

How can we achieve inclusive neighbourhood change?

Predicting Gentrification

A study into planning application features that can help to signal early warnings of gentrification*

Author: Mariia Shapovalova

Date: April, 2023

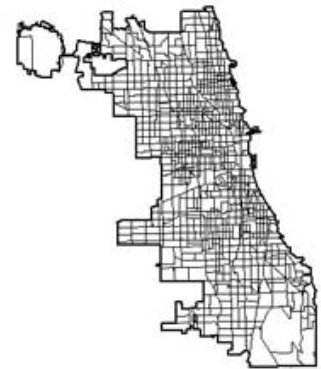
Research Aim:

Gentrification refers to a demographic transformation of a neighbourhood in which lower-income groups are displaced by the influx of residents from a higher social class. Gentrification has been studied extensively in recent years due to its increasing occurrence in major cities, with notable examples of Hackney in London or Wicker Park in Chicago. However, while gentrification can contribute to local economic development and revitalisation, it often results in the displacement of long-term residents, cultural erosion, and rising living costs, making it a highly contested and complex issue.

Due to its increasing occurrence in major cities, gentrification has been studied extensively in the quest to achieve inclusive and sustainable urban development. However, in most cases, gentrification is evaluated after it occurs. To address this issue, researchers have been developing methodologies that predict gentrification before it happens. To contribute to this discourse, this project creates a machine-learning workflow to analyse building permits in Chicago. By predicting gentrification and understanding the factors that contribute to it, policymakers and communities can take proactive measures to mitigate its negative impacts and promote equitable and inclusive urban development.

Datasets:

This project will be exploring 2 type of datasets: Chicago Planning Application Dataset (2006 to present, updated daily) and US Census Income Datasets (one for each year from 2010 to 2021).



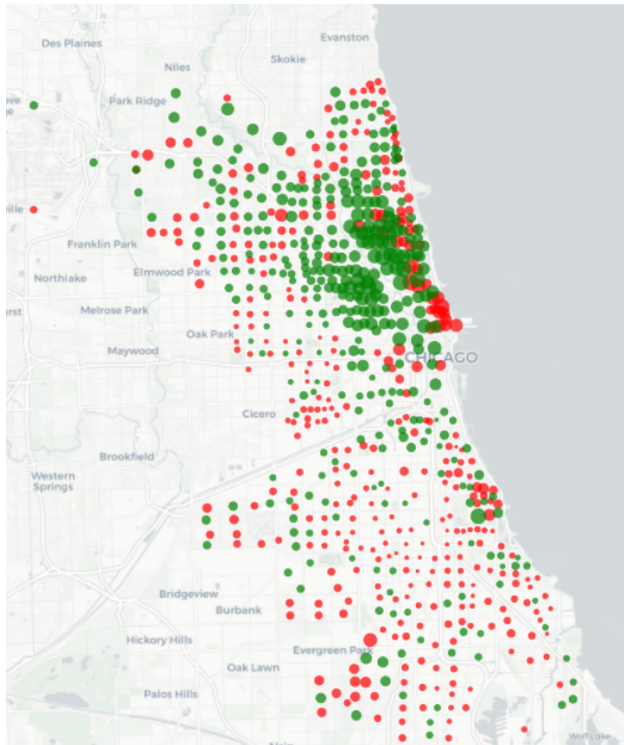
Main Cleaning Steps:

- Matching Census Tracts between all datasets
- Merging the smallest Census Tract

How can we define gentrification?

A metric combining absolute and percentage income change was developed in the process, to help capture census tracts with various income levels

Based on this metric the following Chicago areas are considered to have been gentrified from 2010 to 2021



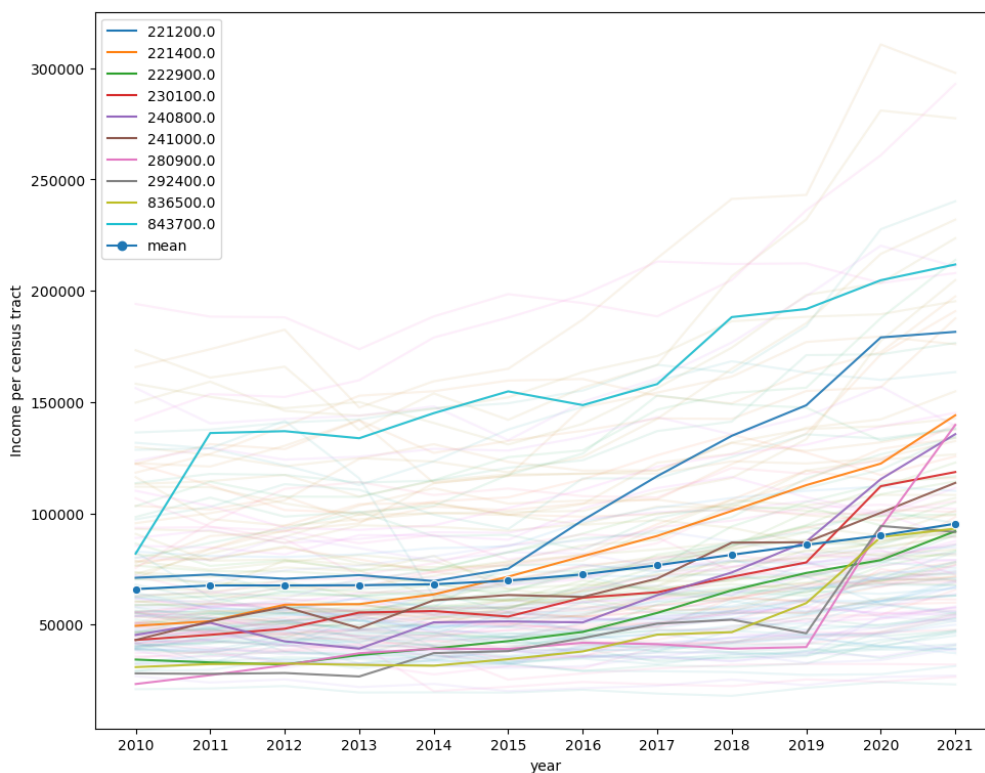
$$\log(|Income_{t2} - Income_{t1}|) * \frac{(Income_{t2} - Income_{t1})}{Income_{t1}}$$

where:

- $Income_{t1}$: The income in the initial year (t1)
- $Income_{t2}$: The income in the final year (t2)
- $|Income_{t2} - Income_{t1}|$: The absolute change in income from t1 to t2
- $\frac{(Income_{t2} - Income_{t1})}{Income_{t1}}$: The percentage change in income from t1 to t2

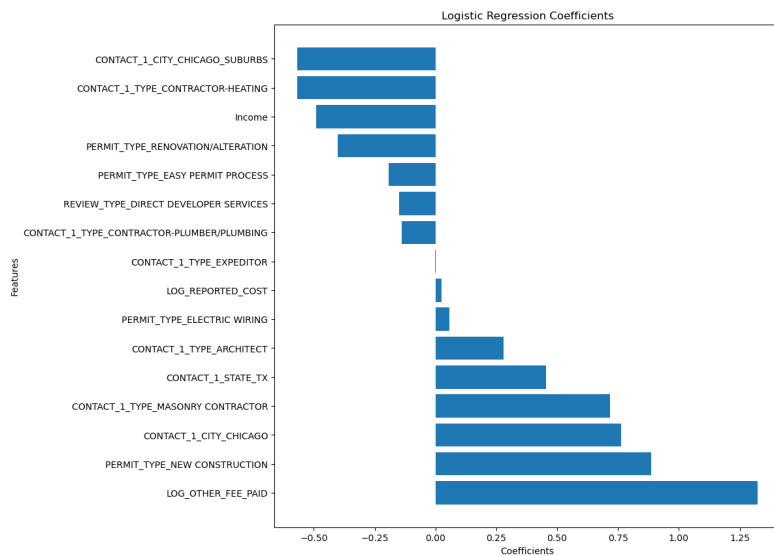
● gentrified
● not_gentrified

Mean Income Change for Census Tracts with the Highest Income Change from 2010 to 2021

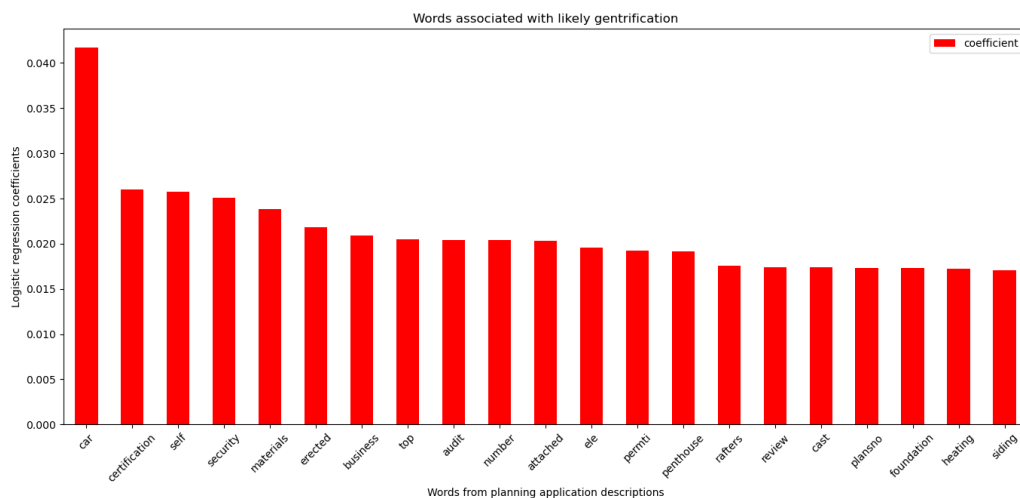


What features could be used to predict gentrification?

1. Numeric and One Hot Encoded features extracted from planning application dataset



2. NLP Analysis of Project Descriptions



Summary and Next Steps:

- The main next step would be to improve the accuracy of the target variable at predicting gentrification by introducing other variables, such as education or crime levels
- It would also be beneficial to consider adjusting the classification threshold split and consider introducing more classes
- while splitting data using media, might at first seem to be a problem, simplification, in reality, might also make it more difficult to split data, since the majority of datapoints would be around the median
- While numerical and categorical features were explored extensively, Work Description column has the potential to be investigated much further
- running initial topic analysis already offered compelling insights such as that many permits currently classified as Renovation Permits are maintenance-related
- further topic analysis can be helpful to have more control over the assignment of permit types