APPLIED DATA SCIENCE1 ASSIGNMENT REPORT

NAME: **DIVYA MURALEEDHARAN** STUDENT ID: **22030331**

Here I have used three charts for visualization: line chart, bar graph and pie chart. I have examined three sets of data extracted from the World Bank data. The recorded data is filtered and uploaded to OneDrive. Here is the link for the folder containing the datasets:

Assignment1

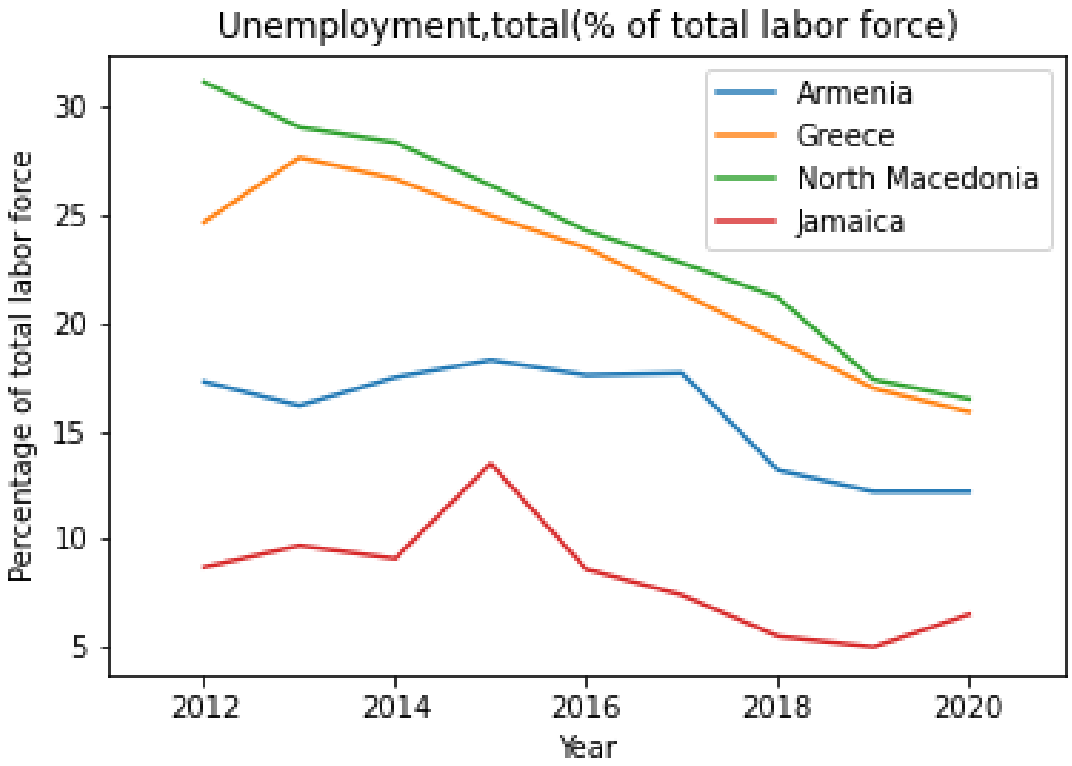
Data source link from world bank is: https://data.worldbank.org/

The dataset selected for the line chart is the overall unemployment rate for the four selected countries. The dataset I chose for the bar chart is the mortality rate for five countries in 2018. The dataset for the pie chart is the prevalence of tuberculosis (per 100,000 population) in 2020.

The GitHub link of my repository that contains python code is given below:

https://github.com/divyamuraleedharan12/divyaProjects

ANALYSIS OF UNEMPLOYMENT, TOTAL (% OF TOTAL LABOR FORCE)

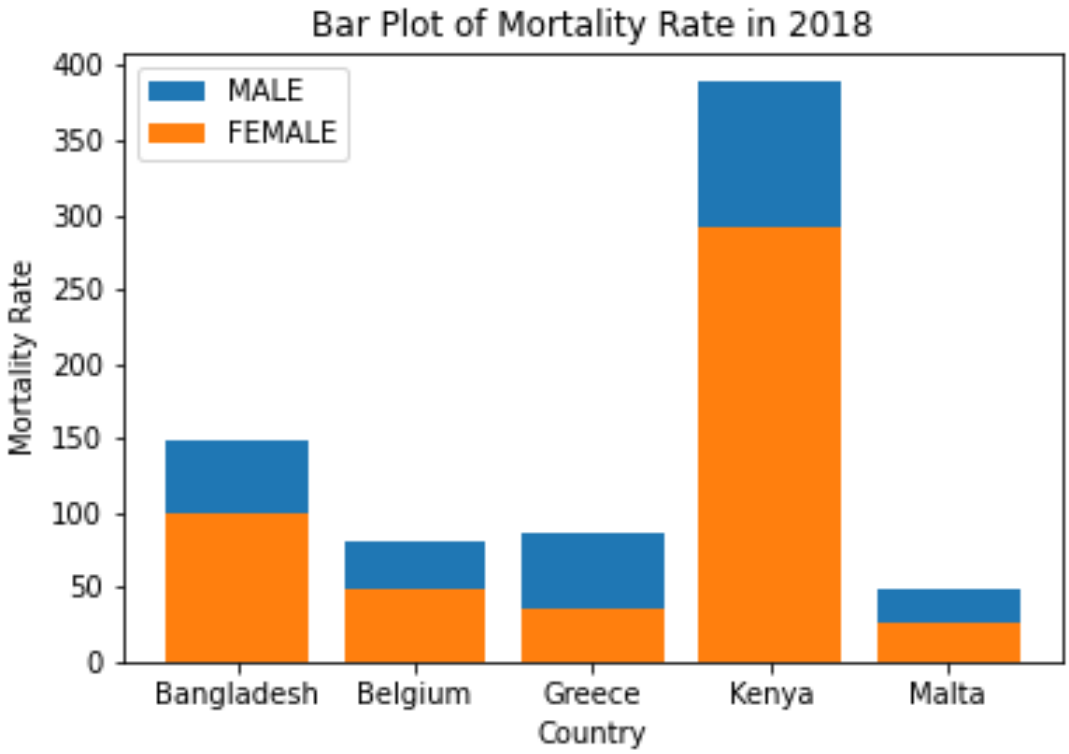


Line plots are used to show trends in data over time periods. It is useful for presenting numerical data and where the data points have a natural order.

The line plot above illustrates the percentage of the total workforce in four different countries from 2012 to 2020. It was observed that North Macedonia initially had the highest percentage compared to other countries. Jamaica and Armenia show an upward trend in the proportion of unemployed. The unemployment rate fell in subsequent years up to 2020. Jamaica had an unemployment rate above 5% in 2012 and 2014, but peaked at 15% between 2014 and 2016, after which the rate fell to around 5% in 2020. While the unemployment rate was between 15% and 20% from 2012 to 2018, it has fallen to less than a fifth in 2020. Although the labor force was 25% and 30% respectively in 2012. It has reached almost a quarter in 2020, in Greece and North Macedonia at the same time.

I choose a line plot so I can see the differences between the highest and lowest unemployment rates in four different countries.

ANALYSIS OF MORTALITY RATE OF ADULTS IN 2018



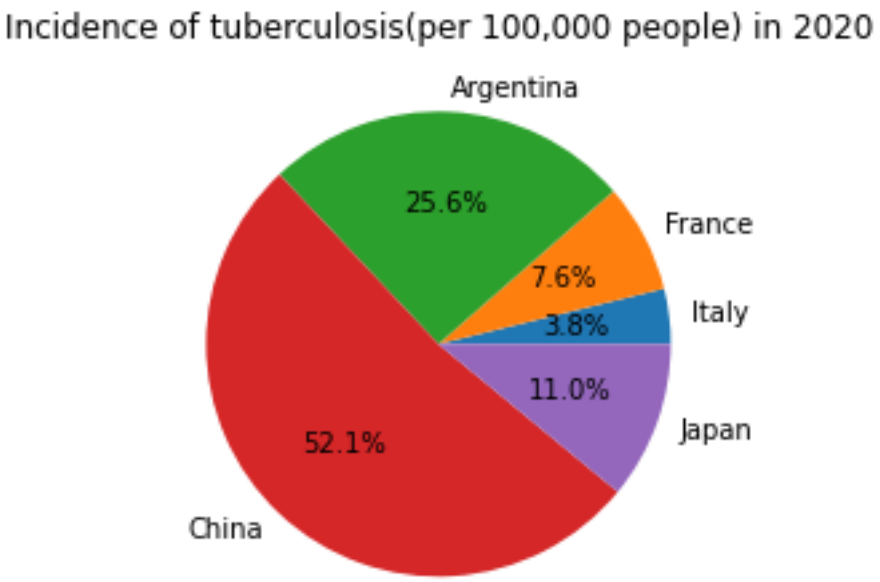
Bar graph are used to display segments of information. It is easy to understand, widely used, and capable of showing changes over time.

The provided bar graph depicts the death rates for males and females in five different regions in 2018. From this graph we can conclude that Kenya had the highest mortality rate for men and women around 400, while Malta had the lowest death rate and others were below 150. The death rate in Bangladesh is about 150 for both boys and girls, but the figure drops to less than 100 in Belgium and Greece, and Malta has a male-female death rate of about 50, while Kenya

has the highest number of mortality rate in Bangladesh has in 2018, there are 300 women and another 100 men.

I choose this bar graph because it is easy to interpret death rates for males and females in different regions.

ANALYSIS OF INCIDENCE OF TUBERCULOSIS (PER 100,000 PEOPLE)



Pie charts can be used to display population percentages and percentages at specific time points.

The given pie chart represents the tuberculosis in the five different countries mentioned above in 2020. It can be seen that half of the population in China had tuberculosis whilst France, Italy and Japan has less that one-fifth of the tuberculosis case rate in China in year 2020. In Argentina, the incidence of tuberculosis is one in four per 100,000 people. In 2020, France, Italy and Japan are 7.6%, 3.8% and 11% respectively. In Argentina, a quarter of the population suffers from tuberculosis, while China accounts for half of the tuberculosis cases, that is 52.1% in 2020.

I chose this pie chart; it is easy to see between different countries at the same time without getting confused.