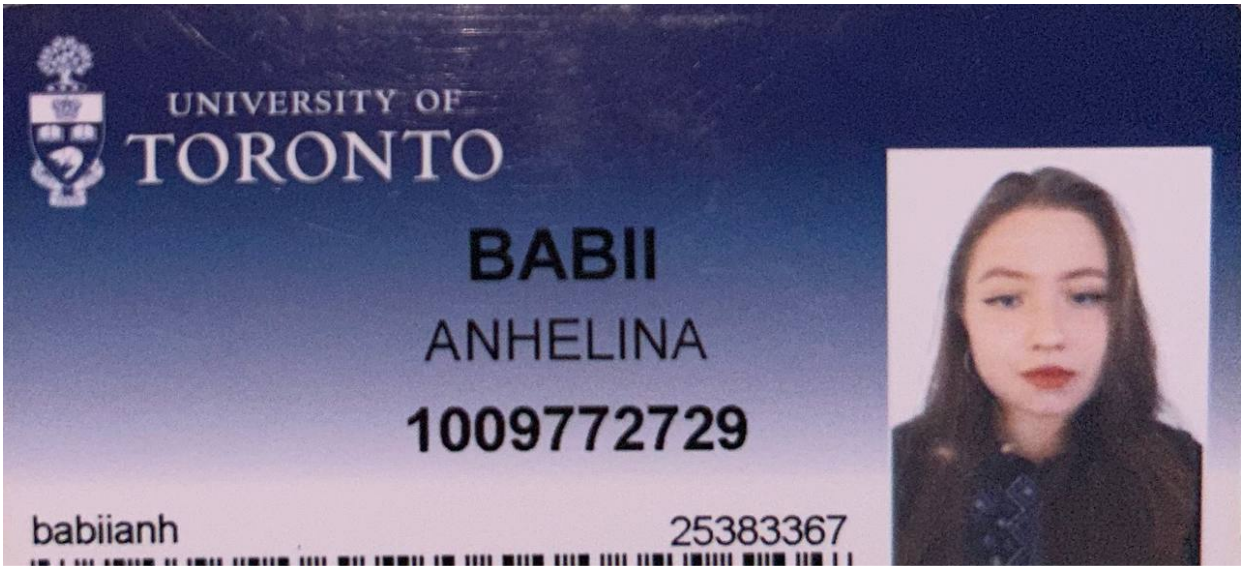
Video URL: https://play.library.utoronto.ca/watch/263924d9c4f71c67f90f4996d638c4af (https://play.library.utoronto.ca/watch/263924d9c4f71c67f90f4996d638c4af)

In [10]:
```
from IPython.display import Image
Image(filename='photo_5307980885640989520_y.jpg')
```

Out[10]:



How asthma percentage and mental health annual visits are connected to sociodemographic factors in the neighbourhoods of Toronto?

Name: Anhelina Babii

Tutorial: TUT601

Introduction

- organise the obtained information, vizualize it properly and analyze using the statistical instruments;
- databases used:
  - census_2016_Income_neighb_LHIN.xlsx,
  - socdem_2016_LangAtHome_Neighb_LHIN.xlsx,
  - ChildrenAndYouth_neighb_2015_LHIN.xls;
- all the information is represented as maps as the data grouped by Toronto neighbourhoods.

Methods

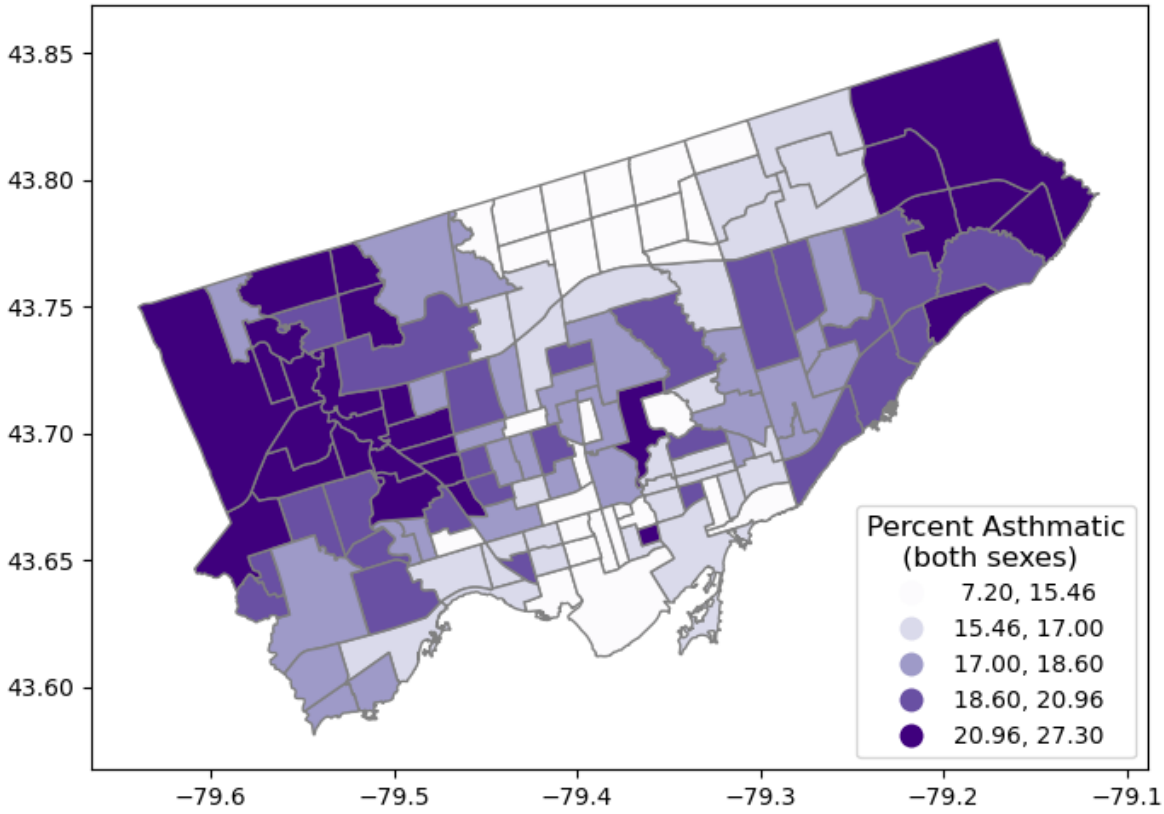
- choosing the data needed for the research by creating new data frames that have only the necessary information;
- linking the chosen data with map data of Toronto neighbourhoods;
- representing the information in map plots for the best visualization;
- calculating p-value to statistically support the obtained plots.

Results

In [8]:

```
import numpy as np
import pandas as pd
import geopandas as gpd
import matplotlib.pyplot as plt
import mapclassify
import esda
import splot
import libpysal as lps
import contextily as cx
import xlrd
from scipy import stats

# MAP
nbrhd = gpd.GeoDataFrame.from_file("Neighbourhoods - historical 140.geojson")

important_spat_cols = nbrhd.columns[[6, 5, 11]]
colnames_spat = {important_spat_cols[0]: 'name',
           important_spat_cols[1] : 'nbrhd_spat_id',
           important_spat_cols[2] : 'geometry'}

nbrhd_simple = nbrhd.copy()

nbrhd_simple = nbrhd_simple[important_spat_cols]
nbrhd_simple.rename(columns = colnames_spat, inplace=True)
nbrhd_simple["Neighbid"] = nbrhd_simple["nbrhd_spat_id"].astype(int)


# ASTHMA
fname = 'ChildrenAndYouth_neighb_2015_LHIN.xls'
sname = 'CAY_Asthma_2015'

asthma_neighb = pd.read_excel(fname, sheet_name = sname, header = 13)

important_cols = asthma_neighb.columns[[0, 8, 9, 10]]
colnames = {important_cols[0]: 'Neighbid',
            important_cols[1] : 'male_pct',
            important_cols[2] : 'female_pct',
            important_cols[3] : 'both_pct'}

asthma_rates = asthma_neighb.copy()
asthma_rates = asthma_rates[important_cols]
asthma_rates.rename(columns = colnames, inplace=True)

nbrhd_simple2 = nbrhd_simple.merge(asthma_rates, on="Neighbid")

#both sexes
fig, axes = plt.subplots(1, 1, figsize = (8,8))
nbrhd_simple2.plot(column='both_pct', scheme='quantiles',
                   k=5, cmap='Purples', edgecolor='grey',
                   ax = axes, legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Percent Asthmatic\n
                                'title_fontsize': 12,'fontsize': 10})
```

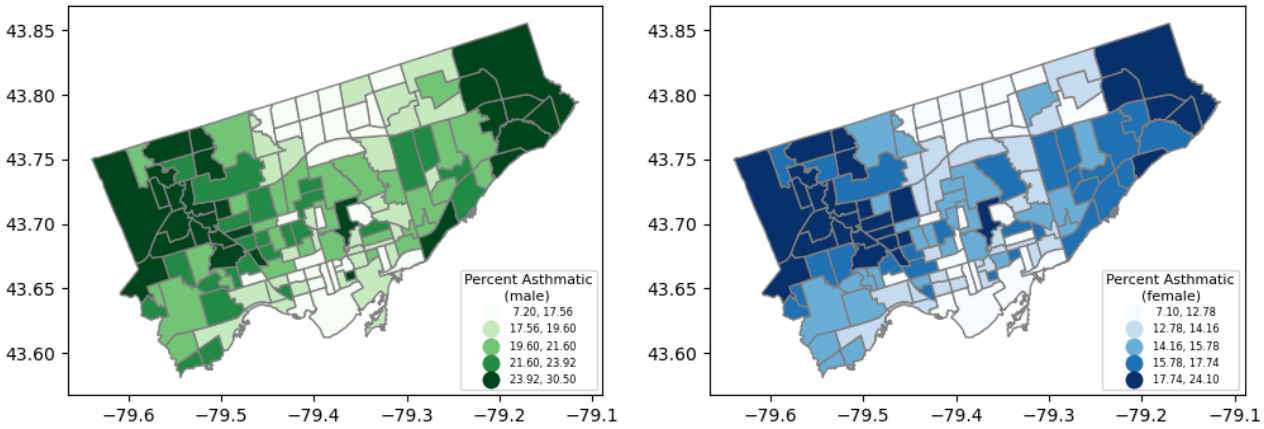
Out[8]: <AxesSubplot: >



In [2]:

```
# male
fig, axes = plt.subplots(1, 2, figsize = (12,12))
nbrhd_simple2.plot(column='male_pct', scheme='quantiles',
                   k=5, cmap='Greens', edgecolor='grey',
                   ax = axes[0], legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Percent Asthmatic\n
                                'title_fontsize': 8,'fontsize': 6})

# female
nbrhd_simple2.plot(column='female_pct', scheme='quantiles',
                   k=5, cmap='Blues', edgecolor='grey',
                   ax = axes[1], legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Percent Asthmatic\n
                                'title_fontsize': 8,'fontsize': 6})
```

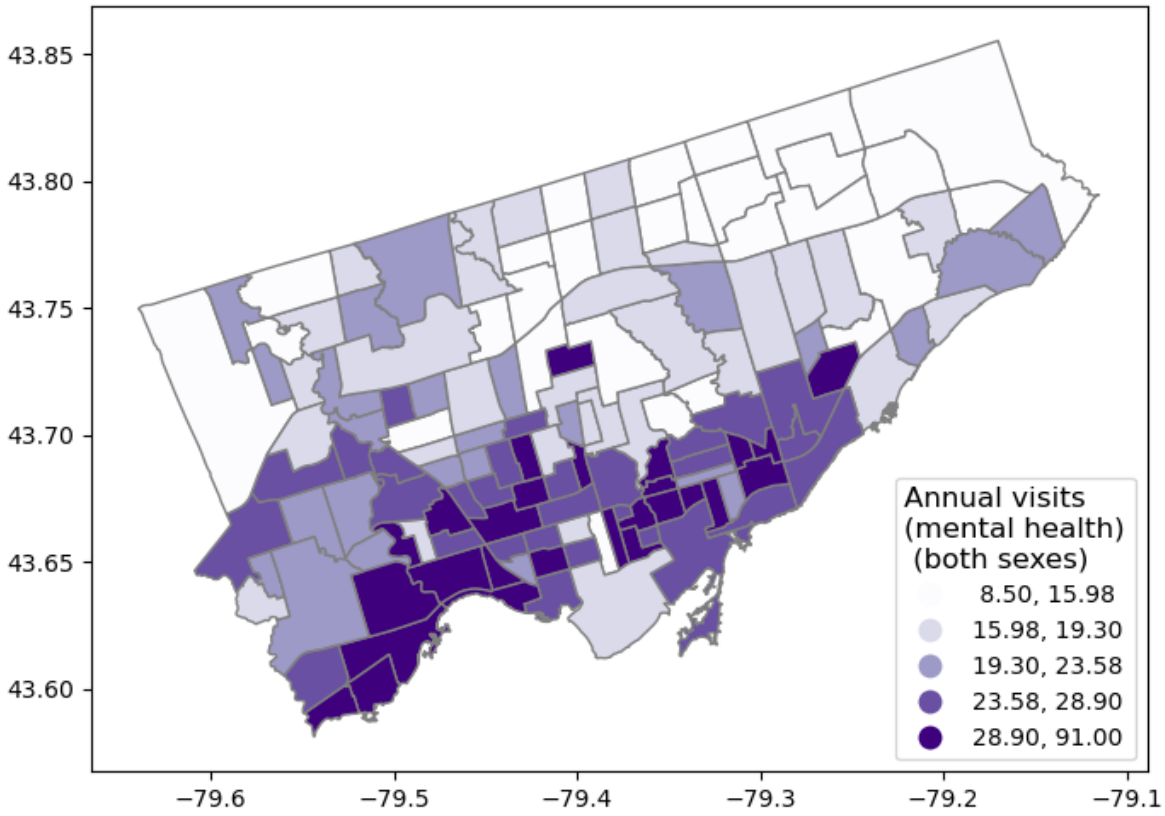
Out[2]: <AxesSubplot: >



In [3]:

```
# MENTAL HEALTH + ADDICTION
mental_neighb = pd.read_excel(fname, sheet_name = 'CAY_MentalHealth_Addiction_EDv', header = 12)

important_cols1 = mental_neighb.columns[[0, 8, 10, 11]]
colnames1 = {important_cols1[0]: 'Neighbid',
             important_cols1[1] : 'male_pct',
             important_cols1[2] : 'female_pct',
             important_cols1[3] : 'both_pct'}

mental_rates = mental_neighb.copy()
mental_rates = mental_rates[important_cols1]
mental_rates.rename(columns = colnames1, inplace=True)

nbrhd_simple3 = nbrhd_simple.merge(mental_rates, on="Neighbid")

# both sexes
fig, axes = plt.subplots(1, 1, figsize = (8,8))
nbrhd_simple3.plot(column='both_pct', scheme='quantiles',
                   k=5, cmap='Purples', edgecolor='grey',
                   ax = axes, legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Annual visits\n(mental
                                'title_fontsize': 12,'fontsize': 10})
```

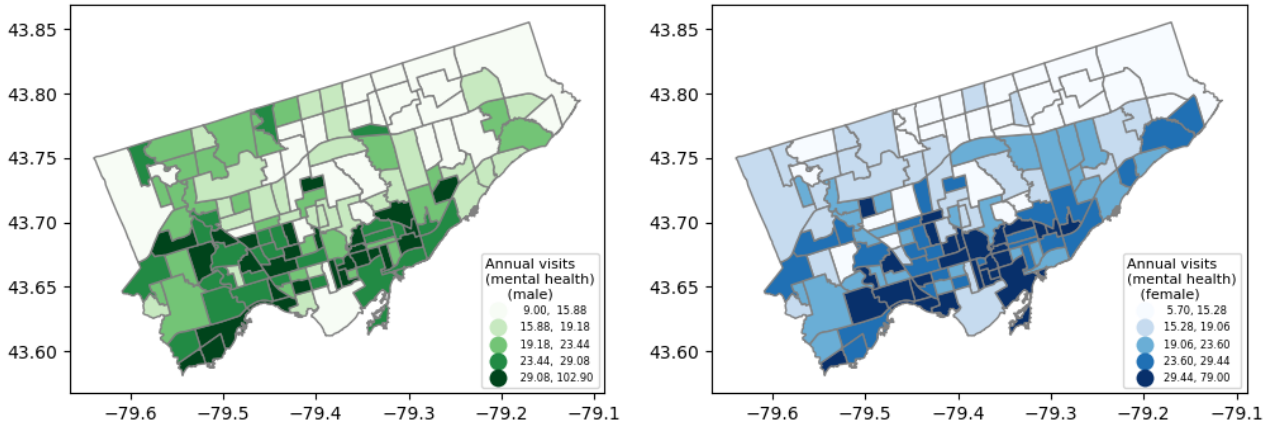
Out[3]: <AxesSubplot: >



In [4]:

```
# male
fig, axes = plt.subplots(1, 2, figsize = (12,12))
nbrhd_simple3.plot(column='male_pct', scheme='quantiles',
                   k=5, cmap='Greens', edgecolor='grey',
                   ax = axes[0], legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Annual visits\n(mental
                                'title_fontsize': 8,'fontsize': 6})

# female
nbrhd_simple3.plot(column='female_pct', scheme='quantiles',
                   k=5, cmap='Blues', edgecolor='grey',
                   ax = axes[1], legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Annual visits\n(mental
                                'title_fontsize': 8,'fontsize': 6})
```

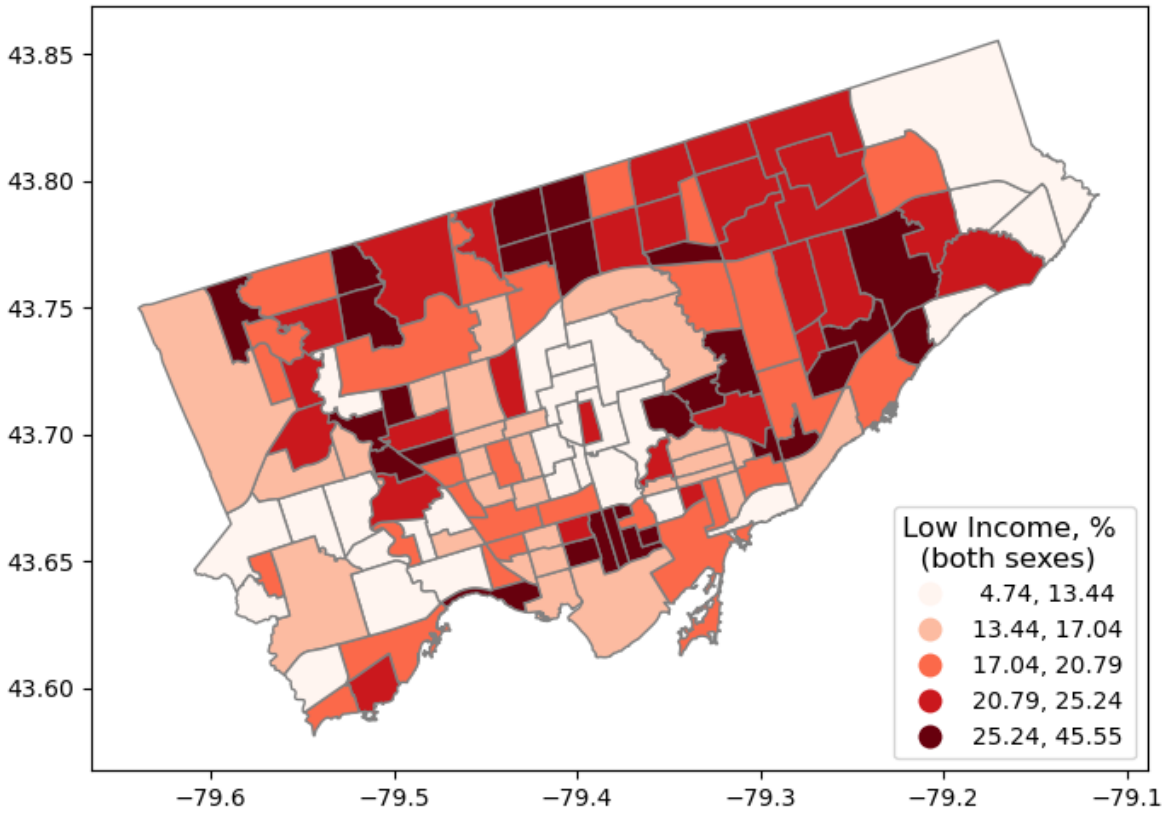
Out[4]: <AxesSubplot: >



In [5]:

```
# LOW INCOME
fnamee = 'census_2016_Income_neighb_LHIN.xlsx'
lowinc_neighb = pd.read_excel(fnamee, sheet_name = 'Income - Persons', header = 13)

important_cols2 = lowinc_neighb.columns[[0, 20]]
colnames2 = {important_cols2[0]: 'Neighbid',
             important_cols2[1] : 'both_pct'}

lowinc_rates = lowinc_neighb.copy()
lowinc_rates = lowinc_rates[important_cols2]
lowinc_rates.rename(columns = colnames2, inplace=True)

nbrhd_simple4 = nbrhd_simple.merge(lowinc_rates, on="Neighbid")

# both sexes
fig, axes = plt.subplots(1, 1, figsize = (8,8))
nbrhd_simple4.plot(column='both_pct', scheme='quantiles',
                   k=5, cmap='Reds', edgecolor='grey',
                   ax = axes, legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Low Income, % \n  (bot
                                'title_fontsize': 12,'fontsize': 10})
```

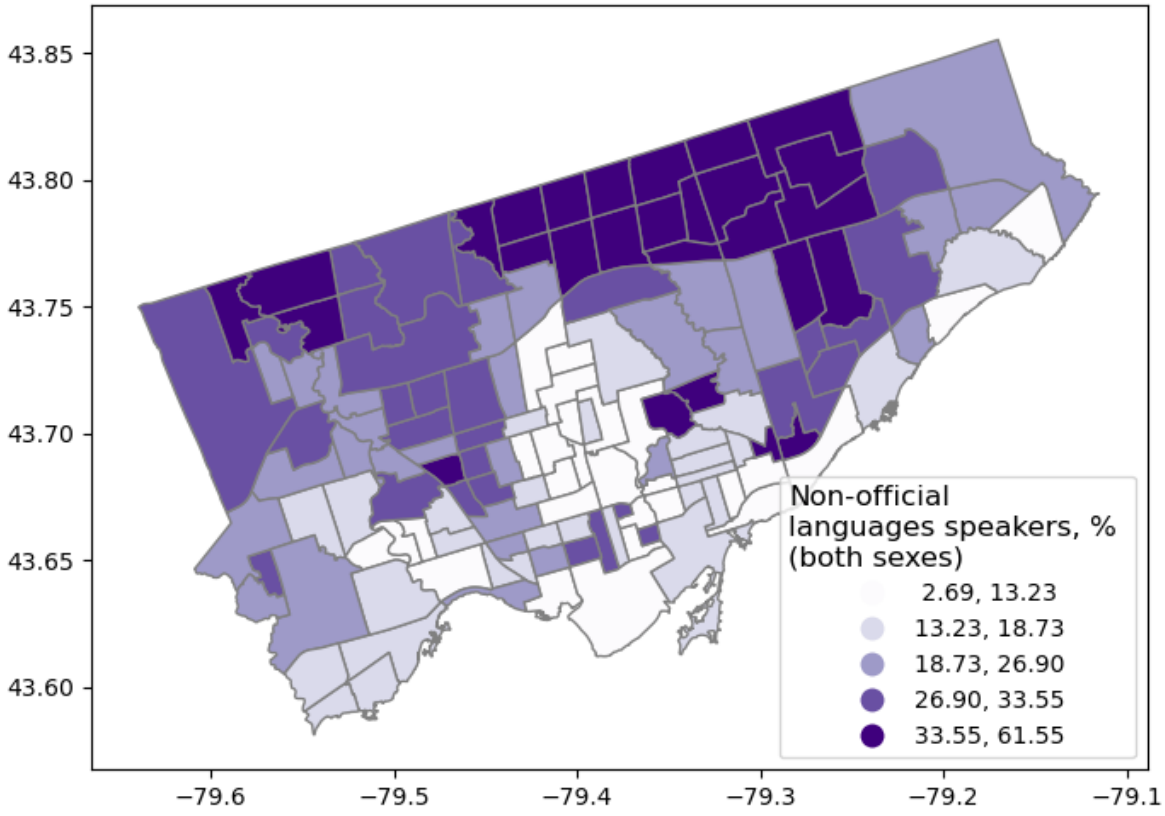
Out[5]: <AxesSubplot: >



In [6]:

```
# NON-OFFICIAL LANGUAGES
fnamee1 = 'socdem_2016_LangAtHome_Neighb_LHIN.xlsx'
lang_neighb = pd.read_excel(fnamee1, sheet_name = 'Neighb_Toronto_LangAtHome_RANK', header = 15)

important_cols3 = lang_neighb.columns[[0, 2, 10]]
colnames3 = {important_cols3[0]: 'Neighbid',
             important_cols3[1] : 'AllPop',
             important_cols3[2] : 'NonOffLangPop'}

lang_rates = lang_neighb.copy()
lang_rates = lang_rates[important_cols3]
lang_rates.rename(columns = colnames3, inplace=True)

lang_rates['Percent'] = round((lang_rates['NonOffLangPop']/lang_rates['AllPop'])*100, 2)

nbrhd_simple5 = nbrhd_simple.merge(lang_rates, on="Neighbid")

# both sexes
fig, axes = plt.subplots(1, 1, figsize = (8,8))
nbrhd_simple5.plot(column='Percent', scheme='quantiles',
                   k=5, cmap='Purples', edgecolor='grey',
                   ax = axes, legend=True,
                   legend_kwds={'loc': 4, 'bbox_to_anchor': (1, 0) ,'title': 'Non-official\nlanguage
                                'title_fontsize': 12,'fontsize': 10})
```

Out[6]: <AxesSubplot: >



In [14]:

```
# P-VALUE COMPARISON
asthma = nbrhd_simple2['both_pct']
mental_health = nbrhd_simple3['both_pct']
low_income = nbrhd_simple4['both_pct']
nonoff_lang = nbrhd_simple5['Percent']


print("Asthma x Low Income correlation: " + str(stats.ttest_ind(asthma,low_income)[1]))
print("Asthma x % of non-official language speakers correlation: " + str(stats.ttest_ind(asthma,nor
print("Mental health and addiction x Low Income correlation: " + str(stats.ttest_ind(mental_health,
print("Mental health and addiction x % of non-official language speakers correlation: " + str(stats
```

```
Asthma x Low Income correlation: 0.03444591697917437
Asthma x % of non-official language speakers correlation: 1.4146433970960525e-08
Mental health and addiction x Low Income correlation: 0.0005964228311205681
Mental health and addiction x % of non-official language speakers correlation: 0.4195521049648651
7
```

Conclusion

- The main idea of this project was to find any correlation between certain health conditions(asthma and mental health + addiction) or find out that there is none.

- It was predictable just by looking at the map plots that the correlation will not be very strong, but it indeed turned out to be very low. Only in one case a decent correlation is noticed and it can be seen very well on the plots as well (mental health and non-official languages correlation).

- There was no missing values, which is always good for any project, but the main challenge was the fact that the age groups were not all the same. Which is why, in my opinion, the results might be not very representative.