

TEAM 17 FINAL PROJECT

Jesse Choi, Darren Leong, Nicole Liu, Sungwoo Noh



TEAM MEMBERS

JESSE CHOI

Data Analyst

SUNGWOO NOH

Data Architect

DARREN LEONG

Project Manager

NICOLE LIU

Data Engineer

Recommendation

Goal: Find the most unrestrictive policies that keep the growth rate of deaths below 1% and the growth rate of new cases below 3% on a 30-day rolling average?

Our recommendation are policies:

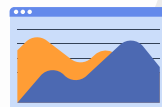
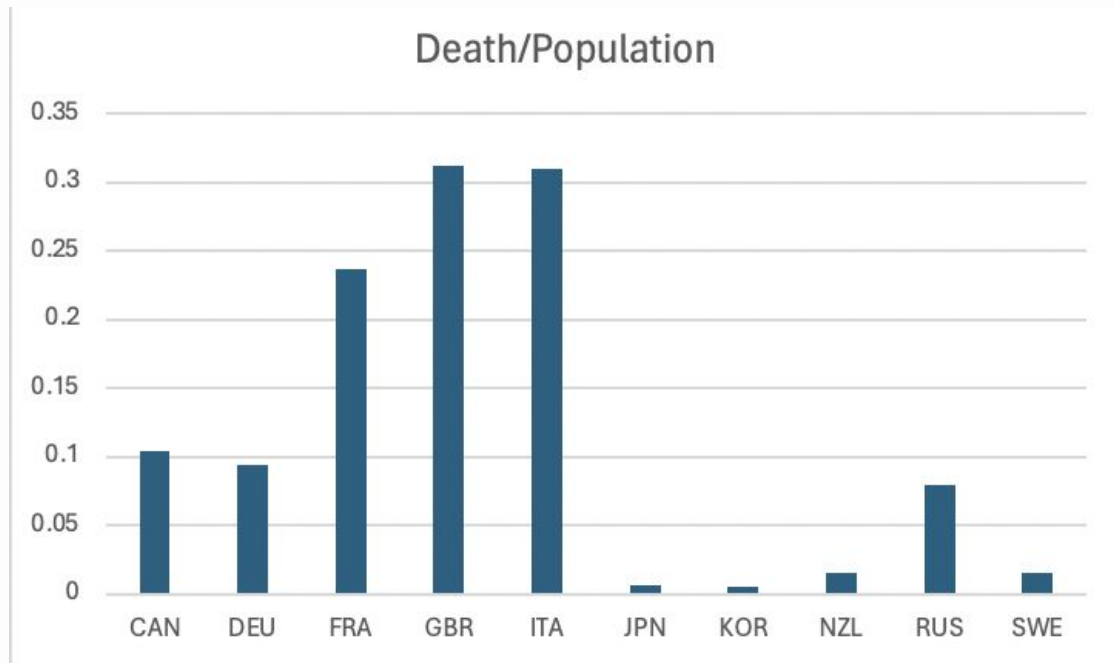
C2: Workplace Closing

C5: Closing Public Transport

C6: Stay at Home Requirements

C2, C5, C6 are the least restrictive policies that consistently keep growth rates of cases and deaths below 3% and 1% respectively, in spring (March–May).

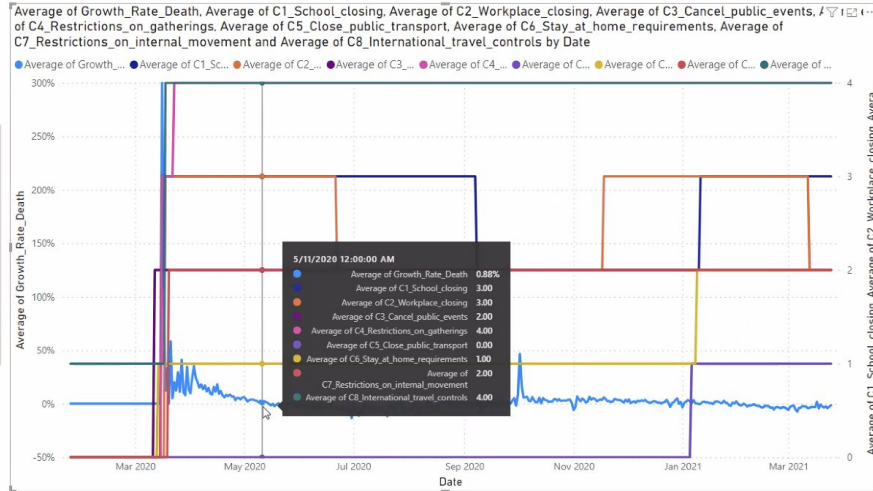
Death/Population Ratio



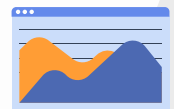
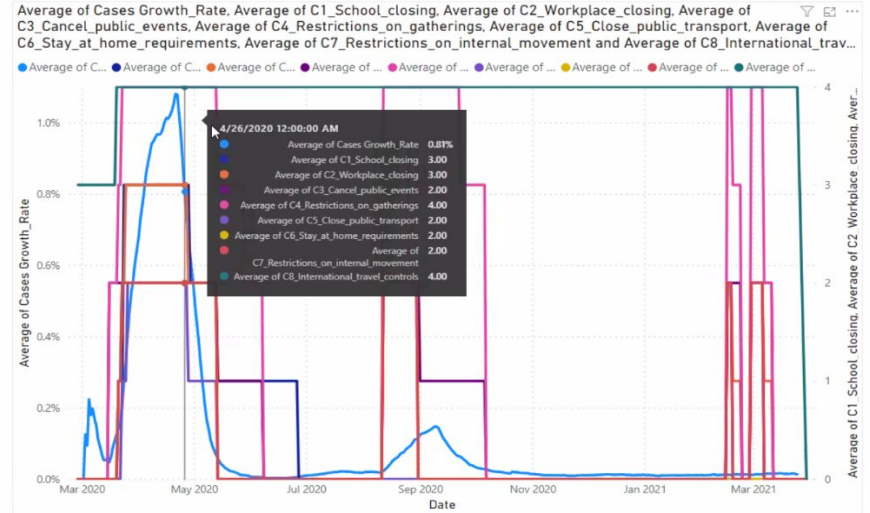
Growth Rate of Death/Policies



Canada: Growth Rate of Death with Policies v.s. Date



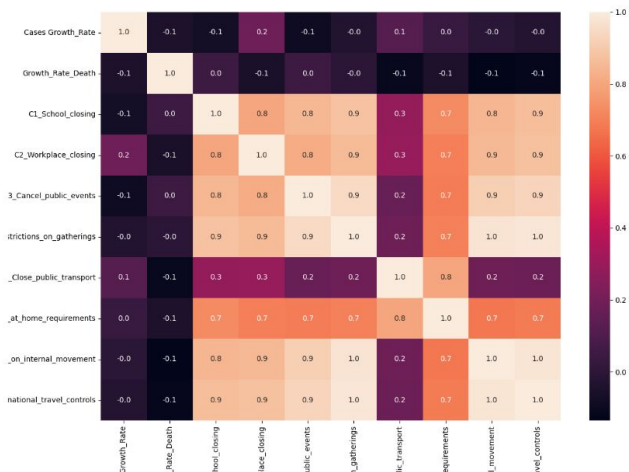
New Zealand: Growth Rate of Death with Policies v.s. Date



Correlation Matrices

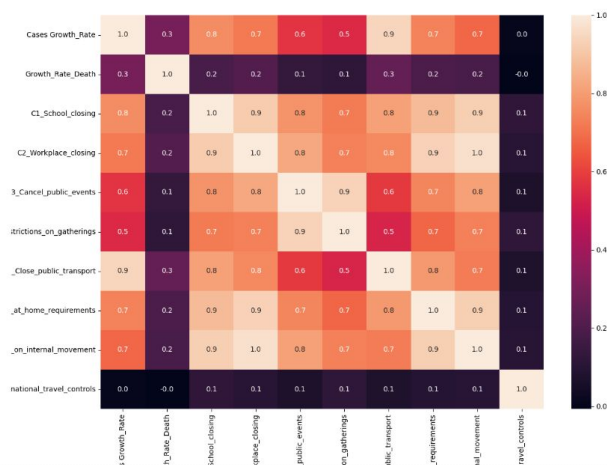
Canada:

Correlation Between Growth Rates of Cases and Deaths to Each Policy



New Zealand:

Correlation Between Growth Rates of Cases and Deaths to Each Policy



Very high correlations between independent variables...

Multiple Linear Regression for Canada



Cases Growth:

```
Regression Coefficients: [-0.03609658 0.00817035 -0.1535177 -0.1535177 -0.08547517 0.00817035  
-0.0064673 -0.03609658]  
explained_variance: 0.6861  
r2: 0.675  
MAE: 0.0343  
MSE: 0.0019  
RMSE: 0.0438
```

Deaths Growth

```
Regression Coefficients: [-0.02537604 -0.03448393 0.16666667 0.16666667 -0.21263234 -0.03448393  
0.00758746 -0.02537604]  
explained_variance: 0.5479  
r2: 0.5439  
MAE: 0.064  
MSE: 0.0114  
RMSE: 0.107
```

Multiple Linear Regression for New Zealand



Multiple Linear Regression for New Zealand

Cases Growth:

```
Regression Coefficients: [-6.28779710e-03 8.14685894e-02 -7.95026130e+12 7.95026130e+12  
-1.24752376e-01 8.42185913e-02 -9.60715828e-03 0.00000000e+00]  
explained_variance: 0.0553  
r2: 0.055  
MAE: 0.056  
MSE: 0.0068  
RMSE: 0.0824
```

Deaths Growth

```
Regression Coefficients: [-1.77794009e-02 4.53013140e-02 2.96062551e+11 -2.96062551e+11  
1.13035230e-02 -1.35137377e-04 1.75469088e-02 0.00000000e+00]
```



Recommendation

Goal: Find the most unrestrictive policies that keep the growth rate of deaths below 1% and the growth rate of new cases below 3% on a 30-day rolling average?

Our recommendation are policies:

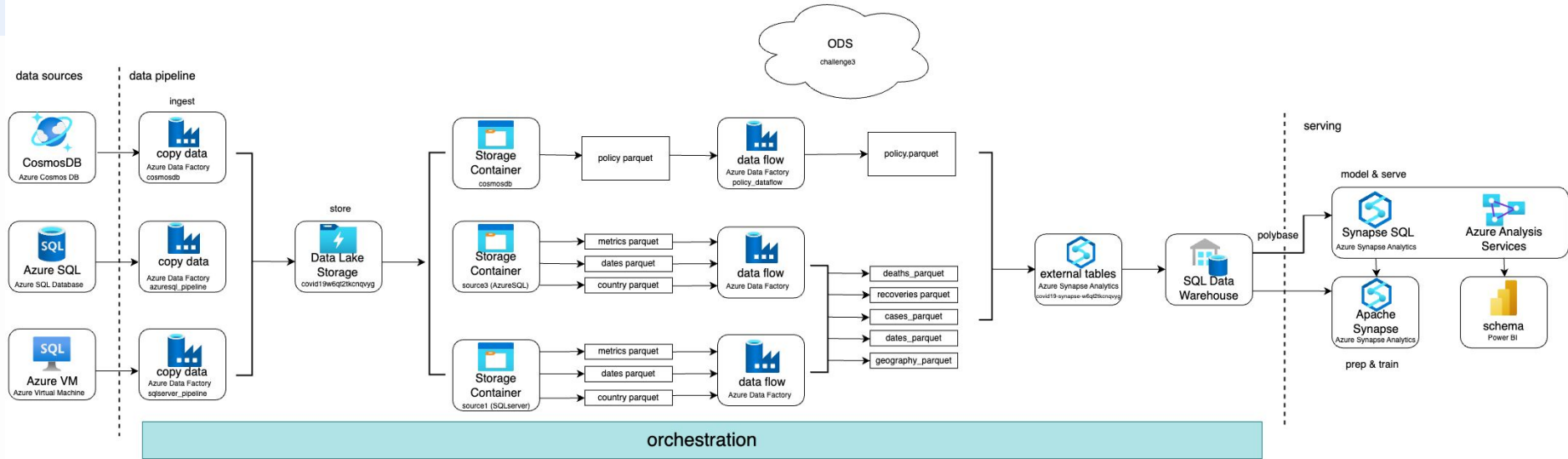
C2: Workplace Closing

C5: Closing Public Transport

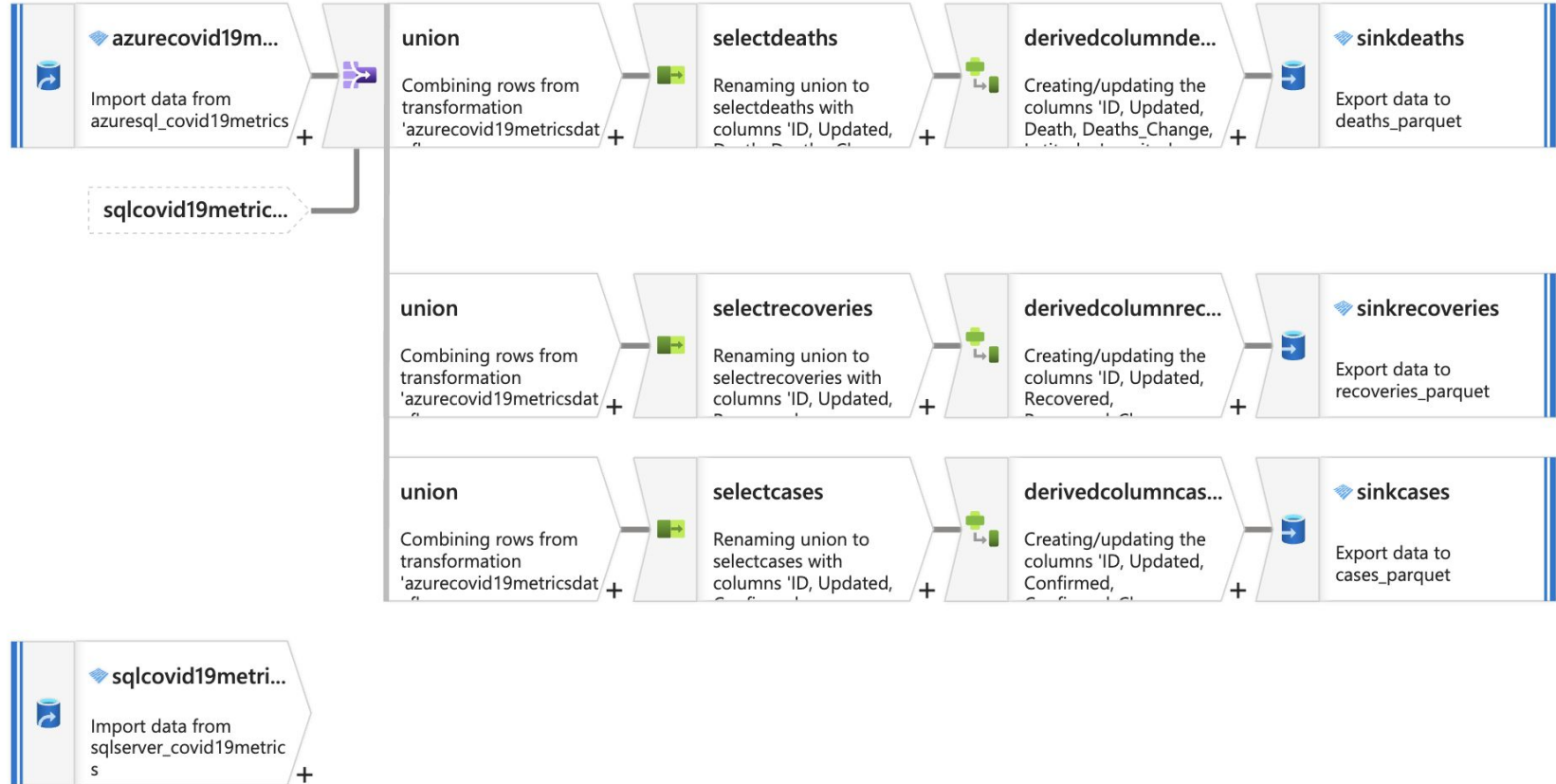
C6: Stay at Home Requirements

C2, C5, C6 are the least restrictive policies that consistently keep growth rates of cases and deaths below 3% and 1% respectively, in spring (March–May).

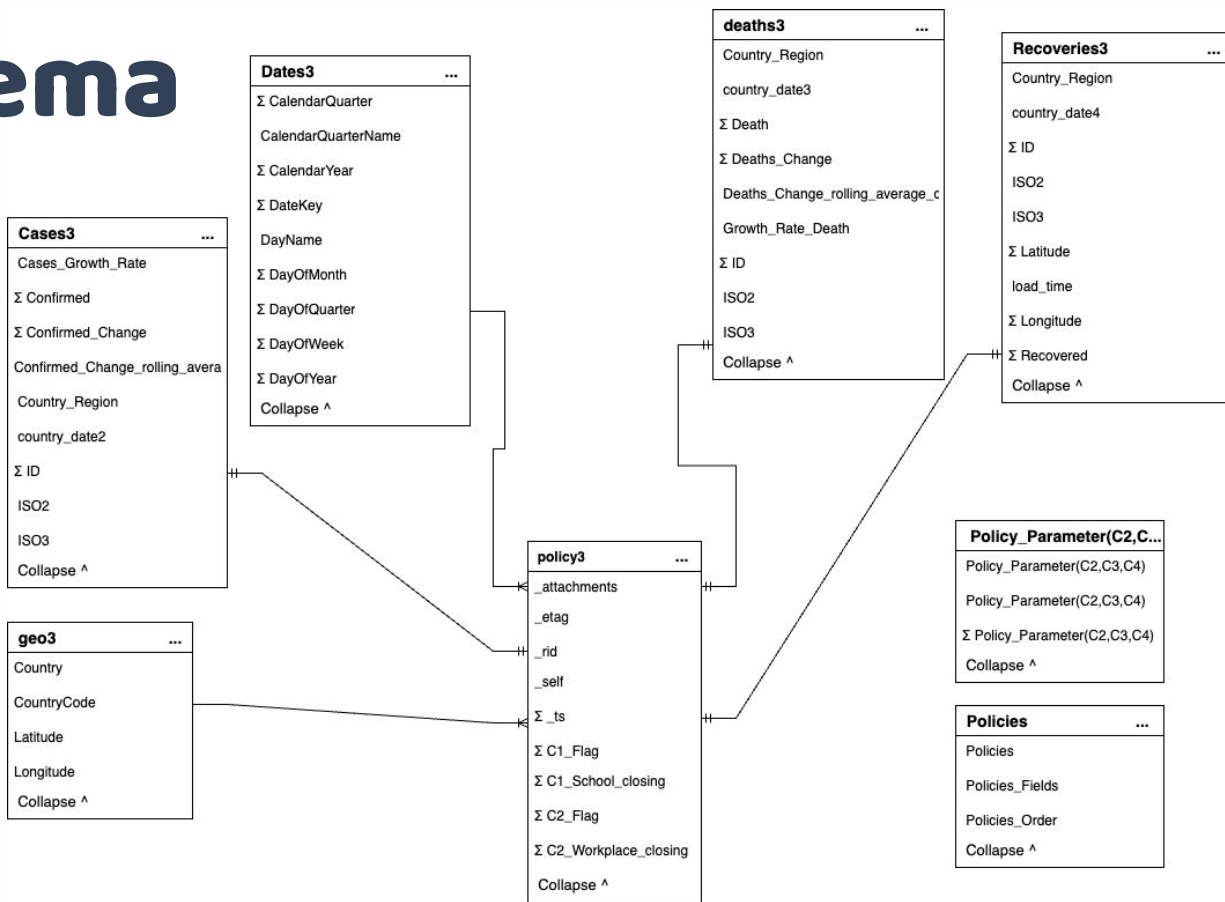
architecture



data flow



schema



limitations

- none countries have a similar population size to Caladan (3.2 mil)
- some countries had more restrictions than others in winter so death rate decreased in spring
- Policies in Caladan may not be directly applicable due to cultural, economic, or health infrastructure differences

future improvements

- figure out how to normalize all the data from the countries & make it comprehensive
- employing more advanced machine learning models could enhance the predictiveness of the project



thanks! :D