



Impact of Weather Variables and Roles of Human Population in Vector Borne Diseases

Hello!

We are aspiring Softies

17pw13 - Harisai Pravin

17pw32 - Sathwik Ch



In God we trust, all others bring data



When you Visualize, then you materialize.



*Without Data you're just another person
with opinion*

State Wise Rainfall in India

Dataset from data.gov.in

https://data.gov.in/catalog/district-rainfall-normal-mm-monthly-seasonal-and-annual-data-period-1951-2000?filters%5Bfield_catalog_reference%5D=87155&format=json&offset=0&limit=6&sort%5Bcreated%5D=desc

The data refers to details on climatological normals of district wise rainfall (in mm) calculated with the data for the period 1951-2000.

Multiple States with it's rainfall each year.

State Wise Climate in India

Dataset Description

The dataset was used with the help of the worldweatheronline.com API and the `wwo_hist` package.

The datasets contain hourly weather data from 01-01-2009 to 01-01-2020. The data of each city is for more than 10 years.

This data was used to visualize climate changes in different cities and how it effects on other datasets.

Air Pollution

Dataset Description

The dataset contains air quality data and AQI (Air Quality Index) at hourly and daily level of various stations across multiple cities in India.

The data has been made publicly available by the Central Pollution Control Board: <https://cpcb.nic.in/> which is the official portal of Government of India. They also have a real-time monitoring app: https://app.cpcbcr.com/AQI_India/

Malaria Cases in India

Dataset from National Vector Borne Disease control programme (NVDCDP)

www.nvbdc.gov.in

<https://nvbdc.gov.in/Doc/mal-situation-Jan13.pdf>

The dataset has malaria cases, deaths for 5 contiguous years from 2009 to 2015.

Dengue Cases in India

Dataset from National Vector Borne Disease control programme (NVDCDP)

www.nvbdcp.gov.in

<https://nvbdcp.gov.in/index1.php?lang=1&level=1&sublinkid=5776&lid=3690>

The dataset has dengue cases, deaths for 5 contiguous years from 2009 to 2015.

It was highly noisy and required full clean.

Dataset extracted from pdf in annual report.

1.

Modelling and Data Analysis

Understanding of the raw data

Cleaning



NULL value removal

The NULL values in data had to be taken care of as it would affect when building our final model.



Visualization & Analysis

Visualization plays a major part in analysing the data. Every year and state performance has been visualized here.



Feature Selection

Although no embedded or filter approach is used yet, We manually understood the data and removed the unnecessary ones.
Normalization check



Correlation

To get in depth knowledge about the relationship between feature and to check if PCA should be used or not.



Grouping data & Sampling

With so much of information given, there is a need to club them all together. Here a kind of under sampling was used by taking their mean.



Data Cleanup

The noisy data had to be cleared to get a cleared dataset that can be used for processing.

Challenges Faced

TYPE of vis.

With lot of Datavis in the market choosing the best was challenge.

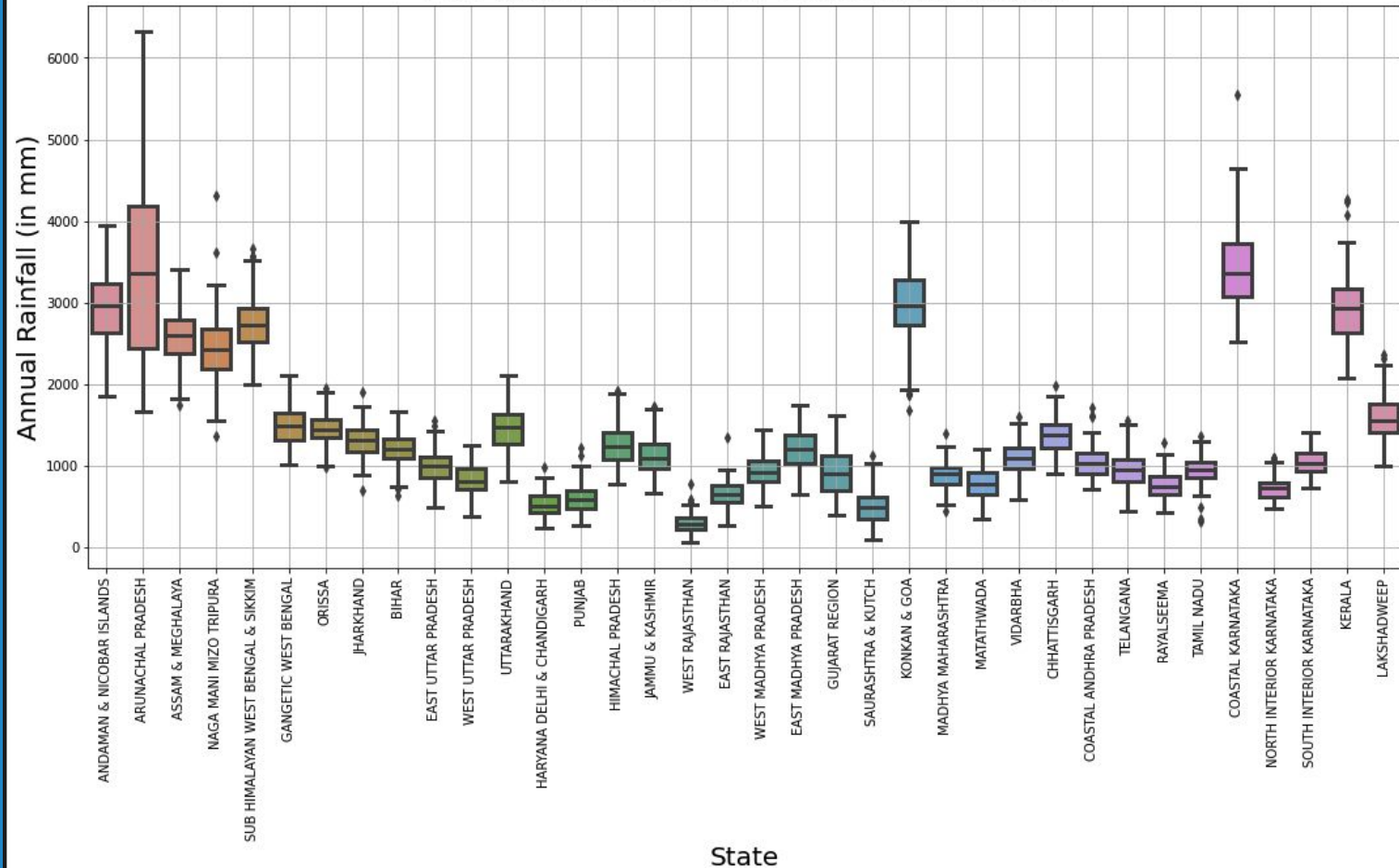
Heterogeneous dataset

Since the dataset is collected from lot of sources, each of the dataset is different from the rest, even the basic names are changed.

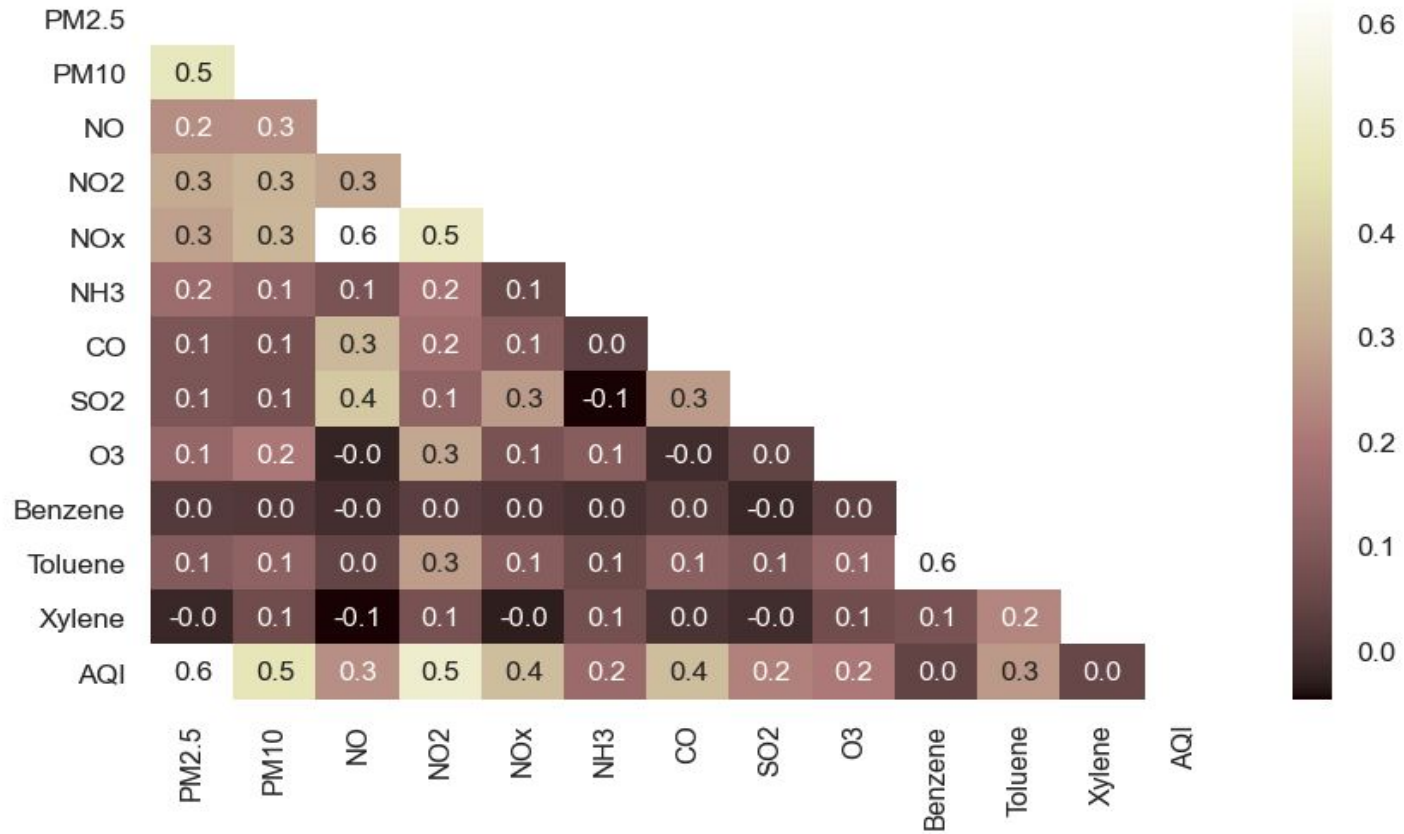


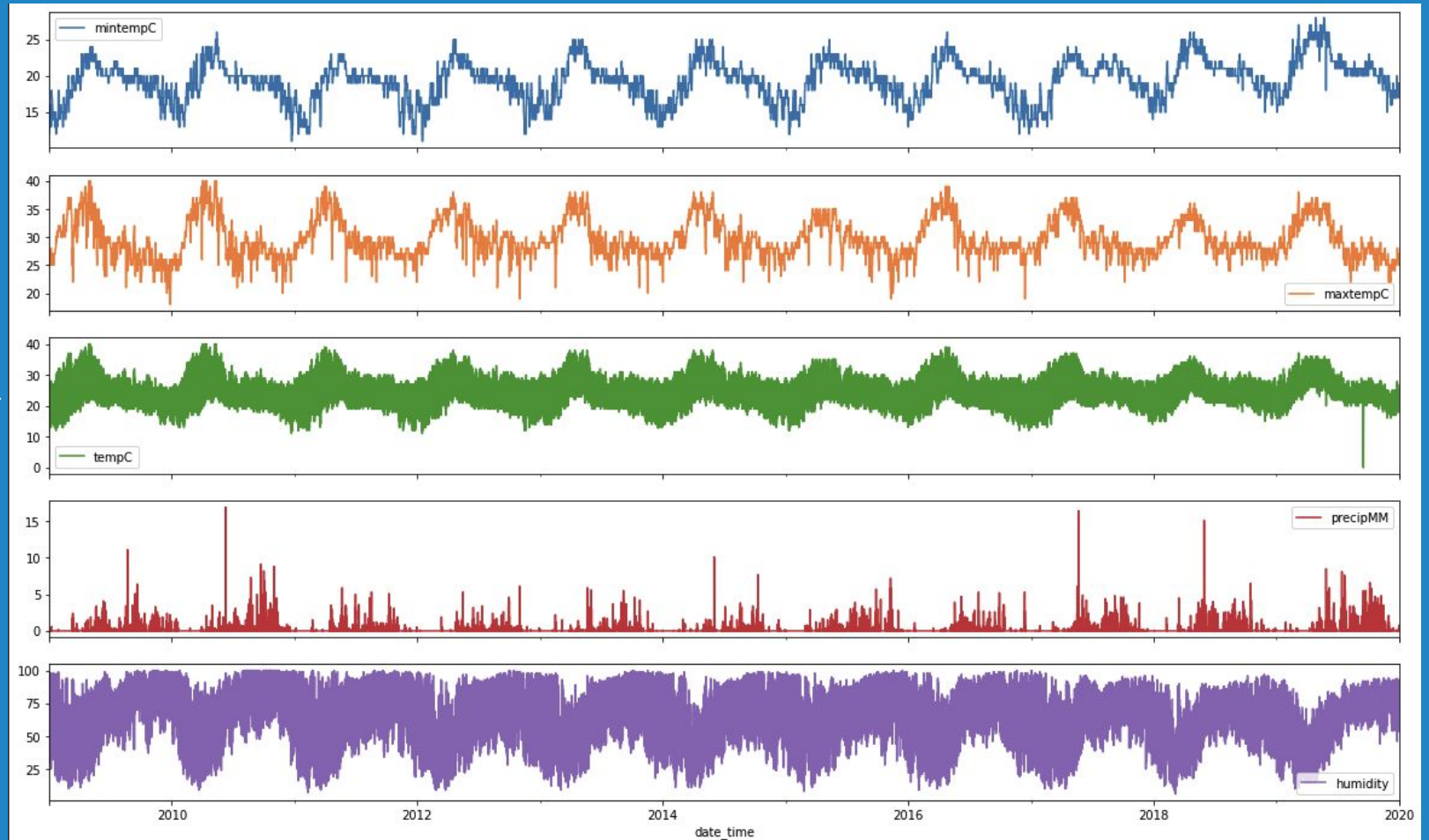
Glimpse of what we
achieved

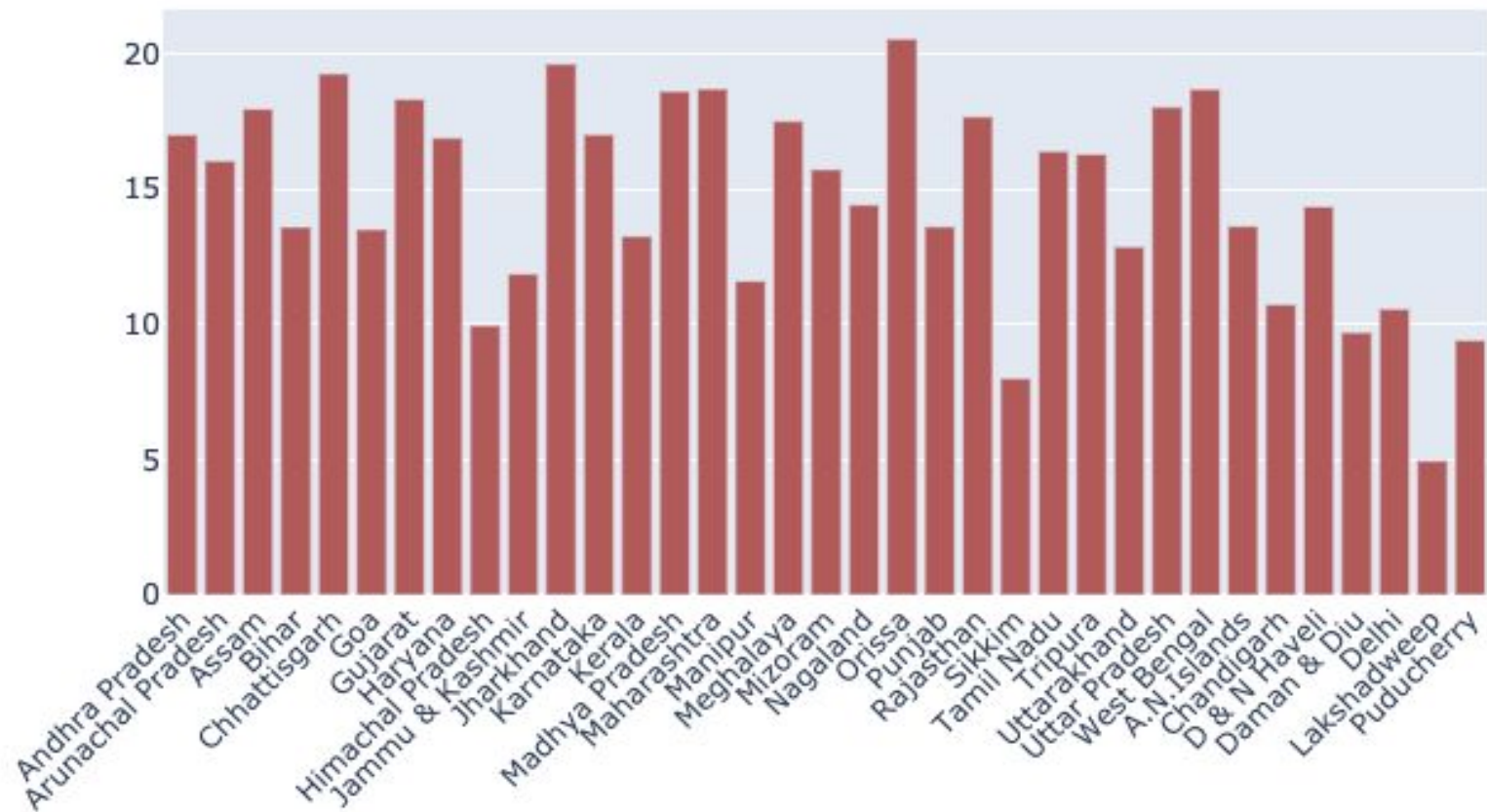
Annual Rainfall in States of India



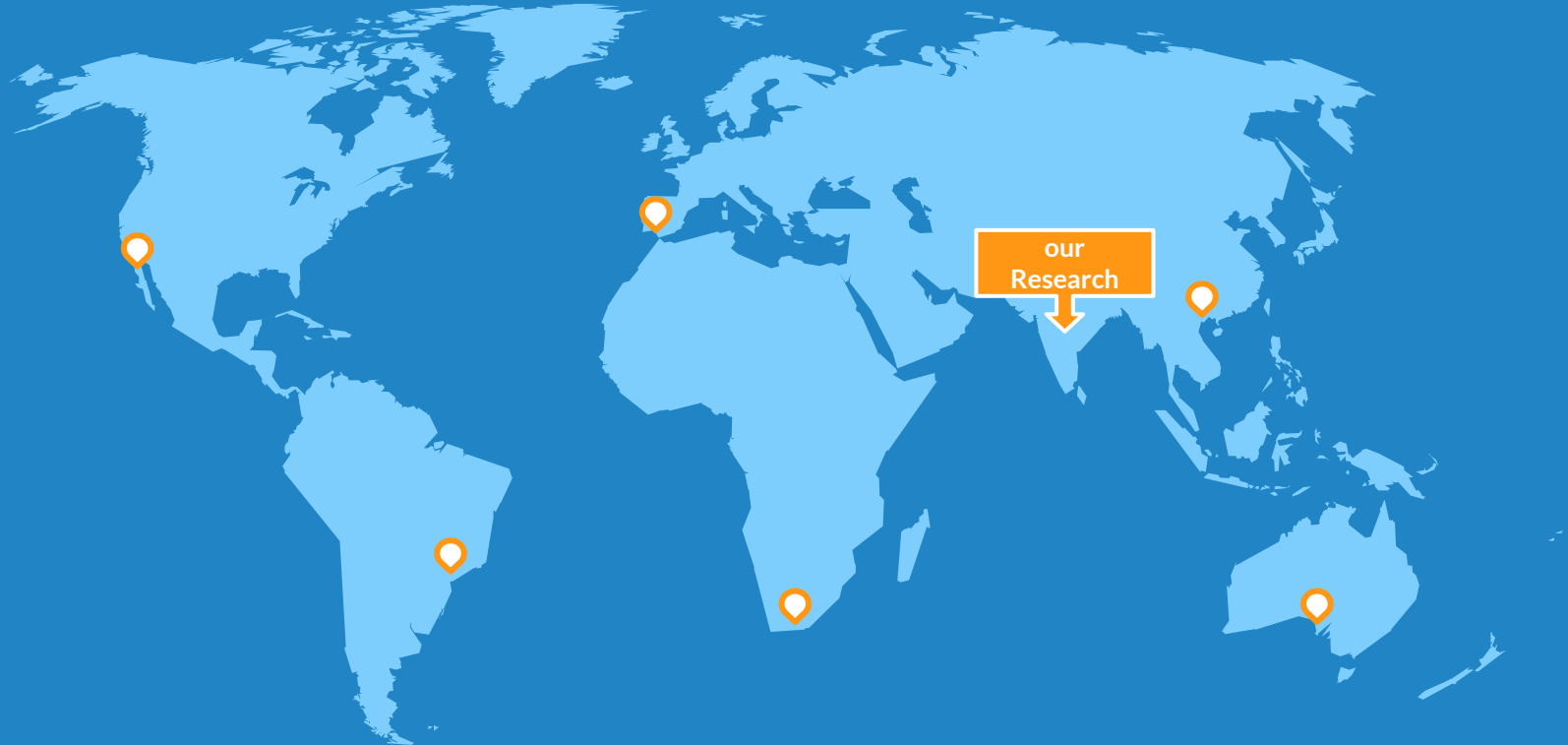
Correlation Analysis







Geospatial data For India on dev.



THANK YOU

Creativity is allowing yourself to make mistakes. Design is knowing which ones to keep.

<Good Design is Good Business/>