

# Importing Libraries

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

## Importing Dataset

```
In [2]: Titanic = pd.read_csv("D:/Datasets/train.csv")
```

## Exploratory Data Analysis

```
In [3]: ## 1.) Finding the first rows and columns of data set ##
```

```
In [4]: Titanic.head(10)
```

Out[4]:	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
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 [5]: ## 2.) Finding the last roes and columns of the data set ##
```

```
In [6]: Titanic.tail(10)
```

Out[6]:	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.8958	NaN	S
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.5167	NaN	S
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.5000	NaN	S
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.0500	NaN	S
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	NaN	Q
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

```
In [7]: ## 3.) Finding the number of rows and columns of the data set ##
```

```
In [8]: Titanic.shape
```

```
Out[8]: (891, 12)
```

```
In [9]: ## 4.) Finding the Data Types of the data set ##
```

```
In [10]: Titanic.dtypes
```

```
Out[10]: PassengerId .... int64
Survived ..... int64
Pclass ..... int64
Name ..... object
Sex ..... object
Age ..... float64
SibSp ..... int64
Parch ..... int64
Ticket ..... object
Fare ..... float64
Cabin ..... object
Embarked ..... object
dtype: object
```

```
In [11]: ## 5.) Finding the unique value in the data set ##
```

```
In [12]: Titanic.nunique()
```

```
Out[12]: PassengerId ... 891
Survived ..... 2
Pclass ..... 3
Name ..... 891
Sex ..... 2
Age ..... 88
SibSp ..... 7
Parch ..... 7
Ticket ..... 681
Fare ..... 248
Cabin ..... 147
Embarked ..... 3
dtype: int64
```

```
In [13]: ## 5a.) Checking the unique values of SibSp ##
```

```
In [14]: Titanic['SibSp'].unique()
```

```
Out[14]: array([1, 0, 3, 4, 2, 5, 8], dtype=int64)
```

```
In [15]: ## 5b.) Checking the unique values of Pclass ##
```

```
In [16]: Titanic['Pclass'].unique()
```

```
Out[16]: array([3, 1, 2], dtype=int64)
```

```
In [17]: ## 5c.) Checking the unique values of Cabin ##
```

```
In [18]: Titanic['Cabin'].unique()
```

```
Out[18]: array([nan, 'C85', 'C123', 'E46', 'G6', 'C103', 'D56', 'A6',
.....: 'C23 C25 C27', 'B78', 'D33', 'B30', 'C52', 'B28', 'C83', 'F33',
.....: 'F G73', 'E31', 'A5', 'D10 D12', 'D26', 'C110', 'B58 B60', 'E101',
.....: 'F E69', 'D47', 'B86', 'F2', 'C2', 'E33', 'B19', 'A7', 'C49', 'F4',
.....: 'A32', 'B4', 'B80', 'A31', 'D36', 'D15', 'C93', 'C78', 'D35',
```

```

..... 'C87', 'B77', 'E67', 'B94', 'C125', 'C99', 'C118', 'D7', 'A19',
..... 'B49', 'D', 'C22 C26', 'C106', 'C65', 'E36', 'C54',
..... 'B57 B59 B63 B66', 'C7', 'E34', 'C32', 'B18', 'C124', 'C91', 'E40',
..... 'T', 'C128', 'D37', 'B35', 'E50', 'C82', 'B96 B98', 'E10', 'E44',
..... 'A34', 'C104', 'C111', 'C92', 'E38', 'D21', 'E12', 'E63', 'A14',
..... 'B37', 'C30', 'D20', 'B79', 'E25', 'D46', 'B73', 'C95', 'B38',
..... 'B39', 'B22', 'C86', 'C70', 'A16', 'C101', 'C68', 'A10', 'E68',
..... 'B41', 'A20', 'D19', 'D50', 'D9', 'A23', 'B50', 'A26', 'D48',
..... 'E58', 'C126', 'B71', 'B51 B53 B55', 'D49', 'B5', 'B20', 'F G63',
..... 'C62 C64', 'E24', 'C90', 'C45', 'E8', 'B101', 'D45', 'C46', 'D30',
..... 'E121', 'D11', 'E77', 'F38', 'B3', 'D6', 'B82 B84', 'D17', 'A36',
..... 'B102', 'B69', 'E49', 'C47', 'D28', 'E17', 'A24', 'C50', 'B42',
..... 'C148'], dtype=object)

```

```
In [19]: ## 6a.) Checking the number of counts in Cabin varisble #
```

```
In [20]: Titanic.Cabin.value_counts()
```

```

Out[20]: C23 C25 C27 ... 4
         G6 ... 4
         B96 B98 ... 4
         D ... 3
         C22 C26 ... 3
         ..
         E17 ... 1
         C32 ... 1
         C50 ... 1
         C110 ... 1
         A36 ... 1
         Name: Cabin, Length: 147, dtype: int64

```

```
In [21]: ## 6b.) Checking the number of counts in Name varisble ##
```

```
In [22]: Titanic.Name.value_counts()
```

```

Out[22]: Silven, Miss. Lyyli Karoliina ... 1
         Novel, Mr. Mansouer ... 1
         Patchett, Mr. George ... 1
         Spencer, Mrs. William Augustus (Marie Eugenie) ... 1
         Banfield, Mr. Frederick James ... 1
         ..
         Moutal, Mr. Rahamin Haim ... 1
         Lefebre, Miss. Jeannie ... 1
         Bowerman, Miss. Elsie Edith ... 1
         Andersson, Mr. August Edvard ("Wennerstrom") ... 1
         Cacic, Miss. Marija ... 1
         Name: Name, Length: 891, dtype: int64

```

```
In [23]: ## 6c.) Checking the number of counts in Ticket varisble ##
```

```
In [24]: Titanic.Ticket.value_counts()
```

```

Out[24]: 1601 ... 7
         CA. 2343 ... 7
         347082 ... 7
         CA 2144 ... 6
         3101295 ... 6
         ..
         237565 ... 1
         S.O.P. 1166 ... 1
         323951 ... 1
         345777 ... 1
         SC 1748 ... 1
         Name: Ticket, Length: 681, dtype: int64

```