Title: District Wise Rainfall in India (1951-2000)

Author: Idhika Nikhil Vaidya

PRN: 22070243021
Course: MSc Data Science and Spatial Analytics
Institute: Symbiosis Institute of Geoinformatics

Subject: Data Driven Governance
Assignment 1

CERTIFICATE OF ORIGINALITY

This is to certify, that the project submitted by me, Idhika Vaidya is an outcome of my original and independent work. I have duly acknowledged all the sources from which the ideas and extracts have been taken.

Name of Author (s): Idhika Vaidya
Designation: Student at SIG
Affiliated Institution: Symbiosis Institute of Geoinformatics
Title of the project: District Wise Rainfall in India
E-mail: 22070243021@sig.ac.in

Acknowledgement

I would like to express my sincere gratitude to all the individuals who contributed to the successful completion of this project for Data Driven Governance.

First and foremost, I would like to thank my teacher Moushami Dasgupta for her guidance and support.Her insights and feedback were instrumental in shaping the direction and focus of the project.

I want to also thank the Symbiosis Institute of Geoinformatics for providing the resources and facilities necessary to carry out the project. Additionally, I am grateful to the Department for their cooperation and assistance throughout the project.

I would like to express my appreciation to my peers for their encouragement and support during the course of this project. Their constructive feedback and suggestions have helped me to improve the quality of the project.

Lastly, I would like to thank my family and friends for their unwavering support, encouragement, and motivation throughout the project. Their belief in me and my abilities inspired me to work harder and strive for excellence.

I want to express my gratitude towards all those who helped and supported me in completing this project for the subject Data Driven Governance.

Index

Introduction

This data set contains monthly rainfall details of 36 meteorological subdivisions of India. It contains the rainfall data for all the districts of India. Sub-division wise monthly data for 50 years from 1951-2000. The rainfall data is measured in mm.Its granularity is monthly. In this project data is visualized using PowerBI with the purpose of deriving meaningful insights.

Metadata

The data value in the columns is the sum of rainfall for the time period 1951-2000. The data is uploaded by the Indian Meteorological Department. The data is taken from a credible government source data.gov.in .
As the data is of previous years, it is static in nature. The integrity of the data is maintained without updating it.

The data has the following 19 columns:

Name Columns:

1. States/Union territories: Names of the States and Union Territories of India
2. District: the name of the districts in each state

Yearly Data:

1. Annual: Rainfall for the entire year

Monthly Data: Rainfall for each month

1. Jan: Rainfall for the month of January
2. Feb: Rainfall for the month of February
3. Mar: Rainfall for the month of March
4. Apr: Rainfall for the month of April
5. May: Rainfall for the month of May
6. Jun: Rainfall for the month of June
7. Jul: Rainfall for the month of July
8. Aug: Rainfall for the month of August
9. Sep: Rainfall for the month of September
10. Oct: Rainfall for the month of October
11. Nov: Rainfall for the month of November
12. Dec: Rainfall for the month of December

Quarterly Rainfall: Rainfall for a particular quarter:

1. Jan-Feb: Rainfall for the first quarter of the year between January and February. This is the beginning of winter.

2. Mar-Apr: Rainfall for the second quarter of the year between March and April. This is the summer season in India.

3. Jun-Sep: Rainfall for the third quarter of the year between June and September. This is the monsoon season in India, most of the states receive the maximum amount of rainfall during monsoon.

4. Oct-Dec: Rainfall for the fourth quarter of the year between October and December. This is the fall season and the beginning of Winter.

Methodology

The data was downloaded into an excel file which was then loaded into Power BI. Power BI was used to create interactive dashboards with multiple graphs as well as charts that describe the data and its various aspects with the aim of obtaining good insights.Columns in the dataset (Annual, Monthly, Quarterly) are visualized using their statistical values (Median, Standard Deviation, Sum).

There are 4 types of visualizations used:

1. Map: the data is spatial, a map was used to make the data comprehensive. The bubble size directly corresponds to the magnitude of the value being presented on the map.

2. Pie charts are used to show distribution of values as percentages of a whole
- Rainfall for state as a percentage of rainfall of India
- Rainfall for a quarter as a percentage of annual rainfall

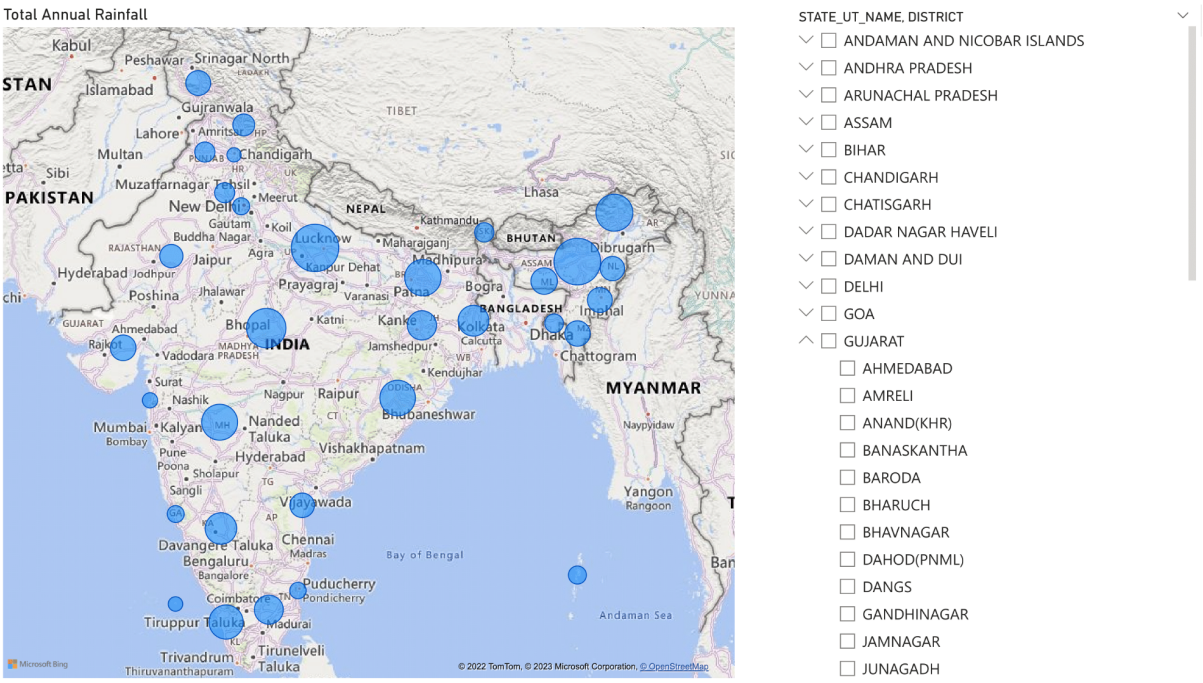
3. Column charts: Direct comparison of values takes place, can easily find the states with the greatest/ smallest values.

4. Flowchart: Explains the median rainfall for districts and states. Visual understanding of the relationship between columns.

Findings

The below visualizations enable a comprehensive understanding of a large dataset about the Rainfall in India.The data is spatial in nature as it has a location component.
Although the inferences mentioned in the report are largely based upon the attributes which are graphically represented, it must be noted that using the tooltips feature of PowerBI servers to show additional details about the represented data.

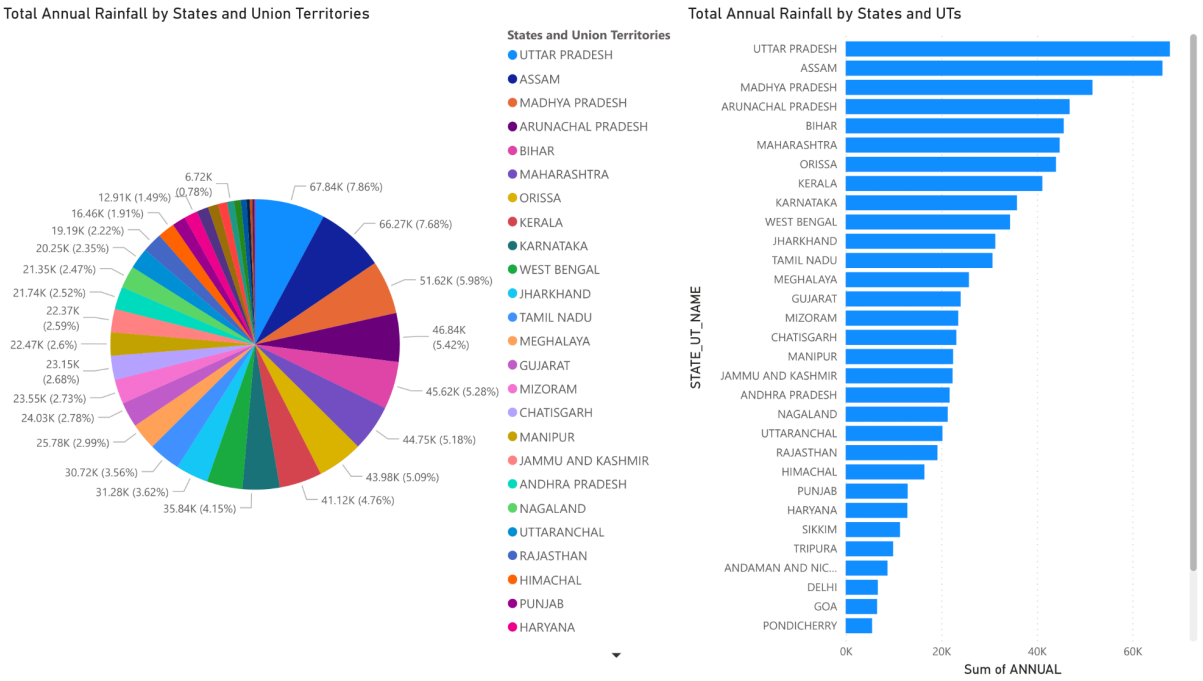
Figure1:



The map depicts the total annual rainfall for each state over the timespan of 50 years (1951-2000).The size of the bubbles shows the amount of rainfall, larger bubbles for more rainfall. We can select districts/ states in the slicer on the left to obtain more specific data.

Inferences: North-eastern part of India receives the most amount of rainfall and the Northern part of India receives overall the least amount of rainfall.

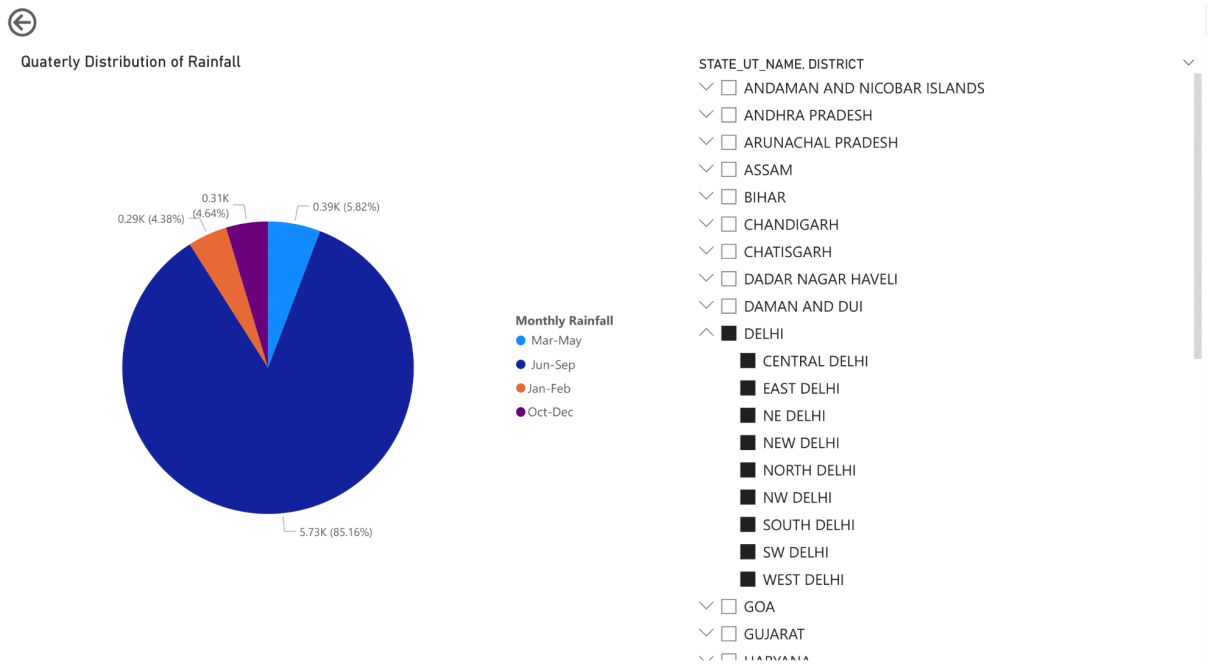
Figure 2:



The pie chart shows the state wise distribution of rainfall for 1951-2000: states with the highest percentage receive the most rainfall. States with the lowest percentage get the least rainfall.
The horizontal bar graph shows the ranking and compares the total annual rainfall received by a state between 1951-2000. We can compare each state's annual rainfall better.

Inferences: Overall the states of India receive somewhat of an equitable distribution of rainfall, except for a few states at the extreme ends of the spectrum.
UP receives 7.86% of the total rainfall received by India which is twice of what is received by Jharkhand.
Uttar Pradesh and Assam receive the heaviest annual rainfall, using this information we can conclude these states are the most prone to floods.
Telangana and Rajasthan have the least rainfall and are thus most prone to droughts.

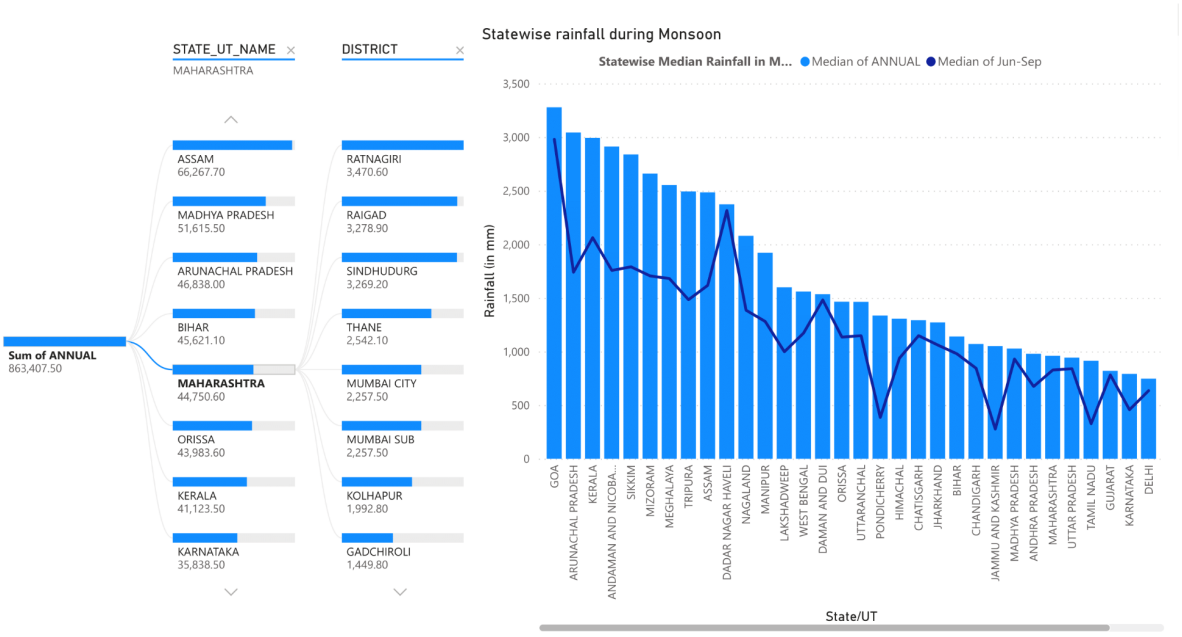
Figure 3:



This pie chart allows us to see the quarterly distribution of rainfall over the 50 years. We can select the state and district for which we want to know the quarterly distribution, using the slicer on the right hand side.

Inferences: In the pie chart for Delhi, we see that the majority of rainfall is received in June to September, which is monsoon. This tells us that there is a need for rain water harvesting to avoid water shortage in the summer/ other seasons.

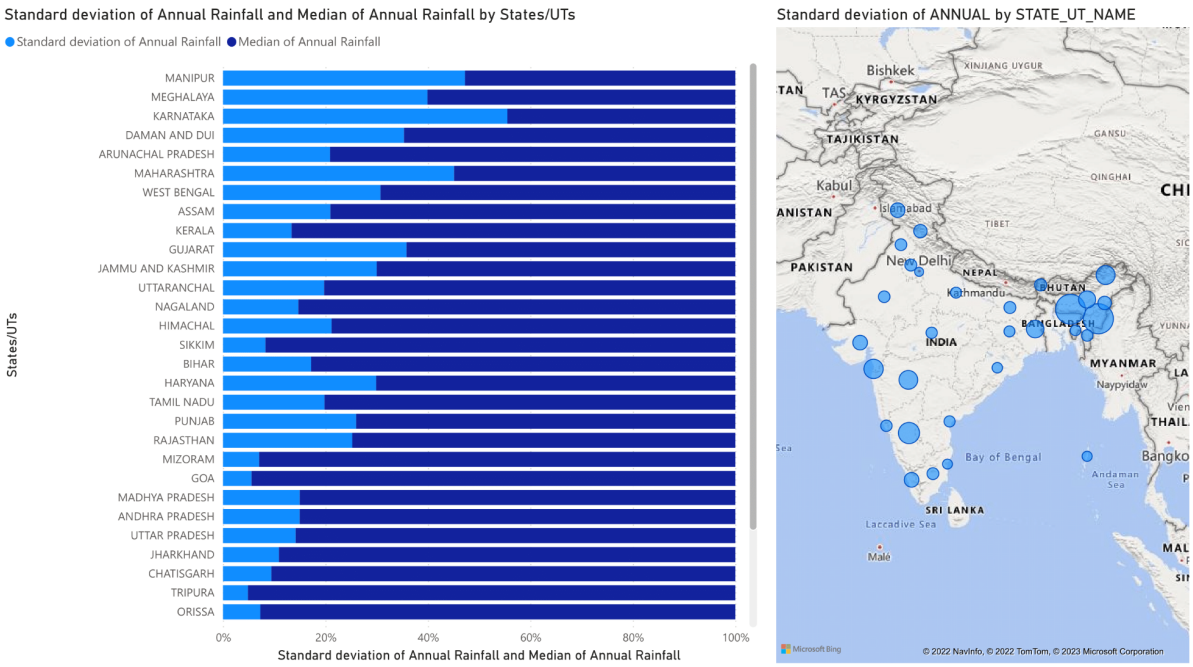
Figure 4:



The figure on the right serves as a data map. The flowchart provides a clear picture of the states and their division in districts. We can select the district/state, to see more information about it on the graph on the right.
The graph depicts the median annual rainfall for a particular state and the dark blue line shows the median rainfall for the monsoon (June to September). We get a fair estimate of what part of the rainfall is received during monsoon. The median shows the center of the distribution i.e rainfall for an average possible year. Median is a suitable value to be considered as the data is skewed.

Inferences: The flowchart shows Thane and Mumbai City get an equivalent amount of rainfall annually.Ratnagiri gets the maximum rainfall annually in Maharashtra.
The graph shows states like Pondicherry and Jammu and Kashmir get a substantial amount of rainfall not only during monsoon but also during other parts of the year.
Madhya Pradesh and Gujarat receive most of the rainfall during monsoon, thus they require rain water harvesting.
In terms of median annual rainfall Goa is the highest, this can be attributed to it being a coastal state.

Figure 5:



The graph on the left shows the Standard Deviation as a percentage of the median rainfall received by a particular state. The states where the Standard Deviation of rainfall is higher the amount of rainfall has more fluctuations.
On the map on the right the bubble size directly correlates to the Standard Deviation in the rainfall received by the state.

Inference: Manipur and Karnataka have high Standard Deviation thus weather prediction is more complex, they are more prone to floods/droughts or water shortages. There is also a negative impact on agricultural produce as most crops require a certain optimum amount of rainfall.

Conclusion

Analysis and exploration of the overall Rainfall data of 50 years has provided some valuable insights. These insights are what is required to take data driven action leading to data driven governance. The government and other bodies can form better and more effective policies if they take information provided by the data analysis.
Droughts and floods can be thus predicted, casualties and damage can be minimized. Better resource allocation can be done during natural calamities. Rainwater Harvesting and water conservation policies can be formed. Rainfall also directly impacts agriculture and availability of drinking water.
Thus,we can conclude that it is essential that data is analyzed and visualized effectively for enabling the overall betterment of society.

References

https://data.gov.in/resource/district-rainfall-normal-mm-monthly-seasonal-and-annual-data-period-1951-2000

https://www.iosrjournals.org/iosr-jestft/papers/vol11-issue%203/Version-3/B1103030719.pdf