

OKNA STATISTIKY DOKOŘÁN ANEB POSTŘEHY JEDNÉ DATOVÉ ANALYTIČKY

Martina Litschmannová

Vysoká škola báňská – Technická univerzita Ostrava

Fakulta elektrotechniky a informatiky, Katedra aplikované matematiky



**Katedra
aplikované matematiky**

50. konferencia slovenských matematikov
Jasná pod Chopkom
24. 11. 2018

Obsah

Část 1

- Analýza dat – Základní pojmy
- Popisná statistika kvalitativního znaku – Tabulky četnosti, vizualizace
- Jak to vypadá v praxi

Část 2

- Popisná statistika kvantitativního znaku – Míry polohy, míry variability, vizualizace, zaokrouhlování

Část 3

- Analýza závislosti dvou kvantitativních znaků – Korelační koeficient, vizualizace

Co je to statistika?

Google – $83 \cdot 10^6$ odkazů (čeština), $1,3 \cdot 10^9$ odkazů (angličtina)

- Teoretická disciplína, která se zabývá metodami sběru a analýzy dat (matematická statistika vs. aplikovaná statistika)
- Číselný údaj „syntetizující“ vlastnosti datových souborů (četnost, průměr, rozptyl, ...)
- Uspořádaný datový soubor (statistika přístupů na web. stránky, statistika střel na branku, statistika nehodovosti, ekonomické statistiky, ...)

[Real Time Statistics Project](#) – pěkná ukázka toho, k čemu lze použít statistické modely

Co vypovídá statistika o jednotlivci?



Donald Trump



podnikatel



politik (prezident)

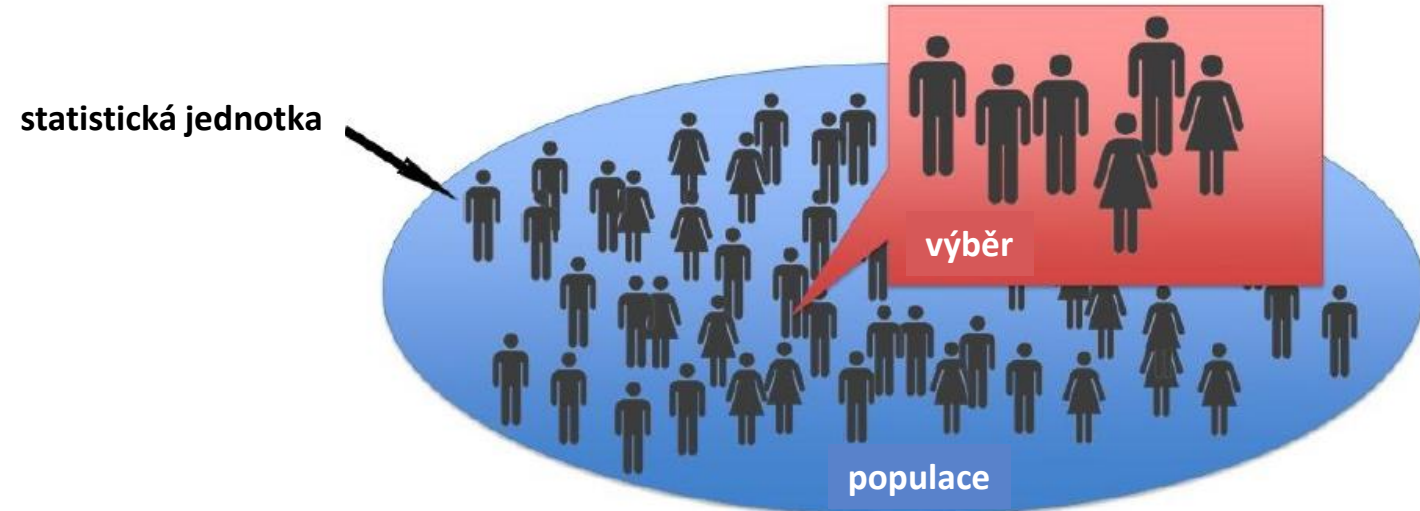


američan

- Statistika nezkoumá jednotlivce jako individualitu, ale jako anonymního nositele některého znaku (činnosti, vlastnosti).

Základní pojmy

- **Populace** (základní soubor) je soubor nějakých prvků, o kterém chceme statistickými metodami něco vypovídat. Definuje se výčtem nebo pomocí zvolené vlastnosti. O každém prvku umíme rozhodnout, zda do populace patří či nikoliv.
- **Výběr** je část dané populace, která má sloužit k odvození závěrů platných pro celou populaci. (Pozor na reprezentativnost výběru!)
- **Statistická jednotka** je prvek populace.
- **Statistický znak (proměnná)** je nějaká měřitelná (zjistitelná) charakteristika statistické jednotky (hmotnost, pohlaví, ...).



Vlastní pokus o výběrové šetření

- Dotazník pro studenty (např. pomocí Google Apps) - <http://goo.gl/forms/Z289s0ALPY>

Dotazník pro výukové účely - Brigáda při studiu

Vážené studentky, vážení studenti,

prosím Vás o vyplnění krátkého anonymního dotazníku na téma brigáda při studiu. Výsledky dotazníkového šetření budou použity pro prezentaci analýzy dat v rámci výuky.

Předem moc děkuji za váš čas a ochotu,
Martina Litschmannová

*Povinné pole

Identifikace respondenta

1. Pohlaví *

Označte jen jednu elipsu.

- ☐ žena
☐ muž

2. Výška (cm) *

.....

3. Váha (kg) *

.....

4. Přivyděláváte si v rámci prezenčního studia na brigádách? *

Označte jen jednu elipsu.

- ☐ ano *Přeskočte na otázku 5.*
☐ ne *Přeskočte na otázku 8.*

Brigáda při studiu

5. Jak často brigádu máte? *

Označte jen jednu elipsu.

- ☐ každý pracovní den
☐ pouze o víkendu
☐ nepravidelně
☐ Jiné:

6. Jak byste svou brigádu charakterizoval(a)? *

Označte jen jednu elipsu.

- ☐ praxe v oboru během studia
☐ "kancelářská" práce mimo obor mého studia
☐ manuální práce
☐ Jiné:

7. Kolik času týdně obvykle věnujete brigádě? *

Uvedte čas v hodinách.

.....

Studium

8. Kolik času týdně obvykle věnujete studiu? *

Uvedte čas v hodinách.

.....

Základní pojmy

- **Populace** (základní soubor) je soubor nějakých prvků, o kterém chceme statistickými metodami něco vypovídat.
- **Výběr** je část dané populace, která má sloužit k odvození závěrů platných pro celou populaci.
- **Statistická jednotka** je prvek populace.
- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika statistické jednotky.

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídaní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.
 - **Kvantitativní znak** – znak, jehož varianty mají číselné hodnoty (má smysl posuzovat rozdíly a poměry)

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.
 - **Kvantitativní znak** – znak, jehož varianty mají číselné hodnoty (má smysl posuzovat rozdíly a poměry)

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.
 - **Kvantitativní znak** – znak, jehož varianty mají číselné hodnoty (má smysl posuzovat rozdíly a poměry)
 - **Kvalitativní znak** – znak, jehož varianty se liší kvalitou (může jít i o číselné hodnoty – např. známka z matematiky)

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.
 - **Kvantitativní znak** – znak, jehož varianty mají číselné hodnoty (má smysl posuzovat rozdíly a poměry)
 - **Kvalitativní znak** – znak, jehož varianty se liší kvalitou (může jít i o číselné hodnoty – např. známka z matematiky)

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídnání dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.
 - **Kvantitativní znak** – znak, jehož varianty mají číselné hodnoty (má smysl posuzovat rozdíly a poměry)
 - **Kvalitativní znak** – znak, jehož varianty se liší kvalitou (může jít i o číselné hodnoty – např. známka z matematiky)
 - **Alternativní (dichotomický) znak** – znak, který nabývá pouze dvou variant

Základní pojmy

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídnání dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- **Statistický znak** je nějaká měřitelná (zjistitelná) charakteristika prvků základního souboru.
 - **Kvantitativní znak** – znak, jehož varianty mají číselné hodnoty (má smysl posuzovat rozdíly a poměry)
 - **Kvalitativní znak** – znak, jehož varianty se liší kvalitou (může jít i o číselné hodnoty – např. známka z matematiky)
 - **Alternativní (dichotomický) znak** – znak, který nabývá pouze dvou variant

**Popisná statistika
aneb
Jak efektivně popsat a vizualizovat data**

Část 1
Kvalitativní znak

Popisná statistika – Kvalitativní znak

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Respondent (proband) – označení statistické jednotky v dotazníkovém šetření

Popište strukturu datového souboru v závislosti na pohlaví respondentů.

Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Varianta znaku	Absolutní četnost	Relativní četnost
x_i	n_i	p_i
x_1	n_1	$p_1 = n_1/n$
x_2	n_2	$p_2 = n_2/n$
x_k	n_k	$p_k = n_k/n$
Celkem:	$n_1 + n_2 + \dots + n_k = n$	1

+ **Modus** (varianta, které znak nabývá s nejvyšší četností)

Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost
muž	66	0,776470588
žena	19	0,223529412
Celkem:	85	1,000000000

Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	66	77,6470588
žena	19	22,3529412
Celkem:	85	100,0000000

Jak zaokrouhlovat relativní četnosti?

1,0 % ... 0,85 osob
0,1 % ... 0,085 osob



Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	66	77,6470588
žena	19	22,3529412
Celkem:	85	100,0000000

} Pozor na zaokrouhlovací chybu!

Jak zaokrouhlovat relativní četnosti?

1,0 % ... 0,85 osob
0,1 % ... 0,085 osob



Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	66	78
žena	19	22
Celkem:	85	100

Pozor na zaokrouhlovací chybu!

Součet musí být 100 %!

Jak zaokrouhlovat relativní četnosti?

1,0 % ... 0,85 osob
0,1 % ... 0,085 osob



Příklad demonstrující problém zaokrouhlovací chyby

TABULKA ROZDĚLENÍ ČETNOSTI		
Typ pasažéra	Absolutní četnosti	Relativní četnosti (%)
Muž	77	37,37864
Žena	85	41,26214
Dítě	44	21,35922
Celkem:	206	100,00000

1 % ... 2,06 osob

0,1 % ... 0,206 osob

Jak zaokrouhlit relativní četnost?



Příklad demonstrující problém zaokrouhlovací chyby

TABULKA ROZDĚLENÍ ČETNOSTI		
Typ pasažéra	Absolutní četnosti	Relativní četnosti (%)
Muž	77	37,4
Žena	85	41,3
Dítě	44	21,4
Celkem:	206	100,1



POZOR na
zaokrouhlovací
chybu!



Příklad demonstrující problém zaokrouhlovací chyby

TABULKA ROZDĚLENÍ ČETNOSTI		
Typ pasažéra	Absolutní četnosti	Relativní četnosti (%)
Muž	77	37,4
Žena	85	41,3
Dítě	44	21,3
Celkem:	206	100,0

Dopočet do 100 %!



Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	?	78
žena	?	22
Celkem:	85	100

Relativní četnosti používejte pouze jako doplněk absolutních četností!



Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	66	78
žena	19	22
Celkem:	85	100

Určete modus proměnné pohlaví.

Popisná statistika – Kvalitativní znak

- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	66	78
žena	19	22
Celkem:	85	100

Modus = muž

Mezi respondenty převažovali muži.

Popisná statistika – Kvalitativní znak

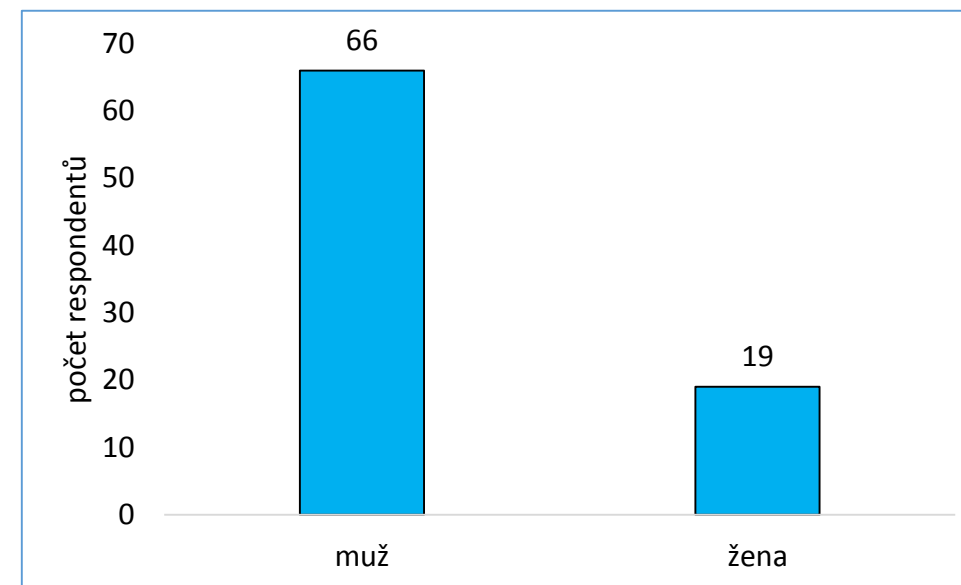
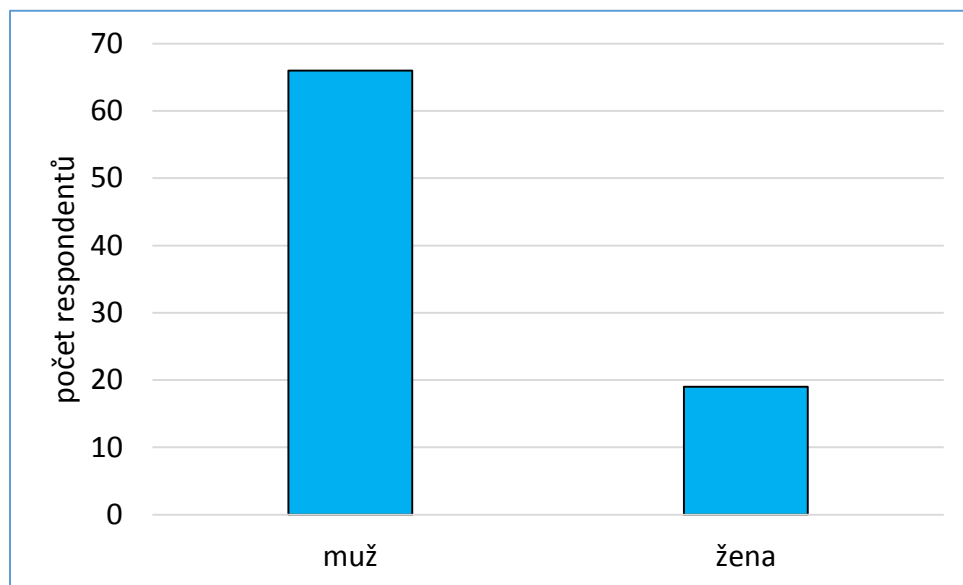
- Tabulka četností

Tabulka četností		
Pohlaví	Absolutní četnost	Relativní četnost (%)
muž	66	78
žena	19	22
Celkem:	85	100

Jak data vizualizovat?

Popisná statistika – kvalitativní znak

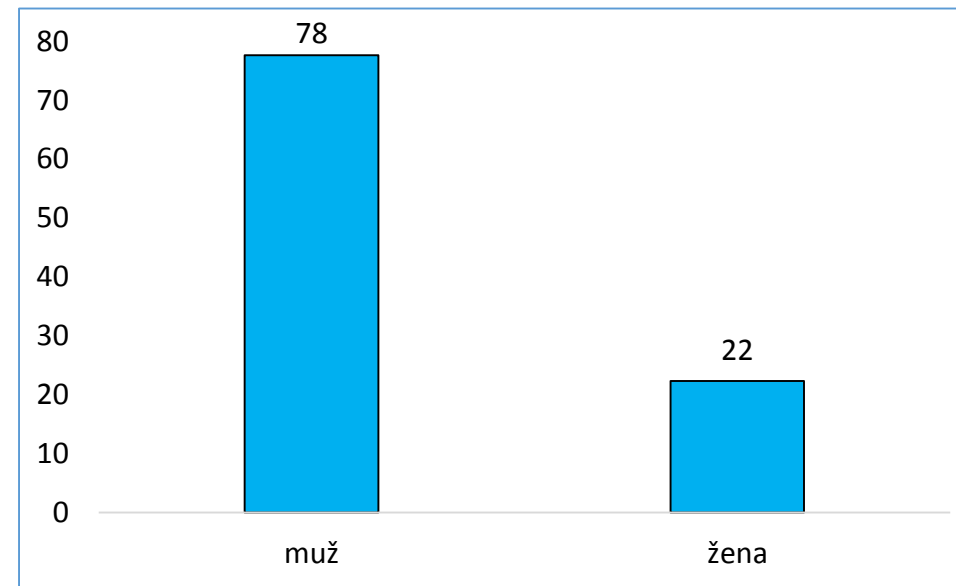
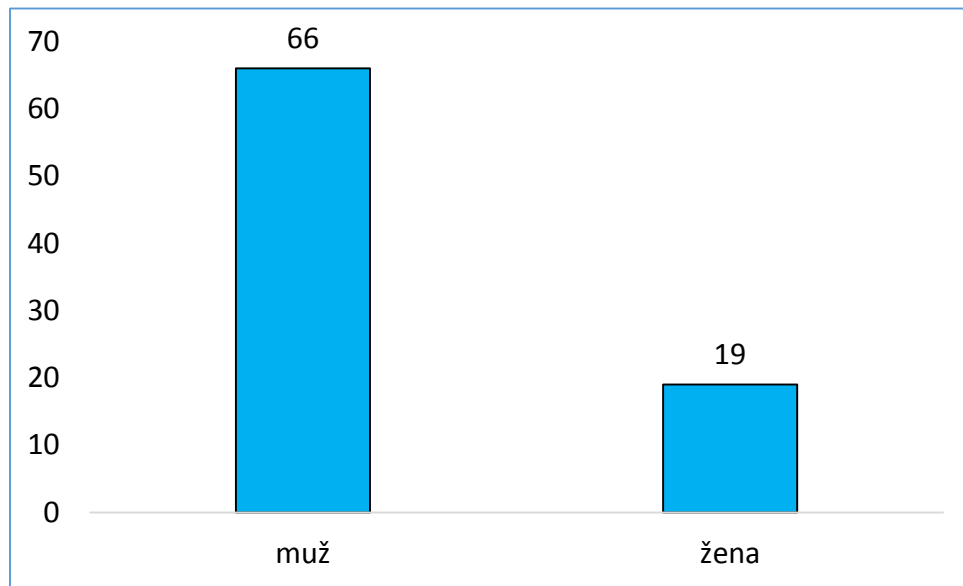
Sloupcový graf (Bar Chart)



- Nejsou-li v grafu uvedeny absolutní (relativní) četnosti, obvykle je nedokážeme „od oka“ přesně odečíst.

Popisná statistika – kvalitativní znak

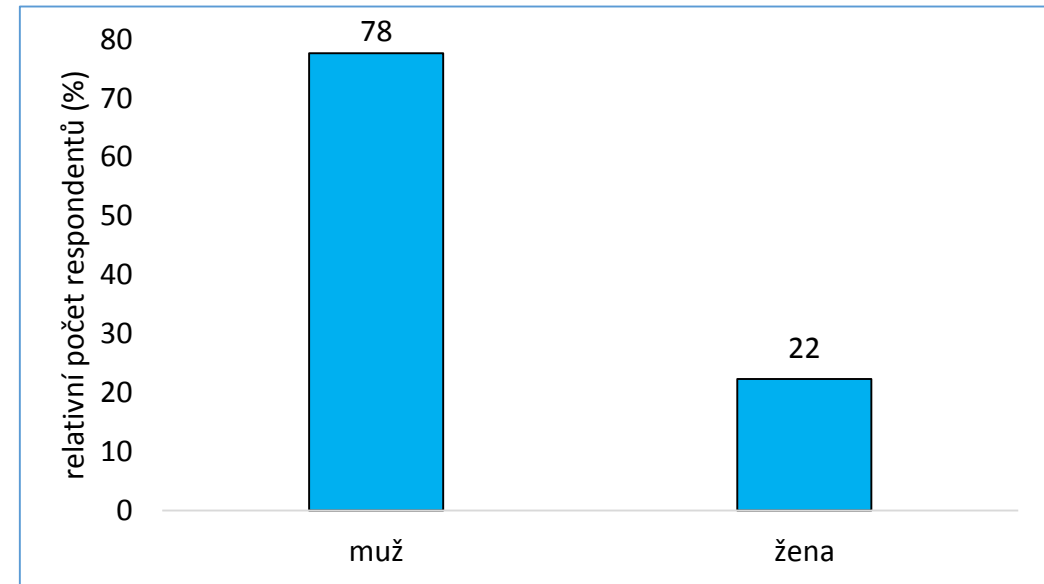
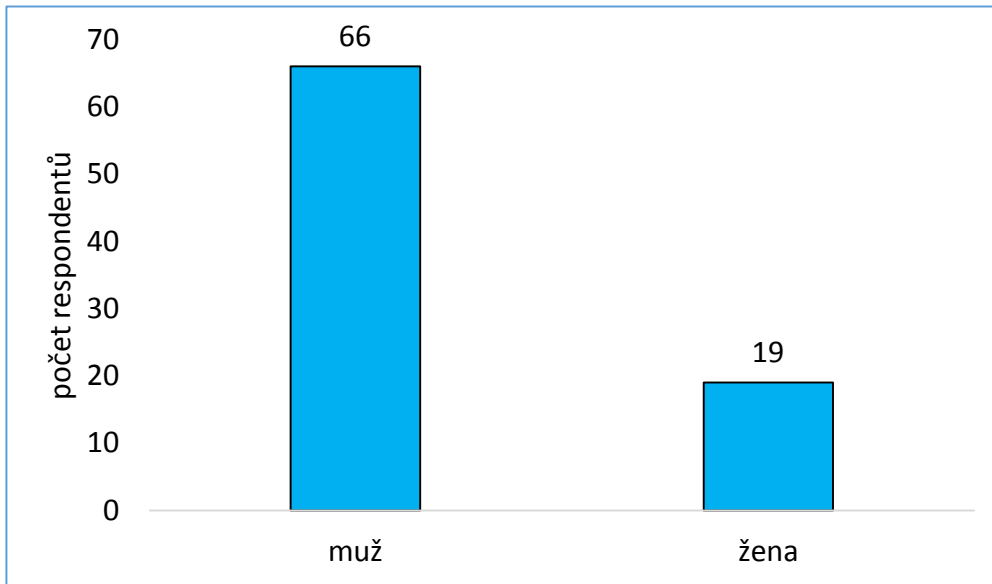
Sloupcový graf (Bar Chart)



- Nejsou-li v grafu uvedeny absolutní (relativní) četnosti, obvykle je nedokážeme „od oka“ přesně odečíst.

Popisná statistika – kvalitativní znak

Sloupcový graf (Bar Chart)



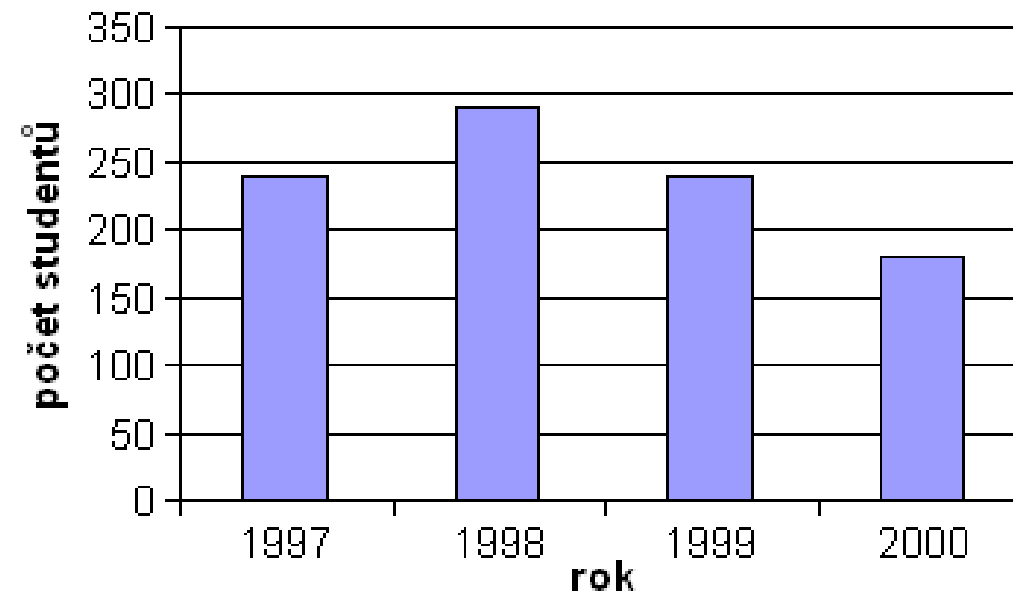
- Nejsou-li v grafu uvedeny absolutní (relativní) četnosti, obvykle je nedokážeme „od oka“ přesně odečíst.
- Pozor na uvádění popisu os!

Popisná statistika – Kvalitativní znak

Zdroj: Srovnávací testy pro žáky 9. tříd

Určete pravdivost tvrzení:

V žádných dvou letech nebyl počet studentů stejný.

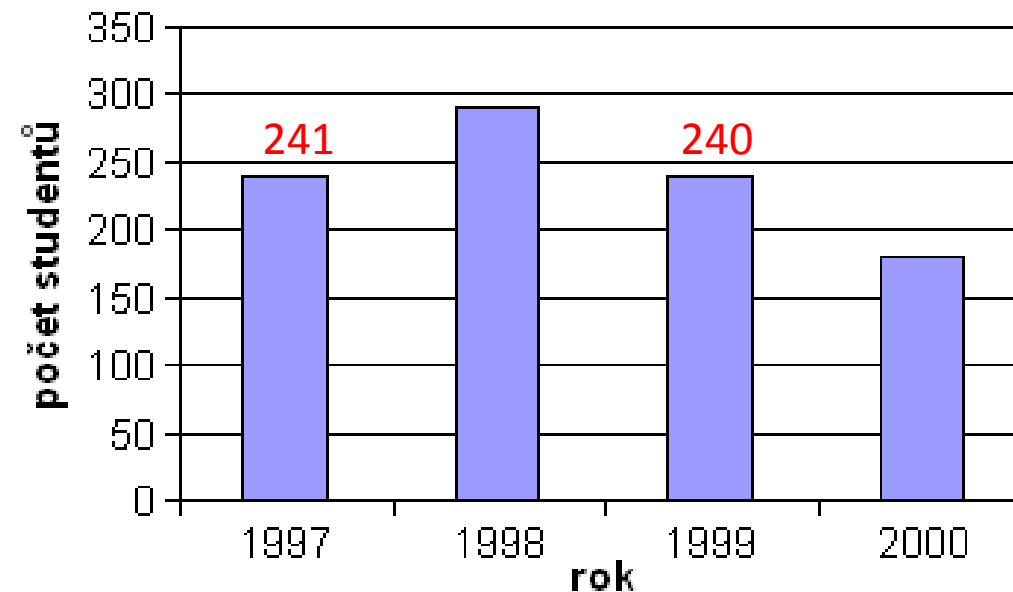


Popisná statistika – Kvalitativní znak

Zdroj: Srovnávací testy pro žáky 9. tříd

Určete pravdivost tvrzení:

V žádných dvou letech nebyl počet studentů stejný.

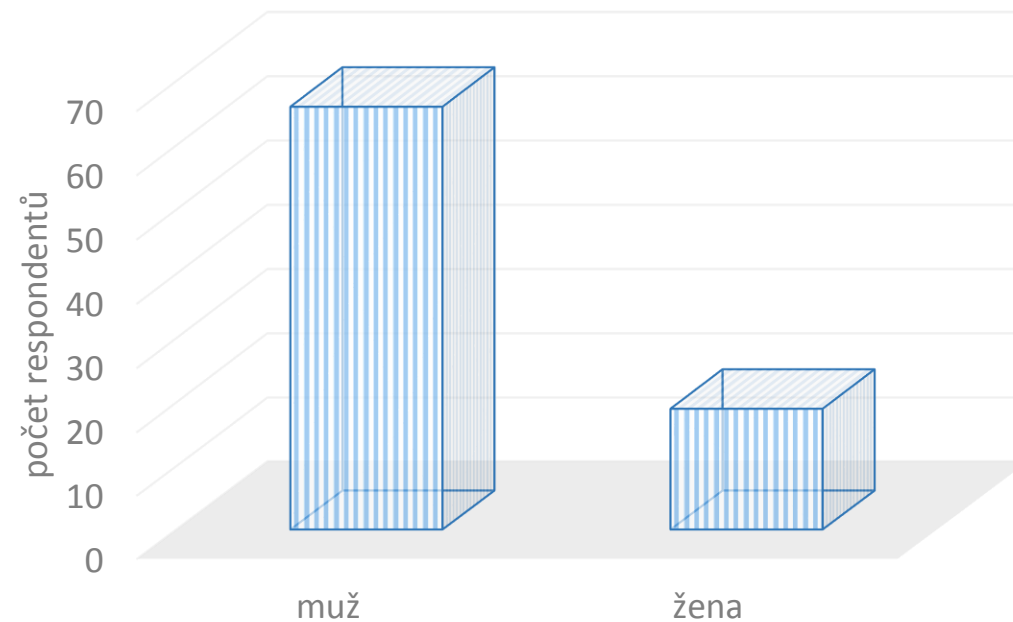


Pozor na omezenou
vypovídací schopnost grafů!



Popisná statistika – Kvalitativní znak

- Sloupcový graf

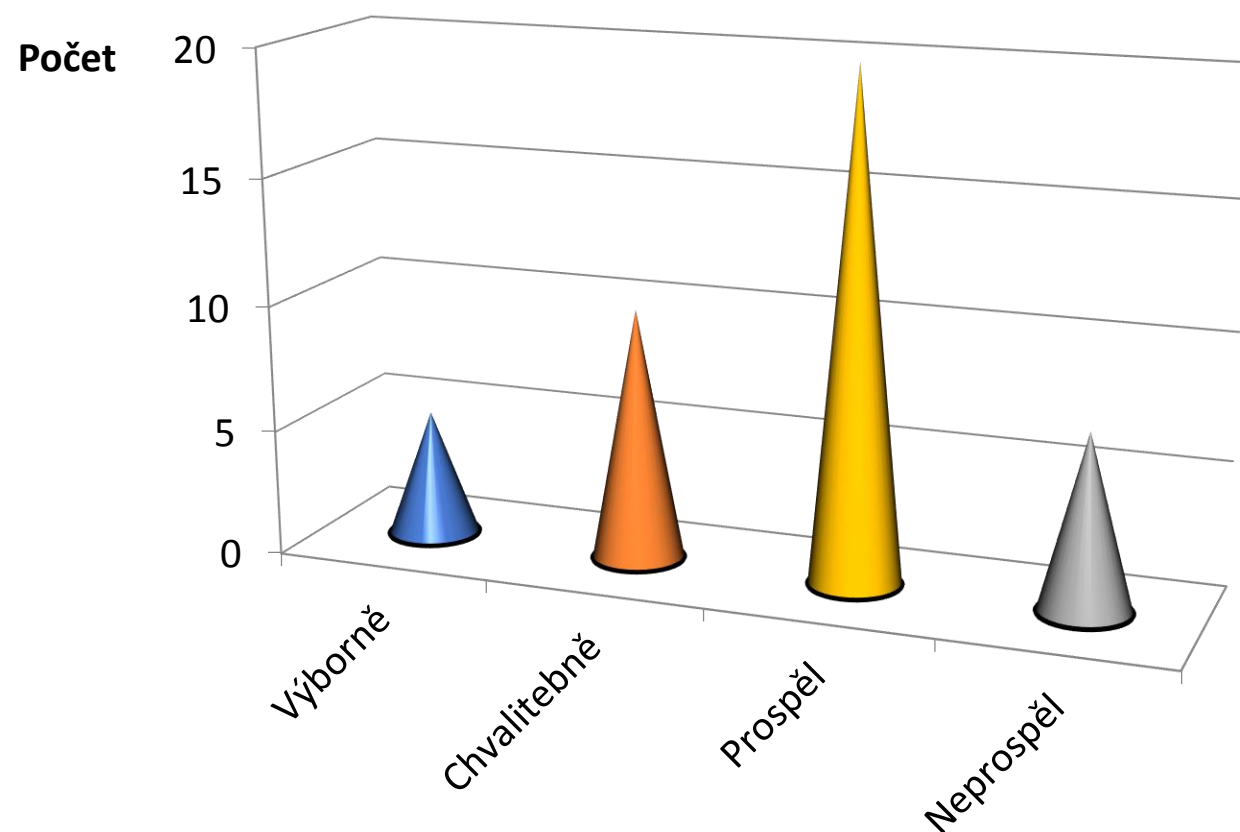


Nezapomínejte, že méně mnohdy znamená více ...



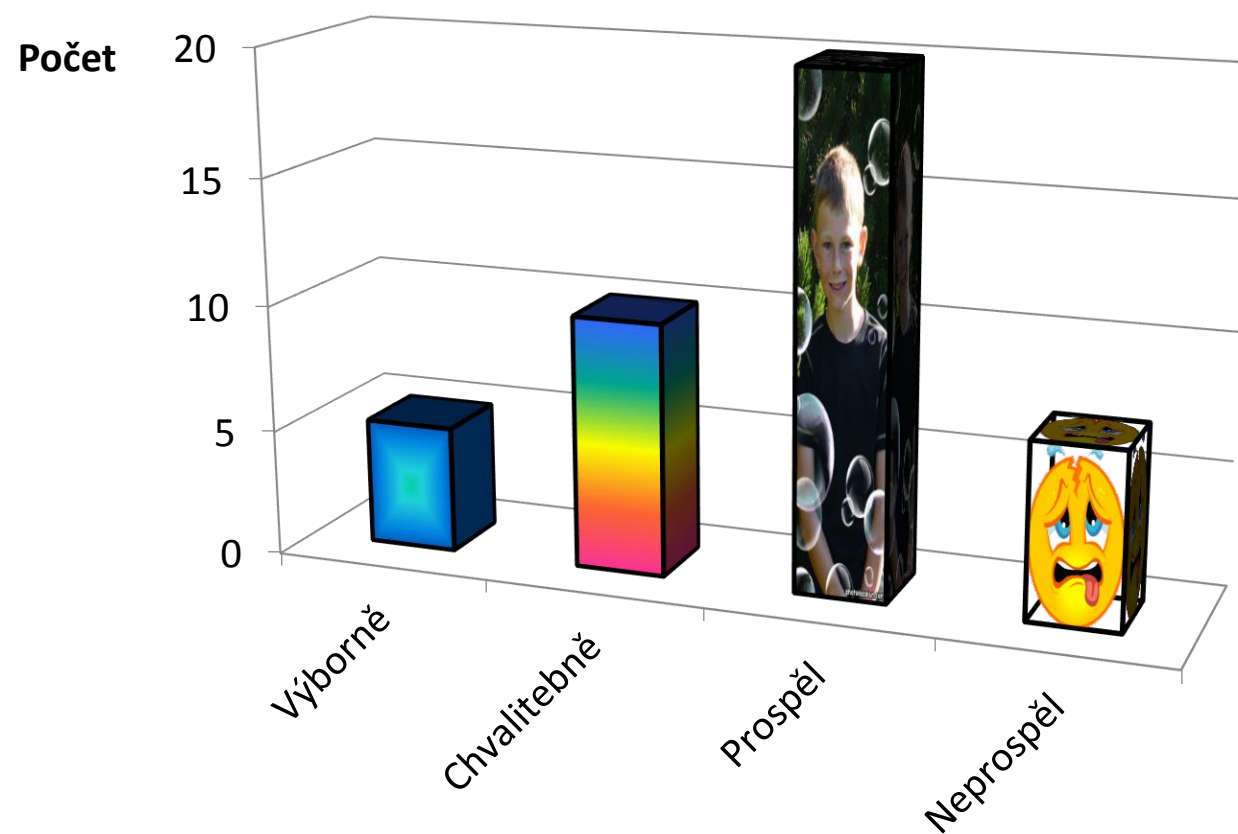
Popisná statistika – Kvalitativní znak

- I takto může vypadat sloupcový graf ...



Popisná statistika – Kvalitativní znak

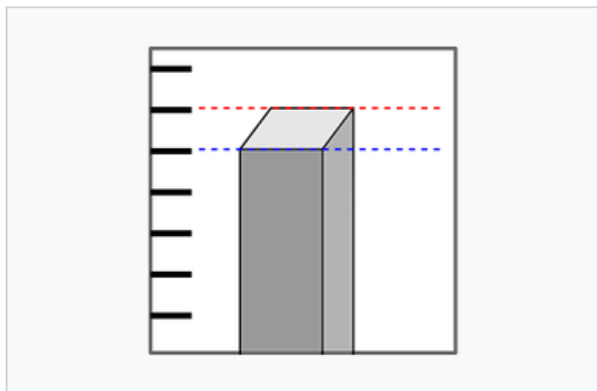
- I takto může vypadat sloupcový graf ...



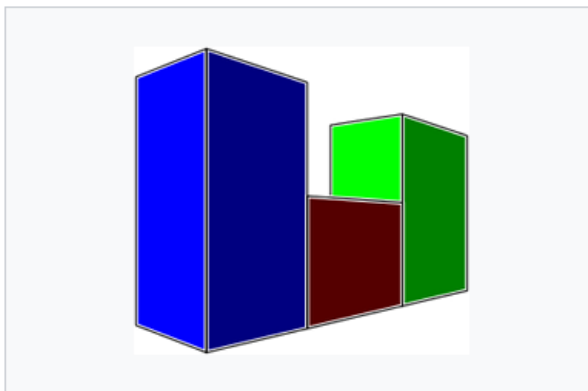
Popisná statistika – Kvalitativní znak

- 3D sloupcový graf

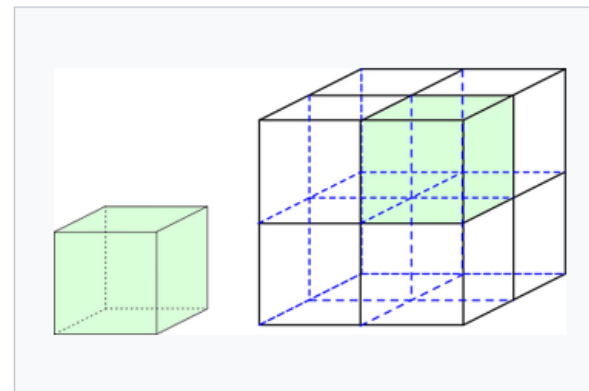
Opravdu musí být v každé efektní prezentaci / publikaci?



3D graf může být pro čtenáře matoucí...



Modrý a zelený sloupec mají stejnou výšku. Odpovídá to tomu, jak obrázek vnímáte?

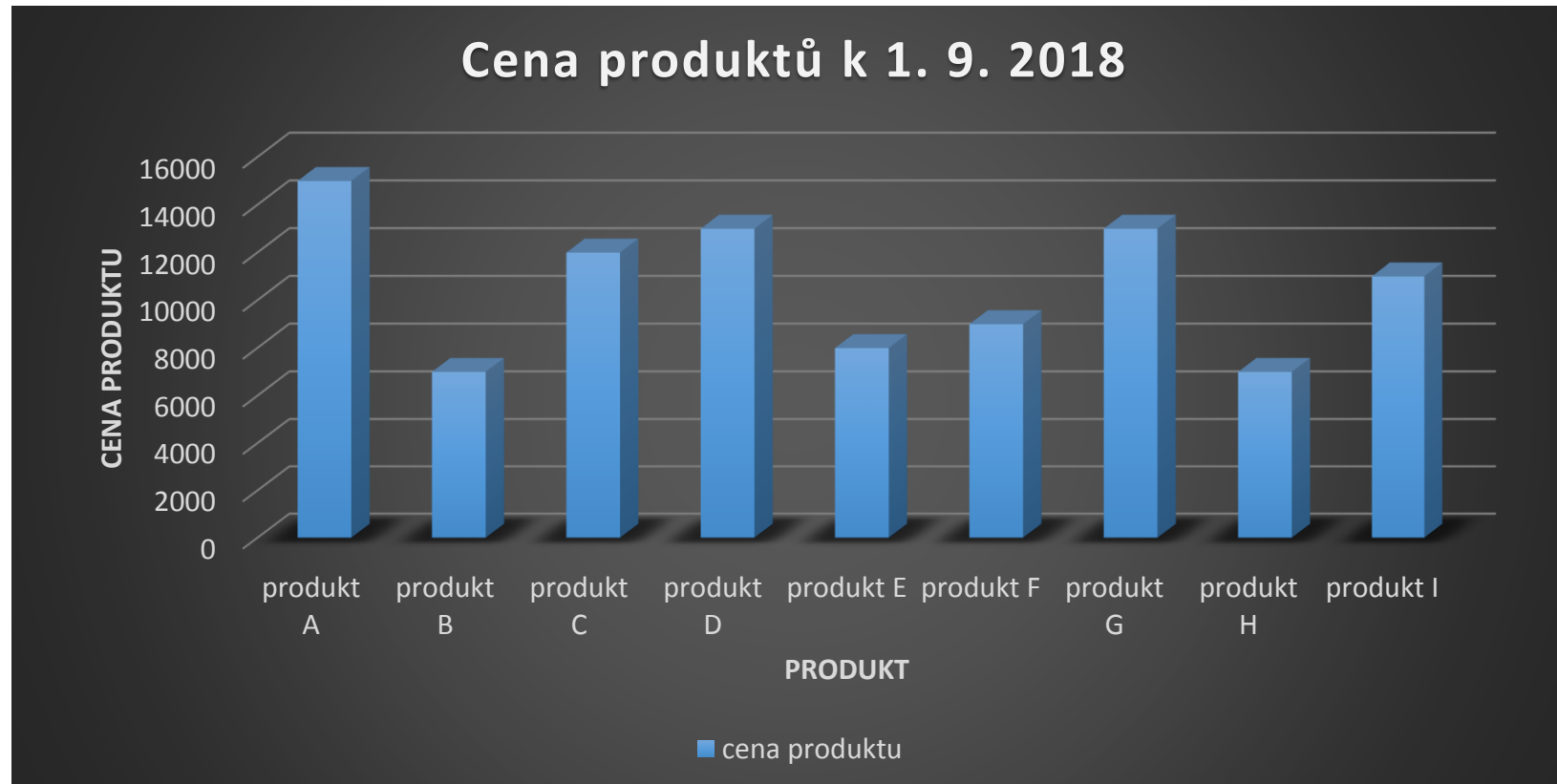


Zvětšíte-li k krát stranu krychle, její objem (to, co vnímáte) se zvětší k^3 krát.

Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

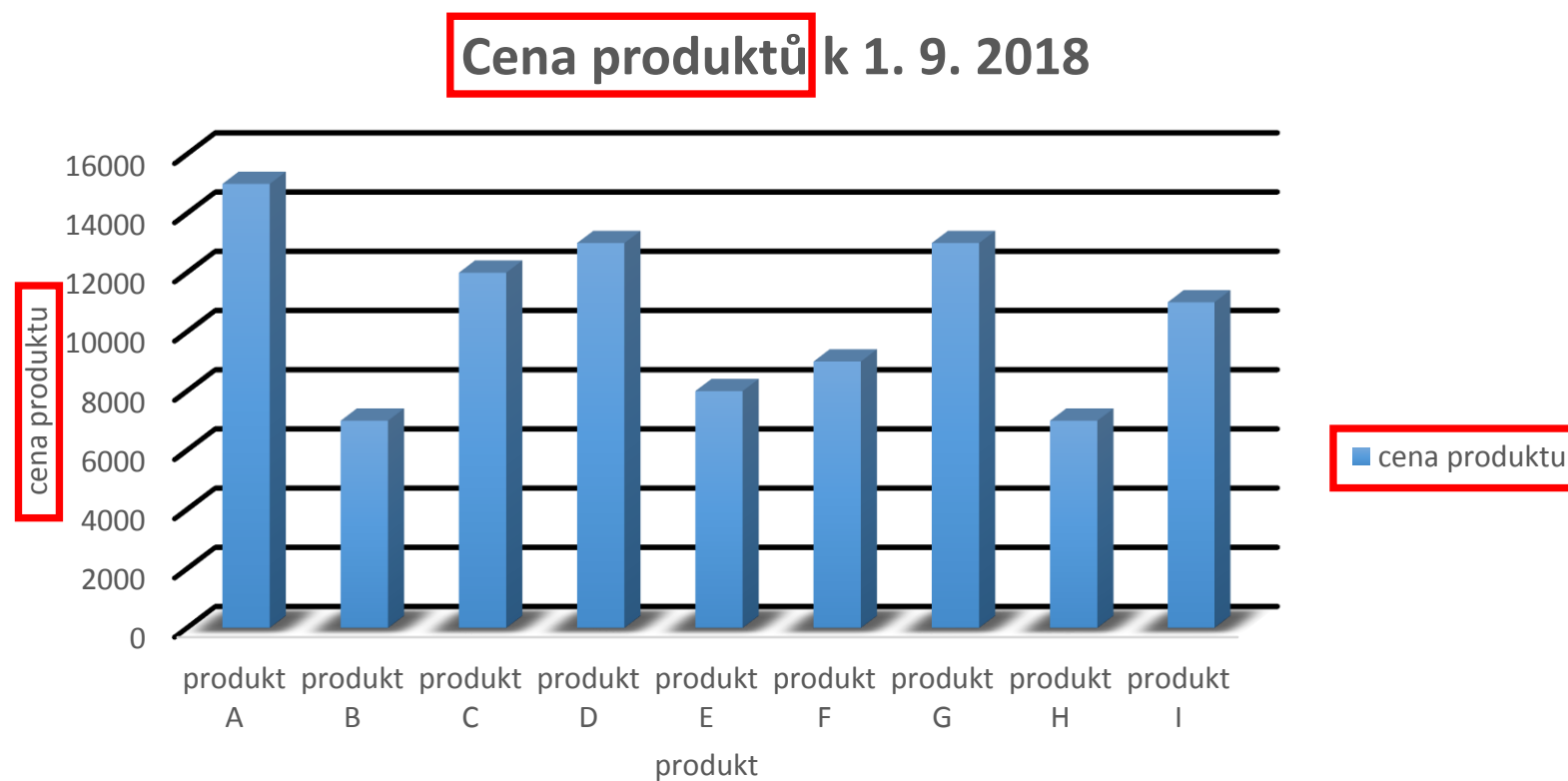
Nepoužívejte barevné pozadí grafů!



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

Neopakujte informace!



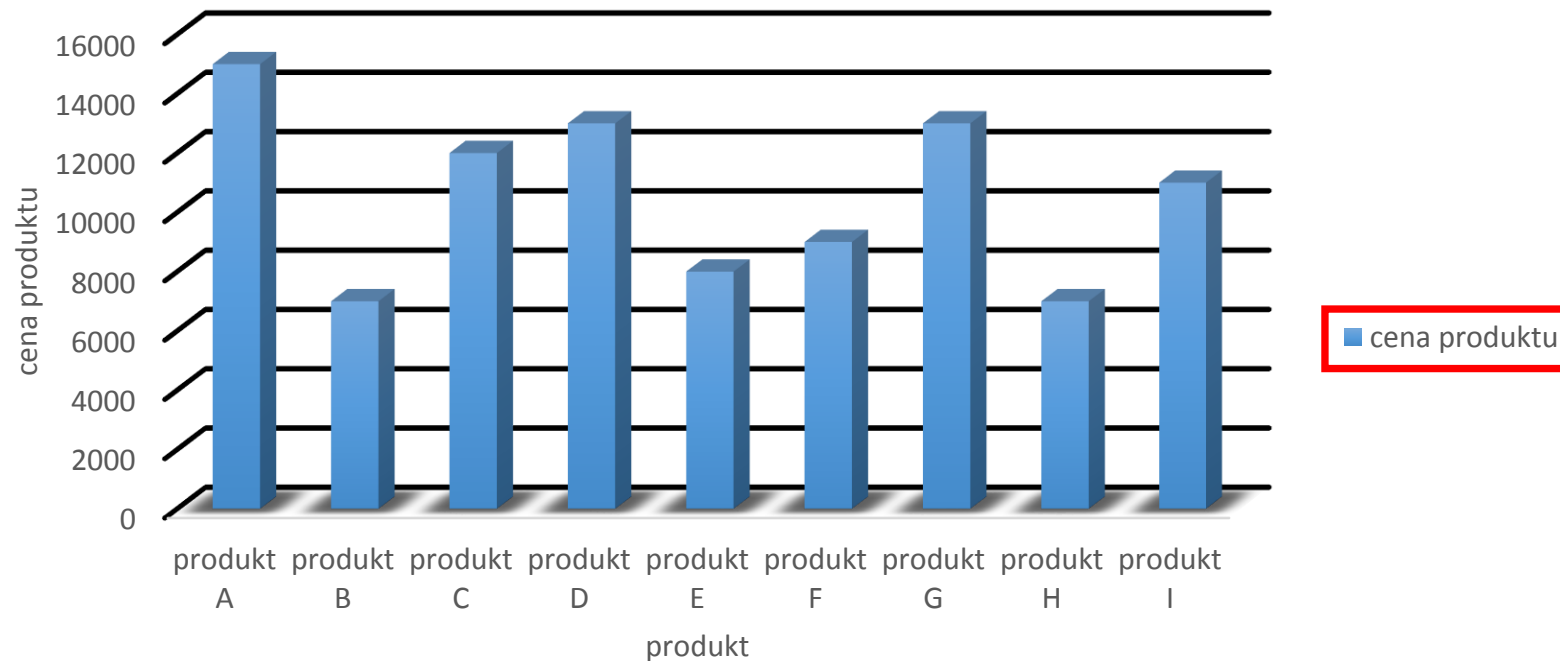
Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

Neopakujte informace!

V tomto případě je legenda zcela nadbytečná. V případě, že ji potřebujete, zvažte její umístění!

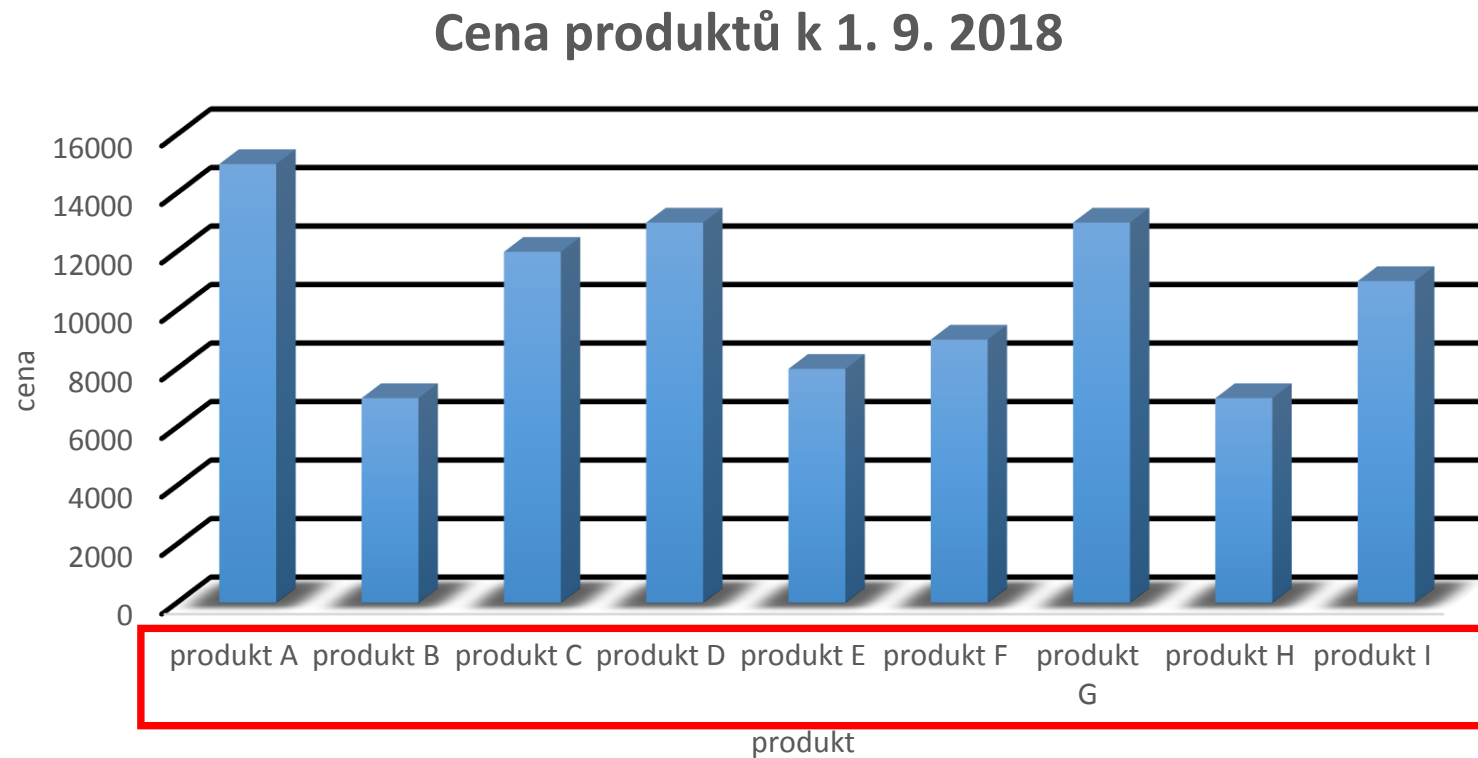
Cena produktů k 1. 9. 2018



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

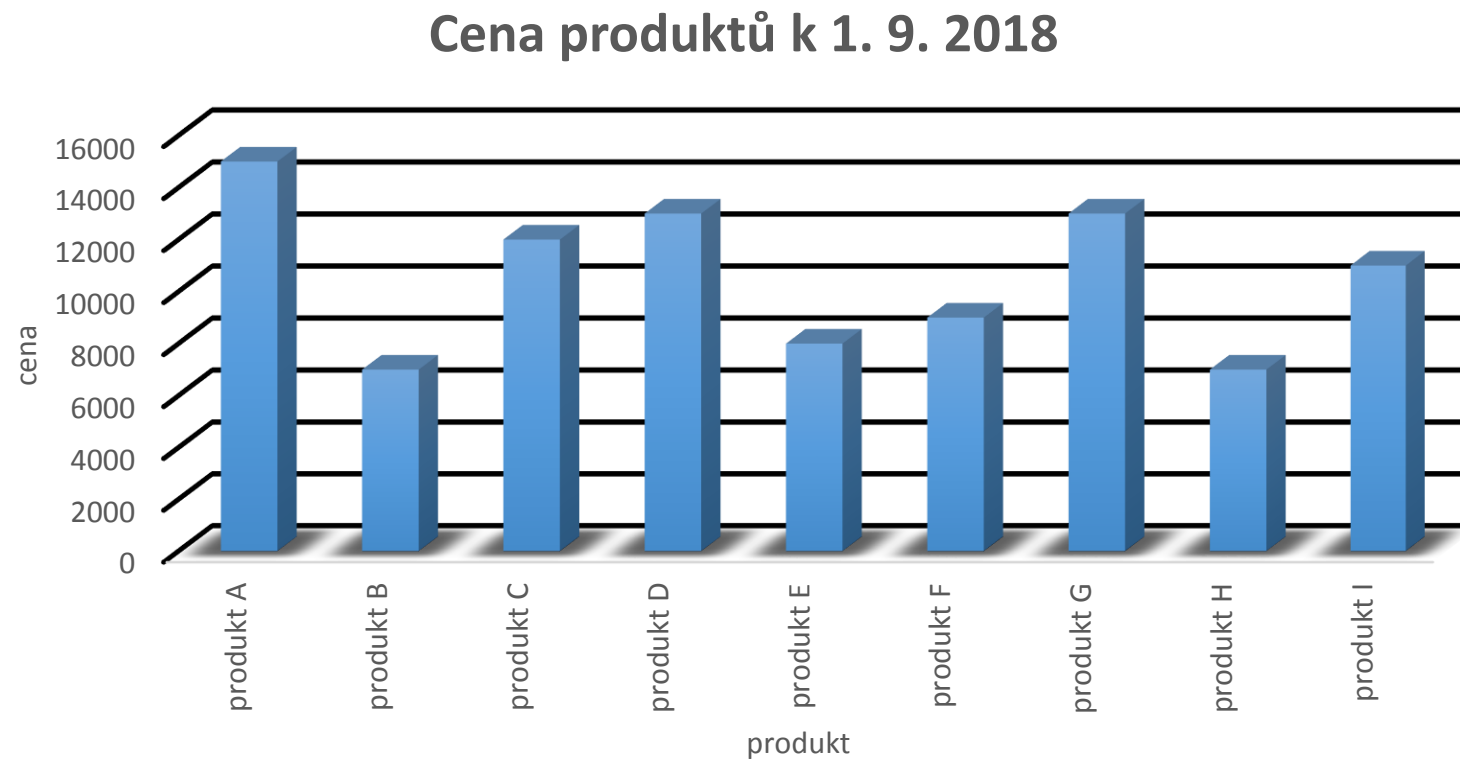
Zvažte úpravu popisku horizontální osy!



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

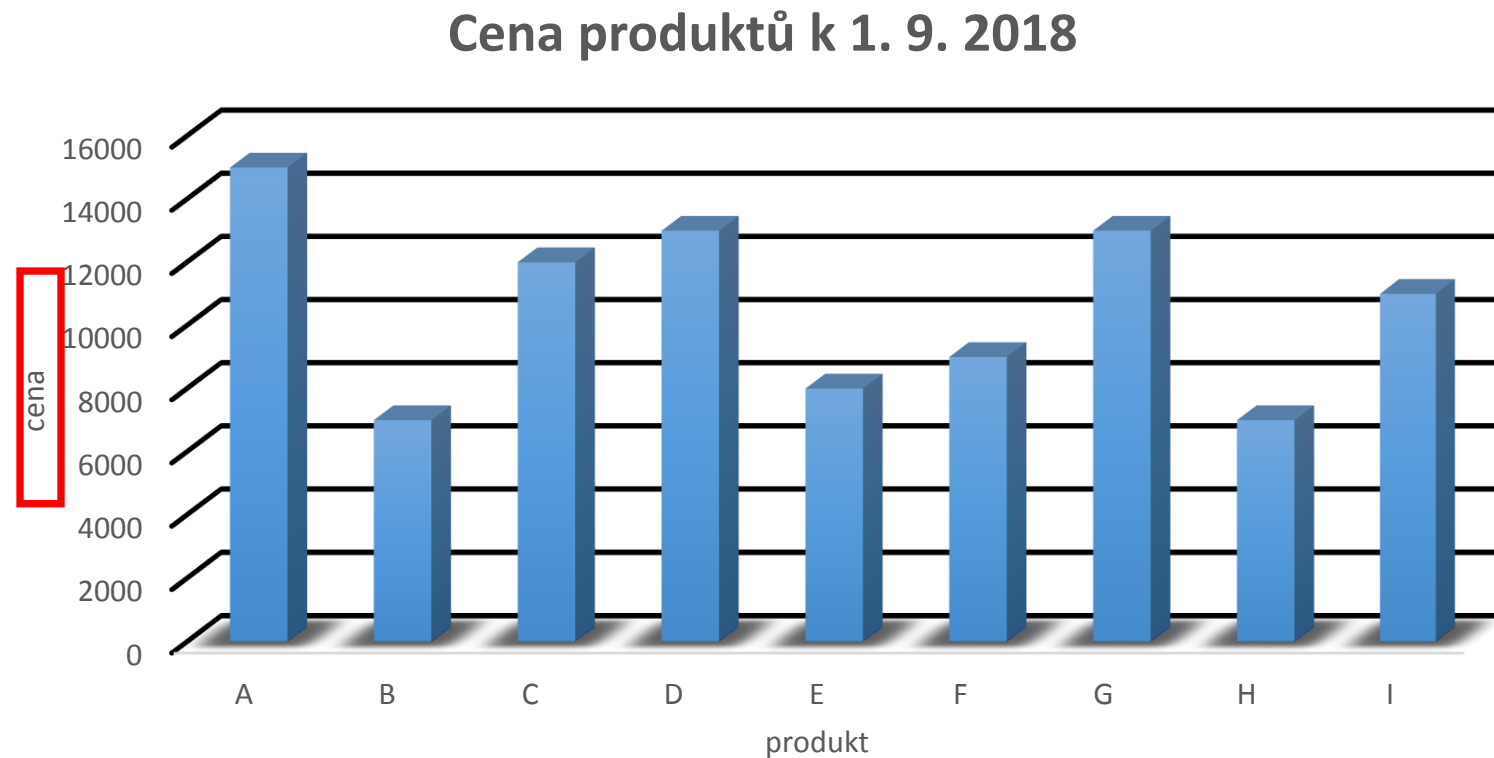
Zvažte úpravu popisku horizontální osy!



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

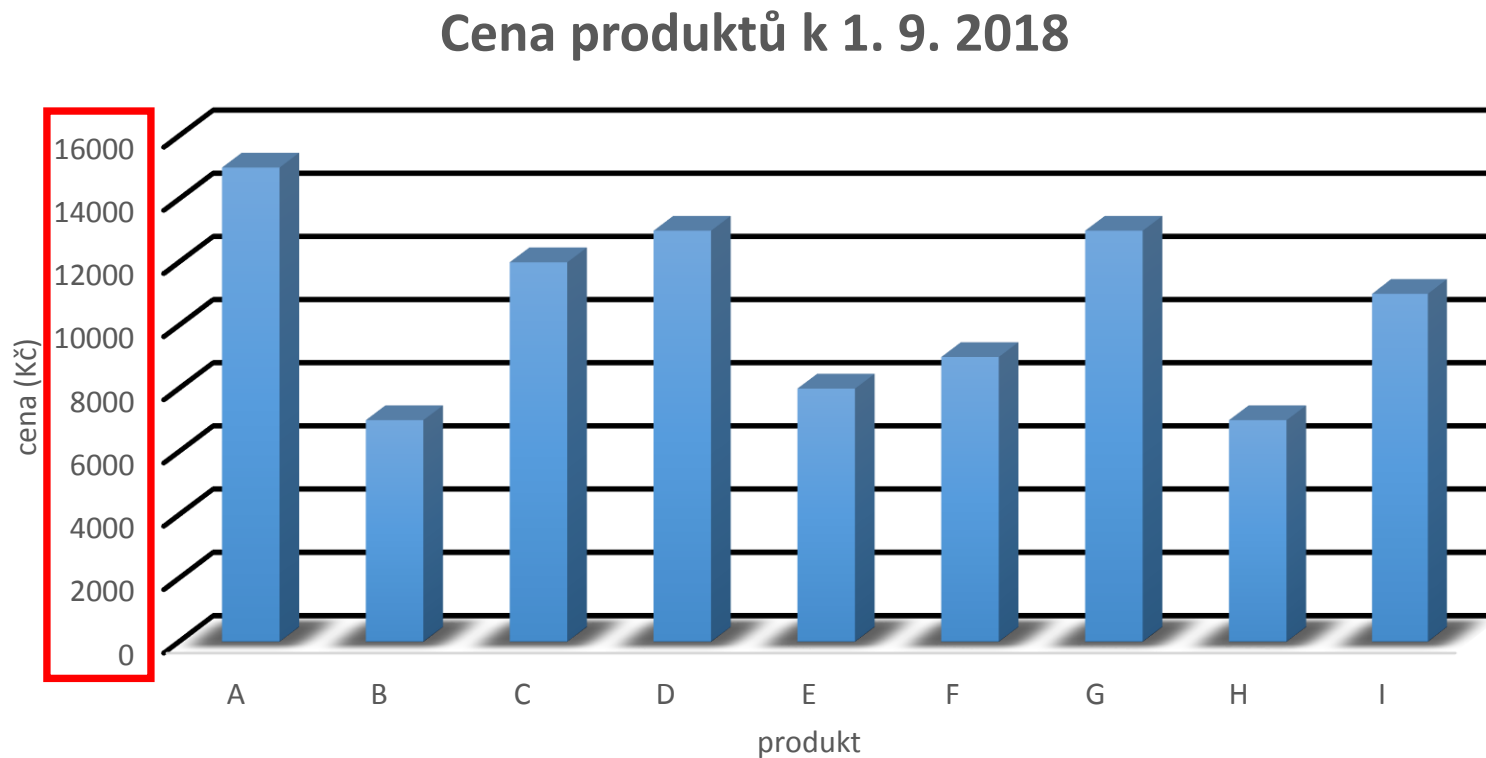
Vždy uvádějte jednotky!



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

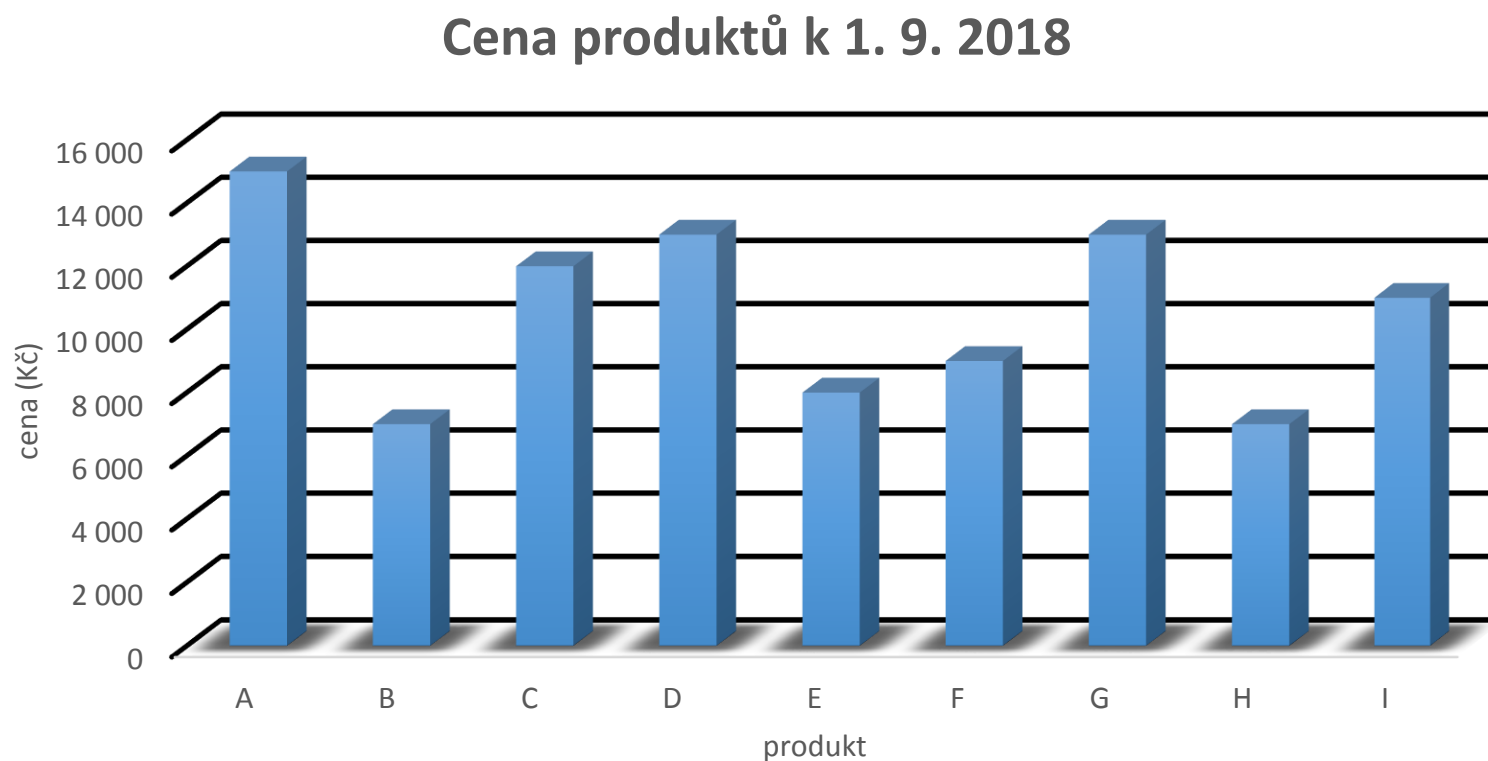
Nepoužívejte neefektivní nuly! (Nebo alespoň použijte oddělovače tisíců!)



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

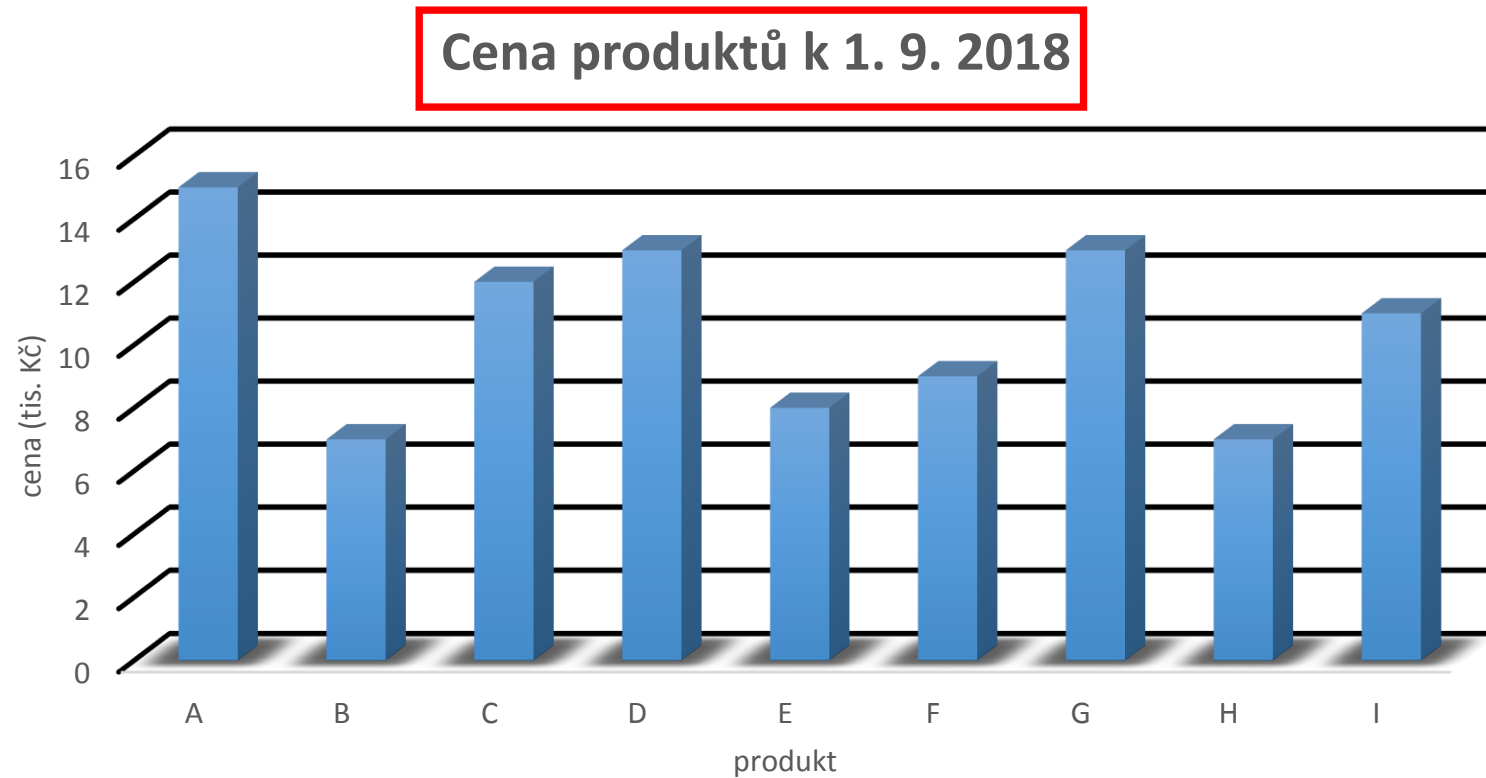
Nepoužívejte neefektivní nuly! (Nebo alespoň použijte oddělovače tisíců!)



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupkové grafy

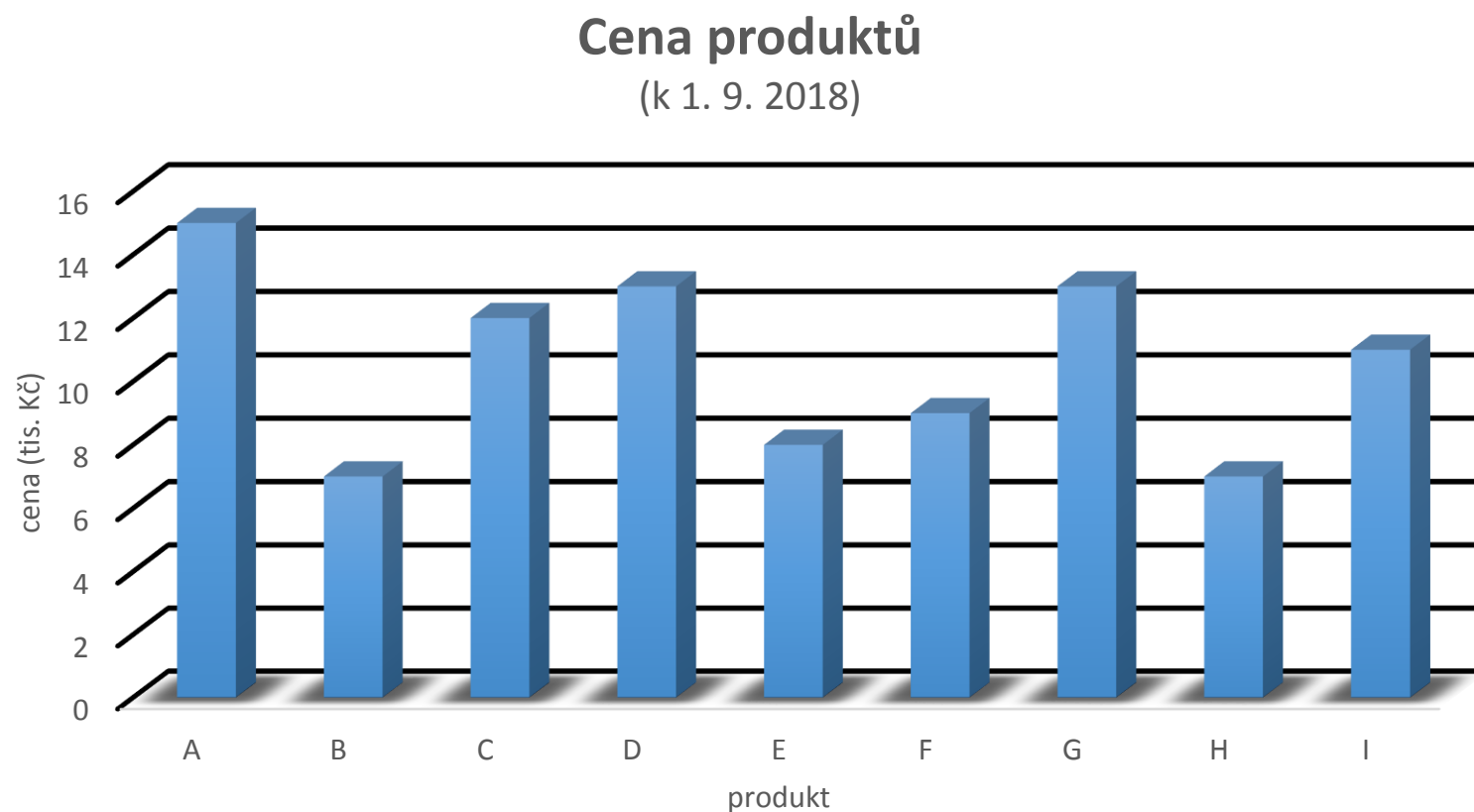
Používejte strukturované nadpisy!



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

Zvažte zvýraznění mřížky!

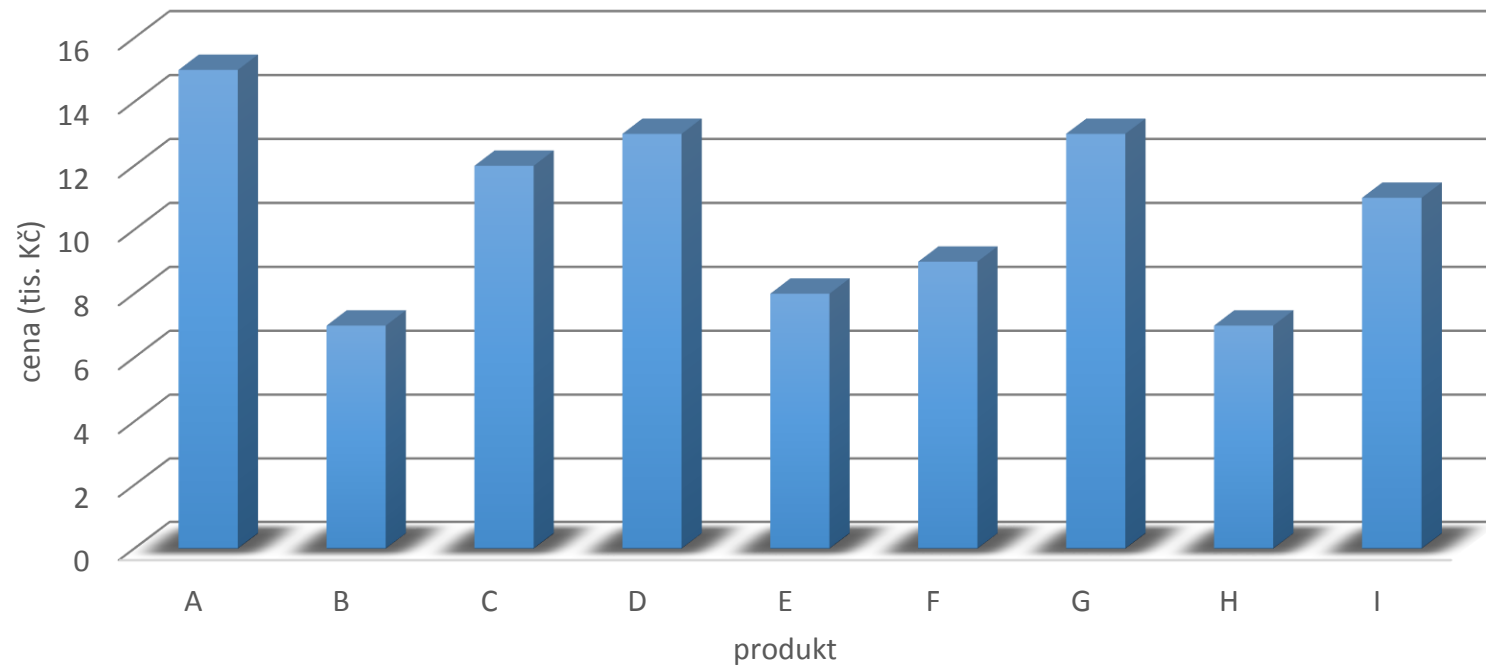


Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

Nepoužívejte 3D grafy!

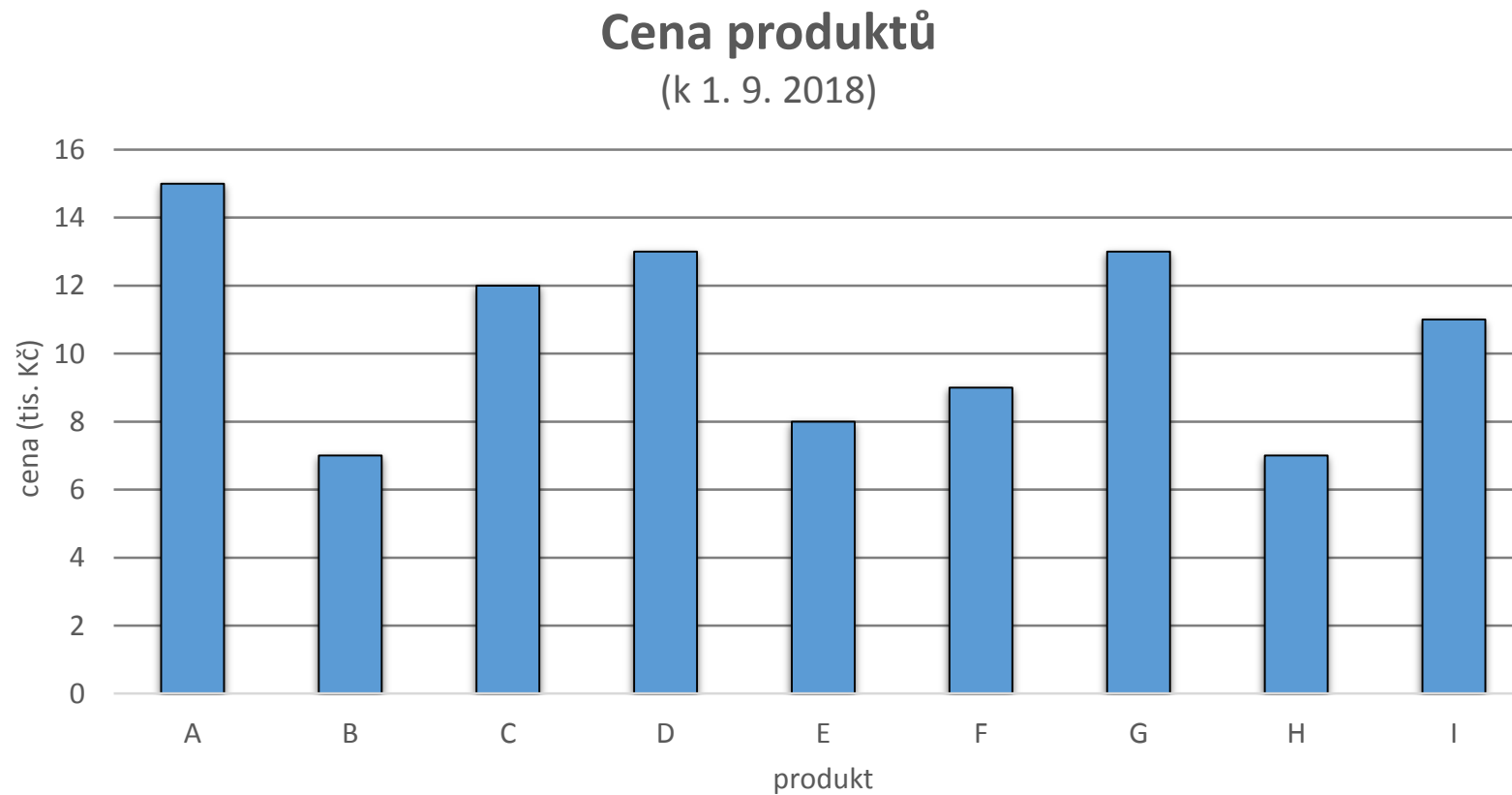
Cena produktů
(k 1. 9. 2018)



Popisná statistika – Kvalitativní znak

- Tipy nejen pro sloupcové grafy

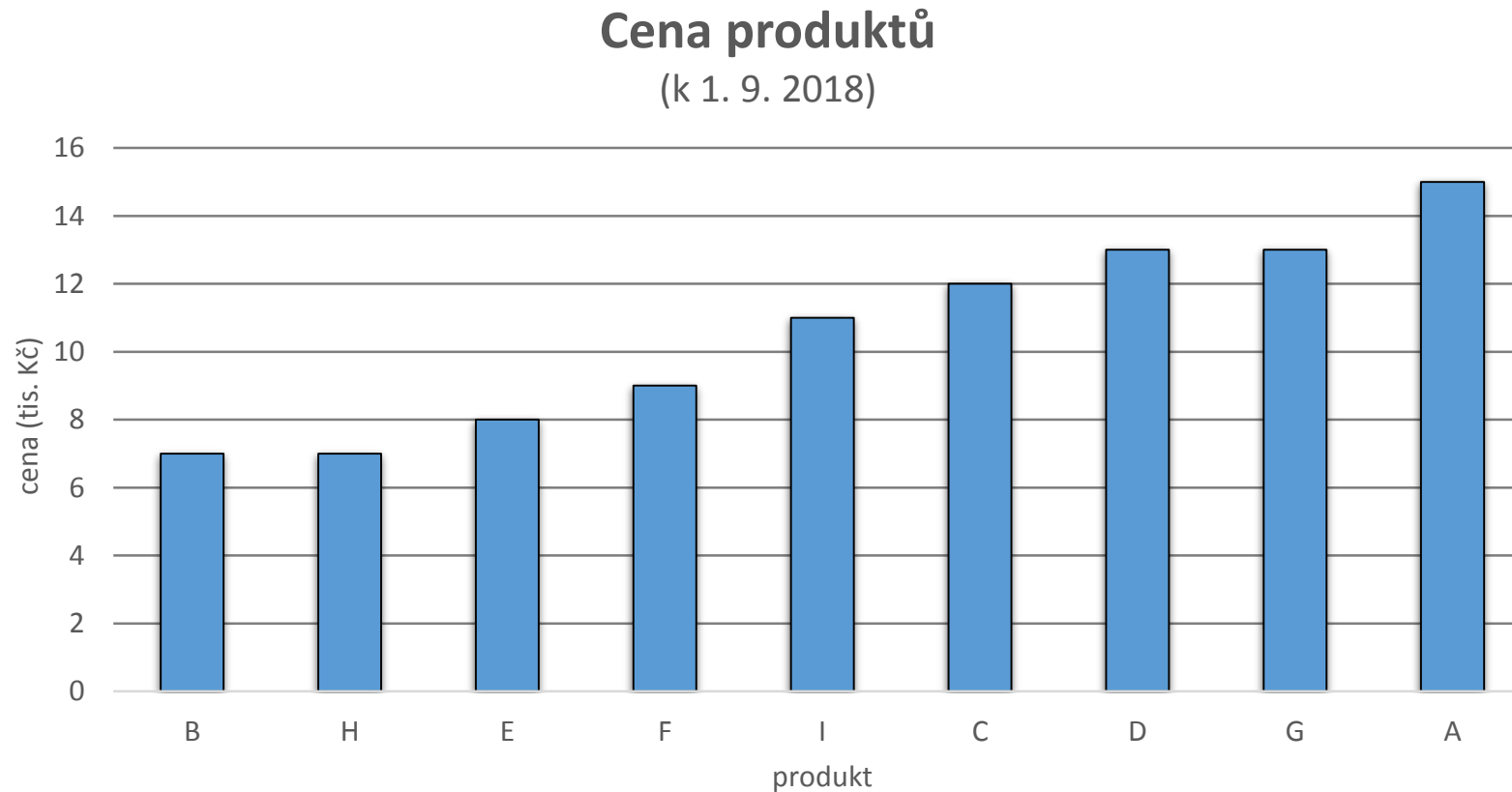
Pokud je to vhodné, seřaďte varianty proměnné!



Popisná statistika – Kvalitativní znak

- **Tipy nejen pro sloupcové grafy**

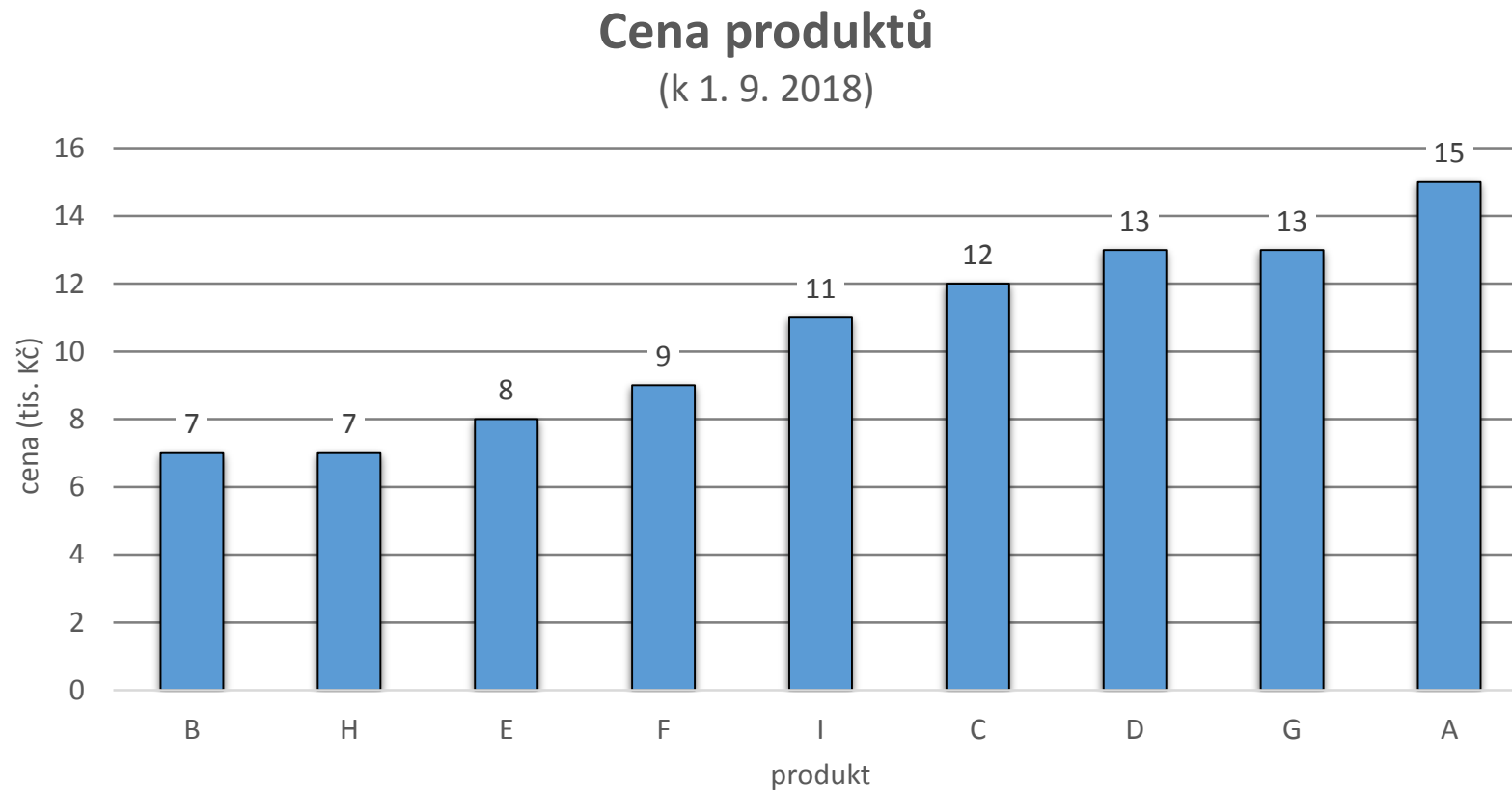
Nejsou-li v grafu uvedeny absolutní četnosti, obvykle je nedokážeme „od oka“ přesně odečíst!

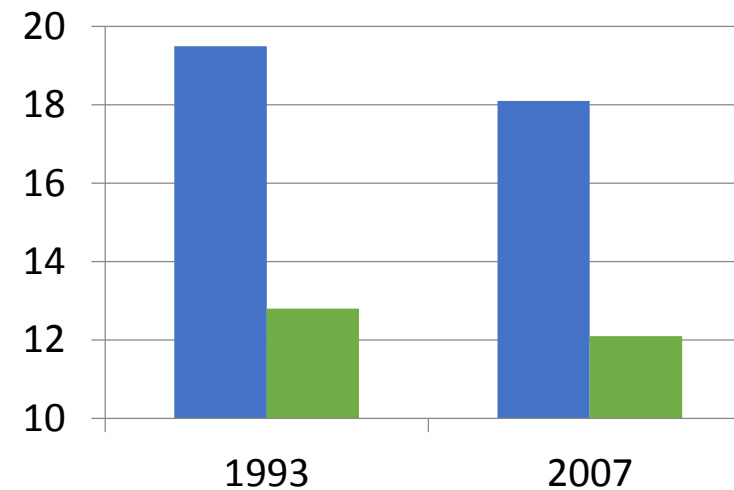
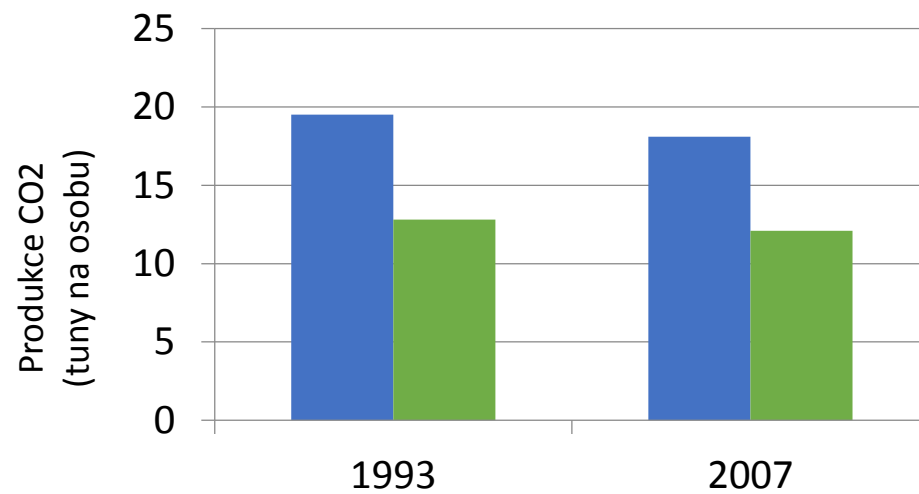


Popisná statistika – Kvalitativní znak

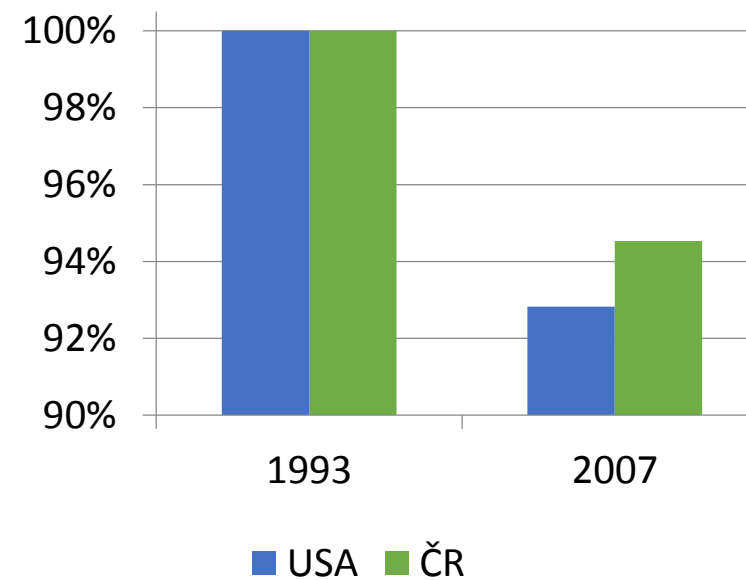
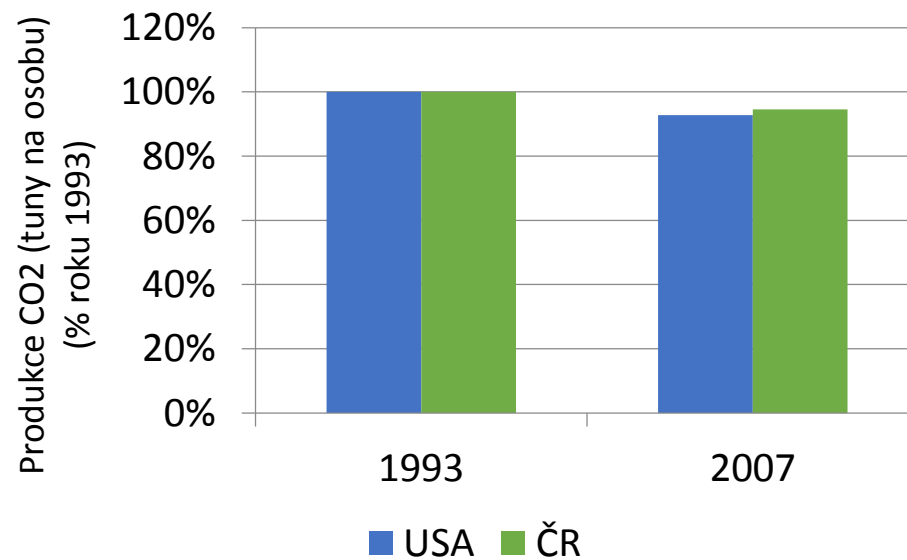
- **Tipy nejen pro sloupcové grafy**

V případě, že graf je součástí publikace, musí obsahovat titulek. V tomto případě obvykle není třeba používat nadpis grafu. Informace obsažena v nadpisu bývá uváděna v titulku.



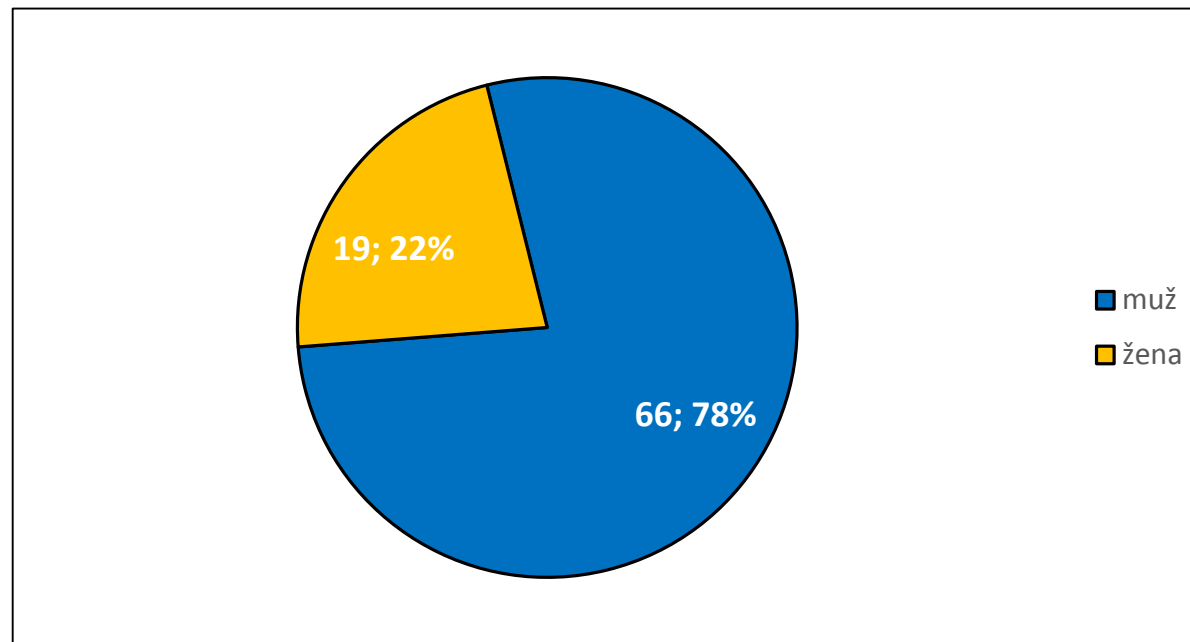


Který z grafů je „správný“?



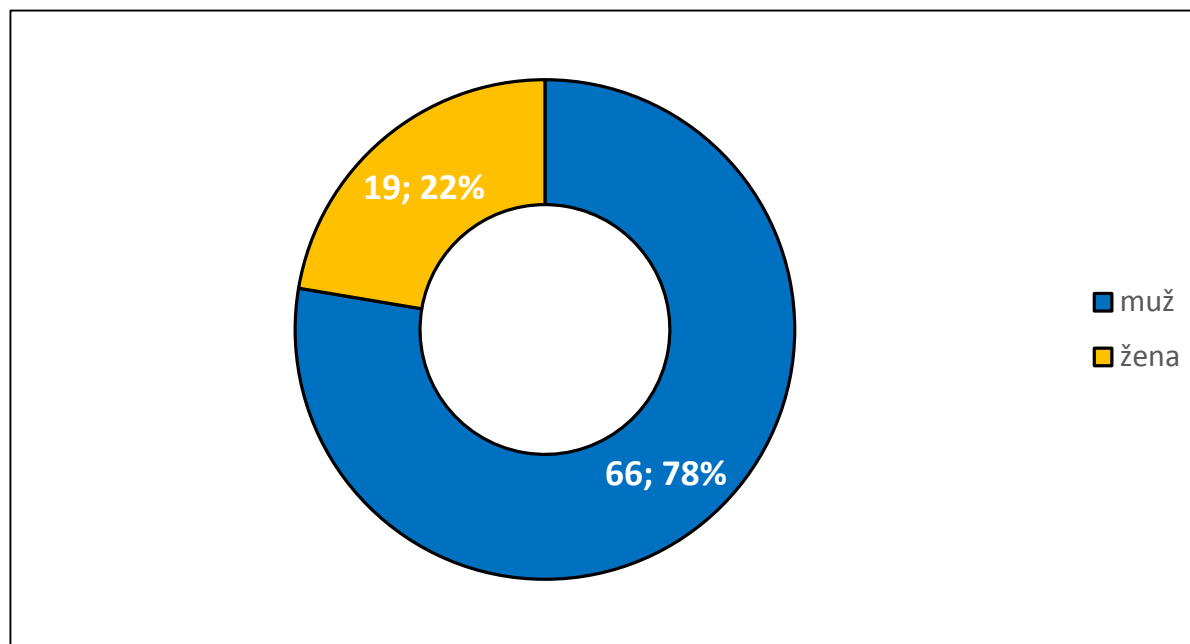
Popisná statistika – Kvalitativní znak

- Výsečový graf



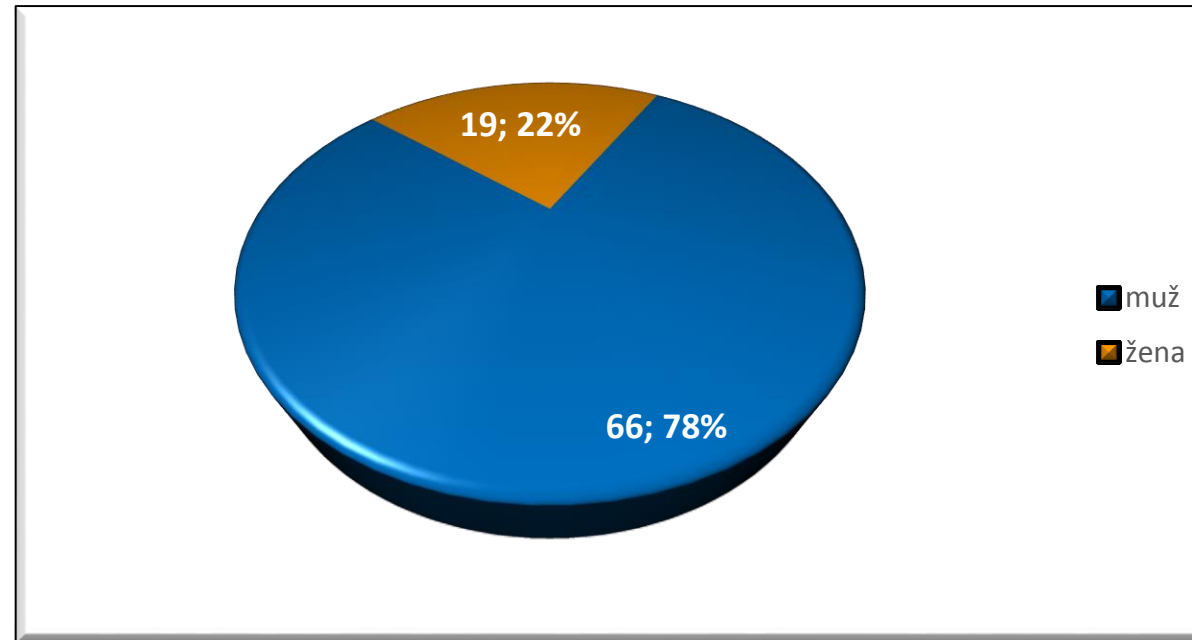
Popisná statistika – Kvalitativní znak

- Prstencový graf



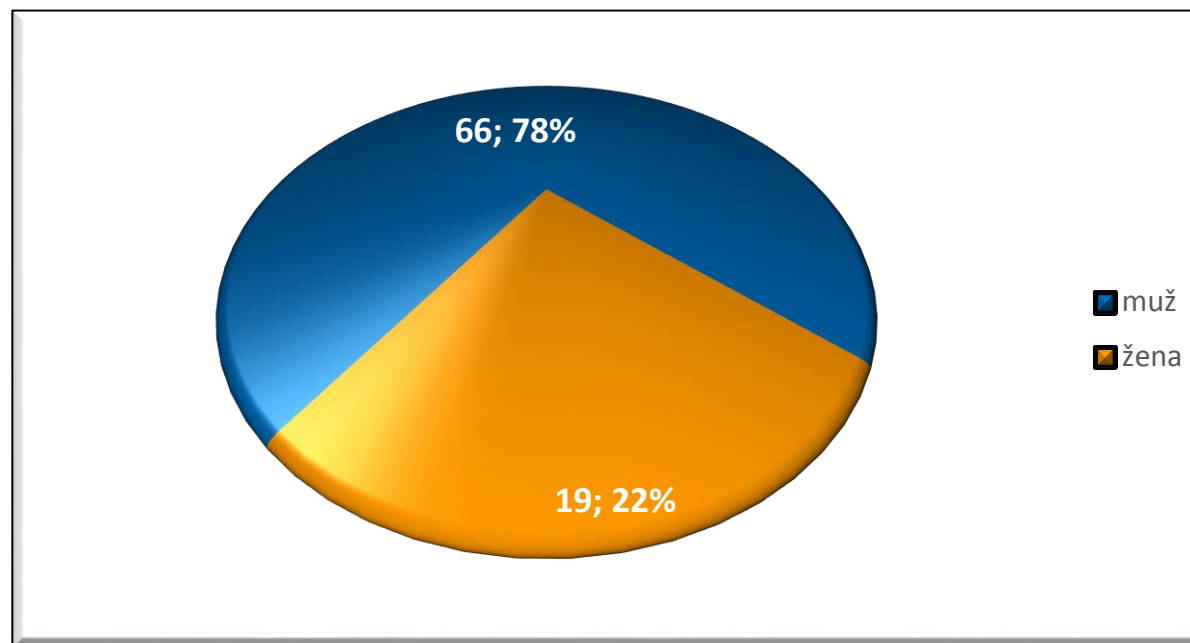
Popisná statistika – Kvalitativní znak

- 3D výsečový graf



Popisná statistika – Kvalitativní znak

- 3D výsečový graf



Pozor na vypovídací
schopnost 3D grafů!

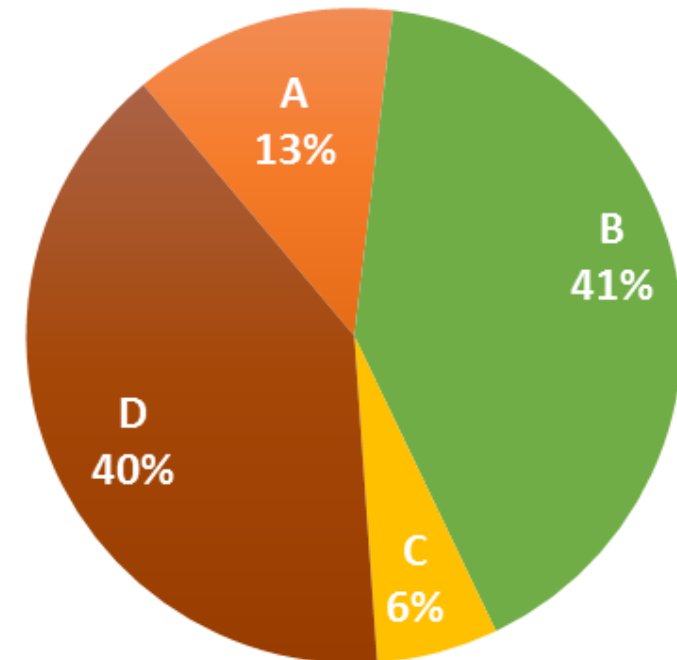
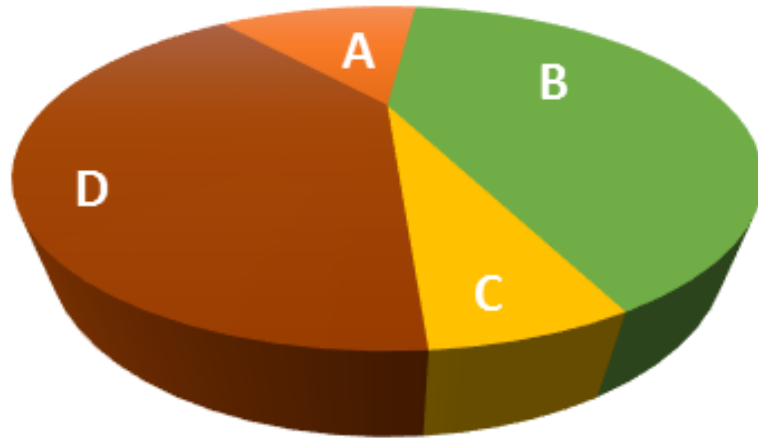


Popisná statistika – Kvalitativní znak

- 3D vs 2D výsečový graf

Jaký je poměr výsečí A a C?

Jaký je poměr výsečí B a D?



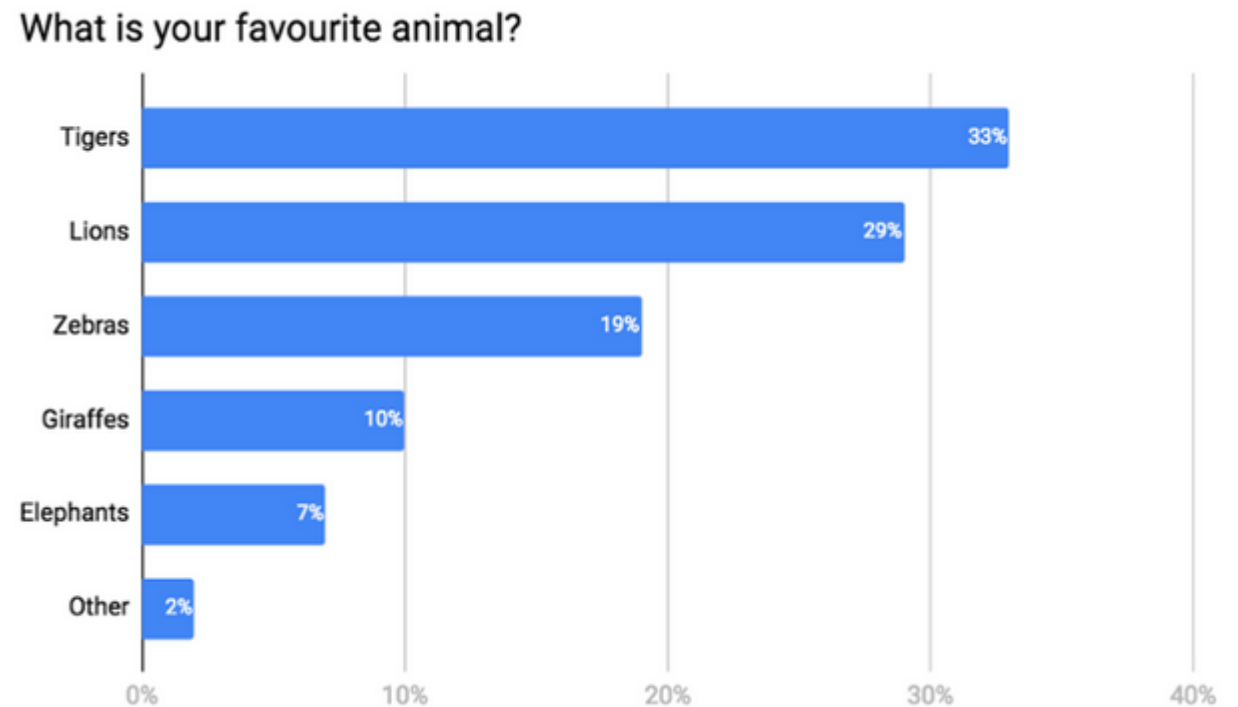
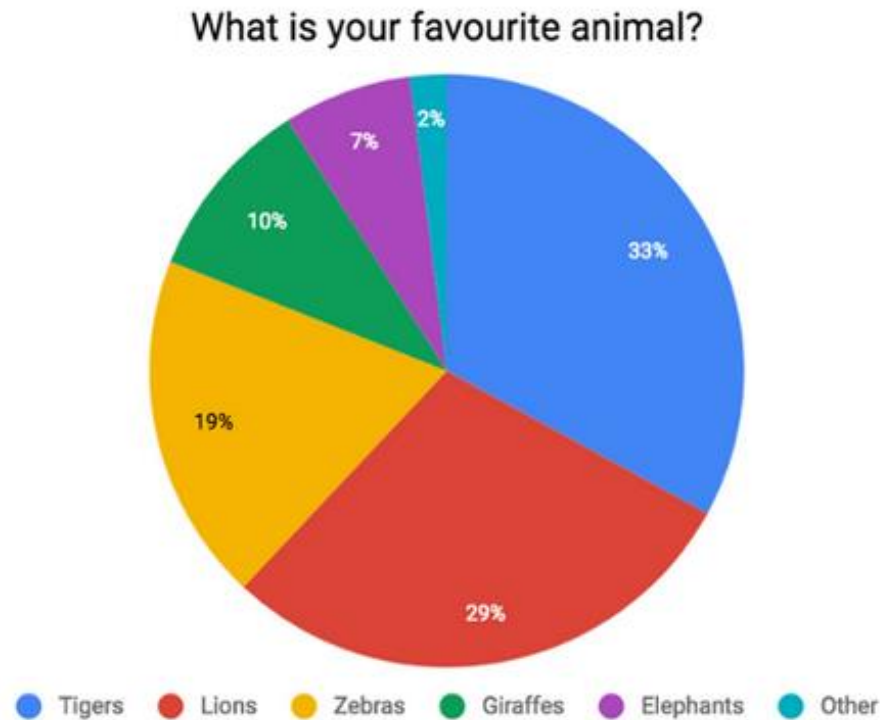
Anketa

Jste pro navýšení hodinové dotace matematiky?



Nezapomínejte, že relativní četnosti byste měli uvádět pouze jako doplněk četností absolutních!

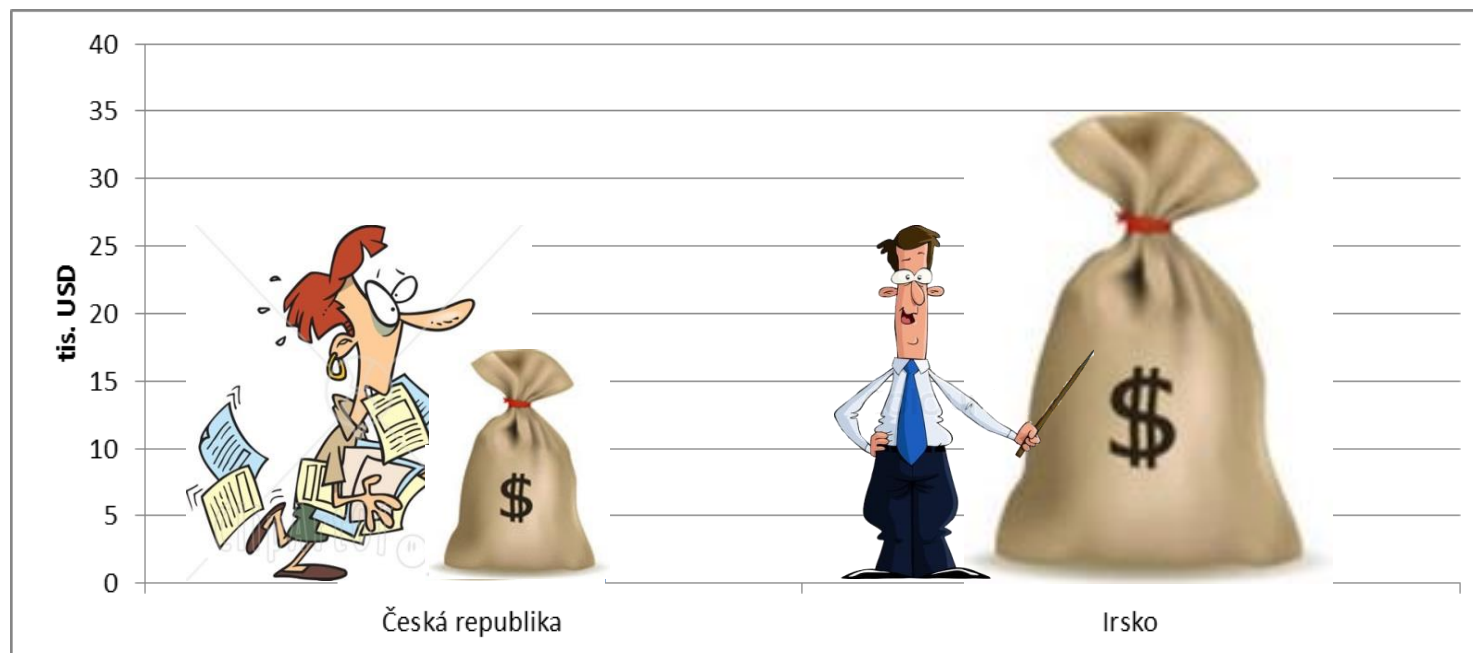
Je výsečový graf tou správnou volbou?



*Are Tigers more popular than Lions? How much more popular are Zebras than Giraffes?
Tigers are more popular than Lions! Zebras are twice as popular as Giraffes!*

Obrázkové grafy – užiteční pomocníci?

Obrázkové grafy



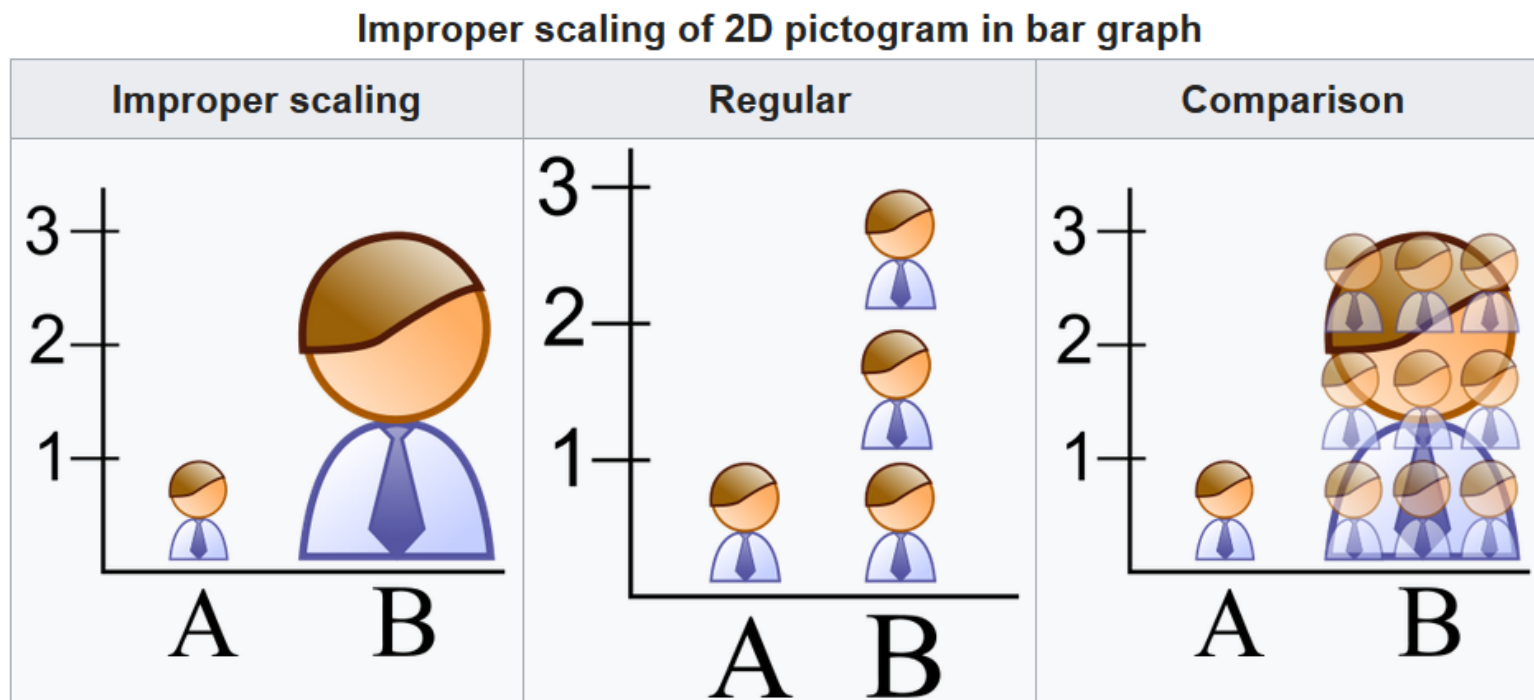
*Srovnání průměrných ročních nástupních platů učitelů středních škol
v ČR (17 244 \$) a Irsku (34 604 \$)*

Obrázkové grafy



*Srovnání průměrných ročních nástupních platů učitelů středních škol
v ČR (17 244 \$) a Irsku (34 604 \$)*

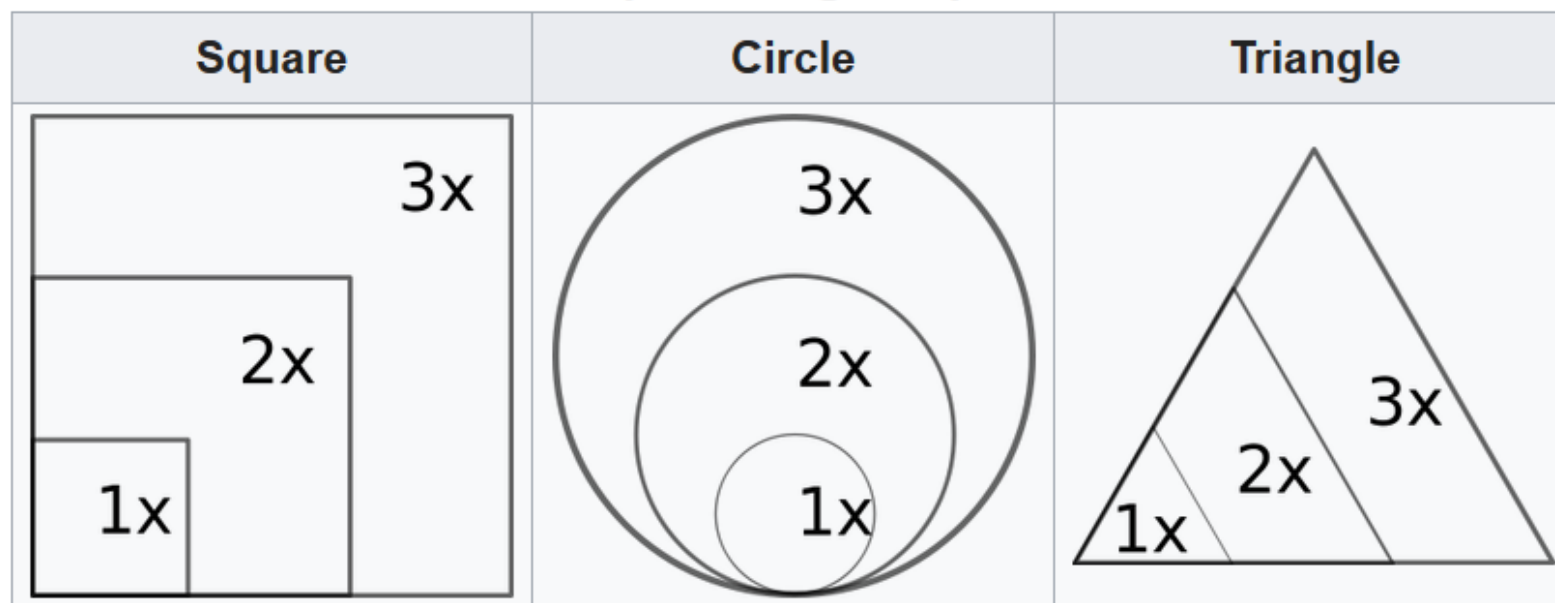
Obrázkové grafy



Zdroj: UTTS, Jessica M. Seeing through statistics. 3rd ed. Belmont, CA: Thomson, Brooks/Cole, c2005. ISBN 0-534-39402-7.

Obrázkové grafy

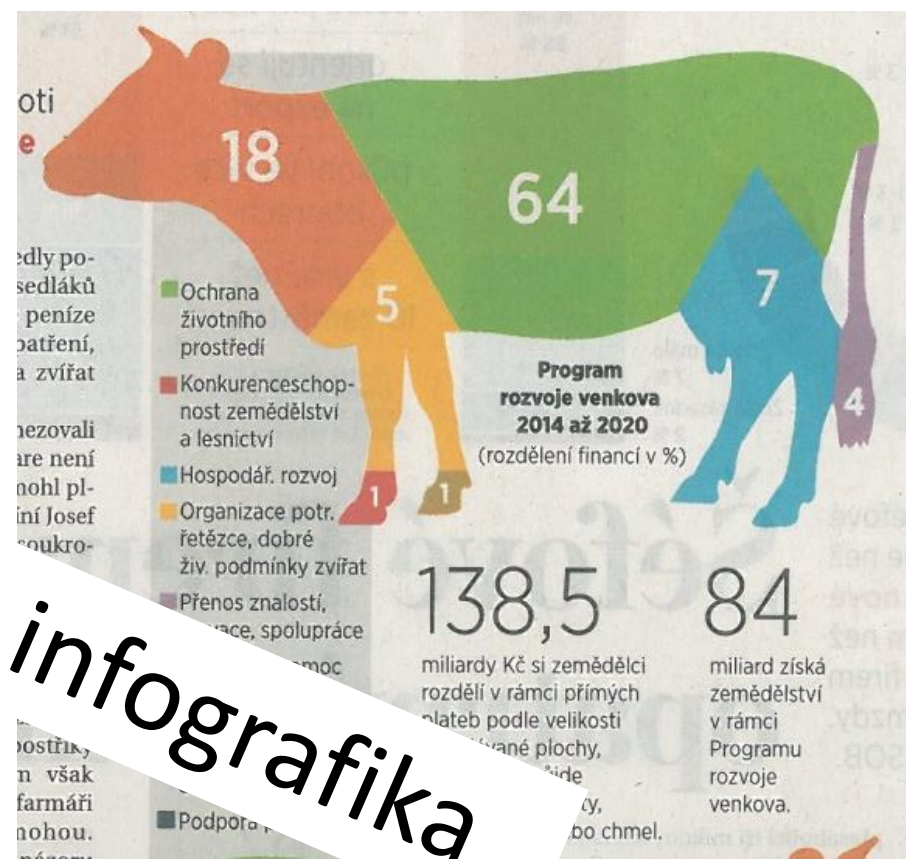
2D shape scaling comparison



Zdroj: UTTS, Jessica M. Seeing through statistics. 3rd ed. Belmont, CA: Thomson, Brooks/Cole, c2005. ISBN 0-534-39402-7.

Pár příkladů z praxe

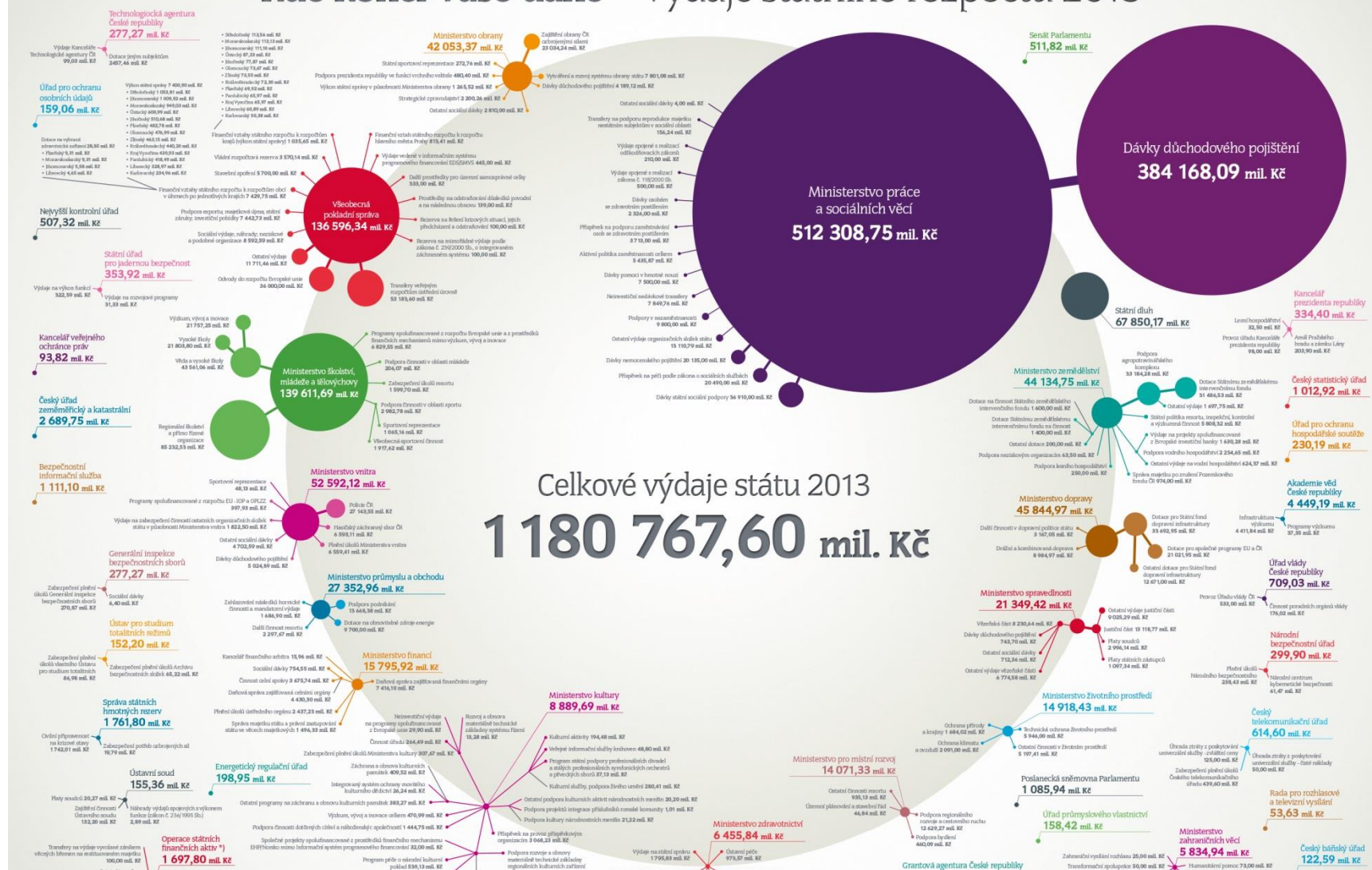
Obrázkové grafy – užiteční pomocníci?



(Zdroj: Mf Dnes, 10. 7. 2014:

Zemědělci si rozdělí miliardy. Krávy a vepři se budou mít lépe.

Kde končí Vaše daně – Výdaje státního rozpočtu 2013



„Úžasná infografika o výdajích státního rozpočtu České republiky v roce 2013“

Zdroj: <http://www.estat.cz/zpravy/informace-k-projektum/kde-konci-vase-dane/>

Celkové výdaje státu 2014
1 211 307,51 mil. Kč
+2,59 %

Dávky důchodového pojištění
381 584,93 mil. Kč
+2,65 %

Ministerstvo práce a sociálních věcí
525 164,13 mil. Kč
+2,51 %

Všeobecná pokladniční správa
144 130,71 mil. Kč
+5,52 %

Ministerstvo financí
16 759,85 mil. Kč
+6,10 %

Ministerstvo zdravotnictví
6 810,66 mil. Kč
+5,50 %

Ministerstvo školství, mládeže a tělovýchovy
137 301,16 mil. Kč
-1,65 %

Ministerstvo zahraničních věcí
5 815,01 mil. Kč
-0,34 %

Ministerstvo životního prostředí
12 958,89 mil. Kč
-13,13 %

Ministerstvo spravedlnosti
22 120,78 mil. Kč
+3,61 %

Český telekomunikační úřad
742,03 mil. Kč
+20,73 %

Český statistický úřad
932,11 mil. Kč
-7,98 %

Český úřad zeměměřičský a katastrální
2 756,23 mil. Kč
+2,47 %

Úřad pro ochranu osobních údajů
154,35 mil. Kč
-2,96 %

Úřad pro rozhlášení a televizní vysílání
54,99 mil. Kč
+2,52 %

Energetický regulační úřad
215,95 mil. Kč
+8,54 %

Úřad pro ochranu hospodářské soutěže
244,02 mil. Kč
+6,00 %

Generální inspekce bezpečnostních sborů
279,72 mil. Kč
+0,89 %

Technická agentura České republiky
2 962,49 mil. Kč
+15,88 %

Grantová agentura České republiky
3 464,55 mil. Kč
+4,69 %

Ministerstvo průmyslu a obchodu
35 006,13 mil. Kč
-27,98 %

Ministerstvo zemědělství
52 170,67 mil. Kč
+18,21 %

Ministerstvo dopravy
43 875,46 mil. Kč
-4,30 %

Ministerstvo pro místní rozvoj
11 710,88 mil. Kč
-16,77 %

Poslanecká sněmovna Parlamentu
1 076,57 mil. Kč
-0,86 %

Správa státních hmotných rezerv
1 969,75 mil. Kč
+11,80 %

Státní úřad pro jadernou bezpečnost
376,30 mil. Kč
+6,32 %

Česká národní banka
107,38 mil. Kč
+0,00 %

Bezpečnostní informační služba
1 152,68 mil. Kč
+3,74 %

Nekýšší kontrolní úřad
500,39 mil. Kč
-1,35 %

Úřad pro ochranu hospodářské soutěže
244,02 mil. Kč
+6,00 %

Ministerstvo školství, mládeže a tělovýchovy
137 301,16 mil. Kč
-1,65 %

Ministerstvo zdravotnictví
6 810,66 mil. Kč
+5,50 %

Ministerstvo zahraničních věcí
5 815,01 mil. Kč
-0,34 %

Ministerstvo životního prostředí
12 958,89 mil. Kč
-13,13 %

Ministerstvo spravedlnosti
22 120,78 mil. Kč
+3,61 %

Český telekomunikační úřad
742,03 mil. Kč
+20,73 %

Český statistický úřad
932,11 mil. Kč
-7,98 %

Český úřad zeměměřičský a katastrální
2 756,23 mil. Kč
+2,47 %

Úřad pro ochranu osobních údajů
154,35 mil. Kč
-2,96 %

Úřad pro rozhlášení a televizní vysílání
54,99 mil. Kč
+2,52 %

Energetický regulační úřad
215,95 mil. Kč
+8,54 %

Úřad pro ochranu hospodářské soutěže
244,02 mil. Kč
+6,00 %

Generální inspekce bezpečnostních sborů
279,72 mil. Kč
+0,89 %

Technická agentura České republiky
2 962,49 mil. Kč
+15,88 %

Grantová agentura České republiky
3 464,55 mil. Kč
+4,69 %

Ministerstvo průmyslu a obchodu
35 006,13 mil. Kč
-27,98 %

Ministerstvo zemědělství
52 170,67 mil. Kč
+18,21 %

Ministerstvo dopravy
43 875,46 mil. Kč
-4,30 %

Ministerstvo pro místní rozvoj
11 710,88 mil. Kč
-16,77 %

Poslanecká sněmovna Parlamentu
1 076,57 mil. Kč
-0,86 %

Správa státních hmotných rezerv
1 969,75 mil. Kč
+11,80 %

Státní úřad pro jadernou bezpečnost
376,30 mil. Kč
+6,32 %

Česká národní banka
107,38 mil. Kč
+0,00 %

Bezpečnostní informační služba
1 152,68 mil. Kč
+3,74 %

Nekýšší kontrolní úřad
500,39 mil. Kč
-1,35 %

Úřad pro ochranu hospodářské soutěže
244,02 mil. Kč
+6,00 %

Ministerstvo školství, mládeže a tělovýchovy
137 301,16 mil. Kč
-1,65 %

Ministerstvo zdravotnictví
6 810,66 mil. Kč
+5,50 %

Ministerstvo zahraničních věcí
5 815,01 mil. Kč
-0,34 %

Ministerstvo životního prostředí
12 958,89 mil. Kč
-13,13 %

Ministerstvo spravedlnosti
22 120,78 mil. Kč
+3,61 %

Český telekomunikační úřad
742,03 mil. Kč
+20,73 %

Český statistický úřad
932,11 mil. Kč
-7,98 %

Český úřad zeměměřičský a katastrální
2 756,23 mil. Kč
+2,47 %

Úřad pro ochranu osobních údajů
154,35 mil. Kč
-2,96 %

Úřad pro rozhlášení a televizní vysílání
54,99 mil. Kč
+2,52 %

Energetický regulační úřad
215,95 mil. Kč
+8,54 %

Úřad pro ochranu hospodářské soutěže
244,02 mil. Kč
+6,00 %

Generální inspekce bezpečnostních sborů
279,72 mil. Kč
+0,89 %

Technická agentura České republiky
2 962,49 mil. Kč
+15,88 %

Grantová agentura České republiky
3 464,55 mil. Kč
+4,69 %

Ministerstvo průmyslu a obchodu
35 006,13 mil. Kč
-27,98 %

Ministerstvo zemědělství
52 170,67 mil. Kč
+18,21 %

Ministerstvo dopravy
43 875,46 mil. Kč
-4,30 %

Ministerstvo pro místní rozvoj
11 710,88 mil. Kč
-16,77 %

Poslanecká sněmovna Parlamentu
1 076,57 mil. Kč
-0,86 %

Správa státních hmotných rezerv
1 969,75 mil. Kč
+11,80 %

Státní úřad pro jadernou bezpečnost
376,30 mil. Kč
+6,32 %

Česká národní banka
107,38 mil. Kč
+0,00 %

Bezpečnostní informační služba
1 152,68 mil. Kč
+3,74 %

Nekýšší kontrolní úřad
500,39 mil. Kč
-1,35 %

Úřad pro ochranu hospodářské soutěže
244,02 mil. Kč
+6,00 %

Ministerstvo školství, mládeže a tělovýchovy
137 301,16 mil. Kč
-1,65 %

Ministerstvo zdravotnictví
6 810,66 mil. Kč
+5,50 %

Ministerstvo zahraničních věcí
5 815,01 mil. Kč
-0,34 %

Ministerstvo životního prostředí
12 958,89 mil. Kč
-13,13 %

Ministerstvo spravedlnosti
22 120,78 mil. Kč
+3,61 %

Český telekomunikační úřad
742,03 mil. Kč
+20,73 %

Český statistický úřad
932,11 mil. Kč
-7,98 %

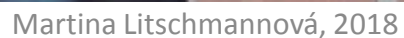
Český úřad zeměměřičský a katastrální
2 756,23 mil. Kč
+2,47 %

Úřad pro ochranu osobních údajů
154,35 mil. Kč
-2,96 %

Úřad pro rozhlášení a televizní vysílání
54,99 mil. Kč
+2,52 %

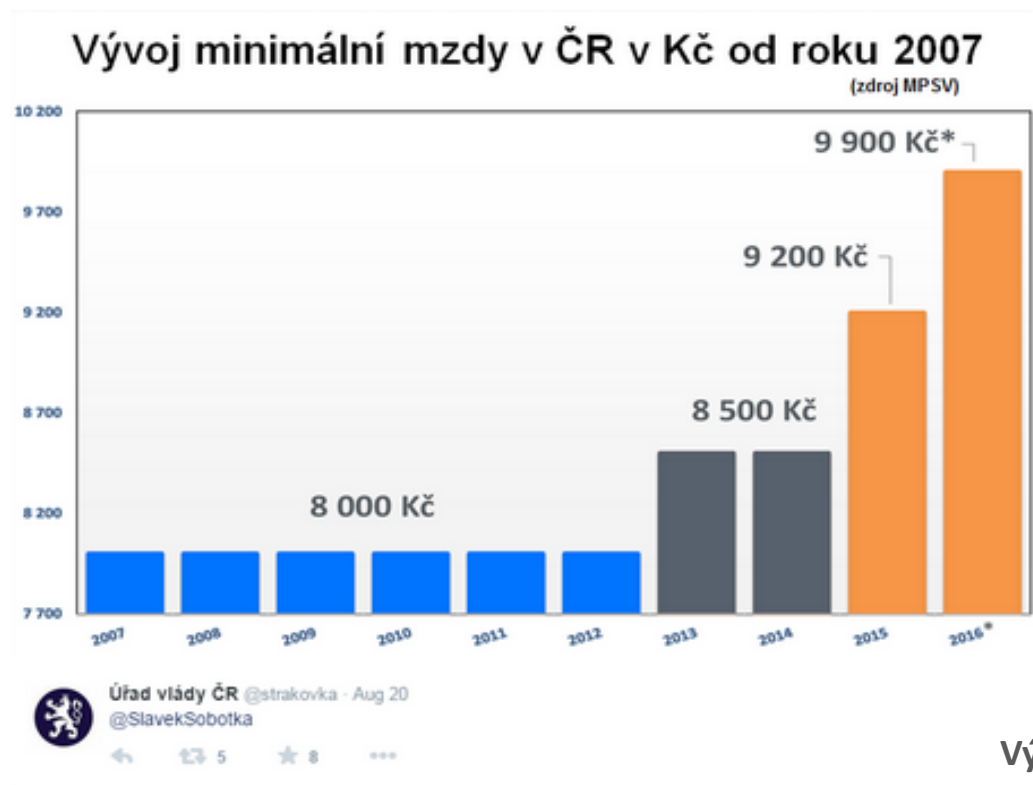
Energetický regulační úřad
215,9

Zdroj: <http://www.estat.cz/zpravy/informace-k-projektum/kde-konci-vase-dane/>



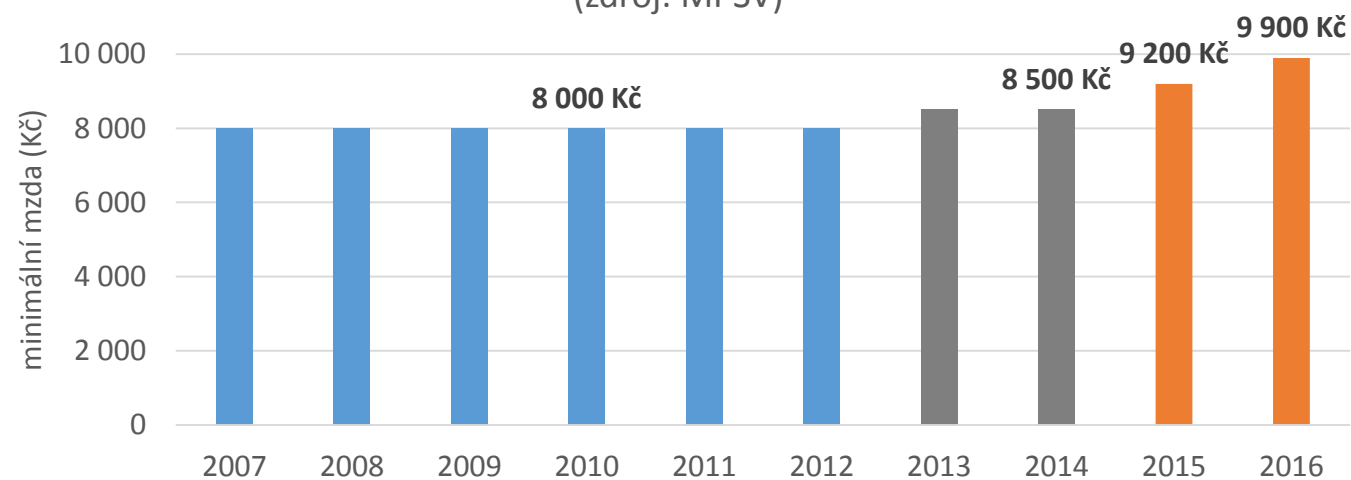
Příklad s klobásou

2. místo: klobása z trhů na náměstí Míru (45 Kč)	
Obsah masa:	79,73 %
Obsah tuku:	24,94 %
Obsah bílkovin:	12,4 %
Obsah vody:	55,56 %
Obsah kolagenu:	2,6 %
Sójová bílkovina:	méně než 0,7 %



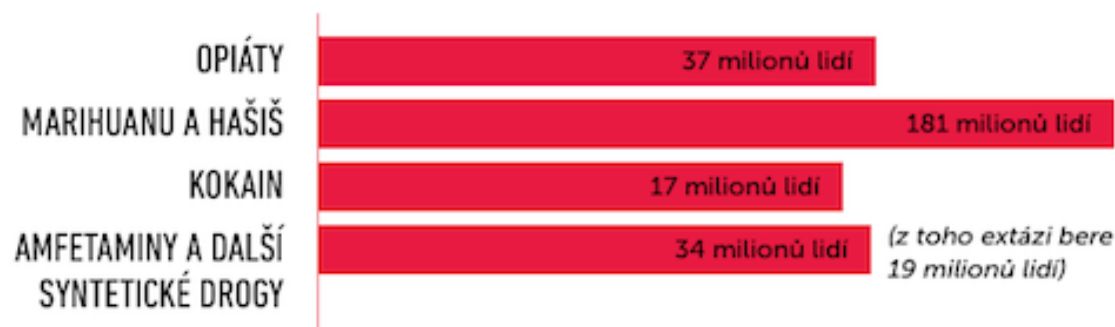
Zdroj: Twitter @strakovka
(20. srpna 2015)

Vývoj minimální mzdy v ČR od roku 2007
(zdroj: MPSV)



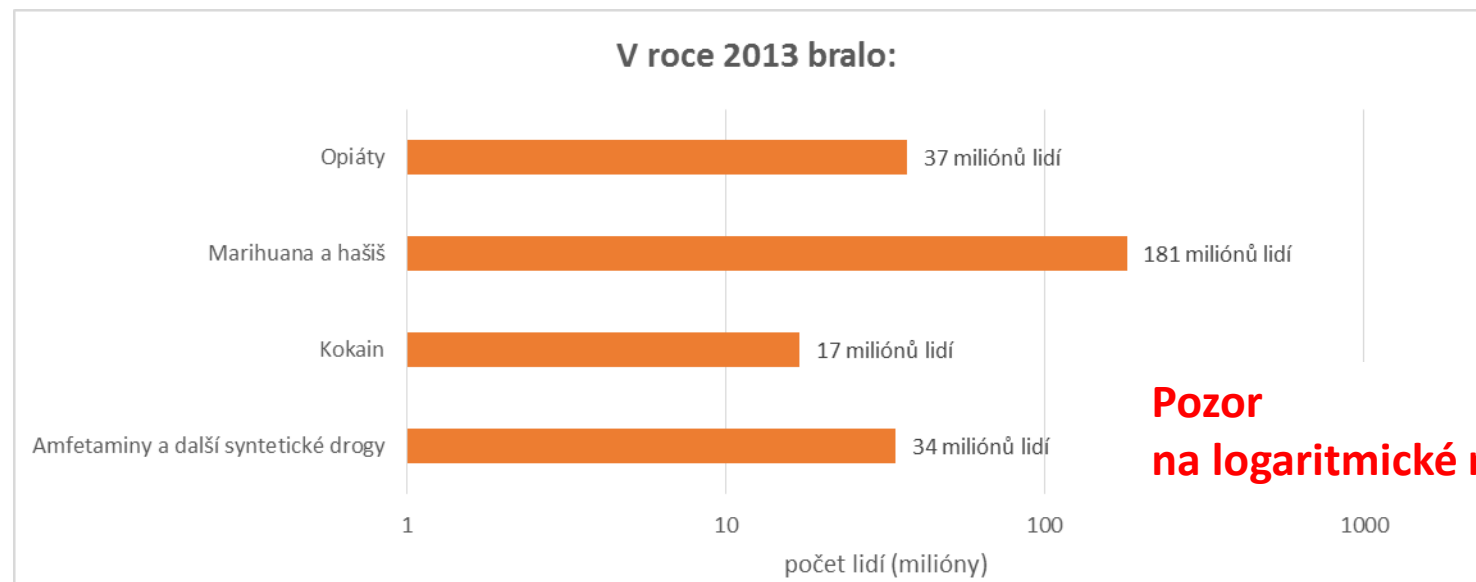
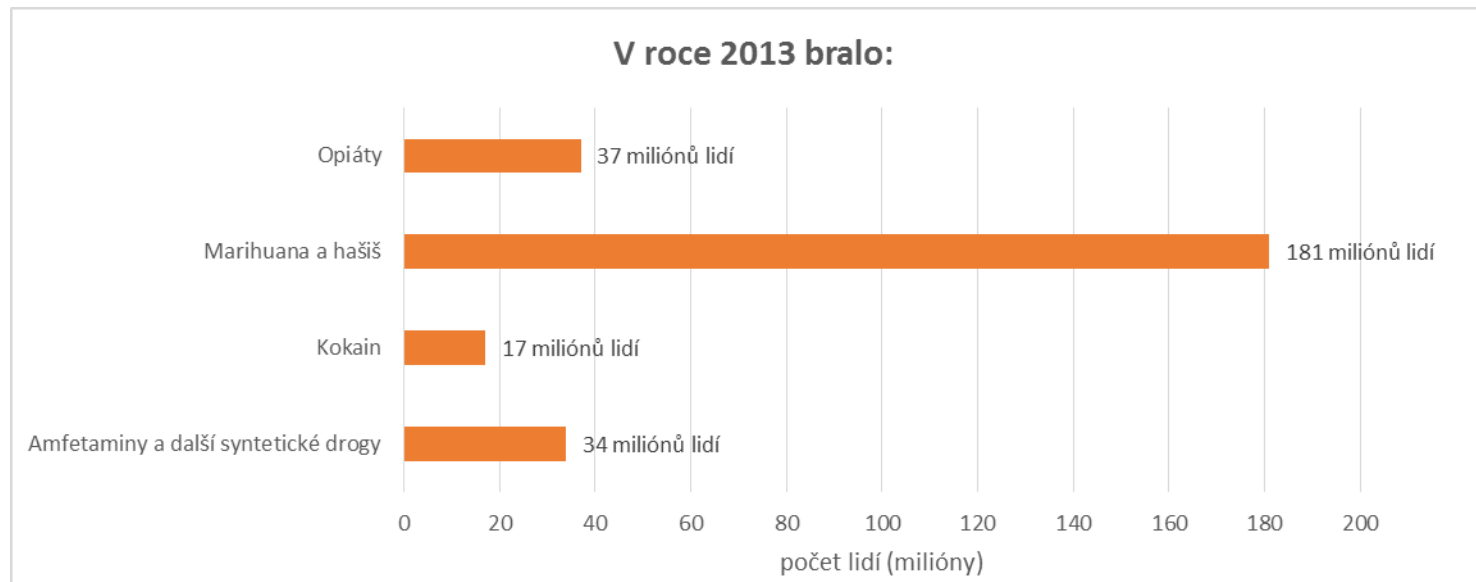
BILANCE ZÁVISLOSTI

V ROCE 2013 BRALO:



- Celkový počet lidí užívajících ilegální drogy se odhaduje za rok 2013 na 246 milionů. To je o 3 miliony více než v roce 2012, ale vzhledem k růstu celkového počtu obyvatel Země se drží podíl konzumentů drog na stabilní úrovni.
- Odhadem asi 27 milionů z těchto uživatelů drog má vážné zdravotní problémy nebo je na drogách závislá. Asi 12 milionů z těchto „problémových“ uživatelů si píchá drogy do žíly. Celkem 1,65 milionu narkomanů, kteří berou drogy injekčně, je HIV-pozitivních.
- Jen každý šestý problémový narkoman má přístup k lékařské péči.
- Ročně umírá na užívání drog 183 000 lidí.

Zdroj: Dotyk, týdeník, 34. číslo, 21. 8. 2015, ISSN: 1805-9465





Bohuslav Sobotka

@SlavekSobotka



Follow

Naše práce na lepších vztazích ČR-Čína má výsledky. Český export do Číny se zvýšil mezi 2013 a 2015 o 8 mld z 37,5 miliardy na 45,5 miliardy

[View translation](#)

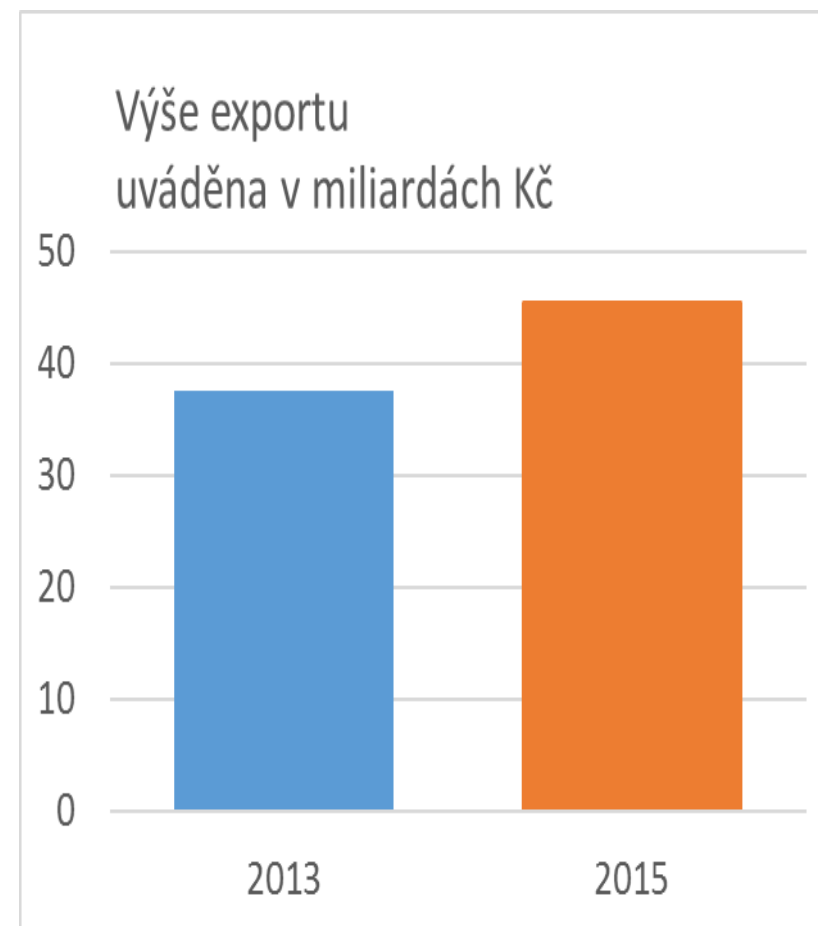


Bohuslav Sobotka
@SlavekSobotka

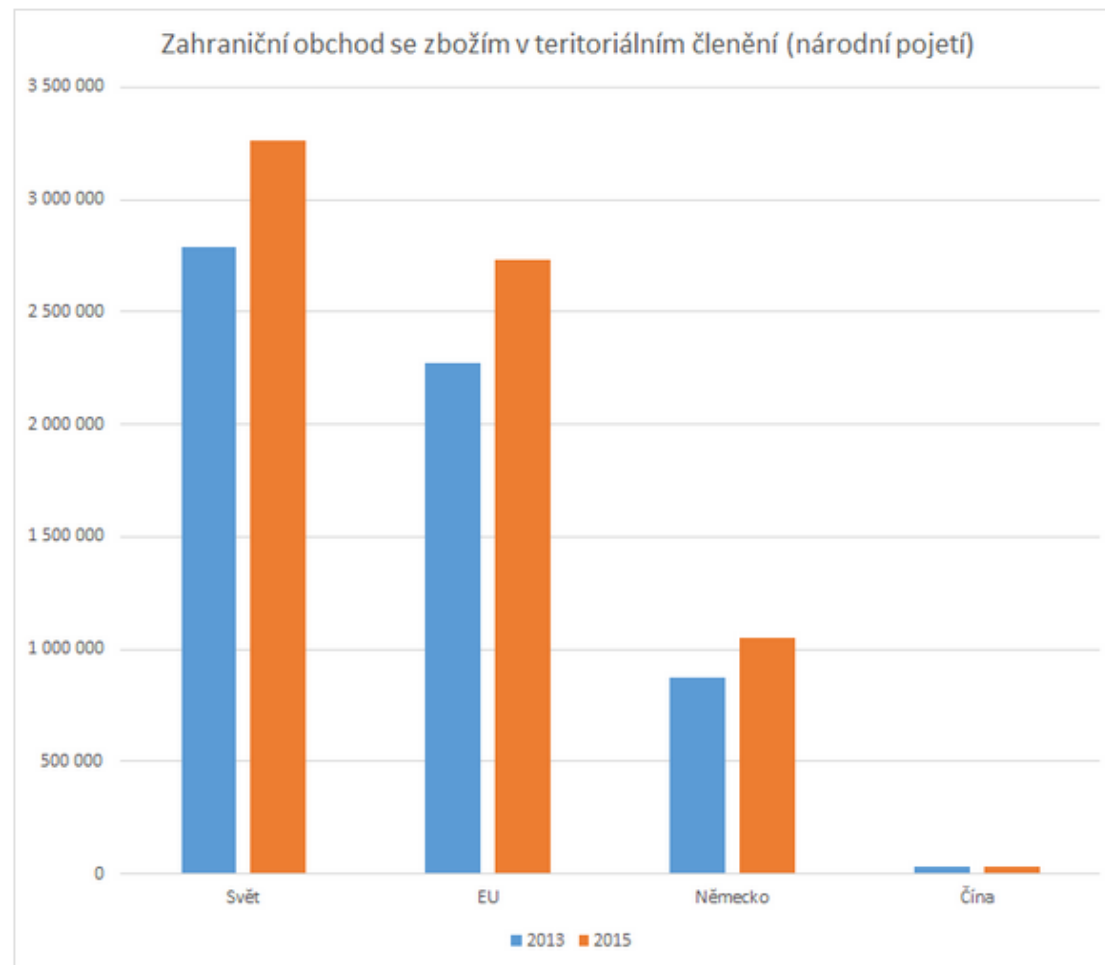


Follow

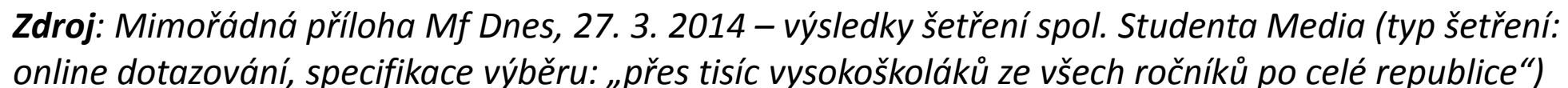
Naše práce na lepších vztazích ČR-Čína má výsledky. Český export do Číny se zvýšil mezi 2013 a 2015 o 8 mld z 37,5 miliardy na 45,5 miliardy



Český export



Zdroj: <https://www.souki.cz/kouzelne-grafy>



Průzkum o představách studentů o budoucím zaměstnání



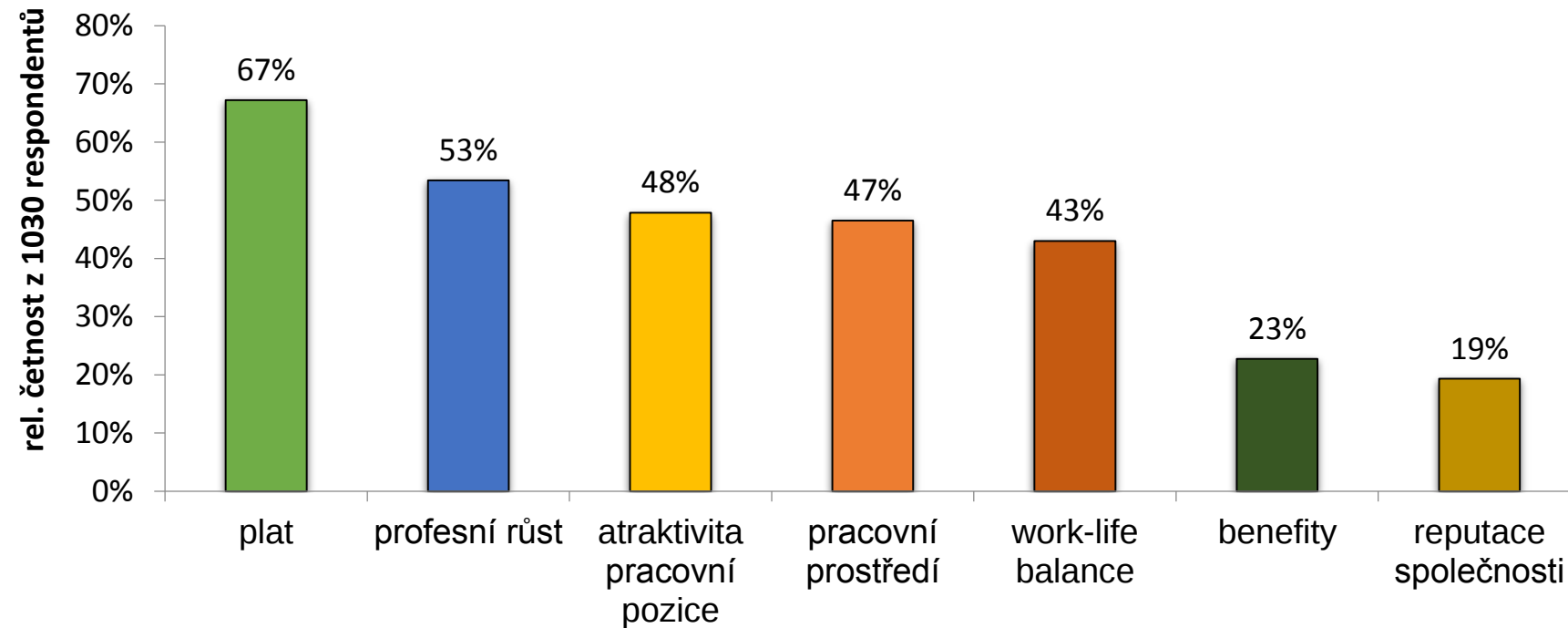
Zdroj: Mimořádná příloha Mf Dnes, 27. 3. 2014 – výsledky šetření spol. Studenta Media (typ šetření: online dotazování, specifikace výběru: „přes tisíc vysokoškoláků ze všech ročníků po celé republice“)

Jak výsledky šetření zobrazit správně?

Co je pro Vás důležité při výběru zaměstnání? (vyberte 3 pro Vás nejdůležitější faktory)			
	četnost	rel. četnost (%)	rel. četnost (%) vzhledem k počtu respondentů
plat	692	22	67
profesní růst	550	18	53
atraktivita pracovní pozice	493	16	48
pracovní prostředí	479	16	47
work-life balance	443	14	43
benefity	234	8	23
reputace společnosti	199	6	19
celkem	3090	100 %	---

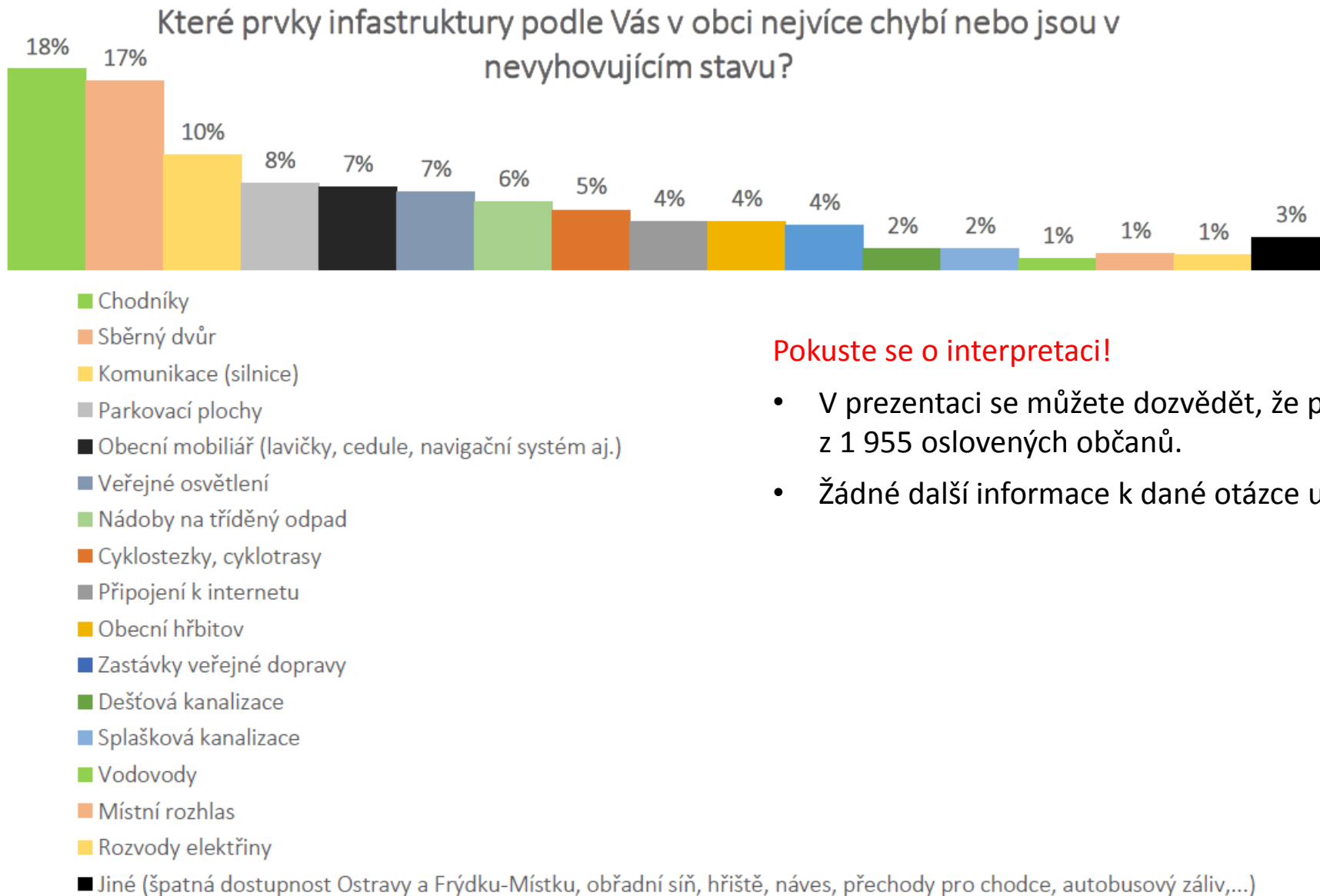
Jak výsledky šetření zobrazit správně?

Co je pro Vás důležité při výběru zaměstnání?



Hodnocení dostupnosti infrastruktury obce

Martina Litschmannová, 2018



Pokuste se o interpretaci!

- V prezentaci se můžete dozvědět, že průzkumu se zúčastnilo 219 (11,20 %) z 1 955 oslovených občanů.
- Žádné další informace k dané otázce uvedeny nejsou.

Informace z vyžádané podrobnější zprávy (na webu není zveřejněna)

Hodnocení infrastruktury obce a dopravy

Nevyhovující či chybějící prvky infrastruktury v obci dle občanů jsou uvedeny v grafu 10. Celkový počet odpovědí na tuto otázku byl 686. Celkem 120 občanů (18 % z celkového počtu odpovědí) je nespokojeno s chodníky v obci, dále 113 občanům (17 % z celkového počtu odpovědí) chybí sběrný dvůr a 71 občanů (10 % z celkového počtu odpovědí) je nespokojeno s komunikacemi (silnicemi) v obci, dále 53 občanů (8 % z celkového počtu odpovědí) s parkovacími plochami, 51 občanů (7 % z celkového počtu odpovědí) s obecním mobiliářem jako jsou lavičky, cedule, navigační systém aj. a 48 občanů (7 % z celkového počtu odpovědí) je nespokojeno s veřejným osvětlením. Všechny odpovědi jsou uvedeny v grafu 10.

Hodnocení dostupnosti infrastruktury obce

Martina Litschmannová, 2018
**Analýza je jedním z podkladů
pro tvorbu
Strategického plánu obce!**

Které prvky infrastruktury podle Vás v obci nejvíce chybí nebo jsou v nevyhovujícím stavu?



- Chodníky
- Sběrný dvůr
- Komunikace (silnice)
- Parkovací plochy
- Obecní mobiliář (lavičky, cedule, navigační systém aj.)
- Veřejné osvětlení
- Nádoby na tříděný odpad
- Cyklostezky, cyklotrasy
- Připojení k internetu
- Obecní hřbitov
- Zastávky veřejné dopravy
- Dešťová kanalizace
- Splašková kanalizace
- Vodovody
- Místní rozhlas
- Rozvody elektřiny
- Jiné (špatná dostupnost Ostravy a Frýdku-Místku, obřadní síň, hřiště, náves, přechody pro chodce, autobusový záliv,...)

Pokuste se o interpretaci!

- V prezentaci se můžete dozvědět, že průzkumu se zúčastnilo 219 (11,20 %) z 1 955 oslovených občanů.
- Žádné další informace k dané otázce uvedeny nejsou.

Informace z vyžádané zprávy:

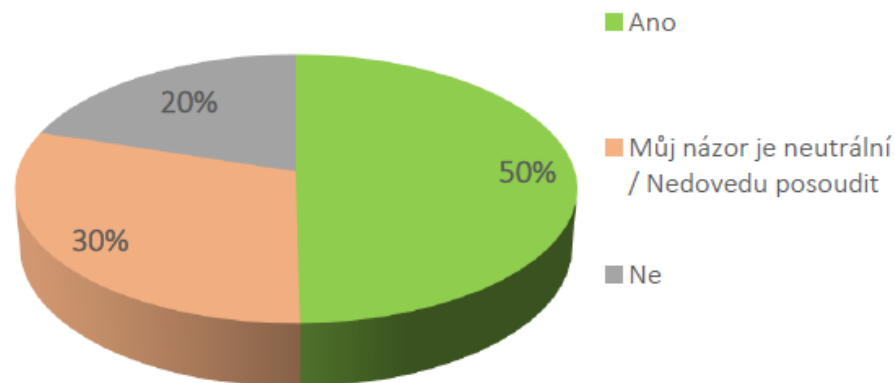
- „Celkový počet odpovědí na tuto otázku byl 686. Celkem 120 občanů (18 % z celkového počtu odpovědí) je nespokojeno s chodníky...”
- Titulek k grafu: „Graf 10 Podíl výskytu odpovědi na vyjádření občanů, které prvky infrastruktury v obci nejvíce chybí nebo jsou v nevyhovujícím stavu dle podílu z celkového počtu odpovědí“

Hodnocení dostupnosti služeb

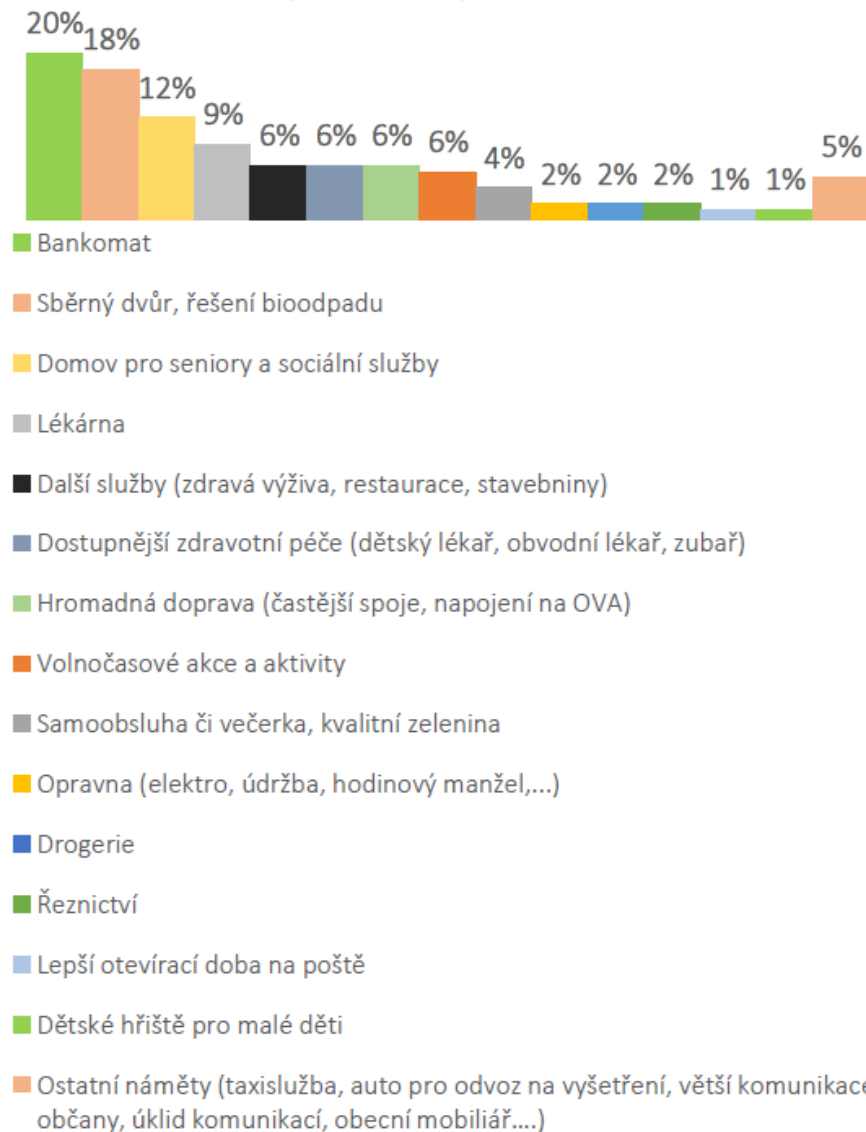
Zdroj:

https://www.krmelin.cz/evt_file.php?file=1673&original=Dotaznikove_setreni_obce_Krmelin_vysledky_10_18.pdf (18. 11. 2018)

Chybí Vám v obci nějaké služby?



Služby, které chybí občanům:



Co je pro respondenty „palčivějším“ problémem – stav chodníků nebo chybějící bankomat?

- V prezentaci se můžete dozvědět, že průzkumu se zúčastnilo 219 (11,20 %) z 1 955 oslovených občanů.
- Žádné další informace k dané otázce uvedeny nejsou.

Informace z vyžádané zprávy:

- „Občané uvedli celkem 160 námětů na chybějící služby v obci. Občanům nejčastěji chybí v obci bankomat (20 % z celkového počtu odpovědí, 32 odpovědí) ...“
- Titulek k grafu: „Graf 6 Vyjádření občanů k otázce, jaké služby jim v obci chybí dle podílů z celkového počtu odpovědí“

**Popisná statistika
aneb
Jak efektivně popsat a vizualizovat data**

Část 2
Kvantitativní znak

Popisná statistika - Kvantitativní znak

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Popište a vizualizujte hmotnost respondentů.

Popisná statistika - Kvantitativní znak

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídnání dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Míry polohy: průměr, kvantily

Míry variability: rozptyl, směrodatná odchylka, variační koeficient

Popisná statistika - Kvantitativní znak

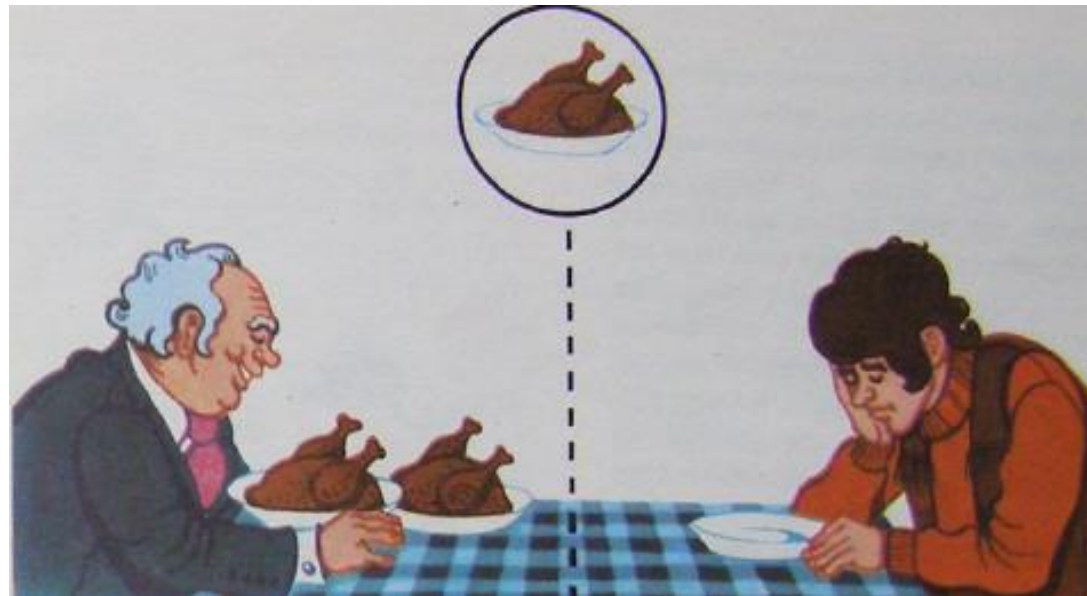
- Míry polohy

- **(Aritmetický) průměr:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

Popisná statistika – kvantitativní znak

Míry polohy

- **Aritmetický průměr:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$



Zdroj: Swoboda Helmut, Moderní statistika, 1977

Popisná statistika – kvantitativní znak

Míry polohy

- Aritmetický průměr: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$



Průměrná produkce kuřat (na osobu):
1,0 (denně)

Popisná statistika – kvantitativní znak

Míry polohy

- Aritmetický průměr: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

Průměrná mzda přivádí Čechy k zuřivosti: Kdo z vás má 27 tisíc?

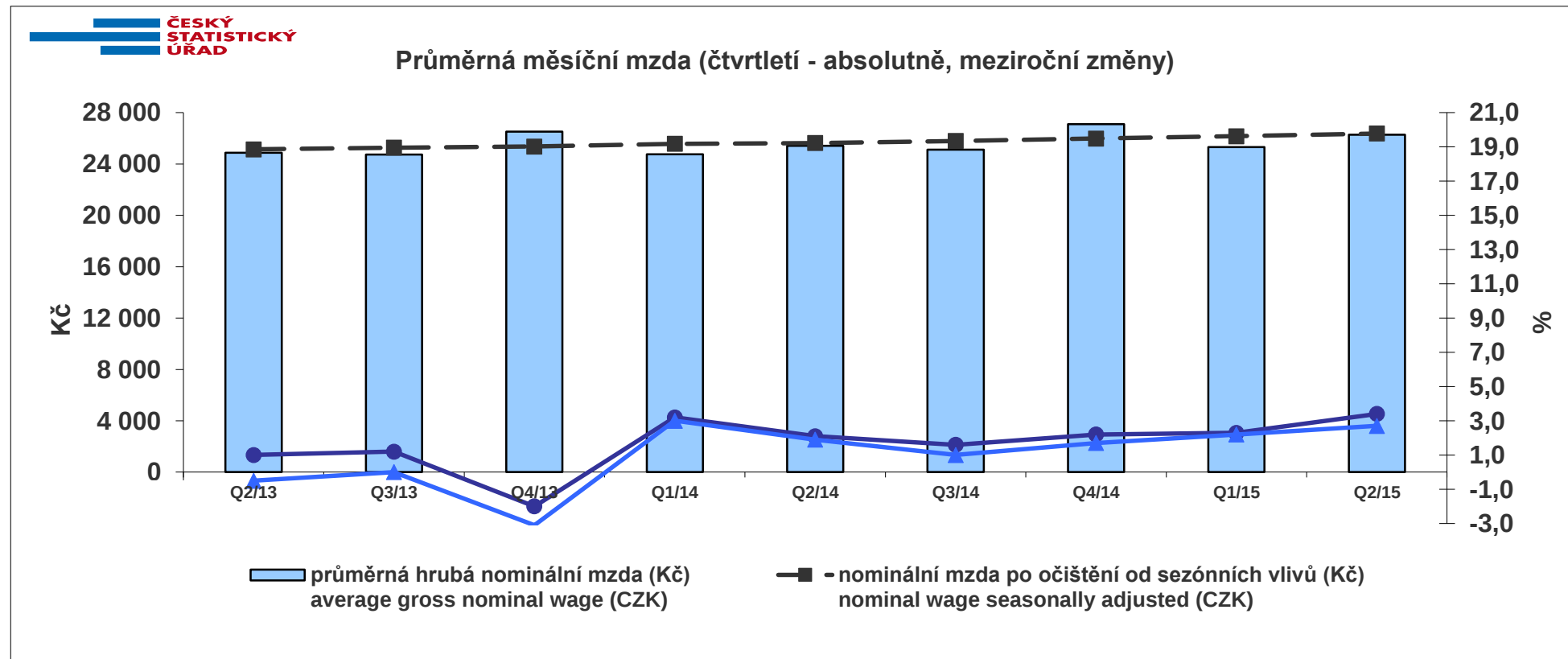


Zdroj: Blesk, 9.4.2013

Popisná statistika – kvantitativní znak

Míry polohy

- Aritmetický průměr: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$



Popisná statistika – kvantitativní znak

Míry polohy

- **Aritmetický průměr:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

POZOR!

- Průměr je číslo, které nemusí patřit do definičního oboru analyzovaného znaku. (např. průměrný počet dětí jedné ženy)

Popisná statistika – kvantitativní znak

Míry polohy

- **Aritmetický průměr:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

V malé vesnici někde v Americe žije 6 lidí, jejichž roční plat je uveden níže.

\$25 000 \$27 000 \$29 000

\$35 000 \$37 000 \$38 000

Určete průměrný plat obyvatel této vesnice.

(\$31 830)

Do vesnice se přistěhoval Bill Gates, jehož roční příjem je \$40 000 000.

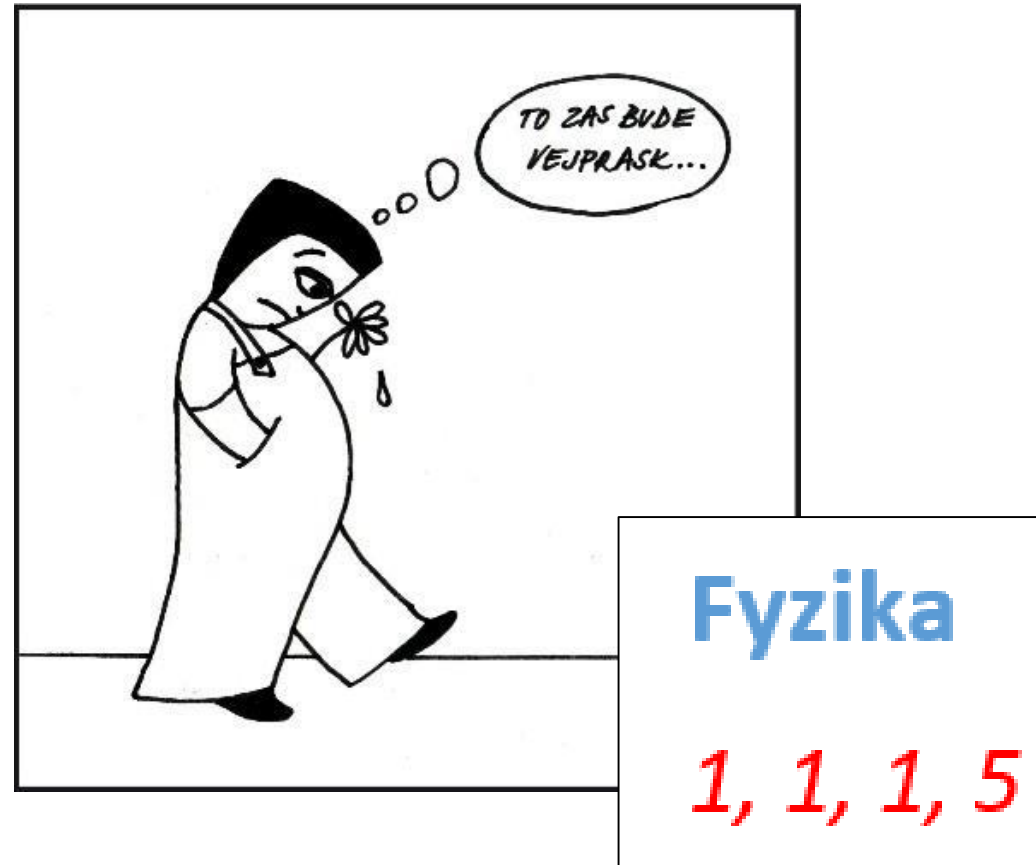
\$25 000 \$27 000 \$29 000

\$35 000 \$37 000 \$38 000 \$40 000 000

Určete průměrný plat obyvatel této vesnice.

(\$5 741 571)

Popisná statistika – kvantitativní znak



Popisná statistika - Kvantitativní znak

- **Míry polohy**

- **(Aritmetický) průměr:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

POZOR!

- Pozor na interpretaci průměru! Nepřisuzujme mu vlastnosti, které nemá!
- Průměr je číslo, které nemusí patřit do definičního oboru analyzovaného znaku. (např. průměrný počet dětí jedné ženy)
- Průměr není rezistentní vůči odlehlým pozorováním!



Popisná statistika – kvantitativní znak

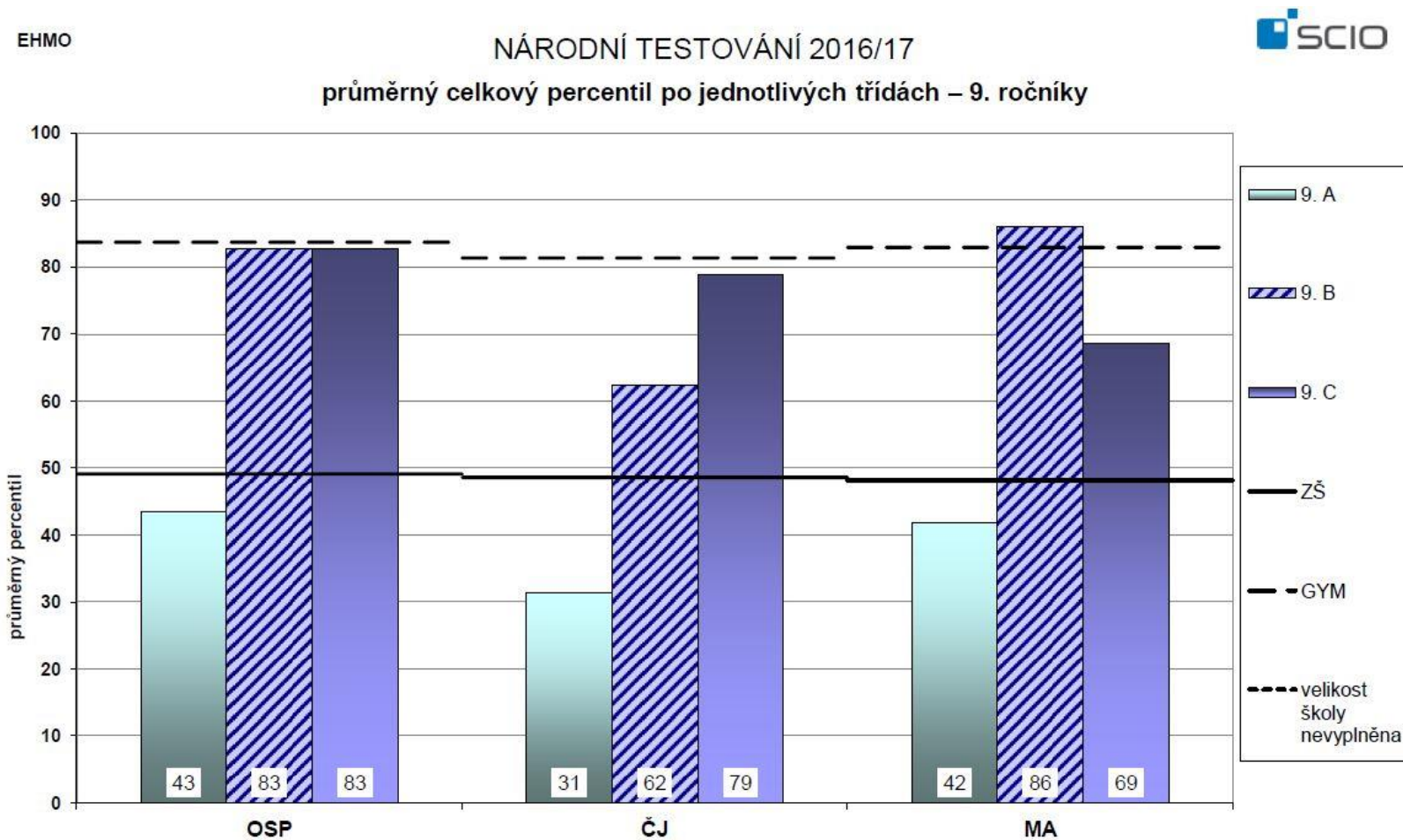
Míry polohy

- **Aritmetický průměr:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
- **Medián *Med*** (50% kvantil – 50 % hodnot je menších nebo rovných mediánů)
- **Dolní kvartil *Q1*** (25% kvantil – 25 % hodnot je menších nebo rovných dolnímu kvartilu)
- **Horní kvartil *Q3*** (75% kvantil – 75 % hodnot je menších nebo rovných hornímu kvartilu)
- **100p% kvantil** – 100p % hodnot je menších nebo rovných 100p% kvantilu

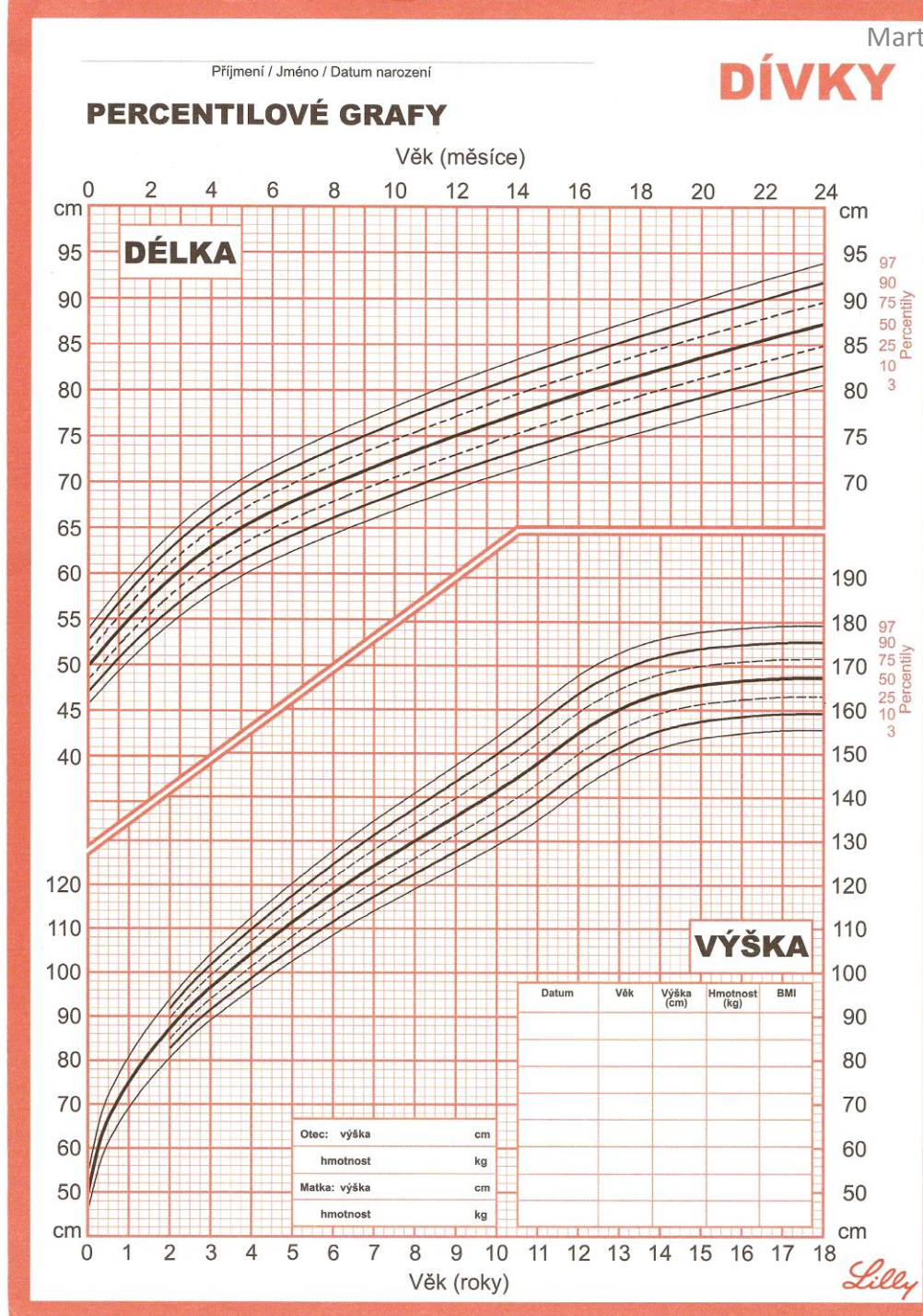
Speciální typy kvantilů:

- Kvartily
- Decily
- Percentily

Kvantily v praxi

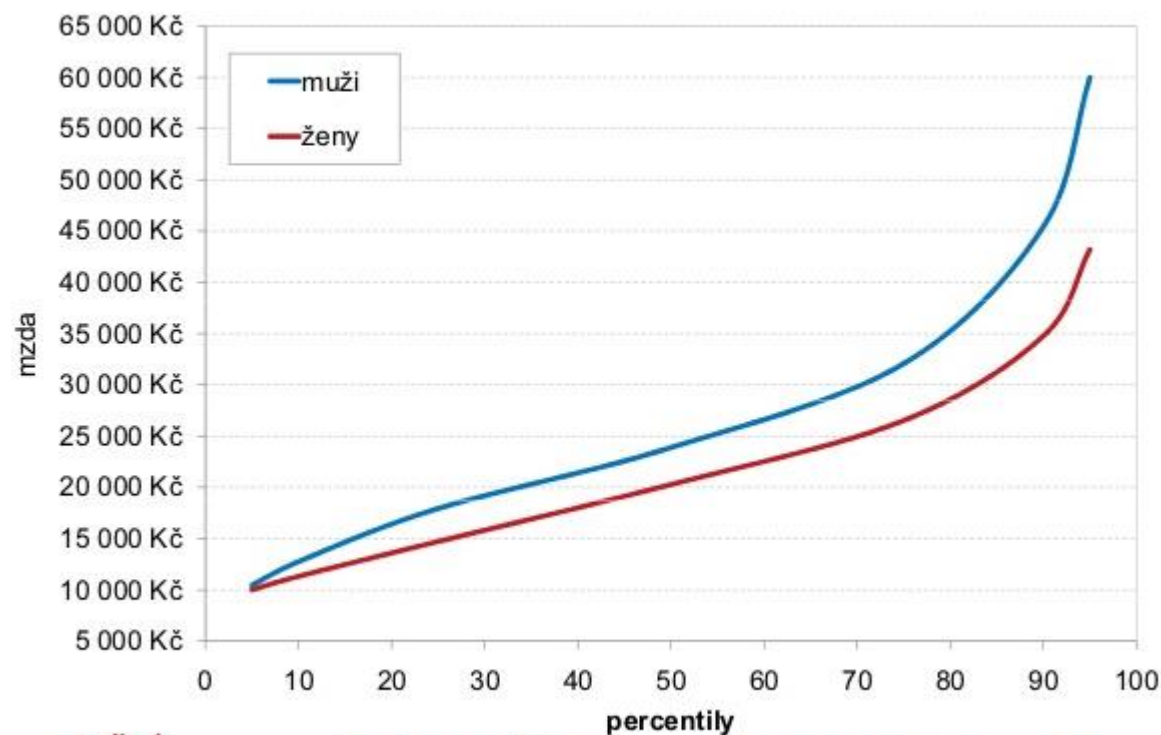


Graf znázorňuje průměrné celkové percentily všech tříd 9. ročníku vaší školy. Zároveň je zde pro porovnání vyznačen průměrný percentil všech žáků 9. ročníku ZŠ a odpovídajícího ročníku víceletých gymnázií.

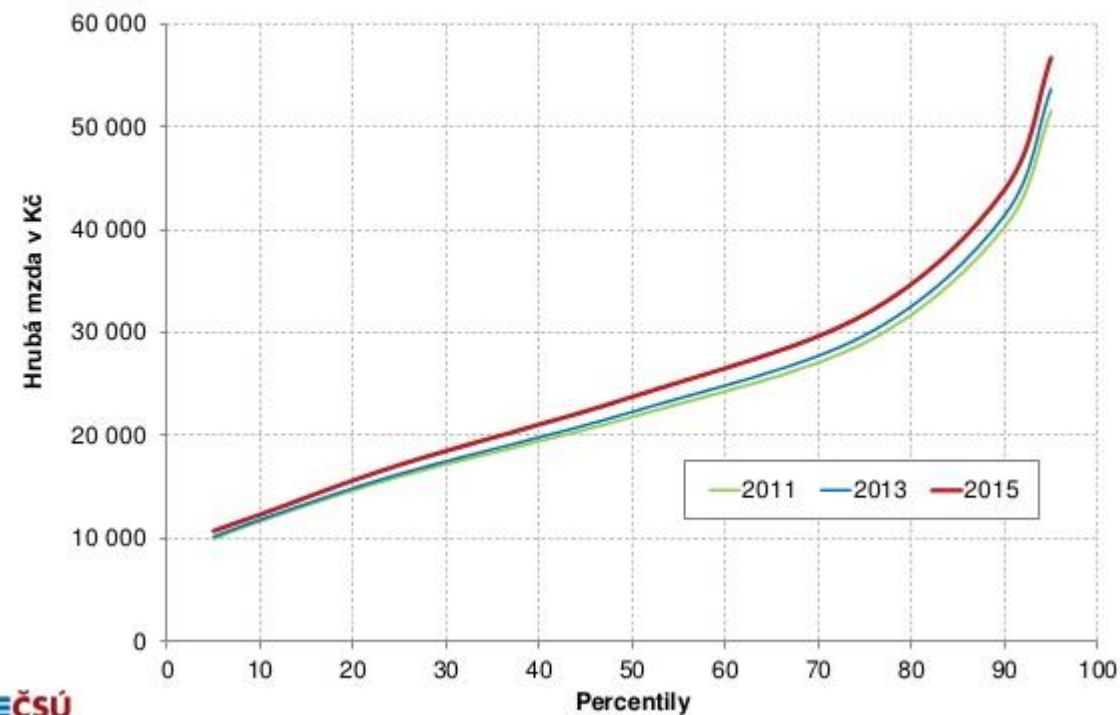


Kvantily v praxi

MZDY MUŽŮ A ŽEN (2012)

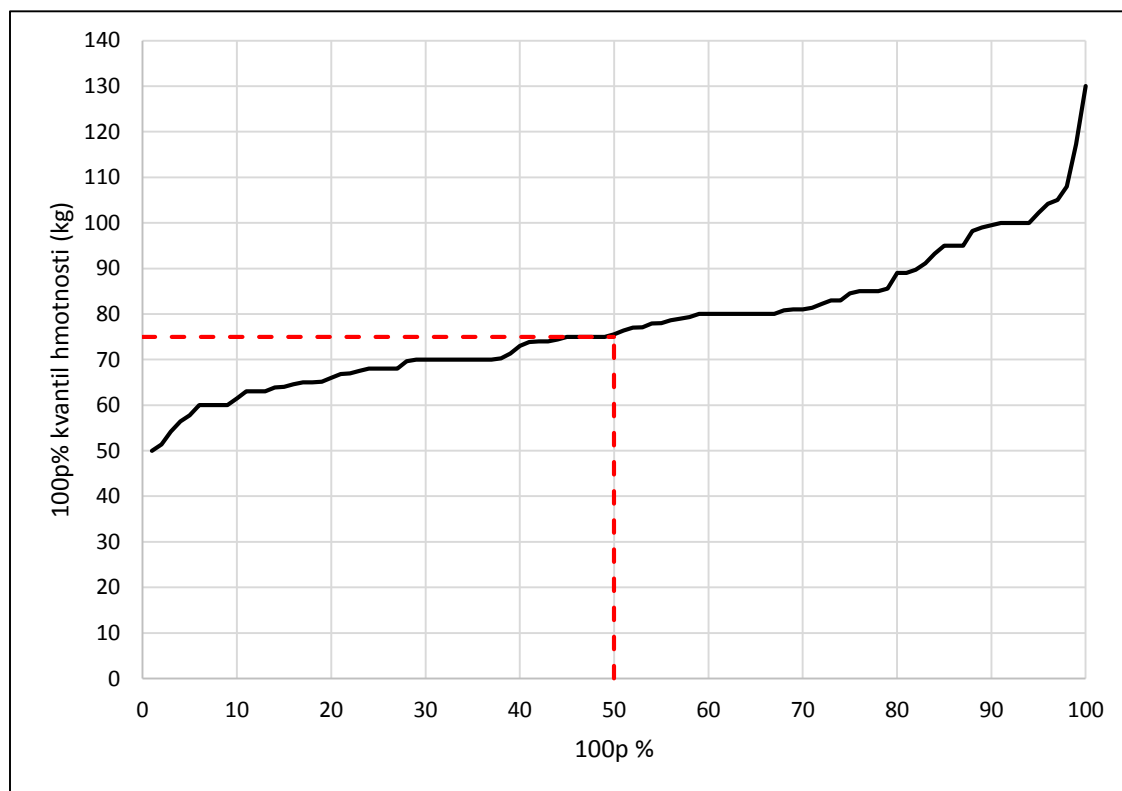


Proměny distribuce mezd I.



Popisná statistika – kvantitativní znak

Vizualizace

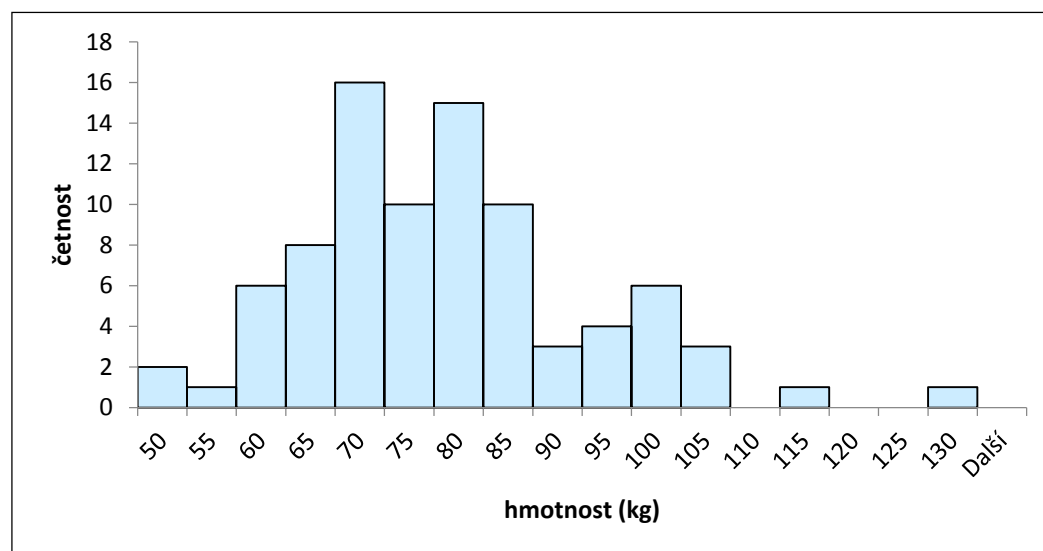


Kvantilová funkce

Míry polohy	Váha (kg)
minimum	50
dolní kvartil	68
průměr	78
medián	76
horní kvartil	85
maximum	130

Popisná statistika - Kvantitativní znak

- Vizualizace

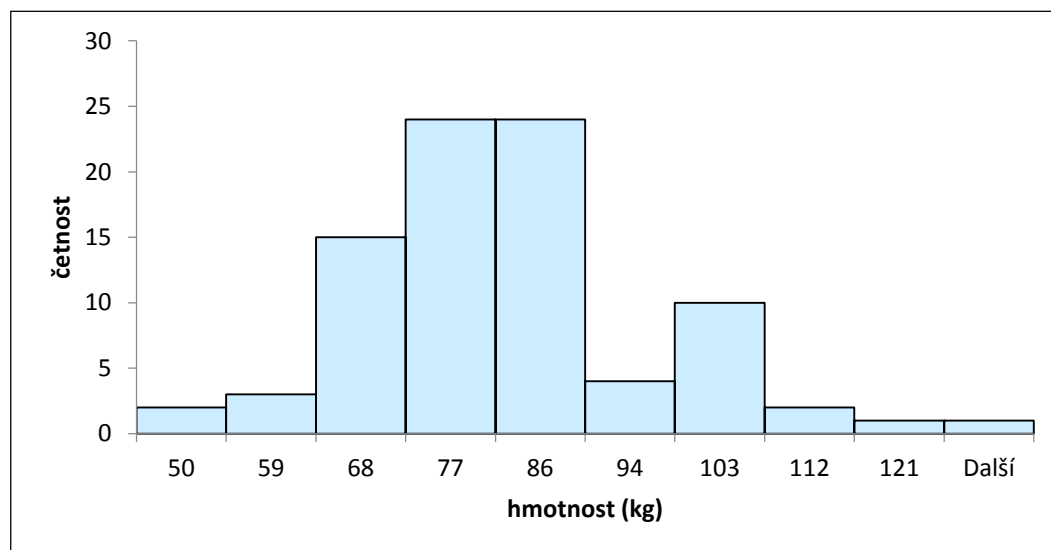


Histogram

Míry polohy	Hmotnost (kg)
Minimum	50
Dolní kvartil	68
Průměr	78
Medián	76
Horní kvartil	85
Maximum	130

Popisná statistika - Kvantitativní znak

- Vizualizace



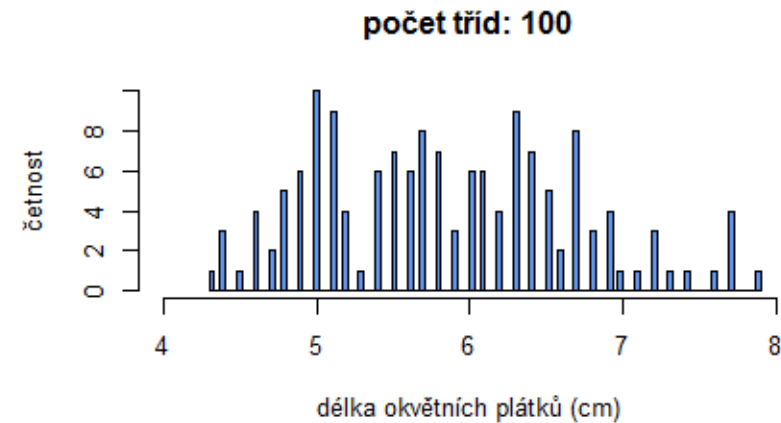
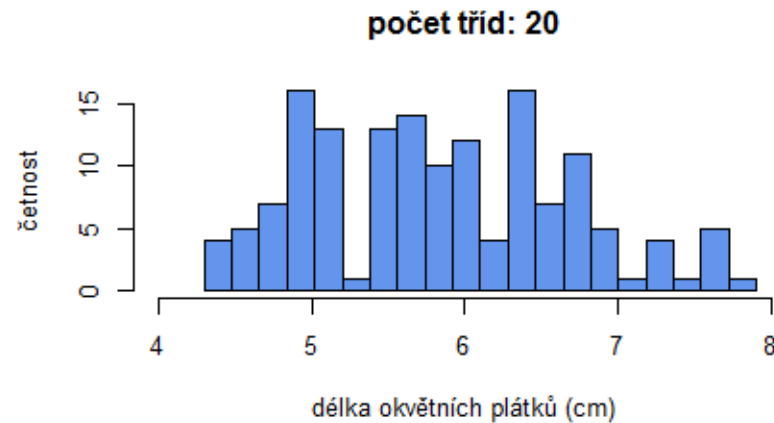
Histogram

Míry polohy	Hmotnost (kg)
Minimum	50
Dolní kvartil	68
Průměr	78
Medián	76
Horní kvartil	85
Maximum	130

Tvar histogramů závisí na počtu tříd (sloupečků)!



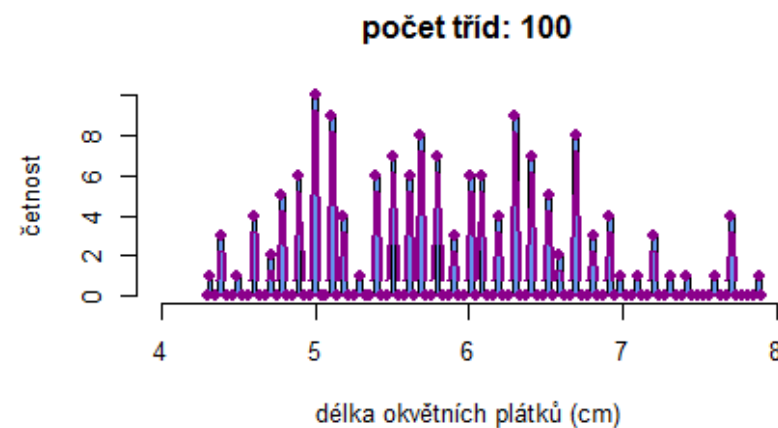
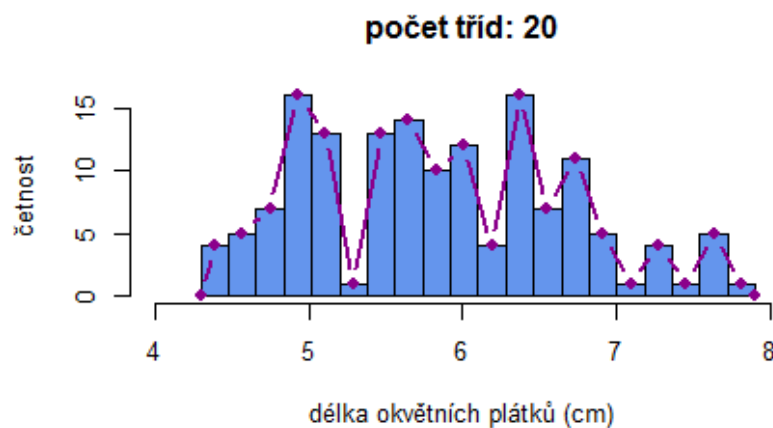
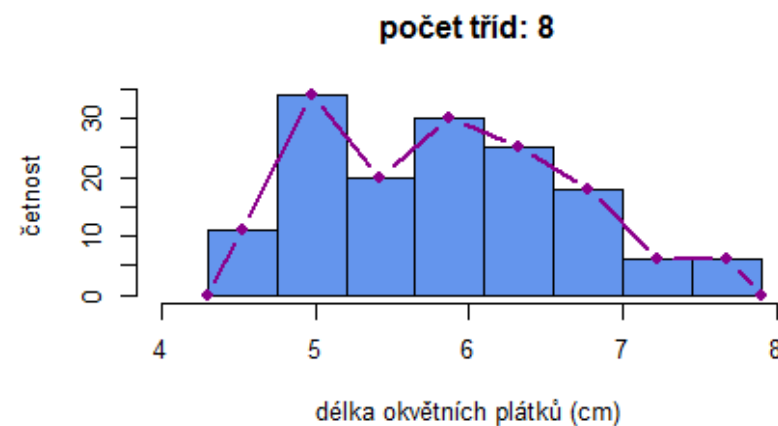
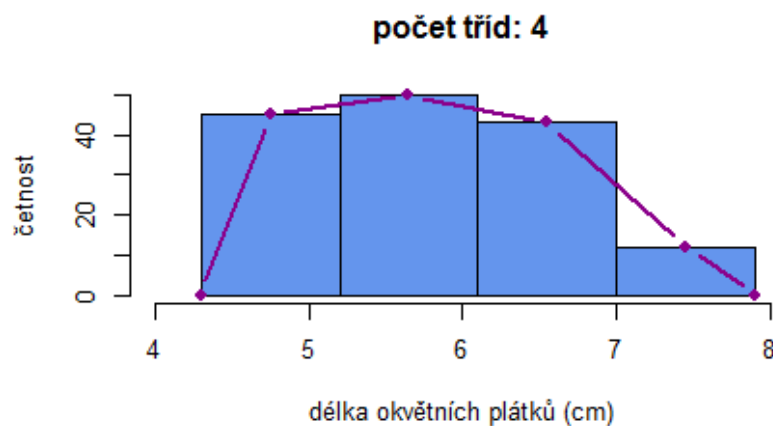
Popisná statistika - Kvantitativní znak



Tvar histogramů závisí na počtu tříd (sloupečků)!



Popisná statistika - Kvantitativní znak

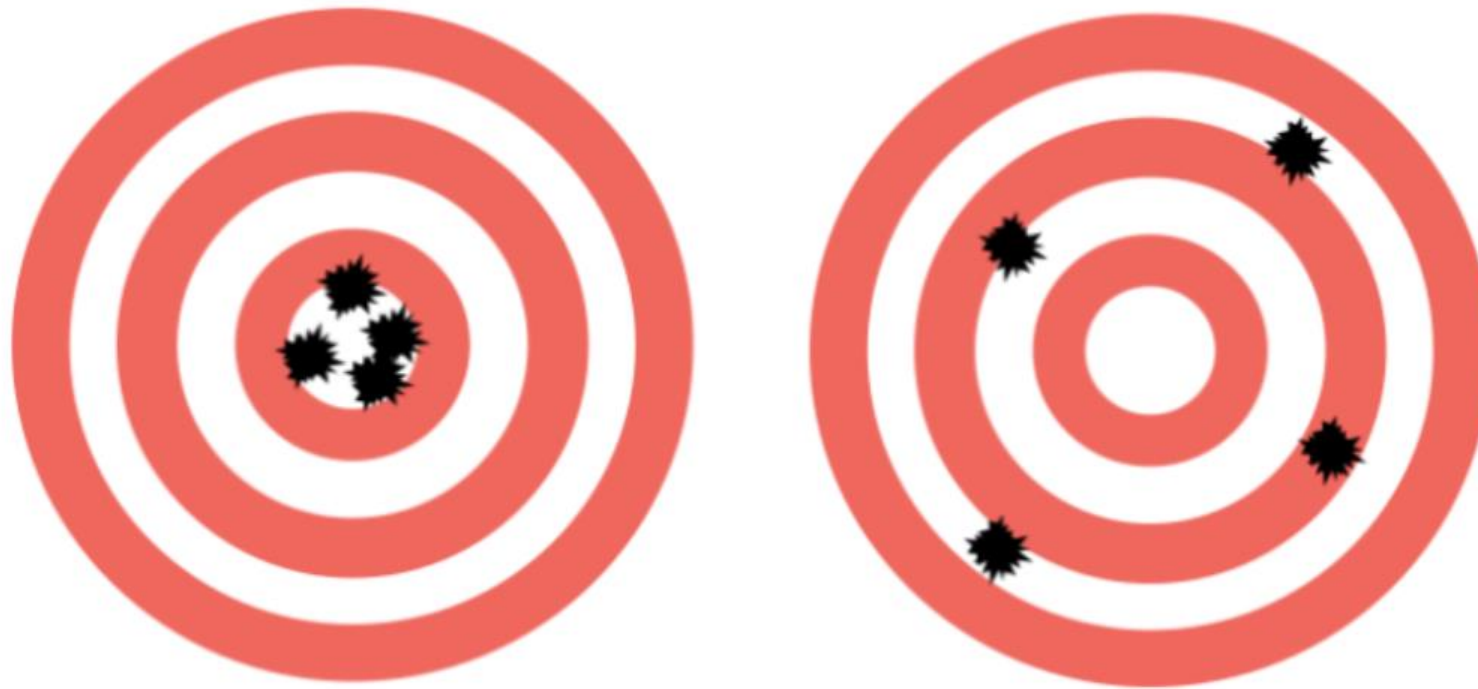


Histogram doplněný o polygon četnosti

Popisná statistika – Kvantitativní znak

Míry variability

K čemu nám jsou dobré?



Zdroj: <http://blackswanfarming.com/>

Popisná statistika – kvantitativní znak

Míry variability

- **Výběrový rozptyl:** $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$
 - POZOR! – Jednotka rozptylu je kvadrátem jednotky analyzovaného znaku.
- **Výběrová směrodatná odchylka:** $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$
 - Neumožňuje srovnání variability znaků s různými jednotkami.
- **Variační koeficient:** $V = \frac{s}{|\bar{x}|} \cdot 100 (\%)$
 - Čím nižší var. koeficient, tím homogennější soubor.
 - $V > 50 \%$ značí silně rozptýlený soubor. (empirické doporučení)
- **Rozpětí:** $R = \max - \min$
- **Interkvartilové rozpětí:** $IQR = \tilde{x}_{0,75} - \tilde{x}_{0,25}$

Proč se pro směrodatnou odchylku někdy používá symbol s a jindy symbol σ ?

Míry variability

- **Výběrová směrodatná odchylka:** $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$
- **Populační směrodatná odchylka:** $\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}}$
 - N ... rozsah populace, tj. pro výpočet musíme mít k dispozici všechna data ze základního souboru (populace), tj. musíme provést úplné šetření.
 - Lze ukázat, že nejlepším odhadem populační směrodatné odchylky je výběrová směrodatná odchylka:

$$\hat{\sigma} \cong s$$

Jakou představu o variabilitě dat nám dává směrodatná odchylka?

Obecně platí tzv. Chebysheva nerovnost

Mají-li data libovolné rozdělení s konečným průměrem (μ) a konečnou sm. odchylkou (σ), pak nejméně $\left(1 - \frac{1}{k^2}\right) \cdot 100$ % variant leží v intervalu $(\mu - k\sigma; \mu + k\sigma)$.

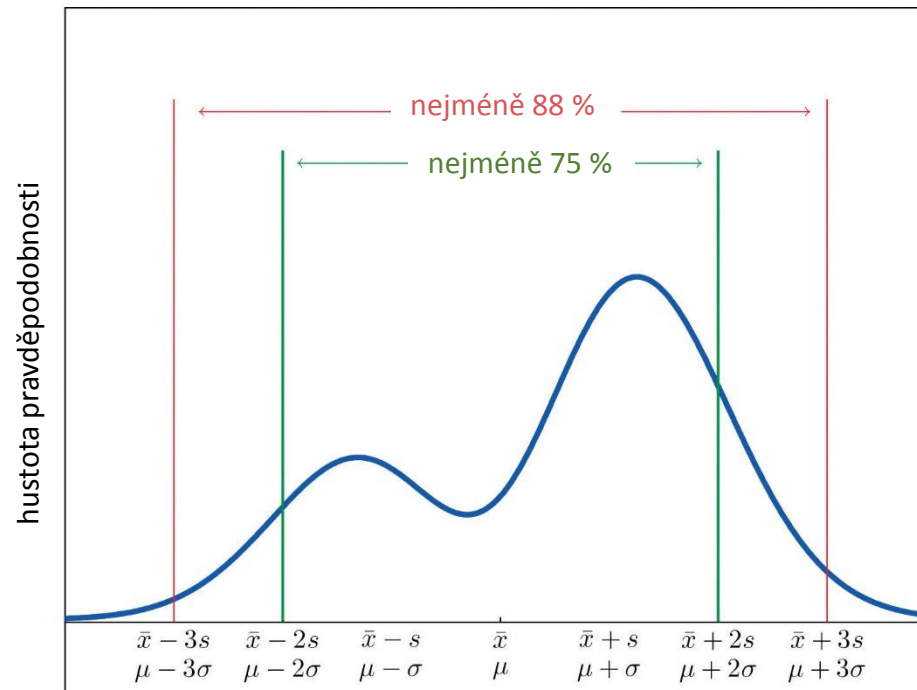
$$\forall k > 0: P(\mu - k\sigma < X < \mu + k\sigma) > 1 - \frac{1}{k^2}$$

k	$P(\mu - k\sigma < X < \mu + k\sigma)$ (tj. pravděpodobnost, že data leží v intervalu $(\mu - k\sigma; \mu + k\sigma)$)
1	> 0 %
2	> 75 %
3	> 88 %

Jakou představu o variabilitě dat nám dává směrodatná odchylka?

Obecně platí tzv. Chebyschevova nerovnost

Mají-li data libovolné rozdělení s konečným průměrem (μ) a konečnou sm. odchylkou (σ), pak nejméně $\left(1 - \frac{1}{k^2}\right) \cdot 100 \%$ variant leží v intervalu $(\mu - k\sigma; \mu + k\sigma)$.



Jakou představu o variabilitě dat nám dává směrodatná odchylka?

Pro data, která mají normální rozdělení platí:

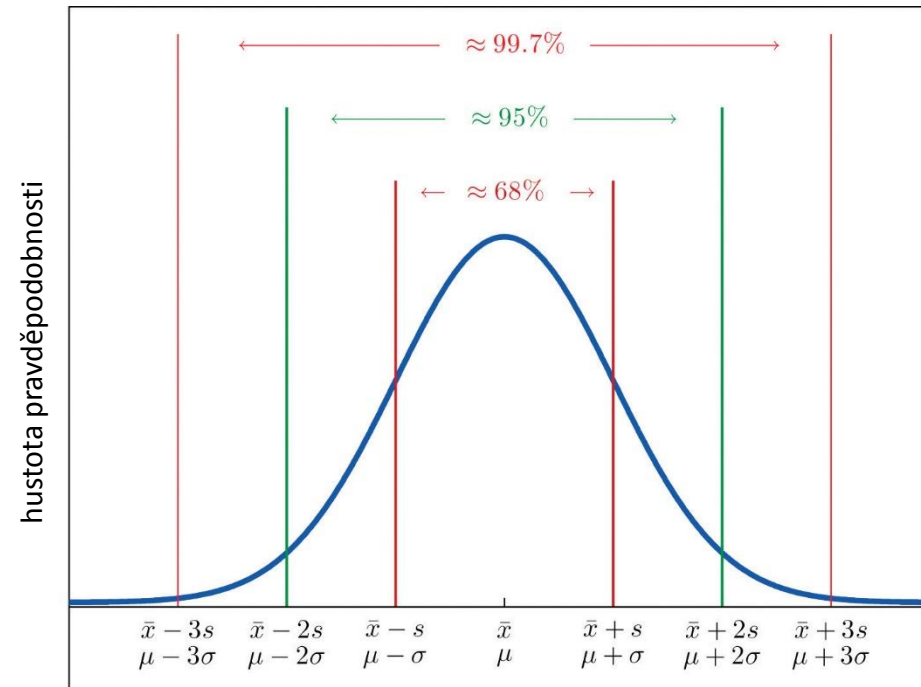
Mají-li data normální rozdělení (obálka histogramu odpovídá Gaussově křivce (zvonovitý tvar)) s konečným průměrem (μ) a konečnou sm. odchylkou (σ), pak:

k	$P(\mu - k\sigma < X < \mu + k\sigma)$ (tj. pravděpodobnost, že data leží v intervalu $(\mu - k\sigma; \mu + k\sigma)$)
1	$\cong 68 \%$
2	$\cong 95 \%$
3	$\cong 99,7 \%$

Jakou představu o variabilitě dat nám dává směrodatná odchylka?

Pro data, která mají normální rozdělení platí:

Mají-li data normální rozdělení (obálka histogramu odpovídá Gaussově křivce (zvonovitý tvar)) s konečným průměrem (μ) a konečnou sm. odchylkou (σ), pak:



Popisná statistika - Kvantitativní znak

Míry variability	Hmotnost (kg)
Rozptyl	215,3893
Směrodatná odchylka	14,67615
Variační koeficient (%)	18,90576

Jak zaokrouhlovat výběrové charakteristiky?

Směrodatnou odchylku zaokrouhlujeme nahoru na 1-2 platné cifry.

Míry polohy zaokrouhlujeme následně na stejný řád.

Popisná statistika - Kvantitativní znak

Jak zaokrouhlovat výběrové charakteristiky?

Směrodatnou odchylku zaokrouhlujeme nahoru na 1-2 platné cifry.

Míry polohy zaokrouhlujeme následně na stejný řád.

Míry polohy	Váha (kg)	zaokrouhleno
minimum	50	
dolní kvartil	68	
průměr	77,62791	
medián	75,5	
horní kvartil	84,5	
maximum	130	
Míry variability		
směrodatná odchylka	14,67615	
variační koeficient (%)	18,90576	

Popisná statistika - Kvantitativní znak

Jak zaokrouhlovat výběrové charakteristiky?

Směrodatnou odchylku zaokrouhlujeme nahoru na 1-2 platné cifry.

Míry polohy zaokrouhlujeme následně na stejný řád.

Míry polohy	Váha (kg)	zaokrouhleno
minimum	50	
dolní kvartil	68	
průměr	77,62791	
medián	75,5	
horní kvartil	84,5	
maximum	130	
Míry variability		
směrodatná odchylka	14,67615	15
variační koeficient (%)	18,90576	

zaokrouhlujeme **nahoru** na 1-2 platné cifry

Popisná statistika - Kvantitativní znak

Jak zaokrouhlovat výběrové charakteristiky?

Směrodatnou odchylku zaokrouhlujeme nahoru na 1-2 platné cifry.

Míry polohy zaokrouhlujeme následně na stejný řád.

Míry polohy	Váha (kg)	zaokrouhleno
minimum	50	
dolní kvartil	68	68
průměr	77,62791	78
medián	75,5	76
horní kvartil	84,5	85
maximum	130	
Míry variability		
směrodatná odchylka	14,67615	15
variační koeficient (%)	18,90576	

zaokrouhlujeme na stejný řád
jako směrodatnou odchylku

Popisná statistika - Kvantitativní znak

Jak zaokrouhlovat výběrové charakteristiky?

Směrodatnou odchylku zaokrouhlujeme nahoru na 1-2 platné cifry.

Míry polohy zaokrouhlujeme následně na stejný řád.

Míry polohy	Váha (kg)	zaokrouhleno
minimum	50	50
dolní kvartil	68	68
průměr	77,62791	78
medián	75,5	76
horní kvartil	84,5	85
maximum	130	130
Míry variability		
směrodatná odchylka	14,67615	15
variační koeficient (%)	18,90576	

← nezaokrouhlujeme
(údaj vybrán z datového souboru)

← nezaokrouhlujeme
(údaj vybrán z datového souboru)

Popisná statistika - Kvantitativní znak

Jak zaokrouhlovat výběrové charakteristiky?

Směrodatnou odchylku zaokrouhlujeme nahoru na 1-2 platné cifry.

Míry polohy zaokrouhlujeme následně na stejný řád.

Míry polohy	Váha (kg)	zaokrouhleno
minimum	50	50
dolní kvartil	68	68
průměr	77,62791	78
medián	75,5	76
horní kvartil	84,5	85
maximum	130	130
Míry variability		
směrodatná odchylka	14,67615	15
variační koeficient (%)	18,90576	18,9

zaokrouhlujeme na desetiny %
(potřebujeme srovnávat s 50 %)

Popisná statistika aneb Jak efektivně popsat a vizualizovat data?

Část 3
Analýza závislosti dvou kvantitativních znaků

Analýza závislosti dvou kvantitativních znaků

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Jak popsat a vizualizovat závislost mezi výškou a hmotností respondentů?

Analýza závislosti dvou kvantitativních znaků

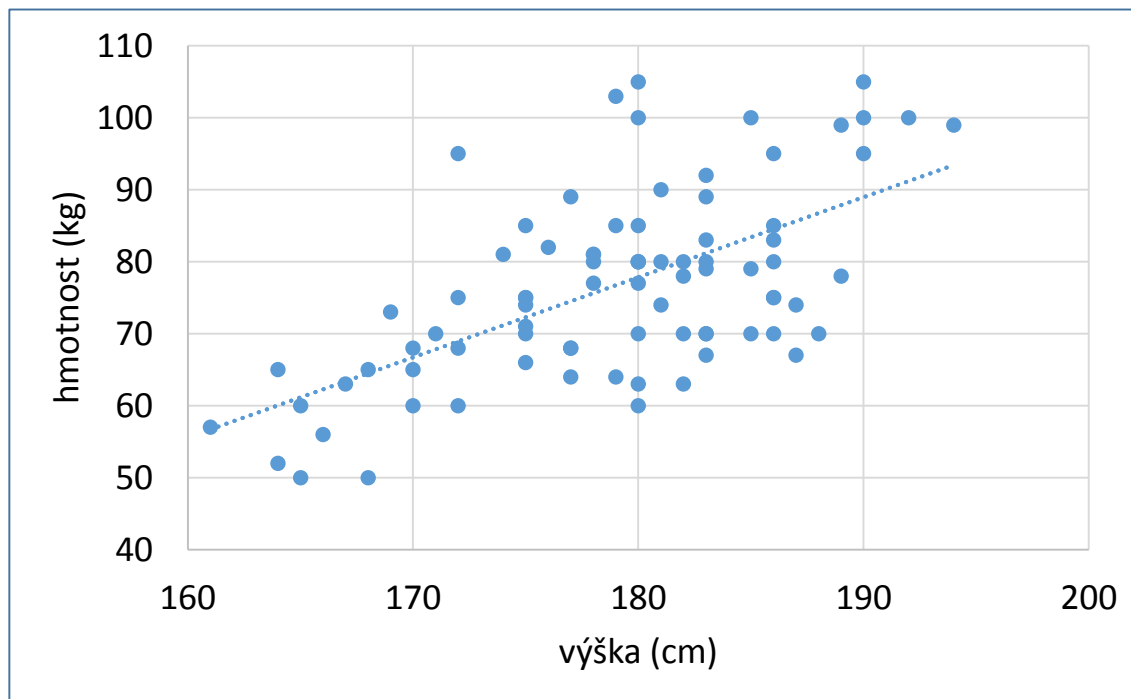
Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

- Bodový graf
- Korelační koeficient

Jak popsat a vizualizovat závislost mezi výškou a hmotností respondentů?

Analýza závislosti dvou kvantitativních znaků

Vizualizace – Bodový graf



Když chceme graficky prezentovat závislost dvou kvantitativních znaků, je pro snadnou interpretaci důležité rozhodnout, který ze znaků je vysvětlující (osa x) a který je vysvětlovaný (osa y).

Analýza závislosti dvou kvantitativních znaků

- Vizualizace – Bodový graf
- Pearsonův výběrový korelační koeficient

$$r(X, Y) = \frac{1}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y},$$

kde

n ... rozsah výběru,

x_i ... i -tá hodnota znaku x (vysvětlující proměnná),

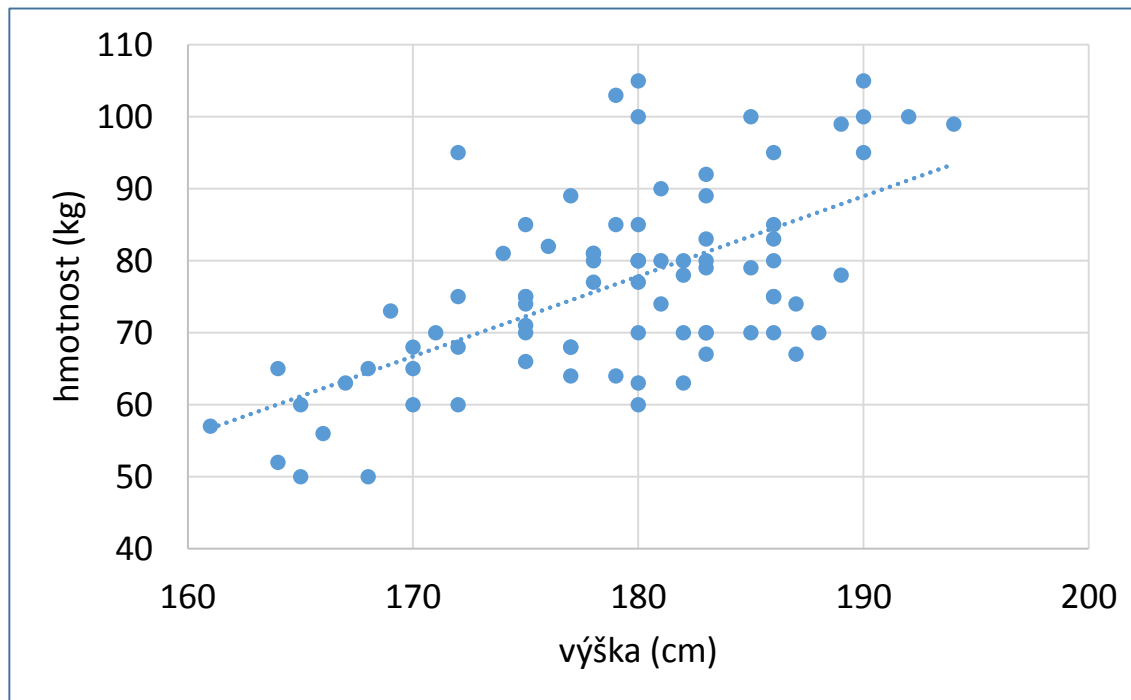
y_i ... i -tá hodnota znaku y (vysvětlovaná proměnná),

$\bar{x}$ ($\bar{y}$) ... průměr znaku x (y),

s_X (s_Y) ... výběrová směrodatná odchylka znaku x (y)

Analýza závislosti dvou kvantitativních znaků

Vizualizace – Bodový graf



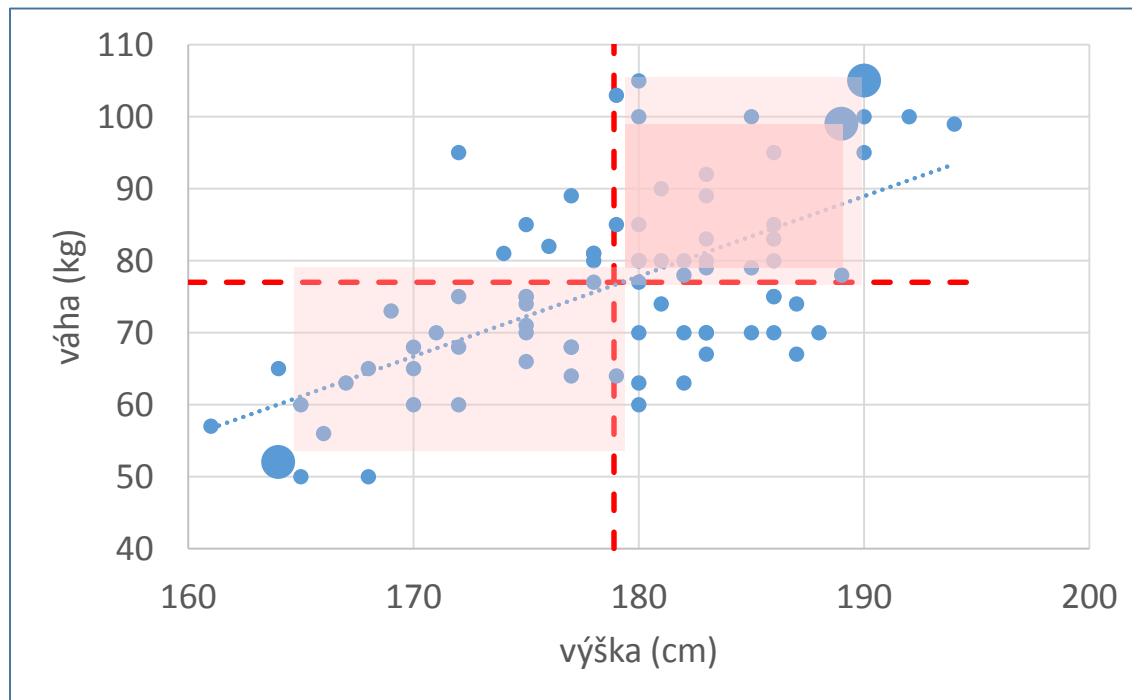
Pearsonův výběrový korelační koeficient:

$$r = \frac{1}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

Analýza závislosti dvou kvantitativních znaků

Vizualizace – bodový graf



Pearsonův výběrový korelační koeficient:

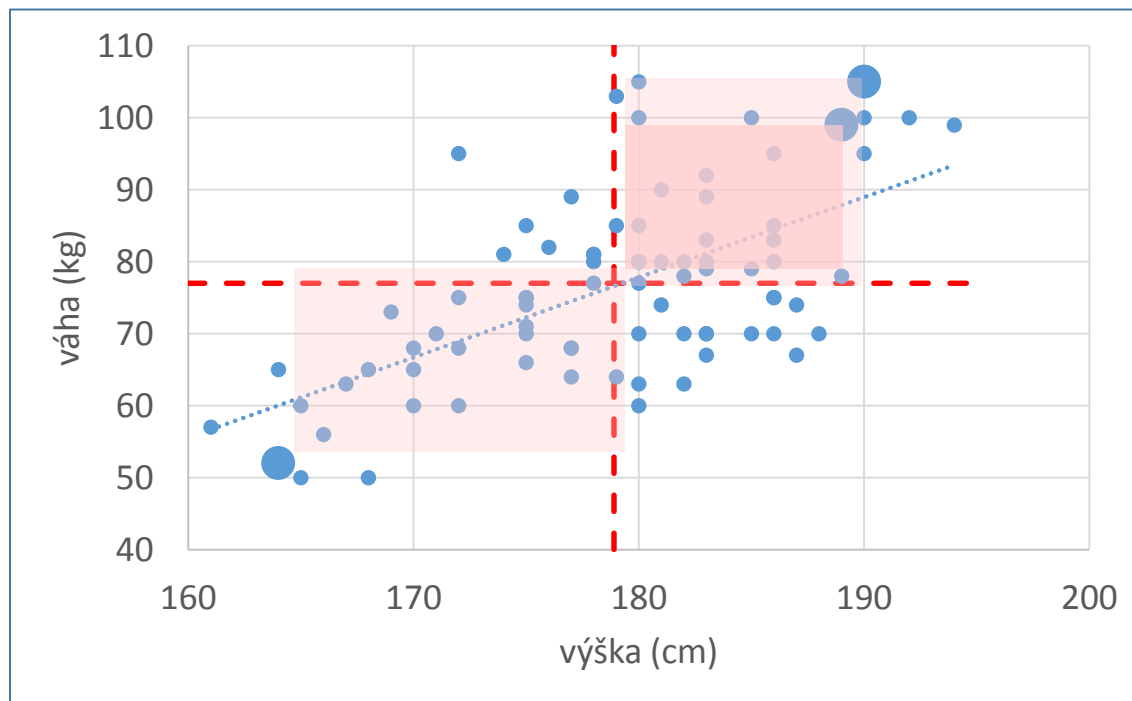
$$r = \frac{1}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

Nabývají-li pro danou statistickou jednotku oba znaky nadprůměrných hodnot, korelační koeficient se zvyšuje.

Analýza závislosti dvou kvantitativních znaků

Vizualizace – bodový graf



Pearsonův výběrový korelační koeficient:

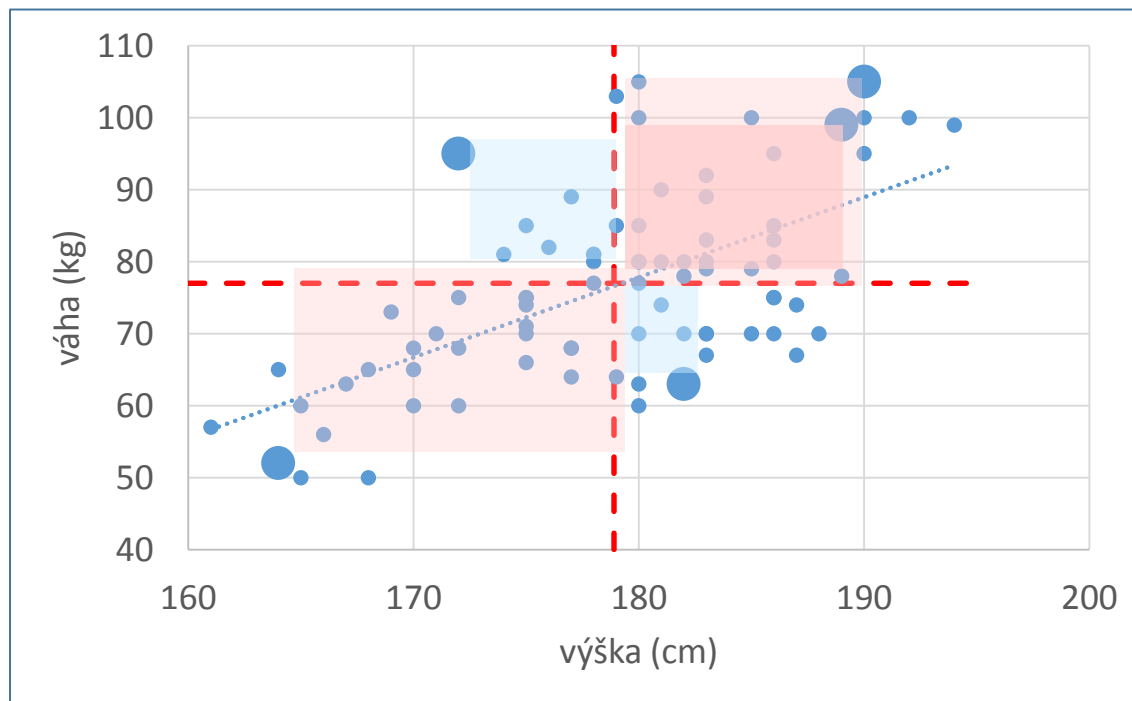
$$r = \frac{1}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

Nabývají-li pro danou statistickou jednotku oba znaky podprůměrných hodnot, korelační koeficient se zvyšuje.

Analýza závislosti dvou kvantitativních znaků

Vizualizace – bodový graf



Pearsonův výběrový korelační koeficient:

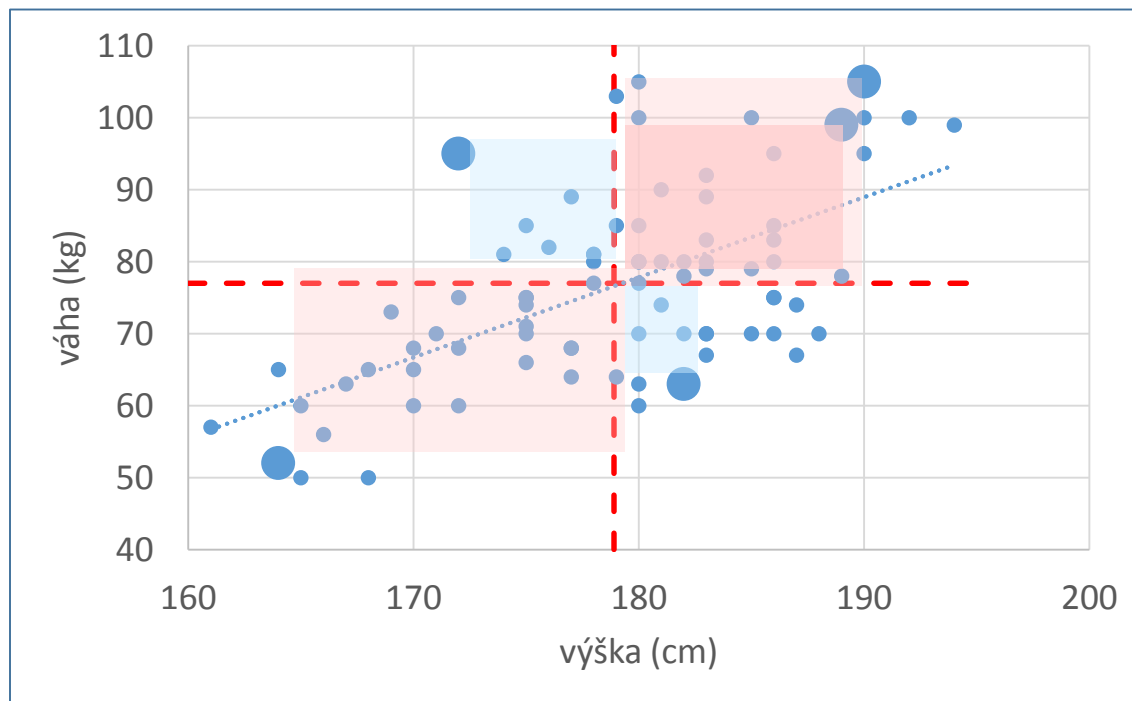
$$r(X, Y) = \frac{1}{n - 1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

Nabývá-li pro danou statistickou jednotku jeden znak nadprůměrné hodnoty a druhý znak podprůměrné hodnoty, korelační koeficient se snižuje.

Analýza závislosti dvou kvantitativních znaků

Vizualizace – bodový graf



Pearsonův výběrový korelační koeficient:

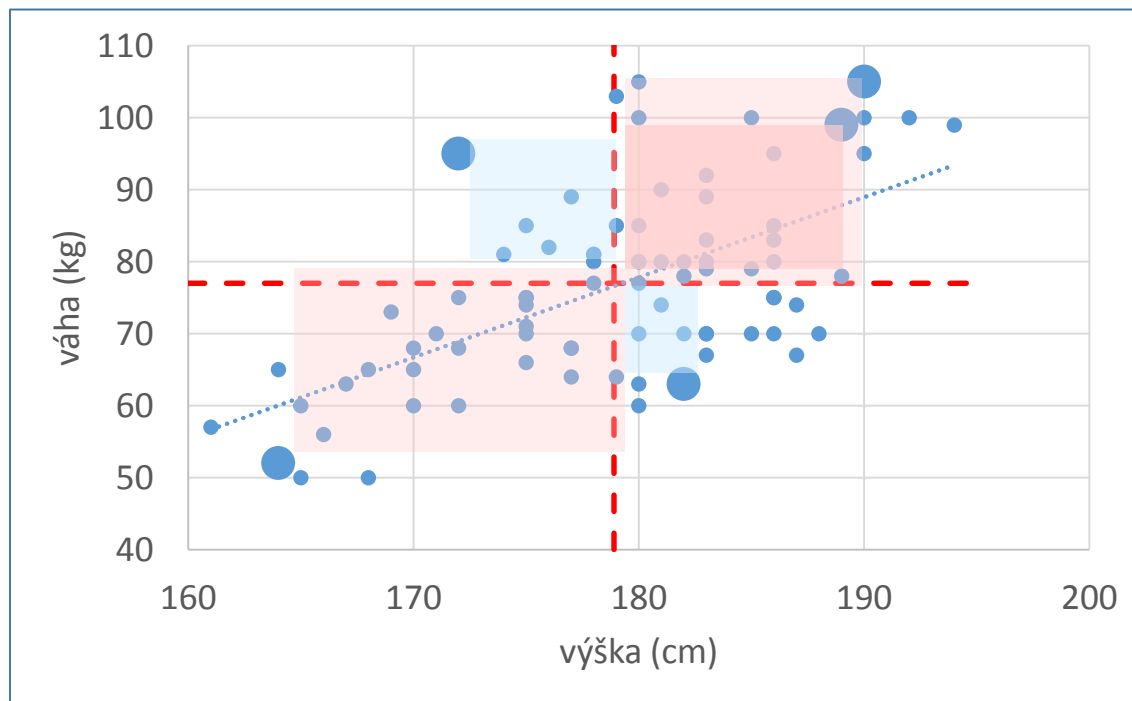
$$r(X, Y) = \frac{1}{n - 1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

Lze ukázat, že korelační koeficient může nabývat hodnot z intervalu $\langle -1; 1 \rangle$.

Analýza závislosti dvou kvantitativních znaků

Vizualizace – bodový graf



Pearsonův výběrový korelační koeficient:

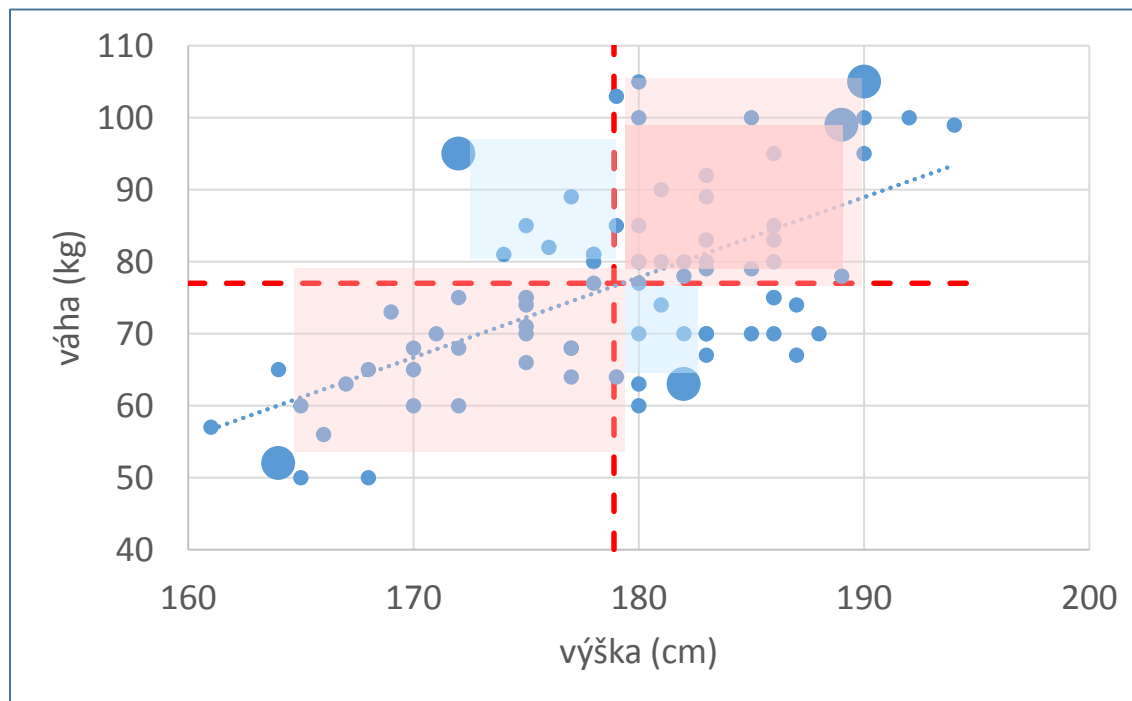
$$r(X, Y) = \frac{1}{n - 1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

$|r| = 1 \Leftrightarrow \forall i \in \{1, 2, \dots, n\}: y_i = ax_i + b$, kde $a, b \in \mathbb{R}$ (mezi proměnnými je lineární závislost)

Analýza závislosti dvou kvantitativních znaků

Vizualizace – bodový graf



Pearsonův výběrový korelační koeficient:

$$r(X, Y) = \frac{1}{n - 1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

Jak „funguje“ korelační koeficient?

Korelační koeficient (Pearsonův) je mírou **lineární** závislosti!!!

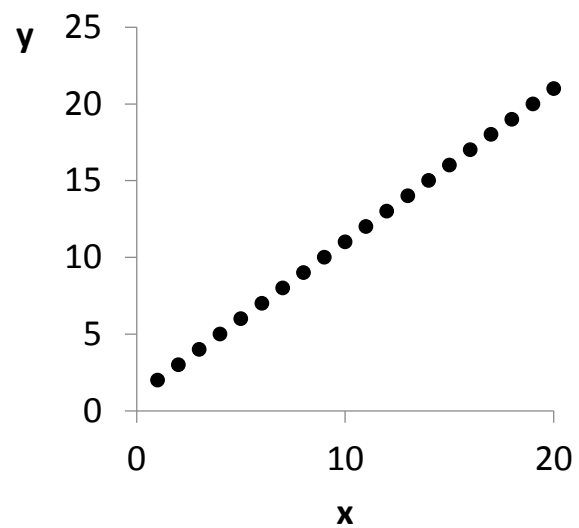
Analýza závislosti dvou kvantitativních znaků

- Vizualizace – Bodový graf
- Pearsonův výběrový korelační koeficient

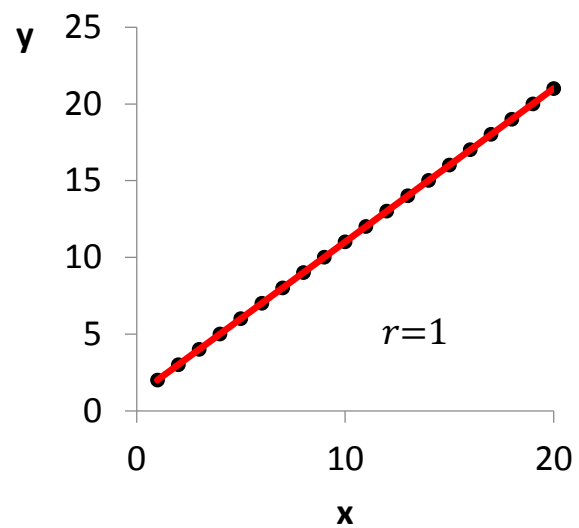
$$r(X, Y) = \frac{1}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X \cdot s_Y}$$

- $-1 \leq r(X, Y) \leq 1$,
- $r(X, Y) = r(Y, X)$,
- $r(X, X) = 1$,
- je-li $r(X, Y) = 0$, říkáme, že X, Y jsou **nekorelované** náhodné veličiny,
- je-li $r(X, Y) > 0$, říkáme, že X, Y jsou **pozitivně korelované** (s rostoucím X roste Y),
- je-li $r(X, Y) < 0$, říkáme, že X, Y jsou **negativně korelované** (s rostoucím X klesá Y),
- je-li $|r(X, Y)| = 1$, pak je mezi X a Y **lineární závislost**.

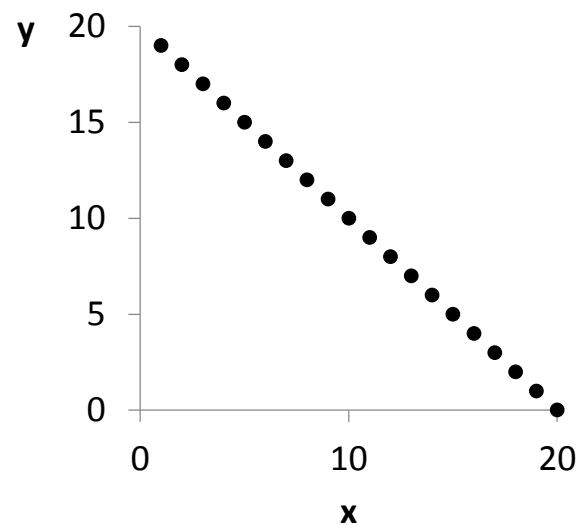
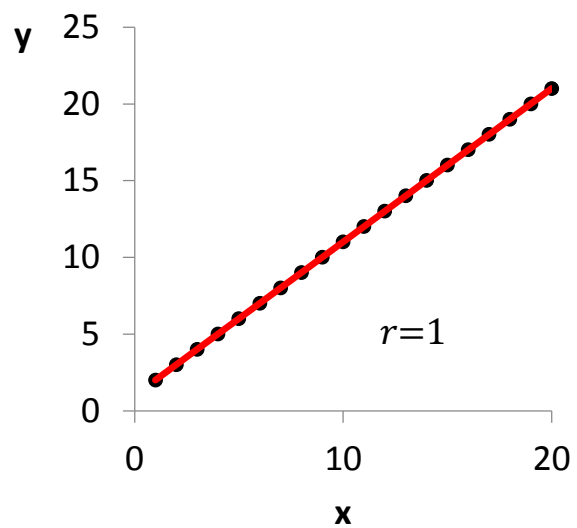
Pearsonův výběrový korelační koeficient



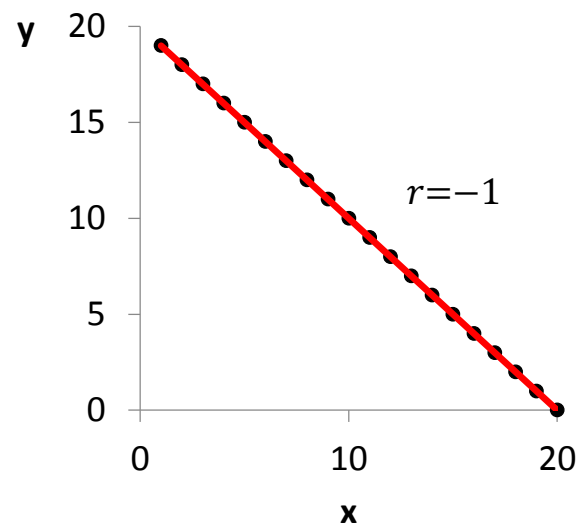
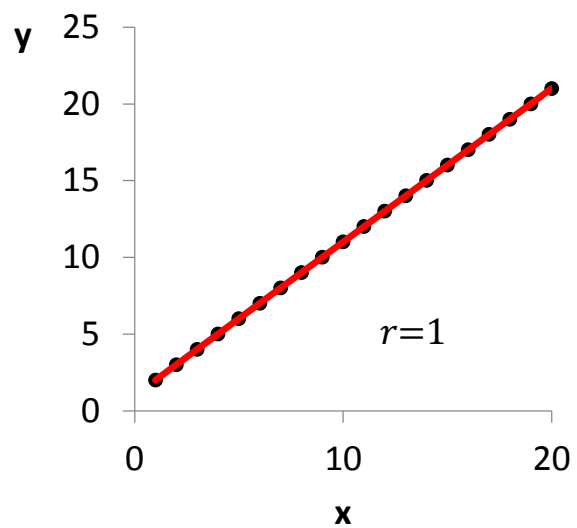
Pearsonův výběrový korelační koeficient



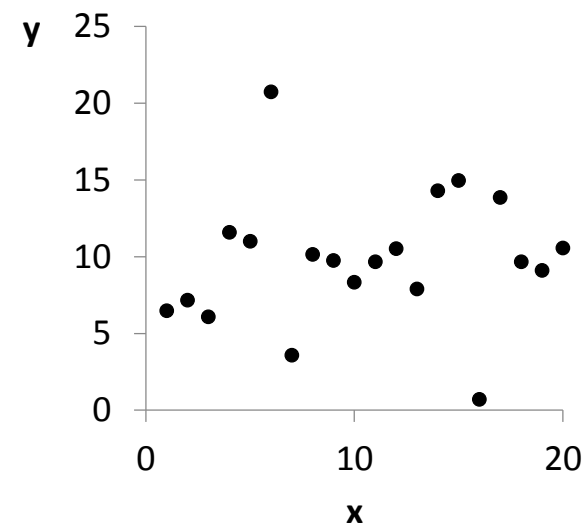
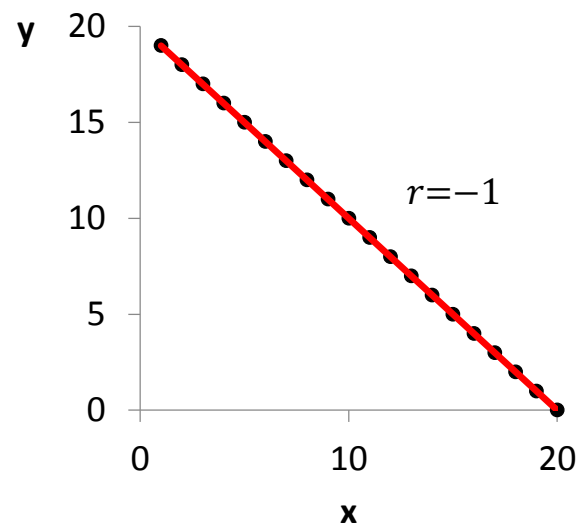
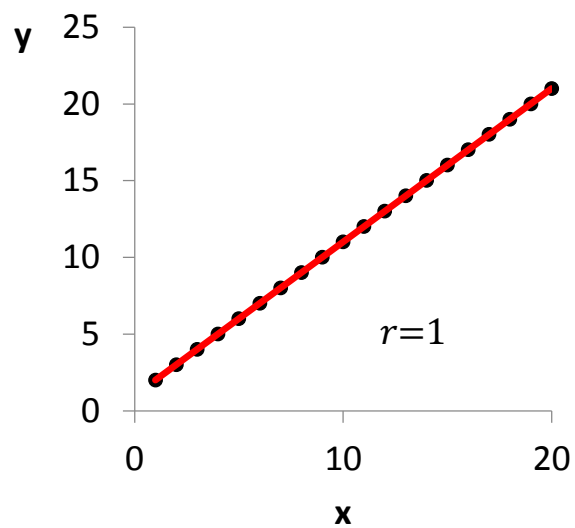
Pearsonův výběrový korelační koeficient



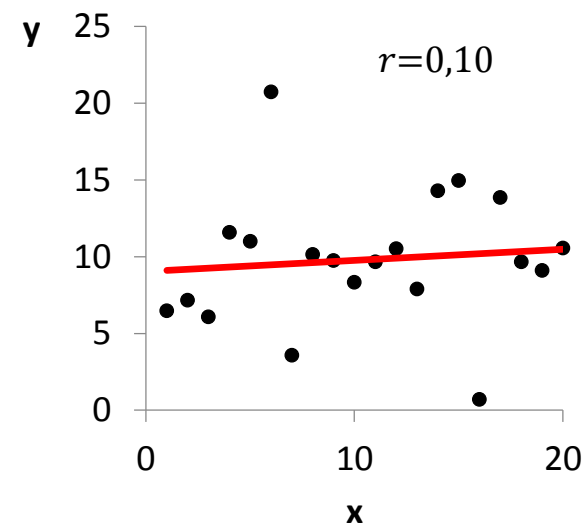
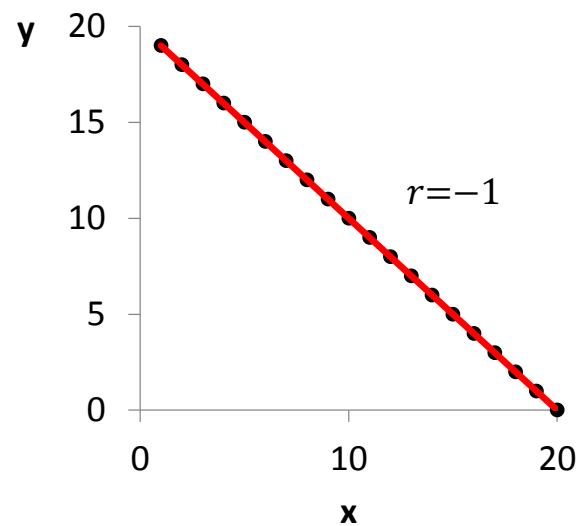
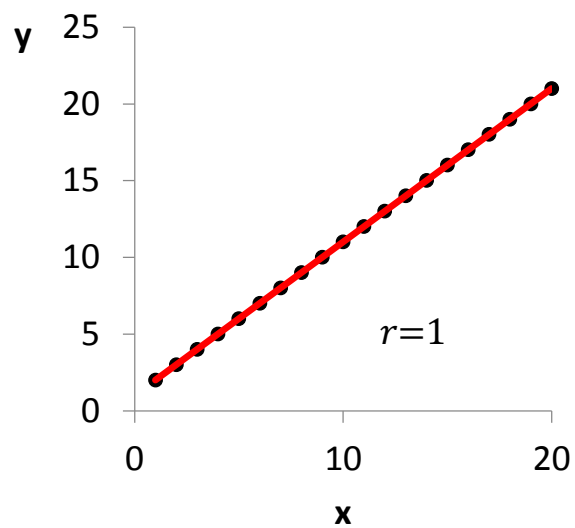
Pearsonův výběrový korelační koeficient



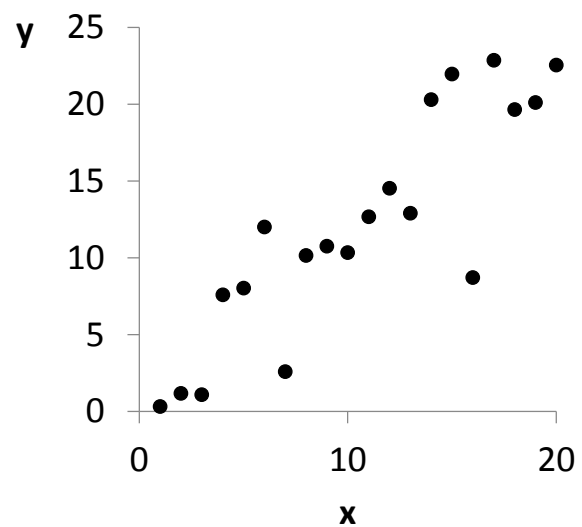
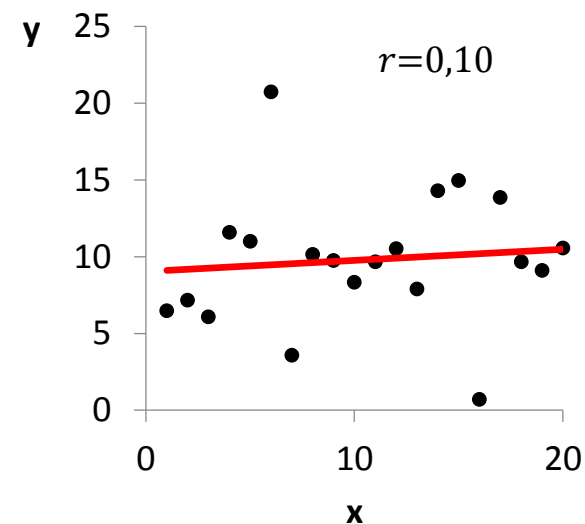
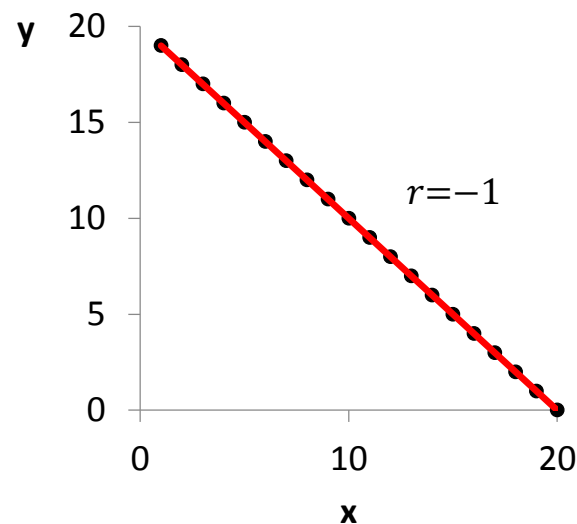
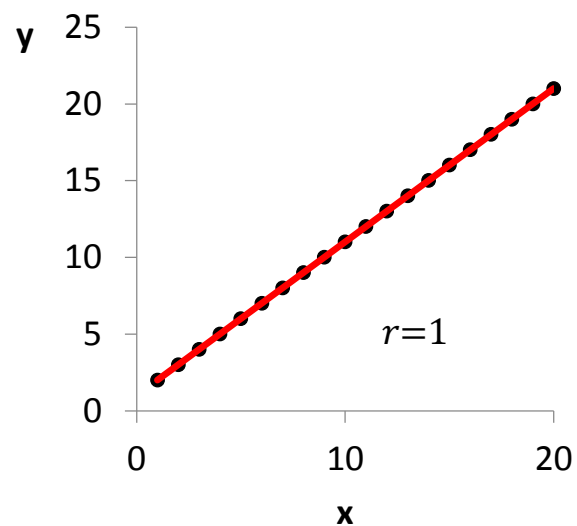
Pearsonův výběrový korelační koeficient



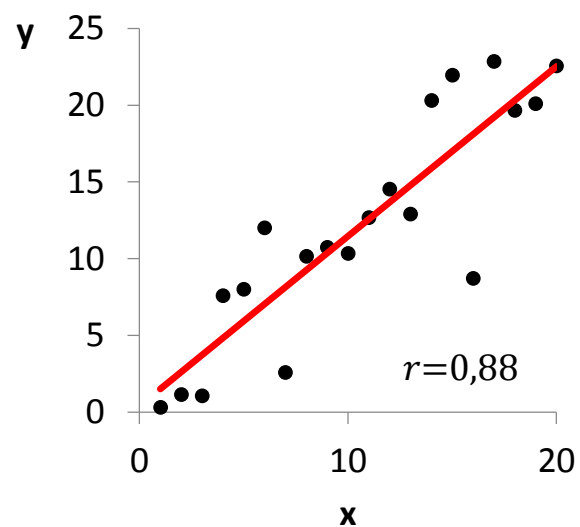
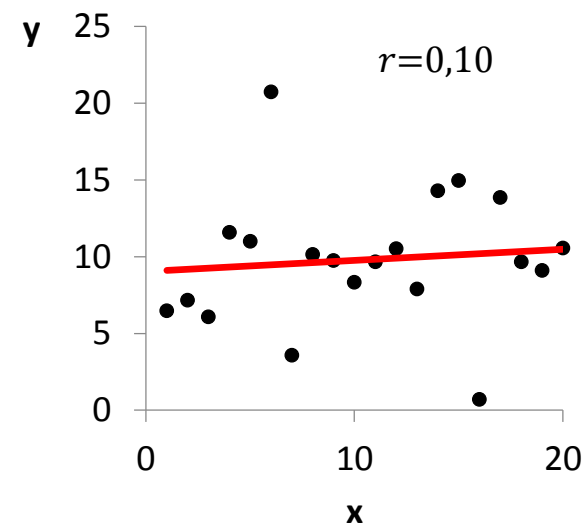
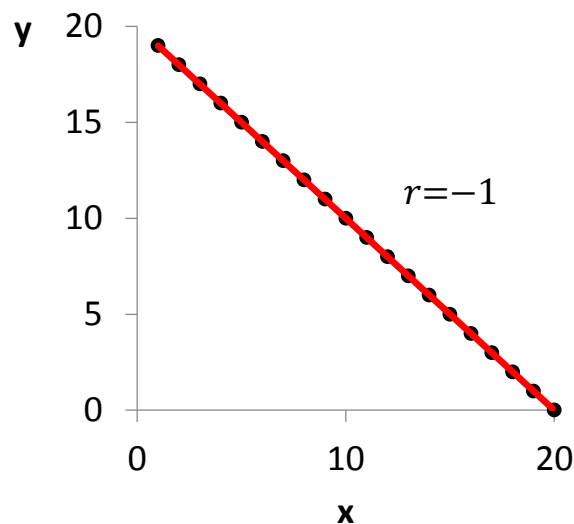
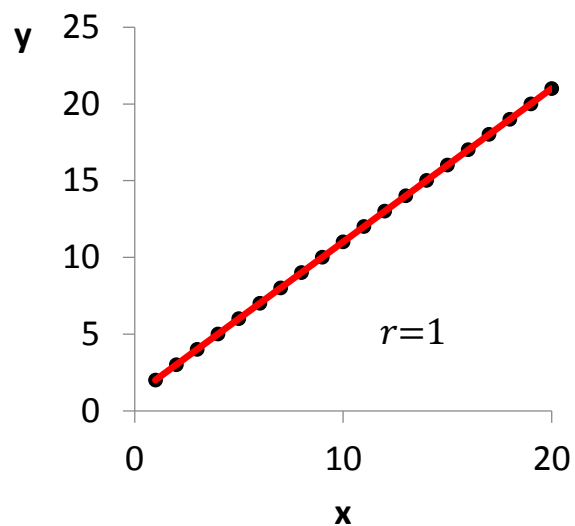
Pearsonův výběrový korelační koeficient



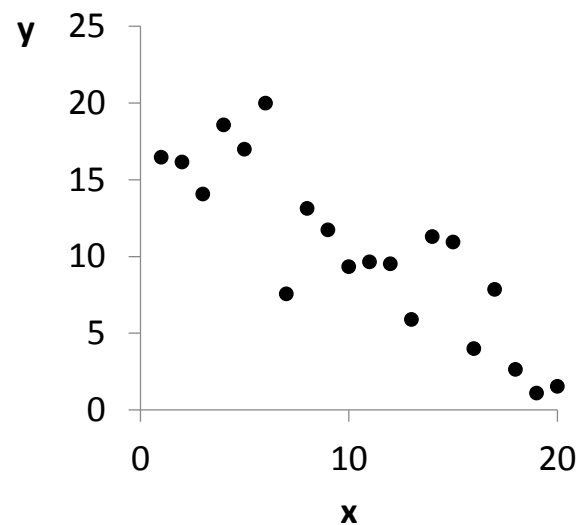
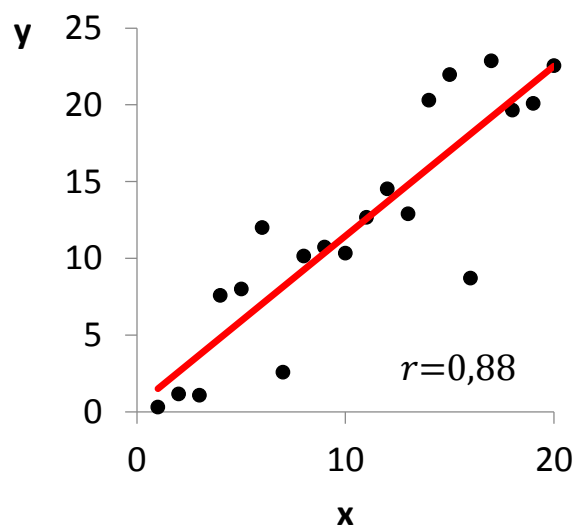
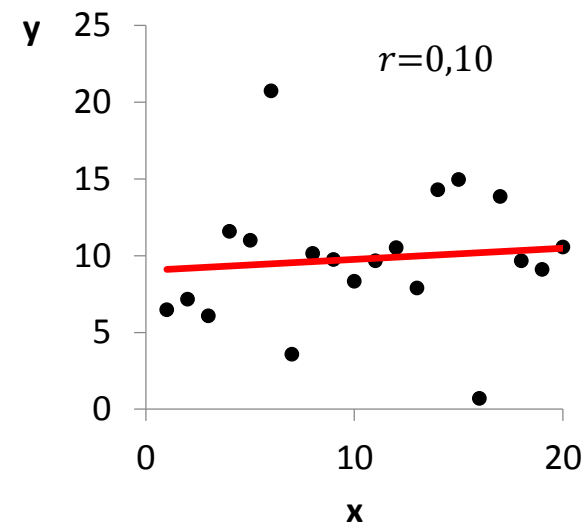
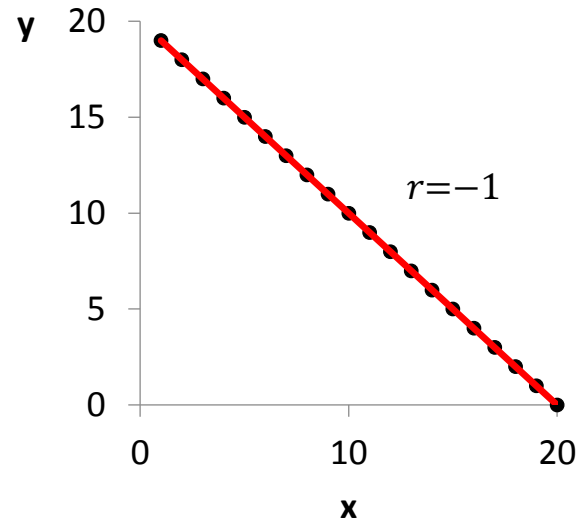
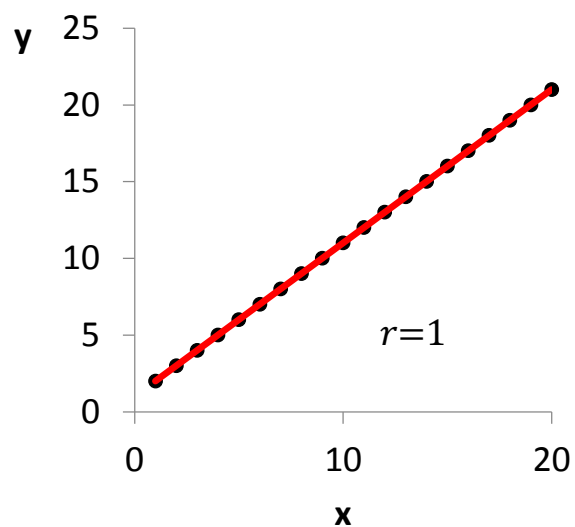
Pearsonův výběrový korelační koeficient



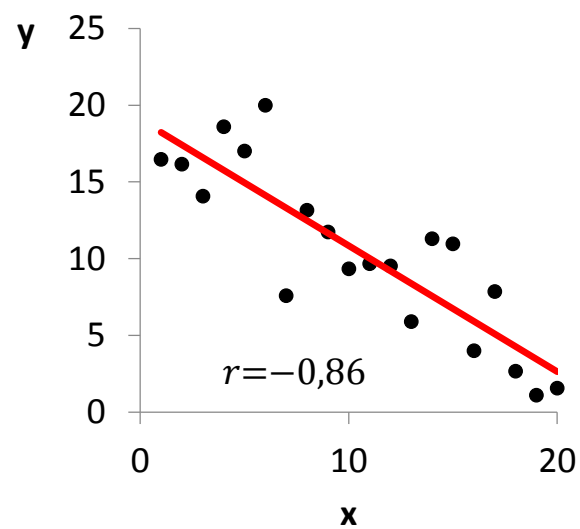
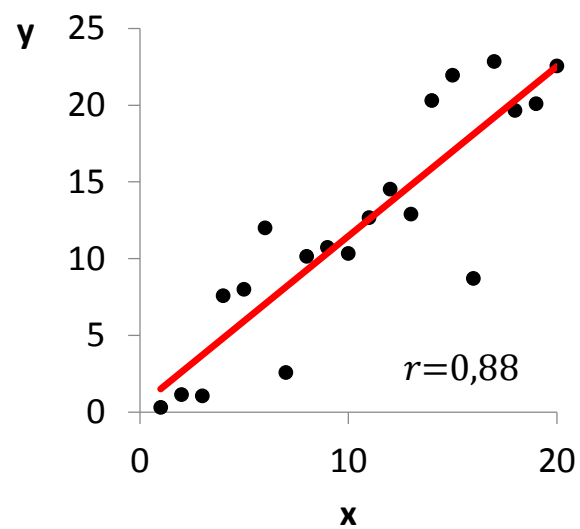
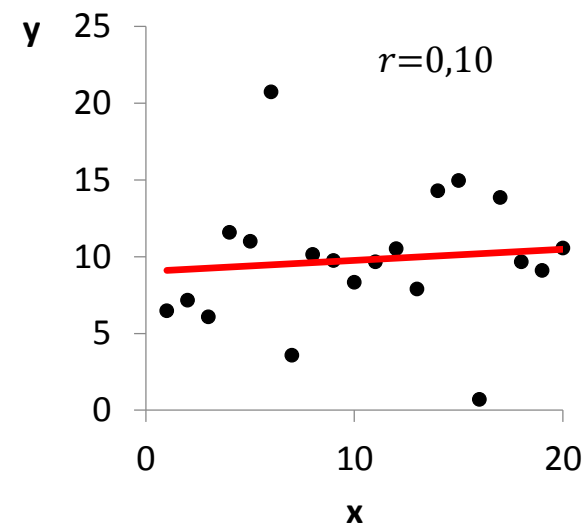
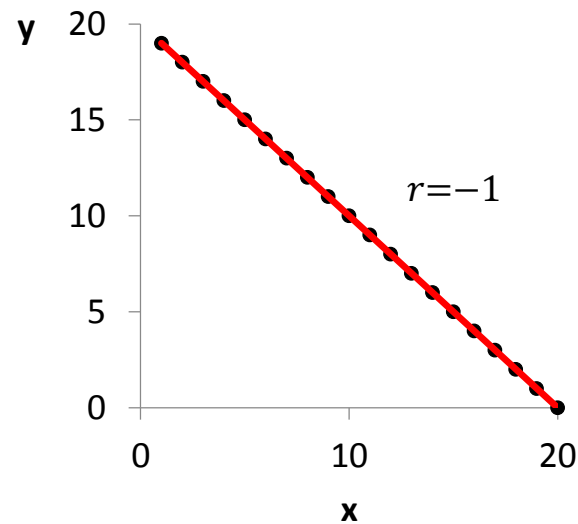
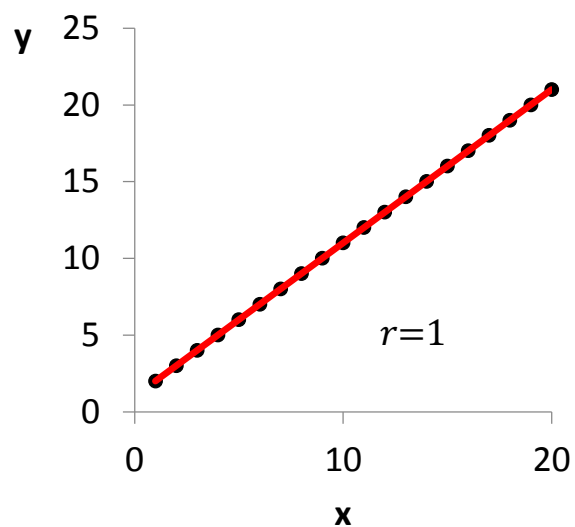
Pearsonův výběrový korelační koeficient



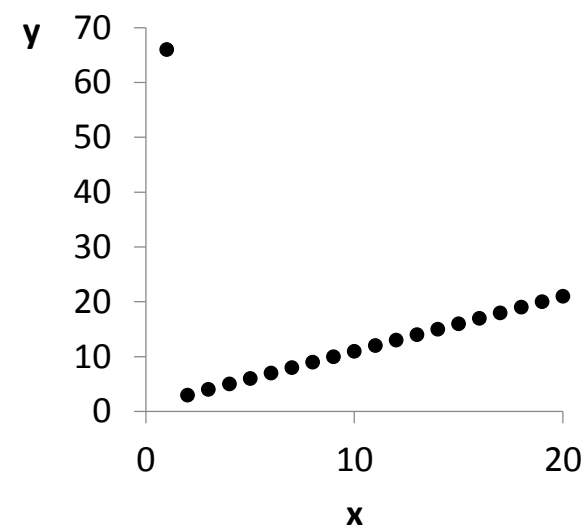
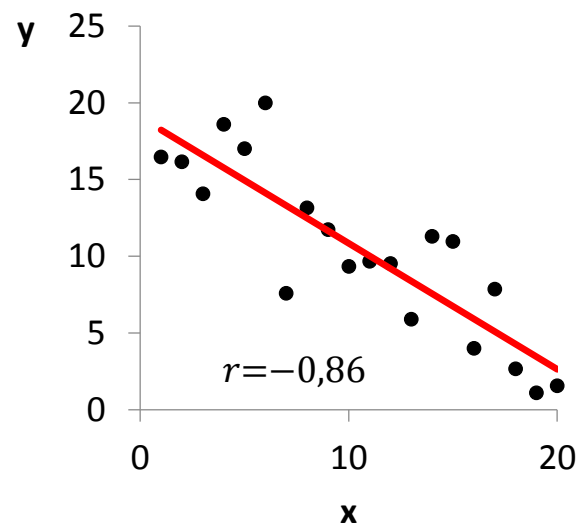
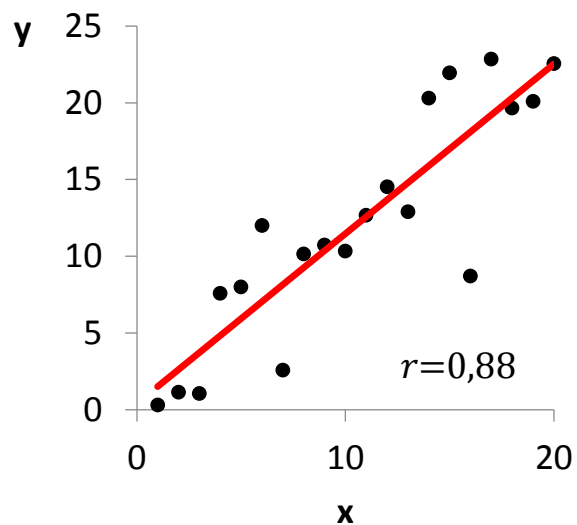
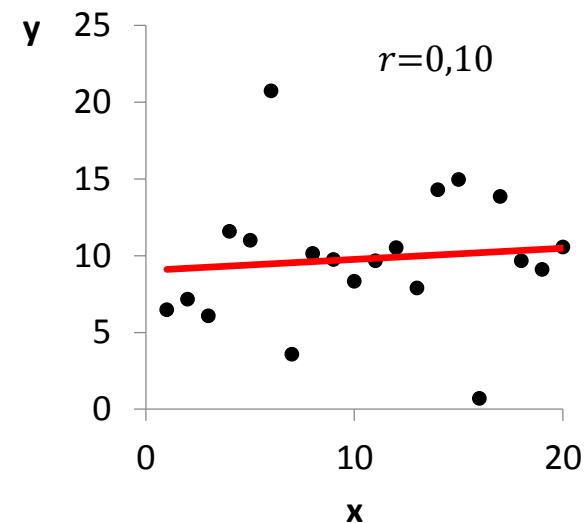
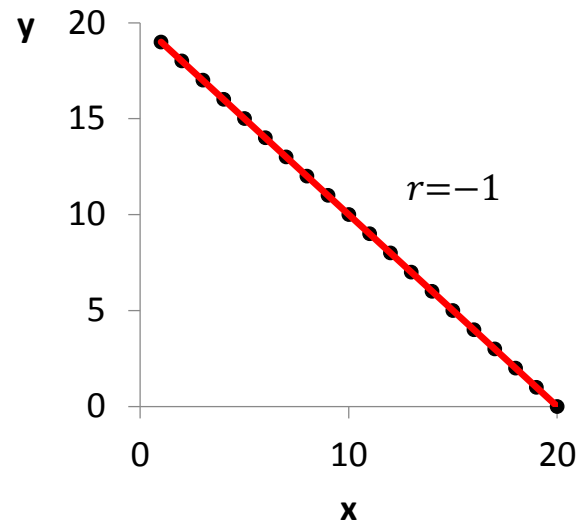
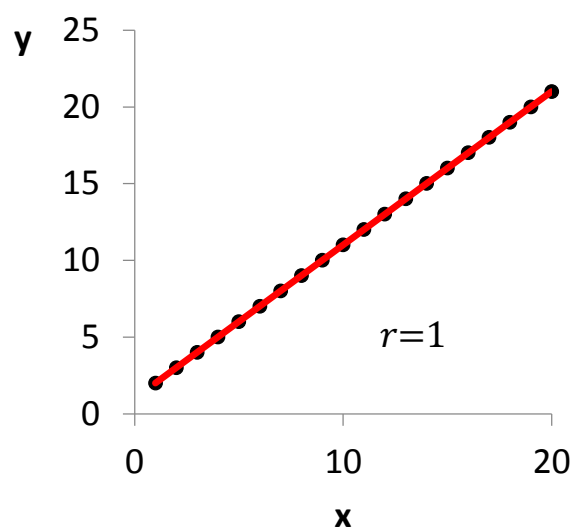
Pearsonův výběrový korelační koeficient



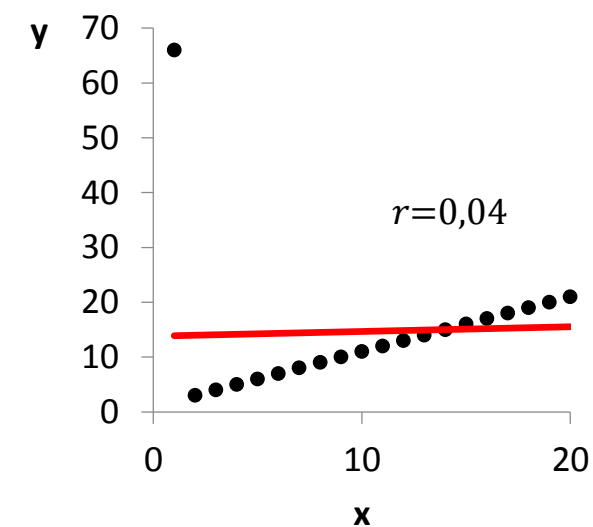
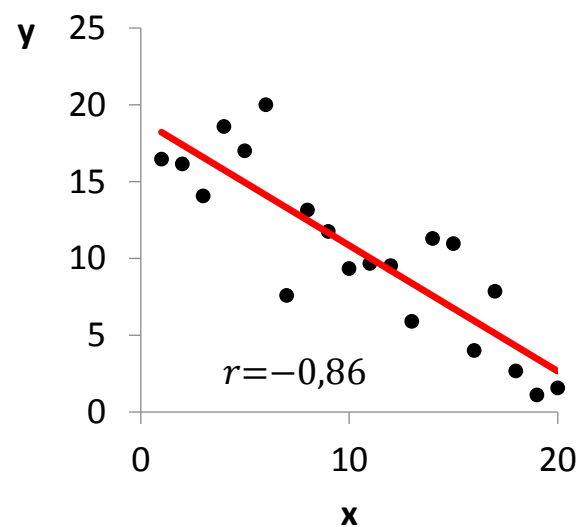
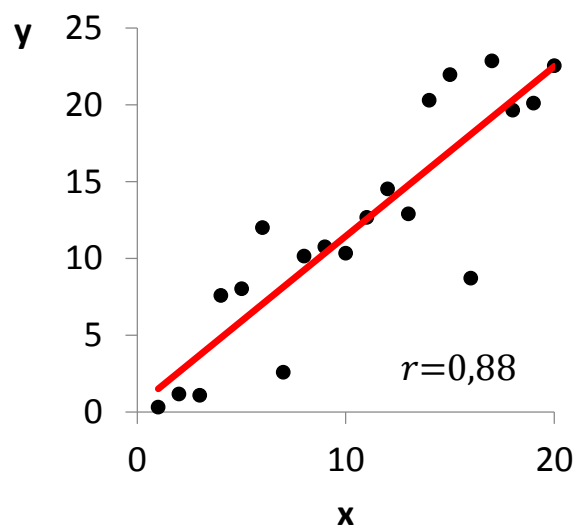
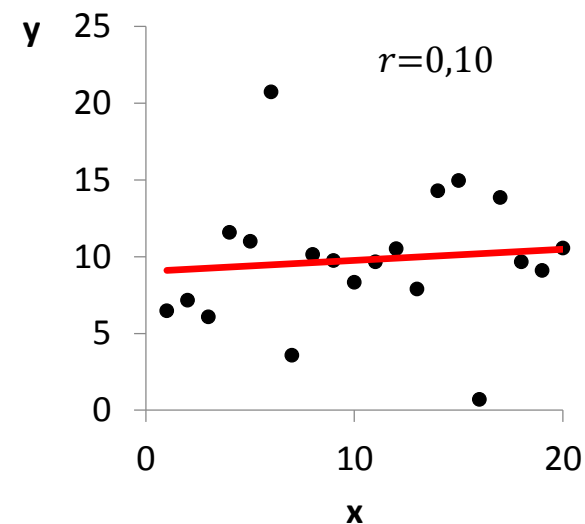
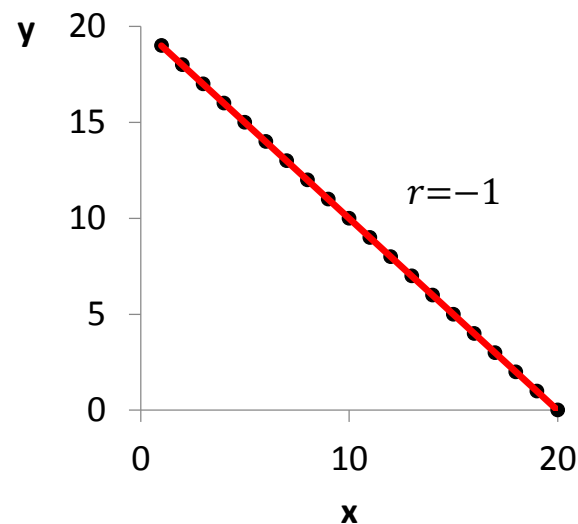
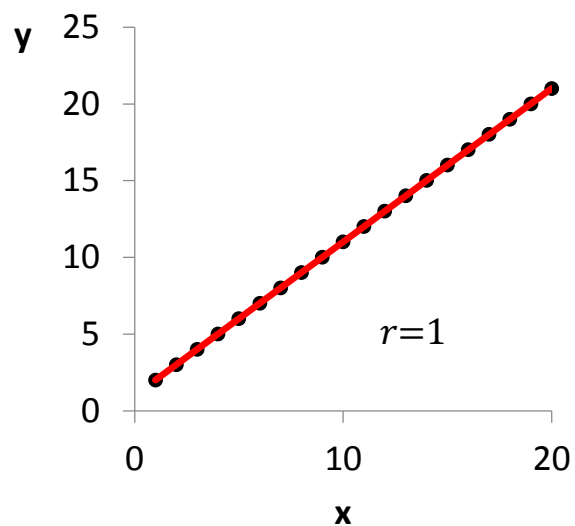
Pearsonův výběrový korelační koeficient



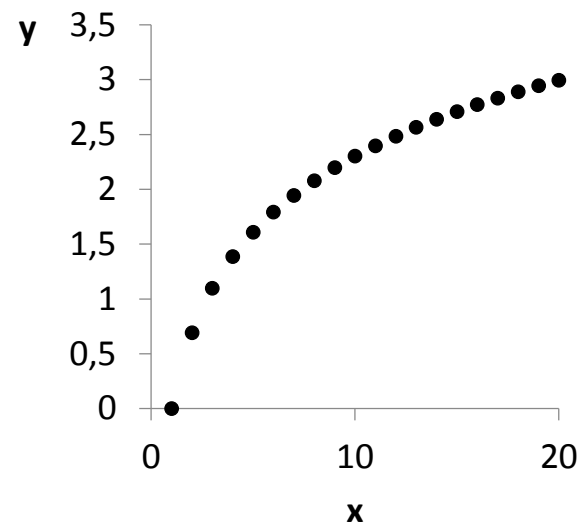
Pearsonův výběrový korelační koeficient



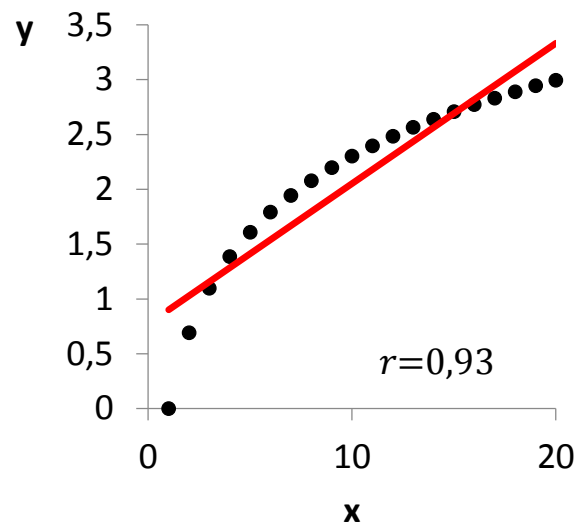
Pearsonův výběrový korelační koeficient



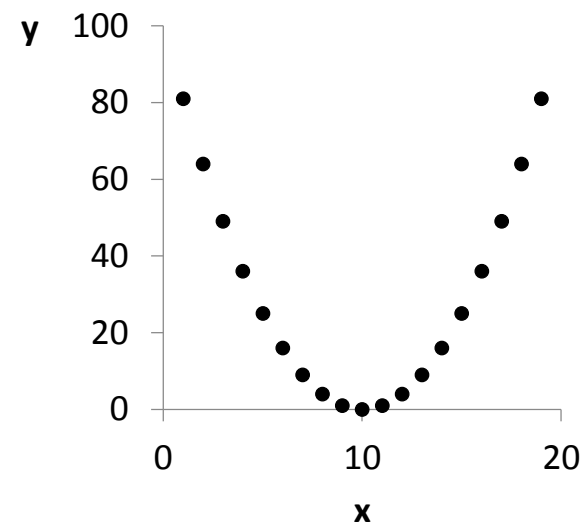
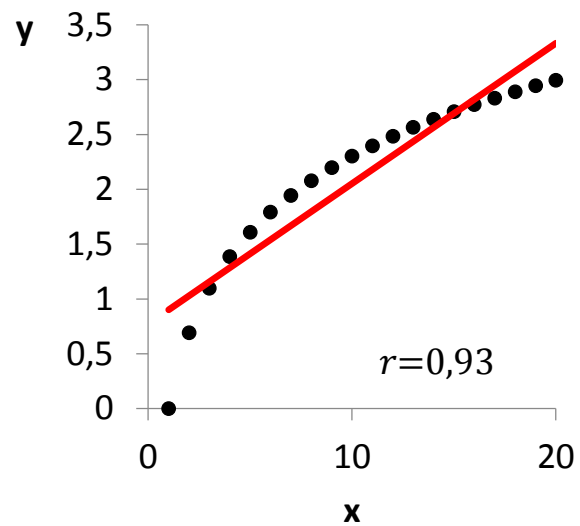
Pearsonův výběrový korelační koeficient



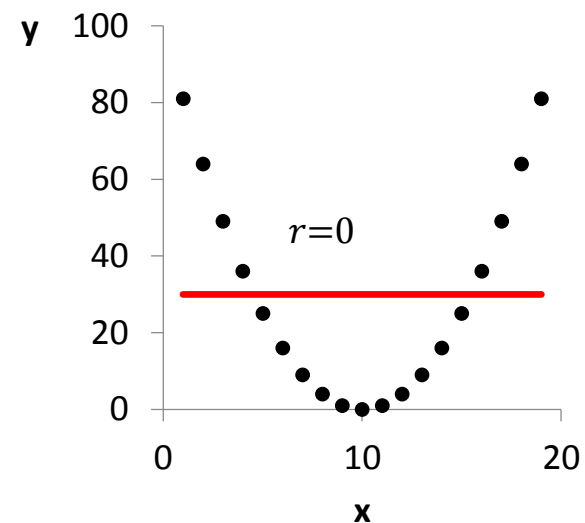
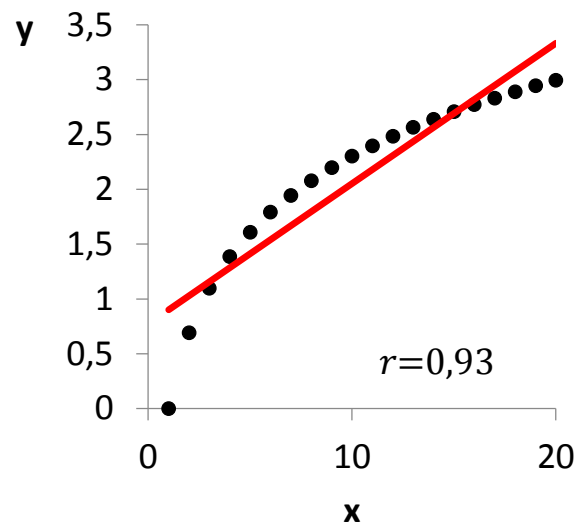
Pearsonův výběrový korelační koeficient



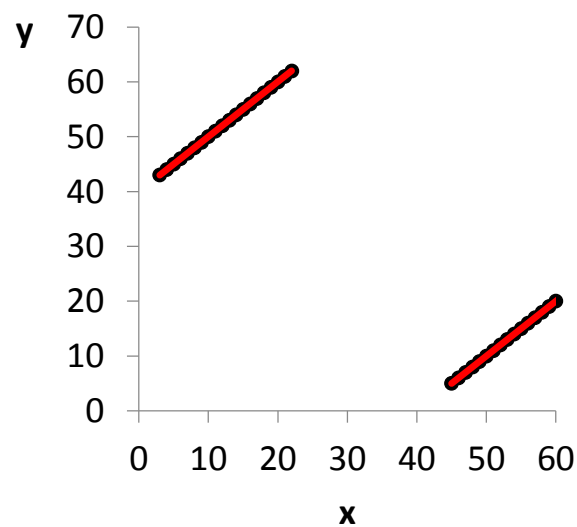
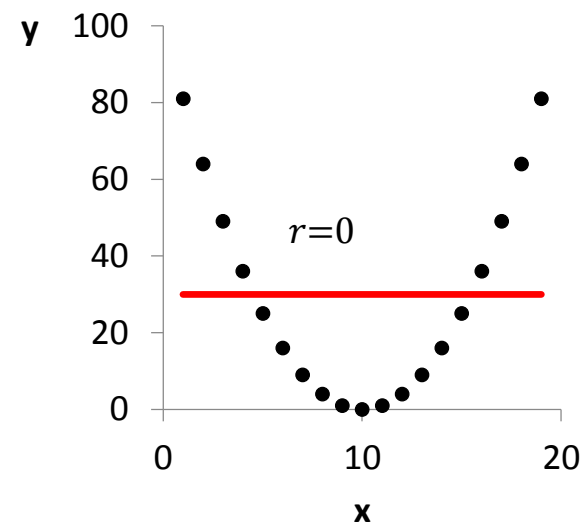
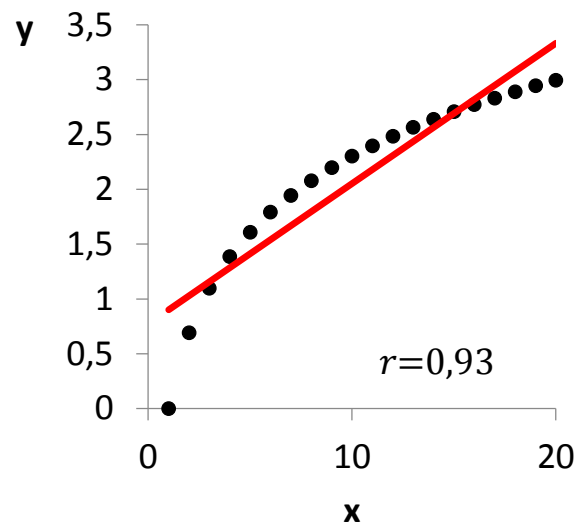
Pearsonův výběrový korelační koeficient



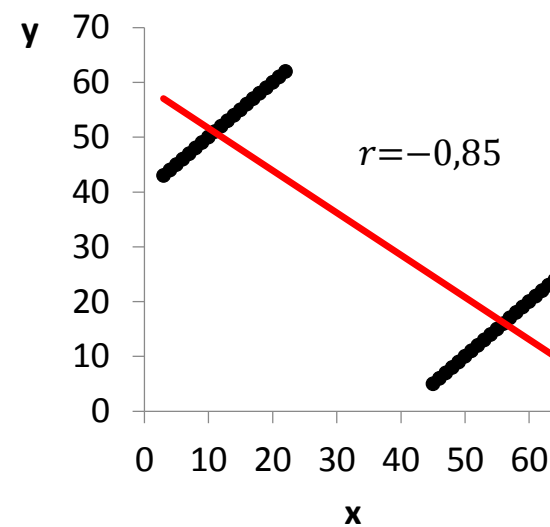
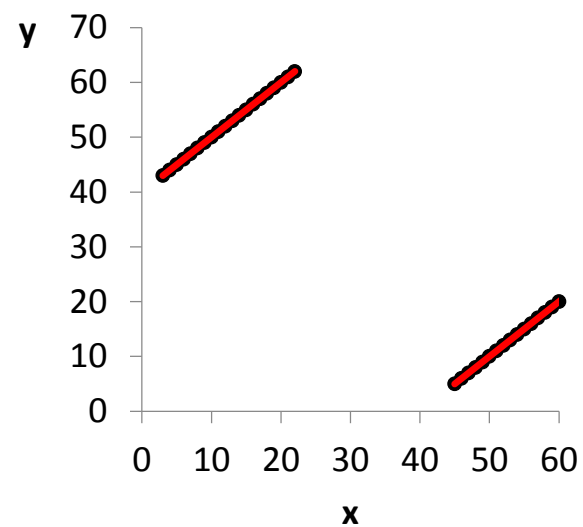
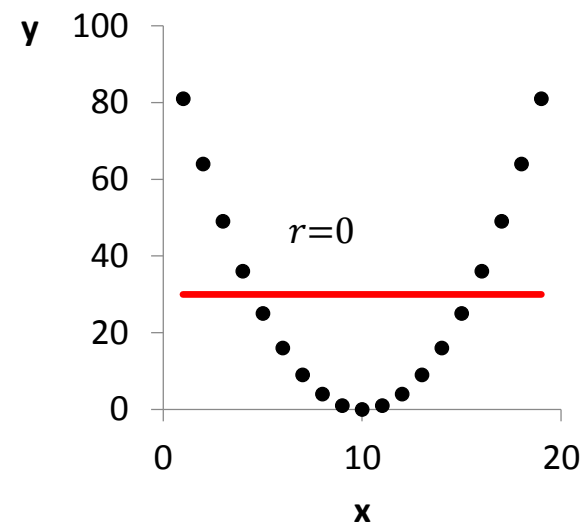
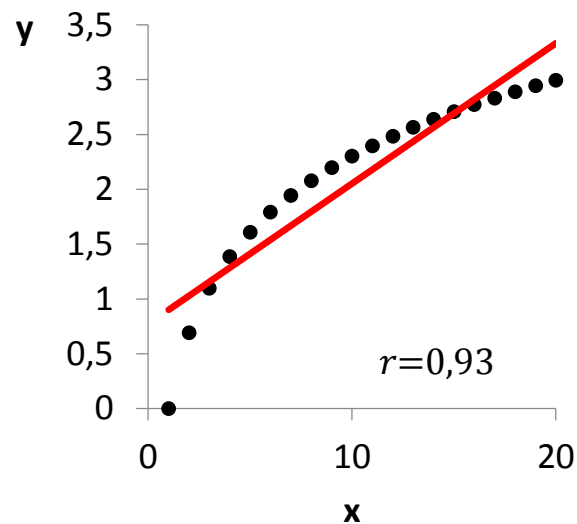
Pearsonův výběrový korelační koeficient



Pearsonův výběrový korelační koeficient



Pearsonův výběrový korelační koeficient

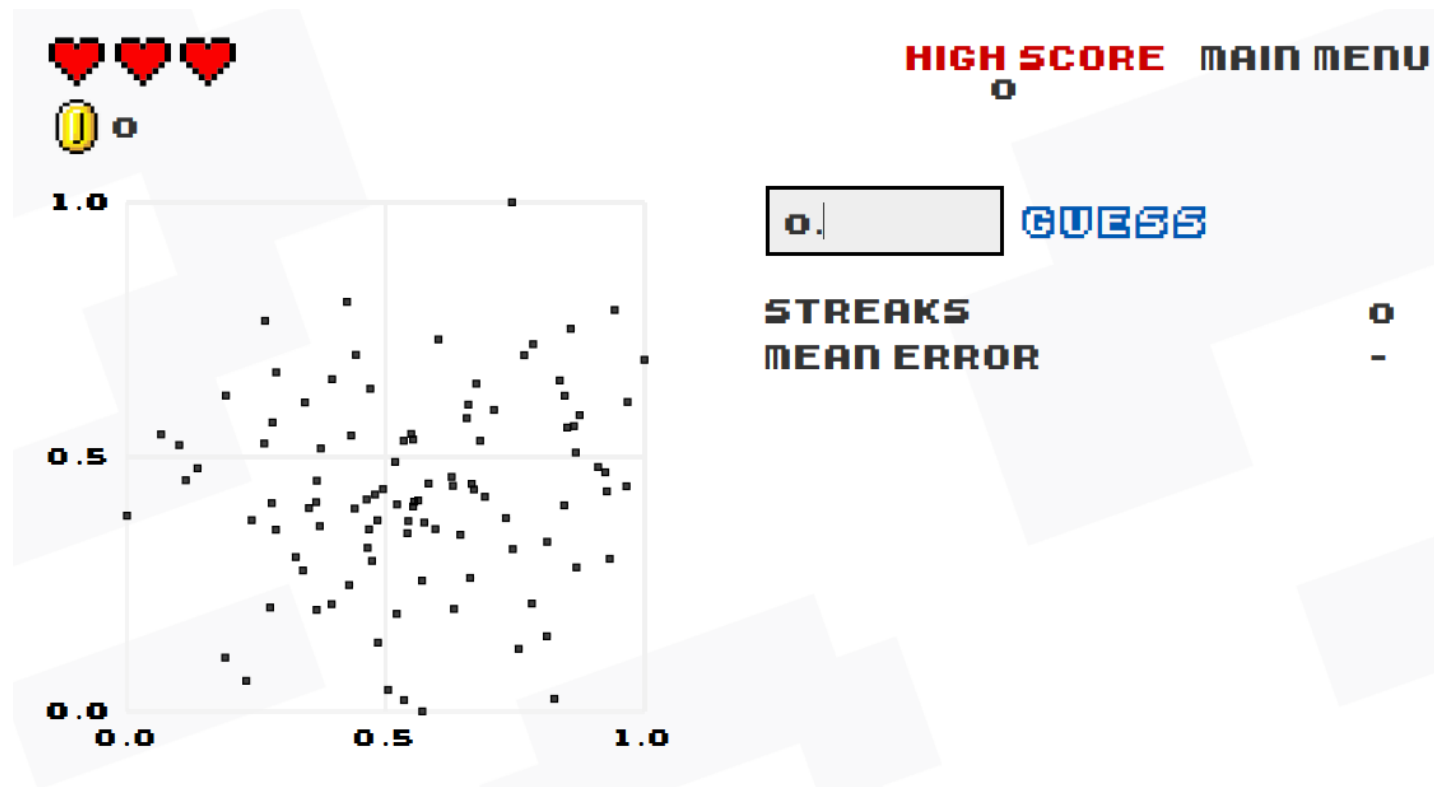


Korelační koeficient

Tip na hru do výuky:

Guess the Correlation Game

<http://guessthecorrelation.com/>

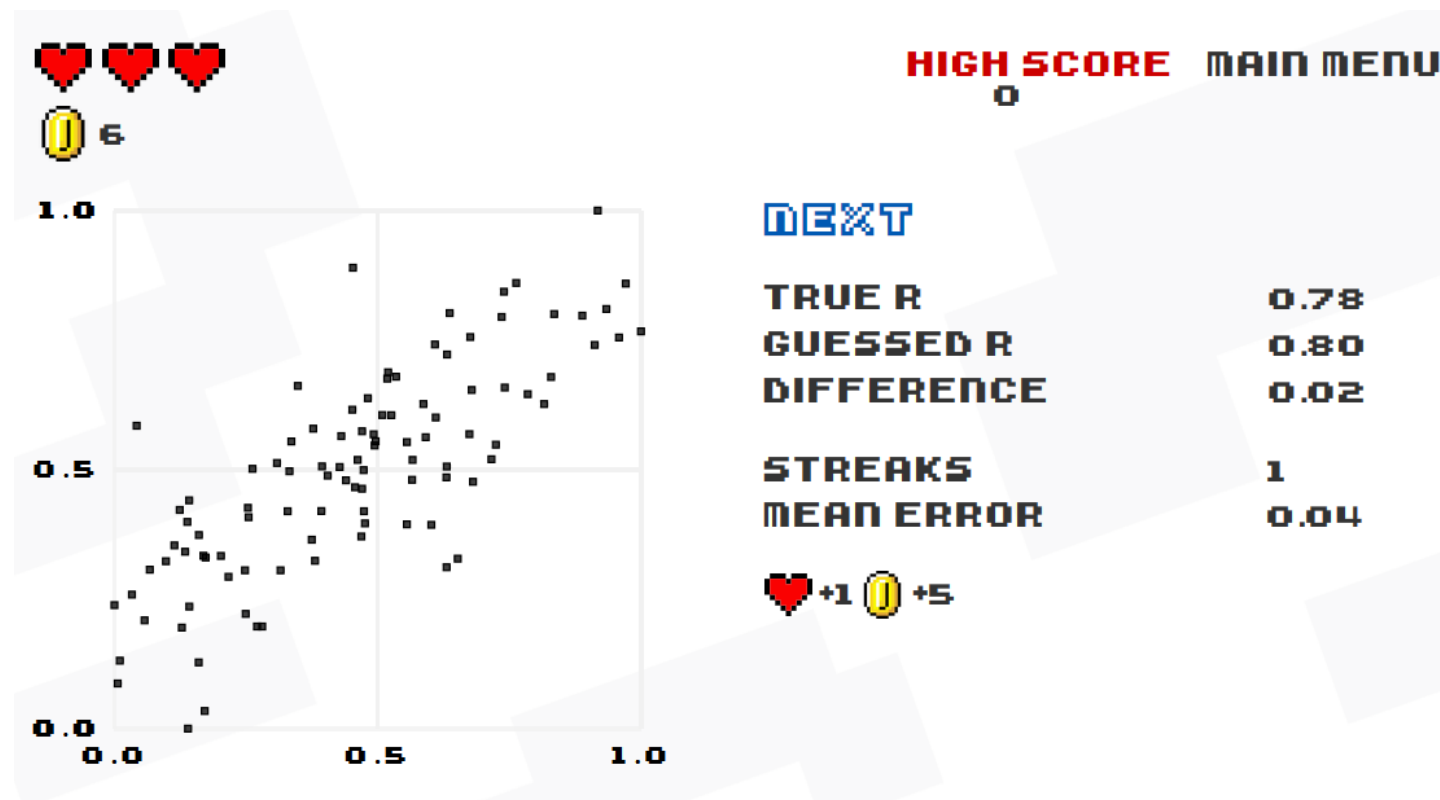


Korelační koeficient

Tip na hru do výuky:

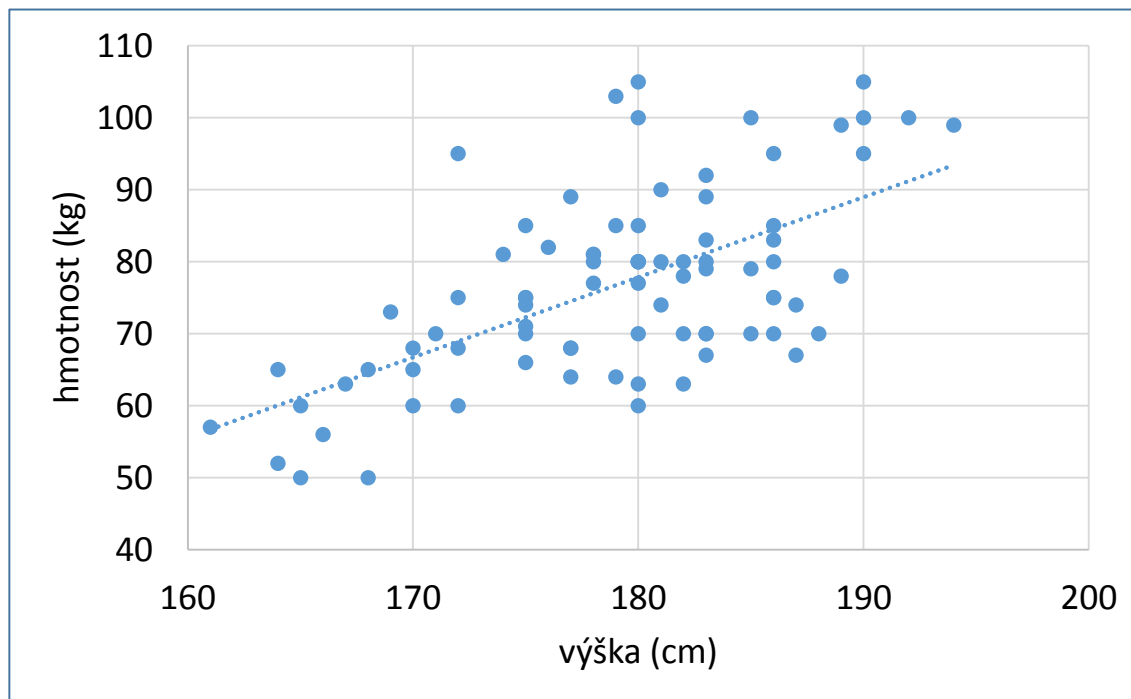
Guess the Correlation Game

<http://guessthecorrelation.com/>



Analýza závislosti dvou kvantitativních znaků

Bodový graf



Pearsonův výběrový korelační koeficient

$$r = 0,6146$$

Analýza závislosti dvou kvantitativních znaků

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Tipněte si, které dva znaky jsou negativně korelovány.

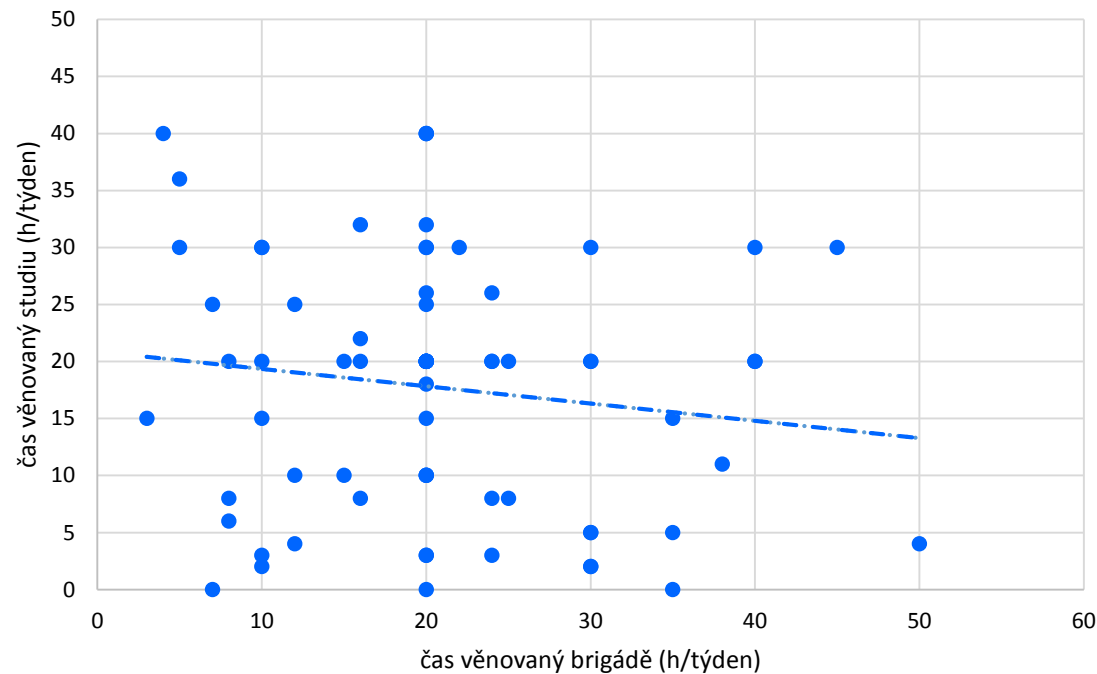
Analýza závislosti dvou kvantitativních znaků

Časová značka	Pohlaví	Výška (cm)	Váha (kg)	Přivyděláváte si v rámci prezenčního studia na brigádách?	Jak často brigádu máte?	Jak byste svou brigádu charakterizoval(a)?	Kolik času týdně obvykle věnujete brigádě?	Kolik času týdně obvykle věnujete studiu?
ID	pohlaví	výška (cm)	váha (kg)	brigáda	frekvence brigády	charakteristika brigády	čas věnovaný brigádě (h/týden)	čas věnovaný studiu (h/týden)
1.4.2016 10:38	muž	180	70	ano	každý pracovní den	praxe v oboru během studia	20	15
1.4.2016 10:41	muž	186	85	ano	nepravidelně	kancelářská práce a na ní navazující práce manuální při realizaci projektů	30	20
1.4.2016 10:41	muž	172	75	ano	nepravidelně	praxe v oboru během studia	5	36
1.4.2016 10:45	žena	166	56	ano	Různě, 2-3 týdně	Hlídní dětí	12	10
1.4.2016 10:52	žena	188	70	ano	3 dny v týdnu	praxe v oboru během studia	24	26

Tipněte si, které dva znaky jsou negativně korelovány.

Analýza závislosti dvou kvantitativních znaků

Bodový graf

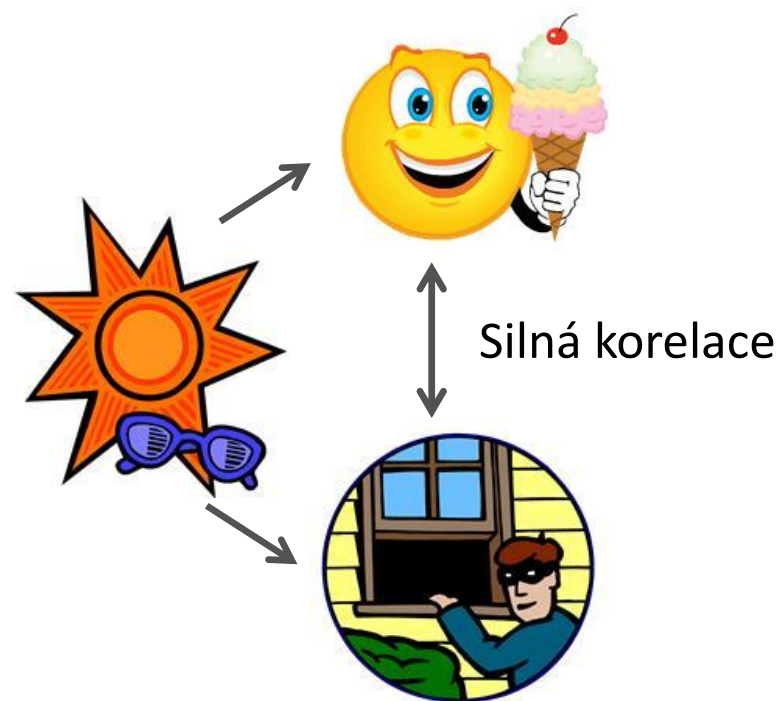


Pearsonův výběrový korelační koeficient

$$r = -0,1370$$

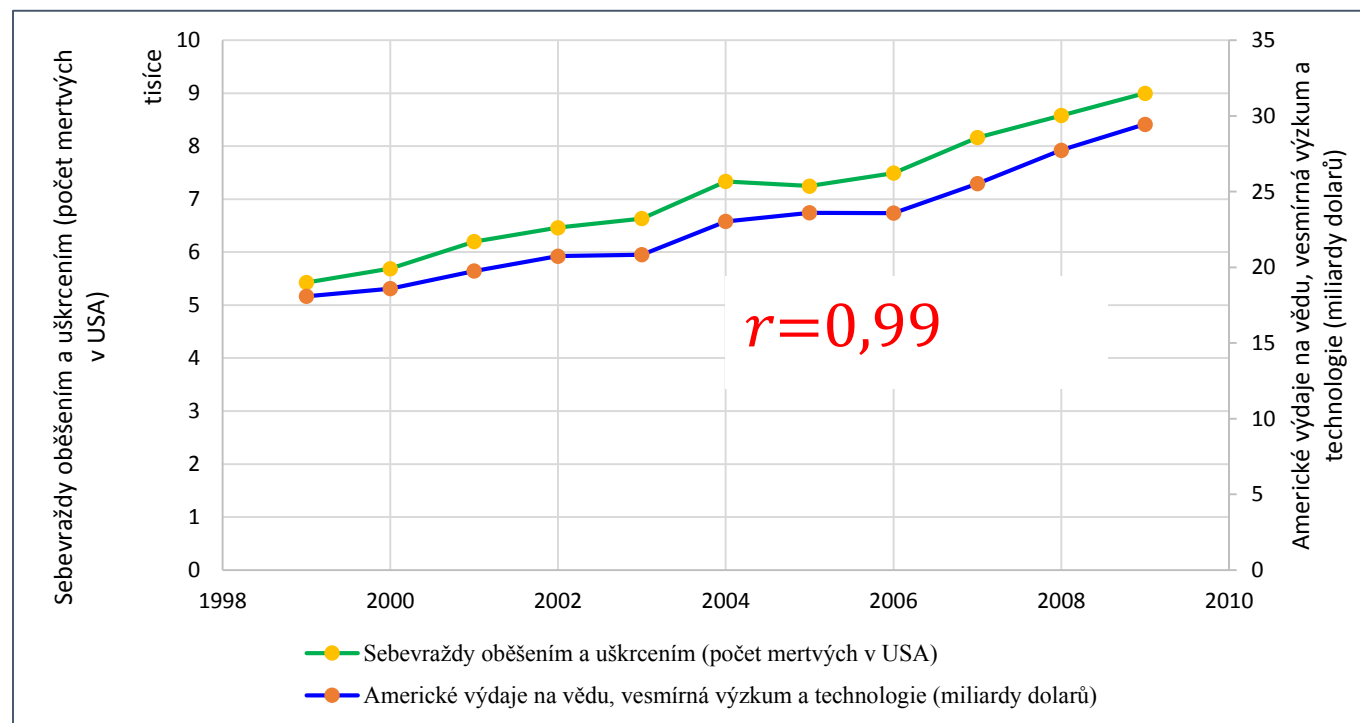
Korelační koeficient

Pokud jsou dvě náhodné veličiny korelované, znamená to pouze to, že jsou lineárně závislé. Nelze z toho však ještě usoudit, že by jedna z nich musela být **příčinou** a druhá **následkem**. To samotná korelovanost nedovoluje rozhodnout.



Korelační koeficient

Pokud jsou dvě náhodné veličiny korelované, znamená to pouze to, že jsou lineárně závislé. Nelze z toho však ještě usoudit, že by jedna z nich musela být **příčinou** a druhá **následkem**. To samotná korelovanost nedovoluje rozhodnout.



Aktuálně.cz

ZPRÁVY

Domácí Zahraničí

ZPRÁVY / ZAHRANIČÍ

K Nobelově ceně dopomáhá čokoláda, naznačuje studie

12. 10. 2012 10:36 **AKTUALIZOVÁNO**

Mezi počtem nobelistů v přepočtu na obyvatele a konzumaci čokolády je souvislost

New York - Počet nositelů Nobelovy ceny v přepočtu na jednoho obyvatele se v jednotlivých zemích odvíjí od spotřeby čokolády.

To není úryvek z reklamy na sladkosti, ale závěr studie publikované v jednom z nejprestižnějších světových lékařských časopisů New England Journal of Medicine.

Nejvíce laureátů Nobelovy ceny mají Švýcaři, kteří jsou zároveň největšími jedlíky oblíbené sladké pochutiny.

Zdroj: <http://zpravy.aktualne.cz/zahranici/k-nobelove-cene-dopomaha-cokolada-naznacuje-studie/r~i:article:760147/>



Korelační koeficient

V praxi se zpravidla hodnota koeficientu korelace interpretuje takto:

Korelační koeficient	Typ lineární závislosti
$ r = 0,0$	neexistující
$ r \in (0,0; 0,3)$	velmi slabá
$ r \in (0,3; 0,7)$	středně silná
$ r \in (0,7; 1,0)$	těsná
$ r = 1,0$	funkční

- Mezi proudem a napětím na odporu byl zjištěn korelační koeficient 0,6.
- Mezi školním prospěchem a pocitem deprese u dětí byl zjištěn korelační koeficient 0,6.

Výsledky interpretujte!



Pozvánka na seminář Škola matematického modelování



Na semináři Škola matematického modelování (ŠKOMAM)

- třídní seminář pro studenty středních škol, který existuje od roku 2005
- má dvě části
 - jedna část se skládá z popularizačních přednášek
 - druhá část je zaměřena na samotné matematické modelování



Katedra
aplikované matematiky

Pozvánka na seminář Škola matematického modelování

Škola matematického modelování (ŠKOMAM)

- co jsme v druhé části v uplynulých letech mimo jiné zvládli?
 - navrhovali jsme optimální karoserii auta tak, aby mělo nejlepší aerodynamické vlastnosti
 - ukázali jsme si, jak upravovat fotografie pomocí nástrojů lineární algebry
 - pomocí diferenciálních rovnic jsme modelovali změnu koncentrace léku v krvi a díky tomu jsme byli schopni navrhnout optimální interval mezi jeho pravidelným podáváním



Katedra
aplikované matematiky

Pozvánka na seminář Škola matematického modelování

Škola matematického modelování (ŠKOMAM)

- 15. ročník se koná **28.-30.1.2019** v Ostravě
- Více zde: <http://skomam.vsb.cz/>



Katedra
aplikované matematiky

DĚKUJI ZA POZORNOST!

martina.litschmannova@vsb.cz

