

DV Project Final Submission

Through our Data Visualisation Project, we are analysing the Database of the Judiciary Cases in India but mainly depicting the distribution of judges across gender and their trends with certain factors, for example such as case duration.

1. India Chloropleths/Bivariate Chloropleths

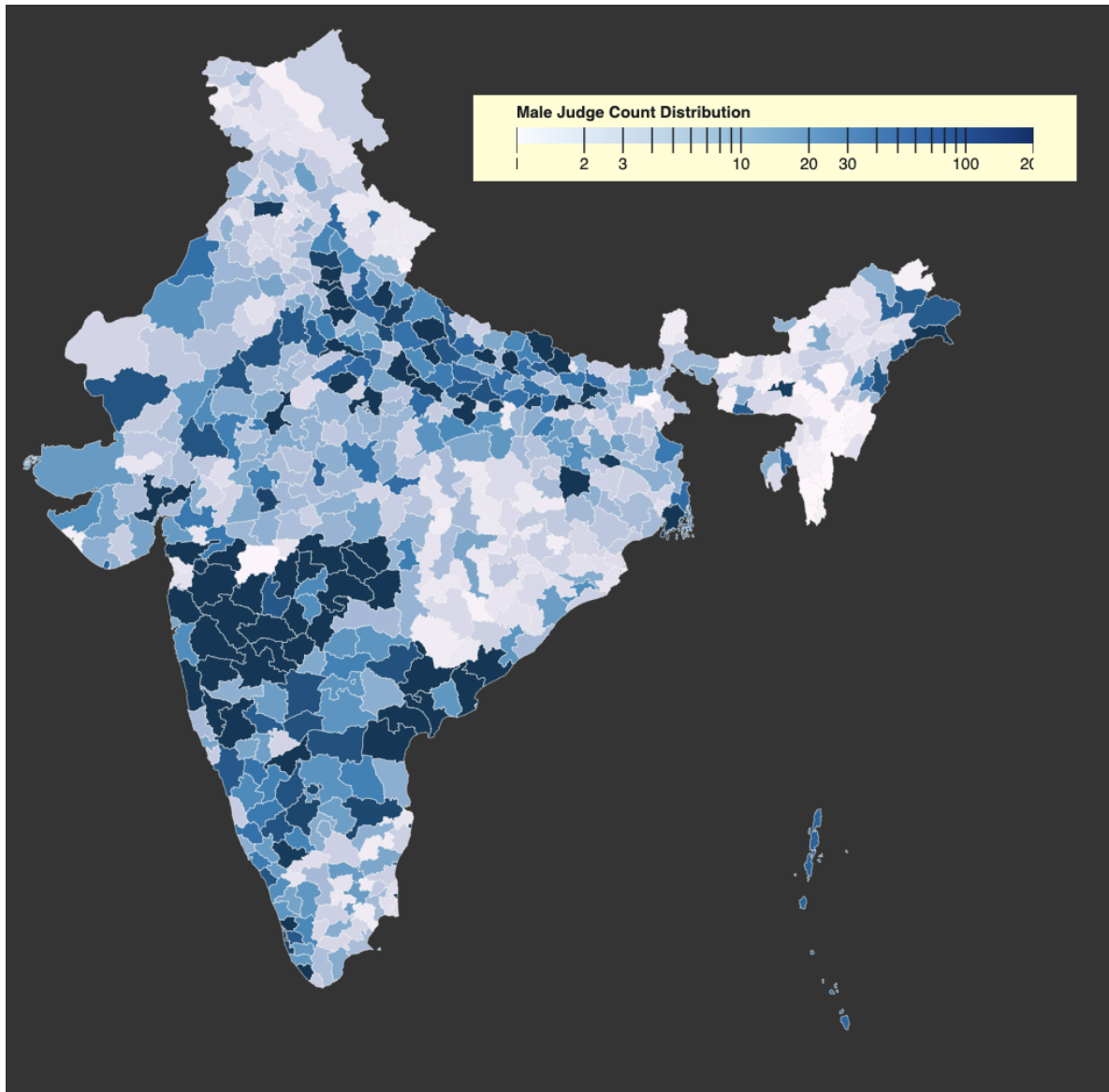
Chloropleths and chloropleth animations make excellent data visualizations because they allow for the representation of data at a geographic level. This can be particularly useful when looking at demographic or other data that varies significantly depending on location. Chloropleth animations, in particular, are an effective way to show changes in data over time, allowing viewers to observe trends and patterns that may not be immediately apparent from static visualizations. Additionally, tooltips and other interactive features are added to chloropleths to allow for more detailed exploration of the data with identifying the specific districts. In this way, it helps tell our story on the Gender of Judges and associated relationships between time taken for cases in India.

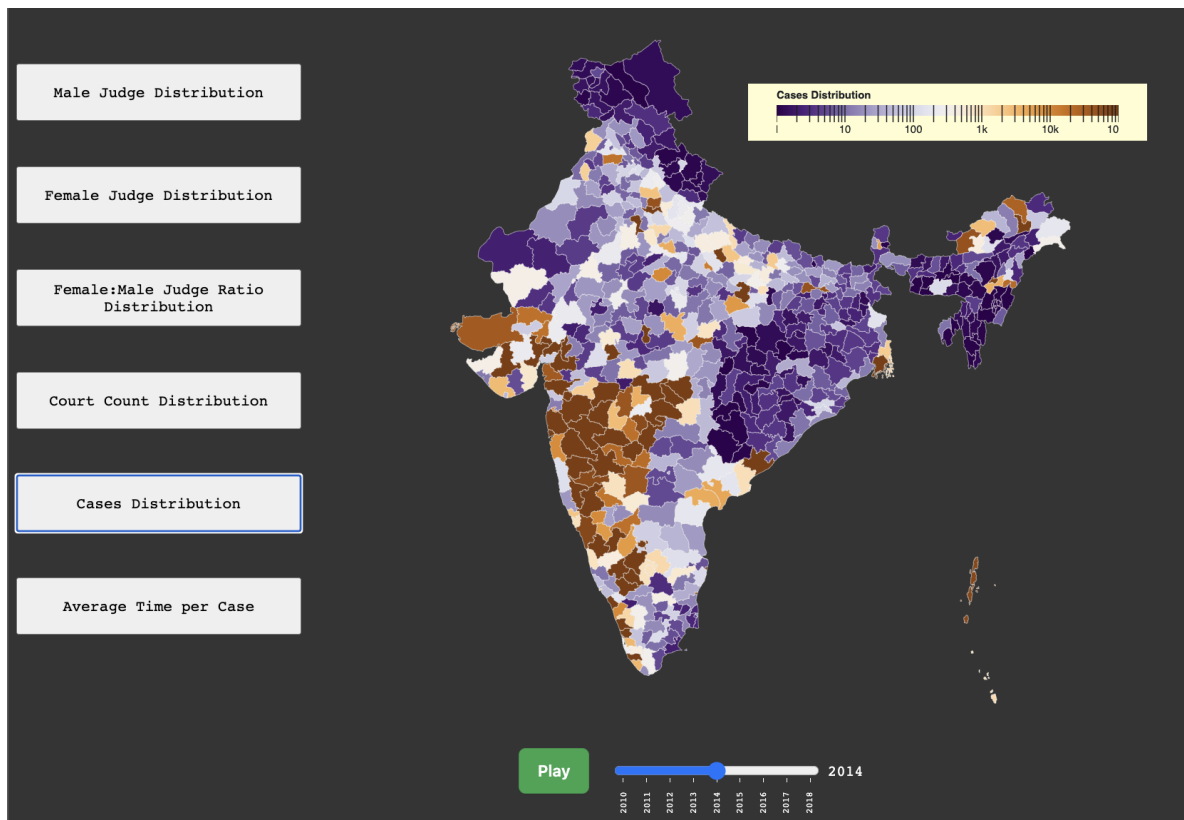
From the Mid submission, I have completed the animations which provide life to showing the change of time from 2010-2018 on the metrics of cases and average length of cases. In this animation, I

have provided options to the user to play/pause and also a timeline which they can directly click on a year (which also pauses the animation) to see that year's data more clearly. This animation helps analysts better find the trends they would need for study and also for general users to get a good feel of the data with its order/chaos and general trends.

I also made a system on the left side of the screen where the user can select the chloropleth visualization of the following: (Green for completed)

- Chloropleth for the Distribution of male judges across India
- Chloropleth for the Distribution of female judges across India
- Chloropleth for the Distribution displaying both male and female judges across India.
- Chloropleth for the Distribution of courts across India
- Chloropleth for the Distribution of cases across India
- Chloropleth for the Distribution of average case times across India





There were 2 parts to the complete dataset required for this visualization: Dataset of topography of India and Data of mapping each District to a Value. While the paths formed by the dataset of India's topography, the colors of each district are then referenced to the value dataset. This was problematic as for many of the datasets, as they were made in different time periods, there were different numbers of districts and they were not numbered in the same manner. In any case, the districts of the datasets were manually fixed in order to create the accurate choropleths. For the bivariate choropleths discussed above, they would be the combination of their according 2 distributions.

2. Bar Chart

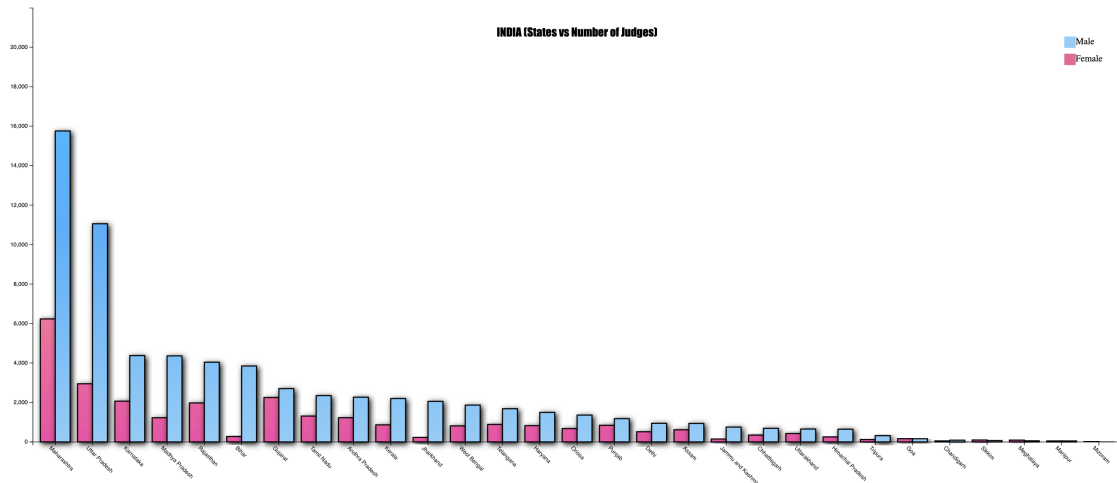
WHAT This visualisation displays the distribution of non-female and female judges across each state in India. Clicking on a specific state bar reveals another bar graph, which showcases the number of non-female and female judges in each district of that state. This feature provides a more detailed view of the distribution of judges within the state.

Each bar represents a state(or a district), and the height of the bar represents the total number of non-female and female judges in that state(or a district). The bars are colour-coded to represent the gender of the judges - blue for non-female judges and pink for female judges. The colour-coding allows the viewer to quickly identify which gender of judges is more predominant in each state. Clicking on a specific state bar reveals another bar chart.

WHY The reason for choosing this visualisation is to clearly show the comparison between the number of non-female and female judges. By presenting this information in a bar chart format, we can easily compare the number of judges in each state and district and observe any differences in gender representation i.e it allows us to identify any disparities in gender representation within the judiciary system. This information can be used to initiate efforts to address any gender imbalance in the recruitment and appointment of judges.

HOW We made this visualisation using d3.js. To generate this visualisation, we had to use multiple CSVs. We first combined the data from judges_clean.csv and cases_district_key.csv using the common variables dist_code and state_code. This resulted in a new CSV file with the following columns: dist_code, state_code,

female_judge from judges_clean.csv, and state_name, district_name from cases_district_key.csv.



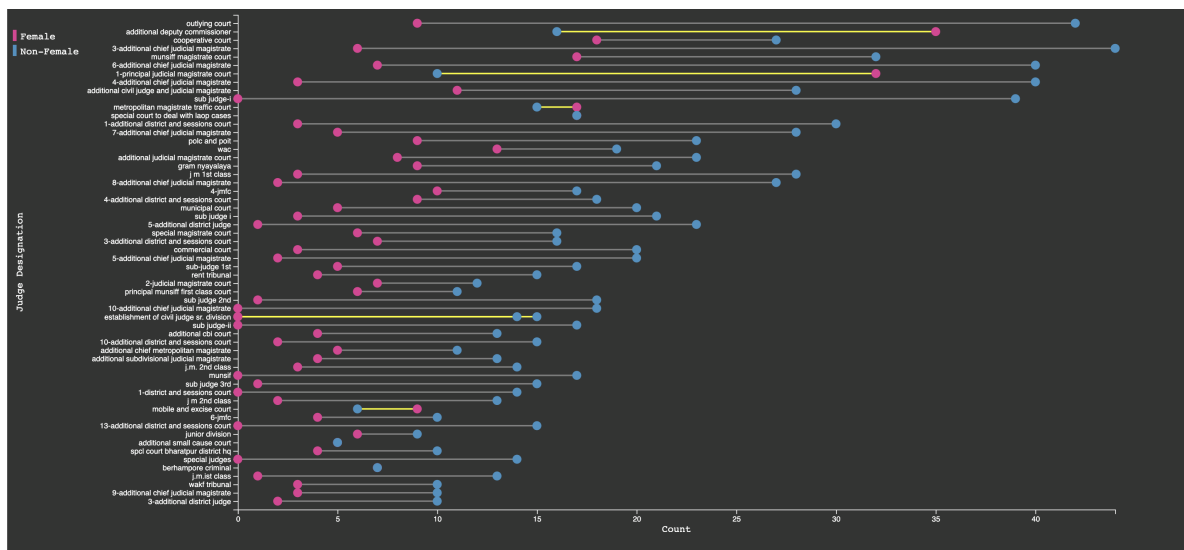
3. Cleveland Dot Plot

WHAT In this visualisation, we show the distribution of non-female and female judges by their designation. The designation of judges will be represented on the vertical axis, and the number of judges will be represented on the horizontal axis. We will also use two sides of the dot plot, one for non-female judges and the other for female judges. Both dots are of a different colour, to easily identify the difference between the number of non-female judges and the number of female judges.

WHY The Cleveland dot plot is an effective visualisation technique for showing the distribution of data across categories, and using two sides of the dot plot for male and female judges allows us to compare their representation across different designations. This visualisation

will help us gain insights into the representation of non-female and female judges in different designations and it will help us identify any potential gaps in gender diversity among judges.

HOW We made this visualisation using d3.js. To generate this visualisation, we used “judges_clean.csv”. Using this CSV, we get the number of Female and Non Female Judges in different designations. The connected scatter plot is an effective visualization technique for showing how two variables are related over time. By connecting the points in the plot, it allows the viewer to see the trend or pattern in the data. In this case, we are interested in showing the trend in the number of female and male judges over time. Additionally, the use of two lines for male and female judges will help us compare their representation in the judiciary.



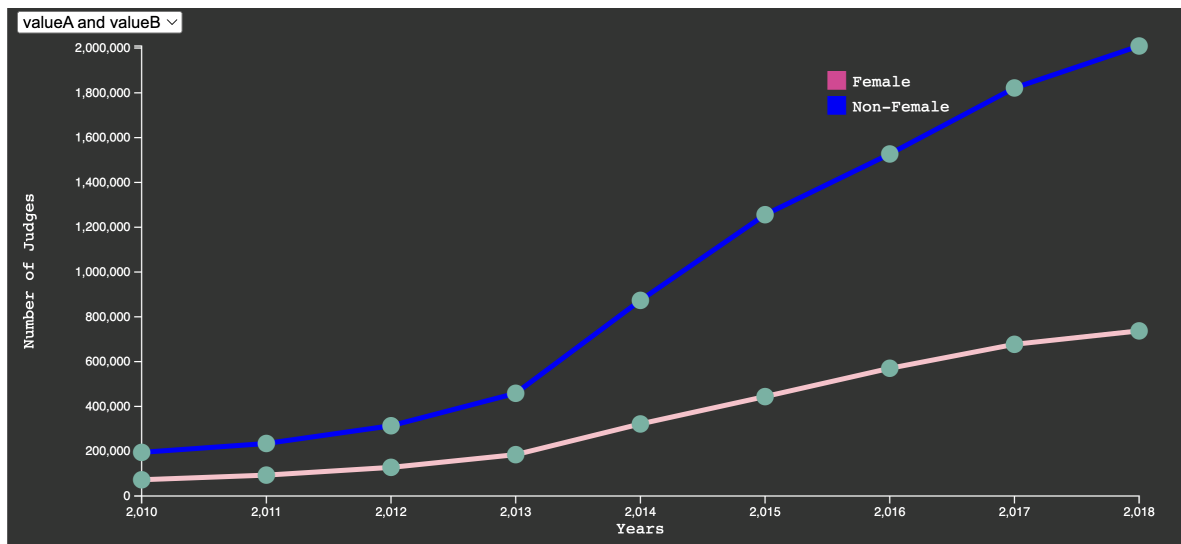
4. Connected Scatter Plot

WHAT A connected scatter plot, shows the relationship between two variables, i.e., here it shows the year on the horizontal axis and the

number of judges on the vertical axis. This visualisation will have two lines, one for non-female judges and the other for female judges, to compare their representation over time. Initially the number of female judges is plotted. Using the drop down we can select to see either the number of female or non-female judges or both over time. The drop-down button provides users the flexibility to choose the line they want to view.

WHY Here, the connected scatter plot is an effective visualisation technique for showing how the number of judges and time are related over time. By connecting the points in the plot, it allows the viewer to see the trend or pattern in the data. We are interested in showing the trend in the number of female and non-female judges over time. Additionally, the use of two lines for non-female and female judges will help us compare their representation in the judiciary and how the number has changed over time. Overall, this visualisation will help us gain insights into the gender diversity and trends of judges over time.

HOW We made this visualisation using d3.js. To generate this visualisation, we used multiple CSVs. We first combine data from “judge_case_merge_key.csv”, “cases_[year].csv” where years range from 2010 to 2018, and “judges_clean.csv”. In the end we will have the number of Female Judges and the number of Non Female Judges for the years 2010-2018.



5. Diverging Stacked Bar Chart



This visualisation basically focuses on the difference between the time duration spent by male/ female judges on a particular position. The left side represents the males and the right side represents the female time spans. Also, The bars represent the different judge positions. As we can clearly see, in some of the positions, females spent more time as compared to men but usually, time spent by the men is higher. We can also see that the number of women judges

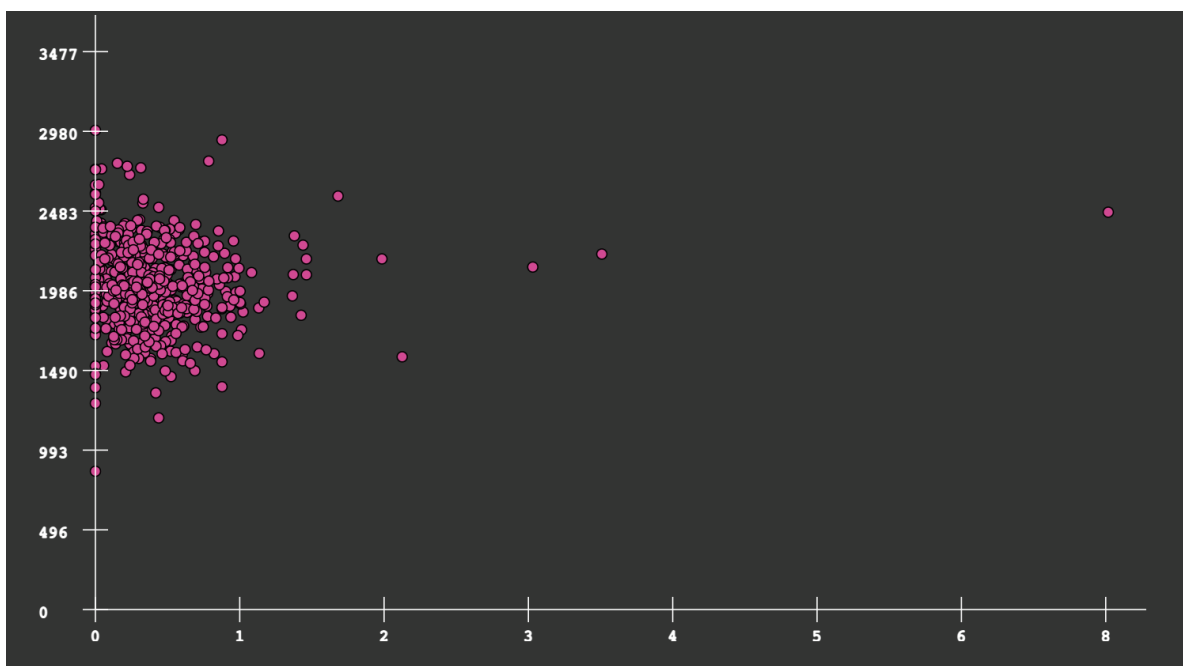
are lesser as in many of the positions, the time spent by women is zero. This clearly indicates the male dominance in the judges.

Every bar is clickable. The total number of judge positions where a lot so, I combined few judge positions in a single bucket and made 10 buckets in total. And these are the ones which are shown in the bar graph. Now, if you click on these bars, it will open up the other bar chart with the ones grouped earlier. the judge positions are sorted alphabetically in order to facilitate easier searching.

To generate this visualisation, we take data from the file `judges_clean.csv`.

6. Scatter Plot

This visualization shows how the average time of a case varies with the ratio of female judges in every district. The X axis marks the ratio of number of female judges to the number of male judges. The Y axis marks the number of days. Every dot represents one district.

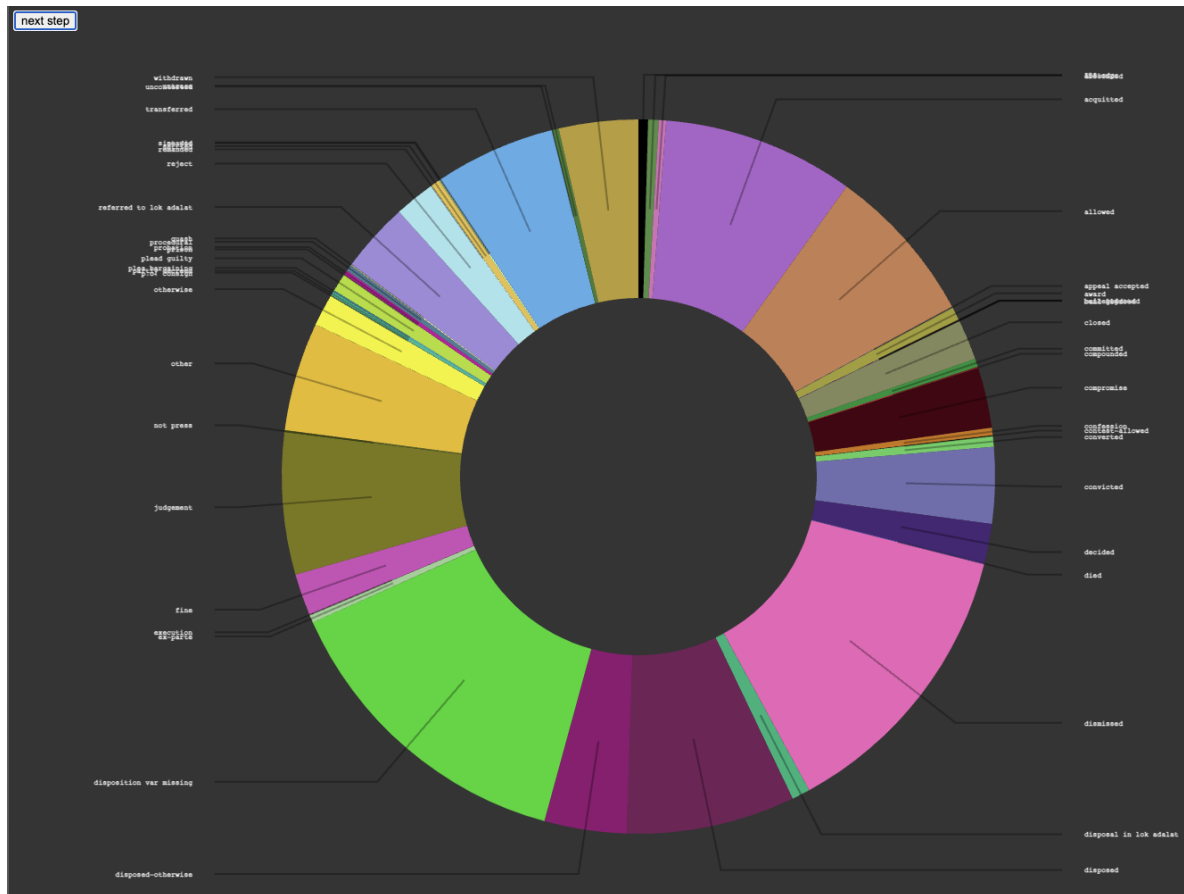


The Scatter Plot visualization has been made, which shows how the average time of a case varies with the ratio of female judges in every district. The X axis marks the ratio of number of female judges to the number of male judges. The Y axis marks the number of days. Every dot represents one district. This visualization will help us gain insights into the effect of gender diversity among judges on the time taken for cases to be resolved in different districts of India.

7. Pie Chart

This was a minor visualization that shows the distribution of most frequent tags associated to each case per year. Within the dataset, there is a file in the keys folder, `disp_name_key.csv`, that is used to reference the nature of the crime's result. This data is split between 50-60 terms per year for every year between 2010-2018. Each term has a frequency which will be the data shown. A pie chart has been made that represents these frequencies to show the distribution of the terms over the years. This has been done through the use of an animation through updates of the pie chart.

A pie chart is a great visualization because it allows viewers to quickly see the proportion of different categories within a dataset. The use of colors and slices makes it easy to visually compare the size of different categories and understand the overall composition of the data. Additionally, pie charts are easy to create and can be used to convey complex data in a simple and concise way. They are particularly useful when trying to show the relationship between different parts of a whole.



Conclusion

We have successfully completed all of the required visualizations for the Data Visualization Project. All of the visualizations have been created using d3.js and are available in the project submission. The visualizations are diverse in nature and use different techniques to depict the data in the most effective way possible. The visualizations provide valuable insights into the gender representation and trends of judges in India, and

can be used to initiate efforts to address any gender imbalance in the recruitment and appointment of judges.

Statements

I, Anuhya Nallapati (2021101076), did the Bar Chart, Cleveland Dot Plot, and the Connected Scatter Plot and wrote the code to make the CSVs for these visualisations.

I, Harshit Aggarwal (2021111015), made the Diverging Stacked Bar Chart, Scatter Plot and made the csvs for most of the visualisations. For making the CSVs, I had to use concepts like DNA sequencing Global alignment, rule based code, etc.

I, Shreyas Reddy Palley (2021101016), made all the Chloropleth Charts of India, including the: Male Judge Count Distribution, Female Judge Count Distribution, Female-Male Judge Count Distribution, and Court Count Distribution as well as the animations of Cases Count Distribution and Average Case Length Distribution. I also made the Pie Chart Visualization. I styled all of the charts/graphs and added the Legends to the Chloropleths.