CORPORACIÓN FAVORITA

Time Series Forecasting

Agenda

Introduction
1

EDA
2

Modelling
3

Conclusion
4

Background

1952

Quito
First Store - LA FAVORITA

Supermarket set into shopping malls

SUPERMERCADOS LA FAVORITA C.A.

1971-79

1983

SUPERMAXI
el placer de comprar

Rename Bodega La Favorita to Supermaxi

March 28
Line of business expanded so well

CORPORACIÓN FAVORITA

2008

Figures of interest

54 Store

Throughout Ecuador

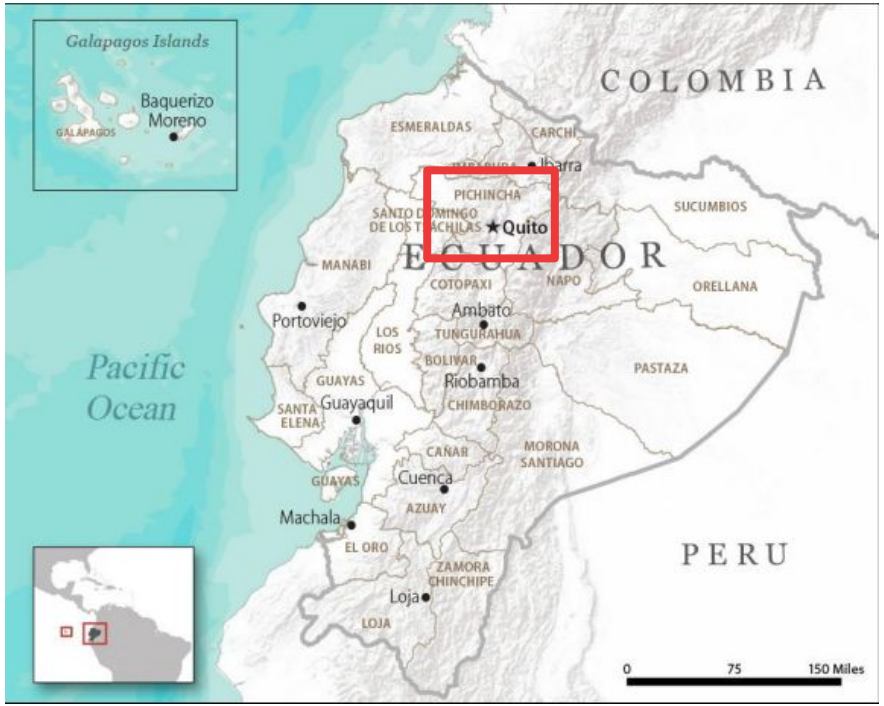
16 State

In Pichincha

Alone has 19 outlets

70 Years

We are committed to Ecuador and its development.



10,000

Employees

Corporación Favorita

288M

Sales

In 2015

3M

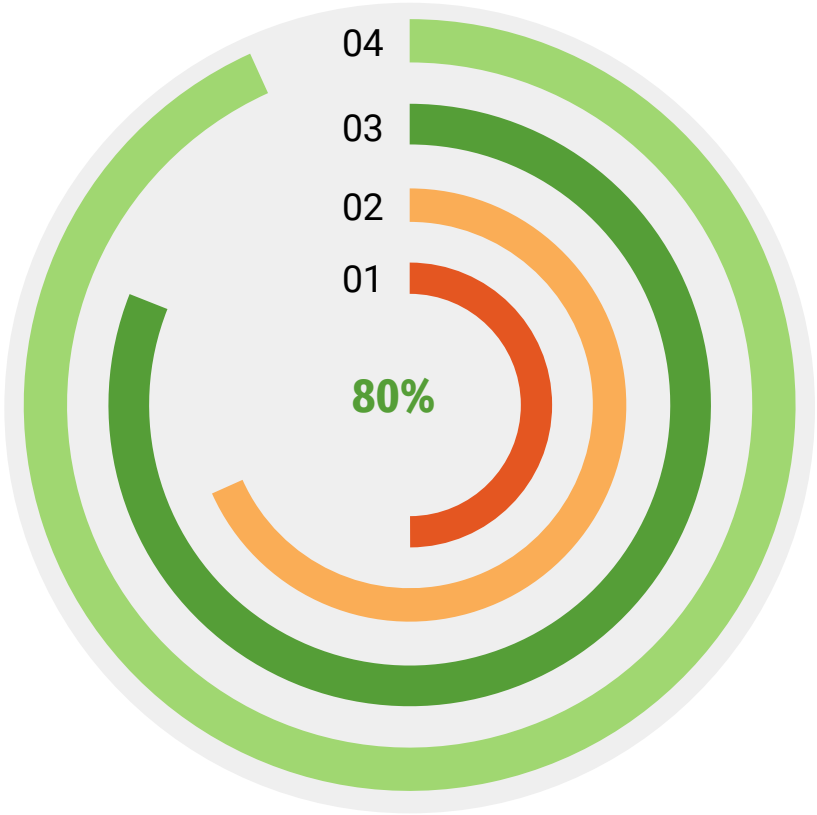
Customer

The client is our passion

Problem Statement

Forecasts

- Especially relevant to brick-and-mortar grocery stores
- Predict a little over, and grocers are stuck with overstocked, perishable goods.
- Guess a little under, lost revenue and upset customers.
- Apply machine learning to help Favorita buyer to ensure enough supplies to sell and not overstock.



EDA
2

Data Sets

Train & Test Set

Sales data since 2013 - 2017
Product Family
Store Code No.

3M Row

Holiday Set

Holiday Types and Nature
Since 2013 - 2017

350

Store & Transactions

Transaction data per store since 2013 - 2017
Store State Classifier, Claster

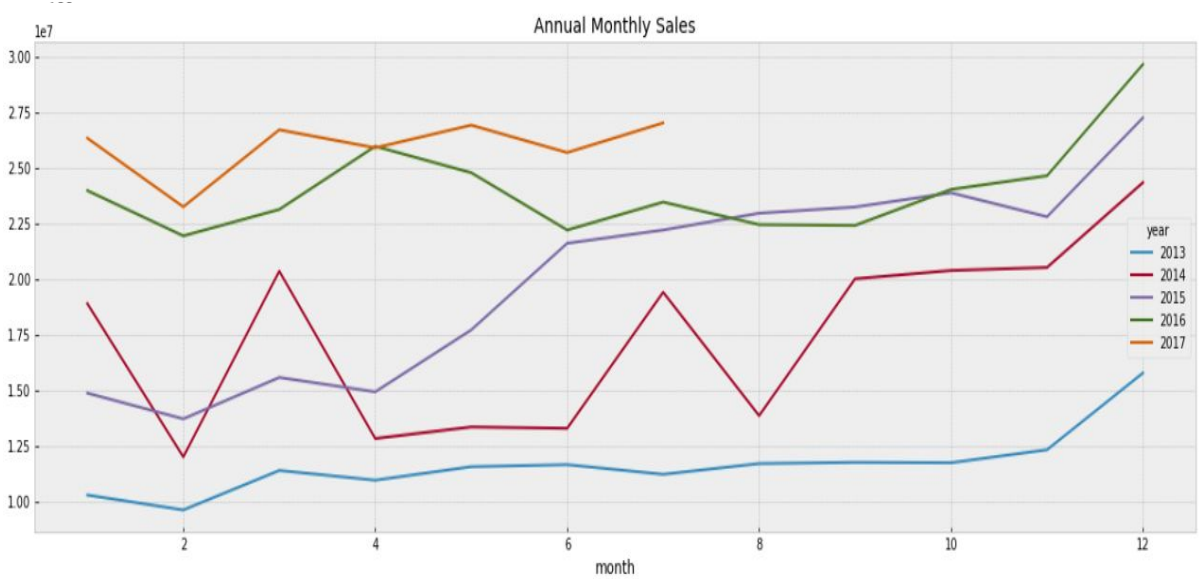
83k

Oil Price

Daily oil price since 2013 - 2017

1175

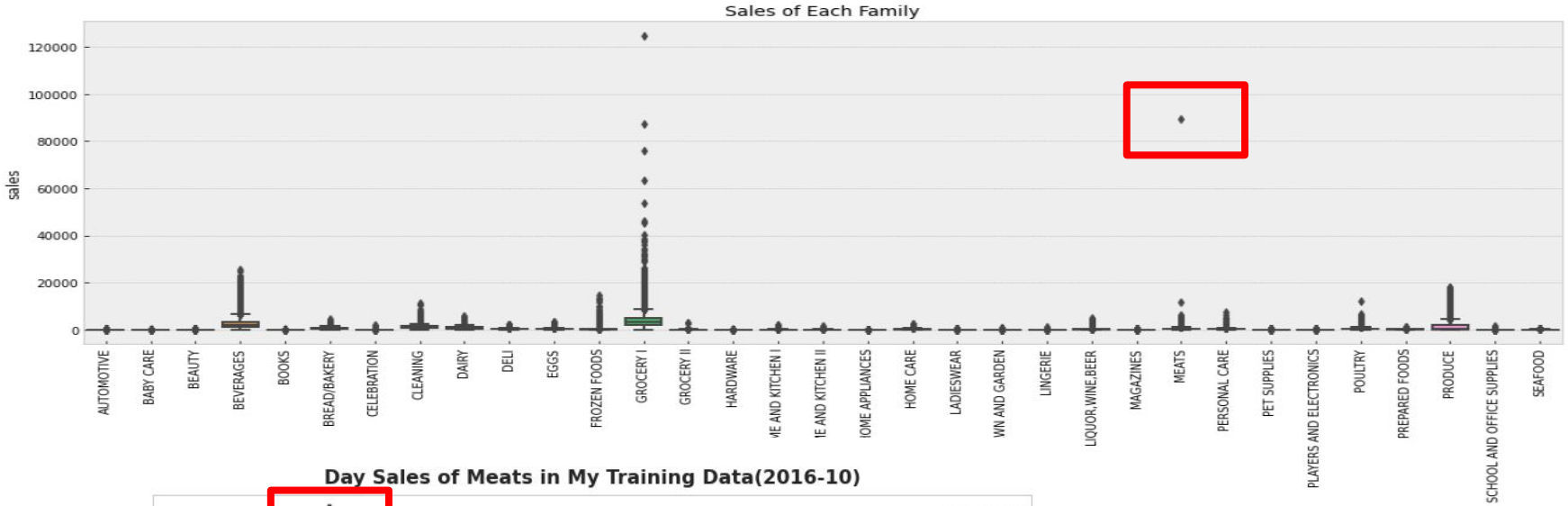
Sales by Annual



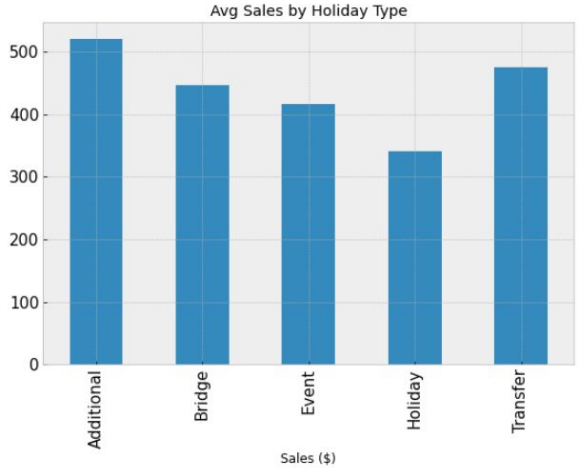
Annual Sales (M)	
2014	209
2015	240
2016	288

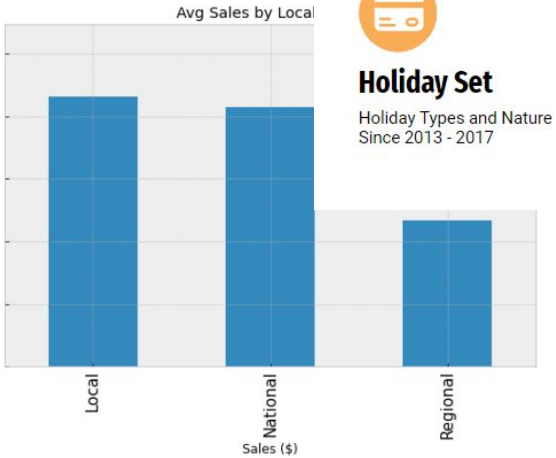
Total sales increase	+ 16%

Sales by Product Family



Sales During Holiday



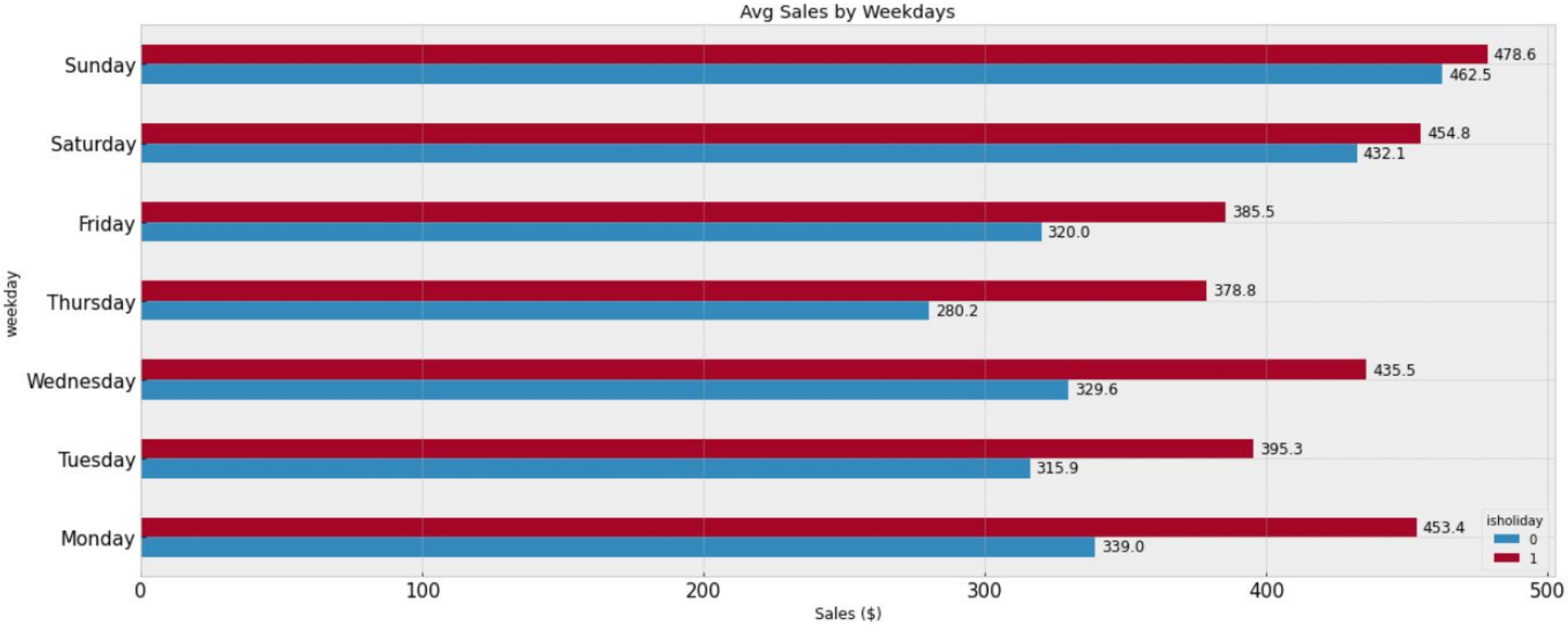


Holiday Set

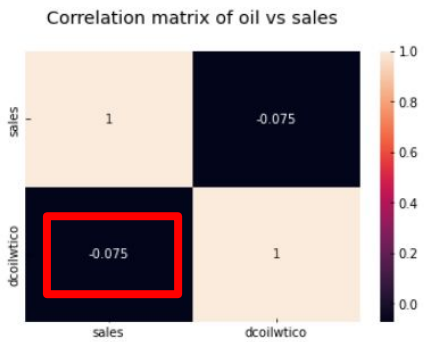
Holiday Types and Nature
Since 2013 - 2017

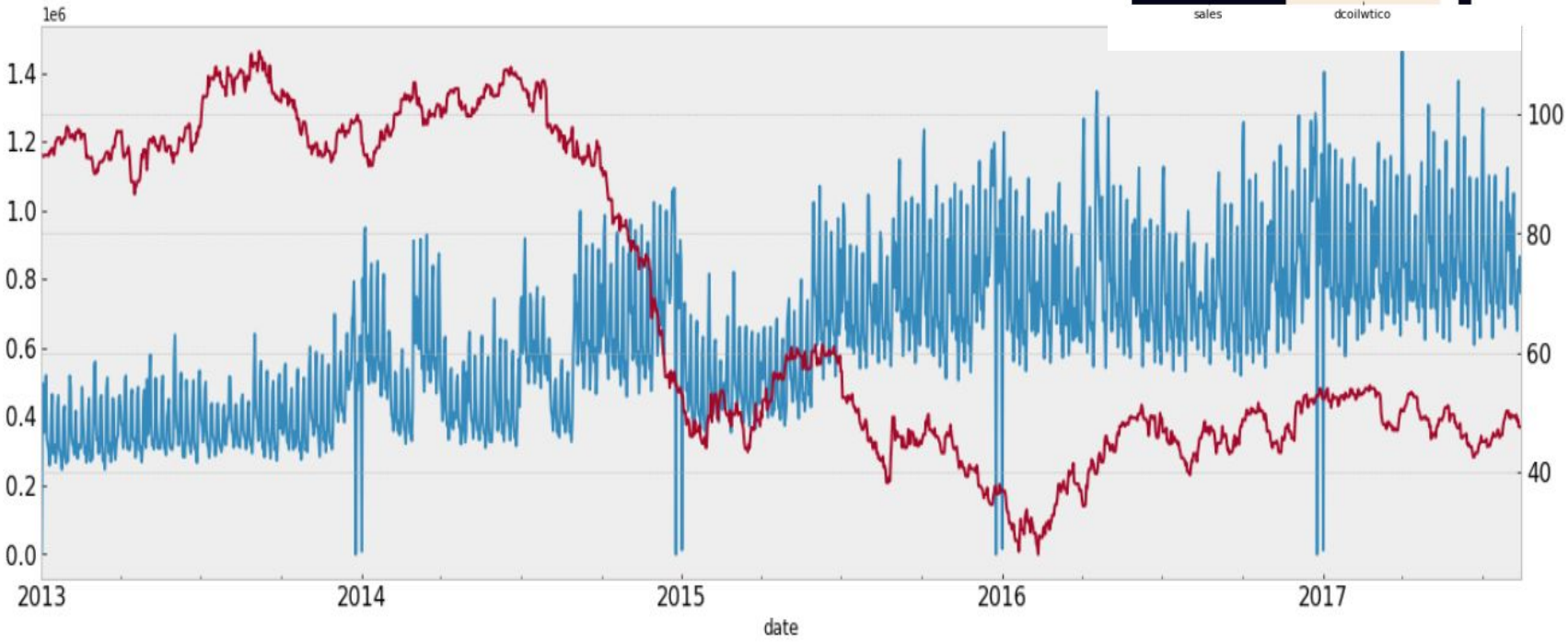
350

Avg Sales by Weekdays



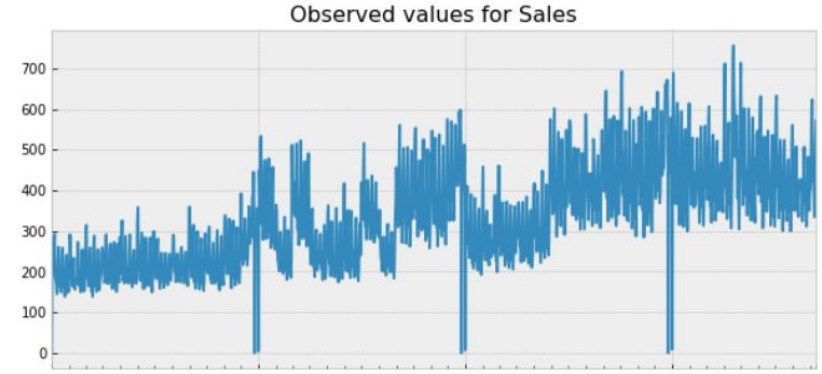
Sales vs Oil Price

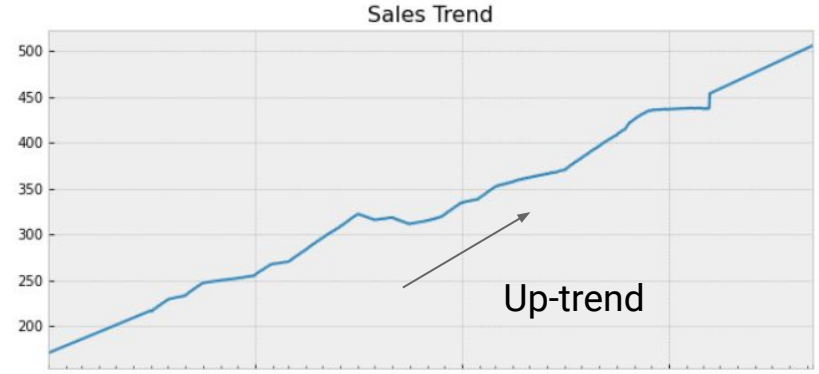


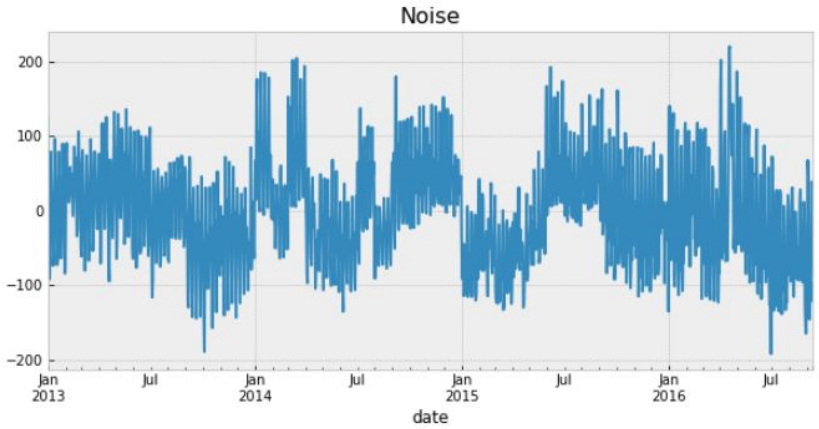


Modelling 3

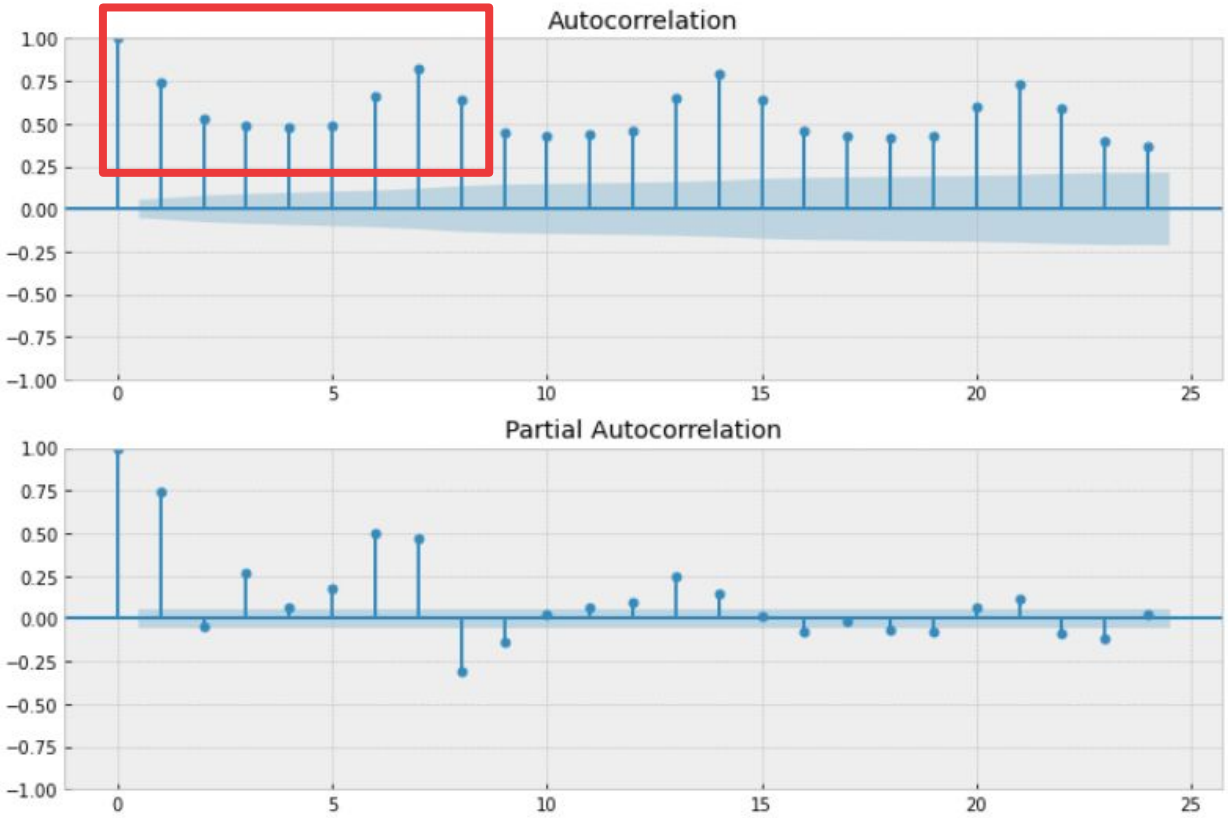
Pattern Recognition for Time Series



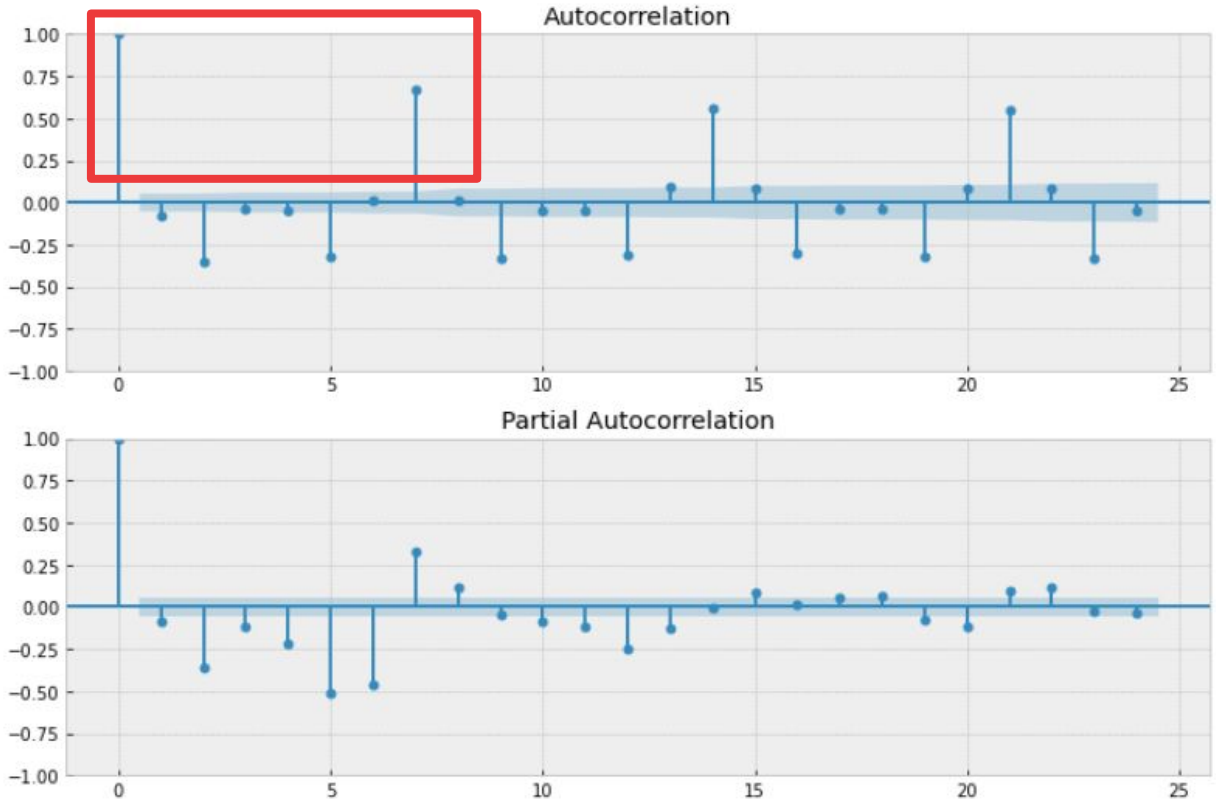




Correlogram



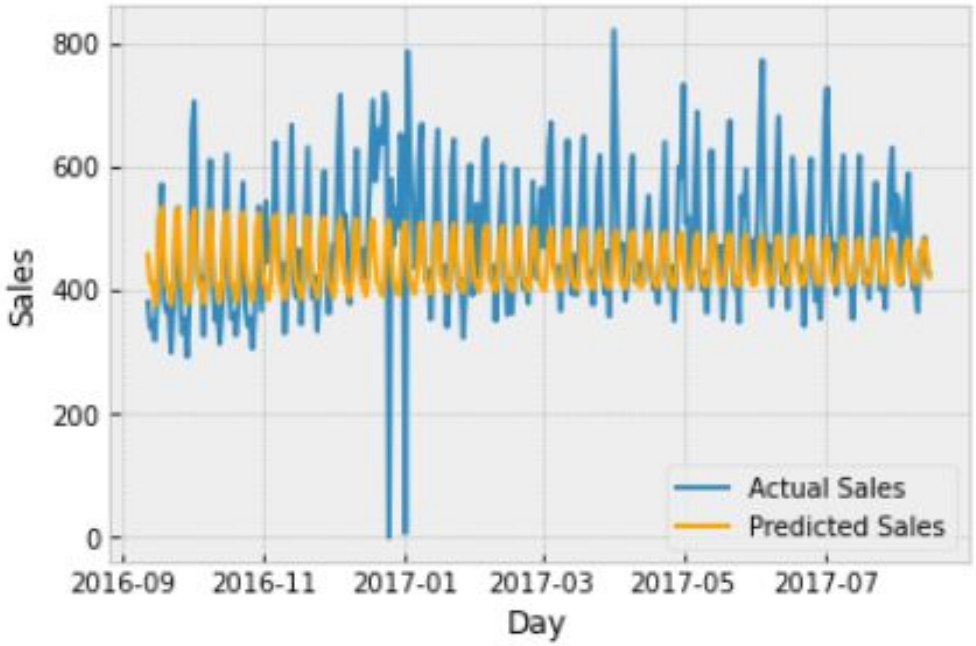
Correlogram



Model 1 - ARIMA

RMSLE	0.4402	AutoRegression Moving Average

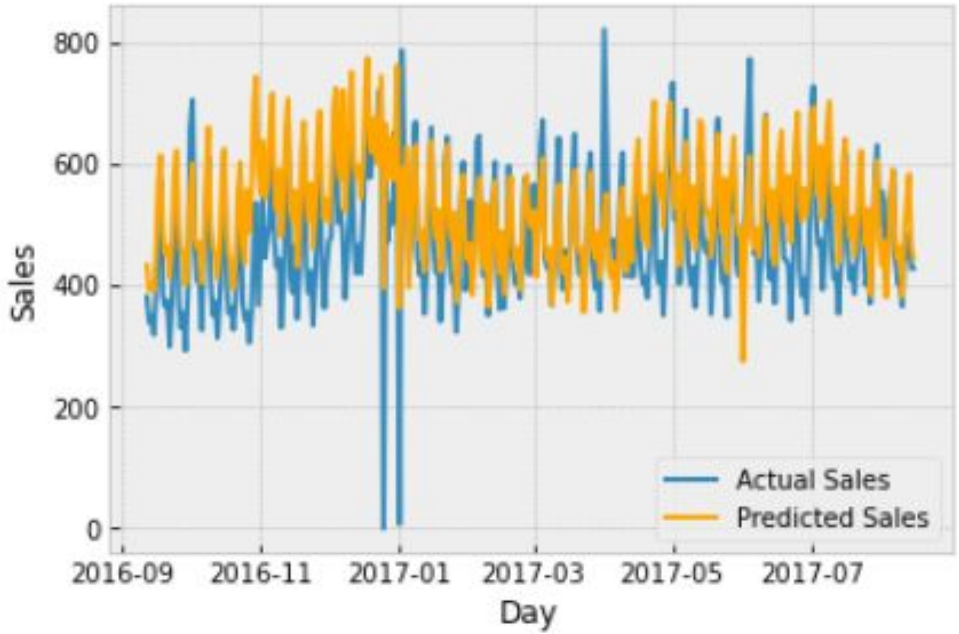
Date & Sales Feature Only



Model 2 - SARIMAX

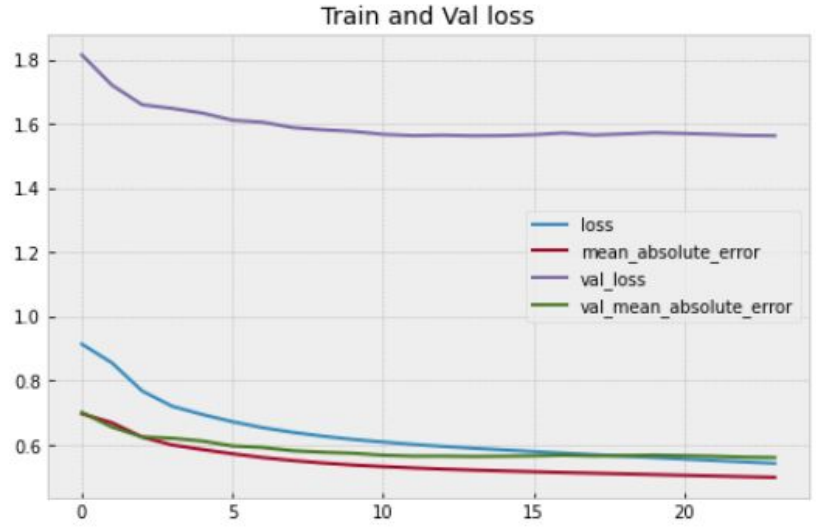
RMSLE	0.4254	Seasonal AutoRegression Moving Average with eXogenous Factor

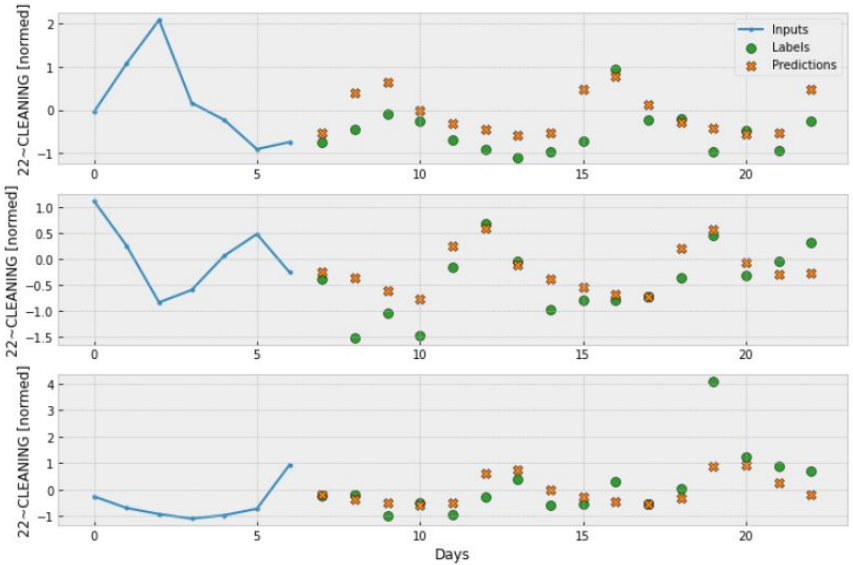
Date, Sales, Onpromotion, isHoliday Features



Model 3 - LSTM

RMSLE	0.564	Long short-term memory



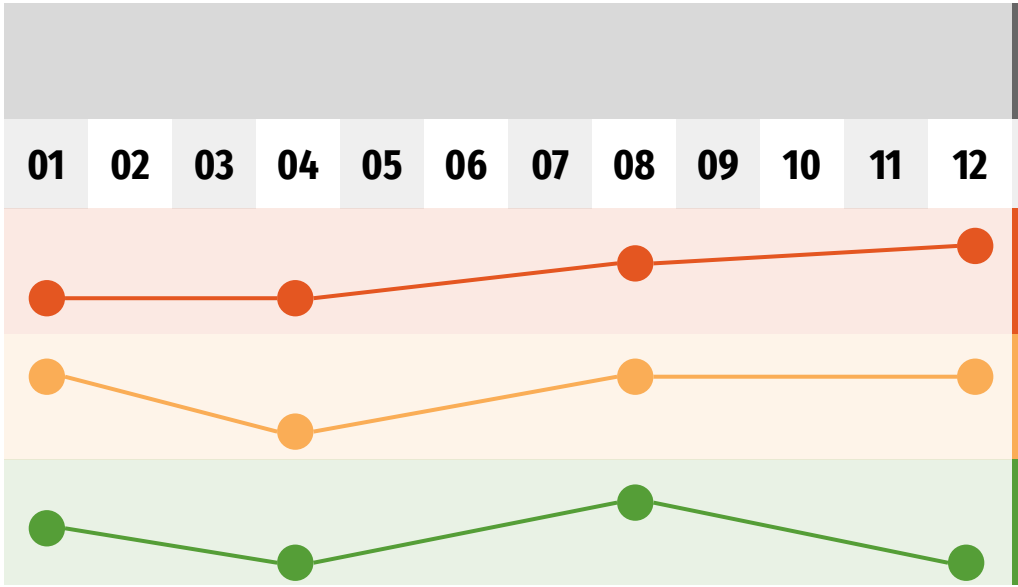


Type of Modelling Used

ARIMA

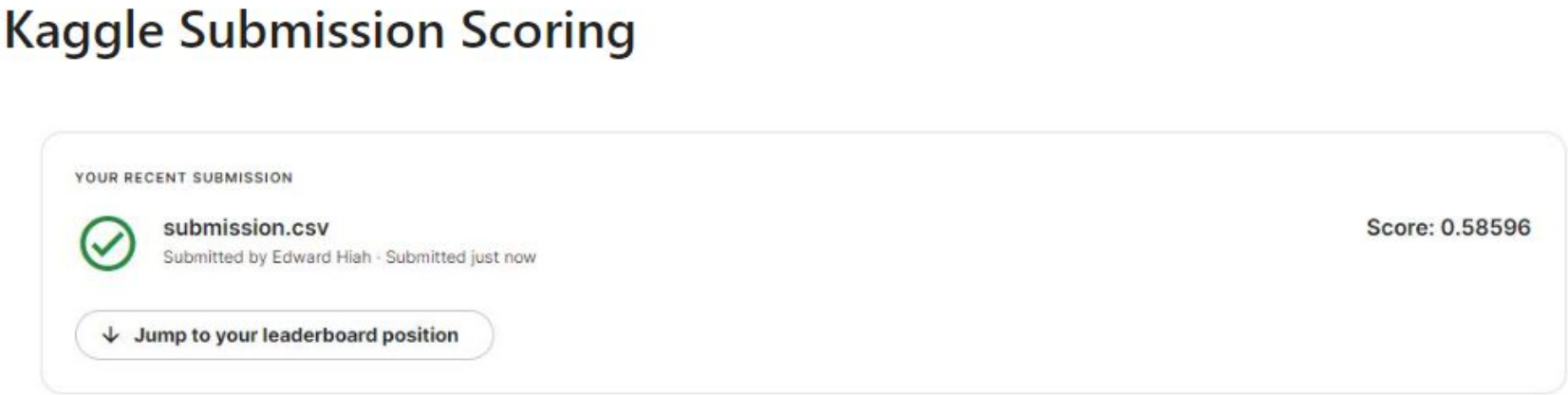
SARIMAX

LSTM



RSMLE	Modelling
0.44	ARIMA
0.42	SARIMAX
.564	LSTM

Kaggle Score



Conclusion

4

Conclusion

- The best results are taken from Univariate LSTM (with rolling window of 1 year)
- The SARIMA model outperformed the LSTM for the long term prediction task, but has it limitation to run by stores and product families.
- LSTM model was more robust, although RMSLE wasn’t very good for Long window period but it did quite well if the window for 1 year.
- Might be further consider XGBoost or LGBM or FB Prophet for further ML forecasting

End of Presentation

Thank You

2022