Data Visualization Assignment(Spatial Charts)

Dr. Ashish Anand **V.Saikiran(194161016)**

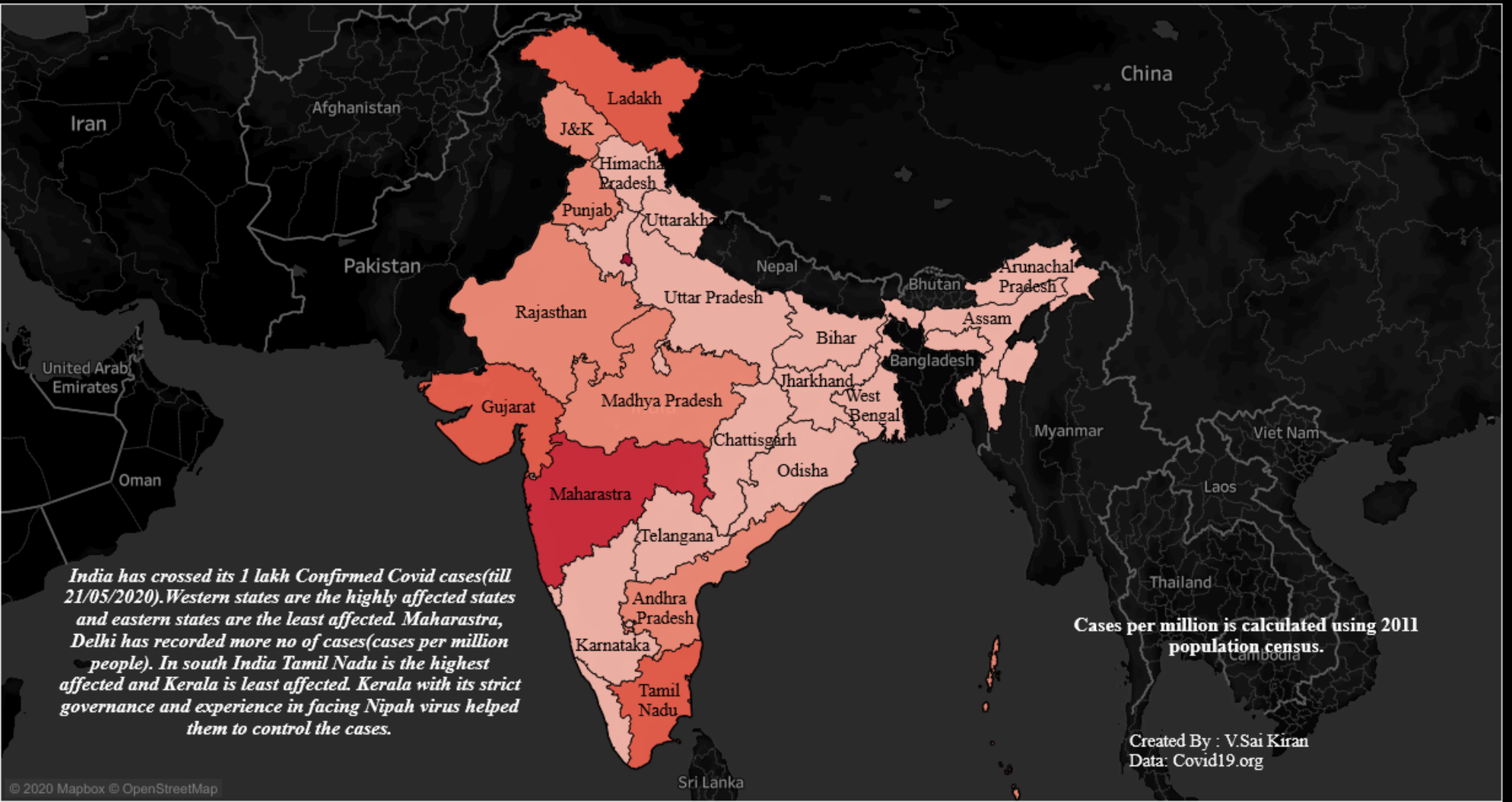
Spatial Charts:

Spatial family of charts allows us to visualize the data which is spread across **different geographical regions** across the world. These charts help us in **mapping spatial patterns** through **overlays** and **distortions.**

Let us discuss the following charts:

1. Choropleth Map
2. Proportional Symbol Map
3. Flow Map
4. Dot Map
5. Prism Map
6. Isarithmic Map

Indian states with their confirmed Covid Cases



Choropleth Map

- Choropleth maps are used to display a **data variable** associated with **distinct**,and defined **geographical areas or spatial regions.**
- Geographical area is well defined, such as **countries**, **states**, **counties** etc.
- The target measurement/variable is directly associated with and continuously relevant across the spatial region. In other words, the raw data should not be used and it must be **normalized** to reveal rates, aggregate statistics or ratios.

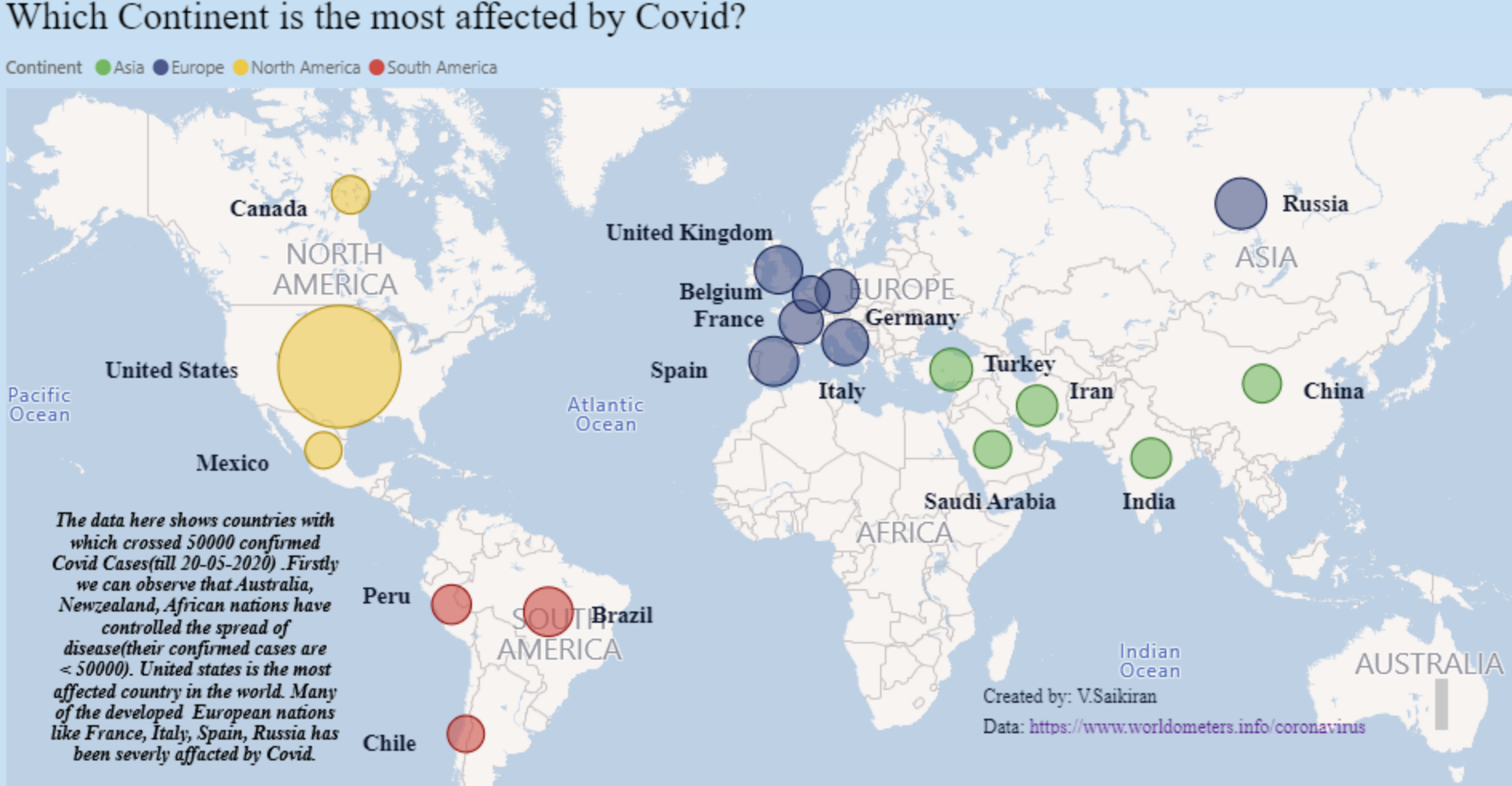
Data and source:
To demonstrate this map I have used the confirmed Covid cases for each in India. To normalize the raw them I have calculated and displayed no of confirmed cases per million. I have scraped the data from covid19.org

Tools used:
Excel and Tableau public

Analysis:

- As the population is **not evenly spread** across the states, comparing the raw data is not a good idea. So we will look into **no. of cases per million** instead of overall confirmed cases per state.
- **Delhi** being the capital city of India and most populated it has more no of cases per million.
- **Mumbai** the financial capital of India and one of the densely populated cities of India currently has the highest no. of cases in India. Its per million cases is greater than **600**.
- Reason for highest number of cases per million cases in Delhi and Mumbai can be
    1. Foreigners first had to visit Delhi or Mumbai as many flights from foreign countries come to either Delhi or Mumbai.
    2. Mumbai and Delhi have many **densely populated** areas which can lead to rapid spread of disease.
- **Western states** like **Rajasthan**, **Gujarat**, **Punjab** and central state **Madhya Pradesh** are highly affected by covid-19.
- **Assam, Arunachal Pradesh, Manipur, Mizoram** are the least affected states located in the eastern region of the country.
- Coming to Southern states **Tamil Nadu** has recorded **300-600** cases per million with major proportion of the cases in the capital city **Chennai**.
- **Kerala** is the state which recorded the **first Covid-19 case** in India. With strict **government policies** and past experiences with **Nipah virus** it has successfully controlled the spread of the virus.

Which Continent is the most affected by Covid?



Proportional Symbol Map

- Similar to choropleth map, **proportional symbol map** is also used to represent quantitative variables associated with a priory well defined **geographical areas.** Proportional symbol map uses proportionally sized shapes (**circles**, **bars**).
- **Proportional symbol maps are** useful when **raw data** needs to be visualized or when raw data cannot be transformed into ratio or proportion (a **requirement** for a **choropleth map**).

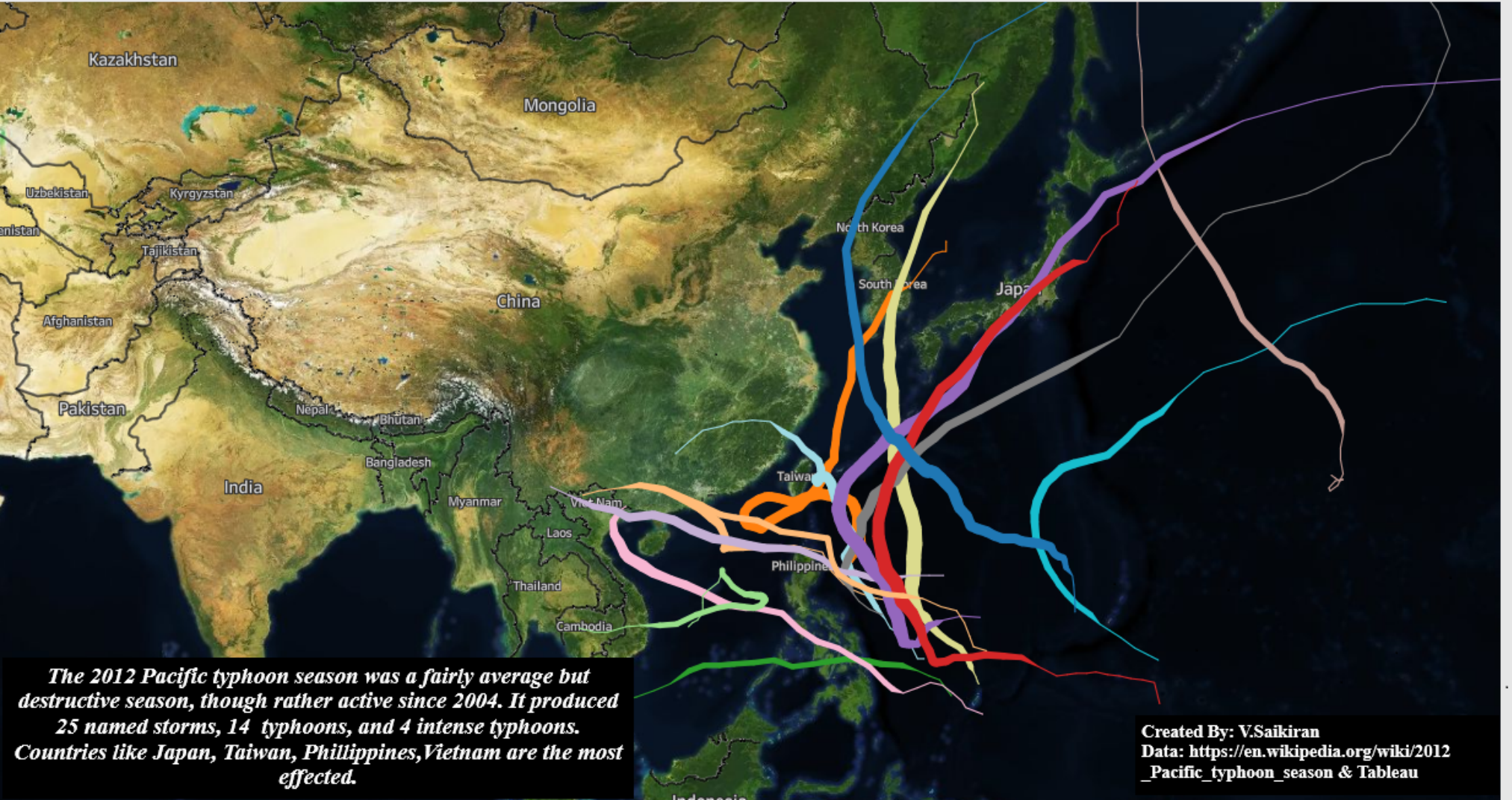
Data and source:
To demonstrate proportional symbol maps, I have considered no. of Confirmed Covid cases across different countries of the world. Data is scraped from worldometers.info.

Tools: PowerBi and Excel

Analysis:

- Covid-19 disease has spread to more than **180 countries** all over the world. Due to this most of the **businesses** across the world came to halt for about two months and it's still continuing.
- Let us look at the most affected countries with more than **50,000 confirmed cases.**
- Firstly, the most powerful country **the United States** has more no. of confirmed cases. People from different parts of the world visit the United states for **educational purposes, tourism, business, jobs** etc. which lead to the spread of disease.
- **The United states** has also got the **densely populated** regions like **New Jersey, Massachusetts, New York** which has recorded more no of cases.
- **Europe** is the continent which is highly affected by Covid-19. We can observe out of **18 countries** which have more than 50,000 confirmed cases **7 countries are from Europe.**
- Developed countries like **France, Italy, Spain, Germany, Russia** have been hugely affected by Covid-19.
- 5 of the Asian countries have made it to our list. **India**, with its **precautions** and **lockdowns** was able to control the spread of the disease despite its **huge population**.
- We can see that none of the **African**, **Australian** countries have cases more than **50,000.** But there is no valid source that says the cases in **Africa** are **contained or unrecorded.**
- **Australia** has taken several **preventive measures** to control the spread of the disease in the country.

2012 Pacific typhoon season - Storm paths and wind speeds



The 2012 Pacific typhoon season was a fairly average but destructive season, though rather active since 2004. It produced 25 named storms, 14 typhoons, and 4 intense typhoons. Countries like Japan, Taiwan, Phillippines,Vietnam are the most effected.

Created By: V.Saikiran
Data: https://en.wikipedia.org/wiki/2012_Pacific_typhoon_season & Tableau

Flow map

- A flow map shows the **characteristics of the movement** or **flow** of a phenomenon across **spatial regions**.
- It is often formed using **line marks** to map flow and combinations of attributes to display the characteristics of this flow.
- The **quantity of flow** can then be represented by the **width** of the line.
- Flow maps generally display characteristics of **origin** and **destination (positions on a map)**, **route, directions (using arrow or tapered line width)** and **categorical classification (colour)**.

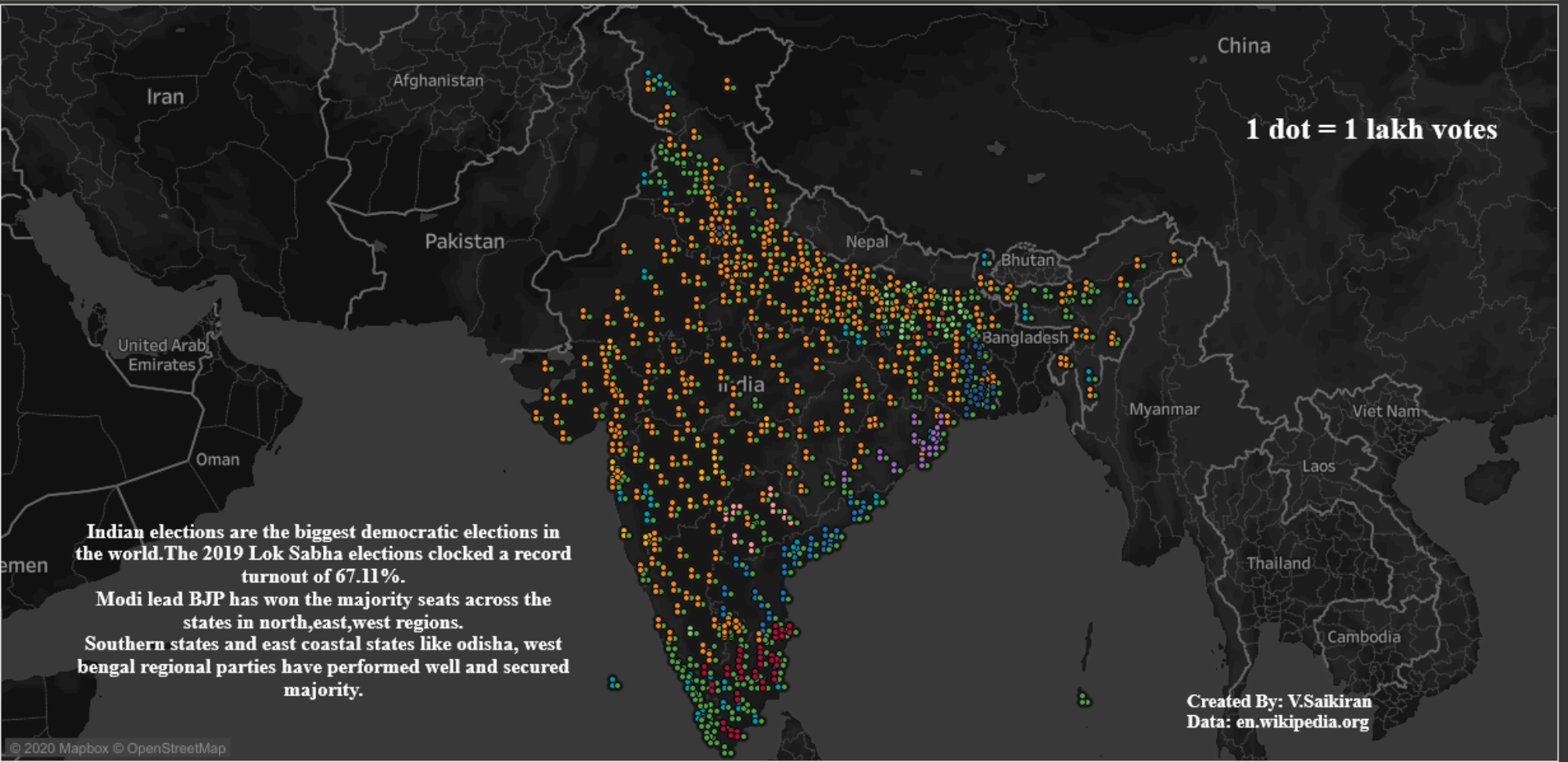
Data and source: To explore this map I used the data for 2012 Pacific typhoon season available on wikipedia page. It has got clear explanations of how and where the storms have started and ended.

Tool : Tableau public

Analysis:

- The **2012 Pacific typhoon season** was a fairly average but **destructive** season, though rather active since **2004**.
- It produced **25 named storms, 14 typhoons, and 4 intense typhoons.**
- It was an event in the annual cycle of tropical cyclone formation, in which tropical cyclones form in the **western Pacific Ocean.**
- From the map we can see that storms start in the ocean and travel all the way through the ocean onto the **eastern countries** and go back into the ocean.
- Eastern countries like the **Philippines, Taiwan, Japan, Vietnam, South and North Korea** experience storms throughout the year.
- Flow maps help us to analyse the path of the storm and the pressure it had at different locations.
- For example **Jelawat** is the longest amongst the storms. It had huge pressure since its origin and continued the same pressure onto the **philippines, Japan** and then the wind pressure has decreased.
- We can see that **Gaemi** is the shortest storm, which had a very less wind pressure.
- Storms like **Guchol, Bolaven, Tembin, Sanba** has maintained almost the same kind of wind pressure throughout their travel.
- The **lengths** of the lines tells us how **long** the storm existed. The **width** of the line explains to us what was the **wind pressure** with which it travelled.

Indian States and Lok Sabha Election Results 2019



Dot Map

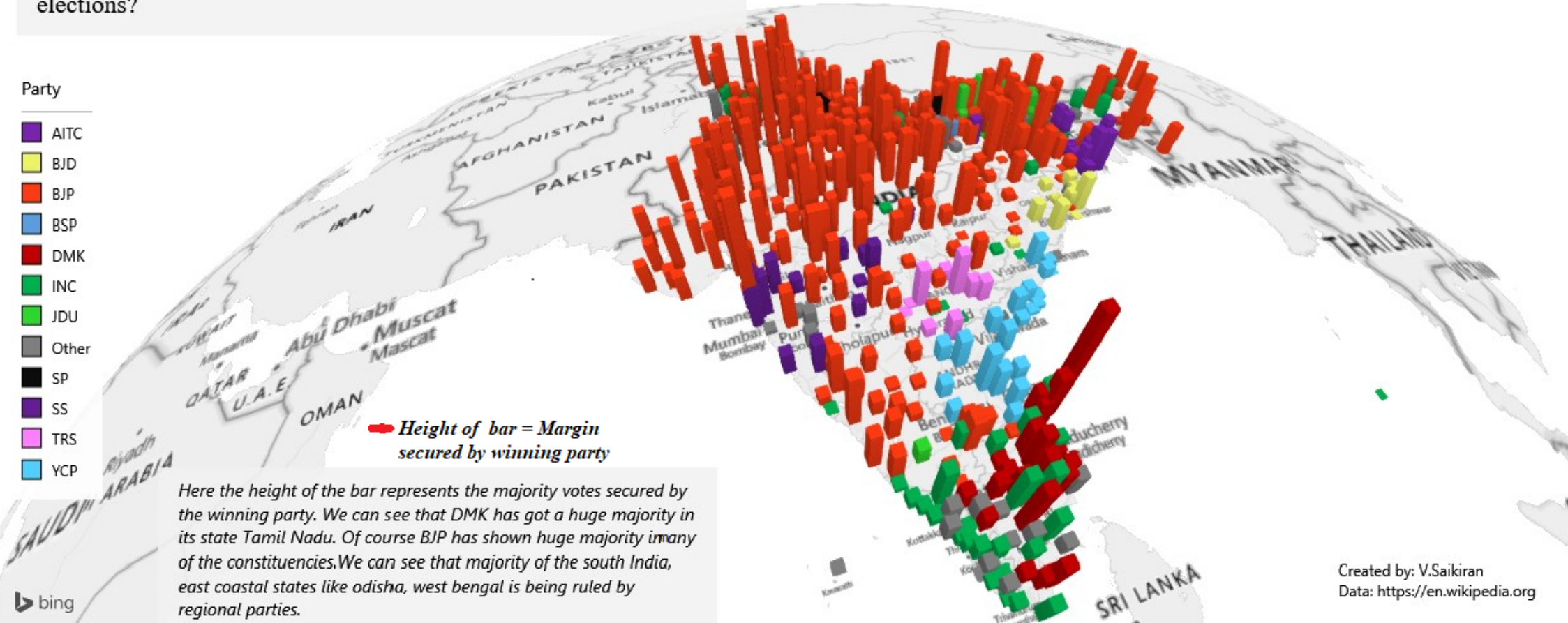
- Dot Map is used to display spatial distributions or densities of data representing a large number of single objects.
- Equally sized dots are used in the visualization. There are two types of Dot Maps:
    - **one-to-one:** one dot or point represents a single count or object,e.g. 1 dot = 1 tree;
    - **one-to-many**: one dot or point represents a particular unit of the object, e.g. 1 dot = 100 people.

Data and source: To explore this map I have selected the Indian Lok sabha elections 2019. Data source : en.wikipedia.org
Tools : Tableau and excel

Analysis:

- **Indian elections** are the **biggest democratic elections** in the world. Many interesting patterns can be found by exploring the data visually.
- For the sake of demonstration I have generated **20-25%** of the data along with the **real data.** Here we deal with a **one-many** kind of relationship where **one dot represents 1 million(approximated for sake of demonstration) votes.**
- If we observe carefully most of the constituencies are filled up with votes favoring **BJP**.
- As **Indian National Congress** and **BJP** are national parties most of the states have some vote share for these two parties.
- BJP has major vote share in all parts of the country except **southern states** and **east coastal states.**
- In states like **Odisha**, **West bengal, Telangana, Andhra Pradesh, Tamil Nadu,** people have more faith in **regional parties**. Each state has their respective regional parties having major vote shares.
- Dot maps give the privilege of observing the **density** of a data variable by adding a **new dimension** which can differentiate dots using **colors**. In our case **different colors** indicate **different parties.**

Whats's the margin secured by winning parties in Lok Sabha 2019 elections?



Prism Map

- **Prism Map** generates a **3d view** corresponding to a **quantitative** variable over geographical regions.
- The quantitative values are represented proportionally in the form of **lines** or **bars**.

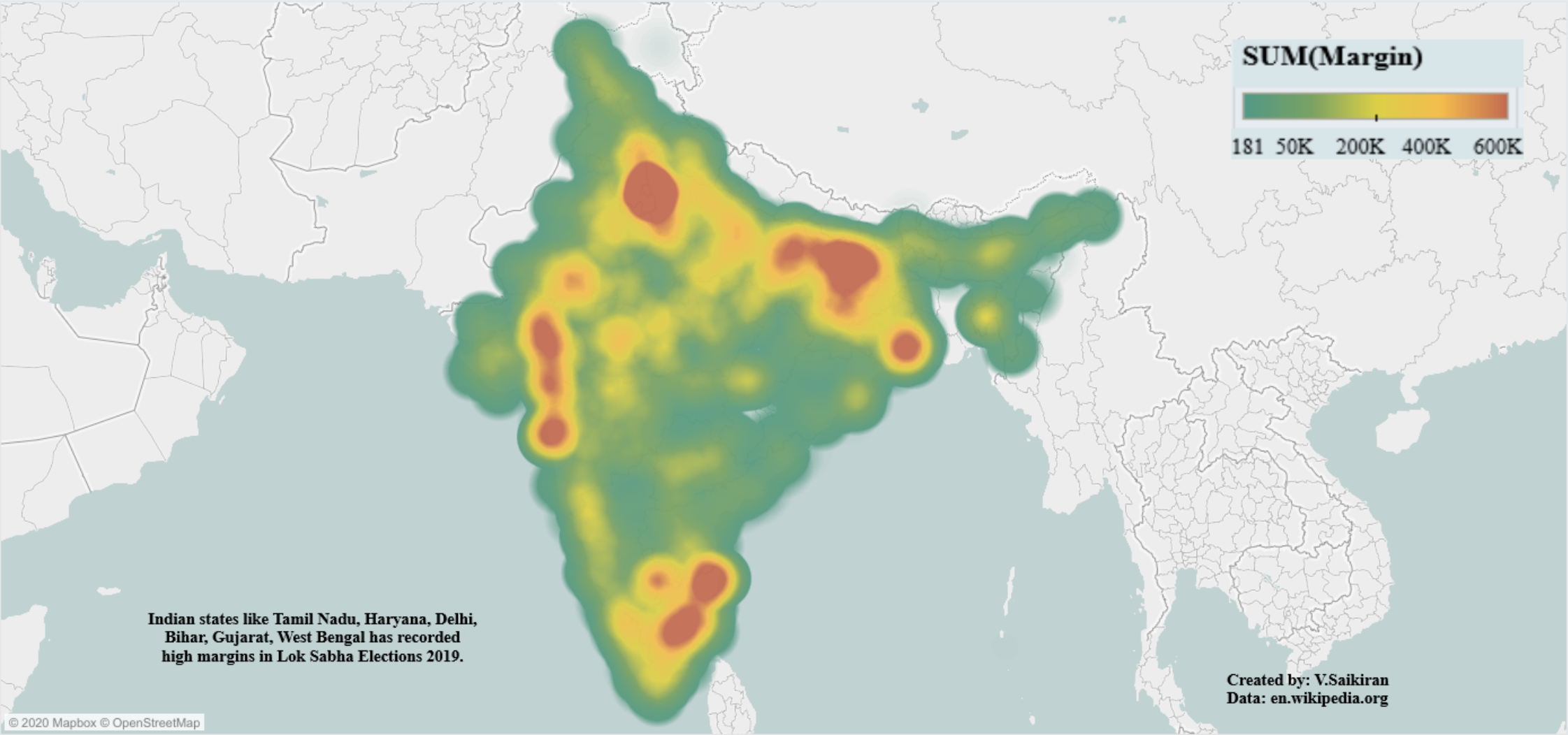
Data and source: To explore this map I have selected the Indian Lok sabha elections 2019. Data source : en.wikipedia.org

Tool : Excel 3d charts

Analysis:

- Till now we have looked into lok sabha elections on the basis of vote share by different parties. We can add one more dimension called **margin(difference between winning party votes and runner up party votes)** to the data and we can enhance the visualisation a bit more using a 3d prism map.
- The **height** of the bar represents the **margin** secured by the **winning party**. In the map we can see that **BJP** has secured high margins in regions of **Haryana**, **Delhi** and **Gujarat**.
- **DMK**, a regional party in **Tamil Nadu**, has secured the **highest margin in the southern states.**
- We can see that in the **northern states BJP** has secured a very good margin in many of the constituencies.
- Plotting a prism map in 2D view has a lot of **drawbacks**
    1. **Smaller values** will be hidden behind the **larger forms**, just as **smaller buildings** are hidden by **skyscrapers** in a city
    2. If the data that we are dealing with spreads across the whole world, we may not be able to see the other side of the globe.
    3. The Height of the bars just allows us to understand that one bar has got more height than the other. But we may not be able to **quantify** the **differences** in heights of the bar numerically.
- Ideally prism maps would be provided with **interactive features** that allow panning around the map region to offer different viewing angles to overcome the perceptual difficulties of judging the dimensions of 3D forms in a 2D view.

Visualizing margins secured in Loksabha Elections 2019 using Heat maps



Geographical Heat Map(Isarithmic Map)

- Isarithmic Map displays distinct spatial surfaces on a map sharing the same quantitative classification.
- In contrast to the Choropleth Map, spatial boundary is not defined by geopolitical boundaries, rather it is regions sharing a certain quantitative value or interval scale.
- These regions are formed by connecting points of similar measurement.
- The relevant quantitative values are colour coded and regions are encoded with the respective colours.

Data and source: To demonstrate this map I have used the Lok sabha elections 2019 data available on wikipedia page.

Tools: Tableau public and Excel

Analysis:

- **Geographical heat maps** are one of the variants of **Isarithmic maps** which helps us to visualize **spatial regions** that **share** the same range of quantitative values.
- Till now we have seen the **Lok sabha election 2019 data** in two forms. One is using **dot maps**(visualized the **winning party** in different constituencies), **prism maps**(visualized the **margins** secured by the winning parties with the help of height of the bars).
- Now the drawback we observed in case of prism maps is that in 2 dimensional views, bars with **low heights** are covered up by the **larger height bars**. And also we were not able to say the quantitative difference in values of the bars.
- We can visualize the **like margin valued** constituencies with the help of **heat maps.** In the chart we observe that a major part of the country has secured a margin under **200K**.
- In the states of **Tamil Nadu, Haryana, Gujarat, Delhi, Bihar, West Bengal** there are **red regions** where the sum of margins are as high as **600k**.