

VAD FINAL PROJECT PROPOSAL

Group Members - Aditya Ghuge, Akshay Avhad, Archit Dukhande, Harshada Daund, Kapil Tare

Project Title

Exploring Campaign Finance Trends over the Last 5 years in Washington State

Project Objective

To analyze campaign financing trends over the last 5 years, focusing on understanding the role of contributions in influencing election outcomes. The analysis will provide insights for political analysts, campaign managers, and advocacy groups, helping them understand the impact of financial contributions on political success.

Data Description

To give some context about the dataset, it basically consists of summarized information about candidate campaigns and political committees by election year, spanning the past 5 years that is from 2023 to 2019. The primary data sources are the campaign reports such as deposit reports (C3), summary reports (C4), registrations (C1/C1pc), and candidate declarations.

Number of rows and columns identified - 18,530 rows and 52 columns

Mentioned below are some of the primary attributes -

1. **election_year**: Year of the election (useful for tracking trends over time).
2. **filer_name**: Name of the filer (could be candidates or committees).
3. **committee_city, committee_state, committee_county**: Geographic data, useful for regional analysis of contributions.
4. **office**: The office being contested, such as senator, governor, etc.
5. **party**: The political party associated with the filer.
6. **contributions_amount**: The amount of money contributed to the campaign or committee.
7. **expenditures_amount**: The amount spent by the campaign or committee.
8. **loans_amount**: Loans taken by the campaign or committee.
9. **independent_expenditures_for_amount**: Independent expenditures supporting the campaign.
10. **independent_expenditures_against_amount**: Independent expenditures against the campaign.
11. **jurisdiction**: Area of governance or influence, useful for mapping contributions to regions.
12. **position**: The position being contested (governor, mayor, etc.).

Currently we are in the process of narrowing down the number of features based on their relevance to the EDA exercise.

Exploratory Data Analysis (Phase Type - Initial Research)

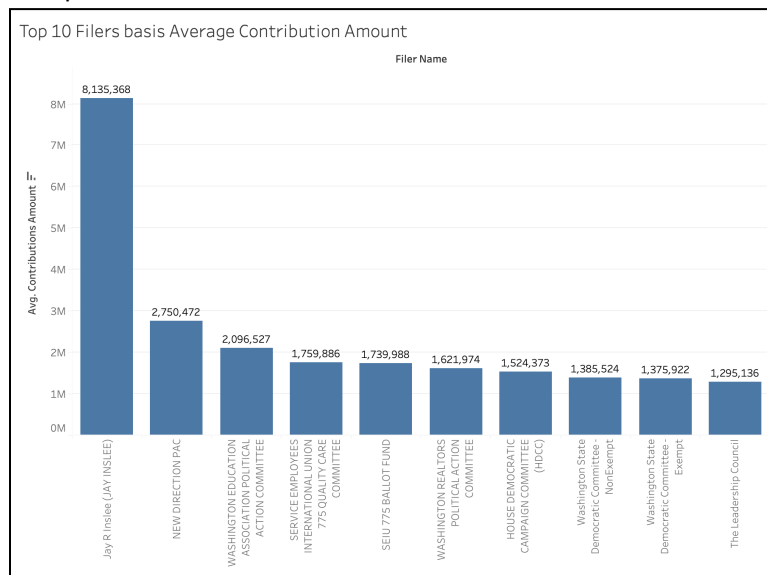
As part of our initial exploration, we identified 10 years of live data on political campaigns and committees, but to ensure data consistency and relevancy, we are focusing on the last 5 years (2023 to 2018) . Initially, the dataset contained **52,000 rows and 76 columns**, but after basic data cleaning steps, we have now reduced the dataset size to **18,530 rows and 52 columns**.

Currently we are in the process of -

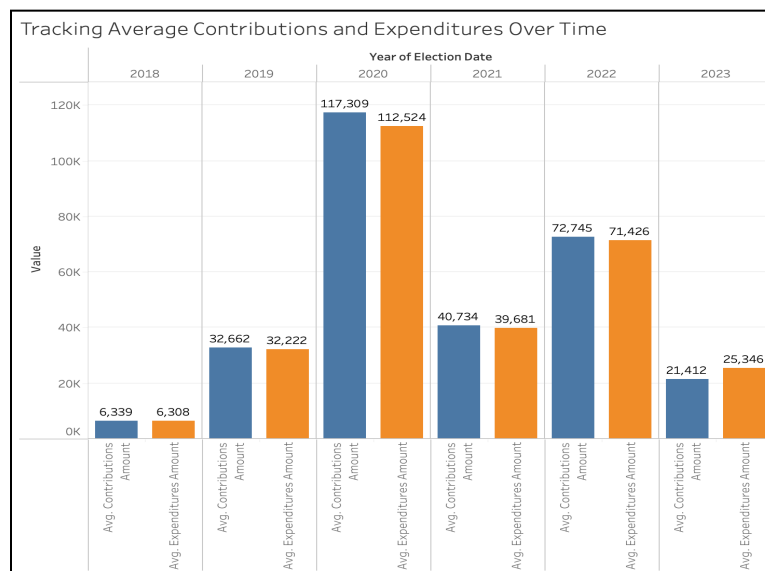
1. Identifying and addressing null values in key columns.
2. Implementing Imputation methods such as mean and median where applicable to ensure data completeness.
3. Performing a broader analysis to identify significant trends, outliers, and patterns in campaign activities, focusing on financial disclosures and contributions.

Below are two visualizations from our ongoing exploration exercise of the dataset -

1. **Top 10 Filers Based on Average Contribution Amount:** here we have highlighted the top 10 filers, which depicted Jay Inslee's campaign as a major outlier with significantly higher contributions compared to others.



2. **Tracking Average Contributions and Expenditures Over Time:** here we are comparing the average contributions and expenditures across the past five election cycles, showing how contributions peak during major election years like 2020.



Stakeholders

Below mentioned are some of the primary stakeholders for the consumption of our Tableau storyboard -

1. **Political Analysts:** Interested in trends, outliers, and patterns in campaign financing over time.
2. **Journalists:** Reporting on political financing, especially large contributors or notable trends.
3. **Campaign Managers:** Focused on optimizing campaign fundraising and expenditure strategies.

Expected EDA Visualizations for the final storyboard

1. **Line Chart** for tracking total contributions by election year to observe trends in campaign contributions over time.
 - **Objective:** Highlight trends and peaks during election cycles.
2. **Bar Chart** for comparing average contributions per candidate by year.
 - **Objective:** Show changes in financial support for candidates over different election cycles.
3. **Treemap** for identifying Breakdown of campaign expenditures by category.
 - **Objective:** Explore different types of spending to show which categories dominate.
4. **Scatter Plot** for identifying correlation between total contributions and election outcomes.
 - **Objective:** Reveal any relationships between financial contributions and candidate success.
5. **Geographical Map** for visualizing contributions by region.
 - **Objective:** Show the regional distribution of campaign contributions, helping to identify geographic trends.
6. **Summary Dashboard** which will consist of KPIs summarizing key metrics such as total contributions, expenditures, and the number of committees/candidates by year.
 - **Objective:** Provide a quick overview of the most important campaign finance metrics at a glance.