

# TUGAS PRESENTASI

---

Fuel Efficiency Prediction Using Machine  
Learning

# ANGGOTA KELOMPOK:

Iqbal Choirul Rasyid  
101012400328

Ariq Bagasputra. S  
101012400136

M Raafi Nurhadi  
101012400338

Zalzabila. S  
101012400269

# LATAR BELAKANG

Proyek ini bertujuan untuk menganalisis dan memprediksi efisiensi bahan bakar mobil berdasarkan karakteristik kendaraan menggunakan pendekatan machine learning. Efisiensi bahan bakar diklasifikasikan menjadi dua kategori yaitu irit dan boros untuk membantu memberikan rekomendasi kendaraan yang lebih ramah lingkungan.

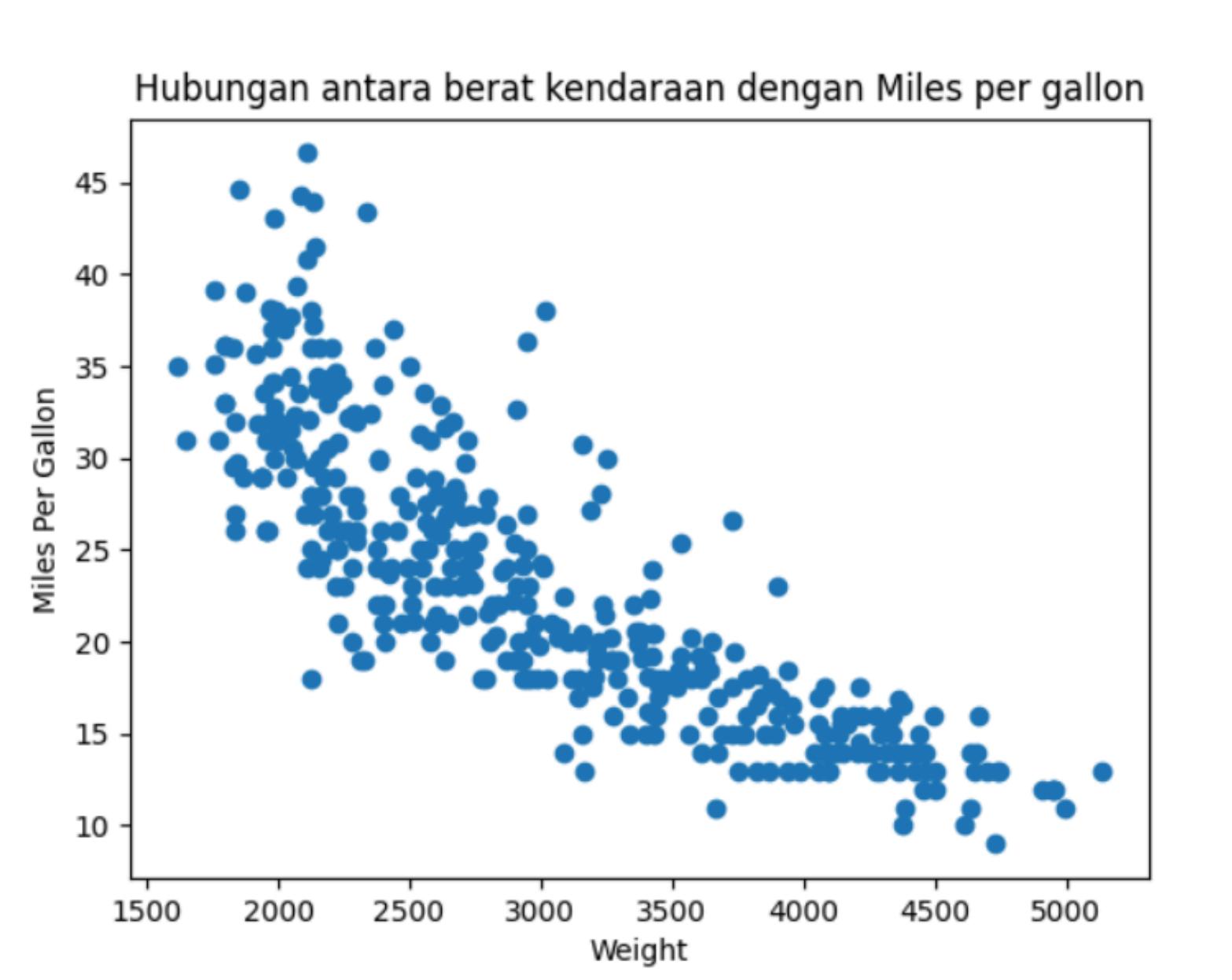
# TUJUAN PENELITIAN

1. Menganalisis pola penggunaan bahan bakar pada mobil
2. Mengidentifikasi karakteristik kendaraan yang mempengaruhi efisiensi bahan bakar
3. Membangun model machine learning untuk memprediksi kategori mobil irit atau boros bahan bakar
4. Memberikan rekomendasi kendaraan yang lebih efisien dan ramah lingkungan.

# DATASET

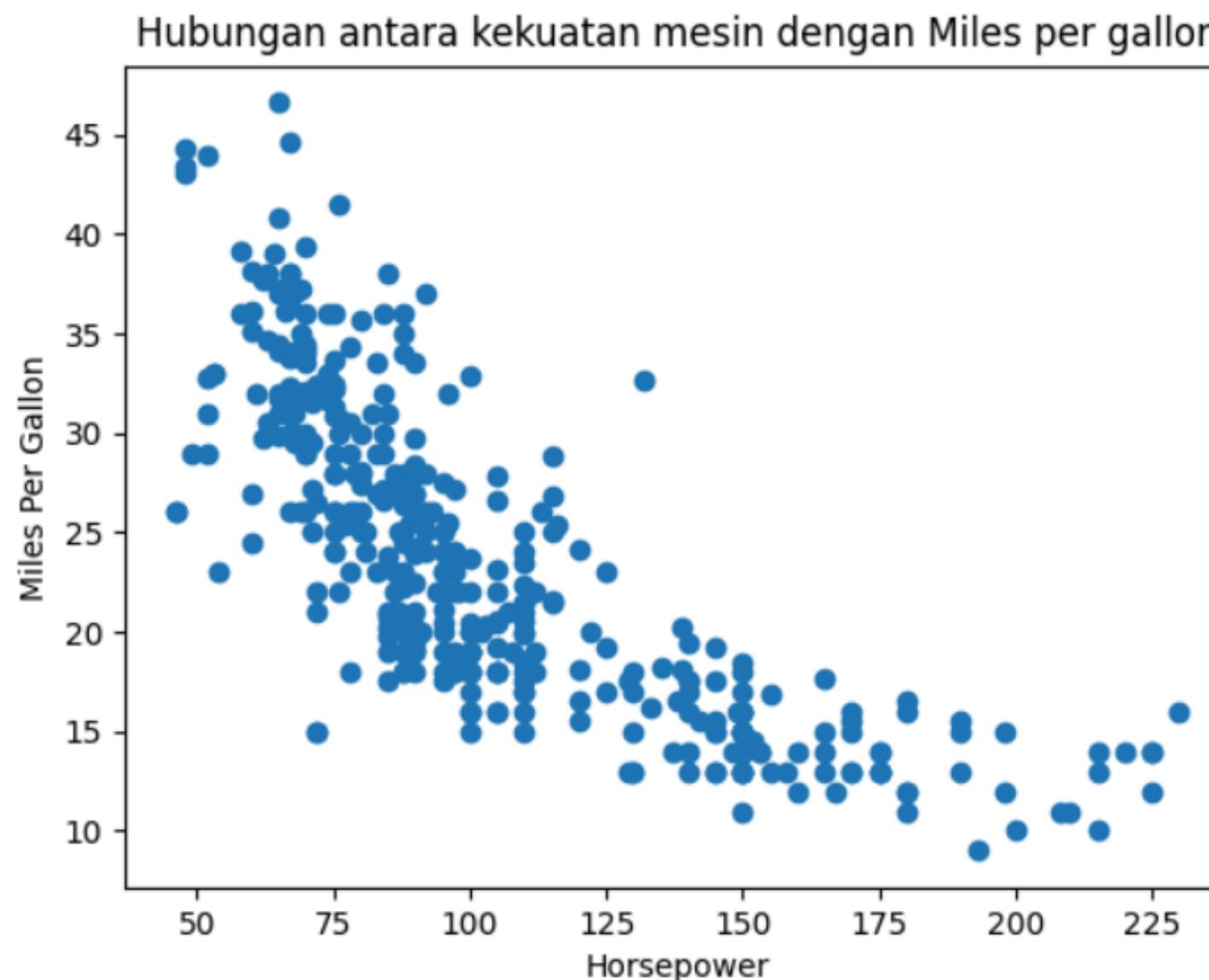
<https://github.com/hananlu/basicPython/blob/master/Dataset/cars2.json>

# VISUAL DATA



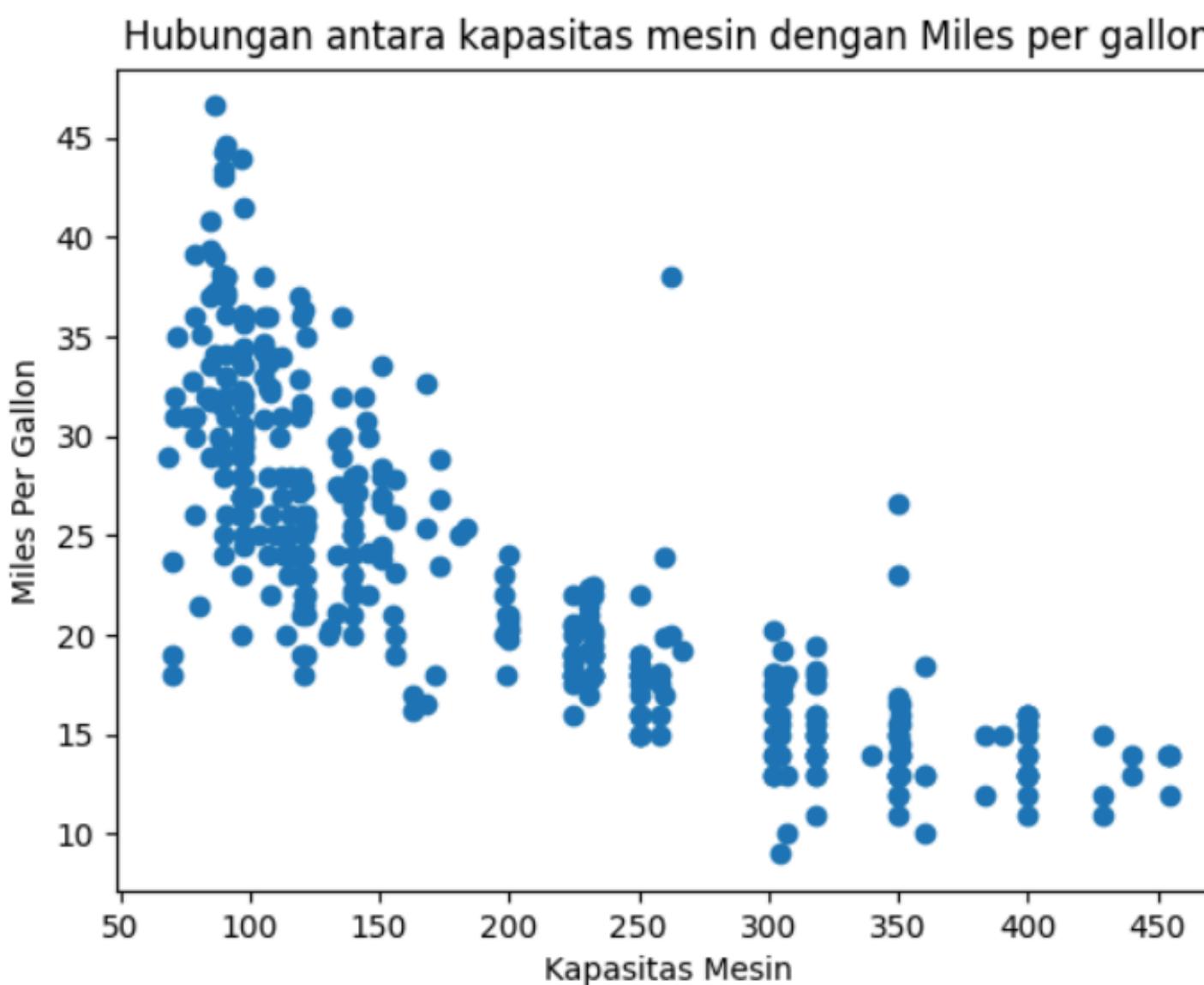
Terdapat hubungan terbalik yang sangat jelas antara berat kendaraan dan efisiensi bahan bakar. Semakin besar nilai Weight (berat kendaraan), maka nilai Miles Per Gallon (jarak tempuh per galon) akan semakin menurun. Kendaraan dengan efisiensi tertinggi didominasi oleh kendaraan ringan yang memiliki berat di kisaran 1.750 hingga 2.250 lbs. Pola akan mengalami penurunan pada berat 1.500–3.000 lbs, saat variasi MPG sangat tinggi, dan menunjukkan faktor lain seperti teknologi mesin sangat berpengaruh pada bobot ringan.

# VISUAL DATA



Pada pola terlihat jelas di mana peningkatan kekuatan mesin (Horsepower) berbanding terbalik dengan efisiensi bahan bakar (MPG). Semakin besar tenaga mesin, maka akan semakin rendah jarak tempuh per galonnya. Efisiensi bahan bakar terbaik di atas 35 MPG akan sangat efisien pada kendaraan dengan tenaga mesin rendah, yaitu di bawah 100 Horsepower. Pada rentang 50–150 HP, terjadi penurunan MPG yang sangat signifikan. Pada kendaraan dengan tenaga di atas 200 HP, efisiensi bahan bakar cenderung melandai dan stabil pada angka yang rendah (rata-rata di bawah 15-20 MPG).

# VISUAL DATA



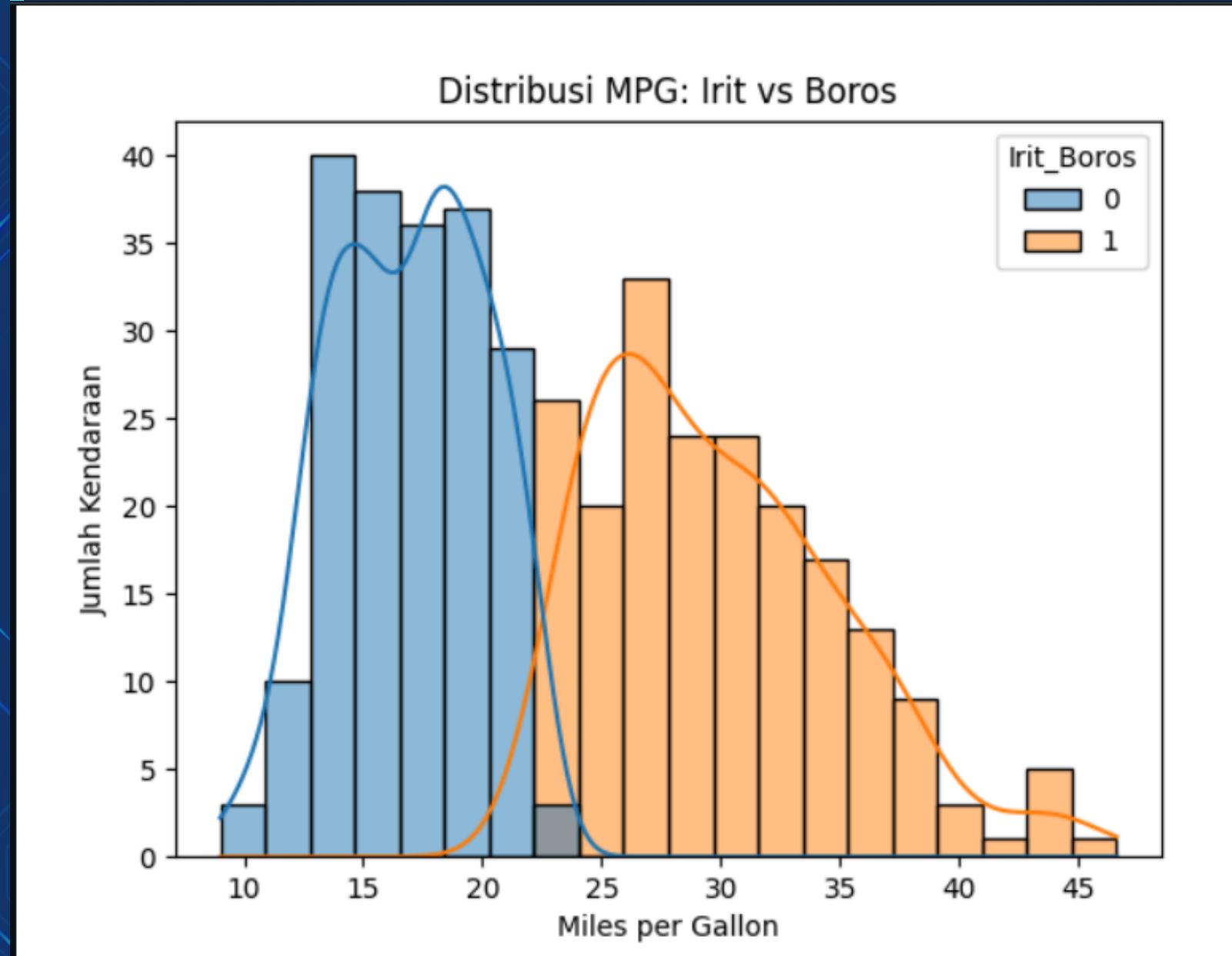
Terdapat hubungan terbalik yang jelas antara kapasitas mesin dan efisiensi bahan bakar. Semakin besar kapasitas mesin suatu kendaraan, maka nilai Miles Per Gallon (MPG) cenderung semakin rendah atau boros. Ketika kapasitas mesin melebihi 300, efisiensi bahan bakar menurun drastis dan cenderung stabil di angka yang rendah, yaitu di bawah 20 MPG. Jadi ketika mesin dengan kapasitas besar memerlukan volume bahan bakar yang lebih banyak untuk setiap siklus pembakaran.

# VISUAL DATA



Heatmap ini menunjukkan bahwa saat semua fitur utama ( Displacement, Horsepower, dan Weight) memiliki korelasi negatif yang sangat kuat terhadap Miles per Gallon (rentang -0,78 hingga -0,83), terdapat korelasi positif yang hampir sempurna antar fitur teknis mesin yaitu hubungan antara Displacement dengan Cylinders mencapai 0,95 dan dengan Weight mencapai 0,93. Hal ini menunjukkan bahwa kendaraan dengan kapasitas mesin besar hampir selalu memiliki jumlah silinder lebih banyak dan bobot yang lebih berat.

# VISUAL DATA



Visualisasi ini membagi kendaraan ke dalam dua kategori berdasarkan efisiensi bahan bakarnya untuk melihat pola sebaran data yaitu: Kategori 0 (Biru/Boros) ketika kendaraan dengan efisiensi rendah, terkonsentrasi di rentang 10 hingga 22 MPG, sedangkan kategori 1 (Oranye/Irit) ketika kendaraan efisien, tersebar luas mulai dari 23 hingga di atas 45 MPG. Kendaraan dalam kelompok Boros paling banyak ditemukan pada angka sekitar 15 MPG. Kendaraan dalam kelompok Irit memiliki puncak populasi di sekitar 25-27 MPG, namun distribusinya lebih melandai ke arah angka yang lebih tinggi.

# PRE PROCESSING

```
import pandas as pd  
import numpy as np  
  
df_raw = pd.read_json("X:\Code\Tubes_Phyton_Kel4\Dataset\Raw\cars2.json")  
df = df_raw.copy()  
baris_duplikat = df[df.duplicated()]  
print("jumlah data duplikat adalah", baris_duplikat.shape)  
print(df.isnull().sum())  
df = df.rename(columns = {"Miles_per_Gallon":"MPG",  
"Horsepower":"Tenaga Mesin", "Weight_in_lbs":"Berat Dalam LBS", "Displacement":"Kapasitas Mesin"})  
df.head(5)  
df = df[['MPG', 'Tenaga Mesin', 'Berat Dalam LBS', 'Kapasitas Mesin']]
```

```
df.dropna(inplace=True)  
median_mpg = df['MPG'].median()  
df['Irit atau Boros'] = np.where(df['MPG'] >= median_mpg, 1, 0)  
df.to_csv("X:\Code\Tubes_Phyton_Kel4\Dataset\Processed\cars_processed.csv", index=False)
```

# MODEL

```
import pandas as pd
import joblib
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
df =
pd.read_csv("X:\Code\Tubes_Phyton_Kel4\Dataset\Processed\cars_processed.csv")
X = df[['Tenaga Mesin', 'Berat Dalam LBS', 'Kapasitas Mesin']]
y = df['Irit atau Boros']
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=1
)
```

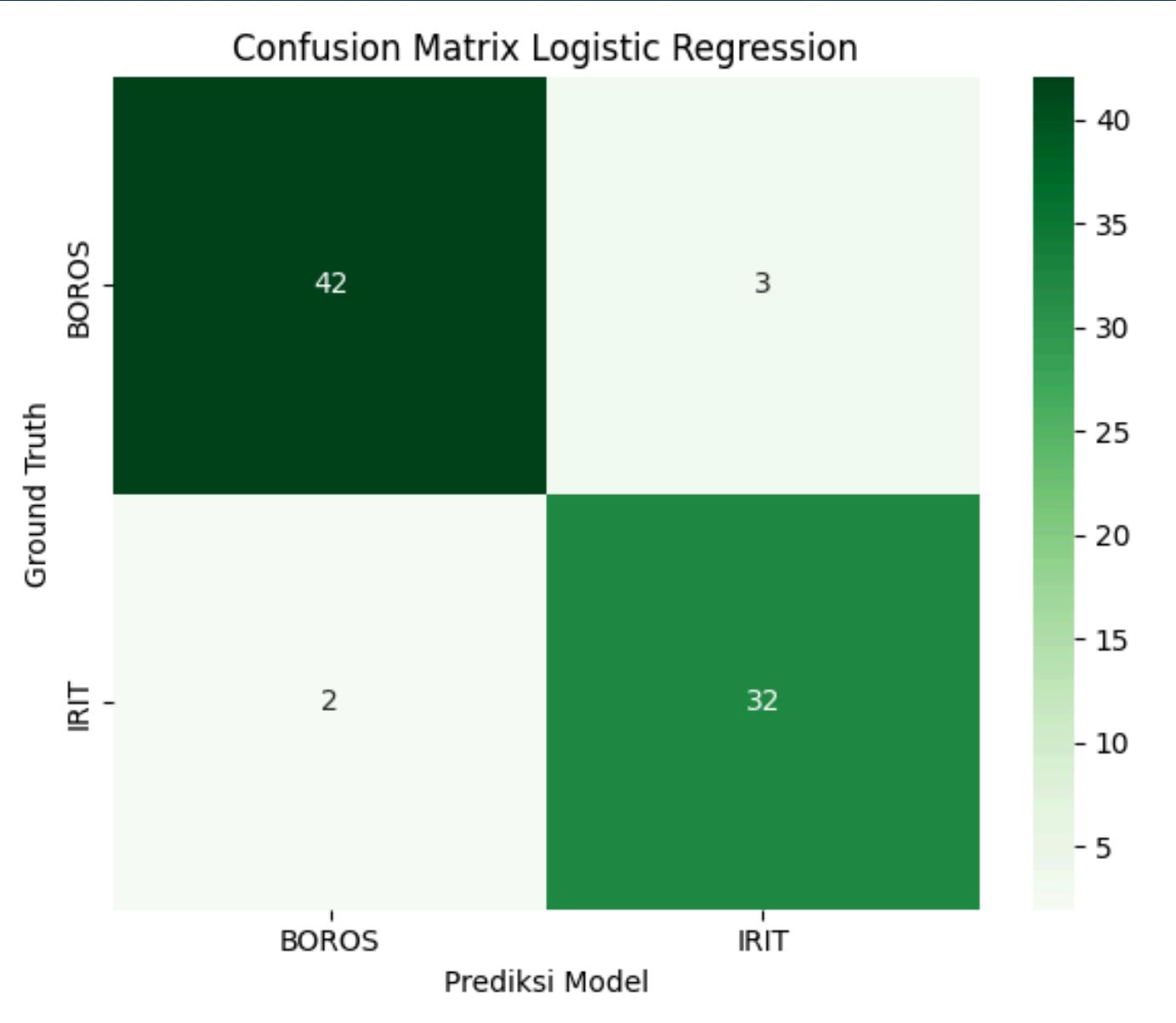
```
model = LogisticRegression(solver='liblinear')
model.fit(X_train, y_train)
accuracy = accuracy_score(y_test, model.predict(X_test))
print(f"Akurasi Model: {accuracy:.2f}")
joblib.dump(model, "model_logistic.joblib")
```

# TEST DATA BARU

```
import joblib
def pengubah_satuan(hp, berat_kg, cc):
    LBS = 2.20462
    CU_IN = 16.3871
    berat_lbs = berat_kg * LBS
    mesin_cu_in = cc / CU_IN
    return hp, berat_lbs, mesin_cu_in
def dataset():
    model      = joblib.load("X:/Code/Tubes_Phyton_Kel4/Model
ML/model_logistic.joblib")
    Nama = input("Masukkan Nama Kendaraan : ")
    hp = float(input("Masukkan horsepower kendaraan (hp): "))
    berat_kg = float(input("Masukkan berat kendaraan (kg): "))
    cc = float(input("Masukkan kapasitas mesin (cc): "))
    hp, berat_lbs, mesin_cu_in = pengubah_satuan(hp, berat_kg, cc)
    analisis = [[hp, berat_lbs, mesin_cu_in]]
    preds = model.predict(analisis)
    if preds[0] == 1:
        hasil = "IRIT"
    else:
        hasil = "BOROS"
print("\n==== HASIL ANALISIS ===")
print(f"Nama Mobil : {Nama}")
print(f"Horsepower : {hp}")
print(f"Berat Kendaraan : {berat_lbs:.2f} lbs")
print(f"Kapasitas Mesin : {mesin_cu_in:.2f} cu in")
print(f"Hasil Prediksi : {hasil}")
return hasil
while True:
    print("\n==== MENU ===")
    print("1. Analisis Menggunakan Model Logistic Regression")
    print("0. Keluar")
    pilihan = input("Masukkan pilihan: ")
    if pilihan == "1":
        dataset()
    elif pilihan == "0":
        print("Program selesai.")
        break
    else:
        print("Pilihan tidak valid.")
```

# CONFUSION MATRIX

Confusion Matrix Logistic Regression



$$\text{Akurasi} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$
$$\frac{32 + 42}{32 + 42 + 3 + 2}$$
$$\frac{74}{79}$$

**Akurasi = 0,93**

$$\text{Presisi} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$
$$\frac{32}{32 + 2}$$
$$\frac{32}{34}$$

**Presisi Irit = 0,94**

# HASIL

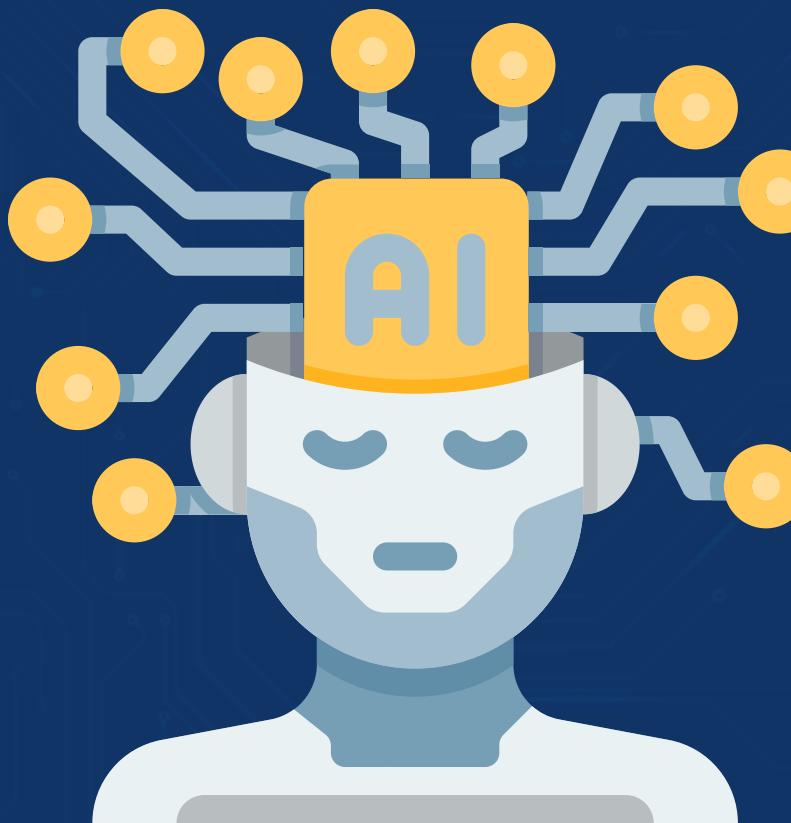
## ==== HASIL ANALISIS ====

Nama Mobil : Toyota Fortuner 2.4 Diesel  
Horsepower : 148.0  
Berat Kendaraan : 4188.78 lbs  
Kapasitas Mesin : 146.03 cu in  
Hasil Prediksi : IRIT

## ==== HASIL ANALISIS ====

Nama Mobil : Mitsubishi Pajero Sport 3.0  
Horsepower : 220.0  
Berat Kendaraan : 4629.70 lbs  
Kapasitas Mesin : 183.07 cu in  
Hasil Prediksi : BOROS

# KESIMPULAN



1. Berat kendaraan dan kapastas memiliki pengaruh yang signifikan terhadap efisiensi bahan bakar
2. Model Logistic Regression mampu mengklasifikasikan mobil irit dan boros dengan performa yang baik
3. Berdasarkan hasil tersebut, kendaraan dengan berat yang ringan dan jumlah silinder lebih sedikit direkomendasikan sebagai pilihan yang lebih hemat bahan bakar dan ramah lingkungan

## SARAN

---

1. Untuk masyarakat yang peduli lingkungan dan sedang mencari mobil dengan konsumsi bahan bakar sedikit, disarankan untuk memilih mobil dengan kapasitas mesin yang rendah seperti Nissan Grand Livina, dan Toyota Avanza.
2. Untuk produsen mobil yang ada di indonesia, disarankan untuk lebih banyak membuat mobil yang ramah lingkungan dan mulailah beralih ke kendaraan listrik demi lingkungan yang lebih sehat.
3. Untuk peneliti yang akan melanjutkan penelitian ini, masih banyak variabel dan fitur yang dapat ditambahkan untuk memperkuat model. Variabel kecil seperti konsumsi bahan bakar di jalan tol dan di jalan biasa, kemacetan pada jalan raya, dan karakteristik pengguna kendaraan dapat mempengaruhi perilaku model dalam menangkap pola di sebuah dataset saat pelatihan.

**SEKIAN  
TERIMAKASIH**