In [9]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [69]:
```
cruise=pd.read_csv('train.csv')
```

In [5]:
```
cruise.head()
```

Out[5]:

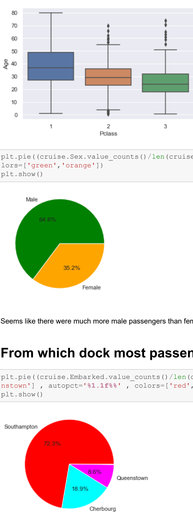
	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 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

In [4]:
```
cruise.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 [70]:
```
sns.boxplot(x='Pclass',y='Age',data=cruise)
```

Out[70]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19c1c5c8f48>
```



In [71]:
```
plt.pie((cruise.Sex.value_counts()/len(cruise.Sex)) , labels=['Male','Female'] , autopct='%1.1f%%' , colors=['green','orange'])
plt.show()
```


Seems like there were much more male passengers than female on Titanic

From which dock most passenger were embarked?

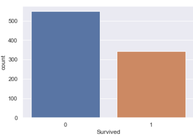
In [72]:
```
plt.pie((cruise.Embarked.value_counts()/len(cruise.Embarked)) , labels=['Southampton','Cherbourg','Queenstown'] , autopct='%1.1f%%' , colors=['red','cyan','magenta'])
plt.show()
```


The most number of passenger were embarked from Southampton and least from Queenstown

Who survived and who did not?

In [73]:
```
sns.set()
sns.countplot(x='Survived',data=cruise)
```

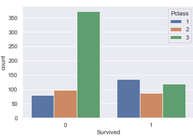
Out[73]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19c1b9c3748>
```



We can conclude from the above plot that over 500 people died on that fateful day.

In [74]:
```
sns.set()
sns.countplot(x='Survived',hue='Pclass',data=cruise)
```

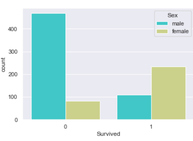
Out[74]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19c1959de08>
```



So from above we can conclude that if you are poor in a cruise which has crashed you are going to die. 0 stands for dead 1 stands for survive

In [75]:
```
sns.countplot(x='Survived',hue='Sex',data=cruise,palette='rainbow')
```

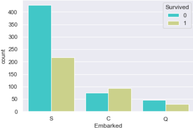
Out[75]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19c1c6a1148>
```



So women were allowed to escape first.

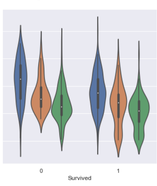
In [76]:
```
sns.countplot(x='Embarked', hue='Survived', data=cruise, palette = 'rainbow')
```

Out[76]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19c1c703288>
```



In [77]:
```
sns.catplot(x='Survived',y='Age',hue='Pclass',data=cruise,kind='violin')
```

Out[77]:
```
<seaborn.axisgrid.FacetGrid at 0x19c1c743648>
```

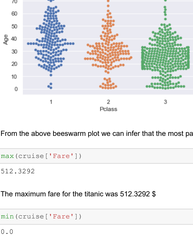


From the above plot we may conclude that the age group of 20-40 survived more.

In [78]:
```
sns.swarmplot(x='Pclass', y='Age', data=cruise)
```

```
c:\users\manish\appdata\local\programs\python\python37\lib\site-packages\seaborn\categorical.py:1311: RuntimeWarning: invalid value encountered in less
  off_low = points < low_gutter
c:\users\manish\appdata\local\programs\python\python37\lib\site-packages\seaborn\categorical.py:1315: RuntimeWarning: invalid value encountered in greater
  off_high = points > high_gutter
```

Out[78]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x19c1d7fa608>
```



From the above beeswarm plot we can infer that the most pasangers were between the age range of 20-40

In [48]:
```
max(cruise['Fare'])
```

Out[48]:
```
512.3292
```

The maximum fare for the titanic was 512.3292 $

In [49]:
```
min(cruise['Fare'])
```

Out[49]:
```
0.0
```

The minimum was 0.0 $

In [60]:
```
cruise[['Fare']].mean().round(2)
```

Out[60]:
```
Fare    32.2
dtype: float64
```

The average fare for a ticket to titanic was 32.20$

In [ ]:
```

```