

Assessing Core Housing Needs in Major Canadian Cities

DATA 604 - Fall 2020

Joel Penner, Janny Tam, Peter Baran

Introduction

The rental market is a key indicator of many important economic topics for neighborhoods, cities, and the health of an entire nation. The housing market has been of great interest to economists in Canada for many years. With hot housing markets come hot rental markets. Renters may have a harder time keeping up with rental pricing due to the increase in house prices, poor job prospects, and finding the right dwelling for themselves and their families.

To assess the availability of adequate housing, we will explore the 2011 and 2016 census data for three major Canadian housing markets: Calgary, Vancouver, and Toronto. All of our data will be acquired from Canada Mortgage and Housing Corporation (CMHC); a crown corporation that acts as Canada's national housing agency. We intend to investigate the differences in housing adequacy between survey zones and within different demographics.

Our investigation will focus on CMHC's measure of 'Core Housing Need' (CHN). CHN is a two-stage indicator that helps to identify who needs housing assistance and thus identifies vulnerable populations. A CHN occurs when housing requires a major repair that the residents cannot afford, when there are not enough bedrooms for the residents, or when the home costs more than the residents can afford. Many of the data tables in our dataset have percentages that represent the percentage of households that qualify for core housing needs.

In the individual queries, our objective is to find notable trends in our three cities using SQL. We will identify survey zones where CHN data shows that adequate housing is hard to find. Finally, we will combine our dataset to compare how each city provides adequate housing for key demographics such as families in need, seniors, and immigrants.

Individual Datasets

Our datasets were downloaded from CMHC's official housing portal. The CMHC has licensed our data 'as is' with free terms of use. Under the CMHC License Agreement, CMHC offers a "revocable worldwide, royalty-free, non-exclusive license to use, reproduce, publish, distribute, or sell the information" (CMHC, 2018). CMHC also reserves the legal right to limit or discontinue our use of their information if we breach their terms including misrepresenting the information provided.

The data sets we chose to focus on were: Average Rent, Vacancy Rate, Structure Type, Average Age (of Primary Household Maintainer), Household Type, Housing Standards, Immigrants, and Income. Each of these data sets are in tabular format and were filtered to include only renters and rows were organized by their survey zone. The tables Average Rent, Vacancy Rate and Structure Type is data on the full population while the other 5 is data for core

housing needs. With 8 tables, for 3 cities, for 2 census years, we have a total of 48 tables. Peter was responsible for looking at the Toronto data sets, Joel was responsible for looking at the Vancouver data sets and Janny was responsible for looking at the Calgary data sets.

One particular challenge with our data set pertains to the lack of detail about specific demographics. With the cells in each row representing the percentage of that demographic which qualifies for CHN, we were unable to ascertain information about the size of the population in that demographic. For instance, we may know that 33% of Immigrants in English Bay qualify for CHN, but we could not know the number of immigrants that live in English Bay. Alternatively, we would know the average household income for a survey zone, but would not be able to know what the average income for seniors in that survey zone is. This problem of generalization made it difficult to draw conclusions from our queries.

Data Exploration

Conveniently all of our data was imported from CMHC meaning they all had similar formatting. Since we are downloading the same datasets for 3 different cities it made data cleaning very simple. In Excel, we deleted all unnecessary information, such as the “Notes” section and row headers to make the import to Python easier. Using the Pandas library, we renamed the column “Unnamed: 0” to “Zone” as it will be our foreign key for our individual queries. Stars were used for data that were not statistically reliable or to protect confidentiality. We decided to replace these stars with numpy NA using the numpy library. Then we deleted columns that had letter codes indicating the reliability of the estimates. Finally, we inserted the Year column with its proper values in our data set. Once we have cleaned our data sets, we used the function “to_csv” to create comma-separated values files that will be ready to use in MySQL. When importing our tables into MySQL, we replaced all the NA values with NULLs. However, one issue we had when importing our data into MySQL was the formatting of the numbers in Excel that were in the thousands. Since fields were being terminated by commas, we would only get the thousands value of some data. To solve this problem, we changed the formatting of the numbers in Excel so that it did not contain commas anymore.

For our group queries, we decided to combine our datasets and include data for Alberta, British Columbia, Ontario and Canada. To our convenience, all of our data was from CMHC meaning that all data sets had the same column names. We simply used the function ‘concat’ in the pandas library to combine all of our data sets. In total, we will have 16 tables: Average Rent, Vacancy Rate, Structure Type, Average Age, Household Type, Housing Standards, Immigrants, and Income for 2011 and 2016. Due to there being a “Downtown” zone in Calgary and Vancouver, we decided to add a column called Jurisdiction that would allow us to distinguish between the downtowns. The combination of the Jurisdiction and Zone will be our composite key. Once again, we used the function “to_csv” to create 16 comma-separated values files. Within Excel, we changed the formatting of the values so that numbers in the thousands do not contain commas.

In our individual queries, we focused on the rental market in cities Calgary, Vancouver, and Toronto. We decided to do five different queries for each city, totalling to 15.

In Calgary, we learned that from 2011 to 2016, a two bedroom structure had the largest average rent increase by \$164, and the largest vacancy rate increase by 5.9%. This makes sense because, with such a large increase in average rent, fewer people are willing to pay that much for a two bedroom and leading to the largest vacancy rate increase in 2016. Next, we focus on the Survey Zones with the highest percentage of Couples with Children with core housing needs in 2011 and 2016. We learned that in 2011, Couples with Children with core housing needs were living in the Chinook Survey Zone and they mainly lived in low rise or high rise apartment buildings. However, in 2016, the Survey Zone with the highest percentage of Couples with Children with core housing needs is now the Northeast and they have transitioned to living in single-detached housing. Our next query focused on if there was a certain Survey Zone that immigrants with core housing needs were mainly in Calgary. We learned that the Survey Zone Southeast had an increase of immigration households in both 1996 and 2001. Lastly, we were curious to see if landlords were not making enough money for the upkeep of properties. We would expect that households with higher average monthly shelter costs would mean that landlords would put more effort into upkeeping the households. However, we see that for all of Calgary the average monthly shelter costs increased but so did the Below Adequacy Standard percentage. Meaning that it did not meet our expectations.

In Vancouver, we first focused on young families. By filtering the data in the 'household type' and 'income' tables, the survey zone 'English Bay' was strongly identified as being unaffordable for this demographic. With a relatively low pre-tax income (<\$30,000) and a high average monthly shelter cost (>\$1000), we can see that couples with children are likely to qualify for CHN. Interestingly, the percentage of couples with children that qualify for CHN drops from 33.8% to 21.1% between census years which could indicate that families moved to somewhere more affordable. Our next area of focus in Vancouver was areas where more than 20% of immigrants qualify for CHN, and immigrants that immigrated more than 15 years ago qualify less than more recent immigrants. Unsurprisingly, this query showed that many survey zones matched these criteria. There are multiple ways to interpret this result, but the optimistic view is that immigrants that stay long enough will be able to find suitable housing eventually. The final query of interest combined the 'structure type' table with the 'household type' table to assess the living conditions for one person households. The largely student-dominated University Endowment Land (UBC Campus) had the lowest proportion of one person households in CHN as many students resorted to living with others to afford their housing, or they live in one of many newly built student housing accommodations. East Hastings, on the other hand, had the highest percentage of one person households in CHN which is likely a result of the homelessness epidemic in that survey zone.

In Toronto, after exploring the data in our individual queries, we focused our exploration on four main aspects, senior's housing needs, housing size for families, immigration and density, and the cost of different sizes of housing. After the exploration, it was found that over the five years between the 2011 and 2016 census, the number of seniors in Toronto that require affordable housing jumps from 39.2% to 52.5 %, which is a significant amount. We also observed that zones with a high number of couples with children, along with single parents, correlate with

zones that have a higher number of residents in need of more affordable housing, up to 88% percent of the population. Zones with a higher amount of structures, density, also seem to have a higher amount of immigrants that require more affordable housing, which ties into the last research question. These zones with higher density seem to be more expensive than those that aren't, giving an idea into locations where those in need of more affordable housing can get more bang for their buck in different areas.

Group Query

The datasets we used for our group queries are our combined data from Toronto, Calgary and Vancouver. We also included data for Alberta, Ontario, British Columbia and Canada. There will be 16 final data frames: "final_age_2011", "final_age_2016", "final_householdtype_2011", "final_householdtype_2016", "final_housingstandard_2011", "final_housingstandard_2016", "final_immigrant_2011", "final_immigrant_2016", "final_income_2011", "final_income_2016", "final_rent_2011", "final_rent_2016", "final_structuretype_2011", "final_structuretype_2016", "final_vacancy_2011" and "final_vacancy_2016". These final datasets were cleaned and combined in Python and can be found under the section 'Cleaning + Combining to make our Final Dataframes' in the Jupyter Notebook.

Average Monthly Shelter Cost

Once our datasets were combined, we first wanted to take a look into the average shelter costs

```
select `Year`, `Zone`, `Average_Monthly_Shelter_Costs` from final_income_2011
where Jurisdiction = 'Vancouver' and Zone = 'Vancouver'
or Jurisdiction = 'Calgary' and Zone = 'Calgary'
or Jurisdiction = 'Toronto' and Zone = 'Toronto'
or Jurisdiction = 'Canada';
```

for each

of our respective cities, along with Canada.

The query above is stating that we will be creating a table looking at the columns Year, Zone, and Average_Monthly_Shelter_Costs. We are creating conditions for Jurisdiction and Zone so that we are only focusing on the three cities and Canada. We found that the average shelter costs in 2011 for Calgary, Toronto, Vancouver and Canada were:

- Canada - \$785
- Calgary - \$997
- Toronto - \$951
- Vancouver - \$921

We also looked at the average shelter costs in 2016 and found the following average shelter costs:

```
select `Year`, `Zone`, `Average_Monthly_Shelter_Costs` from final_income_2016
where Jurisdiction = 'Vancouver' and Zone = 'Vancouver'
or Jurisdiction = 'Calgary' and Zone = 'Calgary'
or Jurisdiction = 'Toronto' and Zone = 'Toronto'
or Jurisdiction = 'Canada';
```

- Canada - \$915
- Calgary - \$1210
- Toronto - \$1096
- Vancouver - \$1044

We were all surprised by this outcome, as we believed that Calgary would have had the lowest of the average between the three cities. One main factor that we think would contribute to the highest shelter costs in Calgary is the pricing of utilities. If we were to continue this project, we would like to explore more into the costs of electricity, fuel, and water to see how these also compare between the three cities.

Families in Need

We wanted to find the survey zones where Couples with Children that have core housing needs, we will call them Families in Need, were meeting national or provincial thresholds. These thresholds are that they are above a certain percentage of Families in Need, below a certain Average Household Income Before Taxes, and above a certain Average Monthly Shelter Cost. We were able to find the national and provincial thresholds from the datasets final_householdtype and final_income. For these queries, we are finding the Year, Jurisdiction, Zone, and information on that zone, such as percentage of Families in Need, Average Household Income Before Taxes, and Average Monthly Shelter Costs.

For Canada, we found the average percentage of Families in Need in 2011 was 21.7% and 20.3% in 2016. The average household income before taxes was \$21,242 in 2011 and \$24,775 in 2016. And the average monthly shelter costs were \$785 in 2011 and \$915 in 2016.

```
SELECT final_householdtype_2011.Year, final_householdtype_2011.Jurisdiction,
final_householdtype_2011.Zone, final_householdtype_2011.Couple_With_Children,
final_income_2011.Average_Household_Income_Before_Taxes, final_income_2011.Average_Monthly_Shelter_Costs
FROM final_householdtype_2011
INNER JOIN final_income_2011
ON final_householdtype_2011.Zone = final_income_2011.Zone
WHERE Couple_With_Children > 21.7 and Average_Household_Income_Before_Taxes < 21242 and Average_Monthly_Shelter_Costs > 785
UNION
SELECT final_householdtype_2016.Year, final_householdtype_2016.Jurisdiction,
final_householdtype_2016.Zone, final_householdtype_2016.Couple_With_Children,
final_income_2016.Average_Household_Income_Before_Taxes, final_income_2016.Average_Monthly_Shelter_Costs
FROM final_householdtype_2016
INNER JOIN final_income_2016
ON final_householdtype_2016.Zone = final_income_2016.Zone
WHERE Couple_With_Children > 20.3 and Average_Household_Income_Before_Taxes < 24775 and Average_Monthly_Shelter_Costs > 915
ORDER BY Zone;
```

From our thresholds listed above, you can see that we are finding the Survey Zones in 2011 that have Couples_With_Children that is greater than 21.7%, Average_Household_Income_Before_Taxes that is less than \$21,242 and Average_Monthly_Shelter_Costs that is above \$785. We are doing the same thing but with 2016 thresholds and using the “UNION” function to combine the data together. When we ran this query, no rows were returned meaning that there were no survey zones in 2011 and 2016 that met Canada’s thresholds.

For Alberta, we found the average percentage of Families in Need in 2011 was 19.2% and 20.2% in 2016. The average household income before taxes was \$25,155 in 2011 and \$30,991 in 2016. And the average monthly shelter costs were \$964 in 2011 and \$1150 in 2016.

```
SELECT final_householdtype_2011.Year,final_householdtype_2011.Jurisdiction,
final_householdtype_2011.Zone, final_householdtype_2011.Couple_With_Children,
final_income_2011.Average_Household_Income_Before_Taxes,final_income_2011.Average_Monthly_Shelter_Costs
FROM final_householdtype_2011
INNER JOIN final_income_2011
ON final_householdtype_2011.Zone = final_income_2011.Zone
WHERE Couple_With_Children > 19.2 and Average_Household_Income_Before_Taxes < 25155 and Average_Monthly_Shelter_Costs > 964
and final_householdtype_2011.Jurisdiction = 'Calgary'
UNION
SELECT final_householdtype_2016.Year,final_householdtype_2016.Jurisdiction,
final_householdtype_2016.Zone, final_householdtype_2016.Couple_With_Children,
final_income_2016.Average_Household_Income_Before_Taxes,final_income_2016.Average_Monthly_Shelter_Costs
FROM final_householdtype_2016
INNER JOIN final_income_2016
ON final_householdtype_2016.Zone = final_income_2016.Zone
WHERE Couple_With_Children > 20.2 and Average_Household_Income_Before_Taxes < 30991 and Average_Monthly_Shelter_Costs > 1150
and final_householdtype_2016.Jurisdiction = 'Calgary'
ORDER BY Zone;
```

From our thresholds listed above, you can see that we are finding the survey zones in 2011 that have Couples_With_Children that is greater than 19.2%, Average_Household_Income_Before_Taxes that is less than \$25,155, and Average_Monthly_Shelter_Costs that is above \$964. Since we are using Alberta thresholds, we are only concerned about survey zones in Calgary, so we include Jurisdiction = ‘Calgary’. We are doing the same thing but with 2016 thresholds and using the “UNION” function to combine the data.

	Year	Jurisdiction	Zone	Couple_With_Children	Average_Household_Income_Before_Taxes	Average_Monthly_Shelter_Costs
►	2011	Calgary	Calgary	21.5	25056	997
	2016	Calgary	Chinook	27.5	30539	1171

We found that in 2011, the Calgary average was the only zone that met these thresholds. This means that the overall average of Calgary with core housing needs is above the Alberta average for the percentage of Families in Need. Compared to the Alberta average, the Calgary average with CHN are paying more for shelter and receiving less household income before taxes. This can also be said for the survey zone Chinook in 2016.

For British Columbia, we found the average percentage of Families in Need in 2011 was 23.7% and 20.3% in 2016. The average household income before taxes was \$22,946 in 2011 and \$25,854 in 2016. And the average monthly shelter costs were \$869 in 2011 and \$915 in 2016.

```
SELECT final_householdtype_2011.Year,final_householdtype_2011.Jurisdiction,
final_householdtype_2011.Zone, final_householdtype_2011.Couple_With_Children,
final_income_2011.Average_Household_Income_Before_Taxes,final_income_2011.Average_Monthly_Shelter_Costs
FROM final_householdtype_2011
INNER JOIN final_income_2011
ON final_householdtype_2011.Zone = final_income_2011.Zone
WHERE Couple_With_Children > 23.7 and Average_Household_Income_Before_Taxes < 22946 and Average_Monthly_Shelter_Costs > 869
and final_householdtype_2011.Jurisdiction = 'Vancouver'
UNION
SELECT final_householdtype_2016.Year,final_householdtype_2016.Jurisdiction,
final_householdtype_2016.Zone, final_householdtype_2016.Couple_With_Children,
final_income_2016.Average_Household_Income_Before_Taxes,final_income_2016.Average_Monthly_Shelter_Costs
FROM final_householdtype_2016
INNER JOIN final_income_2016
ON final_householdtype_2016.Zone = final_income_2016.Zone
WHERE Couple_With_Children > 20.3 and Average_Household_Income_Before_Taxes < 25854 and Average_Monthly_Shelter_Costs > 915
and final_householdtype_2016.Jurisdiction = 'Vancouver'
ORDER BY Zone;
```

From our thresholds listed above, you can see that we are finding the survey zones in 2011 that have Couples_With_Children that is greater than 23.7%, Average_Household_Income_Before_Taxes that is less than \$22,946, and Average_Monthly_Shelter_Costs that is above \$869. Since we are using British Columbia thresholds, we are only concerned about survey zones in Vancouver, so we include Jurisdiction = 'Vancouver'. We are doing the same thing but with 2016 thresholds and using the "UNION" function to combine the data.

	Year	Jurisdiction	Zone	Couple_With_Children	Average_Household_Income_Before_Taxes	Average_Monthly_Shelter_Costs
►	2016	Vancouver	White Rock	20.7	25315	1058

The only zone in Vancouver that met our thresholds was White Rock in 2016. This survey zone has a percentage of Families in Need that is greater than the British Columbia average. They are receiving an average household income before taxes that is below the provincial average while paying an average monthly shelter cost that is above the provincial average.

For Ontario, we found the average percentage of Families in Need in 2011 was 26.9% and 28.1% in 2016. The average household income before taxes was \$23,584 in 2011 and \$26,871 in 2016. And the average monthly shelter costs were \$864 in 2011 and \$988 in 2016.

```

SELECT final_householdtype_2011.Year, final_householdtype_2011.Jurisdiction,
final_householdtype_2011.Zone, final_householdtype_2011.Couple_With_Children,
final_income_2011.Average_Household_Income_Before_Taxes,final_income_2011.Average_Monthly_Shelter_Costs
FROM final_householdtype_2011
INNER JOIN final_income_2011
ON final_householdtype_2011.Zone = final_income_2011.Zone
WHERE Couple_With_Children > 26.9 AND Average_Household_Income_Before_Taxes < 23584 AND Average_Monthly_Shelter_Costs > 864
AND final_householdtype_2011.Jurisdiction = 'Toronto'
UNION
SELECT final_householdtype_2016.Year, final_householdtype_2016.Jurisdiction,
final_householdtype_2016.Zone, final_householdtype_2016.Couple_With_Children,
final_income_2016.Average_Household_Income_Before_Taxes,final_income_2016.Average_Monthly_Shelter_Costs
FROM final_householdtype_2016
INNER JOIN final_income_2016
ON final_householdtype_2016.Zone = final_income_2016.Zone
WHERE Couple_With_Children > 28.1 AND Average_Household_Income_Before_Taxes < 26871 AND Average_Monthly_Shelter_Costs > 988
AND final_householdtype_2016.Jurisdiction = 'Toronto'
ORDER BY Zone;

```

From our thresholds listed above, you can see that we are finding the survey zones in 2011 that have Couples_With_Children that is greater than 26.9%, Average_Household_Income_Before_Taxes that is less than \$23,584, and Average_Monthly_Shelter_Costs that is above \$864. Since we are using Ontario thresholds, we are only concerned about Survey Zones in Toronto, so we include Jurisdiction = 'Toronto'. We are doing the same thing but with 2016 thresholds and using the "UNION" function to combine the data.

Year	Jurisdiction	Zone	Couple_With_Children	Average_Household_Income_Before_T	Average_Monthly_Shelter_Costs
2016	Toronto	Toronto (Central)	30.7	26760	1041

The only survey zone that follows the Ontario thresholds is Toronto (Central) in 2016. That means that the percentage of Families in Need in Toronto (Central) is higher than the provincial average. People in Central Toronto are receiving less than the provincial average for average household income before taxes and they are paying above the provincial average for average monthly shelter costs.

The next query was finding the highest percentage of Families in Need in Canada and seeing if their preference in housing changed from 2011 to 2016. For this query, we will be finding the maximum value of Couple_With_Children on line 73. Then we will be joining the household type table with the maximum value of Couple_With_Children on the column Couple_With_Children. This will find all the household type information on the Zone that has the highest value of Couple_With_Children in Canada. We will do another inner join with this table and structure_type on the column Survey_Zone to get all the structure type data for the Zone with the highest percentage of Couple_With_Children. We will do the same thing for 2016 and use the "UNION" function to combine the data.


```

72 • SELECT final_structurertype_2016.*, survey_max_info.Couple_With_Children as max_percentage FROM
73 (SELECT final_householdtype_2016.* FROM final_householdtype_2016
74 INNER JOIN
75 (select max(Couple_With_Children) as max_family FROM final_householdtype_2016) max_info
76 on final_householdtype_2016.Couple_With_Children = max_info.max_family) survey_max_info
77 inner join final_structurertype_2016
78 ON survey_max_info.Zone = final_structurertype_2016.Zone
79 UNION
80 SELECT final_structurertype_2011.*, survey_max_info.Couple_With_Children as max_percentage FROM
81 (SELECT final_householdtype_2011.* FROM final_householdtype_2011
82 INNER JOIN
83 (SELECT max(Couple_With_Children) as max_family FROM final_householdtype_2011) max_info
84 ON final_householdtype_2011.Couple_With_Children = max_info.max_family) survey_max_info
85 INNER JOIN final_structurertype_2011
86 ON survey_max_info.Zone = final_structurertype_2011.Zone;

```

	Year	Jurisdiction	Zone	Single_Detached	Semi_Detached	Row	Duplex	Low_Rise_Apt.	High_Rise_Apt.	Other	Total	max_percentage
▶	2016	Toronto	Scarborough (East)	1480	1210	95	1460	1440	15525	20	21235	48.3
	2011	Toronto	East York	985	240	120	570	5070	16635	0	23630	47.6

From this query, we can see that in 2011, East York had the highest percentage of Families in Need in all of Canada at 47.6%. The majority of the people that lived in East York were living in high rise apartments leading us to believe that lots of Families in Needs were living in these apartments. In 2016, we see that Scarborough (East) is now the highest percentage of Families in Need in all of Canada at 48.3%. Going from 2011 to 2016, fewer people are living in low rise apartments and are living in either single detached or semi detached houses. There is still a large number of people living in high rise apartment buildings.

Seniors

For our demographic of Senior citizens in core housing need, we wanted to see how our three cities, as well as provinces and Canada as a whole, had changed from 2011 to 2016. This finding would lead to seeing how the province's assistant programs which are made specifically for seniors help with housing needs. In Canada, on average, the annual amount of income an individual receives from the Government of Canada is between \$23,000 to \$39,000. For the older population that does not have any savings, a large majority of one's monthly pay would go towards housing.

```

select final_age_2011.Jurisdiction, final_age_2011.Zone, final_age_2011.`Total_65+`, final_age_2016.`Total_65+`,
ROUND(final_age_2016.`Total_65+` - final_age_2011.`Total_65+`,2) as difference
from
final_age_2011 inner join final_age_2016 on final_age_2011.Zone = final_age_2016.Zone;

```

With this query, we first pull up the age bracket of seniors in need for both 2011 and then 2016 in our zones. We then add a new column into our table that shows the difference between 2016 and 2011. If the difference is positive, it would suggest an increase in the amount of seniors requiring housing needs, and a decrease if it is negative.

Jurisdiction	Zone	Total_65+	Total_65+	difference
Calgary	Calgary	44.6	49.3	4.70
Vancouver	Vancouver	20.2	19.7	-0.50
Toronto	Toronto	39.1	52.5	13.40
Ontario	Ontario	30.6	43.1	12.50
Alberta	Alberta	41.9	45.3	3.40
British Columbia	British Columbia	41.9	42.9	1.00
Canada	Canada	28.9	32.8	3.90

From this query, we can see that Ontario, and specifically Toronto had the highest increase in Seniors in core housing need, while Vancouver had a lower amount of seniors in housing need, and British Columbia had an overall increase of 1 percentage point. Calgary and Alberta seemed to follow the national trend with an increase of 4.70 and 3.4 percentage points respectively.

Jurisdiction	Zone	Total_65+	Total_65+	difference
Vancouver	South Granville/Oak	26.2	21.1	-5.10
Vancouver	North Burnaby	22.8	19.9	-2.90
Vancouver	Langley City and Langley DM	18.5	16.1	-2.40
Vancouver	Maple Ridge/Pitt Meadows	18.4	16.3	-2.10
Vancouver	West End/Stanley Park	31.6	29.6	-2.00
Vancouver	Southeast Burnaby	24	22.2	-1.80
Vancouver	Delta	13.8	12	-1.80
Vancouver	Kitsilano/Point Grey	20.2	18.5	-1.70
Vancouver	Southeast Vancouver	22.9	21.2	-1.70
Vancouver	North Vancouver DM	14.7	13	-1.70
Calgary	Downtown Calgary	62.7	61.5	-1.20

Looking into the Zones that had the lowest amount of increase in the population of Seniors in housing need, we can see that many Vancouver Zones did have a decrease over the last five years, including Downtown Calgary. Looking into the assistance programs for seniors in British Columbia provides many grants for housing improvements for seniors to assist with independent living.

Jurisdiction	Zone	Total_65+	Total_65+	difference
Toronto	Toronto (East)	37.5	61.6	24.10
Toronto	Caledon	31.4	52.2	20.80
Toronto	North York (N.Central)	41.4	61.5	20.10
Toronto	Milton/Halton Hills	31.2	51.1	19.90
Toronto	Aurora, Newmkt, Whit-St.	36.4	56.2	19.80
Toronto	Pickering/Ajax/Uxbridge	33.7	52.7	19.00
Toronto	York	41.1	59	17.90
Toronto	Toronto (Central)	31.5	49.1	17.60
Toronto	North York (Northeast)	33	48.7	15.70
Toronto	North York (Northwest)	41.6	56.6	15.00
Toronto	Scarborough (North)	43.3	57.4	14.10

On the opposite side of the country, we can see a significant increase in the amount of Seniors in need of housing specifically in Toronto. We observed many Zones in Toronto where the increase was in the double-digit, where in Toronto(East), the percentage of the Seniors in housing need almost doubled from 2011 to 2016. Toronto does have its own subsidy program called Rent-Geared-To-Income which provides subsidies to those in need. This was introduced in 2007, where there are currently the following wait times:

- 7 years or more for a bachelor unit
- 12 years for a one-bedroom unit
- 10 years for larger units

We think that due to the number of years it takes for this program to finally apply to those individuals in need, that we may not be able to see the overall trend of the program within the 5 years.

Immigrants

With Canada being a hotspot for immigration, it is important to ensure that Immigrants have access to adequate housing. The first step to solving the problem is identifying the problem itself. In order to quantify an immigrant's struggle to find adequate housing, we defined a new metric called the 'Immigrant Differential'. The immigrant differential calculates the percentage difference for those qualifying for CHN based on the resident's immigration status. The immigrant differential is calculated by taking the percentage of immigrants qualified for CHN minus the percentage of non-immigrants qualified for CHN. A positive immigrant differential would indicate that immigrants in a given survey zone have a harder time finding adequate housing than their non-immigrant neighbors.

```
SELECT final_immigrant_2011.Year, final_immigrant_2011.Jurisdiction, final_immigrant_2011.Zone,
final_immigrant_2011.Immigrant, final_immigrant_2011.Non_Immigrant,
ROUND(final_immigrant_2011.Immigrant - final_immigrant_2011.Non_Immigrant, 2) as Immigrant_diff
FROM final_immigrant_2011
UNION
SELECT final_immigrant_2016.Year, final_immigrant_2016.Jurisdiction, final_immigrant_2016.Zone,
final_immigrant_2016.Immigrant, final_immigrant_2016.Non_Immigrant,
ROUND(final_immigrant_2016.Immigrant - final_immigrant_2016.Non_Immigrant, 2) as Immigrant_diff
FROM final_immigrant_2016;
```

The query shown above will return a list of all survey zones in the data set, the immigrant and non-immigrant columns, and the calculation for the immigrant differential. Using the combined dataset with all three cities as well as provincial and national jurisdictions, we will be able to assess how immigrants compare to non-immigrants in terms of CHN.

	Year	Zone	▲	Immigrant	Non_Immigrant	Immigrant_diff
►	2011	Alberta		27.8	22.2	5.60
	2016	Alberta		28.3	23.6	4.70
	2011	British Columbia		35.1	30	5.10
	2016	British Columbia		33.3	28.7	4.60
	2011	Canada		32.8	24.4	8.40
	2016	Canada		33.2	24.7	8.50
	2011	Ontario		35	26.7	8.30
	2016	Ontario		40	30.1	9.90

The results for national and provincial jurisdictions confirm our expectation that immigrants are struggling to find adequate housing. The national average sits at around +8.5% for both census years which indicates that the average difference in CHN for immigrants and non-immigrants for all survey zones in Canada is 8.5%.

Year	Jurisdiction	Zone	▲	Immigrant	Non_Immigrant	Immigrant_diff
2011	Toronto	Orangeville/Mono		51.6	28.5	23.10
2016	Toronto	Orangeville/Mono		39.3	40.9	-1.60

Perhaps the most interesting survey zone is Toronto's Orangeville/Mono. This survey zone features the highest absolute immigrant differential at +23.1% before its precipitous drop 5 years later to join just 3 other survey zones with a negative differential. This large change is only possible because both demographics saw major changes between censuses which begs the question of what caused this to happen.

Year	Jurisdiction	Zone	▲	Immigrant	Non_Immigrant	Immigrant_diff
2011	Vancouver	Langley City and Langley...		28.5	36.2	-7.70
2016	Vancouver	Langley City and Langley...		35.6	32.3	3.30

In the reverse scenario, Langley City just outside of Vancouver was the only survey zone to change from a negative differential to a positive differential.

Year	Zone	▲	Immigrant	Non_Immigrant	Immigrant_diff
2011	Calgary		27.6	20.2	7.40
2016	Calgary		30	22.7	7.30
2011	Toronto		37.5	25.8	11.70
2016	Toronto		42.5	28.9	13.60
2011	Vancouver		36	29	7.00
2016	Vancouver		34.5	28.2	6.30

Finally, we can see that amongst our cities of interest, Toronto has the maximum immigrant differential of 13.6 followed closely by Calgary and Vancouver.

Discussion

Joel

In my investigation of the rental market of Vancouver, I was mostly unsurprised to see my preconceived notions about the city confirmed. The more central areas of the Greater Vancouver Area had higher rents and worse housing standards because landlords can make more money by renting sub-standard apartments than by renovating or rebuilding. Part of Vancouver's challenge is that investors will buy investment properties in the booming housing market but will leave the unit empty, thus driving down the supply of homes on the market. In a future investigation of Vancouver's housing market, I would love to include data from the 2021 census to see how Vancouver's 2017 'Vacancy Tax' affects the rental market.

The addition of the 'Immigrant Differential' proved to be a useful metric in addressing the disparity in housing needs between immigrants and non-immigrants. The trend for both the national average and most survey zones was that the immigrant differential was positive. This indicates that a higher proportion of immigrants are qualified for core housing needs compared to their non-immigrant neighbours. This result is not surprising as non-immigrants do not face the same struggles of adapting to a new country but it raises the question of how to improve the housing market for vulnerable populations. In future inquiry, I would like to have more data about the population size for each of these groups so that we can be more sure about the conclusions we make.

Janny

My role in this project was to do all the data cleaning and data importing into MySQL. What I learned was that it is a lot more convenient to combine data when it is all from the same source. Instead of having to change the formatting of each individual dataset, I simply had to just concatenate the datasets together. For all data that had stars, I replaced them with numpy NA's thinking that it was essential when importing the csv files to MySQL. However, I realized that it was not necessary and I could have simply just replaced the stars with a string 'NA'. I also ran into the difficulty of importing tables into MySQL that used commas to separate the thousands of values. Even though this could have been done in Python, we simply changed the formatting of our data in Excel so that it did not include it. For next time, I would change the formatting of these values in Python just so that we do not have to do an extra step in Excel as that gives us more potential to damage our data.

My focus for the group queries were Families in Need. We learned that the average monthly shelter costs for Toronto were in between Calgary and Vancouver for both 2011 and 2016. Initially, I thought that because Calgary had the highest average monthly shelter cost, they would have the highest percentage of Families in Need. We were able to see that in 2011, East York in Toronto had the highest percentage of Families in Need at 47.6% and in 2016, Scarborough (East) had the highest percentage of Families in Need at 48.3%. It shows us that even though Toronto does not have the highest average monthly shelter cost, it is the city that does not support Families in Need the most.

I was able to find articles about Calgary facing impacts from both COVID-19 and oil price crashes. With Calgary's economy taking such a big hit and job loss becoming more frequent, it would be interesting to see how the rental market changes for demographics with core housing needs.

Peter

I was surprised about the findings we have with the senior citizen population needing housing in the three cities we have chosen. Knowing how the housing market is in Toronto and Vancouver, I was interested to see how contrasted the differences between Toronto and Vancouver are when it comes to the percentage of seniors in need of housing. In Toronto, there was a 12.5 point increase in the number of seniors in need of better housing, as opposed to Vancouver, where it was only 1 point, and most zones were showing a decline in seniors in need. In researching why the polarized differences between the two cities, it would be interesting to see the impact that these assistance programs have now for the senior demographic.

I also found a lot of information regarding the presence of Airbnb in the Toronto housing market and would love to explore to see the full effect of these policies regarding short term rentals. Airbnb has taken thousands of condos off the market in Toronto's downtown waterfront communities, and I would think that adding these condos back into the market place would have a significant impact on the rental prices in the area.

With the use of SQL, I felt that the exploration of data sets feels limited to what we can do with Python or R, however, it felt like less of a process when it came to the group queries. Once we loaded our datasets into SQL, it felt like our team could easily come together on queries as opposed to working with Python or R. The syntax in SQL also seems pretty more intuitive than it does in other languages, however, it does not seem to have the capabilities to come up with visualizations, which is nice when presenting interesting information.

Conclusion

In summary of our project on Core Housing Needs in three major cities, we believe that we did find some interesting data points that did surprise us in the end. Assuming that we have only heard how hot the Vancouver and Toronto housing market is, the most interesting finding we observed was that Calgary had the highest average monthly shelter cost between the three cities.

From our individual queries, we were able to have some insight into the housing needs of different demographics in each of the cities we have chosen. In Calgary, we were able to see the transition of families in core housing needing transition from low/high rise buildings in the Chinook Zone, into the NorthEast zone and single detached homes. Vancouver's University Endowment Land (UBC Campus) has the lowest single person dwellings with housing need, which would correlate with student housing provided by the University. Lastly, in Toronto, we had noticed that almost 90% of the population needs more affordable housing, which begs the question if the CMHC's limit of 30% of pre-taxed household income is too low of a threshold for this indicator.

Pivoting from our Individual queries to our group queries, we shifted focus from our individual cities and focused on demographics. For Families in Need, we found that using the provincial averages for income, shelter cost, and families in core housing need, each city had a zone in which these provincial averages thresholds were not met. For seniors, government assistant programs shape the vast difference in the number of seniors in housing need between Vancouver and Toronto, where Vancouver seems to be improving the housing conditions in the rental market for seniors, whereas in Toronto, it is getting much worse. Lastly, we see that immigrants have a more difficult time finding adequate housing than their non-immigrant neighbors, where on a national average, there is about an 8.5% point differential between the two groups.

As it becomes increasingly more expensive to live in these three cities, it will be interesting to see where the core housing needs of these groups will be next year, when the 2021 Census will provide an update to these statistics. Also, how will COVID-19 shape the rental market in the coming months? With a drastic change to the economy, it will be interesting to see what programs will be most effective in assisting those in need of adequate housing.

References

Canada, S. (2020, September 13). Government of Canada - CPP retirement pension: How much you could receive. Retrieved December 10, 2020, from <https://www.canada.ca/en/services/benefits/publicpensions/cpp/cpp-benefit/amount.html>

City of Toronto. (2020, October 15). Rent-Geared-to-Income Subsidy. Retrieved December 10, 2020, from <https://www.toronto.ca/community-people/employment-social-support/housing-support/rent-geared-to-income-subsidy/>

Office of Housing and Construction Standards. (2018, September 06). Seniors Housing. Retrieved December 10, 2020, from <https://www2.gov.bc.ca/gov/content/housing-tenancy/seniors-housing>

CMHC Licence Agreement for the Use of Data. (2018, May 30). Retrieved November 04, 2020, from <https://www.cmhc-schl.gc.ca/en/data-and-research/cmhc-licence-agreement-use-of-data>

Understanding Core Housing Need. (2019, August 14). Retrieved November 22, 2020, from <https://www.cmhc-schl.gc.ca/en/data-and-research/core-housing-need>