

```
In [3]: #Importing the libraries to use
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

```
In [4]: #import and read data
dataSet = pd.read_csv('titanic.csv')
```

```
In [5]: #get top five rows
dataSet.head()
```

Out[5]:

	Survived	Pclass	Name	Sex	Age	Siblings/Spouses Aboard	Parents/Children Aboard	Fare
0	0	3	Mr. Owen Harris Braund	male	22.0	1	0	7.2500
1	1	1	Mrs. John Bradley (Florence Briggs Thayer) Cum...	female	38.0	1	0	71.2833
2	1	3	Miss. Laina Heikkinen	female	26.0	0	0	7.9250
3	1	1	Mrs. Jacques Heath (Lily May Peel) Futrelle	female	35.0	1	0	53.1000
4	0	3	Mr. William Henry Allen	male	35.0	0	0	8.0500

```
In [6]: print("To be onboard people paid ",round(dataSet["Fare"].mean(),2), " on av
To be onboard people paid 32.31 on average
```

```
In [7]: #get fare and pclass columns
subSet= dataSet[['Fare','Pclass']]
```

```
In [8]: #get fare mins and maxs grouped by Pclass column
fare_by_class = subSet.groupby(['Pclass']).describe()
fare_by_class
```

Out[8]:

	Fare							
	count	mean	std	min	25%	50%	75%	max
Pclass								
1	216.0	84.154687	78.380373	0.0	30.92395	60.2875	93.5	512.3292
2	184.0	20.662183	13.417399	0.0	13.00000	14.2500	26.0	73.5000
3	487.0	13.707707	11.817309	0.0	7.75000	8.0500	15.5	69.5500

```
In [9]: fare_by_class.iloc[0:0,7]
```

```
Out[9]: Series([], Name: (Fare, max), dtype: float64)
```

```
In [10]: max_1_class = fare_by_class.iloc[[0],[7]]  
min_1_class = fare_by_class.iloc[[0],[3]]
```

```
In [11]: float(max_1_class.iloc[[0][0]])
```

```
Out[11]: 512.3292
```

```
In [12]: float(min_1_class.iloc[[0][0]])
```

```
Out[12]: 0.0
```

```
In [13]: items_by_class = fare_by_class.values[:3]
```

```
In [14]: row1 = items_by_class[0]  
for i in range(len(row1)):  
    max1= row1[-1]  
    min1 = row1[3]  
print('The max fare in class 1 was',max1,'while the min was',min1)
```

The max fare in class 1 was 512.3292 while the min was 0.0

```
In [15]: row2 = items_by_class[1]  
for i in range(len(row2)):  
    max2= row2[-1]  
    min2 = row2[3]  
print('The max fare in class 2 was',max2,'while the min was',min2)
```

The max fare in class 2 was 73.5 while the min was 0.0

```
In [22]: row3 = items_by_class[2]  
for i in range(len(row3)):  
    max3= row3[-1]  
    min3 = row3[3]  
print('The max fare in class 3 was',max3,'while the min was',min3)
```

The max fare in class 3 was 69.55 while the min was 0.0

```
In [23]: population =(dataSet['Survived']== 1).sum()
```

```
Out[23]: 342
```

```
In [24]: proba_class = (dataSet['Pclass'] == 3).sum()/population  
proba_years = (dataSet['Age'] <= 10).sum()/population
```

```
In [25]: print("The probability that a child who is in third class and is 10 years o
```

```
The probability that a child who is in third class and is 10 years old or  
younger survives is 30.0%
```