

```
# 1.) Importing the necessary files.
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import accuracy_score, mean_absolute_error
```

```
# 2.) Loading the cleaned dataset.
df = pd.read_csv("/content/cleaned_student_data.csv")
```

```
# 3.) Converting the categoricals to numericals.
for col in df.columns:
    if df[col].dtype == 'object':
        df[col] = df[col].astype('category').cat.codes
```

```
# 4.) Features & Target.
X = df.iloc[:, :-1]
y = df.iloc[:, -1]

print(" ♦ Features shape:", X.shape)
print(" ♦ Target shape:", y.shape)
```

```
♦ Features shape: (2392, 14)
♦ Target shape: (2392,)
```

```
# 5.) Splitting the dataset.
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)

print("\n✅ Data split into training & testing")
```

✅ Data split into training & testing

```
# 6.) Training the model.
model = LinearRegression()
model.fit(X_train, y_train)
```

```
▼ LinearRegression ⓘ ?
LinearRegression()
```

```
# 7.) Predictions and evaluations.  
predictions = model.predict(X_test)  
  
print("\n ♦ Sample Predictions:")  
print(predictions[:5])  
  
mae = mean_absolute_error(y_test, predictions)  
print("\n 🌈 Mean Absolute Error:", mae)
```

```
 ♦ Sample Predictions:  
[3.70929533 1.96919876 2.67559461 1.13937065 4.79012187]
```

```
 🌈 Mean Absolute Error: 0.5435836860967821
```