

# Predicting Gentrification in Denver, CO

July 5, 2018

Brittany Bennett

Springboard Data Science Career Track, Capstone 1

## Overview

Gentrification is becoming more and more of an issue as we see disadvantages communities pushed out of their neighborhoods while housing prices soar. Regardless of your stance on gentrification, it will be valuable for developers and city officials to understand if a certain neighborhood is gentrifying. Being able to predict gentrification will allow appropriate parties to better plan for the future and potentially protect residents from being displaced.

## Statistical Inference Methods

I wanted to know if census tracts with more new cafes and restaurants more prone to gentrification 5 years later (2011-2016).

**H<sub>0</sub>:** There is not a significant difference in the number of new cafes and restaurants started up around 2011 for census tracts that have been shown to gentrify by 2016 in Denver.

**H<sub>A</sub>:** There IS a significant difference in the number of new cafes and restaurants started up around 2011 for census tracts that have been shown to gentrify by 2016 in Denver.

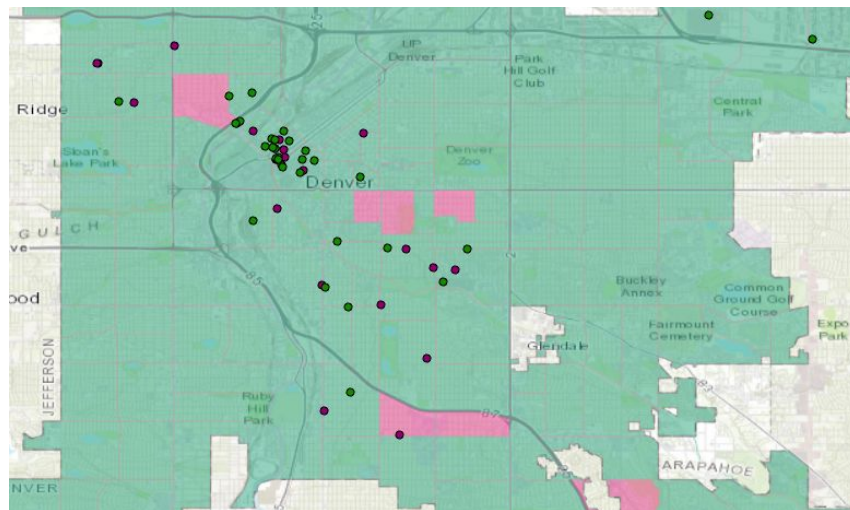


Figure 1: Spatial map of 2010 census tracts in Denver, CO. Pink census tracts are tracts that were calculated to have gentrified from 2011 to 2016. Each green dot is a restaurant established between 2009 and 2011, and each purple dot is a cafe established between 2009 and 2011.

To determine if there was a significant difference, I used a permutation test to compare the difference of two means (the mean of new places of census tracts that had or had not gentrified). I drew 100,00 permutation replicates and determined the alpha level using the code below:

```

diff_of_means = np.abs(np.mean(w.call) - np.mean(b.call))

permutation_replicates = np.empty(100000)

for i in range(len(permutation_replicates)):
    permutation_samples = np.random.permutation(np.concatenate((w.call, b.call)))

    white_perm = permutation_samples[:len(w.call)]
    black_perm = permutation_samples[len(b.call):]

    permutation_replicates[i] = np.abs(np.mean(white_perm) - np.mean(black_perm))

p = np.sum(permutation_replicates > diff_of_means) / len(permutation_replicates)
print('p =', p)

```

My calculations produced an alpha value of  $0.307 > 0.05$ . Given such a high value, I could not reject the null hypothesis. There is no significant difference between the number of new places established around 2011 in census tracts that have gentrified versus those that had not gentrified.