Fundamental Descriptive Statisticso

Ching Heng, Cheng

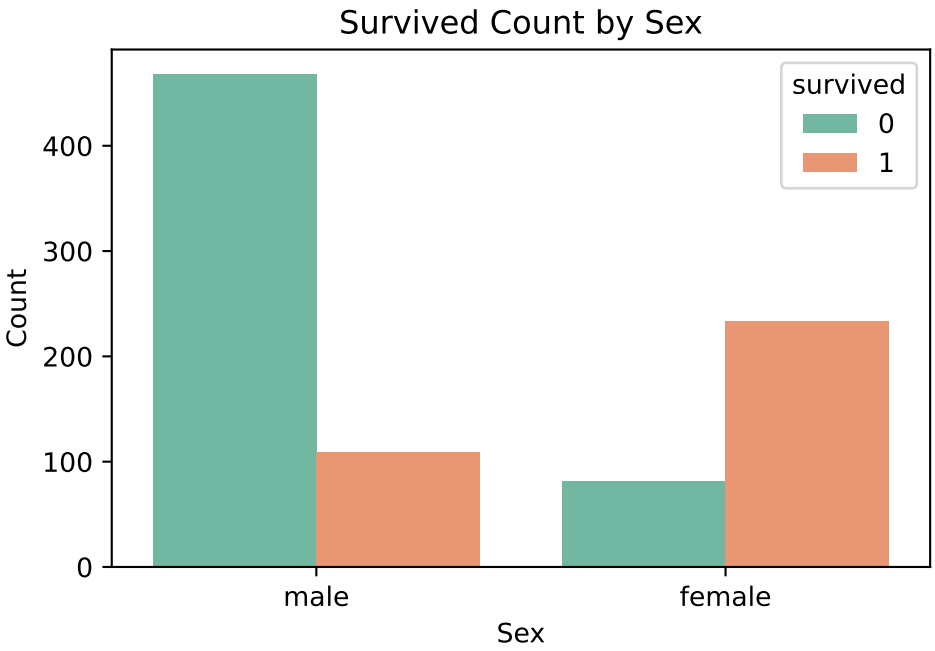
2025-02-27

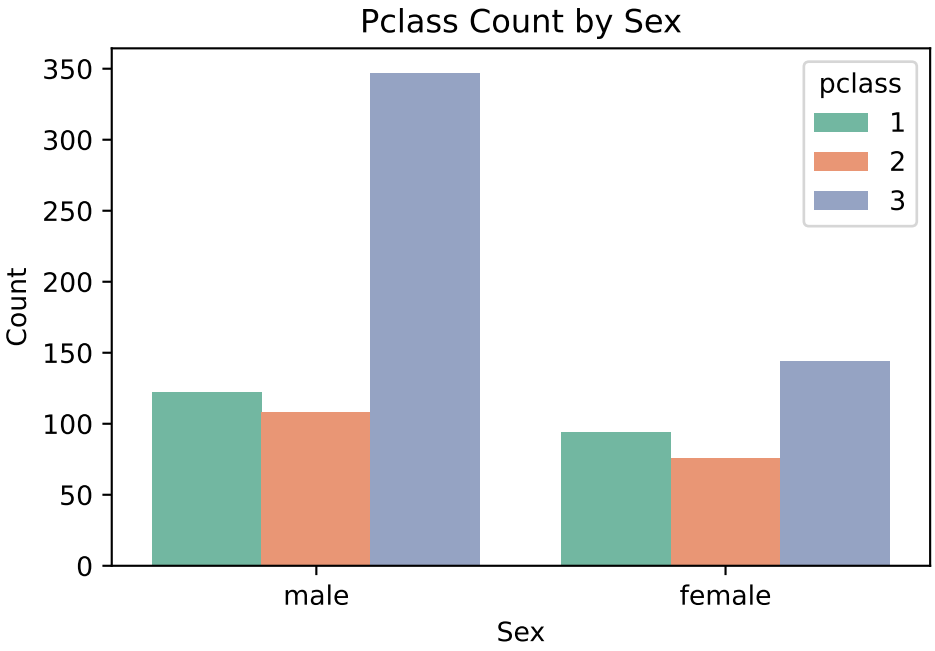
```
import pandas as pd
import numpy as np

data = pd.read_csv('titanic.csv')
df = pd.DataFrame(data)

print("Descriptive Statistics:")
print(df.describe())
```

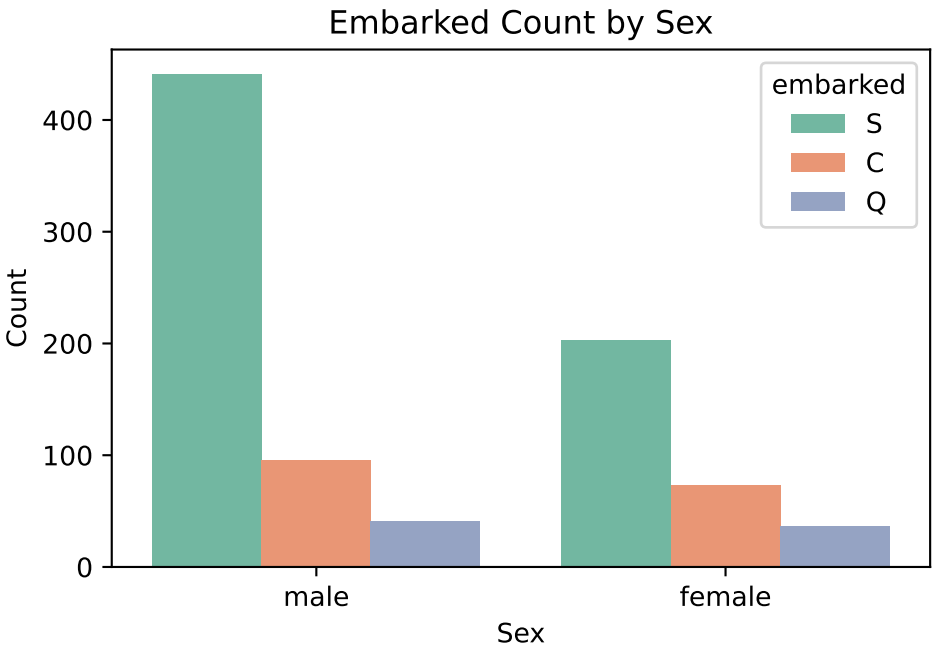
```
Descriptive Statistics:
       PassengerId    Survived      Pclass         Age       SibSp  \
count   891.000000  891.000000  891.000000  714.000000  891.000000
mean    446.000000    0.383838    2.308642   29.699118    0.523008
std     257.353842    0.486592    0.836071   14.526497    1.102743
min       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
max      6.000000  512.329200
```

Count Bar Plot - Sex

```
import seaborn as sns
import matplotlib.pyplot as plt

titanic = sns.load_dataset('titanic')

sns.countplot(x='sex', hue='survived', data=titanic, palette='Set2')

plt.title("Survived Count by Sex")
plt.xlabel("Sex")
plt.ylabel("Count")
plt.show()

sns.countplot(x='sex', hue='pclass', data=titanic, palette='Set2')

plt.title("Pclass Count by Sex")
plt.xlabel("Sex")
plt.ylabel("Count")
plt.show()

sns.countplot(x='sex', hue='embarked', data=titanic, palette='Set2')

plt.title("Embarked Count by Sex")
plt.xlabel("Sex")
plt.ylabel("Count")
plt.show()
```

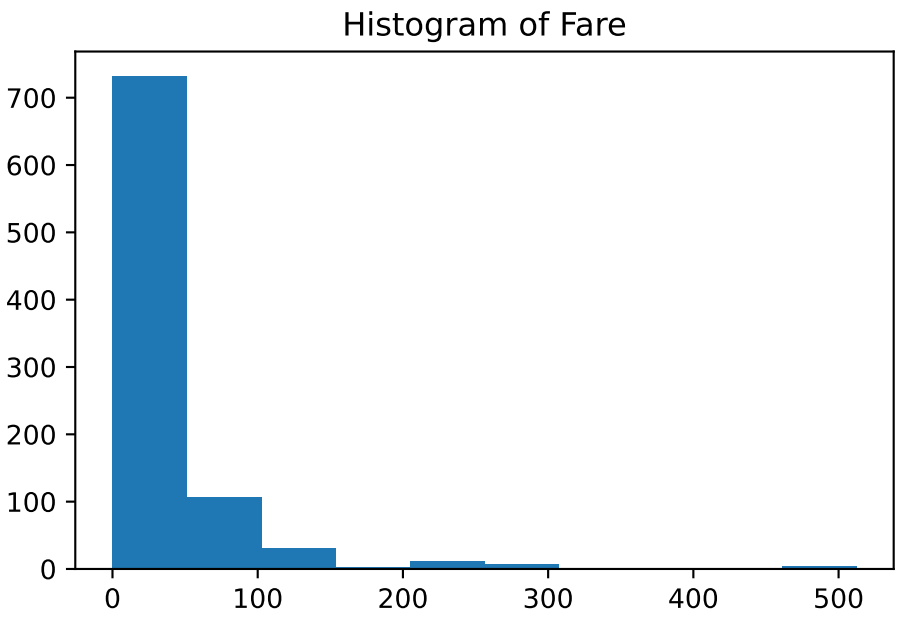

Survived Count by Sex
survived
0
1
400
300
200
100
0
Count
male
female
Sex

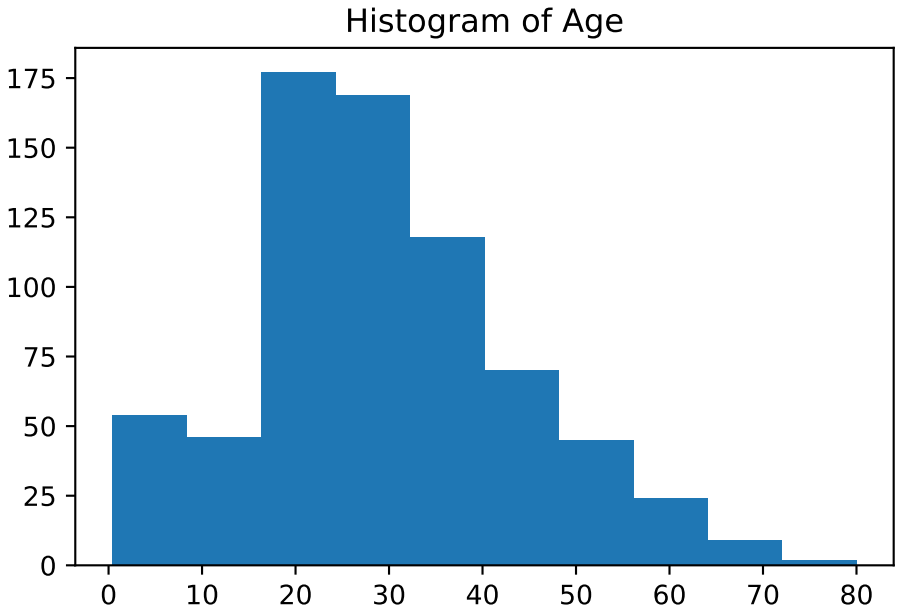

Pclass Count by Sex
pclass
1
2
3
350
300
250
200
150
100
50
0
Count
male
female
Sex



Histogram

```
plt.hist(df['Fare'])
plt.title('Histogram of Fare')
plt.show()

plt.hist(df['Age'])
plt.title('Histogram of Age')
plt.show()
```


Histogram of Fare
700
600
500
400
300
200
100
0
0
100
200
300
400
500


Histogram of Age
175
150
125
100
75
50
25
0
0
10
20
30
40
50
60
70
80