

Information Visualization

Harjinder Singh¹

¹ Assistant Professor Department of Economics G. G.N. Khalsa College, Ludhiana

Abstract: Data today is used in very innovative way. Visualizations of data can be understood by all and presentation of data in the form of visualization can help policy makers to concentrate their efforts to the centre of problem rather than spending time in puzzles of data. Visualization gives a touch of storytelling to the data. One begins by simple visualization and over the period of time numerous visualizations and plots give many different ideas to researchers. The most well-known case of information visualization is Google maps. These are used by almost everyone who possess smart phone today. Fraud detection is another area where use of visualizations is made to detect and highlight the pattern of fraud. Textual and sentimental analysis of data can help to establish pattern in messages which are used in social media today. The analysis of tweets of many famous personalities can give insight into their personalities and their hidden agendas. It is now possible to visualize whole literature into beautiful plots and graphs to analyse them. Basic purpose of this paper is to highlight the point that visualization of data helps in bringing out hidden message in data. Human brain is more perceptible to images than plain data. Visualizations help us in getting more clear idea about our objective. India is already on the verge of becoming world's most populated country by year 2024. The poverty ratio in India is 21% according to 2011 data. This makes for huge number of poor people in India. It will put huge burden on already over stretched Indian resources. Indian cannot wait for rise in income of people for lowering of fertility rate as it will take lots of time. Many countries in spite of having low income have seen tremendous success in lowering of fertility rate. Here, our objective is to find out the countries having very low Fertility Rate even though they have low Gross National Income per capita, so that we can replicate their success in case of lowering Fertility Rate in India. Three variables Fertility Rate, GNI and Region are used to make visualizations. In processing of data "R" an open source software environment for statistical computing and graphics is used. The advantage of using "R" is that it is free and various types of data processing is easily done. There are thousands of statistical and visualizations package are available. Here, I mainly used package "ggplot2" to visualize hidden message pertaining to problem of this paper. An attempt has been made to look at numbers as visualizations to highlight the given objective in the beginning. Visualizations have helped in finding the countries which are having success in controlling. By examining each country central to the objective of this paper many interesting observations have been made. It is not necessary to wait for rise in income as factor to lower fertility rate. 12 countries are having success with lowering of fertility rate and many countries in this group did not wait for rise in income as factor to drive down fertility rate. Good policies and effective implementation brought them success. In some cases, the social structure and religion is not an impediment to the achievement of the objective. Bangladesh, with concentrated efforts both at social and economic level brought down the rising fertility rate. Bangladesh cleverly used women as tool to convince other women to adopt family planning. The economic empowerment of women of Bangladesh through micro credit schemes make them realized the benefits of small family. NGO's in Bangladesh are used very effectively to contribute to these efforts of lowering of fertility rate.

Keywords: Information, Visualization, R, ggplot2, Fertility Rate, Data

I. INTRODUCTION

Big data is most commonly used word in today's age of information explosion. Huge volume of data is generated which outstrips the pace at which this data can be processed. According to conservative estimates 2.5 quintillions of data is generated every day. Every digital device is contributing to this data whether it is computer, cell phone, smart watches, bar code readers or any other device connected to the internet. This massive volume of data creates problem of data noise or data pollution. To get meaningful insight into volume of data we have to eliminate or reduce data noise. Human brain processes and understand information presented through visualizations than through tables, texts or simple statistical data. This ability of human brain to easily understand and filter information through visualizations is demonstrated in this paper.

The early modern history of information visualization starts with William Playfair, an early innovator in statistical graphics. In 1786, he published the commercial and political atlas, which displayed many bar charts, line graphs and histograms representing the economy in England. He followed this up with the first area chart and first pie chart in 1801. In 1857, English nurse Florence Nightingale used information graphics primarily the coxcomb chart, a combination of stacked bar and pie chart to change history and persuade Queen Victoria to improve conditions in military hospitals. Her chart showed the number and causes of deaths during

each month of the Crimean war: preventable diseases in blue, wounds in red, and other causes in black. Charles Joseph Minard (1721-1870) was a french civil engineer produced a chart that tells the story of “Napoleon’s Russian campaign of Russia in 1812”. It was drawn half century afterwards by Charles Minard. Minard’s chart shows six types of information: geography, time, temperature, the course and direction of the army’s movement, and the number of troops remaining. The chart tells the dreadful story with painful clarity, in 1812, the grand army set out from Poland with a force of 422,000, only 100,000 reached Moscow and only 10,000 returned. John snow (15 March 1813 – 16 June 1858) was an english physician used who statistics to illustrate the connection between the quality of the water source and cholera cases which he visualized with of cholera outbreak of 1854. Edward Rolf Tufte, born 1942 (Tufte, er: the visual display of quantitative information, graphics press, 1983) is excellent book on statistical graphics, charts, tables. Theory and practice in the design of data graphics, 250 illustrations of the best (and a few of the worst) statistical graphics, with detailed analysis of how to display data for precise, effective, quick analysis. Design of the high-resolution displays, small multiples. Editing and improving graphics. The data-ink ratio. Time-series, relational graphics, data maps, multivariate designs. Hans gosling (1948 – 2017) was professor of international health at Karolinska Institute, founder and chairman of Gampinder Foundation highlighted many developmental issues with the use of beautiful visualizations.

The basic purpose of this paper is to highlight the point that visualization of data helps in bringing out hidden message in data. Human brain is more perceptible to images than plain data. Visualizations helps us in getting more clear idea about our objective. India is already on the verge of becoming world’s most populated country by year 2024. The poverty ratio in India is 21% according to 2011 data. This makes for huge number of poor people in India. It will put huge burden on already over stretched Indian resources. Indian cannot wait for rise in income of people for lowering of fertility rate as it will take lots of time. Many countries inspite of having low income have seen tremendous success in lowering of fertility rate. Here, our objective is to find out the countries having very low fertility rate even though they have low gross national income per capita, so that we can replicate their success in case of lowering fertility rate in India.

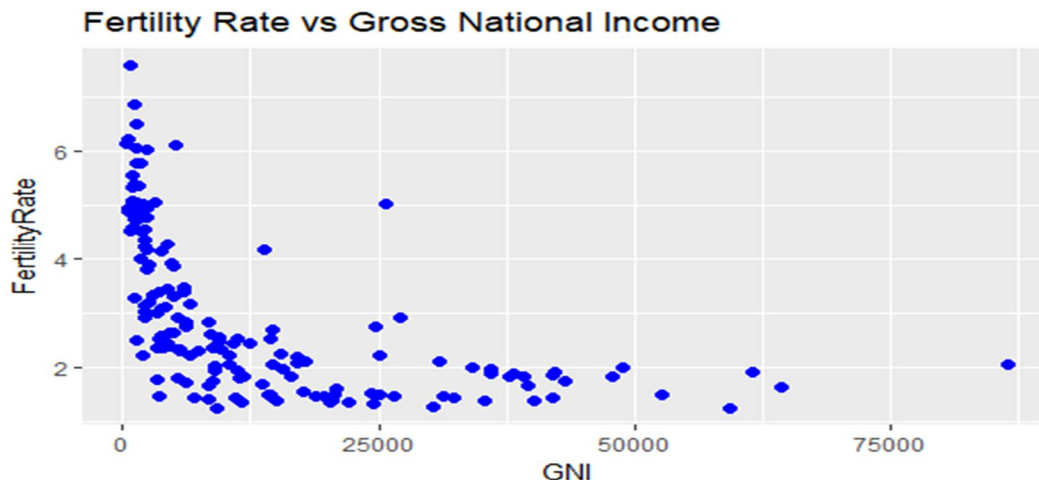
II. RESEARCH METHODOLOGY

This paper uses data from World Health Organization. There are thirteen variables from 194 countries in the dataset like Country, Region, Population, Under15, Over60, Fertility Rate, Life Expectancy, Child Mortality, Cellular Subscribers, Literacy rate, Gross National Income per capita, Primary School Enrolment Male and Primary School Enrolment Female. Out of these thirteen variables, three variables Fertility Rate, GNI and Region are used to make visualizations.

In processing of data “R” an open source software environment for statistical computing and graphics is used. The advantage of using “R” is that it is free and various types of data processing is easily done. There are thousands of statistical and visualizations package are available. Here, we mainly used package “ggplot2” to visualize hidden message pertaining to problem of this paper. Several statistical tools used to bring out maximum and minimum values. For measure of central tendency, we used median as it is not affected by outliers and extreme values.

III. RESULTS AND DISCUSSION

By reading WHO data file which is in .csv format, we first make scatter plot with two variables “Fertility Rate and GNI “we get this plot created with ggplot2.



Each dot in plot represents one of the 194 countries with respect to fertility rate and GNI per capita. There is negative relation between fertility rate and GNI per capita. Countries having high GNI have low fertility rate. This establishes some degree of relation between fertility rate and GNI. There are some countries with very high fertility rate and some with very high GNI per capita. In the next table we will see some basic statistics about fertility rate and Gross National Income per capita.

The Table 1 shows that country having maximum Fertility Rate of 7.58 and GNI per capita of \$720 is Niger. Bosnia and Herzegovina has lowest Fertility Rate with GNI per capita \$9190. Qatar has maximum GNI per capita of \$ 86440 and Fertility Rate of 2.06. Congo is having minimum GNI per capita of \$ 340 with Fertility Rate of 6.15. India has Fertility Rate of 2.53 and GNI per capita of \$3590.

TABLE 1 BASIC DATA ABOUT FERTILITY RATE AND GNI

Country	Fertility Rate	GNI (\$)
Niger	7.58	720
Bosnia & Herzegovina	1.26	9190
Qatar	2.06	86440
Congo	6.15	340
India	2.53	3590

The Table 2 shows Median Fertility Ratio and Median Gross National Income per capita (GNI). The Median Fertility Rate is computed with following code: median (who\$Fertility Rate, na.rm = TRUE) and GNI per capita is computed by median (who\$GNI, na.rm = TRUE).

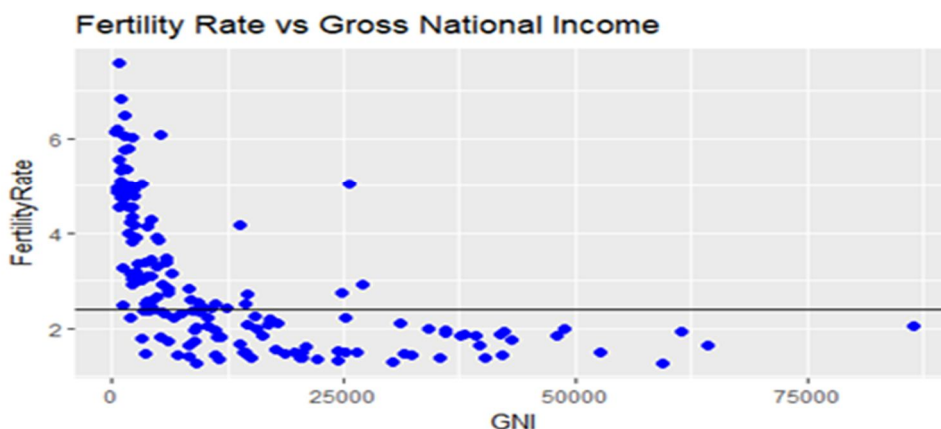
TABLE 2 MEDIAN FERTILITY RATE AND MEDIAN GNI PER CAPITA

Variables	MEDIAN
Fertility Rate	2.4
GNI per capita	\$ 7870

Table shows that median fertility rate is 2.4 and median GNI per capita is \$ 7870 .

The next task is to separate countries by Median Fertility Rate and Median GNI. Countries are first separated by Median Fertility Rate with this code: FR_GNI+geom_hline (yintercept=2.4), with that we get graph below:

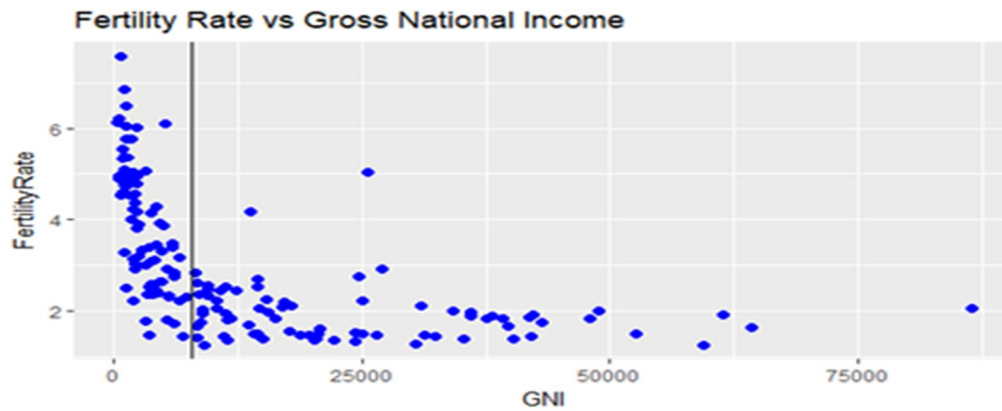
Countries Separated by Median Fertility Rate



The Median Fertility Rate of 2.4 is represented through solid horizontal line running through the graph. It separates countries above and below Median Fertility Rate. In this plot it can be clearly seen that countries above Median Fertility Rate are having low GNI per capita.

In next plot countries are separated by Median GNI per capita with this code: `FR_GNI+geom_vline(xintercept=7870)`, running this code in “R” the plot below is drawn

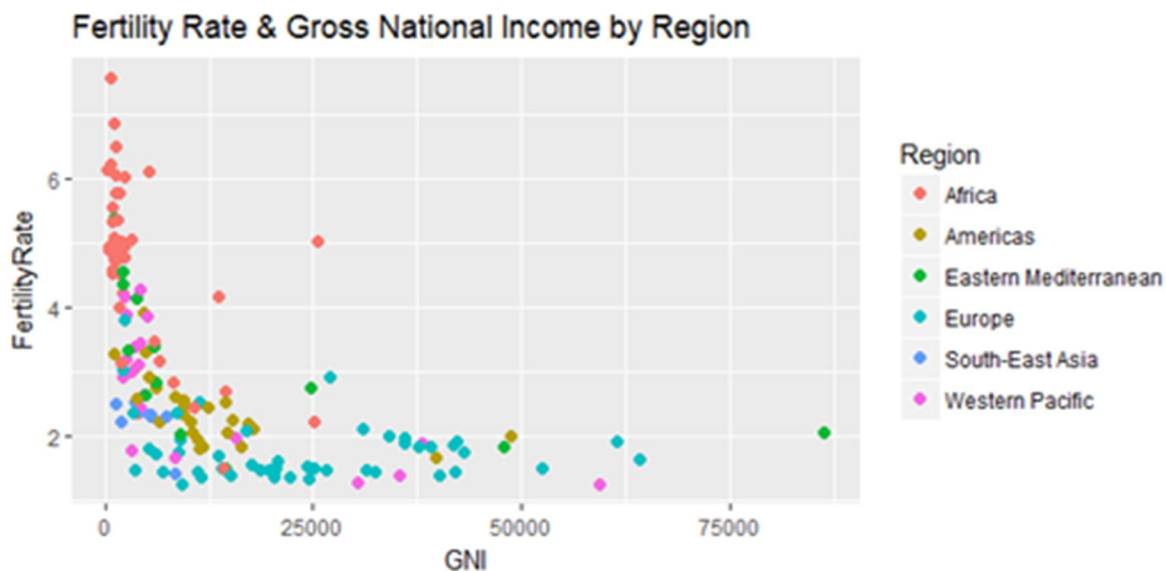
Countries Separated by Median GNI



In plot above the countries are separated by Median GNI per capita of 7870 as vertical solid line running through the plot. This Median vertical line through the plot shows countries having low GNI per capita have high Fertility Rate in general. This clearly establishes some pattern of relationship between Fertility Rate and GNI per capita.

To get Region wise distribution of Fertility Rate and GNI per capita, another variable is introduced at this stage. All countries in data are distributed into six regions such as Africa, Americas, Eastern Mediterranean, Europe, South-East Asia and western Pacific. Through this code: `FR_GNIr= Scatterplot1+geom_point (size=2,shape=19)+ggtitle(“Fertility Rate & Gross National Income by Region”)`.

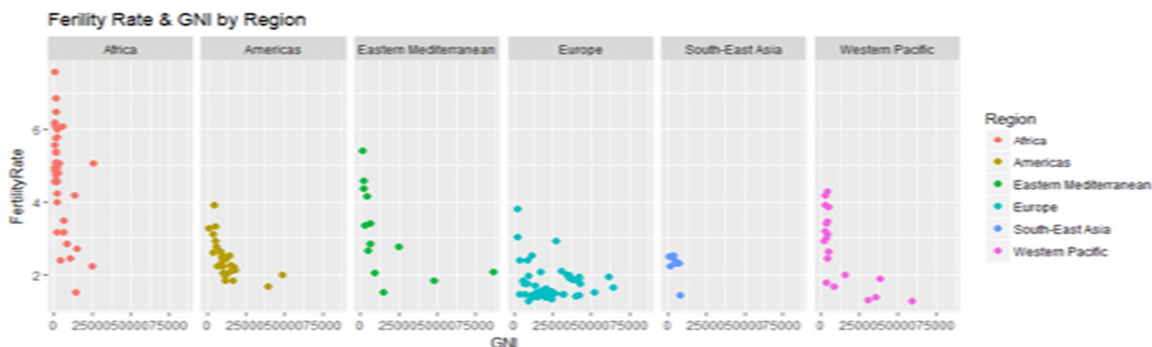
Fertility Rate and GNI Separated by Region



Countries in African region are characterized by very high Fertility Rate and low GNI per capita. Most of countries in European region have low Fertility Rate and High GNI per capita.

Another better way of getting clear view of Fertility Rate and GNI per capita, region-wise is to distribute all the countries in different regions into different regional panels, which is demonstrated through facet plot below.

Fertility Rate and GNI Separated by Region

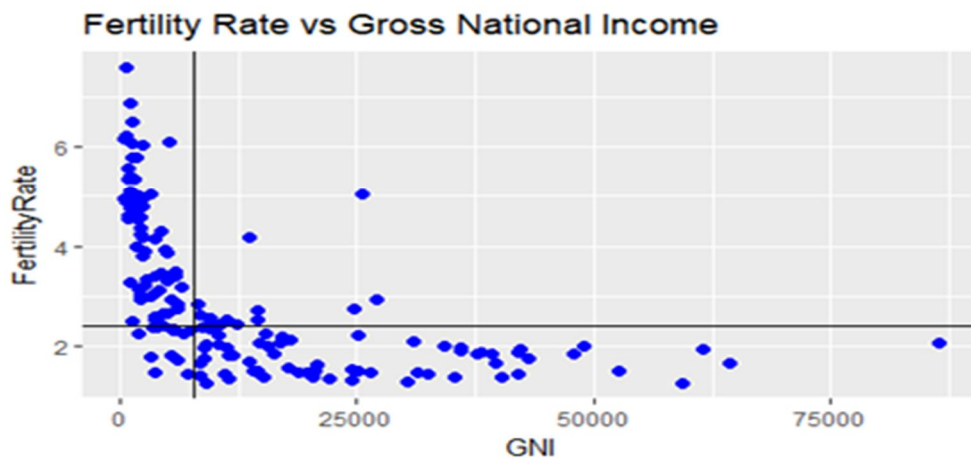


The facet plot above now shows the comparison of different regions with respect to Fertility Rate and GNI per capita. Maximum countries in European Region are reflecting low Fertility Rate, while those in African region are showing high Fertility Rate.

Superimposing, the Median Fertility Rate of 2.4 and Median GNI per capita of \$ 7870 as solid horizontal and vertical line respectively on the scatter plot of Fertility Rate and GNI per capita through code: `FR_GNIr+geom_hline(yintercept=2.4)+geom_vline(xintercept=7870,`

Plot below shows the countries separated by Median Fertility Rate and Median GNI per capita as solid lines running horizontally and vertically through the plot. These solid Median lines separates the countries below Median Fertility Rate and below GNI per capita . Looking at the plot now the hidden pattern of data emerges which is pertaining to our objective of locating countries having low Fertility Rate and Low GNI per capita. In plot above, near the origin of axis we can see the countries encompassed within the intersection of two median lines. These are the countries which are object of interest related to this paper. All these countries in this object of interest are having low Fertility Rate and Low GNI per capita.

Countries separated by Median Fertility rate and GNI



Next, step in the course of this paper is to see how many and which countries are in this object of interest. Running code : `BelowmedC=subset(who, Fertility Rate<2.4 & GNI<7870)` and this code : `BelowmedC[c("Country","GNI","FertilityRate")]` , we get table of countries within object of interest :

TABLE 3

12 COUNTRIES WHOSE FERTILITY RATE IS BELOW MEDIAN OF 2.4 AND GNI PER CAPITA IS BELOW \$ 7870

Sr. No.	Country	GNI (\$ per capita)	Fertility Rate
8	Armenia	6100	1.74
14	Bangladesh	1940	2.24
20	Bhutan	5570	2.32
32	Cape Verde	3950	2.38
55	El Salvador	6640	2.24
65	Georgia	5350	1.82
104	Maldives	7430	2.31
140	Republic of Moldova	3640	1.47
163	Sri Lanka	5520	2.35
182	Ukraine	7040	1.45
188	Uzbekistan	3420	2.38
191	Vietnam	3250	1.79

In Table 3 we get list of countries who are below median Fertility Rate of 2.4 and median GNI \$ 7870. Out of these countries in table Armenia, Georgia, Republic of Moldova, Ukraine and Uzbekistan are parts of erstwhile Soviet Union which had historically low Fertility Rate. The Orthodox Church opposed abortion and thus there were few abortions in Russia until the Revolution. The Bolsheviks instituted an atheism campaign which removed religious leaders from positions of authority. The Soviet Union became the first country to legalize abortion (1920). Soviet women began using abortion as a method of birth control, affecting birth rates. There is also paucity of data from Soviet period. So, in Indian context we cannot replicate or learn from experience of these countries in lowering Fertility Rate.

Cape Verde is a small island country to the coast of Africa in central Atlantic Ocean. The population of Cape Verde is around half million, so Cape Verde's experience and conditions are drastically different from India. It's very small country with very small population.

El Salvador is a country in Central America. It has a very small population of 6.3 million. El Salvador saw great success in controlling Fertility Rate with use of contraceptives from 1978 to 1988 among married women in the age group of 15 – 44 years of age and fertility rate declined from 6.3 to 4.6 within this period. The impressive fall in fertility rate in Salvador was mostly female sterilization led. The 2008 family planning survey documented that female sterilization is still the most common method of contraception in El Salvador, but it also showed a growing use of temporary methods, especially injectable hormones that women must receive every 12 weeks. The percentage relying on inject tables rose from less than 2 percent of users in 1988 to 31 percent in 2008. The reliance on sterilization among young women in El Salvador is striking: Nearly one-fourth of women under age 30 had been sterilized in 2004.

Bangladesh is a typical developing country of South Asia with a large population, low per capita income. The country has dominated agriculture based economy which has moved into the market-based economy in the mid-1970s and received aid under the program of World Bank, IMF, and other donor agencies. The latest 2011 census of Bangladesh estimates a population of 149 million for the country implying an increase of 19 million since the census of 2001. Bangladesh has shown great success in lowering of fertility rate. Bangladesh set an example for countries like India and Pakistan. The Bangladesh family planning program has depended on a large cadre of female outreach workers going door-to-door to provide information, motivate clients, and provide commodities; the program has used mass media to stimulate a change in attitudes about family size. The program both contributed to and benefited from improvements in women's status in Bangladesh during the past 30 years. The contribution of women workers

in conservative Muslim society in Bangladesh is laudable. It broke the barriers of communications to educate women in Bangladesh as women are more receptive for messages from women. Many NGO's funded by World Bank and many other international institutions employed women workers. These women workers moved at grass-root level from dusk to dawn to spread message of better health for women and family planning. The success of micro credit in Bangladesh saw many women setting up small business ventures. Being employed and earning income women in Bangladesh could better convince their husbands about the benefits of small families. Islam does not prohibit birth control measures; it is against infanticide.

Sri Lanka is another country in India's neighbourhood which also witnessed success with falling fertility rate. This success of falling fertility rate in Sri Lanka is attributed to many programme and non-programme factors. In fact, Sri Lanka became only country in the region to achieve replacement fertility of 1.9 in 1998. Four factors have been identified as the most important determinants of Sri Lanka's falling fertility rate, which include proportion married, contraception, post-partum infecund ability, and induced abortions but recently the fertility rate in Sri Lanka is increasing since 1.9 in 1998 to 2.35 around 2010. This interesting phenomenon of rising fertility rate after achieving success with falling fertility rate is great case study for India. Many factors are responsible for this phenomenon. This increase in Fertility rate in Sri Lanka can be attributed to resumption of peaceful life in Sri Lanka after decades of internal strife.

Vietnam's total fertility rate dropped from 5.7 births per woman in 1979 and 3.8 births in 1989 to 2.0 for the period 1998–2002. The years of conflict in Vietnam during 60's and 70's and huge number of displaced people coupled with rigours population policy followed by regime lowered the fertility rate in Vietnam.

Maldives are a group of 1190 small islands in the Indian Ocean south of India. Only 201 of the islands are inhabited. Until recently the country was considered to have high fertility rates for South Asia, although there have been signs of fertility decline since the early 1990s. The total fertility rate of the country, which stood at 6.4 during the early 1990's. All methods of contraception are legal, and the Government supports family planning services. It has also developed educational programmes on population issues, in particular family planning. Decline in fertility can be attributed to several factors, such as the increased use of contraception, improved schooling opportunities in rural areas and the political endorsement of family planning at the highest level.

Bhutan is a country which is more in news because of being number one in world Happiness Report than having low fertility rate. The population of Bhutan is only just over half million (797000). There is acute problem of workforce in Bhutan. So, low fertility rate in Bhutan is rather a bane than boon.

IV. CONCLUSIONS

Through the use of visualizations, the task of data processing and locating hidden messages and trends in data is not new. During prehistoric times, inhabitants of that times used cave paintings to visualize their thoughts. From those cave paintings of prehistoric times to very modern visualizations and info graphics, the art of presenting data or information has touched new height. The proliferation of computers and explosion of data coupled with data processing software both paid and free has given tremendous push to art of visualizing data. In this age of big data one is continuously bombarded with mind boggling quantity of data. Out of this huge quantity of data one has to find new ways and means to present data to masses. Data today is used in very innovative way. Visualizations of data can be understood by all and presentation of data in the form of visualization can help policy makers to concentrate their efforts to the centre of problem rather than spending time in puzzles of data. Visualization give a touch of storytelling to the data. One begins by simple visualization and over the period of time numerous visualizations and plots give many different ideas to researchers.

Los Angeles police department is using visualizations to bring down crime. Data is plotted and visualized in a way to high light certain area with heat maps or other visualization techniques to understand the pattern of crime. In many cases police have been successful in preventing crimes by being at the right place at the right moment. In medical sciences, visualizations are used to highlight diseases affected parts so as to concentrate maximum efforts to that part only. Radiology which is used to shrink and kill cancerous cells is harmful for healthy cells. With the help of data visualizations radiologists direct radiotherapy rays to only highlighted areas. Meteorologists are using visualisation of data to predict weather with degree of accuracy. In political science infographics are used to present information about political trends for the understanding of common masses. The most well-known case of information visualization is Google maps. These are used by almost everyone who possess smart phone today. Fraud detection is another area where use of visualizations is made to detect and highlight the pattern of fraud. Textual and sentimental analysis of data can help to establish pattern in messages which are used in social media today. The analysis of tweets of many famous personalities can give insight into their personalities and their hidden agendas. It is now possible to visualize whole literature into beautiful plots and graphs to analyse them.



Through this paper an attempt has been made to look at numbers as visualizations to highlight the given objective in the beginning. Visualizations have helped in finding the countries which are having success in controlling. By examining each country central to the objective of this paper many interesting observations have been made. It is not necessary to wait for rise in income as factor to lower fertility rate. 12 countries are having success with lowering of fertility rate and many countries in this group did not wait for rise in income as factor to drive down fertility rate. Good policies and effective implementation brought them success. In some cases, the social structure and religion is not an impediment to the achievement of the objective. Bangladesh, with concentrated efforts both at social and economic level brought down the rising fertility rate. Bangladesh cleverly used women as tool to convince other women to adopt family planning. The economic empowerment of women of Bangladesh through micro credit schemes make them realized the benefits of small family. NGO's in Bangladesh are used very effectively to contribute to these efforts of lowering of fertility rate.

In case of El Salvador female sterilization as method to control fertility rate is used very successfully. In Vietnam, the government adopt very strict population policy.

The use of visualizations, graphs and plots huge impact in presentation of data to the policy makers. Instead of making bulky reports full with data, use of info graphic can expedite the formulation and implementation of policy.

REFERENCES

- [1] W. Chang, R Graphics Cookbook, Shroff Publishers & Distributors Private Limited: 396, 2017.
- [2] N. Dangalle, "Demographic transition in Sri Lanka," Sri Lanka Journal of Social Sciences (243): 1-29, 1982.
- [3] I. Naseem, "Rapid fertility decline in Maldives: An assessment," 2004.
- [4] E. Persson, WHO package.
- [5] M. Rahman, J. DaVanzo, and A. Razzaque, "When will Bangladesh reach replacement level fertility? Role of education and family planning services," (2017)
- [6] H. Wickham, and W. Chang, www.ggplot2.org and ggplot2 package
- [7] www.r-project.org
- [8] www.who.int
- [9] An Overview of Population and Development in Vietnam : Population Reference Bureau (www.pbs.org)