untitled

July 6, 2024

```python
[5]: import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     import seaborn as sns
```

```python
[12]: df=pd.read_csv("titanic.csv")
```

```python
[13]: df
```

```
[13]:      PassengerId  Survived  Pclass  \
     0            892         0       3
     1            893         1       3
     2            894         0       2
     3            895         0       3
     4            896         1       3
     ..           ...       ...     ...
     413         1305         0       3
     414         1306         1       1
     415         1307         0       3
     416         1308         0       3
     417         1309         0       3

                                                 Name     Sex   Age  SibSp  Parch  \
     0                                  Kelly, Mr. James    male  34.5      0      0
     1                  Wilkes, Mrs. James (Ellen Needs)  female  47.0      1      0
     2                         Myles, Mr. Thomas Francis    male  62.0      0      0
     3                                  Wirz, Mr. Albert    male  27.0      0      0
     4     Hirvonen, Mrs. Alexander (Helga E Lindqvist)  female  22.0      1      1
     ..                                             ...     ...   ...    ...    ...
     413                            Spector, Mr. Woolf    male   NaN      0      0
     414                   Oliva y Ocana, Dona. Fermina  female  39.0      0      0
     415                   Saether, Mr. Simon Sivertsen    male  38.5      0      0
     416                           Ware, Mr. Frederick    male   NaN      0      0
     417                      Peter, Master. Michael J    male   NaN      1      1

               Ticket      Fare Cabin Embarked
     0         330911    7.8292   NaN        Q
```

1

```
1                363272    7.0000    NaN      S
2                240276    9.6875    NaN      Q
3                315154    8.6625    NaN      S
4               3101298   12.2875    NaN      S
..                  ...       ...    ...    ...
413          A.5. 3236    8.0500    NaN      S
414           PC 17758  108.9000   C105      C
415  SOTON/O.Q. 3101262    7.2500    NaN      S
416               359309    8.0500    NaN      S
417                 2668   22.3583    NaN      C

[418 rows x 12 columns]
```

[14]: `df.head()`

```
[14]:    PassengerId  Survived  Pclass  \
      0          892         0       3
      1          893         1       3
      2          894         0       2
      3          895         0       3
      4          896         1       3

                                          Name     Sex   Age  SibSp  Parch  \
      0                          Kelly, Mr. James    male  34.5      0      0
      1          Wilkes, Mrs. James (Ellen Needs)  female  47.0      1      0
      2                 Myles, Mr. Thomas Francis    male  62.0      0      0
      3                          Wirz, Mr. Albert    male  27.0      0      0
      4  Hirvonen, Mrs. Alexander (Helga E Lindqvist)  female  22.0      1      1

          Ticket     Fare Cabin Embarked
      0   330911   7.8292   NaN        Q
      1   363272   7.0000   NaN        S
      2   240276   9.6875   NaN        Q
      3   315154   8.6625   NaN        S
      4  3101298  12.2875   NaN        S
```

[58]: `df.head(15)`

```
[58]:    PassengerId  Survived  Pclass  \
      0          892         0       3
      1          893         1       3
      2          894         0       2
      3          895         0       3
      4          896         1       3
      5          897         0       3
      6          898         1       3
      7          899         0       2
```

```
8           900      1      3
9           901      0      3
10          902      0      3
11          903      0      1
12          904      1      1
13          905      0      2
14          906      1      1
```

```
                                                   Name  Sex  SibSp  Parch  \
0                                        Kelly, Mr. James    1      0      0
1                           Wilkes, Mrs. James (Ellen Needs)    0      1      0
2                                  Myles, Mr. Thomas Francis    1      0      0
3                                          Wirz, Mr. Albert    1      0      0
4              Hirvonen, Mrs. Alexander (Helga E Lindqvist)    0      1      1
5                                Svensson, Mr. Johan Cervin    1      0      0
6                                      Connolly, Miss. Kate    0      0      0
7                                Caldwell, Mr. Albert Francis    1      1      1
8                    Abrahim, Mrs. Joseph (Sophie Halaut Easu)    0      0      0
9                                    Davies, Mr. John Samuel    1      2      0
10                                        Ilieff, Mr. Ylio    1      0      0
11                                 Jones, Mr. Charles Cresson    1      0      0
12          Snyder, Mrs. John Pillsbury (Nelle Stevenson)    0      1      0
13                                     Howard, Mr. Benjamin    1      1      0
14   Chaffee, Mrs. Herbert Fuller (Carrie Constance…    0      1      0
```

```
        Ticket      Fare Cabin Embarked
0       330911    7.8292   NaN        Q
1       363272    7.0000   NaN        S
2       240276    9.6875   NaN        Q
3       315154    8.6625   NaN        S
4      3101298   12.2875   NaN        S
5         7538    9.2250   NaN        S
6       330972    7.6292   NaN        Q
7       248738   29.0000   NaN        S
8         2657    7.2292   NaN        C
9     A/4 48871   24.1500   NaN        S
10      349220    7.8958   NaN        S
11          694   26.0000   NaN        S
12        21228   82.2667   B45        S
13        24065   26.0000   NaN        S
14   W.E.P. 5734   61.1750   E31        S
```

[16]: `df.shape`

[16]: (418, 12)

[17]: `df.describe()`

```
[17]:        PassengerId    Survived      Pclass         Age       SibSp  \
     count    418.000000  418.000000  418.000000  332.000000  418.000000
     mean    1100.500000    0.363636    2.265550   30.272590    0.447368
     std      120.810458    0.481622    0.841838   14.181209    0.896760
     min      892.000000    0.000000    1.000000    0.170000    0.000000
     25%      996.250000    0.000000    1.000000   21.000000    0.000000
     50%     1100.500000    0.000000    3.000000   27.000000    0.000000
     75%     1204.750000    1.000000    3.000000   39.000000    1.000000
     max     1309.000000    1.000000    3.000000   76.000000    8.000000

                 Parch        Fare
     count  418.000000  417.000000
     mean     0.392344   35.627188
     std      0.981429   55.907576
     min      0.000000    0.000000
     25%      0.000000    7.895800
     50%      0.000000   14.454200
     75%      0.000000   31.500000
     max      9.000000  512.329200
```

```
[21]: df['Survived'].value_counts()
```

```
[21]: 0    266
      1    152
      Name: Survived, dtype: int64
```

```
[25]: sns.countplot(x=df['Survived'], hue=df['Pclass'])
```

```
[25]: <AxesSubplot:xlabel='Survived', ylabel='count'>
```
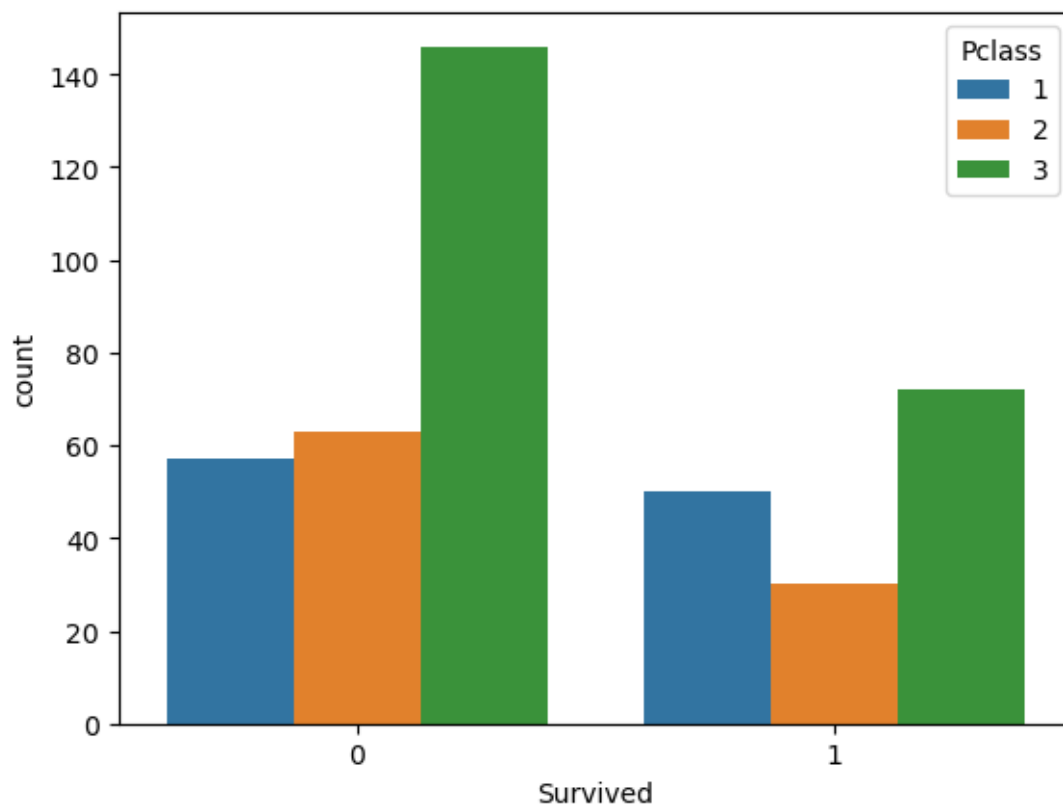
```
[23]: df["Sex"]
```

```
[23]: 0         male
      1       female
      2         male
      3         male
      4       female
              …
      413       male
      414     female
      415       male
      416       male
      417       male
      Name: Sex, Length: 418, dtype: object
```

```
[26]: sns.countplot(x=df['Sex'], hue=df['Survived'])
```

```
[26]: <AxesSubplot:xlabel='Sex', ylabel='count'>
```
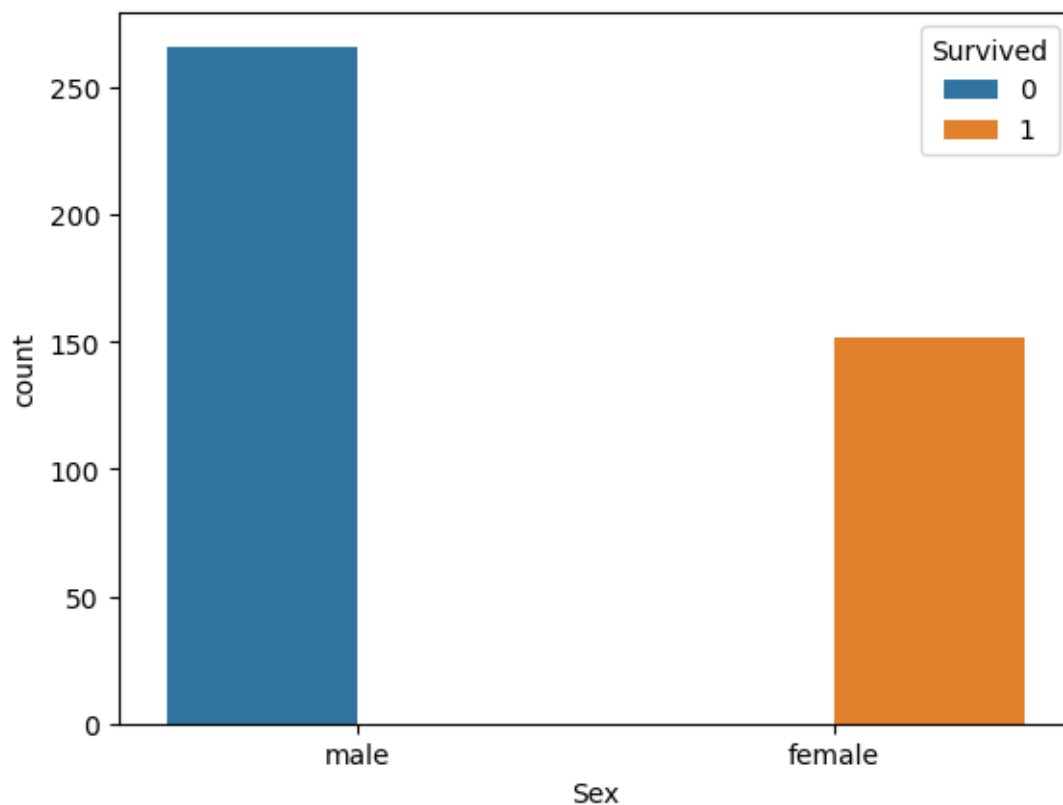
```
[27]: df.groupby('Sex')[['Survived']].mean()
```

```
[27]:         Survived
      Sex
      female       1.0
      male         0.0
```

```
[28]: df['Sex'].unique()
```

```
[28]: array(['male', 'female'], dtype=object)
```

```
[29]: from sklearn.preprocessing import LabelEncoder
      labelencoder = LabelEncoder()

      df['Sex']= labelencoder.fit_transform(df['Sex'])

      df.head()
```

```
[29]:    PassengerId  Survived  Pclass  \
      0          892         0       3
      1          893         1       3
```

```
2                                            894        0        2
3                                            895        0        3
4                                            896        1        3

                                             Name  Sex   Age  SibSp  Parch  \
0                               Kelly, Mr. James    1  34.5      0      0
1               Wilkes, Mrs. James (Ellen Needs)    0  47.0      1      0
2                      Myles, Mr. Thomas Francis    1  62.0      0      0
3                               Wirz, Mr. Albert    1  27.0      0      0
4  Hirvonen, Mrs. Alexander (Helga E Lindqvist)    0  22.0      1      1

      Ticket      Fare Cabin Embarked
0     330911   7.8292   NaN        Q
1     363272   7.0000   NaN        S
2     240276   9.6875   NaN        Q
3     315154   8.6625   NaN        S
4    3101298  12.2875   NaN        S
```

[30]: `df['Sex'], df['Survived']`

[30]:
```
(0      1
 1      0
 2      1
 3      1
 4      0
       ..
 413    1
 414    0
 415    1
 416    1
 417    1
 Name: Sex, Length: 418, dtype: int32,
 0      0
 1      1
 2      0
 3      0
 4      1
       ..
 413    0
 414    1
 415    0
 416    0
 417    0
 Name: Survived, Length: 418, dtype: int64)
```

[31]: `sns.countplot(x=df['Sex'], hue=df["Survived"])`

[31]: <AxesSubplot:xlabel='Sex', ylabel='count'>
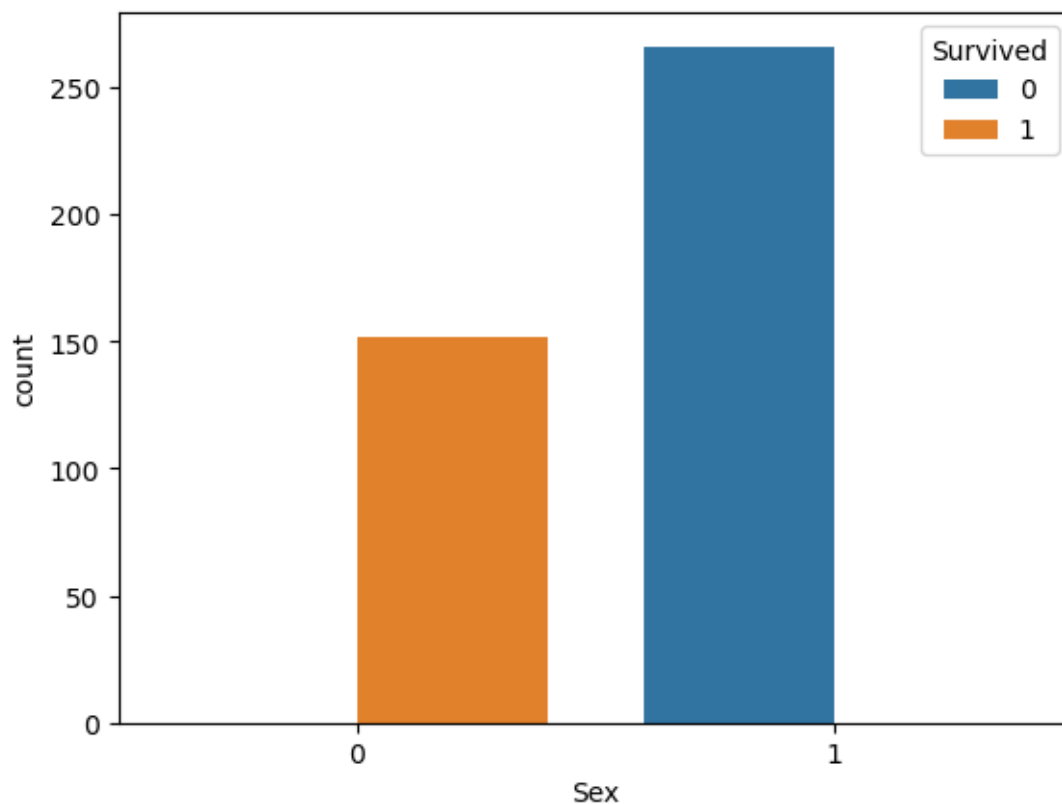
[32]: df.isna().sum()

[32]: PassengerId   0
      Survived   0
      Pclass   0
      Name   0
      Sex   0
      Age   86
      SibSp   0
      Parch   0
      Ticket   0
      Fare   1
      Cabin   327
      Embarked   0
      dtype: int64

[33]: df=df.drop(['Age'], axis=1)

```python
[34]: df_final = df
      df_final.head(10)
```

```
[34]:    PassengerId  Survived  Pclass  \
      0          892         0       3
      1          893         1       3
      2          894         0       2
      3          895         0       3
      4          896         1       3
      5          897         0       3
      6          898         1       3
      7          899         0       2
      8          900         1       3
      9          901         0       3

                                              Name  Sex  SibSp  Parch      Ticket  \
      0                            Kelly, Mr. James    1      0      0      330911
      1            Wilkes, Mrs. James (Ellen Needs)    0      1      0      363272
      2                  Myles, Mr. Thomas Francis    1      0      0      240276
      3                            Wirz, Mr. Albert    1      0      0      315154
      4  Hirvonen, Mrs. Alexander (Helga E Lindqvist)    0      1      1     3101298
      5                 Svensson, Mr. Johan Cervin    1      0      0        7538
      6                        Connolly, Miss. Kate    0      0      0      330972
      7               Caldwell, Mr. Albert Francis    1      1      1      248738
      8     Abrahim, Mrs. Joseph (Sophie Halaut Easu)    0      0      0        2657
      9                     Davies, Mr. John Samuel    1      2      0    A/4 48871

            Fare Cabin Embarked
      0   7.8292   NaN        Q
      1   7.0000   NaN        S
      2   9.6875   NaN        Q
      3   8.6625   NaN        S
      4  12.2875   NaN        S
      5   9.2250   NaN        S
      6   7.6292   NaN        Q
      7  29.0000   NaN        S
      8   7.2292   NaN        C
      9  24.1500   NaN        S
```

```python
[35]: X= df[['Pclass', 'Sex']]
      Y=df['Survived']
```

```python
[36]: 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,
       ↪random_state = 0)
```

```python
[37]: from sklearn.linear_model import LogisticRegression

      log = LogisticRegression(random_state = 0)
      log.fit(X_train, Y_train)
```

[37]: LogisticRegression(random_state=0)

```python
[38]: pred = print(log.predict(X_test))
```

```
[0 0 1 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 1 1 0 1 0 0 0 0 1 0 0 0 1 1 1 1 1 0 0
 1 1 1 1 0 1 1 0 1 0 0 0 0 0 1 1 0 0 1 0 1 0 0 0 1 1 0 0 1 1 1 1 0 0 1 1 1
 1 0 0 1 0 1 0 1 0 0]
```

```python
[39]: print(Y_test)
```

```
360    0
170    0
224    1
358    0
309    1
      ..
100    1
7      0
22     1
68     0
328    0
Name: Survived, Length: 84, dtype: int64
```

```python
[57]: import warnings
      warnings.filterwarnings("ignore")

      res= log.predict([[1,0]])

      if(res==0):
        print("So Sorry! Not Survived")
      else:
        print("Survived")
```

```
Survived
```

[ ]:

[ ]: