A Data-Based Approach to Flattening the Curve

BY: Oluwaseun Ademiloye, Tye Robison, Paul Lee, Vijay Fisch

MEET THE TEAM

Paul Lee

Tye Robison

Vijay Fisch

Oluwaseun Ademiloye

The Question

What are the most unrestrictive policies Caladan can implement to keep the growth rate of deaths below 1% and the growth rate of new cases below 3% on a *30-day rolling average?

Our Answer

The most unrestrictive policies that fit the constraints and we recommend to be used in Caladan are H6_facial_coverings and H7_vaccination_policy.

TABLE OF CONTENTS

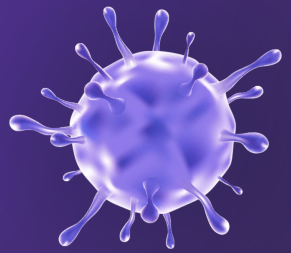

01
Initial Steps

What was the issue?

Objective: Mitigate and prevent the next wave of COVID-19 in Caladan - a midsize commonwealth with a population of 3.2 million.

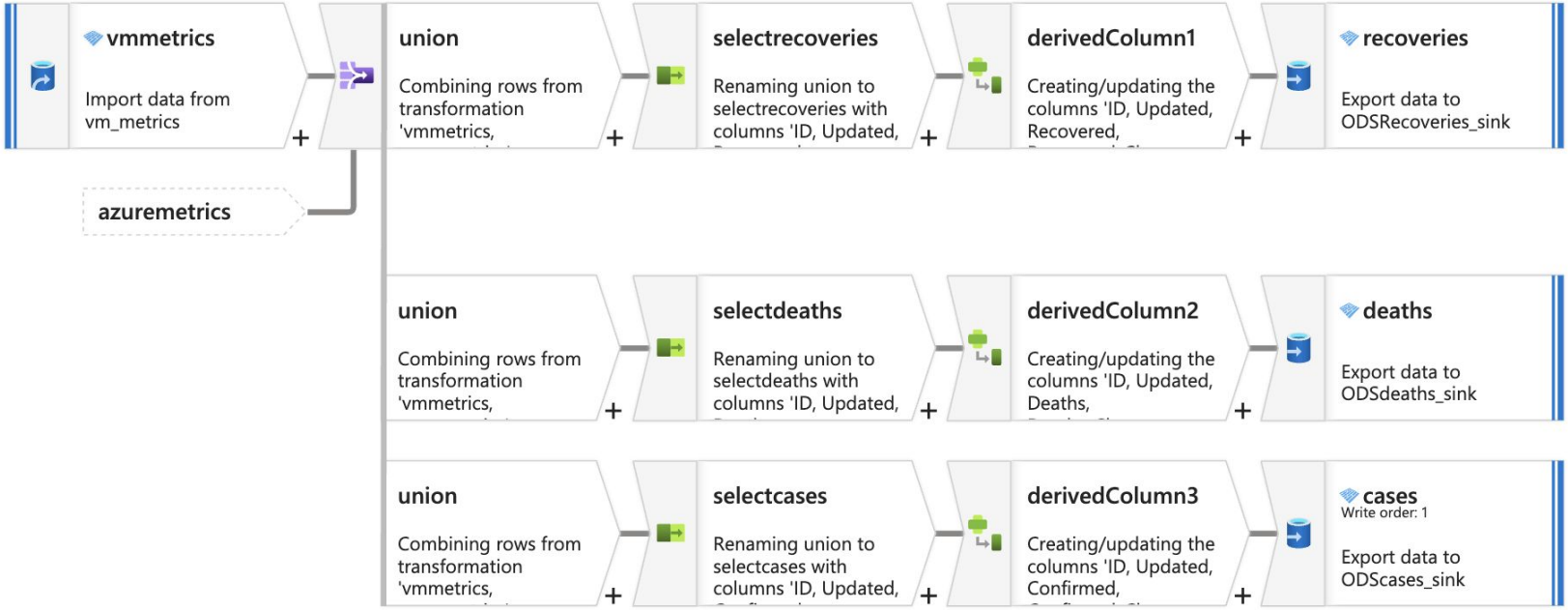
Approach and Strategy:

- Evaluate efficacy of diverse COVID-19 policies from 10 countries.
  - United Kingdom, France, New Zealand, Russia, South Korea, Sweden, Canada, Japan, Italy, and Germany
- Aim to create unrestrictive policies while keeping death growth rate below 1% and new cases growth rate below 3% on a 30-day rolling average
- Analyze correlation between growth and policies implemented by various governments and Establish a functional Data Pipeline.

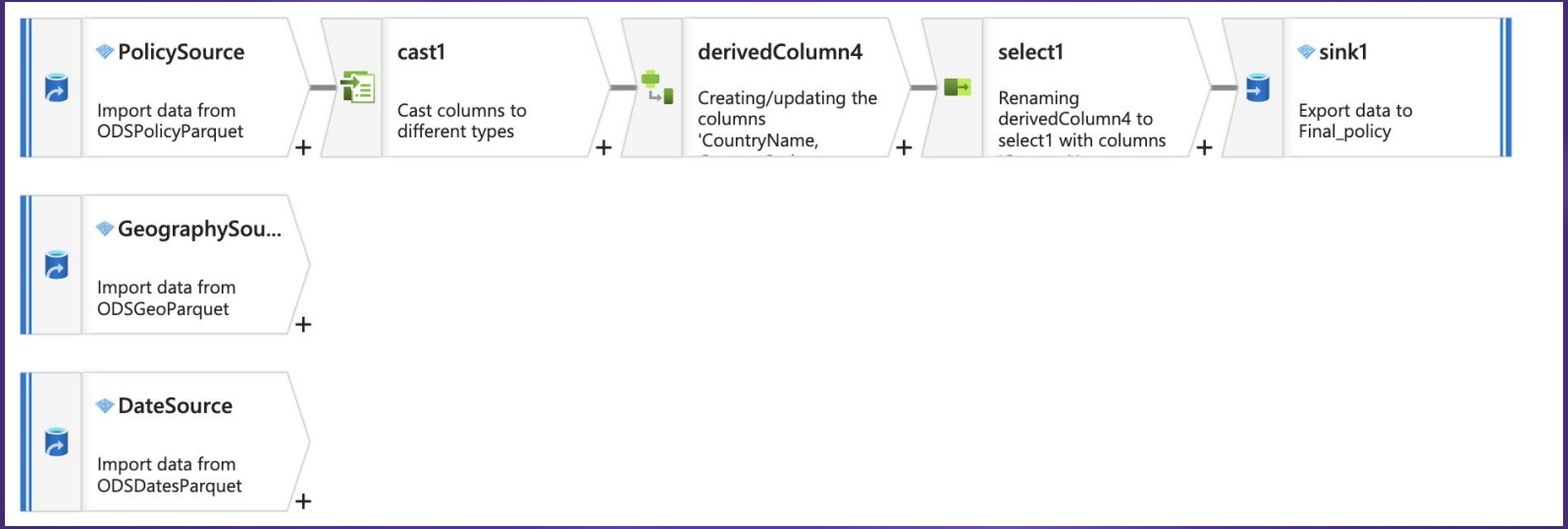



02
Data Flow

Data Flow Pt.1



Data Flow Pt.2



THE OUTPUT

Deaths.parquet

Deleted columns: (country_region, ISO2, latitude, longitude, load time)

Created derived column “UniqueID” to connect on Date & CountryCode

Recoveries.parquet

Deleted columns: (country_region, ISO2, latitude, longitude, load time)

Created derived column “UniqueID” to connect on Date & CountryCode

Cases.parquet

Deleted columns: (country_region, ISO2, latitude, longitude, load time)

Created derived column “UniqueID” to connect on Date & CountryCode

Policy.parquet

Deleted Null Columns: Region name, Region code

Created derived column “UniqueID” to connect on Date & CountryCode

Dates.parquet

Kept all columns

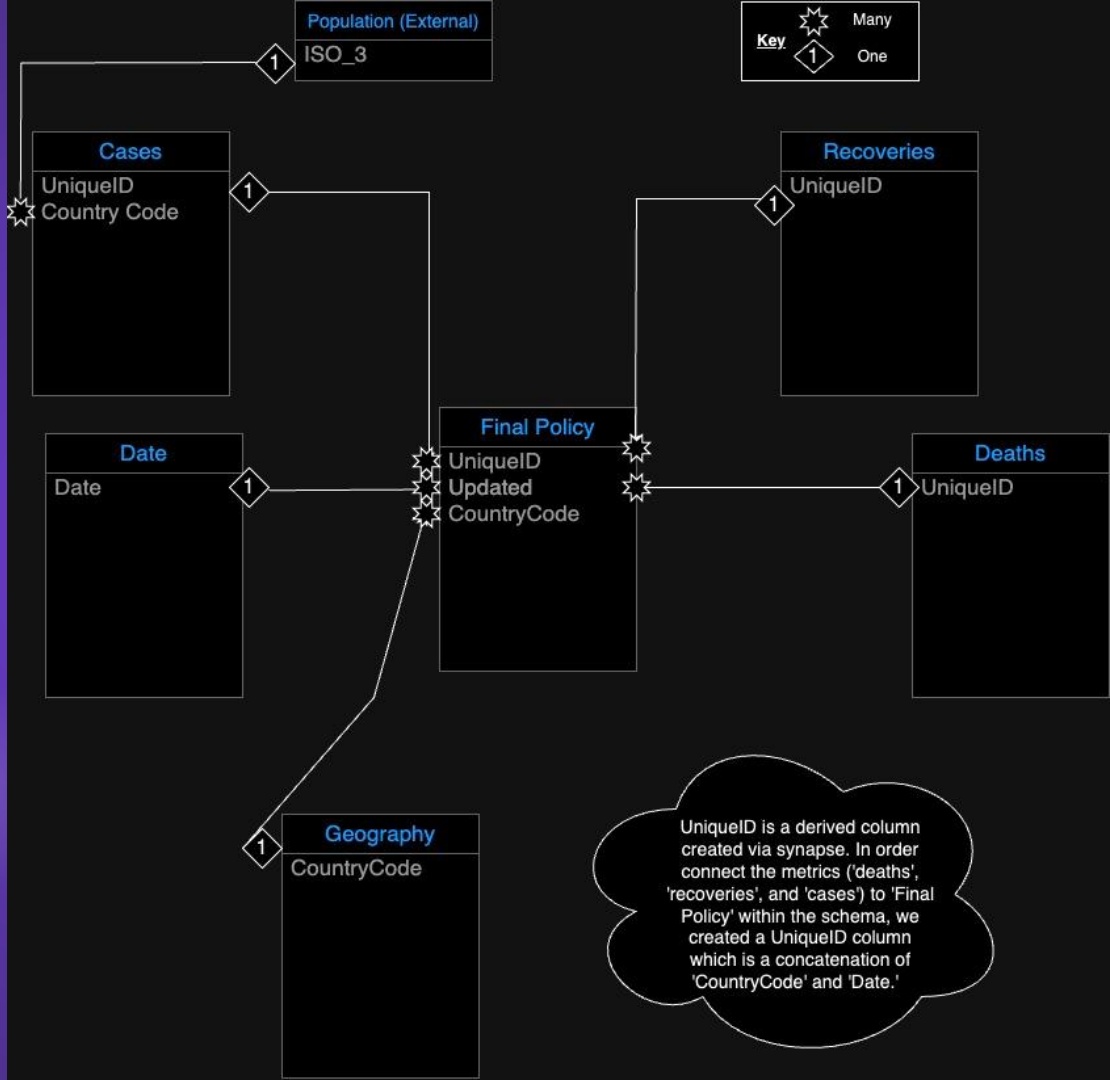
Geography.parquet

Kept all columns

Creating One-to-Many relationships & removed irrelevant/redundant columns


03
Schema Design

THE SCHEMA

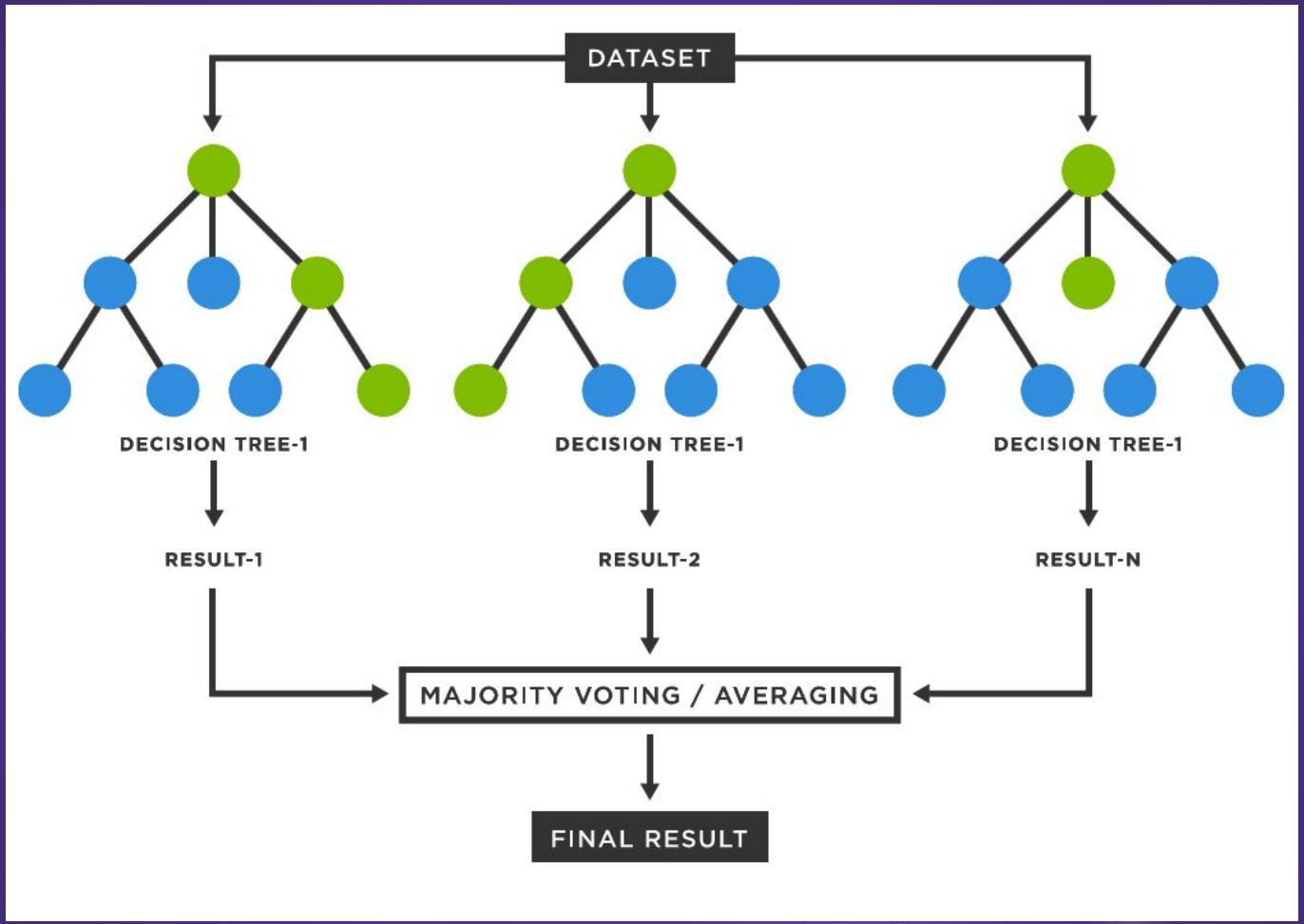

Population (External)
ISO_3
Key
Many
One
Cases
UniqueID
Country Code
Recoveries
UniqueID
Final Policy
UniqueID
Updated
CountryCode
Date
Date
Deaths
UniqueID
Geography
CountryCode
UniqueID is a derived column created via synapse. In order connect the metrics ('deaths', 'recoveries', and 'cases') to 'Final Policy' within the schema, we created a UniqueID column which is a concatenation of 'CountryCode' and 'Date.'


04
Power BI

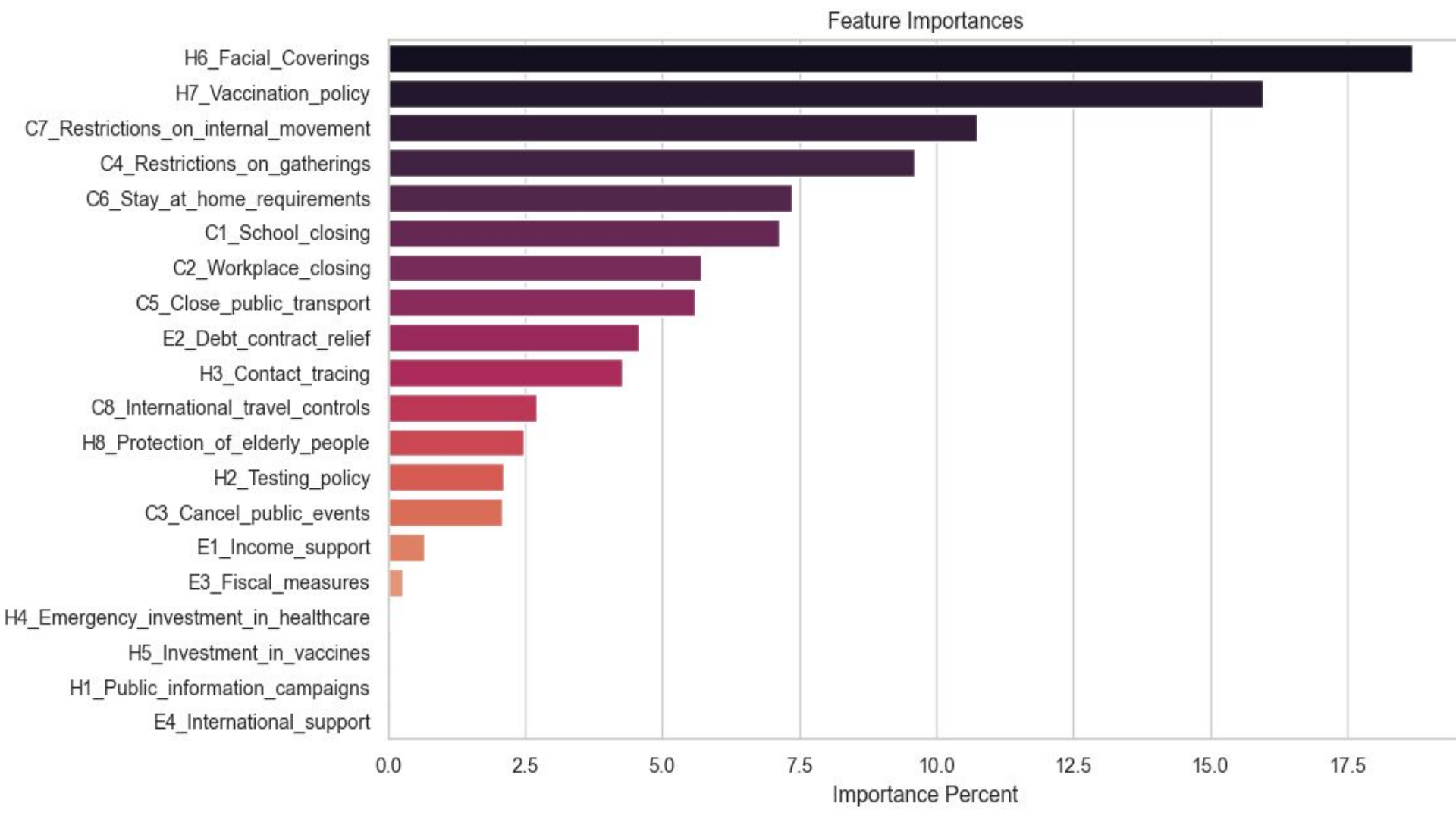

05
Machine Learning
& External Data

MACHINE LEARNING:
RANDOM FOREST REGRESSOR

1) Connected Cases and Policy on UniqueID, eliminated duplicate rows
2) Imported world population data from United Nations, divided confirmed by population to make it proportional to population
3) Isolated columns needed for machine learning analysis (variables are policies, label/target is confirmed_per_capita)
4) Train, test, split (test size 80%)
5) Fit RandomForestRegressor on training data with target as label data, hyperparameter tuning
6) Test it on test data, calculating regressor score (~68%)
7) Run feature importances, graphed using SeaBorn



Source: “Random Forest” by Deniz Gunay on Medium


Feature Importances
H6_Facial_Coverings
H7_Vaccination_policy
C7_Restrictions_on_internal_movement
C4_Restrictions_on_gatherings
C6_Stay_at_home_requirements
C1_School_closing
C2_Workplace_closing
C5_Close_public_transport
E2_Debt_contract_relief
H3_Contact_tracing
C8_International_travel_controls
H8_Protection_of_elderly_people
H2_Testing_policy
C3_Cancel_public_events
E1_Income_support
E3_Fiscal_measures
H4_Emergency_investment_in_healthcare
H5_Investment_in_vaccines
H1_Public_information_campaigns
E4_International_support
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Importance Percent

06

Final Conclusions

CONCLUSIONS

Caladan's suite of optimal policies

- **H7_vaccination_policy (best overall)**
- **H6_facial_coverings**
- One additional policy depending on severity of cases:
  C6_Stay_at_home_requirements

THANKS!

Please give us a good grade!