

```
import pandas as pd
shows=pd.read_csv('/content/tv_shows (1).csv')
type(shows)
shows.head()
```

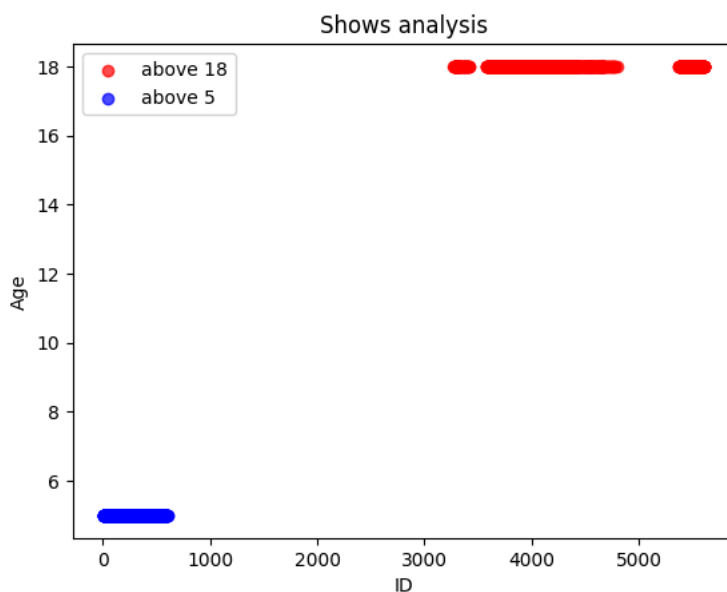
	Unnamed: 0	ID	Title	Year	Age	IMDb	Rotten Tomatoes	Netflix	Hulu	Prime Video	Disney+	Type
0	0	1	Breaking Bad	2008	5	1.1/10	100/100	1	0	0	0	1
1	1	2	Stranger Things	2016	5	1.5/10	96/100	1	0	0	0	1
2	2	3	Attack on Titan	2013	5	1.8/10	95/100	1	1	0	0	1
3	3	4	Better Call Saul	2015	5	1.8/10	94/100	1	0	0	0	1
4	4	5	Dark	2017	5	2.3/10	93/100	1	0	0	0	1

```
shows['IMDb']=shows['IMDb'].str.replace('/10',' ').astype(float)
```

```
x= shows.iloc[:,[1,5]].values
type(x)
y=shows.iloc[:,4].values
type(y)
print(y)
```

```
[ 5  5  5 ... 18 18 18]
```

```
import matplotlib.pyplot as plt
class1=shows[shows.Age == 18]
class2=shows[shows.Age == 5]
plt.title("Shows analysis")
plt.xlabel("ID")
plt.ylabel("Age")
plt.scatter(class1.ID, class1.Age, color= "red", label= "above 18", alpha=0.7)
plt.scatter(class2.ID, class2.Age, color= "blue", label= "above 5", alpha=0.7)
plt.legend()
plt.show()
```



```
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(x,y, test_size=0.4, random_state=1)
```

```
from sklearn.naive_bayes import GaussianNB
classifier= GaussianNB()
```

```
classifier.fit(X_train, Y_train)
GaussianNB()
```

```
▼ GaussianNB
GaussianNB()
```

```
y_pred= classifier.predict(X_test)
```

```
from sklearn.metrics import confusion_matrix
ConfMat=confusion_matrix(Y_test, y_pred)
print(ConfMat)
```

```
[[219  0  0  0  0]
 [ 3 336  3  0  0]
 [ 0  0  5  0  0]
 [ 0  0  3 398  0]
 [ 0  0  0  7 309]]
```

```
from sklearn.metrics import classification_report
print(classification_report(Y_test, y_pred))
```

	precision	recall	f1-score	support
5	0.99	1.00	0.99	219
7	1.00	0.98	0.99	342
13	0.45	1.00	0.62	5
16	0.98	0.99	0.99	401
18	1.00	0.98	0.99	316
accuracy			0.99	1283
macro avg	0.88	0.99	0.92	1283
weighted avg	0.99	0.99	0.99	1283