Economic and Demographic Trends in Downtown Calgary: Supporting the Downtown Revitalization Strategy

Team Members and Role Assignment

Karen Reyes Gonzalez: Handled historical demographic data for Calgary and worked on forecasting models. Karen also introduced an updated methodology, which she clearly explained to the team and partners. Karen also contributed to writing sections of the project reports.

Jesse Tweedie: Managed the project timeline, coordinated communication with the partner organization, and ensured deliverables were met. He contributed to report writing and effectively presented the project objectives, regression models, and key results.

Jessica Silva Martins: Led the development and testing of the VAR model. She presented her findings and some of our visuals to the partners, helped fill gaps in the historical data, and contributed to writing the final report.

Jayden Yuzicapi: Worked on historical data processing, focusing on data cleaning and integration. He assembled the final Python file from team contributions, supported most of the analysis and visualization efforts, contributed to all report writing, and presented at every stage.

Bozhao Wang: Worked with historical dwelling data for Calgary communities, developed visualizations, and contributed to historical modeling. He played a key role in the principal regression analysis for population and dwellings at the community level.

Nihal Amin**:** Worked on historical data for communities and wards, created visualizations, conducted regressions at both ward and community levels, and supported report writing and presentation preparation. Also managed administrative tasks, such as maintaining an organized shared folder for the team.

Project Objective:

The objective of this project is to analyze historical, current, and emerging residential demographic trends in Calgary's downtown communities (Beltline, Downtown Core, East Village, Downtown West, and Eau Claire). By leveraging data from Statistics Canada, the City of Calgary, we aim to understand:

- How is the population distributed across the five communities of interest?
- What indicators best capture population changes in downtown Calgary?
- How are building permits shaping the housing landscape?
- What might the future look like?

The team has formulated a condensed research question that the data analysis will seek to answer: *How have the population, age distribution, household size, and household types in downtown Calgary changed from 2011 to 2025, and what are the demographic projections for the future?*

This project creates value for the City of Calgary by:

- Providing data-driven insights into residential demographics to support targeted investment decisions in downtown communities, such as lowering vacancy office rates and how population could be a leading variable for this vacancy.
- Helping the city understand population shifts to ensure the right return on investment.
- Offering forecasting models to predict future residential trends, enabling proactive planning and resource allocation.

Partner Defining Project Success:

The downtown strategy team is interested in receiving a Python script to generate visualizations, extract insights, and perform forecasting analyses.

Methodology:

We will utilize all available Ward data to understand the relationships among population, dwellings and demographic variables. The demographic data, such as age distribution and

household types, is available every five years, while we have yearly data on total population and the number of occupied dwellings.

To identify how these factors interact, we construct a comprehensive regression model. This model quantifies the relationship between each demographic variable and key predictors: total population, occupied dwellings, and the ratio of population to dwellings. The regression equation takes the form:

$$\text{Demographic} = \beta_0 + \beta_1 \times \text{Population} + \beta_2 \times \text{Dwellings} + \beta_3 \times \left(\frac{\text{Population}}{\text{Dwellings}}\right) + \epsilon$$

In this equation, β1, β2, and β3 are coefficients that reflect the sensitivity of each demographic variable to changes in the predictors. By analyzing these coefficients, we can deduce how shifts in population and housing data influence demographic characteristics across the city.

Once we derive these coefficients, we apply them to estimate demographic changes at the community level on a yearly basis. This application is feasible because we have annual data on population and occupied dwellings for each community. Consequently, even with demographic data only available every five years, our model allows us to predict trends annually.

This approach allows us to forecast demographic changes continuously, rather than relying solely on five-year census updates. By generating annual predictions, the City of Calgary gains valuable insights into demographic trends within specific communities. This methodology is particularly powerful as it leverages available yearly data on population and housing to provide timely forecasts. Additionally, this approach can be enhanced by incorporating other data sets that the partner has access to, such as historical demographic information available for wards prior to 2011. These additional data resources can refine the model and improve its accuracy. Ultimately, the value of our contribution lies in the methodology itself, providing a framework that can be enriched with more data, rather than solely focusing on precision.

Key Developments Since Progress Report 1:

The team worked on spatial and temporal downscaling, aiming to merge different datasets (Community Census & Building Permits and Ward Census & Building Permits) to analyze the relationship between population change and Work Class Group change (demolition, new construction, and improvements).

However, data discrepancies and missing observations limited our ability to run meaningful regressions or estimate coefficients for census years. As a result, we could not reliably use this approach for future forecasts or filling gaps in missing years.

Instead, we agreed to:

- Utilize individual datasets to assess correlations between population and occupied dwellings.
- Track trends in demolitions, new construction, and improvements to understand their impact on downtown housing.
- Examine household composition trends in downtown Calgary using census data from 2011, 2016, and 2021, focusing on changes in household types, including families with children, families without children, and single-person households.

Regression Analysis:

To refine our insights, we will conduct the following regressions:

- Change in occupied dwellings, regressed on change in population.
- Change in demolitions/new construction/improvements, regressed on change in single-family permits.

Equations to be used:

$$\text{Demographic} = \beta_0 + \beta_1 \times \text{Population} + \beta_2 \times \text{Dwellings} + \beta_3 \times \left(\frac{\text{Population}}{\text{Dwellings}}\right) + \epsilon$$

Forecasting:

It was estimated VAR forecasting for the variables: population, dwelling and person per unit for the community Beltline. Because the dataset just had the information by 1968 to 2017 and 202, to have the entire series, an interpolation was done to add the years of 2018, 2019 and 2020. Thereby, before the estimation of VAR, the following tests were performed:

- The unit root test was done in the series to see if they were stationary. Population and dwelling were not stationary, then the first difference was taken to do these series stationary. On the other hand, persons per unit was significant and rejected the hypothesis of unit root.

- The Granger Causality test was estimated to verify if one variable predicts the others, that is, each one causes each other. The results depicted that population causes dwellings and persons per unit. Dwellings cause population and persons per unit and Persons per unit causes population and dwellings.

- The Cointegration test of Johansen illustrates that trace statistics were higher than critical values for two vectors of cointegration, then the series can be cointegrated.

In the last part, the VAR was estimated with forecast until 2030. The VAR showed that population and dwellings will increase by 2030, but the same will not happen with persons per unit that will decrease.

For **Data Sources**, we continue to leverage the following sources:

- Open Calgary – City of Calgary:
  - Provides historical demographic variables (such as population, dwelling counts) at the community level.
  - Other demographic variables (such as household size, total census families, and household type) for the last three census periods (2011, 2016 and 2021)
  - Building permits data sets, and we chose to focus on work process variables that include new, demolished, and renovated buildings from 1999 to 2024 at the community level, which we use as a proxy for population changes.

Prototype testing:

Objectives:

- Validate the accuracy and usability of the Python script, visualizations, and forecasts.
- Gather feedback from the City of Calgary to refine the data product before final delivery.
- Ensure the deliverables align with the city's goals for downtown revitalization and targeted investment decisions.

Structure:

Two Testing Sessions with the downtown strategy team:

A- **Outcomes for the first session (March 28, 2025):**

- Presented the initial prototype.
- Shared demographic visualizations (e.g., population distribution, person per units heatmaps and charts).
- Demonstrated Regression Models for spatial downscaling and discussed challenges in merging two datasets.
- Agreed to focus on individual datasets to better estimate relationships between different variables influencing population changes and downtown demographics.
- Presented early-stage forecasting results.
- The partner found the Beltline-focused forecasting results most useful and indicated that expanding forecasts to other communities was unnecessary.
- Collected feedback via meeting notes and began incorporating it into our analysis (refining forecasting and individual dataset usage).

B- **Pre-Final Delivery Testing Session:**

This session couldn't be conducted due to time constraints.

Results & Outcomes:

The data product include the following features:

Visualizations:

- Heatmaps showing population distribution in the 5 communities downtown (Beltline, Downtown Commercial Core, Downtown East Village, Downtown West, Eau-Claire), permit types, number of types of housing per year, population distribution and household size.
- Charts showing historical trends in population distribution and dwelling counts for the available wards (7 & 8) with an estimated line for the wards.

- Graphs highlighting demographic trends as interpolated for the available years and forecasting for the next 5 years.

Regression Models:

- Analyze the relationship between changes in occupied dwellings and changes in population and demographics across all Calgary wards.
- Apply the estimated coefficients from the ward-level models to project demographic changes in downtown communities.

Forecasting:

- Estimation of VAR forecasting
- Using population prediction growth, dwellings and persons per unit up to 2030.

Results:

- **Building Permits (1999–2024):**
  - Demolitions consistently outpaced new builds and improvements across downtown communities.
  - Demolition activity peaked around 2015, particularly in Beltline, which had the highest overall permit activity among the five downtown communities.
  - Single-family housing permits and demolitions showed negative relationhip while the apartments permits had a positive one.
- **Population trends (Census Years 2011, 2016, 2021)**
  - Population in downtown communities increased from 2011 to 2016.
  - From 2016 to 2021, population growth stabilized, with the trend line appearing almost flat over this period.
- **Population and Demographics:**
  - Age Groups: The dominant household type was single-person households (Household Size 1), consistently the largest across all census years.
  - Gender Distribution: Remained relatively balanced over time.
- **Regression Results (wards level):**
  - Dwelling Count (ΔDwelling) on Population Change:
    - The coefficient of -0.4003 (p-value = 0.005) suggests that as the number of dwellings increases, population change slightly decreases, after controlling for other factors.

        - The negative relationship might reflect the type of dwellings being added (e.g., smaller units attracting individuals without large household growth) or that demolitions could offset the positive impact of new construction.
        - With statistically significant result ($p < 0.01$), the relationship is unlikely due to random chance.
    - Population (Level) on Population Change:
        - The coefficient of -0.255 (p-value = 0.004) indicates that larger existing populations are associated with slightly slower growth.
        - This could suggest that more densely populated communities may have limited capacity for further rapid growth (due to saturation or limited space).
    - Single-Person Households (Household Size 1):
        - The strong positive coefficient of 1.02 (p-value = 0.000) for dwellings indicates that an increase in dwellings is strongly associated with an increase in single-person households.
        - This relationship is highly statistically significant, highlighting the central role of smaller housing types in downtown demographics.
    - Children Aged 5–14:
        - A negative relationship (-0.4003) between dwelling increases and the child population suggests that new developments downtown are not necessarily attracting families with children.
        - However, the population coefficient is positive (0.2995, p-value = 0.000), showing that overall child populations do grow where there is general population growth, but not because of new dwellings specifically. Essentially family growth is tied to broader community growth, not new construction.
    - Multicollinearity Challenges:
        - Using both dwelling counts and population-to-dwelling ratios together in the model introduced multicollinearity, to address this we used:
            - Control variables (like male/female age groups) were added
            - Log transformations were applied to key variables.
        - These methods successfully reduced multicollinearity and improved the stability of the model, although the ultimate solution would be more granular or larger datasets.
        - Forecasting:
            - The VAR showed that population and dwellings will increase by 2030, but the same will not happen with persons per unit that will decrease.