



Are Yelp's First Reviewers' Biased?





Sravanthi Prattipati



Ajinkya Bhavik

Meet the Team

1

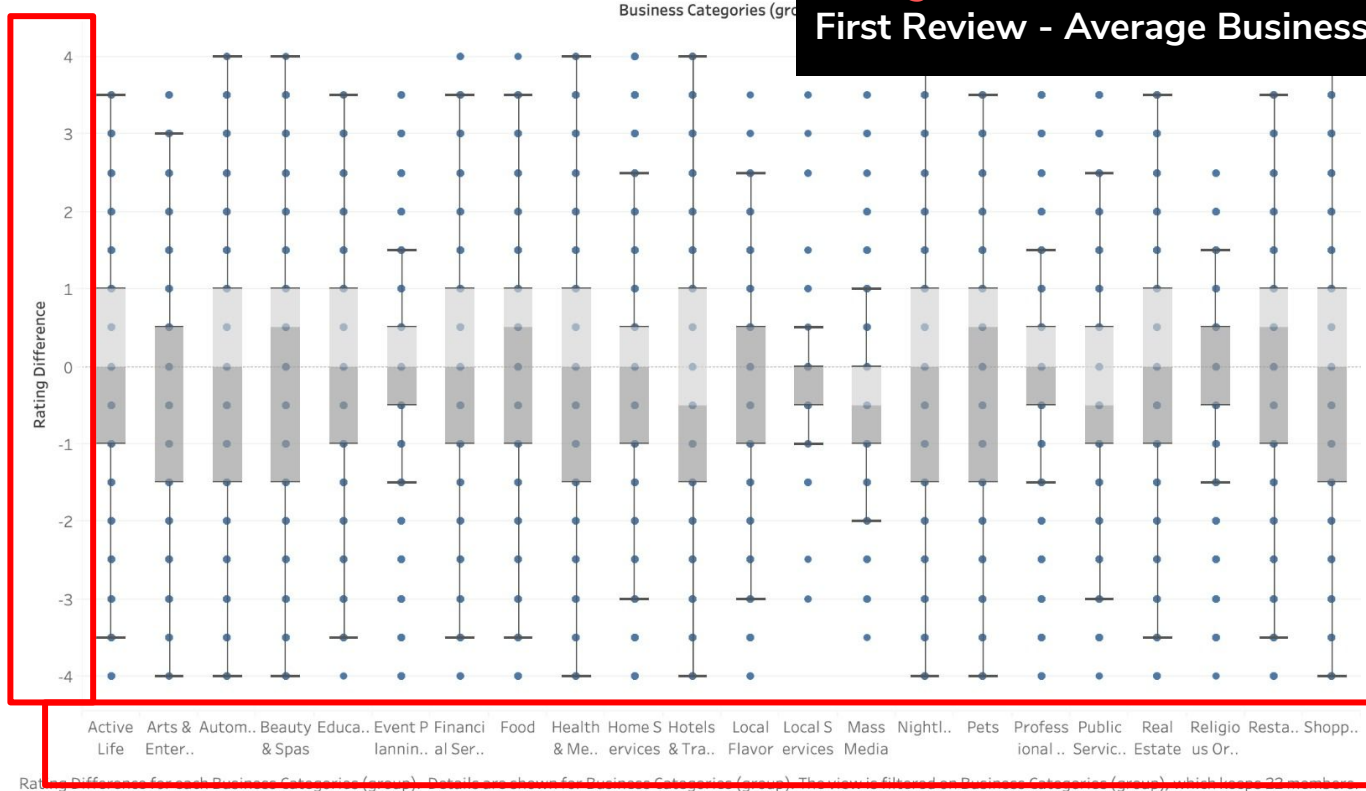
Question



“

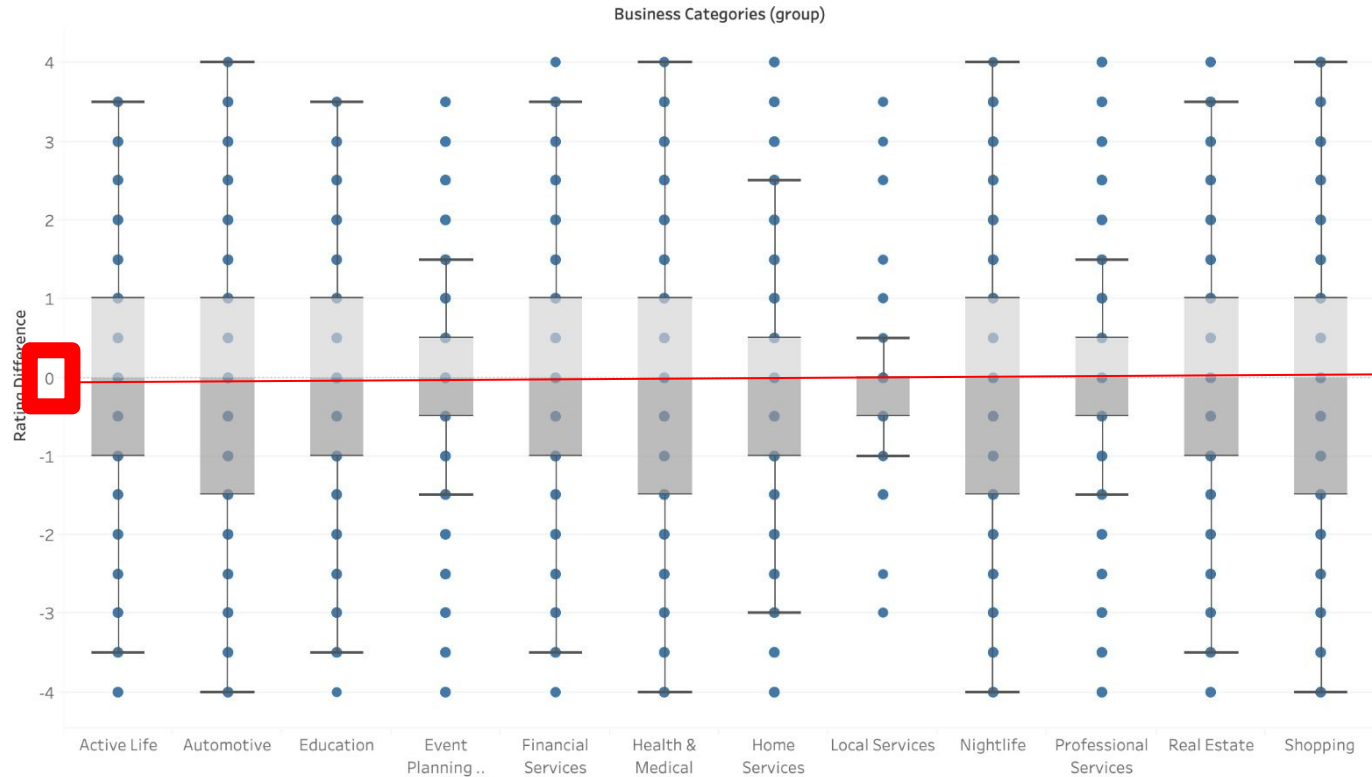
How does Yelp's first reviews compare to the average reviews of the business?

Rating Difference: First Review - Average Business Rating



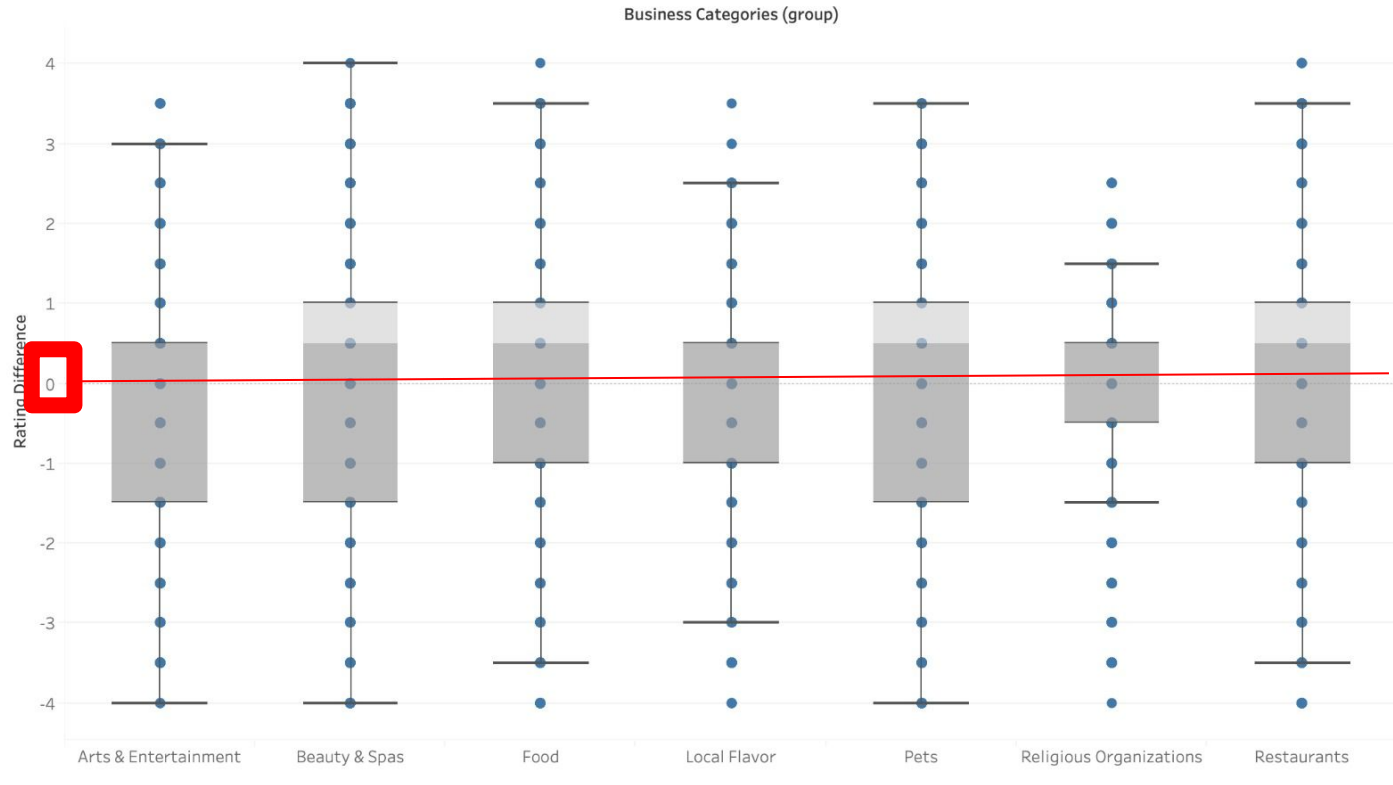
First Review Rating Difference over Different Categories

Rating Difference exactly zero



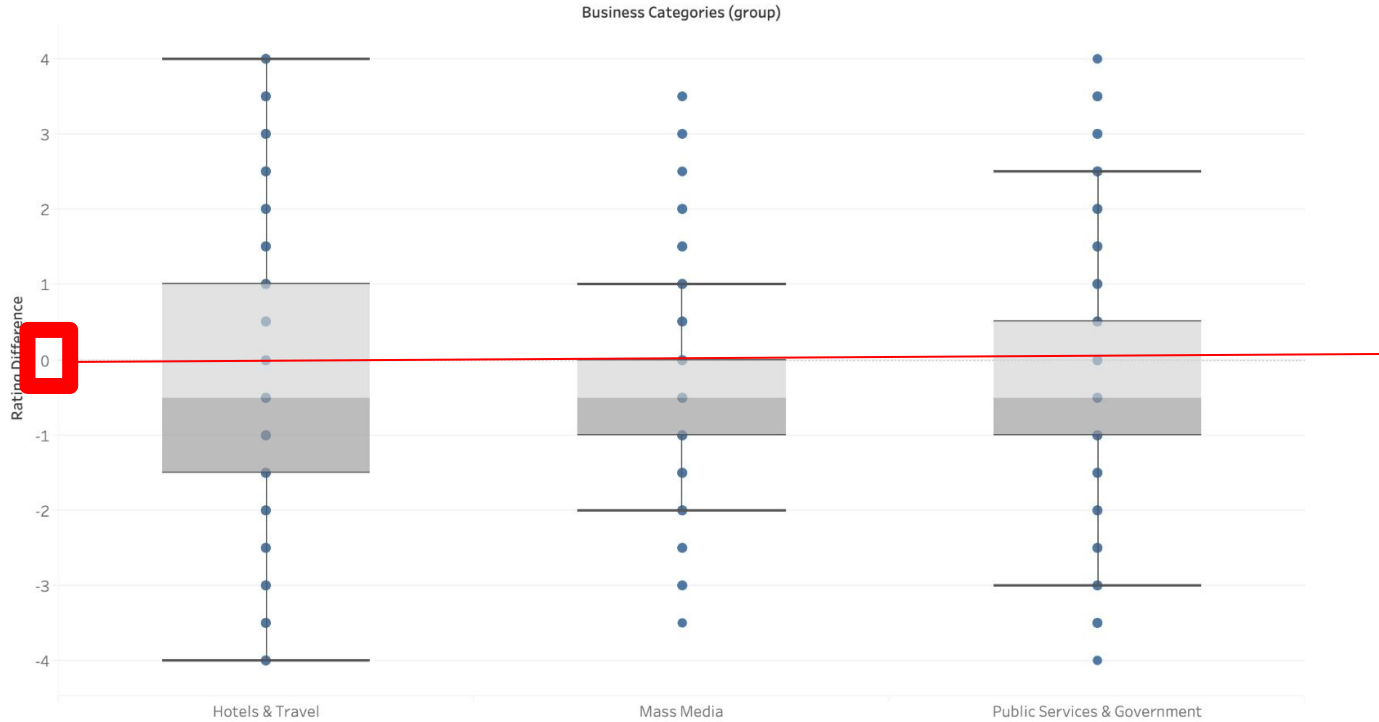
Median = 0; (First Review = Average Review for the Business)

Rating Difference Above



Median > 0; (First Review > Average Review for the Business)

Rating Difference Below



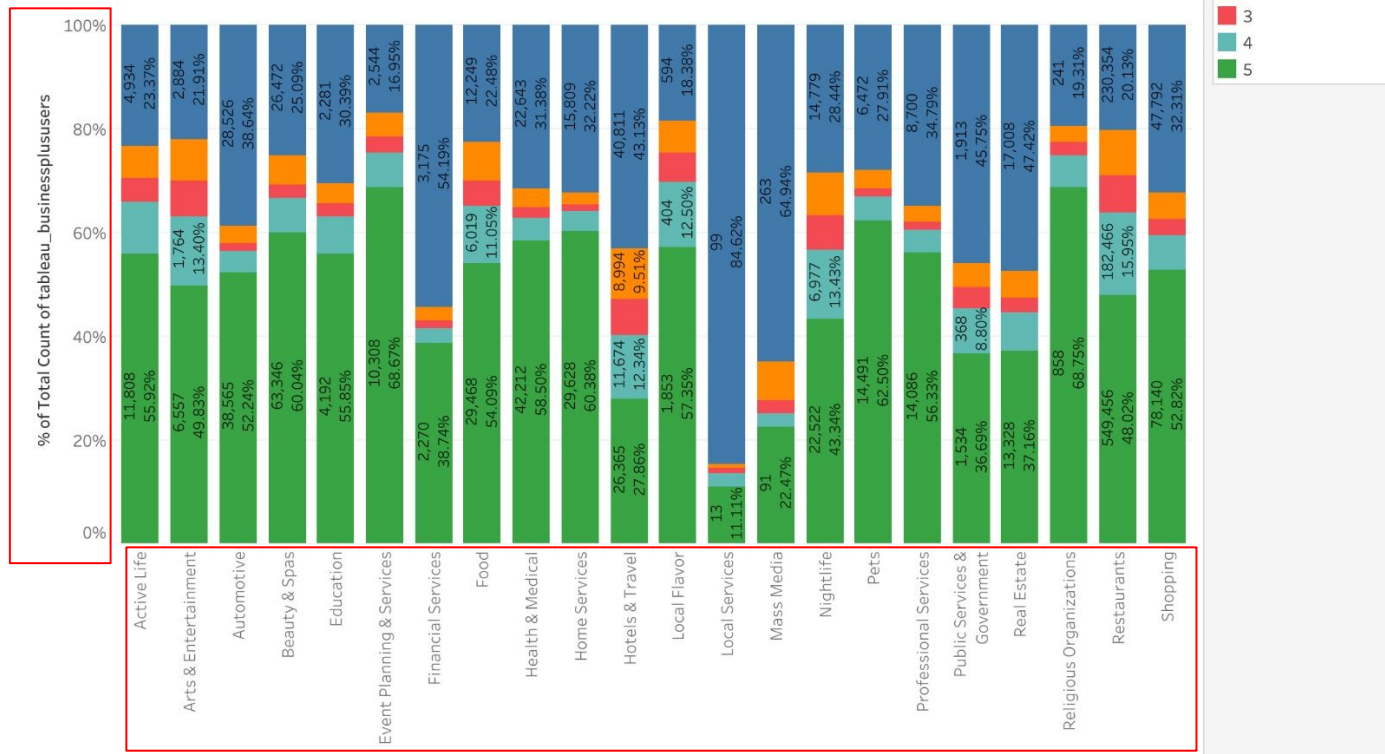
Rating Difference for each Business Categories (group). Details are shown for Business Categories (group). The view is filtered on Business Categories (group), which keeps Hotels & Travel, Mass Media and Public Services & Government.

Median < 0; (First Review < Average Review for the Business)

“

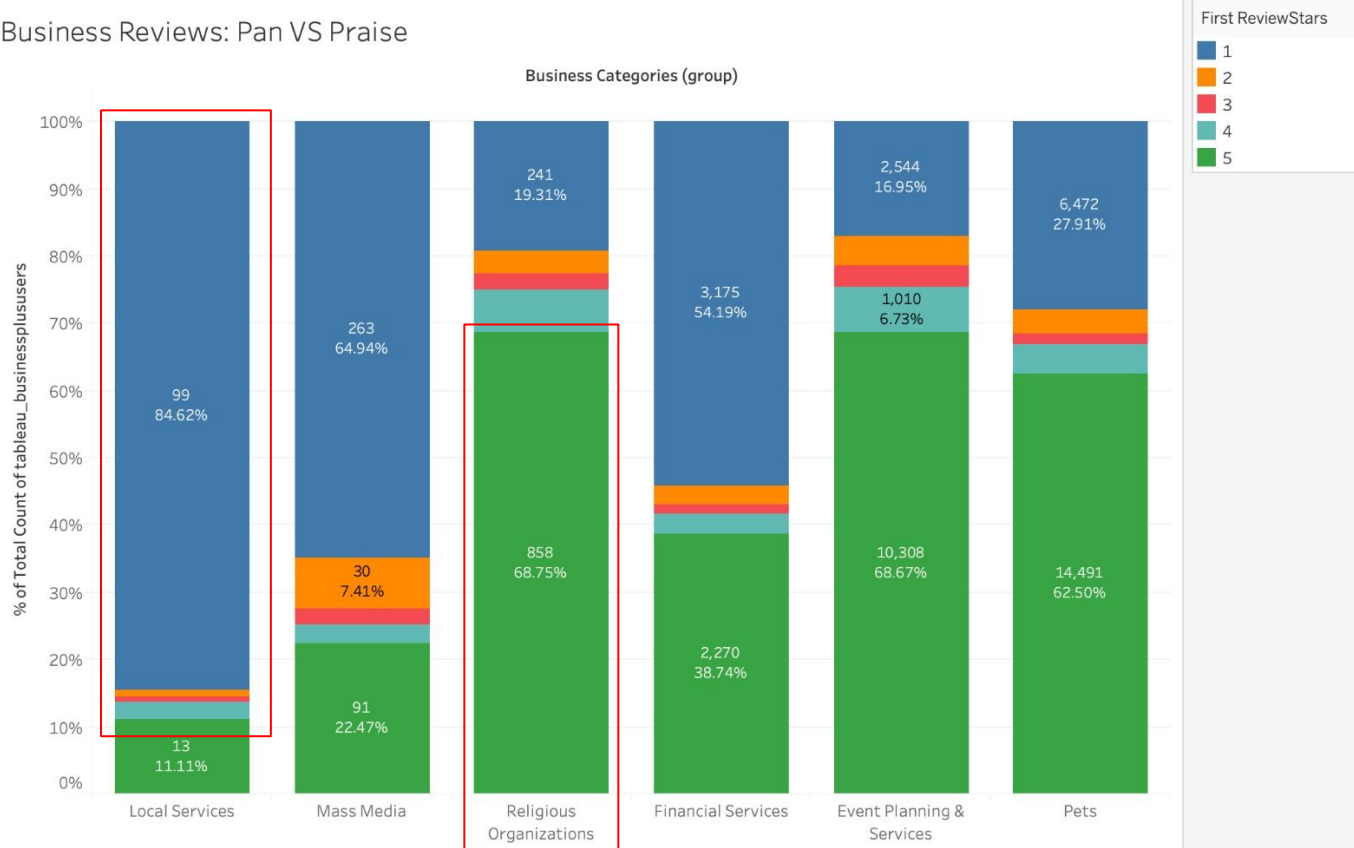
What categories do people review for their first review?

Business vs Percentage Spread of 1st Reviews



First Review Rating Spread Across Different Business Categories

Business Reviews: Pan VS Praise



Top 3 Categories where users join to pan/praise a business

2

Limitations



Databricks Limitations

- Since yelp data is spread out around different countries we were only able to take the metro areas.
- Maintaining multiple tables while switching between pyspark and spark sql was a little time consuming.
- Data Integrity was a bit difficult since the data sizes for different categories were varying but we solved these issues in Tableau
- While attaching the tableau link, databricks didn't allow to show all sheets as an image, but only showed the one which is put a header images in the tableau desktop.



Tableau Limitations

- Tableau took forever to run because of low processing power and higher amounts of data
- Each business is tagged with multiple categories. When we grouped data one business had to go under only one category even if it originally could coexist under several categories. We decided to group them on tableau instead of databricks
- Spinning up and maintaining of clusters esp if connected with tableau since databricks still considered the cluster to be idle. Thus terminating while viz is in process in tableau.



3

Why
should
Yelp care?



“

Are first reviews
accurate
representation of
how the business is
doing?

Analysis

- Overall, it is accurate for:
 - Active life;
 - Automation;
 - Education;
 - Event planning;
 - Financial services;
 - Health and Medical;
 - Home services;
 - Local services
 - Nightlife;
 - Professional Services;
 - Real estate;
 - Shopping



Analysis

- Overall, it is overrated for:
 - Arts and Entertainment;
 - Beauty and Spas;
 - Food;
 - Local flavor;
 - Pets;
 - Religious organization;
 - Restaurant;



Analysis

- Overall, it is underrated for:
 - Hotels and Transportation
 - Mass media
 - Public service



4

What should
Yelp do?



Recommendations



What should Yelp focus on?

Weightage

Yelp should consider to place a lower weight on the first reviews for the categories where the first reviewers tend to overrate or underrate to balance out the bias.

Focus

Yelp can focus on auditing/paying close attention to the businesses who are getting first review ratings that vary highly from their average to ensure that no incentives are being provided by the business for the users to rate the business in a certain way.

Market

Yelp could focus on adding more businesses to the categories that has least # of first reviews to expand into those markets.

Moving Forward

- Steps for future work:
 - Compare user's first review and business average reviews while adding elite users as a filter.
 - Praising and panning ratio in first reviews different for elite users?
 - Work hours timings or locations which affect the rating range of reviews from 1 to 5?



A woman with long brown hair is sitting at a table in a cafe, holding a fork with a small portion of food. She is looking down at the food with a slight smile. The table is set with a yellow mug, a small bowl of food, and some bread. There are green plants and warm lights in the background. The word "Thanks!" is overlaid in large red letters on the left side of the image.

Thanks!