```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

import pickle



# Load dataset (Boston Housing from sklearn - deprecated, so use seaborn)
df = sns.load_dataset('mpg').dropna()

# For simplicity, predict 'mpg' using a few features
df = df[['mpg', 'horsepower', 'weight', 'acceleration']]
df.dropna(inplace=True)

df.head()
```

	mpg	horsepower	weight	acceleration
0	18.0	130.0	3504	12.0
1	15.0	165.0	3693	11.5
2	18.0	150.0	3436	11.0
3	16.0	150.0	3433	12.0
4	17.0	140.0	3449	10.5

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
X = df.drop('mpg', axis=1)
y = df['mpg']

# Split into train/test
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)



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

# Evaluation
y_pred = model.predict(X_test)
print("R2 Score:", r2_score(y_test, y_pred))
print("RMSE:", np.sqrt(mean_squared_error(y_test, y_pred)))
```

```
R2 Score: 0.6510068285906098
RMSE: 4.220523073680038
```

```
with open('model.pkl', 'wb') as f:
    pickle.dump(model, f)
```

```
from google.colab import files
files.download('model.pkl')
```