SYRACUSE
UNIVERSITY

CIS 500 Data Analysis and Decision Testing
Project Proposal
USA Elections Result Analysis

Team Members:

Ayush Saxena asaxen09@syr.edu
Brijesh Kumar Verma bverma@syr.edu
Hemil Anip Shah hashah15@syr.edu

Introduction:
This project aims to analyze 1970-2020 years of USA elections data to uncover insights into past election trends and study what factors are most responsible for General Election Results in USA and later develop a methodology for predicting the outcome of future presidential elections.

Objectives:
To analyze data from the USA general elections spanning from 1976 to 2020 to identify trends, patterns, and influential factors affecting election outcomes, aiming to provide actionable insights into the dynamics of electoral processes and the key determinants shaping election results over time.
This might help the political parties to refine their strategies based on past election data Insights and studies.

Data Collection:
We are using a comprehensive dataset of USA elections spanning past 50 years, including information such as party affiliations, voting percentages, demographic data, etc.
We will ensure data integrity and cleanliness by performing data cleaning and preprocessing steps.
Data will be collected from different open-source resources like MIT lab, Kaggle, US government data.

Data Analysis:

We are dividing our project into 3 major subgroups:

Descriptive Analysis and Exploratory Data Analysis(EDA) -

- Summarize the dataset using descriptive statistics (mean, median, mode, range, standard deviation, etc.).
- Visualize data distributions and trends using histograms, box plots, or line graphs.
- Explore the basic characteristics of election results, voter demographics, and other relevant variables.Identify relationships between variables using correlation analysis.
- Conduct data visualization to uncover patterns, trends, or anomalies.
- Perform clustering or dimensionality reduction techniques to reveal underlying structures within the data.

Predictive Modelling and Inferential Analysis -

- Build predictive models to forecast election outcomes or voter behavior.
- Select appropriate machine learning algorithms (e.g., regression, classification) based on the nature of the analysis.
- Test hypotheses to determine the significance of relationships or differences within the data.

Factorial Analysis-

- Identify key factors or variables that significantly influence election results.
- Conduct factor analysis or principal component analysis (PCA) to uncover latent variables.
- Evaluate the impact of socioeconomic, demographic, or political factors on election outcomes.

Deliverables:

Report: Detailed analysis report summarizing the findings from the dataset and the prediction methodology.
Presentation file: A final PPT file with key insights and trends in the election data.

Conclusion:

In summary, this project is analyzing USA general elections from 1976 to 2020, uncovering trends in voter behavior and influential factors. Through descriptive and predictive analyses, as well as key factors analysis, new insights will be gained regarding demographic shifts, economic influences, and political events impacting election outcomes. The findings contribute to a nuanced understanding of electoral dynamics, offering valuable insights for policymakers and researchers alike.