Name: Bhavesh Kewalramani

Roll No.: A-25

Section: A
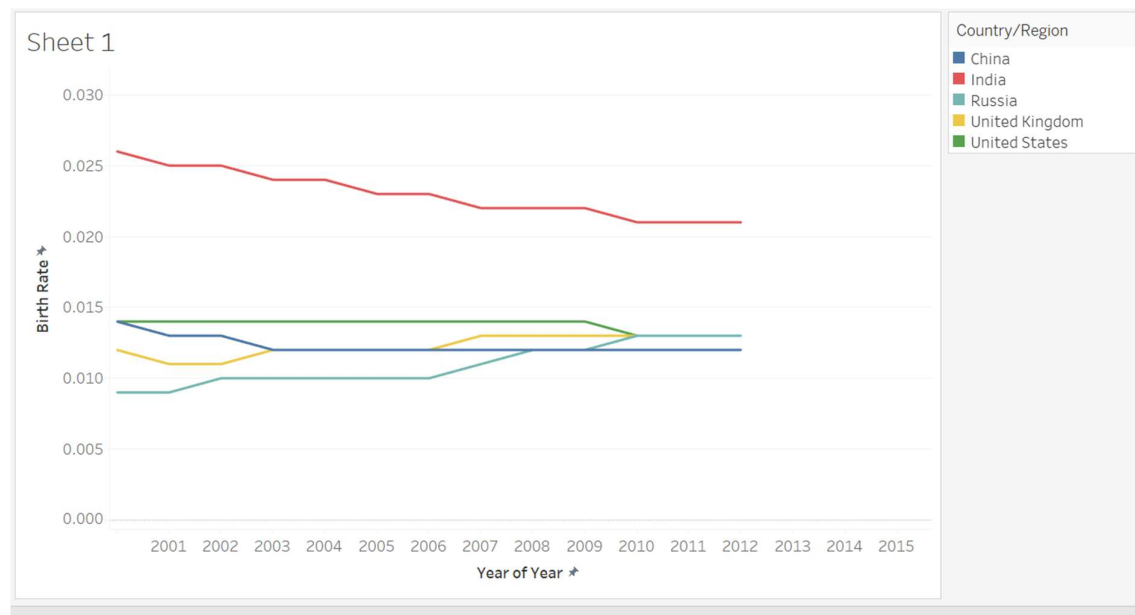
Semester: VII

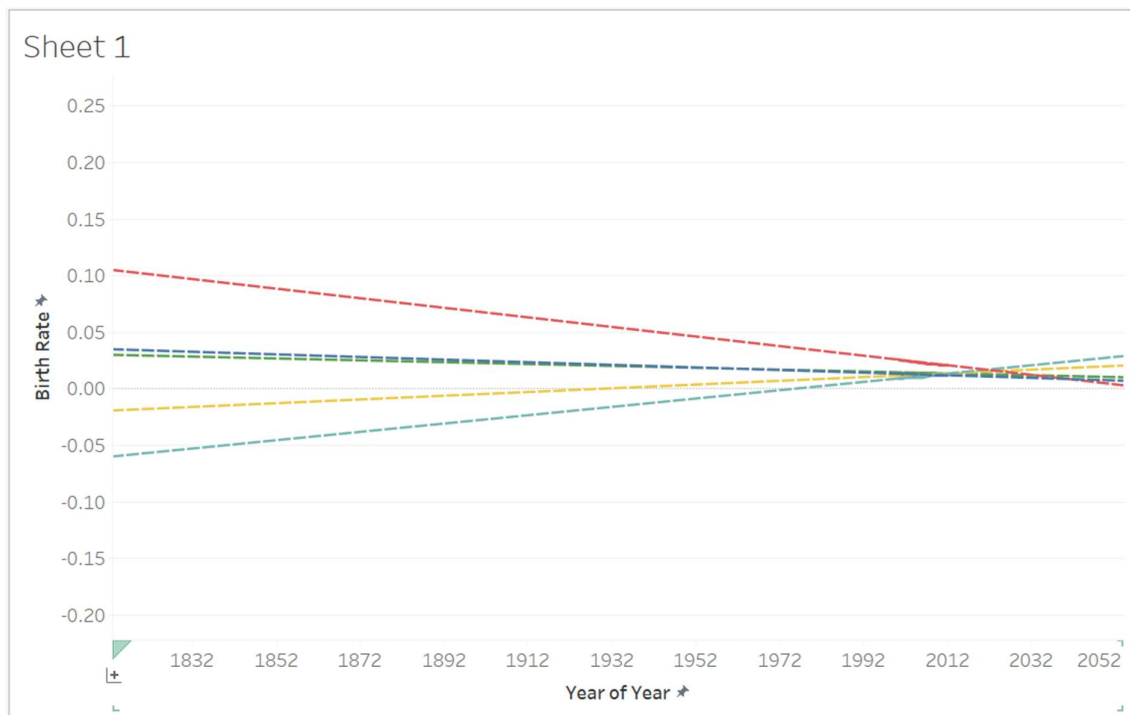**Shift:** I

Batch: A1

PART(A):

Aim: Exploratory data analysis using tableau on population dataset

- Use the same population dataset as that in workshop
- Explain the attributes
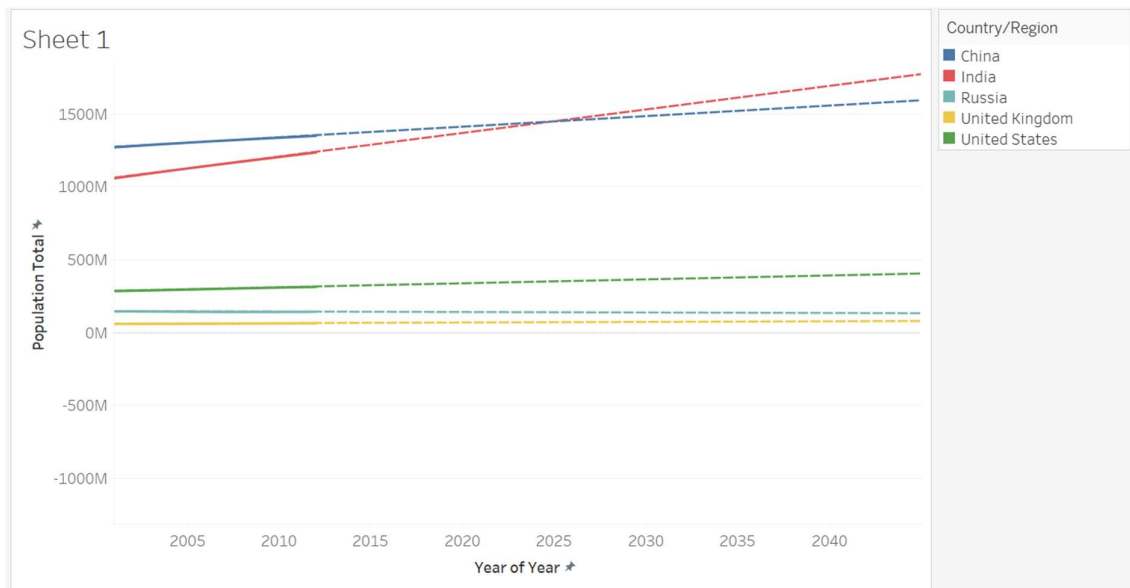- Write the analysis

In Population dataset, we have 8 attributes. These attributes are Birth rate, Population (0-14), Population (15-64), Population (65+), population total, population urban and year and these are ratio attributes. The other one is Country which is nominal attribute. There are records of 30 years for each month. There are around 207 countries in it.

The above plot shows the birth rate in five countries namely China, India, Russia, United Kingdom and United States. We can see how birth date decreased in India till 2013 and in other four it became constant till then.

Sheet 1

Birth Rate ✶

0.25
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.20

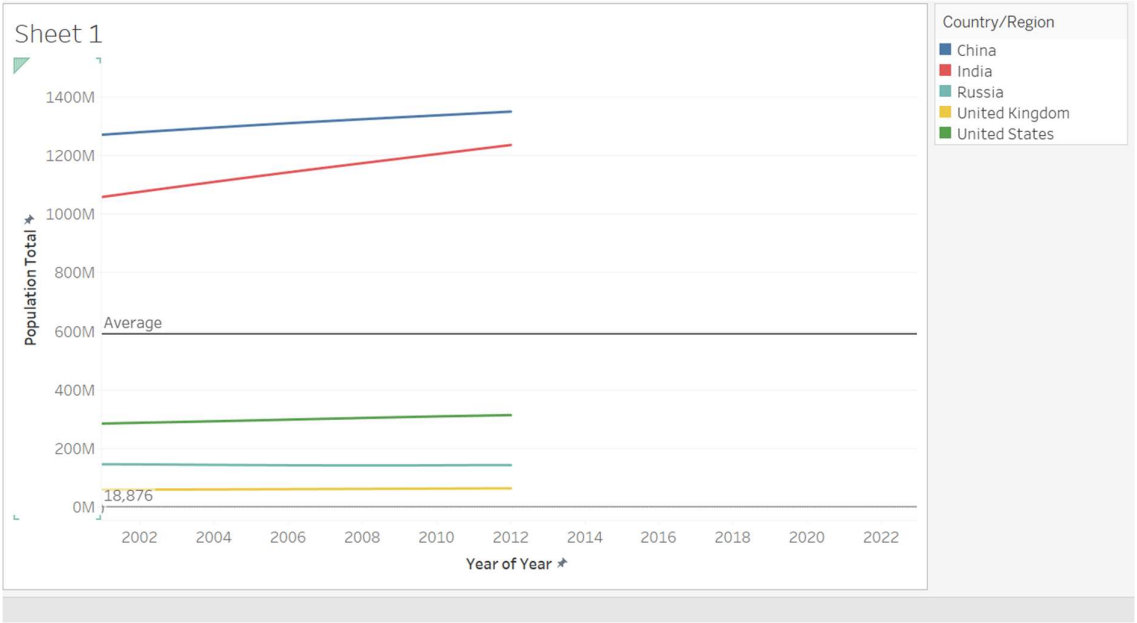1832 1852 1872 1892 1912 1932 1952 1972 1992 2012 2032 2052

Year of Year ✶

Using the Trends we can see that around 2032 the birth rate might become same for the five countries. The birth rate of India will still go on decreasing in future

Sheet 1

Country/Region
■ China
■ India
■ Russia
■ United Kingdom
■ United States

Population Total ✶

1500M
1000M
500M
0M
-500M
-1000M

2005 2010 2015 2020 2025 2030 2035 2040
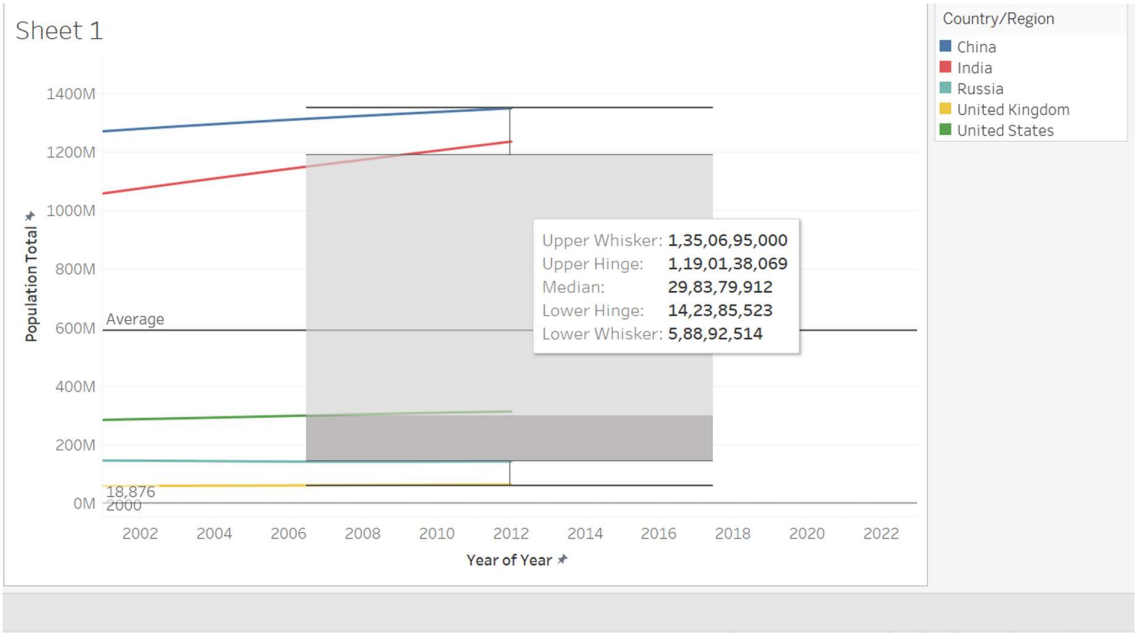
Year of Year ✶

The above trend is plotted for population for the five countries namely China, India, Russia, United Kingdom and United States. From this we can see that till 2025, Population of India will become
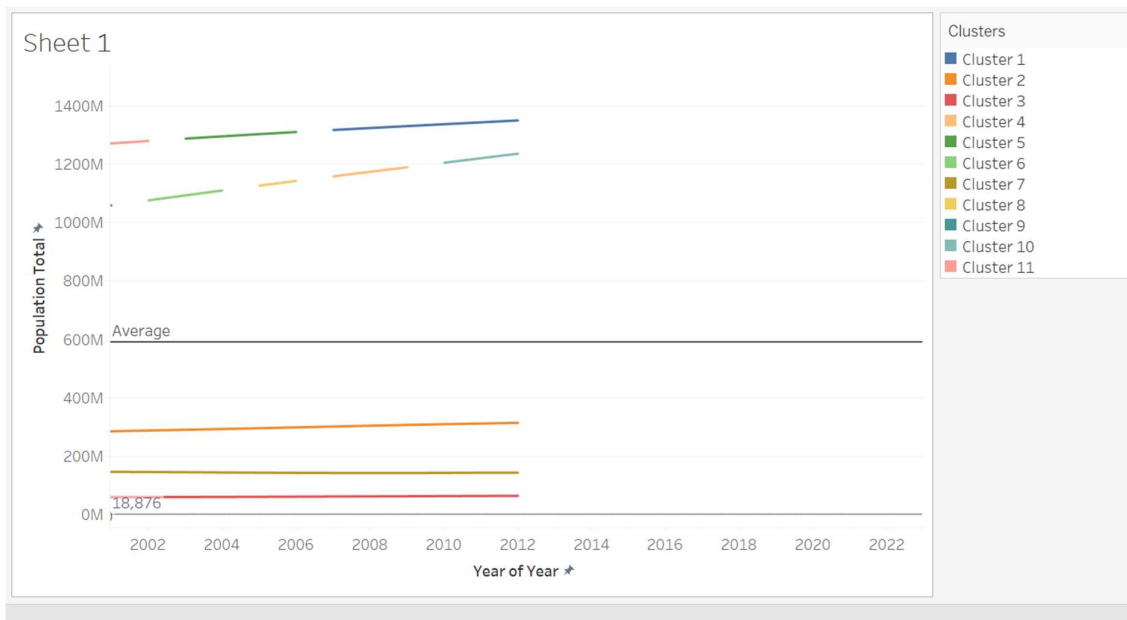
same and then population of India from there will go on increasing and that of China will go on decreasing. We can also see that population if United Kingdom and Russia will almost become constant after 2040.
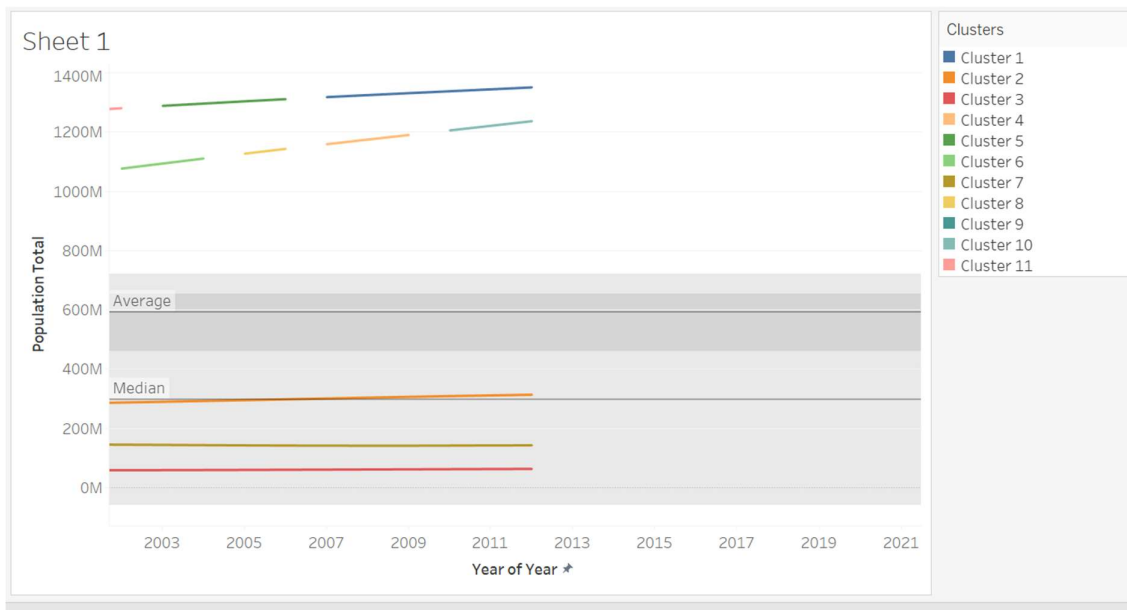

We can see the average population of all these five countries is 60 M.

The boxplot is plotted for all these countries and the median of these is 29 M

When we cluster them, total 11 clusters are formed

The graph shows the average and median of clusters with 95% CLI

From the above graph we can study the trends and see how the population is affected by different factors and how these lie in different clusters based on age and other factors.