

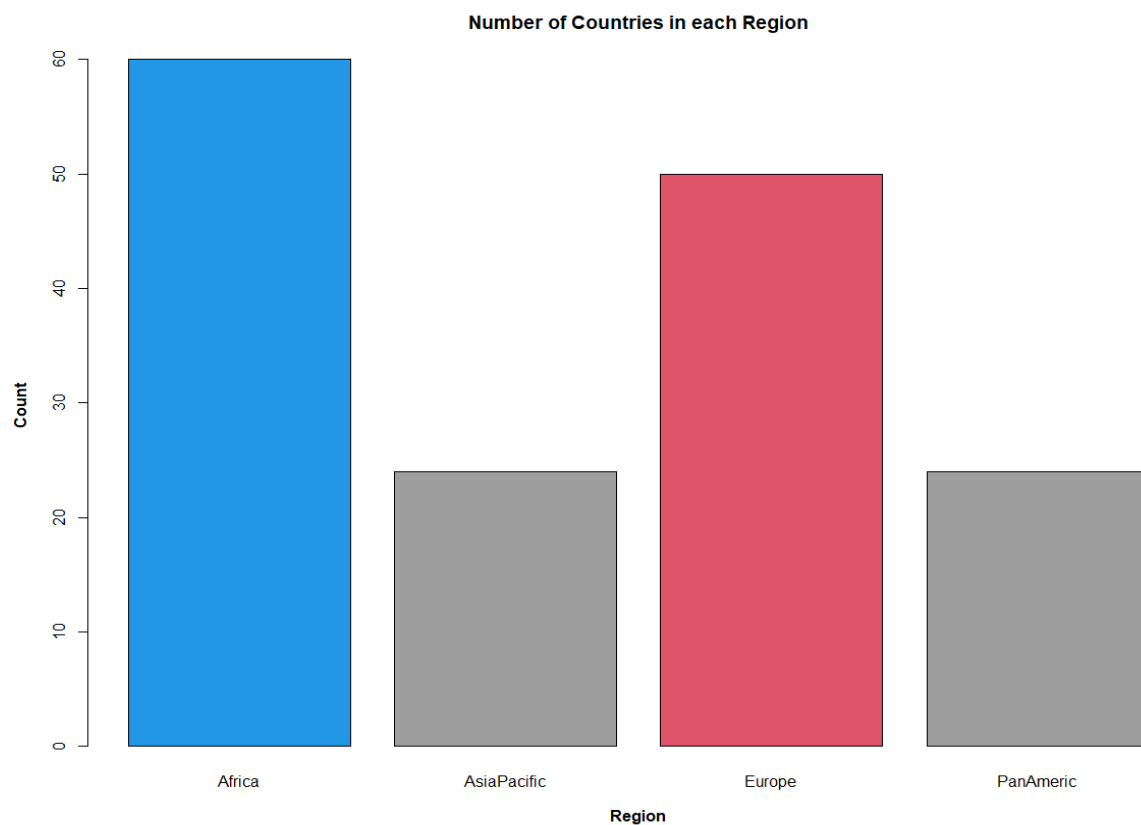
Sang Hwa Lee

INST 314

June 24, 2022

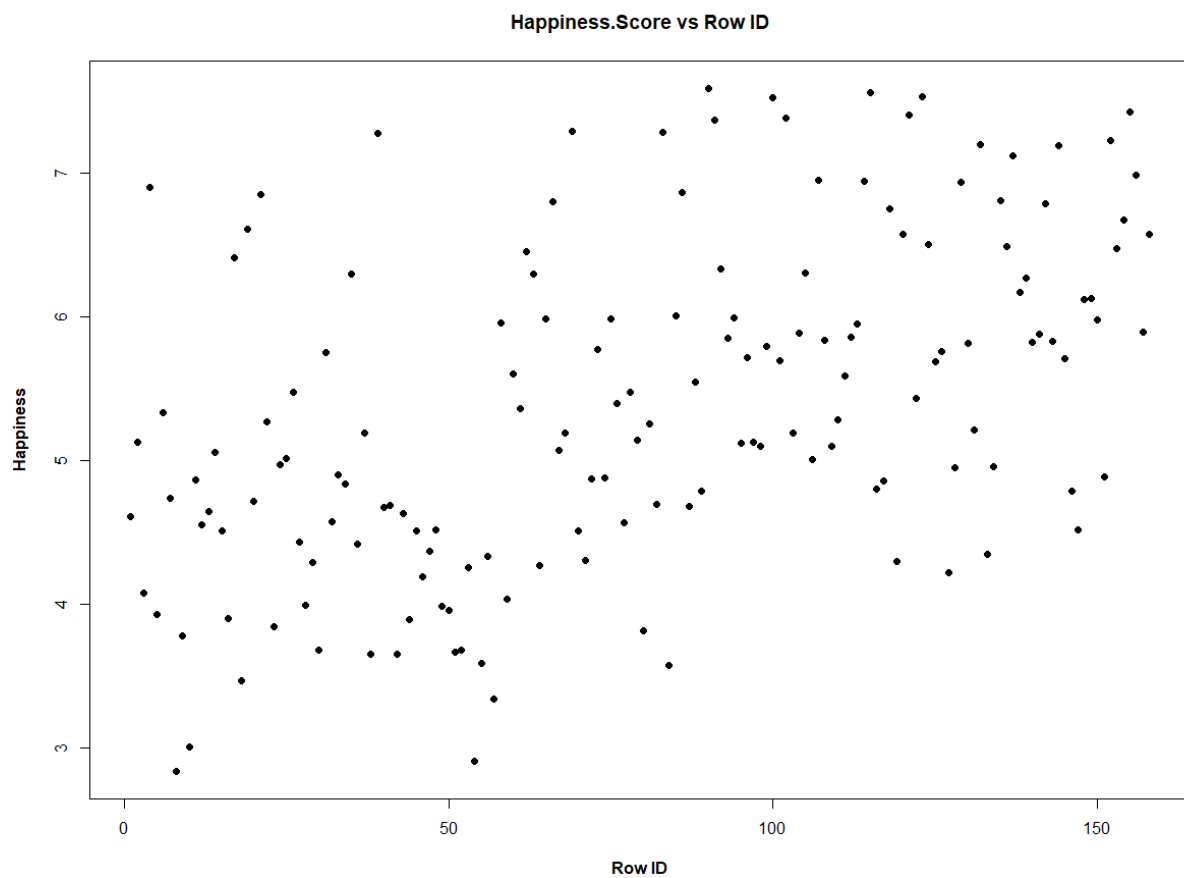
### Assignment 4 - Data Visualization

# Task 3 - What I learn from this chart?



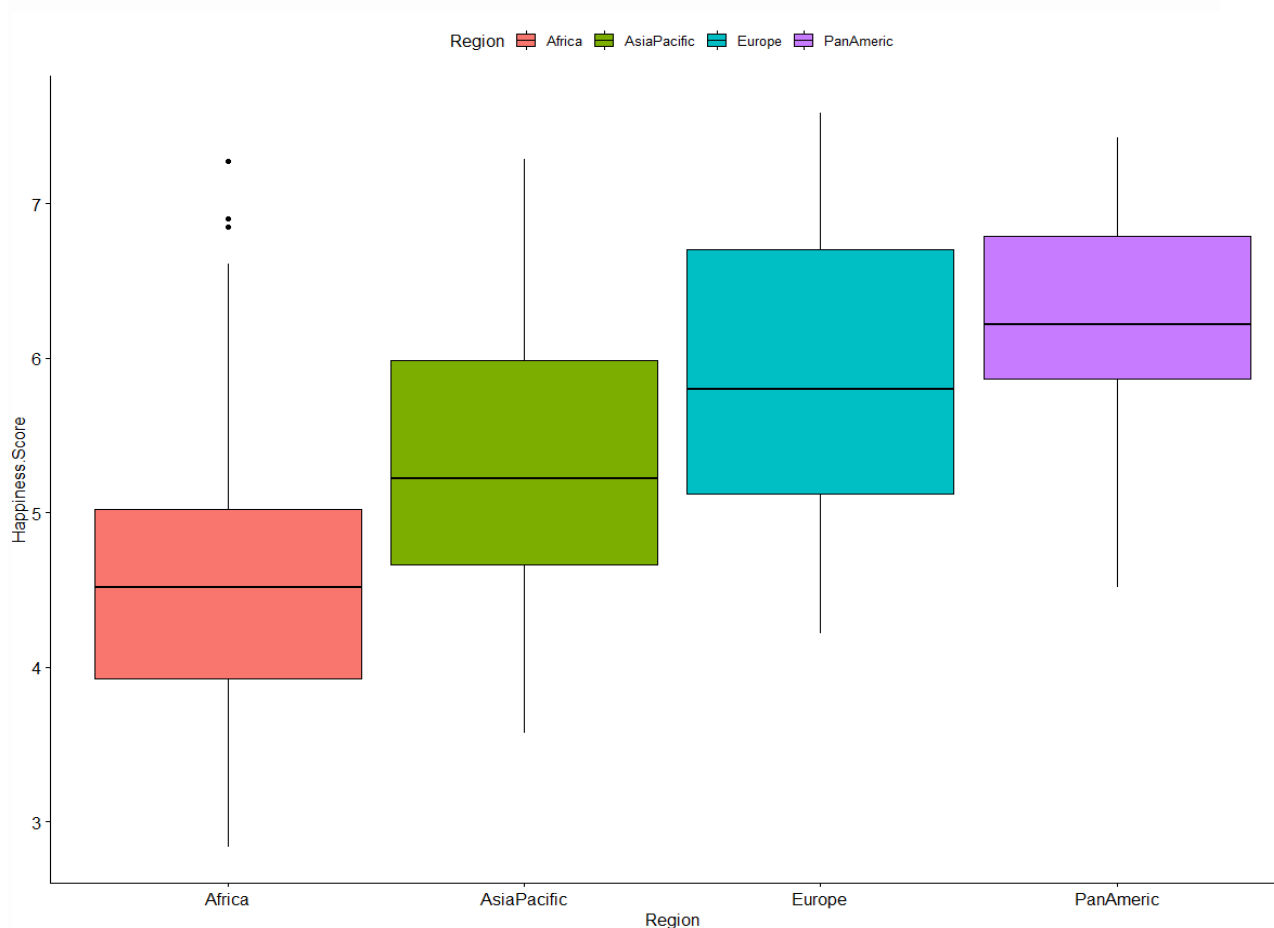
For this task, I created the graph above using the `plot()` function. I learned that I need to define only one axis to create a bar chart. I also learned that to create a bar chart, need to first create a factor of the region column, so that it summarizes the count of each column. What I can see from this chart is that Africa contains the largest amount of data (60). Next up was Europe, AsiaPacific, and PanAmeric.

## Task 4 - What I learn from this chart?



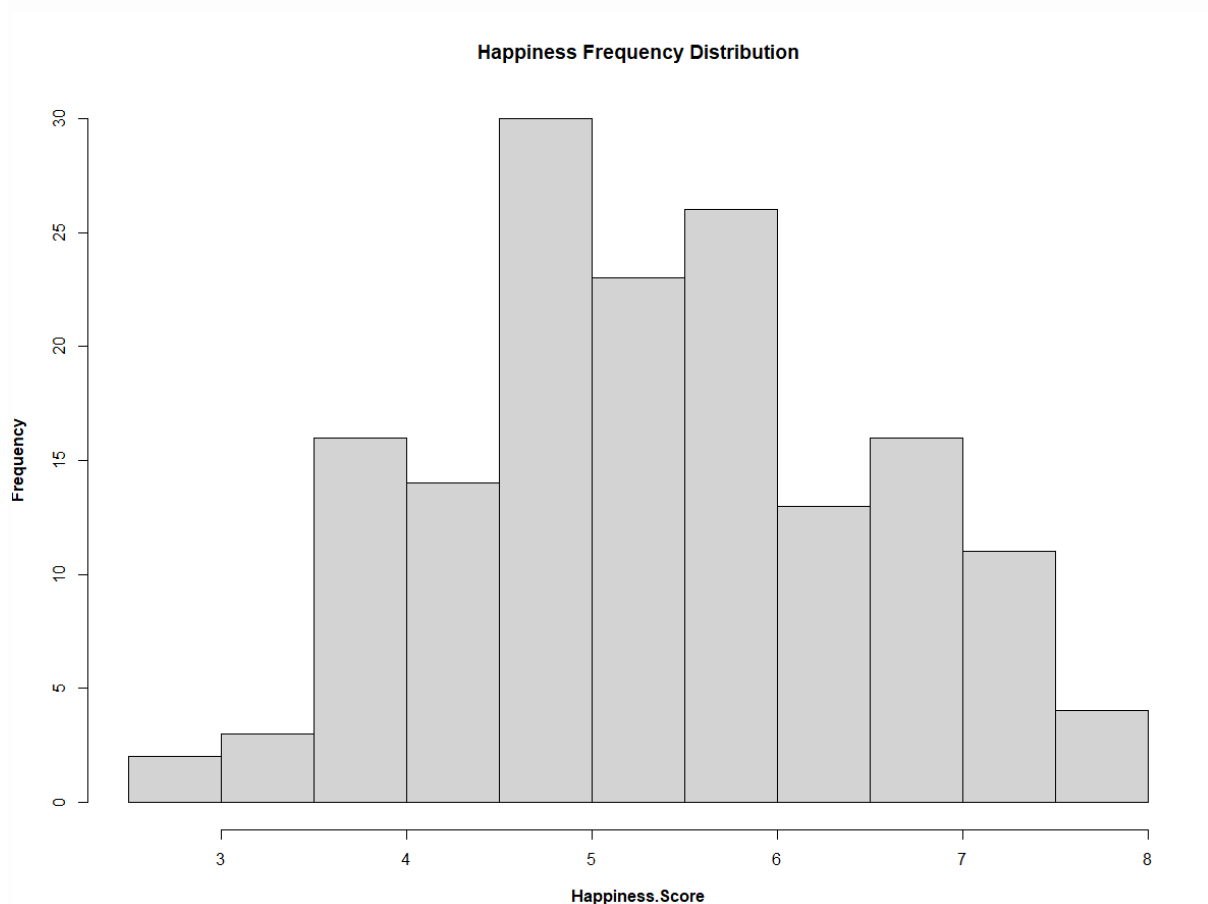
It was a similar process to the previous action 3, but this is a scatter plot. To plot the happiness vs. row ID Firstly, I have created the row ID columns with the `row.names()` function. After creating the ID, I plotted Happiness vs Row ID and also defined xlabel, ylabel, and title. From the above graph, we can see that the happiness values increase as the row ID increases. I can say it has a linear relationship.

## Task 6 - What I learn from this chart?



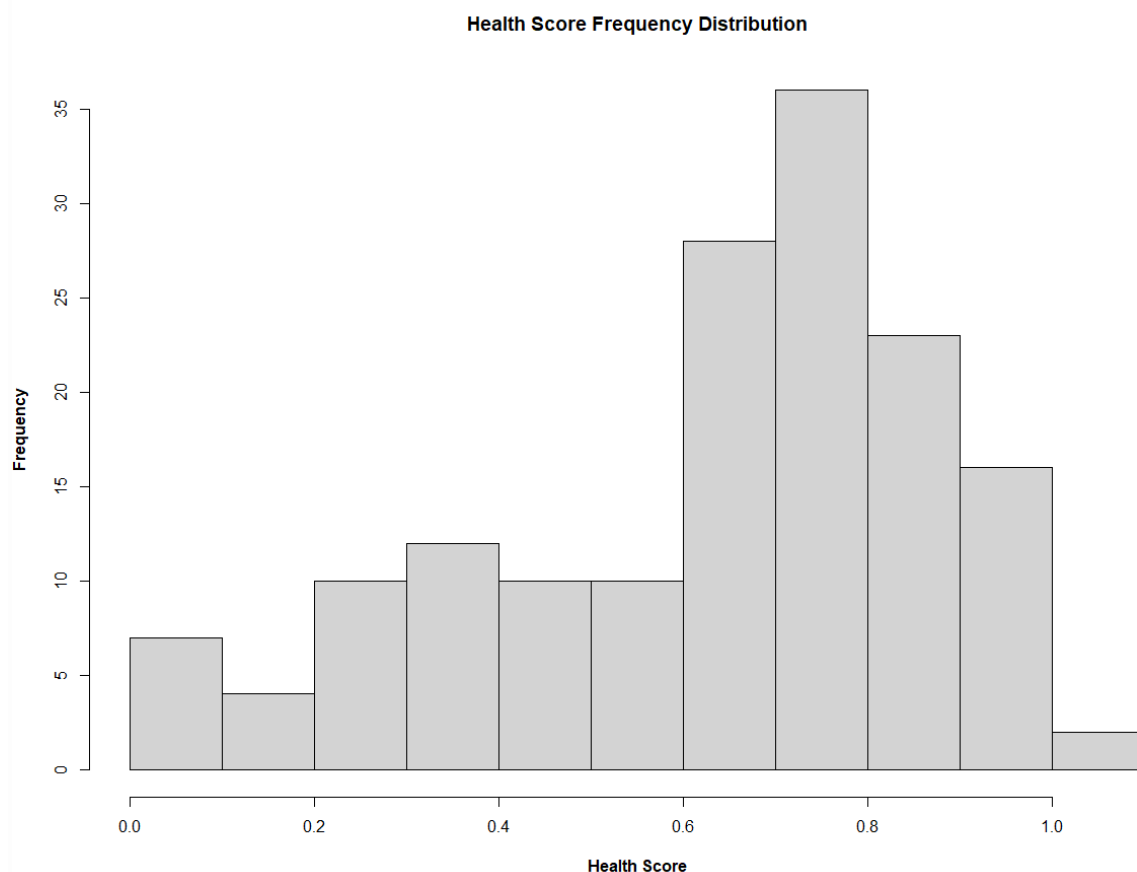
Task 6 is connected task 5. There is little difference from the result of task 5. To make this chart, I need to put happiness and region columns as a relationship in the `plot()` function and R will create desired plot. As I have already created factor for region, we can directly specify region here in `plot()` function. What I found was related to “~”. In the RStudio, I learned that “~” can be used to represent a relationship. In addition, the chart above shows that happiness is the dependent column that you want to calculate, and the region is an independent column.

## Task 7 - What I learn from this chart?



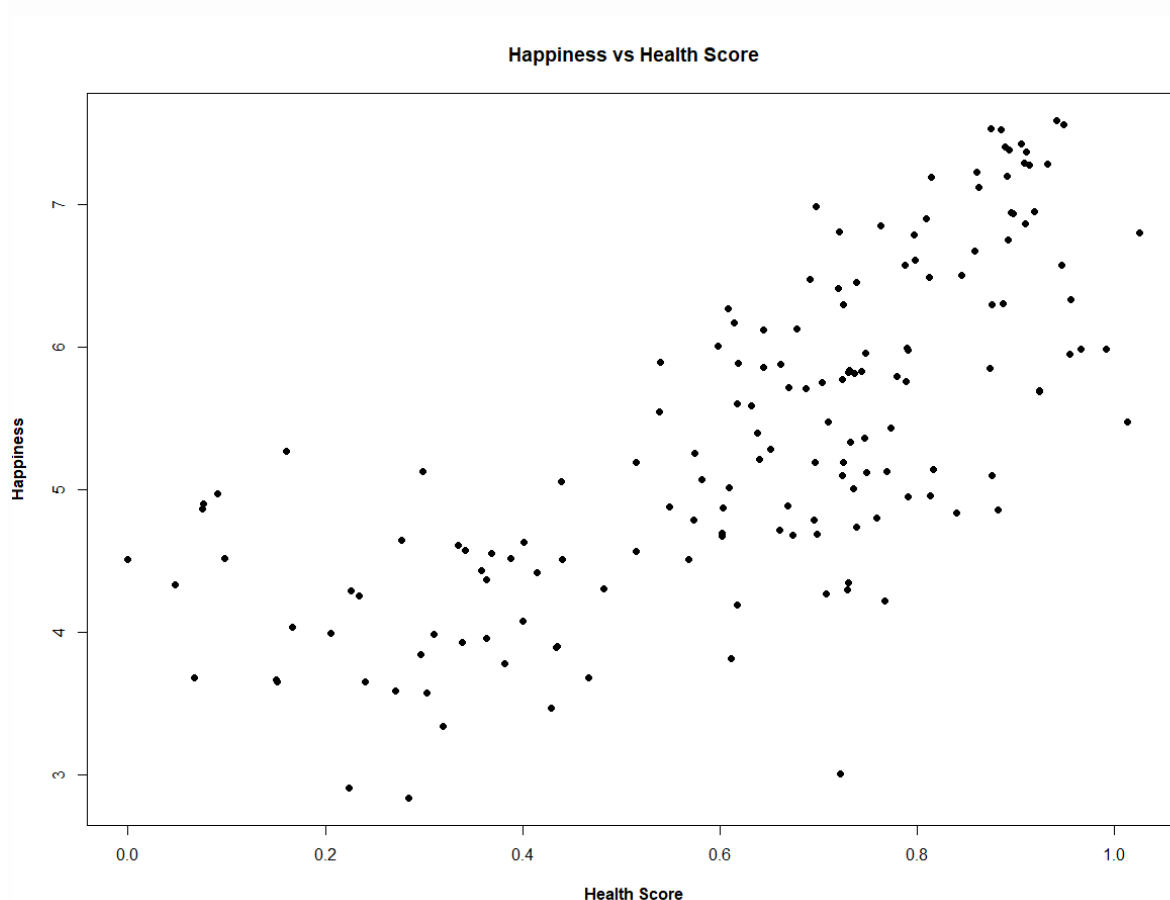
What I learned from this chart is how to make a histogram. I learned that I need to use the hist function for that. I also used the information that needed to define one axis that learned in Task 3 to generate this chart. From the above graph, we can see that the happiness data is uniformly distributed, forming a bell-shaped curve. The maximum of the data looks to be close to the median.

## Task 8 - What I learn from this chart?



The process of task 8 is very similar to the process of 7. But it has different results. If I check at this chart, this is not uniformly distributed data. The graph is skewed to the left and/or includes peaking at the left end. The maximum value of data is the left side of the median, also known as a negatively skewed histogram.

## Task 9 - What I learn from this chart?



To create this chart, I have used the happiness column as the y-axis and the health score as the x-axis. As a result, both columns could be displayed. After that, it is the same as any other graph, defined x-axis, y-axis, and title label. Through this chart, I can check there is a linear relationship between happiness score and health score. I can also say "position relationship." That means when the health score increases, happiness score also increases and vice versa.