# USING PAIRWISE RELATIONSHIP PLOTTING

Affiliation

*Mrs. K. MAHALAKSHMI 1, V NANDHINI PRIYA 2,*

[1] Mrs. K. MAHALAKSHMI,

[1] Assistant professor, [1]Department of B. Com Business Analytics,

[1]PSGR krishnammal College for Women, Coimbatore, India.

muhilm9@gmail.com

[2] Ms. V NANDHINI PRIYA,

[2]UG SCHOLAR, [2]Department of B. Com (Business Analytics),

[2]PSGR krishnammal College for Women, Coimbatore, India.

nandhinivijaykumar010@gmail.com

***Abstract :*** In this current developing society the needs of people in india is increasing and the people below poverty level are also increasing. Most of the family's are dependent on government schemes so to analyze and find the useful strategies which can help the people and also the parties to win the election we used the past election data. The past election results are needed for analyzing the future needs. Visualizing final results and the relationship between each attribute is important for better understanding of the result. The pairwise relationship plotting is used in any projects to show the pairwise relationship between variables in a dataset. Here, pairwise relationship plot is used to show the data that are related in graphical way with the combination of scatterplot.

***Index Terms*** **- Pairwise relationship plot, Scatter plot, Past election data.**

## I. INTRODUCTION

A pairs plot may be a matrix of scatterplots that allows you to understand the pairwise relationship between different variables during a dataset. Once you've prepared a a nice cleaned dataset, the next step is Exploratory Data Analysis (EDA). EDA is that the process of deciding what the info can tell us and that we use EDA to seek out patterns, relationships, or anomalies to tell our subsequent analysis. While there are an almost overwhelming number of methods to use in EDA, one among the foremost effective starting tools is that the pairs plot (also called a scatterplot matrix).

A Pair plots allows us to ascertain both distribution of single variables and relationships between two variables. Pair plots are an excellent method to spot trends for follow-up analysis and, fortunately, are easily implemented in Python. In this we analyses the election voting in response to the growing voting system needs faced by the Country and in case of future regulatory changes and pending legal requirements our current systems are unable to meet. For example, INDIAN NATIONAL ELECTION data which have the past election results ,the party participated in the election and the other details can be used for analyzing and predicting useful strategies for current election system.

The constitution of India has entrusted within the election commission of India the superintendent, direction and control of the whole process for conduct of elections to Parliament and Legislature of every State and to the offices of President and Vice-President of India. Elections are held on the basis of Universal adult franchise. Who may be a citizen of India and not but 18 years aged can register as a voter in electoral roll of India. There is no discrimination on the ground of religion, race, caste, sex or any of them

## II. RELATED WORKS

The Indian Electoral system has been broadly divided into two, they are Direct election supported territorial constituencies and representation by means of one transferable vote.The"First system is followed for holding the election of the members of lok Sabha", state assemblies and union territories assemblies.

The second, election persisted the idea of representation by means of one transferable vote for the President and therefore the Vice-President of India, members of Rajya Sabha and members of Legislative councils. In a country where elections are won on promises of basic entitlements, ensuring that these benefits from wages to pensions reach the targeted beneficiaries has remained a challenge for wealth distributive policies. Add to that the lack of a unified identity tool, and the benefits to residents and governments alike multiply.

Looking at data, We have the finding that there has been no association between growth and electoral performance at the state level over the past several decades. One way to find is to examine the relationship between growth and electoral outcomes by decade, to explore whether it has been changing over time.

Only after disaggregating the relationship between economics and elections by decade does it become clear that a big change has occurred within the period following the year 2000. Unlike in previous decades, the info reveal positive and significant electoral returns to governments ready to generate faster growth rates within the 2000s. Statistical analyses, after controlling for a variety of probably confounding factors, reveal that a 1 per-centage point improvement during a states rate of growth within the 2000s is linked to a 7.5 percent increase in the likelihood that the incumbent party or alliance will be reelected, a 3.3 percentage point gain in seat share, and a positive rise in vote share.29

To reiterate, there is no clear pat-tern of electoral rewards for higher growth in either the 1980s or 1990s; these effects are only evident in the most recent decade.

A recent record called analysis of funds collected and spending done by political parties during elections in 2004 published by association of democratic reforms(ADR) mentions that collectively, for the lok Sabha election held in 2004, 2009 and 2014, political parties disclosed total collection of rs.2355.35 crores Their total expenditure on these elections was Rs. 2466.07 crores with as much as rs.1587.77 crores reportedly spent for lok Sabha 2014 elections alone.

Unofficial estimates by some news agencies 22 indicate that the whole expenditure by candidates and political parties for lok Sabha 2014 elections could possibly be quite rs.30,000 crores. However it's quiet difficult to corroborate the reliability of such unofficial reports. For State Assembly elections over the amount 2004 2015, the ADR report further mentions that political parties disclosed a set of Rs. 3368.06 crores while the entire expenditure as reported for that period was Rs. 2727.79 crores.

To find the strategies that are useful for current election situations using the past election records we can use jupyter notebook which uses Python 3 as back-end tool. Lets make a visualization of final results with pair plot. For doing pairwise relationship plotting we need required library packages to access the features. Seaborn may be a Python data visualization library supported matplotlib. It provides a high level interface for creating attractive and informative statistical graphs. Its plotting functions operate data frames and arrays containing whole datasets and internally perform the required semantic *mapping* and statistical aggregation to produce informative plots. First import the matplotlib & seaborn library.

Tamilnadu assembly elections 2016 shows its extent in democracy by having 42,908,767 as total number of voters, 3727 candidates, 234 constituencies and 65, 616 polling stations. The tamilnadu assembly election 2019 may be a real head-scratcher, with group-specific issues and seat-specific caste and economic concerns.

Then import the INDIAN NATIONAL ELECTION dataset to access the data. The dataset must be preprocessed because the null values are not taken for analysis and cannot be displayed in the visualization. By calling seaborn with a short-hand name sns before mentioning the type of plot will allow us to access the package and make the graph.

In this project the 5 main attributes of the election details are taken to analyze whether the strategies are useful are not. The result of the election with total number of votes they occupied depends on the strategies the parties choose to attract the voters. The name of the parties participating in the election and the number of electors each party have in each state is an very important factor in this. The parties in India have numerous number of electors in different regions in India and each party uses different kind of strategies to help people and secure their votes for the election.

Not every strategy is effective in catching peoples attention. In this dataset we can see the parties using particular strategies which lead them to win the election. Also there is a strategies which failed to work in that economic situation of people to help the electors or the parties to win the election.

In this paper we are only going to Visualize the above looked attributes instead of visualizing the strategies, because the above mentioned attributes are the result of using the strategies. So lets look into the results to define the effectiveness of strategies and which strategies can help in the current economic situation of people in India.

## III. RESEARCH METHODOLOGY

Data visualization could even be a graphical representation of your data, aggregations, and model results summed up in visual elements accessible to the end-user. These visualizations allow users to quickly see trends within the data, outliers, and patterns that exist.

We are going to do this data visualizations using Python programming language. Python offers multiple great graphing libraries that come packed with lots of different features. No matter if you would like to make interactive, live or highly customized plots python has a superb library for you.

Seaborn offers various features like inbuilt themes, color palettes, functions and tools to see univariate, bivariate, rectilinear regression , matrices of knowledge , statistical statistic etc. which lets us to build complex visualizations. We can use the DataComPy library is a package to compare two Pandas Data Frames and provide a human-readable output describing the differences.

The pairwise relationship plotting shows different kinds of variables and their relationships in one graph. It is a visualization of more than two data in a single graph. This plotting feature creates a grid of Axes such each in data will by shared within the y-axis across one row and within the x-axis across one column. So that the multiple data are displayed under one graph. Since, it is a combination of scatterplot it plots all the relationship between numerical variables in pairwise scatterplot.

If the variables tend to extend and reduce together, the association is positive. If one variable tends to extend because the other decreases, the association is negative. If there's no pattern, the association is zero. The type of relationship determines the result of statistical measures and tests that are appropriate. If the variables are related, then the points will fall along a line. The better the correlation, the tighter the points will display the road. In this way the scatterplot shows the relationship between paired data.

After doing the steps on performing preprocessing and using correlation, linear regression and polynomial regression for analyzing , defining and sorting the winning and losing strategies, we got the above mentioned attributes as results. Visualizing the final results are not enough to confirm the accuracy of the result. The need for checking the accuracy level is high when the results are going to be suggested for real-time implementation. So, the accuracy level of three algorithm are also added below for reference.

## IV. RESULTS AND DISCUSSION

### 1.1 ACCURACY TABLE

| | |
|---|---|
| Accuracy value 1 | 1.0 |
| Accuracy value 2 | 0.95 |
| Accuracy value 3 | 0.9499 |
| Accuracy Score | 1.0 |

Accuracy Values of linear and polynomial regression for knowing the accuracy level of analysis.

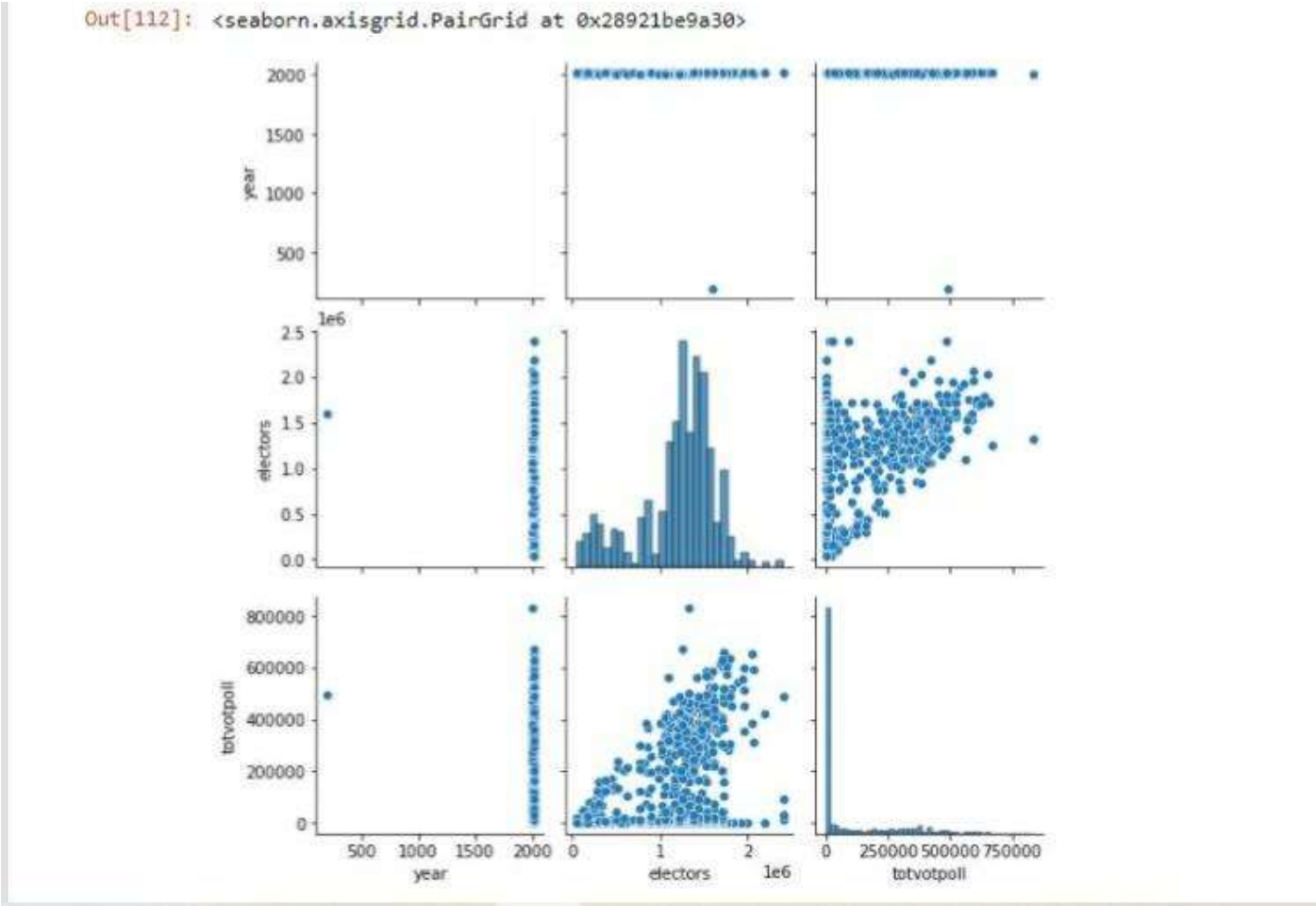


*Fig 1.2 Pairwise relationship of election results*

In the above figure 1.2 the relationship between the 5 main attributes like Year, state name, number of electors, total votes secured by the electors and the election results are visualized. The year of election 2009 and 2014 are taken for analysis. In the two different years the election results were differed. But both years data are related and only the values are different. But the fact the strategies used in 2009 election is also used with different co-strategies in 2014 election. Comparing those results some of the strategies are very useful and the people need this kind of strategy to be satisfied with the electors and the election.

## CONCLUSION

The current economic situation of people in India is growing but not to the level where people below poverty level can get a proper shelter. They are still struggling for food and education. People expecting government to provide more free schemes like free education and job opportunities. Whoever accomplishes or promises to accomplish the need they are expecting will win the election. In this paper, we have seen about visualizing more than two different variables using the pair plot feature with the help of Seaborn library. The year, state name, party participated in the election, number of electors, and the votes secured by each parties using the strategies which satisfied and unsatisfied the need of people are visualized to give a better view about the analysis. The relationship between the year, electors, total votes they secured are displayed in a understandable manner.

### REFERENCE

**[1]** Published by Zach https://www.statology.org/study-register/.

**[2]** Michael Waskom. Created using Sphinx 3.3.1.

**[3]** ® Copyright Analyze-it Software, Ltd. 2021.

**[4]** © 2021 American Society for Quality. All rights reserved.

**[5]** Gautam Chikermane Foreword by Rakesh Mohan© 2018 by Observer Research Foundation

**[6]** Milan Vaishnav., © 2015 Carnegie Endowment for International Peace. All rights reserved.

**[7]** Bibek Debroy1 and Kishore Desai2

**[7]** Dr. M. Bhaskara Raju M.A., Ph.D., Lecturer, Dept. of Political Science, Govt. Degree & P.G.College, Puttur, Chittoor Dist, Andhra Pradesh-517583, India

**[8]** Gilbert Tanner-towardsdatascience

**[9]** Roseday-towardsdatascience