Coursera Stats Video Lectures, Unit 1 vid 1

Vid 1: Data Basics – Source information: Google Transparency Database

Observations

Variables – Types, etc

1 Numerical (quantitative)

1. numbers
2. makes sense to add, subtract, take averages of these values
3. could be continuous, such as height (infinite possibilities within a set range).
4. could be discrete (specific set of numeric values), countable. (ex: number of cars a household owns
   1. count data= discrete variable

2. Categorical (qualitative)

* 1. may be categories or numbers that represent categories (such as 1 for female,2 for male), etc
  2. arithmetic doesn’t make sense on these variables
  3. Ordinal variables: have levels, such as good, neutral bad
  4. regular categorical: no order, example, are you a morning person or a night person

Relationships between variables

1. Associated/Dependent Variables
   1. if two variables plotted on a graph show a trend, they’re considered associated or dependent variables
   2. upward trends are called positive, downward trends are called negative
2. Independent Variables
   1. if a graph of the two variables shows no trend, they’re considered independent

Matrices

1. Rows represent cases, columns represent variables or observations