Finding the best location for opening a new vegan restaurant in LA

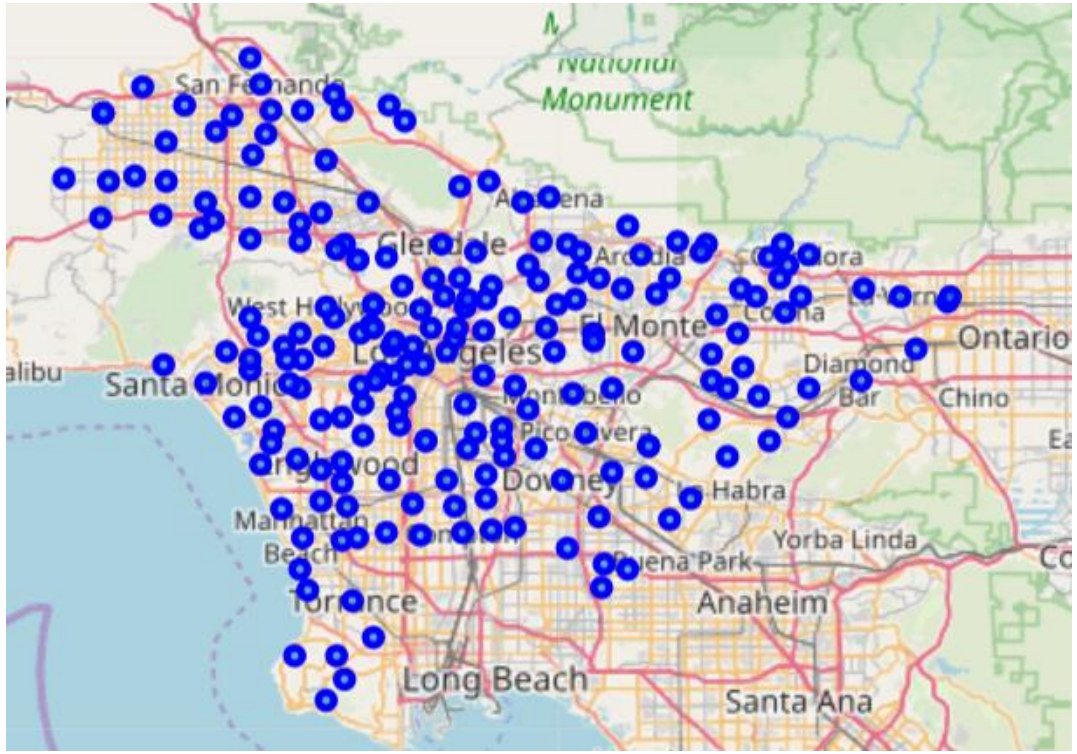
Steffen Buergers

“Vegan restaurants are on the rise” – Forbes magazine

- A vegan diet is being adopted by more and more people in the US, and the demand for vegan restaurants is correspondingly increasing. This is particularly true in large multicultural cities such as Los Angeles.

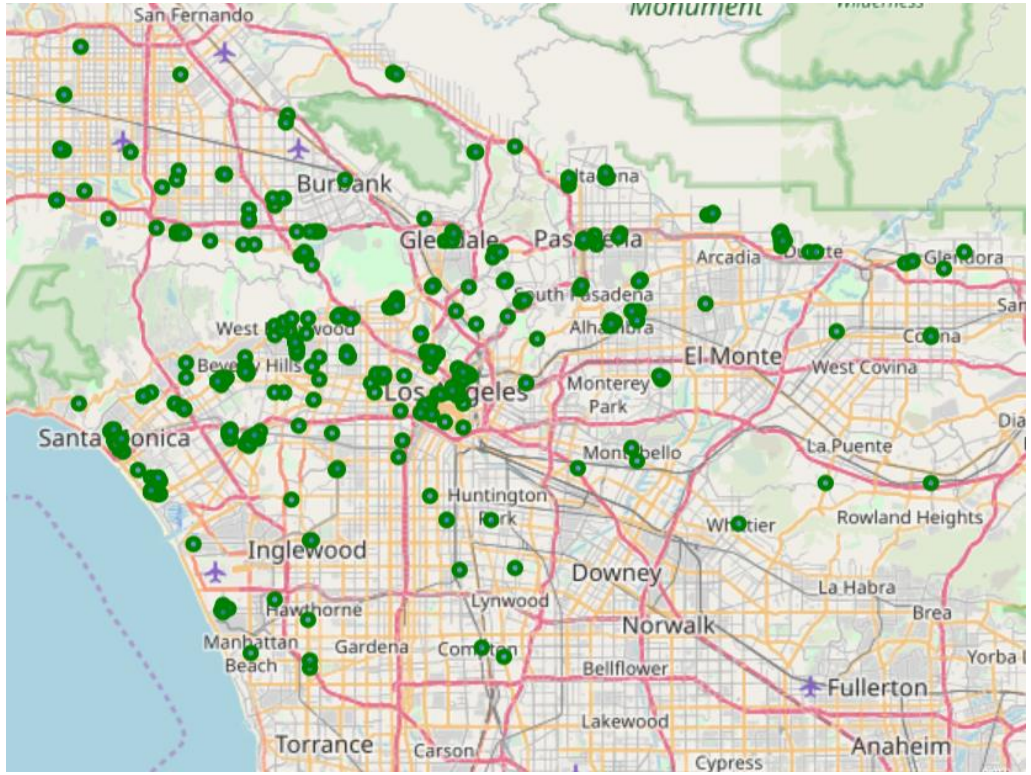
- We will use machine learning leveraging location data from Foursquare and Yelp to try to predict the best location to open a vegan restaurant in LA.

Step 1: Find geospatial coordinates neighborhoods in LA



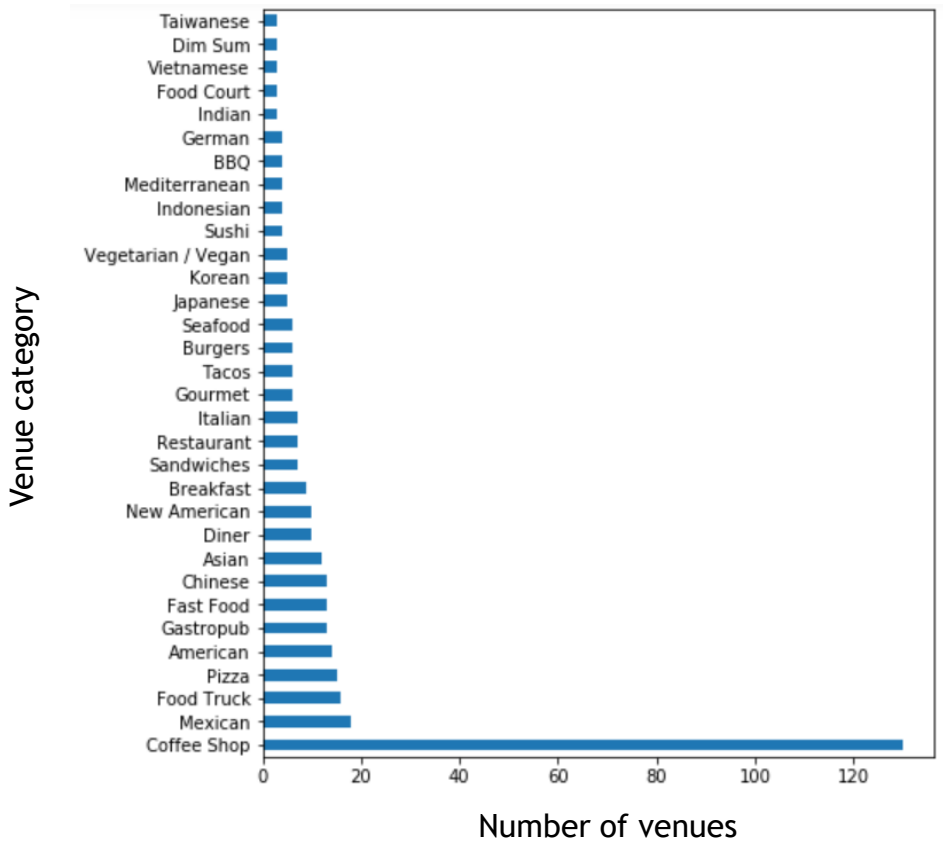
We scraped neighbourhood names from the la-times website.

Step 2: Find vegan venues around these neighborhoods



Vegan venues were obtained via the Yelp API.

The majority of these “vegan” venues are actually coffee shops



Customer ratings on Yelp and Foursquare are surprisingly uncorrelated

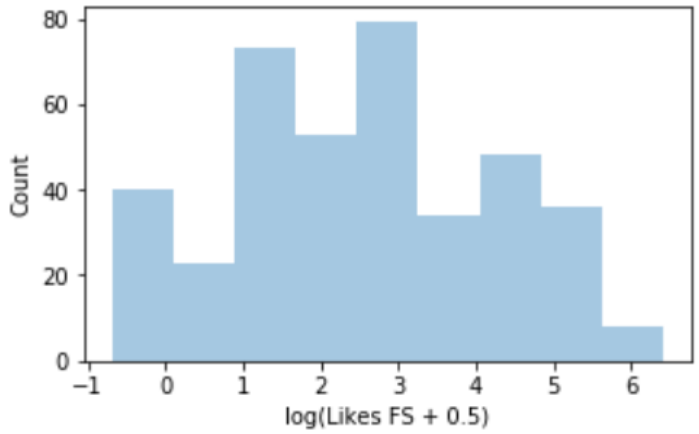
- **Rating** (customer rating on Yelp)
- **Rating FS** (customer rating on Foursquare)
- **Price FS** (price tier from Foursquare)
- **Likes FS** (number of likes on Foursquare)
- **Tips FS** (number of tips on Foursquare)
- **Photos FS** (number of photos posted on Foursquare)

	Rating	Rating FS	Price FS	Likes FS	Tips FS	Photos FS
Rating	1	-0.0444916	0.0118588	-0.157255	-0.132114	-0.174267
Rating FS	-0.0444916	1	0.237514	0.465174	0.356473	0.419765
Price FS	0.0118588	0.237514	1	0.406132	0.422209	0.418841
Likes FS	-0.157255	0.465174	0.406132	1	0.947115	0.961489
Tips FS	-0.132114	0.356473	0.422209	0.947115	1	0.917688
Photos FS	-0.174267	0.419765	0.418841	0.961489	0.917688	1

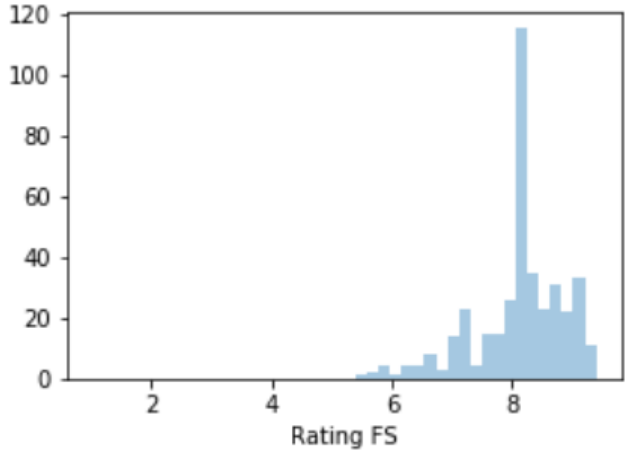
Step 3: Compute a measure of restaurant "success"

Success is a combination of rating and likes:

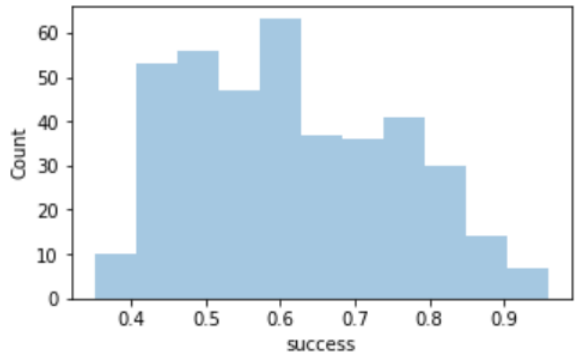
Likes FS



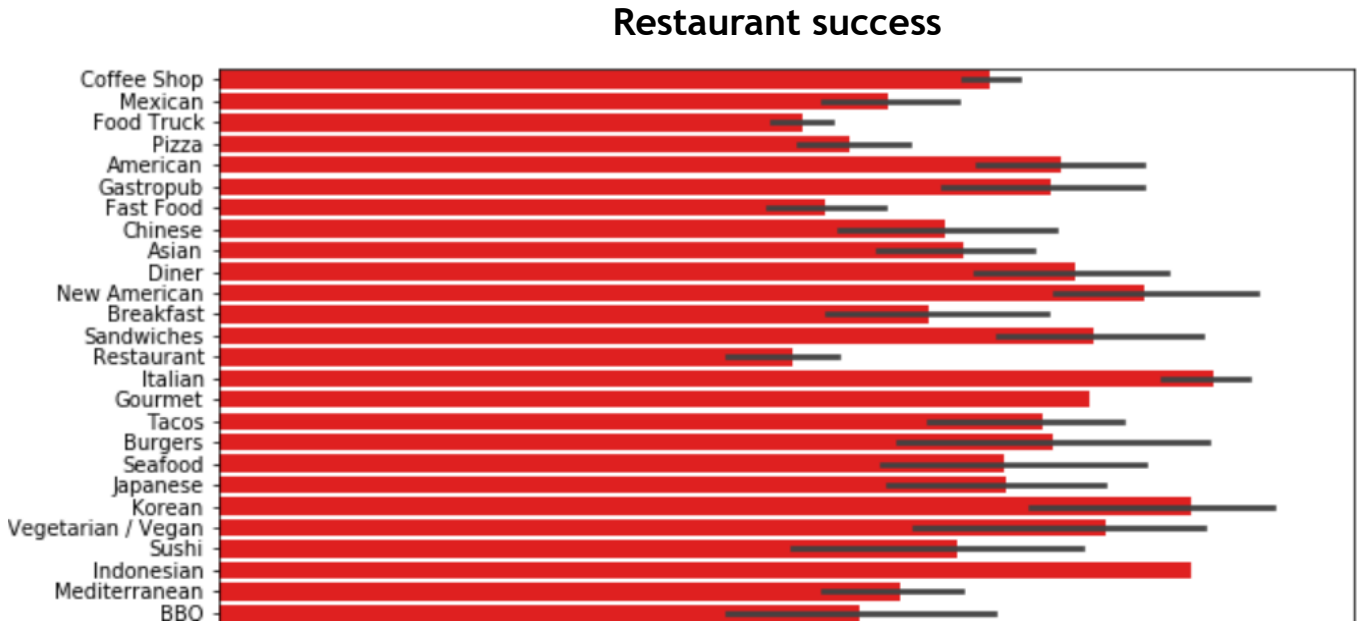
Rating FS



Success



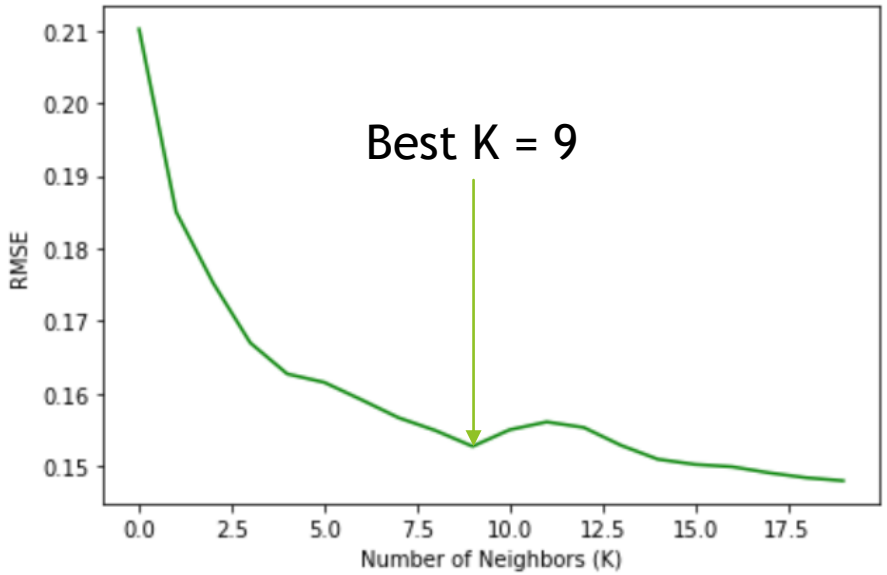
Italian restaurants are most successful



Step 4: Predict restaurant success based on surrounding venues within walking distance

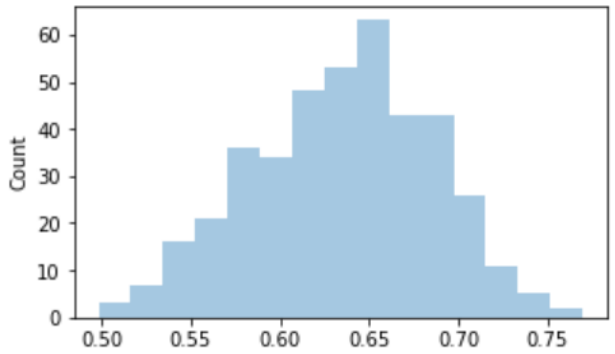
Use K-Nearest Neighbor regression

- Dependent variable: Restaurant success
- Features: Proportion of venue type within walking distance relative to all venues



Step 5: Within a pre-specified area, what is the best location to open a new restaurant?

Dark green = most likely to be successful



Predicted success

Conclusions and future directions

The K-nearest Neighbour model is able to quantify likely success for a vegan venue given its neighbourhood Characteristics.

However, the available data shows large heterogeneity in vegan venue type and it is doubtful that similar characteristics favour all venue categories (e.g. compare coffee shops vs Indian restaurant).

In the future it might be useful to gather a larger dataset over multiple cities to be able to discern the best location for a given restaurant type (e.g. vegan-Italian).

Similarly, it might be worth exploring what features will give the best results and prune the number of features compared to what was done in this analysis.