



Sales Forecasting with Facebook Prophet

Miguel Angel Santana II, MBA

February 2021

A large orange shape on the left side of the slide, consisting of a rectangle with a quarter-circle cutout on its right side.

Introduction

- Methodology
- OSEMN Framework
 - Obtain | Scrub
 - Exploratory Data Analysis
 - Statistical Modeling
 - Validation
 - Interpret Results
- Conclusions & Limitations
- Future Work
- Thank You



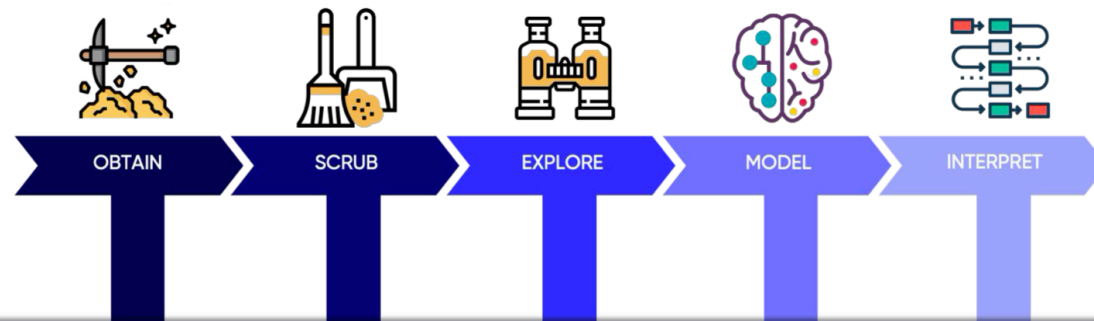


Methodology

- An unknown corporation has provided 1,017,209 sales records inclusive of 1,115 stores over a two-year period.
- Our team will leverage Facebook Prophet to forecast sales in performing stores in order to gain market insight.
- The OSEMN framework is used to process, analyze, model and interpret the results of our analysis.

OSEMN Framework

Data Science Process



O

Gather data from relevant sources

S

Clean data to formats that machine understands

E

Find significant patterns and trends using statistical methods

M

Construct models to predict and forecast

N

Put the results into good use

Originally by Hillary Mason and Chris Wiggins

Obtain | Scrub



Scrubbing / Data Cleaning

The dataset was provided in Udemy course Data Science for Business by Ryan Ahmed.

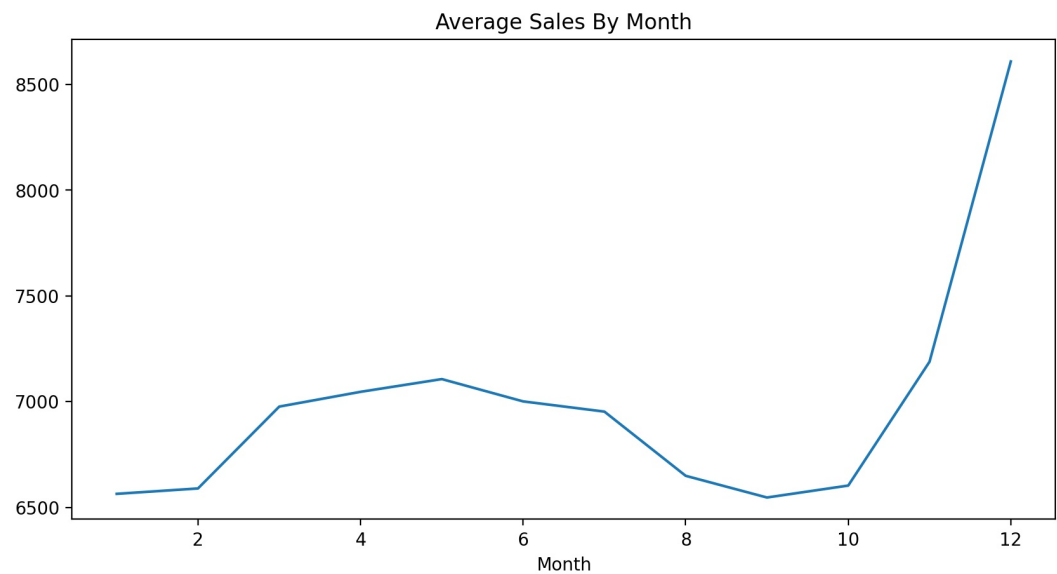


Key Decisions

Null Values were Filled
Dropping Closed Stores
Month, Day and Year Feature Engineering
Joining Data Frames
Extracting School & State Holidays

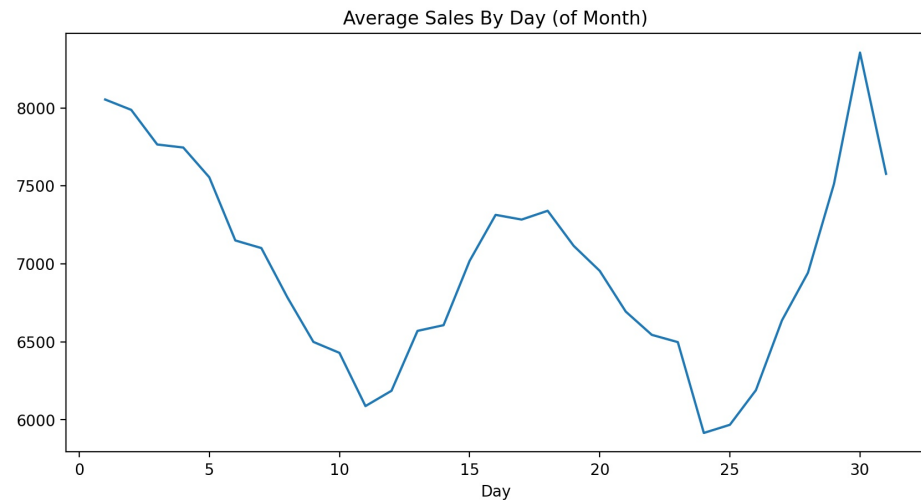
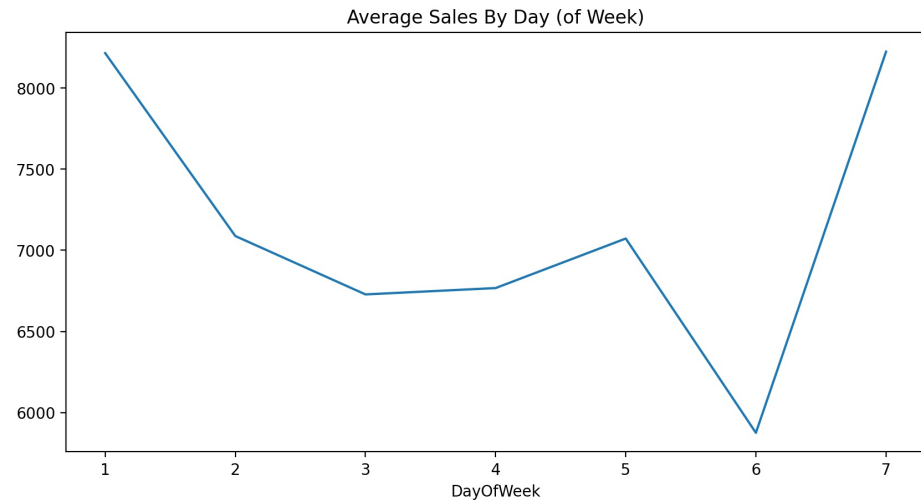
Exploratory Data Analysis

- Average Sales by Month
 - Summer Trend
 - Winter Holiday Trend



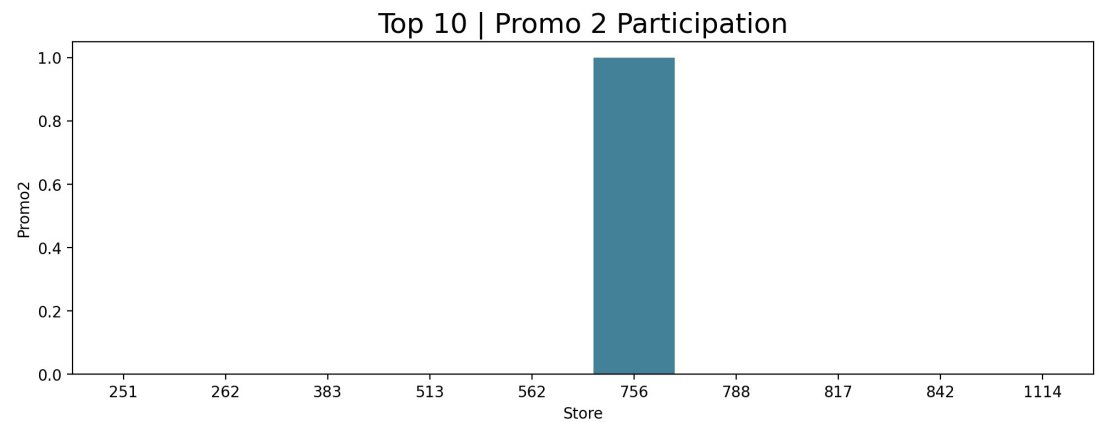
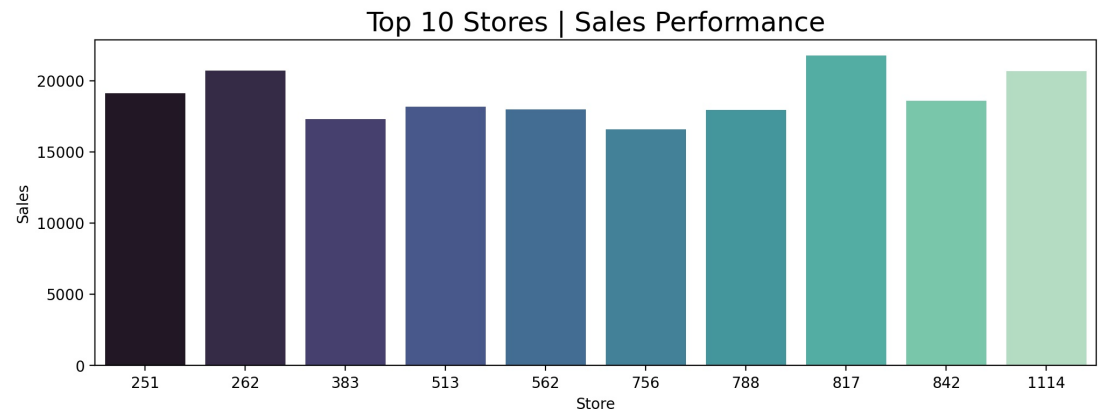
Exploratory Data Analysis Continued

- Average Daily Sales
 - Day of Month
 - Three upticks in sales per month
 - Largest occurs toward the end of the month
 - Day of Week
 - Most sales occur on Sunday and Monday.



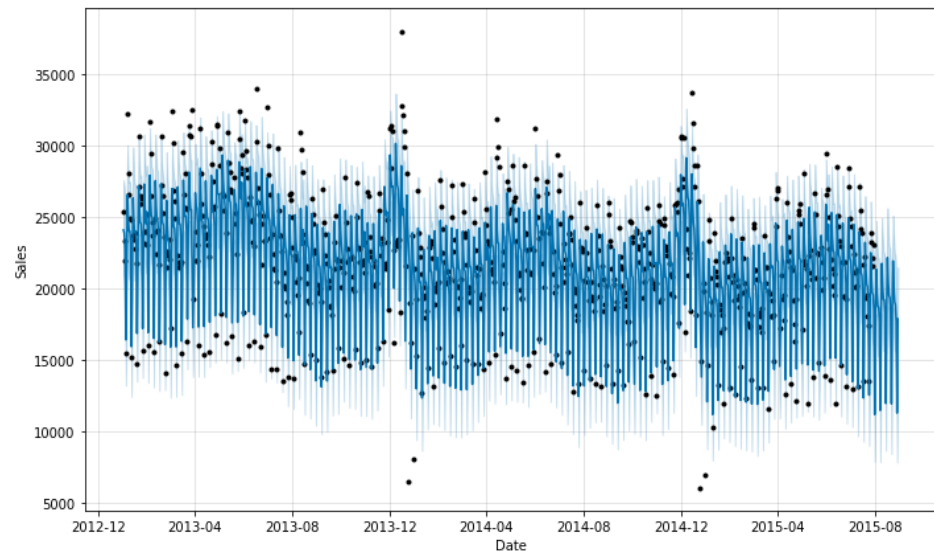
Exploratory Data Analysis Continued

- Top 10 Performing Stores
 - Store 817 has the highest sales average.
 - Only 1 of the 10 stores participated in promotion two.



Statistical Modeling

- Facebook Prophet | Store 817
 - Projected decrease in sales over the next 30 days.
- Validation Metric
 - Root Mean Squared Error
 - Average RMSE: 2025
 - Mean Absolute Percentage Error
 - Average MAPE: 7.8%




Interpret Results | Business Insight

- Performance
 - Our average error metric was less than 10% of the average sales record for store 817.
- Business Insight
 - The average distance of a competitor to a store is 5457.84.
 - The average competitor distance to a top 10 store is 727.
 - 50% of stores participated in promotion two.
 - 1 of 10 top stores participated in promotion two.



Limitations & Future Work

- Limitations

- Data anonymity
 - Product Features
 - Geographic Location
 - Customer Base
 - 6-year-old data
 - Unable to identify trend changes
- 

- Future Work

- Future work should include additional store level data such as employees in store, online sales, product information and geographic information to consider additional external factors through feature engineering.



THANK YOU!

Questions? Miguel Santana | Contact: santana2.miguel@gmail.com

Additional projects can be found on Github.

Username: [miguelangelsantana](#)