**MSU Data Analytics Boot Camp – Project 3 Machine Learning “Game Plan”**

1. Read Kaggle idea / approach
2. Read Slack info provided
3. Look at my dataset
4. Print and review the Kaggle project overview

* Classification scope
* Review lectures and class slides around:
  + Linear Regression
  + Random Forest
  + Regression
* Possible y-values:
  + Claim amount
  + Type of claim
  + Could also try to determine premium based on the inputs
* Possible x-values:
  + Other columns in dataset
* Data considerations:
  + Type = categorical
  + Fraud = categorical
  + Claim = number
  + Premium = number
* So which techniques apply to categorical versus number? Need to determine this
* Will be relatively easier to use my existing data / project as opposed to starting a whole new project
* Kaggle example:
  + Premiums can be done with linear regression
  + Could others work?