

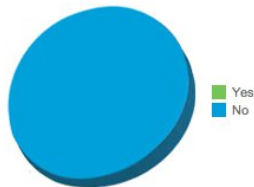
VARIABLES

1 Definitions

2 Issues

3 Recoding

4 Practice



Definitional issues

Mathematical foundations

- **Random variables:** numbers assigned to states
- **Continuous variables:** ranges of values
- **Discrete variables:** sets of values

Dimensions and types

- **Continuous:** infinite/large number of possible values
- **Categorical:** finite/small number of possible values
- **Ordinal:** meaningful ordering of the values

Issues

Measurement

- **Availability** of the data
- **Accuracy** of the measurement
- **Meaningfulness** of the unit

Coding

- **Values:** -1 0.7 9 (numeric), "UK" (string), . (missing)
- **Variable labels:** "Fertility rate"
- **Value labels:** 1 "White" 2 "Black" 3 "Hispanic" ...

Recoding

Rationale

- **Create groups**, e.g. age cohorts, social classes
- **Create dummies**, i.e. binary (0/1), true/false indicators
- **Change encoding** for missing values

Stata commands

- Create variables: `gen`, `tab`, `gen()`, `clonevar`, ...
- Recode values: `recode`, `irecode`, `replace`, ...
- Assign labels: `la var`, `la def`, `la val`, ...

Practice: WVS dataset

“[The government] should implement only the laws of the sharia.”

- Measured on a 5-point scale of agreement
- Measured in 9 African and Asian countries
- Measured between 1999 and 2004

Data:

- World Values Survey (WVS)
- Sample: resident populations aged 15+



Practice session

Class

* Get the do-file for this week.

```
srqm fetch week3.do
```

* Open to read and replicate.

```
doedit code/week3
```

Coursework

- Finish the do-file and read all comments at home.
- Follow instructions on top of the code.
- Form a group and write a common draft do-file.

Exercises

Ex 3.1. European Social Survey 2008

- 1 Load the data.
- 2 Find the variable measuring left/right political positioning.
- 3 How continuous is the measurement?

Ex. 3.2. Quality of Government 2011

- 1 Load the data.
- 2 Find all variables on democracy.
- 3 Which measure seems most reliable?