Capstone Proposal:

* What is the problem you want to solve?
  + How is it possible to predict the real estate value in DC?
* Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis that they wouldn’t have done otherwise?
  + This knowledge can be used to advise real estate developers and investors.
  + This can also be used by the government and NGOs to problem-solve issues arising around high cost-of-living and segregation in the city.
* What data are you using? How will you acquire the data?
  + I found the data from a Kaggle competition, D. C. Residential Properties. It is provided by Open Data DC.
  + https://www.kaggle.com/christophercorrea/dc-residential-properties
* Briefly outline how you’ll solve this problem. Your approach may change later, but this is a good first step to get you thinking about a method and solution.
  + Clean the data
    - Evaluate missing data. Remove or replace with meaningful values.
    - Reorganize, delete, or rename features as needed
  + Exploratory statistical analysis
    - Outlier evaluation and removal if needed
    - Ranges, measures of center
    - Distributions
  + Machine learning applications
    - Explore various predictive models to find which best predict outcomes
      * Linear regression
      * Logisitic regression
      * Clustering
      * Random Forest
  + Presentation
* What are your deliverables? Typically, this includes code, a paper, or a slide deck.
  + Jupyter Notebook