Syracuse University
College of Engineering
& Computer Science

CIS 500: Data Analysis and Decision Testing

US Elections Result Analysis from 1976 to 2020

Ayush Saxena(573647550)
Brijesh Verma(826257344)
Hemil Shah(610010477)

Presentation Outline

- Problem Statement
- Significance
- Problem Description
- Development Methodology
- Results
- Conclusion
- Future Research

Problem Statement

Understanding the dynamics behind USA election results is crucial for political strategists, policymakers, and the public alike. With the vast amount of data spanning from 1970 to 2020, there's a pressing need to dissect and analyze these elections comprehensively. However, the complexity of factors influencing election outcomes presents a formidable challenge. Identifying these factors and understanding their interplay is pivotal for enhancing our understanding of the electoral process.

Significance

- **Informing Political Strategy:** By delving into past election data, we can uncover patterns and trends that inform political strategies. This understanding allows for more effective campaign planning and resource allocation.
- **Enhancing Public Understanding:** Providing insights into the factors shaping election results fosters a better-informed electorate. Understanding the dynamics of elections empowers voters to make more educated choices.
- **Predictive Analysis:** Developing a methodology for predicting future election outcomes based on historical data enables stakeholders to anticipate electoral trends. This predictive capability is invaluable for shaping political campaigns and policy agendas
- **Fostering Democratic Participation:** Through comprehensive analysis, we aim to encourage greater participation in the democratic process. By elucidating the factors influencing election outcomes, we empower citizens to engage meaningfully in the electoral process.
- **Academic and Policy Relevance:** This analysis contributes to academic research and informs policy discussions surrounding electoral reform and governance. By shedding light on past election trends, we can guide future policy decisions aimed at strengthening democratic institutions.

Problem Description

The aim is to perform exploratory data analysis (EDA) on the US presidential election

data from 1970 to 2020, consolidated from all states.

The primary objectives are:

1. Analyze the election results across different political parties over the years.

2. Investigate the state-level voting patterns and trends. We chose this analysis because we wanted to cover most interesting history of USA Elections and to see behaviour of several states over the years and to determine which are the **SWING** states.

3. To look out for the most famous individual who stood in the USA presidency candidacy history.

Existing Methodology

The major part of the problem we experienced is that there are already many hypothesis done on USA Elections, but all of them were mostly targeting one election span say 2016 or 2020. This gave us a reason to demonstrate our data handling capabilities and to do an extensive research on the history of USA Elections spanning 50 years. There is nothing wrong with one year analysis but it would not provide a clear picture of how demographics have changed over the years which is the sole purpose of our analysis.

Developed Methodology

1. Data Preprocessing

- Load and preprocess the US presidential election dataset, handling any missing values or inconsistencies
- Read data from CSV file and performed cleaning and transformation

2. Exploratory Data Analysis

- Used EDA Functions to compute summary statistics, generate visualizations (e.g., histograms, scatter plots), and identify patterns or outliers in the data.
- Visualizing the data using various plots and charts (e.g., line plots, bar charts, scatter plots)

Developed Methodology

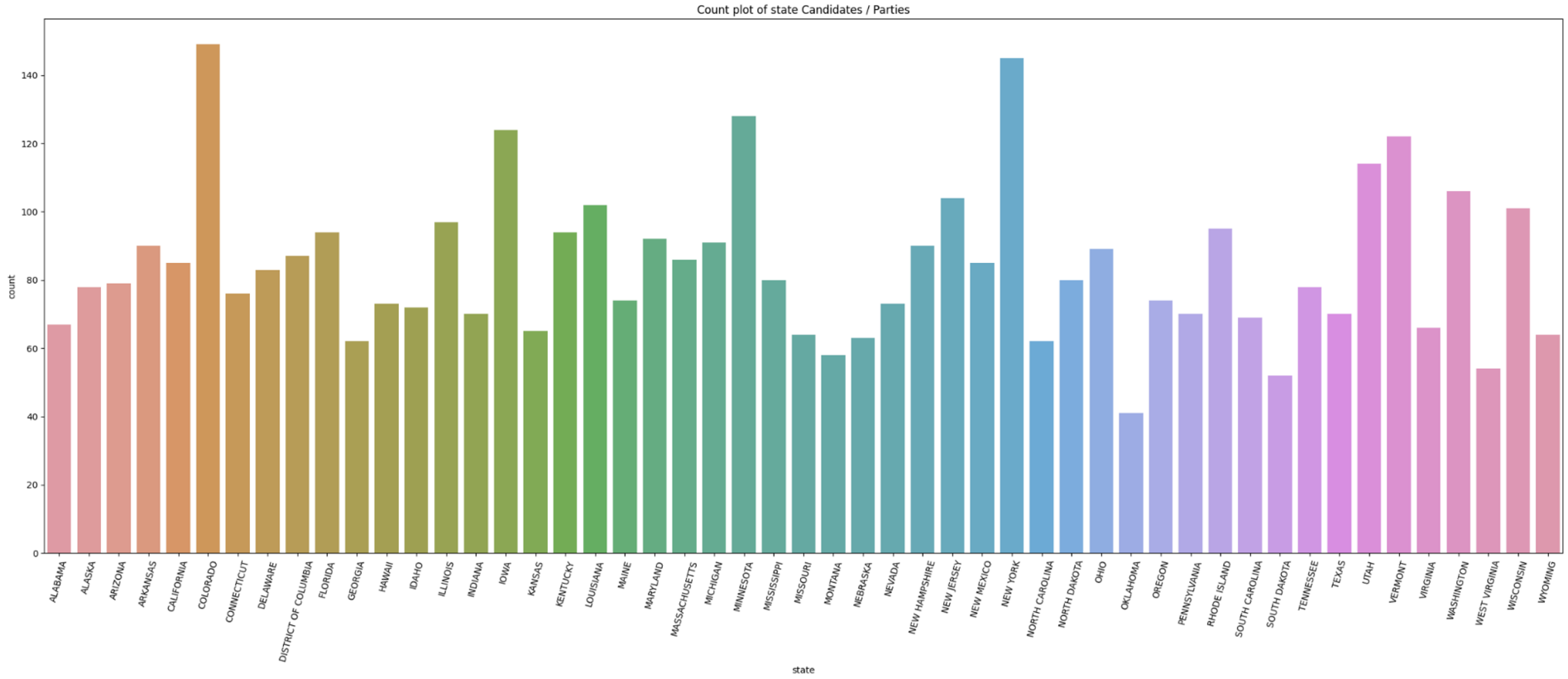
3. Data Visualization

- Leveraged libraries like Plotly and BokehJS to create interactive and animated visualizations, enabling users to explore the data dynamically
  - Line plots to depict the vote share of parties over time
  - Choropleth maps to visualize state-level voting patterns
  - Animated plots to showcase the evolution of voting trends across elections

4. Metric Analysis

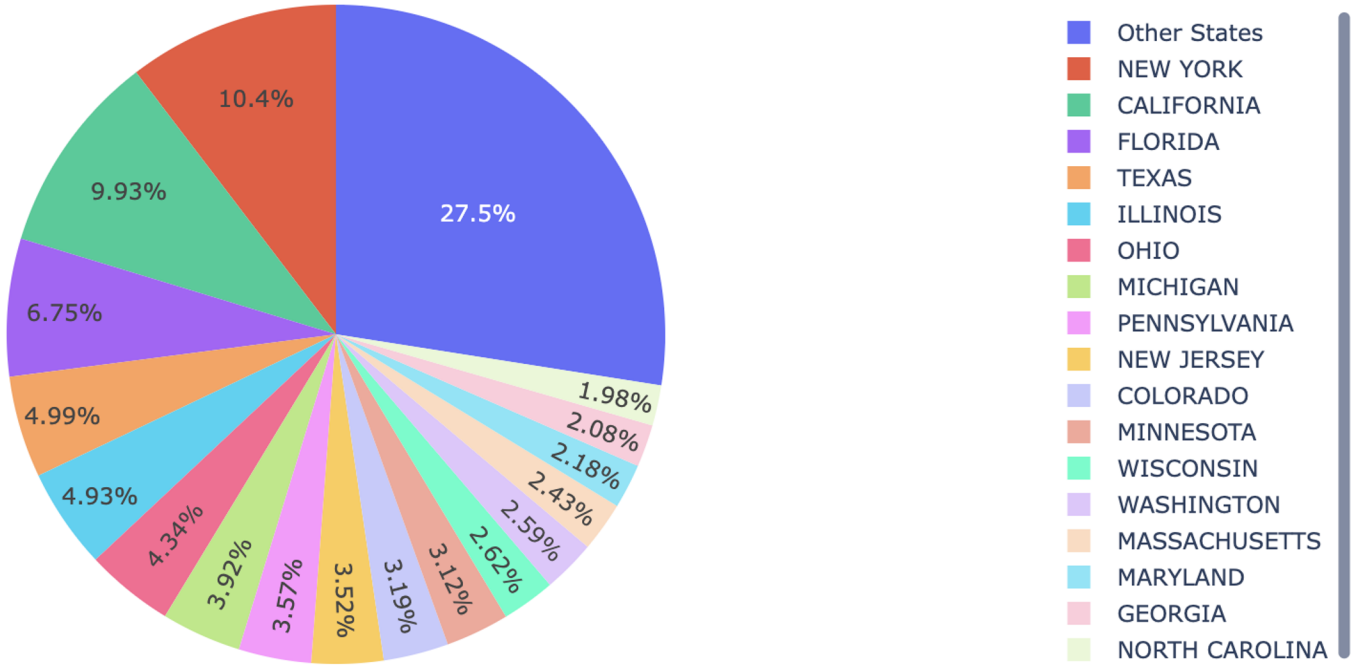
- Employed relevant statistical metrics and techniques to quantify and evaluate the analysis results, such as:
  - Calculating vote shares and vote margins for different parties
  - Measuring the correlation between various factors (e.g., demographics, economic indicators) and voting patterns
  - Performing hypothesis testing to assess the significance of observed trends or differences

Results

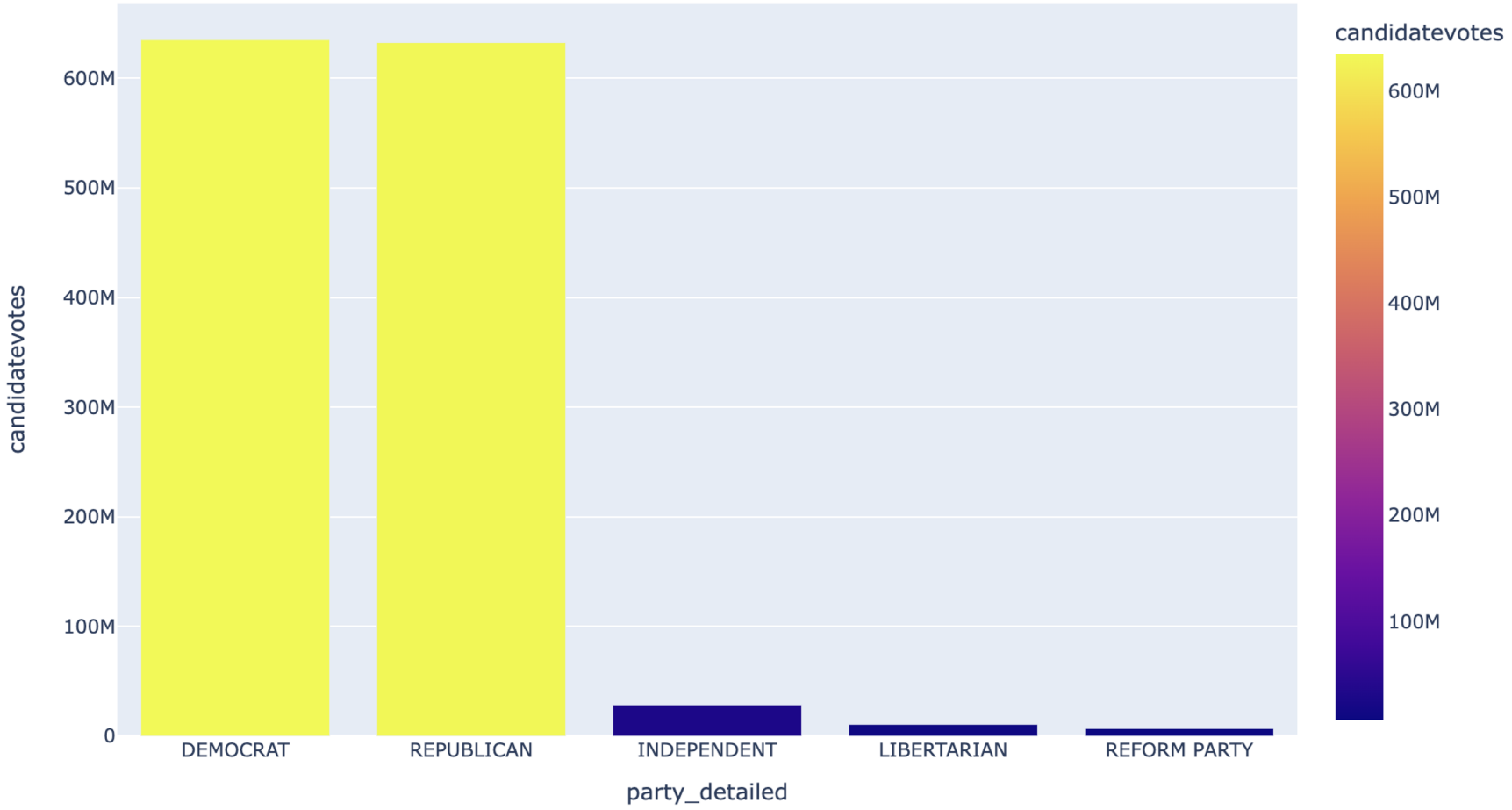


The following count plot shows number of state candidates for each party

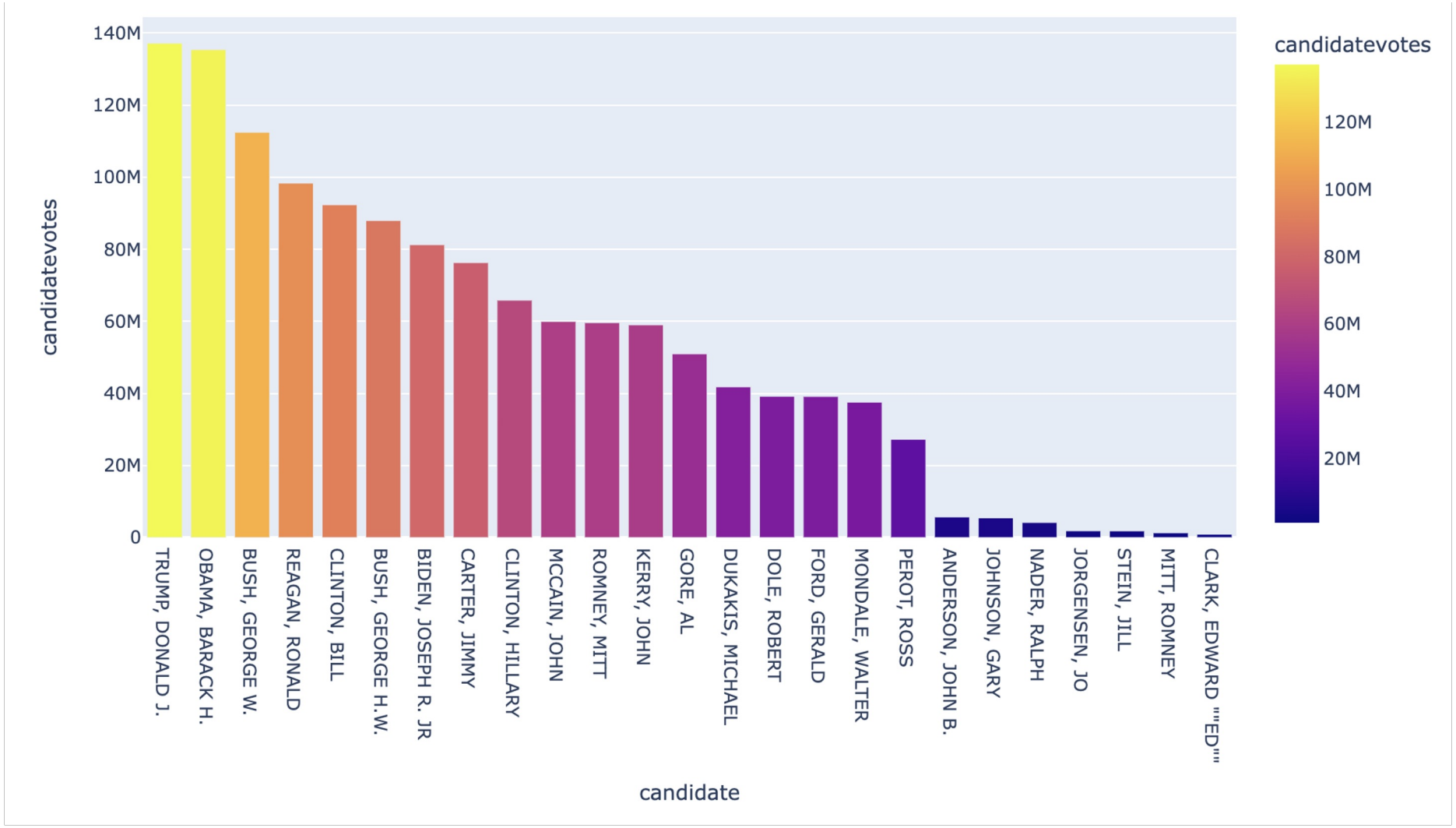
Total Votes



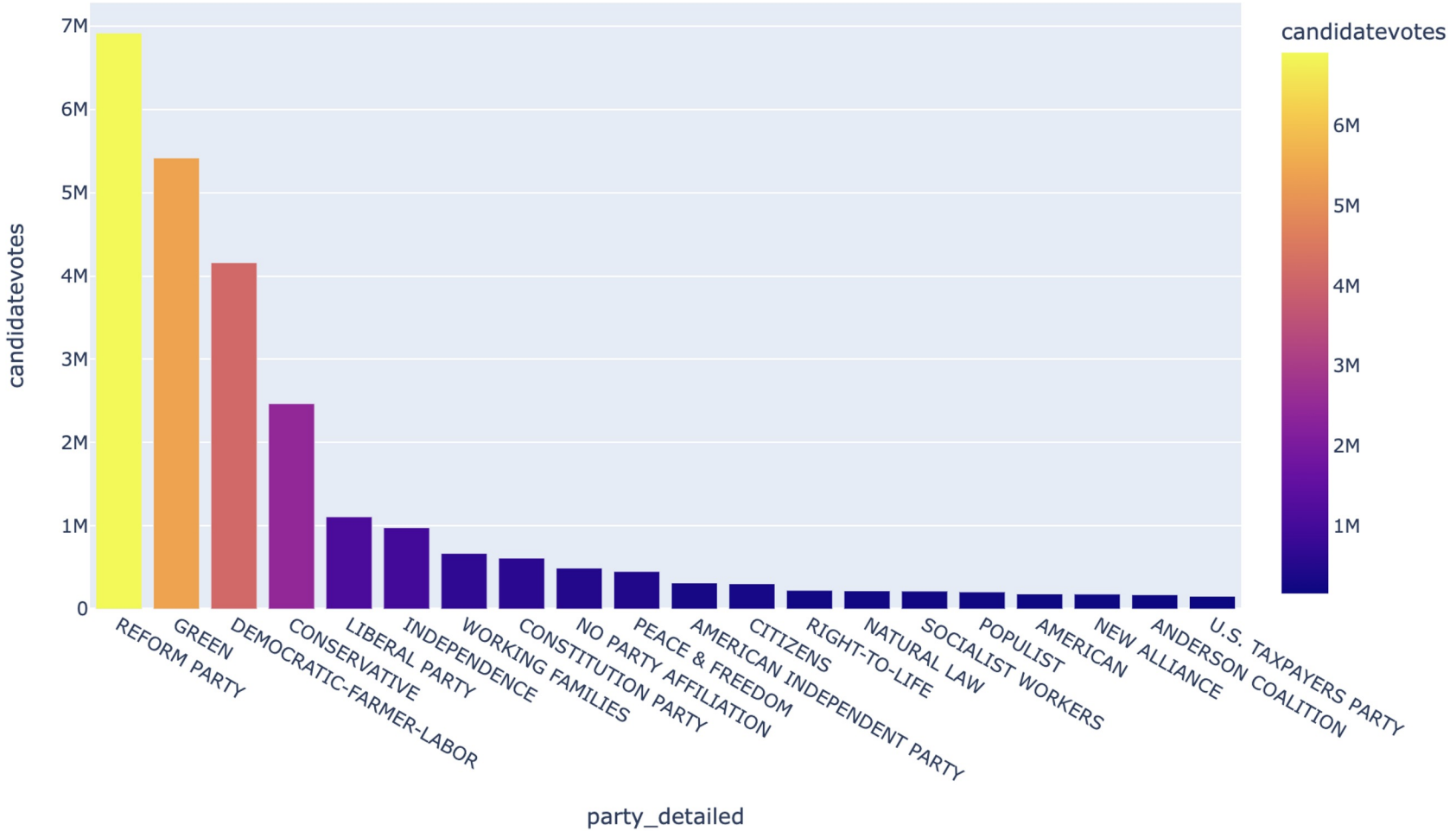
The following pie chart represents division of total votes casted across USA in 50 years



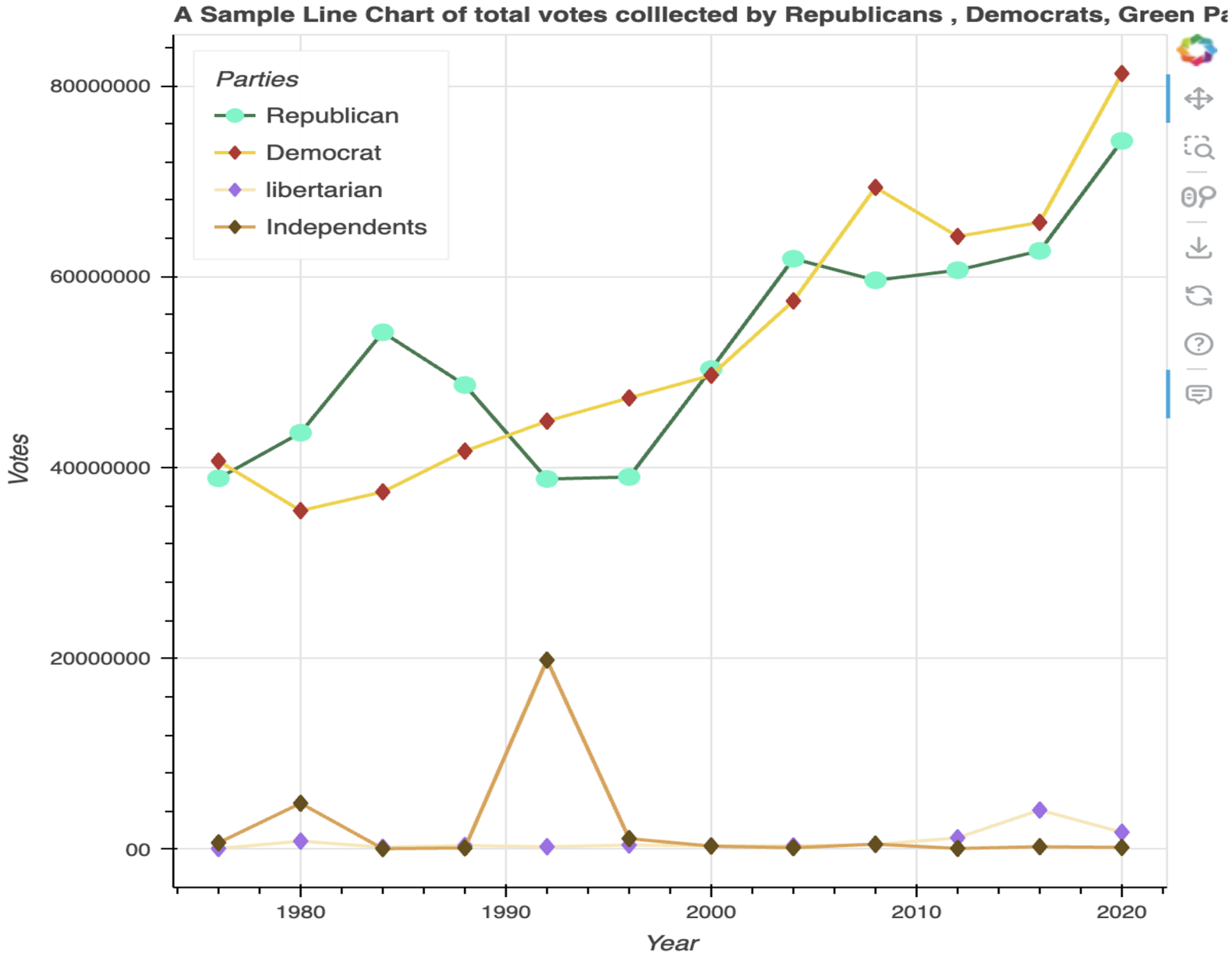
The following bar plot very clearly showcases top 5 dominant parties in USA Election history in span of 50 years

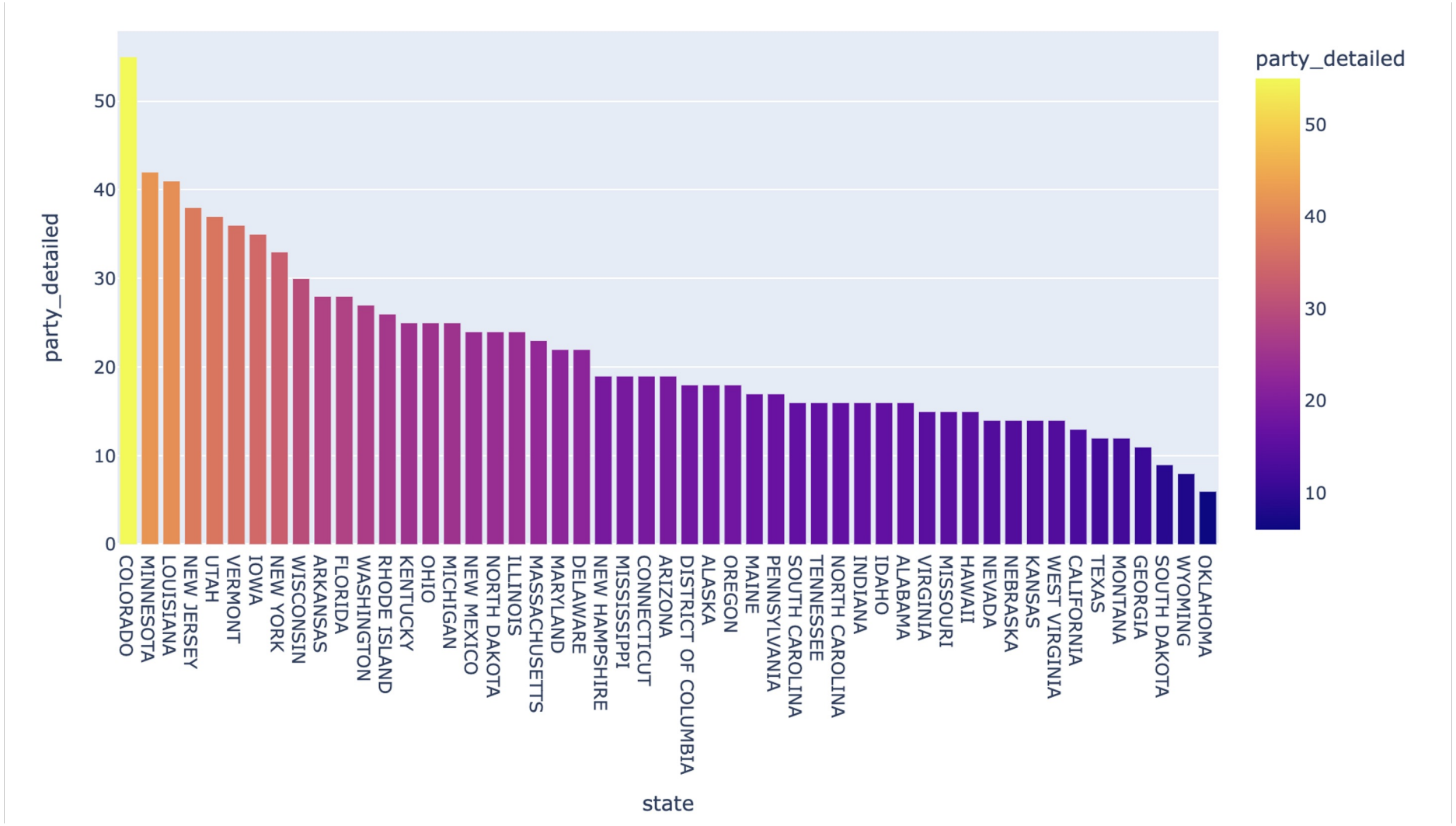


The following bar graph shows which candidates got most votes in USA Election history



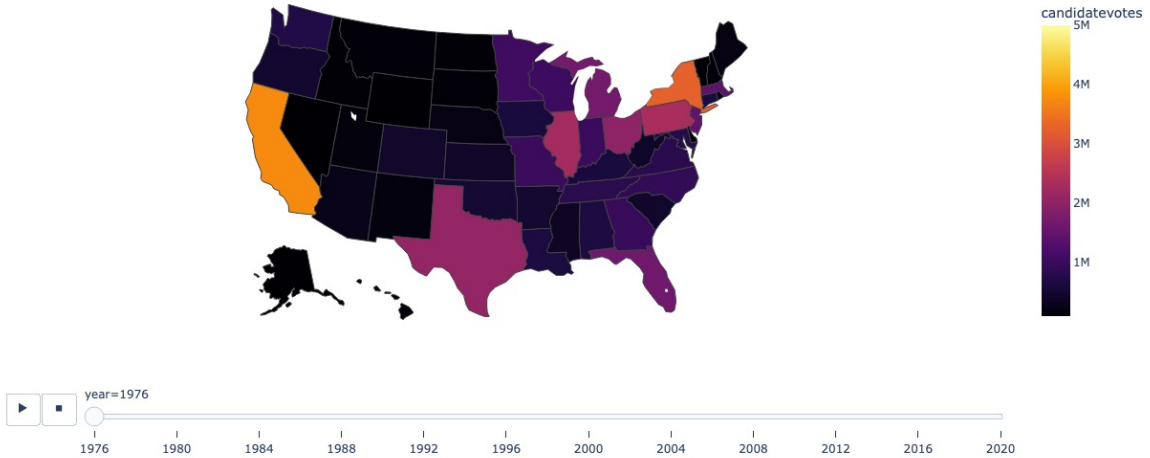
The following plot shows dominance of non-major top 5 parties during the span of 50 years where Reform party is leading by a huge margin


A Sample Line Chart of total votes colllected by Republicans , Democrats, Green Pa
Parties
Republican
Democrat
libertarian
Independents
80000000
60000000
40000000
20000000
00
Votes
1980
1990
2000
2010
2020
Year

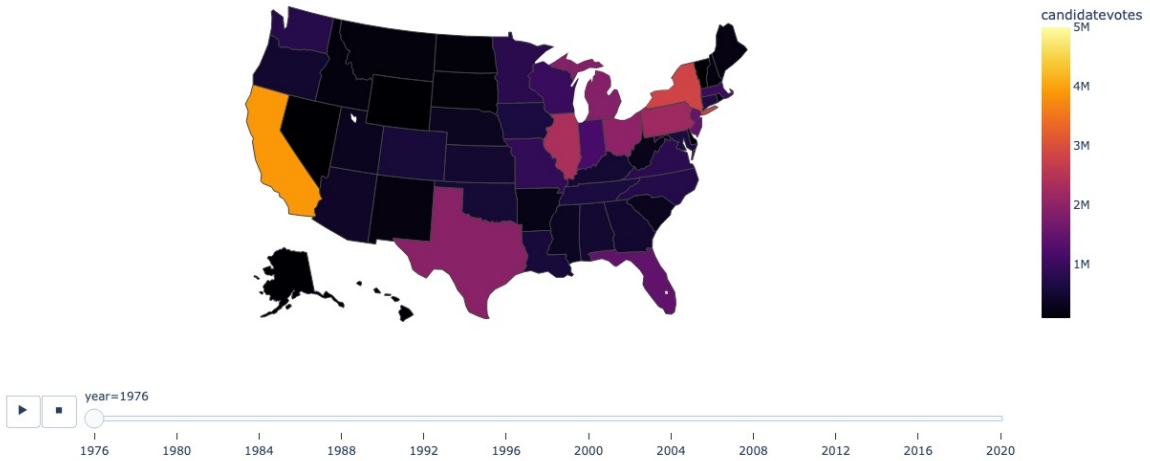


The following bar plot shows number of unique parties with each state aligned in decreasing order

Total Votes by Party - Democrats



Total Votes by Party - Republican



Conclusions

Based on the analysis, we drew many conclusions about the US presidential election trends, party dynamics, and state-level voting behaviors.

1. We found out New York and California are the most important states as far as total contribution in elections is concerned over the span of 50 years.
2. Donald J. Trump has been the most influential leader so far followed by Barack H. Obama.
3. Democrat and Republicans are the only two major parties dominating elections way far than any other party in the history of USA Elections. Although Democrats are found to be slightly ahead then Republicans.
4. Colorado has been the state with most number of unique parties so far followed by New York.
5. There are few battleground states(SWING) which mostly decide the fate of candidates, some of them are Pennsylvania, Wisconsin and Illinois.

Future Research

Although we came up for the solutions of our problem statement, but there is always scope for development, some of them could be:

- **Incorporating additional data sources**, such as demographic data, economic indicators, and polling data, to gain deeper insights into voter behavior.
- **Developing predictive models** to forecast election outcomes based on historical data and relevant factors.
- **Exploring advanced data visualization techniques**, such as interactive dashboards or virtual reality environments, to enhance the presentation and understanding of election data.
- **Investigating potential discrepancies or biases** in the data collection and analysis processes, and proposing methods to mitigate them.

References

Kaggle

MIT Election Lab Dataset

Google.com

Wikipedia

Syracuse University
College of Engineering
& Computer Science

Thank you