

# **“Recipe for Success: Data-driven Strategies for Future Restaurant Owners in San Diego, CA”**

By Tia Page

## **Project Summary**

This project focused on addressing four main objectives that are helpful in providing future San Diego restaurant owners with a good starting point for navigating the city's competitive industry.

The primary goal of identifying the key success factors for restaurants in San Diego was quickly achieved through performing correlation analysis and selecting important features using a RandomForest Classifier machine learning model. Both the analysis and the model showed that the top three indicators of successful restaurants in San Diego are the number of reviews a restaurant receives, the overall rating of the restaurant, and the restaurant's level of affordability. Earlier in this project, a list of criteria was provided for the "success\_level" column/metric; this metric is mainly based on the number of reviews a restaurant has and how high its overall rating is. Due to the lack of financial information made available to the public for each restaurant, the "review\_count" column/metric was used to relatively measure the popularity of restaurants based on how many patrons they attract. It was shown that restaurant's with high review counts tend to also have higher ratings than restaurants with a low volume of reviews.

The second goal of determining appropriate menu prices was achieved through performing EDA on the price\_levels column of the data set as it pertains to various neighborhoods in San Diego and cuisines served at restaurants. Overall, the results showed that the vast majority of restaurants in San Diego tend to stay within an affordable range of menu prices, so it is advisable for anyone opening a new restaurant to try to maintain between a 10 USD to 20 USD average meal price. The results also showed that high menu prices do not have a negative impact on the success of the restaurant. Prior to this project, I hypothesized that restaurants that are too expensive would be negatively impacted. However, insights gathered through correlation analysis using a heatmap and the interactive machine learning model showed that higher prices are linked with highly successful restaurants, as there is a moderately positive correlation between the "price\_level" values and the "success\_level" values. The interactive model also recommends restaurant owners to increase their menu prices as the restaurant's ratings and number of reviews increase; this indicates that the interactive model/recommendation system advises restaurant owners to increase the price of their menu as they become more established and reputable.

The third goal of finding recommend optimal locations in San Diego to establish a restaurant was fulfilled via filtering and sorting the dataset by ranking the top 10 restaurants with the highest overall average success levels, highest average success levels based on the type of cuisine served, and restaurant density. I took a unique approach by incorporating analysis on restaurant density because informing future restaurant owners of the areas in San Diego with the highest density of restaurants allows them to avoid areas that are highly competitive and oversaturated.

In addition, I sorted the data and compiled a list of the top 10 San Diego neighborhoods with the lowest rated restaurants; future restaurants owners can possibly take advantage of the "contrast effect," which is a business psychology term referring to someone receiving praise due to another person's shortcomings. A good example of this would be a new restaurant opening next to a restaurant with a bad reputation, then receiving stellar reviews immediately simply for providing better service than the next door restaurant with terrible reviews. Throughout this analysis, one of the main findings found were that restaurants serving Italian food showed the highest levels of success as there were three different neighborhoods in which Italian restaurants were the most popular.

The fourth and final goal of this project involving finding strategies to enhance customer experience was achieved through conducting sentiment analysis on the reviews in the data set. using the available data, code was written to produce two Word Clouds highlighting the significant words and phrases between the highest rated reviews (4 star and 5 star) and the lowest rated reviews (1 star and 2 star). The Word Clouds showed that the main reason why customers awarded highly rated restaurants with 4 or 5 stars is due to the ambiance/atmosphere of the restaurants, good customer service, and the quality of the food, while customers mainly gave low rated restaurants 1 or 2 stars due to rude/poor customer service, poor tasting food, and the loudness of the restaurant. Most of the words associated (e.g. burger, curbside, etc) with the low rated restaurants indicate that many of them pertain to fast food establishments rather than dine-in restaurants.

In conclusion, this project illuminated several underlying factors that contribute to a restaurant's success in San Diego, California. By considering all elements that go into maintaining a thriving establishment, including those which often go under looked, future restaurant owners can have a good idea of where to start when it comes to embarking on their entrepreneurial journeys. With the overwhelming rise in both costs and competition, this project helps aspiring restaurant owners to stay on track with local trends in the San Diego area so that they can make wise business decisions and keep their future restaurants operating for many years to come.