Term Project Proposal

David D Berberena, Brian Mann

Bellevue University

DSC 630 Predictive Analytics

Andrew Hua

June 16, 2024

Term Project Proposal

The restaurant industry is one of fast-paced, ever-evolving trends that either help investors ride the wave of success or have them drown in the overwhelming depths of missed opportunity. Roughly sixty percent of new restaurants close their doors within the first three years of operation, so while many factors are at play when opening and running a restaurant, it is imperative that owners, operators, and stakeholders are aware of customer trends and wants and have the ability to deliver to the guests to succeed (Letchinger, 2013). Looking higher at those companies who have built multiple large-scale restaurant brands, their shareholders need assurance that the company knows what their customers want to continue long-term increases in revenue. Yet here is where the business problem lies: how can large restaurant conglomerates understand and cater to their customers’ needs when corporate suits never fraternize with the day-to-day guests? Take Bloomin’ Brands for example, or maybe Yum! Brands, or even Darden Restaurants. With so many restaurant brands under their belts with different cuisines and concepts, how can these conglomerates predict which sector will be the most lucrative? Data holds the answer to this problem, as these restaurant conglomerates constantly intake data and process it to make certain that revenue is maximized. Shareholders of these conglomerates would be interested to see how data can predict the revenue of varying sectors of the company, as being aware of what the customers want from the restaurant sectors ultimately leads to informed decision-making resulting in a larger net profit for the company, thereby making the shareholders richer.

# The Data

To directly participate in the predictive analysis aspect of restaurant conglomerates making decisions meant to drive brand revenue, we have accessed a dataset found on Kaggle of simulated data outlining specific factors that affect restaurant revenue on a monthly basis (MrSimple07, 2024). This dataset provides variables such as customer count, average menu price, marketing expenditures, cuisine type, average customer spending, the presence of a restaurant promotion, review count, and monthly revenue. As restaurant revenue is the target variable in this proposal, the other variables will act as predictors to maximize future restaurant revenue growth. There are aspects of the dataset that we will need to specify for the observations to make more sense for the business problem at hand, as we are looking at monthly revenue yet most variables within our dataset do not contain units of measure. For example, the number of customers variable does not specify in what time period this observation is valid. As this data is simulated, we shall place our own units of measure as defined below:

Number of Customers: This variable will be the average hourly customer count for the month

Menu Price: Here will lie the average menu price of all entrée items the restaurant sells

Marketing Marketing Spending: This variable will be represented in thousands of dollars

Average Customer Spending: Here will define the average total check amount per customer

Promotion: We will establish this variable as the presence of a Happy Hour promotion or similar in-store promotion, such as a monthly coupon redeemable upon dine-in.

While these new assumptions have been placed on the dataset, we will be explicit in saying that the results may not be accurate due to the nature of the data being simulated and the missing meanings behind the original dataset. Seeing as we have so many factors affecting restaurant revenue, we can now delve into our model choice: linear regression.

## Model Selection and Performance Metrics

Our choice of linear regression to verify the factors that positively affect restaurant revenue will allow us to see which factors should be presented to shareholders in restaurant conglomerates like Darden Restaurants as growth opportunities. Providing a promising model will create a secure space for informed decision-making, which could lead to increased investments in lucrative business sectors and potential rebranding efforts for struggling restaurants. We will be employing Jupyter Notebook to craft a multiple linear regression model in Python to output results that can tell whether a certain set of factors is more influential in a positive monthly restaurant revenue than the rest. This model assumes that the predictor variables are linearly related to the outcome variable, and while looking visually at the data may suggest one thing, the model will certainly confirm or rebuke the initial assumption. To ensure that we are crafting the model with the highest degree of validation, we will use k-fold cross-validation techniques and regularization models like Lasso and Ridge regression models as part of the hyperparameter tuning process. The goal of the tuning process is to lower the root mean squared error (RMSE) value to the point where the model can confidently output a series of variables that have a significantly positive effect on monthly restaurant revenue for shareholders to see and act on. We are also looking for a coefficient of determination value as close to one as possible, as this means that the model is a good fit for the data and explains most if not all of the variability of the data. To prepare the data for the model, we must simply transform any categorical variables into numerical variables via dummy variables or one-hot encoding. The dataset will then be split into training and test sets and then be placed into the aforementioned models to realize the performance metrics needed to see which factors best predict an increase in monthly restaurant revenue.

**Learning Goals**

Your turn

**Risks and Ethical Implications**

Your turn

### Conclusion

### Your turn

### References

Letchinger, C. (2013, September). The Anatomy of Restaurant Failure: Dead Man Walking. Menu Cover Depot. Retrieved June 12, 2024, from <https://www.menucoverdepot.com/resource-center/articles/restaurant-failure/>

MrSimple07. (2024). Restaurants Revenue Prediction [Data set]. Kaggle. Retrieved June 12, 2024, from <https://doi.org/10.34740/KAGGLE/DSV/7420974>