

Přednáška 2

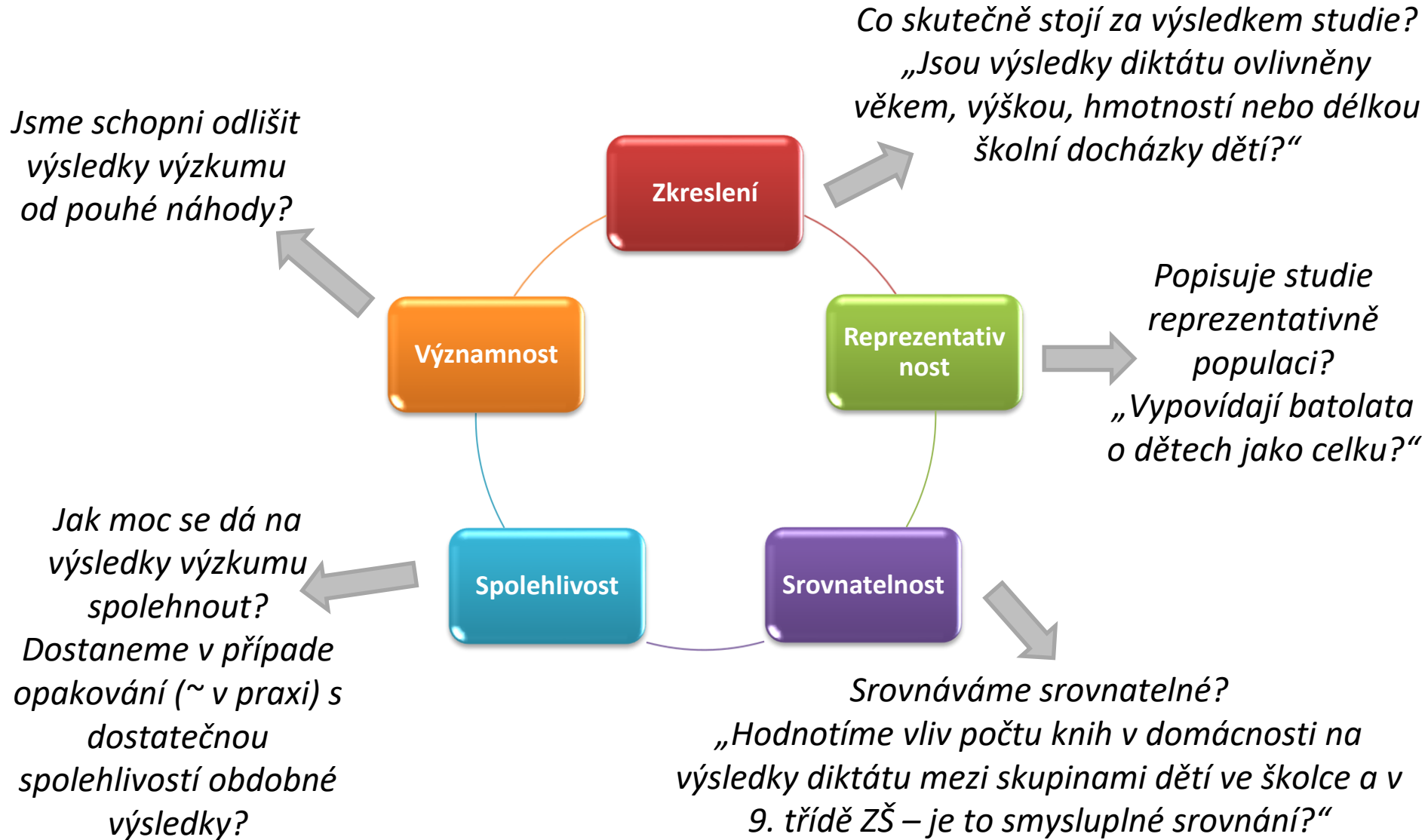
Klíčové principy biostatistiky

Zkreslení, reprezentativnost, srovnatelnost, spolehlivost významnost

Anotace

- Ve statistické analýze biologických a klinických dat musíme vždy nad prováděným výzkumem a jeho výsledky přemýšlet v kontextu 5 klíčových principů biostatistiky.
- Zkreslení – skutečně vidíme to co si myslíme, že vidíme?
- Reprezentativnost – vypovídá naše analýza o skupině objektů, která nás zajímá?
- Srovnatelnost – co ve skutečnosti v analýze srovnáváme?
- Spolehlivost – jak spolehlivé jsou naše výsledky, dají se zopakovat?
- Významnost – jak moc je pravděpodobné, že pozorujeme výsledky pouhé náhody?
- Zanedbání těchto principů může vést k chybné interpretaci výsledků.

Klíčové principy biostatistiky

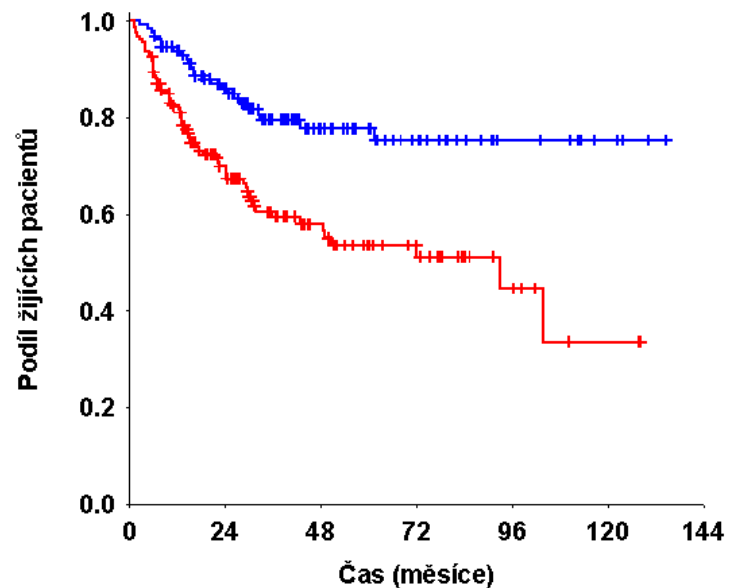


Klíčové principy – zkreslení

- V jakémkoliv hodnocení se snažíme vyhnout zkreslení výsledků („biased results“) – tedy zkreslení výsledků jinými faktory než těmi, které jsou cíli výzkumu.
- Statistické srovnání není nikdy 100% spolehlivé, existuje náhoda a tedy i pravděpodobnost chybného úsudku – to nelze ovlivnit.
- Chceme použít adekvátní metody pro odstranění vlivů, které by zkreslily výsledky a nebyly přítom náhodné (např. zastoupení pohlaví, nadmořská výška).

Klíčové principy – zkreslení

- Co způsobuje rozdíl v saprobním znečištění vodního toku?
- Co způsobuje rozdíl v naměřených biochemických ukazatelích?
- Čím by mohl být způsoben pozorovaný rozdíl v 10letém přežití pacientů?



Léčba?

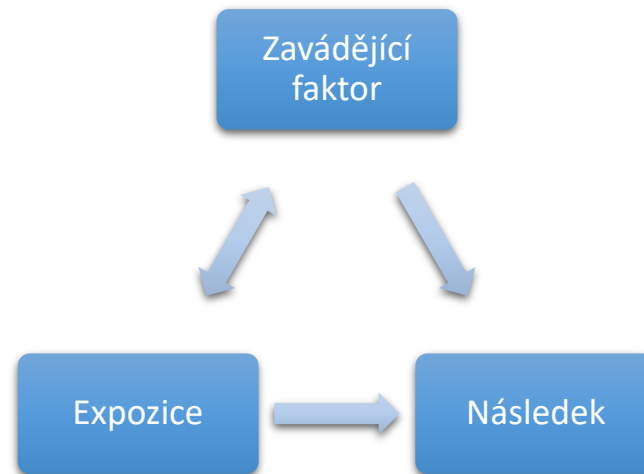
Nějaký prognostický faktor?

Stadium nemoci?

Věk?

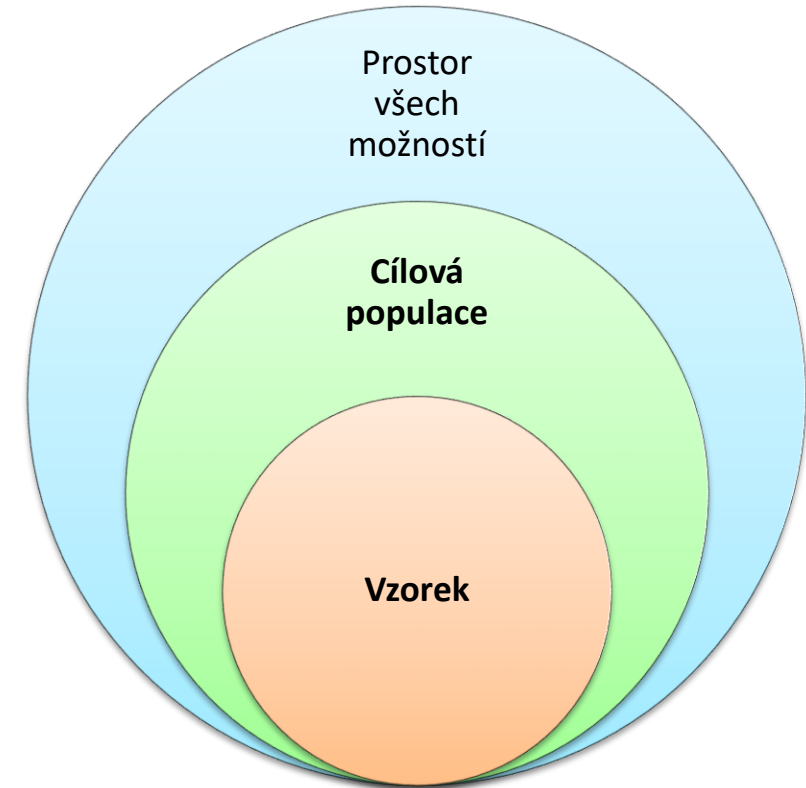
Klíčové principy – zkreslení

- Pojem zavádějící faktor
- Pro zavádějící faktor současně platí, že
 - přímo nebo nepřímo ovlivňuje sledovaný následek,
 - je ve vztahu se studovanou expozicí ,
 - není mezikrokem mezi expozicí a následkem.

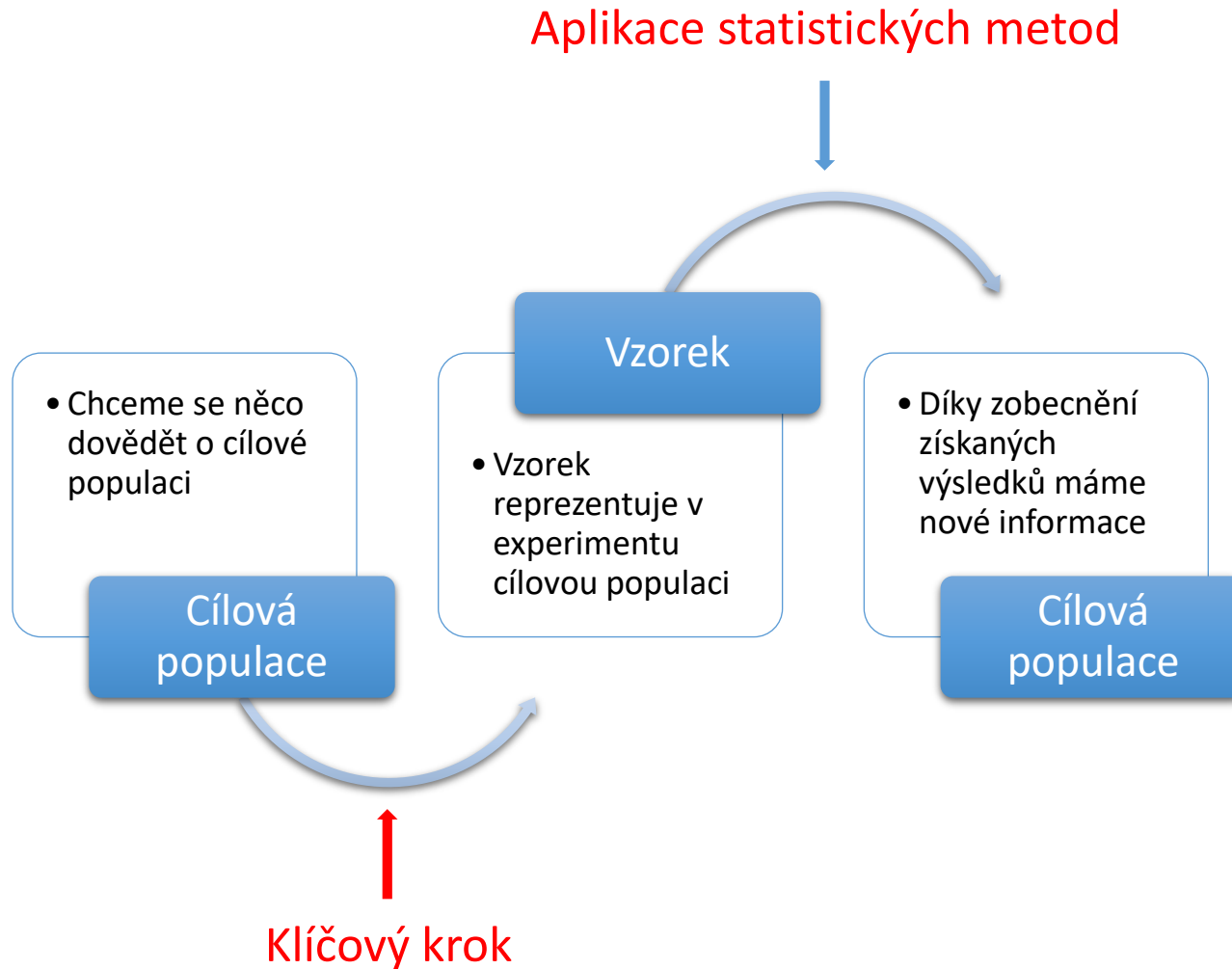


Klíčové principy – reprezentativnost

- Pojem cílová populace – skupina subjektů, o které chceme zjistit nějakou informaci.
- Pojem experimentální vzorek – podskupina cílové populace, kterou „máme k dispozici“.
 - Musí odpovídat svými charakteristikami cílové populaci.
 - Chceme totiž zobecnit výsledky na celou cílovou populaci.
 - Souvislost s náhodným výběrem.



Klíčové principy – reprezentativnost



Klíčové principy – srovnatelnost

- Korektní výsledky při srovnávacích analýzách lze získat pouze při srovnávání srovnatelného.
- V striktně kontrolovaných studiích je srovnatelnost zajištěna randomizací.
- U studií bez randomizace je nutné se tématu srovnatelnosti skupin věnovat.
- Metody adjustace, matching, propensity scores.



Klíčové principy – spolehlivost

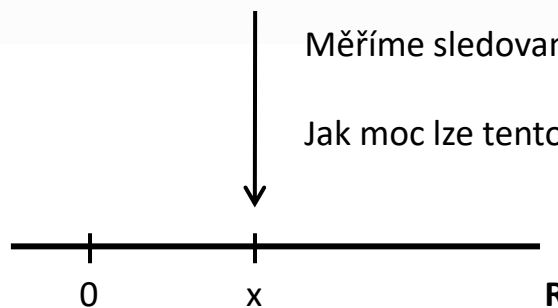
- Ve většině studií nás zajímá kvantifikace sledovaného efektu nebo charakteristiky, obecně náhodné veličiny, ve formě jednoho čísla, bodového odhadu.
- Bodový odhad je však sám o sobě nedostatečný.
- Je nutné ho doplnit intervalovým odhadem, který odpovídá pravděpodobnostnímu chování sledované veličiny, tedy odpovídá určité spolehlivosti výsledku.

Klíčové principy – spolehlivost



Měříme sledovanou veličinu a následně spočítáme odhad.

Jak moc lze tento bodový odhad zobecnit na cílovou populaci?



Klíčové principy – spolehlivost



Opět měříme sledovanou veličinu.

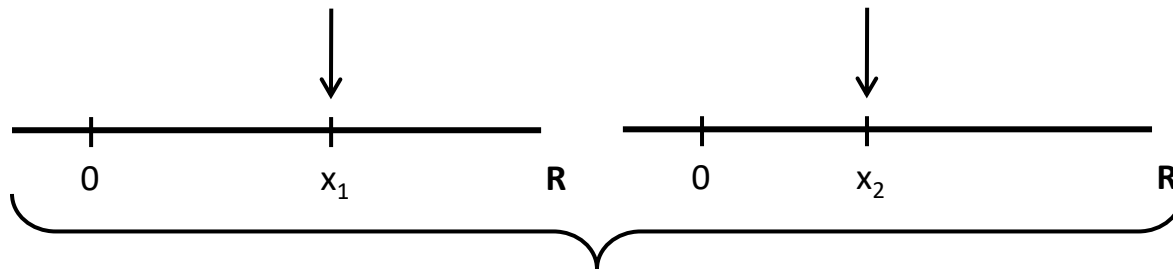
Jaký je rozdíl?

A co když naopak přidáme někoho jiného?

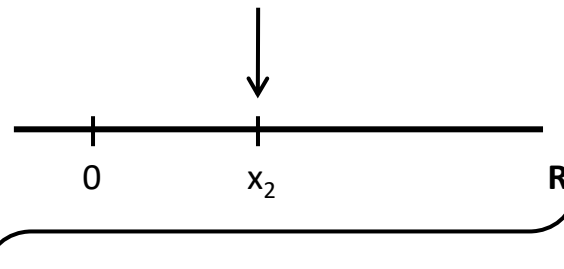


Klíčové principy – spolehlivost

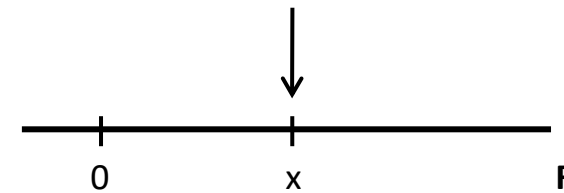
Výběr číslo 1



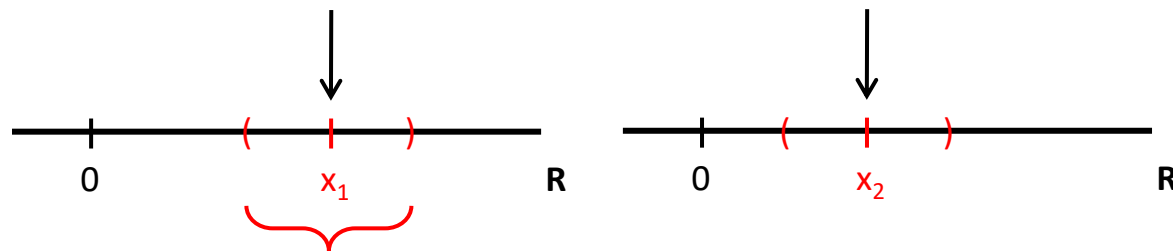
Výběr číslo 2



Celá cílová populace



Pracujeme-li s výběrem z cílové populace, je třeba na základě variability pozorovaných dat spočítat tzv. interval spolehlivosti pro bodový odhad.



Interval spolehlivosti na základě výběru číslo 1.

Umíme-li „změřit“ celou cílovou populaci, nepotřebujeme interval spolehlivosti, protože jsme schopni odhadnout sledovaný parametr přesně – v praxi je tato situace nereálná.

Klíčové principy – významnost

- Analytické výsledky studie nemusí odpovídat realitě a skutečnosti. Statistická významnost jednoduše nemusí znamenat příčinný vztah!
- Statistická významnost pouze indikuje, že pozorovaný rozdíl není náhodný (ve smyslu stanovené hypotézy).
- Stejně důležitá je i praktická významnost, tedy významnost z hlediska lékaře nebo biologa.
- Statistickou významnost lze ovlivnit velikostí vzorku.

Klíčové principy – významnost

		Praktická významnost	
		ANO	NE
Statistická významnost	ANO	OK, praktická i statistická významnost jsou ve shodě.	Významný výsledek je statistický artefakt, prakticky nevyužitelný.
	NE	Výsledek může být pouhá náhoda, neprůkazný výsledek.	OK, praktická i statistická významnost jsou ve shodě.



Statisticky nevýznamný výsledek neznamená, že pozorovaný rozdíl ve skutečnosti neexistuje! Může to být způsobeno nedostatečnou informací v pozorovaných datech!

Příprava dat

Klíčový význam korektního uložení získaných dat

Pravidla pro ukládání dat

Čištění dat před analýzou

Anotace

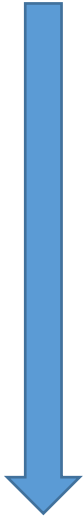
- Současná statistická analýza se neobejde bez zpracování dat pomocí statistických software.
- Předpokladem úspěchu je správné uložení dat ve formě „databázové“ tabulky umožňující jejich zpracování v libovolné aplikaci.
- Neméně důležité je věnovat pozornost čištění dat předcházející vlastní analýze.
- Každá chyba, která vznikne nebo není nalezeno ve fázi přípravy dat se promítne do všech dalších kroků a může zapříčinit neplatnost výsledků a nutnost opakování analýzy.

DATA – ukázka uspořádání datového souboru

Parametry, znaky, charakteristiky, proměnné



Záznamy



Pacient	Clovek	aLeu cell.10 ⁶ /	aTy% %	aSe% %	aNeu% %	aLy% %	aTy cell.10 ⁶ /	aSe cell.10 ⁶ /	aNeu cell.10 ⁶ /	aLy cell.10 ⁶ /	aHtc %	aCLsk mV.s.10 ³	aCLNeus mV.s.10 ³	aCLOZ mV.s.10 ³	aCLNeuO mV.s.10 ³
3	1	4									33	72		32	
4	2	7,6	8	58	66	24	0,6	4,4	5,0	1,8	33	95	19	48	10
8	3	4	3	52	55	40	0,1	2,1	2,2	1,6	22	77	35	33	15
11	4	6,1	5	59	64	35	0,3	3,6	3,9	2,1	33	103	26	49	13
12	5	6,9	3	85	88	9	0,2	5,9	6,1	0,6	37	81	13	45	7
14	6	5,9	15	55	70	19	0,9	3,3	4,1	1,1	32	137	33	61	15
16	7	8	18	75	93	7	1,4	6,0	7,4	0,6	34	151	20	59	8
20	8	9,6	3	72	75	23	0,3	6,9	7,2	2,2	40	77	11	38	5
21	9	6	10	67	77	19	0,6	4,0	4,6	1,1	32	120	26	52	11
22	10	3,3	4	55	59	39	0,1	1,8	2,0	1,3	28	81	42	24	12
37	11	3,8	10	60	70	30	0,4	2,3	2,7	1,1	32	111	42	29	11
38	12	6,4	2	76	78	17	0,1	4,9	5,0	1,1	25	366	73	115	23
39	13	6,8	1	57	58	39	0,1	3,9	3,9	2,7	20	234	59	71	18
49	14	8,5	7	67	74	26	0,6	5,7	6,3	2,2	30	156	25	108	17
51	15	9,3	7	57	64	35	0,7	5,3	6,0	3,3	35	129	21	23	4
52	16	2,2	10	56	66	34	0,2	1,2	1,5	0,7	33	46	30	12	8
55	17	9,9	3	78	81	10	0,3	7,7	8,0	0,1	30	189	24	140	18
56	18	5	2	80	82	13	0,1	4,0	4,1	0,7	26	101	25	54	13
6	1	8,8	11	72	83	12	1,0	6,3	7,3	1,1	44	268	36,6	145	19,9
9	2	9,2	2	66	68	28	0,2	6,1	6,3	2,6	42	168	26,9	76	12,2
13	3	10,0	7	83	90	8	0,7	8,3	9,0	0,8	54	181	20,1	81	9
15	4	9,6	1	75	76	23	0,1	7,2	7,3	2,2	45	343	47	124	16,9
17	5	6,0									45	40		21	

Datová tabulka a její možné problémy

Jednoznačné ID nezbytné pro identifikaci a případné propojení do dokumentace.

Sloupec nesmí obsahovat kombinaci textu a čísel.

Chybně uvedeno datum.

Překlep v názvu kategorie, při zpracování dat se chová jako nová kategorie.

Nereálné odlehle hodnoty, pravděpodobně prohozen věk a výška.

Uvedena 0 zřejmě namísto chybějící hodnoty, je třeba ponechat prázdnou buňku.

Je třeba uvádět v samostatných sloupcích pro diastolický a systolický tlak.

Kombinace dvou možných kategorizací (0/1 nebo N/A), je třeba si vybrat jednu z nich.

ID	Pohlaví	Věk	Výška	Zařazen	Alergie	TKD/TKS
9	M	53	177	13.9.2001	N	80/120
14	M	41	167	10.9.2001	N	75/119
19	M	52	182	14.90.2001	N	91/145
22	M	26	193	17.9.2001	A	78/130
23	MM	53	neznámo	17.9.2001	N	80/120
29	M	23	197	4.10.2001	0	75/119
30	M	58	158	4.10.2001	N	91/145
32	Z	198	45	5.10.2001	N	78/130
33	Z	51	191	5.10.2001	1	80/120
34	M	44	169	5.10.2001	1	75/119
35	Z	22	0	5.10.2001	N	91/145
38	M	42	163	5.10.2001	A	78/130

Zásady pro ukládání dat

- Správné a přehledné uložení dat je základem jejich pozdější analýzy
- Je vhodné rozmyslet si předem jak budou data ukládána
- Pro počítačové zpracování dat je nezbytné ukládat data v tabulární formě
- Nejvhodnějším způsobem je uložení dat ve formě databázové tabulky
 - Každý sloupec obsahuje pouze jediný typ dat, identifikovaný hlavičkou sloupce
 - Každý řádek obsahuje minimální jednotku dat (např. pacient, jedna návštěva pacienta apod.)
 - Je nepřípustné kombinovat v jednom sloupci číselné a textové hodnoty
 - Komentáře jsou uloženy v samostatných sloupcích
 - U textových dat nezbytné kontrolovat překlepy v názvech kategorií
 - Specifickým typem dat jsou datумы u nichž je nezbytné kontrolovat, zda jsou datумы uloženy v korektním formátu
- Takto uspořádaná data je v tabulkových nebo databázových programech možné převést na libovolnou výstupní tabulku
- Pro základní uložení a čištění dat menšího rozsahu je možné využít aplikací MS Office

Vizualizace dat

Typy grafické vizualizace

Rizika desinterpretace grafického zobrazení dat

Anotace

- Prvním krokem v analýze dat je jejich vizualizace.
- Různé typy dat nám umožňující získání představy o rozložení dat, zastoupení kategorií i vztazích proměnných navzájem.
- Prostřednictvím vizualizace získáváme vhled do dat a začínáme vytvářet hypotézy o zákonitostech panujících mezi proměnnými v hodnoceném souboru dat.

V čem vytvářet grafy

- Nejrozličnější software – nejrozličnější možnosti
 - MS Office – základní grafy, snadná editovatelnost, lze invenčně upravit, snadná replikovatelnost výměnou dat
 - R – různé knihovny (např. ggplot) – vyšší vstupní investice, nejrozličnější typy grafů, automatizace
 - SPSS, Statistica – rychlá tvorba velkého množství grafů, mnoho typů grafů
- Kritéria
 - Výběr různých typů grafů
 - Snadnost editace a úpravy vzhledu
 - Snadná replikovatelnost/automatizace/rychlost tvorby grafů

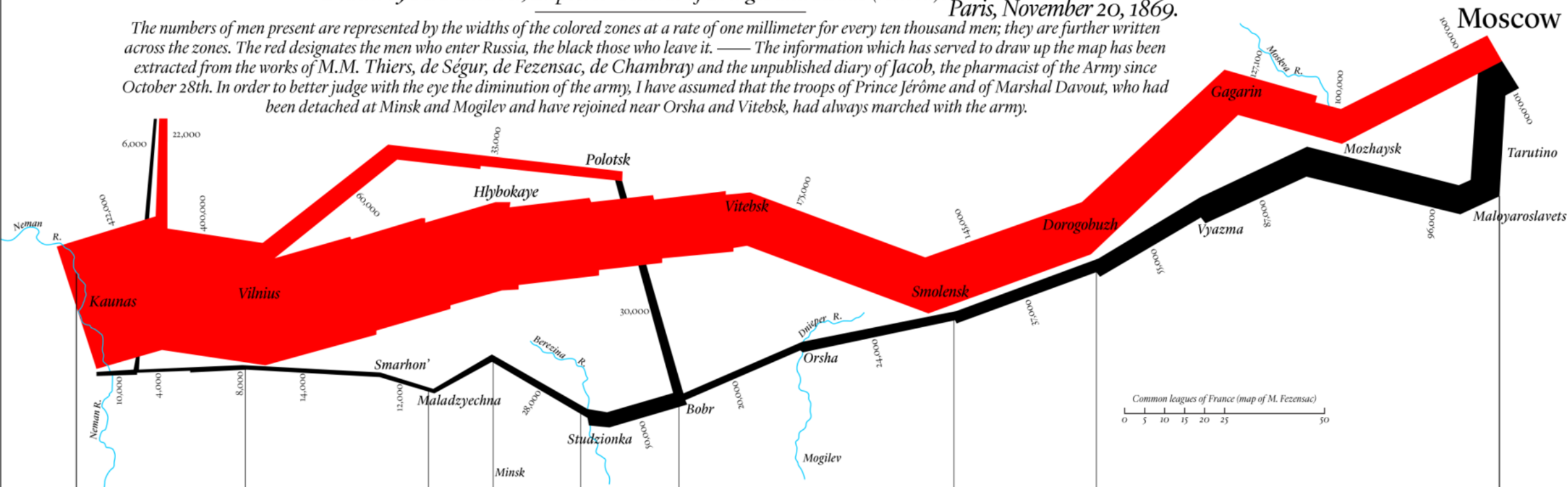
Slavné grafy: Charles Joseph Minard – Napoleonovo tažení do Ruska

Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812 ~ 1813

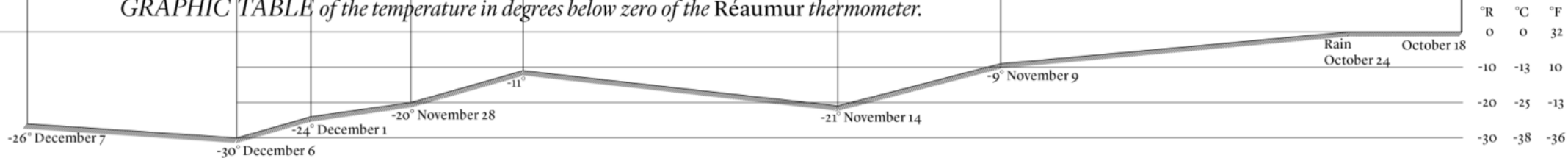
Drawn by M. Minard, Inspector General of Bridges and Roads (retired).

Paris, November 20, 1869.

The numbers of men present are represented by the widths of the colored zones at a rate of one millimeter for every ten thousand men; they are further written across the zones. The red designates the men who enter Russia, the black those who leave it. — The information which has served to draw up the map has been extracted from the works of M.M. Thiers, de Ségur, de Fezensac, de Chambray and the unpublished diary of Jacob, the pharmacist of the Army since October 28th. In order to better judge with the eye the diminution of the army, I have assumed that the troops of Prince Jérôme and of Marshal Davout, who had been detached at Minsk and Mogilev and have rejoined near Orsha and Vitebsk, had always marched with the army.



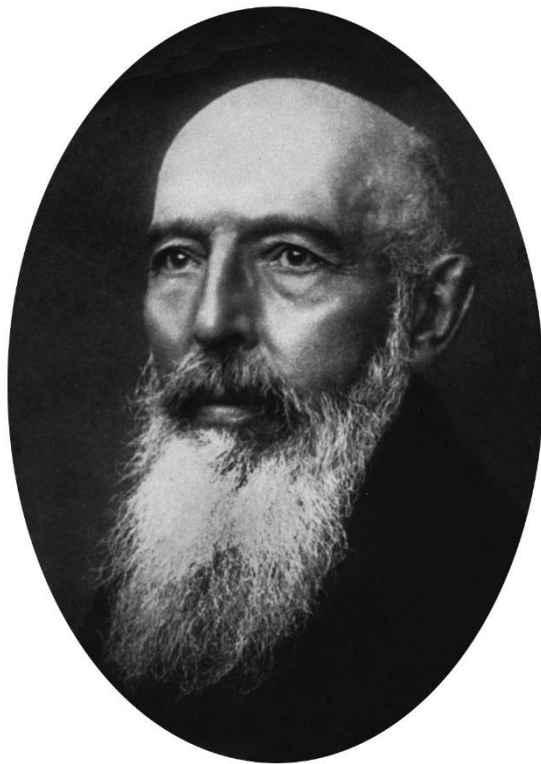
GRAPHIC TABLE of the temperature in degrees below zero of the Réaumur thermometer.



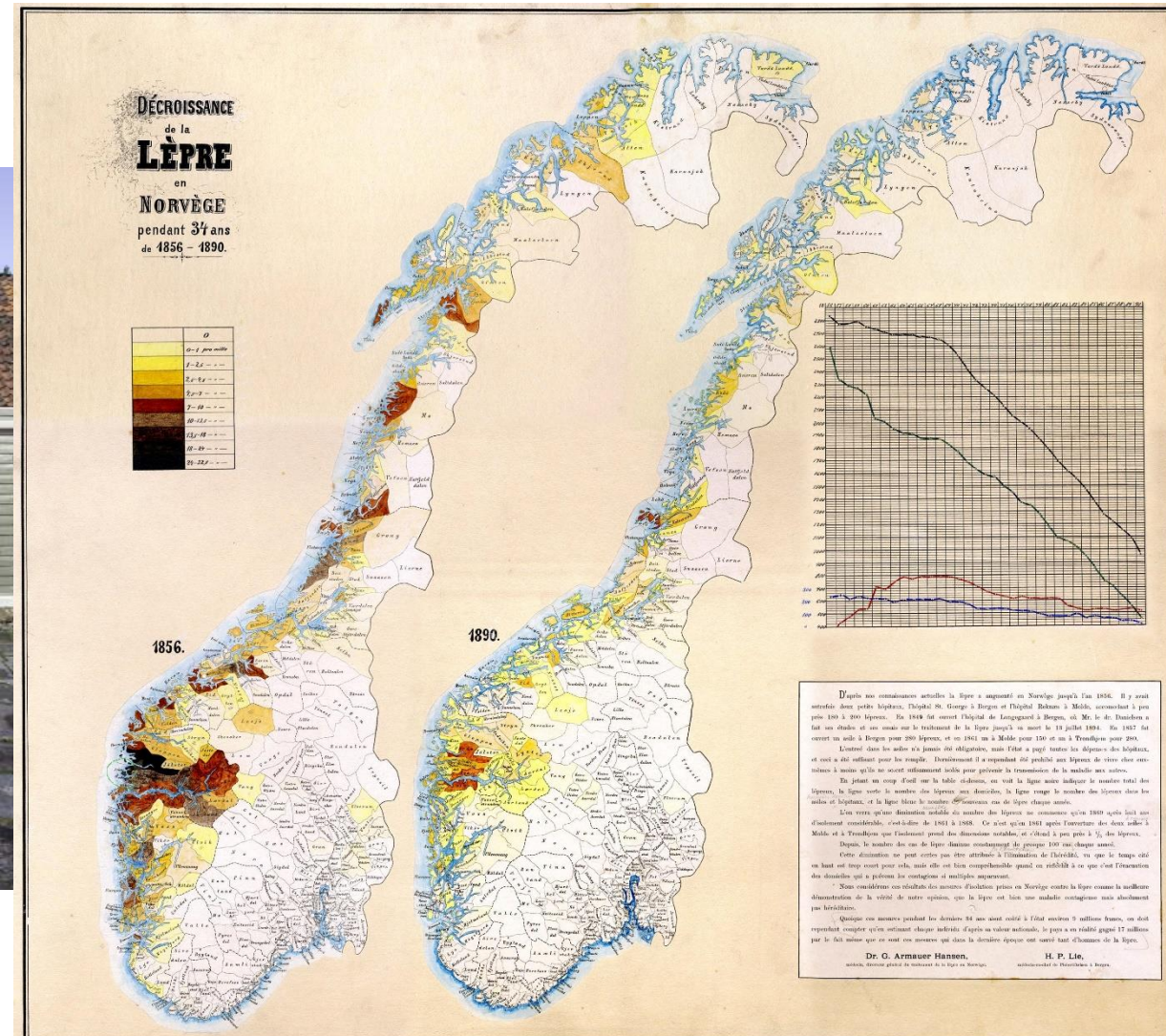
The Cossacks pass the frozen Neman at a gallop.

Slavné grafy: Eradikace lepry v Norsku

- 1856 – národní registr lepry v Norsku založen v Bergenu -> analýza získaných dat -> opatření k eradikaci lepry v Norsku
- Gerhard Armauer Hansen



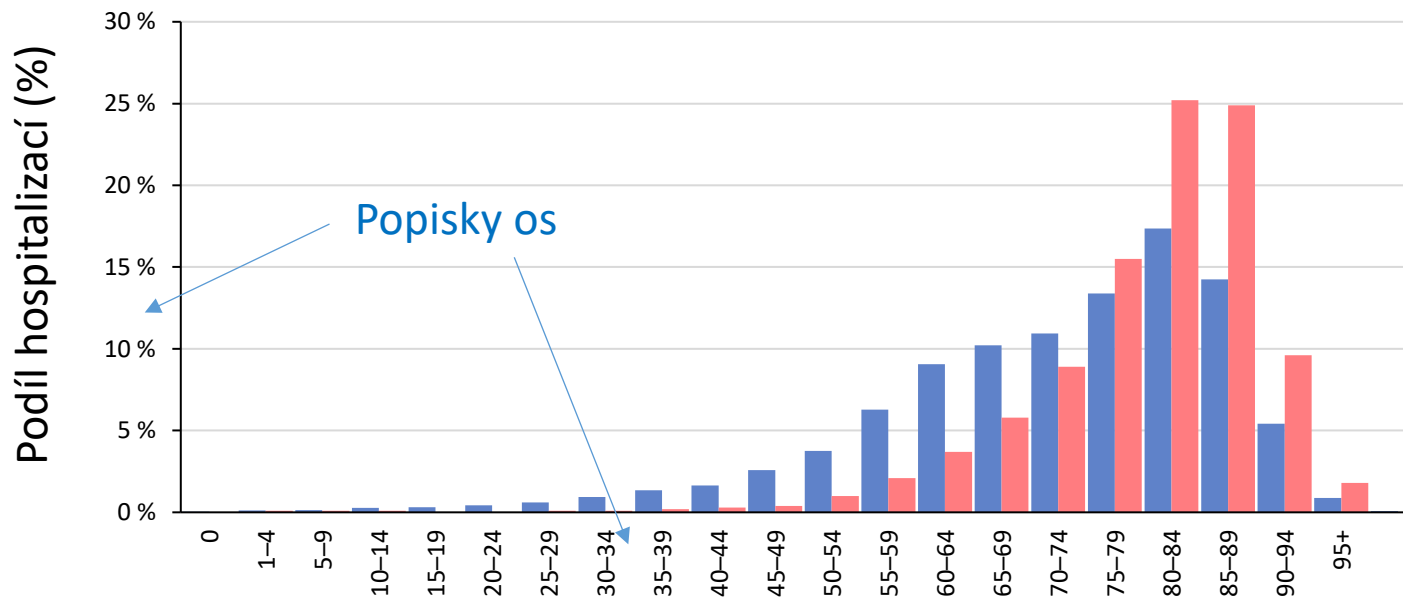
Muzeum lepry v Bergenu



Co nesmí chybět na grafu

- Každý graf musí být jednoznačně popsán – self explained
- Graf, který nic neříká, nemá smysl kreslit !!!

Věková struktura pacientů při zahájení hospitalizace



Nadpis grafu

Popis kategorií grafu

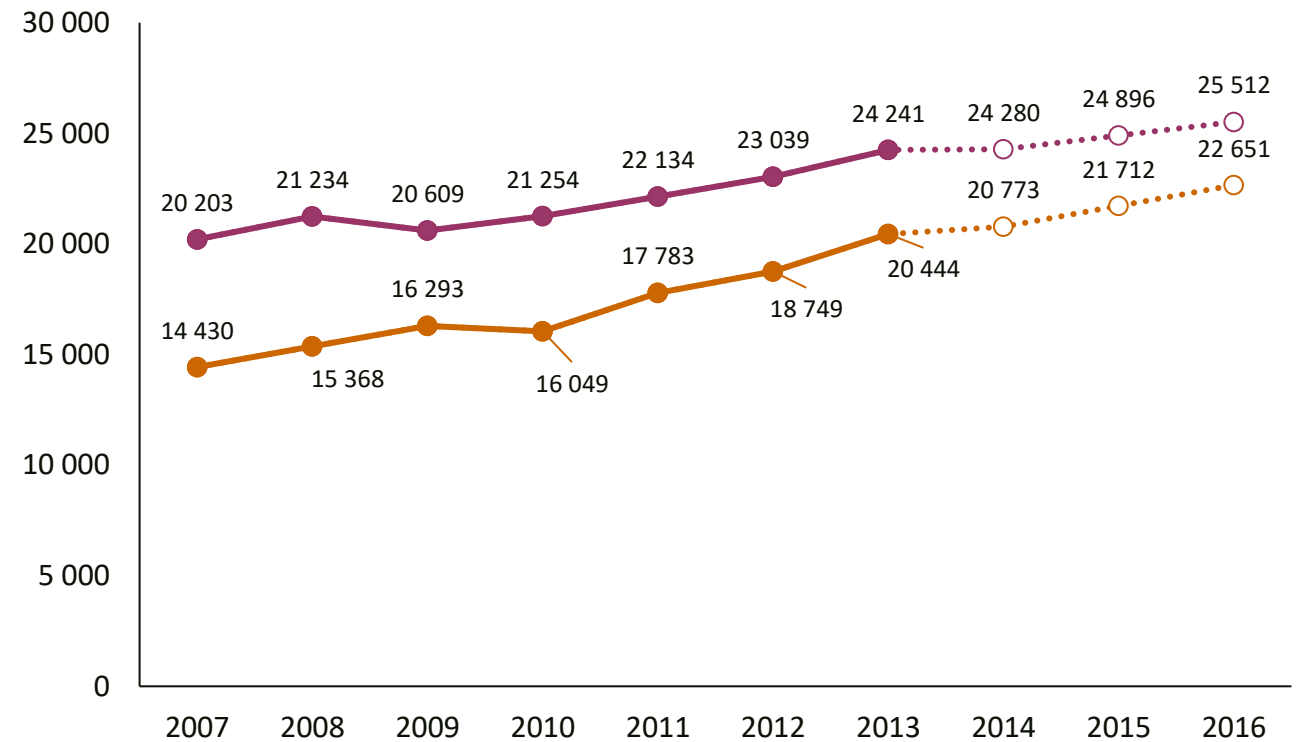
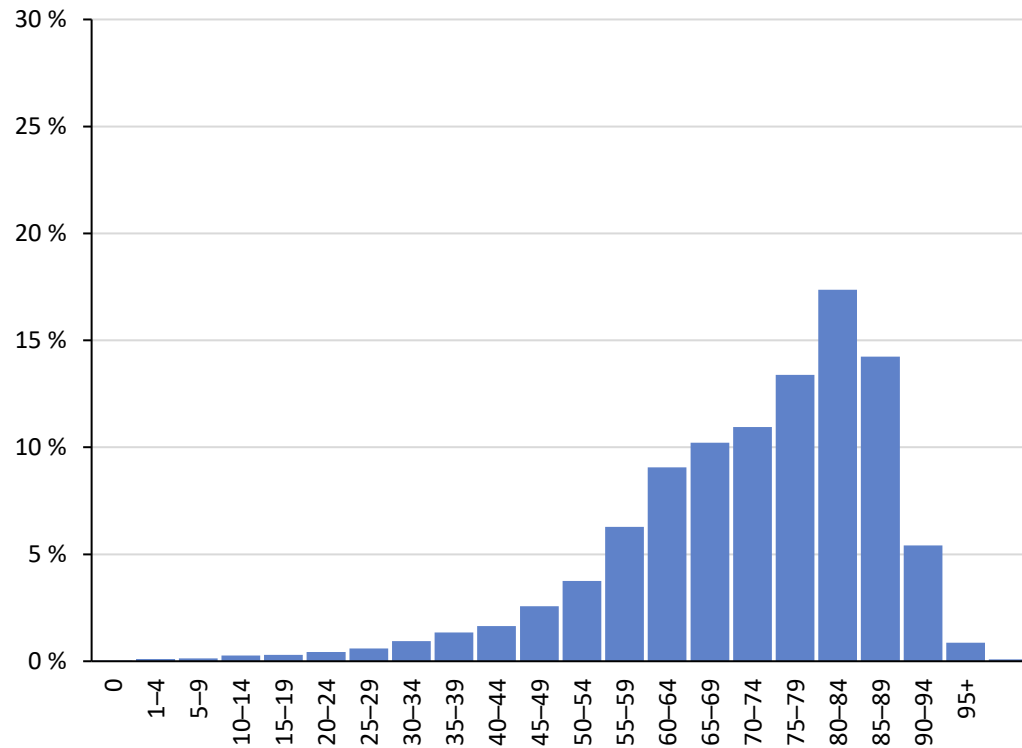
Popisky os

Nadpisy os (včetně jednotek)

Věk při zahájení hospitalizace (roky)

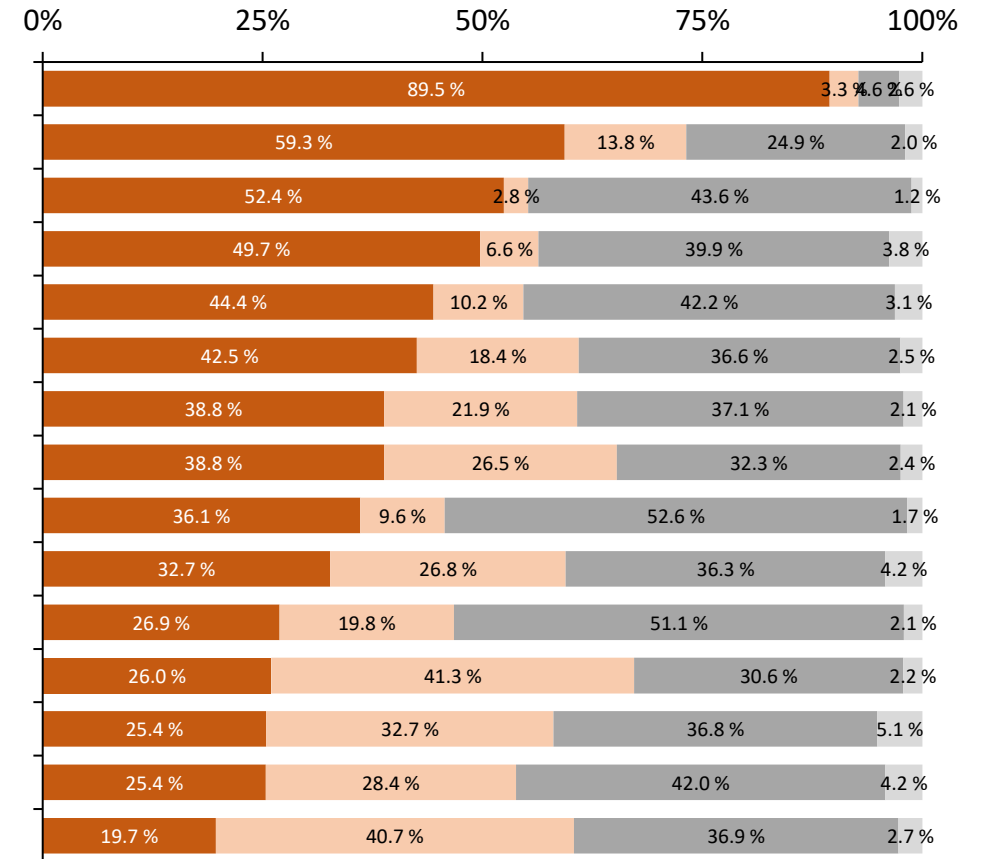
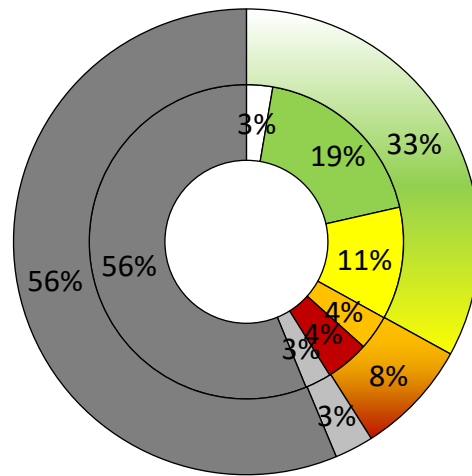
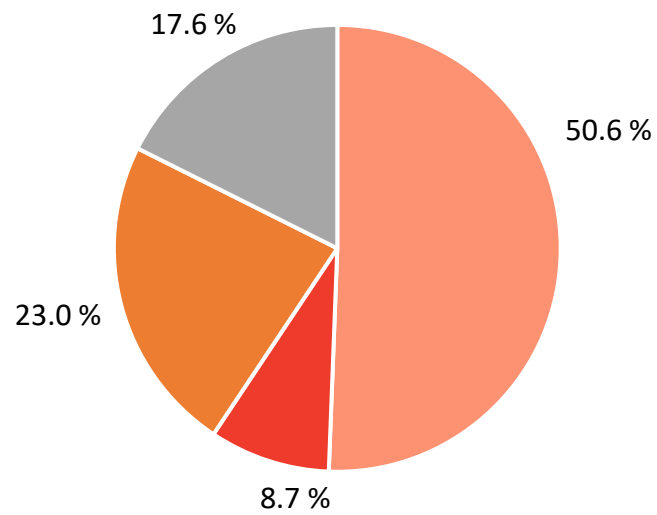
Sloupcové a čárové grafy

- Jednoduchá tvorba, vizualizace absolutních hodnot nebo procent



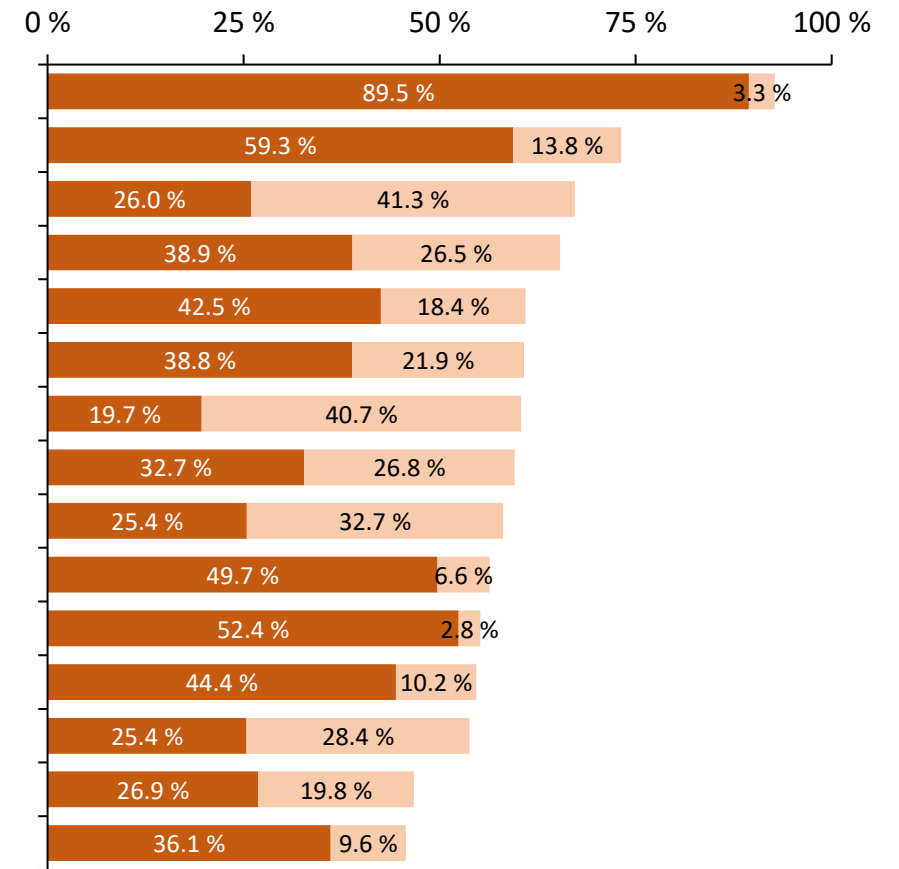
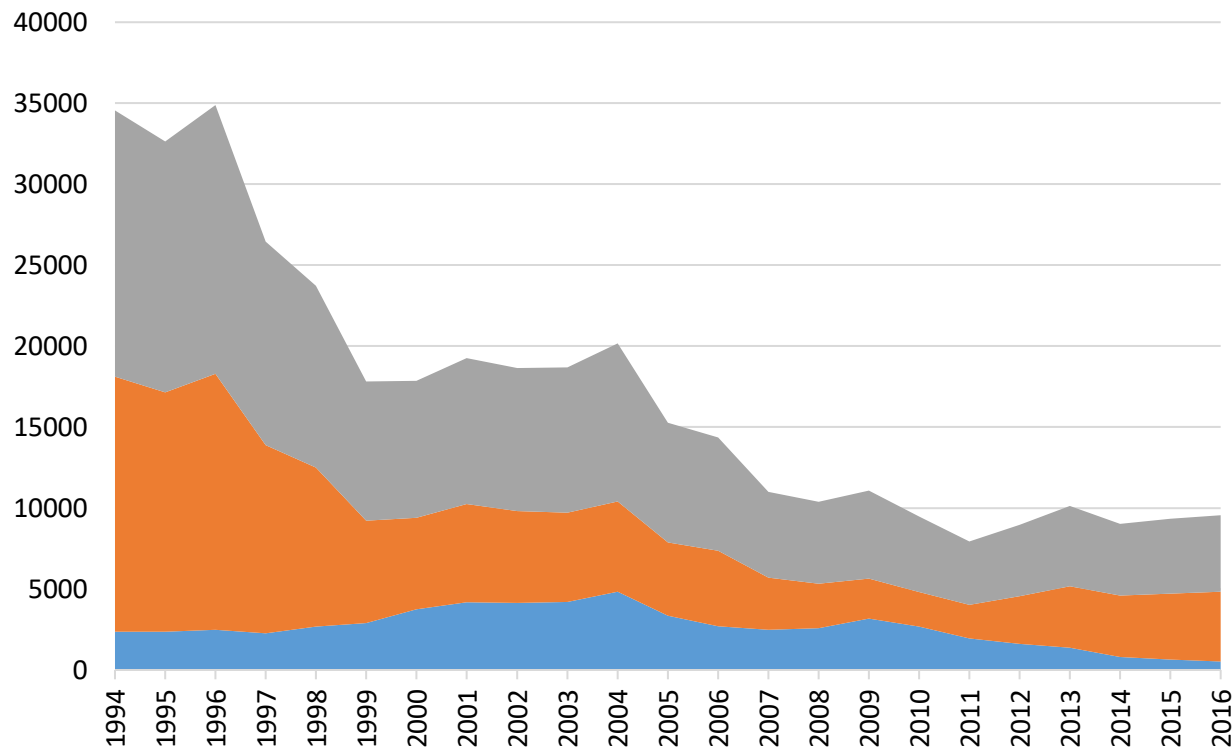
Koláčové a páskové grafy

- Jednoduchá tvorba, vizualizace procent



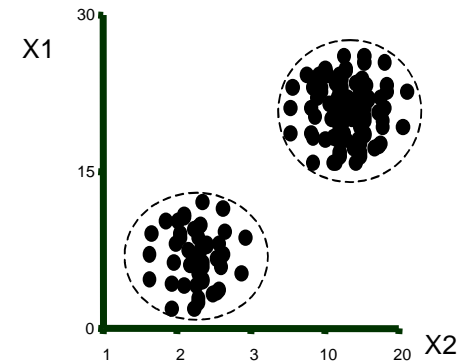
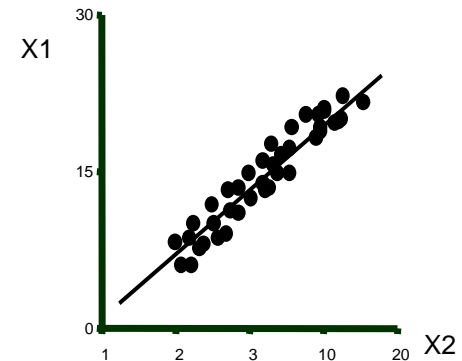
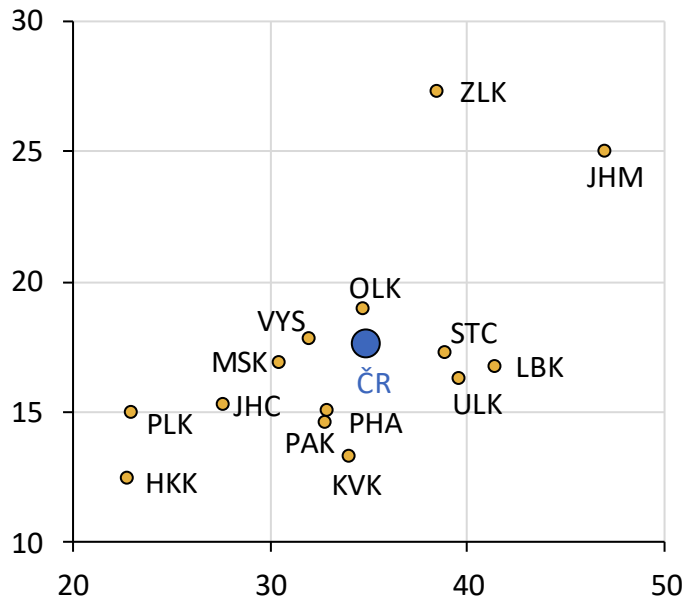
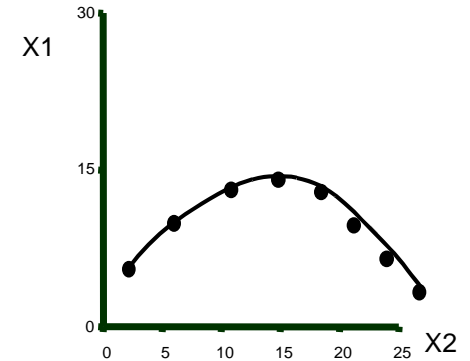
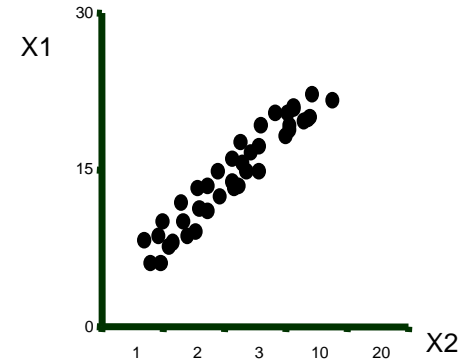
Skládané grafy

- Kumulativní zobrazení více informací



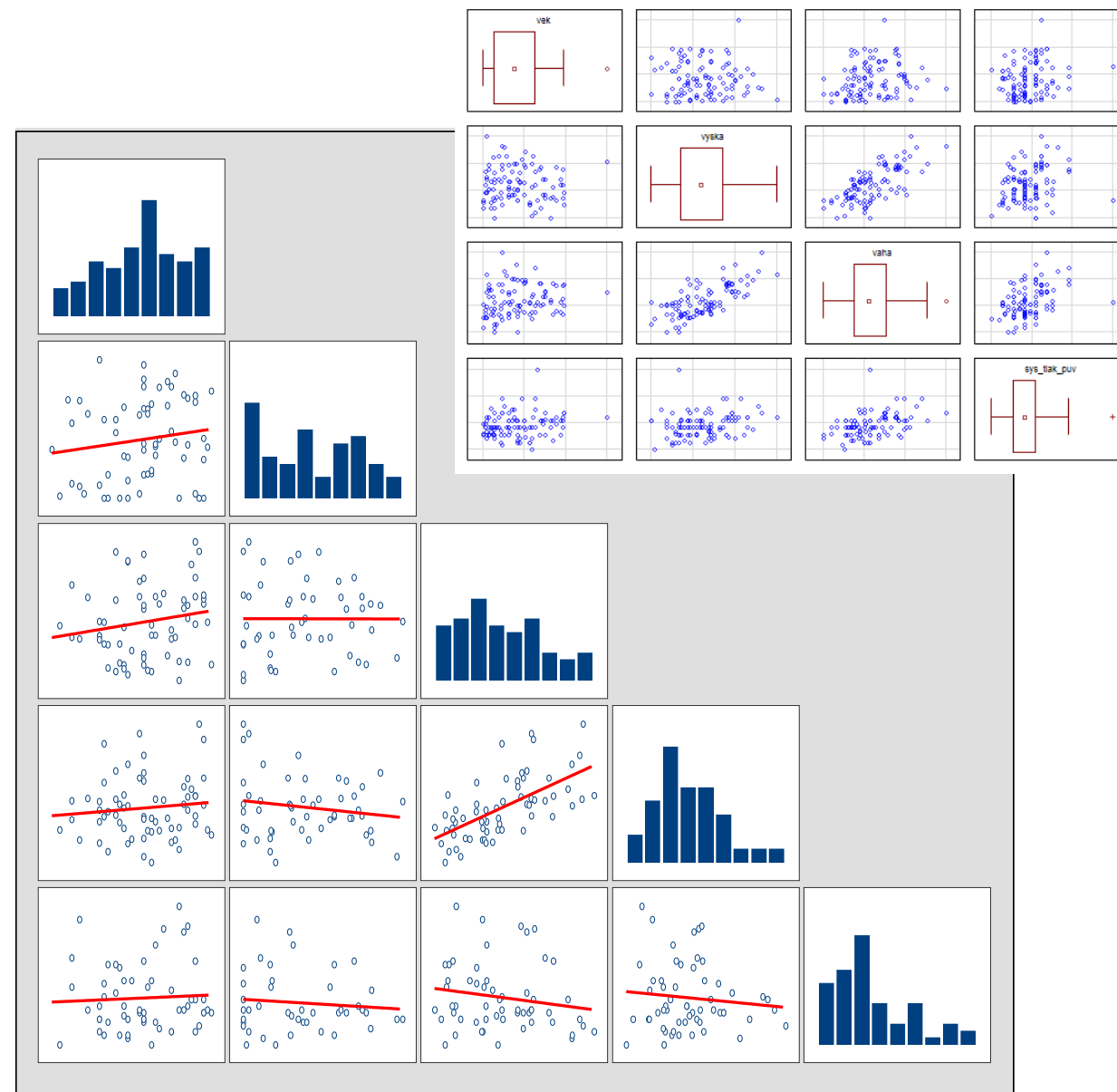
XY graf (scatter plot)

- Popis vztahu dvou spojitých proměnných
- Možnost kategorizace a popisu bodů
- Prokládání modelů do grafů
- Základní graf pro prohlídku dat před korelační a regresní analýzou



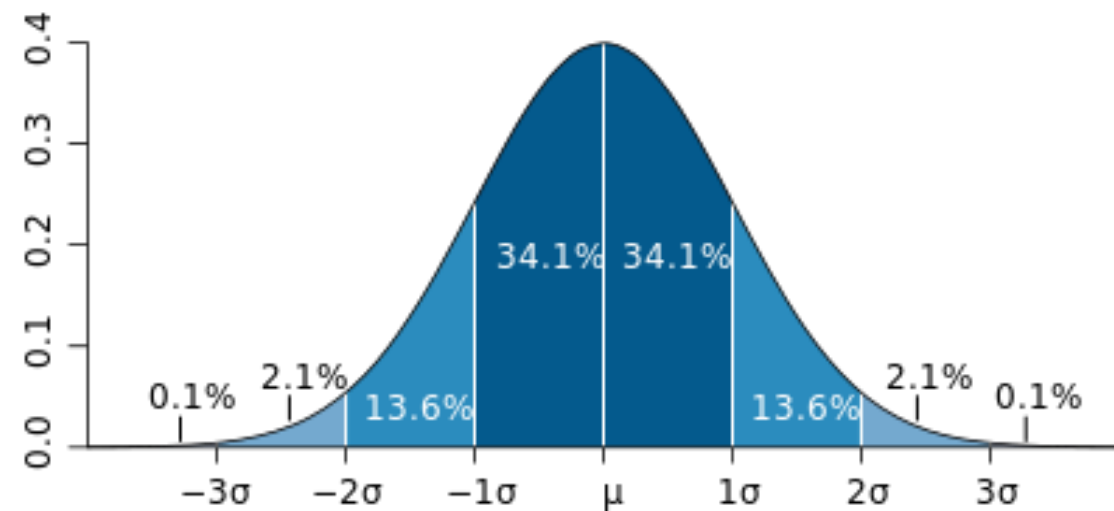
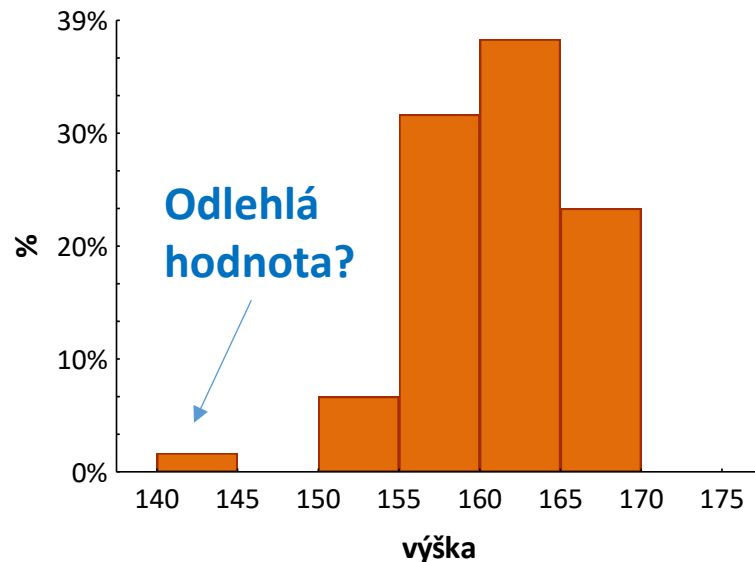
Maticový graf

- Rozšíření xy grafů ve statistických SW
- Současná vizualizace rozložení hodnot (diagonála) a vzájemných vztahů většího počtu spojitých proměnných
- Různé varianty
 - Sada proměnných každý s každým
 - Dvě sady proměnných proti sobě
 - Doplnění o výpočet korelačních koeficientů
- Základní nástroj vizualizace před vícerozměrnou analýzou



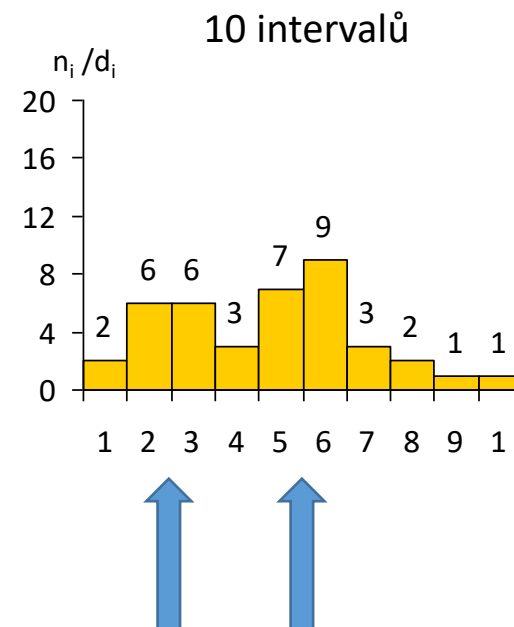
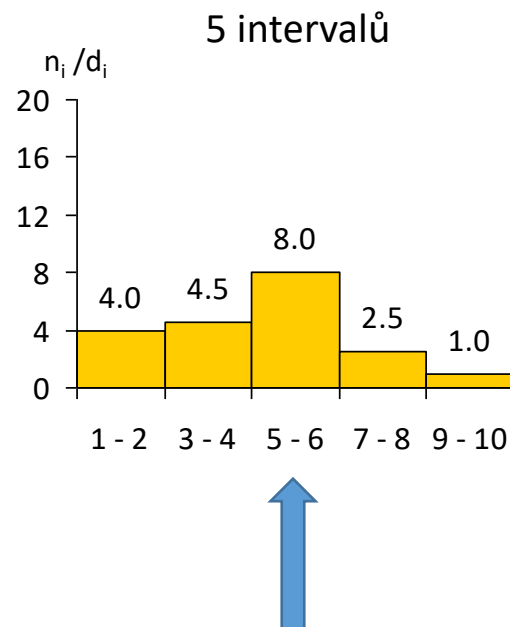
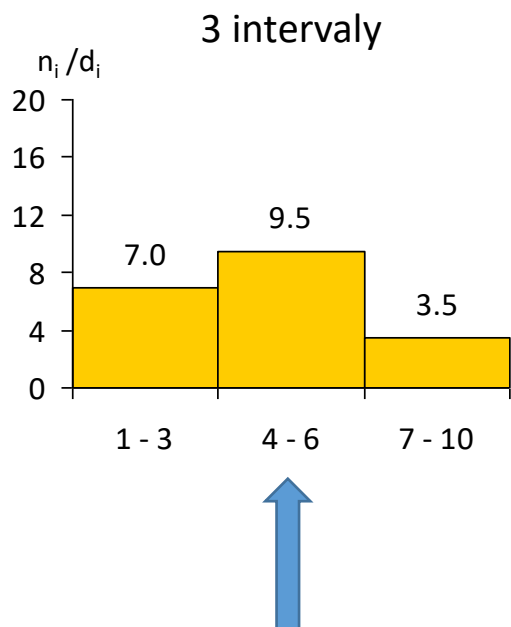
Histogram

- Graf sumarizující rozložení hodnot spojitéch proměnných, úzce spjat s teorií statistických rozdělání
- V klasické formě podobný (ale nikoliv totožný) se sloupcovým grafem
- V praxi se pod názvem histogram často skrývá sloupcový graf (přípustné pokud nevede k dezinterpretaci dat)
- Jeden ze základních grafů pro posouzení rozložení dat



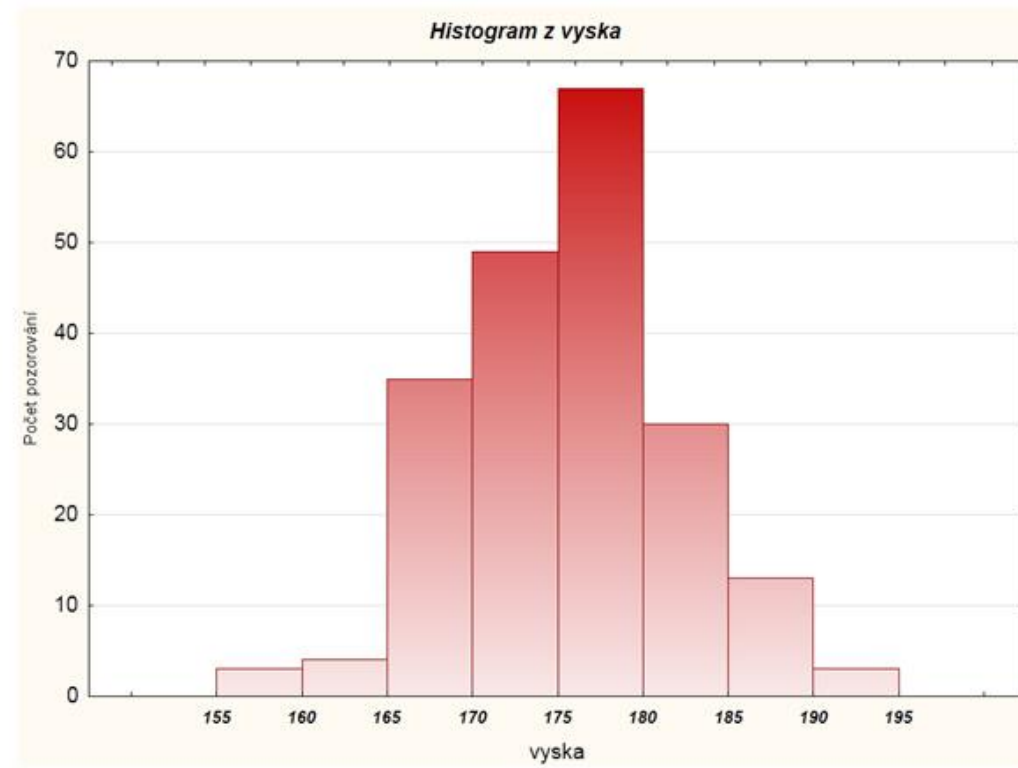
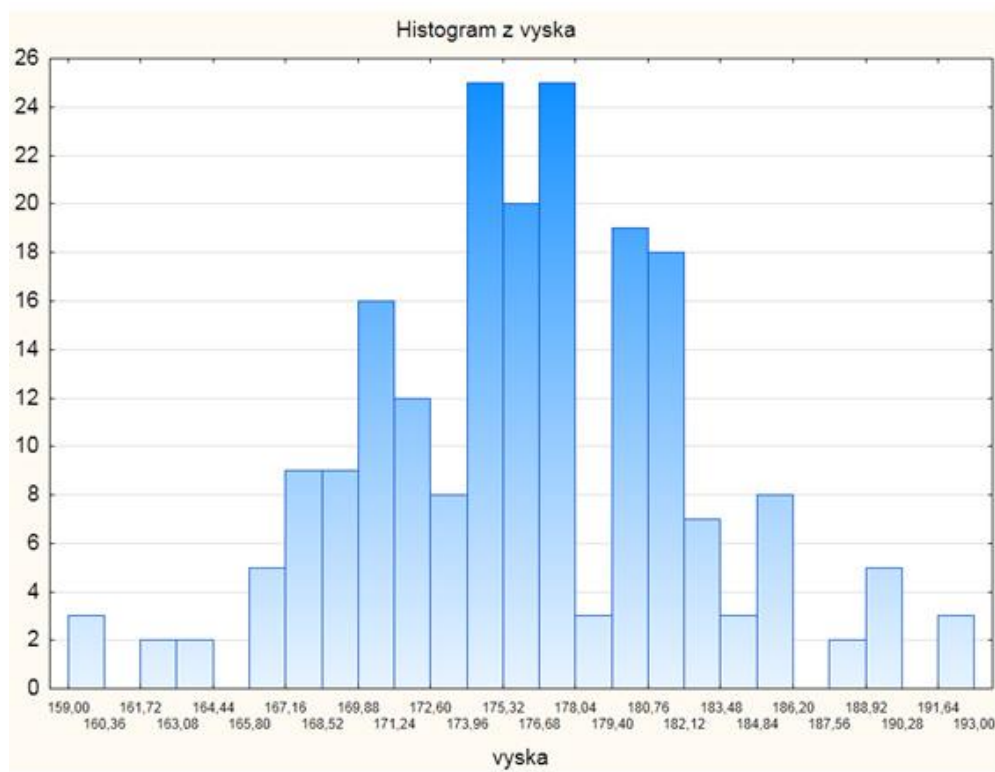
Histogram: vliv kategorizace dat

- Počtem zvolených intervalů v histogramu rozhodujeme o tom, jak bude vypadat. Při malém počtu můžeme přehlédnout důležité prvky v datech, při velkém zase může být informace roztržštěná.



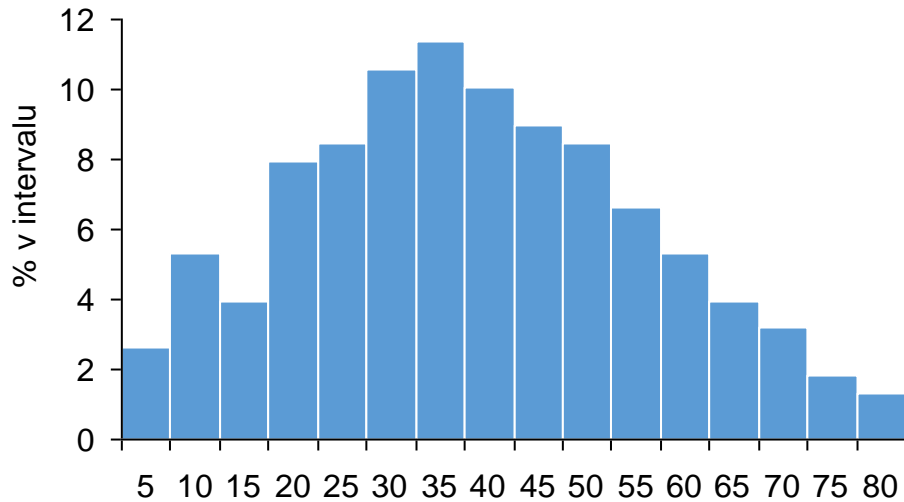
Histogram: vliv kategorizace dat

- Výběr počtu kategorií – důležitý pro interpretaci
- Ruční nebo automatický výběr – různé algoritmy (závisí na velikosti vzorku a variabilitě dat)

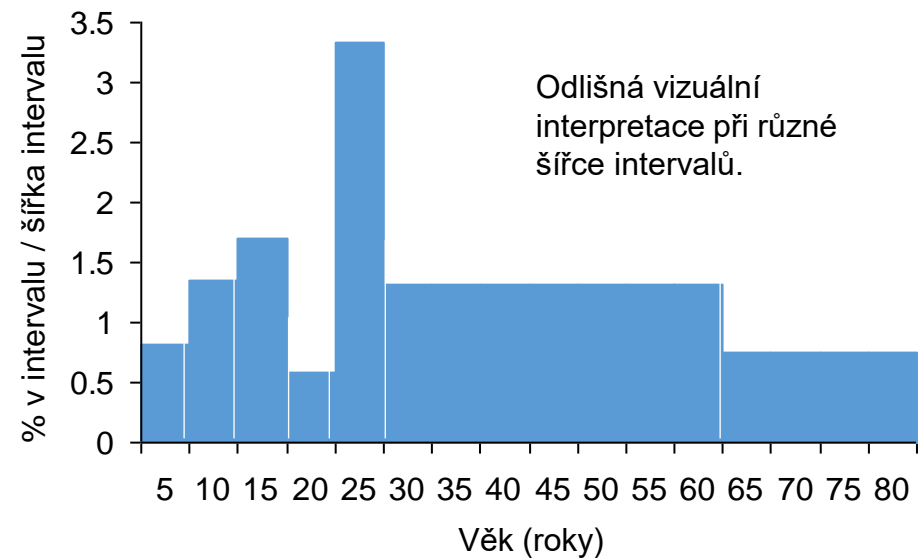
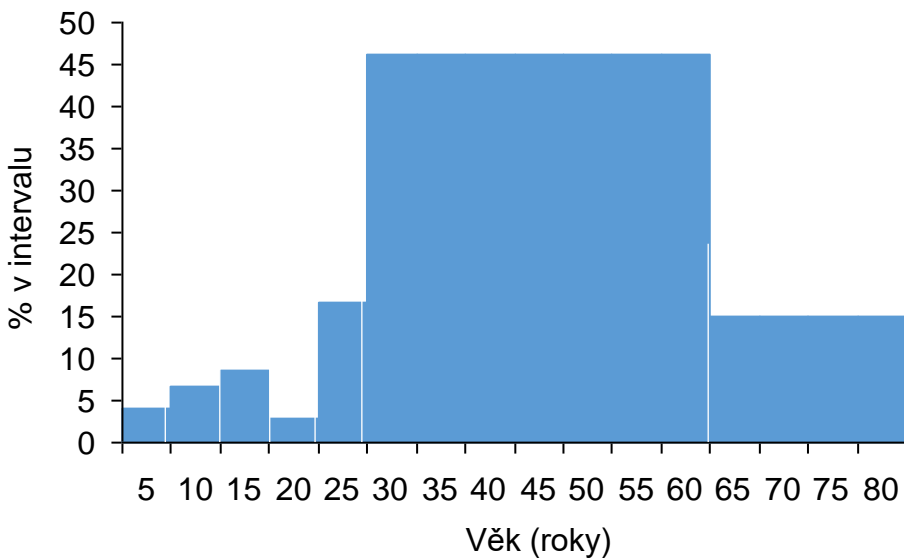
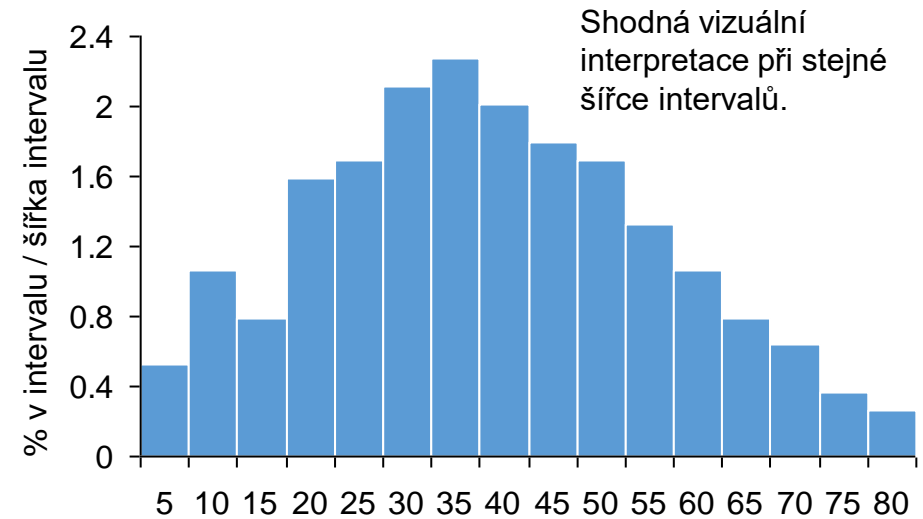


Histogram a sloupcový graf

Sloupcový graf

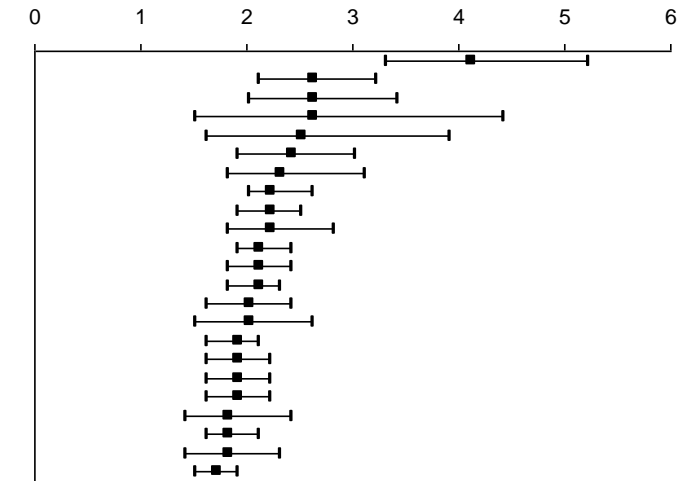
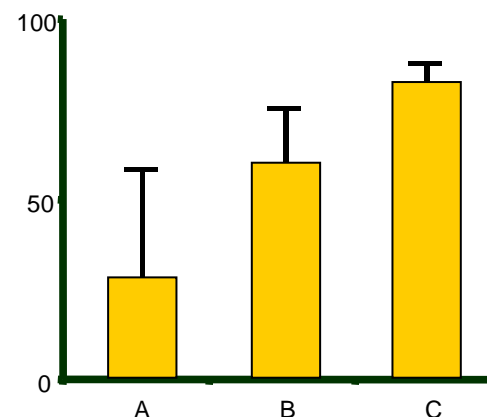
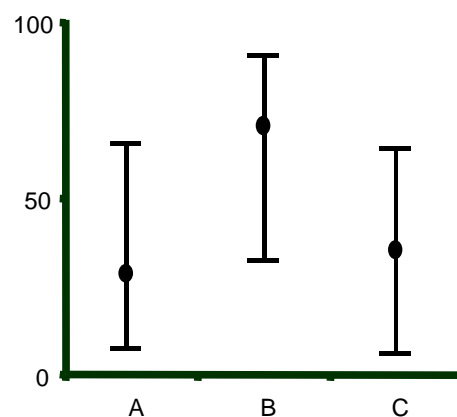
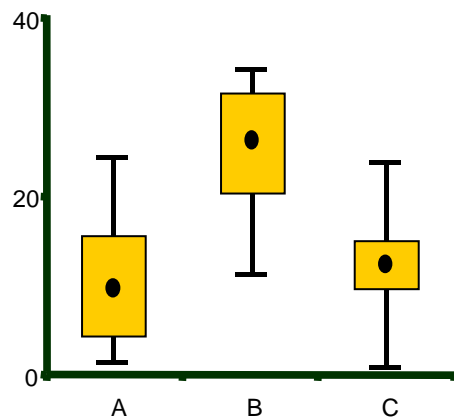


Histogram

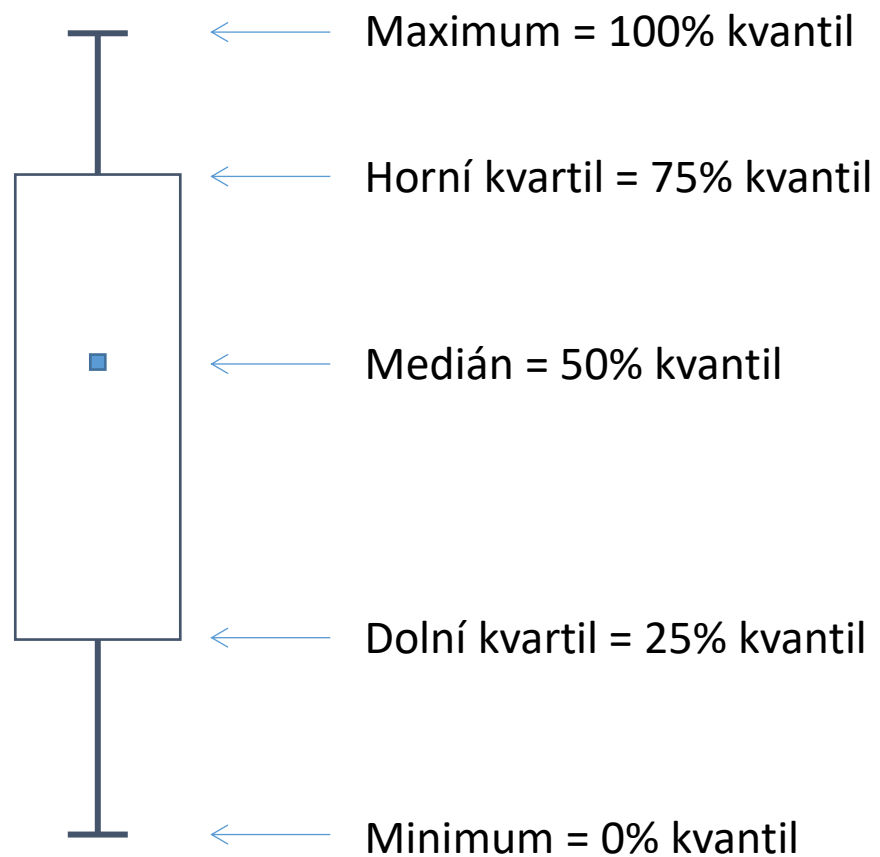


Krabicový graf – box and whisker plot: co to je?

- V analýze dat oblíbený typ grafu umožňující jednoduché srovnání více skupin objektů a hodnocení rozložení dat
- Nejběžnější pro popis spojitých dat, ale využitelný pro libovolné typy dat, které lze popsat střední hodnotou a variabilitou (procenta, regresní koeficienty, odds ratio, risk ratio, hazard ratio atd.)
- Obrovské množství variant



Krabicový graf – box and whisker plot: příklad jedné možné varianty

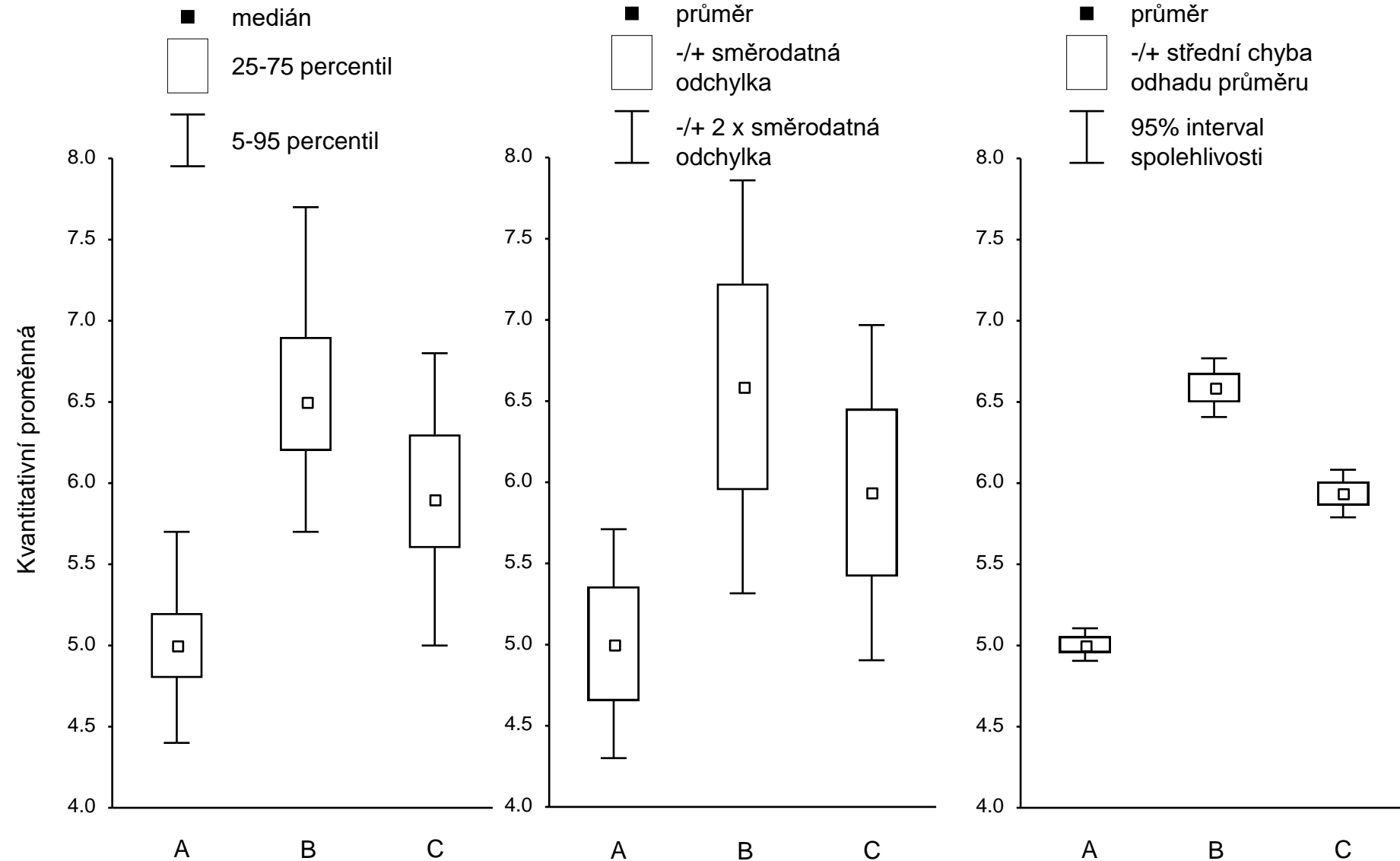


Jednotlivé body grafů mohou obsahovat libovolné popisné statistiky – průměry, směrodatné odchylky, intervaly spolehlivosti, odds ratio, hazard ratio atd.

Počet datových bodů v grafu může být od tří do např. devíti.

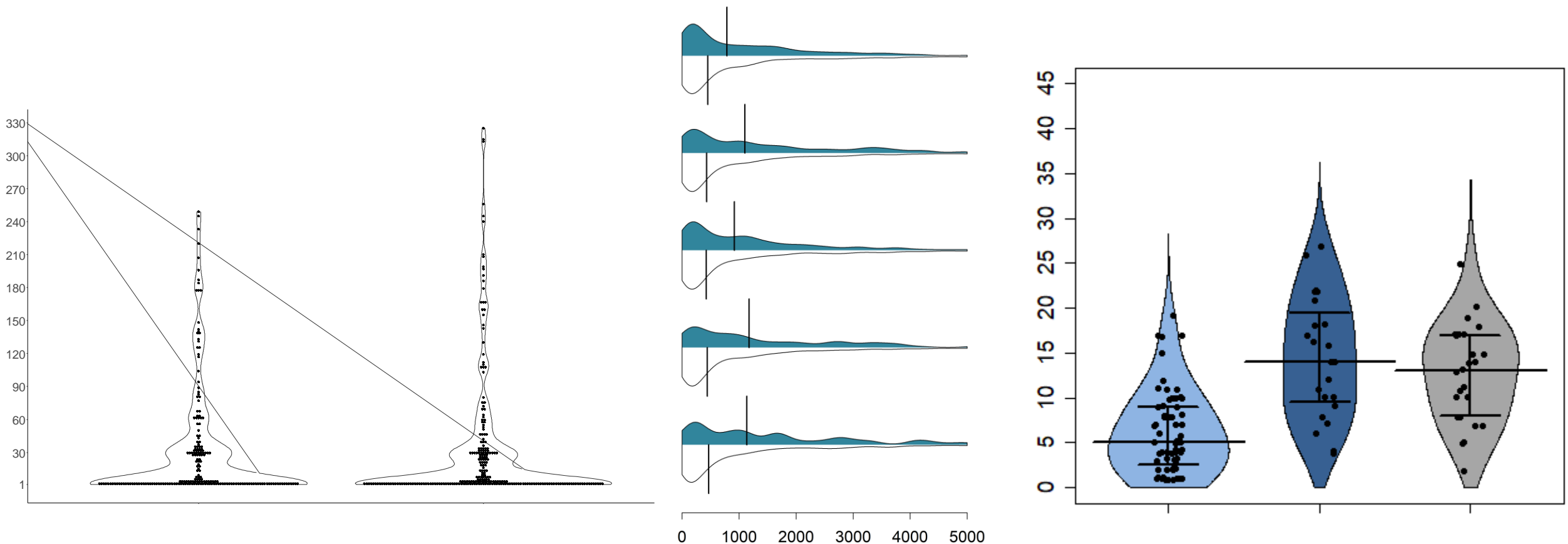
Box and whisker plot a jeho různé varianty I

- Je nezbytné číst popisky
- Různé varianty grafu mohou mít zcela jinou interpretaci



Box and whisker graf a jeho různé varianty II: Violin plot a Beanplot

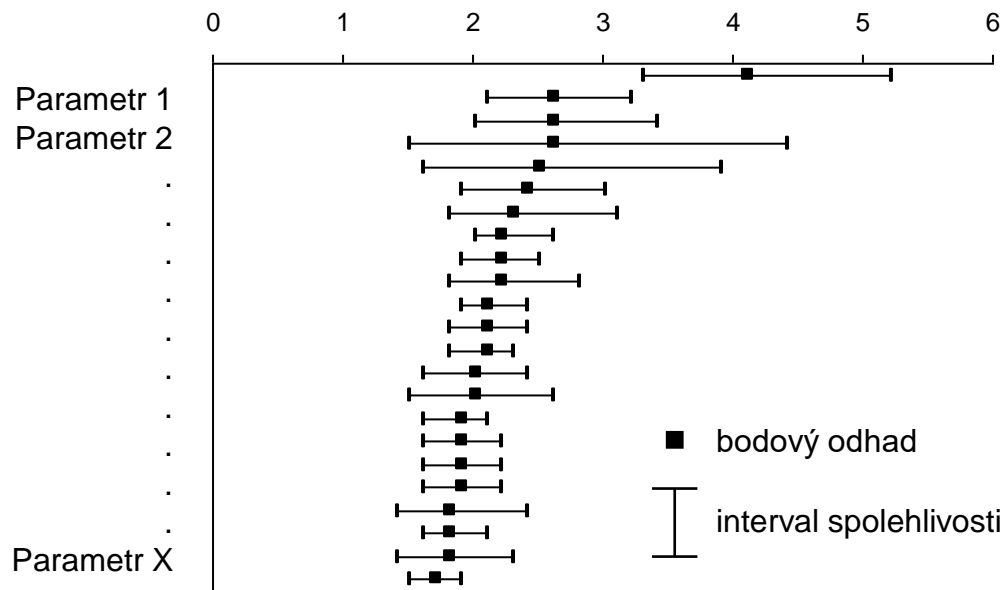
- Kombinace histogramu a box plotu nebo tečkového grafu
- K dispozici v R – např. knihovny beanplot a ggplot2



Box and whisker graf a jeho různé varianty III: Forest plot

- Varianta box and whisker plotu
- Často používaná pro zobrazení regresních koeficientů nebo odds/risk/hazard ratí

Hodnocená charakteristika (průměr, podíl, poměr šancí, relativní riziko, poměr rizik)

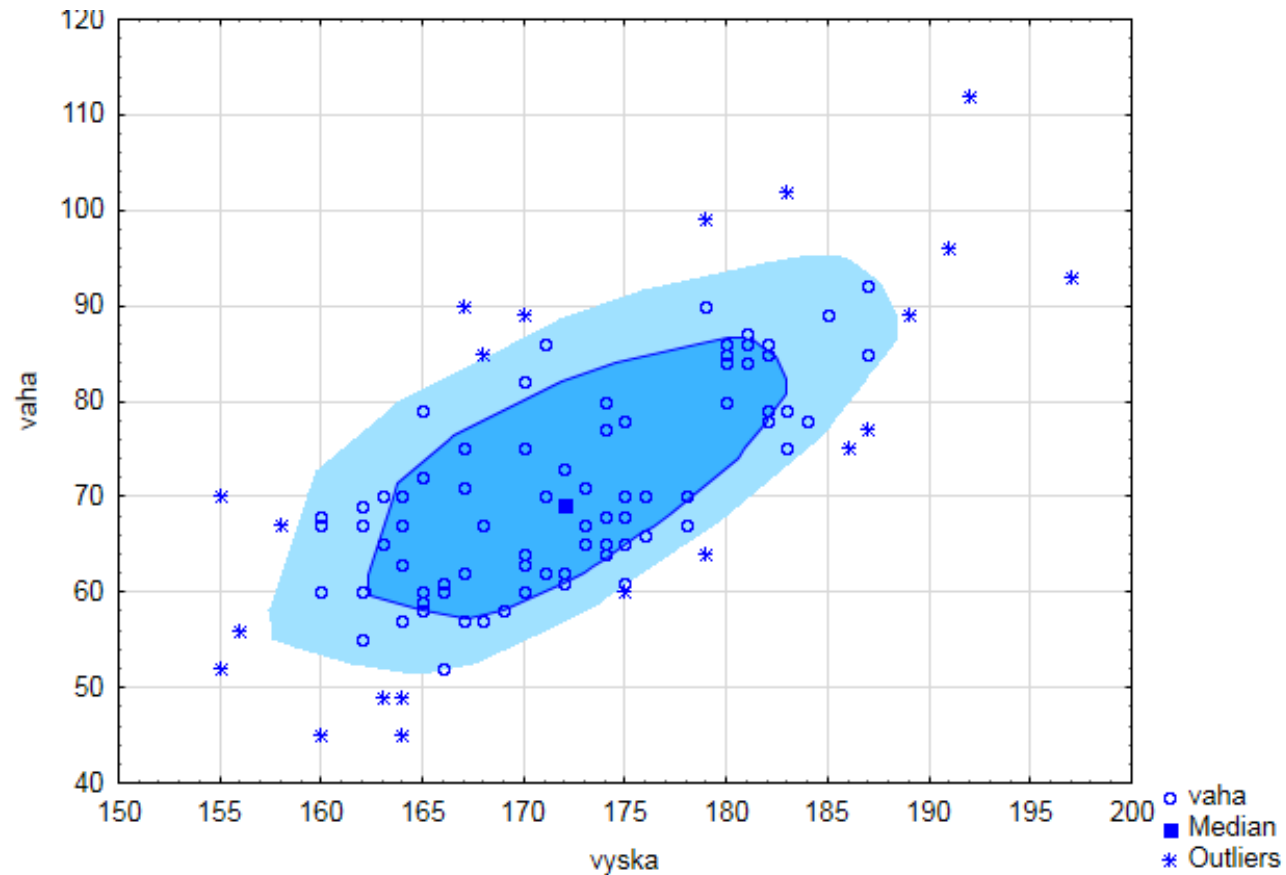


Variable	Subgroup	N		Median PFS (months)		HR
		Placebo-Rd	IRd	Placebo-Rd	IRd	
All patients	ALL	362	360	14.7	20.6	0.742
Age (yrs)	≤65	176	168	14.1	20.6	0.683
	>65-75	125	145	17.6	17.5	0.833
	>75	61	47	13.1	18.5	0.868
ISS stage (stratification factor)	I or II	318	314	15.7	21.4	0.746
	III	44	46	10.1	18.4	0.717
Cytogenetic risk	Standard-risk	216	199	15.6	20.6	0.640
	High-risk	62	75	9.7	21.4	0.543
Number of prior therapies	1	217	224	15.9	20.6	0.832
	2	111	97	14.1	17.5	0.749
	3	34	39	10.2	NE	0.366
Proteasome inhibitor	Exposed	253	250	13.6	18.4	0.739
	Naive	109	110	15.7	NE	0.749
Prior IMiD therapy	Exposed	204	193	17.5	NE	0.744
	Naive	158	167	13.6	20.6	0.700
Refractory to last prior therapy	Yes	55	59	NE	NE	0.712
	No	307	301	14.1	20.6	0.742
Relapsed or refractory	Relapsed	280	276	15.6	18.7	0.769
	Refractory	40	42	13.0	NE	0.784
	Ref & rel	42	41	13.1	NE	0.506

Moreau P et al. ASH 2015, oral presentation. Abstract #727

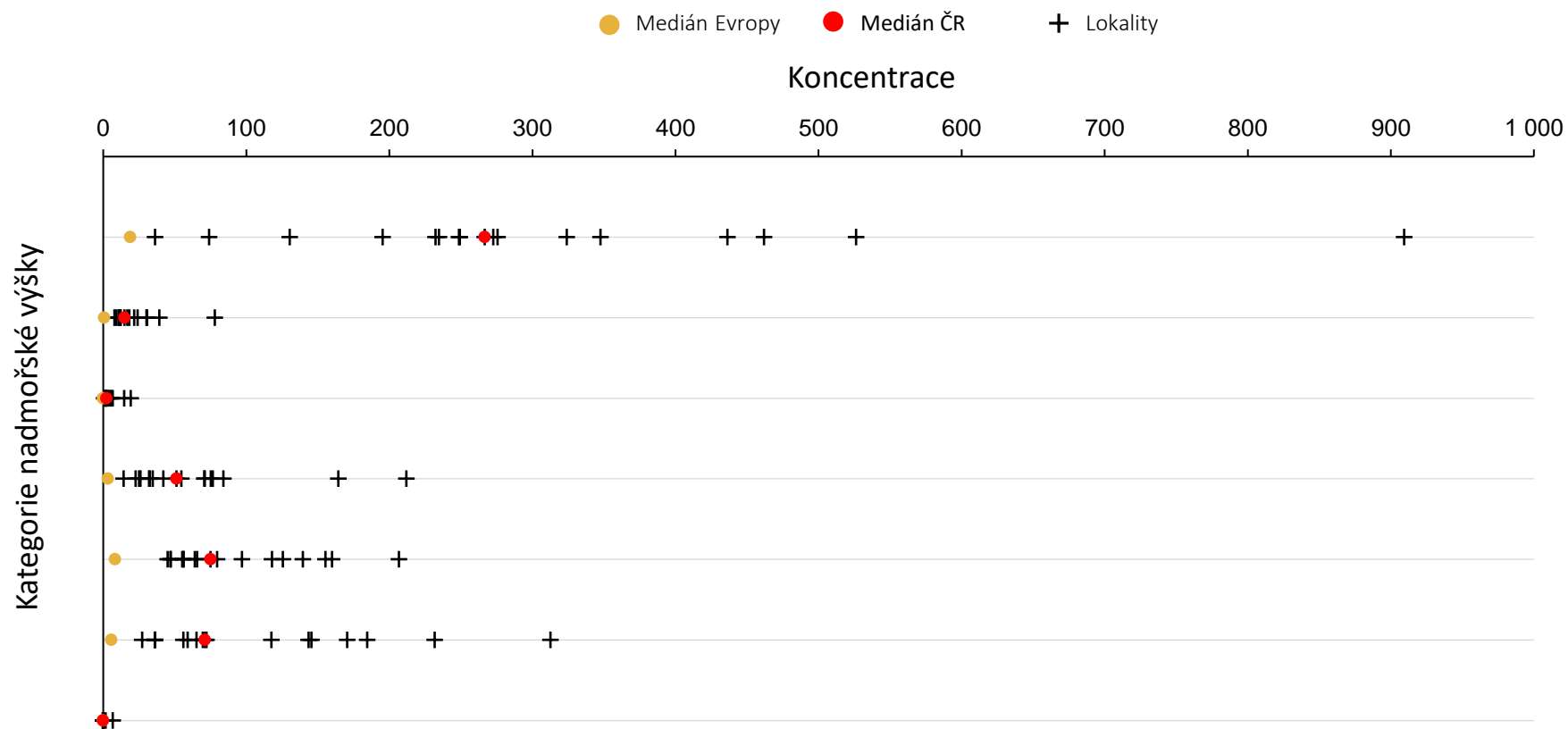
Box and whisker graf a jeho různé varianty IV: Bagplot

- Bagplot = „bivariate boxplot“ (tzn. „dvourozměrný krabicový graf“)



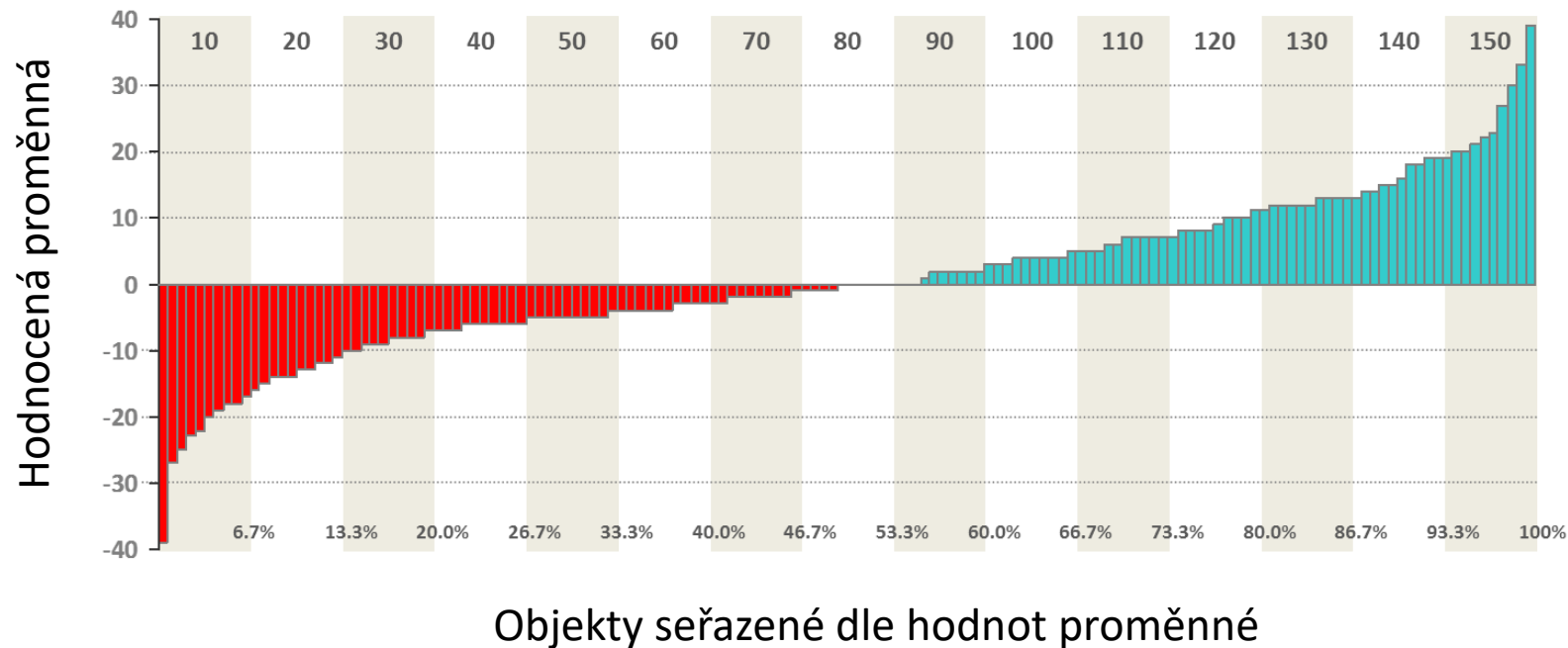
Invenční využití jednoduchých grafů: Korálkový graf

- Lze vytvořit z XY grafu v MS Office
- Velké množství informace na malé ploše



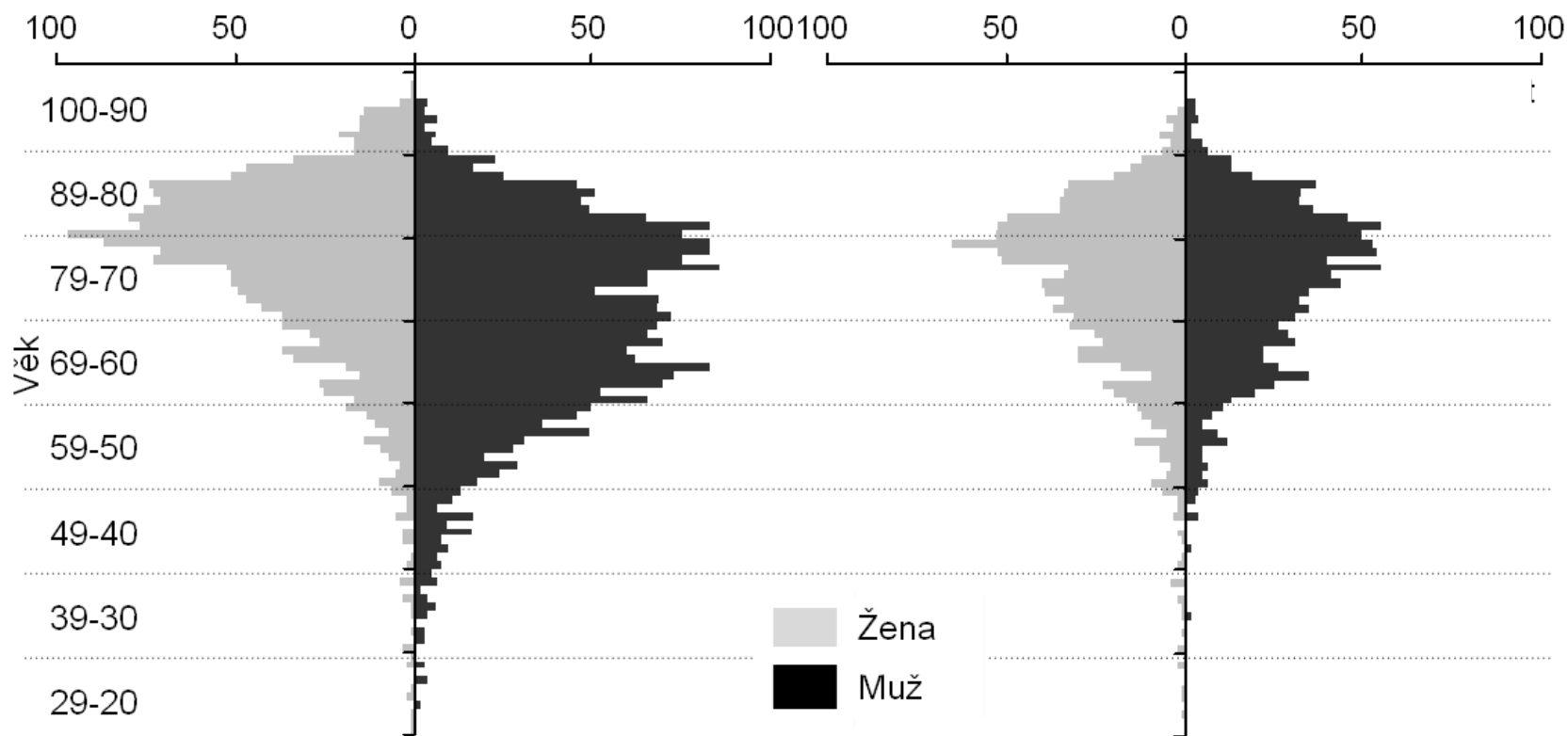
Invenční využití jednoduchých grafů: Waterfall plot

- Vizualizace výsledků individuálních objektů, často u proměnných popisujících změny
- Hodnoty jsou v grafu seřazeny dle velikosti
- Může být doplněn o hodnoty norem, procenta objektů v kategoriích normy apod.



Invenční využití jednoduchých grafů: Demografická pyramida

- Jednoduchý ležatý sloupečkový graf
- Atraktivní vizualizace pro srovnání dvou skupin objektů



Excel – podmíněné formátování jako grafy

- Pro zpřehlednění excelových tabulek je možné využít grafické prvky v jeho buňkách
- Datové pruhy a barevné škály

The image illustrates the application of conditional formatting in Excel. It shows three stages: the 'Podmíněné formátování' (Conditional Formatting) menu, a data table with conditional formatting applied, and the 'Barevné škály' (Color Scales) sub-menu.

Podmíněné formátování menu:

- Pravidla zvýraznění buněk
- Pravidla pro nejvyšší či nejnižší hodnoty
- Datové pruhy**
- Barevné škály
- Sady ikon
- Nové pravidlo...
- Vymazat pravidla
- Spravovat pravidla...

Tabulka s podmíněným formátováním:

	M	N	O	P	Q	R	S	T	U
	10		1	2	3	4	5	6	
	15		2	3	4	5	6	7	
	1		3	4	5	6	7	8	
	5		4	5	6	7	8	9	
	6		5	6	7	8	9	10	
	7		6	7	8	9	10	11	
	1								
	22								

Barevné škály menu:

- Pravidla zvýraznění buněk
- Pravidla pro nejvyšší či nejnižší hodnoty
- Datové pruhy
- Barevné škály**
- Sady ikon
- Nové pravidlo...
- Vymazat pravidla
- Spravovat pravidla...
- Další pravidla...

Excel – grafy v buňkách

- Pro zpřehlednění excelových tabulek je možné využít grafické prvky v jeho buňkách
- Několik typů grafů umožňujících vizualizovat v jedné buňce datové řady
- Základní možnosti editace os a vzhledu

The screenshot shows the Excel interface with the 'Minigrafy' ribbon selected. The ribbon includes options for 'Spojnicový' (Line), 'Sloupcový' (Column), and 'Vzestupy/poklesy' (Up/Down) mini-charts. Below the ribbon, a data table is visible with columns K through W and rows 1 through 6. Each cell in the table contains a numerical value and a small chart icon. A tooltip is displayed over the 'Minigrafy' ribbon, showing the three chart options and a note: 'Minigrafy jsou malé grafy umístěné v samostatných buňkách.'

K	L	M	N	O	P	Q	R	S	T	U	V	W
10	11	12	15	16	19							
6	9	10	12	12	18							
3	5	6	9	9	17							
2	1	2	6	8	13							
-1	-2	-3	4	3	8							
-5	-4	-7	4	0	4							
-5	-7	-8	2	0	2							

Heatmapa

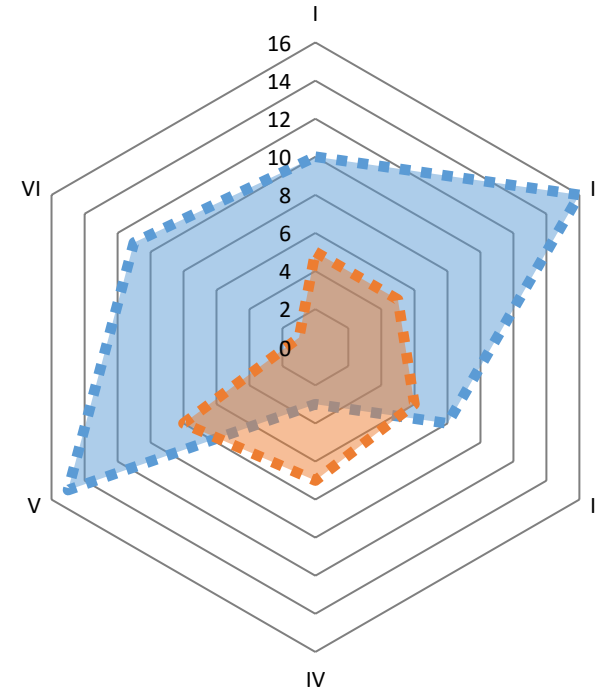
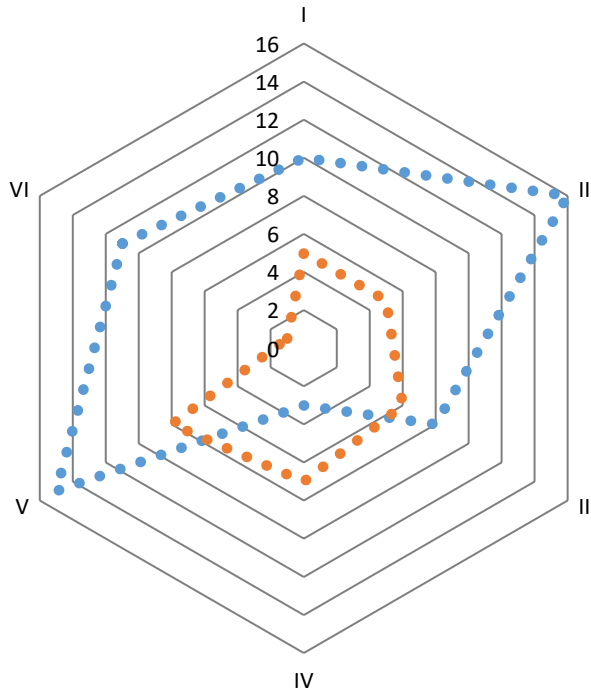
- Druh 3D grafu – osy tvoří dvě proměnné, barva třetí proměnnou
- Lze vytvořit v excelu pomocí podmíněného formátování
- Často ve vícerozměrné analýze pro vizualizaci asociačních matic

Výskyt indikátorového organismu v závislosti na dvou proměnných

Hloubka v cm vs. Koncentrace polutantu	< 60	60-69	70-74	75-79	80-84	85-89	90-94	95-99	100-109	110-119	120+
<= 30	29.8%	29.2%	27.9%	23.0%	20.5%	19.9%	20.6%	22.1%	22.1%	22.9%	23.3%
31-35	29.4%	28.2%	26.5%	22.0%	20.0%	19.5%	20.4%	21.6%	21.8%	22.6%	23.1%
36-39	18.5%	16.3%	15.8%	13.2%	12.9%	14.1%	15.3%	18.2%	20.4%	23.9%	28.4%
40-44	14.6%	14.3%	12.9%	12.0%	14.3%	20.2%	24.5%	22.2%	21.3%	20.2%	25.0%
45-49	12.6%	11.7%	13.0%	15.0%	17.9%	21.4%	22.5%	19.6%	20.3%	21.1%	30.0%
50+	12.2%	11.4%	13.6%	17.5%	22.0%	25.6%	25.9%	20.4%	19.9%	20.3%	31.3%

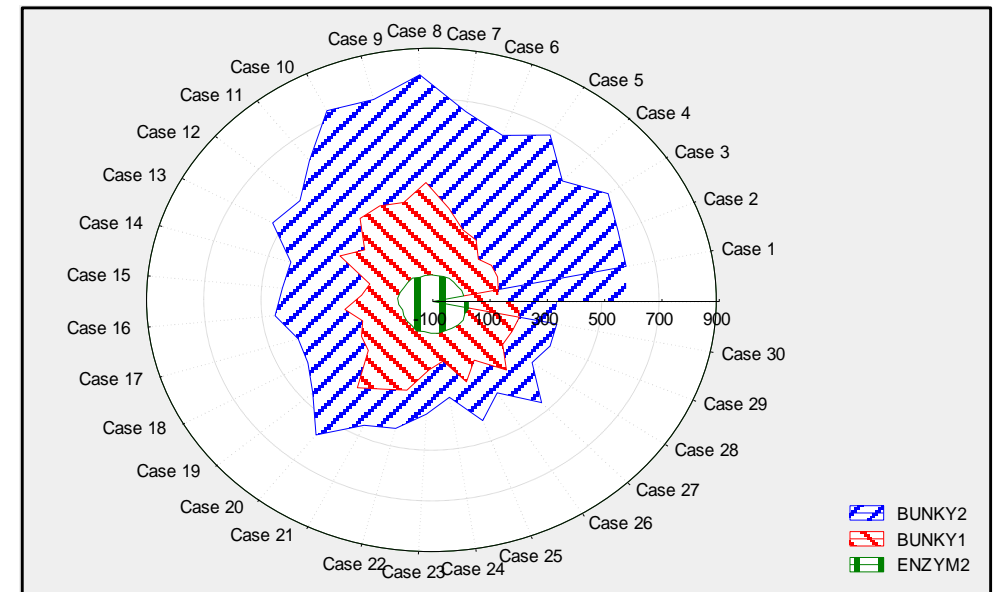
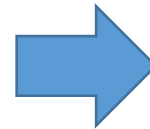
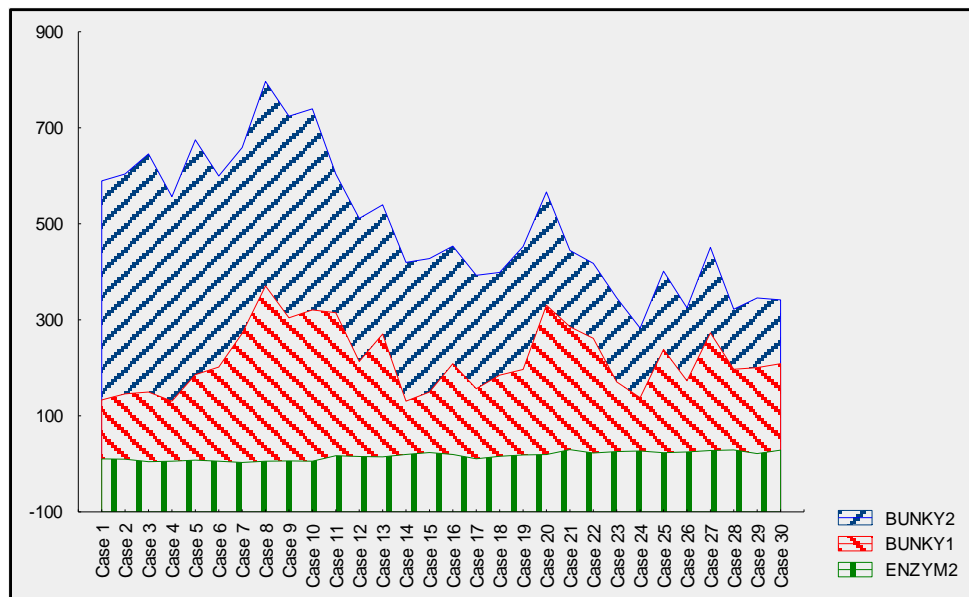
Pavoučí / paprskové grafy

- Vhodné pro srovnání profilů objektů nebo skupin objektů pomocí více proměnných
- Různá grafická forma



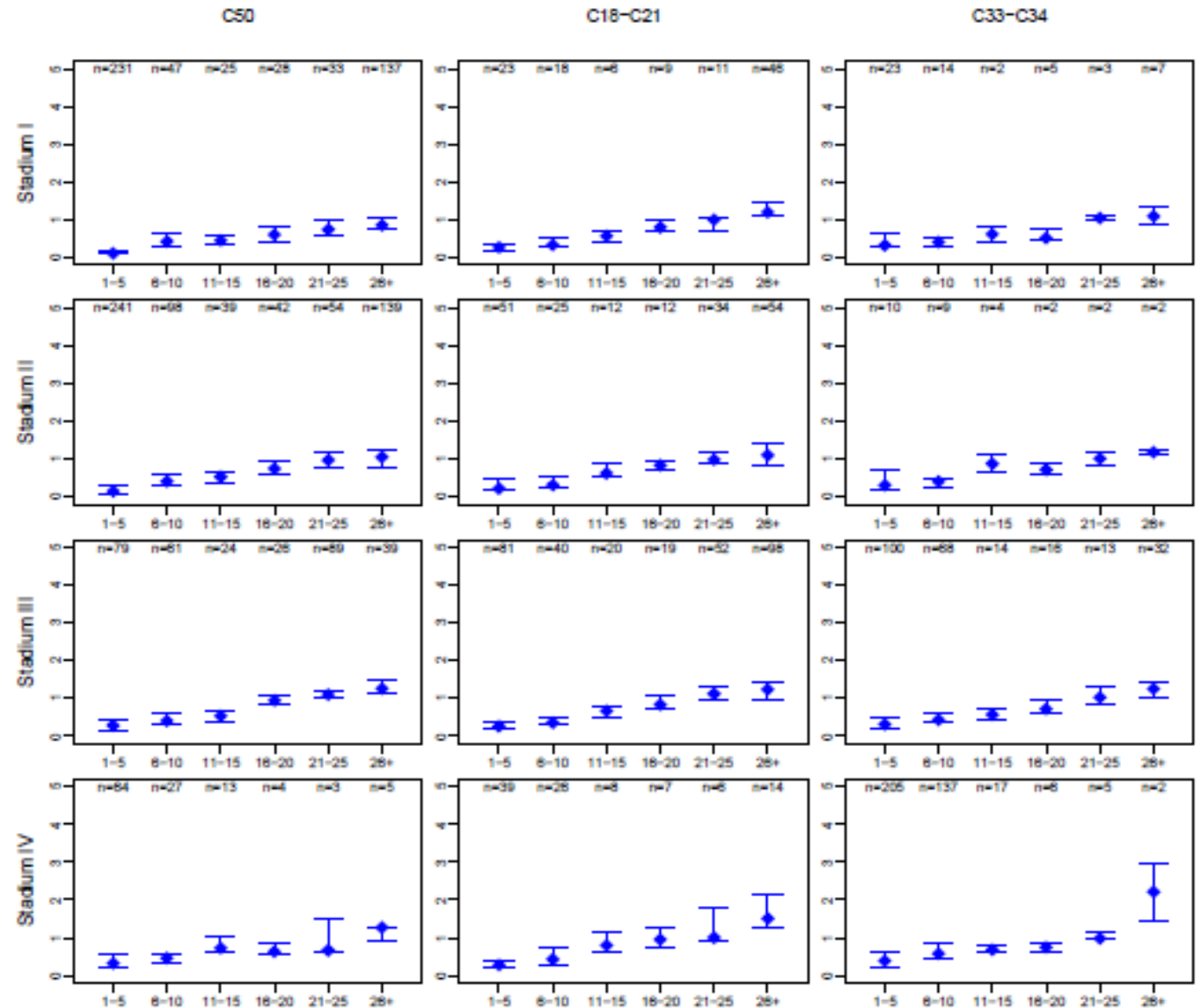
Polární graf

- Obdoba čárového, sloupcového nebo plošného grafu s osou X vynesenu na kružnici
- Vhodný pro cyklická data (cirkadiánní rytmy, sezonalita, směrová statistika pohybu živočichů)



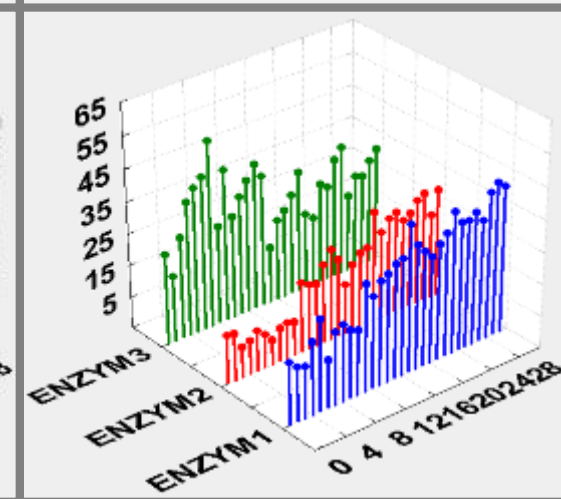
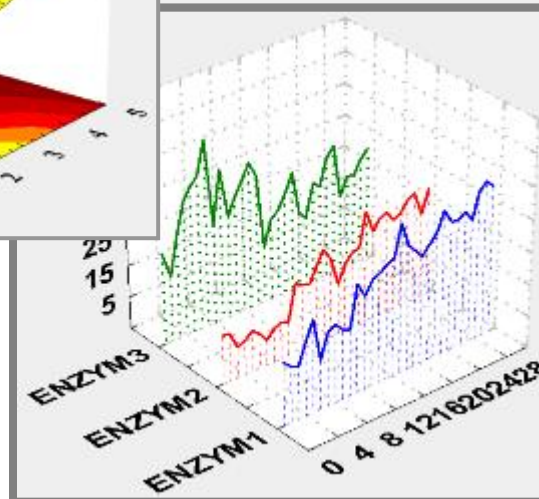
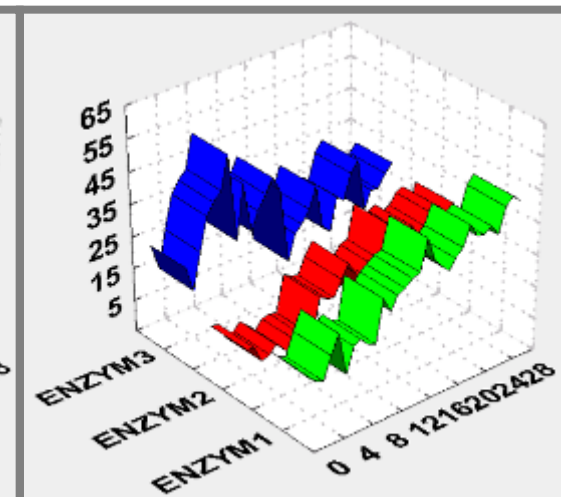
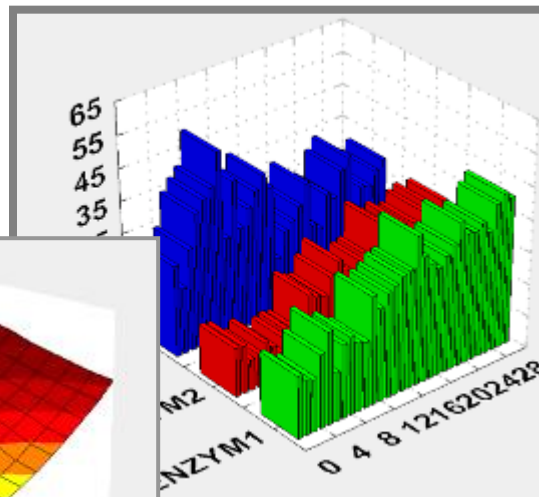
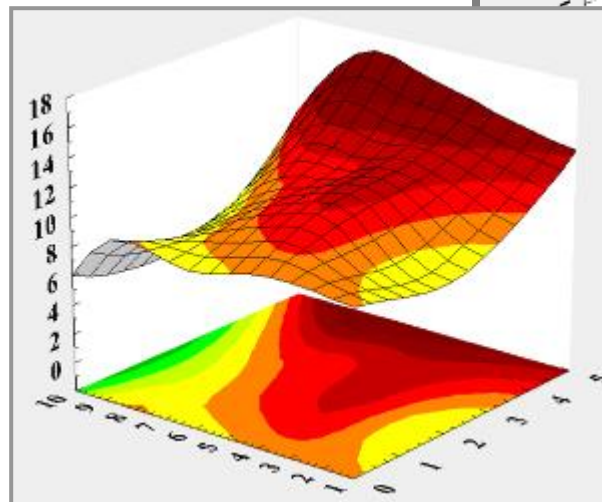
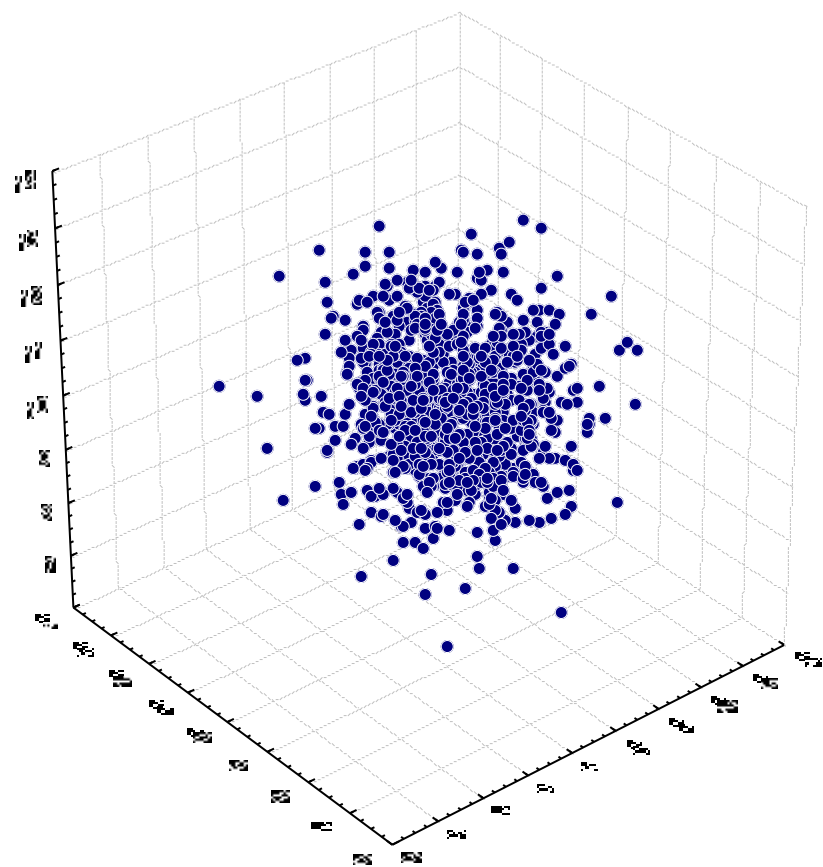
Grafické tabule

- Více grafů tvořících grafickou tabuli
- Možné skládat z různých grafů jednoho nebo více typů
- Prezence velkého množství dat na malém prostoru



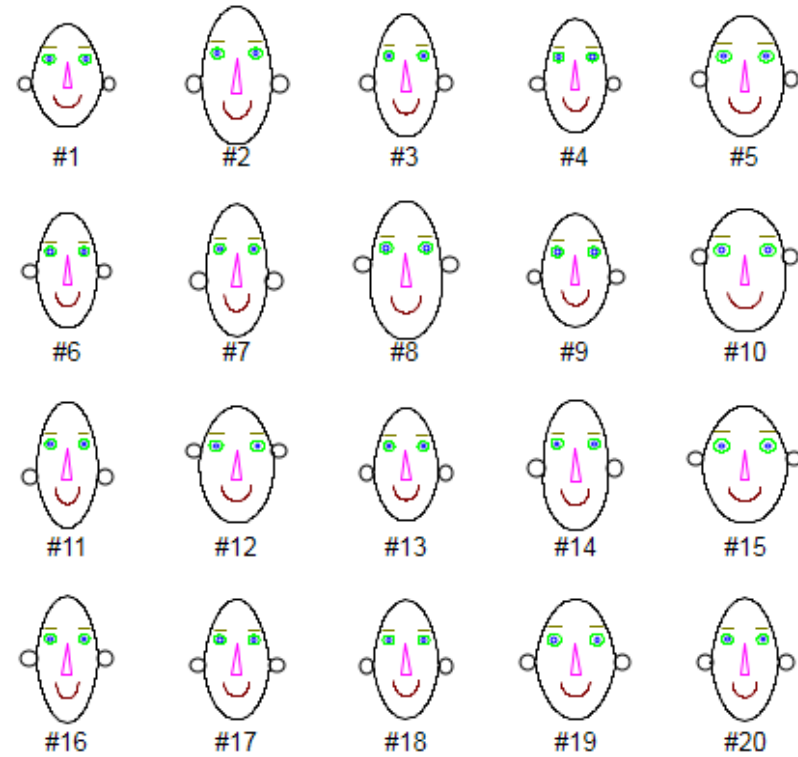
3D grafy

- Mnoho typů
- Velký důraz je třeba klást na interpretovatelnost a smysluplnost

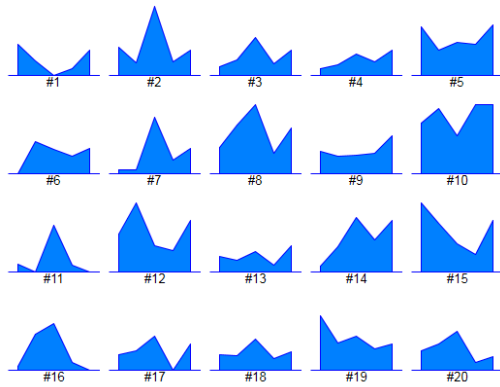


Chernoffovy tváře (ikonové grafy)

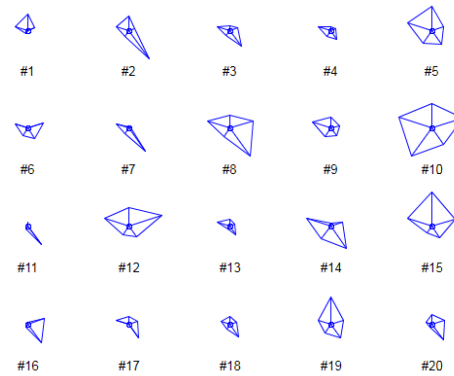
- Jednotlivé proměnné jsou zobrazeny jako rysy tváře
- Patří mezi tzv. ikonové grafy
 - hodnoty znaků znázorněny jako geometrické útvary či symboly
 - každému objektu (subjektu) odpovídá jeden obrazec složený z těchto geometrických útvarů či symbolů
 - umožní vizuálně porovnat, které objekty (subjekty) jsou si podobné



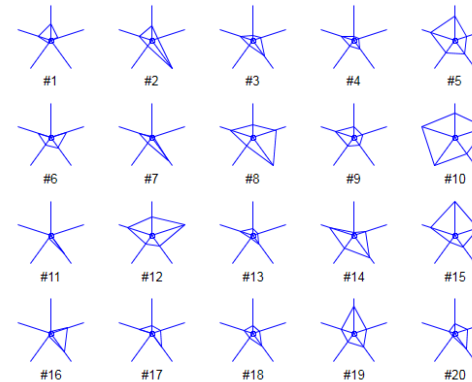
— face/w = vek
 — ear/lev = cel_cholesterol
 — halfface/h = vaha
 — upface/ecc = sys_tlak
 — loface/ecc = dia_tlak



Left to right:
 vek
 cel_cholesterol
 vaha
 sys_tlak
 dia_tlak



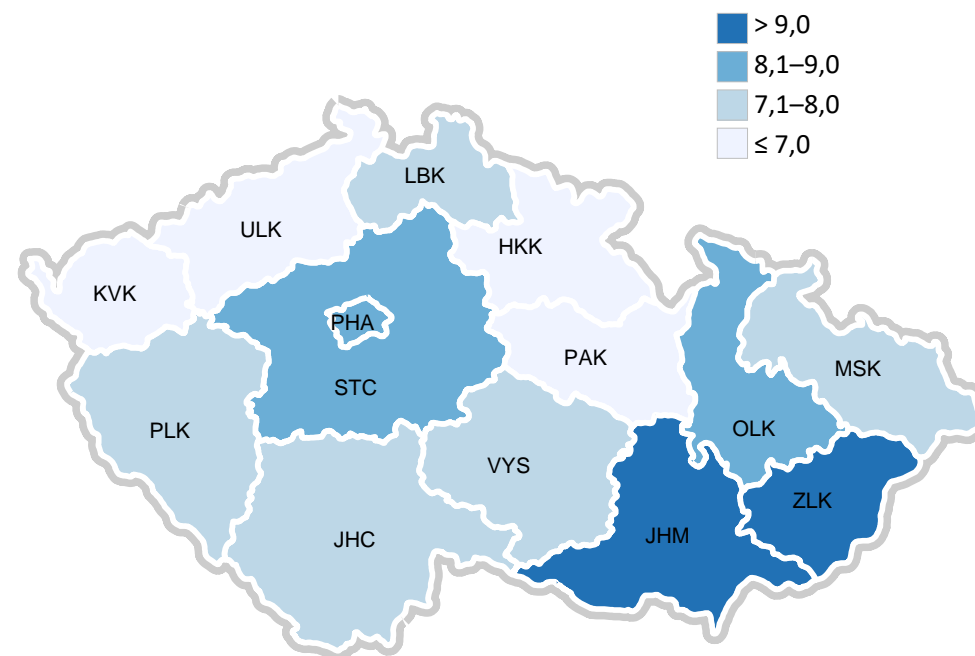
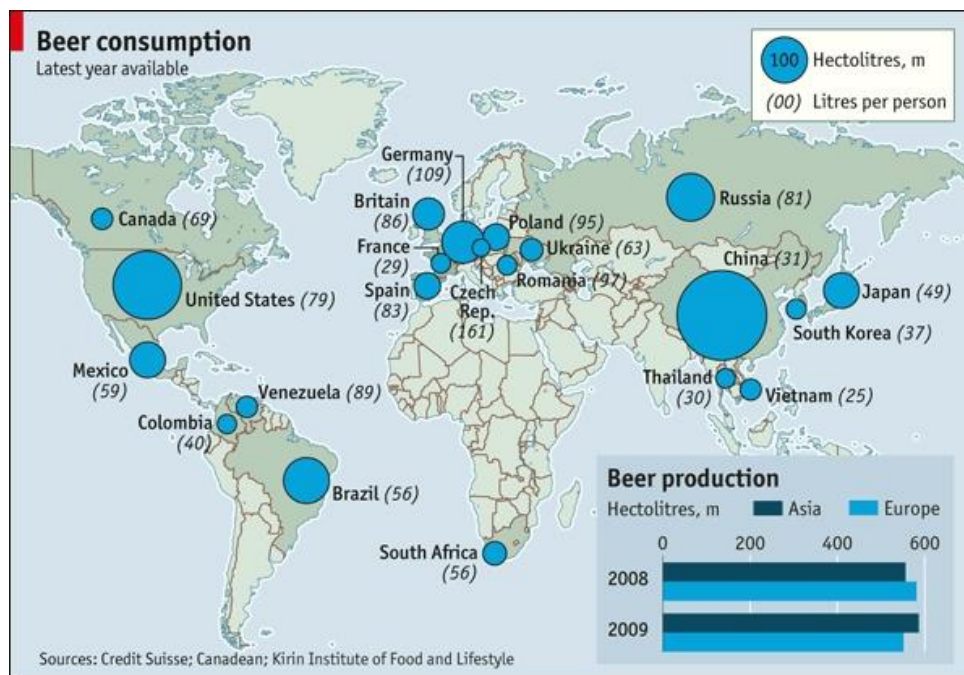
Clockwise:
 vek
 cel_cholesterol
 vaha
 sys_tlak
 dia_tlak



Clockwise:
 vek
 cel_cholesterol
 vaha
 sys_tlak
 dia_tlak

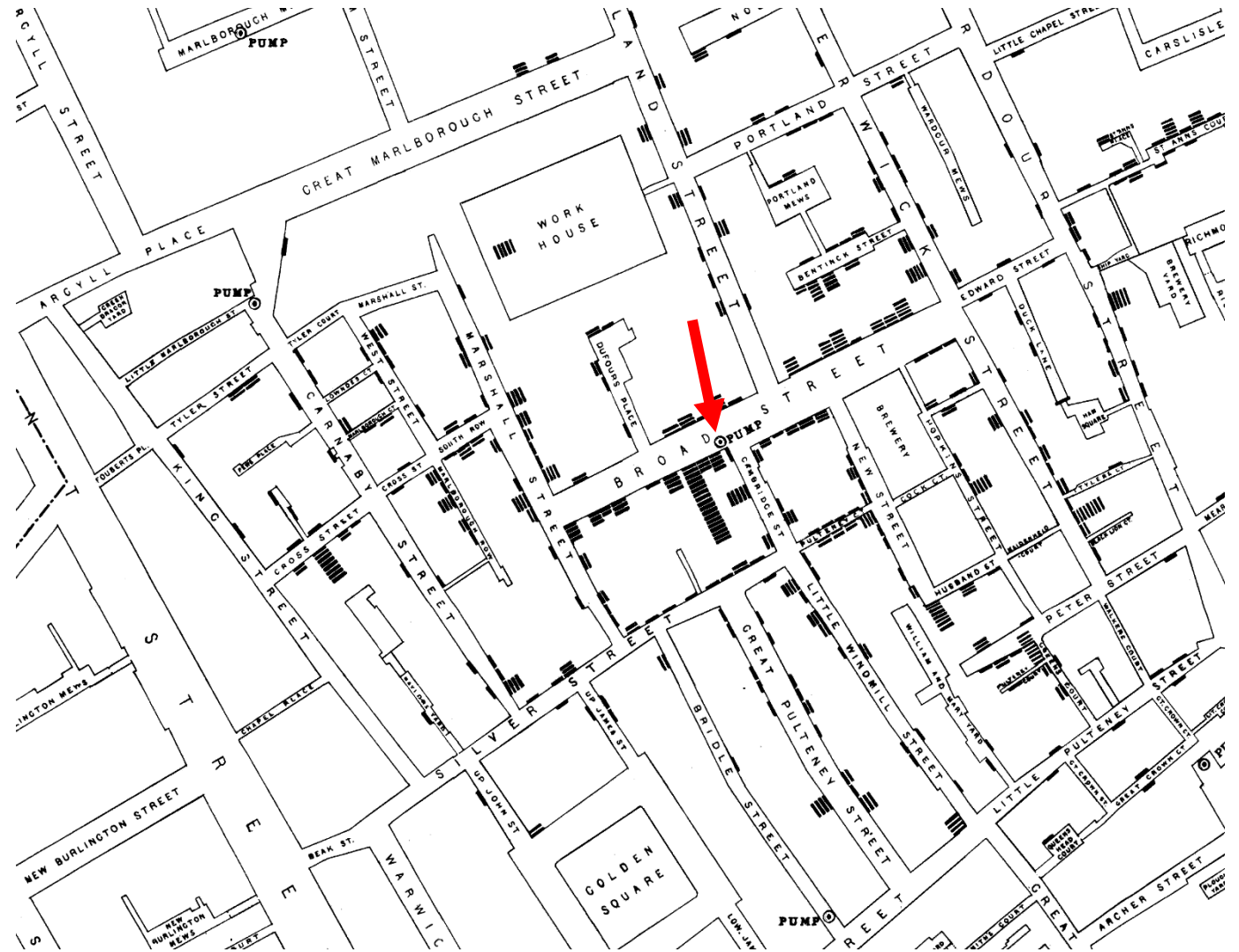
Mapy jsou také grafy

- Samostatná kapitola vizualizace dat
- Obarvení regionů v mapě dle výsledků analýzy nebo přímo vkládání grafů do map
- (sloupcové, koláčové atd.)
- ArcGIS – další z SW dostupných na inet.muni.cz

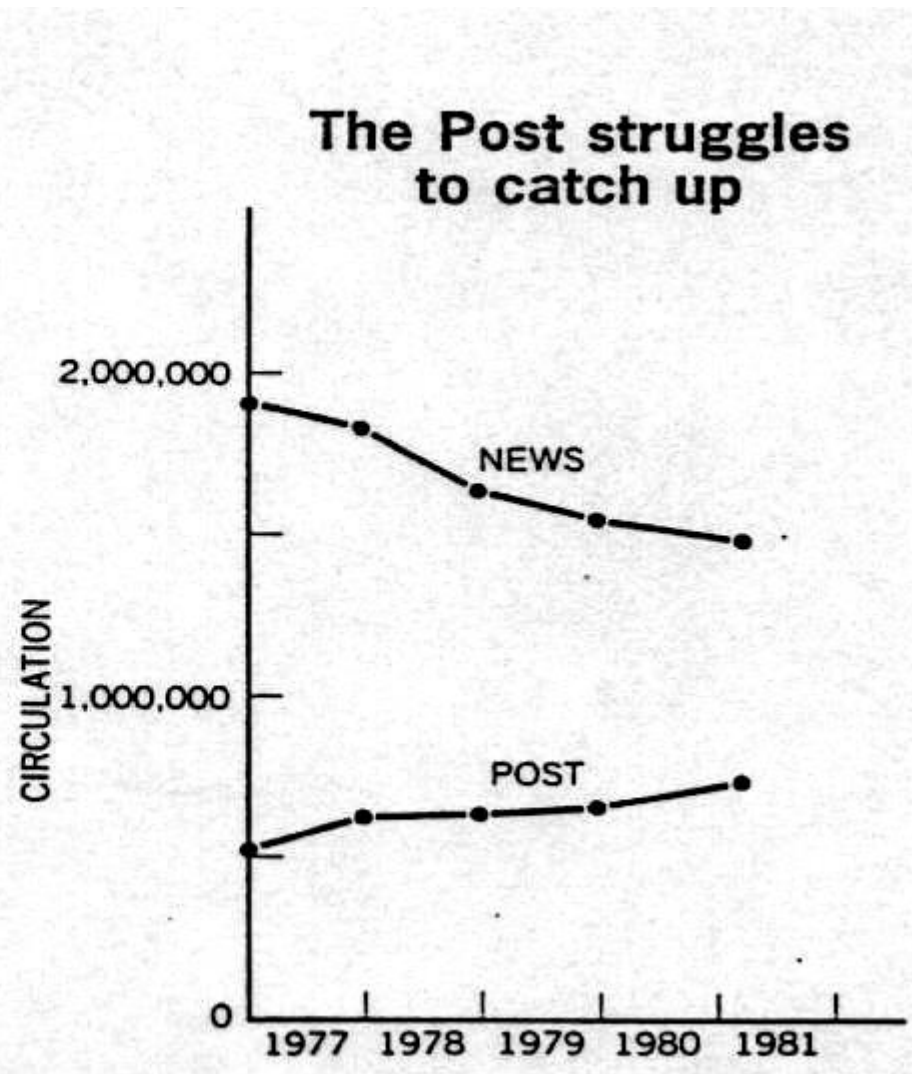
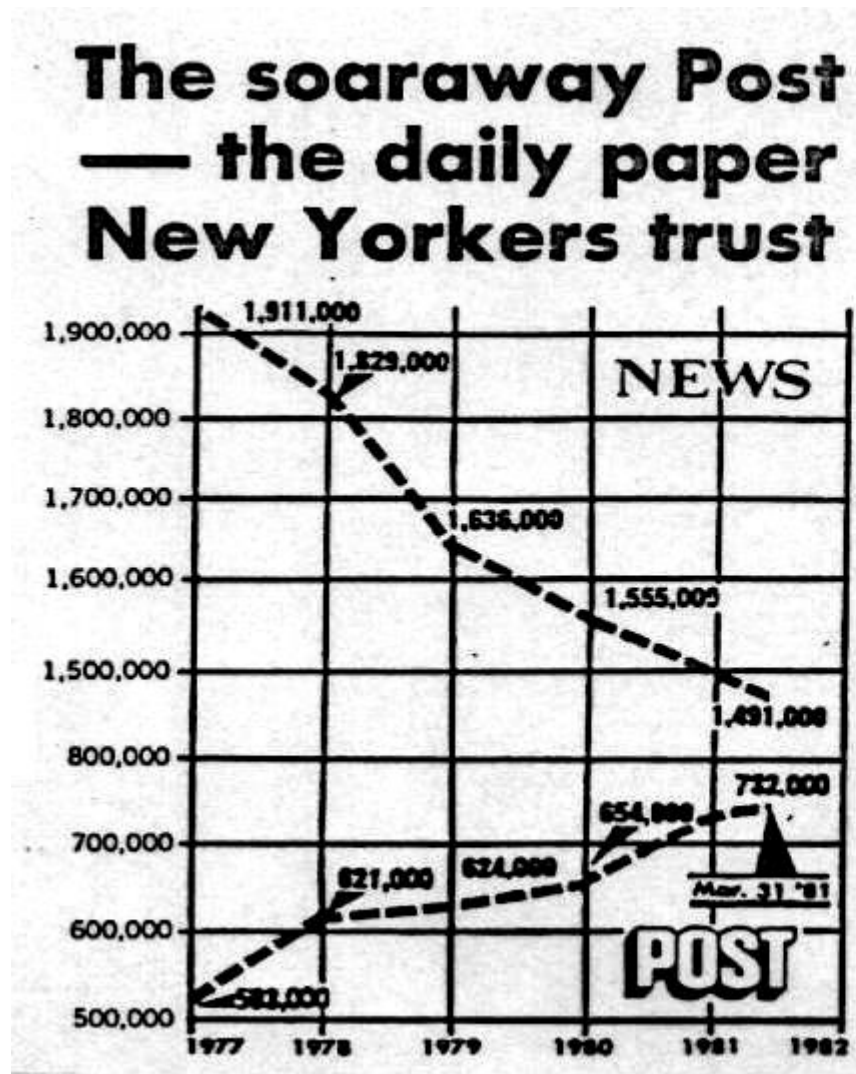


Slavné mapy: John Snow – cholera v Londýně

- 1854 Broad Street cholera outbreak
- Počty případů vyneseny jako černé sloupce dle bydliště obětí
- Identifikace zdroje nákazy – kontaminovaná studně
- Jeden z prvních příkladů prostorové analýzy dat a epidemiologického mapování



Nesprávné použití grafů: rozsah os („nevíme jak nakreslit“)



Nesprávné použití grafů: standardizace os („nevíme co kreslíme“)

