Analysis of Facebook Political Ad Data
Project Report

Joshua Wang

Abstract

Big data analytics has transformed decision-making due to the exponential increase in data generation. Traditional analytics methods are inadequate for handling the scale and speed of big data, requiring us to turn to scalable solutions like distributed systems such as Hadoop and Spark. This report leverages distributed systems to analyse the Facebook advertisement strategies of politicians running for federal parliament in Australia. Using Python and PySpark, this report examines spending patterns and targeted demographics and states based on a dataset spanning four years of Facebook sponsored political advertisements. The analysis reveals trends in advertisement spending over time, demographic targeting, and spending trends by party and candidate success. The findings indicate that there is increased spending and ad volume leading up to the Federal Election, with targeted demographics aligning with population demographics. Successful candidates tended to spend more overall but less per ad, suggesting a strategy favouring a higher volume of ads with smaller budgets yielded better election results. This analysis offers insights into the effectiveness of Facebook advertisement strategies in political campaigns and shows the importance of big data analytics in elections.

Table of Contents

Introduction

Big data analytics has revolutionized decision-making for organizations, driven by the exponential growth in data generation facilitated by the extensive use of digital platforms in daily life. Traditional analytics methods have become inadequate in handling the scale and velocity of big data, and this necessitates scalable solutions.

Distributed systems like Hadoop and Spark offer the scalability required to process vast volumes of data efficiently. By distributing tasks across multiple nodes, these systems enable parallel processing and real-time analysis of streaming data. A practical example that illustrates the need for distributed systems for handling big data analytics is in the operations of the supermarket retailer Walmart. Walmart processes over 200 billion rows of transactional data per week, making distributed systems a necessity to execute analysis at scale. Leveraging big data analytics, Walmart is able to efficiently compete on prices, ensure that locations have the right products that customers want to buy, and even monitor real-time sales data to identify when certain products have not been put onto shelves.

This report will use the same distributed systems that many organisations around the world use to perform big data analytics. With Hadoop and Spark, we will analyse the Facebook advertisement strategies employed by politicians running for federal parliament, and their impact on election success by examining spending patterns and targeted demographics/regions.

Dataset Analytics

The analysis in this report will focus on spending patterns and strategies used by candidates of the 2022 Federal Election. The code for analysis was written in Python and PySpark, and can be found in the appendix.

Datasets

The main dataset for this report consists of 4 years of Facebook sponsored political advertisements in Australia from March 17th 2020 to February 7th 2024. Each of the records in the dataset contain details for the advertisement campaign such as the funding entity, spending, time period of the campaign, number of impressions and details of the demographics.

Additionally, external datasets were also used to obtain information on the candidates and successful electees of the 2022 Federal Election. The lists of all candidates for the 2022 Federal Election was obtained through the Australian Electoral Commission (AEC), whilst lists for the current members of parliaments who were successful in the 2022 Election were acquired through OpenAustralia. During federal elections, Australia votes for candidates for the House of Representatives and the Senate, so there was a total of 4 extra datasets used.

Dates for the state and federal elections in Australia and population data was used throughout the analysis in this report. These were obtained from Wikipedia and Australian Institute of Health and Welfare, respectively.

Data Pre-Processing

The first pre-processing step was standardising candidate names in each of the candidate datasets. The list of all candidates from the AEC included full names and middle names, whilst the list of all electees only included first and last names. All middle names were removed, since these datasets would need to be linked together to identify which candidates were successful, and candidates' names would also need to be matched to their Facebook pages to find the ads run by the candidates.

Additionally, Stuart Robert was a successful candidate from the election, but resigned in 2023 and was succeeded by Cameron Caldwell. This resulted in the dataset of all candidates not including Cameron Caldwell, but his name appeared in the list of electees. Since the focus of this report is to analyse spending by candidates of the 2022 Federal Election, Cameron Caldwell was manually replaced by Stuart Robert in the dataset of electees.

Names of parties were also standardised, as some candidates affiliated with the same party used different names, such as "The Greens" and "Queensland Greens". Furthermore, to identify whether a candidate was successfully elected, an additional column storing a Boolean value was added to the dataset of candidates by comparing the names of all candidates with those in the list of successful candidates.

Next, the dataset of all candidates was joined with the main advertisement dataset. This was performed by checking whether the name of the candidate was found in the name of the funding entity of the post. By doing this, a resulting subset of the ad data with all advertisement campaigns funded by candidates in the 2022 Federal Election was created.

Finally, anomaly identification was performed. By observing sponsored posts with significant amounts of spending, it was found that posts with ID 198359791795746 and 2628325640740192 had duplicated versions with content asking users to follow the page. The two posts also spent $75,162,061 and $68,119,771 respectively, which were extremely high, so these were removed from the dataset.

Analysis of Advertisement Posts Over Time

To analyse trends in advertisement spending by candidates over time throughout Australia, two factors were considered: the number of posts, and the spending on posts. First, to visualise the number of active advertisements over time, the data was grouped by date, and the number of active posts was counted.

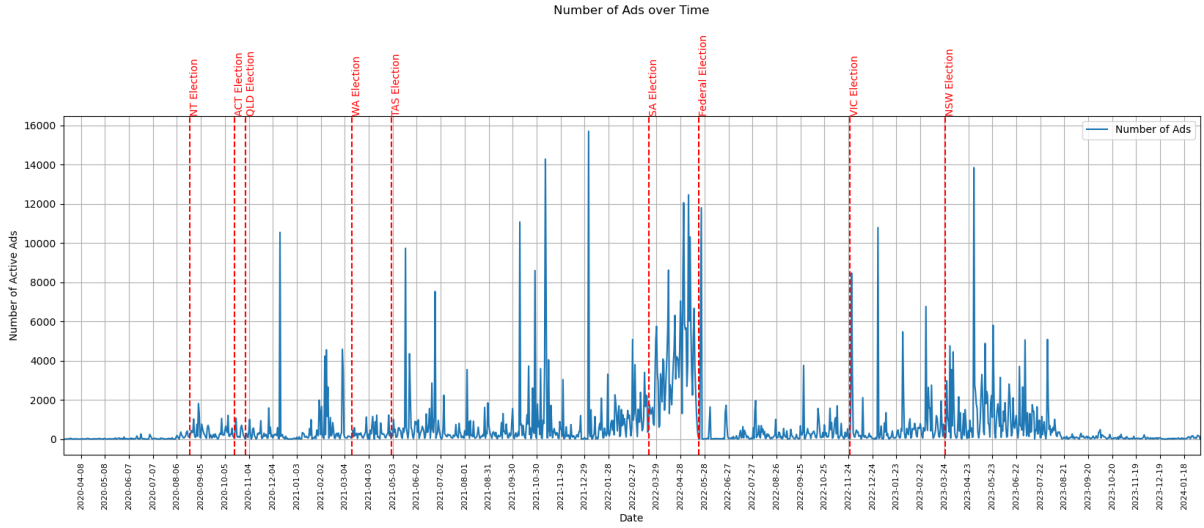

Figure 1: Number of Active Ads Run by Candidates over Time

In Figure 1, we can see the number of active advertisements tends to be mostly random, except from a significant build-up in posts leading up to the Federal Election. This build-up started approximately 6 months in advance from December 2021, and continued increasing until the election in May 2022. To analyse the amount spent on each of these posts, the data was grouped by day and the spending in each day was summed.

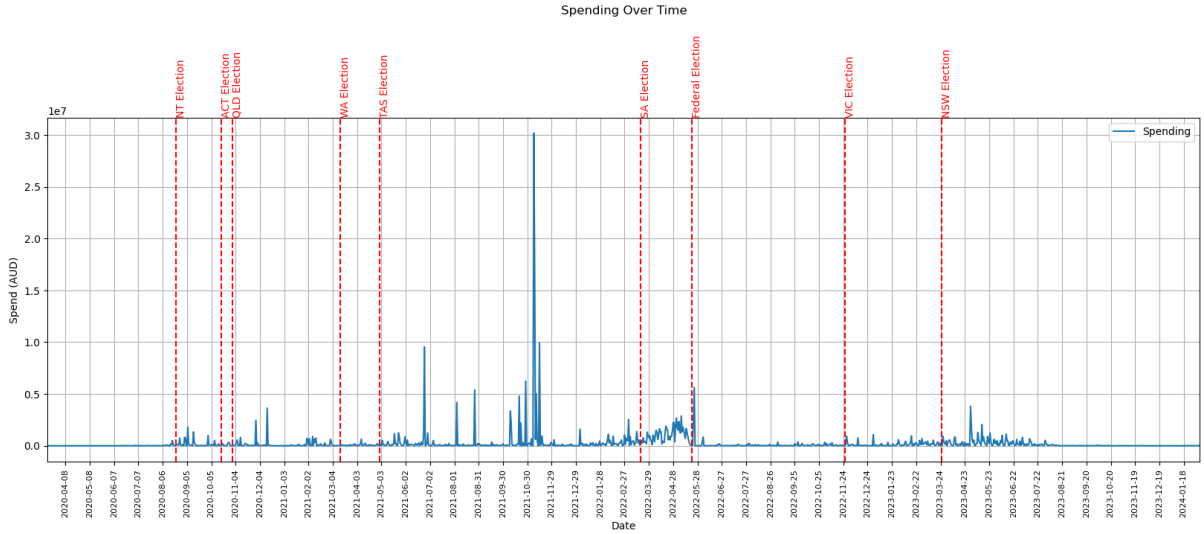

Figure 2: Amount of Spending on Ads by Candidates over Time

The same trend where spending does not follow any specific patterns apart from a build-up of spending prior to the Federal Election can be seen in Figure 2. We can also observe that spending on advertisements leading up to the Federal Election was not necessarily greater, with spending on other ads being greater than those just before the election. An anomaly in the spending is the significant spike during November 2021, and this was the result of a campaign by the One Nation party about COVID-19 vaccines.

In the time series analysis on advertisement spending, we only find patterns indicating more spending and posts prior to the Federal Election. Most other advertisement spending is relatively random and have no correlation with State Elections. This is likely because the candidates being considered in this analysis are those running for the federal parliament, meaning candidates only run posts for their own election and do not run ads in support of their counterparts at the state government level.

Analysis of Targeted Demographics

To analyse the demographics targeted by political advertisements run by candidates, we can first observe the number of impressions that advertisements attracted. The number of impressions in each demographic was first found by multiplying the number of impressions by the percentage of each demographic, before summing and grouping by demographic.

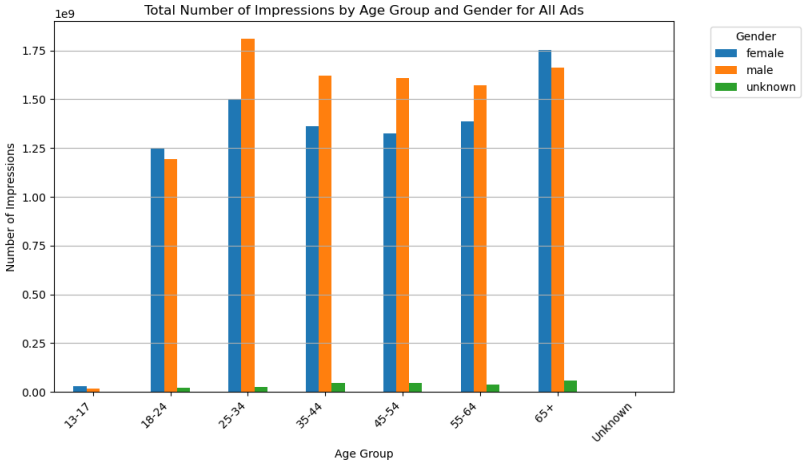

Figure 3: Number of Impressions per Age Group and Gender

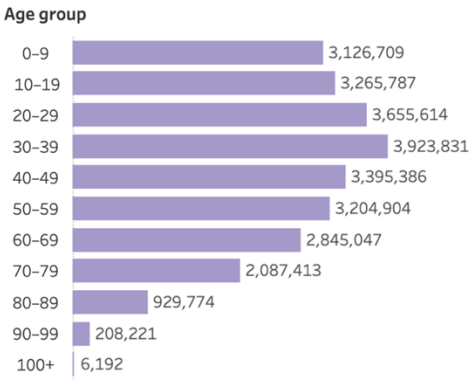

Figure 4: Demographic Snapshot of Australia at June 2023 (Australian Institute of Health and Welfare, 2024)

Figure 3 shows that the political posts target the age groups of 25-34 and 65+ the most often. This finding aligns with the fact that the majority of Australia's population lie within these age groups, which can be observed in the demographic snapshot of Australia in Figure 4.

The spending per state is another aspect of demographics to investigate. In Figure 5, the average spend per ad in each state/territory is plotted against its population. A regression model, shown in red, was also fitted to the data. Overall, there is a trend where the average ad spend by politicians was greater in states with larger populations.

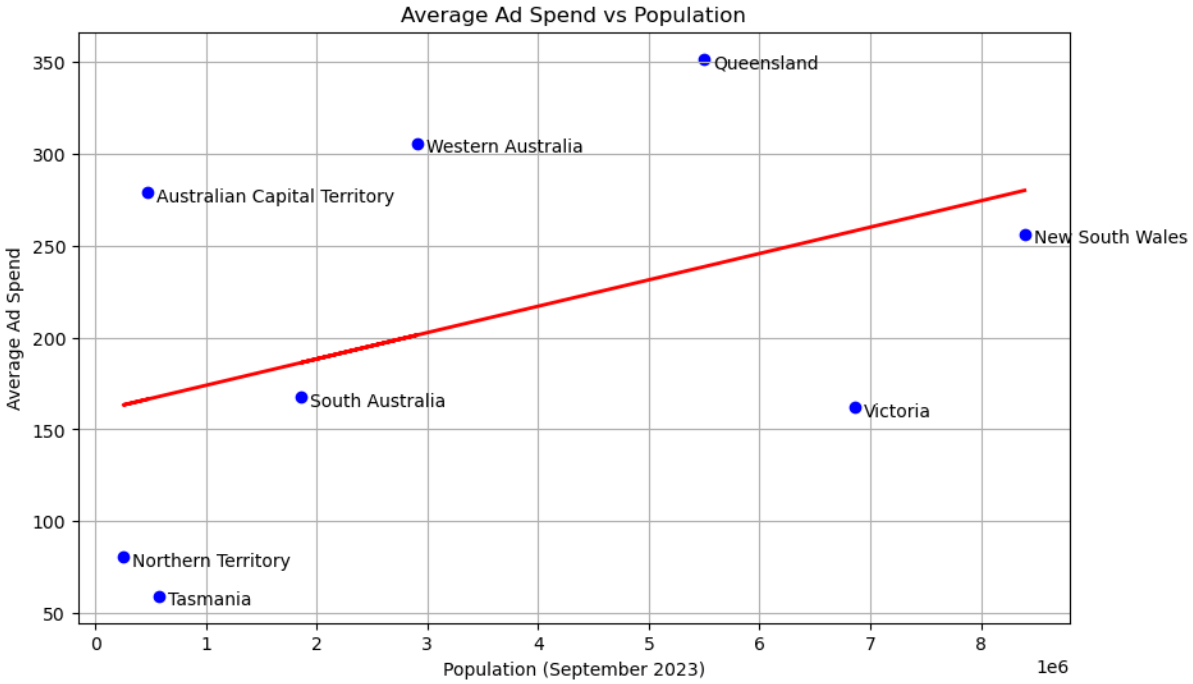

Figure 5: Average Spending per Ad in each State Compared Against Population

Analysis of Spending Trends

To observe if there are any trends in Facebook advertisement spending and election success, we next analyse the average spend by candidates in each party. This was obtained by grouping the advertisements by party, averaging the spending, and then splitting the data based on election results.

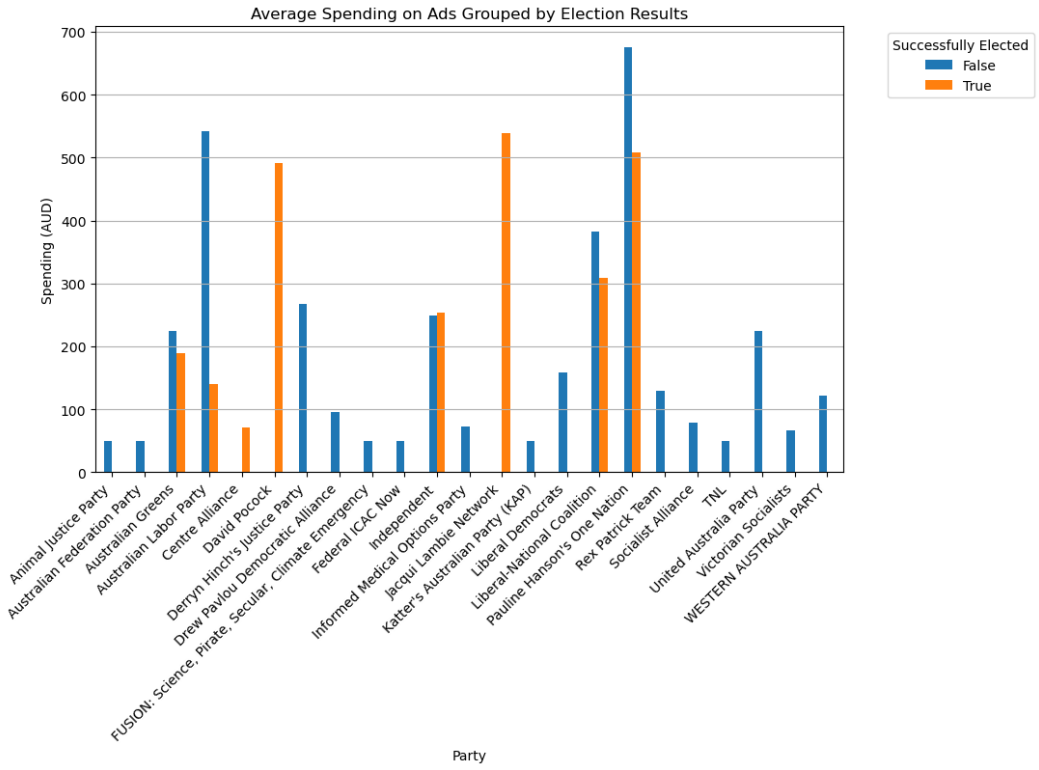

Figure 6: Average Spending per Ad in each Party Aggregated by Election Results

In Figure 6, we can observe that within most parties, the candidates which had a lower spend per ad tended to be successfully elected. This shows that spending more money per post did not necessarily lead to election success, and a more effective strategy may be to run a greater volume of lower-cost advertisements.

Party	Average Spending per Candidate	Average Spending per Elected Candidate
Liberal-National Coalition	$791,194.49	$839,015.21
Australian Labor Party	$711,823.61	$759,803.31
Australian Greens	$56,016.93	$84,626.39
One Nation	$1,197,653.71	$8,420,609.0
Centre Alliance	$81,025.0	$81,025.0
Jacqui Lambie Network	$311,061.5	$311,061.5
David Pocock	$94,204.0	$94,204.0
Independent	$669,013.11	$1,282,160.72

Table 1: Average Spending for each Candidate and Successful Electees per Party

Finally, for parties which had successful candidates, we can analyse spending patterns by comparing the average spending per candidate against the spending per electee. In Table 1, we find that for almost all parties, successful candidates tended to spend more than the average candidate from the same party. This further solidifies the theory that the volume of ads was an important factor, as successful candidates spent more overall (Table 1), despite usually spending less per ad (Figure 6).

Discussion and Conclusions of the Analysis

Overall, the analysis on spending and targeting of advertisements run on Facebook by candidates for the 2022 Australian Federal Election has identified key patterns in spending strategies by candidates of the Federal Election.

Firstly, creation of ad campaigns begins approximately 6 months before the election, and the volume and spending on active ads increases significantly as election day approaches. Patterns in spending by candidates can also only be found during the time period before the election, with all other spending being random and having no correlation with state elections. The results from analysing the targeted demographic was also expected, where candidates targeted age ranges and states that included more people.

Additionally, candidates who spent more were more successful in the election. However, the allocation of the spending was equally important. The optimal strategy was to run a greater number of ads, each of which uses smaller budgets. In future work, it may be interesting to apply the same analysis to candidates for the State Elections to observe if the same spending strategies emerge.

References

Representatives and Senators in Australia's Parliament. OpenAustralia.org. (n.d.). https://www.openaustralia.org.au/

Australian Electoral Commission. (2022, August 15). *2022 Federal Election: Downloads and statistics*. https://www.aec.gov.au/Elections/federal_elections/2022/downloads.htm

Profile of Australia's population. Australian Institute of Health and Welfare. (2024). https://www.aihw.gov.au/reports/australias-health/profile-of-australias-population

Wikimedia Foundation. (2024, March 31). *2022 South Australian state election*. Wikipedia. https://en.wikipedia.org/wiki/2022_South_Australian_state_election

Wikimedia Foundation. (2024, April 13). *2020 Northern Territory general election*. Wikipedia. https://en.wikipedia.org/wiki/2020_Northern_Territory_general_election

Wikimedia Foundation. (2024, April 13). *2022 Victorian state election*. Wikipedia. https://en.wikipedia.org/wiki/2022_Victorian_state_election

Wikimedia Foundation. (2024, April 21). *2020 Queensland state election*. Wikipedia. https://en.wikipedia.org/wiki/2020_Queensland_state_election

Wikimedia Foundation. (2024, April 24). *2021 Tasmanian state election*. Wikipedia. https://en.wikipedia.org/wiki/2021_Tasmanian_state_election

Wikimedia Foundation. (2024, April 25). *2020 Australian Capital Territory General election*. Wikipedia. https://en.wikipedia.org/wiki/2020_Australian_Capital_Territory_general_election

Wikimedia Foundation. (2024, April 26). *2023 New South Wales state election*. Wikipedia. https://en.wikipedia.org/wiki/2023_New_South_Wales_state_election

Wikimedia Foundation. (2024, May 2). *2021 Western Australian state election*. Wikipedia. https://en.wikipedia.org/wiki/2021_Western_Australian_state_election

Wikimedia Foundation. (2024, May 6). *2022 Australian federal election*. Wikipedia. https://en.wikipedia.org/wiki/2022_Australian_federal_election