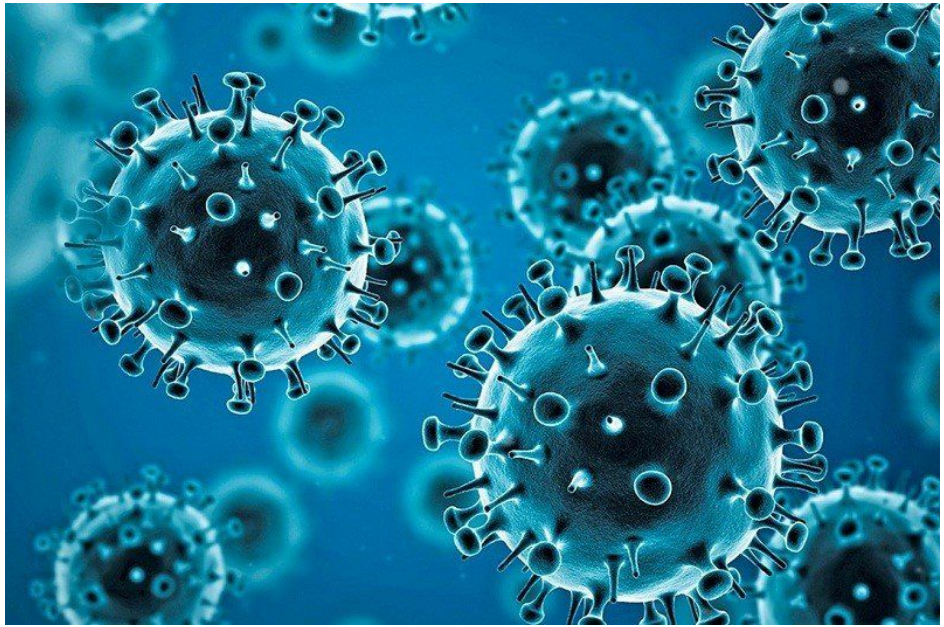
COVID-19 Severity, Vaccine Uptake, and Pre-Existing Conditions in Texas

Group 17: Kevin Jin, Mingyu Sun, Michael Tsang

Introduction and Datasets



1. Caseloads:
    a. JHU CSSE COVID-19 Data Repository → Texas (256 variables, 813 samples)
        i. Negative values deleted
    b. Land Area from The County Information Program, Texas Association of Counties
2. Vaccine Uptake:
    a. JHU CCI CRC COVID-19 Data → Texas (13 variables, 479 samples)
        i. Negative values deleted
3. Pre-Existing Conditions
    a. Texas DHHS COVID-19 Case Fatality and Demographics
    b. CDC Research Study (“Underlying Medical Conditions and Severe Illness Among 540,667 Adults Hospitalized With COVID-19, March 2020–March 2021”)

COVID-19 has resulted in over **420 million cases** and **5.91 million deaths** around the world and devastated the global economy.

We seek to understand the **factors that influence** COVID-19 cases and deaths, focusing on the state of Texas. We also attempt to **predict the course** of future cases and vaccine administrations.

This will improve our understanding of health trends and correlations with respect to pandemics, improving our preparation for future health crises.

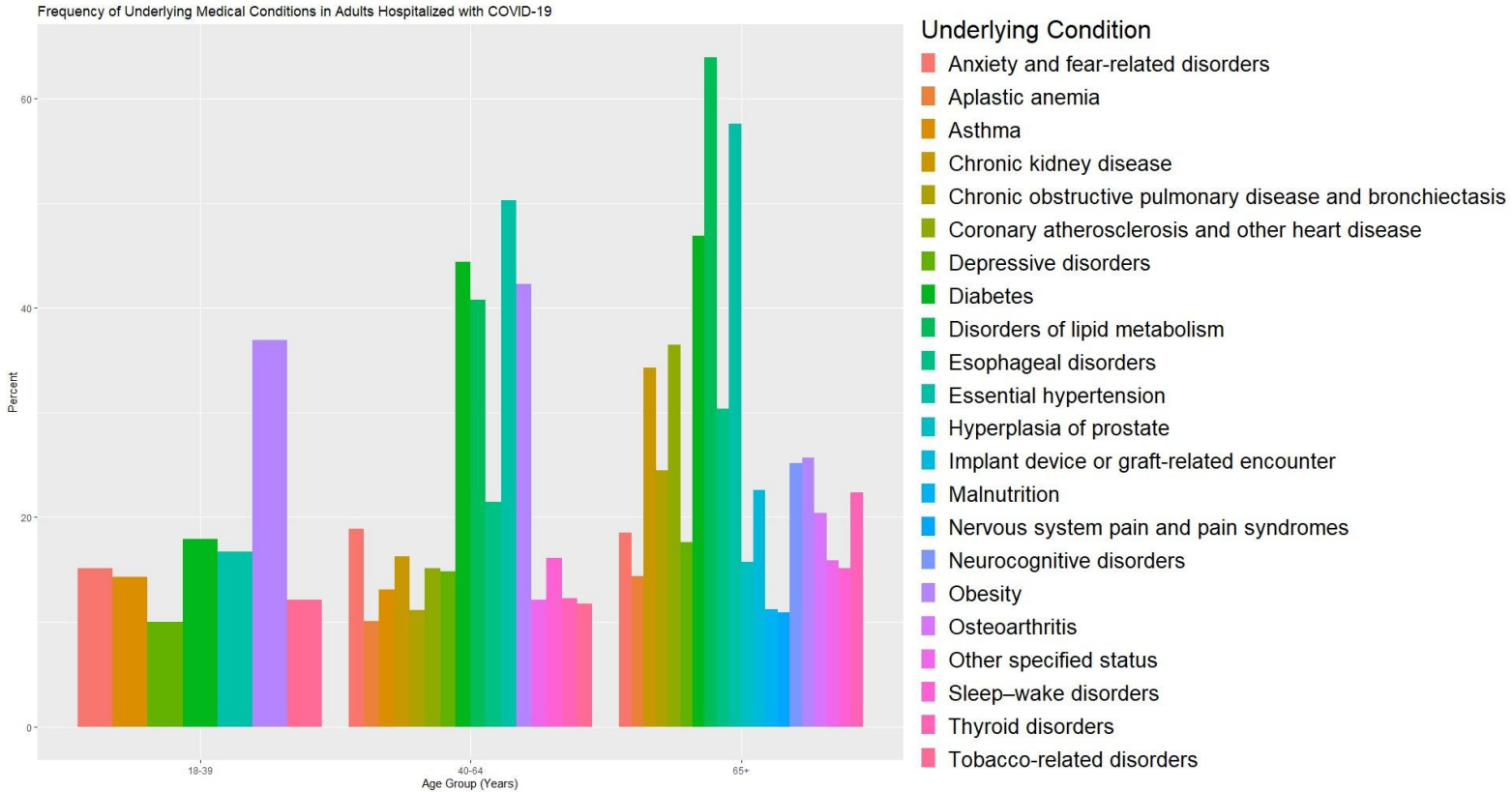
Driving Questions

1. Do other health conditions **increase** the probability of getting hospitalized for COVID? If so, which ones?
2. Are the geographical distribution and population figures in Texas **correlated** with COVID-19 cases and distribution of vaccines?
3. Is the COVID-19 fatality rate **correlated** with the vaccination rate in Texas?
4. Can we predict whether COVID-19 cases will continue to **rise** in the future or go **down**? What about vaccine uptake?

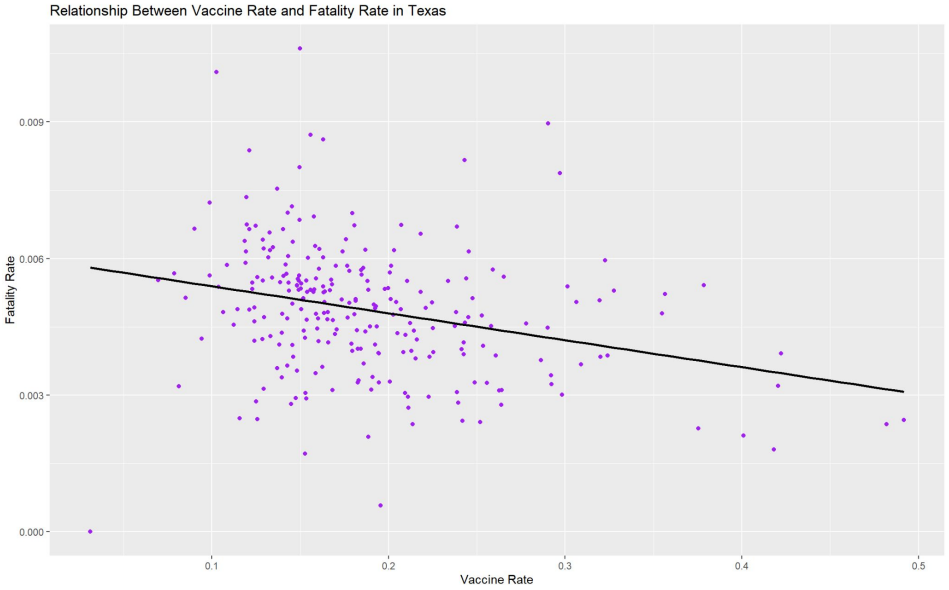
Pre-Existing Conditions and COVID-19

Most Frequent Underlying Conditions:

- Obesity
- Diabetes
- Essential Hypertension



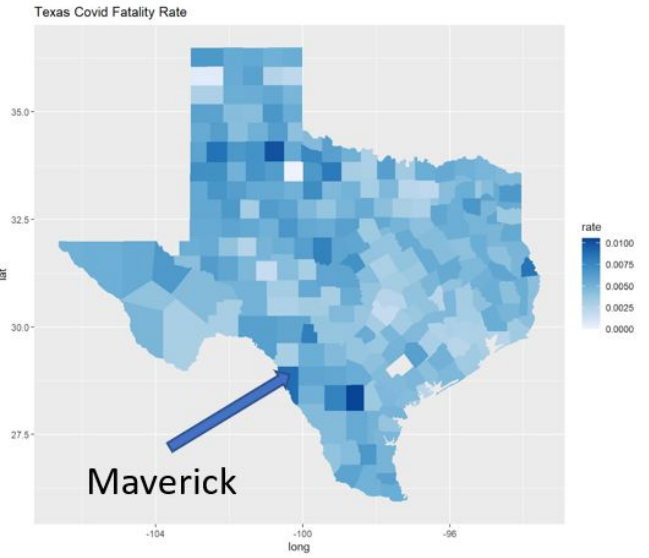
Fatality and Vaccine Rate Correlation

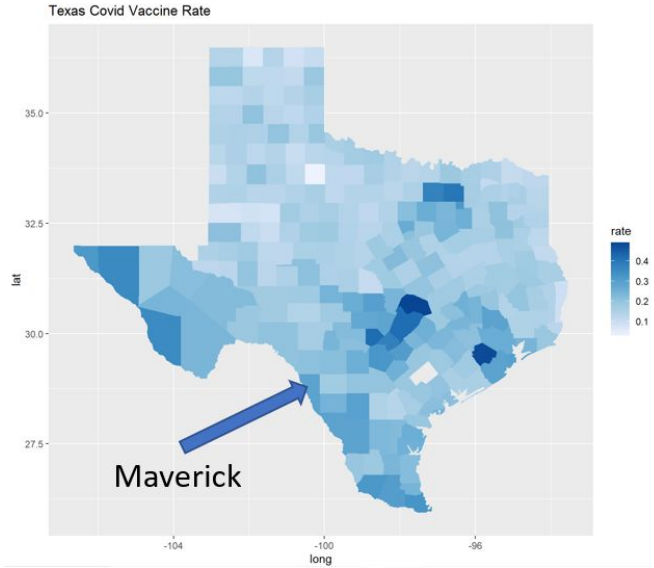


- Fatality Rate and Vaccine Rate are **negative correlated**
- The correlation is -0.3265792
- Unemployment rate was one of the factors causing outliers in this graph

```
Residuals:
      Min         1Q     Median         3Q        Max
-0.0057996 -0.0008199  0.0000568  0.0008057  0.0055164

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                      0.0059842  0.0002572  23.271  < 2e-16
texas_new_combine$vaccine_rate -0.0059252  0.0012766  -4.642 5.59e-06

(Intercept)                     ***
texas_new_combine$vaccine_rate ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.001412 on 250 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.07934,   Adjusted R-squared:  0.07566
F-statistic: 21.54 on 1 and 250 DF,  p-value: 5.588e-06
```

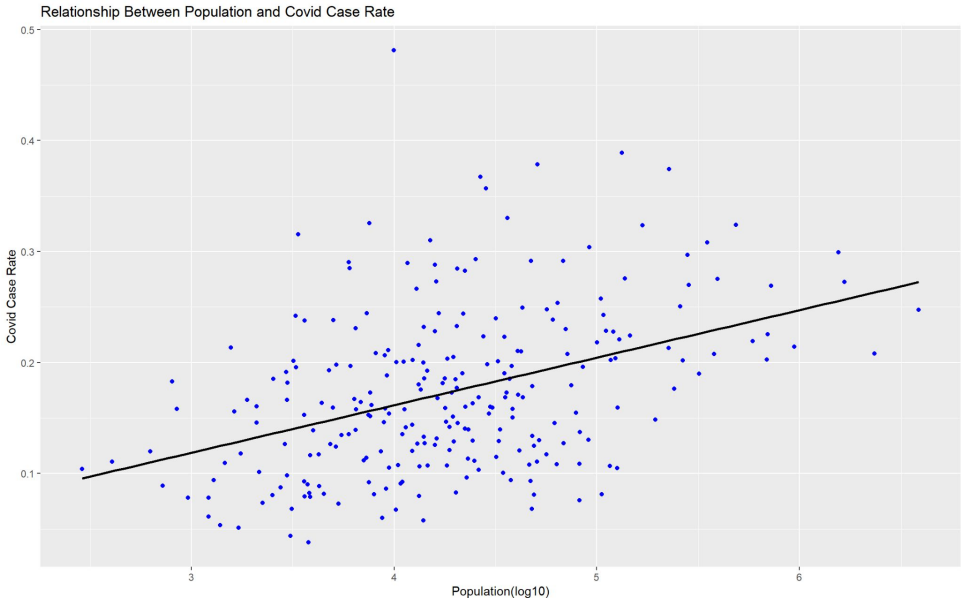
Fatality and Vaccine Rate Maps





- The unemployment rate of Maverick was 15.0%
- People do not have enough money to do treatment
- Population density could be one of the factors that caused outliers

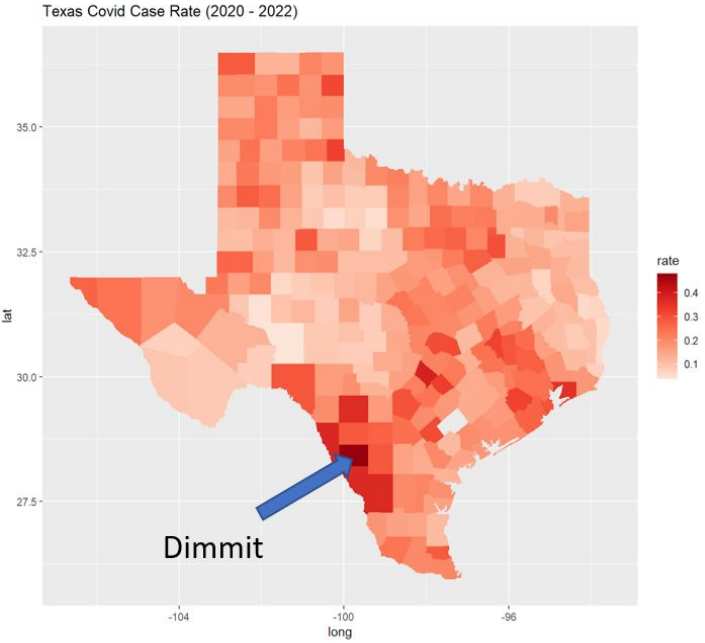
Population Density and COVID Case Rate Correlation

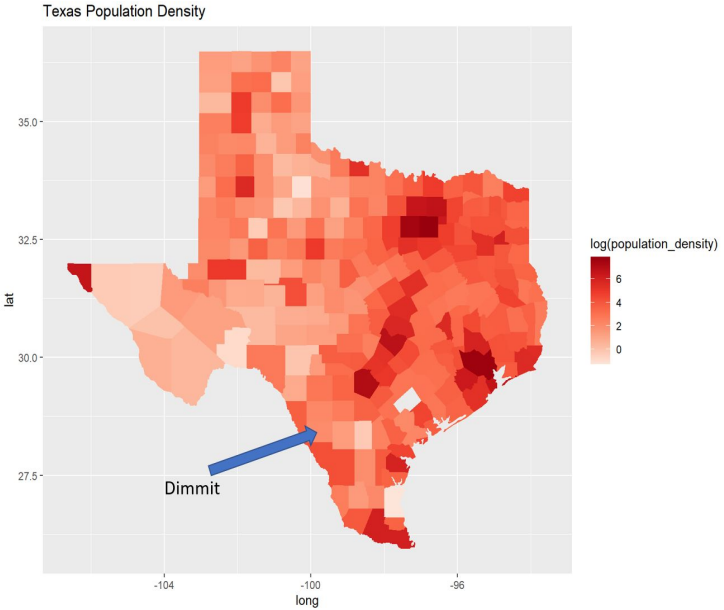


- Population and Covid Case Rate are **positive correlated**
- The correlation is 0.4029131

```
Residuals:
      Min        1Q    Median        3Q       Max
-0.131348 -0.054693 -0.009314  0.037460  0.311810

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.690e-01  4.733e-03  35.699  < 2e-16 ***
texas_new_data$pop 4.780e-08  1.369e-08   3.491  0.00057 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.07237 on 250 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.04647,   Adjusted R-squared:  0.04266
F-statistic: 12.18 on 1 and 250 DF,  p-value: 0.0005696
```

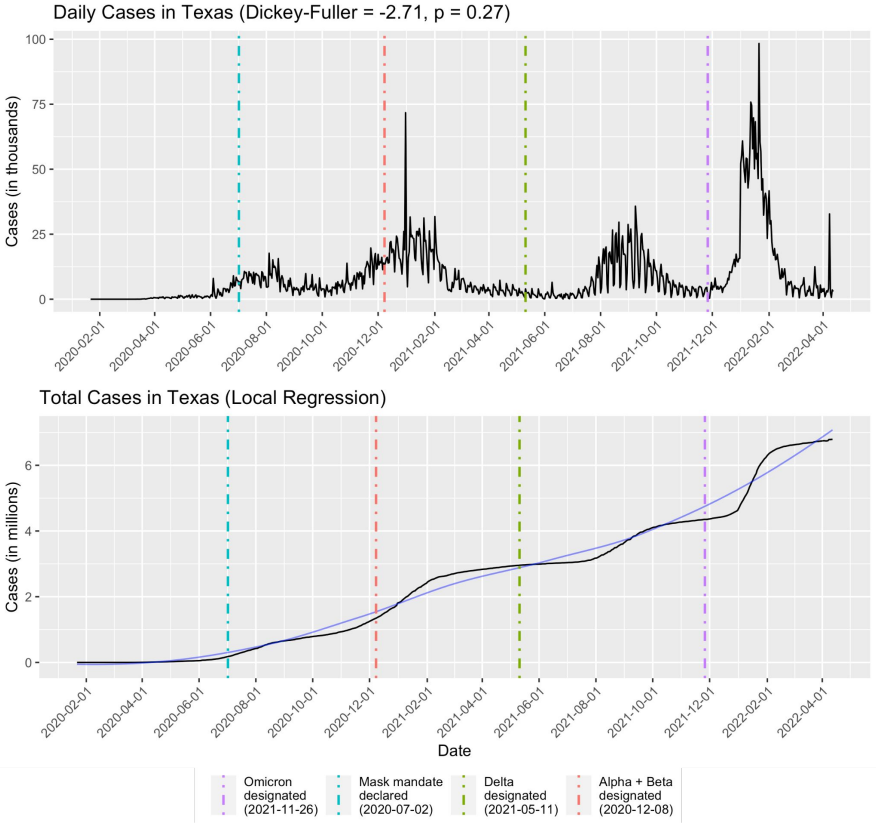
Texas Population Density and COVID Case Maps





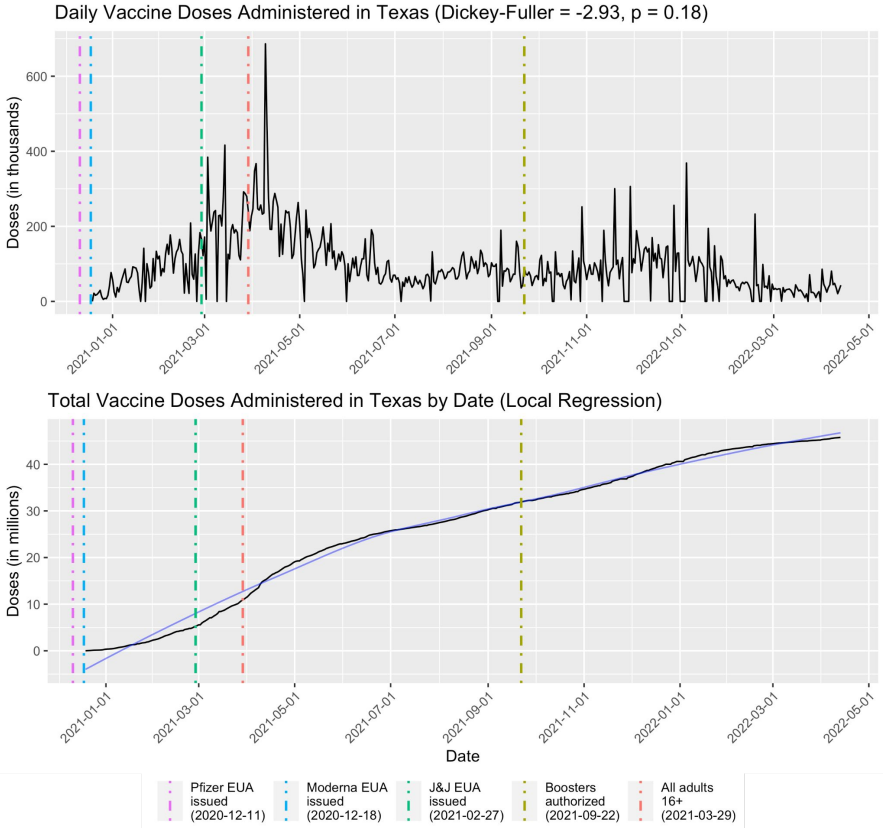
- Random error
- Sampling error
- Mass infection (gathering)
- Super spreaders

Can we predict the course of cases?



- Cases spike irregularly
- Variants can be mapped
- Augmented Dickey-Füller test : -2.71 (p-value: 0.27 > 0.05)
- Suggests a non-stationary time series; forecasting of cases **not recommended**
- Cumulative cases can be fit with LOESS
- Cases appear to be **leveling off, then growing again** (BA.2)

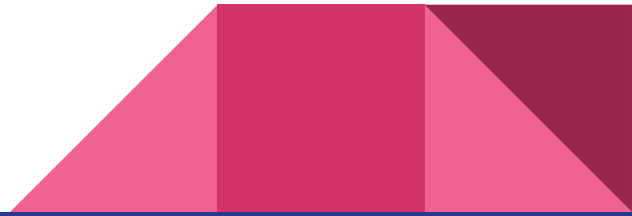
Can we predict the course of vaccine administrations?



- Vaccines spike irregularly
- EUAs and approvals mappable
- Augmented Dickey-Füller test : -2.93 (p-value: 0.18 > 0.05)
- Suggests a non-stationary time series; forecasting of doses **not recommended**
- Cumulative doses can be fit with LOESS
- Vaccines appear to be **leveling off**

Future Directions

1. What kind of people are more likely to catch COVID-19? (Age, Race, Economic Status)
2. Apply conditional models (Bayesian, SMA, WMA, EMA, ARIMA) to attempt forecasting of daily cases and vaccine administrations while accounting for external factors (e.g. local demographics, existing health conditions).
3. What are the other factors that correlate with vaccine rate and fatality rate?



References

1. Underlying conditions, demographics, caseloads, and vaccines:
   a. https://dshs.texas.gov/coronavirus/additionaldata/
   b. https://www.cdc.gov/pcd/issues/2021/21_0123.htm#tables
   c. https://txcip.org/tac/census/morecountyinfo.php?MORE=1005
2. Predicting cases and vaccines:
   a. https://github.com/CSSEGISandData/COVID-19
   b. https://github.com/govex/COVID-19
3. Images:
   a. Coronavirus disease (COVID-19): Vaccines (who.int)