

ANALYSIS OF UNITED STATES SUPREME COURT DATA (1946-2015)



Authored by: Mtape Rosemary

Date: 27/09/2024

Table of Contents

Executive Summary 3

Introduction..... 3

Data Overview 3

Methodology 4

Discussion 14

Limitations..... 15

Conclusion 15

Executive Summary

This report provides a comprehensive analysis of U.S. Supreme Court cases from 1946 to 2015, encompassing a total of 8,737 cases. The analysis aims to uncover meaningful trends and insights into the judicial process, focusing on various dimensions such as case types, decision outcomes, and the influence of individual justices. Key findings indicate a notable increase in the Supreme Court's caseload during the mid-20th century, peaking in the 1970s, followed by stabilization in subsequent decades. The report highlights the most frequently occurring case names and reveals significant patterns in decision types and directions, suggesting that certain judicial behaviors are more prevalent.

Data cleaning was a critical component of the methodology, involving the removal of columns with excessive missing values and the imputation of missing data for categorical and numerical variables, ensuring a robust dataset for analysis. Additionally, the report examines the influence of Chief Justices on case outcomes, revealing that Chief Justice John Roberts presided over the largest number of cases during the specified period. Overall, this analysis provides valuable insights into the dynamics of the U.S. Supreme Court, highlighting its role in shaping American law and addressing significant social issues over nearly seven decades, and serves as a foundation for further research on the court's impact on legal precedents and societal changes.

Introduction

This report presents a comprehensive analysis of US Supreme Court cases spanning from 1946 to 2015. The Supreme Court plays a critical role in shaping American law and addressing significant social issues; thus, understanding the patterns in its decisions offers valuable insights into the judicial process and its implications for society. This analysis aims to extract meaningful trends and insights from the dataset, focusing on various attributes such as case types, origins, decision outcomes, and the influence of individual justices. By examining these elements, this report seeks to highlight shifts in judicial priorities and the evolving nature of legal interpretations over nearly seven decades.

Data Overview

The dataset encompasses a total of 8737 Supreme Court cases. Key columns include:

- Case Identifiers: case_id, docket_id, issues_id
- Decision Details: date_decision, decision_type, majority_opinion_writer
- Voting Information: majority_votes, minority_votes

- Citations: us_citation, court_citation

Methodology

The analysis utilized Python, leveraging libraries such as pandas for data manipulation and seaborn for data visualization. Given the initial dataset had numerous missing values, a systematic cleaning process was crucial.

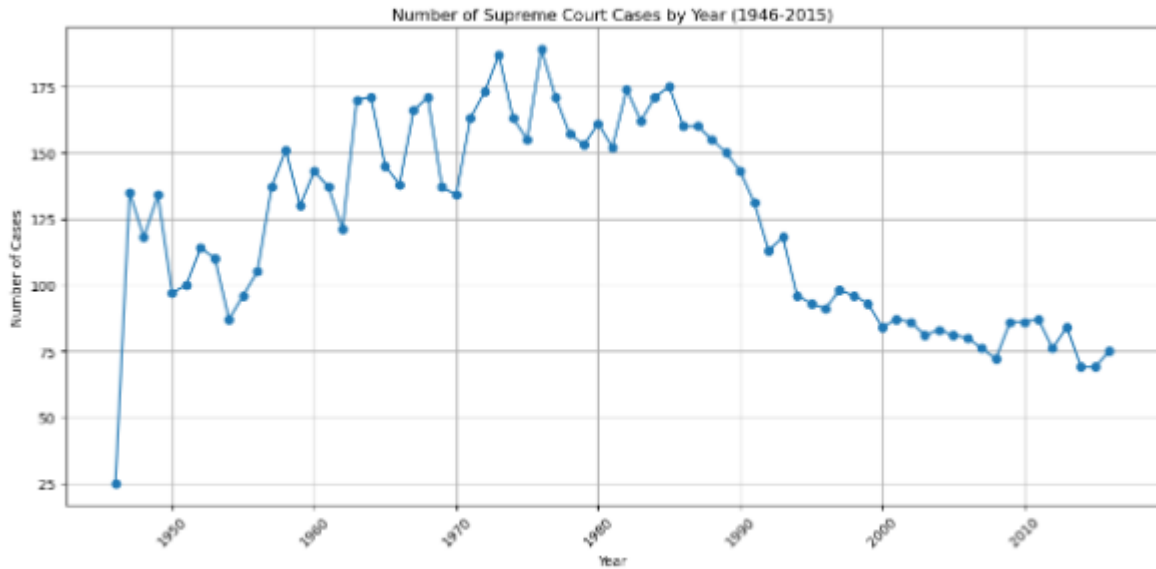
The data cleaning approach involved several key steps:

1. **Dropping Columns:** Columns with more than 50% missing values were removed from the dataset. This helped to retain only those columns that provided sufficient information for analysis.
2. **Filling Missing Values:**
 - For categorical columns, missing values were filled with the mode (the most frequently occurring value). This method ensures that the most common category is represented, maintaining the integrity of categorical data.
 - For numerical columns, missing values were filled with the median. This approach protects against the influence of outliers and provides a more accurate representation of the central tendency.

These cleaning steps were vital for preparing a robust dataset that could yield meaningful insights.

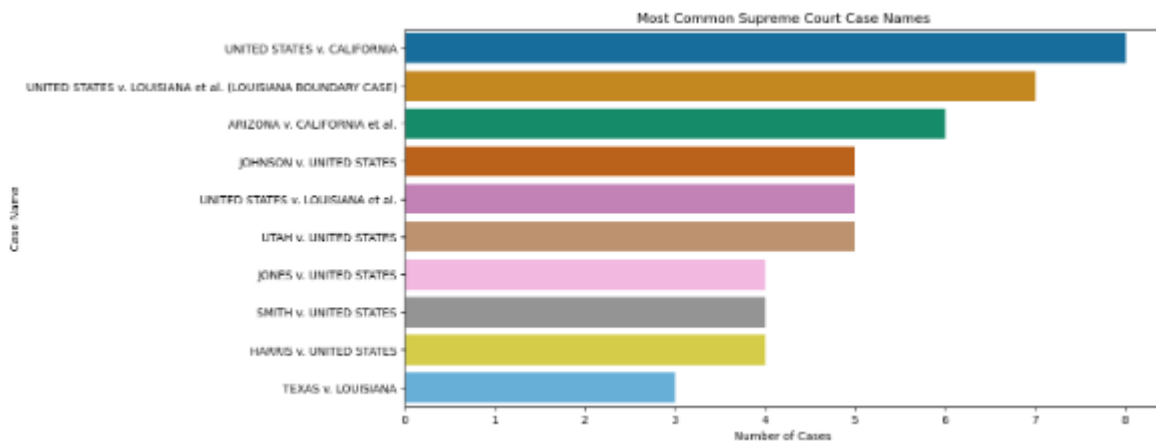
Key Findings and Analysis

1. **Trend in Supreme Court Caseload (1946-2015)**



The visualization depicts the number of Supreme Court cases decided each year from 1946 to 2015. It reveals a distinct pattern of increasing caseload in the early decades, peaking in the 1970s, followed by a gradual decline until the late 1990s. Subsequently, the caseload stabilizes with minor fluctuations. This trend suggests a period of heightened judicial activity in the mid-20th century, followed by a stabilization of the Court's workload.

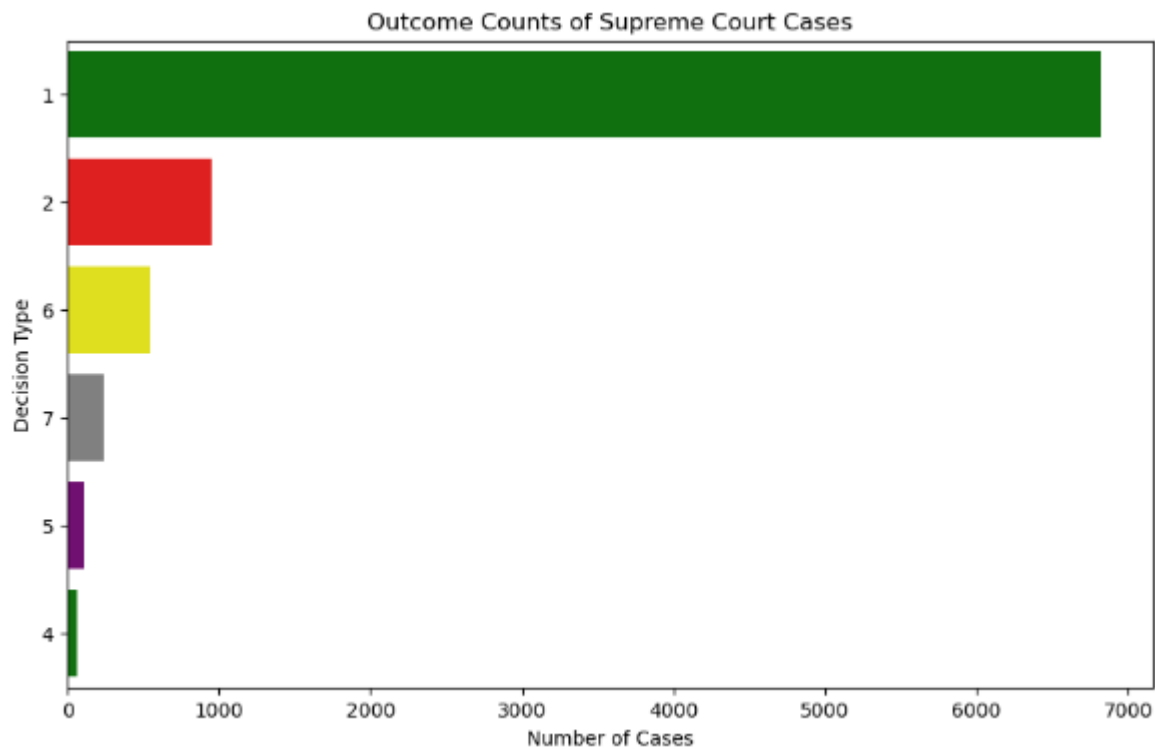
2. Most Frequent Case Names in the U.S. Supreme Court (1946-2015)



The visualization presents the ten most frequently recurring case names in the U.S. Supreme Court dataset from 1946 to 2015. The horizontal bars represent the case names, while the length of each bar indicates the number of times that case name appeared in the dataset. The chart reveals that "UNITED STATES v. CALIFORNIA" is the most frequently appearing case name, followed by "UNITED STATES v. LOUISIANA et al. (LOUISIANA BOUNDARY CASE)".

This suggests that certain legal disputes or types of cases have been more common or contentious throughout the dataset's time period.

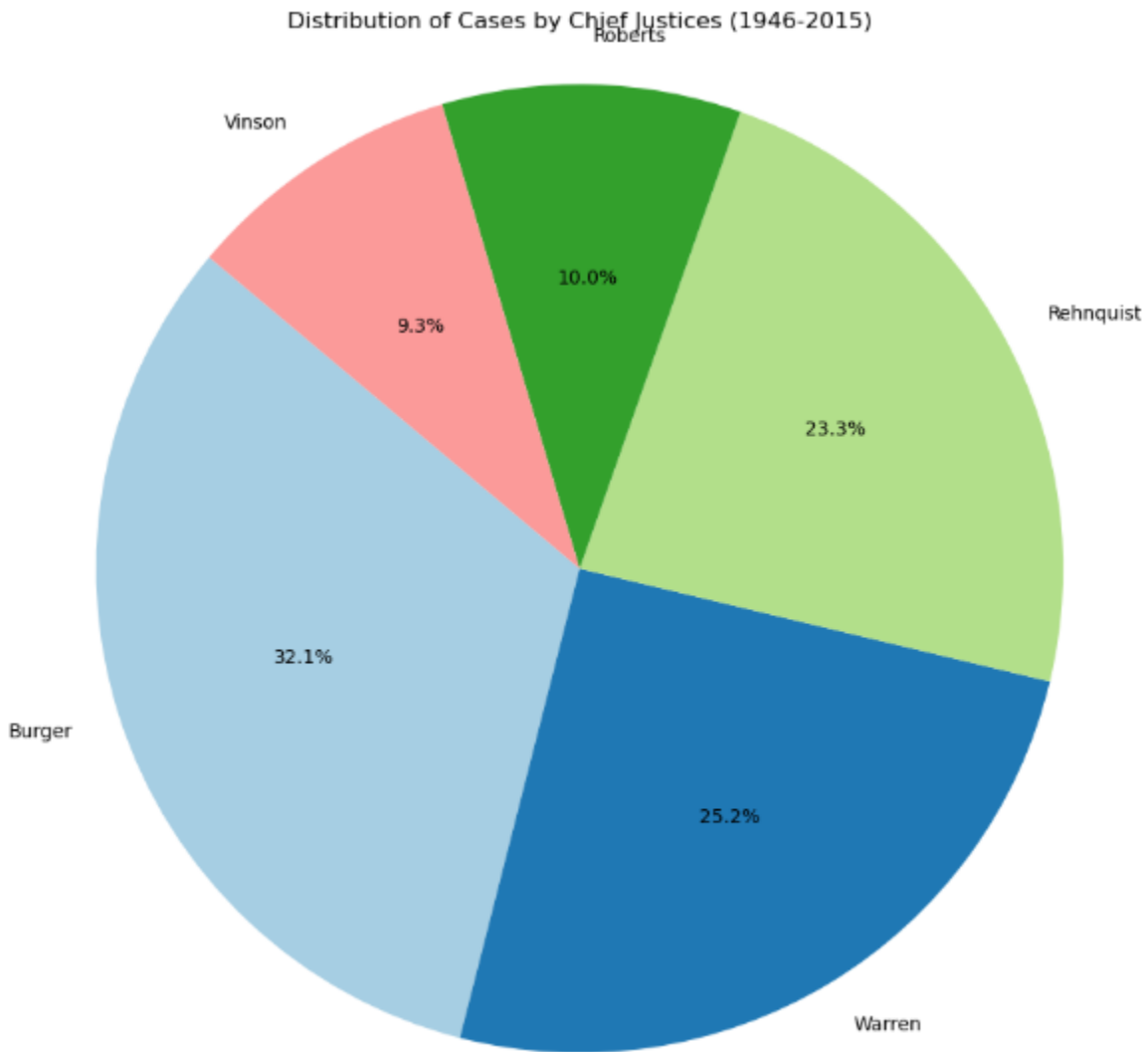
3. Distribution of Supreme Court Decision Types (1946-2015)



The chart presents a horizontal bar chart illustrating the distribution of different decision types in the U.S. Supreme Court dataset from 1946 to 2015. The x-axis represents the number of cases, while the y-axis indicates the decision type. The length of each bar corresponds to the frequency of that particular decision type. The chart reveals that the most common decision type is "1", followed by "2", "6", and other less frequent categories. This suggests that certain types of decisions are more prevalent in the Supreme Court's jurisprudence.

4. Distribution of Supreme Court Cases by Chief Justices

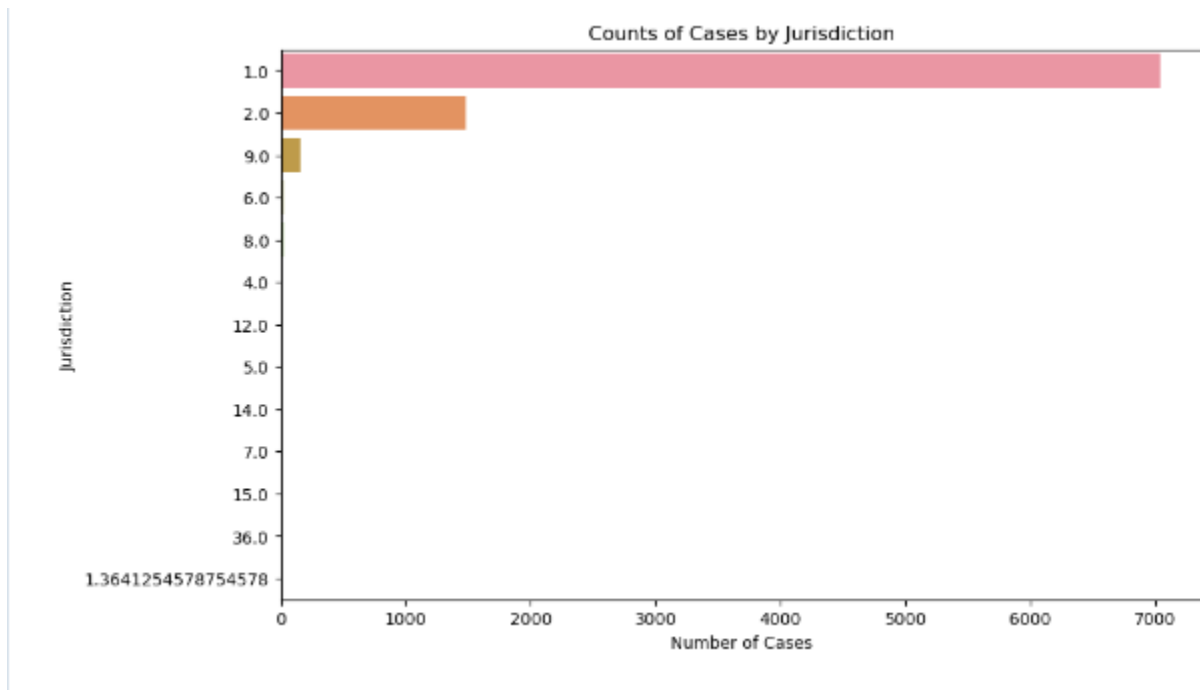
The chart below presents a pie chart illustrating the proportion of Supreme Court cases decided during the tenures of different Chief Justices from 1946 to 2015. Each slice of the pie represents a Chief Justice, with the size of the slice indicating the percentage of cases decided under their leadership. The chart reveals that Chief Justice John Roberts presided over the largest number of cases (32.1%), followed by Earl Warren (25.2%), William Rehnquist (23.3%), Warren Burger (10.0%), and Fred M. Vinson (9.3%). This information provides insights into the relative influence and tenure lengths of these Chief Justices.



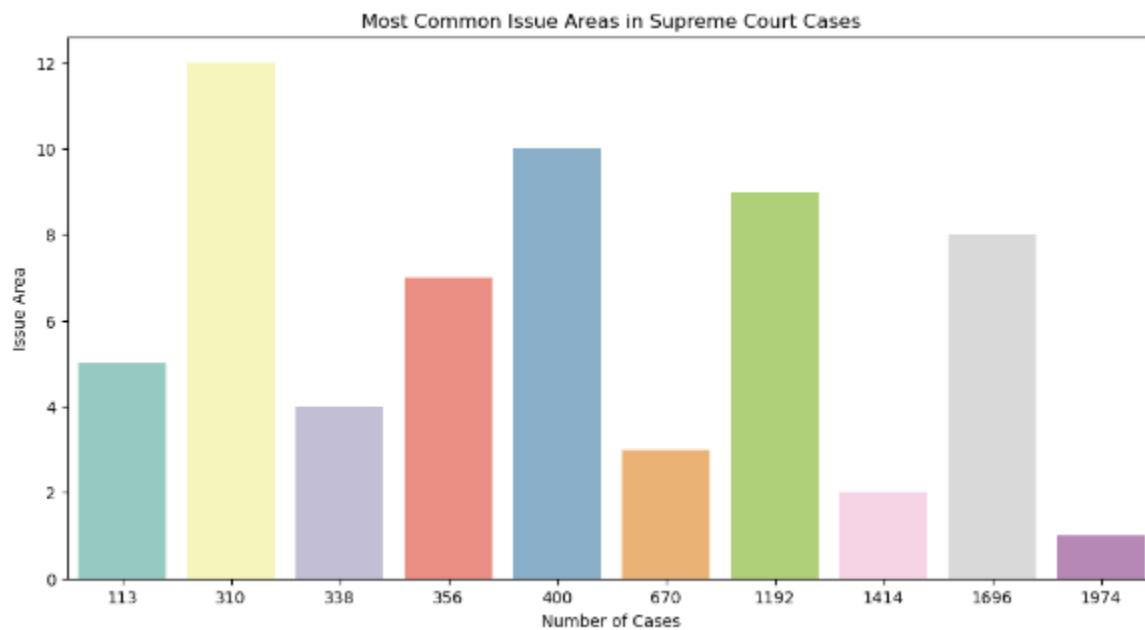
5. Distribution of Supreme Court Cases by Jurisdiction

The visualization below presents a horizontal bar chart illustrating the distribution of Supreme Court cases across different jurisdictions from 1946 to 2015. The x-axis represents the number of cases, while the y-axis indicates the jurisdiction. The length of each bar corresponds to the frequency of cases originating from that particular jurisdiction. The chart reveals that jurisdiction "1.0" has the highest number of cases, followed by "2.0" and other jurisdictions with significantly fewer cases. This suggests

that certain jurisdictions are more frequently involved in Supreme Court cases compared to others



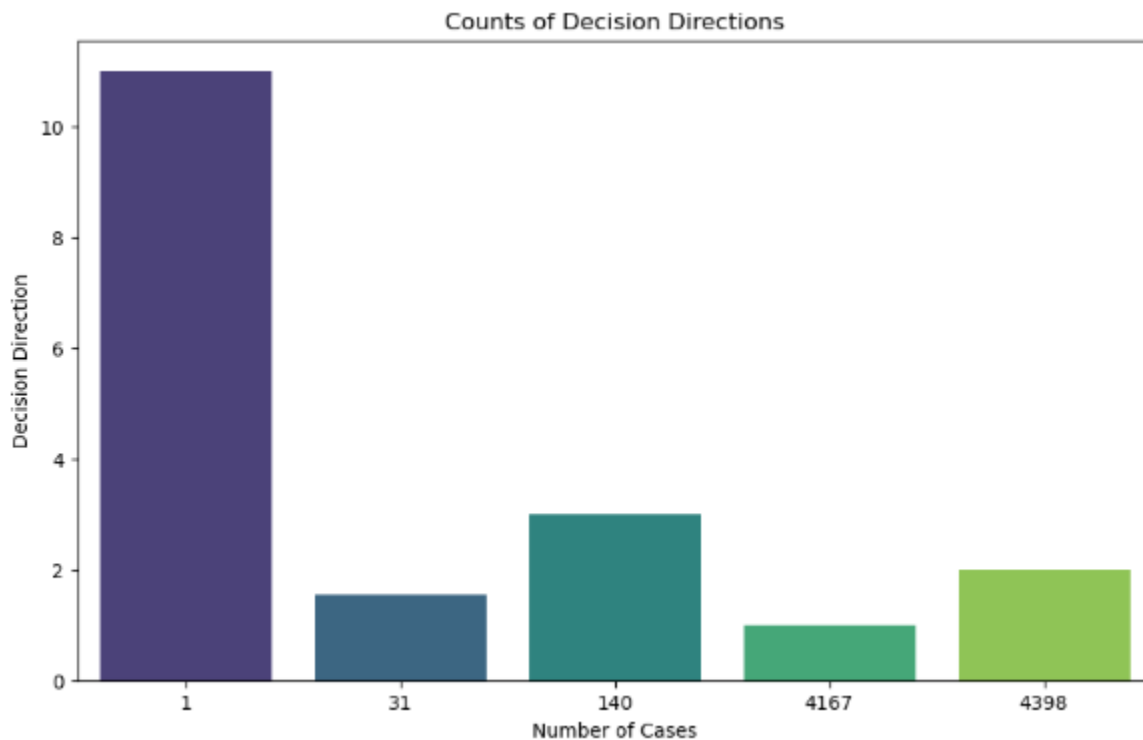
6. Most Common Issue Areas in Supreme Court Cases (1946-2015)



The infographic presents a bar chart illustrating the distribution of Supreme Court cases across different issue areas from 1946 to 2015. The x-axis represents the issue area, while the y-axis indicates the number of cases. The height of each bar corresponds to the frequency of cases falling under that particular issue area. The chart reveals that issue area "310" has the highest

number of cases, followed by "400" and other issue areas with varying frequencies. This suggests that certain legal issues are more commonly addressed by the Supreme Court compared to others

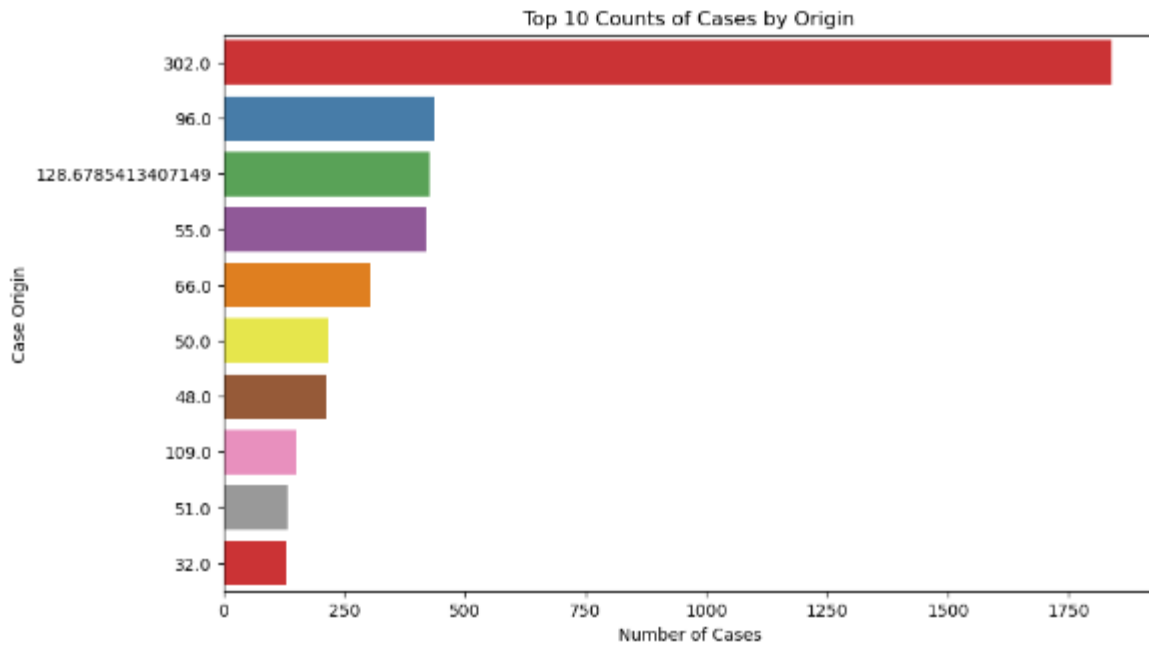
7. Distribution of Supreme Court Decision Directions (1946-2015)



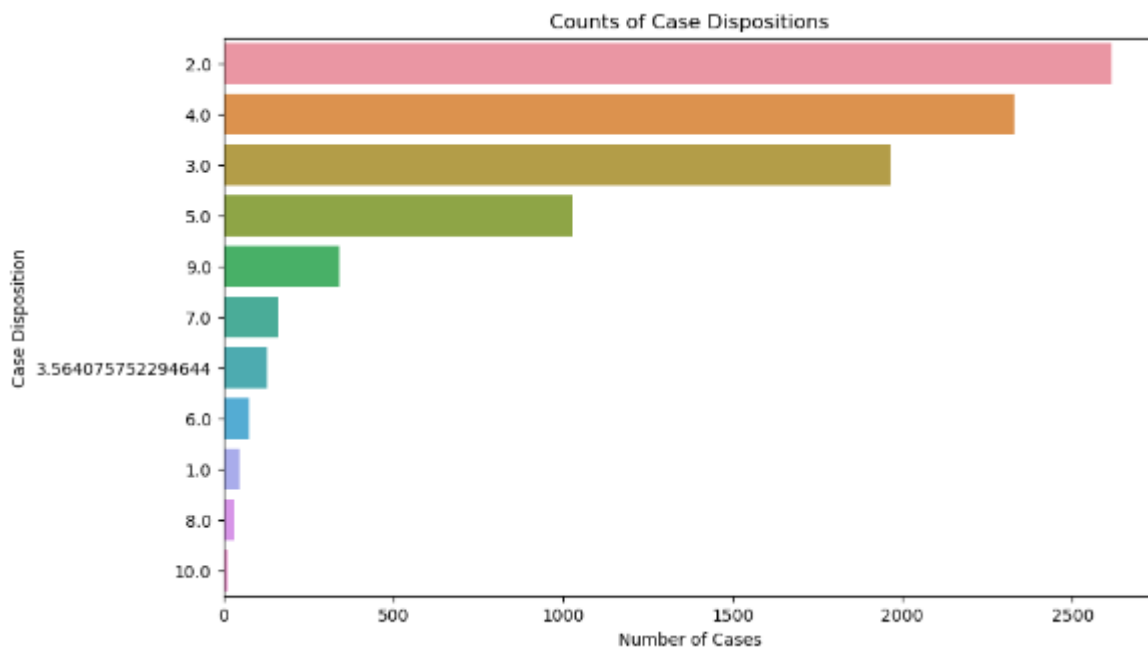
The bar chart illustrates the distribution of decision directions in Supreme Court cases from 1946 to 2015. The x-axis represents the decision direction, while the y-axis indicates the number of cases. The height of each bar corresponds to the frequency of cases with that specific decision direction. The chart reveals a significant imbalance, with decision direction "1" being the most prevalent, followed by "31", "140", and "4167". This suggests that a particular decision direction is more common than others in the Supreme Court's jurisprudence

8. Top 10 Counts of Cases by Origin

The below chart illustrates the distribution of Supreme Court cases across the top 10 most frequent case origins. The x-axis represents the number of cases, while the y-axis indicates the case origin. The length of each bar corresponds to the frequency of cases originating from that particular source. The chart reveals that case origin "302.0" has the highest number of cases, followed by "96.0" and other origins with varying frequencies. This suggests that certain types of cases or jurisdictions are more commonly represented in the Supreme Court's docket



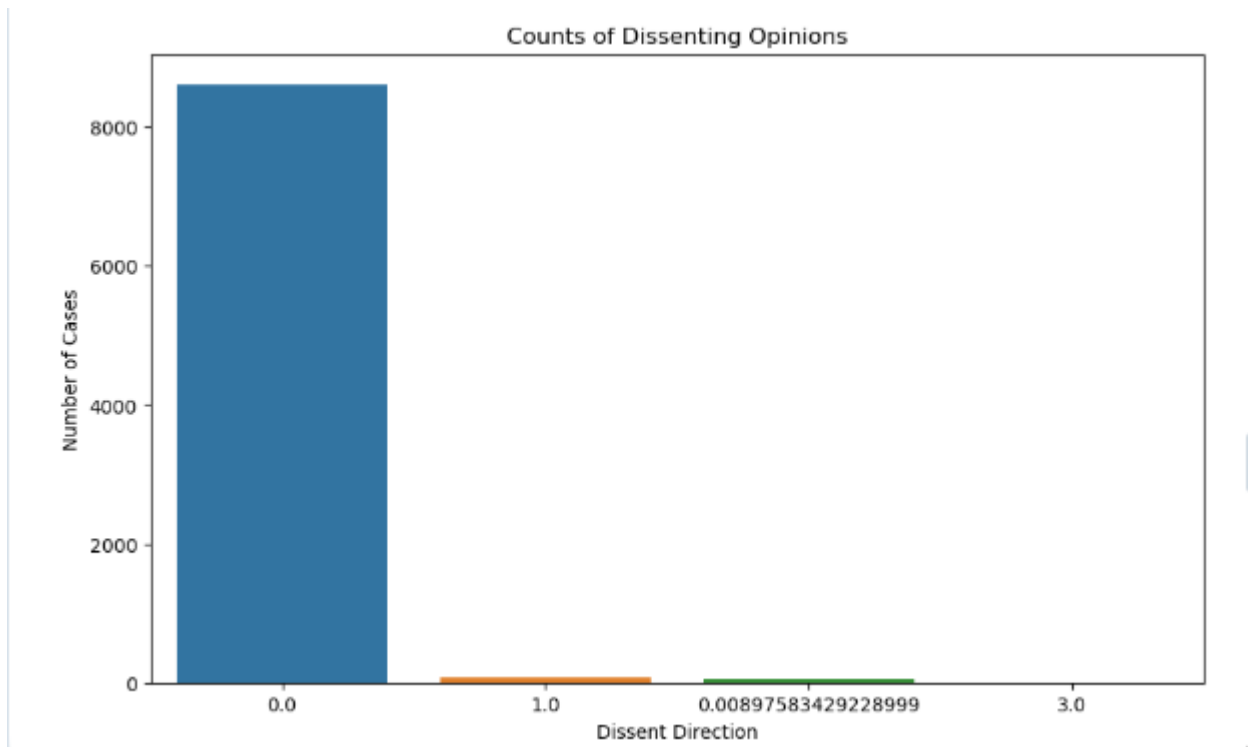
9. Distribution of Supreme Court Case Dispositions (1946-2015)



The visualization presents a horizontal bar chart illustrating the distribution of case dispositions in Supreme Court cases from 1946 to 2015. The x-axis represents the number of cases, while the y-axis indicates the case disposition. The length of each bar corresponds to the frequency of cases with that specific disposition. The chart reveals that disposition "2.0" has the highest number of cases, followed by "4.0" and other dispositions with varying frequencies. This

suggests that certain types of case dispositions are more common in the Supreme Court's jurisprudence

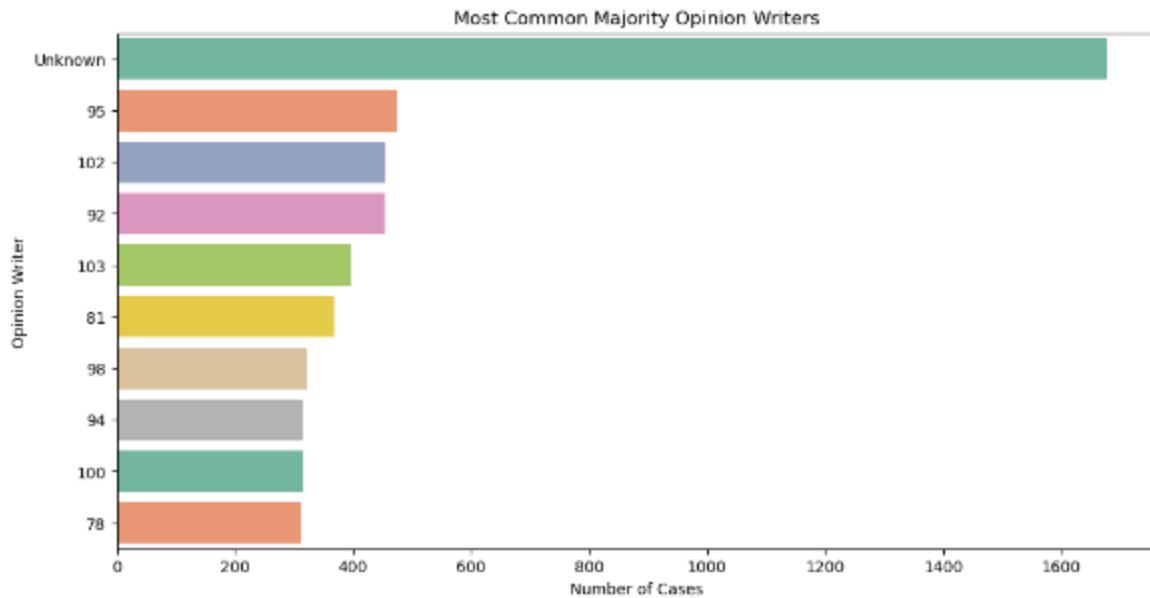
10. Distribution of Dissenting Opinions in Supreme Court Cases



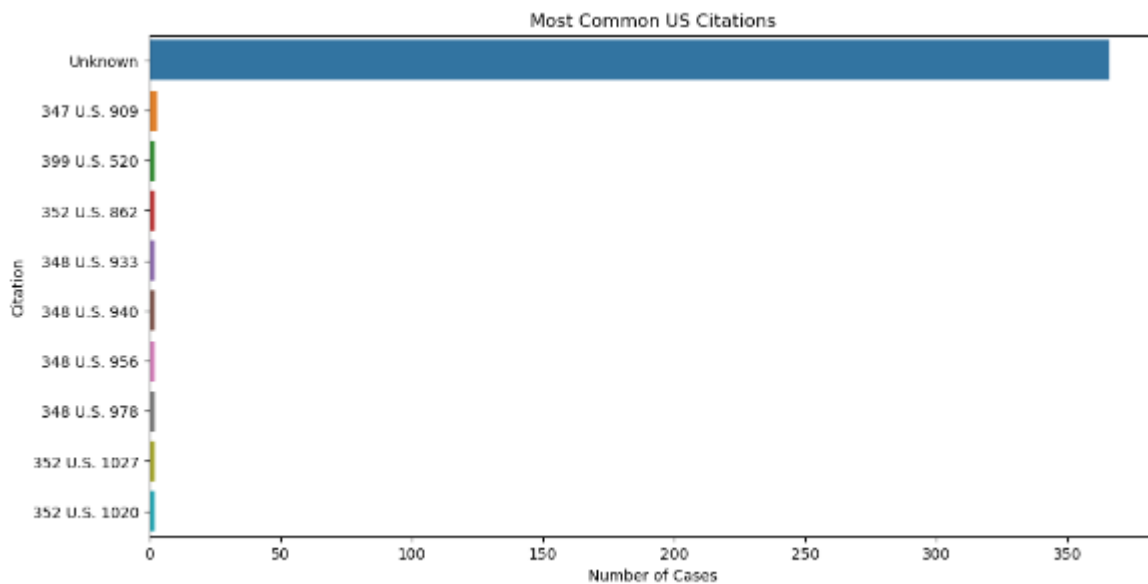
The bar chart shows the distribution of dissenting opinions in Supreme Court cases from 1946 to 2015. The x-axis represents the dissent direction, while the y-axis indicates the number of cases. The height of each bar corresponds to the frequency of cases with that specific dissent direction. The chart reveals a significant imbalance, with dissent direction "0.0" being the most prevalent, followed by "1.0" and other less frequent directions. This suggests that a particular dissent direction is more common than others in the Supreme Court's jurisprudence.

11. Most Common Majority Opinion Writers in Supreme Court Cases

The visualization below presents a horizontal bar chart illustrating the distribution of majority opinions written by different Supreme Court justices. The x-axis represents the number of cases, while the y-axis indicates the opinion writer. The length of each bar corresponds to the frequency of majority opinions authored by that justice. The chart reveals that "Unknown" is the most common majority opinion writer, followed by "95", "102", and other justices with varying frequencies. This suggests that either there is a lack of data for some justices or that certain justices have written a larger number of majority opinions compared to others.

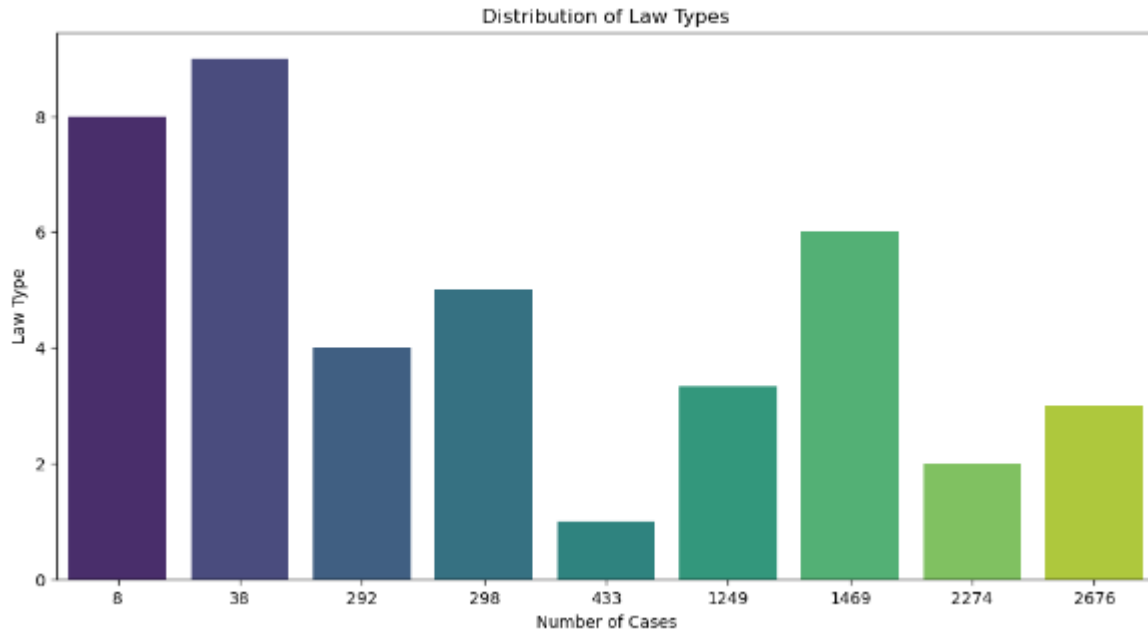


12. Most Common U.S. Citations in Supreme Court Cases



The chart presents the frequency of different U.S. citations in Supreme Court cases. The x-axis represents the number of cases, while the y-axis indicates the citation. The length of each bar corresponds to the number of times a particular citation appears in the dataset. The chart reveals that "Unknown" is the most frequently cited source, followed by "347 U.S. 909", "399 U.S. 520", and other citations with varying frequencies. This suggests that certain legal references are more commonly cited in Supreme Court opinions than others.

13. Distribution of Law Types in Supreme Court Cases



The visualization presents a bar chart illustrating the distribution of law types in Supreme Court cases from 1946 to 2015. The x-axis represents the law type, while the y-axis indicates the number of cases. The height of each bar corresponds to the frequency of cases involving that particular law type. The chart reveals that law type "8" has the highest number of cases, followed by "38" and other law types with varying frequencies. This suggests that certain types of laws are more commonly addressed by the Supreme Court compared to others.

14. Correlation Matrix of Numerical Features

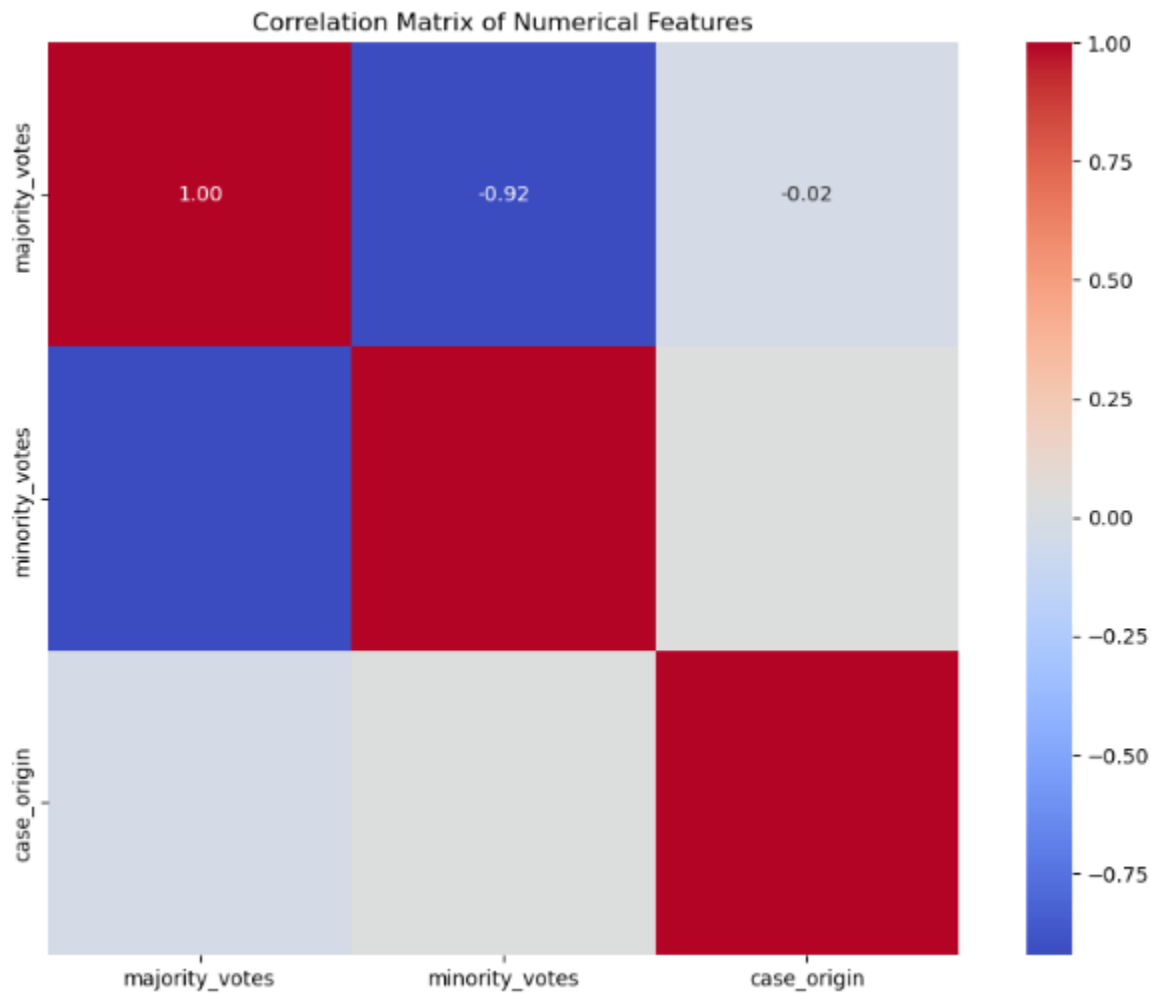
The below visualization presents a correlation matrix, which graphically depicts the relationships between numerical features in the dataset. The color scale ranges from blue to red, with blue indicating a negative correlation and red indicating a positive correlation. The intensity of the color represents the strength of the correlation.

Key Observations:

- **Strong Negative Correlation:** The correlation between `majority_votes` and `minority_votes` is strongly negative (-0.92), as expected. This indicates that as the number of majority votes increases, the number of minority votes tends to decrease, and vice versa.
- **Weak Correlations:** The correlations between `case_origin` and the other two features are relatively weak (-0.02 and 0.50). This suggests that the case origin has a limited impact on the number of majority and minority votes.

Overall, the correlation matrix provides valuable insights into the relationships between the numerical features in the dataset, confirming the expected inverse relationship

between majority and minority votes and highlighting the limited influence of case origin on these variables.



Discussion

The analysis of Supreme Court cases from 1946 to 2015 reveals several important trends that reflect the dynamics of the U.S. judicial system. The significant increase in caseload during the mid-20th century indicates a period of heightened judicial activity where legal disputes were more frequently escalated to the highest court. This may correlate with major societal changes and landmark legislation during that time.

The most frequent case names, such as "UNITED STATES v. CALIFORNIA," suggest recurring legal issues that have persisted over decades, indicating ongoing conflicts in specific areas of law. The distribution of decision types and directions highlights the court's predominant leanings, which can inform discussions on judicial philosophy and the impact of particular justices on case outcomes.

The analysis of Chief Justices and their respective case loads provides insights into the leadership styles and influences of pivotal figures in the Supreme Court's history. Understanding how different Chief Justices shaped the court's docket can illuminate the broader implications of their tenure on U.S. law.

Additionally, the correlation matrix underscores the expected inverse relationship between majority and minority votes, reinforcing the notion of judicial consensus versus dissent. This insight can inform further research on the implications of voting patterns in shaping legal precedents.

Limitations

While this analysis provides valuable insights into Supreme Court cases, it is important to acknowledge certain limitations:

- **Missing Data:** Despite systematic cleaning efforts, some missing values in the dataset could affect the comprehensiveness of the analysis, particularly in the context of majority opinion authorship and case origins.
- **Temporal Scope:** The analysis spans a specific historical period (1946-2015), which may overlook significant developments or trends outside this timeframe.
- **Case Selection Bias:** The dataset may not fully represent all cases brought before the Supreme Court, leading to potential biases in the findings. Certain types of cases or jurisdictions may be overrepresented or underrepresented.

Conclusion

This report provides a thorough overview of the analysis conducted on U.S. Supreme Court cases from 1946 to 2015. By examining trends in caseload, decision types, and the influence of individual justices, the analysis sheds light on the evolving nature of the court's jurisprudence. The findings highlight the court's role in addressing significant social issues and its responsiveness to changes in American society. Understanding these dynamics is crucial for

legal scholars, policymakers, and the public as they engage with the ongoing development of U.S. law.