Biostats Lecture 1: Descriptive Statistics

Public Health 783

Ralph Trane
University of Wisconsin–Madison

Fall 2019

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Learning Objectives

1. Understand why descriptive statistics is important, and useful
2. Know the difference between discrete and continuous variables/data
3. Have some ideas of which summaries and figures are appropriate for different types of data

Descriptive Statistics

- What: the art of describing data with few important measures ('summary statistics')
- Why:
  - know your population!!
  - explore your data
- How: try to get an idea of the distributions of variables included
  - what's a distribution?!
  - what's a variable?!

Data Types

Two general data types:

- Discrete data
  - categorical
    - no natural ordering
    - examples: sex, race, blood type, political orientation, etc.
  - ordinal
    - naturally ordered
    - educational level, age groups, disease severity scales, etc.
  - summarized by
    - frequency counts
    - relative frequencies
- Continuous data
  - numerical
    - examples: age, height, weight, BMI, proportions, etc.
  - infinite (uncountable, actually...) number of potential values
  - summarized by
    - location measures
    - spread/variation measures

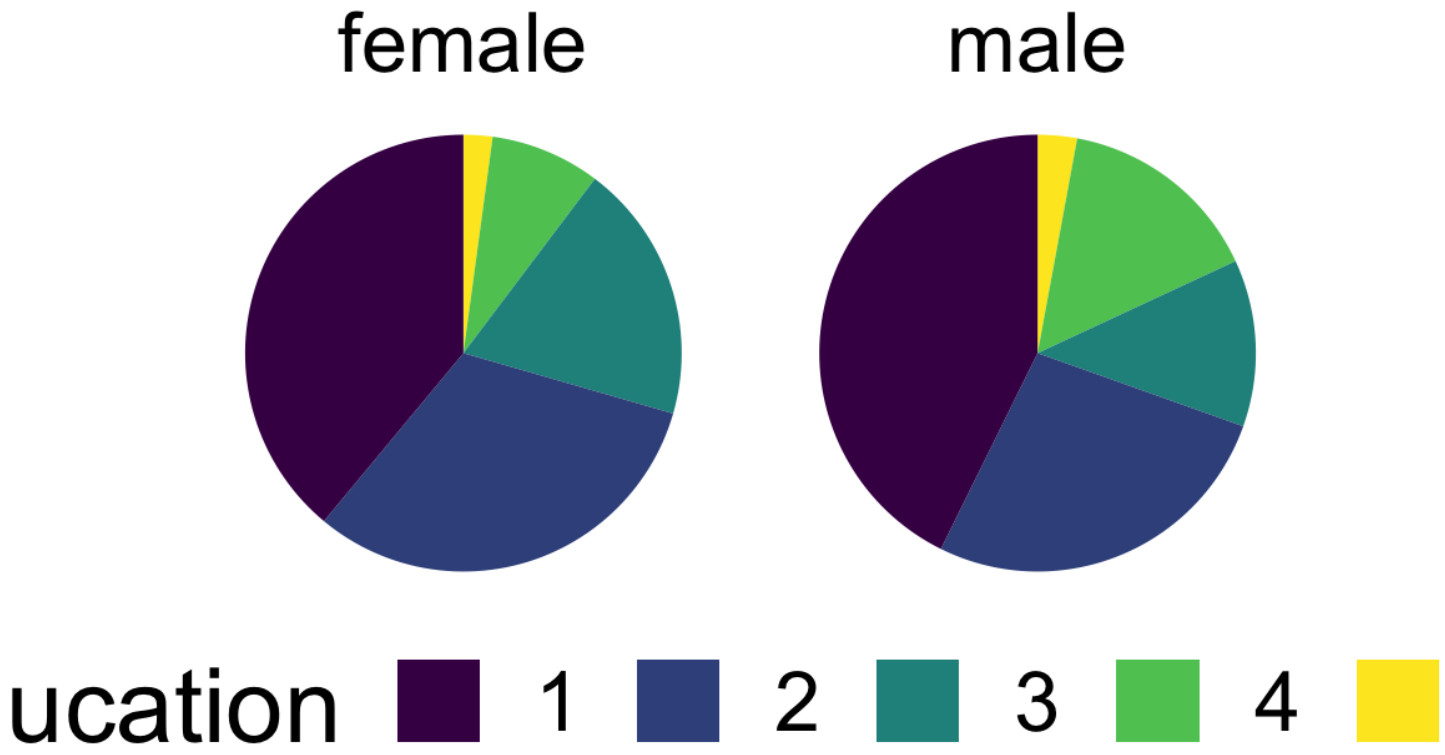
Example

Framingham Heart Study

Quick Detour

What NOT to do with categorical data.

A.K.A. my least favorite chart of all time...



Quick Detour

Why don't I like pie charts?

What we want from figures:

1. self-explanatory
2. important information should be easy to get
3. show trends (if available)
4. only as complicated as the data
5. free of unnecessary complexity
    - irrelavent decorations
    - 3D effects --(**BIG YIKES!!!!!**)

The pie chart violates 2, 4, and 5.

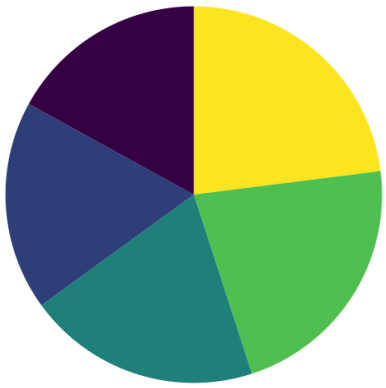
- 5: a bar chart is almost always more appropriate
- 2 and 4: pie charts compare angles. Humans are awful at comparing angles!

Quick Detour

From Wikimedia Commons

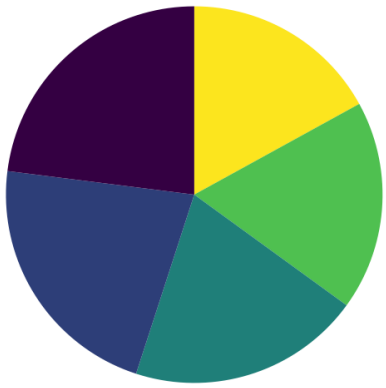
For each of the following, rank the colors in terms of size:

1



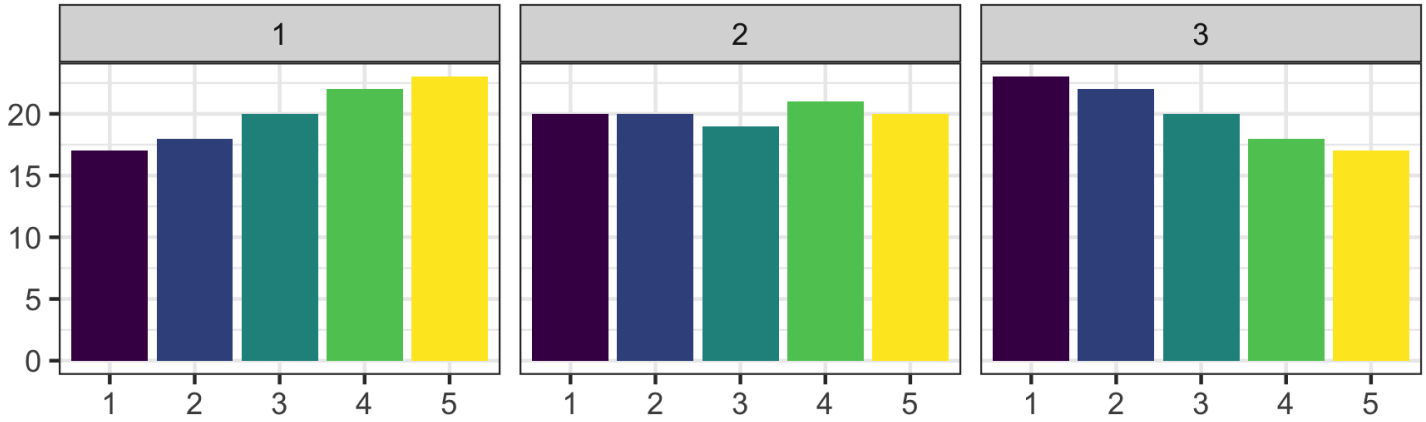
2

3



Quick Detour

For each of the following, rank the colors in terms of size:



Quick Detour

Quick Detour

> "Only pie chart ever allowed:"
>
> 
>
>
> --- Me (24/9/2019)

That was two days ago, though. Things change... I guess.

> "All pies might not be made equal.....?"
>
> --- Me (25/9/2019)

Before using a pie chart, read this, this, and this. This Twitter thread is also great!

Quick Detour