INT375  
PROJECT REPORT  
(Project Semester January–April 2025)  
   
**Constituency Insights: A Dashboard for Election Results Analysis**   
   
Submitted by:  
   
Name: Divyansh  
Registration No: 12318255  
Programme and Section: B.Tech CSE K23GX  
Course Code: INT375  
   
   
Department of Computer Science and Engineering  
Lovely Professional University  
Phagwara, Punjab (India)  
April 2025

**Under the Guidance of**  
   
**Dr. Gargi Sharma (UID: 29439)**  
  
**Lovely School of Computer Science and Engineering**  
**Lovely Professional University, Phagwara**

**CERTIFICATE**

This is to certify that Divyansh, bearing Registration No. 12318255, has successfully completed the project work entitled “Constituency Insights: A Dashboard for Election Results Analysis” as part of the course INT375 during the project semester January–April 2025, under my supervision and guidance.To the best of my knowledge, the present work is the result of the student’s original research, development, and effort.

Signature and Name of the Supervisor  
**Gargi Sharma (UID: 29439)**

**DECLARATION**

I, Divyansh, a student of B.Tech under the CSE/IT Discipline at Lovely Professional University, Punjab, hereby declare that the project work entitled  
“Constituency Insights: A Dashboard for Election Results Analysis”  
submitted in partial fulfillment of the course INT375, is the result of my own intensive work. The content in this report is original, genuine, and has not been copied from any unauthorized source. All efforts and data analysis have been conducted with sincerity and academic integrity.

Date: 11-04-2025  
   
   
Signature: \_\_\_\_\_\_\_\_\_\_\_\_  
RegNo.: 12318255  
Name of the Student: Divyansh

**ACKNOWLEDGEMENT**  
  
I would like to express my sincere gratitude to all those who supported me throughout the completion of this project titled  
**“Constituency Insights: A Dashboard for Election Results Analysis”**

First and foremost, I would like to thank **Gargi Sharma**, my respected project supervisor, for her invaluable guidance, support, and encouragement. Her expert advice, timely feedback, and motivation helped me stay focused and complete this project efficiently.

I would also like to extend my appreciation to the **faculty and staff of the Discipline of CSE/IT, Lovely Professional University**, for creating a learning environment that inspired this research.

A special thanks to my peers, friends, and family for their constant moral support and encouragement throughout the project.

Lastly, I would like to acknowledge the use of the dataset provided by **https://www.eci.gov.in/general-election-to-loksabha-2024-statistical-reports** which served as the foundation for the analysis conducted in this report.

**Divyansh**  
**Registration No.:** 12318255

**TABLE OF CONTENT**

[1. INTRODUCTION 7](#_Toc195369545)

[2.OVERVIEW AND SOURCE OF DATASET 7](#_Toc195369546)

[3. EXPLORATORY DATA ANALYSIS (EDA) PROCESS 9](#_Toc195369547)

[4. ANALYSIS ON DATASET 9](#_Toc195369548)

[i. Summary of Votes and Voter Turnout 9](#_Toc195369549)

[ii. Party-wise Vote Share Analysis 11](#_Toc195369550)

[iii. Candidate Demographics Insights 12](#_Toc195369551)

[iv. Constituency-wise Winning Margins 14](#_Toc195369552)

[v. Correlation Between Variables 15](#_Toc195369553)

[vi. Category-Wise Representation 16](#_Toc195369554)

[vii. NOTA Impact Analysis 17](#_Toc195369555)

[viii. Total Votes Polled vs Valid Votes for Top 10 Constituencies by Total Electors 18](#_Toc195369556)

[5. CONCLUSION 20](#_Toc195369557)

[6. FUTURE SCOPE 20](#_Toc195369558)

[7. REFERENCES 21](#_Toc195369559)

### 1. INTRODUCTION

Elections are a pillar of democratic practices, representing the voice of the masses. This project seeks to extensively explore constituency-wise election outcomes through contemporary data analysis techniques in Python. Utilizing Exploratory Data Analysis (EDA), this research offers a comprehensive examination of voting patterns, party dominance, demographic impact, and voter sentiment. The combination of visualization capabilities strengthens interpretation, clarifies complicated patterns in the data, and facilitates easy presentation of the results. Not only does this project draw significant insights from unprocessed electoral data but also creates a model for open and intelligent political analysis.

### 2.OVERVIEW AND SOURCE OF DATASET

The dataset named provides a wide and detailed perspective on constituency-level election outcomes. It is a fertile ground for grasping many dimensions of the election process, such as voter turnout, party performance, demographic makeup of candidates, and even the mood of the people as intimated by the **NOTA (None Of The Above)** vote. Every row in the dataset corresponds to a candidate running in a specific constituency, thereby covering the multi-candidate nature of Indian parliamentary elections.

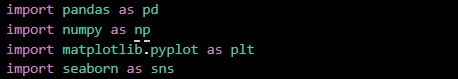
* **Name**: Constituency-wise Detailed Result during 2024
* **Source**: <https://www.data.gov.in/resource/constituency-wise-detailed-result-during-2024>
* **Format**: Comma Separated Values (CSV) file
* **Time Period**: 2024
* **Number of Records**: Over 8,902 entries
* **File Size**: Varies with updates (approx. 1,037 KB)

The data set includes several dimensions of election data, which are discussed below:

* **Geographical and Electoral Scope**: Every record is associated with a particular state and parliamentary constituency, which are represented by state\_name and pc\_name. This enables comparisons at the regional level and dives into the voting patterns at the state level.
* **Candidate Attributes**: It encompasses candidate names, gender, age, and category (e.g., General, SC, ST, OBC). This information allows for analysis of representation in electoral politics—enabling analysis of diversity, inclusivity, and age-related participation patterns.
* **Party Affiliation**: The political party each candidate belongs to is captured in the party\_name column, including prominent national parties, regional parties, and the NOTA option. This enables vote consolidation, fragmentation, and party-wise performance analysis across constituencies.
* **Electorate and Voting Data:** The dataset contains the number of electors in each constituency and a detailed breakdown of votes cast:
  + total\_votes\_polled\_in\_the\_constituency (aggregate of general and postal votes)
  + valid\_votes (excluding invalid ballots)
* **Candidate-wise vote counts:**
  + votes\_secured\_\_\_general (at polling booths)
  + votes\_secured\_\_\_postal (by postal ballots)
  + votes\_secured\_\_\_total (aggregate of general and postal)
* **Performance Metrics**: The dataset contains pre-computed percentages measuring:
  + The proportion of votes secured by a candidate out of the total electors
  + The proportion of votes in comparison to total votes cast
  + The proportion in comparison to total valid votes

These measurements offer a complete means of evaluating candidate performance outside of raw vote totals.

Importing Libraries that are required for my project



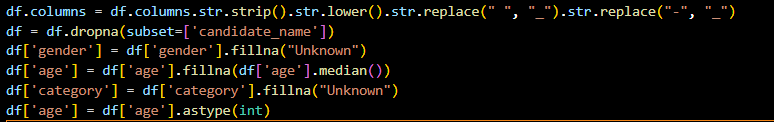
Loading the dataset which is in CSV format



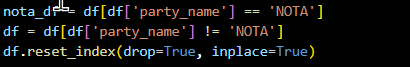
### 3. EXPLORATORY DATA ANALYSIS (EDA) PROCESS

The EDA (Exploratory Data Analysis) phase includes the following key steps:

1. **CLEANING:** Standardized column names by making them lowercase and replacing spaces and dashes with underscores.
2. **HANDLING MISSING VALUES:**
   * Removed rows with missing candidate names
   * Filled missing gender and category with "Unknown"
   * Replaced missing ages with median age
3. **TYPE CONVERSION:** Converted age to integer and ensured all numeric columns were of numeric type.



1. **FILTERING:** Isolated NOTA entries for later analysis and excluded them from the main dataset.

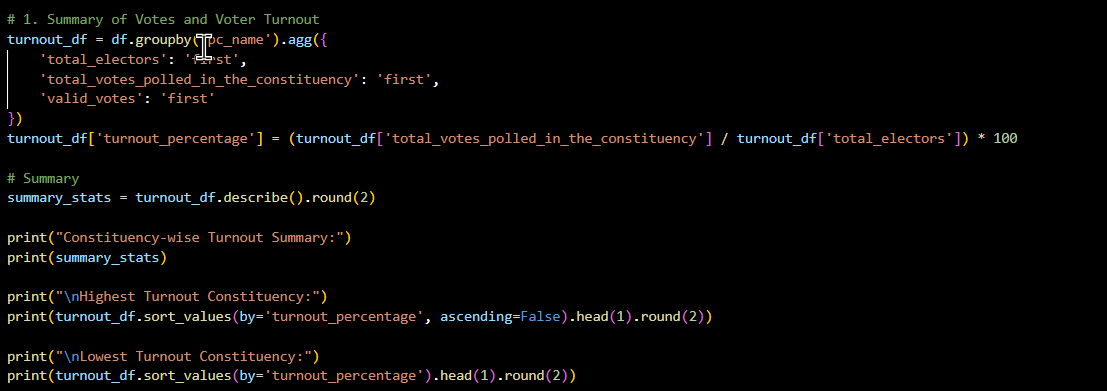


### 4. ANALYSIS ON DATASET

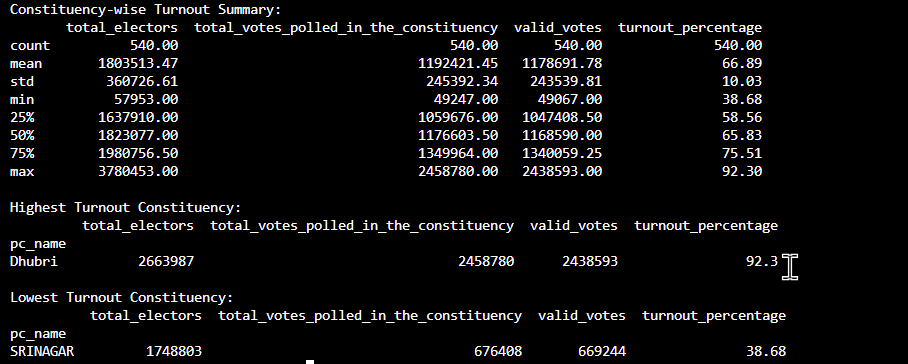
#### i. Summary of Votes and Voter Turnout

* Introduction: Understand how many citizens participated in the voting process.
* General Description: Grouped data by each constituency (pc\_name) and aggregated total electors, total votes polled, and valid votes.
* Functions and Formulas:
  + groupby('pc\_name')
  + agg({}) to extract the first available value
  + Turnout percentage = (votes polled / total electors) \* 100
* Results:
  + Summary statistics such as mean, min, max of turnout percentage
  + Constituencies with highest and lowest turnout identified
* Visualization:
  + Text-based display of summary stats
  + Printed details of constituencies with extreme turnout values

**Code Used:**



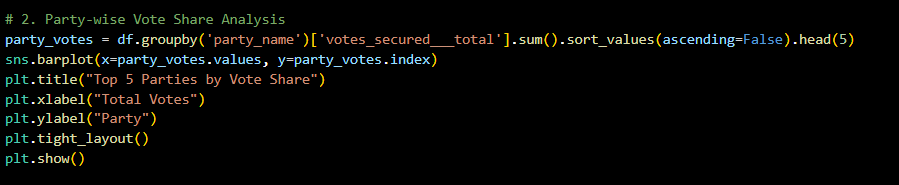
**Output:**



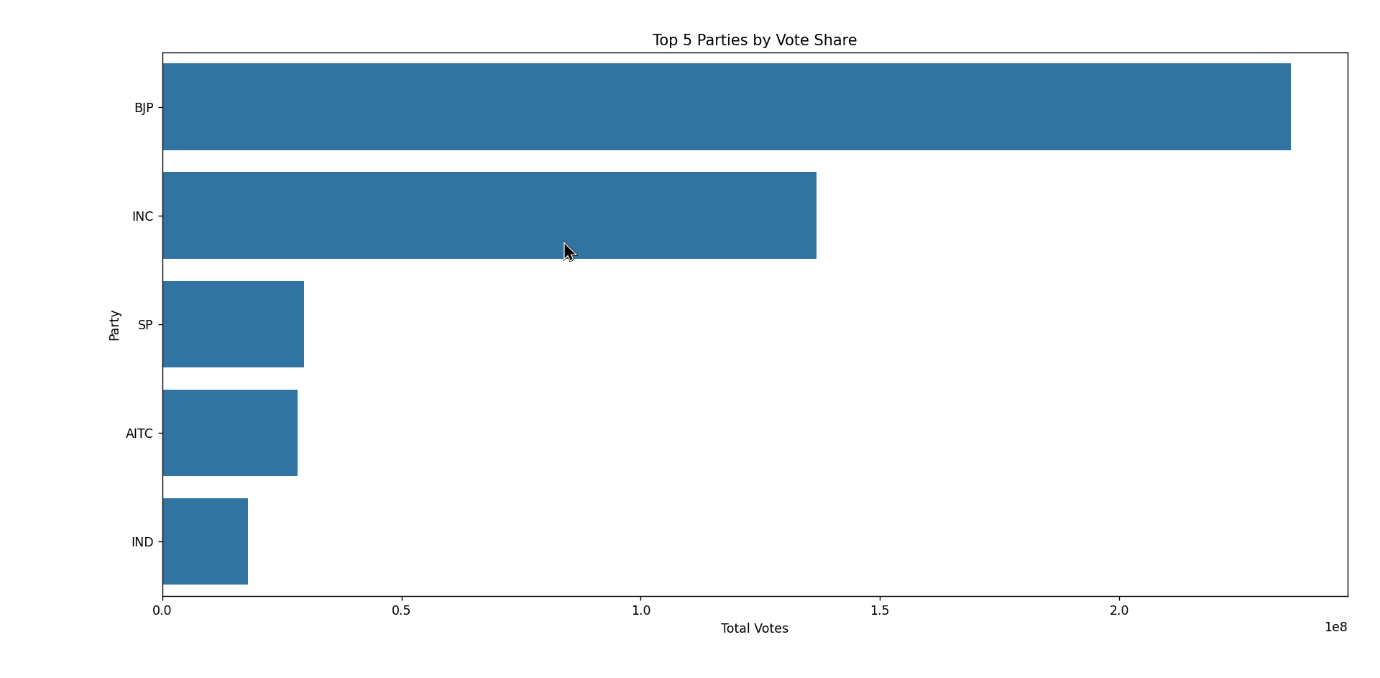
#### ii. Party-wise Vote Share Analysis

* Introduction: Discover which parties secured the most votes.
* General Description: Grouped data by party\_name and summed the total votes received.
* Functions and Formulas:
  + groupby('party\_name')['votes\_secured\_\_\_total'].sum()
  + sort\_values(ascending=False)
* Results:
  + Top 5 political parties based on votes secured
* Visualization:
  + Horizontal bar plot using Seaborn showing top 5 parties

**Code Used:**

****

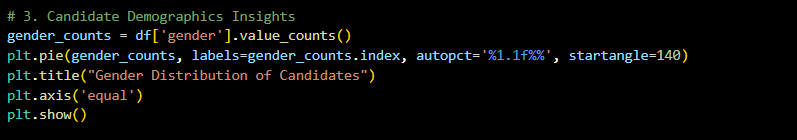
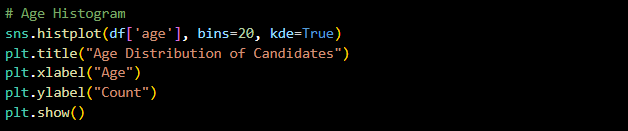
**Output:**

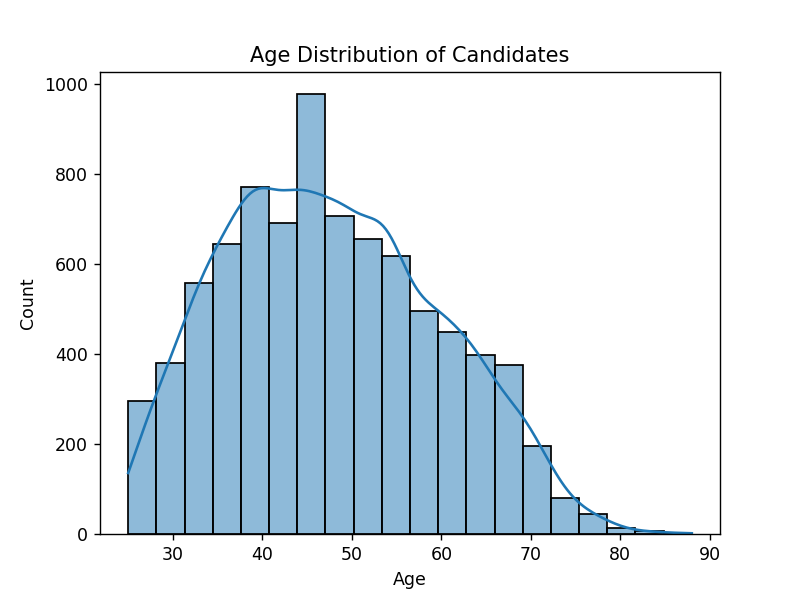
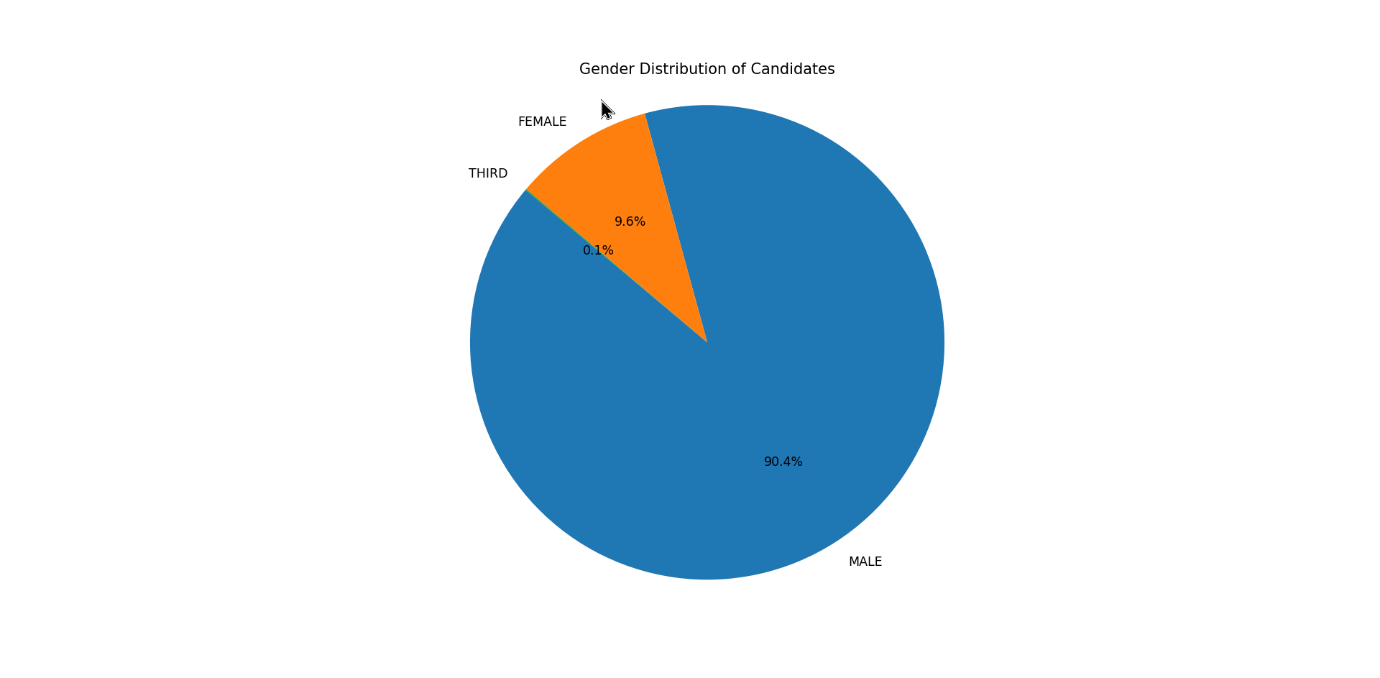
****

#### iii. Candidate Demographics Insights

* Introduction: Gain insights into who the candidates are in terms of gender and age.
* General Description: Analyzed the distribution of gender and age across all candidates.
* Functions and Formulas:
  + Gender: value\_counts()
  + Age: histplot() with KDE
* Results:
  + Majority of candidates were male
  + Age distribution followed a bell-shaped curve with candidates mainly between 35 and 60 years
* Visualization:
  + Pie chart for gender distribution
  + Histogram with KDE for age

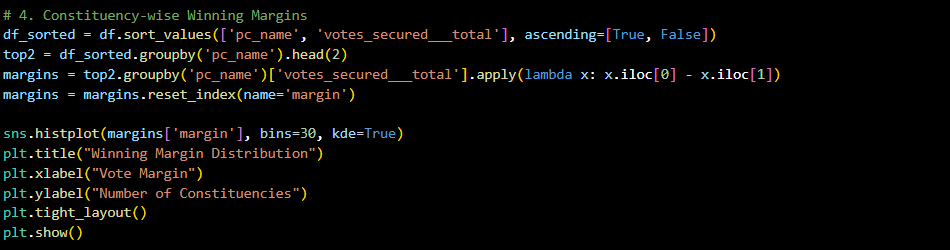
**Code Used:**

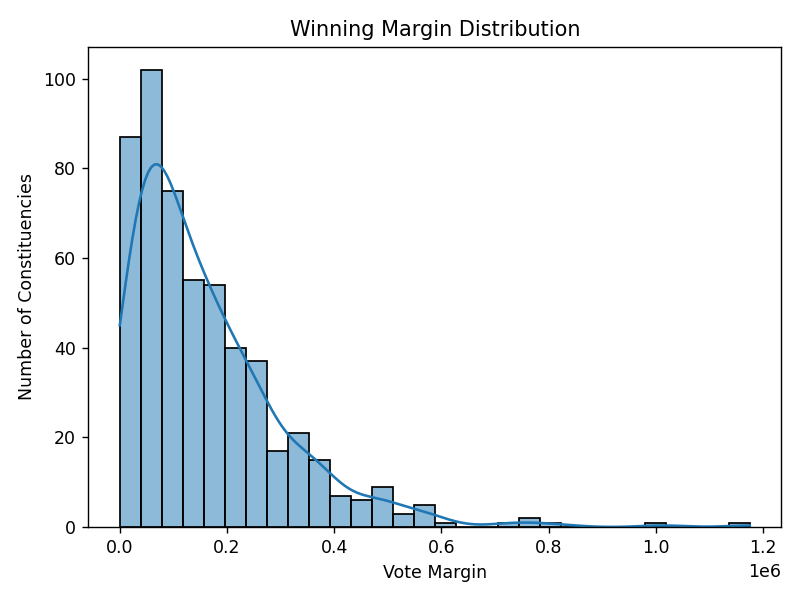
**  
Output:**

****

#### iv. Constituency-wise Winning Margins

* Introduction: Measure how closely contested each election was.
* General Description: Sorted candidates by votes in each constituency and took the difference between the top two.
* Functions and Formulas:
  + sort\_values(['pc\_name', 'votes\_secured\_\_\_total'])
  + groupby().head(2)
  + Winning Margin = First - Second
* Results:
  + Winning margins varied significantly across constituencies
  + Some constituencies had very close races (margins < 100)
* Visualization:
  + Histogram of margins using Seaborn

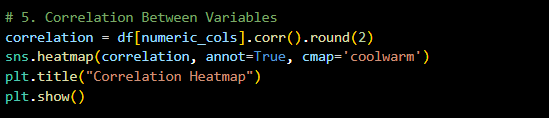
**Code Used: **

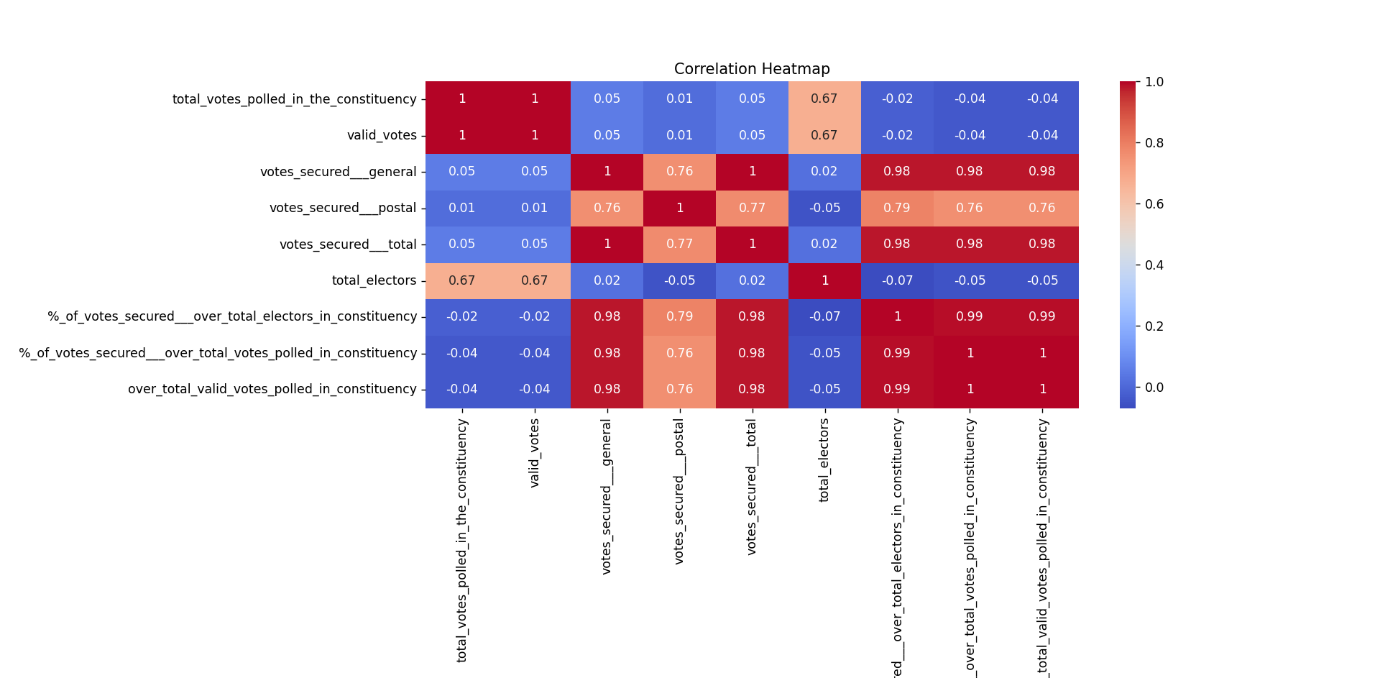
**Output: **

#### v. Correlation Between Variables

* Introduction: Understand how variables such as electors and votes are related.
* General Description: Used a correlation matrix to measure linear relationships.
* Functions and Formulas:
  + df.corr() for Pearson correlation
  + seaborn.heatmap()
* Results:
  + High correlation between total electors and votes polled
  + Other correlations provided additional insights
* Visualization:
  + Annotated heatmap using 'coolwarm' palette

**Code Used:**

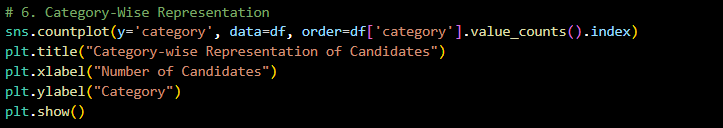
****

**Output: **

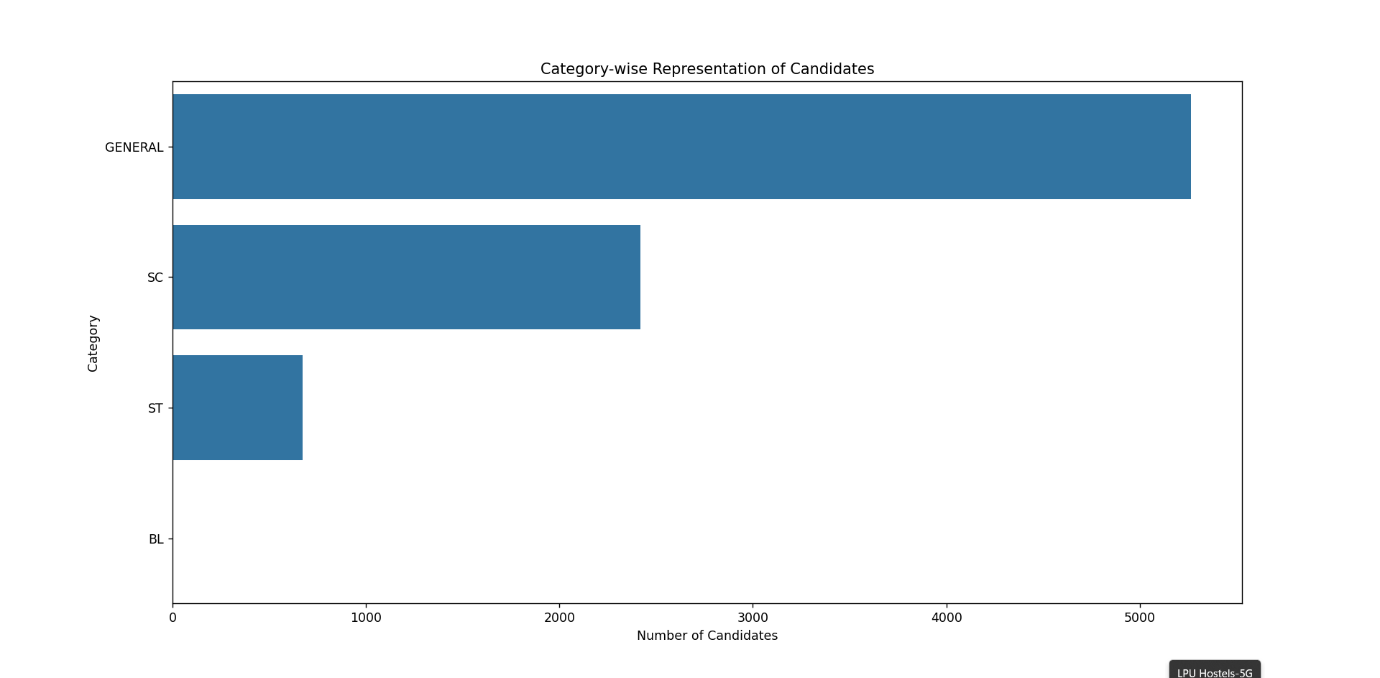
#### vi. Category-Wise Representation

* Introduction: Examine the representation of different social categories.
* General Description: Counted the number of candidates in each category.
* Functions and Formulas:
  + value\_counts() on category column
  + seaborn.countplot()
* Results:
  + General category had the highest number of candidates
  + SC/ST/OBC and Unknown were also present in notable numbers
* Visualization:
  + Bar chart showing number of candidates per category

**Code Used:**

****

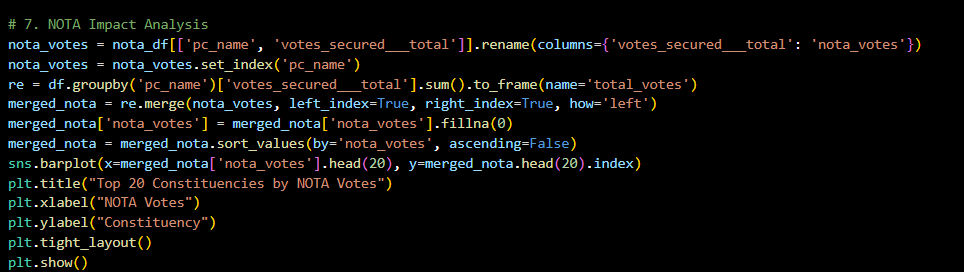
**Output:**

****

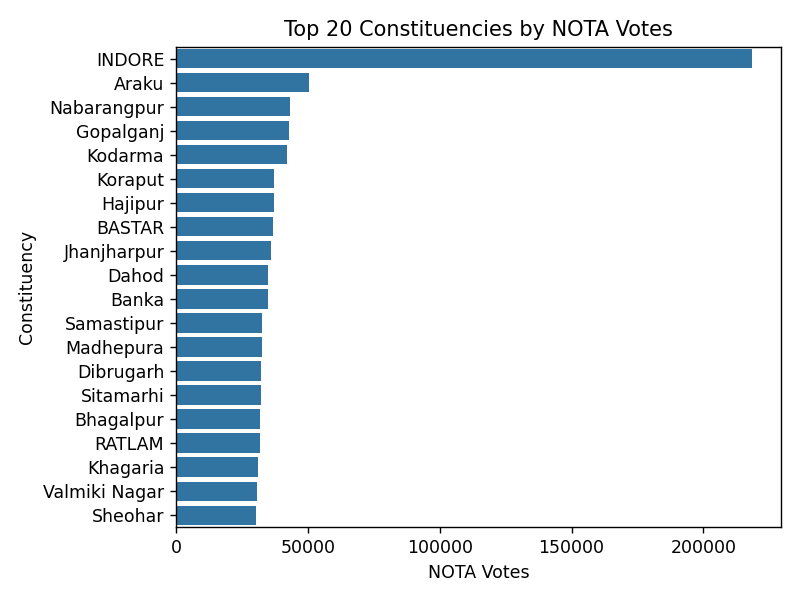
#### vii. NOTA Impact Analysis

* Introduction: Investigate constituencies with high NOTA votes.
* General Description: Isolated NOTA entries, merged them with constituency-level data.
* Functions and Formulas:
  + df[df['party\_name'] == 'NOTA']
  + Merging dataframes using merge()
* Results:
  + Some constituencies had thousands of NOTA votes
  + Indicates public dissatisfaction in certain areas
* Visualization:
  + Bar plot of top 20 constituencies by NOTA votes

**Code Used:**

****

**Output:**

****

#### viii. Total Votes Polled vs Valid Votes for Top 10 Constituencies by Total Electors

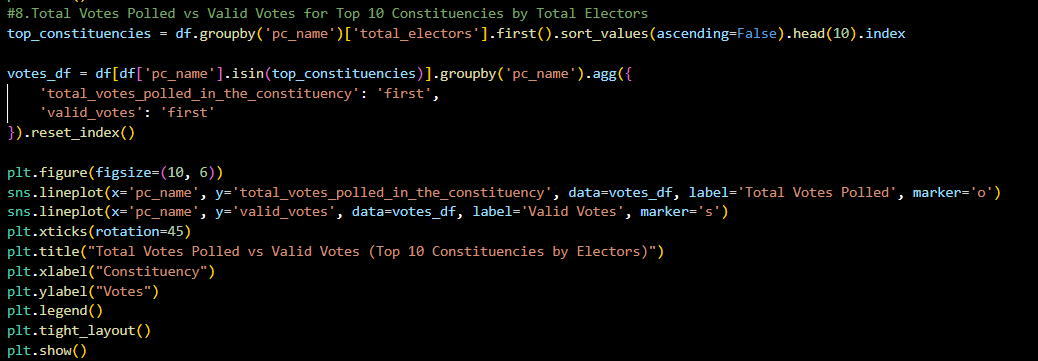
* Introduction

This analysis compares total votes polled and valid votes in the top 10 constituencies with the highest number of electors, helping identify voting effectiveness and ballot rejection rates.

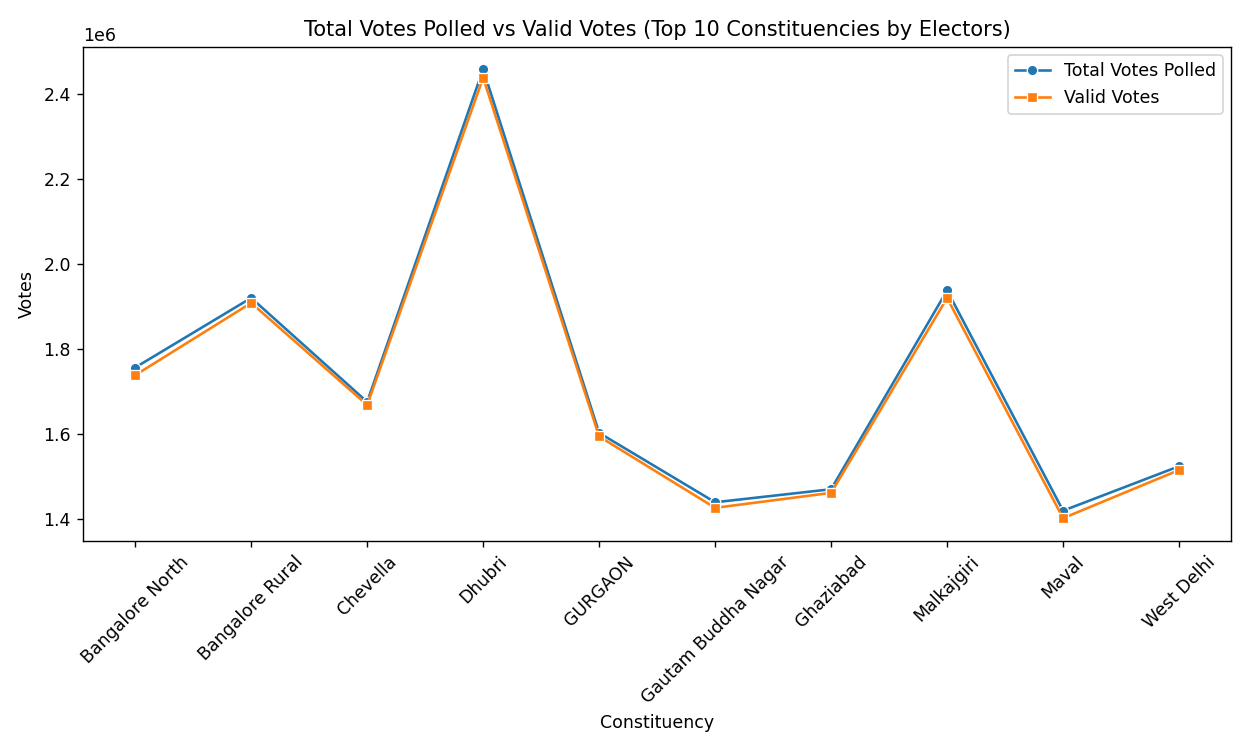
* Description & Tools
  + Data grouped by constituency (pc\_name)
  + Aggregated electors, votes polled, and valid vote
  + Selected top 10 based on total\_electors
  + Used a line chart to visualize total vs valid votes
* Visualization
  + X-axis: Top 10 constituencies by electors
  + Y-axis: Number of votes
  + Two lines: one for Total Votes Polled, one for Valid Votes

This chart helps assess voting efficiency and highlights areas with higher vote rejection.

**Code Used:**

****

**Output:**

****

### 5. CONCLUSION

This analysis provided deep insights into voting trends, party performance, candidate demographics, and voter behavior. High turnout rates reflected active civic engagement, while high NOTA votes pointed to political dissatisfaction. The study successfully demonstrated the power of EDA techniques in deriving meaningful insights from raw electoral data.

### 6. FUTURE SCOPE

**1. Expand analysis to incorporate data from several elections for trend analysis**

By examining data from several election years (e.g., 2014, 2019, 2024), we can determine long-term trends in voter turnout, party performance, and candidate demographics. This assists in understanding how public opinion and electoral behavior have changed over time, and whether particular regions tend to support particular parties or exhibit swings.

**2. Incorporate geographic plots using GeoPandas or Folium**

With geographic libraries such as GeoPandas or Folium, we can visualize election outcomes geographically across India. For instance, showing vote shares or turnout rates on a map aids in identifying regional trends, political stronghold hotspots, or low-turnout areas. It provides spatial context to the data and renders the results more intuitive.

**3. Utilize predictive models to predict probabilities of winning**

With enough historical information, machine learning algorithms (such as logistic regression, decision trees, or ensemble approaches) could be trained to forecast election outcomes. They could incorporate demographics, previous outcomes, voting turnout, and party representation when projecting future winners or margins to benefit political strategists and campaign planners.

**4. Add sentiment analysis with NLP on public conversation and media**

Natural Language Processing (NLP) methods can be applied to social media posts, news stories, and speeches to measure public opinion towards candidates or parties. This enables researchers to compare sentiment with voting patterns, which can help reveal the power of media narratives or online discussion in determining election results.

### 7. REFERENCES

* Election Commission of India: <https://eci.gov.in/>
* Pandas Documentation: <https://pandas.pydata.org/>
* Seaborn Documentation: <https://seaborn.pydata.org/>
* Matplotlib Documentation: <https://matplotlib.org/>
* Kaggle (for related election datasets):<https://www.kaggle.com/>