);
        n2=n2+1;
    elseif a==6
        res_sabado(n3,:)=basedatos(i,:);
        n3=n3+1;
    else
        res_domingo(n4,:)=basedatos(i,:);
        n4=n4+1;
    end
end
end

```

```

%%POR DIAS
%%Lunes
res_lunes_cluster(:,1)=res_lunes(:,7);
res_lunes_cluster(:,2)=res_lunes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_lunes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes';
legend('Cluster Centroid')
curv_res_lunes=C;
%%Martes
res_martes_cluster(:,1)=res_martes(:,7);
res_martes_cluster(:,2)=res_martes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_martes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
    'MarkerSize',15,'LineWidth',3)
title 'Datos Martes';
legend('Cluster Centroid')
curv_res_martes=C;
%%Miercoles
res_miercoles_cluster(:,1)=res_miercoles(:,7);
res_miercoles_cluster(:,2)=res_miercoles(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_miercoles_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on

```

```

plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Miercoles';
legend('Cluster Centroid')
curv_res_miercoles=C;
%%Jueves
res_jueves_cluster(:,1)=res_jueves(:,7);
res_jueves_cluster(:,2)=res_jueves(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_jueves_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Jueves';
legend('Cluster Centroid')
curv_res_jueves=C;
%%Viernes
res_viernes_cluster(:,1)=res_viernes(:,7);
res_viernes_cluster(:,2)=res_viernes(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_viernes_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Viernes';
legend('Cluster Centroid')
curv_res_viernes=C;
%%Sabado
res_sabado_cluster(:,1)=res_sabado(:,7);
res_sabado_cluster(:,2)=res_sabado(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_sabado_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...

```



```

    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Sabado';
legend('Cluster Centroid')
curv_res_sabado=C;
%%Domingo
res_domingo_cluster(:,1)=res_domingo(:,7);
res_domingo_cluster(:,2)=res_domingo(:,9);

% Algoritmo K-Means 20 centroides
rng('default') % For reproducibility
Y2=table2array(res_domingo_cluster);
opts = statset('Display','final');
[idx,C] = kmeans(Y2,20,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Domingo';
legend('Cluster Centroid')
curv_res_domingo=C;

%Dias_de_Lunes_a_Viernes
curv_res_lunvie=[curv_res_lunes;curv_res_martes;curv_res_miercoles;curv_res_
jueves;curv_res_viernes];
% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility
Y2=curv_res_lunvie;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')
hold on
plot(C(:,1),C(:,2),'kx',...
     'MarkerSize',15,'LineWidth',3)
title 'Datos Lunes a Viernes';
legend('Cluster Centroid')
curv_res_lunesaviernes=C;
%% Curva de Carga Lunes a Viernes
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvalunvie(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)

```

```

        curvalunvie(2,:)=C(i,:);
elseif arreglo(3,1)==C(i,1)
        curvalunvie(3,:)=C(i,:);
elseif arreglo(4,1)==C(i,1)
        curvalunvie(4,:)=C(i,:);
elseif arreglo(5,1)==C(i,1)
        curvalunvie(5,:)=C(i,:);
elseif arreglo(6,1)==C(i,1)
        curvalunvie(6,:)=C(i,:);
elseif arreglo(7,1)==C(i,1)
        curvalunvie(7,:)=C(i,:);
elseif arreglo(8,1)==C(i,1)
        curvalunvie(8,:)=C(i,:);
elseif arreglo(9,1)==C(i,1)
        curvalunvie(9,:)=C(i,:);
elseif arreglo(10,1)==C(i,1)
        curvalunvie(10,:)=C(i,:);
end
end
figure
plot(curvalunvie(:,1),curvalunvie(:,2))
title 'Curva de Carga para los dias de Lunes a Viernes';

%%Indice MIA y CDI
ini1=sqrt(((C(10,1)-C(9,1))^2)+((C(10,2)-C(9,2))^2));
ini2=sqrt(((C(9,1)-C(8,1))^2)+((C(9,2)-C(8,2))^2));
ini3=sqrt(((C(8,1)-C(7,1))^2)+((C(8,2)-C(7,2))^2));
ini4=sqrt(((curvalunvie(7,1)-curvalunvie(6,1))^2)+((curvalunvie(7,2)-
curvalunvie(6,2))^2));
ini5=sqrt(((C(6,1)-C(5,1))^2)+((C(6,2)-C(5,2))^2));
ini6=sqrt(((C(5,1)-C(4,1))^2)+((C(5,2)-C(4,2))^2));
ini7=sqrt(((C(4,1)-C(3,1))^2)+((C(4,2)-C(3,2))^2));
ini8=sqrt(((C(3,1)-C(2,1))^2)+((C(3,2)-C(2,2))^2));
ini9=sqrt(((C(2,1)-C(1,1))^2)+((C(2,2)-C(1,2))^2));
d2c=ini4;
sumdt=sum(sumd);
Ind_mia_lunvie=sqrt((sumdt)/10);
Ind_cdi_lunvie=(Ind_mia_lunvie/d2c);
fprintf('El indice MIA es de: %u',Ind_mia_lunvie);
fprintf('El indice CDI es de: %u',Ind_cdi_lunvie);

%Dias_Sabado_y_Domingo
curv_res_sabdom=[curv_res_sabado;curv_res_domingo];
% Algoritmo K-Means 15 centroides
rng('default') % For reproducibility
Y2=curv_res_sabdom;
opts = statset('Display','final');
[idx,C,sumd,D] = kmeans(Y2,10,'Distance','cityblock',...
'Replicates',5,'Options',opts);
figure
gscatter(Y2(:,1),Y2(:,2),idx,'bgm')

```

```

hold on
plot(C(:,1),C(:,2),'kx',...
      'MarkerSize',15,'LineWidth',3)
title 'Datos Sabado y Domingo';
legend('Cluster Centroid')
curv_res_sabadoydomingo=C;
%% Curva de Carga Sabado y Domingo
arreglo=sort(C(:,1));
for i = 1:10
    if arreglo(1,1)==C(i,1)
        curvasabdom(1,:)=C(i,:);
    elseif arreglo(2,1)==C(i,1)
        curvasabdom(2,:)=C(i,:);
    elseif arreglo(3,1)==C(i,1)
        curvasabdom(3,:)=C(i,:);
    elseif arreglo(4,1)==C(i,1)
        curvasabdom(4,:)=C(i,:);
    elseif arreglo(5,1)==C(i,1)
        curvasabdom(5,:)=C(i,:);
    elseif arreglo(6,1)==C(i,1)
        curvasabdom(6,:)=C(i,:);
    elseif arreglo(7,1)==C(i,1)
        curvasabdom(7,:)=C(i,:);
    elseif arreglo(8,1)==C(i,1)
        curvasabdom(8,:)=C(i,:);
    elseif arreglo(9,1)==C(i,1)
        curvasabdom(9,:)=C(i,:);
    elseif arreglo(10,1)==C(i,1)
        curvasabdom(10,:)=C(i,:);
    end
end
figure
plot(curvasabdom(:,1),curvasabdom(:,2))
title 'Curva de Carga para los dias Sabado y Domingo';
%%Indice MIA y CDI
ini1=sqrt(((C(10,1)-C(9,1))^2)+((C(10,2)-C(9,2))^2));
ini2=sqrt(((C(9,1)-C(8,1))^2)+((C(9,2)-C(8,2))^2));
ini3=sqrt(((C(8,1)-C(7,1))^2)+((C(8,2)-C(7,2))^2));
ini4=sqrt(((curvasabdom(7,1)-curvasabdom(6,1))^2)+((curvasabdom(7,2)-
curvasabdom(6,2))^2));
ini5=sqrt(((C(6,1)-C(5,1))^2)+((C(6,2)-C(5,2))^2));
ini6=sqrt(((C(5,1)-C(4,1))^2)+((C(5,2)-C(4,2))^2));
ini7=sqrt(((C(4,1)-C(3,1))^2)+((C(4,2)-C(3,2))^2));
ini8=sqrt(((C(3,1)-C(2,1))^2)+((C(3,2)-C(2,2))^2));
ini9=sqrt(((C(2,1)-C(1,1))^2)+((C(2,2)-C(1,2))^2));
d2c=ini4;
sumdt=sum(sumd);
Ind_mia_lunvie=sqrt((sumdt)/10);
Ind_cdi_lunvie=(Ind_mia_lunvie/d2c);
fprintf('El indice MIA es de: %u',Ind_mia_lunvie);
fprintf('El indice CDI es de: %u',Ind_cdi_lunvie);

```