**Data Retrieval**

**Q1)**

[1] 891 12

**Data Pre-processing**

**Q2)**

29.69912

**Q3)**

Graphical user interface, text, application, email

Description automatically generated

**K-means Clustering**

**Q4)**

**Within Sum of Square**

**Graphical user interface

Description automatically generated**

**Between Sum of Square**

**Graphical user interface

Description automatically generated**

**Q5)**

The desired number of clusters we select for the analysis is **5**

**Tableau / R Integration**

**Q 6)**

**Graphical user interface, application

Description automatically generated**

**Q 7)**

**Graphical user interface, application

Description automatically generated**

**Q 8)**

**Graphical user interface, application

Description automatically generated**

**Q9)**

**Graphical user interface, table

Description automatically generated with medium confidence**

**Q10)**

From Cluster 1, we can infer that the majority of passengers that survived are between the age of 9 and 31. The majority of passengers that did not survive are between the age of 23 and 32. This shows that in cluster 1, the survival probability of kids was higher than adults.

In cluster 1, we can also infer that the majority of passengers that survived travelled with 0 or 1 sibling/spouse. We can also note that cluster 1 mostly contains Male passengers.

**Q11)**

From Cluster 4, we can infer that the majority of passengers that survived are between age 20 and 36. The majority of passengers that did not survive are between the age of 17 and 30. This shows that in cluster 4, the survival probability of passengers below the age of 20 is lesser than the ones above the age of 30.

In cluster 4, we can also infer that the majority of passengers that survived embarked from Cherbourg. Cluster 4 includes both Male and Female passengers, but it majorly contains those Male passengers that survived.

**Q12)**

**Graphical user interface, chart, application, Excel, bar chart

Description automatically generated**

**Q13)**

Cluster 2 and Cluster 4

**Q14)**

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender |  |  |
| Ideal Passenger Class |  |  |
| Ideal Age Category |  |  |
| Ideal Embarked point |  |  |
| Ideal number of siblings |  |  |

**Q15)**

**Chart, bar chart

Description automatically generated**

**Q16)**

For females, cluster 3 has higher probability of survival

**Q17)**

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Male | Female |
| Ideal Passenger Class |  |  |
| Ideal Age Category |  |  |
| Ideal Embarked point |  |  |
| Ideal number of siblings |  |  |

**Q18)**

Chart, waterfall chart

Description automatically generated

**Q19)**

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Female | Male |
| Ideal Passenger Class | Female/Pclass= 2 | Male/Pclass= 3 |
| Ideal Age Category |  |  |
| Ideal Embarked point |  |  |
| Ideal number of siblings |  |  |

**Q20)**

Chart, waterfall chart

Description automatically generated

**Q21)**

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Female | Male |
| Ideal Passenger Class | Female/Pclass=2 | Male/Pclass=3 |
| Ideal Age Category | Female/AgeCat(17-32) | Male/AgeCat(17-32) |
| Ideal Embarked point |  |  |
| Ideal number of siblings |  |  |

**Q22)**

**Chart, waterfall chart

Description automatically generated**

**Q23)**

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Female | Male |
| Ideal Passenger Class | Female/Pclass=2 | Male/Pclass=3 |
| Ideal Age Category | Female/AgeCat(17-32) | Male/AgeCat(17-32) |
| Ideal Embarked point | Female/Southampton | Male/Southampton |
| Ideal number of siblings |  |  |

**Q24)**

Chart, waterfall chart

Description automatically generated

**Q25)**

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Female | Male |
| Ideal Passenger Class | Female/Pclass=2 | Male/Pclass=3 |
| Ideal Age Category | Female/AgeCat(17-32) | Male/AgeCat(17-32) |
| Ideal Embarked point | Female/Southampton | Male/Southampton |
| Ideal number of siblings | Female/SibSp=0 | Male/SibSp=0 |

**Q26)**

The profiles of passengers with best chance of survivability for the two clusters are as follows:

**Cluster 2:** The chance of survival is highest for females of passenger class 2 between the age of 17-32 who embarked from Southampton and travelled with no Siblings/Spouse.

**Cluster 4:** The chance of survival is highest for males of passenger class 3 between the age of 17-32 who embarked from Southampton and travelled with no Siblings/Spouse.