Yelp Ratings & Restaurant Attributes in Orange County

Team 5B: Boram Gaudet, Chris Gu, Jason Jia, Srita Kothuri

BANA 212

December 8, 2025

Abstract

Orange County's diverse food scene inspired our group to analyze how different factors such as food quality, ambience, and customer service affect a restaurant's review. We scraped data from Yelp, a popular crowd sourced review platform, and compared how 10 different restaurants ranging across 5 cuisines differed along each of these attributes. Using statistical analyses, including Analysis of Variance (ANOVA) and Simple Regression Models, we found that the five chosen cuisines did not vary significantly across the three chosen attributes. Additionally, we found that food quality was the factor that contributed most to a restaurant's overall rating.

Introduction

Orange County, California is known for its high-quality and diverse food scene. A 2025 Eater LA review of Orange County's restaurant landscape highlights the region's wide range of cuisines, underscoring the culinary diversity that motivated our choice of cuisines for this analysis: Vietnamese, Korean, Japanese, Mexican, and Italian (Eater LA, 2025). To explore how people experience and evaluate this variety, we are using Yelp, which is one of the most widely used platforms for connecting customers with local businesses. Users can search for restaurants and share their experiences through ratings and written reviews, and Yelp's *review insights* feature provides aggregated data on how often certain attributes, such as ambience or service, are mentioned in positive or negative reviews. Using Yelp ratings and review insights as our dataset, this project examines restaurants in Orange County. Specifically, we focus on how food quality, ambience, and customer service affect overall restaurant ratings and how these attributes differ across cuisine types, which may reveal how various cuisines prioritize different aspects of dining experience. Overall we aimed to answer two main questions with our analysis:

1) Do attributes such as food quality, ambience, and customer service differ significantly between cuisines?
2) Which of the three attributes influence a restaurant's overall rating the most, and does this differ across cuisines?

Gaining insight on these questions can help us formulate suggestions for business owners in Orange County, as well as consumers who are looking to dine at these establishments.

Data Collection and Exploratory Analysis

The data acquisition process was done manually. First, we chose five cuisines that are most popular among restaurants in Orange Country: Vietnamese, Korean, Japanese, Mexican, and Italian. Then, we selected the top 10 Yelp search results for restaurants within each cuisine type, giving us a total sample of 50 restaurants. Next, we recorded the basic details of the restaurant that Yelp provides, including the city the restaurant is located in, price range, review count, and overall rating.

After that, we collected attribute-specific data from Yelp's "Review Insights" section for each restaurant. This section shows, for any chosen attribute, how many reviews mention it and what percentage of those reviews are positive. Because our research focused on ambience, food, and service, we filtered for each of these attributes individually. For example, the Vietnamese restaurant VOX Kitchen has 81% positive reviews among those that mention "ambience."

We compiled all of this information into an Excel sheet and then loaded it into Python for analysis. Below are the final columns we used:

- name: Restaurant name
- cuisine: Cuisine type (Vietnamese, Korean, Japanese, Mexican, or Italian)
- location: City in Orange County where the restaurant is located
- price: Restaurant's price range ($-$$$)
- overall_num_review: Total number of reviews for the restaurant
- ambience_num_review: Total number of reviews mentioning "ambience"
- ambience_rating: Proportion of positive reviews mentioning "ambience"
- food_num_review: Total number of reviews mentioning "food quality"
- food_rating: Proportion of positive reviews mentioning "food quality"
- service_num_review: Total number of reviews mentioning "service"
- service_rating: Proportion of positive reviews mentioning "service"

Before conducting statistical tests, we performed exploratory data analysis to gain an understanding of the distribution of the data through aggregating and visualizing our variables. First, we examined the overall restaurant rating distribution and saw that average ratings ranged from about 3.6 to 5 stars, with most restaurants falling at 4 stars and above (Fig 1). Next, we examined the average proportion of positive reviews mentioning each attribute. Across all 50 restaurants, food quality had the highest percentage of positive mentions, followed closely by ambience, with service receiving slightly lower positive ratings overall (Fig 2). Finally, we analyzed the average proportion of positive reviews mentioning each attribute by cuisine type. The chart suggests that the most positively mentioned attribute may vary across cuisines. For Italian and Mexican restaurants, "ambience" appears most frequently in positive reviews, while for Japanese, Korean, and Vietnamese restaurants, "food quality" receives the highest proportion of positive mentions (Fig 3). From the visual, we can see slight differences in attribute mentions depending on cuisines, and in the next section we test whether these differences are statistically significant.

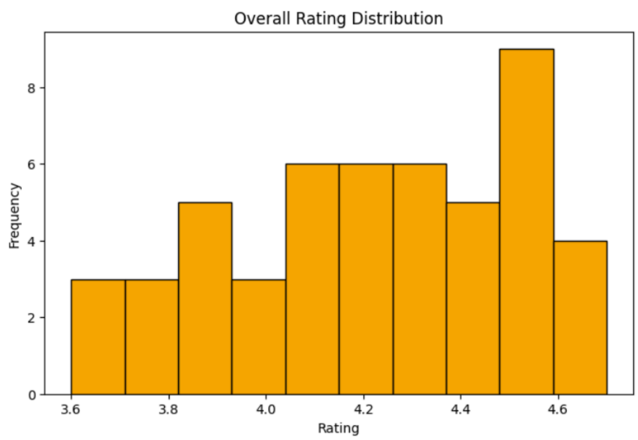


Fig 1. Overall restaurant rating distribution

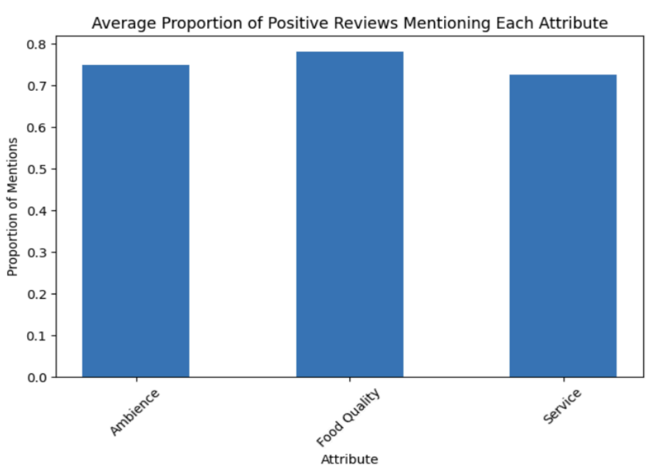


Fig 2. Average proportion of positive reviews mentioning each attribute (ambience, food quality, and service)

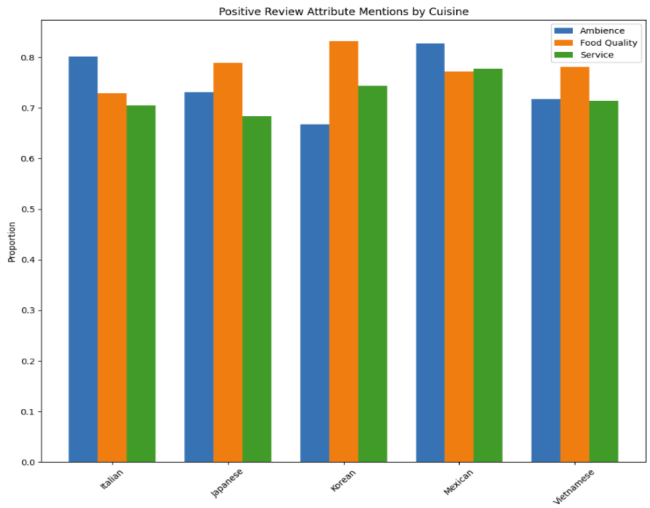


Fig 3. Positive review attribute mentions by cuisine

Data Analysis

ANOVA

To answer our first research question—"Do attributes such as food quality, ambience, and customer service differ significantly between cuisines?"— we conducted a two-tailed, one way Analysis of Variance (ANOVA) test. ANOVA tests are used to compare the averages or proportions of two or more groups of variables. In the case of our project, we compared the average proportion of positive reviews mentioning each of the three attributes. We decided to analyze our data as such because we assumed that if a positive review of a restaurant mentioned one or more of the attributes, they were likely signaling that this was an important and influential factor in their review of the restaurant. We ran four separate ANOVA tests, one for each our attributes as well as overall ratings, to see if these values differed significantly between cuisines:

- Average Proportion of Positive Reviews Mentioning Ambience
- Average Proportion of Positive Reviews Mentioning Food Quality
- Average Proportion of Positive Reviews Mentioning Service Quality
- Average overall rating
    - H_0: $\mu_{Italian} = \mu_{Japanese} = \mu_{Korean} = \mu_{Mexican} = \mu_{Vietnamese}$
    - H_a: $\mu_{Italian} \neq \mu_{Japanese} \neq \mu_{Korean} \neq \mu_{Mexican} \neq \mu_{Vietnamese}$

Before running ANOVA on our data, we checked our assumptions and ran the Shapiro-Wilk test on each attribute column to ensure the normality requirement for ANOVA was met. The Shapiro-Wilk test is used to check for the assumption of a normal distribution in a dataset. Only one of our three attributes, ambience, did not pass the assumption of normality; therefore, we transformed the variable via an arcsine square root transformation.

To run ANOVA, we imported the scipy.stats library on Python and used the f_oneway function to report the f-statistic and p-value of each of the ANOVA tests. This can be seen in Table 1 below.

Attribute Tested	**F-Statistic**	**P-Value**	**Tukey Post Hoc**
Ambience	3.43	0.016	Yes
Food Quality	2.28	0.076	No
Service Quality	0.74	0.569	No
Overall Rating of Restaurant	0.263	0.899	No

Table 1. Statistics for one way ANOVA test conducted on restaurant attributes.

The preliminary ANOVA tests showed that there was no significant difference in the proportion of positive reviews that mentioned food quality and service between cuisines. Additionally, there was no difference in overall rating of restaurants from different cuisines. This was made clear when comparing the plotted values, as can be seen in Figure 4 below.

While we didn't see any difference for most of our attributes, the ANOVA did flag that the average proportion of positive reviews that mentioned ambience did differ between cuisines. We then conducted a post-hoc Tukey test to see where this difference occurred and found that ambience ratings were significantly different between Mexican and Korean restaurants (Fig 4), with a mean difference of 0.16 (p-value = 0.0207). Overall, Mexican restaurants had, on average, a 0.16 higher proportion of positive reviews mentioning ambience than Korean restaurants.

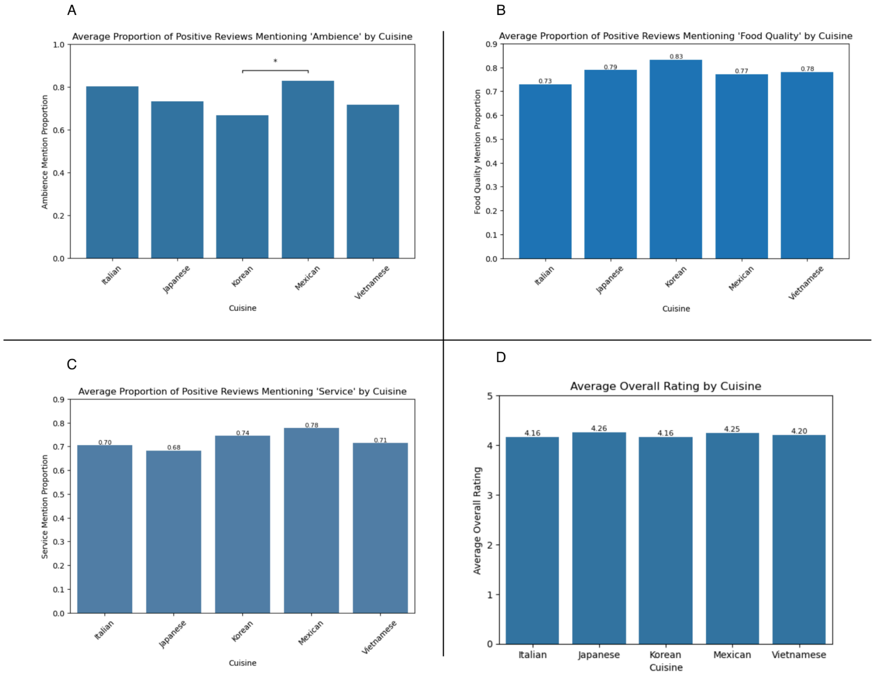


Figure 4. Graphical representation of average positive review mentioning each restaurant attribute (A-C) and overall rating (D). Asterisk represents significance at alpha level of 0.05.

Pearson Correlation Analysis / Simple Regression

To answer the second portion of our research question—“Which of the three attributes influence a restaurant's overall rating the most, and does this differ between cuisines as well?”—we ran Pearson's correlation analysis comparing each cuisine with the three restaurant attributes. Pearson’s correlation tests measure the linear relationship between two variables. It allows us to visualize and understand the strength and direction of association between the restaurant attributes and cuisines. We ran a total of six sets of Pearson’s correlation tests, examining the relationship between overall rating and each of the three attributes within each of the five cuisines, as well as an additional test assessing the relationship between overall ratings (across all cuisines) and the three attributes. The Pearson’s correlation coefficient (r) was then squared to give us the r^2 value, which provides a more interpretable measure of how much variance the attribute accounts for in the corresponding rating.

To conduct the simple regression analysis, we used the stats.pearson function from the scipy.stats library in python. The r and r^2 values from this analysis are reported in Table 2 below.

Cuisine	Ambience (r)	Ambience (r2)	Food Quality (r)	Food Quality (r2)	Service Quality (r)	Service Quality (r2)
Italian	0.67*	0.45*	0.87**	0.76**	0.69*	0.48*
Mexican	0.37	0.14	0.91*	0.83*	0.85*	0.72*
Japannese	0.42	0.18	0.91***	0.83***	0.66*	0.44*
Korean	0.71*	0.5*	0.51	0.26	0.46	0.23
Vietnamese	0.59	0.35	0.67*	0.45*	0.43	0.18
Overall rating for all cuisines	0.48***	0.23***	0.71***	0.5***	0.60***	0.36***

Table 2. Pearson's correlation coefficients and r^2 values for simple regression model. Single asterisk represents significance at alpha level 0.05, double asterisk indicates significance at alpha level 0.01, and triple asterisk indicates significance at alpha level 0.001.

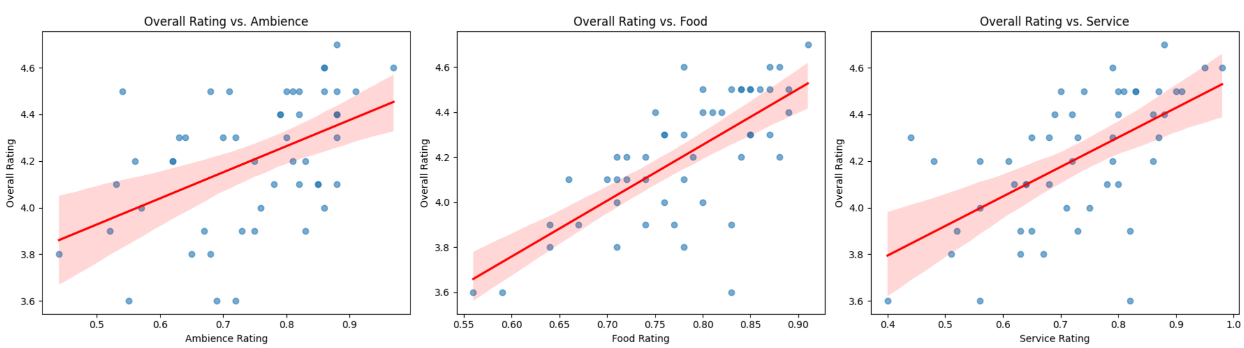


Figure 5. Association between overall rating of all restaurants and restaurant attributes.

The results from the simple regression models show that, for all cuisines except Korean, food quality explains the most variance in customer ratings. This indicates that, for most cuisine types, food quality is the primary driver of how customers rate their dining experience. This pattern is also reflected in the overall rating across all cuisines. The regression outputs further show a strong relationship between food quality and overall rating, as demonstrated by the magnitude of the r^2 values. For Korean cuisine, however, ambience is the strongest predictor of customer ratings, with an r^2 value of 0.5. Figure 5 illustrates the relationships between overall ratings and each restaurant attribute across all cuisines. All three attributes display strong positive correlations with overall rating, suggesting that each plays an important role in shaping customer perceptions. However, food quality consistently emerges as the most influential factor.

Multiple Regression

To round out the analysis, our group also conducted a multiple regression analysis on our data. Once we knew which restaurant attribute (ambience, food quality, and service quality) predicted overall rating the most, we aimed to understand exactly how much each attribute actually impacted rating. Multiple regression is a statistical test that examines how several different independent variables influence an outcome variable, while controlling for the others. Table 3 shows the multiple regression equation for each cuisine.

Italian rating = 1.4199 + (0.9007 x ambience) + (2.8557 x food) + (-0.0896 x service)
Mexican rating = 2.1392 + (0.244 x ambience) + (2.0157 x food) + (0.4537 x service)
Japanese rating = 1.5701 + (0.0468 x ambience) + (3.1632 x food) + (0.2342 x service)
Korean rating = 2.5808 + (1.6138 x ambience) + (0.6943 x food) + (-0.1006 x service)
Vietnamese rating = 1.5015 + (0.7559 x ambience) + (2.8253 x food) + (-0.0701x service)

Table 3. Multiple regression equations across each cuisine.

The resulting equations show that food quality consistently has the strongest positive effect on customer ratings across all cuisines, with the largest impact observed in Japanese and Italian restaurants. Ambience also plays a meaningful role—particularly for Korean and Vietnamese restaurants—while the influence of service quality varies by cuisine and is sometimes negative. This suggests that customers may prioritize food and ambience more heavily when forming their overall impressions. Overall, these regressions demonstrate that although all three attributes matter, their relative importance differs notably across cuisine types.

For example, while food quality is a major predictor for almost all cuisines, it is weighted most heavily for Japanese restaurants (B = 3.16) and least for Mexican restaurants (B = 2.0157). Additionally, Japanese restaurants appear to face less pressure regarding ambience, as reflected by their low ambience (B = 0.0468), indicating that customers judge these establishments primarily on food quality rather than atmosphere.

Discussion

To summarize, our goal was to understand how customer satisfaction differs among restaurants in Orange County, California, by analyzing Yelp reviews across three key attributes: food quality, ambience, and service quality. Using a combination of statistical methods, including ANOVA, simple regression, and multiple regression, we addressed two main questions: (1) Do these attributes differ significantly across cuisines? and (2) Which attributes most strongly influence a restaurant's overall rating, and does this influence vary by cuisine?

Our results showed that the five cuisines did not differ significantly in their overall Yelp ratings, nor in the average proportion of positive reviews mentioning food quality or service. The only significant difference we found was in ambience: Mexican restaurants had a higher proportion of positive ambience-related reviews compared to Korean restaurants, with a difference of about 0.16.

The regression analyses further highlighted that food quality is the strongest driver of overall ratings for nearly all cuisines. The one exception was Korean restaurants, where ambience played a larger role, consistent with the ANOVA findings. Although the degree of influence varied by cuisine, food quality was consistently a significant predictor of higher ratings across the board.

Business Implications

These findings offer important insights for restaurant owners and managers in Orange County. First, because food quality is the strongest driver of overall ratings across nearly all cuisines, restaurants should prioritize investments in ingredients, menu development, and kitchen staff training. Enhancing the quality of food is likely to yield the greatest improvement in customer satisfaction and online ratings.

Second, Korean restaurants may benefit from allocating more resources toward enhancing ambience. Since ambience played a uniquely strong role in both regression and review-based analyses for this cuisine, improvements in interior design, decor, lighting, and music could positively influence customer ratings.

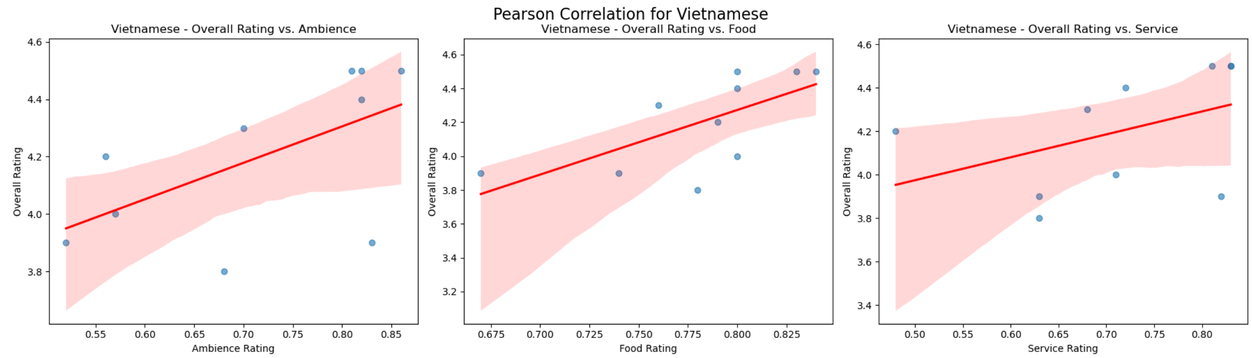
Limitations

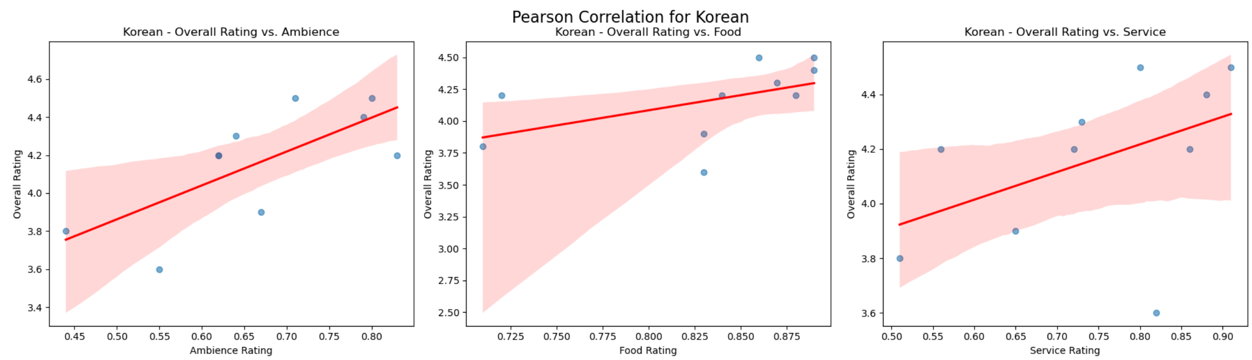
The first and most apparent limitation in our project is the size of our dataset. Because we manually collected attribute-level summary data from Yelp rather than full review texts, our final data set included only 50 restaurants (10 per cuisine). This gave us a broad overview across cuisines; however, the information for each restaurant was limited to aggregating proportions (e.g., proportion of positive reviews mentioning food, ambience, or service), not individual review data. Since this makes it difficult to capture the full variability in customer experiences, it restricts the statistical power of our analysis. If this project were to move forward, it would be valuable to gather a more robust data set, ideally thousands of individual reviews across the same five cuisines. This would allow for more accurate sentiment analysis and increase reliability of our findings.

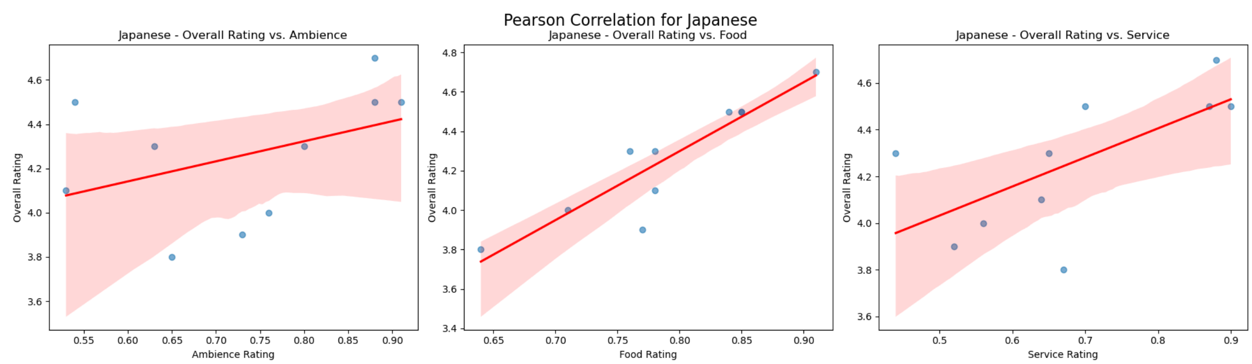
Secondly, because our data focused solely on restaurants in Orange County, our conclusions regarding ambience, food quality, and service quality cannot be generalized to restaurants or customer reviews in other areas of California or the United States. Orange County is home to a unique and diverse population; therefore, customers in other regions may prioritize certain factors differently.

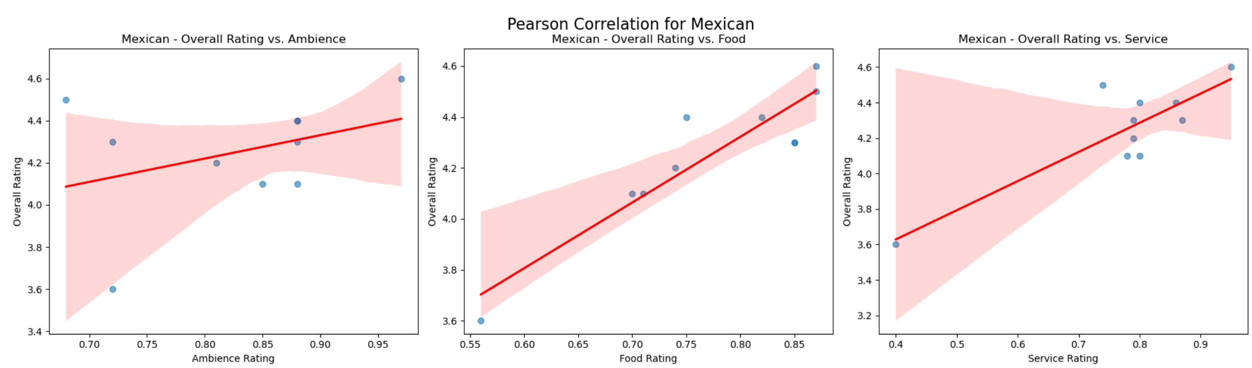
Lastly, although we considered three main restaurant attributes, there are more factors that influence a customer's experience. For example, wait time, price, and region all play important roles in customer satisfaction and Yelp ratings. Therefore, if we were to continue this project, it would be valuable to incorporate these additional factors.

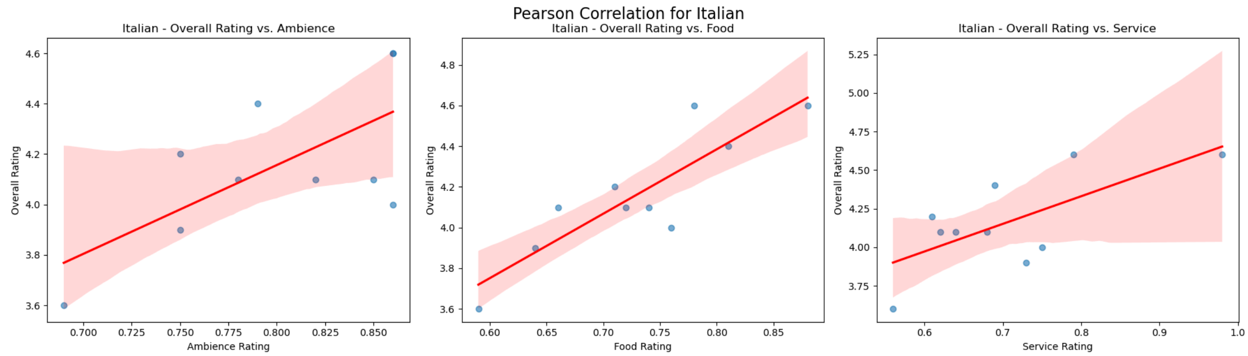
Appendix A: Additional Figures

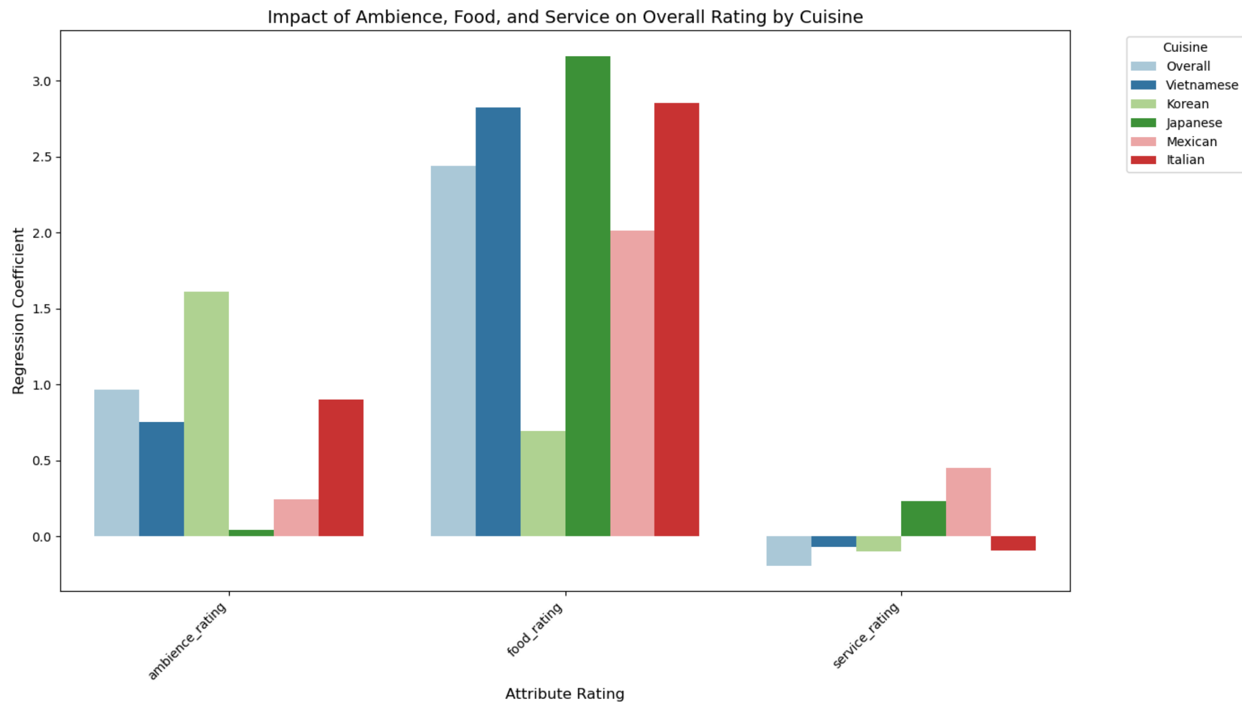




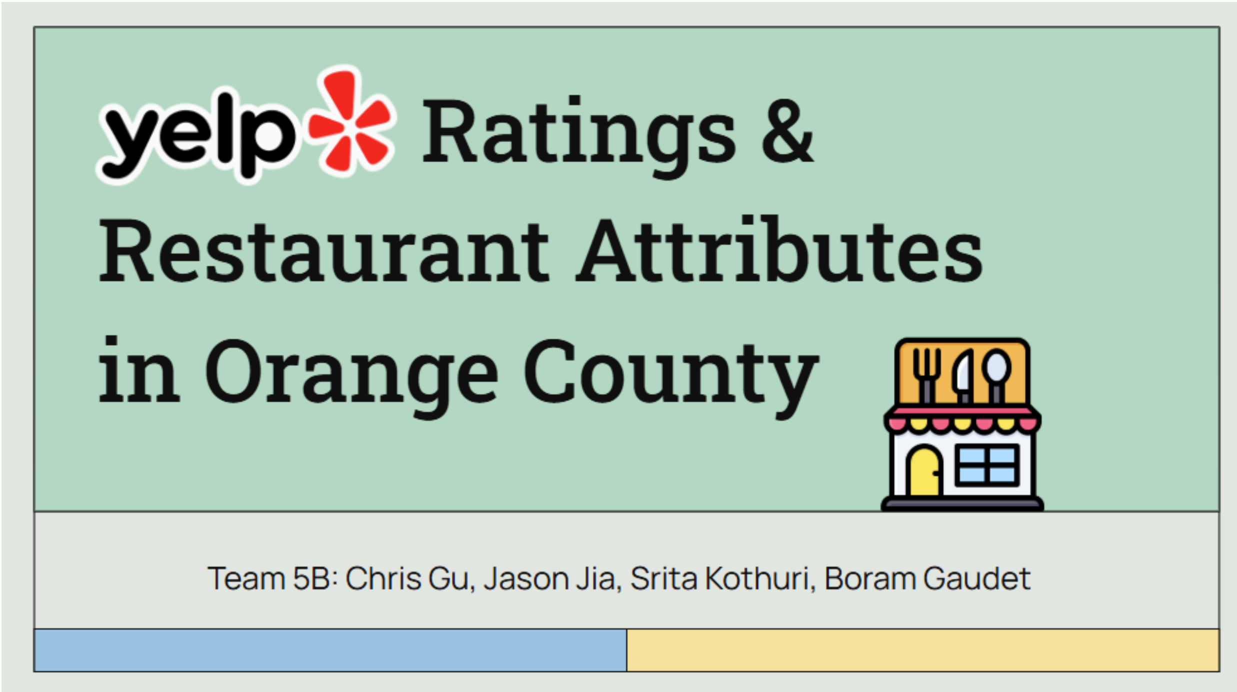


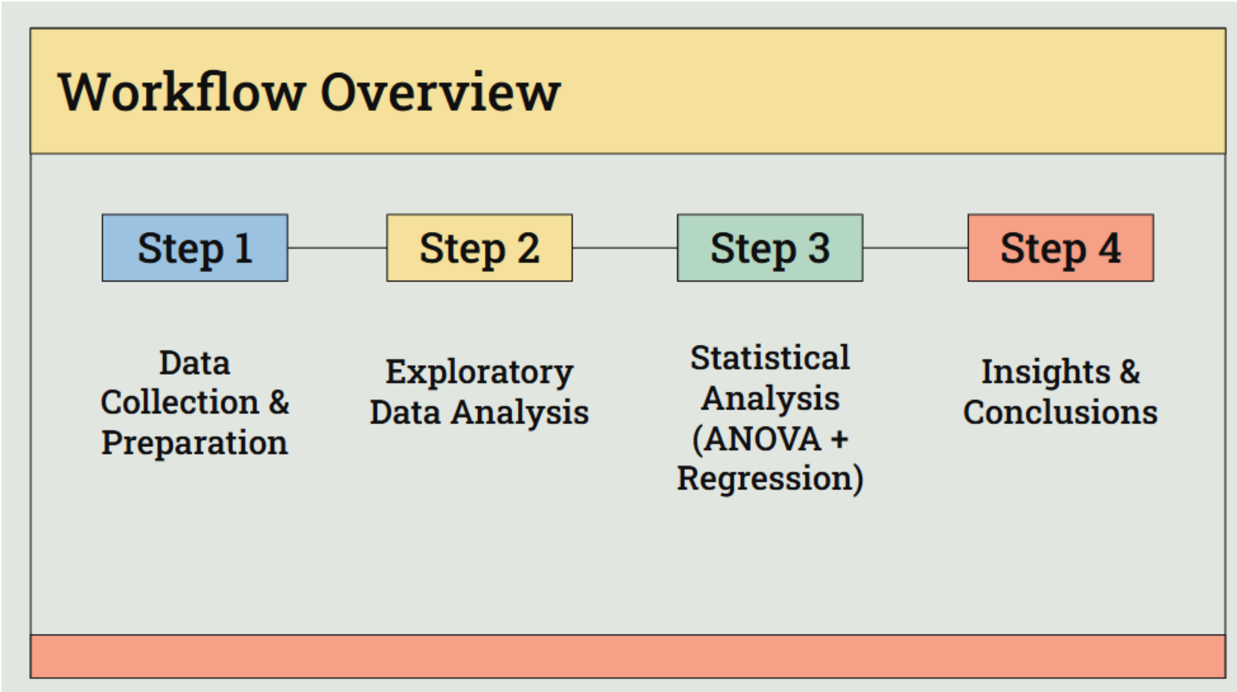


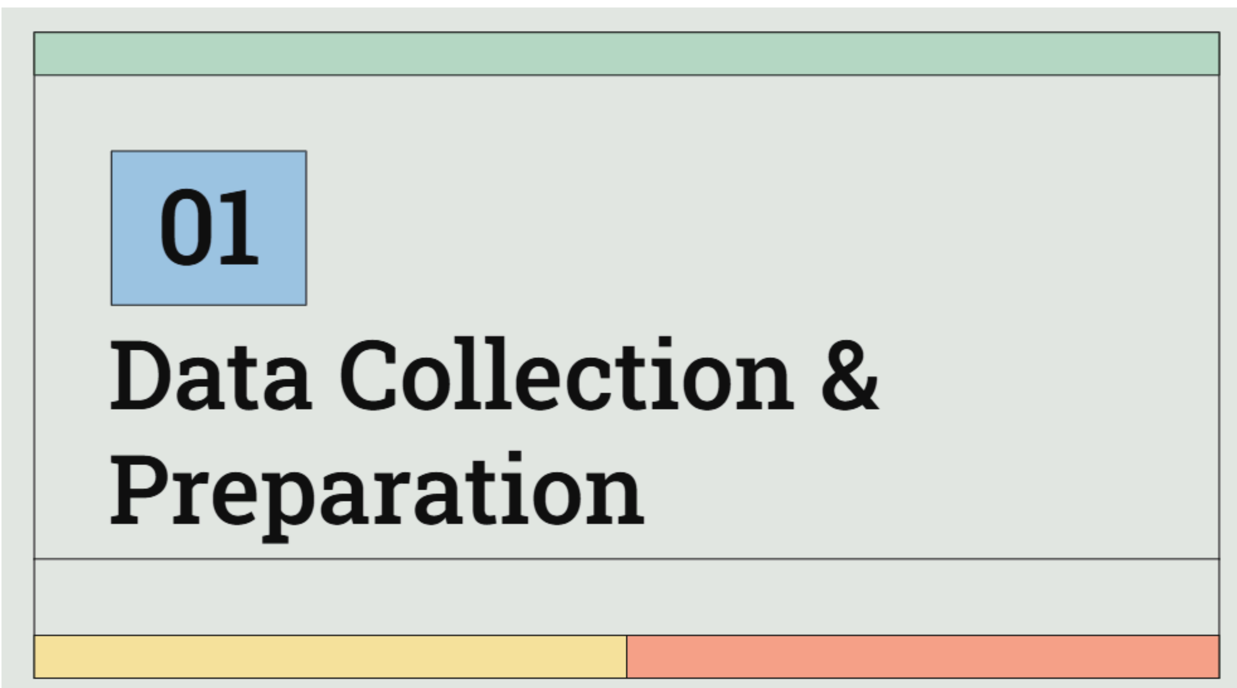

Pearson Correlation for Italian
Italian - Overall Rating vs. Ambience
Italian - Overall Rating vs. Food
Italian - Overall Rating vs. Service
Overall Rating
Ambience Rating
Food Rating
Service Rating


Impact of Ambience, Food, and Service on Overall Rating by Cuisine
Regression Coefficient
Attribute Rating
ambience_rating
food_rating
service_rating
Cuisine
Overall
Vietnamese
Korean
Japanese
Mexican
Italian

Appendix B: Presentation Slides

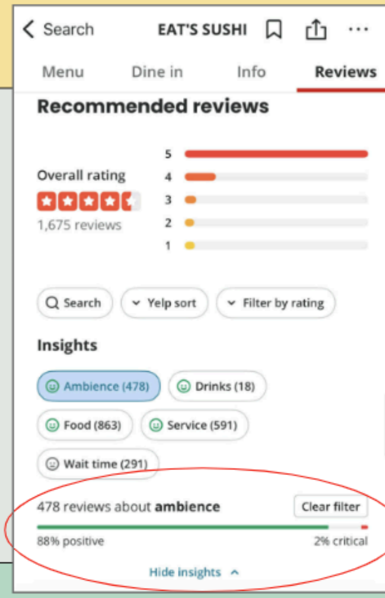



Workflow Overview
Step 1
Step 2
Step 3
Step 4
Data Collection & Preparation
Exploratory Data Analysis
Statistical Analysis (ANOVA + Regression)
Insights & Conclusions


01
Data Collection & Preparation

Data Collection & Preparation

- Selected 10 Yelp restaurants per cuisine (Vietnamese, Korean, Japanese, Mexican, Italian)
  - 50 restaurants total
- Collected basic Yelp details: city, price range, review count, overall rating
- Pulled attribute data from Yelp's Review Insights (mentions within positive reviews):
  - Ambience
  - Food quality
  - Service
- Compiled the dataset into Excel, then loaded into Python for analysis



Dataset Snapshot

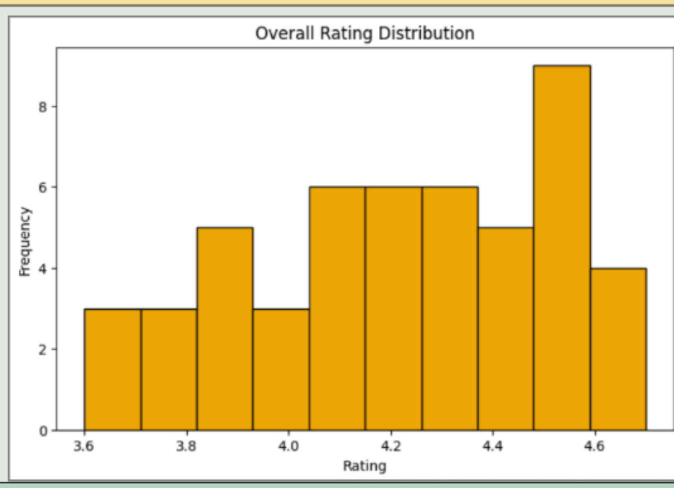
index	name	cuisine	location	price	overall_num_review	overall_rating	ambience_num_review	ambience_rating	food_num_review	food_rating	service_num_review	service_rating
0	VOX Kitchen	Vietnamese	Fountain Valley	$$	7964	4.5	846	0.81	1677	0.84	1194	0.83
1	NEP Café	Vietnamese	Fountain Valley	$$	4537	4.5	1035	0.86	1870	0.83	1314	0.83
2	Brodard Restaurant	Vietnamese	Fountain Valley	$$	3924	3.8	423	0.68	908	0.78	583	0.63
3	Pho 79 Restaurant	Vietnamese	Garden Grove	$$	3769	4.2	199	0.56	581	0.79	381	0.48
4	Oc & Lau Restaurant	Vietnamese	Graden Grove	$$	3273	4	267	0.57	560	0.8	363	0.71
5	Sup Noodle Bar	Vietnamese	Irvine	$$	3074	4.3	350	0.7	947	0.76	662	0.68
6	Nguyen's Kitchen	Vietnamese	Orange	$$	2714	3.9	99	0.52	242	0.67	132	0.63
7	Nep Café	Vietnamese	Irvine	$$	1759	4.5	949	0.82	1531	0.8	1111	0.81
8	Pho Ba Co	Vietnamese	Irvine	$$	1752	3.9	51	0.83	123	0.74	98	0.82
9	Au Lac	Vietnamese	Fountain Valley	$$	1743	4.4	51	0.82	111	0.8	84	0.72
0	Bakjeong	Korean	Irvine	$$$	5472	4.5	462	0.8	1248	0.86	1360	0.91
1	BCD Tofu House	Korean	Irvine	$$	3519	3.9	167	0.67	428	0.83	319	0.65
2	All That Barbecue	Korean	Irvine	$$$	2933	3.8	118	0.44	219	0.71	210	0.51
3	Mo Ran Gak Restaurant	Korean	Garden Grove	$$	2636	4.5	165	0.71	497	0.89	337	0.8

- ambience_rating = Proportion of positive reviews mentioning restaurant's ambience
- food_rating = Proportion of positive reviews mentioning restaurant's food quality
- service_rating = Proportion of positive reviews mentioning restaurant's service

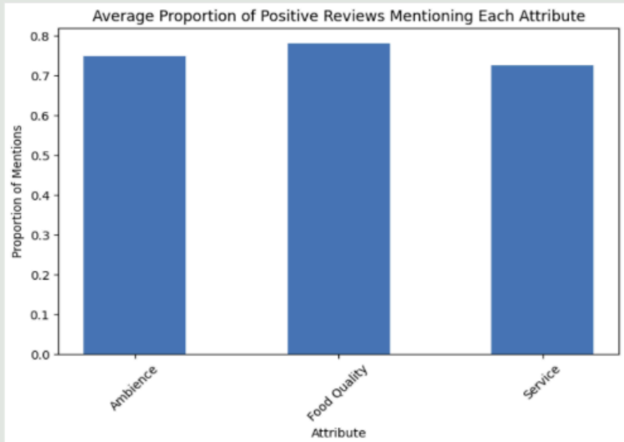
02

Exploratory Data Analysis

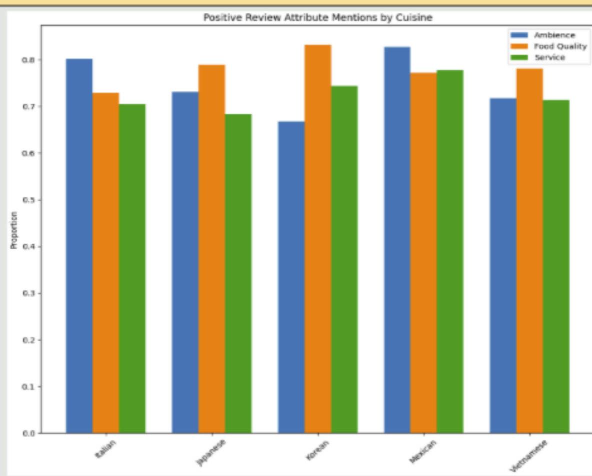
Overall Restaurant Rating Distribution



Attribute Mention Proportions (Overall)



Positive Review Attribute Mention by Cuisine



We will break each of these attributes down in the following slides

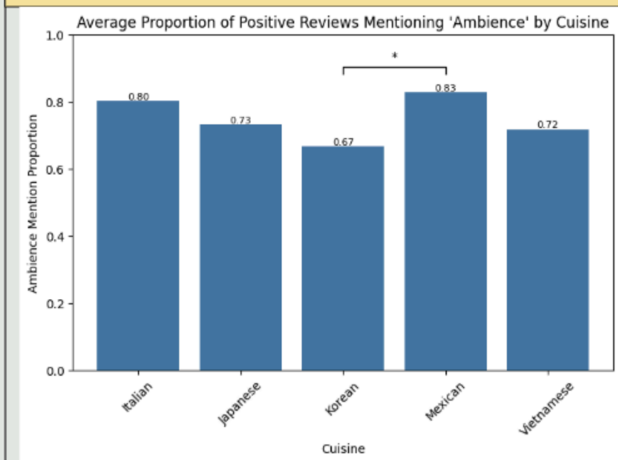
03

Statistical Analysis

Statistical Analysis: ANOVA

- ANOVA approach
  - ANOVA run on each attribute
  - Two-tailed test
  - Purpose: Assess if average proportion of attribute mentions differed **significantly** between cuisines (pairwise)
- Assumptions & Prep
  - Normality checked with Shapiro Wilks Test
  - Ambience attribute was not normally distributed
    - Applied arcsine square root transformation
- Why ANOVA?
  - Identify attribute differences across cuisines
  - Highlights what customers notice most per cuisine

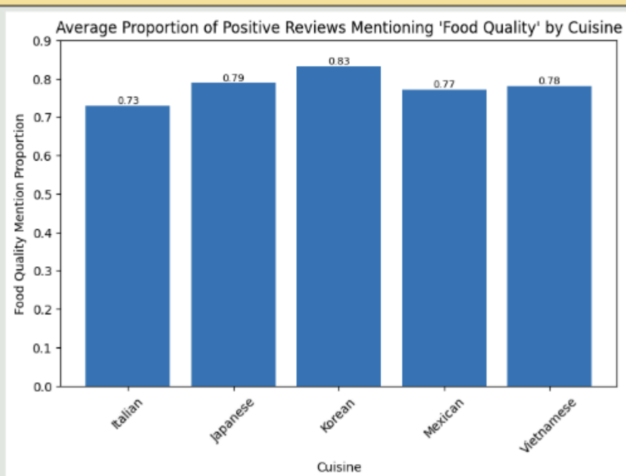
ANOVA: Ambience Rating



- **Significant difference** in proportion of reviews mentioning ambience by cuisine
  - F(2,50) = 3.429 , p =0.015
- Post Hoc (Tukey): Significant difference only between Korean vs. Mexican
  - p=0.021

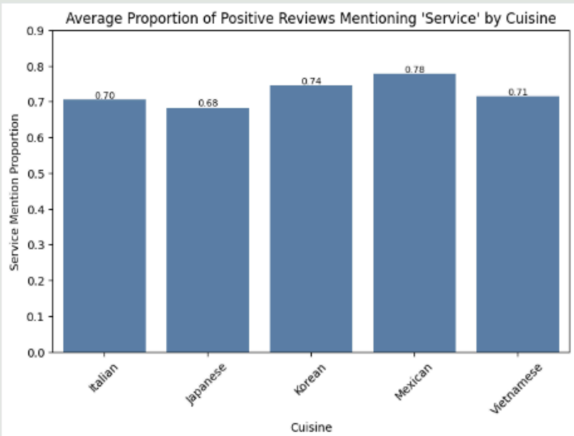
Note: this is the transformed data

ANOVA: Food Quality



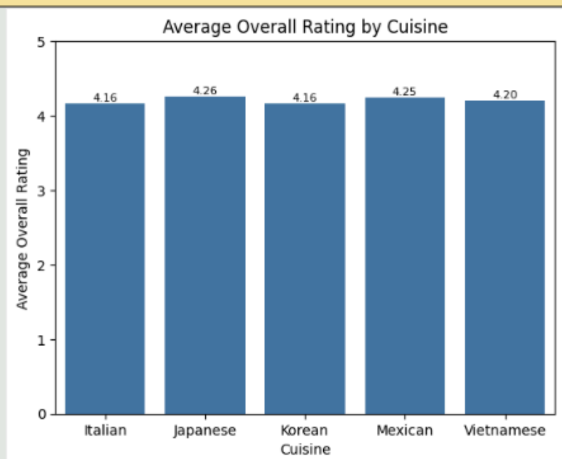
- **No significant difference** in proportion of reviews mentioning food quality by cuisine
  - F(2,50) = 2.27 , p =0.076

ANOVA: Service Rating



- **No significant difference** in proportion of reviews mentioning service by cuisine
  - F(2,50) = 0.74 , p =0.569

ANOVA: Overall Rating By Cuisine



- **No significant difference** in overall ratings between each cuisine
  - F(2,50) = 0.26 , p =0.90

Statistical Analysis: Regression

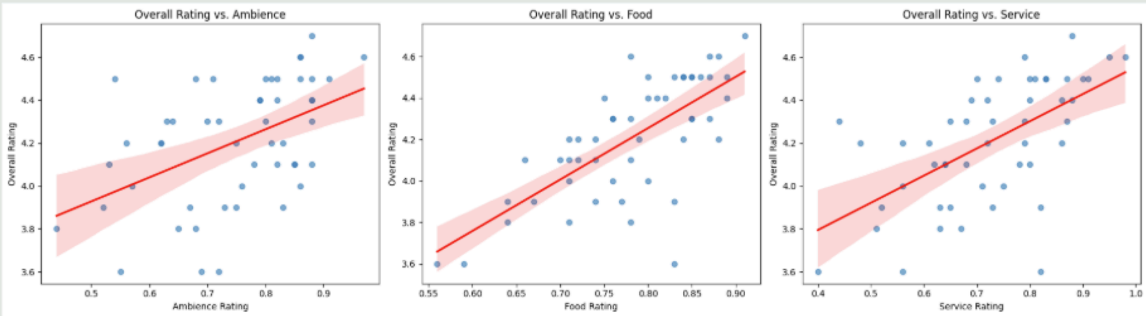
- Goal: Understand which attribute explains the most variance in overall rating of a cuisine

Regression: Each Cuisine

Italian	Japanese	Korean	Vietnamese	Mexican
Ambience: $r^2 = 0.45$*	Ambience: $r^2 = 0.18$	**Ambience: $r^2 = 0.50$***	Ambience: $r^2 = 0.35$	Ambience: $r^2 = 0.14$
Food Quality: $r^2 = 0.76$**	**Food Quality: $r^2 = 0.83$*****	Food Quality: $r^2 = 0.26$	**Food Quality: $r^2 = 0.45$***	**Food Quality: $r^2 = 0.85$*****
Service: $r^2 = 0.48$*	Service: $r^2 = 0.44$*	Service: $r^2 = 0.21$	Service: $r^2 = 0.18$	Service: $r^2 = 0.72$**

*indicates p-value < 0.05 ; ** indicates p-value < 0.01; *** indicates p-value < 0.001

Regression: Overall



$r^2 = 0.23^{***}$ | $r^2 = 0.50^{***}$ | $r^2 = 0.35^{***}$

*indicates p-value < 0.05 ; ** indicates p-value < 0.01; *** indicates p-value < 0.001

04

Insights & Conclusions

Insights & Conclusions

1) No difference in average overall ratings across cuisines
2) Significant higher proportion of ambience mentions in **Mexican** restaurants vs. **Korean** restaurants
    a) Suggests that **ambience** is sought after in Mexican restaurants
3) Food Quality explains most variation in overall rating (4/5 cuisines)

Limitations

1) Limited sample size and reliance on summary-level Yelp data
2) Sample restricted to Orange County restaurants
    a) Potential self-selection bias
3) Focused on only three main attributes (food, service, ambience)
    a) Many other factors may influence restaurant experience

Thank You

Bibliography

Eater LA. (2025, October 31). The 38 essential Orange County restaurants. https://la.eater.com/maps/best-essential-restaurants-orange-county-california