 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	---	--

CIENCIA DE DATOS APLICADA



Entrega No. 1 Proyecto CDA

INTEGRANTES

Nombre	Correo Electrónico
Juliana Andrea Galeano Caicedo	ja.galeanoc1@uniandes.edu.co
Juan Nicolás Estepa Guzmán	j.estepa@uniandes.edu.co
Santiago Rodríguez Cruz	s.rodriquez52@uniandes.edu.co
Harvy Benítez Amaya	h.benitez@uniandes.edu.co

INGENIERÍA DE SISTEMAS Y COMPUTACIÓN
 MAESTRÍA EN INGENIERÍA DE LA INFORMACIÓN
 UNIVERSIDAD DE LOS ANDES
 BOGOTÁ D. C.



 <p>Universidad de los Andes Facultad de Ingeniería</p>	<p>Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401</p>	<p>Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015</p>
---	--	---

TABLA DE CONTENIDO

1.	DEFINICIÓN DE LA PROBLEMÁTICA Y ENTENDIMIENTO DEL NEGOCIO	3
1.1.1	Objetivo del proyecto	4
1.1.2	Métricas de negocio	4
2.	IDEACIÓN	5
3.	<i>RESPONSIBLE</i>	6
4.	ENFOQUE ANALÍTICO	6
4.1.1	Pregunta de negocio	6
4.1.2	Técnicas propuestas	7
4.1.3	Métricas de evaluación	7
5.	RECOLECCIÓN DE DATOS	8
6.	ENTENDIMIENTO DE LOS DATOS	9
7.	CONCLUSIONES INICIALES.....	12

 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	--	--


1.DEFINICIÓN DE LA PROBLEMÁTICA Y ENTENDIMIENTO DEL NEGOCIO

El Área de Gobierno y Monitoreo del Dato del Ministerio de Minas y Energía (MinEnergía), junto con el Área de Gestión de la Información de la Unidad de Planeación Minero-Energética (UPME), enfrenta una problemática significativa relacionada con las limitaciones de los modelos actuales para realizar pronósticos precisos de demanda energética a corto plazo (1, 3 y 6 meses) desagregados a nivel regional. Estas limitaciones resultan en sobrecostos debido a la subestimación y sobreestimación de la demanda, impactando negativamente la planeación y operación del sistema energético. Sin embargo, existe una oportunidad clara para mejorar esta situación mediante el uso de Ciencia de Datos, que podría facilitar pronósticos más precisos al incorporar variables exógenas y manejar eventos atípicos, optimizando así la gestión energética y reduciendo costos.

La estrategia del MinEnergía y UPME para pronosticar la demanda de energía en Colombia incluye el uso de modelos estadísticos y econométricos que analizan datos históricos y factores económicos como el PIB y la población. Se desarrollan diferentes escenarios de crecimiento, se consulta a actores clave del sector energético, y se realizan revisiones periódicas de los pronósticos para ajustarlos a cambios en las condiciones del país. Además, se incorporan tecnologías emergentes y se realizan análisis regionales para entender mejor las dinámicas locales, lo que permite una planificación más precisa del sistema energético.

El Sector Eléctrico Colombiano se rige fundamentalmente por las Leyes 142 y 143 de 1994, las cuales, definen el régimen de prestación de servicios públicos desde las actividades de generación interconexión, transmisión, distribución y comercialización de energía eléctrica. A su vez, se definen las funciones de las principales entidades del sector y su estructura y gobernanza. La estructura institucional del sector eléctrico colombiano se basa en varias entidades clave que desempeñan funciones específicas: el Ministerio de Minas y Energía (MinEnergía), que formula políticas y regula el sector; la Unidad de Planeación Minero Energética (UPME), encargada de la planificación y proyección de la demanda y oferta de energía; la Comisión de Regulación de Energía y Gas (CREG), que regula las tarifas y las condiciones de acceso al sistema; el Operador del Sistema (XM), responsable de la operación del sistema eléctrico y la gestión del mercado mayorista; y las Empresas de Servicios Públicos, que generan, transmiten, distribuyen y comercializan la energía eléctrica. Esta estructura busca garantizar un suministro eficiente y sostenible de energía, promover la competencia en el mercado y asegurar el acceso universal a servicios de energía.

Si bien el MinEnergía encabeza la dirección del sector eléctrico colombiano y la UPME es la entidad encargada de la planeación, corresponde a los operadores del sistema (XM

 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	--	--

/ Expertos en mercado) recolectar y generar datos para predecir la demanda de energía eléctrica en Colombia a través de un sistema integral que incluye el monitoreo en tiempo real de la generación y el consumo de electricidad en todo el país.

1.1.1 Objetivo del proyecto

Desarrollar un modelo analítico que utilice datos históricos y variables exógenas para predecir la demanda de energía a corto y mediano plazo en Colombia, mejorando la precisión de los pronósticos actuales.

1.1.2 Métricas de negocio


Mejorar la precisión de los pronósticos actuales: XM se enfoca en calibrar constantemente sus modelos de pronóstico mediante el uso de datos actualizados y técnicas avanzadas de análisis. Esto permite identificar errores y ajustar los modelos para reflejar con mayor exactitud las dinámicas del consumo eléctrico en Colombia.

***Nota:** Es importante aclarar que, en este momento no se tiene acceso a los modelos de pronóstico actuales utilizados por XM ni a información relacionada con la precisión de los mismos. Sin embargo, el intermediario del MinEnergía aún se encuentra gestionando el acceso a esta información, por lo tanto, se ha decidido establecer como métrica de negocio principal tomar la tolerancia de error de pronóstico establecida por el CNO, que es un tema netamente regulatorio y se desconoce su cumplimiento en los modelos actuales.*

Tolerancia de error de pronóstico $\leq 5\%$: El Consejo Nacional de Operación (CNO), conformado por generadores, transmisores y distribuidores de energía, supervisa las desviaciones de pronóstico de los operadores de red y establece una tolerancia de error de pronóstico $\leq 5\%$, de lo contrario, estas deben ser reportadas a la Superintendencia de Servicios Públicos, puesto que, generan costos adicionales para el sistema.

Reducción de costos operativos: La mejora en la precisión de los pronósticos contribuye a la reducción de costos operativos al minimizar pérdidas por sobreproducción o falta de suministro. Esto optimiza la generación y distribución de energía, permitiendo reinvertir recursos en infraestructura y tecnología.

Aumento de la confiabilidad del sector eléctrico: Un sistema que puede anticipar las necesidades de consumo con precisión es menos propenso a enfrentar apagones o déficit en la oferta, lo que genera mayor confianza entre los usuarios y el sector empresarial. Además, una operación estable y confiable del sector eléctrico contribuye a la seguridad energética del país que es esencial para el desarrollo económico y social y, que su vez, atrae inversiones y promueve el uso eficiente de la energía, beneficiando a todos los actores involucrados.

 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	--	---

2. IDEACIÓN

El producto consiste en un modelo analítico predictivo que se ofrece a través de una API. Como corresponde a los operadores del sistema (XM / Expertos en mercado) recolectar y generar datos para predecir la demanda de energía eléctrica en Colombia, estos podrían ser considerados como los “usuarios finales” del modelo. Sin embargo, es necesario aclarar que los datos que generará el modelo y que por cuestiones regulatorias deben ser de acceso público (*posiblemente a través del portal web <https://www.integrarme.gov.co>*), pueden ser utilizados por diversos actores del sector energético, incluidos el MInEnergía, empresas de generación, transmisión y distribución, así como las entidades reguladoras CREG y UPME.

Ilustración 1. Diagrama de componentes




Fuente: Elaboración propia.

Actualmente, el modelo se basa en técnicas de series de tiempo, específicamente regresión lineal multivariada, lo que permite realizar pronósticos sobre la demanda energética. Sin embargo, el enfoque presenta varios desafíos:

- Las limitaciones en los pronósticos a corto plazo dificultan la capacidad de respuesta ante cambios repentinos en la demanda
- La baja granularidad de los modelos impide obtener pronósticos precisos a nivel regional o por sectores específicos
- La complejidad en el manejo de variables exógenas, como factores económicos y climáticos, complica la calibración del modelo.

Estos "dolores" resaltan la necesidad de un enfoque más avanzado y flexible que pueda mejorar la precisión y utilidad de las predicciones en el sector energético.

El desarrollo del nuevo modelo analítico predictivo sigue un proceso estructurado que abarca varios requerimientos esenciales. En primer lugar, se inicia con la recopilación de datos históricos, que son fundamentales para entender patrones y tendencias en la

 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	---	--

demanda energética. Luego, se procede al procesamiento de estos datos, seguido de su división en conjuntos que permitirán la construcción y entrenamiento del modelo. Este proceso incluye la validación del modelo para asegurar que cumpla con los estándares necesarios antes de pasar a la evaluación final, donde se determinará su efectividad y precisión.

Para la implementación del modelo, se requieren diversos componentes tecnológicos. Esto incluye un servidor dedicado para el entrenamiento del modelo, que garantice el rendimiento adecuado durante el procesamiento de grandes volúmenes de datos. Además, se necesitará un entorno de desarrollo que facilite la creación y ajuste del modelo, así como una interfaz de programación de aplicaciones (API) que permita su integración y acceso por parte de los usuarios. Por último, se proporcionará documentación detallada para facilitar el uso e implementación del modelo, asegurando que todos los actores involucrados puedan beneficiarse de sus capacidades analíticas de manera efectiva.


3. RESPONSIBLE

El modelo analítico predictivo para el sector energético se basa en datos que son de dominio público y que provienen de entidades reguladas con la legislación vigente en Colombia. No existen temas regulatorios adicionales que limiten la privacidad de estas fuentes de datos, ya que, están respaldadas por el marco normativo vigente y que está diseñado para fomentar la transparencia y el acceso a la información. En este contexto, el uso de datos de dominio público para el modelo analítico no solo es legal, sino que también se alinea con los principios de transparencia y acceso a la información que rigen en Colombia. Esto garantiza que el modelo se construya sobre bases sólidas, utilizando información que está disponible para todos y que puede contribuir significativamente a la toma de decisiones informadas en el sector energético, sin vulnerar la privacidad de los datos personales.

4. ENFOQUE ANALÍTICO

4.1.1 Pregunta de negocio

¿Cómo podemos utilizar técnicas avanzadas de analítica y Machine Learning para generar pronósticos más precisos y granulares de la demanda de energía eléctrica en Colombia para horizontes de 1, 3 y 6 meses, aprovechando datos históricos de consumo y variables exógenas relevantes, con el fin de optimizar la toma de decisiones en procesos claves como la planeación de la generación, la operación del sistema interconectado y la gestión comercial de las empresas de energía?

 <p>Universidad de los Andes Facultad de Ingeniería</p>	<p>Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401</p>	<p>Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015</p>
---	--	---

4.1.2 Técnicas propuestas

La propuesta del modelo analítico predictivo se centra en el uso de redes neuronales recurrentes LSTM (Long Short-Term Memory), pertenecientes a la categoría de deep learning. Las LSTM son especialmente efectivas para la predicción de series temporales, ya que pueden capturar patrones temporales complejos y mantener información relevante a lo largo del tiempo, gracias a su arquitectura que permite regular el flujo de información mediante mecanismos de "puertas".

Además, las LSTM se adaptan a cambios en los patrones de datos y manejan de manera efectiva información ruidosa, lo cual es crucial en un entorno dinámico como el consumo energético. Esta capacidad no solo mejora la precisión de los pronósticos a corto plazo, sino que también proporciona una mayor robustez frente a la variabilidad de los datos, lo que resulta en predicciones más fiables y útiles para predecir la demanda de energía diaria en un horizonte de 30 días y, de manera recursiva ofrecer datos de demanda diaria en intervalos de 1, 3 y 6 meses, desagregados por región.

4.1.3 Métricas de evaluación

Error Absoluto Medio (MAE): Mide el error promedio en las predicciones, indicando cuán cerca están las estimaciones de los valores reales. Un MAE bajo sugiere alta precisión en las predicciones.

Raíz del Error Cuadrático Medio (RMSE): Identifica desviaciones significativas en las predicciones al penalizar errores grandes, lo que ayuda a resaltar fallos críticos en el modelo.

Porcentaje de Error Absoluto Medio (MAPE): Evalúa la precisión de las predicciones en términos porcentuales, permitiendo comparaciones entre diferentes períodos y regiones.

Coefficiente de Determinación (R^2): Evalúa la proporción de la varianza en la demanda que es explicada por el modelo; un valor cercano a 1 indica un buen ajuste.

Error Cuadrático Medio (MSE): Similar al RMSE, mide el promedio de los errores al cuadrado, proporcionando una visión detallada de la distribución de los errores.

Gráficos de Residuos: Analizan los residuos del modelo para detectar patrones; un patrón aleatorio sugiere un buen ajuste, mientras que patrones sistemáticos indican la necesidad de ajustes.

Pruebas de Estacionariedad: Verifican si la serie temporal es estacionaria, lo cual es crucial para la precisión del modelo; la no estacionariedad puede afectar negativamente las predicciones.

5. RECOLECCIÓN DE DATOS

A continuación, se presentan las fuentes de datos seleccionadas y se describe su utilidad en el modelo analítico predictivo de la demanda de energía.

Tabla 1. Fuentes de datos

VARIABLE OBJETIVO				
NOMBRE	FUENTE	COMPONENTE	TIPO DE DATO	DESCRIPCIÓN
Demanda de energía	XM (Operador del Sistema / Expertos en Mercado)	Energético	Datos estructurados	Un (1) dato por día por región.

VARIABLES INDEPENDIENTES				
NOMBRE	FUENTE	COMPONENTE	TIPO DE DATO	DESCRIPCIÓN
Proyección de la población	DANE	Demográfica	Datos estructurados	Un (1) dato anual por ciudad.
IPC	DANE	Económica	Datos estructurados	Un (1) dato mensual por ciudad.
PIB	DANE	Económica	Datos estructurados	Un (1) dato anual por ciudad.
Temperatura seca	IDEAM	Climática	Datos estructurados	Temperatura media diaria por estación de monitoreo.
Calendario festivos Colombia	Nager.Date	Calendario	Datos estructurados	Días festivos y fechas especiales que alteren la demanda de forma temporal.

Fuente: Elaboración propia.


Demanda de energía: Esta variable proporciona datos históricos de consumo que reflejan patrones y tendencias en el uso de energía. Analizarla permite ajustar el modelo a comportamientos pasados y prever cambios futuros en la demanda.

Proyección de la población: La población es un factor crítico que impacta directamente en la demanda energética. A medida que la población crece, también lo hace el consumo de energía, especialmente en sectores residenciales y comerciales. Incluir esta variable permite captar cómo los cambios demográficos influyen en la demanda.

Índice de Precios al Consumidor (IPC): El IPC refleja la inflación y el costo de la vida, lo que puede afectar la capacidad de los consumidores para gastar en energía. Un aumento en el IPC podría llevar a una disminución en el consumo energético, especialmente en segmentos vulnerables. Esta variable ayuda a entender la relación entre la economía y la demanda.

Producto Interno Bruto (PIB): El PIB es un indicador clave del crecimiento económico. Un PIB en crecimiento suele correlacionarse con un aumento en la actividad industrial y comercial, lo que generalmente eleva la demanda de energía. Incorporar el PIB permite capturar la dinámica entre el crecimiento económico y el consumo energético.

Temperatura seca: Las variaciones en la temperatura afectan significativamente la demanda energética, especialmente en climas donde la calefacción y la refrigeración son

 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	---	--

necesarias. Incorporar datos climáticos ayuda a anticipar picos en la demanda durante períodos de temperaturas extremas.

Calendarios festivos Colombia: Los días festivos pueden influir en los patrones de consumo, ya que durante estos períodos la actividad comercial y el comportamiento del consumidor pueden cambiar notablemente. Tener en cuenta estos días permite ajustar las proyecciones de demanda para reflejar estas variaciones estacionales.

6. ENTENDIMIENTO DE LOS DATOS

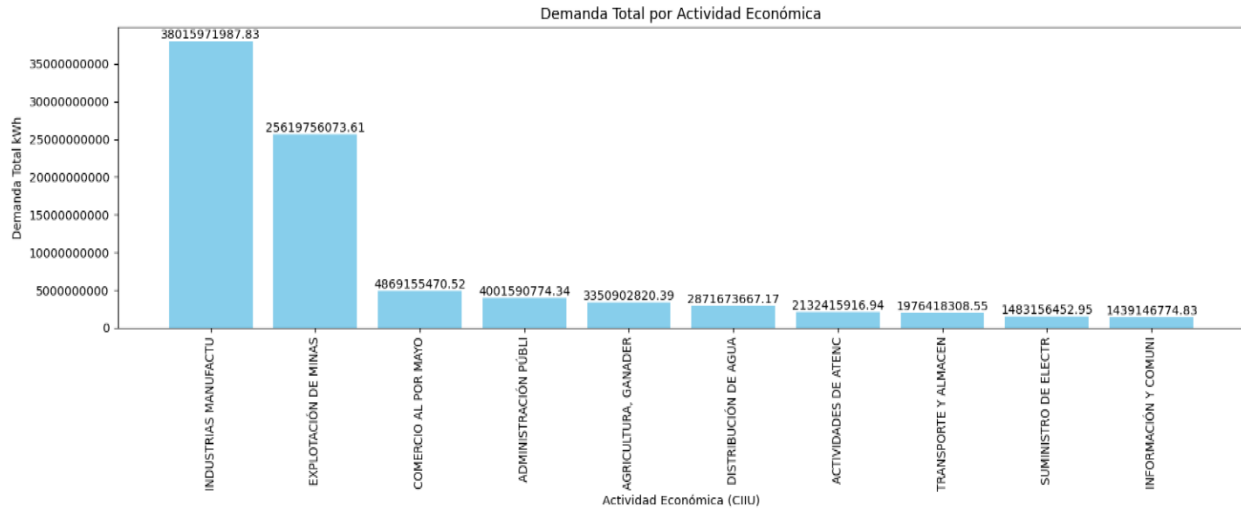
Se realizó análisis no gráfico univariado para cada una de las variables (*variable objeto y variables independientes*) revisando las estadísticas como la desviación estándar, media, máximos y mínimos, ect. Así mismo, se hizo análisis no gráfico bivariado para el entender el comportamiento de la variable objetivo en relación con cada una de las variables independientes, utilizando las técnicas estadísticas mencionadas con anterioridad.

Por otra parte, se realizó análisis gráfico univariado y bivariado para comprender el comportamiento de cada una de las variables y, la relación entre la variable objetivo y las variables independientes, respectivamente. Haciendo uso de diagramas de barra, box plot, gráficas de líneas, etc.

Adicionalmente, se ejecutaron técnicas de análisis de calidad para los datos de cada variable, tales como, validación de valores nulos, transformaciones de cada columna, sumar los valores de demanda horaria para obtener un dato de demanda por día, mapear las regiones del DANE con respecto a las usadas por XM para realizar análisis, etc.

A continuación, se presentan algunos gráficos relacionados con la variable demanda de energía, si se desea conocer un análisis más técnico y detallado de cada una de las variables, el lector puede remitirse al siguiente link:

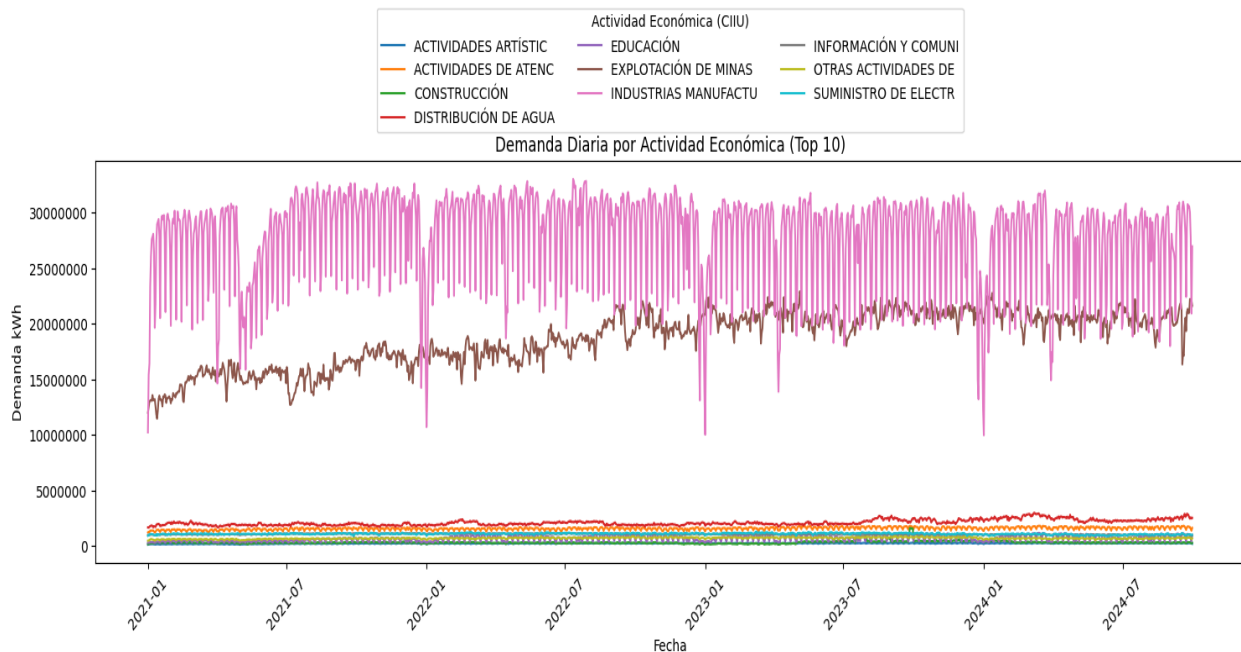
Ilustración 2. Demanda total por actividad económica



Fuente: Elaboración propia.

De acuerdo con la *Ilustración 2*, se puede observar que las actividades económicas que generan mayor demanda de energía en el país son las Industrias Manufactureras y Explotación de minas y canteras.

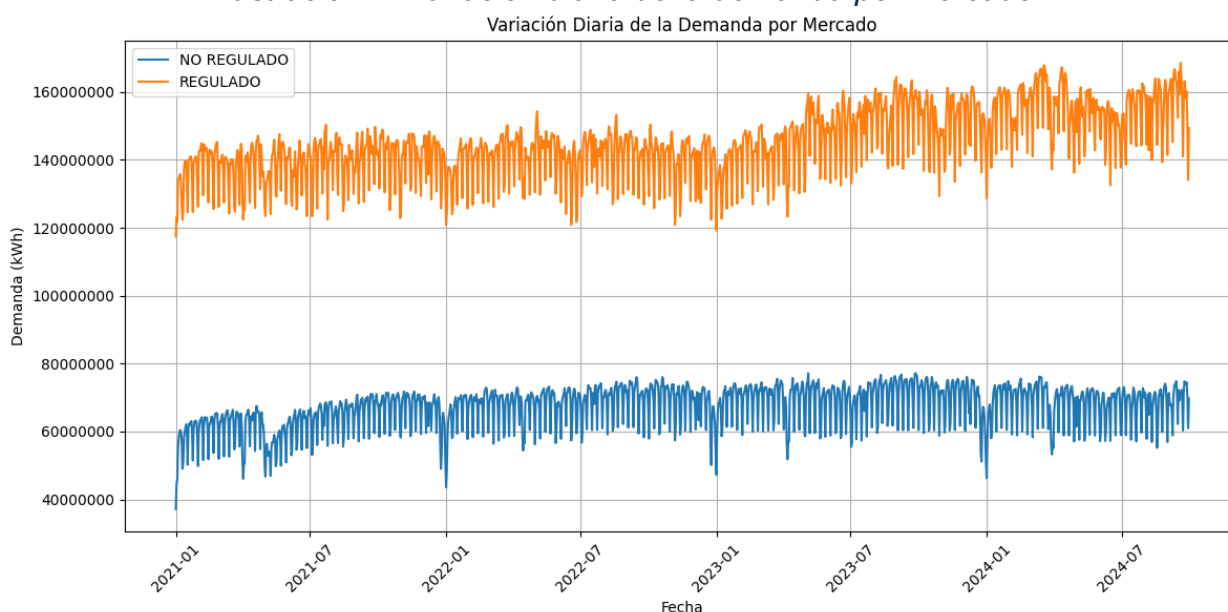
Ilustración 3. Demanda diaria por actividad económica (Top 10)



Fuente: Elaboración propia.

Según la *Ilustración 3*, se puede observar que la actividad económica *Industrias Manufactureras* presenta una alta variabilidad con respecto a *Explotación de minas y canteras*. Adicionalmente, las *Industrias Manufactureras* presentan una leve tendencia de reducción de la demanda de energía, mientras que, para la para *Explotación de minas y canteras* se ve un pronunciado crecimiento de la demanda de energía en los últimos años.

Ilustración 4. Variación diaria de la demanda por mercado

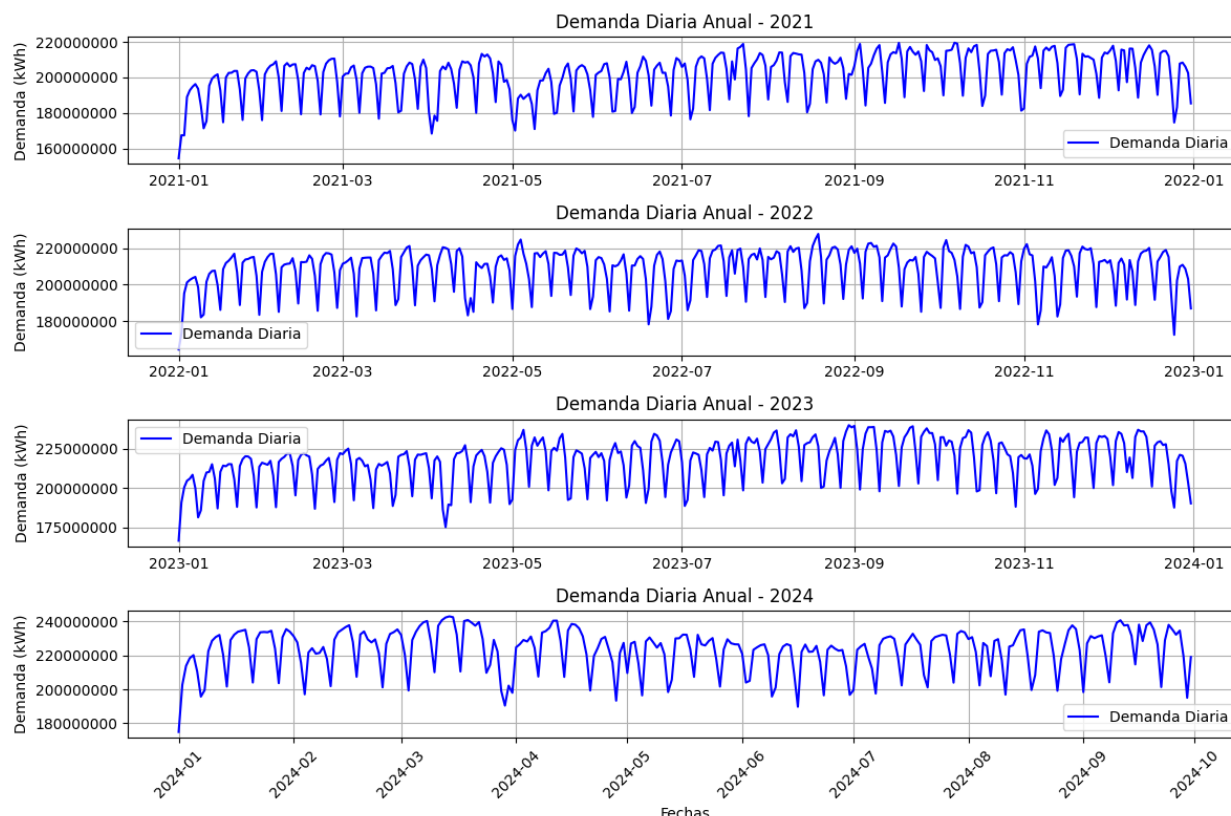


Fuente: Elaboración propia.

Conforme a la *Ilustración 4*, se puede observar que el mercado regulado (*principalmente residencial*) demanda aproximadamente el doble de energía que el mercado no regulado (*principalmente industrial y comercial*) y, que el mercado regulado presenta una tendencia de crecimiento mayor con respecto al mercado no regulado.



Ilustración 5. Demanda diaria anual




Fuente: Elaboración propia.

Por último, en la *Ilustración 5* se puede observar que existe un patrón en el comportamiento de la demanda de energía tanto a nivel semanal como mensual que se repite a lo largo de los años, lo cual, sugiere que cierto grado de estacionalidad en la demanda de energía, es decir, su comportamiento es recurrente y predecible para intervalos de tiempos regulares.

7.CONCLUSIONES INICIALES

- ✓ Las regiones con mayor consumo medio de energía son Bogotá y Cundinamarca, con aproximadamente 21.9 millones de kWh, seguidas de Antioquia con 14.7 millones de kWh.
- ✓ Otras regiones con consumo notable incluyen el Caribe Mar y el Caribe Sol, ambos superando los 13 millones de kWh.
- ✓ La variabilidad en el consumo de energía es considerable, siendo Bogotá y Cundinamarca las que muestran la mayor desviación estándar, lo que indica una fluctuación significativa en su demanda. En contraste, regiones como Chocó y

 Universidad de los Andes Facultad de Ingeniería	Departamento de Ingeniería de Sistemas y Computación Maestría Maestría en Ingeniería de Información MINE-4105_1: SOLUCIONES DE DATOS EN LA NUBE Semestre 2024-20 Horario: Lunes 6:00 a 9:00 p.m., SD_401	Acreditación institucional de alta calidad 10 años Mineducación Resolución 582 del 9 de enero de 2015
---	--	--

Guaviare presentan un consumo más estable, con desviaciones estándar menores.

- ✓ Las regiones sin clasificar tienen un consumo significativamente bajo, sugiriendo que los datos de estas áreas pueden no estar completamente representados o que la actividad económica es mínima.
- ✓ En regiones con mayor actividad económica, como "Bogotá - Cundinamarca" y "Antioquia", se observa una correlación positiva entre el crecimiento económico y la demanda de energía, lo que también se refleja en un aumento del IPC.
- ✓ Regiones como "Caribe Sol" y "Cali - Yumbo - Puerto Tejada" muestran una demanda relativamente estable, indicando que tienen necesidades energéticas constantes, posiblemente impulsadas por patrones de consumo industrial o residencial.
- ✓ Algunas regiones con características mixtas urbanas y rurales exhiben alta variabilidad en la demanda a lo largo de los meses, especialmente durante temporadas agrícolas, lo que sugiere que el consumo de energía está influenciado por sectores específicos en lugar de una inflación generalizada.
- ✓ En 2024, se observa que las temperaturas alcanzan niveles más altos que en años anteriores, con picos que superan los máximos registrados anteriormente. Esta variabilidad sugiere que la temperatura podría no ser una variable confiable para el modelo, ya que no muestra un comportamiento estable para realizar proyecciones precisas.
- ✓ La mayor demanda de energía se registra en las ciudades capitales, lo que indica que no existe una relación directa entre la temperatura promedio y la demanda energética. Esto sugiere que otros factores, además de la temperatura, influyen en el consumo de energía en estas áreas.
- ✓ Los días "No Festivos" muestran una demanda más alta y variabilidad en comparación con los días "Festivos y Domingos".
- ✓ Los "Domingos y Festivos" tienen demandas promedio similares, lo que sugiere un patrón de consumo similar en esos días.
- ✓ Es recomendable considerar la variabilidad en la demanda, especialmente en días "No Festivos", para el análisis y la planificación futura de la oferta de energía.
- ✓ Se debe realizar una reunión de socialización con el cliente (*MinEnergía*) para dar a conocer los hallazgos encontrados en esta primera etapa para de análisis exploratorio de los datos.
- ✓ Se debe realizar la búsqueda de más variables y validar su relación con la demanda de energía, puesto que, algunas de las variables seleccionadas en la primera etapa no aportan de forma significativa al comportamiento de la variable objetivo.
- ✓ Por último, se debe proceder con el ajuste necesario de los datos (*limpieza*) para que puedan ser utilizados en el modelo de series de tiempo LSTM, puesto que, este ajuste varía en comparación con un modelo tradicional.