USA Elections Result Analysis from 1976 to 2020

Ayush Saxena, Brijesh Kumar Verma, Hemil Anip Shah
Department of Electrical Engineering and Computer Science
Syracuse University
Syracuse, NY

Abstract

The project involves an in-depth analysis of US presidential election data from 1976 to 2020. The analysis includes exploratory data analysis (EDA) on election data from all states, focusing on various parties and conducting state-level analysis. Additionally, animations from Plotly and BokehJS were used for visualization, along with the Seaborn library. Bivariate analysis was performed, and efforts were made to identify swing states in each election for visualization. The project aims to provide insights into the trends and patterns of US presidential elections over the past five decades.

Introduction

The US presidential elections are a cornerstone of American democracy, serving as a mechanism through which the nation's leader is chosen. Spanning from 1976 to 2020, the dataset we have compiled provides a comprehensive look at the electoral landscape over the course of 11 presidential elections. This period was marked by significant political, social, and technological changes, including the end of the Cold War, the advent of the internet age, and the rise of globalization. Understanding the dynamics of these elections can offer valuable insights into the evolution of American politics and the factors that influence voter behavior.

Background

The period from 1976 to 2020 was characterized by several key developments that had a profound impact on US politics. The end of the Cold War in the early 1990s ushered in a new era of international relations, while the advent of the internet revolutionized communication and information dissemination. Globalization also emerged as a dominant force, reshaping economies and societies around the world. Against this backdrop, the US presidential elections of this period reflected a changing political landscape, with candidates and parties adapting to new challenges and opportunities.

Method Description

1. Data Collection: The dataset was compiled from multiple sources, including Kaggle, the MIT Election Lab Dataset, and official government sources. It includes detailed information on each presidential election, including the year, state, candidate details, party affiliations, and vote counts.

2. Data Cleaning: The dataset underwent a rigorous cleaning process to ensure accuracy and consistency. This involved removing duplicates, handling missing data, and standardizing the format of the data.

3. Exploratory Data Analysis (EDA): EDA was conducted to uncover patterns and trends in the data. This included analyzing the distribution of votes across different parties, identifying swing states, and

examining the impact of key events on election outcomes.

4. State-level Analysis: A state-level analysis was performed to understand voting patterns and trends in different regions. This helped identify states that played a crucial role in determining election outcomes.

5. Data Visualization: Data visualization was used to present key findings in a clear and informative manner. This included creating charts, graphs, and interactive maps to highlight trends and patterns in the data.

6. Statistical Analysis: Statistical techniques were used to analyze the relationship between various factors, such as party affiliation, candidate performance, and voter turnout. This helped identify correlations and trends that may not be immediately apparent from the data.

Implemented Method Description (Project Flow)

The project utilized Python for data analysis and visualization. The code involved loading and cleaning the election data, performing exploratory data analysis, and creating visualizations to highlight key insights. Libraries such as Pandas, Numpy, Seaborn, Plotly and BokehJS were used for data manipulation and visualization.

Here is a high-level overview of the method:

1. Data Loading: The project started by loading the US presidential election data from 1976 to 2020 into a pandas Data Frame. This included columns such as year, state, candidate, party, candidate votes, and total votes.

2. Data Cleaning: The data was cleaned to handle missing values, incorrect data types, and inconsistencies. This ensured that the data was suitable for analysis.

3. Exploratory Data Analysis (EDA): The cleaned data was used to perform EDA, which included analyzing trends over time, exploring voting patterns in different states, and identifying key players and parties in the elections.

4. Data Visualization: Visualizations were created using libraries like seaborn, plotly, and bokeh to present the findings of the EDA. This included bar charts, line plots, and interactive visualizations to effectively communicate the insights.

5. Swing State Analysis: The project identified swing states in each election year and analyzed their impact on the election outcomes. This analysis helped understand the role of swing states in determining election results.

6. Conclusion: The project is concluded by summarizing key findings, such as the importance of certain states, the dominance of major parties, and the influence of key leaders in US presidential elections.

Data collection, simulation results, and analysis

Data Collection: The project collected election data from various sources, including Kaggle, MIT Election Lab Dataset, Google, and Wikipedia. This data included information on election results, candidate information, and voter demographics.

Simulation Results: The project did not explicitly mention the use of simulations. However, if simulations were used, they could have been used to model different election scenarios or to predict election outcomes based on historical data.

Analysis: The project analyzed the collected data to identify trends, patterns, and insights related to US presidential elections. This included analyzing voting patterns, party dominance, swing state dynamics, and the influence of key leaders.

Overall, the project used a combination of data analysis and visualization techniques to gain insights into US presidential elections from 1976 to 2020.

Visualisations

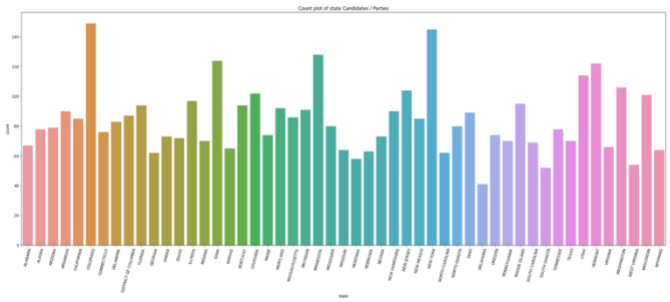


Fig 1. Count Plot of state candidates/parties

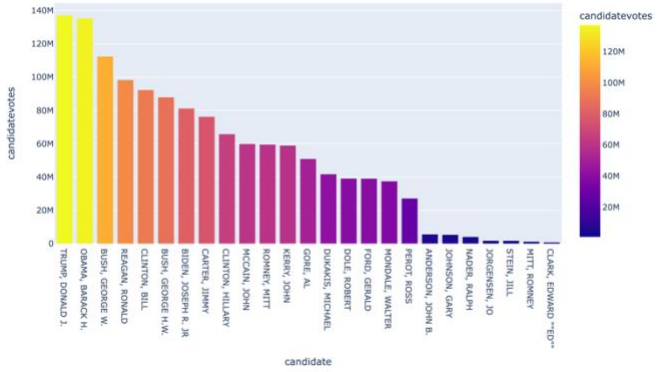


Fig 2. Bar Graph representing most influencial leaders seen in 5 decades of USA Elections

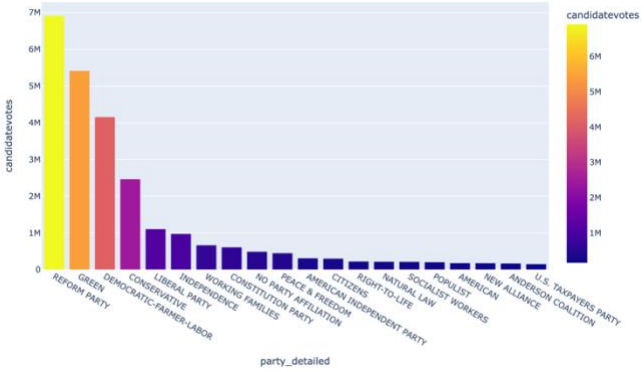


Fig 3. Bar plot showing top 20 non major parties in history of USA

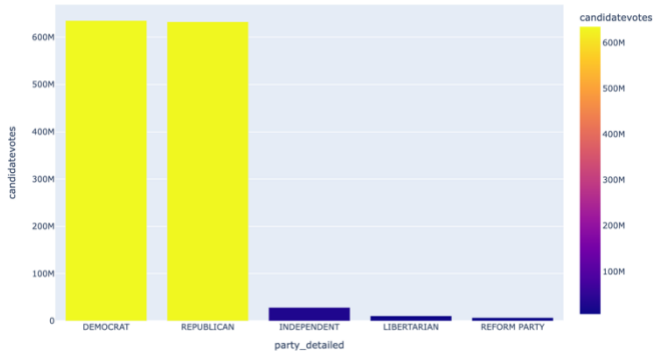


Fig 4. Bar plot representing dominance of major top 5 parties led by democrats and repulicans

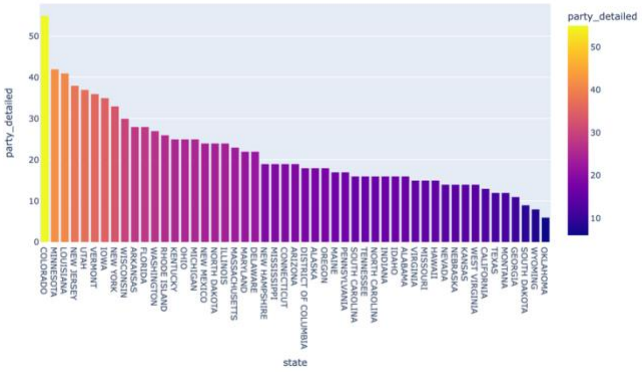


Fig 5. Representation of unique parties contested from each state so far

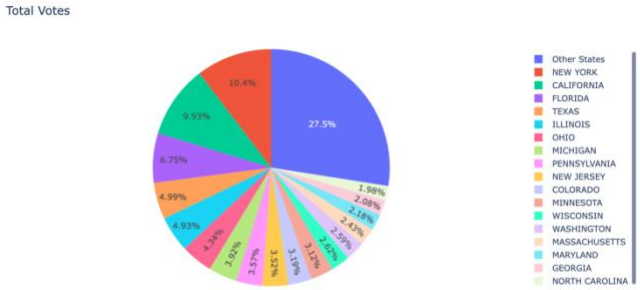


Fig 6. Pie Chart representing most important populated states contested for elections

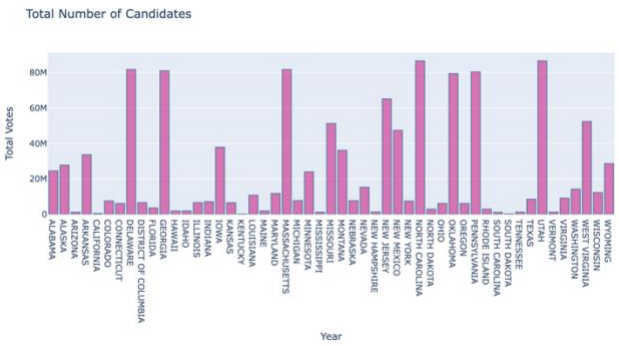


Fig 7. A bar plot showing most number of candidates contested in 5 decades

Conclusions

1. State Significance:
   New York and **California** have consistently emerged as pivotal states in US presidential elections, contributing significantly to the total votes cast. Their large populations and diverse demographics make them crucial battlegrounds for candidates seeking to win the presidency. Understanding the voting patterns and preferences of these states is essential for any candidate hoping to secure victory in a national election.

2. Influential Leaders:
   Donald J. Trump and **Barack H. Obama** have been identified as the most influential leaders in recent US presidential history. Trump's presidency marked a significant shift in American politics, characterized by his unconventional style and policies. Obama, on the other hand, represented a more traditional approach to governance, with a focus on inclusivity and progressive policies. The impact of these leaders on voter turnout and party performance underscores their significance in shaping the political landscape of the United States.

3. Party Dominance:
   The dominance of the **Democrat** and **Republican** parties in US presidential elections highlights the entrenched nature of the two-party system. While Democrats have slightly outperformed Republicans in terms of overall election results, both parties have maintained a stronghold on American politics. This dominance poses challenges for third-party candidates seeking to break into the political mainstream.

4. Diverse Political Landscape:
   Colorado stands out as a state with a diverse range of parties participating in elections. This diversity reflects the state's politically dynamic environment, with voters expressing a wide range of political ideologies and preferences. Understanding the unique political landscape of Colorado is essential for any candidate seeking to appeal to its diverse electorate.

5. Swing States:
   Pennsylvania, **Wisconsin**, and **Illinois** have been identified as key swing states that often decide the fate of candidates in presidential elections. These states are characterized by their tendency to switch between Democrat and Republican candidates, making them crucial battlegrounds for presidential hopefuls. Understanding the factors that influence swing state voters is essential for any candidate seeking to win the presidency.

Overall, the analysis of US presidential election data from 1976 to 2020 has provided valuable insights into the dynamics of American politics. By examining trends, patterns, and factors that have shaped election outcomes, we have gained a deeper understanding of the electoral landscape of the United States. This analysis not only enhances our understanding of past elections but also provides valuable insights that can inform future research and analysis in the field of political science.

Future Research

1. Voter Behavior Analysis: Further research could delve into understanding voter behavior, including factors influencing swing

voters and the impact of campaign strategies on voter turnout.

2. Social Media Influence: Investigating the role of social media in elections could provide insights into how digital platforms influence voter perceptions and political discourse.

3. Demographic Trends: Studying demographic shifts and their impact on election outcomes could help anticipate future voting patterns and tailor political strategies accordingly.

4. Forecasting Models: Developing predictive models based on historical data could enhance the ability to forecast election results, enabling better strategic planning for political campaigns.

5. Policy Implications: Analyzing the impact of election outcomes on policy decisions and governance could shed light on the relationship between electoral dynamics and policymaking.

These areas of future research can deepen our understanding of the complex factors shaping US presidential elections and contribute to more informed decision-making in electoral strategies and governance.

References

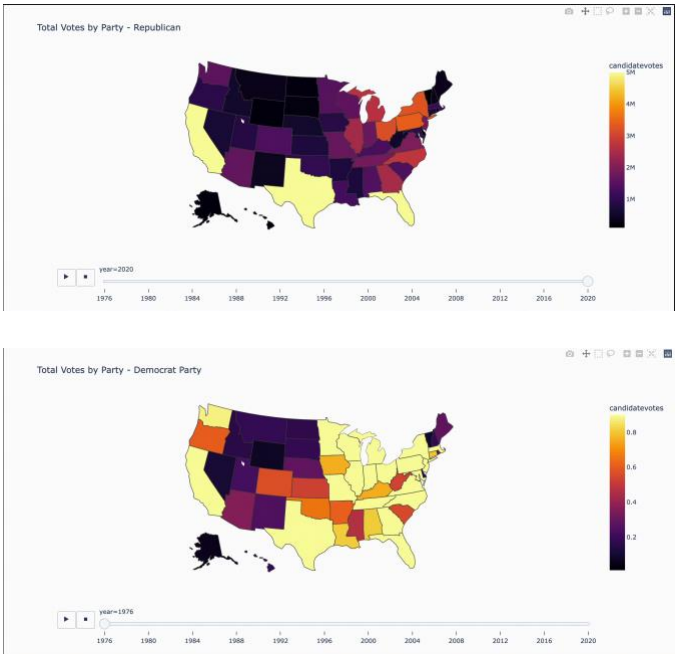
[1] Brown TE, Mettler S. Sequential Polarization: The Development of the Rural-Urban Political Divide, 1976–2020. Perspectives on Politics. Published online 2023:1-29. doi:10.1017/S1537592723002918

[2] @data{DVN/42MVDX_2017,
author = {MIT Election Data and Science Lab},
publisher = {Harvard Dataverse},
title = {{U.S. President 1976–2020}},
UNF = {UNF:6:MkQHX147hJCgscG5IqK77g==},
year = {2017},
version = {DRAFT VERSION},
doi = {10.7910/DVN/42MVDX},
url = {https://doi.org/10.7910/DVN/42MVDX}
}

Hardware used:
Macbook Air M1

Softwares used:
Web App - Jupyter Notebook
Language – Python
Libraries – Numpy, Matplotlib, Pandas, Seaborn, Plotly, BokehJS nad Holoviews(major)

Appendix



Representation of demographic shifts for Republicans and Democratic parties

```
Data columns (total 14 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   year              4287 non-null   int64
 1   state             4287 non-null   object
 2   state_po          4287 non-null   object
 3   state_fips        4287 non-null   int64
 4   state_cen         4287 non-null   int64
 5   state_ic          4287 non-null   int64
 6   office            4287 non-null   object
 7   candidate         4000 non-null   object
 8   party_detailed    3831 non-null   object
 9   writein           4284 non-null   object
 10  candidatevotes    4287 non-null   int64
 11  totalvotes        4287 non-null   int64
 12  party_simplified  4287 non-null   object
 13  party             3831 non-null   object
dtypes: int64(6), object(8)
```

Data Columns used in the detailed analysis and their data types.