# **Covid, Market Volatility and Liquidity (Progress Report)**

**Natasha Karwal, Minh Nguyen, Abdul Tarawally**

## Introduction

The dataset that we are using is quite extensive. We are in the process of trimming the dataset since not all of the predictors will be needed to address our question concerning the effect of Covid-19 on the financial market volatility. We are currently running 3 different models with 3 different response variables to test volatility. The response variables are vix.close (Volatility Index close price), vix.open (Volatility Index open price) which might be more useful since investors would want to know market volatility at the start of each day, and vix.change.percent (Volatility Index percent change between each observation – newly created predictor). Below are some details, broken up into accomplishments, goals, roadblocks and lesson learned.

## Accomplishments / Initial Findings

* Descriptive Analytics – Basic statistics (mean, range, sd, etc) and Exploratory Data Analysis (EDA).
  + Plots were generated using using vix.close as the response variable
  + Curvature and different spread suggested transformation.
  + Histogram for most quantitative shows various gaps with non-normal distributions and skewness with long tails. Stockpreviousrate/stock return show somewhat pattern of normality.
* Data Transformation and Creation - Categorical variables (category, week, company, result) were transformed to factor using as.factor() and stock close, vix close, confirmed total, and deaths total were lagged by 1. Lagging data could help reduce serial correlation causes by the timeseries nature of our dataset. We will run a few tests such as the Durbin-Watson (DW) Test in R to measure the level of serial correlation of our models.
  + Using the transformed variables to generate basic visual analysis to highlight the impact on market volatility and stock fluctuations with the onset of covid confirmed cases on various companies.
  + Stockpreviousrate/stock, previous (stock close price lagged by 1), change.percent (stock % change lagged by 1), result (2 levels, up/down, categorical variable expressing if the stock close price increased or decreased), vix.change (raw vix change using vix.previous – vix.close), vix.previous (vix close value lagged by 1), vix.change.percent (vix % change lagged by 1), vix.result (2 levels, up/down, categorical variable expressing if the vix close increased or decreased), confirmed.previous (confirmed total lagged by 1), confirmed.change.percent (confirmed total % change lagged by 1), deaths.previous (deaths total lagged by 1), and deaths.change.percent (deaths total % change lagged by 1),
* Simple regression between vixclose and confirmedtotal
  + The residual plotting shows a possible outlier with no apparent pattern of a horizontal band or linearity. The QQ plots also show cluster of data in intervals not falling on the line and moving away at the ends. The R2 value being 30% confirms the model not being favorable.
* Mulitple regression
  + Test run model did not provide a good result, even after the inclusion of another predictor(Stockpreviousrate/stock return). The QQ plotting appear closer to the line in comparison to the previous model. Variable transformation using sqrt(vixclose) did result in improvement. R^2 to 41% and QQ plotting appears closer from last modelling.
  + Model 1 (Adjusted R-squared: 0.8591) – vix.open: lm(vix.open ~ category + open + previous + vix.previous + confirmed.previous + deaths.previous, data = weekly.model)
  + Model 2 (Adjusted R-squared: 0.9176) – vix.close: lm(vix.close ~ category + open + high + low + close + previous + vix.previous + confirmed.total + confirmed.change.percent + deaths.total + deaths.change.percent, data = weekly.model)
  + Model 3 (Adjusted R-squared: 0.6982) – vix.change: lm(vix.change ~ category + open + high + low + close + previous + change.percent + vix.change.percent + confirmed.total + confirmed.previous + confirmed.change.percent + deaths.total + deaths.previous + deaths.change.percent, data = weekly.model)

## Goals / Next Steps

* Variable and model selection are our main focus for the next few days.
  + We will use a stepwise approach for variable selection.
* Finalize the presentation deck and visual supports.