Comparing Neighborhoods of New York City and Toronto

Introduction

Given the current rate of greenhouse gas emissions (GGEs), the global average temperature might hit 2 degrees Celsius above pre-industrial levels within 15 years, a threshold that will likely cause serious harm (Mann, 2014). The transportation sector is the largest culprit of GGEs and stands for over a fourth of emissions (EPA, n.d.).

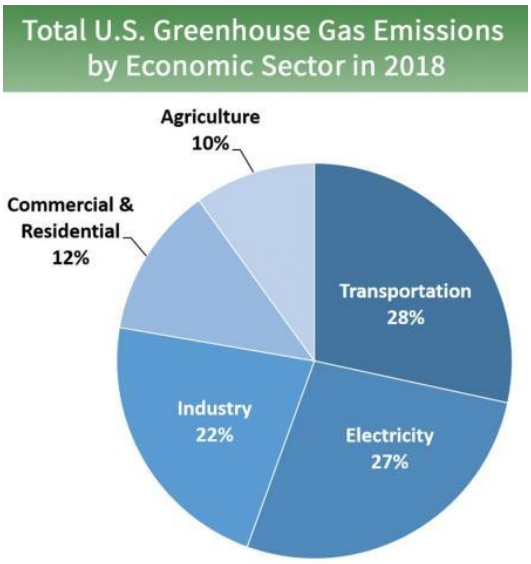

Figure 1: Pie chart of US GGEs by sector (EPA, n.d.).

A place for improvement could be business-related travels, for instance real estate scouting. Is there a way to assess a neighborhood in one city that is analogous to a neighborhood in another city? And if so, to which accuracy? Building such a model could be useful to real estate investors who cannot or does not want to travel but would still like to assess the neighborhood and its characteristics.

Data

I will use the New York data we used in week 3 and the Toronto dataset in week 4. Then I will use Foursquare to get data on nearby venues. Features will include different venues in the close distance to each neighborhood, e.g. Yoga Studio, ATM, Lebanese Restaurant.

Methodology

I will use K-means classifier to group the neighborhoods in both cities into clusters. Thus, I can use these clusters to find similar neighborhoods. In addition, I can use the data on number of

venue types as a sorting key. I.e. I can find a similar neighborhood in the other city with a specific requirement to nearby venues.

Results

Discussion

Conclusion

References

EPA. (n.d.). Sources of Greenhouse Gas Emissions. *United States Environmental Protection Agency*. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

Mann, M. E. (2014). Earth Will Cross the Climate Danger Threshold by 2036. *Scientific American*. https://www.scientificamerican.com/article/earth-will-cross-the-climate-danger-threshold-by-2036/