

```

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

# Sample car dataset
data = {
    "Year": [2010, 2012, 2015, 2018, 2020, 2011, 2016, 2019],
    "Mileage": [80000, 60000, 40000, 30000, 20000, 75000, 35000,
25000],
    "EngineSize": [1.2, 1.5, 1.6, 2.0, 2.2, 1.3, 1.8, 2.0],
    "Price": [300000, 400000, 550000, 800000, 1000000, 320000, 700000,
900000]
}

df = pd.DataFrame(data)
df

{
  "summary": {
    "name": "df",
    "rows": 8,
    "fields": [
      {
        "column": "Year",
        "properties": {
          "dtype": "number",
          "std": 3,
          "min": 2010,
          "max": 2020,
          "num_unique_values": 8,
          "samples": [
            2012, 2011, 2010
          ],
          "semantic_type": "\",
          "description": "\"
        }
      },
      {
        "column": "Mileage",
        "properties": {
          "dtype": "number",
          "std": 23058,
          "min": 20000,
          "max": 80000,
          "num_unique_values": 8,
          "samples": [
            60000, 75000, 80000
          ],
          "semantic_type": "\",
          "description": "\"
        }
      },
      {
        "column": "EngineSize",
        "properties": {
          "dtype": "number",
          "std": 0.35856858280031817,
          "min": 1.2,
          "max": 2.2,
          "num_unique_values": 7,
          "samples": [
            1.2, 1.5, 1.3
          ],
          "semantic_type": "\",
          "description": "\"
        }
      },
      {
        "column": "Price",
        "properties": {
          "dtype": "number",
          "std": 269201,
          "min": 300000,
          "max": 1000000,
          "num_unique_values": 8,
          "samples": [
            400000, 320000, 300000
          ],
          "semantic_type": "\",
          "description": "\"
        }
      }
    ],
    "type": "dataframe",
    "variable_name": "df"
  }
}

X = df[["Year", "Mileage", "EngineSize"]]
y = df["Price"]

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.25, random_state=42
)

```

```
model = LinearRegression()
model.fit(X_train, y_train)

LinearRegression()

y_pred = model.predict(X_test)

print("Mean Squared Error:", mean_squared_error(y_test, y_pred))
print("R2 Score:", r2_score(y_test, y_pred))

Mean Squared Error: 1741315806.092752
R2 Score: -0.08832237880796989

# New car details
new_car = pd.DataFrame(
    [[2017, 45000, 1.8]],
    columns=["Year", "Mileage", "EngineSize"]
)

predicted_price = model.predict(new_car)
print("Predicted Car Price:", int(predicted_price[0]))

Predicted Car Price: 777966
```