

# Quantitative methods

## Lesson 3

Daróczi Gergely

Corvinus University of Budapest, Hungary

2011 February 22



# Outline

- 1 Replay!
  - Types of research topics
  - Ecological fallacy
  - Types of variables and attributes
- 2 Relation between variables
  - Visual examples
  - Exact types
  - Further examples
- 3 Stages of Social Research
- 4 Preparation of Research Design
  - Conceptualization
  - Operationalization
- 5 Levels of Measurement
  - Examples
- 6 Final examination questions

# Types of research topics

Definition, examples

## **Define the type of the following topics:**

- Homeless people living at Budapest
- Reintegrating homeless people (?)
- Popular books in Hungary
- Living costs in London
- Why is it so freaking expensive to live in London?

**What is (the difference between) exploratory, descriptive and explanatory studies?**

# 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.

## Replay!

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



# Types of variables and attributes

Towards methodology

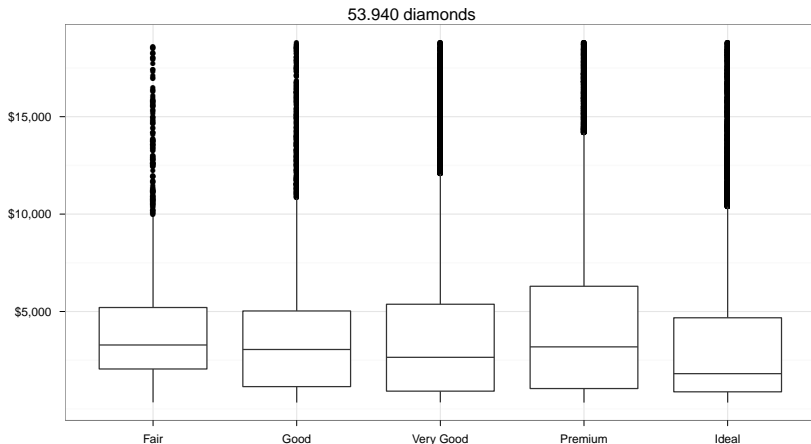
## Types of variables

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

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

# Relation between variables

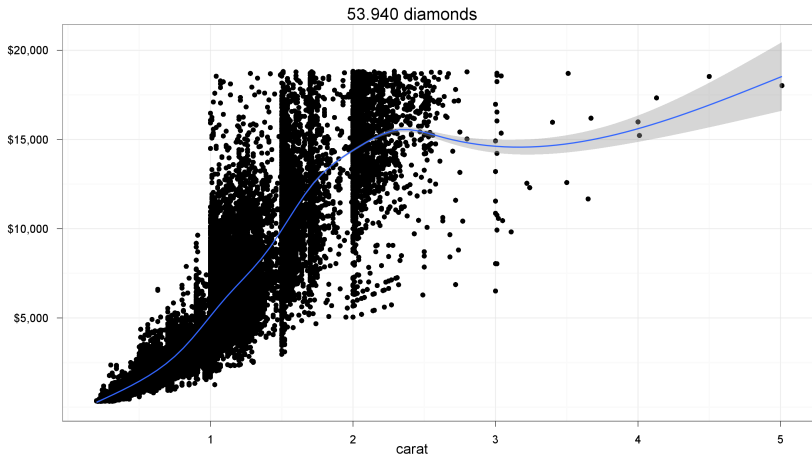
## A visual example



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

# Relation between variables

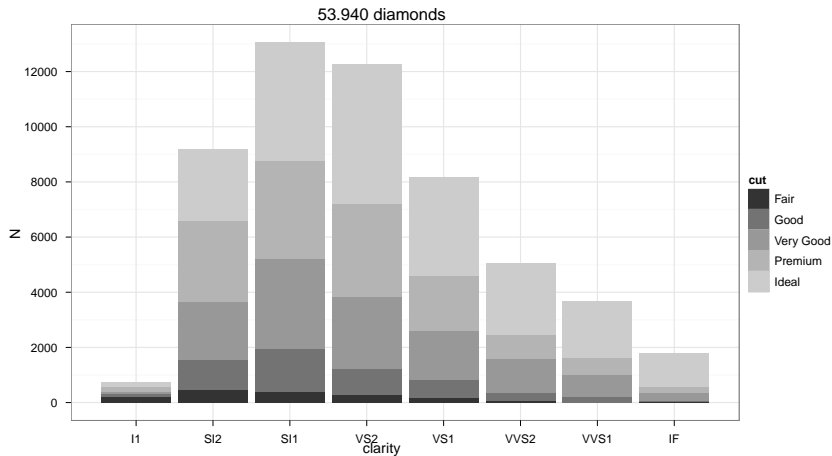
## A visual example



```
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

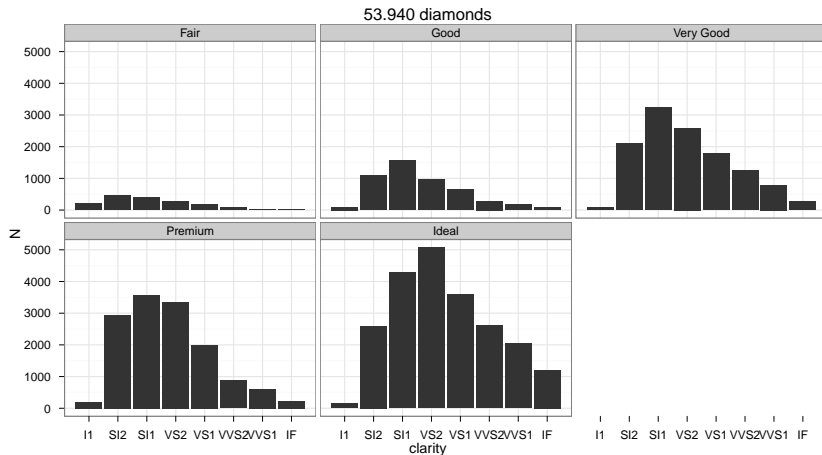


```
ggplot(diamonds, aes(clarity, fill=cut)) + geom_bar() + ylab("N") +  
theme_bw() + opts(title="53.940 diamonds")
```



# Relation between variables

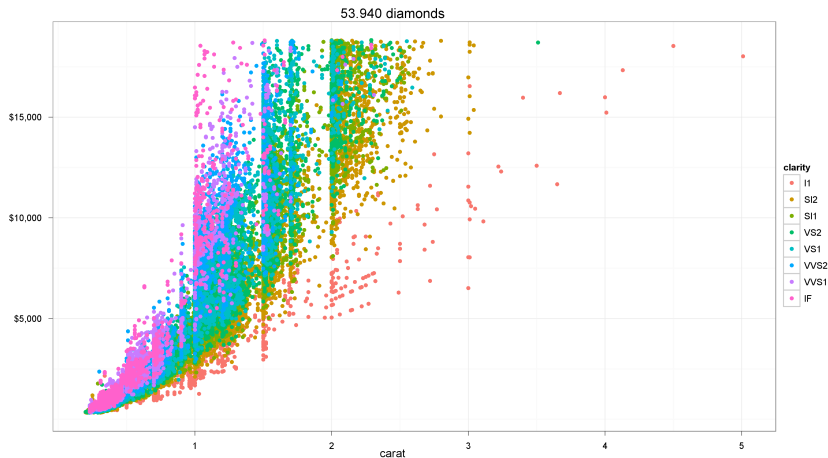
## A visual example



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

# Relation between variables

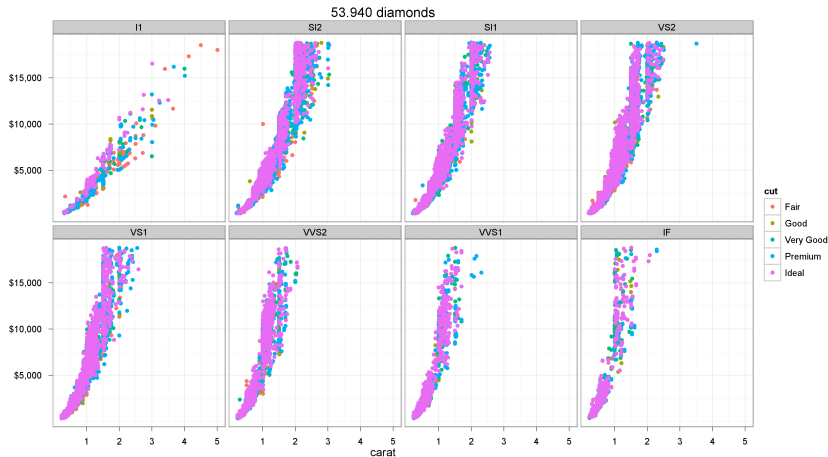
## A visual example



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

# Relation between variables

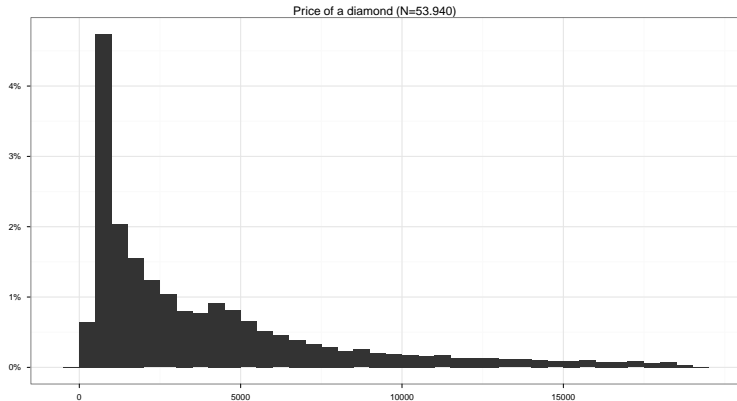
## A visual example



```
ggplot(diamonds, aes(carat, price, color=cut)) + geom_point() + ylab('') + facet_wrap(~ clarity, nrow=2) +  
scale_y_continuous(formatter="dollar") + theme_bw() + opts(title="53.940 diamonds")
```

# Test your knowledge!

## Reliability and validity



A survey was taken place about diamonds available for sale on the Internet. The respondents gave the above prices.

**What do you think of the reliability and validity of this research?**

# Types of variables and attributes in practice

## Relationship between variables

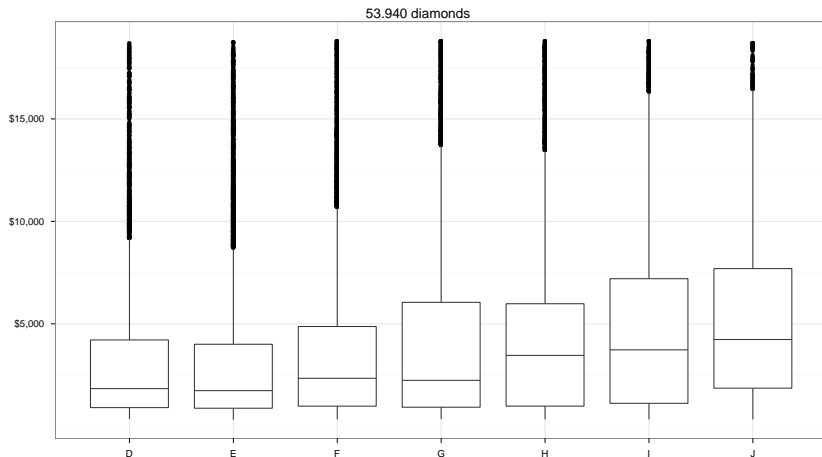
### Possible relationship between variables:

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

### STATISTICALLY SIGNIFICANT

# Relation between variables

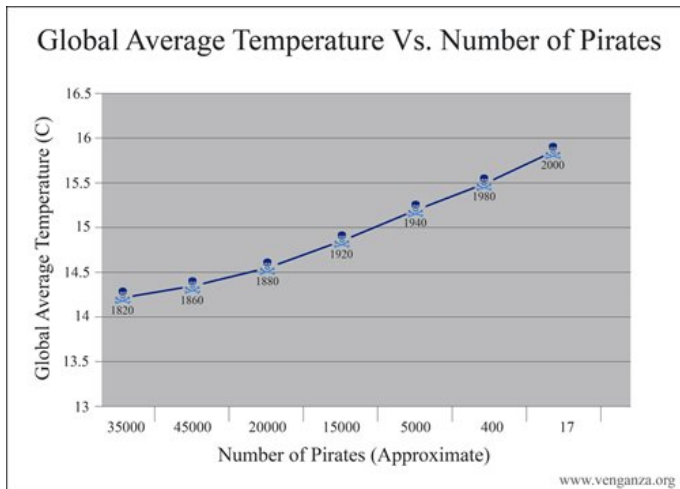
## Direction of influence



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

# Relation between variables

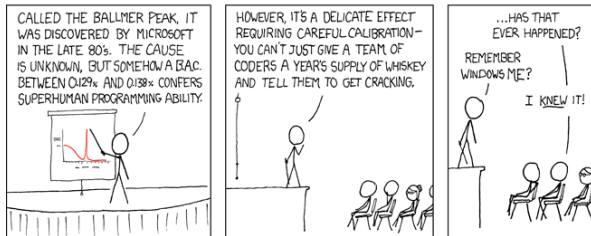
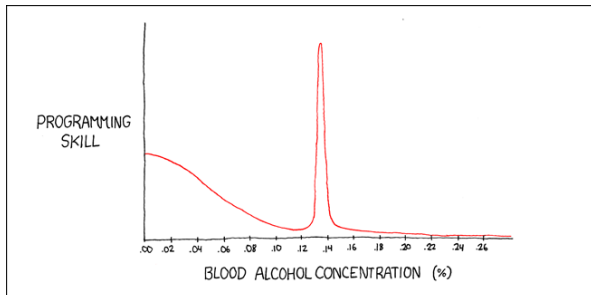
High correlation



A high correlation can be pointed out. So what?

# Relation between variables

No correlation

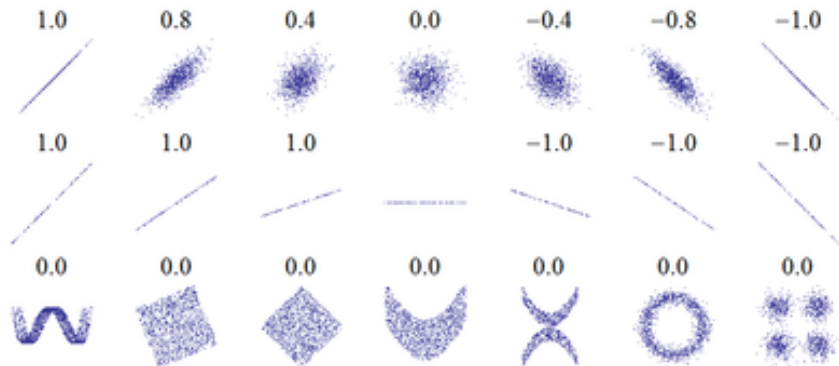


Source: <http://xkcd.com/323/>



# Relation between variables

## Correlations

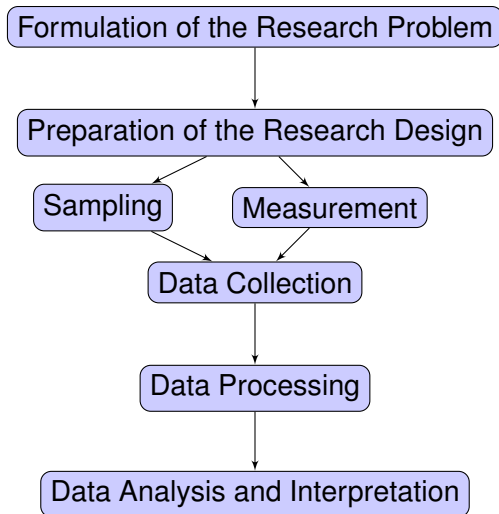


Positive (direct:  $R = 1$ ), negative (inverse:  $R = -1$ ), linear, curvilinear and uncorrelated ( $R = 0$ ) relationships

$R$ : correlation coefficient

# Stages of Social Research

A flowchart





# Conceptualization

A catchy example

## **Let us make a research about Friendship!**

- „Friends have all things in common.” (Plato)
- „Misfortune shows those who are not really friends.” (Aristotle)
- „What is a friend? A single soul in two bodies.” (Aristotle)
- „A friend to all is a friend to none.” (Aristotle)
- „One loyal friend is worth ten thousand relatives.” (Euripides)
- „My best friend is the one who brings out the best in me.” (H. Ford)
- „In a friend you find a second self.” (Isabelle Norton)
- „A friend should be a master at guessing and keeping still.”  
(Nietzsche)

# Operationalization

## A catchy example (continued)

Do you have a best friend?

- Yes, I have one or two best friends with whom I share almost everything.
- Yes, I have several friends whom I consider to be my best friend.
- No, I don't have a best friend.

Why do we need a friend?

- We need someone to confide into.
- We need someone who can listen to all our tantrums.
- We need someone with whom we can have fun.
- All of the above.
- We don't really need friends.

Source: <http://www.samplequestionnaire.com/mcgill-friendship-questionnaire.html>

# Levels of Measurement

Qualitative and quantitative variables in depth

## Qualitative variables:

- Nominal: exhaustive labels with no intersect (mutual exclusivity) 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

# Levels of Measurement

## Examples

**Determine the level of measurement of the following variables!**

- Gender
- Education
- Salary
- IQ
- Scholastic record
- Place of birth
- Favorite color

# Final examination questions

## Comprehensive exam

Singleton, R. A. Jr. and Bruce C. Straits (1999): *Approaches to Social Research*. Third Edition. Oxford University Press: New York/Oxford.

### Questions:

- 1 What is reliability? How do the main rules concerning the order of survey questions improve the reliability and validity of survey data? (pp. 113-117, 292-296)
- 2 What is meant by probability sampling? How do stratification and multistage cluster sampling affect sampling errors? Why? (pp. 141-142, 145-156)
- 3 What are the main types of non-probability sampling? Explain why these types do not meet the criteria of probability samples. (pp. 157-169)
- 4 What factors affect the desired sample size? (pp. 163-169)



It was a pleasure!

Daróczi Gergely  
*daroczi.gergely@btk.ppke.hu*