

# Demography of Education in Toronto's neighbourhoods with Highest and Lowest Income

23/01/2020

## Abstract

what was done, what was found, and why this matters. Descriptive, no prediction

## Wellbeing Toronto - Demographics: NHS Indicators

The data set package we delve into is *Wellbeing Toronto - Demographics: NHS Indicators*, which provides information from the 2011 National Household Survey (NHS) and presents a profile of Toronto. Data is broken down by the 140 neighbourhoods that make up the City of Toronto. With this data set package, we first identify the two neighbourhoods that have the highest income and the lowest income and then analyze the education demographics in these two neighbourhoods to understand difference of education attainment between these two neighbourhood. We also identify the median household income and the average education levels of the city to be compared with the two neighbourhoods. There are two data sets in this package used in the analysis: *Demographics NHS Indicators 2011* and *Education NHS Indicators 2011*. Additionally, we incorporate the *Wellbeing Toronto - Housing* data set solely for extracting the names of the neighbourhoods; no housing indicators are used in our analysis.

## Demographics NHS Indicators 2011 data set

*Demographics NHS Indicators 2011* shows the profile of each neighbourhood regarding population breakdown by ethnicity, household income, education level, employment, age and sex. The indicators we are interested in are *Median household total income* and *Median after-tax household income*. These two columns show the median household income of every neighbourhood before and after tax in Canadian dollars.

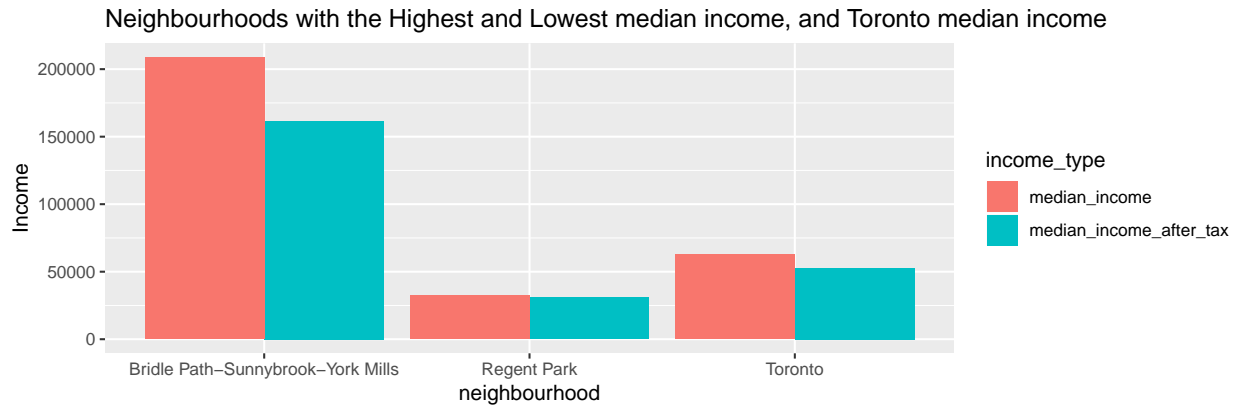
From this data set, we identify the neighbourhood with the highest median household income (Bridle Path-Sunnybrook-York Mills) and the neighbourhood with the lowest median household income (Regent Park). We also calculate the median household income of all neighbourhoods in the city of Toronto as summary statistics to be compared with the two neighbourhoods, appended to the last row.

Table 1: Median household income before and after tax.

neighbourhood	median_income_after_tax	median_income
Bridle Path-Sunnybrook-York Mills	161448	208674
Regent Park	30794	32172
Toronto	52660	62730
### Median Total Household Income Gr	aph	

```
#Plotting the graph
ggplot(graph1_df_resshaped, aes(x = neighbourhood, y = as.numeric(income_value), fill = income_type)) +
  # Specify that we want a bar graph:
  geom_bar(stat="identity", position=position_dodge()) +
```

```
# Add titles and axis labels:
labs(title =
      "Neighbourhoods with the Highest and Lowest median income, and Toronto median income",
      x = "neighbourhood",
      y = "Income")
```



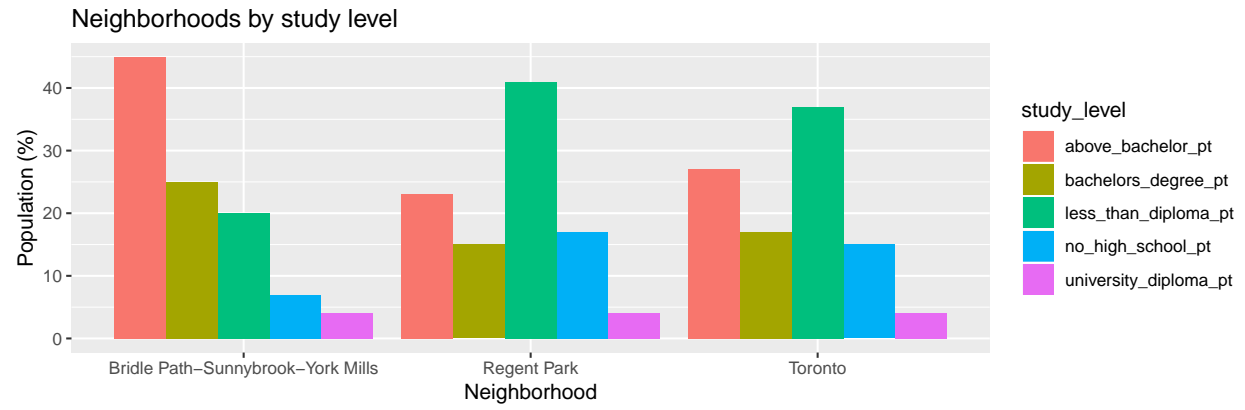
## Education NHS Indicators 2011

*Education NHS Indicators* is provided for the population of age 15 and over living in private households. The data set presents the distribution of people's most advanced education level in each neighbourhood. Because the population in each neighbourhood is different, we can not compare the percentage of education attainment by the absolute values provided in the data set. Therefore, we transform the data to relative values. The transformation is done by aggregating the number of population in all education levels and dividing it by the number in each education level.

neighbourhood	no_high_school_pt	above_bachelor_pt	university_diploma_pt
Bridle Path-Sunnybrook-York Mills	7	45	4
Regent Park	17	23	4
Toronto	15	27	4

## Education Levels Percentage Graph

```
#Plotting the graph
ggplot(
  graph2_df_percentage_long, aes(x = neighbourhood, y = value,
                                fill = study_level)) +
  # Specify that we want a bar graph:
  geom_bar(stat="identity", position=position_dodge()) +
  # Add titles and axis labels:
  labs(title = "Neighborhoods by study level",
       x = "Neighborhood",
       y = "Population (%)")
```



Authors: Ke-Li Chiu, Diego Mamanche Castellanos