

Canadian Federal Election Data Visualizer

CPSC 547 2025: Project proposal

Ricky Curry (rickycurry@gmail.com)

Gale Chen (galechen07@gmail.com)

Abstract

[This space intentionally left blank]

1 Introduction

Canada's federal elections and its candidate demographics offers a continuous record of how the country has grown, politically, socially, and geographically. Since the confederation of Canada in 1867, the country has had 45 federal elections. In that span of time, Canada has changed in numerous ways: it has grown from four provinces to ten provinces and three territories; its House of Commons has grown its initial 180 seats to 343; and its members of parliament have diversified to reflect its diverse population, especially in terms of gender expression and sexual identity.

Representation is crucial to the democratic process [1]. The progression of representation in Canada is reflected not only in who could vote, but also in who could stand for election. The gradual inclusion of women, indigenous peoples, racialized communities, and LGBTQ2S+ candidates reflects the contemporary struggles for representation and equity. Yet, disparities still remain. While the proportion of visible minority candidates across the major Canadian political parties continues to trend upwards (20.1% in 2025, versus 26.5% total population proportion as of the 2021 census), the elected proportion of women (30.3%), indigenous (3.5%), and openly LGBTQ2S+ (0.9%) candidates continues to fall well short of their population share [2]. As a country, we continue to wrestle with questions of proportional representation, indigenous self-governance, and gender parity in politics. Studying long-term trends in candidate demographics can help identify where progress has been made and where systemic barriers persist.

This project aims to visualize the trajectory of Canadian federal elections throughout their historical development. It utilizes two primary datasets: (1) *Candidates*: a dataset detailing the biographical information of Canadian federal election candidates from 1867 to 2021 [3]; and (2) *Federal Electoral Districts (FED)*: a dataset detailing the geography of federal electoral districts over the same period [4]. The resulting tool will enable users to interactively explore Canada's electoral, geographic, and political history.

2 Related Works

Several previous papers have utilized individual visualizations to communicate findings related to candidate demographics. Sevi's candidate dataset [3] is accompanied by an article [5] wherein she explores a variety of election candidacy questions that could be at least partly answered by this data, including the effects of incumbency, gender, occupation, and party affiliation. In doing so, she utilizes some simple static visualizations, including line charts and scatterplots. Lapointe et al. [6] use dot plots with error bars representing 95% confidence intervals to visualize the relationship between various minority identities and degrees of defeat, concluding that minority candidates were

3 Data and Task Abstractions

3.1 Dataset Information

We obtain our data from two sources. First, *Candidates* data covering the biographical information of Canadian federal election candidates from 1867 to 2021 [3], and second, *Federal Electoral Districts* data detailing the geographic boundaries of federal electoral districts (FEDs) over the same period [4].

3.2 Data Abstraction

The tabular *Candidates* dataset consists of around 46 500 items, each representing a specific candidate-district-election combination. The raw dataset has 32 attributes, detailed in Appendix A. Several attributes of particular interest to us are binary-categorical, such as indigeneity, sexuality (openly LGBTQ2S+ versus not), and election type (regular versus by-election). Several more are categorical with 3 to dozens of categories such as FED name, gender, country of birth, province, and party affiliation. Some attributes are quantitative, including election date, year of birth, and number of candidates in the running.

The *FED* dataset consists of named geospatial boundary data corresponding to historical electoral districts. There are 18 separate representation orders (ROs) corresponding to specific years, each spaced roughly one decade apart, to reflect census-backed federal redistricting mandates. The number of FEDs per RO ranges from 189 to 338. Two further details are worth noting. First, sometimes seats in the House of Commons are added during by-elections, usually the result of adding a new FED that occupies previously un-districted space, or resulting from splitting previous FEDs into more and smaller FEDs. Second, up until the 1966 RO, there were several FEDs with multiple representatives, so the number of seats was often larger than the number of FEDs. Further details are provided by Taylor et al. [10].

3.3 Task Abstraction

3.3.1 Who

The intended users of our visualization tool are a theoretical group of people who are interested in Canadian federal election history and topics relating to political representation of minorities. We (the authors) appoint ourselves as proxies for such users.

3.3.2 Actions

Our tool will support users asking questions such as, “what did my FED look like, geographically and politically, 100+ years ago?”, “is there an overall pattern between candidate demographics and party identity?”, “how common is it for a candidate to be elected with significantly less than a majority of votes (due to the first-past-the-post electoral system)?”, and, “what patterns underlie the temporal and geographic distribution of fringe parties?” (expanded and abstracted in Table 1). Furthermore, we conjecture that users will spend time freely exploring the dataset if presented with appropriate visualization tools, and would therefore benefit from tools that provide customization options.

Question	Task Abstraction
(1) What did my FED look like (politically, geographically) 100+ years ago?	Browse, discover
(2) Is there an overall pattern between candidate demographics and party identity?	Compare
(3.1) How common is it for a candidate to be elected with significantly less than a majority of votes (due to the first-past-the-post electoral system)?	Summarize, identify extremes
(3.2) Does this happen in some places more often than others?	Compare
(4) What patterns underlie the temporal and geographic distribution of fringe parties?	Explore

Table 1. Example domain-level tasks and their corresponding abstractions.

Most or all of these questions deal with similar domains of geography and history (time), so it is conceivable that some individual views could support multiple data types. For example, a visualization that arranges time along the x-axis could support various quantitative attributes along the y-axis, such as overall proportion of women candidates, or number of candidates run by each party. Adapting one view to multiple datasets with common encoding idioms simplifies both the implementation and the user learning curve. Similarly, a map displaying FEDs can use colour encoding for a wide variety of categorical and quantitative data to help identify geographic patterns in the data.

Solution

4.1 Tools

We intend to implement our solution using the D3 framework due to its powerful flexibility and our prior use experience. In particular, D3 natively supports animated transitions, which we hope to apply to visualizing FED changes throughout history (e.g. we can show how one or more FEDs evolve spatially over time, usually subdividing into more and smaller FEDs as time progresses).

We will also need to carry out some cleanup of our source data, which we plan to undertake using Python and the Pandas and GeoPandas¹ Python modules. We have observed issues such as duplicate FED names, e.g. in RO 1867 there two “Ottawa”s, where one is actually “County of Ottawa” and the other is “City of Ottawa”; text formatting non-uniformity, e.g. certain ROs’ FED names have accented characters that don’t load correctly with default settings in Python; and capitalization issues, e.g. capitalization of FED names that is inconsistent throughout both datasets. Sometimes the actual FED names change between official RO updates, e.g. 38 FEDs changed their names in 2004, between the 2003 and 2013 ROs [11], meaning that we need to consider how to take this effect into account for both candidate-FED matching and for user-facing display purposes.

¹ <https://geopandas.org/en/v1.1.1/>

4.2 Results/Solutions

The interface for this project will likely consist of a dashboard with multiple interactive and static visualizations. We are still considering which specific tasks to address with the visualizations, as well as the possibility of stacking tasks such that similar tasks share the same visualization, with a UI element which allows the user to switch between them.

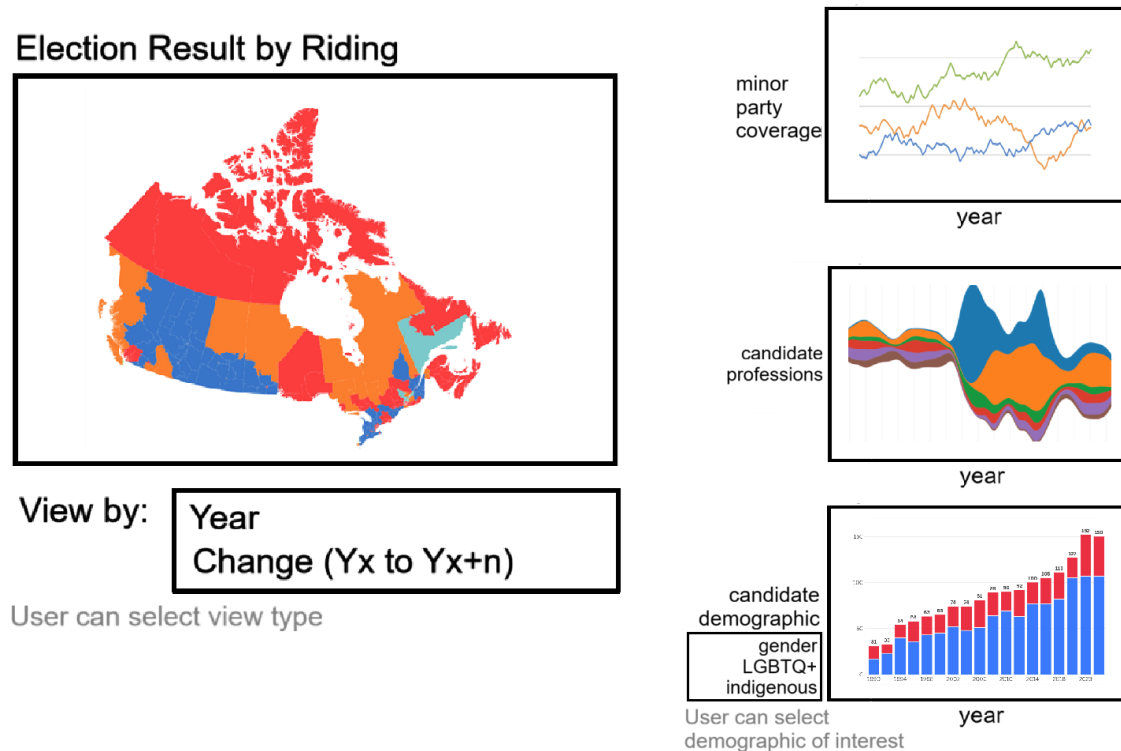


Figure 3. Low fidelity prototype of dashboard.

As shown in Figure 3, some potential visualizations we may implement include:

- 1) A choropleth map that allows users to view three ways: by singular year, by comparison between two years, and by animation over a range of years.
 - a) Liberal use of insets will help users see details in regions with high FED density.
- 2) A multiple line graph that allows users to compare how various quantitative attributes have changed over time for each party.
- 3) A streamgraph, possibly vertically normalized, that allows users to compare various proportional changes over time and between groups (e.g. candidate occupation category).
- 4) A stacked bar plot that allows users to view the change in low-cardinality candidate demographics over time.

4.2.1 Usage Scenarios

We present two of our domain-level user tasks as hypothetical usage scenarios based on our current conception of the tool and its features.

4.2.1.1 Visualizing a FED throughout history

This scenario explores question (1) from section 3.3.2: “What did my FED look like (politically, geographically) 100+ years ago?” The user starts by locating the FED of interest, either

by selecting it on the map or by finding it using a text search widget. Once the FED is selected, it is highlighted on the map, and the map view zooms to envelop the FED with a specified outer margin. In this view, the user is still able to pan and zoom the map view in case they desire more geographic context.

When one or more FEDs are in the “selected” state, the date slider UI widget is marked up with additional informational context (as with scented widgets [12]). For example, dates where the selected FED(s) split to become multiple FEDs, otherwise changed shape in some way, or flipped to elect a different party, will be indicated by the presence of corresponding glyphs, so that users can navigate directly to dates of particular interest. The data encoded by the colour channel in the FED boundary shape mark can be directly changed by additional UI widgets, possibly including elected party, win margin, number of running candidates, etc. Additionally, while in the “selected” state, the other views will be linked so as to filter their data to reflect only the current selection. This way, users can learn more about their specific FEDs of interest.

4.2.1.2 Party and sub-party popularity dynamics

This scenario explores question (4) from section 3.3.2: “What patterns underlie the temporal and geographic distribution of fringe parties?” The user begins at the streamgraph view. Using UI widgets, they configure the view to display party vote share distribution over time. By default, the least popular parties will be aggregated into one stream segment. By comparing this segment to the others, each corresponding to one of the major parties, they can infer the relative popularity (or lack thereof) of the fringe parties at any point in time. This view will be linked to the map view such that when the user hovers over a location on the streamgraph, the party and year (gleaned from the date closest to the cursor’s x-axis coordinate) will be used to display the corresponding party’s popularity in that election year in each FED. This will use a sequential single-hue colour scheme to highlight the regions where the party receives the most support.

In cases where parties are aggregated together into one segment, such as with the fringe parties, clicking on a segment will expand that group into its individual constituents and re-scale the y-axis to only the selected group. When a segment can be expanded, the cursor will change to a “pointing finger” cursor to indicate that interaction is available. Expanding segments into their constituent items allows users to see smaller trends that are not visible when scaled against the larger segments.

5 Milestones

Table 2 outlines our proposed milestones, their associated target deadlines, and the team members assigned to the task.

Milestone	Date	Person hours	Team members
All necessary data preprocessing is identified and specified	2025/10/25	4	Gale, Ricky
All data preprocessing is complete	2025/10/31	10	Ricky
Geographic view MVP Data can be displayed as choropleth map (e.g. vote proportion FED wins)	2025/11/12	15	Ricky, Gale

Time can be adjusted via UI element (e.g. timeline slider)			
Streamgraph view MVP - Can display various data over time - Major/minor party vote proportion - Candidate employment categories	2025/11/12	5	Ricky, Gale
Stacked bar chart MVP - Can display various data - Proportion of <50% vote wins by year, broken down by party	2025/11/12	5	Gale
Line plot MVP - Can display various data - Candidate count by party by time	2025/11/12	5	Gale
Implement linking between map and relevant plot views - Filtering data to selected FEDs on map - Highlighting FEDs from selected data on plots	2025/11/20	10	Ricky
Polish geographic view - Insets - Zoom mechanism - User can swap data intuitively	2025/12/01	10	Ricky
Polish streamgraph view - Can expand/collapse individual streams to focus on constituent groups - Navigation UI - User can swap data intuitively	2025/12/04	10	Gale
Polish stacked bar chart view User can swap data intuitively	2025/12/04	10	Gale
Polish line plot view	2025/12/04	10	Gale
Implement historical region animation	2025/12/10	20	Ricky, Gale
Final presentation - Prepare slides - Rehearse	2025/12/11	10	Ricky, Gale
Write final report	2025/12/15	15	Ricky, Gale

Table 2. Milestone timeline.

Discussion

[This space intentionally left blank]

Future work

[This space intentionally left blank]

Conclusions

[This space intentionally left blank]

Bibliography

- [1] D. Plotke, “Representation is democracy,” *Constellations*, vol. 4, no. 1, pp. 19-34, 1997.
- [2] J. H. Black, A. Griffith, “The diversity of candidates and MPs stalled for some groups in this election,” *Policy Options*, May 30, 2025. [Online], Available: <https://policyoptions.irpp.org/2025/05/diversity-federal-election/>. [Accessed Oct. 17, 2025].
- [3] S. Sevi, 2019, “Who runs? Canadian federal and Ontario provincial candidates since 1867. *Canadian Journal of Political Science* 54(2): 471-476.” Harvard Dataverse, doi: 10.7910/DVN/ABFNSQ.
- [4] Z. Taylor, J. Lucas, J. P. Kirby, C. M. Hewitt, Mar. 2023, “Replication data for ‘Canada’s federal electoral districts, 1867–2021: New digital boundary files and a comparative investigation of district compactness.’ *Canadian Journal of Political Science*,” Borealis, doi: <https://doi.org/10.5683/SP3/4E8DCR>
- [5] S. Sevi, “Who runs? Canadian federal and Ontario provincial candidates from 1867 to 2019,” *Canadian Journal of Political Science*, vol. 54, no. 2, pp. 471-476, 2021. doi: 10.1017/S0008423920001213.
- [6] V. Lapointe, B. Ferland, and L. Turgeon, “Still sacrificial lambs? Yes! Minority groups in Canadian federal elections, 2015–2021,” *Elect. Stud.*, vol. 87, p. 102717, Feb. 2024, doi: 10.1016/j.electstud.2023.102717.
- [7] M. A. Hadi, F. H. Fard, and I. Vrbik, “Geo-spatial data visualization and critical metrics predictions for Canadian elections,” *2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Aug. 2020, pp. 1-7. doi: 10.1109/CCECE47787.2020.9255824.
- [8] N. A. S. Abdullah, M. N. Mohamed Idzham, S. Aliman, and Z. Idrus, “Malaysia election data visualization using hexagon tile grid map,” *Soft Computing in Data Science*, B. W. Yap, A. H. Mohamed, and M. W. Berry, Eds. Singapore, Springer, 2019, pp. 364-373. doi: 10.1007/978-981-13-3441-2_28.
- [9] “16 ways to visualize US elections data.” Flourish. Accessed: Oct. 17, 2025. [Online]. Available: <https://flourish.studio/blog/report-on-elections-with-flourish/>
- [10] Z. Taylor, J. Lucas, J. P. Kirby, and C. M. Hewitt, “Canada’s federal electoral districts, 1867–2021: New digital boundary files and a comparative investigation of district compactness,” *Canadian Journal of Political Science*, vol. 56, no. 2, pp. 451-467, 2023. doi: 10.1017/S0008423923000185.
- [11] Supreme Court of Canada, 37th Parliament, 3rd Session. (2004, May 14). *Bill C-20, Chapter 19, An Act to change the names of certain electoral districts*. [Online]. Available: https://laws-lois.justice.gc.ca/eng/AnnualStatutes/2004_19/page-1.html
- [12] W. Willett, J. Heer, M. Agrawala, “Scented widgets: Improving navigation cues with embedded visualizations,” *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, vol. 13, no. 6, pp. 1129-1136, 2007.

Appendix

Table A. Data abstractions for *Candidates* dataset.

Column Name	Meaning	Data type	Cardinality
id	Unique candidate ID	Categorical	27658
parliament	Which sitting of parliament	Ordinal	[1, 44]
year	Year candidate ran	Quantitative	147
type_elxn	Type of election (general, by-election)	Categorical	2
elected	Whether the candidate was successful	Categorical	2
candidate_name	Full name (LAST, First)	Categorical	28067
edate	Election date	Quantitative	[1867/08/07, 2021/09/20]
incumbent	Whether the candidate was incumbent	Categorical	2
gender	Candidate gender	Categorical	3
birth_year	Candidate birth year	Quantitative	[1798, 1998]
country_birth	Candidate country of birth	Categorical	27
lgbtq2_out	Candidate is openly LGBTQ2S+	Categorical	2
indigenousoptions	Candidate is indigenous	Categorical	2
occupation	Candidate primary occupation	Categorical	7599
lawyer	Whether the candidate was a lawyer	Categorical	2
censuscategory	Occupation census category	Categorical	11
riding_id	Riding ID	Categorical	342
riding	Riding name	Categorical	1250
province	Province	Categorical	13
votes	Raw votes received	Quantitative	[0, 71535]
peprcent_votes	Percentage of total votes received	Quantitative	[0, 100]
acclaimed	Whether candidate is running	Categorical	2

	uncontested		
switcher	Candidate switched parties for this election	Categorical	2
multiple_candidacy	Candidate ran in multiple ridings	Categorical	2
party_raw	Candidate's party name	Categorical	162
party_minor_group	Candidate's minor party group	Categorical	50
party_major_group	Candidate's major party group	Categorical	8
gov_party_raw	Resulting governing party name	Categorical	4
gov_minor_group	Resulting governing party minor name	Categorical	2
gov_major_group	Resulting governing party major name	Categorical	2
num_candidates	Number of candidates in the given riding for the given election	Quantitative	[1, 21]