Engineering Economics

GEE CODING PROJECT

Table of Contents

ABOUT THE DATASET

ERA5 MONTHLY AGGREGATES - LATEST CLIMATE REANALYSIS PRODUCED BY ECMWF / COPERNICUS CLIMATE CHANGE SERVICE

Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset.ERA5 MONTHLY provides aggregated values for each month for seven ERA5 climate reanalysis parameters: 2m air temperature, 2m dewpoint temperature, total precipitation, mean sea level pressure, surface pressure, 10m u-component of wind and 10m v-component of wind. Additionally, monthly minimum and maximum air temperature at 2m has been calculated based on the hourly 2m air temperature data. Monthly total precipitation values are given as monthly sums. All other parameters are provided as monthly averages.

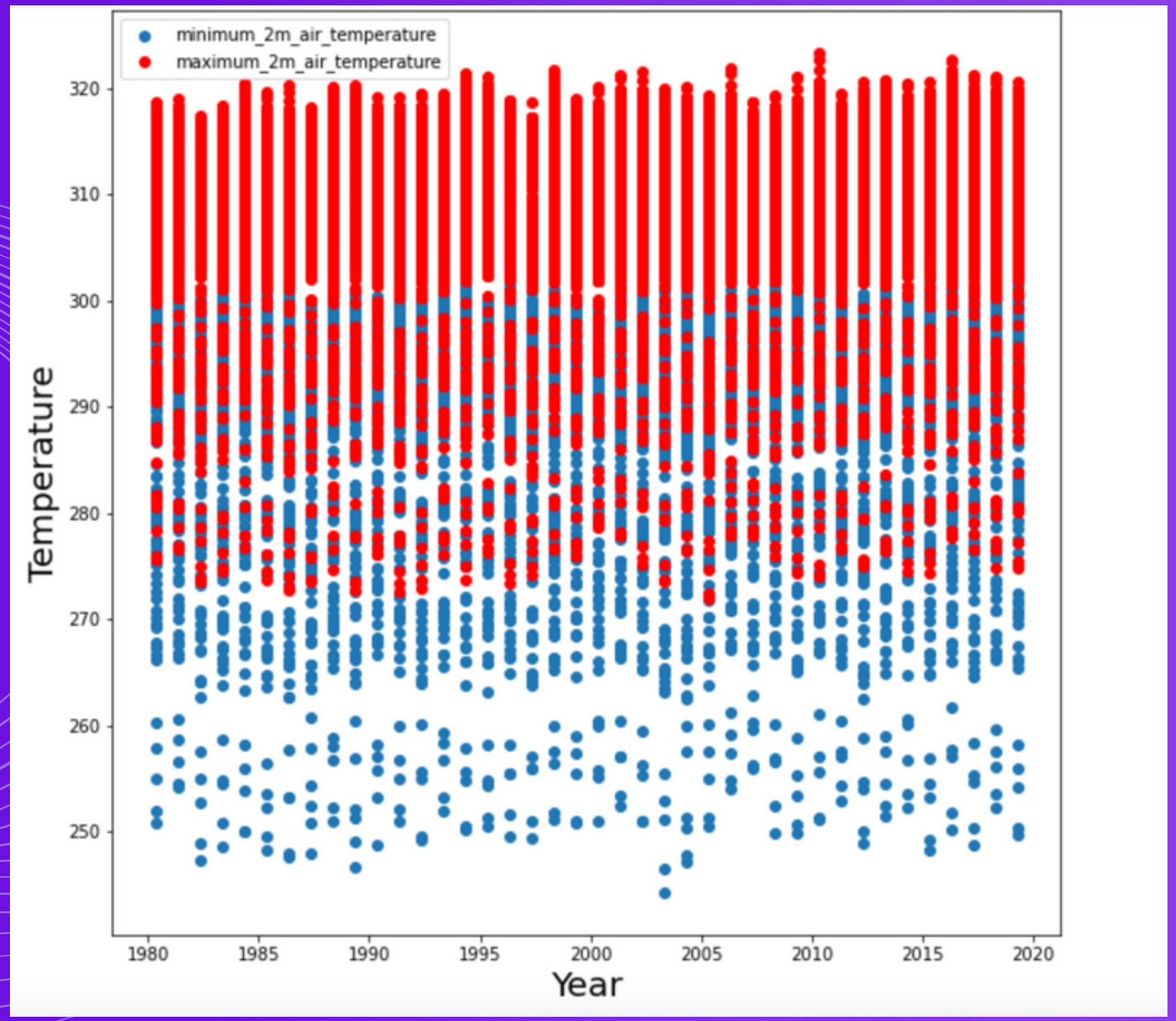
	DIST91_ID	NAME	STATE_UT	date	maximum_2m_air_temperature	minimum_2m_air_temperature
0	444.0	DADRA_&_NAGAR_HAVELI	DADRA_&_NAGAR_HAVELI	1980-01-01	303.363086	287.336563
1	446.0	DIU	DAMAN_&_DIU	1980-01-01	300.782562	290.148163
2	445.0	DAMAN	DAMAN_&_DIU	1980-01-01	299.633728	292.205414
3	447.0	DELHI	DELHI	1980-01-01	298.071065	274.780852
4	443.0	CHANDIGARH	CHANDIGARH	1980-01-01	296.485575	275.245342

UNIVARIATE ANALYSIS

SCATTER PLOT

Scatter plot depicting the variation in minimum temperature and maximum temperature accross India for the month of may

There has only been very minute variations from the general trend. The tempertures have had a slight increase .



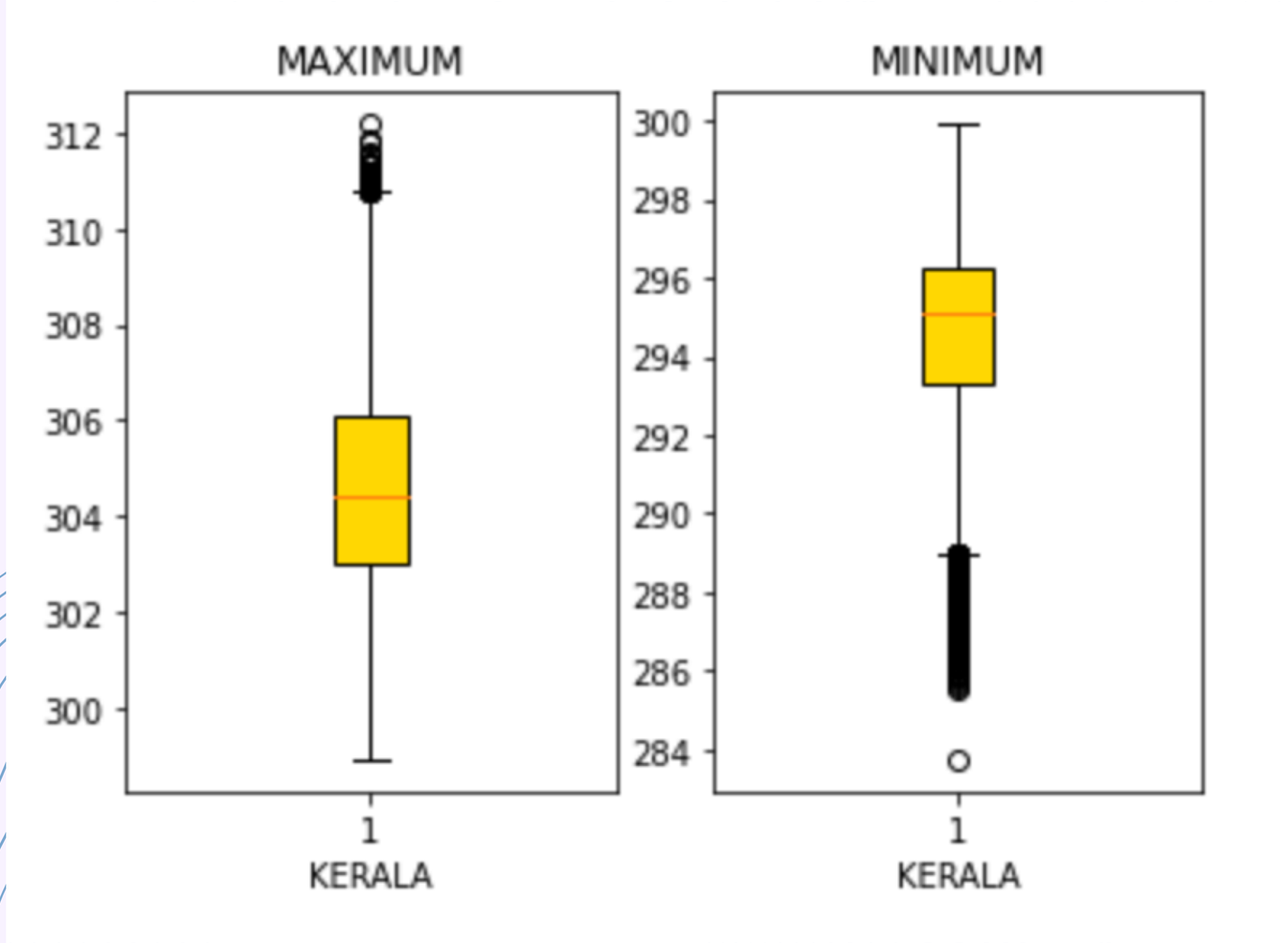
Haneem E M

BOX PLOT

The box plot shows the distribution of numerical data and skewness through displaying the data quartiles (or percentiles) and averages.

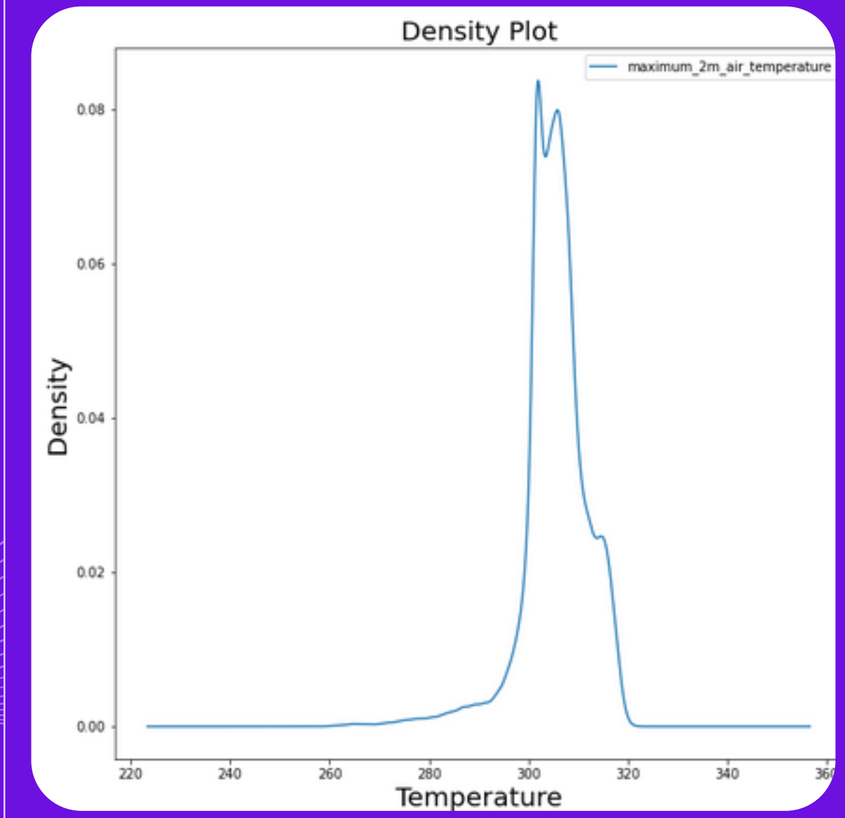
- The box covers the value from 25th percentile to 75th percentile
- The lower end of the box marks 25th percentile value of temperatures.
- The upper end of the box marks 75th percentile value of temperatures.
- The Red line inside the box gives the median of temperatures.
- The small horizontal lines marks the lowest and highest values.

Box plot of temperature values of state Kerala



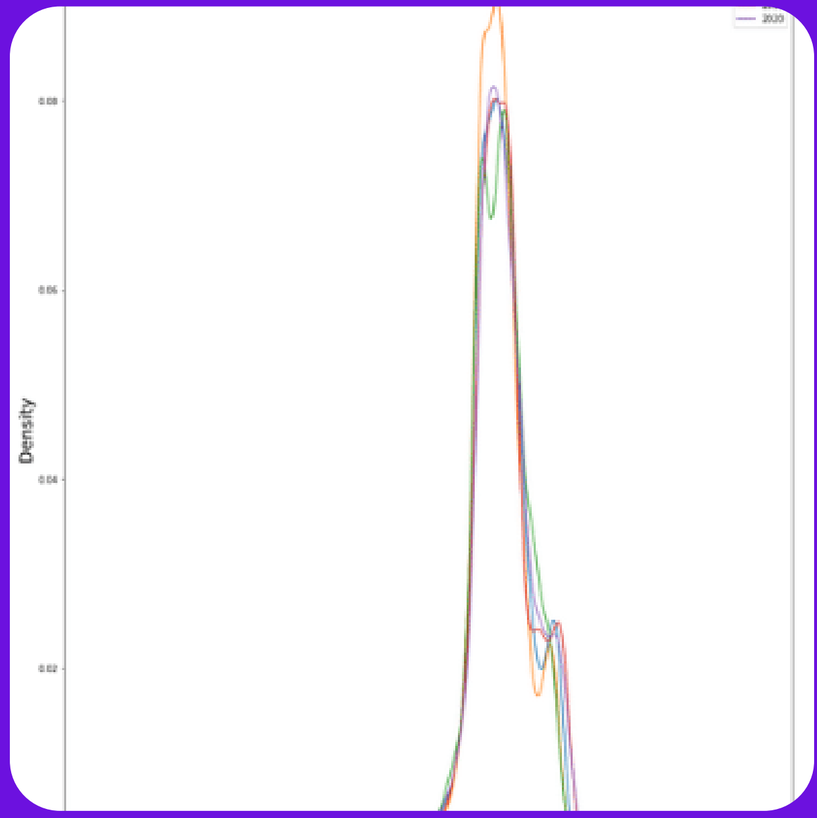
DENSITY GRAPH

Density plots are used to observe the distribution of a variable in a dataset. It plots the graph on a continuous interval or time-period. This is also known as Kernel density plot. Density plots are a variation of Histograms. ... The peaks of a Density Plot help display where values are concentrated over the interval.



BY TAKING ALL STATES

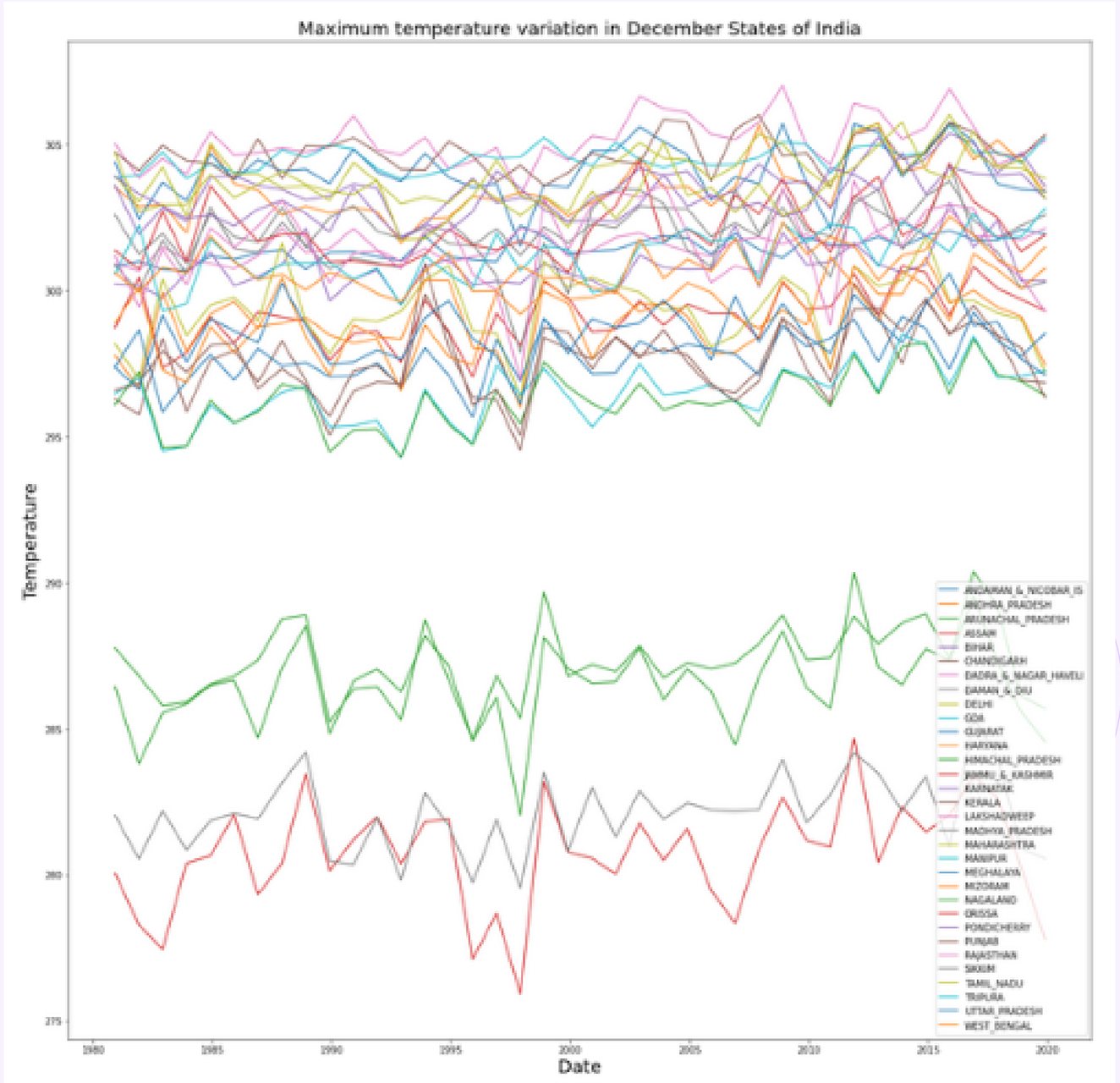
- On an average India's average maximum temperaturn is 300 K



BY YEAR WISE

- India's average temperature is 300 K
- There was a steep rise in states having 300 K during the year 1990

A Vishnusankar



LINE GRAPH

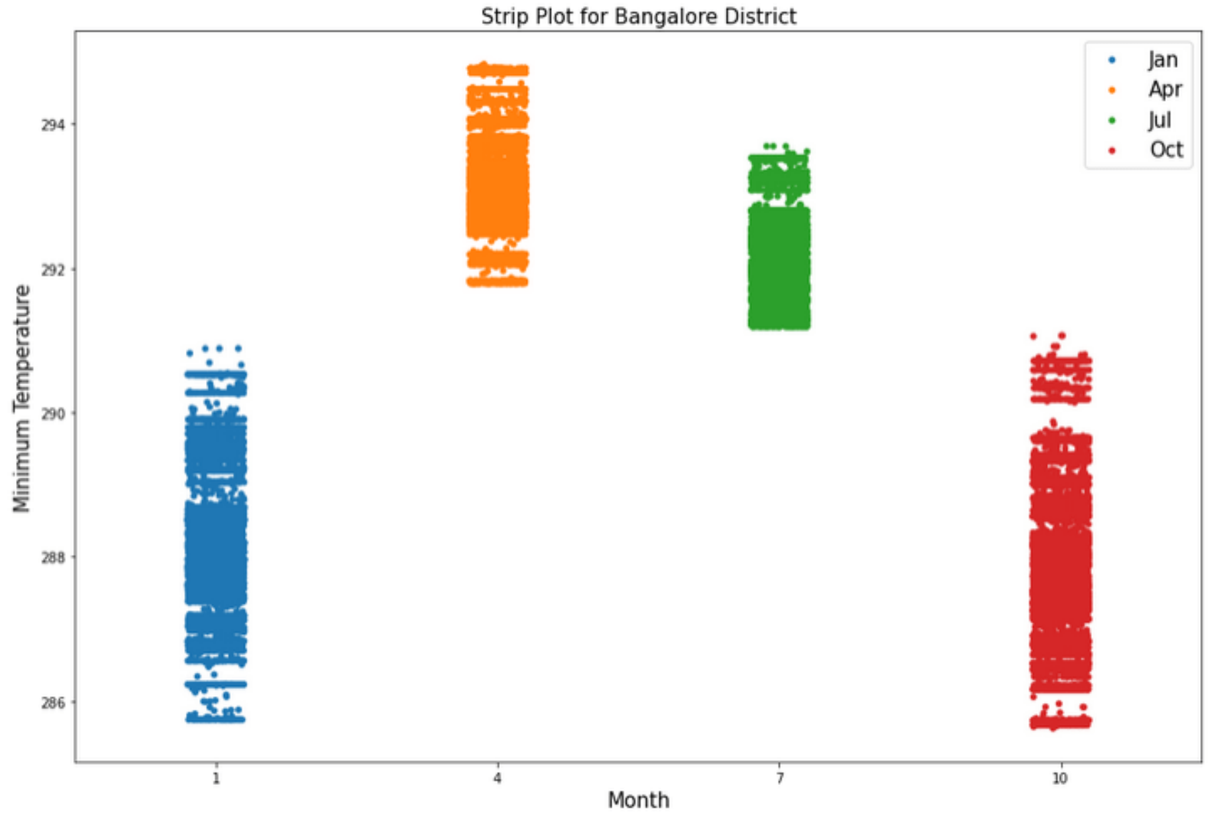
By plotting line graph of maximum temperature of all states in the month of December, we can see that :

- Himachal Pradesh and Jammu & Kashmir have the lowest temperature.
- Every year, the temperature curve is increasing
- Rest all states seems to have very high temperature

A Vishnusankar

Strip Plot

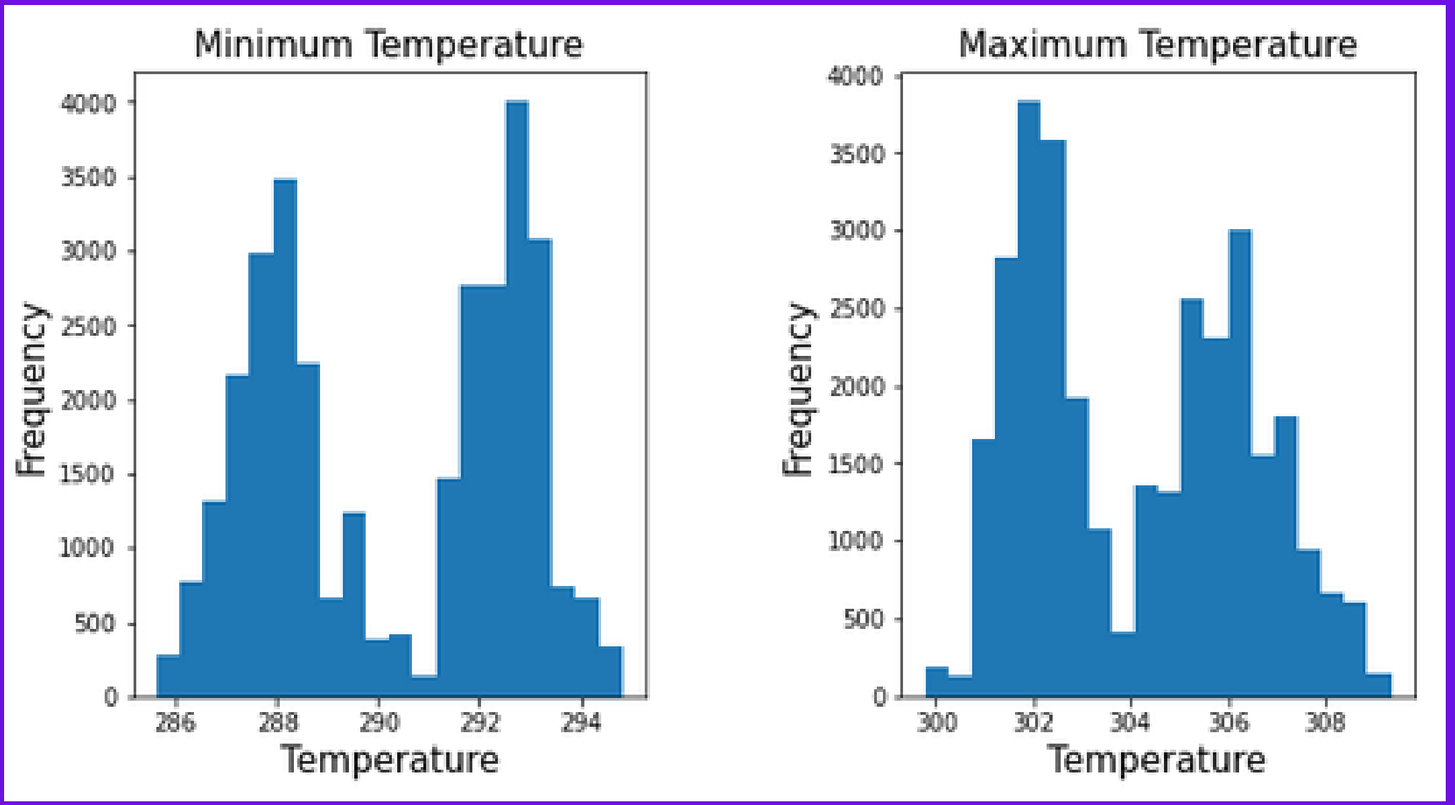
By analyzing bangalore dataset for 40 years,plotting strip plot , we can analyze that minimum temperatur is highest during the months of April and lowest during the months of January and October



A Vishnusankar

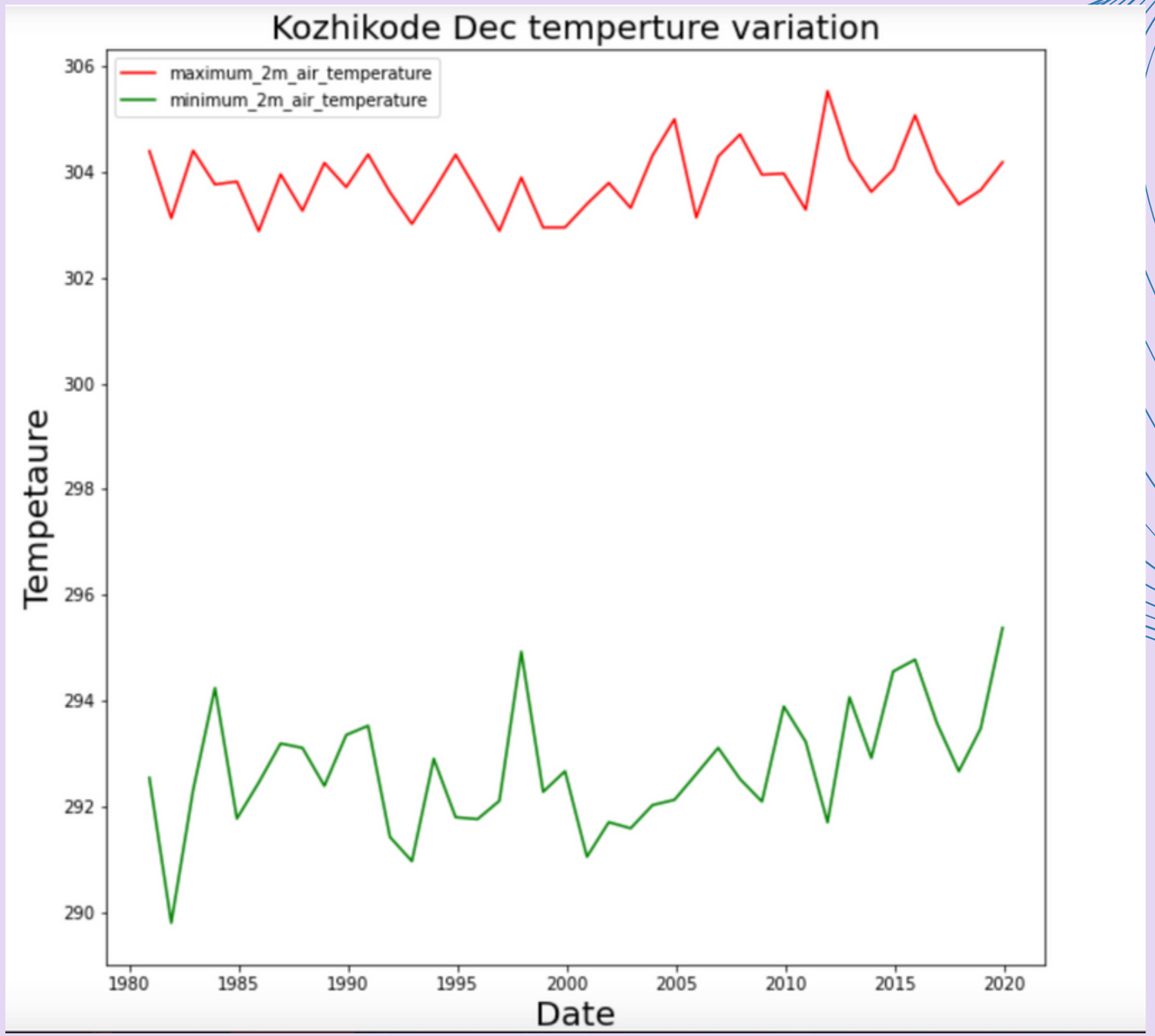
Histogram

Histogram is plotted for bangalore dataset to count the wards having same minimum and maximum temperature



By analysing the graph we can draw the conclusion that most wards are having minimum temperature of 293 K and maximum temperature of 302 K

A Vishnusankar



TEMPERATURE vs YEAR

This line plot shows minimum and maximum values of temperature for Kozhikode district from 1980 - 2020 for the month December

The value of Maximum temperature can be seen varying but Maximum temperature of 1980 and 2020 has mostly remained the same. But on the contrary the Minimum temperature value is rising and highly volatile

Haneem E M

BIVARIATE ANALYSIS

DATASETS ANALYSED

-> ERA5 Monthly aggregates - Latest climate reanalysis produced by ECMWF / Copernicus Climate Change Service

->MCD19A2.006: Terra & Aqua MAIAC Land Aerosol Optical Depth Daily 1km

Date Range : 2005 - 2019

Haneem E M

Temperature Dataset

STATE_UT	date	maximum_2m_air_temperature	minimum_2m_air_temperature
GOA	2005-01-01	303.652585	293.474574
GOA	2005-02-01	304.045929	293.718666
GOA	2005-03-01	303.967547	295.207613
GOA	2005-04-01	305.541901	296.849751
GOA	2005-05-01	305.040244	298.603507

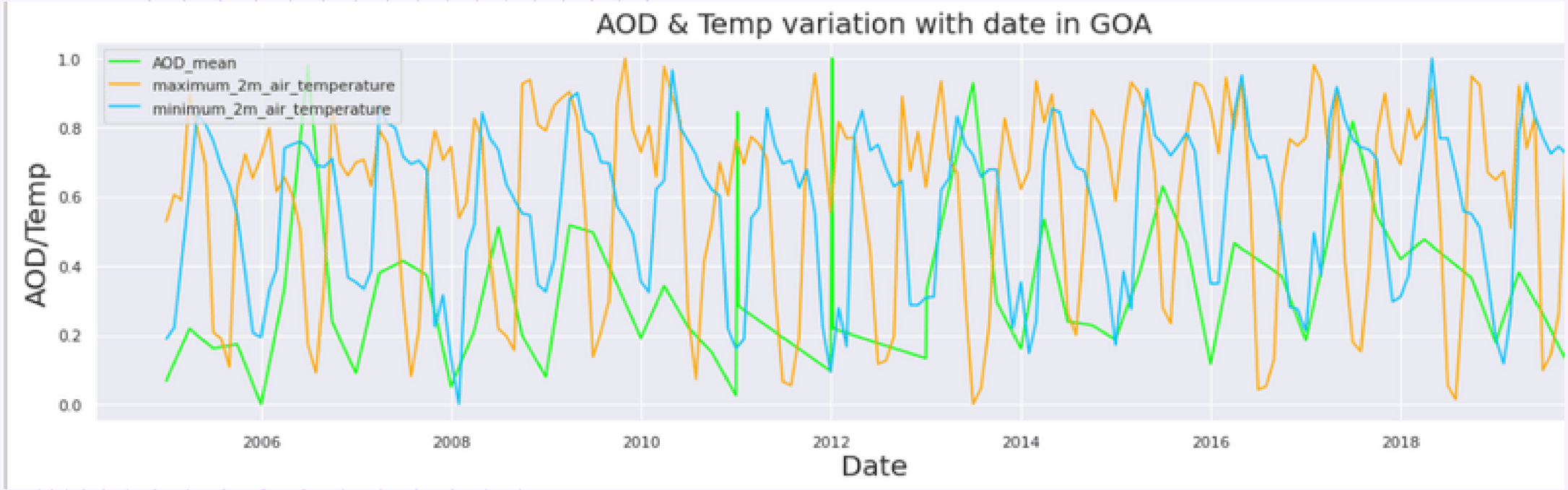
AOD Dataset

	STATE_UT	date	AOD_mean
533	GOA	2005-01-01	373.237458
534	GOA	2005-04-01	524.256071
535	GOA	2005-07-01	467.441401
536	GOA	2005-10-01	480.090491
537	GOA	2006-01-01	305.607830

State wise averaged datasets.

Line Plot of Temperature & AOD Variation in GOA

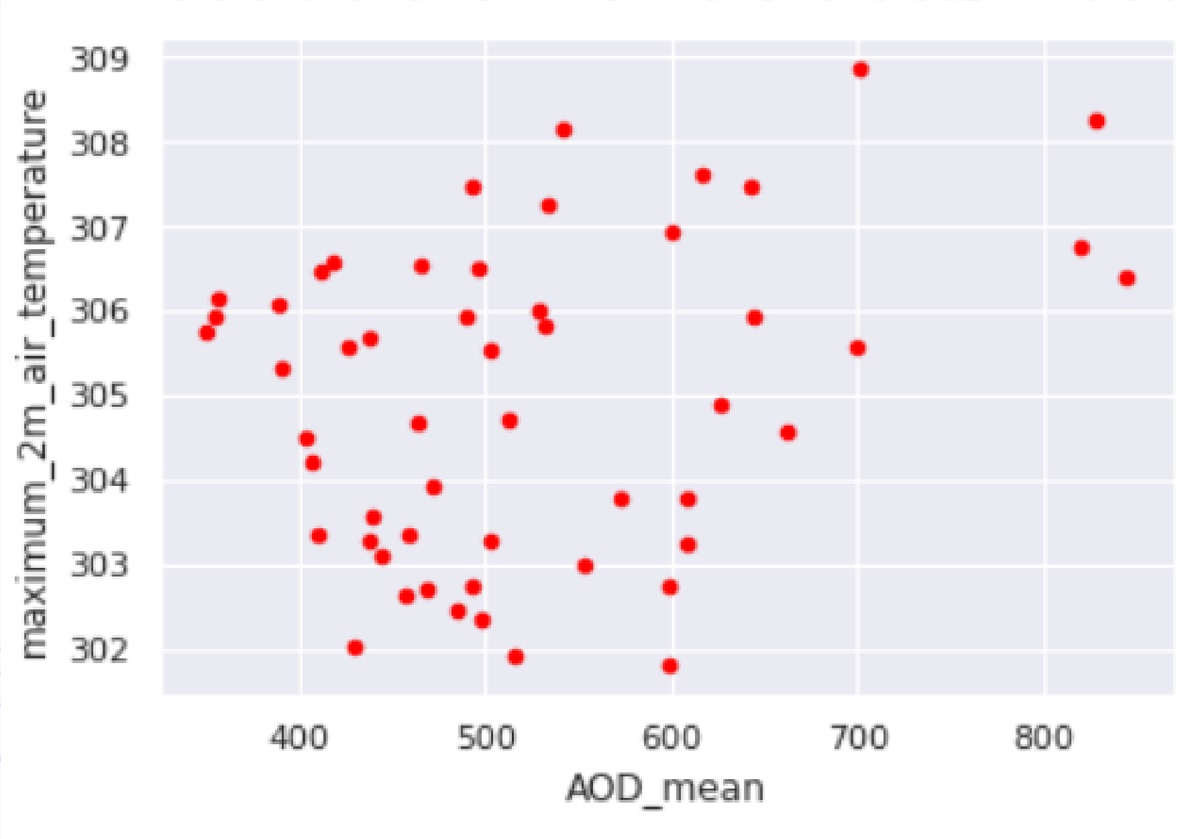
Data values of both datasets were scaled to range (0,1) to plot on single graph.. Could not derive any major conclusions from this graph



SCATTER PLOT
AOD VS MAX_TEMPERATURE

Both Datasets were merged together to plot this scatter plot

This graph shows a higher temperature on higher values of AOD



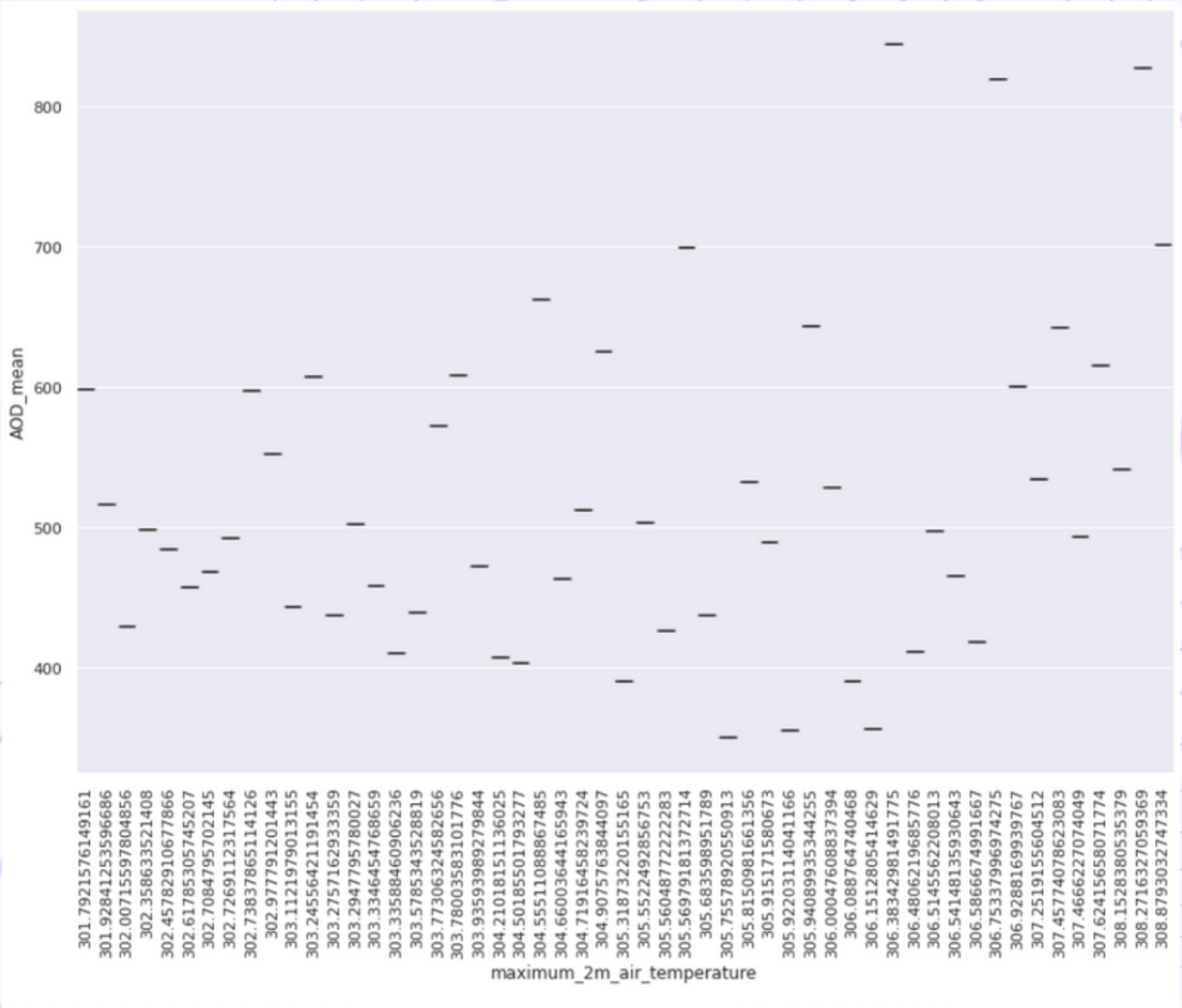
A Vishnusankar

BOX PLOT

AOD VS MAX_TEMPERATURE

Almost similar observation as on scatter plot can be observed here also

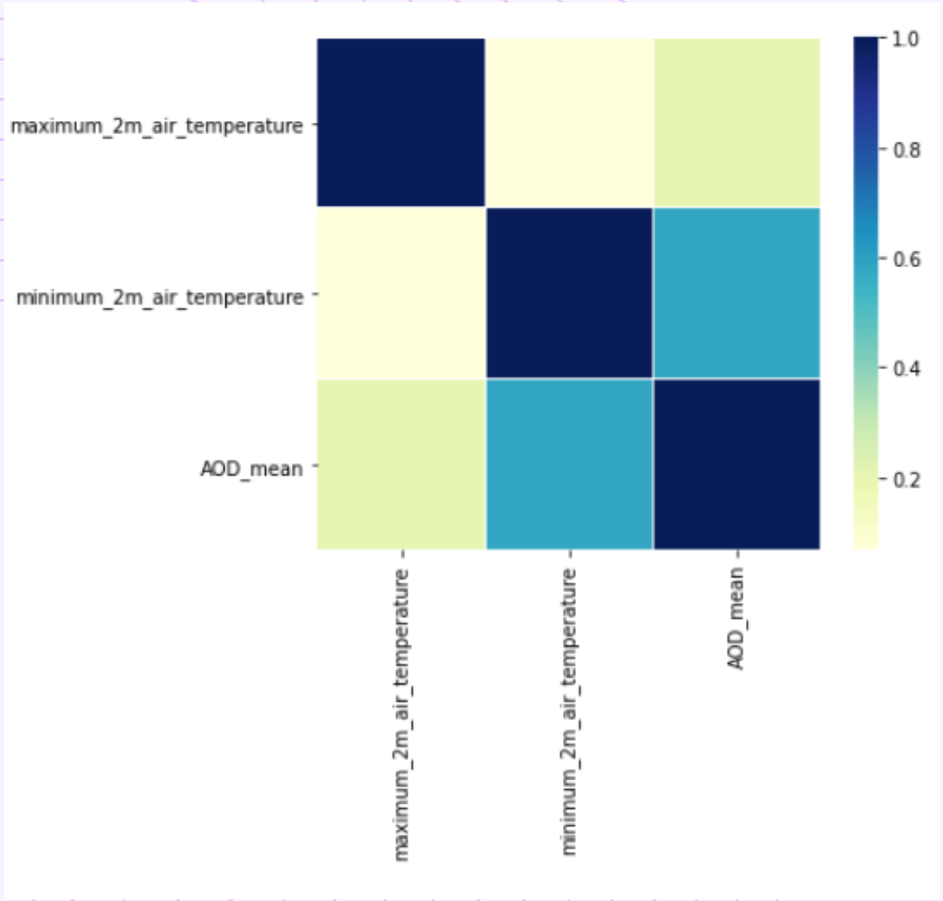
Higher temperatures correspond to higher AOD values



GEE CODING PROJECT 2020

Jaideep Kumar P M

CORRELATION MATRIX & HEATMAP



Correlation analysis confirms these datasets are not closely related.

maximum_2m_air_temperature	minimum_2m_air_temperature	AOD_mean
1.000000	0.068860	0.207547
0.068860	1.000000	0.582406
0.207547	0.582406	1.000000

Jaideep Kumar P M

MULTIVARIATE ANALYSIS

DATASETS ANALYSED

-> TEMP : ERA5 Monthly aggregates - Latest climate reanalysis produced by ECMWF / Copernicus Climate Change Service

-> AOD : MCD19A2.006: Terra & Aqua MAIAC Land Aerosol Optical Depth Daily 1km

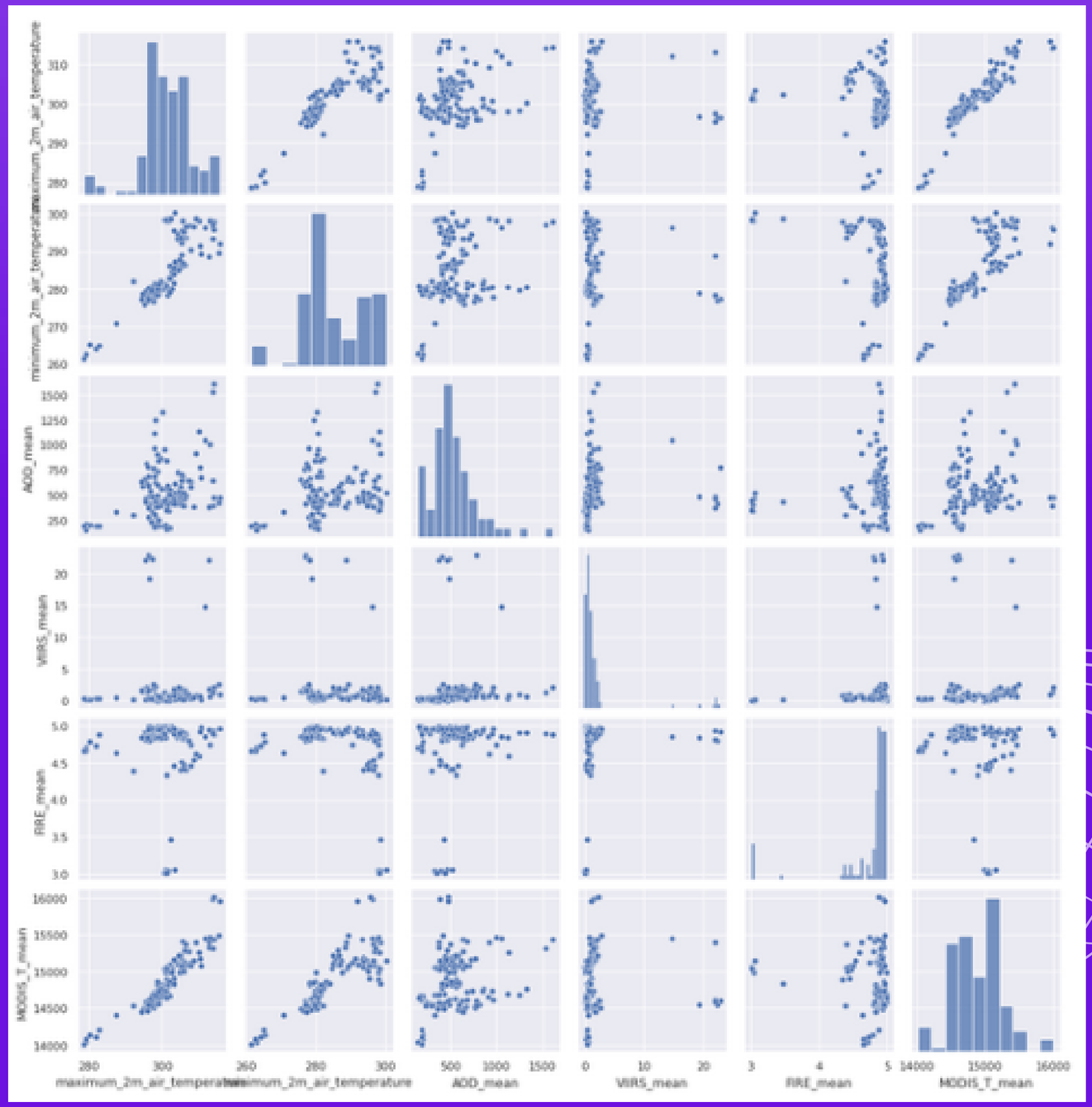
-> VIIRS : Stray Light Corrected Nighttime Day/Night Band Composites Version 1

-> MODIS : MOD16A2: MODIS Global Terrestrial Evapotranspiration 8-Day Global 1km

-> FIRE : MOD14A1.006: Terra Thermal Anomalies & Fire Daily Global 1km

Date Range : 2014 - 2019

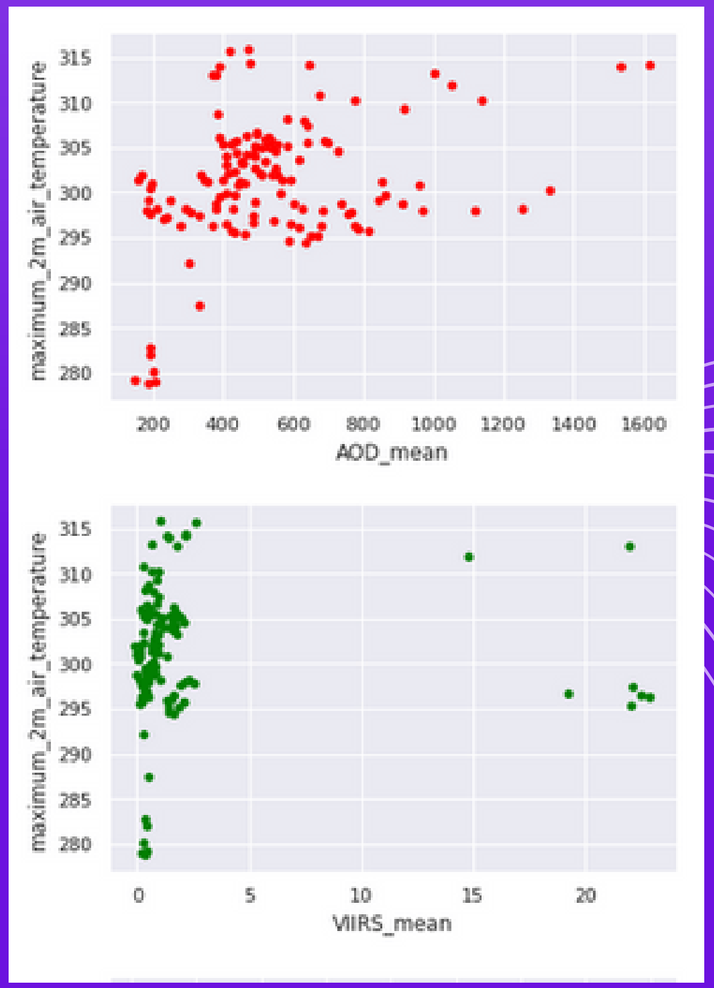
Jaideep Kumar P M

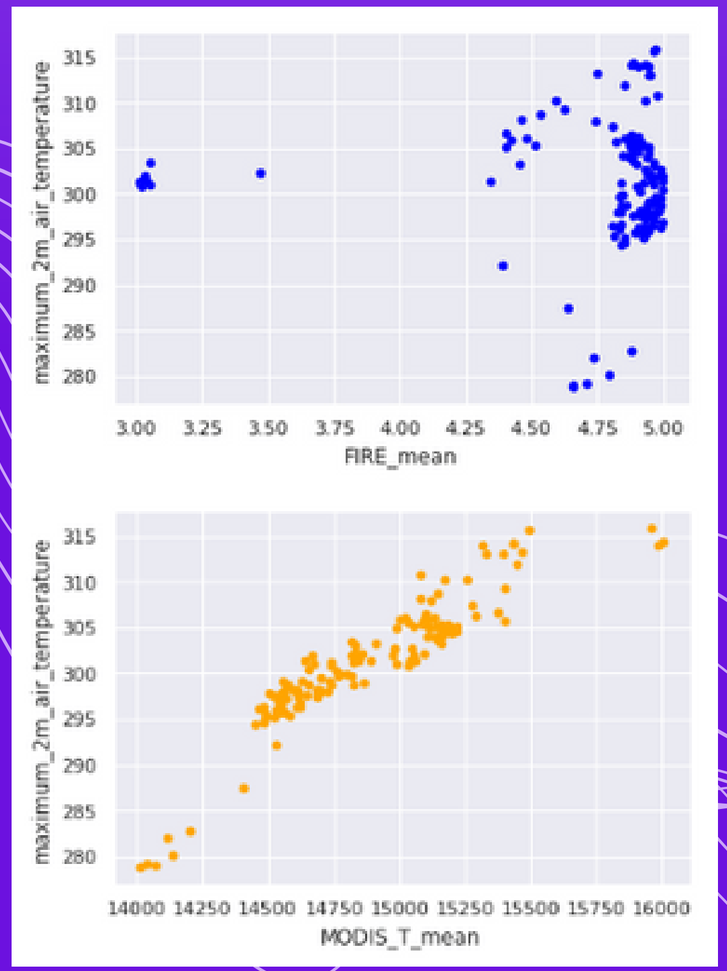


PAIR PLOT

All 5 Datasets were merged together to plot pairwise plot

Jaideep Kumar P M





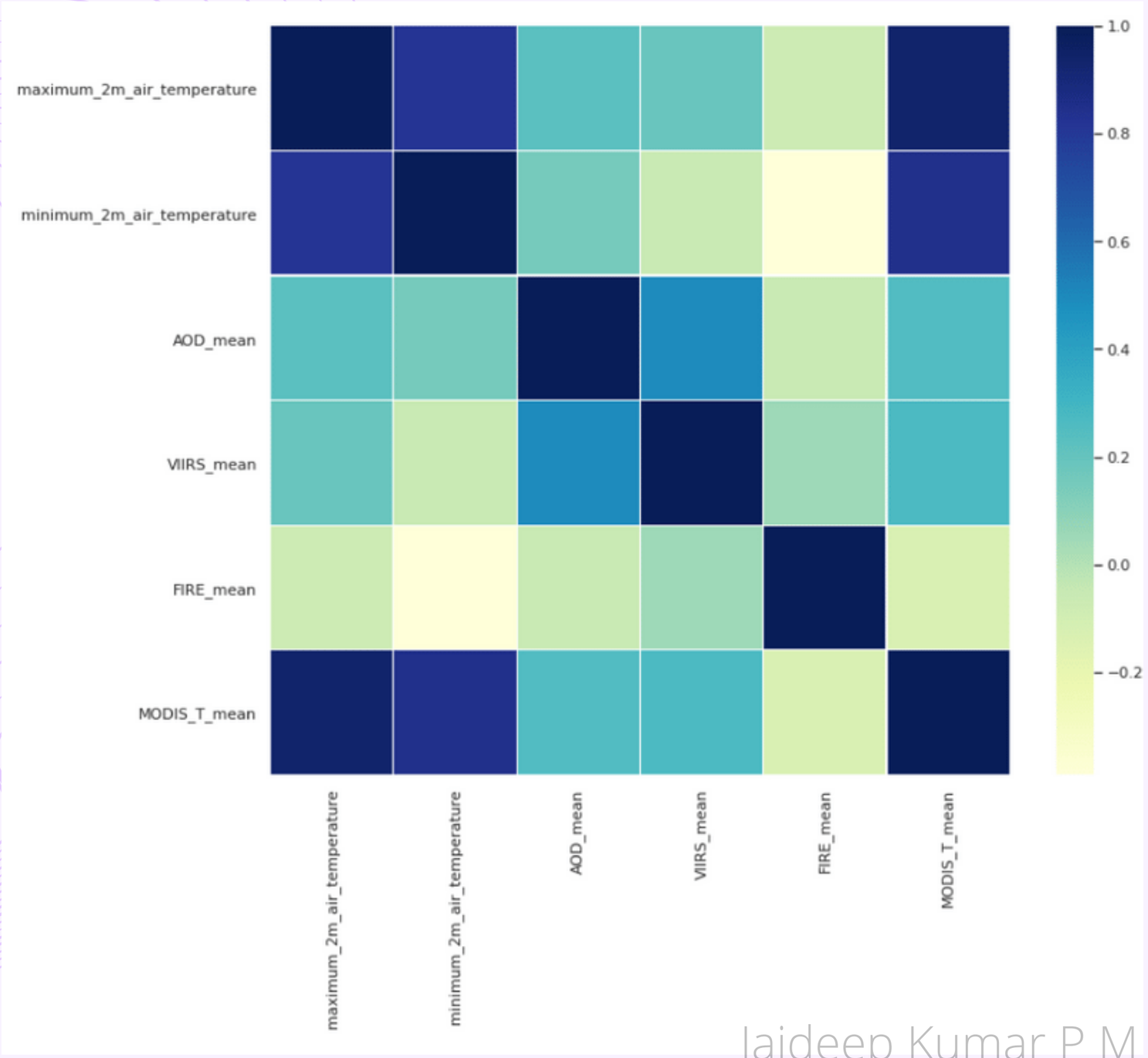
Pair wise plots of air_temperature with other datasets

Jaideep Kumar P M

CORRELATION ANALYSIS ON MULTIPLE DATASETS

Relation between five merged datasets is analysed using correlation matrix. Using the correlation matrix Heatmap is plotted

Variables that has highest correlation with ERA5 air temperature is MODIS_T which is the surface temperature variable which is predictable
Other ones having slight correlation are AOD and VIIRS(Night Light data - which is not reasonable)



Jaideep Kumar P M

Visualizing India & Bangalore

Adarsh Joshy

GEOPANDAS

GeoPandas is an open source project to make working with geospatial data in python easier. GeoPandas extends the datatypes used by pandas to allow spatial operations on geometric types.

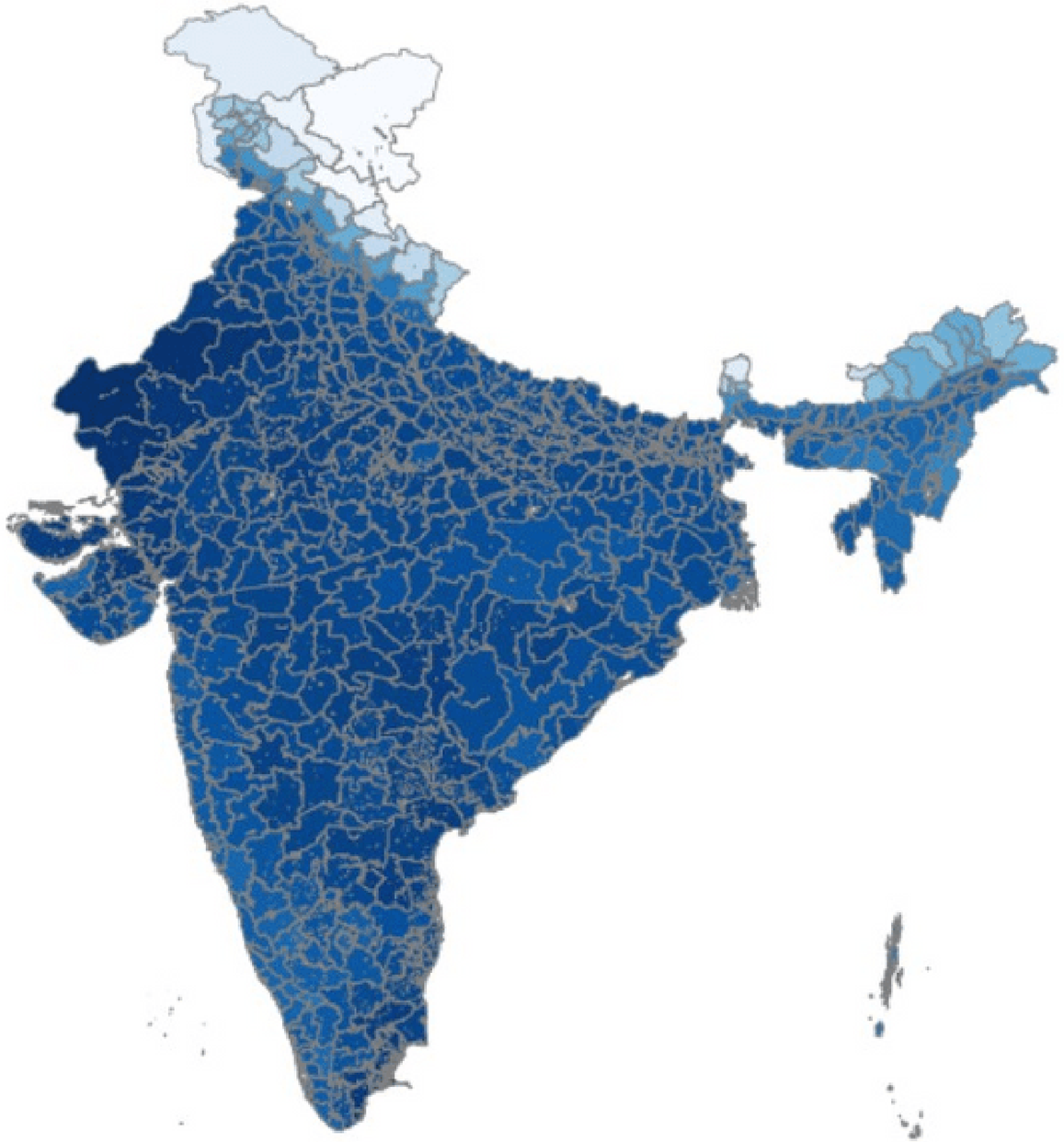
Install command:
!pip install geopandas

Import command:
import geopandas as gpd

File:
GEE_Visualization.ipynb

Plotting map of India using geopandas

Adarsh Joshy



Plotting average maximum air temperature of each district using geopandas

EXTENDING THE IDEA FURTHER!

We merged the shape file provided with the data extracted to plot the air temperature with varied color density dependending on the air temperature of a particular district.

Adarsh Joshy

Merging!

- Geopandas identifies and plots the **geometry** of the particular district / ward
- The **color density** is determined by the corresponding merged temperature data

Extracted dataset

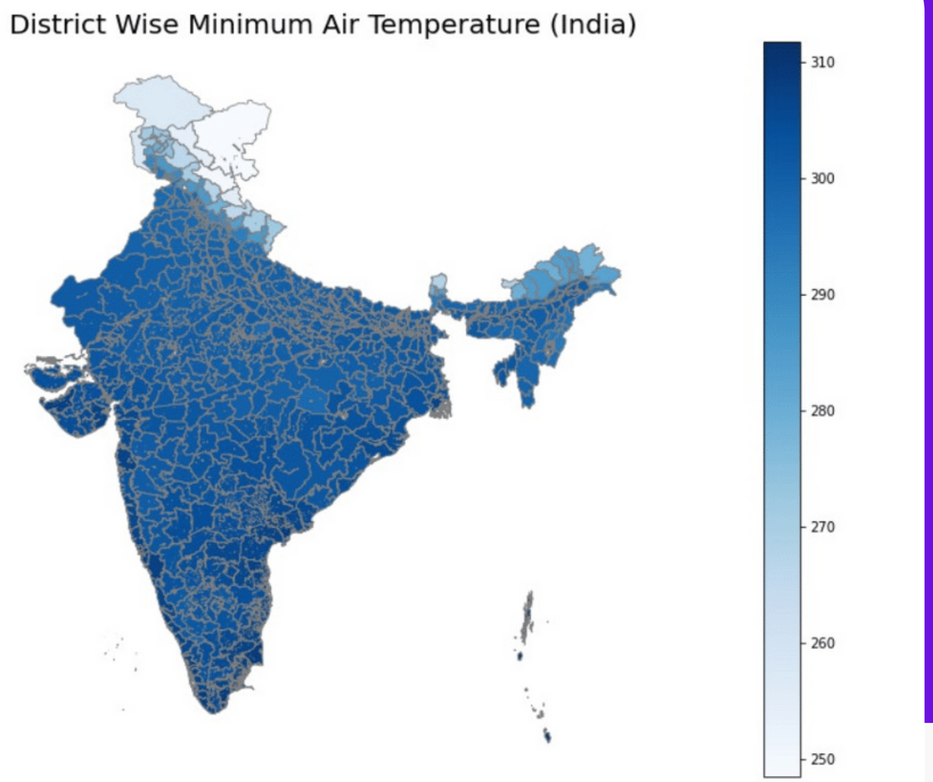
NAME	DIST91_ID	maximum_2m_air_temperature	minimum_2m_air_temperature
ADILABAD	19.0	309.127730	292.214277
AGRA	380.0	308.695673	289.329998
AHMADNAGAR	236.0	307.644744	290.298326
AHMEDABAD	113.0	310.366161	292.457588
AIZAWL	271.0	303.675486	288.137167

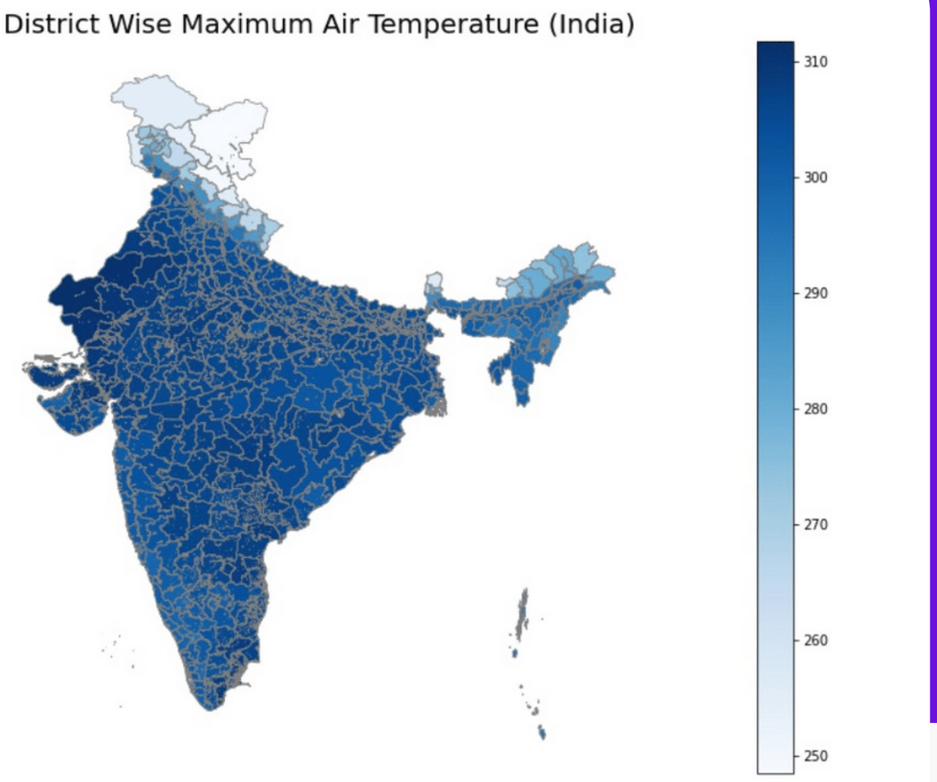
After merging with shape file

	DIST91_ID_x	NAME	STATE_UT	TOT_AREA	R_AREA	U_AREA	P_R_AREA	P_U_AREA	TOT_HOUSE	R_NM_HOUSE	U_NM_HOUSE
0	9999.0	DATA NOT AVAILABLE	JAMMU_&_KASHMIR	0.0	0.0	0.0	0.0	0.0	0	0	0
1	453.0	LADAKH	JAMMU_&_KASHMIR	0.0	0.0	0.0	0.0	0.0	0	0	0
2	454.0	KARGIL	JAMMU_&_KASHMIR	0.0	0.0	0.0	0.0	0.0	0	0	0

Adarsh Joshy

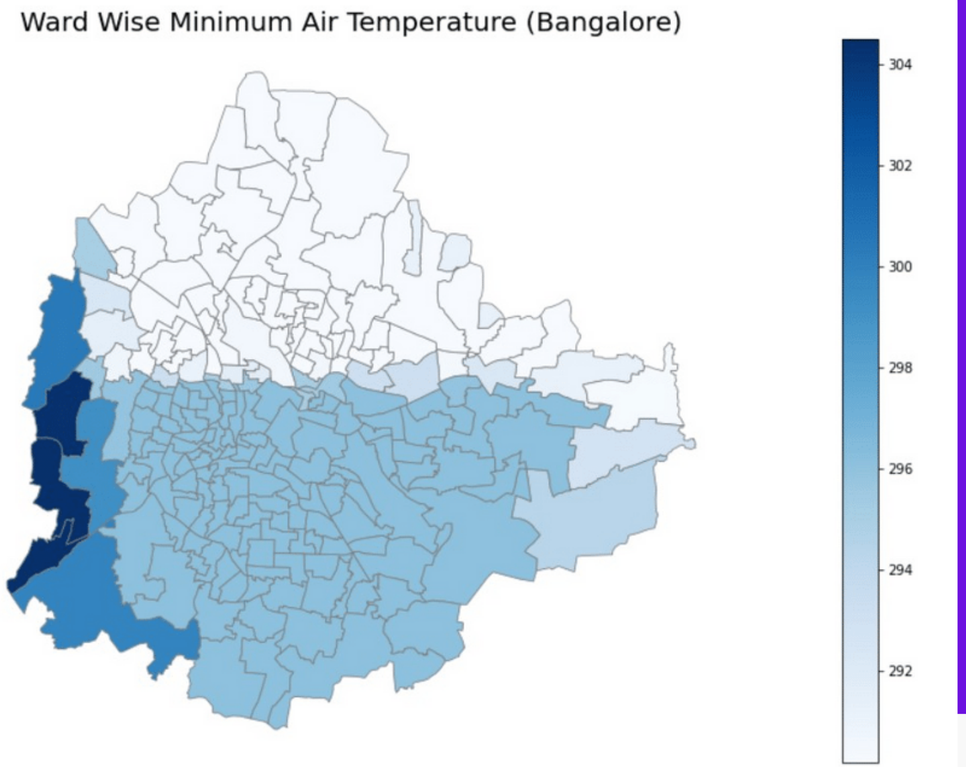
OBSERVATIONS

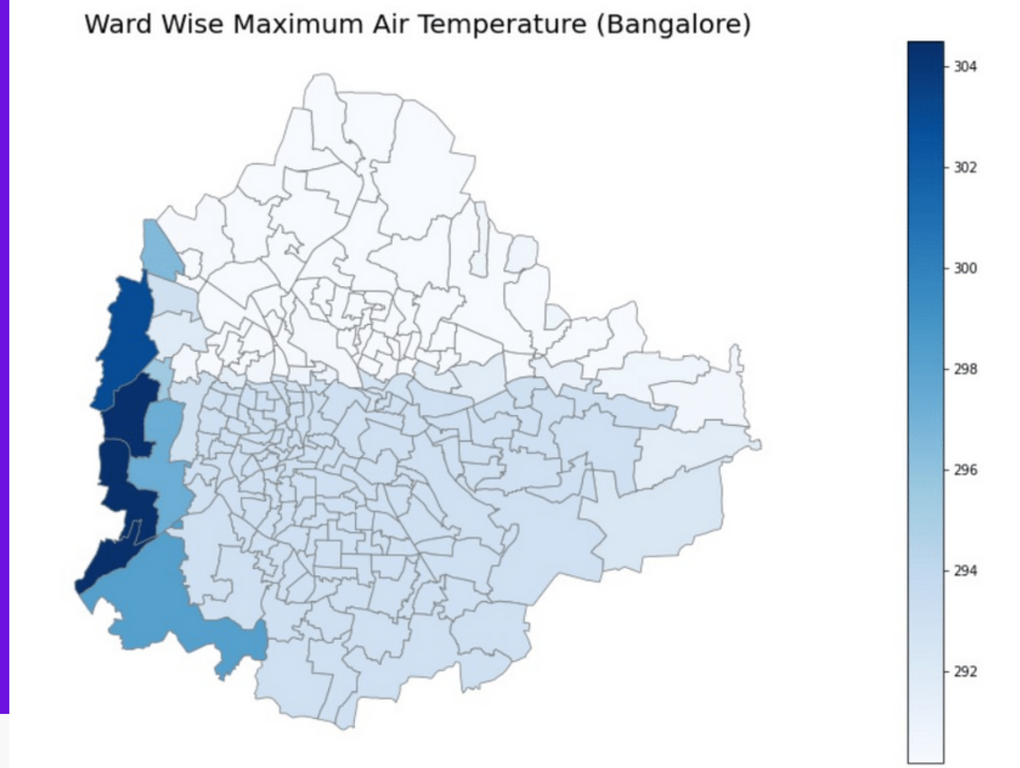




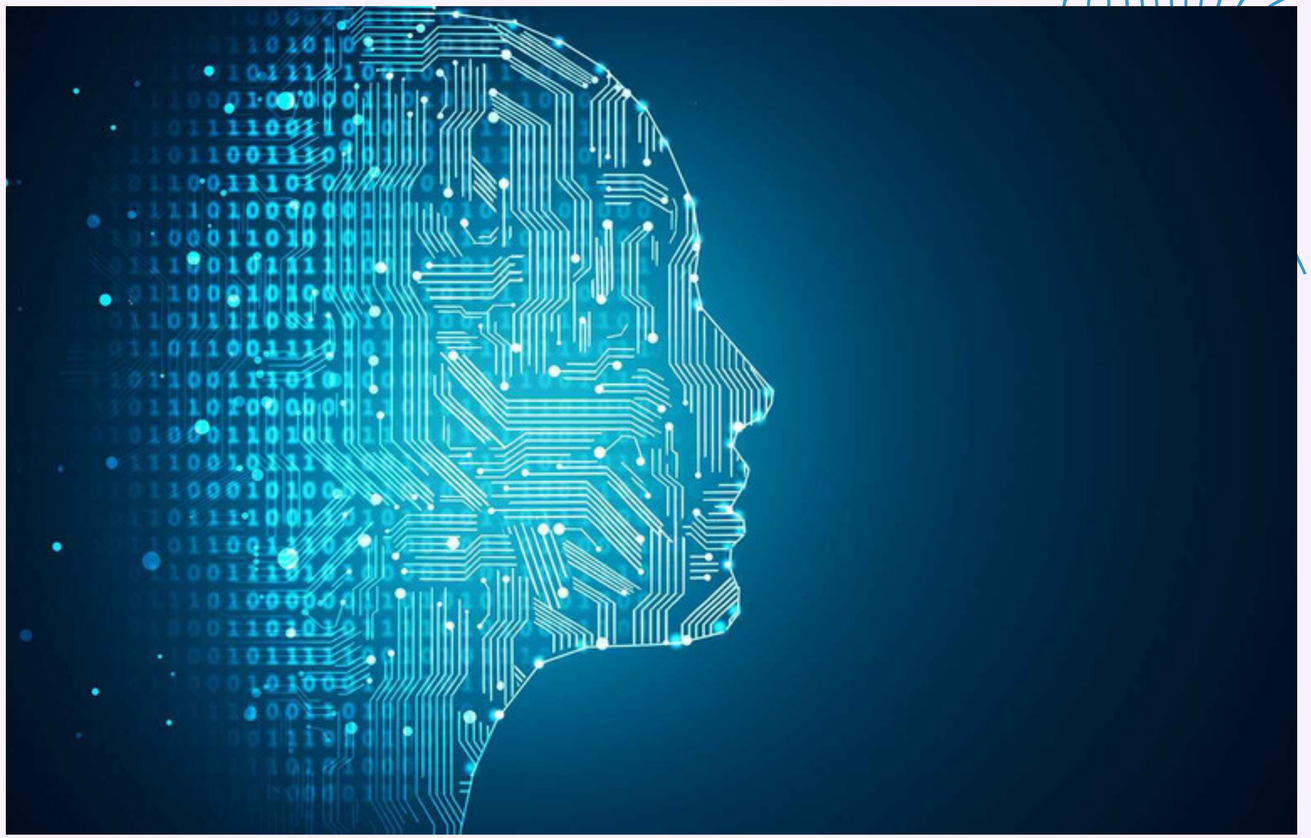
- Temperature varies from **250 K** to **310 K**
- The average maximum temperature in the northern most regions is drastically **lower** than the other parts of the country.

OBSERVATIONS





- Temperature varies from **292 K** to **304 K**
- The average temperature in the upper half of bangalore is **lower.**
- The western region has relatively **higher** temperature than other regions.



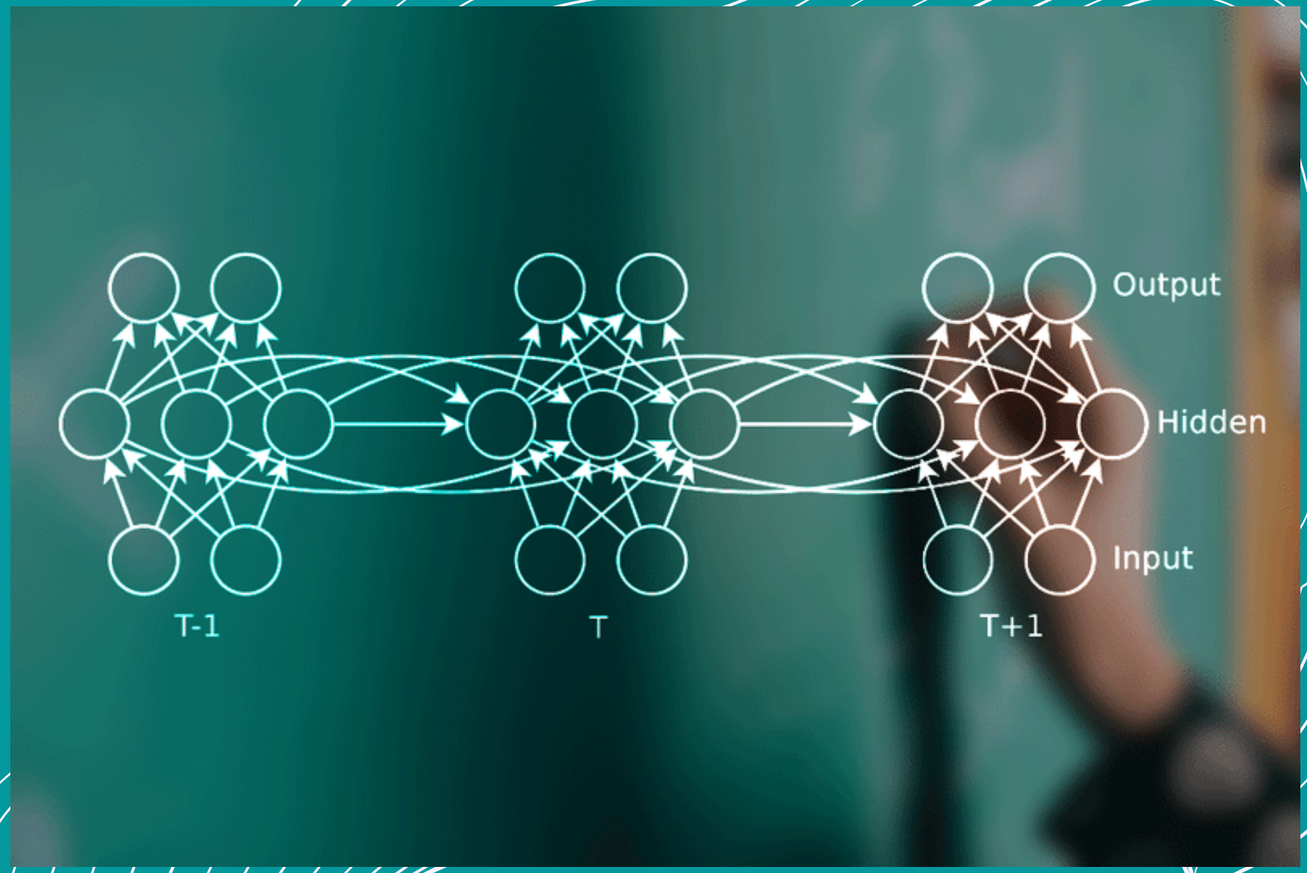
MACHINE LEARNING

With advances in Data Science and Technology, we are able to get precise forecasts of the weather in almost every location around the world. Data collected from weather stations and satellites are used to analyze and predict the meteorological disasters caused by extreme weather. According to research, based on observations of the weather in the past we can predict the weather in the future. Machine Learning can be used to forecast weather with historical weather data. Remember that the weather predicted is an educated guess!

ARJUN B L

RECURRENT NEURAL NETWORK(RNN) WITH KERAS

We used recurrent neural networks (RNN), which is a class of neural networks that is powerful for modeling sequence data. It is prominent in the field of NLP (Natural language processing).
Recurrent Neural Network(RNN) are a type of Neural Network where the output from previous step are fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is Hidden state, which remembers some information about a sequence.



SUMMARY OF THE RNN MODEL

```
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
bidirectional_2 (Bidirection (None, 30, 60)            7680
_________________________________________________________________
dropout_8 (Dropout)          (None, 30, 60)            0
_________________________________________________________________
lstm_9 (LSTM)                (None, 30, 30)            10920
_________________________________________________________________
dropout_9 (Dropout)          (None, 30, 30)            0
_________________________________________________________________
lstm_10 (LSTM)               (None, 30, 30)            7320
_________________________________________________________________
dropout_10 (Dropout)         (None, 30, 30)            0
_________________________________________________________________
lstm_11 (LSTM)               (None, 30)                7320
_________________________________________________________________
dropout_11 (Dropout)         (None, 30)                0
_________________________________________________________________
dense_2 (Dense)              (None, 4)                 124
=================================================================
Total params: 33,364
Trainable params: 33,364
Non-trainable params: 0
_________________________________________________________________
```

Our RNN model provides the next 4 values of forecast of temperature based on 30 values of historical temperature data.

PREDICTIONS

Made 3 different RNN models for prediction.

BANGALORE MEAN TEMPERATURES

We used the quarterly extracted data with the help of Bangalore Shape files given. The data was cleaned, prepared and was fitted to the RNN Model 1.

KERALA MEAN TEMPERATURE

Kerala mean temperature data was obtained from the main ERA-5 Monthly Dataset. The data was cleaned, prepared and was fitted to the RNN Model 2.

INDORE MAXIMUM TEMPERATURE

Indore Maximum temperature data was obtained from the main ERA-5 Monthly Dataset. The data was cleaned, prepared and was fitted to the RNN Model 3.

ACCURACY OF THE MODELS

MODEL 1

The below are the predicted values

```
print(predicted_temperature[0])
```

```
[298.21646 299.73602 297.93655 296.91956]
```

Now lets see the actual values

```
acutal_value = (test_set_sample[30:34])
```

```
acutal_value
```

```
array([[297.59],
       [300.38],
       [296.74],
       [294.47]])
```

Now finding the error between actual and predicted data

```
error1 = predicted_temperature[0] - acutal_value[0]
```

```
print(error1)
```

```
[ 0.62646118  2.14602295  0.34655396 -0.67044434]
```

MODEL 2

The below are the predicted values

```
predicted_temperature
```

```
array([[300.18274, 300.5763 , 300.68668, 298.6599 ]], dty
```

Seeing the actual value to compare

```
test_set_sample[30:34]
```

```
array([[299.37],
       [299.94],
       [300.47],
       [297.62]])
```

Finding the error

```
error = predicted_temperature[0] - test_set_sample[30:34]
```

```
print(error)
```

```
[0.81274 0.6363  0.21668 1.0399 ]
```

MODEL 3

```
predicted_temperature[0]
```

```
array([306.89902, 307.5554 , 307.065  , 305.6764 ],
```

Actual values

```
actual_value = (test_set_sample[34:38])
```

```
actual_value
```

```
array([[307.522407  ],
       [303.71404966],
       [304.23127174],
       [305.16312889]])
```

Finding the error

```
error = predicted_temperature[0] - actual_value[0]
```

```
print(error)
```

```
[-0.62338967  0.0329824  -0.45740456 -1.8460154 ]
```

The above error values indicates that all 3 of our models performs really well

FUTURE PREDICTIONS

BANGALORE DISTRICTS NEXT 4 QUATERLY MEAN TEMPERATURES

Predicting future temperature for next 4 quartely month

```
testing_sample_scaled = sc.fit_transform(test_set_sample[4:34])
```

```
testing_sample_scaled = np.reshape(testing_sample_scaled,(testi

predicted_temperature = regressor.predict(testing_sample_scaled
```

```
predicted_temperature = sc.inverse_transform(predicted_temperat
```

```
print(predicted_temperature[0])
```

```
[297.91928 299.58896 297.62598 296.5048 ]
```

The above are the predicted mean values for the next quarterly months for banglore

KERALA STATE NEXT 4 MONTHLY MEAN TEMPERATURES

```
testing_sample_scaled = sc.fit_transform(test_set_sample[70:100
```

```
testing_sample_scaled = np.reshape(testing_sample_scaled,(testir

predicted_temperature = regressor.predict(testing_sample_scaled
```

```
predicted_temperature = sc.inverse_transform(predicted_temperatu
```

```
print(predicted_temperature[0])
```

```
[300.33813 301.1022  300.9345  299.91898]
```

The above values are the predicted mean temperature of Kerala for next 4 months, the values are in kelvin scale

INDORE DISTRICT NEXT 4 MONTLY MAXIMUM TEMPERATES

Predicting Future Temperature with current data

```
testing_sample_scaled = sc.fit_transform(test_set_sample[31:61])
```

```
testing_sample_scaled = np.reshape(testing_sample_scaled,(testir

predicted_temperature = regressor.predict(testing_sample_scaled
```

```
predicted_temperature = sc.inverse_transform(predicted_temperatu
```

```
predicted_temperature[0]
```

```
array([307.0098 , 306.61124, 304.31796, 303.4949 ], dtype=floa
t32)
```

*NOTE : ALL THE PREDICTED VALUES ARE IN KELVIN SCALE

CONCLUSION

Our story

All 3 model's predicted values are indicating a slight increase in the temperature. Also all the analysis both quantitative and qualitative indicates that the temperature values are slightly increasing over the time. The change is almost clearly noticeable in the recent decade. This would mostly due to the human contribution to this long term crisis. During analysis, it was observed that density plot has a significant peak in the year 1990. 1990 was the warmest year since comparable record-keeping began in the middle of the 19th Century. 1990 was a year in which there was El Nino effect. El Niño is a climate pattern that describes the unusual warming of surface waters in the eastern tropical Pacific Ocean. El Nino conditions arise from the episodic warming of the tropical Pacific off Peru and lead to surges of extreme drought and heat in the Northern and Southern hemispheres. Acceleration of the planetary greenhouse effect point to higher temperatures in the 1990s and confirm that pollution of the atmosphere is already causing global warming. Preliminary evidence shows that temperatures in the lower atmosphere were warmer in 1990 than in any year since record-keeping began in 1958.

CLIMATE CHANGE IS REAL.

THANK YOU!

PRESENTED BY

GROUP 7

- A VISHNUSANKAR - B180355CS
- ADARSH JOSHY - B180098CS
- ARJUN B J - B180287CS
- HANEEM E M - B180603CS
- JAIDEEP KUMAR P M - B180049CS