When Will Customer Buy: An Evaluation of sales and what takes out money from customers pocket?

Rahul Gupta - DSC 680 - Winter 2020

https://github.com/rahulgupta271/DSC680_Project_1_Market_analytics_holiday_trend

Which Domain?

This project will concentrate on the patterns in retail sales and the correlations that influence shifts. I don't have any background in the retail industry, other than in retail analytics, but I'm fascinated in how traders decide when, apart from the usual Black Friday/Cyber Monday sales, they should put what products on sale.

Reference for this project are mentioned below:

Nicasio, Francesca. August 20, 2019. “Retail Analytics: How to Use Data to Win More Sales and Customers.” https://www.vendhq.com/blog/how-retailers-can-use-data-to-boost-productivity-customer-service-sales/

Bobriakov, Igor. July 22, 2018. “Top 10 Data Science Use Cases in Retail.”

https://medium.com/activewizards-machine-learning-company/top-10-data-science-use-cases-in-retail- 6483accc6042

Alloy Client Solutions. March 21, 2019. “Retail analytics trends in 2019 and beyond.”

https://medium.com/alloytech/retail-analytics-trends-in-2019-and-beyond-c671783c67b7

Jones, Joshua. March 2, 2018. “Fast Forward: How retailers will use data and analytics to succeed in 2018 and beyond.” https://www.strategywise.com/fast-forward-how-retailers-will-use-data-and-analytics- to-succeed-in-2018-and-beyond/

Hufford, Jillian. June 3, 2019. “How to Perform Sales Trend Analysis For Your Retail Business.”

https://www.nchannel.com/blog/how-to-perform-sales-trend-analysis/

Prevedere.com. N.D. “10 Mistakes to Avoid to Better Leverage Retail Analytics and Forecasting.”

https://www.prevedere.com/leveraging-retail-analytics-and-forecasting/
Sisene.com. N.D. “What is Retail Analytics?”
https://www.sisense.com/glossary/retail-analytics/

Chatterjee, Sharmistha. May 27, 2019. “Traditional vs Deep Learning Algorithms in Retail Industry.”

https://towardsdatascience.com/traditional-vs-deep-learning-algorithms-in-retail-industry-i-b7b7f86793d4

Maitra, Sarit. October 18, 2019. “Clustering & Machine Learning Combination in Sales Prediction.”

https://towardsdatascience.com/clustering-machine-learning-combination-in-sales-prediction- 330a7a205102

Bosier, Fabian. November 12, 2019. “Pandas’ groupby explained in detail.”

https://towardsdatascience.com/pandas-groupby-aggregate-transform-filter-c95ba3444bbb

Which Data?

The dataset I'm going to use was posted to Kaggle.com through Manjeet Singh two years ago and is located at:

https://www.kaggle.com/manjeetsingh/retaildataset#Features%20data%20set.csv

This dataset covers weekly sales at 45 separate retailers from February 2010 to October 2012. This statistic provides measures as to whether or not a certain department has substantial marks or profits leading up to and through the indicated week.

Research Questions? Benefits? Why analyzes these data?

My key analysis concern is What drives sales?" "Drilling deep, I'm going to try to find clear explanations for high and low sales between different retailers and between different divisions within those stores..

The benefit of this analysis is to obtain insight as to whether a store's consumer base is expected to make a purchase and what if any, the store can do to affect such sales. Obviously having the power to control revenue has one big advantage: improved profits.

I can visually review the results, see some apparent patterns and see where each department has flourished. Then I'm going to try to classify the associations that influence revenue. Most definitely, I'm going to try heatmaps and uncertainty matrixes and see how the features apply to each other.

My motivation in this area is to chat to a friend who works for an analytics company focused specifically on retail analytics

What Method?

At first, I'm going to use R and/or Python for visualization. I'm more likely to attempt to build visualizations of Power BI, when many

businesses are using Power BI and searching for analysts who are fluent in it. I'm more inclined to use R to discover correlations, as I've always found it easier to use than Python.

Potential Issues?

For me, one of the greatest problems I see would be that I could not identify any real meaningful associations, as my intuition would be that there are actually a variety of correlating characteristics. This is going to go off track if I concentrate more on one thing than the others and lose sight of time control.

Concluding Remarks

I'm interested in getting a career in analytics, and as retail analytics appears to be a growing focus in the world of data science, I decided to discuss this field in order to gain more insight in learning what stores should do to influence buyers to buy from them.