JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH

Pedagogická fakulta

Katedra matematiky

Bakalářská práce

Statistika v příkladech pro vzdělávání na středních školách

Vypracovala: Pavla Kaiseršatová

Vedoucí práce: doc. RNDr. **Vladimíra Petrášková**, Ph.D.

České Budějovice 2017

Prohlášení

Prohlašuji, že svoji bakalářskou práci na téma Statistika v příkladech pro vzdělávání na středních školách jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě, elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích .... ……………………………

Poděkování

Ráda bych touto cestou poděkovala doc. RNDr**. Vladimíře Petrášková,** Ph.D. za odborné vedení a podnětné rady.

Statistika v příkladech pro vzdělávání na středních školách

Statistics in the examples for secondary school education

Anotace

Cílem práce je vytvořit učební materiál základů statistiky, který bude určen pro žáky 2. stupně základní školy a studenty střední školy. Práce bude zaměřena na následující téma: charakteristika polohy – aritmetický průměr, modus a medián; charakteristika variability – rozptyl a směrodatná odchylka.

Abstract

The aim of this thesis is to create didactic material of basic statistics which is intented for the secondaty and high school students. The work will be focused on the following topic: measure of location – arithmetic mean, mode and median; measure of variability – variance and standart deviation.

Obsah

Úvod

Cílem této bakalářské práce je nastínit a přiblížit základy statistiky žákům základní školy a studentům střední školy prostřednictvím uceleného a přehledného učebního materiálu a ukázkových řešených příkladů z běžného života. Pokusím se vysvětlit na jednotlivých příkladech, co znamená aritmetický průměr, geometrický průměr, harmonický průměr, modus, medián, směrodatná odchylka, rozptyl, jak se v praxi počítá a k čemu nám slouží. Obsah práce vychází z programu RVP pro střední školy.

Proč jsem si zvolila toto téma bakalářské práce? Vysvětlení je jednoduché. Podle mého názoru je celkově matematika velmi obávaný předmět a na střední škole málo oblíbený. Je potřeba vytvářet přehledné materiály pro usnadnění pochopení všech matematických témat. A toto je i mým cílem, vytvořit přehledný učební materiál s ukázkovými příklady pro lepší pochopení základů statistiky i pro samostudium.

1. Co je to statistika

Statistika je věda, která se zabývá sběrem, tříděním a zpracováním dat. Vychází ze shromážděných statistických údajů. Slovo *statistika* pochází z latinského „status“, což znamená *stav*. Se statistickými údaji se setkáváme v běžném životě každý den. V novinách se setkáváme s daty zpracovanými do tabulek i grafů. Statisticky jsou zpracovány sportovní výsledky, nehody, dlouhodobý stav počasí, výsledky ve škole, porodnost i úmrtnost. Každý člověk je zahrnut do statistického souboru při sčítání lidu (Zwerenz 2015).

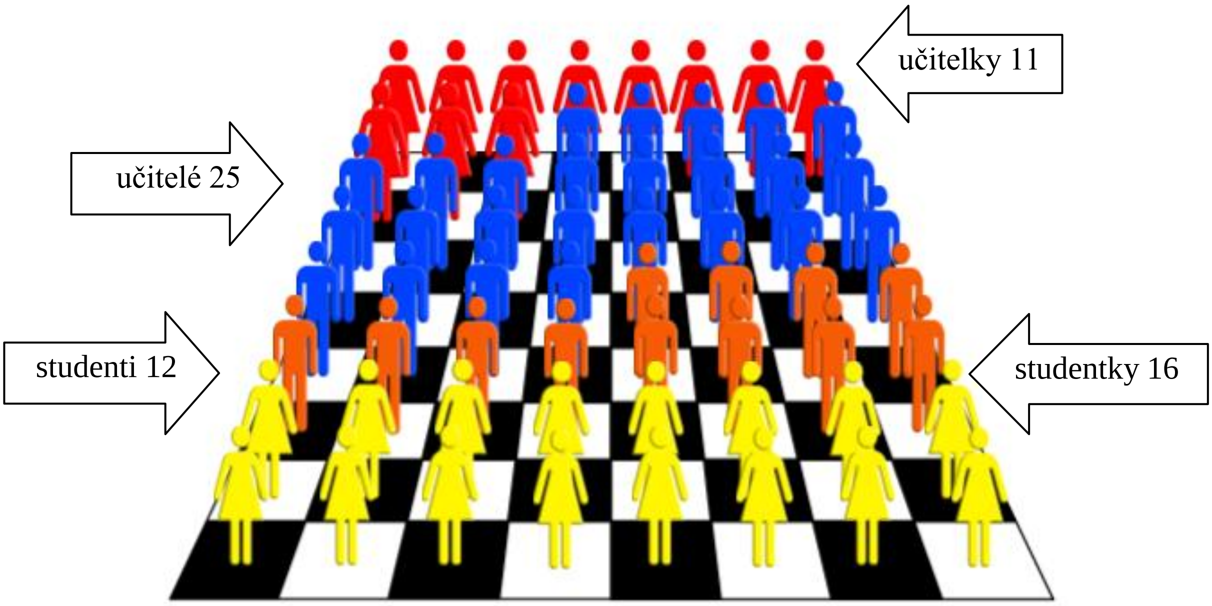
1.1. Základní pojmy

Na začátek si pouze ve stručnosti představíme pár základních pojmů, se kterými se budeme ve statistice potkávat a podrobněji si je rozebereme v následujících kapitolách. **Statistická data** – jsou údaje o tzv. statistických jevech, např. společenských, přírodních, technických aj.

Statistický soubor – je soubor osob, věcí, událostí, jevů aj. shromážděných na základě určité společné vlastnosti.

Statistické jednotky – jsou jednotlivé prvky statistického souboru.

Rozsah souboru – je počet všech prvků statistického souboru (Calda a Dupač 1993).



Obrázek č. 1 - Soubor dat. Znázornění tříd osob ve škole.

1.2. Sběr a třídění dat

Nasbírané informace označujeme jako data. Obvykle se jedná o číselné údaje. Data se musí dále před samostatnou prezentací rozdělit do skupin neboli tříd, aby seznamy údajů dávaly nějaký smysl. Data dále prezentujeme přehledným způsobem v podobě tabulek nebo grafů. Nezpracovaným nasbíraným informacím říkáme syrová data.

Všechna data je třeba uváženě nasbírat a roztřídit před samotnou prezentací a analýzou dat. Nejběžnější způsob sběru dat je průzkum. Určité skupině lidí je položena otázka týkající se jejich názorů, návyků či preferencí právě v podobě dotazníku. Jejich odpovědi představují syrová data, která pak třídíme do tabulek a znázorňujeme pomocí grafů.

Dotazník bývá sestaven nejčastěji formou více výběrového testu, kdy si vybíráte odpověď z předem daných možností. Odpovědi se poté snadněji rozdělují do jednotlivých tříd dat (Vordermanová 2015).

Data můžeme dělit dle typu měření na nominální, ordinální a metrická.

Nominální data mají význam kvality, někdy se mluví o kvalitativních datech. Nedají se seřadit, nemají žádnou velikost. Příkladem je pohlaví, rodinný stav nebo krevní skupina. Zvláštní skupinou jsou alternativní data, která nabývají pouze hodnot Ano a Ne. Lze u nich vypočítat modus. (Podrobněji viz kapitola 3.2 Modus)

Ordinální data jsou podobná jako nominální, také nenabývají číselných hodnot, ovšem dají se seřadit. U každé vybrané dvojice lze jednoznačně určit, která hodnota je větší a která menší. Příklad je nejvyšší dosažené vzdělání. Lze u nich vypočítat modus a medián. (Podrobněji viz kapitola 3.2 Medián)

Metrická data neboli kardinální jsou číselná data, která se dělí na intervalová – lze u nich určit o kolik je jedna hodnota větší či menší než druhá a poměrová – u nichž lze vypočítat kolikrát je jedna hodnota menší či větší než druhá. Patří sem například věk, cena výrobku, roční příjem aj. (IASTAT 2001).

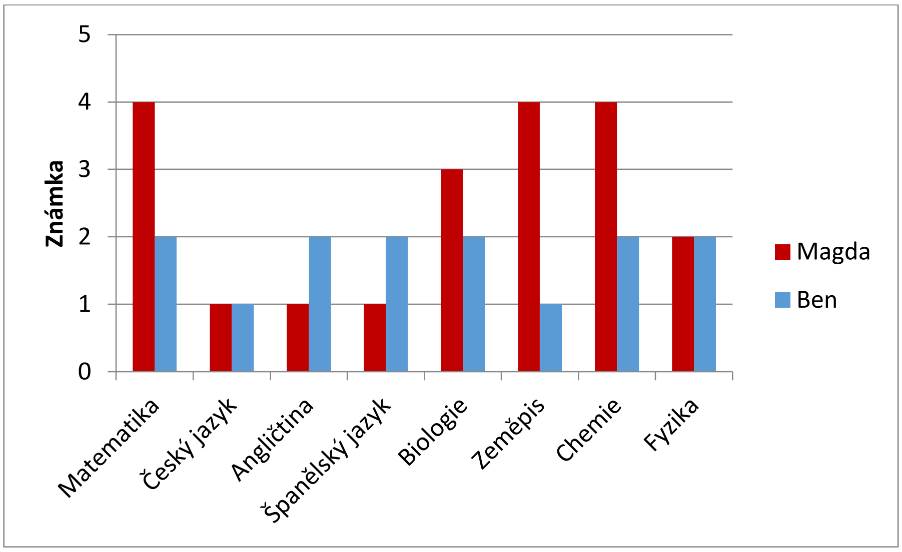
Ukázka

Rozdělení dat znázorňuje rozpětí metrických dat. Z tabulek obsahující data lze sestavit grafy, které znázorňují rozpětí jednotlivých souborů dat. Z grafů se tedy dá vyčíst, jak jsou data rozdělena. Můžeme vyčíst nejvyšší a nejnižší hodnoty, rozpětí dat, neboli rozložení do úzkého nebo širokého rozpětí.

Tabulka č. 1 – Studijní výsledky Magdy a Bena

Předmět	**Magdiny výsledky**	**Benovy výsledky**
Matematika	4	2
Český jazyk	1	1
Angličtina	1	2
Španělský jazyk	1	2
Biologie	3	2
Zeměpis	4	1
Chemie	4	2
Fyzika	1	2

Rozpětí vypočítáme tak, že u příslušného souboru dat odečteme nejnižší hodnotu od nejvyšší hodnoty.



Obrázek č. 2 – Graf k tabulce č. 1

V běžném životě se například setkáváme s uváděním maximální rychlosti internetového připojení, například 300 Mb/s. Ovšem tato informace může být zkreslující. Lepší představu o přesnější rychlosti, kterou můžeme očekávat, nabízí průměrná rychlost. Chceme-li si ovšem udělat ucelený obraz, potřebujeme rozpětí a rozložení dat (Vordermanová 2015).

1.3. Prezentace dat

1.3.1. Tabulka dat

Tabulka dat je nejjednodušší způsob, jak prezentovat statistická data. Informace jsou seskupeny do tříd, což názorněji odhaluje trendy, které data charakterizují. Tabulku lze použít pak ke konstrukci grafů. Statistická tabulka vedle absolutních četností obsahuje relativní četnosti a kumulativní četnosti (Vordermanová 2015).

Tabulka č. 2 – Tabulka dat

třída	četnost
Třída 1	4
Třída 2	8
Třída 3	6
Třída 4	4
Třída 5	5

1.3.2. Četnost, Relativní Četnost, Kumulované Četnosti

Při statistickém šetření se zpravidla vyšetřuje řada znaků, které nás zajímají jak každý zvlášť, tak i ve vzájemném vztahu. Ve většině šetření je počet různých hodnot sledovaného znaku menší než počet jednotek tohoto statistického souboru tzn., že několik různých statistických jednotek stejného souboru nabývá stejných hodnot znaku. Například dva lidé mohou mít stejný věk nebo stejnou výšku. Výsledkem šetření je seznam jednotek s udanou hodnotou znaku u každé z nich. Jsou-li jednotky v seznamu očíslovány *1, 2, 3,…., n*, pak jim odpovídající hodnoty znaku x označíme $x_1, x_2, x_3, ..., x_n$. Pokud znak *x* nabývá určitého počtu *r* různých hodnot, označujeme je $x_1^*, x_2^*, x_3^*, ..., x_r^*$. Pro každou možnou hodnotu x_j^* pak zjistíme, kolikrát se vyskytla mezi $x_1, x_2, x_3, ..., x_n$.

Hovoříme o tzv. četnosti hodnot sledovaného znaku. **Absolutní četnost** nebo jen četnost hodnoty x_j je počet výskytů této hodnoty ve sledovaném souboru. Četnost hodnoty x_j označujeme n_j, kde $j \epsilon \{1, 2, 3, \ldots, k\}$.

Součet četností všech možných *n* hodnot znaku se rovná počtu všech jednotek souboru.

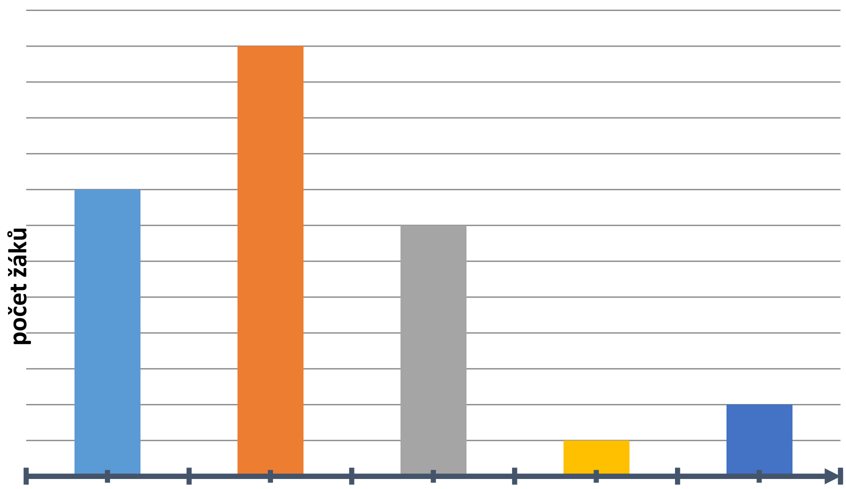
$$n_1 + n_2 + n_3 \ldots + n_k = \sum_{j=1}^{k} n_j = n$$

Příklad 1

Ve třídě 8. B je celkem 30 žáků. V následující tabulce máme zaznamenáno, jakou známku dostalo kolik žáků na vysvědčení z matematiky.

Tabulka č. 3 – Tabulka dat k příkladu 1

známka	počet žáků
Výborně	8
Chvalitebně	12
Dobře	7
Dostatečně	1
Nedostatečně	2

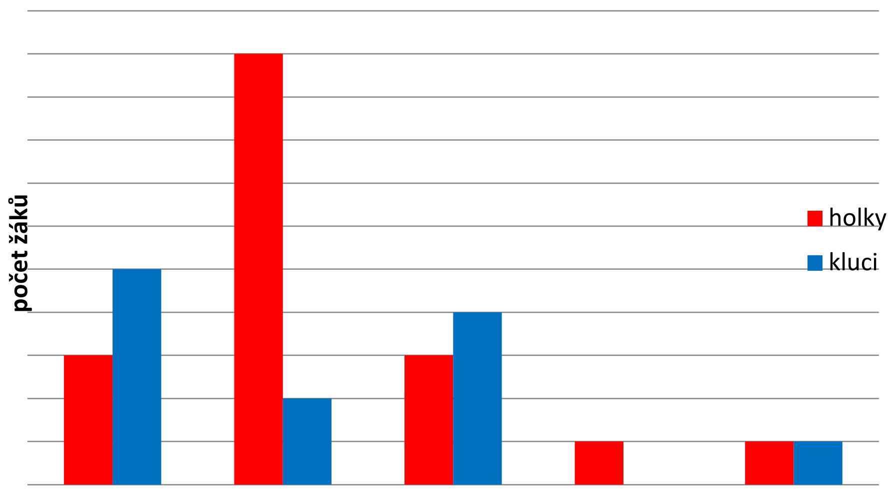


Obrázek č. 3 – Znázornění příkladu 1 ve sloupcovém grafu

Tyto četnosti můžeme dále rozdělit do podskupin. Například rozdělení známek mezi dívky a kluky. Nyní potřebujeme již dva sloupce. Tabulka zobrazuje skupinové rozložení četnosti.

Tabulka č. 4 – Skupinové rozložení četnosti.

známka	Počet dívek	Počet kluků
Výborně	3	5
Chvalitebně	10	2
Dobře	3	4
Dostatečně	1	0
Nedostatečně	1	1



Obrázek č. 4 – Skupinové rozložení četností ve sloupcovém grafu

Pokud nás nezajímá počet žáků, kteří jakou známku dostali, ale chceme vědět procentuální zastoupení známek žáků, poslouží nám **relativní četnost**.

Příklad 2

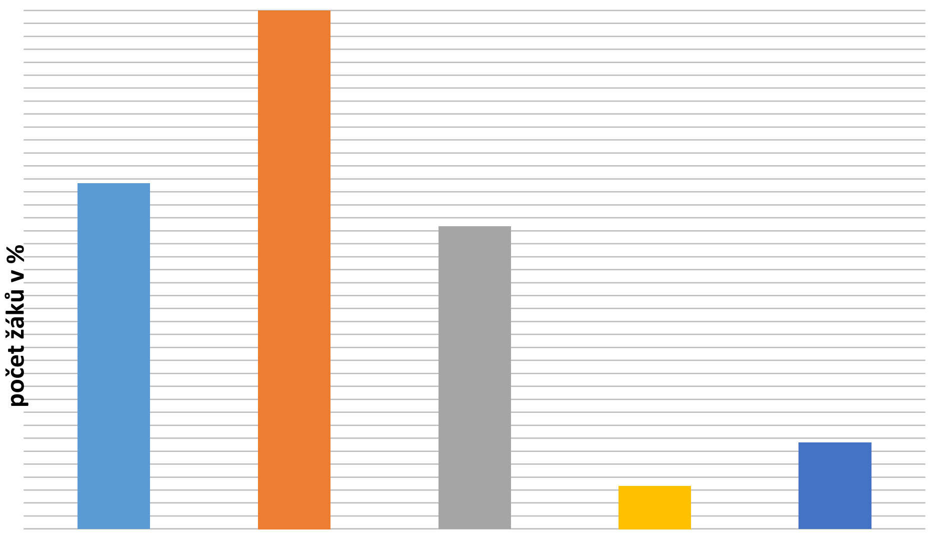
Máme 30 žáků. Počet označíme; $n=30$. Relativní četnost dostaneme tak, že absolutní četnost vydělíme n. Z předchozí tabulky vidíme, že známku chvalitebně dostalo 12 žáků. Relativní četnost tedy vypočítáme jako $12 \div 30$. Výsledkem relativní četnosti je 0,4. Pokud chceme pro lepší představivost výsledek v procentech, výsledek vynásobíme 100, tedy 40%.

Jednotlivé výsledky opět můžeme zapsat do tabulky.

Tabulka č. 5 – Tabulka relativních četností

Známka	Absolutní četnost	Relativní četnost	Relativní četnost v %
Výborně	8	0,2666	26,66%
Chvalitebně	12	0,4	40%
Dobře	7	0,2333	23,33%
Dostatečně	1	0,0333	3,33%
Nedostatečně	2	0,0666	6,66%

Nebo zakreslit grafem.



Obrázek č. 5 – Sloupcový graf relativních četností

Kumulativní četnost udává úhrnou četnost jednotivých vzestupně uspořádaných hodnot statistického znaku ve statistickém souboru.

Na prvním řádku je komulativní četnost rovna absolutní četnosti a na posledním řádku musí být komulativní četnost shodná s celkovým počtem.

Ve druhém řádku jsme kumulativní četnost 20 získali tak, že jsme sečetli absolutní četnost prvního a druhého řádku, tedy 8+12. Ve třetím řádku je kumulativní četnost 27, tedy 8+12+7, atd.

Tabulka č. 6 – Tabulka kumulativních četností

známka	Počet žáků	Kumulativní četnost
Výborně	8	8
Chvalitebně	12	20
Dobře	7	27
Dostatečně	1	28
Nedostatečně	2	30

Kumulativní četnost lze vyjádřit absolutně i relativně. Opět v prvním řádku musíme mít kumulativní relativní četnost shodnou s relativní četností a v poslední řádku musí být kumulativní relativní četnost rovna 1, tedy 100%.

Kumulativní relativní četnost můžeme vypočítat následovně.

$$P_1 = p_1$$
$$P_2 = p_1 + p_2$$
$$\ldots$$
$$P_K = \sum_{i=1}^{K} p_i = 1$$

kde P_i je kumulativní relativní četnost a K je počet obměn znaku, v našem případě K=5 (matematika 2014).

Tabulka č. 7 – Tabulka kumulativních relativních četností

známka	**Relativní četnost**	**Kumulativní relativní četnost**
Výborně	0,266	0,266
Chvalitebně	0,4	0,666
Dobře	0,233	0,899
Dostatečně	0,033	0,932
Nedostatečně	0,066	1

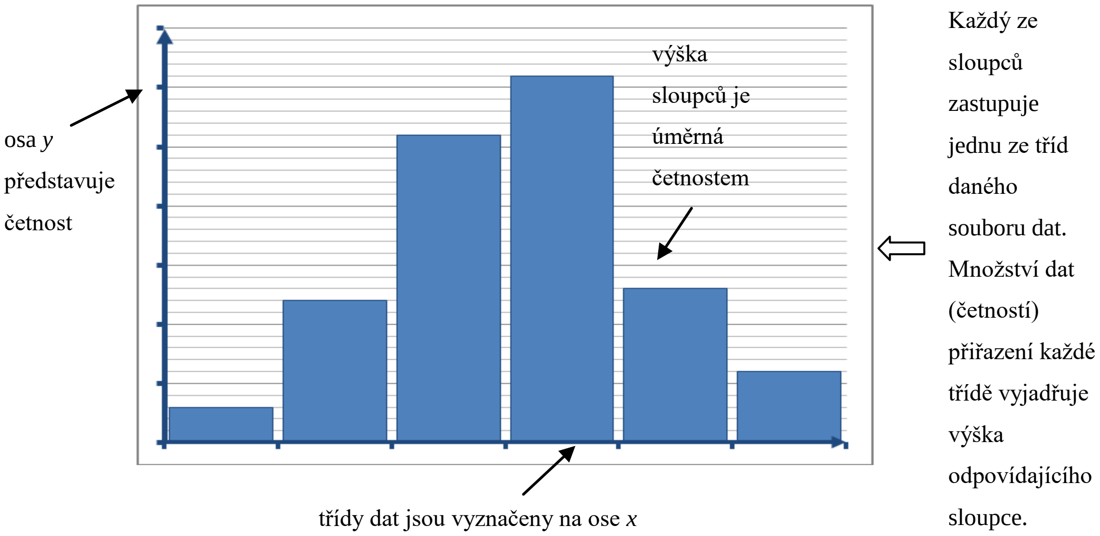
Nyní si můžeme do jedné tabulky shrnout všechny četnosti.

Tabulka č. 8 – Souhrnná tabulka četností

	Četnost		***Kumulativní četnost***	
známka	*absolutně*	*Relativně v %*	*absolutně*	*relativně v %*
Výborně	*8*	*26,6%*	*8*	*26,6%*
Chvalitebně	*12*	*40%*	*20*	*66,6%*
Dobře	*7*	*23,3%*	*27*	*89,9%*
Dostatečně	*1*	*3,3%*	*28*	*93,2%*
Nedostatečně	*2*	*6,6%*	*30*	*100%*
Celkem	*30*	*100%*		

1.3.3. Sloupcové grafy

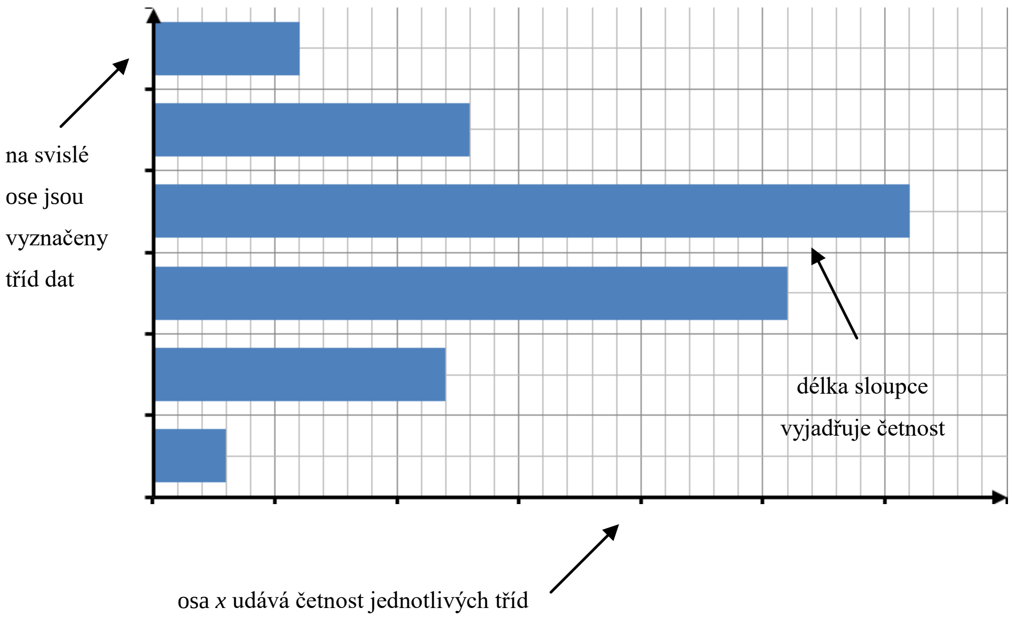
Další metodou prezentace dat jsou sloupcové grafy, které zobrazují soubor dat graficky. Statistický graf představuje snadněji pochopitelný a přehlednější způsob prezentace dat než tabulka. Sloupce neboli pruhy různých délek odpovídají četnostem jednotlivých tříd souboru. Pruhy mohou být nakresleny svisle i vodorovně. Sloupcový graf poskytuje rychlý přehled o poměrech jednotlivých hodnot. Na první pohled vypadá jako histogram, ale vlastnosti jsou jiné. Při sestrojování sloupcového grafu je důležité zvolit si vhodné měřítko. Poté si pojmenujeme osy podle sloupců tabulky a označíme stupnicemi odpovídajícími datům v tabulce. Informace získáme jasně a rychle z výšek sloupců. Přesnější hodnotu pak odečteme ze svislé osy grafu.



Obrázek č. 6 – Sloupcový graf svislý

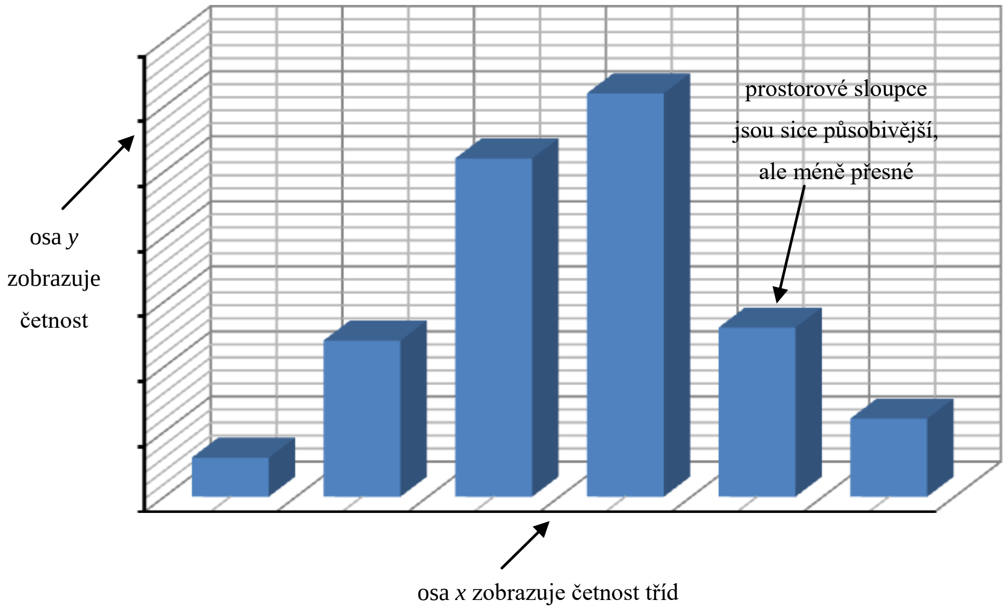
Sloupcový graf může mít několik podob. Sloupce lze zobrazit vodorovně, prostorově, ve skupinách po dvou nebo jako skládaný sloupcový graf. V každém druhu velikost sloupce odpovídá četnosti uvedené pro příslušnou třídu dat.

U tohoto grafu, obr. 7, jsou sloupce zobrazeny vodorovně. Významy os jsou prohozeny. Četnosti pro každou třídu dat lze nyní odečítat na ose *x*.



Obrázek č. 7 – Sloupcový graf vodorovný

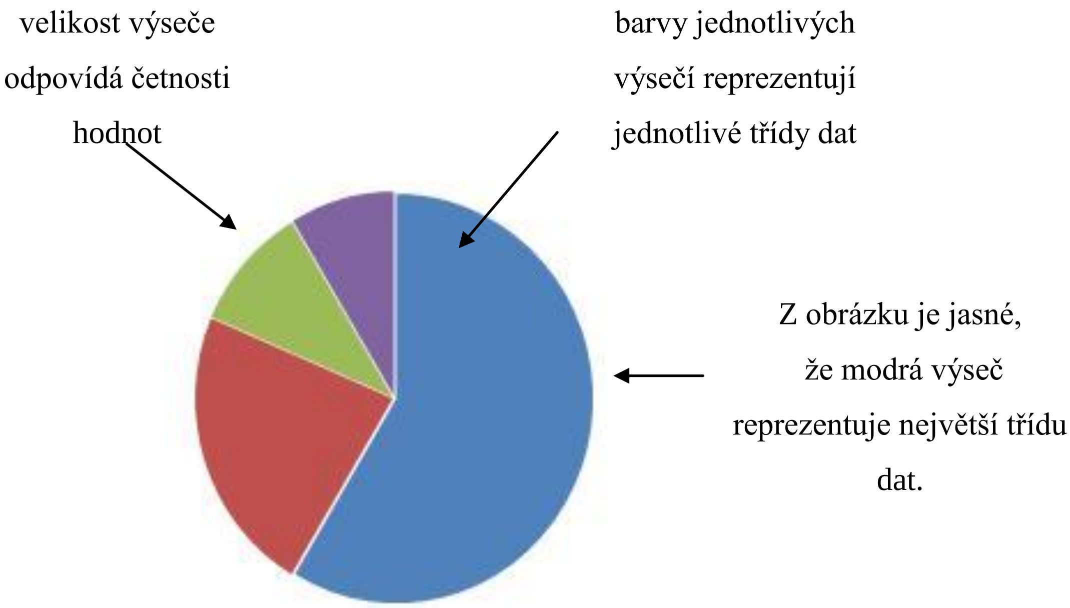
Prostorový neboli trojrozměrný sloupcový graf, obr. 8, je sice vizuálně působivější, ale může být zavádějící. Díky perspektivně se zdá, že horní plochy sloupců vyjadřují dvě hodnoty četnosti. Avšak skutečnou četnost ukazuje přední plocha sloupce (Vordermanová 2015).



Obrázek č. 8 – Prostorový neboli trojrozměrný sloupcový graf

1.3.4. Výsečové grafy

Výsečové grafy, můžeme se setkat s názvem kruhový či koláčový graf, znázorňují data pomocí kruhových výsečí, kde každá výseč reprezentuje příslušnou třídu souboru dat. Výsečové grafy jsou velmi často používány pro svůj okamžitý vizuální účinek. Velikost každé z výsečí jasně prozrazuje poměrné velikosti jednotlivých tříd statistického souboru, což znamená, že údaje lze porovnat snadno a rychle.



Obrázek č. 9 – Výsečový graf

Velikost úhlů jednotlivých výsečí spočítáme následujícím vztahem.

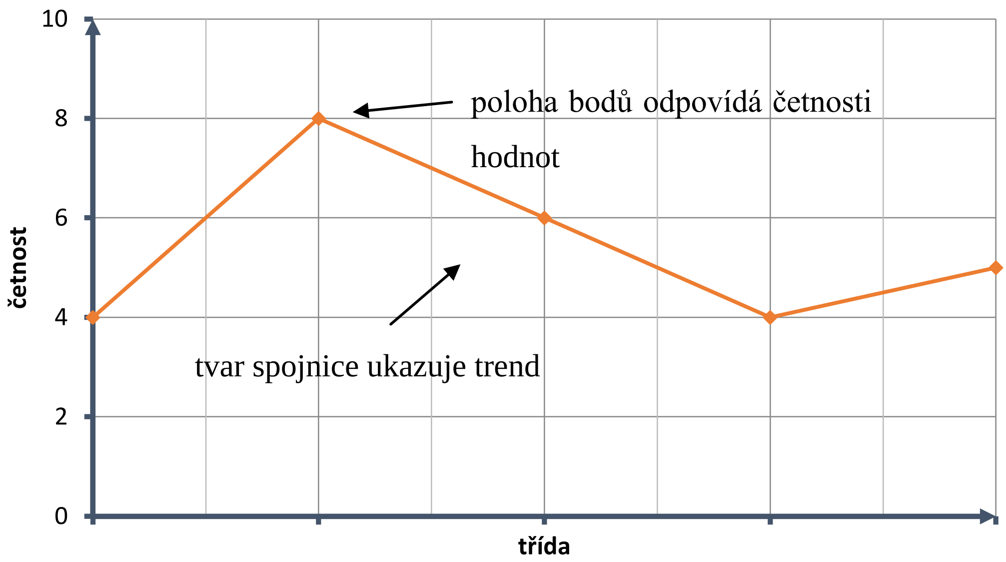
$$úhel = \frac{četnost\ třídy}{celková\ četnost} \times 360°$$

Abychom získali všechny informace potřebné k výpočtu velikosti úhlu, sestavíme si tabulku četností, kde nezapomeneme na celkovou četnost.

Plocha kruhu představuje celý soubor. 360° představuje 100 % a kruhová výseč, která zaujímá 1 % celku má středový úhel 3,6°(Vordermanová 2015).

1.3.5. Spojnicové grafy

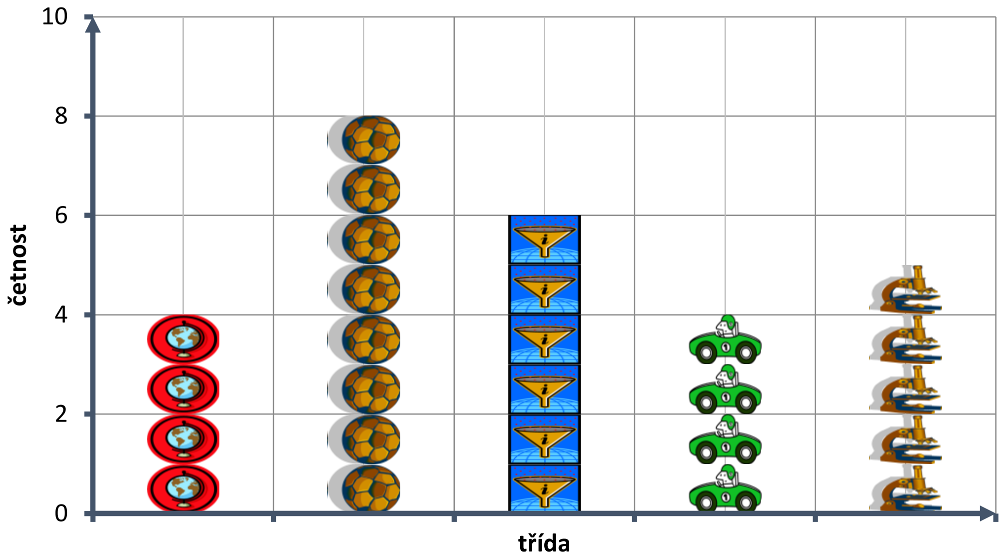
Spojnicový graf znázorňuje data lomenou čárou tvořenou úsečkami. Spojnicový graf je jedním ze způsobů přesné a přehledné prezentace dat. Využívá se zejména k znázornění dat v časovém období. Třídy jsou na grafu vyznačeny na ose x a četnost na ose y. Jednotlivé body představují četnosti jednotlivých tříd a jednotlivé úsečky představují trendy (Vordermanová 2015).



Obrázek č. 10 – Spojnicový graf nebo-li polygon

1.3.6. Obrázkový graf – piktogram

Piktogram představuje nejjednodušší typ sloupcového grafu. Každý obrázek reprezentuje množství dat (Vordermanová 2015).



Obrázek č. 11 – Obrázkový graf - piktogram

2. Statistické veličiny

2.1 Statistický soubor, jednotka a znak

Statistické údaje neboli statistická data jsou údaje, které jsou typické pro statistiku, o společenských, přírodních, technických a jiných skutečnostech, o tzv. statistických jevech. Tyto jevy vždy zkoumá na dostatečně rozsáhlém souboru případů a hledá vlastnosti jevů, které se projevují v souboru případů, nikoli na jednom případě. Statistický soubor tedy je soubor osob, věcí, událostí, jevů, časových období apod., jehož jednotlivé prvky se nazývají statistické jednotky. Rozsah souboru je potom počet všech prvků statistického souboru, který obvykle označujeme písmenem *n*.

Příklad užití statistiky ve škole

Třídní učitelka chce zjistit počet zameškaných hodin u každého z 22 žáků třídy 9. C.

• Statistickým souborem je množina všech žáků této třídy. Jejich společnou vlastností je, že jsou žáky uvedené třídy.

- Statistický znak je „počet zmeškaných hodin“
- Zameškané hodiny žáků jiných tříd třídní učitelku nezajímají.
- Hodnota znaku je číselný údaj, počet zameškaných hodin.

• Statistickou jednotkou je každý žák této třídy.

• Rozsah souboru je počet žáků třídy, tedy $n = 22$.

Statistický znak je vyšetřovaná vlastnost všech statistických jednotek tvořících statistický soubor. Statistické jednotky vyšetřujeme z hlediska zvoleného znaku nebo několika zvolených znaků. Znak může být kvantitativní nebo kvalitativní.

Hodnoty kvantitativního znaku lze vyjádřit číslem, například výška postavy, měsíční příjem, hmotnost, známky. Je-li něco popsáno číslem, ještě neznamená, že se jedná o kvantitativní znak. Například tramvaje jsou očíslovány, ale jedná se pouze o číselné označení linek.

Hodnoty kvalitativního znaku lze vyjádřit slovním popisem, například druh nemoci, pohlaví, národnost. Speciálním případem jsou kvalitativní znaky, které mohou nabývat pouze dvou variant, ty nazýváme alternativní znaky, například pohlaví – muž/žena, výsledek přijímací zkoušky – přijat/nepřijat. Kvalitativní znak můžeme dále rozdělit na spojitý a nespojitý znak.

Spojitý znak může nabývat jakékoli reálné hodnoty v jistém intervalu, který je omezen nejmenší a největší hodnotou, například výška člověka, hmotnost jednoho rohlíku, čas v běhu na 60 m. Hodnot, kterých znak může nabývat, je teoreticky nekonečně mnoho. Nespojitý neboli diskrétní znak může nabývat jen některých hodnot, například počet bodů z testu, konfekční velikost oděvu, počet dětí v rodině.

U každé jednotky zjistíme a zaznamenáme hodnotu znaku, což jsou jednotlivé údaje znaku. Hodnoty znaku musí být stanoveny tak, aby byly navzájem neslučitelné a jedna z nich vždy musela nastat. Hodnotu znaku obvykle označujeme x_j (Embden 2014).

Příklad 1

Při určování výšky vybrané skupiny lidí je statistickým znakem výška člověka, hodnotou znaku je číselná hodnota, která vyjadřuje příslušnou výšku.

Tabulka č. 9 – Tabulka dat k příkladu 1

statistická jednotka

statistický soubor

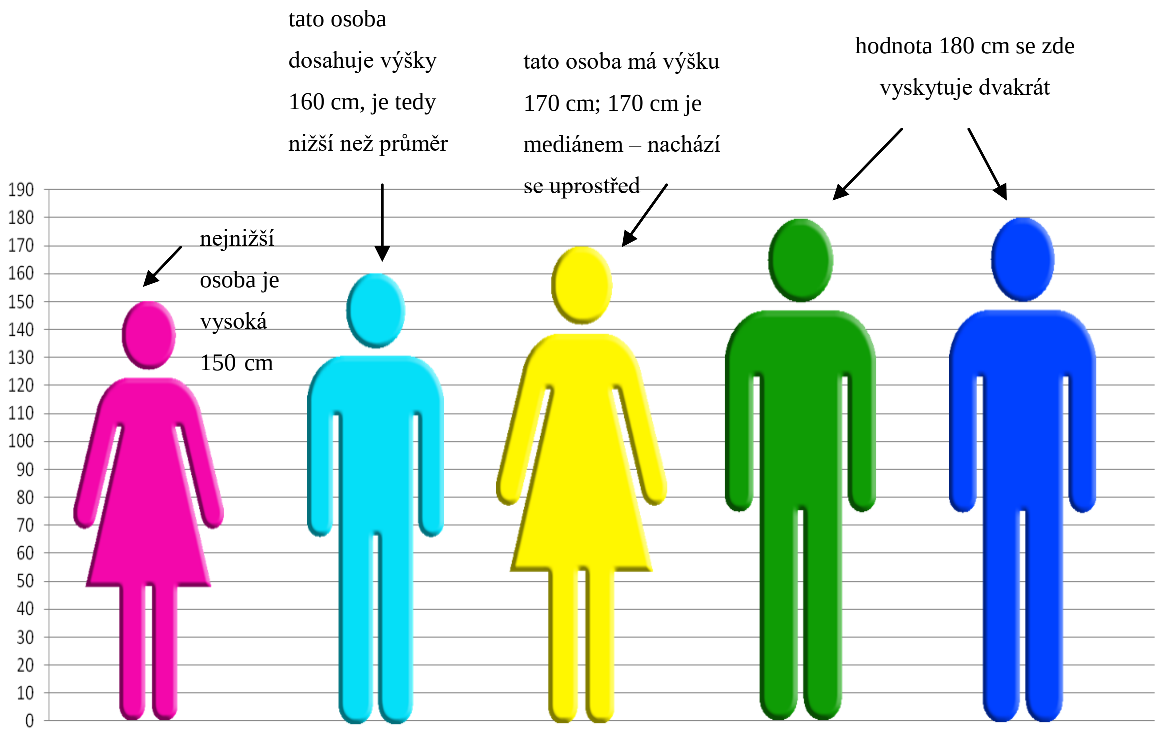
statistický znak

hodnota znaku

Osoba	X_j – výška [cm]
Tatínek	180
Maminka	170
Dcera	150
Syn	160

3. Charakteristika polohy

Charakteristiky polohy vyjadřují určitou střední hodnotu, nejčastěji aritmetický průměr, modus a medián. Vybraný ukazatel by měl co nejvýstižněji charakterizovat polohu četností v daném souboru. Slouží k porovnávání dvou i více souborů (Hazardní hry 2007).



Obrázek č. 12 – Charakteristiky polohy

3.1. Průměry hodnot kvantitativních znaků

Průměr představuje „střední“ hodnotu souboru metrických dat. Jedná se o typickou hodnotu, která reprezentuje celý soubor. Ve statistice slovo průměr uslyšíte nejčastěji. Máme několik druhů průměrů, například aritmetický průměr, geometrický průměr, vážený průměr a harmonický průměr. Každý z nich nám poskytuje o datech mírně odlišnou informaci. V některých případech jsou vhodné doplňkové průměrné hodnoty – modus a medián. V běžném životě se s ním setkáváme dnes a denně, především se jím rozumí aritmetický průměr sledovaných hodnot. Průměrné hodnoty znaku se též nazývají charakteristiky polohy znaku. (Hazardní hry 2007).

3.1.1. Aritmetický průměr

Jedná se o nejčastěji užívanou charakteristiku polohy znaku x. S aritmetickým průměrem se setkáváme v běžném životě nejčastěji. Dalo by se říci, že téměř denně. Zajímáme se o průměrný plat, průměrný počet lidí, průměrná výška, váha, průměrný prodej nejrůznějších výrobků a děti ve škole si vypočítávají průměrnou známku z daného předmětu. Aritmetický průměr se značí $\overline{x}$ a je dán vzorcem:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$průměr = \frac{součet\ všech\ hodnot}{celkový\ počet\ hodnot}$$

Součet zjištěných hodnot znaku u všech jednotek souboru vydělený počtem všech jednotek souboru (Polák 2014; Calda a Dupač 1993).

Příklad 1

Vypišme si výšky osob z obr…Dostaneme následující čísla: 150, 160, 170, 180 a 180. Zjistěte aritmetický průměr jejich výšek.

Řešení

1. Seřaďte všechny hodnoty podle velikosti a zjistěte počet hodnot ve výčtu.

150, 160, 170, 180, 180 ve výčtu je pět čísel

2. Sečtěte všechny hodnoty v tomto souboru.

150 + 160 + 170 + 180 + 180 = **840** součet všech hodnot

3. Vydělte tento součet počtem hodnot.

$840 \div 5 = 168$

$\overline{x} = 168$

Aritmetický průměr výšek osob je 168 cm.

Počítáme-li aritmetický průměr z tabulky rozdělení četností, musíme každou hodnotu x_j^* vynásobit její četností.

$$\overline{x} = \frac{1}{n}\sum_{j=1}^{r} x_j^* n_j$$

Příklad 2

Údaje z obrázku č. 11 si přepíšeme do následující tabulky. Nyní spočítáme opět aritmetický průměr použitím vzorce pro aritmetický průměr.

Osoba	**Maruška**	**Janička**	**Adam**	**Honzík**	**Martin**
výška	**150**	**160**	**170**	**180**	**180**

Tabulka č. 10 – Zadání příkladu 2

1. Vzorec pro aritmetický průměr

$$\overline{x} = \frac{1}{n}\sum_{j=1}^{r} x_j^* n_j$$

2. Nyní do něj dosadíme

výška 180 cm se vyskytuje dvakrát

$$\overline{x} = \frac{1\times150 + 1\times160 + 1\times170 + 2\times180}{5}$$

ve výčtu je pět hodnot

3. $\overline{x} = 168$
4. Průměrná výška osob je 168 cm.

3.1.2. Geometrický průměr

Geometrický průměr podává věrohodnější výsledky u tzv. růstových veličin. Ty se používají zejména v národohospodářství například, když chceme zjistit průměrné tempo růstu HDP[1] nebo cen určitého zboží za několik let zpětně. Geometrický průměr pro soubor s kladnými daty vyjde menší nebo roven průměru aritmetickému. Rovnost nastane pouze pro $x_1 = x_2 = ... x_n$.

Geometrický průměr, pro soubor $x_1 = x_2 = ... x_n$ kladných dat, se vypočte tak, že se všechny hodnoty vynásobí a poté odmocní. Takže geometrický průměr vypočteme takto

$$\overline{x}_G = \sqrt[n]{x_1 \cdot x_2 \cdot .... \cdot x_n} = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}$$

Všechny hodnoty mezi sebou vynásobíme a uděláme n-tou odmocninu (Polák 2014; Calda a Dupač 1993).

Příklad 1

Hrubý domácí produkt v ČR za posledních pět let vyvíjel následovně:

Tabulka č. 11 – Změna HDP

Rok	**Změna HDP oproti minulému roku v %**
2012	+1
2013	-1
2014	+2
2015	+4,3
2016	+2,3

(Křovák a kol. 2008).

[1] HDP – hrubý domácí produkt

Řešení

1. Každý rok je jako základna k následujícímu roku, například rok 2013 je 100% a v roce 2014 HDP vzrostlo o 2%, tzn., že jsme na 102% minulého roku (102% = 1,02)
2. Spočtěme vzrůsty HDP oproti výchozímu roku a dostaneme hodnoty 1,01; 0,99; 1,02; 1,043; 1,023
3. Nyní dosadíme do vzorce pro výpočet geometrického průměru

$$\bar{x} = \sqrt[5]{1,01 \cdot 0,99 \cdot 1,02 \cdot 1,043 \cdot 1,023}$$
$$\bar{x} = \sqrt[5]{1,0882}$$
$$\bar{x} = 1,017$$

$$1,017 \cdot 100 = 101,7\ \%$$
$$101,7\% - 100\% = \underline{\underline{1,7\%}}$$

4. Můžeme tedy říci, že HDP vzrostlo o 8,82% , vypočteno jako $\left[(1,0882 \cdot 100) - 100\right]$ a průměrný roční růst HDP je o 1,7%.
5. Odpověď: Hrubý domácí produkt v ČR rostl průměrným ročním tempem 1,7%.

Příklad 2

Cena kakaa za 1 kg se v letech 2013 – 2016 vyvíjela následovně. První rok stoupla o 20, 16 %, druhý rok klesla o 31, 38% a třetí rok ještě klesla o 17, 77%. Vypočítejte průměrný roční růst cen kakaa (Křovák a kol. 2008).

Řešení:

$$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot .... \cdot x_n}$$
$$\bar{x} = \sqrt[3]{1,2016 \cdot 0,6862 \cdot 0,8223}$$
$$\bar{x} = \sqrt[3]{0,678}$$
$$\bar{x} \doteq 0,8785$$

Průměrný pokles ceny kakaa za poslední tři roky……. (0,678·100)-100 = 32,2%

Průměrný roční pokles ceny kakaa……………………. (0,8785·100)-100 = 12,15%

Zde vidíme, že cena kakaa za poslední tři roky klesla o 32,2% a průměrný roční pokles je 12,15%.

3.1.3. Harmonický průměr

Harmonický průměr kladných hodnot statistického souboru je definován jako převrácená hodnota aritmetického průměru převrácených hodnot, tzn. podíl počtu členů a součtu převrácených hodnot zadaných členů.

$$\bar{x}_H = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1}+\frac{1}{x_2}+\ldots+\frac{1}{x_n}\right)} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$$

Harmonický průměr se využívá k zjištění průměrné délky času potřebné k provedení nějakého úkonu, který provádí současně několik osob či strojů. Například při výpočtu průměrné rychlosti na úsecích stejné délky.

Výsledek harmonického průměru je vždy menší nebo roven geometrickému průměru (Polák 2014; Calda a Dupač 1993).

Příklad 1

Traktor orá pole, které je mírně do kopce. Zdola nahoru jede průměrnou rychlostí 10 km/h. Z kopce má rychlost o 5km/h vyšší. Jaká je průměrná rychlost traktoru?

$$\bar{x}_H = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1}+\frac{1}{x_2}+\ldots+\frac{1}{x_n}\right)}$$

$$\bar{x}_H = \frac{1}{\frac{1}{2}\left(\frac{1}{10}+\frac{1}{15}\right)}$$

$$\bar{x}_H = \frac{1}{\frac{1}{2}\left(\frac{3+2}{30}\right)}$$

$$\bar{x}_H = \frac{1}{\frac{1}{2}\left(\frac{5}{30}\right)}$$

$$\bar{x}_H = 1 \div \frac{1}{12}$$

$$\bar{x}_H = \frac{1}{1}\cdot\frac{12}{1}$$

$$\bar{x}_H = 12 km/h$$

Traktor jede průměrnou rychlostí 12km/h.

Příklad 2

Tři kamarádky se chtějí nechat učesat na maturitní ples. Všechny tři si nechávají dělat stejný účes. První kadeřnici účes trval 1,5 hod, druhé kadeřnici 2 hodiny a třetí kadeřnici 2,5 hod. Jaký průměrný čas zabrala kadeřnici tvorba maturitního účesu?

$$\bar{x}_H = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1}+\frac{1}{x_2}+\ldots+\frac{1}{x_n}\right)}$$

$$\bar{x}_H = \frac{1}{\frac{1}{3}\left(\frac{1}{90}+\frac{1}{120}+\frac{1}{150}\right)}$$

$$\bar{x}_H = \frac{1}{\frac{1}{3}\left(\frac{20+15+12}{1800}\right)}$$

$$\bar{x}_H = \frac{1}{\frac{1}{3}\left(\frac{47}{1800}\right)}$$

$$\bar{x}_H = 1 \div \frac{47}{5400}$$

$$\bar{x}_H = \frac{1}{1}\cdot\frac{5400}{47}$$

$$\bar{x}_H \doteq 115\,\text{min} = 1\,\text{hod}\,55\,\text{min}$$

Kadeřnicím by maturitní účes trval průměrně 1hod 55min.

3.1.4. Klouzavý průměr

Klouzavý průměr vyjadřuje trendy v datech za určité časové období.

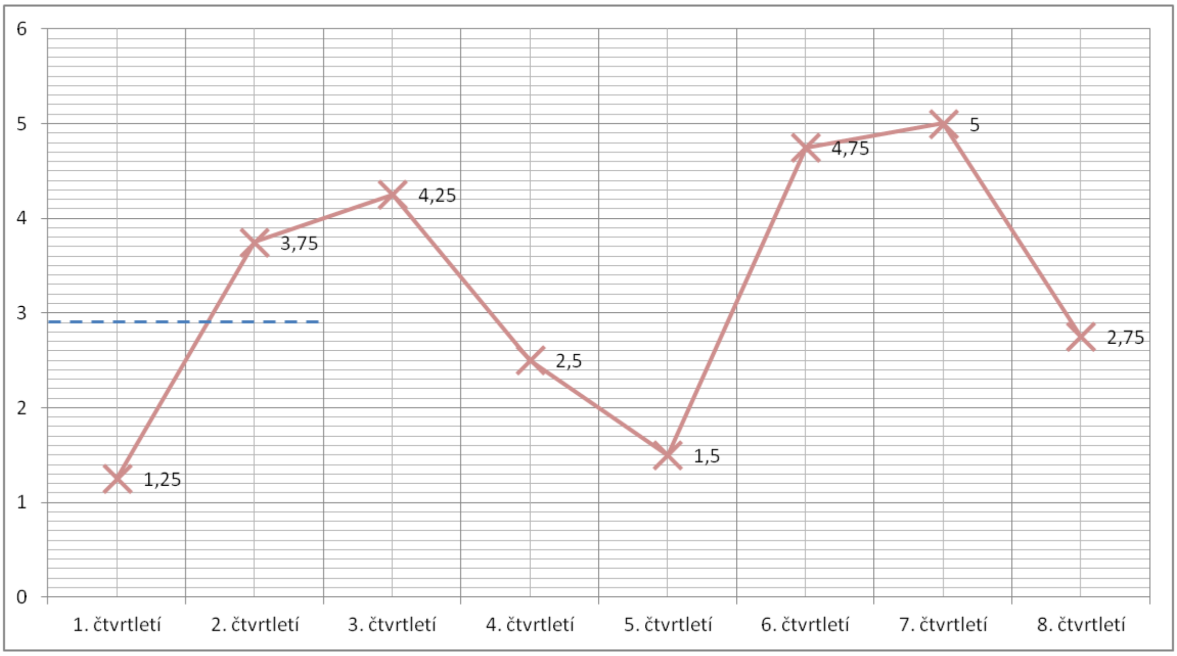
Tabulka zachycuje prodej zmrzliny za období dvou let. Každý rok je rozdělen do čtvrtletí (Polák 2014; Calda a Dupač 1993).

Tabulka č. 12 – Prodej zmrzliny

	První rok				Druhý rok			
čtvrtletí	1.	2.	3.	4.	5.	6.	7.	8.
Prodej v tisících kusů	1,25	3,75	4,25	2,5	1,5	4,75	5,0	2,75

Tabulku zobrazíme jako spojnicový graf, kde na ose y je prodej v tisících kusů a na ose x doba ve čtvrtletích.

Graf prodeje ukazuje prudké výkyvy (červená křivka) a žlutá křivka představuje klouzavý průměr, který vyjadřuje trend za dva roky (Vordermanová 2015).



Obrázek č. 13 – Znázornění klouzavého průměru

Klouzavý průměr vypočteme následujícím způsobem. Z údajů v tabulce vypočteme aritmetický průměr pro každé čtvrtletí.

$$\bar{x} = \frac{1,5 + 4,75 + 5 + 2,75}{4}$$
$$\bar{x} = \frac{14}{4}$$
$$\bar{x} = 3,5$$

Průměr za 1. – 4. čtvrtletí

$$\bar{x} = \frac{1,25 + 3,75 + 4,25 + 2,25}{4}$$
$$\bar{x} = \frac{11,75}{4}$$
$$\bar{x} \doteq 2,94$$

Průměrná hodnota za první čtvrtletí je 2,94. Tuto hodnotu zaneseme do grafu na osu y.

Průměr za 2. – 5. čtvrtletí

$$\bar{x} = \frac{3,75 + 4,25 + 2,5 + 1,5}{4}$$
$$\bar{x} = \frac{12}{4}$$
$$\bar{x} = 3$$

Průměr za 3. – 6. čtvrtletí

$$\bar{x} = \frac{4,25 + 2,5 + 1,5 + 4,75}{4}$$
$$\bar{x} = \frac{13}{4}$$
$$\bar{x} = 3,25$$

Hodnotu opět zaneseme do grafu na osu y.

Průměr za 4. – 7. čtvrtletí

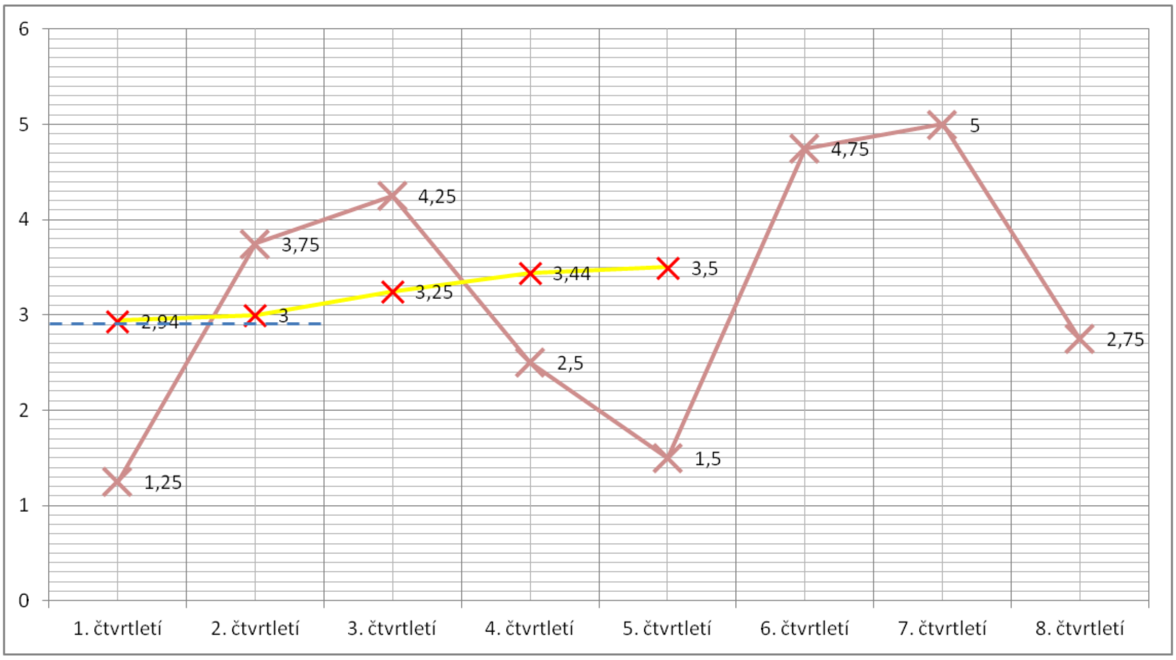
$$\bar{x} = \frac{2,5 + 1,5 + 4,75 + 5}{4}$$
$$\bar{x} = \frac{13,75}{4}$$
$$\bar{x} \doteq 3,44$$

Průměr za 5. – 8. čtvrtletí

$$\bar{x} = \frac{1,5 + 4,75 + 5 + 2,75}{4}$$

$$\bar{x} = \frac{14}{4}$$

$$\bar{x} = 3,5$$



Obrázek č. 14 – Znázornění klouzavého průměru

3.2. Modus

Modus je další charakteristika polohy, která se používá pro číselná i nečíselná data, u kterých nelze spočítat průměr.

Modus je hodnota znaku, která má v souboru největší zastoupení, neboli má v souboru největší četnost. Značíme jej *Mod(x)* (Polák 2014; Calda a Dupač 1993).

Příklad 1

Pro příklad použijeme tabulku *č.3*

Tabulka č. 13 – Tabulka pro příklad 1

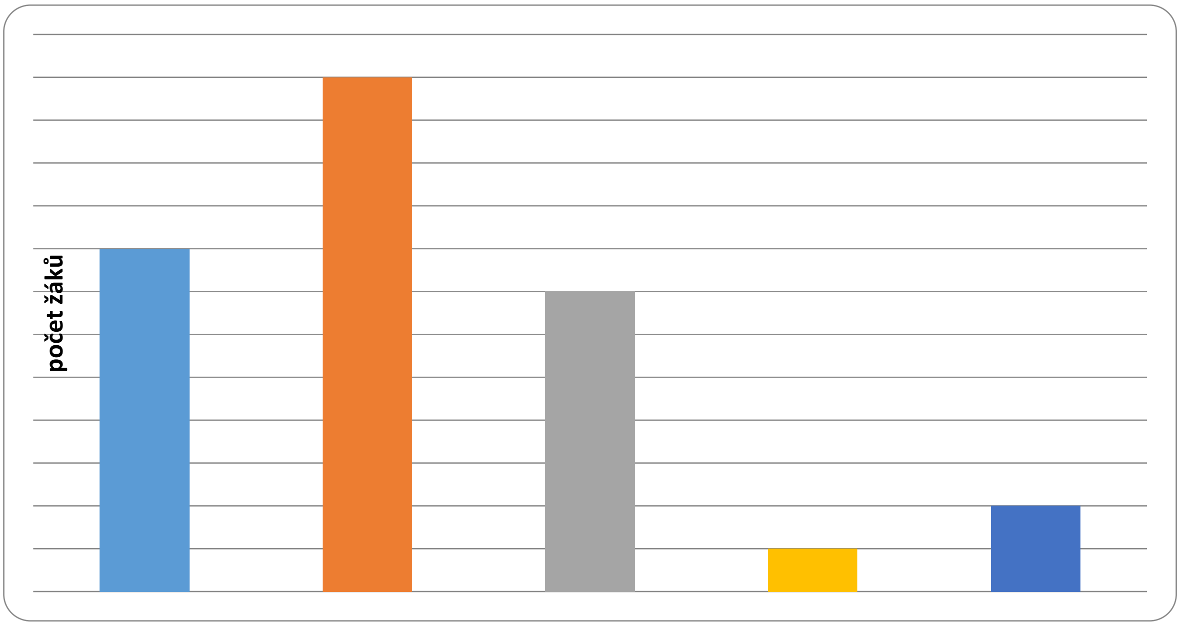
známka	počet žáků
Výborně	8
Chvalitebně	12
Dobře	7
Dostatečně	1
Nedostatečně	2

Tabulka zachycuje známky žáků z daného předmětu. Jaký je modus známek v této třídě?

Modus je hodnota znaku s největší četností. V této třídě je nejvíce zastoupena známka chvalitebně, kterou dostalo 12 žáků.

Mod(x) = chvalitebně

Ve sloupcovém grafu se modus ukáže jednoznačně, a to jako nejvyšší sloupec.



Obrázek č. 15 – Grafické znázornění příkladu 1

Se stejnou četností se může vyskytnout více hodnot, tzn., že modus nemusí být vždy jednoznačně určen.

Příklad 2

Ve třídě děti odpovídaly, kolik PC mají v domácnosti. Odpovědi jsme zaznamenali do neúplné tabulky.

Tabulka č. 14 – Tabulka pro příklad 2

Počet PC v domácnosti	0	1	2	3	4
Počet odpovědí	5	x	x	1	0

Průměrný počet PC v domácnosti je 1,2. Modus ovšem nelze jednoznačně určit, byl by roven 1 nebo 2.

Dále se můžeme zeptat, kolik dětí bylo celkově dotazováno?

Počet dotazovaných dětí vypočteme jednoduchou rovnicí.

$$\frac{0\cdot 5+1\cdot x+2\cdot x+3\cdot 1+4\cdot 0}{6+x+x}=1,2$$

$$\frac{3+3x}{6+2x}=1,2$$

$$0,6x=4,2$$

$$x=7$$

$$x_{(1)}\le x_{(2)}\le \ldots \le x_{(n)}$$

$$\bar{x}=\frac{2+2}{2}$$

$$\bar{x}=7,4$$

Vypočítané četnosti můžeme dopsat do tabulky a určit počet dotazovaných dětí.

Tabulka č. 15 – Výsledek příkladu 2

Počet PC v domácnosti	0	1	2	3	4
Počet odpovědí	5	7	7	1	0

Bylo dotazováno 20 dětí.

Dle definice můžeme modus určit i takových souborů, které neobsahují číselné údaje.

Příklad 3

Mějme soubor s těmito prvky oblíbenosti barev:

růžová, bílá, černá, červená, růžová, růžová, modrá, červená, růžová, bílá

Určete modus.

Tabulka č. 16 – Postup příkladu 3

Barva	bílá	růžová	černá	modrá	červená
Četnost	2	4	1	1	2

Mod(x) = růžová

3.3. Medián

Medián dělí soubor podle velikostí seřazených číselných hodnot znaku na dvě poloviny. Je to tedy číslo, které se nachází přesně uprostřed číselné řady.

Medián je prostřední hodnota znaku, pokud máme soubor, v němž jsou hodnoty znaku uspořádány vzestupně, platí $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$.

Medián značíme Med(x).

Mylně si mnoho lidí myslí, že aritmetický průměr je číslo, které je přesně uprostřed souboru dat, prostřední číslo je ovšem medián.

Výhoda mediánu je, že není ovlivněn extrémními hodnotami. Proto se často používá tam, kde aritmetický průměr nedává vhodné výsledky (Polák 2014; Horenský a kol. 2015; Calda a Dupač 1993).

Příklad 1

Mějme soubor hodnot: 2, 2, 2, 2, 29

Med(x) = 2

Stejně tak Mod(x) = 2.

$\bar{x} = 7,4$

Medián je zde vhodnějším ukazatelem než aritmetický průměr.

Má-li soubor lichý počet prvků, je mediánem prostřední hodnota.

Příklad 2

Využijeme opět tabulku známek č. 3.

Tabulka č. 17 – Zadání příkladu 2

známka	počet žáků
Výborně	8
Chvalitebně	12
Dobře	7
Dostatečně	1
Nedostatečně	1

Určete medián daného souboru.

Pro určení mediánu je potřeba si všechny známky vypsat a seřadit.

1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 5

Medián je prostřední hodnota.

medián

Med(x) = 2

Pokud má soubor sudý počet prvků, mediánem je aritmetický průměr prostředních dvou hodnot.

Příklad 3

Tabulka č. 18 – Zadání příkladu 3

známka	počet žáků
Výborně	8
Chvalitebně	12
Dobře	7
Dostatečně	1
Nedostatečně	2

Opět je potřeba známky vypsat a seřadit.

1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 5,5

Nyní máme dvě prostřední hodnoty.

Uděláme jejich aritmetický průměr.

$$\bar{x} = \frac{2+2}{2}$$
$$\bar{x} = 2$$

V tomto případě je medián stále 2.

Med (x) = 2.

Při určování mediánu z tabulky je potřeba zohlednit četnosti jednotlivých hodnot. Častou chybou bývá, že je za medián považována hodnota v prostředním sloupci tabulky.

Tabulka č. 19 – Zadání příkladu 3

známka	počet žáků
Výborně	8
Chvalitebně	12
Dobře	7
Dostatečně	1
Nedostatečně	2

Známka dobře je sice prostřední sloupec tabulky, ale neurčuje střední hodnotu souboru. Musíme zohlednit četnost jednotlivých známek (Polák 2014; Horenský a kol. 2015; Calda a Dupač 1993).

4. Charakteristiky variability

Charakteristiku polohy označujeme číslo, kolem kterého jednotlivé hodnoty znaku kolísají, např. aritmetický průměr, modus a medián. Velikost tohoto kolísání vyjadřují charakteristiky proměnlivosti znaku. Například kdybychom zkusili zvážit několik sáčků mouky, která má vážit 1kg, zjistili bychom, že každý sáček váží jinak, mají odchylku od určité hodnoty (Horenský a kol. 2015; Hazardní hry 2007).

4.1. Rozpětí

Jediná charakteristika variability, která se neváže k žádné charakteristice polohy. K určení rozpětí nám postačí znát mezní hodnoty znaku – maximum a minimum. Rozpětí je tedy rozdíl mezi nejvyšší a nejnižší hodnotou. Používá se nejen v matematice, ale i mezi hudebníky nebo v managementu – rozpětí not, hlasu, věkové rozpětí, kolísání tělesné teploty a počet pracovníků podřízených jednomu manažerovi. Variační rozpětí samo neposkytuje dostatek informací o vyšetřovaném souboru. Pokud se v souboru nacházejí extrémní hodnoty, je tato charakteristiky nevhodná. V popisné statistice málo používána.

Máme-li n hodnot znaku (x_1, x_2, ..., x_n), které uspořádáme vzestupně, pak variační rozpětí, značeno R, je rozdíl největší a nejmenší hodnoty znaku, tedy

$$R = x_{\max} - x_{\min}$$

(Horenský a kol. 2015).

Příklad 1

Byla zkoumána výdrž baterie u nových mobilních telefonů stejného typu a značky. Byly naměřeny tyto hodnoty v hodinách:

8, 24, 72, 9, 20, 35, 18

$$x_{\min} = 8$$
$$x_{\max} = 72$$
$$R = 72 - 8$$
$$R = 64 hodin$$

4.1.1. Kvartilové rozpětí

Kvartilové rozpětí je definováno jako rozdíl mezi horním a dolním kvartilem a vymezuje střední polovinu hodnot dat. Řadí se mezi percentilové rozpětí. Jinými slovy se jedná o rozdíl mezi hodnotou třetího a prvního kvartilu neboli mezi 75. a 25. percentilem. Kvartilové rozpětí udává šířku intervalu, ve kterém leží 50% prostředních hodnot uspořádaného souboru. Značí se R_Q (Horenský a kol. 2015; Polák 2014).

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2$$

$$s_x^2 = \frac{(30,20-30,07)^2 + (30,00-30,07)^2 + (29,90-30,07)^2 + \ldots + (30,15-30,07)^2 + (30,30-30,07)^2}{10}$$

$$s_x^2 = \frac{0,2789}{10}$$

$$s_x^2 = 0,02789$$

$$sx = 0,167$$

$$R_Q = x_{0,75} - x_{0,25}$$

nebo

$$R_Q = Q_3 - Q_1$$

4.2. Odchylka

4.2.1. Absolutní odchylka

Absolutní odchylku značíme d a její hodnota je vždy nezáporné číslo. Geometricky se jedná o vzdálenost hodnoty od hodnoty $\bar{x}$. Absolutní odchylka hodnoty x_j od aritmetického průměru $\bar{x}$ je rovna absolutní hodnotě rozdílu těchto hodnot, platí tedy:

$$d_j = |x_j - \bar{x}|$$

(Horenský a kol. 2015).

Příklad 1

Dlouhodobá průměrná teplota vzduchu na Šumavě v lednu je -15 °C. Maximální teplota, která zde byla v lednu naměřena bylo 3,5°C v roce 2008 a minimální teplota byla -34,8°C v roce 2012. Jaká je absolutní odchylka těchto teplot?

Absolutní odchylka maximální naměřené teploty od průměru:

$$d_j = |3,5 - (-15)|$$
$$d_j = 18,5°C$$

Absolutní odchylka minimální naměřené teploty od průměru:

$$d_j = |-34,8 - (-15)|$$
$$d_j = 19,8°C$$

Příklad 2

Dvě studentky se učily na maturitu. Počet hodin, který strávily učením během týdne je následující.

První studentka – 5, 2, 4, 10, 1, 3, 3

Druhá studentka – 5, 5, 3, 4, 3, 4, 4

Vypočítejte absolutní odchylky počtu hodin strávených studiem a z vypočítaných údajů vyvoďte závěr týkající se způsobu přípravy obou dívek na maturitu.

Vypočítáme-li aritmetický průměr přípravy každé studentky, zjistíme, že je stejný, a to 4 hodiny.

Tabulka č. 20 – Absolutní odchylky studentek

První studentka

$\bar{x} = 4$

Počet hodin	Odchylka		Absolutní odchylka
5	1		1
2	-2		2
4	0		0
10	6		6
1	3		3
3	-1		1
3	-1		1

Druhá studentka

$\bar{x} = 4$

Počet hodin	Odchylka	Absolutní odchylka
5	1	1
5	1	1
3	-1	1
4	0	0
3	-1	1
4	0	0
4	0	0

Obě dívky se sice učily v průměru stejný počet hodin, ovšem kolísavost okolo aritmetického průměru je rozdílná. Zatímco první studentka se od průměrné doby odchyluje až o 6 hodin, druhá studentka maximálně o 1 hodinu. Druhá studentka se učí stabilněji a pravidelněji.

4.2.2. Průměrná absolutní odchylka

Pokud bychom měli velké množství naměřených hodnot, bylo by obtížné a náročné vypočítávat, porovnávat a vyhodnotit takové množství vypočtených odchylek. Nedokázali bychom ani popsat celý soubor jednou číselnou hodnotou. Z toho důvodu zavádíme průměrnou absolutní odchylku, kterou značíme $\bar{d}$.

Průměrná absolutní odchylka je aritmetický průměr absolutních odchylek hodnot znaku všech prvků daného souboru od aritmetického průměru, platí tedy:

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n}\left|x_j - \bar{x}\right|$$

(Horenský a kol. 2015).

4.2.3. Relativní odchylka

Relativní odchylka také používaná v praxi nám určuje podíl průměrné absolutní odchylky a příslušného aritmetického průměru. Vyjadřuje se většinou v procentech. Značíme ji r. Jedná se o číslo, které nemá jednotku a tak nám umožňuje porovnání souborů, v nichž mají hodnoty znaků různé jednotky, např. měření v milimetrech a metrech.

Platí tedy:

$$r = \frac{\overline{d}}{\overline{x}}$$

Ve fyzice se pro relativní odchylku používá termín relativní chyba. Relativní chyba měření vyjadřuje přesnost měření. Laboratorní měření ve školních podmínkách považujeme za přesné, pokud je relativní chyba menší než 1% (Horenský a kol. 2015).

4.2.4. Kvartilová odchylka

Kvartilová odchylka udává průměrnou vzdálenost mezi dvěma kvartily. Jinými slovy je definovaná jako polovina kvartilového rozpětí. Značí se *Q* (Horenský a kol. 2015)

$$Q = \frac{R_Q}{2}$$

4.2.5. Směrodatná odchylka

Směrodatná odchylka představuje nejčastější charakteristiku variability. V praxi se s ní setkáváme častěji než s rozptylem a je definována jako kvadratický průměr odchylek hodnot znaku od jejich aritmetického průměru. Směrodatná odchylka je vyjádřena ve stejných jednotkách jako sledovaný znak, např. pokud jsou metrická data v jednotkách *m*, pak výsledek směrodatné odchylky bude v jednotkách m^2. Jednoduše řečeno je směrodatná odchylka druhá odmocnina z rozptylu. Značíme ji s_x (Horenský a kol. 2015; Calda a Dupač 1993).

$$s_x = \sqrt{s_x^2} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2}$$

4.3. Rozptyl

Rozptyl nám ukazuje, jak se hodnoty pohybují okolo střední hodnoty, jak jsou okolo rozptýleny. Patří mezi nejpoužívanější charakteristiky variability. Rozptyl se používá pro srovnávání dvou a více souborů dat. Rozptyl je definován jako aritmetický průměr druhých mocnin odchylek hodnot znaku od aritmetického průměru. Pro rozptyl použijeme tento vzorec, pokud máme neroztříděná data, nemáme je roztřízeny do intervalů.

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2$$

Rozptyl i směrodatná odchylka vyjadřují, jak jsou jednotlivé hodnoty vyrovnané, jak jsou vzdálené od průměru. Druhou odmocninu používáme, protože rozptyl nemá stejnou jednotku jako měřená veličina (Calda a Dupač 1993; Horenský a kol. 2015).

Příklad

Studenti při hodině fyziky vážily olovněnou kuličku. Měření zopakovali desetkrát a dostali tyto navážené hodnoty v gramech:

měření	výsledek v gramech
1	30, 20
2	30, 00
3	29, 90
4	30, 20
5	30, 10
6	30, 00
7	30, 05
8	29,80
9	30, 15
10	30, 30

Vypočtěte směrodatnou odchylku a rozptyl.

1. Nejprve si vypočteme aritmetický průměr

$$\bar{x} = \frac{30,20 + 30,00 + 29,90 + 30,20 + 30,10 + 30,00 + 30,05 + 29,80 + 30,15 + 30,30}{10}$$

$$\bar{x} = \frac{300,7}{10}$$

$$\bar{x} = 30,07$$

2. Nyní spočítáme rozptyl

$$s_x^2 = \frac{\overbrace{(30,20-30,07)^2}^{(x_1-\bar{x})^2} + (30,00-30,07)^2 + (29,90-30,07)^2 + \ldots + \overbrace{(30,15-30,07)^2}^{(x_9-\bar{x})^2} + (30,30-30,07)^2}{\underbrace{10}_{\text{ve výčtu máme deset hodnot – deset měření}}}$$

$$s_x^2 = \frac{0,206}{10}$$

$$s_x^2 = 0,0206$$

3. Rozptyl je tedy 0, 0206
4. Směrodatnou odchylku vypočteme jako druhou odmocninu rozptylu

$$s_x = \sqrt{0,0206}$$

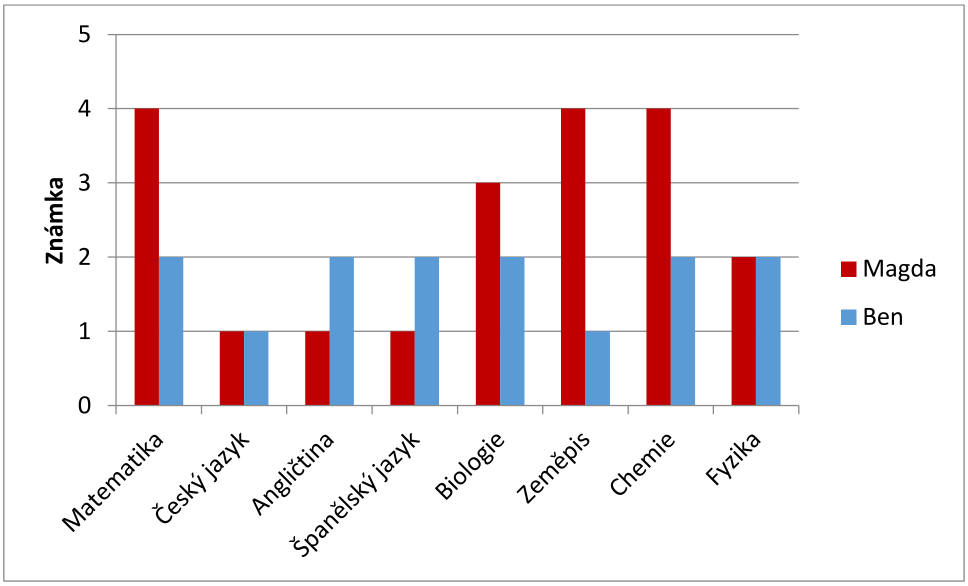
$$s_x \doteq 0,14$$

Směrodatná odchylka je 0,14 g^2.

Příklad 2

Pro následující příklad použijeme tabulku č. 1.

Ben a Magda dostaly na vysvědčení následující známky, které máme zaznamenány v grafu.



Obrázek č. 16 – Zadání příkladu 2

Vypočtete aritmetický průměr, rozptyl a směrodatnou odchylku.

1.

Aritmetický průměr známek Magdy

$$\bar{x} = \frac{3\times4+3+2+3\times1}{8}$$

$$\bar{x} = \frac{20}{8}$$

$$\bar{x} = 2,5$$

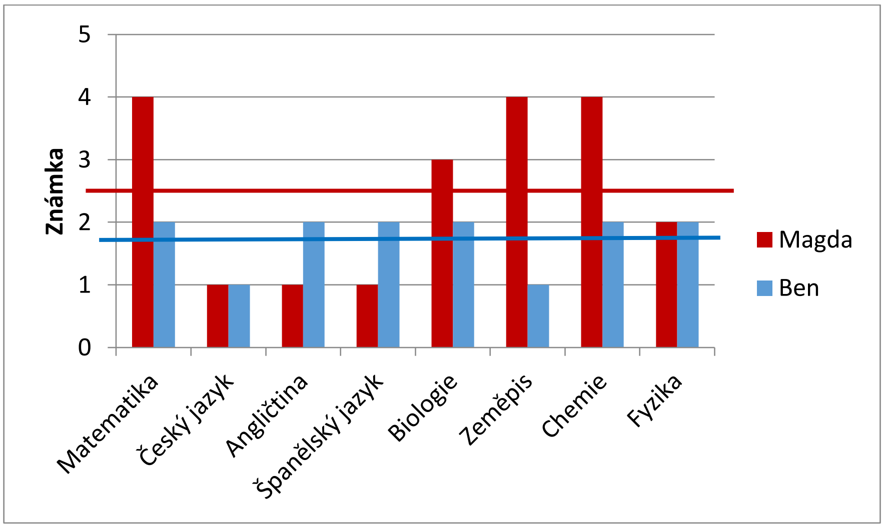
Aritmetický průměr známek Bena

$$\bar{x} = \frac{6\times2+2\times1}{8}$$

$$\bar{x} = \frac{14}{8}$$

$$\bar{x} = 1,75$$

V následujícím grafu vidíme, že Benovy výsledky se pohybují okolo aritmetického průměru, kdežto Magdiny jsou více vzdálené o aritmetického průměru jejích známek.



Obrázek č. 17 – Znázornění aritmetických průměrů v grafu

2. Nyní spočítáme rozptyl.

Magda

$$s_x^2 = \frac{(4-2,5)^2 + (1-2,5)^2 + (1-2,5)^2 + (1-2,5)^2 + (3-2,5)^2 + (4-2,5)^2 + (4-2,5)^2 + (2-2,5)^2}{8}$$

$$s_x^2 = \frac{14}{8}$$

$$s_x^2 \doteq 1,75$$

Ben

$$s_x^2 = \frac{(2-1,75)^2 + (1-1,75)^2 + (2-1,75)^2 + (2-1,75)^2 + (2-1,75)^2 + (1-1,75)^2 + (2-1,75)^2 + (2-1,75)^2}{8}$$

$$s_x^2 = \frac{1,5}{8}$$

$$s_x^2 \doteq 0,19$$

Zde vidíme, že rozptyl u Magdy je větší než jedna, kdežto u Bena je téměř zanedbatelný. U Magdy jsou známky více rozptýlené okolo aritmetického průměru.

3. Na závěr spočítáme směrodatnou odchylku.

Magda

$$s_x = \sqrt{1,75}$$

$$s_x \doteq 1,32$$

Ben

$$s_x = \sqrt{0,19}$$

$$s_x \doteq 0,44$$

Závěr

Cílem této bakalářské práce bylo sestavit ucelený a přehledný učební materiál pro žáky 2. stupně základní školy a studenty střední školy ze základů statistiky.

Začátkem jsme si definovali pojmy spjaté se statistikou, vysvětlili jsme si co to statistika je, kde se s ní setkáváme a sběr a prezentaci dat. Poté následovalo vysvětlení jednotlivých pojmů s řešeným příkladem. Při sestavování učebního materiálu bylo vycházeno z programu RVP pro střední školy.

Seznámili jsme se s tématy, jako je aritmetický průměr, modus, medián, rozptyl nebo směrodatná odchylka.

Danou prací bych se ráda dále zabývala ve své diplomové práci, kde bych ji rozvinula o didaktický materiál pracovních listů s příklady ze statistiky propojené s občanskou výchovou, jak pro 2. stupeň základní školy, tak pro střední školy.

Seznam použité literatury

CALDA, Emil a Václav DUPAČ. Matematika pro gymnázia: Kombinatorika, pravděpodobnost, statistika. Praha: Prometheus s.r.o., 1993.

ČESKÝ STATISTICKÝ ÚŘAD. Školní statistická ročenka 2008. Praha: Scientia, 2008

EMDEN, Helmut van. Statistik ohne Albträume: Ene Einführung für Biowissenschaftler. Weinheim: Wiley-VCH, 2014.

GIBILISCO, Stan. Statistika bez předchozích znalostí: Průvodce pro samouky. Brno: Computer Press, 2007.

HORENSKÝ, Radek, a kol. Matematika pro střední školy 8. díl: Kombinatorika, pravděpodobnost, statistika. Brno: Didaktis, 2015.

PAVELKA, Jindřich. Hazardní hry [online]. 2007-2017 [cit. 2017-6-20].

Dostupné z: http://www.hazardni-hry.eu/statistika/

POLÁK, Josef. Didaktika matematiky: Jak učit matematiku zajímavě a užitečně. Plzeň: Fraus, 2014.

ŘEZANKOVÁ, H., L. MAREK a M. VRABEC. Interaktivní učebnice statistiky: Typy proměnných [online]. 2001 [cit.2017-05-22].

Dostupné z: http://iastat.vse.cz/typy_promennych.html

VORDERMANOVÁ, Carol. Matematika: Spolu to zvládneme. Praha: Slovart, 2015.

ZWERENZ, Karlheinz. Statistik: Einführung in die computergestützte dateanalyse. München: Oldenbourg, 2015.