Holidays and COVID-19

Allen Lang (al8he)

CS 4774 Final Project Presentation

12/8/2020

Motivation

- Holiday season approaches
- Time of frequent travel/gatherings
- Determine effect of holiday travel/festivities on COVID-19 cases

Background: Pandemic in US

- Implemented quarantine measures in response to COVID-19
- Heated debate in US about “re-openings”
- Studies show COVID-19 spreads faster indoors rather than outdoors
- Concerns about winter time and the holiday season

Lockdown protests opposing COVID-19 restrictions began in mid-April as people grew weary of polices that they felt were unnecessary.

George Floyd incident sparked nationwide protests, pushing BLM back to the forefront of news. Protests began late-May/early-June.

Background: Forecasting

- High demand in many sectors
  - Economics and finance
  - Weather and geosystems
- Low supply of quality forecasting
  - Lack of experts
  - Difficult to tune/understand models
- Contemporary methods: ARIMA, ETS

Background: Prophet

"We are, in effect, framing the forecasting problem as a curve-fitting exercise" – Taylor and Letham

- Facebook Core Data Science Team Python module
- Simplification of time-series statistical analysis

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Trend
- Non-periodic changes
- Logistic or linear
- Change points

Seasonality
- Daily, weekly, monthly, yearly, etc.
- Fourier Series

Holiday and Events
- Common, predictable shocks without periodicity
- Custom or built-in holidays by country

Error term
- Assumed to be normally distributed

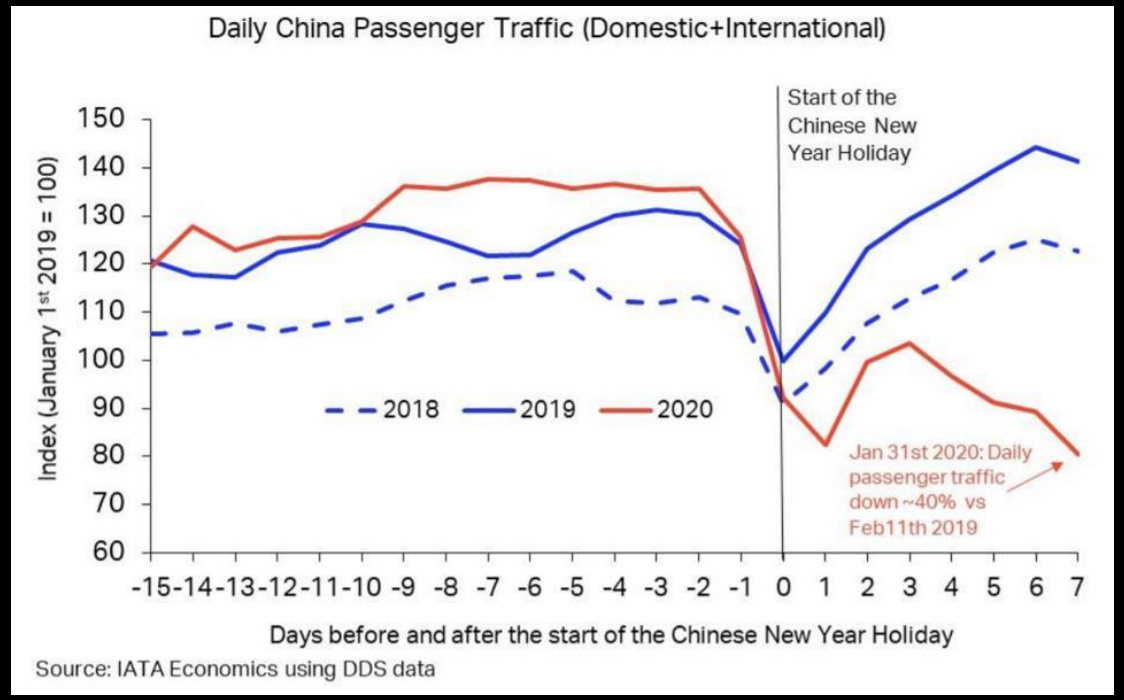
Related Work

- COVID-19 forecasting
  - Compilation of models from various groups
  - Compartmental model of infectious diseases
  - Accounting for human factors
- Prophet forecasting
  - Wikipedia web page visits
  - Stock price predictions
  - Biking usage rates

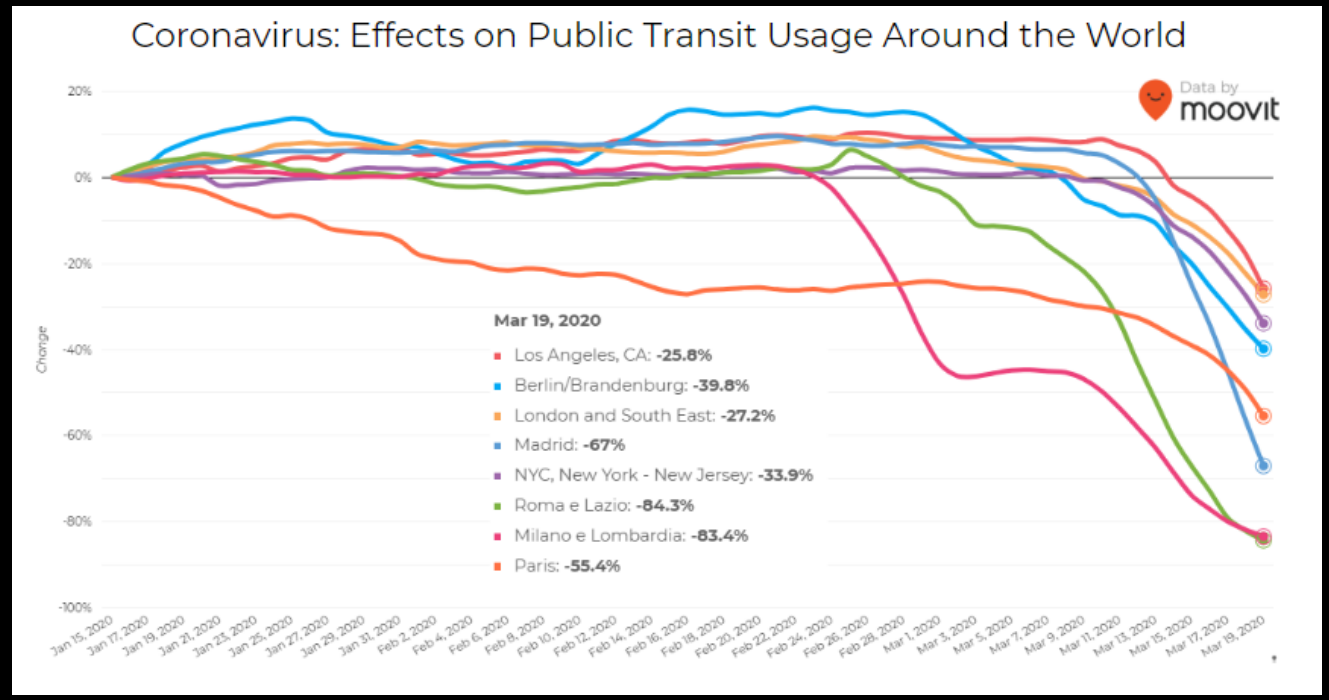
Target Task

- Claim: holidays/special events lead to greater increase in COVID-19
- Identify patterns/trends in COVID-19 statistics during and after a holiday period

Intuitive Figure



Effect of COVID pandemic on daily movement in China



Effect of COVID pandemic on daily movement across different cities

Proposed Solution and Contribution

- Construct Prophet time-series model
- Fine tune model with hyperparameters
- Introduce outside data sets

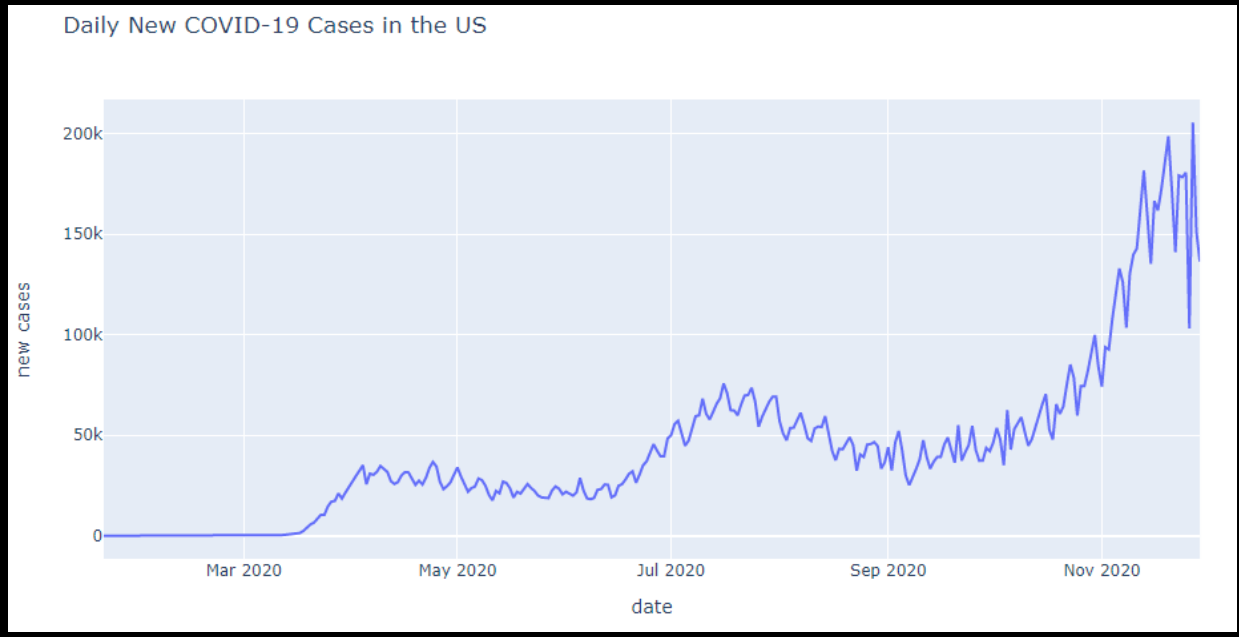
1. Build/Tune model that accounts for holidays and special events when forecasting COVID-19 cases
2. Better understanding of movements during pandemic, holidays, and holidays during pandemic

Implementation

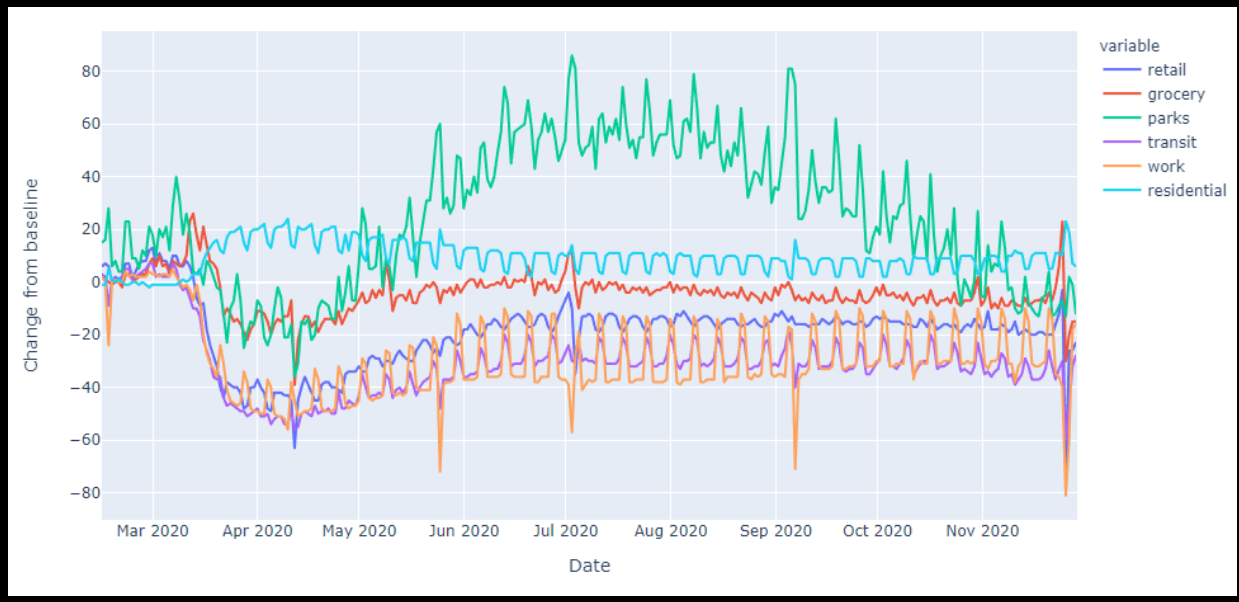
- Cleaning of Data with Numpy/Pandas
- Construction of baseline model with Prophet
- Hyperparameter tuning, figures with Plotly
  - Seasonality/Trends
  - Holidays
  - External regressors
- Prediction forecast

Data Summary

1. New York Times COVID-19 cases/deaths
2. National holidays across the world
3. Stringency index of national governments
4. Google Mobility Reports



Daily new cases of COVID-19 in the United States.

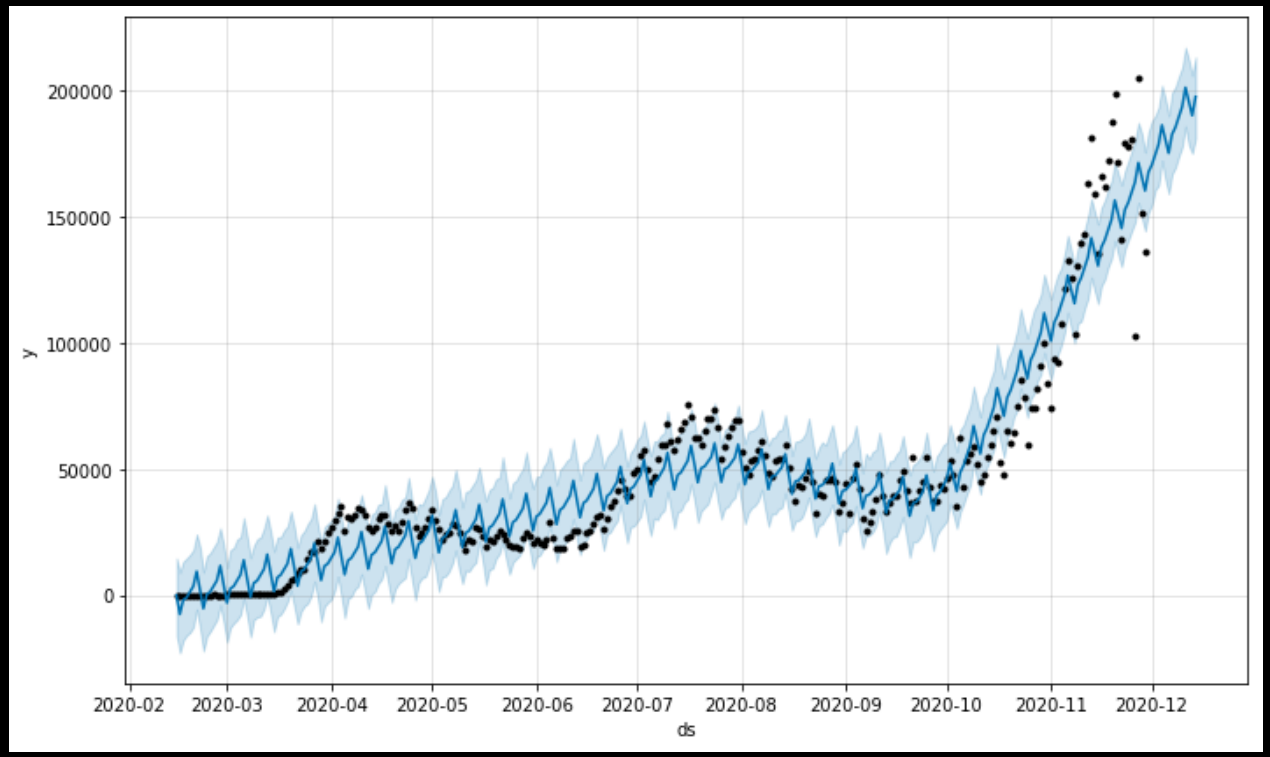


Google Mobility data during the COVID-19 pandemic.

1. https://www.kaggle.com/paultimothymooney/nytimes-covid19-data
2. https://github.com/dr-prodigy/python-holidays
3. https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker
4. https://www.google.com/covid19/mobility/

Results/Analysis: Baseline

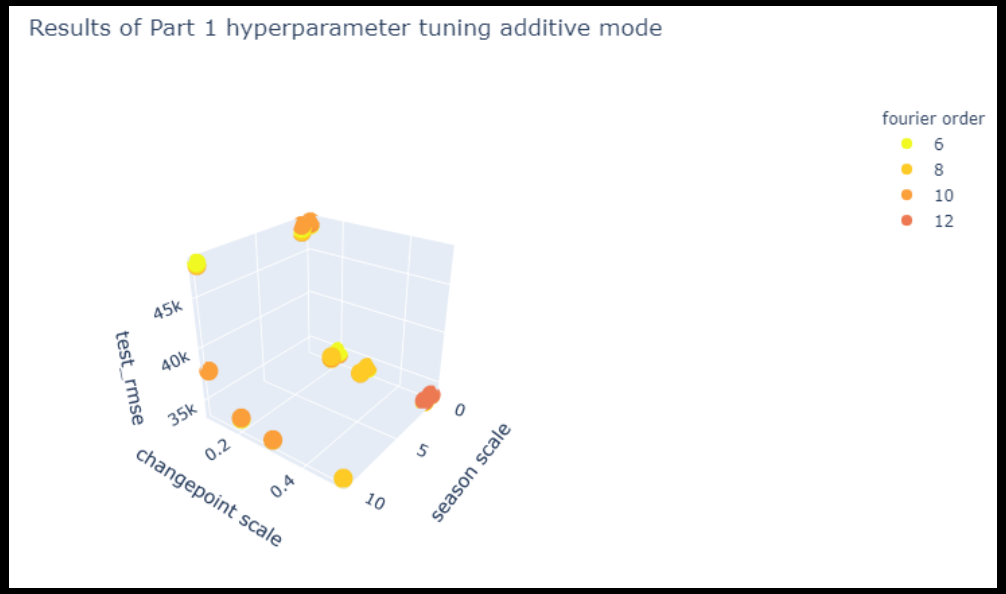
- Prophet model with COVID-19 new case data
- February 15th to November 29th
- Train RMSE: 11186.34
- Test RMSE: 73059.23



Forecasted model with no parameters

Results/Analysis: Tuning without Holidays

- Seasonality: type of seasonality
- Seasonality scale: flexibility of seasonality
- Changepoint scale: flexibility of trend between changepoints
- Fourier order: order of the Fourier series used for seasonality



Result of hyperparameter tuning for additive mode during part 1.

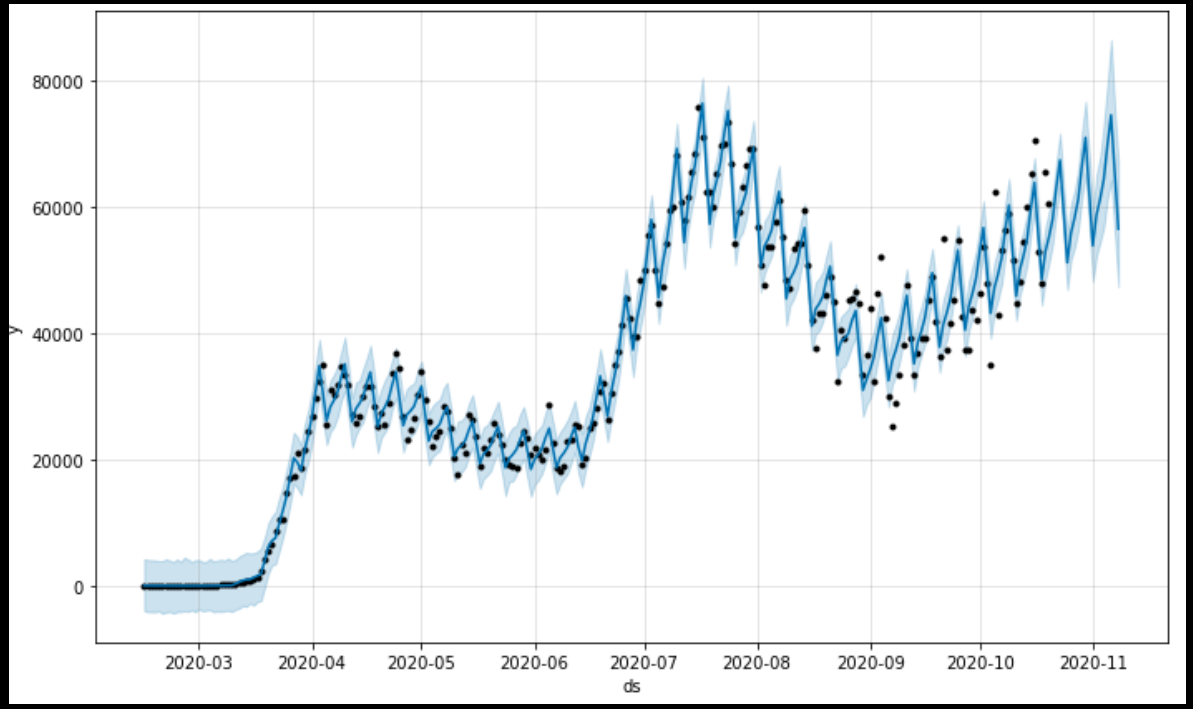
```
Seasonality types: ['additive', 'multiplicative']
Seasonality scale: [0.1, 1, 10]
Changepoint scale: [0.1, 0.2, 0.3, 0.5]
Fourier Order: [6, 8, 10, 12]
```

Hyperparameter grid for part 1

Results/Analysis: Tuning without Holidays

- Best parameters:
  - Seasonality: multiplicative
  - Seasonality scale: 10
  - Changepoint scale: 0.5
  - Fourier order: 12

- RMSE:
  - Train: 3171.63
  - Test: 32726.62



Forecasted model using best parameters for part 1.

Results/Analysis: Tuning with Holidays

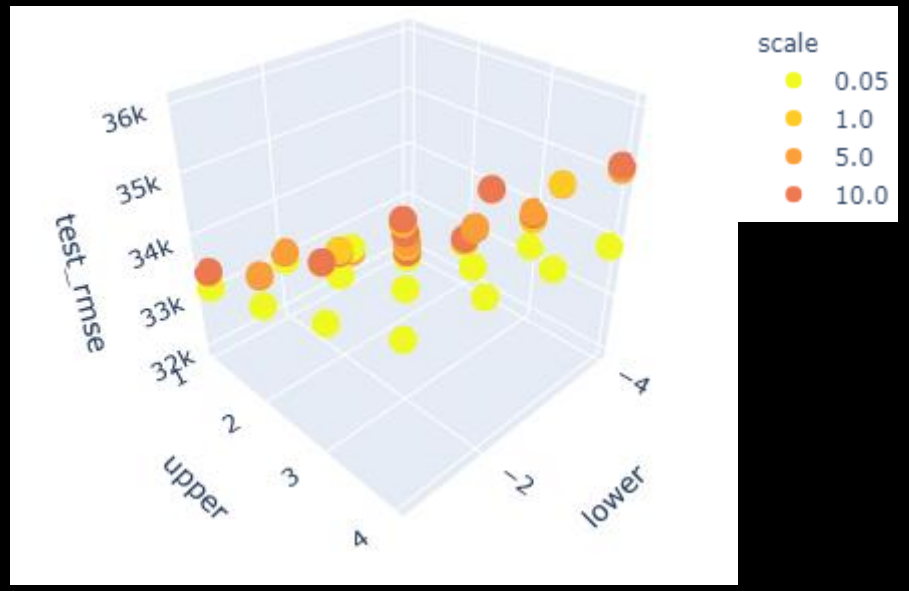
- Lower window: how far the holiday goes backward
- Upper window: how far the holiday goes forward
- Priority Scale: importance of each holiday as a "shock" to model

Results/Analysis: Tuning with Holidays

- Part 2: Tuning with only generated holidays

	ds	holiday
0	2020-02-17	Washington's Birthday
1	2020-05-25	Memorial Day
2	2020-07-04	Independence Day
3	2020-09-07	Labor Day
4	2020-10-12	Columbus Day
5	2020-11-11	Veterans Day
6	2020-11-26	Thanksgiving
7	2020-12-25	Christmas Day

Holidays used in part 2.



Results of hyperparameter tuning for part 2

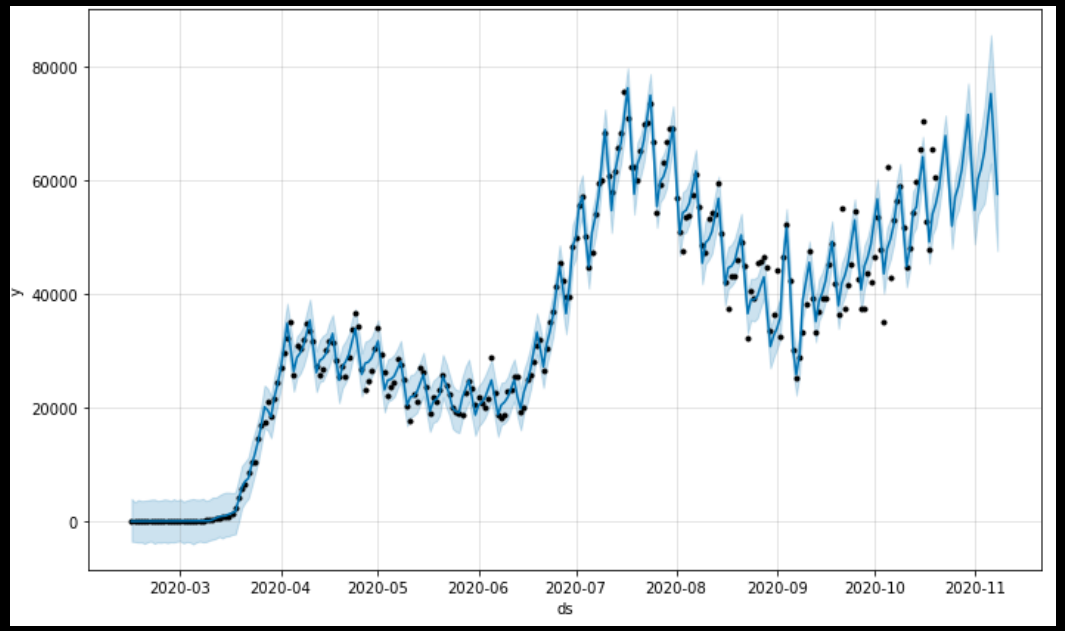
```
Lower window: [-1, -2, -3, -4]
Upper window: [1, 2, 3, 4]
Holiday scale [0.05, 1, 5, 10]
```

Hyperparameter grid for part 2 and 3

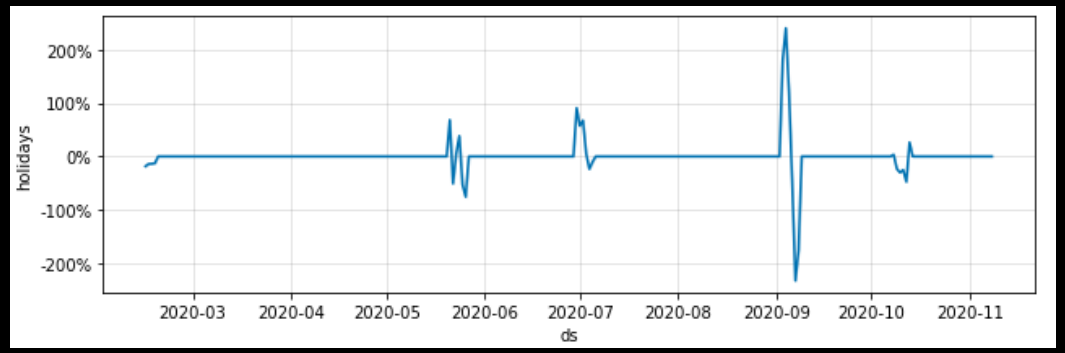
Results/Analysis: Tuning with Holidays

- Best parameters:
  - Lower window: -4
  - Upper window: 1
  - Holiday Prior Scale: 5

- RMSE:
  - Train: 2886.28
  - Test: 32057.55



Forecasted model using best parameters for part 2.



Holidays influence on forecasting for part 2.

Results/Analysis: Tuning with Holidays

- Part 3: Tuning with addition of special events

Stringency index for the US over the pandemic. Lessening of measures highlighted in red.

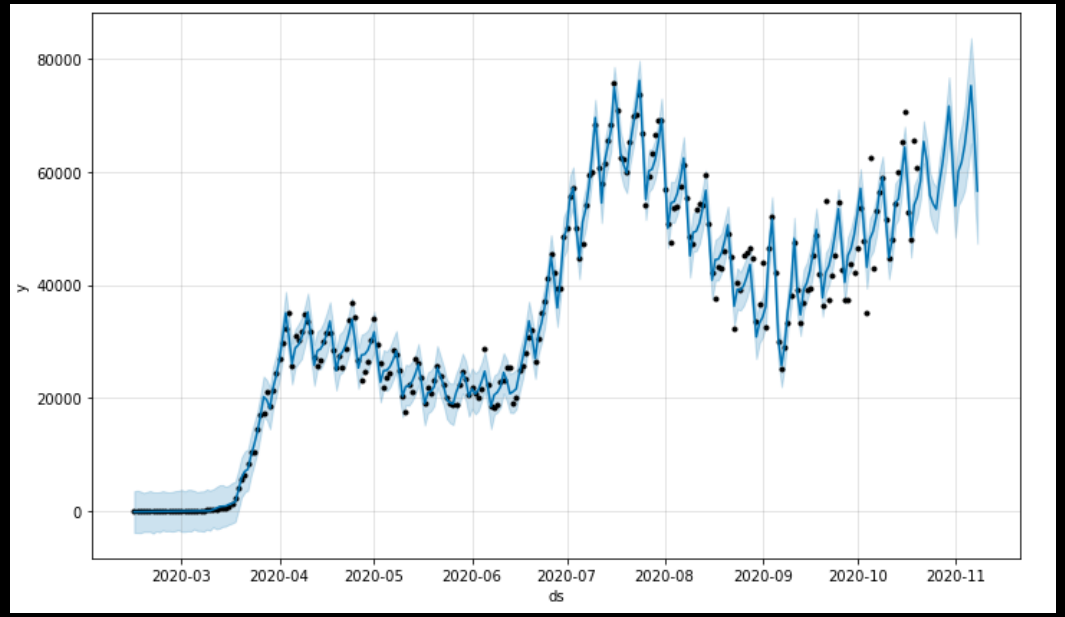
	ds	holiday
0	2020-02-17	Washington's Birthday
1	2020-05-25	Memorial Day
2	2020-06-01	BLM Protests
3	2020-06-15	Lessening of Quarantine Measures
4	2020-07-04	Independence Day
5	2020-07-20	Lessening of Quarantine Measures
6	2020-09-07	Labor Day
7	2020-09-12	Lessening of Quarantine Measures
8	2020-10-12	Columbus Day
9	2020-10-26	Lessening of Quarantine Measures
10	2020-10-31	Halloween
11	2020-11-03	Election Day
12	2020-11-11	Veterans Day
13	2020-11-26	Thanksgiving
14	2020-12-25	Christmas Day

Holidays used in part 3.

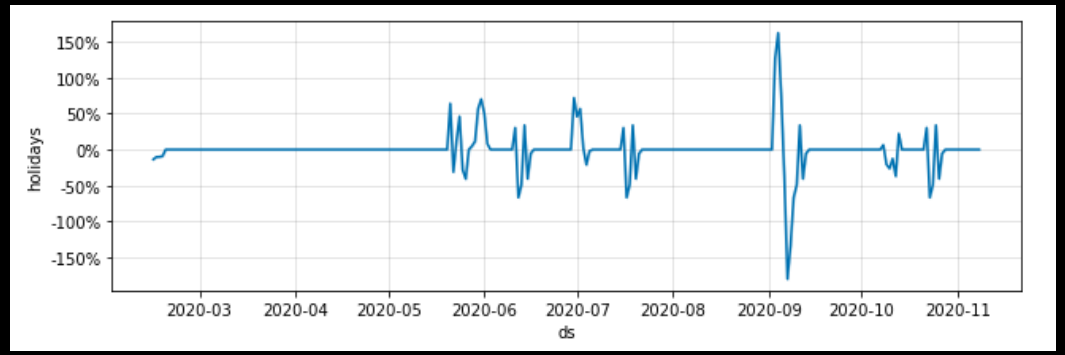
Results/Analysis: Tuning with Holidays

- Best parameters:
  - Lower window: -4
  - Upper window: 1
  - Holiday Prior Scale: 10

- RMSE:
  - Train: 2786.01
  - Test: 32712.98



Forecasted model using best parameters for part 3.

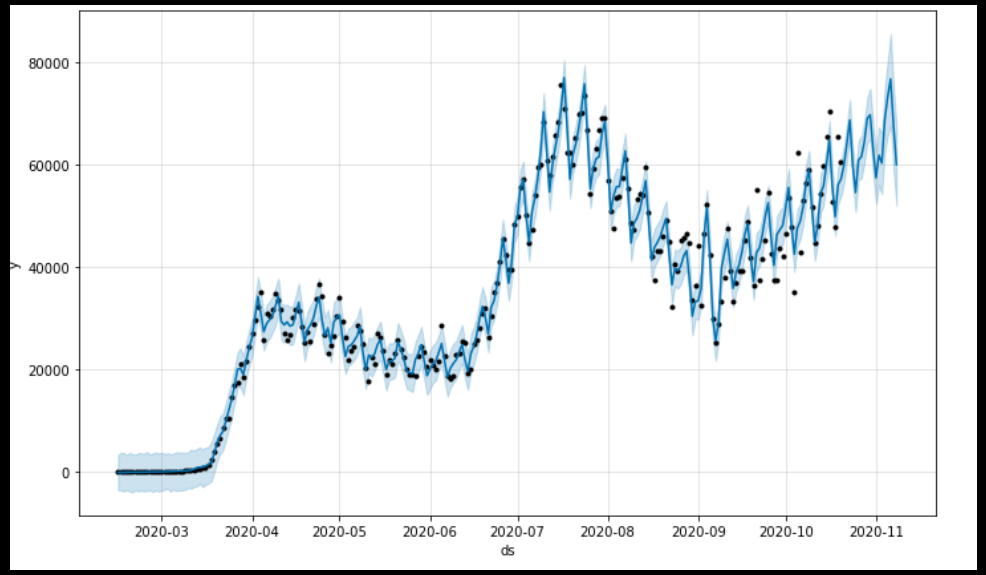


Holidays influence on forecasting for part 3.

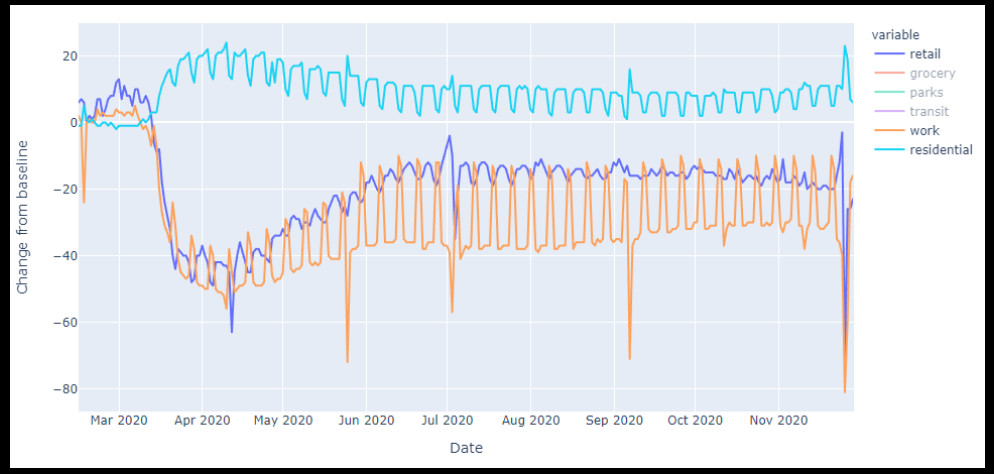
Results/Analysis: Tuning with Regressors

- Best regressors:
  - Residential: multiplicative, 5
  - Work: multiplicative, 1
  - Retail: multiplicative, 1

- RMSE:
  - Train: 2850.94
  - Test: 30731.03



Forecasted model using best parameters for part 5.



Retail, work, and residential mobility trends

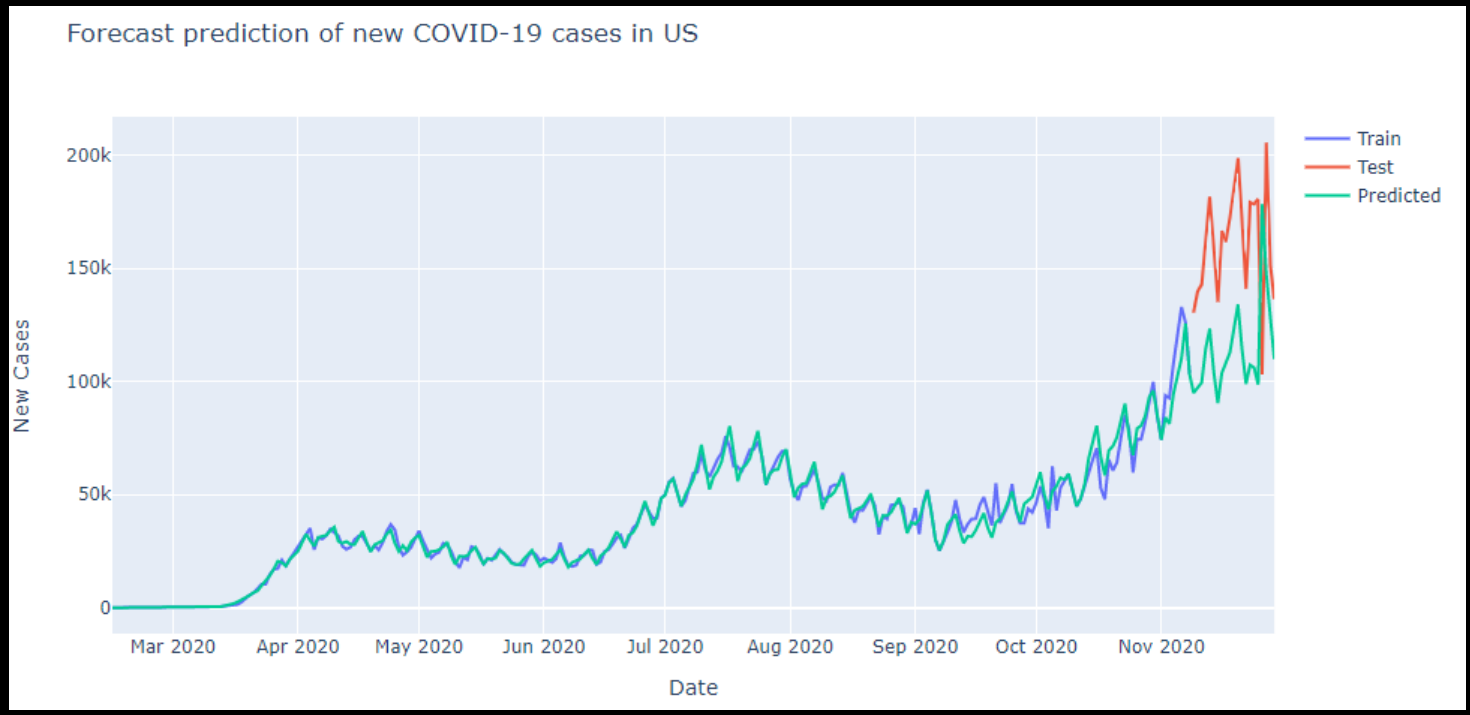
Results/Analysis: Tuning summary

Results of hyperparameter tuning

Tuning Stage	Train RMSE	Test RMSE
Part 1: Without holidays	3171.63	32726.62
Part 2: With holidays	2886.28	32057.55
Part 3: With holidays/events	2786.01	32712.98
Part 4: With holidays/events manual	3024.23	32700.53
Part 5: With regressors	2850.94	30731.03

Results/Analysis: Prediction

- Baseline:
  - Train RMSE: 11186.34
  - Test RMSE: 73059.23
- Final model:
  - Train RMSE: 4137.23
  - Test RMSE: 56454.53



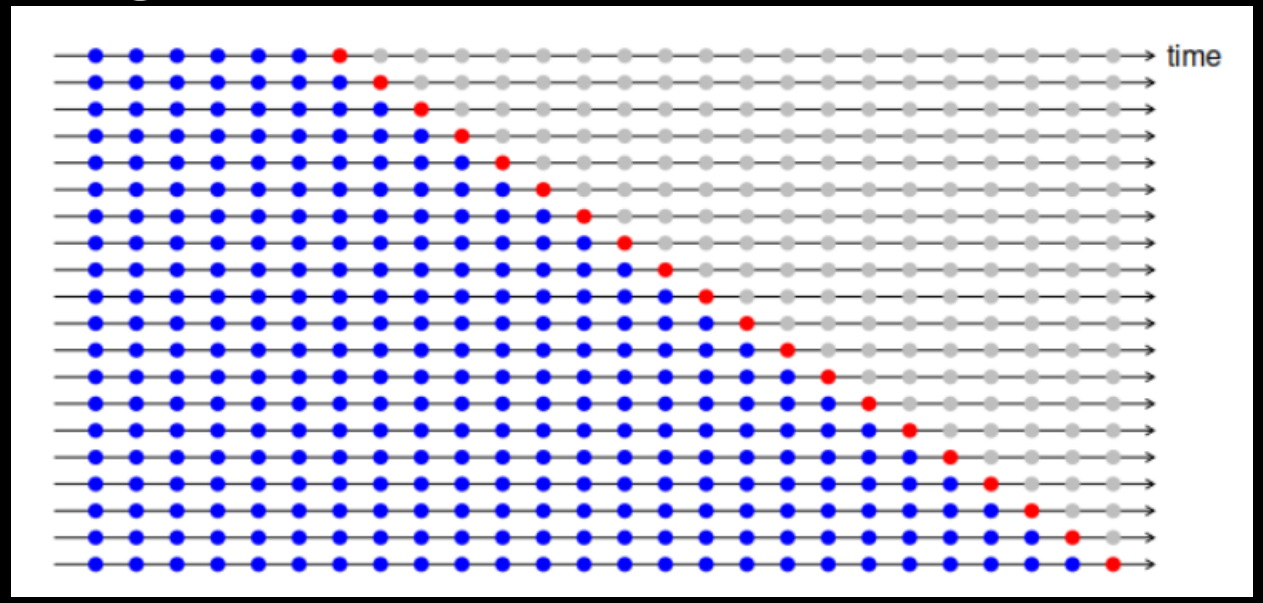
Predicted forecast model with tuned model.

Conclusion: Observations

- Inclusion of national holidays improved model performance
- Stark contrast between lower and upper window parameters
- Positive response to other time-series data

Future Work: Extensions within Prophet

- Metrics: looking into percent-error based
- Tuning methods: Prophet's built-in cross validation
- Holidays: precise way to optimizing each holiday's parameters
- Regressors: other time-series data like temperature



Graphic explaining cross validation for a time series

Future Work: Extensions outside Prophet

- Benchmarking: comparison with other methods
- Classical statistical methods
  - ARIMA
  - ETS
  - Holt-Winters
- Neural networks
  - LSTM, RNN
  - MLP

Thank You!

Professor Qi, Zhe, Jack, Arshdeep

References

- https://commons.wikimedia.org/wiki/File:CIMG_0355_(49799414598).jpg
- https://en.wikipedia.org/wiki/File:Minneapolis_05-28-20_(49947863357).jpg
- https://facebook.github.io/prophet/
- https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html
- https://covid19-projections.com/
- https://nbviewer.jupyter.org/github/nicolasfauchereau/Auckland_Cycling/blob/master/notebooks/Auckland_cycling_and_weather.ipynb
- https://www.kaggle.com/paultimothymooney/nytimes-covid19-data
- https://github.com/dr-prodigy/python-holidays
- https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker
- https://www.google.com/covid19/mobility/
- https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194889