**Assignment 7 – Integrated Analysis with R**

1. **Data Retrieval**

Q1) Output on Console Window:

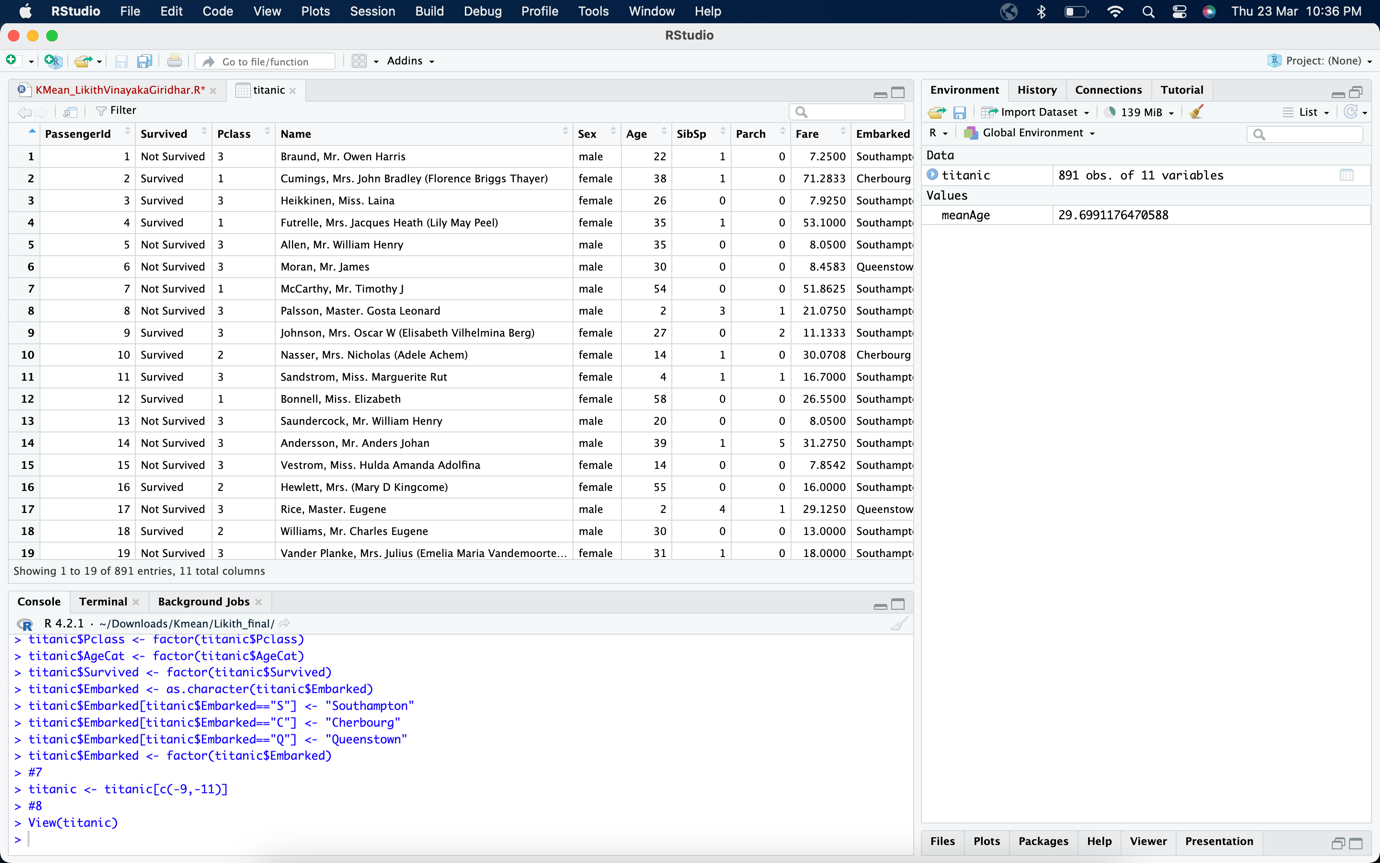
> dim(titanic)

[1] 891 12

1. **Data Pre-Processing**

Q2) The Average Age that we get is 29.69912

Q3)



1. **K-Means Clustering**

Q4)

Graphical user interface, application

Description automatically generated

Graphical user interface

Description automatically generated

Q5) The desired number of clusters we have chosen for the analysis is **5** as it is the elbow point in the plot.

1. **Tableau / R Integration**

Q6)

Graphical user interface, application

Description automatically generated

Q7)

Graphical user interface, application

Description automatically generated

Q8)

Graphical user interface, application

Description automatically generated

Q9)

Calendar

Description automatically generated

Q10)

Cluster 1 suggests that most of the survivors were aged between 8.5 and 30, while the majority of those who did not survive were aged between 23 and 31. This implies that in Cluster 1, the likelihood of survival was greater for children than adults. Additionally, it can be inferred from Cluster 1 that the majority of survivors travelled with no more than one sibling or spouse, and the cluster was predominantly composed of male passengers.

Q11)

Based on Cluster 4, it can be deduced that most survivors were aged between 18.5 and 35, while the majority of those who did not survive were aged between 15.5 and 30. This suggests that in Cluster 4, passengers under the age of 20 had a lower chance of survival compared to those above 30. Moreover, it can be inferred from Cluster 4 that most survivors embarked from Cherbourg. The cluster consists of both male and female passengers, but the majority are male survivors.

Q12)

Graphical user interface, chart, Excel, bar chart

Description automatically generated

Q13) Cluster 2 & Cluster 4

Q14)

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

Q15)

Chart, Excel, bar chart

Description automatically generated

Q16) For a female, cluster 3 has the highest 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 | Male | Female |
| Ideal Passenger Class | Male/PClass = 1 | Female/PClass = 1 |
| Ideal Age Category |  |  |
| Ideal Embarked point |  |  |
| Ideal number of siblings |  |  |

Q20)

Chart, waterfall chart

Description automatically generated

Q21)

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Male | Female |
| Ideal Passenger Class | Male/PClass = 1 | Female/PClass = 1 |
| Ideal Age Category | Male/AgeCat (33-48) | Female/AgeCat (17-32) |
| Ideal Embarked point |  |  |
| Ideal number of siblings |  |  |

Q22)

Chart, waterfall chart

Description automatically generated

Q23)

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Male | Female |
| Ideal Passenger Class | Male/PClass = 1 | Female/PClass = 1 |
| Ideal Age Category | Male/AgeCat (33-48) | Female/AgeCat (17-32) |
| Ideal Embarked point | Male/Southampton | Female/Cherbourg |
| Ideal number of siblings |  |  |

Q24)

Chart, waterfall chart

Description automatically generated

Q25)

|  |  |  |
| --- | --- | --- |
|  | Cluster 2 | Cluster 4 |
| Ideal Gender | Male | Female |
| Ideal Passenger Class | Male/PClass = 1 | Female/PClass = 1 |
| Ideal Age Category | Male/AgeCat (33-48) | Female/AgeCat (17-32) |
| Ideal Embarked point | Male/Southampton | Female/Cherbourg |
| Ideal number of siblings | Male/SibSp = 0 | Female/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 males of passenger class 1 between the age of 33-48 who embarked from Southampton and travelled with no Siblings/Spouse.

**Cluster 4:** The chance of survival is highest for females of passenger class 1 between the age of 17-32 who embarked from Cherbourg and travelled with no Siblings/Spouse.