# **Purpose**

The Data Cleaning Team (DCT) was charged prepping the raw source data for use in the creation of the “WeVote” Industry Project (WeVote) database as part of the Bow Valley College Tech Initiative Program.

This included the following:

1. Acquisition of raw data from data sources provided both by the client (source links) and WeVote management team (downloaded data files)
2. Assessing the quality of the raw data (fidelity, completeness, spelling, formatting, etc.)
3. Cleaning the raw data to remove any defects identified in Step 2)
4. Merging all the data into a single CSV formatted table (due to time constraints and internal functional issues, two CSV files were ultimately created rather than just one)
5. Formatting the final CSV tables to meet requirements established by the Development Team (DEV) as established in their data attributes and naming standards documentation

# **Project Roles**

DCT had a total of four members comprised of a Team Lead and three Analysts. Those members and their role descriptions are as follows:

**Ian Sidebottom - Team Lead**

Coordinated data cleaning team meetings and relayed issues/concerns to/from management team through regular management meetings and calls.

Performed initial QA of prepared data tables as part of cleaning process.

Coordination of the data cleaning process documentation and tracking of requirements/issues and solutions decided. Performed initial QA on risk assessment matrix.

**Jamie Schmidt - Lead Analyst**

Accumulated, cleaned, and merged raw data from source files into final two CSV files delivered to the Development Team.

With the support of fellow team members, assembled the master tables, performed calculations for certain fields, converted source file formats, adjusted appearances of string fields, and ensured assigned data types were applied to all. Errors discovered by QA and any changes requested by Development Team were implemented and the tables updated promptly. Primary developer of Data Cleaning’s risk assessment matrix.

**Hai Qing Huang - Analyst**

Assisted with ensuring raw data is correct, consistent, and usable. This included manual collection of data as needed, formatting of the data to comply with the business requirements, identifying and correcting errors in data, deleting unnecessary data, and quantifying risks associated with the potential un-removable errors in the data set.

**Elisabeth Gorrill - Analyst**

Participated in meetings and team discussions, responded to and assisted with team members’ queries. Assisted in data cleaning actions, including the removal of unwanted observations and fixing structural errors in the tables.  Provided cleaned data to the rest of the project teams for usage in the context of business requirements.

# **Data Sources**

A total of five, public data sources were used in the creation of the final WeVote CSV tables. One was an online database while two others were online visualizations of database information. These three sources provided the majority of the raw data used in WeVote. A further two were online maps and were used to establish relationships between communities, voting stations, and wards.

All five are published on the official City of Calgary website as outlined below:

1. Calgary [2017 Municipal Election Results by Voting Station](https://data.calgary.ca/Government/2017-Official-Election-Results-by-Voting-Station/atsy-3a4w)
2. Calgary [2017 Municipal Election Voter Turnout](https://www.calgary.ca/content/dam/www/election/documents/2017electionstats/voter-turnout-2017-vsd.pdf)
3. Calgary [Community Profiles](https://www.calgary.ca/csps/cns/research-and-strategy/community-profiles/community-profiles.html)
4. Calgary [2017 Municipal Election Voting Station Boundaries Interactive Map](https://data.calgary.ca/Government/Voting-Subdivisions-effective-October-16-2017-/nbbr-yfwc)
5. Calgary [New Ward Boundaries Map](https://www.calgary.ca/content/dam/www/engage/documents/Recommended_Ward_Boundaries_Map.png)

# **Initial Assessment of Raw Data**

The general consensus among DCT was that these data sources would most likely require minimal cleaning. This assumption was based on the fact that all three data sources were publicly available on the city’s website. As such, they were viewed as “official” resources and presumed to have already been cleaned and formatted for both human and computer consumption. DCT was simply repurposing the data for use in the WeVote.  
  
While this assumption held true for the 2017 Municipal Election Results data set, it did not for the other two. The 2017 Municipal Election Voter Turnout data set had one prominent error and the Community Profiles data set revealed a plethora of errors and inconsistencies.

Furthermore, as the data cleaning process progressed, problems (as opposed to errors) arose relating to the ward boundary changes implemented by the City of Calgary between the 2017 municipal election and the upcoming 2021 municipal election.

In later sections of this document, all of these issues will be itemized along with the remedies DCT debated and which of those were ultimately chosen as solutions.

# **Data File Format Conversions**

The two visualizations data sets were only available to WeVote in PDF format. With no access to the underlying databases on which these PDF files were built, DCT needed to convert them to a more database-friendly format, namely CSV.

Converting files from PDF to CSV is notoriously finicky, especially when using free, open-source conversion software. Even paid converters can struggle to make the PDF to CSV conversion perfectly. DCT used both types over the course of multiple attempts to convert these data files.

The most common error with such conversions is that the data of interest is not actually converted into multiple, comma-deliminated columns but rather is amalgamated into a single, multi-valued cell. This state is of little use for populating table columns and is often impossible to split into columns with ease.

Further still, even when a successful conversion is made, the targeted data must still be plucked from the resulting CSV files, transposed and reformatted, and finally inserted into the WeVote data files. This data manipulation was performed using Microsoft Excel, a tedious but ultimately successful task.

All this data manipulation is ripe with potential for the introduction of errors into the data set. Great care was taken by DCT members to prevent this from happening with some degree of quality assurance applied to underpin the accuracy of the data. Ultimately, though, this is the data cleaning step with the highest risk potential for artificially inserted errors into the data set.

# **2017 Municipal Election Results Data - Issues and Resolutions**

As was hoped, the 2017 Municipal Election Data set proved to be an exceptional clean and high-quality data set. The caveat to this statement is that DCT did not have access to the original voting results data source upon which the City of Calgary data set was created.

DCT (nor the QA team) had any means of confirming that the vote tabulations used in this data set are, in fact, correct. A “leap of faith” was required by the entire WeVote project team, as well as the client, that the voting results are 100% accurate.

Otherwise, no errors in the data source could be found. Spelling was perfect and formatting sound, though certain fields would be altered to meet Development Team’s requirements.

Two concerns were nonetheless identified:

* the inclusion of a Ward 0 (City of Calgary has 14 official wards numbered 1 through 14)
* the presence of seven separate vote types (Advance, Hospital, Mail-in, Outside Separate, Regular, Special, and Traveling)

## Ward 0 Issue

DCT determined that “Ward 0” was a catch-all for mayoral votes (and only mayoral votes) cast with advance, hospital, mail-in, and traveling vote types. This was problematic in that these votes could not be allocated to a specific ward, let alone specific voting stations and/or communities.

Refer to the following discussion and resolution of the Vote Types Issue which also eliminated the Ward 0 Issue.

## Vote Types Issue

That seven different vote types existed was not itself an issue. Rather, it was the inequality of these vote types that were of concern.

Regular and Special (i.e. seniors homes) vote types are cast only on election day and are tabulated by voting station. If a voter is a resident of Woodbine, for example, they vote at the voting station assigned to Woodbine residents and their votes are tabulated as results for that Woodbine voting station.

WeVote capitalizes on these associations between communities and voting stations in order to provide demographic information relative to voter behaviour on a community level of detail. Advance, Hospital, Mail-in, and Traveling vote types, however, are cast prior to election day and do not have the granularity necessary for the WeVote project’s mandate.

For example, a resident of Silver Springs (a NW community in Ward 1) may work downtown and cast an Advance vote type on their lunch break (downtown core is Ward 7). That voter will identify their Ward of residence when voting, but no information is obtained as to which specific community they live in. And since they are using a voting station in an entirely different ward, there is no way for that vote to be correctly associated with a specific community of residence in the WeVote database.

Two solutions were proposed by WeVote and discussed with the client, regarding this issue:

* create a secondary database with only ward level definition to be used only with Advance, Hospital, Mail-in, and Traveling vote types
* remove these four vote types, focusing the project solely on Regular and Special vote types

Creating a secondary database was feasible, though outside the initial scope of WeVote. While Hospital, Mail-in, and Traveling vote numbers represent a small proportion of overall votes, Advance vote results are not insignificant in number. Eliminating them entirely could introduce biases into predictions users derive from the voting/demographic relationships provided in WeVote.

DEV produced a preliminary ERD that would have accommodated this secondary database but upon further discussion with the client, it was determined that all non-Regular, non-Special vote types would not be included in the final WeVote project database.

The removal of Advance, Hospital, Mail-in, and Traveling vote types also eliminated the Ward 0 Issue since it was a construct of these non-Regular, non-Special vote types. When other vote types were removed from the project, Ward 0 ceased to be an issue.

The remaining vote type, Outside Separate, is a niche vote type limited to separate school trustee voters in three small communities on the city’s easter border (Chestermere, Indus, and Langdon). The vote results for these communities are especially small and community profiles for these localities are not available. Both the WeVote project team and the client agreed to remove this vote type from the data.

# **2017 Municipal Election Voter Turnout Data - Issues and Resolutions**

A significant issue was identified in the 2017 Municipal Election Voter Turnout Data during the quality assurance phase of the project conducted by the WeVote QA team. Some of the Special voting stations (seniors homes) showed a voter turnout greater than 100%, which is impossible.  
  
Upon further investigation by DCT, it was determined that of the 68 individual Special voting stations:

* 48 did not provide a total number of enumerated voters and therefore percentage voter turnout could not be calculated
* 16 did provide a total number of enumerated voters and the number of actual voters was less than enumerated voters, allowing for a valid percentage voter turnout to be calculated
* 4 did provide a total number of enumerated voters but the number of actual voters was greater than enumerated voters, resulting in an illogical percentage voter turnout calculation

Three possible resolutions to this issue were discussed by DCT:

* Leave the data as is and include a warning in WeVote documentation
* Remove all voter turnout data for all Special voting stations
* Nullify the four Special voting type with the illogical voter turnout calculation

Due to the absurdity of the voter turnout calculation in some instances (e.g. voting station 1433 had voter turnout of 1450%), keeping the data as is and including a warning in documentation was quickly discarded as a viable option by DCT.

In the grand schemed of things, these Special voting type numbers are small by comparison to the Regular voting type numbers. Removing them completely, for all Special voting stations, was a legitimate, viable solution. And with only 16 out of a total of 68 such voting stations having “valid” voter turnout calculations, this was seriously considered as a solution by DCT.

However, it was eventually decided to nullify the four offending Special voting stations. With 48 such null values already in the data, an additional 4 was not viewed as an egregious change to the data set. Furthermore, with age distribution representing a prominent segment of the community demographics data, keeping the 16 Special voting stations with valid voter turnout calculations, though small, was deemed more valuable than removing them since they represent a specific age group of voters.

**Important** – It remains vital to note that since 4 Special voting stations had a demonstrably erroneous “enumerated voters” number, it remains unclear if the “enumerated voters” numbers for the 16 Special voting stations with valid voter turnout calculations are indeed correct. This is a risk that remains in the WeVote database. It is small enough that it is unlikely to affect overall results derived from the database, but it nonetheless exits.

The following table summarizes the Special voting stations that were nullified and the remaining Special voting stations which potentially suspect enumerated voters data.

|  |  |
| --- | --- |
|  | Voting Station Number |
| 4 Special Voting Stations with Erroneous Voter Turnout Calculation | 634, 830, 1031, 1433 |
| 16 Special Voting Stations with Valid Voter Turnout Calculation | 131, 133, 232, 431, 730, 736, 737, 738, 832, 837, 838, 1135, 1136, 1231, 1331, 1431 |

# **Community Profiles - Issues**

By far, the majority of the issues identified by DCT in the data sources used in the WeVote project were found in the Community Profiles.

These profiles were the sole source of all the community demographics data used in the final data set from which the WeVote database was created. The profiles are published in PDF format and downloadable from the City of Calgary website. They are created from detailed Government of Canada 2016 census data compiled by the city via access provided by the [Community Data Program](https://communitydata.ca/).

Myriad demographic data are presented in these profiles, with identical data values often shown in multiple locations. Additionally, calculated percentage values are often included alongside numeric values. With the WeVote database table design calling for both numeric and percentage fields, these profiles offered DCT an opportunity to scrape both numeric and percentage values directly without the need for our own calculations.

During the cleaning and compilation process, it became clear to DCT that numerous inconsistencies existed within these Community Profiles. These issues are as follows:

## Population Sums Issue

Many of the Community Profiles have errors in population summations for total population, total by gender, and total by age group. The persistent and varied nature of this issue is best shown in this example using the [Bel-Aire community profile](C://Users/Stephanie/Downloads/bel-aire.pdf):

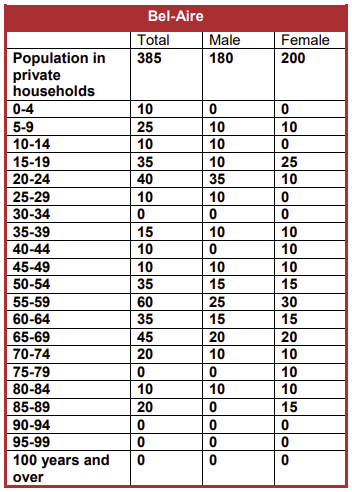


Figure - Screen grab directly form profile PDF

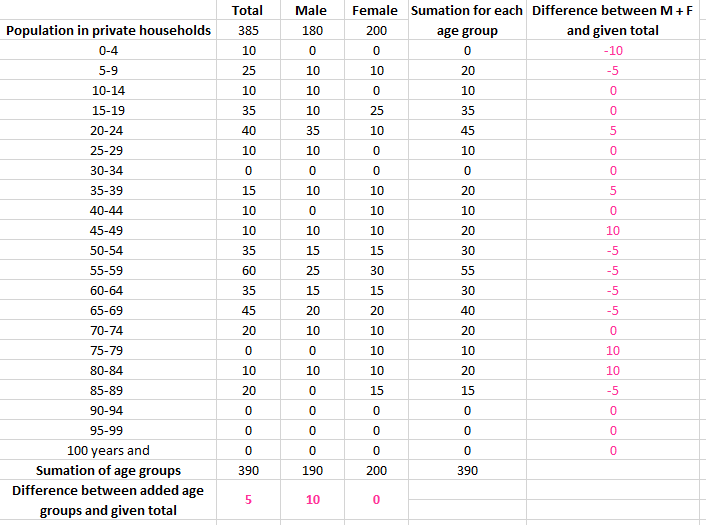


Figure - Calculations performed by DCT member in Excel spreadsheet

Notice the following discrepancies in the above data:

* The TOTAL POPULATION shown = **385** butTOTAL MALE + TOTAL FEMALE = **380**
* The TOTAL MALE POPULATION shown = **180** but SUM of MALE ages = **190**
* The TOTAL of all the individual age groups MALE and FEMALE = **390**
* The TOTAL in several individual age categories have errors (e.g. 0-4, 50-54, 85-90)

## Total Private Households Issue

The total number of private households was also found to be contradictory within different sections of several Community Profiles. Again, an illustrated example using the [Eagle Ridge community profile](C://Users/Stephanie/Downloads/eagle-ridge.pdf):

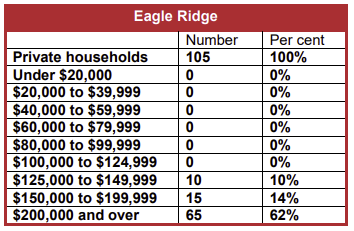


Figure - Screen grab from income section of profile

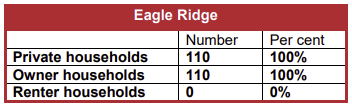


Figure - Screen grab from home ownership section of profile

Notice the following discrepancies in the above data:

* TOTAL PRIVATE HOUSEHOLDS in the INCOME data = **105**
* TOTAL PRIVATE HOUSEHOLDS in the OWNERSHIP/RENTAL data = **110**
* SUM of individual INCOME groups = **90**

## Total Census Families Issue

More inconsistencies are found in the way the number of families are summed. The variance in these errors is detailed using both Eagle Ridge and [North Haven Upper community profiles](https://www.calgary.ca/csps/cns/social-research-policy-and-resources/community-profiles/north-haven-upper-profile.html):

Diagram

Description automatically generated

Figure - Screen grabs from families sections of profiles

Notice the following discrepancies in the above data:

* Line 2 + Line 5 should = Line 1
  + For Eagle Ridge it does, for North Haven Upper it does not
* Line 3 + Line 4 should = Line 2
  + For Eagle Ridge it does not, for North Haven Upper it does
* Line 3 + Line 4 + Line 5 should = Line 1
  + For Eagle Ridge it does not (5 too many), for North Haven Upper it does not (5 too few)

## Cumulative Percentages Issue

As an internal quality assurance measure during the data assembly phase, DCT summed the percentage values of individual constituents in each demographic category for each community. Since these percentage values had been taken directly from the Community Profiles rather than having been calculated from the number data (and with so many discrepancies in that number data), DCT wanted to ascertain if any additional errors were present in the percentage data.  
  
There were, in fact, many although most were minimal (i.e. 99% or 101% instead of the accurate 100%). However, there were a few cases where more significant discrepancies in these cumulative percentage calculations (e.g. income for Bayview =88%). The figure below presents just a portion of the communities but shows the regularity and variance of this issue in the community profiles:

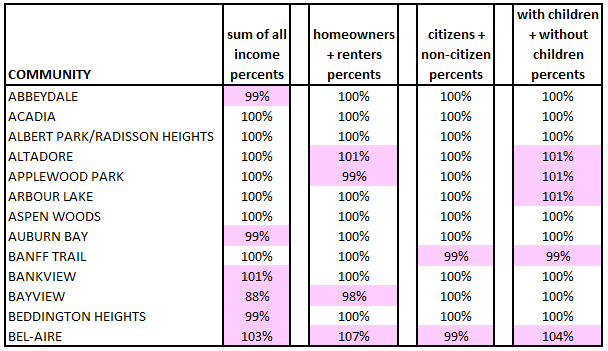


Figure - Calculated in Excel by DCT member

**Important** – The population data used in the WeVote data set was limited to voting age demographics (i.e. 18 and older) only. The percentage for each age group is calculated relative to the total population of that community. With citizens under 18 excluded, the cumulative percentage of the age groups used in WeVote could not equal 100% regardless. This does not mean that similar errors were not present in the WeVote community population percent values, they were just not visible.

# **Community Profiles - Resolutions**

All of these issues reveal a singular data risk: an unreliable data source. It is clear that the data presented in the City of Calgary Community Profiles is riddled with errors and discrepancies. The cause of these errors is unclear, but it is worth noting that the totals miscalculations always show an error of plus/minus a multiple of 5. The persistence of this fact is surely more than a coincidence however DCT could not determine the reason for such an oddly specific error.

With these issues having such a pervasive impact on the WeVote data set, resolutions were discussed both within DCT and then further with DEV and WeVote project management. The following options were discussed:

* Leave the data as is
* Petition client to pay for access to the raw data through the Community Data Program
* Recalculate totals and percentages using the constituent data (where applicable)

DCT was in favour of the first option: leaving the data as is. The reasoning was three-fold.

First, it’s all publicly available data published on the official City of Calgary website. There is no lack of transparency here in the sense that the client, or any other end user, can obtain and view the data themselves and see the noted discrepancies. None of the errors was generated internally during the WeVote data cleaning process nor are we hiding or “painting over” these errors.

Second, DCT had no way of determining the cause of these errors in the Community Profiles without access to the raw data from which they were built. To remove randomly occurring errors without fully understanding their origin did not resonate with DCT as a scientifically valid means of addressing the error.

Third, with only a very few exceptions, the bulk of the discrepancies/errors were insignificant. A calculation error +/- 5 on sums in the hundreds or thousands is not going to radically influence results derived from the database. Nor is a cumulative percentage 1% or 2% above/below the expected 100%.

Ideally, DCT would have had access to the raw data on which the Community Profiles were created, enabling the team to identify and correct the errors. Or, DCT could have simply extracted correct values for all required demographic parameters without the extra step of using the Community Profiles.  
  
Though all agreed this was undoubtedly the best solution (though itself not guaranteed to be error free), the cost and time constraints of WeVote within the confines of the Bow Valley College Tech Initiative program precluded this resolution from being implemented.

**Important** – If the client determines at some point in the future that more precise demographic data is necessary for the proper functioning and reliability of this database, then DCT recommends obtaining the raw data rather than using these prefabricated community profiles.

DEV, and to a lesser extent, project management, voiced significant concerns with leaving the data as is, preferring to recalculate all totals and percentages based on constituent values. DCT noted that doing this has inherent risks since we do not know where the errors come from in the original data. There is no guarantee that the constituent values are “correct” and that only totals were calculated incorrectly.

Ultimately, DCT acquiesced to the DEV’s wishes and proceeded to recalculate all demographic data points, such as totals and percentages, using constituent data. This eliminated all inconsistencies in values such as total population and it ensured all cumulative percentages equalled 100%.

**Important** – The Data Cleaning Team, and by extension the WeVote Project Team, make no claims that the demographic data presented in the database is accurate.

# **Community Profiles – Seton Anomaly**

One anomaly DCT purposely introduced to the WeVote database involves the community of Seton, a newer community in Calgary’s southeast. During the 2017 Municipal Election, Seton residents were considered part of Auburn Bay for voting purposes. However, a single Special voting station existed inside the community of Seton, separate from the Regular voting station shared with Auburn Bay.  
  
DCT therefore decided to include Seton as an individual community in the demographic data set despite there not being a Seton community profile in the Community Profiles source data. The result is that Seton currently exists in the WeVote database as an entity but, as yet, has no demographic data available.

There remains some uncertainty within DCT as to whether this was the correct course of action. Since it is a Special voting station, the impact on the overall validity of the WeVote database is insignificant. As WeVote evolves over coming election cycles, this anomaly will evaporate as Seton becomes an established and recognized community with published demographics data.

# **Multi-Community Voting Stations – Issues and Resolutions**

An issue surfaced during the initial attempts to merge the two primary data sets (vote results and demographics). Not all communities vote at individual, community-specific voting stations. Many do, but in some situations more than one community votes at a single, designated voting station. Furthermore, the vote results tabulated for that designated voting station do not differentiate between the two (or more) communities assigned to it.

DCT recognized that in these situations, those voting results would need to be divvied up between the two (or more) communities that voted at a particular station. How to do this appropriately was unclear and led to significant discussion within DCT as well as with DEV.

The obvious solution would be to split votes using a percentage based on a weighted average of total population for each community. The problem with such a solution is that it assumes all communities voting at a specific voting station vote identically. This might not be the case if two communities with significantly different economic demographics vote at the same station, for example.

Ultimately, after much discussion between teams and further consultation with the client, it was established that in situations where multiple communities vote at a single voting station, the eventual project interface will return all communities using said voting station even if the user only indicated a single community.

So, for example, Crestmont and Valley Ridge both vote at voting station 112. If the user identifies Crestmont as their community of interest, the interface will present the complete vote results for voting station 112 alongside the demographic data for both Crestmont and Valley Ridge since this is the data that can be reliably compared.

# **Ward Boundary Changes – Issues and Resolutions**

A similar issue arose once the new ward boundaries were accounted for in the data sets. The reallocation of some communities to new wards presents a unique problem when comparing 2017 voting results in a 2021 (and beyond) frame of reference.

There are 4 situations where 2 or more communities shared a single voting station in 2017 but have since been split into separate wards for the 2021 election (and beyond). For example, Tuxedo Park and Winston Heights Mountview were both in Ward 7 and both voted at Voting Station 712 in the 2017 election. After the ward boundary changes, Tuxedo Park remains in Ward 7 but Winston Heights Mountview has moved to Ward 4.

This is an added wrinkle to the issue described above. With demographic data established on community boundaries and vote results established on voting station boundaries, splitting voting stations between wards complicates matters since the original project mandate has no intent of mixing communities from neighbouring wards.

The presumed users are future candidates, most likely councillors and trustees, who by definition only stand for election in specific wards (1 ward for councillors, 2 wards for trustees). A candidate running for Ward 4 councillor in 2021 will not be interested in vote results for Ward 7 (i.e. Tuxedo Park) but because Winston Heights Mountview has moved from Ward 7 to Ward 4, part of the voting results at Voting Station 712, which was in Ward 7 in 2017, are of interest.

How to reconcile this issue was discussed but ultimately it was determined to be beyond the scope of Phase I. How to resolve this issue will be up to the Phase II team once it is better understood in terms of database and dashboard functionality. The client will also be better positioned to comment on their needs in these few, unique situations.

As such, no data allocation was performed. All required fields to make necessary calculations are included in the final CSV files, most notably the New Ward and Old Ward entities. An issue-specific resolution, however, has not, as yet, been introduced.

|  |  |
| --- | --- |
|  | Impacted Communities and Voting Stations |
| Voting Stations “Split” by Ward Boundary Changes prior to 2021 Municipal Election | 712, 802, 1104, 1207 |
| Communities That Changed Wards Prior to 2021 Municipal Election | Britannia, Douglasdale Glen, Downtown West End, Eau Claire, Elbow Park, Elboya, Erlton, Garrison Green, MacEwan Glen, Mission, Parkhill, Rideau Park, Riverbend, Roxboro, Sandstone Valley, Spruce Cliff, Wildwood, Winston Heights Mountview |

# **Changes to Data Formatting**

The final stage for DCT was formatting the data within the final CSV files to meet the standards set out in the DEV’s Data Attributes document. This required modifications to the appearance of some data including that obtained from the otherwise untouched 2017 Municipal Election Results data set.

That original data set included unique formatting for the candidate names; their surname first (all in caps) followed by their first and middle/nickname (if desired) names in proper form. This is how candidate names appear on actual voting ballots and it made some sense to be kept that way in the database. As such, DCT did not initially change the candidate name field despite it not being best practice.

DEV’s focus, however, was on keeping the database as small as possible from a storage standpoint. Nice as the unique formatting looked (and DCT made community names all caps as well for consistency), it does require database data types that consume more storage than the minimal possible.

Ultimately, database storage arguments won the day and DCT modified the master data files as much as possible to meet the criteria established in DEV’s documentation. In addition to reordering candidate names and removing all caps, superfluous characters like hyphens and slashes were also removed where justified (proper names of candidates including hyphens were exempted from this rule).

# **Outstanding Data Risks**

Despite best efforts of DCT, outstanding risks to data veracity remain. Any data manipulation comes with the inherent risk of human error. Cutting and pasting, transposition, file conversion, and even sorting all leave the data vulnerable to mistakes.

DCT did its best to mitigate all such errors to the best of their ability. Recognized issues are catalogued in this document and the applied remedies have been discussed along with the reasoning for implementing them.

Several introduced errors were caught by the robust quality assurance regimen applied to the final CSV data files by the WeVote QA team and corrections made.

Still, a few minor glitches made their way into the database and were identified by DEV while attempting to build the database tables. Those were corrected directly within the Sql Server environment without DCT’s involvement though DCT did confirm the corrections.

# **Implication of Data Risks**

The cumulative implication of all these data risks, both know and unknown, is that any decisions and/or actions taken based solely on results provided by the WeVote project require appropriate caution. WeVote is a tool, and like any tool, proper and careful use is critical.

This is particularly important when making predictions of future voter behaviour based on community demographics and past voting behaviour. The vast majority of errors in the underlying data are minimal in size and impact, but users need to understand the potential additive effect they can have on results and any conclusions derived from those results. Many minute errors can prove just as detrimental to results as a single, obvious error.

Furthermore, communities are continually in flux. With ever-evolving community demographics, the WeVote project becomes less reflective of voter proclivities as the underlying database ages. Future updates are vital in limiting the impact of these risks.

# **Final Aside**

Although the deadline for transfer to Development had passed, DCT did manage to create a single, master data file as initially envisioned when the project began. Unfortunately, it was too late for use in creation of the database, but it remains as proof of concept that such a table was possible.