Coursera Stats Video Lectures – Sampling and Source of Bias

Wk 1, Video 3

1. Census vs Sample
   1. Sample : a part of the population
      1. benefit: economical and easy to gauge the whole population
      2. sampling a small population and measuring is called exploratory analysis
      3. Inferences can be drawn from the sample population
      4. sampling should be truly random to properly infer about the population
   2. Census: the whole population
      1. disadvantage: some individuals are hard to locate/measure
      2. populations rarely stay static, so it’s not possible to get a perfect measure
2. Sources of Bias in studies
   1. convenience bias: individual who are easy to access are more likely to be included in the sample
   2. non-response bias: if only a non-random fraction of the population responds to a survey, then the sample no longer represents the population
   3. Voluntary response bias: occurs when only those who volunteer to respond because they have strong opinions on an issue (such as political polls)
3. Sampling Methods
   1. simple random sample
      1. each case is equally likely to be selected
      2. such as drawing names from a hat
   2. stratified sample
      1. divide the population into homogenous strata, then randomly draw from each strata
      2. such as polling from each gender
   3. cluster sample
      1. divide population into clusters, then randomly select some clusters, then sample all observations within clusters
   4. multistage sample
      1. divide population into clusters, then randomly select some clusters
      2. then randomly sample from each cluster
      3. such as selecting random but similar neighborhoods within a city then polling in the selected neighborhoods
      4. very economical