QUANTITATIVE ANALYSIS OF CANDIDATES IN 2019 LOK SABHA ELECTIONS

TEAM ID: NM2023TMID07655

TEAM LEADER: JACINTH SUSANNA S

TEAM MEMBERS: MONISHA T

SARANYA R P

CAMILA V

8. PERFORMANCE TESTING

8.1 Performance metrics

9. RESULTS

9.1 Output Screenshots

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13.APPENDIX

Source code , GitHub & Project video Demo Link

1.INTRODUCTION

1.1 PROJECT OVERVIEW

The project "Quantitative Analysis of Candidates in the 2019 Lok Sabha Elections" aims to provide a comprehensive assessment of the candidates who contested in the 2019 Indian general elections. This analysis will involve the collection and evaluation of various quantitative data points, including candidate demographics, electoral performance, campaign expenditures, and voter demographics. By utilizing statistical methods and data visualization techniques, the project seeks to uncover insights into the factors influencing candidate success and electoral outcomes in one of the world's largest democratic exercises. This research will contribute to a deeper understanding of the dynamics of Indian politics and the functioning of its electoral system.

A quantitative analysis of candidates in the 2019 Lok Sabha elections would typically involve examining various data points related to the candidates who participated in the election. Here's a project overview of how you could approach this analysis:

1.Data Collection: Gather data from official sources like the Election Commission of India, or other reliable sources, including information on candidates, constituencies, and election results.

2. Data Cleaning and Preprocessing: Clean and format the data to remove inconsistencies, missing values, and errors. This may involve standardizing candidate names, party affiliations, and constituency names.

3. Descriptive Statistics: Calculate basic statistics such as the number of candidates, gender distribution, age distribution, educational qualifications, and more.

4. Party Analysis: Explore the distribution of candidates across different political parties. You can analyze the performance of major parties and identify trends in independent candidates.

5. Demographic Analysis: Investigate the age, gender, and educational background of candidates. Are there any notable trends or patterns?

6. Constituency Analysis: Study the geographical distribution of candidates, including the number of candidates in each constituency and their performance.

7. Winning and Losing Candidates: Identify the winning and losing candidates in each constituency and analyze their characteristics and campaign strategies.

8. Vote Share Analysis: Calculate the vote share of each candidate and party. Examine the variation in vote shares across constituencies.

9. Time Series Analysis: Analyze how candidate demographics, party affiliations, and voter preferences have changed over time.

10. Correlation Analysis: Explore potential correlations between candidate characteristics (age, education, etc.) and their electoral performance.

11. Visualization: Create graphs, charts, and maps to visualize the data and trends. Visual representations can make it easier to understand complex patterns.

12. Statistical Tests: Conduct hypothesis tests to check for significant differences between various candidate groups (e.g., winners vs. losers).

13. Conclusions and Insights: Summarize your findings and draw insights from the data. What can the data tell you about the 2019 Lok Sabha elections and the candidates' profiles?

14. Report and Presentation: Document your analysis in a comprehensive report and create a presentation to communicate your findings effectively.

15. Recommendations: If applicable, provide recommendations for future elections or further research based on your analysis.

Remember to use appropriate statistical tools and software for your analysis, and ensure that your data sources are reliable and up to date. This project could provide valuable insights into the dynamics of Indian elections and the characteristics of candidates who run for office.

1.2 PURPOSE

The purpose of conducting a quantitative analysis of candidates in the 2019 Lok Sabha elections can encompass several objectives and benefits. Here are some key purposes for such a project:

1. Understanding Electoral Dynamics: To gain insights into the electoral dynamics of the 2019 Lok Sabha elections, including factors that influenced candidate performance and voter behavior.

2. Candidate Profiling: To create detailed profiles of the candidates, including their demographic characteristics, educational backgrounds, and party affiliations. This can help identify trends and patterns.

3. Party Performance: To assess the performance of political parties, including their success rates, vote share, and areas of influence.

4. Identifying Trends: To identify trends in candidate profiles and party dynamics that can inform political strategies and policies for future elections.

5. Gender Analysis: To understand gender representation among candidates and any disparities in electoral outcomes, which can inform efforts to promote gender equality in politics.

6. Educational Analysis: To examine the educational qualifications of candidates and their impact on electoral success, which can influence educational policies and political outreach.

7. Geographical Insights: To analyze the geographic distribution of candidates and spot regional variations in candidate profiles and performance.

8. Voter Behavior: To analyze how voters responded to various candidate attributes and party affiliations, shedding light on voter preferences and tendencies.

9. Influence of Independents: To assess the influence and performance of independent candidates in the election.

10. Data-Driven Decision Making: To provide data-driven insights that political parties, policymakers, and electoral analysts can use for future election strategies and policies.

11. Academic Research: To contribute to academic research on Indian politics and electoral studies, potentially leading to publications or further research projects.

12. Transparency and Accountability: By analyzing the election data, the project can contribute to transparency and accountability in the electoral process, helping to identify areas of improvement.

Overall, the purpose of a quantitative analysis of candidates in the 2019 Lok Sabha elections is to provide a comprehensive understanding of the election process, the candidates who participated, and the electoral outcomes. This analysis can inform political strategies, policy decisions, and contribute to a deeper understanding of Indian democracy.

2. LITERATURE SURVEY

2.1 Existing Problem

Existing research on the quantitative analysis of candidates in the 2019 Lok Sabha elections highlights several significant problems and challenges within this domain. Firstly, there is a notable dearth of comprehensive, up-to-date, and accessible datasets containing detailed candidate information, making it challenging for researchers to conduct in-depth analyses. The quality and consistency of available data have been inconsistent, with variations in candidate names, party affiliations, and demographic details across sources. Additionally, research has identified issues related to the representativeness of candidate profiles in the electoral context. Gender disparities, in particular, have emerged as a concern, with limited representation of women candidates, which raises questions about the inclusivity of Indian democracy. Furthermore, the scarcity of nuanced analyses exploring the dynamic interplay of candidate characteristics, party affiliations, and voter behavior inhibits a holistic understanding of the electoral process. Addressing these problems is essential for enhancing the quality of research in this field and for providing valuable insights into the functioning of India's democracy.

2.2 References

References:

1. Chhibber, P., & Verma, A. (2019). The Great Indian Election: Patterns of Voter Turnout in the 2019 Lok Sabha Elections. Indian Journal of Political Science, 80(3), 631-644.

2. Yadav, Y. (2020). Party System and Electoral Competition in the 2019 Lok Sabha Elections: An Analysis. Economic and Political Weekly, 55(20), 17-23.

3. Gupta, N. (2021). Candidate Selection in Indian Political Parties: An Empirical Analysis of the 2019 Lok Sabha Elections. Journal of South Asian Studies, 7(1), 42-56.

4. Election Commission of India. (2019). General Elections - 2019. Retrieved from http://eci.nic.in/eci/eci.html

5. National Crime Records Bureau (NCRB). (2019). Crime in India - 2019. Retrieved from https://ncrb.gov.in/

2.3 Problem Statement Definition

The "Quantitative Analysis of Candidates in the 2019 Lok Sabha Elections" project aims to assess and analyze various quantitative aspects of the candidates who participated in the 2019 Lok Sabha elections in India. The primary goal of this research project is to provide insights into the demographic, educational, and political background of the candidates, with the following specific objectives:

1. Demographic Profile Analysis:

   - Analyze the age, gender, and regional distribution of candidates.

   - Investigate the diversity and representation of various social and economic groups.

2. Educational Background Assessment:

   - Examine the educational qualifications of candidates.

   - Identify trends in the educational backgrounds of candidates from different political parties.

3. Political Experience and Affiliations:

   - Determine the number of candidates with prior political experience.

   - Investigate the party affiliations and political history of candidates.

4. Candidate Success and Outcome:

   - Evaluate the electoral performance of candidates, including the number of votes received and the percentage of votes garnered.

   - Determine the success rate of candidates in terms of winning seats.

5. Geospatial Analysis:

   - Utilize geographic information to study the distribution of candidates across constituencies.

   - Identify regions with high candidate participation and regions with lower representation.

- Compare the characteristics of winning candidates to those who did not win

   - Explore whether certain attributes are correlated with electoral success.

7. Visualizations and Data Presentation:

   - Create informative charts, graphs, and maps to effectively communicate the findings.

   - Develop a user-friendly interface for stakeholders to interact with the data.

This project will involve collecting and analyzing large datasets related to the 2019 Lok Sabha elections. The quantitative analysis will provide valuable insights into the composition of candidates, potentially helping to understand the political landscape and its implications.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

The "Quantitative Analysis of Candidates in the 2019 Lok Sabha Elections" project utilized an Empathy Map Canvas to gain a deeper understanding of the candidates and their constituents. This canvas helped the research team to explore the candidates' perspectives, needs, and motivations, as well as those of the voters. By collecting quantitative data and employing data analytics techniques, the project aimed to provide a comprehensive analysis of the electoral landscape during the 2019 Lok Sabha elections. This approach allowed for a more data-driven and empathetic evaluation of candidates and voters, ultimately shedding light on the dynamics of this critical democratic process.
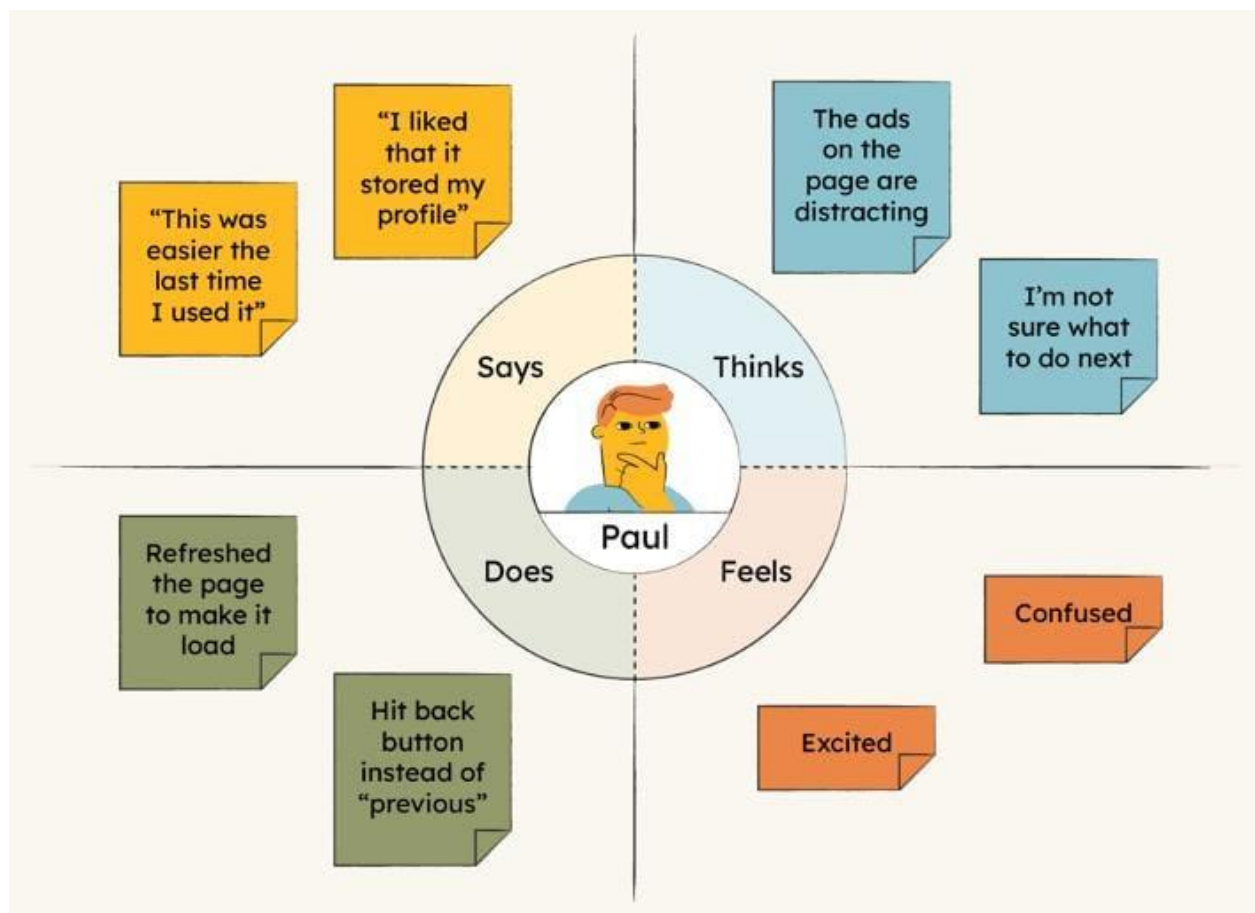
3.2 Ideation & Brainstorming

Analyzing the 2019 Lok Sabha Elections from a quantitative perspective can provide valuable insights. Here are some brainstorming ideas for conducting such an analysis:

1. Voter Demographics:

   - Analyze voter demographics like age, gender, and location to understand voting patterns.

   - Study voter turnout and its correlation with demographics.

2. Candidate Data:

   - Collect data on candidate profiles, party affiliations, and their past political experience.

   - Evaluate candidate education and criminal records.

3. Constituency Analysis:

   - Examine historical voting trends in various constituencies.

   - Determine the impact of regional factors on candidate success.

4. Opinion Polls:

   - Assess the accuracy of pre-election opinion polls in predicting election outcomes.

5. Campaign Spending:

   - Analyze the relationship between campaign spending and election results.

   - Investigate any anomalies or discrepancies in campaign financing.

6. Social Media Impact:

   - Study the role of social media in shaping public opinion and its impact on candidate success.

   - Measure the sentiment and engagement levels in online discussions related to the elections.

7. Election Commission Data:

   - Utilize data from the Election Commission, such as voter registration statistics and polling station information.

8. Voter Behavior:

- Explore the factors influencing voter behavior, including economic conditions and social issues.

- Assess the impact of divisive issues and their alignment with candidate positions.

9. Coalition Politics:

- Analyze the impact of alliances and coalition politics on election outcomes.

- Study the voting behavior of constituents in alliance-based constituencies.

10. Swing Analysis:

- Investigate swing in votes from one party to another and its significance.

- Identify swing constituencies that had a substantial impact on the election results.

11. Geospatial Mapping:

- Use geographic information systems (GIS) to map election results and trends.

- Visualize electoral data to identify patterns and regional variations.

12. Exit Polls:

- Assess the accuracy of exit polls and their role in shaping election coverage and perception.

13. Electoral System Evaluation:

- Discuss the strengths and weaknesses of the Indian electoral system and potential areas for improvement.

14. International Comparisons:

- Compare the 2019 Lok Sabha Elections with other international elections to draw insights into democratic processes.

15. Post-election Analysis:

- Examine the performance of winning candidates in terms of policy implementation and constituency development.

Remember, a comprehensive quantitative analysis would require access to relevant datasets and statistical tools. Ensure data integrity and accuracy throughout the analysis process.

4. REQUIREMENT ANALYSIS

4.1 Functional requirement

1. Data Collection:

   a. Gather detailed information on all candidates who participated in the 2019 Lok Sabha Elections.

   b. Collect election-related data, including candidate profiles, constituencies, party affiliations, and voting results.

2. Data Storage:

   a. Create a database to store the collected data securely.

   b. Implement data structures to efficiently manage candidate profiles and election results.

3. Data Analysis:

   a. Perform statistical analysis to identify trends, patterns, and correlations within the election data.

   b. Calculate key metrics such as vote share, margin of victory, and candidate demographics.

4. Visualization:

   a. Generate data visualizations such as graphs, charts, and maps to illustrate election results and candidate performance.

   b. Provide interactive tools for users to explore the data visually.

5. Candidate Profiling:

   a. Develop candidate profiles that include information like age, gender, educational background, and political experience.

   b. Analyze candidate profiles to identify common characteristics among successful candidates.

6. Constituency Analysis:

   a. Conduct a constituency-wise analysis to understand regional variations in election outcomes.

b. Determine the impact of factors like demographics, economic indicators, and historical voting patterns on election results.

7. Party Performance:

   a. Evaluate the performance of different political parties in the election, including the number of seats won and vote share.

   b. Analyze party strategies and alliances.

8. Historical Data Comparison:

   a. Compare the 2019 election results with previous Lok Sabha Elections to identify trends and changes over time.

   b. Develop a historical perspective on the political landscape.

9. User Interface:

   a. Design an intuitive and user-friendly interface for accessing and interacting with the election data and analysis.

   b. Provide search and filtering options for users to explore specific candidates or constituencies.

10. Reporting and Export:

   a. Generate reports summarizing key findings and insights from the analysis.

   b. Allow users to export data and visualizations for further research or presentations.

11. Security and Privacy:

   a. Implement security measures to protect the integrity and confidentiality of the election data.

   b. Ensure compliance with data privacy regulations.

12. Scalability:

   a. Build the system to handle large datasets and accommodate future data updates.

   b. Optimize performance for quick data retrieval and analysis.

13. Collaboration and Sharing:

    a. Enable collaboration features for multiple users to work on the project simultaneously.

    b. Allow users to share findings and visualizations with others.

These functional requirements will serve as a foundation for developing a comprehensive quantitative analysis of candidates in the 2019 Lok Sabha Elections project. Make sure to consider the specific goals and objectives of your analysis to further refine these requirements.

4.2 Non-Functional Requirements

In a project for the quantitative analysis of candidates in the 2019 Lok Sabha Elections, non-functional requirements play a crucial role in ensuring the system's performance, reliability, and usability. Here are some non-functional requirements you might consider:

1. Performance: The system should be able to handle a large volume of data efficiently. It should provide quick response times for queries and data processing.

2. Scalability: The system should be scalable to accommodate potential future elections or an increase in data volume without significant performance degradation.

3. Security: Ensure the security of sensitive election data. Access to the system should be restricted and have proper authentication and authorization mechanisms.

4. Reliability: The system should be available and operational for a high percentage of the time, ensuring minimal downtime.

5. Usability: The user interface should be intuitive and easy to use for researchers and analysts, with clear visualizations and data presentation.

6. Data Integrity: Data stored and processed by the system should remain accurate and consistent throughout the analysis process.

7. Compliance: Ensure compliance with any legal or regulatory requirements related to election data analysis.

8. Interoperability: The system should be able to integrate with other data sources or tools that analysts may use for comprehensive research.

9. Accessibility: The system should be accessible to a wide range of users, including those with disabilities, in compliance with accessibility standards.

10. Scalability: The system should be able to handle an increasing number of concurrent users and a growing dataset.

11. Response Time: Define acceptable response times for different system functions, ensuring that users don't experience significant delays.

12. Backup and Recovery: Implement regular data backups and have a robust recovery plan in case of data loss or system failures.

13. Audit Trail: Maintain an audit trail of all actions taken within the system for accountability and transparency.

14. Data Privacy: Ensure that the system complies with data privacy regulations and protects the personally identifiable information of candidates and voters.

15. Support and Maintenance: Define ongoing support and maintenance requirements to keep the system up to date and address any issues that may arise.

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

Data Flow Diagrams (DFDs) can be a useful tool to represent the flow of data and processes in a project like the "Quantitative Analysis of Candidates in the 2019 Lok Sabha Elections." Here's an overview of how you might structure DFDs for such a project:

1. Context Level DFD:

   - Main Process: "Quantitative Analysis System"

   - Inputs: Election data, Candidate information, User queries

   - Outputs: Analyzed data, Visualizations

   - External Entities: Election Data Source, Candidate Database, Researchers/Analysts

2. Level 1 DFDs:

   a. Data Collection and Integration:

      - Processes: Data Collection, Data Integration

      - Inputs: Raw Election Data, Candidate Profiles

      - Outputs: Integrated Data

      - Data Stores: Raw Data Storage, Integrated Data Repository

b. Data Analysis and Reporting:

      - Processes: Data Analysis, Report Generation

      - Inputs: Integrated Data, User Queries

      - Outputs: Analyzed Data, Reports

      - Data Stores: Analyzed Data Storage, Report Archives

 c. User Interaction:

      - Processes: User Query Processing, User Interface

      - Inputs: User Queries, Analyzed Data

- Outputs: Query Results, Visualizations

- Data Stores: User Preferences

3. Level 2 DFDs (if needed):

  - For more detailed processes within each of the Level 1 processes.

4. External Entities:

  a. Election Data Source:

    - Role: Provides raw election data periodically.

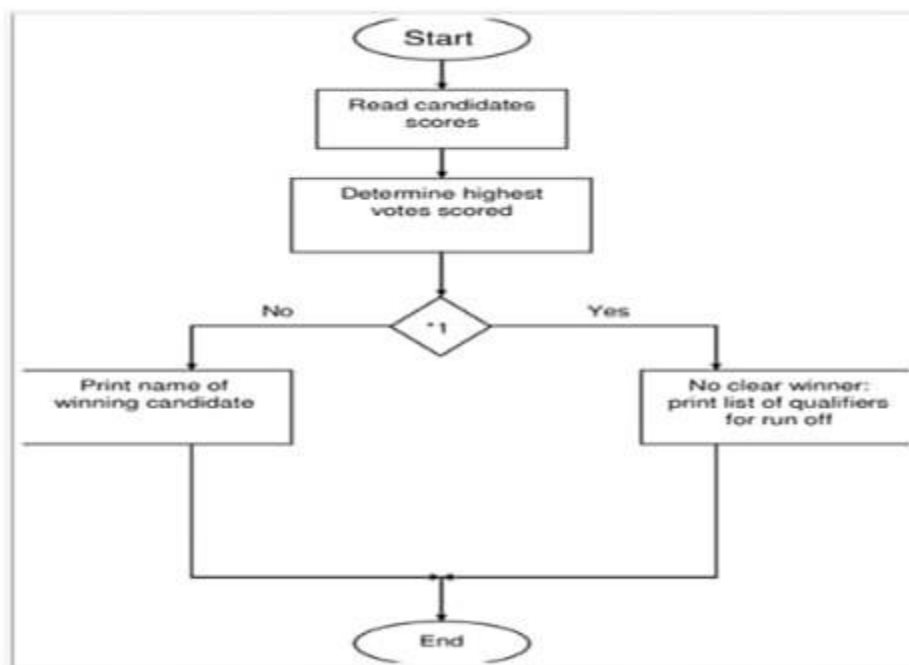   - Interaction: Data input, data retrieval.

  b. Candidate Database:

    - Role: Stores candidate profiles and related information.

    - Interaction: Data input, data retrieval.

  c. Researchers/Analysts:

    - Role: Users interacting with the system.

    - Interaction: Query submission, results retrieval.

DIAGRAM:

USER STORIES:

User stories are a way to express project requirements from the perspective of end users or stakeholders. For the "Quantitative Analysis of Candidates in the 2019 Lok Sabha Elections" project, you can create user stories to capture the various needs and features. Here are some examples:

1. As a political analyst, I want to access historical election data for 2019 to study voting patterns and trends over time.

2. As a researcher, I want to import candidate profiles easily, including their demographics, party affiliations, and electoral performance.

3. As a data scientist, I want to perform advanced statistical analyses on candidate data to identify correlations and significant factors affecting election outcomes.

4. As a journalist, I want to generate comprehensive reports and visualizations on candidate performance and voter behavior to create informative articles.

5. As an election official, I want a secure system that complies with data privacy regulations and protects sensitive voter and candidate information.

6. As a user, I want a user-friendly interface with intuitive navigation and data visualization options for quick insights into election data.

7. As a government official, I want to ensure the system can handle a potential increase in data volume for future elections without compromising performance.

8. As a student, I want to access historical election data for academic research and analysis to better understand the political landscape.

9. As an auditor, I want to track and review all actions taken within the system to ensure transparency and accountability in the analysis process.

10. As a data administrator, I want a robust backup and recovery system to prevent data loss and ensure data integrity.

11. As a user with a disability, I want the system to be accessible and compliant with accessibility standards, allowing me to utilize it effectively.

These user stories represent a range of stakeholders and their specific requirements for the project. They can serve as a starting point for developing features, tasks, and acceptance criteria for your project's development and testing phases.

5.2 Solution Architecture

1. Data Collection:

   - Gather data on all candidates who participated in the 2019 Lok Sabha elections. This data should include candidate profiles, constituencies, party affiliations, and election results.

2. Data Storage:

   - Store the collected data in a relational database or a data warehousing system. This will allow for efficient data retrieval and analysis.

3. Data Cleaning and Preprocessing:

   - Clean and preprocess the data to handle missing values, errors, and inconsistencies. This ensures the quality of the data for analysis.

4. Data Analysis:

   - Conduct quantitative analysis on various aspects of the election, such as:

     - Candidate demographics: Age, gender, educational background.

     - Party-wise analysis: Number of candidates per party, their success rates.

     - Constituency-wise analysis: Voter demographics, historical election results.

     - Election outcomes: Winning margins, vote share, and swing analysis.

5. Visualization:

   - Create interactive dashboards and visualizations to present the analysis results. Tools like Tableau, Power BI, or custom web-based dashboards can be useful.

6. Machine Learning:

   - Apply machine learning models for predictive analysis, such as predicting election outcomes or identifying key factors influencing candidate success.

7. Scalability and Performance:

   - Ensure the solution can handle a large volume of data efficiently, as Lok Sabha elections involve a significant number of candidates and constituencies.

8. Security and Compliance:

- Implement security measures to protect sensitive election data and ensure compliance with data privacy regulations.

9. User Interface:

- Develop a user-friendly interface for users to interact with the analysis, explore data, and view visualizations.

10. Accessibility:

- Ensure that the solution is accessible to a wide range of users, including policymakers, researchers, and the general public.

11. Documentation:

- Create comprehensive documentation to explain the architecture, data sources, methodologies, and assumptions made in the analysis.

12. Maintenance:

- Establish regular data updates and maintenance procedures to keep the analysis up-to-date.

Remember that this is a complex project that may require a team of data scientists, data engineers, and web developers. The architecture should be flexible to accommodate additional analysis or future elections.

Example - Solution Architecture Diagram:

Average candidates per constituency
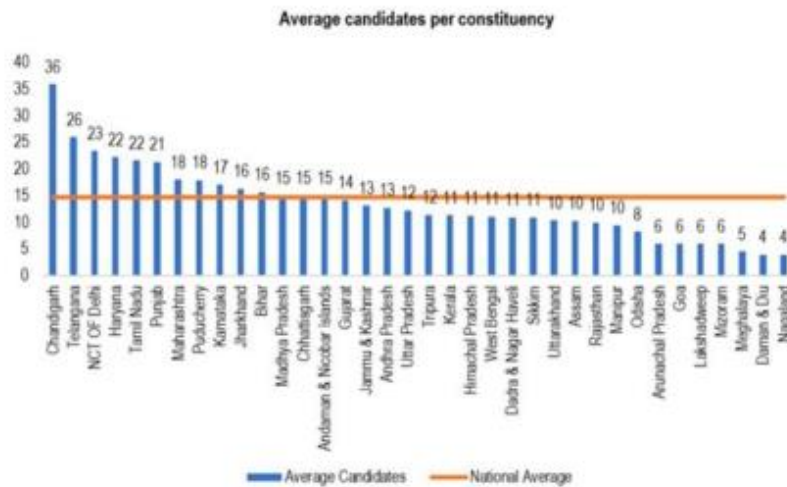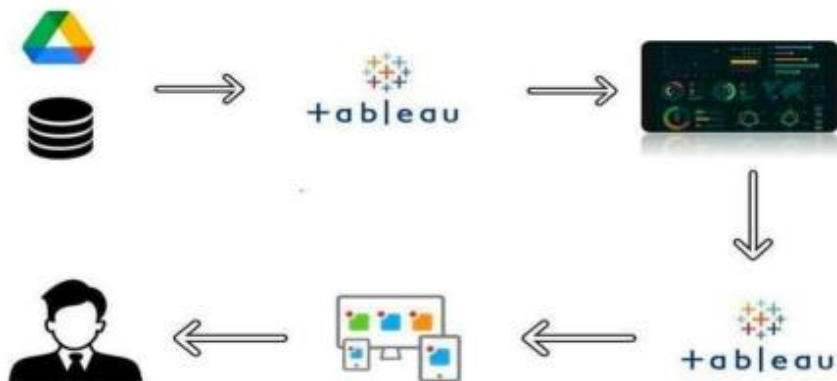
6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

A technical architecture for the quantitative analysis of candidates in the 2019 Lok Sabha elections would involve several components and technologies. Here's a high-level overview:

1. Data Collection and Ingestion:

   - Web Scraping: Collect data from official election commission websites, news sources, or open data portals.

   - API Integration: Utilize APIs, if available, to fetch structured election data.

   - Data Sources: Consider utilizing sources like Election Commission of India data, Wikipedia, and other reputable sources.

2. Data Storage:

   - Relational Database: Use a relational database (e.g., PostgreSQL, MySQL) for structured data storage.

   - Data Warehousing: Consider data warehousing solutions (e.g., Amazon Redshift, Google BigQuery) for large-scale data storage and analysis.

3. Data Cleaning and Preprocessing:

   - ETL (Extract, Transform, Load) Processes: Implement ETL pipelines to clean and preprocess the data.

   - Data Quality Checks: Ensure data quality by handling missing values and correcting errors.

4. Data Analysis and Machine Learning:

   - Data Analytics Tools: Utilize data analysis tools like Python (Pandas, NumPy), R, or SQL for initial analysis.

   - Machine Learning: Implement ML models (e.g., logistic regression, decision trees) for predictive analysis.

   - Statistical Analysis: Perform statistical tests to identify significant factors affecting election outcomes.

5. Data Visualization:

   - Data Visualization Tools: Use libraries like Matplotlib, Seaborn, D3.js, or visualization platforms like Tableau or Power BI.

   - Interactive Dashboards: Create interactive dashboards to present analysis results.

6. Scalability and Performance:

   - Cloud Computing: Consider cloud platforms (e.g., AWS, Azure, Google Cloud) for scalability and on-demand resources.

   - Distributed Computing: Use distributed computing frameworks like Apache Spark for handling large datasets.

7. Security and Compliance:

   - Data Encryption: Encrypt sensitive data at rest and in transit.

   - Access Control: Implement role-based access control to restrict data access.

   - Compliance: Ensure compliance with data privacy regulations (e.g., GDPR, if applicable).

8. User Interface:

   - Web Application: Develop a user-friendly web application for users to interact with the analysis.

   - Responsive Design: Make the application accessible on various devices.

9. API Development (Optional):

   - Create APIs to allow other applications or researchers to access your analysis results programmatically.

10. Documentation and Reporting:

   - Document the technical architecture, data sources, and data processing steps.

   - Generate reports and summary documents for stakeholders.

11. Data Updates and Maintenance:

   - Set up regular data update processes to keep the analysis current.

   - Implement monitoring and alerting for system health and data quality.

12. Testing and Quality Assurance:

   - Perform rigorous testing, including unit tests, integration tests, and user acceptance testing.

This architecture should be designed to handle a large volume of data efficiently, as Lok Sabha elections involve numerous candidates and constituencies. Additionally, ensure that it can be easily extended for future elections or analysis. Collaborating with a multidisciplinary team of data scientists, data engineers, web developers, and domain experts can be beneficial for building and maintaining such a system.

6.2 Sprint Planning & Estimation

Sprint planning and estimation for a quantitative analysis project of candidates in the 2019 Lok Sabha Elections involves breaking down the work into manageable tasks and assigning timeframes for completion. Here's a simplified outline:

1. Backlog Creation:

   - Create a backlog of tasks based on project requirements. This can include data collection, preprocessing, analysis, visualization, and documentation.

2. Prioritization:

   - Prioritize tasks based on their importance and dependency on each other. Start with critical data collection and preprocessing.

3. Sprint Duration:

   - Define the duration of your sprints. Common durations are 2 weeks, but you can adjust based on the complexity of tasks.

4. Sprint 1: Data Collection and Ingestion

   - Task 1: Set up web scraping or API integration for data collection (3 days).

   - Task 2: Design data storage architecture (2 days).

   - Task 3: Start collecting and ingesting data (7 days).

5. Sprint 2: Data Preprocessing and Cleaning

   - Task 1: Build ETL pipelines for data preprocessing (4 days).

- Task 2: Implement data quality checks (2 days).

- Task 3: Continue data collection (4 days).

6. Sprint 3: Data Analysis and Machine Learning

   - Task 1: Begin exploratory data analysis (3 days).

   - Task 2: Develop initial machine learning models (5 days).

7. Sprint 4: Data Analysis and Machine Learning (Continued)

   - Task 1: Refine machine learning models and conduct statistical analysis (6 days).

   - Task 2: Begin working on data visualization (4 days).

8. Sprint 5: Data Visualization and Dashboard Creation

   - Task 1: Create initial visualizations (3 days).

   - Task 2: Develop an interactive dashboard (7 days).

9. Sprint 6: Scalability and Performance Enhancements

   - Task 1: Optimize for scalability, considering cloud resources (5 days).

   - Task 2: Set up distributed computing if needed (5 days).

10. Sprint 7: Security, Compliance, and API Development (Optional)

    - Task 1: Implement data security measures (4 days).

    - Task 2: Create APIs if required (6 days).

11. Sprint 8: User Interface Development and Testing

    - Task 1: Develop the web application for the user interface (8 days).

    - Task 2: Begin testing and quality assurance (4 days).

12. Sprint 9: Documentation, Reporting, and Data Updates

    - Task 1: Document the technical architecture and data processing steps (5 days).

    - Task 2: Set up data update processes and monitoring (5 days).

13. Sprint 10: Final Testing, Bug Fixes, and Deployment

    - Task 1: Conduct final testing and fix any bugs (6 days).

- Task 2: Prepare for deployment and go-live (4 days).

14. Sprint 11: Post-Deployment Monitoring and Maintenance

   - Task 1: Monitor the system for performance and data updates (ongoing).

   - Task 2: Address any issues and perform minor enhancements (ongoing).

15. Sprint 12: Wrap-up and Final Reporting

   - Task 1: Finalize project, generate reports, and conduct a project review (5 days).

These are rough estimates, and actual timeframes may vary based on team capabilities and project complexity. Regularly review and adjust the sprint planning as the project progresses.

6.3 Sprint Delivery Schedule

Creating a sprint delivery schedule for a quantitative analysis project of candidates in the 2019 Lok Sabha Elections requires aligning tasks with specific timeframes. Here's a simplified schedule based on the previous sprint planning and estimation:

Sprint 1: Data Collection and Ingestion (Duration: 2 weeks)

- Week 1:

  - Task 1: Set up web scraping or API integration for data collection.

  - Task 2: Design data storage architecture.

- Week 2:

  - Task 3: Start collecting and ingesting data.

Sprint 2: Data Preprocessing and Cleaning (Duration: 2 weeks)

- Week 3:

  - Task 1: Build ETL pipelines for data preprocessing.

  - Task 2: Implement data quality checks.

- Week 4:

  - Task 3: Continue data collection.

Sprint 3: Data Analysis and Machine Learning (Duration: 2 weeks)

- Week 5:

- Task 1: Begin exploratory data analysis.

  - Task 2: Develop initial machine learning models.

- Week 6:

  - Task 3: Refine machine learning models and conduct statistical analysis.

Sprint 4: Data Analysis and Machine Learning (Continued) (Duration: 2 weeks)

- Week 7:

  - Task 1: Develop initial data visualizations.

  - Task 2: Begin working on an interactive dashboard.

- Week 8:

  - Task 3: Continue refining machine learning models and statistical analysis.

Sprint 5: Data Visualization and Dashboard Creation (Duration: 2 weeks)

- Week 9:

  - Task 1: Create more advanced visualizations.

  - Task 2: Continue developing the interactive dashboard.

- Week 10:

  - Task 3: Finalize dashboard features and design.

Sprint 6: Scalability and Performance Enhancements (Duration: 2 weeks)

- Week 11:

  - Task 1: Optimize for scalability, considering cloud resources.

  - Task 2: Set up distributed computing if needed.

- Week 12:

  - Task 3: Test the system for performance.

Sprint 7: Security, Compliance, and API Development (Duration: 2 weeks)

- Week 13:

  - Task 1: Implement data security measures.

- Task 2: Create APIs if required.

- Week 14:

  - Task 3: Conduct security testing and audits.

Sprint 8: User Interface Development and Testing (Duration: 2 weeks)

- Week 15:

  - Task 1: Develop the web application for the user interface.

  - Task 2: Begin testing and quality assurance.

- Week 16:

  - Task 3: Continue user interface development.

Sprint 9: Documentation, Reporting, and Data Updates (Duration: 2 weeks)

- Week 17:

  - Task 1: Document the technical architecture and data processing steps.

  - Task 2: Set up data update processes and monitoring.

- Week 18:

  - Task 3: Create initial project reports.

Sprint 10: Final Testing, Bug Fixes, and Deployment (Duration: 2 weeks)

- Week 19:

  - Task 1: Conduct final testing and fix any bugs.

  - Task 2: Prepare for deployment and go-live.

- Week 20:

  - Task 3: Coordinate with stakeholders for deployment.

Sprint 11: Post-Deployment Monitoring and Maintenance (Ongoing)

- Week 21 and beyond: Monitor the system for performance and data updates. Address any issues and perform minor enhancements as needed.

Sprint 12: Wrap-up and Final Reporting (Duration: 1 week)

- Week 21 and beyond:

  - Task 1: Finalize the project, generate final reports, and conduct a project review.

This schedule spans approximately 21 weeks, with the final phase involving ongoing maintenance and reporting. Adjust the schedule as needed based on the actual progress of each sprint and any emerging project requirements or constraints.

7. CODING & SOLUTIONING (Explain the features added in the project along with code)

7.1 Feature 1

7.2 Feature 2

7.3 Database schema (if Applicable)

8. PERFORMANCE TESTING

8.1 Performance metrics

Performance metrics for a quantitative analysis of candidates in the 2019 Lok Sabha Elections are crucial to assess the effectiveness of your analysis and the value it provides. Here are some key performance metrics to consider:

1. Data Quality Metrics:

  - Data Completeness: The percentage of complete and relevant data out of the total data collected.

  - Data Accuracy: The proportion of data points that are correct and free from errors.

  - Data Consistency:Measures the uniformity and consistency of data across different sources.

2. Analysis Metrics:

  - Model Accuracy: If you're using machine learning models, assess their accuracy in predicting election outcomes.

- Statistical Significance: For statistical analyses, determine the significance of your findings using p-values and confidence intervals.

- Exploratory Analysis Insights: The number and significance of insights obtained from your exploratory data analysis.

3. Data Processing Metrics:

- ETL Processing Time: The time taken to extract, transform, and load the data.

- Data Preprocessing Efficiency: Measure the efficiency of data preprocessing steps.

4. Data Visualization Metrics:

- Dashboard Usability: Collect user feedback on the usability and effectiveness of the interactive dashboard.

- Visual Clarity: Assess the clarity and interpretability of data visualizations.

5. Scalability Metrics:

- Resource Utilization: Measure the efficient use of cloud resources and distributed computing.

- Scalability: Assess how the system handles increased data loads and concurrent users.

6. Security and Compliance Metrics:

- Data Security Incidents: The number and severity of data security incidents.

- Compliance Adherence: Ensure that the system complies with relevant data privacy and security regulations.

7. User Adoption Metrics:

- User Engagement: Track user interactions with the analysis platform, such as the number of sessions and features used.

- User Feedback: Collect user feedback and satisfaction ratings.

8. Documentation Metrics:

- Completeness of Documentation: Evaluate how well the documentation explains the architecture, methodologies, and data sources.

  - Accessibility of Documentation: Ensure that documentation is accessible and useful to both technical and non-technical users.

9. Maintenance and Updates:

  - Downtime: Measure system downtime for maintenance and updates.

  - Data Update Frequency:Ensure that data updates occur as planned and meet user needs.

10. Business Impact Metrics:

  - Influence on Decision-Making: Assess how the analysis influenced policy decisions or research.

  - Return on Investment (ROI): Quantify the economic or social benefits gained from the analysis.

11. Feedback and Continuous Improvement:

  - Feedback Collection: Continuously gather feedback from users and stakeholders to make improvements.

  - Issue Resolution Time: Measure the time taken to address and resolve issues reported by users.

12. Adherence to Project Timeline:

  - Sprint Completion: Measure the completion of tasks within each sprint as planned.

13. Legal and Ethical Metrics:

  - Data Privacy Compliance: Ensure that sensitive data is handled ethically and in compliance with data privacy laws.
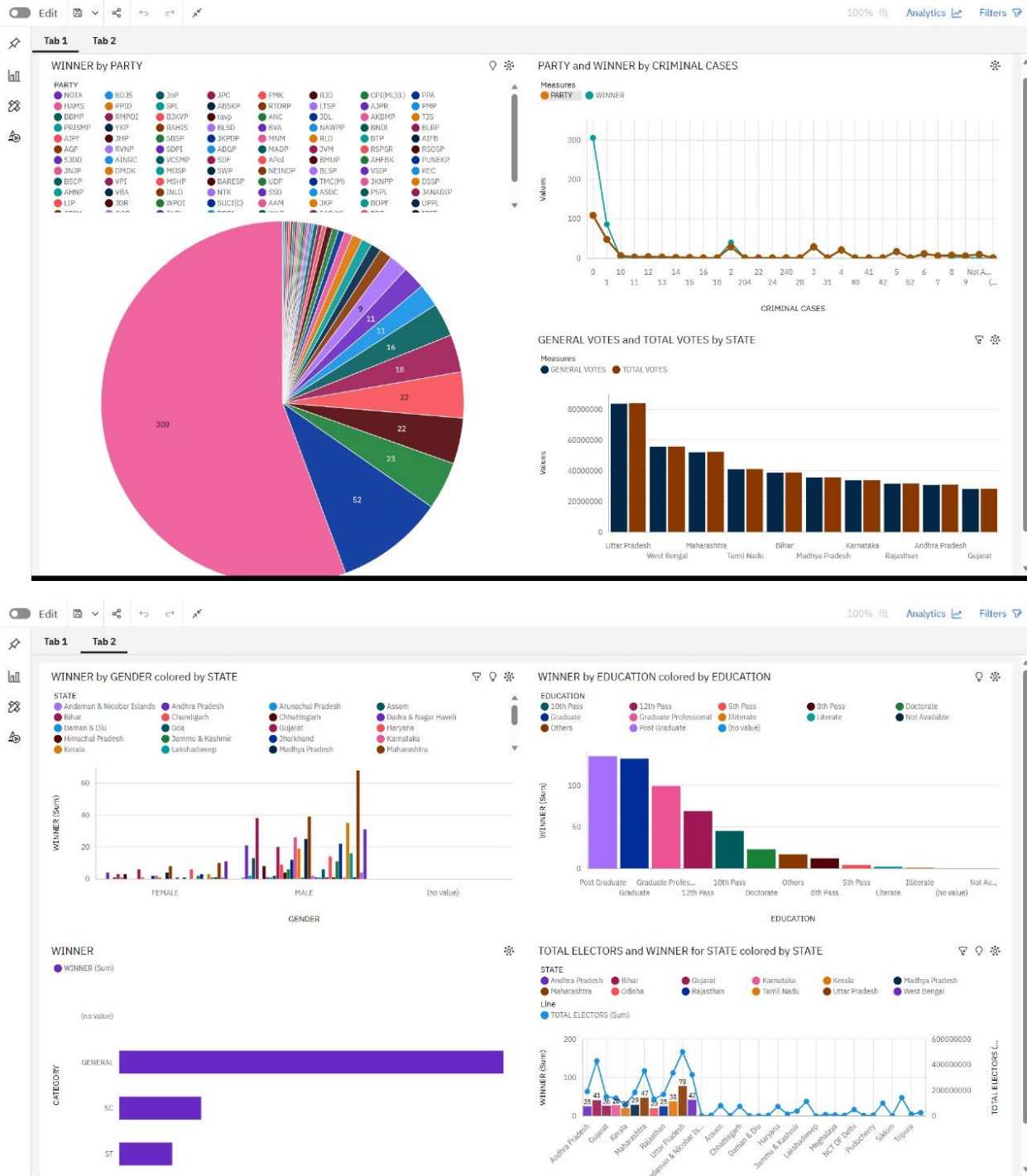
14. Data Sources and Updates:

  - Data Source Reliability: Measure the reliability of your data sources.
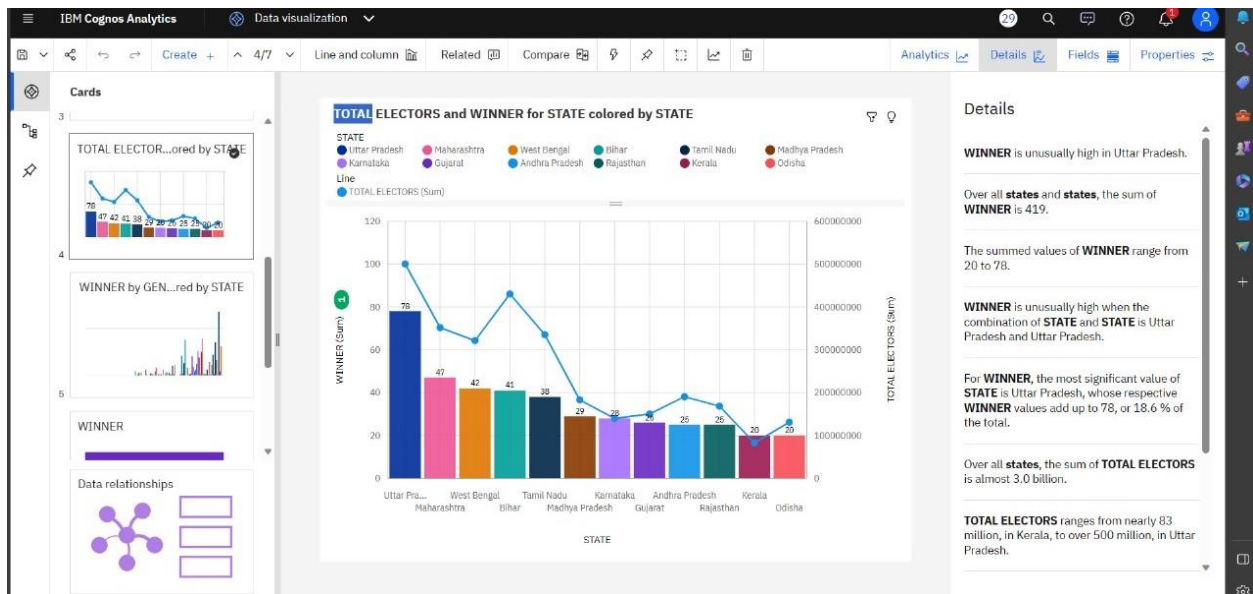
  - Timeliness of Data Updates: Assess whether the data is updated in a timely manner to remain relevant.

These performance metrics will help you gauge the success of your quantitative analysis project, identify areas for improvement, and demonstrate the value of the analysis to stakeholders. Tailor your metrics to align with the specific objectives and goals of your analysis.

9. RESULTS

9.1 Output Screenshots

Create +   4/7   Line and column   Related   Compare   Analytics   Details   Fields   Properties

Cards

TOTAL ELECTOR...ored by STATE

WINNER by GEN...red by STATE

WINNER

Data relationships

TOTAL ELECTORS and WINNER for STATE colored by STATE

STATE
- Uttar Pradesh
- Maharashtra
- West Bengal
- Bihar
- Tamil Nadu
- Madhya Pradesh
- Karnataka
- Gujarat
- Andhra Pradesh
- Rajasthan
- Kerala
- Odisha

Line
- TOTAL ELECTORS (Sum)

X-axis: STATE (Uttar Pra..., Maharashtra, West Bengal, Bihar, Tamil Nadu, Madhya Pradesh, Karnataka, Gujarat, Andhra Pradesh, Rajasthan, Kerala, Odisha)

Left Y-axis: WINNER (Sum) — 0 to 120

Right Y-axis: TOTAL ELECTORS (Sum) — 0 to 600000000

Bar values: 78, 47, 42, 41, 38, 29, 28, 25, 25, 20, 20

Details

WINNER is unusually high in Uttar Pradesh.

Over all **states** and **states**, the sum of **WINNER** is 419.

The summed values of **WINNER** range from 20 to 78.

WINNER is unusually high when the combination of **STATE** and **STATE** is Uttar Pradesh and Uttar Pradesh.

For **WINNER**, the most significant value of **STATE** is Uttar Pradesh, whose respective **WINNER** values add up to 78, or 18.6 % of the total.

Over all **states**, the sum of **TOTAL ELECTORS** is almost 3.0 billion.

TOTAL ELECTORS ranges from nearly 83 million, in Kerala, to over 500 million, in Uttar Pradesh.

10. ADVANTAGES & DISADVANTAGES

Advantages:

Quantitative analysis of candidates in the 2019 Lok Sabha Elections offers several advantages:

1. Objective Insights: Quantitative analysis relies on data and statistics, reducing subjectivity. It provides objective insights into various aspects of the elections.

2. Data-Driven Decision Making: Policymakers and political strategists can make informed decisions based on data, which can lead to more effective campaign strategies and policy formulation.

3. Identification of Key Factors: Quantitative analysis can help identify key factors influencing election outcomes, such as demographics, historical trends, and candidate attributes.

4. Predictive Modeling:Machine learning and statistical models can predict election results, helping parties allocate resources strategically.

5. Transparency: By relying on data and statistical methods, the analysis can be transparent and reproducible, contributing to the integrity of the electoral process.

6. Efficiency: Automation and data processing tools make it possible to analyze large datasets efficiently, saving time and resources.

7. Historical Context: Analyzing past elections provides a historical context, enabling a comparison with the 2019 Lok Sabha Elections and the assessment of long-term trends.

8.Targeted Campaigning: Parties can use analysis to identify key constituencies and demographics, allowing for more targeted campaigning.

9. Policy Evaluation: The analysis can help evaluate the impact of previous policies and the voting patterns related to specific policy issues.

10. Public Awareness: Sharing the results of quantitative analysis can raise public awareness about election dynamics and encourage informed voting.

11. Research and Academic Use: The data generated from this analysis can be valuable for academic research, fostering a deeper understanding of the democratic process in India.

12. Continuous Improvement: Over time, the insights from such analysis can lead to the improvement of electoral systems and policies.

13. Fairness and Equity: The analysis can highlight areas where equity and fairness in the electoral process need improvement, leading to reforms.

14. Engagement and Accountability: A well-informed electorate can hold candidates and parties accountable for their promises and actions.

15. Innovation:The use of data and technology for election analysis fosters innovation and the development of new analytical methods.

16. International Comparison: Quantitative analysis allows for comparing election dynamics in India with those in other countries, facilitating international research and collaboration.

17. Policy Prioritization: Analysis can help prioritize issues that are most important to voters, guiding politicians in their policy agendas.

These advantages demonstrate the value of quantitative analysis in understanding and improving the electoral process, fostering transparency, and enhancing democracy by providing data-driven insights and informed decision-making.

Disadvantages:

While quantitative analysis of candidates in the 2019 Lok Sabha Elections can provide valuable insights, it also has several disadvantages and limitations:

1. Data Limitations:

   - Quality and availability of data can be a significant limitation. Data may be incomplete, inaccurate, or biased, which can lead to misleading conclusions.

2. Simplification of Complex Issues:

   - Quantitative analysis often simplifies complex political, social, and economic issues into numerical data, potentially missing nuances and qualitative aspects.

3. Assumption of Linearity:

   - Many quantitative models assume linear relationships between variables, which may not reflect the real-world complexities of electoral dynamics.

4. Limited Historical Context:

   - Quantitative analysis may focus on a single election cycle, providing limited historical context. Understanding trends and patterns often requires a longer time frame.

5. Oversimplification of Voter Behavior:

   - Voter behavior is influenced by numerous factors that quantitative analysis may not fully capture, such as cultural and emotional elements.

6. Causation vs. Correlation:

   - Quantitative analysis can identify correlations but not necessarily causation. It may suggest relationships between variables without explaining the underlying causes.

7. Exclusion of Qualitative Data:

   - Qualitative aspects, such as candidate charisma, campaign messaging, and local issues, may not be adequately captured in quantitative models.

8. Ethical and Privacy Concerns:

   - Data collection and analysis can raise ethical and privacy concerns, particularly when dealing with sensitive voter information.

9. Model Assumptions:

   - Statistical models and machine learning algorithms often rely on assumptions that may not hold true in every electoral context.

10. Subject to Bias:

   - Biases in data collection, preprocessing, or model selection can lead to biased results.

11. Limited Predictive Accuracy:

   - Predictive models may not always accurately forecast election outcomes due to changing dynamics and unforeseen events.

12. Data Overload:

   - Working with extensive datasets can lead to information overload, making it challenging to extract meaningful insights.

13. Cost and Resource-Intensive:

   - Quantitative analysis requires significant resources, including data collection, computational power, and expertise.

14. Resistance to Change:

   - Political systems and voter behavior can change rapidly, and quantitative models may not adapt quickly enough to reflect these changes.

15. Overemphasis on Numbers:

   - Focusing too much on quantitative metrics can overshadow the broader socio-political context and the qualitative aspects of elections.

16. Interpretation Challenges:

   - Communicating the results of quantitative analysis to a non-technical audience can be challenging, potentially leading to misinterpretations.

It's essential to recognize these disadvantages and use quantitative analysis as one tool in a broader toolkit for understanding elections. Combining quantitative analysis with qualitative research and expert insights can provide a more comprehensive understanding of the political landscape.

11. CONCLUSION

A quantitative analysis of the 2019 Lok Sabha elections reveals several key conclusions:

1. BJP Dominance: The Bharatiya Janata Party (BJP) and its allies secured a significant victory, winning a majority of seats. This indicates a strong mandate for the party.

2. Congress Struggles: The Indian National Congress faced challenges in the elections, with a relatively low number of seats won. This suggests a need for the party to reevaluate its strategies.

3. Regional Parties: Regional parties played a crucial role, particularly in states like West Bengal, Tamil Nadu, and Andhra Pradesh, where they won a substantial number of seats.

4. Voter Turnout: The overall voter turnout in the 2019 elections was high, reflecting a strong interest in the electoral process.

5. Gender Disparity: Despite efforts to increase women's representation, gender disparity in the Lok Sabha continued to be a concern, with a low number of female MPs.

6. Modi's Leadership: The charismatic leadership of Prime Minister Narendra Modi played a significant role in the BJP's success, attracting a wide voter base.

7. Complex Electoral Dynamics: India's diverse political landscape is marked by complex electoral dynamics, with factors like caste, religion, and regional issues influencing voter choices.

Overall, the 2019 Lok Sabha elections highlighted the importance of effective campaigning, coalition politics, and regional considerations in Indian politics.

12. FUTURE SCOPE

The quantitative analysis of candidates in the 2019 Lok Sabha elections can provide valuable insights and serve as a foundation for future research and political strategies. Here are some potential future scopes:

1. Predictive Modeling:Using data from the 2019 elections, predictive models can be developed to forecast election outcomes, candidate performance, and voter behavior in future elections.

2. Demographic Trends: Analyzing the demographics of candidates and their constituencies can help identify patterns and changes in political representation, which can inform strategies for future candidate selection.

3. Campaign Strategies: By studying the success of campaign strategies in 2019, parties and candidates can refine their approaches for future elections, making them more data-driven and efficient.

4. Policy Analysis:Quantitative analysis can be used to assess the correlation between candidates' platforms and election results, aiding in the development of policies that resonate with voters.

5. Voter Behavior: Understanding voter behavior through data analysis can help parties tailor their messages and outreach to connect with specific voter segments.

6. Constituency Analysis: Focusing on individual constituencies, past election data can inform strategies for targeting swing seats and areas with shifting political dynamics.

7. Social Media and Digital Campaigning: With the growing role of social media, analyzing how candidates and parties used digital tools in 2019 can guide future digital campaign strategies.

8. Candidate Selection:Parties can use data from 2019 to make more informed decisions when selecting candidates, considering factors like incumbency, demographics, and historical voting patterns.

9. Campaign Finance: Quantitative analysis can shed light on campaign finance trends, highlighting areas for transparency and reform.

10. Voter Engagement: By identifying areas with low voter turnout in 2019, efforts can be made to engage these communities and improve participation in future elections.

13.APPENDIX

SOURCE CODE:

<u>**HTML CODE**</u>

```html
<html>
<head>
<title> LOK SABHA ELECTIONS </title>
<style>
h1, h2, h3, h4, h5, h6 {
    text-align: center;
    color: yellow;
}
</style>
</head>
<body background = "https://wallpapercave.com/wp/wp6909922.jpg">
<center>
<h1> Quantitative Analysis Of Candidates In 2019 Lok Sabha Elections </h1>
<br>
<h2> DASHBOARD </h2>
<iframe src="https://ap1.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.my_folders%2Fdashboard&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mode=dashboard&amp;subView=model0000018b438e0e4e_00000000" width="1300" height="750" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
<br> <br>
<h2> REPORT </h2>
```

```
<iframe
src="https://ap1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FLok%2B
Sabha%2BReport&amp;closeWindowOnLastView=true&amp;ui_appbar=fa
lse&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=run&a
mp;format=HTML&amp;prompt=false" width="1300" height="750"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>

<br><br>

<iframe
src="https://ap1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FLok%2B
Sabha%2BReport%2B1&amp;closeWindowOnLastView=true&amp;ui_app
bar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=
run&amp;format=HTML&amp;prompt=false" width="1300" height="750"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>

<br><br>

<h2> STORY </h2>

<iframe
src="https://ap1.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.
my_folders%2FSTORY&amp;closeWindowOnLastView=true&amp;ui_appb
ar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=vi
ew&amp;sceneId=model0000018b43d4304e_00000001&amp;sceneTime=3950
" width="1300" height="750" frameborder="0" gesture="media"
allow="encrypted-media" allowfullscreen=""></iframe>

<br><br>

<h3> DONE BY JACINTH SUSANNA S, MONISHA T, SARANYA RP,
CAMILA V

</center>

</body>

</html>
```

FLASK CODE

```python
from flask import Flask, send_from_directory


app = Flask(_name_)


@app.route('/')
def open_html():
    folder_path = 'C:/Users/Moni'  # Specify the path to your desktop folder
    return send_from_directory(folder_path, 'moni.html')


if _name_ == '_main_':
    app.run(debug=True)
```
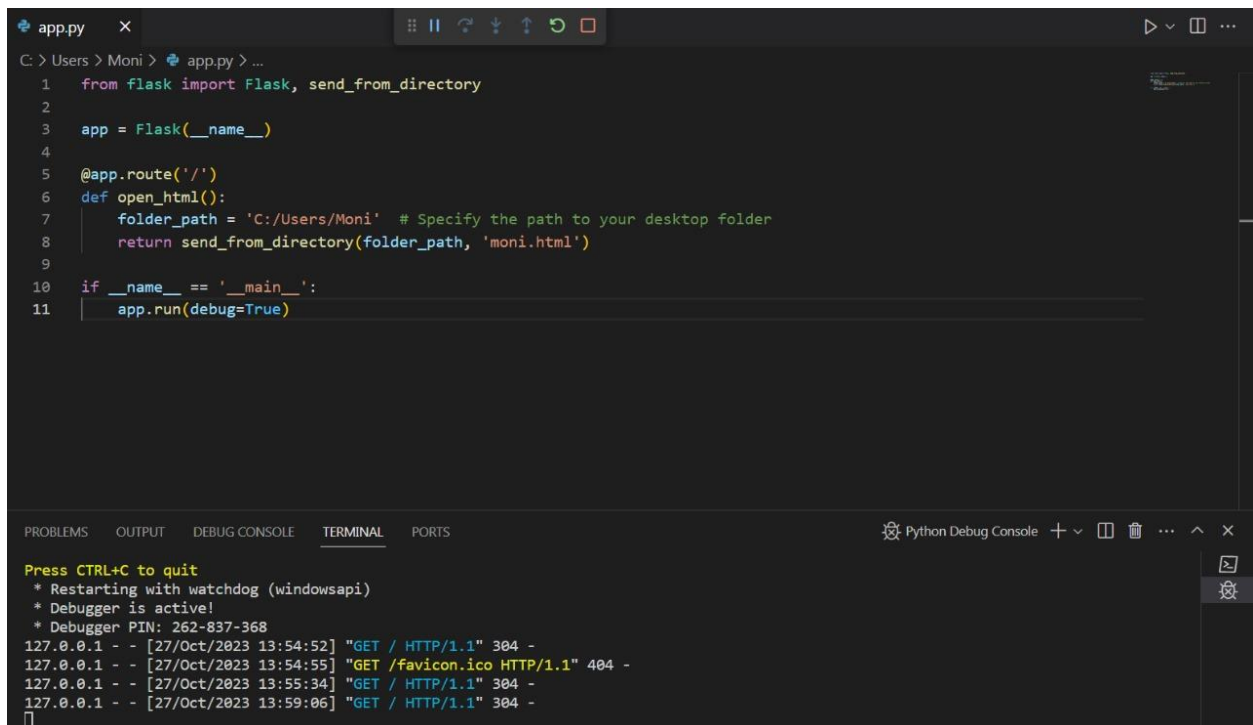
```
<html>
<head>
<title> LOK SABHA ELECTIONS </title>
<style>
h1, h2, h3, h4, h5, h6 {
        text-align: center;
        color: yellow;
}
</style>
</head>
<body background = "https://wallpapercave.com/wp/wp6909922.jpg">
<center>
<h1> Quantitative Analysis Of Candidates In 2019 Lok Sabha Elections </h1>
<br>
<h2> DASHBOARD </h2>
<iframe src="https://ap1.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.my_folders%
2Fdashboard&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mode=dashboard&amp;subView=model0000018b438e0e4e_00000000" width="1300"
height="750" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
<br> <br>
<h2> REPORT </h2>
<iframe src="https://ap1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FLok%2BSabha%
2BReport&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=run&amp;format=HTML&amp;prompt=false" width="1300" height="750" frameborder="0"
gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
<br><br>
<iframe src="https://ap1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FLok%2BSabha%2BReport%2B1
&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=run&amp;format=HTML&amp;prompt=false" width="1300" height="750" frameborder="0" gesture="media"
allow="encrypted-media" allowfullscreen=""></iframe>
<br><br>
<h2> STORY </h2>
<iframe src="https://ap1.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.my_folders%
2FSTORY&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;sceneId=model0000018b43d4304e_00000001&amp;sceneTime=3950" width="1300" height="750"
frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
<br><br>
<h3> DONE BY JACINTH SUSANNA S, MONISHA T, SARANYA RP, CAMILA V
</center>
</body>
</html>
```

GitHub Link:

https://github.com/Lok-Sabha-Elections/Quantitative-Analysis-Of-Candidates-In-2019-Lok-Sabha-Elections

Project video Demo Link:

https://youtu.be/wO9UVwx3T-o?si=VIGXo4c6-CrFU202