CHAPTER 1

1. INTRODUCTION

The current climatology for India, 2007-2021, drawn from observable, historical data. In order to evaluate future climate scenarios and expected change, information should be used to establish a strong understanding of current climatic circumstances. Data for the present climatology can be visualized using regional variation, the seasonal cycle, or a time series. Both annual and seasonal data are accessible for analysis.

1.1 OBJECTIVES

- To know the trend of climate in India
- To inspect the gradual change in temperatures
- To find out best cities for a great summer vacation based on certain factors

1.2 DATASETS

Three Datasets have been used to attain the above goals through our project. The datasets were collected from Kaggle Platform.

Dataset 1: Indian Summer 2007-11.

The dataset has daily temperatures, dew, humidity, wind speed and more for the top 15 populous cities of India for the period between 2007 to 2011

Attributes: 20

Tuples: 6733

https://www.kaggle.com/datasets/akashram/indian-summer-over-the-years

Dataset 2: Indian Summer 2012-2021

The dataset has daily temperatures, dew, humidity, wind speed and more for the top 15 populous cities of India for the period between 2012 to 2021.

Attributes: 20

Tuples: 6733

https://www.kaggle.com/datasets/akashram/indian-summer-over-the-years

Dataset 3: Indian Cities

Each row includes a city's latitude, longitude, state, and other variables of interest.

https://www.kaggle.com/datasets/parulpandey/indian-cities-database

1.2.1 DETAILS OF DATA

The following are the units of measures for the attributes used in our project

tempax tempmin temp feelslikemax feelslikemin feelslike Dew	Celcius Celcius Celcius Celcius Celcius Celcius Celcius
Wind Speed WindDirection	2 minute average of wind speed, in KM/HR 2 minute average of wind direction- DEGREES
Visibility	KMs
Moonphase	FULL CYCLE 0 â€“ new moon 0-0.25 â€“ waxing crescent 0.25 â€“ first quarter 0.25-0.5 â€“ waxing gibbous 0.5 â€“ full moon 0.5-0.75 â€“ waning gibbous 0.75 â€“ last quarter 0.75 -1 â€“ waning crescent

CHAPTER 2

POWER QUERY

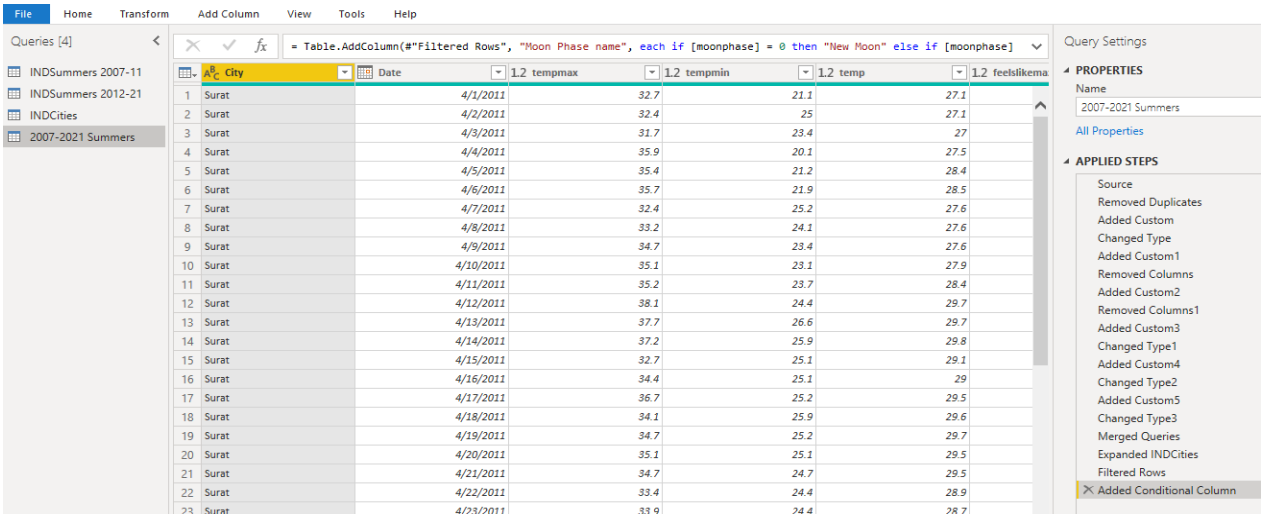
Steps followed to start the process of data cleaning

2.1 LOADING, TRANSFORMING AND CLEANING DATA

1. Clicking on Load Data and importing the CSV files into power bi
2. Transforming the data and combing the Indian summer datasets using append option
   Appended Dataset: 2007-2021 Summers
3. Removing the blank and duplicate values

2.2 MERGING THE DATASETS

1. Using merge option, joined Indian Cities data set with "2007- 2021 Summers" using City column
2. Later created required columns for future purposes.

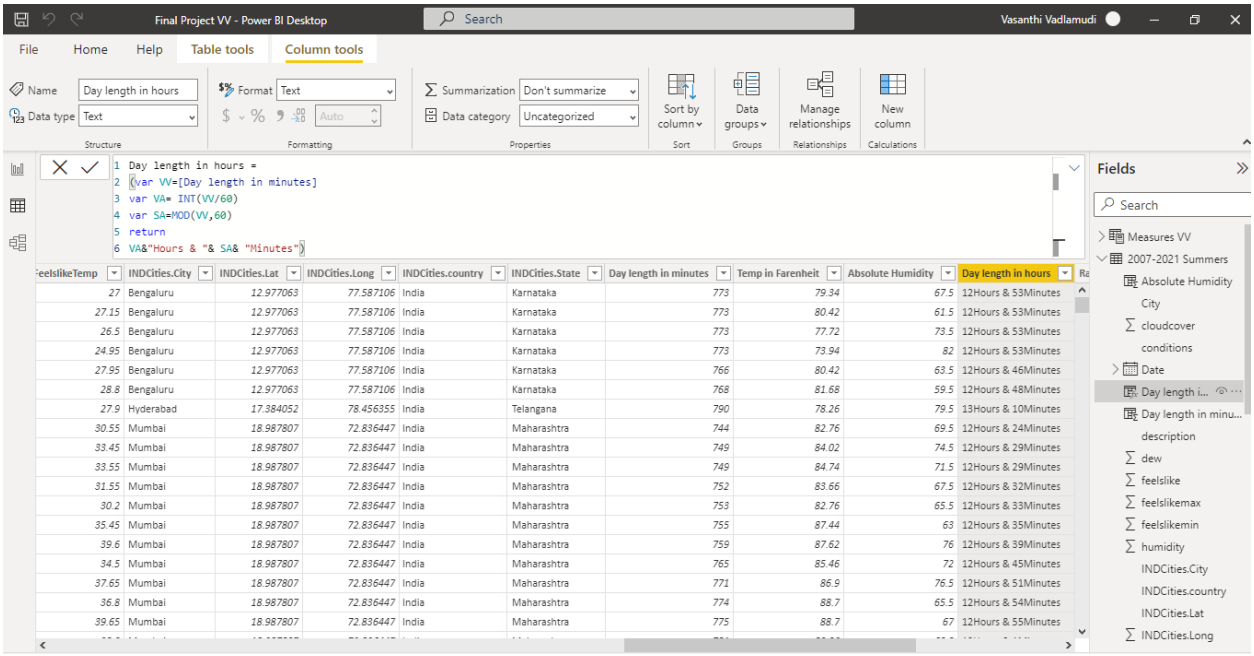


2.3 CALCULATED COLUMS

Length of Day in Hours:

To determine length of day in hours using the "Day length in minutes"custom column

```
Day length in hours =
(var VV=[Day length in minutes]
var VA= INT(VV/60)
var SA=MOD(VV,60)
return
VA&"Hours & "& SA& "Minutes")
```

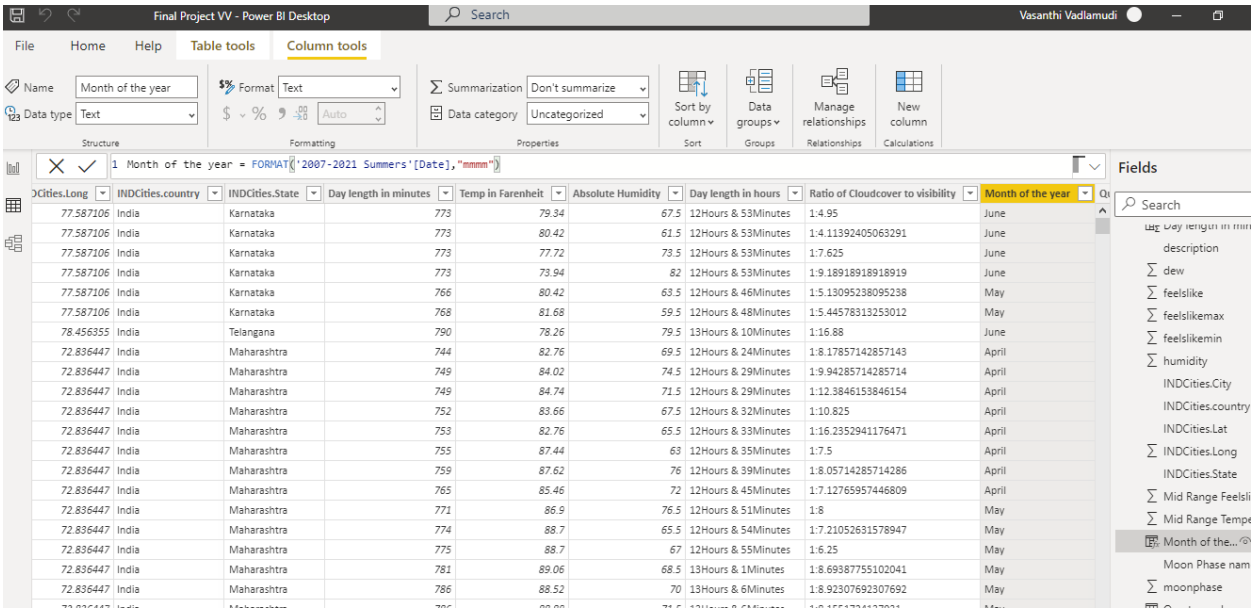


FeelslikeTemp	INDCities.City	INDCities.Lat	INDCities.Long	INDCities.country	INDCities.State	Day length in minutes	Temp in Farenheit	Absolute Humidity	Day length in hours
27	Bengaluru	12.977063	77.587106	India	Karnataka	773	79.34	67.5	12Hours & 53Minutes
27.15	Bengaluru	12.977063	77.587106	India	Karnataka	773	80.42	61.5	12Hours & 53Minutes
26.5	Bengaluru	12.977063	77.587106	India	Karnataka	773	77.72	73.5	12Hours & 53Minutes
24.95	Bengaluru	12.977063	77.587106	India	Karnataka	773	73.94	82	12Hours & 53Minutes
27.95	Bengaluru	12.977063	77.587106	India	Karnataka	766	80.42	63.5	12Hours & 46Minutes
28.8	Bengaluru	12.977063	77.587106	India	Karnataka	768	81.68	59.5	12Hours & 48Minutes
27.9	Hyderabad	17.384052	78.456355	India	Telangana	790	78.26	79.5	13Hours & 10Minutes
30.55	Mumbai	18.987807	72.836447	India	Maharashtra	744	82.76	69.5	12Hours & 24Minutes
33.45	Mumbai	18.987807	72.836447	India	Maharashtra	749	84.02	74.5	12Hours & 29Minutes
33.55	Mumbai	18.987807	72.836447	India	Maharashtra	749	84.74	71.5	12Hours & 29Minutes
31.55	Mumbai	18.987807	72.836447	India	Maharashtra	752	83.66	67.5	12Hours & 32Minutes
30.2	Mumbai	18.987807	72.836447	India	Maharashtra	753	82.76	65.5	12Hours & 33Minutes
35.45	Mumbai	18.987807	72.836447	India	Maharashtra	755	87.44	63	12Hours & 35Minutes
39.6	Mumbai	18.987807	72.836447	India	Maharashtra	759	87.62	76	12Hours & 39Minutes
34.5	Mumbai	18.987807	72.836447	India	Maharashtra	765	85.46	72	12Hours & 45Minutes
37.65	Mumbai	18.987807	72.836447	India	Maharashtra	771	86.9	76.5	12Hours & 51Minutes
36.8	Mumbai	18.987807	72.836447	India	Maharashtra	774	88.7	65.5	12Hours & 54Minutes
39.65	Mumbai	18.987807	72.836447	India	Maharashtra	775	88.7	67	12Hours & 55Minutes

Month:

To generate a column having the respective month

```
Month of the year = FORMAT('2007-2021 Summers'[Date],"mmmm")
```

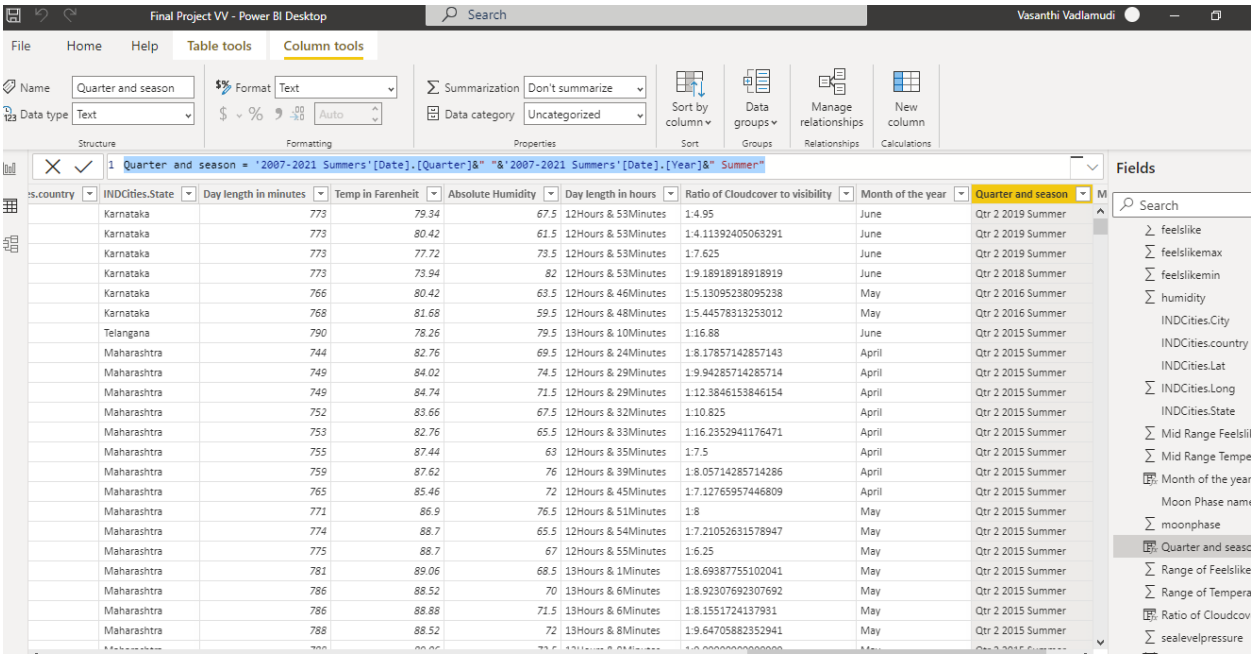


INDCities.Long	INDCities.country	INDCities.State	Day length in minutes	Temp in Farenheit	Absolute Humidity	Day length in hours	Ratio of Cloudcover to visibility	Month of the year
77.587106	India	Karnataka	773	79.34	67.5	12Hours & 53Minutes	1:4.95	June
77.587106	India	Karnataka	773	80.42	61.5	12Hours & 53Minutes	1:4.11392405063291	June
77.587106	India	Karnataka	773	77.72	73.5	12Hours & 53Minutes	1:7.625	June
77.587106	India	Karnataka	773	73.94	82	12Hours & 53Minutes	1:9.18918918918919	June
77.587106	India	Karnataka	766	80.42	63.5	12Hours & 46Minutes	1:5.13095238095238	May
77.587106	India	Karnataka	768	81.68	59.5	12Hours & 48Minutes	1:5.44578313253012	May
78.456355	India	Telangana	790	78.26	79.5	13Hours & 10Minutes	1:16.88	June
72.836447	India	Maharashtra	744	82.76	69.5	12Hours & 24Minutes	1:8.17857142857143	April
72.836447	India	Maharashtra	749	84.02	74.5	12Hours & 29Minutes	1:9.94285714285714	April
72.836447	India	Maharashtra	749	84.74	71.5	12Hours & 29Minutes	1:12.3846153846154	April
72.836447	India	Maharashtra	752	83.66	67.5	12Hours & 32Minutes	1:10.825	April
72.836447	India	Maharashtra	753	82.76	65.5	12Hours & 33Minutes	1:16.2352941176471	April
72.836447	India	Maharashtra	755	87.44	63	12Hours & 35Minutes	1:7.5	April
72.836447	India	Maharashtra	759	87.62	76	12Hours & 39Minutes	1:8.05714285714286	April
72.836447	India	Maharashtra	765	85.46	72	12Hours & 45Minutes	1:7.12765957446809	April
72.836447	India	Maharashtra	771	86.9	76.5	12Hours & 51Minutes	1:8	May
72.836447	India	Maharashtra	774	88.7	65.5	12Hours & 54Minutes	1:7.21052631578947	May
72.836447	India	Maharashtra	775	88.7	67	12Hours & 55Minutes	1:6.25	May
72.836447	India	Maharashtra	781	89.06	68.5	13Hours & 1Minutes	1:8.69387755102041	May
72.836447	India	Maharashtra	786	88.52	70	13Hours & 6Minutes	1:8.92307692307692	May

Quarter and Season:

To generate a column having the entire details of the year such as the quarter, year, season

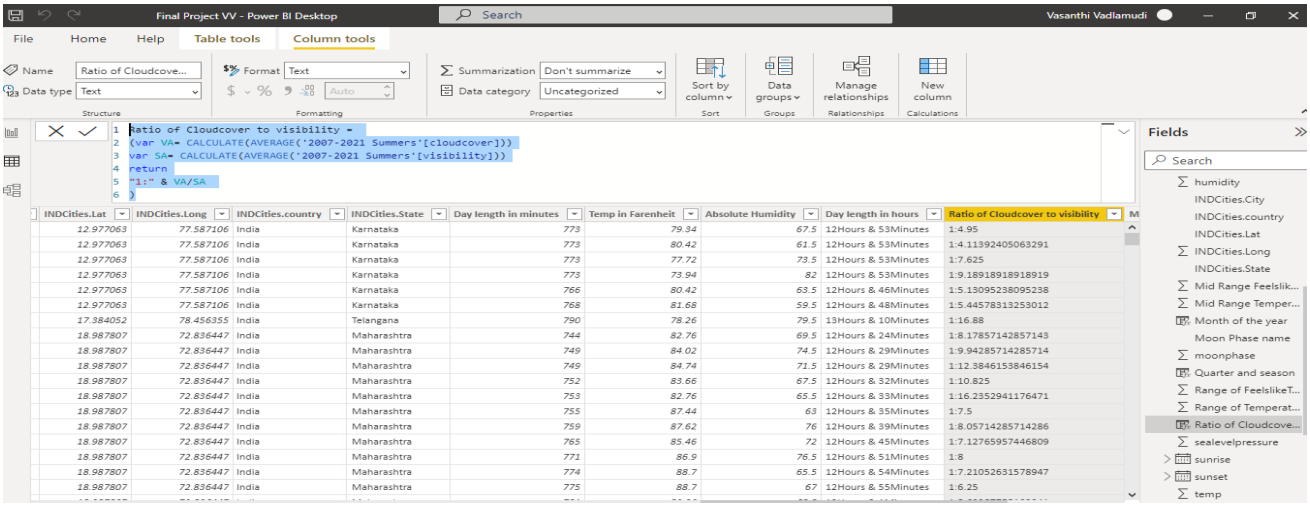
```
Quarter and season = '2007-2021 Summers'[Date].[Quarter]&" "&'2007-2021
Summers'[Date].[Year]&" Summer"
```



Ratio of Cloud cover to visibility:

Since the cloud cover and visibility are related to one another, here we calculated the ratio to know the exact proportions

```
Ratio of Cloudcover to visibility =
(var VA= CALCULATE(AVERAGE('2007-2021 Summers'[cloudcover]))
var SA= CALCULATE(AVERAGE('2007-2021 Summers'[visibility]))
return
"1:" & VA/SA
)
```



2.4 MEASURES

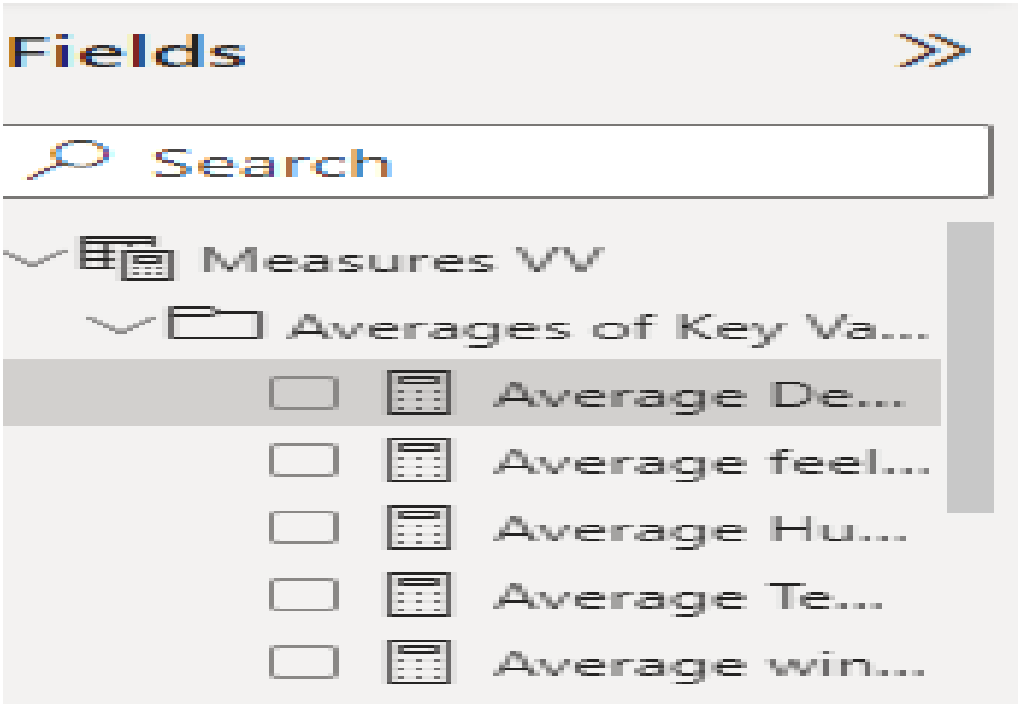
There was a measures table with the name Measures VV, where it has two folders

1. Average function for of Key attributes
2. Essential Measures

AVERAGE FUNCTION FOR KEY ATTRIBUTES

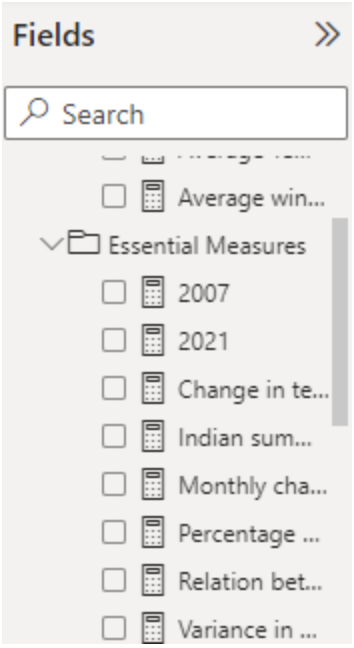
Measures using **Average** function for dew, humidity, temperature, windspeed, feels like temperature per city

```
Average Dew per city = AVERAGEX( KEEPFILTERS(VALUES('2007-2021
Summers'[City])),CALCULATE(AVERAGE('2007-2021 Summers'[dew]))
)
```



ESSENTIAL MEASURES

These are the most vital measures which are helpful for logical representation of visuals



Change in temperature every year:

```
Change in temperature everyear =
IF(
    ISFILTERED('2007-2021 Summers'[Date]),
    ERROR("Time intelligence quick measures can only be grouped or filtered by the Power BI-provided date hierarchy or primary date column."),
    VAR VV =
        CALCULATE(
            AVERAGE('2007-2021 Summers'[temp]),
            DATEADD('2007-2021 Summers'[Date].[Date], -1, YEAR)
        )
    RETURN
        DIVIDE(AVERAGE('2007-2021 Summers'[temp]) - VV, VV)
)
```

Indian summer average temperature:

```
Indian summer average temp = AVERAGE('2007-2021 Summers'[Temp in Farenheit])
```

Monthly change in temperature:

```
Monthly change in temperature =
IF(
    ISFILTERED('2007-2021 Summers'[Date]),
    ERROR("Time intelligence quick measures can only be grouped or filtered by the Power BI-provided date hierarchy or primary date column."),
    VAR vv =
```

```
        CALCULATE(
            AVERAGE('2007-2021 Summers'[temp]),
            DATEADD('2007-2021 Summers'[Date].[Date], -1, MONTH)
        )
    RETURN
        DIVIDE(AVERAGE('2007-2021 Summers'[temp]), vv, vv)
)
```

Percentage Change in Temperature 2007-21:

```
2007 =
CALCULATE(AVERAGE('2007-2021 Summers'[temp]),FILTER('2007-2021 Summers','2007-2021
Summers'[Date].[Year]=2007))
2021 =
CALCULATE(AVERAGE('2007-2021 Summers'[temp]),FILTER('2007-2021 Summers','2007-2021
Summers'[Date].[Year]=2021))

Percentage Change in Temperature 2007-21 =
DIVIDE('Measures VV'[2021],'Measures VV'[2007],0)-1
```

Then change result by changing format to percentage

Relation between Windspeed and sealevel pressure:

```
Relation between Windspeed and sealevel pressure =
VAR __CORRELATION_TABLE = VALUES('2007-2021 Summers'[City])
VAR __COUNT =
    COUNTX(
        KEEPFILTERS(__CORRELATION_TABLE),
        CALCULATE(
            AVERAGE('2007-2021 Summers'[windspeed])
                * AVERAGE('2007-2021 Summers'[sealevelpressure])
        )
    )
VAR __SUM_X =
    SUMX(
        KEEPFILTERS(__CORRELATION_TABLE),
        CALCULATE(AVERAGE('2007-2021 Summers'[windspeed]))
    )
VAR __SUM_Y =
    SUMX(
        KEEPFILTERS(__CORRELATION_TABLE),
        CALCULATE(AVERAGE('2007-2021 Summers'[sealevelpressure]))
    )
VAR __SUM_XY =
    SUMX(
        KEEPFILTERS(__CORRELATION_TABLE),
```

```
        CALCULATE(
            AVERAGE('2007-2021 Summers'[windspeed])
                * AVERAGE('2007-2021 Summers'[sealevelpressure]) * 1.
        )
    )
VAR __SUM_X2 =
    SUMX(
        KEEPFILTERS(__CORRELATION_TABLE),
        CALCULATE(AVERAGE('2007-2021 Summers'[windspeed]) ^ 2)
    )
VAR __SUM_Y2 =
    SUMX(
        KEEPFILTERS(__CORRELATION_TABLE),
        CALCULATE(AVERAGE('2007-2021 Summers'[sealevelpressure]) ^ 2)
    )
RETURN
    DIVIDE(
        __COUNT * __SUM_XY - __SUM_X * __SUM_Y * 1.,
        SQRT(
            (__COUNT * __SUM_X2 - __SUM_X ^ 2)
                * (__COUNT * __SUM_Y2 - __SUM_Y ^ 2)
        )
    )
```

Variance in windspeed:

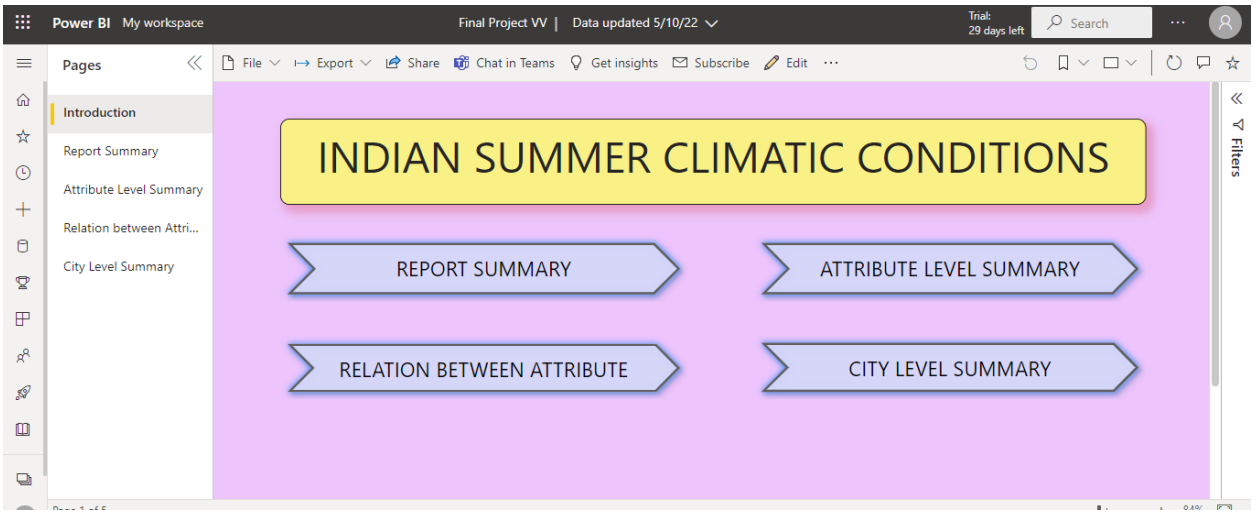
```
Variance in windspeed =
VARX.P(
    KEEPFILTERS(VALUES('2007-2021 Summers'[City])),
    CALCULATE(AVERAGE('2007-2021 Summers'[windspeed]))
)
```

CHAPTER 3
REPORT VISUALS

The report has five (5) pages.

1. Introduction
2. Report Summary
3. Attribute level summary
4. Relation between attributes
5. City level summary

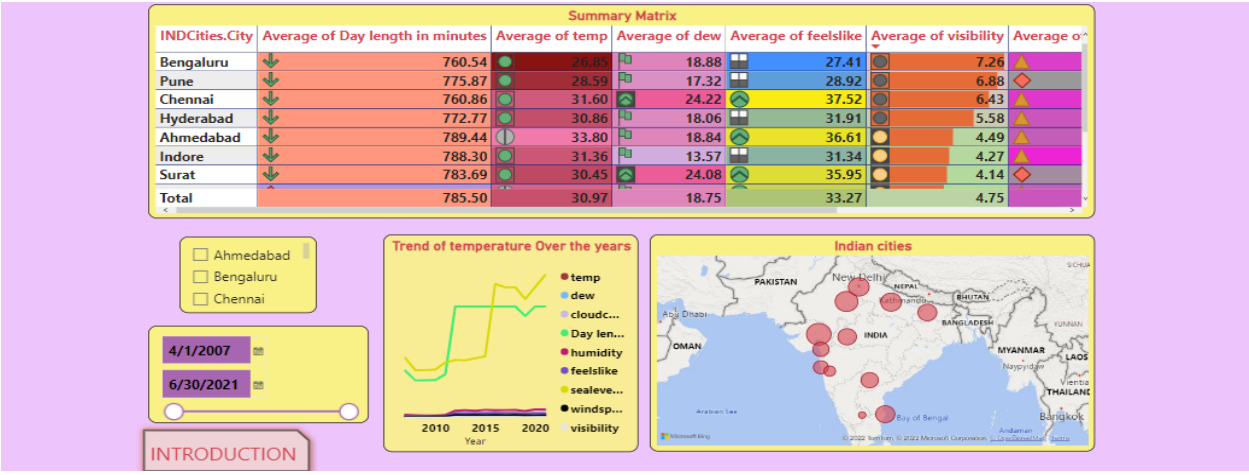
3.1 INTRODUCTION PAGE



This page has navigation controls for the rest of the pages using button in shape of chevron arrow.

3.2 REPORT SUMMARY

Here you can observe the overall Indian city's climatic conditions



This page has 5 visuals in total

Types of visuals in this page

1. Matrix
2. Slicers
3. Line chart
4. Map

Summary Matrix

This have almost all the important climate influencing attributes sorted using conditional formatting to understand the maximum, minimum and middle values of each attribute. It is formatted using data bars, color and icons.

City and Date Slicers

Useful to find the details for a particular city at a particular period of time to have better and easy understanding.

Trend of temperature over the years

In this line chart we interpreted some other attributes along with temperature to understand the trend of climate over the years.

OBSERVATION 1:

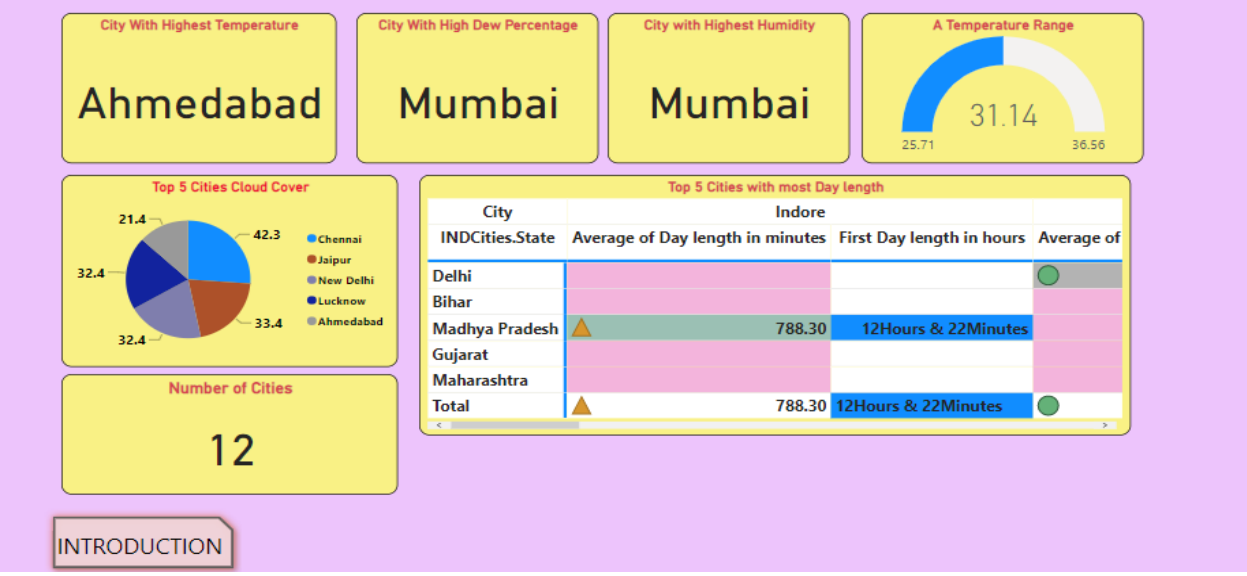
From this graph we can understand that over the last 10 years the temperature is being fluctuated while with other key influencers such as humidity, dew, sea-level pressure is increasing.

Indian Cities

From this map we visually see compare the heat in atmosphere in the city by looking at the size of bubbles as it is filtered based on the average temperature of the cities.

3.3 ATTRIBUTE LEVEL SUMMARY

In this page you can see the conclusions obtained for certain attributes.



This page has 7 visuals, where 4 of them being Card Visuals

Types of visuals in this page

1. Card
2. Pie Chart
3. Matrix
4. Gauge

Card Visuals

Using the card visuals, we displayed the cities with highest

Temperature- Ahmedabad

Dew- Mumbai

Humidity- Mumbai

OBSERVATION 2:
We can clearly say that Mumbai is the most suffocating city in India during summer due to high dew and humidity.

Reason: It have a lot of beaches

One more is number of cities that were being considered in this project to analyze.

Top 5 Cities Cloud Cover

This pie charts explains the percentage of cloud cover for top 5 cities based on temperature.

OBSERVATION 3:

Although having high temperatures in summer Chennai, it also have the highest cloud cover. So, there would be a cloudy climate for most of the summer in Chennai.

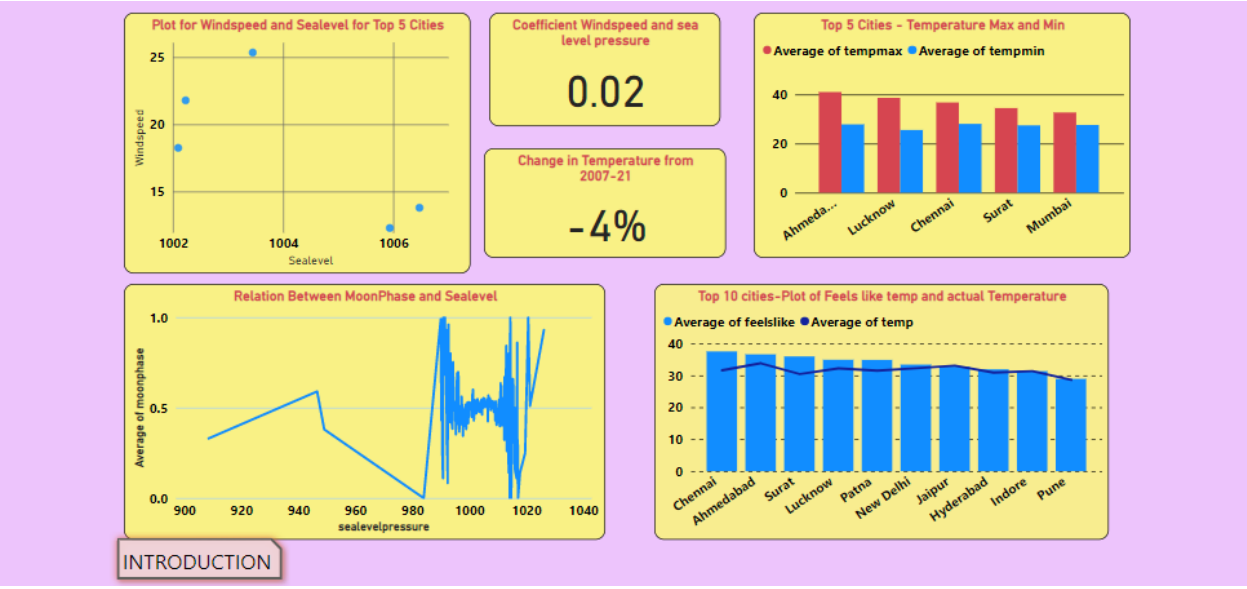
Top 5 cities based on Day length

As the title says with top 5 cities based on length of day (sunrise to sunset). It was also conditionally formatted based on max and min values.

Temperature Range

This displays the usual mid-range of temperature and gauge indicator points at 31 degrees Celsius as average range of heat.

3.4 RELATION BETWEEN ATTRIBUTES



In this page there are 5 visuals

Types of Visuals used in this page

1. Scatter plot
2. Card visual
3. Clustered Column chart
4. Line Chart
5. Line and clustered column chart

Plot for windspeed and sea-level pressure for top 5 cities

This scatter plot represents the relation between wind speed and sea level pressure for top 5 cities filtered based on temperature.

OBSERVATION 4:

From the above graph we can understand that sea-level pressure and wind speed are inversely correlated.

Note: There are also proven theories on this concept

Change in temperature 2007-2021

It displays that there is a change of -4% in temperature, which explains our trend analysis of climate presented in our REPORT SUMMARY page.

Relation between Moon phase and sea level

From the line graph we can say that moon phase has an impact on sea level pressure which is known as lunar barometric pressure.

OBSERVATION 5:
We can clearly say that the sea-level pressure is almost high most of the days when it is a full moon phase-0.5.

Minimum and Maximum Temperatures

This bar chart represents the minimum and maximum temperatures for top 5 cities

OBSERVATION 6:
We can say that Ahmedabad has the highest maximum temperature where Chennai has the least minimum temperature

Actual Temperature and feels like Temperature

In this chart the line represents the actual temperature, and the bar represents feels like temperature. The graph is plotted only for top 10 hottest cities in summer.

OBSERVATION 7:
Chennai has the highest difference between the actual temperature and the feels like temperature, where Pune has similar temperatures.

3.5 CITY LEVEL ANALYSIS

This is the page which can answer the question of best place to live in summer.

Types of visuals used

1. Slicer
2. Tree map
3. Multirow card

Date and Temperature Slicer

This is helpful if we want to know the best cities to live in a particular period and to adjust the required temperature

Top 10 Cities with mild temperature

This tree map provides the information required for the cities with mild temperatures in summer based on the requirement (you can adjust by using slicers).

Best Places to live in Summer

This multi row card displays the top 5 cities or states to live based on temperature, dew, humidity, wind speed, feels like temperature.

OBSEREVATION 8:

Bangalore is the top place to live moderate climatic conditions.

This is actually true because usually Bangalore never gets too hot, and it is known for its pleasant climate throughout the year.

CONCLUSION

From the project we made, all the above eight (8) observations are made. Bengaluru was observed as the best city to live or visit due to its moderate climatic conditions. Even few geographical correlations were observed. This model can be further developed using a better dataset and applying more visuals in the future.