**BANJOKO OLASUMBO {Assignment3}**

For this exploratory data analysis using visualization, we import packages

such as pandas, numpy, seaborn, and matplotlib.

using read\_csv and url we import data set from gapminder:

df.describe() and .head() gives a brief view of how our data looks like and its summary statistics

df.column allows me to access the names of the column in our data frame.

using this code below \s{\d}:

# pivot table creation

table = df.pivot\_table(index=['country'], columns=['year'], values=['lifeExp'])

table.plot()

print(table)

we create a pivot table consisting of:

lifeExp ...

year 1952 1957 ... 2002 2007

country ...

Afghanistan 28.801 30.33200 ... 42.129 43.828

Albania 55.230 59.28000 ... 75.651 76.423

Algeria 43.077 45.68500 ... 70.994 72.301

Angola 30.015 31.99900 ... 41.003 42.731

Argentina 62.485 64.39900 ... 74.340 75.320

Australia 69.120 70.33000 ... 80.370 81.235

Austria 66.800 67.48000 ... 78.980 79.829

(Life expectancy, year, and country)

{Plots with seaborn}

The main plot consists of country by year grouping of life expectancy in anonymous

color of red and purple, with purple denoting lower life expectancy across countries over the year and red denoting high life expectancy, vis a vis.

**Ploting seaborn and matplotlib**

Using matplotlib.pyplot package, to combine with seaborn, I declare title,

font-size, and position axis to accommodate more features into seaborn command for visualization.

Then console.log the command below:

sns.heatmap(table, annot=True,fmt="",cmap='RdYlGn',linewidths=0.30,ax=ax)

plt.show()

The annot implies text will be written on the cell of the table, if labels exist for them

fmt changes the axis values to empty in this case column axis, if it has "f" it changes to floats

cmap changes our heat map to Red as Rd, Yellow as Yl and Green as Gn.

linewidth gives the thickness of line to be 0.30, ax to be axis specified through labeling.

show plot is from pyplot to show the figure = plt.show()

Alternatively {sectioning}

I subset the required variables for analysis into one data-set called hmp1,

using [[]] double sub-setting bracket to select the columns of interest.

hmp1 = df[['country','year','lifeExp']]

#pivoting

from the dataframe(df), I create a pivot\_table that assigns rows/index as country and values as lifeExpec.

Subsequently, I view the first 20 element of my data-frame and specify the figure size to plot using this cmd below:

# pivot with various variables

hmp2 = pd.pivot\_table(hmp1,values='lifeExp',index=['country'],columns='year')

print(hmp2.head(20))

plt.figure(figsize=(8,12))

sns.heatmap(hmp2, cmap="RdBu").get\_figure().savefig('hmp2.png')

links to visualisation below:

<https://python3-5.trinket.io/python3-generated/9u0oogvi/trinket_plot.png>

<https://python3-5.trinket.io/python3-generated/k4v7k3xr/trinket_plot.png>

<https://python3-5.trinket.io/python3-generated/dmbwqaor/heatmap.png>

<https://python3-5.trinket.io/python3-generated/pptoi529/hmp2.png>

<https://python3-5.trinket.io/python3-generated/w4620529/hmp2.png>

