

# CSC343 Project Presentation

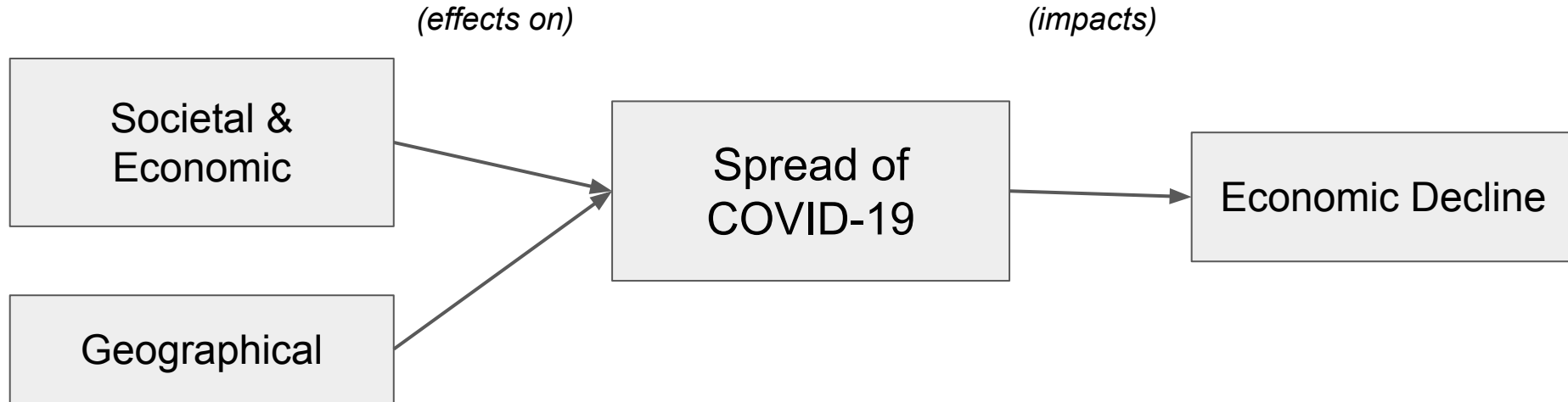
Group 0047

Yubo Gao (1005024493)

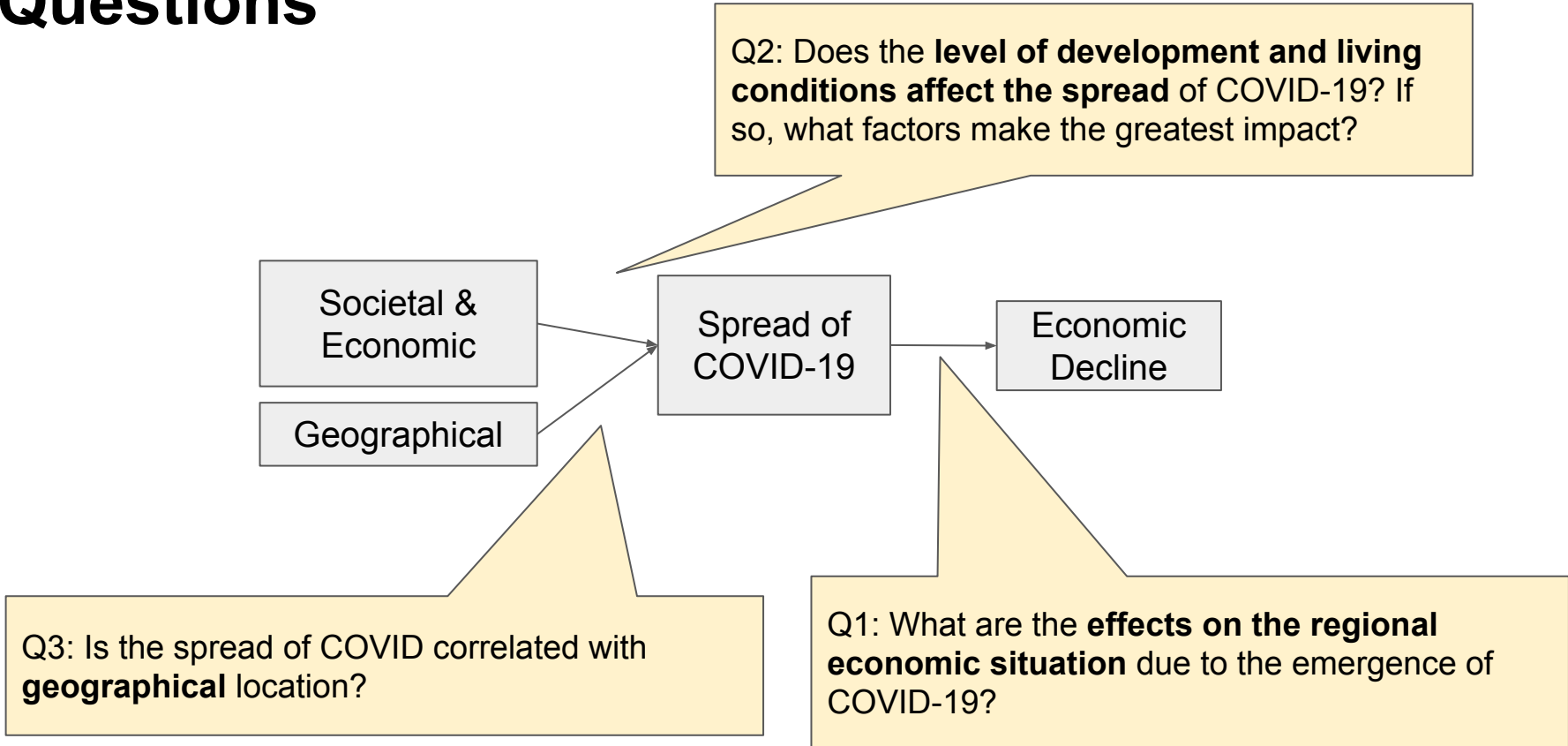
Tian Yu (1004750655)

# Domain

We investigate the factors affecting the spread of COVID-19 and its negative impacts on the economy.

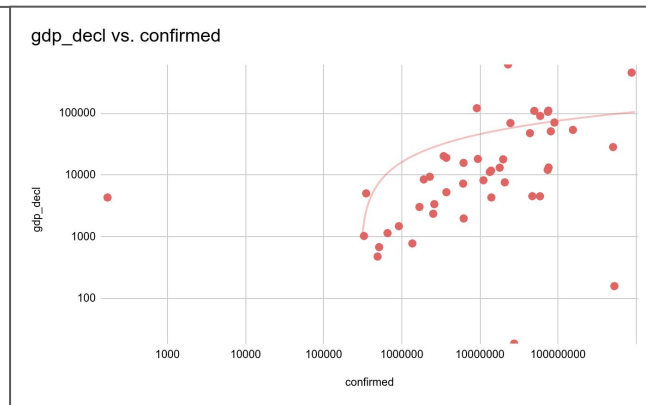
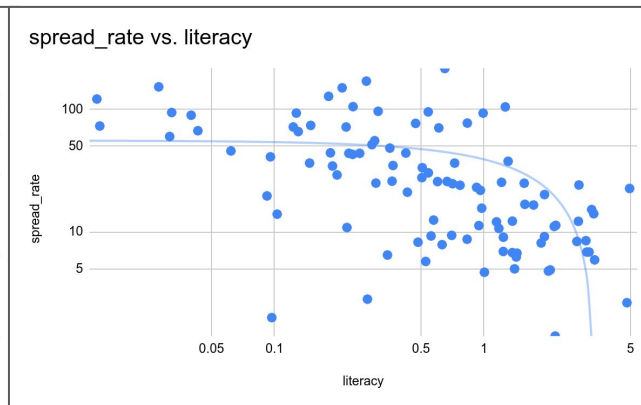
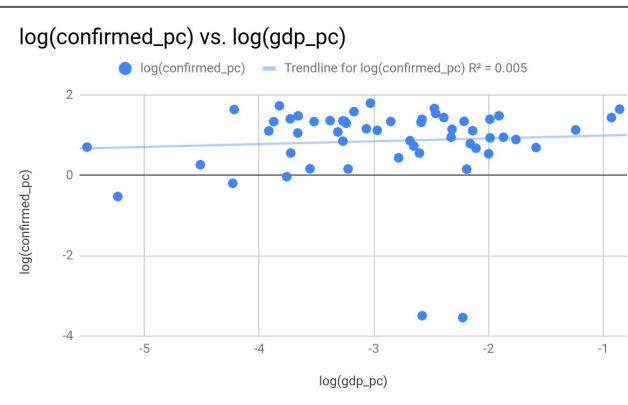


# Questions



# Major Findings

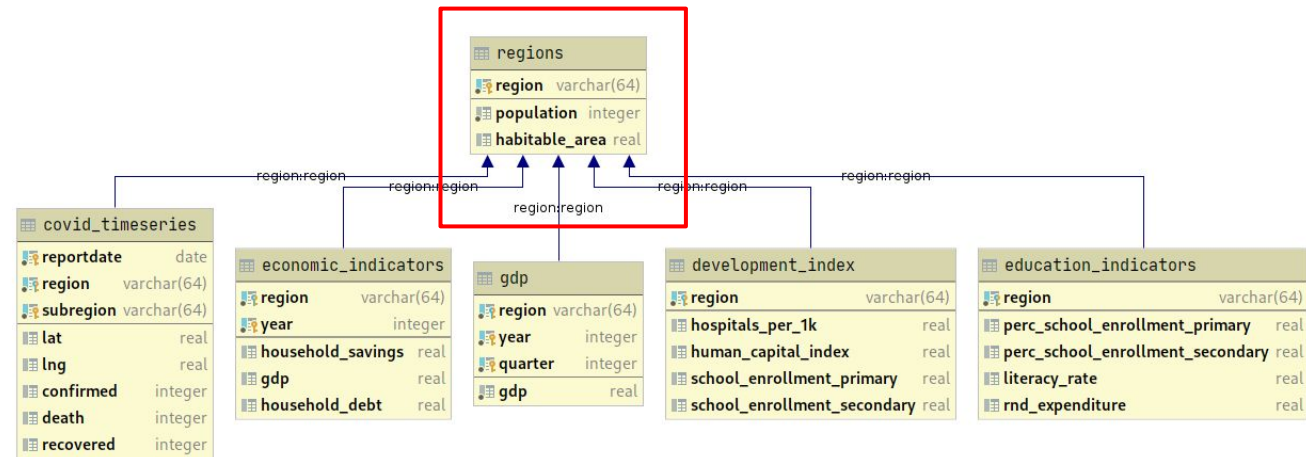
- Societal / educational factors have a much larger impact on preventing spread than economic ones (Queries 1b, 1c).
- Significant decline in economy proportional to scale of confirmed cases.



# Evolution

## Added “regions” table:

- to reconcile inconsistent and incomplete region names across different data sources
- restructured to include static regional information such as population.



# Lessons

Use standardized codes instead of names whenever possible:

- We faced consistency issues with region names.
- Use ISO 3166 country codes (CAN, USA, FRA, etc.) instead.
- Increases efficiency by decreasing amount of redundant data in the DB.

Use estimations in place of missing data:

- We face data sparsity issues in most of the downloaded datasets.
- Better to estimate during analysis rather than during data import (integrity)
- We used the difference of quarters (either two or three) to estimate the spread rate

**Questions?**