

PS2_{Quick}

Sam Quick

January 2019

1 Main Tools of a Data Science

****Six broad categories discussed in class:****

1. Measurement We discussed how important measurement is in forming policies and insights, as you cannot construct a policy "without first measuring what the policy is meant to accomplish."
2. Statistical Programming Languages -We discussed three languages that will be utilized in this class: R, Python, and Julia.
3. Web Scraping We discussed the different types of Data scraping such as utilizing APIs or Parsing
4. Handling Large Data Sets Under this topic we talked about utilizing RDDs and SQL to handle data sets that are far too large to do otherwise
5. Visualization Under this section we talked about the different ways to visualize data through the three coding languages we discussed above. R - ggplot2; Python - matplotlib; Julia - plot.jl
6. Modeling Now we discussed the modeling of data that we will do to make sense of the data. I am sure we will utilize a lot of topics learned in Econometrics.