

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra statistiky



Bakalářská práce

**Návrh a statistické vyhodnocení dotazníku pro průzkum
trhu**

Tereza Hadravová

© 2013 ČZU v Praze

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Katedra statistiky
Provozně ekonomická fakulta

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Hadravová Tereza

Provoz a ekonomika

Název práce

Návrh a statistické vyhodnocení dotazníku pro průzkum trhu

Anglický název

Market research questionnaire design and statistical analysis

Cíle práce

Cílem práce je pojednat o možnostech, jak se tvoří dotazník a následně správně statisticky zpracuje. V teoretické části proto budou přehledně představeny teoretické přístupy k dané problematice. V navazující části praktické budou poznatky aplikovány na konkrétní dotazníkové šetření. Bude navržen dotazník, který bude sloužit ke zkoumání postojů středoškolských studentů ke studiu. Následně bude dotazník distribuován mezi respondenty a výsledky statisticky zpracovány.

Metodika

Práce se bude zabývat výběrovým dotazníkovým šetřením. V první části budou představena teoretická východiska na základě analýzy odborné literatury. Ve druhé části bude proveden vlastní průzkum, kterého se zúčastní zhruba 150 žáků 4-letých maturitních oborů ze SOŠ a SOU Kladno. Data budou zpracována pomocí programu IBM SPSS Statistics 20. Výsledky budou představeny na výstupech programu, tzn. na tabulkách a grafech. Metody budou zvoleny v závislosti na typu proměnných a typu pravděpodobnostního rozdělení. Výsledky statistické analýzy budou komentovány.

Harmonogram zpracování

Literární rešerše: 02/2012 - 11/2012

Vytvoření dotazníku a příprava průzkumu: 12/2012 - 01/2013

Provedení průzkumu a vyhodnocení dat: 02/2013

Rozsah textové části

30 - 40 stran

Klíčová slova

výběrové dotazníkové šetření, návrh dotazníku, statistická analýza, postoje středoškolských studentů ke studiu

Doporučené zdroje informací

KOZEL, Roman. Moderní marketingový výzkum: nové trendy, kvantitativní a kvalitativní metody a techniky, průběh a organizace, aplikace v praxi, přínosy a možnosti. 1. vyd. Praha: Grada, 2006, 277 s. ISBN 80-247-0966-X.

PECÁKOVÁ, Iva. Statistika v terénních průzkumech. 2. dopl. vyd. Praha:

Professional Publishing, 2011, 236 s. ISBN 978-80-7431-039-3.

PECÁKOVÁ, Iva, Ija NOVÁK a Jan HERZMANN. Pořizování a vyhodnocování dat ve výzkumech veřejného mínění. Vyd. 3. Praha: Oeconomica, 2004, 146 s. ISBN 80-245-0753-6.

RAAB-STEINER, Elisabeth a Michael BENEŠCH. Der Fragebogen: Von der Forschungsidee zur SPSS / PASW-Auswertung. 2., aktualisierte Aufl. Wien: UTB, 2005. ISBN 978-382-5284-060.

ŘEZANKOVÁ, Hana. Analýza dat z dotazníkových šetření. 3. aktualiz. vyd. Praha: Professional Publishing, 2011, 223 s. ISBN 978-807-4310-621.

Vedoucí práce

Nešetřilová Helena, doc. RNDr., CSc.

Termín odevzdání

březen 2013

doc. RNDr. Bohumil Kába, CSc.
Vedoucí katedry



prof. Ing. Jan Hron, DrSc., dr.h.c.
Děkan fakulty

V Praze dne 20.2.2013

Čestné prohlášení

Prohlašuji, že svou bakalářskou práci „Návrh a statistické vyhodnocení dotazníku pro průzkum trhu“ jsem vypracovala samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autorka uvedené bakalářské práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušila autorská práva třetích osob.

V Praze dne 14.3.2013



Poděkování

Ráda bych touto cestou poděkovala vedoucí své bakalářské práce doc. RNDr. Heleně Nešetřilové CSc. za její ochotu, vstřícný přístup a cenné rady. Děkuji také vedení školy SOŠ a SOU Kladno za pomoc při provádění dotazníkového šetření a v neposlední řadě i žákům, kteří dotazníky vyplnili.

Market research questionnaire design and statistical analysis

Souhrn

Práce se zabývá možnostmi návrhu dotazníku a jeho následného statistického vyhodnocení v rovině teoretické i praktické. Teoretická východiska popisují proces tvorby dotazníku od formulace výzkumné otázky, přes výběr respondentů, až po strukturu samotného výzkumného nástroje se zaměřením na otázky a typ proměnných a škál. Na návrh navazuje pojednání o zpracování dat získaných v rámci šetření. Z široké škály statistických metod jsou vybrány ty, které jsou vhodné pro výběrové dotazníkové šetření. Jedná se o nástroje statistiky popisné (tabulky rozdělení četností, míry polohy, variability a koncentrace) a především analytické (kontingence, korelace, neparametrické testy – Mannův-Whitneyho U-test a Kruskalův-Wallisův test). Část s vlastním zpracováním prezentuje dotazník navržený pro žáky SOŠ a SOU Kladno a výsledky šetření, které proběhlo v prvním týdnu února 2013, a které zkoumalo postoje žáků ke škole, maturitě a hlavně přípravě do školy. Data byla zpracována pomocí software *IBM SPSS Statistics 20*. Výsledky jsou přehledně prezentovány pomocí tabulek a grafů, ke kterým je vždy připojen komentář a informace týkající se vybrané metody a jejích předpokladů.

Summary

This bachelor thesis deals with the different options of questionnaire design and its subsequent statistical analysis in theory as well as in practice. The theoretical part describes the process of the questionnaire design from the formulation of a research question over the selection of respondents to the structure of the questionnaire itself. It focuses on the various types of questions, variables and scales. Following the design, there is a part explaining how to analyse the data won from the survey. Out of the wide range of statistical methods, it concentrates on those which are suitable for a sample survey with a questionnaire. In particular, it deals with descriptive methods (frequencies tables, measures of central tendency, variability and concentration) and most importantly analytical methods (contingency, correlation, non-parametric tests – Mann-Whitney U-test and Kruskal-Wallis test). The next part presents a questionnaire designed for high school students from a vocational school in Kladno (SOŠ a SOU Kladno) and the results of the sample survey which was carried out in the first week of February 2013. The questionnaire examined students' attitudes towards their school, the final leaving exam and their preparations for school. The data were analysed using *IBM SPSS Statistics 20*. The outcomes are presented in tables and diagrams followed by comments and information regarding the chosen method and its requirements.

Klíčová slova: výběrové dotazníkové šetření, návrh dotazníku, statistická analýza, postoje středoškolských studentů ke studiu

Keywords: sample survey, questionnaire design, statistical analysis, high school students' attitudes towards their studies

Obsah

1	Úvod.....	11
2	Cíl práce a metodika	12
3	Teoretická východiska	13
3.1	Co je to dotazník	13
3.2	Návrh dotazníku.....	14
3.2.1	Formulace výzkumné otázky	14
3.2.2	Výběrový soubor.....	15
3.2.3	Struktura dotazníku.....	18
3.3	Zpracování dotazníku	22
3.3.1	Vytvoření kódového klíče.....	23
3.3.2	Čištění dat	23
3.3.3	Popis datového souboru	24
3.3.4	Analýza závislostí	27
3.3.5	Porovnávání skupin – Neparametrické testy pro nezávislé výběry	30
4	Vlastní zpracování a zhodnocení výsledků.....	33
4.1	Dotazník pro žáky SOŠ a SOU Kladno	33
4.2	Zpracování dat	34
4.2.1	Struktura výběrového souboru.....	34
4.2.2	Charakteristika jednotlivých proměnných	36
4.2.3	Analýza závislostí	44
4.2.4	Porovnávání skupin.....	51
5	Závěr	57
	Seznam použitých zdrojů.....	59
	Přílohy.....	61

Seznam použitých tabulek

Tabulka 4.1: Celkový počet respondentů	35
Tabulka 4.2: Rozdělení četností u proměnné Pohlaví	35
Tabulka 4.3: Rozdělení četností u proměnné Ročník studia	35
Tabulka 4.4: Míry polohy, variability a koncentrace pro proměnnou Doba přípravy do školy.....	36
Tabulka 4.5: Rozdělení četností pro proměnnou Kategorie doby přípravy do školy	36
Tabulka 4.6: Rozdělení četností u proměnné Tendence v intenzitě učení.....	37
Tabulka 4.7: Absolutní četnosti využití jednotlivých míst k přípravě do školy	38
Tabulka 4.8: Míry polohy a variability pro proměnnou Spokojenost se studijními výsledky	40
Tabulka 4.9: Míry polohy a variability pro proměnnou Obavy z maturity	41
Tabulka 4.10: Rozdělení četností pro proměnnou Příprava na maturitu	42
Tabulka 4.11: Míry polohy, variability a koncentrace pro proměnnou Doba přípravy do školy.....	43
Tabulka 4.12: Rozdělení četností pro kategorizovanou proměnnou Doba cesty do školy..	43
Tabulka 4.13: Výsledky Chí-kvadrát testu o nezávislosti pro proměnné Pohlaví a Tendence v intenzitě učení	44
Tabulka 4.14: Výsledky Chí-kvadrát testu o nezávislosti pro proměnné Pohlaví a Příprava doma.....	45
Tabulka 4.15: Hodnoty Cramérova V a ϕ pro proměnné Pohlaví a Příprava doma.....	45
Tabulka 4.16: Kontingenční tabulka pro proměnné Pohlaví a Příprava na maturitu	46
Tabulka 4.17: Výsledky Chí-kvadrát testu o nezávislosti pro proměnné Ročník a Příprava doma (rekódovaná)	47
Tabulka 4.18: Výsledky Fisherova exaktního testu pro proměnné Pohlaví a Příprava u známých nebo kamarádů (rekódovaná)	49
Tabulka 4.19: Kontingenční tabulka s adjustovanými rezidui pro proměnnou Pohlaví a zrekódovanou proměnnou Příprava u známých nebo kamarádů	49
Tabulka 4.20: Hodnoty Cramérova V a ϕ pro proměnnou Pohlaví a zrekódovanou proměnnou Příprava u známých nebo kamarádů.....	49

Tabulka 4.21: Kendallovo τ_b pro vybrané dvojice proměnných (zeleně jsou označeny statisticky významné vztahy na hladině významnosti 1%, žlutě na hladině významnosti 5%).....	50
Tabulka 4.22: Testy normality pro proměnnou Doba přípravy do školy	50
Tabulka 4.23: Testy normality pro proměnnou Doba přípravy do školy podle ročníku studia.....	51
Tabulka 4.24: Levenův test na homogenitu rozptylů proměnné Doba přípravy do školy v jednotlivých ročnících.....	51
Tabulka 4.25: Výsledky Kruskalova-Wallisova testu pro proměnnou Doba přípravy do školy podle ročníku studia	52
Tabulka 4.26: Výsledky Mannova-Whitneyho testu pro 1. a 4. ročník.....	52
Tabulka 4.27: Výsledky Mannova-Whitneyho testu pro 2. a 4. ročník.....	52
Tabulka 4.28: Výsledky Mannova-Whitneyho testu pro 3. a 4. ročník.....	52
Tabulka 4.29: Testy normality pro proměnnou Doba přípravy do školy podle pohlaví	53
Tabulka 4.30: Levenův test na homogenitu rozptylů proměnné Doba přípravy do školy v jednotlivých ročnících.....	53
Tabulka 4.31: Výsledky Mannova-Whitneyho testu pro proměnnou Doba přípravy do školy podle pohlaví.....	54
Tabulka 4.32: Výsledky Mannova Whitneyho testu pro proměnnou Obavy z maturity podle pohlaví	55
Tabulka 4.33: Výsledky Kruskalova-Wallisova testu pro proměnnou Obavy z maturity podle ročníku studia.....	55

Seznam použitých grafů

Graf 4.1: Rozdělení četností proměnné Doba přípravy do školy	37
Graf 4.2: Výšečový graf znázorňující, jaký způsob přípravy do školy vyhovuje žákům nejvíce.....	39
Graf 4.3: Výšečový graf pro proměnnou Vztah k oboru studia.....	39
Graf 4.4: Sloupcový graf pro proměnnou Spokojenost se studijními výsledky	40
Graf 4.5: Sloupcový graf pro proměnnou Obavy z maturity	41
Graf 4.6: Graf pro rozdělení četností proměnné Doba cesty do školy	42
Graf 4.7: Shlukový sloupcový graf pro přípravu doma podle pohlaví	45
Graf 4.8: Krabičkový graf zobrazující dobu přípravy do školy podle ročníku studia	52
Graf 4.9: Krabičkový graf zobrazující dobu přípravy do školy podle pohlaví.....	54
Graf 4.10: Krabičkový graf zobrazující obavy z maturity podle pohlaví	55
Graf 4.11: Krabičkový graf zobrazující obavy z maturity podle ročníku studia	55

1 Úvod

„Byl to odvážný počin, požádat, tak jako jsem to udělal v roce 1874 každého člena Královské společnosti, který zastával nějaký významný post, aby zodpověděl značné množství otázek potřebných pro můj úmysl, z nichž několik se týkalo náboženství a jiných delikátních záležitostí“,¹ popsal sir Francis Galton (1822-1911), který je považovaný za vynálezce dotazníku, první použití tohoto výzkumného instrumentu. Ačkoli si to Galton tehdy pravděpodobně neuvědomoval, vytvořil výzkumný nástroj, který je dodnes hojně využíván v různých oblastech; např. ve vědě (sociologie, psychologie), ale také ve sférách spíše komerčních, kde se pomocí dotazníků zkoumají názory a postoje zákazníků a i samotný trh. Galtona by jistě překvapily i moderní způsoby provádění průzkumů, které se nedrží jen tradiční papírové formy, ale využívají i telefon a nejnověji počítače a internet.²

Co ale vedlo k takovému rozmachu v používání a oblibě dotazníků a průzkumů? Iva Pecáková vidí příčinu v povaze člověka, jehož přirozeností je sbírat informace, aby na jejich základě mohl činit rozhodnutí. Tuto svoji tezi ilustruje Pecáková na příkladech z antického Řecka, kde se průzkum mínění uplatňoval v procesu ostrakizace, tedy vyhnání toho, kdo podle názoru občanů ohrožoval athénskou demokracii.³ V současnosti roste zájem o dotazníky a další formy průzkumu veřejného mínění a trhu v souvislosti s usilováním o tzv. informační společnost.⁴

Kupředu šel nejen vývoj dotazníků, ale také způsoby jejich zpracování. Díky výpočetní technice jsou doby ručního zpracování již několik desetiletí pryč.⁵ Data se dnes běžně vyhodnocují pomocí specializovaného statistického software; za jeden z nejrozšířenějších je považován program *IBM SPSS Statistics* upotřebený i v této bakalářské práci.⁶ Jak je vidět, dotazníky jsou velice živým výzkumným nástrojem, se kterým je i do budoucna potřeba počítat.

¹ GALTON, F. *Memories of my life*, s. 291. Předklad autorky z originálu: „It was a daring undertaking, to ask as I did, in 1874, every Fellow of the Royal Society who had filled some important post, to answer a multitude of Questions needful for my purpose, a few of which touched on religion and other delicate matters.“

² KOZEL, R. *Moderní marketingový výzkum: nové trendy, kvantitativní a kvalitativní metody a techniky, průběh a organizace, aplikace v praxi, přínosy a možnosti*. s. 142-145

³ PECÁKOVÁ, I. *Statistika v terénních průzkumech*. s. 9

⁴ viz např. web Rada vlády pro konkurenceschopnost a informační společnost. [online]. [cit. 5.3.2013]. Dostupné z: <http://www.vlada.cz/cz/ppov/rvis/rada-vlady-pro-informacni-spolecnost-73372/>

⁵ LEONHART, R. *Datenanalyse mit SPSS*. s. 14

⁶ PECÁKOVÁ, cit. d., s. 13

2 Cíl práce a metodika

Cílem práce je pojednat o způsobech tvorby dotazníku a o možnostech jeho následného metodicky správného vyhodnocení. Vhodným poskládáním poznatků by měly na povrch vyplynout nové zajímavé souvislosti nejen v oblasti teorie, ale také ve výsledcích z praktického příkladu šetření názorů středoškolských žáků. Záměrem autorky je představit v práci přehledně nejdůležitější součásti dotazníku, tj. otázky a typy proměnných a škál, a zároveň popsat dotazník jako celek, včetně metod výběru respondentů. Za účelem provedení názorné ukázky dotazníkového šetření se navrhne konkrétní dotazník, který bude zjišťovat postoje středoškolských žáků ke studiu, způsob, jakým přistupují k přípravě do školy a jejich spokojenost s oborem a studijními výsledky. Vytvořený dotazník se předloží žákům SOŠ a SOU Kladno, kteří budou představovat respondenty pro průzkum. Cílem je získaná data zpracovat, použít při tom vhodné metody a výsledky představit.

Jako podklady pro vlastní zpracování autorky byla sepsána teoretická východiska, na základě kterých bylo postupováno. Analyzována byla odborná statistická literatura, a kde to bylo možné, byly přístupy různých autorů kontrastovány. Mezi stěžejní publikace pro tuto práci patří díla autorů Andyho Fielda, dvojice Elisabeth Raab-Steinerové a Michaela Benesche, Ivy Pecákové a Hany Řezankové. Část vlastního zpracování je pojata jako výběrové dotazníkové šetření. Pro jeho provedení vznikl papírový dotazník, který vyplnilo během výuky v prvním týdnu nového pololetí 156 žáků 4-letých oborů SOŠ a SOU Kladno. Získaná data byla přepsána do programu *IBM SPSS Statistics 20*. Pomocí zmíněného programu byla provedena analýza metodami, které byly vybrány jako vhodné pro výběrové dotazníkové šetření a které jsou blíže popsány v části zabývající se teoretickými východisky. Jedná se tedy o popis souboru jako celku, o charakteristiku jednotlivých proměnných pomocí tabulek rozdělení četnosti, grafů a v některých případech jsou pro dobrou ilustraci použity též popisné charakteristiky. Na základě informací zjištěných z deskripce proměnných jsou vybrány oblasti pro zjištění vztahů mezi proměnnými (kontingence a korelace). V poslední části se nachází porovnávání vybraných skupin. Vždy jsou otestovány podmínky pro provedení testu (typ rozdělení, homogenita rozptylů, typ proměnných), vybrán vhodný test a jeho výsledky jsou interpretovány. V případě statisticky významných rozdílů a vztahů jsou výsledky doprovázeny grafem a další analýzou. Veškeré tabulky i grafy jsou opatřeny komentářem.

3 Teoretická východiska

3.1 Co je to dotazník

Na úvod je potřeba definovat pojem, který je pro účely této bakalářské práce stěžejní - dotazník. Jako nejvhodnější se nabízí definice autorů Raab-Steinerové a Benesche, která není globální, ale profiluje dotazník jako nástroj především výzkumů veřejného mínění a marketingových výzkumů: „Pro naše účely se dotazník týká mínění, názorů a postojů k tématům. Do výsledků se má promítnout pouze to, na co je osoba konkrétně dotázána.“⁷ Pecáková, Novák a Herzmann ve své publikaci klasifikují dotazník jako jeden z postupů zjišťování údajů. Rozlišují metody nstandardizované, kterých využívá kvalitativní výzkum, a standardizované, jejichž cílem je získat číselné údaje vhodné pro statistické operace. Dotazník řadí do skupiny standardizovaných způsobů, protože je pro něj typický shodný způsob získání údajů od všech dotázaných.⁸ Kromě dotazníku se mezi standardizované postupy řadí anketa a standardizovaný rozhovor. Posledně jmenovaný se po formální stránce blíží dotazníku, neboť se odehrává podle předem definovaného scénáře se standardizovanými otázkami i odpověďmi, ale za přítomnosti tazatele.⁹ Rozdíl mezi dotazníkem a anketou spočívá především v tom, jakým způsobem jsou předávány dotazovaným a následně od nich vybírány. „**Dotazník** je určen k **předání** pro vyplnění konkrétním osobám, od kterých se dotazníky vybírají, a to **organizovaně** tak, aby každý respondent dotazník určitě obdržel, resp. jej odevzdal.“¹⁰ U ankety do značné míry nezáleží výběr respondentů na výzkumníkovi, jelikož je anketní lístek distribuován velkému počtu osob, které ale anketu nemusí odevzdat. Typickým rysem je nízká návratnost a hlavně to, že není zcela jasné, kdo patří do základního souboru.¹¹ Roli zde hraje tzv. samovýběrový efekt.¹²

⁷ RAAB-STEINER, E., BENESCH, M. *Der Fragebogen*, s. 43. Překlad autorky z originálu: „Für unsere Zwecke geht es um die Erfassung von Meinungen, Einstellungen, Positionen zu Themen oder Sachverhalten. Der Fragebogen wird als Forschungsinstrument zu deren Erkundung eingesetzt. Es soll und kann in die Ergebnisse nur einfließen, worüber die Person konkret befragt wurde.“

⁸ PECÁKOVÁ, I., NOVÁK, I., HERZMANN, J. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. s. 23

⁹ NOVÝ, I., SURYNEK, A.. *Sociologie pro ekonomy a manažery*. s. 269

¹⁰ Tamtéž, s. 267

¹¹ Tamtéž, s. 267-268

¹² PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 23

3.2 Návrh dotazníku

3.2.1 Formulace výzkumné otázky

Na první pohled by se mohlo zdát, že dotazník je jen sada otázek, které výzkumníka napadly, a které bez většího rozmyslu sepsal do určité formy. Jak ale poznamenává ve své knize *Analýza dat z dotazníkových šetření* Hana Řezanková, nejde zdaleka o nějakou triviální záležitost.¹³ Právě naopak, za dotazníkem stojí vždy velké množství práce a rozmyšlení. Celý proces začíná formulováním výzkumné otázky, která dává smysl a účel celému dotazníku. Tento krok zdůrazňují autoři Raab-Steinerová a Benesch, podle kterých je zásadní vědět, co chceme vědět a tedy zkoumat.¹⁴ Nabízí také možnost, jak k takovému poznatku dojít. Při formulaci výzkumné otázky pro dotazník, který patří většinou do kvantitativního výzkumu, tzn., kvantifikuje jevy, je možno použít metod kvalitativního výzkumu a podněty získat např. v hloubkových rozhovorech s experty nebo s jednotlivci z populace, kterou zkoumáme.¹⁵ Na základě získaných informací a odborné literatury se pak definují relevantní proměnné, resp. znaky, na které se dotazník zaměří.¹⁶ „Toto pořadí není nedůležité, neboť redukuje nebezpečí, že se výzkum zabývá něčím lehce zjištělným a ne informacemi, které jsou k tématu opravdu potřeba.“ zdůrazňují v knize *Dotazník: od idey k vyhodnocení v SPSS* autoři Raab-Steinerová a Benesch.¹⁷

V předešlém odstavci byl zmíněn pojem populace, který je významný jak pro výzkumnou otázku, tak i pro návrh, zpracování a především vyhodnocení dotazníku. Populace neboli základní soubor (o populaci lze mluvit hlavně tehdy, týká-li se výzkum osob)¹⁸ jsou všechny statistické jednotky, zde tedy lidé, které jsou nositeli definovaných znaků, které u žádných jiných jednotek nelze najít. „Předmětem výzkumu jsou však takové znaky, v nichž se jednotlivé jednotky mohou lišit.“¹⁹ Pro potřeby výzkumu je tedy nutné vymezit základní soubor „[...] věcně, v čase i v prostoru [...]“.²⁰ Určí-li se základní soubor nesprávně, vyvstávají problémy se zobecňováním výsledků šetření, který bylo provedeno

¹³ ŘEZANKOVÁ, H. *Analýza dat z dotazníkových šetření*. s. 13

¹⁴ RAAB-STEINER, BENESCH, cit. d., s. 45

¹⁵ Tamtéž, s. 46

¹⁶ Tamtéž, s. 46

¹⁷ Tamtéž, s. 46. Překlad autorky z originálu: „Diese Reihenfolge ist nicht unwesentlich, da sie die Gefahr reduziert, in einer Untersuchung etwas zu erfragen, was vermutlich leicht abfragbar ist und nicht das, was an Information wirklich zum Thema brauchbar ist.“

¹⁸ PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 7

¹⁹ Tamtéž, s. 8

²⁰ PECÁKOVÁ, cit. d., s. 16

pouze na části populace, na celý základní soubor.²¹ Konkrétními příklady základních souborů mohou být všichni studenti České zemědělské univerzity, všichni předplatitelé časopisu Respekt nebo také všichni obyvatelé České republiky.

3.2.2 Výběrový soubor

Základní soubory jsou často velmi rozsáhlé a čítají tisíce (studenti ČZU) až miliony (obyvatelé ČR) osob. Z důvodů praktických, finančních i časových proto není výhodné prošetřit celou populaci, ale jenom její část – tzv. výběrový soubor, výběr nebo vzorek.²² Pecáková uvádí ještě další důvod: „Moderní statistika totiž disponuje postupy, které umožňují provádět za určitých podmínek pro základní soubor zobecnění skutečností zjištěných v jeho vzorku.“²³ Díky statistickým metodám je tedy možné dojít k výsledkům, které se v kvalitě dokážou vyrovnat vyčerpávajícímu šetření. Nicméně takovéto výsledky nezaručí jakýkoli výběrový soubor; je potřeba dbát na reprezentativnost výběru. Ve výběru by se měla odrážet struktura základního souboru tak, aby ho bylo možné označit jako „zmenšeninu“ populace.²⁴ Na druhou stranu se autoři Pecáková, Novák a Herzmann shodují s Raab-Steinerovou a Beneschem v tom, že dokonalé zmenšeniny dosáhnout nelze. I když totiž výběrový soubor odráží strukturu populace například podle věku, už tomu tak nemusí být podle pohlaví, vzdělání a dalších znaků. Je velmi obtížné, ne-li nemožné dosáhnout shody ve vícerozměrných strukturách.²⁵ Je to způsobeno také tím, že populaci neznáme, nemáme tudíž většinou informace o rozložení znaků v populaci. Přesto je snaha o co nejvěrnější „kopii“ základního souboru žádoucí, protože platí, že čím lépe odráží výběrový soubor populaci, tím přesněji lze na populaci usuzovat.²⁶

Velikost výběrového souboru

Významnou roli v přesnosti zobecňování na základní soubor hraje i rozsah výběrového souboru, tedy kolika respondentů je potřeba se dotázat. Různí autoři popisují různé možnosti. Zde bude popsán přístup, který ve své publikaci *Moderní marketingový výzkum* používá Radim Kozel. Ten rozlišuje tři odlišné přístupy: slepý odhad, nákladový a

²¹ KOZEL, R., MYNÁŘOVÁ, L., SVOBODOVÁ, H. *Moderní metody a techniky marketingového výzkumu*. s. 90

²² PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 7

²³ PECÁKOVÁ, cit. d., s. 16

²⁴ PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 61

²⁵ Tamtéž, s. 61-62

²⁶ RAAB-STEINER, BENESCH, cit. d., s. 16

statistický přístup. První jmenovaný je založen na zkušenostech výzkumníka, který si rozsah vzorku určí podle toho, s čím se v praxi setkal či co mu napovídá intuice. Nebezpečím je nedostatečná praxe, která zde implikuje horší výsledky. Nákladový přístup je vysoce pragmatický, neboť v centru jeho zájmu jsou náklady na jednoho respondenta. Celková suma peněz vyčleněná na provedení výzkumu se vydělí náklady na jednoho dotázaného a výsledek je rozsah výběrového souboru. Ani tento přístup není optimální, co se výsledků týče, nicméně faktory zdrojů nelze podceňovat. Poslední přístup (statistický) přináší nejpřesnější výsledky. „Rozsah výběru závisí na **variabilitě** výskytu zkoumaného znaku v základním souboru, na požadované míře **přesnosti** odhadu a koeficientu **spolehlivosti**“.²⁷ Kozel také uvádí jako příklad dva vzorce, pomocí kterých lze dojít ke konkrétnímu rozsahu:

$$A) n \geq \frac{(z^2 \times p \times q)}{\Delta^2}$$

nebo

$$B) n \geq \frac{(z^2 \times \delta^2)}{\Delta^2},$$

kde n je minimální počet dotázaných, z je koeficient spolehlivosti (nabývá hodnot 1 = 68,3% spolehlivost, 2 = 95,4% spolehlivost anebo 3 = 99,7% spolehlivost), p a q jsou relativní četnosti dotázaných, kteří by se přiklonili k první (p) nebo ke druhé (q) variantě (neznáme-li tyto četnosti, použijeme poměr „půl na půl“, tedy p i q rovno 0,5), δ je směrodatná odchylka od průměru, který ovšem musíme znát a Δ označuje maximální přípustnou chybu, kterou si sami zvolíme.²⁸

Typy výběru

Kromě velikosti výběru hraje také roli způsob výběru. Raab-Steinerová a Benesch k tomu podotýkají, že ani sebevětší vzorek nespasí průzkum, jsou-li osoby špatně vybrány a jako příklad uvádí: „Když bychom chtěli zjistit například něco o průměrném příjmu studenta a dotázali bychom se pěti tisíc studujících, dostali bychom silně pokřivené výsledky, kdyby tito studenti z větší části studovali při zaměstnání, tzn., měli plnohodnotný příjem“.²⁹ Jak již bylo výše zmíněno, narážíme zde na problém tzv. reprezentativnosti vzorku. Abychom docílili co největší shody mezi rozložením znaků v populaci a ve

²⁷ KOZEL, cit. d., s. 159

²⁸ Tamtéž, s. 159-160

²⁹ RAAB-STEINER, BENESCH, cit. d., s. 16

výběrovém souboru, nabízí se různé způsoby, jak vzorek vybrat. V základní rovině se rozlišují výběry pravděpodobnostní (neboli náhodné) a nepravděpodobnostní (neboli záměrné). Zásadním rozdílem spočívá v tom, že u pravděpodobnostních výběrů lze určit, s jakou pravděpodobností bude statistická jednotka zařazena do výběrového souboru, přičemž tato pravděpodobnost může a nemusí být pro všechny jednotky stejná, ale musí být známá.³⁰ Statistické modely počítají právě s pravděpodobnostními výběry, a proto se aplikace stejných modelů na záměrné výběry setkává s opodstatněnou kritikou.³¹ Pecáková charakterizuje záměrné výběry následovně: „Výběr jednotek probíhá s neznámými pravděpodobnostmi, nelze říci, zda pořízené odhady konvergují k populačním charakteristikám a jaké vlastnosti mají výběrové chyby. Nejsou totiž známy a kontrolovány zákonitosti jejich vzniku“.³²

Na základním rozdělení se shodují všichni autoři, ale při jemnějších členění jednotlivých typů výběrů se již částečně liší a to nejčastěji v tom, že jejich výčet není vyčerpávající a mezi nepravděpodobnostními výběry bývá jako jediný zástupce jmenován kvótní výběr.³³ Pro účely této práce proto bylo zvoleno členění podle autorky Ivy Pecákové, které je použito v její knize Statistika v terénních průzkumech.³⁴

1. Pravděpodobnostní výběry:

- a. **Prostý náhodný výběr:** Je to základní a nejjednodušší typ pravděpodobnostního výběru, ve kterém mají všechny jednotky stejnou pravděpodobnost, že budou vybrány. Hodí se především v případě, že nejsou k dispozici informace o rozdělení důležitých vlastností v populaci.³⁵
- b. **Oblastní (stratifikovaný) výběr:** Základní soubor se rozdělí do vrstev (strat), které se nepřekrývají a jsou vnitřně sourodé ale vzájemně odlišné. Z jednotlivých vrstev se pak vybírají prvky do vzorku.³⁶
- c. **Skupinový výběr:** Tento typ výběru se použije v případě, když základní soubor sestává ze skupin, které jsou „[...] vnitřně různorodé, ale vzájemně se velmi podobají a navíc jsou i přibližně stejně velké [...]“.³⁷ V praxi

³⁰ PECÁKOVÁ, cit. d., s. 17

³¹ PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 16

³² PECÁKOVÁ, cit. d., s. 20

³³ viz např. PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 14-16 nebo RAAB-STEINER, BENESCH, cit. d., s. 17-18

³⁴ PECÁKOVÁ, cit. d., s. 17-22

³⁵ RAAB-STEINER, BENESCH, cit. d., s. 17

³⁶ Tamtéž, s. 17. Překlad autorky z originálu: „Diese Vorgangsweise macht natürlich nur dann Sinn, wenn die Schichtungsvariable einen hohen Zusammenhang mit dem eigentlich interessierendem Untersuchungsmerkmal hat.“

³⁷ PECÁKOVÁ, cit. d., s. 19

bývají takovéto skupiny přirozeně k dispozici, což je nejlépe vidět na typickém příkladu tříd ve školách.³⁸

- d. **Víceřádný výběr:** Doporučuje se hlavně při práci s velkým základním souborem, který se pro usnadnění práce rozdělí do několika skupin, které se pak dále dělí – vznikne tedy více řádnů. Z malých skupinek na nejnižším stupni se vyberou již konkrétní jednotky do výběrového souboru.

2. Nepravděpodobnostní výběry:

- a. **Příležitostný výběr:** Příkladem tohoto typu výběru je náhodné kontaktování chodců na ulici.³⁹
- b. **Samovolný výběr (anketa):** I zde není populace jednoznačná. Obecně mají tendenci se do ankety zapojovat lidé, kteří mají o téma zájem, ať už v pozitivním nebo negativním smyslu.⁴⁰
- c. **Nabalovací (snowball) výběr:** Jak vyplývá z názvu, výběr probíhá na principu nabalování – již oslovení respondenti doporučí další.⁴¹
- d. **Úsudkový výběr:** Zde se do výběru zásadním způsobem promítá osoba výzkumníka, který si jednotky vybírá na základě svého úsudku, že jde o vhodné osoby splňující kritéria.⁴²
- e. **Typický výběr:** „Při použití typického výběru jsou [...] vybírány jednotky považované někým (expertem) v dané populaci za typické“.⁴³
- f. **Kvótní výběr:** Tento často kritizovaný, ale přesto hojně rozšířený typ, uvedl do praxe George Gallup. Cílem je vybrat takový soubor, který by byl miniaturou zkoumané populace. Volí se proto kvótní kritéria, které jednotky musí splnit, aby byly zařazeny do výběru. Z praktických důvodů je ale počet kritérií omezen a populace se se vzorkem shoduje jen v několika znacích.⁴⁴

3.2.3 Struktura dotazníku

Aby dotazník správně plnil svoji funkci, je potřeba dobře promyslet jeho strukturu a celkový vzhled, aby dotazník už na první pohled neodrazoval od vyplnění, ale naopak

³⁸ PECÁKOVÁ, cit. d., s. 19

³⁹ Tamtéž, s. 20

⁴⁰ KOZEL, cit. d., s. 159

⁴¹ URBAN, Lukáš. *Sociologie trochu jinak*. s. 67

⁴² KOZEL, cit. d., s. 158

⁴³ PECÁKOVÁ, cit. d., s. 20

⁴⁴ PECÁKOVÁ, NOVÁK, HERZMANN, cit. d., s. 15

respondenta zaujal.⁴⁵ Z ryze praktického hlediska je potřeba vzít v úvahu, na koho dotazník cílí. Pokud jsou to starší lidé, je například potřeba zvolit písmo větší a dobře čitelné, pro děti se hodí doplnit otázky ilustracemi či schémata.⁴⁶ Ačkoli by se grafické zpracování rozhodně nemělo podceňovat, bude na tomto místě probrána hlavně struktura dotazníku, resp. co by měl dotazník obsahovat.

Nejpodstatnější část dotazníku tvoří otázky. Nicméně je nevhodné, aby byly tím jediným, co dotázaný v dotazníku najde. Na začátek se proto zařazuje část úvodní, která by podle Raab-Steinerové a Benesche měla obsahovat následující body:

- Představení osoby, popř. instituce, která výzkum provádí;
- téma dotazníku;
- za jakým účelem je průzkum prováděn;
- prosba o kompletní vyplnění;
- prosba o upřímné odpovědi a upozornění, že nejsou ani správné ani špatné odpovědi;
- záruka anonymity;
- poděkování za vyplnění.⁴⁷

Podle Radima Kozla může být poděkování zmíněno ještě jednou v závěru dotazníku.⁴⁸

Otázky v dotazníku

To nejpodstatnější na dotazníku jsou otázky, které mají přímý vliv na statistické zpracování a kvalitu výsledků. Základem každé otázky je, aby byla pro dotázaného srozumitelná, a aby na ni dotázaný byl schopen a ochoten odpovědět.⁴⁹ Autoři Raab-Steinerová a Benesch vytvořili seznam dalších sedmnácti doporučení, jak správně otázky formulovat, aby splňovaly již zmíněné tři základní zásady. Vyzdvihnout je potřeba hlavně to, že otázky by měly být krátké, ale zároveň pregnantní a autor by se měl vyhnout sugestivním, stereotypním nebo stigmatizujícím formulacím.⁵⁰

⁴⁵ ŘEZANKOVÁ, cit. d., s. 14

⁴⁶ PECÁKOVÁ, cit. d., s. 28

⁴⁷ RAAB-STEINER, BENESCH, cit. d., s. 49-50

⁴⁸ KOZEL, cit. d., s. 166

⁴⁹ PECÁKOVÁ, cit. d., s. 28

⁵⁰ RAAB-STEINER, BENESCH, cit. d., s. 50-51

Pro detailnější probrání typů otázek bylo pro tuto práci zvoleno členění opět podle autorky Ivy Pecákové, které se ale do značné míry shoduje s členěním Romana Kozla.⁵¹ Ten používá jako členící kritéria obsah, formu a účel⁵², přičemž Pecáková přidává ještě pořadí.

1. **Otázky podle pořadí:** Pro pořadí otázek je nejdůležitější jejich logický sled, resp. plynulá návaznost. Na druhou stranu by předchozí otázky neměly ovlivňovat ty následující.⁵³ Na konec se doporučují klást tzv. „identifikační otázky (zjišťující například pohlaví, věk, zaměstnání ap.), které by mohly narušit ochotu respondenta ke spolupráci [...]“.⁵⁴
2. **Otázky podle účelu:** Roman Kozel je dále člení na otázky nástrojové, které pomáhají navázat kontakt s dotázaným; mohou respondenty třídit (např. kuřák nebo nekuřák) a tím větvit dotazník; týkají se demografických údajů; nebo se jedná o otázky kontrolní, které mohou odhalit podvody při vyplňování odpovědí. Další podskupinou jsou výsledkové otázky, které se ptají respondenta na jeho názory, zkušenosti a postoje. Poslední skupinou jsou pomůcky, které mívají nejčastěji podobu obrázku nebo tabulky a doplňují slovní otázku.⁵⁵
3. **Otázky podle obsahu:** Rozlišují se zde dva typy otázek, a to přímé a nepřímé. „U přímých otázek je zcela zřejmý účel dotazu a respondent vědomě odpovídá na to, na co je tázán (*Jakého jste vyznání?*)“.⁵⁶ Tyto otázky ale mohou právě kvůli své přímosti působit nepříjemně a respondent nemusí být ochotný odpovědět. Vhodným přeformulováním do nepřímé podoby je možno přimět respondenta, aby odpověď přeci jen poskytl. Pecáková uvádí jako příklad přímou otázku „[...] *Bijete své děti?* [...]“ , kterou je lepší položit jako nepřímou například ve znění „*Někteří rodiče bijí své děti – co tomu říkáte?*“.⁵⁷
4. **Otázky podle formy:** Z tohoto hlediska dělíme otázky na otevřené, uzavřené a polouzavřené, resp. polootevřené. V otevřených otázkách je dotázaným dána možnost samostatně formulovat odpověď. Jejich hlavní výhodou je volnost v odpovídání, což je zároveň i nevýhodou v tom smyslu, že pro některé dotázané je

⁵¹ PECÁKOVÁ, cit. d., s. 28-32

⁵² KOZEL, cit. d., s. 167-173

⁵³ ŘEZANKOVÁ, cit. d., s. 14

⁵⁴ PECÁKOVÁ, cit. d., s. 28

⁵⁵ KOZEL, cit. d., s. 166-171

⁵⁶ PECÁKOVÁ, cit. d., s. 30

⁵⁷ Tamtéž, s. 30

samostatná formulace z různých důvodů náročná. Ze statistického hlediska je nutno podotknout, že zpracování je náročné.⁵⁸ Naopak pro statistickou analýzu nejvhodnější jsou uzavřené otázky, kde jsou respondentovi nabídnuty předpřipravené odpovědi. Může se ale stát, že výzkumník nevyčerpá veškeré možné odpovědi, a proto se používají i otázky polouzavřené, které většinou pod možností odpovědi „jiné“ nabízí vlastní odpověď jako v otázce otevřené.⁵⁹

Typy proměnných a škál

Už při navrhování dotazníku je potřeba se zamýšlet nad fází následující, tedy nad vyhodnocením, nad operacemi, které budou s daty prováděny a v jakém programu.⁶⁰ Prostředkem je v tomto případě vytvoření struktury datového souboru. Podle Hany Machkové se musí již v návrhu určit „[...] typy proměnných, škály hodnot, označení chybějících údajů [...]“.⁶¹ Pokud se výzkumník rozhodne zabývat se těmito záležitostmi až při analýze dat, svou práci si tak zbytečně komplikuje.⁶²

Na tomto místě se opět dostáváme k významu otázek v dotazníku, jelikož právě ty ovlivňují, jak bude datová struktura vypadat, konkrétně s jakými proměnnými (veličinami) a škálami se bude pracovat.⁶³ Pro výzkum trhu a veřejného mínění je typické, že kromě veličin, jejichž měření je nasnadě (např. věk, váha, počet dětí), se měří nejčastěji motivace, postoje nebo názory respondentů. Zde přichází na řadu tzv. operacionalizace, což je postup, pomocí kterého se dá odmítání nebo souhlas s nějakým názorem udělat měřitelným. V rámci operacionalizace se určí typ proměnné a přiřadí se jí odpovídající škála.⁶⁴ V následujícím textu bude pojednáno právě o typech proměnných a škálách, pomocí kterých se měří. Co se týká členění proměnných a škál, dosud citovaní autoři shodně rozeznávají tři základní typy (nominální, ordinální a kardinální) a většinou ještě rozlišují dvě podskupiny kardinálních, což jsou proměnné, resp. škály intervalové a poměrové.

⁵⁸ RAAB-STEINER, BENESCH, cit. d., s. 48

⁵⁹ PECÁKOVÁ, cit. d., s. 31

⁶⁰ MACHKOVÁ, H. *Mezinárodní marketing*. s. 56

⁶¹ Tamtéž, s. 56

⁶² Tamtéž, s. 56

⁶³ PECÁKOVÁ, cit. d., s. 32

⁶⁴ RAAB-STEINER, BENESCH, cit. d., s. 22

- **Nominální proměnné a škály:** O hodnotách nominální proměnné lze říct pouze to, jestli se liší nebo ne. Při vyhodnocování lze u nominálních proměnných pracovat pouze s jejich četnostmi, což analýzu velmi omezuje.⁶⁵
- **Ordinální proměnné a škály:** Máme-li dvě hodnoty ordinální proměnné, můžeme určit, která hodnota je položena výše a která níže, ale nelze určit velikost rozdílu mezi hodnotami. Možnosti vyhodnocení jsou již bohatší než u ordinálních proměnných, lze spočítat např. medián nebo variační rozpětí.⁶⁶
- **Kardinální proměnné a škály:** „[...] představují určitý počet konstantních měrných jednotek použitých pro vyjádření stupně sledovaného znaku, případně vyjadřují počet nějakých objektů či osob (počet zaměstnanců firmy)“.⁶⁷ V rámci kardinálních proměnných se rozlišují proměnné intervalové a poměrové.⁶⁸

3.3 Zpracování dotazníku

Ve chvíli, kdy je potřebný počet dotazníků vyplněn, přichází na řadu zpracování získaných dat. Raab-Steinerová a Benesch ve své knize o dotaznících s potěšením poznamenávají, že doby, kdy se data zpracovávala bez pomoci počítačů, jsou již naštěstí dávno pryč.⁶⁹ Výzkumníci si mohou v současnosti vybrat z celé řady specializovaných programů, mezi které patří např. *IBM SPSS Statistics*, *SAS*, *SYSTAT*, *Statistica*, *Stata* a další.⁷⁰ Rainer Leonhart do seznamu programů zahrnuje i *Excel* z balíčku *Microsoft Office*.⁷¹ Autoři Raab-Steinerová a Benesch ho ale za validní statistický program nepovažují a argumentují tím, že je to v první řadě tabulkový kalkulační editor a má jen některé přidané statistické funkce.⁷² V tomto textu bude blíže pojednáno o programu *IBM SPSS Statistics* verze 20 (zkráceně *SPSS*), kterého bude následně využito i na příkladu zpracování konkrétního dotazníku.

⁶⁵ KOZEL, MYNÁŘOVÁ, SVOBODOVÁ, cit. d., s. 6

⁶⁶ ŘEZANKOVÁ, cit. d., s. 49

⁶⁷ PECÁKOVÁ, cit. d., s. 33

⁶⁸ RAAB-STEINER, BENESCH, cit. d., s. 27

⁶⁹ RAAB-STEINER, BENESCH, cit. d., s. 64

⁷⁰ LEONHART, cit. d., s. 22-24

⁷¹ Tamtéž, s. 24

⁷² RAAB-STEINER, BENESCH, cit. d., s. 64

3.3.1 Vytvoření kódového klíče

Prvním krokem zpracování dat z dotazníku je jejich zadání do programu a vytvoření datové matice, která je uspořádána následujícím způsobem: „Řádky jsou vymezeny pro odpovědi jednotlivých respondentů, sloupce obsahují odpovědi na jednotlivé otázky, resp. na jejich části“.⁷³ Před zadáním dat do řádků se definují proměnné: pojmenují se a jednotlivým odpovědím se přiřadí většinou číselné kódy, se kterými je počítač schopen pracovat a které zároveň ulehčí přepis údajů do matice.⁷⁴ Takto vznikne tzv. kódový klíč, který je dobré si pro přehlednost zaznamenat na jeden prázdný dotazník.⁷⁵ V souvislosti s číselným kódováním upozorňuje Hana Řezanková na fakt, že u nominálních proměnných nemají kódy význam čísel a u ordinálních proměnných vyjadřují pouze pořadí.⁷⁶

V programu *SPSS* se proměnné definují v listu *Variable View*. První proměnnou bývá obvykle identifikace konkrétního dotazníku, ze kterého údaje pochází. Očíslování dotazníku je užitečné především při čištění dat, kdy je v případě zadání chybného čísla možné snadno dohledat správnou verzi.⁷⁷ Kromě zadání názvu proměnné se určí její další vlastnosti jako např. typ škály, hodnoty, chybějící údaje ad. Problematika chybějících údajů přichází na řadu ve chvíli, kdy v dotazníku odpověď chybí, respondent vybral možnost „nechci odpovídat“ nebo „nevím“ apod. *SPSS* je přednastaveno tak, že pokud se v listu *Data View* nechá políčko prázdné, je vyhodnoceno jako systémová chyba. Pokud mají ale chybějící údaje obsahovou relevanci, je dobré jim ve *Variable View* ve sloupci *Missing* přiřadit rozličné hodnoty.⁷⁸ Chybějící hodnoty nejsou zahrnovány do analýz a to bez ohledu na jejich definici.⁷⁹

3.3.2 Čištění dat

Před započítím vyhodnocování dat je vhodné data zkontrolovat a opravit tak možné chyby: „Tímto lze předejít nepříjemným překvapením a hlavně špatným výsledkům“.⁸⁰ Kontrole se podrobují jednotlivé proměnné, kdy se zkoumají vybočující hodnoty, tak i

⁷³ ŘEZANKOVÁ, cit. d., s. 21

⁷⁴ LEONHART, cit. d., s. 28

⁷⁵ RAAB-STEINER, BENESCH, cit. d., s. 68

⁷⁶ ŘEZANKOVÁ, cit. d., s. 23

⁷⁷ RAAB-STEINER, BENESCH, cit. d., s. 68

⁷⁸ Tamtéž, s. 72-74

⁷⁹ LEONHART, cit. d., s. 30

⁸⁰ RAAB-STEINER, BENESCH, cit. d., s. 76. Překlad autorky z originálu: „Damit können böse Überraschungen und vor allem falsche Ergebnisse vermieden werden.“

konzistence mezi proměnnými, aby se v souboru nevyskytli např. těhotní muži či desetileté vdané ženy.⁸¹ První typ čištění lze v *SPSS* provést pomocí *Analyze – Descriptive Statistics – Frequencies*, pro druhý typ jsou naopak vhodné kontingenční tabulky. Odhalené nepřesnosti jsou opraveny na základě původních dotazníků.⁸²

3.3.3 Popis datového souboru

První fází samotného zpracování dat je jejich popis neboli deskripce. Cílem je získat přehled o výběrovém souboru a podněty pro následné analýzy. Již z pouhého popisu by měly být důležité skutečnosti na první pohled zřetelné.⁸³ Při deskripci i dalších analýzách dat je nutné brát v úvahu, zda se jedná o proměnnou nominální, ordinální nebo kardinální. Ačkoli lze např. spočítat aritmetický průměr i pro nominální proměnnou, jejímž hodnotám byly pro potřeby vyhodnocení v *SPSS* přiděleny číselné kódy, jedná se o metodickou chybu a výsledek nelze interpretovat.⁸⁴ V textu bude na případná omezení využití upozorňováno. Konkrétní příklady budou uvedeny v části zabývající se zpracováním konkrétního dotazníku.

Tato část textu bude strukturována na základě knihy *Analýza dat z dotazníkových šetření* od Hany Řezankové, jelikož postup této autorky je velice přehledný. V ostatních knihách je postup sice obdobný, nicméně seskupování témat v jednotlivých podkapitolách je odlišné.

Rozdělení četností

Do popisné statistiky patří rozdělení četností, ze kterého lze zjistit, kolik respondentů zvolilo dané kategorie určité proměnné. K zobrazení slouží buď tabulka, nebo vhodný graf. Leonhart poznamenává, že především pro nominální proměnné jsou četnosti významné, jelikož další možnosti popisu jsou oproti jiným typům proměnných omezené.⁸⁵ Intervalové kardinální proměnné je naopak vhodné vzhledem k velkému počtu různých hodnot nejprve rozdělit do kategorií.⁸⁶ V programu *SPSS* lze tabulku rozdělení četností vygenerovat pomocí *Analyze – Descriptive Statistics – Frequencies*. Ve výsledku se

⁸¹ ESSER, E., HILL, P. B., SCHNELL, R. *Methoden der empirischen Sozialforschung*. s. 437

⁸² RAAB-STEINER, BENESCH, cit. d., s. 76

⁸³ Tamtéž, s. 82

⁸⁴ KOZEL, cit. d., s. 95

⁸⁵ LEONHART, cit. d., s. 110

⁸⁶ Tamtéž, s. 110

zobrazí absolutní četnosti (*Frequency*), relativní četnosti (*Percent* a *Valid Percent*)⁸⁷ a kumulativní relativní četnosti (*Cumulative Percent*), které ale mají smysl pouze ordinální a kardinální proměnné.⁸⁸

V knize o dotaznících od autorů Kirchhoffová a kol. nese kapitola o grafech název „Jeden obrázek řekne více než tisíc čísel [...]“.⁸⁹ Tím autoři přesně vystihují hlavní důvod pro použití grafů k zobrazení rozdělení četností: názornost. Andy Field nicméně upozorňuje na nedostatky, které názornost snižují a kterým je potřeba se vyvarovat. Jsou to např. zbytečné využívání 3-D efektu, špatný popis os, rozličné vzory a další prvky, které odvádí pozornost.⁹⁰ Grafy lze v *SPSS* vytvořit různými způsoby, mezi něž patří *Chart Builder* z menu *Graphs*, který obsahuje velmi rozmanitou nabídku nejrůznějších typů grafů. Které jsou ale vhodné pro zobrazení rozdělení četností? „Mezi grafy jsou pro vyjádření výsledků třídění nejpoužívanější různé typy sloupkových či výsečových diagramů, pro spojité číselné proměnné jsou vhodné histogramy“.⁹¹

Popisné charakteristiky

Mezi popisné charakteristiky patří míry polohy, variability a koncentrace. Pro odlišné typy proměnných se používají odlišné statistiky, přičemž ty, které lze spočítat pro „nižší“ typ proměnné (nominální, popř. ordinální), lze použít i pro „vyšší“ typ (kardinální, popř. ordinální). V *SPSS* lze zjistit hodnoty popisných charakteristik následujícím postupem: *Analyze – Descriptive Statistics – Frequencies* a zde v menu *Statistics* vybereme požadované míry. Zde uvedený výčet se zaměří na charakteristiky, které budou následně prakticky využity v další části textu.

Míry polohy

1. **Modus** (\hat{x}) je charakteristika používaná pro nominální proměnné a označuje nejčetnější kategorii. Má-li sledovaná proměnná pouze jeden modus neboli modální kategorii, pak se hovoří o unimodálním rozdělení četností. Modálních kategorií může být ale více, což označujeme jako „[...] *k*-modální rozdělení, kde *k* je počet

⁸⁷ *Percent* a *Valid Percent* se liší v případě, že se mezi daty objevují chybějící odpovědi.

⁸⁸ ŘEZANKOVÁ, cit. d., s. 35-36

⁸⁹ KIRCHHOFF, S., KUHN, S., LIPP, P., SCHLAWIN, S. *Der Fragebogen Datenbasis, Konstruktion und Auswertung*. s. 59

⁹⁰ FIELD, A. *Discovering statistics using SPSS: (and sex, drugs and rock 'n' roll)*. s. 89

⁹¹ PEČÁKOVÁ, cit. d., s. 77

modálních kategorií“.⁹² Přesáhne-li relativní četnost modální kategorie 0,5, pak se jedná zároveň o kategorii majoritní.⁹³

2. **Medián** (\tilde{x}) se využívá pro charakteristiku ordinálních proměnných a také kardinálních proměnných, které nemají normální rozdělení. Výpočet se provádí následovně: hodnoty se sestaví vzestupně do řady. Je-li počet hodnot lichý, medián je identický s číslem uprostřed. Při sudém počtu hodnot se spočítá jako aritmetický průměr dvou prostředních hodnot. Výhodou této míry je její necitlivost na extrém.⁹⁴
3. **Aritmetický průměr** (\bar{x}) je vhodné použít pro popis kardinálních normálně rozdělených proměnných. Jeho nevýhodou je citlivost na extrém, ale i přesto je nejnámější mírou polohy a v podstatě synonymem pro průměr.⁹⁵ Výpočet z původních hodnot se provádí podle následujícího vzorce:

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

kde n je počet respondentů a index i značí kategorie ($i = 1, 2, \dots, K$, kde K je počet kategorií).⁹⁶

Míry variability

1. **Variační rozpětí** (R) je míra udávající rozpětí mezi nejmenší a nejvyšší naměřenou hodnotou v souboru. Na podobném principu funguje i mezikvantilové rozpětí, které udává, kde se nachází prostředních 50 % hodnot. Vypovídací hodnota obou statistik je ale omezená, protože neposkytují informace o hodnotách ležících mezi.⁹⁷ Obě lze využít k charakteristice ordinálních proměnných.⁹⁸
2. **Rozptyl** (s^2) a z něj odvozená směrodatná odchylka a variační koeficient mají smysluplné využití pouze u kardinálních proměnných s normálním rozdělením.⁹⁹ Výpočet z původních hodnot se provádí podle následujícího vzorce:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \cdot 100$$

⁹² ŘEZANKOVÁ, cit. d., s. 46

⁹³ Tamtéž, s. 46

⁹⁴ RAAB-STEINER, BENESCH, cit. d., s. 98

⁹⁵ Tamtéž, s. 96-97

⁹⁶ ŘEZANKOVÁ, cit. d., s. 52

⁹⁷ RAAB-STEINER, BENESCH, cit. d., s. 102-104

⁹⁸ ŘEZANKOVÁ, cit. d., s. 50

⁹⁹ RAAB-STEINER, BENESCH, cit. d., s. 100

¹⁰⁰ ŘEZANKOVÁ, cit. d., s. 53

Ze vzorce vyplývá, že se jedná o „[...] aritmetický průměr ze čtverců odchylek jednotlivých hodnot od průměru [...]“.¹⁰¹

3. **Směrodatná odchylka** (s) se využívá k ohodnocení variability, protože je lépe interpretovatelná než rozptyl, který je udáván ve čtvercích původních jednotek. Způsob výpočtu je tedy nasnadě:

$$s = \sqrt{s^2} .^{102}$$

4. **Variační koeficient** (V_x) je vhodnou mírou variability pro srovnání více souborů, jelikož je relativní; neudává se v měrných jednotkách původní proměnné, ale většinou v procentech. K výpočtu se používá tento vzorec:

$$V_x = \frac{s}{\bar{x}} .^{103}$$

Míry koncentrace

1. **Koeficient šikmosti** posuzuje míru symetrie rozdělení četností. Nabyde-li hodnoty nula, jedná se o souměrné rozdělení, které je v praxi spíše výjimečné. Častěji je rozdělení zešikmeno buď kladně, což znamená, že hodnoty menší než průměr jsou četnější, nebo naopak záporně.¹⁰⁴
2. **Koeficient špičatosti** charakterizuje tvar rozdělení četností v tom smyslu, zda se odpovědi respondentů spíše koncentrují u určité odpovědi nebo jsou rovnoměrně rozloženy. Normální rozdělení je charakterizováno hodnotou nula. Jsou-li hodnoty větší než nula, jedná se o spíše špičaté rozdělení, jsou-li hodnoty menší než nula, pak je rozdělení spíše plošší.¹⁰⁵

3.3.4 Analýza závislosti

Dosud byly proměnné analyzovány samostatně. Navazujícím krokem tedy je posoudit, jak spolu dvojice určitých proměnných souvisí. Základem je vytvoření dvourozměrné tabulky zobrazující rozdělení četností – tzv. kontingenční tabulky. V ní se podobně jako v tabulce rozdělení četností jedné proměnné uvádí četnosti buď v absolutní, nebo relativní podobě. Řezanková uvádí tři možnosti pro relativní četnosti; jsou to: „[...] podíly počítané na základě celého rozsahu souboru (jejich součet se rovná hodnotě 1, resp.

¹⁰¹ SOUČEK, E. *Statistika pro ekonomy*. s. 20

¹⁰² ŘEZANKOVÁ, cit. d., s. 53

¹⁰³ SOUČEK, cit. d., s. 22

¹⁰⁴ Tamtéž, s. 23

¹⁰⁵ KOZEL, cit. d., s. 99

100 v procentním vyjádření), *řádkové podíly* (součet hodnot v každém řádku se rovná hodnotě 1) nebo *sloupcové podíly* (součet hodnot v každém sloupci se rovná hodnotě 1)¹⁰⁶. V souvislosti s kontingenční tabulkou je potřeba zmínit také pojmy sdružené a marginální četnosti. Sdružené četnosti se zaznamenávají do jednotlivých políček tabulky a jsou-li v absolutní hodnotě, vyjadřují počet respondentů, kteří zvolili *i*-tou kategorii znaku *X* a *j*-tou kategorií znaku *Y*, kde *i* (*i* = 1, 2, ..., *R*) značí řádky a *j* sloupce (*j* = 1, 2, ..., *S*).¹⁰⁷ V SPSS lze kontingenční tabulku vytvořit přes *Analyze – Descriptive Statistics – Crosstabs*, kde se následně vybere proměnná zobrazená v řádku a proměnná ve sloupci.

Dvě nominální proměnné

„Závislost dvou nominálních proměnných se nazývá kontingence“.¹⁰⁸ Základem pro její posouzení je chí-kvadrát test o nezávislosti, který je založen na kontingenční tabulce a rozšiřuje chí-kvadrát test dobré shody. Princip fungování vysvětluje Rimarčík následovně: „Test overuje nulovú hypotézu, ktorá tvrdí, že očakávané a zo vzorky získané početnosti sa rovnajú [...], teda dve nominálne premenné sú nezávislé [...]“.¹⁰⁹ Pro ověření nulové hypotézy je potřeba testového kritéria, v tomto případě se jedná o Pearsonovu statistiku chí-kvadrát, jejíž vzorec je následující:

$$\chi_P^2 = n \sum_{i=1}^R \sum_{j=1}^S \frac{(n_{ij} - m_{ij})^2}{m_{ij}},$$

kde *R* je počet řádků, *S* počet sloupců, *n_{ij}* značí četnost v *i*-tém řádku a *j*-tém sloupci, a *m_{ij}* očekávanou četnost v *i*-tém řádku a *j*-tém sloupci.¹¹⁰

Aby bylo možné chí-kvadrát test použít, je potřeba, aby se rozdělení četností ve výběrovém souboru přibližovalo chí-kvadrát rozdělení.¹¹¹ Pro čtyřpolní tabulku znamená tento předpoklad, že žádná z očekávaných četností nesmí klesnout pod 5. Pro větší tabulky je pravidlo mírnější a stanovuje, že „[...] všechny očekávané četnosti by měly být větší než 1 a ne více než 20% očekávaných četností by mělo být menší než 5“.¹¹² Není-li tento předpoklad splněn, použije se Fisherův exaktní test, který se ale obvykle využívá

¹⁰⁶ ŘEZANKOVÁ, cit. d., s. 77

¹⁰⁷ Tamtéž, s. 78

¹⁰⁸ Tamtéž, s. 83

¹⁰⁹ RIMARČÍK, M. *Štatistika pre prax*. s. 171 - 172

¹¹⁰ ŘEZANKOVÁ, cit. d., s. 84

¹¹¹ FIELD, cit. d., s. 690

¹¹² Tamtéž, s. 695. Překlad autorky z originálu: „[...] all expected counts should be greater than 1 and no more than 20% of expected counts should be less than 5.“

především pro menší výběrové soubory a hlavně pro tabulky 2x2, jelikož je náročný na výpočet.¹¹³

Je-li mezi proměnnými zjištěna závislost, lze její míru zjistit kontingenčními koeficienty, mezi které patří např. koeficient ϕ a Cramérovo V. Andy Field doporučuje používat spíše druhý koeficient, jelikož ϕ lze využít pouze pro čtyřpolní tabulku, ve které navíc Cramérovo V dosahuje stejné hodnoty jako ϕ .¹¹⁴ Zde je vzorec pro výpočet Cramérova V:

$$V = \sqrt{\frac{\chi^2_p}{n(q-1)}}$$

kde $q = \min \{R, S\}$.¹¹⁵

Dvě ordinální proměnné

Od ordinálního typu proměnné výše se již nehovoří o kontingenci, ale o korelaci. Hana Řezanková popisuje dva typy korelace: „[...] *pozitivní* (nízkým hodnotám jedné proměnné odpovídají nízké hodnoty proměnné druhé) a *negativní* (nízkým hodnotám jedné proměnné odpovídají vysoké hodnoty druhé proměnné)“.¹¹⁶ Pro účely této práce budou na tomto místě popsány tři míry: Spearmanův koeficient pořadové korelace, Kendallovo τ_b a τ_c .

Jak již vyplývá z názvu Spearmanova koeficientu pořadové korelace, pracuje jeho výpočet s pořadími. Hodnotám obou proměnných jsou tedy nejdříve přiřazena pořadí, na které se uplatní vzorec pro Pearsonův korelační koeficient. Na rozdíl od Pearsonova korelačního koeficientu, se ten Spearmanův hodí pro nenormálně rozdělené soubory. Hodnoty koeficientu spadají do intervalu $(-1; 1)$, přičemž 0 značí lineární nezávislost proměnných, záporné hodnoty negativní korelaci a kladné hodnoty pozitivní korelaci.¹¹⁷

Obě Kendallova τ jsou založena na porovnávání dvojic objektů.¹¹⁸ Field doporučuje použití těchto koeficientů v případě, že se jedná o menší soubor, kde se mnoho případů dělí o stejné pořadí. Naznačuje také preferenci Kendallova τ , ačkoli je Spearmanův koeficient známější.¹¹⁹ Rozdíl mezi τ_b a τ_c spočívá v tom, že τ_c zohledňuje rozměry tabulky.¹²⁰

¹¹³ FIELD, cit. d., s. 690

¹¹⁴ Tamtéž, s. 698

¹¹⁵ ŘEZANKOVÁ, cit. d., s. 86

¹¹⁶ Tamtéž, s. 96

¹¹⁷ ŘEZANKOVÁ, cit. d., s. 98

¹¹⁸ Tamtéž, s. 98

¹¹⁹ FIELD, cit. d., s. 181

Dvě kardinální proměnné

Pro zjištění závislosti mezi dvěma kardinálními proměnnými lze použít Pearsonův korelační koeficient. Jeho omezením nicméně je fakt, že ho lze použít pouze tehdy, jsou-li proměnné normálně rozdělené. V opačném případě se počítá výše popsany Spearmanův koeficient pořadové korelace.¹²¹ Jelikož se také jedná o standardizovaný koeficient, leží jeho hodnoty stejně jako u Spearmanova koeficientu v intervalu $(-1; 1)$.¹²²

3.3.5 Porovnávání skupin – Neparametrické testy pro nezávislé výběry

Poslední typ analýzy, který bude použit při zpracování výběrového dotazníkového šetření, se týká porovnávání souborů podle určité charakteristiky polohy.¹²³ Vzhledem k tomu, že ve výběrovém souboru popisovaném v následující části textu nemají data normální rozdělení, zaměří se výklad na neparametrické testy, pro které tento typ rozdělení není předpokladem. Kromě typu rozdělení mají parametrické testy ještě další předpoklady, jako jsou kardinální typ proměnné a homogenita rozptylů v souborech. Ani tyto pro neparametrické testy neplatí; nicméně proměnná musí být alespoň ordinálního typu.¹²⁴ Princip neparametrických testů je podobný, jako byl popsán u Spearmanova korelačního koeficientu, který se také využívá pro nenormálně rozdělené hodnoty: místo konkrétních údajů se použijí jejich pořadí neboli skóry. Tím ale dochází ke ztrátě informací, a proto lze v literatuře často narazit na tvrzení, že neparametrické testy mají menší sílu než parametrické, tzn., existuje-li efekt, pak ho parametrický test odhalí spíše než test neparametrický – viz např. *Štatistika pre prax*¹²⁵ od Mariána Rimarčika nebo *Výzkumné metody a statistika*¹²⁶ od Iana Walkera. Andy Field jim ale oponuje tím, že tomu tak nemusí být vždy. Neparametrický test je ve srovnání s parametrickým prokazatelně slabší, jsou-li oba testy použity pro soubor, který splňuje předpoklady pro parametrické testy. Na druhou stranu, parametrický test nelze použít pro data, která předpoklady nesplňují a zde jsou proto neparametrické testy nenahraditelné.¹²⁷

¹²⁰ TOUTENBURG, H., HEUMANN, C., SCHOMAKER, M. *Deskriptive Statistik: eine Einführung in Methoden und Anwendungen mit R und SPSS*. s. 125

¹²¹ ŘEZANKOVÁ, cit. d., s. 108

¹²² FIELD, cit. d., s. 173

¹²³ ŘEZANKOVÁ, cit. d., s. 143

¹²⁴ NAUMANN, E., HOFMANN, W., FRIESE, M., RASCH, B. *Quantitative Methoden 2*. s. 145

¹²⁵ RIMARČÍK, cit. d., s. 160

¹²⁶ WALKER, I. *Výzkumné metody a statistika*. s. 154

¹²⁷ FIELD, cit. d., s. 551

Kromě dělení testů na parametrické a neparametrické se ještě používá dělení podle počtu porovnávaných souborů (2 anebo více než 2) a podle toho, jestli jsou výběry závislé nebo nezávislé. Hana Řezanková definuje nezávislé výběry takto: „*Nezávislé výběry* jsou takové, kdy sledujeme určitý znak v členění podle skupin respondentů, přičemž tyto skupiny jsou vytvořeny na základě kategorií vysvětlující proměnné (musí se tedy jednat o proměnnou kategoriální)“.¹²⁸ Typickým příkladem zmiňované kategoriální proměnné je porovnávání hodnot podle pohlaví. Závislé výběrové soubory jsou takové, kdy mezi hodnotami z různých souborů existuje nějaká souvislost, resp. závislost. Klasickým příkladem je několikanásobné dotázání těch samých osob v různých časových okamžicích – např. před začátkem léčby, po pěti týdnech a po dokončení léčby.¹²⁹ V dotazníkovém šetření uskutečněném pro účely této práce se ale závislé výběrové soubory nevyskytují, a proto se text zaměří na testy pro nezávislé výběry.

Mannův-Whitneyho U-test

Mannův-Whitneyho U-test je neparametrickou obdobou t-testu pro dva nezávislé výběry. Před jeho provedením je potřeba provést test na normalitu rozdělení (Kolmogorov-Smirnovův nebo Shapiro-Wilkův test) a test na homogenitu rozptylů (Leveneho test).¹³⁰ U-test, stejně jako t-test, zkoumá, zda mezi soubory existují statisticky významné rozdíly, co se mediánů týče.¹³¹ Testovou statistikou je Mannova-Whitneyova statistika U založená na Wilcoxonově statistice W . Výpočet U je následující:

$$U = W - \frac{n_1(n_1+1)}{2},$$

kde n_1 je rozsah menšího výběru a W je nižší ze součtů pořadových hodnot jednotlivých souborů.¹³² V SPSS lze tento test najít pod *Analyze – Nonparametric Tests – Legacy Dialogs – 2 Independent Samples*. Zde se vybere testovaná proměnná a proměnná, podle které se ta první rozdělí na dvě skupiny. Není-li výběr příliš velký, doporučuje Andy Field zvolit v *Options* exaktní metodu výpočtu p-hodnoty, která je ovšem náročnější na výpočet.¹³³

¹²⁸ ŘEZANKOVÁ, cit. d., s. 143

¹²⁹ RAAB-STEINER, BENESCH, cit. d., s. 115

¹³⁰ FIELD, cit. d., s. 546

¹³¹ NAUMANN, HOFMANN, FRIESE, RASCH, cit. d., s. 145

¹³² ŘEZANKOVÁ, cit. d., s. 158

¹³³ FIELD, cit. d., s. 547

Kruskalův-Wallisův test

Kruskalův-Wallisův test se používá pro analýzu více než dvou nezávislých výběrů. Svůj parametrický protějšek proto nachází v analýze rozptylu. Jeho princip je velice podobný výše popsaném U-testu; opět se jedná o výpočty s pořadími a i zde je nejprve potřeba otestovat normalitu rozdělení a homogenitu rozptylů. Jako testová statistika se používá Kruskalovo-Wallisovo H , které má tento vzorec:

$$H = \frac{12}{N(N-1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N + 1),$$

kde N je celková velikost souboru, R_i je suma pořadí jednotlivých souborů a n_i jejich velikost.¹³⁴ V SPSS lze tento test provést přes Analyze – Nonparametric Tests – Legacy Dialogs – K Independent Samples. I zde je v Options možno nastavit přesnou metody výpočtu p-hodnoty.¹³⁵

Ohledně post-hoc analýzy nemá tento test tak elegantní možnosti jako ANOVA a výpočty se potýkají s problémem mnohonásobných testů, kdy se prudce zvyšuje pravděpodobnost chyby prvního druhu.¹³⁶ Andy Field doporučuje zobrazit si výsledky pomocí krabičkového grafu a na jeho základě vybrat několik dvojic, které se podrobí Mannovu-Whitneyovu testu. Přes tzv. Bonferroniho korekci se vydělí hladina významnosti 5% počtem provedených U-testů a teprve s touto hodnotou α se porovnává z testu získaná p-hodnota.¹³⁷

¹³⁴ FIELD, cit. d., s. 560

¹³⁵ Tamtéž, s. 563

¹³⁶ RAAB-STEINER, BENESCH, cit. d., s. 116

¹³⁷ FIELD, cit. d., s. 565

4 Vlastní zpracování a zhodnocení výsledků

Pro účely této bakalářské práce byl vytvořen dotazník vycházející ze zásad popsaných v části týkající se návrhu dotazníku. Tématem průzkumu byly postoje středoškolských žáků ke studiu, způsob, jakým přistupují k přípravě do školy a jejich spokojenost s oborem a studijními výsledky. Výběrový soubor byl pořízen na SOŠ a SOU Kladno mezi denními studenty 4-letých maturitních oborů. Ačkoli byl průzkum proveden napříč všemi třídami, nejedná se o vyčerpávající šetření vzhledem k tomu, že část studentů v době šetření nebyla ve škole přítomna a to buď z důvodu praxe mimo školu anebo se jednalo o regulérní absenci. Rozhodnutí vytvořit výběrový soubor právě ze žáků SOŠ a SOU Kladno bylo učiněno z několika důvodů, mezi které patří např. technická a časová náročnost klasického pravděpodobnostního výběru a komplikované navazování spolupráce s různými středními školami na Kladensku. Před samotným hlavním průzkumem v Kladně byl proveden malý pilotní průzkum na doladění obsahu dotazníku s několika žáky SPŠE V Úžlabině. I když se jedná o oborově odlišnou školu, pro účely tohoto průzkumu neměl tento fakt velký význam, jelikož otázky se netýkají žádného konkrétního oboru. Účelem pilotního průzkumu bylo především ověřit, zda jsou otázky správně naformulovány a zda jsou žáci schopni a ochotni je zodpovědět. Přínos pilotního průzkumu potvrzují získané výsledky, kde se vyskytuje naprosté minimum chybějících odpovědí (viz kapitola Struktura souboru).

4.1 Dotazník pro žáky SOŠ a SOU Kladno

Dotazník, který v tištěné podobě vyplňovali žáci ze SOŠ a SOU Kladno, je součástí přílohy této práce. Následující text popisuje jeho strukturu, jejíž teoretické principy jsou nastíněny v kapitole Struktura dotazníku. Na začátku dotazníku se nachází jeho úvod, který žákům vysvětluje téma a účel průzkumu. Nejdůležitější součástí jsou instrukce k vyplnění, které byly detailněji rozepsány, aby žáci neměli žádné pochyby, jelikož u dotazníku není možnost dodatečného vysvětlení nejasností. Specifické pokyny k vyplnění jednotlivých otázek jsou poskytnuty před každou otázkou a jsou graficky odlišeny kurzívou. Prvních osm otázek spadá do hlavní části průzkumu; žáci byli tázáni na čas, který věnují studiu, místa, kde se připravují a osoby, se kterými se připravují. Dále se otázky týkaly jejich spokojenosti s oborem, známkami a maturity. Otázku číslo 2 nevyplňovali žáci prvních ročníků, protože se jednalo o srovnání přípravy do školy za více než jeden školní rok.

Závěrečné tři otázky byly tzv. identifikační, které jsou obzvláště důležité pro vyhodnocení, což bylo v dotazníku také řádně zdůrazněno. Celkem žáci vyplňovali 11 otázek. Většina otázek (sedm) byla zodpovězena pomocí ordinální škály a to buď přidělením bodů na stupnici od 0 do 10 nebo zakroužkováním slovní odpovědi. Dvě otázky měly nominální škálu (s kým se žáci učí nejčastěji a identifikace pohlaví) a dalších dvou se jednalo o kardinální proměnné, kdy žáci sami uváděli číselný údaj (počet hodin přípravy do školy, počet minut cesty do školy). Možnost volných odpovědí nebyla do dotazníku zahrnuta, protože tyto otázky zůstávají často bez odpovědi.¹³⁸

4.2 Zpracování dat

Data získaná v průzkumu byla vyhodnocena pomocí metod vhodných pro výběrové dotazníkové šetření. Tento typ vyhodnocení byl vybrán proto, jelikož se nejedná o typický pravděpodobnostní výběr, který by umožňoval usuzovat na populaci. Výhoda tohoto typu zpracování spočívá v tom, že po díky výsledkům statistických analýz je soubor lépe uchopitelný a na povrch vyplývají vztahy, které nemusí být na první pohled zřetelné. Šetření je přínosem i v tom smyslu, že poskytuje odpovědi na otázky, na které se nelze zeptat jinak než anonymním způsobem, protože jinak by mohlo dojít ke zkreslení dat. Výsledky mohou být užitečné také pro vedení školy, které je může využít při vylepšování podmínek studia nebo přístupu k výuce.

Údaje získané z dotazníků byly zpracovány pomocí programu *IBM SPSS Statistics* verze 20 (zkráceně pouze *SPSS*). Pro účely převedení dat do elektronické podoby byl vytvořen kódový klíč, jehož náhled z listu *Variable View* je součástí příloh této práce (viz příloha 2). Na základě definovaných proměnných, byly v listu *Data View* vloženy do programu hodnoty z vyplněných dotazníků (viz příloha 2). Následně byla data zkontrolována a vyčištěna od nesmyslných hodnot. Byla ošetřena problematika chybějících hodnot, které sice figurují v rámci tabulek rozdělení četností, ale neúčastní se zjišťování vztahů mezi proměnnými a porovnávání souborů.

4.2.1 Struktura výběrového souboru

Výběrový soubor se skládá ze žáků SOŠ a SOU Kladno, kteří studují jeden z následujících 4-letých denních oborů zakončený maturitní zkouškou: Veřejnosprávní

¹³⁸ RAAB-STEINER, BENESCH, cit. d., s. 48

činnost, Sociální činnost, Logistické a finanční služby a Obchodník. Průzkum proběhl v prvním týdnu nového pololetí, resp. na začátku února 2013.

Tabulka 4.1: Celkový počet respondentů

		Statistics	
		Pohlaví	Ročník studia
N	Valid	155	156
	Missing	1	0

Zdroj: vlastní zpracování

Celkem se průzkumu účastnilo 156 respondentů z 16 různých tříd, tzn., že se jedná o většinu žáků, kteří v současnosti zmíněné obory studují. Díky vysokému počtu je možno provést kvalitní vyhodnocení. Z dat vyplývá, že na třídu připadá průměrně necelých 10 žáků, kteří jsou ale na většinu obecných předmětů sloučeni s příbuzným oborem (např. obor Veřejnosprávní činnost s oborem Sociální činnost). Na otázku ohledně pohlaví odmítl odpovědět pouze jeden respondent, který proto nebude zařazen do většiny následujících statistických analýz. Naopak ročník uvedli všichni žáci (viz tabulka 4.1).

Tabulka 4.2: Rozdělení četností u proměnné Pohlaví

		Pohlaví			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	žena	112	71,8	72,3	72,3
	muž	43	27,6	27,7	100,0
Total		155	99,4	100,0	
Missing	System	1	,6		
Total		156	100,0		

Zdroj: vlastní zpracování

Tabulka 4.3: Rozdělení četností u proměnné Ročník studia

		Ročník studia			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	první	40	25,6	25,6	25,6
	druhý	33	21,2	21,2	46,8
	třetí	44	28,2	28,2	75,0
	čtvrtý	39	25,0	25,0	100,0
Total		156	100,0	100,0	

Zdroj: vlastní zpracování

Tabulka 4.2 zobrazuje rozdělení četností u proměnné Pohlaví v souboru. Z čísel vyplývá velká převaha dívek nad chlapci: žákyně tvoří 72,3% všech dotázaných. Důvodem převahy dívek je především zaměření oborů, které je pro dívky atraktivnější. Tabulka 4.3 se týká rozdělení četností v jednotlivých ročnících. Z údajů ve sloupci absolutními četnostmi je vidět, že nejslaběji zastoupeným ročníkem je ročník druhý, naopak nejčetnější je ročník třetí. Kumulované relativní četnosti dokládají o něco menší zastoupení prvních dvou ročníků, které dohromady tvoří 46,8%. Nicméně rozdíly jsou vcelku zanedbatelné.

4.2.2 Charakteristika jednotlivých proměnných

V následující části popisu výběrového souboru budou charakterizovány jednotlivé proměnné a to v pořadí, ve kterém se objevují v dotazníku.

Doba přípravy do školy

Tabulka 4.4: Míry polohy, variability a koncentrace pro proměnnou Doba přípravy do školy

Statistics		
Doba přípravy do školy		
N	Valid	155
	Missing	1
Mean		3,345
Median		2,500
Mode		1,0
Std. Deviation		3,234
		1
Variance		10,46
		0
Skewness		1,872
Std. Error of Skewness		,195
Kurtosis		4,943
Std. Error of Kurtosis		,387
Range		20,0
Percentiles	25	1,000
	50	2,500
	75	5,000

Zdroj: vlastní zpracování

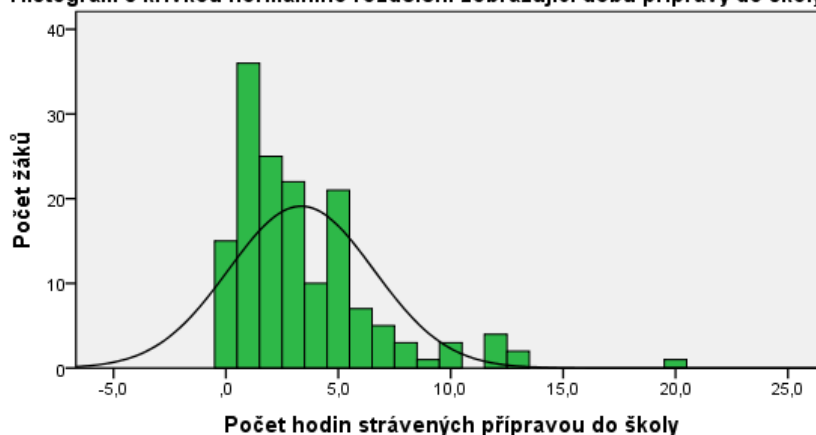
Tabulka 4.5: Rozdělení četností pro proměnnou Kategorie doby přípravy do školy

	Frequency	Valid Percent	Cumulative Percent	
Valid	méně než 0,5 hodiny týdně	15	9,7	9,7
	0,5 - 2,5 hodin týdně	68	43,9	53,5
	3 - 5 hodin týdně	46	29,7	83,2
	5,5 a více hodin týdně	26	16,8	100,0
	Total	155	100,0	
Missing	System	1		
Total		156		

Zdroj: vlastní zpracování

Graf 4.1: Rozdělení četností proměnné Doba přípravy do školy

Histogram s křivkou normálního rozdělení zobrazující dobu přípravy do školy



Zdroj: vlastní zpracování

Doba přípravy do školy je spojitá kvantitativní proměnná, proto je vyobrazena pomocí histogramu, přes který se line křivka normálního rozdělení (graf 4.1). Již na první pohled je vidět, že se výsledky od normálního rozdělení značně odchyľují. Domněnku nenormálního rozdělení pak potvrzují hlavně míry koncentrace (viz tabulka 4.4), které se jinak pohybují kolem 0. Zde ale hodnota míry špičatosti 1,872 ukazuje na kladně zešikmené rozdělení a vysoká kladná míra špičatosti (4,943) na zašpičatělé rozdělení. Proto ze středních hodnot bude brát v úvahu spíše medián (2,5) než aritmetický průměr (3,345). Pro přehlednost byla původní proměnná rozdělena do čtyř kategorií (tabulka 4.5). Podle údajů v tabulce 4.5 více než polovina žáků věnuje přípravě do školy maximálně 2,5 hodin týdně. Naopak těch, kteří se připravují 5,5 hodiny a více je podstatně méně (16,8%).

Tendence v intenzitě učení

Tabulka 4.6: Rozdělení četností u proměnné Tendence v intenzitě učení

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	čím vyšší ročník, tím MÉNĚ studuji studiu věnuji přibližně stále stejně času bez ohledu na ročník	17	10,9	14,7	14,7
	čím vyšší ročník, tím VÍCE studuji	62	39,7	53,4	68,1
	Total	116	74,4	100,0	
	Missing System	40	25,6		
Total		156	100,0		

Zdroj: vlastní zpracování

Na otázku č. dvě, která se týkala tendence v intenzitě učení, odpovídali pouze žáci 2. až 4. ročníků, protože žáci prvního ročníku neměli možnost srovnání. Z toho také vyplývá vysoké číslo chybějících odpovědí, které je totožné s počtem žáků prvního ročníku (viz tabulka 4.6). Zbylých 116 respondentů nejčastěji volilo prostřední neutrální odpověď, že studiu věnují stále stejně času bez ohledu na ročník. Naopak nejméně žáků mělo pocit, že by učení ubývalo. Názorný sloupcový graf je součástí přílohy (č. 3).

Místo přípravy do školy a frekvence jeho využití

Tabulka 4.7: Absolutní četnosti využití jednotlivých míst k přípravě do školy

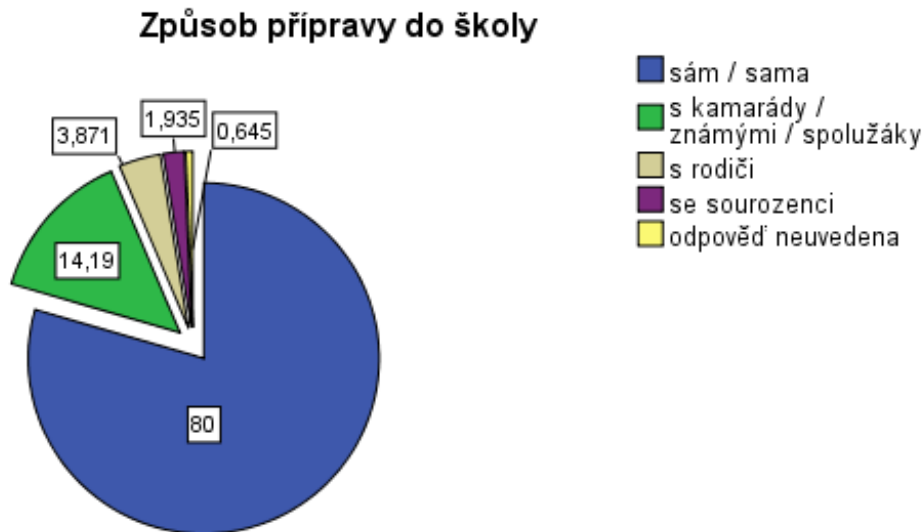
	většinou	často	občas	výjimečně	nikdy
	Count	Count	Count	Count	Count
Připravuji se doma	81	25	27	15	8
... při cestě do nebo ze školy	8	8	37	19	84
... ve škole	41	47	39	13	16
... ve studovně nebo v knihovně	0	0	3	4	149
... se u známých nebo kamarádů	6	1	17	24	108
... jinde	5	3	16	11	121

Zdroj: vlastní zpracování

Otázka č. 2 v dotazníku se týkala míst, kde se žáci učí. Absolutní četnosti jsou zobrazeny v tabulce 4.7, relativní řádkové četnosti jsou v příloze pod položkou č. 4. Vezmou-li se v úvahu pouze možnosti většinou až výjimečně, pak se nejvíce žáků (81) uvedlo, že se většinou připravují doma. 89 žáků většinou nebo často volí k přípravě školu. Ostatní místa jsou využívána poměrně méně, což dobře dokumentuje poslední sloupec, kde jsou uvedeny možnosti, které žáci vůbec nevyužívají. Nejméně je žáky využívána knihovna nebo studovna, kde se nikdy nepřipravovalo 95,5% respondentů.

Způsob přípravy do školy

Graf 4.2: Výšečový graf znázorňující, jaký způsob přípravy do školy vyhovuje žákům nejvíce

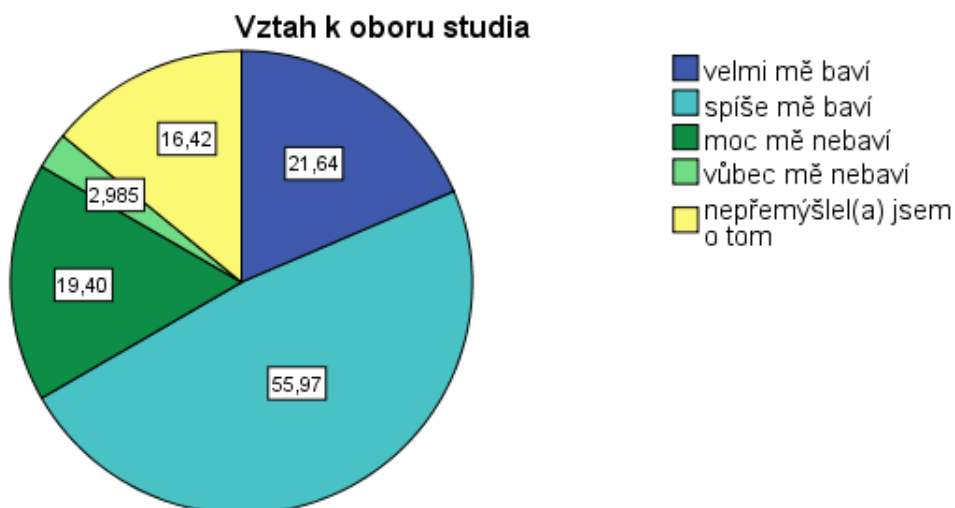


Zdroj: vlastní zpracování

Graf 4.2 zobrazuje relativní četnosti jednotlivých odpovědí na otázku zjišťující, jaký způsob přípravy vyhovuje žákům nejvíce. Drtivá většina (80%) preferuje učit se sám / sama. Druhou nejpočetnější kategorií je učení se s kamarády, známými nebo spolužáky; tuto možnost si vybralo 14,2% respondentů. Jeden respondent odpověď neuvedl.

Vztah k oboru studia

Graf 4.3: Výšečový graf pro proměnnou Vztah k oboru studia



Zdroj: vlastní zpracování

Na otázku, jestli studenty baví jejich obor, odpověděly kladně více než tři čtvrtiny žáků. Zbylou menší část naopak obor nebaví. Poměrně velká skupina žáků (16,42%) nad tím, zda je obor baví nebo nebaví, nezamýšlelo. Výsledky jsou znázorněny pomocí výšečového grafu (graf 4.3).

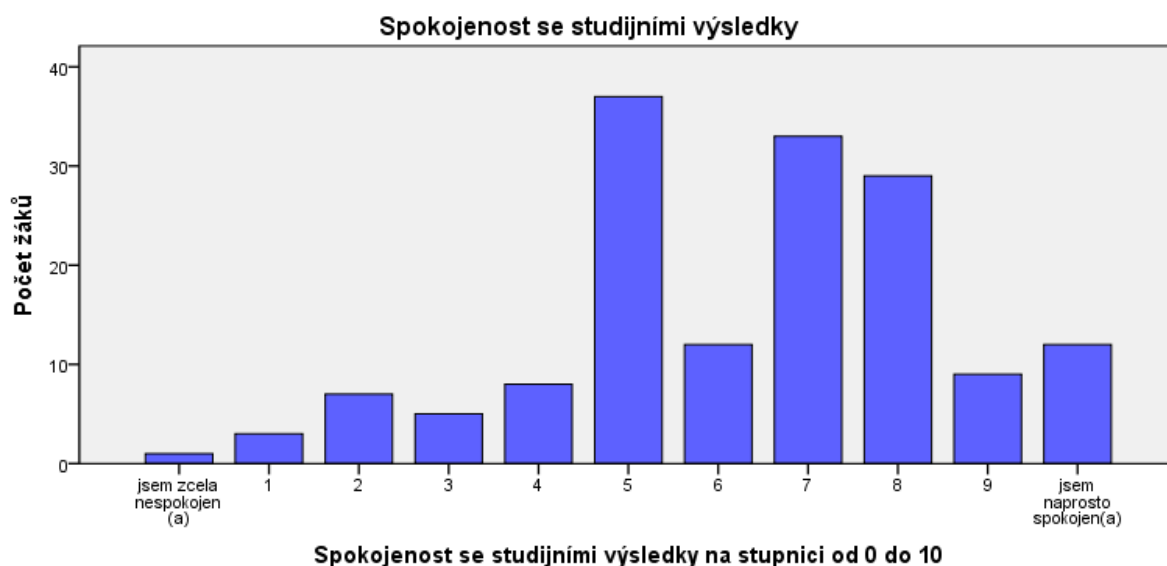
Spokojenost se studijními výsledky

Tabulka 4.8: Míry polohy a variability pro proměnnou Spokojenost se studijními výsledky

Statistics		
Spokojenost se studijními výsledky		
N	Valid	156
	Missing	0
Median		7,00
Mode		5
Percentiles	25	5,00
	50	7,00
	75	8,00

Zdroj: vlastní zpracování

Graf 4.4: Sloupcový graf pro proměnnou Spokojenost se studijními výsledky



Zdroj: vlastní zpracování

Z grafu 4.4 i tabulky 4.8 s popisnými charakteristikami vyplývá, že studenti jsou se svými studijními výsledky vcelku spokojeni. Dokládá to např. hodnota mediánu (7) a fakt, že mezikvantilové rozpětí, tedy interval obsahující 50% odpovědí respondentů, je mezi

hodnotami 5 a 8. Zajímavý je také fakt, že nejčetnější kategorií (modem) je střední kategorie (počet bodů 5), kterou vybralo 37 žáků.

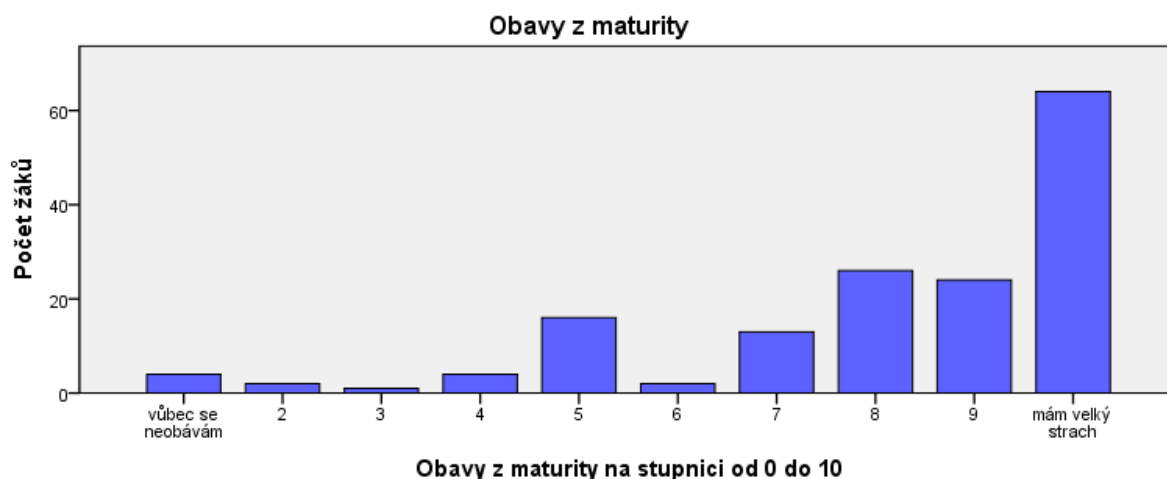
Obavy z maturity

Tabulka 4.9: Míry polohy a variability pro proměnnou Obavy z maturity

Statistics		
Obavy z maturity		
N	Valid	156
	Missing	0
Median		9,00
Mode		10
Percentiles	25	7,00
	50	9,00
	75	10,00

Zdroj: vlastní zpracování

Graf 4.5: Sloupcový graf pro proměnnou Obavy z maturity



Zdroj: vlastní zpracování

U hodnot proměnné Obavy z maturity je už na první pohled zřejmé, že se koncentrují spíše v horní polovině stupnice a ukazují tak na velké obavy z maturitní zkoušky (viz tabulka 4.9 a graf 4.5). Střední hodnota (medián) je dokonce číslo 9, tedy druhá nejvyšší možnost. Nejčetnější kategorií je dokonce ta nejvyšší, kterou zvolilo 64 žáků, což je 41% všech respondentů.

Příprava na maturitu

Tabulka 4.10: Rozdělení četností pro proměnnou Příprava na maturitu

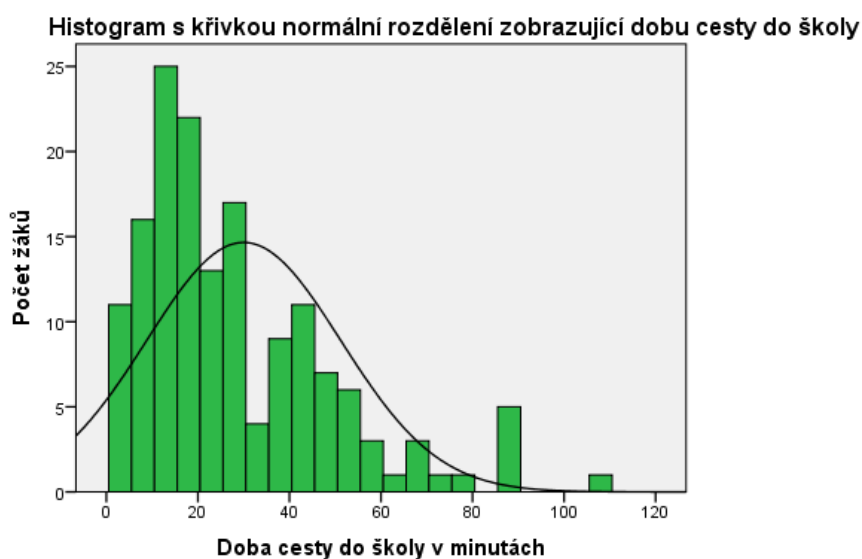
Příprava na maturitu				
	Frequency	Percent	Valid Percent	Cumulative Percent
nejvýše poslední týden (tzv. "svatý") před maturitou	3	1,9	1,9	1,9
asi 1 až 2 týdny	11	7,1	7,1	9,0
asi 2 týdny až 1 měsíc	23	14,7	14,7	23,7
asi 1 až 2 měsíce	37	23,7	23,7	47,4
asi 2 měsíce až půl roku	22	14,1	14,1	61,5
asi celý poslední rok (4. ročník)	60	38,5	38,5	100,0
Total	156	100,0	100,0	

Zdroj: vlastní zpracování

Z možností odpovědí na otázku, jak dlouho předpokládají, že se budou připravovat na maturitu, volili žáci především delší časová období. Více než polovina (52,5%) jich uvedla, že se chystají připravovat minimálně dva měsíce. Největší část studentů (38,5%) se plánuje připravovat dokonce celý poslední (čtvrtý) ročník. Na poslední chvíli, tedy na tzv. „svatý“ týden, plánují nechat přípravu pouze 3 studenti. Zdá se tedy, že žáci berou maturitní zkoušku opravdu vážně. Sloupcový graf vyobrazující absolutní četnosti z tabulky 4.10 je součástí přílohy (č. 5).

Doba cesty do školy

Graf 4.6: Graf pro rozdělení četností proměnné Doba cesty do školy



Zdroj: vlastní zpracování

Tabulka 4.11: Míry polohy, variability a koncentrace pro proměnnou Doba přípravy do školy

Statistics		
Doba cesty do školy		
N	Valid	156
	Missing	0
Mean		30,00
Median		25,00
Mode		15
Std. Deviation		21,229
Variance		450,658
Skewness		1,291
Std. Error of Skewness		,194
Kurtosis		1,589
Std. Error of Kurtosis		,386
Range		107
Percentiles	25	15,00
	50	25,00
	75	43,75

Zdroj: vlastní zpracování

Tabulka 4.12: Rozdělení četností pro kategorizovanou proměnnou Doba cesty do školy

Kategorie doby cesty do školy					
	Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	1 - 10 minut	27	17,3	17,3	17,3
	11 - 20 minut	47	30,1	30,1	47,4
	21 - 30 minut	30	19,2	19,2	66,7
	31 - 40 minut	13	8,3	8,3	75,0
	41 - 50 minut	18	11,5	11,5	86,5
	51 - 60 minut	9	5,8	5,8	92,3
	více než 1 hodina	12	7,7	7,7	100,0
	Total	156	100,0	100,0	

Zdroj: vlastní zpracování

Doba cesty do školy je druhou kvantitativní proměnnou v tomto průzkumu. Lze konstatovat, že se opět nejedná o normální rozdělení četností. Tuto skutečnost lze dobře dokumentovat na histogramu (graf 4.6), kde je vidět, že sloupce nekopírují průběh křivky normálního rozdělení. I hodnoty z tabulky 4.11 potvrzují, že se nejedná o normální rozdělení: rozdělení četností je mírně kladně zešikmené a mírně kladně zašpičatělé. Proto bude kladen důraz spíše na medián (25) než na aritmetický průměr (30). Zajímavé výsledky přináší tabulka 4.12, kde je původní spojitá proměnné rozdělena do kategorií po 10ti minutách s posledním otevřeným intervalem. Grafické znázornění se nachází v příloze (č. 6). Výsledky ukazují na to, že většina žáků bydlí v blízkosti školy, jelikož dvě třetiny se jich dopraví do budovy SOŠ a SOU Kladno do 30 minut. Téměř všichni žáci (92,3%) jsou ve škole do hodiny a pouze 7,69% žáků trvá cesta déle než hodinu.

4.2.3 Analýza závislosti

Vztah mezi pohlavím, resp. ročníkem a tendencí v intenzitě učení

Tabulka 4.13: Výsledky Chi-kvadrát testu o nezávislosti pro proměnné Pohlaví a Tendence v intenzitě učení

Chi-Square Tests				
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	4,712 ^a	2	,095	,104
Likelihood Ratio	5,047	2	,080	,081
Fisher's Exact Test	4,863			,091
Linear-by-Linear Association	3,140 ^b	1	,076	,091
N of Valid Cases	115			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,17.

b. The standardized statistic is -1,772.

Zdroj: vlastní zpracování

První provedená analýza závislosti se týká vztahu mezi pohlavím a tendencí v intenzitě učení. Nulová hypotéza je následující: Mezi chlapci a dívkami není statisticky významný rozdíl v tom, jak vnímají tendence v intenzitě učení. Naproti H_0 stojí hypotéza alternativní: Mezi chlapci a dívkami je statisticky významný rozdíl v tom, jak vnímají tendence v intenzitě učení. Za účel ověření H_0 byl proveden chí-kvadrát test o nezávislosti, který používá Pearsonovu statistiku chí-kvadrát (viz kapitola Analýza závislosti – Dvě nominální proměnné). Hladina významnosti α byla stanovena na úrovni 5%. Z tabulky 4.13 vyplývá, že podmínky testu byly splněny (žádná z očekávaných četností neklesla pod 5). Hodnota p (zvýrazněna žlutě) je ovšem vyšší než α (0,05), a proto nezamítáme H_0 o nezávislosti.

Provedeme-li stejný test tentokrát pro ročník a tendenci v intenzitě učení, získáme obdobné výsledky a nezamítáme nulovou hypotézu o nezávislosti těchto proměnných na hladině významnosti 5%. Z dat vypočtená p-hodnota testové statistiky chí-kvadrát totiž nabývá hodnoty 0,584. I zde jsou podmínky pro test splněny: pouze 1 políčko (11,1%) má očekávanou četnost menší než 5 (viz příloha č. 7).

Vztah mezi pohlavím, resp. ročníkem a místem přípravy do školy

Vztah mezi pohlavím a přípravou doma

Tabulka 4.14: Výsledky Chí-kvadrát testu o nezávislosti pro proměnné Pohlaví a Příprava doma

Chi-Square Tests				
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	28,338 ^a	4	,000	,000
Likelihood Ratio	28,405	4	,000	,000
Fisher's Exact Test	25,649			,000
Linear-by-Linear Association	17,109 ^b	1	,000	,000
N of Valid Cases	155			

a. 2 cells (20,0%) have expected count less than 5. The minimum expected count is 2,22.

b. The standardized statistic is 4,136.

Zdroj: vlastní zpracování

Tabulka 4.15: Hodnoty Cramérova V a ϕ pro proměnné Pohlaví a Příprava doma

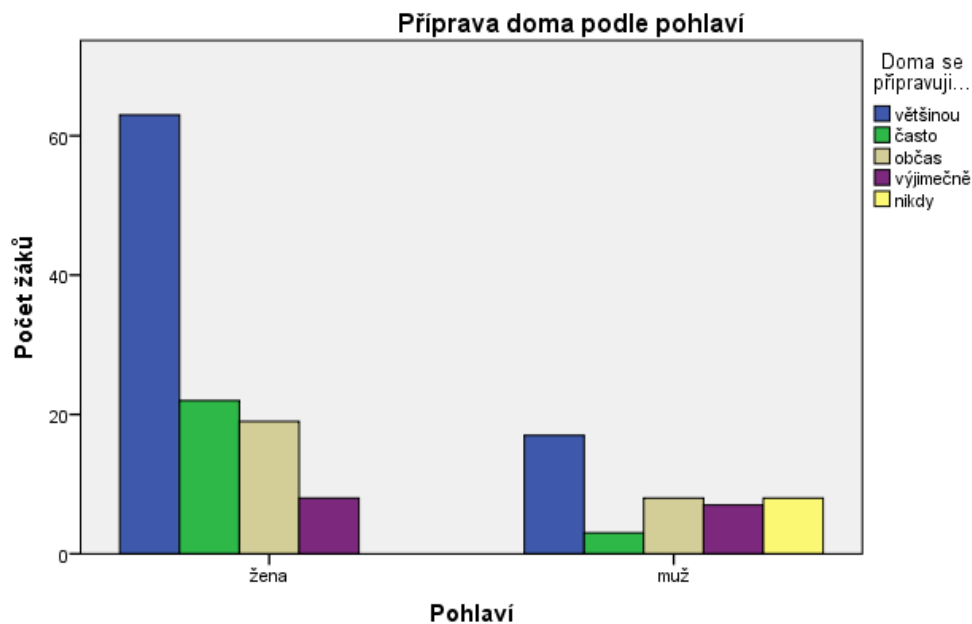
Symmetric Measures				
		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	,428	,000	,000
	Cramer's V	,428	,000	,000
N of Valid Cases		155		

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Zdroj: vlastní zpracování

Graf 4.7: Shlukový sloupcový graf pro přípravu doma podle pohlaví



Zdroj: vlastní zpracování

Tabulka 4.16: Kontingenční tabulka pro proměnné Pohlaví a Příprava na maturitu

Pohlaví * Přípravuji se doma Crosstabulation

		Doma se připravuji...					Total	
		většinou	často	občas	výjimečně	nikdy		
Pohlaví	žena	Count	63	22	19	8	0	112
		Adjusted Residual	1,9	1,9	-,2	-1,7	-4,7	
Pohlaví	muž	Count	17	3	8	7	8	43
		Adjusted Residual	-1,9	-1,9	,2	1,7	4,7	
Total		Count	80	25	27	15	8	155

Zdroj: vlastní zpracování

Závislost mezi proměnnými Pohlaví a Příprava doma byla zkoumána pomocí chí-kvadrát testu. Ačkoli je příprava doma proměnná ordinální, byl vybrán tento test, protože pohlaví je naopak proměnná nominální, a proto pro ni nelze použít postupy vhodné pro ordinální proměnné, což ovšem opačným směrem možné je. Nulová hypotéza pro tento test byla tato: Neexistuje statisticky významná závislost mezi pohlavím a přípravou doma; alternativní hypotéza: Existuje statisticky významná závislost mezi pohlavím a přípravou doma. Hladina významnosti byla opět 5%. Výsledky chí-kvadrát testu v tabulce 4.14 ukázaly, že podmínka pro test byla velmi těsně splněna; 20% políček má totiž očekávanou četnost nižší než 5. P-hodnota je 0,000, což znamená, že výsledek je signifikantní nejen na hladině významnosti 5%, ale i na 1% hladině, a proto se zamítá nulová hypotéza a přijímá alternativní. Tabulka 4.15 zobrazuje hodnotu Cramérova V, které udává sílu závislosti ve výši 0,428. Andy Field používá k hodnocení velikosti efektu stupnici, kde $\pm 0,1$ je malý efekt, $\pm 0,3$ je střední efekt a $\pm 0,5$ je velký efekt.¹³⁹ Na základě této stupnice, lze výsledek 0,428 ohodnotit jako větší střední závislost. Pro bližší určení povahy závislosti slouží adjustovaná rezidua s tabulky 4.16, která se poměřují s číslem 1,96 v absolutní hodnotě. Vyšší čísla pak ukazují na statisticky významný rozdíl.¹⁴⁰ V případě kategorií proměnné Příprava doma je statisticky významný rozdíl především u možnosti „nikdy“, kde adjustovaná rezidua nabývají v absolutní hodnotě velikosti 4,7. U možností „většinou“ a „často“ se hodnota 1,9 blíží hraniční 1,96. Výsledky tedy ukazují na to, že dívky uvedli, že se doma učí častěji než chlapci. Výsledky graficky znázorňuje shlukový graf 4.7, kde je

¹³⁹ FIELD, cit. d., s. 173

¹⁴⁰ KOZEL, MYNÁŘOVÁ, SVOBODOVÁ, cit. d., s. 118

vidět, že u dívek žlutý sloupec „nikdy“ zcela chybí a naopak dominují modrý („většinou“) a zelený („často“) sloupec.

Vztah mezi ročníkem a přípravou doma

Tabulka 4.17: Výsledky Chí-kvadrát testu o nezávislosti pro proměnné Ročník a Příprava doma (rekódovaná)

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4,064 ^a	3	,255
Likelihood Ratio	4,883	3	,181
Linear-by-Linear Association	2,512	1	,113
N of Valid Cases	156		

a. 1 cells (12,5%) have expected count less than 5. The minimum expected count is 4,87.
Zdroj: vlastní zpracování

Při zkoumání závislosti mezi ročníkem a přípravou doma bylo zjištěno, že nejsou splněny podmínky chí-kvadrát testu, jelikož 40% políček kontingenční tabulky mělo očekávané četnosti menší než 5. Proto byly sloučeny první tři a poslední dvě možnosti odpovědí a vznikly tak dvě kategorie: „alespoň občas“ a „výjimečně nebo nikdy“. Po provedení chí-kvadrát testu byla ovšem nezamítnuta nulová hypotéza o nezávislosti obou proměnných (viz p-hodnota 0,255 v tabulce 4.17).

Vztah mezi ročníkem a přípravou na cestě do nebo ze školy, ve škole, ve studovně či knihovně, u známých či kamarádů nebo jinde

Stejně jako v případě zkoumání závislosti mezi ročníkem a přípravou doma, bylo i u zkoumání vztahů mezi ročníkem a přípravou na cestě do nebo ze školy, ve škole, ve studovně či knihovně, u známých či kamarádů nebo jinde potřeba provést sloučení kategorií do stejných dvou skupin jako v předchozím případě, aby byly splněny podmínky o očekávaných četnostech. Poté byl znovu pro každou dvojici proveden chí-kvadrát test o nezávislosti, přičemž výsledky byly obdobné, jelikož ve všech případech nebyla zamítnuta nulová hypotéza o nezávislosti. U dvojice Ročník a Příprava ve studovně nebo v knihovně nebyly podmínky chí-kvadrát testu splněny ani po rekódování, jelikož v 50% políček byly očekávané četnosti nižší než 5. U zbylých dvojic vyšly p-hodnoty následovně: Příprava na

cestě do nebo ze školy (0,130)¹⁴¹, Příprava ve škole (0,452), Příprava u známých nebo kamarádů (0,305), Příprava jinde (0,410). Doposud byla prošetřena všechna místa, kromě cesty do a ze školy, které bude věnován speciální oddíl. Z výsledků zatím vyplývá, že ročník studia nesouvisí s frekvencí využívání různých míst přípravě do školy.

Vztah mezi pohlavím a přípravou na cestě do nebo ze školy ve škole, ve studovně či knihovně, u známých či kamarádů nebo jinde

Podmínky pro chí-kvadrát test o nezávislosti, kterým byly zkoumány výše uvedené dvojice, byly splněny pouze u páru Pohlaví a Příprava ve škole. V ostatních případech byla proměnná týkající se přípravy zrekódována na dvě kategorie „alespoň občas“ a „výjimečně nebo nikdy“, čímž vznikla čtyřpolní tabulka a následně byl použit Fisherův exaktní test. I tak nešlo z důvodu nízkých očekávaných četností provést test u dvojice Pohlaví a Příprava v knihovně nebo studovně, kde 50% políček mělo očekávané četnosti pod 5. Nulová hypotéza o nezávislosti nebyla na základě výsledků testů zamítnuta v těchto případech: Příprava na cestě do nebo ze školy (p-hodnota pro Fisherův exaktní test 0,354), Příprava ve škole (p-hodnota pro Pearsonovu statistiku chí-kvadrát 0,848), Příprava jinde (p-hodnota pro Fisherův exaktní test 0,223). U dvojice Pohlaví a Příprava u kamarádů nebo známých byla pomocí Fisherova exaktního testu zjištěna p-hodnota 0,024 (viz tabulka 4.18). Proto byla na hladině významnosti 5% zamítnuta nulová hypotéza o nezávislosti těchto dvou proměnných a přijata hypotéza alternativní hypotéza, že existuje statisticky významná závislost mezi pohlavím a využíváním přípravy do školy u známých nebo kamarádů. Pomocí Cramérova V i ϕ (jedná se o čtyřpolní tabulku) byla vyhodnocena velikost efektu jako malá (0,186) – viz tabulka 4.20. Na základě adjustovaných reziduí z tabulky 4.19 bylo zjištěno, že u známých nebo kamarádů se častěji připravují dívky.

¹⁴¹ Před testováním proměnné Příprava na cestě do nebo ze školy byly vyfiltrovány pouze hodnoty větší nebo rovny 20 minutám (toto platí i pro zjišťování vztahů mezi Pohlavím a Přípravou na cestě do nebo ze školy popsané v následujícím oddílu). Teprve od této hodnoty je totiž pravděpodobné, že žáci celou cestu neabsolvují pouze pěšky, ale použijí také nějaký dopravní prostředek, ve kterém se mohou učit. Způsob dopravy do školy nebyl v průzkumu zjišťován.

Tabulka 4.18: Výsledky Fisherova exaktního testu pro proměnné Pohlaví a Příprava u známých nebo kamarádů (rekódovaná)

Chi-Square Tests				
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	5,336 ^a	1	,021	
Continuity Correction ^b	4,252	1	,039	
Likelihood Ratio	6,464	1	,011	
Fisher's Exact Test				,024
Linear-by-Linear Association	5,302	1	,021	
N of Valid Cases	155			

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 6,66.

b. Computed only for a 2x2 table

Zdroj: vlastní zpracování

Tabulka 4.19: Kontingenční tabulka s adjustovanými rezidui pro proměnnou Pohlaví a zrekódovanou proměnnou Příprava u známých nebo kamarádů

Crosstab					
		Připravuji se u známých nebo kamarádů (zrekódováno)		Total	
		alespoň občas	výjimečně nebo nikdy		
Pohlaví	žena	Count	22	90	112
		Adjusted Residual	2,3	-2,3	
	muž	Count	2	41	43
		Adjusted Residual	-2,3	2,3	
Total		Count	24	131	155

Zdroj: vlastní zpracování

Tabulka 4.20: Hodnoty Cramérova V a ϕ pro proměnnou Pohlaví a zrekódovanou proměnnou Příprava u známých nebo kamarádů

Symmetric Measures		
	Value	Approx. Sig.
Nominal by Nominal	Phi	,186
	Cramer's V	,186
N of Valid Cases		155

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Zdroj: vlastní zpracování

Vztahy mezi proměnnými Doba přípravy do školy, Vztah k oboru studia, Spokojenost se studijními výsledky, Obavy z maturity a Příprava na maturitu

Tabulka 4.21: Kendallovo τ_b pro vybrané dvojice proměnných (zeleně jsou označeny statisticky významné vztahy na hladině významnosti 1%, žlutě na hladině významnosti 5%)

			Correlations				
			Doba přípravy do školy	Vztah k oboru studia	Spokojenost se studijními výsledky	Obavy z maturity	Příprava na maturitu
Kendall's tau_b	Doba přípravy do školy	Corr. Coeff.	1,000	-,106	,075	,071	,201**
		Sig. (2-tailed)	.	,132	,214	,249	,001
		N	155	133	155	155	155
	Vztah k oboru studia	Corr. Coeff.	-,106	1,000	-,157*	-,043	-,256**
		Sig. (2-tailed)	,132	.	,030	,558	,001
		N	133	134	134	134	134
	Spokojenost se studijními výsledky	Corr. Coeff.	,075	-,157*	1,000	-,151*	,072
		Sig. (2-tailed)	,214	,030	.	,017	,256
		N	155	134	156	156	156
	Obavy z maturity	Corr. Coeff.	,071	-,043	-,151*	1,000	,147*
		Sig. (2-tailed)	,249	,558	,017	.	,024
		N	155	134	156	156	156
	Příprava na maturitu	Corr. Coeff.	,201**	-,256**	,072	,147*	1,000
		Sig. (2-tailed)	,001	,001	,256	,024	.
		N	155	134	156	156	156

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Zdroj: vlastní zpracování

Tabulka 4.22: Testy normality pro proměnnou Doba přípravy do školy

	Tests of Normality					
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Doba přípravy do školy	,175	155	,000	,826	155	,000

a. Lilliefors Significance Correction

Zdroj: vlastní zpracování

Pro zjišťování korelace byly vybrány ordinální proměnné Vztah k oboru studia, Spokojenost se studijními výsledky, Obavy z maturity, Příprava na maturitu a také jedna kardinální proměnná – Doba přípravy do školy. U poslední zmiňované proměnné byly

provedeny testy ověřující normalitu rozdělení – viz tabulka 4.22. Nulová hypotéza, že se jedná o normální rozdělení četností, byla na základě výsledků Kolmogorova-Smirnova testu i Shapirova-Wilkova testu zamítnuta. Proto bylo pro zjišťování korelace využito Kendallova τ_b (tabulka 4.21) a pro srovnání také Spearmanova korelačního koeficientu (viz příloha 8). Statisticky významné vztahy na hladině významnosti α 1% jsou v tabulce 4.21 vyznačeny dvěma hvězdičkami a zvýrazněny zeleně, pokud $\alpha=5\%$, pak jsou vyznačeny jen jednou hvězdičkou a zvýrazněny žlutě. Jedná se o následující vztahy (včetně určení směru korelace):

- Čím více se žák připravuje do školy, tím více se plánuje připravovat na maturitu.
- Čím více žáka baví jeho obor, tím více se plánuje připravovat na maturitu.
- Čím více žáka baví jeho obor, tím je spokojenější se svými studijními výsledky.
- Čím více je žák spokojen se svými studijními výsledky, tím méně se obává maturity.
- Čím více se žák obává maturity, tím více se plánuje na ni připravovat.

Všechny právě uvedené vztahy lze označit spíše za slabé, jelikož se u nich absolutní hodnoty Kendallova τ_b pohybují od 0,147 do 0,256.

4.2.4 Porovnávání skupin

Doba přípravy do školy podle ročníku studia

Tabulka 4.23: Testy normality pro proměnnou Doba přípravy do školy podle ročníku studia

Tests of Normality							
	Ročník studia	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Doba přípravy do školy	první	,262	40	,000	,756	40	,000
	druhý	,234	32	,000	,663	32	,000
	třetí	,250	44	,000	,825	44	,000
	čtvrtý	,138	39	,060	,920	39	,009

a. Lilliefors Significance Correction
Zdroj: vlastní zpracování

Tabulka 4.24: Levenův test na homogenitu rozptylů proměnné Doba přípravy do školy v jednotlivých ročnících

Test of Homogeneity of Variances

Doba přípravy do školy

Levene Statistic	df1	df2	Sig.
,902	3	151	,442

Zdroj: vlastní zpracování

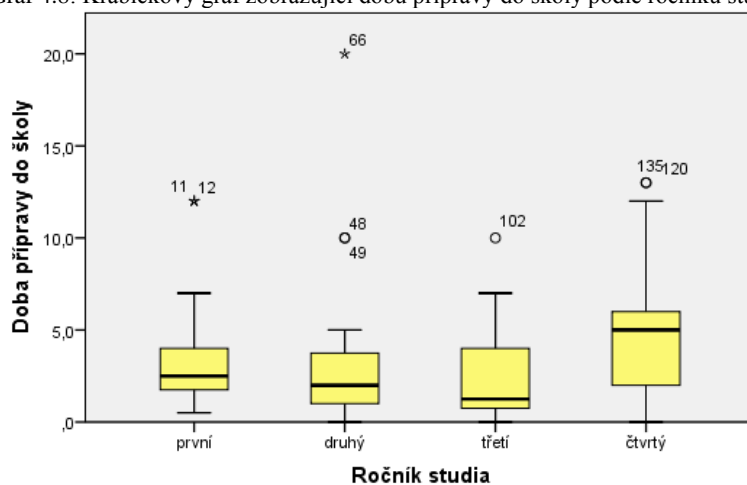
Tabulka 4.25: Výsledky Kruskalova-Wallisova testu pro proměnnou Doba přípravy do školy podle ročníku studia

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Doba přípravy do školy is the same across categories of Ročník studia.	Independent-Samples Kruskal-Wallis Test	,004	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Zdroj: vlastní zpracování

Graf 4.8: Krabičkový graf zobrazující dobu přípravy do školy podle ročníku studia



Zdroj: vlastní zpracování

Tabulka 4.26: Výsledky Mannova-Whitneyho testu pro 1. a 4. ročník

Test Statistics ^a	
	Doba přípravy do školy
Mann-Whitney U	614,000
Wilcoxon W	1434,000
Z	-1,634
Exact Sig. (2-tailed)	,103
Point Probability	,001

a. Grouping Variable: Ročník studia

Zdroj: vlastní zpracování

Tabulka 4.27: Výsledky Mannova-Whitneyho testu pro 2. a 4. ročník

Test Statistics ^a	
	Doba přípravy do školy
Mann-Whitney U	441,500
Wilcoxon W	969,500
Z	-2,121
Exact Sig. (2-tailed)	,034
Point Probability	,000

a. Grouping Variable: Ročník studia

Zdroj: vlastní zpracování

Tabulka 4.28: Výsledky Mannova-Whitneyho testu pro 3. a 4. ročník

Test Statistics ^a	
	Doba přípravy do školy
Mann-Whitney U	527,000
Wilcoxon W	1517,000
Z	-3,043
Exact Sig. (2-tailed)	,002
Point Probability	,000

a. Grouping Variable: Ročník studia

Zdroj: vlastní zpracování

Pro porovnání doby přípravy do školy podle ročníku by byla vhodná analýza rozptylu. Provedením Shapiro-Wilkova testu normality ale bylo zjištěno, že rozdělení proměnné Doba přípravy do školy není ani v jednom z ročníků normální a byly tak zamítnuty nulové hypotézy o normalitě (tabulka 4.23). Druhá podmínka pro ANOVU by sice splněna byla, rozptyly jsou totiž napříč ročníky homogenní, jak vyplývá z Levenova testu (tabulka 4.24). Přesto ale musí být kvůli nenormalitě rozdělení použit Kruskalův-Wallisův neparametrický test. Jeho výsledek je shrnut v tabulce 4.25: na hladině významnosti 5% je zamítnuta nulová hypotéza o tom, že neexistují statisticky významné rozdíly mezi ročníky a je přijata alternativní hypotéza. Pro lepší přehled o výsledcích byl vygenerován krabičkový graf (graf 4.8), ze kterého vzešla domněnka, že rozdíl se týká čtvrtého ročníku, jehož medián, mezikvantilové i variační rozpětí se zdají být vyšší než u ostatních ročníků. Pro post-hoc analýzu byly vybrány tři dvojice, které se podrobily Mannovu-Whitneyovu testu. Hladina významnosti byla upravena tzv. Bonferroniho korekcí: $\alpha 0,05$ se vydělila počtem post-hoc testů, což je 3: $(0,05 : 3 = 0,0167)$. 0,0167 byla tedy hladinou významnosti, se kterou se poměřovaly p-hodnoty post-hoc testů. Nulová hypotéza o nezávislosti mezi ročníky byla zamítnuta pouze v případě třetího a čtvrtého ročníku (viz tabulky 4.26, 4.27 a 4.28). Z krabičkového grafu lze odvodit, že žáci čtvrtého ročníku se v průměru připravují do školy více než žáci třetího ročníku.

Doba přípravy do školy podle pohlaví

Tabulka 4.29: Testy normality pro proměnnou Doba přípravy do školy podle pohlaví

Tests of Normality							
	Pohlaví	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Doba přípravy do školy	žena	,174	111	,000	,863	111	,000
	muž	,276	43	,000	,585	43	,000

a. Lilliefors Significance Correction
Zdroj: vlastní zpracování

Tabulka 4.30: Levenův test na homogenitu rozptylů proměnné Doba přípravy do školy v jednotlivých ročnících

Test of Homogeneity of Variances

Doba přípravy do školy

Levene Statistic	df1	df2	Sig.
,776	1	152	,380

Zdroj: vlastní zpracování

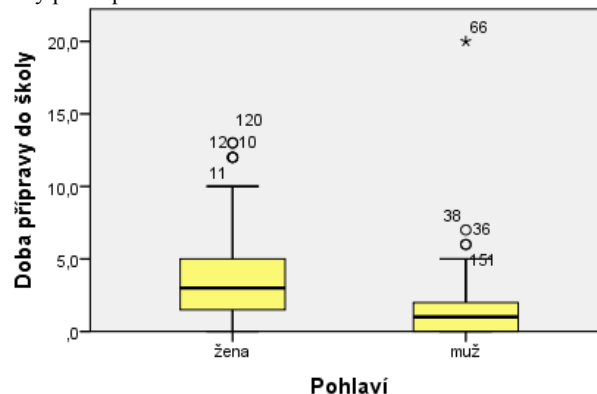
Tabulka 4.31: Výsledky Mannova-Whitneyho testu pro proměnnou Doba přípravy do školy podle pohlaví

Hypothesis Test Summary			
Null Hypothesis	Test	Sig.	Decision
The distribution of Doba přípravy do školy is the same across categories of Pohlaví.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Zdroj: vlastní zpracování

Graf 4.9: Krabičkový graf zobrazující dobu přípravy do školy podle pohlaví



Zdroj: vlastní zpracování

Pro porovnání doby přípravy do školy podle pohlaví byl použit neparametrický Mannův-Whitneyův test. I když jsou rozptyly u obou pohlaví homogenní (viz výsledky Levenova testu v tabulce 4.30), rozdělení četností doby přípravy do školy v ročnících není normální (tabulka 4.29). Jak vyplývá z tabulky 4.31, nulová hypotéza byla testem na hladině významnosti 5% zamítnuta a přijata byla alternativní hypotéza, která říká, že existují významné rozdíly mezi muži a ženami, co se týče doby přípravy do školy. Graf 4.9 tuto skutečnost dobře ilustruje a ukazuje na to, že dívky se do školy připravují více než chlapci, jelikož krabička pro dívky je mnohem výše a výše je také její medián znázorněný tlustou čarou uvnitř krabičky. V Příloze 9 se nachází graf-strom, na kterém jsou rozdíly mezi skupinami také dobře pozorovatelné a také hodnota testové statistiky.

Spokojenost se studijními výsledky podle pohlaví a podle ročníku studia

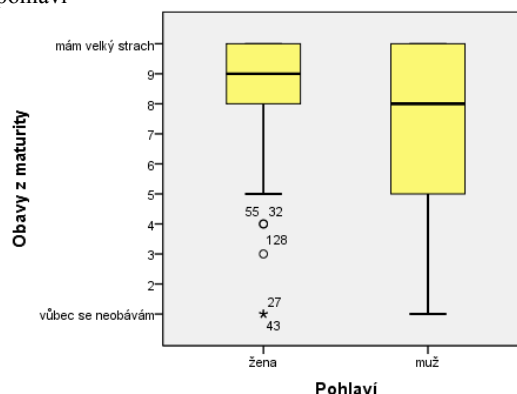
V případě spokojenosti se studijními výsledky se jedná o ordinální proměnnou, a proto byly opět použity neparametrické testy. Pro porovnání mužů a žen byl zvolen Mannův-Whitneyho test a pro porovnání ročníků Kruskalův-Wallisův test. V obou případech nebyla zamítnuta nulová hypotéza o nezávislosti souborů na hladině významnosti 5%. V prvním případě byla totiž z dat vypočtená p-hodnota 0,651, v druhém případě pak 0,371.

Obavy z maturity podle pohlaví a ročníku studia

Tabulka 4.32: Výsledky Mannova Whitneyho testu pro Graf 4.10: Krabičkový graf zobrazující obavy z maturity podle proměnnou Obavy z maturity podle pohlaví

Test Statistics ^a	
	Obavy z maturity
Mann-Whitney U	1680,000
Wilcoxon W	2626,000
Z	-3,033
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Pohlaví
Zdroj: vlastní zpracování



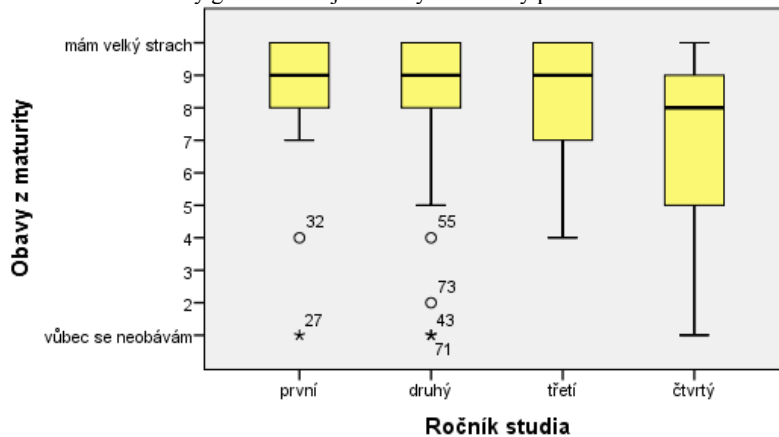
Zdroj: vlastní zpracování

Tabulka 4.33: Výsledky Kruskalova-Wallisova testu pro proměnnou Obavy z maturity podle ročníku studia

Test Statistics ^{a,b}	
	Obavy z maturity
Chi-Square	8,062
df	3
Asymp. Sig.	.045

a. Kruskal Wallis Test
b. Grouping Variable: Ročník studia
Zdroj: vlastní zpracování

Graf 4.11: Krabičkový graf zobrazující obavy z maturity podle ročníku studia



Zdroj: vlastní zpracování

Porovnávání souborů u proměnné Obavy z maturity proběhlo obdobně jako u Doby přípravy do školy, byly použity stejné principy i testy popsané výše. Nejprve budou popsány výsledky pro porovnání podle pohlaví. Na hladině významnosti 5% byla zamítnuta nulová hypotéza o tom, že neexistují významné rozdíly mezi chlapci a dívkami (viz tabulka 4.32). Následně byl vytvořen krabičkový graf (graf 4.10), ze kterého lze vyčíst, že u dívek se odpovědi koncentrovaly u vyšších hodnot, což znamená, že mají z maturity větší obavy než chlapci.

U jednotlivých ročníků bylo také zjištěno, že existují významné rozdíly mezi ročníky (tabulka 4.33). Na základě krabičkového grafu (graf 4.11) byly vybrány tři páry k post-hoc analýze: 1. a 4., 2. a 4., 3. a 4. ročník. Hladina významnosti pro tyto post-hoc

Mannovy-Whitneyovy testy byla po Bonferroniho korekci 0,167. U první dvojice (1. a 4. ročník) byly prokázány statisticky signifikantní rozdíly v obavách z maturity, u ostatních dvojic nikoli. Krabičkový graf ukazuje na to, že žáci prvního ročníku se maturity bojí více než žáci čtvrtého ročníku.

Příprava na maturitu podle pohlaví a ročníku studia

I u porovnávání Přípravy na maturitu podle pohlaví a ročníku studia byl vybrán shodný postup jako u Doby přípravy studia podle pohlaví a ročníku studia. V obou případech byla na hladině α 5% zamítnuta nulová hypotéza, že neexistují statisticky významné rozdíly mezi soubory (viz příloha č. 10 a č. 11). V případě pohlaví bylo zjištěno, že dívky se plánují připravovat na maturitu déle než chlapci (viz graf v příloze č. 12). V případě ročníků byl pomocí tří post-hoc testů odhalen statisticky významný rozdíl mezi třetím a čtvrtým ročníkem. Žáci třetího ročníku se totiž plánují připravovat v průměru déle než žáci čtvrtého ročníku. Situaci ilustruje krabičkový graf v příloze – č. 13. Na něm je vidět, že „čtvrťáci“ jsou ve svých plánech již realističtější a mezikvantilové rozpětí je proto širší a medián spíše nižší.

5 Závěr

V této bakalářské práci byly v části s teoretickými východisky popsány zásady tvorby dotazníků a provádění průzkumů. V textu je přehledně shrnuto, co všechno má dobře vytvořený dotazník obsahovat a bylo také pojednáno o typech otázek a o související problematice proměnných a škál. Na část o návrhu dotazníku navazují kapitoly o jeho zpracování. Zde jsou představeny metody vhodné pro výběrové dotazníkové šetření, které jsou následně využity v části vlastního zpracování. Postupy lze shrnout do tří větších celků: popis souboru a proměnných, analýza závislostí a porovnání souborů.

Vlastní zpracování je věnováno průzkumu postojů žáků ke škole, maturitě a hlavně přípravě do školy. Šetření se uskutečnilo na začátku února 2013 na SOŠ a SOU Kladno. Celkem dotazník vyplnilo 156 respondentů, z nichž většina byly dívky, jelikož škola, kde průzkum probíhá, nabízí obory, které preferují právě dívky. Data byla zpracována v programu *IBM SPSS Statistics* verze 20. Ke zpracování byly použity hlavně metody vhodné pro nenormálně rozdělené proměnné a ordinální škály. Z výsledků vplynuly některé zajímavé poznatky, které by mohlo využít jak vedení školy, tak i žáci samotní, aby se dozvěděli více o názorech své skupiny. Zde je souhrn toho nejpodstatnějšího: 1) Popis proměnných: přes 50% žáků se do školy připravuje v průměru maximálně 2,5 hodiny týdně; žáci 2. až 4. ročníků mají pocit, že přípravě věnují stále stejně času bez ohledu na ročník; drtivá většina žáků (80%) se učí sama; žáci jsou vcelku spokojeni se svými studijními výsledky; obavy z maturity jsou velmi vysoké; 38,5% žáků se plánuje připravovat na maturitu celý poslední ročník; 2/3 žáků se dopraví do školy do 30 minut. 2) Analýza závislostí: frekvence přípravy doma a u známých nebo kamarádů souvisí s pohlavím (dívky se doma i u známých nebo kamarádů připravují častěji), ale nikoli s ročníkem; u ostatních míst přípravy do školy nebyly prokázány statisticky významné vztahy s pohlavím nebo ročníkem; čím více se žák připravuje do školy, tím více se plánuje připravovat na maturitu; čím více žáka baví jeho obor, tím více se plánuje připravovat na maturitu a je také spokojenější se svými studijními výsledky a méně se obává maturity. 3) Porovnávání skupin: doba přípravy do školy souvisí s ročníkem i pohlavím: významný rozdíl byl zjištěn mezi třetáky a čtvrtáky, dívky se připravují více než chlapci; naopak spokojenost se studijními výsledky nesouvisí ani s pohlavím, ani s ročníkem; ohledně obav z maturity se prokázalo, že se jí obávají více dívky a ve srovnání se čtvrtáky také třetáci; u

zmíněných skupin bylo také zjištěno, že se plánují připravovat na maturitu déle. Závěrem lze říci, že cíle bakalářské práce stanovené v úvodu byly splněny.

Seznam použitých zdrojů

1. ESSER, Elke, HILL, Paul B., SCHNELL, Rainer. *Methoden der empirischen Sozialforschung*. 7., völlig überarb. und erw. Aufl. München [u.a.]: Oldenbourg, 2005. ISBN 978-348-6576-849.
2. FIELD, Andy P. *Discovering statistics using SPSS: (and sex, drugs and rock 'n' roll)*. 3rd ed. Los Angeles: SAGE Publications, 2009, xxxii, 821 s. ISBN 978-1-84787-906-6.
3. GALTON, Francis. *Memories of my life*. London: Methuen, 2001. ISBN 978-114-2726-836.
4. KIRCHHOFF, Sabine a kol. *Der Fragebogen Datenbasis, Konstruktion und Auswertung*. 5. Aufl. Wiesbaden: VS, Verl. für Sozialwiss, 2010. ISBN 978-353-1167-886.
5. KOZEL, Roman, MYNÁŘOVÁ, Lenka, SVOBODOVÁ, Hana. *Moderní metody a techniky marketingového výzkumu*. 1. vyd. Praha: Grada, 2011, 304 s. Expert (Grada). ISBN 978-80-247-3527-6.
6. KOZEL, Roman. *Moderní marketingový výzkum: nové trendy, kvantitativní a kvalitativní metody a techniky, průběh a organizace, aplikace v praxi, přínosy a možnosti*. 1. vyd. Praha: Grada, 2006, 277 s. ISBN 80-247-0966-X.
7. LEONHART, Rainer. *Datenanalyse mit SPSS*. Göttingen: Hogrefe, 2010. ISBN 978-380-1721-640.
8. MACHKOVÁ, Hana. *Mezinárodní marketing*. 2., rozš. a přeprac. vyd. Praha: Grada, 2006, 205 s. ISBN 80-247-1678-X.
9. NAUMANN, Ewald a kol. *Quantitative Methoden 2*. 1. Aufl. Berlin [u.a.]: Springer, 2004. ISBN 978-35-40201656.
10. NOVÝ, Ivan, SURYNEK, Alois. *Sociologie pro ekonomy a manažery*. 2., přeprac. a rozš. vyd. Praha: Grada, 2006, 288 s. Manažer. ISBN 80-247-1705-0.
11. PECÁKOVÁ, Iva, NOVÁK, Ilja, HERZMANN, Jan. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. Vyd. 3. Praha: Oeconomica, 2004, 146 s. ISBN 80-245-0753-6.
12. PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 2. dopl. vyd. Praha: Professional Publishing, 2011, 236 s. ISBN 978-80-7431-039-3.

13. RAAB-STEINER, Elisabeth, BENESCH, Michael. *Der Fragebogen: Von der Forschungsidee zur SPSS / PASW-Auswertung*. 2., aktualisierte Aufl. Wien: UTB, 2008. ISBN 978-382-5284-060.
14. RIMARČÍK, Marián. *Štatistika pre prax*. Košice: M. Rimarčík, 2007, 200 s. ISBN 978-80-969813-1-1.
15. ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření*. 3. aktualiz. vyd. Praha: Professional Publishing, 2011, 223 s. ISBN 978-807-4310-621.
16. SOUČEK, Eduard. *Statistika pro ekonomy*. Vyd. 1. Praha: Vysoká škola ekonomie a managementu, 2006. ISBN 978-808-6730-066.
17. TOUTENBURG, Helge, HEUMANN, Christian, SCHOMAKER, Michael. *Deskriptive Statistik: eine Einführung in Methoden und Anwendungen mit R und SPSS*. 7., aktualisierte u. erw. Aufl. Berlin: Springer, 2009. ISBN 978-364-2018-343.
18. URBAN, Lukáš. *Sociologie trochu jinak*. 2., rozš. vyd. Ilustrace Zdeněk Prošek. Praha: Grada, 2011, 271 s. Sociologie (Grada). ISBN 978-802-4735-627.
19. WALKER, Ian. *Výzkumné metody a statistika*. Vyd. 1. Editor Nigel Holt, Robert Lewis. Praha: Grada, 2013, 218 s. Z pohledu psychologie. ISBN 978-802-4739-205.
20. Rada vlády pro konkurenceschopnost a informační společnost. [online]. [cit. 5.3.2013]. Dostupné z: <http://www.vlada.cz/cz/ppov/rvis/rada-vlady-pro-informacni-spolecnost-73372/>

Přílohy

Příloha č. 1: Dotazník pro žáky SOŠ a SOU Kladno

Vážení žáci,

právě před vámi leží dotazník týkající se učení a přípravy do školy. Ráda bych vás tímto požádala o jeho vyplnění. Přečtěte si prosím pozorně otázky a instrukce k jejich zodpovězení a odpovídejte upřímně a podle skutečnosti. Postupujte otázku po otázce tak, abyste žádnou otázku nevynechali (s výjimkou otázky č. 2, kterou prváci nevyplňují). Dotazník je anonymní a výsledky budou použity pro účely mé bakalářské práce.

Děkuji vám za spolupráci!

Tereza Hadravová,
studentka České zemědělské univerzity v Praze

1. Kolik hodin týdně průměrně věnuješ přípravě do školy (včetně psaní domácích úkolů)? *Údaj uveď zaokrouhleně na půlhodiny, tj. např. 3,5 hodiny, 8 hodin apod.*

 hodin

Následující otázku zodpovědí pouze žáci druhého až čtvrtého ročníku (tedy NE prváci).

2. Myslíš si, že čím vyšší ročník školy navštěvuješ, tím více se učíš nebo je to spíše naopak? *Vyber pouze jednu možnost.*
- čím vyšší ročník, tím VÍCE studuji
 - čím vyšší ročník, tím MÉNĚ studuji
 - studiu věnuji přibližně stále stejně času bez ohledu na ročník
3. Kde všude se připravuješ do školy? *Lze uvést více možností. Uveď všechny, které využíváš a uveď také jak často.*

	ANO	NE	většinou	často	občas	výjimečně
doma	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
při cestě do nebo ze školy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ve škole	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ve studovně nebo v knihovně	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
u známých nebo kamarádů	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
jinde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. Který způsob přípravy do školy ti vyhovuje nejvíce? *Vyber pouze jednu možnost.*
- sám / sama
 - s kamarády / známými / spolužáky
 - s rodiči
 - se sourozenci

5. Baví tě obor, který studuješ? *Vyber pouze jednu možnost.*
- velmi mě baví
 - spíše mě baví
 - moc mě nebaví
 - vůbec mě nebaví
 - nepřemýšlel(a) jsem o tom
6. Jak jsi spokojen(a) se svými studijními výsledky? *Vyjádři množstvím přidělených bodů od 0 do 10, kde 0 = jsem zcela nespokojen(a) a 10= jsem naprosto spokojen(a).*

bodů

7. Obáváš se maturity? *Vyjádři množstvím přidělených bodů od 0 do 10, kde 0 = vůbec se neobávám a 10= mám velký strach*

bodů

8. Na maturitu se žáci obvykle připravují intenzivněji než v průběhu studia. Jak dlouho předpokládáš, že se budeš intenzivněji věnovat přípravě na maturitu (procházení maturitních otázek)? *Vyber pouze jednu možnost.*
- nejvýše poslední týden (tzv. „svatý“) před maturitou
 - asi 1 až 2 týdny
 - asi 2 týdny až 1 měsíc
 - asi 1 až 2 měsíce
 - asi 2 měsíce až půl roku
 - asi celý poslední rok (4. ročník)

Na závěr ještě několik důležitých otázek, které jsou zásadní pro vyhodnocení tvých odpovědí.

9. Pohlaví.
- žena
 - muž
10. Ročník studia.
- první
 - druhý
 - třetí
 - čtvrtý

11. Průměrná doba cesty do školy (od bydliště k budově školy). *Uved' v minutách, tj. např. 10 minut, 85 minut, apod.*

minut

Zkontroluj si, prosím, ještě jednou, zda je vyplněný celý dotazník.

Děkuji za vaše odpovědi a váš čas!

Příloha č. 2: List *Variable View* a *Data View* v SPSS se 17 proměnnými z dotazníku

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	id	Numeric	8	0	Identifikace	None	None	3	Right	Scale	Input
2	A1	Numeric	8	1	Doba přípravy do školy	None	None	3	Right	Scale	Input
3	A2	Numeric	8	0	Tendence v intenzitě učení	{-1, čím vyš...	None	3	Right	Ordinal	Input
4	A3.1	Numeric	8	0	Připravuji se doma	{1, většinou...	None	3	Right	Ordinal	Input
5	A3.2	Numeric	8	0	Připravuji se při cestě do nebo ze školy	{1, většinou...	None	3	Right	Ordinal	Input
6	A3.3	Numeric	8	0	Připravuji se ve škole	{1, většinou...	None	3	Right	Ordinal	Input
7	A3.4	Numeric	8	0	Připravuji se ve studovně nebo v knihovně	{1, většinou...	None	3	Right	Ordinal	Input
8	A3.5	Numeric	8	0	Připravuji se u známých nebo kamarádů	{1, většinou...	None	3	Right	Ordinal	Input
9	A3.6	Numeric	8	0	Připravuji se jinde	{1, většinou...	None	3	Right	Ordinal	Input
10	A4	Numeric	8	0	Způsob přípravy do školy	{1, sám / s...	None	3	Right	Nominal	Input
11	A5	Numeric	8	0	Vztah k oboru studia	{1, velmi m...	9	3	Right	Ordinal	Input
12	A6	Numeric	8	0	Spokojenost se studijními výsledky	{0, jsem zc...	None	3	Right	Ordinal	Input
13	A7	Numeric	8	0	Obavy z maturity	{0, vůbec s...	None	3	Right	Ordinal	Input
14	A8	Numeric	8	0	Příprava na maturitu	{1, nejvýše ...	None	3	Right	Ordinal	Input
15	B1	Numeric	8	0	Pohlaví	{1, žena}...	None	3	Right	Nominal	Input
16	B2	Numeric	8	0	Ročník studia	{1, první}...	None	3	Right	Ordinal	Input
17	B3	Numeric	8	0	Doba cesty do školy	None	None	3	Right	Scale	Input

	id	A1	A2	A3.1	A3.2	A3.3	A3.4	A3.5	A3.6	A4	A5	A6	A7	A8	B1	B2	B3
1	1	2,5	.	1	3	4	5	5	5	1	2	8	10	6	1	1	110
2	2	3,0	.	1	5	2	5	5	3	1	2	7	10	6	1	1	90
3	3	3,0	.	1	5	2	5	5	3	1	2	8	9	6	1	1	5
4	4	3,0	.	1	5	2	5	5	3	1	2	8	9	6	1	1	60
5	5	2,5	.	1	5	4	5	5	5	1	1	5	10	6	1	1	15
6	6	2,0	.	1	5	1	5	5	5	1	1	7	9	6	1	1	20
7	7	3,0	.	1	5	2	5	5	5	3	2	7	10	5	1	1	20
8	8	1,0	.	1	3	1	5	5	5	1	2	7	8	3	1	1	65
9	9	1,5	.	1	5	2	5	3	5	1	2	7	8	4	1	1	50
10	11	12,0	.	2	5	4	5	5	5	1	1	10	7	6	1	1	30
11	12	12,0	.	1	5	4	5	3	5	1	1	6	7	6	1	1	15
12	13	12,0	.	1	5	2	5	4	5	1	1	5	10	6	1	1	5
13	14	4,0	.	1	3	1	4	4	1	2	3	5	10	3	1	1	15
14	15	2,5	.	1	5	2	5	2	5	1	9	6	10	3	1	1	15
15	16	1,0	.	4	5	1	5	5	5	1	2	8	7	6	1	1	20
16	17	4,5	.	1	3	4	5	4	5	1	2	5	10	6	1	1	45
17	18	1,5	.	3	5	2	5	5	5	1	9	8	8	6	1	1	10
18	19	2,0	.	3	5	1	5	5	5	1	2	7	8	6	1	1	5
19	20	4,0	.	1	5	1	5	5	3	1	3	2	10	6	1	1	20
20	21	2,5	.	4	5	4	5	5	5	1	4	2	10	2	1	1	55
21	22	5,0	.	1	4	3	4	4	5	1	1	10	8	5	1	1	30
22	23	2,0	.	2	3	2	5	4	5	2	2	5	10	5	1	1	20

Zdroj: vlastní zpracování

Příloha č. 3: Graf rozdělení četností pro proměnnou Tendence v intenzitě učení



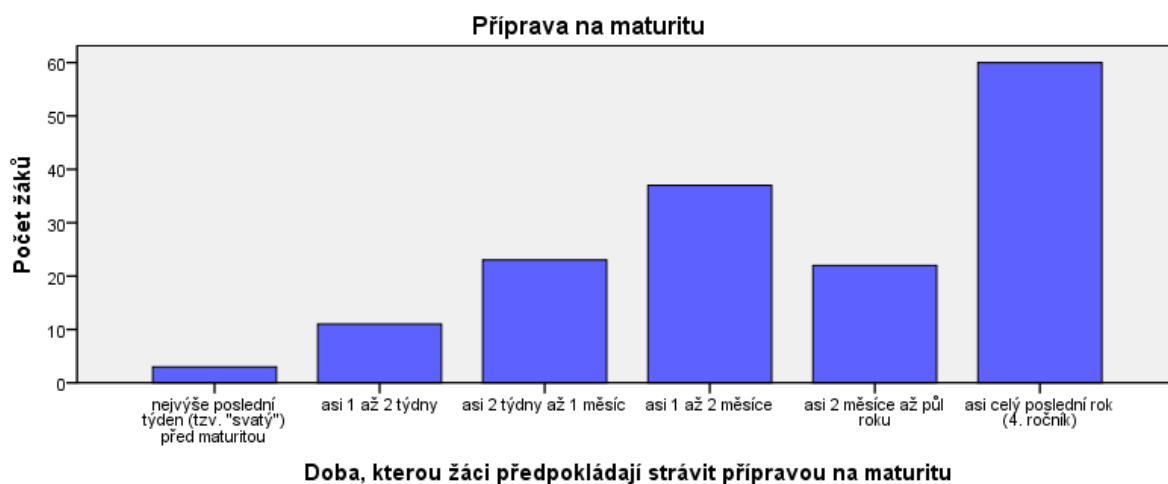
Zdroj: vlastní zpracování

Příloha č. 4: Relativní četnosti využití jednotlivých míst k přípravě do školy (řádková %)

	většinou	často	občas	výjimečně	nikdy
Připravuji se doma	51,9%	16,0%	17,3%	9,6%	5,1%
Připravuji se při cestě do nebo ze školy	5,1%	5,1%	23,7%	12,2%	53,8%
Připravuji se ve škole	26,3%	30,1%	25,0%	8,3%	10,3%
Připravuji se ve studovně nebo v knihovně	0,0%	0,0%	1,9%	2,6%	95,5%
Připravuji se u známých nebo kamarádů	3,8%	0,6%	10,9%	15,4%	69,2%
Připravuji se jinde	3,2%	1,9%	10,3%	7,1%	77,6%

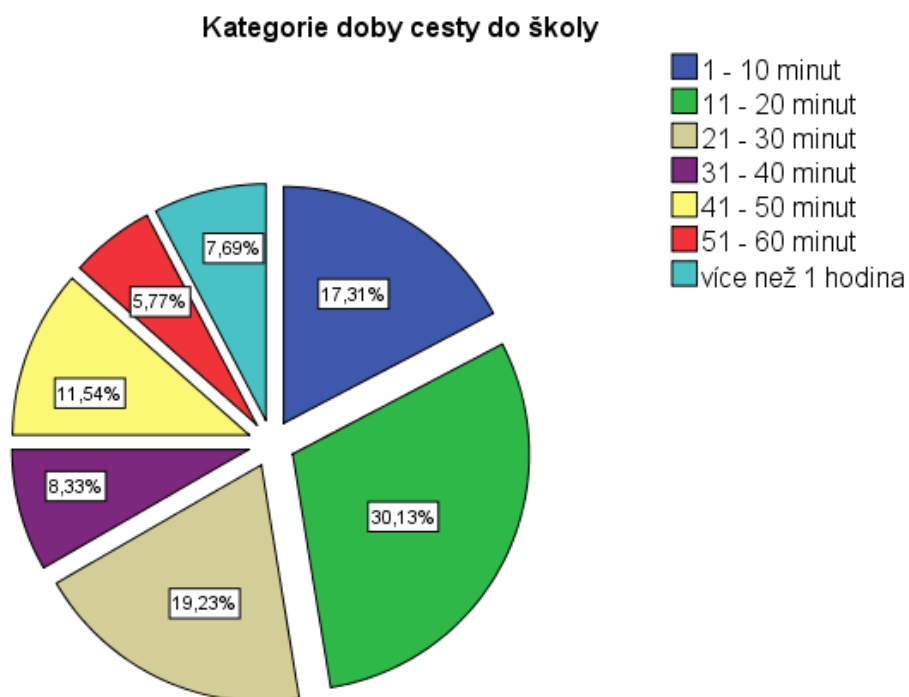
Zdroj: vlastní zpracování

Příloha č. 5: Sloupcový graf znázorňující absolutní četnosti pro proměnnou Příprava na maturitu



Zdroj: vlastní zpracování

Příloha č. 6: Výšečový graf zobrazující dobu, kterou žáci stráví na cestě do školy



Zdroj: vlastní zpracování

Příloha č. 7: Výsledky Chí-kvadrát testu nezávislosti pro proměnné Ročník a Tendence v intenzitě učení

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	2,872 ^a	4	,580	,584
Likelihood Ratio	2,935	4	,569	,581
Fisher's Exact Test	2,940			,574
Linear-by-Linear Association	,499 ^b	1	,480	,534
N of Valid Cases	116			

a. 1 cells (11,1%) have expected count less than 5. The minimum expected count is 4,84.

b. The standardized statistic is ,706.

Zdroj: vlastní zpracování

Příloha č. 8: Spearmanův koeficient pořadové korelace pro vybrané dvojice proměnných (zeleně jsou označeny statisticky významné vztahy na hladině významnosti α 1%, žlutě na hladině významnosti α 5%)

Correlations

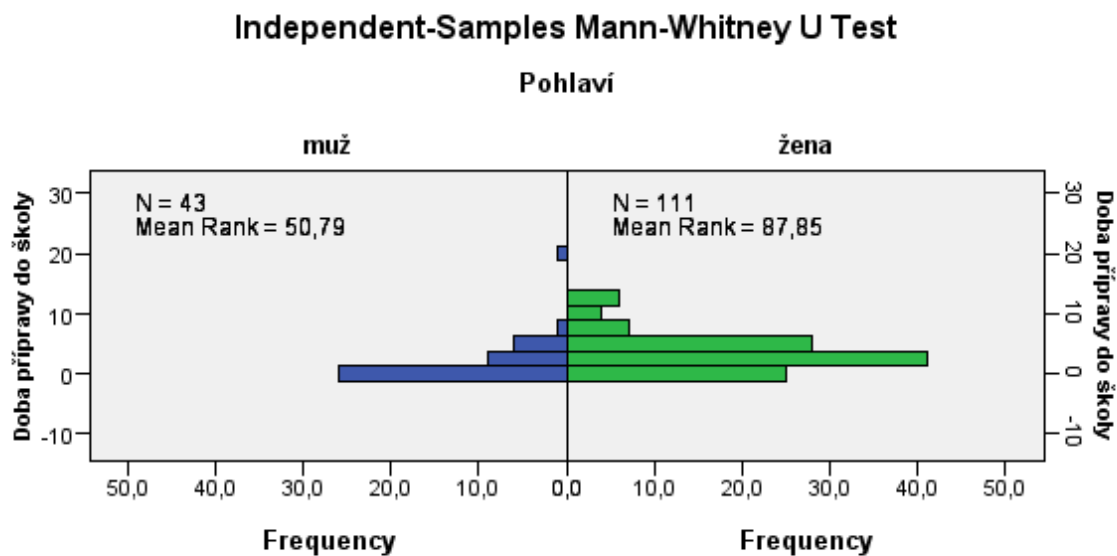
		Doba přípravy do školy	Vztah k oboru studia	Spokojenost se studijními výsledky	Obavy z maturity	Příprava na maturitu
Spearman's rho	Doba přípravy do školy	1,000	-,132	,093	,093	,261**
			,130	,252	,249	,001
		155	133	155	155	155
	Vztah k oboru studia	-,132	1,000	-,185*	-,050	-,293**
		,130		,033	,565	,001
		133	134	134	134	134
	Spokojenost se studijními výsledky	,093	-,185*	1,000	-,189*	,088
		,252	,033		,018	,277
		155	134	156	156	156
	Obavy z maturity	,093	-,050	-,189*	1,000	,181*
		,249	,565	,018		,024
		155	134	156	156	156
Příprava na maturitu	,261**	-,293**	,088	,181*	1,000	
	,001	,001	,277	,024		
	155	134	156	156	156	

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Zdroj: vlastní zpracování

Příloha č. 9: Detailní výsledky a graf k Mannovu-Whitneyho testu pro Dobu přípravy do školy podle pohlaví



Total N	154
Mann-Whitney U	1 238,000
Wilcoxon W	2 184,000
Test Statistic	1 238,000
Standard Error	246,751
Standardized Test Statistic	-4,654
Asymptotic Sig. (2-sided test)	,000

Zdroj: vlastní zpracování

Příloha č. 10: Výsledky Mannova-Whitneyho testu pro proměnnou Příprava na maturitu podle pohlaví

	Příprava na maturitu
Mann-Whitney U	1828,500
Wilcoxon W	2774,500
Z	-2,411
Asymp. Sig. (2-tailed)	,016

a. Grouping Variable: Pohlaví

Zdroj: vlastní zpracování

Příloha č. 11: Výsledky Kruskalova-Wallisova testu pro proměnnou Obavy z maturity podle ročníku studia

	Příprava na maturitu
Chi-Square	15,666
df	3
Asymp. Sig.	,001

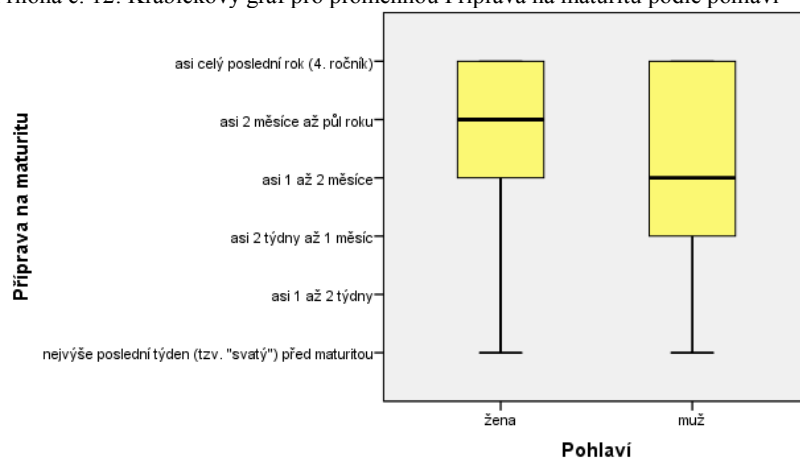
a. Kruskal Wallis Test

b. Grouping Variable: Ročník

studia

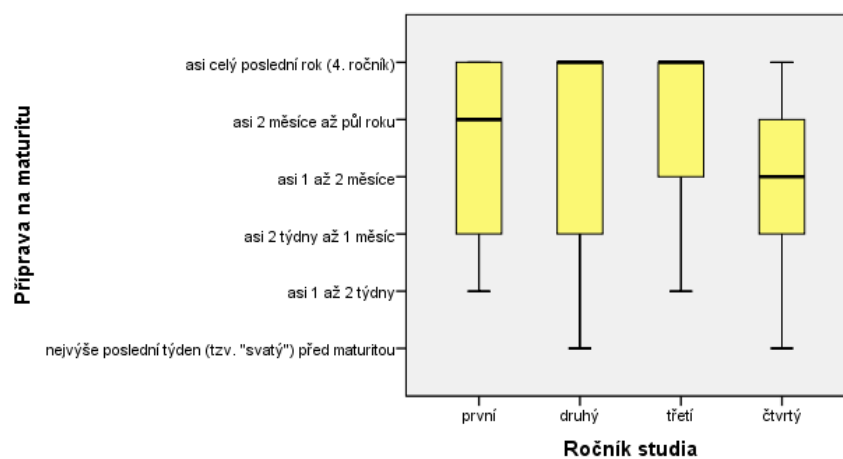
Zdroj: vlastní zpracování

Příloha č. 12: Krabičkový graf pro proměnnou Příprava na maturitu podle pohlaví



Zdroj: vlastní zpracování

Příloha č. 13: Krabičkový graf pro proměnnou Příprava na maturitu podle ročníku studia



Zdroj: vlastní zpracování