Quantitative methods
Week #2

Gergely Daróczi

Corvinus University of Budapest, Hungary

17 February 2012

BUDAPESTI
CORVINUS
EGYETEM

Outline

1. Types of error
   - Measurements
   - Gone Hunting

2. Elements of Research Design
   - Possible goals of research
   - Units of analysis
   - Ecological fallacy
   - Ecological fallacy
   - Variables and attributes
   - Types of variables and attributes
   - Relation between variables

3. Levels of Measurement
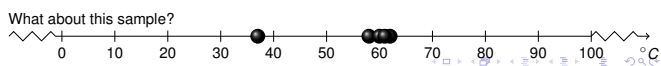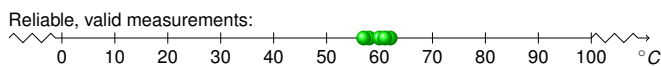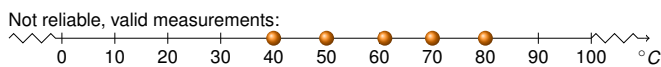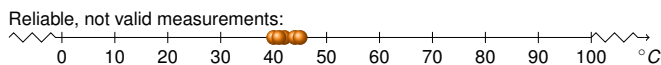   - Relationship between variables

4. Stages of Social Research

Measurements
Types of error

Measuring the temperature of a bowl of water kept at 60°C

Not reliable, not valid measurements:

Reliable, not valid measurements:

Not reliable, valid measurements:

Reliable, valid measurements:

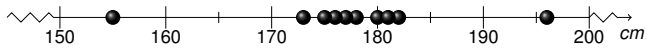What about this sample?
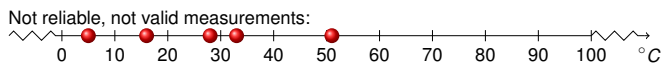
Gone Hunting
Types of error
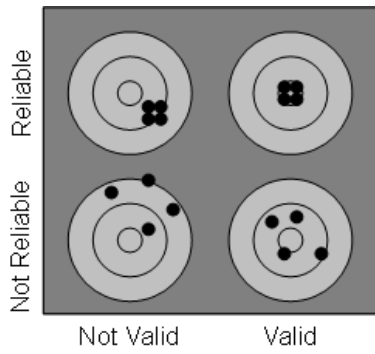
Source: customerthink.com

Social research
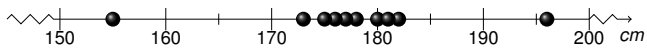### Types of error

We measured the height of 10 students:

Social research
Types of error

We measured the height of 10 students:

Source: http://www.wallpaperhere.com/

Notes

Notes

Notes

Possible goals of research
Brainstorming

Which of the followings would make a good topic for research?

- How many homeless people lives in Budapest?
- What is the most popular way of suicide?
- What color is a banana?

Notes

Possible goals of research
Example of topics seeming meanless

A simple time series analysis:

Banana timeline

Notes

Possible goals of research
Example of topics seeming meanless

A simple time series analysis:

Banana timeline

Color names if you're a girl...

Color names if you're a guy...

Source: http://blog.xkcd.com/2010/05/03/color-survey-results/

Notes

Possible goals of research
Brainstorming

Which of the followings would make a good topic for research?

- How many homeless people lives in Budapest?
- What is the most popular way of suicide?
- What color is a banana?
- How much does a pile of apples cost?
- What kind of relationship can be found between bodyweight and hair-color?
- Which is the best book in the world?

What could be the main factors of choosing a topic?

Definition: exploratory vs. descriptive vs. explanatory study

---

Units of analysis
Units, cases, objects

What entities are described and compared below?

- Are older people more afraid of crime than younger people?
- Does economic development lower the birth rate?
- Which Hungarian soccer team has the utmost fan club?
- The longer the emgagement period, the longer the marriage.
- What is the most prestigious occupation?

Do we always ask the units of analysis?

Units of analysis
Ecological fallacy and so

We found, that 2 million tourists arrived to Budapest in 2010. It was also known, that 2.5 million tourists visited the lake Balaton in the same year.

Could we state that ...

- the inhabitants of Budapest are not as friendy as people living at the lake Balaton?
- Budapest has less spectacle than the lake Balaton?
- the lake is a bigger tourist goal?
- tourists think that lake Balaton is a lot nicer place than Budapest?
- the beach of Balaton is lot more crowded than the streets of Budapest?

What can be said based on the above data?

Ecological fallacy
A classical example

Emile Durkheim (1897): *Le Suicide*

- Suicide rates are higher in men than women.
- Suicide rates are higher for those who are single than those who are married.
- Suicide rates are higher for people without children than people with children.
- Suicide rates are higher among Protestants than Catholics and Jews.
- Suicide rates are higher among soldiers than civilians.

> **Definition**
> The ecological fallacy occurs when you make conclusions about individuals based only on analyses of group data.

Variables and attributes
A concrete example

How old are you?

- >18
- 19-24
- 25-30
- 31-40
- 41-100
- 100<

What is variable and what is attribute in the above example?

Variables and attributes
Test your knowledge!

Research topics:

- Are older people more afraid of crime than younger people?
- Does economic development lower the birth rate?
- Which Hungarian soccer team has the utmost fan club?
- The longer the emgagement period, the longer the marriage.
- What is the most prestigious occupation?

What is variable and what is attribute in the above example?

Types of variables and attributes
Towards methodology

Types of variables

- explanatory variables
  - dependent variables
    - qualitative variables
    - quantitative variables
  - independent variables
    - qualitative variables
    - quantitative variables
- extrameous variables
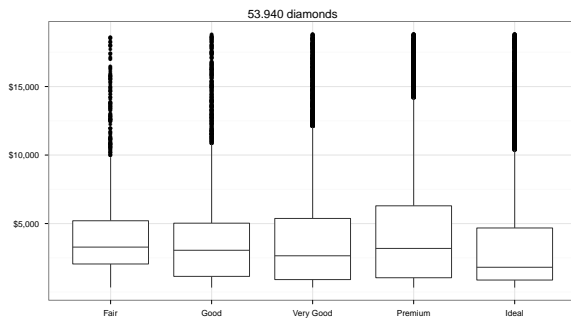 - control variables
  - other variables

Let's make up some examples based on the above list!

Relation between variables
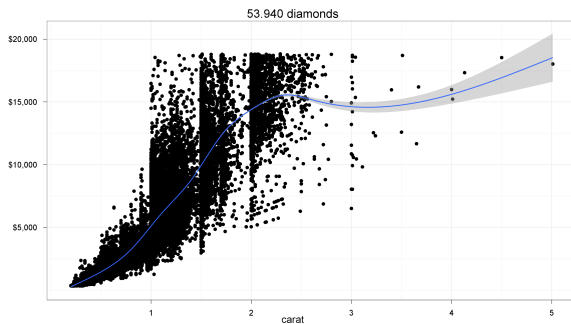### A visual example

53.940 diamonds

```
ggplot(diamonds, aes(cut, price)) + geom_boxplot() + xlab('') + ylab('') +
scale_y_continuous(formatter="dollar") + theme_bw() + opts(title="53.940 diamonds")
```

Relation between variables
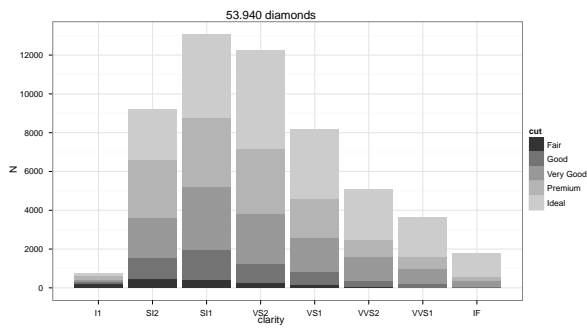### A visual example

53.940 diamonds

```
ggplot(diamonds, aes(carat, price)) + geom_point() + geom_smooth() + ylab('') +
scale_y_continuous(formatter="dollar") + theme_bw() + opts(title="53.940 diamonds")
```

Relation between variables
A visual example

53.940 diamonds


```
ggplot(diamonds, aes(clarity, fill=cut)) + geom_bar() + ylab("N") +
theme_bw() + opts(title="53.940 diamonds")
```
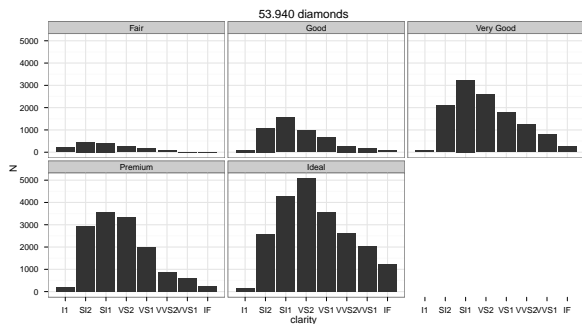
Relation between variables
A visual example

53.940 diamonds


```
ggplot(diamonds, aes(clarity)) + geom_bar() + ylab("N") + facet_wrap(~ cut) +
theme_bw() + opts(title="53.940 diamonds")
```

Levels of Measurement
Qualitative and quantitative variables in depth

Qualitative variables:

- Nominal: exhaustive labels with no intersect (mutual exlcusivity) not in a specific order
- Ordinal: an (possible) ordered variable with exhaustive labels not intersecting

	Nominal	Ordinal	Interval	Ratio
Classification	X	X	X	X
Rank order		X	X	X
Equal intervals			X	X
Nonarbitrary zero				X

Quantitative variables:

- Interval: equal distances between the ordered labels (numbers)
- Ratio: a scale with a zero point

Notes

Types of variables and attributes in practice
Relationship between variables

Possible relationship between variables:

- association,
- correlation,
- spurious relationship,
- influence,
- direction of influence,
- **causality**.

What is hypothesis?

Stages of Social Research
A flowchart
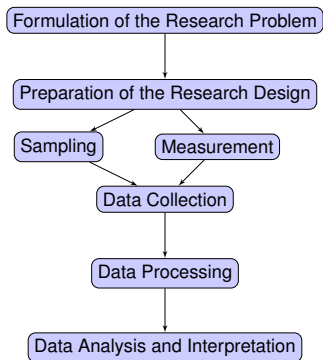
```
Formulation of the Research Problem
                │
                ▼
Preparation of the Research Design
          │            │
          ▼            ▼
      Sampling     Measurement
          │            │
          └────┬───────┘
               ▼
        Data Collection
               │
               ▼
        Data Processing
               │
               ▼
  Data Analysis and Interpretation
```

It was a pleasure!

Daróczi Gergely
daroczi.gergely@btk.ppke.hu