

Additional Census Data Analysis

November 26, 2024

```
[1]: # import all required libraries and file downloaded after completing Part 2 -  
      ↪Toronto Demographic Data  
import pandas as pd  
df = pd.read_csv('Central_Toronto_Census_Data.csv')
```

```
[2]: # see shape of file  
df.shape
```

```
[2]: (9, 14)
```

```
[3]: # get list of all the columns in file  
df.columns
```

```
[3]: Index(['Postal Code', '15 to 19 years - Counts - Total',  
          '20 to 24 years - Counts - Total', '25 to 29 years - Counts - Total',  
          'Between 7 a.m. and 7:59 a.m. - Counts - Total',  
          'Between 8 a.m. and 8:59 a.m. - Counts - Total',  
          'Bicycle - Counts - Total', 'Population, 2021 - Counts - Total',  
          'Public transit - Counts - Total',  
          'Total - Main mode of commuting - Counts - Total',  
          'Total - Place of work status - Counts - Total',  
          'Total - Time leaving for work - Counts - Total',  
          'Walked - Counts - Total', 'Worked at home - Counts - Total'],  
          dtype='object')
```

```
[4]: # calculating percentage of people that leave early in the morning to  
      ↪understand our target of working professionals  
df["percentage_leaving_early_morning"] = ((df["Between 7 a.m. and 7:59 a.m. -  
      ↪Counts - Total"] + df["Between 8 a.m. and 8:59 a.m. - Counts - Total"]) /  
      ↪df["Total - Time leaving for work - Counts - Total"]*100)  
result = df[["Postal Code", "percentage_leaving_early_morning"]]  
  
print(result)
```

	Postal Code	percentage_leaving_early_morning
0	M4N	59.053834
1	M4P	52.487008
2	M4R	54.442344

3	M4S	54.366960
4	M4T	53.385417
5	M4V	54.098361
6	M5N	52.962963
7	M5P	49.640288
8	M5R	51.271534

```
[5]: # calculating percentages to understand ways of commuting in the area
df["percentage_bicycle"] = ((df["Bicycle - Counts - Total"]) / df["Total - Main_
mode of commuting - Counts - Total"]*100)
df["percentage_walked"] = ((df["Walked - Counts - Total"]) / df["Total - Main_
mode of commuting - Counts - Total"]*100)
df["percentage_transit"] = ((df["Public transit - Counts - Total"]) / df["Total_
Main mode of commuting - Counts - Total"]*100)

df["percentage_commuting"] = ((df["Bicycle - Counts - Total"] + df["Walked -
Counts - Total"]+ df["Public transit - Counts - Total"])/ df["Total - Main_
mode of commuting - Counts - Total"]*100)
df["percentage_car"] = 100 - df["percentage_commuting"]
df["number_car"] = df["Total - Main mode of commuting - Counts - Total"] -
(df["Bicycle - Counts - Total"] + df["Walked - Counts - Total"]+ df["Public
transit - Counts - Total"])
result1 = df[["Postal_
Code","percentage_bicycle","percentage_walked","percentage_transit",
"percentage_commuting","percentage_car" ]]

print(result1)
```

	Postal Code	percentage_bicycle	percentage_walked	percentage_transit \
0	M4N	1.305057	9.624796	23.001631
1	M4P	2.598367	12.249443	38.010393
2	M4R	2.079395	10.018904	31.190926
3	M4S	2.166554	11.103588	38.727150
4	M4T	3.125000	13.020833	29.166667
5	M4V	3.060109	17.267760	26.010929
6	M5N	1.728395	7.407407	25.308642
7	M5P	2.980473	9.352518	26.310380
8	M5R	7.465135	23.625923	25.266612

	percentage_commuting	percentage_car
0	33.931485	66.068515
1	52.858203	47.141797
2	43.289225	56.710775
3	51.997292	48.002708
4	45.312500	54.687500
5	46.338798	53.661202
6	34.444444	65.555556

7	38.643371	61.356629
8	56.357670	43.642330

```
[6]: # calculating percentage of population working from home to understand target
      ↪market lifestyle
df["percentage_WFH"] = ((df["Worked at home - Counts - Total"]) / df["Total -
      ↪Place of work status - Counts - Total"]*100)
result2 = df[["Postal Code", "percentage_WFH"]]

print(result2)
```

	Postal Code	percentage_WFH
0	M4N	54.178886
1	M4P	51.965994
2	M4R	56.623586
3	M4S	56.313794
4	M4T	60.486322
5	M4V	54.941860
6	M5N	46.822309
7	M5P	52.360097
8	M5R	55.149982

```
[7]: # calculate the percentage of 15-29 age group in the neighbourhood populations
df["percentage_15_to_29"] = ((df["15 to 19 years - Counts - Total"] + df["20 to
      ↪24 years - Counts - Total"]+ df["25 to 29 years - Counts - Total"]) /
      ↪df["Population, 2021 - Counts - Total"]*100)
df["sum_15_to_29"] = (df["15 to 19 years - Counts - Total"] + df["20 to 24
      ↪years - Counts - Total"]+ df["25 to 29 years - Counts - Total"])
result3 = df[["Postal Code", "sum_15_to_29", "percentage_15_to_29"]]

print(result3)
```

	Postal Code	sum_15_to_29	percentage_15_to_29
0	M4N	2730	17.000872
1	M4P	5465	21.810273
2	M4R	2285	19.187169
3	M4S	5700	18.534174
4	M4T	1545	14.953542
5	M4V	3130	16.240336
6	M5N	2990	18.509348
7	M5P	3670	18.543783
8	M5R	5445	20.784823

```
[8]: # export to CSV for graph making
df.to_csv("Extra_Data_Analysis.csv", index=False)
```