Canadian Federal Election Data Visualizer

CPSC 547 2025: Project Update

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 candidate representation within 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 how the demographics of federal election candidates have changed over time.

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 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. [5] 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

overrepresented in hard-to-win districts, and more likely to be nominated in districts where their parties experienced landslide loss in the previous election.

A few information visualization papers have also investigated the design of visualization tools for conveying election data. Both Hadi et al. [6] and Abdullah et al. [7] identify a common problem in that election data is reported in static tabular format, which is too dense for the layperson to understand and draw meaningful conclusions. Both converge to similar solutions involving interactive map-based visualizations. Abdullah [7] proposes a hexagon tile grid map for Malaysian election data, which additionally allows users to view riding racial demographics (Figure 1). Hadi [6] develops an interactive tool that allows users to select a Canadian election type and election year, whose results are displayed on a choropleth map of Canada. Users can further explore specific questions of interest via conventional plotting techniques generated by the tool.

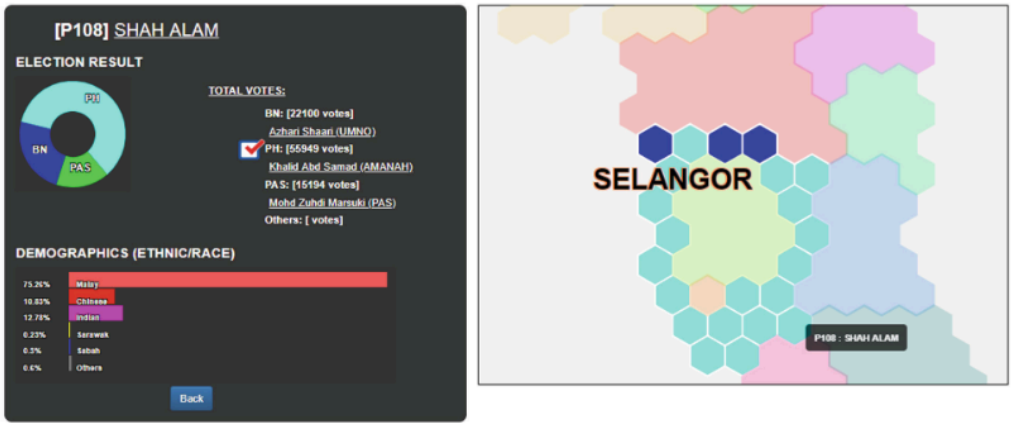

Fig. 1. Hexagon tile grid map of Malaysian election result and ethnic distribution. From Abdullah et al. [8].

Fig. 2. Treemap of European parliament structure by party group, then country [8].

Outside of the realm of academic research, there is vibrant discussion around the wide range of possibilities for visualizing election outcomes. Some other specific approaches include swing arrow

maps [8] to indicate changing voter support between subsequent elections and treemaps to depict hierarchical relationships within political parties and geographic units (Figure 2).

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 primarily support discovery regarding candidate and winner demographics. Users will be able to visualize information on the following attributes of interest (AOI): gender, sexuality, indigeneity, age, occupation, incumbency, margin of victory, and party membership. They will be able to explore these demographics as a statistic per year and change over a range of years. We have provided a sample of tasks we expect users to perform in Table 1.

Task	Task Abstraction
1. What did the FEDs look like, geographically, in a given year?	Browse, Explore
2. What are the election years where a candidate/winner demographic changed the most?	Identify, Discover
3. How do candidate/winner demographics compare between two years?	Compare
4. At a glance, what is the general pattern of demographic change over the entirety of Canada's political history?	Summarize

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

Solution

4.1 First Iteration

In our first iteration of a solution, we proposed a dashboard with multiple interactive and static visualizations. Tasks were stacked such that similar tasks share the same visualization, with a UI element which allows the user to switch between them.

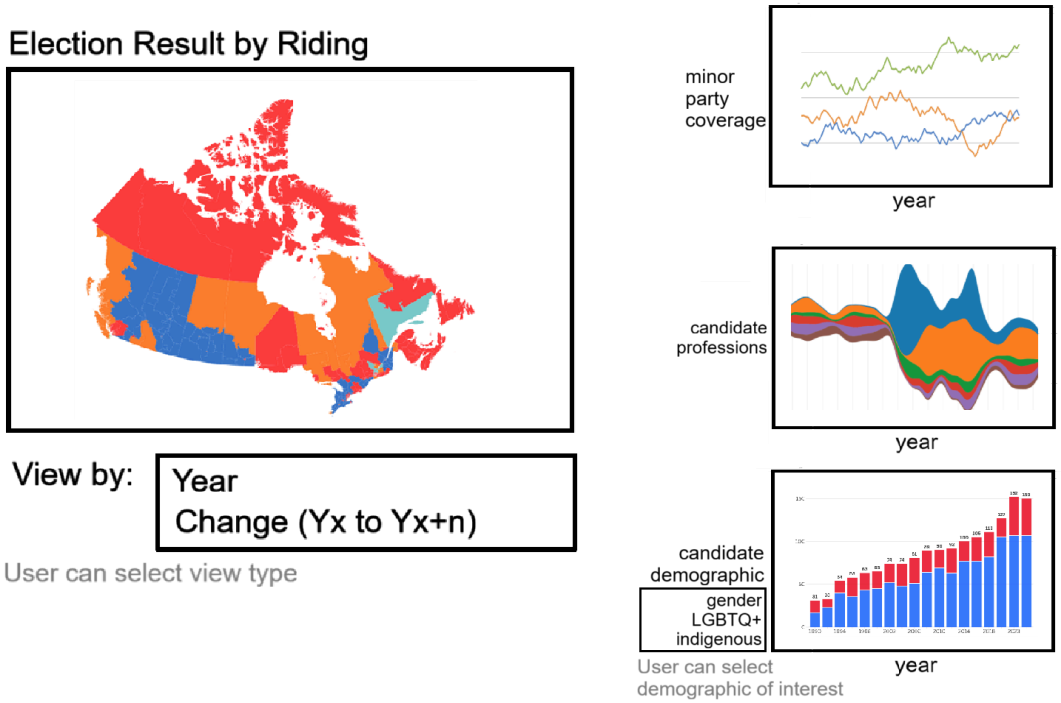

Fig. 3. Low fidelity prototype of dashboard.

As shown in Figure 3, some potential visualizations we considered included:

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 Current Iteration

In our current solution (Fig. 4), we propose a dashboard with three main components. On the bottom is a timeline slider that allows the user to control the range of years being viewed. Underneath the timeline slider is a heatmap that summarizes the parliament-by-parliament change for each AOI. There are three linked interactive views visualizing candidate demographics: a choropleth map, small multiple bar plots, and one main bar plot. By default, the AOI will be the number of candidates, and the slider will be set to uni-slider selecting the latest election year available, 2021. Compared to Fig. 4, our final interface aims to have the heat map aligned to each parliament tick, our drop down AOI selector should include LGBTQ+, and the large bottom right bar plot should allow a drop down menu on the x-axis to toggle between viewing demographics as a function of political party, and occupation census category.

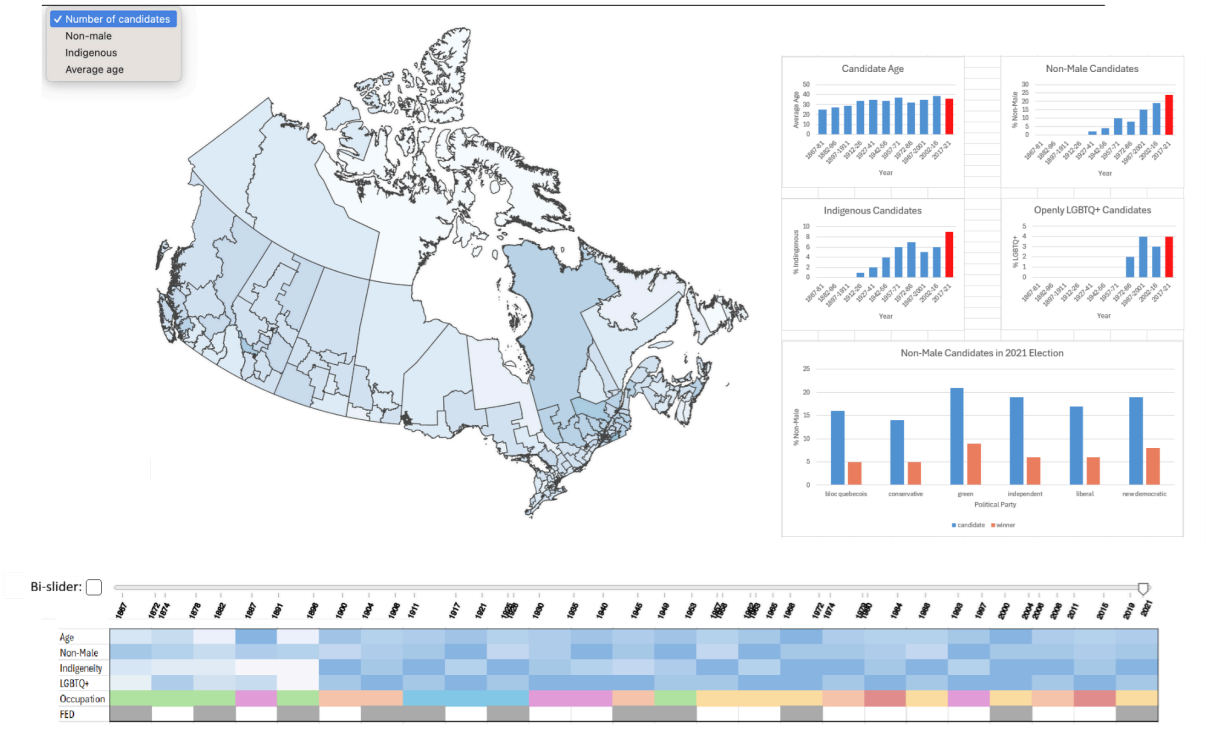

Fig. 4. Mock up of current interface.

4.2.1 Attribute of Interest, Slider & Heat Map

The slider allows the user to select either one year or two years. The years on the slider are mapped to each parliament. Given the large number of parliaments and candidate AOI, the tool required a high level summary view of the data. Therefore, the heat map under the slider shows the degree of change for each demographic trait from parliament to parliament. FED change is indicated by alternating blocks of grey and white. By referring to the heat map, the user may identify time

periods of interest and adjust the slider accordingly. The user can use the drop down menu on the top left to select which trait should be visualized on the choropleth map and main bar plot.

4.2.2 Visualizing FED geography

On the left pane is the choropleth map of the FEDs. When the user selects one year on the slider, the map simply shows the prevalence in percentage of the selected demographic trait in each FED. Sequential colour schemes are used for quantitative demographic traits, with higher prevalence being assigned darker colours. When the user selects two years on the slider, and the two years share the same RO, the map uses diverging colour schemes to represent the change for that demographic trait from year 1 to year 2 for each FED. If the two years do not share the same RO, the map simply shows the outlines of the FEDs for the two years superimposed. The later RO is foregrounded with a darker luminance and separate hue.

In designing the visualization of FED geography for a single year, we considered many options. We were cautious of simple choropleth maps due to the high density of small FEDs in urban centers which are confounding due to their disproportionately small visual perceptual footprint [9]. Alternatively, we considered the hex tile map, with each hexagon representing a standard population number (Fig. 5(a)), but a satisfactory automated process for converting geographic shape data to equal-area hexagonal cartograms does not appear to be readily available, and it would be highly time-consuming to spend time on a non-automated process for the 18 individual ROs in our dataset.. Lastly, we considered the population-proportional cartogram (Fig. 5(b)), but this would be too distorted to be a useful geographic map analogue. For the sake of simplicity, we opted for the choropleth map.

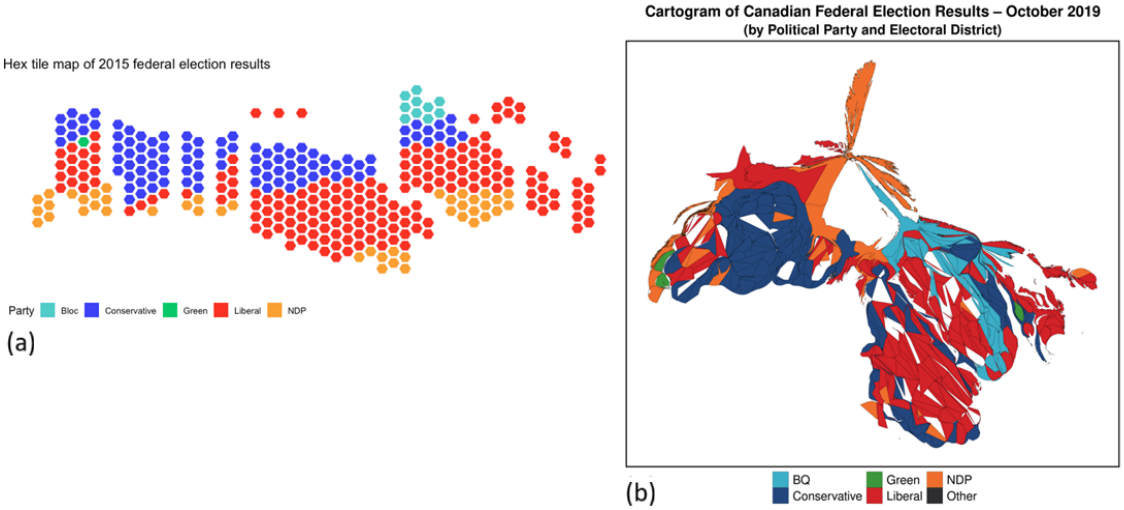

Fig. 5. Visualizations of Canadian election data. (a) Hex tile map. (b) Population-proportional cartogram. [10], [11]

Another complex question was how to handle visualization of candidate demographics on the map, as a change over a range of years. We considered a juxtaposed view of two choropleths, but due to the high latency of rendering even one choropleth view, we determined that it was not worth the time cost. Once we decided to include only one choropleth, we had to decide how to show the changing FEDs on one map. The obvious option was to show the borders of both the old and new FED superimposed, with luminance and hue double encoding for FED. However, given our interest in also colouring in the FEDs to indicate candidate demographic change, we determined that this would be too visually cluttered. Instead, we decided that if a user selects two years with different FEDs, the

map would simply show the outlines of the FEDs for the two years, without any demographic data projected.

4.2.3 Bar Plot

On the right pane are four small multiple bar plots that visualize the candidate age, gender, indigeneity and sexuality over the slider-specified range of years. When the user changes the selected years on the slider, the plot's x-axis (time) should rescale to put into focus the selected period. When the user selects one year on the slider, the small multiple bar plots highlight the corresponding bar. On the bottom of the pane, a larger double bar plot visualizes the selected demographic trait for both candidates and winners as a function of a categorical trait, i.e., party or occupation.

4.3 Usage Scenario

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

4.3.1 Non-Male Candidates and Winners

Imagine a user who is interested in understanding the patterns in the number of non-male election candidates and winners. First, they would select the AOI of "non-male" on the top left corner. The choropleth map would update to show the percentage of non-male candidates in the latest (2021) election. To identify the bucket of years where the first non-zero value appears, they would refer to the small bar plot on the top right showing the percentage of non-male candidates over time. Once they have identified the first year of non-zero value, they would select a bi-slider option, and move the sliders to view the period from that year to 2021. The map would update to show the percent change in non-male candidates within that time period. To further identify a year of interest, the user may refer to the heat map below the slider. They may identify one year where there was a large change in representation of non-male candidates. To see more details, they might return to the uni-slider view, select that year, and refer to the larger double bar plot. By default, the bar plot will show the breakdown of non-male candidates grouped by political party, but they may use the drop down menu on the x-axis to view by occupation census group instead.

4.3.2 Understanding Representation Change in FED Overtime

Imagine a user who is interested in understanding how representation in their FED has changed over time. 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. The user would be able to refer to the FED heatmap to identify years where the selected FED split to become multiple FEDs, or otherwise changed shape in some way. Once the user has selected the time period of interest using the bi-slider, The map would update to show superimposed FED outlines of the two tail years. To change the AOI encoded by the colour channel, the user would use the drop down AOI selector. Additionally, while in the "selected" state, the bar plot comparing candidate and winner demographics will update to reflect only the current selection.

5 Implementation

5.1 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 GeoPandas[1] Python modules.

5.2 Data Processing

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. We developed a bespoke data processing pipeline using Python to fix these issues, and to ensure one-to-one mappings between candidates and their correct FEDs and ROs by incorporating additional information about federal election dates found on a defunct Parliament of Canada webpage, accessed through the WayBack Machine [12]

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 [13], meaning that we need to consider how to take this effect into account for both candidate-FED matching and for user-facing display purposes. The simplest approach would be to choose one single name for each FED and RO, ignoring any mid-RO renaming, thus exchanging historical specificity for ease of implementation.

We converted our geographic shape data from its original EPSG:3348 coordinate system into EPSG:4326 using built-in GeoPandas methods for ease of manipulation in D3’s geography module. In the near future we will need to algorithmically simplify our geographic data shapes in order to render them more efficiently, since Canada’s coastal geography is highly convoluted and leads to very low frame rates and interface responsiveness. Early experiments with simplification resulted in various unacceptable deficiencies. Using one approach, shared FED boundaries were simplified differently, resulting in slightly overlapping or non-conjoining geometry. With another approach, some regions were rendered invalid, possibly due to re-entrant geometry, and could not be rendered by D3.

6 Milestones

Table 2 outlines our proposed milestones, their associated target deadlines, and the team members assigned to the task. Blue indicates completed tasks, and orange indicates in progress tasks.

Milestone	Date	Person hours	Team members
Individual Pitches	2025/09/26	4	Gale, Ricky
Idea Selection Meeting	2025/10/07	2	Gale, Ricky

[1] https://geopandas.org/en/v1.1.1/

Proposal Writing	2025/10/19	5	Gale, Ricky
All necessary data preprocessing is identified and specified	2025/10/25	4	Gale, Ricky
All data preprocessing is complete	2025/10/31	10	Ricky
Discussion of Interface Decisions	2025/10/06	4	Ricky, Gale
D3 tool familiarization	2025/11/10	5	Gale
Geographic view MVP • Data can be displayed as choropleth map (e.g. vote proportion FED wins) • Time can be adjusted via UI element (e.g. timeline slider)	2025/11/12	15	Ricky, Gale
Updated Interface	2025/11/12	2	Gale
Proposal Update Writing	2025/11/12	10	Gale, Ricky
Bar chart MVP • Can display various data ○ Proportion of <50% vote wins by year, broken down by party	2025/11/20	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 stacked bar chart view • User can swap data intuitively	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, doi: 10.1111/1467-8675.00033.

[2] “The diversity of candidates and MPs stalled for some groups in this election.” Accessed: Nov. 12, 2025. [Online]. Available: https://policyoptions.irpp.org/2025/05/diversity-federal-election/

[3] S. Sevi, “Who Runs? Canadian Federal and Ontario Provincial Candidates since 1867. Canadian Journal of Political Science 54(2): 471-476.” Harvard Dataverse, Apr. 11, 2024. doi: 10.7910/DVN/ABFNSQ.

[4] Z. Taylor, J. Lucas, J. P. Kirby, and C. M. Hewitt, “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, Mar. 27, 2023. doi: 10.5683/SP3/4E8DCR.

[5] 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.

[6] M. A. Hadi, F. H. Fard, and I. Vrbik, “Geo-Spatial Data Visualization and Critical Metrics Predictions for Canadian Elections,” in *2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Aug. 2020, pp. 1–7. doi: 10.1109/CCECE47787.2020.9255824.

[7] N. A. S. Abdullah, M. N. Mohamed Idzham, S. Aliman, and Z. Idrus, “Malaysia Election Data Visualization Using Hexagon Tile Grid Map,” in *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.

[8] “16 ways to visualize US elections data | Flourish.” Accessed: Nov. 12, 2025. [Online]. Available: https://flourish.studio/blog/report-on-elections-with-flourish/

[9] T. Munzner, *Visualization analysis & design*. in AK Peters visualization series. Boca Raton London New York: CRC Press, 2015. doi: 10.1201/b17511.

[10] A. jd, “Canada–Election 2019,” Clouds, Clocks, and Sitting at Tables. Accessed: Nov. 12, 2025. [Online]. Available: https://josipdasovic.com/tag/canada-election-2019/

[11] *Installing mapcan*. Accessed: Nov. 12, 2025. [Online]. Available: https://cran.r-project.org/web/packages/mapcan/readme/README.html

[12] “History of Federal Ridings since 1867.” Accessed: Nov. 12, 2025. [Online]. Available: https://web.archive.org/web/20110420154718/http://www2.parl.gc.ca/Sites/LOP/HFER/hfer.asp?Language=E&Search=G

[13] L. S. Branch, “Consolidated federal laws of Canada, An Act to change the names of certain electoral districts.” Accessed: Nov. 12, 2025. [Online]. Available: https://laws-lois.justice.gc.ca/eng/AnnualStatutes/2004_19/page-1.html

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
indigenousorigins	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 uncontested	Categorical	2
switcher	Candidate switched parties for this	Categorical	2

	election		
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]