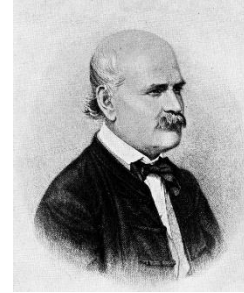
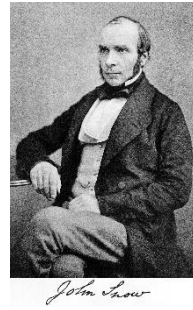
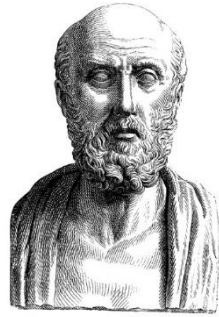


# Epidemiologické studie

Mgr. Aleš Peřina, Ph. D.

# Z historie



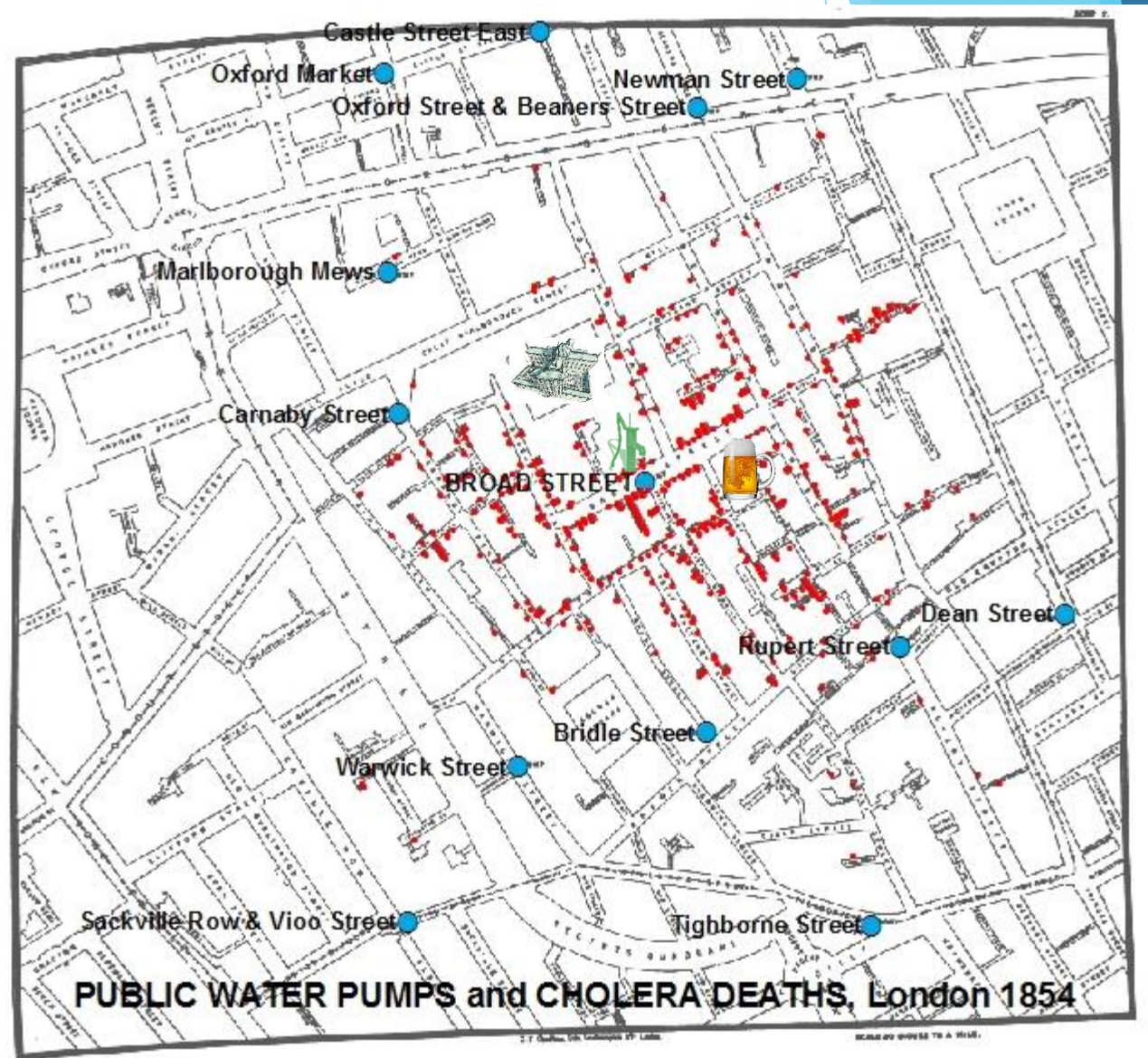
- ▶ **Hippokratés z Kósu:** soustavným sledováním a racionální úvahou nad výsledky se zasloužil o základy moderní medicíny
- ▶ **John Snow** (Londýn, 1854): analýzou místních souvislostí odhalil ohnisko epidemie cholery
- ▶ **Ignaz Semmelweis** (1818 - 1865): všiml si, že větší výskyt puerperální sepse je při domácích porodech a proto nařídil dezinfekci rukou, i když původce onemocnění objevil Luis Pasteur až 1879.
- ▶ Polovina 20. století: postinfekční éra, i když s rozvojem laboratorních metod vyšetřování prostředí nastal příklon k „uvěření“ výsledku laboratoře (otázka stanovitelnosti a objektivity).

## Londýnská epidemie cholery, John Snow, 1854

Mělké infiltrační studny pro jímání  
říční vody ve městě bez kanalizace.

Žádní nemocní v blízkém pivovaru  
ani v blízkém chudobinci s vlastní  
hlubokou studnou.

Pro svoje chuťové vlastnosti byla  
studna na *Broad Street* vyhledávána  
obyvateli i ze vzdálenějších oblastí.

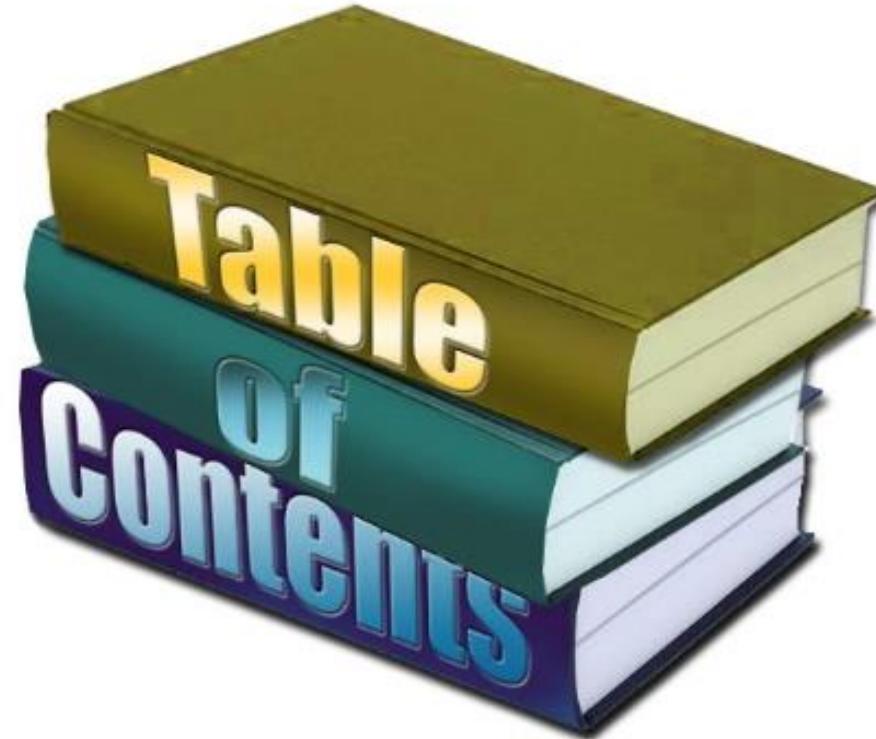


# Epidemiologie

- ▶ Epidemiologií se rozumí studium distribuce a determinant zdravotně významných jevů a událostí v definovaných populacích a využití tohoto studia k řešení zdravotních problémů.
  - ▶ Infekční
  - ▶ Neinfekční
- ▶ Přesah do nových oblastí
  - ▶ Klinická epidemiologie: měření efektu léčby, rozvoj medicíny založené na důkazu (Evidence Based Medicine)
  - ▶ Měření zdravotních služeb
- ▶ *Epidemiologie odpovídá na otázku „kolik?“, kvalitativní výzkum na otázku „proč?“.*

# Jak na to?

- ▶ Výzkumná otázka
- ▶ Volba proměnných
- ▶ Volba typu studie
- ▶ Volba souborů
- ▶ Pilotní studie
- ▶ Měření
  - ▶ Velikost nebo frekvence jevu
- ▶ Statistické zpracování
  - ▶ Deskriptivní (jednorozměrná analýza)
  - ▶ Analytické (dvou- a více rozměrná analýza)
- ▶ Interpretace výsledků
- ▶ Nejistoty



# Nosná myšlenka

- ▶ Co studovat?
  - ▶ Biologický, chemický, fyzikální, psychologický, sociální, behaviorální... faktor
- ▶ Úvaha o expozici, aneb „cesta do hlubin člověka“.
- ▶ Lze očekávat následek?
  - ▶ Hypotéza
    - ▶ Na základě literární rešerše (pokud možno)
  - ▶ Ukazatele míry
    - ▶ V metrické soustavě (výška, hmotnost, obvod pasu, glykémie na lačno...)
  - ▶ Ukazatele frekvence
    - ▶ Incidence nemoci, Počet zemřelých
  - ▶ Riziko je pravděpodobnost (ale i emoce)



# Zdroje dat

- ▶ Rutinně sbíraná data o nemocnosti, úmrtnosti
  - ▶ Rutinní evidence ze statistiky UZIS, ČSÚ, registry nemocí, úrazů, transplantací, lékařské záznamy praktických a odborných lékařů, záznamy o úmrtích
  - ▶ Základ tzv. ekologických (korelačních) studií
- ▶ Vlastní výzkum
  - ▶ Dotazník
  - ▶ Interview
  - ▶ Laboratorní analýza

# Strategie deskriptivní statistiky

- ▶ **Spojité proměnná**
  - ▶ Ordinální: když 2 je více než 1
  - ▶ Intervalová: rozpětí ordinálních hodnot, např. skóre v testu
  - ▶ Indexy: relativní číslo
- ▶ **Nespojitá (kategorická) data**
  - ▶ Nominální stupnice: somatotyp, krevní skupina, barva očí...

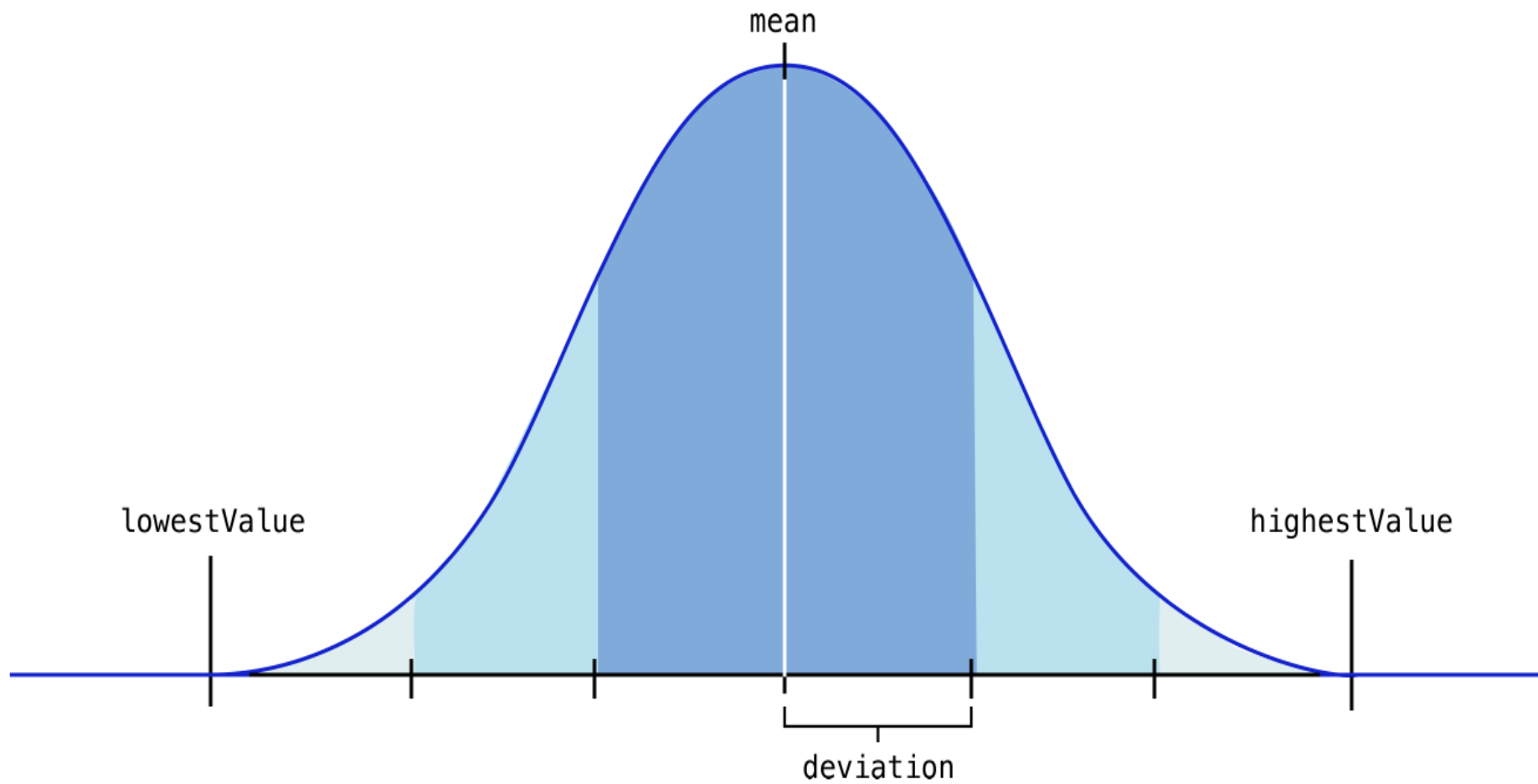


# Deskripce spojité proměnné

- ▶ Průměr: součet množiny čísel vydělený počtem těchto čísel, ukazatel centrální tendence
  - ▶ =PRŮMĚR(A2:A6)
- ▶ Medián: prostřední hodnota, 50. percentil v řadě čísel seřazených vzestupně
  - ▶ =MEDIAN(A2:A6)
- ▶ Modus: nejčastěji se vyskytující hodnota v souboru
  - ▶ =MODE(A2:A7)
- ▶ Minimum: nejmenší hodnota v souboru
  - ▶ =MIN(A2:A6)
- ▶ Maximum: největší hodnota v souboru
  - ▶ =MAX(A2:A6)

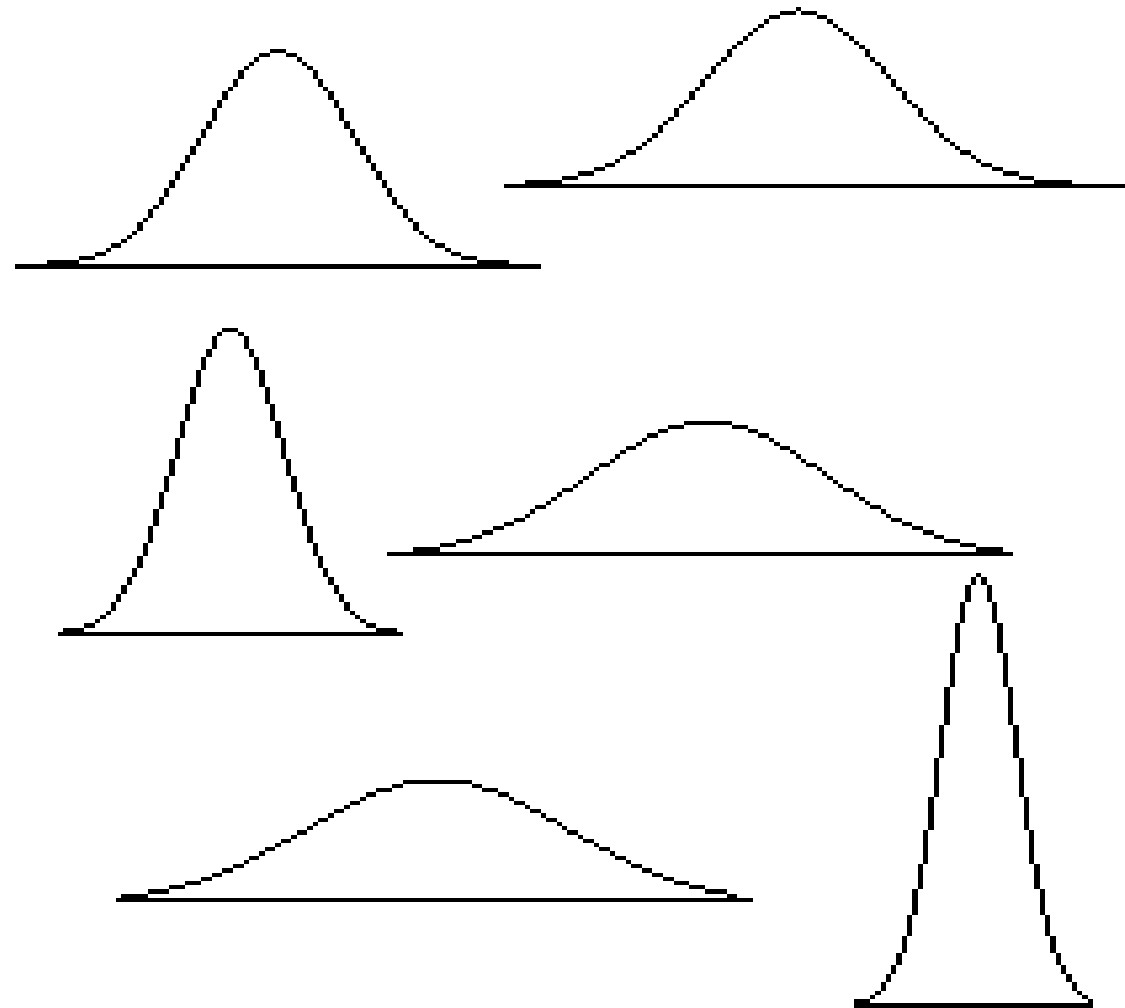
U symetrického rozložení platí: průměr = medián = modus

# Průměr



## Průměr

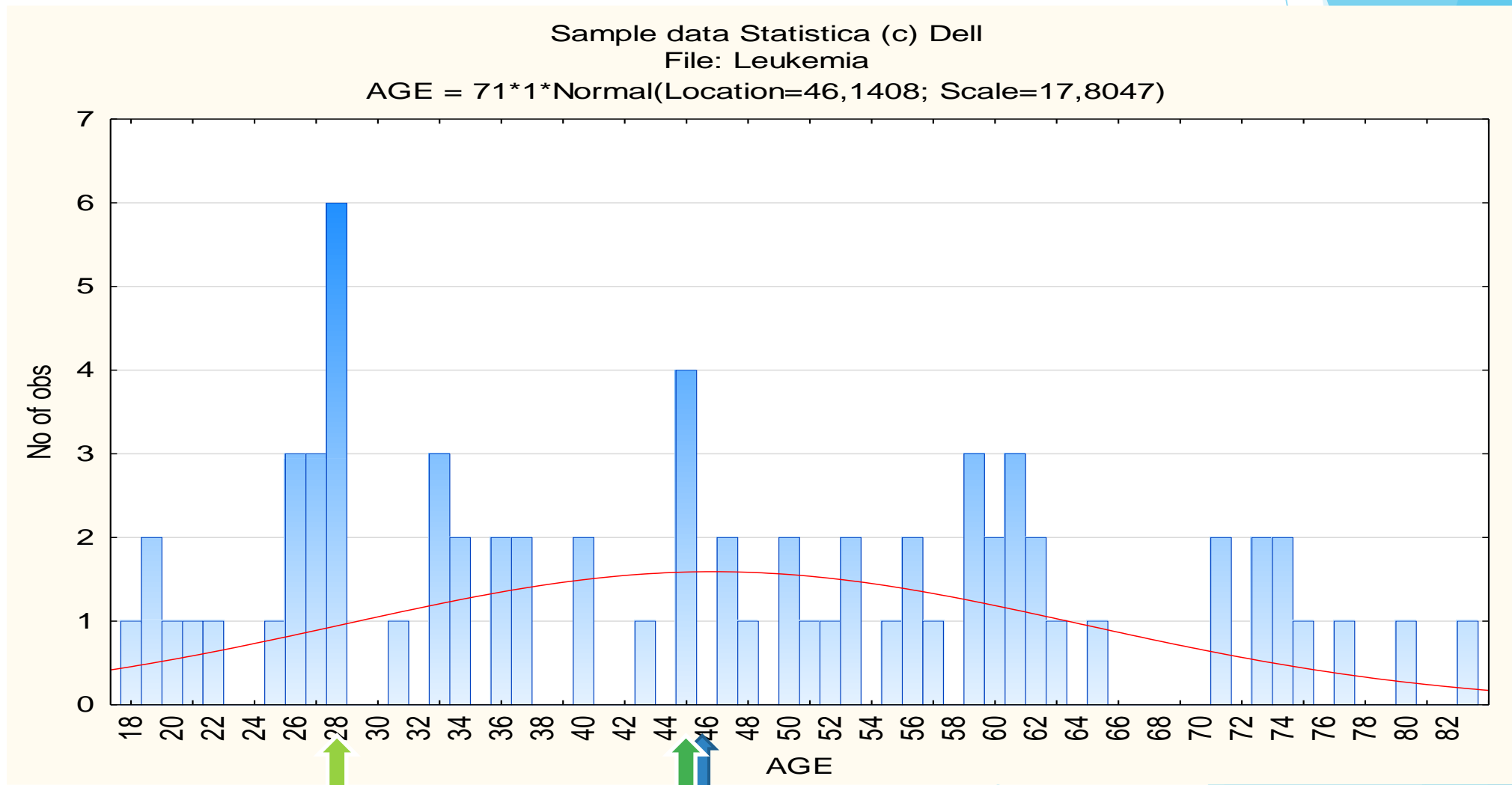
- ▶ Je opodstatněný u dat rovnoměrně rozložených kolem střední hodnoty.
- ▶ Doplněný o SMODHC.
- ▶ Při vychýlených rozloženích je zavádějící



# Průměr

Sample data Statistica (c) Dell								
File: Leukemia								
Variable	Valid N	Mean	Median	Mode	Frequency of Mode	Minimum	Maximum	Std.Dev.
AGE	71	46,14085	45,00000	28,00000	6	18,00000	83,00000	17,80473

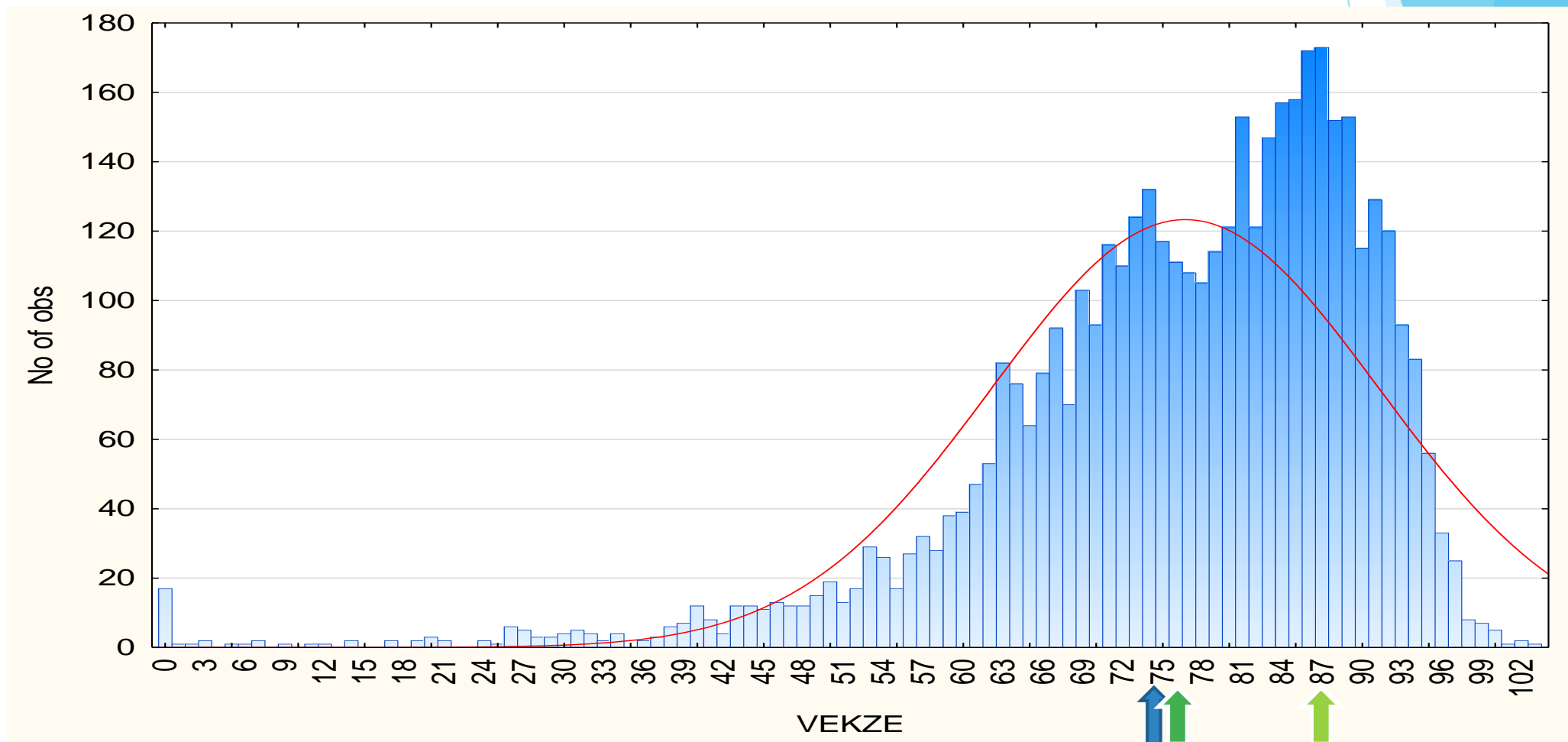
# Průměr



# Průměr

Variable	Mortalita v oblasti vlivu JETE							
	Valid N	Mean	Median	Mode	Frequency of Mode	Minimum	Maximum	Std.Dev.
VEKZE	4473	76,30673	79,00000	87,00000	173	0,00	103,0000	14,46791

# Průměr: počet zemřelých podle věku





# Deskripce nespojitých (kategoriálních) dat

- ▶ Četnost: počet výskytu hodnoty dané kategorie (např. věkový interval od - do)
- ▶ Relativní četnost: počet výskytu hodnot vztažených ke zvolenému základu (100 %)
- ▶ Kumulativní četnost: součet hodnot sledované kategorie a všech nižších kategorií

# Deskripce nespojitých (kategoriálních) dat

Count	Percent	Cumulative Percent
101	2,3	2,3
122	2,7	5,0
102	2,3	7,3
218	4,9	12,1
151	3,4	15,5
70	1,6	17,1
87	1,9	19,0
111	2,5	21,5
96	2,1	23,7
169	3,8	27,4
119	2,7	30,1
1042	23,3	53,4
993	22,2	75,6
1092	24,4	100,0
0	0,0	100,0

# Požadavky na vstupní údaje

- ▶ Expoziční proměnná
  - ▶ Spotřeba potravin, frekvenční dotazník, spotřeba vody, fyzická aktivita, množství kontaminující látky v potravinách nebo v pitné vodě
- ▶ Znak
  - ▶ Zdraví
  - ▶ Nemoc: definice diagnostických znaků (*inclusion* kritéria), *exclusion* kritéria k vyloučení zdrojů statistických bias (příště 😊 )
  - ▶ Smrt
- ▶ Nemocnost a úmrtnost
  - ▶ Celková
  - ▶ Specifická podle příčiny nebo podle věku

# Požadavky na vstupní údaje

## ▶ Validita

- ▶ Popisuje studovaný znak situaci, kterou hodláme studovat?
  - ▶ Lze z frekvence konzumace potravin usuzovat na zdravotní stav nebo jen na to, kdo má co rád?
  - ▶ Lze na základě výsledků laboratorního vyšetření definovat případ (inclusion kritérium), nebo má laboratorní vyšetření pouze pomocný charakter?
    - ▶ Srov. Glykemie na lačno vs. CRP

## ▶ Senzitivita a specificita

- ▶ Senzitivita: úspěšnost záchytu sledovaného znaku (100 % = žádné falešně negativní)
- ▶ Specificita: schopnost testu odlišit případy, které znak nemají (100 % = žádné falešně pozitivní)

# Senzitivita a specificita

	Má pozitivní výsledek testu	Má negativní výsledek testu
Je nemocný/-á	Skutečně pozitivní (SENZITIVITA)	Falešně negativní
Je zdravý/-á	Falešně pozitivní	Skutečně negativní (SPECIFICITA)

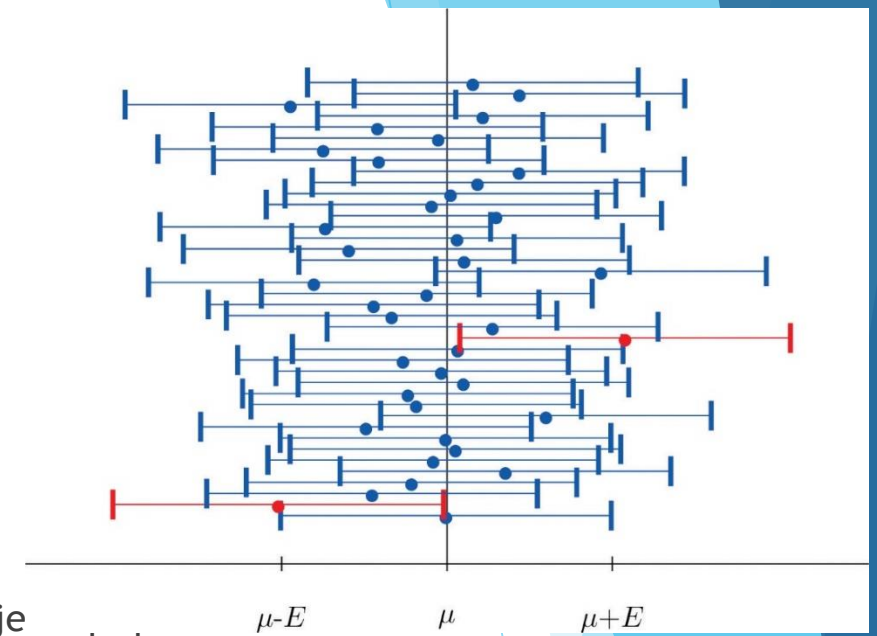
## Ročenka UZIS

1.7.1 Zemřelí a úmrtnost podle příčin smrti – muži  
Deaths and mortality rate by causes of death – males

Kód diagnózy (MKN-10)	Název kapitoly, diagnózy	Zemřelí		Standardizovaná úmrtnost
		absolutně	na 100 000 mužů	
				1/2
A00–B99	I. Některé infekční a parazitární nemoci	947	18,2	13,8
A15–A16	z toho: tbc dýchacího ústrojí	18	0,3	0,3
A17–A19	jiná tbc	3	0,1	0,0
A83–A86	virová encefalitida	7	0,1	0,1
A80–A82, A87–A89	ostatní virové infekce centrální nervové soustavy	3	0,1	0,1
B15–B19	virová hepatitida	18	0,3	0,3
C00–D48	II. Novotvary	15 485	297,4	221,1
C33–C34	z toho: zhoubný novotvar průdušnice, průdušky a plice	3 639	69,9	50,9
D50–D89	III. Nemoci krve, krvetvorných orgánů a některé poruchy mechanismu imunity	100	1,9	1,5
E00–E90	IV. Nemoci endokrinní, výživy a přeměny látek	2 025	38,9	29,5
F00–F99	V. Poruchy duševní a poruchy chování	686	13,2	10,2
G00–G99	VI. Nemoci nervové soustavy	1 408	27,0	20,9
H00–H59	VII. Nemoci oka a očních adnex	-	-	-
H60–H95	VIII. Nemoci ucha a bradavkového výběžku	1	-	-
I00–I99	IX. Nemoci oběhové soustavy	22 840	438,6	334,2
I05–I09	z toho: chronické revmatické choroby srdeční	98	1,9	1,4
I10	esenciální (primární) hypertenze	523	10,0	7,6
I11–I15	jiné hypertenzní nemoci	799	15,3	11,8
I21–I22	akutní a pokračující infarkt myokardu	2 673	51,3	38,9
I20, I23–I25	ostatní ischemické nemoci srdeční	9 209	176,8	134,8
I30–I52	jiné formy srdečního onemocnění	4 129	79,3	61,1
I60–I69	cévní nemoci mozku	3 633	69,8	52,6
J00–J99	X. Nemoci dýchací soustavy	4 392	84,3	63,7
J09–J11	z toho: chřipka	80	1,5	1,2
J12–J18	zánět plic	1 675	32,2	25,0
K00–K93	XI. Nemoci trávicí soustavy	2 734	52,5	41,8
L00–L99	XII. Nemoci kůže a podkožního vaziva	91	1,7	1,3
M00–M99	XIII. Nemoci svalové a kosterní soustavy a pojivové tkáně	82	1,6	1,2
N00–N99	XIV. Nemoci močové a pohlavní soustavy	750	14,4	10,9
O00–O99	XV. Těhotenství, porod a šestinedělí	-	-	-
P00–P96	XVI. Některé stavy vzniklé v perinatálním období	103	2,0	2,8

# Incidence a prevalence

- ▶ Incidence: nově vzniklé případy během studie
  - ▶ Osoboočas: počet účastníků studie násobený délkou studie
    - ▶ 30 osob během dvouleté kohortové studie je 60 osoboroků, incidence je osoboroků, tj. 0,167/osoborok (vyjádření rizika).
- ▶ Prevalence: existující případy k určitému datu (bodová) nebo období (intervalová). Vyjadřuje se jako proporce v populaci (obv. %).
  - ▶ Nižší o vyléčené nebo zemřelé
  - ▶ Nižší o počty případů, které se vyskytly po stanoveném datu.
  - ▶ Vyšší o případy, které přešly např. do chronicity
  - ▶ Vhodný ukazatel ke získání přehledu a formulaci hypotéz.
- ▶ Incidence a prevalence se vyjadřují v relativních číslech vztažených k vhodnému základu (1000 osob, 100 tis. osob, v závislosti na frekvenci jevu a velikosti populace).

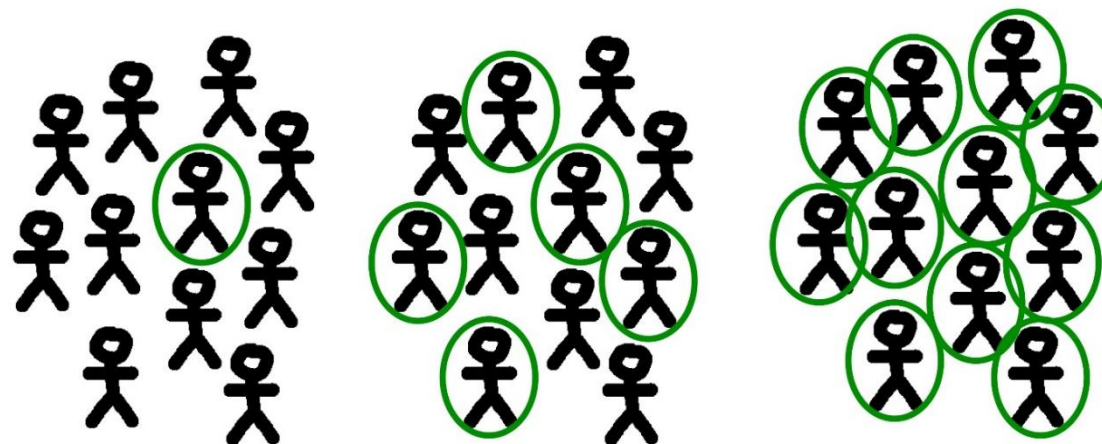




# Studovaná populace

- ▶ Reprezentativní výběr: mající shodné charakteristiky se souborem základním
  - ▶ Charakteristiky věku, pohlaví, ekonomické aktivity
- ▶ Kritéria
  - ▶ Randomizace (nahodilost)
  - ▶ Stratifikace (rozvrstvení)

Measurement > Sampling (more than one measurement)



Measurement

Sample

Population

not representative

representative

# Návrh (*design*) studie

- ▶ Deskriptivní studie (zpravidla univariační)
- ▶ Analytická studie (zpravidla bivariační)
  - ▶ Průřezová (cross-sectional)
  - ▶ Ekologická
  - ▶ Případů a kontrol (case - control)
  - ▶ Kohortová
- ▶ Kontrolovaný experiment
  - ▶ Na populační úrovni prakticky nerealizovatelný
  - ▶ Výjimky: jaderná havárie v Černobylu (1986) jako nechtěný „experiment“; velmi dobře popsaná expozice umožňující studovat následky ionizujícího záření na zdraví populaci

# Měření vztahu dvou (a více) proměnných

## ► Design:

- Ekologická studie: obvykle na základě rutinních statistik
- Studie případů a kontrol: problematická definice **případu**; následek je znám, pátráme po expozičních odlišnostech
- Kohortová studie: problematický způsob definice kohort, kontrolujeme expozici a pátráme po následku.

## ► Příklad

2X2 TABULKA	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	205	129	334
Anemie NE	89	86	175
CELKEM	294	215	509

# Pozorované a očekávané četnosti

Pozorované četnosti	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	205	129	334
Anemie NE	89	86	175
CELKEM	294	215	509

Očekávané četnosti =CHITEST (Excel) P = 0,02	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	193 = $294 \times 334 / 509$ - 12	141 + 12	334
Anemie NE	101 + 12	74 - 12	175
CELKEM	294	215	509

# Měření velikosti efektu pomocí ukazatelů velikosti rizika

Pozorované četnosti	Zdroje Fe nedostatečné	Zdroje Fe dostatečné	CELKEM
Anemie ANO	205	129	334
Anemie NE	89	86	175
CELKEM	294	215	509

**Ekologická studie:**  
Korelace incidence případů anémie v různých územích vzhledem k prodeji masa.

**Relativní riziko (kohorta)**

$$205/294 = 0,70$$

$$125/215 = 0,60$$

$RR = 0,70/0,60 = 1,17$   
1,17-krát nebo o 17 % více

**Poměr šancí (Odds Ratio, případy)**

$$205/89 = 2,30$$

$$120/86 = 1,50$$

$OR = 2,30/1,50 = 1,53$   
O 53 % větší šance.

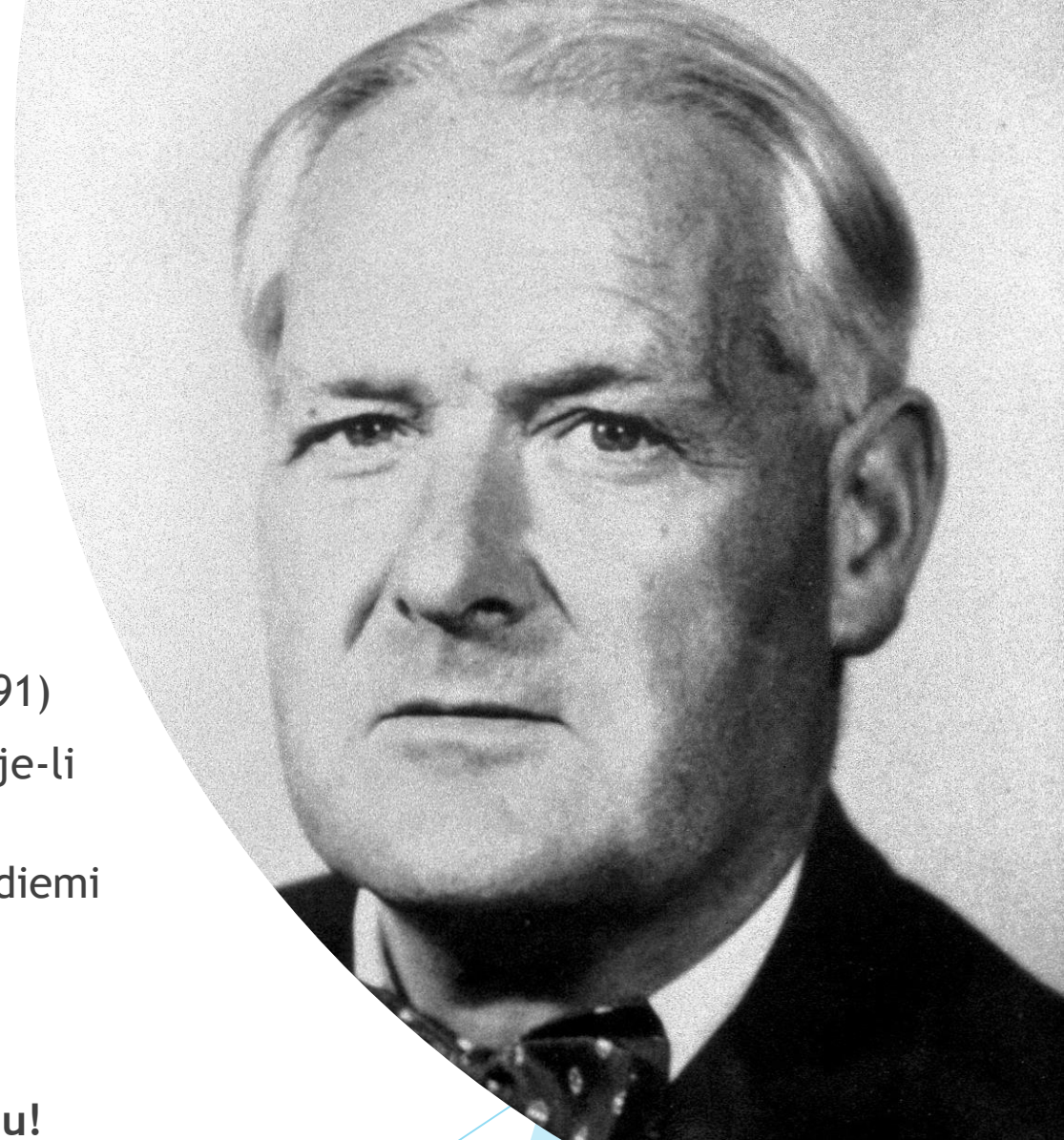
## Obecný formát 2x2 (čtyřpolní) tabulky

<b>Exponovaní Nemocní</b>	<b>Neexponovaní Nemocní</b> <i>Falešně pozitivní</i>
<b>Exponovaní Zdraví</b> <i>Falešně negativní</i>	<b>Neexponovaní Zdraví</b>

RR nebo OR = 1 ... není efekt  
RR nebo OR > 1 ... rizikový faktor  
RR nebo OR < 1 ... ochranný faktor

# Epidemiologie v hodnocení zdravotních rizik

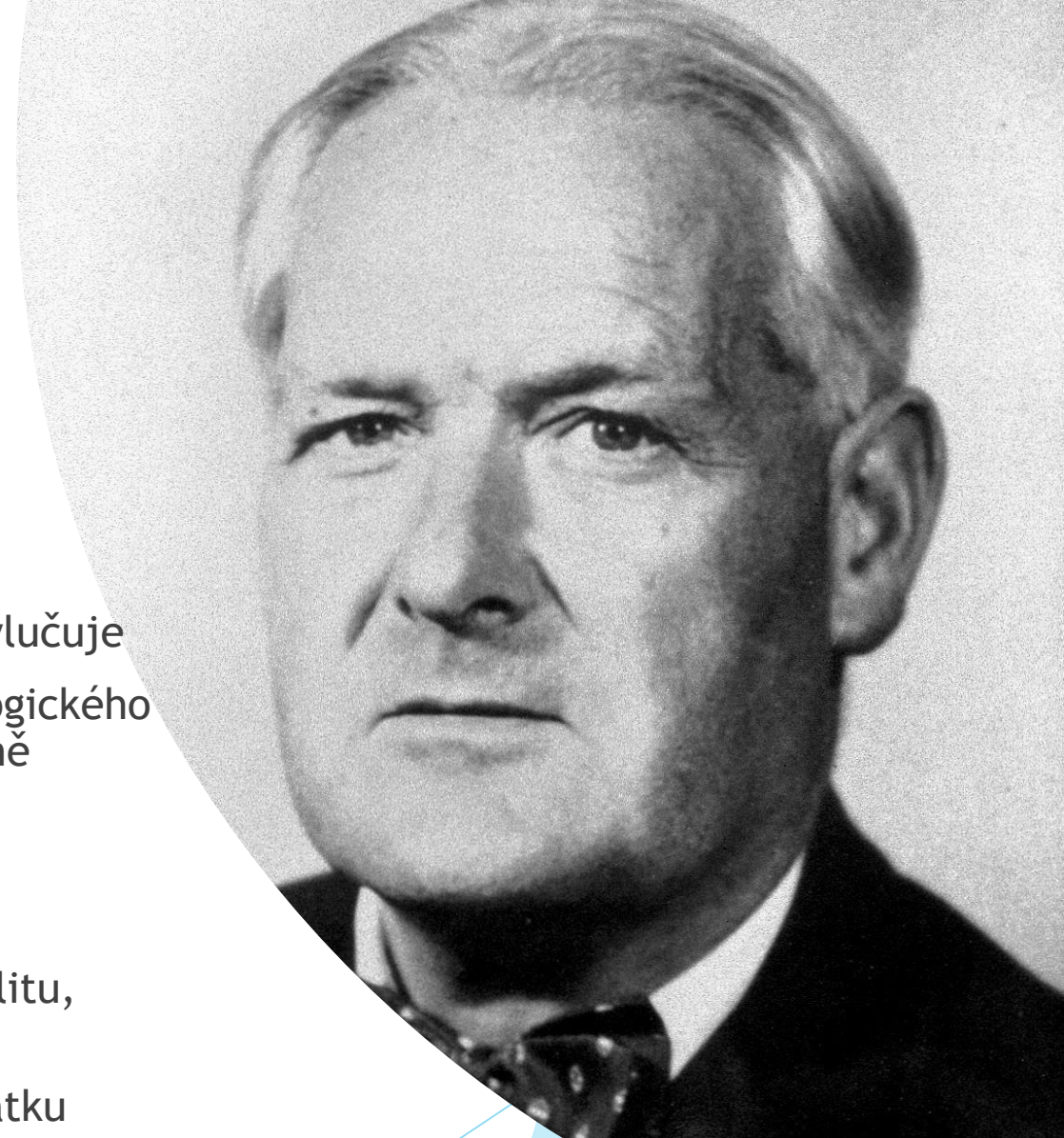
- ▶ Úskalí: přenositelnost výsledků
- ▶ Interní validita epidemiologické studie
- ▶ Hillova kritéria kauzality (sir Austin Bradford Hill, 1897 - 1991)
  - ▶ **Síla asociace:** ani slabá asociace nevylučuje kauzalitu, je-li oslabena nerozpoznanými confoundery
  - ▶ **Konzistence:** avšak nekonzistentnost s jinými epid. studiemi nevylučuje kauzalitu, efekt se může dostavovat jen za zvláštních okolností
  - ▶ **Specifitu účinku** kauzalita nepředpokládá
  - ▶ **Časová posloupnost** expozice a následku je podmínkou!

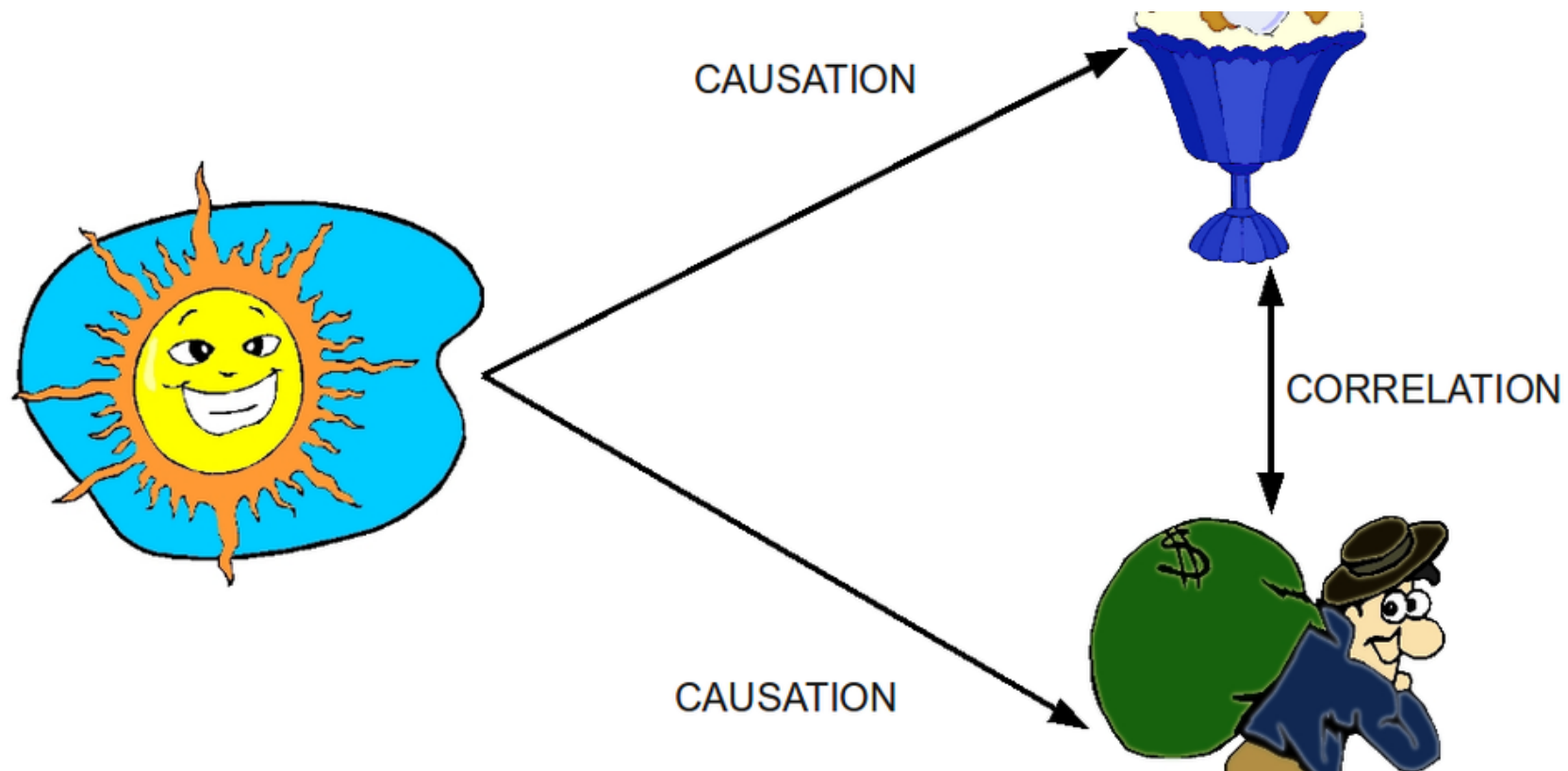




# Epidemiologie v hodnocení zdravotních rizik II.

- ▶ Hillova kritéria kauzality ... pokrač.
  - ▶ **Biologický gradient:** avšak jeho absence kauzalitu nevylučuje
  - **Biologická přijatelnost:** avšak neschopnost patofyziologického vysvětlení jevu může být jen důsledkem aktuální úrovně vědeckého poznání
  - **Koherence:** inkompatibilita se "zavedenými" teoriemi nevylučuje kauzalitu
  - **Experimentální důkaz:** jeho absence nevylučuje kauzalitu, neboť experimentu mohou bránit též etické důvody
  - **Analogie:** její absence může být jen projevem nedostatku vědecké představitosti



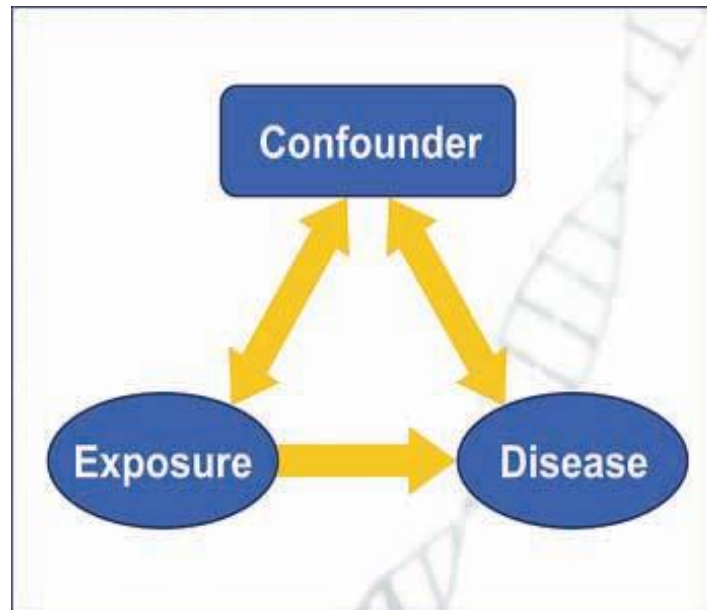


## Formální a logická kauzalita.

Spotřeba zmrzliny matematicky koreluje s nárůstem kriminality ve městě, ale ve skutečnosti je nárůst kriminality spojen s nástupem turistické sezóny.

# Bias

- ▶ Základní typy
  - ▶ Informační: na co se ptát a jak?
  - ▶ Selekční: zdrojem bývá nízká reprezentativnost vzorku
  - ▶ Confounding: spolupůsobící či závadějící faktory



# Závěr

- ▶ Základem vědecké práce jsou
  - A. Informace o stavu řešení dané vědecké otázky
  - B. Dostupnost dat pro vlastní vědeckou práci
- ▶ Data jsou cenná a některá i osobního charakteru. Data je proto třeba chránit před zneužitím a vytěžit z nich maximum nových informací (*data mining*), které jsou využitelné pro další vědecký výzkum nebo praxi.