Stat 157: Applied Stats for Computing
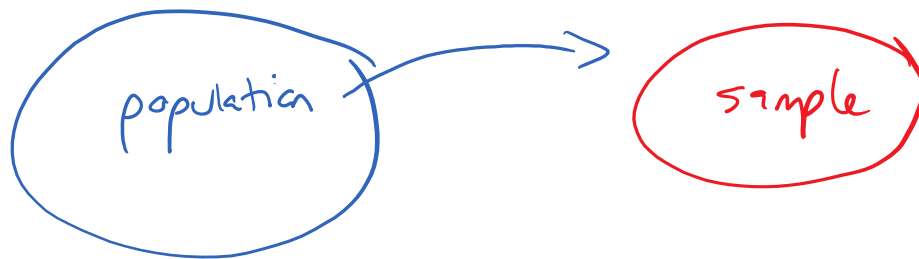
Chapter 1:   Describing Data
Section 1.1:   Variables and Data

statistics ≡   a branch of applied mathematics concerned with the collection and interpretation of data

ideas of collecting data:

population ≡   the entire set of measurements of interest

note: sometimes not practical or possible to collect data on the entire population

sample ≡   a subset of the measurements of interest

⇒ must ensure as far as possible

that the sample is <u>representative</u>
↗

the sample should look like
the population

two types of statistics:

descriptive — procedures used to summarize
and describe the important
characteristics of a set of data

inferential — procedures used to draw
conclusions or make predictions
about a population based on
a sample

variable ≡ a characteristic that either

① changes over time

② changes for different individuals or objects
under consideration

2019/10/22

example: the height of an individual tree
measured over a period of years
— ①

the height of all trees within a
certain area at a particular time

- (2)

experimental unit - individual or object on which
a variable is measured

_____

univariate data - result of a single variable
measured on experimental units

(height of a tree)

bivariate data - two variables

(height and diameter of a tree)

multivariate data - more than two

_____

qualitative variable - measure is a quality or
characteristic

⇒ does not result in a numerical value

examples: political party
nationality
favourite food

quantitative variable - measure is a numerical quantity

examples: height, speed

number of students in a class

for quantitative variables, two types:

discrete   vs.   continuous

discrete ~ can only have finite or
            countable number of values

example of finite:
    values can only be one of
        { 3.75, 5.21, 8.32, 9.21 }

example of countable:

    shoe size: ... 11, 11½, 12, 12½, 13, ...

Continuous ~ can be any real number

    examples: speed
                 mass
                 length

note: although height, for example, is
      in theory a continuous variable

    practically speaking, we usually round
     our measurement to a certain precision

(certain number of decimal places, or to the nearest millimetre)

due to

→ limitations of your measuring instrument

→ limitations on object being measured (fuzziness of a tennis ball)

variable

qualitative

quantitative

discrete

continuous