Team 15

Johnny, Roshni, Chloe, and Claire

Executive Summary

Our project's primary goal is to recommend effective COVID-19 policies for Caladan, aiming to keep new case and death growth rates below 3% and 1%, respectively, on a 30-day rolling average. We utilize the Kimball model for data exploration, incorporating statistical analysis, external data sources such as global population demographics, and Power BI for visualization. For short-term impact, we propose hybrid work and travel restrictive policies. In the long term, we recommend mandating face covering and testing requirements and focusing on protecting elderly people. These measures aim to reduce transmission rates and ensure sustainable pandemic management while considering Caladan's unique context.

Our target metric of success was the growth rate of the 30-day rolling average of both cases and deaths. We compared policy data against the growth rate in three main ways to determine the most important policies. First, we looked at each policy's correlation coefficients between growth rates and policy restriction levels. By creating our own correlation metric and using the PowerBI key influencer tool, we could narrow down the policies with the strongest correlations (both positive and negative). Next, we graphed the trends of growth rates and the policy restrictions for all countries over time, allowing us to visualize patterns we may not have encountered before. We were able to restrict our visuals to specific time intervals and countries, which helped to identify differences in patterns.. Lastly, to translate policy effectiveness into a metric, we calculated the rate at which a country’s growth rates were below our target thresholds. This statistic helped select each policy's least restrictive level, while effectively maintaining growth rates below 3% and 1%. By creating correlation coefficients, graphical analysis of the growth rates against policies, and calculating the average amount of time a country met our thresholds, we were able to use these statistics to determine the least restrictive policies as a whole.

Our recommended policies to consistently enforce are face coverings (mandatory indoors), testing (mandatory for covid symptoms), and elderly protection (maximum elderly prioritization). We decided to recommend these policies because our key influencers graphic shows that the average growth rate for deaths goes down when these policy levels go up. We illustrated this for face covering by comparing Great Britain and Germany, both mid-sized European countries, in the time frame around June 2020 to March 2021. As Great Britain increased its face-covering restrictions from level 1 to level 3, the death growth rate decreased dramatically. On the other hand, Germany remained at a level 2 throughout this period, and its death growth rate peaks at 2.3%, nearly double that of Great Britain. As for testing policy, we compared the threshold rates between Sweden and France and found that only a level 2 restriction was needed. To illustrate the impact of elderly protection, we compared Canada and Russia in the same time frame, from June 2020 to March 2021. Canada has a protection level of 3 during this entire period, while Russia’s elderly protection level fluctuates anywhere from 0 to 3. Russia’s maximum growth rate for deaths was 2.48%, while Canada saw a much lower peak death growth rate of 0.86%. Overall, it is evident that these policies need to be maintained at higher restrictions in order to achieve our goal.

Our recommended policies to initially enforce are the closure of workplaces and the restriction of international travel. These two policies are driving factors of transmission since they involve lots of human interaction, but once the growth rates are controlled under the threshold, the policies can be relaxed. First, we suggest that workplace closure should be implemented at level 3 (enforce most workplace closures), and relaxed to level 1 (only recommend closures) once growth rates are under control. There is a strong positive correlation between workplace closure and growth rates. This suggests that workplace restrictions are very high initially, but as the growth rates go down, the restrictions also relax. Countries that use stronger restrictions initially see more success in bringing down their growth rates. For example, Russia and Canada are able to control their growth rates very quickly, using an initial level of 3. However, we see that Russia and Canada can also maintain their growth rates in the future despite using policy levels of 1 and 3, respectively. We see similar trends with international travel controls, where there is a strong positive, and countries with higher initial restrictions saw the most future success. For example, Russia and Sweden started with strict policy levels at 4 and 3, respectively, quickly lowering both growth rates. However, while both countries keep growth rates under control, Russia achieves this by relaxing its policies, whereas Sweden maintains its strict level. This shows that we are able to relax our policy as well, thus providing the justification to initially enforce at level 3 (ban most arrivals) and eventually enforce at level 2 (allow quarantine arrivals). Both workplace closures and international travel demonstrate that initial policies are very important, but they should be relaxed to create the least restrictive environment.

In conclusion, we were able to find 5 policies that Caladan should focus on enforcing. Through the analysis of the correlation coefficient, threshold rates, and graph trends, the least restrictive policies were calculated to optimally keep the growth rates of cases below 3% and the growth rates of deaths below 1%.

Project Breakdown:

Johnny:

Challenge 1: Extracted Cosmos DB data

Challenge 2: Created data flows for Policies, Date, and Geography

Challenge 3: Created schema design and fact tables

Challenge 4: Created statistics for analysis (rolling average, growth rates, corr. coefficient, threshold rates). Organized and created many visuals. Helped to write the essay and make slides.

Claire

Challenge 1: Created storage account in team resource group and a container for each source.

Challenge 2: Explored datasets.

Challenge 3: Helped with brainstorming schema ideas. Translated schema into the model view on Power BI.

Challenge 4: Created visuals for policies we selected, cleaned up titles and labels on Power BI. Helped write the essay. Exported ARM template. Configured our Data Factory with Github.

Roshni

Challenge 1: Configured SHIR access to help get the on-premise SQL server data.

Challenge 2: Explored datasets to find anomalies and some duplicates. Configured Github settings.

Challenge 3: Transferred the data into ODS, created external tables with Chloe, and helped brainstorm and draw out potential schema layouts.

Challenge 4: Designed exploratory visualizations that led to creating our recommendation strategy. Aided in creating some policy dashboards and cleaning them up. Created the slides and visualized our data flow.

Chloe

Challenge 1: Extract SQL Server and SQL Database data.

Challenge 2: Create data flows for Metrics, Date, and Countries and normalize as needed.

Challenge 3: Transfer the data into ODS and created external tables

Challenge 4: Looked for external data, fixed the duplicate dataset and the rolling average, and helped to write the essay.