**DATA VISUALIZATION**

**Prof. Johannes Han Reichgelt, PhD**

**Final project report**

**on**

**LITERACY RATE**

* **Sai Swetha Somi**

**(U04880400)**

**Introduction:**

Literacy rate is nothing but the percentage of population who are proficient in reading and writing. It is the essential ability for society to participate in social and economic activities, access the information. High literacy rates are linked to higher educational attainment, expansive economic prospects, and better health outcomes. On the other hand, low literacy rates can result in poverty, social marginalization, and restricted access to necessities like education and healthcare.

The availability and standard of education, socioeconomic situations and cultural perceptions of literacy are only a few of the variables that might have an impact on literacy rates. Governments and development organizations prioritize increasing literacy rates as one of their developing objectives.

Literacy has no single definition as many nations and organizations use various methodologies to measure literacy, which makes it difficult to determine literacy rates. Nevertheless, despite the difficulties, literacy rates continue to serve as a crucial parameter of a population’s level of education and as an important component of the promotion of economic and social development.

**Aim of the Project:**

The aim of creating dashboards and developing visualizations for literacy rates is to promote greater understanding of the importance of literacy and to compare literacy rates across different countries and regions, by which we can identify areas that need attention and intervention to increase the literacy rate, promote economic and social development. By presenting literacy rates in visual format, it would be easier for one to identify the trends and patterns observed on how literacy rates changed over time.

**Data sources:**

I have gathered data from two different sources, global dataset from world bank -education-literacy which is an indicator for the development and the other one is from data.gov, extracted data from word file (source file) and performed cleaning on data. There are another two sheets which is literacy file and economy data formatted file.

After cleaning the data and excluding the null values, joined both the datasets to create visualizations on how literacy rates have evolved over time and what factors played major role in affecting the literacy rates and observed the trend patterns over time.

**Dashboard 1:**

A picture containing text, diagram, screenshot, line

Description automatically generated

* Dashboard 1 consists of 5 visualizations.
* Different age groups in a country.
* Average Literacy rate for each country from the years 2010-2018
* Average country wise literacy rate
* Highest or lowest literacy rate for every country
* Comparing literacy and economy rate by country over time.

**Dashboard2:**

* This consists of 5 visualizations.
* Economy rate of each country within a region
* Economy rate for each country over gender
* Number of records for each region over time
* Animation on Average Literacy rate for every year
* Animation on Economy rate for each country

A picture containing text, diagram, screenshot, plot

Description automatically generated

**Analysis on each visualizations:**

From 2010 to 2018, the data depicts trends in the frequency of the number of records gathered from various locations. With a peak of 168 in 2014, records were most common in Latin America and the Caribbean, while they were fewest in Oceania (12 in 2012 and 14). An all-time high of 204 records were gathered in Sub-Saharan Africa in 2018.

A peak of 59 records collected in Central and Southern Asia in 2011, whereas 96 in Northern Africa and Western Asia in 2017. The most erratic pattern was in Eastern and South-Eastern Asia, which reached 72 in 2010, fell to 9 in 2012, and then rose once again to a peak of 57 in 2018. A peak of 156 records are noted in Europe and Northern America in 2011, while there is no data for 2018 yet.

A picture containing text, screenshot, parallel, line

Description automatically generated

A picture containing text, plot, line, diagram

Description automatically generated

The above animation depicts information about average literacy rate for every year i.e. from 2010 to 2018. I have added the country name as filters, where we can select the country. This visualization is after selecting all countries around the world. As we can see the patterns , there is downfall and rise in the literacy rate from the year 2010 – 2014 and there’s a constant increase in literacy rate from 2014-2016 and then there is a fall.

A picture containing screenshot, text, diagram, circle

Description automatically generated

For this visualization, according to the data the economy rate spent over education for both female and male is depicted, where females having a slightly higher rate compared to males. Even though the girl childs are majorly focused are for education., the economy spent over the education for both the males and females are almost equal as there is no gender disparity in education.

A screen shot of a graph

Description automatically generated with medium confidence

This visualization represents the economy rate for each county which means how much money is spent over education for each country is animated.

**Research Questions:**

1. **How have literacy rates changed over time in a particular country or region?**

This research questions answers about how those literacy rates changed over time for a country or specific region and it depicts that the importance of improving literacy rates can lead to a better job opportunity, improved health care and increased economic growth.

As we can see that there are several countries that have the highest literacy rate, such as Armenia, Belarus, Estonia, Italy, Jordan, and Kazakhstan. These countries have a literacy rate greater than or equal to 98%. There are also countries with lowest literacy rates such as Afghanistan, Burkina Faso, Chad, Guinea, and Liberia that have a literacy rate of less than 40%. We can also analyze that the literacy rates have improved over time for many countries like Bangladesh, literacy rate has increased from 55% in 2010 to 69% in 2018. In Ecuador, the literacy rate has gone upto 90% from 88% and similarly for Honduras.

A screenshot of a computer

Description automatically generated with medium confidence

Fig 1: Average Literacy Rate (2010-2018)

1. **What is the relationship between literacy rates and economic development?**

From the analysis we can see literacy and economy rates of Afghanistan and Bangladesh from the year 2011 to 2018.

* Afghanistan has an average literacy rate in 2011 was 31% that went up to 38% in 2018. Whereas the economy rate was 38.08 in 2011 and increased a little to 38.40% in 2018.
* In Bangladesh, the average literacy rate increase from 55% in 2011 to 69% in 2018and economy rate was 25.59 in 2011 that gradually decreased to 23.38 in 2018 with fluctuations over years.
* There might be many reasons to affect the literacy rate in Afghanistan when compared to Bangladesh, even though the economy spent over education in Afghanistan is more than the economy spent over the education in Bangladesh. Factors like size of the population, education system, infrastructure and international aid might have played a role in decreasing literacy rate corresponding to economic development for both these countries.

A picture containing text, line, diagram, plot

Description automatically generated

Fig 2: Literacy and Economy rate by country over time

1. **How does literacy proficiency vary among different age groups within a given country and what factors may contribute to these variations?**

* From the analysis, we can say that the literacy rates vary among different age groups within a country. The highest literacy rate is observed among the age group of 15-24 with an average of 83% and 25-64 group has a lower literacy rate with 54% and 65+ age group has the lowest rate with an average of 31%.
* There may be several factors contributing to these variations in literacy rates among different groups. Younger groups have had greater access to education and resources than older age groups, resulting in higher levels of literacy. Another factor is the quality of education provided across different age groups, with younger age groups receiving comprehensive educational experiences. It is necessary for identifying areas where target educational and literacy programs may be needed.

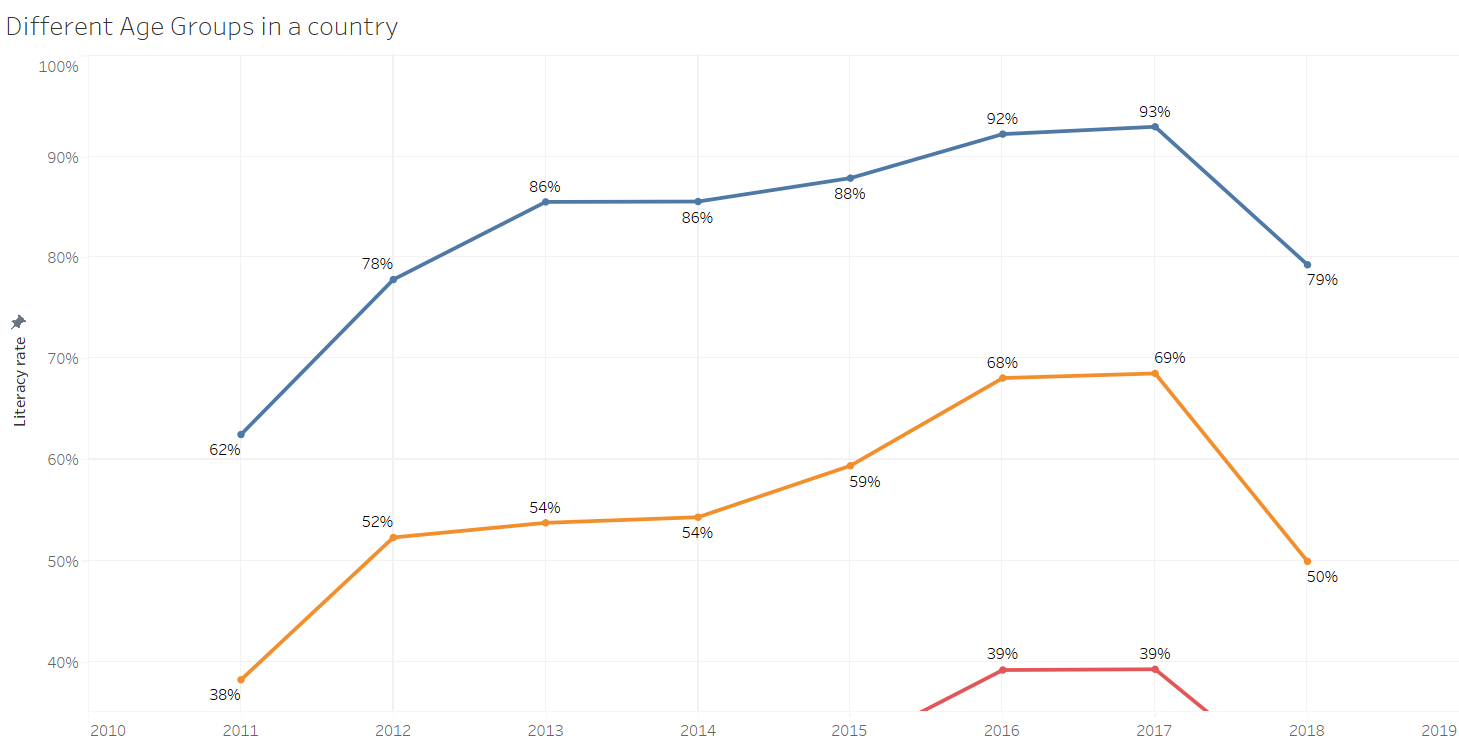


Fig 3: Age groups in a country

1. **What are the trends and patterns in average country wide literacy rates over time?**

* I have used a Map graph, to depict the average Literacy rates in different countries around the world. The intensity of the colour in the graph is high if there is high average literacy rate and vice versa.
* Kazakhstan and Russian Federation has the highest average literacy rate with 0.997 and 0.9966 respectively

A map of the world

Description automatically generated

Fig 4:Average Literacy rate by country

1. **What are the factors that contribute to the highest and lowest literacy rates for each country?**

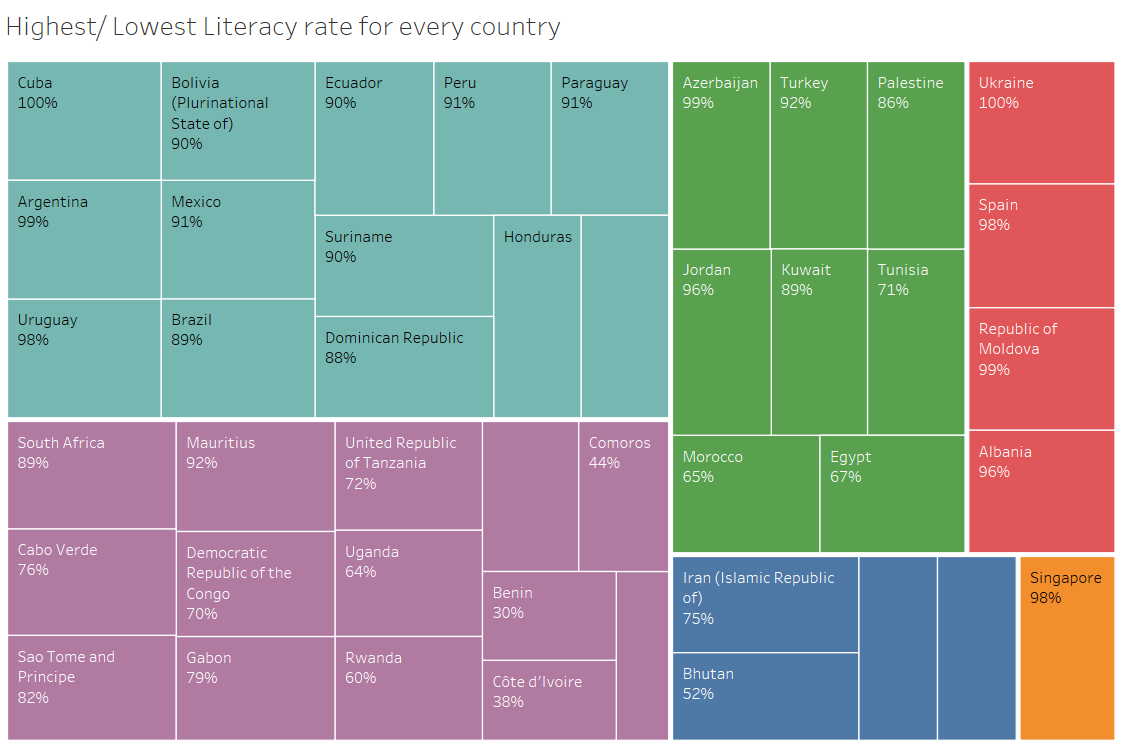


Fig 5: Min/Max Literacy rate by Country

This visualization shows an analysis about the average and maximum literacy rates for various countries in different regions. It depicts that Singapore has the highest literacy rate of 98% in Eastern and South-Eastern Asia, Ukraine has the highest literacy rate of 100% in Europe and Northern America. When it comes to lowest literacy rates, Benin has the lowest average literacy rate of 30% in Sub Saharan Africa. It highlights the disparities in literacy rates over different regions and countries.

1. **How do the economic indicators of countries within a given region correlate with their literacy rates and what implications might have for policies and programs aimed at improving literacy and promoting economic development?**

A screenshot of a computer

Description automatically generated

Fig 6: Economy rate of each country within a region

This visualization shows the literacy rates for various countries in different regions like central and southern Asia, Eastern and south-Eastern Asia regions and countries belonging to that particular region and their economy rate, Economy rate is that which says that the expenditure or economy spent over education and programs or policies implement in that particular region or the country to improve literacy, an education development indicator. From the above scatter plot, we can say that

* + Uzbekistan and Afghanistan has the highest and lowest literacy rates in central and southern Asia region.
  + Brunei Darussalam and Myanmar ranks the highest and lowest in eastern and southeastern Asia region.
  + For Europe and northern America region, Latvia and Cuba ranks the highest and lowest with minute difference.

**Future Research Questions:**

1. How do literacy rates vary based on different socio-economic indicators such as income?
2. How is literacy going to be correlated with the level of access to technology and the internet?
3. How do literacy rates in rural and urban areas compare with each other and what factors contribute to the differences?
4. How has COVID-19 impacted the literacy rates, particularly in Least Developing Countries (LDC)?

**Conclusion:**

* From the first visualization, we can see that literacy rates vary greatly among different countries. This indicates a significant disparity in educational opportunities and resources available to individuals in different parts of the world. It's encouraging to see that some countries have made progress in improving their literacy rates over time, but there is still much work to be done to ensure equal access to education for all.
* The second visualization shows the relationship between literacy rates and economic development in Afghanistan and Bangladesh. It's clear that investing in education is crucial for overall economic growth. While Bangladesh has seen some improvement in literacy rates, it has also experienced a decline in its economy rate. Afghanistan, on the other hand, has been making slow progress in both areas.
* The third visualization highlights the differences in literacy rates across age groups within a country. It's not surprising to see that younger age groups tend to have higher literacy rates, as they have had greater access to education and resources. However, it's concerning to see that the 65+ age group has the lowest rate, as this suggests that there may be a lack of educational opportunities for older individuals.
* The fourth visualization provides insight into the average literacy rates of various countries around the world. Kazakhstan and the Russian Federation have the highest rates among the countries analyzed, indicating that they have been successful in providing widespread access to education and resources.
* The fifth visualization emphasizes the importance of investing in education and literacy programs in order to promote social and economic development. It's clear that low literacy rates can hold individuals and communities back from reaching their full potential, and improving literacy rates can lead to better opportunities and quality of life.
* Finally, the sixth visualization shows the correlation between literacy rates and economy rate in different regions. It's not surprising to see that countries with higher economy rates tend to have better literacy rates, as investing in education and literacy programs can lead to greater economic development. This highlights the need for policymakers to prioritize education as a key factor in promoting economic growth and development.