

Bakalářský studijní program: **Ekonomika a management**

Studijní obor: **Účetnictví a finanční řízení podniku**

**Způsoby získávání statistických dat.  
Praktické porovnání různých postupů.**

BAKALÁŘSKÁ PRÁCE

Autor: **Marie TRUČKOVÁ**

Vedoucí bakalářské práce: **Mgr. Milan KŘÁPEK**

**Znojmo, 2015**

### ***Čestné prohlášení***

Prohlašuji, že jsem tuto bakalářskou práci na téma: *Způsoby získávání statistických dat, praktické porovnání různých postupů* vypracovala zcela samostatně, pouze za odborného vedení a konzultací s vedoucím bakalářské práce Mgr. Milanem Křápkem, a s použitím uvedené literatury.

Ve Znojmě dne 20. 4. 2015

.....

Marie TRUČKOVÁ

### ***Poděkování***

Velice děkuji vedoucímu bakalářské práce Mgr. Milanovi Křápkovi za cenné odborné rady a připomínky, které mi poskytl v průběhu vzniku této bakalářské práce. Zároveň děkuji ostatním lidem, bez jejichž morální podpory by tato práce nevznikla.



## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Autor	<b>Marie TRUČKOVÁ</b>
Bakalářský studijní program	Ekonomika a management
Obor	Účetnictví a finanční řízení podniku
Název	<b>Způsoby získávání statistických dat. Praktické porovnání různých postupů</b>
Název (v angličtině)	<b>Ways of obtaining statistical data. A practical comparison of different methods</b>

Zásady pro vypracování:

Cíl práce: Cílem práce je analýza a porovnání různých způsobů získávání statistických dat, a to zejména formou dotazování a dále takhle získaná data zhodnotit a následně doporučit nejvýhodnější z těchto způsobů získávání statistických dat, a to vše pro různé skupiny otázek.

Postup práce:

1. Výběr dotazovaných respondentů, tvorba a realizace dotazníku
2. Získání statistických dat formou dotazování (sběr dat)
3. Zpracování a analýza získaných dat
4. Interpretace zjištěných dat - vyhodnocení dotazníku a rozbor výsledků
5. Zhodnocení a následná doporučení

Metody: tvorba dotazníku, korelační analýza, shluková analýza, analýza výsledků

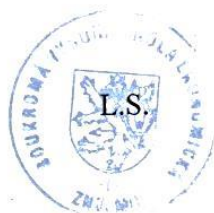
Rozsah práce: 40 - 55

Seznam odborné literatury:

1. BÍLKOVÁ, Diana a kol. *Pravděpodobnost a statistika*. 1. vyd. Plzeň: Vydavatelství a nakladatelství Aleš Čeněk, 2009. 658 s. ISBN 978-80-7380-224-0.
2. KOZEL, Roman a kol. *Moderní metody a techniky marketingového výzkumu*. 1. vyd. Praha: Grada Publishing, 2011. 304 s. ISBN 978-80-247-3527-6.
3. ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření*. 1. vyd. Praha: Professional Publishing, 2007. 212 s. ISBN 978-80-86946-49-8.

Datum zadání bakalářské práce: duben 2014

Termín odevzdání bakalářské práce: duben 2015



  
Marie TRUČKOVÁ  
student

  
Mgr. Milan KRÁPEK  
vedoucí bakalářské práce

  
doc. Ing. Hana BŘEZINOVÁ, CSc.  
garant studijního oboru

  
doc. Ing. Hana BŘEZINOVÁ, CSc.  
rektorka SVŠE Znojmo

## **ABSTRAKT**

Bakalářská práce na téma: *Způsoby získávání statistických dat. Praktické porovnání různých postupů*, se zabývá způsoby získávání statistických dat, a to prostřednictvím dotazování.

Teoretická část popisuje důležité pojmy, které se týkají zejména metod sběru statistických dat s důrazem na metodu dotazování.

V praktické části této bakalářské práce jsou analyzovány výsledky dotazníkového šetření, a to prostřednictvím statistických hypotéz. Na základě výsledků těchto hypotéz jsou následně navržena doporučení, která by mohla pomoci při různých průzkumech trhu.

**Klíčová slova:** dotazník, respondent, statistická data, statistické šetření, statistická hypotéza

## **ABSTRACT**

The bachelor thesis: *Ways of obtaining statistical data. Practical comparison of different methods*, deals with various ways of obtaining statistical data by the means of questioning.

The theoretical part describes essential terms concerning methods of collecting statistical data with focus on the questioning method.

In the practical part of this thesis the results of a questionnaire survey are analyzed by the means of statistical hypotheses. On the basis of these hypotheses recommendations are made which could help at various kinds of market research.

**Key words:** a questionnaire, a respondent, statistical data, statistical research, statistical hypothesis

# OBSAH

1	ÚVOD.....	9
2	CÍL PRÁCE A METODIKA.....	10
3	TEORETICKÁ ČÁST.....	11
3.1	Informace (data).....	11
3.2	Statistický soubor.....	11
3.3	Základní a výběrový soubor.....	11
3.4	Typy proměnných.....	12
3.5	Statistická šetření.....	13
3.5.1	Testování statistických hypotéz.....	14
3.6	Korelační analýza.....	17
3.7	Shluková analýza.....	17
3.8	Metody sběru statistických dat.....	18
3.8.1	Pozorování.....	18
3.8.2	Experiment.....	18
3.8.3	Dotazování.....	19
3.9	Příprava dotazníkového šetření.....	23
3.10	Shrnutí teoretické části.....	29
4	PRAKTICKÁ ČÁST.....	30
4.1	Typy dotazování.....	30
4.2	Dotazník.....	30
4.3	Způsob zpracování získaných dat.....	32
4.4	Osobní otázky.....	33
4.4.1	Předpokládané hypotézy.....	33
4.4.2	Dokazování předpokládaných hypotéz.....	33
4.5	Finanční otázky.....	47

4.5.1	Předpokládané hypotézy .....	47
4.5.2	Dokazování předpokládaných hypotéz .....	47
4.6	Návrhy a doporučení .....	54
4.7	Shrnutí praktické části .....	56
5	ZÁVĚR .....	57
6	SEZNAM POUŽITÉ LITERATURY .....	59
7	SEZNAM OBRÁZKŮ, GRAFŮ A TABULEK .....	61
8	SEZNAM PŘÍLOH .....	63



# 1 ÚVOD

Získávání statistických dat je v dnešní společnosti velice důležité. Způsobem získávání statistických dat se zabývají různé subjekty (jako např. soukromé firmy, organizace, stát, ale i jednotlivci, tedy podnikatelé apod.), které chtějí prozkoumat, jak si jejich firma stojí oproti své konkurenci, nebo chtějí vědět názor a preference obyvatel, či zjistit veřejné mínění občanů ČR atd. Statistické průzkumy jsou v dnešní době nepostradatelné, jelikož na jejich základě se dá např. určit, průměrný měsíční zdanitelný příjem v ČR, spotřeba domácností, míra inflace, struktura obyvatelstva a další.

Nejpoužívanější formou různých průzkumů je právě dotazníkové šetření. Právě dotazník se považuje za základní a nepostradatelný prvek celého statistického šetření. Dotazníkové šetření se tedy využívá zejména při průzkumech trhu. S jakoukoliv formou dotazníku se v běžném životě již pravděpodobně setkal každý člověk.

Dotazování lze provádět online, tedy prostřednictvím internetu, dále osobním dotazováním, telefonickým dotazováním, či písemným dotazováním. Každý typ dotazování má své výhody i nevýhody. Existence více typů dotazování totiž vede k tomu, že si každý tazatel zvolí nejvýhodnější způsob dotazování, případně pro porovnání volí např. i více typů zároveň, a to v závislosti na tom, jaká data a od jaké skupiny respondentů chce získat, dále volí dle preferencí nebo podle své finanční situace.

K tomu, aby byla data získaná z dotazníku reprezentativní, nestačí pouhé sestavení jakéhokoliv dotazníku. Před začátkem sběru dat je nutné, aby si tazatel určil cíl svého výzkumu a tohoto cíle se držel po celou dobu trvání průzkumu. Dále je nutné, aby měl respondent vždy možnost zvolit si, zda chce na danou otázku odpovědět, jelikož dotazník sestavený způsobem nucení respondentů odpovídat, vede pouze k nerepresentativním výsledkům daného průzkumu.

K tomu, aby byly zjištěné výsledky co nejvíce spolehlivé a pravdivé, je zapotřebí znát a respektovat pravidla přípravy, stavby, průběhu a v neposlední řadě také vyhodnocení, tedy správné analyzování nasbíraných dat.

## 2 CÍL PRÁCE A METODIKA

Hlavním cílem bakalářské práce na téma: *Způsoby získávání statistických dat. Praktické porovnání různých postupů*, bylo porovnat internetové a osobní dotazování a tím zjistit, zda je výhodnější online dotazování, tedy dotazování prostřednictvím internetu, nebo naopak, zda je výhodnější dotazování osobní, a to vše pro různé skupiny otázek (osobní, všeobecné a finanční). Z čehož vyplývá, že hlavním cílem práce bylo zjistit, které z těchto typů dotazování je pro dané skupiny otázek reprezentativnější.

Cílem práce bylo tedy zjistit, zda je pro kategorii osobních otázek výhodnější dotazovat se přes internet, či prostřednictvím osobního dotazování a totéž zhodnotit i pro kategorii finančních otázek.

Dalším cílem práce bylo správně analyzovat data získaná z daných druhů dotazování a na základě těchto zjištěných výsledků navrhnout doporučení, kterými by se mohli řídit ostatní tazatelé při provádění různých statistických průzkumů.

Veškeré informace i poznatky, které byly použity v teoretické části, byly získány z odborné literatury, tedy z knižních zdrojů, elektronických zdrojů a zahraničních zdrojů.

Potřebné informace důležité pro zpracování praktické části byly získány z dotazníkového šetření. Toto dotazníkové šetření probíhalo přes internet v čase od 1. 10. 2014 do 31. 12. 2014, a dále probíhalo prostřednictvím osobního dotazování od 3. 11. 2014 do 14. 11. 2014.

Pro zpracování práce byly použity následující metody:

1. Metoda dotazování, tedy metoda tvorby dotazníku
2. Zpracování údajů prostřednictvím kontingenčních tabulek
3. Statistické hypotézy
4. Shluková analýza
5. Korelační analýza
6. Grafická vyhodnocení
7. Analýza výsledků

## 3 TEORETICKÁ ČÁST

Nejprve je nutné vymezit některé pojmy, které jsou pro tuto práci velmi důležité.

### 3.1 Informace (data)

Nejdůležitějším pojmem pro každá statistická šetření jsou statistická data, z nichž se následně vytvářejí informace. Informace můžeme charakterizovat jako uspořádaná data, která můžeme koupit buď v podobě studií, nebo je většinou najdeme v různých tabulkách a grafech. Každá informace by měla splňovat určitá kritéria, a to především úplnost, pravdivost, srozumitelnost, objektivnost, aktuálnost, spolehlivost atd. (Kozel, 2011, s. 50-51)

### 3.2 Statistický soubor

Statistickým souborem nazýváme množinu, která má určité shodné vlastnosti. Příkladem může být množina osob, zvířat či věcí. (Bílková a kol., 2009, s. 14)

*Prvky statistického souboru nazýváme **statistické jednotky**. Počet statistických jednotek statistického souboru se nazývá **rozsah** statistického souboru.* (Bílková a kol., 2009, s. 14)

### 3.3 Základní a výběrový soubor

#### Základní soubor

Základnímu souboru se někdy též říká populace. Základní soubor je většinou velmi rozsáhlý, a proto není v praxi prakticky vůbec uskutečnitelné, aby byly prozkoumány vlastnosti všech prvků v daném souboru. (Bílková a kol., 2009, s. 14)

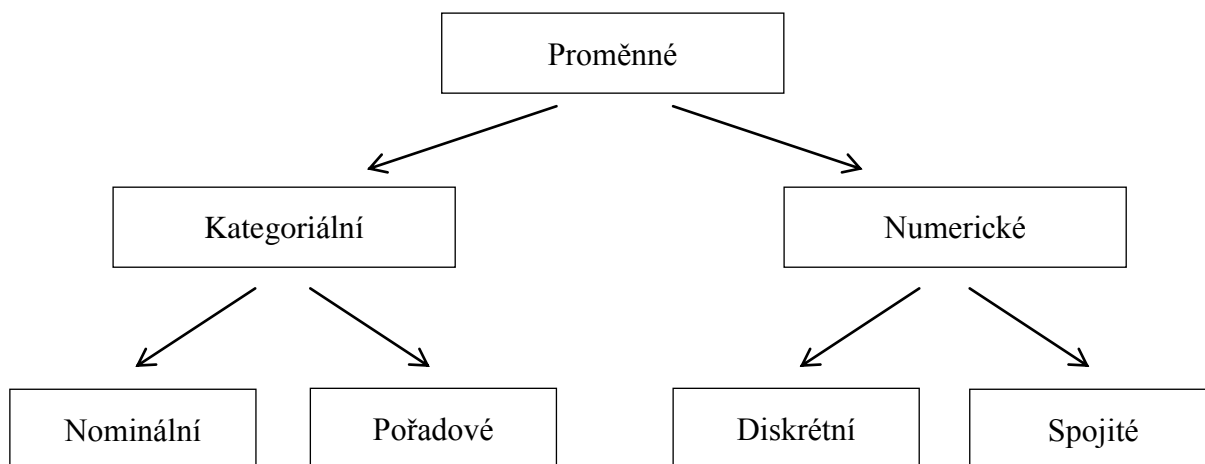
#### Výběrový soubor

Výběrovým souborem rozumíme soubor prvků, které vybereme ze základního souboru a následně tyto statistické jednotky zkoumáme. Výběrový soubor by měl co možná nejlépe reprezentovat základní soubor. (Bílková a kol., 2009, s. 14)

### 3.4 Typy proměnných

Proměnné rozlišujeme podle toho, zda je vyjadřujeme v číslech nebo slovně.

**Graf 1 - Rozdělení statistických proměnných podle typu**



**Zdroj:** Práce autora s použitím informací z (Křápek, 2013, s. 11)

#### **Kategoriální proměnné**

Někdy tyto proměnné nazýváme také slovní. Příkladem může být výsledek zkoušky u studentů, kde si zvolíme tyto obměny: výborně, velmi dobře, dobře, nevyhověl. (Bílková a kol., 2009, s. 14)

- **Nominální kategoriální proměnné**

U těchto obměn nelze určit pořadí daných znaků. (Bílková a kol., 2009, s. 14)

- **Pořadové kategoriální proměnné**

U těchto proměnných známe pořadí, což znamená, že je můžeme seřadit od nejnižší obměny po nejvyšší. Dané výsledky můžeme mezi sebou následně srovnávat. (Bílková a kol., 2009, s. 14)

#### **Numerické proměnné**

Tyto proměnné se rovněž nazývají číselné. Příkladem numerické proměnné je počet členů v domácnosti. Numerické proměnné dělíme na diskrétní proměnné a spojité proměnné. (Krejšlová, 2008, s. 9)

- **Diskrétní numerické proměnné**

*Na daném intervalu nabývají pouze izolovaných číselných hodnot. Nejčastěji se jedná o přirozená, nebo celá čísla. (Křápek, 2013, s. 10)*

- **Spojité numerické proměnné**

Na určitém intervalu nabývají reálné číselné hodnoty. (Křápek, 2013, s. 10)

### **3.5 Statistická šetření**

Jde především o vlastnosti statistických souborů, které jsou vyjádřeny určitými proměnnými a nabývají daných statistických hodnot. Právě tyto hodnoty (data) jsou velice důležité pro statistická zkoumání. Tato data rozlišujeme na dva typy, a to na primární data a sekundární data. Primárními daty rozumíme data, která získáváme sami tím, že je sami zjišťujeme. Sekundární data vždy odněkud převezmeme. Jedná se tedy o převzaté hodnoty proměnných. Hlavním účelem statistického šetření je získání statistických dat, která jsou potřebná k prozkoumání určitých hromadných jevů<sup>1</sup> (Bílková a kol, 2009, s. 17)

Výsledkem statistického šetření bývá velké množství číselných údajů, které nejsou většinou přehledné a musí se setřídit, což znamená, že se jedná o rozdělení jednotek souboru do takových skupin, aby co možná nejlépe vznikly charakteristické vlastnosti zkoumaných jevů. (Hindls, Hronová, Seger, 2004, s. 17-18)

#### **Vyčerpávající statistická šetření**

Zjišťujeme hodnoty daných proměnných u všech statistických jednotek základního souboru. (Bílková a kol, 2009, s. 18)

- **Úplné statistické šetření**

Jde především o takové statistické šetření, při kterém byly prozkoumány všechny statistické jednotky daného souboru. (Bílková a kol, 2009, s. 18)

- **Neúplné statistické šetření**

Při neúplném statistickém šetření dochází k tomu, že nebyly prozkoumány veškeré statistické jednotky, které měly být prozkoumány. Pokud je rozsah těchto neprošetřených jednotek

---

<sup>1</sup> Jevy, které se vyskytují mnohokrát a můžou se opakovat (Křápek, 2013, s. 7)

vzhledem ke zkoumanému souboru velmi malý, dá se v tomto případě tato neúplnost tolerovat. (Bílková a kol, 2009, s. 18)

### **Nevyčerpávající statistická šetření**

U těchto šetření se již předem počítá s tím, že se hodnoty daných proměnných budou zjišťovat pouze z výběrového statistického souboru. Nevyčerpávající statistické šetření se rovněž dělí na úplné a neúplné. (Bílková a kol, 2009, s. 18)

- **Nerepresentativní nevyčerpávající statistická šetření**

Příkladem takových šetření bývají obvykle ankety. Výsledky z daného zkoumání neodpovídají základnímu souboru, tzn., že nerepresentují. (Bílková a kol, 2009, s. 19)

- **Reprezentativní nevyčerpávající statistická šetření**

Výsledky se mohou zobecnit a mohou odpovídat vlastnostem celého základního souboru. (Bílková a kol, 2009, s. 19). Cílem této bakalářské práce je zjistit, jakým způsobem provést reprezentativní šetření.

### **3.5.1 Testování statistických hypotéz**

Testovací postup platný pro každý z dále popsaných testů je následující:

1. Výběr vhodného testu
2. Zformulování nulové a alternativní hypotézy
3. Výpočet hodnoty testového kritéria
4. Určení kritického oboru podle tvaru alternativní hypotézy  $H_1$
5. Vyvození závěru (Bílková a kol, 2009, s. 441)

#### **Hypotéza o střední hodnotě**

Hypotéza o střední hodnotě se využívá tehdy, pokud je cílem určit tvrzení, že střední hodnota je rovna určitému číslu, přičemž  $H_0: \mu = \mu_0$ . Pro tuto hypotézu platí: oboustranná alternativní hypotéza:  $H_1: \mu \neq \mu_0$ , pro kterou je kritický obor  $U \in \left(-\infty; u_{\frac{\alpha}{2}}\right) \cup \left(u_{1-\frac{\alpha}{2}}; \infty\right)$ , levostranná alternativní hypotéza:  $H_1: \mu < \mu_0$ , pro kterou kritický obor náleží  $(-\infty; u_{\alpha})$  a pravostranná alternativní hypotéza:  $H_1: \mu > \mu_0$ , pro kterou kritický obor náleží  $(u_{1-\alpha}; \infty)$ .

Testové kritérium:  $U = \frac{\bar{x} - \mu_0}{\hat{s}_x}$ . (Bílková a kol, 2009, s. 440-445)

Tato hypotéza bude v práci použita například pro porovnání průměrného věku respondentů, či pro porovnání, zda zjištěný průměrný měsíční zdanitelný příjem odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu v ČR.

### Hypotéza o relativní četnosti

Pokud je potřeba určit platnost tvrzení, že je relativní četnost určitého znaku A rovna číslu  $\pi_0$ , potom tedy platí:  $H_0: \pi = \pi_0$ , oboustranná alternativní hypotéza  $H_1: \pi \neq \pi_0$ , pro kterou je kritický obor  $U \in (-\infty; u_{\frac{\alpha}{2}}) \cup (u_{1-\frac{\alpha}{2}}; \infty)$ , levostranná alternativní hypotéza  $H_1: \pi < \pi_0$ , pro kterou je kritický obor  $(-\infty; u_{\alpha})$  a pravostranná alternativní hypotéza  $H_1: \pi > \pi_0$ , pro kterou je kritický obor  $(u_{1-\alpha}; \infty)$ . Testové kritérium:  $U = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$ .

(Bílková a kol, 2009, s. 445-446).

Tato hypotéza bude v práci využita zejména k ověření, zda se zjištěné procentní zastoupení respondentů rovná skutečnému procentnímu zastoupení obyvatel ČR, a to hlavně u pohlaví, nezaměstnanosti, či u procentního zastoupení studentů.

### Hypotéza o rovnosti středních hodnot

Tato hypotéza se využívá, pokud existují dva náhodné výběry ze dvou různých souborů. Cílem je určit, zda se střední hodnoty těchto základních souborů rovnají, tedy:  $H_0: \mu_1 = \mu_2$ . Oboustranná alternativní hypotéza  $H_1: \mu_1 \neq \mu_2$ , pro kterou platí kritický obor hodnot  $U \in (-\infty; u_{\frac{\alpha}{2}}) \cup (u_{1-\frac{\alpha}{2}}; \infty)$ , levostranná alternativní hypotéza  $H_1: \mu_1 < \mu_2$ , pro kterou je kritický obor  $(-\infty; u_{\alpha})$  a pravostranná alternativní hypotéza  $H_1: \mu_1 > \mu_2$ , pro kterou je kritický obor  $(u_{1-\alpha}; \infty)$ . Testové kritérium pro  $n_1 > 30$  a  $n_2 > 30$ :  $U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1'^2}{n_1} + \frac{s_2'^2}{n_2}}}$ .

(Bílková a kol, 2009, s. 446-452)

Tato hypotéza bude v práci použita zejména k ověření, zda se například průměrný věk respondentů jednoho rozdělení, rovná průměrnému věku respondentů druhého rozdělení.

## Hypotéza o rovnosti relativních četností

Předpokládá se náhodný výběr velkého rozsahu, kde:  $H_0: \pi_1 = \pi_2$ . Oboustranná alternativní hypotéza tedy je  $H_1: \pi_1 \neq \pi_2$ , pro kterou je kritický obor  $U \in \langle u_{1-\alpha}; \infty \rangle$ , levostranná alternativní hypotéza je  $H_1: \pi_1 < \pi_2$ , pro kterou je kritický obor  $U \in (-\infty; -u_{1-\alpha})$  a pravostranná alternativní hypotéza  $H_1: \pi_1 > \pi_2$ , kde kritický obor  $U \in$  všem hodnotám

větším než  $u_{1-\alpha}$  a menším než  $-u_{1-\alpha}$ . Testové kritérium:  $U = \frac{P_1 - P_2}{\sqrt{\bar{P}(1-\bar{P})}} * \sqrt{\frac{n_1 * n_2}{n_1 + n_2}}$ , kde:

$$P_1 = \frac{M_1}{n_1}, P_2 = \frac{M_2}{n_2}, \bar{P} = \frac{M_1 + M_2}{n_1 + n_2} = \frac{n_1 * P_1 + n_2 * P_2}{n_1 + n_2}$$

(Bílková a kol, 2009, s. 452-453)

Tato hypotéza bude v práci použita například ke zjištění, jaký typ dotazování zvolit, aby respondenti odpovídali častěji, tedy aby se docílilo reprezentativních výsledků daného šetření.

## Test dobré shody ( $\chi^2$ )

Slouží k ověření, zda získaná data mohou pocházet z náhodné veličiny s daným rozdělením.

$H_0$ : Rozdělení základního souboru je ...,  $H_1$ : Neplatí hypotéza  $H_0$ . Testové kritérium:

$$G = \sum_{i=1}^k \frac{(n_i - n\pi_{0,i})^2}{n\pi_{0,i}}, \text{ kde } \pi_{0,i} \text{ udává teoretické pravděpodobnosti obměn, } n_i \text{ četnosti}$$

Pokud přesně známe rozdělení pak má statistika  $G$  rozdělení chí-kvadrát s  $k-1$ .

Kritický obor:  $G \in \langle \chi_{1-\alpha}^2; \infty \rangle$ . Test je průkazný pokud pro každé  $i$  platí  $n\pi_{0,i} > 5$ . (Bílková a kol, 2009, s. 457-459)

Tento test bude v práci využit zejména ke zjištění, zda určité rozdělení respondentů odpovídá skutečnému rozdělení obyvatel v ČR. Dále bude test použit například k určení, zda je věkové rozložení respondentů stejné pro dané typy dotazování, či zda jsou respondenti ze stejných vzdělanostních skupin.



### 3.6 Korelační analýza

Modely, v nichž se předpokládá, že  $n$  napozorovaných dvojic, trojic, čtveřic, atd. jsou hodnoty vícerozměrné náhodné veličiny, se nazývají **korelační modely** a analýza dat pomocí takových modelů se nazývá **korelační analýza**. (Hindls, Hronová, Novák, 2000, s. 76) Korelační analýza se tedy zabývá mírou závislosti náhodných dat a popisuje vztahy mezi danými veličinami. (Bílková a kol, 2009, s. 520-603). To, zda jsou proměnné  $X$  a  $Y$  závislé ověřujeme  **$\chi^2$ -testem nezávislosti**, kde  $H_0$ : proměnné jsou nezávislé,  $H_1$ : proměnné jsou závislé. Testové kritérium:  $\chi^2 = \sum_{i=1}^k \sum_{j=1}^l \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$ , kritický interval je:  $(\chi^2_{1-\alpha}; \infty)$ , stupně volnosti daného rozdělení se určí jako:  $\nu = (k - 1)(l - 1)$ . K určení míry závislosti slouží **Pearsonův koeficient**:  $C = \sqrt{\frac{\chi^2}{n + \chi^2}}$ . Tento koeficient nabývá hodnot  $(0; 1)$ . Pokud je nulový, pak jsou proměnné nezávislé a naopak pokud se výsledná hodnota blíží jedné, pak je mezi proměnnými silná závislost. (Bílková a kol, 2009, s. 510 - 513).

Korelační analýza bude v práci využita zejména k ověření, zda existuje závislost mezi věkem a pohlavím respondentů. Pokud bude zjištěno, že existuje závislost mezi věkem a pohlavím, tak bude dále pomocí Pearsonova koeficientu určeno, jak silná závislost mezi těmito proměnnými existuje.

### 3.7 Shluková analýza

Shluk je skupina objektů, které jsou si navzájem podobné a rozdílné od objektu do této skupiny nepatřících. Tato analýza se zabývá metodami, pomocí kterých sdružuje data s podobnými vlastnostmi do shluku. Shluková analýza je nástroj datové analýzy, který řídí různé objekty do shluků tak, že podobnost dvou objektů náležících do jedné skupiny je maximální, zatímco podobnost s objekty mimo tento shluk je minimální. Tzv. shlukováním je možné najít určité vztahy mezi danými objekty bez dalšího vysvětlování, či interpretace. Právě shluková analýza má své využití v různých průzkumech trhu – používá se pro zpracování dat s více proměnnými, které jsou získány průzkumy. (Kučera, Shluková analýza, © 2008)

Shluková analýza bude v práci řešena pouze graficky. Konkrétně bude graficky znázorňovat rozdíl mezi respondenty, kteří odpovídali přes internet, a to dle jejich věku a pohlaví.

## 3.8 Metody sběru statistických dat

Hlavními metodami sběru primárních dat je především dotazování, dále pozorování a v neposlední řadě také experiment. Nejvíce využívanou metodou je v dnešní době dotazování. Tato kapitola se bude především zabývat metodou dotazování a pouze okrajově nastíní metodu pozorování a experimentu.

### 3.8.1 Pozorování

Pozorování provádí většinou odborný pracovník, který je řádně vyškolen. Pozorovatel musí být nezávislý a také objektivní. Pozorování se uskutečňuje bez aktivní účasti pozorovaného člověka. (Foret, Stávková, 2003, s. 41)

Jedním ze způsobů rozlišení může být např. **pozorování nezúčastněným pozorovatelem**, což znamená, že se daný subjekt děje neúčastní, ale pouze tento děj sleduje. Pozorovatel totiž nechce nijak děj ovlivnit, či narušit. Dalším typem je **pozorování zúčastněné**, při kterém se pozorovatel děje osobně účastní – je jeho součástí. Tento typ se využívá zejména v různých pracovních skupinách, kam se pozorovatel začlení a skupinu zkoumá. (Pauknerová a kol., 2012, s. 55)

Dále dělíme pozorování podle stupně standardizace, a to na pozorování **standardizované** – zde jsou přesně stanoveny jevy i kategorie, do kterých bude pozorovatel dané skutečnosti zaznamenávat. Dále rozlišujeme pozorování **nestandardizované** – zde je určen pouze cíl pozorování. Pozorovatel má tzv. volnou ruku v rozhodování. (Foret, Stávková, 2003, s. 47)

### 3.8.2 Experiment

*Experimentální metody sledují vliv jednoho jevu (nezávisle proměnná) na druhý (závisle proměnná), a to v nově vytvořené situaci.* (Foret, Stávková, 2003, s. 48)

Experiment se od pozorování liší zejména tím, že při něm dochází ke kontrolovanému zásahu do průběhu sledovaného jevu. Rozlišujeme dva základní typy experimentů – experiment **přirozený**, který je podobný pozorování. Příkladem tohoto experimentu může být např. cílevědomá změna podmínek ovlivňování pracovníků – změnou pravidel odměňování, či změnou stylu vedení a tak dále. A dále také experiment **laboratorní**, který se vyznačuje tím, že má vědomou účast zkoumaných osob a převážně probíhá v uměle vytvořených laboratořích. Samotná příprava experimentu vyžaduje splnění určitých kroků, jako je

například: stanovení cílů, výběr experimentálních osob, poskytnutí základní instrukce osobám zúčastněným a vyhodnocení výsledků experimentu. (Pauknerová a kol., 2012, s. 57-58)

V praktickém životě se experimentů využívá velice málo, jelikož jsou tyto výzkumy velmi časově náročné a také velmi nákladné. (Simová, 2005, s. 65)

### **3.8.3 Dotazování**

*Podstatou dotazování je pokládání otázek dotazovaným (respondentům), které se uskutečňuje pomocí nástrojů (dotazníku nebo záznamových archů) a vhodně zvoleného kontaktu s nositelem informací – dotazovaným. Dotazování může probíhat přímou nebo nepřímou komunikací s respondentem.* (Simová, 2005, s. 65)

Dotazování se liší podle stupně standardizace:

#### **1. Standardizovaný rozhovor**

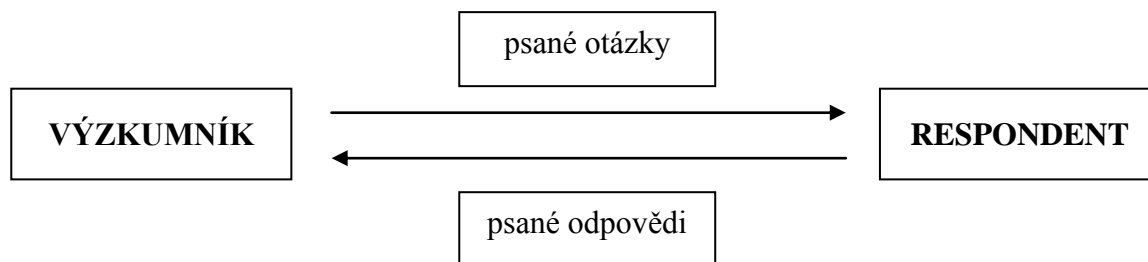
Obsahuje přesně formulované otázky, které jsou předkládány každému respondentu ve stejném znění i ve stejném pořadí. Tazatel předkládá dotazovanému otázky dle předem připraveného protokolu. (Pauknerová a kol., 2012, s. 60)

#### **2. Nestandardizovaný rozhovor**

Je oproti předchozímu typu rozhovoru mnohem volnější a pružnější. Většinou se nedodrží žádné dané schéma a tazatel má k dispozici pouze jakousi hrubou osnovu rozhovoru, což ale předpokládá speciální výcvik ve vedení takového rozhovoru. (Pauknerová a kol., 2012, s. 61)

Na následujících obrázcích je jasně a přehledně zobrazen rozdíl mezi přímým písemným dotazováním a nepřímým písemným dotazováním, čili zprostředkovaným dotazováním.

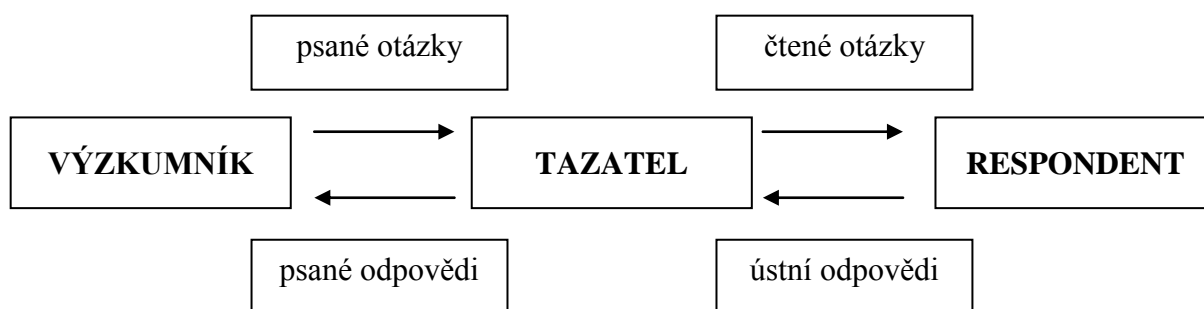
**Obrázek 1 - Přímé písemné dotazování**



**Zdroj:** Práce autora s použitím informací z (Foret, Stávková, 2003, s. 32)

Z výše uvedeného obrázku vyplývá, že při přímém písemném dotazování pokládá výzkumník otázky respondentovi a respondent naopak poskytuje svoje odpovědi zpět přímo tazateli. Jedná se tedy o dotazování, které probíhá pouze mezi výzkumníkem a respondentem.

**Obrázek 2 - Zprostředkované dotazování**



**Zdroj:** Práce autora s použitím informací z (Foret, Stávková, 2003, s. 3)

Z výše uvedeného obrázku vyplývá, že při zprostředkovaném dotazování vstupuje mezi výzkumníka a respondenta tazatel, který vystupuje v roli tzv. zprostředkované osoby.

### 3.8.3.1 Typy dotazování

#### Osobní dotazování (ústní dotazování)

Osobní dotazování patří k základním typům dotazování. Dochází zde k osobnímu styku a přímé komunikaci s danými respondenty (*face to face*). Hlavní výhodou této metody je právě osobní kontakt, při kterém vzniká **přímá zpětná vazba** mezi tazatelem a respondentem. Tato přímá zpětná vazba je velice důležitá, jelikož slouží např. k upřesnění některých otázek, které nejsou respondentům jasné, dále také může tazatel určitým způsobem motivovat respondenty k odpovědím. (Kozel, 2011, s. 175)

U této metody je velmi důležité, jakým dojmem působí tazatel na respondenty. Dotazování totiž raději reagují na někoho, kdo je milý, příjemný a má vhodné společenské vystupování, než na osobu, která má nevhodné společenské vystupování a je nepříjemná. Tento způsob dotazování má tzv. nejvyšší návratnost odpovědí, ale zároveň je také velice nákladný, ať už finančně, či časově. (Kozel, 2011, s. 175)

Při osobním dotazování se dříve využívalo především papírových dotazníků. Tento způsob se nazývá **P+P**<sup>2</sup>. Toto tradiční vyplňování papírových dotazníků je v dnešní době spíše nahrazováno tzv. metodou **CAPI**<sup>3</sup>, která je neustále více a více využívána. CAPI představuje dotazování za pomoci multimediálních přenosných počítačů, při kterém se využívá elektronický dotazník a zároveň dochází k osobnímu kontaktu a odpovědi se zaznamenávají přímo do počítače. *Prakticky se jedná o nejkvalitnější metodu realizace kvantitativních výzkumů.* (Kozel, 2011, s. 175). Pro potřeby této bakalářské práce bude v praktické části použit způsob P+P, tedy data získaná od respondentů budou zaznamenávány papírově.

Osobní dotazování probíhá v domácnostech, na veřejných místech, tzn. na ulicích nebo v obchodních střediscích či prodejnách, kde se vyskytuje velké množství lidí. Podle toho, kde probíhá dotazování, se zvolí délka a složitost dotazníku. Například dlouhý dotazník se využívá tam, kde se předpokládá, že na odpovědi budou mít respondenti více času, tzn. v domácnostech. Naopak kratší a jednodušší dotazníky volíme při dotazování na ulicích či nákupních střediscích. (Krejslová, 2008, s. 12)

Mezi nejvýznamnější výhody tedy patří:

- Vysoká návratnost vyplněných dotazníků
- Dané pořadí otázek, které umožňuje určitou gradaci
- Možnost předkládání určitých vzorků atd. (Boučková a kol. 2003, s. 64)

Mezi nejvýznamnější nevýhody patří:

- Vyžaduje tazatele
- Relativně drahé (Boučková a kol. 2003, s. 64)

---

<sup>2</sup> paper + pencil

<sup>3</sup> Computer Assisted Personal Interviewing

## **Telefonické dotazování**

Telefonické dotazování je velice využívanou metodou sběru dat, avšak v poslední době je stále více nahrazováno tzv. online výzkumy. Nejvíce se tento způsob dotazování využívá při zpětném volání zákazníkům, dále při dotazování ve firmách a při všech dalších typech volání z databází. Hlavní výhodou je především rychlost získávání potřebných údajů. Další výhodou je, že pokud respondenta nezastihneme na telefonu, hovor můžeme kdykoliv zopakovat jindy. Pokud bychom srovnali telefonické dotazování a osobní dotazování, dospějeme k závěru, že při telefonickém dotazování není potřeba vynaložit tak vysoké náklady. Mezi nevýhody můžeme zařadit vysoké nároky na soustředění dotazovaných – doba hovoru by neměla přesáhnout deset minut. Otázky, které tazatel klade respondentům, by měly být jednoduché, s minimálním množstvím variant odpovědí. (Kozel, 2011, s. 176)

## **Online dotazování**

Při online dotazování se zjišťují informace prostřednictvím e-mailů nebo také prostřednictvím webových stránek. Jedná se o nejmladší způsob zjišťování určitých dat. Mezi největší výhodu můžeme řadit zejména nižší finanční a časovou náročnost. Rozesílání e-mailů je mnohem rychlejší, než rozesílání klasických dopisů. Zpracování odpovědí je jednodušší, jelikož jsou všechny odpovědi v elektronické formě. Jednou z největších nevýhod je fakt, že existuje mnoho lidí, kteří nevěří moderním technikám, a tudíž se bojí vyplňovat jakékoliv osobní údaje online. (Kozel, 2011, s. 177) Online dotazování existuje několik druhů, které rozlišujeme podle toho, jakým způsobem přistupujeme k výběru respondentů.

V dnešní době existují např. tyto přístupy:

- Nahodilé výběry na webových stránkách firem, či výzkumných agentur
- Nahodilé výběry na sociálních sítích
- a další, ... (Kozel, 2011, s. 178)

## **Písemné dotazování**

Písemnému dotazování se také říká **dotazování poštou**. Jedná se o nejméně rozšířený typ dotazování. Dotazník bývá zpravidla přikládán k nějakému výrobku. Výhodou jsou vcelku nízké náklady a respondent má dostatek času na zodpovězení otázek. Velkou nevýhodou je nízká návratnost, která má za následek narušení reprezentativnosti výzkumu a tím se prodlužuje doba celkového statistického šetření. (Kozel, 2011, s. 178) Mezi další nevýhodu

lze zařadit fakt, že tento způsob dotazování vyžaduje seznam adres dotazovaných (Boučková a kol., 2003, s. 64)

*Výběr vhodné techniky šetření závisí zejména na povaze zjišťovaných informací a na jejich potřebném rozsahu, na charakteru respondentů, na časových a finančních možnostech. Každá z uvedených technik má své výhody a nevýhody, které je třeba předem pečlivě zvážit. V praktickém využití se jednotlivé techniky navzájem kombinují, obecně nelze o žádné říci, že je nejvhodnější.* (Foret, Stávková, 2003, s. 47)

V praktické části této bakalářské práce bude použito online, tedy internetové dotazování a dále také osobní dotazování. Dle stupně standardizace bude dotazování probíhat jako standardizovaný rozhovor.

### **3.9 Příprava dotazníkového šetření**

Před začátkem každého sběru dat z dotazníkových šetření je důležité stanovení následujících kroků:

1. Stanovení cíle
2. Určení způsobu dotazování
3. Určení cílové skupiny respondentů
4. Konstrukce jednotlivých otázek
5. Konstrukce celého dotazníku

(Kozel, 2011, s. 187)

Při dotazníkovém průzkumu jde především o sběr dat, kde se tazatel snaží popsat, porovnat, nebo vysvětlit individuální a společenské znalosti, pocity, hodnoty, preference, či chování dotazovaných. Existuje mnoho metod, jak získat informace o lidech, ale průzkum je pouze jeden. (Fink, 2013, s. 2-3)

***Dotazník** je jedním z nejběžnějších nástrojů pro sběr dat pro různé typy průzkumů. Skládá se ze série otázek, jejichž cílem je získat názory a fakta od respondentů.* (Dotazník-online, © 2007)

Nejdůležitější je, aby dotazník nebyl nudný a aby byl vizuálně zajímavý.

Důležité je zaměřit se především na:

- Srozumitelnost
- Přehlednost
- Jednoduchost
- Jazykovou správnost
- Grafickou úpravu

(Dotazník-online, © 2007)

### **Délka a rozsah vyplňování dotazníku**

Obvykle se uvádí, že by měl dotazník obsahovat 40 – 50 otázek a délka vyplňování by se měla pohybovat kolem 20 minut. Z praktického hlediska je však toto pravidlo většinou nereálné, jelikož na takové množství otázek je ochotný odpovídat pouze velmi dobře motivovaný respondent. Více než 20 otázek většinou odrazuje a délka vyplňování by neměla překročit 10 minut. (Dotazník-online, © 2007)

### **Výběr respondentů**

Vždy je velmi nepravděpodobné, že bychom mohli dotazovat všechny členy základního souboru, tedy populace. Dotazování celého základního souboru by bylo drahé, namáhavé a mnohdy také zcela nemožné, z tohoto důvodu tedy volíme tzv. výběrový soubor (určitý vzorek respondentů). Nejdříve je nutné rozhodnout, **koho** budeme dotazovat, dále **způsob**, kterým budeme dotazovat a nakonec **kolik** respondentů budeme dotazovat. Čím více respondentů bude dotazováno, tím spolehlivější budou výsledky. Na začátku je tedy velmi důležité určit cílovou skupinu dotazovaných a také zvážit obsah tématu a jeho vztah k cílové skupině dotazovaných respondentů. (Kozel, 2011, s. 192-194)

### **Formulace dotazů**

Otázky můžeme dělit do dvou základních skupin:

1. Otázky, které se týkají názorů a chování respondentů
2. Otázky, které se využívají za účelem získávání jiných údajů – např. demografických

Velice důležité je, aby byly otázky formulovány srozumitelně a také jednoznačně. U většiny dotazů se respondentům nabízejí různé možnosti odpovědí. Tyto otázky nazýváme uzavřené.



Uzavřené otázky dále dělíme na alternativní – nabízejí dvě varianty odpovědi a dále selektivní – nabízejí více, než dvě varianty odpovědi. (Řezanková, 2007, s. 11)

Vedle výše zmiňovaných otázek uzavřených existují samozřejmě také otázky otevřené. Škála hodnot se v tomto případě vytváří až potom, co respondent odpoví. Existuje i možnost tyto dva typy kombinovat a tím vznikají polootevřené či polouzavřené otázky, které spočívají v tom, že si respondent buď vybere z daných možností, nebo má možnost uvést svoji odlišnou variantu. (Řezanková, 2007, s. 12)

U tzv. otevřených odpovědí jde zejména o to, že respondenti mohou odpovídat na danou otázku vlastními slovy. Tyto typy otázek nemají omezený rozsah odpovědí, jejich hranice jsou omezeny pouze respondenty samými, kteří tvoří vlastní odpovědi. Co se týče uzavřených otázek, jde o tzv. nejpobulárnější způsob dotazování. (Church, Waclawski, 2007, s. 65-67)

**Tabulka 1 - Výhody a nevýhody otevřených a uzavřených otázek**

	VÝHODY	NEVÝHODY
OTEVŘENÉ OTÁZKY	Respondent není nijak omezen ve svých odpovědích	Odpovědi se tvoří delší dobu
	Každá odpověď je originální	Je potřebná následná kategorizace odpovědí
	Umožňují získání odpovědí, které tvůrce dotazníku nemusely napadnout	Složitá interpretace získaných dat
UZAVŘENÉ OTÁZKY	Jednoduché a rychlé vyplnění odpovědi	Omezují možnost svobodného názoru respondentů
	Větší ochota respondentů spolupracovat	Odpovědi nemusí vystihovat názory respondenta
	Variety odpovědí pomáhají respondentovi lépe formulovat jeho názory	Mohou vnučovat odpověď

**Zdroj:** Práce autora s použitím informací z (Kozel, 2011, s. 214-215)

## **Konstrukce otázek**

Neexistuje přesný návod, jaké otázky se mají a jaké nemají vyskytovat v dotazníku. Nejdůležitější je, aby byla každá otázka srozumitelná. Tazatel by měl při přípravě otázek do dotazníku dodržovat určitá pravidla, jako jsou například:

1. Ptát se přímo
2. Ptát se jednoduše
3. Užívat jednovýznamová slova
4. Ptát se konkrétně
5. Užívat krátké otázky
6. Vyloučit nepříjemné otázky
7. Vyloučit negativní otázky
8. Umožnit neposkytnout odpověď
9. Dodržovat slušnost
10. Snižovat citlivost otázek

(Kozel, 2011, s. 202-204)

## **Škály odpovědí**

Odpovědi respondentů jsou hodnotami, které jsou z určité škály. (Řezanková, 2007, s. 14) Škálování tedy slouží k měření a vyjádření postojů daných respondentů. Při škálování je důležité zvolit správný počet kategorií. V praxi je doporučeno zadávat 5 – 7 kategorií. Nižší počet kategorií se nedoporučuje, a to z důvodu omezení při dalších analýzách. Vyšší počet kategorií se také nedoporučuje, jelikož se tím může zkomplikovat výběr odpovědi respondentů. Každá škála by měla nabízet stejný počet pozitivních i negativních odpovědí. Pokud budeme zařazovat i neutrální odpovědi, tak je zde velká pravděpodobnost, že respondenti budou volit právě tyto odpovědi. (Krejslová, 2008, s. 20-21)

Odpovědím na uzavřené otázky jsou přiřazeny slovní, či číselné kódy. Příkladem statistických znaků charakterizovaných pomocí kategorií mohou být:

- Národnost (česká, slovenská, polská, ...)
- Úroveň vzdělání (základní, středoškolské, vysokoškolské, ...)
- Počet dětí (0, 1, 2, 3, ...)

Výše uvedené příklady znázorňují určité vztahy mezi kategoriemi. Některé kategorie lze uspořádat a jiné nikoliv. V prvním případě je nelze uspořádat, ve druhém případě je můžeme uspořádat a ve třetím případě je lze uspořádat a navíc u těchto hodnot můžeme vypočítat jejich rozdíl. (Řezanková, 2007, s. 19)

Odpovědi na otevřené otázky jsou zachyceny v původní podobě. V případě polootevřených nebo polouzavřených otázek se na jednu otázku pohlíží jako na dvě. První otázka je tedy uzavřená a druhá je otevřená. (Řezanková, 2007, s. 20)

#### **U odpovědí by mělo být zajištěno:**

1. Validita – co nejvěrnější zachycení skutečnosti
2. Realiabilita – spolehlivost
3. Jednoznačnost – výstižné a jednoduché věty
4. Srozumitelnost – volit jazyk určený cílové skupině respondentů
5. Stručnost – krátké a výstižné věty
6. Nepoužívat sugestivní otázky – otázky, které svou formulací nutí respondenta k určité odpovědi
7. Vyhnout se tzv. Haló-efektu – řada příbuzných otázek za sebou, odpověď se přenáší z jedné otázky do ostatních

(Dotazník-online, © 2007), (Řezanková, 2007, s. 12)

#### **Zdroje chyb při aplikaci metod šetření**

- Chyby v důsledku odmítnutí dotazování
- Stupeň pochopení otázky
- Schopnost a ochota respondenta odpovídat pravdivě
- Vyjadřovací schopnost respondenta
- Chyby při zpracování odpovědí

(Zbořil, 1994, s. 47-50)

Veškeré výše zmíněné chyby přispívají k nereprezentativním výsledkům statistických šetření. Je tedy důležité, aby byly tyto chyby co nejvíce eliminovány. Chybám v důsledku odmítnutí dotazování lze zabránit tím, že tazatel zvolí pro konkrétní otázky vhodný způsob dotazování. Například pro otázky finančního charakteru zvolí raději online dotazování, jelikož je zde vyšší pravděpodobnost, že respondent bude ochoten odpovídat. Aby respondent odpovídal

pravdivě, je zde nutné zvolit dle pohlaví respondenta, zda bude průzkum provádět muž, či žena. Chybám při zpracování odpovědí lze zabránit tím, že tazatel zvolí vhodné metody, které použije pro zpracování konkrétních dat.

### **Interpretace a prezentace výsledků výzkumu**

Při interpretaci výsledků je především důležité dodržovat objektivitu. Výsledky výzkumu lze prezentovat písemně nebo ústně, přičemž písemná prezentace je považována za základní a nejvíce využívanou. (Zbořil, 1994, s. 69-71). Výsledky výzkumu získané z dotazníkového šetření se obvykle interpretují pomocí statistických metod, a to zejména za použití různých statistických testů.

### **Výhody a nevýhody dotazníku**

**Tabulka 2 - Výhody a nevýhody dotazníku**

<b>VÝHODY (+)</b>	<b>NEVÝHODY (-)</b>
Velice levná metoda průzkumu	Obtížnost získávání respondentů
Jednoduché zpracování a vyhodnocení	Respondent může vyplňovat nepravdivé informace
Jedna z nejméně dotěrných metod průzkumu	Obtížnější komunikace s respondenty

**Zdroj:** Práce autora s použitím informací z: (Dotazník-online, © 2007)

Z výše uvedené tabulky vyplývá, jaké výhody a nevýhody má dotazníkové šetření.

### 3.10 Shrnutí teoretické části

Na začátku teoretické části jsou uvedeny a také vysvětleny důležité pojmy, jako např.: co to jsou informace, statistický soubor, jaký je rozdíl mezi základním a výběrovým souborem, jaká je definice korelační a shlukové analýzy a další. Dále je zde uvedeno, jaké existují typy proměnných, tedy fakt, že typy proměnných se dělí na kategoriální a numerické proměnné. A následně se kategoriální proměnné dělí na nominální a pořadové a numerické proměnné se dělí na diskrétní a spojité.

Pro statistická šetření jsou důležitá tzv. statistická data, která se dělí do dvou kategorií, a to na primární, která zjišťujeme sami a dále na sekundární, která odněkud převezmeme. Výsledkem statistického šetření je množství údajů, které se musí nejprve seřadit a následně zpracovat. Příprava dotazníkového šetření zahrnuje stanovení cíle, určení způsobu dotazování, určení cílové skupiny respondentů, konstrukci jednotlivých otázek a konstrukci celého dotazníku.

Důležitou podkapitolou je podkapitola věnující se metodám sběru statistických dat. Statistická data můžeme získávat pozorováním, experimentem a v neposlední řadě také dotazováním. Dotazování uskutečňujeme osobně, telefonicky, online či písemně.

Další samostatná podkapitola se věnuje problematice dotazníku. Kdy dotazník lze charakterizovat jako nástroj pro sběr dat pro různé typy průzkumů. Dotazník by měl být srozumitelný a přehledný. Doporučená délka dotazníku by neměla překročit 20 otázek a rozsah vyplňování by neměl být delší, jak 10 minut. Na začátku se zvolí způsob dotazování a kolik respondentů budeme dotazovat. Dotazník by měl obsahovat jak otevřené, tak i uzavřené otázky.

Odpovědi respondentů jsou hodnotami, které jsou z určité škály, kdy škálování slouží k měření a vyjádření postojů daných respondentů. Výsledky výzkumu se prezentují zejména písemně.

Další důležitou podkapitolou je samozřejmě také podkapitola věnující se testování statistických hypotéz. V této podkapitole je popsán obecný postup, podle kterého se statistické hypotézy testují a dále jsou zde uvedeny veškeré důležité vzorce, které se pro toto testování používají.

## **4 PRAKTICKÁ ČÁST**

### **4.1 Typy dotazování**

#### **Online dotazování**

Nejprve byl sestaven dotazník na internetové stránce: [www.survio.com](http://www.survio.com) a následně byl odkaz na dotazník rozeslán prostřednictvím sociální sítě facebook a prostřednictvím emailu známým, kteří odkaz rozesílali dál svým známým a tak vznikl určitý řetězec respondentů. Online dotazování, tedy přes internet probíhalo od 1. 10. 2014 do 31. 12. 2014.

#### **Osobní dotazování**

Osobní dotazování probíhalo od 3. 11. 2014 do 14. 11. 2014. Osobní dotazování se uskutečnilo ve dvou městech, a to v Třebíči – kraj Vysočina a ve Znojmě – Jihomoravský kraj. V Třebíči probíhalo osobní dotazování v centru města, tedy na Karlově náměstí a ve Znojmě probíhalo dotazování na Masarykově náměstí.

Pro účely osobního dotazování bylo vytisknuto 100 dotazníků. Informace z osobního dotazování byly tedy zaznamenávány papírovou formou.

### **4.2 Dotazník**

#### **Počet odpovědí z dotazníkového šetření**

Celkem bylo od respondentů získáno 200 odpovědí. Z toho 100 odpovědí bylo získáno prostřednictvím internetu a dalších 100 odpovědí bylo získáno osobním dotazováním. Osobním dotazováním bylo získáno 50 odpovědí v Třebíči a dalších 50 odpovědí ve Znojmě. Tazatelem ve Znojmě byl muž a v Třebíči prováděla dotazování žena. Daný počet respondentů byl určen záměrně pro lepší porovnání osobního a online dotazování.

#### **Výběr respondentů**

Před začátkem sběru dat byl stanoven minimální věk respondentů. Odpovídat mohli všichni respondenti starší 15 let, přičemž maximální věk nebyl omezen.

Dotazník obsahoval 18 otázek. Tyto otázky by se daly rozdělit na 3 podkategorie: osobní otázky, všeobecné otázky a finanční otázky. V dotazníku byly použity otevřené i uzavřené otázky. Otázky, které dotazník obsahoval, se dají rozdělit podle typu otázek na kategoriální i numerické, přičemž převládají právě kategoriální proměnné.

Dotazník se skládal z následujících otázek:

#### Osobní otázky

1. Pohlaví
2. Věk
3. Místo narození
4. Místo bydliště (název města/vesnice)
5. Rodinný stav
6. Počet dětí
7. Typ bydlení
8. Nejvyšší dosažené vzdělání
9. Zaměstnání

#### Všeobecné otázky

10. Nejvíce vyhledávaný typ dovolené/rekreace
11. Nejvíce upřednostňovaný druh dopravy
12. Jak často cestujete do zahraničí (za jakýmkoliv účelem)
13. Ohodnoťte známkou jako ve škole úroveň školství v ČR

#### Finanční otázky

14. Za co nejvíce utrácíte peníze
15. Kolik Kč činí Váš měsíční zdanitelný příjem (měsíční hrubá mzda)
16. Kolik Kč měsíčně činí v průměru Vaše osobní výdaje
17. Využíváte v osobním životě úvěrů/půjček
18. Aktuální výše úvěrů/půjček

Pro analýzu získaných dat byly nejvíce vypovídající právě otázky osobní a finanční, a naopak otázky všeobecné sloužily zejména jako tzv. „vsuvka“. Tato vsuvka měla za úkol oddělit otázky osobního a finančního charakteru a zamezit tak tomu, aby dotazník nevyvolal v respondentech pocit, že dané otázky jsou postaveny pouze na sběru velmi osobních dat.

Pokud by totiž dotazník obsahoval pouze velmi osobní otázky, mohlo by to vést k neochotě respondentů odpovídat.

### **4.3 Způsob zpracování získaných dat**

Aby mohla být získaná data vyhodnocena, byly vytvořeny různé kontingenční tabulky, a to ke každé z daných otázek. Poté byla data uspořádaná v daných kontingenčních tabulkách vyhodnocena. Před samostatným zpracováním získaných dat byly u podkategorie osobních a finančních otázek vytvořeny hypotézy, u kterých by se z nějakého důvodu dalo předpokládat nebo očekávat, že platí. U podkategorie všeobecných otázek nebyly zpracovány žádné hypotézy, jelikož jak již bylo zmíněno, tato podkategorie není pro danou analýzu získaných dat důležitá, ani přínosná.

Pro dokázání či zamítnutí některých předpokládaných, tedy alternativních hypotéz byl použit statistický software QC-Expert od firmy TriloByte Statistical Software, s.r.o.

Veškerá skutečná data, byla získána ze stránek Českého statistického úřadu, a to konkrétně z různých ročenek, grafů, ze sčítání obyvatel a podobně. Data odpovídají 4. čtvrtletí roku 2014. Takto získaná data byla použita za účelem porovnání se získanými daty prostřednictvím dotazování.

Práce se bude dále zaměřovat na podkategorii osobních a finančních otázek, kde pro všechny zkoumané hypotézy byla zvolena hladina významnosti  $\alpha = 5 \% = 0,05$ . Pro dokazování veškerých hypotéz byl zvolen postup odpovídající postupu uvedenému v podkapitole *3.5.1 Testování statistických hypotéz*: Nejprve byl zvolen vhodný test, dále byla zformulována nulová a alternativní hypotéza, poté bylo vypočítáno testové kritérium, následně byl určen kritický obor dle tvaru alternativní hypotézy a nakonec byl vyvozen závěr.



## 4.4 Osobní otázky

### 4.4.1 Předpokládané hypotézy

1. Při dotazování prostřednictvím internetu je průměrný věk dotazovaných respondentů pod průměrem věku obyvatel ČR
2. Při dotazování přes internet je průměrný věk respondentů nižší, než při osobním dotazování
3. Pohlaví dotazujícího má vliv na zkreslení vybíraného vzorku, a to ve prospěch osob druhého pohlaví
4. Internetové dotazování respektuje skutečné rozdělení osob dle pohlaví ve společnosti
5. Na internetové dotazování, v porovnání s osobním dotazováním odpovídají častěji studenti a nezaměstnaní, a to ve větší míře, než by odpovídalo vhodně zvolenému vzorku
6. Při porovnání internetového a osobního dotazování se předpokládá, že respondenti budou z rozdílných vzdělanostních skupin
7. Dotazování přes internet v porovnání s osobním dotazováním, lépe odpovídá skutečnému rozdělení společnosti dle počtu obyvatel v obci (dle velikosti obce)

### 4.4.2 Dokazování předpokládaných hypotéz

#### 4.4.2.1 Hypotéza 1: Při dotazování prostřednictvím internetu je průměrný věk dotazovaných respondentů pod průměrem věku obyvatel ČR

**Tabulka 3 - Porovnání průměrného věku z internetového dotazování se skutečností**

<b>Průměrný věk z internetového dotazování</b>	31,93
<b>Skutečný průměrný věk v ČR</b>	41,50

**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Z výše uvedeného vyplývá, že průměrný věk z internetového dotazování je 31,93 a skutečný průměrný věk v ČR činí 41,50.

Byla provedena **hypotéza o střední hodnotě**, kde:  $H_0: \mu = 41,50$ ,  $H_1: \mu < 41,50$ , testové kritérium:  $t = 6,43$  a kritický obor:  $(1,6604; \infty)$

Z tohoto výsledku vyplývá, že testové kritérium náleží do kritického oboru. Na hladině významnosti 5 % se podařilo potvrdit hypotézu  $H_1$  – průměrný věk respondentů, který byl získaný prostřednictvím internetového dotazování je pod průměrem věku obyvatel ČR.

Pro podrobnější porovnání zastoupení respondentů dle věkových kategorií přes internet se skutečným zastoupením obyvatel ČR je nutné provést další testování. Pro porovnání jednotlivých věkových skupin byl použit statistický **test dobré shody**.

**Tabulka 4 - Procentní zastoupení obyvatel dle věkových kategorií**

VĚKOVÉ KATEGORIE		INTERNET	SKUTEČNOST
	15 – 19	20,00 %	6,75 %
	20 – 24	25,00 %	7,90 %
	25 – 29	11,00 %	8,43 %
	30 – 34	8 %	10,15 %
	35 – 39	10 %	10,39 %
	40 – 44	5 %	8,21 %
	45 – 49	5 %	8,11 %
	50 – 54	6 %	7,77 %
	55 – 59	3 %	8,79 %
	60 – 64	1 %	8,73 %
	65 – 69	5 %	6,58 %
	70 – 74	0 %	4,54 %
75 - 79	1 %	3,64 %	

**Zdroj:** Práce autora s použitím dat získaných z Českého statistického úřadu

Ve výše uvedené tabulce je znázorněn rozdíl mezi procentním zastoupením obyvatel získaným z online dotazování a skutečným procentním zastoupením obyvatel v ČR, a to vždy dle určité věkové kategorie.

Byl tedy proveden test dobré shody, kde:  $H_0$ : Věkové rozdělení respondentů přes internet odpovídá skutečnému rozdělení obyvatel ČR,  $H_1$ : Neplatí hypotéza  $H_0$ , testové kritérium: chí-kvadrát ( $\chi^2$ ) = 84,55, kritický obor: (21,0607;  $\infty$ )

Z toho plyne, že testové kritérium náleží do kritického oboru a shoda poměrů se tedy zamítá. Na hladině významnosti 5 % byla zamítnuta hypotéza  $H_0$ . Bylo zamítnuto, že by věkové rozdělení respondentů přes internet odpovídalo skutečnému rozdělení obyvatel ČR. Na základě výše provedeného testu lze tvrdit, že věkové rozdělení respondentů přes internet neodpovídá skutečnému rozdělení obyvatel ČR.

#### 4.4.2.2 Hypotéza 2: Při dotazování přes internet je průměrný věk respondentů nižší, než při osobním dotazování

Bylo provedeno porovnání dvou výběrů, kde jeden výběr obsahoval věkové údaje získané přes internet a druhý výběr obsahoval věkové údaje získané osobním dotazováním.

Bylo zjištěno, že průměrný věk respondentů, který byl získán internetovým dotazováním, odpovídá hodnotě **31,93** a průměrný věk respondentů, který byl získán prostřednictvím internetového dotazování, činí **31,69**.

Pro ověření této hypotézy byl použit **test o rovnosti středních hodnot**, kde:  $\mu_1$  = průměrný věk respondentů přes internet,  $\mu_2$  = průměrný věk respondentů získaný osobním dotazováním,  $H_0: \mu_1 = \mu_2$ ,  $H_1: \mu_1 < \mu_2$ , testové kritérium:  $U = 0,20$ , kritický obor:  $(-\infty; -1,6449)$

Testové kritérium nenáleží do kritického oboru. Byl tedy proveden test shody středních hodnot, při kterém bylo zjištěno, že průměry i daná rozdělení jsou shodná. Tento výsledek tedy vyvrací hypotézu, která říká, že při dotazování prostřednictvím internetu je průměrný věk respondentů nižší, než při osobním dotazování. Z toho vyplývá, že při internetovém dotazování je průměrný věk dotazovaných respondentů stejný, jako při osobním dotazování.

Pro podrobnější zkoumání je nutno provést další testování.

**Tabulka 5 - Procentní zastoupení respondentů dle věkových kategorií**

VĚKOVÉ KATEGORIE	INTERNET	OSOBNĚ
	15 – 19	20,00 %
20 – 24	25,00 %	20,00 %
25 – 29	11,00 %	15,00 %
30 – 34	8,00 %	12,00 %
35 – 39	10,00 %	11,00 %
40 – 44	5,00 %	2,00 %
45 – 49	5,00 %	3,00 %
50 – 54	6,00 %	5,00 %
55 – 59	3,00 %	6,00 %
60 – 64	1,00 %	4,00 %
65 – 69	5,00 %	2,00 %
70 – 74	0,00 %	0,00 %
75 - 79	1,00 %	0,00 %

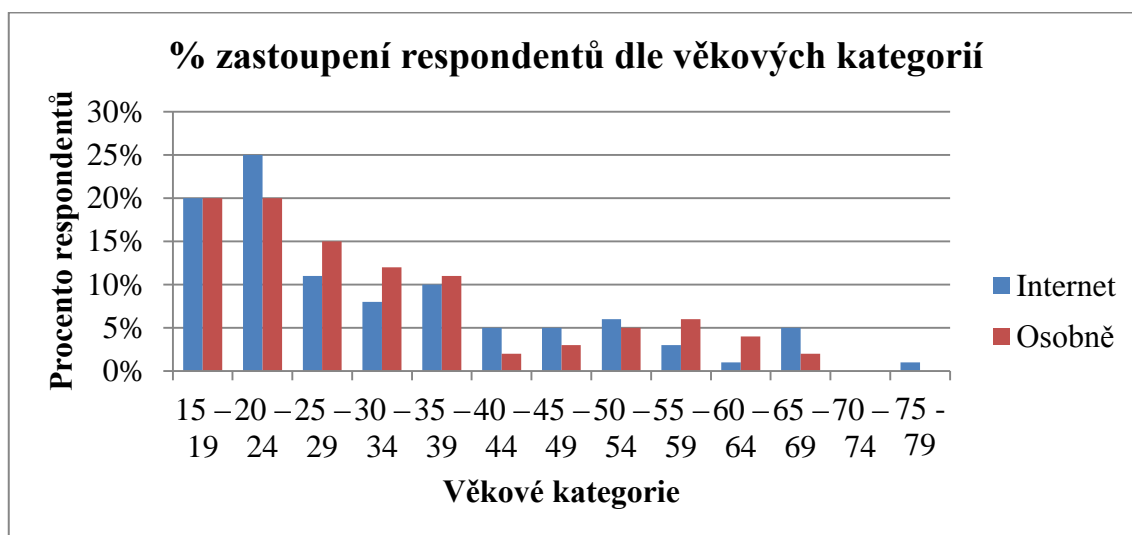
**Zdroj:** Práce autora

Ve výše uvedené tabulce je znázorněn rozdíl mezi procentním zastoupením obyvatel získaným z online dotazování a procentním zastoupením obyvatel získaným prostřednictvím osobního dotazování, a to vždy dle určité věkové kategorie.

Byl proveden **test dobré shody**, kde:  $H_0$ : Věkové rozložení respondentů je stejné pro internetové i osobní dotazování,  $H_1$ : Neplatí hypotéza  $H_0$ , Testové kritérium: chí-kvadrát ( $\chi^2$ ) = 22,12, kritický obor:  $(19,6751; \infty)$

Bylo zjištěno, že testové kritérium náleží do kritického oboru a tudíž se shoda poměrů zamítá. Na hladině významnosti 5 % bylo zamítnuto, že by věkové rozložení respondentů z internetového a osobního dotazování bylo shodné. Z toho vyplývá, že věkové rozložení respondentů z internetového dotazování je jiné, než věkové rozložení respondentů získané osobním dotazováním.

**Graf 2- Porovnání procentuálního zastoupení respondentů dle věkových kategorií**

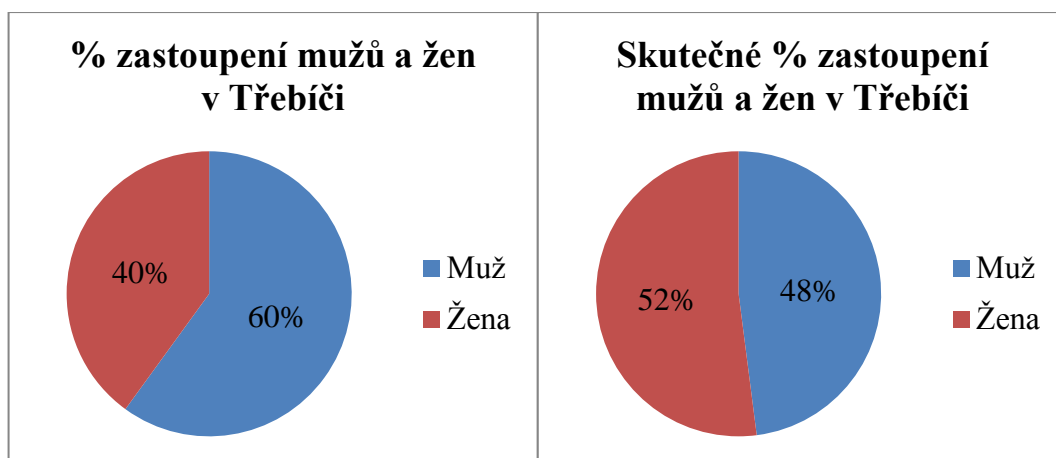


**Zdroj:** Práce autora

Výše uvedený graf zobrazuje rozdíl v procentuálním zastoupení respondentů dle věkových kategorií, a to mezi osobním a internetovým dotazováním. Z grafu je patrné, že věkové rozložení respondentů z internetového dotazování je jiné, než věkové rozložení respondentů získané osobním dotazováním.

#### 4.4.2.3 Hypotéza 3: Pohlaví dotazujícího má vliv na zkreslení vybíraného vzorku, a to ve prospěch osob druhého pohlaví

Graf 3 - Porovnání skutečného a zjištěného zastoupení mužů a žen v Třebíči



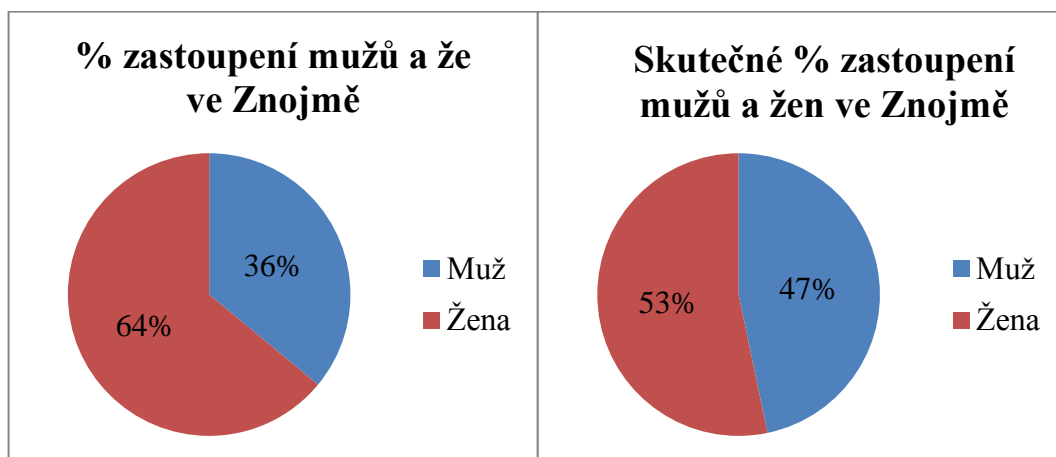
**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Z výše uvedeného grafu vyplývá, že při osobním dotazování, které probíhalo v Třebíči, odpovědělo 60 % mužů a 40 % žen, přičemž ve skutečnosti tvoří muži 48 % obyvatel Třebíče a ženy 52 % všech obyvatel Třebíče.

Pro porovnání byla provedena **hypotéza o relativní četnosti**, tedy:  $H_0: \pi = 0,48$ ,  $H_1: \pi > 0,48$ , testové kritérium:  $U = 1,70$ , kritický obor:  $(1,6449; \infty)$

Bylo zjištěno, že testové kritérium spadá do kritického oboru. Na hladině významnosti 5 % bylo tedy potvrzeno, že při osobním dotazování v Třebíči mělo pohlaví tazatele vliv na zkreslení vybíraného vzorku ve prospěch osob druhého pohlaví.

**Graf 4 - Porovnání skutečného a zjištěného zastoupení mužů a žen ve Znojmě**



**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Z výše vyobrazeného vyplývá, že při osobním dotazování, které probíhalo ve Znojmě, odpovědělo 36 % mužů a 64 % žen, přičemž ve skutečnosti ženy tvoří ve Znojmě 53 % a muži 47 % obyvatel.

Pro porovnání byla provedena **hypotéza o relativní četnosti**, tedy:  $H_0: \pi = 0,47$ ,  $H_1: \pi < 0,47$ , testové kritérium:  $U = -1,56$ , kritický obor:  $(-\infty; -1,6449)$

Testové kritérium spadá do kritického oboru, a tudíž lze na hladině významnosti 5 % potvrdit, že při osobním dotazování ve Znojmě mělo pohlaví tazatele vliv na zkreslení vybíraného vzorku ve prospěch osob druhého pohlaví.

V Třebíči se dotazovala žena, z čehož podle výše graficky vyjádřených výsledků vyplývá, že žena se záměrně více dotazuje opačného pohlaví, tedy mužů a tím dochází ke zkreslení vybíraného vzorku dotazovaných respondentů. Obdobně je to i u tazatele muže, který se dotazoval ve Znojmě a z výše uvedených výsledků vyplývá, že muž se mnohem raději dotazuje žen, tedy opačného pohlaví a tím samozřejmě taktéž dochází ke zkreslení daného vzorku dotazovaných respondentů.

Z daných výsledků možno předpokládat, že tazatel muž i tazatelka žena si vybírají toho, koho se budou dotazovat stejným způsobem. Pro potvrzení tohoto předpokladu byl proveden **test o rovnosti relativních četností**, kde:  $H_0: \pi_1 = \pi_2$ ,  $H_1: \pi_1 > \pi_2$ , testové kritérium:  $U = 2,40$ , kritický obor:  $(1,6449; \infty)$ , kde  $\pi_1$  odpovídá hodnotě 60 %, tedy získanému procentnímu

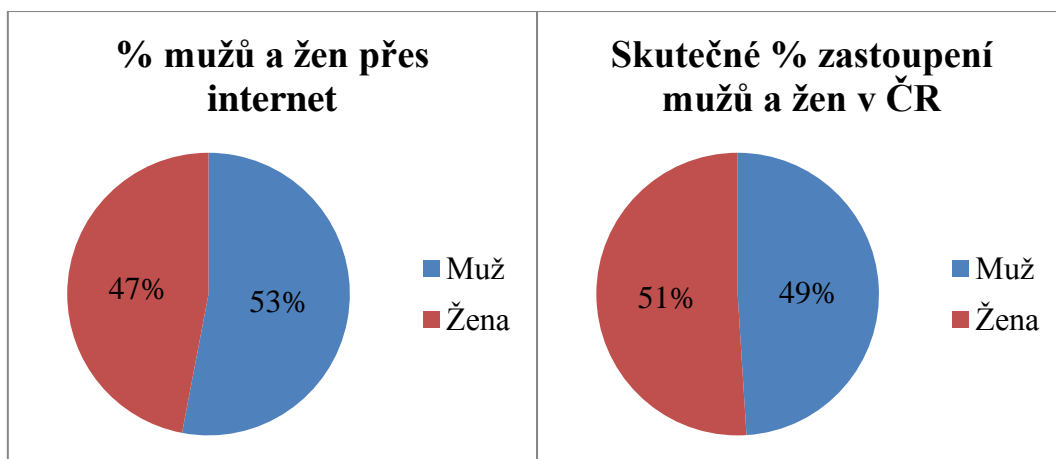
zastoupení mužů v Třebíči a  $\pi_2$  odpovídá hodnotě 36 %, čili zjištěnému procentnímu zastoupení mužů ve Znojmě.

Testové kritérium náleží do kritického oboru. Na 5% hladině významnosti se podařilo zamítnout nulovou hypotézu a prokázalo se, že pokud je tazatelkou žena, ptá se více opačného pohlaví, tedy mužů. Z čehož vyplývá, že pokud se dotazuje muž, ptá se převážně žen a méně mužů.

Dále byl proveden další **test o rovnosti relativních četností**, kde:  $\pi_1 = \pi_2$ ,  $\pi_1 \neq \pi_2$ , testové kritérium:  $U = -0,32$ , kritický obor:  $(1,6449; \infty)$ , kde  $\pi_1$  odpovídá hodnotě 60 %, tedy získanému procentnímu zastoupení mužů v Třebíči a  $\pi_2$  odpovídá hodnotě 64 %, tedy zjištěnému procentnímu zastoupení žen ve Znojmě. Testové kritérium nenáleží do kritického oboru. Na 5% hladině významnosti se podařilo prokázat, že tazatel muž i tazatelka žena si vybírají potenciální respondenty stejným způsobem. Z čehož vyplývá, že míra preferencí u obou pohlaví je totožná.

#### 4.4.2.4 Hypotéza 4: Internetové dotazování respektuje skutečné rozdělení osob dle pohlaví ve společnosti

**Graf 5 – Porovnání procentního zastoupení mužů a žen z internetového dotazování se skutečností**



**Zdroj:** Práce autora s použitím informací získaných z Českého statistického úřadu

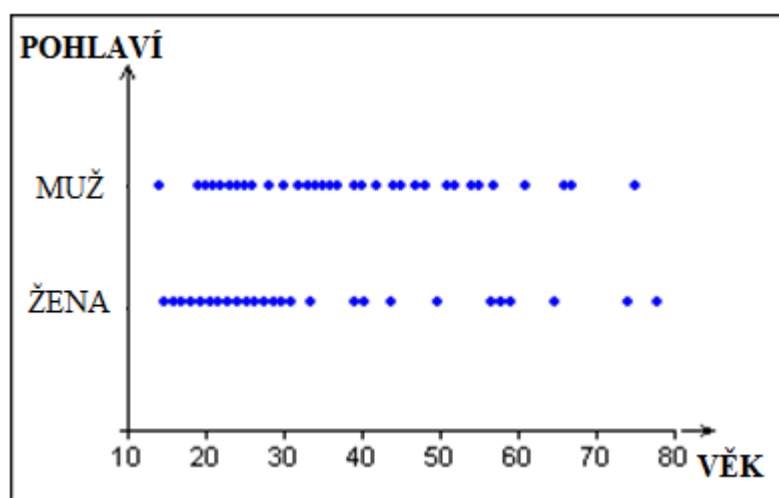
Z výše uvedeného grafu vyplývá, že přes internet odpovědělo 53 % mužů a 47 % žen, přičemž ve skutečnosti se ČR skládá ze 49 % mužů a 51 % žen.

Byla provedena **hypotéza o relativní četnosti**, kde:  $H_0: \pi = 0,49$ ,  $H_1: \pi \neq 0,49$ , testové kritérium:  $U = 0,80$ , kritický obor:  $(-\infty; -1,96) \cup (1,96; \infty)$

Testové kritérium nenáleží do kritického oboru. Na 5% hladině významnosti byla zamítnuta alternativní hypotéza  $H_1$ . Z výsledků tedy vyplývá, že internetové dotazování respektuje skutečné rozdělení osob dle pohlaví ve společnosti.

Dále byla provedena **shluková analýza**, která má za cíl zobrazit rozdíl mezi respondenty, kteří odpovídali prostřednictvím internetu, a to zejména dle jejich věku a pohlaví.

**Graf 6 – Respondenti, kteří odpovídali prostřednictvím internetu – znázornění dle věku a pohlaví**



**Zdroj:** Práce autora za použití programu QC-Expert

Z grafu vyplývá, že ženy odpovídají přes internet nejvíce ve věku 15 – 30 let a dále je z grafu patrné, že muži oproti ženám odpovídají více rovnoměrně. Nejvíce odpovídají muži ve věkové kategorii 19 – 27 let a dále 31 – 39 let.

**4.4.2.5 Hypotéza 5: Na internetové dotazování, v porovnání s osobním dotazováním, odpovídají častěji studenti a nezaměstnaní, a to ve větší míře, než by odpovídalo vhodně zvolenému vzorku**

**Tabulka 6 - Procentní přehled studentů a nezaměstnaných**

	STUDENTI	NEZAMĚŠTNANÍ
INTERNET	36,00 %	32,00 %
OSOBNĚ	28,00 %	4,00 %
ČR	13,86 %	7,50 %

**Zdroj:** Práce autora s použitím informací získaných z Českého statistického úřadu



Výše uvedená tabulka udává, kolik procent studentů a nezaměstnaných odpovědělo prostřednictvím dotazování a také kolik procent obyvatel ČR tvoří ve skutečnosti nezaměstnaní a studenti.

Byly provedeny **hypotézy o relativních četnostech**:

**1. Porovnání relativní četnosti mezi procentním zastoupením studentů získaným z internetového dotazování a skutečným procentním zastoupením studentů v ČR, kde:**

$H_0: \pi = 0,1386$ ,  $H_1: \pi > 0,1386$ , testové kritérium:  $U = 6,41$ , kritický obor:  $\langle 1,6449; \infty \rangle$

Testové kritérium tedy náleží do kritického oboru, a proto lze na 5% hladině významnosti potvrdit, že na internetové dotazování odpovídají častěji studenti, a to ve větší míře, než by odpovídalo vhodně zvolenému vzorku.

**2. Porovnání relativní četnosti mezi procentním zastoupením nezaměstnaných získaným z internetového dotazování a skutečným procentním zastoupením nezaměstnaných v ČR, kde:**

$H_0: \pi = 0,075$ ,  $H_1: \pi > 0,075$ , testové kritérium:  $U = 6,58$ , kritický obor:  $\langle 1,6449; \infty \rangle$

Testové kritérium náleží do kritického oboru. Na hladině významnosti 5 % se podařilo potvrdit hypotézu  $H_1$ , tedy fakt, že prostřednictvím online dotazování odpovídají častěji nezaměstnaní, a to ve větší míře, než by odpovídalo vhodně zvolenému vzorku.

**3. Porovnání relativní četnosti mezi procentním zastoupením studentů získaným z osobního dotazování a skutečným procentním zastoupením studentů v ČR, kde:**

$H_0: \pi = 0,1386$ ,  $H_1: \pi > 0,1386$ , testové kritérium:  $U = 2,89$ , kritický obor:  $\langle 1,6449; \infty \rangle$

Testové kritérium tedy náleží do kritického oboru, a proto lze na 5% hladině významnosti potvrdit, že prostřednictvím osobního dotazování odpovídají častěji studenti, a to ve větší míře, než by odpovídalo vhodně zvolenému vzorku.

**4. Porovnání relativní četnosti mezi procentním zastoupením nezaměstnaných získaným pomocí osobního dotazování a skutečným procentním zastoupením nezaměstnaných v ČR, kde:**

$H_0: \pi = 0,075$ ,  $H_1: \pi < 0,075$ , testové kritérium:  $U = -0,94$ , kritický obor:  $(-\infty; -1,6449)$

Testové kritérium tedy nenáleží do kritického oboru hodnot. Z toho vyplývá, že na 5% hladině významnosti byla zamítnuta alternativní hypotéza  $H_1$ . Lze tedy tvrdit, že prostřednictvím osobního dotazování je procentní zastoupení nezaměstnaných respondentů stejné, jako skutečné procentní zastoupení nezaměstnaných v ČR, a tedy toto procentní zastoupení nezaměstnaných, odpovídá vhodně zvolenému vzorku.

Dále byly provedeny **hypotézy o rovnosti relativních četností:**

**1. Porovnání relativní četnosti mezi procentním zastoupením studentů získaným z internetového dotazování a procentním zastoupením studentů, které bylo získáno osobním dotazováním, kde:**

$H_0: \pi_1 = \pi_2$ ,  $H_1: \pi_1 > \pi_2$ , testové kritérium:  $U = 1,45$ , kritický obor:  $(1,6449; \infty)$

Kde:  $\pi_1$  odpovídá hodnotě 36 %, což je procentní zastoupení studentů z internetového dotazování a  $\pi_2$  odpovídá hodnotě 28 %, což je procentní zastoupení studentů získané osobním dotazováním.

Testové kritérium nenáleží do kritického oboru. Na hladině významnosti 5 % nebyla zamítnuta nulová hypotéza, a tedy platí, že studenti odpovídají stejně často na internetové dotazování, tak i prostřednictvím osobního dotazování.

**2. Porovnání relativní četnosti mezi procentním zastoupením nezaměstnaných získaným z internetového dotazování a procentním zastoupením nezaměstnaných, které bylo získáno osobním dotazováním, kde:**

$H_0: \pi_1 = \pi_2$ ,  $H_1: \pi_1 > \pi_2$ , testové kritérium:  $U = 2,57$ , kritický obor:  $(1,6449; \infty)$

Kde:  $\pi_1$  odpovídá hodnotě 32 %, což odpovídá procentnímu zastoupení nezaměstnaných z internetového dotazování a  $\pi_2$  odpovídá hodnotě 4 %, tedy procentnímu zastoupení studentů z osobního dotazování.

Testové kritérium náleží do kritického oboru, a proto se dá prohlásit, že na 5% hladině významnosti se zamítá nulová hypotéza a přijímá se alternativní hypotéza. To znamená, že nezaměstnaní odpovídají mnohem častěji přes internet, než prostřednictvím osobního dotazování.

Na závěr lze tedy předpokládanou hypotézu potvrdit pouze částečně, a to takto: Potvrdilo se, že na internetové dotazování v porovnání s osobním odpovídají nezaměstnaní častěji, přičemž studenti odpovídají stejně často přes internet i prostřednictvím osobního dotazování.

#### 4.4.2.6 Hypotéza 6: Při porovnání internetového a osobního dotazování se předpokládá, že respondenti budou z rozdílných vzdělanostních skupin

**Tabulka 7 – Procentní vyjádření stupně dosaženého vzdělání**

STUPEŇ VZDĚLÁNÍ		INTERNET	OSOBNĚ	SKUTEČNOST
	ZÁKLADNÍ	23,00 %	20,00 %	14,91 %
	VYUČENÍ V OBORU (BEZ MATURITY)	33,00 %	33,00 %	28,01 %
	STŘEDOŠKOLSKÉ (S MATURITOU)	25,00 %	21,00 %	17,22 %
	VYSOKOŠKOLSKÉ	19,00 %	26,00 %	10,58 %

**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Z výše uvedené tabulky vyplývá, že rozdíly mezi internetovým a osobním dotazováním jsou znatelné zejména u vysokoškolského vzdělání, kde přes internet odpovědělo celkem 19 % respondentů a u osobního dotazování odpovědělo 26 % respondentů. Naopak co se týče vyučení v oboru (bez maturity), je zde viditelná shoda, a to ve výši 33 %. Při porovnání online dotazování, či osobního dotazování se skutečností, jsou zde již na první pohled viditelné rozdíly.

Pro potvrzení či zamítnutí výše zmíněné hypotézy byly provedeny **testy dobré shody**:

### **1. Porovnání internetového a osobního dotazování**

$H_0$ : Respondenti jsou ze stejných vzdělanostních skupin,  $H_1$ : Neplatí hypotéza  $H_0$ ,  
testové kritérium: chí-kvadrát ( $x^2$ ) = 3,0965, kritický obor:  $\langle 7,8147; \infty \rangle$

Testové kritérium nenáleží do kritického oboru, a tudíž na hladině významnosti 5 % nebyla nulová hypotéza zamítnuta. Z výsledků tedy vyplývá, že se nezamítá, že by respondenti byli ze stejných vzdělanostních skupin.

### **2. Porovnání internetového dotazování se skutečností**

$H_0$ : Respondenti jsou ze stejných vzdělanostních skupin,  $H_1$ : Neplatí hypotéza  $H_0$ ,  
testové kritérium: chí-kvadrát ( $x^2$ ) = 15,4945, kritický obor:  $\langle 7,8147; \infty \rangle$

Testové kritérium náleží do kritického oboru. Na hladině významnosti 5 % bylo prokázáno, že se shoda poměrů zamítá. Zamítá se tedy, že by respondenti byli ze stejných vzdělanostních skupin.

### **3. Porovnání osobního dotazování se skutečností**

$H_0$ : Respondenti jsou ze stejných vzdělanostních skupin,  $H_1$ : Neplatí hypotéza  $H_0$ ,  
testové kritérium: chí-kvadrát ( $x^2$ ) = 25,9305, kritický obor:  $\langle 7,8147; \infty \rangle$

Testové kritérium náleží do kritického oboru. Na hladině významnosti 5 % bylo prokázáno, že se shoda poměrů zamítá. Zamítá se tedy, že by respondenti byli ze stejných vzdělanostních skupin.

Bylo zjištěno, že při porovnání online a osobního dotazování jsou respondenti ze stejných vzdělanostních skupin. Ovšem při porovnání internetového či osobního dotazování se skutečností, bylo prokázáno, že respondenti jsou z rozdílných vzdělanostních skupin.

**4.4.2.7 Hypotéza 7: Dotazování přes internet v porovnání s osobním dotazováním, lépe odpovídá skutečnému rozdělení společnosti dle počtu obyvatel v obci (dle velikosti obce)**

**Tabulka 8 – Procentní zastoupení respondentů dle počtu obyvatel v obci**

POČET OBYVATEL		INTERNET	OSOBNĚ	SKUTEČNOST
	<b>DO 199</b>		9,00 %	3,00 %
<b>200 – 499</b>		11,00 %	19,00 %	6,21 %
<b>500 – 999</b>		11,00 %	11,00 %	8,31 %
<b>1 000 – 1 999</b>		14,00 %	5,00 %	8,66 %
<b>2 000 – 4 999</b>		10,00 %	6,00 %	10,23 %
<b>5 000 – 9 999</b>		6,00 %	2,00 %	8,84 %
<b>10 000 – 19 999</b>		8,00 %	1,00 %	9,69 %
<b>20 000 – 49 999</b>		12,00 %	53,00 %	11,63 %
<b>50 000 – 99 999</b>		5,00 %	0,00 %	11,39 %
<b>100 000 A VÍCE</b>		14,00 %	0,00 %	23,40 %

**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Z výše uvedené tabulky vyplývá, že prostřednictvím osobního dotazování odpovídalo největší procento dotazovaných z měst o velikosti 20 000 – 49 999 obyvatel, a to 53 % obyvatel, naopak přes internet odpovídají respondenti více rovnoměrně, tedy jak z menších, tak i z větších obcí či měst. Pro porovnání je v tabulce zobrazeno i to, jaké je skutečné rozdělení společnosti v ČR dle počtu obyvatel v obci.

Pro potvrzení či zamítnutí výše zmíněné hypotézy byly provedeny **testy dobré shody**:

**1. Porovnání dat získaných z internetového dotazování se skutečnými daty zjištěnými z Českého statistického úřadu**

$H_0$ : Procentní zastoupení respondentů dle počtu obyvatel je stejné, jako skutečné procentní zastoupení obyvatel ČR,  $H_1$ : Neplatí hypotéza  $H_0$ , testové kritérium: chí-kvadrát ( $\chi^2$ ) = 49,7666, kritický obor:  $\langle 16,9190; \infty \rangle$ , P-hodnota =  $1,19175243 \cdot 10^{-7}$

Testové kritérium spadá do kritického oboru, a tudíž se shoda poměrů zamítá. Na hladině významnosti 5 % se zamítá, že by procentní zastoupení respondentů dle počtu obyvatel bylo stejné, jako skutečné procentní zastoupení obyvatel ČR.

## **2. Porovnání dat získaných prostřednictvím osobního dotazování se skutečnými daty zjištěnými z Českého statistického úřadu**

$H_0$ : Procentní zastoupení respondentů dle počtu obyvatel je stejné, jako skutečné procentní zastoupení obyvatel ČR,  $H_1$ : Neplatí hypotéza  $H_0$ , testové kritérium: chí-kvadrát ( $\chi^2$ ) = 226,6964, kritický obor:  $(16,9190; \infty)$ , P-hodnota = 0

Testové kritérium náleží do kritického oboru. Shoda poměrů se tedy zamítá. Na 5% hladině významnosti se zamítá, že by procentní zastoupení respondentů dle počtu obyvatel bylo stejné, jako skutečné procentní zastoupení obyvatel ČR.

Pro určení, které z daných rozdělení lépe odpovídá skutečnému rozdělení společnosti, je důležitá právě výše zmíněná P-hodnota. Čím větší P-hodnota, tím podobnější jsou daná rozdělení. Z toho vyplývá, že při porovnání internetového dotazování se skutečností je P-hodnota větší, než při porovnání osobního dotazování se skutečností.

Lze tedy tvrdit, že internetový dotazník v porovnání s osobním dotazováním lépe odpovídá skutečnému rozdělení společnosti dle velikosti obcí, ve kterých respondenti bydlí, avšak stále není úplně vhodný.

## 4.5 Finanční otázky

### 4.5.1 Předpokládané hypotézy

1. Při porovnání internetového a osobního dotazování budou respondenti na dané otázky ve větší míře odpovídat spíše prostřednictvím internetu a naopak se předpokládá, že u osobního dotazování budou mít dotazování s finančními otázkami problém a tudíž na ně nebudou ochotni odpovídat, a to zvláště v závislosti na pohlaví a věku
2. Průměrný měsíční zdanitelný příjem získaný internetovým dotazováním lépe odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu v ČR, než osobní dotazník
3. Při osobním rozhovoru mají respondenti tendenci více zkreslovat finanční skutečnosti, než při internetovém dotazování. Zvláště v závislosti na pohlaví dotazujícího a pohlaví a věku respondenta

### 4.5.2 Dokazování předpokládaných hypotéz

**4.5.2.1 Hypotéza 1: Při porovnání internetového a osobního dotazování budou respondenti na dané otázky ve větší míře odpovídat spíše prostřednictvím internetu a naopak se předpokládá, že u osobního dotazování budou mít dotazování s finančními otázkami problém a tudíž na ně nebudou ochotni odpovídat, a to zvláště v závislosti na pohlaví a věku**

**Tabulka 9 - Procentní vyjádření respondentů, kteří neodpověděli na dané otázky**

	<b>INTERNET</b>	<b>OSOBNĚ</b>
<b>ZA CO NEJVÍCE UTRÁCÍTE PENÍZE</b>	0 %	0 %
<b>MĚSÍČNÍ ZDANITELNÝ PŘÍJEM</b>	3 %	50 %
<b>PRŮMĚRNÉ MĚSÍČNÍ OSOBNÍ VÝDAJE</b>	3 %	22 %
<b>VYUŽÍVÁTE ÚVERŮ/PŮJČEK</b>	3 %	46 %
<b>AKTUÁLNÍ VÝŠE ÚVERŮ/PŮJČEK</b>	3 %	66 %

**Zdroj:** Práce autora

Pro získání nezkráceného procentního výsledku, který určí, kolik procent dotazovaných mělo s vyplněním daných otázek problém, je nejprve nutné určit, kolik osob nevyplnilo alespoň jednu z těchto finančních otázek.

**Tabulka 10 - Počet osob, které neodpověděli alespoň na jednu z finančních otázek**

	INTERENT	TŘEBÍČ	ZNOJMO	TŘEBÍČ + ZNOJMO
<b>POČET OSOB, KTERÉ NEODPOVĚDĚLI ALESPŮŇ NA JEDNU Z FINANČNÍCH OTÁZEK</b>	5	34	37	71

**Zdroj:** Práce autora

Přes internet neodpovědělo alespoň na jednu z uvedených finančních otázek 5 respondentů. Při osobním dotazování bylo zjištěno, že v Třebíči neodpovědělo alespoň na jednu danou otázku 34 respondentů a ve Znojmě 37 respondentů. Celkově tedy u osobního dotazování došlo k tomu, že neodpovědělo 71 osob.

Z výše uvedeného zjištění vyplývá, že dotazování prostřednictvím internetu je u těchto finančních otázek mnohem lepší volbou, jelikož zde neodpovědělo pouze 5 % respondentů. Naopak u osobního dotazování se nedocílí reprezentativních výsledků, jelikož na jednotlivé otázky neodpovědělo vysoké procento dotazovaných osob, a to 71 %.

Aby mohla být výše uvedená hypotéza potvrzena či vyvrácena, je nutné, aby byl proveden **test rovnosti relativních četností**, kde:  $H_0: \pi_1 = \pi_2$ ,  $H_1: \pi_1 < \pi_2$ , testové kritérium:  $U = -3,03$ , kritický obor:  $(-\infty; -1,6449)$ , kde  $\pi_1$  odpovídá hodnotě 5 %, což určuje počet osob, kteří neodpověděli alespoň na jednu otázku přes internet, a  $\pi_2$  odpovídá hodnotě 71 %, tedy počet respondentů, kteří neodpověděli alespoň na jednu otázku u osobního dotazování.

Testové kritérium náleží do kritického oboru. Na hladině významnosti 5 % se podařilo prokázat, že při porovnání internetového a osobního dotazování odpovídají respondenti na dané otázky ve větší míře právě prostřednictvím internetu a naopak se potvrzuje, že u osobního dotazování mají dotazování s finančními otázkami problém a tudíž na ně nejsou ochotni odpovídat.

Nyní je nutno určit, zda u osobního dotazování existuje závislost mezi věkem a pohlavím respondentů. Pro tento účel byla provedena **korelační analýza**, a to tedy konkrétně  **$\chi^2$ -testem nezávislosti**, kde:  $H_0$ : Proměnné jsou nezávislé,  $H_1$ : Proměnné jsou závislé. Tento test byl proveden zvlášť pro Znojmo a zvlášť pro Třebíč.



**Tabulka 11 - Počet respondentů dle věku a pohlaví, kteří neodpověděli alespoň na jednu otázku – Znojmo**

VĚKOVÉ KATEGORIE		MUŽ	ŽENA	CELKEM
	15 – 29	4	12	16
	30 - 59	8	10	18
	60 A VÍCE	1	2	3
	<b>CELKEM</b>	13	24	<b>37</b>

**Zdroj:** Práce autora

Z tabulky vyplývá, že alespoň na jednu otázku neodpovědělo celkem 37 osob, z toho 13 mužů a 24 žen. Dále tabulka popisuje počet mužů a žen dle věkových kategorií. Ve Znojmě tedy neodpovědělo alespoň na jednu otázku 72,22 % mužů a 75 % žen.

Kritický interval odpovídá intervalu  $(0,10259; \infty)$ , testové kritérium: 1,41, počet stupňů volnosti: 2. Bylo zjištěno, že testové kritérium náleží do kritického intervalu a tudíž lze potvrdit, že na hladině významnosti 5 % bylo zjištěno, že jsou dané proměnné (věk, pohlaví) závislé. Pro zjištění míry závislosti byl použit **Pearsonův koeficient**, který odpovídá hodnotě 0,1644. Mezi proměnnými existuje mírná závislost.

**Tabulka 12 - Počet respondentů dle věku a pohlaví, kteří neodpověděli alespoň na jednu otázku – Třebíč**

VĚKOVÉ KATEGORIE		MUŽ	ŽENA	CELKEM
	15 – 29	8	5	13
	30 - 59	13	6	19
	60 A VÍCE	2	0	2
	<b>CELKEM</b>	23	11	<b>34</b>

**Zdroj:** Práce autora

Z tabulky vyplývá, že alespoň na jednu otázku neodpovědělo celkem 34 osob, z toho 23 mužů a 11 žen. Dále tabulka popisuje počet mužů a žen dle věkových kategorií. V Třebíči neodpovědělo alespoň na jednu otázku 76,67 % mužů a 55 % žen.

Kritický interval odpovídá intervalu  $(0,10259; \infty)$ , testové kritérium: 1,18, počet stupňů volnosti: 2. Bylo tedy zjištěno, že testové kritérium náleží do kritického intervalu a tudíž lze na hladině významnosti 5 % prohlásit, že jsou dané proměnné (věk, pohlaví) závislé. Dále byl

taktéž použit **Pearsonův koeficient**, který odpovídá hodnotě 0,1715. Mezi proměnnými existuje mírně silná závislost.

Z tabulek vyplývá, že pokud se na otázky finančního charakteru dotazuje žena, odpovídají častěji ženy a naopak, pokud se dotazuje muž, odpovídají častěji muži. Toto tvrzení může být ovšem zkreslené tím, že pokud se táže muž, jsou většinou dotazovány ženy, a z toho důvodu jich také více neodpovědělo a naopak, pokud se dotazuje žena, jsou většinou dotazováni muži, proto jich neodpovědělo více v porovnání se ženami. Ve Znojmě totiž bylo dotazováno 36 % mužů a 64 % žen, a v Třebíči bylo dotazováno 40 % žen a 60 % mužů.

#### 4.5.2.2 Hypotéza 2: Průměrný měsíční zdanitelný příjem získaný internetovým dotazováním lépe odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu v ČR, než osobní dotazník

**Tabulka 13 – Přehled průměrného měsíčního zdanitelného příjmu**

	INTERNET	OSOBNĚ	SKUTEČNOST
<b>PRŮMĚRNÝ MĚSÍČNÍ ZDANITELNÝ PŘÍJEM</b>	29 371 Kč	22 722 Kč	27 200 Kč

**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Bylo zjištěno, že průměrný měsíční zdanitelný příjem získaný prostřednictvím internetového dotazování činí 29 371 Kč. Průměrný měsíční zdanitelný příjem získaný osobním dotazováním činí 22 722 Kč. Dále bylo zjištěno, že skutečný průměrný měsíční zdanitelný příjem za 4. čtvrtletí 2014 činil v ČR 27 200 Kč.

Pro hlubší porovnání je nutné provést **hypotézy o střední hodnotě:**

#### 1. Porovnání průměrného měsíčního zdanitelného příjmu získaného z internetového dotazování se skutečným průměrným měsíčním zdanitelným příjmem v ČR, kde:

$H_0: \mu = 27\,200$ ,  $H_1: \mu \neq 27\,200$ , testové kritérium:  $t = 0,90$ , kritický obor:  $(2,0423; \infty)$

Testové kritérium tedy nenáleží do kritického oboru. Na hladině významnosti 5 % byla zamítnuta alternativní hypotéza, a z toho důvodu lze tvrdit, že průměrný měsíční zdanitelný příjem získaný z internetového dotazování odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu v ČR.

## 2. Porovnání průměrného měsíčního zdanitelného příjmu získaného osobním dotazováním se skutečným průměrným měsíčním zdanitelným příjmem v ČR

$H_0: \mu = 27\,200, H_1: \mu < 27\,200$ , testové kritérium:  $t = 1,87$ , kritický obor:  $(1,7396; \infty)$

Testové kritérium náleží do kritického oboru. Rozdíl mezi danými průměry je významný. Z toho vyplývá, že na 5% hladině významnosti bylo potvrzeno, že průměrný měsíční zdanitelný příjem získaný osobním dotazováním je menší, než skutečný průměrný měsíční zdanitelný příjem v ČR.

Bylo potvrzeno, že průměrný měsíční zdanitelný příjem získaný internetovým dotazováním lépe odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu v ČR, než osobní dotazník. Z výsledků vyplývá, že při osobním rozhovoru, v porovnání s internetovým dotazováním, jsou získávány nižší hodnoty průměrných zdanitelných příjmů.

### 4.5.2.3 Hypotéza 3: Při osobním rozhovoru mají respondenti tendenci více zkreslovat finanční skutečnosti, než při internetovém dotazování. Zvláště v závislosti na pohlaví dotazujícího a pohlaví respondenta

Pro účely této hypotézy bude zkoumána finanční otázka týkající se průměrného měsíčního zdanitelného příjmu.

**Tabulka 14 - Průměrný měsíční zdanitelný příjem dle pohlaví**

	INTERNET	TŘEBÍČ	ZNOJMO	SKUTEČNOST
MUŽ	31 084 Kč	29 450 Kč	nezjištěno	29 846 Kč
ŽENA	23 498 Kč	nezjištěno	17 600 Kč	22 825 Kč

**Zdroj:** Práce autora s použitím informací z Českého statistického úřadu

Z výše uvedené tabulky vyplývá, že průměrný měsíční zdanitelný příjem získaný přes internet je u mužů 31 084 Kč a u žen 23 498 Kč. Dále z tabulky vyplývá, že zjištěný průměrný měsíční zdanitelný příjem u mužů v Třebíči činí 29 450 Kč a u žen nebyl měsíční příjem zjištěn, jelikož 5,55 % tvoří nezaměstnané ženy, studentky tvoří 44,44 % a 44,44 % neodpovědělo vůbec. V Třebíči odpověděla pouze jedna žena, jejíž měsíční zdanitelný příjem činí 15 000 Kč. Jelikož odpověděla pouze jedna žena, není tento údaj reprezentativní, a tudíž jej nelze brát v úvahu. Ve Znojmě činí průměrný měsíční zdanitelný příjem u žen 17 600 Kč a u mužů nebyl příjem zjištěn, jelikož zde odpovědělo 6,67 % nezaměstnaných, 33,33 % studentů a zbytek, tedy 60 % neodpovědělo vůbec. Ve skutečnosti, tedy v ČR činí průměrný

měsíční zdanitelný příjem u mužů 29 846 Kč a u žen 22 825 Kč. Pro účely zkoumání dané hypotézy nebudou brány v úvahu hodnoty nezjištěného průměrného měsíčního zdanitelného příjmu mužů ve Znojmě a žen v Třebíči, jelikož zde není k dispozici reprezentativní výsledek, který by se dal zkoumat.

Pro porovnání byly provedeny **hypotézy o střední hodnotě**:

**1. Porovnání průměrného měsíčního zdanitelného příjmu z internetového šetření se skutečností u mužů**

$H_0: \mu = 29846, H_1: \mu \neq 29846$ , testové kritérium:  $t = 0,45$ , kritický obor:  $(2,0287; \infty)$ .

Na hladině významnosti 5 % bylo zjištěno, že průměrný měsíční zdanitelný příjem mužů, získaný prostřednictvím internetu, odpovídá skutečnému průměrnému zdanitelnému příjmu mužů v ČR.

**2. Porovnání průměrného měsíčního zdanitelného příjmu z internetového šetření se skutečností u žen**

Byla provedena hypotéza o střední hodnotě, kde:  $H_0: \mu = 22825, H_1: \mu \neq 22825$ , testové kritérium:  $t = 0,16$ , kritický obor:  $(2,4469; \infty)$ .

Na hladině významnosti 5 % bylo zjištěno, že průměrný měsíční zdanitelný příjem žen, získaný prostřednictvím internetu, odpovídá skutečnému průměrnému zdanitelnému příjmu žen v ČR.

**3. Porovnání průměrného měsíčního zdanitelného příjmu z osobního dotazování (Znojmo) se skutečností u žen**

Byla provedena hypotéza o střední hodnotě, kde:  $H_0: \mu = 22825, H_1: \mu < 22825$ , testové kritérium:  $t = 2,10$ , kritický obor:  $(1,8595; \infty)$ .

Na hladině významnosti 5 % bylo zjištěno, že průměrný měsíční zdanitelný příjem žen, získaný osobním dotazováním ve Znojmě, je nižší, než skutečný průměrný měsíční zdanitelný příjem žen v ČR.

#### **4. Porovnání průměrného měsíčního zdanitelného příjmu z osobního dotazování (Třebíč) se skutečností u mužů**

Byla provedena hypotéza o střední hodnotě, kde:  $H_0: \mu = 29846$ ,  $H_1: \mu \neq 29846$ , testové kritérium:  $t = -0,12$ , kritický obor:  $(2,3646; \infty)$ .

Na hladině významnosti 5 % bylo zjištěno, že průměrný měsíční zdanitelný příjem mužů, získaný osobním dotazováním v Třebíči, odpovídá skutečnému průměrnému měsíčnímu příjmu mužů v ČR.

Výše předpokládanou hypotézu lze tedy z části potvrdit a z části vyvrátit, a to konkrétně takto: Z výše uvedených zjištění vyplývá, že průměrný měsíční zdanitelný příjem získaný internetovým dotazováním odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu ve skutečnosti v ČR, a to jak u mužů, tak u žen. Dále bylo zjištěno, že pokud je tazatel ženského pohlaví a dotazuje se mužů na otázku zdanitelného příjmu, tak muž ženě odpovídá podle pravdy, jelikož bylo zjištěno, že průměrný měsíční zdanitelný příjem mužů v Třebíči odpovídá skutečnému průměrnému měsíčnímu zdanitelnému příjmu mužů v ČR. Ze zjištění dále vyplývá, že pokud je tazatel mužského pohlaví a dotazuje se žen, tak ženy mají tendenci říkat, že jejich průměrný měsíční zdanitelný příjem je nižší, než ve skutečnosti opravdu je, a tím tedy dochází ke zkreslení zjištěných údajů.

## 4.6 Návrhy a doporučení

Na základě testování předpokládaných hypotéz byly vyvozeny závěry, které jsou důležité pro potenciální tazatele, kteří mají zájem získat určitý návod, kterým by se měli řídit při různých průzkumech trhu. Na základě těchto návrhů a doporučení lze vyvodit, zda je lepší pro určité otázky využívat internetového dotazování, nebo použít raději osobní dotazování.

Pomocí shlukové analýzy bylo zjištěno, že pokud se tazatel bude dotazovat prostřednictvím online dotazování, měl by očekávat, že nejvíce respondentek (žen) bude odpovídat ve věku 15 – 30 let a respondentů (mužů) ve věkové kategorii 19 – 27 let a dále 31 – 39 let.

Pokud se průzkum trhu bude zabývat nezaměstnaností, kdy tazatel bude chtít zjistit jakékoliv údaje o nezaměstnaných, je výhodnější provádět průzkum prostřednictvím internetového dotazování, jelikož v porovnání s osobním dotazováním, odpovídají nezaměstnaní mnohem častěji přes internet.

Za předpokladu, že tazatel bude chtít zjistit co nejvíce údajů o studentech – zde záleží pouze na samotném tazateli, který způsob dotazování využije, jelikož bylo zjištěno, že pro tento účel lze využít jak internetové, tak i osobní dotazování a výsledek bude reprezentativní. Studenti totiž odpovídají stejně často přes internet i prostřednictvím osobního dotazování

V případě, že tazatel bude chtít získat ničím nezkreslené informace jak o ženách, tak i o mužích, doporučuje se, aby byl tým tazatelů smíšeného pohlaví, to tedy znamená, aby průzkum prováděla alespoň jedna žena a alespoň jeden muž. Bylo totiž prokázáno, že tazatel muž i tazatelka žena si vybírají dotazované respondenty stejným způsobem, a to tak, že pokud se dotazuje žena, ptá se záměrně mužů a naopak, pokud se dotazuje muž, ptá se záměrně žen a tím dochází ke zkreslení vybíraného vzorku ve prospěch osob druhého pohlaví.

Pokud se průzkum bude zabývat pohlavím a tazatel bude chtít získat nezkreslený vzorek respondentů, tedy takový vzorek, který bude odpovídat skutečnému rozdělení společnosti, doporučuje se využít internetového dotazování. Bylo totiž zjištěno, že internetové dotazování respektuje skutečné rozdělení osob dle pohlaví ve společnosti.

Tazatel by měl předpokládat, že pokud bude použito osobní dotazování pouze ve větším městě, tak na něj budou odpovídat ve větší míře právě obyvatelé daného města, tudíž hlavní skupina respondentů bude tvořena z obce větší, než například 20 000 obyvatel. Doporučuje se tedy zvolit toto osobní dotazování za předpokladu, že k danému průzkumu bude zapotřebí

získat údaje od respondentů, kteří bydlí v konkrétním městě či obci. Internetové dotazování v tomto případě není příliš vhodné, jelikož by respondenti odpovídali v rámci celé ČR.

Internetové dotazování se dále doporučuje i za předpokladu, že tazatel bude chtít znát finanční situaci respondenta, jako například jaký je jeho průměrný měsíční zdanitelný příjem, zda využívá úvěrů/půjček, jaká je aktuální výše úvěrů/půjček, jaké jsou jeho průměrné měsíční osobní výdaje a další. Pokud se totiž dotazování uskutečňuje osobně, mají respondenti problém na tyto otázky odpovídat.

Pokud ovšem bude chtít tazatel prostřednictvím osobního dotazování získat co nejvíce reprezentativní údaje pouze o zdanitelném příjmu, doporučuje se, aby dotazování u mužů prováděla žena, jelikož bylo prokázáno, že muž ženě odpovídá podle pravdy a tak tedy nedochází ke zkreslování zjištěných údajů. Naopak se nedoporučuje, aby průzkum u žen prováděl muž, jelikož bylo zjištěno, že žena muži odpovídá záměrně zkresleně a tudíž jsou takto získaná data nerepresentativního charakteru.

**Tabulka 15 – Vhodnost internetového a osobního dotazování**

<b>INTERNETOVÉ DOTAZOVÁNÍ JE VHODNÉ PRO ZÍSKÁNÍ REPREZENTATIVNÍCH ÚDAJŮ O:</b>	<b>OSOBNÍ DOTAZOVÁNÍ JE VHODNÉ PRO ZÍSKÁNÍ REPREZENTATIVNÍCH ÚDAJŮ O:</b>
Respondentech v rámci celé ČR	Respondentech z konkrétní lokality (města, obce)
Studentech	Studentech
Finanční situaci respondentů	Respondentech v určité věkové kategorii
Průměrném měsíčním zdanitelném příjmu v ČR	Průměrném měsíčním zdanitelném příjmu mužů, za předpokladu, že tazatel bude žena
Mužích i ženách (rozdělení mužů a žen bude odpovídat skutečnému rozdělení mužů a žen v ČR)	Mužích i ženách, za předpokladu, že tým tazatelů bude smíšeného pohlaví (jeden muž, jedna žena)
Nezaměstnaných respondentech	

**Zdroj:** Práce autora

Z výše uvedené tabulky vyplývá, jaký typ dotazování je vhodný pro konkrétní průzkum trhu.

## 4.7 Shrnutí praktické části

Začátek praktické části se věnuje základním informacím: Pro praktickou část byly zvoleny dva způsoby dotazování, a to internetové a osobní. Nejprve bylo nutné sestavit dotazník, který obsahoval celkem 18 otázek. Internetové dotazování probíhalo od 1. 10. 2014 do 31. 12. 2014 a osobní dotazování probíhalo od 3. 11. 2014 do 14. 11. 2014. Osobní dotazování se uskutečnilo v Třebíči, kde se dotazovala žena a ve Znojmě, kde se dotazoval muž. Celkem bylo získáno 200 odpovědí. Online dotazováním bylo získáno 100 odpovědí. Prostřednictvím osobního dotazování bylo získáno taktéž 100 odpovědí, přičemž 50 odpovědí bylo získáno v Třebíči a dalších 50 ve Znojmě.

Dotazník se skládá ze tří podkategorií, a to z osobních otázek, všeobecných otázek a finančních otázek. Pro osobní i finanční otázky byly zvoleny předpokládané, tedy alternativní hypotézy.

Dále se praktická část věnuje dokazování předpokládaných hypotéz. Tyto předpokládané hypotézy řeší otázky, které se týkají zejména:

- Věku
- Pohlaví
- Nezaměstnanosti
- Nejvyššího dosaženého vzdělání
- Rozdělení společnosti dle počtu obyvatel v obci
- Otázek finančního charakteru, se zaměřením zejména na průměrný měsíční zdanitelný příjem

Tyto hypotézy mají za cíl určit, zda je pro dané otázky výhodnější osobní dotazování, či internetové dotazování. Právě těmto zjištěním se věnuje samostatná podkapitola: Návrhy a doporučení. Tyto návrhy a doporučení mají za úkol pomoci potenciálním tazatelům při průzkumech trhu.



## 5 ZÁVĚR

Bakalářská práce na téma: *Způsoby získávání statistických dat. Praktické porovnání různých postupů*, se zabývala zjištěním, zda je výhodnější či více vypovídající získávání statistických dat za pomoci internetového dotazování, nebo spíše prostřednictvím osobního dotazování, a to vše pro různé skupiny otázek.

V teoretické části byly shrnuty veškeré důležité informace týkající se způsobů získávání statistických dat. Dále zde bylo rovněž uvedeno, jakým způsobem se vyhodnocuje dotazníkové šetření a jaké hypotézy se při tomto vyhodnocování používají.

Celá praktická část se věnovala dokazování předpokládaných hypotéz. Celkem bylo otestováno 10 komplexních hypotéz. Tyto komplexní hypotézy se často testovaly pomocí celé řady jednodušších hypotéz. Dané hypotézy byly testovány prostřednictvím různých testů. Celkem bylo provedeno 32 testů. Tyto testy měly za úkol danou hypotézu buď potvrdit, nebo zamítnout. Na základě výsledků z daných analýz byly vytvořeny návrhy a doporučení. Tyto návrhy a doporučení byly vytvořeny za účelem pomoci potenciálním respondentům, kteří budou provádět jakýkoliv průzkum trhu. Tyto návrhy jsou totiž jakýmsi manuálem, podle kterého by se mohli tazatelé řídit. Návrhy a doporučení se týkají jak otázek osobního, tak i finančního charakteru.

Mezi návrhy, které byly vyvozeny, patří například: Pokud se bude průzkum trhu zabývat nezaměstnaností, za výhodnější se ukázalo dotazování přes internet, jelikož nezaměstnaní tímto způsobem odpovídají mnohem častěji. Internetové dotazování respektuje skutečné rozdělení osob dle pohlaví ve společnosti. Dalším faktem je, že respondenti mají při osobním dotazování problémy odpovídat na finanční otázky, tudíž se doporučuje online dotazování. Pohlaví tazatele má vliv na zkreslení vybíraného vzorku respondentů, a to ve prospěch osob druhého pohlaví, a proto je výhodné sestavit smíšené týmy tazatelů. Ukázalo se, že studenti odpovídají stejně často přes internet i prostřednictvím osobního dotazování, proto si zde tazatel může zvolit metodu dotazování. Pro získání pravdivých údajů o průměrném měsíčním zdanitelném příjmu se doporučuje, aby dotazování u mužů prováděla žena, jelikož bylo zjištěno, že muž ženě odpovídá podle pravdy, naopak žena muži uvádí zkreslené informace, a tím dochází k nereprezentativním výsledkům daného průzkumu.

Na závěr byla vytvořena tabulka, která shrnuje, jaký typ dotazování je vhodný pro konkrétní průzkum trhu.

Internetové dotazování je vhodné pro získání údajů o: respondentech v rámci celé ČR, studentech, finanční situaci respondentů, průměrném měsíčním zdanitelném příjmu v ČR, mužích i ženách – toto rozdělení odpovídá skutečnému rozdělení mužů a žen v ČR a také nezaměstnaných respondentech.

Osobní dotazování je vhodné pro získání údajů o: respondentech z konkrétní lokality – město, obec, studentech, respondentech v určité věkové kategorii, průměrném měsíčním zdanitelném příjmu u mužů, za předpokladu, že se dotazuje žena, mužích i ženách, za předpokladu, že tým tazatelů bude smíšeného pohlaví.

Tímto byl splněn cíl této bakalářské práce, neboť byla provedena analýza a porovnání různých způsobů dotazování (osobní, online), bylo provedeno zhodnocení a následně sepsána různá doporučení.

## 6 SEZNAM POUŽITÉ LITERATURY

### KNIŽNÍ ZDROJE:

BÍLKOVÁ, Diana a kol. 2009. *Pravděpodobnost a statistika*. 1. vyd. Plzeň: Aleš Čeněk, 639 s. ISBN 978-80-7380-224-0

BOUČKOVÁ, Jana a kol. 2003. *Marketing*. 1. vyd. Praha: C. H. Beck, 432 s. ISBN 80-7179-577-1

FORET, Miroslav; STÁVKOVÁ, Jana. *Marketingový výzkum: jak poznávat své zákazníky*. 1. vyd. Praha: Grada Publishing, 2003. 159 s. ISBN 80-247-0385-8

HINDLS, Richard; HRONOVÁ, Stanislava; NOVÁK, Ilja. *Metody statistické analýzy pro ekonomy*. 2. vyd. Praha: Management Press, 2000. 259 s. ISBN 80-7261-013-9

HINDLS, Richard; HRONOVÁ, Stanislava; SEGER, Jan. *Statistika pro ekonomy*. 5. vyd. Praha: Professional Publishing, 2004. 415 s. ISBN 80-86419-59-2

KOZEL, Roman a kol. *Moderní metody a techniky marketingového výzkumu*. 1. vyd. Praha: Grada Publishing, 2011. 304 s. ISBN 978-80-247-3527-6

PAUKNEROVÁ, Daniela a kol. *Psychologie pro ekonomy a manažery*. 3. vyd. Praha: Grada Publishing, 2012. 264 s. ISBN 978-80-247-3809-3

ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření*. 1. vyd. Praha: Professional Publishing, 2007. 216 s. ISBN 978-80-86946-49-8

SIMOVÁ, Jozefína. *Marketingový výzkum*. Liberec, 2005. Technická univerzita v Liberci, 121 s. ISBN 80-7372-014-0

ZBOŘIL, Kamil. *Marketingový výzkum*. Praha: Vysoká škola ekonomická v Praze, 1994. 106 s. ISBN 80-7079-389-9

## **INTERNETOVÉ ZDROJE:**

DOTAZNÍK-ONLINE, © 2007. Dotazník je... *Dotazník-online: jak na dotazník* [online]. [cit. 2014-09-06]. Dostupné z: <http://www.dotaznik-online.cz/index.htm>

KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008. 74 s. Bakalářská práce. ZÁPADOČESKÁ UNIVERZITA V PLZNI. Vedoucí práce Ing. Kateřina Vokáčová. [online]. [cit 2014-08-21]. Dostupné z:

<https://nit.felk.cvut.cz/~dark/Petr/Ivana/Dotazkiny%20-%20metody/BP%20-%20Dotaznikove%20setreni.pdf>

KŘÁPEK, Milan. *Statistika – distanční studijní opora*. 1. vyd. SVŠE Znojmo, 2013. 116 s. ISBN 978-80-87314-40-1 [online]. [cit 2014-08-21]. Dostupné z:

<http://www.svse.cz/uploads/File/statistika%20Ix.pdf>

KUČERA, Jiří. *Shluková analýza*. ©2008, [online]. [cit 2014-09-21]. Dostupné z: [http://is.muni.cz/th/172767/fi\\_b/5739129/web/web/main.html](http://is.muni.cz/th/172767/fi_b/5739129/web/web/main.html)

## **ZAHRANIČNÍ ZDROJE:**

FINK, Arlene. *How to Conduct Surveys: A Step-by-Step Guide*. 5. vyd. USA: SAGE Publications, 2013. 200 s. ISBN 978-1-4522-0387-4

CHURCH, Allan H.; WACLAWSKI, Janine. *Designing and Using Organizational Surveys: A Seven-Step Process*. 1. vyd. New York: HB Printing, 2001. 320 s. ISBN 0-7879-5677-5

## **7 SEZNAM OBRÁZKŮ, GRAFŮ A TABULEK**

### **SEZNAM OBRÁZKŮ**

Obrázek 1 - Přímé písemné dotazování.....	20
Obrázek 2 - Zprostředkované dotazování.....	20

### **SEZNAM GRAFŮ**

Graf 1 - Rozdělení statistických proměnných podle typu .....	12
Graf 2- Porovnání procentuálního zastoupení respondentů dle věkových kategorií.....	36
Graf 3 - Porovnání skutečného a zjištěného zastoupení mužů a žen v Třebíči .....	37
Graf 4 - Porovnání skutečného a zjištěného zastoupení mužů a žen ve Znojmě.....	38
Graf 5 – Porovnání procentního zastoupení mužů a žen z internetového dotazování se skutečností .....	39
Graf 6 – Respondenti, kteří odpovídali prostřednictvím internetu – znázornění dle věku a pohlaví .....	40

### **SEZNAM TABULEK**

Tabulka 1 - Výhody a nevýhody otevřených a uzavřených otázek .....	25
Tabulka 2 - Výhody a nevýhody dotazníku.....	28
Tabulka 3 - Porovnání průměrného věku z internetového dotazování se skutečností.....	33
Tabulka 4 - Procentní zastoupení obyvatel dle věkových kategorií .....	34
Tabulka 5 - Procentní zastoupení respondentů dle věkových kategorií .....	35
Tabulka 6 - Procentní přehled studentů a nezaměstnaných.....	40
Tabulka 7 – Procentní vyjádření stupně dosaženého vzdělání .....	43
Tabulka 8 – Procentní zastoupení respondentů dle počtu obyvatel v obci.....	45
Tabulka 9 - Procentní vyjádření respondentů, kteří neodpověděli na dané.....	47

Tabulka 10 - Počet osob, které neodpověděli alespoň na jednu z finančních otázek.....	48
Tabulka 11 - Počet respondentů dle věku a pohlaví, kteří neodpověděli alespoň na jednu otázku – Znojmo.....	49
Tabulka 12 - Počet respondentů dle věku a pohlaví, kteří neodpověděli alespoň na jednu otázku – Třebíč.....	49
Tabulka 13 – Přehled průměrného měsíčního zdanitelného příjmu .....	50
Tabulka 14 - Průměrný měsíční zdanitelný příjem dle pohlaví.....	51
Tabulka 15 – Vhodnost internetového a osobního dotazování.....	55

## **8 SEZNAM PŘÍLOH**

Příloha 1 - Dotazník k bakalářské práci

# PŘÍLOHY

## Příloha 1 - Dotazník k bakalářské práci

Dobrý den,

jmenuji se Marie Tručková a jsem studentkou SVŠE Znojmo - obor Účetnictví a finanční řízení podniku.

Věnujte prosím několik minut svého času vyplnění následujícího dotazníku, který je součástí mé bakalářské práce na téma: *Způsoby získávání statistických dat. Praktické porovnání různých postupů*

Vyberte vždy pouze jednu možnost odpovědi.

Pokud na některou otázku nebudete chtít odpovídat, tak nemusíte - stačí vybrat možnost nechci odpovídat. Povinná je pouze otázka č. 1 - pohlaví.

Předem Vám mockrát děkuji za vyplnění.

**Důležité:** Minimální věk respondenta musí být 15 let

S pozdravem

Marie Tručková

1. Pohlaví:
  - a. Muž
  - b. Žena
  
2. Věk:
  - a. Zde uveďte: .....
  - b. Nechci odpovídat
  
3. Místo narození:
  - a. Zde uveďte: .....
  - b. Nechci odpovídat



4. Místo bydliště: (Název města/vesnice)
  - a. Zde uveďte: .....
  - b. Nechci odpovídat
  
5. Rodinný stav:
  - a. Svobodný/Svobodná
  - b. Ženatý/Vdaná
  - c. Rozvedený/Rozvedená
  - d. Vdovec/Vdova
  - e. Jiná možnost – v tomto případě uveďte: .....
  - f. Nechci odpovídat
  
6. Počet dětí:
  - a. Zde uveďte: .....
  - b. Nemám děti
  - c. Nechci odpovídat
  
7. Typ bydlení:
  - a. Rodinný dům – ve vlastnictví
  - b. Rodinný dům – v nájmu
  - c. Byt – ve vlastnictví
  - d. Byt – v nájmu
  - e. Jiná možnost – v tomto případě uveďte: .....
  - f. Nechci odpovídat
  
8. Nejvyšší dosažené vzdělání:
  - a. Bez vzdělání
  - b. Základní
  - c. Vyučení v oboru (bez maturity)
  - d. Středoškolské (s maturitou)
  - e. Vysokoškolské
  - f. Jiná možnost – v tomto případě uveďte: .....
  - g. Nechci odpovídat
  
9. Zaměstnání: (př.: podnikatel, sekretářka, číšník, ...)
  - a. Zde uveďte: .....
  - b. Jsem nezaměstnaný/nezaměstnaná
  - c. Jsem student/studentka
  - d. Nechci odpovídat

10. Nejvíce vyhledávaný typ dovolené/rekreace:
- Moře
  - Stanování
  - Lázně
  - Hory
  - Jiná možnost – v tomto případě uveďte: .....
  - Nechci odpovídat
11. Nejvíce upřednostňovaný druh dopravy:
- Automobil
  - Autobus
  - Vlak
  - Jiná možnost – v tomto případě uveďte: .....
  - Nechci odpovídat
12. Jak často cestujete do zahraničí? (Za jakýmkoliv účelem)
- Jedenkrát měsíčně
  - Jedenkrát ročně
  - Vícekrát měsíčně
  - Vícekrát ročně
  - Jiná možnost – v tomto případě uveďte: .....
  - Nechci odpovídat
13. Ohodnoťte známkou jako ve škole úroveň školství v ČR:
- 1 = výborná úroveň
  - 2 = chvalitebná úroveň
  - 3 = dobrá úroveň
  - 4 = dostatečná úroveň
  - 5 = nedostatečná úroveň
  - Nechci odpovídat
14. Za co nejvíce utrácíte peníze?
- Potraviny
  - Oblečení
  - Kosmetika
  - Elektronika
  - Dovolená
  - Jiná možnost – v tomto případě uveďte: .....
  - Nechci odpovídat

15. Kolik Kč činí Váš měsíční zdanitelný příjem (měsíční hrubá mzda)?

- a. Zde uveďte: .....
- b. Nepracuji
- c. Nechci odpovídat

16. Kolik Kč měsíčně činí v průměru Vaše osobní výdaje?

Zaokrouhľujte na celé tisíce podle matematických pravidel.

(Např.: 1 320 Kč => 1 000 Kč)

- a. Zde uveďte: .....
- b. Nechci odpovídat

17. Využíváte v osobním životě úvěrů/půjček?

- a. Ano – vyberte jednu z možností:
  - a. Na bydlení
  - b. Na studium
  - c. Na automobil
  - d. Jiná možnost: .....
- b. Ne
- c. Nechci odpovídat

18. Aktuální výše úvěrů/půjček:

- a. Zde uveďte: .....
- b. Nemám žádný úvěr ani půjčku
- c. Nechci odpovídat