OnlineRetail Data Set Analysis

By Martell Tardy

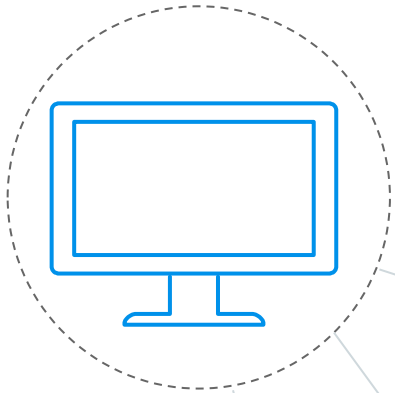
Pennsylvania State University
IE 575: Predictive Analytics

December 13, 2016

Introduction

OnlineRetail Data Set

Through an in depth analysis of the OnlineRetail data set the UK based non-store online retail agency is hoping to gain insight into their top performing products, trends in transaction history, and the forecasting analytics to improve sales moving forward. The data received from the UK based non-store online retail agency contained transaction and demographic history for each purchase during the specific period of interest.

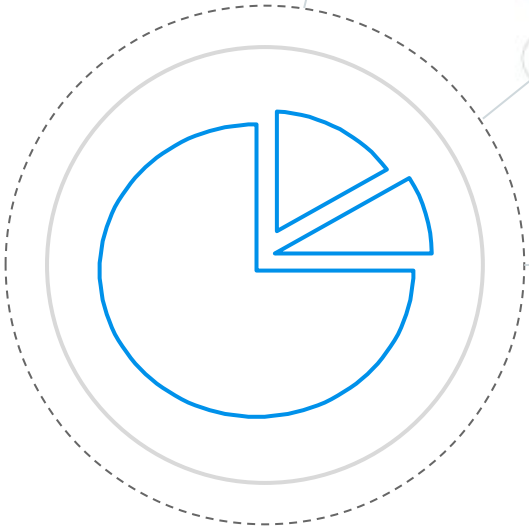


*How **recently** are transactions?*

*How **often** do transactions happen?*

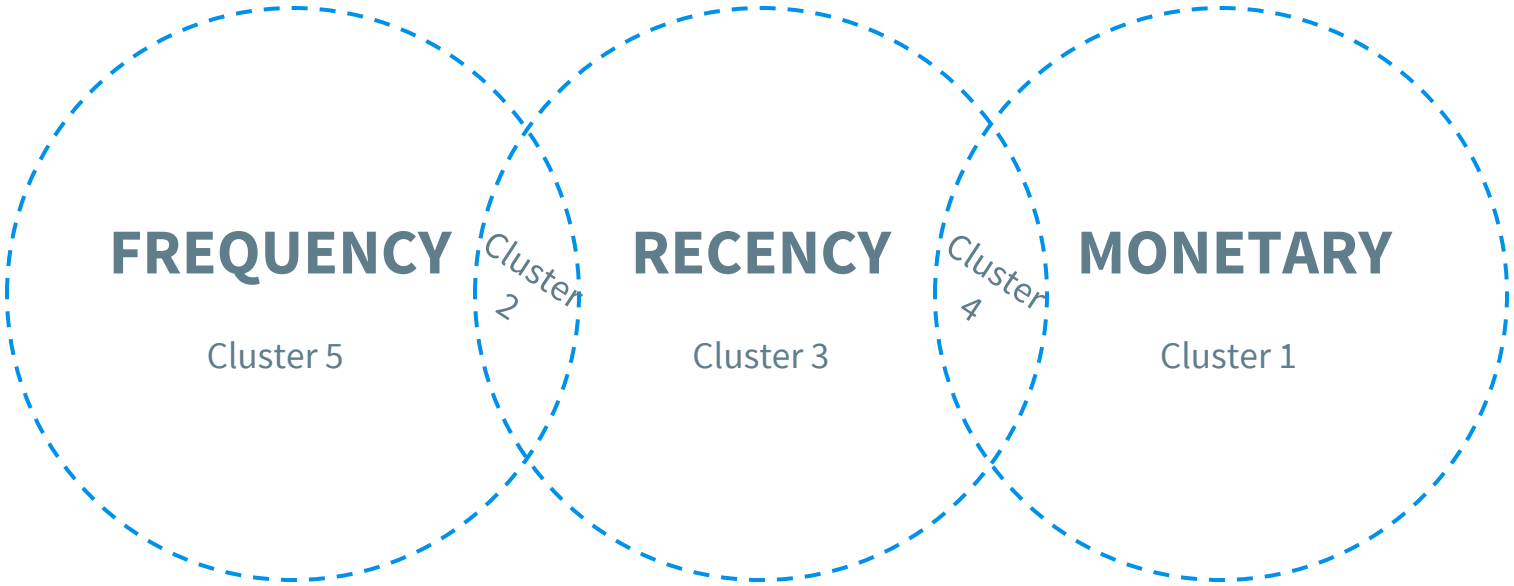
*How much is **spent** (each time on average) during a transaction?*

RFM Analysis



A method used to generate Recency, Frequency, and Monetary understanding of transaction clusters

Five Clusters: RFM Analysis



Five clusters were identified using the RFM Analysis Process.
The measurements of Frequency value >=3000, Recency value <=200, and a Monetary value>=2000 were weighed to determine which circle(s) a cluster most closely represented.

Predictive Methods

Methods used to forecast sales and identify trends emerging during a specific time span

Comparison of Predictive Methods

Moving Average

Used to estimate future values at time t by averaging values of the time series within k periods of t

Time Span

Used to get a relative visual of how each variable has performed over time in comparison to one another

Exponential Smoothing

Used when the data has no trend or seasonal patterns. Unlike a moving average, this technique gives greater weight to the most recent observations of the time series.

ARIMA

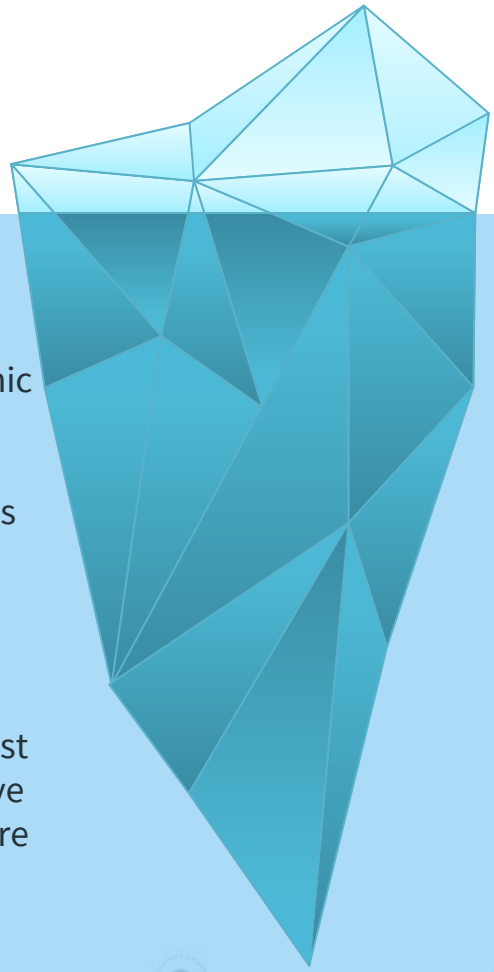
Used as a filter that tries to separate the signal from the noise, and the signal is then extrapolated into the future to obtain forecasts.

After reviewing the results of the four models the MA method provided the most robust understanding and visual of the data. In the MA model there are clear trends relating to each of the three variables and these observation are echoed in the other methods in a less complete fashion.

Next Steps

Recommendations, observations, and best practices that can be applied to improve revenue moving forward

Observations From This Report versus Potential



Clusters and RFM

The agency should use the five transaction clusters in further analysis, cluster demographic breakdown, and tracking of experimental strategies can be executed by the UK based non-store online retail agency to increase sales within each target group and potentially the development of the ideal transaction cluster.

10 Best Selling Items

This agency should create a real-time rolling list of their best sellers and the relationship the five transaction clusters maintain with this list. Here is where this agency has a true opportunity to increase sales and see their experiential strategies to move the needle in real-time.

Moving Average Model

The agency should review more closely the findings from the Moving Average model. Observations pertaining to the time of year certain Items were seen in transaction and the Quantity of each transaction could be seen. This agency should review their pricing, product availability, and website optimization in the summer months since in each forecasting model there was a decline in the trend during this season.

Conclusion

OnlineRetail Data Set

The UK based non-store online retail agency was looking to improve future sales through the analysis of their retail data from 2010/12/01 to 2011/12/09. Through this report the online retail agency gained insight into their top performing products, trends in transaction history, and recommendation on ways to improve sales moving forward. With the results and recommendations from this report, the UK based non-store online retail agency is well on their way to increasing further sales and customer retention.