



Pedagogická
fakulta
Faculty
of Education

Jihočeská univerzita
v Českých Budějovicích
University of South Bohemia
in České Budějovice

Jihočeská univerzita v Českých Budějovicích
Fakulta pedagogická
Katedra matematiky

Bakalářská práce

Sbírka příkladů ze statistiky pro střední školy

Vypracovala: Nováková Aneta
Vedoucí práce: doc. RNDr. Vladimíra Petrášková, Ph.D.

České Budějovice Rok 2021

JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH

Pedagogická fakulta

Akademický rok: 2019/2020

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Aneta NOVÁKOVÁ**
Osobní číslo: **P18857**
Studijní program: **B7507 Specializace v pedagogice**
Studijní obor: **Matematika se zaměřením na vzdělávání
Zeměpis se zaměřením na vzdělávání**
Téma práce: **Sbírka příkladů ze statistiky pro SŠ**
Zadávající katedra: **Katedra matematiky**

Zásady pro vypracování

Cílem bakalářské práce je sestavit sbírku řešených příkladů ze statistiky, která bude určena studentům středních škol. Obsahové zaměření sbírky bude následující:

- statistický soubor, jednotka, znak (nominální, ordinální, metrické),
- rozložení četností (absolutní, relativní, kumulativní absolutní, kumulativní relativní) a jeho grafické znázornění (polygon, histogram, ogiva),
- aritmetický, harmonický a geometrický průměr,
- modus, medián, kvartily,
- směrodatná odchylka, relativní odchylka, variační koeficient, kvartilová odchylka.

Práce bude obsahovat jednak řešené příklady, jednak neřešené příklady. Řešení příkladů bude s využitím tabulkového procesoru EXCEL.

Studentka:

- prostuduje RVP pro gymnázia a SOŠ, dále ŠVP vybrané školy gymnaziálního typu a vybrané SOŠ
- prostuduje vybrané učebnice určené pro SŠ, které jsou tematicky zaměřené na statistiku
- prostudování statistických funkcí v MS Excel
- sestaví sbírku příkladů ze statistiky, která bude obsahovat i její vlastní příklady

Rozsah pracovní zprávy: **dle dohody**

Rozsah grafických prací: **dle dohody**

Forma zpracování bakalářské práce: **tištěná**

Seznam doporučené literatury:

ČERMÁK, Pavel a Petra ČERVINKOVÁ. Odmaturuj! z matematiky: přehled středoškolského učiva. Vyd. 3. opr. Brno: Didaktis, 2004, 208 s. Odmaturuj!. ISBN 80-735-8014-4.

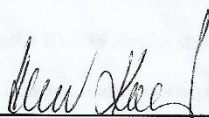
KUBEŠOVÁ, Naděžda a Eva CIBULKOVÁ. Matematika: přehled středoškolského učiva. 1. vyd. Třebíč: Petra Velanová, 2006, 239 s. Maturita (Petra Velanová). ISBN 80-868-7303-X.

PETRÁNEK, Oldřich, Emil CALDA a Petr HEBÁK. Matematika pro střední odborné školy a studijní obory středních odborných učilišť. 5. vyd. Praha: Prometheus, 1997, 148 s. Učebnice pro střední školy (Prometheus). ISBN 80-719-6040-3.

HINDLS, R., HRONOVÁ, SEGER a FISCHER. Statistika pro ekonomy. 8. vyd. Praha: Professional Publishing, 2007, 415 s. ISBN 978-80-86946-43-6.
CALDA, Emil a Václav DUPAČ. Matematika pro gymnázia. 5. vyd. Praha: Prometheus, 2008, 170 s. Učebnice pro střední školy (Prometheus). ISBN 978-80-7196-365-3.
MELICHAR, Jan a Josef SVOBODA. Slovní úlohy v učivu matematiky 1. Stupně základní školy: Poznámky k didaktice matematiky. Katedra matematiky Pedagogické fakulty Univerzity J. E. Purkyně v Ústí nad Labem, 2002 k nalezání na http://www.pf.ujep.cz/files/KMA_poznamkydidamat16.pdf
HRACH, Karel a Václav DUPAČ. Sběrka úloh ze statistiky. Vyd. 1. Ústí nad Labem: FSE UJEP, 2006, 65 s. Učebnice pro střední školy (Prometheus). ISBN 80-704-4845-8.
CALDA, E., DUPAČ, V. Kombinatorika, pravděpodobnost, statistika. Učebnice pro gymnázia. Prometheus, 1993.

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

Datum zadání bakalářské práce: **11. prosince 2019**
Termín odevzdání bakalářské práce: **30. dubna 2021**



doc. RNDr. Helena Koldová, Ph.D.
děkanka

LS.



doc. RNDr. Vladimíra Petrášková, Ph.D.
vedoucí katedry

Prohlášení

Prohlašuji, že svoji bakalářskou práci na téma Sbíрка příkladů ze statistiky pro střední školy 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 jejich 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

Tímto bych ráda poděkovala vedoucí své bakalářské práce doc. RNDr. Vladimíře Petráškové, Ph.D. za její trpělivost, odbornou pomoc a cenné rady, které mi pomohly ke zpracování bakalářské práce. Dále bych chtěla poděkovat mé rodině za trpělivost a podporu během celých studií.

Anotace

Cílem této bakalářské práce je sestavení sbírky příkladů ze statistiky pro studenty středních škol. Dále se tato práce zabývá vysvětlením základních pojmů souvisejících s danými příklady. Jsou zde také uvedeny příklady k procvičení, kde lze ověřit správné řešení na konci každé kapitoly.

Klíčová slova

statistika, četnost, charakteristiky polohy, charakteristiky variability, příklady k procvičení

Annotation

The goal of this bachelor thesis is to compile a collection of examples from statistics for secondary school's students. This thesis is also focused on explanation of basic related concepts with the given examples. There are also examples to exercise, where you can verify the correct solution at the end of each chapter.

Keywords

statistics, frequency, characteristics of position, characteristics of variability, examples to exercise

OBSAH

Úvod.....	11
Cíl a metodika	12
1. Statistický soubor, znak, jednotka	13
2. Rozdělení četností a jejich grafické znázornění	17
2.1. Absolutní četnost	17
2.2. Relativní četnost	17
2.3. Kumulativní absolutní četnost.....	18
2.4. Kumulativní relativní četnost	19
2.5. Třídní intervaly a tabulka četností.....	20
2.6. Grafické znázornění četností	22
2.6.1. Sloupcový diagram	22
2.6.2. Spojnicový diagram	22
2.6.3. Histogram.....	23
2.6.4. Kruhový diagram	24
3. Charakteristiky polohy.....	32
3.1. Aritmetický průměr	32
3.1.1. Vážený aritmetický průměr	33
3.1.2. Průměr počítaný z dat roztříděných intervalů.....	34
3.2. Geometrický průměr.....	35
3.3. Harmonický průměr	36
3.4. Modus	37
3.5. Medián.....	37
3.6. Kvantily	38
3.6.1. Percentily a decily.....	39
3.6.2. Tercily	39
3.6.3. Kvartily	39
4. Charakteristiky variability	47
4.1. Rozpětí.....	47
4.2. Mezikvartilové rozpětí	48
4.3. Mezikvartilová odchylka	50
4.4. Rozptyl	50
4.5. Směrodatná odchylka	52

4.6. Variční koeficient	52
5. Závěr	57
6. Literatura a zdroje:.....	58

SEZNAM TABULEK

Tabulka 1: Tabulka absolutních četností.....	17
Tabulka 2: Rozdělení absolutních a relativních četností.....	18
Tabulka 3: Kumulativní absolutní četnosti	19
Tabulka 4: Kumulativní relativní četnosti.....	20
Tabulka 5: Rozdělení četností	21
Tabulka 6: Rozdělení četností	24
Tabulka 7: Výsledné hodnoty objednaného jídla.....	25
Tabulka 8: Počty uchazečů na vysokou školu.....	26
Tabulka 9: Výsledky z čtvrtletní práce z 9.A.....	27
Tabulka 10: Výsledné hodnoty objednaného jídla.....	27
Tabulka 11: Počty studentů hlásící se na vysokou školu	29
Tabulka 12: Znázornění četností ušitých šatů	29
Tabulka 13: Infrastruktura silniční dopravy v ČR v roce 2019	31
Tabulka 14: Tabulka rozdělení četností výšky dětí.....	33
Tabulka 15: Znázornění četností ušitých šatů	37
Tabulka 16: Tabulka hodnot potřebných ke vzorovému příkladu	40
Tabulka 17: Výsledky z písemné práce z českého jazyka	41
Tabulka 18: Míra nezaměstnanosti v České republice.....	41
Tabulka 19: Výsledky SCIO testu.....	43
Tabulka 20: Počet zaměstnanců s průměrnou mzdou	43
Tabulka 21: Výchozí data k výpočtu rozpětí	48
Tabulka 22: Data potřebná pro výpočet mezikvartilového rozpětí.....	49
Tabulka 23: Počet prodaných zmrzlin ve stáncích.....	51
Tabulka 24: Data potřebná pro výpočet směrodatné odchylky.....	52
Tabulka 25: Data potřebná pro výpočet variačního koeficientu	53
Tabulka 26: Hodnoty potřebné pro výčtu k příkladu číslo 1	54

SEZNAM OBRÁZKŮ

Obrázek 1: Výsledky písemné práce z českého jazyka z kapitoly 2.1	22
Obrázek 2: Výsledky písemné práce z českého jazyka z kapitoly 2.1	23
Obrázek 3: Výška dětí na dětském táboře z kapitoly 2.5	23
Obrázek 4: Volba starosty	24
Obrázek 5: Graf počtu domácích mazlíčků.....	25
Obrázek 6: Výsledky z čtvrtletní práce z 9.A	27
Obrázek 7: Absolutní četnosti objednaného jídla	28
Obrázek 8: Relativní četnosti objednaného jídla	28
Obrázek 9: Znárodnění uchazečů z různých typů škol.....	29
Obrázek 10: Sloupcový diagram ušitých šatů	30
Obrázek 11: Polygon četností ušitých šatů	30
Obrázek 12: Infrastruktura silniční dopravy v ČR v roce 2019	31

Úvod

Pro svou závěrečnou bakalářskou práci jsem si zvolila téma Sbíрка příkladů pro střední školy. Toto téma jsem si vybrala z důvodu, že mě statistika velmi bavila a oslovila na střední škole, a tak jsem si díky bakalářské práci chtěla rozšířit své vědění o statistice. Dalším důvodem, proč jsem zvolila toto téma, byla výzva pro mě samotnou, vytvořit sbírku příkladů, kterou budou jednou počítat žáci středních škol.

Bakalářská práce se skládá ze čtyř kapitol, které obsahují teoretický základ, řešené příklady a příklady k procvičení.

V první kapitole dochází k vysvětlení základních pojmů, kterými jsou statistický soubor, statistický znak a statistická jednotka. S těmito pojmy se setkáváme v průběhu celé bakalářské práce.

Ve druhé kapitole se snažím vysvětlit rozdělení četností a jejich grafické znázornění. Tato kapitola se dále dělí na šest podkapitol, kdy v prvních pěti podkapitolách popisují jednotlivé četnosti. Poslední podkapitola vysvětluje grafické znázornění četností, které byly výše v kapitole popsány.

Ve třetí kapitole řeším problematiku charakteristik polohy. Tyto charakteristiky polohy nám ukazují polohu statistického znaku na číselné ose. Opět je tato kapitola rozdělena do šesti podkapitol. Každá podkapitola se zabývá jednotlivými charakteristikami polohy, kterými jsou aritmetický průměr, geometrický průměr, harmonický průměr, modus, medián a kvantily.

V poslední kapitole se zabývám charakteristikami variability. Charakteristiky variability nám popisují míru kolísání neboli proměnlivost hodnot ve statistickém znaku. Kapitola je rozdělena do šesti podkapitol, kdy každá podkapitola, jako je tomu ve třetí kapitole, popisuje jednotlivé charakteristiky variability. Mezi charakteristiky variability řadíme rozpětí, mezikvartilové rozpětí, mezikvartilovou odchylku, rozptyl, směrodatnou odchylku a variační koeficient.

Cíl a metodika

Cílem mé práce bylo vytvořit komplexní sbírku příkladů, která by svým obsahem byla zaměřena pro studenty středních škol. Sbíрка je koncipována následovně: teoretický základ dané problematiky, vzorový příklad a příklady k procvičení.

Při zpracování sbírky jsem vycházela jednak z Rámcového vzdělávacího programu pro střední odborné školy a Rámcově vzdělávacího programu pro gymnázia, jednak z učebnic určených pro střední vzdělávání. Z těchto učebnic jsem se nejvíce inspirovala učebnicí Kombinatorika, pravděpodobnost a statistika (Caldá, Dupač, 2008), která je určena pro gymnázia, a která dle mého mínění názorně vysvětluje danou problematiku. Při tvorbě příkladů jsem především vycházela z dat Českého statistického úřadu¹.

¹ <https://www.czso.cz/>

Praktická část – Sbírnka příkladů

1. Statistický soubor, znak, jednotka

Statistika jako obor se zaměřuje shromažďováním, porovnáváním a vyhodnocováním poznatků, které získala na dostatečně velkém souboru objektů. Za objekty můžeme považovat nejrůznější věci z běžného lidského života, např. automobily, mobilní telefony, oblečení a jiné, dále ze zástupců živé přírody pak brouky, savce a ryby. Rovněž sem spadají i předměty neživé přírody jako je hlína, stromy a řeky ale i lidé jakožto obyvatelstvo určitého etnika, zaměstnanci jedné firmy, studenti středních škol nebo obyvatelé mluvící daným jazykem. Cílem statistiky je ze získaných údajů sestavit ucelený obraz o celku, ze kterého soubor pochází a popsat zákonitosti tohoto celku (Robová a kol., 2013).

Než budeme pokračovat, shrneme si důležité pojmy ze základů statistiky, které potřebujeme znát do dalších kapitol.

Statistickým souborem nazýváme množinu předmětů nebo subjektů roztríděných z hlediska jejich určité společné vlastnosti zvané *znak*.

Za statistickou jednotku považujeme prvek souboru, kterého vybrané vlastnosti jsou předmětem zkoumaného statistického problému. Statistickou jednotkou může být člověk, událost, příroda, zvíře, věc, budova, území apod (Petránek, 2011).

Statistické jednotky zkoumáme z pohledu jednoho znaku nebo ze skupiny zvolených znaků. Například u jehličnatých stromů můžeme zkoumat stáří jednotlivých stromů a výšku stromů. V tomto případě zkoumáme dva znaky – stáří a výšku stromů statistického souboru – jehličnaté stromy. Statistickou jednotkou je jehličnatý strom.

Rozsah souboru – počet všech prvků statistického souboru, značíme jej n .

Zkoumané znaky dělíme na kvantitativní znaky a kvalitativní znaky. Za kvantitativní znaky považujeme znaky, které se vyjadřují číselnou velikostí, např. známka z testu, váha, výška, hrubá měsíční mzda apod. Hodnoty kvalitativního znaku udávají kvalitu, která je většinou zapsaná slovním popisem. Kvalitativní znaky se dále dělí na alternativní, které udávají opačné jevy, např. žena – muž, prospěl – neprospěl,

zdravý – nemocný a znaky které můžou mít více možností hodnot, např. barva očí, náboženství, druh stromu, národnost apod (Kubešová, Cibulková, 2011).

Vzorový příklad: Na pracovní pozici prodavačky do obchodního řetězce se přihlásila skupinka 22 lidí. Zaměstnavatel se na pohovoru ptal jednotlivých uchazečů na věk, vzdělání, délku praxe v oboru a flexibilitu práce. Získané informace si zapisoval do tabulky a na základě srovnání vybral 5 vhodných uchazečů na vybranou pozici. Určete, co je statistickým souborem, statistickou jednotkou a zkoumaným znakem? U znaku určete, zda jde o kvalitativní nebo kvantitativní znak.

Řešení: Statistickým souborem je skupinka 22 lidí, od kterých zaměstnavatel získává potřebné informace. Statistickými jednotkami jsou jednotliví lidé hlásící se na pozici prodavačky a statistické znaky jsou v našem případě čtyři – věk – kvantitativní znak, vzdělání – kvalitativní znak, délka praxe – kvantitativní znak a flexibilita – kvalitativní znak.

Příklady k procvičení

Příklad 1.: Firma vyrábí 173 plastových výrobků na jednom stroji za hodinu. Každých 15 minut vezme zaměstnanec náhodný nově vzniklý plastový výrobek a položí jej na váhu. Získané informace o váze zapisuje do archivů, které slouží ke zpětné kontrole nadřízeného. Napište, co by mohlo být statistickým souborem, statistickou jednotkou a statistickým znakem. U znaku určete, zda jde o kvalitativní nebo kvantitativní znak.

Příklad 2.: Studentka druhého ročníku stření školy sbírá knihy, ke každé knize si zapisuje do čtenářského deníku několik pro ni důležitých věcí. Ve svém čtenářském deníku má zapsaný žánr, počet stran, od koho knihu dostala dále charakterové vlastnosti hlavních postav a v neposlední řadě si zapisuje, jak se jí kniha líbila. Určete, co se považuje za statistický soubor, statistickou jednotku a statistický znak. U znaku určete, zda jde o kvalitativní nebo kvantitativní znak.

Příklad 3.: Roku 2011 se uskutečnilo doposud poslední sčítání lidu, domů a bytů. Zamyslete se, co by mohlo být statistickými jednotkami a statistickými znaky, když budeme brát, že sčítání lidu, domů a bytů je statistickým souborem. Uveďte minimálně dva příklady statistické jednotky, dva příklady kvantitativního statistického znaku a dva příklady kvalitativního statistického znaku.

Výsledky

Příklad 1.: Za statistický soubor považujeme 173 plastových výrobků. Statistickými jednotkami jsou jednotlivé plastové výrobky a kvantitativním statistickým znakem v tomto případě je váha plastových výrobků.

Příklad 2.: Statistickým souborem je celá sbírka knih, která není více specifikována. Statistickými jednotkami jsou jednotlivé knihy sbírky a statistických znaků máme pět, jedná se o žánr – kvalitativní znak, počet stran – kvantitativní znak, darující osobu – kvalitativní znak, charakterové vlastnosti – kvalitativní znak a libost/nelibost knihy – kvalitativní znak.

Příklad 3.: Statistickými jednotkami můžou být osoby, byty, domy nebo například domácnosti. Kvantitativním statistickým znakem můžeme považovat věk obyvatelstva a například počet členů v domácnosti. Mezi kvalitativní znaky patří v našem případě například pohlaví obyvatelstva, národnost lidu, vybavenost bytu apod.

2. Rozdělení četností a jejich grafické znázornění

V této kapitole se budeme zabývat četnostmi statistického znaku, který označíme písmenem x a který může nabývat hodnot x_1, x_2, \dots, x_n , kdy n je rovno rozsahu souboru. V mnoha případech se hodnoty znaku opakují, z čehož vyplývá, že počet rozdílných hodnot znaku k je menších než celkový počet jednotek souboru n , $k \leq n$ (Kubešová, Cibulková, 2011, str. 211).

2.1. Absolutní četnost

Absolutní četnost, v některých literaturách označována pouze jako četnost, nám ukazuje, kolikrát se hodnota x_j statistického znaku vyskytuje ve statistickém souboru. Značí se n_j a součet všech možných absolutních četností se rovná počtu všech jednotek souboru, $n_j = n$ (Calda, Dupač, 2008, str. 132-133).

Vzorový příklad: Ve třídě 7.A psali žáci písemnou práci z českého jazyka. Znamky, které žáci obdrželi, byli následující 2, 1, 3, 5, 1, 1, 4, 2, 2, 3, 3, 3, 4, 2, 1, 1, 1, 2, 3, 3, 1, 3, 3, 2, 3, 1, 3. Sestavte tabulku absolutních četností.

Řešení:

Tabulka 1: Tabulka absolutních četností

známka	1	2	3	4	5
absolutní četnost n_j	8	6	10	2	1

2.2. Relativní četnost

Relativní četnost značíme v_j a je dána podílem absolutní četnosti n_j hodnoty znaku x_j a rozsahem souboru n . Relativní četnost tedy můžeme zapsat jako: $v_j = \frac{n_j}{n}$ (Calda, Dupač, 2008, str. 133).

Zpravidla se udává v procentech a součet všech relativních četností se rovná číslu 1, eventuálně 100 %.

Vzorový příklad: Z údajů, na které můžete nahlédnout v tabulce 1, doplňte tabulku rozdělení četností o relativní četnost. Relativní četnost запиšte v % a zaokrouhlete na dvě desetinná místa.

Řešení: Nejprve si musíme vypočítat kolik žáků chodí celkem do třídy 7.A. Počet žáků vypočítáme součtem všech absolutních četností, tedy: $8 + 6 + 10 + 2 + 1 = 27$. Součtem absolutních četností jsme zjistili, že do třídy 7.A chodí 27 žáků, $n=27$. Relativní četnosti vypočítáme následujícím způsobem, například pro známku 1 počítáme $\frac{8}{27} \times 100$, z čehož nám vyjde, že relativní četnost pro známku 1 je 29,63 %. Analogicky vypočítáme zbylé četnosti a zapíšeme do tabulky 2. Pro naší zpětnou kontrolu si spočítáme součet relativních četností, který se musí rovnat 100 %. V našem případě tedy: $29,63 + 22,22 + 37,04 + 7,41 + 3,7 = 100$.

Tabulka 2: Rozdělení absolutních a relativních četností

známka	1	2	3	4	5	součet
absolutní četnost n_j	8	6	10	2	1	27
relativní četnost v_j v %	29,63	22,22	37,04	7,41	3,70	100 %

2.3. Kumulativní absolutní četnost

Kumulativní absolutní četnost nám udává hodnotu kvantitativního znaku, kterou získáme součtem četnosti určité hodnoty x_j a četností všech hodnot menších. Tudíž poslední kumulativní absolutní četnost bude rovna rozsahu souboru n (Robová a kol., 2013).

Vzorový příklad: Z údajů, na které můžete nahlédnout do tabulky 2, doplňte tabulku rozdělení četností o kumulativní absolutní četnost.

Řešení: Na prvním řádku kumulativní četnosti bude hodnota absolutní četnosti pro známku 1. Další řádek bude výsledek součtu absolutní četnosti pro známku 1 a 2. Touto metodou analogicky spočítáme všechny ostatní řádky. Zpětnou kontrolou nám může být poznatek, že poslední řádek se musí rovnat rozsahu souboru n , v našem případě 27.

Správné výsledky a postup nám ukazuje následující tabulka 3:

Tabulka 3: Kumulativní absolutní četnosti

Známka	Absolutní četnost	Kumulativní absolutní četnost	Postup
1	8	8	= 8
2	6	14	= 8 + 6
3	10	24	= 8 + 6 + 10
4	2	26	= 8 + 6 + 10 + 2
5	1	27	= 8 + 6 + 10 + 2 + 1
Celkem	27	/	/

2.4. Kumulativní relativní četnost

Kumulativní relativní četnost vypočítáme podílem kumulativní četnosti a rozsahu souboru n . Tak jako u relativní četnosti se zpravidla udává v procentech a poslední kumulativní relativní četnost se rovná číslu 1, eventuálně 100 % (Robová a kol., 2013).

Vzorový příklad: Použijte údaje z tabulky 2 a 3 a vytvořte tabulku všech výše definovaných četností. Kumulativní relativní četnost zapište v % a zaokrouhlete na jedno desetinné místo.

Řešení:

- I. Způsob – Kumulativní relativní četnost vypočítáme následujícím způsobem, například pro známku 1 počítáme $\frac{8}{27} \times 100$, což se rovná 29,6 %. Tímto analogicky stejným způsobem dopočítáme hodnoty pro zbylé známky.
- II. Způsob – Ve druhém způsobu využijeme vypočítané údaje relativních četností z tabulky 2, kdy kumulativní relativní četnost pro známku 1 bude rovna relativní četnosti pro známku 1. Další řádek bude výsledek součtu relativní četnosti pro známku 1 a 2. Pro známku 3 počítáme součet relativních četností pro známku 1, 2 a 3, nebo můžeme použít výsledek z předchozího řádku, ke kterému přičteme relativní četnost pro známku 3. Takto můžeme dopočítat zbylé hodnoty. Zpětnou kontrolou nám může být poznatek, že poslední řádek se musí rovnat 100 %.

Tabulka 4: Kumulativní relativní četnosti

Známka	Absolutní četnost	Relativní četnost v %	Kumulativní četnost	Kumulativní relativní četnost
1	8	29,6	8	29,6
2	6	22,2	14	51,9
3	10	37,0	24	88,9
4	2	7,4	26	96,3
5	1	3,7	27	100,0
Celkem	27	100	/	/

2.5. Třídící intervaly a tabulka četností

V případě, že statistický soubor je příliš velký a hodnoty se od sebe liší pouze o malé rozdíly např. při měření výšky, pak můžeme hodnoty sobě blízké sjednotit do intervalu např. po deseti centimetrech. Počet intervalů vypočítáváme pomocí Sturgesova pravidla, kdy počet intervalů je celé číslo k , které je co nejbližší hodnotě $1 + 3,3 \log n$, kdy n je počet dat ve statistickém souboru. Z čehož dostaneme rovnici: $k = 1 + 3,3 \log n$. Nejjednodušším rozdělením hodnot znaků do intervalů pro další zpracovávání dat je vhodné, rozdělit intervaly o stejné šíři a aby každá hodnota ze souboru patřila právě do jednoho intervalu. Šířku intervalu vypočítáme ze vztahu: $l \geq \frac{\max(x_i) - \min(x_i)}{k}$, kdy v čitateli je rozdíl maximální a minimální hodnoty x_i ze statistického souboru dat a ve jmenovateli počet intervalů k . Výslednou šířku intervalů zaokrouhlujeme na celé číslo nebo vhodně zaokrouhlené číslo, se kterým se bude dobře počítat. Intervalové rozdělení četností se obvykle graficky znázorňuje pomocí histogramu četností nebo sloupcovým skupinovým diagramem (Robová a kol., 2013).

Vzorový příklad: Na dětský tábor se přihlásilo 30 dětí ve věku od 10 do 14 let. Každému dítěti v den příjezdu byla změřena výška postavy v cm. Naměřené údaje jsou uvedeny níže:

152 160 149 154 157 148 162 142 152 147 149 161 165 140 141
150 150 145 148 152 139 143 146 151 149 153 161 157 136 145.

Uvedené údaje zpracujte do tabulky rozdělení četností pomocí třídících intervalů. Dále vypočítejte relativní, kumulativní absolutní a kumulativní relativní četnost. Relativní a kumulativní relativní četnost zapište v % a zaokrouhlete na jedno desetinné místo.

Řešení: Při tvorbě intervalů je důležité, aby každá hodnota souboru spadala do nějakého intervalu. Jako první se podíváme, kolik máme celkem dat v souboru. V našem případě máme 30 dat v souboru, tedy $n = 30$. Pomocí Sturgesova pravidla vypočítáme, kolik intervalů bude naše řešení obsahovat. Počet dat dosadíme do rovnice $k = 1 + 3,3 \log n$, kdy nám vyjde $k = 5,87$, což zaokrouhlíme na 6. Teď už víme, že naše řešení bude mít 6 intervalů. Jako další se musíme podívat jaká nejmenší hodnota se nám v souboru vyskytuje, a od té vytváříme první interval. V našem případě nejmenší hodnotou je 136 centimetrů, tudíž první interval bude začínat od 136 centimetrů. Největší naměřenou hodnotou je 165 centimetrů, kterou bude končit poslední interval. Aby se nám dále dobře pracovalo, vytvoříme všechny intervaly stejně velké. Šířku intervalů určíme ze vztahu $l \geq \frac{\max(x_i) - \min(x_i)}{k}$. V našem případě to vypadá následovně: $l \geq \frac{165-136}{6}$, kdy $l = 4,8\bar{3}$ a to zaokrouhlíme na 5. V našem případě tedy zvolíme intervaly o šířce 5-ti centimetrů. První interval bude od 136 centimetrů do 140 centimetrů, druhý od 141 centimetrů do 145 centimetrů, analogicky budeme pokračovat až k poslednímu intervalu, který bude od 161 centimetrů do 165 centimetrů. Když máme zvolené intervaly, s kterými budeme pokračovat, vypíšeme absolutní četnost, která udává, kolik naměřených hodnot se nachází v každém intervalu. Dále vypočítáme relativní, kumulativní absolutní a kumulativní relativní četnost, kterou už umíme vypočítat z předešlých kapitol. Výsledná tabulka 5 rozdělení četností bude vypadat následovně:

Tabulka 5: Rozdělení četností

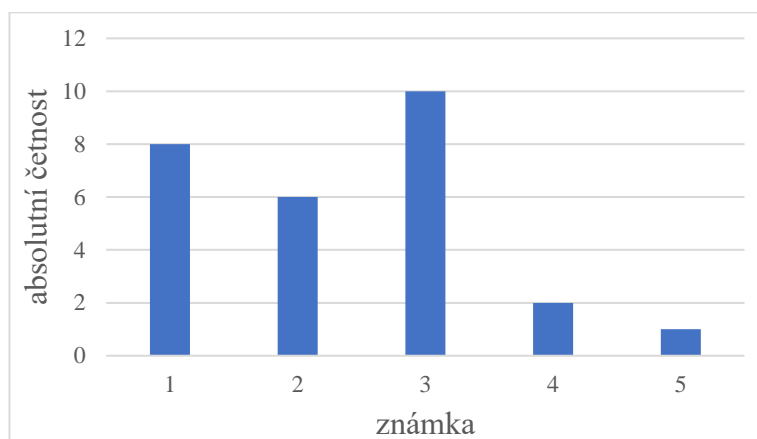
Výška dětí	Absolutní četnost	Kumulativní četnost	Relativní četnost v %	Kumulativní relativní četnost v %
136–140	3	3	10,0	10,0
141–145	4	7	13,3	23,3
146–150	10	17	33,3	56,7
151–155	6	23	20,0	76,7
156–160	3	26	10,0	86,7
161–165	4	30	13,3	100,0
Celkem	30	/	100,0	/

2.6. Grafické znázornění četností

Jak jsme již výše poznamenali, četnosti zaznamenáváme do takzvané tabulky rozdělení četností, kterou lze znázornit graficky. Zjištěné hodnoty můžeme znázornit několika grafy – sloupcovým diagramem, spojnicovým diagramem (polygon četností), histogramem a kruhovým diagramem (Calda, Dupač, 2008).

2.6.1. Sloupcový diagram

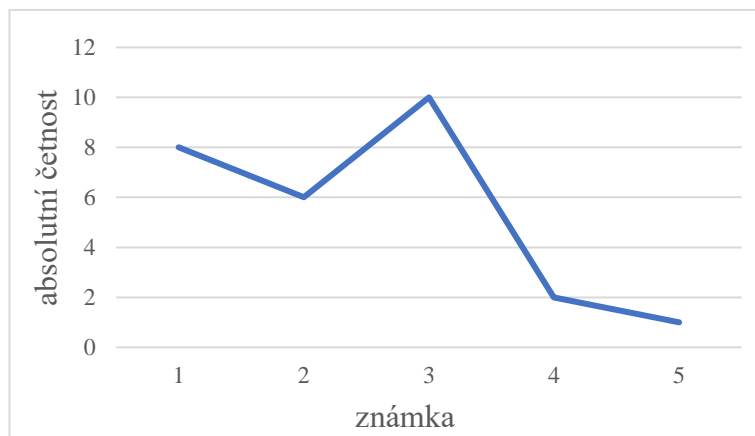
Sloupcový diagram vytvoříme tak, že na osu x nanese průměrnou hodnotu šířky intervalu kvantitativního znaku a na osu y odpovídající četnosti, mezi jednotlivými sloupci jsou mezery. Na následujícím obrázku 1 bude znázorněn sloupcový diagram pro výsledky písemné práce z českého jazyka pro příklad z kapitoly 2.1.



Obrázek 1: Výsledky písemné práce z českého jazyka z kapitoly 2.1

2.6.2. Spojnicový diagram

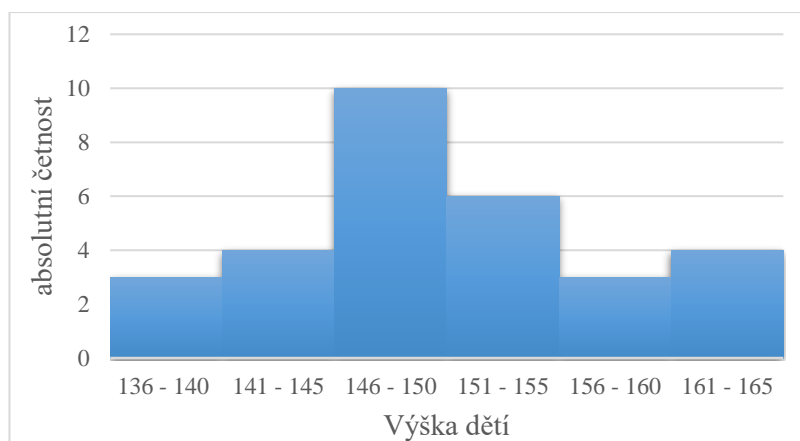
Spojnicový diagram (polygon četností) sestavíme spojením bodů, kdy na osu x nanášíme hodnoty kvantitativního znaku a na osu y jim odpovídající četnost. Vzniklé body spojíme lomenou čarou. Pokud se hodnota kvantitativního znaku rovná nule, pak daný bod leží na ose x . Na následujícím obrázku 2 bude znázorněn spojnicový diagram pro výsledky písemné práce z českého jazyka pro příklad z kapitoly 2.1.



Obrázek 2: Výsledky písemné práce z českého jazyka z kapitoly 2.1

2.6.3. Histogram

Histogram vytvoříme tak, že na osu x nanese celou šířku intervalu a na osu y odpovídající četnosti. V případě, že bychom nezvolily intervaly stejně dlouhé, potom by se četnostem rovnal plošný obsah sloupků nikoliv výška sloupků. Zpravidla ho používáme u intervalového rozdělení četností (Caldá, Dupač, 2008). Na následujícím obrázku 3 bude znázorněn histogram pro výšku dětí na dětském táboře pro příklad z kapitoly 2.5.



Obrázek 3: Výška dětí na dětském táboře z kapitoly 2.5

2.6.4. Kruhový diagram

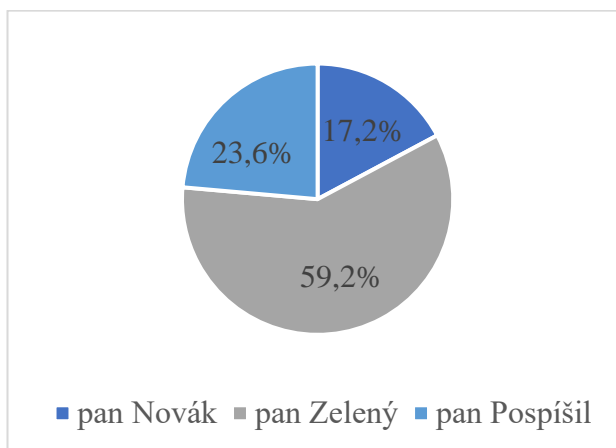
Kruhový diagram používáme při znázorňování kvalitativního znaku. Graf sestavíme tak, že různým hodnotám znaku odpovídají kruhové výseče, jejichž plošné obsahy jsou úměrné relativním četnostem (Caldá, Dupač, 2008).

Vzorový příklad: Občané nejmenovaného městečka, ve kterém žije 725 obyvatel si volili nového starostu. Na výběr měli ze tří kandidátů. Volby nakonec dopadli následovně: 125 občanů zvolilo pana Nováka, 429 občanů zvolilo pana Zeleného a 171 občanů hlasovalo pro pana Pospíšila. Vytvořte tabulku rozdělení četností a z vypočítaných hodnot sestavte kruhový diagram.

Řešení:

Tabulka 6: Rozdělení četností

jméno kandidáta	pan Novák	pan Zelený	pan Pospíšil
absolutní četnost n	125	429	171
relativní četnost v %	17,2	59,2	23,6



Obrázek 4: Volba starosty

Příklady k procvičení

Příklad 1.: Učitel opravuje čtvrtletní písemné práce z třídy 9.A. Výsledky práce si zapisuje: 1, 3, 2, 1, 1, 3, 1, 2, 3, 5, 1, 5, 2, 1, 3, 1, 2, 2, 2, 4.

Sestavte tabulku rozdělení četností a pak z ní sestavte polygon četností.

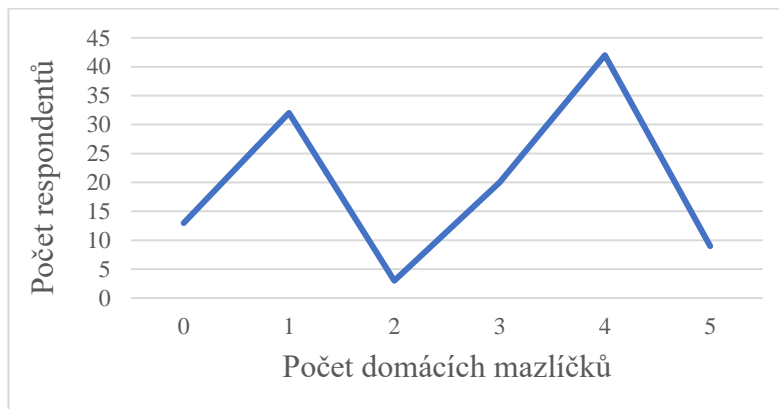
Příklad 2.: Do restaurace přišlo během hlavního obědového menu 300 zákazníků. Číšníci si během dne zapisovali, kdo si objednal, jaké jídlo a výsledné hodnoty zanesli do tabulky 7:

Tabulka 7: Výsledné hodnoty objednaného jídla

Název jídla	Rizoto	Svíčková	Caesar salát	Guláš	Pečené kuře
Absolutní četnost		75	116	34	63
Relativní četnost v %	4				21

Doplňte chybějící údaje do tabulky a vytvořte kruhové digramy pro oba typy četností.

Příklad 3.: U 119 respondentů jsme se ptali, kolik mají doma domácích mazlíčků. Ze zjištěných informací jsme sestavili polygon četností viz. Obrázek 5:



Obrázek 5: Graf počtu domácích mazlíčků

- A) Kolik respondentů má nejvíce tři domácí mazlíčky?
- B) Kolik respondentů má právě čtyři domácí mazlíčky?
- C) Kolik respondentů má nejméně dva domácí mazlíčky?

Příklad 4.: Na vysokou školu bylo přijato 540 studentů z různých typů středních škol. Počty uchazečů uvádí následující tabulka 8:

Tabulka 8: Počty uchazečů na vysokou školu

Typ střední školy	Počet uchazečů
Gymnázium	201
Obchodní akademie	134
SOŠ	168
SOU	37
Celkem	540

Vypočítejte relativní četnosti (zaokrouhlete na jedno desetinné místo) a výsledky znázorněte pomocí kruhového digramu.

Příklad 5.: Švadlena za měsíc ušila 59 šatů. Velikosti ušitých šatů jsou následující:

38, 36, 40, 42, 44, 34, 38, 38, 36, 34, 36, 40, 44, 44, 42, 40, 36, 36, 38, 38, 36, 40, 34, 42, 46, 40, 42, 34, 38, 36, 38, 42, 42, 40, 40, 40, 46, 38, 36, 38, 40, 34, 36, 36, 40, 42, 38, 38, 34, 38, 40, 42, 40, 36, 36, 40, 36, 44, 36.

- A) Vypracujte tabulku rozdělení četností, vypočítejte kumulativní četnost, relativní četnost a kumulativní relativní četnost.
- B) Sestrojte pro získaná data sloupcový diagram a polygon četností.

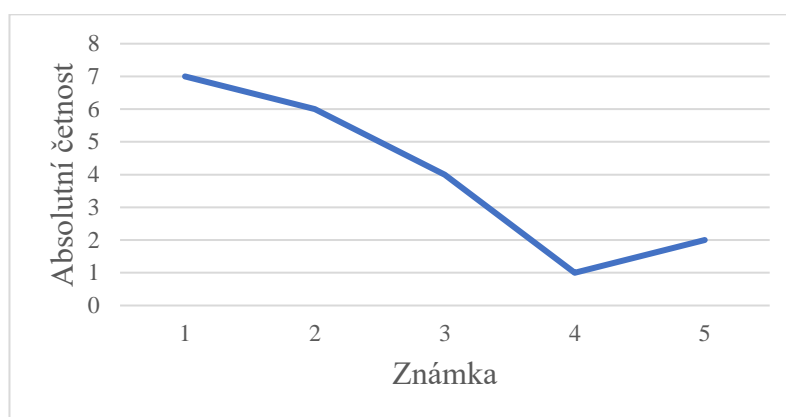
Příklad 6.: Sestavte kruhový diagram, kde poukážete na infrastrukturu silniční dopravy v České republice v roce 2019. Přičemž silnice rozpracujete pro silnice I. třídy, silnice II. třídy a silnice III. třídy. Veškeré potřebné informace naleznete na Českém statistickém úřadě.

Výsledky

Příklad 1.:

Tabulka 9: Výsledky z čtvrtletní práce z 9.A

Známka	1	2	3	4	5
Absolutní četnost	7	6	4	1	2
Relativní četnost v %	35	30	20	5	10

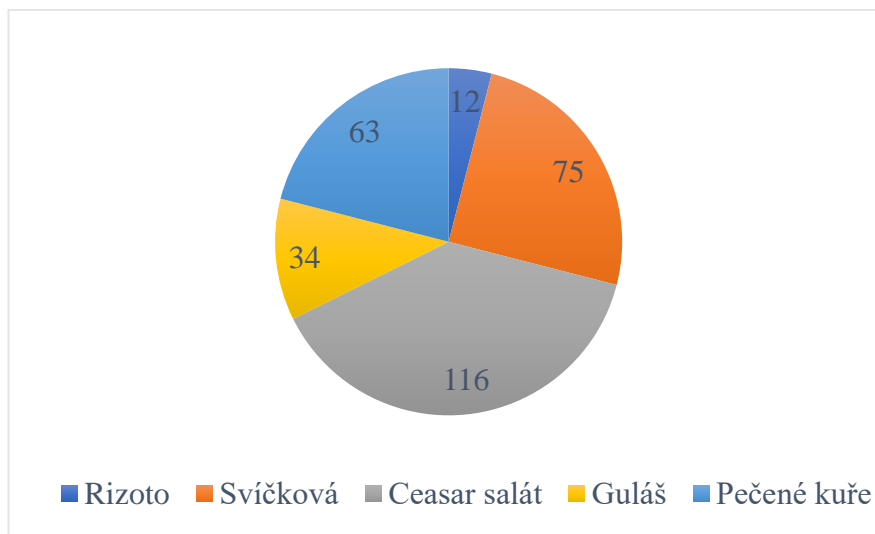


Obrázek 6: Výsledky z čtvrtletní práce z 9.A

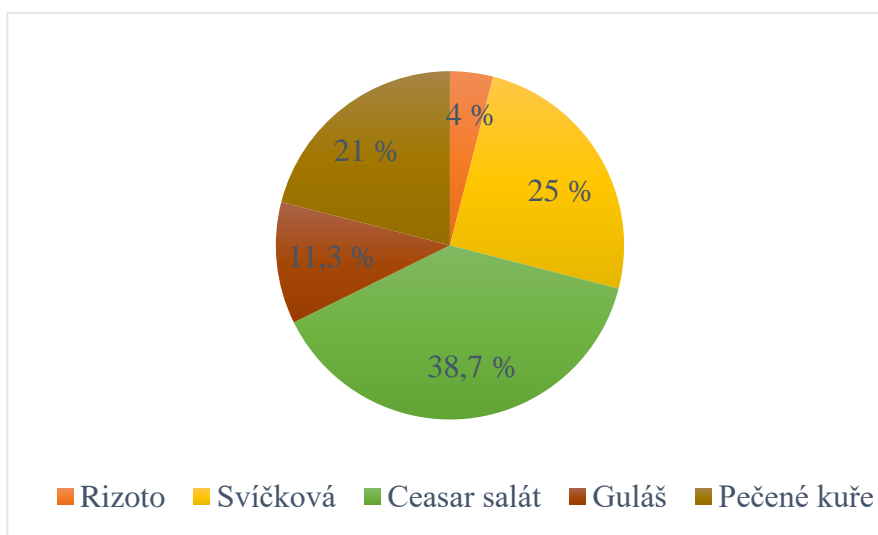
Příklad 2.:

Tabulka 10: Výsledné hodnoty objednaného jídla

Název jídla	Rizoto	Svíčková	Caesar salát	Guláš	Pečené kuře
Absolutní četnost	12	75	116	34	63
Relativní četnost v %	4	25	38,7	11,3	21



Obrázek 7: Absolutní četnosti objednaného jídla



Obrázek 8: Relativní četnosti objednaného jídla

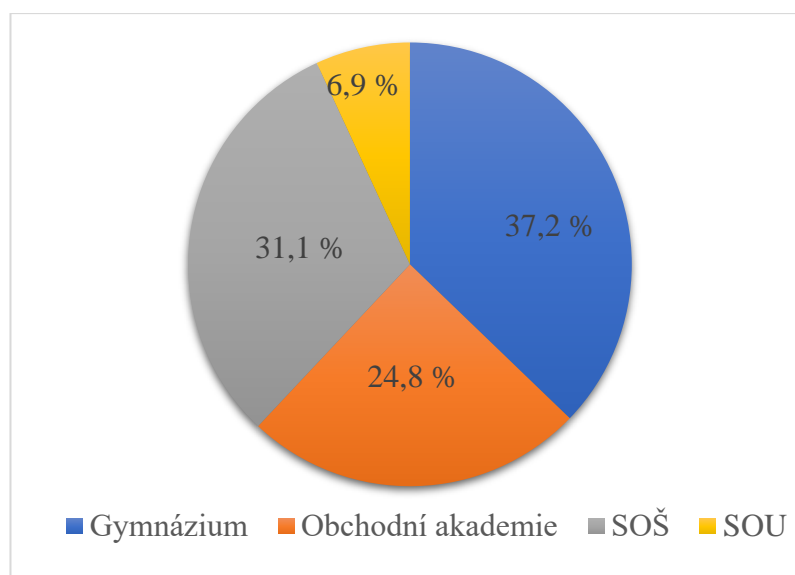
Příklad 3.:

- A) 63 respondentů má doma nejvíce tři domácí mazlíčky
- B) 42 respondentů má právě čtyři domácí mazlíčky
- C) 74 respondentů má doma nejméně dva domácí mazlíčky

Příklad 4.:

Tabulka 11: Počty studentů hlásící se na vysokou školu

Typ střední školy	Počet uchazečů	Relativní četnost v %
Gymnázium	201	37,2
Obchodní akademie	134	24,8
SOS	168	31,1
SOU	37	6,9
Celkem	540	100



Obrázek 9: Znázornění uchazečů z různých typů škol

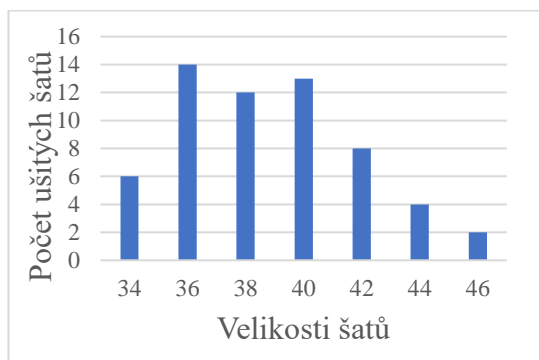
Příklad 5.:

A)

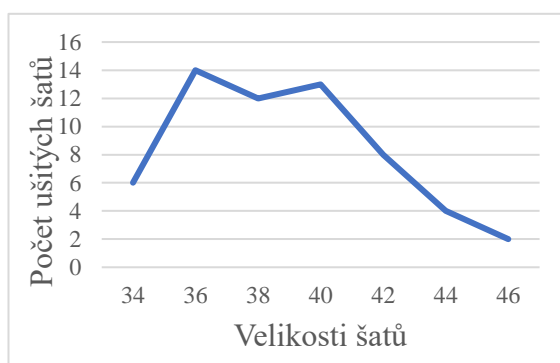
Tabulka 12: Znázornění četností ušitých šatů

Velikost šatů	Absolutní četnost	Kumulativní četnost	Relativní četnost v %	Kumulativní relativní četnost v %
34	6	6	10,2	10,2
36	14	20	23,7	33,9
38	12	32	20,3	54,2
40	13	45	22	76,2
42	8	53	13,6	89,8
44	4	57	6,8	96,6
46	2	59	3,4	100
Celkem	59	/	100	/

B)



Obrázek 10: Sloupcový diagram ušitých šatů



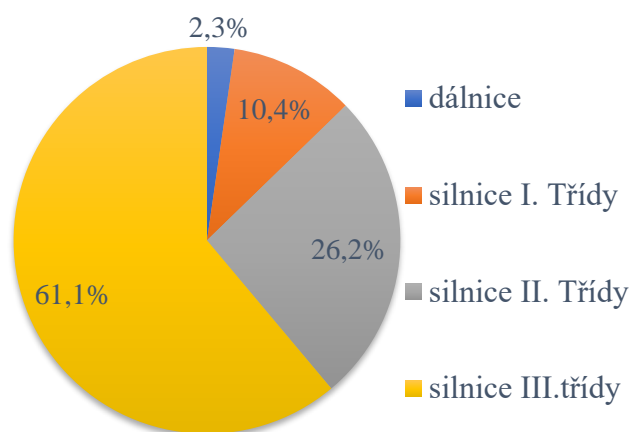
Obrázek 11: Polygon četností ušitých šatů

Příklad 6.:

Tabulka 13: Infrastruktura silniční dopravy v ČR v roce 2019

Rok	Délka silnic a dálnic							
	Celkem	z toho evropská silniční síť typu E	Dálnice v provozu	Rychlostní komunikace	Silnice ²⁾			
					celkem	silnice I. třídy	silnice II. třídy	silnice III. třídy
2019	55 768	2 629	1 276	0	54 492	5 826	14 585	34 081

Zdroj: Český statistický úřad



Obrázek 12: Infrastruktura silniční dopravy v ČR v roce 2019

3. Charakteristiky polohy

V této kapitole se budeme zabývat charakteristikami polohy kvantitativního znaku, který označíme písmenem x . Charakteristiky polohy označujeme někdy jako střední hodnotu znaku, díky kterým vidíme polohu znaku na číselné ose. Mezi charakteristiky polohy řadíme aritmetický průměr, geometrický průměr, harmonický průměr, medián, modus a kvantily (Calda, Dupač, 2008).

3.1. Aritmetický průměr

Aritmetický průměr značíme \bar{x} , což čteme jako „ x s pruhem“. Aritmetický průměr vypočítáme jako součet všech hodnot znaku x jednotek statistického souboru n dělený počtem všech jednotek statistického souboru n . Aritmetický průměr můžeme tedy zapsat jako:

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n), \text{ tj: } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{Calda, Dupač, 2008, str. 137}).$$

Vzorový příklad: Na dětský tábor se přihlásilo 30 dětí ve věku od 10 do 14 let. Každému dítěti v den příjezdu byla změřena výška postavy v cm. Naměřené údaje jsou uvedeny níže:

152 160 149 154 157 148 162 142 152 147 149 161 165 140 141
150 150 145 148 152 139 143 146 151 149 153 161 157 136 145.

Vypočítejte průměrnou výšku dětí v centimetrech.

Řešení: Průměrnou výšku dětí vypočítáme jako podíl, kdy v čitateli bude součet všech výšek dětí a ve jmenovateli počet všech dětí. Výsledná rovnice vypadá následovně:

$$\bar{x} = \frac{152 + 160 + 149 + \dots + 157 + 136 + 145}{30}$$

$$\bar{x} = 150,1\bar{3}$$

3.1.1. Vážený aritmetický průměr

V případě, že chceme počítat aritmetický průměr z tabulky rozdělení četností, pak musíme každou hodnotu x_j^* vynásobit její absolutní četností n_j . Vážený aritmetický

průměr zapíšeme vzorcem: $\bar{x} = \frac{1}{n} \sum_{j=1}^r x_j^* n_j$ (Calda, Dupač, 2008, str. 137).

Vzorový příklad: Na dětský tábor se přihlásilo 30 dětí ve věku od 10 do 14 let. Každému dítěti v den příjezdu byla změřena výška postavy v cm. Naměřené údaje jsou uvedeny níže:

152 160 149 154 157 148 162 142 152 147 149 161 165 140 141
150 150 145 148 152 139 143 146 151 149 153 161 157 136 145.
Vytvořte tabulku rozdělení četností, z které následně vypočítejte průměrnou výšku dětí.

Řešení: Jako první sestavíme tabulku rozdělení četností pro naše údaje.

Tabulka 14: Tabulka rozdělení četností výšky dětí

výška dětí	absolutní četnost n_j	výška dětí	absolutní četnost n_j	výška dětí	absolutní četnost n_j
136	1	146	1	153	1
139	1	147	1	154	1
140	1	148	2	157	2
141	1	149	3	160	1
142	1	150	2	161	2
143	1	151	1	162	1
145	2	152	3	165	1

Potom z vytvořené tabulky vypočítáme průměrnou výšku dětí v centimetrech.

$$\bar{x} = \frac{136 \cdot 1 + 139 \cdot 1 + 140 \cdot 1 + \dots + 161 \cdot 2 + 162 \cdot 1 + 165 \cdot 1}{30}$$

$$\bar{x} = 150,1\bar{3}$$

Průměrná výška dětí nám vychází úplně stejně, jako když jsme ji počítali rovnou z údajů, které jsme znaly ze zadání příkladu.

3.1.2. Průměr počítaný z dat roztríděných intervalů

Pokud chceme počítat aritmetický průměr z třídních intervalů a tabulky četností, pak vždy při výpočtu počítáme se středem intervalu. Tento způsob je vhodný, pokud počítáme s malými čísly, které se dobře počítají z paměti. V případě, že se v souboru objevují velká čísla, která jsou pro rychlý ruční výpočet zcela nepohodlné a člověk si musí psát své mezi výpočty. Tak počítáme s tzv. metodou vhodně zvoleného počátku. Hodnoty znaku x zapíšeme ve tvaru $x = a + bu$, kde a a b jsou konstanty, pak bude tento vztah platit i pro průměry, kdy $\bar{x} = a + b\bar{u}$ (Caldá, Dupač, 2008, str. 138).

Vzorový příklad: Na dětský tábor se přihlásilo 30 dětí ve věku od 10 do 14 let. Každému dítěti v den příjezdu byla změřena výška postavy v cm. Naměřené údaje jsou uvedeny níže:

152 160 149 154 157 148 162 142 152 147 149 161 165 140 141
150 150 145 148 152 139 143 146 151 149 153 161 157 136 145.

Vypočítejte průměrnou výšku dětí v centimetrech, pro výpočet použijte údaje z tabulky číslo 5.

Řešení:

A) pomocí středu intervalu

$$\bar{x} = \frac{138 \cdot 3 + 143 \cdot 4 + 148 \cdot 10 + 153 \cdot 6 + 158 \cdot 3 + 163 \cdot 4}{30}$$

$$\bar{x} = \frac{4510}{30}$$

$$\bar{x} = 150,3$$

B) pomocí metody vhodně zvoleného počátku

Konstantu a zvolíme 138, což je středová hodnota našeho prvního intervalu, konstanta b se bude rovnat 5, protože naše intervaly máme určené po pěti centimetrech. V našem případě nám tudíž vznikne rovnice, kdy $x = 138 + 5u$. Hodnotám 138, 143, 148, 153, 158 a 163 znaku x budou odpovídat hodnoty 0, 1, 2, 3, 4, 5 a 6 znaku u . Pomocný průměr \bar{u} vypočítáme následovně:

$$\bar{u} = \frac{0 \cdot 3 + 1 \cdot 4 + 2 \cdot 10 + 3 \cdot 6 + 4 \cdot 3 + 5 \cdot 4}{30} = \frac{74}{30} = 2,4\bar{6}$$

V momentě, kdy máme vypočítaný pomocný průměr \bar{u} , můžeme dosadit do rovnice $\bar{x} = a + b\bar{u} = 138 + 5 \cdot 2,4\bar{6} = \underline{150,3}$.

Na začátku této kapitoly jsme počítali s původními daty, kdy nám průměrná výška dětí vyšla 150,13. Když jsme k výpočtu použili zaokrouhlené hodnoty na středy intervalů, vyšla nám průměrná výška dětí 150,3. Proto je třeba si uvědomit, že když budeme počítat se středy intervalů, tak se nám výsledek vždy bude o něco málo lišit.

3.2. Geometrický průměr

Geometrický průměr se počítá jen pro kladná čísla na rozdíl od aritmetického průměru, označujeme ho symbolem \bar{z}_G . V běžném životě se používá převážně v geografii na výpočet průměrného tempa růstu nebo například v národohospodářských časových řadách, které nám říkají, jakým tempem se vyvíjí průmyslová nebo zemědělská výroba. Pokud bychom si jednotlivá období označili jako 0, 1, 2, ..., n pak jim odpovídající hodnoty znaku jsou $x_0, x_1, x_2, \dots, x_n$. Pak můžeme říct, že $y_1 = x_1 - x_0, y_2 = x_2 - x_1, \dots, y_n = x_n - x_{n-1}$ jsou přírůstky za jednotlivá období. Průměrný přírůstek tak můžeme vypočítat jako $\bar{y} = \frac{x_n - x_0}{n}$. Někdy se setkáváme s tím, že potřebujeme znát pouze hodnotu za jedno období, čímž se myslí průměr podílů hodnot za dvě po sobě jdoucí období, tedy $z_1 = \frac{x_1}{x_0}, z_2 = \frac{x_2}{x_1}, \dots, z_n = \frac{x_n}{x_{n-1}}$. Geometrický průměr můžeme tedy zapsat jako:

$$\bar{z}_G = \sqrt[n]{z_1 \cdot z_2 \cdot \dots \cdot z_n} = \sqrt[n]{\frac{x_n}{x_0}}. \text{ Zpravidla se hodnoty udávají v procentech, ale nemusí}$$

tomu být tak vždy (Calda, Dupač, 2008, str. 140-141).

Vzorový příklad: Nejmenovaná firma vykázala za posledních 7 let následující hodnoty: 1,0; 1,8; 2,6; 3,0; 3,4; 3,7; 4,2 a 5,1, tyto hodnoty představují její čistý zisk v milionech korun. Vypočítejte průměrné roční tempo růstu za uvedených 7 let.

Řešení: Nejprve musíme vypočítat jednotlivá roční tempa růstu. Hodnotu 1,0 zapíšeme jako x_0 , analogicky takto zapíšeme všechny hodnoty. První období z_1 vypočítáme podle vztahu: $z_1 = \frac{x_1}{x_0}$, tedy $z_1 = \frac{1,8}{1,0} = 1,8$. Analogicky vypočítáme zbylá období: $z_2 = \frac{2,6}{1,8} = 1,4$

$$z_3 = \frac{3,0}{2,6} \doteq 1,15 \quad z_4 = \frac{3,4}{3,0} = 1,1\bar{3} \quad z_5 = \frac{3,7}{3,4} \doteq 1,09 \quad z_6 = \frac{4,2}{3,7} \doteq 1,14 \quad z_7 = \frac{5,1}{4,2} \doteq 1,21;$$

když máme vypočítané jednotlivá období, pak z těchto hodnot můžeme vypočítat geometrický průměr. Geometrický průměr vypočítáme ze vztahu $\bar{z}_G = \sqrt[n]{z_1 \cdot z_2 \cdot \dots \cdot z_n}$, v našem případě tedy $\bar{z}_G = \sqrt[7]{1,8 \cdot 1,4 \cdot 1,15 \cdot 1,1\bar{3} \cdot 1,09 \cdot 1,14 \cdot 1,21} \doteq 1,26$.

Avšak mnohem jednodušším způsobem je vypočítat průměrné roční tempo růstu ze vztahu: $\bar{z}_G = \sqrt[n]{\frac{x_n}{x_0}}$, kdy za x_n dosadíme poslední hodnotu souboru 5,1 a za x_0 první hodnotu souboru 1,0. Pak tedy dostáváme $\bar{z}_G = \sqrt[7]{\frac{5,1}{1,0}} \doteq 1,26$. Ať už volíme jakýkoliv způsob, vždy nám musí vyjít stejný výsledek. Průměrné roční tempo firmy činí 1,26 %.

3.3. Harmonický průměr

Harmonický průměr značíme \bar{x}_H . Harmonickým průměrem kladných čísel x_1, x_2, \dots, x_n chápeme převrácenou hodnotu aritmetického průměru převrácených hodnot.

Harmonický průměr můžeme zapsat jako: $\bar{x}_H = \frac{1}{\frac{1}{n} \left(\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n} \right)}$ (Caldá, Dupač, 2008, str. 141).

Vzorový příklad: V kavárně mají tři kávovary. První kávovar vyrobí kávu za 3 minuty, druhý kávovar vyrobí kávu za 4 minut a třetí kávovar vyrobí kávu za 12 minut. Vypočítejte, za jak dlouho se v průměru vyrobí jedna káva v kavárně.

Řešení: Ze zadání známe tři hodnoty, první hodnotu $x_1 = 3$, druhou hodnotu $x_2 = 4$ a poslední hodnotu $x_3 = 12$. Průměrnou výrobu jedné kávy vypočítáme pomocí vztahu

$\bar{x}_H = \frac{1}{\frac{1}{n} \left(\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n} \right)}$. Za neznámou n dosadíme číslo 3, protože průměrnou dobu

počítáme z tří kávovarů. Po dosazení dostáváme: $\bar{x}_H = \frac{1}{\frac{1}{3} \cdot \left(\frac{1}{3} + \frac{1}{4} + \frac{1}{12} \right)} = 4,5$. Průměrná doba

na výrobu jedné kávy je 4 minuty a 30 sekund.

3.4. Modus

Modus znaku x je charakteristika polohy, která nám označujeme hodnotu znaku x s největší absolutní četností souboru n , značíme jej $Mod(x)$ (Calda, Dupač, 2008, str. 142).

Vzorový příklad: Z údajů, na které můžete nahlédnout v tabulce 12, určete modus.

Řešení: Z tabulky rozdělení četností vidíme, že švadlena ušila nejvíce šatů ve velikosti 36, proto $Mod(x) = 36$.

3.5. Medián

Medián znaku x označuje prostřední hodnotu znaku, pokud jsou hodnoty x_1, x_2, \dots, x_n uspořádané od nejmenší po největší. Značí se $Med(x)$. Pokud splňujeme podmínku, že $x_1 \leq x_2 \leq \dots \leq x_n$, pak platí: $Med(x) = x_{(\frac{n+1}{2})}$, je-li n liché číslo a

$Med(x) = \frac{1}{2} \left(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)} \right)$, je-li n sudé číslo. Medián zpravidla používáme u souborů, kdy nám některé hodnoty hodně vyčnívají z řady (Calda, Dupač, 2008, str. 142).

Vzorový příklad: Z údajů, na které můžete nahlédnout v tabulce 12, vypočtete medián.

Řešení: Připomeneme si výslednou tabulku, která vypadá následovně:

Tabulka 15: Znázornění četností ušitých šatů

Velikost šatů	Absolutní četnost	Kumulativní četnost	Relativní četnost v %	Kumulativní relativní četnost v %
34	6	6	10,2	10,2
36	14	20	23,7	33,9
38	12	32	20,3	54,2
40	13	45	22	76,2
42	8	53	13,6	89,8
44	4	57	6,8	96,6
46	2	59	3,4	100
Celkem	59	/	100	/

Ze sloupce s absolutní četností vidíme, že náš soubor má celkem 59 hodnot, což je liché číslo, proto použijeme vzorec $Med(x) = x_{(\frac{n+1}{2})}$. Za konstantu n dosadíme hodnotu 59 a následně dostaneme rovnici $Med(x) = x_{(\frac{59+1}{2})} = x_{30}$. Zjistili jsme, že medián se rovná

hodnotě znaku x_{30} . Tuto hodnotu snadno zjistíme ze sloupce kumulativní četnosti, kde vidíme, že hodnota znaku x_{30} se rovná 38. $\text{Med}(x) = 38$.

3.6. Kvantily

Kvantily rozdělují statistické soubory na několik zhruba stejně velkých částí. Pokud budeme mít libovolné číslo p , které bude nabývat hodnot od 0 do 100, pak p -procentním kvantilem statistického znaku bude hodnota \tilde{x}_p . Pro hodnotu \tilde{x}_p platí, že nejméně p -procent prvků má menší hodnotu nebo rovnu \tilde{x}_p a zároveň zbytek prvků (tedy $(100-p)\%$) má hodnotu větší nebo rovnu \tilde{x}_p . Hodnotu \tilde{x}_p určujeme z uspořádaného souboru hodnot, který je ve tvaru $x_1 \leq x_2 \leq \dots \leq x_n$. Nejprve musíme určit pořadí, které značíme písmenem r , na kterém se nachází náš hledaný kvantil. Pořadí vypočítáváme pomocí vztahu: $r = \frac{p}{100} \cdot (n + 1)$. Pokud bude r celé číslo, pak se hledaný kvantil bude přímo rovnat hodnotě na pozici r , tedy $\tilde{x}_p = x_r$. Pokud r nebude celé číslo, pak r určíme zaokrouhlením na nejbližší menší číslo, které označujeme písmenem c . Zbytek, který nám zbude po zaokrouhlení značíme z a platí pro něj vztah: $z = r - c$. Hledaný kvantil \tilde{x}_p bude ležet mezi hodnotou x_c a hodnotou $x_{(c+1)}$. Kvantil \tilde{x}_p vypočítáme pomocí vztahu: $\tilde{x}_p = x_c + z \cdot [x_{(c+1)} - x_c]$ (Robová a kol., 2013, str. 172).

Vzorový příklad: Tradičního závodu v malé vesničce na Benešovsku se zúčastnilo 16 závodníků, kteří měli na svých dresech svá oblíbená čísla, která jsou následující: 19 40 27 43 3 17 25 33 77 99 13 7 31 57 61 24. Určete 25, 40 a 90procentní kvantil pro tyto data, tedy čísla \tilde{x}_{25} , \tilde{x}_{40} a \tilde{x}_{90} .

Řešení: Nejprve musíme uspořádat data podle velikosti. Uspořádaný soubor vypadá následovně: 3 7 13 17 19 24 25 27 31 33 40 43 57 61 77 99. Nyní už budeme počítat jednotlivé kvantily.

Pro $p = 25$ je $r = \frac{25}{100} \cdot (16 + 1) = 4,25$, což zaokrouhlíme na 4 se zbytkem 0,25, tedy $c = 4$ a $z = 0,25$. Potom tyto výsledky dosadíme do vztahu $\tilde{x}_p = x_c + z \cdot [x_{(c+1)} - x_c]$ v našem případě tedy $\tilde{x}_{25} = x_4 + 0,25 \cdot [x_5 - x_4] = 17 + 0,25 \cdot (19 - 17) = 17,5$.

Pro $p = 40$ je $r = \frac{40}{100} \cdot (16 + 1) = 6,8$, tedy $c = 6$ a $z = 0,8$. Potom $\tilde{x}_{40} = x_6 + 0,8 \cdot [x_7 - x_6] = 24 + 0,8 \cdot (25 - 24) = 24,8$.

Pro $p = 90$ je $r = \frac{90}{100} \cdot (16 + 1) = 15,3$, a tedy $\tilde{x}_{90} = x_{15} + 0,3 \cdot [x_{16} - x_{15}] = 83,6$.

3.6.1. Percentily a decily

Ve statistickém souboru můžeme definovat tzv. percentily, které dělí statistický soubor na setiny. Percentilů je ve statistickém souboru celkem 99 a označují se jako $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{99}$. Percentily se hodně používají při vyhodnocování SCIO testů, kdy se bodové skóre převede na percentily. Pokud má účastník umístění na 85. percentilu, znamená to, že 85 % účastníků mělo horší výsledek z testu a zbylých 15 % účastníků je buď lepších nebo stejných jako on. Místo 85. percentilu můžeme zvolit jakýkoliv od 0 do 100 (Robová a kol., 2013).

Obdobně můžeme definovat i tzv. decily, které dělí soubor na desetiny. Celkem máme 9 decilů a označují se jako $\tilde{x}_{10}, \tilde{x}_{20}, \dots, \tilde{x}_{90}$.

3.6.2. Tercily

Tercily dělí statistický soubor na třetiny. První třetina nám odděluje první třetinu hodnot statistického souboru od zbývajících dvou třetin. Druhá třetina nám odděluje první dvě třetiny hodnot statistického souboru od poslední třetiny hodnot statistického souboru. V praxi se používají jen zřídka.

3.6.3. Kvartily

Kvartily jsou hodnoty, které od sebe oddělují čtvrtiny vzestupně uspořádaných hodnot statistického souboru. Jsou tedy celkem tři. Prvním kvantilem, který označujeme jako *dolní kvartil*, rozumíme číslo \tilde{x}_{25} , který odděluje první čtvrtinu hodnot statistického souboru od zbývajících čtvrtin hodnot. Dalším kvantilem je prostřední kvartil = medián, tedy číslo \tilde{x}_{50} a odděluje první polovinu hodnot od druhé poloviny neboli první dvě čtvrtiny od zbylých dvou čtvrtin. Třetím kvantilem, který nazýváme *horní kvartil* je číslo \tilde{x}_{75} a odděluje první tři čtvrtiny hodnot statistického souboru od poslední čtvrtiny hodnot statistického souboru (Robová a kol., 2013, str. 173).

Vzorový příklad: Určete dolní, prostřední a horní kvartil, pokud máte k dispozici tyto údaje uvedené v tabulce 16:

Tabulka 16: Tabulka hodnot potřebných ke vzorovému příkladu

Hodnota znaku	Absolutní četnost	Kumulativní absolutní četnost
1	3	3
2	6	9
3	1	10
4	2	12
5	12	24
6	4	28
7	4	32
8	6	38
9	8	46
10	15	61
11	7	68
12	2	70
13	1	71
14	7	78
15	3	81
Celkem	81	/

Řešení: Pro $p = 25$ je $r = \frac{25}{100} \cdot (81 + 1) = 20,5$; tedy $c = 20$ a $z = 5$.

Potom $\tilde{x}_{25} = x_{20} + 0,5 \cdot [x_{21} - x_{20}] = 5 + 0,5 \cdot (5-5) = 5$

Pro $p = 50$ je $r = \frac{50}{100} \cdot (81 + 1) = 41$, což je celé číslo, pak $\tilde{x}_{50} = x_{41} = 9$.

Pro $p = 75$ je $r = \frac{75}{100} \cdot (81 + 1) = 61,5$; tedy $c = 61$ a $z = 5$.

Potom $\tilde{x}_{75} = x_{61} + 0,5 \cdot [x_{62} - x_{61}] = 10 + 0,5 (11-10) = 10,5$.

Příklady k procvičení

Příklad 1.: U deseti vysokoškolských studentů jsme zjišťovali, kolik peněz utratí za volnočasové aktivity během jednoho měsíce. Jejich odpovědi byly následující: 1500 Kč, 750 Kč, 3700 Kč, 1200 Kč, 2350 Kč, 450 Kč, 3000 Kč, 5200 Kč, 700 Kč a 1100 Kč. Vypočítejte, kolik peněz v průměru utratí vysokoškolský student za volnočasové aktivity.

Příklad 2.: Pan Veselý si během dubna dvakrát týdně zapisoval venkovní teplotu. Naměřené údaje jsou ve °C a jsou následující: 5,6; 12,7; 9,4; 3,4; 5,6; 10,4; 10,8; 17,7 a 6,9. Vypočítejte průměrnou teplotu za měsíc duben. Data byla použita z Českého hydrometeorologického ústavu.

Příklad 3.: Z tabulky 17 vypočtete průměrnou známku, modus a medián z písemné práce z českého jazyka.

Tabulka 17: Výsledky z písemné práce z českého jazyka

známka	1	2	3	4	5
absolutní četnost n_j	8	6	10	2	1

Příklad 4.: Průměrná míra nezaměstnanosti vyjadřuje procentní změnu nezaměstnanosti za daný rok oproti průměrné nezaměstnanosti za rok předešlý. Následující tabulka 18 udává hodnoty průměrné míry nezaměstnanosti v procentech v České republice za roky 2010–2019.

Tabulka 18: Míra nezaměstnanosti v České republice

2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
7,3	6,7	7	7	6,1	5	4	2,9	2,2	2

Zdroj: Český statistický úřad

Vypočítejte průměrnou míru nezaměstnanosti za celé období.

Příklad 5.: Pan Kovář jel ze svého bydliště do hlavního města pro náhradní součástky, které potřeboval ke své práci, jelikož vyrazil v době největší dopravní špičky jel průměrnou rychlostí pouze 42 km/h. Při zpáteční cestě jel už v pozdějších hodinách, kdy silnice byly téměř prázdné, a proto jeho průměrná rychlost byla 77 km/h. Vypočítejte průměrnou rychlost pana Kováře do hlavního města a zpět domů.

Příklad 6.: Město vypsaló výběrové řízení na zakázku na výrobu lávky přes řeku v městském parku. O tuto zakázku se přihlásily celkem čtyři firmy. První firma označme ji *A* by lávku stihla vytvořit za 56 dní. Druhá firma *B* by vytvořila lávku za 70 dní, třetí firma *C* by zakázku stihla za 81 dní a poslední firma *D* by lávku vytvořila za 43 dní.

a) Vypočítejte, za kolik dní by byla zakázka v průměru vytvořena.

b) Jak dlouho by trvala zakázka, kdyby firmy lávku tvořily společně?

Příklad 7.: Hrubý domácí produkt tzv. HDP v České republice za posledních 6 let rostlo následovně: 2 %, 4,3 %, 2,3 %, 4,6 %, 2,9 %, 2,4 %. Vypočítejte průměrný růst hrubého domácího produktu v České republice. Data byla použita z Českého statistického úřadu.

Příklad 8.: Skupina natěračů natírá určitou část plotu. Dvanáct natěračů natře část plotu za 4 minuty, pět natěračů ji natře za 10 minut a tři natěrači za 6 minut. Kolik minut připadá průměrně na jednoho natěrače?

Příklad 9.: Vypočítejte 33. percentil z výšek dívek, které se zúčastnily soutěže krásy. Měřením jsme získali následující výsledky, které jsou vyjádřené v centimetrech: 171 175 183 170 172 175 166 178 180 175 187.

Příklad 10.: Podnikatel prodával jeden kus svého zboží za 100 Kč. Po nějaké době se rozhodl své zboží zdražit o 15 % na 115 % hodnoty. Avšak se zvyšující se nabídkou stejného zboží se podnikatel rozhodl zdražit zboží o dalších 30 % na 130 % již zvýšené hodnoty. O kolik procent průměrně obchodník zdražil při jednom zdražení?

Příklad 11.: Lékař během jednoho pracovního dne měřil u svých devíti pacientů hladinu cholesterolu v krvi. Naměřené hodnoty byly následující: 3,1 3,8 4,5 2,0 1,9 5,0 3,3 4,2 3,7. Vypočítejte horní a dolní kvartil.

Příklad 12.: V České republice se nachází celkem 25 chráněných krajinných oblastí. Následující hodnoty udávají rok zřízení: 1973 1980 1981 1989 1991 1976 1972 2005 1955 1969 1968 1976 1978 1972 1990 1976 1956 1969 1976 1991 1974 1963 1979 1970 1991. Vypočítejte 85. percentil, 15 percentil a medián.

Příklad 13.: SCIO testů na základní škole se zúčastnilo 45 žáků devátého ročníku. Za každou správnou odpověď žáci získávají 1 bod, za špatnou odpověď se jim část bodu odečítá. Test se skládá celkem ze 40 otázek, což je i maximální možný zisk bodů. Výsledné body jednotlivých žáků jsou zaznamenány v následující tabulce 19:

Tabulka 19: Výsledky SCIO testu

Žák	Body	Žák	Body	Žák	Body
1.	45	16.	28,25	31.	15
2.	43,5	17.	28	32.	14,75
3.	42	18.	27	33.	14,5
4.	40	19.	26,75	34.	13,25
5.	39,25	20.	26,5	35.	13
6.	39	21.	25,5	36.	12,75
7.	37,5	22.	24,25	37.	11,5
8.	35	23.	23	38.	9
9.	34,75	24.	22,75	39.	8,75
10.	34,25	25.	22,5	40.	8,25
11.	32	26.	22	41.	6,75
12.	31,75	27.	21	42.	5
13.	31,75	28.	19,75	43.	4,25
14.	30,25	29.	18	44.	2,75
15.	30	30.	17,5	45.	1

Vypočítejte 85 a 33 percentil.

Příklad 14.: Nejmenovaná firma má tři pobočky ve třech různých městech. Ke každé pobočce udává počet zaměstnanců a průměrnou měsíční mzdu zaměstnanců dané pobočky v Kč. Údaje jsou uvedeny v následující tabulce 20:

Tabulka 20: Počet zaměstnanců s průměrnou mzdou

Pobočka	Počet zaměstnanců	Průměrná mzda
Pobočka A	15	25 432
Pobočka B	22	27 671
Pobočka C	9	21 899

Vypočítejte průměrný měsíční příjem celé firmy.

Výsledky

Příklad 1.:

$$\bar{x} = \frac{1500 + 750 + 3700 + 1200 + 2350 + 450 + 3000 + 5200 + 700 + 1100}{10}$$

$$\bar{x} = 1995 \text{ Kč}$$

Příklad 2.:

$$\bar{x} = \frac{5,6 + 12,7 + 9,4 + 3,4 + 5,6 + 10,4 + 10,8 + 17,7 + 6,9}{9}$$

$$\bar{x} = 8,25 \text{ °C}$$

Příklad 3.:

$$\bar{x} = \frac{8 \cdot 1 + 6 \cdot 2 + 10 \cdot 3 + 2 \cdot 4 + 1 \cdot 5}{27}$$

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

$$\text{Mod}(x) = 3$$

$$\text{Med}(x) = x_{\left(\frac{27+1}{2}\right)}$$

$$\text{Med}(x) = x_{14} = 2$$

Příklad 4.:

$$\bar{z}_G = \sqrt[10]{7,3 \cdot 6,7 \cdot 7 \cdot 7 \cdot 6,1 \cdot 5 \cdot 4 \cdot 2,9 \cdot 2,2 \cdot 2}$$

$$\bar{z}_G = 4,54$$

Příklad 5.:

$$\bar{x}_H = \frac{1}{\frac{1}{2} \cdot \left(\frac{1}{42} + \frac{1}{77}\right)}$$

$$\bar{x}_H \doteq 54,35 \text{ km/h}$$

Příklad 6.:

$$a) \bar{x}_H = \frac{1}{\frac{1}{4} \cdot \left(\frac{1}{56} + \frac{1}{70} + \frac{1}{81} + \frac{1}{43} \right)}$$

$$\bar{x}_H = 59 \text{ dní}$$

$$b) t = \frac{1}{\frac{1}{56} + \frac{1}{70} + \frac{1}{81} + \frac{1}{43}}$$

$$t = 14,8 \text{ dne}$$

Příklad 7.:

$$\bar{z}_G = \sqrt[6]{2 \cdot 4,3 \cdot 2,3 \cdot 4,6 \cdot 2,9 \cdot 2,4}$$

$$\bar{z}_G = 2,93$$

Příklad 8.:

$$\bar{x}_H = \frac{12 + 5 + 3}{12 \cdot \frac{1}{4} + 5 \cdot \frac{1}{10} + 3 \cdot \frac{1}{6}}$$

$$\bar{x}_H = 4 \text{ minuty}$$

Příklad 9.:

$$\text{Pro } p = 33 \text{ je } r = \frac{33}{100} (11 + 1)$$

$$\tilde{x}_{33} = 3,96 \doteq 4$$

Hledaný 33. percentil se nachází na 4. pozici uspořádaného souboru výšek, proto 33. percentil je roven 172 centimetrům, což znamená, že 33 % dívek soutěže je buď menších nebo rovno 172 centimetrům a 67 % dívek je buď větších nebo stejné výšky.

Příklad 10.:

$$\bar{z}_G = \sqrt[2]{1,15 \cdot 1,3} = \sqrt[2]{1,495}$$

$$\bar{z}_G \doteq 1,22$$

Příklad 11.:

$$\text{Pro } p = 25 \text{ je } r = \frac{25}{100} (9 + 1) = 2,5$$

$$\tilde{x}_{25} = 2,55$$

$$\text{Pro } p = 75 \text{ je } r = \frac{75}{100} (9 + 1) = 7,5$$

$$\tilde{x}_{75} = 4,35$$

Příklad 12.:

$$\text{Pro } p = 85 \text{ je } r = \frac{85}{100} (25+1)$$

$$\tilde{x}_{85} = \tilde{x}_{22} = 1991$$

$$\text{Pro } p = 15 \text{ je } r = \frac{15}{100} (25+1)$$

$$\tilde{x}_{15} = \tilde{x}_4 = 1968$$

$$\text{Med}(x) = x_{\left(\frac{25+1}{2}\right)}$$

$$\text{Med}(x) = x_{13} = 1976$$

Příklad 13.:

$$\text{Pro } p = 85 \text{ je } r = \frac{85}{100} (45+1)$$

$$\tilde{x}_{85} \doteq 39$$

$$\text{Pro } p = 33 \text{ je } r = \frac{33}{100} (45+1)$$

$$\tilde{x}_{33} \doteq 15$$

Příklad 14.:

$$\bar{x} = \frac{25\,432 \cdot 15 + 27\,671 \cdot 22 + 21\,899 \cdot 9}{15 + 22 + 9}$$

$$\bar{x} = 25\,811,58 \text{ Kč}$$

4. Charakteristiky variability

V této kapitole se budeme zabývat charakteristikami variability, které nám pomůžou podrobněji popsat statistický soubor. Charakteristiky variability jsou čísla, které nám charakterizují míru kolísání neboli proměnlivost hodnot ve statistickém znaku. Mezi charakteristiky variability řadíme rozpětí, mezikvartilové rozpětí, mezikvartilová odchylka, rozptyl, směrodatnou odchylku a variační koeficient (Robová a kol., 2013).

4.1. Rozpětí

Rozpětím chápeme rozdíl mezi maximální hodnotou a minimální hodnotou statistického znaku, značíme jej velkým písmenem R . Rozpětí tedy můžeme zapsat jako $R = x_n - x_1$, kdy x_n je poslední neboli maximální hodnota v uspořádaném souboru dat a x_1 je první neboli minimální hodnota v uspořádaném souboru dat. Nevýhodou této charakteristiky je, že je závislá pouze na první a poslední hodnotě znaku, ostatní hodnoty nejsou vůbec promítnuty do rozpětí. Tudíž se nám v rozpětí dost často projevují extrémní (odlehlé) hodnoty v souboru dat (Robová a kol., 2013, str. 175).

Vzorový příklad: Z údajů, na které můžete nahlédnout do tabulky 16, vypočítejte rozpětí.

Řešení: Připomeneme si tabulku, která vypadá následovně:

Tabulka 21: Výchozí data k výpočtu rozpětí

Hodnota znaku	Absolutní četnost	Kumulativní absolutní četnost
1	3	3
2	6	9
3	1	10
4	2	12
5	12	24
6	4	28
7	4	32
8	6	38
9	8	46
10	15	61
11	7	68
12	2	70
13	1	71
14	7	78
15	3	81
Celkem	81	/

Nejprve se musíme podívat, jestli máme data v souboru uspořádaná podle velikosti. Což na první pohled vidíme, že máme, proto výpočet rozpětí bude pouhým dosazením do vztahu: $R = x_n - x_1$. V našem případě dostaneme: $R = 15 - 1 = 14$. Rozpětí se tedy v našem příkladě rovná 14.

4.2. Mezikvartilové rozpětí

Mezikvartilové rozpětí značíme *MKR* a chápeme jím rozdíl mezi horním a dolním kvantilem. Mezikvartilové rozpětí můžeme zapsat: $MKR = \tilde{x}_{75} - \tilde{x}_{25}$. Mezikvartilové rozpětí se používá, když jsou ve statistickém znaku extrémní (odlehle) hodnoty (Robová a kol., 2013, str. 175).

Vzorový příklad: Z údajů, na které můžete nahlédnout do tabulky 16, vypočítejte mezikvartilové rozpětí.

Řešení: Připomeneme si tabulku, která vypadá následovně:

Tabulka 22: Data potřebná pro výpočet mezikvartilového rozpětí

Hodnota znaku	Absolutní četnost	Kumulativní absolutní četnost
1	3	3
2	6	9
3	1	10
4	2	12
5	12	24
6	4	28
7	4	32
8	6	38
9	8	46
10	15	61
11	7	68
12	2	70
13	1	71
14	7	78
15	3	81
Celkem	81	/

Abychom mohli vypočítat mezikvartilové rozpětí, potřebujeme znát hodnotu dolního a horního kvartilu. Dolní a horní kvartil už umíme vypočítat z předešlé kapitoly.

Pro $p = 25$ je $r = \frac{25}{100} \cdot (81 + 1) = 20,5$; tedy $c = 20$ a $z = 5$.

Potom $\tilde{x}_{25} = x_{20} + 0,5 \cdot [x_{21} - x_{20}] = 5 + 0,5 \cdot (5-5) = 5$.

Pro $p = 75$ je $r = \frac{75}{100} \cdot (81 + 1) = 61,5$; tedy $c = 61$ a $z = 5$.

Potom $\tilde{x}_{75} = x_{61} + 0,5 \cdot [x_{62} - x_{61}] = 10 + 0,5 (11-10) = 10,5$.

Teď už jen hodnoty \tilde{x}_{25} a \tilde{x}_{75} dosadíme do vztahu: $MKR = \tilde{x}_{75} - \tilde{x}_{25}$. V našem případě tedy dostaneme:

$MKR = 10,5 - 5 = 4,5$. Mezikvartilové rozpětí se v tomto příkladě rovná 4,5.

4.3. Mezikvartilová odchylka

Mezikvartilovou odchylku značíme $Q(x)$ a rozumíme ji jako polovinu rozdílu mezi horním a dolním kvartilem. Mezikvartilovou odchylku můžeme zapsat: $Q(x) = \frac{\tilde{x}_{75} - \tilde{x}_{25}}{2}$. Mezikvartilová odchylka se používá, když jsme pro charakteristiku polohy zvolili medián, a když se ve statistickém znaku vyskytují extrémní (odlehlé) hodnoty.

Vzorový příklad: Ptali jsme se 15 lidí na jejich čistý měsíční příjem v tisících korun. Odpovědi byly následující: 19 23 29 15 22 35 100 41 32 17 24 33 30 40 37. Vypočítejte z údajů mezikvartilovou odchylku.

Řešení: Jako první musíme soubor dat uspořádat od nejmenší po největší hodnotu znaku. Uspořádaný soubor dat bude vypadat následovně: 15 17 19 22 23 24 29 30 32 33 35 37 40 41 100. Když máme uspořádaný soubor, tak musíme vypočítat dolní a horní kvartil. Pro $p = 25$ je $r = \frac{25}{100} \cdot (15 + 1) = 4$, tedy $\tilde{x}_{25} = x_4 = 22$. Pro $p = 75$ je $r = \frac{75}{100} \cdot (15 + 1) = 12$, tedy $\tilde{x}_{75} = x_{12} = 37$. Teď už jenom chybí dosadit vypočítaný dolní a horní kvartil do vztahu pro mezikvartilovou odchylku. V našem případě bude vztah vypadat následovně: $Q(x) = \frac{22+37}{2} = 29,5$. Mezikvartilová odchylka se rovná 29,5.

4.4. Rozptyl

Rozptyl používáme jako charakteristiku variability v případě, že jsme pro charakteristiku polohy zvolili aritmetický průměr. Rozptylem rozumíme průměr druhých mocnin odchylek od zmiňovaného aritmetického průměru. Rozptyl značíme symbolem s_x^2 a zapíšeme ho vztahem:

$s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$. Pokud bychom chtěli počítat rozptyl

z tabulky rozdělení četností, pak rozptyl zapíšeme jako: $s_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$

(Caldá, Dupáč, 2008, str. 143).

Vzorový příklad: Následující tabulka 23 nám znázorňuje počet prodaných zmrzlin ve dvou stáncích A a B během pěti po sobě jdoucích pracovních dnů.

Tabulka 23: Počet prodaných zmrzlin ve stáncích

	Pondělí	Úterý	Středa	Čtvrtek	Pátek
Stánek A	34	40	55	19	27
Stánek B	22	34	64	29	26

Vypočítejte rozptyl prodaných zmrzlin ve stáncích.

Řešení: Nejprve musíme vypočítat aritmetický průměr pro obě linky.

$$\bar{x}_A = \frac{34 + 40 + 55 + 19 + 27}{5} = 35 \qquad \bar{x}_B = \frac{22 + 34 + 64 + 29 + 26}{5} = 35$$

Vidíme, že je aritmetický průměr pro oba stánky stejný a je roven 35. Teď vypočítáme

podle vztahu: $s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ rozptyl pro každý stánek zvlášť.

Stánek A:

$$s_x^2 = \frac{1}{5} [(34-35)^2 + (40-35)^2 + (55-35)^2 + (19-35)^2 + (27-35)^2]$$

$$s_x^2 = \frac{1}{5} (1 + 25 + 400 + 256 + 64)$$

$$s_x^2 = 149,2$$

Stánek B:

$$s_x^2 = \frac{1}{5} [(22-35)^2 + (34-35)^2 + (64-35)^2 + (29-35)^2 + (26-35)^2]$$

$$s_x^2 = \frac{1}{5} (169 + 1 + 841 + 36 + 81)$$

$$s_x^2 = 225,6$$

Z výsledků je patrné, že z hlediska rozptylu je variabilita stánku B daleko větší než u stánku A.

4.5. Směrodatná odchylka

Směrodatná odchylka je druhou odmocninou rozptylu. Největší výhodou směrodatné odchylky je, že výsledek je uveden v těch samých jednotkách, jako jsou uvedeny hodnoty

znaku. Značíme symbolem s_x a zapíšeme vztahem:
$$s_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$
.

Směrodatnou odchylku tedy můžeme chápat, že má stejný rozměr jako znak. (Caldá, Dupáč, 2008, str. 144).

Vzorový příklad: Z údajů, na které můžete nahlédnout v tabulce 23, vypočítejte směrodatnou odchylku.

Řešení: Připomeneme si tabulku, která vypadá následovně:

Tabulka 24: Data potřebná pro výpočet směrodatné odchylky

	Pondělí	Úterý	Středa	Čtvrtek	Pátek
Stánek A	34	40	55	19	27
Stánek B	22	34	64	29	26

Práci nám usnadní výsledky rozptylů, které jsme počítali v předchozím příkladě. Pro stánek A nám vyšel rozptyl: $s_x^2 = 149,2$ a pro stánek B $s_x^2 = 225,6$. Z definice víme, že směrodatná odchylka se vypočítá jako druhá odmocnina rozptylu.

Stánek A:

$$s_x = \sqrt{s_x^2}$$

$$s_x = \sqrt{149,2} = 12,21$$

Stánek B:

$$s_x = \sqrt{225,6} = 15,01$$

4.6. Variační koeficient

Variační koeficient značíme v_x a udává podíl směrodatné odchylky od aritmetického průměru. Variační koeficient zpravidla udáváme v procentech a zapíšeme ho vztahem:

$$v_x = \frac{s_x}{\bar{x}} \cdot 100 \%$$
. Variační koeficient používáme, pokud chceme použít charakteristiku

variability znaku vyjádřenou bezrozměrným číslem, avšak musíme dávat pozor, jestli hodnoty znaku x nabývají pouze nezáporných hodnot. V případě, že hodnoty znaku x nabývají záporných hodnot, pak variační koeficient nemá smysl. (Calda, Dupač, 2008, str. 144).

Vzorový příklad: Z údajů, na které můžete nahlédnout v tabulce 23, vypočítejte variační koeficient. K výpočtu využijte výsledků z předešlých výpočtů.

Řešení: Připomeneme si tabulku, která vypadá následovně:

Tabulka 25: Data potřebná pro výpočet variačního koeficientu

	Pondělí	Úterý	Středa	Čtvrtek	Pátek
Stánek A	34	40	55	19	27
Stánek B	22	34	64	29	26

Z předešlých výpočtů k tomuto příkladu už víme, že aritmetický průměr se pro oba stánky rovná 35. Směrodatná odchylka pro stánek A se rovná $s_x = 12,21$ a směrodatná odchylka pro stánek B se rovná $s_x = 15,01$. Variační koeficient vypočítáme jako podíl směrodatné odchylky od aritmetického průměru.

Stánek A:

$$v_x = \frac{s_x}{\bar{x}} \cdot 100 \%$$

$$v_x = \frac{12,21}{35} \cdot 100 \%$$

$$v_x \doteq 34,9$$

Stánek B:

$$v_x = \frac{s_x}{\bar{x}} \cdot 100 \%$$

$$v_x = \frac{15,01}{35} \cdot 100 \%$$

$$v_x \doteq 42,9$$

Příklady k procvičení

Příklad 1.: Z příložené tabulky 12, vypočítejte rozpětí, mezikvartilové rozpětí a mezikvartilovou odchylku ušitých šatů.

Tabulka 26: Hodnoty potřebné pro výčtu k příkladu číslo 1

Velikost šatů	Absolutní četnost	Kumulativní četnost	Relativní četnost v %	Kumulativní relativní četnost v %
34	6	6	10,2	10,2
36	14	20	23,7	33,9
38	12	32	20,3	54,2
40	13	45	22	76,2
42	8	53	13,6	89,8
44	4	57	6,8	96,6
46	2	59	3,4	100
Celkem	59	/	100	/

Příklad 2.: Při namátkové kontrole 100 výrobků pracovníci zjistili, že 77 výrobků nemá žádnou závadu, osm výrobků má jednu závadu, pět výrobků má dvě závady, devět výrobků má tři závady a jeden výrobek má závad dokonce pět. Určete směrodatnou odchylku a variační koeficient závad výrobků.

Příklad 3.: Vraťme se k *příkladu 2* z kapitoly 3. Pan Veselý si během dubna dvakrát týdně zapisoval venkovní teplotu. Naměřené údaje jsou ve °C a jsou následující: 5,6; 12,7; 9,4; 3,4; 5,6; 10,4; 10,8; 17,7 a 6,9. Z předchozího počítání víte, že průměrná teplota za měsíc duben byla 8,25 °C. Vypočítejte směrodatnou odchylku a variační koeficient.

Příklad 4.: Vraťme se k *příkladu 9* z kapitoly 3. Soutěže krásy se zúčastnilo 11 dívek, u kterých jsme měřili jejich výšku. Měřením jsme získali následující výsledky, které jsou vyjádřené v centimetrech: 171 175 183 170 172 175 166 178 180 175 187. Vypočítejte rozpětí, mezikvartilové rozpětí a mezikvartilovou odchylku.

Výsledky

Příklad 1.:

$$R = 46-34$$

$$R = 12$$

$$MKR = \tilde{x}_{75} - \tilde{x}_{25}$$

$$\text{Pro } p = 25 \text{ je } r = \frac{25}{100} \cdot (59 + 1) = 15, \tilde{x}_{25} = x_{15} = 36$$

$$\text{Pro } p = 75 \text{ je } r = \frac{75}{100} \cdot (59 + 1) = 45, \tilde{x}_{75} = x_{45} = 40$$

$$MKR = 40-36$$

$$MKR = 4$$

$$Q(x) = \frac{40-36}{2}$$

$$Q(x) = 2$$

Příklad 2.:

$$\bar{x} = \frac{77 \cdot 0 + 8 \cdot 1 + 5 \cdot 2 + 9 \cdot 3 + 1 \cdot 5}{100} = 0,5$$

$$s_x^2 = \frac{1}{100} [77 \cdot (0-0,5)^2 + 8 \cdot (1-0,5)^2 + 5 \cdot (2-0,5)^2 + 9 \cdot (3-0,5)^2 + 1 \cdot (5-0,5)^2]$$

$$s_x^2 = 1,09$$

$$s_x = \sqrt{1,09}$$

$$s_x = 1,04$$

$$v_x = \frac{1,04}{0,5} \cdot 100 \%$$

$$v_x = 208 \%$$

Příklad 3.:

$$s_x^2 = \frac{1}{9} \left[(5,6-8,25)^2 + (12,7-8,25)^2 + (9,4-8,25)^2 + (3,4-8,25)^2 + (5,6-8,25)^2 + \right. \\ \left. + (10,4-8,25)^2 + (10,8-8,25)^2 + (17,7-8,25)^2 + (6,9-8,25)^2 \right]$$

$$s_x^2 = 17,88$$

$$s_x = \sqrt{17,88}$$

$$s_x = 4,23$$

$$v_x = \frac{4,23}{8,25} \cdot 100 \%$$

$$v_x = 51,27 \%$$

Příklad 4.:

$$R = 187-166$$

$$R = 21$$

$$MKR = \tilde{x}_{75} - \tilde{x}_{25}$$

$$\text{Pro } p = 25 \text{ je } r = \frac{25}{100} \cdot (11 + 1) = 3, \tilde{x}_{25} = x_3 = 171$$

$$\text{Pro } p = 75 \text{ je } r = \frac{75}{100} \cdot (11 + 1) = 9, \tilde{x}_{75} = x_9 = 180$$

$$MKR = 180-171$$

$$MKR = 9$$

$$Q(x) = \frac{180-171}{2}$$

$$Q(x) = 4,5$$

5. Závěr

Cílem předkládané bakalářské práce bylo sestavit komplexní sbírku příkladů z vybraných témat vědní disciplíny statistika. Soustředila jsem se na témata vymezená v RVP pro různé typy středních škol. Příklady byly voleny tak, aby popisovaly problémy běžného života. Při zpracování sbírky jsem řadu témat konzultovala s učiteli středních škol a dospěla jsem k závěru, že na středních školách není tomuto okruhu matematiky věnován dostatek času. Dle mého názoru tato situace není příznivá pro vzdělání žáků středních škol, protože se statistikou se setkávají téměř každý den, aniž by si to uvědomovali. Například v současné době tak aktuální sčítání obyvatelstva.

Doufám, že tato práce pomůže žákům středních škol lépe pochopit některá témata z oblasti statistiky, se kterými se setkali ve škole.

6. Literatura a zdroje:

1. BOUCNÍK, Pavel, Jiří HERMAN, Peter KRUPKA a Jaromír ŠIMŠA. Odmaturuj! z matematiky 3: sbírka řešených příkladů. Vyd. 1. Brno: Didaktis, 2004, 248 s., ISBN 80-7358-010-1.
2. CALDA, Emil a Václav DUPAČ. Kombinatorika, pravděpodobnost, statistika. Učebnice pro gymnázia. Praha: Prométheus, 1993, 170 s., ISBN 978-80-7196-365-3.
3. Český hydrometeorologický ústav; dostupné z <https://www.chmi.cz/>
4. Český statistický úřad; dostupné z <https://www.czso.cz/>
5. KUBEŠOVÁ, Naděžda a Eva CIBULKOVÁ. Matematika: přehled středoškolského učiva. 2. vyd. Třebíč: Petra Velanová, 2006, 239 s. Maturita (Petra Velanová), ISBN 978-80-86873-05-3.
6. PETRÁNEK, Oldřich, Emil CALDA a Petr HEBÁK. Matematika pro střední odborné školy a studijní obory středních odborných učilišť. 6. vyd. č. 4. Praha: Prométheus, 2011, 148 s. Učebnice pro střední školy (Prometheus), ISBN 978-80-7196-407-0.
7. ROBOVÁ, Jarmila, Martin HÁLA a Emil CALDA. Komplexní čísla, kombinatorika, pravděpodobnost a statistika. Matematika pro střední školy. 1.vyd. Praha: Prométheus, 1900, 235 s., ISBN 978-80-7196-425-4.