# Final Project Plan

# Ben Karabinus

# COMP 4442

# Fall Quarter 2021

# Project Team

- Ben Karabinus

# Possible Methods for Analysis

## Method 1 – Gradient Boosting Machines:

My top choice for methods of analysis to use in the final project is Gradient Boosting Machines. I would like to use one of the R libraries listed below to build, tune and train a model to predict home prices.

- gbm
- xgboost
- h20

## Method 2 – Principal Components Analysis

My second choice for methods of analysis to use in the final project is Principal Component Analysis. I would like to use the R library listed below to carry out an analysis of what factors contribute most significantly to the prices of homes across different regions of the US.

- stats

## Method 3 – Simple-Effects Analysis

My third choice for methods of analysis to use in the final project is Simple Effects Analysis. I don't know much about simple-effects analysis. I'm not sure which R libraries I would use.

## Data/Data Domain

I would like to use a combined data set consisting of the Zillow Home Value Index (ZHVI) and data pulled from the Bureau of Economic Analysis (BEA). The combined dataset will include typical home values for metropolitan statistical areas in the United States and socioeconomic factors for metropolitan statistical areas in the United States (per capita income, cost of living, etc.

References:

Zillow home Value Index

Bureau of Economic Analysis