

# There's No Place Like Home

How one young company found the perfect place to  
put down roots

# Introduction

- ▶ A fast-growing web content management company began with no central headquarters and all employees working remotely. They've now grown to a size where a physical headquarters location makes sense.

# Business Problem

- ▶ Management strongly believe a brick-and-mortar headquarters will be beneficial.
- ▶ Want the new location to help them retain current staff and attract excellent new talent.
- ▶ Not able to lease space in one of the major US cities.
  - ▶ expense of real estate.
  - ▶ overall cost of living.

# Business Problem

- ▶ Strategy:
  - ▶ Emulate the highly-rated working conditions at top-tier competitors in their industry, in a city and neighborhood that is affordable for their young, growing company.

# Business Problem

## ► Assumptions:

- Renovating in a less expensive city is more cost-effective than leasing in a high cost-of-living city.
- The quality of a work location is determined by the quality of the neighborhood in which it is located.
- For a work location, the quality of a neighborhood is determined by the of service-oriented businesses, recreational, and cultural opportunities available.
- Top-tier competitors' neighborhoods represent ideal mixes of these venues that attract top-notch talent.

# Business Problem

- ▶ Groundwork
  - ▶ Engage a data science consultant to identify cities and neighborhoods offering the best environment for the headquarters.
  - ▶ Evaluate by:
    - ▶ Cost
    - ▶ Population
    - ▶ Commuting Time

# Business Problem

- ▶ Process:
  - ▶ Gather Foursquare venue data for top-tier competitors' neighborhoods
  - ▶ Train a machine learning model to recognize those neighborhoods
  - ▶ Use the model to identify similar neighborhoods in candidate cities

# Data

- ▶ **Target Cities List** : Narrow the list of target cities by using the [2018 U.S. News & World Report Best Places Rankings](#), filtered as follows:
  - ▶ Population size < 1,000,000
  - ▶ Average Rent < \$1000.00
  - ▶ Median Home Price < \$390,000
  - ▶ Commute Time < 20 minutes

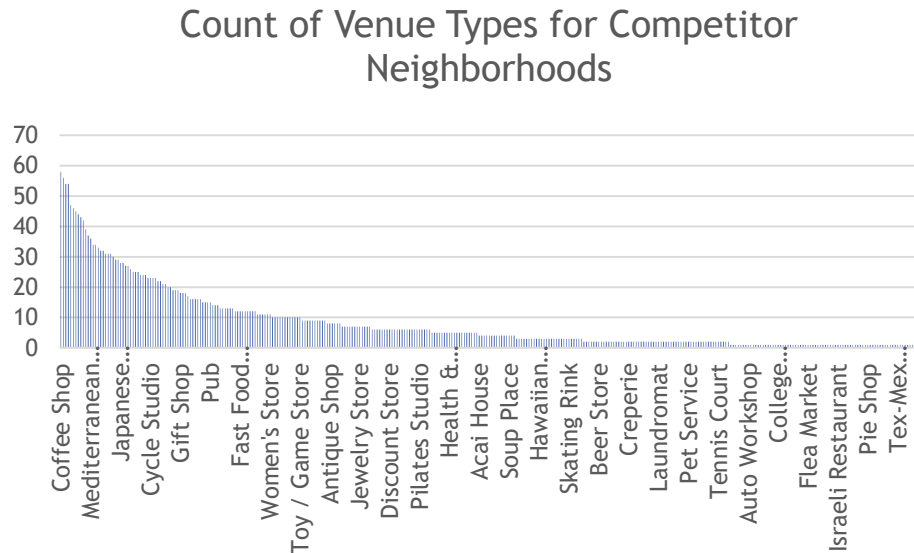


# Data

- ▶ **Candidate city Foursquare Data:** Full city records, by neighborhood, of Foursquare data from the top 10 in the U.S. News & World Report Best Places Rankings, filtered as above.
- ▶ **Competitors:** US Companies appearing appearing in the [“EContent 100” for 2018](#).
- ▶ **Training / Test Data - Positive Examples:** Foursquare data for the address of each corporate headquarters.
- ▶ **Training / Test Data - Negative Examples:** Foursquare, data from neighborhoods where there are no Competitor companies.

# Methodology

- ▶ Foursquare data for top-tier competitors:
- ▶ Venue data from within 500 meters of the street address of company headquarters
- ▶ Venue counts fell off rapidly:



# Methodology

- ▶ Competitor neighborhood venue observations:
  - ▶ Only 11 venues appeared in more than half the neighborhoods.
  - ▶ The most frequent 50 venues appeared in only 18 (22%) of the neighborhoods.
  - ▶ The most frequent 100 venues appeared in only 9 (11%) of the neighborhoods.
- ▶ Training and test sets filtered to top 25 most frequently occurring venues
  - ▶ Include venues that appear in 35% of the competitors' neighborhoods
  - ▶ Low enough number of features to minimize overhead in training model.

# Methodology

- ▶ Candidate City Foursquare data
  - ▶ List of neighborhoods and their GPS coordinates gathered manually from Google and Google Maps
  - ▶ Foursquare venue data processed in the same manner as the test/training data.
  - ▶ Filtered / Standardized to the same 25 venues as the test/training data.

# Machine Learning Model

- ▶ A logistic regression model, outputting probabilities was trained with the competitor data.
- ▶ Data for all 10 candidate cities was input, and the results combined into one result set.
- ▶ Top 10 candidate neighborhoods were as follows:

City	Neighborhood	Probability
Madison	High Crossing	0.890728
Omaha	Aksarben Village	0.839298
Wichita	Sleepy Hollow	0.814578
Des Moines	East Village	0.810072
Wichita	A Price Woodard	0.807952
Des Moines	Downtown	0.749342
Greenville	Downtown	0.737687
Madison	Eken Park	0.728206
Madison	Capitol	0.724861
Omaha	Midtown	0.715809

# Results - top ten neighborhoods

- ▶ Most recommended city - Madison - Three neighborhoods
- ▶ Wichita, Des Moines, Omaha - Two neighborhoods each
  - ▶ However, Wichita had two neighborhoods above 80% probability
- ▶ Greenville - One neighborhood
- ▶ 11<sup>th</sup> place was a tie between two Omaha neighborhoods
- ▶ After 11<sup>th</sup> place, the probability fell below 70%

# Conclusion & Observations

- ▶ Recommend the 5 cities in the top 10 to the client for further study.
- ▶ Avenues to pursue for additional accuracy:
- ▶ Additional parameter values in the models used
- ▶ More powerful model (e.g. neural network)
  - ▶ Would likely require additional computing power
- ▶ Additional features - Based on the rapid decline in frequency of venues, the limit of increased accuracy per added feature would quickly be reached