**Project 2 Proposal**

September 16, 2021

**Group 5**

Bradley Barker, Cody Gardner, Yawavi Koudjonou (Vi)

**Situation**: The CDC site offers many datasets on COVID-19, but we want to combine and clean datasets in order to show a greater picture.

**Questions**:

* What is the correlation between vaccination count and COVID deaths per state?
* What states have the highest vaccination count?
* What states have the lowest vaccination count?
* What is the correlation between hourly wage and COVID vaccinations per state?
* What is the correlation between hourly wage and COVID deaths per state?

Findings

* After reviewing the data, we see a correlation between State’s that have a lower hourly wage and the number of vaccinations that get administered in that state. The data shows that if a state has a low hourly wage, then the vaccination count will typically be low as well.

**Data Sources**:

* CDC Covid Datasets:
  + COVID-19 Deaths per 100k: <https://covid.cdc.gov/covid-data-tracker/#cases_deathsper100k>
  + Total Vaccination Doses Administered per 100k: <https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-rate-total>
* US Bureau of Labor Statistics:
  + Modeled Wage Estimates: <https://www.bls.gov/mwe/>

**Final Production Database**

* Postgres (relational database)