Effects of Public Transportation on Businesses

MICHAEL PRETKO

COURSERA CAPSTONE PROJECT

MARCH 7, 2020

How should you choose the location of a new business?

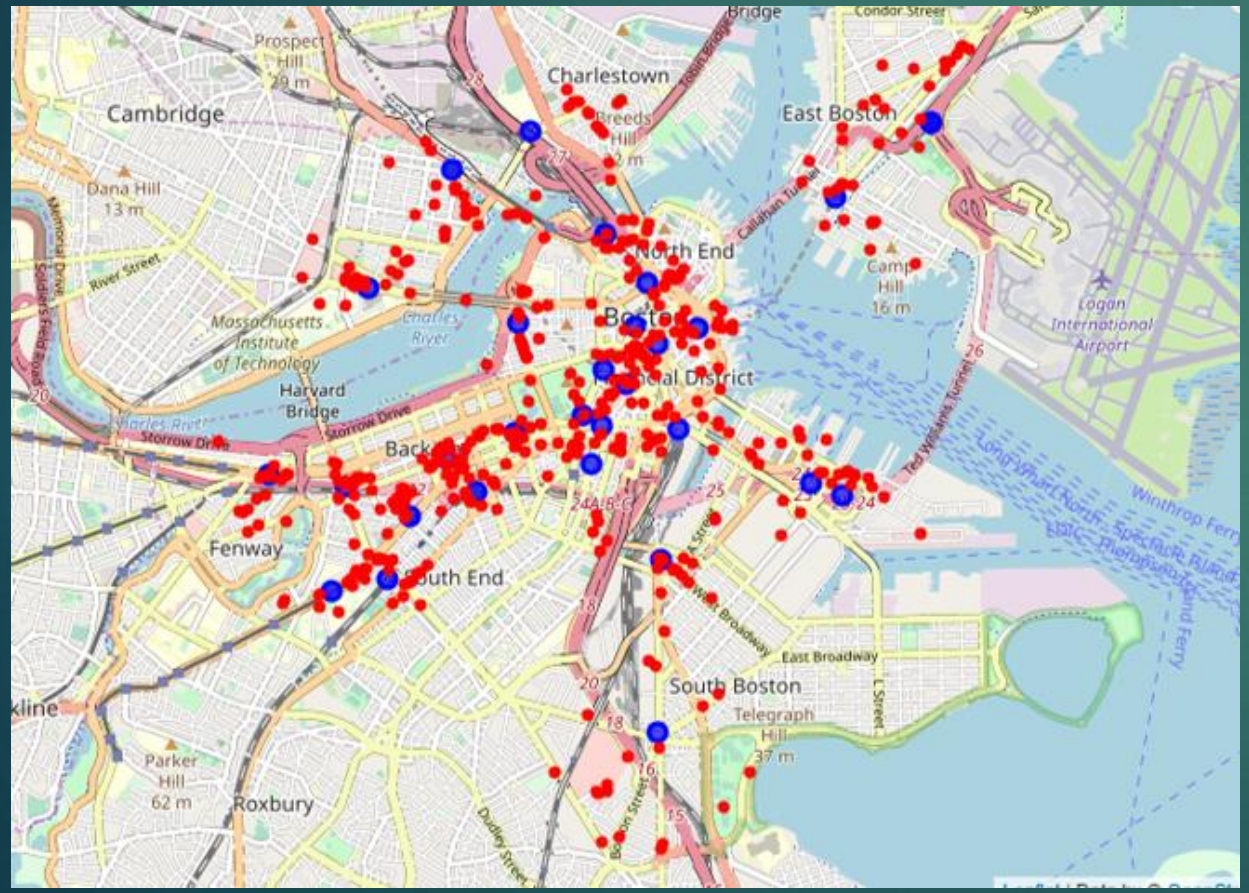
- In densely populated urban areas, many people get around primarily via public transportation

- Is distance to public transportation important in selecting the location of a new business?
  - Does proximity to public transportation help or hurt businesses, and is there an ideal distance?
  - Does the answer depend on the type of business, or is the answer universal?
  - Are the trends strong enough for public transportation to be a key factor in selecting a location?

Description of Data

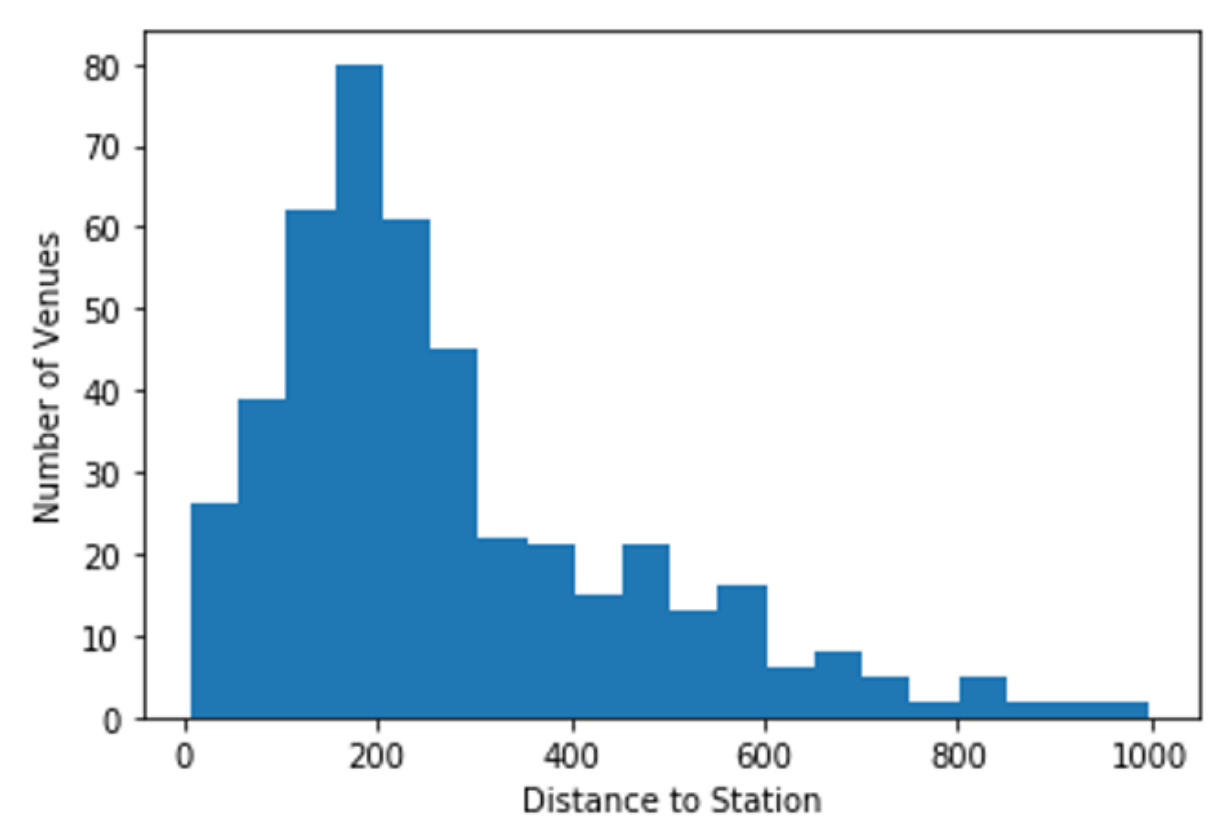
- We focus on top venues in Boston, finding their distance from the nearest subway station.
- We use the Foursquare API to obtain a list of all metro stations in the urban core of Boston (about 30 stations within a 3km radius)
- We find the top 20 venues within 1km of each station, eliminating duplicates as necessary, and finding distance to nearest station
- Premium Foursquare calls are used to obtain the ratings and number of likes for each of these top venues

Top Venues are Clustered Around Metro Stations



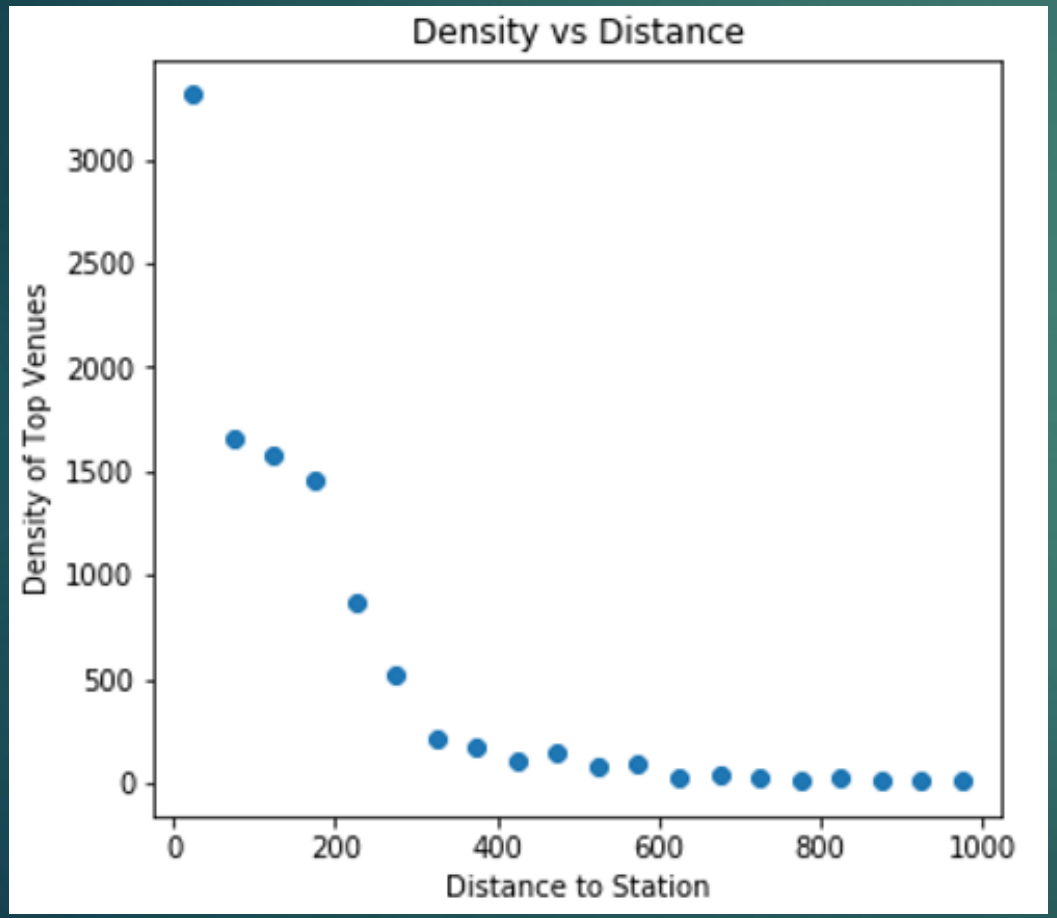
- Using Folium, we visualize the top venues in Boston (red dots) along with the metro stations (blue dots)
- In dense urban core, top venues are clustered close to the nearest station
- In less densely populated areas (e.g. South Boston), top venues tend to be more evenly spread out

Top Venues are Clustered Around Metro Stations



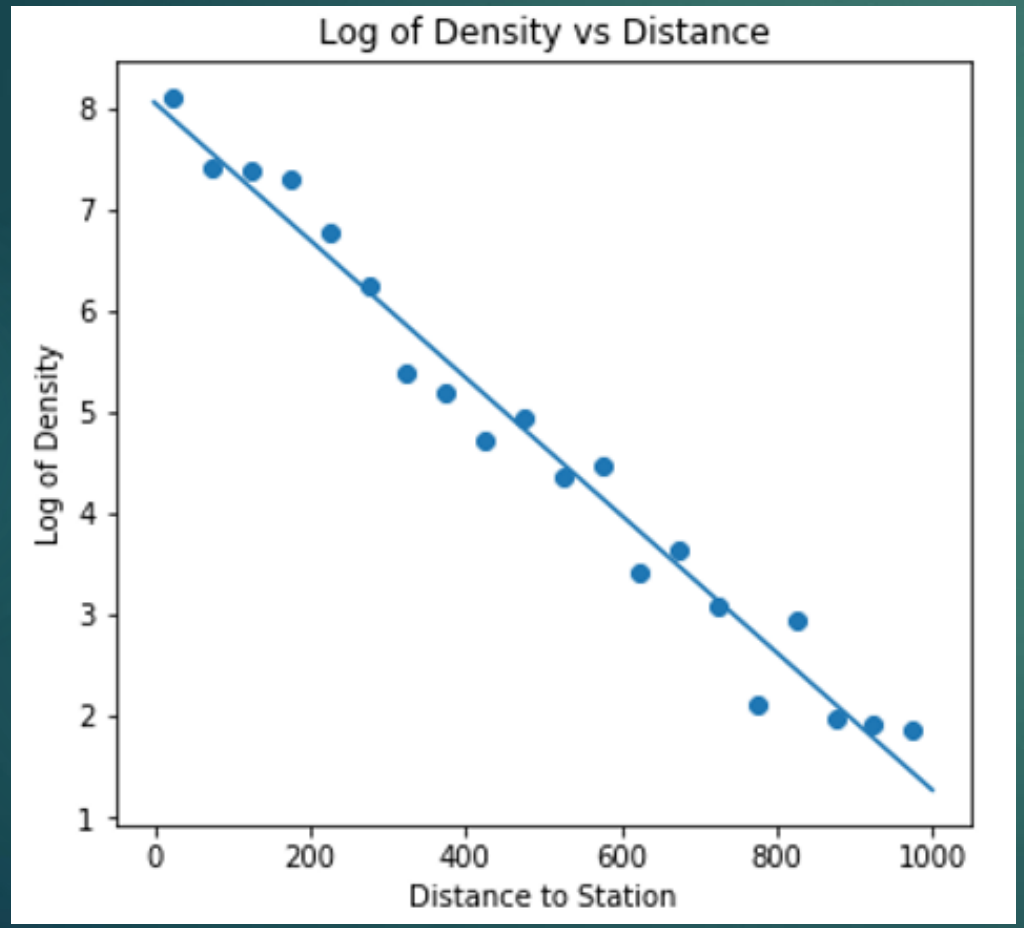
- Histogram of top venues indicates most lie between 100-300 meters away from nearest station
- Plot is slightly misleading due to growth of area with distance
- More accurate visualization requires plot of density of venues

Top Venues are Clustered Around Metro Stations



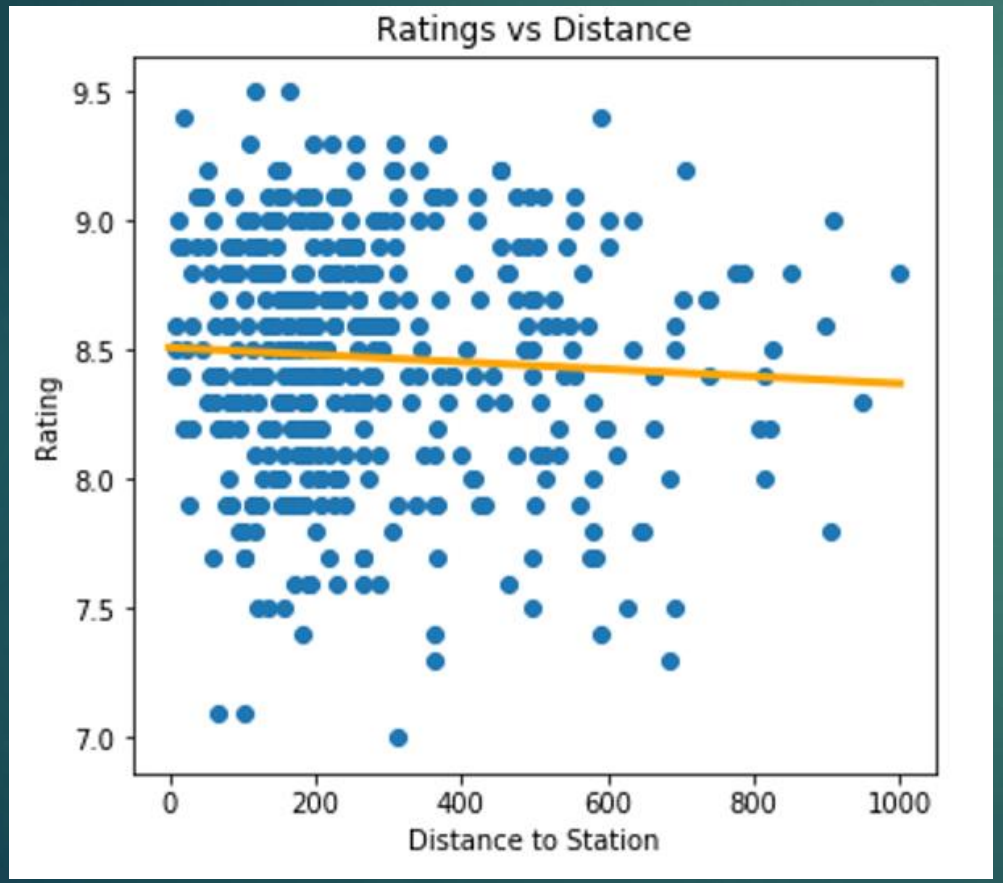
- Density of top venues falls off sharply (and monotonically) as a function of distance from station
- To find the best functional dependence, it is useful to plot the logarithm of the density instead

Top Venues are Clustered Around Metro Stations



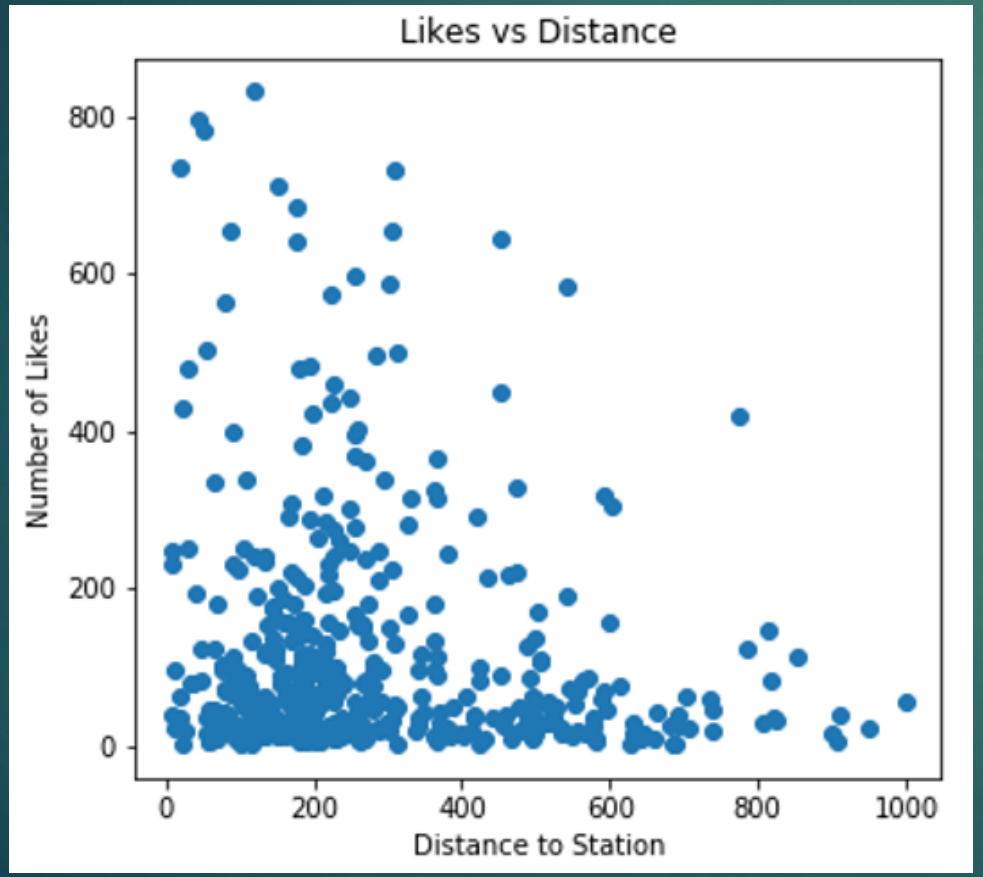
- Logarithm of density decreases linearly with distance (R-square = 0.97)
- Indicates exponential decay of density
- Identifies distance from metro stations as a crucial factor in the success of businesses

Distance from Metro Influences Quantity of Likes, Not Ratings



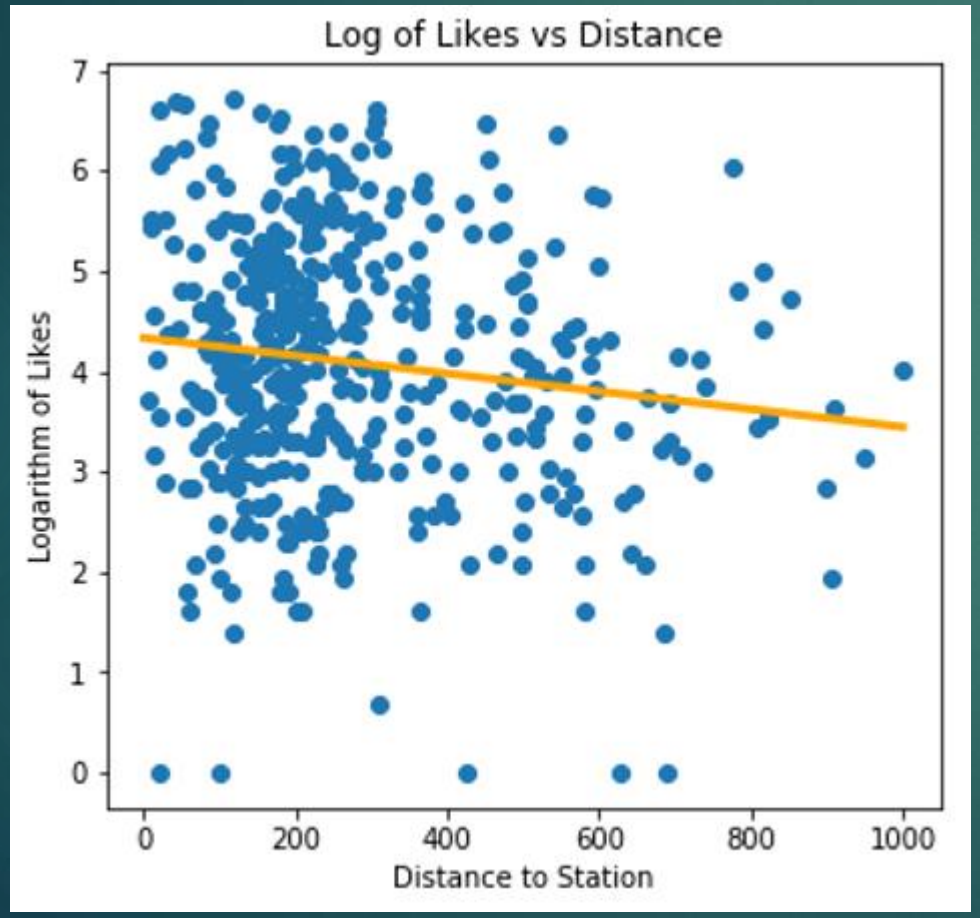
- Ratings of top venues experience only a slight decrease with distance
- No significant difference in quality between top venues close to metro and venues farther away

Distance from Metro Influences Quantity of Likes, Not Ratings



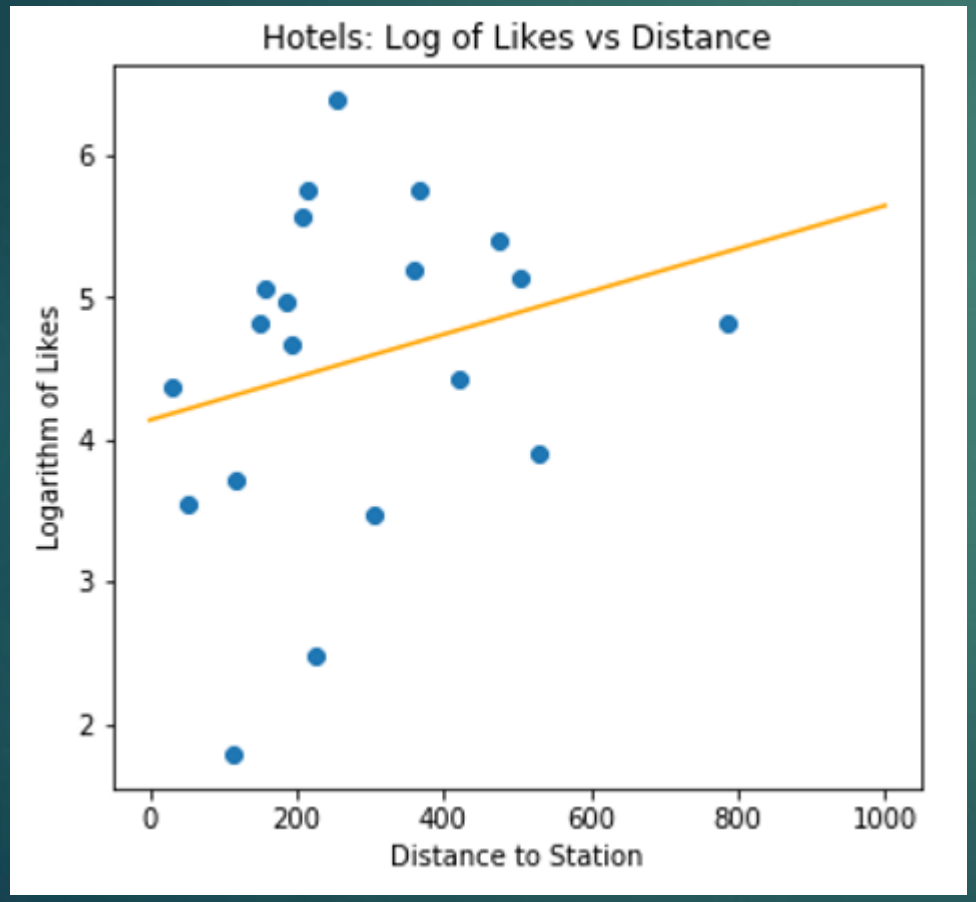
- Average number of likes drops off dramatically as a function of distance from the metro
- Curve is nonlinear and features greater fluctuations at small distances
  - Remedied by studying logarithm

Distance from Metro Influences Quantity of Likes, Not Ratings



- Logarithm of likes has more uniform fluctuations, and is roughly consistent with a linear fit
- Exponential decrease of likes as one moves away from metro stations
- Indicates distance from metro is a strong predictor of business success

Some Types of Businesses are Outliers



- Majority of businesses are more successful when closer to the metro
- A few businesses, such as hotels, exhibit opposite trend, doing better as distance increases

Conclusions and Outlook

- Success of businesses typically decays exponentially as distance from metro stations increases
- For the majority of business types, business success can be improved by being located as close to metro stations as possible
  - A small number of businesses, such as hotels, obey opposite trend
- Many important future directions:
  - Incorporate effects of rent (presumably higher near metro)
  - Study smaller cities and other modes of transportation (e.g. buses)
  - Get concrete data on the revenue of businesses