

Sentimental Indicators: Analyzing the Impact of Online Reviews on Restaurant Success

Mustafa Elahi

May 21, 2024

Abstract

In the digital age, online reviews on platforms like Yelp play a critical role in shaping the success of businesses, particularly in the restaurant industry. This capstone project investigates the correlation between the sentiments expressed in online restaurant reviews and key indicators of economic success, such as ratings and business continuity. Utilizing the Yelp Open Dataset, this study employs sentiment analysis alongside econometric and machine learning techniques, including Ordinary Least Squares (OLS) regression and clustering methods, to decode the linguistic patterns within reviews and their association with restaurant longevity and customer ratings.

Our findings reveal that while positive reviews significantly correlate with higher ratings, their impact on a restaurant's operational status is less straightforward, suggesting a complex interplay between customer feedback and business success. The project not only enhances our understanding of the influence of consumer sentiment on the hospitality industry but also underscores the nuanced role of digital word-of-mouth in the modern economic landscape.

Data Sources and Description

This study uses the Yelp Open Dataset, which includes millions of reviews, business profiles, and user interactions from various global cities. The dataset includes business information and reviews specifically filtered for restaurants, allowing for detailed insights into consumer behavior and business performance in the restaurant industry.

Methodology

We conducted sentiment analysis on the review texts to categorize each review as either 'POSITIVE' or 'NEGATIVE'. This binary classification facilitates the quantitative assessment of review sentiments. Ordinary Least Squares (OLS) regression was then used to model the relationship between review sentiments and overall restaurant ratings.

Results

The results from the OLS regression and clustering analyses provide insights into how review sentiments correlate with business success indicators. Graphs and tables are used to illustrate these relationships and provide a visual summary of the data.

OLS Regression Analysis

The results indicate that the coefficient for positive sentiment is 0.4539, which is statistically significant with a P-value less than 0.001. This suggests that positive sentiments in reviews are associated with a higher star rating by approximately 0.454 points on a 5-point scale. The high t-value for the sentiment variable confirms robustness in this finding. The R-squared value of the model is 0.101, implying that about 10.1% of the variance in restaurant star ratings can be explained by the sentiment expressed in reviews alone.

Clustering Analysis

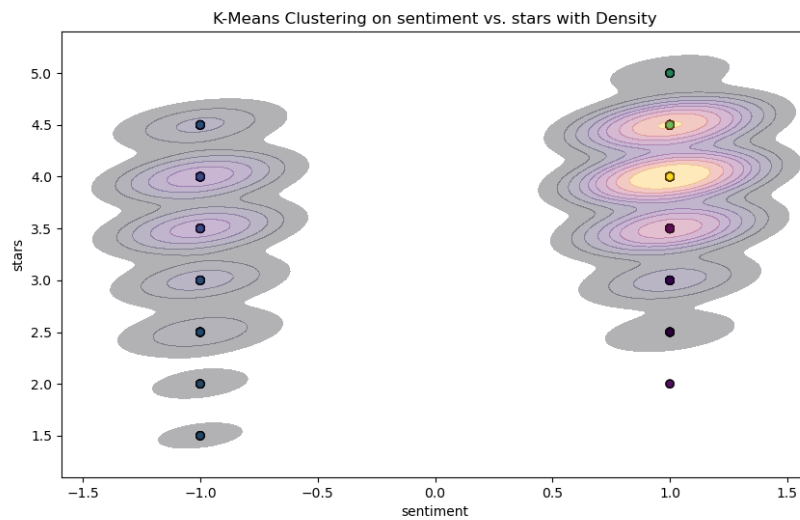


Figure 1: Clustering output showing the grouping of restaurants based on review sentiments and their operational status.

Conclusion

The study confirms that positive review sentiments are correlated with higher ratings, yet they do not solely determine a restaurant's survival, suggesting a more intricate relationship between customer feedback and restaurant success.