File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

OUTLINE

TIMELINE

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All Restart Clear All Outputs | Jupyter Variables Outline ... .venv (Python 3.11.9)

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
data=pd.read_csv("C:\intern.py\Insurance.csv")
data
```

[90] Python

	age	sex	bmi	children	smoker	region	charges	insuranceclaim
0	19	0	27.900	0	1	3	16884.92400	1
1	18	1	33.770	1	0	2	1725.55230	1
2	28	1	33.000	3	0	2	4449.46200	0
3	33	1	22.705	0	0	1	21984.47061	0
4	32	1	28.880	0	0	1	3866.85520	1
...	...	...	...	...	...	...	...	...
1333	50	1	30.970	3	0	1	10600.54830	0
1334	18	0	31.920	0	0	0	2205.98080	1
1335	18	0	36.850	0	0	2	1629.83350	1
1336	21	0	25.800	0	0	3	2007.94500	0
1337	61	0	29.070	0	1	1	29141.36030	1

1338 rows × 8 columns

```
data.isnull()
```

[91] Python

	age	sex	bmi	children	smoker	region	charges	insuranceclaim

0 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

07:58 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All Restart Clear All Outputs | Jupyter Variables Outline ... .venv (Python 3.11.9)

[91]

```
data.isnull()
```

Python

	age	sex	bmi	children	smoker	region	charges	insuranceclaim
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...
1333	False	False	False	False	False	False	False	False
1334	False	False	False	False	False	False	False	False
1335	False	False	False	False	False	False	False	False
1336	False	False	False	False	False	False	False	False
1337	False	False	False	False	False	False	False	False

1338 rows × 8 columns

[92]

```
data.isnull().sum()
```

Python

```
age          0
sex          0
bmi          0
children     0
smoker       0
region       0
```

OUTLINE

TIMELINE

0 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

07:59 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All Restart Clear All Outputs | Jupyter Variables Outline ... .venv (Python 3.11.9)

[92]

```
data.isnull().sum()
```

Python

```
age             0
sex             0
bmi             0
children        0
smoker          0
region          0
charges         0
insuranceclaim  0
dtype: int64
```

[93]

```
feature_columns=['age','sex','bmi','children','smoker','region','charges']
x=data[feature_columns]
x
```

Python

	age	sex	bmi	children	smoker	region	charges
0	19	0	27.900	0	1	3	16884.92400
1	18	1	33.770	1	0	2	1725.55230
2	28	1	33.000	3	0	2	4449.46200
3	33	1	22.705	0	0	1	21984.47061
4	32	1	28.880	0	0	1	3866.85520
...	...	...	...	...	...	...	...
1333	50	1	30.970	3	0	1	10600.54830
1334	18	0	31.920	0	0	0	2205.98080
1335	18	0	36.850	0	0	2	1629.83350

OUTLINE

TIMELINE

English (United States)
English (India)

To switch input methods, press Windows key + space.

⊗ 0 ⚠ 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:01 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All Restart Clear All Outputs | Jupyter Variables Outline ... .venv (Python 3.11.9)

[93]

```
feature_columns=['age','sex','bmi','children','smoker','region','charges']
x=data[feature_columns]
x
```

Python

	age	sex	bmi	children	smoker	region	charges
0	19	0	27.900	0	1	3	16884.92400
1	18	1	33.770	1	0	2	1725.55230
2	28	1	33.000	3	0	2	4449.46200
3	33	1	22.705	0	0	1	21984.47061
4	32	1	28.880	0	0	1	3866.85520
...	...	...	...	...	...	...	...
1333	50	1	30.970	3	0	1	10600.54830
1334	18	0	31.920	0	0	0	2205.98080
1335	18	0	36.850	0	0	2	1629.83350
1336	21	0	25.800	0	0	3	2007.94500
1337	61	0	29.070	0	1	1	29141.36030

1338 rows × 7 columns

[94]

```
y=data['insuranceclaim']
y
```

Python

```
0       1
1       1
2       0
3       0
```

OUTLINE

TIMELINE

0 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:01 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All Restart Clear All Outputs | Jupyter Variables Outline ... .venv (Python 3.11.9)

[94]
```python
y=data['insuranceclaim']
y
```
Python

```
0       1
1       1
2       0
3       0
4       1
       ..
1333    0
1334    1
1335    1
1336    0
1337    1
Name: insuranceclaim, Length: 1338, dtype: int64
```

[95]
```python
sns.countplot(x="insuranceclaim",hue="charges",data=data)
```
Python

```
<Axes: xlabel='insuranceclaim', ylabel='count'>
```

OUTLINE

TIMELINE

0 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:01 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

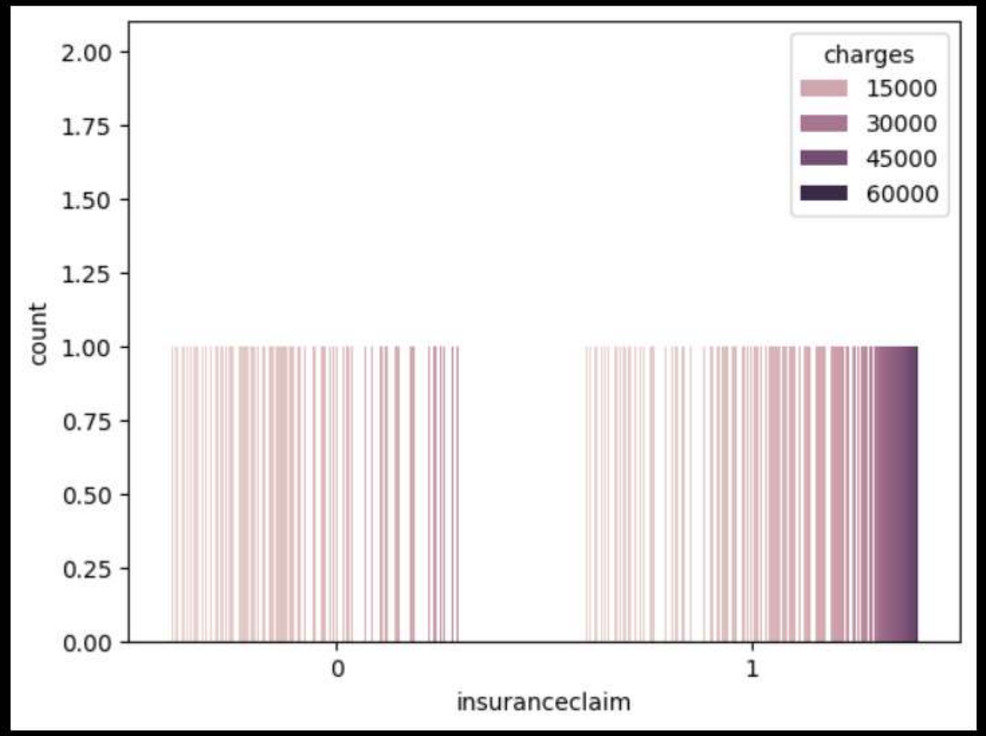
Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All | Restart | Clear All Outputs | Jupyter Variables | Outline | ... | .venv (Python 3.11.9)

```
sns.countplot(x="insuranceclaim",hue="charges",data=data)
```

[95] Python

```
<Axes: xlabel='insuranceclaim', ylabel='count'>
```



Snipping Tool

Screenshot copied to clipboard
Automatically saved to screenshots folder.

Mark-up and share

```
from sklearn.model_selection import train_test_split
```

OUTLINE

TIMELINE

⊗ 0 ⚠ 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:01 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All | Restart | Clear All Outputs | Jupyter Variables | Outline ... .venv (Python 3.11.9)

Generate + Code + Markdown

```
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score
from sklearn.preprocessing import LabelEncoder
```

[114] Python

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2)
```

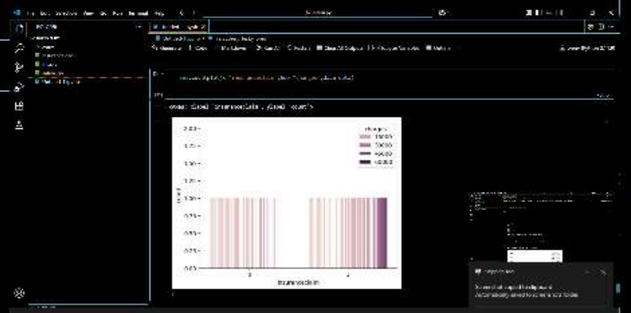
[116]

```
from sklearn.neighbors import KNeighborsClassifier
classifier=KNeighborsClassifier(n_neighbors=3)
classifier.fit(x_train,y_train)
y_pred=classifier.predict(x_test)
from sklearn.metrics import accuracy_score
accuracy_score(y_test,y_pred)
```

[117]

```
0.6156716417910447
```

```
model_knn = KNeighborsClassifier(n_neighbors=3)
```



Snipping Tool

Screenshot copied to clipboard
Automatically saved to screenshots folder.

Mark-up and share

OUTLINE

TIMELINE

0 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:01 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

OUTLINE

TIMELINE

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate | Code | Markdown | Run All | Restart | Clear All Outputs | Jupyter Variables | Outline | .venv (Python 3.11.9)

```
model_knn = KNeighborsClassifier(n_neighbors=3)
model_knn.fit(x_train, y_train)
y_pred_knn = model_knn.predict(x_test)
print("KNN Accuracy:", accuracy_score(y_test, y_pred_knn))
```

[118] Python

```
KNN Accuracy: 0.6156716417910447
```

```
from sklearn.ensemble import RandomForestClassifier
model_rf =RandomForestClassifier(n_estimators=100, random_state=42)
model_rf.fit(x_train, y_train)
y_pred_rf = model_rf.predict(x_test)
from sklearn.metrics import accuracy_score
accuracy_score(y_test,y_pred)
print("Random Forest Accuracy:", accuracy_score(y_test,y_pred))
```

[119] Python

```
Random Forest Accuracy: 0.6156716417910447
```

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score
model = DecisionTreeClassifier()
model.fit(x_train, y_train)
y_pred = model.predict(x_test)
print("Decision Tree Accuracy:", accuracy_score(y_test, y_pred))
```

[121] Python

```
Decision Tree Accuracy: 0.9738805970149254
```

0 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:02 24-07-2025

File Edit Selection View Go Run Terminal Help

intern.py

EXPLORER

INTERN.PY

- .venv
- Insurance.csv
- Iris.csv
- salary.csv
- Untitled-1.ipynb

Untitled-1.ipynb

Untitled-1.ipynb > r2_score(y_test,y_pred)

Generate + Code + Markdown | Run All | Restart | Clear All Outputs | Jupyter Variables | Outline | .venv (Python 3.11.9)

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score
model = DecisionTreeClassifier()
model.fit(x_train, y_train)
y_pred = model.predict(x_test)
print("Decision Tree Accuracy:", accuracy_score(y_test, y_pred))
```

[121] Python

```
Decision Tree Accuracy: 0.9738805970149254
```

```
reg=LinearRegression()
print("MDE",mean_absolute_error(y_test,y_pred))
```

[127] Python

```
MDE 0.026119402985074626
```

Generate + Code + Markdown

```
print("MSE",mean_squared_error(y_test,y_pred))
```

[128] Python

```
MSE 0.026119402985074626
```

```
r2_score(y_test,y_pred)
```

[129] Python

```
0.8903885480572598
```

OUTLINE

TIMELINE

0 ⚠ 5

Cell 68 of 68

25°C Light rain

Search

ENG IN

08:02 24-07-2025