Project Proposal

1. Me
2. I chose my project topic by looking through Data.gov and found data on food affordability in the US. Using this data is important to me because it is only for female-headed households which is the type of household I was raised in, and I want to see how affordable/unaffordable food is for moms to purchase in the country.
3. - Research question: Can I predict the ratio of average cost of food to income for different regions in California based on socio-economic and demographic features?

* Where the data is from: <https://catalog.data.gov/dataset/food-affordability-fc448>
* Elements to be used:
  + Feature engineering for non-tabular data
  + Linear models, KNN, random forests, SVM, maybe another
  + Develop and tune hyper-params and tuning deep learning model
  + Cross validation techniques
  + SHAP/other to explore feature importance
  + Cluster analysis
* Timeline:
  + Explore data
  + Feature engineering
  + Dat visualizations
  + Nov 1: EDA
  + Implement additional elements not already included
  + Nov 20: Rough Draft
  + Fix issues
  + Dec 11: Final Draft