Predicting Petroleum Products Demand

Spark Ed Case Study, Phillips 66 & NCSA Datathon

Patrick Liu, Chengbo Li, Xinkai Zhao

Table of Contents

- Executive Summary

- Introduction

- Methodology

- Results

- Conclusion

Executive Summary

- Summary of methodologies
  - Data Wrangling
  - Models
    - Time Series Model- ARIMA
    - Time Series Model- Prophet
- Summary of all results
  - MAPE of Models
  - Prediction Graphs

Introduction

- Project background and context

  - Provide a forecast of US demand from Oct 22 to Mar 23

- Problems

  - What are potential factors shifting consumer demand?

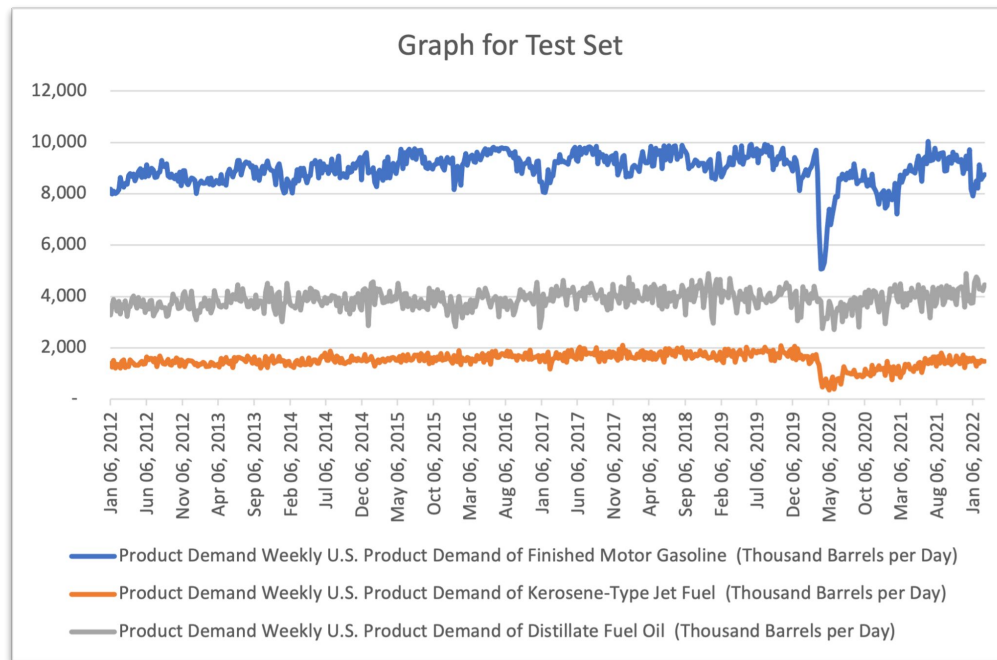 - What can Phillips 66 expect the market demand to be?
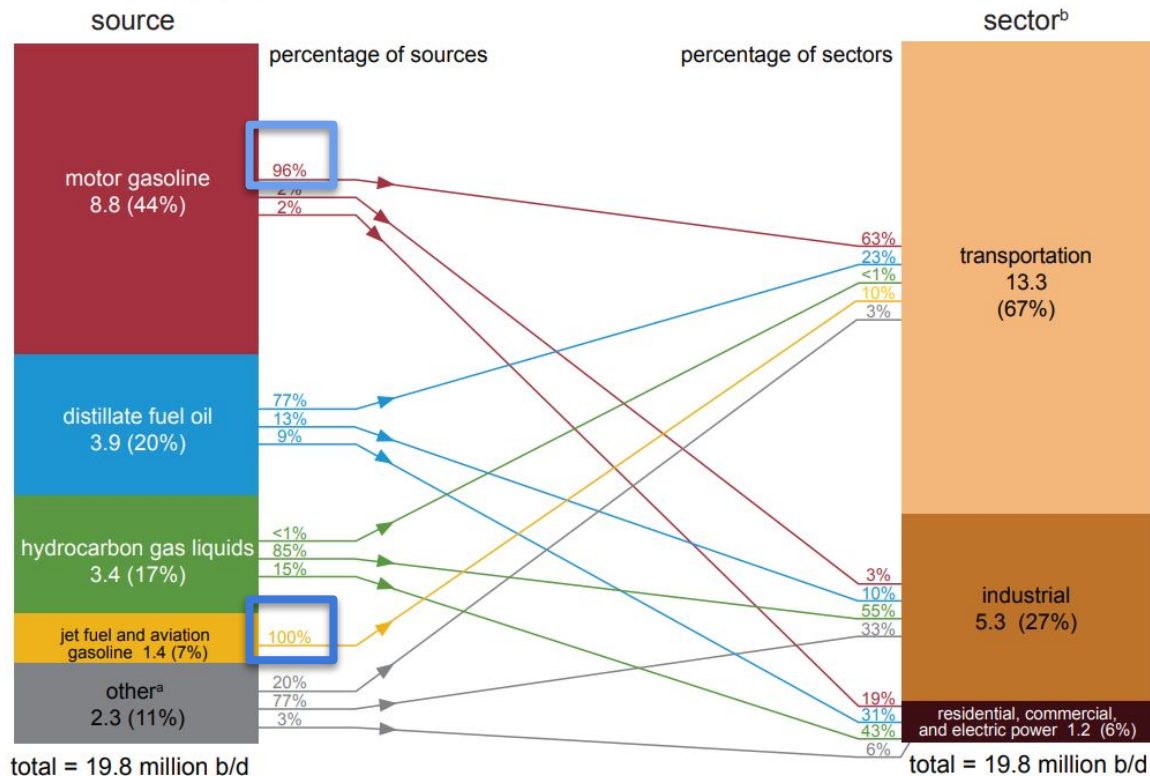
Methodology

Section 1

Data Collection

- Provided Test Set ->

- Notice:

  - Impact of Covid-19

U.S. petroleum products consumption by source and sector, 2021

million barrels per day (b/d)

source	percentage of sources	percentage of sectors	sector[b]

source

motor gasoline
8.8 (44%)

distillate fuel oil
3.9 (20%)

hydrocarbon gas liquids
3.4 (17%)

jet fuel and aviation
gasoline 1.4 (7%)

other[a]
2.3 (11%)

percentage of sources

96%
2%
2%

77%
13%
9%

<1%
85%
15%

100%

20%
77%
3%

percentage of sectors

63%
23%
<1%
10%
3%

3%
10%
55%
33%

19%
31%
43%
6%

sector[b]

transportation
13.3
(67%)

industrial
5.3 (27%)

residential, commercial,
and electric power 1.2 (6%)

total = 19.8 million b/d

total = 19.8 million b/d

eia

Data Collection
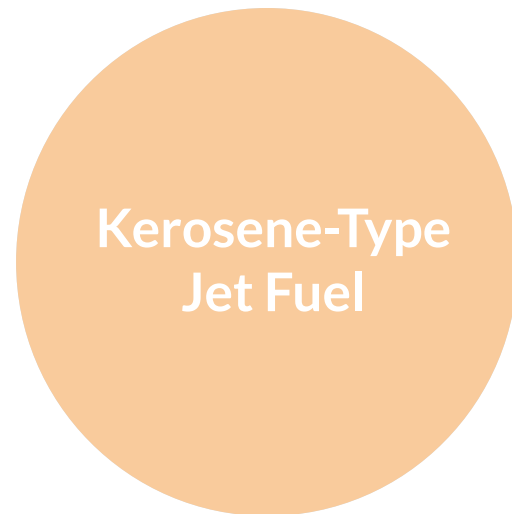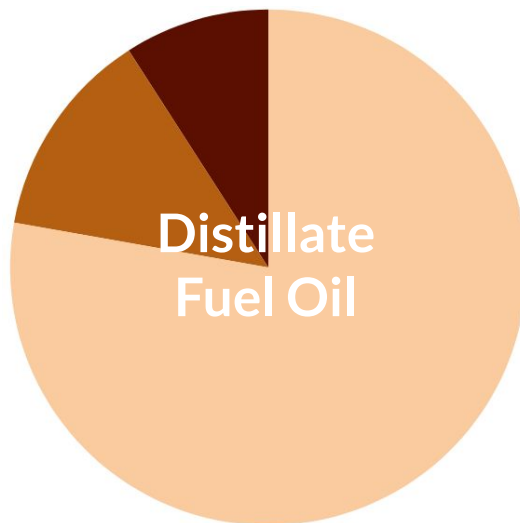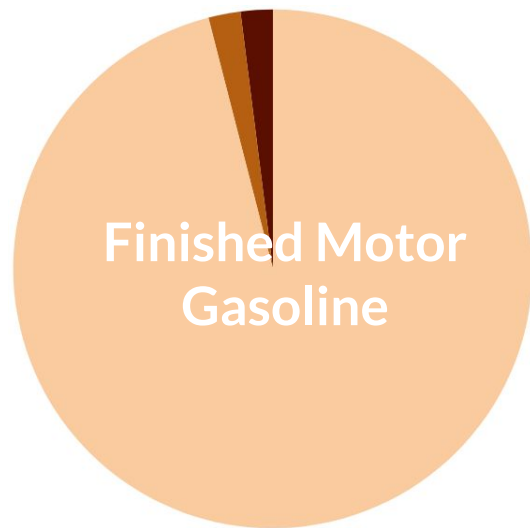
Transportation ● Industrial ● Residential, Commnercial, and Electric Power

Predictors

Finished Motor Gasoline

- Vehicle Miles
- Total Sales of Motor Vehicles

Distillate Fuel Oil

- Rail Freight Intermodal
- Rail Passenger Miles
- Heating Degree Days
- Electricity generated from petroleum
- Vehicle Miles

Kerosene-Type Jet Fuel

- Airplane Total Travel Distance

Different Predictors

Deriving with Different Predictors

- **Gas Price**

    - The law of demand

    - Since oil has <u>low elasticity of demand</u>, its demand is not quite responsive to price changes

    - *Not including price as predictor*
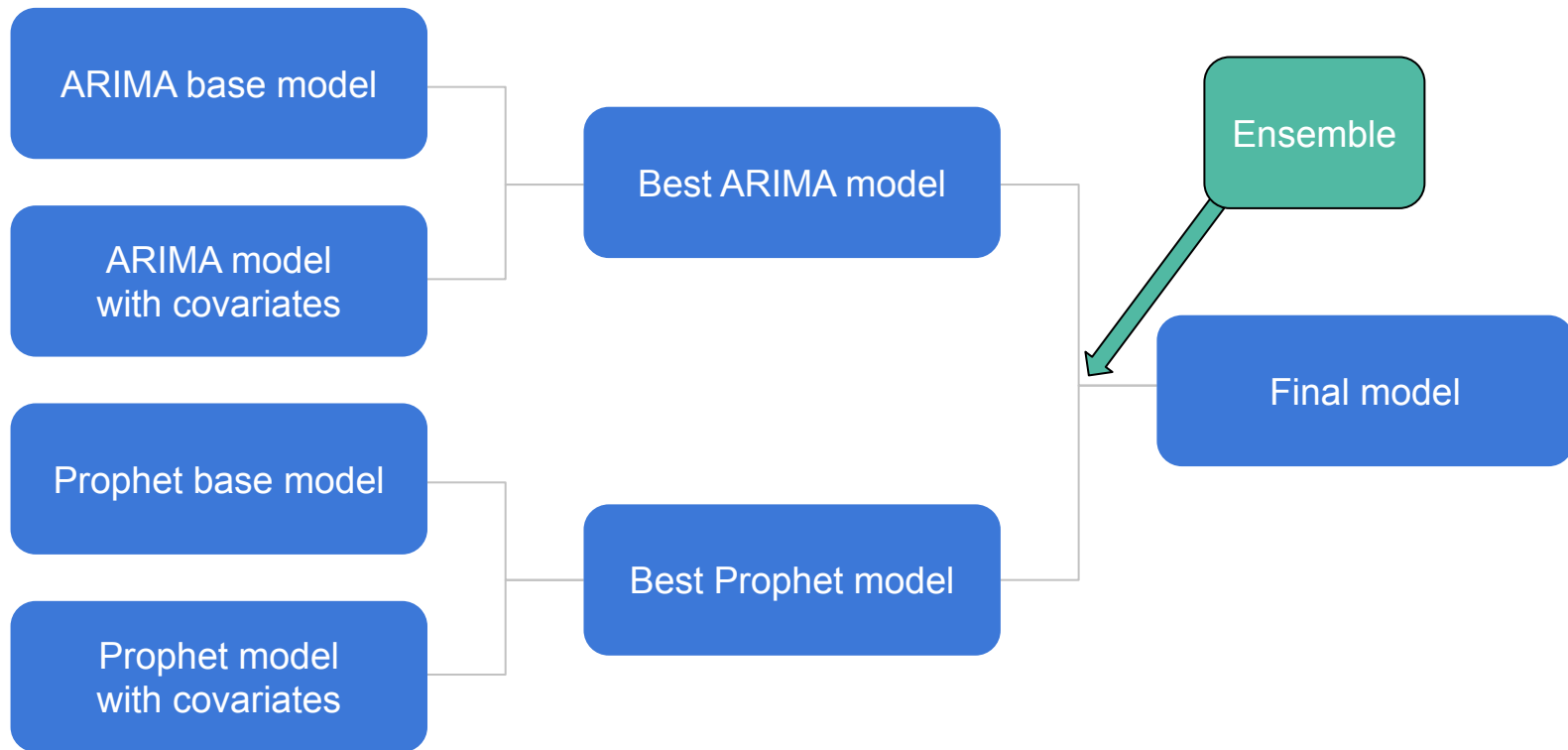
Demand Shifters

- Common Shifter of Demand
  - Policy Change/ Prevailing of pandemic

- Demand of Individuals
  - Employment/ Unemployment Rate
- Demand of Companies
  - Consumer Demand of different manufacturing products
  - Supply chain management

PHILLIPS 66

PROVIDING ENERGY.
IMPROVING LIVES.

Model Overview

Time Series Model - ARIMA

ARIMA(p,d,q)(P,D,Q)[m]

- Non-seasonal part (p,d,q)
  - AutoRegressive Integrated Moving Average
    - Autoregressive (p): linear combination of past variables
    - Moving average (q): weighted moving average of the past few forecast errors
    - Differencing (d): differences between consecutive observations
- Seasonal part (P,D,Q)
  - Similar, but with backshift of seasonal period
- Number of observations per year (m)

ARIMA with Alteryx Designer Snapshot
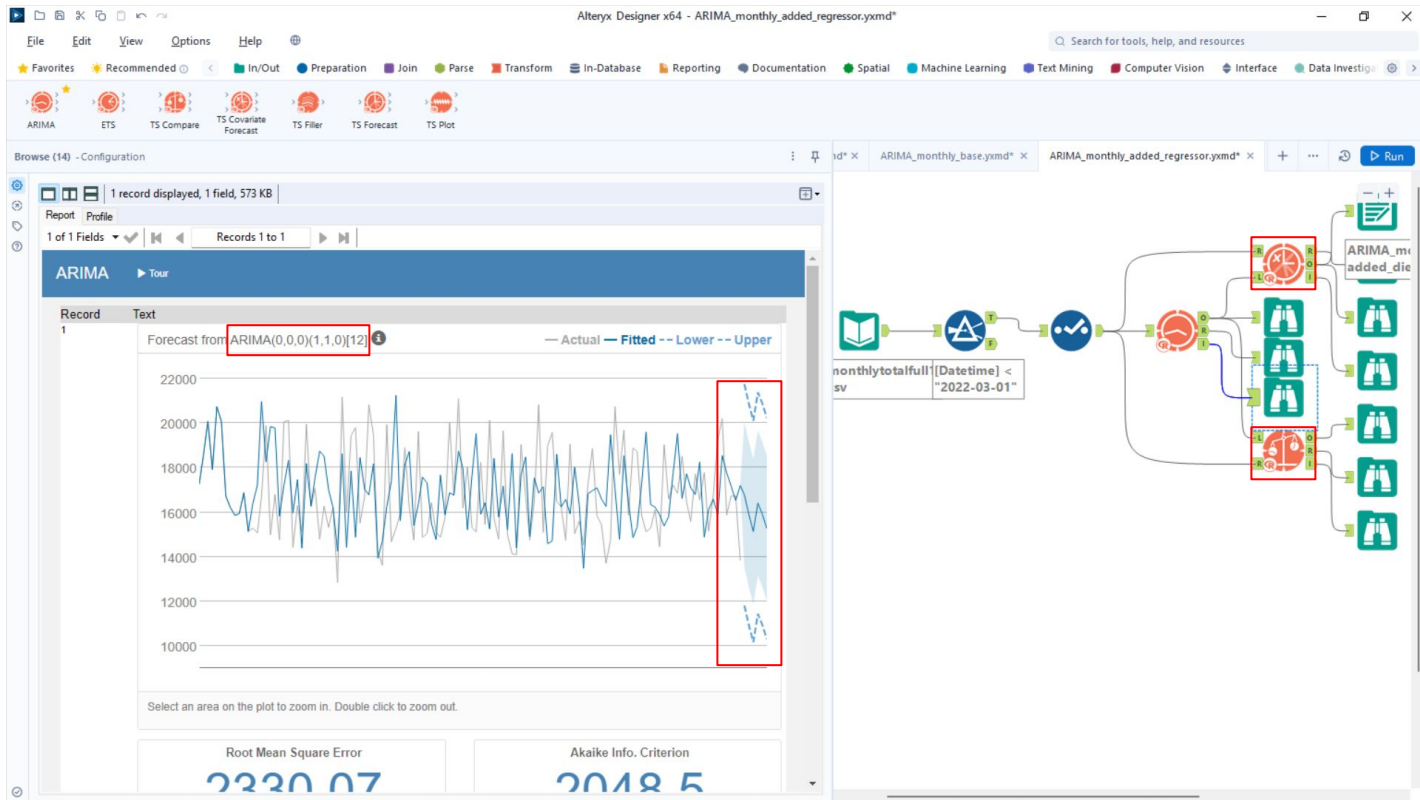
Time Series Model - Prophet

- Additive model where non-linear trends are fit with yearly, weekly, daily seasonality
- Works well with seasonal effects
- Needs several seasons of historical data
- Robust to shift in trend and outliers
- Open source by Facebook

https://github.com/facebook/prophet

Model Performance Evaluation Metrics

MAPE: Mean Absolute Percentage Error

$$M = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

M = mean absolute percentage error

n = number of times the summation iteration happens

A_t = actual value

F_t = forecast value

Insights

Section 2

Closer Look at the Product Demand

- **Covid-19**
- The sudden gap mainly caused by Covid-19 public health policies:
    - **[01/30/20, 02/03/22]** Federal quarantine order
    - **[03/15/20]** Public school locks down
    - **[03/19/20]** Non-mandatory stay-at-home order
- Prediction of this time period may be **lower** than other time period, regarding the presence of Covid

2020 Season1-2
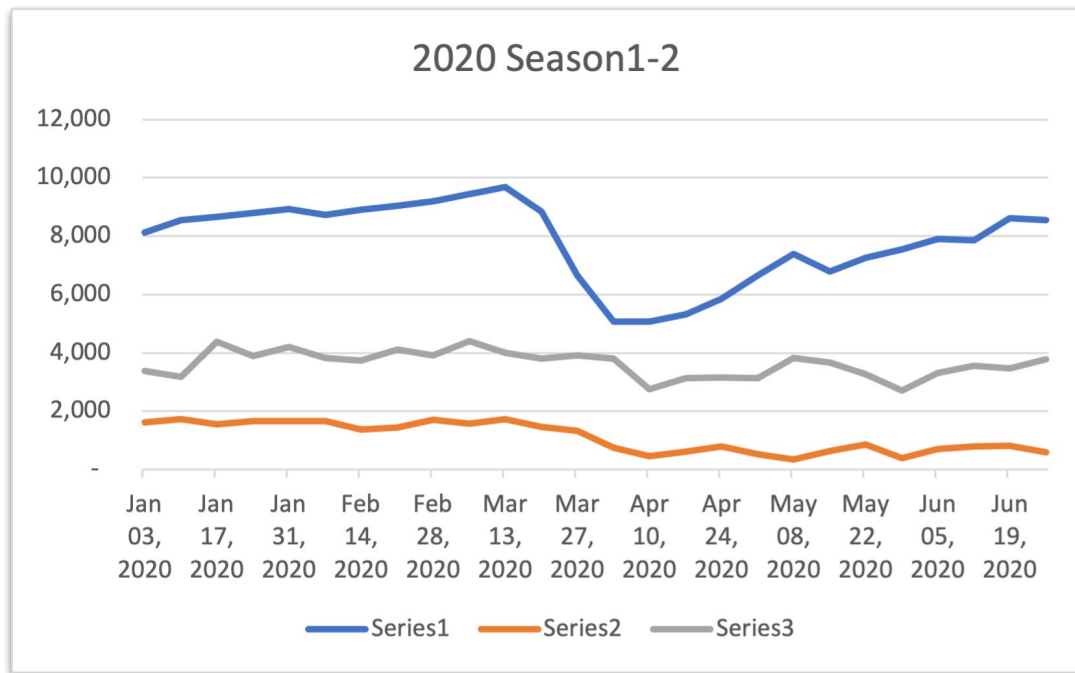
Trend in Covariates
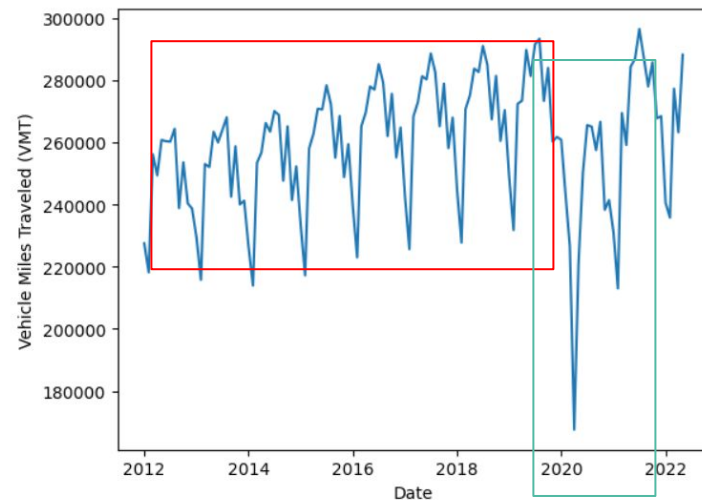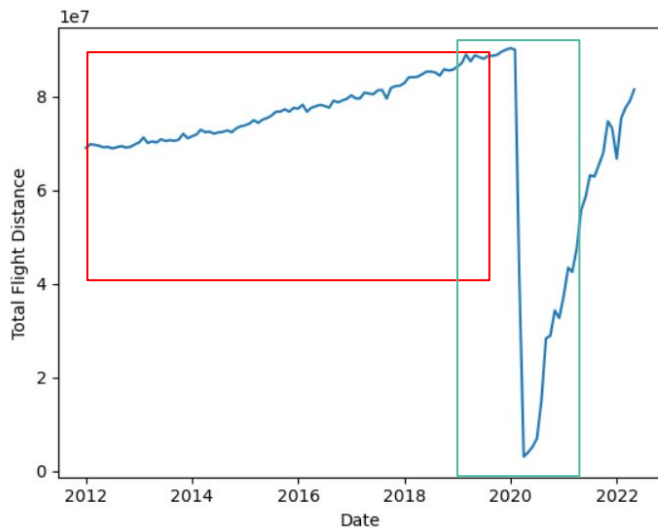
Similarities: Sudden decrease due to COVID
Differences: Seasonalities

Results

Section 3

Model Performance - Finished Motor Gasoline
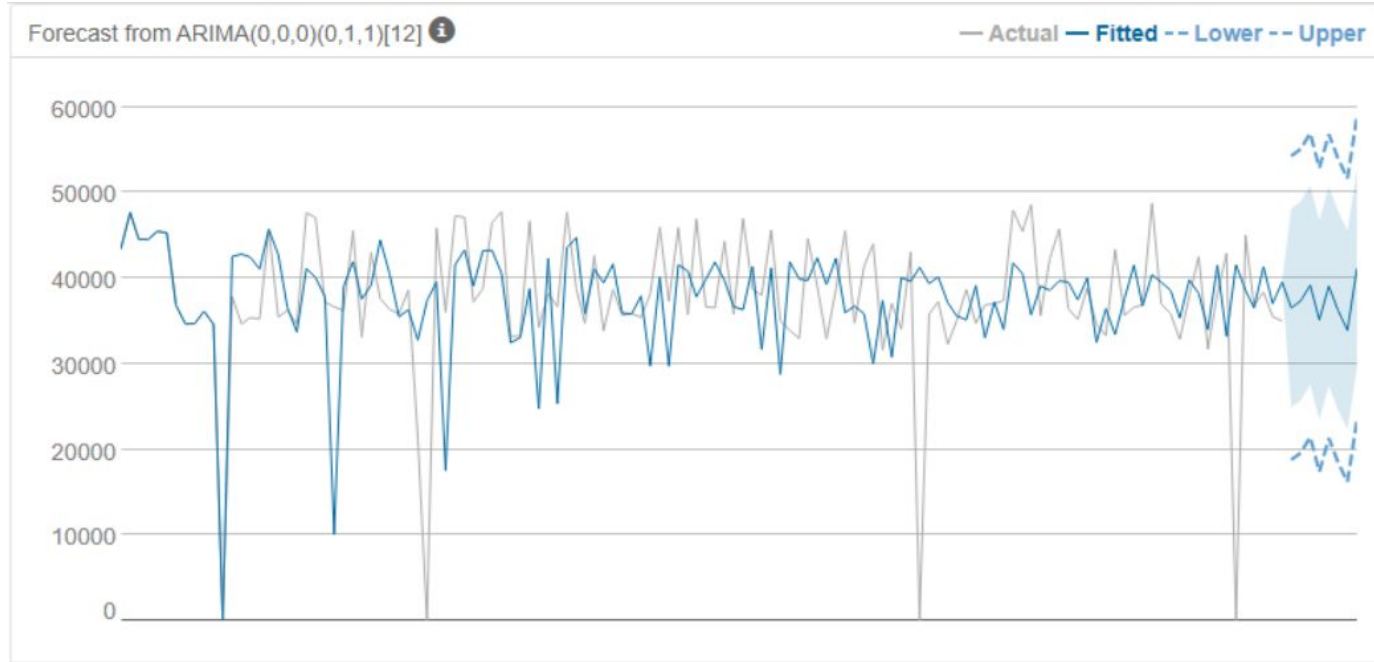
MAPE
- ARIMA base model: 13.81%
- ARIMA model with covariates: 11.89%
- Prophet base model: 15.12%
- Prophet model with covariates: 15.66%

MAPE of the different models

Finished Motor Gasoline - ARIMA with Covariates

Forecast from ARIMA(0,0,0)(0,1,1)[12]
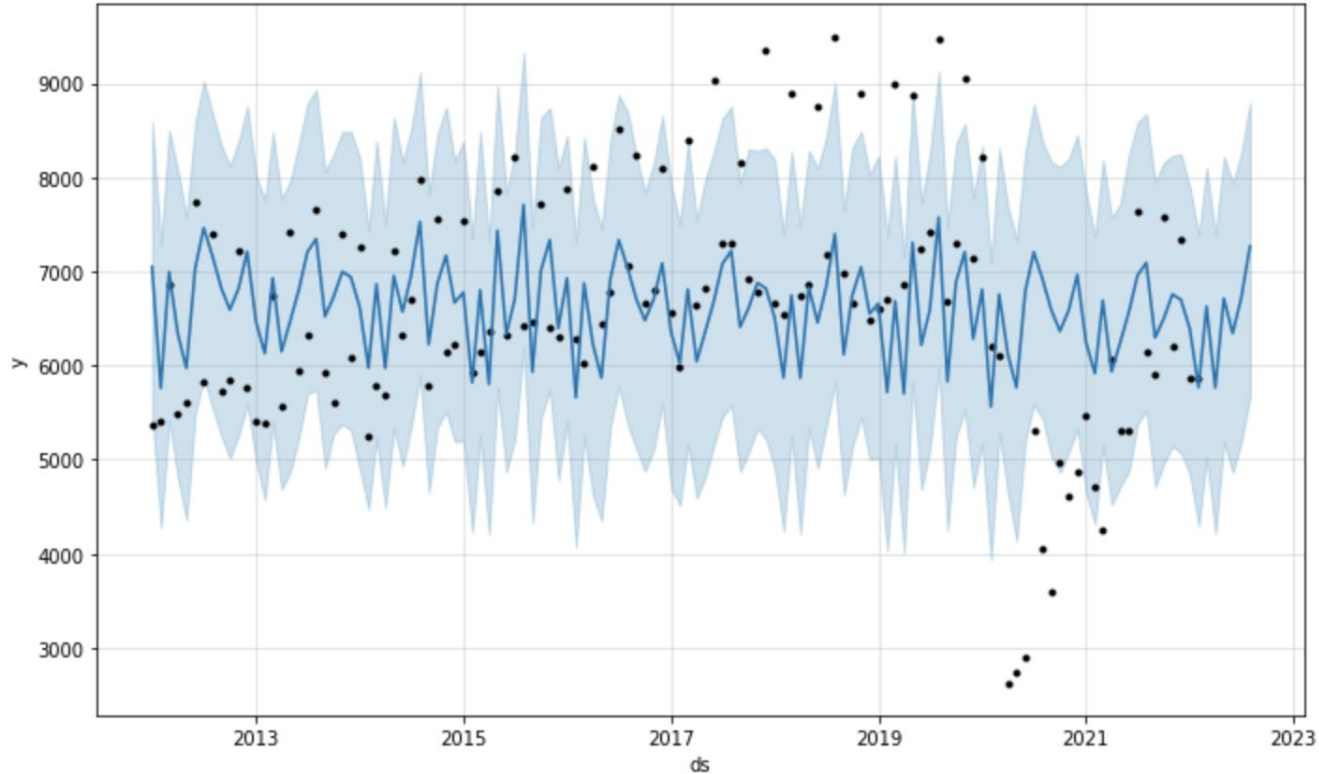
Model Performance - Kerosene-type jet fuel

MAPE
- ARIMA base model: 26.76%
- ARIMA model with covariates: 12.09%
- Prophet base model: 11.25%
- Prophet model with covariates: 14.81%
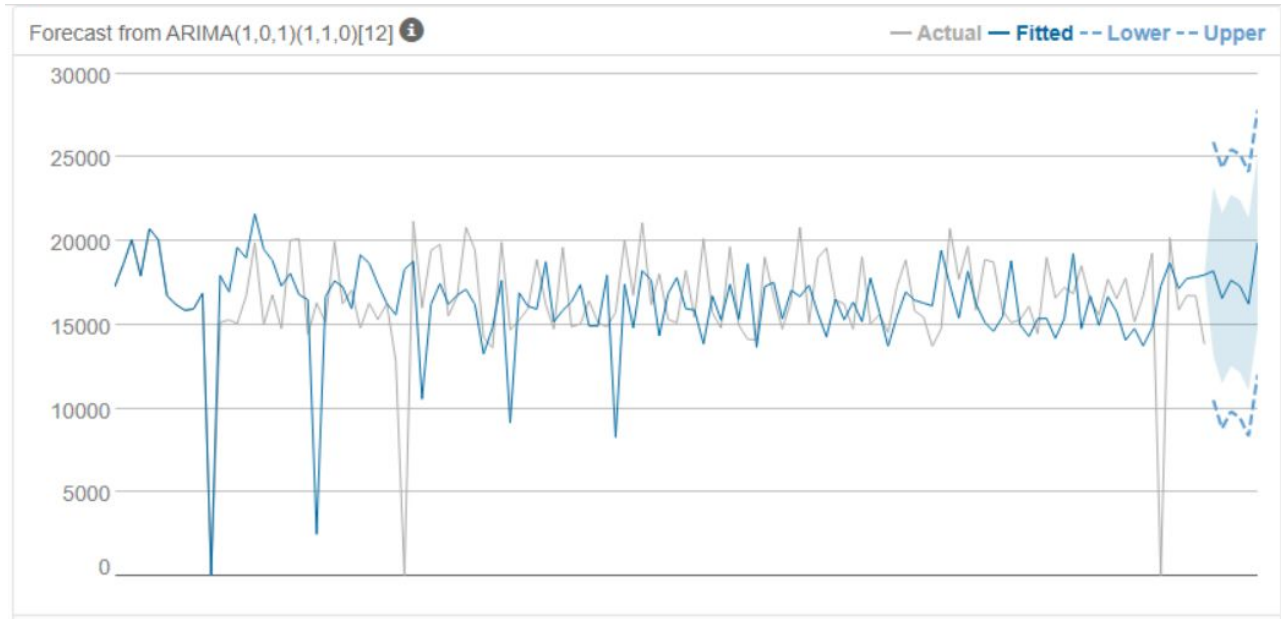
Kerosene-type jet fuel - Prophet Base Model

Model Performance - Distillate Fuel Oil

MAPE
- ARIMA base model: 10.27%
- ARIMA model with covariates: 9.16%
- Prophet base model: 16.05%
- Prophet model with covariates: 17.20%

Distillate Fuel Oil - ARIMA with Covariates

Prediction

Product Demand	Team 2		
Date	Monthly U.S. Product Demand of Finished Motor Gasoline (Thousand Barrels per Day)	Monthly U.S. Product Demand of Kerosene-Type Jet Fuel (Thousand Barrels per Day)	Monthly U.S. Product Demand of Distillate Fuel Oil (Thousand Barrels per Day)
Oct 01, 2022	34277.86639	6817.901255	15712.83109
Nov 01, 2022	39149.13412	7409.818318	19728.29671
Dec 01, 2022	36289.09148	6765.543343	10095.53105
Jan 01, 2023	33938.08566	6638.399633	19156.53406
Feb 01, 2023	41520.01859	6094.401022	16042.74379
Mar 01, 2023	33429.27938	7366.52388	17717.89816

Conclusion

- In general, ARIMA model works better with the petroleum product demand data
- The sudden decrease of demand due to COVID-19 pandemic tend to drive future prediction lower
- Limitation and future work:
  - Limited availability of weekly data for current predictors
  - Feature engineering for political events, society turmoil, extreme weather, etc.

Q & A

Welcome to ask us any questions

Thank you

Team 2