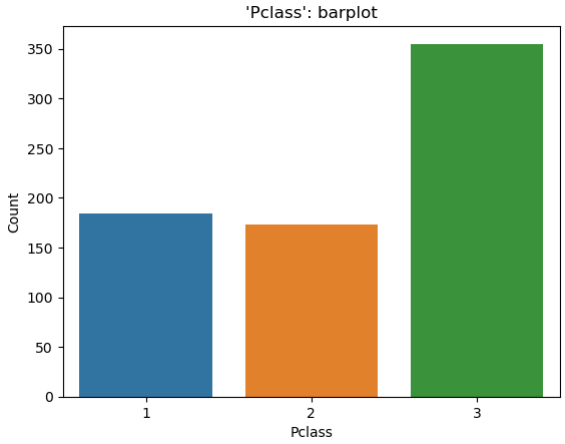
Summary ‘titanic’ project

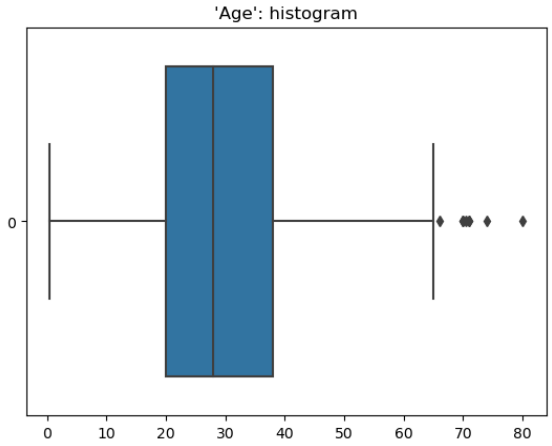
EDA

Passenger class



The majority of people who traveled on the Titanic, specifically 355 individuals, were in third class.

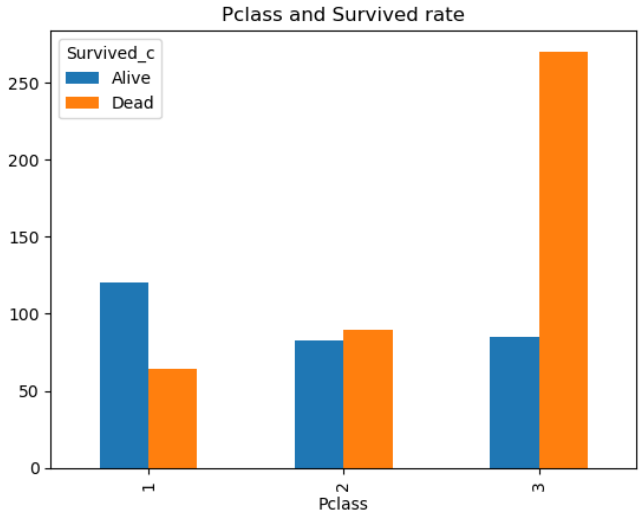
Age



The median age is 28. Additionally, the first quartile (Q1) is 20 and the third quartile (Q3) is 38.

Hypothesis

Determine if the survival rate is associated to the class of passenger:



$H0$: *The survival rate and passenger class are indepedent*

Ha: *The survival rate and passenger class are not independent*

Pclass	Survived rate	
	Alive	Dead
1	120	64
2	83	90
3	85	270

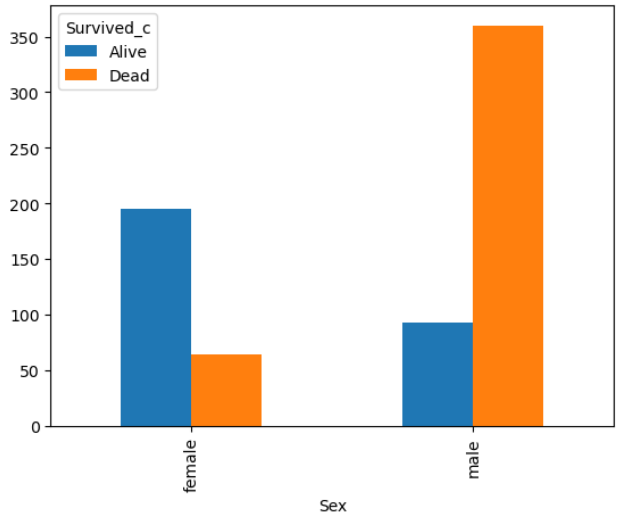
$Chi - square\ statistic$: 91.0807454879102

$P - value$: $1.6675060315554516e - 20$

$Degrees\ of\ freedom$: 2

Since 91.08 is greater than the DP. We reject H0. In consequence, the survival rate and passenger class are not independent.

Determine if the survival rate is associated to the gender:



$H0$: $The\ survival\ rate\ and\ gender\ are\ indepedent$

Ha: $The\ survival\ rate\ and\ gender\ class\ are\ not\ indepedent$

Sex	Survived rate	
	Alive	Dead
Female	195	64
Male	93	360

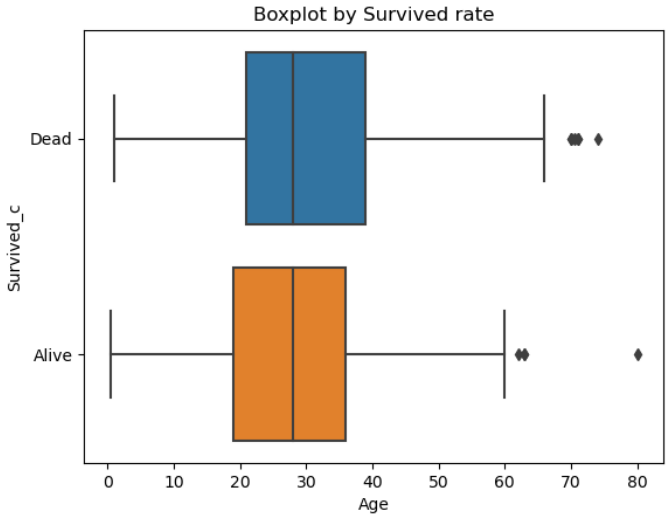
$Chi - square\ statistic$: 202.86944877617123

$P - value$: $4.939416685451492e - 46$

$Degrees\ of\ freedom$: 1

Since 202.86 is greater than the DP. We reject H0. In consequence, the survival rate and 'sex' are not independent.

Determine the survival rate is associated to the age:



	Q1	Q2	Q3
Dead	21	28	39
Alive	19	28	36

Despite some outliers, the median of people who died and the median of who lived are the same. Besides, the first and third quartiles are not so different. In conclusion, the age and the survival rate are not related.