Modeling Covid-19 in UK

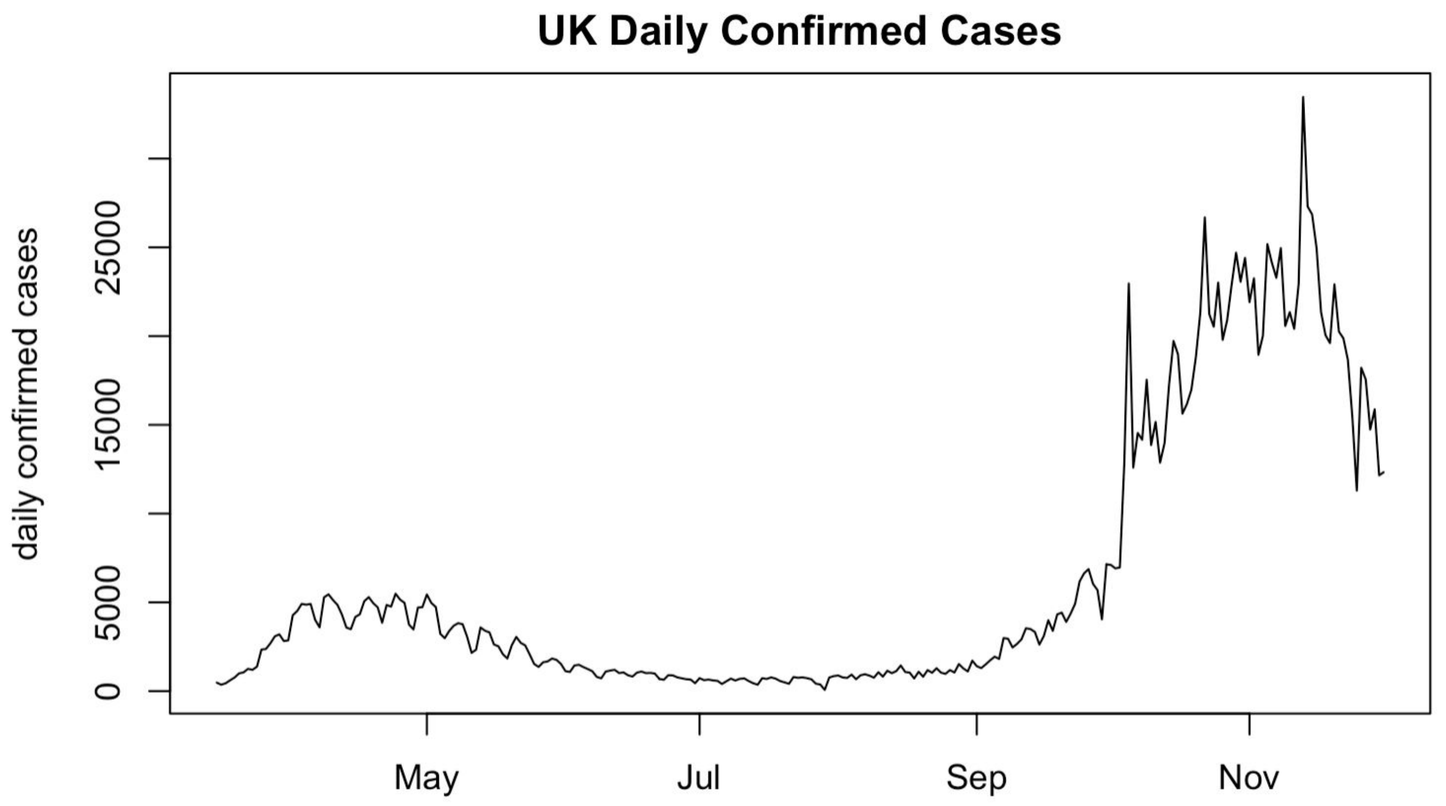
Andrew Hong, Kezhen Zhao, Lily Qiao, Yuhao Chen
Supervisors: Rebecca Nugent, Peter Freeman
12/16/2020

Project Objective

Covid 19 has been disrupting human health as well healthcare all over the world, and we can expect it to continue affecting us next year.

Objectives:

- Find potential leading indicators
- Make inferences on UK Covid data
- How to predict hotspots



* Source: *ECDC*. Consistent with *OurWorld Data*.

Leading Indicators - Mobility

Leading Indicators - Mobility

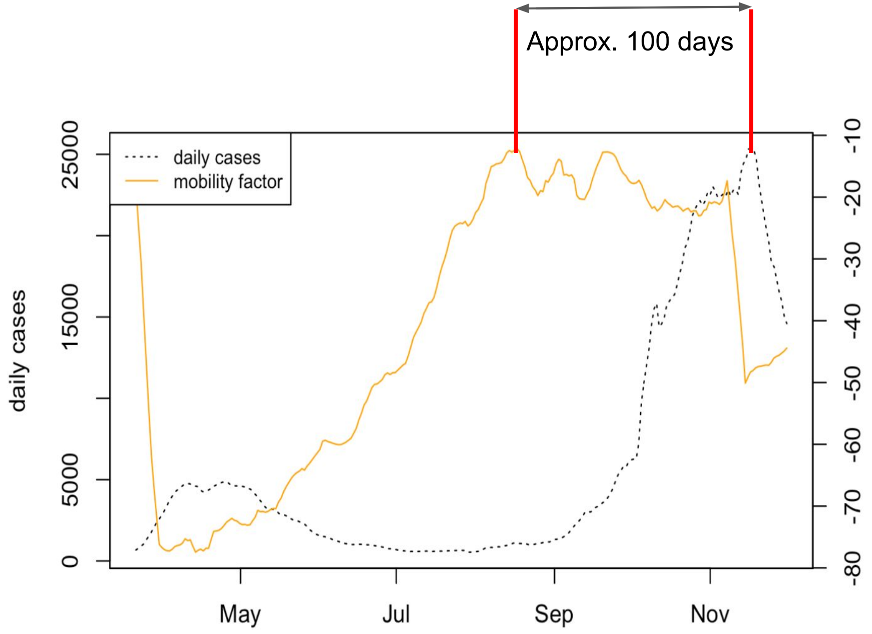
Report mobility patterns, i.e. movement trends (% change from pre-covid baseline) over time by geography, across different categories of places such as:

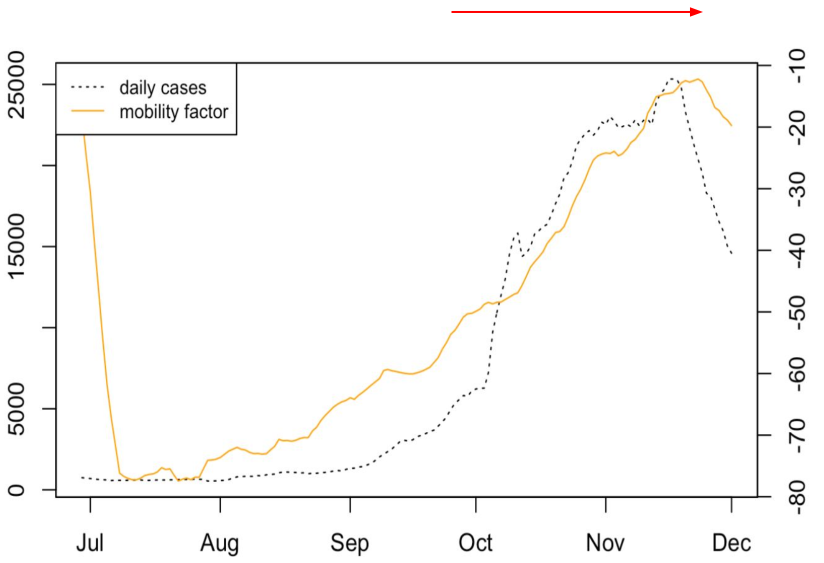
- retail and recreation,
- groceries and pharmacies,
- parks,
- transit stations,
- workplaces, and
- residential.

Using MA7 (Moving Averages) to account for any recording delays over the week.

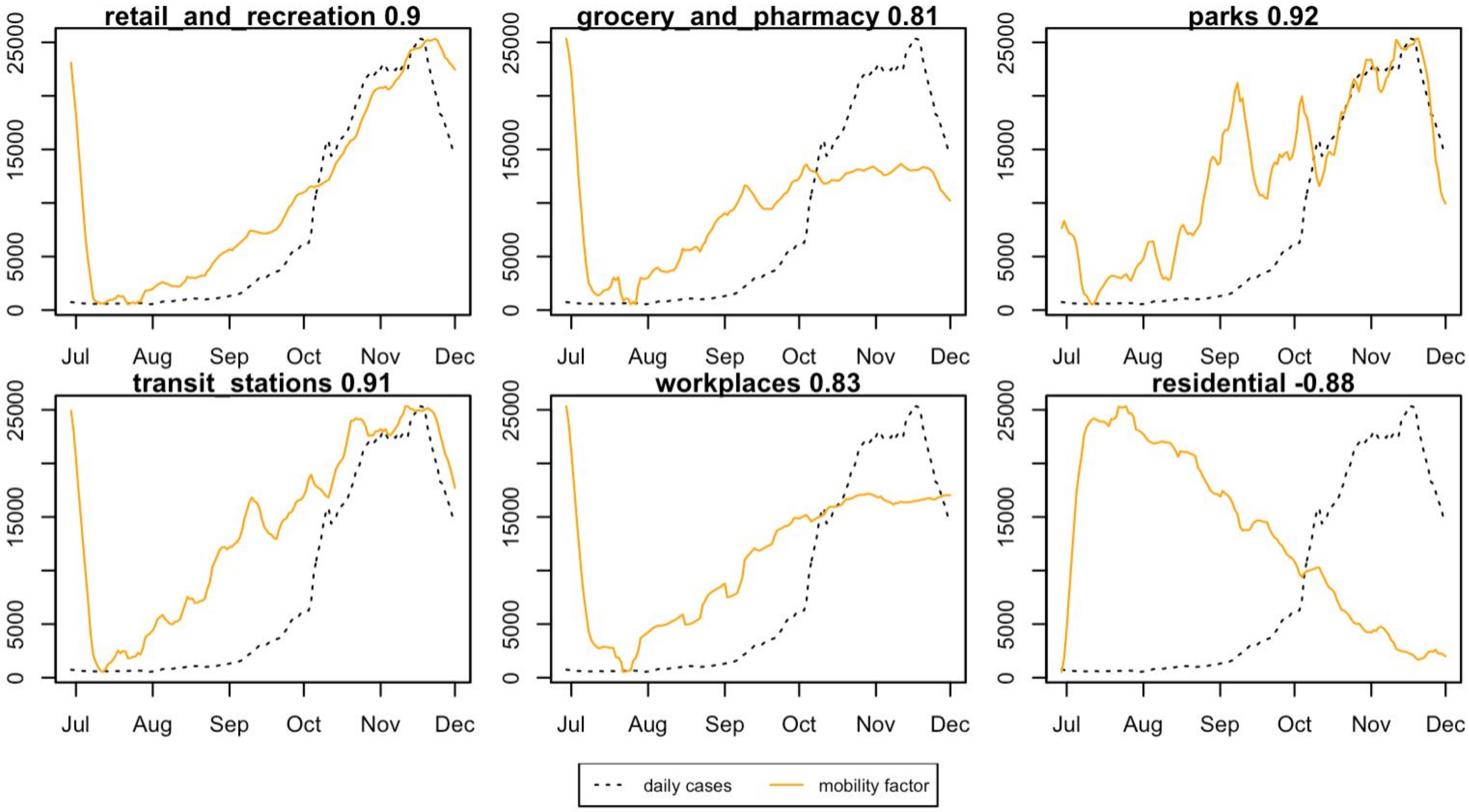
*Mobility data from *Google Cloud BigQuery*. Covid data from *ECDC*.

	Optimal Lag <int>	Spearman Cross-Correlation <dbl>
retail_and_recreation	100	0.8959089





Correlation between mobility and daily cases



*Mobility data from *Google Cloud BigQuery*. Covid data from *ECDC*.

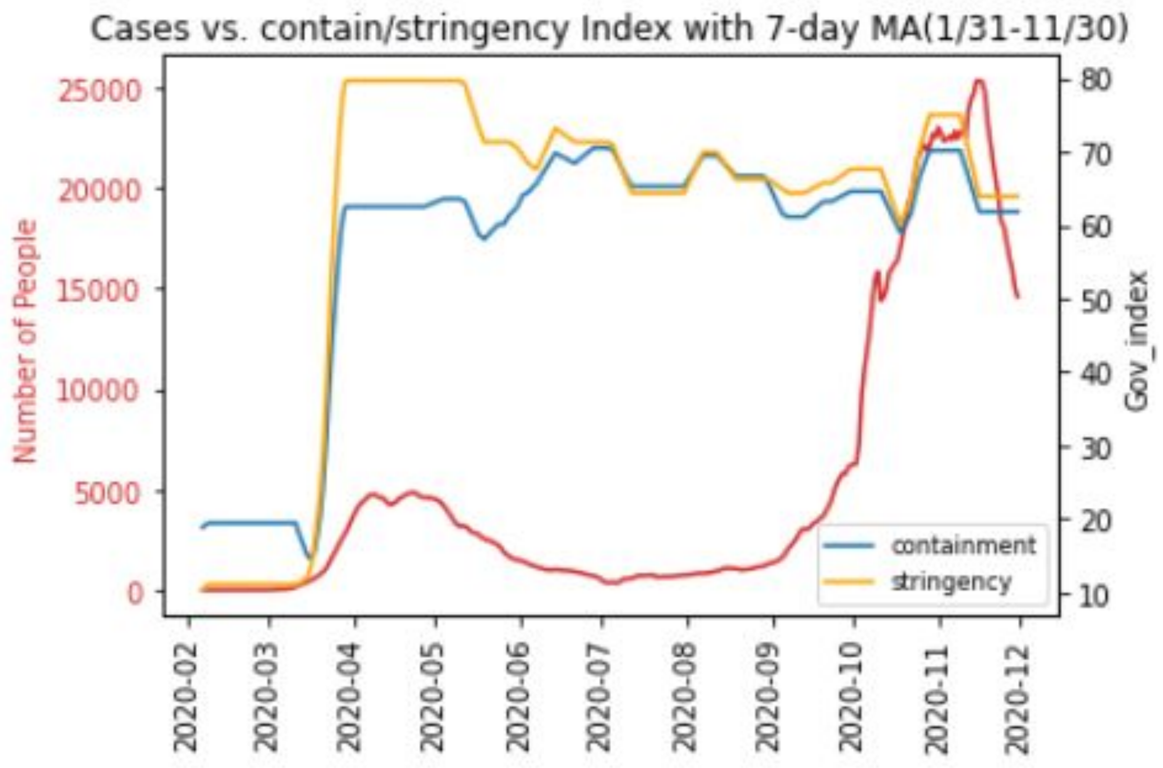
Leading Indicators - Gov. Policy Index

Leading Indicators - Gov. Policy Index

ID	Name	Type	Targeted/ General?
Containment and closure			
C1	School closing	Ordinal	Geographic
C2	Workplace closing	Ordinal	Geographic
C3	Cancel public events	Ordinal	Geographic
C4	Restrictions on gathering size	Ordinal	Geographic
C5	Close public transport	Ordinal	Geographic
C6	Stay at home requirements	Ordinal	Geographic
C7	Restrictions on internal movement	Ordinal	Geographic
C8	Restrictions on international travel	Ordinal	No
Economic response			
E1	income support	Ordinal	Sectoral
E2	debt/contract relief for households	Ordinal	No
E3	fiscal measures	Numeric	No
E4	giving international support	Numeric	No
Health systems			
H1	Public information campaign	Ordinal	Geographic
H2	Testing policy	Ordinal	No
H3	Contact tracing	Ordinal	No
H4	Emergency investment in healthcare	Numeric	No
H5	Investment in Covid-19 vaccines	Numeric	No
H6	Facial coverings	Ordinal	Yes
Miscellaneous			
M1	Other responses	Text	No

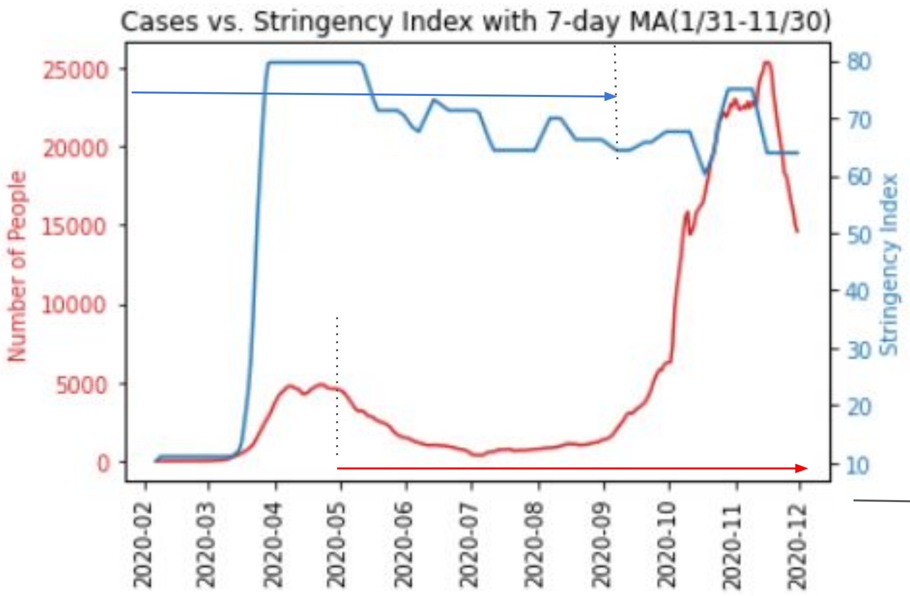
Index Name	C1	C2	C3	C4	C5	C6	C7	C8	E1	E2	H1	H2	H3	H6
Government Response Index	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Containment and Health Index	x	x	x	x	x	x	x	x			x	x	x	x
Stringency Index	x	x	x	x	x	x	x	x			x			
Economic Support Index									x	x				

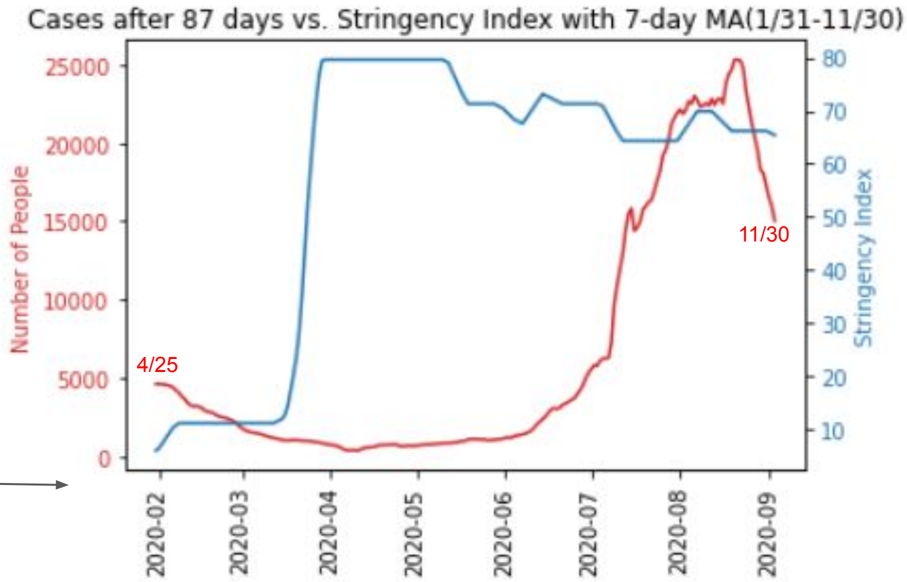
Source: Oxford COVID-19 Government Response Tracker (OxCGRT)



Stringency = C1~C8 + H1
Containment = Stringency + H2+H3+H6

Potential Lag effect:
Overall decreasing trend in indexes (June - Oct.) ---> increase in daily cases (Aug. - Nov.)
Overall increasing trend in indexes (Oct. - Nov.) ---> decrease in daily cases (Nov - now)



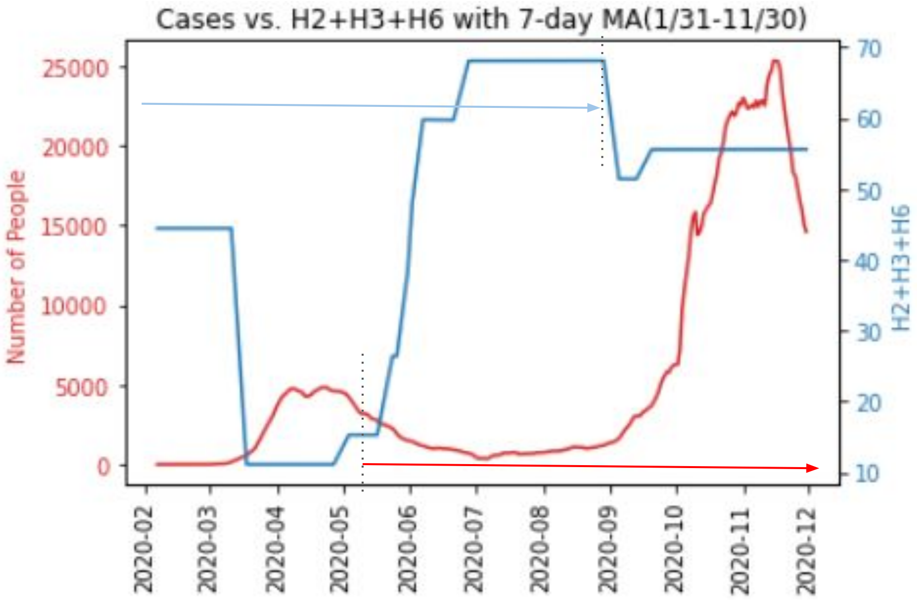


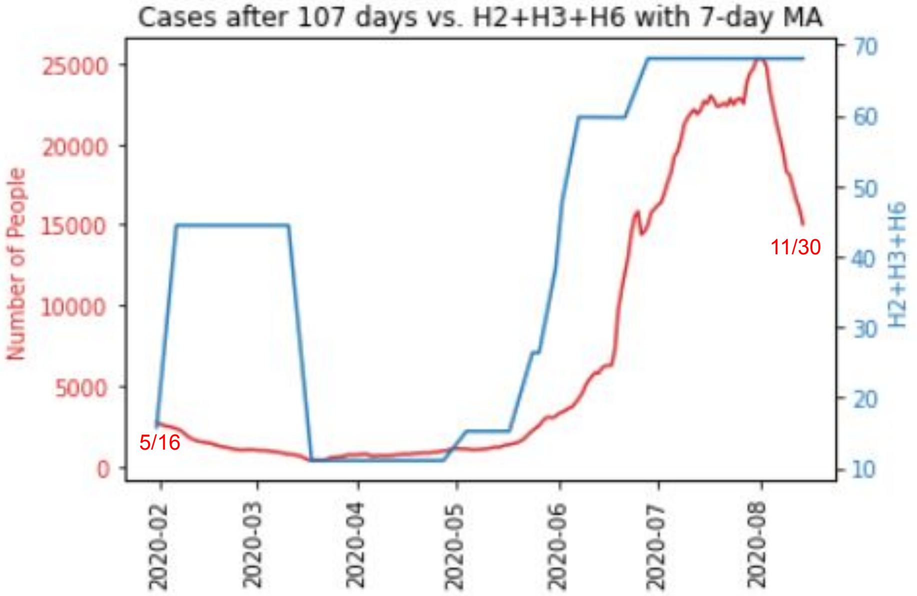

```
best lag:  87
window:  7
spearman corr:  0.613081328509765
```

0-lag correlation: 0.45

Observation: cases after 87 days are corresponding better to the changes (reversely) in stringency index now.

H2: testing policy
H3: contact tracing
H6: facial coverings



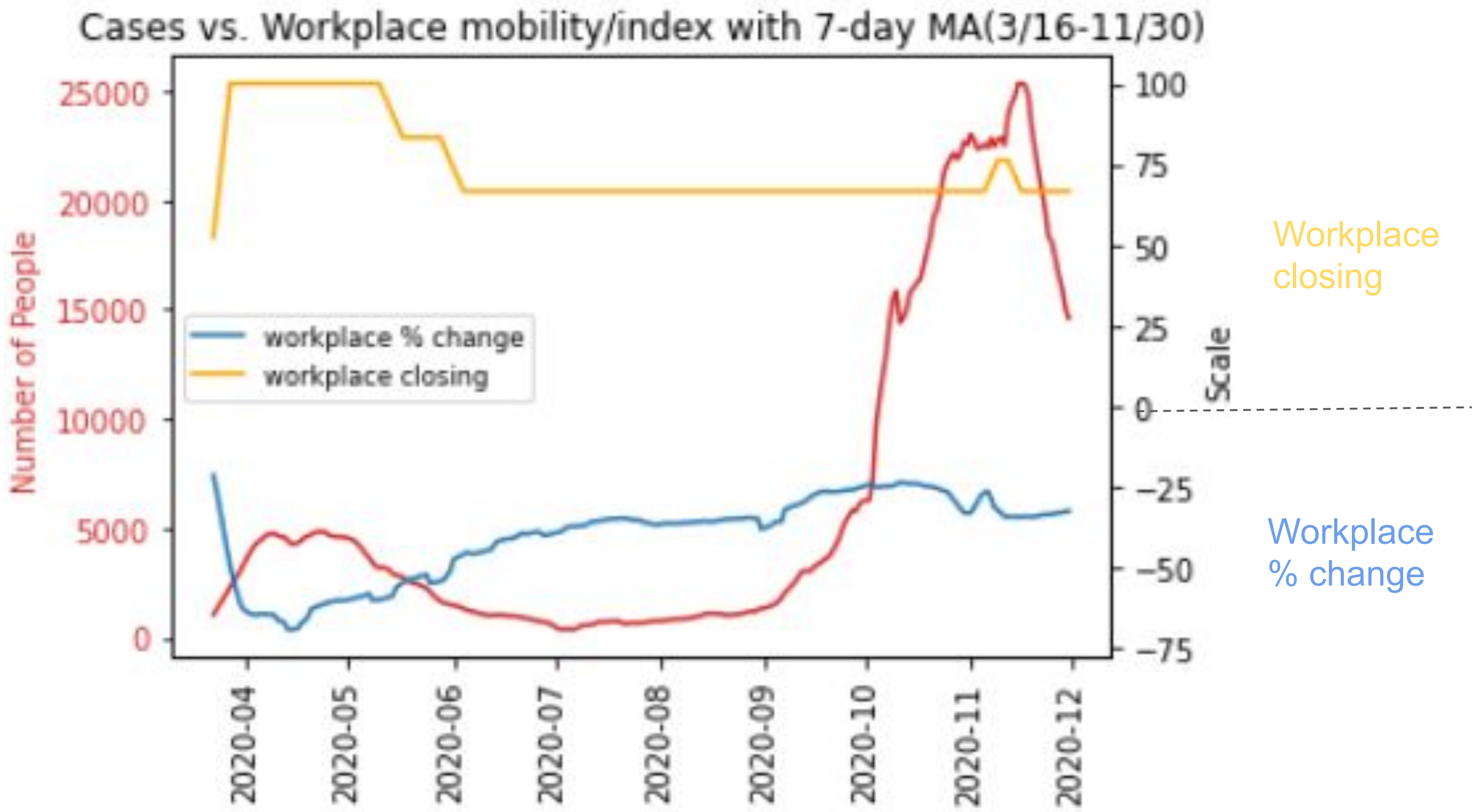


0-lag Correlation: -0.18

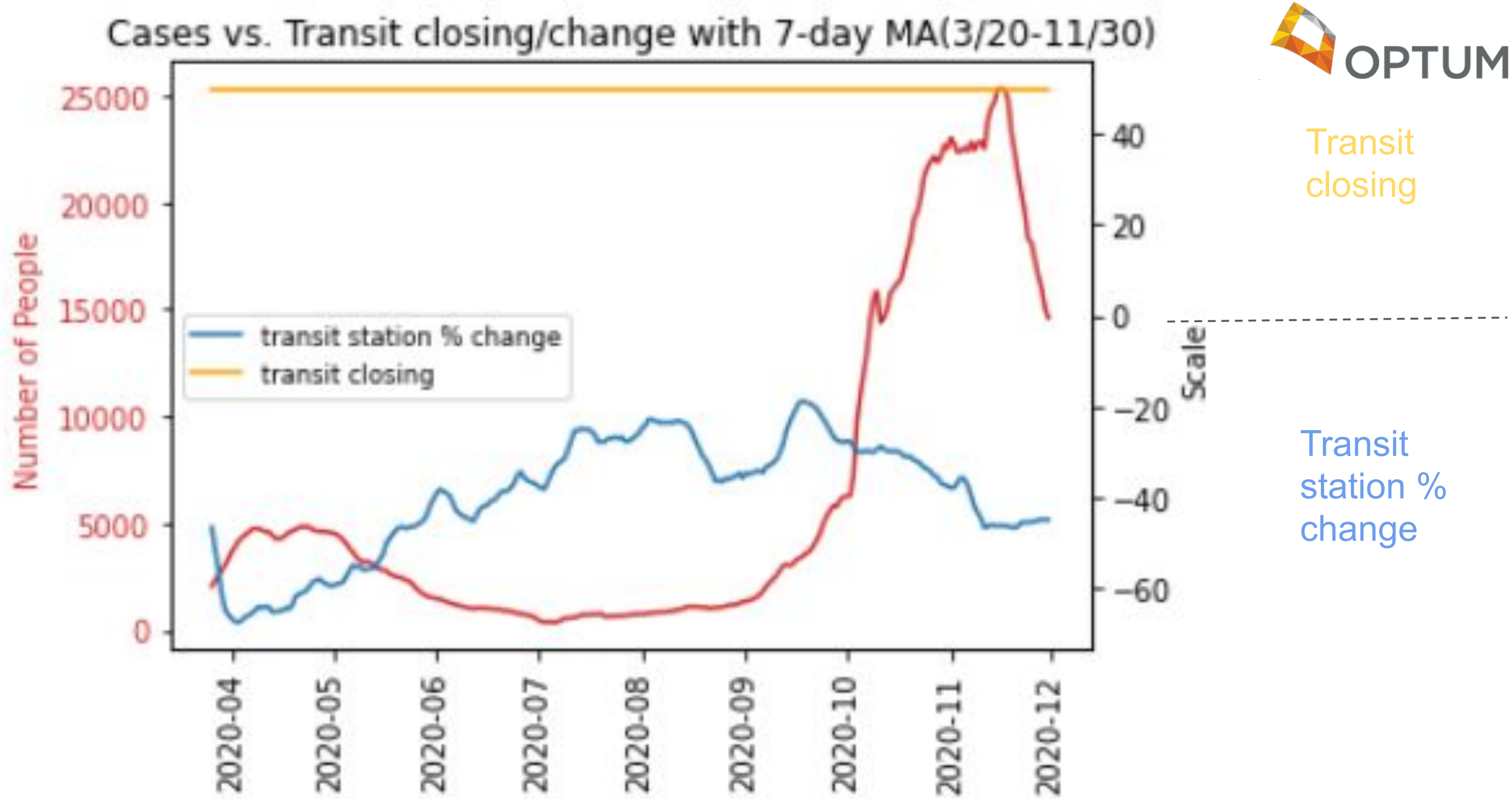
```
best lag:  107
window:  7
spearman corr:  -0.8727348723480409
```

Combined Leading Indicators

Gov. policy indicator	**Mobility**
Workplace closing	workplace % change
Transit closing	transit stations % change
Stay at home requirement	residential % change

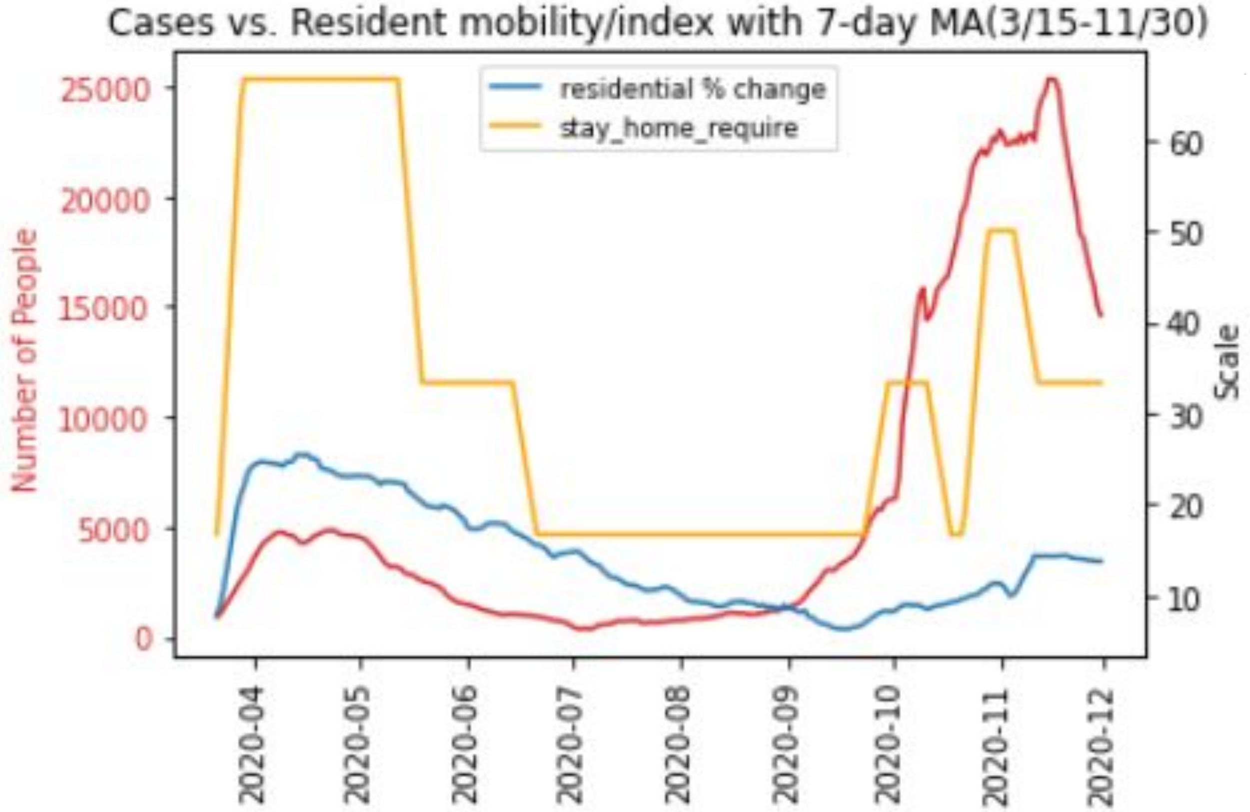



```
Correlation between workplace_closing and workplace % change:  -0.673173880562277
Correlation between workplace_closing and workplace % change (MA=7):  -0.7357487707371038
```



This is due to transit_closing index (C5) remains unchanged during the period

```
Correlation between transit_closing and transit stations % change:  nan
Correlation between transit_closing and transit stations % change (MA=7):  nan
```



OPTUM™

```
Correlation between stay_home_requirement and residential % change:  0.5994838285111471
Correlation between stay_home_requirement and residential % change (MA=7):  0.7319404429433023
```

ARIMA

ARIMA model

What is it?

- We can use trends in the past to predict Covid cases in the future

ARIMA = Auto Regressive Integrated Moving Average

- **Auto Regressive** - Fitting a regression based on several previous timesteps
- **Integrated -** Accounts for data not being stationary (constant mean and variance)
- **Moving Average -** Forecasting based on errors from previous time step predictions

Train/test

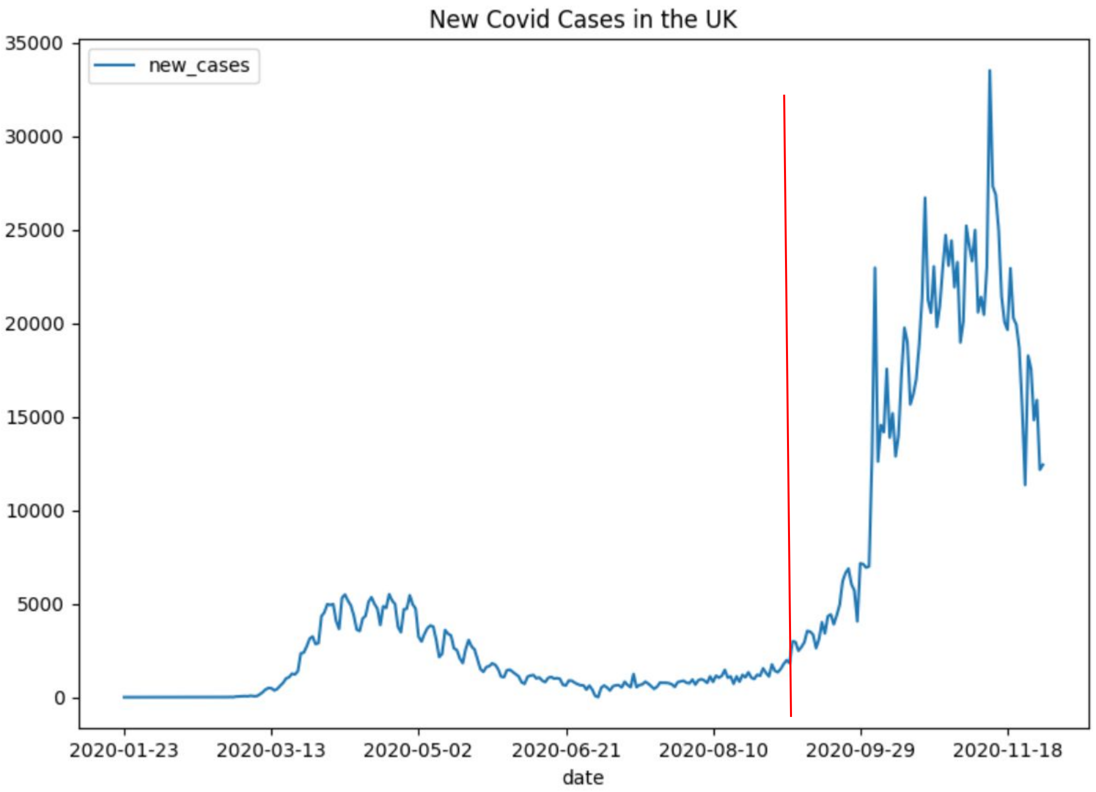
Train: 75%: 1/23 - 9/12

Test: 25%: 9/13 - 11/30

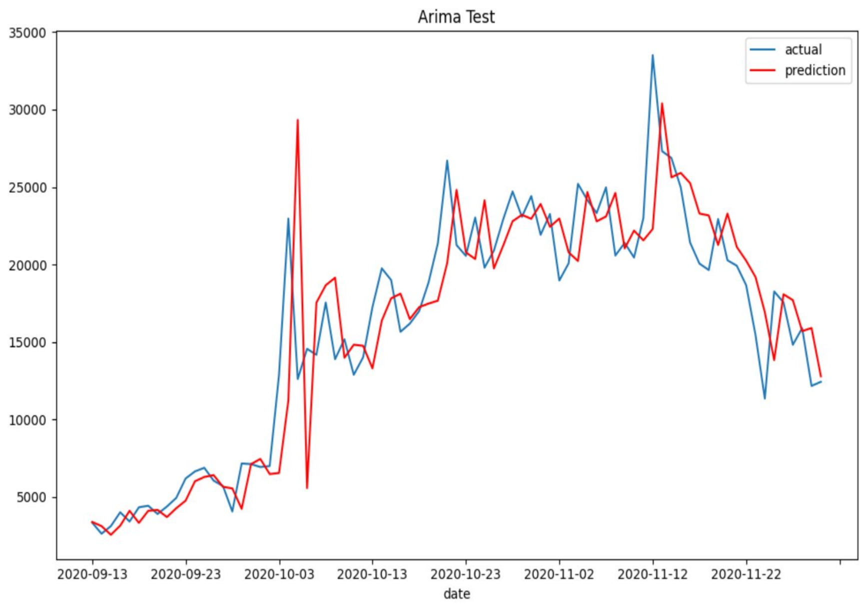
P (auto regressive) = 6

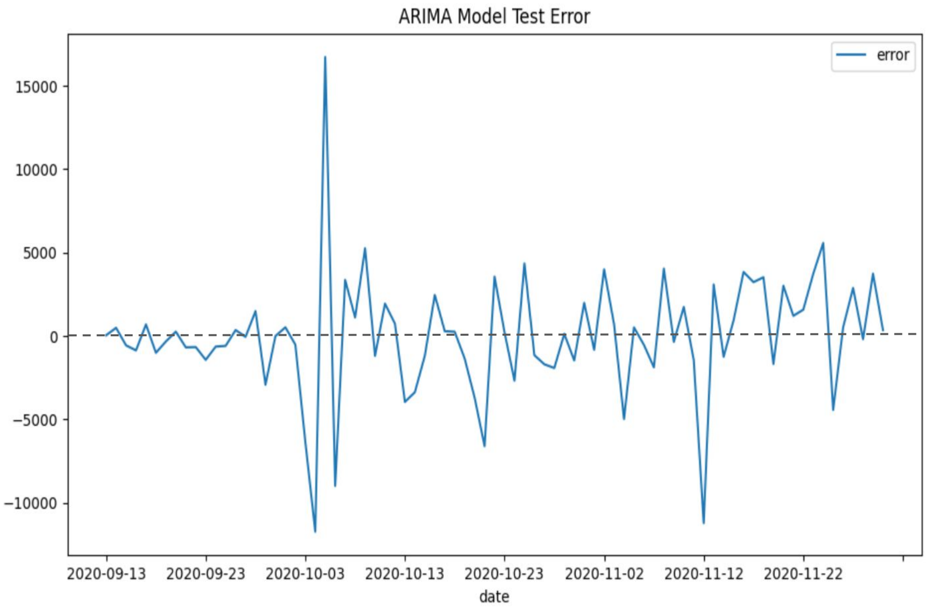
D (integrated) = 1

Q (MA window) = 6



Model Results

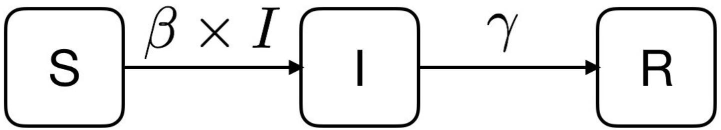


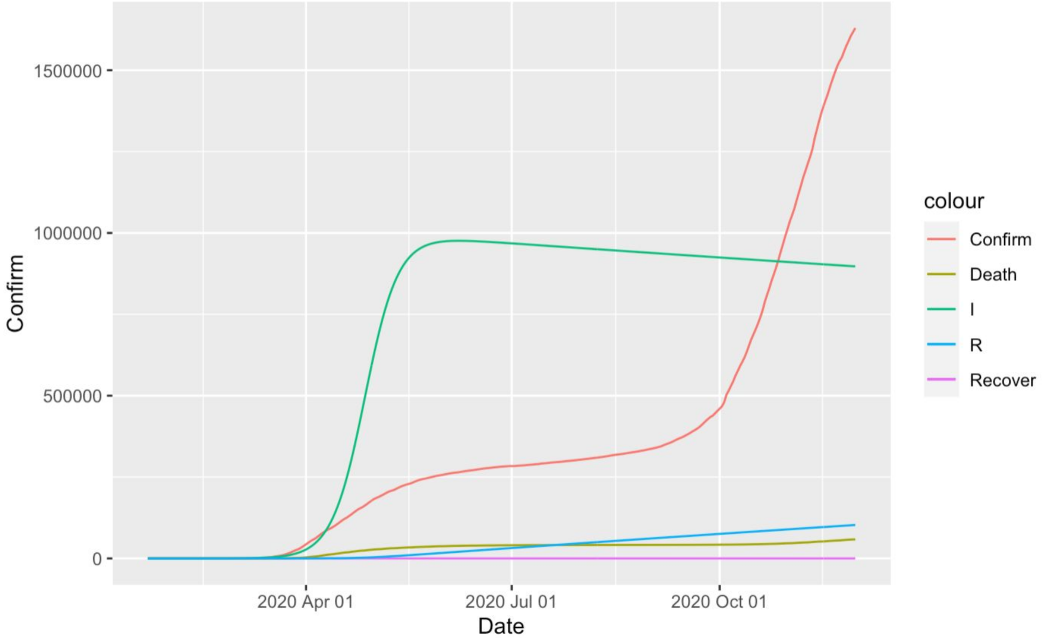


SIR and More

SIR Model

- Modelling vs Data Mining
- JHU Novel Covid Data: include confirm, death, recover data in country level
  - Fit parameters beta and gamma to minimize RMSE
  - Poor fit due to no recover data for UK after April 13th
  - Model Assumptions not hold: close country
- Next step:
  - Adjust the model
  - Data mining approach: Bi-LSTM





Citations: The SIR Model for Spread of Disease - The Differential Equation Model. (n.d.). Retrieved December 15, 2020, from https://www.maa.org/press/periodicals/loci/joma/the-sir-model-for-spread-of-disease-the-differential-equation-model

Future Improvements and Acknowledgements:

- Leading indicators: Mobility and Government Indexes (using MA7) are highly correlated with daily cases after according for a lag.
- ARIMA: Perform well in fitting the data.
- SIR: More complete data, and model adjustments are needed to improve performance.
- Other Models: Currently doing research, may try it in the future.

We cannot express enough thanks to the Optum Team and CMU faculties for their continued support and encouragement regarding this project: Danita Kiser, Paul Nielsen and Professor Rebecca. The completion of this project could not have been accomplished without the support of them.

Thank You!

Any Questions?