

# Yelp Reviews in Santa Barbara

Justin Liu, Venkataratnam Vasabathula, John Trinh

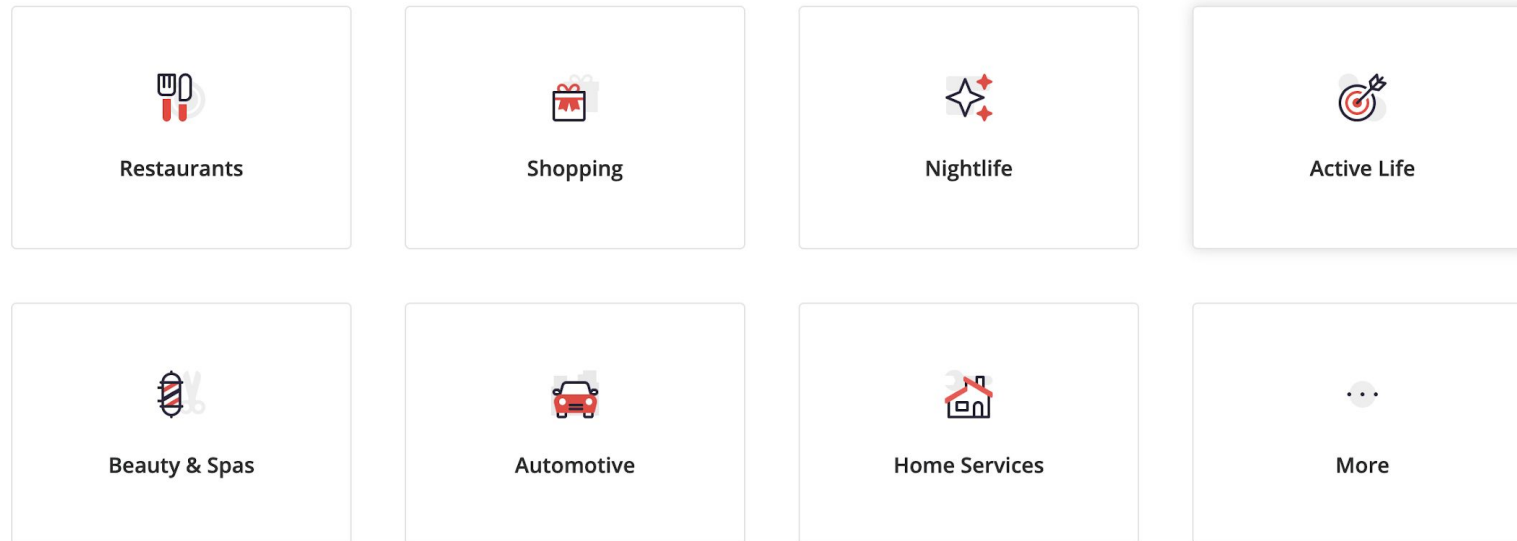
# Data Cleaning

- We checked the dataset for duplicates
- These were the main columns of concern:
  - Business IDs
  - Business Name and Address
- The duplicate: Broad Street Oyster Company.
- This data point had two entries listed
  - It is significant because is part of Santa Barbara and
  - This business was split up into two entries

# Data Cleaning Challenge 1

- There were over 1000 categories to select from
  - This was hard to decide to look at
  - Decided to use the top 7 categories from Yelp

## Categories

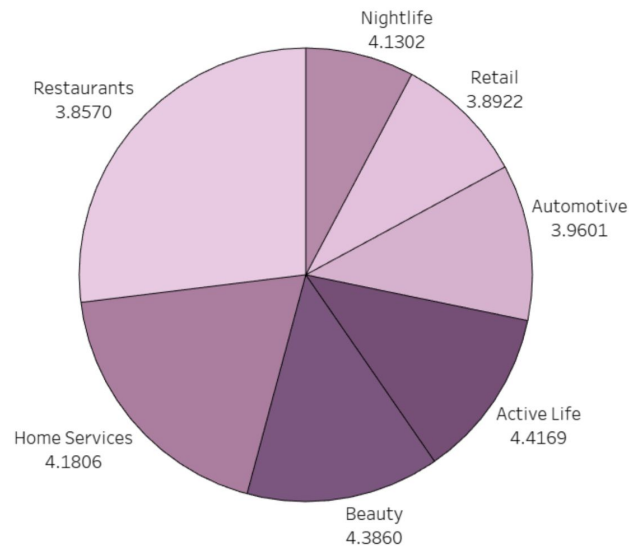


# Data Cleaning Challenge 2

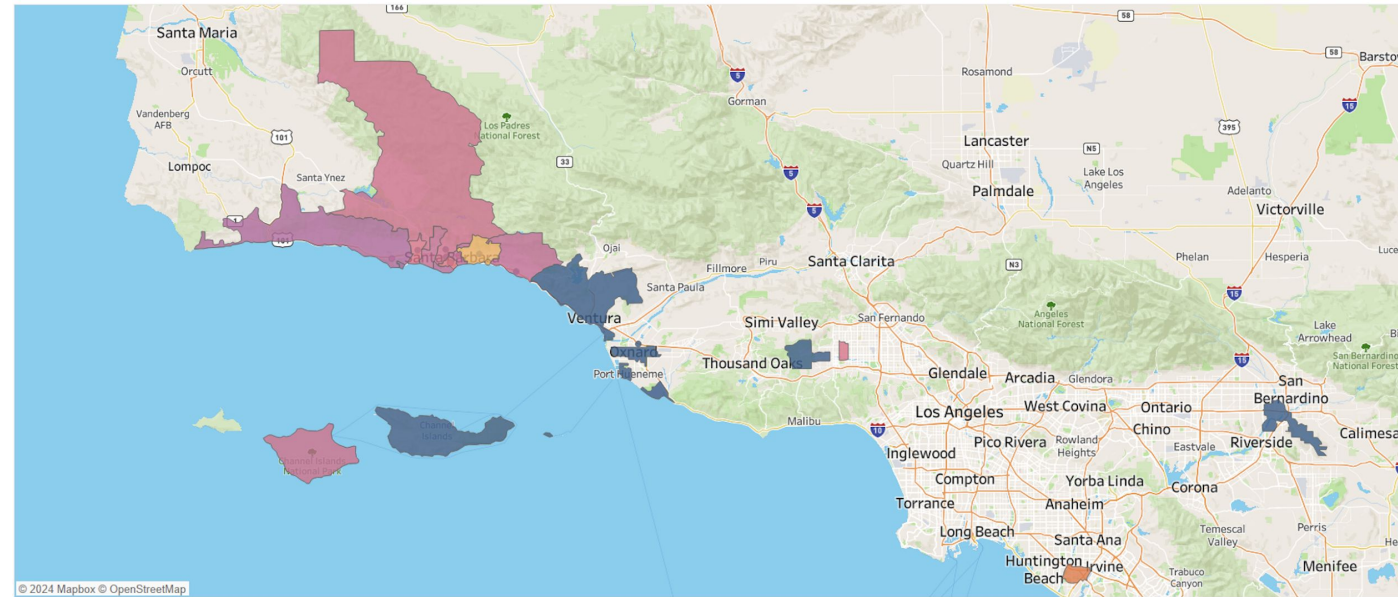
- There were over 10 cities listed for all the businesses within the dataset
  - However, most of the businesses lie within Santa Barbara County
  - Data shown is in Santa Barbara County

# Data Visualization

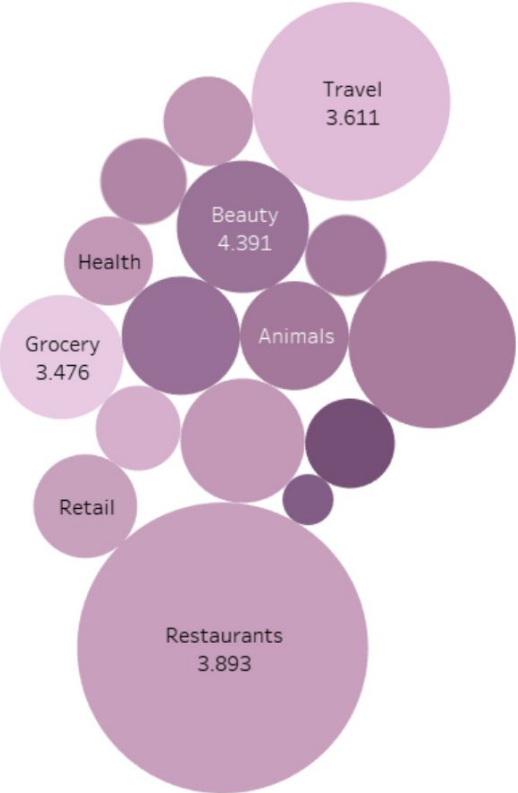
Average Stars by Category



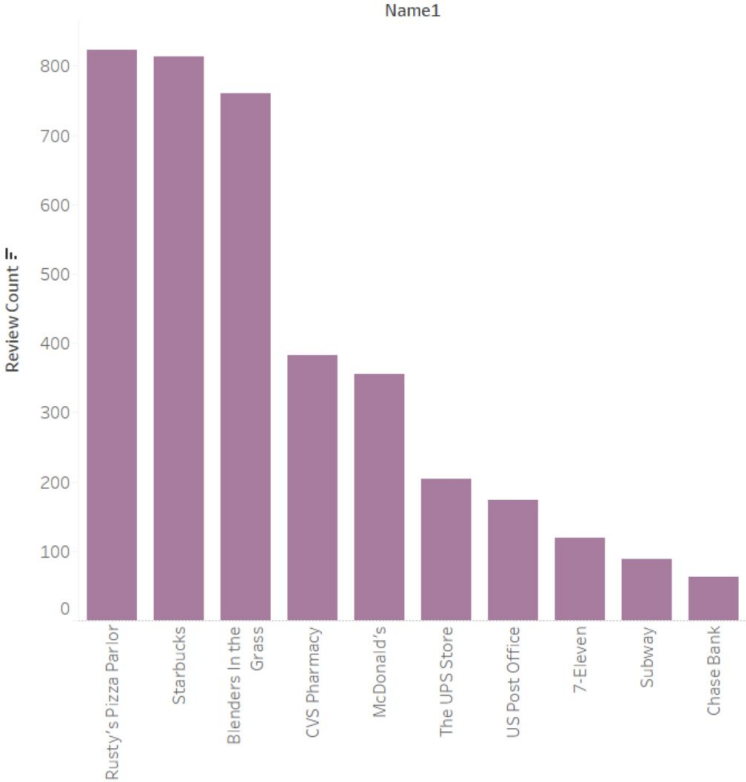
Map is for average restaurant price range for each postal code



The marks are labeled by Categories and average of stars and size shows avrage of review count



Top 10 Business by review count



# Data Storytelling

- We now have to determine what makes a business “successful”
- We used two metrics for each business:
  - Average stars (or average ratings)
  - Number of reviews for each business.
- We argue this: Nightlife business that is \$ or \$\$, within Downtown Santa Barbara, you will find people coming into your business