A

AISANYM
NURBAULINA

Analysis of victims and survivors of passengers in the Titanic

data analysis
final project

Question we need to anwer:

1. How many survived passengers? and what gender they were?
2. What the average age for each gender?
3. What information can u give about Age ?
4. Occurance of age in each pclass.

```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

FIRST STEPS

```python
titanic = pd.read_csv('tested.csv')
titanic.head()
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	0	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S

```
#describing how many columns and rows have table
titanic.shape
```

```
# columns of the table
titanic.columns
```

```
# sums of nulls in each column
titanic.isnull().sum()
```

```
69]

PassengerId      0
Survived         0
Pclass           0
Name             0
Sex              0
Age             86
SibSp            0
Parch            0
Ticket           0
Fare             1
Cabin          327
Embarked         0
dtype: int64
```

Analyse 1

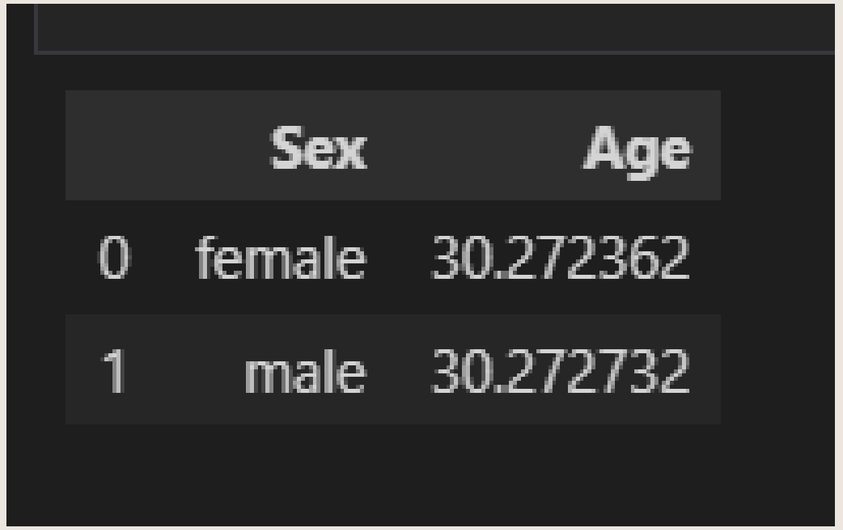
Sum of survived people

by this anayse we knowed that

```
There were  152  survived passengers from the titanic
```

Analyse 2

Average for each gender category



	Sex	Age
0	female	30.272362
1	male	30.272732

by this analyse we knowed that

In each gender type there are approximately same average. but average of male is greater for 0.013 than female

Analyse 3

coefficent of survived by pclasses

	Pclass	Survived
0	1	0.467290
2	3	0.330275
1	2	0.322581

by this analyse we knowed that

Most of survived passengers are from 1st class

Analyse 4

Number of survived males and females

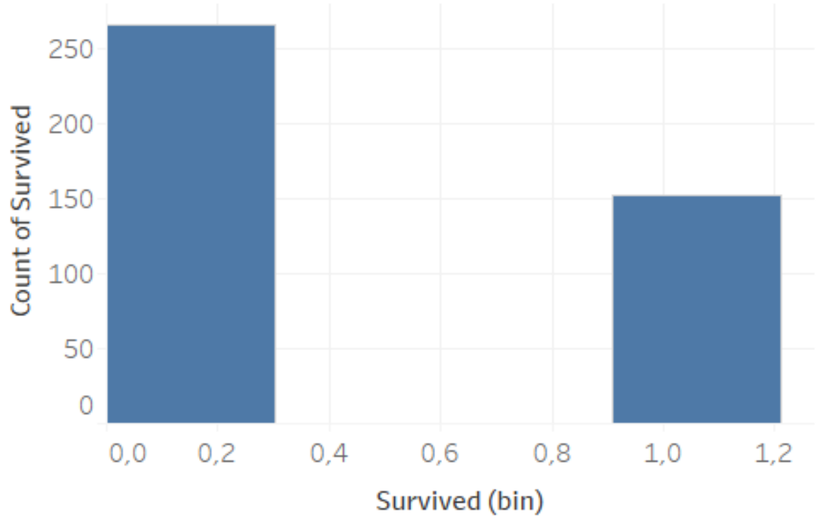
Survived	0	1
Sex		
female	NaN	152.0
male	266.0	NaN

By this analyse we knowed that

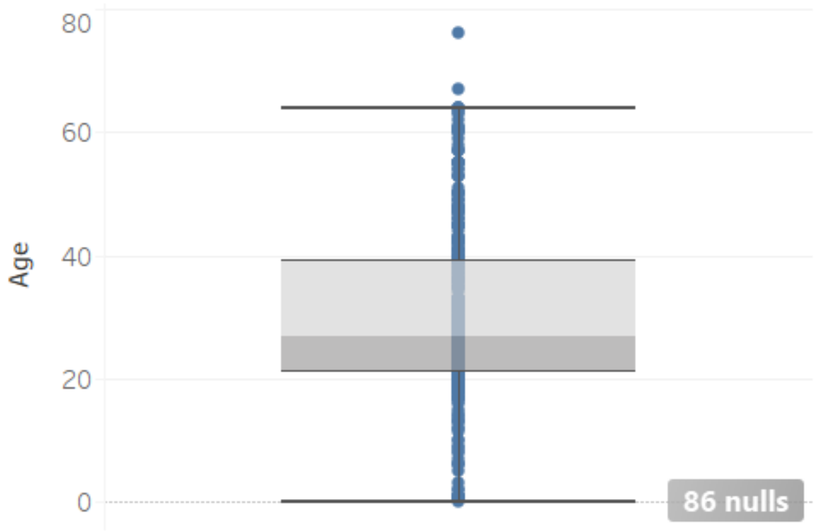
All survaved people were females and all males were died

Dashboard 1

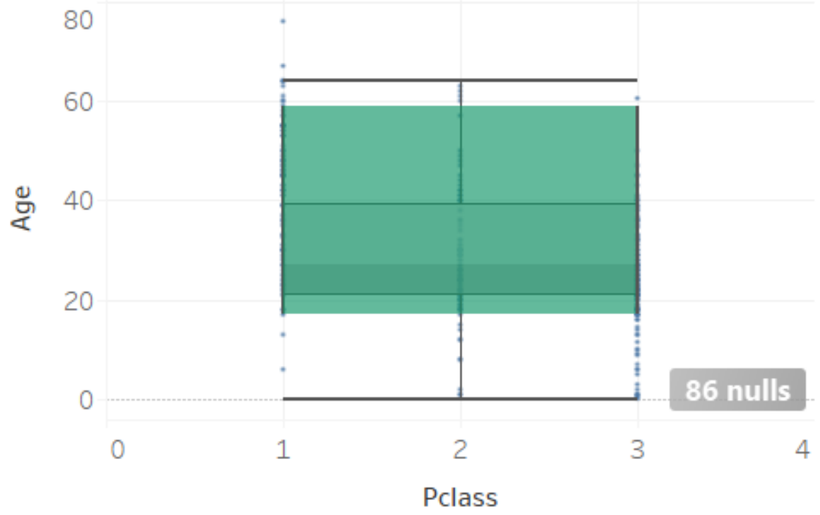
Sheet 4



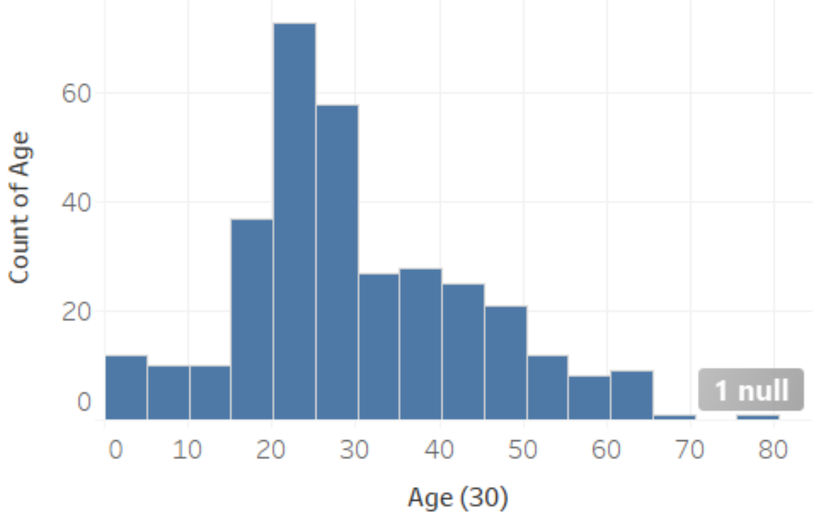
Sheet 2



Sheet 3

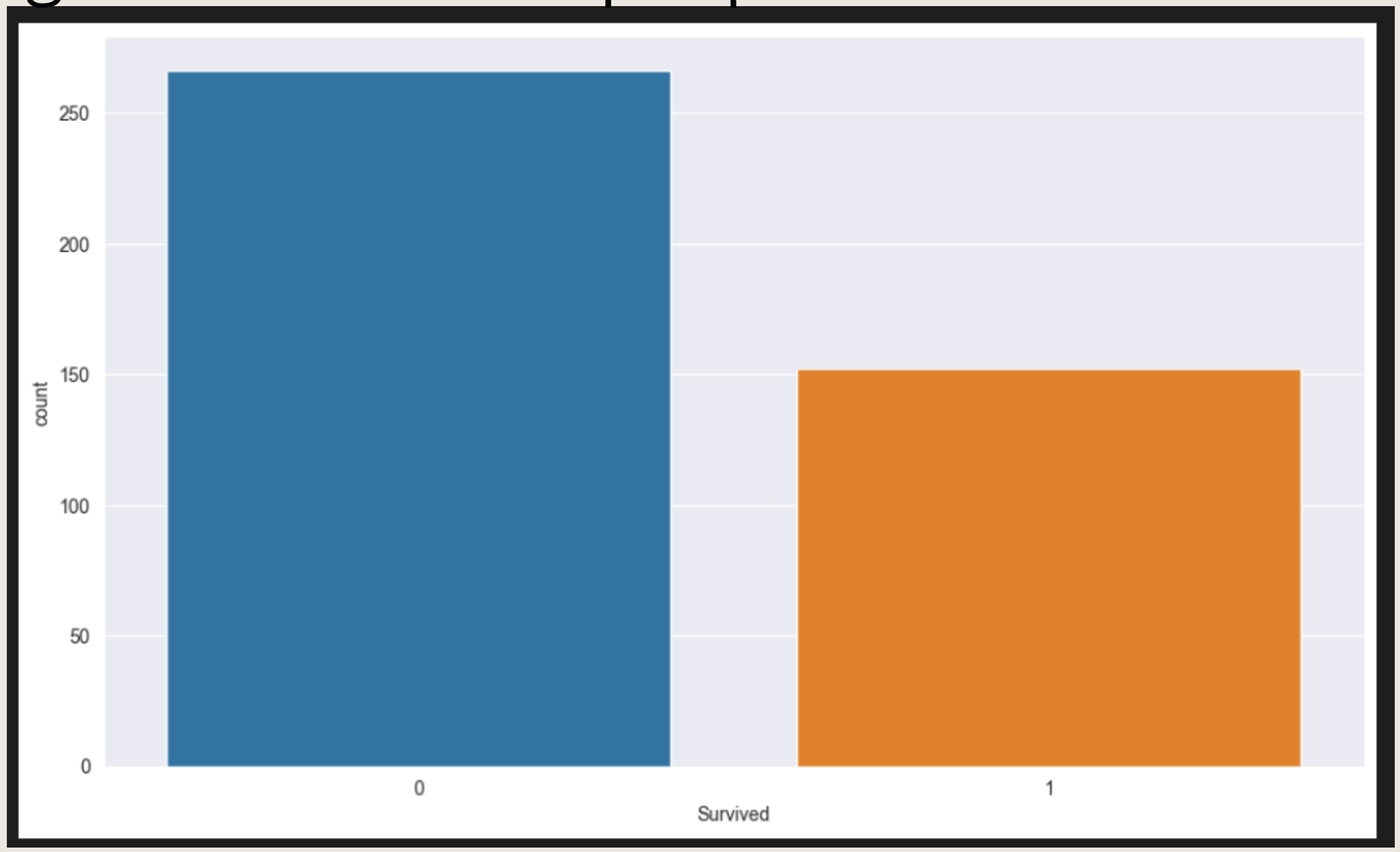


Sheet 1



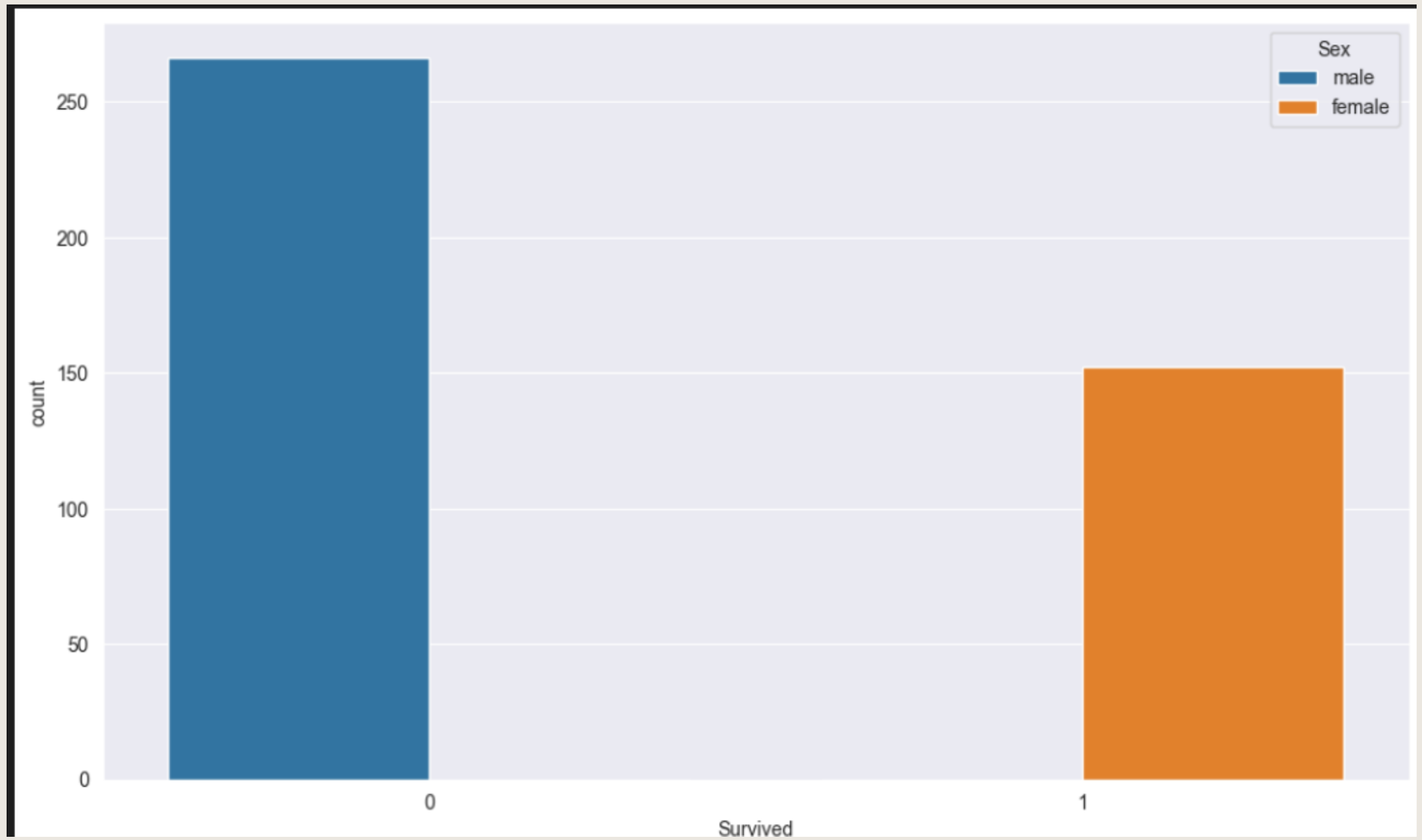
Analyse 5

showing hist of survived people



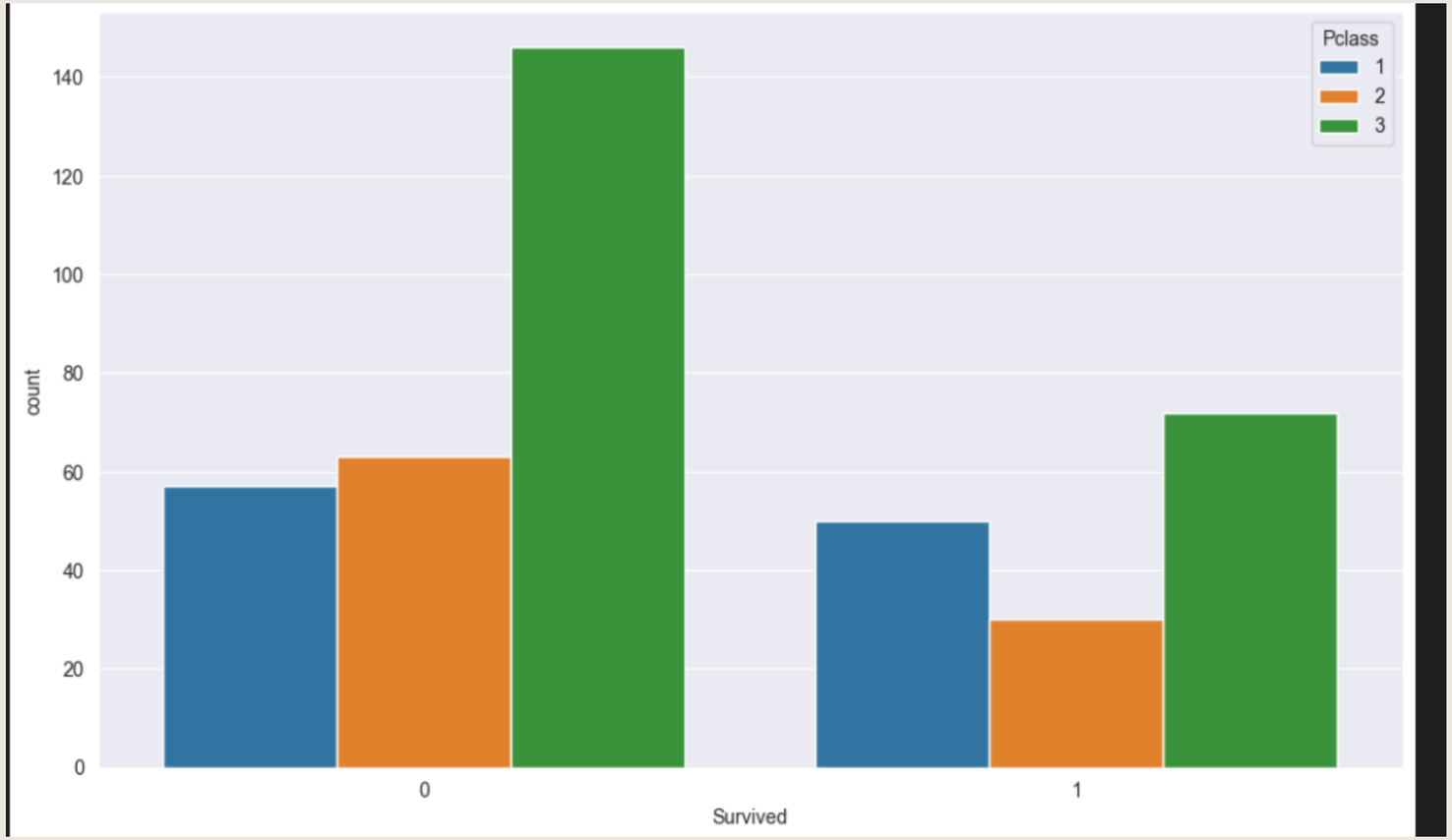
Analyse 6

showing survived statuses for each gender



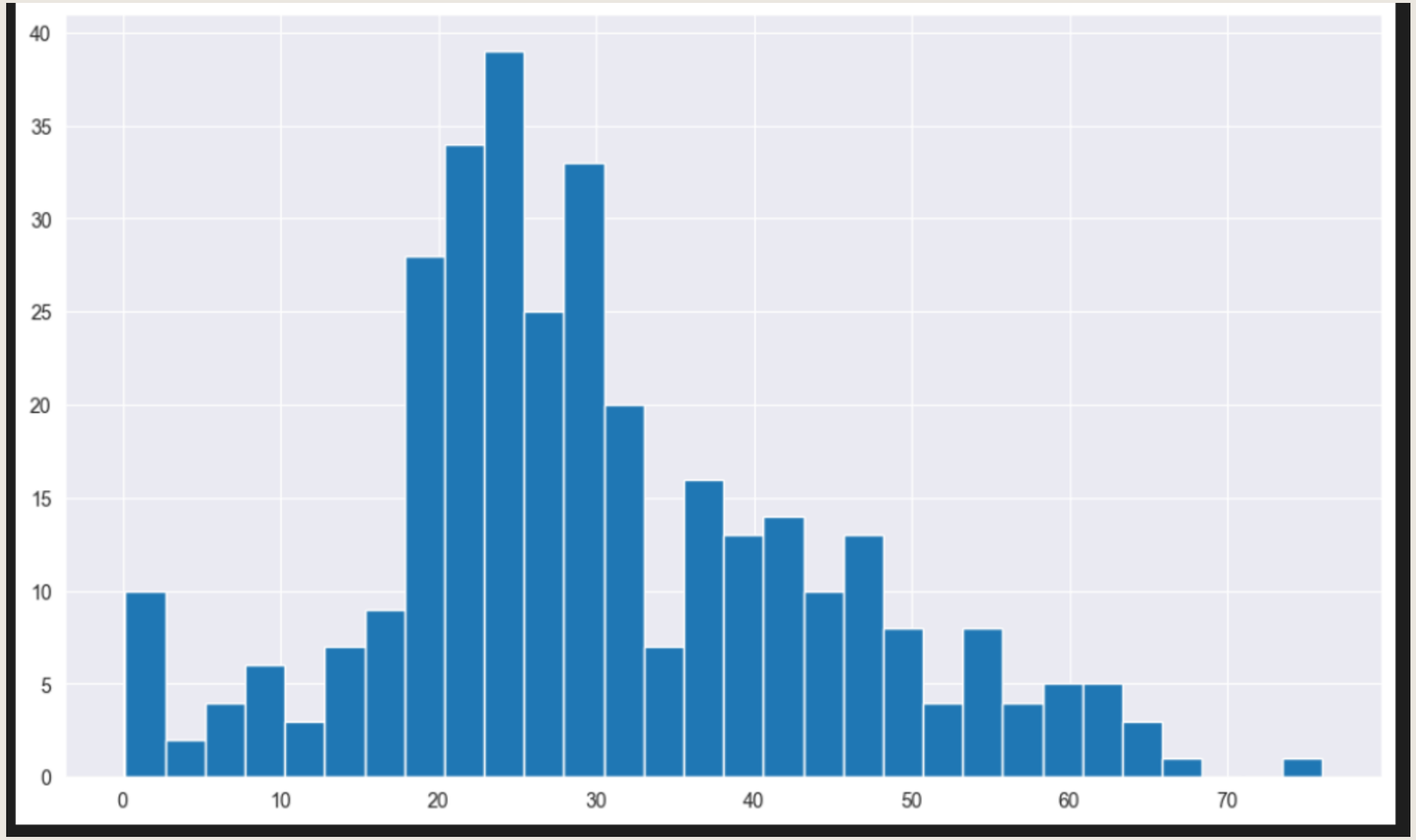
Analyse 7

showing survived statuses for each pclass



Analyse 8

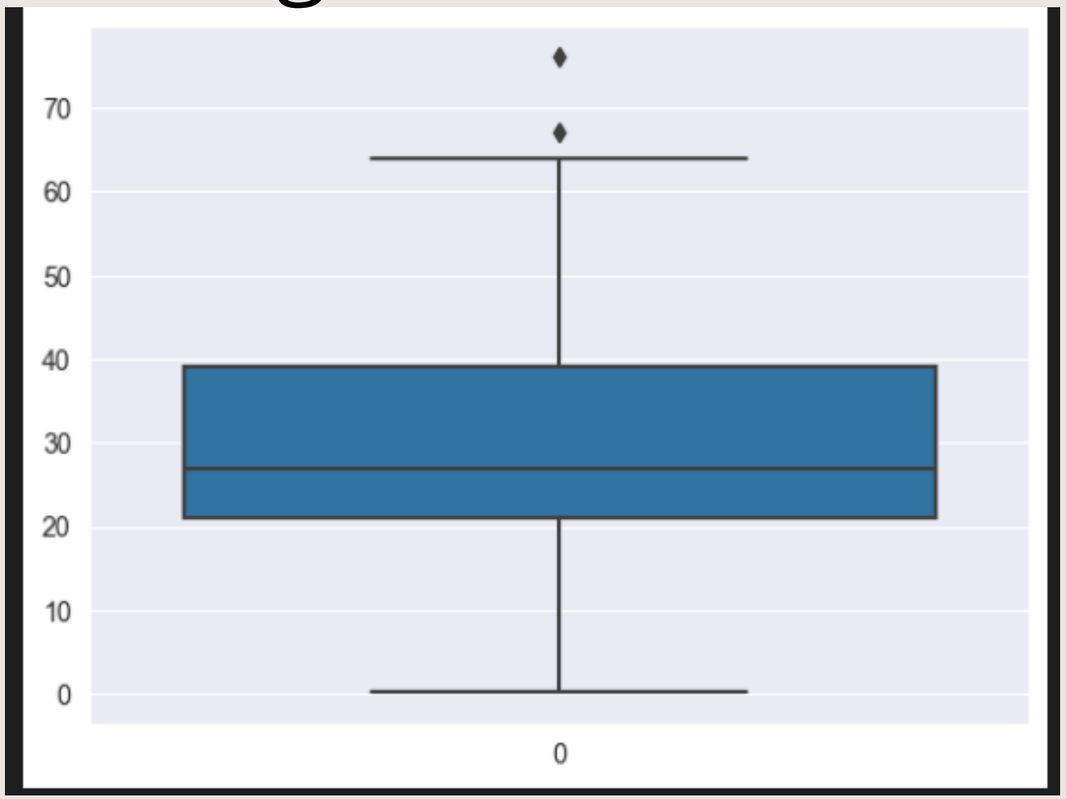
histogram of age



Analyse 9

information about age

count	332.000000
mean	30.272590
std	14.181209
min	0.170000
25%	21.000000
50%	27.000000
75%	39.000000
max	76.000000



Number of passengers who was <18

41

Analyse 10

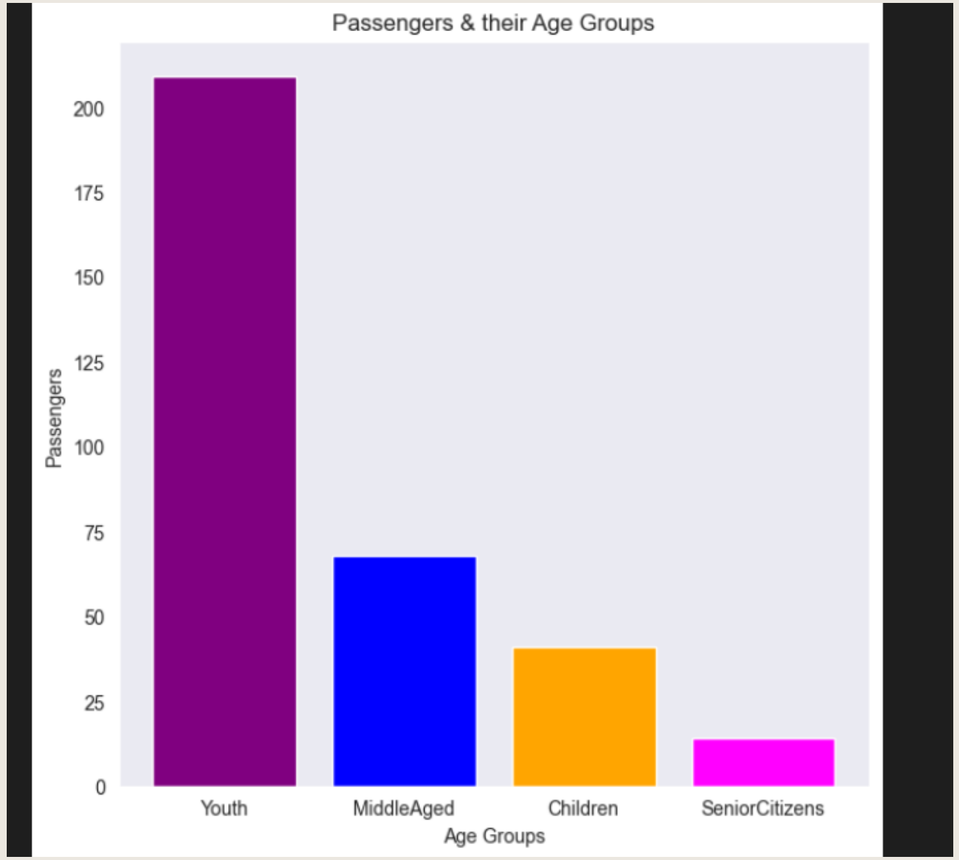
dividing to categories by age

labels = ["Children","Youth","MiddleAged","SeniorCitizens"]

bins = [-np.inf,17, 39, 59, np.inf]

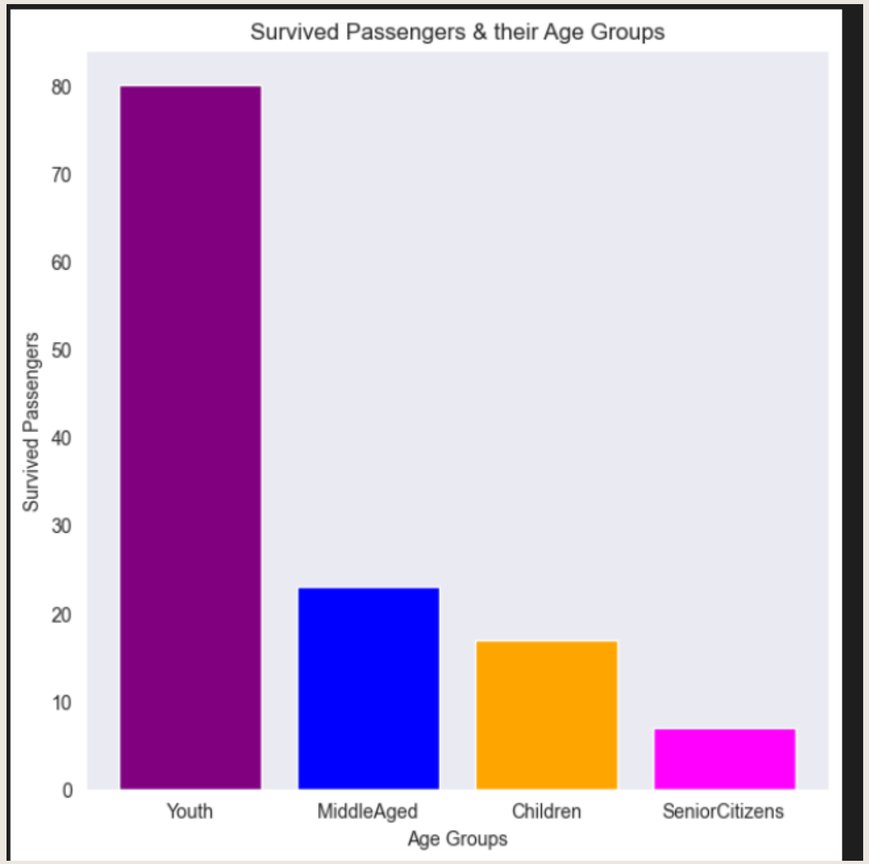
Analyse 11

Passengers & their Age Groups



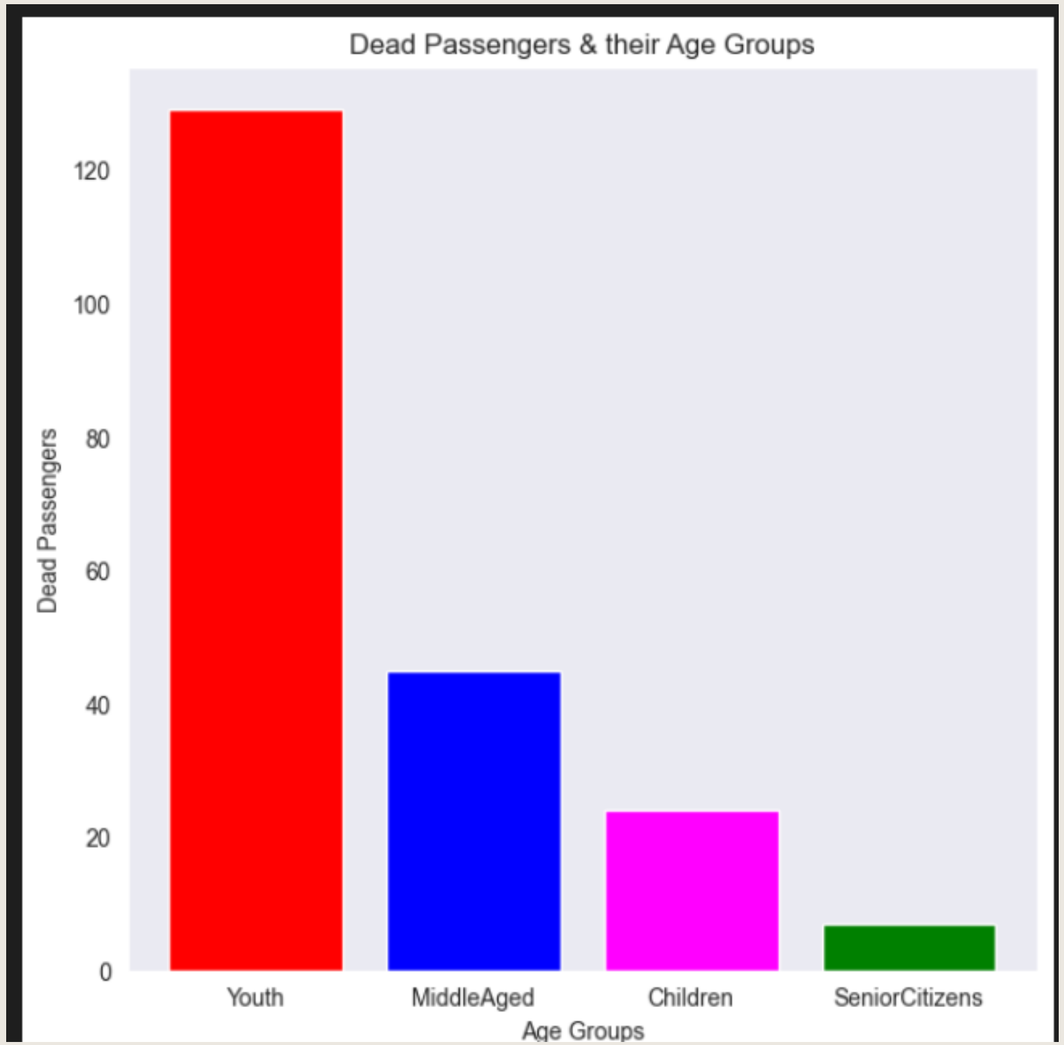
Analyse 12

Survived passengers and their age group



Analyse 13

Dead Passengers & their Age Groups

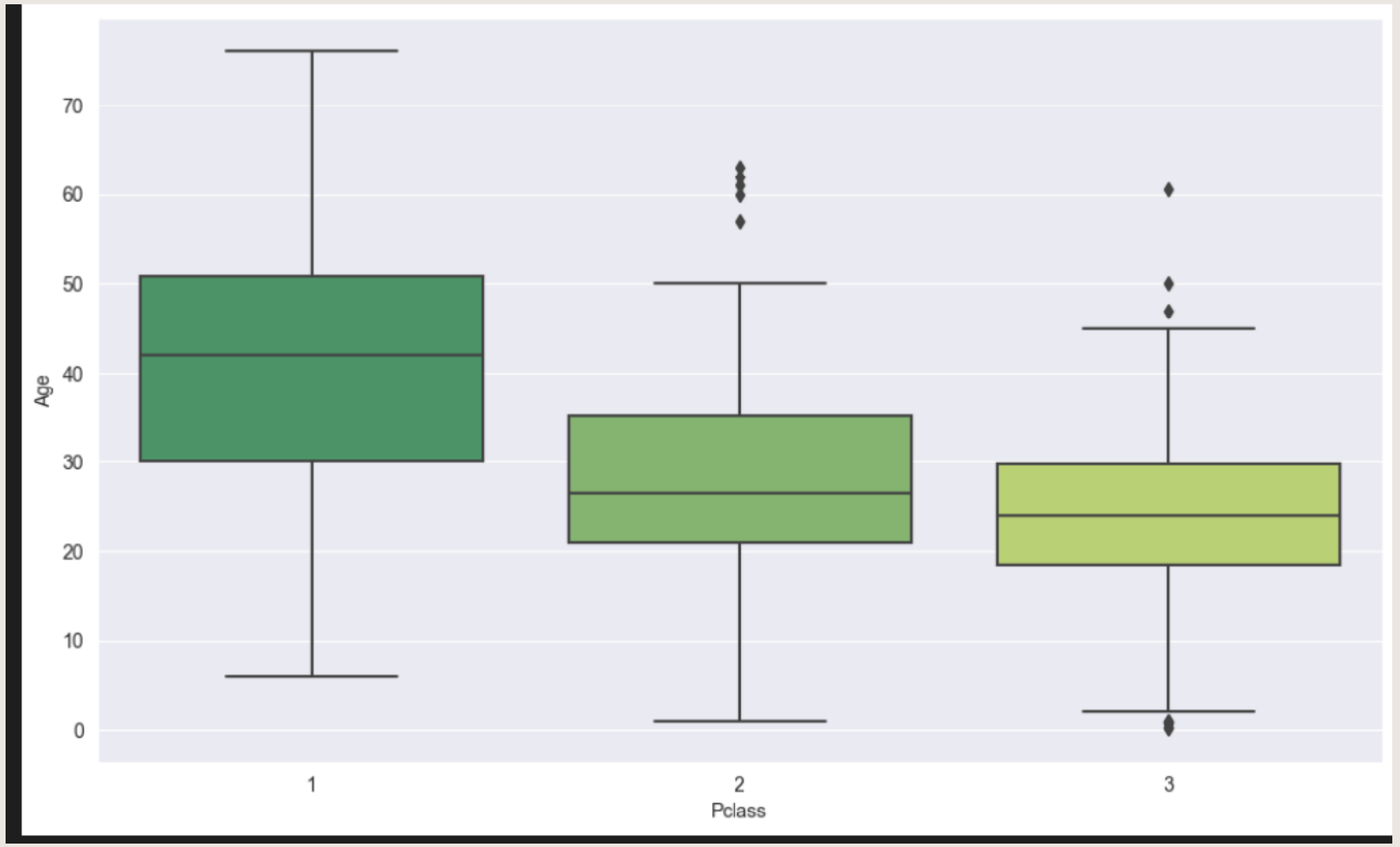


Analyse 14

Showing ages in each pclass

We can see the wealthier passengers in the higher classes tend to be older, which makes sense.

By the analyse above, in the 1st class most of people are older than other classes.



Analyse 15

filling nan values in age

by analyse before, we can see that the wealthier people are older than other classes. filling nan values by this analyse if pclass 1 =37, and 29, 24

Analyse 16

how many males and females had cabins

```
..   47  males had cabins and  44  females had cabin
```

Analyse 17

Showing Embarked column as pie

```
titanic['Embarked'].unique()
```

✓ 0.3s

```
array(['Q', 'S', 'C'], dtype=object)
```