**Project Title: Reddit GameStop Analysis -** Does Reddit message Traffic affect GameStop pricing

**Team Members:** Carter Kioski, Claire Wong, Ed Ober

**Project Description / Outline**

A group of Reddit users have driven up the price of the GameStop stock over the past year, collaborating through the Reddit application and purchasing stock. Our theory is that message volumes on Reddit would precede movements of the GameStop stock price. Our approach will be to create a model of price by day and hour against the message volume, lagged by a range of hours.

**Research Questions to Answer**

H0 : Reddit message volume has **no** effect on GameStop stock price within 3 hours

HA : Reddit message volume has an effect GameStop stock price within 3 hours

**Datasets to be used**

* **Messaging Volume: (Reddit and/or Twitter API):** To provide message volume by day and hour for the specific group, and for those messages containing a variant of $GME, GameStop, or potentially other names / tags
* **Yahoo Stock Market History: (Kaggle or Yahoo file):** hourly stock prices by day for the GameStop and AMC stocks.

**Rough Breakdown of tasks**

* Identify and validate data sources to be used
  + Messaging volume: Reddit API, Twitter API
    - Data by Date and Hour
    - Identifiable to the Reddit Group or the activity
    - Text is indexed or can be scanned for keywords for GameStop
  + Stock Market history: Yahoo, Kaggle
    - Data by Date and Hour
    - For GameStop - $GME
    - Pricing for the period: Open, Close, High, Low
* Outline of the analysis
  + Overview of the “story”
  + Data analysis steps
    - ETL, particularly of the message data
    - Merging and verification
    - Modeling
    - Analysis of results
    - Redos
  + Define set of graphs showing the story
  + Documentation
  + Integrate graphs and documentation
    - Highlights / lowlights
    - Results