

```
main.py × insurance_data.csv +
main.py > ...
1 import pandas as pd
2 from matplotlib import pyplot as plt
3 from sklearn.model_selection import train_test_split
4 from sklearn.linear_model import LogisticRegression
5
6
7 df = pd.read_csv("insurance_data.csv")
8 print(df.head())
9 plt.scatter(df.age, df.bought_insurance, marker='+', color='red')
10 X_train, X_test, y_train, y_test = train_test_split(df[['age']], df.bought_insurance, train_size=0.8)
11
12 print(X_test)
13 print(y_test)
14
15 model = LogisticRegression()
16 model.fit(X_train, y_train)
17
18 y_predicted = model.predict(X_test)
19 print(y_predicted)
20
21 print(model.predict_proba(X_test))
22 print(model.score(X_test, y_test))
```

> Console × Shell +

Run Ask AI 11s on 11:55:26, 05/09 ✓

```
   age  bought_insurance
0    22                0
1    25                0
2    47                1
3    52                0
4    46                1
   age
7    60
5    56
3    52
23   45
0    22
12   27
[1 1 1 1 0 0]
[[0.03243911 0.96756089]
 [0.05670928 0.94329072]
 [0.09731155 0.90268845]
 [0.23049819 0.76950181]
 [0.89588972 0.10411028]
 [0.80571369 0.19428631]]
0.8333333333333334
```