

**UNIVERZITA JANA AMOSE KOMENSKÉHO PRAHA**

**MAGISTERSKÉ PREZENČNÍ STUDIUM  
2013 - 2015**

**DIPLOMOVÁ PRÁCE**

**Veronika Doskočilová**

**Výběrová šetření - teorie a praxe**

Praha 2015

Vedoucí diplomové práce: RNDr. Jaroslav Huk, CSc.

**JAN AMOS KOMENSKY UNIVERSITY PRAGUE**

**MASTER FULL-TIME STUDIES  
2013 - 2015**

**DIPLOMA THESIS**

**Veronika Doskočilová**

**Sample surveys: The theory and the practise**

Prague 2015

The Diploma Thesis Work Supervisor: RNDr. Jaroslav Huk, CSc.

## **Prohlášení**

Prohlašuji, že předložená diplomová práce je mým původním autorským dílem, které jsem vypracovala samostatně. Veškerou literaturu a další zdroje, z nichž jsem při zpracování čerpala, v práci řádně cituji a jsou uvedeny v seznamu použitých zdrojů.

Souhlasím s prezenčním zpřístupněním své práce v univerzitní knihovně.

V Praze dne

*Jméno autorky .....*

## **Poděkování**

Chtěla bych velmi poděkovat vedoucímu práce RNDr. Jaroslavu Hukovi, CSc., za intenzivní spolupráci a obrovskou ochotu zodpovědět všechny mé dotazy při vypracovávání diplomové práce. Jeho pomoc, konzultace a veškeré rady byly pro mou práci jednoznačným přínosem.

## **Anotace**

Diplomová práce se zabývá výběrovým šetřením. Zkoumá je z pohledu teoretického, ale i z pohledu praktického. Diplomová práce popisuje, jaké možnosti jsou v teoretické oblasti výběrových šetření a jaké možnosti jsou nejvíce uplatňovány při výběrových šetřeních v organizacích zabývajících se výběrovým šetřením v České republice.

## **Klíčová slova**

číselné škály, dosběr, kvótní výběr, nepravděpodobnostní výběr, non-response, pravděpodobnostní výběr, škály, výběr vzorku, výběrová šetření

### **Annotation**

The diploma thesis is focused on surveys. It deals with theoretical and practical point of view. It tries to describe possibilities of surveys in theory and in the same time it tries to describe the most using ways of surveys in organizations in Czech republic.

### **Key words**

non-response, number scale, quota scale, sampling methods, stochastic supplementary data collection, survey, surveys,

# OBSAH

|                                             |    |
|---------------------------------------------|----|
| ÚVOD.....                                   | 9  |
| TEORETICKÁ ČÁST .....                       | 11 |
| 1 VÝBĚROVÁ ŠETŘENÍ.....                     | 12 |
| 2 PRAVDĚPODOBNOSTNÍ VÝBĚR.....              | 14 |
| 2.1 Prostý náhodný výběr.....               | 15 |
| 2.2. Oblastní výběr .....                   | 17 |
| 2.3. Skupinový výběr .....                  | 18 |
| 2.4. Vícetupňový výběr .....                | 19 |
| 3 NEPRAVDĚPODOBNOSTNÍ VÝBĚR .....           | 21 |
| 3.1. Anketa .....                           | 22 |
| 3.2. Snowball.....                          | 23 |
| 3.3. Úsudkové výběry .....                  | 25 |
| 4 METODY SBĚRU DAT .....                    | 27 |
| 4. 1 Pravidla tvorby dotazníku .....        | 28 |
| 4.2 Otázky dle pořadí .....                 | 30 |
| 4.3 Otázky dle účelu.....                   | 32 |
| 4.4 Otázky dle obsahu .....                 | 33 |
| 4.5 Otázky dle formy.....                   | 34 |
| 5 MĚŘENÍ A ŠKÁLY .....                      | 36 |
| 5.1 Stevensova typologie .....              | 37 |
| 5.2 Bodovací škály.....                     | 38 |
| 5.3 Likertova metoda .....                  | 39 |
| 5.4 Sémantický diferenciál.....             | 40 |
| 5.5 Metoda zdánlivě stejných intervalů..... | 42 |
| 5.6 Skalogram .....                         | 43 |
| 5.7 Metoda párových srovnání.....           | 44 |
| 5.8 Zlomkové škály .....                    | 46 |
| 6 PRÁCE S DATY .....                        | 47 |
| 6.1 Míry polohy (úrovně).....               | 48 |
| 6.2 Míry variability .....                  | 53 |
| 6.3 Chyby při sběru dat .....               | 55 |

|                                          |    |
|------------------------------------------|----|
| 6.4 Doplnění dat .....                   | 57 |
| 6.5 Publikování dat .....                | 58 |
| PRAKTICKÁ ČÁST .....                     | 60 |
| 7 CÍL, OKRUHY OTÁZEK A HYPOTÉZY .....    | 60 |
| 8 METODIKA .....                         | 62 |
| 8.1. Hlubkové rozhovory.....             | 62 |
| 8.2 Harmonogram výzkumu.....             | 64 |
| 9 VÝSLEDKY VÝZKUMU .....                 | 65 |
| 9.1. Nejužívanější typy sběru dat .....  | 65 |
| 9.2. Sudé versus liché škály .....       | 67 |
| 9.3. Používání číselných škál .....      | 69 |
| 9.4. Problémy při sběru dat .....        | 70 |
| 9.5. Doplnování dat .....                | 71 |
| 10 INTERPRETACE VÝSLEDKŮ A DISKUZE ..... | 73 |
| ZÁVĚR .....                              | 78 |
| SEZNAM POUŽITÝCH ZDROJŮ .....            | 79 |
| SEZNAM OBRÁZKŮ, GRAFŮ A TABULEK.....     | 82 |
| SEZNAM PŘÍLOH.....                       | 83 |



# ÚVOD

Výběrová šetření jsou v současné době používána ve všech organizacích zabývajících se marketingovými, sociologickými a dalšími průzkumy v České republice. Jde o aktuální téma, které má své zastoupení v teorii i v praxi.

Autorka práce zvolila uvedené téma zejména vzhledem k předchozímu studiu statistiky. Vzhledem k této skutečnosti se rozhodla propojit tuto znalost statistiky se znalostmi nabytými zde, zejména pak z oblasti marketingu a sociologie. Propojení těchto oborů shledává autorka za velmi přínosné.

Cílem práce je porovnat teoretické možnosti realizace výběrových šetření s metodami prakticky využívanými institucemi v ČR. Cíle této práce mají přispět k většímu povědomí o tom, zda a jakým způsobem se liší některé postupy a metody sběru dat v rámci teoretické a praktické roviny.

V teoretické části práce se autorka zaměřuje na teoretické poznatky, definující postupy výběru vzorku při výběrových šetřeních a jejich pestré možnosti. Dále se zaměřuje na postupy při sběru dat a metody tvorby dotazníku. Autorka zmiňuje i postupy a metody škálování dat při průzkumech. Zaměřuje se i na následné zpracování dat, a uvádí možnosti, jakými se data následně vyhodnocují. Na závěr teoretické části se autorka zaměřuje na potíže při celém šetření, jako jsou chyby při sběru dat a možnosti jejich následného doplnění. Teoretická část je zakončena kapitolou o publikování výsledných dat.

V praktické části je definován výzkumný cíl a hypotézy. Dále je podrobně rozepsána metodika použitá při realizaci praktické části a harmonogram výzkumu. Následně jsou předloženy výsledky výzkumu a poté jejich interpretace spojená s diskuzí výsledků. poté autorka shrnuje závěry výzkumu a předkládá i možnosti dalšího výzkumu na základě zjištěných informací.

Pro zpracování praktické části byla použita metoda hloubkových rozhovorů. Osloveni byli odborníci z organizací v ČR, které se zabývají výběrovým šetřením.

Při zpracování neměla autorka potíže se sháněním podkladů a materiálů. Rovněž nenastaly potíže ani při praktické části při kontaktování vhodných odborníků na hloubkové rozhovory.

## TEORETICKÁ ČÁST

Statistická šetření jsou základem každého zkoumání a jsou tedy jeho nedílnou součástí. Ve své podstatě je statistické zjišťování forma získávání a zpracování údajů o určitých vlastnostech či znacích daného zkoumaného souboru osob. Teprve na základě statistických šetření lze provádět výběry, úsudky či jakékoli analýzy a závěry. I statistická šetření však mají svá pravidla.<sup>1</sup>

Nejprve je nutné vymezit tzv. předmět zkoumání. Tím se rozumí statistický soubor, na němž bude zkoumání prováděno. Může se jednat například o ženy na mateřské dovolené ve věku 20–30 let, u kterých se poté dále bude zkoumat oblíbenost určitých televizních pořadů či periodik atd. Poté je nutné vymezit obsah zkoumání. Tím se zároveň vymezí i statistické znaky, které budou předmětem šetření. Může to být například sledovanost určitého pořadu, čtenost určitého periodika či preference nějaké značky.<sup>2</sup>

Až na základě provedení těchto fází probíhá samostatné šetření a poté i analýza získaných dat. To, jak snadno či jak těžko se výsledná data zpracovávají, se odvíjí od způsobu prováděného výběrového šetření.<sup>3</sup>

Nejnákladnější a zároveň nejpodrobnější statistická šetření se nazývají úplná statistická šetření. V případě tohoto druhu šetření se prošetří všechny jednotky v celém souboru.<sup>4</sup> Příkladem vyčerpávajícího statistického šetření je například Sčítání lidu, domů a bytů (SLDB). To se koná vždy jednou za 10 let a v rámci tohoto zjišťování se prošetří všichni občané České republiky. Kromě SLDB ale probíhají i jiná úplná statistická zjišťování, a to například zjišťování zemědělských podniků, tzv. Agrocensus, či soupis maloobchodních prodejen atd.<sup>5</sup>

---

<sup>1</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 1.* 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 5-6. ISBN 80-707-9191-8.

<sup>2</sup> Tamtéž, s. 5-6.

<sup>3</sup> LÖSTER, Tomáš, Hana ŘEZANKOVÁ a Jitka LANGHAMROVÁ. *Statistické metody a demografie.* 1. vyd. Praha: VŠEM, 2008, s. 10 - 12. ISBN 978-80-86730-40-0.

<sup>4</sup> SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru.* Olomouc, 2011, s. 6-7. Diplomová práce. Univerzita Palackého v Olomouci.

<sup>5</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 1.* 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 5-6. ISBN 80-707-9191-8.

Má práce ale není zaměřena na úplná statistická zjišťování, ale naopak na neúplná, či výběrová statistická zjišťování. To jsou taková, u nichž se předem stanoví, že se prošetří pouze malá část ze všech jednotek. V tomto případě se tedy z celé populace ČR vybere pouze určité procento osob a statistická šetření probíhají pouze na těchto osobách. Výběr tohoto souboru je však velmi složitou procedurou. Zároveň v současné době existuje i mnoho možností, jak tento soubor vytvořit.<sup>6</sup>

## 1 VÝBĚROVÁ ŠETŘENÍ

Výběrové šetření je druh neúplného statistického zjišťování. Patří mezi nejdůležitější druhy výběrového šetření.

Výběrová šetření se zakládají na tom, že nejsou prováděna na celkové populaci, ale naopak se z této populace vybere předem určený počet jednotek a na tom se pak provádí toto šetření. Z výsledků, které se získají na základě takto vybraných dat, se pak usuzuje výsledek celého původního souboru, ačkoli do šetření nebyly zahrnuty všechny jednotky z tohoto souboru.<sup>7</sup>

Tato výběrová šetření se dělí na pravděpodobnostní a nepravděpodobnostní výběrové šetření. Pravděpodobnostní výběr je charakterizován tím, že jednotky vybírané ze základního souboru jsou vybírány zcela náhodně a nezávisle na sobě i na rozhodnutí zadavatele. Pravděpodobnostní šetření mají několik postupů a metod, jak dosáhnout náhodného výběru jednotek.<sup>8</sup>

Naproti tomu nepravděpodobnostní výběr využívá rozhodnutí zadavatele či odborného pracovníka, který sám určí, jaké jednotky se do souboru zařadí. I u tohoto typu výběru existuje několik postupů a metod, jak takové výběry provádět.<sup>9</sup>

---

<sup>6</sup> SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru*. Olomouc, 2011, s. 6-7. Diplomová práce. Univerzita Palackého v Olomouci.

<sup>7</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 7. ISBN 80-245-0753-6.

<sup>8</sup> SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru*. Olomouc, 2011, s. 6-8. Diplomová práce. Univerzita Palackého v Olomouci.

<sup>9</sup> KREPL, Jan. *Výběr z konečných populací v ekonomických úlohách*. Praha, 2014, s. 6. Bakalářská práce. Univerzita Karlova v Praze.

Při provádění výběrového šetření musíme nejprve určit základní soubor. Tento soubor tvoří všechny jednotky, tedy například celá populace ČR. V případě provádění výběrového šetření pouze v nějaké firmě ale tvoří základní soubor pouze zaměstnanci této firmy, atd.<sup>10</sup>

Pokud bude výběr prováděn z tohoto základního souboru, je nutné si nejdříve opatřit tzv. oporu výběru. To může být například jmenný seznam všech osob v základním, souboru, tedy zaměstnanců ve firmě, či jmenný seznam všech osob v ČR. Získání opory výběru je mnohdy velmi náročná činnost a někdy se ani nepovede tuto oporu získat celou.<sup>11</sup>

Z této opory se poté vytvoří výběrový soubor, na kterém je pak prováděno ono zjišťování. Právě tento výběrový soubor se získává pomocí pravděpodobnostního či nepravděpodobnostního výběru.<sup>12</sup>

Často se stává, že při výběrovém statistickém zjišťování nebyly prošetřeny některé jednotky, se kterými se ale původně počítalo. Například pomocí pravděpodobnostního výběru byli vybráni určití zaměstnanci podniku, aby se vyjádřili, ale ti odmítli. V tomto případě se jedná o zjišťování, jež je v určitém směru vadné. Někteří autoři však toto zjišťování mylně označují jako neúplné.<sup>13</sup>

---

<sup>10</sup> KAFKOVÁ, Pavla. *Statistika a její využití (zneužití) při tvorbě informací*. Jindřichův Hradec, 2013, s. 21 - 22. Diplomová práce. Vysoká škola ekonomická.

<sup>11</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 13. ISBN 80-245-0753-6.

<sup>12</sup> KAFKOVÁ, Pavla. *Statistika a její využití (zneužití) při tvorbě informací*. Jindřichův Hradec, 2013, s. 21-22. Diplomová práce. Vysoká škola ekonomická.

<sup>13</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 11. ISBN 04-326-80.

## 2 PRAVDĚPODOBNOSTNÍ VÝBĚR

Pravděpodobnostní výběr se využívá zejména v případech, kdy by úplné zjišťování bylo abnormálně neefektivní, či nešlo vůbec provést. Důležitým hlediskem je zde také čas. Pokud by úplné zjišťování trvalo příliš dlouho, je vhodnější použít právě neúplná zjišťování.<sup>14</sup>

Mají samozřejmě i jisté nevýhody, spojené především s dodatečnými náklady na náročnější přípravu a na zesílení kontroly pro eliminaci případných chyb. Svou roli zde hrají také složitější postupy při vyhodnocování získaných dat. Výsledné odhady se také nekryjí zcela přesně se skutečnými hodnotami zkoumaného souboru. Je pak nutné zajistit snížený výběrové chyby.<sup>15</sup>

Tyto výběry jsou založeny na principu, že u každé jednotky ze základního souboru je dopředu již známé, s jakou pravděpodobností bude zahrnuta do výběrového souboru. To ale znamená, že zde musí být znalost základního souboru a každého jeho prvku.<sup>16</sup>

Předpoklad, že pravděpodobnosti v základním souboru jsou stejné a nám známé, je pro úspěch těchto výběrů klíčový. V případě, že by toto pravidlo neplatilo, nemohl by tento výběr být dobrým podkladem a nezajišťoval by ani reprezentativnost vzorku. Na základě toto by se nedaly dělat závěry a úsudky pro celek.<sup>17</sup>

Pravděpodobnostní výběry se užívají zejména pro svou spolehlivost, kterou lze snadno ověřit. Pravděpodobnostní výběr totiž musí splňovat některé základní parametry.

Tím prvním je tzv. konzistentnost. To znamená, že odhad musí v případě zvyšování rozsahu výběru konvergovat skutečné hodnotě. Tedy že čím větší je rozsah výběru, tím více se odhad blíží skutečné reálné hodnotě.<sup>18</sup>

Druhým předpokladem je tzv. nevychýlenost. To znamená, že při každém rozsahu výběru není průměrná hodnota ani nadhodnocována, ale ani podhodnocována.<sup>19</sup>

---

<sup>14</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 14. ISBN 04-326-80.

<sup>15</sup> Tamtéž, s. 14.

<sup>16</sup> Tamtéž, s. 14.

<sup>17</sup> Tamtéž, s. 19.

<sup>18</sup> ČERMÁK, Václav. *Vybrané kapitoly z teorie výběrových šetření*. 1. vyd. Praha: Státní pedagogické nakladatelství, 1974, s. 5. ISBN 17-131-74.

Třetím předpokladem je fakt, že přesnost těchto odhadů lze změřit. A to například pomocí výběrové chyby, či pomocí intervalových odhadů určit interval, ve kterém se téměř jistě nachází skutečná hodnota.<sup>20</sup>

## 2.1 Prostý náhodný výběr

Prostý náhodný výběr patří mezi nejznámější metodu pravděpodobnostních výběrů. Jedná se o výběr jednotek, které mají všechny stejnou pravděpodobnost, že budou vybrány do vzorku. Tato metoda je nejlepší volbou v případě teoretických úvah, například o populačních charakteristikách, ale i o velikosti výběrových chyb. Velkou nevýhodou prostého náhodného výběru je však jeho poměrně velká nákladnost. Velké náklady spojené s tímto druhem výběru jsou zapříčiněny zejména složitými výběrovými postupy, které jsou ale v tomto případě často nutností.<sup>21</sup>

V případě tohoto výběru také záleží na tom, zda pravděpodobnost výběru jednotek je stejná, či není stejná.<sup>22</sup>

V případě, že je stejná pravděpodobnost výběru jednotek do výběrového souboru, je nejjednodušší technikou výběru losování pořadových čísel. K tomu lze použít například generátor náhodných čísel. Pokud je navíc v souboru vyloučeno jakékoli systematické uspořádání, je možné zvolit tzv. systematický výběr, který je nejméně náročný.<sup>23</sup>

Pokud ale nastane případ, kdy pravděpodobnost výběru jednotek do výběrového souboru není stejná, nelze tento postup použít. V tomto případě jsou totiž různé pravděpodobnosti zapříčiněny například různou velikostí vybíraných jednotek, nebo

---

<sup>19</sup> ČERMÁK, Václav. *Vybrané kapitoly z teorie výběrových šetření*. 1. vyd. Praha: Státní pedagogické nakladatelství, 1974, s. 5. ISBN 17-131-74.

<sup>20</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. Praha: 1. vyd. SNTL, 1980, s. 19 - 20. ISBN 04-326-80.

<sup>21</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1. vyd. Praha: VŠE, 1995, s. 11. ISBN 80-7079-737-1.

<sup>22</sup> ČERMÁK, Václav. *Vybrané kapitoly z teorie výběrových šetření*. 1. vyd. Praha: Státní pedagogické nakladatelství, 1974, s. 6-7. ISBN 17-131-74.

<sup>23</sup> SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru*. Olomouc, 2011, s. 8. Diplomová práce. Univerzita Palackého v Olomouci.

obecně nějakou doplňující informací, na základě které jsou tyto pravděpodobnosti stanoveny. Tento případ je ale ještě komplikovanější, a proto se v praxi příliš neužívá.<sup>24</sup>

Důležité je také zmínit, zda se jedná o výběr s vracením či bez vracení. Tedy jestli zde existuje možnost, že vybraná jednotka bude znovu zahrnuta do souboru a bude tedy moci být znovu vybrána.<sup>25</sup>

V případě výběru s vracením jsou jednotlivé tahy zcela nezávislými pokusy. Všechny jednotky mají zároveň stejnou pravděpodobnost, že budou vytaženy. Vzhledem k tomu, že se jedná o výběr s vracením, se základní soubor nijak nemění a lze z něj vybírat prakticky do nekonečna. Tento postup je vhodné využít zejména v případě teoretických úvah.<sup>26</sup>

V případě výběru bez vracení je ale situace jiná. Jednotlivé tahy se v tomto případě stávají závislými pokusy a pravděpodobnost, zda bude jednotka vybrána, se mění. Tato pravděpodobnost se zvyšuje s každým dalším tahem. Zároveň se s každým dalším tahem zmenšuje i základní soubor.<sup>27</sup>

### **Techniky náhodného výběru**

Náhodný výběr lze provést pomocí několika technik. Použití jednotlivých technik závisí na rozsahu opory výběru a na tom, zda je opora výběru k dispozici či ne.<sup>28</sup>

Jednou z metod náhodného výběru je losování. Patří mezi nejjednodušší techniky, jakým se náhodný výběr provádí. Používá se zejména v případech, kdy pravděpodobnost výběru jednotlivých jednotek v souboru je stejná. Z toho vyplývá, že v případě losování je nutné mít k dispozici oporu výběru. Losování probíhá velmi jednoduše. Všechny jednotky, či jejich zástupci (například pořadová čísla) se promíchají a následně se vylosuje požadovaný počet jednotek. Použít jej lze jak v případech výběru s vracením, tak i v případě výběru bez vracení. Nejedná se však

---

<sup>24</sup> SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru*. Olomouc, 2011, s. 8. Diplomová práce. Univerzita Palackého v Olomouci.

<sup>25</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 11. ISBN 80-7079-737-1.

<sup>26</sup> ČERMÁK, Václav. *Vybrané kapitoly z teorie výběrových šetření*. 1. vyd. Praha: Státní pedagogické nakladatelství, 1974, s. 5-8. ISBN 17-131-74.

<sup>27</sup> Tamtéž, s. 5-8.

<sup>28</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 12. ISBN 80-7079-737-1.



o univerzální metodu náhodného výběru. Zvláště v případě rozsáhlých souborů je provedení této metody téměř nemožné.<sup>29</sup>

Další metodou je tzv. výběr pomocí náhodných čísel. Ten se používá v případě, kdy pravděpodobnost vybrání jednotek v souboru není stejná. I zde je potřeba mít k dispozici oporu výběru. Poté se každé jednotce, která se nachází v základním souboru, přiřadí pořadové číslo. Pomocí generátoru náhodných čísel je vytažen potřebný počet jednotek. Ty jsou poté zařazeny do výběrového souboru. Tato metoda je jednodušší zejména v případě rozsáhlých základních souborů, kde metoda losování selhává. Je ale stále značně pracná a proto se využívá spíše systematický výběr.<sup>30</sup>

Systematický výběr se od předchozích typů značně liší. Na rozdíl od předchozích metod si umí velmi dobře poradit s rozsáhlými soubory. Navíc pro provedení této metody není potřeba opora výběru. Podmínkou je pouze, aby jednotky základního souboru byly řazeny zcela objektivně, tedy podle žádného klíče. Poté se vylosuje první jednotka a od ní se postupuje dále, přičemž se vybírá vždy k-tá jednotka od náhodně vylosované.<sup>31</sup>

## 2.2. Oblastní výběr

Náhodný výběr má ale i svá složitější uspořádání, a to například oblastní výběr. A ten funguje následovně.

Základní soubor je nejprve nutné rozdělit do několika skupin (tzv. straty). Tyto skupiny by měly mít shodnou vnitroskupinovou variabilitu, což znamená, že uvnitř by skupina měla obsahovat jednotky, které jsou si z nějakého hlediska velmi podobné. To znamená, že skupina je homogenní. Tyto skupiny si ale podobné navzájem nejsou, a jsou tedy heterogenní. Na základě toho je tedy mnohem významnější tzv. meziskupinová variabilita.<sup>32</sup>

---

<sup>29</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část I.* 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 27. ISBN 80-707-9191-8.

<sup>30</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I.* 1. vyd. Praha: VŠE, 1995, s. 12. ISBN 80-7079-737-1.

<sup>31</sup> Tamtéž, s. 12.

<sup>32</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I.* 1. vyd. Praha: VŠE, 1995, s. 11. ISBN 80-7079-737-1.

Poté se provede náhodný výběr z každé skupiny. Nejčastější je tzv. proporcionální výběr. To znamená, že výběrové rozsahy jednotlivých skupin jsou přímo úměrné velikostem daných oblastí. Konečný soubor je vytvořen všemi prvky, které byly ze skupin vybrány. V praxi lze použít výběr se stejnými, ale i nestejnými pravděpodobnostmi. Popřípadě lze použít i kombinaci obojího. Jedná se ale o složitější metody, které se příliš neuplatňují.

Na základě tohoto postupu lze dosáhnout mnohem lepších úsudků, než při použití prostého náhodného výběru.<sup>33</sup>

V praxi jsou tyto skupiny tvořeny například věkovými skupinami, kdy v jedné skupině jsou osoby např. 15–20 let a ve druhé skupině osoby 30–35 let. Osoby uvnitř skupiny se příliš neliší, ale naopak skupiny navzájem se liší výrazně. Procento vybraných prvků se v praxi často odvíjí od dostupnosti dat či finančních nákladů.<sup>34</sup>

Cílem vytvoření těchto skupin je zamezení vzniku výběrových skupin, které by mohly odhad silně nadhodnotit či podhodnotit. Pomocí oblastního výběru se tak minimalizuje velikost výběrových chyb.<sup>35</sup>

Využívání oblastního výběru má několik odůvodnění. Nejčastěji se stává, že zkoumaný znak u určitých jednotek kolísá mnohem méně, než je tomu v celém základním souboru. Rozdělí se tedy soubor na několik skupin dle sledovaného znaku a v každém se provede již zmíněný výběr. Dalším důvodem však mohou být i finance. Snaha o vytvoření souboru, jehož prozkoumání bude méně nákladné.<sup>36</sup>

## 2.3. Skupinový výběr

V případě, že je základní soubor příliš rozsáhlý a čítá například až miliony osob, nelze užít prostý náhodný výběr. V tomto případě se užívá skupinový výběr. Jeho použití má dvě základní příčiny. První příčinou použití skupinového výběru je velká rozptýlenost

---

<sup>33</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 3*. 1. vyd. Praha: Vysoká škola ekonomická, 1998-1999, s. 5-8. ISBN 80-245-0003-53.

<sup>34</sup> Tamtéž, s. 5-8.

<sup>35</sup> Tamtéž, s. 5-8.

<sup>36</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 3*. 1. vyd. Praha: Vysoká škola ekonomická, 1998-1999, 5 - 8. ISBN 80-245-0003-53.

jednotek. Na prošetření všech jednotek by bylo nutné vynaložit velkou sumu peněz a mnoho času, a tak se využívá skupinový výběr. Druhou příčinou je neexistence evidence jednotek, na základě které by bylo možné vytvořit oporu výběru. V praxi je například k dispozici pouze seznam tříd určité základní školy, ale nikoli jména žáků. Stejně tak v případě obyvatelstva není k dispozici seznam skutečně bydlících osob na určité adrese, ale pouze seznam domů či bytů.

Skupinový výběr se provádí podle následujících pravidel.<sup>37</sup>

Důležité je zajistit, aby skupiny byly pokud možno stejně velké, či se ve velikosti lišily pouze minimálně. Podstatným faktorem je také vnitřní různorodost skupiny, ale vzájemná podobnost skupin mezi sebou. Také variabilita mezi skupinami musí být co nejmenší, což je přesně naopak, než jak tomu bylo u oblastního výběru. Poté, co jsou tyto parametry splněny, se náhodně vybírá určitý počet skupin, ve kterém se poté prošetří všechny jednotky.<sup>38</sup>

V praxi se takovými skupinami myslí například školní třída, rodina, zemědělský závod, ale i obce či okresy. Z formálního hlediska lze ale za skupiny považovat i úseky časových řad, jakožto skupiny navzájem na sebe navazujících období. Podstatné je, aby navzájem byly tyto vybrané skupiny přibližně stejně velké. Tento postup v praxi významně snižuje náklady spojené s šetřením a snižuje i časovou náročnost na provedení celého šetření. Finanční úspory je pak možné věnovat například na rozšíření celého výběru o další skupiny. Tento postup se používá v rámci nahrazení ztrát na přesnosti odhadů. Paradoxně ale tento postup nepřináší nijak výrazně lepší výsledky.<sup>39</sup>

## 2.4. Vícetupňový výběr

Vícetupňový výběr se používá zejména v případech, kdy není možné získat úplnou oporu pro provedení výběru. Základní soubor může i v tomto případě, stejně jako u skupinového výběru, čítat až miliony jednotek. V tomto případě tedy není možné vybrat několik tisíc jednotek pomocí náhodného či oblastního výběru. Kromě

---

<sup>37</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 3.* 1. vyd. Praha: Vysoká škola ekonomická, 1998-1999, 5 - 8. ISBN 80-245-0003-53.

<sup>38</sup> Tamtéž, s. 32 – 33.

<sup>39</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 3.* 1. vyd. Praha: Vysoká škola ekonomická, 1998-1999, s. 31-33. ISBN 80-245-0003-53.

nedostatečné opory by tento výběr stál vysokou částku peněz. Nelze použít ani skupinový výběr, jelikož například skupiny bytů by byly příliš malé a bylo by jich stále mnoho a skupiny krajů či okresu zase naopak moc velké a příliš variabilní. Technika vícestupňového výběru je založena na principu, že všechny prvky jsou uspořádány dle určitého hierarchického popisu. Každý prvek základního souboru je tedy možné na základě toho nějak popsat a zařadit do příslušné skupiny. V praxi může být takovou skupinou například město, které se poté dělí na bloky, jednotlivé domy a poté i na jednotlivé domácnosti.<sup>40</sup>

Po seřazení jednotek do skupin probíhá postupně na každém stupni náhodný výběr skupin a v posledním stupni náhodný výběr jednotek. Teprve tyto náhodně vybrané jednotky se prošetřují. Díky tomu lze i z dost velkých skupin vybrat poměrně malý počet prvků.<sup>41</sup>

Velkou výhodou tohoto typu výběru je relativní ekonomická nenáročnost. Například oproti oblastnímu výběru je prostorová rozptýlenost jednotek menší, a tím se snižují i náklady na získání těchto dat.<sup>42</sup>

Zvláštním případem vícestupňového výběru je tzv. výběr skupin. Ten probíhá ve dvou stupních a má jiná pravidla. V prvním stupni se ze základního souboru vyberou skupiny jednotek pomocí náhodného výběru. Ve druhém stupni se ve vybraných skupinách prošetří všechny jednotky, na rozdíl od předchozího postupu. Rozdílem oproti klasickému vícestupňovému výběru je ještě vyšší prostorová koncentrace jednotek a tím i menší náklady na získání dat.<sup>43</sup>

---

<sup>40</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 307. ISBN 04-326-80.

<sup>41</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část I*. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 23. ISBN 80-707-9191-8.

<sup>42</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 307. ISBN 04-326-80.

<sup>43</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1. vyd. Praha: VŠE, 1995, s. 13. ISBN 80-7079-737-1.

### 3 NEPRAVDĚPODOBNOSTNÍ VÝBĚR

Nepravděpodobnostní výběr se od pravděpodobnostního zásadně liší. Předně je zde důležitý fakt, že v případě nepravděpodobnostního výběru se zde zásah nějakého výzkumného pracovníka do procesu vybírání respondentů či jednotek. Proces selekce tedy neprobíhá čistě náhodně a vzhledem k tomu, že jsou některé jednotky preferované před ostatními je velmi obtížné, prakticky nemožné, zajistit reprezentativnost jakéhokoli vzorku. Problémem může být také obecně malé návratnost v rámci těchto metod.<sup>44</sup>

Pro používání nepravděpodobnostních metod při výzkumu je často hlavním důvodem časová nenáročnost a finanční úspora. Výzkumný pracovník zde má možnost ovlivnit místo i čas výběru respondentů, má možnost vybrat dokonce samotné respondenty a i počet respondentů, které do výzkumu zahrne. Na základě toho má možnost ovlivnit i finanční prostředky, na kterých je výzkum závislý.<sup>45</sup>

Výzkumný pracovník tak nejčastěji volí, dle něj nejvhodnější, místo a čas pro sběr dat, což má ale přímý důsledek na nereprezentativnost. Nikdy totiž není zaručeno, že respondenti, kteří se nachází v určitý čas na určitém místě, představují ucelený průřez kompletní strukturou obyvatelstva.<sup>46</sup>

Může dát také přednost jedné vlastnosti respondenta před jinou. Například rozlišuje mezi blondýnami a brunetami. Čímž opět selektuje na základě jednoho sledovaného znaku, ale nezaručuje si tím reprezentativnost vzorku.

V rámci nepravděpodobnostních výběrů však nemusí jít za každou cenu o snahu vytvořit reprezentativní vzorek. Cílem může být i soupis obyvatel s určitými charakteristikami, jako například zálibou v nějakém sportu či příslušností k nějakému náboženství.<sup>47</sup>

---

<sup>44</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 13. ISBN 80-7079-737-1.

<sup>45</sup> ŠKRAŇKA, Ivan. *Statistické metody v marketingovém výzkumu*. Brno, 2009, s. 25. Diplomová práce. Vysoké učení technické v Brně.

<sup>46</sup> VÝTISK, Radek. *Pokročilé statistické metody v marketingovém výzkumu*. Brno, 2008, s. 17-18. Diplomová práce. Masarykova univerzita.

<sup>47</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 11-19. ISBN 04-326-80.

Metod nepravděpodobnostního výběru je několik. Každá metoda má svůj způsob zpracování i praktické využití v mnoha situacích. Každá je něčím specifická, a tudíž jsou všechny tyto metody využitelné. A to i v případě, že nejsou založeny na reprezentativním vzorku.<sup>48</sup>

### 3.1. Anketa

Anketa patří k nejznámějším typům výběru. Metoda sběru dat je velmi jednoduchá. Určitému okruhu osob, například v nějaké instituci, se rozešlou dotazníky se sestavenými otázkami a žádostí o vyplnění. Následně se vybrané odpovědi zanalyzují a vytvoří se nějaký závěr. Ten je ale zcela nereprezentativní, a to hned z několika důvodů.<sup>49</sup>

Ankety mají všeobecně malou návratnost, pohybující se okolo 30 % a to i v případech, kdy je dotazník posílán adresně přímo vybraným osobám. Již z tohoto pohledu je zřejmé, že charakteristiky, které se vypočítají z obdržených odpovědí, nebudou odpovídat realitě. Jakékoli průměry či poměry tedy nelze považovat za obecně platné a není možné je zobecnit na celou populaci. Důležité je si uvědomit, že i mezi odpovědi a odmítnutím a dotazovanou skutečností bývá určitý vztah. Například osoby s vysokým příjmem neradi odpovídají na výši příjmu. Může se jednat i o utajované příjmy, ke kterým se osoba nechce přiznávat nebo obecně nechce přiznat své příjmy, například z důvodu nějakých obav. S větší finanční částkou je spojena například i záliba ve starožitných předmětech, a tudíž i tito lidé svůj koníček nepřiznají, a to právě kvůli finančním částkám.<sup>50</sup>

I přes zjevný fakt, že závěry získané touto metodou nemají vypovídající hodnotu, dají se přesto využít. Určitý význam ankety může spočívat například v zajímavosti odpovědí. Velmi často jsou totiž požádáni například slavné osobnosti, či odborníci daných

---

<sup>48</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 11-19. ISBN 04-326-80.

<sup>49</sup> SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru*. Olomouc, 2011, s. 7. Diplomová práce. Univerzita Palackého v Olomouci.

<sup>50</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 1*. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 7. ISBN 80-707-9191-8.

problematik, aby anketu též vyplnili. Tyto odpovědi se dají zkoumat z kvalitativního hlediska.<sup>51</sup>

Ankety jsou také velmi často využívány v rámci populárních médií, jako například časopisy či noviny. Takovou klasickou anketou pak může být například Zlatá hokejka či Zlatý slavík. Může se ale jednat i o anketu zveřejněnou v rámci specificky zaměřeného časopisu, například právě o starožitnostech.<sup>52</sup>

### 3.2. Snowball

Snowball<sup>53</sup>, neboli metoda sněhové koule, se využívá zejména v případech, kdy není možné vytvořit soupis celkové populace či by bylo vytvoření soupisu velmi nákladné, nebo je tato populace neznámá.<sup>54</sup>

Tato metoda funguje na principu, kdy se výběrový soubor vytváří postupně. Již jednou vybraná jednotka navrhne další jednotku. Ta poté navrhne další atd. Tento proces trvá až do okamžiku, kdy všechny odkazy již vedou k jednou dotázaným jednotkám. Tedy až do okamžiku, kdy je výběrový soubor teoreticky nasycen.<sup>55</sup>

V rámci této metody je velmi vhodné využít i prostředky internetu. Díky tomu je možné dosáhnout i na jinak těžko dosažitelné populace, zejména při výzkumech ohledně víry či sexuální orientace.<sup>56</sup>

Tato metoda jako jediná z pravděpodobnostních i nepravděpodobnostních metod dokáže těžit z komplikované a propletené sítě vztahů, což ji činí nepostradatelnou v rámci specifických výzkumů.<sup>57</sup>

---

<sup>51</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 11-19. ISBN 04-326-80.

<sup>52</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 16. ISBN 80-7079-737-1.

<sup>53</sup> Poprvé použita Colemanem v r. 1958

<sup>54</sup> PITOŇÁK, Michal. *Queer geografie Česka: Heteronormativita prostoru*. Praha, 2011s. 44-47. Diplomová práce. Univerzita Karlova v Praze.

<sup>55</sup> KREPL, Jan. *Výběr z konečných populací v ekonomických úlohách*. Praha, 2014, s. 8. Bakalářská práce. Univerzita Karlova v Praze.

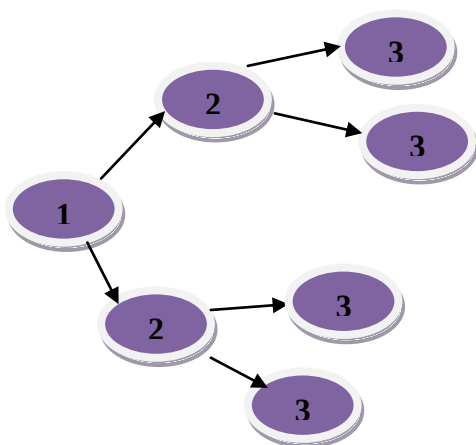
<sup>56</sup> PITOŇÁK, Michal. *Queer geografie Česka: Heteronormativita prostoru*. Praha, 2011s. 44-47. Diplomová práce. Univerzita Karlova v Praze.

<sup>57</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 15. ISBN 80-7079-737-1.

V praxi se snowball využívá v případech, kdy je potřeba oslovit respondenty se specifickým zájmem či například specifického náboženství. Soupisy těchto osob nejsou nikde k dostání, a tudíž by bylo velmi nákladné tyto soupisy vytvářet běžnou cestou. Místo toho je zvolen postup, kdy je nutné znát alespoň jednu osobu, která daná kritéria splňuje. Tedy například se hlásí k jehovistům. Tato osoba pak doporučí další osobu, která se také hlásí k jehovistům a takto pokračuje řetězec až do chvíle, kdy je dostatečný počet respondentů pro výzkum. Na tomto vytvořeném vzorku je pak aplikována výzkumná metoda.<sup>58</sup>

Vzhledem ke skutečnosti, že se jedná o nepravděpodobnostní výběr, je zde samozřejmě silně zpochybněna reprezentativnost takového výsledku. Výsledky se ale dají zobecnit na danou populaci osob, na které byl výzkum vymezen. Dá se tedy na základě toho usuzovat, jaké charakteristiky má tato úzce vymezená populace, nikoli však obecně celá populace ČR.<sup>59</sup>

Obrázek 1: Snowball



Zdroj<sup>60</sup>

Obrázek 1 znázorňuje průběh snowballu, kdy jeden respondent doporučuje dalšího.

<sup>58</sup> PITOŇÁK, Michal. *Queer geografie Česka: Heteronormativita prostoru*. Praha, 2011s. 44-47. Diplomová práce. Univerzita Karlova v Praze.

<sup>59</sup> Tamtéž, s. 44-47.

<sup>60</sup> Vlastní zpracování.



### 3.3. Úsudkové výběry

Na rozdíl od anket, které jsou zařazeny do běžných periodik jako zpestření pro čtenáře či posluchače, úsudkové výběry provádí zkušený pracovník. Ten sám určí, jaké jednotky se průzkumu zúčastní a jaké nikoli. Jednotky vybírá dle předem dohodnutého scénáře tak, aby vybraní respondenti ve svém souhrnu umožnili provést zamýšlené zkoumání.<sup>61</sup>

Počet jednotek, které jsou do průzkumu zahrnuty, je zpravidla dán předem, nebo je určen například finančními možnostmi instituce, jež výběr provádí. Kritéria, dle kterých jsou respondenti osloveni, jsou velmi subjektivní.<sup>62</sup>

U úsudkového výběru je zpravidla kladen požadavek na reprezentativnost výběrového souboru, což ale není jednoduché splnit. Zvláště v tomto případě, kdy jednotky oslovuje a vybírá pracovník, je zabezpečení reprezentativnosti obtížné.<sup>63</sup>

Mezi úsudkové soubory patří tzv. typický výběr a kvótní výběr.<sup>64</sup>

V případě typického výběru volí odborný pracovník takové jednotky, o kterých se domnívá, že jsou pro daný soubor typické. To znamená, že hodnoty, které u nich bude sledovat (například věk), jsou blízké průměru nebo modu ostatních jednotek.<sup>65</sup>

Další variantou úsudkového výběru je kvótní výběr neboli metoda dokonalého průřezu. V rámci této metody je důležité zajistit, aby se shodovala struktura souboru vzhledem k několika podstatným znakům. Těmito znaky mohou být například věk či pohlaví. Podle těchto znaků je soubor roztríděn do několika skupin. Počet osob vybraných do skupiny je přímo úměrný velikosti této skupiny. O tom, jaká jednotka bude zařazena do skupiny, rozhoduje opět zkušený pracovník. Důležité jsou zde instrukce, které dostane.

---

<sup>61</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 13. ISBN 80-7079-737-1.

<sup>62</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část I*. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 8-9. ISBN 80-707-9191-8.

<sup>63</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část I*. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 8-9. ISBN 80-707-9191-8.

<sup>64</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 13. ISBN 80-7079-737-1.

<sup>65</sup> ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část I*. 1. vyd. Praha: Vysoká škola ekonomická v Praze, 1999, s. 8-9. ISBN 80-707-9191-8.

Jsou zde také instalovány určité mechanismy, aby se eliminovala subjektivita pracovníka. I přesto však není zaručena shoda v dalších znacích (například výše platu, bydliště atd.). Tento způsob je díky svému provedení nejlevnější a zároveň je velmi rychlý a jednoduchý.<sup>66</sup>

---

<sup>66</sup> Tamtéž, s. 8-9.

## 4 METODY SBĚRU DAT

Postupů pro sběr dat je mnoho. Rozhodnutí, jakou metodu sběru dat je vhodné použít, závisí na více faktorech. Například zda tyto jevy jsou nebo nejsou přímo pozorovatelné, zda jsou vyvolané uměle, anebo jsou zcela spontánní, ale také záleží na jejich relativní stabilitě či naopak dynamice.<sup>67</sup>

Sběr dat probíhá nejčastěji dvěma základními způsoby. Jeden ze způsobů, jak získat data, je **pozorování**. V tomto případě je zkoumána většinou existence a bytí. Hlavní výhodou této metody je, že poskytuje přesné údaje, které jsou navíc objektivní. A to díky tomu, že tato data jsou získávána přímo pomocí smyslového vnímání a ne pomocí zprostředkovatele, například prostřednictvím vyjádření jiné osoby. Aby však bylo dosaženo přesných a objektivních údajů, musí být pozorovaný jev jednoznačně vymezen a projevy tohoto jevu musí být zaznamatelné. Výsledkem pozorování je pak zjištění, zda tento jev nastal či nenastal. Lze také vyjádřit počet, kolikrát jev nastal anebo nenastal.<sup>68</sup>

**Př.:** Pomocí pozorování lze zkoumat například frekvenci kliknutí na reklamní banner na internetu. Zde lze ideálně aplikovat, zda jev nastal (uživatel klikl na reklamní banner), či nenastal (uživatel neklikl na reklamní banner). A případně lze dále uvést, kolikrát uživatel na tento banner klikl atd.

V případě pozorování jevů, které jsou složitější, lze jejich pozorování zjednodušit pomocí kategorií, do nichž může jev patřit.<sup>69</sup>

**Př.:** V tomto případě lze zkoumat například reakci na reklamu v kině a to pomocí kategorií reakcí, která slouží pro zmíněné zjednodušení. Mezi tyto kategorie reakcí může patřit například – žádná reakce, smích, potlesk, hvízdání, atd.

Pozorování samotných jevů může probíhat různými způsoby, z nichž každá může přinést odlišné výsledky. Jedním z nich je pozorování jevu souběžně s tím, jak jev

---

<sup>67</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 25. ISBN 978-80-86946-74-0.

<sup>68</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 68. ISBN 80-7226-917-8.

<sup>69</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 25. ISBN 978-80-86946-74-0.

probíhá. Druhá možnost je pozorovat jev s odstupem, kdy se zjišťují následky. Pozorování také může probíhat skrytě, anebo naopak zjevně. V případě zjevného pozorování osoby dopředu vědí, že jsou pod dohledem pozorovatele. V případě skryté varianty nikoli.<sup>70</sup>

Nejčastější formou sběru dat v terénu je ale **dotazování**. V tomto případě je zkoumáno vědomí.<sup>71</sup> Dotazování je nejpoužívanější formou, jelikož šetření, které probíhají u různých osob, většinou nejsou přímo pozorovatelná. A právě z toho důvodu se používá dotazník. V tomto případě jsou údaje získávány pomocí cíleně vytvořených otázek, aby tazatel získal co nejpresnější odpověď na zkoumaný jev. Dotazník může být veden tváří v tvář, ale také písemně či telefonicky. Forma dotazníku hraje velkou roli v oblasti kvality získaných informací. Stejně velkou roli také zastává osobnost tazatele.<sup>72</sup>

Existují samozřejmě i jiné metody sběru dat. Ty ale vždy vychází z těchto dvou základních metod.<sup>73</sup>

#### 4. 1 Pravidla tvorby dotazníku

Každý dotazník má svá pravidla, aby bylo možné data z něj analyzovat tím správným způsobem, a aby špatná konstrukce dotazníku nezapříčinila jeho odmítnutí respondentem, či nepochopení a špatné vyplnění.<sup>74</sup>

Typů dotazníků je několik. Jedním z velmi používaných typů dotazníků je dotazník **CAPI**<sup>75</sup>. V případě dotazování pomocí CAPI má tazatel u sebe například notebook

---

<sup>70</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 68. ISBN 80-7226-917-8.

<sup>71</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 25. ISBN 978-80-86946-74-0.

<sup>72</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 84-86. ISBN 80-7226-917-8.

<sup>73</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 24-29. ISBN 80-245-0753-6.

<sup>74</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 15. Bakalářská práce. Západočeská univerzita v Plzni.

<sup>75</sup> ComputerAssistedPersonalInterviewing

a získaná data může ihned odeslat ke zpracování do výzkumné instituce.<sup>76</sup> Zde jsou respondenti vybíráni náhodně. Oporu výběru zde může tvořit například databanka adres bytů v České republice. Vzorek je v tomto případě reprezentativní, jelikož je kladen důraz na zastoupení dle regionu a místa bydliště. Tímto způsobem se zjišťuje například velikost prodaného nákladu titulů a popřípadě i čtenost titulů.<sup>77</sup>

Další variantou dotazníku je dotazník typu **CATI/CAMI**<sup>78</sup>. V tomto případě je dotazník veden po telefonu a otázky tak musí být co nejjednodušší, aby jim respondent na druhé straně rozuměl. Zároveň také tazatel nemůže zabíhat příliš do detailů a zbytečně se tak zdržovat, jelikož tím narůstají náklady na konkrétní šetření.<sup>79</sup> Velkou výhodou této metody je, že umožňuje elektronicky zapsat odpovědi přímo do dotazníku. Otázky v tomto typu jsou standardizované. Tímto způsobem se zjišťuje například poslechovost rozhlasu. Oporou výběru je v tomto případě databáze telefonních linek v ČR. K tomu jsou však často doplněny i osobními rozhovory. Ty slouží k tomu, aby z výzkumu nebyli vyřazeni lidé, kteří nepoužívají žádnou telefonní linku.<sup>80</sup>

Levnější variantou je rozesílání dotazníků poštou. Tímto způsobem lze získat informace od velkého počtu respondentů, a to s výrazně nižšími náklady. Velkým rizikem je ale možnost zkreslení. Osoba, která dotazník vyplňuje, má v tomto případě relativně dost času, tudíž se může poradit s okolím, volí tedy odpovědi co nejpříjemnější, a ne ty dle svého uvážení. Samozřejmě také klesá procento návratnosti dotazníků.<sup>81</sup> Do určité míry mu lze předejít opakovaným urgováním respondenta s žádostí o vyplnění. V pokročilejších verzích může respondent vyplnit takovýto dotazník na počítači, přičemž otázky si může přečíst, nebo je slyší ve sluchátkách. Tato metoda se nazývá **CASI**<sup>82</sup>.

---

<sup>76</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 87. ISBN 80-7226-917-8.

<sup>77</sup> TRAMPOTA, Tomáš a Martina VOJTĚCHOVSKÁ. *Metody výzkumu médií*. 1. vyd. Praha: Portál, 2010, s. 189-190. ISBN 978-80-7367-683-4.

<sup>78</sup> ComputerAssistedTelephone/Mobile Interviewing.

<sup>79</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 87. ISBN 80-7226-917-8.

<sup>80</sup> TRAMPOTA, Tomáš a Martina VOJTĚCHOVSKÁ. *Metody výzkumu médií*. 1. vyd. Praha: Portál, 2010, s. 192-193. ISBN 978-80-7367-683-4.

<sup>81</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1. vyd. Praha: VŠE, 1995, s. 18. ISBN 80-7079-737-1.

<sup>82</sup> ComputerAssistedSelf-Interviewing

Ať už jsou data pro další výzkumy získávána jakoukoli metodou, vždy by měly splňovat podmínku validity a reliability. Podmínka validity znamená, že pomocí dotazníku je skutečně zjišťováno to, co je zamýšleno zjistit. Nejsou zde také systematické chyby, například vážení zboží na neseřazené váze nemůže být považováno za validní. Takto zjištěný údaj naopak nemá žádnou vypovídající hodnotu.<sup>83</sup> Podmínka reliability znamená, že postup použitý ke sběru dat není zatížen náhodnými chybami a není zde zkreslení. Pokud by byl takový postup opakován, mělo by být dosaženo stejných hodnot. Například stejná úroveň znalostí by měla být vždy ohodnocena stejnou známkou či stejným počtem bodů.<sup>84</sup>

Výskytu těchto chyb lze zamezit důkladným vyladěním dotazníku před jeho spuštěním. Důležité je také dotazník pečlivě připravit a dát si o to více záležet, pokud bude dotazník vyplňován respondentem bez pomoci tazatele. V tomto případě velmi pomáhá tzv. pilotní průzkum. Ten spočívá ve vyzkoušení celého procesu nanečisto, nejlépe na menším vzorku respondentů. Díky pilotnímu průzkumu je možné odhalit nedostatky a před oficiálním spuštěním je co nejefektivněji odstranit.<sup>85</sup>

Velkou roli při tvoření dotazníku hraje i forma dotazníku. Důležité jsou i malé detaily, jako velikost či typ písma, použité barvy či nákresy a schémata atd.<sup>86</sup>

Aby bylo dosaženo kvalitně připraveného dotazníku, je nutné odpovědět si na 3 otázky. Zda bude respondent **rozumět** dané otázce, zda bude respondent **schopen** na ni odpovědět a zda bude **ochoten** poskytnout odpověď.<sup>87</sup>

## 4.2 Otázky dle pořadí

Každý dotazník by měl mít přesně stanovené pořadí otázek, jelikož každá otázka má svůj účel. Špatné uspořádání otázek v dotazníku může mít fatální důsledek ve formě nevyplnění dotazníku respondentem.<sup>88</sup>

---

<sup>83</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 103-109. ISBN 80-7226-917-8.

<sup>84</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 14. Bakalářská práce. Západočeská univerzita v Plzni.

<sup>85</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1. vyd. Praha: VŠE, 1995, s. 18. ISBN 80-7079-737-1.

<sup>86</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 103-109. ISBN 80-7226-917-8.

<sup>87</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 23. Bakalářská práce. Západočeská univerzita v Plzni.

Na začátku dotazníku by měly být informace, které respondentovi objasní důvod sběru dat a účel dotazníku. Vhodné je také zdůraznit respondentovi, že jeho účast v šetření je důležitá. Na těchto úvodních informacích v podstatě závisí úspěch dotazníku. Pokud bude mít respondent pocit, že dotazník není nijak důležitý, nebo že není důležité, aby se ho on sám účastnil, nemusí dotazník vůbec vyplnit.<sup>89</sup>

Otázky v dotazníku by na sebe měly volně navazovat, aby udržely pozornost respondenta. Pokud bude dotazník koncipován chaoticky a otázky budou přeskakovat z jednoho tématu na druhé, respondent se začne ztrácet. Pro celkový kontext je tedy nezbytné, aby byl dotazník tvořen logicky.<sup>90</sup>

Velmi prospěšné je také vytvořit tzv. baterii otázek v případě, že respondent bude odpovídat na několik stejně formulovaných otázek a bude mít i stejnou nabídku možností.<sup>91</sup>

**Př.:** Otázka v dotazníku bude zjišťovat frekvenci sledování televizních stanic. Ideální v tomto případě je vytvořit tabulku různých televizních stanic a k nim respondent zaškrtně, zda se dívá pravidelně, nebo občas či vůbec. Pokud by byl respondent tázán na každou televizní stanici zvlášť, mohl by snadno ztratit pozornost a od vyplňování dotazníku upustit.

Tabulka 1: Baterie otázek

| Vyznačte prosím u každé stanice, jak často ji sledujete. | Pravidelně | Občas | Nesleduji |
|----------------------------------------------------------|------------|-------|-----------|
| TV Nova                                                  | •          |       |           |
| TV Prima                                                 |            |       | •         |
| ČT Sport                                                 |            | •     |           |

Zdroj:<sup>92</sup>

<sup>88</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I.* 1.vyd. Praha: VŠE, 1995, s. 19. ISBN 80-7079-737-1.

<sup>89</sup> ZBOŘIL, Kamil. *Marketingový výzkum: metodologie a aplikace.* 2. vyd. Praha: Oeconomica, 2003, s. 51. ISBN 80-245-0615-7.

<sup>90</sup> RAMBALOVÁ, Hana. *Marketingový výzkum.* Brno, 2011, s. 23-24. Bakalářská práce. Masarykova univerzita.

<sup>91</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I.* 1.vyd. Praha: VŠE, 1995, s. 19. ISBN 80-7079-737-1.

<sup>92</sup> Vlastní zpracování.

Jsou i typy otázek, které je vhodně zařadit až na úplný závěr dotazníku. Mezi ně patří otázky osobního charakteru, jako například věk, pohlaví, stupeň dokončeného vzdělání a další. V případě umístění těchto otázek na začátek dotazníku je zde risk, že respondent odmítne odpovídat na celý dotazník. Naopak při zařazení otázek na úplný závěr nebude mít respondent s odpovědí na tyto otázky větší problém, jelikož už strávil čas s vyplňováním celého dotazníku.<sup>93</sup>

### 4.3 Otázky dle účelu

Kromě pořadí, které je z hlediska návratnosti dotazníku vhodně dodržovat, je důležitý také účel otázek, které jsou v dotazníku obsaženy.

Nejdůležitější v celém dotazníku jsou otázky **meritorní**. Tyto otázky jsou zaměřeny na hlavní problém, jsou jádrem celého šetření a zodpovězení těchto otázek je klíčové pro celý výzkum. Tyto otázky pak mohou být dále doplněny či objasněny, aby jim respondent plně porozuměl. Doplnění či objasnění se dosahuje pomocí různých pomůcek, jako jsou grafy, obrázky či schémata.<sup>94</sup>

V dotazníku je obsažen také určitý počet **pomocných** otázek, které slouží k tomu, aby se rozhovor ubíral správným směrem. Může také obsahovat **větvící** otázky. Ty slouží k rozdělení respondentů do předem určených skupin. Tyto skupiny pak mohou dále odpovídat na zcela odlišné otázky. Například dle toho, na které televizní stanice se respondent dívá.<sup>95</sup>

Dalšími účelovými otázkami jsou filtrační otázky. Pomocí nich dotazník filtruje osoby, které jsou pro dotazník potřebné tím, že poskytnou určitou informaci, či nikoli. Například respondent sledující TV Nova může poskytnout odpovědi ohledně této televizní stanice.<sup>96</sup>

---

<sup>93</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 21-22. Bakalářská práce. Západočeská univerzita v Plzni.

<sup>94</sup> ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření: (druhé vydání)*. 2. vyd. Praha: Professional Publishing, 2010, s. 13 - 14. ISBN 978-80-7431-019-5.

<sup>95</sup> Tamtéž, s. 13-14.

<sup>96</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 18. Bakalářská práce. Západočeská univerzita v Plzni.



Kompletní dotazník by měl být co nejstručnější. Neměl by tedy obsahovat otázky, které jsou zcela zbytečné, výzkumu neposlouží a dotazník je jimi zbytečně prodlužován. Mezi tyto otázky patří například ty, které nijak nesouvisí s výzkumem, dotazují se na vedlejší informace, otázky, jejichž odpovědi budou mít velmi nespolehlivý charakter, duplicitní otázky a otázky s velmi malou variabilitou odpovědí.<sup>97</sup>

Při správném koncipování dotazníku lze však některým nevhodným a nepřínosným otázkám zamezit. Například ujasnění cílu dotazníku a účelu otázek lze zamezit tomu, aby dotazník obsahoval otázky mimo cíle výzkumu. Bohužel se nelze zcela vyhnout otázkám, na které dotazník přinese nekvalitní odpovědi. Je však vhodné použít **kontrolní** otázku.<sup>98</sup> To je taková, mezi níž a základní otázkou by měla být jasná souvislost. Uvede-li například respondent, že sleduje ČT Sport pravidelně a poté uvede výčet důvodů, proč tuto stanici nesleduje, je zde jasný rozpor. Díky kontrolním otázkám lze vyzorovat například otázky, které nejsou respondentovi zcela srozumitelné. V krajním případě mohou tyto otázky také pomoci odhalit podvody, jako je doplňování nevyplněných údajů.<sup>99</sup>

## 4.4 Otázky dle obsahu

Z hlediska obsahové stránky se otázky dělí primárně na **přímé** a **nepřímé**.<sup>100</sup>

Pokud má před sebou respondent v dotazníku přímou otázku, je zcela zjevné, na co odpovídá. Taková otázka může například znít: „Jaké televizní stanice sledujete?“. Respondentovi však nemusí být příjemné odpovídat na přímé otázky. Velmi často se může stát, že respondent v tomto případě nezvolí odpověď, s jakou by se ztotožnil, ale tu, jež mu přijde společensky přijatelnější. Tedy například napíše ty televizní stanice, které bude vnímat jako společensky nejpopulárnější.<sup>101</sup>

---

<sup>97</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 29. ISBN 978-80-86946-74-0.

<sup>98</sup> Tamtéž, s. 20.

<sup>99</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 30. ISBN 978-80-86946-74-0.

<sup>100</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 71. ISBN 80-7226-917-8.

<sup>101</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1. vyd. Praha: VŠE, 1995, s. 21. ISBN 80-7079-737-1.

Tento dopad lze zmírnit, pokud se mírně upraví formulace dané otázky. Například tak, že respondentovi nabídneme otázku, jež by ho zařazovala do skupiny osob jednajících podobně. Otázka by tedy zněla: „Jaké televizní stanice jsou ve vaší domácnosti sledovány?“<sup>102</sup>

Účinnějším způsobem jak zmírnit dopad přímé otázky je otázka nepřímá. V tomto případě pak respondentovi není zcela jasné, co je otázkami sledováno. Lze také použít metodu, při které je do otázek promítnuto chování jiných osob. V případě, že je použito jméno konkrétní osoby, ale může dojít ke zkreslení. Pokud je otázka položena například tak, že se ptá na souhlas či názory nějakého politika, může se respondent ztotožnit s osobou politika a vyjadřuje pak souhlas či nesouhlas s tímto politikem a nikoli s uvedeným názorem na věc.<sup>103</sup>

V průzkumech trhu se pak využívají speciální metody a postupy, které přinášejí informace na základě nějakých podnětů, které zdánlivě nesouvisí s předmětem šetření. Jedná se například o metody, kdy je od respondenta vyžadováno, aby reagoval na fotografii či dialog. Respondent tak má možnost řadit značky výrobků dle zadaných kritérií, vybírá dle jeho uvážení nejlepší koncept z několika jemu předložených možností nebo vyjadřuje své představy o uživateli konkrétního výrobku.<sup>104</sup>

## 4.5 Otázky dle formy

Otázky dle formy dělíme na uzavřené a otevřené.<sup>105</sup>

Uzavřené otázky jsou vyžadovány zejména u kvantitativních metod. V případě uzavřených otázek respondent pouze vybírá odpovědi buď ze dvou možností, nebo z více možností. Pokud vybírá z okruhu dvou možností, jedná se o alternativní otázky. Pokud vybírá z více možností, jedná se o selektivní otázky.<sup>106</sup>

---

<sup>102</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 30. ISBN 978-80-86946-74-0.

<sup>103</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 72. ISBN 80-7226-917-8.

<sup>104</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 31. ISBN 978-80-86946-74-0.

<sup>105</sup> ŠKRAŇKA, Ivan. *Statistické metody v marketingovém výzkumu*. Brno, 2009, s. 31. Diplomová práce. Vysoké učení technické v Brně.

<sup>106</sup> Tamtéž, s. 31.

Varianty odpovědí v případě uzavřených otázek by měly pokrýt celou škálu možných odpovědí. Nemělo by se tudíž stávat, že respondent se neztotožní s žádnou z odpovědí, která je mu nabízena, ale nebude mít možnost odpovědět mimo tuto nabídku. Zároveň by ale výčet odpovědí neměl být příliš dlouhý, aby pak respondent nevybíral pouze například z několika prvních či posledních možností. Pokud by byl výčet odpovědí dlouhý, lze přikročit k variantě otázky polouzavřené. V takovém případě má respondent možnost, v případě že se jeho názor neztotožnil s žádným z uvedených, vypsát do kolonky svůj vlastní. Místo této možnosti lze také použít odpověď typu „nevím“, aby otázka zůstala uzavřená.<sup>107</sup>

V případě, že dotazník obsahuje selektivní otázky, může respondent dle povahy otázky zvolit buď jednu, anebo i více odpovědí. Tato možnost se nazývá multiple response. Vícenásobné odpovědi jsou však náročnější na zpracování a vyžadují tudíž speciální analytické postupy.<sup>108</sup>

Pokud má respondent naopak odpovídat na otevřené otázky, formuluje si odpovědi sám. I v tomto případě však nemusí být pro respondenta snadné odpovědět. Například otázky na měsíční příjem vnímá mnoho respondentů jako velmi citlivé a osobní, a proto je neradi uvádí. Možností, jak tomu předejít, je například určitá kategorizace, kdy respondent neuvede přímou výši svého měsíčního příjmu, ale vybere si interval, v jakém se jeho příjem nachází.<sup>109</sup>

V případě otevřených otázek je nutné důsledně formulovat otázku. Špatně formulovaná otázka znesnadní nejen samotné dotazování, ale i následné zpracování získaných informací.<sup>110</sup>

Dotazování ale může probíhat i formou rozhovorů. V tomto případě se jedná o dotazování, které probíhá tváří v tvář respondentu a tazatele.

---

<sup>107</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 22. ISBN 80-7079-737-1.

<sup>108</sup> BÁRTOVÁ, Hilda. *Marketingový výzkum trhu*. 1.vyd. Praha: Economia, 1991, s. 107. ISBN 80-85378-09-4.

<sup>109</sup> PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 23. ISBN 80-7079-737-1.

<sup>110</sup> BÁRTOVÁ, Hilda. *Marketingový výzkum trhu*. 1.vyd. Praha: Economia, 1991, s. 107. ISBN 80-85378-09-4.

## 5 MĚŘENÍ A ŠKÁLY

Veškeré údaje zjišťované v dotazníku se musí nějakým způsobem klasifikovat. Je to nutný předpoklad k tomu, aby se data dala nějakým způsobem změřit a následně dále zpracovat. V případě, kdy je v dotazníku uveden dotaz například na věk, váhu, výši platu apod., je způsob měření veskrze snadný. Respondent odpoví na otázku jednoznačně. V dotazníku se ale mohou objevit i otázky, na které není odpověď takto jednoznačná, jelikož z velké míry závisí na posouzení a pocitech respondentů. Například: "Jste spokojeni s...". Tyto znaky se již měří obtížně a nastává zde problém, jak změřit tyto znaky zjednodušeně, ale zároveň kvalitně, aby odpovídaly skutečnému pocitu respondenta.<sup>111</sup>

Zjednodušením jsou v tomto případě škály. Pomocí nich lze na jedné stupnici vyjádřit zkoumaný jev, který zahrnuje jednu proměnou, ale více znaků. Ale nalézt vhodnou škálu či stupnici, aby bylo možné sesbíraná data efektivně analyzovat, je velmi obtížné. Už samotná tvorba určité stupnice představuje obtížný úkol. Často bývá tvorba vhodné stupnice pracným výzkumným úkolem.<sup>112</sup>

Pouze pro ilustraci, například na škále "Jste spokojeni s..." bude uvedeno "Jsem spokojen" "nevím" "nejsem spokojen". Možností může být samozřejmě více. Čím více možností má respondent na výběr, tím lépe se dokáže ztotožnit s jednou z možností.

Díky tomuto systému škálování je možné změřit pocity respondentů velmi jednoduše. Každý respondent se ztotožní s nějakou z odpovědí a výsledná data bude možné zpracovat o mnoho jednodušeji.<sup>113</sup>

V případě, že pro zkoumání daného jevu nepostačuje pouze jednorozměrná škála, jak bylo uvedeno výše, lze vzít v úvahu i další dimenze. V takovém případě se bude jednat

---

<sup>111</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 24-25. ISBN 978-80-86946-74-0.

<sup>112</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 35. ISBN 80-245-0753-6.

<sup>113</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 20-21. Bakalářská práce. Západočeská univerzita v Plzni.

o škálu dvourozměrnou či vícerozměrnou. Ale vzhledem k tomu, že již vytvoření jednorozměrné škály není jednoduché, jedná se v tomto případě o velmi těžký úkol.<sup>114</sup>

## 5.1 Stevensova typologie

Stevensova typologie je jedním z nejjednodušších způsobů, jak klasifikovat proměnné. Tato typologie vznikla ve 40. letech minulého století a dnes je používána pouze v omezené míře. Pomocí této typologie je možné dělit proměnné do 3 skupin. Nominálních, ordinálních a kardinálních proměnných, jelikož pro každou proměnnou platí jiná pravidla. Například ne u každé proměnné lze hovořit o její vyšší či nižší hodnotě vzhledem k jiným hodnotám.<sup>115</sup>

Nominální proměnná umožňuje pouze označení objektů a slouží zejména k usnadnění třídění dat. Tato hodnota je většinou vyjádřena slovy. Typickým příkladem je rodinný stav. Zde není možné určit ani pořadí, ani hodnoty měřit. V případě, že je hodnota nominální proměnné vyjádřena číslem, jedná se pouze o funkci kódu, který zde slovo zastupuje.<sup>116</sup>

Ordinální proměnné mohou být vyjádřeny číslem, ale i slovně. Pomocí této proměnné se určuje zejména pořadí. Typickým příkladem pro ordinální proměnnou je vzdělání. V tomto případě vzdělání vyjadřuje určitý stupeň, ale nelze přesně vyjádřit, o kolik je vysoká škola vyšší hodnotou než střední škola.<sup>117</sup>

Kardinální proměnná označuje taková data, která jsou nějakým způsobem měřitelná. Ve většině případů je tedy hodnotou kardinální proměnné nějaké číslo. Zároveň je možné v případě kardinálních proměnných přesně určit, o kolik se jedna hodnota liší od druhé. Ideálním příkladem je tedy výše platu či věk. Tento typ proměnné však umožňuje vyjádřit data i pomocí intervalu. Kardinální stupnice ale není příliš používána.<sup>118</sup>

---

<sup>114</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 35. ISBN 80-245-0753-6.

<sup>115</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 32-33. ISBN 978-80-86946-74-0.

<sup>116</sup> ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření: (druhé vydání)*. 2. vyd. Praha: Professional Publishing, 2010, s. 22. ISBN 978-80-7431-019-5.

<sup>117</sup> Tamtéž, s. 22.

<sup>118</sup> Tamtéž, s. 22.

Stevensova typologie se v dnešní době příliš nevyužívá. Její velkou nevýhodou je značné zjednodušení proměnných. Pomocí této typologie není možné vytvořit stupnici hodnot a data následně měřit pomocí měřících škál. Proto se přistupuje k jiným metodám, které umožňují lepší třídění a analýzu dat.<sup>119</sup>

## 5.2 Bodovací škály

Bodovací škály jsou nejpoužívanější a nejběžnější metodou v oblasti průzkumů. Jejich použití je velmi jednoduché a na základě toho jsou tyto bodovací škály velmi oblíbené. Ale i tvoření bodových škál vyžaduje určité promyšlené postupy a zacházení.<sup>120</sup>

Bodovací škály jsou tvořeny jednoznačně uspořádanými hodnotami. Tyto hodnoty mohou být vyjádřeny číselně, ale i slovně. Nejčastěji se používá 5 až 7 hodnot pro dostatečně široké spektrum odpovědí. Tato škála může být tedy tvořena čísly 1 až 7, ale také hodnocením „výborně“, „velmi dobře“, „dobře“, „dostatečně“, „nedostatečně“.<sup>121</sup>

Ve fázi tvoření bodovací škály je nutné zvážit především vhodný počet škálových hodnot. Velmi praktické je zvolit lichý počet hodnot. Díky tomu je možné doprostřed stupnice umístit neutrální postoj. Za neutrálním postojem a před ním se pak umístí shodný počet hodnot.<sup>122</sup> V případě číselné bodovací škály je vhodné čísla doplnit nějakým popiskem. V tomto případě je bodovací škála tvořena formou ordinální stupnice. Avšak používání čísel k vyjádření hodnot může vést k tomu, že vzdálenost mezi číslicemi je vnímána jako identická. Následkem toho je pak proměnná chápána a zpracovávána jako kardinální.<sup>123</sup>

Kromě slovního a číselného vyjádření však existuje i mnoho dalších variant, jak vytvořit bodovací škálu. Může být tvořena například graficky, kdy respondent volí na škále několika možností od rozesmátého obličej po zamračený, či od prázdné nádoby

---

<sup>119</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 34. ISBN 978-80-86946-74-0.

<sup>120</sup> Tamtéž, s. 35.

<sup>121</sup> TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 15. Bakalářská práce. Vysoká škola ekonomická v Praze.

<sup>122</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 35. ISBN 978-80-86946-74-0.

<sup>123</sup> TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 14. Bakalářská práce. Vysoká škola ekonomická v Praze.

k postupně se zaplňující, dle toho, nakolik souhlasí s daným výrokem. Další možností je také možnost vyjádřit stupeň souhlasu na úsečce, jejíž krajní body obsahují dva naprosto protichůdné názory (například názor zcela souhlasím a zcela nesouhlasím). V tomto případě pak lze použitou stupnici kvalifikovat jako kardinální poměrovou a dle toho ji pak dále analyzovat.<sup>124</sup>

Kromě těchto dvou základních postupů škálování existují i speciální šalovací postupy.

### 5.3 Likertova metoda

Likertova metoda je speciální škálovací postup a patří mezi ty nejrozšířenější. Pomocí této metody jsou respondentovi předloženy jednoduché výroky ohledně zkoumané skutečnosti. Respondent poté na pětistupňové škále zaškrtně takové číslo, které se nejvíce shoduje i s jeho názorem. Poté se celkové body sečtou u jednotlivých respondentů. To pak představuje škálové ohodnocení sledovaného jevu.<sup>125</sup> Na základě tohoto postupu je možné označit získanou stupnici za kardinální.<sup>126</sup> V případě, že číslo 1 znamená naprostý souhlas a číslo 5 naprostý nesouhlas, znamená to, že respondenti s nejnižším skóre mají velmi kladný vztah ke sledované skutečnosti. Naopak respondenti s nejvyšším skóre jsou velmi nespokojeni, či mají negativní vztah ke zkoumané skutečnosti.<sup>127</sup>

V případě jemnější stupnice lze použít více otázek, tedy více stimulů. Je ale nutné zohlednit, zda má takové rozšíření smysl a zda přispěje k lepšímu či detailnějšímu změření sledované problematiky. V opačném případě je respondent zbytečně zatěžován otázkami, které nemají pro další analýzu žádný smysl.<sup>128</sup> K odhalení těchto zbytečně zatěžujících stimulů lze použít tzv. Likertův koeficient diferenciací. Ten nabývá hodnot od 0 do 1, přičemž čím nižší je, tím méně respondenty diferencuje, což znamená, že

---

<sup>124</sup> TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 14. Bakalářská práce. Vysoká škola ekonomická v Praze.

<sup>125</sup> SRNA, Martin. *Marketingový výzkum spokojenosti zákazníků*. Brno, 2008, s. 33. Bakalářská práce. Vysoké učení technické v Brně.

<sup>126</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 37. ISBN 978-80-86946-74-0.

<sup>127</sup> HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 109. ISBN 80-7226-917-8.

<sup>128</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 37. ISBN 978-80-86946-74-0.

jsou respondenti zbytečně zatěžováni. V případě, že tedy koeficient diferenciacie vyjde blízký nule, je vhodné tento konkrétní stimul z dotazníku odstranit, aby se zvýšila jeho efektivita.<sup>129</sup>

Tabulka 2: Likertova stupnice

| Stupeň souhlasu             | 1 | 2 | 3 | 4 | 5 |
|-----------------------------|---|---|---|---|---|
| Pravidelně sleduji televizi |   |   | • |   |   |
| Pravidelně čtu časopisy     | • |   |   |   |   |
| Pravidelně poslouchám rádio |   |   |   |   | • |

Zdroj<sup>130</sup>

Takto vypadá použití Likertovy metody v praxi. Respondent se jednoduše rozhodne na nabízené stupnici. Poté se jednotlivé body sečtou a zjistí se, k jakému médiu mají respondenti nejbližší, nebo které naopak vůbec nesledují. Případně se dle koeficientu diferenciacie zjistí, zda nějaká otázka zbytečně nezatěžuje respondenty a tím nepřináší žádný užitek.

## 5.4 Sémantický diferenciál

Sémantický diferenciál využívá antonyma k tomu, aby respondent vyjádřil svůj názor. Pro co nejlepší výsledky by měl být tvořen co největším počtem stimulů, které nějakým způsobem charakterizují daný objekt. Pro každou vlastnost tohoto objektu jsou poté stanovena antonyma. Například „snadno dostupný“ versus „nedostupný“. Tato antonyma jsou zároveň krajními body stupnice, většinou pěti či sedmibodové. Zároveň je ze stupnice jasné patrné, jaký stupeň vlastnosti odpovídá jakým číslům na stupnici. Například „snadno dostupný“ vyjadřuje číslo 1, naopak „nedostupný“ vyjadřuje číslo 5.

<sup>129</sup>PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 45. ISBN 80-245-0753-6.

<sup>130</sup>Vlastní zpracování



Respondent na základě předložených informací volí dle svého uvážení. Zdá-li se mu například daný objekt snadno dostupný v okrese bydliště, zvolí číslo 1.<sup>131</sup>

Důležité je správně formulovat otázky a orientovat je vždy stejným směrem. Tedy příznivé hodnocení musí být na jedné straně diferenciálu a nepříznivé hodnocení vždy na druhé. V případě špatné konstrukce diferenciálu je znemožněna následná analýza, jelikož by výsledná čísla neměla žádný vypovídající charakter.<sup>132</sup>

Pomocí sémantického diferenciálu lze změřit názor jednotlivce na daný objekt, ale také ho lze využít pro porovnání dvou objektů. Tento způsob použití je častější. Hodnocení stimulů jednotlivého objektu je sečteno za všechny respondenty pomocí průměru či mediánu a je porovnáno s výsledky jiného objektu. Pomocí grafického vyjádření je pak snadné určit, jaký objekt má lepší vlastnosti.<sup>133</sup>

Tabulka 3: Sémantický diferenciál

| Objekt                    | 1 | 2 | 3 | 4 | 5 |                             |
|---------------------------|---|---|---|---|---|-----------------------------|
| <b>Snadno dostupné</b>    |   |   |   |   |   | <b>Zcela nedostupné</b>     |
| <b>Cenově příznivé</b>    |   |   |   |   |   | <b>Cenově nepříznivé</b>    |
| <b>Příjemné prostředí</b> |   |   |   |   |   | <b>Nepříjemné prostředí</b> |

Zdroj<sup>134</sup>

Pokud není z nějakého důvodu možné či výhodné použít právě sémantický diferenciál, lze využít jeho obdobu, Stapelovu škálu. Tou se rozumí symetrická stupnice, která neobsahuje neutrální bod. Tato stupnice má na rozdíl od sémantického diferenciálu i záporné hodnoty. Její rozsah může být například od -5 do 5. Dolní mez představuje

<sup>131</sup> SRNA, Martin. *Marketingový výzkum spokojenosti zákazníků*. Brno, 2008, s. 31. Bakalářská práce. Vysoké učení technické v Brně.

<sup>132</sup> KRČMAŘÍKOVÁ, Monika. *Možnosti využití metody sémantického diferenciálu k získávání informací důležitých pro manažerské rozhodování*. Nymburk, 2007, s. 16-17. Bakalářská práce. Vysoká škola ekonomická v Praze.

<sup>133</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 39. ISBN 978-80-86946-74-0.

<sup>134</sup> Vlastní zpracování.

nejhorší možné vlastnosti a naopak horní mez představuje nejlepší možné vlastnosti. Žádné meze ale nejsou vyjádřeny slovně. Respondent pak pomocí Stapelovy škály vyjadřuje určité vlastnosti nějakého objektu. Například zdá-li se mu daný objekt spíše nedostupný v jeho okolí, zvolí číslo -3. Stapelova škála se vyhodnocuje stejným způsobem jako sémantický diferenciál. Po vyhodnocení je tedy rovněž možné ihned říci, jaký objekt zaostává v jakých vlastnostech a v jakých naopak vyniká.<sup>135</sup>

## 5.5 Metoda zdánlivě stejných intervalů

Tato metoda byla v českém jazyce poprvé popsána v roce 1986.<sup>136</sup> Metoda zdánlivě stejných intervalů vychází z předpokladu, kdy lze sledované znaky rozčlenit na určitý počet kategorií či stupňů. Je vhodné pracovat s ní právě v případech, kdy není možné použít metodu párového srovnání. K tomu dochází v případech, kdy je nutné zpracovat velký počet stimulů, což metoda párových srovnání neumí. Zde se porovnává jeden stimul s jiným stimulem, nebo s více stimuly. Tímto způsobem lze zpracovat i větší počet stimulů.<sup>137</sup> Tato metoda využívá ordinální škálu. Tedy škálu, na které lze vyjádřit stupeň či pořadí, ale nelze určit, o kolik je jedna hodnota vyšší než jiná. Pořadí na této ordinální škále je vyjádřeno přirozenými čísly, nejčastěji od 1 do 5 a jsou zároveň vnímána jako intervalová škála. Respondent je poté požádán, aby například určil důvěryhodnost některého média právě na této škále. Výsledné škálové hodnocení vzniká váženým aritmetickým průměrem těchto pořadových čísel, kde jsou vahami právě četnosti případů zařazené do této kategorie. Na základě těchto výpočtů poté vznikne nová stupnice, kde již můžeme veškeré vlastnosti seřadit dle průměru, jakého

---

<sup>135</sup> SRNA, Martin. *Marketingový výzkum spokojenosti zákazníků*. Brno, 2008, s. 32. Bakalářská práce. Vysoké učení technické v Brně.

<sup>136</sup> KRČMAŘÍKOVÁ, Monika. *Možnosti využití metody sémantického diferenciálu k získávání informací důležitých pro manažerské rozhodování*. Nymburk, 2007, s. 13. Bakalářská práce. Vysoká škola ekonomická v Praze.

<sup>137</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 40. ISBN 978-80-86946-74-0.

dosáhli. V tomto případě již můžeme říci, o kolik je jedna vlastnost lepší, či horší než jiná, a proto lze tuto výslednou stupnici považovat za kardinální.<sup>138</sup>

V praxi vypadá tato metoda následovně.

Tabulka 4: Metoda zdánlivě stejných intervalů

| Objekt | Kategorie |   |   |   |   |
|--------|-----------|---|---|---|---|
|        | 1         | 2 | 3 | 4 | 5 |
| A      |           | 5 |   | 2 |   |
| B      | 2         |   |   | 3 |   |
| C      |           |   | 1 |   |   |

Zdroj<sup>139</sup>

Tabulka představuje zjednodušenou verzi použití metody párového srovnání v praxi. Čísla 1 - 5 představují důvěryhodnost jednotlivých médií. Písmena A - C představují jednotlivá média. Z vyplněné tabulky je patrné, že médium A bylo pětkrát zařazeno do kategorie 2 (spíše důvěryhodný) a dvakrát zařazeno do kategorie 4 (spíše nedůvěryhodný). Takto bude postupováno i u dalších dvou médií. Poté se váženým aritmetickým průměrem vypočte výsledná důvěryhodnost jednotlivých médií.

## 5.6 Skalogram

Skalogram je také jinak nazýván Guttmanova škála. Vznikla, jelikož na základě předchozích zkušeností vyplynulo, že je příliš unáhlené a riskantní dělat závěry o názorech a mínění populace na základě jedné otázky na danou problematiku. Respondenti sami velmi často neumí určit, zda jejich postoj je kladný či záporný, a volí proto spíše neutrální odpovědi. Skalogram tedy vznikl, aby bylo možné změřit míru sledovaného jevu pomocí více otázek formulovaných pokaždé trochu jinak, kdy se všechny týkají jednoho sledovaného jevu.<sup>140</sup>

<sup>138</sup> Tamtéž, s. 40.

<sup>139</sup> Vlastní zpracování

<sup>140</sup> KRČMARÍKOVÁ, Monika. *Možnosti využití metody sémantického diferenciálu k získávání informací důležitých pro manažerské rozhodování*. Nymburk, 2007, s. 13-15. Bakalářská práce. Vysoká škola ekonomická v Praze.

Skalogram je vytvořen tak, že kladná odpověď respondenta na nějakou otázku znamená s velkou pravděpodobností, že odpověděl kladně i na všechny předchozí otázky. Stane-li se toto ve všech případech a respondent opravdu odpoví kladně na všechny otázky, jedná se o perfektní skalogram. Sestavení skalogramu ale bývá velmi složitou záležitostí. Pokud je navíc použito většího počtu otázek, není možné obejít se bez počítače, jelikož následná analýza je velmi náročná.<sup>141</sup>

Klasickým příkladem využití skalogramu je například měření míry rasismu. Každý člověk by na přímou otázku: „Jste rasista?“, odpověděl negativně. Na základě této odpovědi by se tedy dalo usuzovat, že v České republice nejsou rasisti, nebo je jich jen velmi málo. Ovšem pomocí skalogramu se dá změřit míra rasismu, která by nikdy nebyla odhalena, kdyby se používala pouze výše uvedená jednoduchá otázka. Skalogram klade konkrétnější otázky, pod kterými si respondent jasně představí modelovou situaci a dokáže na ni odpovědět. V případě rasismu tedy otázka může znít: „Vadí vám soused s jinou barvou pleti?“, „Pokud byste byl spřízněn s osobou jiné barvy pleti, vadilo by vám to?“ Takto respondent odpoví na větší počet otázek, aby byly výsledky co nejpodrobnější. Pomocí těchto otázek lze již změřit, nakolik je respondent rasistický.

## 5.7 Metoda párových srovnání

Veškeré metody, které byly zmíněny výše, jsou založeny na principu přímého hodnocení. Metoda párových srovnání je ale založena na principu vzájemného srovnání jednotek vzhledem k nějakému měřenému jevu. Všechna pořadí jsou tedy vytvořena na základě porovnávání dvojic. Je ale nutné určit přiměřený počet dvojic, aby nebylo jejich uspořádání příliš obtížné.<sup>142</sup> Tato metoda vznikla v roce 1927 a má široké využití.<sup>143</sup>

---

<sup>141</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 40-43. ISBN 978-80-86946-74-0.

<sup>142</sup> TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 11. Bakalářská práce. Vysoká škola ekonomická v Praze.

<sup>143</sup> KRČMARÍKOVÁ, Monika. *Možnosti využití metody sémantického diferenciálu k získávání informací důležitých pro manažerské rozhodování*. Nymburk, 2007, s. 13. Bakalářská práce. Vysoká škola ekonomická v Praze.

Pomocí této metody se předem stanoví nějaké kritérium a na základě něho se poté porovnávají dvojice jednotek. Poté lze analýzou určit, jakému objektu byla dána přednost před jinými. Výsledné četnosti jsou uspořádány do tabulky, jejíž rozměry jsou určeny počtem všech sledovaných jednotek. Tabulky tedy mohou být opravdu velké.<sup>144</sup>

Tato metoda tak umožňuje porovnávat některé podněty vícekrát než jednou. Nebo například určit, že veškeré podněty budou porovnány pouze se třemi dalšími a nikoli se všemi podněty. Tato metoda je však velmi náročná a vyžaduje přesně definovat podněty a kritéria pro jejich posuzování. Riziko také představuje fakt, že se z této metody může stát přímé měření postoje. Je tedy nutné předem důkladně proškolen pracovníky, jež budou srovnání provádět. Velkou nevýhodou je také značná pracnost provedení a tím i vysoké náklady na celé šetření.<sup>145</sup>

Tabulka 5: Metoda párových srovnání

|          | <b>A</b> | <b>B</b> | <b>C</b> |
|----------|----------|----------|----------|
| <b>A</b> | -        | 0        | 1        |
| <b>B</b> | 0        | -        | 1        |
| <b>C</b> | -1       | 1        | -        |
| $\Sigma$ | -1       | 1        | 2        |

Zdroj<sup>146</sup>

V praxi vypadá tato metoda následovně. Do tabulky se vepíše, s jakým výsledkem skončilo porovnávání jednotlivých stimulů mezi sebou. Přičemž 0 znamená, že porovnání skončilo remízou, 1 znamená výhru a -1 znamená prohru. Poté pomocí součtu zjistíme, jakou výhodu mají jednotlivé značky před jinými. Neboli v kolika procentech má jednotlivá značka přednost před ostatními značkami u jednotlivých spotřebitelů. Kromě toho lze ale určit i to, jak moc jsou si některé značky podobné, podle toho, na jakém místě skončily.

<sup>144</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 37-38. ISBN 80-245-0753-6.

<sup>145</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 40-41. ISBN 978-80-86946-74-0.

<sup>146</sup> Vlastní zpracování

## 5.8 Zlomkové škály

Zlomkové škály jsou další metodou, kterou lze změřit získaná data. I tato metoda je založena na vzájemném srovnání objektů. V případě použití této škály je nejprve vybrán nějaký stimul a tomu je přiřazeno určité bodové hodnocení, například 100 bodů. Další hodnocení pak vychází ze srovnání s tímto stimulem.<sup>147</sup> Respondent tedy ví, že například televizní stanice ČT Sport je hodnocena 100 body. Poté má sám přiřadit body jiným sportovním stanicím například dle důvěryhodnosti, či kvality podávaných informací. Respondent tedy může preferovat jinou televizní stanici, a té dá samozřejmě vyšší hodnocení. Naopak jiným, neoblíbeným televizním stanicím, dá nižší hodnocení. Výsledná stupnice je ale velmi závislá právě na vybraném základu, s kterým respondent porovnává vše ostatní. Bohužel nelze očekávat, že v případě změny tohoto základu bude i respondentovo hodnocení změněno v odpovídajícím poměru. Výslednou stupnici je tedy možné chápat pouze jako ordinální.<sup>148</sup>

Nevýhodou této škály je možnost výskytu extrémních hodnot. Respondent může nějakou sportovní stanici považovat za naprosto jedinečnou a přiřadit jí tak hodnocení i 10 000 bodů.<sup>149</sup>

Tomuto se snaží předcházet škály konstantního součtu. Tyto škály fungují na stejném principu, jako zlomkové škály. Díky svým omezením však zabraňují tomu, aby se v hodnocení vyskytovaly velmi extrémní hodnoty, jako právě u zlomkových škál.<sup>150</sup>

V případě škál konstantního součtu má respondent k dispozici předem stanovený počet bodů, například 100. Poté má tyto body přiřadit určitým objektům. To, kolik bodů přiřadí či nepřiřadí nějakému objektu, závisí čistě na respondentovi. Pokud by ale stejně jako v předchozím případě, preferoval například jednu sportovní stanici, bude jí moci v tomto případě přiřadit maximálně těch 100 bodů, jelikož více nemá k dispozici.<sup>151</sup>

---

<sup>147</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 40-41. ISBN 80-245-0753-6.

<sup>148</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 40-41. ISBN 80-245-0753-6.

<sup>149</sup> TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 13-14. Bakalářská práce. Vysoká škola ekonomická v Praze.

<sup>150</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 40-41. ISBN 80-245-0753-6.

<sup>151</sup> TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 13-14. Bakalářská práce. Vysoká škola ekonomická v Praze.

## 6 PRÁCE S DATY

Po klasifikaci a rozčlenění získaných dat pomocí škál je nutné s daty i nadále pracovat. Pouhé rozčlenění dat nestačí k tomu, aby se vytvořil nějaký závěr aplikovatelný na celou populaci. Tato data se tedy většinou zpracují do tabulek rozdělení četností. Data se ale také často zpracovávají pomocí různých statistických charakteristik, které při vhodném zvolení umožňují kvalitní interpretaci získaných výsledků. Tyto charakteristiky poté měří vlastnosti jednorozměrných, ale i dvourozměrných či vícerozměrných rozdělení četností. V této práci jsou uvedeny pouze základní charakteristiky, se kterými se běžně pracuje při výzkumech veřejného mínění.<sup>152</sup>

Nejběžněji se při analýze dat z výzkumu veřejného mínění využívají míry polohy a variability. Kromě nich ale existují i například míry špičatosti. Ty mají výpočet založený na poměru rozpětí kvartilů a variačního koeficientu. Použitím měr špičatosti lze snadno odhalit extrémní hodnoty v souboru. Jinak se také může nazývat míra koncentrace.

Lze měřit i míry symetrie, nebo jinak míry šikmosti. Tato míra pomáhá určit, jakým směrem je sledovaná proměnná rozložena. Rozděluje se dále na šikmost kladnou, neboli pravostrannou. V tom případě je většina naměřených hodnot pod průměrem. Opakem je šikmost záporná, neboli též levostranná, kdy se naopak většina naměřených hodnot nachází nad průměrem. Míru tohoto rozložení lze poté určit tzv. koeficientem šikmosti.<sup>153</sup>

Míry šikmosti a špičatosti jsou však náročnější na výpočet a vyžadují tudíž více času na zpracování. Tím vzrůstají i případné náklady na celkovou analýzu výsledků získaných při sběru dat. Proto se ve většině případů užívá k charakteristice nejjednodušší metody, a sice míry polohy a variability, jež mají velkou vypovídající hodnotu a umožňují

---

<sup>152</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 22 - 23. ISBN 80-86946-16-9.

<sup>153</sup> LÖSTER, Tomáš, Hana ŘEZANKOVÁ a Jitka LANGHAMROVÁ. *Statistické metody a demografie*. 1. vyd. Praha: VŠEM, 2008, s. 27-28. ISBN 978-80-86730-40-0.

základní charakteristiku získaných dat.<sup>154</sup> V této práci jsou dále podrobně rozebrány pouze míry polohy a míry variability.

## 6.1 Míry polohy (úrovně)

Srovnávání a popisování několika proměnných pouze pomocí tabulek rozdělení četností je velmi náročné a pracné. Výsledné práce poté ani nejsou dostatečně přehledné. Proto se na charakteristiku a analýzu získaných dat využívají míry polohy. Míry polohy umožňují shrnout tyto informace, které jsou v datech obsaženy a vyjádřit ji v koncentrovanější formě pomocí určitých charakteristik. Těchto charakteristik není mnoho, ale velmi přesně popisují základní rysy sledovaných hodnot. Míry polohy jsou jinak také nazývány míry centrální tendence.<sup>155</sup>

Úroveň sledovaného znaku je považována za jeho základní rys. Měří se pomocí různých druhů středních hodnot. Jedná se o jednoduché číselné charakteristiky, které napomáhají ke zjednodušení a zobecnění všech hodnot obsažených v souboru. Míry polohy právě takové zjednodušení představují. Zevšeobecňují a zastupují jednotlivé hodnoty sledovaných znaků. Tyto míry umožňují jednoduché srovnání polohy dvou, ale i více rozdělení četností. Zároveň je pomocí nich možné srovnávat úrovně dvou či více souborů.<sup>156</sup>

Mezi míry polohy patří různé druhy průměrů neboli středních hodnot, různé druhy kvantilů, medián, modus.<sup>157</sup>

Tyto hodnoty určitým způsobem charakterizují střední hodnotu souboru, okolo které ostatní hodnoty kolísají. Kromě nich však existují i kvantily, které neudávají povědomí o středu souboru, ale přesto jsou pro analýzu souboru podstatné.<sup>158</sup>

---

<sup>154</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 22 - 23. ISBN 80-86946-16-9.

<sup>155</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 22 - 25. ISBN 80-86946-16-9.

<sup>156</sup> Tamtéž, s. 22-25.

<sup>157</sup> SRNA, Martin. *Marketingový výzkum spokojenosti zákazníků*. Brno, 2008, s. 40. Bakalářská práce. Vysoké učení technické v Brně.

<sup>158</sup> Tamtéž, s. 40.



Míry polohy často nabývají i hodnot, jakých sledovaný znak z logiky věci ani nenabývat nemůže. Například udává průměrný počet pacientů kteří onemocní nějakou nemocí jako 20,2 pacienta ročně. I přes tuto skutečnost ale poskytují důležitý zdroj informací o analyzovaném souboru a mají určitou vypovídající hodnotu o daném souboru. Kromě toho lze na základě těchto charakteristik určit i případné chování tohoto souboru.<sup>159</sup>

## Průměry

Nejčastěji používané charakteristiky z oblasti míry úrovně jsou průměry. Pomocí průměru se počítají střední hodnoty ze všech jednotek v souboru. Nejdůležitějšími průměry jsou aritmetický, harmonický, geometrický a kvadratický. Pro účely výzkumu veřejného mínění se však využívá téměř výhradně pouze aritmetický průměr a to ve svých dvou podobách.<sup>160</sup>

V prvním případě se využívá prostý aritmetický průměr. Ten je dokonce nejvyužívanější ze všech průměrů. Jeho výpočet je také nejjednodušší. Hodnoty nejsou nijak uspořádány a při prostém aritmetickém průměru se zjištěné hodnoty vydělí počtem pozorování, respektive počtem respondentů. Tím se získá základní představa o tom, jak respondenti v průměru hodnotili, či jak v průměru reagovali na daný podnět.<sup>161</sup>

Pokud ale nastane druhý případ, tedy že hodnoty mají nějakou váhu, tak není možné použít pouze prostý aritmetický průměr. Je nutné využít aritmetický průměr vážený. Ten ve výpočtu zahrnuje i fakt, že různé varianty sledovaného znaku mají různou důležitost.<sup>162</sup> Například právě u škálovací metody zdánlivě stejných intervalů se použije vážený aritmetický průměr, protože každému subjektu byla dána jiná váha dle toho, kolik respondentů jej označilo na stupnici jako důvěryhodného / nedůvěryhodného. Při využití prostého aritmetického průměru by výsledné hodnoty nedávaly příliš smysl.<sup>163</sup>

---

<sup>159</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 22 - 25. ISBN 80-86946-16-9.

<sup>160</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 29. ISBN 80-86946-16-9.

<sup>161</sup> LÖSTER, Tomáš, Hana ŘEZANKOVÁ a Jitka LANGHAMROVÁ. *Statistické metody a demografie*. 1. vyd. Praha: VŠEM, 2008, s. 19-20. ISBN 978-80-86730-40-0.

<sup>162</sup> KAISAROVÁ, Petra. *Základy popisné statistiky v pedagogickém výzkumu*. Brno, 2009, s. 36. Bakalářská práce. Masarykova univerzita.

<sup>163</sup> Tamtéž, s. 36.

Kromě aritmetického průměru existují i další průměry, které mají své využití zejména v popisné statistice, v časových řadách atd. Geometrický průměr se používá zejména v oblastech, kde se počítají indexy. Kvadratický průměr je naopak velmi užitečný při výpočtech koeficientů růstu v oblasti časových řad. Harmonický průměr se zase využívá při výpočtu průměru časové vzdálenosti. Pro účely analýzy dat z výzkumu veřejného mínění ale své uplatnění příliš nenachází, a proto se v této oblasti ani příliš neužívají.<sup>164</sup>

## Kvantily

Kromě měr polohy lze pro změření získaných dat využít také kvantily. To jsou hodnoty, pomocí kterých je celý soubor hodnot rozdělen na dvě části. První část obsahuje takové hodnoty, které jsou menší, nebo stejné, než zvolený kvantil. Druhá část naopak obsahuje pouze hodnoty, které jsou větší či stejné, jako zvolený kvantil. Ve výzkumech veřejného mínění se kvantil uplatní v rámci změření, kolik osob zaujímalo konkrétní postoj na konkrétní škále.<sup>165</sup>

Nejpoužívanější je 50% kvantil, jinak také nazývaný medián. Medián určuje prostřední hodnotu celého souboru hodnot. Pokud soubor obsahuje lichý počet jednotek, je medián vždy tou prostřední hodnotou. Pokud ale soubor obsahuje sudý počet jednotek, je medián mezi dvěma prostředními jednotkami a tudíž je nutné udělat z těchto dvou jednotek průměr. Před hledáním mediánu je nutné seřadit celý soubor hodnot od nejmenší jednotky po tu největší a teprve poté hledat medián. Tím je poté určen, která jednotka či který znak rozděluje celý soubor jednotek na dvě stejně velké poloviny. Například při vyhodnocování otázky, ve které byl respondent dotázán na svůj příjem, bude medián představovat konkrétní plat, který rozděluje soubor platů na dvě poloviny. První polovina má plat menší nebo stejný, než je medián a druhá polovina má plat větší nebo stejný než je medián.<sup>166</sup>

---

<sup>164</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 32 - 33. ISBN 80-86946-16-9.

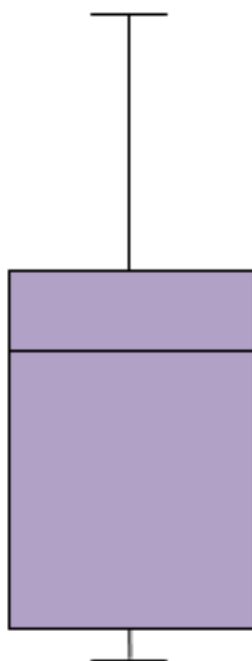
<sup>165</sup> KAISAROVÁ, Petra. *Základy popisné statistiky v pedagogickém výzkumu*. Brno, 2009, s. 36-37. Bakalářská práce. Masarykova univerzita.

<sup>166</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 27. ISBN 80-86946-16-9.

Používá se zejména v případech, kdy se v souboru jednotek vyskytují velmi extrémní hodnoty a průměr by tak nabýval neprůkazných hodnot.<sup>167</sup>

Medián se jinak nazývá také prostřední kvartil. Kvartily se dále rozděluje na horní a dolní kvartil. Horní kvartil odděluje 75 % nejnižších hodnot v souboru od zbývajících 25 %. Dolní kvartil odděluje 25 % nejnižších jednotek od zbytku hodnot v souboru.<sup>168</sup>

Graf 1: Krabičkový graf zobrazující kvantily



Zdroj<sup>169</sup>

---

<sup>167</sup> KAISAROVÁ, Petra. *Základy popisné statistiky v pedagogickém výzkumu*. Brno, 2009, s. 36-37. Bakalářská práce. Masarykova univerzita.

<sup>168</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 27. ISBN 80-86946-16-9.

<sup>169</sup> Vlastní zpracování

Graf 1 znázorňuje tzv. krabičkový graf. Odlehlé hodnoty od grafu označují minimum a maximum. Ohraničení grafu zobrazuje zmiňovaný horní a dolní kvartil. Uvnitř grafu je znázorněn medián.

Kromě kvartilů se používají ještě decily a percentily. Decily dělí soubor na deset stejně velkých částí a percentily na 100 stejně velkých částí. Tyto kvantily jsou již velmi podrobné a využívají se spíše při velmi podrobných analýzách.<sup>170</sup>

## Modus

Další mírou polohy, kterou lze využít v rámci charakteristiky sledovaných znaků, je modus. Ten označuje nejčastěji uvedenou odpověď na otázku položenou respondentovi. Modus je tedy nejčetnější kategorií v celém souboru.<sup>171</sup> Při otázce například na plat jsou tedy platy seřazeny dle četností výskytu a plat, který byl uveden nejčastěji, je zároveň modus v celém souboru.

Modus se využívá zejména v případech, kdy se porovnávají nějaké číselné údaje, tedy například právě v otázce platu. Kromě toho jej také lze využít v případech, kdy se vyhodnocují nominální či ordinální data. Tam totiž nelze použít žádný z průměrů jako charakteristiku. V případech, kdy se vyhodnocují data jako "souhlasím" "spíše souhlasím" "spíše nesouhlasím" "nesouhlasím", je modus výborná charakteristika získaných dat.<sup>172</sup>

Modus může mít různé podoby. Může se jednat o modus unimodální. To znamená, že v celém souboru je pouze jeden modus, a to právě jedna nejčetnější kategorie. Tedy například v souboru, kde odpovědělo 20 lidí, že sleduje TV NOVA, 23 lidí, že sleduje TV Prima a 60 lidí, že nesleduje žádnou televizi a 30 lidí sleduje TV SPORT, je právě tato kategorie osob nesledujících žádnou televizní stanici modem celého souboru. Unimodální data se mohou v praxi objevit běžně při statistickém šetření.<sup>173</sup>

---

<sup>170</sup> Tamtéž, s. 27.

<sup>171</sup> LÖSTER, Tomáš, Hana ŘEZANKOVÁ a Jitka LANGHAMROVÁ. *Statistické metody a demografie*. 1. vyd. Praha: VŠEM, 2008, s. 22. ISBN 978-80-86730-40-0.

<sup>172</sup> KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 27. Bakalářská práce. Západočeská univerzita v Plzni.

<sup>173</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 29. ISBN 80-86946-16-9.

Modus však může být i bimodální. Takový jev nastává v případě, kdy jsou v souboru dva vrcholy. Tedy v případě, že data vypadají například takto: 20, 23, 30, 20, 18, 23, 40, 20. Poté jsou v tomto souboru dva vrcholy, a sice čísla 30 a 40. Bimodální data mohou v praxi nastat při běžném šetření.<sup>174</sup>

Kromě toho však soubor modus obsahovat vůbec nemusí. V tom případě se jedná o tzv. amodální data. V tomto případě není žádná kategorie četnější než ostatní. Všechny kategorie tedy mají stejnou četnost. Tento jev je ale v praxi velmi vzácný a příliš často nenastává.<sup>175</sup>

## 6.2 Míry variability

Dalším způsobem charakteristiky dat jsou míry variability. Ty pomáhají určit, jak těsně či naopak jak vzdáleně jsou od sebe hodnoty získaných dat. Někdy mohou být hodnoty například těsněji koncentrovány okolo průměru, někdy mohou být naopak rovnoměrně rozděleny. Toto se nedá zjistit z běžného aritmetického průměru. Právě proto je vhodné v rámci analýzy dat využít měření variability. Na základě měřené variability pak lze zpětně posoudit, jak kvalitní a vypovídající hodnotu má aritmetický průměr.<sup>176</sup>

Obecně je však vypovídající hodnota aritmetického průměru vyšší, pokud jsou hodnoty rozděleny rovnoměrně. Čím koncentrovanější jsou hodnoty v určitém bodě, tím menší vypovídající hodnotu průměr má.<sup>177</sup>

V praxi se často využívají **míry absolutní variability**. Ty vyjadřují variabilitu sledovaného znaku ve stejných měrných jednotkách, v jakých je vyjadřován sledovaný znak.<sup>178</sup>

Nejjednodušší mírou absolutní variability je **variační rozpětí**. Vypočte se pouze odečtením nejmenší hodnoty od té největší hodnoty v celém souboru. Její výpočet je rychlý a její interpretace velmi snadná. Je to však také nejhrubší míra variability a její

---

<sup>174</sup> Tamtéž, s. 29.

<sup>175</sup> Tamtéž, s. 26 – 27.

<sup>176</sup> Tamtéž, s. 35.

<sup>177</sup> KAISAROVÁ, Petra. *Základy popisné statistiky v pedagogickém výzkumu*. Brno, 2009, s. 36.

Bakalářská práce. Masarykova univerzita.

<sup>178</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 35. ISBN 80-86946-16-9.

vypovídající hodnota je velmi malá. Nevýhodou je také možnost velmi extrémních či odlehklých hodnot v souboru. Ty se poté výrazně projeví. Variační rozpětí také nic neříká o variabilitě uvnitř celého souboru.<sup>179</sup>

Proto se častěji využívá **rozptyl**, jelikož umožňuje hlubší zkoumání variability na získaných datech. Velikost rozptylu je přímo závislá na variabilitě všech hodnot, které se v souboru nachází. Rozptyl navíc měří variabilitu nejen kolem aritmetického průměru, ale také vzájemnou variabilitu jednotlivých hodnot znaku mezi sebou.<sup>180</sup>

Rozptyl má však tu nevýhodu, že je vždy interpretován je čtvercích použité měrné jednotky. Tedy například u zkoumání platu respondentů jsou tyto platy interpretovány v Kč<sup>2</sup>. Proto se používá **směrodatná odchylka**. Směrodatná odchylka je odmocnina z rozptylu a má mnohem jednodušší interpretaci.<sup>181</sup>

Mezi další míry absolutní variability patří i kvantilové odchylky, jež definují průměr sousedních kvantilů. Obdobně lze použít i decilovou či percentilovou odchylku.<sup>182</sup>

Kromě výše uvedených metod lze také použít relativní míry variability. Ty se používají zejména v situacích, kdy hodnoty znaku nejsou uvedeny ve stejných měrných jednotkách. Použitím **variačního koeficientu** se tento problém odstraní, jelikož variační koeficient dává do poměru absolutní variabilitu a medián či aritmetický průměr.<sup>183</sup>

---

<sup>179</sup> STIEBNER, Ondřej. *Prozkoumávání jednorozměrných dat*. Olomouc, 2008, s. 22-23. Diplomová práce. Univerzita Palackého v Olomouci.

<sup>180</sup> LÖSTER, Tomáš, Hana ŘEZANKOVÁ a Jitka LANGHAMROVÁ. *Statistické metody a demografie*. 1. vyd. Praha: VŠEM, 2008, s. 24-25. ISBN 978-80-86730-40-0.

<sup>181</sup> Tamtéž, s. 26.

<sup>182</sup> HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, s. 35. ISBN 80-86946-16-9.

<sup>183</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 35. ISBN 80-245-0753-6.

### 6.3 Chyby při sběru dat

Při sběru dat, ať už pomocí dotazníku či pozorování, dochází k chybám. Mohou nastat pouze dva druhy chyb. Prvním je tzv. **náhodná chyba**, která s velkou pravděpodobností ovlivní závěry o populaci. Ta může vzniknout již při tvorbě dotazníku. Pokud je variabilita zkoumané proměnné příliš vysoká, stane se, že prošetřením jednoho výběrového souboru získáme odlišné výsledky, než prošetřením jiného souboru. V tomto případě se výsledek zjištěný ze vzorku populace neshoduje s populací. Tímto vzniká **náhodná výběrová chyba**. Tuto chybu lze snížit, pokud se zvýší rozsah výběrového souboru, tedy počet osob, které budou osloveny v rámci šetření. Zabránit jejich vzniku ale nelze.<sup>184</sup>

Druhým typem chyby je tzv. **systematická chyba**, která tyto závěry o populaci přímo zkresluje.<sup>185</sup> Především se jedná o **nevýběrovou systematickou chybu**. Ta může vzniknout v jakékoli fázi sběru dat a to především díky špatnému způsobu zjišťování dat. Tomu je možné zabránit například lepším připravením tazatele. Při odstraňování chyb ale není možné použít stejnou metodu, jako u výběrových chyb. V tomto případě by při zvětšení rozsahu výběru došlo i ke zvětšení systematické chyby. Tyto chyby lze eliminovat pouze pečlivou přípravou před samotným šetřením a následným pečlivým zpracováním.<sup>186</sup>

Nejčastěji vzniká systematická chyba **špatnou představou o chování populace**. Pokud je například měřena sledovanost určité televizní stanice ve všední dny v Praze, nelze závěr z tohoto šetření aplikovat na zbytek České republiky, jelikož v jiných částech republiky se mohou lidé chovat zcela odlišným způsobem.<sup>187</sup>

I když je populace správně vymezena, neznamená to eliminaci systematických chyb. Další riziko tkví v **nereprezentativním výběru**. Nejčastější příčinou bývá zkreslení typu **non-response**. To znamená, že tazatelovi se nepodařilo zastihnout danou osobu,

---

<sup>184</sup> KAFKOVÁ, Pavla. *Statistika a její využití (zneužití) při tvorbě informací*. Jindřichův Hradec, 2013, s. 23. Diplomová práce. Vysoká škola ekonomická.

<sup>185</sup> PEČÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 47. ISBN 978-80-86946-74-0.

<sup>186</sup> KAFKOVÁ, Pavla. *Statistika a její využití (zneužití) při tvorbě informací*. Jindřichův Hradec, 2013, s. 24. Diplomová práce. Vysoká škola ekonomická.

<sup>187</sup> PEČÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 48. ISBN 978-80-86946-74-0.

nebo tato osoba odmítla účast v průzkumu. Systematická chyba pak vzniká v případě, že se hodnoty sledovaného jevu významně liší u respondentů a u osob, které tento údaj poskytnout odmítly. Této chybě ale lze do určité míry předejít, a to zejména důkladnou přípravou před samotným šetřením. Je vhodné pečlivě zvolit čas a místo zjišťování a tomu přizpůsobit i způsob, jakým budou data zjišťována. Dále také řádně proškolit tazatele a opakovat pokusy o získání dat od respondentů. Nejčastěji tak vybraná domácnost dostane písemné upozornění o návštěvě tazatele a v případě nezastihnutí je následně kontaktována jiným tazatelem znovu. V jiném případě jsou respondenti kontaktováni telefonicky, než je jim zaslán například elektronický dotazník.<sup>188</sup>

Velmi nevhodné řešení je použít náhradníky za respondenty, které se nepodařilo zastihnout, či nechtěli odpovídat. Již tento krok může být zdrojem zkreslení. Daleko přijatelnější je již v začátku počítat s určitým procentem non-response a navýšit rozsah šetření. K tomu je ale nutné mít předem představu o podílu tohoto jevu.<sup>189</sup>

Systematická chyba může vzniknout i špatným způsobem zjišťování. V tomto případě může chyba vzniknout na straně tazatele, ale i respondenta.<sup>190</sup>

Na straně respondenta je nepochopení otázky, či neochota odpovědět. Může se také snažit ukázat v lepším světle a neodpovídat podle pravdy. Vliv na správnost údajů může mít ale i příliš dlouhý dotazník, nezajímavý obsah a především nesrozumitelné otázky.<sup>191</sup>

Zkreslení může způsobit i tazatel. Zejména jeho nedostatečná příprava a špatné chování k respondentům. Velkým rizikem jsou i tazatelské podvody, pokud nejsou dostatečně kontrolovány.<sup>192</sup>

Pokud se data povede shromáždit tak, aby eliminovali systematické chyby, neznamená to, že již žádné chyby vzniknout nemohou. Systematické chyby se mohou objevit i v konečné fázi, kdy se již zpracovávají data. Je tak nutné dávat pozor při přepisování dat

---

<sup>188</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 12-19. ISBN 04-326-80.

<sup>189</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 49. ISBN 978-80-86946-74-0.

<sup>190</sup> KAFKOVÁ, Pavla. *Statistika a její využití (zneužití) při tvorbě informací*. Jindřichův Hradec, 2013, s. 24. Diplomová práce. Vysoká škola ekonomická.

<sup>191</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 49. ISBN 978-80-86946-74-0.

<sup>192</sup> ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 12-19. ISBN 04-326-80.



z papírové podoby do elektronické. Zejména v této fázi lze chyby velmi efektivně odstranit.<sup>193</sup>

## 6.4 Doplnění dat

V situaci, kdy v sesbíraných dotaznících chybí data či odpovědi na nějaké otázky, je vhodné vybrat způsob doplnění dat. Jedná se o běžný statistický jev, a proto je dobré s ním již dopředu počítat. Možností, jak se vypořádat s chybějícími daty, je pět.<sup>194</sup>

První možností je vycházet z podílu odpovědí. To znamená, že výsledné získané odpovědi se podělí, aby se zjistilo, jaký podíl zaujímají určité odpovědi. Se stejným poměrem se poté doplní i chybějící odpovědi tak, aby odpovídaly původně zjištěným podílům.<sup>195</sup>

Druhá možnost je použít systém náhradníků. Tedy získat chybějící informace od jiných, náhradních respondentů, kteří mají stejné charakteristiky, jako předchozí respondenti.<sup>196</sup> Jak ale bylo uvedeno výše, tento systém řešení je velmi nevhodný a sám může být zdrojem zkreslení.<sup>197</sup>

Třetí možností je použít tzv. Politzovo-Simmonsovo schéma či jeho různá zdokonalení. Tento způsob je přímo určený k tomu, aby se s ním nahrazovaly nezískané odpovědi. Většinou se užívá pro případ, kdy osoba nebyla zastižena doma, tedy tzv. "not-at-home". V případě užití této metody se předpokládá, že odpověď osob, které nebyly doma zastiženy je stejná, jako těch, kteří se charakterizovali jako "málokdy doma".<sup>198</sup>

Čtvrtou možností je využití apriorních znalostí pomocí bayesovské statistiky. Ta vychází z toho, že již před provedením samotného šetření je něco známo o rozdělení četností. Což je do určité míry pravda. Například před volebním průzkumem je skutečně

---

<sup>193</sup> PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 49. ISBN 978-80-86946-74-0.

<sup>194</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 105-108. ISBN 80-245-0753-6.

<sup>195</sup> Tamtéž, s. 105-108.

<sup>196</sup> ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření: (druhé vydání)*. 2. vyd. Praha: Professional Publishing, 2010, s. 31. ISBN 978-80-7431-019-5.

<sup>197</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 105-108. ISBN 80-245-0753-6.

<sup>198</sup> Tamtéž, s. 105-108.

již něco málo známo o politických postojích. Dají se také využít výsledky z minulého šetření. Tato metoda je velmi složitá, a proto není velmi používána. Na druhou stranu je ale velmi přesná a vysoce efektivní.<sup>199</sup>

Poslední možností je pokusit se znovu o získání údajů. Tedy pokusit se respondenty znovu zastihnout a získat chybějící odpovědi. Úspěch této metody je ale čistě náhodný.<sup>200</sup>

## 6.5 Publikování dat

Výsledkem celého procesu výběrového šetření je nějaký závěr a s ním spojená i publikace výsledků, neboli dat, zjištěných na základě tohoto šetření. Publikace těchto dat se výrazně liší na základě toho, kdo je zadavatelem konkrétních projektů. Tyto zadavatele lze přehledně rozdělit do třech základních skupin.

V prvním případě se jedná o výzkum vedený z vlastní iniciativy konkrétní výzkumné instituce. Tímto způsobem například vzniká většina výzkumů v Centru pro výzkum veřejného mínění. Toto centrum zpracovává pravidelně každý rok několik šetření, která se opakují. Díky tomuto je také možné dlouhodobě sledovat vývoj určitých jevů. Závěry těchto výzkumů jsou veřejně prezentovány a jsou tak přístupny i široké veřejnosti.<sup>201</sup>

V druhém případě se jedná o výzkumy, které si zadávají soukromé firmy. Mezi tyto firmy patří například instituce státní správy, vědecká pracoviště atd.<sup>202</sup> Zde se zpravidla zkoumá ekonomická efektivita, analýza monitoringu médií atd. Tyto průzkumy jsou ve většině případů komerční a výsledky zpravidla nejsou zpřístupněny široké veřejnosti. Naopak se stávají součástí know how jednotlivých organizací, které je nadále využívají ke svým ekonomickým či jiným cílům. Někdy se však firma může rozhodnout, že

---

<sup>199</sup> PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 108. ISBN 80-245-0753-6.

<sup>200</sup> Tamtéž, s. 108.

<sup>201</sup> BEBUTOVÁ, Erika. *Spolupráce celostátních deníků a agentur pro výzkum veřejného mínění v České republice*. Praha, 2011, s. 20. Diplomová práce. Univerzita Karlova v Praze.

<sup>202</sup> Tamtéž, s. 20.

závěry výzkumu zveřejní a výzkum tak slouží pro její prezentaci směrem k veřejnosti.<sup>203</sup>

Třetí je oblast akademického výzkumu, kde se zpracovávají průzkumy týkající se například pracovních rutin, odkrývají skryté hodnotové soustavy atd. Tyto výzkumy jsou ve většině případů veřejné a snaží se své závěry publikovat a zpřístupnit co nejširšímu okruhu veřejnosti. Závěry jsou publikovány nejčastěji v odborných periodících či knihách. Tyto výzkumy jsou velmi často financovány z veřejných zdrojů a zjištění, která přinesou, slouží především k obohacení široké veřejnosti, ale i vědecké komunity.<sup>204</sup>

---

<sup>203</sup> TRAMPOTA, Tomáš a Martina VOJTĚCHOVSKÁ. *Metody výzkumu médií*. 1. vyd. Praha: Portál, 2010, s. 12-13. ISBN 978-80-7367-683-4.

<sup>204</sup> Tamtéž, s. 13.

# PRAKTICKÁ ČÁST

## 7 CÍL, OKRUHY OTÁZEK A HYPOTÉZY

Cílem práce bylo porovnat teoretické možnosti realizace výběrových šetření s metodami prakticky využívanými institucemi v ČR. Cíle této práce mají přispět k většímu povědomí o tom, zda a jakým způsobem se liší některé postupy a metody sběru dat v rámci teoretické a praktické roviny.

Diplomová práce neprověřuje všechny faktory výběrových šetření. V rámci praktické části je zaměřena pouze na některé faktory. Některé tyto teoretické poznatky byly shrnuty do následujících hypotéz.

### **Hypotéza č. 1**

*Při výběru respondentů je nejpoužívanější metoda kvótního výběru.*

Tato hypotéza byla zvolena na základě faktu, že pravděpodobnostní metody výběru jsou velmi nákladné, a to jak finančně, tak časově. Velmi těžko se sestavuje databáze, ze které je poté možno náhodný výběr provést. Kvótní výběr tak umožňuje levnější a mnohem dostupnější sběr dat.

### **Hypotéza č. 2**

*Při používání škál u výběrových šetření se upřednostňují škály se sudými počty členů.*

Škály se sudými počty členů umožňují v rámci výběru přimět respondenta, aby se přiklonil na jednu ze stran. Usnadňuje tak v mnohém následné zpracování dat. Při použití lichého počtu členů na škále hrozí nebezpečí, že respondenti se zařadí do středního členu, čímž jeho názor zaniká.

### **Hypotéza č. 3**

*Pokud se při výběrových šetřeních využívají číselné škály, jsou to nejčastěji celočíselné škály.*

Hypotéza byla zvolna vzhledem k teoretickým poznatkům, na základě kterých jsou celočíselné škály mnohdy méně finančně náročné, co se týče realizace. Celočíselné škály jsou také pro respondenta přehlednější

#### **Hypotéza č. 4**

*Nejzávažnějším problémem při sběru dat je problematika non-response.*

Hypotéza vycházela z faktu, že non-response je velmi závažná problematika a dnes, kdy se ochota lidí k vyplňování dotazníků snižuje, je riziko, že dotazník bude odmítnut. Stejně tak se snižuje ochota lidí uvádět o sobě údaje, jež považuje za citlivé, například výši příjmu. Tím tedy vznikají v dotazníku nevyplněné otázky.

#### **Hypotéza č. 5**

*Doplnění chybějících dat probíhá nejčastěji pomocí dosběru dat prostřednictvím náhradníků.*

Vzhledem k problematice non-response se následně musejí řešit i chybějící data. Použití systému náhradníků dle různých postupů by mělo být standardem.

## 8 METODIKA

### 8.1. Hloubkové rozhovory

V rámci dotazování bylo vybráno 5 odborníků, kteří následně absolvovali hloubkový rozhovor, aby popsali metody a způsoby sběru dat přímo v jejich organizaci.

Všichni tito odborníci se přímo či nepřímo podílejí na vytváření metodik či dotazníků pro účely výběrového šetření. Dále mají znalosti z praxe ohledně vytváření škál a škálovacích postupů při tvorbě dotazníků. Mají praktické zkušenosti s nejčastějšími potížemi při sběru dat a způsoby nahrazování dat v případě neúplných dotazníků či neochotě respondentů účastnit se průzkumů. Vzhledem ke svým dlouholetým praktickým zkušenostem přímo z oboru byli schopni popsat, jak probíhá výběrové šetření v praxi a tím pomoci k potvrzení či vyvrácení hypotéz v rámci této práce

Těmito odborníky byly Mgr. Pavel Šimoník, ředitel služeb klientům ve společnosti STEM/MARK, a. s., paní Ing. Kamila Burešová, pracující jako analytička ve společnosti STEM, s. r. o., pan Mgr. Jindřich Krejčí, Ph.D., který pracuje jako vedoucí oddělení a zástupce pro vědeckou a projektovou činnost Sociologického ústavu AV ČR, v. v. i., pan Mgr. Ing. Martin Zelený, Ph.D., který pracuje jako ředitel Odboru šetření domácností na Českém statistickém úřadu a pan PhDr. Stanislav Hampl, který pracuje jako manažer sociálního a politického výzkumu v PPM Factum research, s. r. o.

Pan Mgr. Pavel Šimoník vystudoval Filozofickou fakultu Univerzity Karlovy. Působil 8 let ve společnosti STEM, s. r. o., jako analytik a v současné době pracuje již 13 let jako ředitel služeb klientům ve společnosti STEM/MARK, a. s. V rámci svého působení se účastnil a řídil několik marketingových projektů, které se zaměřovaly na čtenost jednotlivých médií, sledovanost různých televizních kanálů a pořadů, výzkumy konkurenčních stanic apod. Vzhledem ke své pozici ve společnosti STEM/MARK, a. s. má přehled o způsobu výběru vzorků pro účely šetření, ale i o metodách sestavování dotazníků a způsobech nahrazování chybějících dat.

Paní Ing. Kamila Burešová pracuje jako analytička ve společnosti STEM, s. r. o.

Pan Mgr. Jindřich Krejčí, Ph.D. vystudoval Fakultu sociálních věd na Univerzitě Karlově a v současné době pracuje již 16 let jako vedoucí oddělení a zástupce ředitele pro vědeckou a projektovou činnost Sociologického ústavu AV ČR, v. v. i. Vzhledem ke svému zaměstnání má velký přehled o kvalitě výběrových šetření a o způsobech provádění těchto šetření, stejně tak o kvalitě sbíraných dat a metodách výzkumu.

Pan Mgr. Ing. Martin Zelený, Ph.D. vystudoval Statistické a pojistné inženýrství na Vysoké škole ekonomické v Praze, Veřejnou a sociální politiku na Fakultě sociálních věd na Univerzitě Karlově a poté Statistiku na Vysoké škole ekonomické v Praze. V současné době působí již 15 let na Českém statistickém úřadu, kde momentálně pracuje jako ředitel Odboru šetření domácností. Vzhledem ke své pracovní pozici má široké zaměření a zkušenosti z oblasti výběrových šetření a jejich metodik a má široké znalosti z teoretické a praktické oblasti tohoto oboru.

Pan PhDr. Stanislav Hampl, který pracuje jako manažer v oblasti sociálního a politického výzkumu ve společnosti PPM Factum research, s. r. o. Působil také 9 let ve společnosti STEM, s. r. o., kde se zabýval především politickými průzkumy a to jak na zakázku médií, tak i politických stran. Dále také prováděl šetření pro účely Veřejné správy. Vzhledem k jeho dosavadním zkušenostem v praxi má velký přehled o způsobu provádění výběrových šetření, jakožto i způsobech škálování a metodikách výběru dat.

Otázky pro účely rozhovoru byly tvořeny na základě výzkumných hypotéz, které jsou uvedeny výše. Každá hypotéza měla v rozhovoru své zastoupení v podobě minimálně jedné, ale i více otázek a to tak, aby bylo dosaženo co nejobsáhlejší odpovědi na danou problematiku a hypotéza tak mohla být s jistotou potvrzena či vyvrácena.

Data a informace byly sbírány formou expertních rozhovorů. Vzhledem k charakteru otázek byly zvoleny expertní rozhovory s odborníky na danou problematiku. Tato forma byla shledána jako nejpraktičtější a nejúčelnější vzhledem k odbornosti otázek a pravděpodobné neznalosti této problematiky běžnými respondenty .

Respondenti byli získáváni na základě svých zkušeností a pozic, které zastávali v jednotlivých organizacích, které se zabývali určitým typem marketingového výzkumu, výběrového šetření či sběru tak. Všichni oslovení odborníci souhlasili s žádostí

o rozhovor pro účely diplomové práce. Na základě předchozího kontaktu s jednotlivými odborníky byl následně domluven rozhovor.

Expertní rozhovory probíhaly na základě předchozí domluvy s jednotlivými odborníky a vždy osobně. V rámci osobního rozhovoru byly vždy pokládány otázky tak, aby bylo dosaženo co nejdetailnější odpovědi na danou problematiku a na danou hypotézu. Vzhledem k tomu, že všichni odborníci souhlasili s expertním rozhovorem, byla návratnost 100 %.

Ověřování hypotéz probíhalo na základě analýzy dat získaných z provedených rozhovorů.

## **8.2 Harmonogram výzkumu**

Oslovování odborníků na zmíněnou problematiku probíhalo v rámci měsíce prosince 2014, kdy byli osloveni jednotliví experti vždy na základě e-mailové komunikace, nebo prostřednictvím telefonního kontaktu se žádostí o rozhovor pro účely diplomové práce.

Sběr dat probíhal v rámci měsíce ledna 2015 formou expertních rozhovorů. Na základě předchozí domluvy s konkrétními odborníky byly prováděny rozhovory. Tyto rozhovory byly prováděny vždy s každým odborníkem zvlášť a to v takové míře, aby v rámci tohoto rozhovoru odborníci odpověděli na všechny otázky, které byly pro účely diplomové práce podstatné.

Kromě toho byl odborníkům vždy ponechán prostor i na vlastní názory či připomínky, nebo i na další zkušenosti z praxe, které přímo či nepřímo souvisely s otázkami a hypotézami.

Podmínkou pro provedení rozhovoru a sběr dat byly znalosti z praxe, které museli mít všichni, kdo byli v rámci těchto rozhovorů tázáni. Oslovení odborníci se museli přímo či nepřímo podílet na sestavování dotazníků či na formulování podmínek pro sběr dat ve své organizaci, nebo mít přehled a znalosti z praxe, které by mohli formou tohoto rozhovoru poskytnout.



## 9 VÝSLEDKY VÝZKUMU

### 9.1. Nejužívanější typy sběru dat

Jaké typy sběru dat se vůbec užívají v praxi? Hloubkové rozhovory ukázaly, že nejčastěji se využívá zejména kvótní výběr. Podle Pavla Šimoníka, ředitele služeb klientům ve společnosti STEM/MARK, a. s. je jedním z důvodů je neexistující přístup do databáze s adresami, který neexistuje již od druhé poloviny 90. let.

Kamila Burešová, analytička ze společnosti STEM toto potvrzuje: *„Ve STEMu se nejčastěji při kvantitativních průzkumech pracuje s metodou výběru prostřednictvím kvótního výběru, kdy je výsledný soubor respondentů reprezentativní za celou populaci ČR starší 18 let podle kvótních znaků pohlaví, věk, vzdělání, velikostní pásmo obce, region.“*

Shodný názor má i pan Stanislav Hampl, manažer sociálního a politického výzkumu v PPM Factum research, s. r. o. *„V praxi nepoužívanější je jednak kvótní výběr (používá se např. u pravidelného omnibusového šetření, dále u řady regionálních či lokálních šetření).“*

Pavel Šimoník k tomu dodává: *„Pokud se dělá nějaký náhodný výběr, tak se dělá tím způsobem, že se adresy musí sepsat na základě několika kroků, kdy se používá například databáze České pošty, nebo se používá databáze Českého statistického úřadu. Tam se vyberou domy a vchody a v těchto vchodech se zjistí počet bytů. A pak se nějakým náhodným krokem musí udělat soupis adres. Je to finančně velmi nákladné a v současných podmínkách má tento výběr velmi malou návratnost. Používá se ale například u Media projektu, kde se měří čtenost tisku, u kontinuálního výzkumu pro měření sledovanosti televize atd. A někdy se to využívá při výzkumech pro akademický ústav, kde se provádí mezinárodní výzkumy. Tam si oni přímo zadají, že si přejí náhodný výběr a v těchto případech my jej provádíme. Ta návratnost se pohybuje pod 50 %, jelikož nedostupnost domácností je obrovský problém. A v některých metodikách převzatých ze zahraničí je požadavek na nemožnost substituce domácností, které se nepodaří zastihnout. Tudíž se pak velmi těžko dosahuje nějaké návratnosti, musí se dělat další kola výběru. U kontinuálních výzkumů, které se používají pro sledovanost*

*médií, se může pracovat se systémem náhradních domácností. To je ale složitější proces. Navíc tento výběr stojí dvakrát více než kvótní výběr.“* S tím souhlasí i Kamila Burešová, podle které se ve STEMu používá například systematický náhodný výběr velmi málo. A to ze stejného důvodu, a sice velká náročnost a velká finanční nákladnost.

Anketa se například podle Pavla Šimoníka nepoužívá vůbec, nebo jen v určitých krajních případech. Důvodem je zejména její velmi nízká vypovídající hodnota. *„Je to ovlivněno efektem toho, kdo chce odpovídat. Jedná se o velmi aktivní typ lidí, kteří něco velmi chválí a jsou velmi spokojení. Chybí tam ale mainstream<sup>205</sup> těch běžných lidí, kteří na toto nereagují. Ankety se používají spíše při nějakých specifických akcích, kde lidé mohou něco vyplnit. Média si mohou udělat svoje ankety. V klasickém výzkumu ale anketa nemá moc co dělat.“* S tím souhlasí i Jindřich Krejčí, vedoucího oddělení a zástupce pro vědeckou a projektovou činnost Sociologického ústavu AV ČR, v. v. i. Dle něj se nejedná o výběrové šetření a zaměřovat anketu s výběrovým šetřením je dokonce v rozporu s kodifikovanými pravidly výzkumných asociací.

Snowball se používá ve velmi specifických případech. Pokud je zadán výzkum, v rámci kterého je nutné najít 800 osob, které pečují o osobu s příspěvkem na péči, nebo například osoby z určité komunity, například cizince, je snowball ideální volbou. Jindřich Krejčí dodává, že evidence těchto lidí neexistuje a je velmi složité se k nim jinak dostat. Je tedy vhodný zejména v případech, kdy není k dispozici žádná opora výběru. Nikoli však pro obecné šetření populace.

Podle Jindřicha Krejčího však záleží metoda výběru na konkrétní výzkumné situaci. Při výběrových šetřeních v Sociologickém ústavu se většinou liší cílová skupina i jednotka výzkumu. Tudíž zde je k dispozici jiný okruh pro výběrové šetření v organizaci a jiný například pro specifické skupiny v celé společnosti. *„Pokud se ale bavíme o reprezentativních šetřeních populace v ČR, které se provádějí na základě osobních rozhovorů face-to-face, tak dominuje kvótní výběr. V akademickém sektoru nicméně dominují různé postupy pravděpodobnostních výběrů, zpravidla několikastupňový stratifikovaný výběr. Prostý náhodný výběr v běžné praxi šetření populace realizovat nelze, jelikož není dostupný registr osob.“*

---

<sup>205</sup> Střední proud lidí, kteří nereagují na ankety a nikdy se jich neúčastní.

Jiná situace ale panuje na Českém statistickém úřadu. Pan Martin Zelený, ředitel Odboru šetření, uvádí: „V *"oficiální statistice"*, tj. v doméně statistických úřadů a institucí, jednoznačně dominují pravděpodobnostní formy výběru. Spíše zřídka a u specifických šetření se lze setkat s kvótní formou výběru. U šetření v domácnostech se sběrem dat tazateli v terénu se většinou provádí výběr dvou- či vícestupňově, tj. na prvním stupni se vybere určitý počet tzv. primárních výběrových jednotek (primary sampling units - PSU<sup>206</sup>) a v nich se poté vybírají byty nebo jednotlivci pro šetření. Cílem je větší koncentrace výběru v terénu a časově a nákladově efektivnější práce tazatelů - i když tato forma vícestupňového výběru má zpravidla o něco horší statistické vlastnosti (pro stejný počet vyšetřených respondentů produkuje odhady s větší variabilitou, vyjádřenou např. intervaly spolehlivosti). Častá je i tzv. stratifikace - tj. výběr se předem ve fixních proporcích rozdělí do podskupin základního souboru, tzv. strat - např. podle krajů, velikosti obcí, atp.“

Toto potvrzuje i Stanislav Hampl: „Stratifikovaný vícestupňový pravděpodobnostní (náhodný) výběr se používá zejména pro výzkumy prováděné pro akademickou sféru nebo obecně pro klienty, kteří mají vyšší nároky na metodologickou stránku šetření.“

Ani anketa či snowball se zde nevyužívají v tzv. oficiální statistice. A to z důvodu, že by nevyhověly základním kvalitativním požadavkům na výběrová šetření, která Český statistický úřad má.

Naopak ale zde má zastoupení i vyčerpávající šetření, kdy se prošetří všechny subjekty. Jak ale Martin Zelený dodává, je to pouze v tom případě, že je jednotek málo, například několik desítek.

## 9.2. Sudé versus liché škály

Průzkum pomocí hloubkových rozhovorů ukázal, že situace se škálami není zdaleka jednoznačná. Podle Pavla Šimoníka panovala taková situace vždy. „Škály jsou velmi diskutované téma od doby, co vznikly. Neustále se vedou debaty o tom, jaké škály

---

<sup>206</sup> Primární výběrové jednotky vybírané ve vícestupňovém náhodném výběru. Z nich se poté vybírají konkrétní jednotky pro účely výzkumu.

*použít. Zda jsou lepší sudé či liché, krátké či dlouhé škály. Zda čtyřstupňové, pětistupňové nebo desetistupňové.*“

Dodává ale, že zde existují určitá pravidla. Pokud se přejímají určité metody z metodik používaných v zahraničí, vždy se užívají delší škály a to pro hodnocení čehokoliv, ať se jedná o hodnocení spokojenosti, nebo důležitosti, apod. Zejména velké zahraniční firmy si při zadávání průzkumu přejí široké škály. V těchto případech se vždy jedná o škály sudé, a to na stupnici 1 až 10. Tyto škály se využívají i v Sociologickém ústavu, kde se poté zkoumá i souvislost mezi jednotlivými proměnnými.

*„Pro běžné sociologické šetření se spíše používají škály 1 - 4. Jedná se o typickou ordinální škálu, kde je "velmi spokojen" spíše spokojen" "spíše nespokojen" "velmi nespokojen". „ To potvrzuje i Kamila Burešová ze STEMu. „Pro zjištění míry souhlasu nebo třeba spokojenosti se ve STEMu nejčastěji používají škály se sudým počtem variant bez přímého nabízení možnosti „nevím“ (samozřejmě ji respondent uvést může a tazatel takovou odpověď zapíše, někdy je tato odpověď kvůli povaze otázky přímo nabízena).*“

Lichý počet škál je však také velmi využíváný. Hrozí zde ale riziko, že respondenti označí prostřední odpověď jako střední cestu. Podle Pavla Šimoníka pak nastává problém. *„Je to pro výzkumy spíše limitem a nikoli výhodou. Potřebujeme z člověka dostat nějaký názor a postoj. Pokud mu pomocí liché škály umožníme únik, tak jej často využije. Spoustu lidí samozřejmě nemá na danou věc názor a pak je lepší nechat je odpovědět "nevím" " neumím posoudit". Škály jsou tedy obecně problémem. Neplatí však pravidlo, že by se liché škály nesměly používat.*“

V Sociologickém ústavu se druhy škál posuzují u každé otázky zvlášť. Vždy záleží na tom, co je pro konkrétní otázku důležitější, respektive jestli je důležitější dostat z respondenta názor, nebo naopak je důležité zjistit i střední hodnotu. Univerzální doporučení tu neexistuje. S tím se shoduje i Martin Zelený, který dodává, že jistě záleží i na počtu odpovědí, charakteru otázky a významu jednotlivých možností, včetně jejich slovního popisu.

Potvrzuje to i Stanislav Hampl. *„Používají se škály se sudým i lichým počtem, záleží na účelu (jestli jde o to, aby se názory rozdělily na pozitivní a negativní, nebo nás zajímají*

*právě i ty neutrální). Vždy je třeba si být vědom rizika středové kategorie (že může sloužit jako úniková). Používají se tedy oba druhy, zhruba půl napůl. “*

### 9.3. Používání číselných škál

Číselné škály mají v průzkumech a výběrových šetřeních své zastoupení. Na základě hloubkových rozhovorů jsou nejpraktičtější celočíselné škály. Z nich se nejčastěji využívají škály známkovací, kde respondent udělí známku jako ve škole. Pavel Šimoník toto řešení považuje za ideální. *„Lidem je to velmi blízké a umí si to zařadit. Zámka 3 není střed škály, protože 3 už není dobrá známka.“* I ve STEMu se dle Kamily Burešové nejčastěji uchylují k číselným škálám v podobě školních známek.

Zde se však používají i jiné škály. *„Používá se také výběr mezi dvěma protikladnými charakteristikami, výroky (např. sémantický diferenciál, příklad: dvojice „moderní – zastaralý“, škála od 1 do 7, pokud se respondent jednoznačně přiklání k variantě moderní, vybírá 1, pokud mu zcela vyhovuje varianta zastaralý, volí 7, neutrální názor je 4).“*

Celočíselné škály se využívají i ve společnosti Factum. *„Používají se nejvíce celočíselné škály. Číselné škály u postojových otázek se používají a většinou právě s lichým počtem. Ani sudý počet ale u číselné škály neeliminuje tendenci zařadit se do středu. Např. u škály 1-10 bývá respondenty jako střed vybírána pětka, i když středem není. „*

V Sociologickém ústavu jsou realizovány průzkumy veřejného mínění, které slouží k prezentaci v médiích. Proto musí mít škály popsány všechny body, aby každá výseč například na koláčovém grafu měla nějaký význam. *„Pokud už se používají pouze číselné, resp. škály s popsány pouze krajními body, pak téměř vždy s celými čísly, někdy od 0, někdy od 1, zcela výjimečně např. -5 až +5. Výjimečně také procenta („kolik procent podle Vás...).“* Procentuální škály se výjimečně využívají i ve STEM/MARKu, například v oblasti politických průzkumů.

Naopak na Českém statistickém úřadu se škály nepoužívají vůbec. Pokud už se přikročí k užití škál, je to spíše v otázkách příjmu respondentů, například jakou zátěž představují pro respondenta náklady na bydlení. Vzhledem k faktu, že zde se neprovádějí

sociologické či marketingové průzkumy, nevyužívají se zde o hodnotící otázky s možnostmi variant.

## 9.4. Problémy při sběru dat

Hloubkové rozhovory ukázaly, že výběrový šetření mají i své problémy. Nejčastěji se jedná o potíže při sběru dat, nikoli při jejich následném zpracovávání. Podle Pavla Šimoníka je největším problémem dostupnost domácností. *„To je mnohem větší problém, než jakékoli metodické problémy. Nalezení a korektní dotázání respondenta je největším problémem. Dnes bojujeme s tím, že lidé zavírají dveře tazatelům před nosem, nechťejí je pouštět do domů. Poměrně efektivní jsou telefonní výzkumy, kde respondent je ochotný na kratší dobu absolvovat rozhovor. Ale u face to face je největším problémem nevyšetřená populace. Máme sice kvótu nebo uděláme náhodný výběr, ale unikne nám 90% populace, která není ochotná se s námi bavit. Poté se to cykluje atd.“*

Stejný problém potvrzuje i Kamila Burešová, podle které se navíc ochota respondentů k účasti na výzkumu velmi snižuje.

Podle Jindřicha Krejčího nastává daleko větší problém u non-response. Označuje jej dokonce na nejzávažnější problém, ačkoli dodává, že problémů při sběru dat je více. *„Jinak je problém samozřejmě hlavně formulace otázky a odpovědi. Při výběru postupu (firmy pro) sběr dat se hledí především na postup výběru (jaká je opora výběru, kolik výběrových kroků a jaké, kolik dotazníků dělá jeden tazatel, jaká je kontrola tazatelů atd.)“*

Stejný názor zastává i Martin Zelený z Českého statistického úřadu. Non-response rovněž označuje na nejvýznamnější faktor v rámci výběrových šetření.

Kromě non-response se i Český statistický úřad potýká s řadou dalších potíží. *„Kromě toho je významná věc i tzv. measurement error<sup>207</sup> - tedy situace nepřesných nebo nesprávných odpovědí respondenta - ať již záměrně (zatajování / podhodnocení spotřeby např. u alkoholu, cigaret, neúplné údaje o příjmech atd.) nebo nezáměrně - respondent neví, nevzpomene si na všechny případy nebo naopak do sledovaného*

---

<sup>207</sup> Nepřesné či nesprávné odpovědi respondenta, které následně zkreslují výsledky průzkumu.

*období zařadí i případy vyskytnuvší se předtím, nebo vůbec nepochopí, na co se ho otázka ptá.“*

## **9.5. Doplnování dat**

Potíže s výběrovým šetřením se následně musejí nějakým způsobem řešit. Chybějící údaje se musí určitým mechanismem nahradit, nelze je nechat nevyplněné. Ve STEM/MARKu využívají dvě základní metody k doplnění chybějících dat. *„Používají se buď koeficienty navýšení, nebo došetření. Pak se musí použít systém náhradníků a udělat další kolo výběru. Znovu poslat tazatele na místo, aby se pokusili zastihnout domácnosti. Je to ale velmi drahé a je snaha se tomu vyhnout.“* Je zde tedy velká preference koeficientů navýšení a to hlavně z důvodu velké finanční nákladnosti v případě došetřování populace.

Pavel Šimoník a Kamila Burešová se shodují, že někdy ani data jindy dosbírat nelze. *„Pokud se jedná například o kontinuální výzkum, který je zaměřen na sledování čtenosti novin a časopis, nelze data dosbírat jindy. Tam se sbírají data v měsíčních či čtvrtletních cyklech a není možnost je dosbírat jindy. Tam se to pak musí hlídat už v průběhu sběru. Na to už dnes existují přístroje a programy, které monitorují stav toho sběru. Počty navštívených, odmítnutých atd.“* U kvótního výběru ve STEMu není ani substituce povolena.

V Sociologickém ústavu však řeší situaci nadvýběrem. *„Doplnění dat probíhá nadvýběrem, tedy dodatečným došetřováním populace. U kvótního šetření dělá výběr v podstatě tazatel, který má předepsání, kolik osob musí oslovit dle zadaných charakteristik. Když ho někdo odmítne, tak v rámci předepsané výběrové procedury osloví jiného člověka se stejnou charakteristikou. Data pak chybí jen v případě, že tazatel úkol nesplní. To může být dáno i tím, že úkol je nastaven nerealisticky.“*

Stejným způsobem se řeší doplnění dat i ve společnosti Factum.

Na Českém statistickém úřadě probíhá doplnění dat dle přísných pravidel a na základě protokolů. *„Ve standardu zpráv o kvalitě pro jednotlivá šetření se zpravidla vyžaduje protokolovat, jaké jednotky byly jakými nahrazeny - tj. kontrolovat celý proces tak, aby*

*nedocházelo k nekontrolované systematické náhradě jednoho typu respondentů jinými, ochotnějšími (např. typy lidí, kteří jsou častěji doma). Jinak hrozí systematické vychýlení výsledků. Konkrétně u nás substituce není povolena a pracujeme s pevným výběrem.“*

Existuje zde i speciální metoda pro případ, že chybějících údajů není mnoho. Jmenuje se hot-deck.<sup>208</sup> „Náhodným výběrem vybereme "dárce" s údaji podobnými imputovanému respondentovi. Další specifickou oblastí jsou imputační modely na rozklíčování položek udaných respondentem pouze ve formě úhrnu - např. situace, kdy respondent uvede celkové náklady na bydlení, ale ne konkrétní položky na teplou vodu, topení atp. Nebo náklady na zájezd jako celek, bez položek na dopravu a ubytování. Tady jde o modely "na míru" konkrétní situaci, s využitím portfolia pomocných proměnných z dotazníku - např. kam zájezd byl, na kolik dní, jakým dopravním prostředkem, atp.“

---

<sup>208</sup> Speciální metoda pro doplnění chybějících údajů. Využívaná zejména v případech, kdy chybějících dat není velké množství.



## 10 INTERPRETACE VÝSLEDKŮ A DISKUZE

Na počátku praktické části bylo stanoveno 5 hypotéz, které byly prověřovány pomocí hloubkových rozhovorů s odborníky. V této části práce bude rozvedeno, zda a do jaké míry se tyto hypotézy potvrdily. Zároveň bude tato část práce sloužit pro diskuzi nad výsledky.

### Hypotéza č. 1

*Při výběru respondentů je nejpoužívanější metoda kvótního výběru.*

Z hloubkových rozhovorů vyplynulo, že kvótní výběr je v jednotlivých organizacích jednoznačně používán, ačkoli se jedná o nepravděpodobnostní typ výběru.

Jedním z důvodů je jeho častého používání je finanční faktor. Tato metoda vyjde v porovnání s pravděpodobnostní metodou náhodného výběru až o polovinu levněji. V případě použití kvótního výběru také odpadá potíže se sestavováním databáze obyvatel, která dnes prakticky neexistuje a velmi těžko se dává dohromady.

Výhoda kvótního výběru spočívá i v jeho návratnosti. Na rozdíl od náhodného výběru, kde se návratnost pohybuje pod 50 %, u kvótního výběru může tazatel vybírat, dokud kvótu nenaplní.

Kvótní výběr se nejčastěji využívá v rámci pravidelných omnibusových šetření, v regionálních, ale i lokálních průzkumech a při různých typech marketingových průzkumů. Nejčastěji pracují odborní pracovníci s kvótami na věk, vzdělání, velikosti pásma bydliště, obec a případně region.

Celkem čtyři z pěti dotázaných odborníků uvedli, že kvótní výběr v jejich organizaci jednoznačně dominuje. Pouze na Českém statistickém úřadě jednoznačně dominuje pravděpodobnostní typ výběru. Jedná se většinou o prostý náhodný výběr či vícestupňový výběr. Důvodem je, že v tomto případě se jedná o jakousi oficiální statistiku a data z Českého statistického úřadu se dále používají k různým zpracováním. V této organizaci mají také přístup k databázi obyvatel, na rozdíl od jiných typů organizací.

**Vzhledem k výše uvedenému je jednoznačné, že první hypotéza se potvrdila.**

### **Hypotéza č. 2:**

*Při používání škál u výběrových šetření se upřednostňují škály se sudými počty členů.*

Během hloubkových rozhovorů vyšlo najevo, že situace ohledně škál není zcela jednoznačná. Odborníci ze STEMu a STEM/MARKu se shodují, že v jejich organizaci jsou nejpoužívanější typy škál ty se sudými počty členů. Délka škály pak záleží na konkrétní situaci. Pokud se například přejímají metodiky ze zahraničí, či se vytváří průzkum pro zahraničního klienta, pracuje se obvykle se sudými a delšími škálami 1 - 10. Naopak pro běžné sociologické či marketingové výzkumy se častěji používají kratší škály o velikosti 1 - 4.

Běžně se ale využívají i slovní škály, kde figurují výroky jako "Spokojen", "Spíše spokojen", "Spíše nespokojen", "Nespokojen". V případě širší škály jsou pak tyto výroky více rozděleny, aby si respondent vybral citlivěji, s jakým výrokem se ztotožňuje.

Naopak odborníci ze Sociologického ústavu a Českého statistického úřadu se shodují na tom, že není jednoznačná odpověď na to, jaké škály jsou u nich nejpoužívanější. Liché škály zajistí větší spolehlivost a pokryjí i část názorů, které spadají přesně doprostřed škály. Naopak sudé škály umožní rozdělit názory na dva protichůdné, ale existuje zde velké riziko, že prostřední kategorie přitáhne velký počet respondentů. To se stává v případech, kdy respondent nechce ukázat vyhraněný názor a raději zvolí střední cestu, nebo příliš nerozumí problematice a nemá konkrétní názor. Záleží tedy na každé konkrétní otázce.

Ve společnosti FACTUM jsou obě metody používány se stejnou intenzitou a to právě dle toho, zda je potřeba tyto názory rozdělit, nebo je naopak pro výzkum důležitá i střední kategorie s neutrálními názory.

**Vzhledem k tomu že dva z odborníků uvedli, že používají sudé typy škál, dva z odborníků uvedli, že nedovedou rozhodnout, jaké typy škál se v jejich organizaci**

**používají častěji, a jeden odborník uvedl, že používají škály půl napůl, nelze v případě této hypotézy rozhodnout.**

### **Hypotéza č. 3**

*Pokud se při výběrových šetřeních využívají číselné škály, jsou to nejčastěji celočíselné škály.*

Z hloubkových rozhovorů vyplynulo, že celočíselné škály se jednoznačně využívají. Zejména pak pro hodnocení v oblasti spokojenosti zákazníků s určitou službou, či obecně spokojenosti zákazníků.

Velmi často se v rámci celočíselných škál využívá systém školních známek. V rámci tohoto hodnocení spokojenosti zákazníka ohodnotí zákazník službu na stupnici 1 - 5 jako ve škole. Lidem je tato metoda velmi blízká a umějí si ji zařadit daleko snáze, než jinou bodovací škálu. Navíc hodnoty získané pomocí systému známkování bývají trochu odlišné. Zámka 3 zde nefiguruje jako prostřední hodnota. Zákazníci ji často vnímají již jako špatné hodnocení, nikoli jako střed stupnice.

Běžně se ale využívají i číselné škály 1 - 4, nebo 1 - 10. Zde panuje situace, že při použití škály 1 - 10 se respondenti více řadí na páté místo, které vnímají jako onen střed, ačkoli se o střed stupnice nejedná.

Pouze na Českém statistickém úřadě se číselné škály nevyužívají vůbec. Důvod jejich nepoužívání v této organizaci je prostý. Český statistický úřad se nezabývá sociologickými ani marketingovými výzkumy, tudíž tyto škály nepotřebuje využívat.

**Vzhledem k faktu, že se 4 odborníci shodují na tom, že v případě využívání číselných škál se využívají nejvíce celočíselné škály, tato hypotéza se potvrdila.**

### **Hypotéza č. 4**

*Nejzávažnějším problémem při sběru dat je problematika non-response.*

Z hloubkových rozhovorů vyplynulo, že dotazovaní odborníci považují problematiku non-response za skutečně závažnou. Někteří jej považují za nejvýznamnější faktor celého šetření.

Rovněž se také potýkají s velkou částí ne-kontaktu. To je situace, kdy se nepodaří kontaktovat nikoho, například při průzkumech, které se provádějí na konkrétních adresách. Velmi špatnou návratnost a nezastižení respondenta považují za zásadní zejména odborníci ze STEMu a STEM/MARKu.

Velmi se také snižuje ochota respondentů k provedení rozhovoru. Často je tak velmi malá a špatná návratnost jednotlivých dotazníků.

**Vzhledem k faktu, že 3 odborníci uvedli jako nejzávažnější problém problematiku non-response, tato hypotéza se prokázala.**

### **Hypotéza č. 5**

*Doplnění chybějících dat probíhá nejčastěji pomocí dosběru dat prostřednictvím náhradníků.*

Na základě provedených rozhovorů se většina odborníků shoduje, že dosběrem, či jinak také systémem náhradníků, se řeší většina chybějících dat. To probíhá různými metodami. Jednou z nich je například metoda hot-deck, kdy je v rámci dodatečného došetřování vybrán touto metodou náhodný dárce, který má charakteristiky podobné respondentovi, jež odmítl odpovídat. Tato metoda se využívá zejména tam, kde jsou minimální procenta chybějících údajů.

Ovšem v některých případech toto použít nelze. Pokud se jedná o kontinuální výzkumy a zkoumá se například čtenost jednotlivých periodik či časopisů, sbírají se data v určitých cyklech. Po dokončení tohoto cyklu již není možné data dosbírat, pokud chybí.

V některých typech průzkumů ani není povoleno doplňování dat.

Ve STEM/MARKu naopak používají nejvíce koeficienty navýšení, a to zejména kvůli jejich finanční nenáročnosti a naopak velké finanční náročnosti dodatečných došetření v terénu. Zde panuje snaha se dodatečnému dosběru vyhnout.

**Vzhledem k tomu, že 4 odborníci se shodli, že chybějící data se získávají dodatečným dosběrem, tato hypotéza se potvrdila.**

Teoretických možností při realizaci výběrového šetření je mnoho. V praxi jich však až tolik používaných není. Na základě ověřených hypotéz lze říci, že v oblasti výběru vzorku dominuje ve vybraných organizacích kvótní výběr. Ostatní výběry jsou také používány, ovšem nikoli v takové míře. Velmi užívaný je i pravděpodobnostní výběr a jeho podskupiny jako vícestupňový výběr, ovšem zejména v oblasti akademických výzkumů, kde tyto výzkumy slouží jako podklady pro další oblasti bádání. Své místo má v praxi i snowball, který je používán zejména v případech, kdy není možnost dostat se k určité části populace, jako jsou osoby, které pečují o nějakou osobu, homosexuálové apod., jak bylo uvedeno v podkapitole 3.2. Velmi málo oblíbená je anketa, a to právě pro své nereprezentativní výsledky. Další metody výběru vzorku jsou již užívány v naprosto minimální míře, potažmo vůbec.

Rovněž se při realizaci výběrových šetření uplatní i škálování, a to jak se sudým, tak lichým počtem. Délka škály je odvislá od typu průzkumu a závisí často na jejím zadavateli. Zároveň jsou často používány i celočíselné škály, sémantický diferenciál a procentuální typy škál. Ty patří mezi nejhojněji využívané. Naopak skalogram či metoda párových srovnání nebyla jmenována vůbec, a mají tedy mizivé zastoupení. Stejně tak i Likertova škála nebyla odborníky zmiňována mezi preferované druhy škál.

Potíže při sběru dat jsou zapříčiněny nejvíce problematikou non-response. Téměř všichni odborníci se shodují, že jde o nejzávažnější chybu, která je dokonce závažnější, než metodické pochybení. Chyby, které vznikají špatnou představou o populaci či náhodné chyby tedy nejsou zdaleka tak časté, jako právě non-response, se kterým se dnešní marketingové a sociologické průzkumy potýkají.

## ZÁVĚR

Cílem práce bylo porovnat teoretické možnosti realizace výběrových šetření s metodami prakticky využívanými institucemi v ČR. Cíle této práce mají přispět k většímu povědomí o tom, zda a jakým způsobem se liší některé postupy a metody sběru dat v rámci teoretické a praktické roviny.

Vzhledem k výše uvedenému se lze domnívat, že cíle práce byly splněny. Teoretické možnosti uvedené v teoretické části práce byly formou hypotéz konfrontovány s praktickým využitím v různých organizacích v ČR, které se zaměřují na výběrová šetření. Na základě těchto hypotéz je možno všimnout si, že ne všechny metody výběru jsou v praxi využívány a některé mají prioritu před jinými.

Na základě těchto výsledků je možné udělat si základní představu o využívání teoretických možností, které nabízí statistika v oblasti výběrových šetření s těmi, které jsou skutečně v praxi využívány.

Pro oblast dalšího zkoumání je velice zajímavá problematika measurement error, která významně ovlivňuje výsledky výběrových šetření a marketingových, ale i jiných průzkumů. Jedná se o problém s nesprávnými odpověďmi respondenta. Respondent tyto odpovědi může uvést špatně záměrně, či nezáměrně. V případě záměrného uvedení špatné odpovědi se jedná například o podhodnocení spotřeby alkoholu, cigaret atd. Může se jednat i o zkreslené údaje ohledně příjmu, jelikož zde stále panuje obava sdělovat svůj příjem. V případě nezáměrného uvedení špatné odpovědi se jedná o situace, kdy respondent neví, jak má odpovědět, nevzpomene si na všechny případy, které by měl zmínit, či do sledovaného období zařadí i nějaká data navíc.

Tento jev je velmi zajímavý a v rámci dalších průzkumů by mohl být prověřován, např. na kolik je tento jev běžný, a v jaké míře ovlivňuje celkové výsledky průzkumu.

Velmi zajímavá je i problematika tazatelů, kteří mají možnost ovlivnit charakter sbíraných dat. Dle zkušeností odborníků se často stávalo, že tazatel si odpovědi a respondenty vymýšlel. Díky moderním technologiím, které nahrávají i rozhovory face to face, se na mnohá odhalení přišlo. Stále však existují způsoby, jak při sběru dat podvádět a ovlivnit tím výsledek celého šetření.

## SEZNAM POUŽITÝCH ZDROJŮ

### Seznam použitých českých zdrojů

BÁRTOVÁ, Hilda. *Marketingový výzkum trhu*. 1.vyd. Praha: Economia, 1991, s. 107. ISBN 80-85378-09-4.

BEBUTOVÁ, Erika. *Spolupráce celostátních deníků a agentur pro výzkum veřejného mínění v České republice*. Praha, 2011, s. 20. Diplomová práce. Univerzita Karlova v Praze.

ČERMÁK, Václav. *Výběrová statistická zjišťování*. 1. vyd. Praha: SNTL, 1980, s. 11. ISBN 04-326-80.

ČERMÁK, Václav. *Vybrané kapitoly z teorie výběrových šetření*. 1. vyd. Praha: Státní pedagogické nakladatelství, 1974, s. 5. ISBN 17-131-74.

ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 1*. vyd. 1. Praha: Vysoká škola ekonomická v Praze, 1999, s. 5. ISBN 80-707-9191-8.

ČERMÁK, Václav a Michal VRABEC. *Teorie výběrových šetření: Část 3*.1.vyd. Praha: Vysoká škola ekonomická, 1998-1999, s. 5-8. ISBN 80-245-0003-53.

HAGUE, Paul N. *Průzkum trhu: [příprava, výběr vhodných metod, provedení, interpretace získaných údajů]*. 1. vyd. Brno: Computer Press, 2003, s. 68. ISBN 80-7226-917-8.

HINDLS, Richard, Stanislava HRONOVÁ, Jan SEGER a Jakub FISHER. *Statistika pro ekonomy*. 7. vyd. Praha: Professional Publishing, 2006, 22 - 23. ISBN 80-86946-16-9.

KAFKOVÁ, Pavla. *Statistika a její využití (zneužití) při tvorbě informací*. Jindřichův Hradec, 2013, s. 21-22. Diplomová práce. Vysoká škola ekonomická.

KAISAROVÁ, Petra. *Základy popisné statistiky v pedagogickém výzkumu*. Brno, 2009, s. 36. Bakalářská práce. Masarykova univerzita.

KRČMAŘÍKOVÁ, Monika. *Možnosti využití metody sémantického diferenciálu k získávání informací důležitých pro manažerské rozhodování*. Nymburk, 2007, s. 13. Bakalářská práce. Vysoká škola ekonomická v Praze.

KREISLOVÁ, Gabriela. *Dotazníkové šetření*. Plzeň, 2008, s. 14. Bakalářská práce. Západočeská univerzita v Plzni.

KREPL, Jan. *Výběr z konečných populací v ekonomických úlohách*. Praha, 2014, s. 8. Bakalářská práce. Univerzita Karlova v Praze.

LÖSTER, Tomáš, Hana ŘEZANKOVÁ a Jitka LANGHAMROVÁ. *Statistické metody a demografie*. 1. vyd. Praha: VŠEM, 2008, s. 10 - 12. ISBN 978-80-86730-40-0.

PECÁKOVÁ, Iva. *Statistické aspekty terénních průzkumů I*. 1.vyd. Praha: VŠE, 1995, s. 11. ISBN 80-7079-737-1.

PECÁKOVÁ, Iva. *Statistika v terénních průzkumech*. 1. vyd. Praha: Professional Publishing, 2008, s. 24-25. ISBN 978-80-86946-74-0.

PECÁKOVÁ, Iva, Ilja NOVÁK a Jan HERZMANN. *Pořizování a vyhodnocování dat ve výzkumech veřejného mínění*. 3. vyd. Praha: Oeconomica, 2004, s. 24-29. ISBN 80-245-0753-6.

PITOŇÁK, Michal. *Queer geografie Česka: Heteronormativita prostoru*. Praha, 2011s. 44-47. Diplomová práce. Univerzita Karlova v Praze.



RAMBALOVÁ, Hana. *Marketingový výzkum*. Brno, 2011, s. 23-24. Bakalářská práce. Masarykova univerzita

ŘEZANKOVÁ, Hana. *Analýza dat z dotazníkových šetření: (druhé vydání)*. 2. vyd. Praha: Professional Publishing, 2010, s. 13 - 14. ISBN 978-80-7431-019-5.

SCHNEIDEROVÁ, Bc. Romana. *Vybrané metody výběru*. Olomouc, 2011, s. 6-7. Diplomová práce. Univerzita Palackého v Olomouci.

SRNA, Martin. *Marketingový výzkum spokojenosti zákazníků*. Brno, 2008, s. 31. Bakalářská práce. Vysoké učení technické v Brně.

STIEBNER, Ondřej. *Prozkoumávání jednorozměrných dat*. Olomouc, 2008, s. 22-23. Diplomová práce. Univerzita Palackého v Olomouci.

ŠKRAŇKA, Ivan. *Statistické metody v marketingovém výzkumu*. Brno, 2009, s. Diplomová práce. Vysoké učení technické v Brně.

TAHOTNÁ, Tatiana. *Vliv způsobu měření na volbu metody statistické analýzy*. Praha, 2008, s. 15. Bakalářská práce. Vysoká škola ekonomická v Praze.

TRAMPOTA, Tomáš a Martina VOJTĚCHOVSKÁ. *Metody výzkumu médií*. 1. vyd. Praha: Portál, 2010, s. 189-190. ISBN 978-80-7367-683-4.

VÝTISK, Radek. *Pokročilé statistické metody v marketingovém výzkumu*. Brno, 2008, s. 17-18. Diplomová práce. Masarykova univerzita.

ZBOŘIL, Kamil. *Marketingový výzkum: Metodologie a aplikace*. 1.vyd. Praha: Vysoká škola ekonomická, 1998, s. 51. ISBN 80-7079-394-5.

# SEZNAM OBRÁZKŮ, GRAFŮ A TABULEK

## Seznam obrázků

|                          |    |
|--------------------------|----|
| Obrázek 1: Snowball..... | 24 |
|--------------------------|----|

## Seznam grafů

|                                                   |    |
|---------------------------------------------------|----|
| Graf 1: Krabičkový graf zobrazující kvantily..... | 51 |
|---------------------------------------------------|----|

## Seznam tabulek

|                                                     |    |
|-----------------------------------------------------|----|
| Tabulka 1: Baterie otázek .....                     | 30 |
| Tabulka 2: Likertova stupnice .....                 | 39 |
| Tabulka 3: Sémantický diferenciál .....             | 40 |
| Tabulka 4: Metoda zdánlivě stejných intervalů ..... | 42 |
| Tabulka 5: Metoda párových srovnání .....           | 44 |

## SEZNAM PŘÍLOH

|                                                                     |             |
|---------------------------------------------------------------------|-------------|
| <b>Příloha A – Rozhovor s Mgr. Pavlem Šimoníkem .....</b>           | <b>I</b>    |
| <b>Příloha B – Rozhovor s Ing. Kamilou Burešovou .....</b>          | <b>VIII</b> |
| <b>Příloha C – Rozhovor s Mgr. Ing. Martinem Zeleným, Ph.D.....</b> | <b>IX</b>   |
| <b>Příloha D – Rozhovor s Mgr. Jindřichem Krejčím, Ph.D.....</b>    | <b>XII</b>  |
| <b>Příloha E – Rozhovor s PhDr. Stanislavem Hamplem .....</b>       | <b>XIV</b>  |

## PŘÍLOHY

### Příloha A – Rozhovor s Mgr. Pavlem Šimoníkem

#### **Jaká metoda výběru respondentů je při výběrových šetřeních v praxi nejpoužívanější?**

Určitě v rámci výzkumů marketingových i sociologických je to v současné době kvótní výběr. Ten důvod je jednoznačný, protože již od druhé poloviny devadesátých let neexistuje přístup do databáze domácností s adresami. Takže pokud se dělá nějaký náhodný výběr, tak se dělá tím způsobem, že se adresy musí sepsat na základě několika kroků, kdy se používá například databáze České pošty, nebo se používá databáze Českého statistického úřadu. Tam se vyberou domy a vchody a v těchto vchodech se zjistí počet bytů. A pak se nějakým náhodným krokem musí udělat soupis adres. Je to finančně velmi nákladné a v současných podmínkách má tento výběr velmi malou návratnost. Používá se ale například u Media projektu, kde se měří čtenost tisku, u kontinuálního výzkumu pro měření sledovanosti televize atd. A někdy se to využívá při výzkumech pro akademický ústav, kde se provádí mezinárodní výzkumy. Tam si oni přímo zadají, že si přejí náhodný výběr a v těchto případech my jej provádíme. Ta návratnost se pohybuje pod 50 %, jelikož nedostupnost domácností je obrovský problém. A v některých metodikách převzatých ze zahraničí je požadavek na nemožnost substituce domácností, které se nepodaří zastihnout. Tudíž se pak velmi těžko dosahuje nějaké návratnosti, musí se dělat další kola výběru. U kontinuálních výzkumů, které se používají pro sledovanost médií, se může pracovat se systémem náhradních domácností. To je ale složitější proces. Navíc tento výběr stojí dvakrát více než kvótní výběr.

Z tohoto důvodu se používá kvótní výběr a to standardně při použití 5 kvótních parametrů. Pohlaví, věk, vzdělání, velikost místa bydliště, kraj. Někdy se kombinují ještě další kvóty, ale kombinace kvót má určitý limit, aby tazatel vůbec dokázal toho respondenta najít. Někdy se tedy může ještě využít kvóta na počet členů domácnosti, aby se eliminovaly jednočlenné domácnosti. Zhruba 80 % výzkumů probíhá na základě kvótního výběru. Ať už se jedná o telefonický výzkum, výzkum face to face, CAPI, na online panelech.

Anketa se nepoužívá vůbec, pouze v nějakých krajních případech, jelikož její vypovídající hodnota je velmi nízká. Je to ovlivněno efektem toho, kdo chce odpovídat. Jedná se o velmi aktivní typ lidí, kteří něco velmi chválí a jsou velmi spokojení. Chybí tam ale mainstream těch

běžných lidí, kteří na toto nereagují. Ankety se používají spíše při nějakých specifických akcích, kde lidé mohou něco vyplnit. Média si mohou udělat svoje ankety. V klasickém výzkum ale anketa nemá moc co dělat.

Snowball se používá, protože když vám někdo zadá výzkum, v rámci kterého máte najít osoby, které pečují o osobu s příspěvkem na péči a musíte jich najít 800, tak nemáte jinou možnost, než hledat pomocí snowballu. Databáze těchto dat neexistuje, musíte jít přes lidi a různé instituce a teprve až natrefíte na někoho, kdo splňuje tyto podmínky, tak vám dá zase kontakt na někoho dalšího.

Nebo pokud hledáte cizince z určité komunity, tak použijete snowball, abyste navázala kontakt s někým podobným.

To bylo k vaší otázce. Stačí vám odpověď takhle?

**Určitě, moc děkuji.**

Jenom dodám, že standardně na českou populaci je velikost vzorku 1000. U větších akademických šetření se sbírají třeba 3000, jelikož tam dochází ke třídění a hledání podskupin a používání určitých multivariačních technik.

**Pokud používáte škály při tvorbě dotazníků, je lepší použít škály se sudými, či lichými počty členů? Například kvůli riziku, že při lichém počtu členů se mnoho respondentů zařadí do prostředního členu, a naopak při sudém počtu členů budou mít pocit, že si musí vybrat nějakou stranu.**

Škály jsou velmi diskutované téma od doby, co vznikly. Neustále se vedou debaty o tom, jaké škály použít. Zda jsou lepší sudé či liché, krátké či dlouhé škály. Zda čtyřstupňové, pětistupňové nebo desetistupňové. Pokud se něco přejímá z metodik, které se používají v cizině, vždy se používají delší škály, a to pro hodnocení čehokoliv. Ať už pro hodnocení spokojenosti, důležitosti atd. se používá škála 1 až 10. Zejména velké nadnárodní firmy si v rámci měření zákaznické spokojenosti přejí široké škály.

Pro běžně sociologické šetření se spíše používají škály 1 - 4. Jedná se o typickou ordinální škálu, kde je "velmi spokojen" spíše spokojen" "spíše nespokojen" "velmi nespokojen".

**Spíše se tedy používají škály se sudými počty členů?**

Ano, spíše s těmi sudými. Liché škály se také objevují v nějakých případech, ale obecně u nich platí, že tendence lidí je vpasovat se doprostřed škály. A to je pro výzkumy spíše limitem a nikoli výhodou. Potřebujeme z člověka dostat nějaký názor a postoj. Pokud mu pomocí liché škály umožníme únik, tak jej často využije. Spoustu lidí samozřejmě nemá na danou věc názor a pak je lepší nechat je odpovědět "nevím" " neumím posoudit". Škály jsou tedy obecně problémem. Neplatí však pravidlo, že by se liché škály nesměly používat.

Na škálách se také nepoužívají čísla, ale spíše výroky. Velmi často se také používá sémantický diferenciál. Je to velmi oblíbená metoda. V nových metodikách se navíc používají tzv. "šoupátka" pomocí kterých se respondenti pohybují na určité ose a nemají tam ani definované hodnoty. Pouze pocitově se zařadí na určitém místě. Tím se snažíme dosáhnout přirozenosti u respondenta.

**A vy tedy ve většině případů využíváte spíše sudé škály, a to ty 4 bodové?**

Ano, já jsem zastáncem toho, že člověk se pohybuje spíše v užším rámci. Já jsem schopen říct, s tímto souhlasím a s tímto nesouhlasím. A možná spíš s tímto souhlasím a s tímto spíše nesouhlasím.

Rozhozením do delší škály nastává problém. Velmi to těžko pak vysvětlujete klientům z komerční sféry, kteří velmi lpí na tom, zda jim někdo dal 8 či 9 na 10 bodové škále ve spokojenosti. Ale pro někoho je 8 pocitově velmi dobré číslo a 10 by nedal nikomu, protože nic není bez vady. Někdo naopak dá 10 s tím, že je vlastně docela spokojen. Další měsíc se pak hodnocení posune o 1 stupeň a klient řeší, jak to, že se jim snížil průměr a co to znamená. Oni si poté tu škálu rozsekají a pracují s určitými koeficienty podle speciálních metodik. I tato škála má tedy něco do sebe, ale když by si měl ale respondent vybrat, zda je spíše spokojen nebo velmi spokojen, tak se s tím pracuje daleko lépe.

My ale používáme kratší škály a sudé škály.

**A používáte v těchto šetřeních také škály s číselnými hodnotami? Jakého druhu? Například zda se používá zlomková škála, nebo je příliš nepraktická a využívá se spíše jiná číselná škála?**

Zlomkové škály se téměř nepoužívají. Často se používá hodnocení podle školních známek. To je lidem velmi blízké a umí si to zařadit. Zámka 3 není střed škály, protože 3 už není dobrá známka. Zde to tedy nějak funguje a používá se i škála formou "napíšte známku" a lidé tam mohou napsat i 1-, 2- atd. Jinak se standardně používají škály od 1 do 4 nebo do 5. A používají se i procentuální škály.

### **A bodovací škály, kde má respondent za úkol rozdělit určitý počet bodů?**

Ano, ty bodovací škály, kde respondent rozděluje balík procent, se také používá. Rozděluje procenta dle důvěryhodnosti, například u politiků. Možností je mnoho, ale používají se spíše osvědčené věci, než nějaké experimenty.

### **Známkování je tedy nejosvědčenější metodou?**

Ano, na hodnocení je nejlepší možnou variantou. Ve výzkumu veřejného mínění se často používá.

Kromě toho se používají i úsměvné či mračící obrázky, což se ale následně stejně musí převrátit na číselné hodnoty, aby se s tímto dalo dále pracovat.

### **A jaké jsou nejběžnější potíže při sběru dat ve výběrových šetřeních? Je to například problematika non-response? Nebo jsou i jiné, mnohem závažnější potíže?**

Největším problémem při sběru dat je dnes jednoznačně dostupnost jednotlivých domácností. To je mnohem větší problém, než jakékoli metodické problémy. Nalezení a korektní dotázání respondenta je největším problémem. Dnes bojujeme s tím, že lidé zavírají dveře tazatelům před nosem, nechť je pouštět do domů. Poměrně efektivní jsou telefonní výzkumy, kde respondent je ochotný na kratší dobu absolvovat rozhovor. Ale u face to face je největším problémem nevyšetřená populace. Máme sice kvótu nebo uděláme náhodný výběr, ale unikne nám 90% populace, která není ochotná se s námi bavit. Poté se to cykluje atd.

Pak jsou i další problémy, u face to face zejména to, jak si to sami tazatelé přizpůsobují. Dnes se již i tyto rozhovory nahrávají. CAPI dotazování umožňuje nahrávání. STEM/MARK byl první kdo tyto nahrávky zavedl a zjistilo se, kolik procent podvodných rozhovorů bylo a kolik

se jich muselo vyřazovat. Často ti velmi úspěšní tazatelé, kteří dělali velké počty rozhovorů, neprošli tímto novým opatřením, protože si tyto rozhovory dělali sami.

### **Takže při nahrávání rozhovorů se zjistilo hodně těchto podvodů?**

Ano, zjistilo se, že si to vyplňují sami nebo mluví s někým starším či mladším, než bylo uvedeno. Toto opatření je tedy jakousi revolucí v kontrolách tazatelů při výzkumech Face to face. Začínají se používat i GPS souřadnice, aby se zjistilo, kde se tazatel pohybuje. Zda není pouze na jednom místě. Těchto metod je mnoho.

Nejspolehlivějším výzkumem jsou ale telefonické rozhovory, protože jsou pod plnou kontrolou. Tazatel nemá šanci do toho vnést něco svého. Musí se držet skriptu a je slyšet vše co říká. Nemá šanci ani vybrat respondenta, protože ten se vybírá náhodně z databáze. I zde je však riziko, že respondent se odmítne s tazatelem bavit.

Proti vznikají panely, kde jsou respondenti nějakým způsobem odměňováni za to, že se toho průzkumu zúčastní. Jsou to všechny online výzkumy, ale i panely domácností, které se dotazují face to face. My je odměňujeme za to, že s námi chtějí spolupracovat.

S tou malou návratností pak souvisí i problém nenaplnění kvót a obtížného shánění některých skupin. Například osoby se základním vzděláním se shánějí těžce. Jejich podíl klesá, a i když jsou parametry nastaveny tak, aby jejich podíl odpovídal podílu v populaci, tak se velmi těžko shánějí. Pokud hledáte tohoto člověka, jsou to často buď důchodci a spíše ženy, které nešly dále studovat a byly doma. Nebo mladí lidé, kteří ještě nemají středoškolské vzdělání.

Dalším problémem je i to, že některé osoby nejsou ochotny na určité otázky odpovídat. Tím pak vznikají chybějící odpovědi v dotaznících. Týká se to například příjmů. Je to velmi citlivé téma. I když se dnes situace mění k lepšímu oproti stavu v 90. letech, kdy panovala obrovská obava z toho, kde to bude použito. Ta obava ale panuje i dnes a musíme to řešit například kategorizací, kdy dáváme kategorie například do 10 000, nad 13 000 Kč atd., Lidé jsou pak schopnější se zařadit. Pokud tyto data chybí, dá se to následně modelovat podle ostatních odpovědí, například podle ekonomicko-sociálního statusu, který ten člověk má, podle profese atd., ale je to zásah do dat a není to už čistý dotazník.

Dotazníky by měly být krátké a svižné, ale u akademických dotazníků se jedná o hodinové dotazování. Což je problém, protože se respondent soustředí prvních 10 až 20 minut. Tam je



nutné pak ty respondenty motivovat a spíše na ně platí nějaká odměna než vidina toho, že se něco někde zlepší ve společnosti.

Takže jsou různé otázky, které považují respondenti za choulostivé, kromě příjmu?

Dnes se dá dotazovat na všechno. Dělali jsme průzkum i ohledně sexuálního chování a není to ve společnosti takové tabu. Je to výzkum, který se nemůže dotazovat přes tazatele samozřejmě. Mnohem efektivnější je dotazování na panelu, nebo telefonicky, ale spíše v omezené míře. Dělali jsme výzkumy na viagru a na fyzické dysfunkce. Tam jsme dávaly dotazníky zalepené v obálce a respondent odpovědi zase vložil do obálky, aby je nikdo jiný nečetl.

Choulostivé jsou spíše otázky týkající se rodinných záležitostí a rodinného stavu. Ale spíše otázky ohledně korupce a zdraví, farmacie, nemoci atd. K těmto otázkám se pak musí přistupovat velmi opatrně. Nejvíce choulostivá je však otázka na příjem.

### **Nejproblematictější je tedy vůbec dostat se do domácností?**

Ano, dostat se do domů a bytů je nejtěžší. Tazatel zvoní a snaží se tam dostat a je považován za podomního prodejce. Spíše se lidé pozvou někam telefonicky a zaplatí se jim za to, aby tu byla ta návratnost.

Problém je i v tom, že data se různě vrací s různou návratností z regionů. Pak se používají koeficienty navýšení a dopředu se již ví, že v Praze je nejnižší návratnost. Tam se musí udělat více kol. Praha je v tomto abnormálním. Ochota lidí odpovídat je zde nejmenší. Nejvíce ochotní lidé jsou naopak na vesnicích, na venkově, Moravě, jižních Čechách.

Když nabídnete odměnu, je nejvyšší ochota v severních Čechách a v chudších regionech. Je tedy nutné pracovat s více faktory.

Snažíme se posílat i avizní lístky s tím, že přijde tazatel. Někomu to pomůže a zavolá nám, aby se dotázal, zda nejsme podvodníci. To hlavně v případě, že se jedná o dotazování na vybavenost domácnosti. Na to je mezinárodní metodika, která řadí lidi podle toho, co vlastní. Je ale zastaralá. To, zda mají respondenti sekačku či barevnou televizi je dnes zcela nepodstatné. Dnes se spíše zjišťuje tablet, smartphone, počet aut atd. Někoho naopak avizní lístky odradí a nechce s tazatelem mluvit vůbec.

### **A když je tedy návratnost malá, jak se řeší doplnění dat?**

Používají se buď koeficienty navýšení, nebo došetření. Pak se musí použít systém náhradníků a udělat další kolo výběru. Znovu poslat tazatele na místo, aby se pokusili zastihnout domácnosti. Je to ale velmi drahé a je snaha se tomu vyhnout.

Více se tedy používají koeficienty navýšení, je to snazší a méně finančně nákladné. Někdy ani nelze dosbírat data v jiném termínu, pokud se jedná například o kontinuální výzkum, který je zaměřen na sledování čtenosti novin a časopisů. Tam se sbírají data v měsíčních či čtvrtletních cyklech a není možnost je dosbírat jindy. Tam se to pak musí hlídat už v průběhu sběru. Na to už dnes existují přístroje a programy, které monitorují stav toho sběru. Počty navštívených, odmítnutých atd.

Dnes se při dotazování používají i smartphony, čemuž říkáme SAPA, a u těch se nestává, že by výsledkem byl neúplný dotazník. Ten program to nedovolí.

Dnešní trend spěje k tomu, aby bylo vše pod kontrolou a aby to bylo co nejjednodušší. Aby byla větší rychlost dotazování. Dříve trval výzkum i měsíc, dnes někdo zavolá s tím, že potřebuje za 3 dny výsledky a my musíme ta data sebrat, zpracovat a odeslat do těch 3 dnů. Takže požadavky na zjištění a zpracování dat se neustále zkracují.

### **Takže dnes není problém sehnat za 3 dny respondenty a vyhodnotit dotazník?**

Neuděláte to face to face, ale je možné to zorganizovat na malém počtu respondentů v rámci nějakého bleskového šetření. Skvěle to funguje na panelu, tam je 60 000 panelistů, zadáte si cílovou skupinu a oni jsou do 24 hodin schopni reagovat. Poté se to vyhodnotí. Klient může i průběžně vidět, jak respondenti odpovídají, což je i pro něj určitá forma kontroly.

Data jdou v podstatě sebrat i za den, speciální software pak ty data rovnou formuje do grafů.

Například výzkum pro Českou televizi, kdy televize zavolá odpoledne s tím, že potřebuje zjistit reakci lidí například na útoky ve Francii. Dotazník se připraví během 2 hodin a pak se druhý obvolá 500 lidí. Oni potřebují výsledky do večerních zpráv, takže se to musí stihnout. Ale u velkých akademických dotazování se to takhle praktikovat nedá.

## **Příloha B – Rozhovor s Ing. Kamilou Burešovou**

### **Jaké typy sběru dat používáte ve vaší organizaci?**

Ve STEMu se nejčastěji při kvantitativních průzkumech pracuje s metodou výběru prostřednictvím kvótního výběru, kdy je výsledný soubor respondentů reprezentativní za celou populaci ČR starší 18 let podle kvótních znaků pohlaví, věk, vzdělání, velikostní pásmo obce, region. Systematický náhodný výběr se používá (vzhledem ke své náročnosti a nákladnosti) méně často. Samozřejmě pro specifické účely mohou být využity i jiné metody (např. výběr cílený na určité skupiny obyvatel nebo určité regiony, anketa apod.).

### **Používáte ve vaší organizaci sudé či liché škály?**

V dotazníku se používají různé škály v závislosti na tématu otázky. Pro zjištění míry souhlasu nebo třeba spokojenosti se ve STEMu nejčastěji používají škály se sudým počtem variant bez přímého nabízení možnosti „nevím“ (samozřejmě ji respondent uvést může a tazatel takovou odpověď zapíše, někdy je tato odpověď kvůli povaze otázky přímo nabízena).

### **Používáte také nějaké druhy číselných škál?**

Škály s číselnými hodnotami se rovněž používají, většinou pro hodnocení. Nejčastěji pak prostřednictvím školních známek. Používá se také výběr mezi dvěma protikladnými charakteristikami, výroky (např. sémantický diferenciál, příklad: dvojice „moderní – zastaralý“, škála od 1 do 7, pokud se respondent jednoznačně přiklání k variantě moderní, vybírá 1, pokud mu zcela vyhovuje varianta zastaralý, volí 7, neutrální názor je 4).

### **S jakými problémy se při sběru dat setkáváte?**

Potíže při sběru dat souvisejí s typem výběru. Při náhodném výběru je samozřejmě základním problémem zajištění dostatečné návratnosti, která souvisí s odmítáním rozhovoru, obtížností zastihnout konkrétního respondenta. V kvótním výběru není problém odmítnutí tak zásadní, i když podle zkušeností tazatelů z terénu se ochota k účasti na výzkumu mezi lidmi snižuje.

### **Jak probíhá doplnění chybějících dat?**

V nejčastěji prováděném kvótním výběru problém s doplněním chybějících dat není. Při náhodném výběru záleží na konkrétním zadání tohoto výběru (někdy substituce výpadků návratnosti v zadání není povolena).

## **Příloha C – Rozhovor s Mgr. Ing. Martinem Zeleným, Ph.D.**

### **Jaká metoda výběru respondentů je při výběrových šetřeních v praxi nejpoužívanější?**

V "oficiální statistice", tj. v doméně statistických úřadů a institucí, jednoznačně dominují pravděpodobnostní formy výběru. Spíše zřídka a u specifických šetření se lze setkat s kvótní formou výběru. U šetření v domácnostech se sběrem dat tazateli v terénu se většinou provádí výběr dvou- či vícestupňově, tj. na prvním stupni se vybere určitý počet tzv. primárních výběrových jednotek (primary sampling units - PSU) a v nich se poté vybírají byty nebo jednotlivci pro šetření. Cílem je větší koncentrace výběru v terénu a časově a nákladově efektivnější práce tazatelů - i když tato forma vícestupňového výběru má zpravidla o něco horší statistické vlastnosti (pro stejný počet vyšetřených respondentů produkuje odhady s větší variabilitou, vyjádřenou např. intervaly spolehlivosti). Častá je i tzv. stratifikace - tj. výběr se předem ve fixních proporcích rozdělí do podskupin základního souboru, tzv. strat - např. podle krajů, velikosti obcí, atp. Snowball typy náboru respondentů nebo ankety se v oficiální statistice nepoužívají, nevyhověly by základním kvalitativním požadavkům na výběrová šetření. Maximálně jako podpora expertních odhadů. V podnikové statistice jsou ale leckdy velké jednotky zařazovány do šetření automaticky, a výběr se dělá jen u menších subjektů (s méně zaměstnanci, s menším obratem). V některých specifických odvětvích s vlastním zjišťováním se může zjišťovat i vyčerpávajícím způsobem - všechny subjekty, pokud jich je málo - např. několik desítek.

**Při používání škálování u těchto šetření, například v dotazníku, je lepší použít škály se sudými, či lichými počty členů? Například kvůli riziku, že při lichém počtu členů se mnoho respondentů zařadí do prostředního členu, a naopak při sudém počtu členů budou mít pocit, že si musí vybrat nějakou stranu.**

Co vím, výzkum u subjektivně laděných škál podporuje hypotézu, že škály s lichým počtem možností vedou k častějším volbám prostřední "neutrální" kategorie a škály se sudým počtem položek mohou mít výhodu. Záleží ale jistě i na počtu odpovědí, charakteru otázky a významu jednotlivých možností, včetně jejich slovního popisu.

**Používají se v těchto šetřeních také škály s číselnými hodnotami? Jakého druhu?**

**Například zda se používá zlomková škála, nebo je příliš nepraktická a využívá se spíše jiná číselná škála?**

Pokud pominu prosté označení možností čísly, typ škál, které máte na mysli, nepoužíváme. Jednak je zaměření našich šetření spíše na "tvrdší" kvantitativní data - s pouze omezeným počtem subjektivnějších otázek vyžadujících ordinální škálu - typu jak celkově vycházíte s příjmem, jakou zátěž pro vás představují náklady na bydlení. Málokdy jde o otázky hodnotící více variant a vyjadřující postoje respondenta k více variantám. To je častější spíše v sociologickém nebo psychologickém výzkumu nebo v marketingu.

**Jaké jsou nejběžnější potíže při sběru dat ve výběrových šetřeních? Je to například problematika non-response? Nebo jsou i jiné, mnohem závažnější potíže?**

Non-response je určitě nejvýznamnějším faktorem - kromě odmítnutí rovněž v té části nekontaktní - tj. situace, kdy se respondenta / nikoho v šetřeném bytě nepodaří vůbec kontaktovat. Kromě toho je významná věc i tzv. measurement error - tedy situace nepřesných nebo nesprávných odpovědí respondenta - ať již záměrně (zatajování / podhodnocení spotřeby např. u alkoholu, cigaret, neúplné údaje o příjmech atd.) nebo nezáměrně - respondent neví, nevzpomene si na všechny případy nebo naopak do sledovaného období zařadí i případy vyskytnuvší se předtím, nebo vůbec nepochopí, na co se ho otázka ptá.

**Jakým způsobem nejčastěji probíhá doplnění chybějících dat? Je to třeba pomocí systému náhradníků, nebo nějakého statistického vzorce?**

"Náhradníci", resp. tzv. substituce souvisí spíše s non-response než s imputací dat. V některých šetřeních se substituce - např. povolený kontakt jiného bytu v domě pokud první domácnost odmítne - používá, ale měly by pro ni platit striktní pravidla a jde o věc ošemetnou. Ve standardu zpráv o kvalitě pro jednotlivá šetření se zpravidla vyžaduje protokolovat, jaké jednotky byly jakými nahrazeny - tj. kontrolovat celý proces tak, aby nedocházelo k nekontrolované systematické náhradě jednoho typu respondentů jinými, ochotnějšími (např. typy lidí, kteří jsou častěji doma). Jinak hrozí systematické vychýlení výsledků. Konkrétně u nás substituce není povolena a pracujeme s pevným výběrem.

U imputací hodnot v neúplných, nekompletních dotaznících je metod celá řada, od primitivních v podobě nahrazení chybějící hodnoty průměrem ostatních pozorování až po sofistikované vícefázové systémy regresních modelů. Míra této tzv. položkové non-response je u nás zpravidla nižší - respondenti spíše buď odmítnou celé šetření, pokud ne, už zpravidla rozhovor dokončují.

Pro dořešení situace s relativně nízkým procentem chybějících pozorování používáme nejčastěji tzv. hot-deck metody s náhodným výběrem "dárce" údaje podobného imputovanému respondentovi. Další specifickou oblastí jsou imputační modely na rozklíčování položek udaných respondentem pouze ve formě úhrnu - např. situace, kdy respondent uvede celkové náklady na bydlení, ale ne konkrétní položky na teplou vodu, topení atp. Nebo náklady na zájezd jako celek, bez položek na dopravu a ubytování. Tady jde o modely "na míru" konkrétní situaci, s využitím portfolia pomocných proměnných z dotazníku - např. kam zájezd byl, na kolik dní, jakým dopravním prostředkem, atp.

## **Příloha D – Rozhovor s Mgr. Jindřichem Krejčím, Ph.D.**

**Jaká metoda výběru respondentů je při výběrových šetřeních v praxi nejpoužívanější? (Například je to metoda prostého náhodného výběru či nějakou jeho modifikaci? Nebo například snowball či anketa?)**

Metoda výběru vždy odpovídá výzkumné situaci. Většinou se totiž liší cílové skupiny i jednotky výzkumu. Jiný okruh metod máte k dispozici pro šetření v organizaci a jiný pro nějaké specifické skupiny ve společnosti. A jiný zase pro šetření obecné populace. Velmi také záleží na opoře výběru, která je k dispozici.

Pokud se bavíme o reprezentativních šetřeních populace v ČR, které se provádějí na základě osobních rozhovorů face-to-face, tak dominuje kvótní výběr. V akademickém sektoru nicméně dominují různé postupy pravděpodobnostních výběrů, zpravidla několikastupňový stratifikovaný výběr. Prostý náhodný výběr v běžné praxi šetření populace realizovat nelze, jelikož není dostupný registr osob.

Snowball je vhodný pro specifické výzkumné situace, kdy není k dispozici žádná opora výběru, ale nikoli pro obecné šetření populace.

Ankety jsou zcela nerepresentativní, tady se nejedná o výběrové šetření. Zaměňovat anketu za výběrová šetření je v rozporu s kodifikovanými pravidly výzkumných asociací.

**Při používání škálování u těchto šetření, například v dotazníku, je lepší použít škály se sudými, či lichými počty členů? Například kvůli riziku, že při lichém počtu členů se mnoho respondentů zařadí do prostředního členu, a naopak při sudém počtu členů budou mít pocit, že si musí vybrat nějakou stranu.**

Lichý počet odpovědí zajišťuje vyšší reliabilitu, sudý počet odpovědí umožňuje dichotomizaci odpovědí, navíc střední hodnota může občas „přitáhnout“ příliš velký podíl respondentů. U každé otázky zvlášť se tudíž zvažuje co je pro ni konkrétně důležitější, resp. který problém je v tom kterém případě závažnější. Univerzální doporučení není.

**Používají se v těchto šetřeních také škály s číselnými hodnotami? Jakého druhu? Například zda se používá zlomková škála, nebo je příliš nepraktická a využívá se spíše jiná číselná škála?**

Průzkumy veřejného mínění jsou realizovány většinou pro jednoduchou prezentaci v médiích, takže používají spíše škály s popsánymi všemi body, aby každý výsek koláčového grafu něco věcně znamenal. Pokud už se používají pouze číselné, resp. škály s popsánymi pouze krajními body, pak téměř vždy s celými čísly, někdy od 0, někdy od 1, zcela výjimečně např. -5 až +5. Výjimečně také procenta („kolik procent podle Vás...“)

Vědecké výzkumy používají častěji číselné škály s popsánymi pouze krajními body (a to obvykle 10, nebo 11 bodové), protože jde především o analýzu souvislostí mezi proměnnými.

**Jaké jsou nejběžnější potíže při sběru dat ve výběrových šetřeních? Je to například problematika non-response? Nebo jsou i jiné, mnohem závažnější potíže?**

Non-response je nejzávažnější problém. Jinak je problém samozřejmě hlavně formulace otázky a odpovědí. Při výběru postupu (firmy pro) sběr dat se hledí především na postup výběru (jaká je opora výběru (coverage), kolik výběrových kroků a jaké, kolik je celkem sampling points, kolik dotazníků dělá jeden tazatel, jaká je kontrola tazatelů atd.)

**Jakým způsobem nejčastěji probíhá doplnění chybějících dat? Je to třeba pomocí systému náhradníků, nebo nějakého statistického vzorce?**

Pokud se jedná o pravděpodobnostní výběr, tak na začátku se provede výběr a ne každý, kdo je v tomto výběru, je dosažitelný nebo odpoví. Dotazník může vypadnout i z jiného důvodu. Výběr se musí plánovat dostatečně velký, aby se splnilo zadání. Když se naplánuje špatně, musí se udělat dovýběr a pokračovat ve sběru dat.

U kvótního šetření dělá výběr v podstatě tazatel, který má předepsání, kolik osob musí oslovit dle zadaných charakteristik. Když ho někdo odmítne, tak v rámci předepsané výběrové procedury osloví jiného člověka se stejnou charakteristikou. Data pak chybí jen v případě, že tazatel úkol nesplní. To může být dáno i tím, že úkol je nastaven nerealisticky.

Doplnění dat probíhá nadvýběrem, tedy dodatečným došetřováním populace.



## **Příloha E - rozhovor s PhDr. Stanislavem Hamplem**

### **Jaké typy sběru dat používáte ve vaší organizaci?**

V praxi nejpoužívanější je jednak kvótní výběr (používá se např. u pravidelného omnibusového šetření, dále u řady regionálních či lokálních šetření), jednak stratifikovaný vícestupňový pravděpodobnostní (náhodný) výběr (zejména pro výzkumy prováděné pro akademickou sféru nebo obecně pro klienty, kteří mají vyšší nároky na metodologickou stránku šetření). Prostý náhodný výběr lze použít jen u některých speciálních výzkumů, kdy máte jednu databázi, z níž můžete vybrat. Snowball je metoda výběru, která není příliš často používána, týká se jen případů, kdy není jiná možnost dostat se na další respondenty a kontakty na ně, než se doptávat těch stávajících. Anketa se v rámci praxe výzkumných agentur v podstatě nepoužívá, protože nezaručuje reprezentativní vzorek (existuje zde tzv. efekt samovýběru).

### **Používáte ve vaší organizaci sudé či liché škály?**

Používají se škály se sudým i lichým počtem, záleží na účelu (jestli jde o to, aby se názory rozdělily na pozitivní a negativní, nebo nás zajímají právě i ty neutrální). Vždy je třeba si být vědom rizika středové kategorie (že může sloužit jako úniková). Používají se tedy oba druhy, zhruba půl napůl.

### **Používáte také nějaké druhy číselných škál?**

Číselné škály u postojových otázek se používají a většinou právě s lichým počtem. Ani sudý počet ale u číselné škály neeliminuje tendenci zařadit se do středu (např. u škály 1-10 bývá respondenty jako střed vybírána pětka, i když středem není). Zlomkovou škálu jsem snad ještě neviděl. Používají se nejvíce celočíselné škály.

### **S jakými problémy se při sběru dat setkáváte?**

Problematika non-response je určitě velmi závažná potíž, asi by se dala označit za nejzávažnější. Aktuálním problémem, který se teď často řeší, je také kombinace různých módů dotazování (face-to-face s papírovým dotazníkem (PAPI), s počítačem (CAPI), telefonicky (CATI) nebo online (CAWI)) v rámci jednoho výzkumu.

### **Jak probíhá doplnění chybějících dat?**

Chybějící data se doplňují „dosběrem“, tedy dodatečným sběrem dat.

## **BIBLIOGRAFICKÉ ÚDAJE**

**Jméno autora: Bc. Veronika Doskočilová**

**Obor: Sociální a mediální komunikace**

**Forma studia: Prezenční**

**Název práce: Výběrová šetření - teorie a praxe**

**Rok: 2015**

**Počet stran textu bez příloh: 69**

**Celkový počet stran příloh: 14**

**Počet titulů českých použitých zdrojů: 28**

**Počet titulů zahraničních použitých zdrojů: 0**

**Počet internetových zdrojů: 0**

**Počet ostatních zdrojů: 0**

**Vedoucí práce: RNDr. Jaroslav Huk, CSc.**