

# *ANALISIS EFISIENSI KONSUMSI BAHAN BAKAR*

Disusun oleh Kelompok 19 :

**Kevin Febrian Widhiarta(24031554025)**

**Davan Juheins Tololiu (24031554157)**

Dosen pengampu :

**Dinda Galuh Guminta, M.Stat.**

**Belgis Ainatul Iza, S.Si., M.Mat.**

# *Pendahuluan*

---

Efisiensi bahan bakar menjadi perhatian utama, karena terdapat **masalah lingkungan** (emisi gas rumah kaca) dan **ekonomi** (biaya operasional kendaraan).

Keputusan pembelian kendaraan seringkali dipengaruhi oleh faktor teknis dan eksternal.

## *Rumusan Masalah*

---

1. Bagaimana data teknis kendaraan yang kompleks dan data historis harga minyak dan bensin dapat diintegrasikan dan distandarisasi untuk menciptakan dataset terstruktur?
2. Bagaimana tren rata-rata Efisiensi BBM (km/L) dari tahun ke tahun berkorelasi dengan perubahan harga minyak mentah global dan harga bensin?
3. Bagaimana data kendaraan berintegrasi terhadap efisiensi untuk menentukan merek kendaraan dengan efisiensi serta inefisiensi(boros)?
4. Bagaimana mengetahui trend efisiensi BBM terhadap brand kendaraan?
5. Bagaimana hubungan harga minyak mentah dan harga BBM terhadap efisiensi kendaraan?

# *Tujuan*

---

1. **Mengintegrasikan** data spesifikasi kendaraan, data efisiensi BBM, dan data historis harga minyak mentah berdasarkan Tahun.
2. **Membersihkan data** dan mengkonversi satuan efisiensi BBM.
3. Melakukan **eksplorasi data** untuk mengidentifikasi korelasi antara variabel teknis fitur dengan efisiensi BBM, serta korelasi temporal dengan harga minyak global.
4. **Menganalisis** efisiensi kendaraan dari tahun ke tahun berdasarkan data historis untuk mengambil kesimpulan apakah terjadi peningkatan atau penurunan efisiensi dalam periode tertentu.
5. Mengetahui merek kendaraan **paling efisien** dan **boros** berdasarkan analisis merek terhadap efisiensi kendaraan.
6. Melakukan **analisis korelasi** harga minyak mentah terhadap efisiensi kendaraan.

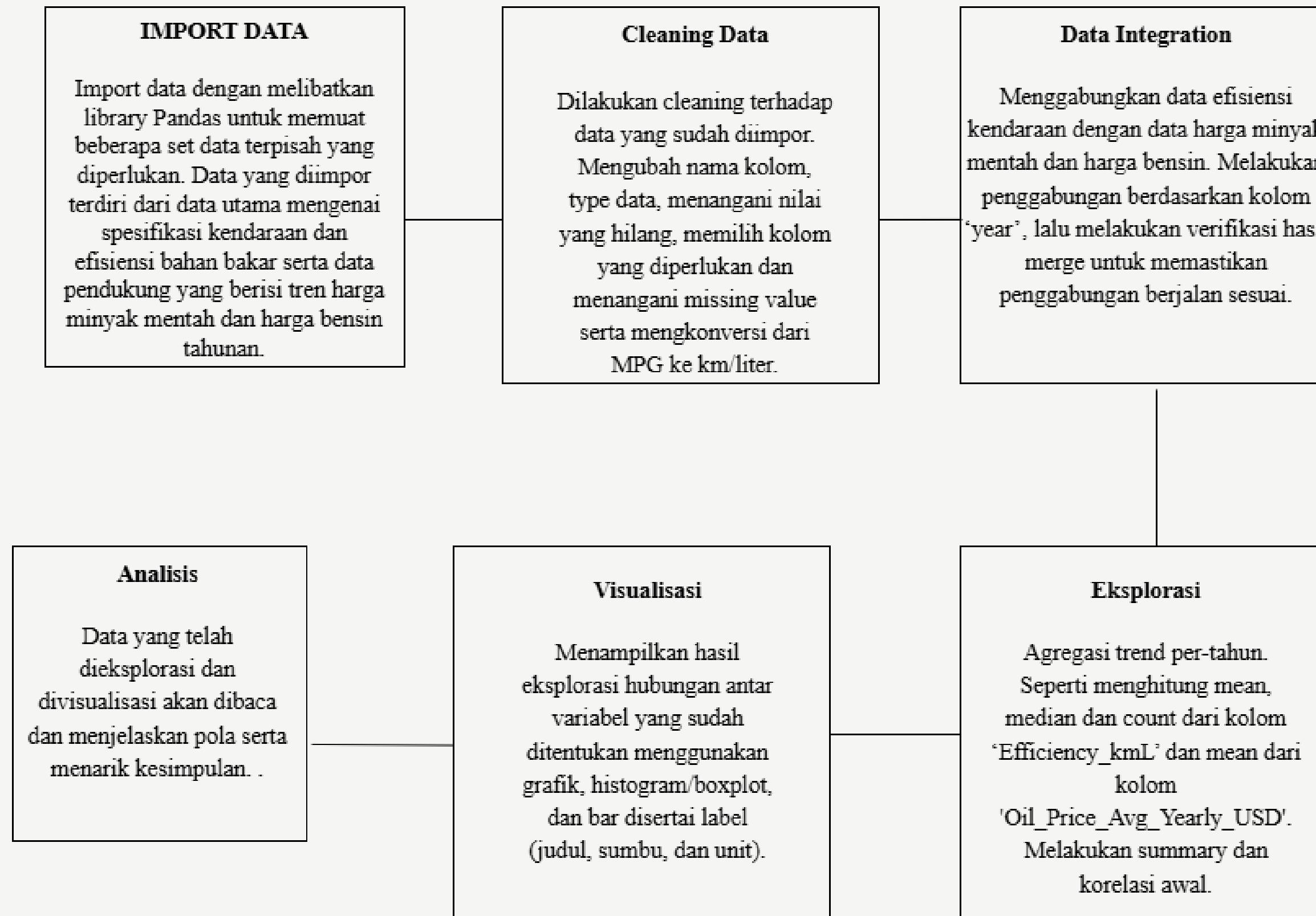
# *Manfaat*

---

- 1. Bagi Konsumen:** Memberikan wawasan berbasis data untuk **membandingkan efisiensi berbagai model kendaraan** sebelum membeli.
- 2. Bagi Industri Otomotif: Mengidentifikasi fitur teknis** (misalnya, jenis transmisi, ukuran mesin) yang paling berkorelasi dengan efisiensi tinggi.
- 3. Bagi Kebijakan Energi:** Menyediakan dasar data untuk studi **dampak fluktuasi harga energi terhadap tren efisiensi** yang ditawarkan pabrikan.

# Pipeline

---



# Metode Proses

Import Data

```
import pandas as pd
import numpy as np

FILE_KENDARAAN = 'vehicles.csv'
FILE_HARGA_MINYAK = 'wti_oil_prices.csv'
FILE_HARGA_BENSIN = 'retail_gasoline_prices.csv'

try:
    df_kendaraan = pd.read_csv(FILE_KENDARAAN, low_memory=False)
    print("Data Kendaraan Dimuat:", df_kendaraan.shape)
except Exception as e:
    print(f"Error loading Data Kendaraan: {e}")

try:
    df_minyak_bumi = pd.read_csv(FILE_HARGA_MINYAK)
    print("\nData Harga Minyak mentah Dimuat:", df_minyak_bumi.shape)
except Exception as e:
    print(f"Error loading Data Harga Minyak: {e}")

try:
    df_bensin = pd.read_csv(FILE_HARGA_BENSIN, sep=';', skiprows=4)
    print(f"\nData Harga Bensin Dimuat: {df_bensin.shape}")
except Exception as e:
    print(f"Error loading Data Harga Bensin: {e}")
```

# Metode Proses

Clean Data

## Data Kendaraan

```
# Penanganan Missing Values dan Standardisasi/Konversi Satuan
if 'df_kendaraan' in locals() or 'df_kendaraan' in globals():
    print("df_kendaraan DataFrame found.")
    print("Columns in df_kendaraan:", df_kendaraan.columns.tolist())

    df_kendaraan_to_clean = df_kendaraan.copy()
    cols_to_keep = ['year', 'make', 'model', 'hpv', 'comb08', 'co2']

    missing_cols = [col for col in cols_to_keep if col not in df_kendaraan_to_clean.columns]
    if missing_cols:
        print(f"Error: The following columns are missing from df_kendaraan: {missing_cols}")
    else:
        df_kendaraan_clean = df_kendaraan_to_clean[cols_to_keep].dropna(subset=['hpv']).copy()

        # MPG (Miles Per Gallon) ke km/Liter (1 mil = 1.60934 km, 1 galon = 3.78541 liter)
        KONVERSI MPG TO KML = 0.425143707
        df_kendaraan_clean['Efficiency_km/L'] = df_kendaraan_clean['comb08'] * KONVERSI MPG TO KML

        df_kendaraan_clean.drop(columns=['comb08', 'co2'], inplace=True)
        df_kendaraan_clean = df_kendaraan_clean[(df_kendaraan_clean['year'] >= 2016) & (df_kendaraan_clean['year'] <= 2025)].copy()
else:
    print("df_kendaraan DataFrame not found. Please run the data loading cell first.")

df_kendaraan_clean
```

# Metode Proses

Clean Data

## Harga Minyak Mentah

```
df_minyak_bumi['date'] = pd.to_datetime(df_minyak_bumi['date'])

df_minyak_bumi['Year'] = df_minyak_bumi['date'].dt.year

df_minyak_bumi_clean = df_minyak_bumi.groupby('Year')['close'].mean().reset_index()
df_minyak_bumi_clean.rename(columns={'close': 'Oil_Price_Avg_Yearly_USD'}, inplace=True)

print("\nHarga Minyak Bumi Tahunan:\n")
df_minyak_bumi_clean
```

# Metode Proses

Clean Data

## Harga Bensin

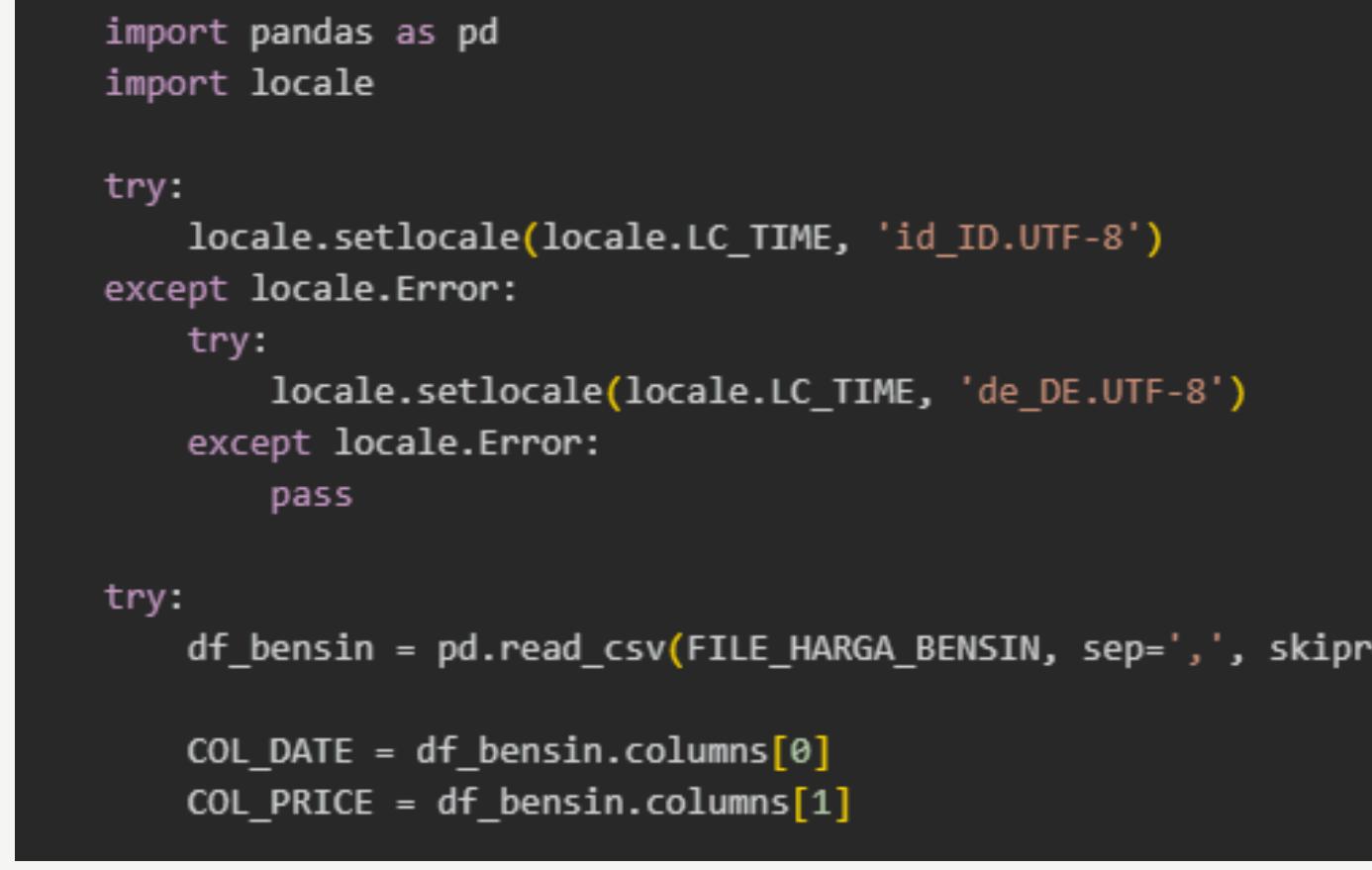
```
import pandas as pd
import locale

try:
    locale.setlocale(locale.LC_TIME, 'id_ID.UTF-8')
except locale.Error:
    try:
        locale.setlocale(locale.LC_TIME, 'de_DE.UTF-8')
    except locale.Error:
        pass

try:
    df_bensin = pd.read_csv(FILE_HARGA_BENSIN, sep=',', skiprows=4)

    COL_DATE = df_bensin.columns[0]
    COL_PRICE = df_bensin.columns[1]

```



```
try:
    df_bensin['Date'] = pd.to_datetime(df_bensin[COL_DATE], format='%b %Y')
except ValueError:
    df_bensin['Date'] = pd.to_datetime(df_bensin[COL_DATE], format='%b-%y')

df_bensin['Year'] = df_bensin['Date'].dt.year

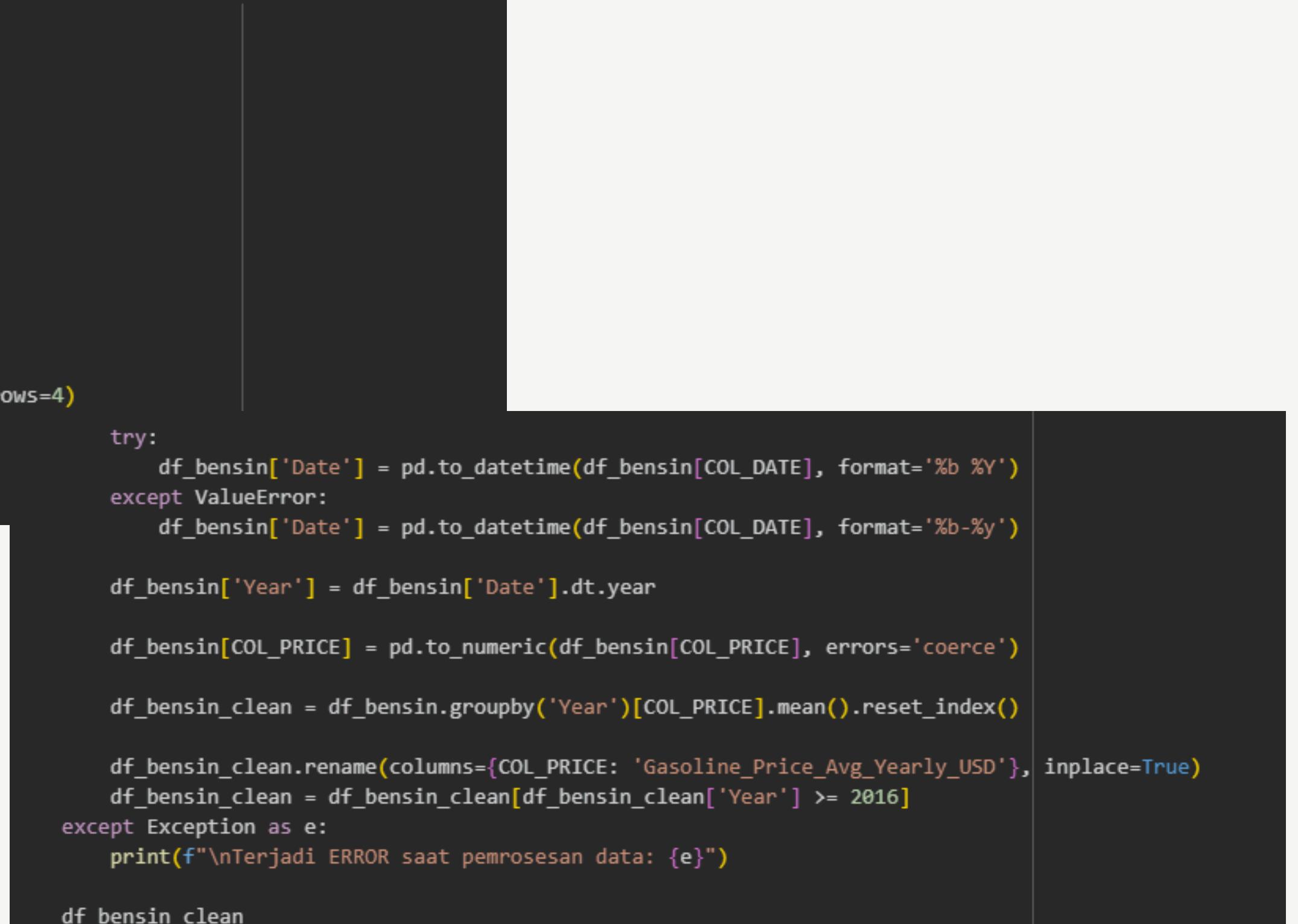
df_bensin[COL_PRICE] = pd.to_numeric(df_bensin[COL_PRICE], errors='coerce')

df_bensin_clean = df_bensin.groupby('Year')[COL_PRICE].mean().reset_index()

df_bensin_clean.rename(columns={COL_PRICE: 'Gasoline_Price_Avg_Yearly_USD'}, inplace=True)
df_bensin_clean = df_bensin_clean[df_bensin_clean['Year'] >= 2016]
except Exception as e:
    print(f"\nTerjadi ERROR saat pemrosesan data: {e}")

df_bensin_clean

```



# Metode Proses

Integrasi Data

```
df_kendaraan_clean.rename(columns={'year': 'Year'}, inplace=True)

df_integrated = pd.merge(df_kendaraan_clean, df_minyak_bumi_clean, on='Year', how='left')

df_integrated = pd.merge(df_integrated, df_bensin_clean, on='Year', how='left')

df_integrated
```

# Metode Proses

Eksplorasi

## Menghitung Biaya Bahan Bakar Tahunan

- Jarak tempuh tahunan rata-rata: Kita akan menggunakan 15.000 km sebagai standar.
- Konversi satuan: harga bensin dari USD per galon ke USD per liter (1 galon = 3.78541 liter).

```
ANNUAL_DISTANCE_KM = 15000

GALLON_TO_LITER = 3.78541

df_fuel_cost = df_integrated.copy()

df_fuel_cost.dropna(subset=['Gasoline_Price_Avg_Yearly_USD'], inplace=True)

df_fuel_cost['Gasoline_Price_Avg_Yearly_USD_per_Liter'] = \
    df_fuel_cost['Gasoline_Price_Avg_Yearly_USD'] / GALLON_TO_LITER

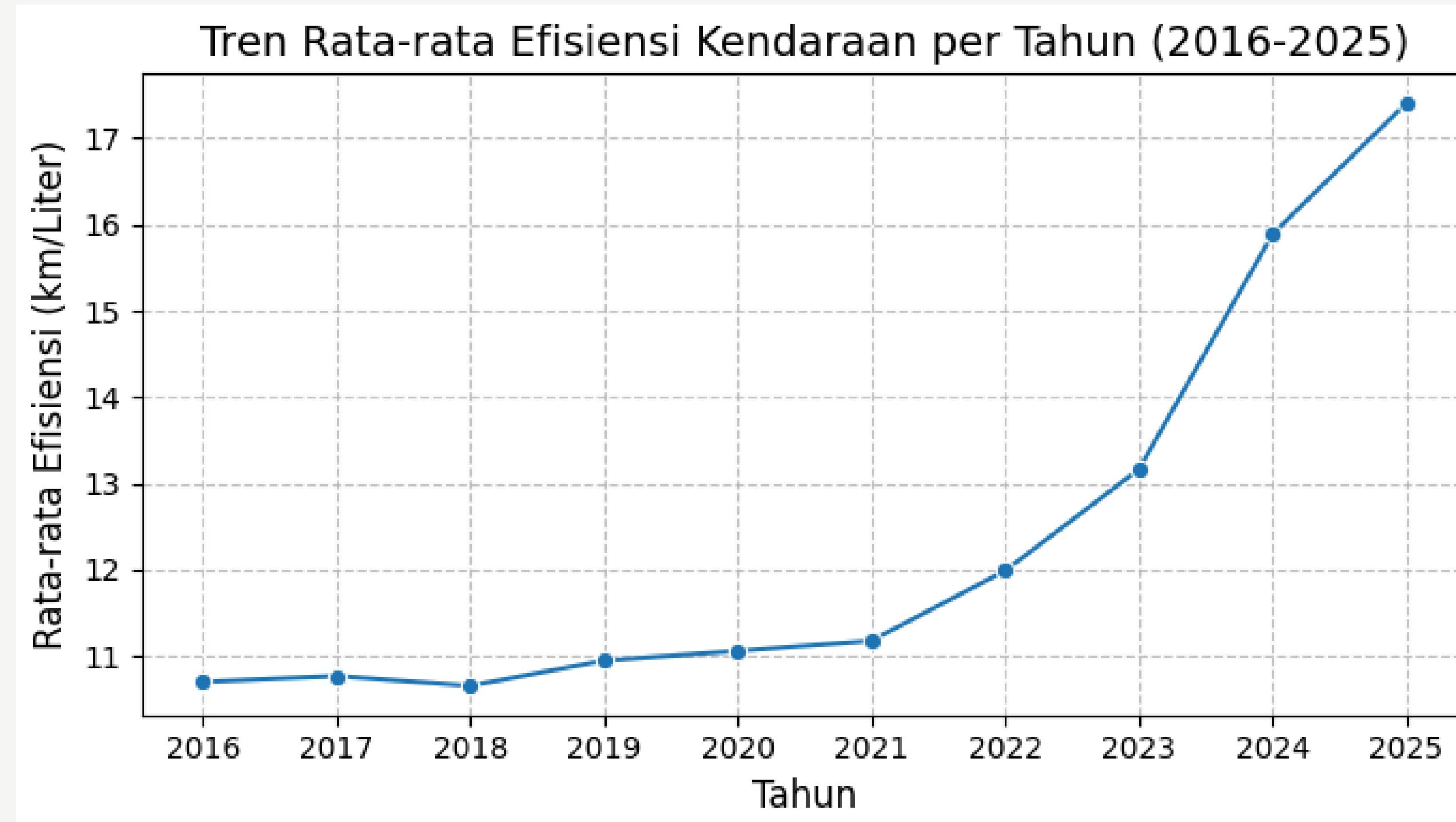
df_fuel_cost['Annual_Fuel_Consumption_Liters'] = \
    ANNUAL_DISTANCE_KM / df_fuel_cost['Efficiency_km/L']

df_fuel_cost['Annual_Fuel_Cost_USD'] = \
    df_fuel_cost['Annual_Fuel_Consumption_Liters'] * df_fuel_cost['Gasoline_Price_Avg_Yearly_USD_per_Liter']

df_fuel_cost
```

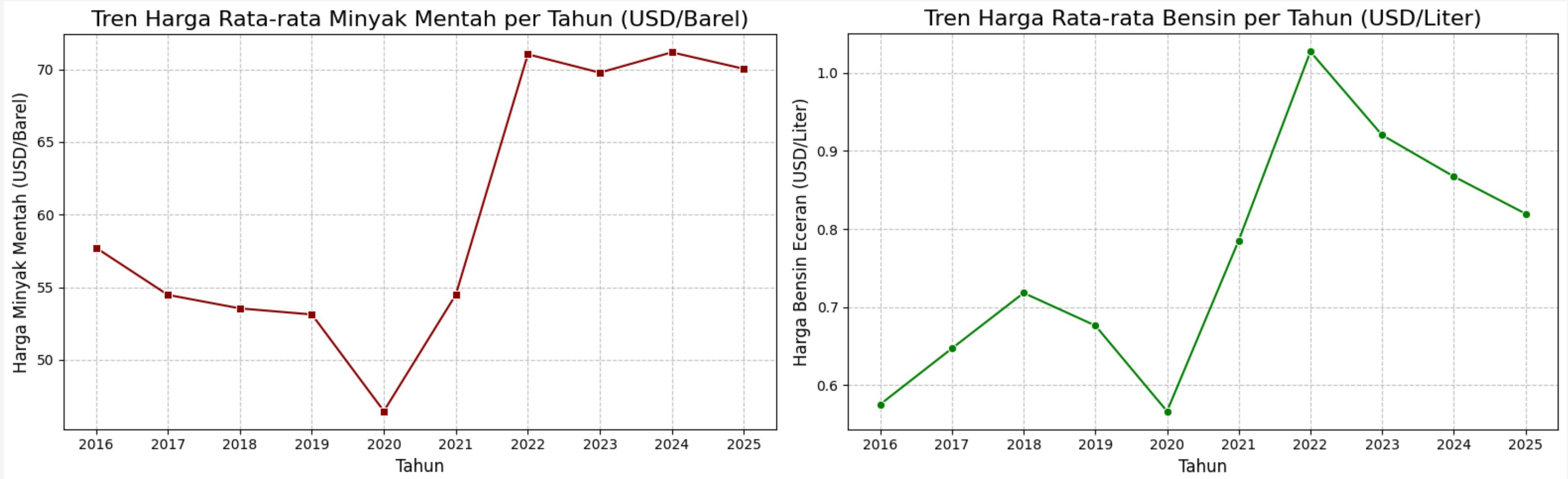
# Visualisasi dan Analisis

Tren Efisiensi Tahunan



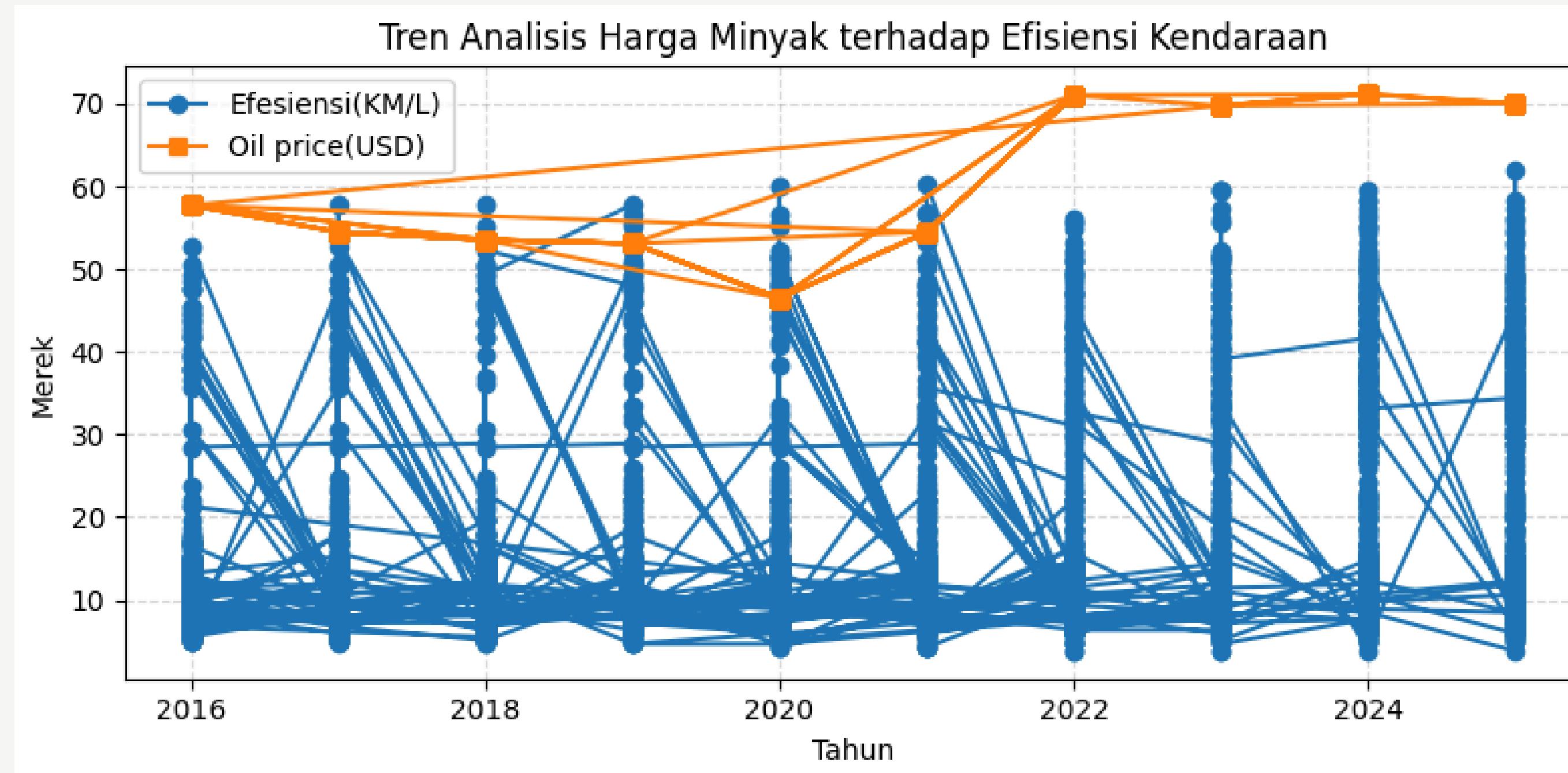
# Visualisasi dan Analisis

Tren Harga Minyak Mentah dan Bensin



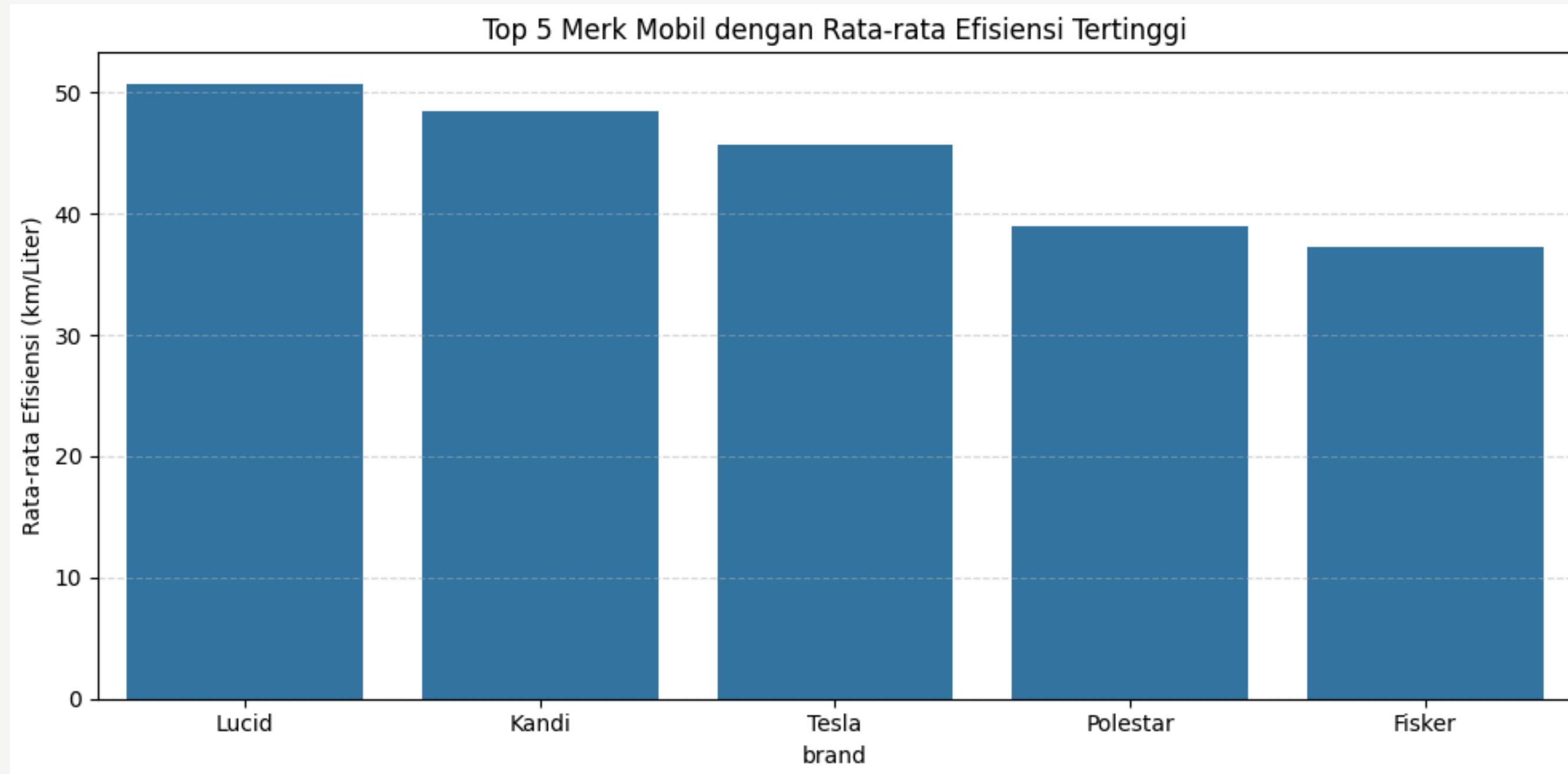
# Visualisasi dan Analisis

Tren Harga Minyak Mentah terhadap Efisiensi Kendaraan



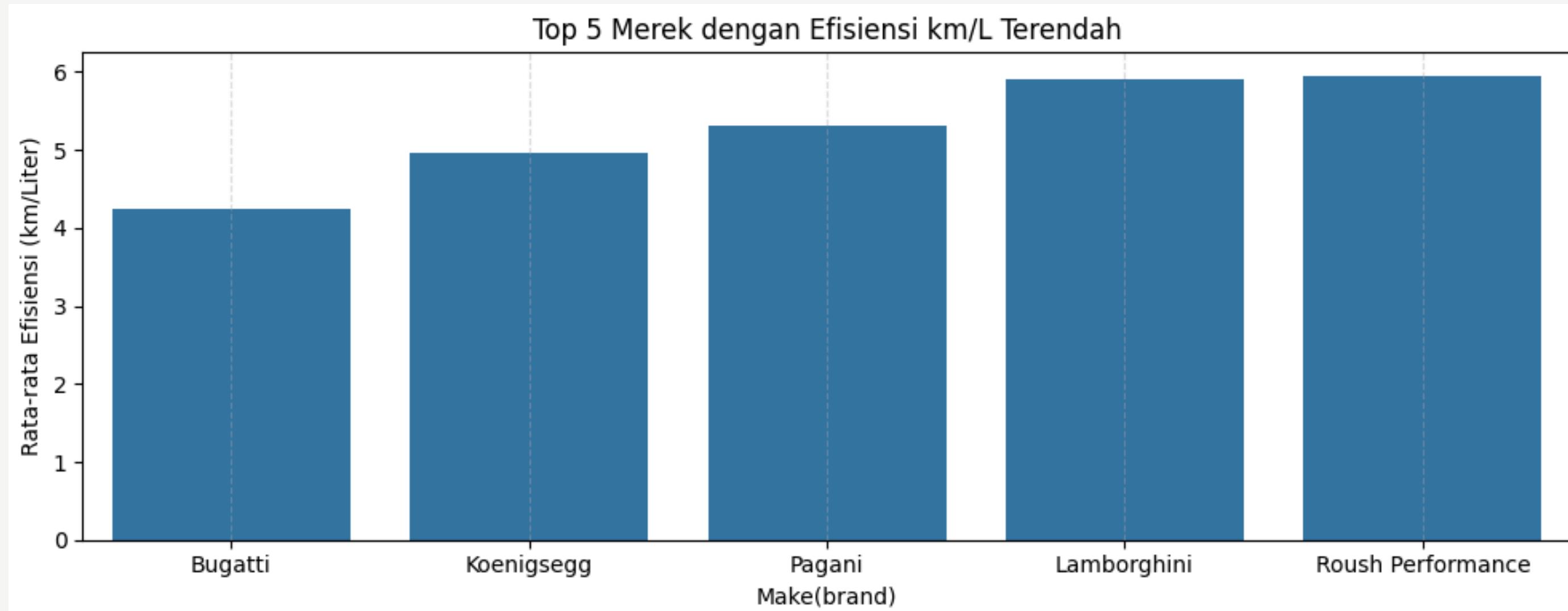
# Visualisasi dan Analisis

---



# Visualisasi dan Analisis

---



# *Hasil dan Analisa*

---

Berdasarkan proses pengolahan data efisiensi kendaraan

1. Pada rentang 2016-2025 tren efisiensi kendaraan cenderung **mengalami peningkatan**.
2. Harga minyak mentah memiliki korelasi yang **lemah** terhadap efisiensi kendaraan.
3. Dalam hal efisiensi, terdapat beberapa merek kendaraan yang masuk dalam kategori inefisiensi(boros).
4. Faktor teknologi dan regulasi **lebih dominan** dibandingkan harga minyak mentah dalam membentuk trend efisiensi kendaraan.
5. Tiap tahun, trend efisiensi **mengalami peningkatan bertahap**, tidak terdapat lonjakan yang signifikan.
6. Cleaning data dan integrasi data dapat meningkatkan kualitas analisis.

# *Kendala dan Rencana Tindak Lanjut*

Data yang akan diproses masih terdapat keterbatasan ketersediaan data

Membangun model regresi time series untuk **memprediksi tingkat efisiensi** yang optimal di masa depan.

Hasil analisis juga dapat dijadikan sebagai landasan **perencanaan strategis** dan investasi jangka panjang bagi industri kendaraan.

*Thank you*  
— For your attention