Mockup for Project Proposal

Author: Angel Claudio

# Original Requirements

Submit a short (1-2 page) project proposal, including a description of the problem, why it's interesting, and what other approaches have been tried. Then there needs to be a discussion on what the student's hypothesis is and how the student's specific solution will improve or solve the problem.

# Requirements Breakdown

## In Summary

* 1–2-page project proposal.
* Description of the problem and why it’s interesting.
  + What other approaches have been tried.
* Discussion on what is the hypothesis – in other words what will the approach be and what do we expect.
  + How do we believe this solution will improve or solve the problem?

# Problem Description

Data Science has been a core field during the covid-19 pandemic in order to try and find control through predictability by using patterns found in historic data.

The [SEIR](https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology#The_SEIR_model) (Susceptible 🡪 Exposed 🡪 Infectious 🡪 Recovered) simulation model has been one of the popular [ways](https://c2m-africa.shinyapps.io/togo-covid-shiny/) of forecasting in what state the population of a give area in terms of infected and those not infected or recovered.

Configuring the equation to be accurate is a challenge when trying to predict for instance what New York’s recovered, healthy, and sick count will be over time due to the immunization phases and requirements that are to follow for the rest of this year.

# Hypothesis

We intend on using a time series which has been [shown](https://www.washingtonpost.com/graphics/2020/world/corona-simulator/) to be effective for simulating SEIR. Using NYC’s requirement for vaccinations and target goals we could simulate how the turnout may be and study early adopters of the vaccination such as [Russia](https://graphics.reuters.com/world-coronavirus-tracker-and-maps/countries-and-territories/russia/) to further tweak variables in our equation for realistic expectations.

# Team Plan

**Stage 1:**

Everyone: Sync up on choosing a dataset to use and what the approach will be to solving the problem.

**Stage 2:**

2 Persons (Group 1):

* ETL and EDA on data. Analysis conclusion.

2 Persons (Group 2):

* Documenting the resources, problem, and approach from Stage 1.
* Formatting and editing of work from stage 2 of the ETL and EDA.

**Stage 3:**

2 Persons (Group 3):

* Model building, testing, and prediction. Results analysis.

2 Persons (Group 2):

* Document formatting for results in Stage 3. Wrap up in conclusion. Export in a summarized version to Power Point for possible presentation tool.