

Úvod do sociologického výzkumu

Hynek Jeřábek

Praha, Carolinum 1992

Stručný obsah:

Úvodem.....	7
1. Věda, vědecký výzkum	8
2. Problémy a hypotézy.....	11
3. Pojmy a konstrukty.....	20
4. Znaky a proměnné.....	25
5. Populace, výběry, reprezentativita.....	34
6. Uspořádání výzkumu. Typy výzkumných projektů.....	46
7. Terénní sběr dat, jeho fáze a techniky.....	53
8. Metody pozorování.....	58
9. Metody dotazování.....	65
10. Studium dokumentů (Sekundární analýza dat).....	73
11. Dotazování s využitím počítače.....	79
12. Kódování a příprava dat pro analýzu.....	83
13. Formulace úloh pro statistickou analýzu dat.....	100
14. Prezentace výsledků sociologického výzkumu.....	109
15. Etické a profesionální zásady sociologa výzkumníka.....	121
16. Proces rozhodování v sociologickém výzkumu.....	125
Literatura:.....	131
Otázky:.....	134
Osnova.....	138
Poznámky.....	140

Podrobný obsah

Úvodem.....	7	7.1. Sběr dat v procesu sociologického výzkumu.....	53
1. Věda, vědecký výzkum.....	8	7.2. Důležité fáze sběru dat.....	53
1.1. Základní vymezení vědy.....	8	7.2.1. Konspektování literatury a konzultace s odborníky.....	53
1.2. Čtyři metody poznání.....	9	7.2.2. Získání vstupu do sociálního systému.....	53
1.3. Definice vědeckého výzkumu.....	10	7.2.3. Seznámení s respondenty a prostředím.....	54
1.4. Postup výzkumníka.....	10	7.2.4. Předvýzkum nebo pilotáž.....	54
2. Problémy a hypotézy.....	11	7.2.5. Formulace otázek, stanovení pozorovaných elementárních jevů.....	55
2.1. Od výzkumného námětu k výzkumnému problému.....	11	7.2.6. Záznam a kódování informace.....	56
2.2. Výběr a formulace výzkumného problému.....	12	7.2.7. Kontrola záznamu, testování validity a spolehlivosti.....	56
2.3. Kritéria problémů a vytýčení problémů.....	13	7.2.8. Organizace a překódování dat.....	56
2.4. Hypotézy.....	14	7.3. Hlavní techniky sběru dat.....	56
2.5. Důležitost formulace hypotéz.....	16	8. Metody pozorování.....	58
2.6. Možnosti problémů a hypotéz.....	16	8.1. Charakteristika pozorování.....	58
2.7. Příklad formulace problému a hypotéz.....	17	8.2. Druhy pozorování.....	58
3. Pojmy a konstrukty.....	20	8.2.1. Nezúčastněné a zúčastněné pozorování.....	58
3.1. Teorie a empirie v sociologickém výzkumu.....	20	8.2.2. Standardizované a nestandardizované pozorování.....	58
3.2. Pojmy a konstrukty.....	21	8.2.3. Skryté a zjevné pozorování.....	59
3.2.1. Definice pojmu.....	21	8.3. Obor použití pozorování.....	59
3.2.2. Příklady sociologických pojmů.....	21	8.3.1. Charakter zkoumaných jevů.....	59
3.2.3. Vytváření pojmů.....	21	8.3.2. Charakter výzkumné akce.....	60
3.3. Definice.....	22	8.4. Příprava a zásady standardizovaného pozorování.....	61
3.4. Operacionalizace.....	22	8.4.1. Plán pozorování.....	61
4. Znaky a proměnné.....	25	8.5. Výhody a nevýhody zúčastněného pozorování.....	63
4.1. Definice znaku - proměnné.....	25	8.5.1. Problémy a nevýhody zúčastněného pozorování.....	63
4.2. Příklady znaků - proměnných.....	25	8.5.2. Výhody zúčastněného pozorování.....	63
4.3. Klasifikace znaků - proměnných.....	26	9. Metody dotazování.....	65
4.3.1. Klasifikace z hlediska vztahů, o nichž znak vypovídá (<i>Formální klasifikace</i>).....	27	9.1. Rozhovor.....	65
4.3.2. Klasifikace z hlediska místa ve výzkumném projektu.....	32	9.1.1. Standardizovaný rozhovor s dotazníkovým archem.....	65
5. Populace, výběry, reprezentativita.....	34	9.1.2. Rozhovor (částečně standardizovaný) se soupisem zjišťovaných faktů.....	65
5.1. Smysl výběrového šetření. Populace, cílový, základní a výběrový soubor.....	34	9.1.3. Nestruturovaný rozhovor.....	66
5.2. Reprezentativita výběrového souboru.....	36	9.1.4. Pravidla rozhovoru.....	67
5.3. Chyba výběru, vychýlení a přesnost, standardní chyba výběru.....	37	9.2. Dotazník jako nástroj sběru dat.....	68
5.3.1. Standardní chyba výběru.....	39	9.2.1. Příprava dotazníku.....	68
5.4. Pravděpodobnostní výběry.....	39	9.2.2. Pravidla uspořádání dotazníku.....	69
5.4.1. Náhodný výběr a jeho výhody.....	39	9.2.3. Formulace otázek.....	69
5.4.2. Typy pravděpodobnostního výběru.....	40	9.3. Dotazník zasílaný poštou (Mailed Questionnaire)(Postal Questionnaire).....	70
5.5. Ostatní typy výběrů - nepravděpodobnostní.....	42	9.3.1. Výhody [Moser-Kalton 1971:257-9].....	70
5.5.1. Kvótní výběr (Srovnaj [Moser-Kalton 1971:127-137]).....	42	9.3.2. Nevýhody (Omezení) [Moser-Kalton 1971:260-261].....	71
5.5.2. Systematický výběr.....	44	9.3.3. Fáze dotazování poštou (podle E.Kaneové [1985:86-87]).....	71
5.5.3. Výběr úsudkem.....	44	9.3.4. Průvodní dopis k dotazníku.....	71
5.5.4. Výběr nabalováním.....	45	9.3.5. Návratnost při poštovním dotazování.....	72
6. Uspořádání výzkumu. Typy výzkumných projektů.....	46	9.3.6. Faktory ovlivňující návratnost.....	72
6.1. Členění výzkumných projektů.....	46	10. Studium dokumentů (Sekundární analýza dat).....	73
6.2. Cíle výzkumného šetření.....	46	10.1. Typy dokumentů.....	73
6.3. Typy výzkumných projektů.....	48	10.2. Operace s dokumenty.....	73
6.3.1. Neexperimentální výzkumné projekty, mezi něž řadíme případovou studii, korelační výzkum, srovnávací výzkum a longitudinální výzkum:.....	48	10.3. Obory použití studia dokumentů.....	73
6.3.2. Jednoduché experimentální výzkumné projekty, kam řadíme kvaziexperiment, ex-ante - ex-post experiment a experiment s kontrolní skupinou:.....	48	10.3.1. Přednosti studia dokumentů.....	74
6.3.3. Faktoriální experimentální výzkumné projekty, jejichž smyslem je zjišťovat samostatný nebo společný vliv několika proměnných (faktorů).....	49	10.3.2. Nevýhody studia dokumentů.....	74
6.4. Kritéria hodnocení experimentálních výzkumných projektů (Srovnaj [Labovitz-Hagedorn 1976: 56-60]).....	49	10.4. Sekundární analýza.....	75
6.5. Kontrola vlivu vnějších proměnných.....	50	10.4.1. Zdroje dat pro sekundární analýzu.....	75
6.6. Vnitřní a vnější validita.....	52	10.4.2. Výhody sekundární analýzy.....	75
7. Terénní sběr dat, jeho fáze a techniky.....	53	10.4.3. Nevýhody sekundární analýzy.....	75
		10.5. Analýza osobních dokumentů.....	75
		10.5.1. Osobní dokumenty jako zdroj dat.....	76
		10.5.2. Klasické studie osobních dokumentů.....	76
		10.6. Obsahová analýza.....	77
		10.6.1. Postup obsahové analýzy.....	77
		11. Dotazování s využitím počítače.....	79
		11.1. CADI - pořizování dat s využitím počítače.....	79
		11.2. CATI - telefonní počítačové dotazování.....	80

11.3. CAPI - osobní počítačové dotazování.....	81	14.3.14. Mnohojednotkový graf - multiple unit chart	119
11.4. Budoucnost - CAMI - dotazování elektronickou poštou.....	82	14.3.15. Trojrozměrný graf.....	119
12. Kódování a příprava dat pro analýzu.....	83	14.3.16. Korelační graf.....	119
12.1. Matematizující a spekulativní pojetí sociologie.....	83	14.3.17. Korespondenční graf, graf podobnosti, vzdálenosti.....	119
12.2. Počítače v sociologickém výzkumu při přípravě a analýze dat.....	83	15. Etické a profesionální zásady sociologa výzkumníka.....	121
12.2.1. Co počítače sociologii přinesly?.....	84	15.1. Co je a co není "etické", "morální"?.....	121
12.2.2. Jak počítače ovlivňují způsob myšlení sociologa výzkumníka?.....	84	15.2. Jak je možno zachovat mravní normy?.....	121
12.3. Kódování.....	85	15.3. Tři příklady projektů, které se prohrály proti etice sociologické výzkumné práce (podle [Burges 1984:185-189]).	123
12.3.1. Příprava kódovacího klíče.....	85	15.4. Základní body etického kodexu Americké sociologické asociace (podle [Bailey 1978: 447nn] Appendix, Part B: Code of Ethics of the American Sociological Association).....	123
12.4. Příprava (pořizování a čišťení) dat.....	85	15.5. Podmínky publikace výsledků výzkumu.....	124
12.4.1. Obecné zásady přípravy dat.....	86	16. Proces rozhodování v sociologickém výzkumu.....	125
12.4.2. Fáze přípravy dat.....	87	16.1. Výzkumný problém.....	125
12.5. Data vstupují do analýzy.....	91	16.2. Pojmový rámec.....	125
12.5.1. Možnosti uspořádání dat. Typy vstupních dat.....	91	16.3. Populace a výběrový soubor.....	126
12.5.2. Práce s vynechanými údaji.....	96	16.4. Výzkumný projekt - uspořádání výzkumu.....	126
12.5.3. Modifikace a transformace dat, konstrukce nových znaků.....	97	16.5. Metody sběru dat.....	127
13. Formulace úloh pro statistickou analýzu dat.....	100	16.6. Příprava a čišťení dat.....	128
13.1. Úloha deskripce.....	100	16.7. Analýza a interpretace.....	128
13.1.1. Četnosti kategorizovaných znaků.....	100	16.8. Prezentace a závěry.....	129
13.1.2. Statistické charakteristiky spojených znaků.....	101	Literatura:.....	131
13.2. Úloha komparace.....	103	Otázky:.....	134
13.2.1. Porovnání souhrnných charakteristik.....	103	Osnova.....	138
13.2.2. Porovnání statistických rozložení.....	104	Poznámky.....	140
13.3. Měření asociací a korelační analýza.....	104		
13.3.1. Lineární a nelineární vztahy.....	104		
13.3.2. Tabulka korelací a její typy.....	105		
13.3.3. Explorační analýza závislostí.....	105		
13.4. Další metody a možnosti mnohorozměrné analýzy dat.....	106		
13.4.1. Regresní analýza.....	106		
13.4.2. Faktorová analýza (explorační a konfirmační).....	106		
13.4.3. Seskupovací analýza (cluster analysis).....	107		
13.4.4. Pěšinková analýza (Path Analysis).....	107		
13.4.5. Strukturní modelování.....	107		
13.4.6. Mnohorozměrné škálování.....	108		
13.4.7. Logaritmicko-lineární modelování.....	108		
14. Prezentace výsledků sociologického výzkumu.....	109		
14.1. Zásady prezentace dat z výzkumu.....	109		
14.1.1. Okruh čtenářů.....	109		
14.1.2. Účel publikace.....	109		
14.1.3. Forma sdělení a způsob publikace.....	110		
14.1.4. Styl.....	110		
14.1.5. Obsah výzkumné publikace.....	111		
14.1.6. Mravní odpovědnost.....	111		
14.2. Struktura výzkumné publikace.....	111		
14.2.1. Titul.....	112		
14.2.2. Obsah.....	112		
14.2.3. Text.....	113		
14.2.4. Bibliografie.....	114		
14.3. Grafická prezentace, grafy, schémata.....	115		
14.3.1. Vývojový čárový graf - line chart.....	116		
14.3.3. Srovnávací sloupkový graf - srovnávací histogram - comparative bar chart.....	116		
14.3.5. Srovnávací složený graf - comparative composite chart.....	117		
14.3.6. Rozvinutý sloupkový graf - extended bar chart.....	117		
14.3.7. Kruhový velikostní graf - kruhový diagram.....	117		
14.3.8. Kruhový dělený graf (koláčový graf) - pie chart.....	117		
14.3.9. Graf rozptýlení (krabičkový graf) - box chart.....	117		
14.3.10. Stromový graf - dendrogram.....	118		
14.3.11. Šipkový graf - causal diagram and flowgraph.....	118		
14.3.12. Znaménkové schéma.....	118		
14.3.13. Blokový diagram - block diagram.....	119		

Zájem o sociální data týkající se našeho každodenního života, počtu nezaměstnaných, výše životních nákladů rodiny, prosperity státu, blahobytu nebo bídy jeho obyvatel nekončí pro mnohé občany přečtením novin nebo poslechem politického zpravodajství. Klademe si otázky nad výsledky sociologických výzkumů. Mnohdy uvažujeme o tvrzeních, která jsou nám autoritativně předkládána jako výsledky průzkumů veřejného mínění. Snažíme se lépe rozumět komentářům k číslům, tabulkám a grafům. Chtěli bychom si ověřit, že je to skutečně tak a že nás redaktor nebo výzkumník nešidí.

Někteří dokonce cítíme potřebu sami provést výzkumné šetření, vytvořit vlastní dotazník, pozorovat chování lidí a vysvětlit, co jsme zjistili. Potřebujeme však také porozumět výsledkům, ke kterým došli ve výzkumu jiní.

Co lze očekávat od příručky, kterou jste právě otevřeli? Základní informace o tom, jak se provádí sociologický výzkum. Jak postupovat od začátku do konce výzkumné akce. Knížku lze přirovnat k automapě po jedné ze zemí sociologického bádání, po průzkumech prováděných klasickými technikami. Existuje však mnohé v sociologickém výzkumu, co se nevejde do stručného úvodu. Výzkumník, který se vydává mimo hlavní cesty, používá méně běžných metod, nutně potřebuje také speciální mapy, podrobnější příručky a učebnice. Seznam literatury na některé z nich odkazuje.

Kromě klasického sociálního průzkumu existují další výzkumná odvětví sociologie, výzkumná teritoria vyžadující mapy, na nichž by byly zakresleny jiné splavné toky a úrodné nížiny, existuje kvalitativní výzkum, přístup naturistický, paradigma interpretativní sociologie. Uvádím proto přehled základních učebnic kvalitativního výzkumu v poznámkách.

1. Věda, vědecký výzkum

1.1. Základní vymezení vědy

Věda je složitý myšlenkový systém, v němž můžeme rozlišit:

teorii = je pokusem o vysvětlení příčin stavu a příčin změn pozorovaného jevu;

metodologii = je soustavou pravidel stanovujících postupy, jimiž lze k vysvětlení dojít anebo je ověřit;

poznatky = jsou sumou zjištění, která byla učiněna.

PŘ.: Teorie - vysvětluje pravidelné střídání fází Měsíce
 Metodologie - soubor postupů, jímž lze teorii ověřit;
 v našem případě jsou to astronomická pozorování nebo fyzikální experiment
 Poznatek - skutečnost, že Měsíc obíhá kolem Země, je osvětlován slunečním svitem a toto světlo odráží k Zemi.

Můžeme rozlišit dvě pojetí vědy:

a) statické, uvažující vědu jako soubor faktů, jako současný stav vědění a jeho doplňování,

b) dynamické, považující vědu za činnost, kterou dělají vědci, pojetí heuristické (poznávací)

Zrod moderní vědy je datován počátkem novověku
 = koncem 16.stol./ počátkem 17.století

Významní	Francis BACON	(1561-1626)
filozofové	René DESCARTES	(1596-1650)
	John LOCKE	(1632-1704)
	David HUME	(1711-1776)

Významní	Mikuláš KOPERNÍK	(1473-1543)
přírodovědci	Jan KEPLER	(1571-1630)
	Galileo GALILEI	(1564-1642)
	Isaac NEWTON	(1642-1727)

bodů mimo nás ... metoda musí být taková, aby konečný závěr každého člověka byl stejný. Taková je metoda vědy...".¹⁾

Teprve v polovině 19. století byla moderní věda schopna se sjednotit na základních požadavcích své metodologie:

Základní postuláty metodologie vědy

(Ve formulaci Claude Bernarda 1815-1878)

(srovnej podrobněji viz. Možný-Rabušic 1989: 6-10)

1. Fakta:

Vědec se musí bez výhrad podrobit tlaku faktů a nesmí nikdy váhat obětovat jakoukoli teorii, jestliže fakta svým dosahem ruší její hlavní smysl.

2. Determinismus:

Účinná vědecká zkoumání musejí vycházet z předpokladu přísného determinismu jevů, z odhalování vztahů mezi příčinami, podmínkami a jevy.

3. Věda filozoficky neutrální:

Věda je absolutně neutrální v poměru k filozofickým otázkám.

4. Pokus falzifikovat:

Žádná vědecká myšlenka nemůže být dost pevná, jestliže vědec zná toliko fakta, která potvrzují jeho hypotézu, leč nevěnoval-li stejně (lépe však více) úsilí na objevení faktů, které by ji vyvracely (tj. nevystavil-li ji falzifikujícím pokusům)

1.2. Čtyři metody poznání

Charles Peirce (americký filosof) rozlišuje čtyři obecné způsoby poznání (způsoby fixování přesvědčení):
(srovnej viz. [Kerlinger 1972:21nn.])

- 1) Metoda tradice: Lidé se drží pravdy, protože ji vždy znají jako pravdu
- 2) Metoda autority: Je to metoda pevně stanoveného názoru
PŘ.: Jestliže to říká bible, je tomu tak.
- 3) Metoda a priori: Tvrzení akceptovaná "aprioristy" "jsou ve shodě s rozumem" (intuice)
- 4) Metoda vědy:
Ch.Peirce:"...je nutné, abychom našli metodu, pomocí níž lze určit naše názory ne něčím lidským, ale pomocí nějakého pevného

1.3. Definice vědeckého výzkumu

"vědecký výzkum je systematické, kontrolované, empirické a kritické zkoumání hypotetických výroků o předpokládaných vztazích mezi přirozenými jevy." [Kerlinger 1972: 27]

systematický a kontrolovaný - znamená kritickou důvěru ve výzkumné výsledky

empirické vědecké bádání - znamená, že vědec musí podrobit svůj názor vnějšímu ověření

1.4. Postup výzkumníka²⁾

- 1) prvním krokem je dostat nápad
Dewey:"...existuje nesnáze, zmatek, obtížná situace, pochybnost, emoční napětí..."
- 2) ve druhém kroku je formulován problém, situace je pojmově ztvárňována s využitím studia literatury
- 3) ve třetím kroku je formulována hypotéza
hypotéza = konstatování opírající se o domněnky, prozatímní (zkušební) výrok o vztazích mezi dvěma nebo více pozorovanými jevy nebo proměnnými (někdy nepozorovanými) (jestliže se vyskytne to a to, pak to vede k těm a těm výsledkům)
- 4) čtvrtým krokem je usuzování - dedukce; vědec dedukuje důsledky z hypotézy, kterou formuloval
- 5) pátým krokem je ověřování, testování důsledků, pozorování, experiment; podstatou testování je ověřování vztahů vyjádřených v hypotéze, hypotézu neověřujeme přímo - ověřujeme dedukované implikace z hypotézy

2. Problémy a hypotézy

2.1. Od výzkumného námětu k výzkumnému problému

Výzkumný námět vzniká jako nápad, obtížná situace. Nejprve výzkumník pociťuje pochybnost, nesnáze, objeví se nápad. Připomeňme si přístup J.Deweye v práci k analýze reflexivního myšlení: říká, že vzniká emoční napětí, existuje nesnáze, zmatek. Je třeba vyhledat literaturu pojednávající o dané otázce, prostudovat výsledky dosud provedených výzkumů. Vhodné je vést diskuse s odborníky na dané téma, rozhovory s dalšími lidmi, kteří znají prostředí a problémy zkoumané oblasti. Mnohou inspirací je možno najít v knihách, nejen v odborné literatuře, ale také v krásné literatuře. Řadou aktuálních otázek se zabývá publicistika ve zprávách a komentářích v denním tisku i v rozsáhlejších článcích v kulturně politických týdenících a měsícnících.

Existuje mnoho zdrojů výzkumných námětů:

- individuální zkušenost
- odborné publikace (knihy a časopisy)
- krásná literatura
- publicistika
- osobní rozhovory
- výsledky předchozích výzkumů
- sdílené hodnoty
- teorie (výzkumný problém je odvozován ze souhrnu tvrzení této teorie)

Teoreticky nejzávažnější ideje jsou odvozovány z výsledků předchozích výzkumů a z diskusí s kolegy zabývajícími se příbuznými otázkami.

PŘ.: Výzkumný problém odvozený ze sdílených hodnot výzkumníka. Když si někdo myslí, že demokratický systém je nejlepší, může ověřovat hypotézu o tom, že demokraticky strukturované malé skupiny budou mít největší úspěch při řešení úloh, nebo že budou mít nejvyšší stupeň spolupráce, soudržnosti nebo vzájemné důvěry. (Srovnej viz. [Labovitz-Hagedorn 1976:24n].)

Správně projektovaný výzkum poskytne výsledky, které nebudou ovlivněny výzkumníkovými hodnotami, i když jeho hodnoty ovlivnily výběr výzkumného problému.

Předchozí výzkumné výsledky nebo statistické závislosti často vedou k dalšímu výzkumu, který má již teoretické cíle, teoretické ambice.

Např.: Emile Durkheim si všiml, že v Dánsku (většinově protestantském) je větší podíl sebevražd než např. v Itálii (která je převážně katolická). Z těchto pozorování Durkheim neodvodil prosté tvrzení, že protestanti mají větší podíl sebevražd než katolíci, nýbrž pokračoval v úvaze až k formulaci teoretického tvrzení - hypotézy o vlivu sociální integrace nebo skupinové solidarity na sebevražednost.

Jedno ze základních tvrzení Durkheimovy teorie zní, že:

"podíl sebevražd je nepřímě úměrný ke stupni sociální integrace";

³⁾

2.2. Výběr a formulace výzkumného problému

Vědec na základě studia literatury, osobními rozhovory, pozorováním, studiem výsledků předchozích výzkumů zpřesňuje výzkumný námět a formuluje výzkumný problém.

Co je to problém: "Problém je tázací věta nebo výrok, který se ptá: Jaký vztah existuje mezi dvěma nebo více proměnnými?" [Kerlinger 1972:32]

Problém: je výrok formulovaný ve formě otázky

Jaký volit postup formulace výzkumného problému?

1. seznámíme se s dosud dosaženými výsledky týkajícími se vymezené oblasti,
2. objasníme si rozměry zkoumané problematiky,
3. hledáme doporučení dosud použitých postupů řešení,
4. snažíme se vyvarovat chyb dosavadních postupů,
5. pokusíme se o samostatné řešení postavených otázek,
6. pokusíme se o formulaci nových otázek.

Překvapivá zjištění nebo nečekané asociace (souvislosti) mohou vést k užitečným výzkumným myšlenkám.

PŘ.: Robert K.Merton uvádí příklad situace, v níž zjišťoval ve dvou odlišných sídelních útvech počet dívek, které ohlírají rodičům dítě (baby sitter). Podle objektivních statistických údajů bylo ve velkém městě mnohem více dívek vhodných k takové službě mladým rodičům než v malé příměstské komunitě. Přesto podle názoru rodičů byla situace opačná: V malé příměstské komunitě Craftownu odhadovali rodiče relativně větší výběr dívek než ve velkoměstě. Posouzení situace tedy záviselo nikoli jen na pouhém počtu, ale do značné míry na důvěře. Merton uzavírá: *...vnímání závisí na struktuře mezilidských vztahů...*⁴¹

Neočekávané výsledky mohou vést k užitečným odhalením.

PŘ.: Samuel A.Stouffer - v často citovaném klasickém výzkumu Americký voják ⁵¹ z období druhé světové války pozoroval malé rozdíly ve spokojenosti s životem v armádě v USA a na bojišti v Evropě, přitom však se ukázaly jako významné zcela jiné, *neočekávané rozdíly: vojáci v jednotkách s vyšším podílem vyšších šarží byli méně spokojeni, (měli pocit, že mají menší šanci k povýšení než v jednotkách s nižším podílem vyšších důstojníků); S.Stouffer zavádí vysvětlující pojem **relativní deprivace**: vojáci v jednotkách s vysokým podílem vyšších šarží se cítili deprivováni relativně k těm, jejichž jednotky měly méně vysokých hodností - měli pocit, že nemohou tak snadno dosáhnout povýšení.*

Výzkumník má představu, co chce zkoumat. Měl by

- seznámit se s dosavadním poznáním a se současným stavem výzkumu v dané oblasti,
- diskutovat výzkumnou ideu s kolegy,
- prostudovat literaturu, která o předmětu byla napsána.

Prospěch z tohoto postupu pro něj vyplyne z toho, že:

- 1) může se vyvarovat chyb a slepých uliček (ušetří čas a půjde efektivněji k cíli);
- 2) objasní si základní rozměry zkoumané problematiky (na některé by mohl zapomenout);
- 3) může najít doporučení různých způsobů, jak mohou být měřeny, pozorovány, zkoumány dané otázky.

2.3. Kritéria problémů a vytýčení problémů

- 1) Problém by měl vyjadřovat vztah mezi dvěma nebo více proměnnými (znaky)

proměnná (znak)= symbolické vyjádření vlastnosti, které svými hodnotami vyjadřuje stavy vlastnosti a vztahy mezi hodnotami vyjadřuje vztahy v realitě (srovnej [Řehák 1972:617])

- 2) Problém by měl být formulován jasně a jednoznačně v tázací formě; v praxi se s označením "problém" setkáváme nesprávně také v situacích, kdy se jedná pouze o vymezení tématu zkoumání
- 3) Tvrzení, která vyplývají z problému, by měla být ověřitelná objevují se také nevědecké otázky, které nevypovídají o vztazích, metafyzické otázky, jejichž pojmy a konstrukty je obtížné definovat tak, aby je bylo možno měřit; tvrzení nejsou tedy vědecky ověřitelná.

2.4. Hypotézy

Výzkumník na základě problému nebo problémů formuluje hypotézy.

Hypotéza: **je podmíněný výrok o vztahu mezi dvěma nebo více proměnnými**

Kritéria dobrých hypotéz:

- 1) hypotézy jsou výroky o vztazích mezi proměnnými
- 2) hypotézy obsahují proměnné, které lze zjišťovat a měřit
- 3) vztahy mezi proměnnými lze ověřovat.

PŘ.: Pedagogicko-psychologický výzkum

H: Studium ve skupině přispívá ke zvýšení výkonu.

proměnné: studium ve skupině (zjistitelné, měřitelné)
zvýšení výkonu (měřitelné)

Diskuse kritérií:

- | | |
|------------------------------------|-----|
| 1) Existuje vztah mezi proměnnými? | ANO |
| 2) Lze proměnné zjišťovat, měřit? | ANO |
| 3) Lze vztah ověřit? | ANO |

PŘ.: Výzkum čtenářů časopisu Mladý svět

H: Čtenářský zájem o časopis Mladý svět s věkem klesá.

proměnné: čtenářský zájem
věk

Diskuse kritérií:

- | | |
|--|-----|
| 1) Existuje vztah mezi proměnnými věk a čtenářský zájem? | ANO |
| 2) Lze proměnné zjišťovat, měřit? | |
| věk | ANO |
| čtenářský zájem - počet předplatitelů | ANO |
| - dotazem na počet čtenářů ve věkové skupině | ANO |
| - dotazem, jakou část MS čtou čtenáři ve věkové skupině | ANO |
| 3) Lze vztah ověřit? - tabulkou závislosti | ANO |

PŘ.: Výzkum pracovníků státního statku

H: Dělníci si vydělávají víc než technicko-hospodářští pracovníci

proměnné: příjem ("vydělávají si")
kategorie pracovníků

Diskuse kritérií:

- | | |
|---|-----|
| 1) Existuje vztah mezi proměnnými kategorie pracovníků a příjem | ANO |
| 2) Lze proměnné zjišťovat, měřit? | |
| kategorie pracovníků podle skutečného zařazení | ANO |
| příjem - podle výplatní pásky | ANO |
| - dotazem | ANO |
| - za skupiny podle evidence | ANO |
| 3) Lze vztah ověřit? - tabulkou průměrů platů | ANO |

PŘ.: Výzkum spokojenosti

H: Spokojenost lidí závisí na jejich životním prostředí.

NEVHODNÁ - nutno přeformulovat

H': Lidé žijící v lepším životním prostředí jsou spokojenější.

H": Lidé žijící na venkově se častěji usmívají než lidé ve městě.

proměnné: H' spokojenost
kvalita životního prostředí
H" četnost úsměvů
typ bydliště

Diskuse kritérií:

	H	H'	H''
1) Existuje vztah mezi proměnnými	URČEN	ANO	ANO
2) Lze proměnné zjišťovat, měřit?	NEPŘESNĚ	NEPŘÍMO	NEPŘÍMO POZOROVAT
3) Lze vztah ověřit?	NE	NEPŘÍMO	POZOROVÁNÍM

2.5. Důležitost formulace hypotéz

Proč jsou hypotézy tak důležité? (Srovnej viz. [Kerlinger 1972:34-35])

- Hypotézy jsou pracovní nástroje teorie
 - jsou odvozovány z teorie
 - potvrzené nebo nepotvrzené obohacují a rozvíjejí teorii
- Hypotézy mohou být ověřovány
 - izolovaná fakta neověřujeme
 - hypotézy jsou tvrzení o vztazích a vztahy ověřujeme

==> hypotézy ověřujeme

 - je to jeden z hlavních úkolů výzkumu
 - je to základní způsob vědecké práce.
- Hypotézy umožňují člověku vyjít z jeho subjektivity
 - hypotézy jsou vytvořeny člověkem
 - hypotézy existují a je možno prověřit jejich pravděpodobnou pravdivost nebo nepravdivost bez ohledu na hodnoty a mínění člověka

2.6. Možnosti problémů a hypotéz

(Srovnej [Kerlinger 1972:35-37])

- Problémy a hypotézy usměrňují bádání.*
Vztahy vyjádřené v hypotézách říkají ve svých důsledcích výzkumníkovi, co má dělat
- Problémy a hypotézy umožňují odvozovat z hypotéz dílčí tvrzení (odvozená, implikovaná)/ důsledky, které jsou již ověřitelné.*
Jejich ověřením je původní hypotéza podpořena.
- Problém nemůže být vědecky řešen, pokud jej neredukujeme na formu hypotézy.*

Dobře postavená hypotéza
 buď 1. může být ověřována
 nebo 2. z ní můžeme odvodit jiné ověřitelné hypotézy.
 V obou případech ověřujeme vztahy.

- 4) *Problémy a hypotézy rozvíjejí vědecké poznání tím, že pomáhají badateli podpořit nebo oslabit teorii.*
 Je-li formulována předem hypotéza, (která má vztah k teorii), jejím potvrzením nebo odmítnutím je podpořena nebo oslabena nebo v krajním případě zamítnuta teorie. Jestliže byl zjištěn vztah mezi proměnnými nahodile, bez předem stanovené hypotézy, může být náhodný nebo dokonce falešný.
- 5) *Badatel, který nevysloví hypotézu o vztahu předem, neumožní faktům cokoli dokázat nebo vyvrátit.*
 Užítí hypotézy se podobá hraní sázkové hry.
 Hypotéza existuje = Pravidla jsou stanovena předem.
 Sádky jsou dány předem.

 Ve výzkumu = Je předem stanoveno, které vztahy zkoumáme a k jakému se přikloníme závěru, když nabudou určité síly nebo hodnoty. Když vybíráme z předem neznámé množiny vztahů jen některý, který právě vyšel významný, nehrajeme čestně podle pravidel.

2.7. Příklad formulace problému a hypotéz

PŘÍKLAD: Mají studenti zájem o metody sociologického výzkumu?

- 1) Nesnáž, obtížná situace, emoční napětí profesora způsobené tím, že studenti nepřišli na přednášku:

Studenti nepřišli na přednášku	----->	Studenti nemají zájem o metody sociologického výzkumu
--------------------------------	--------	---

- 2) Pochybnost, nápad:

- Čím vším mohlo být způsobeno to, že studenti nepřišli na přednášku?
- Čím se projevuje zájem studentů o metody sociologického výzkumu?

- 3) Logický rozbor situace:

Studenti mají -----> a) studenti navštěvují přednášky zájem o metody

sociologického -----> b) studenti samostatně studují výzkumu literaturu o metodách sociologického výzkumu

-----> c) studenti se podílejí na praktickém výzkumu

-----> d) studenti se pečlivě připravují na semináře

-----> e) studenti aktivně vystupují na seminářích

- 4) Formulace problému: Jaké existují vztahy mezi zájmem studenta o metody sociologického výzkumu a jeho aktivitami při studiu metod sociologického výzkumu?

Model 1	ZÁJEM	STUDIJNÍ	--X sledování
		AKTIVITA	--X studijních
	test zájmu		--X aktivit
			--X

H: Vysoký zájem o metody se objevuje současně s vysokou studijní aktivitou: (zahrnuje v sobě hypotézy H₁, H₂ a H₃).

H₁: Kdo má vyšší zájem, má v důsledku toho vyšší studijní aktivitu.

H₂: Kdo má vyšší studijní aktivitu, má v důsledku toho i vyšší zájem.

H₃: Vysoký zájem a vysoká studijní aktivita jsou důsledkem společné příčiny.

Model 2

V průběhu studia se prohlubuje vztah mezi zájmem o studium a studijní aktivitou. Růst zájmu souvisí s růstem studijní aktivity.

ZÁJEM	----->	STUDIJNÍ AKTIVITA
	<-----	
test zájmu v čase T1	test zájmu v čase T2	sledování studijní aktivity v časové řadě

H₄: Vysoký zájem vede ke zvyšování studijní aktivity.

H₅: Vysoká studijní aktivita vede ke zvyšování zájmu.

ZÁJEM
testy zájmu
v čase

STUDIJNÍ AKTIVITA
sledování studijních
aktivit v časové řadě

ZNALOSTI
testy znalostí
v čase

H₀: Vyšší zájem má za následek vyšší studijní aktivitu.

H₁: Vyšší studijní aktivita má za následek vyšší znalosti.

H₂: Vyšší znalosti mají za následek vyšší zájem.

Záleží na výzkumníkovi, který model si zvolí, které hypotézy bude ověřovat. Tomu pak přizpůsobí výběr zjišťovaných skutečností, metody zkoumání i způsob zpracování úlohy.

3. Pojmy a konstrukty

3.1. Teorie a empirie v sociologickém výzkumu

Vědci pracují na dvou úrovních poznání:

- na úrovni pojmů = t.j. v oblasti TEORIE a
- na úrovni pozorování = t.j. v oblasti EMPIRIE

Při podrobnějším rozlišení lze v oblasti teorie dále rozlišit rovinu teorií, zákonů a zákonitostí a rovinu problémů, hypotéz a modelů.

K postupu poznání je třeba vědecké práce v obou těchto úrovních:

- v pozorování, zjišťování, měření, rozboru faktů, t.j. ve sféře EMPIRIE a
- v konstrukci teorií, hypotéz, ve formulaci pojmů a jejich operacionalizaci, t.j. ve sféře TEORIE.

Albert Einstein v roce 1930 napsal:

"...poznání nemůže rozkvést z holé empirie. Takový rozkvět je možný pouze ze srovnání toho, co je vymyšleno, s tím, co je pozorováno".⁶⁾

M.Cohen v roce 1956 formuluje obdobnou myšlenku:

"...Bez nějaké vůdčí ideje nevíme, jaká fakta sbírat. Nemáme-li co dokazovat, nemůžeme určit, co je významné a co nevýznamné".⁷⁾

Václav Černík⁸⁾ v rovině filosofie a obecné metodologie rozeznává:

- 1) Empirickou poznávací činnost (pozorování, měření) **E**
- 2) Teoretickou poznávací činnost prvního druhu (formulace problémů, konstrukce modelů a ověřování hypotéz) **T₁**
- 3) Teoretickou poznávací činnost druhého druhu (výstavba teorií, formulace zákonů) **T₂**

Sociologický výzkum v převážné většině případů probíhá na úrovni formulace problémů, konstrukce modelů a ověřování hypotéz, t.j. na úrovni teoretické poznávací činnosti prvního druhu **T₁** a na úrovni pozorování, dotazování, zjišťování a měření, t.j. na úrovni empirické poznávací činnosti **E**. Oba tyto druhy výzkumných činností se střídají a většinou několikrát opakují dříve než je možno obohatit sociologickou teorií (zkoumání zákonitostí).

Základními stavebními kameny teorie jsou problémy, modely a hypotézy. K jejich výstavbě, formulaci, ověřování používáme v teoretické poznávací činnosti pojmy a konstrukty.

Empirická poznávací činnost se obrací k jevům, ke společenské skutečnosti. V průběhu pozorování, měření, přípravy k nim a jejich vyhodnocování jsou v sociologii užívány znaky - proměnné a jsou zjišťovány vztahy (relace) mezi nimi.

3.2. Pojmy a konstrukty

3.2.1. Definice pojmu

pojem : termín nebo symbolické vyjádření, které reprezentuje podobnosti v jinak rozdílných jevech.

V pojmech - vyjadřujeme výzkumné problémy
- formulujeme hypotézy
- vytváříme (formulujeme) teorie

3.2.2. Příklady sociologických pojmů

Sociologie je bohatá na pojmy, s nimiž pracuje (např. status, role, norma, hodnota, interakce, relativní deprivace). Pokud jsou měřeny, mohou být použity jako znak - proměnná.

Příklad pojmu: STATUS

hrubý index statusu = klasifikace prestiže povolání
vysoký status = vysoká prestiž povolání (např. lékař)
nízký status = nízká prestiž povolání (např. prostitutka)

podrobný index statusu = klasifikace podle:

příjmové kategorie
kategorie povolání
nejvyššího dosaženého vzdělání
společenské prestiže (úcty)

3.2.3. Vytváření pojmů

Vývoj pojmů vyžaduje procesy zobecňování (generalizace) a abstrakce.

zobecňování : je proces odvozování principu z množství zkušeností;
každý pojem a každá proměnná (sociologický znak)
nutně představují
abstrakci - t.j. obsahují pouze vybrané rysy jevu

3.3. Definice

Možnost porozumění jedné osoby druhé je závislá na zachování stejného významu užitých symbolů. Významy symbolů jsou určovány jejich definicí. Definice říká, co termín znamená.

Definice je výrok o úmyslu užit symbol určitým způsobem, v určitém významu.

Jedna z možných definic sebevraždy:
Sebevražda = zabýt se nebo nevyužít možnosti k záchraně, která je použitelná.

Definice vznikají dohodou (říkáme, že jsou arbitrární). Žádná definice není správná nebo špatná, pravdivá nebo nepravdivá.

PŘ.:

Říci, že láska je sexuální vztah a nic víc není nepravdivé nebo chybné tvrzení. ta slova mají význam sdělení, co jejich prostřednictvím chci stanovit jako význam použitého termínu. Definice, např. nalezené ve slovníku, vyjadřují konvenční významy slov.

3.4. Operacionalizace

Technicky rozeznáváme dva druhy (typy) definic: *nominální (konstitutivní) a operacionální definice.*

Nominální definice jsou definice, v nichž dané slovo je vyjádřeno synonymem - jiným výrazem, jehož význam je znám, byl stanoven, určen již dříve. Výhodou takové definice je, že usnadňuje komunikaci (dorozumívání).

Operacionální definice je dána stanovením kroků v pozorování nebo měření daného jevu.

K problematice operacionální definice se váže pojem *reliabilita* (česky *spolehlivost*). *Reliability* je dosaženo, když dva nebo několik

kvalifikovaných pozorovatelů dojde ke shodným výsledkům při měření
téhož jevu.

PŘ.: měření výšky postavy dané osoby:

Operacionální definice výšky postavy osoby = počet cm od hlavy
k patě měřená ve stoje zády ke svislé stěně.

PŘ.: test inteligence:

Operacionální definice inteligence = to, co měří test IQ.

4. Znaký a proměnné

4.1. Definice znaku - proměnné

Znak v sociologii = proměnná = symbolické vyjádření vlastnosti, které zároveň vyjadřuje stavy této vlastnosti svými hodnotami a vztahy mezi těmito stavy relacemi mezi svými hodnotami. Musí splňovat tyto podmínky:

- má alespoň 2 hodnoty (diskriminabilita - rozlišitelnost)
- ke každému stavu vlastnosti existuje hodnota znaku (úplnost)
- dvě různé hodnoty znaku nemohou odpovídat jednomu stavu vlastnosti (jednoznačnost).

Toto vymezení je definicí znaku v širším pojetí. (Srovnej [Řehák 1972:617]).

Znak = proměnná (úzké pojetí) = je měřitelná dimenze pojmu nebo měřitelný pojem mající dvě nebo více hodnot pro různé jednotky nebo pro různá období.

4.2. Příklady znaků - proměnných

Vlastnost	Znaky
Věk	Věk v letech 01 - 98 let; 99 = více let Pětileté věkové kategorie: 01 = 15-19 let 02 = 20-24 let 03 = 25-29 let 04 = 30-34 let 05 = 35-39 let 06 = 40-44 let 07 = 45-49 let 08 = 50-54 let 09 = 55-59 let 10 = 60-64 let 11 = 65-69 let 12 = 70-74 let 13 = 75-79 let

	14 = 80 a více
	88 = do 14 let
	99 = neví
	00 = neodpověděl
Tři věkové skupiny	1 = do 14 let
	2 = 15-59 Mu 15-54 Že
	3 = 60 + Mu 55 + Že

Profesní skupina	Jednotná klasifikace zaměstnání	jednomístný kód 0 - 9 dvoumístný kód 00 - 99 trojmístný kód 001 -999 + doplňkové kódy
	Mezinárodní klasifikace ISCO	kódy profesí a skupin profesí: jednomístný kód 0 - 9 dvoumístný kód 00 - 99 trojmístný kód 000 -999 čtyř a vícemístný kód
Sociální vrstva	Kategorie podle výše měsíčního příjmu na osobu	1 nedostatečný příjem (do 500,-Kčs na os.) 2 velmi nízký příjem (500 - 750,-Kčs na os.) 3 nízký příjem (751 -1000,-Kčs na os.) 4 střední příjem (1001 -1500,-Kčs na os.) 5 vysoký příjem (1500 -2000,-Kčs na os.) 6 velmi vysoký příjem (nad 2000,-Kčs na os.)
	Kategorie podle subjektivního zařazení	VV = vyšší-vyšší vrstva VS = vyšší-stř. vrstva NS = nižší-stř. vrstva NN = nižší-nižší vrstva

4.3. Klasifikace znaků - proměnných

4.3.1. Klasifikace z hlediska vztahů, o nichž znak vypovídá (Formální klasifikace)

Základní třídění:

znaky nominální (klasifikační)

PŘ.: pohlaví, profese, typ televizního pořadu, typ rodiny

znaky pořadové (ordinální)

PŘ.: školní klasifikace, spokojenost s prací, stupeň souhlasu

znaky intervalové

PŘ.: výsledek testu IQ, časový údaj

znaky poměrové

PŘ.: sledovanost televizních pořadů, počet dětí

Znaky nominální (klasifikační)

Pro tento typ znaku je charakteristické, že nemá uspořádané hodnoty. Pro nominální znak nemá smysl výrok "hodnota a je vyšší než hodnota b ". Nominální znaky rozlišují množinu navzájem různých hodnot, mezi nimiž jsou uvažovány pouze vztahy totožnosti a rozdílnosti. Jednotka může mít hodnotu znaku buď a nebo b nebo c , ale jen jednu z nich. Hodnoty znaku pro dvě různé jednotky X a Y mohou být buď stejné $x = a$, $y = a$ nebo různé $x = a$, $y = b$, jinak ale o jejich vztahu nic dalšího nevíme. Statistik by řekl, že nominální znak "generuje rozklad na třídy ekvivalence, z nichž každá obsahuje právě všechny prvky, které jsou k sobě navzájem vztahy popsanou relací". Do nich jsou zařazováni zkoumané objekty, např. respondenti.

PŘ.1: Místo bydliště

- A - Čechy
- B - Morava
- C - Slezsko
- D - Slovensko

PŘ.2: Pohlaví

- 1 - muž
- 2 - žena

PŘ.3: Odvětví

- 1 - průmysl
- 2 - zemědělství
- 3 - obchod
- 4 - služby
- 5 - státní správa
- 6 - peněžnictví a pojišťovnictví
- 7 - školství a věda
- 8 - kultura
- 9 - zdravotnictví
- 0 - ostatní

Znaky pořadové (ordinální)

Pořadové znaky rozlišují kromě totožnosti a rozdílnosti hodnot také jejich uspořádání. O každých dvou hodnotách

$a = b$ platí buď $a > b$ nebo $a < b$ (relace úplnosti). Dále platí, že pokud $a > b$ pak neplatí současně $b > a$ (relace asymetrie). A konečně platí, že pokud $a > b$ a současně $b > c$ pak také $a > c$ (relace tranzitivity). Matematickým jazykem vyjádřeno: "...na množině zkoumaných objektů je generován rozklad, jehož všechny třídy jsou úplně uspořádány".

Jako příklady pořadových znaků je možno uvést stupnici vyjadřující stupeň souhlasu, míru spokojenosti, pozitivitu nebo negativitu postoje, stupnici vyjadřující četnost prováděných činností, kategorie příchozích podle včasnosti příchodu a pod. Hodnoty znaků mohou být vyjádřeny čísly nebo číslicemi, ale také písmeny nebo jinými symboly :

PŘ.1:

- A - zcela souhlasím
- B - částečně souhlasím
- C - částečně nesouhlasím
- D - zcela nesouhlasím

PŘ.2:

- +2 - velmi spokojen
- +1 - částečně spokojen
- 0 - ani spokojen ani nespokojen
- 1 - částečně nespokojen
- 2 - velmi nespokojen

PŘ.3:

- 1 - denně
- 2 - několikrát týdně

- 3 - jednou týdně
- 4 - 2-3 krát za měsíc
- 5 - jednou měsíčně
- 6 - několikrát ročně
- 7 - jednou ročně
- 8 - méně často
- 9 - vůbec

PŘ.4

- a = mezi prvními, v předstihu
- b = ještě včas, v termínu
- c = opožděně, těsně po termínu
- d = mezi posledními, dlouho po termínu
- e = vůbec

Znaky intervalové

Pro tento typ znaků jsou typické číselné hodnoty, které vyjadřují nejen totožnost a rozdílnost hodnot, jako u nominálních znaků, a nejen pořadí, jako u pořadových znaků, ale také vzdálenosti. Např. pro jednotky A, B, C, D, $a = 2$, $b = 8$, $c = 1$ a $d = 3$, řekneme, že A je vzdáleno od B

$(8-2=6)$ třikrát více než C od D $(3-1=2)$.

PŘ.1: Doba příchodu na plovárnu

- 10 - do 10 hodin
- 11 - po 10 do 11 hodin
- 12 - po 11 do 12 hodin
- 13 - po 12 do 13 hodin
- 14 - po 13 do 14 hodin
- 15 - po 14 do 15 hodin
- 16 - po 15 do 16 hodin
- 17 - po 16 do 17 hodin
- 18 - po 17 do 18 hodin
- 19 - po 18 do 19 hodin
- 20 - po 19 do 20 hodin
- 21 - po 20 hodině

PŘ.2: Rok nástupu do zaměstnání

- 60 - do roku 1960
- 61 - v roce 1961
- 62 - v roce 1962

- 63 - v roce 1963
- . - .
- . - .
- . - .
- 89 - v roce 1989
- 90 - v roce 1990
- 91 - v roce 1991
- 92 - v roce 1992
- 0 - nebyl zaměstnán

PŘ.3: Rok nástupu do zaměstnání - kategorizovaný

- 1 - do roku 1955
- 2 - 1956 - 1960
- 3 - 1961 - 1965
- 4 - 1966 - 1970
- 5 - 1971 - 1975
- 6 - 1976 - 1980
- 7 - 1981 - 1985
- 8 - 1986 - 1990
- 9 - 1991 a později
- 0 - nebyl zaměstnán

Znaky poměrové

Poměrový znak má ve srovnání se znakem intervalovým navíc schopnost vyjadřovat svými číselnými hodnotami přímo poměr mezi hodnotami znaku a , b pro dvě porovnávané jednotky A a B. Můžeme říci, že rodina A má dvakrát více členů než rodina B, nebo že otec je dvakrát starší než syn. Poměrový znak má přirozený počátek, ke kterému jsou vztahovány všechny další hodnoty znaku. To je také důvodem pro možnost srovnávat kolikrát dále od nuly je jedna hodnota znaku než hodnota druhá.

PŘ.1: Počet dětí - vyjádřený číslem

Hodnoty znaku "počet dětí" pro rodiny A,B,C jsou:

$a=6$

$b=2$

$c=3$.

Rodina A má trojnásobek dětí než rodina B a dvojnásobek dětí než rodina C.

PŘ.2: Kapesné - vyjádřené v Kčs měsíčně

učeň 1	-	1000
učeň 2	-	1100
učeň 3	-	1400
učeň 4	-	1600
učeň 5	-	1000
učeň 6	-	1100
student 1	-	250
student 1	-	600
student 1	-	350
student 1	-	450
student 1	-	200
student 1	-	550

Kapesné učně č.1 je 1000 Kčs a kapesné studenta č.1 je 250 Kčs, tj. kapesné učně č.1 je čtyřikrát vyšší. Průměrné kapesné učně činí 1200 Kčs. Průměrné kapesné studenta je pouze 400 Kčs. Můžeme provést srovnání, že kapesné učně je třikrát vyšší než kapesné studenta.

Podrobné třídění:

prostý nominální znak

(je nominální znak, jehož počet hodnot se shoduje s počtem jednotek, každá jednotka má unikátní hodnotu znaku) PŘ.: křestní jména dětí v jedné rodině;

nominální znak

částečně uspořádaný znak

(je znak, jehož část hodnot je úplně uspořádaná a tvoří pořadový znak, avšak obsahuje hodnoty, které nelze vzájemně uspořádat) PŘ.: ;

prostý ordinální znak

(je pořadový - ordinální znak, jehož počet hodnot se shoduje s počtem jednotek, každá jednotka má unikátní hodnotu znaku a jednotky jsou úplně uspořádané)

PŘ.: pořadí příchodu do sálu za předpokladu, že osoby přicházejí po jedné za sebou;

ordinální znak

ordinální znak s počátkem

(je pořadový - ordinální znak, jehož jedna hodnota tvoří přirozený počátek stupnice, tj. neutrální hodnotu nebo nulovou hodnotu) PŘ.:

ordinální znak s uspořádanou metrikou

(Coombsova varianta)

(tvoří mezistupeň mezi pořadovým a intervalovým znakem, známe uspořádání všech hodnot a navíc známe pořadí vzájemných vzdáleností všech dvojic hodnot, znak vzniká jako výsledek Coombsovy škálovací techniky)

PŘ.: V konkrétním armádním sboru lze úplně uspořádat hodnosti důstojníků a navíc lze pomocí výzkumu zjistit, jaká míra zásluh je třeba k povýšení z jedné hodnosti na druhou, vyšší. Nejobtížnější povýšení na bezprostředně následující vyšší hodnost může být z podplukovníka na plukovníka, dále z plukovníka na generálmajora, pak z kapitána na majora atd. Nejsnazší může být postup z podporučíka na poručíka. Seřazeny jsou nejen hodnosti, ale i všechny vzájemné relace mezi nimi.

intervalový znak

poměrový znak

absolutní znak

(má přirozenou interpretaci počátku i měřítka)

4.3.2. Klasifikace z hlediska místa ve výzkumném projektu

A) Nezávisle a závisle proměnné

PŘ.: <i>nezávisle proměnná</i>		<i>závisle proměnná</i>
předpokládaná příčina	->	předpokládaný účinek
předpovídáme z ní		předpovídáme ho
(ovlivňující)		(ovlivňovaná)
věk dítěte	->	výška, váha
prestiž, příjem, povolání, vzdělání	->	status
spokojenost v práci	->	pracovní výkon
inteligence	->	prospěch žáka
vzdělání otce	->	vzdělání syna

B) Aktivní a určované nezávisle proměnné

PŘ.: aktivní nezávisle proměnná určovaná nezávisle proměnná
(můžeme s ní aktivně (nemůžeme s ní aktivně
manipulovat) manipulovat)
(lze ovlivnit) (lze jen volit, kontrolovat)

a) náročnost, zatížení v testu a) pohlaví, věk
b) druh léku, množství léku b) stav pacienta
c) způsob tréninku c) věk, váha, výška sportovce
d) metody, podněty d) inteligence, vloh, rasa

závisle proměnná:

a) schopnost reakce na náročné situace
b) výsledek léčení
c) atletické výkony
d) úspěch, výdrž

C) Zjevné (manifestní) a skryté (latentní) proměnné

PŘ.: zjevná (manifestní) proměnná skrytá (latentní) proměnná
(lze přímo zjišťovat) (lze zjišťovat jen nepřímo)

výsledky testů IQ -> inteligence
příjem, povolání, projevy úcty,
dosažený stupeň vzdělání -> společenský status

5. Populace, výběry, reprezentativita

5.1. Smysl výběrového šetření. Populace, cílový, základní a výběrový soubor

Sociologie vypovídá o společenských jevech, populacích osob, souborech jednotek, které vytvářejí společenská seskupení jako jsou rodiny, obce, podniky, sociální skupiny, národy, státy a pod. V sociologickém výzkumu nezjišťujeme údaje o všech jednotkách, u nichž lze daný jev sledovat. Součástí naprosté většiny výzkumů je výběrové šetření, které slouží k volbě dílčího souboru jednotek, na němž je provedeno zkoumání. Výsledky takového zkoumání jsou pak zobecňovány na soubor zahrnující všechny uvažované jednotky.

Soubory jednotek nazýváme populace a úkolem výběrového šetření je zajistit, aby výsledky platné pro zvolený zkoumaný podsoubor byly zobecnitelné na celý soubor.

Populace = soubor jednotek, o nichž vypovídá výzkum.

Vzhledem k tomu, že některé jednotky souboru, pro něž chceme vyslovit závěry sociologického výzkumného šetření jsou pro výzkum nedostupné, rozlišujeme mezi cílovou populací = to je populace, na niž bychom chtěli vztáhnout naše výzkumné závěry, a základním výzkumným souborem, což je nejširší zachytitelný soubor jednotek.

Cílová populace = soubor jednotek, pro které chceme vyslovit závěry.

Základní populace = (základní soubor) soubor jednotek, které v dané situaci zastupují cílovou populaci.

PŘ.: cílová populace = soubor obyvatel ČSFR
základní soubor = obyvatelé ČSFR mimo dlouhodobě žijících v cizině

PŘ.: cílová populace = žáci základních škol v ČSFR
základní soubor = žáci základních škol mimo těch, kteří jsou dlouhodobě v nemocničním ošetření nebo v cizině, např. s rodiči

PŘ.: cílová populace = rodiny s dětmi
základní soubor = alespoň jeden z rodičů pobírá příspěvek na některé dítě

PŘ.: cílová populace = města v České republice
základní soubor = města zapsaná v lexikonu měst a obcí v České republice

PŘ.: cílová populace = zločiny na území hlavního města Prahy
základní soubor = jen zločiny evidované orgány policie,
prokuratury, soudů

Výzkumné šetření většinou nelze, z ekonomických a praktických důvodů, provádět na všech jednotkách základního souboru resp. cílové populace. Musíme provádět výběr jednotek ze základního souboru. V každém jednotlivém šetření, které není prováděno na celé populaci (jednotek), zastupuje základní soubor výběrový soubor (výběr, setkáme se také s označením vzorek, z anglického sample).

Výběrový soubor = je soubor jednotek (pro něž jsou k dispozici data nebo pro něž tato data získáváme), který zastupuje základní soubor ve výzkumu. Výsledky pro něj zjištěné mají být zobecnitelné na základní soubor.

Abychom byli schopni zvolit výběrový soubor, jehož výsledky by zastoupily údaje za nedostupný celek, používáme tzv. oporu výběru. Tou je buď seznam jednotek základního souboru nebo několikastupňový předpis výběrových kroků zajišťující předepsanou šanci být vybrán pro všechny uvažované jednotky základního souboru.

Opora výběru = seznam jednotek cílové nebo základní populace, z něž je vybírán zkoumaný výběrový soubor nebo jednoznačný výběrový předpis, kterým je tento soubor získán

Účelem výběrového souboru je poskytnout údaje, které by bylo možno zobecnit na celý základní soubor.

Zkoumat populaci výběrového souboru je:

- levnější
- rychlejší (zvládnutelné v kratším čase)
- je možno zajistit standardnější provedení
- často ani technicky nelze zkoumat celý základní soubor

Důvody výběrového šetření:

některé populace není možno zkoumat celé
a) všechny jejich jednotky nejsou dostupné

- b) na výzkum není tolik času
- c) finanční důvody - výzkum by byl příliš drahý
Např. nikdo není v současné době schopen organizovat výzkum všech obyvatel světa.
- d) některé typy zkoumaných problémů vyžadují zničení zkoumaných jednotek v průběhu výzkumu
Př.: bariérové zkoušky automobilů
zkoušky odolnosti proti korozi a pod.
- e) větší přesnost při výzkumu problémů, při nichž je obtížné zastihnout všechny členy populace
(např. rozptýlení v nedostupné oblasti)

5.2. Reprezentativita výběrového souboru

Úkolem výběrového šetření je vybrat zkoumané jednotky tak, aby údaje o nich zároveň přinášely informaci také o všech jednotkách, které do výběru nebyly zařazeny. Kvalita a spolehlivost zastupování základního souboru výběrovým souborem je vyjadřována pojmem reprezentativita.

Reprezentativita:

výběr je reprezentativní v té míře, nakolik odráží charakteristiky populace.

- Např. - jestliže 10 % populace má blahobyť, výběr by měl ukazovat stejné procento
- jestliže 50 % domácností má jeden osobní automobil a 8 % dva, ve výběru by tomu mělo být také tak.

Závažným krokem sociologického výzkumu je přesné, jasné (zřetelné) vymezení populace, základního souboru. Dobře definovaná populace usnadňuje provedení výběru a je daleko pravděpodobnější, že se podaří stanovit reprezentativní výběrový soubor z dobře definovaného základního souboru.

Reprezentativitu výběru je obtížné, ne-li vůbec nemožné, kontrolovat. Často je výběr prováděn právě proto, aby byla odhadnuta neznámá hodnota charakteristiky pro celou populaci. Při hodnocení reprezentativity tedy často není s čím srovnávat. Za ukazatele reprezentativity výběru se všeobecně považuje spolehlivost výběrového postupu.

Reprezentativita závisí na:

- přesnosti vymezení populace
- adekvátnosti výběru
- heterogenitě (homogenitě) populace
(srovnej [Labovitz-Hagedorn 1976:46-47])

Spolehlivost výběru roste, jestliže je populace správně definována. Např. jestliže existuje seznam jmen všech mužů zaměstnaných v Praze nebo seznam všech studentů letos nově přijatých na vysoké školy v ČSFR.

Adekvátnost výběru je důležité uvažovat jen u velmi malých výběrů. Adekvátní výběr musí být dostatečně rozsáhlý, aby umožňoval výzkumníkovi důvěřovat závěrům získaným statistickými prostředky. Např. při zkoumání sexuální potence mužů můžeme sotva důvěřovat zobecnění výsledků čtyř mužů z Československa na celou populaci mužů ve státě; nebo při kontrole kvality broskví balených do lísek nemůžeme mít důvěru ve výběr jedné lísky jako reprezentanta celého miliónu takových lísek broskví.

Reprezentativita závisí také na stupni homogenity populace: Náhodný výběr 100 vojáků základní vojenské služby reprezentuje tuto skupinu populace homogenní z hlediska věku a pohlaví patrně lépe než náhodný výběr 100 televizních diváků celou diváckou obec.

Je jistě snazší pořídit vzorek z barelu nafty než vytvořit výběr z populace všech obyvatel USA nebo Indie. V chemické laboratoři vystačíme s jednou kapkou nafty, abychom zjistili chemické složení celého barelu, u populace nám bude činit potíže národní a etnická příslušnost, jazyk, kultura a mnoho dalších charakteristik vázaných na různým způsobem diferencované území rozlehlého státu.

5.3. Chyba výběru, vychýlení a přesnost, standardní chyba výběru

Údaje zjištěné výběrovým šetřením se liší od skutečné hodnoty, kterou bychom naměřili v rámci celého základního souboru. Rozdíl naměřené hodnoty od této teoretické, v naprosté většině případů neznámé, hodnoty nazýváme chyba výběru.

Chyba výběru může být náhodná, t.j. taková chyba, která se s velikostí výběrového souboru zmenšuje, jejíž výše při opakování šetření na dalších stejně velkých výběrových souborech vybraných stejným postupem z téhož základního souboru se nemění. Je způsobena jen skutečností, že výběrový soubor nikdy nereprezentuje základní

soubor přesně. Náhoda, s níž jsou do našeho konkrétního výběru zařazeny dané konkrétní jednotky základního souboru a ne jiné, které v něm také mohly být, způsobuje tento hodnotový rozdíl – náhodnou chybu.

Chyba výběru však může být také systematická. Je to chyba, která se projevuje v situacích, kdy je porušena náhodnost výběru. Např. některé jednotky mají pravděpodobnost zařazení do výběru systematicky vyšší. Důvodem může být nepřesnost vymezení populace, nesprávná nebo neúplná opora výběru, neadekvátní nebo nesprávně provedený výběr.

Systematická chyba narušuje reprezentativitu výběru. Výsledek zatížený touto chybou se liší od skutečné hodnoty platné pro základní soubor o hodnotu zvanou vychýlení.

Vychýlení (posunutí) (bias):

je veličina, o niž se liší hodnota výběrového průměru od populačního průměru, a to zpravidla opakovaně ve všech výběrech provedených stejným způsobem podle stejného předpisu. (Srovnej [Moser-Kalton 1971:65])

Kromě vychýlení resp. nevychýlení odhadu populačního průměru určujeme také přesnost měření, přesnost pozorování.

Přesnost je určována stupněm rozptýlení pozorovaných (naměřených) hodnot kolem výběrového průměru. Jestliže je rozptýlení naměřených hodnot malé, je přesnost vysoká, je-li naopak rozptýlení velké, je přesnost nízká.

Při zjednodušeném pohledu rozeznáváme čtyři možné případy výsledků výběrového šetření (viz. [Moser-Kalton 1971:68n]):

- A) - velké vychýlení + malá přesnost
- B) - velké vychýlení + velká přesnost
- C) - žádné vychýlení + malá přesnost a
- D) - žádné vychýlení + velká přesnost.

Vždy, pokud nám to praktické okolnosti šetření umožní, dáváme přednost přesnému nevychýlenému pozorování (D).

V ostatních případech je třeba se rozhodnout, zda reprezentativně výsledku méně uškodí systematické vychýlení (tj. opakovaně se projevující rozdíl mezi výběrovým a populačním průměrem) nebo snížená přesnost (tj. vysoký rozptyl naměřených hodnot kolem výběrového průměru a tedy malá spolehlivost odhadu hodnoty populačního průměru). Univerzální předpis v tomto případě neexistuje.

Při přesném měření s vychýlením se pokoušíme odhadnout směr, ev. velikost tohoto vychýlení a tím zkvalitnit výsledky. Při nevychýleném

měření je naší snahou zvyšovat přesnost, tj. snižovat výběrovou chybu, např. vhodným výběrovým postupem nebo zvětšováním výběru.

Nevychýlený postup odhadu (unbiased estimate):

Jestliže pro daný výběrový postup je očekávaná hodnota konečného odhadu rovna populačnímu parametru (tj. většinou neznámé hodnotě charakteristiky pro celou populaci), je postup odhadu (estimate) nevychýlený (unbiased).

Příklad (prostý náhodný výběr):

(Srovnej [Moser-Kalton 1971:63-69])

Populace = čtyři muži ve věku 15,17,18,22 let.

Průměrný věk = populační průměr = 18 let.

Výběry dvojic..... výběrové odhady průměrného věku

15 a 17	16,0	
15 a 18	16,5	výběrová
15 a 22	18,5	distribuce
17 a 18	17,5	průměrů
17 a 22	19,5	
18 a 22	20,0	let
Součet	=====	108,0	

Očekávaná hodnota průměru 108,0 : 6 = 18,0 let

Závěr: Prostým náhodným výběrem dospějeme
k nevychýlenému odhadu průměru

5.3.1. Standardní chyba výběru

Jak vyplývá z předchozího příkladu, potřebujeme míru, která by nám řekla, jak se od sebe liší odhady získané jednotlivými výběry provedenými stejným postupem (za předpokladu použití pravděpodobnostního - náhodného výběru). Tato důležitá míra se jmenuje standardní chyba průměru S.E. (standard error of the mean). Je mírou přesnosti odhadu, není mírou vychýlení.

5.4. Pravděpodobnostní výběry

(Srovnej [Moser-Kalton 1971: 79-117])

5.4.1. Náhodný výběr a jeho výhody

Způsob, jakým lze provádět výběr jednotek pro výzkumné šetření, je v podstatě dvojitý: náhodný - pravděpodobnostní výběr a záměrný -

nepravděpodobnostní výběr. Rozdíl mezi nimi spočívá v tom, že v prvním případě (u náhodného výběru) je stanovena pravděpodobnost výběru každé z částí základního souboru do souboru výběrového. Důsledkem toho je, že jsme schopni matematickým výpočtem odhadnout velikost chyby, která je způsobena výběrovým šetřením. U záměrného výběru nemáme pro odhad chyby matematicko- statistickou oporu.

5.4.2. Typy pravděpodobnostního výběru

Náhodné - pravděpodobnostní výběry jsou založeny na snaze zachovat stejnou šanci všem jednotkám nebo stejně velkým skupinám jednotek, aby byly vybrány a reprezentovaly populaci. Náhodnost výběrové procedury brání vychýlení výběru.

Jednotlivé typy pravděpodobnostního výběru se mezi sebou liší způsobem, jak se snaží:

- zvýšit praktickou použitelnost výběrového postupu,
- zvýšit přesnost výběru.

Rozeznáváme různé typy pravděpodobnostního výběru: prostý náhodný výběr, stratifikovaný náhodný výběr, skupinkový výběr, víceúrovňový náhodný výběr.

Prostý náhodný výběr:

všechny podsoubory dané velikosti mají stejnou šanci být realizovány jako výběr.

Vybíráme, buď losem nebo pomocí tabulek náhodných čísel nebo jiným postupem zajišťujícím neovlivnění výběrového postupu zásahem výzkumníka, jednotky buď ze seznamu nebo z množiny čísel zastupující seznam jednotek základního souboru.

Stratifikovaný náhodný výběr:

- 1) Základní soubor je rozdělen do podsouborů, které nazýváme strata. Strata jsou takové podmnožiny základního souboru, jejichž jednotky mají v rámci strata společnou vlastnost nebo více společných vlastností, a strata se mezi sebou liší zpravidla více než jejich jednotky uvnitř.
- 2) V každém stratu provedeme prostý náhodný výběr.

Tímto postupem se zvyšuje přesnost pozorování (měření), snižuje se předpokládaný rozptyl naměřených hodnot v rámci strat. Tím se také snižuje náhodnost. Je-li správně určen, odhadnut, skutečný podíl

každého strata na základním souboru, lze odhady populačních hodnot z naměřených výběrových hodnot zpřesnit.

Např. názory mužů a žen jednotlivých věkových kategorií nebo vzdělanostních skupin obyvatelstva na otázky ekonomické reformy, reformy trestního práva nebo úpravy státního uspořádání se pravděpodobně liší. Výběr tedy uspořádáme stratifikovaně podle věkových skupin respondentů mužů a žen. Musíme však znát podíly obyvatelstva těchto demografických skupin v základním souboru abychom mohli stanovit podíl jednotlivých strat na výběru. Pokud jsme postupovali správně, snížili jsme možnou chybu, kterou by mohlo způsobit menší nebo naopak výrazně vyšší zastoupení některé kategorie obyvatelstva ve výběru v případě realizace prostého náhodného výběru.

Stratifikovaný náhodný výběr může být

proporcionální:

- velikost výběru z vrstvy (strata) je úměrná velikosti populace této vrstvy

nebo neproporcionální:

- užívá se, pokud je významně odlišný rozptyl v některé vrstvě nebo pokud je třeba zajistit za účelem srovnávací analýzy dostatečný, (tj. většinou vyšší) počet zkoumaných jednotek v některých stratach

Skupinkový výběr: 1. a 2. krok

Víceetapový náhodný výběr: společný postup

- 1) Základní soubor je (většinou přirozeně) rozdělen do podsouborů, které v tomto případě nazýváme skupinky. Na rozdíl od vrstev (strat) tvoří skupinky podsoubory, o nichž předpokládáme, že jsou v zásadě vzájemně zastupitelné, nijak výrazně se od sebe neliší a rozdílnost mezi jednotkami se skrývá spíše uvnitř skupinek a je způsobena jinými charakteristikami než příslušností ke skupinkám.
- 2) Ze základního souboru vybereme náhodným postupem jen některé ze skupinek.

Skupinkový výběr: 3. krok samostatně

(cluster sampling)

- 3) Vybrané skupinky, tj. všechny jejich jednotky, šetříme vyčerpávajícím způsobem. Vybrané skupinky tedy reprezentují

ostatní skupinky (a jejich jednotky), které se nedostaly do výběru.

Víceetapový náhodný výběr: 3. krok samostatně
(multi-stage sampling)

- 3) Ze souboru vybraných skupinek provedeme náhodně další stupně výběru;
 - buď přímo výběr jednotek
 - nebo nejprve výběr podskupinek a v nich výběr jednotek.
- Vybrané jednotky z těchto skupinek nebo podskupinek reprezentují celý základní soubor.

5.5. Ostatní typy výběrů - nepravděpodobnostní

5.5.1. Kvótní výběr (Srovnej [Moser-Kalton 1971:127-137])

Z ostatních výběrů, tzv. nepravděpodobnostních, bývá v sociologickém výzkumu nejčastěji užíván kvótní výběr. Název tohoto typu výběru je odvozen od kvót, kterými je vymezen rozsah výběru jednotlivých, nejčastěji demografických nebo sociálních skupin. Kvóta je číselný údaj udávající tazateli kolik má dotazovat žen, kolik mužů, nebo kolik má vyhledat obyvatel s ukončeným jen základním vzděláním, kolik s maturitou a kolik s vysokoškolským vzděláním apod.

Postup kvótního výběru:

- 1) Zvolíme výběrové charakteristiky, pro něž budou stanoveny kvóty. V praxi volíme nejvýše pět znaků. Jako vhodné, použitelné bývají označovány relativně snadno zjistitelné údaje, které jsou buď viditelné nebo je dotázaný ochoten je v situaci prvního kontaktu přiznat a neuvedou ho do nežádoucích rozpaků. Zároveň to bývají hlediska, která populaci dosti významným způsobem rozlišují.

Jako kvóty se nejčastěji používají znaky:

pohlaví, věková kategorie, dokončený stupeň vzdělání, rodinný stav, charakter bydliště, skupina povolání, sociální skupina; podle předmětu šetření mohou být zvoleny také další, jako jsou počet sourozenců, počet dětí, vlastnictví určitého předmětu, např. osobního auta, národnost a jiné.

2) Pro zvolené kvóty vyhledáme v dostupných statistických pramenech kvantitativní oporu pro rozhodnutí, jak velké mají být podíly jednotlivých kategorií dotázaných podle vybraných znaků.

Jestliže je například ve sledované populaci stejně početné zastoupení mužů a žen, budou kvóty pro výběr mužů a žen stejné; najdeme-li ve statistické ročence údaj, že v dané věkové skupině, kterou šetříme, je podíl vysokoškoláků mezi muži jedna sedmina a mezi ženami jedna desetina, přizpůsobíme velikost kvóty právě těmto údajům: budeme tedy vybírat ze sedmdesáti mužů deset s ukončeným vysokoškolským vzděláním a ze sedmdesáti žen sedm, které získaly vysokoškolský diplom.

3) Rozhodneme se, zda budeme zadávat kvóty pro výběr nezávisle na sobě nebo ve vzájemných vazbách. Každá z obou možností má své výhody a nevýhody.

- Jestliže ponecháme kvóty pro výběr bez vzájemné závislosti, bude pro tazatele snazší najít příslušný stanovený počet osob s určenými charakteristikami a dotázat je. Může se však stát, že nahodile zkreslí obraz o zkoumané populaci z toho důvodu, že vybere respondenty s netypickými kombinacemi hodnot znaků nebo naopak všechny jen v nejčastěji se vyskytujících kombinacích. Takto kvótně vybraný soubor může pak obsahovat sice požadovaných 25 mužů a 25 žen a také splňovat požadavek na 30 osob se základním, 15 s úplným středoškolským a 5 s vysokoškolským vzděláním; může se však stát, že všech pět vysokoškoláků budou muži, čtrnáct z patnácti středoškoláků budou mladé ženy apod.
- Jestliže však naopak náš nárok na vzájemné propojení kvót bude přílišný, znesnadníme práci tazatelům natolik, že tím můžeme také ohrozit spolehlivost dat nebo počet uskutečněných rozhovorů (návratnost výzkumné akce). Příkladem může být požadavek vyhledat na malé vesnici dvě vdané vysokoškolačky ve věku 25–34 let, svobodného muže se základním vzděláním ve věku 35–50, vdovu nad 50 let se středním vzděláním apod. Tento požadavek, který lze ve velkém městě s určitým úsilím splnit, nabývá značně na obtížnosti na malé vsi, kde může být složení obyvatel jiné, než by odpovídalo nějaké průměrné představě.

4) Stanovíme výběrové kvóty pro každého tazatele tak, aby pokrývaly potřebný počet a strukturu dotázaných. Skupiny s výběrovými charakteristikami, o nichž předpokládáme, že budou obtížně dostupné, ve výběru nahodnotíme. Nepodaří-li se všem tazatelům

získat potřebný počet např. žen v domácnosti nebo vysokoškoláků apod., vyšší kvóta tuto nepřesnost sníží.

5.5.2. Systematický výběr

Systematický výběr je významným typem nepravidelnostního výběru, který se některými svými znaky blíží výběru pravděpodobnostnímu. Je založen na zásadě výběru jednotek ze základního seznamu podle pevně zvoleného kroku. Chceme-li vybrat desetinu souboru, vybíráme ze základního souboru každou desátou jednotku.

Zásadní rozdíl oproti pravděpodobnostnímu výběru spočívá v tom, že většinou není s dostatečnou jistotou zaručeno náhodné pořadí jednotek v souboru. Tak se může stát, že zvolený výběrový krok souvisí se skrytou pravidelností uspořádání jednotek v seznamu, z něž vybíráme. Vybrané jednotky tak mohou mít některou vlastnost významně častěji než by odpovídalo jejímu zastoupení v celé populaci.

Bývá uváděn příklad výběru domů ve vilové čtvrti, kdy délka zvoleného výběrového kroku, např. každý desátý dům, se může nahodile shodovat s pravidelně se opakujícím jevem, např., že každý desátý dům je postaven na pozemku rohové parcely a je tedy odlišný od ostatních devíti desetin domů. Taková situace není častá v českých městech, bývá však uváděna jako příklad pravidelnosti u městských celků v USA. Výběrem s tímto systematickým krokem získáme tedy buď soubor majitelů rohových domů nebo výběrový soubor, v němž vůbec nebudou zastoupeni. V obou případech se zvyšuje chyba a výběr neodpovídá pravidlům pro výběry pravděpodobnostní.

5.5.3. Výběr úsudkem

Ve speciálních případech, zvláště pro malé výběry, může výzkumník dobře obeznámený se zkoumanou problematikou použít výběr úsudkem. Pracovník manželské poradny může pro experiment navrhnout soubor manželských párů, které podle jeho zkušeností reprezentují typické manželské dvojice obracející se na manželskou poradnu. Kriminolog může zodpovědně vybrat malý vzorek vězňů, o nichž lze říci, že reprezentují kategorii odsouzených s vyšší pravděpodobností budoucí recidivy v případě propuštění. Obecně, v situaci, kdy se rozhodujeme pro výběr překračující několik desítek osob, lze však použití výběru úsudkem označit za vysoce riskantní. Chybí objektivní měřítko reprezentativity

takového souboru a výsledky z něj získané nelze většinou spolehlivě zobecnit.

5.5.4. Výběr nabalováním

Výběr nabalováním (snowball sampling) má využití přede-vším při výzkumu speciálních skupin obyvatelstva, pro něž neexistují seznamy ani spolehlivá opora výběru. Hledáme např. lidové léčitele, programátory, sběratele starožitností, bývalé studenty obchodní akademie a podobně. Předpokládáme, že každý z nich zná alespoň některé ze svých kolegů a že nám bude ve výběru nápomocen.

V prvním výběrovém kroku se obrátíme na jistý zárodečný soubor jednotek (jednotlivců nebo organizací) se žádostí o informace o dalších jednotkách vhodných pro zařazení do výběru. Předpokládáme, že na základě zjištěných adres nebo referencí se v druhém kroku dotážeme již širšího okruhu osob nebo institucí a s jejich pomocí rozšíříme výběrový soubor ještě významněji.

Jak napovídá název výběru, využívající průměr s valící se sněhovou koulí, výběr roste až do určité velikosti tím rychleji, čím větší počet osob v daném kole dotážeme. Počet kol je závislý na velikosti zárodečné skupiny, požadované celkové velikosti výběru i na počtu nově získaných adres v každém z následujících kol. Nepřibývají-li nová jména a adresy se začínají opakovat, výběr raději ukončíme.

6. Uspořádání výzkumu. Typy výzkumných projektů.

Sociologický výzkum založený na zkoumání vybraných jednotek je možno uspořádat mnoha způsoby podle účelu, k němuž mají být data využita a podle míry reprezentativity, kterou od výsledků požadujeme. V literatuře jsou typy výzkumných projektů, jednotlivé způsoby uspořádání výzkumu označovány pojmem Research Design, který volně překládáme jako uspořádání výzkumu nebo výzkumný projekt. (Termín research design je však třeba rozlišit od termínu projekt výzkumu, kterým označujeme podrobný plán provedení výzkumné akce.)

6.1. Členění výzkumných projektů

Výzkumné projekty členíme na:

případové studie (monografie) [Case Study],
při nichž zkoumáme jednotlivé případy, vyjimečné nebo naopak typické jednotky, používáme pro výzkum malých komunit

průzkumy (pozorování) [Surveys, Investigations],
při nichž zkoumáme jev tak, jak probíhá, neovlivňujeme ho, zjišťujeme na základě šetření výběrového souboru populační charakteristiky

experimenty [Experiments],
při nichž ovlivňujeme nezávisle proměnné a měříme vliv těchto změn jak se projevuje u závislých proměnných, měříme závisle proměnné (srovnej [Labovitz-Hagedorn 1976: 54-70]).

6.2. Cíle výzkumného šetření

Výběr výzkumného uspořádání je určován mimo jiné cílem výzkumného šetření. Jednotlivé výzkumné cíle se liší z hlediska náročnosti, která je na výzkum kladena.

Cílem výzkumného šetření může být:

popis (deskripce), který používáme v situaci, kdy o zkoumaném objektu mnoho nevíme; zjišťujeme buď přímo vybrané populační charakteristiky nebo shromažďujeme fakta o zkoumaném jevu. Deskriptivní výzkum často slouží jako příprava k výzkumu s náročnějšími cíli.

zjišťování souvislosti, závislosti proměnných (korelace), které tvoří nejčastější výzkumný záměr, často však nepřesně formulovaný;

srovnání (komparace) populací, nejčastěji zemí, regionů, tedy srovnání mezinárodní nebo mezioblastní. Při tomto druhu srovnání je obtížné odstranit rozdíly v chápání stejných pojmů v různých zemích a kulturách.

Př.:

porovnávané mládež několika zemí a jednou z podstatných rozlišujících proměnných je dosažené vzdělání; je třeba si uvědomit, že se mezi sebou liší vzdělávací systémy těchto zemí, společenské systémy, tradice i kultura obyvatel. Vysokoškolské vzdělání v USA, ve Velké Británii, v Německu, Maďarsku, Československu a Sovětském svazu představují do značné míry rozličné pojmy lišící se jak mírou vnitřní diferenciací v jednotlivých zemích, tak i obsahem, náplní učiva a v neposlední řadě v jednotlivých oborech také úrovní dosažených znalostí a vědomostí.

odhalování příčinných (kauzálních) vztahů, patří k nejcennějším a nejobtížnějším zaměřením výzkumné práce; takový výzkum musí být důkladně připraven, často bývá předcházen jedním nebo několika šetřeními s deskriptivním, korelačním nebo komparačním poznávacím zaměřením; jde v něm o přijetí nebo odmítnutí hypotézy nebo řady hypotéz o vztazích mezi zkoumanými proměnnými, o vysvětlení souvislostí mezi nezávislými a závislými proměnnými; v takto orientovaném výzkumném šetření jde často o to zjistit příčinu, nebo celý řetězec příčin, souvislosti nebo souvislostí, o nichž víme již z předchozích šetření;

zjišťování vývojových procesů, které patří k méně čtým a vysoce náročným výzkumným cílům; je k němu třeba opakovaných výzkumných šetření, při nichž je dotazován buď panel respondentů, tj. opakovaně dotázaný soubor osob, nebo je celé výzkumné šetření organizováno v několika časových etapách v odstupu i několika let jako tzv. longitudinální výzkum; jsou využívány časové řady údajů, je však třeba si být vědom možnosti posunu významu jednotlivých skutečností a změně obsahu zjišťovaných pojmů za dlouhé zkoumané období.

Př.:

Zkoumáme-li proměny vybavenosti domácností předměty dlouhodobé spotřeby, v průběhu let se budou měnit předměty, které jsou považovány za standard vybavení (radiopřijímač, pračka v roce 1960, televizor, chladnička, vysavač, osobní auto v roce 1980), v různých zemích bude existovat různý standard (mikrovlnná trouba, mraznička, video, myčka nádobí, sušička prádla, osobní počítač a pod.) a také předmět označený stejným

názvem může nabývat kvalitativně velice rozdílných poloh: např. osobní automobil (za deset, sto nebo pět set tisíc Kčs).

6.3. Typy výzkumných projektů

Z hlediska uspořádání výzkumu rozlišujeme několik typů výzkumných projektů:

6.3.1. Neexperimentální výzkumné projekty, mezi něž řadíme případovou studii, korelační výzkum, srovnávací výzkum a longitudinální výzkum:

- případová studie
se zabývá výzkumem jedné vyčleněné skupiny, komunity, jednotky, většinou bez nároku na širší zobecnitelnost na celou populaci
- korelační výzkum,
jehož cílem je odhalení souvislostí mezi proměnnými, souvislostí většinou vyjádřených jen korelacemi bez znalosti směru ovlivňování, bez prokazatelných příčinných souvislostí
- srovnávací výzkum
zaměřený na porovnání podsouborů, populací z hlediska vybraných charakteristik
- longitudinální výzkum
soustřeďující se na zkoumání delšího časového období, většinou několika let, na sledování zvolené populace systémem několika vzájemně spjatých nebo navazujících výzkumných šetření

6.3.2. Jednoduché experimentální výzkumné projekty, kam řadíme kvaziexperiment, ex-ante - ex-post experiment a experiment s kontrolní skupinou:

- kvaziexperiment
se vyznačuje zjišťováním skutečností jen po zvolené události, tj. bez možnosti zjistit, zda zaznamenaný jev je způsoben touto událostí nebo jiným bližším jevem
- experiment (ex-ante - ex-post)
je založen na zjišťování stavu sledovaného jevu předběžně před (ante) i následně po (post) události, jejíž vliv uvažujeme, předpokládáme, testujeme; nevýhodou tohoto typu experimentálního uspořádání je skutečnost, že časový soulad

změn sledovaného jevu s výskytem uvažované události ještě nemusí znamenat, že ona událost zjištěnou změnu způsobila, může jít o náhodnou časovou souslednost

- experiment s kontrolní skupinou (před i po) odstraňuje námitku často vznášenou proti předchozímu typu výzkumného uspořádání, kontrolní skupina je vystavena všem působícím vlivům s výjimkou události, jejíž vliv posuzujeme, tak je možno rozlišit, zda změny způsobuje tato událost nebo souhrn ostatních okolností probíhajícího jevu.

6.3.3. Faktoriální experimentální výzkumné projekty, jejichž smyslem je zjišťovat samostatný nebo společný vliv několika proměnných (faktorů).

Náročnější výzkumná situace se promítá do větších nároků na počet zkoumaných jednotek, na jemnější rozlišení situací spojených s izolovaným i kombinovaným působením zkoumaných faktorů a dalších proměnných. Ve faktoriálních experimentálních výzkumných projektech sledujeme několik nezávisle a několik závisle proměnných, rozlišujeme také několik kontrolních skupin. Ve speciální literatuře jsou rozpracovány způsoby, jak uskutečnit výběr jednotek tak, aby vzájemně kombinoval samostatný i společný vliv všech uvažovaných faktorů a přitom extrémně nevzrostly nároky na počet zkoumaných jednotek, na velikost výběrového souboru.

6.4. Kritéria hodnocení experimentálních výzkumných projektů (Srovnej [Labovitz-Hagedorn 1976: 56-60])

- 1) Je možná kontrola shodného složení experimentálních a kontrolních skupin?
Př.: Srovnatelnost by nebyla zajištěna např. pokud by experimentální skupina byla vybrána z městského a kontrolní skupina z venkovského obyvatelstva, experimentální ze studentů a kontrolní z učňů a pod.
- 2) Jsou pod kontrolou vnější faktory?
Někdy nevíme nakolik se uvažované skupiny liší, zda přímo jako výběrový faktor při zařazení do některé ze skupin nepůsobí významná proměnná, kterou neuvažujeme a jejíž vliv bychom tedy opomněli.

Př.: Nakolik se například liší lidé s cukrovkou a bez ní z hlediska věku, pohlaví, způsobu života, životního prostředí svých předků a pod.? Vliv těchto vnějších faktorů se snažíme výběrovými postupy všemožně snižovat.

- 3) Lze provádět analýzu změn pretest - posttest po jednotkách?
Nejsou jednotky zkoumané před událostí významně odlišné od jednotek zkoumaných po sledované události?
- 4) Je výzkumné uspořádání reprezentativní? Reprezentuje výběr základní soubor?
Př.: Chování většiny lidí v laboratorních podmínkách je odlišné od jejich běžných projevů. Nelze tedy experimentální skupinu zkoumat v laboratoři a kontrolní skupinu mimo ni. Nelze také výsledky laboratorního experimentu nekriticky zobecňovat na přirozené podmínky života lidí, v nichž mohou platit jiné než laboratorně naměřené hodnoty.

Je možno si položit otázku: Jak je možno kontrolovat, zjišťovat, event. ovlivňovat vnější faktory?

6.5. Kontrola vlivu vnějších proměnných

Existují v zásadě tři způsoby, jak kontrolovat vliv vnějších proměnných při výzkumu: vyrovnávání, vážení při analýze a znáhodňování. Vyrovnávání a znáhodňování je častější při experimentálním výzkumu, vážení při analýze častěji užíváme v neexperimentálním výzkumném uspořádání.

Vyrovnávání (matching) mezi experimentálním a kontrolním souborem

Cílem tohoto způsobu kontroly je vybrat pro experiment sobě co nejpodobnější experimentální a kontrolní soubory. Jejich podobnost zajišťujeme dvěma způsoby:

- vyrovnáváním v párech, kdy ke každé jednotce určitého typu zařazené do experimentální skupiny najdeme jednotku stejného typu pro kontrolní skupinu. Tento způsob vyrovnání je poměrně náročný na výběr, protože sleduje několik charakteristik zkoumaných osob (objektů) ve vzájemných kombinacích.
Př.: Vytváříme experimentální a kontrolní skupinu a vyrovnáváme v párech jejich složení z hlediska pohlaví, věku, vzdělanostní skupiny a rodinného stavu, protože se domníváme, že všechny tyto charakteristiky ovlivňují sledovaný jev ve vzájemných kombinacích. Znamená to, že k jedné svobodné vysokoškolačce do

30 let v experimentálním souboru vybereme svobodnou vysokoškolačku ze stejné věkové skupiny do souboru kontrolního, že ke třem vyučeným ženatým mužům ve věku 41 až 50 let v experimentálním souboru najdeme tři muže se stejnými charakteristikami pro kontrolní soubor atd. až získáme zrcadlové obrazy souborů, které si budou párově odpovídat z hlediska všech čtyř sledovaných charakteristik.

- vyrovnáváním v rozloženích za soubory, při němž nevyhledáváme určité jednotky totožné z hlediska několika charakteristik, ale spokojujeme se s identickým rozložením souborů z hlediska těchto charakteristik aniž bychom sledovali jejich kombinace.

Př.: Do experimentálního i do kontrolního souboru jsme zařadili stejné počty mužů a žen, stejné počty svobodných, ženatých nebo vdaných, rozvedených, vdov nebo vdovců, stejné počty osob v jednotlivých věkových kategoriích a stejné počty osob s různým stupněm dokončeného vzdělání. Neurčovali jsme však již kombinace těchto charakteristik. Může se tedy stát, že v experimentálním souboru budou ve větší míře zastoupeny mladé svobodné vysokoškolačky a středně staří ženatí vyučení muži a v kontrolním souboru bude naopak více mladých svobodných vysokoškoláků a středně starých vdaných vyučených žen. Přitom z hlediska všech jednotlivých izolovaně posuzovaných identifikací budou soubory identické.

Vážení při analýze (adjustment)

Po provedení sběru dat zjistíme, že soubory, za něž bychom rádi srovnali výsledky, se významně liší svým složením. Nezbyvá nám než pokusit se o změnu váhy, kterou přisuzujeme údajům za jednotlivé skupiny jednotek. Jestliže původně měla každá jednotka váhu jedna, po převážení bude deseti jednotkám (např. svobodným mladým mužům, které je obtížné při dotazování zastihnout) přiřazena váha patnácti jednotek (každé jednotce váha 1.5) a naopak jiným deseti jednotkám (např. vdaným ženám středního věku) bude přiřazena váha pouhých pěti jednotek (každé jednotce váha 0.5). Simulujeme tak situaci, při níž podily dotázaných ve výběru odpovídají jejich skutečným poměrům v populaci.

Znáhodňování (randomization)

Náhodně přiřazujeme jednotlivé osoby (zkoumané jednotky) jednou do experimentální, jednou do kontrolní skupiny. Používáme k tomu tabulek náhodných čísel, hodu mincí, kostkou a pod.

6.6. Vnitřní a vnější validita

Při přípravě a provádění výzkumů, zvláště experimentů sledujeme platnost (validitu) výsledků. V literatuře [viz.např. Moser-Kalton:233] je rozlišována vnitřní (interní) validita a vnější (externí) validita.

Vnitřní validita se týká otázky, zda je použita vhodná míra pro zjišťování vlivu experimentální proměnné. Měříme opravdu to, co chceme měřit?

Vnější validita se týká zobecnitelnosti výsledků zjištěných ve výběrovém šetření na širší populaci. Platí to, co jsme zjistili pro výběrový soubor, také pro celou populaci?

7. Terénní sběr dat, jeho fáze a techniky

7.1. Sběr dat v procesu sociologického výzkumu

Etapa sběru dat zaujímá v procesu sociologického výzkumu místo mezi fázemi teoretické přípravy a operacionalizace a fázemi analýzy dat a prezentace výsledků z výzkumu.

Schéma

1. Teoretický rozbor problematiky
2. Formulace problému a tvorba hypotéz
volba znaků a indikátorů
3. Terénní sběr dat a SOUČASNĚ volba uspořádání výzkumu
výběrový soubor
4. Příprava a čistění (kontrola) dat
5. Analýza dat a interpretace
6. Prezentace výsledků a závěry

7.2. Důležité fáze sběru dat

Nan Lin [1976:197nn] rozlišuje osm fází sběru dat v empirickém sociologickém šetření od konspektování literatury a konzultací s odborníky až po organizaci a překódování dat do podoby vhodné pro hromadné zpracování.

7.2.1. Konspektování literatury a konzultace s odborníky

Výzkumník se seznamuje s dosud dosaženými výsledky ve výzkumu zvolené problematiky, s použitými znaky, testovanými hypotézami a vztahy mezi proměnnými zkoumanými a prokázanými v předchozích výzkumech. Studuje literaturu k řešeným problémům a porovnává své představy a hypotézy s poznatky publikovanými v odborných statích, monografiích, výzkumných zprávách i sociologické publicistice.

7.2.2. Získání vstupu do sociálního systému

Způsob, jakým výzkumník vstupuje do terénu, v němž chce provádět výzkum, může do značné míry ovlivnit jeho budoucí možnosti při sběru dat, snadnost nebo naopak obtíž při přístupu k datům, ochotu ke spolupráci, kterou ve zkoumaném prostředí nalezneme.

Musí se v zásadě rozhodnout pro jednu z rolí ve vztahu ke zkoumanému sociálnímu systému. Buď bude vystupovat jako identifikovaný se systémem, jako "insider", nebo jako nezávislý na systému, jako tzv. "outsider".

A) Výhody nezávislého výzkumníka:

- má více času (neztrácí čas formálním schvalováním svých výzkumných kroků);
- může získat otevřenější a méně zkreslené výpovědi;
- bude mít větší svobodu odmítnout okrajové otázky, které nemíří k jádru výzkumného problému (více výzkumné svobody);
- výzkumník má více předpokladů zachovat si objektivitu (nebrání mu identifikace se systémem).

B) Nevýhody nezávislého výzkumníka:

- dotázaní si mohou udržovat odstup, respondenti si nejsou jisti účelem výzkumu, objeví se nedůvěra, stylizace odpovědí a chování;
- vzhledem k vnější objektivitě výzkumu není umožněno zkoumat intimní a neformální stránky společenského systému;
- nezávislý výzkumník hůře získává podporu od autoritativního systému.

Být nezávislým výzkumníkem má více výhod při výzkumu formálních systémů než neformálních systémů.

7.2.3. Seznámení s respondenty a prostředím

Výzkumník musí nechat na sebe působit běžné zvyky, chování respondentů, uvědomit si, jak myslí, vzájemně komunikují, jaké mají vystupování, vyjadřování, gestikulaci. To mu umožní lépe zvolit strategii sběru dat a plánovat způsoby, jak co nejlépe postihnout sociální prostředí ve výzkumu.

7.2.4. Předvýzkum nebo pilotáž

Úkolem předvýzkumu je ověřit všechny výzkumné nástroje a etapy sběru dat, výběrového šetření i analýzy. Pouhé ověření nástroje (např. dotazníku) (pilotáž, pretest) stačí jen vyjimečně, pokud je výzkumník

zevrubně seznámen se sociálním prostředím a souvislostmi zkoumání. I tehdy je však jeho povinností prověřit výzkumný nástroj a provést nácvik se spolupracovníky.

Předvýzkum je povinnost – pokud se výzkumník předvýzkumu vyhne, měl by výslovně zdůvodnit proč a obhájit předem své rozhodnutí.

Výhody předvýzkumu:

1. poskytuje příležitost konečného vyzkoušení a změny výzkumného nástroje;
2. výzkumník má příležitost seznámit se v úplnosti se sociálním prostředím, v němž bude provádět hlavní šetření;
3. má příležitost předem vyzkoušet, jak bude vypadat analýza sebraných dat, odstranit včas řadu nepředvídaných problémů, měl by vždy provést alespoň dílčí analýzu dat z předvýzkumu;
4. má možnost podívat se, jak se osvědčily zvolené proměnné, zda získaná data budou vhodná pro testování vybraných hypotéz a přijímané teorie;
5. výzkumník má poslední možnost zastavit výzkumné šetření a vynaložit značné finanční náklady efektivněji.

7.2.5. Formulace otázek, stanovení pozorovaných elementárních jevů

Správně položená otázka musí být teoreticky (vědecky) relevantní (významná) a empiricky smysluplná. Musí mít vztah k systému pojmů zkoumaného prostředí i k systému pojmů testované teorie.

Jedním z nejčastějších hříchů akademického výzkumu je kladení otázek, jejichž využitelnost pro řešení některého z položených problémů nebo pro testování některé z hypotéz je neprokazatelná, je ponechána náhodě. Takové otázky zbytečně zatěžují výzkum, znamenají ztrátu času, kterého by bylo možno využít ke zpřesnění důležitější části informace. Druhou často se opakující chybou je nejednoznačnost položených otázek. Části dotázaných jsou otázky nesrozumitelné nebo je chápou v jiném významu, než bylo výzkumníkem míněno. Při volbě elementárních jevů pro pozorování se jedná o obdobný proces. Jeho obtíže však spočívají spíše v interpretaci pozorovaného a v nesnadnosti najít vztah mezi jednotlivými elementárními jevy a pojmy užívanými teorií.

Propojení teorie a jejích pojmů s elementárními fakty zjišťovanými při sběru dat a využití empirických zjištění pro odpovědi

na otázky formulované jazykem teorie tvoří patrně nejobtížnější úkol sociologického výzkumu.

7.2.6. Záznam a kódování informace

Výzkumník rozhoduje o tom, která z technik sběru dat by měla být užita. Záznam informace probíhá v situaci, v níž rozeznáváme výzkumný nástroj, dotázaného a toho, kdo zaznamenává informaci. U některých technik zaznamenává informaci sám dotázaný. Pokud tomu tak není, do vztahu dotázaný – nástroj může rušivě vstoupit pozorovatel, tazatel nebo experimentátor.

Důležitý je trénink, instruktáž spolupracovníků a kontrola záznamu.

7.2.7. Kontrola záznamu, testování validity a spolehlivosti

Výzkumník by měl se spolupracovníky projít – zkontrolovat prvních pár dotazů, pozorování. Chybějící nebo nesprávná data by měl dohledat nebo opravit s respondentem.

Když je sběr skončen, výzkumník by měl prověřit 5% sebraných údajů (dotazníků, pozorování) přímým kontaktem (např. telefonátem dotázanému a pod.). Prověřit by měl ty, kteří odpověděli i ty, kteří neodpověděli.

7.2.8. Organizace a překódování dat

Poslední, osmou, fází sběru dat je organizace a překódování dat do podoby vhodné pro hromadné zpracování dat (s využitím počítače). Někdy sem bývá zařazována také kontrola dat. Doporučuji však kontrolu a čistění dat uvažovat jako zvláštní etapu výzkumného šetření = přípravu dat. Vlastním obsahem organizace a překódování dat je pak převedení dat na záznamové médium počítače, jejich spolehlivé uložení a dokumentace.

7.3. Hlavní techniky sběru dat

Členění metod podle Labovitze a Hagedorna [1976:71-87]:

1. Metody, které získávají údaje z odpovědí dotázaného:
 - dotazníkové šetření (questionnaire)
 - standardizovaný rozhovor (schedule)
 - nestandardizovaný rozhovor (interview guide)
2. Pozorování - pozorovatelem člověk
 - mechanický záznam pozorovaného
3. Zkoumání materiálních výtvorů (stop)

Všechny tyto metody sběru dat mohou získávat data buď
 A) z primárního zdroje nebo
 B) ze sekundárního zdroje

Členění metod podle Nan Lina [1976:201n.n.]:

- 1) metoda pozorování
- 2) dokumentačně historická metoda
- 3) metoda průzkumu (Survey) - metoda dotazníkového šetření
 - metoda rozhovoru
- 4) experimentální metoda

Hlavní techniky sběru dat podle Dismana [1969: 151nn]

- 1) Pozorování
- 2) Rozhovor ___ metody
- 3) Dotazníkové šetření / dotazování
- 4) Studium dokumentů

8. Metody pozorování

8.1. Charakteristika pozorování

- 1) Pozorováním jsou sledovány jevy, které jsou přímo dostupné smyslovému vnímání.
- 2) Pozorováním jsou sledovány jevy, které nebyly vyvolány zásahem výzkumníka přímo v průběhu pozorování.

Z těchto charakteristik výzkumné situace vyplývající malý počet transformací má za následek malé zkruslení výsledků.

8.2. Druhy pozorování

8.2.1. Nezúčastněné a zúčastněné pozorování

Pozorovatel při nezúčastněném pozorování je posuzovatelem, je oddělen od situace (neúčastní se jí) a používá předem zvolený soubor znaků a jejich kategorií, do nichž zařazuje a kóduje jednotlivé elementární jevy - výroky nebo chování.

Při zúčastněném pozorování je pozorovatel součástí sociálního organismu, sociálního prostředí, které sleduje (zkoumá), je účastníkem společenských vazeb, je začleněn do situace a také výsledky jeho pozorování jsou podmíněny způsobem jeho začlenění do sledovaného sociálního prostředí.

8.2.2. Standardizované a nestandardizované pozorování

Při standardizovaném pozorování:

- jsou přesně vypočítány a definovány elementární jevy,
- jsou určeny třídy hodnot, do nichž jsou jevy klasifikovány,
- jsou přesně stanoveny znaky a jejich kategorie,
- výsledky standardizovaného pozorování různých jednotek různými pozorovateli jsou srovnatelné.

Při nestandardizovaném pozorování:

- je určen pouze cíl pozorování,
- nejsou vůbec nebo jen minimálně formalizovány nástroje sběru dat,

- v průběhu sběru dat je umožněno upravit volbu sledovaných, elementárních jevů eventuálně upravit i navrženou hypotézu,
- způsob sběru dat umožňuje použití v přípravné fázi výzkumu,
- jsou kladeny vysoké požadavky na pozorovatele - výzkumníka, jedná se o jednu z nejnáročnějších technik sběru dat.

- pozorujeme elementární jevy přímo dostupné smyslovému vnímání,
- jevy záměrně nevyvoláváme
- kromě použití pozorování v experimentu nesledujeme jevy řídké, těžko předpověditelné
- pozorování může postihnout dynamiku (relativně) rychle se měnících jevů (v dimenzích minut, hodin) v její kontinuitě, ke sledování tohoto typu jevů se pozorování hodí lépe než jiné techniky
- pozorování je nenahraditelné u jevů vázaných na strukturu vztahů ve skupinách

8.2.3. Skryté a zjevné pozorování

Při skrytém pozorování (skutečně fyzicky skrytém nebo sociálně skrytém):

- pozorovatel přijímá roli řadového člena skupiny nebo některou z rolí, které jsou obvykle součástí sledovaného sociálního prostředí
- je pozorovateli odhaleno přirozené sociální prostředí a může sledovat jevy cizímu okolí skryté
- je velkou chybou, když je skrytý pozorovatel odhalen, vede to zpravidla ke ztrátě důvěry k výzkumnému týmu i celému výzkumu a k naprostému znemožnění výzkumné práce v daném terénu na dlouhou dobu.

Při zjevném pozorování:

- pozorovatel neskrývá svou roli sociologa výzkumníka a předpokládá, že si prostředí na jeho výzkumné aktivity přivykne a nebude jimi ovlivněno
- v běžných situacích nemusí působit přítomnost pozorovatele rušivě
- v případech, kdy předmětem sledování jsou společenské jevy související s dodržováním nebo porušováním společenských norem, může být zjevnou přítomností pozorovatele sledovaný jev rušen.

Příklad klasifikace chování při standardizovaném pozorování: Balesův systém kategorií pro klasifikaci chování účastníků diskutující skupiny⁹⁾

Kódujte pozorované chování účastníků diskutující skupiny!

1. projevuje solidaritu	kladné emoce
2. ukazuje uvolnění	- " -
3. souhlasí	- " -
4. dává návrhy	pokusy o řešení
5. vyjadřuje mínění	- " -
6. dává (poskytuje) orientaci	- " -
7. vyžaduje orientaci	otázky
8. ptá se po mínění ostatních	- " -
9. vyžaduje návrhy	- " -
10. nesouhlasí	záporné emoce
11. ukazuje napětí	- " -
12. projevuje antagonismus	- " -

Bales navíc charakterizoval dvojice elementárních pozorovaných interakčních jevů z hlediska jejich zaměření na:

- integraci	1,12
- vyrovnání napětí	2,11
- společné rozhodnutí	3,10
- ovlivňování-kontrolu	4, 9
- hodnocení	5, 8
- orientaci	6, 7

8.3. Obor použití pozorování

8.3.1. Charakter zkoumaných jevů

8.3.2. Charakter výzkumné akce

Nestandardizované pozorování je většinou nenahraditelné pro prvou orientaci v neznámé problematice a pro předvýzkum, pro některé monografické studie a pro rychlé orientační operativní studie.

Pokud chceme vysvětlit význam jevů a sledujeme posloupnost rychle po sobě následujících událostí, používáme pozorování (zvláště zúčastněné). Také pokud se nám jedná o porozumění chování nějaké speciální skupiny, o vysvětlení podstatného sociálního problému v současnosti.

Pozorování obecně se hodí lépe pro akce s menšími nároky na rozsah výběru a tím i na počet pozorovatelů.

8.4. Příprava a zásady standardizovaného pozorování

Do pozorování - jeho přípravy - vstupujeme s podrobnou pracovní hypotézou.

- A) stanovíme elementární jevy dostupné empirickému zkoumání;
- B) připravíme plán pozorování;
- C) připravíme technické pomůcky, terén a pozorovatele;
- D) provedeme pozorování;
tyto fáze opakujeme pro předvýzkum a po jeho zpracování a zhodnocení pro hlavní výzkumnou akci.

8.4.1. Plán pozorování

- A) PŘEDMĚT POZOROVÁNÍ = co má být pozorováno
 - B) ZPŮSOB POZOROVÁNÍ = jakým způsobem budou jevy sledovány
 - C) KÓDOVÁNÍ POZOROVANÝCH JEVU = do jakých kategorií budou pozorované jevy zařazovány
 - D) ZÁZNAM POZOROVÁNÍ = jak bude prováděn záznam pozorovaného
- A) Stanovení předmětu pozorování
- a) stanovíme elementární jevy, tj. dílčí sledovatelné skutečnosti, které mají podobu dílčích úseků činnosti nebo výskytu pozorovatelných faktů; pro posouzení významu sledované skutečnosti je třeba znát širší kontext sledovaného jevu;
 - b) stanovíme v případě potřeby indikátory elementárních jevů, tj. sledovatelné skutečnosti, fakta činnosti, které vypovídají o sledovaném elementárním jevu nepřímým jako jeho ukazatele; právě tak jako elementární jev, musí být i jeho indikátory definovány

- zcela jednoznačně, aby bylo pro každého pozorovatele rozhodnutelné, zda jev v daném případě nastal nebo nenastal;
 - c) určíme jedince (event. jiné pozorovatelné jednotky), na nichž bude elementární jev sledován; pokud je do akce zapojeno více pozorovatelů v témže terénu, je třeba přesně (jednoznačně) rozdělit = určit, kterou část sledovaných jednotek kdo sleduje (např. řada 2,3,4);
 - d) stanovíme čas, ve kterém se bude pozorování realizovat, tj.
 - období, ve kterém se výzkumná akce bude konat a
 - momenty, ve kterých bude pozorovaný jev zaznamenáván, např. každých 5 minut nebo v situaci výskytu jevu.
- Každou nečekanou zkreslující událost je třeba bezpodmínečně zaznamenat - tj. také pro její záznam vyčlenit zvláštní rubriku.

B) Způsob pozorování

- a) určíme, zda budeme provádět skryté nebo zjevné pozorování;
- b) při zjevném pozorování stanovíme roli pozorovatele;
- c) při skrytém pozorování stanovíme, jak se bude ukrytí realizovat:
 - jako - technické maskování
 - společenské maskování.

C) Kódování pozorovaných jevů

Kódujeme způsobem:

- a) jev nastal - jev nenastal
 - b) kódy = kategorie znaku vyjadřují stavy, kterých může jev nabýt.
- Kódování do kategorií zjednodušuje záznam.
 - Je třeba kódovat všechny možné stavy.
 - Kategorie je třeba definovat jednoznačně, ze stejného hlediska.
 - Definujeme nejen kategorie znaku ale také klíč - návod, který usnadní správné a jednoznačné kódování.

D) Záznam pozorování

Z průběhu pozorování zaznamenáváme:

- a) zpřesňující informace (o předmětu pozorování)
- b) místo a čas pozorování (dobu začátku a konce pozorování)
- c) kdo provedl pozorování a záznam
- d) výskyt náhodných (neočekávaných) faktorů
- e) připomínky k interpretaci, metodické připomínky

- f) vlastní informace o výsledcích pozorování
- slovní popis jevů
 - stenografický záznam
 - mechanický záznam obrazu (zvuku) (foto, film, video, magnetofon)

Pro záznam pozorování používáme pozorovací arch nebo registrační přístroje.

8.5 Výhody a nevýhody zúčastněného pozorování

8.5.1. Problémy a nevýhody zúčastněného pozorování

- a) nedostatek spolehlivosti (reliability) vyplývající z nahodilosti pozorování uskutečněných výzkumníkem - není pravděpodobné, že by dva výzkumníci zjistili a zaznamenali stejné události stejným způsobem, projevuje se "výběrové vnímání";
- b) zúčastněný pozorovatel může ovlivnit sledovaný subjekt svou přítomností, dochází ke změně chování a tím ke změně obrazu;
- c) současná role omezuje možnosti pozorovatele: např. řadový účetní v bance, jehož roli výzkumník přijal, stěží může získávat osobní informace od bankovního ředitele;
- d) pozorovatel svým začleněním do skupiny ztrácí objektivitu;
- e) většina pozorovatelů musí pasivně čekat až nastane událost, kterou se rozhodli pozorovat - je možné, že událost, o jejíž sledování nám jde nejvíce, v době pozorování nenastane.

8.5.2. Výhody zúčastněného pozorování

- a) pozorování se koná v "přirozených" podmínkách;
- b) pozorovatel může sledovat emocionální reakce sledovaného subjektu, což může vést k užitečným hypotézám;
- c) pozorováním v dlouhém časovém rozpětí lze získat velké množství informací;
- d) pozorovatel může zaznamenat kontext, který dává význam respondentovu vyjádření názorů a hodnot;
- e) pozorovatel může navázat užší kontakty se zkoumaným prostředím a může tedy klást "důvěrné" otázky, které by jinak u tazatele zvnějšku nepřipadaly v úvahu.

9. Metody dotazování

Mezi metody dotazování řadíme rozhovor a dotazníkové šetření. Obě skupiny metod mají řadu příbuzných rysů z hlediska přípravy výzkumného nástroje i nároků na jeho použití. Nejvíce společných vlastností nalézáme mezi standardizovaným rozhovorem s dotazníkovým archem a dotazníkovým šetřením, při němž je dotazník osobně předáván a sbírán. Definitivní rozhodnutí o použití jedné z těchto metod provádíme často až na základě předvýzkumu.

9.1. Rozhovor

9.1.1. Standardizovaný rozhovor s dotazníkovým archem

- je to nejvíce formalizovaná forma rozhovoru;
- od každé osoby je vyžadována stejná informace;
- každému jsou kladeny přesně stejně formulované otázky;
- otázky jsou v přesně stejném pořadí.

Standardizované interview (rozhovor) je nejvhodnější použít:

- když je dotazován velký počet osob;
- když je soubor dotazovaných homogenní z hlediska životního stylu a filozofie, náhledu a zkušeností (kultury);
- když opravdu víme, co je důležité se ptát a jak se na to ptát.

9.1.2. Rozhovor (částečně standardizovaný) se soupisem zjišťovaných faktů

- je založen na předpokladu, že je možno získat stejnou informaci od různých skupin dotázaných rozdílně formulovanými otázkami kladenými navíc v některých případech i v rozdílném pořadí.

E.Kaneová [Kane 1985:63-64] uvádí jako příklad této techniky zjišťování postojů k užívání drog u různých skupin obyvatelstva:

- u legálních uživatelů drog,
- u lékařů,
- u dozorcích úředníků a
- u prostého obyvatelstva.

Každá z těchto skupin používá jiných pojmů a termínů; posuzuje drogy z jiného úhlu pohledu:

- policisté z hlediska toho, co je legální/ ilegální;
- lékaři uvažují medicínské charakteristiky drogy;
- uživatelé posuzují způsob užívání a cenu drog.

Při této technice rozhovoru vznikají větší problémy s kódováním ve srovnání se standardizovaným rozhovorem. Je třeba ověřit, že se u různých skupin dotázaných jedná o stejné skutečnosti pouze vyjádřené jiným způsobem, je třeba zajistit srovnatelnost okrajových kategorií sledovaných znaků.

9.1.3. Nestrukturovaný rozhovor

- není předem určena struktura rozhovoru, znění ani pořadí otázek, výzkumník se nesnaží získat stejnou informaci od každé dotazované osoby;
- pro každou další otázku je používáno pojmů a členění dotázaného;
- nestrukturovaný rozhovor je často užíván na počátku výzkumného projektu;
- nestrukturovaný rozhovor je vhodné užít při zkoumání osob z odlišného kulturního prostředí s výrazně odlišnou hodnotovou strukturou.

Příklad: (podle Kane [1985:64-65])

Zjišťujeme vztahy sester k pacientům v psychiatrické léčebně. Náhodného setkání se zdravotní sestrou využijeme k nestrukturovanému rozhovoru. Jedním z jeho cílů bude zjistit, jak sestry spontánně podle svých hledisek rozlišují pacienty, jaké typy pacientů rozlišují.

Rozlišujeme:

1. přístup klasifikace zevnějšku - umožňuje porovnání s jinými zcela odlišnými skupinami (mezikulturní srovnání), kategorie tvoří výzkumník;
2. přístup klasifikace zevnitř - charakterizuje, jak skupina sama vidí a rozčleňuje svět; popis chování a vztahů zevnitř závisí na odhalení přirozených kategorií činnosti; přístup lze uplatnit i v rámci vlastní kultury a společnosti.

Příklad:

Výzkum studentů univerzity:

A) zvnějšku je možno kategorizovat:

- a) studenti 1. ročníku, 2. ročníku, 3. ročníku,
- b) studenti oborů humanitních/ přírodovědných,
- c) studenti denního/ studenti dálkového studia.
- lze srovnat mezi univerzitami

B) zvnitřku: studenti byli požádáni, aby řekli, do jakých typů nebo skupin by rozdělili studenty své univerzity.

E.Kaneová uvádí těchto pět skupin, které byly jmenovány:

- 1) lidé, kteří nikdy nezvednou nos od knihy,
- 2) ti, kteří stále dělají potíže
- 3) slušní lidé,
- 4) velcí světáci
- 5) cizí studenti
- lze lépe porozumět vnímání studentského světa samotnými studenty.

9.1.4. Pravidla rozhovoru

Rozhovor ovlivňuje celá řada okolností, z nichž některé může výzkumník ovlivnit:

- věk, - pohlaví, - sociální postavení, - způsob vystupování tazatele, - jeho hlas, - gesta, - osobní charakteristiky;
- v zásadě platí, čím je tazatel blíže dotázanému v sociálním postavení, věku a zájmech, tím větší má šanci uspět v rozhovoru.

Na počátku rozhovoru:

1. tazatel se představí;
2. uvede název organizace nebo skupiny, kterou reprezentuje;
3. stručně (dvěma větami) vyjádří, čeho se výzkum týká, aby dotázaný pochopil význam výzkumu obecně a pokud možno i význam pro něj;
4. sdělí dotázanému, jak byl vybrán, stručně vysvětlí výběrovou metodu;
5. ujistí dotázaného, že jím uvedené údaje jsou důvěrné a že nebude možno je použít v souvislosti s jeho jménem, že nikde nebude jeho jméno uvedeno;

Vlastnosti tazatele (tazatelky) :
(podle Moser-Kalton [1971:285-6])

1. čestný (pocitivý) a svědomitý (až puntičkářsky);
2. mít zájem o práci a o zkoumané otázky;
3. být pečlivý a přesný;
4. být přizpůsobivý okolnostem;
5. osobnost ani příliš agresivní ani příliš společenská (příjemné, živé a přitom obchodní vystupování je ideální);
6. inteligence a vzdělání ne mimořádné - je třeba, aby inteligence byla dostatečné k pochopení a vyplnění složitých instrukcí a adaptaci na situaci a respondenta;
7. ve věku zpravidla 25-45 let, osoby mimo toto věkové rozpětí bývají častěji odmítány;
8. ženy mívají přednost z řady důvodů: - jsou ochotny pracovat večer a o víkendech, výhodou je jejich osobní vzhled, nejsou nikdy neupravené a jsou vždy elegantní;
9. tazatelé ze střední společenské vrstvy mají výhodu v možnosti přístupu k dotázaným z nižších i vyšších společenských vrstev;
10. je třeba mít zřetelné vyjadřování (někdy je výhodou i krajový přízvuk).

Tazatelé jsou podrobováni dosti přísnému výběru a kontrole a procházejí speciálním tréninkem.

9.2. Dotazník jako nástroj sběru dat

9.2.1. Příprava dotazníku

- nejprve na základě přípravné fáze výzkumu sestavíme seznam toho, co chceme vědět;
- připravíme první formulaci otázek (E.Kaneová [1985:80nn] radí používat systém dvou druhů kartiček):
 - 1.druh - napíšeme, co chceme zjistit;
 - 2.druh - formulujeme otázku;
- otázky uspořádáme do bloků podle tématických okruhů;
- upravíme hladké přechody mezi bloky otázek;
- doplníme úvodní oslovení dotázaného a identifikační otázky;
- provedeme pilotáž dotazníku - ověřujeme formulaci otázek, jejich srozumitelnost, kategorie odpovědí.

9.2.2. Pravidla uspořádání dotazníku

(podle Kaneové [1985:74-76])

1. první stránka obsahuje údaje administrativní povahy, je zde umístěno mimo jiné číslo dotazníku, které odpovídá jménu v seznamu, který není volně přístupný, jeho obsah je důvěrný;
2. pořadí otázek:
 - a) nejprve zařazujeme širší snadné otázky neosobní povahy, úvod dotazníku by měl respondentu zaujmout, chceme ho přimět ke spolupráci;
 - b) ve střední části dotazníku zařazujeme méně zajímavé otázky, mezi tématickými bloky řadíme spojovací otázky;
 - c) závěrečná část dotazníku může obsahovat citlivé nebo osobní dotazy, otevřené otázky;
 - d) zcela na závěr (pokud je to nezbytné) řadíme dotazy typu "zeptej se a uteč", snižujeme tím riziko vyššího procenta nedokončených rozhovorů nebo nevyplněných dotazníků.
3. výběr odpovědí:
 - a) uzavřené otázky (dotázaný musí zvolit jednu z nabízených odpovědí);
 - b) otevřené otázky (dotázaný použije své vlastní formulace);
 - c) polouzavřené otázky (jsou nabídnuty kategorie odpovědí a mimo to je poskytnuta možnost odpovědět volně mimo);
4. filtry - jsou zvláštním typem otázky, který rozděluje dotázané na ty, kterým budou položeny následující otázky a ostatní, kteří na ně nebudou odpovídat nebo budou odpovídat na jiné, jinak formulované.

9.2.3. Formulace otázek

Existuje velké množství běžných nedostatků ve formulaci otázek. Nejčastější typy nevhodných otázek podle E.Kaneové [1985:78,79] jsou tyto:

1. Dvojitá otázka
Př.: "Jste šťasten ve svém manželství a práci?"
2. Chybná nabídka odpovědí
Př.: "Vaše vlasy jsou žluté, purpurové, zelené nebo modré?"
3. Zjišťuje všechno najednou
Př.: "Vyjmenujte prosím všechna místa, kde jste v posledních pěti letech pracoval, zastávané profese a funkce v nich, váš plat a proč jste odešel!"

4. Mlhavá otázka
Př.: "Chodíte často tančit?"
5. Všeobjímající otázka
Př.: "Co si myslíte o Středním Východě?"
6. Otázka v žargónu (hantýrce)
Př.: "Zdá se vám, že váš manžel má sebe-aktualizovanou autonomní strukturu osobnosti?"
7. Hypotetická otázka
Př.: "Jaký druh vzdělání byste chtěla aby mělo vaše dítě?"
8. Zavádějící otázka
Př.: "Proč jste šťastný jako voják základní služby?"
9. Otázka o názoru někoho jiného
Př.: "Myslíte si, že vaše rodiče potěšila výstavba nové školy?"
10. Otázka s výpadkem (předpokládá skutečnost, která však není univerzální)
Př.: "1.Kterému investičnímu fondu jste svěřil své kupóny?
2.Využijete možnosti odprodeje svých podílů fondu po roce za desetinásobek?"

9.3. Dotazník zasílaný poštou (Mailed Questionnaire)(Postal Questionnaire)

9.3.1. Výhody [Moser-Kalton 1971:257-9]

1. U populací řídké rozptýlených ve velkých oblastech je (při omezených prostředcích) dotazník zasílaný poštou jediným dostupným způsobem sběru dat.
2. Zasílání dotazníku poštou je ekonomicky výhodné. Neplatí to však obecně. Při velmi nízké návratnosti průměrné náklady na jeden dotazník rostou.
3. Rychlá metoda.
4. Odstraňuje nižší přesnost způsobenou vynuceným použitím skupinkového nebo víceúrovňového náhodného výběru při rozhovorech. U poštovního dotazníku není třeba aby dotázaní byli vybíráni jen z některých území nebo oblastí.
5. Někdy je výhodou ponechání možnosti dotázanému konzultovat faktické zjišťované údaje v rámci rodiny a pod.
6. Možnost podívat se do dokumentů.

7. Někteří lidé raději a přesněji odpovídají, když nejsou tváří v tvář neznámému tazateli.

- podle možností připojit dotaz, zda chce dotázaný obdržet výsledky výzkumu.

9.3.2. Nevýhody (Omezení) [Moser-Kalton 1971:260-261]

1. Otázky musí být snadné, relativně prosté, jednoduché.
2. Odpovědi je třeba považovat za konečné.
3. Není vhodné pro otázky, na něž chcí spontánní odpověď.
4. Jednotlivé odpovědi nelze považovat za nezávislé.
5. Výzkumník si nemůže být jist, že dotazník vyplnila správná, tj. vybraná osoba.
6. Odpověď dotázaného není možno doplnit o výsledek pozorování.

9.3.3. Fáze dotazování poštou (podle E.Kaneové [1985:86-87])

1. Zašleme dopis se žádostí o spolupráci, vysvětlením účelu výzkumu spolu s dotazníkem a oznámkovanou obálkou pro odpověď.
2. Když začíná slábnout návrat prvních vyplněných dotazníků, je třeba poslat první upomínací dopis těm, kteří neodpověděli.
3. Po uplynutí doby potřebné na odpověď (týdne, deseti dnů) od zaslání prvního upomínacího dopisu je třeba zaslat druhý upomínací dopis spolu s kopií dotazníku a jinou oznámkovanou obálkou pro návrat dotazníku těm, kteří ještě neodpověděli.
4. Všem, kteří odpověděli, poslat děkovný dopis.

9.3.4. Průvodní dopis k dotazníku

Průvodní dopis k dotazníku zasílanému poštou by měl obsahovat:

- účel výzkumu;
- označení instituce nebo skupiny, která výzkum organizuje a financuje;
- údaj o tom, jak byl dotázaný vybrán;
- výzva ke spolupráci;
- příslib důvěrného zacházení s údaji (pokud je třeba);
- pokyny, jak vyplnit dotazník;

9.3.5. Návratnost při poštovním dotazování

Pro většinu poštovních dotazníků platí, že mají mnohem nižší návratnost než jaká je u rozhovorů. Ti, kteří odpověděli, se pravděpodobně více nebo méně liší od těch, kteří neodpověděli. Následkem toho vzniká chyba, v podstatě systematické vychýlení výsledků od hodnot očekávaných v celé uvažované populaci.

Pro obecnou populaci je obvyklá návratnost při poštovním dotazování 30-40%. Někdy však dosahuje jen např. 10%. Velmi zřídka je možno v tématu, které je dotázaným blízké, dosáhnout až 90% návratnosti.

Na dotazník neodpovídají nejčastěji lidé s nižším vzděláním z méně kvalifikovaných profesí.

9.3.6. Faktory ovlivňující návratnost

Faktory, které nelze ovlivnit:

- výzkumná instituce
- zkoumaná populace
- předmět výzkumu

Faktory, které lze ovlivnit:

- délka dotazníku: nadměrně dlouhý dotazník snižuje návratnost;
- pilotáž: na základě pilotáže lze vyloučit nesrozumitelné a neobratně formulované otázky;
- dobře formulovaný průvodní dopis by měl zbavit obav respondenta a motivovat ho k vyplnění dotazníku;
- zvýšení návratnosti lze dosáhnout také tím, že na upomínací dopis doplníme ručně psanou poznámku a tím snížíme neosobnost kontaktů s dotázaným.

10. Studium dokumentů (Sekundární analýza dat)

10.1. Typy dokumentů:

1. osobní dokumenty vytvořené z iniciativy autora:
zpravidla vypovídá o svých zážitcích, názorech, postojích a pod.;
náleží sem dopisy, diáře, autobiografie, zprávy před sebevraždou,
přiznání nebo doznání;
2. úřední dokumenty, které vznikly se záměrem shromáždit určitý
soubor údajů; např. statistiky, úřední výkazy, soudní akta;
3. tištěné a další materiály hromadného předávání informací (noviny,
časopisy, knihy, fotografie, filmy, videozáznamy a některé
výtvořiny materiální kultury);

Všechny dokumenty je možno dále rozdělit podle toho, kdo je zaznamenal na:

- A) původní (které zaznamenal očitý svědek);
- B) druhotné (které zaznamenal někdo jiný, kdo mluvil se svědkem, četl jeho původní záznam a pod.).

10.2. Operace s dokumenty

Při využívání dokumentů pro sociologický výzkum rozeznáváme tři skupiny operací (podle Dismana [1969:325] a podle Baileyho [1978:273]):

1. výběr dokumentů a získání přístupu k nim;
2. kritické zhodnocení dokumentů z hlediska využitelnosti pro zkoumání uvažovaného sociologického problému;
3. kódování a analýza z hlediska potřeb výzkumu.

10.3. Obory použití studia dokumentů

- přímé informace jsou nedostupné
 - fyzicky (z důvodů časových, územních a p.)
 - politicky
 - finančně
- studium dokumentů umožňuje analýzu delšího časového období
- je organizačně méně náročné a méně nákladné než ostatní techniky

10.3.1. Přednosti studia dokumentů

1. Je možno studovat i výpovědi jinak nedostupných subjektů (např. již dávno zemřeli).
2. Výpověď není reakcí na dotaz výzkumníka. Není tedy tímto dotazem vyvolána ani zkreslena.
3. Studium dokumentů umožňuje longitudinální analýzu (např. analýza tisku za období několika desítek let).
4. Velikost výběru (počet studovaných dokumentů) může být větší než u pozorování nebo experimentu, rozsah výběru jako u rozhovoru nebo dotazníkového šetření.
5. Záznam byl proveden spontánně.
6. Tvůrce dokumentu snáze uvede doznání (přiznání).
7. Studium dokumentů je relativně levnější než terénní průzkum prováděný formou rozhovoru nebo dotazníkového šetření.
8. Některé dokumenty vytvořené profesionály (např. novináři) mají vysokou kvalitu záznamu.

10.3.2. Nevýhody studia dokumentů

1. Účel, pro nějž byl záznam pořízen, jej modifikuje. Vzniká vychýlení oproti ideálnímu stavu zjišťování.
2. Dokumenty nemají stejnou životnost. Dochází k ztrátám, jejichž rozsah se liší podle zdroje dokumentů.
3. Výzkumníkovi chybí informace potřebné k porozumění dokumentu. Jeho autor nepředpokládá využívání dokumentu výzkumníkem a údaje jsou neúplné.
4. Některé dokumenty jsou nedostupné.
5. Vzniká výběrové vychýlení způsobené skutečností, že méně vzdělaní lidé méně často tvoří psané dokumenty.
6. Nelze sledovat neverbální chování.
7. Nestejný formát dokumentů. Informace nejsou zaznamenávány stejným způsobem a ve stejném čase.
8. Vznikají potíže s kódováním. Za účelem standardizace byla vyvinuta speciální kvantitativní metoda analýzy dokumentů - obsahová analýza.
9. Srovnatelnost údajů v čase je omezená.

10.4. Sekundární analýza

10.4.1. Zdroje dat pro sekundární analýzu

- archivní data
- oficiální statistická data
- databanky statistických údajů
- datové publikace z výzkumů
- datové soubory z výzkumů na počítači

10.4.2. Výhody sekundární analýzy

1. Ušetří čas a peníze.
2. Méně zasahuje do soukromí než např. dotazovací techniky.
3. Je snazší provádět srovnávací analýzy:
 - mezinárodní srovnání;
 - analýzy trendů;
 - časové srovnání obdobných jevů ve dvou nebo několika okamžicích.

10.4.3. Nevýhody sekundární analýzy

1. Některá potřebná data jsou nedostupná a nelze je dodatečně doplnit.
2. Původní data mohou obsahovat chyby, které výzkumník provádějící sekundární analýzu není schopen odhalit.

10.5. Analýza osobních dokumentů

Jedná se o techniku podobnou zúčastněnému pozorování. Má výhodu spontánnosti záznamu informace. Odráží velmi často hloubku vcítění tvůrce osobního dokumentu do situace. V protikladu k strukturované a neosobní obsahové analýze představuje analýza osobních dokumentů metodu nestrukturovanou a osobní.

10.5.1. Osobní dokumenty jako zdroj dat

Jak vyplývá z názvu, osobní dokumenty mají často osobní až soukromou hodnotu a jsou většinou důvěrné povahy. Je tedy obtížné zajistit si jejich dostupnost. Extrémně obtížné je sebrat jich velký soubor. Musíte prokázat, že jste profesionální sociolog a udat důvěryhodný účel analýzy. Měl(a) byste slíbit výsledky výzkumu. Samozřejmý je slib uchovat dokumenty jako důvěrné a z nich získaná data užívat v souladu se zásadami výzkumnické etiky.

10.5.2. Klasické studie osobních dokumentů

Thomas - Znaniecki: Polský sedlák v Evropě a v Americe.

= nejslavnější studie osobních dokumentů v sociologickém výzkumu. ¹⁰⁾

Je založena na velkém počtu osobních dopisů psaných polskými přistěhovalci v Americe a jejich příbuznými v Polsku. Autoři charakterizují nejčastěji vyměňované listy mezi emigranty a jejich rodinami v Polsku jako "pozdravné dopisy" a provádějí jejich typologii na:

1. ceremoniální, tj. nahrazující neúčast člena širší rodiny na svatbě, pohřbu nebo obdobné příležitosti, kdy se obvykle schází celá rodina;
2. informativní, tj. podávající podrobnou informaci o životě chybějícího člena rodiny;
3. citové, které mají posílit solidaritu rodiny;
4. literární, často psané ve verších, které složí k přečtení nebo přednesu při slavnostních rodinných příležitostech a nahrazují vystoupení váženého člena rodiny;
5. obchodní.

Práce Thomase a Znanieckého obsahuje ve dvou svazcích dopisy a komentáře k nim, ukazující důležitost dopisů pro život rodin polských sedláků a význam rozboru těchto dopisů pro studium struktury rodin polských vystěhovalců i jejich příbuzných v Polsku. Práce je příkladem využití kvalitativního přístupu v sociologickém výzkumu a zároveň ukazuje možnosti případové studie z hlediska hloubky vzhledu do společenské reality významné skupiny obyvatel.

10.6. Obsahová analýza

Hlavním cílem obsahové analýzy je využít pro analýzu verbální (psaný) nekvantitativní dokument a převést (transformovat) jej na kvantitativní údaje.

Definice (Berelson): ¹¹⁾

Obsahová analýza je výzkumná technika pro objektivní, systematický a kvantitativní popis sledovatelného (manifestního) obsahu komunikace. ("jakéhokoli symbolického chování")

Obsahová analýza je:

strukturovaná analyticko-dokumentační technika, v níž výzkumník

1. nejprve vytváří soubor vzájemně se vylučujících a vyčerpávajících kategorií, které mohou být použity pro rozbor dokumentů a
2. pak zaznamenává četnosti výskytu jednotlivých kategorií v dokumentu.

10.6.1. Postup obsahové analýzy

1. Výběr dokumentů
2. Volba kategorií
3. Volba záznamové jednotky
 - slovo (nejmenší jednotka)
 - téma (námet)
 - charakter (např. sociální status, etnická příslušnost)
 - věta nebo paragraf
 - položka (celá jednotka)
4. Volba kontextové jednotky (větší jednotky, která zahrnuje záznamovou jednotku a pomáhá jednoznačně určit význam v němž je užitá)
5. Stanovení systému kvantifikace
 - výskyt (kódujeme jen ano nebo ne)
 - frekvence (kolikrát se jednotka vyskytla)
 - rozsah (plocha)
 - síla, intenzita

Příklad pro volbu kategorií:

Berelson a Salter ¹²⁾ kategorizují hodnoty vyjadřované v populárním příběhu do dvou skupin podle jejich orientace:

Apel na city (na "srdce")
Heart goals

1. romantická láska
2. sňatek
3. idealismus
4. emoční bezpečí
5. patriotismus
6. dobrodružství

Apel na rozum (na "hlavu")
Head goals

1. řešení konkrétního problému
2. sebeprosazování
3. peníze a materiální statky
4. ekonomické a sociální jistoty
5. moc a dominance

11. Dotazování s využitím počítače

Výpočetní technika našla své využití v sociologickém výzkumu v masovějším měřítku až v šedesátých letech. Použití se však dlouhou dobu omezovalo na zpracování dat. Proto také ve sféře sběru dat se počítače nejprve uplatnily při pořizování dat, předstupni strojně početního zpracování. Teprve s nástupem osobních počítačů se počalo plně rozvíjet odvětví sociologického výzkumu, které dnes již bez rozpaků můžeme nazvat počítačové dotazování.

11.1. CADI - pořizování dat s využitím počítače

Computer Assisted Data Input, doslova vstup dat s využitím počítače, je soubor operací, metoda využití počítače k přenosu údajů co nejkratší cestou od jejich zdroje do paměti počítače.

Tento postup nahrazuje klasické pořizování dat, při němž existuje primární zdroj dat, nejčastěji dotazník vyplněný dotázaným nebo tazatelem. Údaje z dotazníku kódovač přenesl do děrovacího formuláře, sloužícího jako podklad pro děrování. Každý řádek děrovacího formuláře vyděruje děrovačka do jednoho osmdesáti- nebo devadesáti- sloupcového štítku. Po ní následuje kontrola děrování - tzv. přezkoušení. Údaje z formuláře jsou děrovány podruhé, avšak ne fyzicky, pouze pro kontrolu, a zjištěné rozdíly jsou podkladem pro opravu chybného děrování.

Tento složitý a značně chybový postup je při pořizování dat s využitím počítače značně zjednodušen. Snahou je jednak ušetřit lidskou administrativní práci, jednak zkrátit řetěz proměn, jimiž data procházejí a především snížit v důsledku toho množství chyb, které dlouhý řetěz operací do dat vnáší.

Nejprostší zjednodušení fáze pořízení dat spočívá v nahrazení operace děrování děrných štítků a jejich přezkoušení přímým vstupem dat klávesnicí do počítače nebo terminálu. Tato změna v zacházení s daty bývá většinou doplněna naprogramováním automatických kontrol přípustných kódů pro jednotlivé znaky a jejich skupiny (baterie). Tak dochází jednak ke snížení počtu operací s daty, jednak ke spojení přenosu dat s jejich kontrolou. Snižuje se pracnost této etapy prací. Práce se zrychlují a zbývá více času na podrobnější kontrolu a opravy podle původního datového zdroje. Snižuje se množství chyb v datech, především je méně chyb sekundárních, tj. těch, které při klasické

práci s daty vznikají dodatečně při prepisech, děrování a přezkoušení dat.

Větší změnu přineslo zavedení čtecích zařízení. Jejich prvotní podoba vycházela z optického snímání záznamu na speciálních štítcích. Tužkou bylo začerněno okénko symbolizující číselný kód nebo skupina okének symbolizující písmeno. Náročnější snímání napsaných číslic nebo písmen vyžadovalo jejich vyplnění černou tužkou předepsaným typem písma do speciálního čtecího formuláře. Snímací zařízení však hlásí chybu při každé větší odchylce od vzoru a bezchybné čtení dat je na těchto strojích spíše výjimkou. Vyžaduje vyškolené pracovníky pro přípravu čtecích formulářů. Nejnovější způsob čtení vstupních dat snímacím zařízením, a jistě ne poslední, je použití tzv. scannerů, tj. přidavných zařízení počítače vybavených programem pro rozpoznávání obrazců. Scanner snímá tištěné i psané písmo a převádí je přímo do paměti počítače. Umožňuje využití nejrůznějších formátů vstupního formuláře. Scanner jsou buď ruční, snímající menší plochu nebo celostránkové, snímající a přenášející celé obrazové strany graficky věrně a při použití vhodných programů i bez chyb.

11.2. CATI - telefonní počítačové dotazování

V osmdesátých letech jsou vyvíjeny počítačové programy slučující výhody textového editoru, kterým je připravován dotazovací formulář, a programu pro vstup a kontrolu dat. Program slouží pro přípravu dotazování, t.j. dotazníku i celého průběhu dotazovací situace včetně všech variant, filtrů a kontrol. Připravený program užívá tazatel k dotazování telefonem. Tazatel usedne před počítač, vytočí telefonní číslo dotázaného a podle monitoru počítače klade otázky a zaznamenává varianty odpovědí. Předem naprogramované kontroly nepřekročitelných minimálních a maximálních hodnot jednotlivých znaků i logické kontroly vazeb mezi hodnotami jednotlivých znaků umožňují upozornit na chybu v okamžiku, kdy ji lze napravit. Kontrolní otázkou se lze vrátit ke sporné odpovědi, výsledný záznam splňuje ihned po skončení sběru dat podmínky kontrol. Navíc je tazatel programem veden i složitými bludišti filtračních otázek, které by při klasickém dotazování nebylo možno pro složitost dotazníku vůbec použít. Tento způsob dotazování je tedy rychlejší, způsobuje méně chyb a umožňuje složitější strukturu dotazovacího plánu než klasická technika.

Nevýhoda telefonického počítačového dotazování spočívá v omezené dostupnosti značné části potenciálních respondentů. V naší populaci není možno na bázi souboru majitelů telefonních koncesí vytvořit

reprezentativní soubor obyvatel ČSFR ani jeho regionů. Již dnes však lze dotazovat firmy nebo organizace, úřady různých oblastí státní správy a pod. V úvahu je také třeba vzít stále rostoucí síť telefonních nahrávačů. Slouží nejen jako náhrada nepřítomného telefonního účastníka ale také jako informační filtr. Volanému vybranému účastníkovi je třeba zanechat vzkaz, který jej bude motivovat k účasti na výzkumu. Nepodaří-li se nám jej tímto dvouminutovým vzkazem přesvědčit, je velmi pravděpodobné, že s ním pro výzkum nemůžeme počítat.

Ještě nedávno používané skeptické argumenty považující za podstatnou překážku většího rozšíření telefonického dotazování s využitím počítače v našich podmínkách vysokou cenu počítačů a programů jsou téměř bezpředmětné. Výzkumné firmy jsou počítači relativně dobře vybaveny a programy dnes již také nejsou rozhodujícím problémem. Tím však ještě nějakou dobu zůstane nereprezentativnost telefonní sítě.

11.3. CAPI - osobní počítačové dotazování

Nevýhodu počítačového dotazování telefonem spočívající v nerovnoměrném zastoupení různých skupin obyvatelstva mezi majiteli telefonního přístroje zcela odstraňuje CAPI (Computer Assisted Personal Interviewing) - dotazování s přenosným počítačem (osobní počítačové dotazování). Struktura rozhovoru, všechny otázky i větve dotazování, kontroly a filtry jsou nahrány v paměti přenosného počítače nebo na jeho disketě. Tazatel navštíví dotázaného v bytě nebo kdekoli jinde, nezávisle na telefonní síti, a záznam rozhovoru provede na místě do přenosného kufříkového počítače (laptopu).

Je-li systémem osobního počítačového dotazování prováděn rozsáhlý výzkum, na němž spolupracuje mnoho tazatelů, je třeba je všechny vybavit přenosnými počítači. Je pravda, že jsou v současné době ještě dražší než stolní osobní počítače. Další náklady vyžaduje programové vybavení pro tuto metodu sběru dat. Jako vhodný program přichází v úvahu v Nizozemí vyvinutý BLAISE.¹³⁾ Počítačové dotazování je tedy nákladné.

Na druhé straně však tato technika umožňuje dosáhnout nové kvality ve výzkumu. Jsou výzkumné akce, jejichž výsledky je potřeba znát krátce po sběru dat. Osobní počítačové dotazování minimalizuje časové ztráty a zdá se být nejvhodnější metodou pro tyto potřeby. Sebraná zkontrolovaná data mohou být tentýž den jako datová matice spolehlivě přenesena telefonem nebo elektronickou poštou do ústředí k

centrálnímu zpracování. Analytické požadavky je možno znát předem, a proto základní výsledky šetření mohou být známy, při perfektní organizaci, tentýž nebo nejdéle následující den. Klasická technika sociologického dotazování není této rychlosti při zajištění stejné spolehlivosti schopna.

Přesto však metody počítačového dotazování mají jistá omezení. Nejsou vhodné pro situace s vysokou neurčitostí, pro zkoumání nových jevů, pro výzkumné situace, v nichž je třeba použít nestandardizovaných výzkumných technik. Doménou jejich použití je standardizovaný rozhovor, opakovaně prováděná šetření s důrazem na rychlý sběr a následné vyhodnocení dat.

11.4. Budoucnost - CAMI - dotazování elektronickou poštou

Computer assisted mailing interviewing (CAMI) - počítačové poštovní dotazování je předpokládám věcí nedaleké budoucnosti. Jeho předpokladem je vybavenost dotazovaných osobním počítačem. To je sice pro širokou populaci obyvatelstva opravdu věcí vzdálené budoucnosti, alespoň v naší zemi a viděno optikou generace dnešních čtyřicátníků. Optimistický pohled mladé generace nebo pohled za oceán však svědčí pro menší skepsi.

Již dnes je možno si představit počítačové poštovní dotazování na speciální populaci prodejců počítačů a výpočetních programů, manažerů velkých firem, výzkumných ústavů, výzkumných laboratoří univerzit, významných klinik, center kupónové privatizace atp. Program s pokyny pro jeho instalaci a obsluhu včetně dotazníku a prosby o odpověď by mohl být zaslán na počítačové disketě (dnes) nebo elektronickou poštou (v cizině běžně již dnes, u nás v nedaleké budoucnosti) a vyplněný vrácen stejnou cestou. Jistěže by návratnost ve srovnání s rozhovory poklesla, náklady by však výrazně klesly. Při klasickém poštovním dotazování a osobních rozhovorech platí přeci obdobná relace v návratnosti a nákladech. Ostatní charakteristiky počítačového dotazování by při této obměně metody platily také. Rozdíl je možno očekávat v rychlosti sběru dat. Poštou dopravované diskety by akcí neúměrně zdržovaly, pravděpodobněji je tedy varianta dotazování elektronickou poštou až se jí podaří dostatečně rozšířit.

12. Kódování a příprava dat pro analýzu

12.1. Matematizující a spekulativní pojetí sociologie

Již od počátku vývoje sociologie po celé 19. století a po část století 20. stojí proti sobě:

- spekulativní pojetí sociologie založené na badatelské práci filozofické a historické a
- matematizující pojetí sociologie využívající statistických údajů a posléze výzkumů a výpočetních postupů samočinných počítačů.

Námítky, že pokus o zachycení sociální skutečnosti v číslech je nepřipustným redukcionismem, má své kořeny v nedostatečné rozvinutosti kvantitativní metodologie a techniky a v pouze nedokonalé deskripci, k níž dospívala podstatná část dosavadních sociologických výzkumů. Nejde však o to, prohlásit všechno, co není spočítatelné za nevýznamné, ale o to učinit všechno, co je významné, spočítatelným. To je jeden ze způsobů, jak vymežit metodologický přístup využívající k poznávání měřicí nástroje, škály a kvantifikaci, statistickou analýzu a výpočetní techniku.

Vedle tohoto pojetí sociologie, sociologického výzkumu, nadále existuje v poznávání společenské reality přístup nebo dokonce řada přístupů, které upřednostňují celostní pohled na lidský svět, nesnaží se o postižení jeho měřitelných dimenzí a jejich vztahů.

Pochopitelně přijetí metodologického přístupu matematizující sociologie ovlivňuje způsob práce sociologa výzkumníka. Prostředky výpočetní techniky umožňují rozsáhlé, do hloubky směřující zpracování zjištěných údajů. Sám přístup i použitá technika ovlivňují i způsob myšlení sociologa.

12.2. Počítače v sociologickém výzkumu při přípravě a analýze dat

Trvalo, můžeme říci, desítky let, nežli sociolog získal přímý přístup k samočinnému počítači. Teprve v posledním desetiletí můžeme hovořit jednoznačně o přímém ovládnutí a využívání počítače sociologem. Děroštitková technika sloužící k přípravě a přenosu dat, k jejich třídění a posléze i k výpočtům pomocí speciálních programů tvořila první stupeň na této cestě. Dalším stupněm byly soubory programů vytvořené pro sálové počítače. K práci s nimi sociolog potřeboval radu a pomoc specialisty programátora nebo statistika. Analytickým nástrojem v rukou sociologa se počítač stává teprve masovým rozšířením

systemů programů (SPSS, SAS, BMDP a dalších) pro zpracování hromadných dat. Při jejich ovládnutí sám sociolog zadává příkazy počítači. Původně takto připravený program ve speciálním jazyce symbolických příkazů zprostředkovával počítači operátor sálového počítače, později terminálová síť tohoto počítače. V posledních letech již sociolog sám ovládá osobní počítač vybavený programem pro přípravu dat i pro jejich náročnou statistickou analýzu.

12.2.1. Co počítače sociologii přinášejí?

- urychlují zpracování dat;
- rozšiřují (spolehlivě a pohotově) paměť výzkumníka;
- velký rozsah jejich paměti umožňuje zpracovat ohromné soubory dat;
- umožňují rozsahem dat větší zobecnění, zvyšují jeho spolehlivost;
- počítače umožňují odhalit struktury vztahů, pravidelnosti v datech;
- počítače vybavené speciálními programy umožňují konstruovat a simulovat modely sociální skutečnosti;
- výsledky počítačů nejsou ovlivněny únavou, počítač nemá lidských slabín (má však zase slabiny stroje).

12.2.2. Jak počítače ovlivňují způsob myšlení sociologa výzkumníka?

1. Nutí sociologa, aby formuloval své cíle, problémy a požadavky naprosto jasně a jednoznačně. Vylučují každou neurčitost.
2. Vytvářejí (počítače) iluzi klamné přesnosti. Precizně zpracují i nespolehlivá data.
3. Počítače nabízejí nezkušenému výzkumníkovi velikou spoustu tabulek, grafů a čísel, v níž ztrácí svůj původní cíl (pokud jej měl). Čeho se z výsledků chopí, bývá pak dílem náhody a pragmatického řešení.
4. Programové vybavení počítače je zhmotněnou (petrifikovanou) metodologií. Kvalifikované využití této metodologie bez chyb však vyžaduje důkladnější logicko-matematickou, statistickou a metodologickou přípravu každého sociologa výzkumníka, který výsledky připravené počítačem využívá.
5. Sociolog výzkumník je veden logikou hlubší statistické analýzy, kterou počítače umožňují, k formulaci nových výzkumných otázek a k zevrubnějšímu využití výzkumem zjišťovaných údajů.

12.3. Kódování

Operace kódování může být připravena předem. Každá z uzavřených otázek standardizovaného dotazníku má předtištěná políčka, do nichž dotázaný nebo tazatel vyplňují kód (číselné označení) odpovědi. Tuto operaci můžeme však svěřit také kódérům. Dotazování zapíše odpověď mimo připravená políčka. Jejich záznam varianty odpovědi je kódovačem (kódérem) zkontrolován a přenesen do předtištěného políčka.

Náročnější část práce v rámci operace kódování však spočívá v kódování otevřených otázek. Úloha výzkumníka spočívá v navržení kategorií, do nichž jsou zařazovány jednotlivé volně vyjádřené odpovědi.

Práci při operaci kódování obecně dělíme na přípravu kódovacího klíče a vlastní kódování podle jeho pokynů. Kódovací klíč připravuje člen výzkumného týmu. Kódování je možno svěřit zaškoleným kódovačům.

12.3.1. Příprava kódovacího klíče

1. Rozdělení variability znaku do tříd se musí řídit cílem výzkumu. Musí odpovídat reálné struktuře zkoumaného jevu ve zkoumané populaci.
2. Použité klasifikační kategorie musí být vyčerpávající. Tedy třídy kategorizace musí pokrývat celé universum.
3. Třídy kategorizace se nesmí překrývat.
4. Kódovací klíč musí obsahovat jednoznačná kritéria bez nerozhodnutelných případů.
5. Použité kategorie musí být na téže rovině obecnosti.
6. Musí být použit jeden klasifikační princip.
7. Z praktických důvodů použitelnosti kódu jako třídícího znaku je třeba rozdělit kategorie dostatečně rovnoměrně nebo podle přijatého statistického modelu.

12.4. Příprava (pořizování a čišění) dat

Cílem přípravy dat je převést údaje z jejich fyzické podoby záznamu kódů nebo verbálních záznamů na dotaznicích, formulářích nebo jiných záznamových prostředcích do formy záznamu na paměťových

nosičích počítače. Druhým cílem přípravy dat je odstranění chyb, které v procesu tvory a přenosu dat mohly nastat a většinou také nastaly. Třetím úkolem přípravy dat je přenesení co největšího množství doprovodných údajů o datech do podoby využitelné při jejich zpracování.

12.4.1. Obecné zásady přípravy dat

1. Předem sestavíme plán přípravy dat.

Plán musí obsahovat všechny operace, které je třeba postupně vykonat při kontrole, přenosu, čišění a dokumentaci dat. Měl by respektovat podobu prvotních záznamů a míru jejich spolehlivosti, použité programové prostředky pro přenos, kontrolu a záznam dat i potřebnou výslednou podobu dat vstupujících do analýzy. Součástí plánu je také časový harmonogram zohledňující data převzetí prvotních dokladů, potřeby analýzy a termín odevzdání závěrečné zprávy z výzkumu.

2. Dbáme na prevenci chyb.

Ve výzkumné práci obecně, a při přípravě dat dvojnásob, platí, že nejlevnější chyba je ta, která se nestala. Řadu chyb můžeme odhalit pečlivou kontrolou ve fázi kódování dat, čitelným záznamem dat a jejich přehledným uložením.

3. Provádíme opakovanou kontrolu v průběhu přípravy dat.

Důvodem pro toto opatření je ověřený fakt, že řada chyb vzniká v průběhu přípravy dat, při přenosech a transformacích dat.

4. Dokumentujeme prováděné změny v datech.

Chyba v datech může vzniknout také nesprávnou interpretací kódů. Změníme-li například z důvodů stejné orientace stupnic škálu jednoho ze znaků na škálu s opačnou orientací, kód pro ano na kód pro ne a opačně, musíme tuto změnu pečlivě dokumentovat, jinak je omyl v interpretaci téměř zaručen.

5. Maximum chyb odstraňujeme v prvních fázích přípravy dat.

Čím později v řetězci přenosů chybu v datech odhalíme, tím obtížnější je zjistit její původ a tím méně informací máme o původním obsahu chybou změněné informace. Buď nám zabere více času oprava podle původního dokladu nebo ztrácíme cennou

informaci. Může se dokonce stát, neodhalíme-li chybu v prvotních fázích, že v agregované podobě dat nebude již odhalitelná vůbec.

Př.: Představme si například údaje o příjmech členů domácnosti.

Vyskytují-li se v souboru údajů chyby týkající se příjmů části osob odkázaných na sociální podporu, tyto chyby ovlivní výpočet příjmů domácností. Tak se stane, že údaje o příjmech všech domácností, v nichž tyto osoby žijí, budou nespolehlivé.

6. Dodržujeme pořadí fází přípravy dat.

Pořadí fází přípravy dat je dáno používanými programovými prostředky a změnami ve formě dat, které daná fáze vyvolá.

- Dříve než připravíme definitivní strukturu datového záznamu musíme si být jisti, že nám v prvotních dokladech důležitý údaj nechybí, že jsou všechny potřebné údaje převedeny do podoby kódů, které lze přenést do počítače.

- Dříve než provedeme samotný záznam dat, přenos dat na záznamové médium, musíme si být jisti, že jsou dokumenty úplné a že jsou správně uspořádány. Nesmí např. chybět některý z formulářů, stránka v jednom z dotazníků, stránky dotazníků by neměly být přeházeny a pod.

7. Neprovádíme analýzu před dokončením přípravy dat.

Nejčastější chybou málo zkušeného výzkumníka je snaha co nejdříve, ještě před skončením přípravy dat, vytvořit pracovní soubor a poskytnout zadavateli předběžné výsledky výzkumu. Většinou jde jen o vybrané údaje o několika znacích, několika otázkách. Není-li dokončena kontrola datového souboru a nejsou-li provedeny všechny opravy, výsledky neposkytujeme. Jen tak si můžeme být jisti, že nezpochybníme vypovídací schopnost dat celého finančně náročného výzkumného šetření.

12.4.2. Fáze přípravy dat

1. Evidence a kontrola prvotních dokladů (dotazníků, formulářů).

Nakódované dotazníky, záznamy o rozhovoru nebo pozorování, formuláře pro pořízení dat, musí mít všechny standardní podobu z hlediska pořadí a struktury zaznamenané informace. Musí být použito stejného kódového označení pro celý soubor jednotek a pokud některý formulář nebo jeho část chybí, musí být nahrazena dokumentem odpovídající struktury záznamu aby se předešlo

chybám. Osvědčilo se rozdělit soubor dokladů do podsouborů podle přirozeného členění výzkumného souboru dat.

2. Příprava struktury datového záznamu výzkumného souboru

K záznamu (pořízení) dat se dnes používá programového vybavení, které je nejčastěji součástí systému programů pro analýzu dat. V sociologii je nejrozšířenějším takovým programovým prostředkem systém programů SPSS¹⁴⁾ a jeho část pro přípravu dat nazvaná Data Entry. Úkolem této fáze přípravy dat je převést informace o struktuře záznamu dat v prvotním dokladu (dotazníku, záznamu nebo formuláři) do parametrických údajů programu Data Entry nebo přímo SPSS.

3. Záznam dat, přenos dat na paměťové médium, do počítače.

Záznam dat, také pořízení dat je operace, jejíž časová náročnost je přímo úměrná počtu jednotek, o nichž informaci ukládáme a pochopitelně také délce záznamu o jedné jednotce. Dnes se nejčastěji data vkládají z primárních záznamů přímo do osobního počítače event. do terminálu sálového počítače. Tím je snížen počet operací s daty a tedy i chybovost pořizování dat.

V šedesátých a sedmdesátých letech bylo nejčastěji užíváno k záznamu dat děrných štítků, které bylo třeba vyděrovat a pro kontrolu "vyděrovat" ještě jednou, tj. "přezkoušet". Děrný štítek obsahoval 80 sloupců, každý z nich pro záznam jednoho nejčastěji číselného kódu, jedné číslice. Záznam pro jednu výzkumnou jednotku pak sestával i z několika štítků. (Vždy však stejného počtu štítků pro každou jednotku.) Teprve tyto štítky byly za pomoci programu, který dešifroval obsah jednotlivých sloupců štítků, nahráván do paměti počítače.

Struktura záznamu výzkumného souboru, předem připravená v programu Data Entry, umožňuje vyškolenému pracovníkovi vkládat posloupnost údajů z formulářů pro jednoho respondenta jako jednu záznamovou jednotku o předem známé délce a struktuře. Další respondent nebo další statistická jednotka tvoří novou záznamovou jednotku, nový případ. Datové podsoubory zahrnující údaje o podsouborech jednotek je možno tvořit současně na několika počítačích a pak jednoduše spojit.

4. Kontrola úplnosti a uspořádání dat.

Jednak před vstupem do počítače, jednak v samotném procesu pořizování dat a jednak kontrolou dat v počítači, zjišťujeme, zda struktura záznamů o všech jednotkách je stejná. Při tom lze využít programového vybavení. Program Data Entry umožňuje

stanovit počet záznamových míst (sloupců) pro každý znak, typ znaku (číselný nebo alfanumerický), povolený rozsah hodnot, který kódy nesmějí překročit a pochopitelně i délku celého záznamu. Další možností je připravit strukturu záznamu dat v podobě předlohy tvaru použitého formuláře nebo stránek dotazníku. Výhodou je, že ten, kdo pořizuje data, vidí před sebou na obrazovce stránku dotazníku vzhledově identickou se stránkou fyzického dotazníku, z něž data opisuje. Tím je usnadněna optická kontrola prováděné práce a snižuje se tím množství chyb při záznamu dat.

5. Čištění dat, opravy chyb.

Nejspolehlivější kontrolou při pořizování dat je dvojí nezávisle na sobě provedený záznam dat ze stejného souboru formulářů. Porovnáním odstraníme případné překlepy.

Dále v počítači provádíme logické kontroly. Známe-li nutně existující pravidelnost, souvislost mezi hodnotami dvou nebo několika znaků, v počítači ji zkontrolujeme pro všechny statistické jednotky.

- Víme např., že spolu souvisí nejvyšší dosažené vzdělání a věk dotázaného: že vysokoškolského vzdělání nelze z důvodu přijímacích pravidel dosáhnout před 22. rokem věku.
- Můžeme zkontrolovat, zda svobodný dotázaný nemá v dotazníku uvedeny údaje o své manželce; pokud má, je důvod předpokládat, že je ženatý, rozvedený nebo snad vdovec.

Chyby opravujeme přednostně podle primárních dokladů - dotazníků, formulářů. Jen pokud údaj neznáme ani nemůžeme spolehlivě odvodit z ostatních dat, doplníme na místo platného údaje kód pro vynechávaná data (missing values). Tím sdělujeme počítači, že danou statistickou jednotku má pro chybou zatížený znak vynechat z výpočtu koeficientu nebo z tabulky.

6. Popis a dokumentace souboru dat.

Soubor dat je určen několika základními údaji. Jsou potřebné k jeho vyhledání, archivaci, k analýze dat i k rozhodnutí o jeho dalším využití. Kromě údajů o celém souboru dat je třeba určit také údaje o každém jednotlivém znaku. Konečně dokumentujeme doplňkové údaje o souboru, tj. údaje o způsobu pořízení dat a o příbuzných souborech..

Údaje o souboru dat:

1. jméno souboru (identifikátor, pod kterým soubor hledám)

2. místo uložení souboru (disketa, knihovna)
3. popis souboru (podrobnější označení obsahu souboru)
4. počet znaků a jejich jména
5. pořadí znaků v souboru
5. počet jednotek (event. členění na podsoubory)

Údaje o jednotlivých znacích:

1. jméno znaku (identifikátor, pod kterým znak hledám)
2. rozsah a formát znaku (typ znaku)
3. hodnoty znaku (pro kategorizovaný znak)
4. popis znaku (podrobnější označení znaku pro tabelace)
5. popisy hodnot znaku (označení hodnot pro tabelace)
6. kód (kódy) pro vynechávané hodnoty

Dokumentace doplňkových údajů o souboru:

1. uspořádání výběru
2. typ výzkumného uspořádání (Research Design)
3. podrobnosti o metodě sběru dat
4. charakteristika chyb a způsobu jejich oprav

Dokumentace doplňkových údajů o příbuzných souborech:

1. soubory dat z předvýzkumu
2. soubory dat z minulých srovnatelných výzkumů
3. mezinárodní soubory dat s možností komparace

Dokumentace provedených transformací

(nejlépe přímo v podobě formule, předpisu pro transformaci)

7. Zkušební výpočty a ladění programů.

Již na datech z předvýzkumu a později na cvičném podsouboru výzkumných dat jsou prováděny první výpočty sloužící převážně k metodickým cílům, k dokončení kódování, ke kategorizaci znaků resp. jejich překódování, k přípravě nových transformovaných znaků, k vyzkoušení vhodných tabelací a k přípravě mnohorozměrných statistických analýz. Výsledky těchto výpočtů nezveřejňujeme.

Programy pro hlavní fázi analýzy dat připravujeme v případě rozsáhlých datových souborů z úsporných důvodů na menším, cvičném souboru dat a na tomto souboru programy ladíme, tj. vylepšujeme parametry programů, hledáme jejich optimální varianty.

12.5. Data vstupují do analýzy

12.5.1. Možnosti uspořádání dat. Typy vstupních dat.

Do počítače mohou data vstupovat v mnoha různých uspořádáních. Záleží na tom, zda nahráváme přímo výsledky dotazování nebo pozorování jednotlivých jednotek podle předem zvolených znaků, pak se jedná o primární data, nebo data již zpracovaná do tabulek nebo dokonce jen matice koeficientů vyjadřujících nepodobnosti, vzdálenosti nebo podobnosti mezi objekty nebo proměnnými.

Tyto datové struktury je třeba přizpůsobit požadavkům analytické metody na jejich uspořádání vhodné pro vstup do výpočtu. Tak vznikají další typy dat.

Primární data

Základní matice dat je uspořádána tak, že její řádky tvoří jednotlivé případy, statistické jednotky, například dotázaní jednotlivci nebo města, o nichž jsme shromáždili údaje. Každý její sloupec představuje údaje o hodnotách jednoho ze sledovaných znaků postupně pro všechny jednotky.

Například řádky základní matice tvoří žáci základní školy a sloupce matice údaje o jejich výšce, váze, prospěchu z matematiky, počtu zameškaných hodin, výsledku inteligenčního testu, povolání otce, povolání matky, vzdělání otce a vzdělání matky a údaj o projeveném zvláštním nadání.

Pro některé úlohy má smysl vytvořit matici, jejíž řádky tvoří údaj o zjišťovaném znaku a jejíž sloupce odpovídají hodnotám za jednotlivé jednotky. Tato matice vznikne transponováním základní matice dat (tj. jejím překlopením podél hlavní diagonály) a říkáme jí tedy transponovaná základní matice dat. Používáme ji například pro konstrukci typologie znaků pomocí seskupovací analýzy.

Zjišťujeme vztahy mezi vlastnostmi pohádkových postav ve světě dětské fantazie. Jako statečný byl označen Honza, princ Bajaja, králevic, Rumcajs, Kašpárek, Plaváček a další hrdinové. Jako zlý byl označen Troll, černokněžník, drak, ježibaba, baba Jaga, čarodějnice a další pohádkové bytosti. Další posuzované vlastnosti mohou být moudrost, rychlost, lakota, čestnost, ochota pomáhat, laskavost, rozhodnost, vytrvalost, lstivost, ukrutnost, spravedlivost a další charakterové rysy pohádkových postav. Každá z řádek transponované základní matice dat vypovídá o jedné vlastnosti a každý sloupec

reprezentuje jednu z pohádkových postav. Výsledkem analýzy pak může být zjištění vlastností, které se u pohádkových postav nejčastěji objevují společně a tvoří typ kladného hrdiny.

Četnostní data

V sociologii často pracujeme s kategorizovanými znaky. V takovém případě zjišťujeme rozdělení jednotek do tabulky podle hodnot jednoho nebo několika znaků. Zajímá nás, kolik jednotek zkoumaného souboru má danou hodnotu znaku nebo kombinaci hodnot dvou nebo několika znaků. Vytváříme tedy četnostní data. Rozeznáváme četnostní data jednorozměrná, dvourozměrná a vícerozměrná.

Jednorozměrná četnostní data sestávají z četností, tedy počtů nebo podílů jednotek v jednotlivých hodnotách kategorizovaného znaku. Vznikla vlastně jako výsledek třídění jednotek do těchto kategorií označených hodnotami znaku. Rozeznáváme absolutní četnosti, což jsou přímo počty jednotek, které třídíme, relativní četnosti vyjádřené jako podíl jednotek v dané kategorii znaku, kumulativní absolutní četnosti vyjadřující počty jednotek ve všech uspořádaných kategoriích až po uvažovanou a konečně kumulativní relativní četnosti vypovídající o podílu kumulativně načteného souboru jednotek z uvažovaného celku. Relativní četnosti mohou být vyjádřeny jako podíl z celku, kterému je přisouzena hodnota 1 (číslo mezi nulou a jedničkou) nebo jako procento, tedy podíl z celku, který je vyjádřen jako 100% (číslo mezi 0 a 100%).

Dvourozměrná četnostní data jsou tvořena tabulkou, jejíž řádky jsou označeny hodnotami jednoho z třídících znaků a jejíž sloupce označují hodnoty druhého třídícího znaku. V jednotlivých polích tabulky jsou zaznamenány počty nebo podíly jednotek, které splňují obě podmínky současně, tj. jednotek, které mají hodnotu řádkového třídícího znaku rovnou jeho řádkové hodnotě a hodnotu sloupcového třídícího znaku rovnou hodnotě v příslušném sloupci. Také dvourozměrná četnostní data mohou být absolutní nebo relativní, prostá nebo kumulovaná.

U dvourozměrných četnostních dat rozeznáváme navíc tzv. relativní řádkové četnosti a relativní sloupcové četnosti. Slouží zejména pro srovnání rozložení podsouborů definovaných kategoriemi jednoho třídícího znaku podle hodnot druhého třídícího znaku. Relativní řádkové četnosti jsou tedy podíly jednotek určitého řádku s danou hodnotou sloupcového třídícího znaku ze souboru všech jednotek tohoto řádku. Obdobně relativní sloupcové četnosti jsou podíly jednotek určitého sloupce s danou hodnotou řádkového třídícího znaku ze souboru

všech jednotek tohoto sloupce. Mezi sebou můžeme srovnávat relativní řádkové četnosti jednotlivých podsouborů tvořených řádky tabulky. Nezávisle na tom můžeme také mezi sebou srovnávat relativní sloupcové četnosti jednotlivých podsouborů tvořených sloupci tabulky.

Dvourozměrné tabulky absolutních četností, celkových relativních četností, relativních řádkových četností a relativních sloupcových četností můžeme prezentovat jako samostatné tabulky nebo mohou být obsaženy ve společné tabulce určené stejnými řádky a sloupci. Potom v každém políčku tabulky např. první číslo označuje četnost absolutní a druhé celkovou relativní četnost.

Časové řady

Typem dat se zvláštním určením jsou časové řady. Tento typ dat vyjadřuje vývoj určitého, většinou kvantitativního, ukazatele v několika po sobě následujících časových okamžicích nebo časových intervalech. Jedná-li se o vývoj jediného ukazatele a o jeden sledovaný objekt, jde o jednorozměrnou časovou řadu. Jedná-li se o vývoj téhož ukazatele u několika objektů, jde o časovou řadu dvourozměrnou. Pochopitelně datová struktura může být složitější. Pro jeden nebo pro několik objektů můžeme sledovat paralelní vývoj několika ukazatelů v čase. Nejčastěji jsou časové řady uváděny ve tvaru, kdy řádka vyjadřuje údaje pro jeden objekt a jeden sledovaný ukazatel. V jednotlivých sloupcích následují hodnoty ukazatele pro časové okamžiky nebo intervaly v pořadí jak následovaly za sebou. Je-li objektů nebo ukazatelů více, údaje ve sloupci pod sebou odpovídají témuž časovému okamžiku nebo intervalu.

Agregovaná data

Sociologický výzkum pracuje se soubory statistických jednotek různé velikosti a různého řádu. Shromažďujeme údaje o jednotlivcích, rodinách, cenových, společně hospodařících i bytových domácnostech, o domovních blocích, městských urbanech, katastrech, městských čtvrtích a obvodech, obcích, katastrálních územích zahrnujících střediskovou a několik spádových obcí, o volebních okrscích, volebních obvodech, okresech, regionech, krajích, republikách, o jednotlivých závodech, pracovištích, místně odloučených provozních jednotkách, podnicích, celých rezortech, o výrobních oborech a odvětvích, o jednotlivých povoláních a o skupinách povolání, o politických stranách a jejich koalicích, o školách, jejich žácích, studentech a učitelích, jednotlivých ročnících a třídách těchto škol, o územně, věkem žáků nebo oborově vymezených souborech, typech a stupních škol.

Data o jednotkách nejružnějšího stupně obecnosti je možno použít pro výpočet charakteristik za jednotku vyšší úrovně. Tak je možno z příjmů všech jednotlivých členů domácnosti spočítat příjem domácnosti jako celku. Tento ukazatel považujeme za agregovaný, vytvořený agregací (seskupením, sloučením) z dílčích údajů o příjmech jednotlivých členů domácnosti. Z údajů o ekonomických výsledcích závodů lze vypočítat agregovaný údaj o celém podniku. Výsledky výzkumu žáků mnoha různých škol, pokud byl výběrový postup vhodně navržen, lze agregovat do menšího množství údajů charakterizujících tyto školy nebo jejich ročníky resp. třídy.

Agregovaná data jsou tedy data o jednotkách vyššího řádu zjištěná nebo vypočtená z dat za části těchto jednotek, z údajů o jednotlivých jednotkách nižšího řádu, z nichž agregáty sestávají. Data vypovídají o celcích na základě údajů o částech.

Relační data

Některé analytické metody, zvláště metody mnohorozměrné analýzy dat, vyžadují vstupní data v předpracované podobě ve tvaru, kdy nevelké množství číselných údajů obsahuje silně zhuštěnou informaci. Relační data plní tuto úlohu a shrnují v jedné tabulce údaje o souboru proměnných, o jejich vztazích zjištěných pro soubor analyzovaných jednotek. Jiná podoba relačních dat soustřeďuje v tabulce o jednotkách údaje o vztazích mezi jednotkami z hlediska mnoha zkoumaných znaků. Relační data tedy vyjadřují relace, vztahy, buď relace mezi znaky, proměnnými realizované na souboru jednotek nebo naopak relace, vztahy mezi jednotkami realizované prostřednictvím mnoha dílčích dimenzí a ukazatelů mnoha znaků, proměnnými.

Rozeznáváme několik typů těchto relačních dat. Abychom mohli snáze popsat rozdíly mezi nimi, uvedeme vymezení několika základních pojmů:

Vzdálenost = definujeme ji mezi jednotkami, přičemž v jednom rozměru se rovná rozdílu hodnot znaku mezi jednotkami. Ve dvou rozměrech je možno vypočítat nebo změřit vzdálenost mezi jednotkami jako klasickou eukleidovskou vzdálenost. Ta je dána druhou odmocninou z rozdílu čtverců mezi vzdálenostmi v obou rozměrech. Předpokládáme nezávislost obou proměnných, které tvoří jakoby navzájem k sobě kolmé souřadnice dvojrozměrného vztahu a vzdálenost v dvojrozměrném prostoru je za tohoto předpokladu dána jako přepona pravoúhlého trojúhelníka. To je však jen jedna z možností. Vzdálenost ve dvojrozměrném prostoru můžeme definovat také v metrice "městských bloků", v tzv.

Manhattanské metrice. Vzdálenost mezi dvěma místy v centru New Yorku je třeba překonat po pravouhlých souřadnicích. Rovná se tedy součtu vzdálenosti ve směru rovnoběžně spolu jdoucích ulic "avenues" a vzdálenosti ve směru napříč jdoucích ulic "streets". Obecně je tato "manhattanská" vzdálenost součtem vzdálenosti mezi jednotkami v jednom rozměru, tj. z hlediska jednoho znaku, a vzdálenosti ve druhém rozměru, tedy z hlediska druhého znaku. Pochopitelně otázkou zůstává volba vhodného měřítka.

Podobnost = Podobnost může být dána v jednoduchém případě počtem společných charakteristik. Je možno si představit, že jsou si podobní dva lidé, kteří jsou stejného pohlaví, věku, vzrůstu, váhové kategorie, vzdělání, sociálního postavení, majetku, příjmové kategorie, rodinného stavu, náboženského vyznání, inteligence, mravního přesvědčení, a z hlediska vzhledu např. barvy vlasů, tvaru obličeje, tvaru úst, barvy očí, výšky čela a tvaru nosu. Mohou se sice lišit v některé další podstatné charakteristice, ale je jistě možno předpokládat, že si budou podobnější, než dvojice lidí, kteří se ve většině z těchto charakteristik budou mezi sebou lišit. A na podobném principu je měřena podobnost mezi jednotkami v mnohorozměrném prostoru sociálních charakteristik.

Nepodobnost = Nepodobnost je nejčastější druh relace, který zjišťujeme u vztahů mezi znaky, proměnnými. Mezi míry nepodobnosti řadíme celou řadu koeficientů vztahu dvou proměnných, například známý Pearsonův koeficient lineární korelace. Platí vztah, že čím vyšší nepodobnost proměnných, tím nižší je koeficient mezi nimi.

Relační data mohou tedy mít podobu matice vzdáleností, matice podobností nebo matice nepodobností. Tato matice může být symetrická, čtvercová - to je v případě, že v řádcích a sloupcích matice jsou uvedeny stejné proměnné nebo stejné jednotky. Druhou možností je "obdélníková" matice, která obsahuje v řádcích jeden soubor proměnných a ve sloupcích druhý soubor proměnných.

Symetrická čtvercová matice nepodobností nebo podobností bývá někdy nahrazena jen maticí trojúhelníkovou z toho důvodu, že trojúhelník koeficientů pod diagonálou je zrcadlovým obrazem trojúhelníku nad diagonálou. Na diagonále vyjadřující vztah proměnné k sobě samé je maximální podobnost vyjádřena jedničkami, minimální nepodobnost nebo minimální vzdálenost nulami.

Asymetrická "obdélníková" matice mívá nejčastěji podobu tzv. korelačního pásu. Ten vypovídá o relacích (vyjádřených korelačními koeficienty) několika vybraných proměnných (nejčastěji identifikací) k ostatním proměnným.

Relační data slouží pro vstup do faktorové analýzy, seskupovací analýzy, mnohorozměrného škálování v jeho metrické i nemetrické podobě a do dalších speciálních multivariačních technik.

12.5.2. Práce s vynechávanými údaji

V průběhu sběru dat, při jejich kódování i ve fázi pořizování dat dochází k celé řadě chyb, projevuje se i nedostatek potřebných informací ve standardním tvaru. Najdeme-li a odstraníme chybu, není vždy možné nahradit chybný údaj údajem správným, protože jej prostě neznáme a dodatečně ho v některých případech ani nelze zjistit. Chybějící hodnotu znaku pro zkoumanou jednotku nahrazujeme zvláštním kódem. Protože při statistických výpočtech používajících daného znaku jednotku vynecháváme z analýzy, nazýváme takovéto kódy vynechávaná data (missing data). Hodnoty znaků vyjadřující, že údaj chybí, neznáme jeho správnou hodnotu nebo že nebude zařazen do analýzy, jsou vynechávané hodnoty (missing values).

Existuje ještě jeden důvod, proč definujeme hodnotu znaku jako vynechávanou. Přestože hodnota znaku je správně určena, jeho hodnota tedy nenahrazuje chybně zaznamenaný nebo neznámý údaj, v konkrétním výpočtu potřebujeme soubor jednotek definovaný zvolenou hodnotou nebo hodnotami předem známého znaku vynechat z analýzy. Děje se tak z celé řady většinou interpretačních důvodů. Chceme například vynechat extrémní hodnocení nebo naopak nezařadit do výpočtu respondenty, kteří se vyjádřili neurčitě, nebo kteří nemají ustálený názor a kolísají v odpovědích. Takovéto vynechání z analýzy je většinou omezeno na řešení určité dílčí úlohy a je pouze součástí definice této úlohy. Nemusí se stát trvalou součástí definice dat.

Můžeme tedy rozlišovat vynechávaná data na trvale vynechávaná a přechodně vynechávaná. Rozlišení je však, jako mnoho jiných rozhodnutí prováděných v sociologickém výzkumu, relativní.

Analytické tabulky generované výpočetními programy a programovými systémy jako je například SPSS ¹⁴⁾ nabízejí variantně různé způsoby práce s vynechávanými údaji. Někdy tisknou vedle sebe tabulky dvě. Jednu s relativními četnostmi spočtenými pro celý, neredukovaný soubor jednotek a druhou pro soubor bez vynechávaných dat.

12.5.3. Modifikace a transformace dat, konstrukce nových znaků

Kategorizace znaků

Rozlišujeme znaky spojité a znaky diskrétní, s konečným počtem hodnot. Z praktického hlediska některé znaky s konečným počtem hodnot plní vlastně úlohu spojitých znaků, zejména pokud rozsah hodnot je velký a stupnice je dostatečně podrobná. Údaje o výšce postavy v centimetrech, o váze v kilogramech, věku v letech, příjmu v Kčs, o rozloze pozemku v arech nebo v hektarech apod. je vhodné pro účely tabelací vyjádřit variantně v kategorizované podobě.

V takovém případě definujeme nový znak s menším počtem hodnot, než měl původní, a stanovíme předpis jímž hodnotám ze zvolených intervalů původního znaku přiřadíme zvolené hodnoty nově zkonstruovaného znaku. Přitom se snažíme zachovat maximum možných relací, které v původním znaku platily. Byl-li původní znak poměrový nebo intervalový, snažíme se, aby si hodnoty nového znaku zachovaly tytéž vlastnosti. Pokud byl původní znak pořadový (ordinální), kategorie nového znaku reprezentují intervaly ve stejném pořadí. Pozorně zkontrolujeme, jakých hodnot nabývá nový znak pro krajní body intervalů kategorizovaného znaku. Pečlivě transformujeme také vynechávané hodnoty.

Kategorizaci můžeme provést jmenovitým výčtem hodnot původního znaku, pro něž je definována hodnota kategorizovaného znaku. Druhý, častější způsob je stanovení spojitého intervalu jeho krajními body. Třetím používaným způsobem je aritmetická operace zaokrouhlování, jímž je společně pro celý definiční obor znaku určeno pravidlo, jak transformovat hodnoty. Kterou z hodnot zaokrouhlit ještě dolů a kterou již nahoru. Obdobou tohoto způsobu je tzv. "ořezávání" (truncation). Zvolíme úroveň počtu platných míst před nebo za desetinnou čárkou a zbylý číselný údaj vynecháme (ořežeme). Tak z původních 1,556 vznikne 1, z původních 2,001 ale také z 2,999 vznikne 2 atd.

Konstrukce nového znaku z jediného znaku

Rozložení hodnot původního znaku může být překážkou pro jeho hlubší rozbor, pro jeho použití v náročnějších technikách mnohorozměrné analýzy dat. Rozložení je většinou nesymetrické, šikmé a metoda analýzy vyžaduje pokud možno zvonovité rozložení dat blízké normálnímu. V takovém případě volíme transformaci, která umožní zhustit hodnoty na vzdálenějším méně obsazeném konci stupnice a

rozlišit číselné hodnoty na opačném konci škály. K tomu slouží nejčastěji buď logaritmická nebo logitová transformace.

Logaritmická transformace: $y = \ln x$

Logitová transformace: $y = \ln \frac{x}{1-x}$

Aritmetická transformace - konstrukce znaku ze skupiny znaků

Aritmetická transformace využívá k výpočtu hodnoty nově vytvářeného znaku aritmetických operací sčítání, odčítání, násobení, dělení, umocňování a odmocňování a některých speciálních funkcí, které provádí s hodnotami jednoho nebo několika znaků.

Nejjednodušším takto konstruovaným znakem je konstanta. Znak nabývá pro všechny jednotky stejné hodnoty: $y = a$.

Jednoduchou konstrukcí je kopie znaku: $y = x$.

Častým případem je součtový znak: $Y = A + B + C + D$.

Často bývá používán také rozdíl, součin nebo podíl:

$$Y = P - Q \quad Y = R * S \quad Y = P / Q.$$

Ze složitějších transformací je možno jmenovat aritmetický a geometrický průměr z několika znaků, součet čtverců dvou znaků a jeho odmocnina a další aritmetické výrazy libovolné složitosti.

Součtová transformace

Zvláštním případem konstrukce znaku je součtová transformace, jíž tvoříme nový znak, jehož hodnota vyjadřuje počet výskytů zvoleného kódu v baterii znaků, například počet kladných odpovědí na množinu stejně strukturovaných otázek nebo počet správně zodpovězených otázek v testu apod.

Logická transformace

Konstrukce nových znaků je založena na kombinaci hodnot několika měřitelných, zjišťovaných znaků. Některé vztahy mezi znaky vyjadřujeme aritmetickým výrazem, jiné je vhodné vyjádřit logickými podmínkami. Hodnoty nově konstruovaného znaku jsou podmíněny společnou, vzájemně se doplňující nebo vylučující existencí vybraných hodnot dvou nebo několika znaků.

Výsledný znak je většinou konstruován několika logickými transformacemi. Každá z podmínek tvoří jednu z větví logického stromu. V konečné kombinaci se navzájem doplňují a vylučují. Tak lze vytvořit i značně členitý logicky konstruovaný znak, jehož hodnoty mohou vhodně modelovat teoretické konstrukce. Logická konstrukce znaků je jedním z důležitých nástrojů pro testování komplikovaných sociologických teorií.

13. Formulace úloh pro statistickou analýzu dat

Tato kapitola ať není chápána jako pokus nestatistika učít statistiku. Takové ambice si neklade. Jde jen o pohled sociologa opřený o zkušenosti z výzkumné praxe, o můstek, jistě nedokonalou lávku mezi pojmy a procedurami výzkumné dílny sociologa a nástroji statistika. Skutečný návod k práci se statistickými nástroji čtenář nalezne v zasvěcených knihách odborníků statistiků.

13.1. Úloha deskripce

Nejzákladnější úlohou, s níž se setkáváme při statistické analýze dat je popis (deskripce) souboru dat. Soubor jednotek charakterizujeme z hledisek jednotlivých znaků a jejich kombinací, a to prostřednictvím tabulek četností (frekvencí) kategorizovaných znaků a výpočtem souhrnných statistických charakteristik spojených i kategorizovaných znaků za soubory.

13.1.1. Četnosti kategorizovaných znaků

Pro kategorizované znaky (tj. znaky s rozumně omezeným počtem kategorií, které lze třídit, tabelovat) spočteme rozložení četností podle hodnot znaku. Tvoříme tabulky třídění prvního stupně. Z primárních dat provádíme výpočet absolutních četností, které tvoří základ pro ostatní typy četnostních dat. Známe tedy počty jednotek pro každou z hodnot kategorizovaného znaku. Známe také celkový počet jednotek v souboru. Jako podíl jednotek z celku spočteme tedy relativní četnosti a jako postupně načítanou sumu těchto podílů kumulativní relativní četnosti.

Třídění kombinujeme také pro dvojice znaků. Tím vytváříme tabulky druhého stupně, a to buď absolutních nebo relativních četností, tj. buď počtů jednotek nebo jejich podílů z celku. Podíly jednotek z řádku dané tabulky představují řádek relativních řádkových četností. Je-li takto zpracována celá tabulka, hovoříme o dvojrozměrné tabulce relativních řádkových četností. Obdobně podíly jednotek ze sloupce analyzované tabulky tvoří sloupec relativních sloupcových četností a takto spočtené sloupce tabulku relativních sloupcových četností.

Také dvojrozměrné tabulky mohou být sestaveny jako kumulativní, tj. vypovídající o podílu jednotek s hodnotou nižší nebo rovnající se

hodnotě sloupcového nebo hodnotě řádkového znaku. Dále je možno v jedné složené tabulce kombinovat údaje o celkových, řádkových i sloupcových relativních četnostech.

Tato četnostní data představují základ interpretace pro všechny typy kategorizovaných znaků. Pro nominální znaky jsou v zásadě nezastupitelná. Pro ostatní typy znaků spolehlivě chrání výzkumníka před chybnými závěry z informačně zhuštěných údajů souhrnných statistických charakteristik.

13.1.2. Statistické charakteristiky spojitých znaků

Pro kategorizované znaky můžeme spočítat tabulku rozložení četností. Pro znaky spojitě, nabývající pro každou jednotku v principu jiné hodnoty, takovou tabulku sestavit prakticky nelze, nemá to analyticky smysl. Máme v zásadě dvě možnosti, jak se s těmito daty blíže seznámit v jejich souhrnné podobě za celý soubor: kategorizaci spojitěho znaku nebo výpočet souhrnné statistické charakteristiky.

Jestliže spojitý znak kategorizujeme, rozdělíme obor jeho hodnot do několika, nejčastěji dvou až deseti intervalů, kategorií a do těchto kategorií pak zařadíme všechny jednotky. Kategorizovaný znak můžeme třídit a tabelovat jak bylo shrnuto v předchozím paragrafu. Výpočet statistické charakteristiky je založen na snaze vyjádřit jedním číslem to podstatné, co platí o celém souboru jednotek z hlediska analyzovaného znaku.

Rozlišujeme dvě hlavní skupiny statistických charakteristik: míry polohy, úrovně nebo střední hodnoty znaku a míry variability, rozptýlení znaku.

Míry polohy, střední hodnoty

Míry polohy vypovídají o nejtypičtější hodnotě znaku nebo o hodnotě, která by nejlépe zastupovala jednotky souboru z hlediska zkoumaného znaku. Pro nominální znaky je to módus, tj. hodnota znaku, kterou má největší počet jednotek souboru. Pokud je takových nejčastěji zastoupených hodnot několik, hovoříme o vícemodálním rozložení, např. o bimodálním. Také ostatní typy kategorizovaných znaků mohou s touto statistikou pracovat. Nemá však význam pro znaky spojitě.

Mírou polohy pro pořadové, ale i pro intervalové a poměrové znaky je medián, tj. řečeno slovy statistika ".střední člen uspořádané statistické řady.". Jednotky souboru jsou seřazeny z hlediska hodnot

charakterizovaného znaku. Výše mediánu je stanovena jako hodnota znaku té jednotky, která leží přesně uprostřed řady jednotek. Pro sudý počet jednotek se medián spočte jako průměr hodnot dvou středních členů řady.

Aritmetický průměr je nejčastěji počítanou statistickou charakteristikou pro číselné (intervalové i poměrové) znaky. Jeho výpočet je založen na myšlence stejného příspěvku každé jednotky společné průměrné hodnotě. Jsou-li v souboru zastoupeny nejčastěji jednotky se střední úrovní hodnot a výrazných odchylek není mnoho, aritmetický průměr vyjadřuje vhodným způsobem střední, pravděpodobnou, očekávanou hodnotu za celý soubor. Jsou-li však zastoupeny jednotky s nízkými a jednotky s vysokými hodnotami, hodnota průměru se pohybuje mezi těmito naměřenými extrémnějšími hodnotami a průměr vyjadřuje neexistující abstrakci, není vhodnou mírou. Ani pro rozložení hodnot soustřeďující většinu jednotek k jednomu pólu stupnice a rozprostírající hodnoty pro zbytek jednotek podél stupnice až k jejím opačným extrémním hodnotám nepředstavuje průměr vhodnou míru. Jeho hodnota je příliš závislá na nepatrných změnách v zastoupení extrémních hodnot u několika jednotek.

Existují ještě další méně známé míry polohy, míry střední hodnoty, např. geometrický průměr, harmonický průměr a další. Výklad jejich použití a podrobnější výklad k uvedeným charakteristikám je třeba hledat v učebnicích statistiky. (viz. např. [Řehák - Řeháková 1979]).

Míry variability (rozptýlení)

Mezi míry variability řadíme statistické charakteristiky vyjadřující nakolik se liší hodnoty posuzovaného znaku pro jednotky souboru. Velké rozdíly mezi hodnotami znaku pro jednotky souboru způsobují vysokou hodnotu míry rozptýlení. Minimální hodnoty bude taková charakteristika nabývat v případě, že všechny jednotky mají stejnou hodnotu znaku.

Jednoduchou charakteristikou variability znaku je variační rozpětí. Je vyjádřeno jako rozdíl mezi minimální a maximální hodnotou znaku v souboru.

Variabilitu je možno měřit nejen pro číselné znaky, ale i pro znaky pořadové a nominální. Nomvar, tj. nominální variance, a dorvar, diskrétní ordinální variance, jsou vhodnými mírami pro nominální a pro ordinální (pořadové) znaky. Snadno z nich lze vypočítat odvozené charakteristiky potřebné pro další rozbor dat. (viz. [Řehák - Řeháková 1979: 67nn, 88nn])

Rozptyl (variance) je definován pomocí odchylek od aritmetického průměru, a to tak, že se sečtou druhé mocniny těchto odchylek a vydělí se počtem případů zmenšeným o jeden. Každá jednotlivá odchylka vstupuje do výpočtu rozptylu svou druhou mocninou. Rozptyl je tedy tím větší, čím dále od průměru se nacházejí jednotlivé hodnoty znaku v souboru. Jsou-li všechny hodnoty v průměru, rozptyl je roven nule.

Směrodatná odchylka (standard deviation) je druhou odmocninou rozptylu.

13.2. Úloha komparace

Jednou z nejčastějších úloh je srovnání souborů nebo podsouborů z hlediska jednoho nebo několika znaků. Porovnáváme populace státních celků, národů, přírozených sociálních skupin nebo uměle vytvořených agregátů. Klademe si za cíl ověřit rozdílnost mezi nimi nebo změřit velikost tohoto rozdílu, nebo naopak je naší snahou prokázat, že rozdíly mezi nimi jsou zanedbatelné, že tvoří z hlediska zkoumané charakteristiky jeden soubor. K tomu používáme nejčastěji porovnání souhrnných statistických charakteristik za podsoubory. Podrobnější informaci poskytuje porovnání statistických rozložení znaku za podsoubory. To je však obtížnější a provádí se zřídka, ve speciálních úlohách.

13.2.1. Porovnání souhrnných charakteristik

Nejběžnějším srovnáním je komparace dvou souborů z hlediska jejich průměrů. Provádíme je např. tzv. t - testem. Oba soubory lze považovat buď za rozdílné, průměry se liší, nebo naopak neprokážeme rozdíl průměrů a soubory považujeme jen za nevýznamně odlišné.

Náročnější operaci provádíme při komparaci několika podsouborů. Sledujeme vliv faktoru, třídící proměnné, na průměrnou hodnotu a na rozptyl hodnot v rámci těchto podsouborů. Analýza rozptylu je založena na porovnání velikostí rozdílů uvnitř jednotlivých podsouborů s rozdíly průměrných hodnot mezi podsoubory. Je-li rozptyl mezi podsoubory významně větší než vnitřní variance uvnitř podsouborů, analýza rozptylu povrdí, že se podsoubory z hlediska třídícího znaku významně liší. Jejím výsledkem je tedy zjištění faktorů, které v rolích třídícího znaku mají vliv na dosahované hodnoty analyzovaného znaku. Tak lze například testovat vliv pohlaví (třídícího znaku) na mzdu (analyzovaný znak) a potvrdit, že pohlaví má významný vliv na

výši vyplácené mzdy, tj. že muži pobírají významně vyšší mzdu než ženy. Dále lze na základě výsledků říci, jaký podíl zjištěných rozdílů sledovaný faktor způsobuje, např., že v daném případě vysvětluje pohlaví 25% rozdílů v příjmu zaměstnanců podniku. Analýzu rozptylu je možno uplatnit jako jednorozměrnou i jako mnohorozměrnou. Při té je porovnáván mezi sebou vliv různých faktorů.

Analýza kovariance je příbuzná metoda k analýze rozptylu. Jestliže základem analýzy rozptylu je analýza vlivu třídícího znaku na základě komparace podsouborů daných jeho hodnotami, analýza kovariance hodnotí vliv třídícího znaku jako celku, jako tzv. kovariančního faktoru.

13.2.2. Porovnání statistických rozložení

Porovnávané podsoubory nespĺňujú vždy podmínky pro srovnání statistických charakteristik, např. průměrů. Jako alternativa se nabízí náročnější metody komparace. Podsoubory je možno porovnávat na základě rozložení četností v rámci sledovaného znaku. Větší množství informací, které vstupují do analýzy, vyžaduje složitější zpracování, přináší však také větší rozlišovací schopnost. Uvedené metody lze nalézt pod označením analýza profilů.

13.3. Měření asociací a korelační analýza

Třetí skupinou úloh je zjišťování souvislostí mezi znaky. Chceme např. zjistit, jaký je vztah mezi kouřením a konzumací alkoholu nebo mezi pozitivním postojem k ekonomické reformě a jednotlivými charakteristikami životní úrovně dotázaného. Zjišťujeme, zda vztah mezi znaky je významný. K tomu slouží celá řada statistických testů lišících se podle typů znaků, které dáváme do souvislosti. Další velmi častou úlohou je měření síly tohoto prokázaného vztahu. Přitom je třeba rozlišovat, zda se jedná o souvislost mezi dvěma nominálními znaky, mezi znaky pořadovými nebo mezi znaky kardinálními. Každé úrovni odpovídá jiná nabídka koeficientů.

13.3.1. Lineární a nelineární vztahy

Zkoumáme-li souvislost intervalových nebo poměrových, tedy kvantitativních znaků, má smysl rozlišovat, zda mezi sledovanými znaky

existuje lineární vztah, který je v ideální podobě vyjádřitelný přímkou. Těsnost tohoto vztahu, tj. stupeň, v němž vztah odpovídá zobrazení na přímku, vyjadřuje Pearsonův koeficient lineární korelace. Teoreticky nabývá hodnot od -1.00 do +1.00. Nabývá-li v praxi hodnot 0.7, 0.8, 0.9, nebo dokonce vyšších, jedná se bezpochyby o silnou pozitivní lineární souvislost mezi zkoumanými znaky. (Při stejných vysokých záporných hodnotách jde o negativní lineární vztah.) Je-li korelační koeficient nižší, neznamená to, že mezi znaky souvislost není. Může se jednat o souvislost nelineární, kterou je třeba měřit jinak, složitějším způsobem.

13.3.2. Tabulka korelací a její typy

Ve výzkumu se často setkáme se zpracováním celých skupin příbuzných znaků, kterým říkáme baterie. Vyjadřují rozmanité činnosti, postoje nebo hodnocení jednotlivých dílčích ukazatelů. V takovém případě bývá výsledek analýzy vyjadřován tabulkou korelačních koeficientů. Vysoké hodnoty koeficientů vyjadřují silný vzájemný lineární vztah, a to kladné hodnoty kladný, záporné hodnoty záporný. Tabulky mohou být symetrické, čtvercové, vyjadřující vzájemné vztahy znaků v takové baterii. Jiným typem korelačních tabulek jsou tabulky vyjadřující vztahy všech znaků z takové baterie k jiné skupině znaků. Takové obdélníkové tabulce, která nemá vlastnost symetrie, říkáme korelační pás. Často bývá využívána pro rozbor vztahů nějaké baterie činností nebo postojů k identifikačním znakům.

13.3.3. Explorační analýza závislostí

Jedním z postupů, s nímž se setkáváme v praxi sociologického výzkumu je explorační analýza závislostí mezi znaky, zkoumanými proměnnými. Nemá smysl předstírat, že je tento postup výjimkou, že většina výzkumníků předem sestavuje modely a testuje hypotézy. Právě proto, že se jedná o přístup relativně rozšířený, je třeba upozornit na jeho možnosti a limity, na úskalí spojená s jeho užíváním.

Výpočet matice korelačních koeficientů pro baterie analyzovaných znaků je jedním ze způsobů, jak realizovat tuto strategii analýzy dat. V takovém případě množina koeficientů, které překračují předem stanovenou hranici, nebo odstupňované hranice, vypovídá o významných zjištěných souvislostech. Výhrady mohou být v principu dvojí. Jednak koeficienty nezahrnou jiné než lineární nebo jim blízké vazby mezi proměnnými. Jednak nalezená empirická souvislost mezi dvěma měřeními

znaky ještě vůbec nemusí znamenat víc než jen důsledek jiných souvislostí, které se v tomto vysokém koeficientu pouze odrážejí. Např. v minulých letech často odhalovaná závislost mnoha znaků na vzdělání odrážela celou řadu jiných, mnohdy podstatnějších souvislostí, které jen nebyly zařazeny do plánu pozorování a zjišťování.

13.4. Další metody a možnosti mnohorozměrné analýzy dat

Existuje množství statistických metod analýzy dat. Jejich velká skupina zkoumá vztahy mezi mnoha znaky současně. Takovým metodám říkáme metody mnohorozměrné analýzy dat. Každá z těchto metod má za sebou jiný matematicko-statistický model, má jiné požadavky na charakter proměnných, které do ní vstupují, klade jiné nároky na výzkumníkovu schopnost formulovat smysluplnou sociologickou úlohu a přeložit ji do řeči čísel a matematických symbolů. Z hlediska sociologa výzkumníka se metody liší zejména modelovou představou úlohy, pro jejíž řešení mají sloužit.

13.4.1. Regresní analýza

Jejím cílem je zjistit, které měřitelné nezávisle proměnné a jakou měrou způsobují změny závisle proměnné. Jednorozměrná a mnohorozměrná regrese, lineární a nelineární, přímá nebo po krocích prováděná regresní analýza jsou jen základní větve a možnosti tohoto přístupu.

13.4.2. Faktorová analýza (explorační a konfirmační)

Jejím cílem je zjistit, jaké jsou hlavní pojmenovatelné zdroje, faktory zjišťovaných rozdílů mezi zkoumanými jednotkami - tak je tomu u explorační faktorové analýzy. Chceme najít to společné, co rozlišuje respondenty, skupinu proměnných tvořící svými vzájemnými těsnými vazbami faktor, který lze jednak pojmenovat a interpretovat, jednak použít k vysvětlení nalezených rozdílů v datech. Takových faktorů hledáme celou řadu a v naprosté většině úloh o nich předpokládáme, že jsou navzájem nezávislé.

Konfirmační faktorová analýza si klade náročnější cíle. Jak napovídá název, nesnaží se jen odhalit vztahy (jako její starší jednodušší sestra). Jejím cílem je ověřit předem postavenou hypotézu, model vztahů měřených proměnných a jejich společných faktorů - skrytých (latentních) proměnných. Než přistoupíme k provádění konfirmační faktorové analýzy, musíme vědět, jaký model vztahů chceme podrobit testování, jaké předpokládáme faktory, jaké vztahy mezi nimi. Náročnější výpočetní postup spjatý s konkrétnější představou výzkumníka před jeho spuštěním poskytují na závěr daleko cennější výsledky než explorační faktorová analýza.

13.4.3. Seskupovací analýza (cluster analysis)

Cílem této skupiny metod je najít skupiny, trsy, hrozny, "klastry" (anglicky clusters), nebo obecněji množiny jednotek, které jsou si uvnitř navzájem podobné a které se mezi sebou významnou měrou liší. Jednotlivé postupy seskupovací analýzy se mezi sebou liší podle toho, jak chápou "blízkost" jednotek a "vzdálenost" mezi skupinami, dále tím, zda množinu jednotek podle zvoleného kritéria postupně dělí, člení, nebo naopak postupně skládají, spojují. Další metody se snaží při zadaném počtu skupinek jednotky mezi těmito skupinkami v následných krocích znovu a znovu vyměňovat tak, aby nové seskupení-rozdělení lépe splňovalo kritéria "blízkosti" jednotek a "vzdálenosti" klastrů.

13.4.4. Pěšinková analýza (Path Analysis)

Metoda dráhových koeficientů, jak bývá někdy nazývána, "path analýza", nebo českým překladem bez nároku na definitivnost pěšinková analýza je metoda, která se pokouší odhalovat kauzální řetězce vztahů mezi soubory měřitelných proměnných. Je použito rozkládání a opětného skládání vazeb vyjadřovaných regresními a korelačními koeficienty mezi souvisejícími proměnnými.

13.4.5. Strukturní modelování

Metoda je zjednodušeně řečeno kombinací konfirmační faktorové analýzy a pěšinkové analýzy. Spojuje hlavní cíl, jímž je prokázání resp. ověření kauzálních (příčinných) souvislostí mezi zkoumanými

proměnnými, se snahou modelově vyjádřit a nepřímo měřit skryté (latentní) proměnné. Patří k matematicky i metodologicky velmi náročným metodám, které navíc vyžadují značnou dávku teoretického vhledu do řešené problematiky. Nejznámější programový prostředek pro aplikaci strukturního modelování je program LISREL.

13.4.6. Mnohorozměrné škálování

Cílem této skupiny metod je odhalení struktury vztahů v datech, a to jednak mezi jednotkami ve zkoumaném souboru, jednak, v jinak formulované úloze, mezi proměnnými - znaky. Metrické nebo nemetrické škálování, ale i příbuzné metody, jako je korespondenční analýza, bývají často spojovány s grafickými výstupy a slouží také přímé interpretaci výsledků na základě grafické prezentace výsledků těchto metod. Korespondenční analýza jde ještě dále, hledá vztahy mezi souborem zjišťovaných proměnných a souborem zkoumaných jednotek.

13.4.7. Logaritmicko-lineární modelování

Tabulky třetího, čtvrtého a vyšších stupňů třídění mají tu nevýhodu, že rychle klesají počty jednotek v jednotlivých polích tabulky vyjadřující vyšší stupně třídění. Logaritmicko-lineární modely jsou pokusem, jak tuto nesnáz odstranit a jak umožnit zkoumat relativně složité závislosti několika kategorizovaných proměnných, jejich samostatný i společný, vzájemně podmíněný nebo naopak vylučující se vliv.

Výčet metod ani strategií analýzy dat neobsahuje celou řadu neméně významných a neméně známých metod. Snahou bylo spíše naznačit než dopovědět. Prostor, který poskytují tato skripta neumožnil širší rozvedení problematiky. Čtenáři necht' laskavě vyhledají podrobnější vysvětlení v učebnicích statistiky a ve speciálních pojednáních týkajících se jednotlivých metod.

14. Prezentace výsledků sociologického výzkumu

14.1. Zásady prezentace dat z výzkumu

Před psaním výzkumné zprávy je třeba určit:

- pro koho je zpráva určena;
- jaký je účel výzkumné zprávy nebo publikace;
- jakým způsobem bude zpráva publikována, jaká bude forma sdělení;
- jakým stylem bude zpráva napsána;
- co vše bude zpráva obsahovat a jaká bude struktura zprávy;
- jaké mravní závazky výzkumný tým převzal a je povinen jim výzkumnou zprávu a publikaci přizpůsobit
(Srovnej [Lin 1976: 357nn])

14.1.1. Okruh čtenářů

- vědci v základním výzkumu:
pak musí být v publikaci kladen důraz na příspěvky, které výzkum přinesl vědeckému zkoumání;
- vědci pracující v aplikovaném výzkumu:
publikace by měla obsahovat informace o tom, jak přispěl výzkum k řešení konkrétního problému;
- praktici, pracovníci soukromých firem, podnikatelé:
publikace by měla uvádět údaje zajímaví jejich zákazníky, zdůraznit, čím je výzkum důležitý pro jejich klienty;
- úředníci státní správy:
neměly by chybět informace, jak mohou výsledky výzkumu pomoci v práci úřadu;
- sponzoři výzkumu (ti, co jej financují):
zajímají se o technické detaily, je třeba uvést, zda studie dosáhla cílů, které si předsevzala;
- veřejnost, ostatní čtenáři:
výsledky výzkumu musí být sděleny co nejsrozumitelnější formou, aby jim rozuměl i čtenář, který není zvyklý číst zprávy z výzkumů.

14.1.2. Účel publikace

Účelem publikace je podmíněn obsah i zvolená forma publikace, důraz, který volíme, hloubka rozboru jednotlivých okruhů zjištěných

faktů. Rozeznáváme spolu s Nanem Linem [1976:359] tři základní účely publikace z výzkumu :

1. shrnout (sumarizovat) všechno, co bylo zjištěno o daném předmětu
2. popsat nově zjištěný poznatek, doplněk informace k existujícímu souhrnu znalostí:
 - buď publikovat doplňující poznatek;
 - nebo zveřejnit nové paradigma, popsat revoluční teorii, prosadit nové pojmy, nové vysvětlení
3. ukázat praktickou užitečnost zjištěných informací, prokázat jejich využitelnost pro praktiky: informuje čtenáře o důsledcích teorií a metod, které mohou být čtenáři užitečné v jeho práci.

14.1.3. Forma sdělení a způsob publikace

Publikace z výzkumu může mít různou formu. Bude se lišit rozsahem, stupněm odbornosti, hloubkou rozboru, stupněm aktuálnosti, pedagogickou vhodností a mnoha dalšími charakteristikami. V zásadě se nejčastěji setkáváme s následujícími formami sdělení, s následujícími publikacemi z výzkumu:

- kniha - monografie
- stať v odborném časopise
- výzkumná nebo technická zpráva
- učebnice
- populární kniha
- článek v populárním časopise
- novinová zpráva

Pro zveřejňování výsledků výzkumu platí některá pravidla, která jsou ve vědeckém společenství běžně dodržována:

- jeden text by měl být v jednom okamžiku nabízen jen jedné redakci;
- výsledky výzkumu by měly být nejprve zhodnoceny kvalifikovaným vědcem a publikovány ve vědeckém časopisu nebo alespoň odborně oponentovány ve formě výzkumné zprávy, a teprve potom je možno dát je k dispozici sdělovacím prostředkům a veřejnosti ve formě novinové zprávy nebo populárního článku, které nutně zjednoduší a mohou v případě omylu nebo metodické nepřesnosti způsobit více škody než užítku.

14.1.4. Styl

Pojednání o výsledcích výzkumu je možno psát

- v první osobě jednotného čísla: "...zjistil jsem..." nebo
- ve 3. osobě v pasivním tvaru: "...bylo prokázáno..."
- stále častěji se však užívá aktivního tvaru: "...data ukazují..." podle Nan Lina [1979:111].

Raději používáme jednodušší větnou skladbu. Složitě strukturovaná souvětí jsou čtenářům nesrozumitelná. Pokud je porozumění jen částečné, může být smysl textu pochopen nesprávně.

Pokud text obsahuje obrázky a grafy, text by měl být pochopitelný i bez nich. Přitom grafy a obrázky jsou srozumitelnější a je lépe jim dát přednost před tabulkami a grafy.

14.1.5. Obsah výzkumné publikace

1. Formulace problému
2. Popis metodiky výzkumu [výběr, uspořádání výzkumu (Research Design) a použité metody]
3. Analýza a interpretace
4. Závěry a diskuse
5. Bibliografie

14.1.6. Mravní odpovědnost

Autoři výzkumné zprávy jsou povinni chránit dotázané a spolupracovníky před nechtěnými vedlejšími účinky publikace výsledků. Je třeba dodržet slib důvěrného zacházení s daty.

Výzkumník je povinen publikovat přesné a úplné informace. Rozumí se tím povinnost dbát, aby se autoři zprávy při poskytování údajů neomezovali jen na pozitivní výsledky, aby smysl uveřejněných výsledků nebyl zkreslován jejich záměrným výběrem nebo zamlčováním. Neúplná prezentace dat s možností jejich dezinterpretace škodí nejen vývoji vědeckého poznání, ale také diskredituje výzkumníka. Může tvořit podstatnou překážku v pokračování započatého výzkumu nebo v jeho opakování nebo prověřování.

14.2. Struktura výzkumné publikace

Jaké oddíly, části, kapitoly by měla obsahovat výzkumná zpráva nebo jiná publikace z výzkumu? V zahraničních odborných časopisech se

ustálila struktura, která je požadována po vědeckém sdělení z výzkumu. Také výzkumná zpráva musí vždy obsahovat některé oddíly, aby byla úspěšně přijata zadavatelem a odborníky. Výzkumná zpráva, článek nebo monografie se od sebe v podrobnostech své struktury liší. Přesto se pokusím jmenovat základní oddíly publikace z výzkumu v obecné podobě a v komentáři pak rozlišit, v kterých případech je nutno daný oddíl zařadit a kdy to není nutné. (Srovnej viz. [Bogardus 1933:213-221] a [Kane 1985: 176-200].)

1. Titul	Titul Abstrakt Předmluva
2. Obsah	Obsah
3. Text	Úvod Metodika a metodologie Vlastní text Závěry Doporučení pro další výzkum
4. Bibliografie	Přílohy Poznámky Slovníček Bibliografie Index

14.2.1. Titul

Titulu publikace věnujeme nejčastěji samostatnou stránku. Není tomu tak u odborných statí, ale je třeba to dodržet u výzkumných zpráv, disertačních prací, učebnic, monografií a pod.

Abstrakt je stručným výtahem, obsahem knihy nebo článku. U odborné statí, kde se užívá nejčastěji, by neměl přesahovat 10 řádek. Bývá přeložen do cizího jazyka, nejčastěji do angličtiny. Charakterizuje v několika větách předmět výzkumu, použité metody a výsledky výzkumu.

Předmluva bývá zařazována obvykle jen u dlouhých prací. Obsahuje sdělení autora čtenářům a zvláštní místo v ní bývá věnováno poděkování spolupracovníkům a těm institucím, které přispěli materiálem, radou nebo pomocí ke zdaru výzkumu nebo k publikaci díla. Pokud by předmluva obsahovala jen tuto část, lze ji nahradit odstavcem nazvaným "Poděkování".

14.2.2. Obsah

V rozsáhlejších pracích, které mají charakter knihy, je třeba uvést obsah. Má podobu stručného výčtu kapitol nebo několikastránkového strukturovaného seznamu oddílů, kapitol, subkapitol a paragrafů s čísly stránek u každého z nich.

V případě potřeby bývá připojen seznam tabulek zařazených do textu a seznam ilustrací. Oba seznamy mají za cíl usnadnit čtenáři přehled v rozsáhlé monografii nebo učebnici. [Kane 1977:183]

14.2.3. Text

Úvod obsahující seznámení s tématem a okolnostmi výzkumu zařazujeme v delších pracích jako samostatnou část uvedenou někdy nadpisem. V kratších textech zařazujeme úvod jako první paragraf vlastního textu zprávy. V delší práci má úvod tyto funkce:

- vysvětlit teoretický základ studie
- podat krátký přehled literatury
- vysvětlit, o čem studie pojednává
- v technických zprávách z výzkumu úvod obsahuje informace o
 - použitých termínech a pojmech
 - historii zkoumaného předmětu
 - důvodech, motivech výzkumu
 - omezeních, která se týkají získaných výsledků.

Kapitola nebo paragraf nazvané "Metodika" nebo "Metodologie" tvoří povinnou součást výzkumné zprávy, odborné stati o výsledcích výzkumu i dalších výzkumných publikací. Sdělují čtenářům, zvláště odborníkům, jakými postupy byly zveřejněné výsledky získány. Patří do publikace především proto, aby zajistily důvěryhodnost uvedených údajů. Kapitola obsahuje údaje o tom, jaký typ výzkumného projektu byl použit, jaký rozsah a jaký typ výběru byly zvoleny, jaké bylo vybráno uspořádání výzkumu, jaké techniky sběru dat byly použity, jaké byly použity výpočetní programy a jakými metodami byla provedena analýza dat a na závěr s jakými problémy se výzkumníci v průběhu výzkumné akce setkali. V méně rozsáhlém textu je možno údaje o metodice zařadit jako zvláštní paragraf do úvodu. Ani v tom případě však nesmí chybět základní údaje o velikosti výběrového souboru, způsobu výběru, metodě sběru dat a použitých programech a analytických metodách.

Vlastní text je účelně členěn do kapitol a paragrafů, které mohou být zakončeny dílčími závěry. Účelem textu je seznámit čtenáře s výzkumným materiálem. Osnova textu může vycházet z osnovy zkoumání:

Příklad:

Název práce: Formální a neformální struktura vztahů mezi majiteli a zaměstnanci nově vznikajících soukromých firem.

Kapitoly vlastního textu výzkumné publikace:

1. Vznik soukromých firem
2. Majitelé firem
3. Zaměstnanci
4. Struktura vztahů v soukromých firmách
 - formální struktura vztahů
 - neformální struktura vztahů
 - vztahy mezi formální a neformální strukturou.

Při formulaci vlastního textu zprávy je možno se řídit některými zásadami [Kane 1985:186]:

- a) použijeme spíše formulace: "materiál byl sbírán v letech 1989 - 1990" než "sbíral jsem...";
- b) vynecháváme hodnotící přídavná jména "krásný", "optimální", "vyjímečný", raději podrobněji popíšeme daný jev a ponecháme hodnocení na čtenáři;
- c) použijeme grafy, tabulky, mapy a ilustrace kdykoli to usnadní sdělení.

Závěry by měly obsahovat krátké shrnutí hlavních myšlenek nebo faktů nebo výsledků analýz a jejich širších důsledků. Nezveřejňujeme závěry, které nemůžeme doložit. Neopomeneme naopak sdělit závěry, které jsou v protikladu k našim očekáváním nebo jsou pro nás nežádoucí.

Doporučení pro další výzkum netvoří povinnou součást výzkumné publikace. Můžeme v něm uvést domněnky, které se nám nepodařilo dokázat, nedostatky vlastního výzkumu, zkušenosti z výzkumu a doporučení, kde lze ve výzkumu pokračovat.

14.2.4. Bibliografie

Přílohy tvoří součást každé výzkumné zprávy nebo rozsáhlejší publikace. Uspořádáme do nich všechny materiál, který je důležitý pro studii, ale byl příliš rozsáhlý, než aby mohl tvořit součást textu kapitol. Jako přílohu (apendix) zařadíme do zprávy dotazník nebo jiný použitý instrument, text průvodního dopisu, rozsáhlejší tabulky, grafy a další dokumentaci výzkumné akce, která je potřebná čtenářům

specialistům pro zhodnocení našich výsledků i pro pokračování ve výzkumu. Do příloh však není ani zdaleka možno dát všechno, co jsme v terénu zjistili ani co jsme spočetli.

Poznámky tvoří další významnou součást odborného textu. E.Kaneová [1985:193] rozlišuje dva druhy poznámek. Odkazovými poznámkami upozorňujeme na materiál, který není náš vlastní. Citujeme-li, používáme k označení uvozovky "... " nebo apostrofy '...'. Také ústní materiály je třeba citovat. Pokud neznáme přesné znění citátu, výrok parafrázujeme, tj. nedáváme jej do uvozovek, ale opatřujeme jej poznámkou. Důvodem odkazových poznámek je přiznat autorství myšlenky tomu, komu náleží a umožnit čtenáři důkladněji se seznámit s originálním textem.

Druhým typem poznámek jsou vysvětlivky, obsahové poznámky. Uplatňují se všude tam, kde by odbočky z hlavní linie vysvětlení rušily hladký průběh textu. Neobsahují však odkazy na jinou práci.

Slovník obsahuje vysvětlení speciálních výrazů a cizích slov obsažených v textu. Ve většině publikací je neobvyklý. Setkáme se s ním v některých učebnicích, příručkách, manuálech. Termíny ve zprávě většinou vysvětlíme přímo v textu. Pokud se však termíny vyskytují opakovaně na mnoha místech studie, připravíme slovníček termínů v abecedním pořadí.

Bibliografie tvoří povinnou součást prakticky všech výzkumných publikací, odborných statí, monografií a dalších textů. Obsahuje abecední seznam prací zmíněných ve zprávě. U každé práce je uveden úplný bibliografický údaj umožňující najít v práci na příslušné straně také citaci, na níž naše studie odkazuje.

Index neboli rejstřík je uváděn obvykle jen v knihách, učebnicích, monografiích. Slouží lepší orientaci čtenáře, který hledá údaje o určitém tématu (k tomu slouží věcný rejstřík), nebo o určité historické osobě (k tomu slouží jmenný rejstřík). Vyjimečně se vyskytují také rejstříky zaměřené na místní názvy nebo další speciální označení.

14.3. Grafická prezentace, grafy, schémata

Výsledky výzkumu je možno popsat komentářem. Je však řada podrobností, které by si vyžádaly příliš mnoho slov pro své vyjádření. Velké množství informací obsahují tabulky právě tak jako souhrnné statistické charakteristiky a koeficienty. Pro běžného čtenáře jsou však málo srozumitelné. Rozumět jim vyžaduje znalosti statistiky a sociologické metodologie.

Oproti tomu názorné a všem srozumitelné, je sdělování výsledků pomocí grafů a schémat. V sociologickém výzkumu se užívá mnoho různých způsobů grafické prezentace výzkumných dat. Některé jsou vhodné k vyjádření vývoje v časové řadě, jiné se hodí ke komparaci hodnot dvou nebo několika veličin, další používáme k porovnávání skladby souborů, ke srovnání, v jakém poměru jsou v nich obsaženy jednotlivé složky. Přehled typů grafického znázornění srovnej s [Bogardus 1933:196-211].

14.3.1. Vývojový čárový graf - line chart

Je vhodný pro znázornění vývoje veličiny v čase.

14.3.2. Sloupkový graf - histogram - bar chart

Rozeznáváme horizontální a vertikální sloupkový graf. Hodí se k prezentaci kategorizovaných dat. Lze pomocí něj porovnat jak vývoj veličiny, tak rozložení četností v souboru podle hodnot kategorizovaného znaku. Výška sloupce může také vyjadřovat hodnotu znaku a jednotlivé sloupce mohou reprezentovat jednotlivé podsoubory.

14.3.3. Srovnávací sloupkový graf - srovnávací histogram - comparative bar chart

Slouží k porovnání vývoje nebo rozložení hodnot znaku ve dvou nebo několika souborech.

14.3.4. Složený sloupkový graf - composite bar chart

Graficky vyjadřuje složení souboru z jeho částí, ukazuje poměr mezi těmito částmi v souboru, každá část je reprezentována poměrnou výškou sloupce vyšrafovanou jiným způsobem.

14.3.5. Srovnávací složený graf - comparative composite chart

Lze jej výhodně využít k porovnání složení dvou nebo několika podsouborů z jejich částí. Srovnáváme v jakých poměrech jsou tyto části zastoupeny v celku každého souboru reprezentovaného složeným sloupcem.

14.3.6. Rozvinutý sloupkový graf - extended bar chart

Pokud jsou rozdíly mezi veličinami, které chceme znázornit pro jednotlivé hodnoty znaku výrazně rozdílné, můžeme ke grafickému znázornění poměru mezi těmito hodnotami využít rozvinutý sloupkový graf. Princip spočívá v tom, že sloupec, který je příliš dlouhý a nemůže pokračovat zalomíme v grafu dolů a zpět a v následujícím řádku jej prodloužíme opačným směrem, a to tak dlouho, až grafické znázornění odpovídá.

14.3.7. Kruhový velikostní graf - kruhový diagram -

circle chart

Velikost kruhu znázorňuje počet jednotek nebo množství, které kruh reprezentuje. Při interpretaci je třeba vzít v úvahu, že poměru mezi četnostmi mohou odpovídat poměry mezi průměry kruhů nebo poměry mezi jejich plochami.

14.3.8. Kruhový dělený graf (koláčový graf) - pie chart

Graf vyjadřuje, v jakém poměru jsou jednotky souboru rozděleny do jednotlivých hodnot znázorněného znaku. Kruhové výseče znázorňují jednotlivé hodnoty znaku a plocha každé výseče je úměrná četnosti s jakou je hodnota zastoupena v souboru. Některá z kruhových výsečí může být zvýrazněna vyčleněním z kruhu.

14.3.9. Graf rozptýlení (krabičkový graf) - box chart

Graf rozptýlení znázorňuje graficky nejdůležitější statistické charakteristiky souboru. Úsečka vyjadřuje hodnoty znaku mezi krajními hodnotami označenými kroužky. Padesát procent četností je vyjádřeno

obdélníkem mezi dolním a horním kvartilem. Další úsečka přetínající tento obdélník označuje medián. Mimo to graf může obsahovat další důležité body zpřesňující informaci o rozptýlení. Někdy bývá pro účely srovnání vedle sebe znázorněno několik grafů rozptýlení pro jednotlivé podsoubory.

14.3.10. Stromový graf - dendrogram

Stromový graf je používán pro znázornění členění souboru na podsoubory. Využíváme jej při typologické proceduře, při seskupovací analýze hierarchického typu (cluster analysis), kdy graf znázorňuje postupné členění souboru na podsoubory. Jindy tentýž graf naopak znázorňuje postupné hierarchické slučování sobě blízkých shluků jednotek. Samotným spojením větví ukazuje sloučení podsouborů a délkou jednotlivých větví hladinu, na níž se podsoubory slučují.

14.3.11. šipkový graf - causal diagram and flowgraph

šipkový graf má nejčastější použití při grafickém vyjádření příčinných (kauzálních) závislostí. Znázorňuje působení příčiny na následek ve směru šipky. Koeficient připsaný u šipky vyjadřuje sílu tohoto působení. šipkové grafy se uplatňují při použití parciálních korelačních koeficientů, v pěšinkové analýze (Path Analysis), v analýze strukturních rovnic (Structure Equation Analysis) a v dalších postupech kauzální analýzy.

14.3.12. Znaménkové schéma

Dvojrozměrná tabulka četností není příliš vhodným způsobem ke znázornění souvislosti mezi dvěma tabelovanými znaky. K tomuto účelu se jako nejnázornější prostředek osvědčilo znaménkové schéma. Znaménka + a - v jednotlivých políčkách stejně velké tabulky znaménkového schématu ukazují, zda četnost v tomto poli tabulky přesahuje "+" nebo nedosahuje "-" četnosti odpovídající třídění dvou nezávislých znaků. Významně vyšší zaznamenaná četnost v dané kombinaci hodnot obou znaků je znázorněna jako " ++ " , vysoce významně vyšší jako " +++ ". Významně nižší četnost jako " -- ", vysoce významně vyšší jako " --- ".

Obdobným typem grafu jako je graf korelační je možno zobrazit korespondenci mezi typy jednotek a hodnotami třídícího znaku a další typy vztahů mezi jednotkami, např. vztahy podobnosti nebo vztahy vzdálenosti mezi jednotkami z hlediska dvou vzájemně nezávislých rozměrů.

14.3.13. Blokovaný diagram - block diagram

Blokový diagram schematicky znázorňuje posloupnost kroků vyjadřovaného postupu a význam těchto kroků, jejich funkci. Druhou možnou funkcí blokového diagramu je znázornění skladby, členitosti složitě analyzované struktury.

14.3.14. Mnohojednotkový graf - multiple unit chart

Graf znázorňuje množství jednotek nebo jinou kvantitativní charakteristiku pro podsoubory. Předem stanovené kvantum je v grafu zastoupeno jedním symbolem (postavou, sudem, stromkem, komínem, hvězdičkou). Poměr mezi množstvím znázorněných symbolů pro jednotlivé podsoubory odpovídá poměrům mezi množstvím jednotek v těchto podsouborech nebo poměrům mezi hodnotami znaku pro tyto podsoubory (např. 1 postava = 10 000 lidí, 3 postavy = 30 000 lidí a pod.)

14.3.15. Trojrozměrný graf

S rozšířením grafických programů pro osobní počítače se začaly objevovat složitější grafy generované počítačem ze vstupních dat. Zvláštní skupinu mezi těmito grafy tvoří trojrozměrné grafy. Je možno jimi vyjádřit srovnání současně ve dvou rozměrech, např. počty jednotek zastoupených pro kombinace hodnot dvou kategorizovaných znaků.

14.3.16. Korelační graf

Vztah mezi dvěma znaky je možno vyjádřit grafem, na němž jednotlivé jednotky v souřadné soustavě dané hodnotami obou znaků představují body. Seskupí-li se body kolem přímky, existuje mezi znaky lineární závislost. Ostatní typy vztahů jsou graficky složitější.

14.3.17. Korespondenční graf, graf podobnosti, vzdálenosti

15. Etické a profesionální zásady sociologa výzkumníka

15.1. Co je a co není "etické", "morální"?

Být morální znamená chovat se podle obecně přijatých profesionálních zvyklostí. Jisté není mravné někoho poškodit (např. zkoumaného, pozorovaného, dotázaného).

Eileen Kaneová uvádí oddíl své studie věnovaný etice příklady rozhodování výzkumníka v situaci, kdy si musí vybrat, zda a nakolik přísně bude hájit abstraktní princip a nebo zda a nakolik přijme nabízející se řešení usnadňující mu cestu k splnění jeho představ. [Kane 1985:211]

1. Zkoumáte oblast, v níž je hodně lidí odkázáno na sociální podporu. Víte, že stěží seženete dostatek prostředků na výzkum. Peníze na výzkum vám nabídne organizace, o níž víte, že podporuje rasismus a že by byla schopna použít vaše data mimo vědecký rámec. Rozhodnete se přijmout peníze, když víte, že jinak se výzkum neuskuteční?
2. Vaším úkolem je zjistit, zda nový systém rozdělení pravomocí v podniku zlepšil postoje dělníků k továrně. prokázat se vám to nepodařilo, ale myslíte si, že systém je demokratičtější a spravedlivější. Abyste mohl doporučit tento nový systém, nemusíte nic falšovat, měnit ve svých datech. Pouze některé výsledky nezveřejníte. továrna bude jistě lidštějším místem pro pracující. Uděláte to?
3. Zkoumáte, jak prakticky lidé jednají v případě ohrožení. Oznámíte lidem v obchodním domě, že v jednom z oddělení je časovaná nálož, která hrozí výbuchem. Tato vaše informace bude užitečná pro trénink (návlek) lidí, jak se zachránit před reálným nebezpečím. Udělal, udělala byste to?

Mravní odpovědnost je založena na kombinaci citlivosti, ohleduplnosti a zdravého selského rozumu. Zčásti je vymezena právními předpisy. Tuto mravní odpovědnost lze rozlišit na:

1. odpovědnost k tomu, koho zkoumáš;
2. odpovědnost k vědě, profesi;
3. odpovědnost ke sponsorovi a zadavateli výzkumu.

[Kane 1985:211-214]

15.2. Jak je možno zachovat mravní normy?

Jak je možno zachovat mravní normy v situacích, kdy zájem výzkumníka je v rozporu se zájmem zkoumaného objektu (podle [Bailey 1978:382-386]):

1. studiem zvířat;
 - nestudujeme lidi abychom je výzkumem nepoškodili, neznáme vedlejší účinky našich experimentů, a proto provádíme výzkum na zvířatech;
2. počítačovými simulacemi
 - namísto studia reálných objektů modelujeme zkoumané procesy na počítači; tak lze na základě umělé napodobovaných dějů, mechanismů odhadovat pravděpodobné chování lidí v situacích ohrožení jejich života nebo zdraví;
3. studujeme jev za podmínek, kdy negativní vliv působí bez přičinění výzkumníka
 - příkladem může být studium vlivu přírodních katastrof nebo havárií na člověka, jeho chování bezprostředně po události;
4. necháme působit negativní vliv v malé intenzitě a po velmi krátkou dobu
 - je-li prokázáno, že škodlivost vnějších podnětů pro člověka je při nízké intenzitě jen mizivá, můžeme obhájit výzkum působení těchto vlivů v malé intenzitě a po velmi krátkou dobu, pokud tím pomůžeme řešit naléhavější problémy společnosti a člověka;
5. předem informujeme dotázaného (zkoumaného) o možných negativních vlivech a zajistíme si jeho souhlas
 - pokud negativní působení zkoumaných vlivů není životu nebezpečné a pokud je značná šance předejít na základě výsledků výzkumu větším škodám, hledáme dobrovolníky, které včas a úplně informujeme o možných negativních vlivech pokusu a kteří přesto souhlasí se svou účastí;
6. na vybrané otázky nevyžadujeme odpověď od všech dotázaných
 - na otázky, které zasahují příliš do soukromí dotázaného, nevyžadujeme odpověď od všech respondentů; provedeme náhodný podvýběr, jemuž tyto otázky do dotazníku zařadíme, ostatním je vynecháme;
7. negativní vliv na zkoumané osoby je ospravedlnitelný tím, že výsledky výzkumu budou mít významně vyšší hodnotu ve srovnání s momentálním vyrušením respondentů
8. zaručíme zachování soukromí osob publikováním jen agregovaných dat.

15.3. Tři příklady projektů, které se prohřešily proti etice sociologické výzkumné práce (podle [Burges 1984:185-189])

1. Projekt Kamelot ¹⁵⁾
 - byla porušena důvěryhodnost výzkumného týmu, který získával v zemích Jižní Ameriky bez vědomí vlád těchto zemí data o společenských procesech, sociální struktuře a aktivitách sociálních skupin. Tyto údaje měly sloužit ne vědeckému poznání ale k použití americkou výzvědnou službou k potlačení lidových hnutí.
2. Studie homosexuálního chování
 - bylo porušeno soukromí studovaných osob; Land Humprey [1970] prováděl zúčastněné pozorování homosexuálů a mimo to pod záminkou studia kulturních zájmů a činností tytéž osoby dotazoval na řadu osobních údajů, které využil v kombinaci s prvními daty.
3. Studie komunity ve Springdale
 - nebyla dodržena anonymita zkoumaných osob; výsledky zkoumání na malém městě byly zveřejněny v neagregované podobě a ze souvislosti bylo možno identifikovat jednotlivé osoby. ¹⁶⁾

Tyto tři americké výzkumné projekty ilustrují některá hlavní etická dilemata sociologického výzkumu.

15.4. Základní body etického kodexu Americké sociologické asociace (podle [Bailey 1978: 447nn] Appendix, Part B: Code of Ethics of the American Sociological Association)

1. Zachovávat vědeckou objektivitu výzkumu
2. Nepřekračovat své schopnosti a kompetence
3. Respektovat práva zkoumaných osob na soukromí a důstojnost
4. Chránit zkoumané před osobním poškozením
5. Zaručit důvěrné zacházení s výzkumnými daty
6. Zajistit nezkreslenou prezentaci výzkumných výsledků
7. Nezneužívat role výzkumníka k jiným (neprofesním) účelům
8. Poděkovat všem spolupracovníkům
9. Nezatajit žádné zdroje finanční podpory výzkumu a možná zkreslení

10. Povinnost zveřejnit každé ovlivnění výsledků sponsorem nebo uživatelem či zadavatelem
11. Povinnost nepřijmout finanční podporu výzkumu, která by odporovala těmto etickým principům a ukončit spolupráci na výzkumu, který principům odporuje

15.5. Podmínky publikace výsledků výzkumu

(podle [Bailey 1978:449nn] Appendix, Part C: Code of Professional Ethics and Practices of the American Association for Public Opinion Research)

Minimální soubor údajů, které je třeba uvést při každé publikaci výsledků výzkumu:

1. Kdo financoval výzkum
2. Přesné znění položených otázek
3. Definice populace, z níž byl prováděn výběr
4. Velikost výběru, způsob výběru a návratnost
5. Jak odhadnout výběrovou chybu
6. Které výsledky jsou založeny jen na části výběru
7. Zda dotazování probíhalo osobně, telefonicky, poštou; doma nebo na ulici
8. Časové zařazení doby dotazování vzhledem k významným událostem.

16. Proces rozhodování v sociologickém výzkumu

Sociologický výzkum spočívá v řetězu rozhodnutí, které určují postup práce i dosažené výsledky. V každém rozhodovacím kroku má sociolog možnost otevřít nebo naopak uzavřít množství strategií dalšího postupu. Výzkumník musí mimo jiné rozhodnout jaký problém bude řešit, v jakém pojmovém rámci se bude pohybovat, jakým způsobem zvolené pojmy vyjádří ve znacích a jak je bude měřit a zjišťovat, na jaké populaci a pomocí jakého výběru bude získávat data, jaký zvolí typ výzkumného projektu, jaké metody sběru dat a analytické procedury se rozhodne použít, jak bude výsledky interpretovat a pomocí jakých prostředků je bude prezentovat ve výsledcích. To vše ovlivní výslednou podobu a míru poznání zvolené části sledované skutečnosti i konečný úspěch sociologického výzkumu. (Srovnej závěrečnou kapitolu práce Labovitz-Hagedorn [1976: 114-130].)

16.1. Výzkumný problém

Prvním klíčovým místem rozhodování výzkumníka je volba výzkumného problému. Závisí na dosavadních znalostech, zkušenostech a schopnosti sociologa položit otázku, která je nejen důležitá a aktuální, ale také zodpověditelná prostředky sociologického výzkumu.

Sociolog se před započítím řešení problému zabývá studiem literatury. Konzultuje s odborníky. Seznamuje se s výsledky výzkumů, které na dané nebo příbuzné téma dosud byly provedeny. Snaží se začít s výzkumem s tak velkým množstvím znalostí, jak je jen možné.

16.2. Pojmový rámec

Výzkumný problém je vždy řešen v předem zvoleném pojmovém rámci. Sociolog formuluje problém pojmy sociologické teorie, vytváří model, v jehož rámci lze daný problém řešit. S použitím zvolených obecných pojmů vytváří teoretické hypotézy a hledá k nim výroky nižšího řádu, pracovní hypotézy, které lze testovat, ověřovat na rovině empirie. K obecným pojmům konstruuje výzkumník znaky a indikátory, jejichž prostřednictvím jsou uvažované pojmy zjišťovány a měřeny, často jen nepřímou. Pracovní hypotéza je pak na rozdíl od obecné teoretické hypotézy vytvořena jako výrok o vztahu měřitelných a zjišťovaných znaků.

Rozhodovací kroky v této fázi jsou velmi odpovědné především proto, že prakticky vždy se výzkumník rozhoduje pro redukci: snižuje stupeň obecnosti uvažovaných vztahů z obecné teoretické roviny na rovinu empiricky zjišťovaných faktů, redukuje množství sledovaných souvislostí na soubor takových, o nichž předpokládá v položených hypotézách, že jsou ty nejdůležitější, rozhoduje se pod tlakem

výzkumných možností pro zkoumání jen některých vztahů, měřitelných, zjištělých souvislostí.

16.3. Populace a výběrový soubor

Další z řetězce výzkumných rozhodnutí je podmíněno nutností zkoumat jen část populace, o níž výzkum vypovídá. Nejčastěji potřebujeme výsledky výzkumu zobecnit na celou populaci. V takové situaci musíme zajistit, aby výběrový soubor reprezentoval celek, o němž vypovídáme. Nákladnost a organizační náročnost pravděpodobnostních výběrů, které jediné spolehlivě zaručují splnění tohoto úkolu, vedou výzkumníka někdy k rozhodnutí ustoupit od požadavku zaručené reprezentativnosti výběru. Toto rozhodnutí má pochopitelně své důsledky ve fázi interpretace dat.

Použijeme-li například kvótní výběr, omezujeme se při zajišťování reprezentativity jen na zvolené kvótní znaky, a to jen v použitých výběrových kombinacích. Ostatní hlediska reprezentativity jsou ponechána bez kontroly. Závisí na tazatelích a na konkrétních podmínkách výzkumné situace v terénu, jakou míru zobecnitelnosti nám výběr zaručí.

Překračovat rámec zobecnitelnosti výsledků daný výběrovým mechanismem není z profesionálního hlediska přípustné a spekulace jdoucí za rámec dat musí být teoreticky zakotveny již ve východiscích výzkumu. Rozhodnutí o způsobu a technice výběru je tedy v tomto ohledu vlastně rozhodnutí o možnostech závěrů ze získaných dat.

16.4. Výzkumný projekt - uspořádání výzkumu

Způsob uspořádání výzkumu (Research Design) představuje důležité rozhodnutí. Jinak budou sbírána data v rámci výzkumné sondy - případové studie, jinak při řízeném experimentu a jinak půjde-li o reprezentativní výběrové šetření s celostátní působností. A nejen, že budou data jinak sbírána. Jinak bude prováděn výběr zkoumaných jednotek, odlišný bude použitý výzkumný nástroj, technika a organizace sběru dat, jiné budou nároky na finanční zabezpečení výzkumu, jiné budou možnosti zobecnění výsledků. Rozhodnutí o cílech, a tedy i o uspořádání výzkumu je třeba provést co nejdříve. Zcela jistě před volbou výzkumných nástrojů pro sběr dat a v těsné souvislosti s volbou výběrového uspořádání.

Rozhodujeme se pro deskriptivní, explanační nebo prognostické cíle výzkumu. A to jistě ovlivní, co budeme zjišťovat a jak to budeme zjišťovat. Máme omezené finanční zdroje a také tomu musíme přizpůsobit

svá rozhodnutí. Časové možnosti výzkumu, počet spolupracovníků, technické zázemí, to vše souvisí s rozhodováním o typu výzkumného uspořádání. Volbu uspořádání výzkumu ovlivní také množství znalostí, které již máme k dispozici. Nesmíme však zapomenout, že rozhodující váhu v našem rozhodování by měl mít výzkumný záměr, cíl výzkumu.

Chceme-li například popsat životní podmínky rómského obyvatelstva na Mostecku, neuspořádáme za tím účelem celostátní reprezentativní dotazníkové šetření. Soustředíme se na danou lokalitu, na populaci Romů a jistě nevystačíme s dotazníkem. Naopak, chceme-li předpovědět výsledky voleb do parlamentu, musíme se pokusit o reprezentaci celé populace státu. Chceme-li znát reakce skupiny na nepopulární rozhodnutí jejího vedoucího, můžeme uspořádat experiment. Reakce celého národa na nepopulární rozhodnutí vlády je však dost obtížné tímto způsobem zkusmat.

Rozhodujeme-li o výzkumném uspořádání, rozhodujeme o pravděpodobném typu výběru, o finálních nárocích, nejpravděpodobnějších nástrojích sběru dat, o potřebě spolupracovníků, technického zázemí, o časových nárocích výzkumu. Význam rozhodnutí je tedy klíčový a v mnohém jej bez značných ztrát nelze v průběhu výzkumu revidovat. Je tedy třeba mu věnovat patřičnou pozornost již na počátku výzkumu.

16.5. Metody sběru dat

Rozhodujeme se mezi přímou nebo nepřímou technikou zjišťování, mezi technikou formalizovanou, standardizovanou nebo naopak nestandardizovanou. Vybíráme techniku, která respondenta ovlivňuje nebo raději volíme neagresivní tzv. "unobstrusive" techniku, tj. pozorování nebo některou jinou techniku, při níž se nesnažíme respondentovi vnutit svůj pohled na svět. Zůstáváme mimo svět lidí, které studujeme nebo se stáváme jeho součástí, používáme zúčastněnou techniku. Volbu alternativy mezi těmito variantami vždy provází důsledek, v němž se mísí výhody s nevýhodami. V jednotlivých kapitolách byly obě stránky každého takového výběru podrobně diskutovány. Obecnou preferenci doporučit nelze. Lze však doporučit kombinaci několika různých technik, pokud to okolnosti výzkumu umožňují. Dojdeme-li různými, mezi sebou výrazně se lišícími postupy k obdobnému výsledku, naše přesvědčení o jeho nenáhodnosti může být mnohonásobně vyšší.

16.6. Příprava a čištění dat

Rozhodnutí o bezchybnosti dat a možnosti jejich veřejné prezentace a použití k ověřování výzkumných hypotéz je významným krokem, který přes veškerý tlak okolností není možné uspěchat. Etapa přípravy dat ke strojně početnímu zpracování a kontroly těchto dat, včetně nutných oprav v datech je ve většině výzkumných šetření nepříjemně dlouhá. Aktuální data zastarávají. Zadavatel se snaží získat alespoň předběžné, nejzajímavější výsledky, základní údaje pokud možno ihned, co nejdříve. Právě v tomto tlaku je skryto velké nebezpečí pro výzkumný tým. Právě tyto předběžné, první, netrpělivě očekávané výsledky našeho výzkumu je často to jediné, co má svůj celospolečenský dopad. Podle jejich kvality, platnosti a bezchybnosti je náš výzkum veřejnosti posuzován. Nepřesnost, nebo také podstatná chyba, přehlédnutá ve spěchu nebo dokonce způsobená nedokonalým dodržením kontrol v časovém stresu, mohou způsobit neúspěch celé výzkumné akce a ztrátu důvěry zadavatele i veřejnosti ve výzkumný tým.

Před skončením etapy přípravy dat a před dokončením kontrol a oprav proto údaje ani výsledky výzkumu nezveřejňujeme. Abychom splnili očekávání spojovaná s našim výzkumem, musíme této etapě věnovat dostatek pozornosti. Naší snahou bude její zkrácení, maximální urychlení, ale ne na úkor kvality dat, spolehlivosti kontrol a úplnosti provedených oprav. Rozhodnutí o zveřejnění výsledků výzkumu bude následovat bez zbytečných odkladů ihned po skončení oprav a nezbytných analýz. A mělo by být spojováno s důvěrou ve zveřejňovaná data a jejich platnost.

16.7. Analýza a interpretace

Základní možnosti analýzy a tedy i interpretace jsou určovány řetězcem předchozích výzkumných rozhodnutí. Volba problému, hypotéz, výběru a výzkumného uspořádání značně zužuje prostor pro volbu metod a postupů analýzy. Dalším limitem může být použitá metoda sběru dat a v ní aplikované výzkumné instrumenty. V procesu analýzy dat jde především o to, skutečně využít analytické možnosti, které sebraná data poskytují. Rozhodnutí pro metodu analýzy dat musí odpovídat typu dat a požadavkům metody. Splnění těchto pravidel by mělo být samozřejmostí. Častěji než z neznalosti předpokladů a omezení jednotlivých metod a postupů bývají pravidla aplikace daných metod porušována z lehkomyšlné přezíravosti, ze snahy použít efektní, často módní techniku nebo prostě z neznalosti jiných, pro danou úlohu vhodných analytických postupů. Rozhodnutí použít danou metodu by tedy

mělo být opřeno o splnění jejich předpokladů a dále o přesvědčení, že vhodnější metodu nemáme k dispozici.

Druhé, kvalitativně významnější rozhodnutí v etapě analýzy dat a jejich interpretace, je rozhodnutí nepokračovat dále v analýze. Měli bychom si při něm uvědomit, že téměř vždy, ve velké většině sociologických výzkumů, zůstává značná část sebrané informace neanalyzována. Mnoho vazeb mezi zjišťovanými znaky nebylo podrobeno rozboru buď vůbec nebo jen povrchně. Rozhodnutí ukončit pro tuto fázi analýzu dat může být rozhodnutím konečným, a aniž si to připouštíme, v mnoha výzkumech jím skutečně bývá. Je to tedy rozhodnutí významné.

16.8. Prezentace a závěry

Rozhodnutí prezentovat výsledky výzkumu je přirozené. Důležité je rozhodnout se o podobě této prezentace, jejím rozsahu, určení, tj. volbě čtenářské populace. Často bývá prezentace výsledků z výzkumu prováděna postupně, nejprve v podobě odborné stati nebo referátu na konferenci, dále v ucelenější, širší podobě výzkumné zprávy a konečně jako vědecká monografie většinou překračující výsledky jediného výzkumu a reagující na literaturu a ostatní výzkumy k danému problému. Rozhodnout je tedy třeba o době i způsobu prezentace, o spoluautorech, případně o vhodnosti zveřejnění některého z výsledků jen odborné nebo i laické veřejnosti.

Prezentace výsledků výzkumu musí vzít v úvahu také etické aspekty. Zachování anonymity, kterou jsme dotázaným přislíbili, může nabývat nejrůznějších konkrétních podob. Zveřejnění výsledků sociometrického šetření o vztazích na pracovišti, které nám mohou připadat anonymní, zúčastěným osobám poskytuje daleko více informací a může někoho z nich, aniž bychom chtěli, poškodit. Právě tak výsledky dotazování v malém městě, kde třicetiletý rozvedený lékař je zcela určitá jedinečná osoba, vedoucí pracovník s více než sto podřízenými je jediný ředitel ve městě a pracovnice ve spojích je místní poštovní doručovatelka, není možno prezentovat stejně bezstarostně jako údaje o průměrech za stovky a tisíce dotázaných v celostátním výběrovém šetření.

Rozhodnutí o prezentaci výsledků je v neposlední řadě často také rozhodnutí politické. Může mít význam pro utváření veřejného mínění, v některých případech může ovlivnit rozhodování lidí, např. před volbami, v průběhu stávky, v období ekonomických nebo mocenských zvratů. Jako sociologové máme tedy vysokou odpovědnost při rozhodování, jakým způsobem, kdy a komu své výsledky sdělíme, jakou formou je zveřejníme, jaké závěry zdůrazníme.

Literatura:

Adamec, Č. :

Formulace otázek a konstrukce dotazníku.

In: Dotazník v sociologickém výzkumu.

Liblice, Československá sociologická společnost při ČSAV
1976, s. 3- 22 (Reprint Praha, FSV UK 1991)

Babbie, E.R. :

The Practice of Social Research.

Belmont- California USA, Wadsworth Publ.Comp.

(3 th ed.) 1983

Bailey, K.D. :

Methods of Social Research.

New York, Free Press (2nd ed.) 1978

Bogardus, E. S. :

Introduction to Social Research.

Los Angeles - , Suttonhouse Ltd. 1933

Bulmer, M. :

Sociological Research Methods: An Introduction.

London-, Macmillan (2nd ed.) 1984

Burgess, R.G. :

In the Field.

London, Unwin-Hyman 1984

Disman, M. :

Terénní sběr informací.

In: Vybrané techniky sociologického výzkumu.

Praha, Svoboda 1969, s.151-345

Hakim, C. :

Research Design.

London, Allen-Unwin 1987

Handbuch der empirischen Sozialforschung. (Hrsg.R.König)

Band I., Erster Teil

Zweiter Teil: Grundlegende Methoden und

Techniken der Empirischen Sozialforschung.

Hirner, A. :

Ako sociologicky analyzovat ' ?

Bratislava, Ústav školských informácií 1976

Kalous, J. :

Průvodce přípravou empirického výzkumu pro pedagogy a psychology.

(int. tisk)

Praha, Pedagogický ústav J.A.Komenského ČSAV 1983

Kane, Eileen: *Doing Your Own Research.*

London-New York, Marion Boyars 1985

Kerlinger, F.N. :

Základy výzkumu chování.

Praha, Academia 1972

Labovitz, S.- Hagedorn, R. :

Introduction to Social Research.

New York-, McGraw-Hill 2nd ed. 1976

Lamser, V. :

Základy sociologického výzkumu.

Praha, Svoboda 1966

Lin Nan:

Foundations of Social Research.

New York, McGraw-Hill 1976

Mc Neil, P. :

Research Methods.

London, Routledge, 2nd ed. 1990

Metody a techniky sociologického výzkumu I. Principy, struktura a strategie empirického výzkumu. (skripta FF UK)

Praha, SPN 1982

Moser, C.A. - Kalton, G. :

Survey Methods in Social Investigation.

London-, Heinemann (2nd ed.) 1971

Možný, I.: *Zásady zpracování dat pro sociology.* (skripta) Brno, UJEP
1980

Možný, I.- Rabušic, L. :

Úvod do metodologie vědy pro sociology. (skriptum)

Brno, UJEP - fak.filos. 1989

Noelleová, E. :

Výzkum veřejného mínění.

Praha, Svoboda 1969

Petrusek, M. :

Sociometrie.

Praha, Svoboda 1969

Příručka pro sociology. (kolektiv sovětských autorů)

Praha, Svoboda 1980

Řehák, J. :

K pojmu znak v sociologii.

= Sociologický časopis 1972, č. 6, s. 615-625

Řehák, J. - Řeháková, B. :

Analýza kategorizovaných dat v sociologii.

Praha, Academia 1986

Sudman, S. - Bradburn, N.M. :

Asking Questions. A Practical Guide to Questionnaire Design.

San Francisco - , Jossey-Bass 1982

Otázky:

Co je to hypotéza? Uveďte příklad!

Co je to posunutí, vychýlení (bias)?

Co je to výzkumný problém? Jak je formulován?

Co je to znak (proměnná) v sociologii?

Jaké druhy náhodného výběru rozeznáváme?

Co jsou to kvóty? Kde je užíváme?

Jaké jsou nevýhody případové studie?

Tři výhody kvótního výběru:

Tři nevýhody kvótního výběru:

Co je to panel?

Co je to výběrová chyba? Jak vzniká?

Co je to stratifikovaný náhodný výběr? Jak se liší od kvótního výběru?

Co je to skupinkový náhodný výběr? Jak se liší od prostého náhodného výběru?

K čemu slouží předvýzkum? Kdy jej provádíme?

Vyjmenujte libovolné 4 chyby experimentálního projektu!

Co je to longitudinální výzkum?

Co je to vyrovnávání (matching) a k čemu se užívá?

Co je to znáhodňování (randomization) a k čemu se užívá?

Co je to prostý ordinální znak? Uveďte příklad!

Co je to prostý nominální znak? Uveďte příklad!

Co je to intervalový znak? Uveďte příklad!

Vyjmenujte druhy znaků (proměnných)!

Sestavte hypotézu pro proměnné: sociální vrstva, pohlaví, příjem!

Sestavte hypotézu pro proměnné: sociální skupina, vzdělání, příjem!

Sestavte hypotézu pro proměnné: frustrace, agrese, inteligence!

Sestavte hypotézu pro proměnné: hodnoty, majetek, počet dětí!

Sestavte hypotézu pro proměnné: počet sourozenců, věk rodičů, počet dětí!

Sestavte hypotézu pro proměnné: počet sourozenců, vzdělání, počet dětí!

Sestavte hypotézu pro proměnné: náboženské vyznání, pomoc bližnímu, lidský soucit!

Sestavte hypotézu pro proměnné: nacionalismus, agrese, inteligence!

Sestavte hypotézu pro proměnné: majetek, příjem, melouchy!

Sestavte hypotézu pro proměnné: majetek, styky, výkon v zaměstnání!

Navrhněte hodnoty poměrového znaku pro vlastnost "návštěvnost kina"!

Navrhněte hodnoty intervalového znaku pro vlastnost "čtenářská aktivita"!

Navrhněte hodnoty kardinálního znaku pro vlastnost "příjmová kategorie"!

Navrhňte hodnoty intervalového kategorizovaného znaku pro vlastnost "věková kategorie"!
Navrhňte hodnoty nominálního znaku pro vlastnost "způsob získávání časopisu Mladý svět"!
Navrhňte hodnoty nominálního znaku pro vlastnost "rodinný stav"!
Navrhňte hodnoty nominálního znaku pro vlastnost "kategorie povolání"!
Navrhňte hodnoty nominálního znaku pro vlastnost "stranická příslušnost"!
Navrhňte hodnoty nominálního znaku pro vlastnost "oblíbený hrdina dětských pohádek"!
Navrhňte hodnoty ordinálního znaku pro vlastnost "spokojenost s mistrem"!
Navrhňte hodnoty ordinálního znaku pro vlastnost "připravenost k obětem"!
Navrhňte hodnoty ordinálního znaku pro vlastnost "nejvyšší stupeň dosaženého vzdělání"!
Navrhňte hodnoty ordinálního znaku pro vlastnost "důvěryhodnost informačních pramenů"!

Sestavte operacionální definici pojmu "inteligence"!
Sestavte operacionální definici pojmu "agrese"!
Sestavte operacionální definici pojmu "sociální vrstva"!
Sestavte operacionální definici pojmu "student"!
Sestavte operacionální definici pojmu "učitelka"!
*Sestavte operacionální definici pojmu "milionář"!
*Sestavte operacionální definici pojmu "vítěz parlamentních voleb"!
*Sestavte operacionální definici pojmu "malá sociální skupina"!
*Sestavte operacionální definici pojmu "chudoba"!
*Sestavte operacionální definici pojmu "bohatství"!
*Sestavte operacionální definici pojmu "nezaměstnaný"!
*Sestavte operacionální definici pojmu "agresor"!
*Sestavte operacionální definici pojmu "mluvčí OF"!

Vyjmenujte techniky sběru dat!

Vyjmenujte a charakterizujte druhy pozorování!

Vyjmenujte a charakterizujte základní techniky dotazování!

Co je to metoda obsahové analýzy? Zařaďte a charakterizujte!

Charakterizujte možnosti sekundární analýzy! Srovnajte její výhody a nevýhody oproti primární analýze dat!

Příprava standardního rozhovoru! Vyjmenujte etapy!

Vyjmenujte pravidla zúčastněného pozorování!

Jmenujte způsoby využití počítače při sběru dat!

Vyjmenujte fáze sociologického výzkumu a specifikujte místo terénního sběru dat!

**Vyjmenujte hlavní techniky sběru dat!

Uveďte čtyři součásti plánu standardizovaného pozorování!

*Vyjmenujte důležité fáze sběru dat!

&Charakterizujte v protikladu standardizované a nestandardizované pozorování!

Charakterizujte situaci skrytého a zjevného pozorování a roli výzkumníka v nich!

&Charakterizujte obor použití metody pozorování!

*Uveďte 3 výhody zúčastněného pozorování!

&Uveďte 3 nevýhody (problémy) zúčastněného pozorování!

*Co je obsahem záznamu standardizovaného pozorování?

*Vyjmenujte a charakterizujte metody dotazování!

**Popište úvodní fázi rozhovoru!

**Vyjmenujte vhodné vlastnosti tazatele (tazatelky)!

**Uveďte 4 typy nevhodných (chybně formulovaných) otázek! Vysvětlete!

*Vyjmenujte fáze sběru dat při dotazování poštou!

**Specifikujte obsah průvodního dopisu pro dotazování poštou!

**Charakterizujte oblast použití standardizovaného rozhovoru!

*Co je to pilotáž?

**Pravidla uspořádání dotazníku:

*Uveďte 3 výhody (přednosti) poštou zasílaného dotazníku!

Uveďte 3 nevýhody (omezení) poštou zasílaného dotazníku!

*Obsah průvodního dopisu k dotazníku zasílanému poštou:

*Jmenujte 3 faktory ovlivňující návratnost poštovního dotazníku!

*Vyjmenujte typy dokumentů vhodné k využití v sociologickém výzkumu!

**Co je to sekundární analýza?

**Oblasti a důvody použití studia dokumentů v sociologickém výzkumu:

**3 přednosti studia dokumentů ve srovnání s ostatními metodami sociologického výzkumu:

*3 nevýhody studia dokumentů ve srovnání s ostatními metodami sociologického výzkumu:

**Jmenujte způsoby využití počítače při sběru sociologických výzkumných dat!

**Jmenujte 2 výhody dotazování s využitím počítače!

**Jmenujte 2 nevýhody dotazování s využitím počítače!

*Jaké jsou úlohy přípravy a čištění dat?

*Vyjmenujte možnosti uspořádání (typy) dat vstupujících do analýzy!

**Co je to transformace dat? Uveďte příklad!

*Jmenujte míry (statistiky, ukazatele) pro jednorozměrný popis dat!

- **Jmenujte úlohy pro dvou a vícerozměrnou analýzu dat!
- *Jmenujte některé techniky mnohorozměrné analýzy dat!
- *Vyjmenujte závazné části (oddíly) výzkumné zprávy (publikace z výzkumu)!
- **Uvedte způsoby grafického porovnání hodnot kardinálního znaku za podsoubory!
- **K čemu se užívá histogram (sloupkový graf) a k čemu krabičkový graf (BOX CHART)?
- **K čemu se užívá dendrogram (stromový graf) a k čemu kruhový dělený graf (PIE CHART)?
- **K čemu se užívá znaménkové schéma a k čemu čárový graf (LINE CHART)?
- **Vyjmenujte 5 ze zásad etického kodexu sociologa!
 - **Jmenujte 4 z bodů povinného obsahu výzkumné zprávy!
 - *Jmenujte fáze rozhodování v sociologickém výzkumu!

Příklad pojmu: VÝKON ROLE

Vypovídá pouze o těch vzorcích chování, které vyžaduje daná role, ne jiná

- ROLE STUDENTA:
- sleduje přednášky
 - dělá si poznámky
 - účastní se práce v seminářích
 - studuje odbornou literaturu
 - připravuje se na zkoušky
 - skládá zkoušky

Pouze tyto aktivity jsou obsaženy v roli studenta. Další činnosti, které student dělá (např. činnosti v rodině, s rodiči a sourozenci, aktivity volného času, partnerské vztahy, práce o prázdninách), nejsou do této role zahrnuty. Výkon role je abstrakce ze všech činností studenta jen těch, které jsou určeny touto dílčí rolí.

Osnova

1. Vědecký přístup, problémy a hypotézy

- Věda a vědecký přístup
- Výběr a formulace vědeckého problému
- Hypotézy a problém operacionalizace

2. Vědecké pojmy, znaky a proměnné

- Vědecké pojmy a konstrukty
- Znak v sociologii, typy znaků
- Znak, proměnná, indikátor
- Výstavba a testování sociologické teorie

3. Populace, výběry, reprezentativita

- Populace, opora výběru, základní a výběrový soubor
- Náhodný výběr a jeho výhody
- Typy a způsoby výběru, vícestupňový náhodný výběr
- Reprezentativita, chyby, vychýlení výběru
- Velikost výběrového souboru
- Kvótní výběr a jeho omezení

4. Výzkumné projekty

- Cíle výzkumného projektu a kritéria jeho hodnocení
- Výzkumný projekt a kontrola vlivu proměnných
- Varianty výzkumného projektu
- Případová studie, korelační výzkum, panelový výzkum, experimentální projekt

5. Terénní sběr dat. Metody pozorování

- Sběr dat v procesu sociologického výzkumu
- Důležité fáze sběru dat
- Hlavní techniky sběru dat
- Charakteristika pozorování a jeho druhy
 - nezúčastněné a zúčastněné
 - standardizované a nestandardizované
 - skryté a zjevné
- Obor použití pozorování
- Příprava a realizace standardizovaného pozorování
- Výhody a nevýhody zúčastněného pozorování

6. Metody dotazování

- Charakteristika rozhovoru
- Druhy rozhovorů - standardizovaný a nestandardizovaný
- Příprava standardizovaného rozhovoru
- Průběh rozhovoru

- Charakteristika dotazníkového šetření
- Typy dotazníkového šetření a jejich použití
- Příprava dotazníkového šetření

7. Ostatní metody sběru dat, obsahová analýza, sociometrické metody

- Studium dokumentů
- Zásady obsahové analýzy
- Sociometrické techniky
- Škálování
- Tvorba a používání položkových testů
- Vybrané metody kvalitativní analýzy

8. Dotazování s využitím počítače

- Klasické a počítačové dotazování - rozdíly, výhody, nevýhody
- CADI - záznam dat pomocí počítače
- CAPI - osobní dotazování s využitím počítače
- CATI - telefonní dotazování s využitím počítače
- program BLAISE - základní charakteristika

9. Principy analýzy a interpretace dat

- Typy vstupních dat
- Jednorozměrný popis dat
- Komparace souborů
- Korelační analýza
- Další mnohorozměrné techniky analýzy dat
- Kauzální modelování

10. Prezentace výsledků sociologického výzkumu

- Ústní prezentace a argumentace
- Písemná prezentace - projekt a závěrečná zpráva
- Grafická prezentace výsledků
- Důkazy, doklady, ilustrace
- Interpretace a závěry výzkumu

13. Etické a profesionální zásady sociologického výzkumu

- Dosavadní poznání jako východisko výzkumu
- Kritická analýza a prověřování dosažených výsledků
- Metodologická skepse a tvořivá invence
- Zásady citací a uvádění pramenů
- Etická tabu výzkumníka
- Zjistit, prověřit, publikovat

Bales, Robert F.:

Interaction Process Analysis.
Reading, Mass: Addison-Wesley Publ., 1950

Poznámky

- ¹⁾ Buchler, J., ed.: Philosophical Writings of Peirce. Dover 1955, p.18
- ²⁾ podle Dewey, J.: How we Think. Boston, Heath 1933, p.106nn
- ³⁾ Hezkou ukázkou postupného rozvíjení teoretických úvah na toto téma - vhléd do dílny metodologa a zároveň teoretika nalezneme v knize: Stinchcombe, A.L.: Constructing Social Theories. New York -, Harcourt 1968. Srovnej viz. [Močn'-Rabužic 1989: 41-51, kapitola 6.
- ⁴⁾ Viz. Merton, R.K.: Social Theory and Social Structure. Glencoe, Ill., The Free Press 1957
- ⁵⁾ Stouffer, S.A.: The American Soldier. Vol.1. Princeton, N.J., Princeton Univ.Press 1949
- ⁶⁾ Einstein, A.: Johann Kepler. Frankfurter Zeitung, 9.11.1930
- ⁷⁾ Cohen, M.: A Preface to Logic. New York, Meridian 1956, p.148
- ⁸⁾ Černík, V.-Farkašová, E.-Viceník, J.: Teória poznania. 1.vyd. Bratislava, Pravda 1980, str.170-235
- ⁹⁾ Bales, Robert F.: Interaction Process Analysis. Reading, Mass: Addison-Wesley Publ., 1950
- ¹⁰⁾ Thomas, W.I. - Znaniecki, F.: The Polish Peasant in Europe and America, Vol.I. and II.. Boston, Badger 1918
- ¹¹⁾ Berelson, B.: Content Analysis. In: Handbook of Social Psychology (edited by Gardner Lindzey) Cambridge, Ma: Addison-Wesley 1954 Berelson poučil Cartwrightova pojmu "symbolické chování" z jeho práce Cartwright, D.P.: Analysis of Qualitative Material" In: Research Methods in the Behavioral Sciences (Festinger-Katz eds.) New York, Holt, Rinehart & Winston 1953
- ¹²⁾ Berelson, B.- Salter, Patricia: Majority and Minority Americans: An Analysis of Magazine Fiction. Public Opinion Quarterly 10, 1946:168-90
- ¹³⁾ Program BLAISE byl vyvinut v Netherlands Central Bureau of Statistics NCBS v osmdesátých letech a je užíván v Evropě i v USA pro značnou variabilitu poskytovaných možností a pro snadnou ovladatelnost.
- ¹⁴⁾ Program SPSS/PC+ 4.0 je systém pro statistickou analýzu dat užíván jako jeden z nejrozšířenějších po celém světě pro statistickou analýzu, ale i přípravu a transformace dat ze sociálních výzkumů a dalších primárních pramenů na osobních počítačích. Existuje jeho varianta SPSSX používaná na sálových počítačích. SPSS/PC+ 4.0 Base Manual for the IBM PC/XT/AT and PS/2, Marija Noružis/SPSS Inc., Chicago, SPSS Inc. 1991. V rámci systému programů SPSS je používán program DATA ENTRY sloužící pro vstup dat a jejich vybavení informacemi potřebnými pro jejich obsluhu a zpracování. Používá se také pro čištění a opravy dat.
- ¹⁵⁾ Zmiňuje se o něm studie L.Horowitze (ed.) The Rise and Fall of Project Camelot. Cambridge, Mass., MIT Press 1967
- ¹⁶⁾ [Vidich, A.J.- Bensman, J.: Small Town in Mass Society. New York, Princeton University Press 1958, (2 nd edn NY, Princ. 1968)]

