Cross-sectional Data
DPR 190

Miles D. Williams

Data for Political Research

Cross-sectional data

What is it?

Cross-sectional data

A dataset:

```
## # A tibble: 5 x 5
##   ccode1 n_events fatalmin fatalmax country
##    <dbl>    <dbl>    <dbl>    <dbl> <chr>
## 1    517       27       16      425 Rwanda
## 2    490       15       16      262 Congo - Kinshasa
## 3    290        5       11       20 Poland
## 4    702       15        2        2 Tajikistan
## 5    200       73       24       71 United Kingdom
```

Cross-sectional data

Tidy cross-sectional data:

- Each *row* is a unique unit of observation
- Each *column* is a variable for a unique unit
- Each *cell* contains only one value

Cross-sectional data

Not a cross-sectional dataset:

```
## # A tibble: 5 x 6
##   country     continent  year lifeExp      pop gdpPercap
##   <fct>       <fct>     <int>   <dbl>    <int>     <dbl>
## 1 Afghanistan Asia       1952    28.8  8425333      779.
## 2 Afghanistan Asia       1957    30.3  9240934      821.
## 3 Afghanistan Asia       1962    32.0 10267083      853.
## 4 Afghanistan Asia       1967    34.0 11537966      836.
## 5 Afghanistan Asia       1972    36.1 13079460      740.
```

Cross-sectional data

If a unit makes **makes more than one appearance** in the data, the data is either not tidy or not cross-sectional.

Uses of cross-sectional data

- What's a benefit of cross-sectional data?
- What's a limitation?

Uses of cross-sectional data

- Efficient, snap-shot, case comparisons, generate hypotheses
- What's a limitation?

Uses of cross-sectional data

- Efficient, snap-shot, case comparisons, generate hypotheses
- Apples-to-oranges, selection bias

Table of Contents