Note: What follows below is a somewhat simplified “prototype” of what a group write-up might look like. This is stylized and intentionally vague in parts, but the goal of this template is to give you a sense of what a good write-up might look like in terms of organization. Start with the big picture, zero in quickly on the key concepts, data, etc… that you will be using. Provide a very brief and high-level overview of your data processing, without getting into the nitty gritty details.

**--**

**Looking at COVID-19 Data**

[group member list]

What factors made the COVID-19 pandemic worse in some states than in others? The COVID-19 pandemic has emerged in the last year as the defining public health crisis of a generation, and understanding what factors lead to better population outcomes will help societies cope with current and future challenges.

Our group conducted an analysis of state-level data extracted from the COVIDcast system to examine drivers of severity of the pandemic. By looking at data on confirmed COVID-19 cases and data from the Amazoogle Cloud Societal Mobility Tracking system, we identified three key relationships that explained observed state-level variation in the incidence of COVID-19.

The available data on COVID-19 is very detailed, providing information on each state at the daily level, but required processing and merging with other data sources to provide key insights into state-level variation in disease outcomes. Specifically, we downloaded the Amazoogle Cloud dataset from the COVIDcast API (<https://cmu-delphi.github.io/delphi-epidata/api/covidcast.html>) and matched each record by state and week. Data were only available at the county level, so we aggregated to the state level, and additionally, we defined new predictor variable at the state level by computing an average across all counties.

State level variation in COVID-19 incidence has been very large. Our analysis showed that [key fact 1], and, as is shown in Figure 1 (slide 3 of presentation), [key fact 2]. These findings were reinforced by a linear regression model that we ran:

[model equation here]

This model provided additional quantitative evidence to support the idea that the Amazoogle Cloud data was strongly associated with COVID-19 incidence, after controlling for other state-level factors. Specifically, we observed that [interpretation of a model coefficient].

Supported by additional analyses that found X, Y, and Z, we concluded that state-level COVID-19 incidence appears to have been strongly influenced by [factor A] and not as strongly by [factor B]. Additional analyses that look more closely at [factor C] or deal with [some major additional tangled mess of data that we didn’t have a chance to get to] might help untangle additional important messages about the COVID-19 pandemic.