1. We uploaded/inserted the train.csv file to Jupyter notebook and pandas for analysis.

```
[9]: df = pd.read_csv(r'C:\Users\W10\Downloads\train.csv')
```

2. We used describe() function to get the count etc information.

```
[10]: df.describe()
```

[10]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

3. We used Info() function to get non-null values information.

```
[11]: df.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
```

4. We used value_counts() function to get all the relevant information about the passengers.

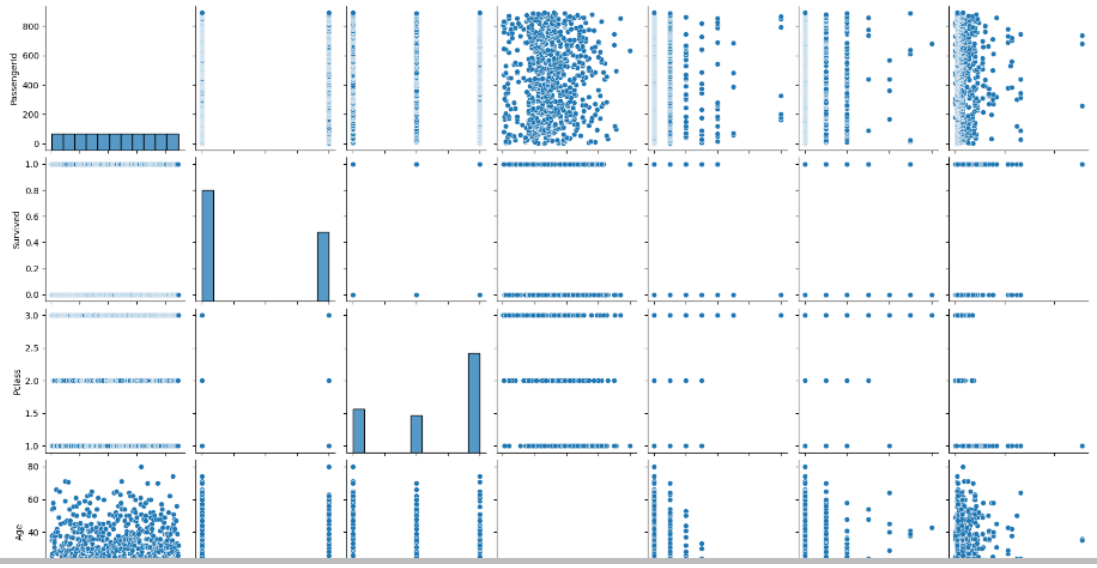
```
[12]: df.value_counts()
```

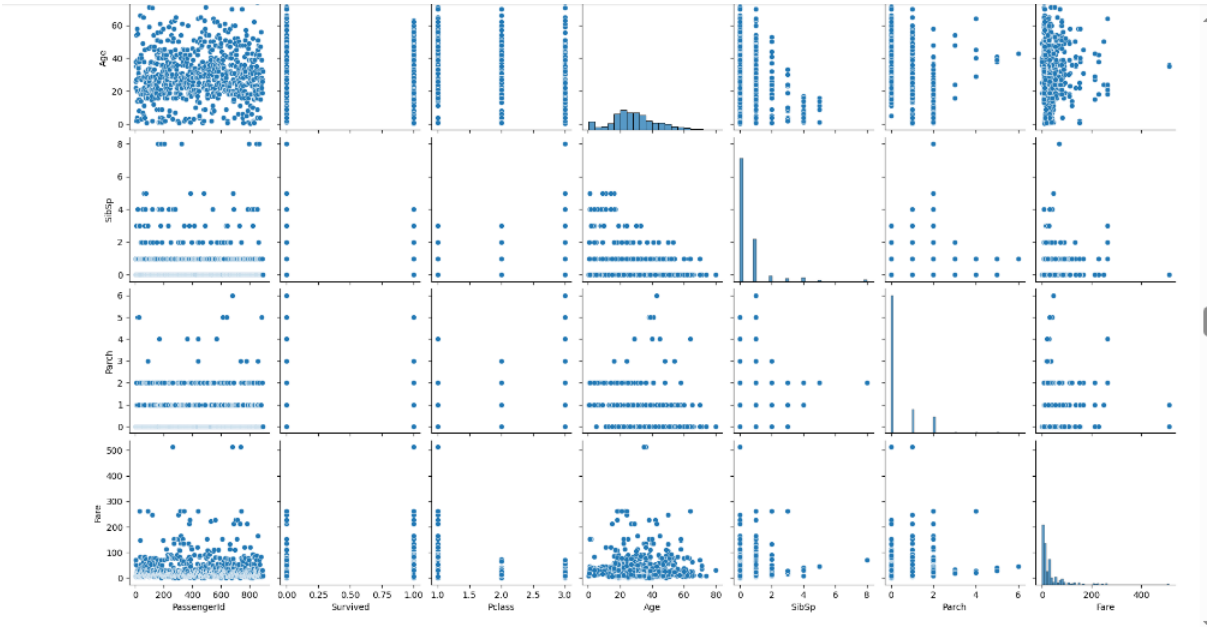
```
[12]: PassengerId  Survived  Pclass  Name                                                Sex     Age   SibSp  Parch  Ticket    Fare
      Cabin        Embarked
      2            1         1       Cumings, Mrs. John Bradley (Florence Briggs Thayer) female  38.0  1      0      PC 17599  71.2
      833  C85           C          1
      4            1         1       Futrelle, Mrs. Jacques Heath (Lily May Peel)        female  35.0  1      0      113803    53.1
      000  C123          S          1
      7            0         1       McCarthy, Mr. Timothy J                             male    54.0  0      0      17463     51.8
      625  E46           S          1
      11           1         3       Sandstrom, Miss. Marguerite Rut                     female  4.0   1      1      PP 9549   16.7
      000  G6            S          1
      12           1         1       Bonnell, Miss. Elizabeth                            female  58.0  0      0      113783    26.5
      500  C103          S          1

      ..
      872          1         1       Beckwith, Mrs. Richard Leonard (Sallie Monypeny)    female  47.0  1      1      11751     52.5
      542  D35           S          1
      873          0         1       Carlsson, Mr. Frans Olof                            male    33.0  0      0      695       5.00
      00   B51 B53 B55   S          1
      880          1         1       Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)       female  56.0  0      1      11767     83.1
      583  C50           C          1
      888          1         1       Graham, Miss. Margaret Edith                        female  19.0  0      0      112053    30.0
      000  B42           S          1
      890          1         1       Behr, Mr. Karl Howell                               male    26.0  0      0      111369    30.0
      000  C148          C          1
      Name: count, Length: 183, dtype: int64
```

5. We used pairplot() functions to see the figures in graphical format

```
[4]: sns.pairplot(df)
```

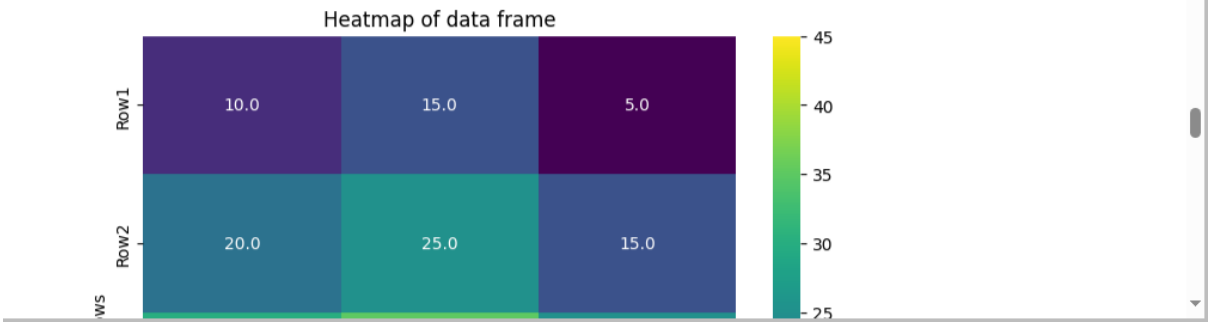
```
[4]: <seaborn.axisgrid.PairGrid at 0x230dba9c050>
```





6. We took heatmap() example to show it in figure.

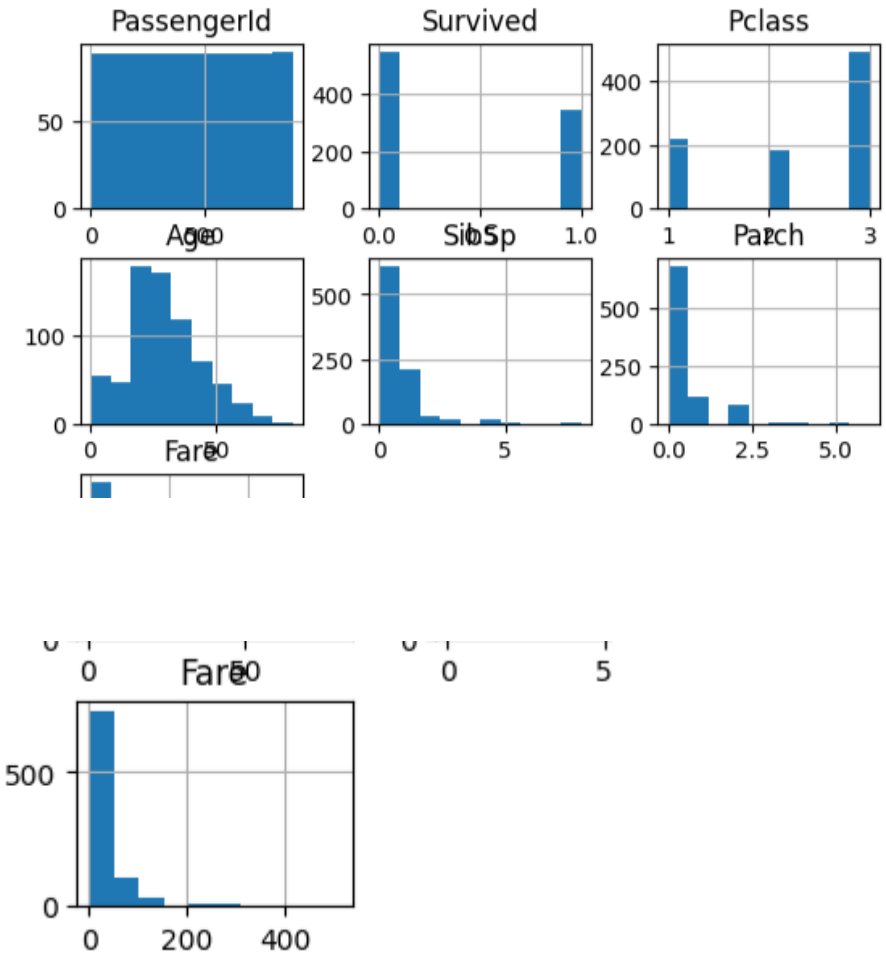
```
[14]: data={
          'Column_A': [ 10, 20, 30, 40],
          'Column_B': [ 15, 25, 35, 45],
          'Column_C': [ 5, 15, 25, 35]
      }
      df1 = pd.DataFrame(data, index=['Row1', 'Row2', 'Row3', 'Row4'])
      plt.figure(figsize=(8,6))
      sns.heatmap(df1, annot=True, cmap='viridis', fmt=".1f")
      plt.title('Heatmap of data frame')
      plt.xlabel('Columns')
      plt.ylabel('Rows')
      plt.show()
```



7. We used hist() function to see it in histogram.

```
[15]: df.hist()
```

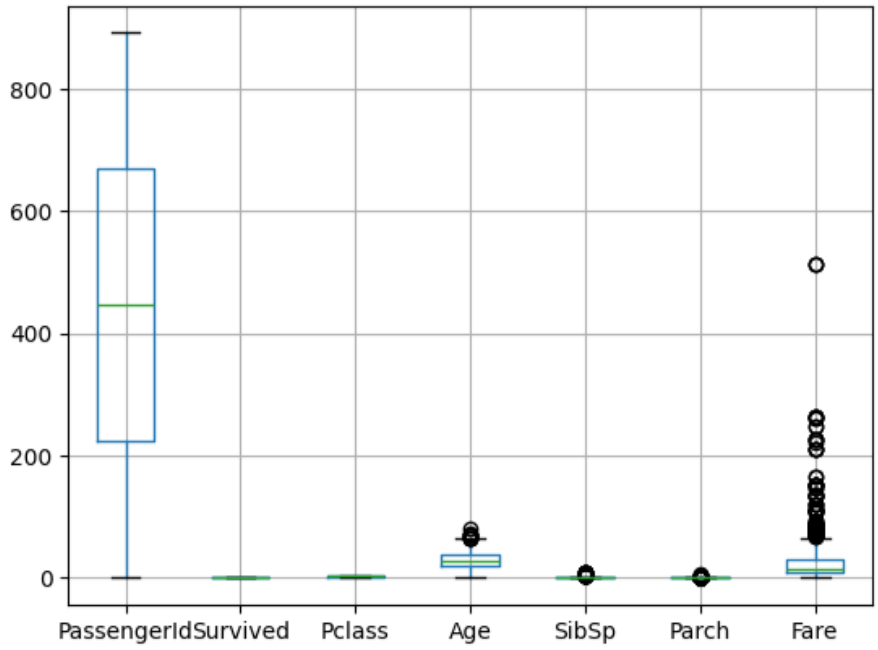
```
[15]: array([[<Axes: title={'center': 'PassengerId'}>,
              <Axes: title={'center': 'Survived'}>,
              <Axes: title={'center': 'Pclass'}>],
             [<Axes: title={'center': 'Age'}>,
              <Axes: title={'center': 'SibSp'}>,
              <Axes: title={'center': 'Parch'}>],
             [<Axes: title={'center': 'Fare'}>, <Axes: >, <Axes: >]],
            dtype=object)
```



8. We used boxplot() function.

```
[16]: df.boxplot()
```

```
[16]: <Axes: >
```



9. We used scatterplot() function.

```
[18]: df.plot.scatter(x='PassengerId', y = 'Fare')
```

```
[18]: <Axes: xlabel='PassengerId', ylabel='Fare'>
```