

```
In [1]: import pandas as pd
```

```
In [2]: df=pd.read_csv('f:/dataset/analysis/titanic.csv')  
df
```

Out[2]:

	PassengerId	Survived	Pclass	Name	gender	Age	SibSp	Parch	Ticket	I
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0
...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7

891 rows × 12 columns

In [3]: `df.Survived.value_counts()`

```
Out[3]: Survived
0      549
1      342
Name: count, dtype: int64
```

```
In [8]: s_0=549/891
s_1=342/891
print(s_1)
```

```
0.3838383838383838
```

```
In [4]: df.Pclass.value_counts()
```

```
Out[4]: Pclass
3      491
1      216
2      184
Name: count, dtype: int64
```

```
In [6]: df[df.Survived==1].Pclass.value_counts()
```

```
Out[6]: Pclass
1      136
3      119
2       87
Name: count, dtype: int64
```

```
In [7]: sur_p1=136/216
sur_p2=87/184
sur_p3=119/491
print(sur_p1,sur_p2,sur_p3)
```

```
0.6296296296296297 0.47282608695652173 0.24236252545824846
```

```
In [9]: df.gender.value_counts()
```

```
Out[9]: gender
male      577
female    314
Name: count, dtype: int64
```

```
In [10]: df[df.Survived==1].gender.value_counts()
```

```
Out[10]: gender
female    233
male      109
Name: count, dtype: int64
```

```
In [11]: sur_male=109/577
sur_female=233/314
print(sur_male,sur_female)
```

```
0.18890814558058924 0.7420382165605095
```

```
In [12]: df[df.Age>70].shape
```

```
Out[12]: (5, 12)
```

```
In [14]: df[df.Age>70].Survived.value_counts()
```

```
Out[14]: Survived
0      4
1      1
Name: count, dtype: int64
```

```
In [15]: df[df.Age>70]
```

```
Out[15]:
```

	PassengerId	Survived	Pclass	Name	gender	Age	SibSp	Parch	Ticket	Fare
	96	0	1	Goldschmidt, Mr. George B	male	71.0	0	0	PC 17754	34.6
	116	0	3	Connors, Mr. Patrick	male	70.5	0	0	370369	7.7
	493	0	1	Artagaveytia, Mr. Ramon	male	71.0	0	0	PC 17609	49.5
	630	1	1	Barkworth, Mr. Algernon Henry Wilson	male	80.0	0	0	27042	30.0
	851	0	3	Svensson, Mr. Johan	male	74.0	0	0	347060	7.7

```
In [18]: df[df.Age>60].shape
```

```
Out[18]: (22, 12)
```

```
In [16]: df[df.Age>60].Survived.value_counts()
```

```
Out[16]: Survived
0      17
1       5
Name: count, dtype: int64
```

```
In [20]: df[df.Age<10].shape
```

```
Out[20]: (62, 12)
```

```
In [21]: df[df.Age<10].Survived.value_counts()
```

```
Out[21]: Survived
1      38
0      24
Name: count, dtype: int64
```

```
In [24]: df[(df.Age<10) & (df.Survived==0)].Pclass.value_counts()
```

```
Out[24]: Pclass
3      23
1       1
Name: count, dtype: int64
```

```
In [25]: df[(df.SibSp==0) & (df.Parch==0)].shape
```

```
Out[25]: (537, 12)
```

```
In [26]: df[(df.SibSp==0) & (df.Parch==0)].Survived.value_counts()
```

```
Out[26]: Survived
0      374
1      163
Name: count, dtype: int64
```

```
In [27]: df.Fare.min()
```

```
Out[27]: 0.0
```

```
In [28]: df.Fare.max()
```

```
Out[28]: 512.3292
```

```
In [29]: df.Fare.mean()
```

```
Out[29]: 32.204207968574636
```

```
In [30]: df[df.Fare>df.Fare.mean()].Survived.value_counts()
```

```
Out[30]: Survived
1      126
0       85
Name: count, dtype: int64
```

```
In [40]: df[df.Fare.between(270,515)]
```

```
Out[40]:
```

	PassengerId	Survived	Pclass	Name	gender	Age	SibSp	Parch	Ticket	Fare
258	259	1	1	Ward, Miss. Anna	female	35.0	0	0	PC 17755	512.3292
679	680	1	1	Cardeza, Mr. Thomas Drake Martinez	male	36.0	0	1	PC 17755	512.3292
737	738	1	1	Lesurer, Mr. Gustave J	male	35.0	0	0	PC 17755	512.3292

```
In [41]: df[df.Fare==0].Survived.value_counts()
```

```
Out[41]: Survived
         0      14
         1       1
         Name: count, dtype: int64
```

```
In [42]: df.Cabin.count()
```

```
Out[42]: 204
```

```
In [43]: 891-204
```

```
Out[43]: 687
```

```
In [44]: df.count()
```

```
Out[44]: PassengerId      891
         Survived        891
         Pclass         891
         Name           891
         gender         891
         Age            714
         SibSp          891
         Parch          891
         Ticket         891
         Fare           891
         Cabin          204
         Embarked       889
         dtype: int64
```

```
In [ ]:
```