import numpy as np
import pandas as pd
df=pd.read\_csv("/content/titanic\_data.csv")
df

₽		PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	
	0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	,
	1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th	female	38.0	1	0	PC 17599	7
	2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	
	3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	5
	4										•

df1=df.groupby("Survived")['Survived'].count()
df1

Survived 0 549 1 342

Name: Survived, dtype: int64

df.head()

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cat
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	N
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs	female	38.0	1	0	PC 17599	71.2833	С

df.tail()

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.00	NaN
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.00	B42
				lohnoton							

df.columns

df.shape

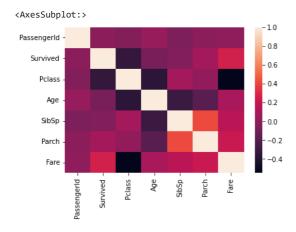
df.describe()

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

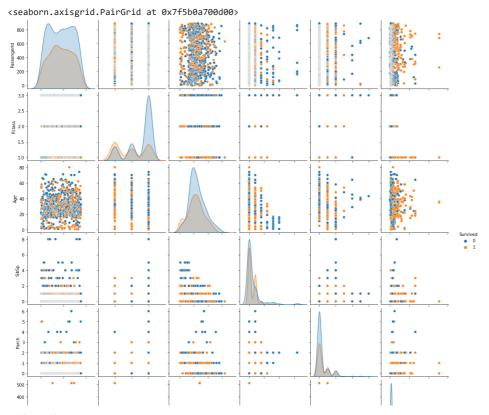
df.corr()

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
Passengerld	1.000000	-0.005007	-0.035144	0.036847	-0.057527	-0.001652	0.012658
Survived	-0.005007	1.000000	-0.338481	-0.077221	-0.035322	0.081629	0.257307
Pclass	-0.035144	-0.338481	1.000000	-0.369226	0.083081	0.018443	-0.549500
Age	0.036847	-0.077221	-0.369226	1.000000	-0.308247	-0.189119	0.096067
SibSp	-0.057527	-0.035322	0.083081	-0.308247	1.000000	0.414838	0.159651
Parch	-0.001652	0.081629	0.018443	-0.189119	0.414838	1.000000	0.216225
Fare	0.012658	0.257307	-0.549500	0.096067	0.159651	0.216225	1.000000

import seaborn as sns
sns.heatmap(df.corr())



sns.pairplot(df,hue='Survived')



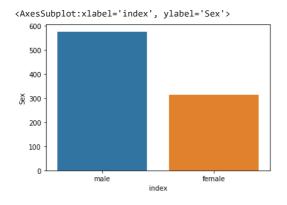
df.isna().sum()

PassengerId Survived Pclass 0 Name 0 0 Sex 177 Age SibSp 0 Parch 0 Ticket 0 Fare 0 Cabin 687 Embarked dtype: int64

sexx=df['Sex'].value\_counts().reset\_index()
sexx

	ındex	Sex
0	male	577
1	female	314

import seaborn as sns
sns.barplot(x='index',y='Sex',data=sexx)

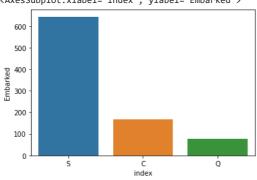


embarked=df['Embarked'].value\_counts().reset\_index()
embarked

	index	Embarked
0	S	644
1	С	168

sns.barplot(x='index',y='Embarked',data=embarked)

<AxesSubplot:xlabel='index', ylabel='Embarked'>



df['Cabin'].value\_counts()

B96	B98		4
G6			4
C23	C25	C27	4
C22	C26		3
F33			3
			• •
E34			1
E34 C7			1 1
			_
C7			1

Name: Cabin, Length: 147, dtype: int64

dummy=pd.get\_dummies(df[['Sex','Embarked']],drop\_first=True)
dummy

	Sex_male	Embarked_Q	Embarked_S
0	1	0	1
1	0	0	0
2	0	0	1
3	0	0	1
4	1	0	1
886	1	0	1
887	0	0	1
888	0	0	1
889	1	0	0
890	1	1	0

891 rows × 3 columns

dfe=pd.concat([df,dummy],axis=1)
dfe

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	(
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
				Allon Mr							

```
dfe.columns
```

## dfe.isna().sum()

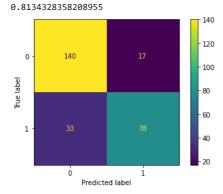
```
PassengerId
Survived
Pclass
                 0
Name
                  0
Sex
                177
Age
SibSp
                 0
Parch
Ticket
                 0
                 0
Fare
Cabin
                687
Embarked
Sex_male
                 0
Embarked_Q
Embarked_S
dtype: int64
                 0
                  0
```

dfe=dfe.drop(['Name','Sex','Embarked','Cabin','Ticket'],axis=True)
dfe

```
PassengerId Survived Pclass Age SibSp Parch
                                                               Fare Sex_male Embarked_Q Embarke
                                      3
                                         22.0
                                                             7.2500
                                                                                        0
                     2
       1
                               1
                                      1
                                        38.0
                                                  1
                                                         0 71.2833
                                                                           0
                                                                                       0
dfe.isna().sum()
     PassengerId
                      a
     Survived
                      a
     Pclass
                      0
                    177
     Age
     SibSp
                      0
     Parch
                      0
                      0
     Fare
     Sex_male
     Embarked_Q
                      0
     Embarked S
                      0
     dtype: int64
      ممه
                                      2 22 0
                                                         0 77500
                   001
dfe['Age']=dfe['Age'].fillna(dfe['Age'].mode()[0])
dfe.isna().sum()
     PassengerId
     Survived
     Pclass
                    0
                    0
     Age
     SibSp
                    0
     Parch
                    0
     Fare
                    a
     Sex_male
                    a
     Embarked_Q
                    0
     Embarked_S
                    0
     dtype: int64
x=dfe.drop(['Survived'],axis=True).values
Х
                     3.,
                                      1.,
                                            0.,
                                                  1.],
     array([[ 1.,
                         22., ...,
               2.,
                     1.,
                                      0.,
                                            0.,
                          38., ...,
                                                  0.],
              3.,
                     3.,
                          26., ...,
                                      0.,
                                            0.,
                                                  1.],
            [889.,
                     3.,
                          24., ...,
            [890.,
                     1.,
                          26., ...,
                                      1.,
                                            0.,
                                                  0.],
                                                  0.]])
                         32., ...,
            Γ891..
                     3..
                                      1.,
                                            1.,
y=dfe.iloc[:,1].values
     array([0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1,
            1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0,
            1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1,
                                                                 0,
                                                                     0,
            1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0,
                                                                 1, 1,
            1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1,
                                                                    1.
            0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1,
                                                                 1,
            0, 1, 0,
                     0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0,
                                                                 0,
                                                                     1,
              0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0,
                                                                 1,
              0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1,
               0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1,
                                                                     0,
            1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0,
              0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
            0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0,
              0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0,
            0, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0,
            1,
              0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0,
            0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1,
                                                                     0,
              1, 1,
                    0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0,
                                                                    1,
              0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0,
                                                                 1,
                                                                     0,
              0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1,
                                                                  0, 1,
               0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1,
            0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1,
              0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,
            1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0,
            0.
              0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0,
            1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1,
              1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0,
            0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 1,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 1,\ 1,\ 0,\ 0,
                                                                 1, 0,
               0, 1,
                     0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0,
                                                                  0,
              0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0,
                                                                  0,
              0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1,
              0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0,
            0, 0, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0,
              1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0,
                                                                 1, 0, 0, 1,
            0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0,
```

```
0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0,
            0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0,
            0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1,
            0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1,
            1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1,
            1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0])
dfe.dtypes
     PassengerId
                      int64
     Survived
                      int64
     Pclass
                      int64
                    float64
     SibSp
                      int64
     Parch
                      int64
     Fare
                    float64
     Sex_male
                      uint8
     Embarked_Q
                      uint8
     Embarked S
                      uint8
     dtype: object
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.30,random_state=42)
x_train
y_test
     array([1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1,
            1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0,
            0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 0,\ 1,\ 0,\ 1,\ 1,\ 1,\ 1,\ 0,\ 1,\ 1,\ 0,\ 0,
                     1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1,
            0, 0, 1,
                     0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1,
            1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1,
            0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 0,
            1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1,
            1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0,
            0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1,
            0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1,
            0. 0. 0. 11)
from sklearn.preprocessing import StandardScaler
scaler=StandardScaler()
scaler.fit(x_train)
x_train=scaler.transform(x_train)
x_test=scaler.transform(x_test)
x_train
x_test
     array([[ 1.01391396, 0.80326712, -0.31932616, ..., 0.72077194,
              -0.31117678, -1.65922031],
            [-0.02533086, -0.41730706, 0.21160173, ..., 0.72077194, -0.31117678, 0.60269272],
            [ 1.51814016, 0.80326712, -0.62271353, ..., 0.72077194, -0.31117678, 0.60269272],
            [0.04010307, -1.63788124, 2.79039434, ..., 0.72077194,
            -0.31117678, 0.60269272],
[ 0.20946149, 0.80326712, -0.85025405, ..., 0.72077194,
              -0.31117678, 0.60269272],
            [-0.05997235, -1.63788124, -0.01593879, ..., 0.72077194, -0.31117678, 0.60269272]])
from sklearn.neighbors import KNeighborsClassifier
model1=KNeighborsClassifier(n neighbors=19)
model1.fit(x_train,y_train)
y_pred=model1.predict(x_test)
y_pred
     array([0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0,
            0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0,
            0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1,
            0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1,
            0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,
                     0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0,
            1, 0, 1,
            0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1,
            0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0,
            0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1,
            1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0,
            0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1,
            0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0,
            0, 0, 0, 0])
```

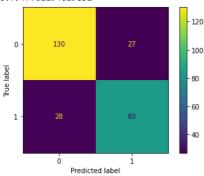
```
from sklearn.metrics import classification_report,confusion_matrix,accuracy_score,ConfusionMatrixDisplay,confusion_matrix
report=classification_report(y_test,y_pred)
con_mat=confusion_matrix(y_test,y_pred)
cm1=ConfusionMatrixDisplay(con_mat)
cm1.plot()
score=accuracy_score(y_test,y_pred)
score
```



#naive\_bayes
from sklearn.naive\_bayes import GaussianNB
model2=GaussianNB()
model2.fit(x\_train,y\_train)
y\_pred2=model2.predict(x\_test)
y\_pred2

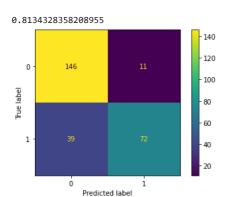
from sklearn.metrics import classification\_report,confusion\_matrix,accuracy\_score,ConfusionMatrixDisplay,confusion\_matrix
report=classification\_report(y\_test,y\_pred2)
con\_mat=confusion\_matrix(y\_test,y\_pred2)
cm1=ConfusionMatrixDisplay(con\_mat)
cm1.plot()
score=accuracy\_score(y\_test,y\_pred2)
score

## 0.7947761194029851



#svm
from sklearn.svm import SVC
model3=SVC()
model3.fit(x\_train,y\_train)
y\_pred3=model3.predict(x\_test)
y\_pred3

from sklearn.metrics import classification\_report,confusion\_matrix,accuracy\_score,ConfusionMatrixDisplay,confusion\_matrix
report=classification\_report(y\_test,y\_pred3)
con\_mat=confusion\_matrix(y\_test,y\_pred3)
cm1=ConfusionMatrixDisplay(con\_mat)
cm1.plot()
score=accuracy\_score(y\_test,y\_pred3)
score

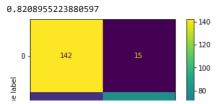


from sklearn.metrics import classification\_report,confusion\_matrix,accuracy\_score
report=classification\_report(y\_test,y\_pred)
print(report)

	precision	recall	f1-score	support
0	0.81	0.89	0.85	157
1	0.82	0.70	0.76	111
accuracy			0.81	268
macro avg	0.82	0.80	0.80	268
weighted avg	0.81	0.81	0.81	268

#random forest
from sklearn.ensemble import RandomForestClassifier
model4=RandomForestClassifier()
model4.fit(x\_train,y\_train)
y\_pred5=model4.predict(x\_test)
y\_pred5

from sklearn.metrics import classification\_report,confusion\_matrix,accuracy\_score,ConfusionMatrixDisplay,confusion\_matrix
report=classification\_report(y\_test,y\_pred5)
con\_mat=confusion\_matrix(y\_test,y\_pred5)
cm1=ConfusionMatrixDisplay(con\_mat)
cm1.plot()
score=accuracy\_score(y\_test,y\_pred5)
score



#decition tree
from sklearn.tree import DecisionTreeClassifier
model5=DecisionTreeClassifier()
model5.fit(x\_train,y\_train)
y\_pred6=model4.predict(x\_test)
y\_pred6

from sklearn.metrics import classification\_report,confusion\_matrix,accuracy\_score,ConfusionMatrixDisplay,confusion\_matrix
report=classification\_report(y\_test,y\_pred6)
con\_mat=confusion\_matrix(y\_test,y\_pred6)
cm1=ConfusionMatrixDisplay(con\_mat)
cm1.plot()
score=accuracy\_score(y\_test,y\_pred6)
score

