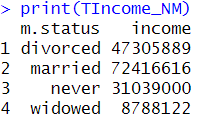
**1.Summarizing Data**

**Question 1.1: Summary Table**

**Create a table to show the total income by each category of marital status**



Here we can see the distribution of total income based on the marital status. As we can see the person who is married has the highest amount of income and the person who is widowed has the lowest amount of income.

**Question 1.2: Calculate the mean rounded by two decimals**

**1)Calculate the mean age of respondents born in Asia**



Here we can see that the mean age of respondents born in Asia is 40.96 which is \

**2) Calculate the mean age of respondents born in Asia weighted by the number of children they have.**

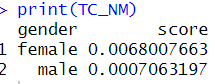


The mean age of respondents from Asia weighted by the number of children they have is 40.61.

As we can see the weighted mean decreases a bit than the normal mean but it is ponly a slight decrease not a large one.

**Question 1.3: Table Comparison**

**Create a table to show the mean score on the political awareness test for males compared to females**



As we can infer, the mean score on the political awareness test for males and females is very close but based on the Table Comparison Females have a better mean score than Males.

**Question 1.4:**

**Calculate the 34th and 63rd percentiles of percentage of time taken on the test**

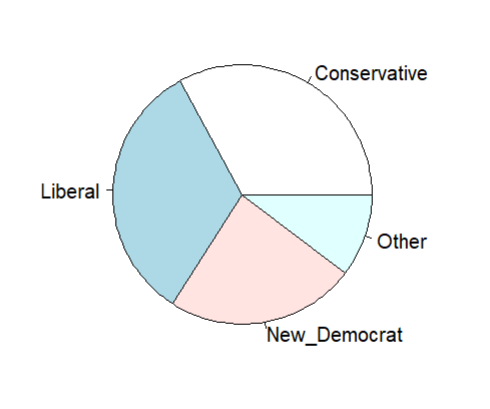
****

Here we can check that the 34th quantile and the 63rd quantile are 0.20 and 0.61 respectively. This provides us with insights on the time taken to complete the test among different respondents.

**3. Organizing Data**

**Question 2.1: Pie Chart**

**a) Create a pie chart showing the number of respondents by Political Affiliation.**



This showcases the Pie Chart showing the number of respondents by Political Affiliation

**b) Which Political Affiliation contain the most respondents (remember each row of your study file represents one respondent)**

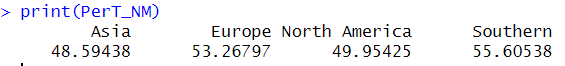
Here as we can see from the Pie Chart, Liberals seem to possess the most respondents when it comes to Political Affiliation.

**c) Which Political Affiliation has the fewest respondents?**

From the Pie Chart, the other category has the fewest respondents. Hence, it can be assumed that other consists of many smaller parties and hence didn’t form the largest number.

**Question 2.2: Summary Table**

**a) Create a table that shows the percentage of respondents from each Region that are in the Treatment group.**



Here we can see that the Southern Region has the greatest number of respondents when it comes to the Treatment Group.

**b) Which region has the highest percentage of people in the Treatment group?**

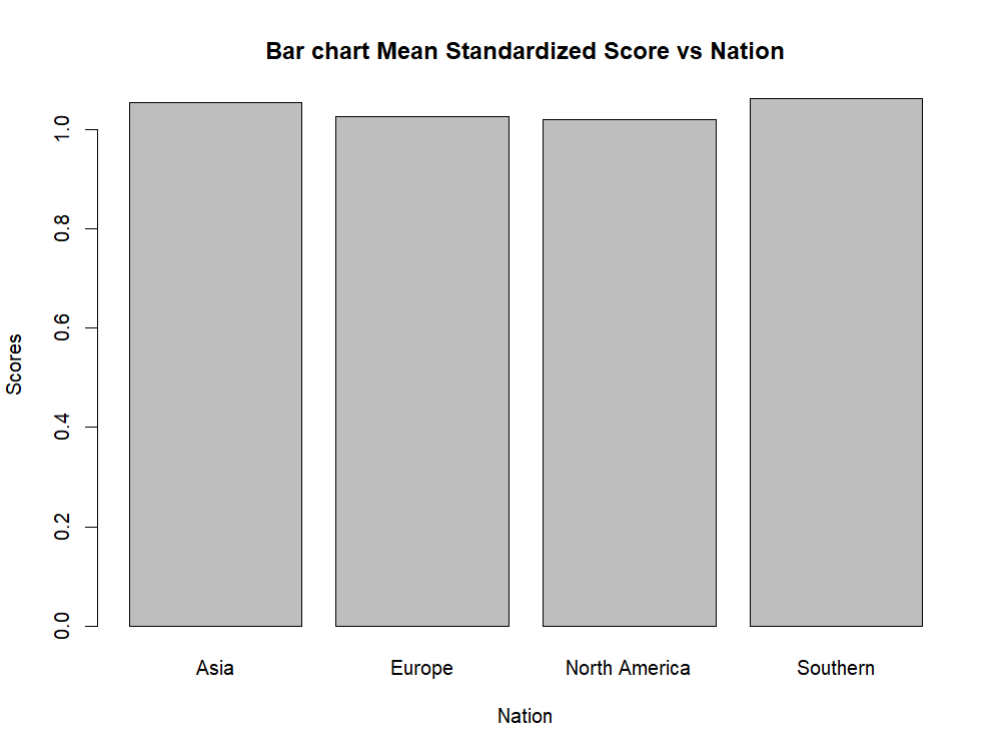
As we can see from the Table Summary Southern region is the region with the highest percentage of people who belong to the Treatment Group.

**c) Which region has the lowest percentage of people in the Treatment group?**

Asia region has the lowest percentage of people who belong to the Treatment group.

**Question 2.3: Bar Chart**

**a) Create a bar chart showing the mean Standardized Test Score on the Political Awareness Test for each Region**.



This is the Bar Chart between the mean Standardized Score and the Nation. As we can see the mean scores are pretty similar across the nations but there are clear highest and lowest.

**b) Which Region has the lowest mean score?**

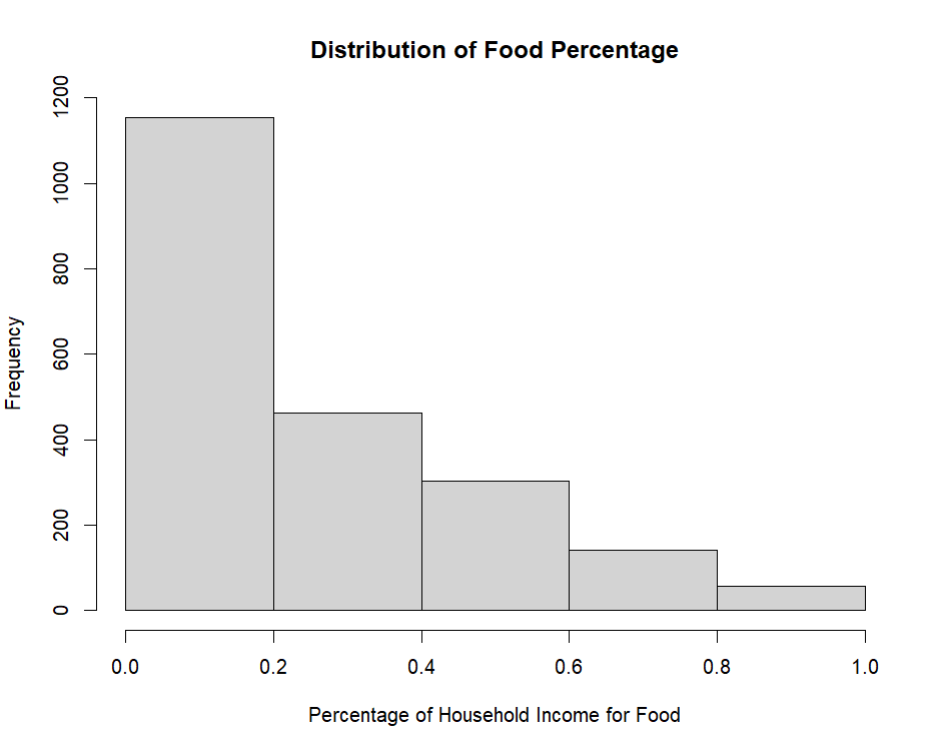
According to the Bar Graph the lowest mean score nation is North America.

**c) Which region has the highest mean score?**

According to the Bar Graph the highest mean Standardized score is with the Southern Region.

**Question 2.4: Histogram**

**a) Create a histogram with 5 bins showing the distribution of the percentage of household income going to food.**



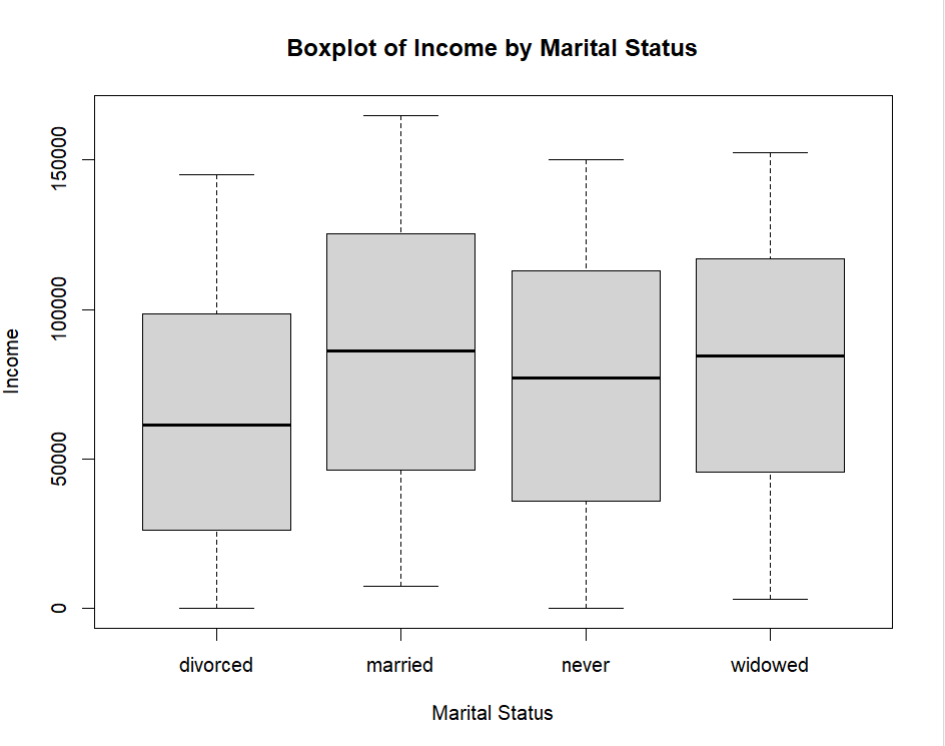
This is the Histogram showing the distribution of income which goes to Food.

**b) Which range of values has the highest frequency?**

As we can see from the Histogram 0 to 20% has the highest range of value and frequency.

**Question 2.5 Box Plots**

**a) Create a sequence of box plots showing the distribution of income separated by marital status.**



**b) According to the charts, which marital status has the highest average income?**

According to the Box Plots married people have the highest average income when compared to other marital status.

**c) Which marital status has the lowest average income?**

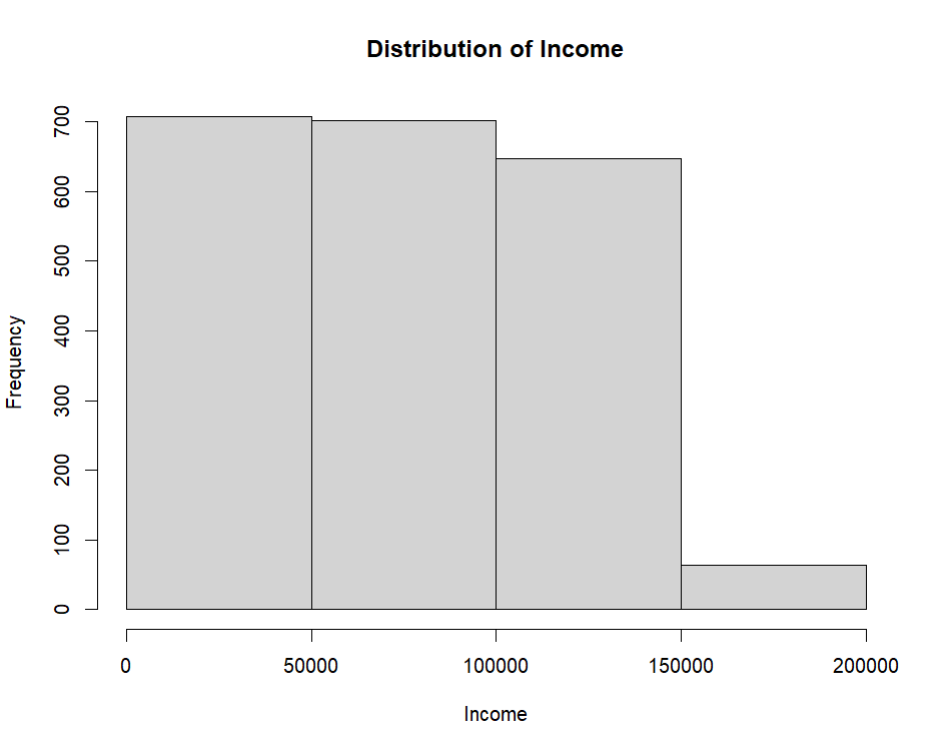
Divorced marital status has the lowest average salary when compared to other marital statuses.

**d) Which marital status has the greatest variability in income?**

Married marital status have the greatest variability when it comes to income.

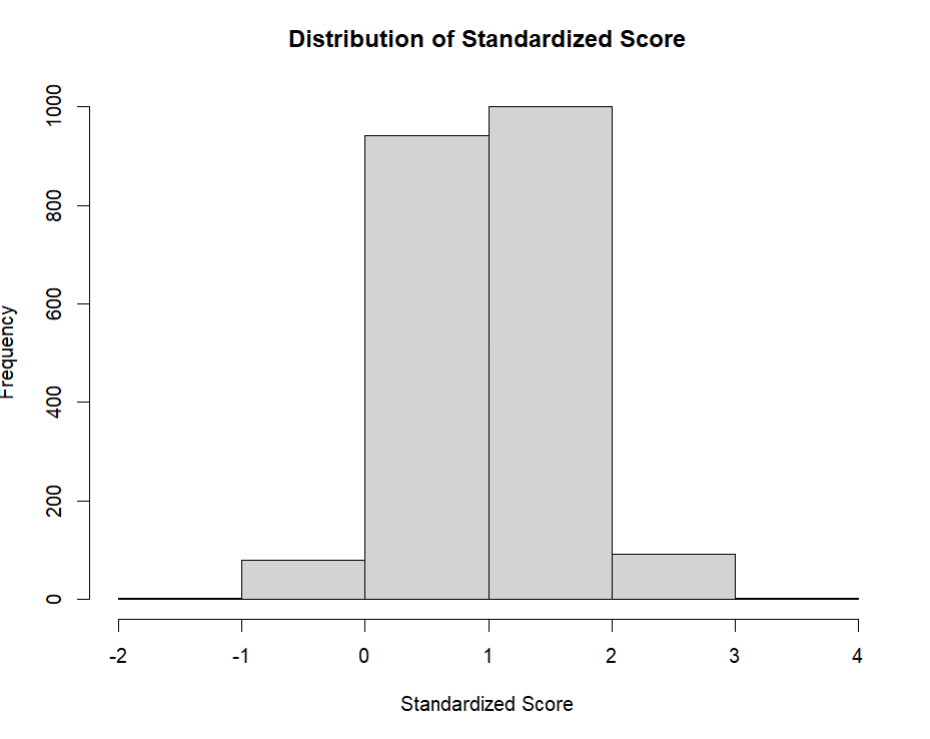
**Question 2.6: Scatter Plots**

**a) Create a histogram for income.**



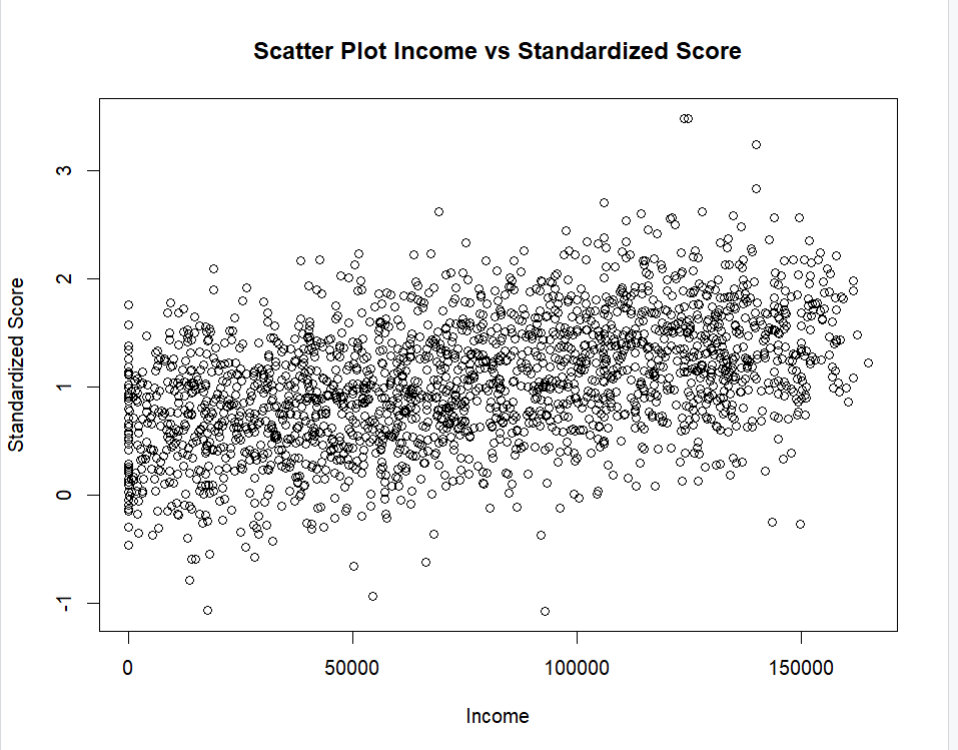
This is the histogram for income. It showcases the distribution of income and there are a lot of people when it comes to income between 0 and 150000.

**b) Create a histogram for standardized score.**



This is the Histogram for the distribution of Standardized Score and as we can see most of the scores lie between 0 and 2.

**c) Create a scatter plot showing the relationship between the income and standardized score. (note: income should be on the x-axis, standardized score should be the y-axis)**



This is the Scatter Plot between Standardized Score and Income. Here we can see there are a lot of results but there is a positive correlation as the graph is trending slightly upwards. The Correlation Score for this comes out to be 0.456841

**d) Calculate a correlation coefficient between these two variables. What conclusion you draw from it?**



Here as we can see this is the correlation score and it is positive.