

WHERE TO OPEN A NEW PERUVIAN RESTAURANT

Adolfo G. Ramirez-Aristizabal



EXPANDING A BUSINESS

- Good food and good service make a restaurant successful
- A local restaurant owner seeks to expand their business
- The success of the second location is crucial to attract investors to commit to a chain restaurant
- We use Four Square data to refine options for expansion into new locations



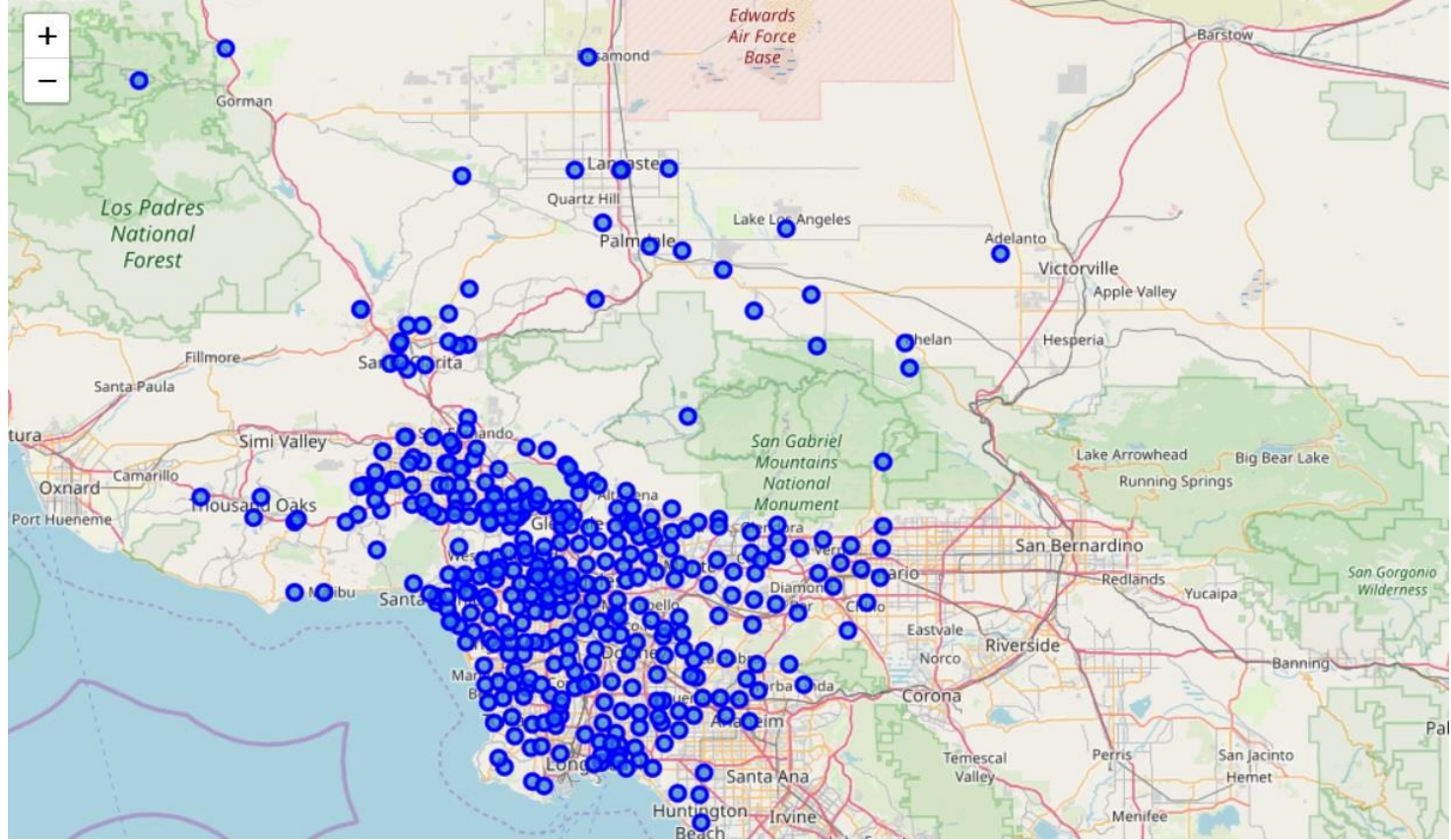
WHAT ARE THE CHALLENGES

- The new restaurant location will be in Los Angeles County
- The challenge is the size and diversity of venues that are already here
- The business will need to consider competition and relevant locations
- Peruvian food is popular but there is room in the market for a restaurant chain



THE DATA

- Our first data set comes from lacounty.gov
- Contains 370 cities with data columns [Zipcode, Latitude/Logitude, City]
- Plotted is a geographic map of the city locations from the data set



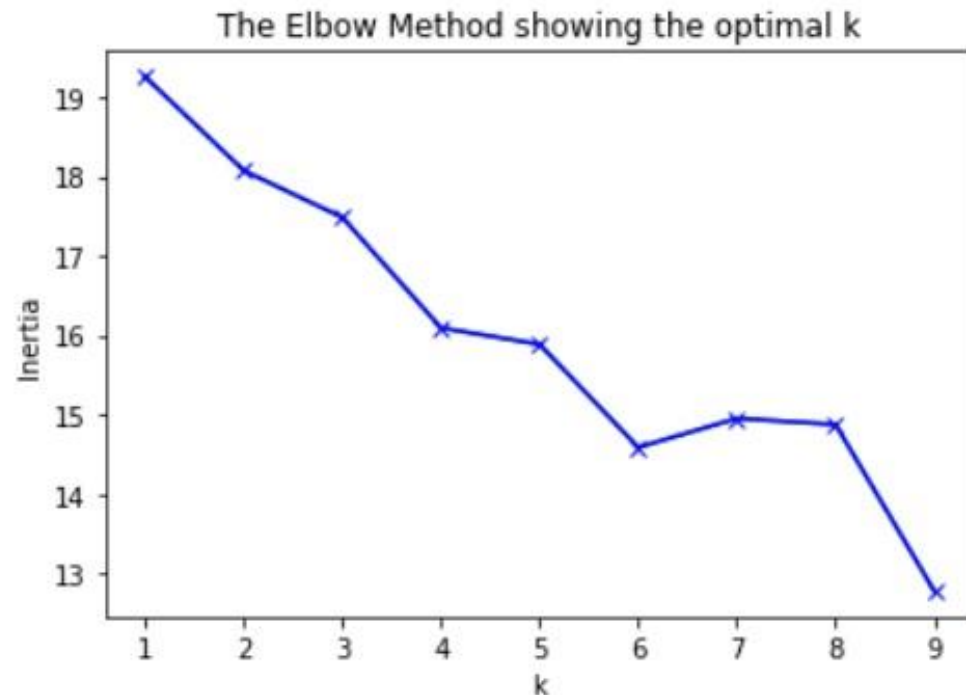
FOUR SQUARE DATA

- We use Four Square API and reference our LA county data to retrieve nearby venues
- Columns from the retrieved data include [Venue Name, Venue Location, and Venue category]
- We were able to retrieve a total of 7296 venues from the 370 cities in LA county



CLUSTERING AS IS

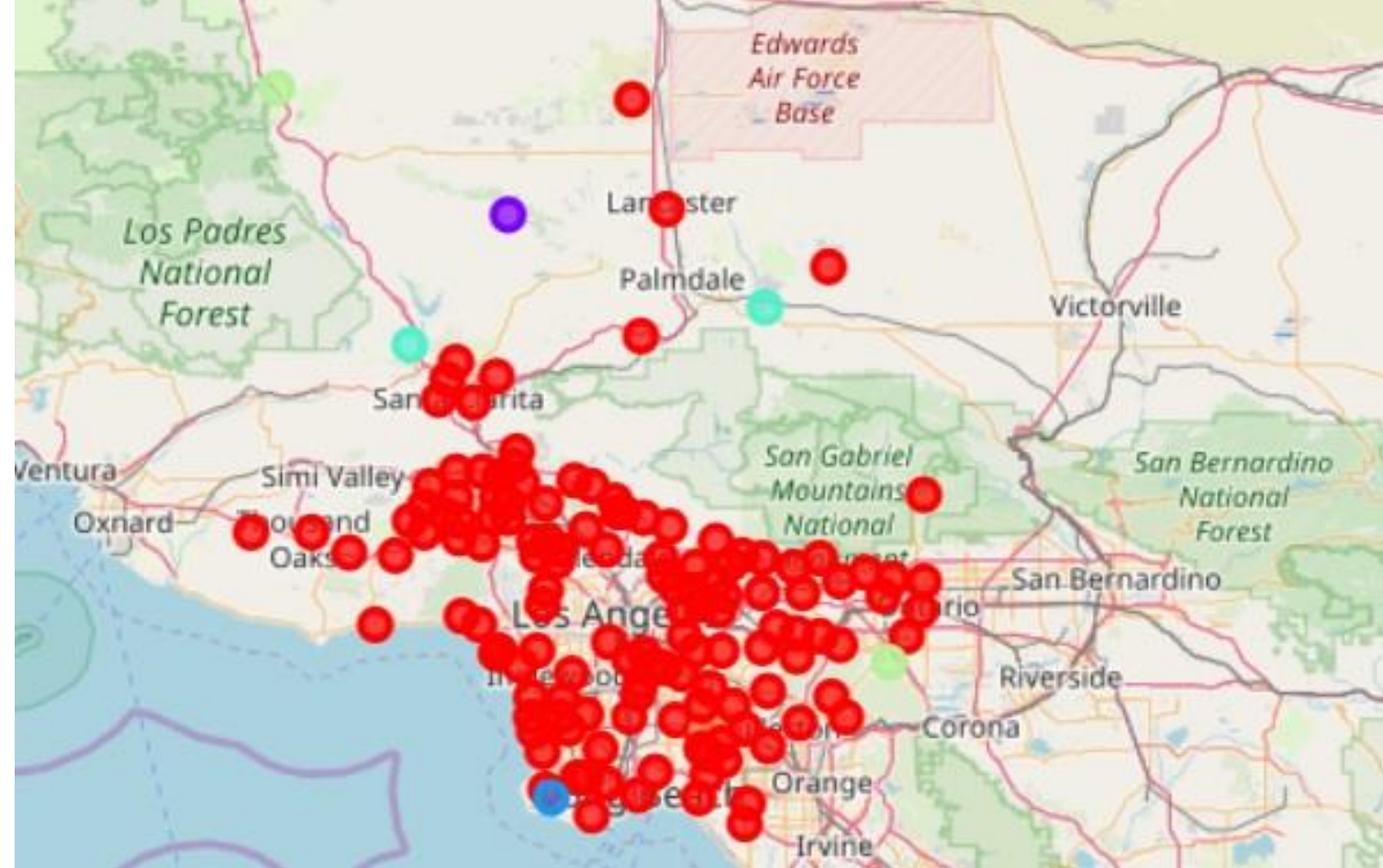
- Without any prior analysis, we take the retrieved venue data and cluster venues
- This includes all venues which include gyms, supermarkets, coffee shops etc..
- Using the elbow method we choose a $k=6$ parameter



Supermarket'
Information
Restaurant'
Shop 'Hotel' Center'
Coffee
Japanese
Lake' Tourist

CLUSTERING AS IS

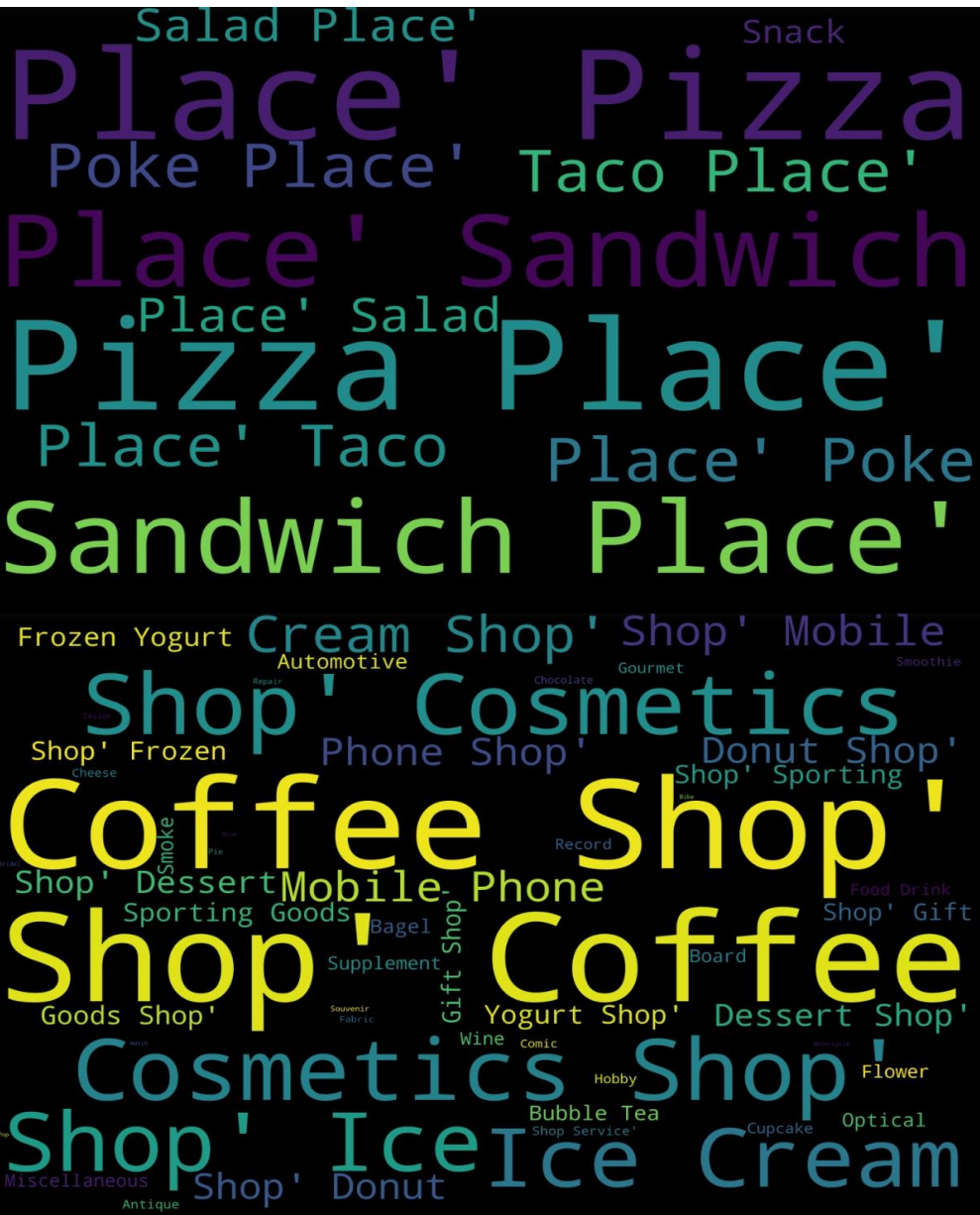
- Red cluster is for food related venues
- Other clusters do not seem very informative in their associations
- The food cluster is too broad for our search purposes



SEMANTIC CLUSTERING

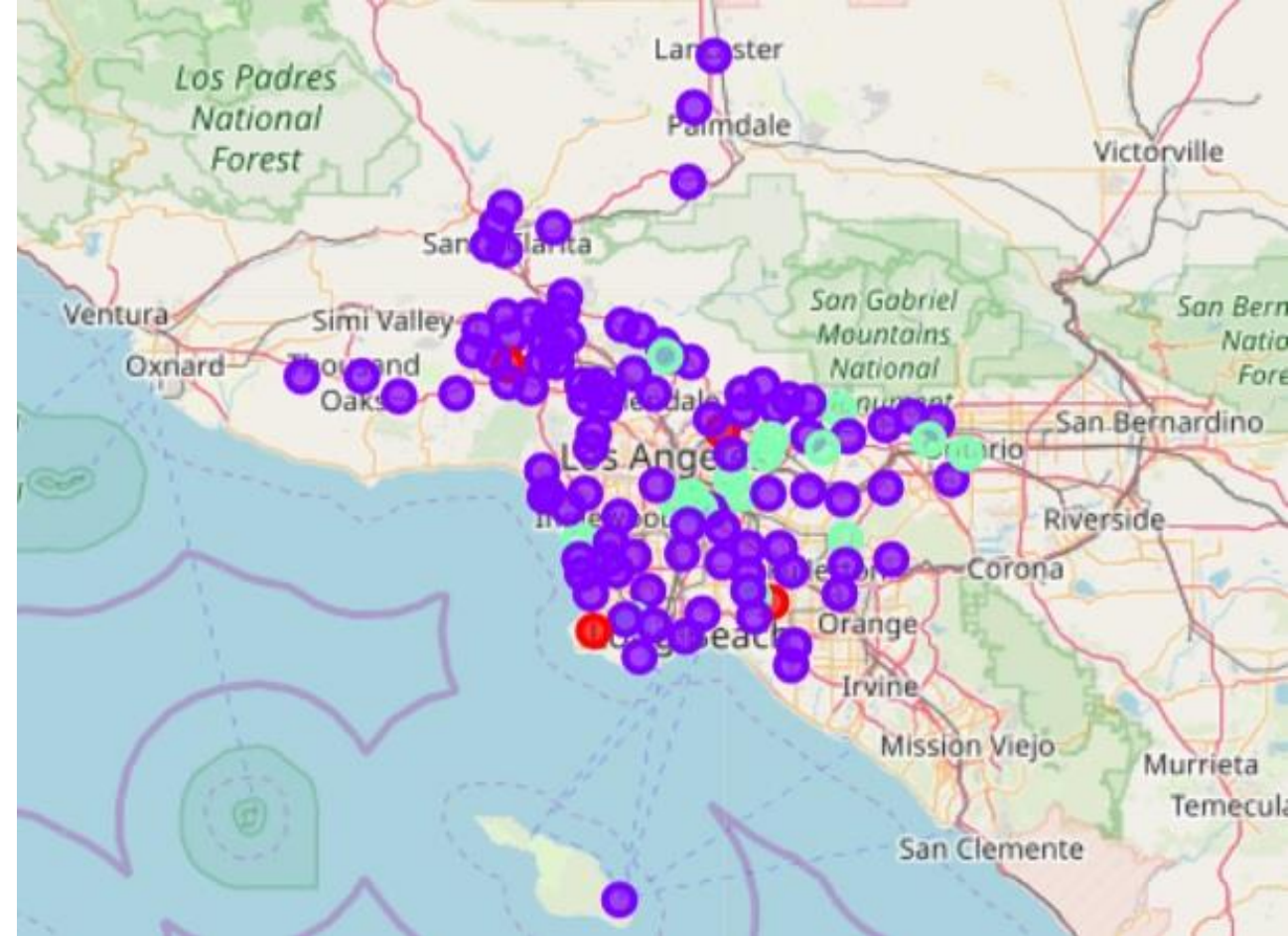
- To remedy our first run issues we use the venue category data to cluster our venues
- Tokenization of text data was done via Term Frequency Inverse Document (TFID)
- This allows us to use semantic associations in the venue description to better group venues
- New clusters now give us relevant associations of venue types

SEMANTIC CLUSTERING



SEMANTIC CLUSTERING

- We now have a restaurant venue cluster separated from other food venues
- We now add more resolution to our grouping and do K-means on just the restaurant cluster
- Our restaurant cluster gives us three new restaurant categories based on restaurant type frequencies across cities
- Fast Food-Asian(Purple),Mexican-Asian (Cyan), Diverse (Red)



INSIGHTS

- The project succeeded in providing a method to explore relevant locations for a new restaurant opening
- We provide 3 options for the business owner and investors to think about
- The 3 options balance out the scale of restaurant competition and location relevance
- Further inferential statistics can be done but is not needed to understand the quality of a possible new location
- We recommend that these 3 new data clusters be used to retrieve supplemental information about consumer traffic and social media interaction for brand growth



THANK YOU

