Notes: MS 204 Chapter 1 part I

Overview

- Data basics
- Data collection principles
- Sampling and experiments

Data basics

```
library(tidyverse)
library(gapminder)
gapminder %>% head(4)
```

```
## # A tibble: 4 x 6
##       country continent  year lifeExp      pop gdpPercap
##        <fctr>    <fctr> <int>   <dbl>    <int>     <dbl>
## 1 Afghanistan      Asia  1952  28.801  8425333  779.4453
## 2 Afghanistan      Asia  1957  30.332  9240934  820.8530
## 3 Afghanistan      Asia  1962  31.997 10267083  853.1007
## 4 Afghanistan      Asia  1967  34.020 11537966  836.1971
```

- Elements of a data frame

- Variable types

Aside Using examples not yet discussed, identify one of each of the four types of variables above

Data collection principles

- Anecdotal evidence
- Census
- Observational Studies
- Experiments

Aside Make a table with +'s and -'s to each of the above

Sampling bias types:

- Non-response
- Voluntary response
- Convenience sample
- Survivorship bias

Examples: Alf Landon versus FDR

Why sample? Exploratory analysis to inference

Obtaining good samples

- Simple random sample
- Stratified sample
- Cluster sample

Obtaining good experiments

- Placebo effect and placebo treatments
- Blocking