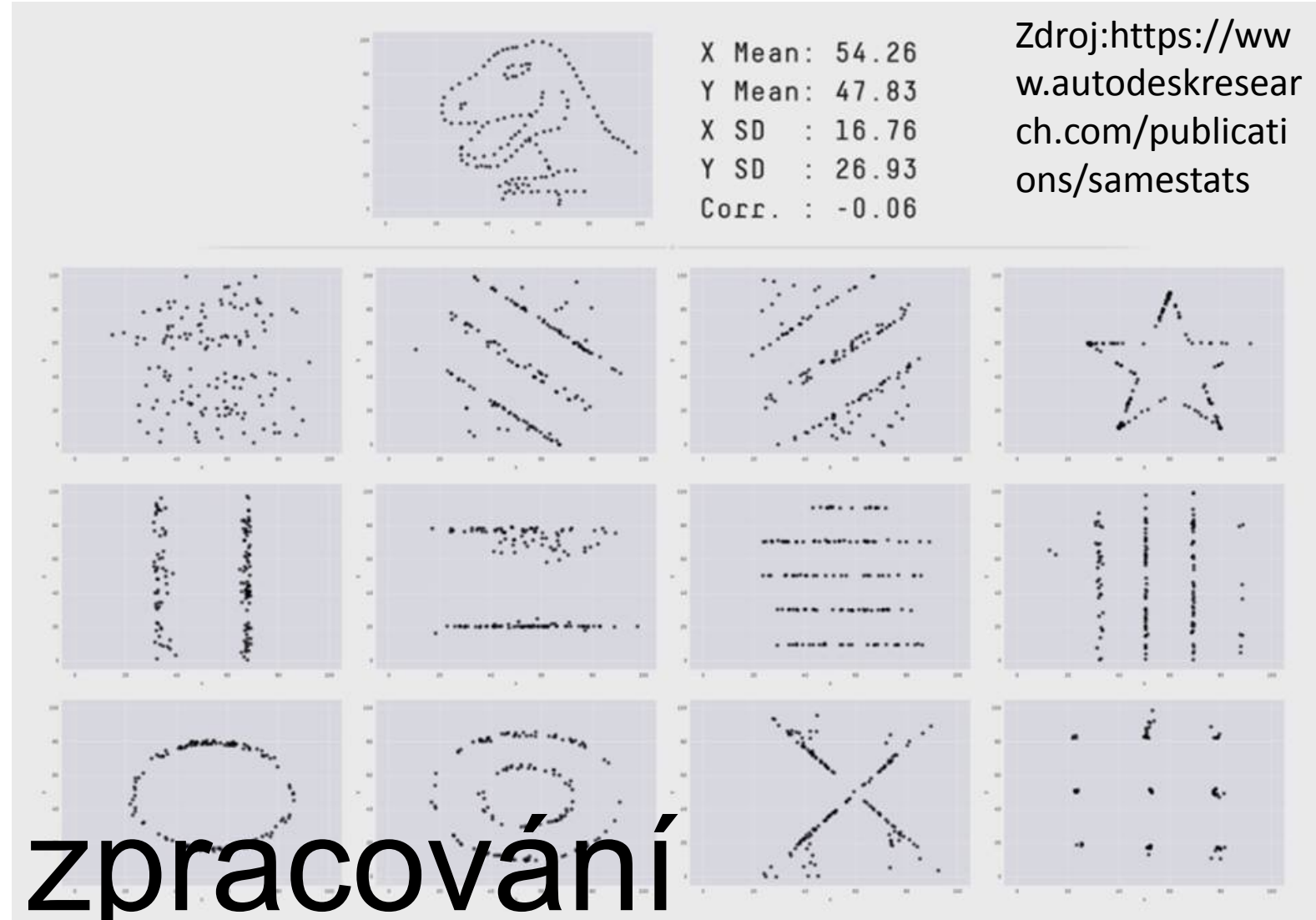


Zdroj: <https://www.autodeskresearch.com/publications/samestats>



Kvantitativní zpracování dat

Zhodnocení sběru dotazníků

Hlavní body

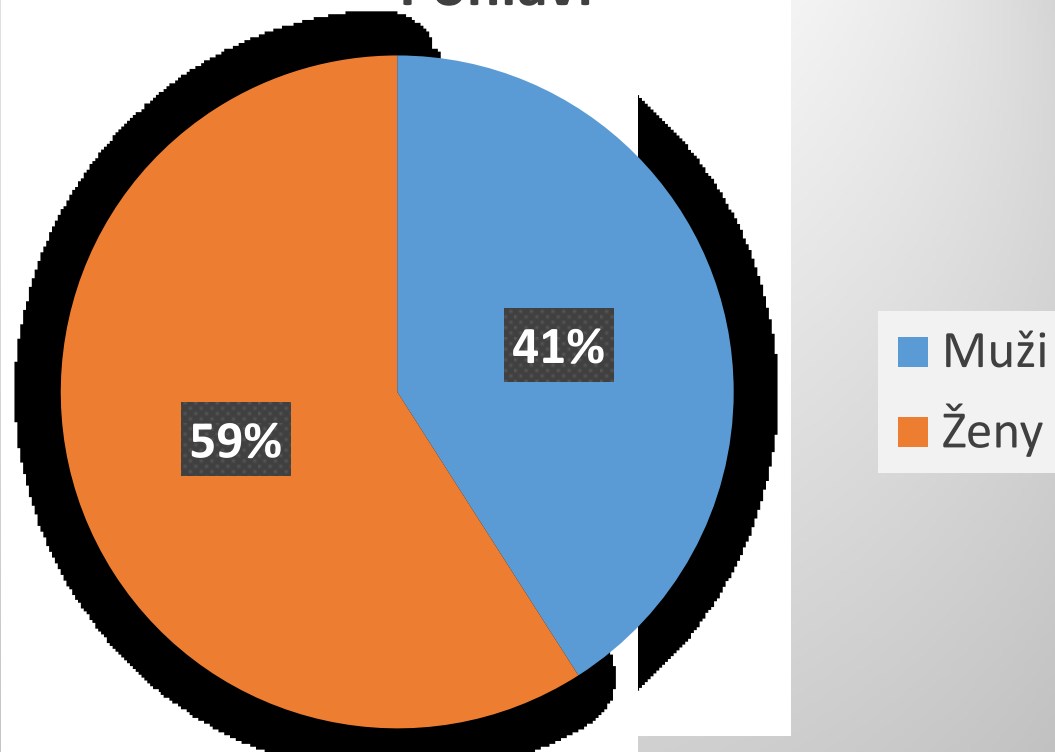
- Naučte se alespoň základy statistiky a jak číst statistické závěry
- Na většinu stačí Excel, maximálně R+ nebo SPSS
- Nechte si poradit od někoho kdo se vyzná
- **POL593 Kvantitativní přístupy v politologii**
 - <https://is.muni.cz/auth/predmet/fss/podzim2016/POL593>

Otázky:

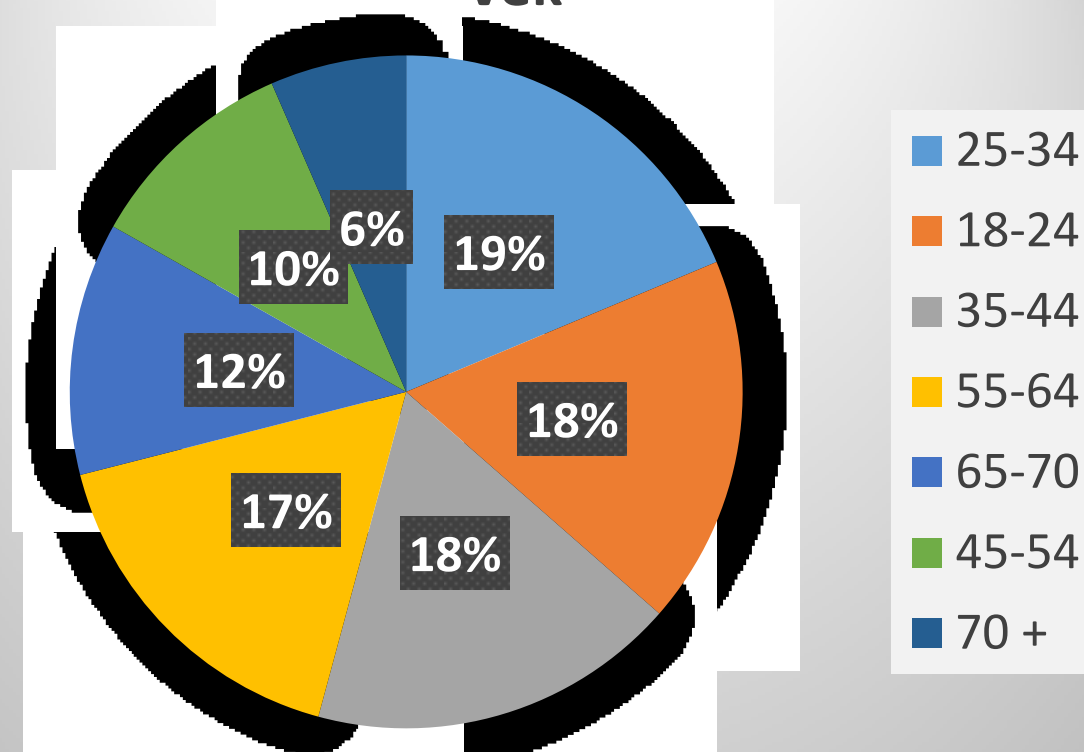
- Jaké druhy proměnných máme? (N, O, I(K))
- Co to znamená, že má něco pozitivní korelaci?

Koho jste oslovovali (četnosti)

Pohlaví

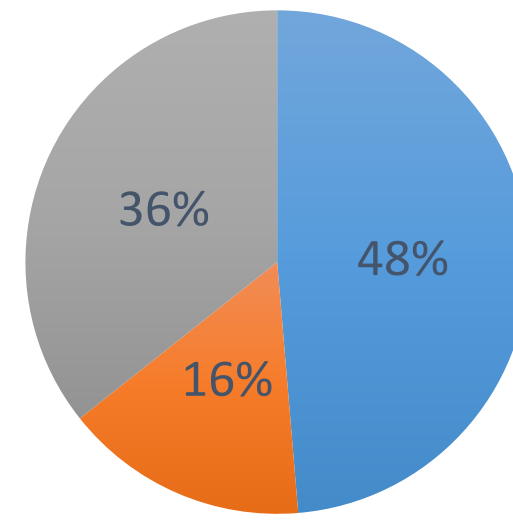
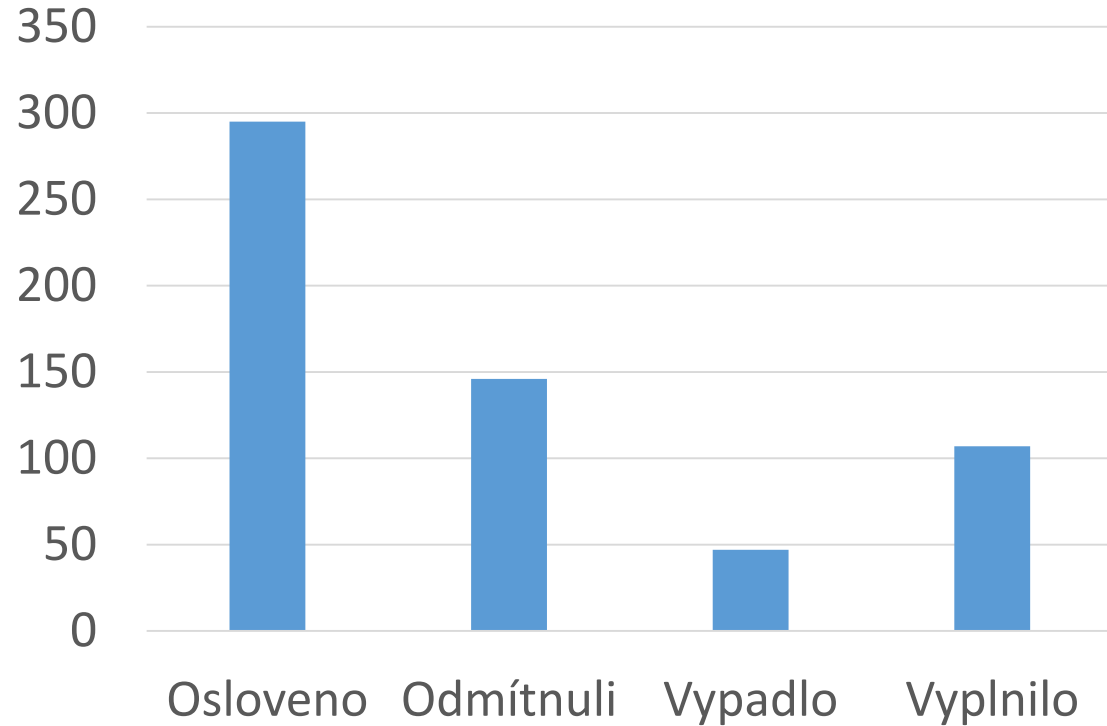


Věk

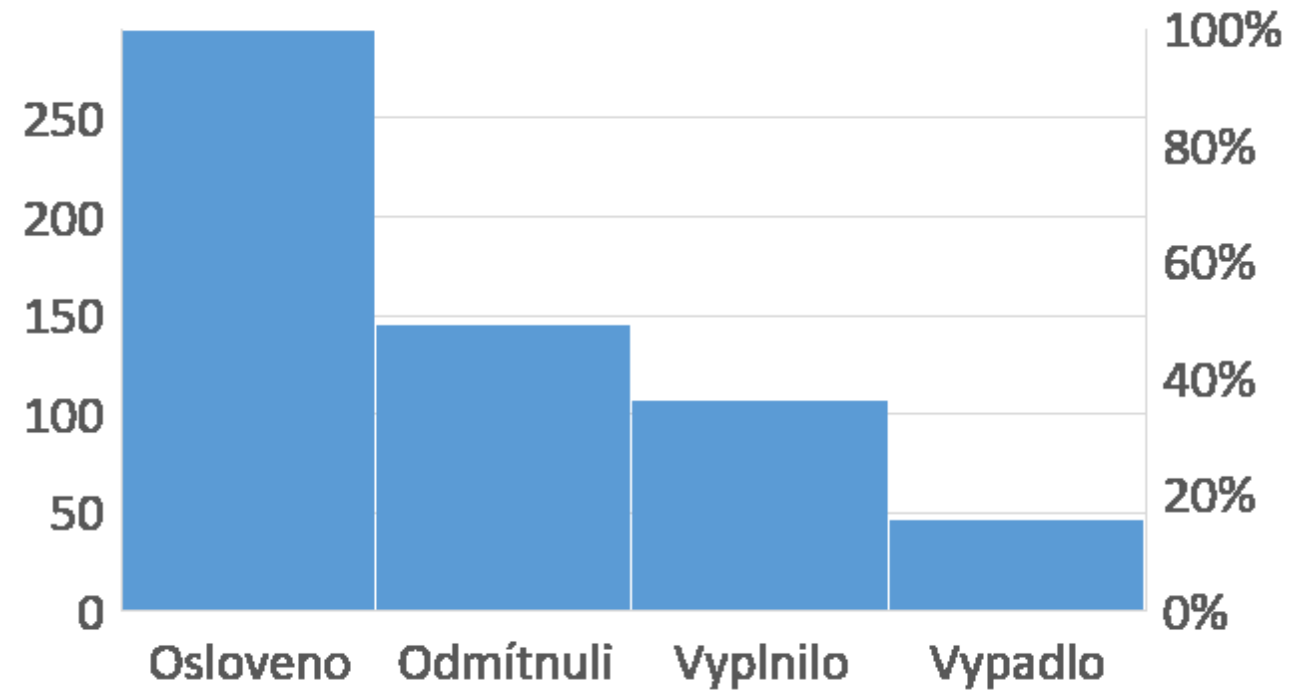


https://docs.google.com/spreadsheets/d/15gM525oTTeK3mNYET_Q_IKg7dQJB727PhKTWhT_FBaA/edit?usp=sharing

| | | |
|-----------|-----|---------|
| Osloveno | 295 | |
| Odmítnuli | 146 | 49,49 % |
| Vypadlo | 47 | 15,93 % |
| Vyplnilo | 107 | 36,27 % |



■ Odmítnuli ■ Vypadlo ■ Vyplnilo



Jak získáváme data?

BOX 16.1

Question formats requiring (a) single-transfer coding and (b) double-transfer coding

(a) How many children are there in your school?

| | | | | | | | |
|----------|-------|-------|-------|-------|-------|--------|----------|
| under 40 | 40-49 | 50-59 | 60-69 | 70-79 | 80-89 | 90-100 | over 100 |
| code 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

enter code ()

(b) How many children are there in your school?

(please circle)

| | | | | | | | |
|----------|-------|-------|-------|-------|-------|--------|----------|
| under 40 | 40-49 | 50-59 | 60-69 | 70-79 | 80-89 | 90-100 | over 100 |
|----------|-------|-------|-------|-------|-------|--------|----------|

(response has then to be translated into appropriate code)

| číslo skupiny | S1 | S1.1 | S2 | S3 | Q1 | Q1.1 | Q1.2 | Q1.3 | Q4.1 | Q4.2 | Q4.3 | Q4.4 | Q4.5 | Q5 | Q6 | Q7 |
|---------------|----|------|----|----|----|------|------|------|------|------|------|------|------|----|----|----|
| A2 | a | b | F | F | B | b | | | 4 | 5 | 3 | 3 | 5 | 5 | 4 | |
| A3 | a | b | E | F | B | a | B | 3 | 1 | 5 | 1 | 5 | 5 | 5 | 4 | |
| A4 | a | b | B | M | C | b | | | 4 | 5 | 4 | 4 | 5 | 1 | 1 | |
| A5 | a | b | E | F | B | b | | | 5 | 5 | 2 | 5 | 5 | 3 | 3 | |
| A6 | a | b | G | M | A | b | | | 2 | 5 | 2 | 5 | 5 | 3 | 1 | |
| A7 | a | b | A | F | B | b | | | 2 | 5 | 2 | 3 | 5 | 3 | 3 | |
| A8 | a | b | G | F | A | b | | | 1 | 1 | 1 | 5 | 5 | 3 | 5 | |
| A9 | a | b | A | M | D | b | | | 5 | 5 | 5 | 1 | 5 | 1 | 1 | |
| B1 | A | B | F | M | D | B | | | 5 | 5 | 2 | 5 | 5 | 4 | 1 | |
| B10 | A | B | F | M | D | B | | | 4 | 4 | 2 | 1 | 1 | 1 | 1 | |
| B2 | A | B | C | M | D | B | | | 5 | 5 | 5 | 5 | 5 | 4 | 4 | |
| B3 | A | B | G | F | A | A | D | 5 | 1 | 4 | 1 | 5 | 5 | 2 | 5 | |
| B4 | A | B | C | M | D | B | | | 5 | 5 | 5 | 5 | 5 | 1 | 1 | |
| B5 | A | B | C | M | B | B | | | 5 | 5 | 4 | 5 | 5 | 1 | 1 | |
| B6 | A | B | F | F | - | B | | | 5 | 5 | 4 | 1 | 5 | 3 | 4 | |
| B7 | A | B | E | F | D | A | B | 5 | 1 | 1 | 1 | 1 | 1 | 5 | 5 | |
| B8 | A | B | F | F | C | B | | | 5 | 5 | 4 | 5 | 5 | 1 | 1 | |

Chybí data?

- Co dělat když chybí data? Nejlepší nic nevymýšlet – prostě chybí.
- Začít se ptát proč nám chybí taková data, která potřebujeme
- Chybějící data nekódujeme „0“ – může dojít ke zkreslení (škála 1-100)
- Je rozdíl mezi:
 - Nevím
 - Ne
 - Nemůžu odpovědět
 - Nemůžu odpovědět v mezích předdefinovaných odpovědí

Vyčistění dat po zadání

- Jak předejít chybám v přepisu? Nejlépe když data zadávají dva výzkumníci a pak se provede rozdílový test
- Jak předejít „překlikům“ – frekvenční analýza jednotlivých proměnných
- Kontingenční tabulky
- Vyřadím všechny pochybná data (neuspokojivé odpovědi na kontrolní otázky, přerušení sběru dat, podezření z podvodu atd.)

Popisná statistika

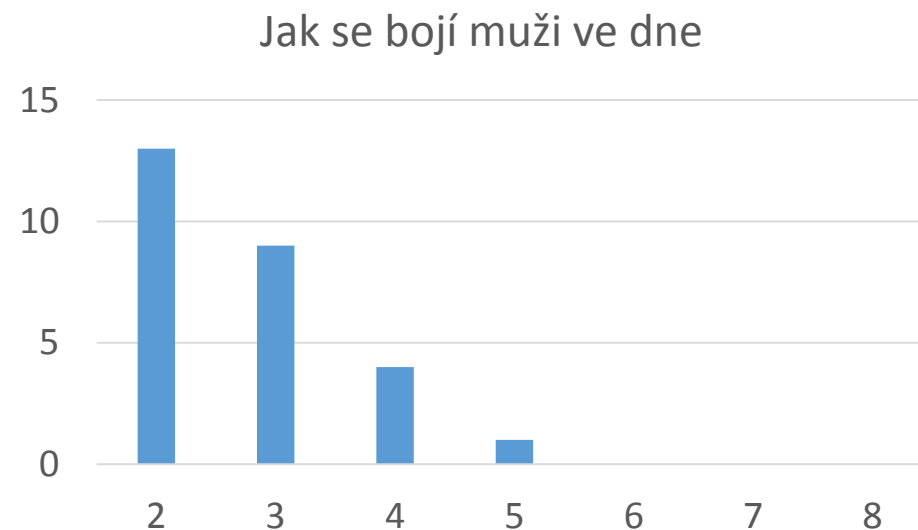
- Četnosti
- Centrální tendence
 - Modus (nominální, ordinální, kardinální)
 - Medián (ordinální, kardinální)
 - Aritmetický průměr (kardinální)

- Jak se bojí ženy ve dne?

- Aritmetický průměr – 4,46
- Medián - 4
- Modus - 4

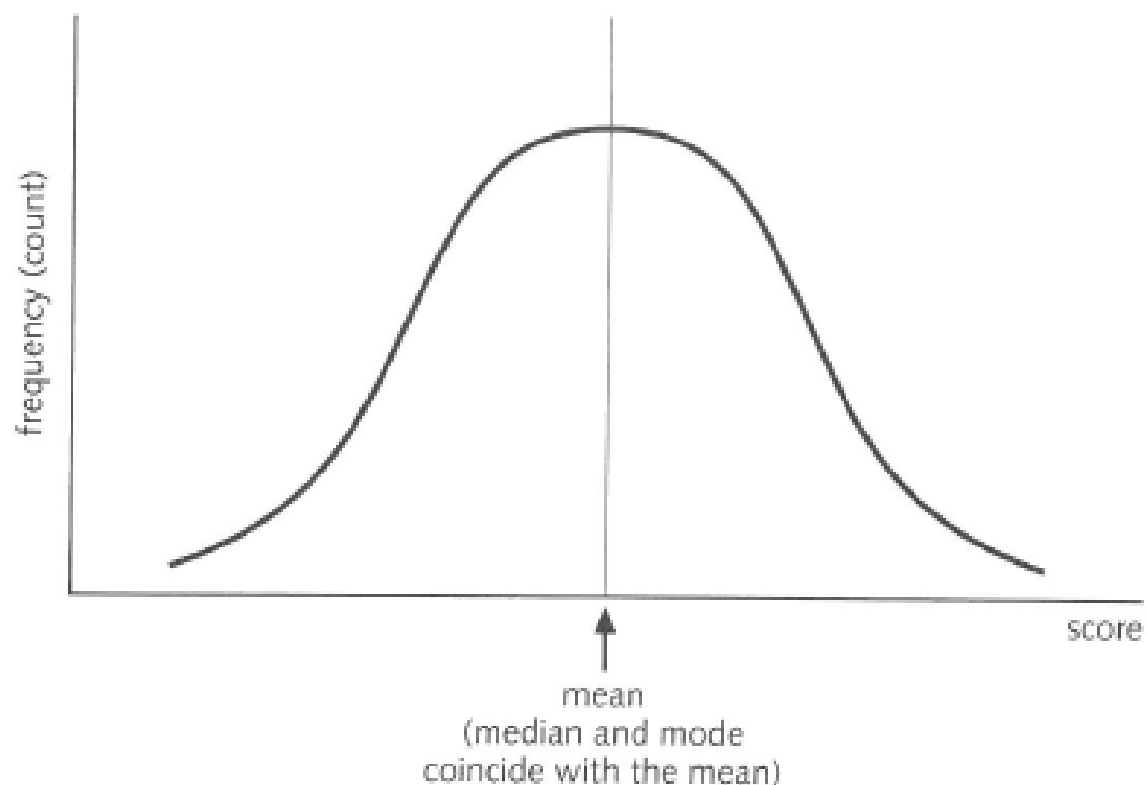
Jak se bojí muži ve dne?

- Aritmetický průměr – 2,79
Medián - 2
Modus - 2



Dvě města: Extrémov a průměrov (Taleb)

- Průměrná výška
- Průměrná váha
- Průměrná mzda
- Průměrné IQ

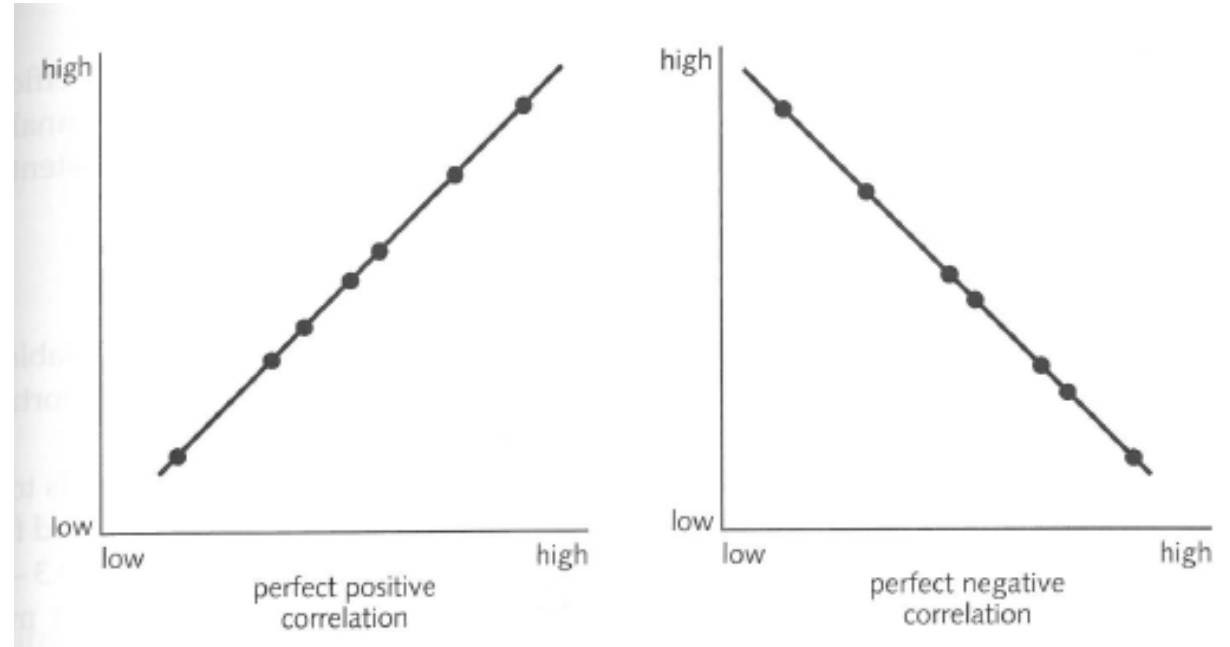


Korelace

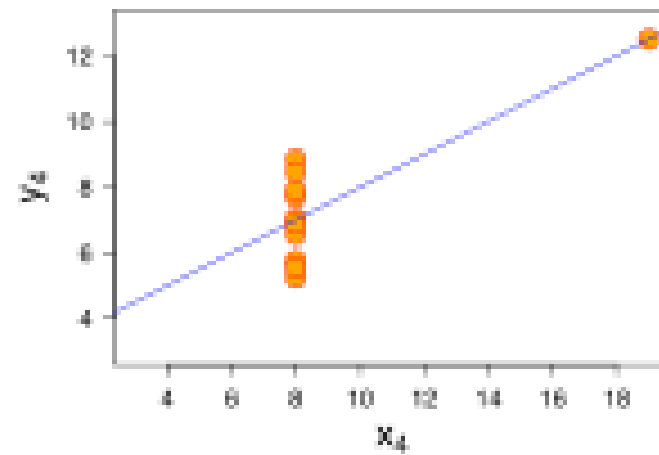
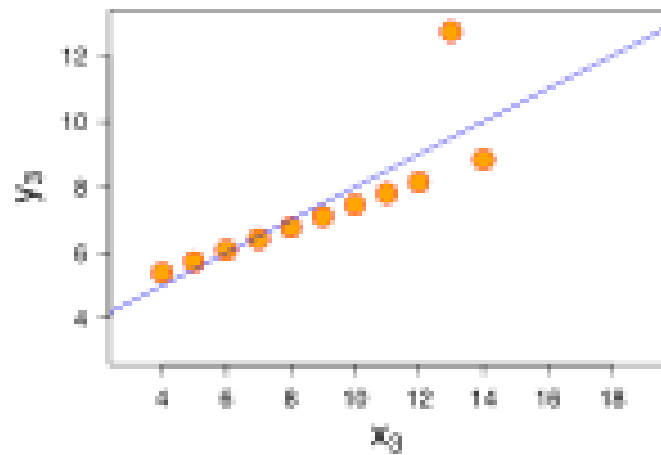
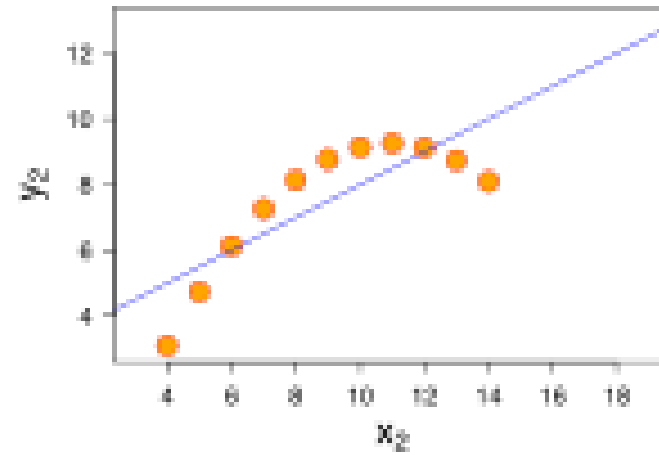
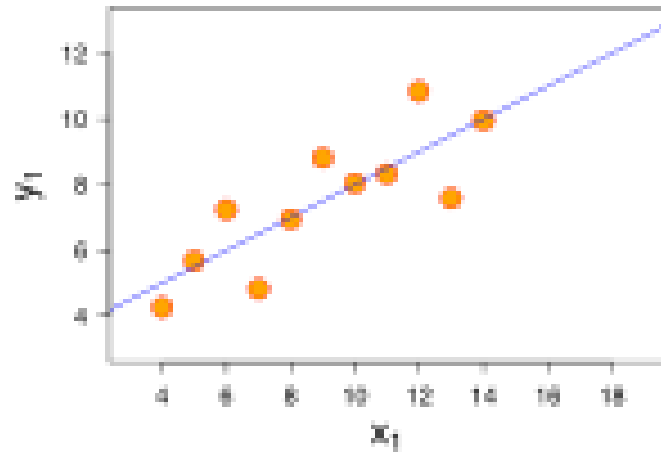
Korelace znamená vzájemný vztah mezi dvěma procesy nebo veličinami.

Pokud se jedna z nich mění, mění se korelativně i druhá a naopak. Pokud se mezi dvěma procesy ukáže korelace, je pravděpodobné, že na sobě závisejí.

Nelze z toho však ještě usoudit, že by jeden z nich musel být příčinou a druhý následkem. To samotná korelace nedovoluje rozhodnout, protože korelace neimplikuje kauzalitu.



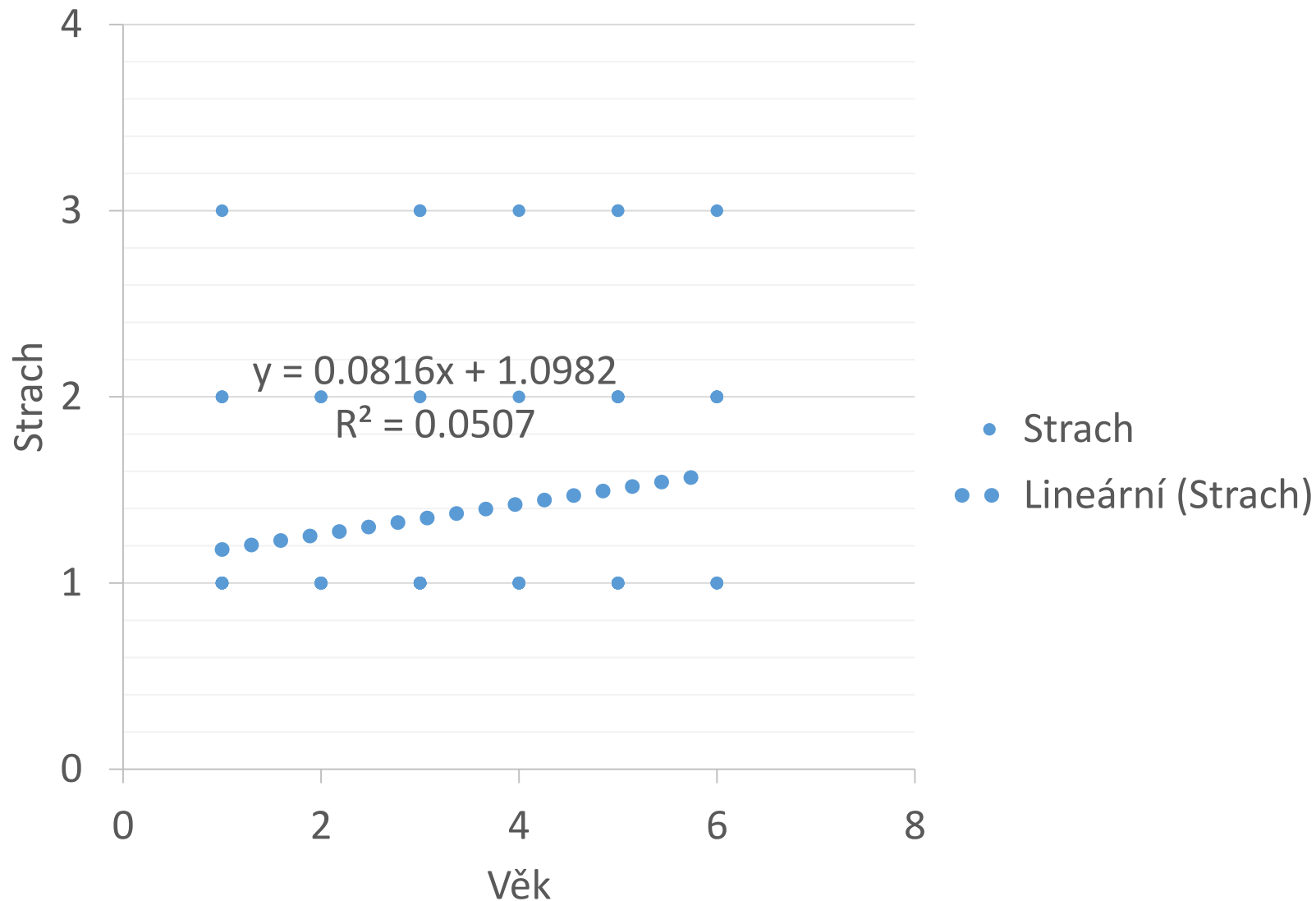
Příklad korelace 0,816



Korelace Strachu v závislosti na věku

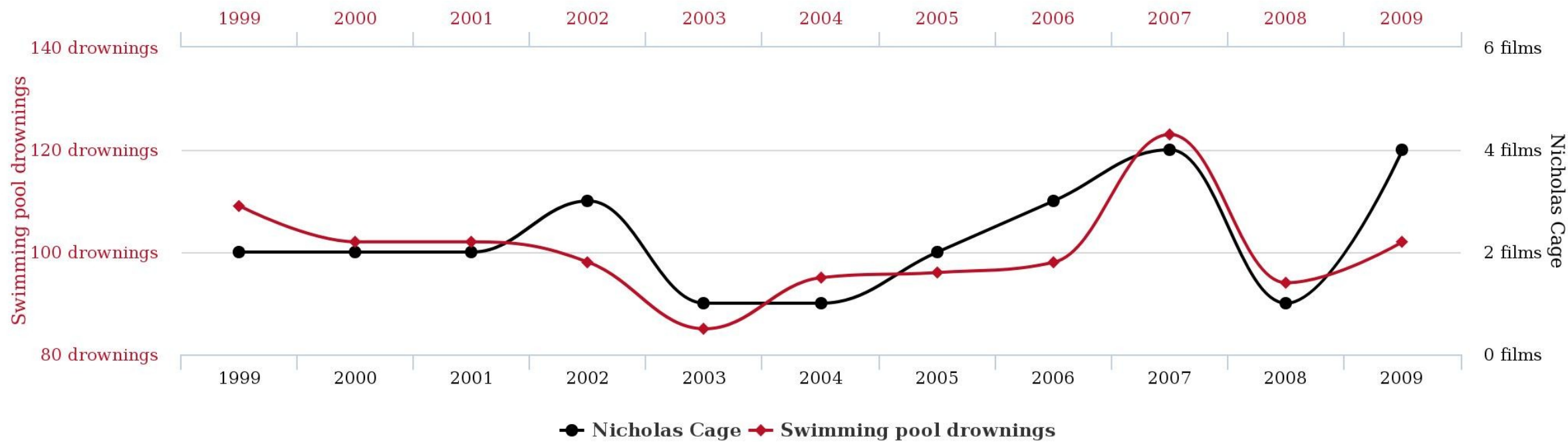
| | Věk | Strach |
|--------|------|--------|
| Věk | 1 | |
| Strach | 0,23 | 1 |

23% případů vysvětluje trend výsledné rovnice (slabá pozitivní korelace)



Proč si neplést korelaci a kauzalitu 😊

Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in



Divorce rate in Maine

correlates with

Per capita consumption of margarine



US spending on science, space, and technology correlates with Suicides by hanging, strangulation and suffocation

