THE SCIENCE OF WINNING U.S. PRESIDENTIAL ELECTIONS

Overview

Data analysis is central to the operation of almost all electoral systems through the relation of two things; the interaction between the punctiform distribution of voters (and their political choices) and the complexity of the election system structure. This invariably affects the outcome of the election results. The following information argues how understanding of such a system can impact partisan strategies to bring desired fate using data.

BUSINESS UNDERSTANDING

Business overview

We seek to aid a candidate in becoming the U.S President.
The winner of the election will be the candidate winning the most Grand Electors.
Grand Electors are attributed at the state level: In each of the 51 states, there is a given number of Grand Electors to win and the presidential candidate receiving the most local votes wins all the Grand Electors in that state.
Because the number of grand electors is not exactly proportional to the population,some states can be prioritized to increase the return on investment of the campaign.We are provided with 2 tables: one giving the number of Grand Electors per state, the other the population per state.
Our task is to identify the states that should be prioritized to win the election, with a smart but simple algorithm.

Business Success Criteria

Objectives

1. To aid a presidential candidate become president using data analysis of victor requirements.
2. To acquire data required for a presidential candidate to win the elections in a clean and simple format.
3. To identify the states that should be prioritized to win the election with highest grandelector number and minimal votes.
4. To find out the relation of the statorial population to the numbers of grand electors allocated in conjunction to winning the election.
5. To sort out the most valuable states while maintaining the running total of grand electors at minimal number required.

Data Mining goals

Expected benefits to meeting the objectives:

1. Save on campaign finance.
2. Acquire insight on approach to get required results.
3. To get the threshold we need to win the presidential election.
4. To make our priority list accurate.
5. To get a large number of grand electors by convincing a small number of people to vote.

Assessing the situation

Available data:

a. Grand electors by State. [Link]
b. Population by State. [Link]

Risk factors involved:

Data to be used is based purely on previous elections practices. However this can be ignored as the election trend in consideration is less likely to change.

Assumptions

1. The population dataset is currently up to date
2. The candidate shall acquire the majority votes in the states recommended.
3. Grand elector distribution dataset is correct.

Project Plan

The overview plan for the project is as follows;

Stage	Steps	Effect
Link data	Joining grand elector and population tables.	Data will be easier to work with.
Clean data	Shorten District of Columbia to “DC”	Requirement by Candidate will be met.

Construct the data	Compute ratio between grand electors and population	Attain population represented by one grand elector per state.
Integrate the data	Order states by decreasing ratio per capita,compute the running total,compute half of total grand electors in the country	Gain priority list,acquire threshold needed for winning the election.
Build the model	Filter out sorted list of states in order to keep the necessary ones	Enable evaluation of results.

DATA UNDERSTANDING

Data description

The existing data entails a list of number of grand electors allocated to each state and a list the population of each state. This data does not contain missing values and is arranged in a presentable format. However the population dataset seems to contain extra values that are not necessary. Both datasets contain State columns in different cases.

Both datasets have vital information in achieving our objectives and will have to be merged to a single clean table.

DATA PREPARATION

Selecting data:
Data from both datasets have to be merged into a single table for easier data cleaning.

Excluding data:
States from the population table that are not included in the Grand electors table states will have to be excluded as there are no such states.

Data cleaning:

States from all rows indicated as “District Of Columbia” will have to be renamed to “DC”.

Constructing new data:
A column with the ratio between the number of grand electors and the population can then be created. This is done by dividing the population per state by the number of grand electors per state.

We can then create our priority list by ordering the states by decreasing ratio of Grand Electors per capita.

We can then compute the running total of Grand Electors in that sorted list. This can be done by summing up all grand electors in the country.

The running total of Grand Electors will then have to be divided into two so as to attain the actual number of Grand Electors required to win the election.

MODEL TEST DESIGN

Our model test is created by filtering our sorted list of states in order to keep only the (top) ones enabling us to reach the computed threshold.This gives us our target list. Our target list should include the states whereby the total number of Grand Electors is equal to slightly above half of the total number of Grand Electors in the entire country.

EVALUATION

The results of our model are in a clear and presentable form. The study proves that less than a third of the total number of states is required to win the elections. A total of eleven states are needed.
If a candidate were to secure the majority votes in these states then he or she would automatically gain more than half of the running total of Grand Electors.It is important to note that these eleven states do hold more than half of the total country population.

DEPLOYMENT

Presidential candidate wishing to use the results of this study will have to devote the majority of campaign forces to attaining majority votes in the stated states.

The recommended states in our priority list are:
Carlifornia, Texas, Florida, New York, Illinois, Pennsylvania, Ohio, Georgia, Michigan, North Carolina and New Jersey
In that order.

FINAL REVIEW

Our study has concluded that the proposed methodology can guarantee accurate results if replicated in the deployment period.
It was also observed that the candidate would have to target either:

a. 41 out of the 51 states to acquire a minimum of 270 Grand electors and 147 million voters.(This option targets the least number of voters to be convinced.)
b. 11 out of the 51 states to acquire a minimum of 270 Grand electors and 183 million voters to be convinced. (This option targets the least number of states needed to be prioritized in the campaign.)

The Grand electors system ensures that all parts of the country are involved in selecting the president and without it areas with lower population would have been completely ignored leaving small towns marginalized. However, the same system gives too much power to certain states and allows the presidential election to be decided by a handful of states without worrying about the partisans actual popular vote totals.