

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
import math

t=pd.read_csv('titanic.csv')
t.head(10)
```

```
Out[1]:
```

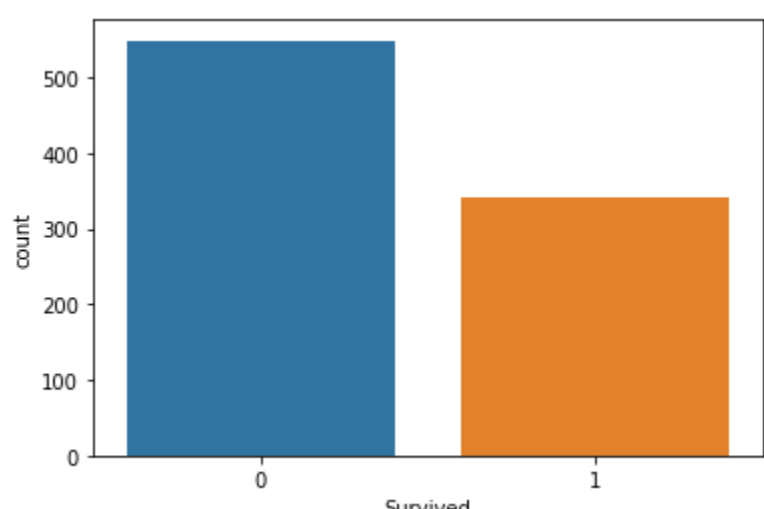
	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	NaN	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	NaN	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	NaN	C

```
In [2]: print("# of passengers in original data:"+str(len(t.index)))

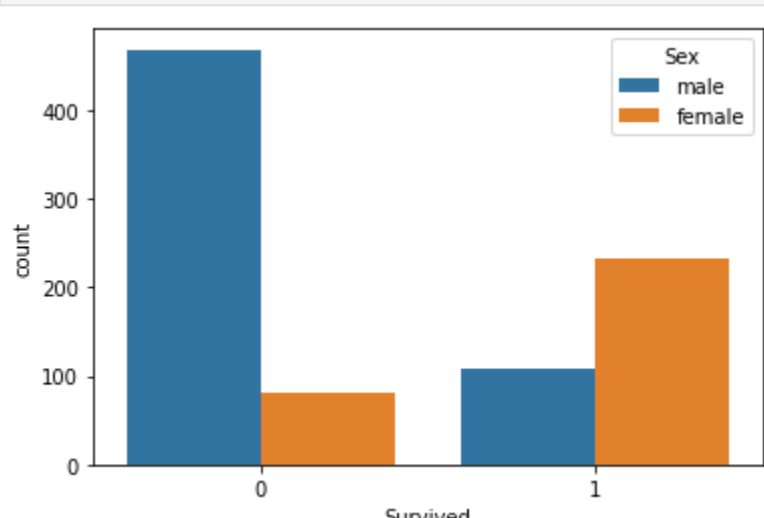
# of passengers in original data:891
```

Analyzing Data

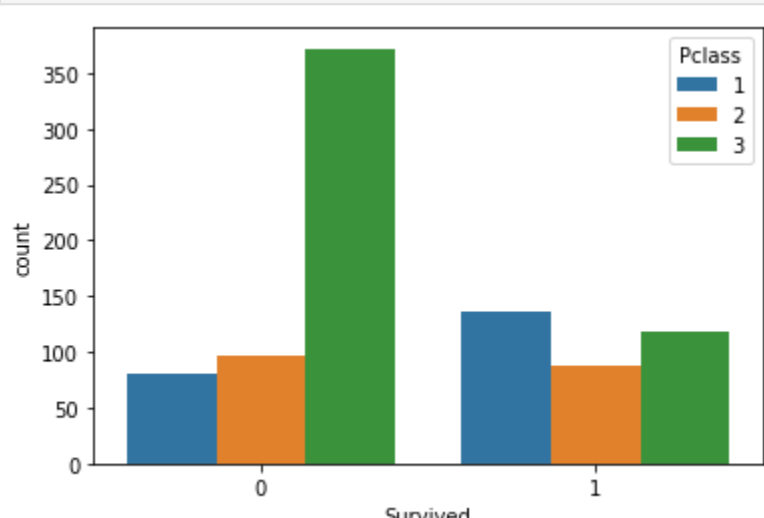
```
In [3]: sns.countplot(x='Survived',data=t)
plt.show()
```



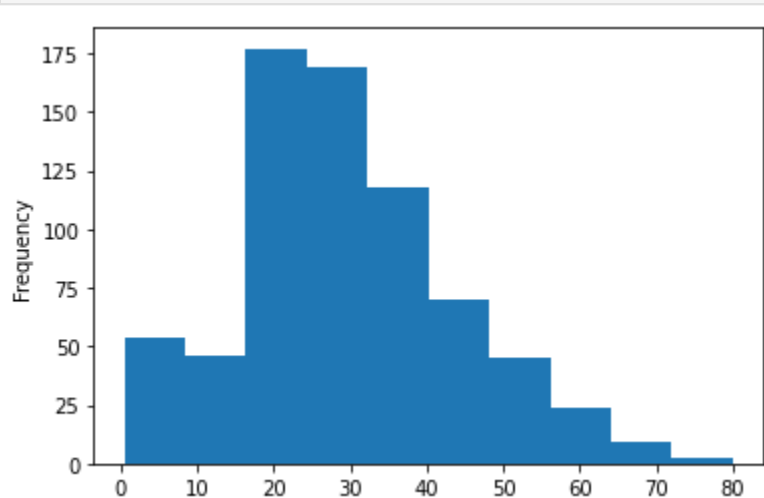
```
In [4]: sns.countplot(x='Survived',hue='Sex',data=t)
plt.show()
```



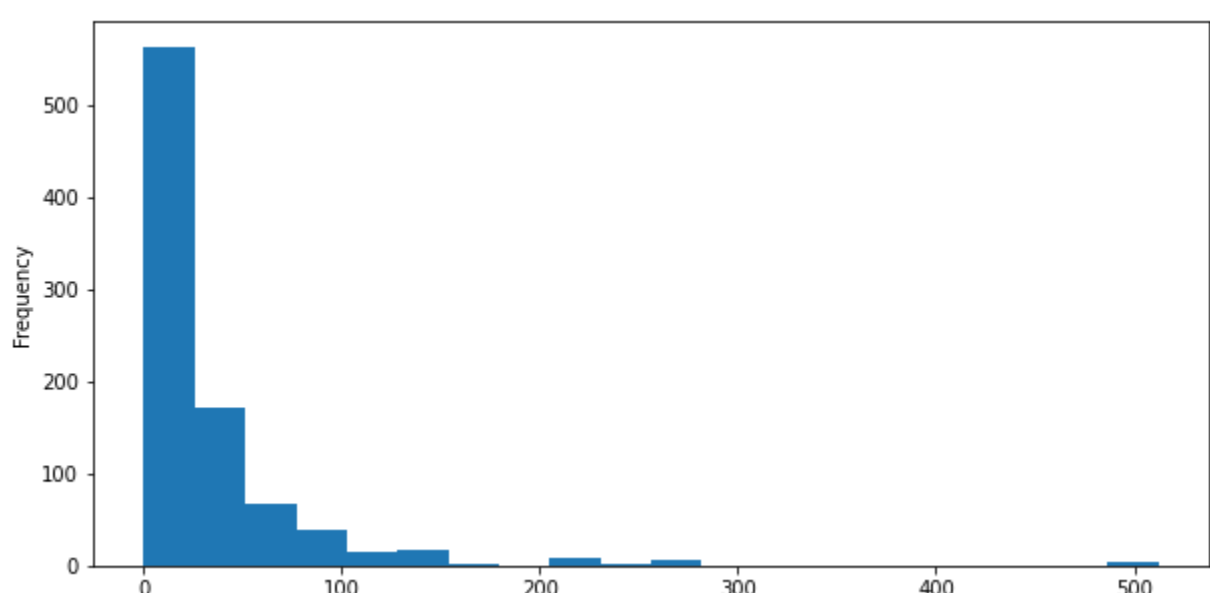
```
In [5]: sns.countplot(x='Survived',hue='Pclass',data=t)
plt.show()
```



```
In [6]: t['Age'].plot.hist()
plt.show()
```



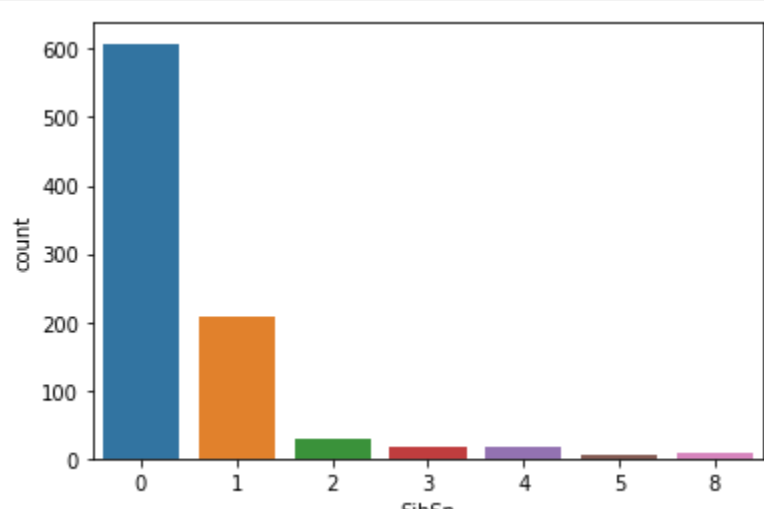
```
In [7]: t['Fare'].plot.hist(bins=20,figsize=(10,5))
plt.show()
```



```
In [8]: t.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
#   column      Non-Null Count  Dtype
---  -
0   PassengerId  891 non-null     int64
1   Survived     891 non-null     int64
2   Pclass       891 non-null     int64
3   Name         891 non-null     object
4   Sex          891 non-null     object
5   Age          714 non-null     float64
6   SibSp        891 non-null     int64
7   Parch        891 non-null     int64
8   Ticket       891 non-null     object
9   Fare         891 non-null     float64
10  Cabin        204 non-null     object
11  Embarked     889 non-null     object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

```
In [9]: sns.countplot(x='SibSp',data=t)
plt.show()
```



```
In [10]: t.isnull()
```

```
Out[10]:
```

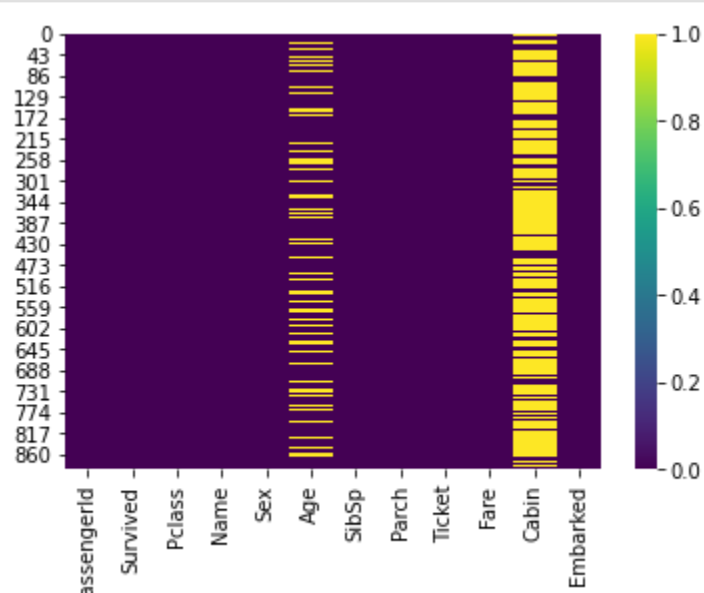
	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	False	False	False	False	False	False	False	False	False	False	True	False
1	False	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	True	False
3	False	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	True	False
...	...	...	...	...	...	...	...	...	...	...	...	...
886	False	False	False	False	False	False	False	False	False	False	True	False
887	False	False	False	False	False	False	False	False	False	False	False	False
888	False	False	False	False	True	False	False	False	False	False	True	False
889	False	False	False	False	False	False	False	False	False	False	False	False
890	False	False	False	False	False	False	False	False	False	False	True	False

891 rows × 12 columns

```
In [11]: t.isnull().sum()
```

```
Out[11]: PassengerId    0
Survived              0
Pclass                0
Name                  0
Sex                   0
Age                  177
SibSp                 0
Parch                 0
Ticket                0
Fare                  0
Cabin                687
Embarked              2
dtype: int64
```

```
In [12]: sns.heatmap(t.isnull(),cmap='viridis')
plt.show()
```



```
In [13]: pd.get_dummies(t['Sex'],drop_first=True)
```

```
Out[13]:
```

	male
0	1
1	0
2	0
3	0
4	1
...	...
886	1
887	0
888	0
889	1
890	1

891 rows × 1 columns

```
In [14]: a=pd.get_dummies(t['Embarked'],drop_first=True)
a.head(5)
```

```
Out[14]:
```

	Q	S
0	0	1
1	0	0
2	0	1
3	0	1
4	0	1

```
In [ ]:
```