Exploratory Data Analysis on World Literacy Rates (UNESCO) along with visualizations through R

Presented by Niraj

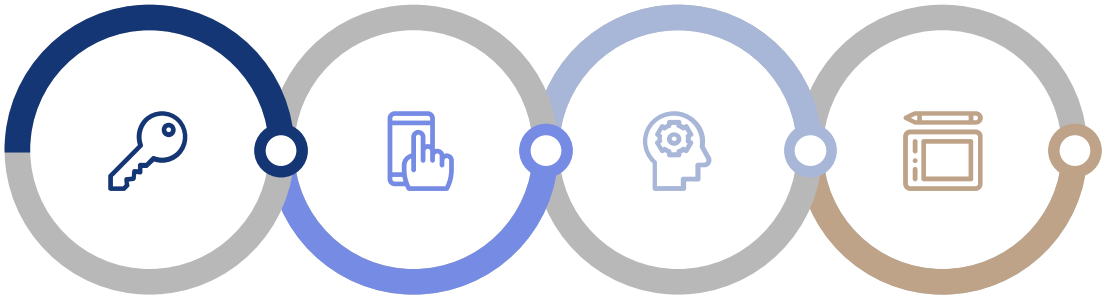
Understanding and Know about data set:- https://data.world/makeovermonday/2019w46

Access

UNSEO data download

Understand

Columns and Rows Understanding



Use

Use in EDA using R

Create

Conclusion and Discussion

Flow of the Presentation

Objective

Set up an objective to find something interesting

Research

Check the data and undestand relation between col. and row

Querry Set up

Set the questions as per objective to find some insights

Cleaning

Data Cleaning checking any outliners, Null, Duplicate

Data Mining

Check the data and its relationship using the R and set the result

Visualization

Visualization the results

Thinking

Think why it happens!

Conclusion

Concluion from the research

Import the packages and Library in R

01

install.packages('ggmap')

Reterive and visualize spatial data from Google maps, Stamen Maps etc

02

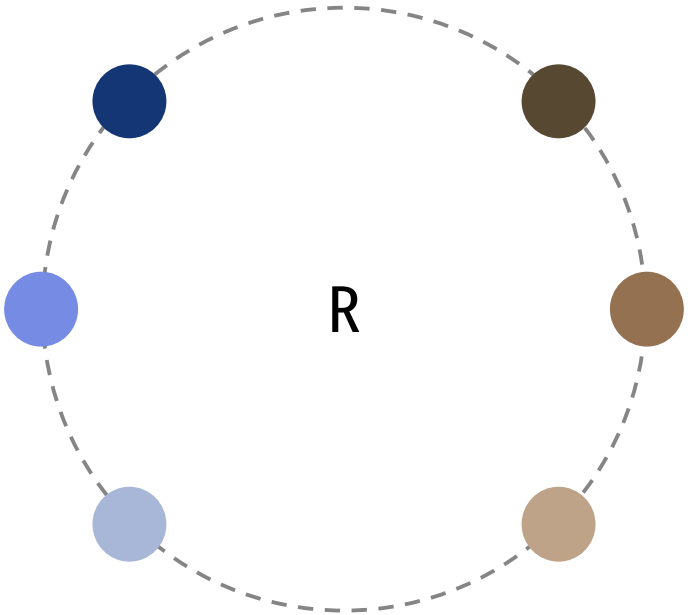
install.packages("ggrepel")

Process of avoiding overlapped texts in plots

03

library(ggalt)

The Earth is the only planet that harbours life



04

install.packages('ggplot2')

Data visulaization package in the R

05

install.packages("ggalt")

Extra Coordinate Systems, 'Geoms', Statistical Transformations, Scales and Fonts for 'ggplot2'

06

library(ggrepel)

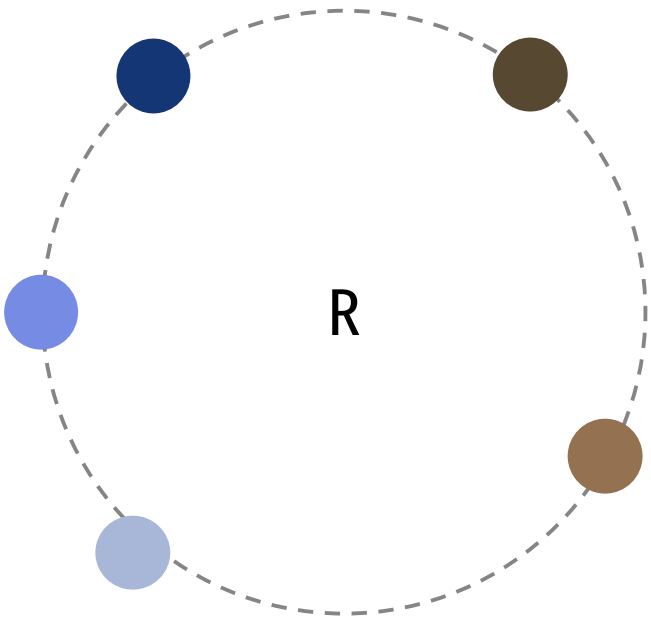
Provides geoms for ggplot2 to repel overlapping text labels

Continued.....

07
library(dplyr)
an ecosystem of several libraries designed to work together by representing data in common formats

08
library(readxl)
makes it easy to get data out of Excel and into R

09
library(ggmap)
Download open sourced map objects



10
library(ggplot2)
Data Visualization

05
library(scales)
internal scaling infrastructure

Objective of the Project:-

Visualization and Conclusion

Conclued the assepts of why it is happening

Gender

Gender ‘Male’ and ‘Female’ Literacy rate

Age Group

Which Age Group is educated

Literacy Rate- Region and Country

Variation in the Literacy Rate

Querry Set up:-

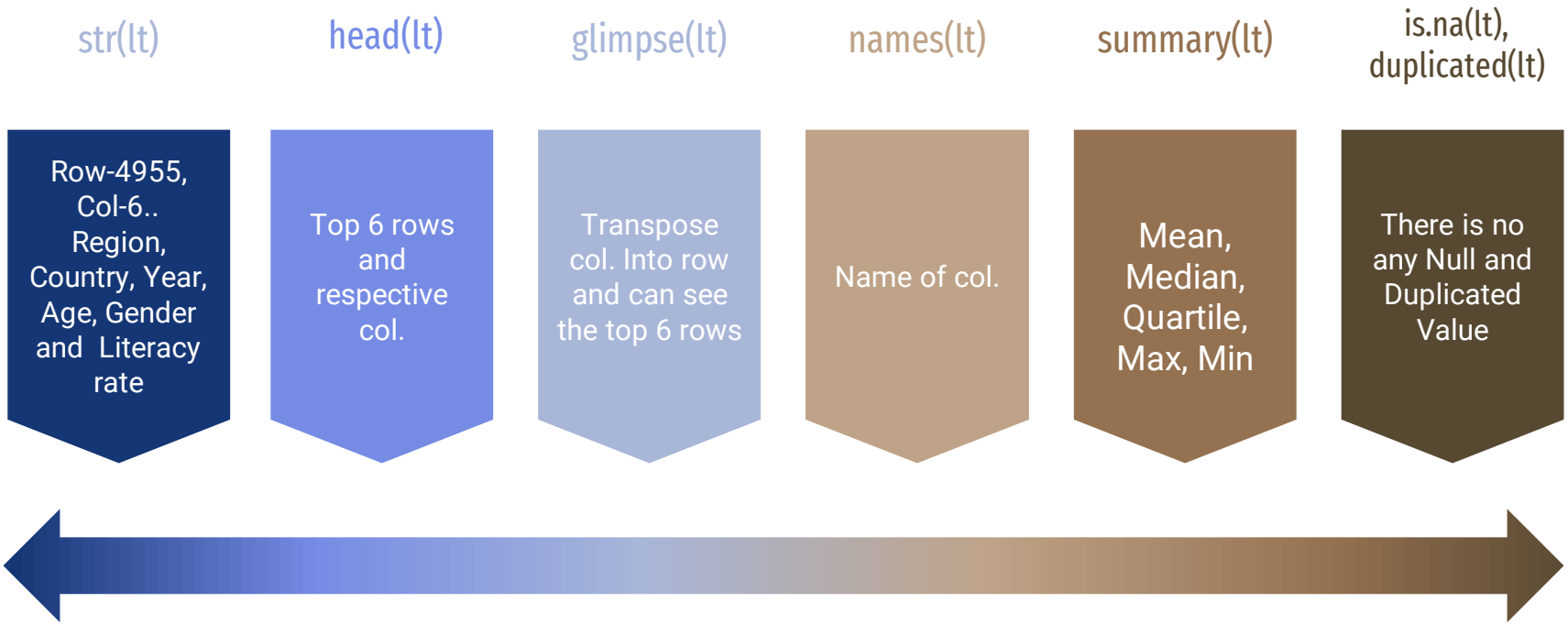
1. Top and Bottom Literacy Rate
2. Literacy rate 2010-2018
3. Focus on India data- Gender
4. High discrepancy between Male and Female
5. Country[female>male] literacy rate
6. Country[male>female] literacy rate
7. Year wise Literacy rate
8. Distribution of Literacy rate among Gender
9. Visulization in MAP

Know about the data and its structures

Region Name	No of Country
Central and Southern Asia	14
Eastern and South-Eastern Asia	14
Europe and Northern America	23
Latin America and the Caribbean	31
Northern Africa and Western Asia	19
Oceania	8
Sub-Saharan Africa	46
7- Regions	**155- Country**

Year	Age Range
2010- 2018	15+, 15-24, 25-64, 64+

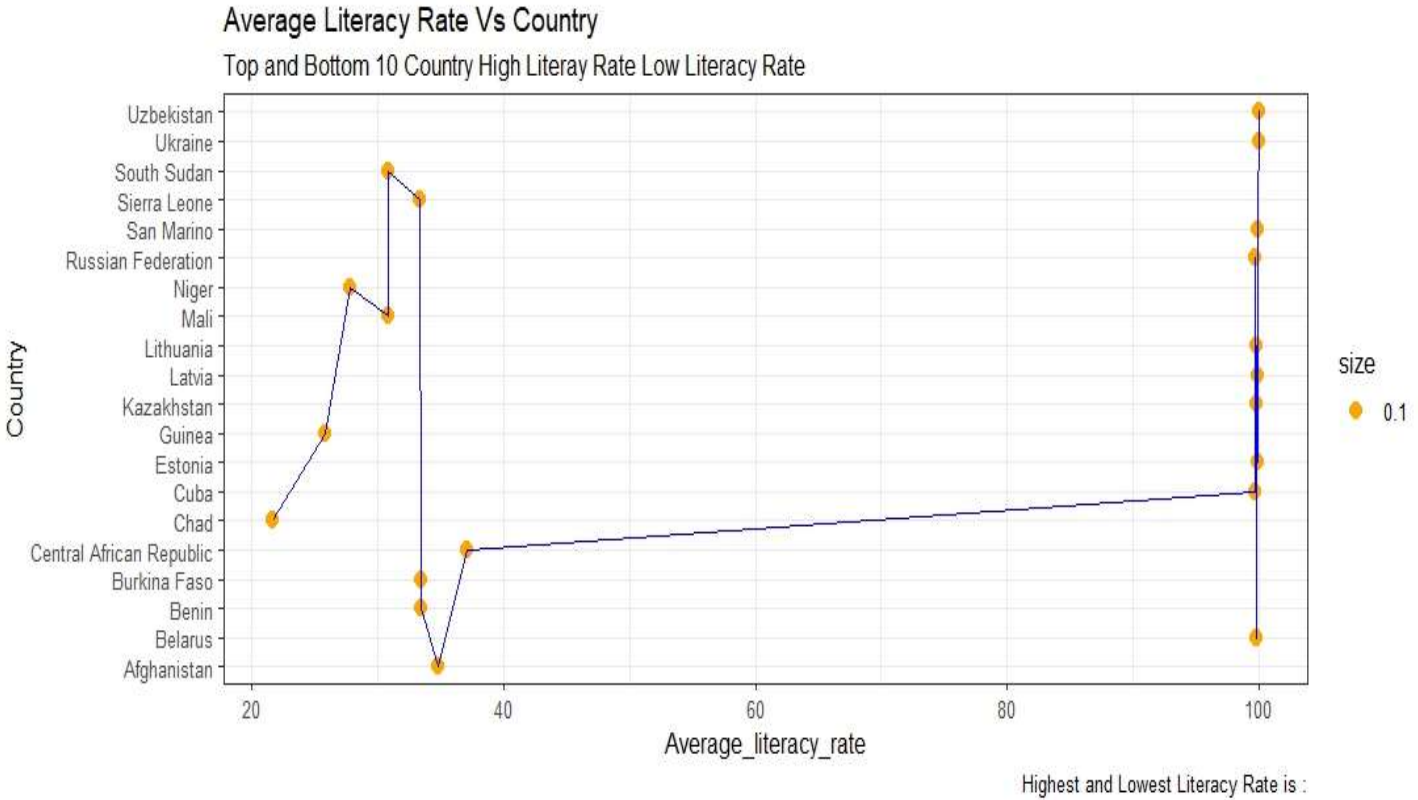
Check and Know about the data:-



Data Mining as per set objective

1-Top and Bottom 10 Country High and Low Literacy Rate

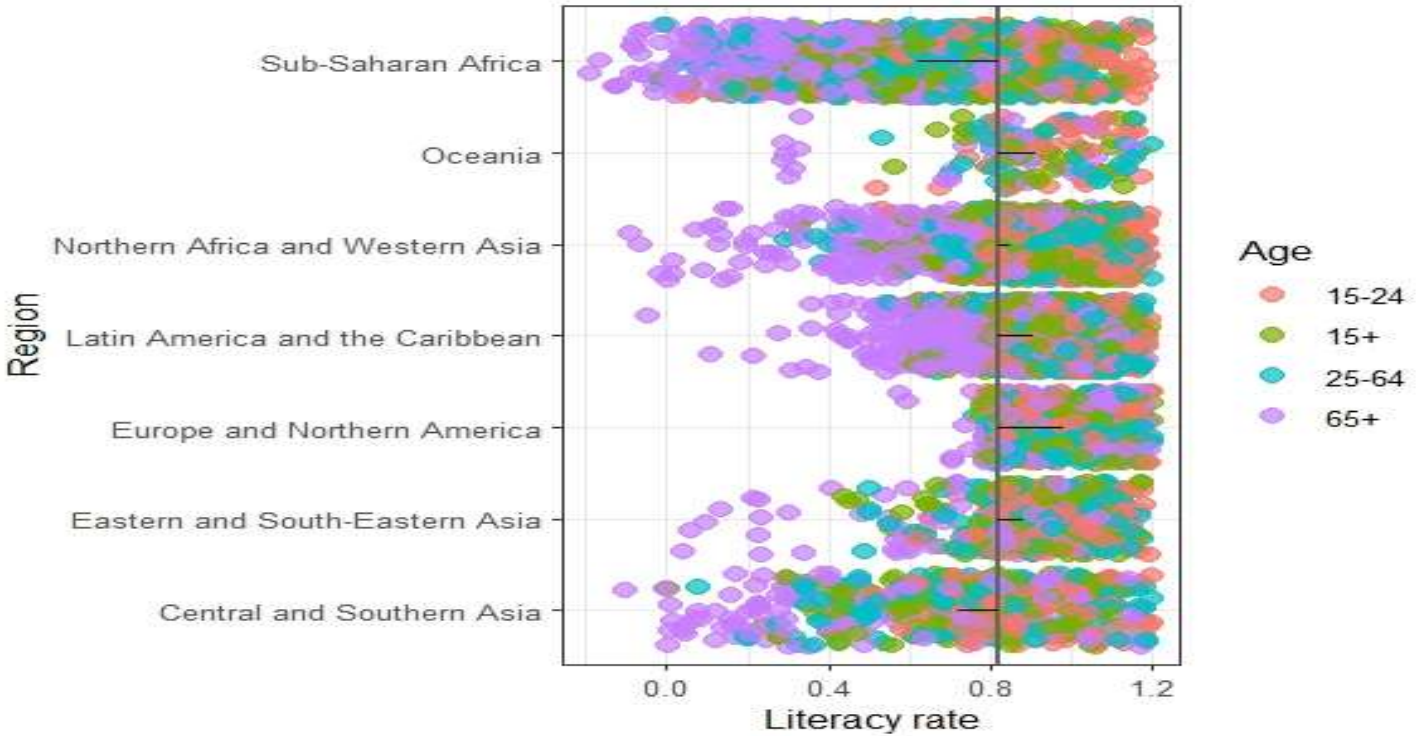
1. Point and Line



Data Mining as per set objective

2- Distribution of literacy rate across regions by age

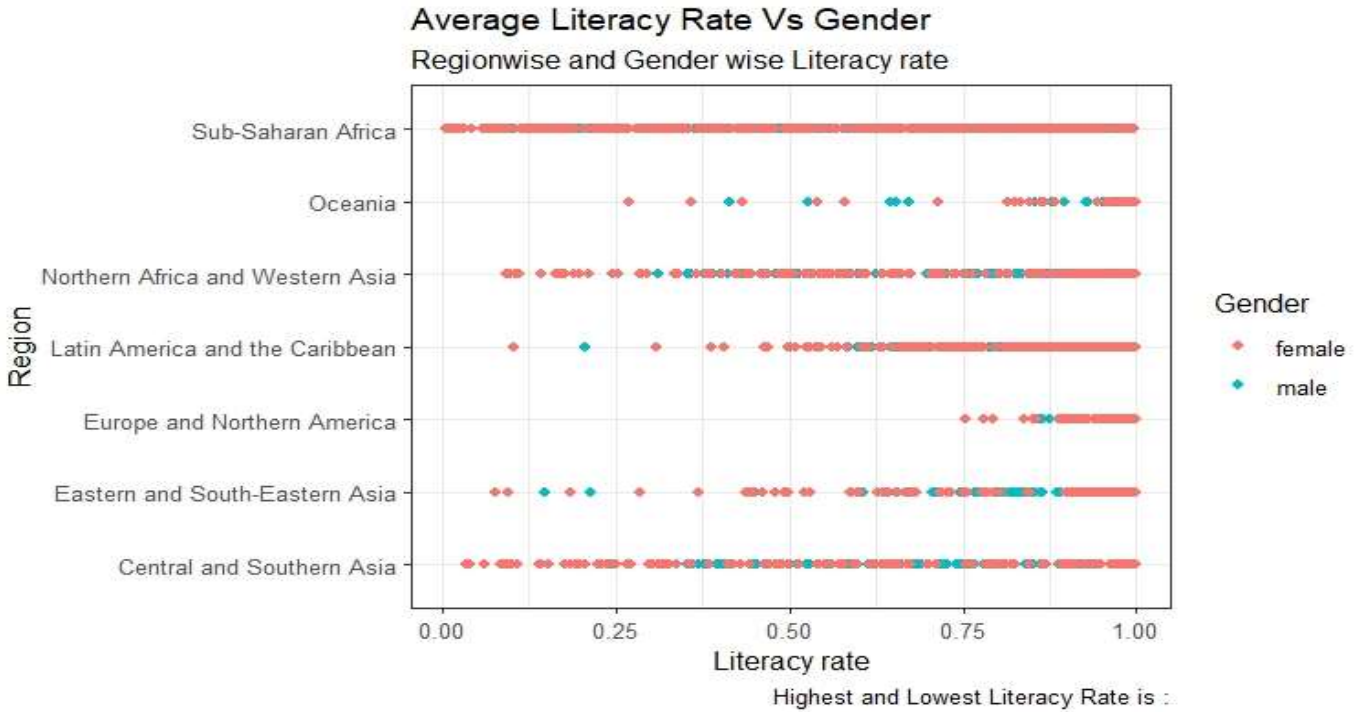
1. Scatter



Data Mining as per set objective

3- Region Wise Gender Wise Literacy rate

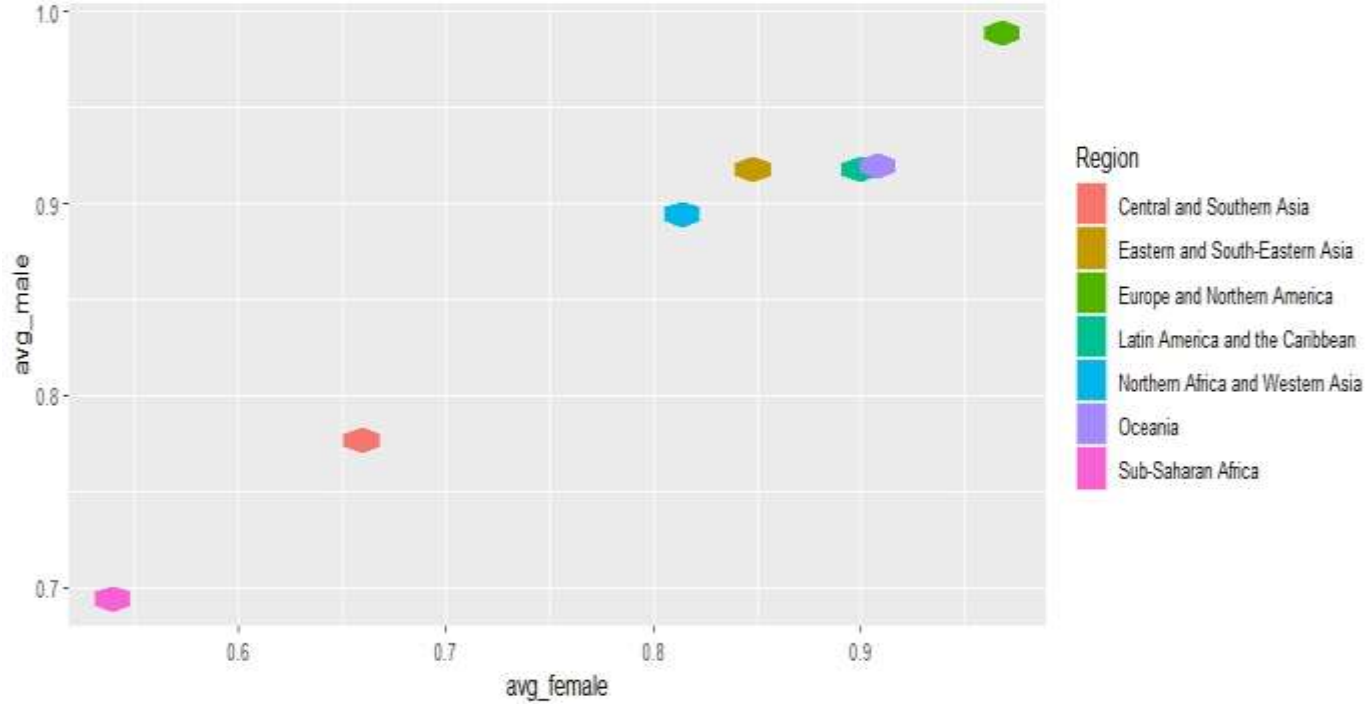
1. Point, Stack



Data Mining as per set objective

4- Region Wise Gender Wise Literacy rate

1. Point, hexa



Data Mining as per set objective

5- Year Wise Literacy Rate

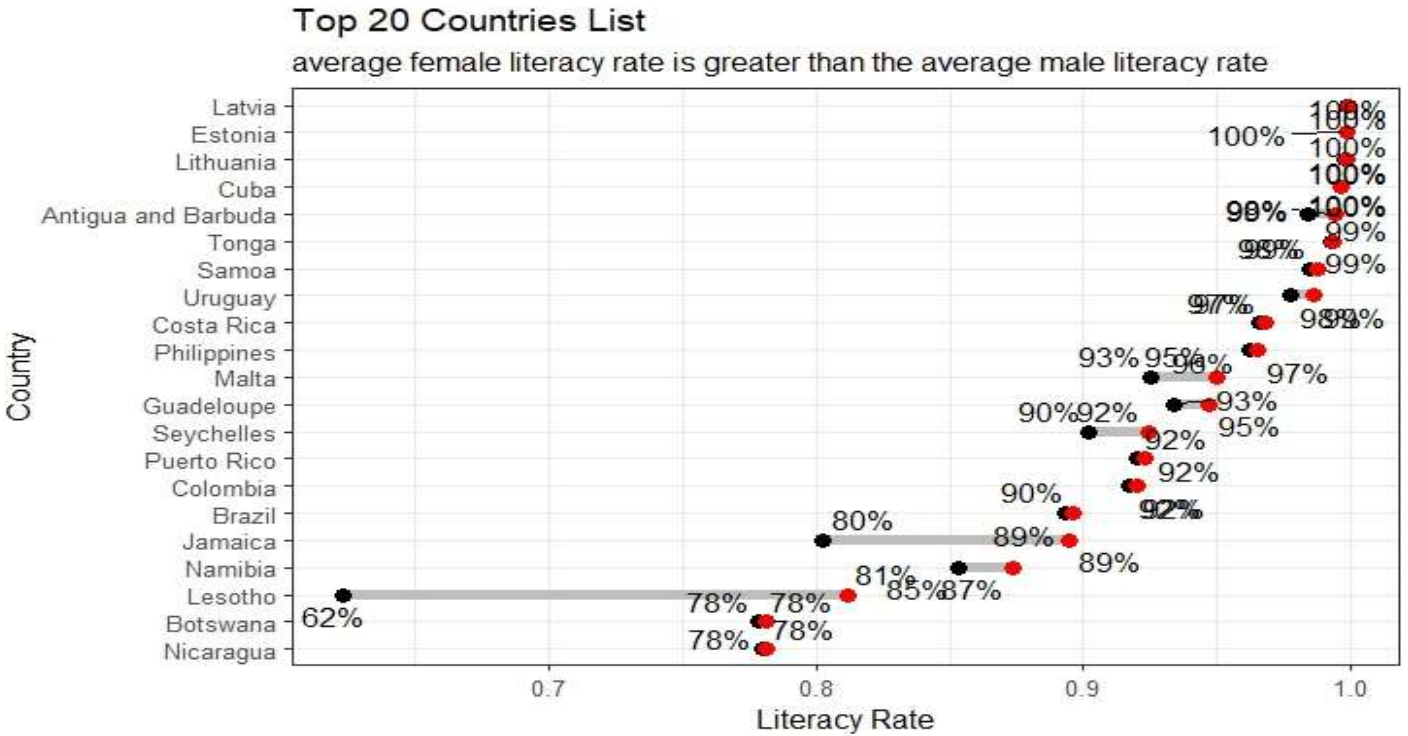
1. Table

Region	avg_2010_1	avg_2011_1	avg_2012_1	avg_2013_1	avg_2014_1	avg_2015_1	avg_2016_1
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
Central and S…	0.800	0.514	0.577	0.706	0.830	0.804	0.865
Eastern and S…	0.857	0.833	0.979	0.946	0.864	0.885	0.885
Europe and No…	0.984	0.975	0.980	0.967	0.985	0.978	0.983
Latin America…	0.905	0.904	0.899	0.890	0.909	0.909	0.892
Northern Afri…	0.818	0.862	0.831	0.812	0.880	0.901	0.907
Oceania	0.560	0.979	0	0.993	0.868	0.945	0
Sub-Saharan A…	0.631	0.684	0.621	0.487	0.627	0.574	0.511

Data Mining as per set objective

6- Countries where the average female literacy rate is greater than the average male literacy rate

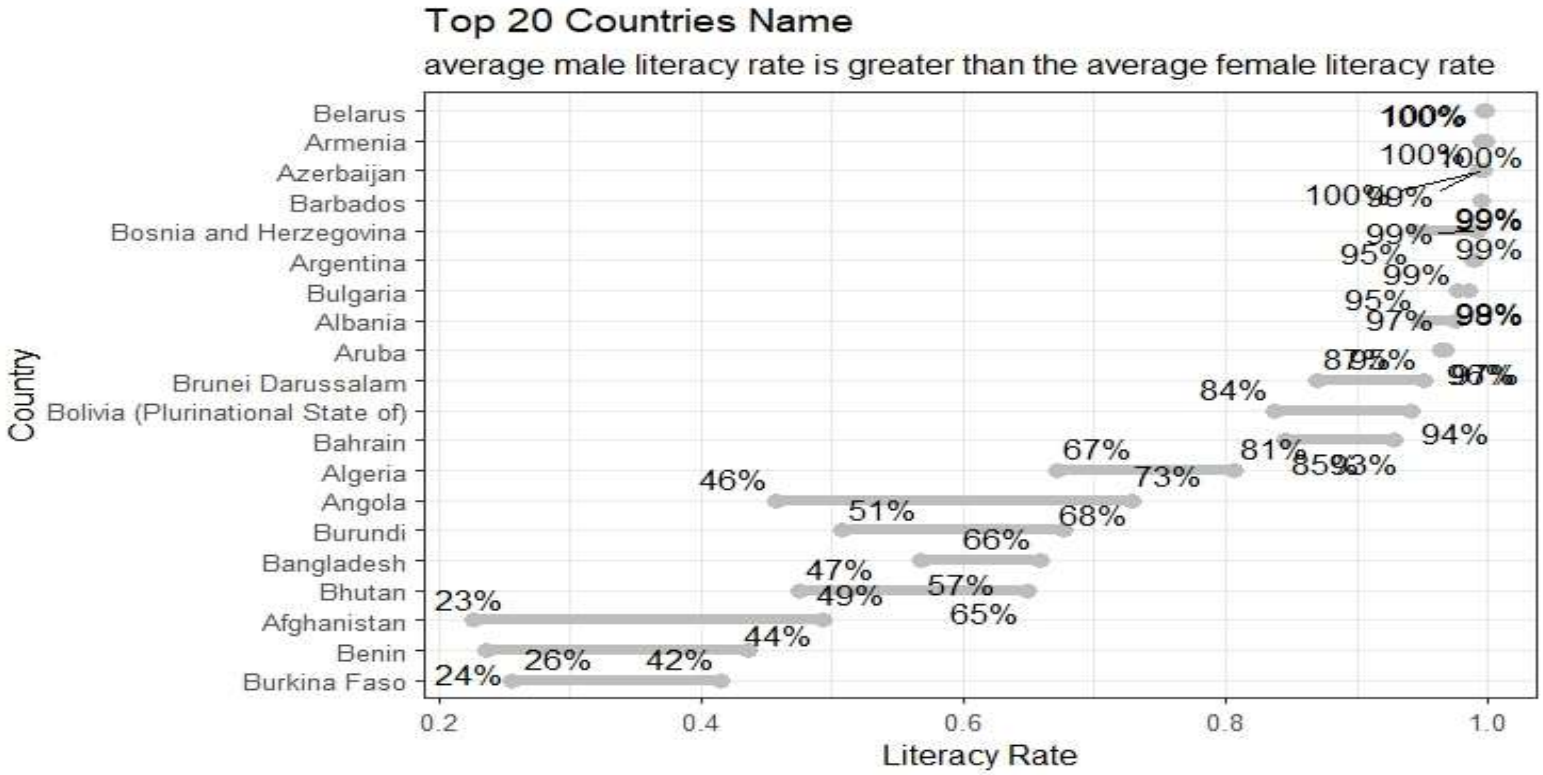
1. Dumbel



Data Mining as per set objective

7- Countries where the average male literacy rate is greater than the average female literacy rate

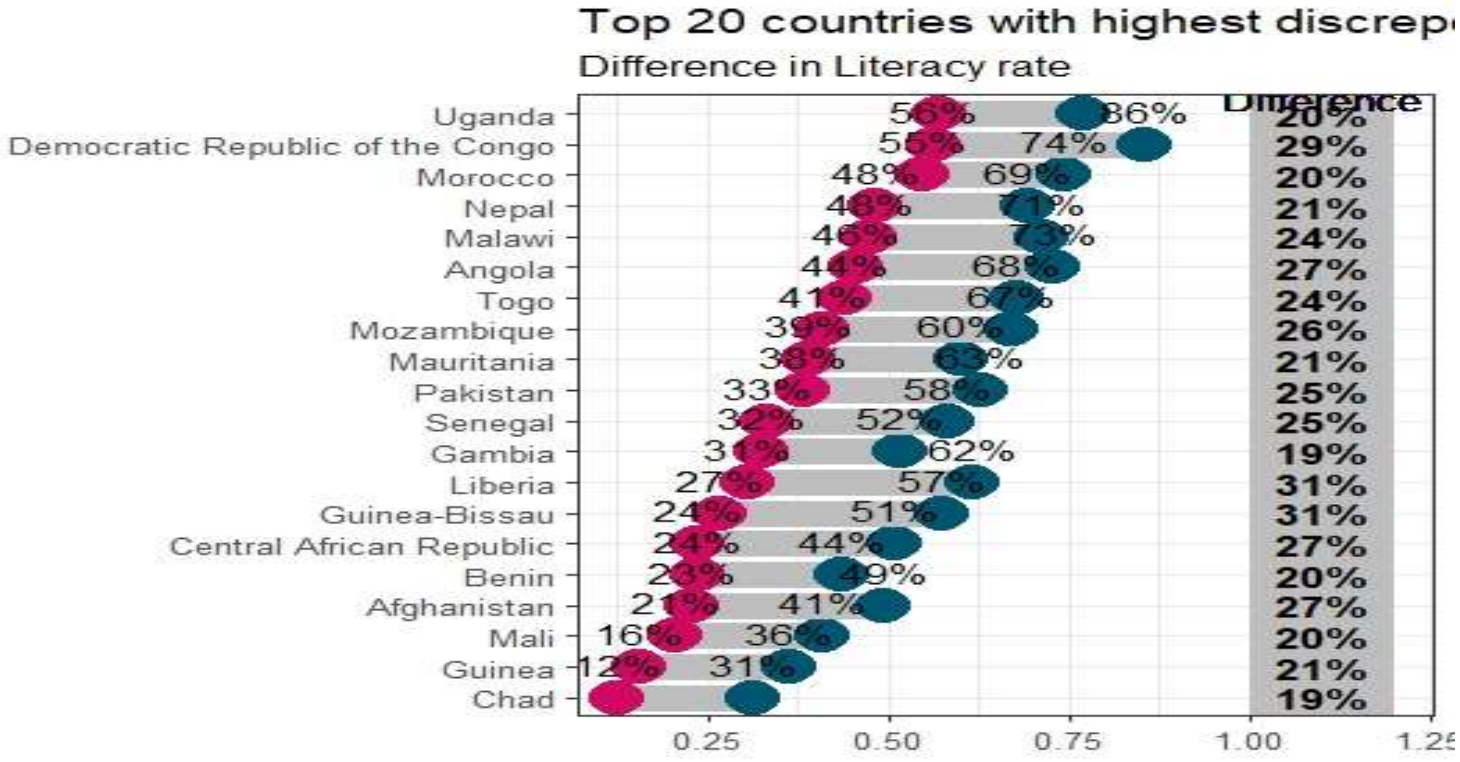
1. Dumbel



Data Mining as per set objective

8- Countries with the highest discrepancy between male and female literacy rates

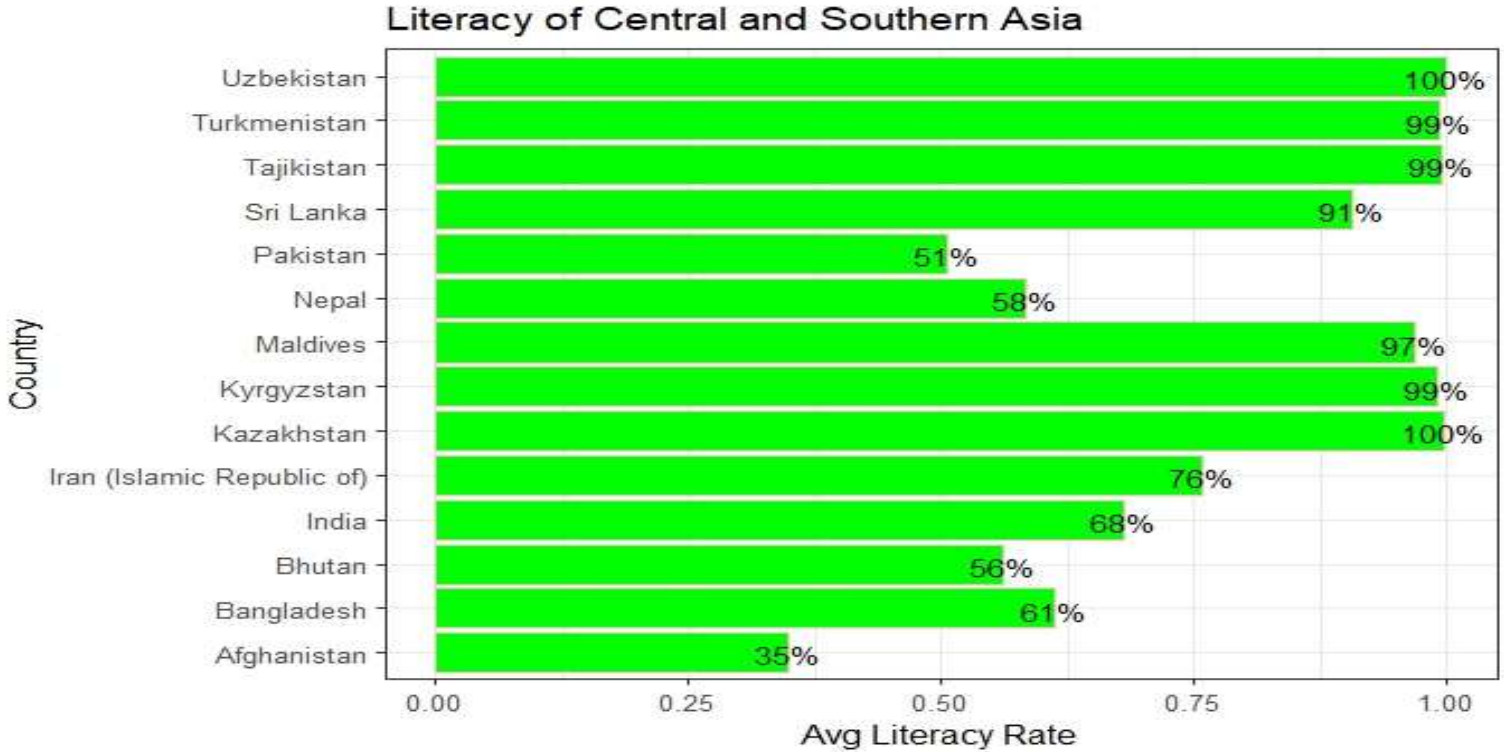
1. Dumbel
2. Female- red
3. Male- blue



Data Mining as per set objective

9- Country India in the Region- 9th Position- Central and Southern Asia

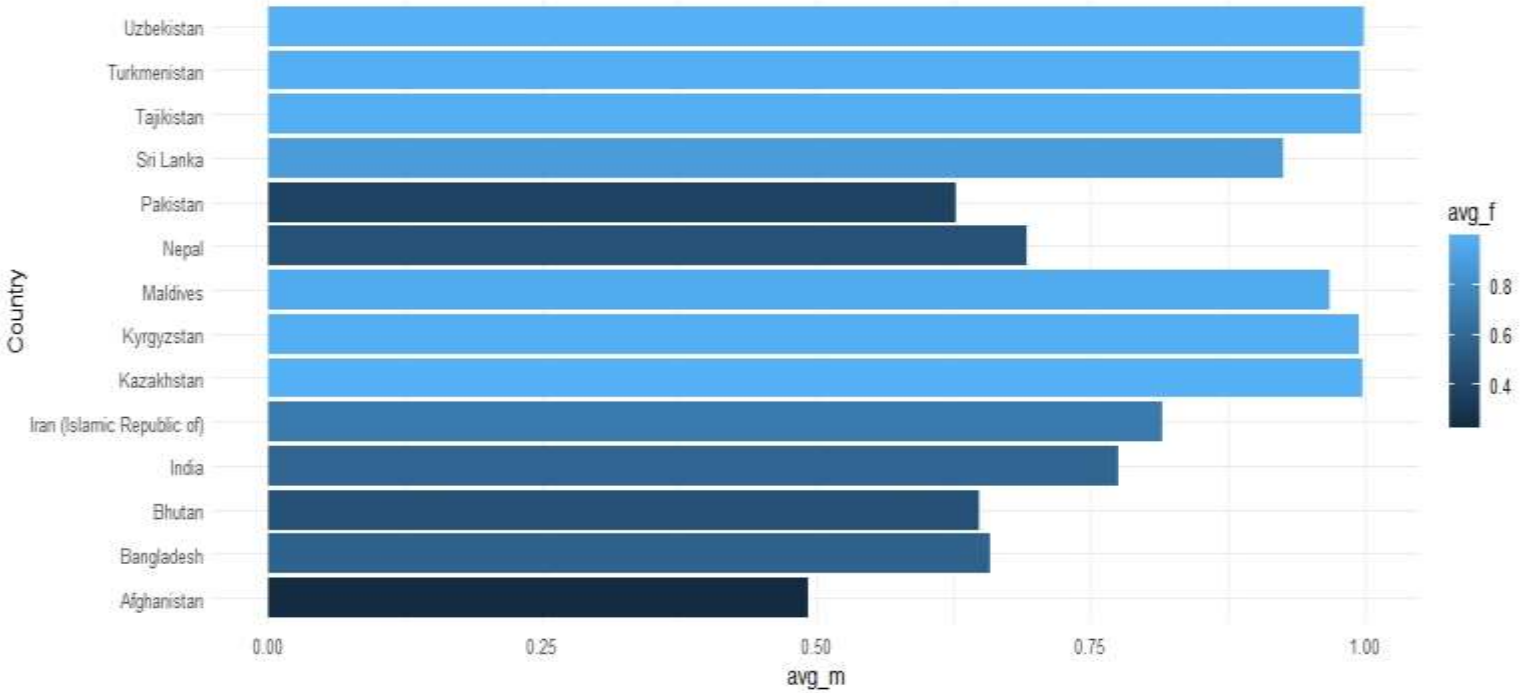
1. Bar



Data Mining as per set objective

10- Male and Female Literacy Rate in Central and Southern Asia

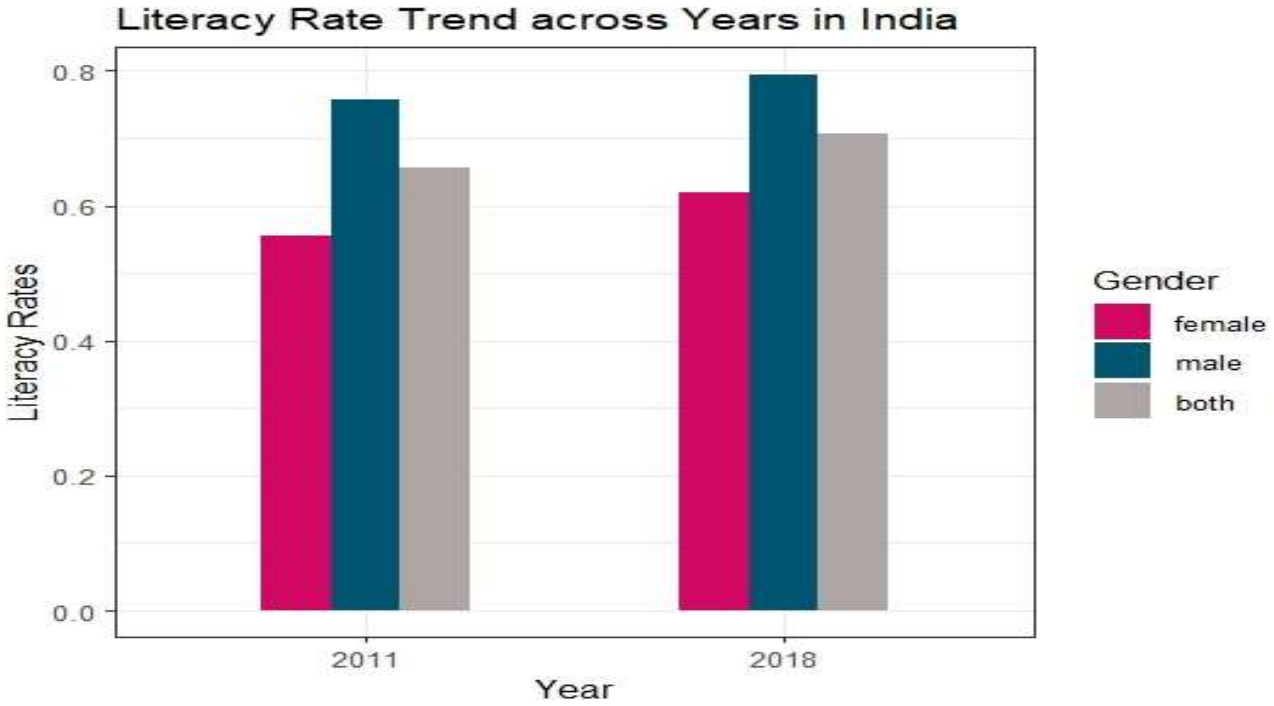
1. Bar



Data Mining as per set objective

11- Literacy Rate by Gender in India by Year

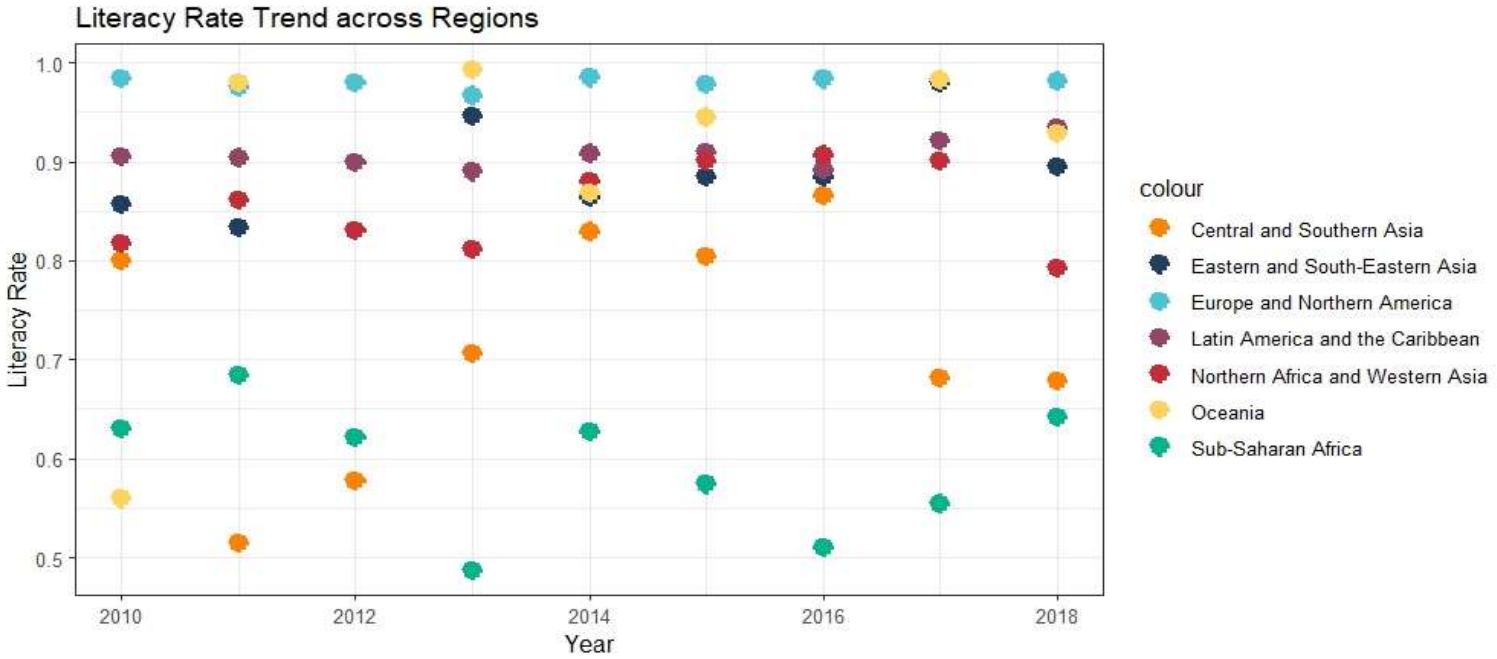
1. Bar



Data Mining as per set objective

11- Literacy rate across regions from 2011 to 2018

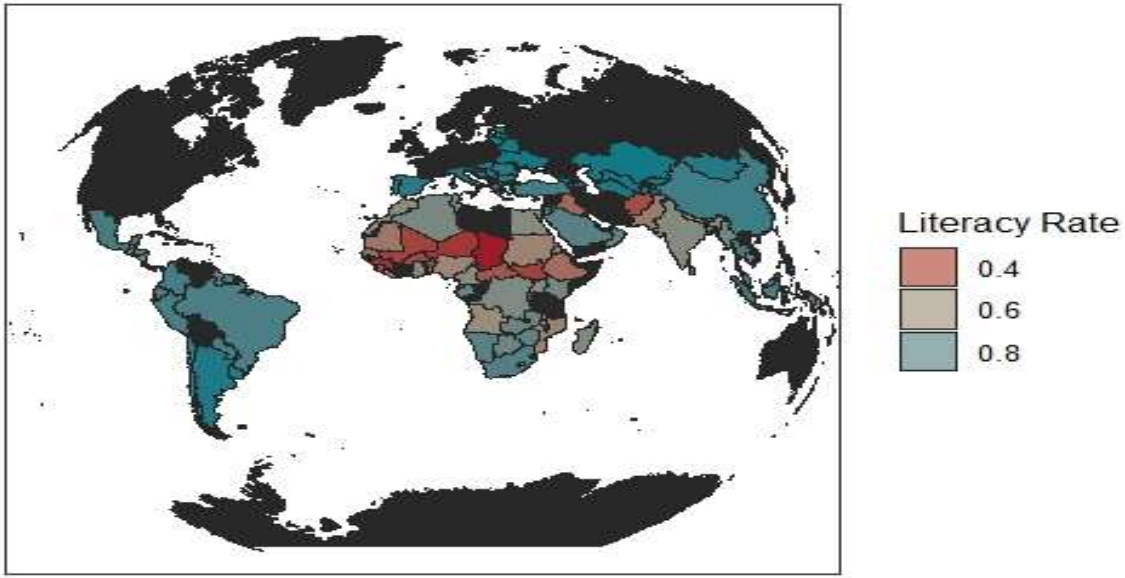
1. Scattered



Data Mining as per set objective

11- Visualization Global Literacy rates on a Map

1. Scattered



Conclusions- Global

Avg Lowest Literarcy Rate

Afganistan- **37%**

Avg Highest Literacy Rate

Uzbekistan, Ukraine- **99%**

Age Group Having Lowest Lit. Rate

65+ Years of Age-Found in Sub-Saharan Africa- **20 to 40%**

Age Group Having Highest Lit. Rate

25-64 Years of Age-Found in Europe and North America- **80 to 95%**

Reason

Political Instability, deep rooted of 'Taliban' armed forces, and traditional practices

Reason

Free Education Policy

Reason

poverty and lack of Government commitment

Reason

access to education and socio-economic conditions

Conclusions- Global

Avg Lowest: Female Lit. Rate

Sub-Saharian Africa- **25 to 50%**

Avg Highest: Female Lit. Rate

Latin America and the Caribbean- **75 to 95 %**

Avg Lowest: Male Lit. Rate

Central and Southern Asia- **37 to 50%**

Avg Highest: Male lit.Rate

Oceania- **70 to 95%**

Reason

historical, cultural, and socioeconomic inequalities

Reason

Robost Administrative and insititutional reforms

Reason

Large families in rural regions are unable to afford education, and most children cannot attend primary schools

Reason

access to education and socio-economic conditions

Conclusions- Global

Avg Female Lit. Rate> Avg Male Lit. rate

Lesotho- **16 % difference**

Avg Female Lit. Rate> Avg Male Lit. Rate

Lativa- **0.1% difference**

Avg Male Lit. Rate> Avg Female Lit. Rate

Angola- **22% difference**

Avg Male Lit. Rate> Avg Female Lit. Rate

Belarus- **0.1 % difference**

Reason

Govt encourages to Female Education

Reason

Govt enforces Policy to have better education systems

Reason

Low level of income

Reason

Govt more budget on Education

Conclusions- Global

Highest disaparities in male and female Lit. Rate and Lit. Rate- 20 to 55%

Liberia and Guinea-Bissau- **31% difference,**

Highest disparities in male and female Lit. Rate and Lit. Rate 10 to 30%

Chad – **19% difference**

Highest disparities in male and female Lit. Rate and Lit. Rate- 60 to 80%

Uganda- **20% difference**

Reason

violence against women, unpaid care, early marriage

Reason

high population growth, lack of financial resources, poverty, conflict, regional insecurity leading to population displacement, and the implications of climate change

Reason

Poverty

Conclusions- Asia

Asia: Lowest Literacy Rate

Afganistan- **35%**

Asia: Highest Literacy Rate

Uzbekistan – **100%**

Asia: India Literacy Rate

India-**68%, ranks 9th in Asia**

Reason

The resurgence of the Taliban has precipitated a dire educational crisis in the country, Denial of education for women

Reason

Govt decision of free and compulsory schooling

Reason

economic inequities, gender discrimination, caste prejudice, and technological hurdles

Conclusions- India

India: Female Literacy Rate

Female- **57% in 2011 and improved by 4% in 2018**

India: Male Literacy Rate

Male – **75% in 2011 and improved by 4% in 2018**

India: Overall Literacy Rate

India- **65 % in 2011 and improved by 4% in 2018**

Reason

economic inequities, gender discrimination, caste prejudice, and technological hurdles

Govt Actions

The National Policy on Education (NPE) is a policy formulated by the Government of India to promote and regulate education in India

Govt Previous Actions

Sarva Shiksha Abhiyan (SSA) is implemented as India's main Programme for universalizing elementary education, Mid Day Meal

Conclusions- Literacy Rate in World Vs Year

World: Over 9 years

Uncertainty seen in African Continent- [40% to 69%]

World: Over 9 years

Constant seen in Latin America and Caribbean- [90 % to 94%]

World: Over 9 years

Increment seen Central and Southern Asia- [50% to 80%]

Reason

Poverty, lack of infrastructure, and political instability have all contributed to low enrollment and high dropout rates, especially for girls

Reason

guarantee inclusive, equitable and quality education and promotes learning opportunities throughout life for all

Reason

Govt Polices, Right to Education Act

Thank You