

31115_Assignment8

February 28, 2022

```
In [2]: import seaborn as sb
import matplotlib.pyplot as mp
import pandas as pd
```

```
In [5]: data = pd.read_csv('./titanic.csv')
```

```
In [6]: print(data.corr())
```

	Passengerid	Age	Fare	Sex	sibsp	zero \
Passengerid	1.000000	0.025799	0.031128	-0.013406	-0.055224	NaN
Age	0.025799	1.000000	0.178182	-0.053663	-0.189972	NaN
Fare	0.031128	0.178182	1.000000	0.185681	0.160349	NaN
Sex	-0.013406	-0.053663	0.185681	1.000000	0.109609	NaN
sibsp	-0.055224	-0.189972	0.160349	0.109609	1.000000	NaN
zero	NaN	NaN	NaN	NaN	NaN	NaN
zero.1	NaN	NaN	NaN	NaN	NaN	NaN
zero.2	NaN	NaN	NaN	NaN	NaN	NaN
zero.3	NaN	NaN	NaN	NaN	NaN	NaN
zero.4	NaN	NaN	NaN	NaN	NaN	NaN
zero.5	NaN	NaN	NaN	NaN	NaN	NaN
zero.6	NaN	NaN	NaN	NaN	NaN	NaN
Parch	0.008942	-0.125851	0.221635	0.213125	0.373587	NaN
zero.7	NaN	NaN	NaN	NaN	NaN	NaN
zero.8	NaN	NaN	NaN	NaN	NaN	NaN
zero.9	NaN	NaN	NaN	NaN	NaN	NaN
zero.10	NaN	NaN	NaN	NaN	NaN	NaN
zero.11	NaN	NaN	NaN	NaN	NaN	NaN
zero.12	NaN	NaN	NaN	NaN	NaN	NaN
zero.13	NaN	NaN	NaN	NaN	NaN	NaN
zero.14	NaN	NaN	NaN	NaN	NaN	NaN
Pclass	-0.038354	-0.377908	-0.558683	-0.124617	0.060832	NaN
zero.15	NaN	NaN	NaN	NaN	NaN	NaN
zero.16	NaN	NaN	NaN	NaN	NaN	NaN
Embarked	-0.051518	-0.065125	-0.239212	-0.099408	0.066055	NaN
zero.17	NaN	NaN	NaN	NaN	NaN	NaN
zero.18	NaN	NaN	NaN	NaN	NaN	NaN
2urvived	-0.331493	-0.055862	0.173786	0.404020	-0.014375	NaN

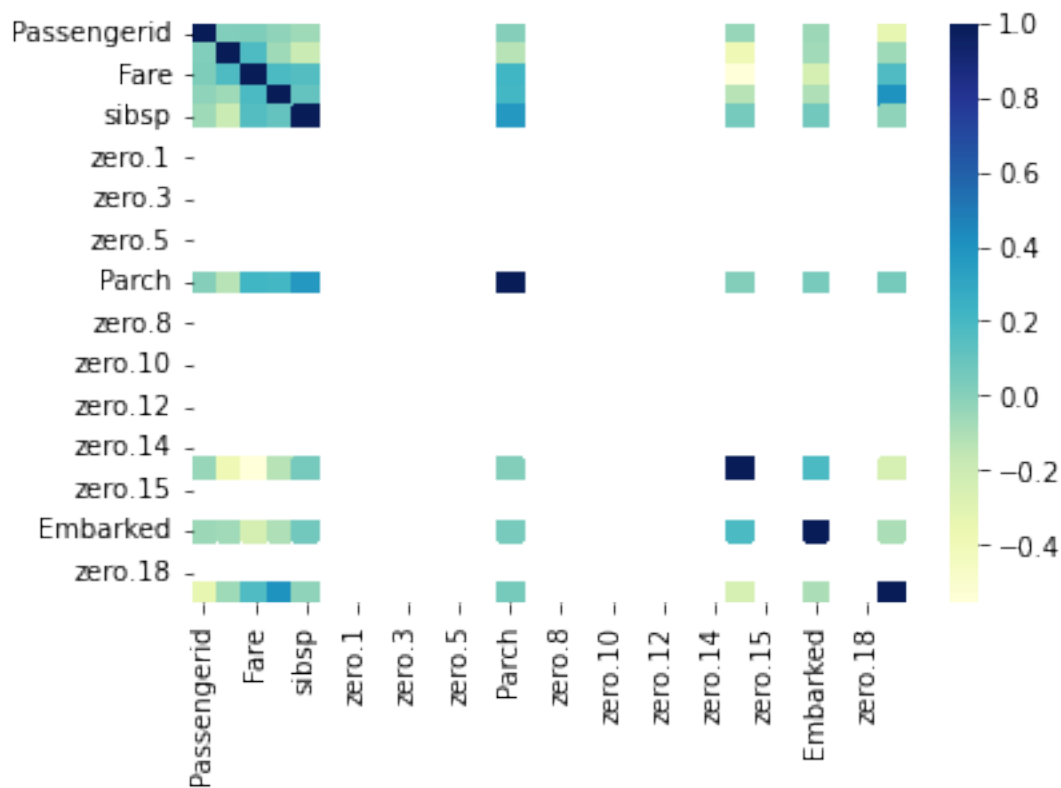
	zero.1	zero.2	zero.3	zero.4	...	zero.12	zero.13	zero.14	\
Passengerid	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
Age	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
Fare	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
Sex	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
sibsp	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.1	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.2	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.3	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.4	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.5	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.6	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
Parch	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.7	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.8	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.9	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.10	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.11	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.12	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.13	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.14	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
Pclass	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.15	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.16	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
Embarked	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.17	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
zero.18	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	
2urvived	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	

	Pclass	zero.15	zero.16	Embarked	zero.17	zero.18	2urvived
Passengerid	-0.038354	NaN	NaN	-0.051518	NaN	NaN	-0.331493
Age	-0.377908	NaN	NaN	-0.065125	NaN	NaN	-0.055862
Fare	-0.558683	NaN	NaN	-0.239212	NaN	NaN	0.173786
Sex	-0.124617	NaN	NaN	-0.099408	NaN	NaN	0.404020
sibsp	0.060832	NaN	NaN	0.066055	NaN	NaN	-0.014375
zero	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.1	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.2	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.3	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.4	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.5	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.6	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Parch	0.018322	NaN	NaN	0.045216	NaN	NaN	0.054908
zero.7	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.8	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.10	NaN	NaN	NaN	NaN	NaN	NaN	NaN

zero.11	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.12	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.13	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.14	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Pclass	1.000000	NaN	NaN	0.187349	NaN	NaN	-0.244686
zero.15	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.16	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Embarked	0.187349	NaN	NaN	1.000000	NaN	NaN	-0.093225
zero.17	NaN	NaN	NaN	NaN	NaN	NaN	NaN
zero.18	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2survived	-0.244686	NaN	NaN	-0.093225	NaN	NaN	1.000000

[28 rows x 28 columns]

```
In [7]: # you can add ', annot=True' at end for annotations
dataplot = sb.heatmap(data.corr(), cmap="YlGnBu")
```



```
In [8]: data.Fare
```

```
Out[8]: 0      7.2500
        1     71.2833
```

```

2          7.9250
3         53.1000
4          8.0500
...
1304        8.0500
1305       108.9000
1306        7.2500
1307        8.0500
1308       22.3583
Name: Fare, Length: 1309, dtype: float64

```

```

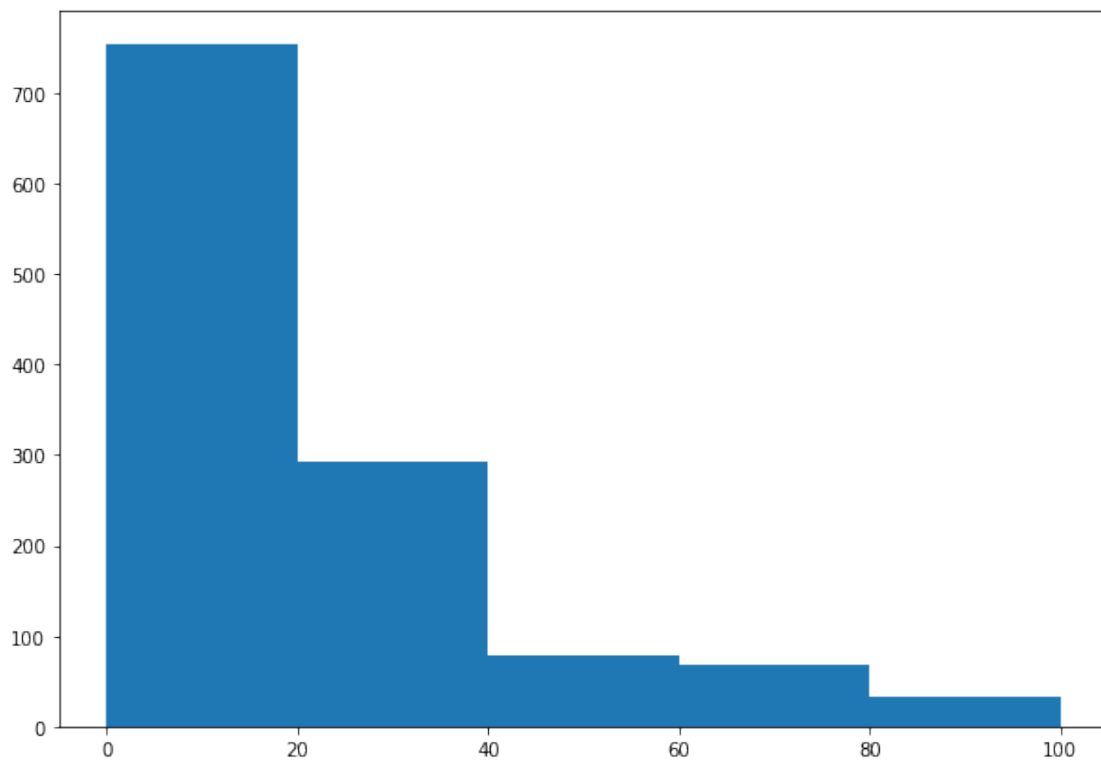
In [9]: # figsize allows us to modify the width and height of the bars
        # bins are the bars

        # fig : This method return the figure layout.
        # ax : This method return the axes. Axes object or array of Axes objects.

        fig, ax = mp.subplots(figsize =(10, 7))
        ax.hist(data.Fare, bins = [0, 20, 40, 60 , 80 , 100])

Out[9]: (array([753., 292.,  78.,  69.,  33.]),
        array([ 0, 20, 40, 60, 80, 100]),
        <BarContainer object of 5 artists>)

```



```
In [10]: data.head()
```

```
Out[10]:
```

	Passengerid	Age	Fare	Sex	sibsp	zero	zero.1	zero.2	zero.3	\
0	1	22.0	7.2500	0	1	0	0	0	0	
1	2	38.0	71.2833	1	1	0	0	0	0	
2	3	26.0	7.9250	1	0	0	0	0	0	
3	4	35.0	53.1000	1	1	0	0	0	0	
4	5	35.0	8.0500	0	0	0	0	0	0	

	zero.4	...	zero.12	zero.13	zero.14	Pclass	zero.15	zero.16	Embarked	\
0	0	...	0	0	0	3	0	0	2.0	
1	0	...	0	0	0	1	0	0	0.0	
2	0	...	0	0	0	3	0	0	2.0	
3	0	...	0	0	0	1	0	0	2.0	
4	0	...	0	0	0	3	0	0	2.0	

	zero.17	zero.18	2urvived
0	0	0	0
1	0	0	1
2	0	0	1
3	0	0	1
4	0	0	0

```
[5 rows x 28 columns]
```