Summary ‘titanic’ project

# EDA

## Passenger class

A bar chart with different colored squares

Description automatically generated

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

## Age

A graph with a blue rectangle and black lines

Description automatically generated

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:

A graph of a number of people

Description automatically generated with medium confidence

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

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:

A graph of a person and person

Description automatically generated

|  |  |
| --- | --- |
| Sex | Survived rate |
| Alive | Dead |
| Female | 195 | 64 |
| Male | 93 | 360 |

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:

A graph with a box plot and a number of numbers

Description automatically generated with medium confidence

|  |  |  |  |
| --- | --- | --- | --- |
|  | 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.