7. VÝSLEDKY VÝZKUMU

V předešlých kapitolách jsme kromě poznatků o výzkumu obecně a výzkumu sociálním (viz 1 a 2) absolvovali první etapu zkoumání, totiž jeho přípravu (viz 3 a 4), a etapu druhou, jeho realizaci. Tu zastupovalo seznámení se základními a též s několika speciálními způsoby shromažďování údajů (viz 5 a 6). Zbývá nám tedy etapa třetí a poslední, vyhodnocování údajů.

V textu jsem opakovaně zmiňoval, že se zaměřuji pouze na základní informace. Nyní to platí dvojnásob. Povaha následující problematiky je totiž taková, že její trochu bližší popis by si vyžádal objasnění dalších souvislostí, ty opět další podrobnosti atd., až by došlo na rozsáhlé vysvětlování. Toho se ale chci vyvarovat, neboť hlavním účelem výkladu je přispět především k pochopení podstaty a základního smyslu operací, kterými data procházejí.

7.1 VYHODNOCOVÁNÍ KVANTITATIVNÍCH DAT

Otázky vyhodnocení dat získaných kvantitativními postupy patří v publikacích, které se tímto způsobem zkoumání zabývají, ke standardním kapitolám. Pojednávají je v různé míře podrobnosti a většinou v nich nechybí ani informace o **využití statistiky a případně i počítačových programů**. Netřeba se však jakkoli obávat – těmto aspektům se budu věnovat skutečně minimálně a případné zájemce tradičně odkazuji na v tomto směru poměrně bohatou literaturu.

Z literární reprezentace, kterou průběžně zmiňuji, mohu doporučit z těch starších titulů například Lamsera (1966, s. 258–326), Kerlingera (1972, s. 582–663), Zicha (1976, s. 214–252) či Nowaka (1975, s. 261–270), speciálně se věnujícího statistické indukci, též názorného Kapra a Šafáře (1969, s. 143–200). Kromě toho i dosud nezmíněné texty, spíše statistické monografie, ovšem relativně dobře srozumitelné, Lamsera a Růžičku (1970) a příspěvek Nováka, Hátleho aj. (1969) do už citovaného Perglerova sborníku. Z novějších jsou to mj. nám již známí – Disman (1993, s. 180–278), Jeřábek (1992, s. 90–119), Punch (2008, s. 82–84), velice přehledný a podrobný Chráska (2007, s. 35–149), nepochybně také Hendlova monografie o zpracování kvantitativních dat (Hendl, 2004) a řada dalších autorů. (Hledání literárních pramenů může též začít u Maříková, Petrusek, Vodáková aj., 1996, s. 1433–1434, s. 1318–1320, s. 56, s. 619–620, kde jsou odkazy na další relevantní hesla i významné, především ale cizojazyčné publikace.)

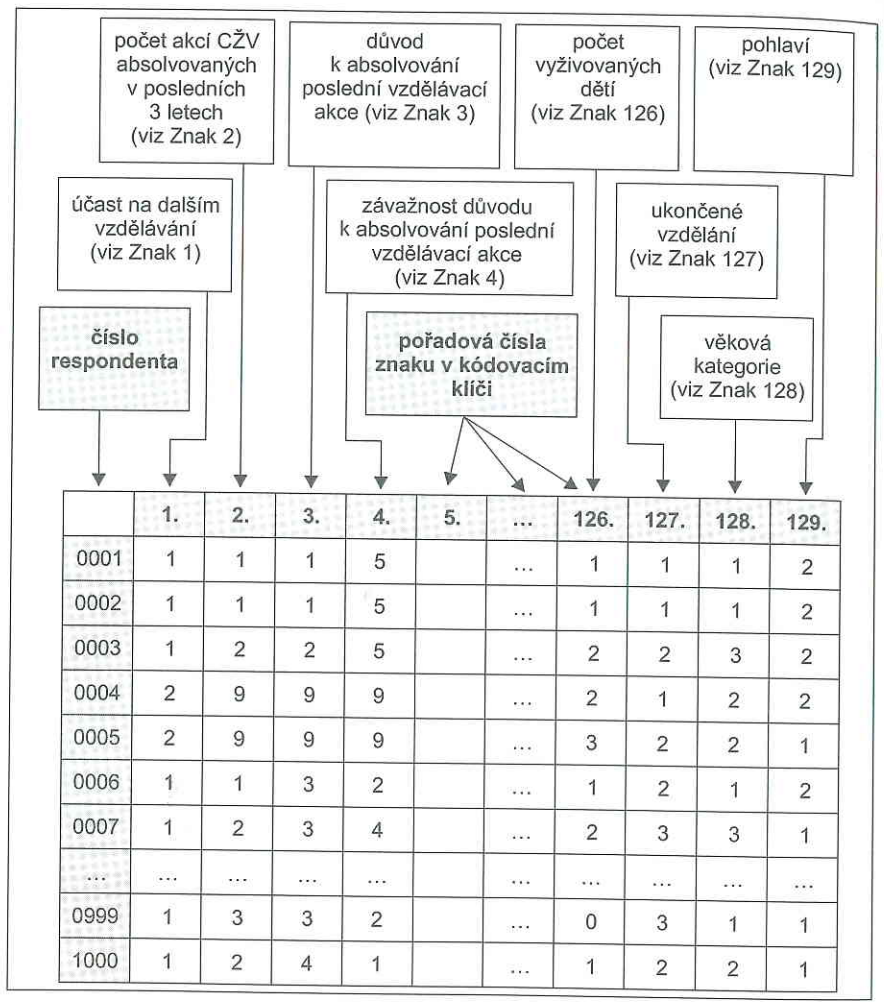
Kvantitativní data standardně procházejí následujícím procesem: **kódování → třídění 1. stupně → úpravy znaků → třídění 2. stupně** (případně dalších). U těchto fází se postupně zastavím.

7.1.1 KÓDOVÁNÍ

Kódování kvantitativních, tj. standardizovaných dat probíhá tak, že **jednotlivým variantám/kategoriím každého znaku jsou přiřazovány symboly, obvykle číselné**. (Pojem kódování se užívá i při zpracovávání kvalitativních údajů, má zde ale poněkud odlišný význam, viz 7.2.) **Materiálem pro kódování** tedy jsou odpovědi respondentů, kteří v rámci rozhovoru (viz 5.2.4) nebo vyplňování dotazníku (5.2.5) vybírali z předem stanovených variant. Také záznamy frekvence různých prvků jisté pozorovací kategorie (viz 5.1) či údaje o výskytu určitých podob záznamové jednotky (5.3). Sbírají-li údaje **pozorovatelé či tazatelé**, díky programům instalovaným v počítačích často **kódují data sami přímo v terénu** (viz CAPI v 5.2.4). V tradiční situaci kódéři probírají záznamové archy, vyplněné dotazníky atd. a „ťukají" údaje do počítače ve výzkumném centru.

Varianty znaků se kódují podle předem připraveného tzv. kódovacího klíče. (Zásadní rozdíl od kvalitativního výzkumu, v němž se kódy průběžně tvoří až při sběru údajů.) Ten určuje **způsob kódování** každého ze znaků a též mu vymezuje **standardní pořadové místo** mezi ostatními znaky (určitý sloupec). Ukázku klíče přináší schéma 6 na následující straně. Pro různé druhy znaků (viz 3.3.2) zde pak má **kód hodnotu čistě symbolickou** (u kvalitativních – pohlaví, důvod absolvování vzdělávací akce), případně jistou **míru pořadovosti** (u ordinálních – závažnost důvodu k absolvování vzdělávací akce, vzdělání) nebo **číselnou** (u kardinálních – počet vyživovaných dětí, počet absolvovaných akcí). Ze schématu je patrné, že pro ilustraci

a jistou kontinuitu užívám příklad, v němž jde o data/znaky z již dříve uváděného fiktivního zkoumání účasti respondentů na CŽV, patrně technikou rozhovoru či dotazníku (počátek zmínek o něm viz v 3.3.1 a v tab. 2). Tento příklad průběžně připomínám i níže v textu, proto hovořím hlavně o otázkách a odpovědích, byť data získáváme i z jiných zdrojů (viz předchozí odstavec).



	1.	2.	3.	4.	5.	...	126.	127.	128.	129.
0001	1	1	1	5		...	1	1	1	2
0002	1	1	1	5		...	1	1	1	2
0003	1	2	2	5		...	2	2	3	2
0004	2	9	9	9		...	2	1	2	2
0005	2	9	9	9		...	3	2	2	1
0006	1	1	3	2		...	1	2	1	2
0007	1	2	3	4		...	2	3	3	1
...	...	...	...	...		...	...	...	...	...
0999	1	3	3	2		...	0	3	1	1
1000	1	2	4	1		...	1	2	2	1

Pokračování

Pokračování

Kódovací klíč (uvedené znaky viz tab. 2 v 3.3.1):

Znak 1 – účast na dalším vzdělávání (*POZOR filtr!*) ano = 1 ne = 2 (*potom znaky 2, 3, 4 = 9!*)	**Znak 126 – počet vyživovaných dětí** počet dětí = uvedený počet číslem
Znak 2 – počet akcí CŽV v posledních 3 letech počet akcí = uvedený počet číslem	**Znak 127 – ukončené vzdělání** základní = 1 středoškolské = 2 vysokoškolské = 3
Znak 3 – důvod k absolvování poslední vzdělávací akce zvýšení kvalifikace = 1 rekvalifikace = 2 mimoprac. zájem = 3 jiné důvody = 4	**Znak 128 – věková kategorie** do 35 let = 1 36–50 let = 2 51 a více = 3
Znak 4 – závažnost důvodu k absolvování poslední vzdělávací akce zcela závažný = 5 spíše závažný = 4 spíše nezávažný = 2 zcela nezávažný = 1	**Znak 129 – pohlaví** muž = 1 žena = 2
atd. = Znak 5–125	**konec kódovacího klíče**

Schéma 6 *Ukázka kódovacího klíče a matice zakódovaných dat*

Pravidla pro kódování a tedy pro tvorbu kódovacího klíče nejsou jednoduchá (srv. např. Hendl, 2004, s. 76–79). Pro náš účel zmíním pouze několik důležitých bodů. Předně se doporučuje **nekódovat chybějící data „0“**, která potom výpočty (zvláště u kvantitativních ukazatelů) zkresluje a má zůstávat vyhrazena hodnotám čistě nulovým, jako počet dětí aj. (Viz u schématu 6 kódovací klíč pro Znak 1, variantu 2 a v tabulce respondenty 0004 a 0005 s devítkami a respondenta 0999 s 0 dětí.)

Dále je třeba vyřešit **kódování otázek s tzv. vynucenou volbou** (viz 5.2.3), tj. hodnocení (škálování postoje) v takovém tvaru, že střední kategorie chybí (obvykle „3“, viz schéma 6, Znak 4). Pokud je takových více a s jejich známkami se dále operuje (sčítají se, průměrují atp.), hodnota výpočtu je lépe interpretovatelná, **kódují-li se nikoli 1–2–3–4, ale 1–2–4–5** (podrobněji ale do otázek škálování nechci zabíhat).

Konečně se též musí **určit kódování u všech polytomických výčtových otázek** (též u výčtových s pořadím, oba typy 5.2.2). Nemá-li tato otázka příliš variant, lze **pro každou jejich kombinaci stanovit samostatný kód**, tvoří-li ji **větší počet kategorií, je nutné kódovat ji jako baterii dichotomických otázek** (viz následující kurzivu). To vše je třeba při tvorbě kódovacího klíče uvažovat, mj. tedy i **kolik nakonec takovýto znak vlastně zabere sloupců**. (Jeden pro každou dichotomickou otázku a u výčtových s pořadím ještě jeden sloupec pro pořadí!) Z těchto důvodů **bývá počet znaků většinou znatelně vyšší než počet otázek**.

Za příklad si vezměme ze schématu 6 Znak 3 – důvod k absolvování poslední vzdělávací akce. Pokud by instrukce zněla „Vyberte nejsilnější důvod/y, který/é vás vedl/y k absolvování poslední vzdělávací akce" apod., respondent by mohl vybrat i více důvodů, šlo by totiž o otázku výčtovou. Její kódování by pak nebylo 1–2–3–4, jak máme teď, ale kromě těchto kódů dále 5 = 1 + 2; 6 = 1 + 3; 7 = 1 + 4; 8 = 2 + 3; 9 = 1 + 2 + 3 atd. I tady, jak je patrno, dojdeme ke značnému počtu variant znaku. U některých se ovšem tímto způsobem ani postupovat nedá. Například v mediálních výzkumech klasická otázka na vybavení domácnosti zahrnuje cca 50 položek (auto, motorka, kolo, chata, chalupa, až po počítač, tiskárnu, topinkovač, kočku, psa, kanárka a jinou faunu). V tomto případě se tedy otázka kóduje jako n-tice, konkrétně padesátice, odpovědí dichotomických 1 = ano / 2 = ne, takže zabere 50 sloupců tvořených jedničkami a dvojkami.

Zakódovaná a nakonec do počítače vkládaná data si lze představit jako rozsáhlou **virtuální matici N krát n**, kde N = počet řádků, čili velikost výběrového souboru (počet respondentů), a n = počet sloupců, neboli počet zkoumaných znaků (viz schéma 6). Jak plyne z předchozích zmínek, jejich analýzy se v dnešní době už ve valné většině provádějí na počítačích pomocí speciálních **statistických programů** (STATISTICA Cz, SPSS ad.). Některé jsou více, jiné méně uživatelsky přátelské, v každém případě lze ale doporučit **konzultace s odborníkem** (jejich zkušeným uživatelem, statistikem apod.) o **relevantním využití jejich možností i o volbě adekvátních statistických operací**, tj. vhodných testů, korelačních koeficientů atd. (viz 7.1.3 a 7.1.4).

7.1.2 TŘÍDĚNÍ 1. STUPNĚ

V prvním kroku analýzy dat, v třídění 1. stupně, se zjišťuje, kolikrát se v určitém sloupci, tj. u nějakého znaku, vyskytuje určitý kód, tedy některá z kategorií znaku. Jinak řečeno – **pro každý znak se identifikuje četnost výskytu jeho jednotlivých variant**. Tak se ukáže, kolik je v souboru mužů a žen, kolik osob s jakým vzděláním, kolik jich neabsolvovalo žádnou vzdělávací akci, kolik uvedlo jaký důvod k účasti na poslední vzdělávací akci atd. Tyto údaje jsou udávány ve faktických číslech, v tzv. **absolutních četnostech**, i procentech, v tzv. **relativních četnostech**. Často se prezentují v tabulkách, popř. lze použít i grafů. Ukázky těchto tabulek přináší následující tabulka.

Tab. 10 *Ukázky tabulek třídění 1. stupně*

Znak 3 – důvod k absolvování poslední vzdělávací akce		
varianta	ač	rč = %
1 – zvýšení kvalifikace	300	30
2 – rekvalifikace	400	40
3 – mimopracovní zájem	270	27
4 – jiné důvody	30	3
celkem respondentů	1 000	100

Znak 127 – ukončené vzdělání		
varianta	ač	rč = %
1 – základní	400	40
2 – středoškolské	500	50
3–vysokoškolské	100	10
celkem respondentů	1 000	100

Vysvětlivky: – znaky a kódovací klíč viz schéma 6
– kvůli zjednodušení neuvažována jinak běžná kategorie „neodpověděl/a"

Zkratky: ač = absolutní četnosti; rč = relativní četnosti, tj. %

Dlužno poznamenat, že užívání relativních četností je u souborů do velikosti sta prvků, nebo pohybující se těsně kolem ní, většinou zbytečné a někdy může působit až komicky. Disman tuto nezřídka se objevující tendenci ilustruje na příkladu, který se ve výzkumných kruzích traduje jako údajný citát z jedné závěrečné zprávy: „Po aplikaci preparátu B se 33,3 % kuřat uzdravilo, 33,3 % uhynulo a o zbývajících 33,3 % nejsme schopni poskytnout uspokojující informaci. Dosud se nám nepodařilo to třetí kuře chytit." (Disman, 1993, s. 186)

Dalšími údaji v 1. stupni třídění jsou **základní statistické míry znaků**, které z předchozích dat vycházejí. **Porovnáváním četnosti jednotlivých variant znaku se ukazuje podoba jeho rozložení**, tzv. distribuce případů

(např. odpovědí respondentů). Sleduje se, zda všechny kategorie určitého znaku jsou obsazeny rovnoměrně či nikoli. Zda s nárůstem hodnoty kategorií (týká se pořadových a hlavně kvantitativních znaků, např. vzdělání, počtu dětí aj.) roste nebo klesá počet případů. Popř. zda jde o rozložení blížící se tzv. normálnímu (v němž se nejčastěji objevují střední hodnoty a čím jsou menší nebo větší, tím jejich četnost klesá) aj. Ukazatelů popisujících rozložení je bohatá paleta a jako obvykle zmíním jen několik základních.

Jedním je **modus** (Mo), též modální hodnota znaku, tj. ta, co se u něj objevuje nejčastěji. (V našem šetření je modální důvod k absolvování vzdělávací akce varianta „rekvalifikace" a modálním vzděláním středoškolské, viz tab. 10.) **Modus lze určit u všech třech typů znaků** (viz 3.3.2), tj. u kardinálních, ordinálních i nominálních. (U těchto konkrétně není ani možné jinou statistickou míru použít.)

Medián (Me), čili mediánová hodnota, rozděluje soubor seřazený podle hodnoty znaku, která byla u každého prvku zjištěna, na dvě poloviny – v jedné jsou prvky/osoby s hodnotou stejnou nebo nižší než medián a ve druhé s hodnotou stejnou či vyšší. Je-li ve hře výše hodnot, znamená to tedy, že **Me** lze zjišťovat především u **kardinálních znaků a v omezené míře u znaků ordinálních** (viz následující kurzivu), ale u znaků nominálních pochopitelně nikoli (viz 3.3.2).

*Je-li počet jednotek souboru **lichý**, například 1001, mediánová hodnota určitého znaku bude ta, kterou vykáže 501. člen řady. (Ten, který se pomyslně nachází uprostřed souboru seřazeného podle velikosti hodnot tohoto znaku, a nalevo i napravo od něho stojí 500 osob.) Pokud je počet prvků souboru **sudý** (např. 1000), mediánem je průměr hodnot, které vykazují oba prostřední členové řady (500. a 501.). **U kvantitativních znaků je tedy identifikace Me bez problémů.***

***Tak i u pořadových znaků při lichém počtu prvků.** (Když náš modelový soubor, který bude mít nyní 1001 osob, z nichž 650 má základní, 250 střední a 101 vysokoškolské vzdělání, seřadíme podle jeho výše, mediánem bude hodnota 501. člena této řady, tj. vzdělání základní.) **U sudého počtu prvků je medián identifikovatelný** v případech, pokud **dva střední členové řady vykazují stejné varianty znaku.** (V souboru 1000 osob, kdy 650 má základní, 250 střední a 100 vysokoškolské, je 500. a 501. se vzděláním základním, tj. mediánovou hodnotou.) **Jestliže ale nesou rozdílné hodnoty, nelze Me určit.** (Při rozložení 500 vzdělání základní, 300 střední a 200 vysokoškolské mají 500. a 501. člen řady rozdílné varianty vzdělání, jeden základní a druhý středoškolské, z nichž pochopitelně nelze průměr vypočítat.)*

Dalšími mírami podobnými mediánu jsou **kvartily**, decily, percentily aj. (obecně kvantily), které obdobným způsobem rozdělují soubor na **čtyři**, deset, sto aj. počet stejných částí a umožňují distribuci hodnot ještě lépe a přesněji vymezit. **Tři kvartily** (tzv. dolní, prostřední a horní, též 1., 2. a 3., z nichž 2., tj. prostřední, je vlastně medián) např. zřetelně ukazují rozložení tzv. extrémních, tj. nestředových hodnot, tedy 25 % souboru spadajících pod dolní/1. kvartil a 25 % respondentů patřících nad horní/3. kvartil (viz použití v delfské metodě, 6.4). V těchto ohledech jsou kvartily poměrně názorné a bývají často prezentovány také v grafické podobě (např. jako tzv. krabičkové grafy).

Aritmetický průměr (M nebo $\bar{x}$) je obecně známá míra, představující součet všech zjištěných hodnot dělený jejich počtem. Je tedy aplikovatelná **jen u kardinálních znaků** (popř. u kvantifikovaných škálových hodnot). Nevýhodou M, narozdíl od Me, je jeho značná citlivost vůči extrémním hodnotám. (Vyjádřená tradičním příkladem, kdy jeden sní celé kuře, druhý slintá hladem, ale statisticky vzato oba měli půlku kuřete.) M též často udává hodnotu, která se ve skutečnosti ani nevyskytuje (např. 1,53 dítěte) nebo se v měření vůbec neobjevila (např. u řady 2, 4, 5, 7, 8, 10 je M = 36 : 6 = 6).

Proto je vhodným **doplňkem průměru směrodatná odchylka** (s), která ukazuje na **míru jeho homogenity**, neboť měří vzdálenost jednotlivých konkrétních měření od vypočítané hodnoty průměru: **Rozptyl je průměrná kvadratická odchylka měření od zjištěného průměru, směrodatná odchylka pak odmocninou tohoto rozptylu.** Když průměr vznikl z totožných hodnot (např. 3, 3, 3, 3, 3, 3; M = 3), je s = 0, roste s oscilací reálných hodnot kolem průměru (2, 2, 2, 4, 4, 4; M = 3, s = 1) a je maximální pro situace, kdy se jedná o průměr z krajních hodnot (1, 1, 1, 5, 5, 5; M = 3, s = 2).

7.1.3 ÚPRAVY ZNAKŮ

Podle dat z 1. stupně třídění (absolutní i relativní četnosti, analýzy rozložení hodnot znaků) se dále **znaky případně upravují k třídění 2. stupně** (nebo dalších, viz 7.1.4). Někdy se v této souvislosti hovoří o tzv. transformaci či **rekategorizaci znaků**, která kromě jiných zahrnuje většinou také následující případy.

Mezi údaji se často vyskytují **škálové znaky** (o jejich kódování 7.1.1), s kterými bylo pracováno nejen v rámci celého souboru, ale též u jednotlivých respondentů (každý z nich obdržel součet svých škálových hodnot, jejich průměr apod.). Půjdou-li tyto **nově vzniklé znaky** do dalších analýz, je

třeba z nich **vytvořit nové kategorie (např. jako určité intervaly těchto obdržených hodnot) a těm přidělit nové kódy**.

Některé znaky vstupovaly do 1. stupně třídění se značným, dokonce **extrémním počtem různých variant**, například uváděl-li respondent přesně svůj věk či měsíční příjem. Tím se sice dal získat poměrně plastický obraz o rozložení hodnot těchto znaků (viz 7.1.2), ale pro další operace je nezbytné u každého **stanovit přijatelný počet kategorií**. To se obvykle provede vytvořením intervalů, které ovšem musejí reflektovat koncepci úkolu, formulované hypotézy apod. (V našem případě jsme vymezili věkové kategorie již v kódovacím klíči, viz schéma 6. Pokud bychom se ale ptali na přesný věk respondentů, nyní bychom museli tyto údaje kategorizovat – třeba tak, jak už máme.)

Úprava znaku je také nutná, jestliže **některá z jeho kategorií, jak vyplyne z 1. stupně třídění, vykazuje velmi nízkou četnost**. Ta se totiž ve 2. stupni třídění ještě dále rozmělní, neboť tento znak bude tříděn s nějakým jiným. (Například bude v souboru pouze 10 vysokoškoláků, a má se srovnávat výše vzdělání s velikostí místa bydliště, kde je 9 kategorií velikosti. Ve vyšších stupních třídění se dokonce jedná o porovnávání tří a více znaků.) Následkem toho by se v některých polích tabulky mohly objevit četnosti zcela miniaturní. Jestliže však statistické testy, které budou dále použity, jako podmínku své korektnosti stanovují pro jednotlivá pole tabulky určité minimální četnosti (např. jen ve 20 % polí může být hodnota nižší než 5 a v žádném poli nesmí být nula), je třeba tuto situaci eliminovat. Není jiný způsob než onu **málo obsazenou kategorii sloučit s jinou**, ovšem tak, **aby ta nově vzniklá byla smysluplně definovatelná**. U kvantitativních a pořadových znaků se zákonitě slučují **sousední kategorie**. (Z původních variant 5, 6, 7, 8 dětí vznikne nová kategorie „5 a více dětí", z variant vzdělání základní, středoškolské, vysokoškolské se utvoří dvě kategorie, „základní" a „středoškolské a vyšší".) U znaků nominálních se musí najít nějaký **obecnější pojem**, který by slučované varianty zahrnul, zastřešil, například zahradník, zedník, instalatér, kominík se sloučí do kategorie „manuální profese". (Na těchto případech se názorně ukazuje, že některé problémy s miniaturními četnostmi může částečně řešit také dostatečně velký výběrový soubor, jak už jsem uvedl v 4.5, včetně příkladu v kurzivě.)

Konečně je třeba též **zpracovat otázky polouzavřené a volné**, pokud se budou v rámci zkoumání vůbec vyhodnocovat (mělo by jich být minimum, o obou typech viz počátek 5.2.2). Ty nešlo zakódovat už v terénu, jelikož jejich kategorie se vytvářejí až nyní, na základě výpovědí respondentů, takže logicky nemohly být součástí kódovacího klíče. Údaje, výroky se shromáždí (vypíšou z dotazníků, či vytisknou, jsou-li zapsány v elektronické podobě jako u CATI, CAPI, CAMI, viz 5.2.4 a 5.2.5, atp.) a **hledají se pro ně nějaké obecnější definovatelné kategorie**, které zahrnou co nejvíce případů, tzn. které budou jejich **„společným jmenovatelem"** (zastřešující pojmy, třídy a skupiny podobných fenoménů aj., jako při slučování variant nominálních znaků, viz výše).

7.1.4 TŘÍDĚNÍ 2. STUPNĚ

Třídění 2. stupně vlastně představuje **porovnávání rozložení určitého znaku v jednotlivých podsouborech, které jsou v rámci zkoumaného souboru vytvářeny variantami dalšího znaku**. Jinými slovy se kategorie vzniklé 1. stupněm třídění dále člení podle ještě jiného znaku. V našem případě to může být například výše vzdělání tříděná dle důvodů vedoucích k absolvování poslední vzdělávací akce. (Neboť jedna z hypotéz, viz 3.3.3, třeba další rozpracování H3, mohla znít v tom smyslu, že osoby s různým stupněm vzdělání měly také různé důvody, proč vzdělávací akci absolvovaly.) **Jaké znaky se mají vzájemně srovnávat** (proměnné nezávislé se závislými, o nich viz ke konci 3.3.2) tedy **vyplývá z hypotéz**. (Při extenzivnějším přístupu, zvláště pokud není znaků příliš mnoho, se někdy srovnává každý s každým, což sice není striktně exaktní, ale občas to přináší zajímavé a netušené informace.) Srovnávání rozložení dvojic znaků podává počítačový program v tzv. **kontingenčních tabulkách** (samozřejmě je možné použít kromě nich i vhodné grafické vyjádření). Pro zmíněné znaky vzdělání a důvod absolvování vzdělávací akce by taková tabulka vypadala třeba jako tabulka 11.

Tabulka přináší řádkové a sloupcové četnosti, absolutní i relativní. (Někdy pro větší přehlednost bývají k dispozici tabulky dvě, jedna pro četnosti řádkové, druhá pro sloupcové.) **Sloupcové četnosti** ukazují kolik případů jednoho sloupce (kolik % z něj) se podílí na prvním řádku, kolik na druhém řádku atd., v našem případě tedy kolik osob se základním vzděláním (kolik % z nich) sdělilo důvod 1, kolik důvod 2 apod. **Řádkové četnosti** pak říkají, kolika případy z jednotlivých sloupců (% z nich) je tvořen první řádek, kolika druhý atd., v našem případě kolik osob (% z nich), co uvedly důvod 1 (zvýšení kvalifikace), má vzdělání základní, kolik středoškolské ad. Z těchto údajů například zjistíme, že polovinu těch, co uvedli jako důvod rekvalifikaci, tvoří osoby se středním vzděláním, u důvodu zvýšení kvalifikace jsou jich dvě třetiny se vzděláním základním, 70 % vysokoškoláků uvádí jako

důvod mimopracovní zájem atd. (Tady je příhodný moment požádat čtenáře/ku, aby si jako cvičení takto postupně prošel/la celou tabulku.)

***Tab. 11** Tabulka třídění 2. stupně s řádkovými a sloupcovými četnostmi (vzdělání x důvod k absolvování poslední vzdělávací akce)*

Znak 3	Znak 127						součet řádků		celkový součet	
	1–ač	sl%	2–ač	sl%	3–ač	slp%	Σač	*Σřád%*	ač-ř/N	*%ř/N*
1–ač	200	50,0	90	18,0	10	10,0	300		300	
řád%		*67,0*		*30,0*		*3,0*		*100,0*		*30,0*
2–ač	140	35,0	250	50,0	10	10,0	400		400	
řád%		*35,0*		*62,5*		*2,5*		*100,0*		*40,0*
3–ač	50	12,5	150	30,0	70	70,0	270		270	
řád%		*18,5*		*55,5*		*26,0*		*100,0*		*27,0*
4–ač	10	2,5	10	2,0	10	10,0	30		30	
řád%		*33,3*		*33,3*		*33,3*		*100,0*		*3,0*
Σač	400		500		100		1 000			
Σslp%		100,0		100,0		100,0				
ač-sl/N	400		500		100				1 000	
%sl/N		40,0		50,0		10,0				100,0

Vysvětlivky (pro přehlednost neuvažována jinak běžná kategorie „neodpověděl/a"):

Znak 3 – důvod k absolvování poslední vzdělávací akce:	**Znak 127 – ukončené vzdělání:**
1 = zvýšení kvalifikace	1 = základní
2 = rekvalifikace	2 = středoškolské
3 = mimopracovní zájem	3 = vysokoškolské
4 = jiné důvody	

Zkratky: ač = absolutní četnosti; slp% = sloupcová procenta; *řád%* = řádková procenta (kurzivou); Σ = suma, součet; ač-ř = absolutní četnosti řádků; ač-sl = absolutní četnosti sloupců; %ř = procenta řádků (kurzivou); %sl = procenta sloupců; /N = z celkového počtu

Pro názornost jsou četnosti v tabulce 11 odlišeny celkem markantně, ale v praxi je třeba dávat si velký pozor na výroky, že je něčeho více či méně, že se něco vyskytuje zřídka a něco naopak častěji aj. Zvláště pokud jde o **statistickou tzv. relační hypotézu** („Vysokoškoláci častěji absolvují vzdělávací akce z mimopracovních důvodů" atp.), kterou je vždy nutné **statisticky**

ověřit. Určitě se při tom nedá **vystačit s prostým, optickým srovnáváním** četností jednotlivých variant znaků, jejich rozložení apod., a na jeho základě usuzovat na vztahy mezi znaky, na podobnosti a rozdíly mezi určitými podsoubory atd. Takovéto analýzy je nezbytné provádět pomocí statistických nástrojů, prostřednictvím různých korelačních koeficientů a statistických testů, které případné vztahy, podobnosti, rozdíly ad. prověřují z hlediska jejich statistické průkaznosti.

***POZOR** – nevylučuji, že následující odstavce mohou každého statistika nebo toho, kdo zná tyto otázky trochu blíže, rozplakat, popř. rozesmát. Bohužel jsem ale nenalezl jiný způsob, jak celou problematiku testování, které je věnována nejedna monografie, ve zkratce vysvětlit.*

Podstatou statistických testů je tzv. **testování statistických hypotéz**. Pro různé druhy znaků se používají adekvátní testy. Jde-li například o porovnávání dvou kvalitativních znaků, je nutné užít jiný test než v případě srovnávání kvalitativního s pořadovým či kvantitativním znakem atp. (podrobněji viz např. Chráska, 2007; Hendl, 2004; a jiní). Většina testů pracuje s tzv. **nulovou hypotézou**, formulovanou v tom smyslu, že mezi znaky, soubory apod. neexistuje žádná závislost/podobnost atp. a všechny pozorované indicie v tomto směru jsou pouze náhodné. Oproti ní stojí tzv. **alternativní hypotéza**, která nemusí být jejím přísně logickým opakem, avšak určitou závislost předpokládá.

Princip většiny testů spočívá velmi obecně a stručně řečeno v matematickém porovnávání skutečně naměřených hodnot s jakýmsi obecným, teoretickým modelem jejich rozložení. Z něj jsou odvozovány tzv. **kritické hodnoty testu**. Ty jsou pro různé testy stanoveny ve statistických tabulkách a jejich výše se řídí **velikostí souboru** a tzv. **hladinou významnosti testu**. (Tj. mírou rizika, se kterým by mohla být nulová hypotéza tímto postupem zamítnuta, přestože je platná, nebo naopak přijata, i když je neplatná. V sociálních výzkumech se nejčastěji pracuje s hladinami významnosti alfa = 0,05 nebo alfa = 0,01, což představuje 95% nebo 99% pravděpodobnost správnosti rozhodnutí.) Je-li vypočítaná testová charakteristika vyšší než kritická hodnota testu (popř. spadá-li do tzv. oblasti zamítnutí), uvažuje se, že **nulová hypotéza je neplatná** (přesněji řečeno že existují podklady pro její zamítnutí). **Připouští se platnost alternativní hypotézy, totiž že vztah, podobnost, rozdílnost ad. nejsou jen náhodné.**

Počítačová zpracování testů také běžně ukazují, zda lze test považovat za **korektní** (zda byly splněny všechny podmínky k jeho aplikaci, viz mini-

mální četnosti v 7.1.3) a na jaké hladině významnosti je možné vztah mezi znaky považovat za **statisticky prokazatelný**.

Dále se užívají **různé druhy koeficientů** (kontingence, regrese, korelace apod.). Ty indikují, jak je **vztah mezi sledovanými znaky/proměnnými těsný**, v případě testů jak jsou jimi **zjištěné souvislosti silné**. Obvykle se hodnota koeficientů **do 0,3** chápe jako **velmi slabý vztah** (určitá tendence), hodnota **0,3–0,6** jako **středně silný**, zřetelný vztah a **nad 0,6** jako **velmi silný**, zcela jasný vztah. Kromě častého rozpětí 0–1 nabývají některé z koeficientů také hodnot -1 až 1. (Například pořadová korelace, zkoumající, zda se pořadí určitých prvků statisticky liší, kupř. zda jsou mezi muži a ženami rozdíly v pořadí závažnosti důvodů zúčastnit se vzdělávání, jaký je rozdíl v pořadí důvodů při opakujícím se měření atp. Tehdy hodnoty směřující k 1 naznačují shodu v pořadí, k -1 opačné pořadí a pohybující se kolem 0 nezávislost obou pořadí.)

Třídění vyšších stupňů znamená porovnávání více znaků, tedy více podsouborů. **Třídění 3. stupně** by například srovnávalo výši vzdělání a důvod vedoucí k absolvování vzdělávací akce, to vše ještě z hlediska pohlaví. (Zde bychom dostali dvě kontingenční tabulky jako je tabulka 11 – jednu pro muže a druhou pro ženy.) Třídění 4. a dalších stupňů se používá jen zřídka. Jednak jejich výsledky přinášejí stále **komplikovanější informaci**, takže často bývají i **hůře interpretovatelné**. Současně **vyšší stupně třídění** představují rozpady vzorku na menší a menší podsoubory, u nichž pak hrozí nebezpečí minimálních četností (7.1.3). To **zákonitě vyžaduje rozsáhlejší výběry** (viz kurzivu o rozložení ve 4.5), což citelně a často neefektivně navyšuje výzkumné náklady.

7.2 VYHODNOCOVÁNÍ KVALITATIVNÍCH DAT

Způsoby třídění a vyhodnocování kvalitativních údajů jsou **různorodé jako tyto údaje samy**. Častým výchozím materiálem zde bývají, jak už si dokážeme na základě dosavadních poznatků o kvalitativním zkoumání představit, výzkumníkovy poznámky z rozhovorů, videozáznamy z pozorování, audionahrávky diskuzí ve skupinách, písemná vyjádření informantů, záznamy a výpisky ze studia dokumentů aj. Tak jako charakter těchto údajů, ani **jejich zpracování nemůže být standardizované, parametrizované** v té podobě, která je typická pro výše uváděné operace s kvantitativními daty. Existuje ale vcelku bohatá paleta různých třídicích, analytických, vyhodnocovacích ad. postupů, z nichž pro bližší představu některé základní uvedu. Pochopitelně se jim věnuje většina metodologických publikací, které se kvalitativním zkoumáním prioritně zabývají.

Z průběžně užívaného arzenálu pramenů doporučuji kromě Dismana, 1993, s. 311–321, též například monografii Miovského, 2006, s. 195–253, nebo Hendlovu, 2005, s. 207–269. (Vedle nich lze jistě najít mnohé další a neméně přínosné, pouze se o ně neopírám.)

K hlavním rozdílům při zpracovávání dat kvalitativních a kvantitativních patří, že **kvalitativní údaje se urovnávají, pořádají a zčásti třídí** nikoli všechny najednou, po ukončení terénních prací, nýbrž **postupně**, již při jejich průběhu. Nicméně platí, že **směr a rámec vyhodnocování údajů musí reflektovat původní cíle zkoumání, popř. jejich změny**, ke kterým se během realizace flexibilně dospělo.

Pojem **kódování** zde má trochu **odlišný význam** než při práci s kvantitativními daty. Jedná se o **rozčleňování množin a souborů údajů na dílčí celky, segmenty** a jejich následné pojmenovávání a třídění, případně klasifikování a kategorizování. To se během analýz několikrát opakuje a postupně vyjasňuje a zpřesňuje, takže díky tomu lze pracovat se stále rozsáhlejšími datovými soubory. (Pro tyto účely se dá použít také řada speciálních softwarů, například MAX pracující s narativními texty, o nich viz 5.2.4. Nebo ATLAS.ti, nakládající s tabulkami, obrázky, audiozáznamy a propojitelný s SPSS, a mnohé další – více o nich např. Miovský, 2006, s. 213–217.)

V tříděných a neustále přehodnocovaných materiálech hledá výzkumník řešení problému už při terénním získávání dat. Dospívá-li k tzv. **teoretické saturaci, tj. když vyhodnocování materiálu a jeho další sběr již nic nového nepřinášejí**, bývá výzkum, jak už z předchozího víme, zpravidla ukončen (viz závěr 4.5 včetně tab. 4).

Důležitým krokem analýz je **transkripce údajů**, jejich přepis z různých podob (z protokolů, audiozáznamů, videozáznamů apod.) do formy, s kterou bude možné **dále pracovat**, nejčastěji do formy písemné. (To však neznamená, že pouze do „papírové", neboť v elektronickém přepisu lze určitá témata, klíčová slova aj. vyhledávat pomocí speciálních programů rychleji a bezchybněji.) Existuje několik **variant transkripcí** (doslovná, selektivní, s komentářem ad.), které rozhodně **nemají být jen mechanické**, nýbrž **aplikovat již určité koncepční hledisko** (výběr témat, vypouštění nerelevantních pasáží, vlastní komentáře, pokusy o primární klasifikace např. situací, reakcí aj.).

Technika prostého výčtu je určitým „nakročením" k analýzám kvantitativního druhu. V podstatě se jedná, pokud to podstata a logika zkoumaného aspektu dovoluje, o **sledování frekvencí a popř. i intenzity výskytu** určitého fenoménu, prvku, znaku atp. Případně může být sledován v souvislosti s výskytem jiného prvku, i z hlediska širšího kontextu či podmínek jeho manifestace atd.

V tom se tento postup blíží tzv. **hledání vzorců**, tj. hledání určitých opakujících se momentů (např. reakcí, témat, způsobů jednání aj.), které jsou vázány na nějaké souvislosti, osoby, komunikační strategie apod. Výsledkem je jakýsi „obecný příběh", neboli vzorec, jak taková událost probíhá. Nesmí se pochopitelně jen favorizovat fakta a údaje, která konstruovaný příběh podporují, ale je nutné brát v úvahu také data, která jej popírají, a na jejich základě vytvářenou „story" poopravovat.

Dalším přístupem k utřídění údajů je **vytváření kategorií a typologií**. Podobně jako při zpracování volných otázek (viz 7.1.3) se hledají **obecnější kategorie** pro jistým způsobem si podobné prvky, jevy apod. Takové kategorie, pod které je možné zahrnout více případů a zároveň jsou dobře definovatelné, interpretovatelné. **Vytváření typologie** znamená tvorbu nějaké **soustavy typů**, z nichž každý zahrnuje jevy, fenomény, prvky aj. s **podobnou konfigurací** určitých znaků/vlastností (bohatí vysokoškoláci, osoby se základním vzděláním bez zájmu o další vzdělávání atd.).

Poměrně častým způsobem rozborů a vyhodnocování dat, hlavně údajů písemné povahy (i když to není nezbytnou podmínkou) je také **obsahová analýza**, kterou jsem už představil v souvislosti s výkladem studia dokumentů (viz 5.3). Pro kvalitativní data je vhodná především podoba nestandardizovaná, výjimečně nepatrně standardizovaná.

Při rozborech a pořádání údajů je užitečné, pokud to jejich povaha umožňuje, **sestavovat různé shrnující a diferencující tabulky**. Stejně tak **některé vztahy** (mezi jevy, výroky, kontexty, i nově tvořenými typologiemi, konstrukty aj.) **graficky vyjádřit** (obrázky, schémata, grafy apod.).

Pominu další v literatuře uváděné přístupy (kontrasty a srovnávání, faktorování, užívání metafor, rámcovou analýzu, narativní analýzu ad.), z nichž některé jsou si velmi podobné a někdy se jen jinak jmenují. Na závěr zmíním postup zvaný zakotvená teorie.

Přístup označený jako **zakotvená teorie** (grounded theory) je spojen se jmény Strauss, Glaser a posléze Corbinová (blíže viz mj. Hendl, 2005, s. 243–255; Miovský 226–238; případně samotnou monografii o ní, Strauss a Corbinová, 1999). Jedná se o poměrně podrobně propracovaný a dosti komplexní postup při práci s kvalitativními údaji, jehož cílem je **hledání**

specifické teorie platné pro zkoumaný problém (pro vymezenou populaci, pro její postoje ke vzdělávání atp.). Základními prvky jsou zde tzv. **koncepty** (teoretické pojmy), **kategorie** (jako základní prostředky budování teorie) a **propozice** (formulující zobecněné vztahy mezi koncepty a kategoriemi a mezi kategoriemi navzájem). **Analytický postup směřuje ke zjištění ústřední kategorie zkoumaného souboru údajů**, která se potom stává středem hierarchické sítě dalších kategorií. (Identifikace a precizování kategorií se odehrává v několika fázích, které se neustále prolínají a v nichž se materiál opakovaně analyzuje. Objevující se kategorie, které jsou tematicky příbuzné, potom naznačují existenci jedné centrální kategorie, která se stává ústředním tématem datového souboru.) Během analýz se koncepty a kategorie mezi sebou neustále srovnávají, proto se někdy o tomto přístupu hovoří jako o **metodě nepřetržitého porovnávání**.

Při práci s datovým materiálem probíhá **kódování**, které pořádá údaje pro další nakládání s nimi, pro jejich interpretaci, přeskupování, další třídění apod. V zakotvené teorii se rozlišuje trojí druh kódování. Jednak **kódování otevřené**, které se provádí při prvních analýzách údajů. Odhaluje určitá témata a člení je, třídí, kombinuje atp., tvoří prvotní kategorie, které charakterizuje, tj. identifikuje jejich základní vlastnosti. **Axiální kódování** představuje hledání příčin, důsledků, podmínek, interakcí, procesů ad., které odhalují vztahy mezi různými kategoriemi a případně je vzájemně propojují. (To je však třeba nahlížet nikoli na základě údajů, ale v intencích tvořeného teoretického konceptu, z kterého lze odvozovat, které znaky/vlastnosti kterých kategorií je možné smysluplně propojit.) Posledním kódovacím krokem je **selektivní kódování**, jímž začíná integrace výsledků. Znovu se přezkoumávají data a jejich kódy, vyhledávají se stěžejní témata a kategorie, které se mohou stát základem vznikající teorie, tj. pomáhají do ní integrovat také další témata a kategorie a vytvářejí s nimi propojené významové sítě. I v zakotvené teorii končí kódování a sběr dalšího empirického materiálu tehdy, nepřinášejí-li již nové výsledky, tj. jestliže se dosahuje zmiňované **teoretické saturace**, teoretického nasycení.

7.3 SEKUNDÁRNÍ ANALÝZA DAT A METAANALÝZA

Specifický případ práce s daty je tzv. **sekundární analýza dat**. Představuje **další využívání údajů**, např. archivovaných sestav z někdejších vlastních šetření, speciální analýzy publikovaných statistických informací či výsledků výzkumů, datových souborů získaných z jiných i zahraničních pracovišť atp. Nakládá se s nimi jako při předchozích postupech (viz 7.1 a 7.2, které je tak v podstatě možné označit za analýzu primární). **Sekundárně mohou být analyzovány údaje povahy kvalitativní i kvantitativní**. Jde o záležitost relativně levnou, určitě levnější než všechny tyto údaje získávat vlastními silami. Jistou nevýhodou však je, že data mohou být převzata i s případnými chybami, kterých se dopustili předchozí autoři (výzkumníci) a které lze dodatečně těžko odhalit.

Objem vědeckých poznatků získaných rozličnými výzkumy den ode dne narůstá. Pokud výzkumník/výzkumný tým nechce zůstat zcela odtržen od reality výzkumného dění, stále častěji potřebuje přehled o aktuálním stavu poznání zkoumaného procesu, fenoménu, jevu apod. Donedávna se tento problém řešil, a dosud stále řeší, především různými formami **výzkumných rešerší**. Ty jsou ovšem jen **kvalitativním, narativním popisem a zhodnocením** toho, co se ve zkoumané oblasti událo, verbálním shrnutím dosažených výsledků a případných diskusí o nich. Proto se začíná využívat také přístup, který **informace a řady dílčích poznatků o stejném či podobném aspektu, jevu, fenoménu apod. shrnuje a systematizuje kvantitativně, za pomoci statistiky – tzv. metaanalýza**.

Podstata metaanalýzy spočívá ve statistickém vyhodnocování výsledků velkého souboru výzkumných prací a studií na určité téma se záměrem tyto poznatky integrovat. Metaanalýza porovnává značný počet dílčích šetření a pracuje tak s poměrně rozsáhlým souborem zkoumaných objektů. Tím pádem mají její závěry i větší, obecnější platnost. Často ovšem **nejde o prosté srovnávání výsledků** výzkumů (tj. primárních analýz), ale **při práci s nimi se používá řada statistických ukazatelů, testů, koeficientů** atp. (7.1). Např. se v analyzovaných šetřeních zkoumají a vyhodnocují zjištěné korelace mezi jistými znaky/proměnnými, výše jejich hodnot, jejich statistická významnost, přijímání či zamítání určitých hypotéz atd.

Výsledky metaanalýzy tak poskytují výzkumníkům před zkoumáním určitého problému nesmírně důležitou **základní výzkumnou, konceptualizační orientaci**. A za druhé, což je možná ještě cennější, metaanalýza svým systematickým shrnutím **posouvá a prohlubuje poznání určitého problému**, neboť řada jeho aspektů byla dosud skrytá, resp. roztříštěně uložená v různých do té doby vzájemně izolovaných studiích.

O problematice metaanalýzy informuje přehledně např. Chráska, 2007, s. 147–149; velice edukativně Průcha, 1995, s. 69–73; značně podrobně včetně etap provedení a základních statistických operací a testů potom Hendl, 2004, s. 491–528; a další autoři.

7.4 INTERPRETACE DAT

V některých pramenech se lze setkat s názorem, že je vhodné **rozlišovat mezi prezentací dat a jejich interpretací**. Pokládám to za náhled vcelku opodstatněný (viz též popsané diference mezi čtvrtou a pátou částí závěrečné zprávy v 7.5), který je však možné splnit spíše **při analýzách kvantitativních dat**. **Popis** třídění dat a jejich dalších analýz, komentáře tabulek, grafů a provedených operací ad. představuje **prezentaci dat**. **Výklad** zjištěných výsledků (vysvětlení co znamenají, co z nich vyplývá, jaké závěry a jaká opatření atd.) **je interpretací dat**. Tyto dvě roviny práce s daty se ale v analýzách kvalitativních neodlučitelně prolínají (viz předchozí 7.2). U kvantitativních zkoumání je ovšem užitečné (přehledné a jednoznačné) **nejprve prezentovat samotné výsledky**, v jejich úplnosti a zcela objektivně, a **teprve pak k nim vyslovit své vlastní závěry, názory, přesvědčení, domněnky, doporučení** ad.

První část je činností dosti drobnohlednou a přesnou, systematickou, „mravenčí" výzkumnou prací. V interpretaci se již dostává ke slovu určitá zkušenost s nakládáním s daty, ale i kreativita, jistá výzkumná „imaginace", schopnost logického myšlení a současně domýšlení abstraktních dimenzí (například nějakých zjištěných trendů) atd. Třeba však dodat, že pokud se výzkumník myšlenkově „rozmáchne", neměl by přitom ztratit rovnováhu. Interpretace totiž mají také svá určitá úskalí.

Především je třeba mít neustále na paměti, že **sociální výzkumy pracují v dosti specifických podmínkách a s dosti specifickými daty** (viz též 2.1). **Všechno zjištěné je jen jakási tendence, jistý náznak, možný trend**, a na žádné údaje, ať už podpořené statisticky (kvantitativní přístupy), nebo výzkumnou imaginací či empatií (kvalitativní zkoumání) nelze přísahat. Natož tvrdit, že něco nějak naprosto jednoznačně je, protože je to naprosto jednoznačně prokázáno. Povaha dat, způsoby jejich získávání, podstata s nimi

prováděných operací atd. totiž ani zdaleka neumožňují něco takového si myslet a už vůbec ne vyslovovat nahlas. To by byla pouze zpupnost pramenící z nevědomosti. (I když se s tím občas můžeme setkat u řady šetření, například různých marketingových či mediálních.) Zdůrazňuji, že to platí jak pro zjištěné statistické významnosti (viz ostatně 2.2 o podstatě statistické metody), tak tím spíše pro výsledky kvalitativních přístupů.

Tuto jistou nedostatečnost stran reálné, pravdivé reflexe reality ještě umocňuje **princip samotné interpretace**. Například Petrusek (1993, s. 94–95) k ní poznamenává, že výzkumná situace byla výzkumníkem již interpretována při tvorbě výzkumného plánu, při formulacích hypotéz atp. Takže on na základě dat interpretuje skutečnost již jednou interpretovanou, čili provádí interpretaci druhého řádu. A interpretace dat přitom není nic jiného než překlad matematicko-statistické informace do přirozeného jazyka. A ještě ke všemu je intersubjektivně značně nereliabilní.

Současně je nutné stále počítat s tím, že jevy, procesy, prvky, fenomény, vše, co vytváří sociální skutečnost a je předmětem zkoumání, vykazuje značnou **proměnlivost a mnohovrstevnost** apod. Proto každá redukce, kterou jsme byli nuceni provést (viz závěr 3.2), může mít určité následky, které však často nejsme schopni postihnout. Část informace se zákonitě ztrácí, neboť zkoumáme jen některé (vybrané) objekty a u nich také jen některé z bohaté škály proměnných. Pro tyto redukce v praxi neexistuje žádný exaktní algoritmus, vycházejí jen z konceptuálních úvah výzkumníka a jeho adekvátní dekonstrukce zkoumaného problému. Disman (1993, s. 21–27) upozorňuje, že některé **redukce mohou zapříčinit určité interpretační chyby**. (Studentům/kám navrhuji, že jako cvičení mohou následující situace znázornit graficky a vygenerovat pro ně vlastní příklady.)

Jednou z nich je tzv. **nepravá korelace**, kdy nezahrnutý aspekt C ovlivňuje znak A i znak B, takže to vypadá, jako kdyby mezi A a B byla nějaké spojitost. (Disman uvádí dnes už legendární případ „korelace" mezi vysokým výskytem čápů a též malých dětí v určitém regionu. Byť by to bezesporu byl významný vědecký objev, vše je dáno tím, že nezahrnutou proměnnou zde byla vysoká kvalita životního prostředí, která pak přitahuje jak čápy, tak mladé rodiny.)

Další chybou může být tzv. **vývojová sekvence**, kdy se nám zdá, že A ovlivňuje B, ale je tomu tak proto, že na A působí proměnná C, která ovšem nebyla do šetření zahrnuta. (Například osoba se vzdělává, aby uspěla na trhu práce, ale tou původní příčinou zájmu, která nebyla šetřena, byl pocit méněcennosti, jenž zvýšením vzdělání odplynul.)

Chybějící střední člen představuje situaci, kdy vidíme působení A na B, ale už není patrné, že příčinou změn B není bezprostředně A. To totiž působí nejprve na C, které nebylo do šetření zahrnuto, a teprve C působí na B. (U objevené souvislosti, že mladí delikventi pocházejí z bohatých rodin, chybí mezičlánek. Jejich rodiče vydělávají, nemají na potomka čas a ten hledá náhradu vztahů a seberealizaci v partě, bez ohledu na to, zda má hodnoty sociálně eufunkční či disfunkční.)

Konečně **dvojí příčina** znamená, že na B nepůsobí pouze A, jak vidíme, ale také proměnná C, která ovšem nebyla do šetření zahrnuta. (Na dobrých výsledcích studia se nepodílí pouze píle samotného studujícího, ale také jeho okolí, které mu vypracovává seminární práce, píše za něho odborné články apod.)

Tenký led v interpretacích někdy představuje i **nakládání s exaktními údaji**. Za zmínku stojí alespoň dva nejčastější prohřešky: Předně nelze při testování statistických hypotéz (7.1.4) propadat iluzi, že veškeré statisticky významně vysoké koeficienty/těsné vztahy nutně musejí znamenat velké vědecké objevy a přínosy. Často jsou to naprosté banality a než je vypustíme „do světa", je třeba se nad nimi v tomto směru zamyslet. (Připomenu extrém – většinou se jako velice významný ukazuje statisticky průkazně vyšší počet studentů mezi osobami do 18 let, podobně důchodců mezi lidmi nad 60 let atp.) Tady platí, že dalším nezbytným krokem statistických analýz, aby vůbec měly nějaký smysl, je **hledat pro zjištěné statistické souvislosti relevantní interpretace**, jinými slovy jejich **logická vysvětlení v souvislostech sociální skutečnosti**.

Také není adekvátní, především při pouhé prezentaci dat, **užívat pro deskripci hodnotící výrazy**. Vyloučíme-li přímý záměr, děje se tak často asi bezděčně, ale v obou případech to podprahově působí, především na laiky. Je to nepřijatelné obecně, a například v předvolebních šetřeních zvláště. (Formulace „necelých 25 %" zní jinak než „téměř 25 %", obdobně jako „nedosahuje ani 40 %" a „dosahuje bezmála celých 40 %" atp.)

7.5 ZÁVĚREČNÁ ZPRÁVA

Závěrečná zpráva je – obrazně řečeno – nezbytnou tečkou za každou výzkumnou akcí. K její obsahové struktuře lze s jistou nadsázkou říci, že je to z **určité části vlastně projekt** (viz 3.6) **psaný teď nikoli v budoucím, ale v minulém čase**, což je zřetelnější spíše u kvantitativních zkoumání. Obsahy

jednotlivých tematických celků zprávy, jejích hlavních kapitol, níže shrnu s tím, že zpráva o akci kvalitativní se od toho pochopitelně může a nejspíš bude v řadě momentů lišit. (Vzhledem k tomu, co už o kvalitativním a kvantitativním přístupu víme, nepovažuji za nutné tyto diference specifikovat a jejich odvození nechávám na úvaze čtenáři/ce jako domácí úkol.)

K bližšímu seznámení doporučuji z průběžně citované literatury například Kerlingera (1972, s. 666–671), Zicha (1976, s. 252–261), Kapra a Šafáře (1969, s. 207–209), z novějších Pelikána (1998, s. 243–246), Punche (2008, s. 87–98), popř. Hendla, probírajícího obsah jak závěrečné zprávy z výzkumu kvantitativního (Hendl, 2004, s. 529–539), tak kvalitativního (Hendl, 2005, s. 315–333). Dají se samozřejmě najít mnozí jiní autoři, k úvodní orientaci může posloužit také Maříková, Petrusek, Vodáková aj. (1996, s. 1449 a další relevantní hesla).

Podobně jako v projektu **první část závěrečné zprávy, teoretická**, prezentuje zadání a cíle výzkumného úkolu a též jeho konceptuální rámec (viz 3.1, 3.2 nebo 3.4.1). Na základě studia relevantní literatury informuje o dosavadním řešení problematiky a o případné návaznosti na něj. Seznamuje se stěžejními proměnnými a s možnostmi jejich měření (tj. zdůvodní rozhodnutí o volbě designu zkoumání, viz 3.1). Vztahy mezi nimi popíše a pro názornost může i graficky vyjádřit. Dále operacionalizuje užívané pojmy a vymezuje zkoumané znaky včetně jejich variant a indikátorů (3.3.1 a 3.3.2). Uvádí výchozí hypotézy, popř. je konkretizuje do úrovně pracovních či statistických (3.3.3). V kontextu cílů zkoumání, hlavních proměnných a formulovaných hypotéz zdůvodňuje nutné tematické redukce (3.2). V závěru zmiňuje úvahy o zajištění validity a reliability zkoumání (3.5).

Metodologická část přináší informaci o objektech zkoumání, jejich základních charakteristikách a jejich dosažitelnosti, uvádí způsob jejich případného výběru (viz 4) a jeho oprávněnost, jakož i otázku reprezentativity (u zkoumání kvantitativních, viz 3.5 a 4.5). Dále představuje způsob výzkumného přístupu (experimentální či ex post facto, viz začátek 5), včetně konkrétních podob sběru dat (viz 5, popř. 6), argumentů pro jejich volbu a důvodů pro výběr konkrétního místa a času jejich aplikace. Odkazuje na užitou podobu techniky (záznamový arch rozhovoru či pozorování, maketu dotazníku aj.), která je připojena v příloze. Pokud se uskutečnil předvýzkum (tj. u kvantitativních akcí, viz 3.6), shrnou se zde poznatky a zkušenosti, které přinesl. Dále jsou uvedeny postupy při vyhodnocování dat (viz 7) a užité počítačové programy. Posléze se tu probere otázka validity a reliabilita výsledků. Nakonec mohou být diskutovány etické aspekty postupu (zajištění anonymity dotazování, vymezení experimentu atp., viz též 8).

Realizační část informuje o složení výzkumného týmu, o kompetencích členů a o plnění jednotlivých výzkumných etap. Vždy předkládá čerpání finančního rozpočtu výzkumné akce ve struktuře odpovídající projektu a dalším účetním požadavkům.

Kromě těchto tří **obsahuje závěrečná výzkumná zpráva na rozdíl od projektu části další:** prezentaci výsledků a interpretaci dat včetně diskuse, závěrů a doporučení. **Prezentace výsledků popisuje získaný empirický materiál** a provedené analýzy (viz 7.3). Předkládá například 1. stupeň třídění, vysvětluje případné úpravy znaků a s poukazem na hypotézy seznamuje s poznatky, které přinesl 2. a další stupně třídění, viz 7.1.2, 7.1.3 a 7.1.4. (U kvalitativních studií pak prezentuje utříděné údaje a souvislosti, které z nich lze vyvodit atp., viz 7.2.) **S výsledky seznamuje nejen slovními komentáři, ale taktéž prostřednictvím tabulek a vhodných grafů** (např. i tzv. „trojrozměrných"). Všechny informace musejí být přesně uvedeny a přesvědčivě doloženy (odkazem na tabelace, výpočtem aj.). Podle povahy zprávy (viz dále) je možné podle potřeby objasnit některé méně užívané statistické postupy a veličiny. Je-li empirický materiál značně rozsáhlý, doporučuje se prezentaci výsledků rozčlenit na menší, nejlépe tematické celky.

Poslední část zprávy, interpretaci dat, která je zakončena diskusí, závěry a doporučeními, bývá vhodné od předchozí části oddělit (viz též 7.4). To samozřejmě nic nemění na tom, že by mělo jít o postřehy, akcenty a návrhy podpořené analyzovanými daty, logicky zdůvodněné, myšlenkově konzistentní a hlavně takové, které se vztahují k základním cílům výzkumného úkolu, hlavním hypotézám atp. Zde už se nepoužívají ani čísla, ani tabulky, ani grafy. Měly by tu být vysloveny (a případně též vysvětleny) jak zjištěné skutečnosti, tak co se předpokládalo a nakonec nepotvrdilo, stejně jako co nového a překvapivého akce přinesla. Doporučení obsahuje dle povahy zkoumání například návrhy na modifikaci/rozšíření určitých teoretických a/nebo metodologických aspektů (v případě základního výzkumu), popř. návrhy na jistá opatření organizační, koncepční, strukturní, marketingová aj. (při aplikovaném či komerčním šetření).

Obdobně jako u projektu, také u **závěrečné zprávy platí, že pro její konkrétní podobu je rozhodující, pro koho je určena**. Například zpráva o výzkumu základním, který je financován z nějakého vědeckého grantu, by teoretickou a metodologickou část měla mít pojednánu velice podrobně a fundovaně, neboť se týkají samé podstaty výzkumného problému. Naopak v případě komerčního šetření zajímají zadavatele hlavně konkrétní

výsledky a závěry včetně doporučení. Tomu je tedy nutné přizpůsobit strukturu obsahu zprávy – podrobně pojednat tyto dvě pasáže, současně je též dostatečně obecně shrnout (tj. pro detailní i pro rychlé, orientační čtení) a výklad teoretických a metodologických aspektů zkoumání minimalizovat. Nicméně realizační část, především otázka čerpání rozpočtu, zajímá grantovou agenturu podobně intenzivně jako klienta či sponzora.

8. POZNÁMKY K ETICKÝM OTÁZKÁM VÝZKUMU

Etických problémů jsem se mj. dotkl při výkladu pozorování (5.1) a experimentu (5.4). Řadu z nich totiž vyvolává realizace hlavně těchto technik, stejně jako kvalitativní přístupy například v psychologii, sociální psychologii, pedagogice. Tím ale rozhodně nechci říci, že jiných vědních oblastí, jiných technik a kvantitativních postupů se snad etické aspekty netýkají. Jde o problematiku dosti komplikovanou, takže nezbývá, než se omezit jen na několik základních poznámek.

Etické aspekty zkoumání nejsou v metodologických publikacích, zvláště v těch staršího data, běžným tématem. Zmiňují je ale např. Disman (1993), Ferjenčík (2000), Švaříček, Šeďová aj. (2007), Jandourek (2008) a řada dalších. Poněkud podrobněji se jim kromě jiných věnují též Janoušek aj. (1986, s. 226–230), Jeřábek (1992, s. 144–147), Hendl (2005, s. 155–157) a Miovský (2006, s. 275–291), z nichž částečně vycházím a zájemcům o tuto oblast doporučuji.

I když se můžeme setkat s rozdílnými koncepty a úhly pohledu, etické otázky lze zhruba rozdělit do **dvou základních okruhů**. První zahrnuje etiku výzkumné práce jako jedné z podob práce vědecké, druhý etické zásady práce s účastníky výzkumu, tzn. se zkoumanými osobami. To je závažné a citlivé specifikum právě výzkumů sociálních.

Etiku výzkumné práce vymezuje několik pravidel, mezi které patří:

- I přes zaujetí řešeným tématem **výzkumník nesmí ztrácet objektivní a citlivý náhled**, neboť musí rozpoznat a přijmout též takové informace, které nepodporují jeho původní představy (či jim přímo odporují). Jakékoli upravování údajů je nepřípustné a je velmi vážným selháním výzkumníka jako vědce i jako osobnosti.