UNIVERSITY OF HERTFORDSHIRE

**DHANA SRIVALLI GOLUKONDA**

(21087223)

Github Link: <https://github.com/DhanaSrivalli/Applied-DataScience---Visualization-assignment.git>

7PAM2000 APPLIED DATA SCIENCE 1

Assignment 1: Visualization

Professor: **Dr.Ralf Napiwotzki**

Hatfield - 2023

**TABLE OF CONTENTS:**

**Page No.**

Introduction…………………………………………………………………………………. 03

Description………………………………………………………………………………….. 03

Interpretations……………………………………………………………………………….. 03

* 1. Line Graph…………………………………………………………………………... 04
  2. Bar Chart...…………………………………………………………………………. 05
  3. Scatter Plot...………………………………………………………………………... 05

Conclusion…………………………………………………………………………………... 06

List of References…..………………...…………………………………………………...… 07

Appendix……………………………………………………………………………………. 08

**INTRODUCTION:**

In this developing world, countries set goals to maintain a high level of standards for their citizens mostly by providing good food and a healthy lifestyle, reducing unemployment by recruiting new employees and, refilling the vacancies left over, focusing on the growth of literacy and education. The major problem to achieving all these goals is ‘Population Growth’.

The growing world population is one of the biggest problems the world is currently facing, and it will become uncontrollable in the coming future if appropriate measures are not taken immediately to improve it. The world population has recorded the fastest growth in the 20th century. In the 1950s the world population was 2.7 billion people, and by the end of this year, it will exceed 8 billion.

The rapid growth of the population is a burning issue for the upcoming generations. According to the major resources like the United Nations and U.S. Census Bureau published about the ‘growth of population in different territories in the world.’ People are facing issues like food shortages, increasing rates of economic development, illiteracy, and its leads the way to the collapse of the economic growth rate, facing crises issue and burdens for middle-class families to survive.

Most countries are planning to give extra schemes and policies for the family members who controlled their family not more than two kids.

**DESCRIPTION:**

In this report, we will show the growth rate of population for the countries of the United Kingdom, the United States, and Russia. The data is extracted and converting the data into visual presentations like graphs, pie charts, bar graphs, scatter plots, etc.,

Python programming is the best way to solve the problem and find the errors in the data. We use python in various fields to maintain the data up to date. Throughout the world, most companies have widely used this program to keep data safe. It can be used in planning designing and calculation.

**INTERPRETATIONS:**

This report analyses the growth of the population in three various countries like the United States, the United Kingdom, and Russia. The Dataset for this report is taken from Kaggle among the data provided I took three countries that have high growth of populations in the United States and Russia with an average growth of population and the United Kingdom which stays in a safe zone and succeeded in controlling the growth of population.

Nowadays python programming plays a key role in analyzing data from different companies and different projects. Python is a high-level programming language that can easily solve different classes of problems. It has a huge, big library that provides tools that suits various tasks. Mainly we use Python’s scientific Computing libraries like Pandas, Numpy, Scipy, and Matplotlib.

A line graph is the best choice to show the exact growth rate and the difference between the countries. is easily shown in the graph whether it is an increasing or decreasing population in this line graph I placed three different colors red is the United Kingdom and Blue is for the United States and green is for Russia. I extract the data from the world\_population\_dataframe and used a plot for the x and y-axis.

The Bar chart which I took is respective to the age of children less than one year older in different countries of the US, the UK, and Russia. Using a plot bar chart in python differentiating the countries in three different colors and plotting them according to years in the x and y-axis.

Using a scatter plot for the third graph know the visual representation of three different variables and can see the overlapping values. The command used for the scatter plot is subplots.

**a) Line Graph: -**

The below line Graph shows the growth of population in three different countries of the United Kingdom, the United States, and Russia. Here I took the x\_axis as the Year from 1950 to 2020 as per available Data sets. On the other hand, the y\_axis describes the level of population in billion from 0.5 billion to 3.5 billion.

As per the line graph, the United States has huge population growth. In the year 1950 United States stays a 1.5 billion population and rapidly increased and reached a high level of nearly 3.5 billion in the year 2020. Russia maintains an average growth of 1.0 billion to 1.5 billion in the year 1950 to 2020. Similarly, the United Kingdom plays a key role in controlling the growth of the population with 0.5 billion to 0.7 billion approximately.

Chart, line chart

Description automatically generated

Figure .1

**b) Bar Chart: -**

The following bar chart illustrates the increase in the population of children under the age of one in the year 1950 to 2020 for the countries of the United States, the United Kingdom, and Russia. In this chart, you can see the children are divided into three different groups as per their nationality. For the united states, the chart shows yellow in color, Russia in black, and the united kingdom in red color.

As per the chart, the population of children under 1 year of age in the United States and Russia boomed after 1950 while in the United Kingdom, it started to decrease. In the United States, the population hit its peak (4 billion) in the year 1960 and it started to gradually reduce, only to start increasing again in the year 1975. The Russian population has both highs and lows but overall we could see a gradual decrease in the population from 1950 to 2020. The Population in the United Kingdom did not undergo a massive change wherein the average remained around 0.8-1 billion every ten years, unlike the other two countries.

Chart, histogram

Description automatically generated

Figure .2

**c) Scatter Plot: -**

The scatter plot shows the population for the aged people between 90 and 95 years old in the years 1950 to 2020. This US is represented with yellow which has a high survival rate compared to other countries like UK and Russia.

In the below Scatter plot, the graph is taken respectively to age. The below chart shows the population distribution across three countries, the United States, the United Kingdom, and Russia. The x-axis denotes the year from 1950 to 2020 and the y-axis denotes the population measured under 90 to 99 years of age in billions.

Chart, line chart

Description automatically generated

Figure .3

**CONCLUSION:**

Overall, the United States stays at the top in increasing the growth of population and Russia maintains a standard growth in increasing population. Meanwhile, the United Kingdom has succeeded in controlling population growth keep the safe as usual from starting the year 1950 to the end year 2020.

To Conclude that Government must take certain measures to control the population and bring awareness among the people especially those who are illiterate. Conducting a weekly event on this topic helps to look forward to family planning.

**LIST OF REFERENCES:**

* [Introduction | The Growth of World Population: Analysis of the Problems and Recommendations for Research and Training |The National Academies Press](https://nap.nationalacademies.org/read/9543/chapter/3)

<https://nap.nationalacademies.org/read/9543/chapter/>3

* [World-population2023 | Kaggle](https://www.kaggle.com/datasets/dinarkhan/worldpopulation2023?resource=download)

[**https://www.kaggle.com/datasets/dinarkhan/worldpopulation2023?resource=download**](https://www.kaggle.com/datasets/dinarkhan/worldpopulation2023?resource=download)

* [World population to reach 8 billion this year, as the growth rate slows | UN News](https://news.un.org/en/story/2022/07/1122272)

[**https://news.un.org/en/story/2022/07/1122272**](https://news.un.org/en/story/2022/07/1122272)

* [Python Data Science Handbook | Python Data Science Handbook (jakevdp.github.io)](https://jakevdp.github.io/PythonDataScienceHandbook/)

[**https://jakevdp.github.io/PythonDataScienceHandbook/**](https://jakevdp.github.io/PythonDataScienceHandbook/)

* National Academies of Sciences, Engineering, and Medicine. 1963. The Growth of World Population: Analysis of the Problems and Recommendations for Research and Training. Washington, DC: The National Academies Press. https://doi.org/10.17226/9543.

**APPENDIX:**

#!/usr/bin/env python

# coding: utf-8

# In[1]:

#Importing Libraries

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

# In[2]:

#Loading data

world\_population\_dataframe = pd.read\_csv("D:\\Datasets\\Book1.csv")

# In[3]:

#DataFrame

world\_population\_dataframe

# In[4]:

#Grouping Individual country

uk\_grouped = world\_population\_dataframe.groupby(world\_population\_dataframe['Country name'])

uk\_data = uk\_grouped.get\_group("United Kingdom")

uk\_data.head(100)

us\_grouped = world\_population\_dataframe.groupby(world\_population\_dataframe['Country name'])

us\_data = uk\_grouped.get\_group("United States")

us\_data.head(100)

ru\_grouped = world\_population\_dataframe.groupby(world\_population\_dataframe['Country name'])

ru\_data = uk\_grouped.get\_group("Russia")

ru\_data.head(100)

# In[7]:

# Generating Line plot

plt.plot(uk\_data['Year'], uk\_data['Population'] , color='red', linewidth=1.0, label="United Kingdom")

plt.plot(us\_data['Year'], us\_data['Population'] , color='blue', linewidth=1.0, label="United States")

plt.plot(ru\_data['Year'], ru\_data['Population'] , color='green', linewidth=1.0, label="Russia")

plt.title('Population')

plt.xlabel('Year')

plt.ylabel('Population')

plt.legend()

plt.show()

# In[8]:

#Importing Libraries

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

# In[9]:

#Loading data

world\_population\_dataframe = pd.read\_csv("D:\\Datasets\\Book1.csv")

#DataFrame

world\_population\_dataframe

#Grouping Individual country

uk\_grouped = world\_population\_dataframe.groupby(world\_population\_dataframe['Country name'])

uk\_data = uk\_grouped.get\_group("United Kingdom")

uk\_data.head(100)

us\_grouped = world\_population\_dataframe.groupby(world\_population\_dataframe['Country name'])

us\_data = uk\_grouped.get\_group("United States")

us\_data.head(100)

ru\_grouped = world\_population\_dataframe.groupby(world\_population\_dataframe['Country name'])

ru\_data = uk\_grouped.get\_group("Russia")

ru\_data.head(100)

# In[13]:

#Generating Bar plot

plt.figure()

plt.bar(us\_data['Year'], us\_data['Population of children under the age of 1'] , color='yellow', width=0.7, label="United States")

plt.bar(ru\_data['Year'], ru\_data['Population of children under the age of 1'] , color='black', width=0.7, label="Russia")

plt.bar(uk\_data['Year'], uk\_data['Population of children under the age of 1'] , color='red', width=0.7, label="United Kingdom")

plt.xlabel("Year") #x-axis

plt.ylabel("under the age of 1") #y-axis

plt.title("Population of children under the age of 1 based on country wise and year wise") #Title

plt.legend()

plt.show()

# In[16]:

#Generating Scatter plot

fig = plt.subplots(figsize =(8, 5))

plt.scatter(us\_data['Year'], us\_data['Population aged 90 to 99 years'] , color='Yellow', label="United States")

plt.scatter(ru\_data['Year'], ru\_data['Population aged 90 to 99 years'] , color='Blue', label="Russia")

plt.scatter(uk\_data['Year'], uk\_data['Population aged 90 to 99 years'] , color='red', label="United Kingdom")

plt.xlabel("Year") #x-axis

plt.ylabel("Population aged 90 to 99 years") #y-axis

plt.title("Population aged 90 to 99 years") #Title

plt.legend()

plt.show()