Executive Summary

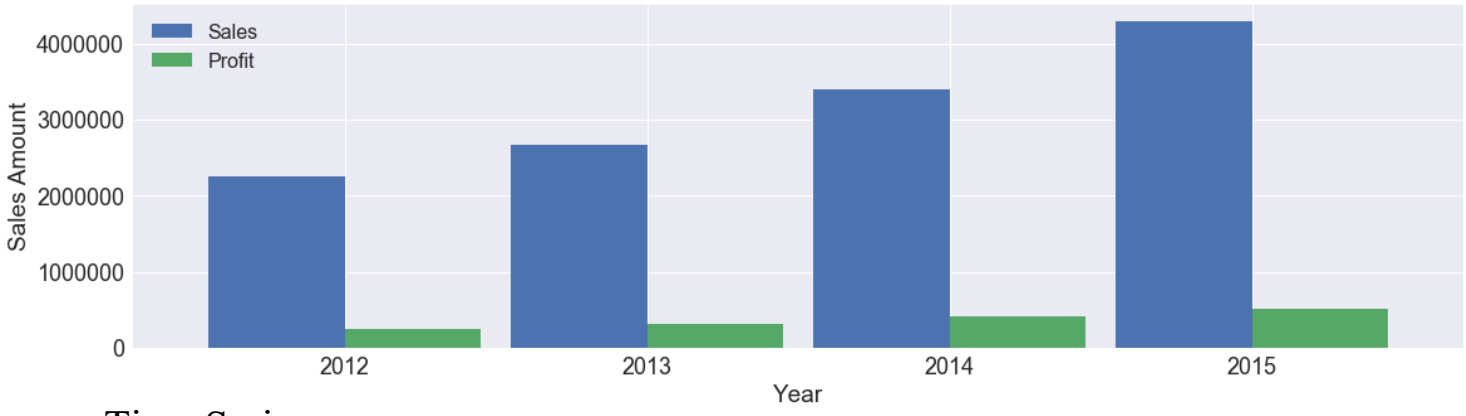
The objective of this project is to help Superstore to maximize its sale. Superstore is a global retail company which has average annual sales around 3.2 million and the products it sells can be divided into three categories, technology, furniture, and office supplies. The dataset I obtain contains Superstore's sales transactions made in 2012 to 2015 and it's available on Tableau Community Forums.

Datasets

- Global Superstore sales transactions from Tableau[1]
- Gender for every customer from genderize.io[2]
- Income level for every country from The World Bank[3]

EDA

In order to maximize the sales, I first look at the historical sales and the sales is growing year by year and so as profit, but the gross profit margin might be too low due to the big difference between sales amount and profit amount.



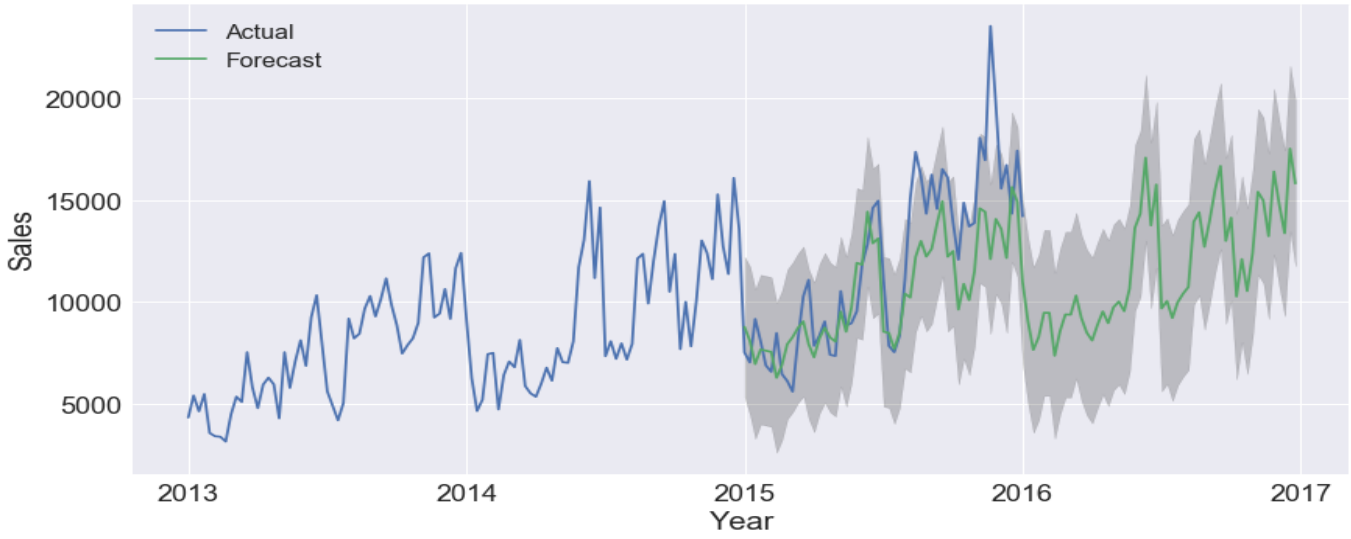
Time Series

Therefore, I built a forecasting model using SARIMA model to try to find out the trend of the sales of the year 2016.

- Best Model: SARIMA (order = (1,1,1), seasonal_order = (1, 1, 0, 52))

[1] https://community.tableau.com/thread/194200
[2] https://genderize.io/
[3] https://data.worldbank.org/

I first used resampling to transform the dataset into weekly data which is also the reason I set the parameter *s* in seasonal_order to 52. In my next step, I checked deviation of different difference levels for weekly data to find a lowest difference level, which is 1 for this dataset, so I can further confirm that I need to set the parameter *d* to 1. Although I can get the idea of potential numbers that I need to use for the other parameters from the autocorrelation plot and the partial autocorrelation plot, I chose to use AIC score to help me find the best model.



According to the forecasting above, the sales are not increasing significantly for the next one year, so my suggestion for Superstore is to focus on maintaining existing customers using the recommender I built for them.

Recommender

My recommender will pull out the transactions the target customer made before and use cosine similarity to find top 10 similar customers. Next it will pull out what those customers have purchased before and then further compare the items with the target customer. The recommender will recommender the item that the target customer hasn't purchased before.

Future Extension

The efficiency of the recommender will have to be test using A/B testing and I will further build a Flask App to allow the web development team and the sales team to access the recommender.