Online Food Dataset (link)

Data

Our dataset contains information about online food orders. There are a total of 389 rows and 12 columns which encompasses demographic information about the customer, the location, and the order details. Each row corresponds to one order where various information is collected about the user and the order. The variables are:

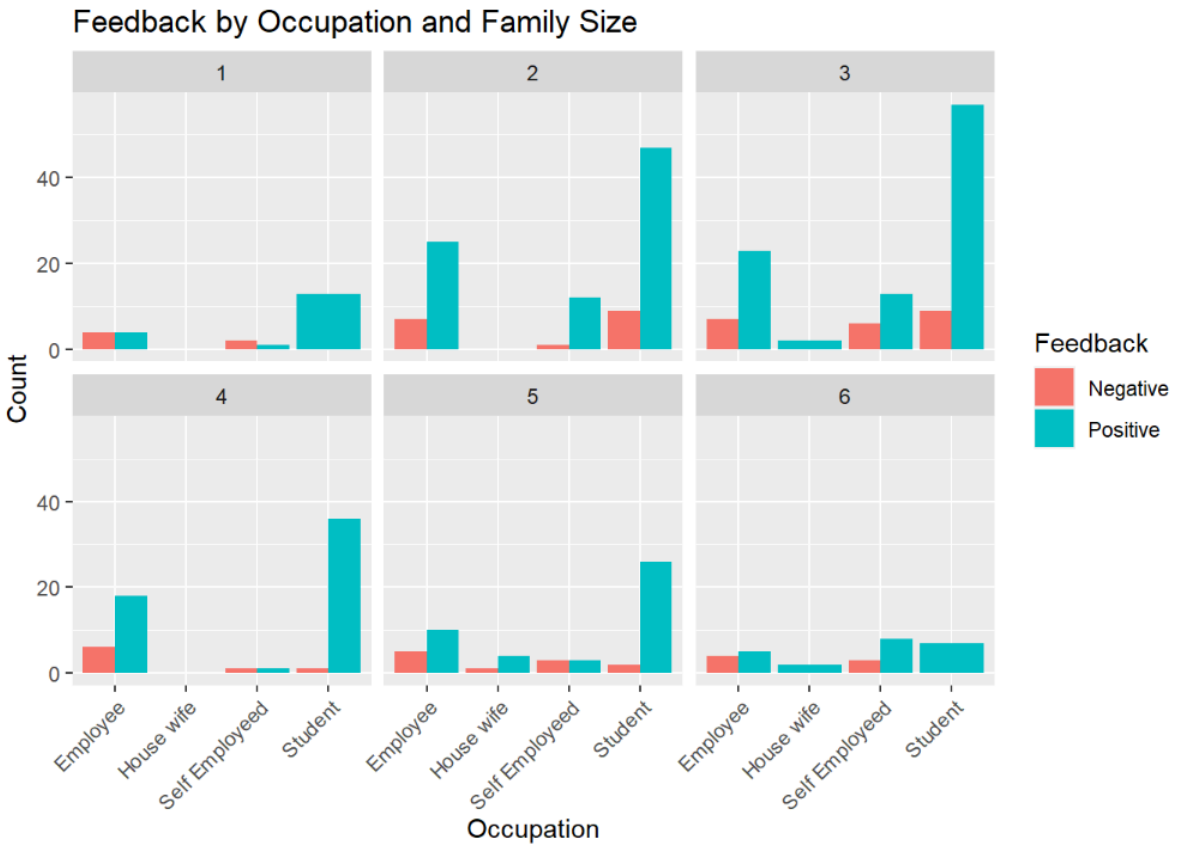
- ➔ Demographic Information
  - ◆ Age
  - ◆ Gender (Male or Female)
  - ◆ Marital Status (Single, Married, etc)
  - ◆ Occupation (Student, Employee, etc)
  - ◆ Monthly Income (No Income, Below 10000, 10001 to 25000, etc)
  - ◆ Educational Qualifications (Graduate, Post Graduate, etc)
  - ◆ Family Size
- ➔ Location
  - ◆ Latitude
  - ◆ Longitude
  - ◆ Pin Code: Pin code of the customer's location
- ➔ Order Details
  - ◆ Output: Received Order or not (Yes or No)
  - ◆ Feedback (Positive or Negative)

Research Questions

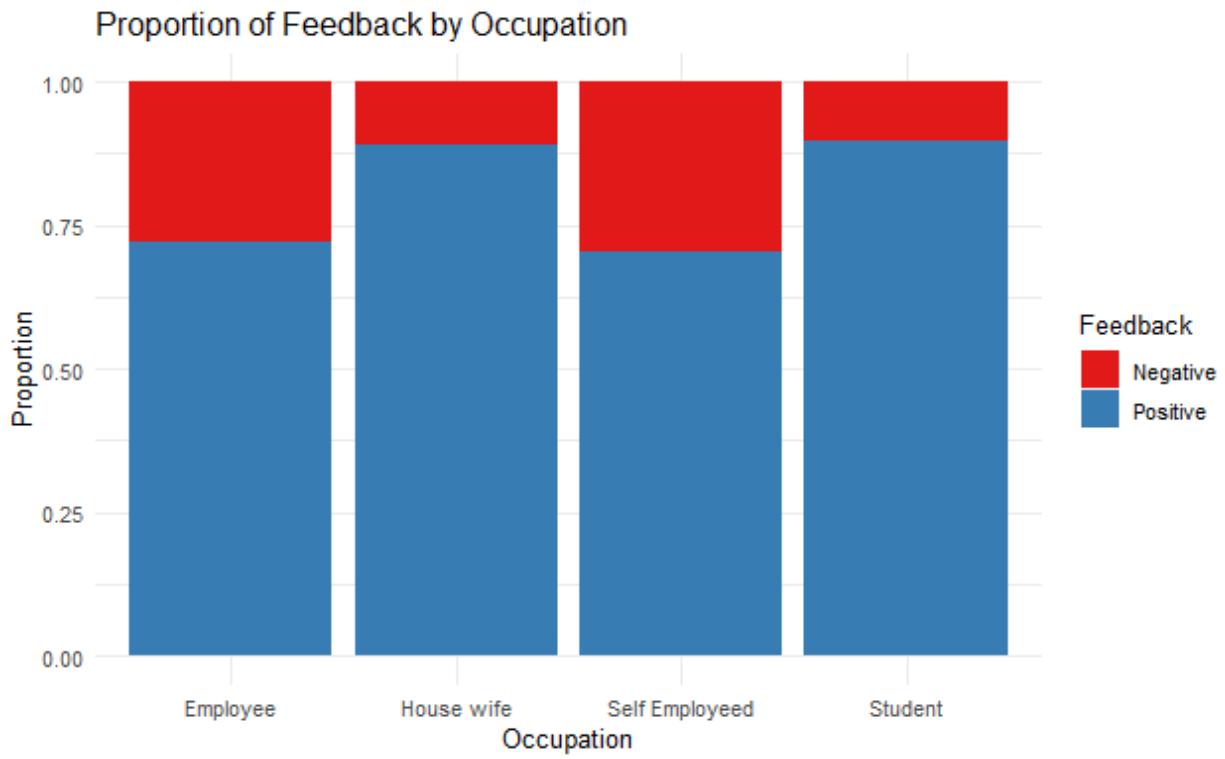
We are interested in answering these three research questions regarding our dataset:

- ➔ Is there a relationship between a customer's occupation and the feedback they provide?
- ➔ How does monthly income level affect the quality of feedback provided by customers?
- ➔ What are the geographical trends in online food ordering? Are certain pin codes or areas more likely to give positive feedback?

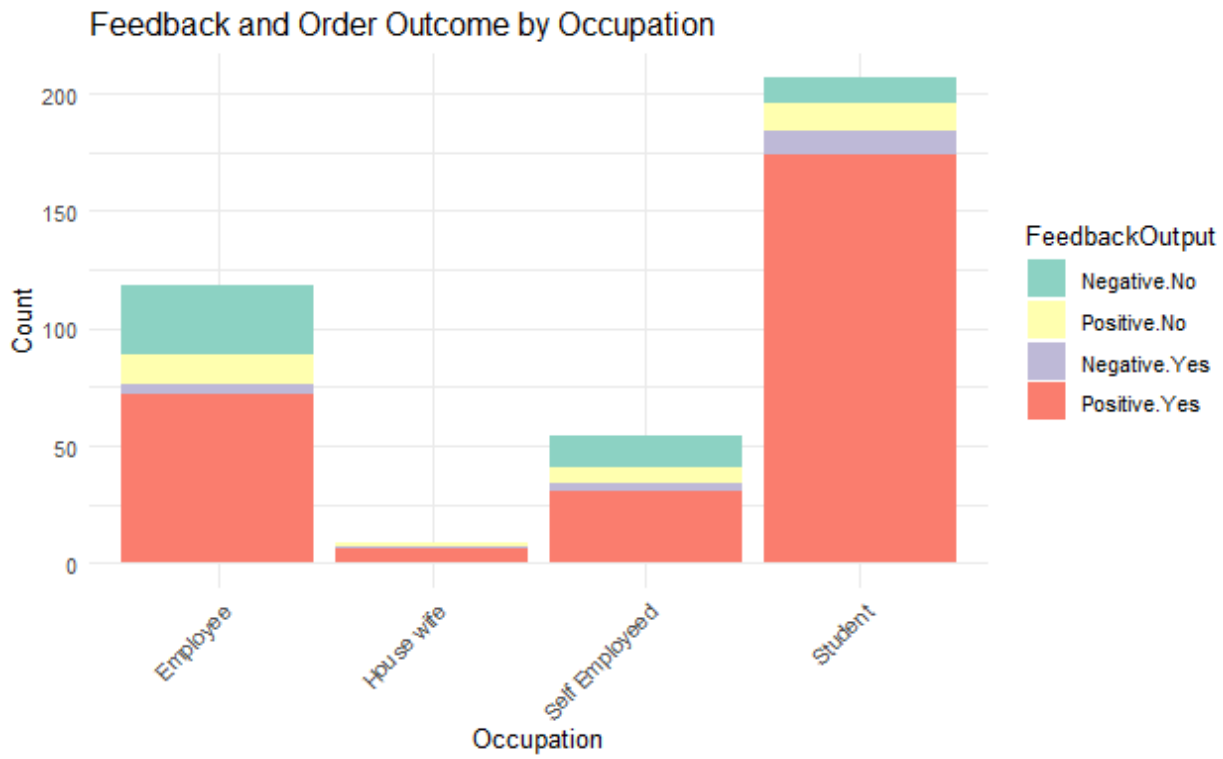
Graphical Analysis



Feedback by Occupation and Family Size: The faceted bar chart reveals that students tend to give more positive feedback compared to other occupations, regardless of family size. Additionally, as family size increases, there seems to be a slight trend towards more feedback, particularly positive. It's noteworthy that the ratio of positive to negative feedback remains fairly consistent across different occupations and family sizes, with positive feedback being dominant. This consistency suggests that occupation and family size, while influential on the volume of feedback, may not necessarily affect the sentiment of the feedback provided.



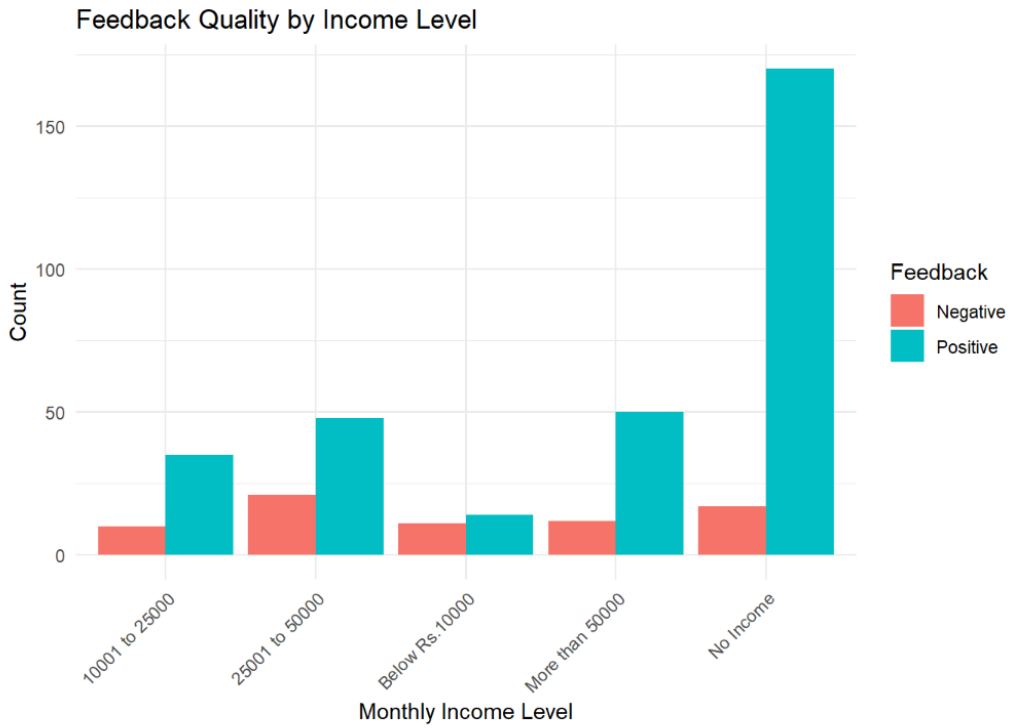
Feedback Proportion by Occupation: The proportion bar chart helps further understand the differences between occupations. It may be harder to compare occupations in the previous graph due to the differences in volume of feedback so comparing the proportion may help understand the differences. Based on the graph, we can see that the proportion of positive feedback does indeed remain fairly consistent across occupations. However, we can notice that students and housewives tend to have a higher proportion of positive feedback compared to employees and the self-employed.



Feedback and Order Outcome by Occupation: Although not part of our research question, another interesting variable to consider is the outcome variable which shows whether or not they have received their order. This may further explain the differences between occupations. There are four possible outcomes:

- Negative.No: Didn't receive order and negative feedback
- Positive.No: Didn't receive order and positive feedback
- Negative.Yes: Did receive order and negative feedback
- Positive.Yes: Did receive order and negative feedback

From the stacked bar chart, we can see that employees are more likely to give negative feedback when they didn't receive the order compared to students who still had about an equal rate of feedback even when they didn't receive the order. This further supports the idea that students are more likely to give positive feedback.



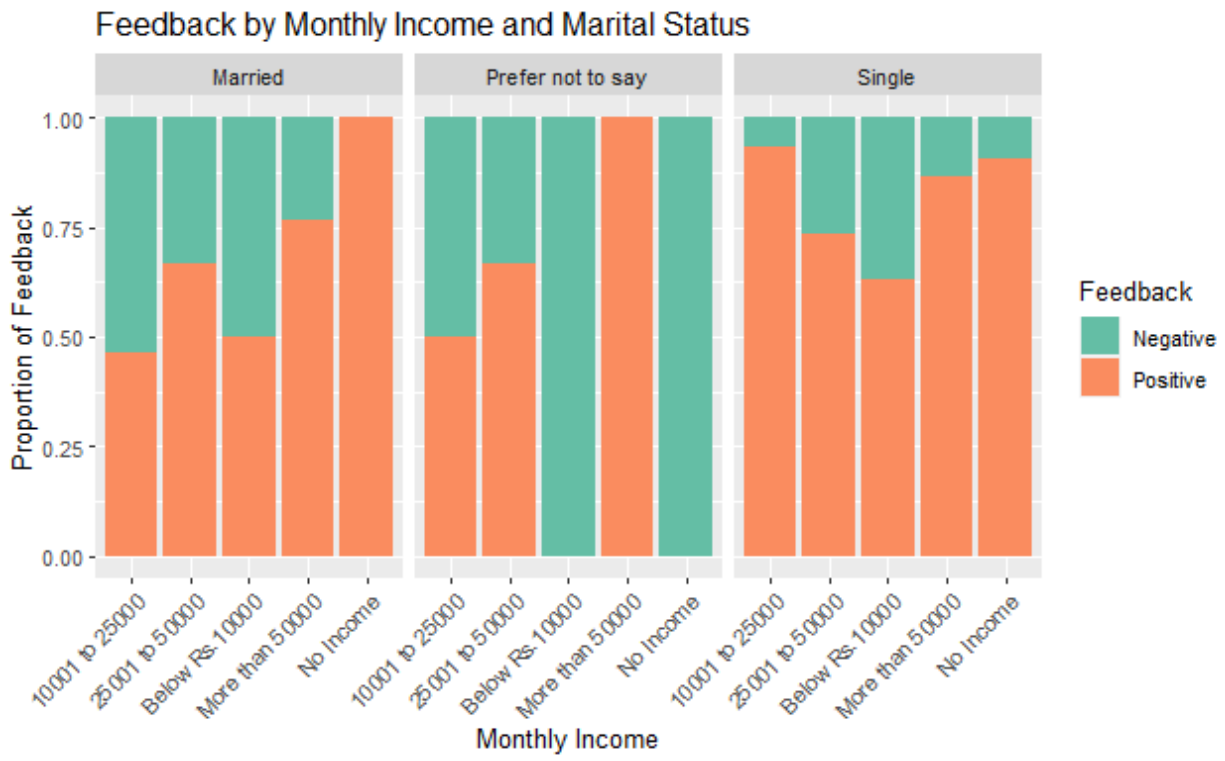
Feedback Quality by Income Level:

The bar chart suggests that positive feedback significantly outweighs negative feedback across all income levels, with the largest volume of feedback coming from customers with no income. This may indicate a high level of satisfaction among users of online food services or a tendency for more frequent users, such as students with no income, to leave feedback. However, the proportion of negative feedback appears slightly higher within the "Below Rs.10000" income group compared to others.

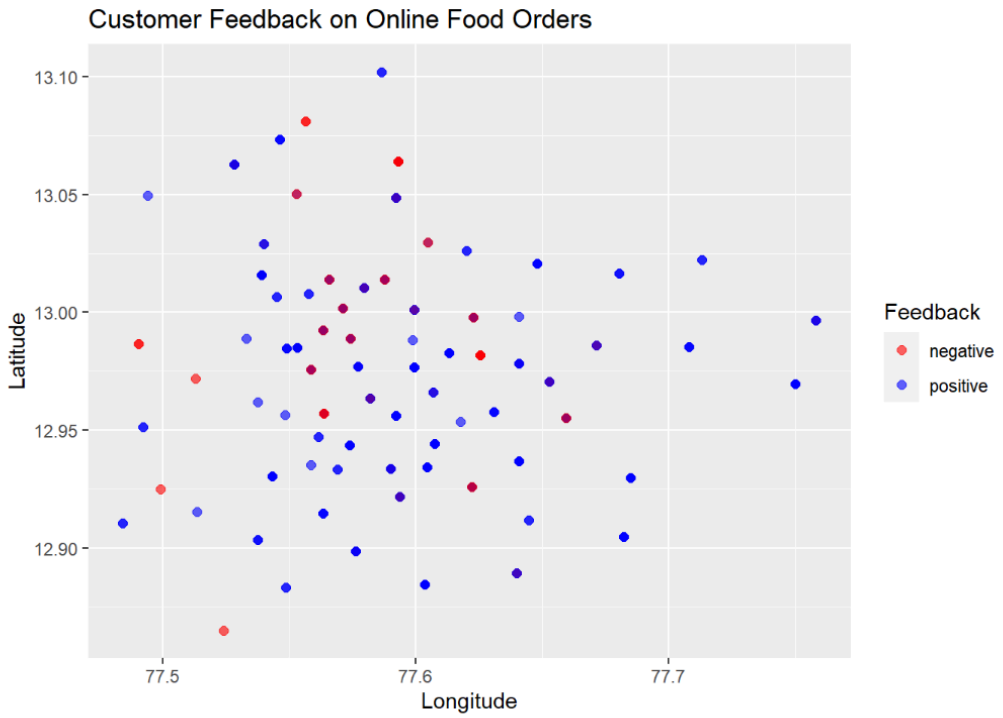
```
##
##  Pearson's Chi-squared test
##
## data:  feedback_income_table
## X-squared = 28.958, df = 4, p-value = 7.974e-06
```

Chi-square test on Feedback Quality by Income Level:

The p-value is much smaller than the common alpha level of 0.05, which suggests that there is a statistically significant association between Monthly Income and Feedback. In other words, the data provide strong evidence that the distribution of feedback (positive or negative) is not the same across different income levels.



Feedback Quality by Income Level and Marital Status: The faceted bar chart of proportion across income levels, separated by marital status, further reveals information about differences between income levels while including marital status which could have an interaction effect. As shown by the graph, we can see that the trends discovered in the previous graphs stay consistent with a few exceptions. For example, although the 'No Income' group tends to have an overwhelmingly positive feedback in 'Married' and 'Single' groups as we have discussed, in the 'Prefer not to say' group, we can see that the 'No Income' group is made up of entirely negative feedback. This potentially suggests that there is a more complex relationship which involves other variables such as marital status.

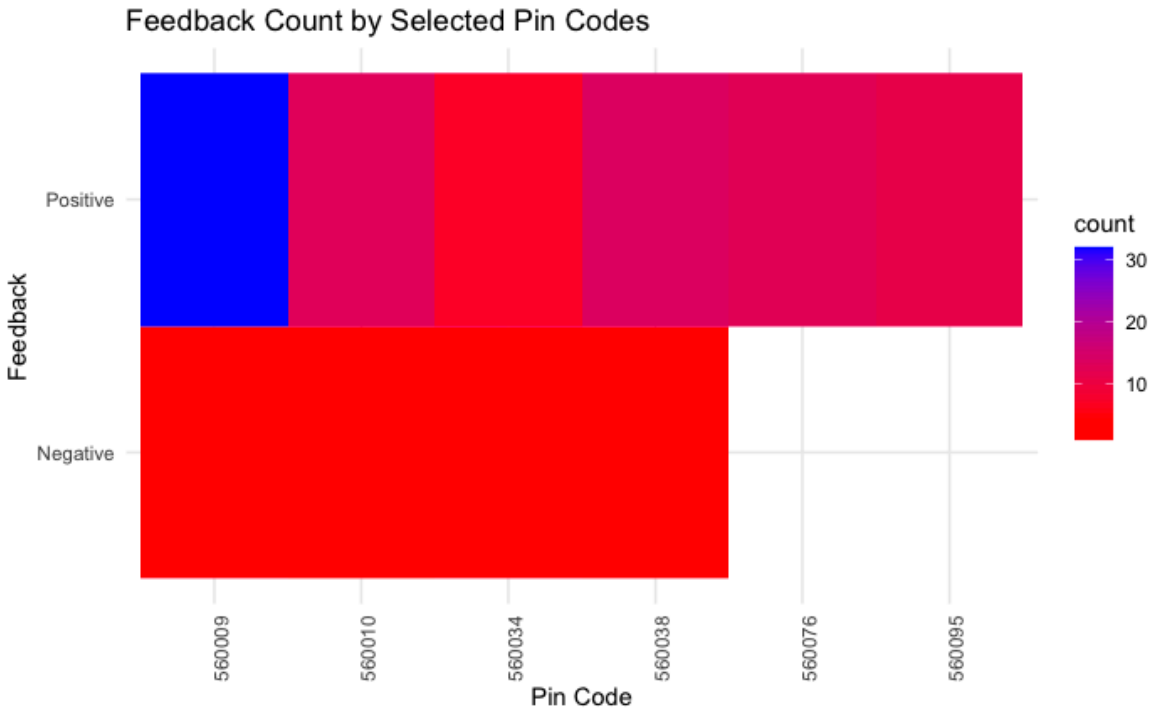


Customer Feedback on Online Food Orders:

The scatter plot shows that both positive and negative feedback is scattered throughout the geographical area, with no apparent concentration of negative feedback in any specific longitude or latitude. This distribution could imply that satisfaction with the food service is not confined to specific neighborhoods or regions within the city. Nevertheless, the scatter plot shows that the instances of negative feedback are less frequent than positive feedback.

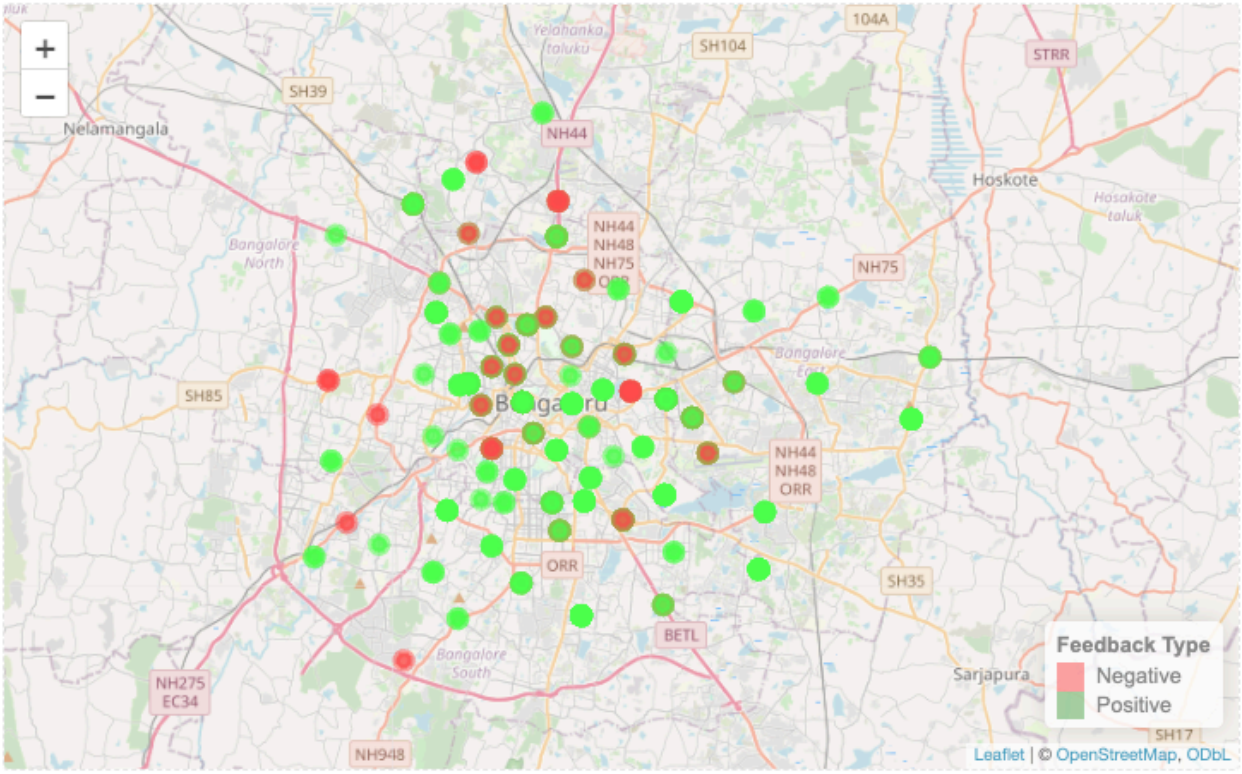
Customer Feedback Based on Area Pin Code:

HeatMap 1:



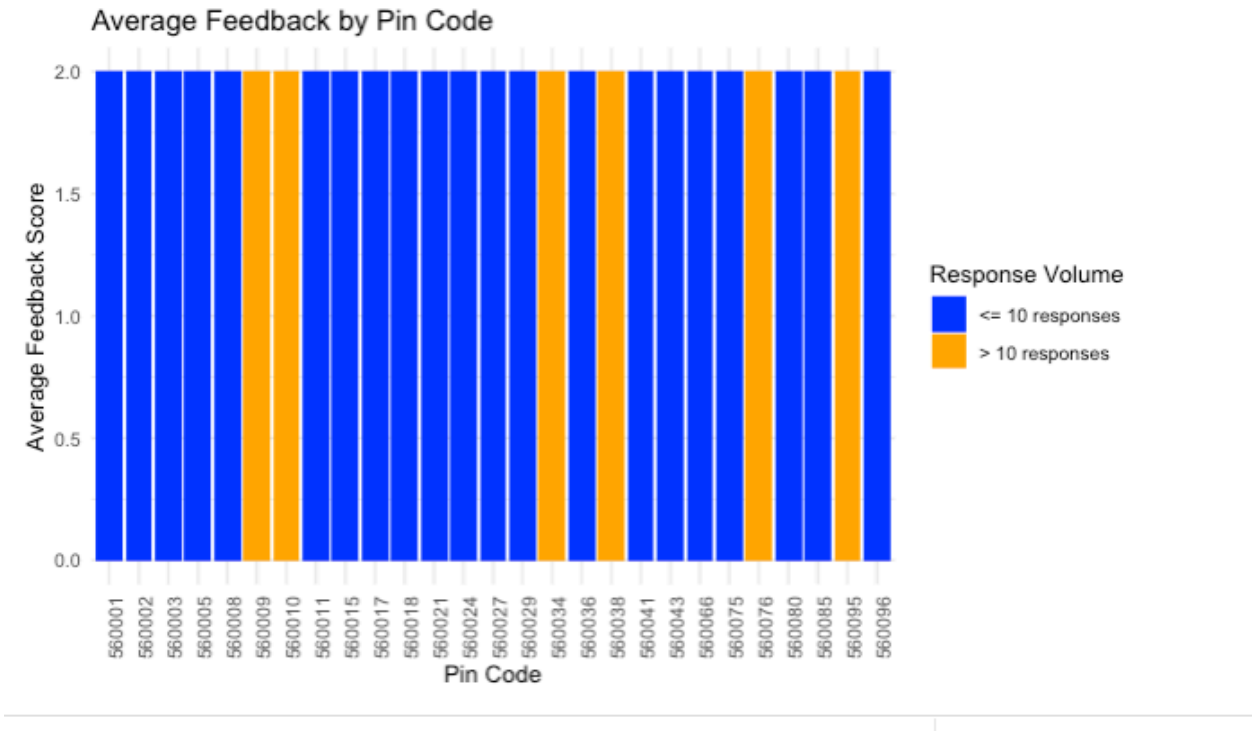
The heatmap visualization of customer feedback by pin code reveals distinct patterns in customer satisfaction across different selected pin codes to narrow down on certain random areas. Areas represented by darker blue tiles indicate higher volumes of positive feedback, suggesting greater customer satisfaction or engagement, while areas with lighter or red tiles highlight regions with less favorable responses or lower engagement levels. This pattern allows us to identify specific pin codes where customer service and product offerings may need improvement or where marketing efforts could be more effectively targeted. By examining these trends, we get to see again how there is less frequent negative feedback than positive feedback.

HeatMap 2:



This interactive map of Bangalore provides a visual representation of customer feedback on online food orders, with green and red markers denoting positive and negative experiences, respectively. A notable observation is the prevalence of green markers, indicating a generally positive reception towards online food services across the city. There's a pronounced cluster of activity in the central region, reflecting a high volume of orders that's characteristic of urban areas with dense populations and presumably a greater number of restaurants offering online delivery. However, the presence of red markers, particularly in clusters, suggests areas where service quality may need to be addressed. Peripheral regions of the map show fewer markers overall, hinting at lower service usage, which could either reflect lower demand or opportunities for market expansion. The absence of a distinct spatial pattern for negative feedback suggests that issues leading to customer dissatisfaction are not constrained to specific neighborhoods and might be more closely associated with individual service providers or isolated delivery issues. This distribution presents an opportunity for businesses to fine-tune their services and logistics, particularly where negative feedback clusters, to enhance customer satisfaction and expand their reach in areas with untapped potential.

HeatMap 3:



The bar chart represents the average feedback scores for online food orders across various pin codes in Bangalore, with a color distinction between areas with less than 10 responses (blue) and those with more than 10 responses (orange). The uniform height of the bars across the chart suggests a consistently positive average feedback score across all represented pin codes, indicative of a generally favorable customer experience with online food delivery services. The predominance of orange bars signals that a substantial number of pin codes have a significant volume of feedback, which lends greater reliability to the average scores represented. Notably, the few blue bars interspersed amongst the orange indicate certain areas with fewer responses; while their average feedback remains positive, these areas might benefit from further analysis or targeted customer engagement strategies to validate the trends with a larger data set. This visualization underscores the success of online food services in the region and points to widespread customer satisfaction, while also highlighting opportunities for businesses to engage with and understand areas with currently lower response volumes.

Conclusion

Our examination of the online food order data has shown insightful revelations about customer feedback dynamics. We have seen that students are more likely to leave positive feedback, suggesting a demographic specific satisfaction or expectation level. When further examining the differences between occupations, students are more likely to give positive feedback compared to employees even when considering different outcomes. Although housewives also have an overwhelmingly positive feedback, due to their low volume of orders, this may require more data for analysis.

In addition, income level also significantly influences feedback quality. Based on our chi-sq test, we can see that there is a statistically significant difference between income level on feedback quality. The no income group had overwhelmingly positive feedback which might follow from our findings regarding students. This underscores the potential impact of economic status on customer satisfaction. One important matter to discuss is how when involving marital status, we noticed that no income group had negative feedback when looking at prefer not to say group which shows that there may be more complex relationships involving other variables.

In contrast to demographic factors, the geographical analysis revealed a uniform distribution of feedback across different regions, suggesting that there aren't significant regional disparities in feedback quality. The lack of negative feedback clustering indicates that negative feedback is likely not due to regional differences.

These insights are useful for analyzing reasoning for negative feedbacks and potentially increasing the rate of positive feedback across all groups. However, it is important to note that this dataset only contained information about 389 orders. Therefore, we suggest potentially collecting more data in order to draw meaningful conclusions. Furthermore, dissecting the interaction between more demographic variables, such as age and gender, could show more complex relationships that affect customer feedback. These directions could contribute significantly to enhancing the strategic planning and service delivery of online food businesses.