

# Visualizing Rainfall Data and its correlation with temperature

## Big Data(CS661) Project Proposal

### 1 Introduction

Rainfall data refers to the rain that falls over a specific region. It is measured using rain gauges or radar systems, and is commonly measured in millimeters (mm) per unit area. It provides us useful information about the spatial (across space) and temporal (across time) distribution of rainfall. The data can be used to compare rainfall across various regions in the same time span. It can also be used to assess risk of floods or droughts. It is an integral part of models which assess the weather data like pressure, temperature etc and predict the climate for future. The data can also be used to identify the patterns and change in patterns of the rainfall.

Visualizing rainfall data can give a clear and concise picture of the rainfall which in turns helps to take various decision related to agriculture and other activities that are affected by rainfall. Other environmental variables such as temperature, pressure, humidity, wind-speed with rainfall data can aid a decision-maker to make informed decisions. Visualization can help to monitor the climate change and assess risks associated with it.

### 2 Data Description

There are two datasets to be used in the project. These datasets are taken from kaggle website [1] :

1. **Monthly Rainfall Data of India of a century** : It contains the amount rainfall in district of all the states in the India per year per month. The dataset consists of records of nearly a century from 1901 to 2002. The data is taken from kaggle [4], and it contains data in csv and excel form.
2. **Temperatures of INDIA** : It contains the temperature (in celsius) in India from 1901 to 2017. It is taken from an indian government website[2]. The data is present in the kaggle[3] in csv format.

### 3 Tasks and Expected Result

We will make a interactive web interface to visualize both the dataset after processing of the data. These are the tasks with the result expected from them in a listed manner :

1. **Data cleaning and processing** : It will involve data aggregation to aggregate data across months across district of the same state for the rainfall data set. The data will be transformed to suit the objective.
2. **Rainfall in whole country** : The task is to make a whole map of India showing the rainfall across all the states kind of heatmap. Filtering would be possible by selecting the time span and rainfall level across the states.

3. **Statewise rainfall** : The task is to get a zoomed in state image showing rainfall across various districts in it. Filtering would be possible by selecting the time span and rainfall level across the states.
4. **Trends** : Task is to make line charts showing average monthly rainfall, year wise cumulative rainfall, seasonal rainfall (rainfall trends across various seasons) to visualize the trends of rainfall.
5. **Correlation between temperature and rainfall data** : Line charts to see the correlation between rainfall and temperature data.

### 4 Future Tasks

If time permits we wish to do a few extra tasks as well :

- Adding more variables with temperature like pressure, humidity, wind-speed to aid in decision making and to get more trends with respect to combination of these variables.
- Predicting rainfall using past rainfall and other environmental variables data using machine learning.

### 5 Division Of Work

We have divided work among ourselves in the following fashion. With working on their specific tasks, team members will aid in other tasks as well whenever required.

1. Abhay Kumar Dwivedi (22111001) : Task 2 & Task 4
2. Hrithik Lal (22111027) : Task 1 & Task 3
3. Sanket Kale (22111052) : Task 3 & Task 5
4. Saqeeb (22111053) : Task 1 & Task 4
5. Shubham Rathore (22111054) : Task 2 & Task 5

### References

- [1] Kaggle: Your machine learning and data science community. <https://www.kaggle.com/>. Accessed: 2023-3-8.
- [2] Open government data (OGD) platform india. <https://data.gov.in/catalog/all-india-seasonal-and-annual-minmax-temperature-series>, January 2022. Accessed: 2023-3-8.
- [3] Venkatesh Garnepudi. Temperatures of INDIA, February 2019.
- [4] Ravi Sane. Monthly rainfall data of india of a century, August 2020.