Quantitative methods

Week #3

Gergely Daróczi

Corvinus University of Budapest, Hungary

24 February 2012

Outline

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?

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!

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

Levels of Measurement

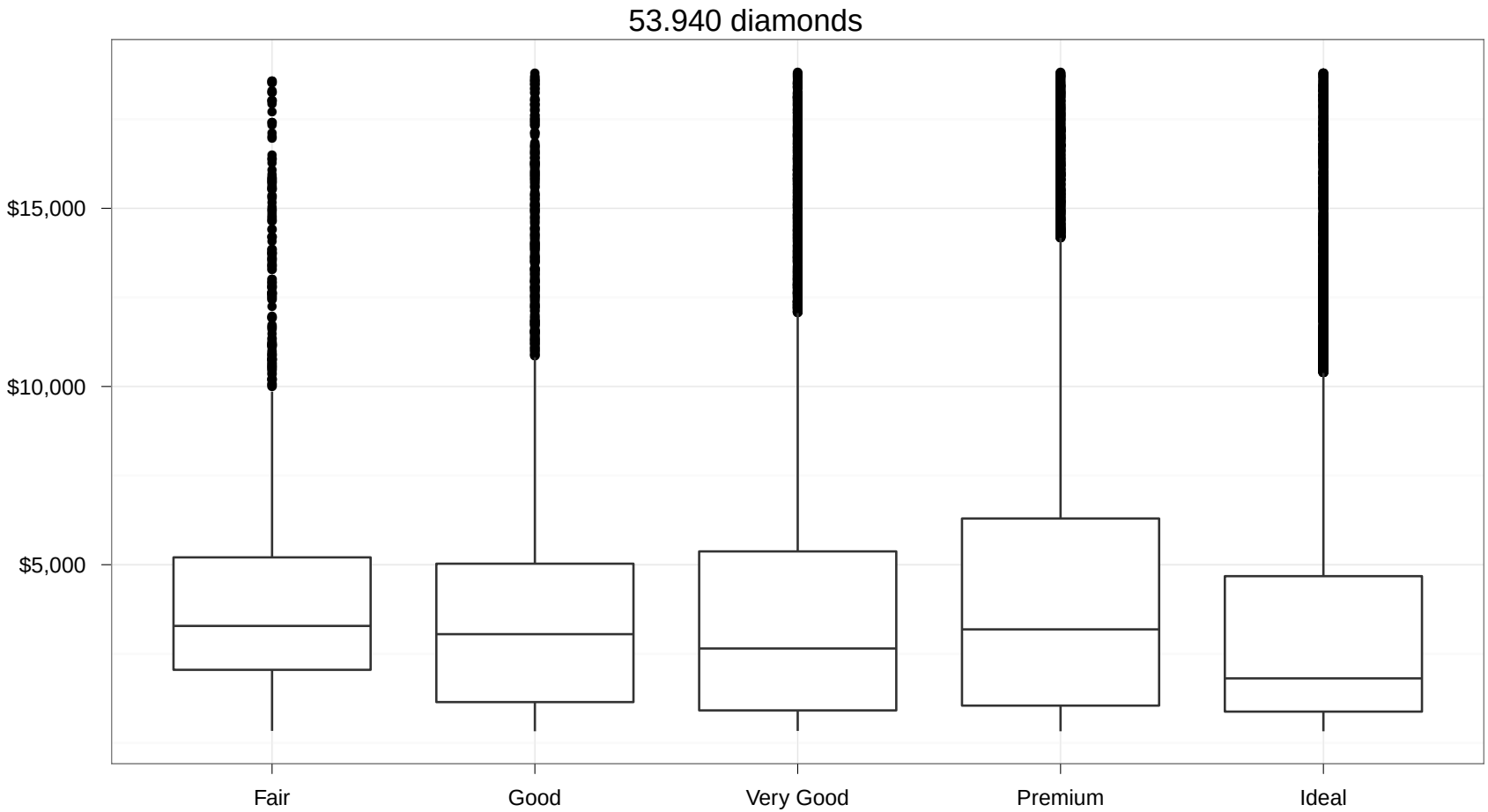
Examples

Determine the level of measurement of the following variables!

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

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")
```

The structure of the demo dataset

ggplot2/diamonds

Prices of 50,000 round cut diamonds

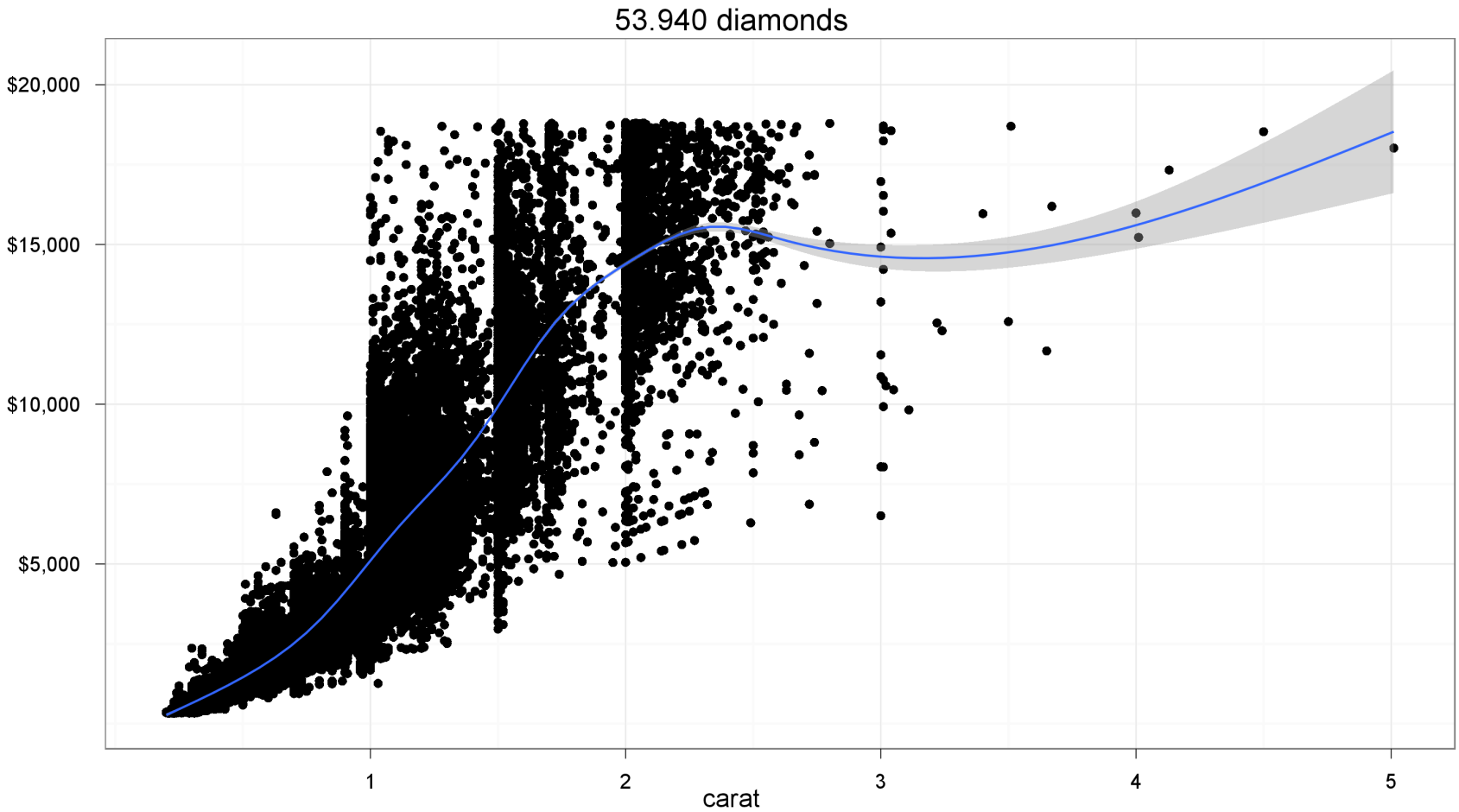
Description:

A dataset containing the prices and other attributes of almost 54,000 diamonds. The variables are as follows:

- price. price in US dollars (\$326--\$18,823)
- carat. weight of the diamond (0.2--5.01)
- cut. quality of the cut (Fair, Good, Very Good, Premium, Ideal)
- colour. diamond colour, from J (worst) to D (best)
- clarity. a measurement of how clear the diamond is (I1 (worst), SI1, SI2, VS1, VS2, VVS1, VVS2, IF (best))
- x. length in mm (0--10.74)
- y. width in mm (0--58.9)
- z. depth in mm (0--31.8)
- depth. total depth percentage = z / mean(x, y) = 2 * z / (x + y) (43--79)
- table. width of top of diamond relative to widest point (43--95)

Relation between variables

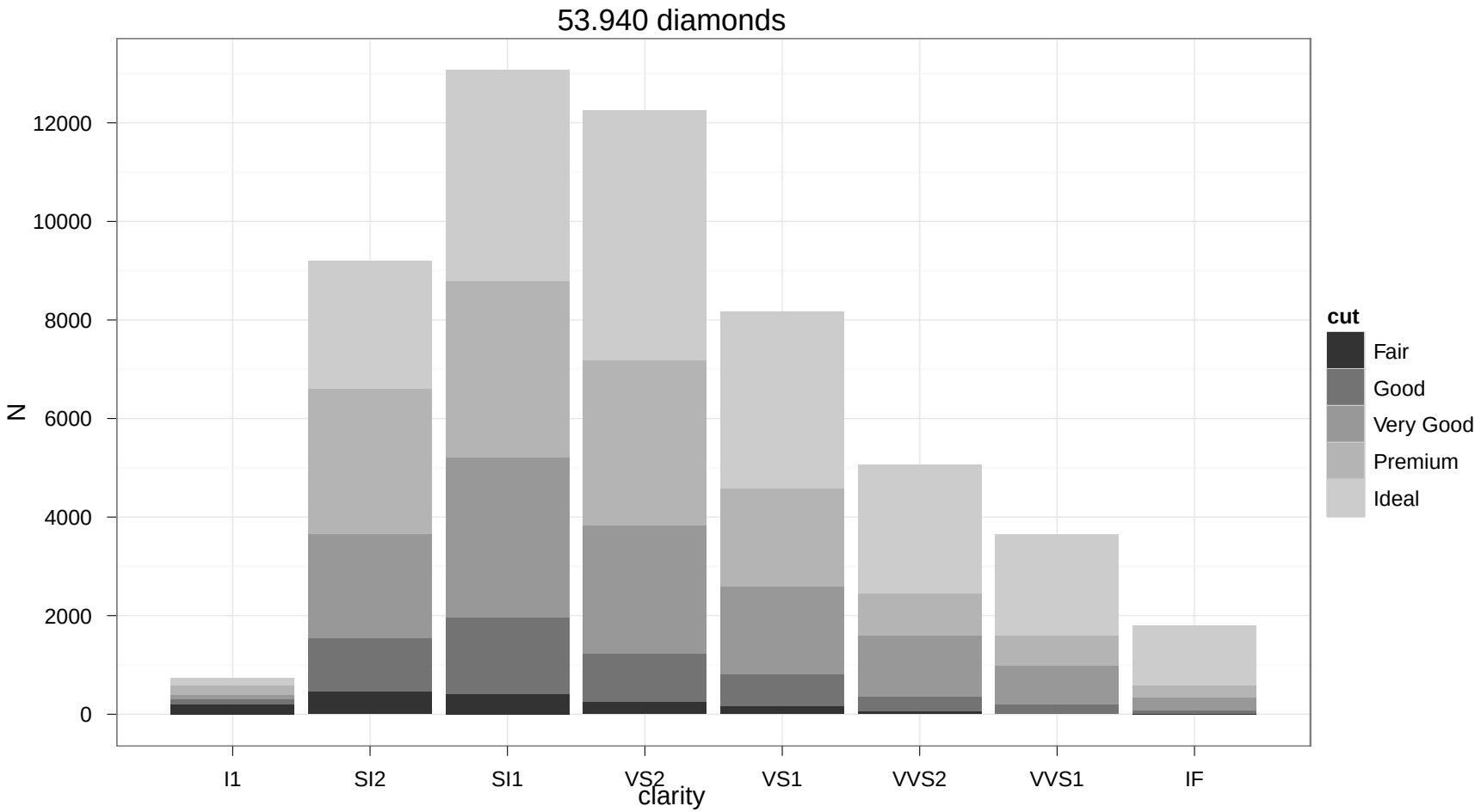
A visual example




```
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

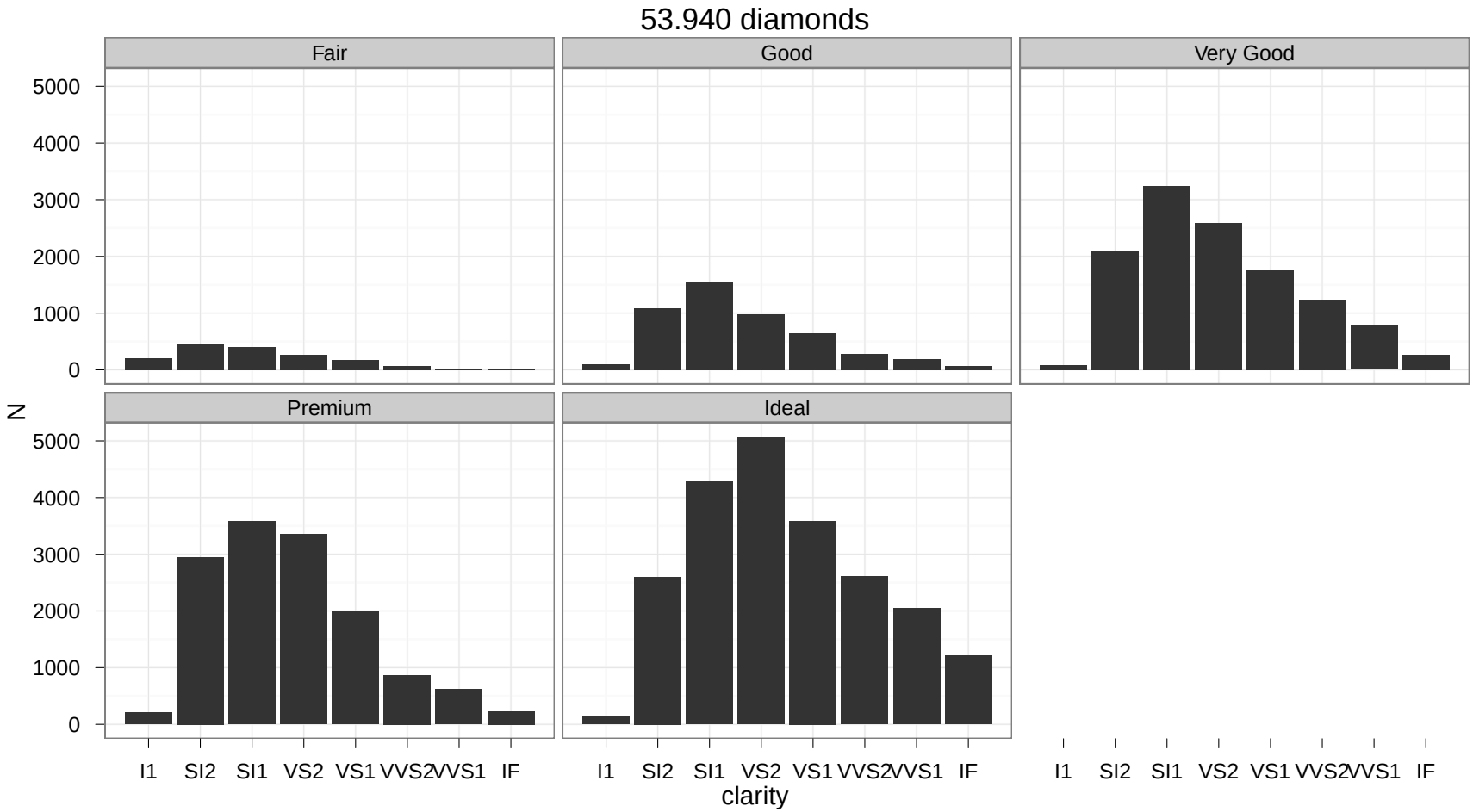
A visual example




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

Relation between variables

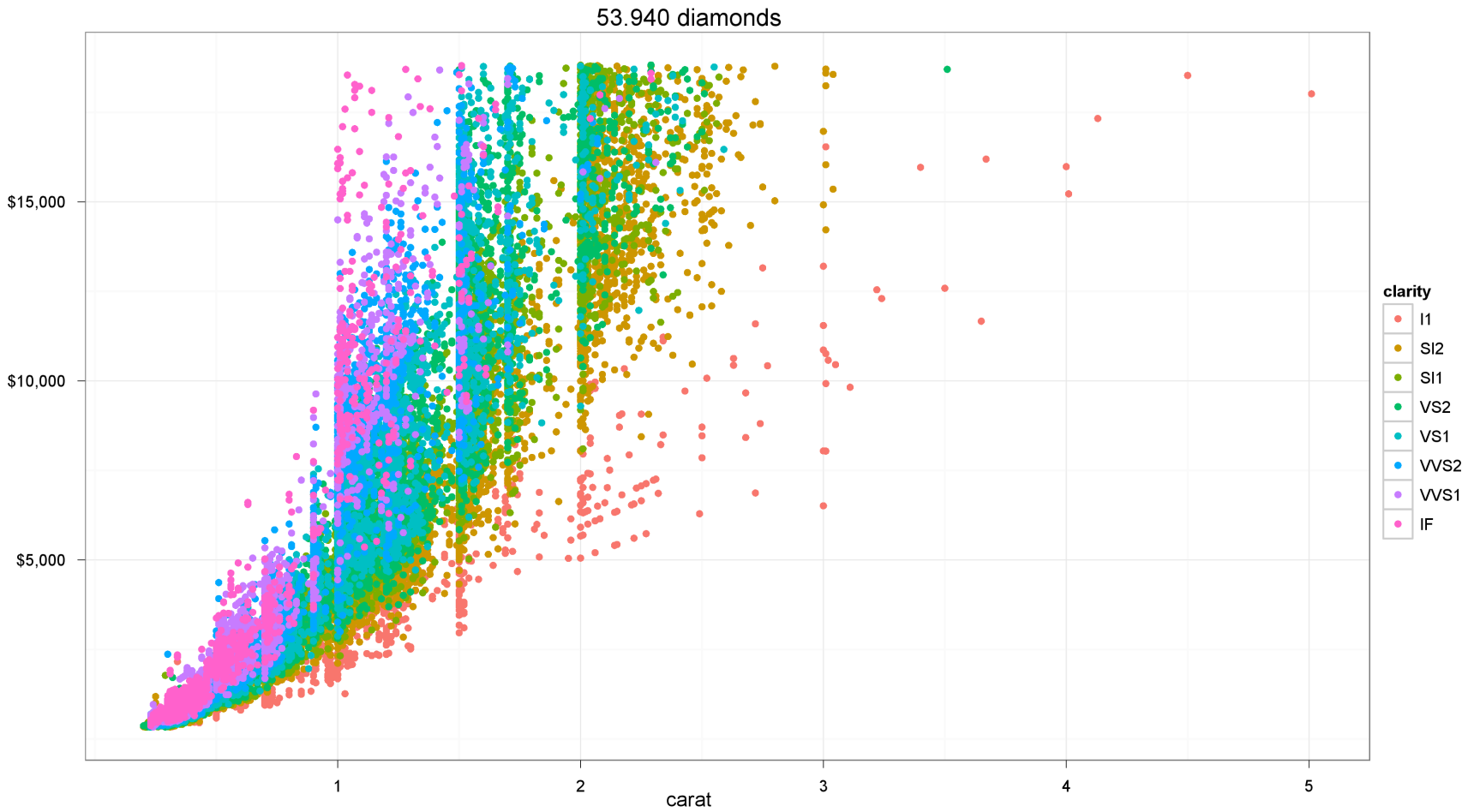
A visual example




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

Relation between variables

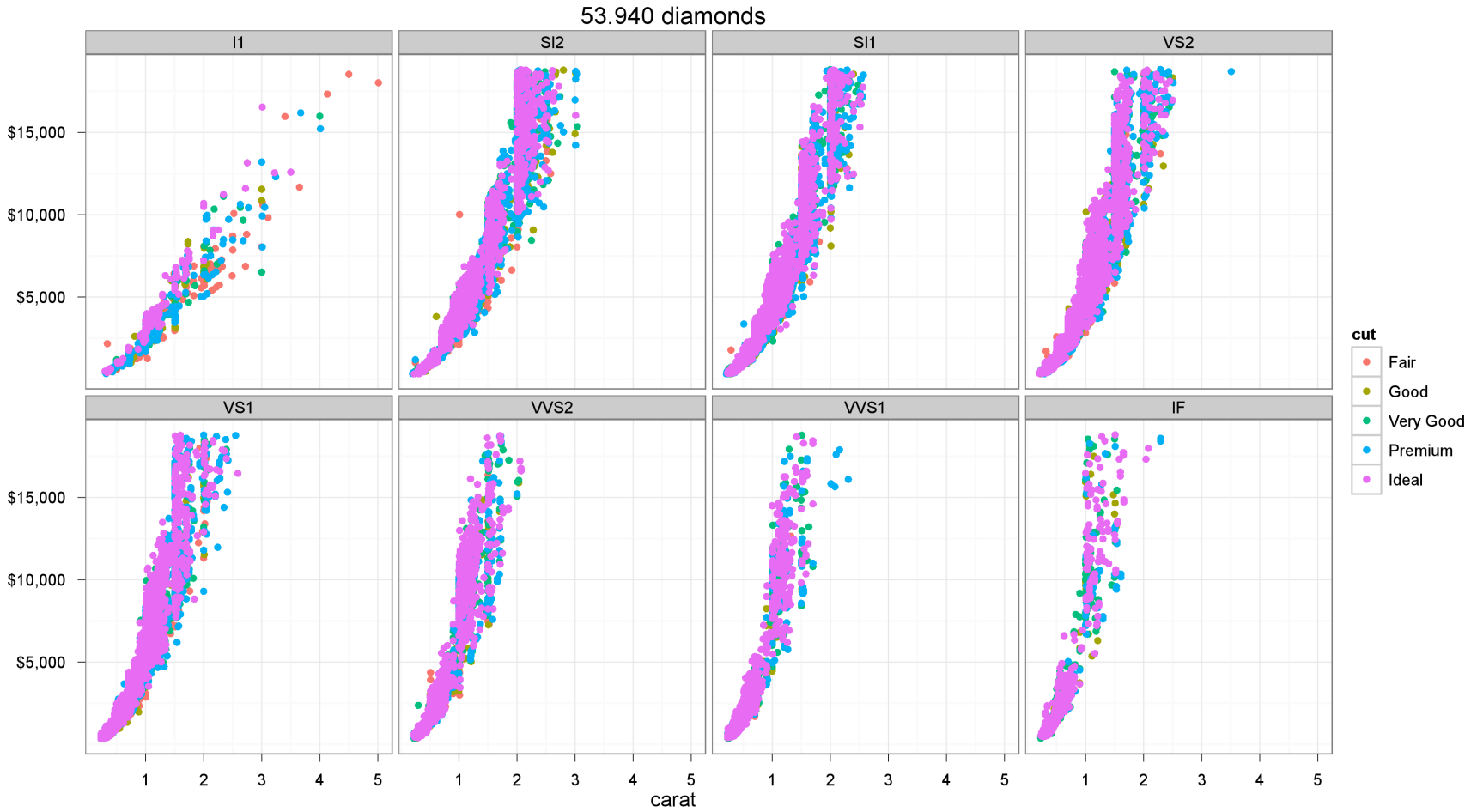
A visual example




```
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

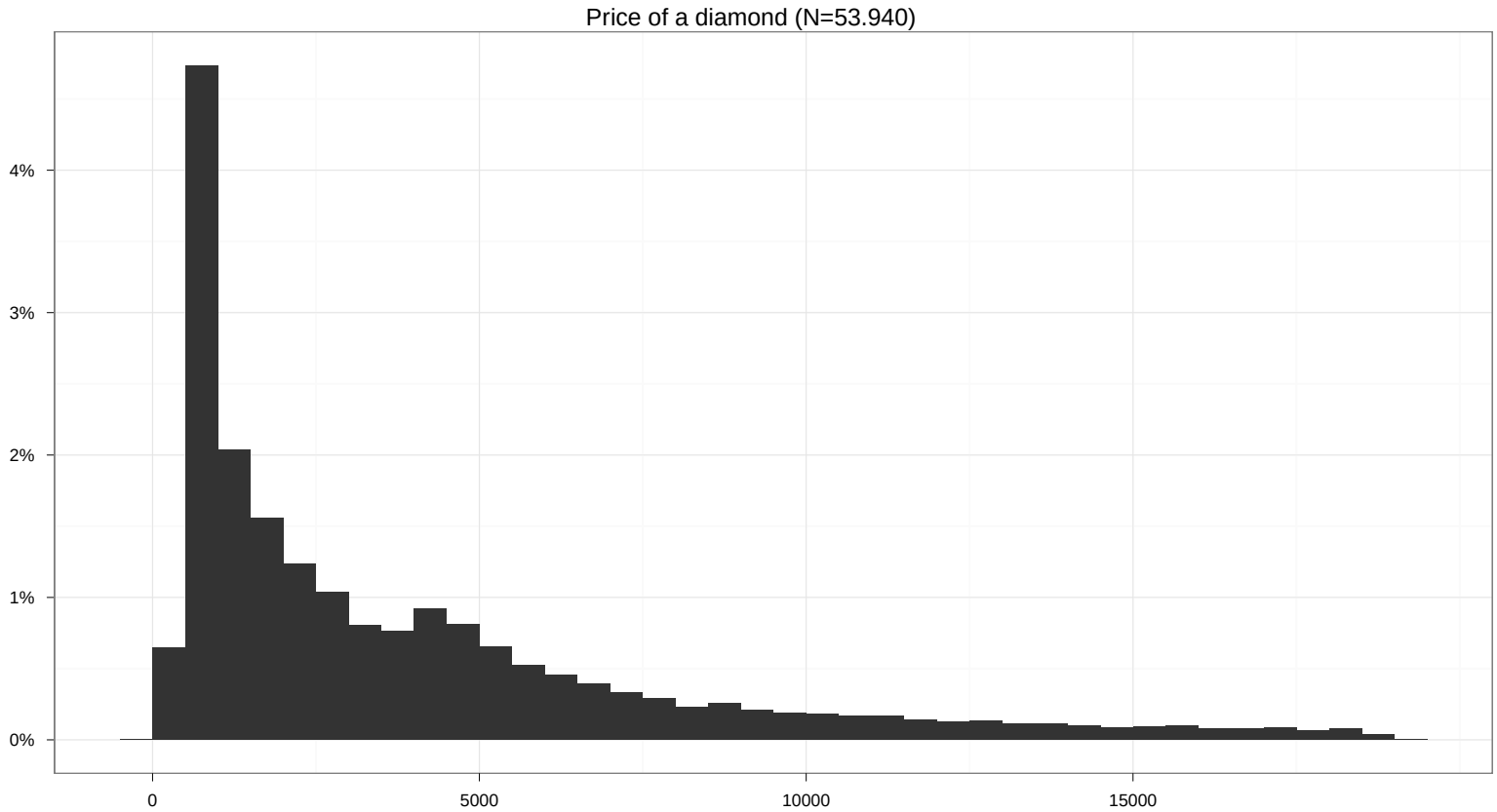
A visual example




```
ggplot(diamonds, aes(carat, price, color=cut)) + geom_point() + ylab('') + facet_wrap(~ clarity,nr
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.

What do you think of the realiability and validity of this research?

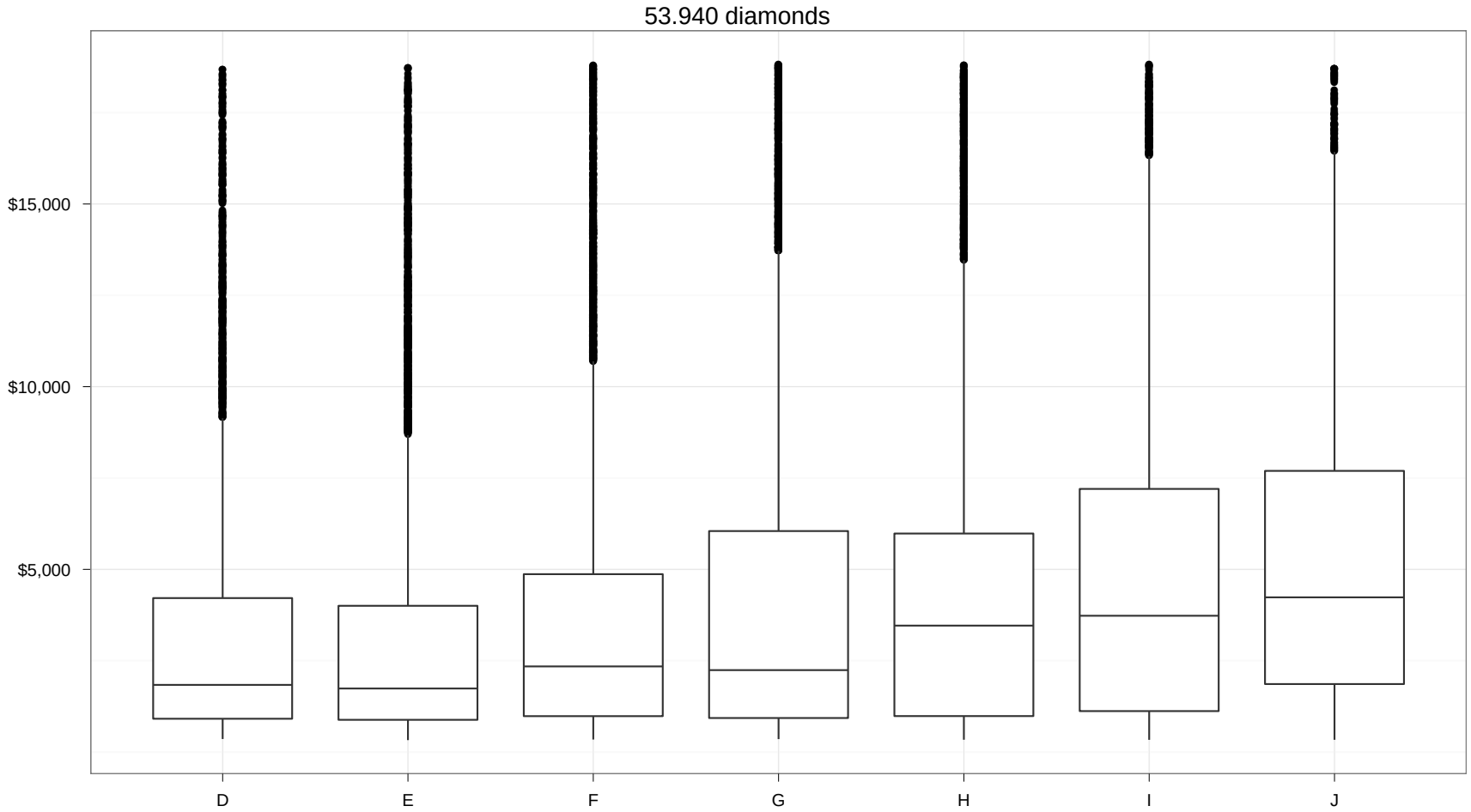
Possible relationship between variables:

- 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")
```

Cheatsheat

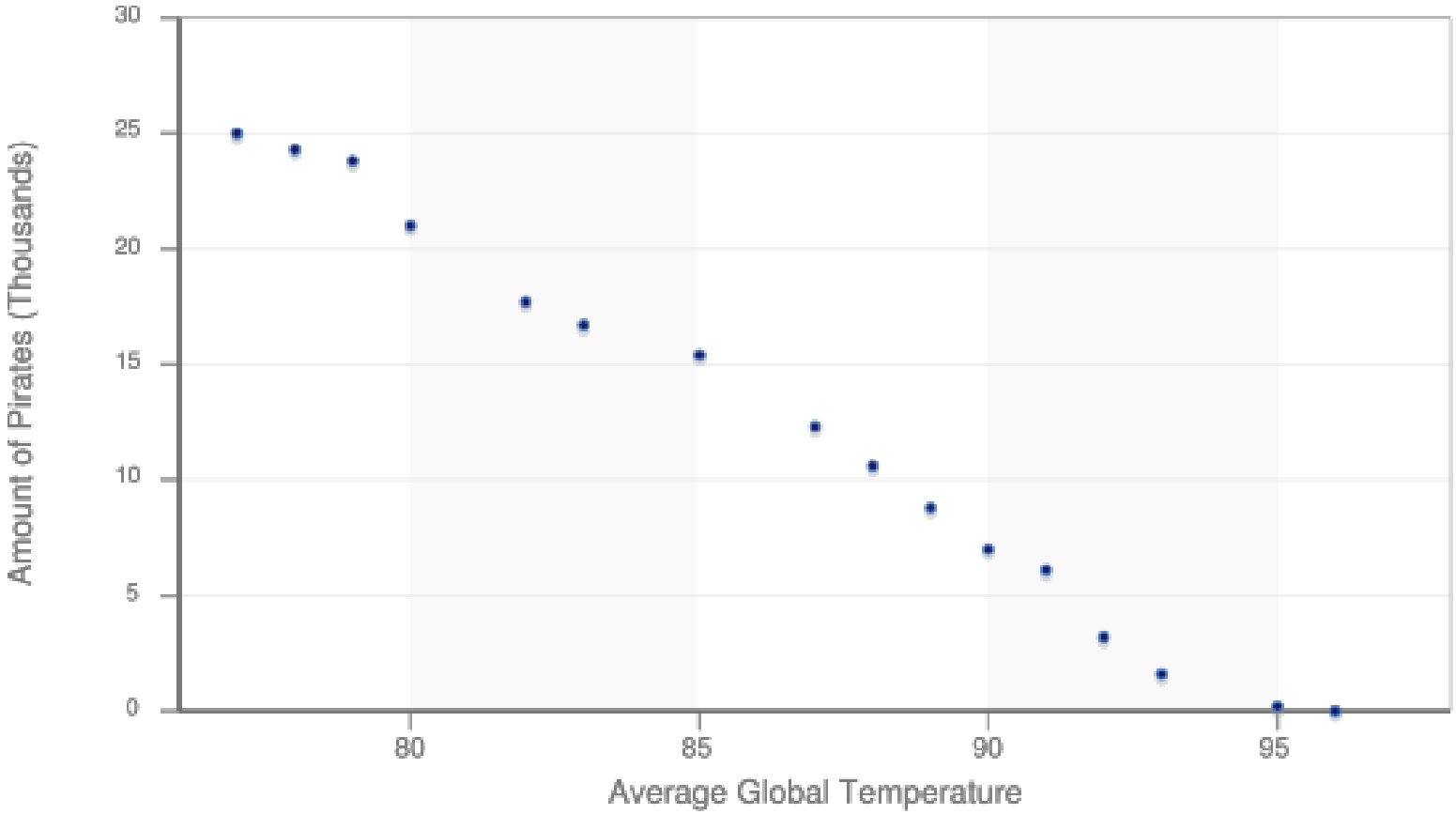
Diamonds' colors

GIA	Status: *current*	AGS	Status: *current*	AGS	Status: *historical: pre 1995*	CIBJO	Status: *current*	IDC	Status: *current*	Scan. D.N.	Status: *current*	Old World Terms	Status: *historical*
grade and description[6]		*grade and electronic colorimeter scale*[7]		*grade and electronic colorimeter scale*[7]		*grade*[8]		*grade and description*[8]		*grade for .50ct and over*[9]	*grade for under .50ct*	*series 1 scale*[8]	*series 2 scale*[8]
D	*Colorless*	0	0–0.49	0	0–0.75	Exceptional white +	Exceptional white +	*Colorless*		River	White	Finest White	Jager
E		0.5	0.5–0.99	1	0.76–1.35	Exceptional white	Exceptional white						River
F		1.0	1.0–1.49	2	1.36–2.00	Rare white +	Rare white +	*Colorless when viewed through the crown*		Top Wesselton		Fine White	
G	*Near Colorless*	1.5	1.5–1.99			Rare white	Rare white						Top Wesselton
H		2.0	2.0–2.49	3	2.01–2.50	White	White			Wesselton		White	Wesselton
I		2.5	2.5–2.99	4	2.51–3.0	Slightly tinted white	Slightly tinted white	*Slightly colored*		Top Crystal	Slightly tinted white	Commercial White	Top Crystal
J		3.0	3.0–3.49	5	3.01–3.75					Crystal		Top silver cape	Crystal
K	*Faint Yellow*	3.5	3.5–3.99	6	3.76–4.5	Tinted white	Tinted white			Top cape	Tinted white	Silver cape	Top cape
L		4.0	4.0–4.49										
M		4.5	4.5–4.99	7	4.51–5.50	Tinted color 1	Tinted color	*Slightly colored to colored*		Cape	Tinted color	Light cape	Cape
N	*Very Light Yellow*	5.0	5.0–5.49			Tinted color 2							Low Cape
O		5.5	5.5–5.99	8	5.51–7.0					Light yellow		Cape	Very light yellow
P		6.0	6.0–6.49										Light yellow
Q		6.5	6.5–6.99										
R		7.0	7.0–7.49	9	7.01–8.5							Dark cape	
S	*Light Yellow*	7.5	7.5–7.99			Tinted color 3				Yellow			
T		8.0	8.0–8.49										
U		8.5	8.5–8.99	10	8.51–10								
V		9.0	9.0–9.49										
W		9.5	9.5–9.99										
X		10	10+	10+									
Y													
Z													

Source: http://en.wikipedia.org/wiki/Diamond_color

Relation between variables

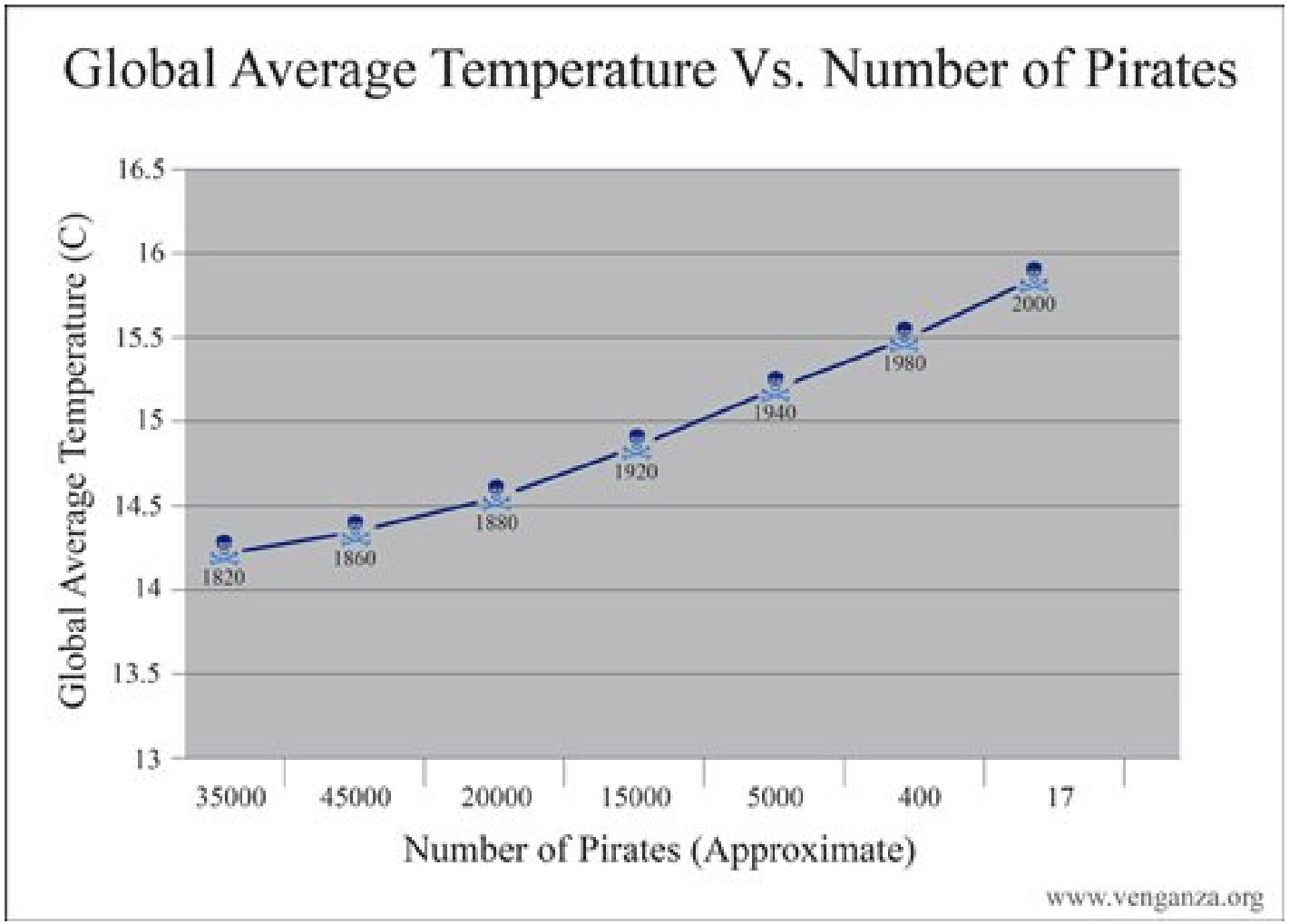
High correlation



A high correlation can be pointed out. Please explain!

Relation between variables

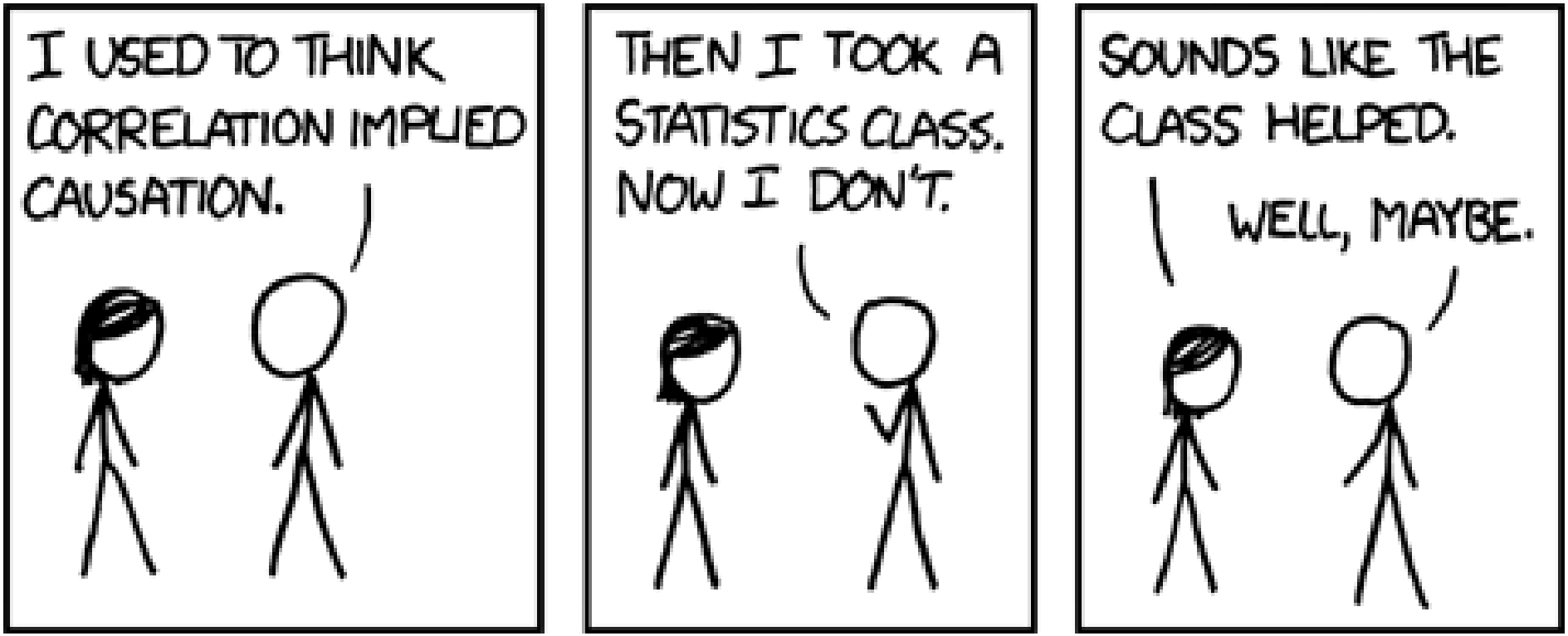
High correlation



A high correlation can be pointed out. So what?

Relation between variables

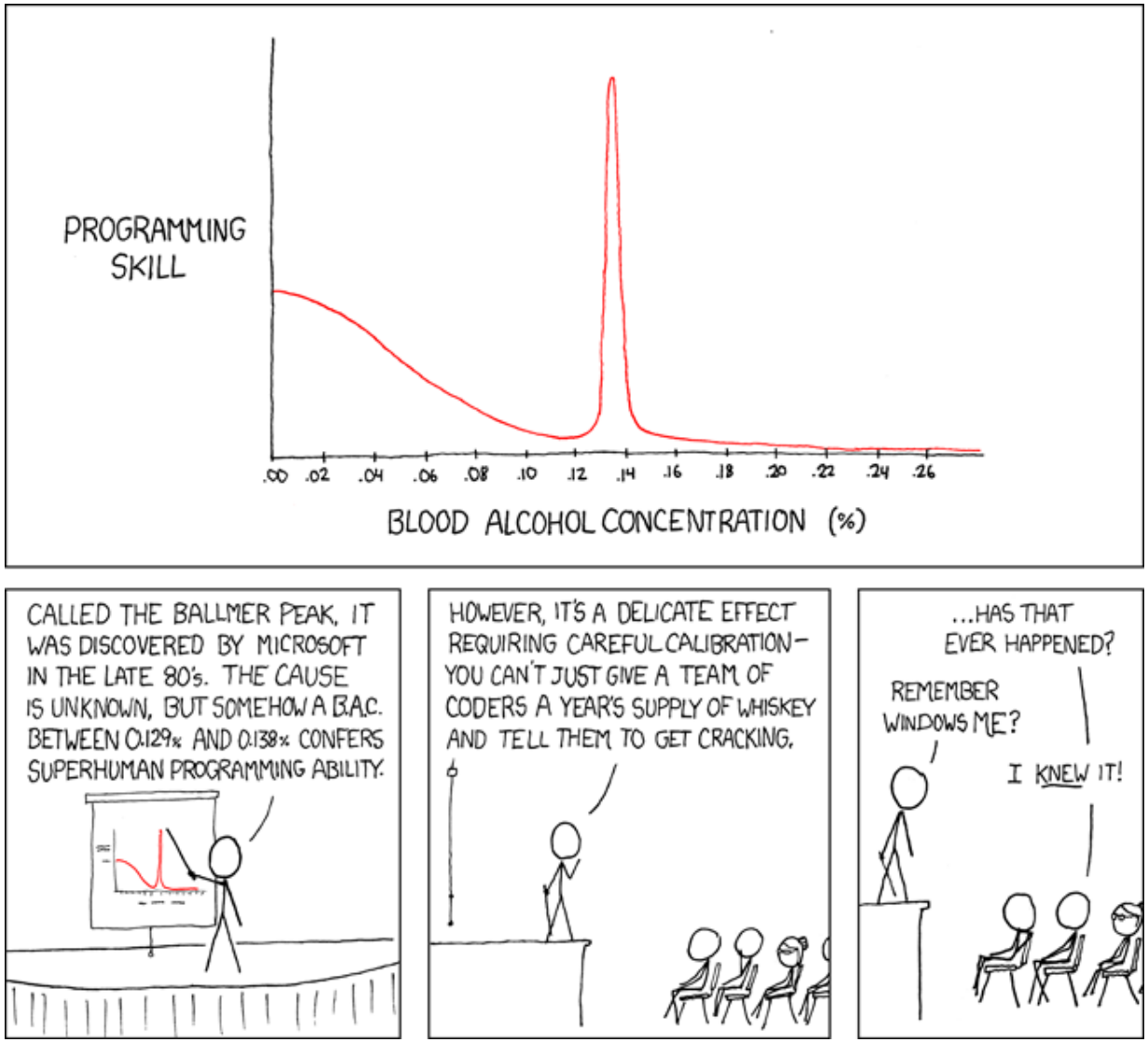
Correlation does not imply causality!



Source:http://xkcd.com/552/

Relation between variables

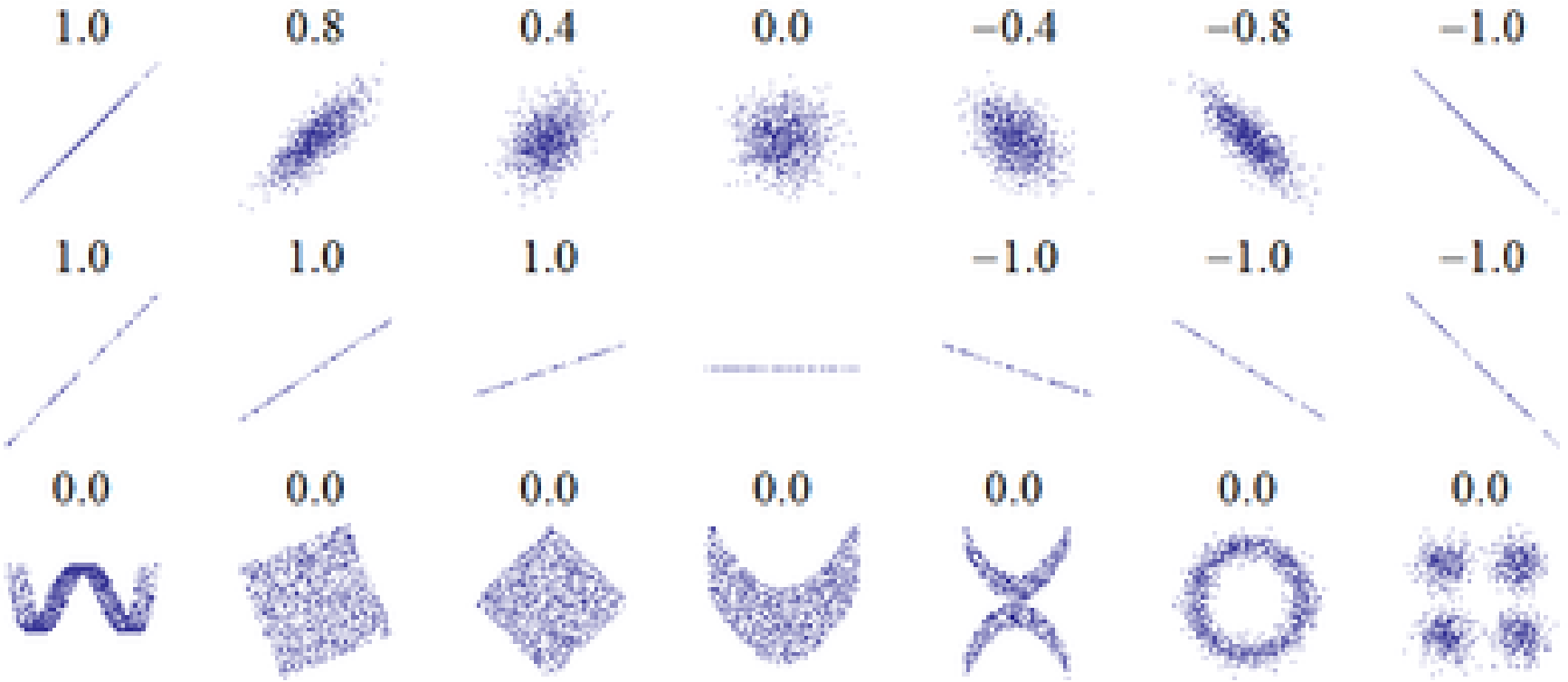
No correlation. No relationship?



Source: *http://xkcd.com/323/*

Relation between variables

Correlation coefficient



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

R: correlation coefficient

Conceptualization:

Definition

Conceptual definition is the process of formulating and clarifying concepts.

Operationalization:

Definition

Operational definition describes the research operations that will specify the value or category of a variable on each case.

Let us make a research about Friendship!

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 tobe my best friend.
- No, I don't have a best friend.

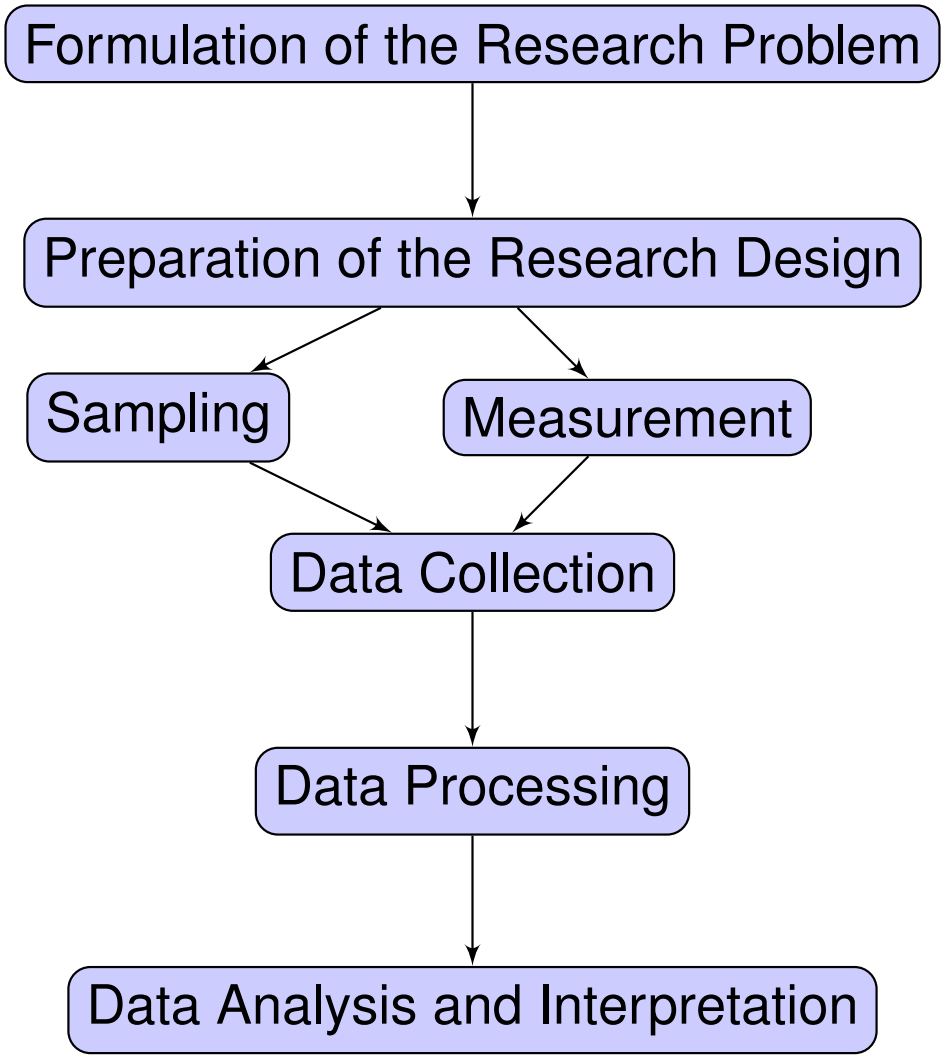
Why do we need a friend?

- 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

Stages of Social Research

A flowchart



It was a pleasure!

Daróczi Gergely
daroczi.gergely@btk.ppke.hu