

Procedimiento Desarrollado para el Análisis de la Demanda Energética en la Industria Manufacturera de Productos Alimenticios

Este análisis tiene como objetivo comprender el comportamiento de la demanda energética en la industria alimentaria, específicamente en los sectores relacionados con la manufactura de productos alimenticios. El proceso se desarrolló en varias etapas que incluyen la obtención de los datos, su transformación, agregación y visualización para facilitar su interpretación. A continuación, se describe cada uno de los pasos realizados:

1. **Conexión a la API de XM:** El análisis comenzó con la conexión a la API de XM utilizando la librería pydataxm, lo que permitió obtener los datos sobre la demanda energética en el sector comercial no regulado, enfocados especialmente en la industria alimentaria.
2. **Selección de los Códigos CIIU Relevantes:** Una vez obtenidos los datos, se filtraron para seleccionar solo aquellos correspondientes a los códigos CIIU de la industria manufacturera alimentaria. Para ello, se definió una lista específica de códigos que cubren las actividades de elaboración de productos alimenticios.
3. **Transformación y Agregación de los Datos:** Los datos fueron procesados para ser agregados a nivel diario. Esto implicó convertir las columnas de hora en un formato adecuado, agrupar los datos por fecha y código CIIU, y calcular la suma de la demanda energética horaria, resultando en la demanda total diaria. Posteriormente, se consolidaron los valores de demanda de cada día para obtener un único valor total por fecha y código CIIU.
4. **Cálculo de Métricas y Porcentajes:** Para facilitar el análisis de la distribución de la demanda energética, se calcularon varias métricas, tales como el promedio de demanda diaria por semana, el porcentaje de la demanda total por cada código CIIU, y la demanda mensual total. Estos cálculos fueron realizados utilizando el método groupby de pandas, lo que permitió organizar los datos por día, semana y mes.
5. **Visualización de los Resultados:** Para representar gráficamente los resultados, se crearon diversas visualizaciones, que incluyen:
 - **Comportamiento de la demanda por hora:** Un gráfico de líneas que muestra la demanda energética promedio por hora.
 - **Distribución porcentual de la demanda por código CIIU:** Un gráfico de barras horizontales que ilustra qué porcentaje del total de la demanda corresponde a cada código CIIU.
 - **Mapa de calor de la demanda energética por mes:** Un gráfico tipo mapa de calor que muestra cómo se distribuye la demanda energética mes a mes para cada código CIIU.
 - **Demanda energética mensual y por día de la semana:** Un gráfico de barras que presenta la demanda total mensual, acompañado de su porcentaje correspondiente.
 - **Promedio de demanda energética por día de la semana:** Un gráfico de barras que muestra el promedio de la demanda por cada día de la semana.

6. **Creación del Diccionario de la Industria Alimentaria:** Como parte del proceso de transformación de los datos, se creó el **diccionario_industria_alimentaria**, el cual mapea los códigos CIIU a descripciones más comprensibles sobre las actividades relacionadas con la industria alimentaria. Esto facilitó la interpretación de los resultados y permitió una mayor claridad en la presentación de los datos.

Agregación Utilizada:

Para obtener los resultados solicitados, se utilizaron las siguientes técnicas de agregación:

- **Agrupación por fecha y código CIIU:** Los datos fueron agrupados por día y por código CIIU, sumando la demanda energética diaria.
- **Sumatoria de horas:** Para cada día, se sumaron las demandas horarias para obtener la demanda total diaria.
- **Promedio por día de la semana:** Se calculó el promedio de la demanda energética para cada día de la semana.
- **Porcentajes de distribución:** Para entender la distribución de la demanda, se calculó el porcentaje que representa cada código CIIU respecto al total anual de demanda.

Dificultades Encontradas:

Durante el desarrollo del análisis, se presentaron varias dificultades que fueron resueltas de la siguiente manera:

1. **Manejo de Datos Temporales:** El manejo adecuado de las fechas y horas fue clave para la correcta agregación y análisis de la demanda. Convertir las columnas horarias y trabajar con series temporales requirió cuidado y atención para garantizar la precisión en los resultados.
2. **Limpieza y Filtrado de Datos:** Aunque los datos obtenidos eran relevantes, se necesitó un filtrado exhaustivo para asegurar que solo se incluyeran los códigos CIIU correspondientes a la industria alimentaria, y que no hubiera datos inconsistentes.
3. **Escalabilidad de las Visualizaciones:** Algunas de las visualizaciones generadas no fueron tan legibles debido a la cantidad de datos presentados. En particular, las etiquetas de los códigos CIIU y los porcentajes en los gráficos de barras se superponían, lo que dificultaba la interpretación. Este problema se podría mejorar mediante la agrupación de códigos o el uso de gráficos interactivos.

Mejoras para el Proceso:

Si tuviera más tiempo para desarrollar el proceso, consideraría implementar las siguientes mejoras:

1. **Automatización de las Consultas:** Para aumentar la eficiencia, se podría automatizar la obtención de datos desde la API de XM, permitiendo que el análisis se actualice automáticamente sin intervención manual.
2. **Optimización de la Visualización Interactiva:** Para facilitar la exploración de los resultados, se podría implementar gráficos interactivos utilizando librerías como plotly o bokeh, lo que permitiría hacer zoom, filtrar o ajustar los gráficos en tiempo real.
3. **Análisis Predictivo:** Se podría extender el análisis utilizando técnicas de análisis predictivo, como modelos de series temporales (ARIMA) o algoritmos de machine learning, para prever la demanda energética futura y ajustar estrategias en tiempo real.

Conclusión:

Este análisis ha permitido comprender mejor el comportamiento de la demanda energética en la industria manufacturera de productos alimenticios. Aunque se presentaron algunas dificultades durante el proceso, las soluciones aplicadas nos proporcionaron resultados valiosos. Con un enfoque en la automatización, la optimización de visualizaciones y el análisis predictivo, este proceso podría mejorar significativamente para análisis futuros.

Todo el material utilizado, incluyendo el código y el **diccionario_industria_alimentaria**, está disponible en el siguiente repositorio de GitHub:

[GitHub - Talento B - Ingeniería de Datos](#)