**PRACTICAL 5**

Create a RMarkdown (.Rmd) document that answers and addresses the following requirements.

1. Getting to know the data:
2. Import the data (<http://becomingvisual.com/rfundamentals/winter_olympic.csv>)
3. View the data
4. Look at column names
5. Look at dimension of data (rows and columns)

| Solution: |
| --- |

1. Data is currently sorted by Rank. Sort data by total medals and country. Assign sorted data to a new data frame. Call it sort\_total.

| Solution: |
| --- |

1. Use describe() function to look at data.
2. If function does not work, first import library: library(psych)

| library  Solution: |
| --- |

1. Look at some statistics
2. What is median of number of gold, silver, bronze and total medals?
3. Also look at the mean and total number of G, S, B and T medals

| Solution: |
| --- |

1. More statistics
2. For Gold, look at summary stats, including: IQR, min, max, mean, var, sd, skew
3. Use summary() and describe(). (May need to install library(psych) )

| Solution: |
| --- |

1. More statistics - subset
2. Redo above statistics, this time group by Region
3. Which region won the highest mean total medals?
4. How many countries are in this Geographic Region?
5. How many countries are in the EUROPE group?
6. What is the max number of medals won? What country won the max?

| Solution: |
| --- |

1. More statistics – correlations
2. Explore correlations between Total medals and number of Gold and Bronze
3. What is the correlation between Rank and Total medals? Is this expected or surprising?

| Solution:     The correlation is negative as higher values of Total tend to be associated with lower values of the Rank. |
| --- |

1. Import the GDP dataset and compute the measures of central tendency for 2017. (Divide by a trillion, and use na.rm = TRUE when computing the measures.)
2. Find the mean
3. Find the median
4. Find the range
5. Find the quantile

| Solution: |
| --- |

\*\*\*\*\*\*\*\*