**Code Documentation for Data Visualization Assignment**

The assignment was to obtain a dataset of certain values about the population of some countries with data pertaining to the people of those countries. I later have to create a pivot table dataframe from the dataset with ‘year’ along the x-axes, ‘continent’ along the y-axes and filled with data from ‘lifeExp’.

Furthermore, I had to plot a heatmap of the pivot table dataframe using the python seaborn visualization library. I saved the plotted figure as Seaborn\_HeatMap.png.

Finally, I had to make a beautiful heatmap visualization of the pivot table dataframe using Tableau. I then saved a screenshot of the heatmap as Tableau\_HeatMap.png

Explaining the code:

First, I imported the necessary libraries needed to work out the assignment

*# importing needed libraries*

import pandas as pd

import seaborn as sns

After importing the libraries, the next thing is the handle reading the data directly in the code and since we can’t get the csv dataset directly, we can get the data through the link given and then save the dataset as a csv file so we can use that csv dataset later in our tableau work:

In the code snippet below, I assigned the url to the dataset to a variable, ‘csv\_url’. This makes it easier to use the link with other methods as compared to having to type the url again. I then read in the dataset into the code using the pd.read\_csv() pandas method.

*# assigning url of dataset to variable 'csv\_url'*

csv\_url = 'https://raw.githubusercontent.com/resbaz/r-novice-gapminder-files/master/data/gapminder-FiveYearData.csv'

read\_data = pd.read\_csv(csv\_url) *# reading dataset*

Next, I saved the read dataset on my local machine for use in Tableau using the .to\_csv() pandas method. I did this because I can’t read the data from the link directly using the Tableau software. I also printed the values of the dataset to make sure it was fully accessible.

*# saving dataset as csv for use in tableau*

save\_data = read\_data.to\_csv('data.csv')

print(read\_data)  *# checking the data*

Here, I made a pivot table dataframe in pandas using the .pivot\_table() method. This pivot table dataframe takes the ‘continent’ as the y-axes(index) and ‘year’ as the x-axes(columns) and is filled with values from ‘lifeExp’. I then printed the pivot table dataframe to make sure it was made in the required form.

*# creating pivot table*

pivot\_table = read\_data.pivot\_table(

*index*='continent', *columns*='year', *values*='lifeExp')

print(pivot\_table)

Finally, I made a heatmap of the pivot table using Seaborn. This was relatively very easy to implement because the seaborn library comes with an sns.heatmap() method that makes the plotting of heatmaps easier and faster. After plotting the heatmap, I got the plotted figure and saved it as ‘Seaborn\_HeatMap.png’.

data\_plot = sns.heatmap(pivot\_table)  *# plotting heatmap*

plot = data\_plot.get\_figure()  *# getting plotted figure for saving*

plot.savefig('Seaborn\_HeatMap.png')  *# saving plot*