```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sn
read=pd.read csv(r"D:\titanic\train.csv")
read
     PassengerId
                   Survived
                             Pclass
0
                1
                2
                                   1
1
                          1
2
                                   3
                3
                          1
3
                4
                                   1
                          1
4
                5
                                   3
                          0
886
             887
                          0
                                   2
                                   1
887
             888
                          1
                                   3
888
             889
                          0
                                   1
             890
                          1
889
890
                                   3
             891
                                                     Name
                                                              Sex
                                                                     Age
SibSp \
                                 Braund, Mr. Owen Harris
                                                             male
                                                                   22.0
1
1
     Cumings, Mrs. John Bradley (Florence Briggs Th...
                                                           female
                                                                   38.0
1
2
                                  Heikkinen, Miss. Laina
                                                           female
                                                                   26.0
0
3
          Futrelle, Mrs. Jacques Heath (Lily May Peel)
                                                           female
                                                                    35.0
1
4
                               Allen, Mr. William Henry
                                                             male
                                                                   35.0
0
. .
                                                               . . .
                                                                     . . .
886
                                   Montvila, Rev. Juozas
                                                             male
                                                                   27.0
0
887
                           Graham, Miss. Margaret Edith
                                                           female
                                                                    19.0
              Johnston, Miss. Catherine Helen "Carrie"
888
                                                           female
                                                                     NaN
1
889
                                   Behr, Mr. Karl Howell
                                                             male
                                                                   26.0
890
                                     Dooley, Mr. Patrick
                                                                   32.0
                                                             male
0
     Parch
                       Ticket
                                   Fare Cabin Embarked
0
                    A/5 21171
                                 7.2500
                                          NaN
         0
                                                      S
                                                      C
1
                     PC 17599
                                71.2833
                                          C85
2
            STON/02. 3101282
                                                      S
                                 7.9250
                                          NaN
```

```
3
                               53.1000
                                         C123
                                                      S
         0
                       113803
4
         0
                       373450
                               8.0500
                                          NaN
                                                      S
                                          . . .
                                                      S
886
         0
                       211536
                               13.0000
                                          NaN
                                                      S
                                          B42
887
                       112053
                               30.0000
         0
                                                      S
                   W./C. 6607
888
         2
                               23.4500
                                          NaN
889
                       111369
                               30.0000
                                         C148
                                                      C
         0
890
         0
                       370376
                                7.7500
                                          NaN
                                                      Q
```

[891 rows x 12 columns]

read.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):

Column	Non-Null Count	Dtype
PassengerId	891 non-null	int64
Survived	891 non-null	int64
Pclass	891 non-null	int64
Name	891 non-null	object
Sex	891 non-null	object
Age	714 non-null	float64
SibSp	891 non-null	int64
Parch	891 non-null	int64
Ticket	891 non-null	object
Fare	891 non-null	float64
Cabin	204 non-null	object
Embarked	889 non-null	object
es: float64(2), int64(5), obj	ect(5)
	PassengerId Survived Pclass Name Sex Age SibSp Parch Ticket Fare Cabin Embarked	PassengerId 891 non-null Survived 891 non-null Pclass 891 non-null Name 891 non-null Sex 891 non-null Age 714 non-null SibSp 891 non-null Parch 891 non-null Ticket 891 non-null Fare 891 non-null Cabin 204 non-null

memory usage: 83.7+ KB

read.describe()

	PassengerId	Survived	Pclass	Age	SibSp	\
count	891.000000	891.000000	891.000000	714.000000	891.000000	
mean	446.000000	0.383838	2.308642	29.699118	0.523008	
std	257.353842	0.486592	0.836071	14.526497	1.102743	
min	1.000000	0.000000	1.000000	0.420000	0.000000	
25%	223.500000	0.000000	2.000000	20.125000	0.000000	
50%	446.000000	0.000000	3.000000	28.000000	0.000000	
75%	668.500000	1.000000	3.000000	38.000000	1.000000	
max	891.000000	1.000000	3.000000	80.000000	8.000000	

	Parch	Fare
count	891.000000	891.000000
mean	0.381594	32.204208
std	0.806057	49.693429
min	0.000000	0.000000
25%	0.000000	7.910400

```
50%
         0.000000
                     14.454200
75%
         0.000000
                     31.000000
         6.000000
max
                    512.329200
read.isnull().sum()
PassengerId
                  0
Survived
                  0
Pclass
                  0
                  0
Name
Sex
                  0
Age
                177
SibSp
                  0
                  0
Parch
Ticket
                  0
                  0
Fare
Cabin
                687
Embarked
                  2
dtype: int64
read['Age'].mean()
29.69911764705882
df=read['Age'].fillna(value=29)
df
0
       22.0
1
       38.0
2
       26.0
3
       35.0
4
       35.0
886
       27.0
       19.0
887
       29.0
888
889
       26.0
890
       32.0
Name: Age, Length: 891, dtype: float64
read['Age']=df
df
0
       22.0
1
       38.0
2
       26.0
3
       35.0
4
       35.0
       . . .
886
       27.0
887
       19.0
888
       29.0
```

```
889
       26.0
890
       32.0
Name: Age, Length: 891, dtype: float64
df.isnull().sum()
0
read.isnull().sum()
PassengerId
                  0
Survived
                  0
Pclass
                  0
Name
                  0
Sex
                  0
Age
                  0
SibSp
                  0
Parch
                  0
Ticket
                  0
Fare
                  0
Cabin
                687
Embarked
                  2
dtype: int64
dd=read.drop('Cabin',axis=1)
dd
     PassengerId
                   Survived
                             Pclass
0
                1
                                   3
                2
                          1
                                   1
1
2
                3
                          1
                                   3
3
                4
                          1
                                   1
4
                5
                          0
                                   3
                                   2
886
             887
                          0
                                   1
887
             888
                          1
888
             889
                          0
                                   3
                                   1
889
             890
                          1
890
             891
                                   3
                                                     Name
                                                               Sex
                                                                     Age
SibSp \
                                 Braund, Mr. Owen Harris
                                                             male 22.0
1
1
     Cumings, Mrs. John Bradley (Florence Briggs Th...
                                                           female
                                                                   38.0
1
2
                                  Heikkinen, Miss. Laina
                                                           female 26.0
0
3
          Futrelle, Mrs. Jacques Heath (Lily May Peel)
                                                                   35.0
                                                           female
1
4
                               Allen, Mr. William Henry
                                                             male 35.0
```

```
. .
886
                                   Montvila, Rev. Juozas
                                                              male 27.0
0
887
                           Graham, Miss. Margaret Edith
                                                            female
                                                                    19.0
0
               Johnston, Miss. Catherine Helen "Carrie"
888
                                                            female
                                                                   29.0
1
                                   Behr, Mr. Karl Howell
889
                                                              male
                                                                    26.0
0
                                     Dooley, Mr. Patrick
890
                                                              male 32.0
0
     Parch
                       Ticket
                                   Fare Embarked
                    A/5 21171
0
                                 7.2500
                                                S
         0
                                                C
1
         0
                     PC 17599
                                71.2833
2
         0
            STON/02. 3101282
                                 7.9250
                                                S
                                                S
3
                       113803
                                53.1000
         0
                                                S
4
         0
                       373450
                                 8.0500
                                                S
                       211536
886
         0
                                13.0000
                                                S
887
         0
                       112053
                                30.0000
         2
                   W./C. 6607
                                                S
888
                                23.4500
                                                C
889
         0
                       111369
                                30.0000
                                                Q
890
                       370376
                                 7.7500
[891 rows x 11 columns]
dd.isnull().sum()
PassengerId
                0
Survived
                0
Pclass
                0
Name
                0
Sex
                0
Age
SibSp
                0
                0
Parch
Ticket
                0
Fare
                0
                2
Embarked
dtype: int64
dd
     PassengerId
                   Survived
                              Pclass \
0
                1
                          0
1
                2
                          1
                                   1
2
                3
                                   3
                          1
3
                4
                          1
                                   1
```

0

```
3
4
                5
                           0
886
              887
                           0
                                    2
887
                                    1
              888
                           1
                                    3
888
              889
                           0
889
              890
                           1
                                    1
                                    3
890
              891
                           0
                                                      Name
                                                                Sex
                                                                       Age
SibSp \
                                  Braund, Mr. Owen Harris
                                                                      22.0
0
                                                               male
1
1
     Cumings, Mrs. John Bradley (Florence Briggs Th...
                                                             female
                                                                      38.0
1
2
                                   Heikkinen, Miss. Laina
                                                             female
                                                                      26.0
0
           Futrelle, Mrs. Jacques Heath (Lily May Peel)
3
                                                             female
                                                                      35.0
1
                                Allen, Mr. William Henry
4
                                                               male
                                                                      35.0
0
. . .
                                    Montvila, Rev. Juozas
886
                                                               male
                                                                      27.0
0
887
                            Graham, Miss. Margaret Edith
                                                             female
                                                                      19.0
               Johnston, Miss. Catherine Helen "Carrie"
                                                             female
888
                                                                      29.0
1
889
                                    Behr, Mr. Karl Howell
                                                               male
                                                                      26.0
0
890
                                      Dooley, Mr. Patrick
                                                                     32.0
                                                               male
0
     Parch
                        Ticket
                                    Fare Embarked
                    A/5 21171
                                  7.2500
0
                                                 C
1
         0
                     PC 17599
                                 71.2833
                                                 S
2
             STON/02. 3101282
                                  7.9250
                                                 S
3
         0
                        113803
                                 53.1000
                                                 S
4
         0
                        373450
                                  8.0500
                                                 S
886
                        211536
                                 13.0000
                                                 S
                        112053
887
                                 30.0000
                                                 S
                   W./C. 6607
888
         2
                                 23.4500
                                                 C
889
         0
                        111369
                                 30.0000
                        370376
                                  7.7500
890
[891 rows x 11 columns]
pd.Categorical(dd['Sex'])
```

```
['male', 'female', 'female', 'male', ..., 'male', 'female',
'female', 'male', 'male']
Length: 891
Categories (2, object): ['female', 'male']
dd['Sex'].replace(['female', 'male'], [0,1], inplace=True)
dd
     PassengerId
                   Survived
                              Pclass
0
                1
                2
                                   1
                           1
1
                                   3
2
                3
                           1
                4
                                    1
3
                           1
4
                5
                           0
                                    3
                           0
                                   2
886
              887
                                   1
887
              888
                           1
                                   3
888
              889
                           0
                                   1
              890
                           1
889
                                   3
890
              891
                                                      Name
                                                            Sex
                                                                   Age
SibSp \
                                 Braund, Mr. Owen Harris
                                                                  22.0
                                                               1
1
     Cumings, Mrs. John Bradley (Florence Briggs Th...
1
                                                                 38.0
1
2
                                  Heikkinen, Miss. Laina
                                                                  26.0
0
3
          Futrelle, Mrs. Jacques Heath (Lily May Peel)
                                                                  35.0
1
4
                                Allen, Mr. William Henry
                                                                 35.0
0
. .
                                                                   . . .
886
                                   Montvila, Rev. Juozas
                                                                 27.0
0
887
                            Graham, Miss. Margaret Edith
                                                                  19.0
               Johnston, Miss. Catherine Helen "Carrie"
888
                                                                  29.0
1
889
                                   Behr, Mr. Karl Howell
                                                                  26.0
890
                                      Dooley, Mr. Patrick
                                                               1 32.0
0
     Parch
                       Ticket
                                   Fare Embarked
0
                    A/5 21171
                                 7.2500
                                                S
         0
                                                C
1
                     PC 17599
                                71.2833
2
             STON/02. 3101282
                                                 S
                                 7.9250
```

3	0	113803	53.1000	S
4	0	373450	8.0500	S
886	0	211536	13.0000	S
887	0	112053	30.0000	S
888	2	W./C. 6607	23.4500	S
889	0	111369	30.0000	С
890	0	370376	7.7500	Q

[891 rows x 11 columns]

ss=dd.drop(['Name','Ticket'],axis=1)
ss

	ssengerId	Survived	Pclass	Sex	Age	SibSp	Parch	Fare
Embarke		0	2	1	22.0	1	0	7 2500
0 S	1	0	3	1	22.0	1	0	7.2500
1	2	1	1	0	38.0	1	0	71.2833
C 2 S	3	1	3	0	26.0	0	Θ	7.9250
S 3 S	4	1	1	0	35.0	1	0	53.1000
4 S	5	Θ	3	1	35.0	0	0	8.0500
886 S	887	0	2	1	27.0	0	0	13.0000
887 S	888	1	1	0	19.0	0	0	30.0000
888 S	889	0	3	0	29.0	1	2	23.4500
889 C	890	1	1	1	26.0	0	0	30.0000
890 Q	891	0	3	1	32.0	0	0	7.7500

[891 rows x 9 columns]

SS

	_	Survived	Pclass	Sex	Age	SibSp	Parch	Fare
Embarke 0		0	3	1	22.0	1	0	7.2500
5	2	1	1	0	38.0	1	0	71.2833
2	3	1	3	0	26.0	0	0	7.9250
	2	1	1	0	38.0	1	Θ	71.2

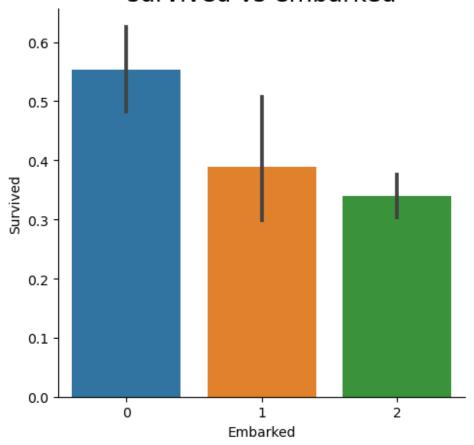
```
3
S
                4
                           1
                                            35.0
                                                        1
                                                                   53.1000
                                    1
                                         0
                                                               0
4
                5
                           0
                                    3
                                             35.0
                                                        0
                                                                    8.0500
                                         1
                                                               0
S
. .
. . .
              887
                           0
                                    2
                                         1
                                            27.0
                                                        0
                                                               0
                                                                   13.0000
886
S
887
              888
                           1
                                    1
                                         0
                                             19.0
                                                        0
                                                               0
                                                                   30.0000
S
                                    3
888
              889
                           0
                                             29.0
                                                        1
                                                               2
                                                                   23.4500
S
889
              890
                           1
                                    1
                                                                   30.0000
                                         1
                                             26.0
                                                        0
C
890
              891
                           0
                                    3
                                         1
                                            32.0
                                                        0
                                                               0
                                                                    7.7500
0
[891 rows x 9 columns]
ss.isnull().sum()
PassengerId
                0
Survived
                0
Pclass
                0
Sex
                0
                0
Age
SibSp
                0
Parch
                0
Fare
Embarked
                2
dtype: int64
ss['Embarked'].value counts()
S
     644
C
     168
0
      77
Name: Embarked, dtype: int64
ss['Embarked']=ss['Embarked'].fillna(value='S')
ss['Embarked']
       S
0
       C
1
       S
2
       S
3
       S
4
886
       S
887
       S
```

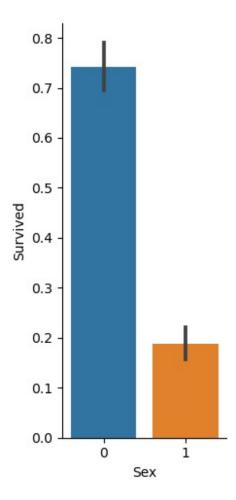
```
888
       S
889
       C
890
Name: Embarked, Length: 891, dtype: object
ss['Embarked'].isnull().sum()
0
SS
     PassengerId Survived Pclass
                                             Age SibSp Parch
                                                                      Fare
                                       Sex
Embarked
                1
                           0
                                    3
                                            22.0
                                                       1
                                                                   7.2500
0
                                         1
                                                               0
S
1
                2
                           1
                                    1
                                            38.0
                                                       1
                                                                  71.2833
C
2
                                    3
                3
                           1
                                         0
                                            26.0
                                                       0
                                                               0
                                                                   7.9250
S
3
                4
                           1
                                            35.0
                                                       1
                                    1
                                         0
                                                               0
                                                                  53.1000
Š
4
                5
                           0
                                    3
                                                                   8.0500
                                         1
                                            35.0
                                                       0
                                                               0
S
. .
              . . .
                                             . . .
                                                             . . .
                                                                       . . .
. . .
              887
                                    2
                                            27.0
                                                       0
                                                                  13.0000
886
                           0
                                         1
                                                               0
S
887
              888
                           1
                                    1
                                         0
                                            19.0
                                                       0
                                                               0
                                                                  30.0000
S
              889
                                    3
888
                           0
                                            29.0
                                                       1
                                                               2
                                                                  23.4500
S
889
              890
                           1
                                    1
                                            26.0
                                                                  30.0000
                                         1
                                                       0
C
890
              891
                                    3
                                         1
                                            32.0
                                                               0
                                                                   7.7500
                           0
                                                       0
[891 rows x 9 columns]
from sklearn import preprocessing
label encoder=preprocessing.LabelEncoder()
ss['Embarked']=label encoder.fit transform(ss['Embarked'])
ss['Embarked'].value counts()
2
     646
0
     168
1
      77
Name: Embarked, dtype: int64
```

SS

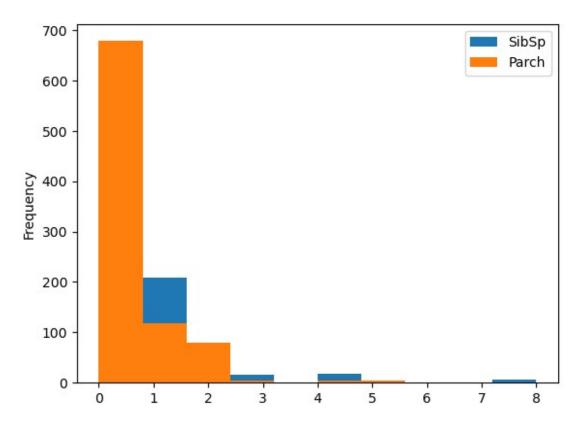
ГI. а	PassengerId	Survived	Pclass	Sex	Age	SibSp	Parch	Fare
0	arked 1	0	3	1	22.0	1	0	7.2500
2	2	1	1	0	38.0	1	0	71.2833
0	3	1	3	0	26.0	0	0	7.9250
2 2 3	4	1	1	0	35.0	1	Θ	53.1000
2 4 2	5	0	3	1	35.0	0	0	8.0500
886	887	0	2	1	27.0	0	0	13.0000
2 887	888	1	1	0	19.0	0	0	30.0000
2 888	889	0	3	0	29.0	1	2	23.4500
2 889	890	1	1	1	26.0	0	0	30.0000
0 890 1	891	0	3	1	32.0	0	0	7.7500
[891	rows x 9 col	umns]						
<pre>sn.catplot(x="Embarked",</pre>								

survived vs embarked





ss.plot(x='Embarked',y=['SibSp','Parch'],kind='hist')
<Axes: ylabel='Frequency'>



ss.corr()

	PassengerId	Survived	Pclass	Sex	Age	
SibSp \ PassengerId	1.000000	-0.005007	-0.035144	0.042939	0.033632	-
0.057527 Survived 0.035322	-0.005007	1.000000	-0.338481	-0.543351	-0.067814	-
Pclass 0.083081	-0.035144	-0.338481	1.000000	0.131900	-0.334974	
Sex 0.114631	0.042939	-0.543351	0.131900	1.000000	0.082949	-
Age 0.232978	0.033632	-0.067814	-0.334974	0.082949	1.000000	-
SibSp 1.000000	-0.057527	-0.035322	0.083081	-0.114631	-0.232978	
Parch 0.414838	-0.001652	0.081629		-0.245489		
Fare 0.159651	0.012658	0.257307	-0.549500	-0.182333	0.093706	
Embarked 0.068230	0.013128	-0.167675	0.162098	0.108262	-0.023464	

Parch Fare Embarked PassengerId -0.001652 0.012658 0.013128 Survived 0.081629 0.257307 -0.167675

Pclass 0.018443 -0.549500 0.162098 Sex -0.245489 -0.182333 0.108262 -0.176486 0.093706 -0.023464 Age SibSp 0.414838 0.159651 0.068230 Parch 1.000000 0.216225 0.039798 Fare 0.216225 1.000000 -0.224719 Embarked 0.039798 -0.224719 1.000000

sn.heatmap(ss.corr())

<Axes: >

