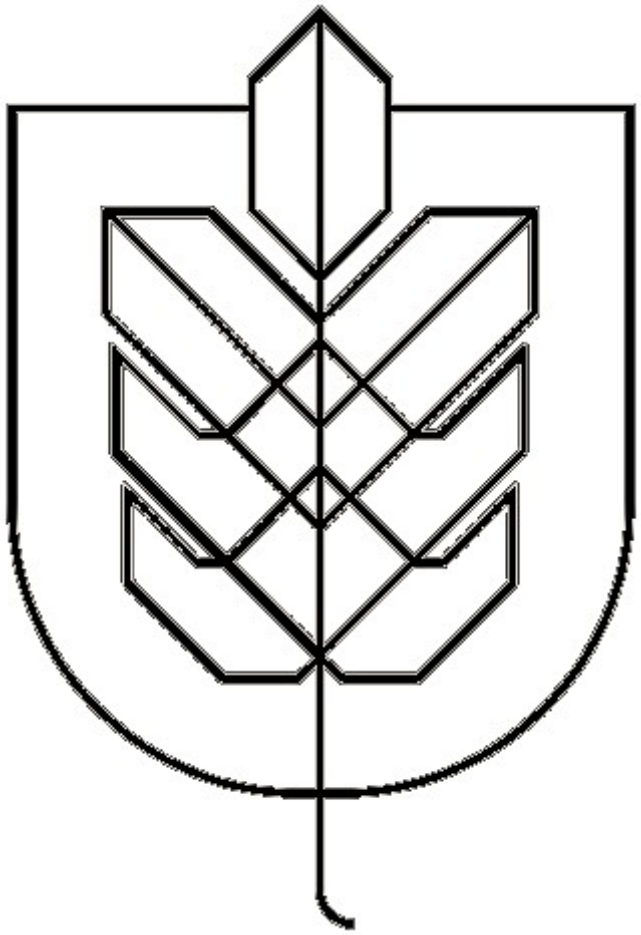
Course Case: Apprentice Chef - Assignment 1



- DAT5303 – Machine Learning
- Professor Chase Kusterer
- Hult International Business School
- Maxime Levintoff

Overview

The Chef Apprentice dataset provided consists of 1946 entries and 29 columns.

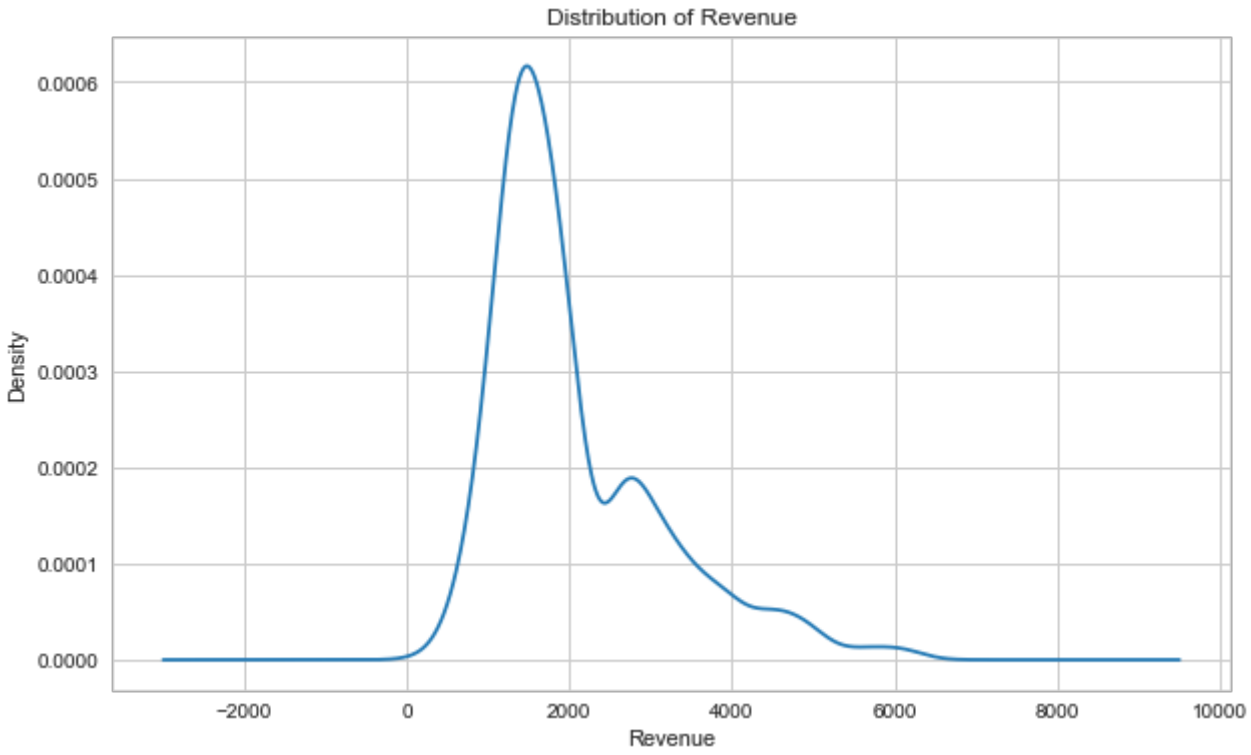
Dataset info:

- Number of variables:29
- Number of observations:1946
- Missing cells:47 (0.1%)

Variable Types:

- Numeric:16
- Categorical:6
- Boolean:5
- Text (Unique):2

- `CANCELLATIONS_BEFORE_NOON` has 667 (34.3%) zeros
- `EARLY_DELIVERIES` has 1167 (60.0%) zeros
- `FAMILY_NAME` has a high cardinality: 1071 distinct value
- `FAMILY_NAME` has 47 (2.4%) missing values
- `FIRST_NAME` has a high cardinality: 1442 distinct values
- `FOLLOWED_RECOMMENDATIONS_PCT` has 155 (8.0%) zeros
- `LATE_DELIVERIES` has 319 (16.4%) zeros
- `TOTAL_PHOTOS_VIEWED` has 1140 (58.6%) zero
- `WEEKLY_PLAN` has 468 (24.0%)



INSIGHTS

- The key performer and the strong contributor to the model were throughout showing strong effect in terms of making revenue:
    - MEDIAN_MEAL_RATING
    - MASTER_CLASS_ATTENDED
    - TOTAL_MEALS_ORDERED
    - CUSTOMER_W_CUSTOMER_SERVICE
    - TASTES_AND_PREFERENCES
    - MOBILE_LOGINS
    - PC_LOGINS

This shows how the revenue is directly being linked through the mentioned parameters. Also, in the analysis, it is quite evident from the graphs, figures and data interpretation that these parameters contribution is the key.

If the median meal rating is high, then the revenue will go high for a meal as more people will tend to order the same.

Similarly, Total meals ordered is anyway strongly relevant to revenue.

Customer interaction with the services in order to make it better in terms of feedback and continuous correction measures. This shows us the importance of having a strong customer service to satisfy customers and increase their likelihood to come back.
Logins generally show the intent of the user as from where they usually tend to order food.

RECOMMEDNATION

- Through the analysis, it is obvious that the median meal rating, the number of master class attended as well as the total number of meals ordered drastically influence the revenue. These parameters usually indicates how loyal and satisfy is a customer. The number of master class attended as well as the total number of meals ordered both represents how loyal is the customer while the medium meal rating shows how happy is the customer with the product. Thus, Apprentice Chef should definitely use these 3 metrics to assess the potential of each customer. They should focus on constantly satisfy these customers as they can serve as ambassadors for the brand. In conclusion, loyalty program should be created to give more benefits to these customers. Apprentice Chef should also get feedback from these customers and ask them what products they would want to have to keep satisfying them. In conclusion, Apprentice Chef has to define is top customers using the 3 metrics and focus on keeping them happy.

FINAL MODEL PERFORMANCE

The best performing model (KNN) gives an R squared value of 74.568 with standardized data.
It runs in 6.08 seconds.

```
The optimal number of neighbors is 14
Training Score: 0.78639
Testing Score: 0.74955
CPU times: user 6.02 s, sys: 39 ms, total: 6.06 s
Wall time: 6.08 s
```

The best performing model (KNN) with unstandardized data and using the optimal number of neighbors (3) gives an R-squared of 99.950.