

Introduction to Data Management Project report

# PROJECT REPORT ON

Indian Candidates for General Election 2019

Dataset of candidates who contested the 2019

Lok Sabha Elections

Submitted by :

Bellapu Narendra

11902977

Section: KM037

Course Code: INT217

Under the Guidance of : Komal Arora ( 17783 )

Lovely School of Computer Science & Engineering

Lovely Professional University, Phagwara

# **DECLARATION :**

I, Bellapu Narendra , student of Computer Science & Engineering under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 14/12/2021

BELLAPU NARENDRA

Registration No: 11902977

**INTRODUCTION :**

Microsoft Excel is a computer application program **written by Microsoft**. It mainly comprises tabs, groups of commands, and worksheets. It stores the data in tabular form and allows the users to perform manipulation operations on them.

Microsoft Excel is one of the most suitable spreadsheet programs that help us to store and represent the data in tabular form, manage and manipulate data, create optically logical charts, and more. Excel provides you the worksheet to create a new document in it. You can save the Excel file with **.xls extension** .

### Indian Candidates for General Election 2019

### Context:

With over 600 Million voters voting for 8500+ candidates across

543 constituencies, the general elections in the world's largest democracy are a potential goldmine of data. While there are existing separate datasets about the votes each candidate received and the personal information of each candidate, there was no comprehensive dataset that included both these information. Thus, this dataset will provide more usability than most existing datasets in this domain.

### Content:

I scraped the website of myneta.info to get the personal information of each candidate (as per their own sworn affidavits) and the website of Election Commission of India to get the data about the votes received. I merged both this datasets to create this comprehensive dataset. Only the candidates who secured at least 1% of the total votes polled in their constituency have been included.

### Acknowledgements:

I have collected the data from MyNeta.info maintained by the Association for Democratic Reforms and the website of Election Commission of India.

### Inspiration:

There are 2 main tasks that can be performed on this dataset: Exploratory Data Analytics to visualize the impact of each feature of the candidate and the use of machine learning to predict the chances of winning of a candidate

**OBJECTIVES :**

After analysing the data given , I have made some objectives

and the objectives are :

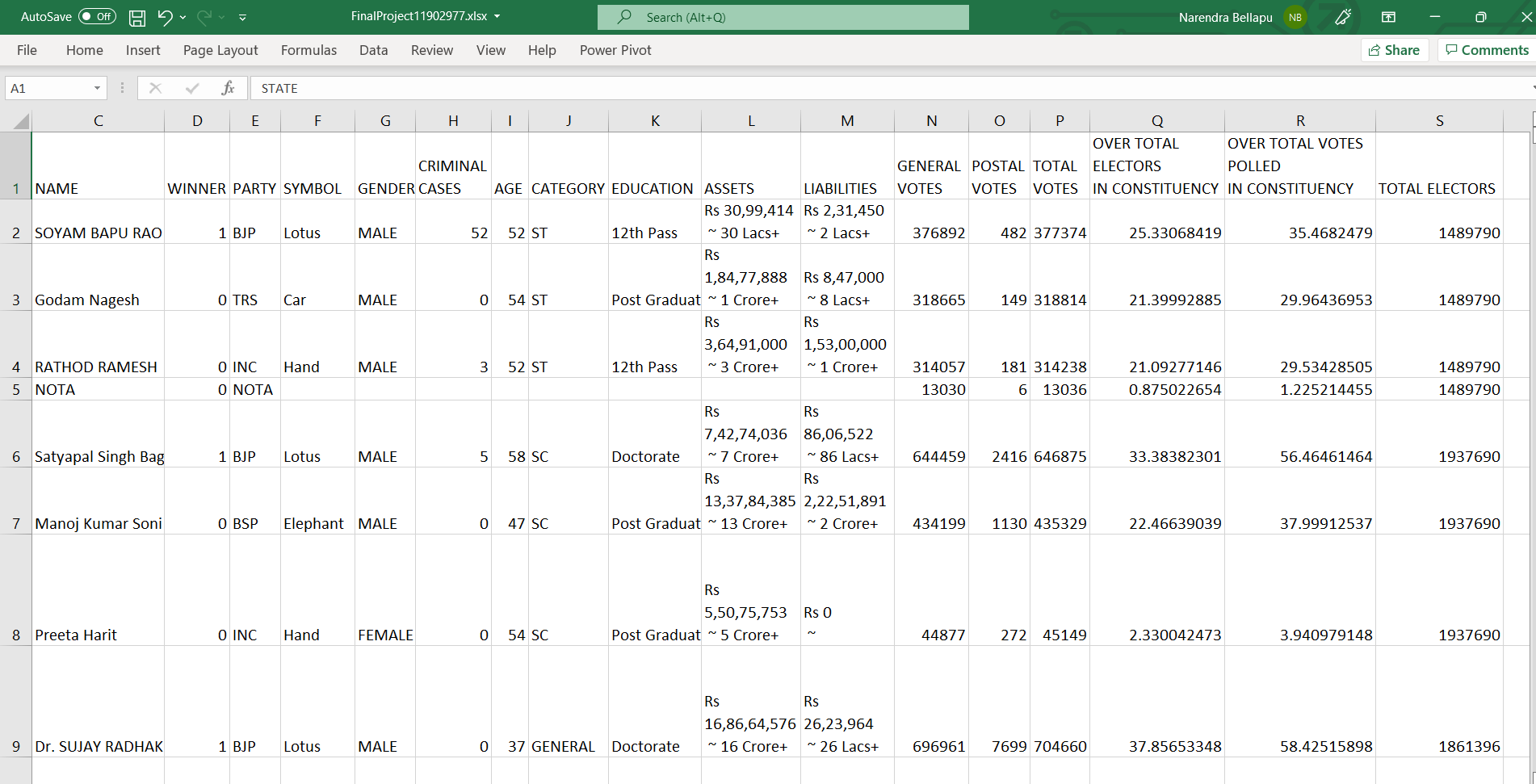
* No.of postal votes for all winners in each state
* Andhra Pradesh politicians who have criminal cases
* Count of politicians in each Education level
* State wise total no.of votes casted in elections
* Category wise analysis of total electors and total votes polled in constituency
* Count of SC politicians in each state

**DATA SET :**

This below data set is about Indian Candidates for General Election 2019 and their columns are :

1. State
2. Constituency
3. Name
4. Winner
5. Party
6. Symbol
7. Gender
8. Criminal Cases
9. Age
10. Category
11. Education
12. Assets
13. Liabilities
14. General Votes
15. Postal Votes
16. Total Votes
17. Over Total Electors in Constituency
18. Over Total Votes Polled in Constituency
19. Total Electors

Sample view of data set



**ANALYSIS ON DATASET:**

**1.No.of postal votes for all winners in each state :**

* After performing this analysis using pivot table we will get the count of postal votes in each state.
* The analysis is based on state,winner and postal votes.

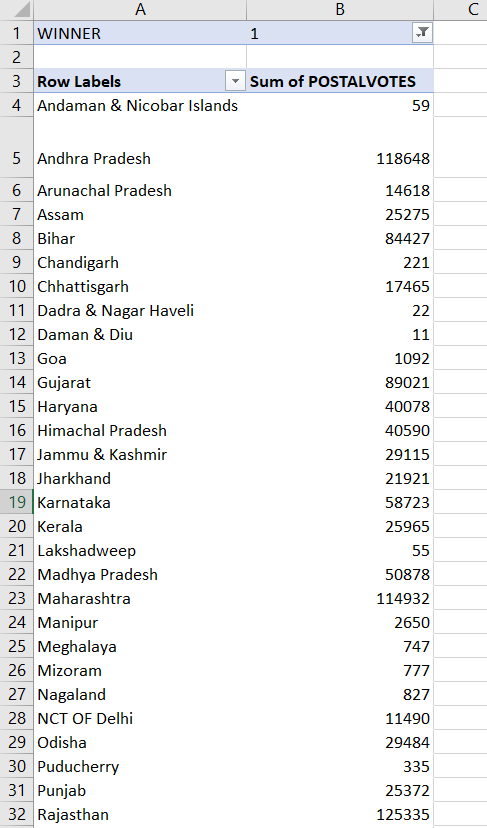
TOOLS USED : 1. Pivot table

2. Pivot chart

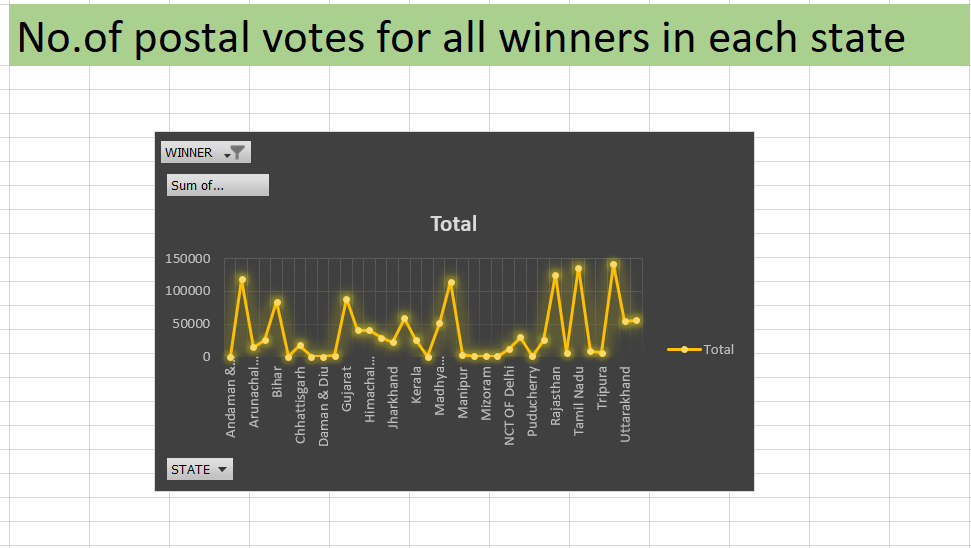
Use pivot table to get count and select required columns

and rows accordingly.

**After analyzing :**



**Visualization:**

****

**2. Andhra Pradesh politicians who have criminal cases:**

* After performing this analysis using pivot table we will get the list of politicians in AP who have criminal cases.
* The analysis based on state,name and criminal cases.

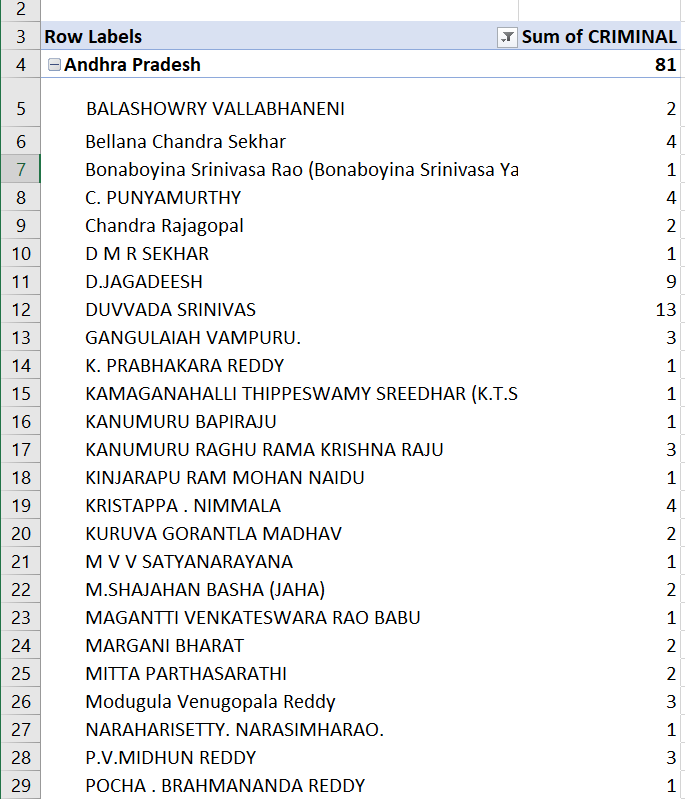
TOOLS USED : 1. Pivot table

2. Pivot chart

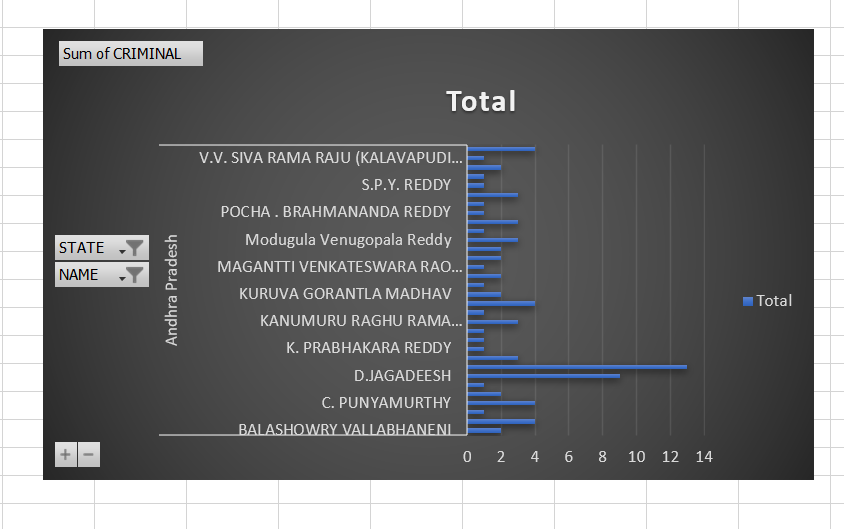
Use pivot table to get count and select required columns

and rows accordingly.

After analyzing :



**Visualization:**



**3. Count of politicians in each education level:**

* After performing this analysis using pivot table we will get the count of candidate names according to education wise.
* The analysis is based on education and name.

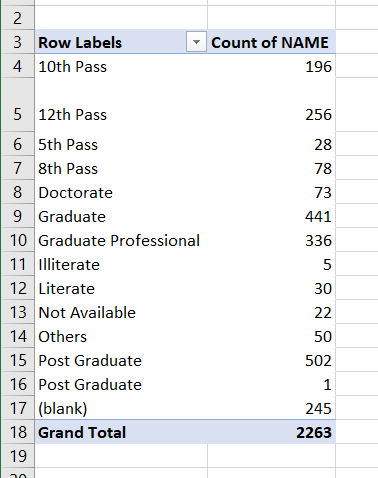
TOOLS USED : 1. Pivot table

2. Pivot chart

Use pivot table to get count and select required columns

and rows accordingly.

After analyzing :



**Visualization:**



**4.State wise total no.of votes casted in elections:**

* After performing this analysis using pivot table we will get the total no.of votes in a state wise.
* The analysis is based on state and total votes.

TOOLS USED : 1. Pivot table

2. Pivot chart

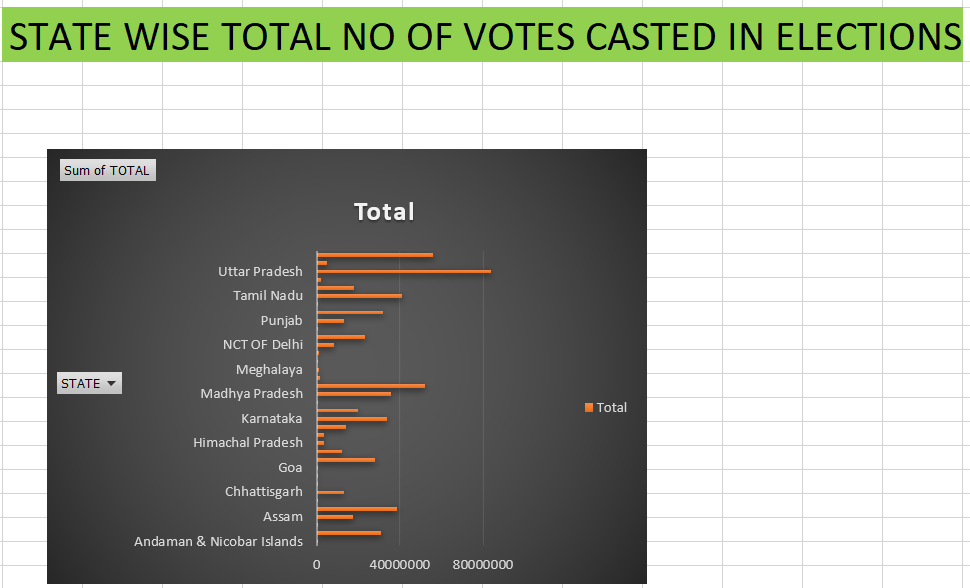
Use pivot table to get count and select required columns

and rows accordingly.

**After analyzing :**



**Visualization:**



**5. Category wise analysis of total electors and total votes polled in constituency:**

* After performing this analysis using pivot table we will get the count of total electors and total votes polled in a constituency.
* The analysis is based on category,over total electors in constituency and total votes polled in constituency.

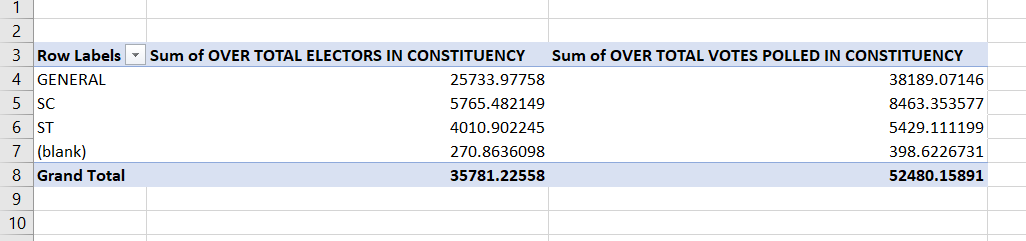
TOOLS USED : 1. Pivot table

2. Pivot chart

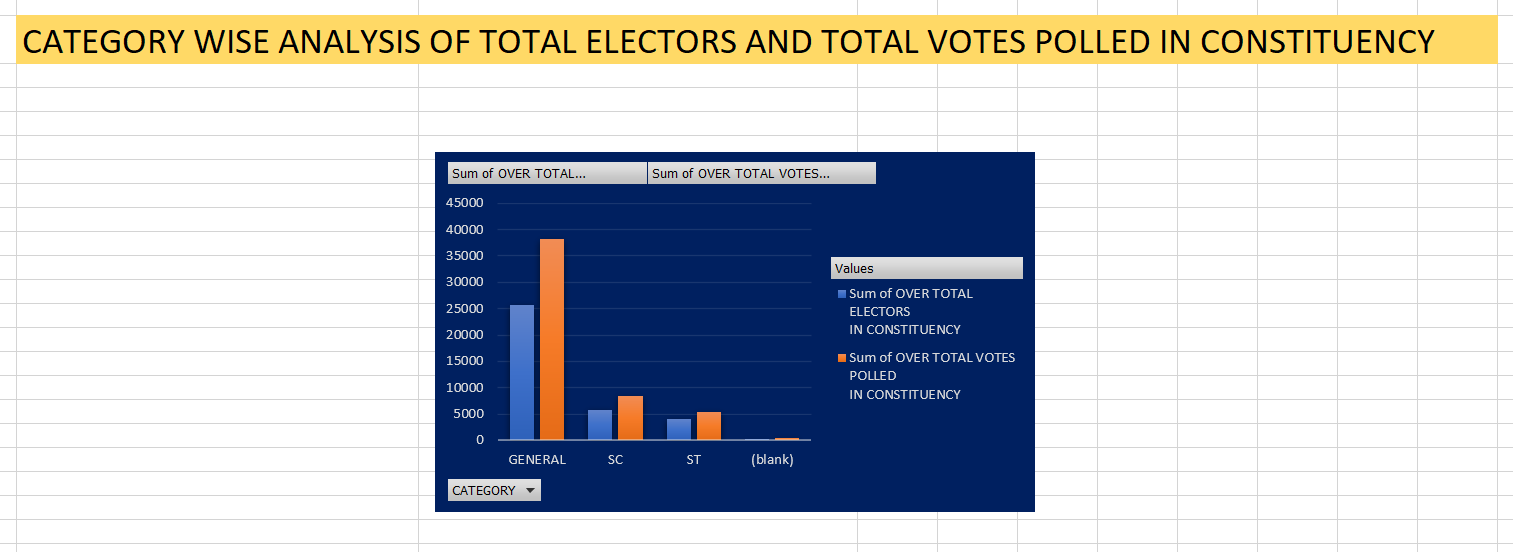
Use pivot table to get count and select required columns

and rows accordingly.

**After analyzing :**



**Visualization:**



**6.Count of SC politicians in each state:**

* After performing this analysis using pivot table we will get the count of SC politicians in each state.
* The analysis is based on category,over total electors in constituency and total votes polled in constituency.

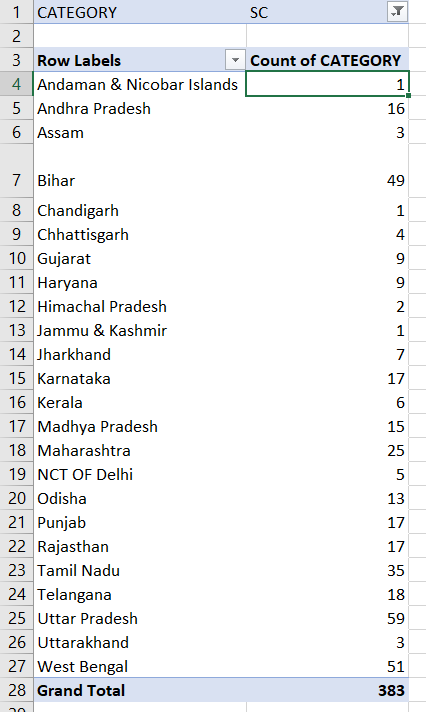
TOOLS USED : 1. Pivot table

2. Pivot chart

Use pivot table to get count and select required columns

and rows accordingly.

**After analyzing :**

****

