**DATA SCIENCE PROJECT REPORT**

(Project Semester August-December 2020)

**Power Consumption in India (2019-2020)**

Submitted by

P Anand Kumar Reddy

Registration No: 11807873

Computer Science and Engineering

Section: KM074

Course Code: INT217

Under the Guidance of

**Ms. Komal Arora**

**Discipline of CSE/IT**

**Lovely School of Computer Science and Engineering**

**Lovely Professional University, Phagwara**



**CERTIFICATE**

This is to certify that P Anand Kumar Reddy bearing Registration no. 11807873 has completed INT217 project titled, **“Power Consumption in India (2019-2020)”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor**

**Designation of the Supervisor**

**School of Computer Science and Engineering**

**Lovely Professional University**

**Phagwara, Punjab.**

Date: 20-11-2020

**DECLARATION**

I, P Anand Kumar Reddy, student of Lovely Professional University under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 20-11-2020 Signature Registration No. 11807873 P Anand Kumar Reddy

**ACKNOWLEDGEMENT**

A project work is a combination of views, ideas, suggestions and contribution of many people. Thus, one of the pleasant parts of writing the report is to thank those who have contributed towards its fulfilment.

I consider it as great privilege to have esteemed Lecturer Ms. Komal Arora as my project guide. I take this opportunity to express my sincere gratitude to him through constant advice and constructive criticism nourished my interest in the subject and provided a free and pleasant atmosphere to work against all odd situations. I avail this opportunity to extend my heart full thanks and deep respect to faculty member for their able guidance during this project.

My gratitude to all those, who responded to my questionnaire in a well-defined manner and helped me acquiring knowledge.

I would like to communicate a deep sense of gratitude to all these people without whom my project would not have been such a great learning experience.

P Anand Kumar Reddy

KM074

Reg No. 11807873

**Lovely Professional University**

**Table of Contents**

**Sr.No. Particulars**

1. [Introduction](#Introduction)
2. [Scope of the Analysis](#Scope_of_the_Analysis)
3. [Source of the Dataset](#Source_of_the_Dataset)
4. [ETL Process](#ETL_Process)
5. [Analysis on Dataset](#Analysis_of_Dataset)
6. Introduction
7. Specific Requirements/Functions and Formulas
8. Analysis Results
9. [List of Analysis with Results](#List_of_Analysis_with_Results)
10. [References](#References)
11. [Bibliography](#Bibliography)

**Introduction**

India is the world's third-largest producer and third-largest consumer of electricity. The national electric grid in India has an installed capacity of 370.106 GW as of 31 March 2020. Renewable power plants, which also include large hydroelectric plants, constitute 35.86% of India's total installed capacity. During the 2018-19 fiscal year, the gross electricity generated by utilities in India was 1,372 TWh and the total electricity generation (utilities and non-utilities) in the country was 1,547 TWh. The gross electricity consumption in 2018-19 was 1,181 kWh per capita.  
In 2015-16, electric energy consumption in agriculture was recorded as being the highest (17.89%) worldwide. The per capita electricity consumption is low compared to most other countries despite India having a low electricity tariff.

In light of the recent COVID-19 situation, when everyone has been under lockdown for the months of April & May the impacts of the lockdown on economic activities have been faced by every sector in a positive or a negative way.  
With the electricity consumption being so crucial to the country, we came up with a plan to study the impact on energy consumption state and region wise.

The dataset is exhaustive in its demonstration of energy consumption state wise.

* **Dataset**

|  |  |
| --- | --- |
| **Column** | **Definition** |
| **States** | **All States in India** |
| **Regions** | **Regions in India** |
| **Latitude and Longitude** | **The Coordinates of States in India** |
| **Dates** | **All dates between 2019 and 2020** |
| **Usage** | **Electricity usage by different states** |

**Scope of the Analysis**

Data is in the form of a time series for a period of 17 months beginning from 2nd Jan 2019 till 23rd May 2020. Rows are indexed with dates and columns represent states. Rows and columns put together, each datapoint reflects the power consumed in Mega Units (MU) by the given state (column) at the given date (row). The data is divided into multiple tables like maximum power usage states and minimum usage states. Data shows that main electricity usage by regions and location in maps by using latitude and longitude.

Power System Operation Corporation Limited (POSOCO) is a wholly-owned Government of India enterprise under the Ministry of Power. It was earlier a wholly-owned subsidiary of Power Grid Corporation of India Limited. It was formed in March 2009 to handle the power management functions of PGCIL.

Since such vast field of data present of the Power Consumption in India there is wide range of scope of the analysis of date. For example:

1. State wise Electricity Usage
2. Date wise Electricity Usage
3. Maximum and Minimum Power Usage States

**Source of the Dataset**

* The dataset is taken from the Kaggle with the name ‘Power Consumption in India’.

<https://www.kaggle.com/twinkle0705/state-wise-power-consumption-in-india>

* Author of the Dataset

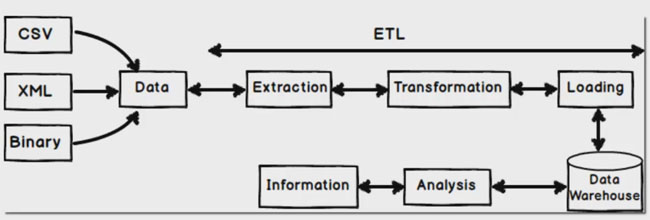
[**Twinkle Khanna**](https://www.kaggle.com/twinkle0705)

* Data last Updated

**June 2020**

**ETL Process**

**ETL** stands for Extraction, Transformation and Loading. It is a **process** in data warehousing to extract data, transform data and load data to final source. **ETL** covers a **process** of how the data are loaded from the source system to the data warehouse. Let us briefly describe each step of the **ETL process**.

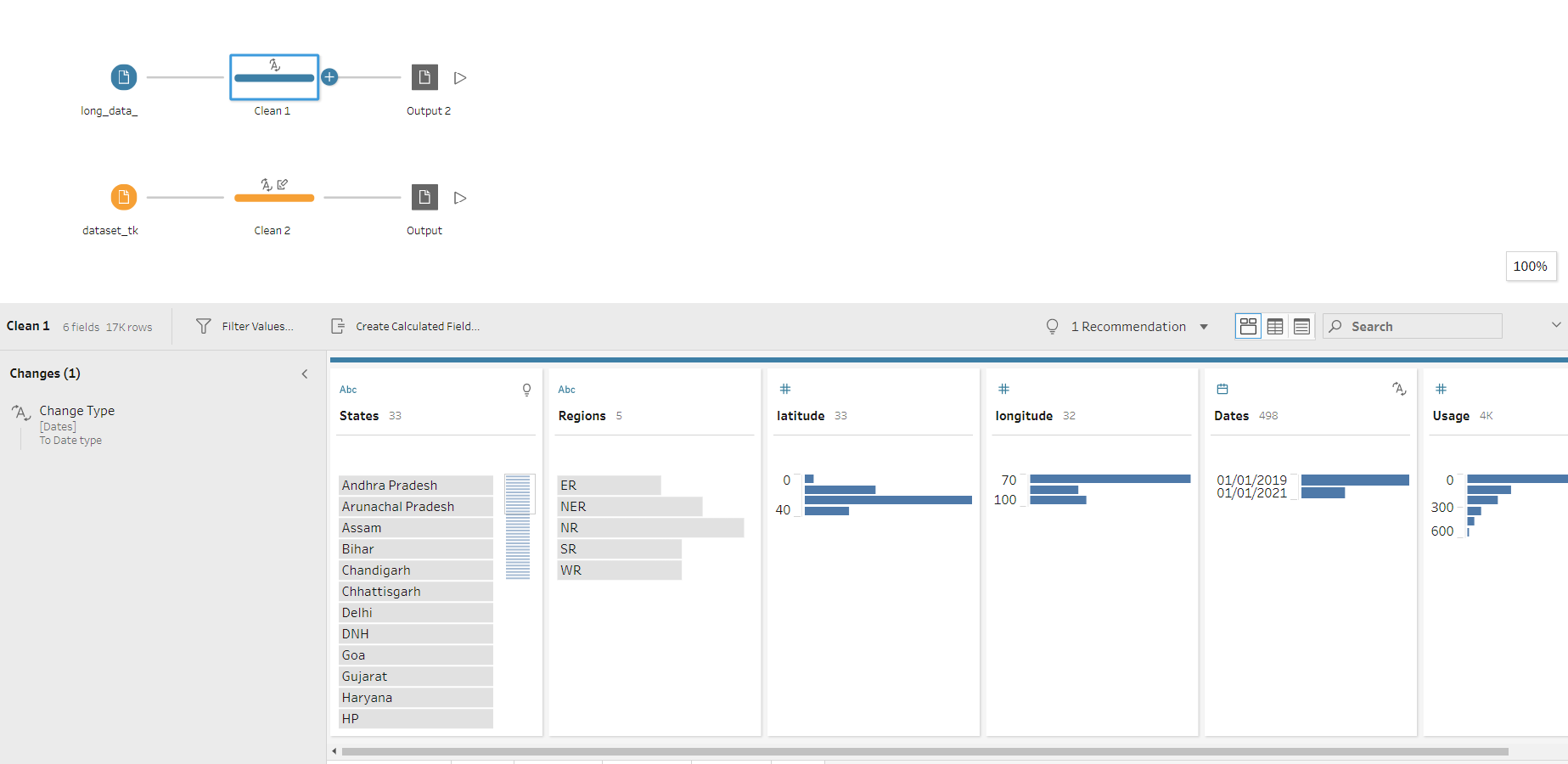


* **Extraction:** Extraction is the first step of ETL process where data from different sources like txt file, XML file, Excel file or various sources collected.
* **Transformation:** Transformation is the second step of ETL process where all collected data is been transformed into same format i.e. format can be anything as per our requirement before loading it to data-warehouse i.e. it may be data-type format, data merge format, splitting format, alphabet joining format, currency format etc.
* **Loading:** Final step of ETL process, the big chunk of data which is collected from various sources and transformed then finally load to our data warehouse.

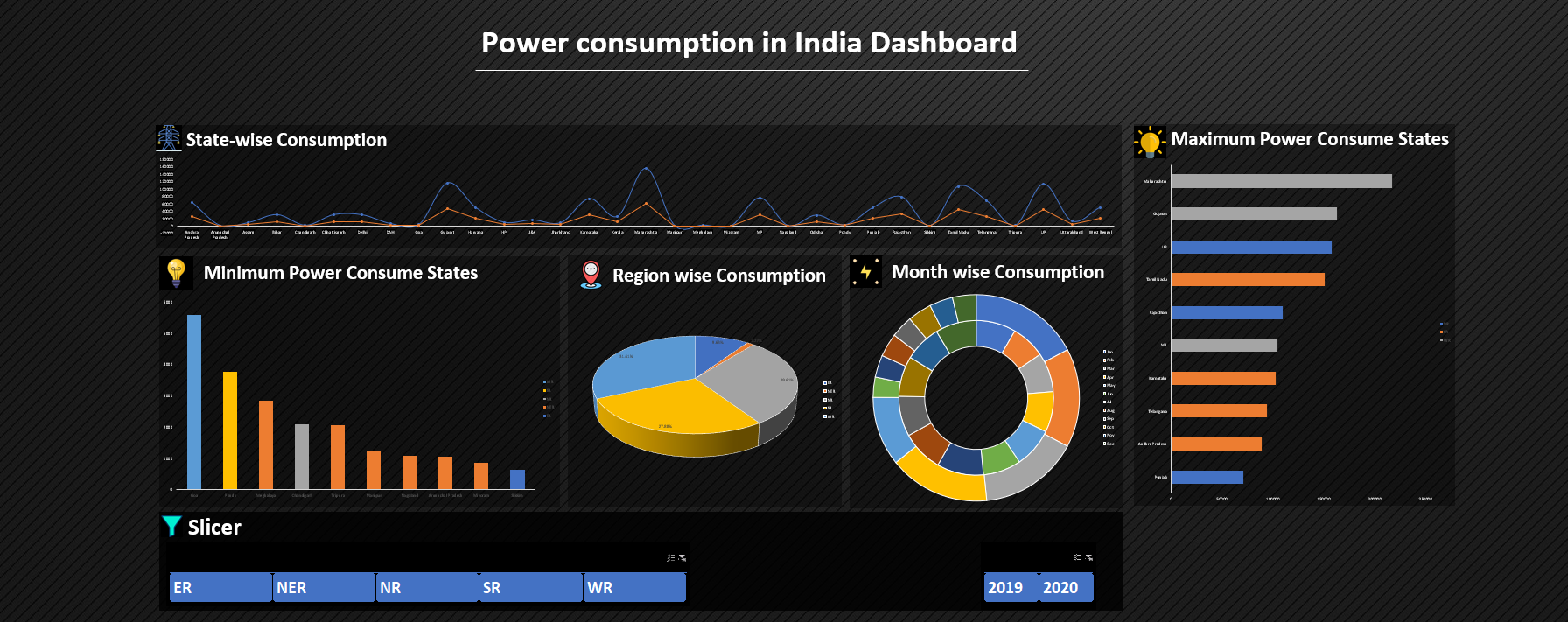
**Analysis of Dataset**

1. **Tableau**
2. Introduction: In this tableau, I made some changes to the data and also some cleaning like change date and removing time.
3. Specific Requirements/Functions and Formulas:

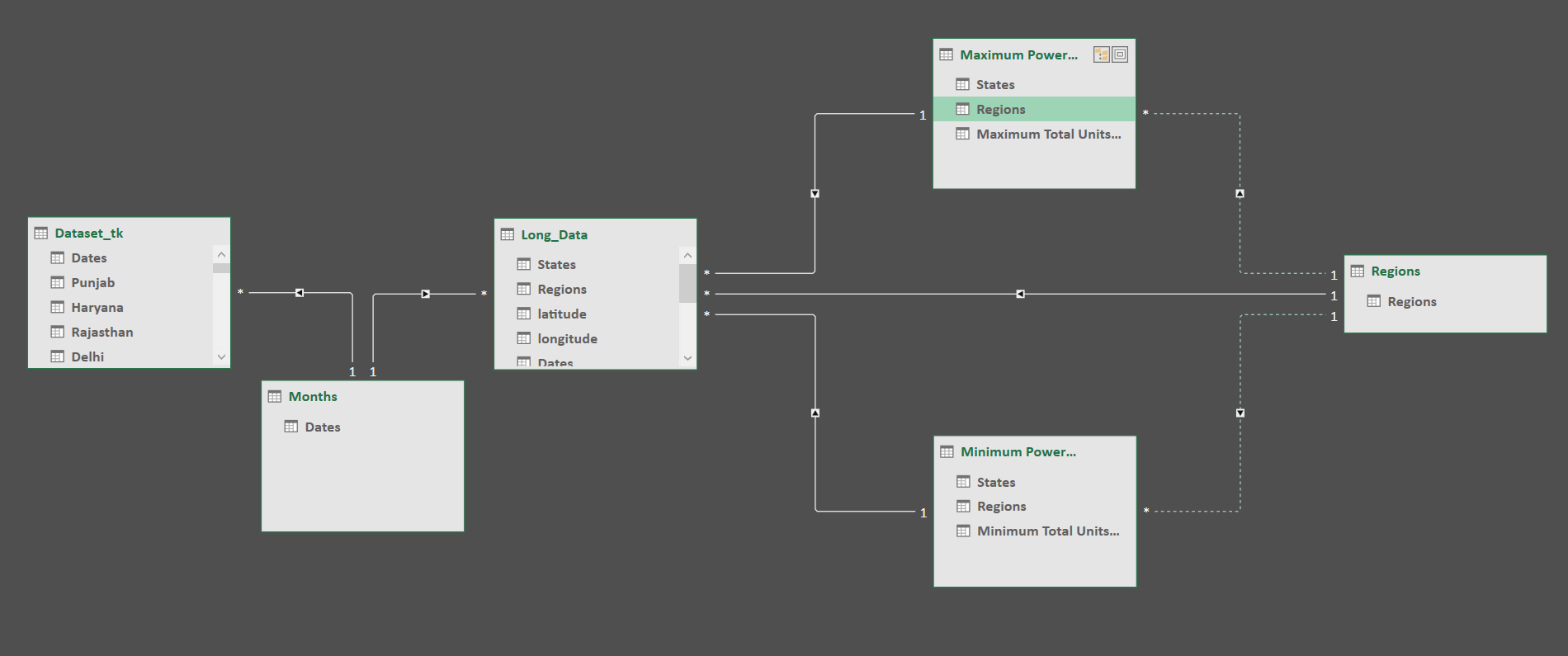
* Tableau Prep



1. **Dashboard**
2. Introduction: It contains all pivot tables, slicers, and hyperlinks. Here we can see the whole data and how it performs when we click the slicers.



1. **Diagram View**
2. Introduction: Diagram View to connect all the tables in excel sheet and I almost connected 4 tables and also extra 2 tables because some of the tables are not connecting.



1. **Maximum Power Consume States:**
2. Introduction: The analysis shows the top 10 maximum power consumption states in India.
3. Specific Requirements/Functions and Formulas:

* Pivot table of States
* Pivot table of Usage
* Stacked Bar Chart

1. Analysis Results:

* Maharashtra state consume 12.7% power. It is the highest Electricity Consumption State.

1. **Minimum Power Consume States:**
2. Introduction: The analysis shows the top 10 minimum power consumption states in India.
3. Specific Requirements/Functions and Formulas:

* Pivot table of states
* Pivot table of Usage
* Stacked Bar Chart

1. Analysis Results:

* Sikkim state consume 0.0379% power. It is lowest Electricity Consumption State.
* In this pandemic time over all Regions Electricity Consumption is decreasing comparatively last year.

1. **State wise Consumption:**
2. Introduction: The analysis shows the power usage comparison between states in India.
3. Specific Requirements/Functions and Formulas:

* Pivot table of years
* Pivot table of states
* Line with Markers Chart

1. Analysis Results:

* In June month usage of power is started decreasing again July month started increasing power usage
* October and November are usually cloudless so that time the usage of Electricity is low.
* Again, July middle started increasing the power consumption

1. **Region wise Consumption**
2. Introduction: The analysis shows the relation between Regions and Sum of Power Usage in India.
3. Specific Requirements/Functions and Formulas:

* Pivot table of Regions
* Pivot table of Sum of Usage
* 3-D Pie Chart

1. Analysis Results:

* In North region UP consume maximum electricity, Chandigarh Consume low electricity in north region.
* In West Region Maharashtra State consume maximum electricity
* In South region Tamil Nadu state consume maximum power and Pondy consume minimum electricity.
* In NER Assam Consume maximum electricity

1. **Month wise Consumption:**
2. Introduction: The analysis shows the relation between Months and Sum of Power Usage in different years in India.
3. Specific Requirements/Functions and Formulas:

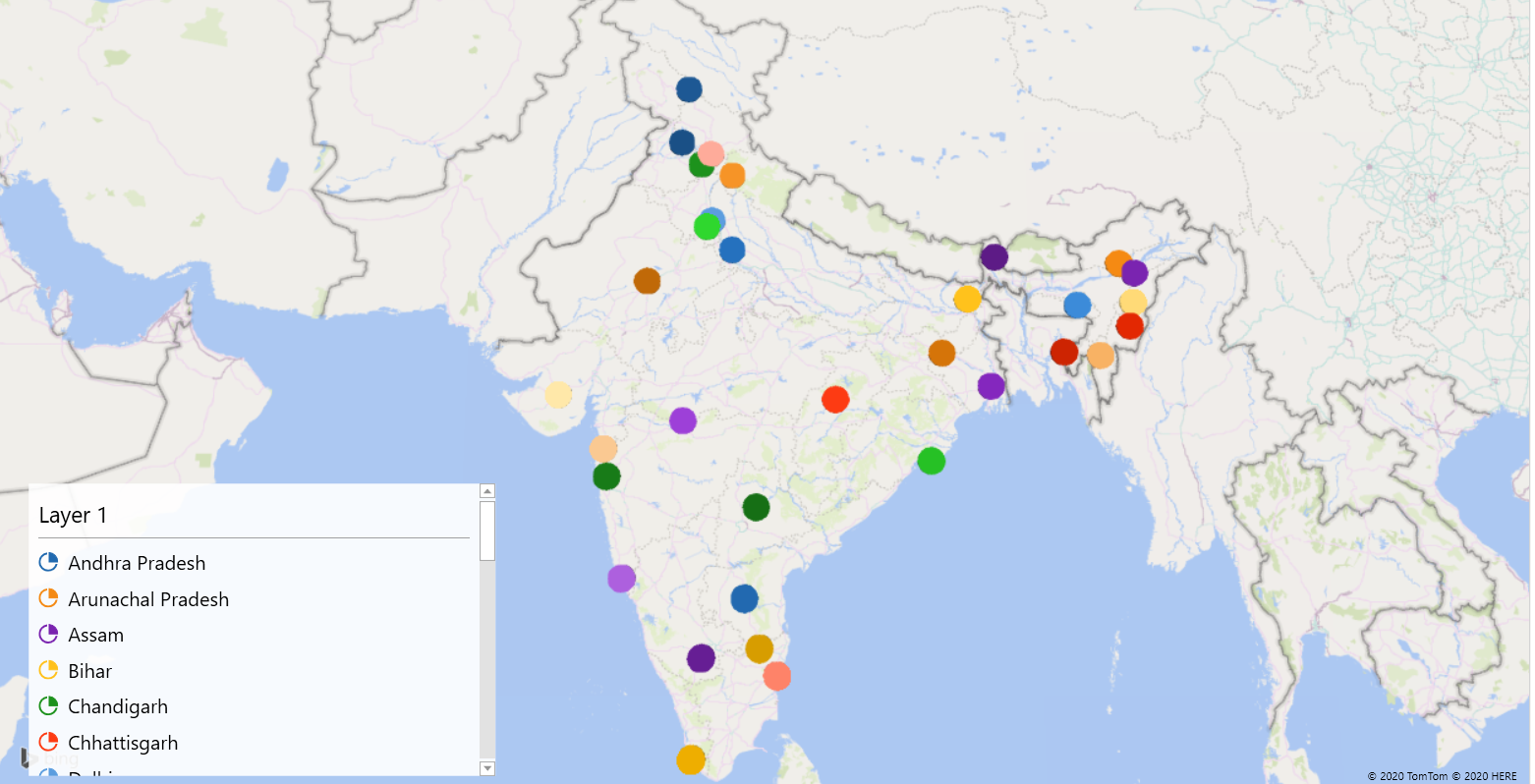
* Pivot table on Months
* Pivot table on Usage in different years
* Doughnut Chart

1. Analysis Results:

* In June month usage of power is started decreasing again July month started increasing power usage
* October and November are usually cloudless so that time the usage of Electricity is low.
* Again, July middle started increasing the power consumption

1. **3D Map**
2. Introduction: The analysis shows the relation between Latitude and Longitude in India.
3. Specific Requirements/Functions and Formulas:

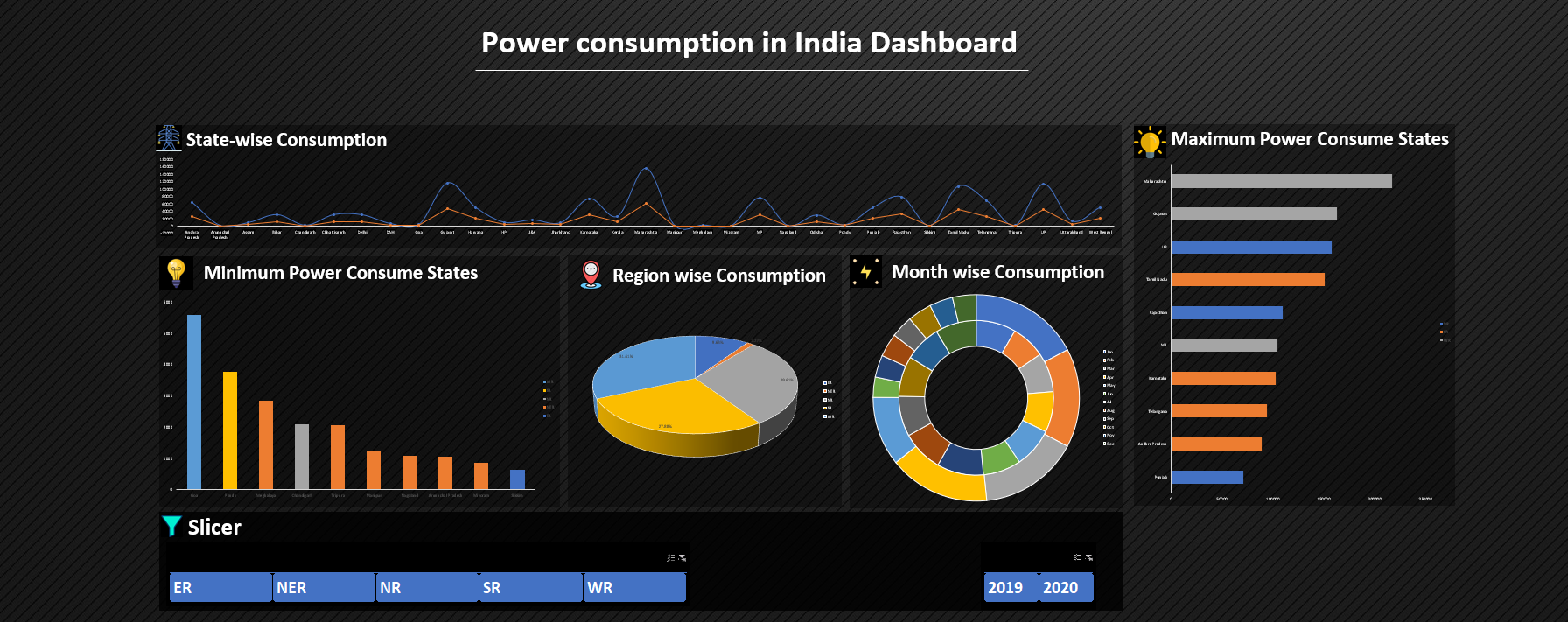
* 3D Map
* Longitude Data
* Latitude Data



1. **Hyperlinks and Slicers**
2. Introduction: Hyperlinks helps to redirect to the respective excel sheet and we have to select to which sheet it has to redirect.

* Slicers helps to change the data in pivot tables like selecting data according to our choice. In my data slicers are Year, Regions.

**Hyperlinks**



Slicers

**List of Analysis with Results**

* Maharashtra state consume 12.7% power. It is the highest Electricity Consumption State.
* Sikkim state consume 0.0379% power. It is lowest Electricity Consumption State.
* Over all North Region consumes high Electricity 27.3%, North east region Electricity Consumption is 21.2 %, South Region and West Region consume almost same Electricity 18.2%, East Region consume minimum Electricity 15.2%.
* In this pandemic time over all Regions Electricity Consumption is decreasing comparatively last year.
* In all Regions, when the summer is starting that time the usage of Electricity is started increasing middle of February the power usage started increasing.
* In June month usage of power is started decreasing again July month started increasing power usage
* October and November are usually cloudless so that time the usage of Electricity is low.
* In North region UP consume maximum electricity, Chandigarh Consume low electricity in north region.
* In West Region Maharashtra State consume maximum electricity.
* Again, July middle started increasing the power consumption
* In South region Tamil Nadu state consume maximum power and Pondy consume minimum electricity.
* In NER Assam Consume maximum electricity.
* Other state consumes very low electricity power (within 200 vol only)

**References**

1. [www.kaggle.com](http://www.kaggle.com)
2. [www.stackoverflow.com](http://www.stackoverflow.com)
3. [www.google.com](http://www.google.com)
4. [www.youtube.com](http://www.youtube.com)
5. [www.github.com](http://www.github.com)

**Bibliography**

1. Microsoft Excel 2016 Bible: The Comprehensive Tutorial Resource by John Walkenbach, Wiley.
2. Fundamentals of Business Analytics by R.N. Prasad, Seema Acharya, Wiley.