

Co všechno znamená NESSTAR?*

Tomáš Čížek**

Sociologický ústav AV ČR, v.v.i.

Martin Vávra**

Sociologický ústav AV ČR, v.v.i.

Abstrakt: *Tento text představuje komplexní úvod k systému NESSTAR. Je určen především (potenciálním) uživatelům tohoto programového balíku, určeného pro publikování, prohledávání a analýzu dat na webu. V úvodní části je popsán vznik tohoto programu, jeho jednotlivé části a také možnosti, které nabízí producentům a depozitorům dat. V druhé části je pak NESSTAR představen jako nástroj pro on-line prohledávání a analýzu dat. Jde současně o uživatelský manuál, který by měl ukázat cestu k sociologickým datům na internetu.*

Data a výzkum - SDA Info 2007, Vol. 1, No. 1: 19-33.

(c) Sociologický ústav AV ČR, v.v.i., Praha 2007.

NESSTAR jako nástroj pro datové archivy

Moderní trend umožněný rozvojem informačních technologií směřuje ke stále snadnější dostupnosti dat. Na této cestě je sice potřebné překonat řadu překážek (jakou je například nutnost uvést data do plného souladu se zákony na ochranu osobních údajů), ale lze doufat, že po ní sociální vědy půjdou stále rychleji.

Jedním ze základních prvků potřebných pro publikaci, sdílení a analýzu dat v prostředí internetu je samozřejmě kvalitní software. I pokud už máme v pořádku rámec pro sdílení dat (který v tomto případě představuje především DDI), stále ještě potřebujeme programy, které jsou schopny tento rámec využít a aplikovat v konkrétní praxi. V roce 1998 (tedy ještě před ustavením iniciativy DDI) byl za

* Tento článek byl připraven s podporou Ministerstva školství, mládeže a tělovýchovy v rámci programu Informační infrastruktura výzkumu v projektu reg. č. 1N04192.

** Veškerou korespondenci posílejte na adresu: Tomáš Čížek, Sociologický ústav AV ČR, v.v.i., Jilská 1, 110 00 Praha 1, e-mail: tomas.cizek@soc.cas.cz; Martin Vávra, Sociologický ústav AV ČR, v.v.i., Jilská 1, 110 00 Praha 1, e-mail: martin.vavra@soc.cas.cz

finanční podpory EU a pod vedením Norwegian Social Science Data Services (NSD) odstartován projekt NESSTAR¹ (Networked Social Science Tools and Resources), který měl vést právě k vývoji praktických aplikací ke sdílení dat. Projekt NESSTAR byl následován projektem FASTER, který dále rozvíjel koncept publikování dat na internetu a jejich on-line analýzy. Oba projekty byly úspěšné, o čemž svědčí i to, že jsou nyní rozvíjeny na komerčním základě a NESSTAR Ltd. vytváří aplikace a webová řešení pro řadu firem.²

Co je dnes, po téměř 10 letech vývoje, NESSTAR? „Předně je to softwarový balík složený z několika modulů, které umožňují provoz internetového serveru poskytujícího data, publikování dat na tomto serveru a prohlížení, kopírování a analýzu dat z jiných počítačů. Dále je NESSTAR komerční společnost, která tento softwarový balík vyvíjí a distribuuje. Za třetí je NESSTAR také systém distribuce datových souborů. A konečně, NESSTAR představuje síť datových archivů, které tento systém implementovaly, a klientů, kteří služeb archivů pomocí systému a softwaru NESSTAR využívají“ [Kalvas 2006]. To je velice široké vymezení. Zde se budeme zabývat pouze programy distribuovanými pod tímto jménem.

Co přesně NESSTAR, nyní chápaný jako balík programů, přináší? „Cílem je nabídnout co největší množství dat on-line v integraci s odpovídajícími metadaty, a to ve spojení se stabilními a funkčními odkazy na relevantní kontextuální informace (odkazy na vědecké publikace, informace o výzkumnících a institucích)“ [Lach 2002].

V zásadě jde nyní o tři základní součásti. První je **NESSTAR server**, který slouží distributorům dat. Jde o rozšíření běžného webového serveru, který obhospodařuje databázi dat a přístup k ní. V původní verzi mohl pracovat NESSTAR pouze v prostředí Windows, což mohlo činit značné komplikace, pokud v dané organizaci fungovaly všechny servery v jiném systému. Z praktických zkušeností při zavádění této verze v SDA je možno zmínit i problém s českými znaky při používání funkce vyhledávání. Nyní existují i verze pro Linux, které již tyto problémy řeší.

Na serveru je možné i definovat a spravovat přístupová práva k datům, takže lze například stanovit k některým souborům zcela volný přístup, včetně jejich stahování, nebo naopak vázat všechny operace na poskytnutí souhlasu s podmínkami užívání. Většina činností spojených s instalováním a údržbou serveru není obtížná a nevyžaduje zvláštní znalosti nad rámec obvyklý u jiných serverů. Se zakoupením licence navíc její držitel získává i právo na technickou podporu od NESSTAR Ltd.

Druhou částí je **NESSTAR Publisher**, nástroj pro přípravu dat a metadat k publikaci. Jde o uživatelsky jednoduchou aplikaci, kde je možné přímo do definovaných políček zapisovat jednotlivé segmenty metadat, a to bez znalosti XML, který je interním formátem DDI. Pracovníci institucí, kde se připravují datové soubory k publikaci, případně pouze k archivování, tedy nemusí ovládat práci s XML asemblery. Data mohou být do *Publisheru* importována v různých formátech (a také samozřejmě mohou být v kterémkoli z těchto formátů exportována, takže je možno jej využívat jako konvertor formátů). U rozsáhlejších souborů je možno definovat a popsat skupiny proměnných tak, aby byl publikovaný soubor pro uživatele přehlednější.

V *Publisheru* je možné i kontrolovat a případně opravovat data (pro to je ovšem výhodnější používat nějaký statistický paket jako například SPSS) a nakonec samozřejmě data i publikovat na NESSTAR serveru.

Konečně třetí nástroj je **NESSTAR Webview**, který slouží uživatelům. Je integrován přímo na serveru a uživatelům stačí mít nainstalován běžný webový prohlížeč. Jde o pokrok narozdíl od předchozích verzí, kdy uživatelé, kteří chtěli využívat všechny funkce systému, museli na svůj počítač instalovat speciální program.

Nelze samozřejmě říci, že systém NESSTAR je bez vad. Zatím to není žádná „sociologická mašina snů“, jak zněl titul jednoho z článků o NESSTARu. Je ovšem dostatečně jednoduchý, aby mohl být používán lidmi, kteří nemají žádné hlubší znalosti IT technologií, a současně i dostatečně sofistikovaný, aby umožňoval publikaci dat a metadat na odpovídající úrovni. Navíc je potřeba porovnávat současnost se situací před jeho implementací, kdy většina archivů žádné on-line řešení pro datové služby neměla nebo pouze velmi jednoduché.

To, že sdílení dat má stále ještě daleko k dokonalosti, není zdaleka jen technický problém, ale často i problém nedostatečné spolupráce, jazykových bariér a rozštěpení evropského výzkumného prostoru do národních komunit. Například to, že NESSTAR funguje do značné míry stále spíše jako řešení pro jednotlivé archivy a méně slouží jejich integraci, lze připsat právě těmto výše zmíněným příčinám. V současnosti už ale funguje CESSDA portál, který představuje rozhraní pro vyhledávání datových zdrojů v jednotlivých národních archivech. Ten, kdo má o data zájem, tedy již nemusí postupně navštěvovat stránky jednotlivých národních archivů a vyhledávat na nich, ale zadá vyhledávaný výraz na jedné stránce a z výsledků vyhledávání se již dostane snadno ke konkrétním výzkumům na stránkách národních archivů (více o tomto článku J. Krejčího v tomto čísle [Krejčí 2007]).

Podmínkou integrace do CESSDA portálu je právě používání systému NESSTAR. Základem úspěchu jakékoli webové technologie je její masivní rozšíření mezi uživateli. Lze říci, že v rámci komunity evropských datových archivů se toto do velké míry již stalo a nadále děje. Nedostatečné je stále ještě rozšíření mimo sociologické datové archivy a zvláště mezi soukromými producenty a distributory sociálněekonomických dat.³

NESSTAR jako nástroj pro prezentaci a analýzu dat

Další část článku představuje základní funkce webového prezentačního rozhraní NESSTAR. Čtenářům přibližuje vzhled a používání jednotlivých funkcí programu.

Vstup do systému (viz obrázek 1) je možný z webové stránky našeho archivu (SDA: <http://archiv.soc.cas.cz>). Hlavní obrazovka programu NESSTAR ukazuje v levém horním rohu datové katalogy uložené v systému. V případě Sociologického datového archivu se jedná o katalog SDA, kde jsou uloženy kvantitativní výzkumy, a katalog Medard, kde se nacházejí popisy kvalitativních výzkumů. Některé výzkumy jsou v systému uloženy i v anglickém jazyce, většina jich je ale pouze v jazyce českém.

NESSTAR umožňuje celou škálu funkcí, které budou v textu dále popsány:

- prohlížení datových souborů a tabulek,
- prohlížení datové dokumentace,
- vytváření kontingenčních tabulek,
- korelační a regresní analýzu,
- grafickou reprezentaci dat a analýz,
- vytváření nových proměnných,
- rekódování stávajících proměnných.

Obrázek 1. Vstup do systému

The image shows a screenshot of the Sociologický datový archiv (SDA) website. The browser address bar shows 'http://archiv.soc.cas.cz/'. The page header includes 'SDA Sociologický datový archiv' and 'Sociologický ústav AV ČR'. A search bar is visible on the right. The left sidebar contains a menu with options: O ARCHIVU, DATA, PROJEKTY, PUBLIKACE, ODKAZY, DEPOZICE, PRÁCE V SDA, MEDARD, SOÚ AV ČR, and ENGLISH. The main content area features a banner with stylized figures and a 'více informací' link. Below the banner, there are three highlighted links: 'Volný vstup do systému NESSTAR', 'Objednání dat - generování objednávkového formuláře', and 'Registrace uživatele do systému NESSTAR'. To the right of these links, a text block explains that registered users have full access to selected Czech data sets from the ISSP program and can access other data sets based on their consent to terms of use. Below this, there is a smaller screenshot of the 'nesstar.soc.cas.cz:8081/webview/' page, which shows a world map and a list of research categories: SDA, Česky, ISSP, IWVM, CVM, EOP - Ekonomická očekávání a postoje, Výzkumy transformace, Výzkumy mládeže, Genderové výzkumy, and Výzkumy bydlení.

Zdroj: Sociologický datový archiv SOÚ AV ČR.

+ Kliknutím na čtvereček vedle názvu katalogu se uživatel dostane k volbě mezi českým a anglickým katalogem a následně do seznamu samotných výzkumů, které jsou uloženy v tematických adresářích (např. ISSP, genderové výzkumy atd.).

Informace o jednotlivých výzkumech jsou v systému NESSTAR ukládány podle mezinárodního standardu DDI.⁴ Základní informace o výzkumech mohou získat i uživatelé, kteří se nechtějí registrovat. Pro plné využití všech funkcí systému je však registrace potřebná.

Registrace (viz obrázek 2) se provádí z webové stránky archivu a uživatel je nucen se zavázat k respektování pravidel zacházení s datovými soubory. Po jejím provedení mohou uživatelé provádět analýzy přímo prostřednictvím NESSTARu, nebo je jim umožněno objednat datové soubory ke stažení. Každému uživateli je přiřazeno jméno a heslo pro přístup ke všem funkcím systému. Přihlašovací informace jsou po skončení registrace posílány automaticky na uvedenou mailovou adresu klienta.

Obrázek 2. Registrace

Adresa <http://archiv.soc.cas.cz/>

SDA Sociologický datový archiv Sociologický ústav AV ČR

Vyhledávání

O ARCHIVU
DATA
PROJEKTY
PUBLIKACE
ODKAZY
DEPOZICE
PRÁCE V SDA

MEDARD
SOÚ AV ČR
ENGLISH

víc informací

Volný vstup do systému NESSTAR

Objednání dat - generování objednávkového formuláře

Registrace uživatele do systému NESSTAR

Registrovaní uživatelé mají přístup k vybraným českým datovým souborům z programu ISSP a mohou získat přístup k dalším souborům na základě souhlasu s podmínkami jejich používání. Bez registrace je možné procházet stránky Webu a informace o archivovaných datech.

SDA Sociological Data Archive
unlocking | data | creating | knowledge

Podmínky:

Pro užívání datových souborů ze Sociologického datového archivu se zavazují dodržet následující obecné podmínky:

- 1) Vydávaná data a materiály použijí pouze pro účely mezinárodního výzkumu nebo výuky. Výzkumné údaje jsou přísně důvěrné. Tyto údaje nebudou nikdy ani se je nebudou poskytnout zjevně a v výzkumných datech a materiálech bude zřejmá důvěrnost těchto údajů ochráněna.
- 2) Každý uživatel musí před použitím datových souborů uzavřít s archivem smlouvu o poskytnutí datových souborů. Smlouva musí být uzavřena před použitím datových souborů.
- 3) Archivní podmínky (období, délka, počet, atd.) jsou v smlouvě uvedeny.
- 4) Během na vědomí, že archiv neodpovídá za spolehlivost dat.
- 5) Archivní podmínky (období, délka, počet, atd.) jsou v smlouvě uvedeny.

Uživatelé, kteří souhlasí s podmínkami, mohou pokračovat v registraci.

Souhlasím s podmínkami

Registrace

Registrační email*:

Jméno*:

Příjmení*:

Instituce*:

útvár/oddělení/fakulta:

Země*:

typ aktivity, který nejvíce odpovídá Vaší situaci*:

akademický výzkum

Zdroj: Sociologický datový archiv SOÚ AV ČR.

Položky hlavního menu každého datového souboru jsou členěny do dvou základních kategorií:

- 1) **Metadata (dokumentace dat)** – informace o výzkumu (viz obrázek 2), datovém souboru, použité metodologii a podmínkách poskytnutí datového souboru. Je zde možné stáhnout dotazník a případné další materiály o výzkumu a datovém souboru.
- 2) **Variable description** (popis proměnných, viz obrázek 3) – zde jsou podle tematických kategorií seřazeny jednotlivé proměnné použité ve výzkumu a jsou uvedeny frekvence jejich kategorií v datovém souboru. V případě, že proměnná představuje odpověď respondenta v dotazníkovém šetření, je zde uvedeno i přesné znění otázky.

Obrázek 3. Study description - popis výzkumu

DESCRIPTION | TABULATION | ANALYSIS

2007

Place of Production
Praha CZ.

Data Distributor

- Sociologický datový archiv Abbreviation: SDA
Affiliation: Sociologický ústav, Akademie věd ČR

Depositor

- Sociologický ústav, Akademie věd ČR Abbreviation: SOU
Affiliation: Czech Academy of Sciences

Series Name
Kontinuální měsíční výzkum CVVM

Bibliographic Citation
Sociologický ústav, Akademie věd ČR (SOU) 2007. Naše společnost 2007. [počítačový soubor] [productent]. Sociologický datový archiv - Sociologický ústav, Akademie věd ČR [archiv].

Topic Classification

- Stranické preference a volební model v březnu 2007
- Důvěra k ústavním institucím a spokojenost s politickou situací
- Důvěra k stranickým představitelům
- Postoje k aférám v politickém a veřejném životě
- Diskuse občanů o politických tématech, představy o vlivech na rozhodování politiků
- Kdy by měl politik odejít ze své funkce?
- Jak jsou na tom Češi s tolerancí?
- Politická kultura v ČR
- Důvěra k některým institucím veřejného života
- Co je pro nás morálně přijatelné a co nikoli?
- Hodnocení informací v médiích

Abstract
Kontinuální měsíční výzkum provedl výzkumný tým CVVM Sociologického ústavu AV ČR při něm zjišťovány postoje a názory občanů.

Date of Collection

Zdroj: Sociologický datový archiv SOU AV ČR.

Obrázek 4. Variable description - popis proměnné

DESCRIPTION | TABULATION | ANALYSIS

Dataset: Naše společnost 2007 - březen

Variable PV.1 : PV.1 Volební účast do Poslanecké sněmovny

Literal Question
Představte si, že by příští týden byly volby do Poslanecké sněmovny. Šel byste volit?

Values	Categories	N	
1	rozhodně ano	280	26.8%
2	spíše ano	307	29.4%
3	spíše ne	135	12.9%
4	rozhodně ne	180	17.2%
8	nená volební právo	58	5.6%
9	neví	85	8.1%
0	BEZ ODPOVĚDI	1	

Summary Statistics

Valid cases 1045
Missing cases 1
This variable is numeric

Zdroj: Sociologický datový archiv SOU AV ČR.

Třetí kategorii představují záložky (*bookmarks*), zde si registrovaní uživatelé mohou ukládat vlastní odkazy na jednotlivé výzkumy nebo na jiné informace a analýzy v systému NESSTAR.

Obrázek 5.
Description, tabulation, analysis



Uživatelům systému jsou přístupné tři hlavní možnosti práce s uloženými daty: DESCRIPTION, TABULATION a ANALYSIS (viz obrázek 5).

Description (popis) – zobrazí veškeré dostupné informace o uložených datech (metadata);

Tabulation (tabulky) – zobrazí kontingenční tabulky (pro využití této funkce je nutná registrace);

Analysis (analýza) – umožňuje provést korelační nebo regresní analýzu (pro využití této funkce je také nutná registrace).

Popis dat (DESCRIPTION) zobrazuje informace o datovém souboru a četnosti proměnných. Kromě četností proměnných jsou uvedeny i základní statistické charakteristiky, jako je počet případů, chybějící odpovědi, maximum, minimum a případně průměr.

Položka TABULATION umožňuje vytvářet jednouchou tabulku četností jednotlivé proměnné nebo kontingenční tabulku pro více proměnných (viz obrázek 6).

Obrázek 6. Tabulation - četnostní tabulky

The screenshot shows the 'TABULATION' interface. On the left, a list of variables is shown with interactive buttons: 'Add to row', 'Add to column', 'Add to layers', and 'Add as measure'. The selected variable is 'IDE.6 Vzdělání'. The central area contains instructions: 'Choose 'Add to column' to place the variable here.' and 'To populate this table you need to select a variable from the browse list, click on it and then add it to row, column or layers, or use it as a measure variable.' Below this, there are dropdown menus for 'Type' and 'Measure'. The resulting contingency table is shown at the bottom.

	Code	Frequency	% of all	% of valid
OV.1 Spokojenost s životem				
OV.1 Spokojenost s životem				
velmi spokojen	1	24	21.2	21.2
spíše spokojen	2	70	61.9	61.9
ani spokojen, ani nespokojen	3	15	13.3	13.3
spíše nespokojen	4	3	2.7	2.7
velmi nespokojen	5	1	0.9	0.9
neví	9	0	0.0	0.0
Total		113	100.0	100.0

Zdroj: Sociologický datový archiv SOÚ AV ČR.

Tabulku vytvoříme kliknutím na TABULATION v hlavní liště programu. V levém sloupci si vybereme proměnnou. Program zde nabízí čtyři následující možnosti:

Add to row – přiřadí proměnnou do řádku tabulky.

Add to column – přiřadí proměnnou do sloupce tabulky.

Add to layers – tabulka je vytvářena pro každou kategorii dané proměnné zvlášť, tato proměnná je pak graficky zvýrazněna.

Add as measure – pokud je spojitá (!) proměnná přidána do prázdné tabulky, program vypočítá statistické charakteristiky této proměnné, jedná se o následující údaje: medián, průměr, minimum, maximum, směrodatná odchylka, suma, počet případů, různé intervaly spolehlivosti a kvantily. Proměnná může být také přidána do tabulky jako míra (*measure*) ostatních proměnných. Například v případě spojitě proměnné věk přidán do tabulky obsahující proměnné stav a pohlaví se zobrazují průměrný věk nebo ostatní charakte-

Obrázek 7. Add as a measure - výpočet statistických charakteristik kategorií

The screenshot illustrates the 'Add as a measure' process in a statistical software interface. It shows a variable list on the left, a context menu for 'IDE.2 Věk', a summary table of statistics for 'IDE.2 Věk', and a pivot table showing the average age by gender and marital status.

IDE.2 Věk Summary Table:

Median	43.84
Mean	44.26
Min	15.00
Max	83.00
S.d.	16.89
Sum	46,247.00
N	1,045
CI 95% min	43.23
CI 95% max	45.28
CI 99% min	42.91
CI 99% max	45.60
Box whisker low	15.00
First quartile	29.22
Third quartile	58.77
Box whisker high	83.00

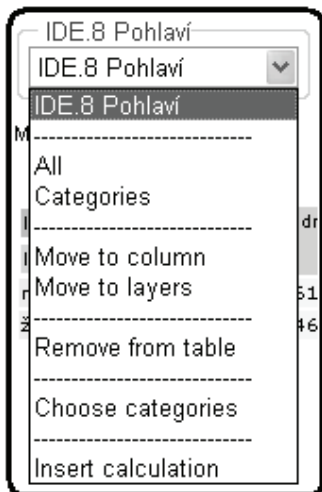
Pivot Table:

IDE.8 Pohlaví	svobodný, svobodná	ženatý, vdaná	rozvedený, rozvedená	vdovec, vdova	neví
muž	24.40	49.33	48.94	65.96	28.00
žena	23.27	46.95	49.96	64.32	-

Zdroj: Sociologický datový archiv SOÚ AV ČR.

ristiky (minimum, maximum, směrodatná odchylka a suma) pro každou kombinaci výše uvedených proměnných. Jako „measure“ lze použít pouze jednu proměnnou.

Obrázek 8.



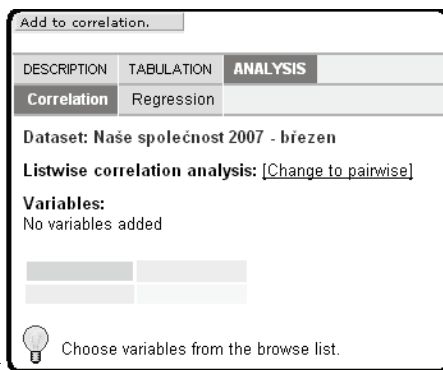
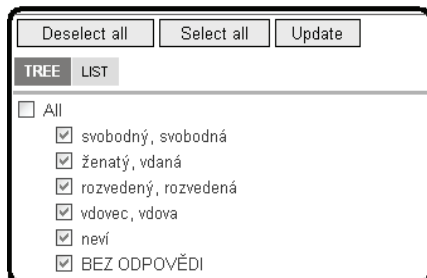
Názvy proměnných se objevují vždy nad tabulkou, případně grafem. Po kliknutí na tento název se rozbalí nabídka obsahující možnosti práce s proměnnou (viz obrázek 8). Ta, podobně jako předchozí volby, umožňuje v tabulce zobrazit údaje za všechny kategorie proměnné dohromady (*All*), za jednotlivé kategorie (*Categories*), umožňuje také přesouvat proměnné z řádku do sloupce a naopak (*Move to column*, *Move to row*). Přidávat proměnné do tabulky lze také přesunem proměnné pomocí stisknutého pravého tlačítka myši z oblasti nad tabulkou do oranžového (sloupec) nebo modrého (řádek) obdélníku, opačným způsobem lze proměnné z tabulky odebírat.

Dále nabídka umožňuje odstranění proměnné z tabulky (*Remove from table*) a zobrazení pouze vybraných kategorií (*Choose categories*). V tomto případě se objeví samostatné dialogové okno, kde lze vybrat jednotlivé kategorie pomocí zaškrtnutí políčka myši, jsou zde možnosti vybrání všech (*Select all*), nebo žádné kategorie (*Deselect all*). Změny se v tabulce objeví po kliknutí na políčko „*Update*“. Pokud je proměnná složena z více úrovní kategorií, lze pomocí „*TREE*“ zobrazit hierarchický strom nebo zvolit abecední seznam kategorií pomocí „*LIST*“ (viz obrázek 9).

Kromě generování tabulek umožňuje systém i pokročilejší statistické operace v oddíle ANALYSIS (viz obrázek 10). Jedná se o korelační a regresní analýzu. Volba „*Correlation*“ vyvolá prázdnou korelační tabulku. Proměnné lze do analýzy přidávat ze seznamu v levém sloupci pomocí volby „*Add to correlation*“.

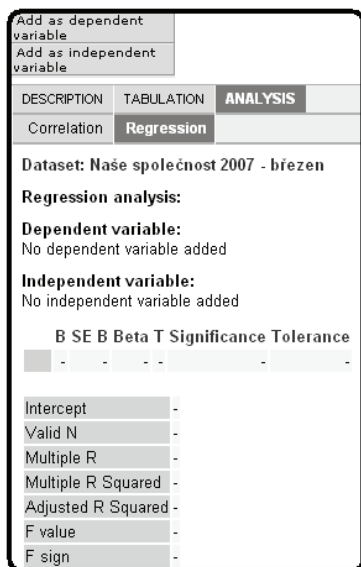
Regresní analýza se provádí obdobně zvolením „*Regression*“ (viz obrázek 11). V levém sloupci zobrazujícím proměnné je pak možno vybrat nezávislou proměnnou (*Add as independent variable*) a proměnné závislé (*Add as dependent variable*).

Obrázek 9.



Obrázek 10.


Obrázek 11.





Kontingenční tabulky a výsledky analýz lze v programu NESSTAR znázornit i vizuálně pomocí více typů grafů. Typ grafu lze zvolit pomocí příslušné ikony v pravém horním rohu obrazovky.


Hlavní menu programu


NESSTAR nabízí tyto možnosti grafického znázornění:


 **Bar chart (sloupcový graf)**, tento typ grafu je dostupný, pokud je v tabulce jedna nebo více proměnných (a nejedná se o jedinou míru);


 **Stacked bar chart (součtový sloupcový graf)**, dostupný, pokud jsou v tabulce dvě a více proměnných (a žádná míra);

 **Pie chart (výšečový graf)**, dostupný, pokud jsou v tabulce dvě a více proměnných (a žádná míra);


 **Box whisker (krabicový graf)**, dostupný, pokud jsou v tabulce dvě a více proměnných, z nichž jedna musí být míra. Pokud je v tabulce přítomna pouze míra, zobrazí její frekvence s normálním rozdělením a kvartily;

 **Confidence interval graph (graf intervalů spolehlivosti)**, dostupný, pokud jsou v tabulce dvě a více proměnných, z nichž jedna musí být míra;

 **Bar chart – mean (sloupcový graf s průměrem)**, dostupný, pokud jsou v tabulce dvě a více proměnných, z nichž jedna musí být míra;

 **Regression chart (regresní graf)**, dostupný pouze po provedení regresní analýzy, nahrazuje ikonu sloupcového grafu;

 **Time series graph (graf časových řad)**, dostupný, pouze pokud je v tabulce přítomna časová proměnná/dimenze;

 **Scatterplot**, dostupný pouze po provedení regresní analýzy, nahrazuje ikonu grafu časových řad.

Funkce NESSTARu jsou většinou dostupné z hlavního menu programu (viz obrázek 12) nacházejícího se v pravém horním rohu okna programu. Je se zde celkem čtrnáct ikon s různými funkcemi, jejichž popis následuje za obrázkem 12.

Přičítání (viz obrázek 14): Funkce *add* umožňuje přičítání konstanty nebo



Obrázek 12.
Hlavní menu
programu



Table view

Zobrazí tabulku (viz popis v předchozí části textu).



Bar chart

Zobrazí graf. Typ grafu závisí na volbě uživatele nebo možnosti použití v rámci prováděné analýzy (podrobněji viz výše).



Time series graph

Popis funkce výše



Scatterplot

Popis funkce výše



Map

Zobrazí mapu. Tato funkce není v současné době v našem archivu dostupná. Pokud bude zpřístupněna, umožní uživateli zobrazit mapové podklady a výstupy k uloženému výzkumu.



Clear

Tato funkce „vyčistí“ okno s tabulkami nebo analýzami.



Weight (survey data only)

Uložené datové soubory mohou obsahovat váhy, tato funkce umožňuje jejich použití. Je také možno použít jako váhy jiné proměnné ze souboru.



Subset (survey data only)

Tato možnost dovoluje uživateli provádět analýzy nebo zobrazení grafů na podsouboru kategorií z proměnných (například pouze za osoby mužského nebo ženského pohlaví).

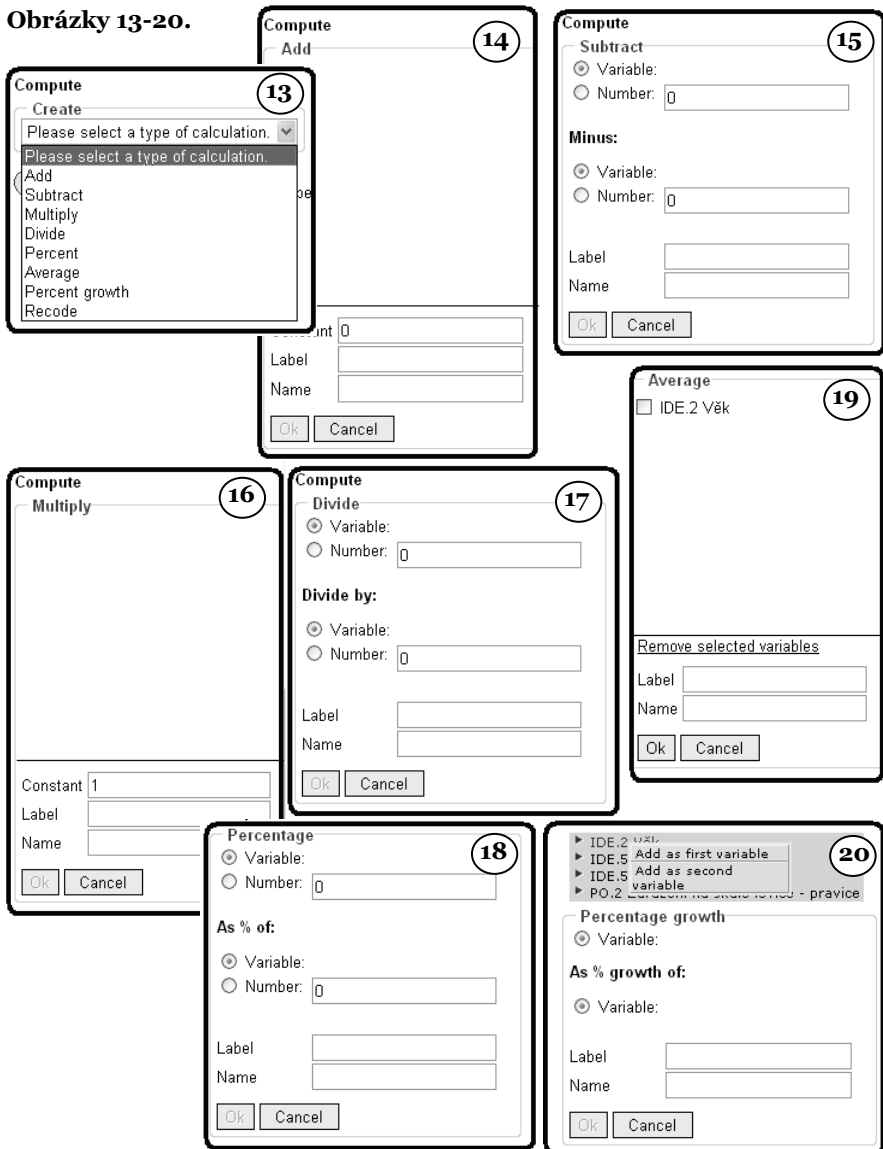
Vytváření a rekódování proměnných



Compute (survey data only)

Systém NESSTAR ve své nejnovější verzi umožňuje i vytváření nových uživatelsky definovaných proměnných. To je možné pomocí v programu předem připravených matematických a statistických operací. Po zvolení ikony „Compute“ se rozbálí nabídka jednotlivých operací. Jméno výsledné proměnné musí být unikátní. Pokud bude zadán název již existující proměnné, program zobrazí chybové hlášení.

Obrázky 13-20.



hodnoty jiné proměnné (případně více proměnných) ke zvolené proměnné. Tato operace (stejně jako všechny níže popsané) se po zadání provede kliknutím na políčko OK, zruší se zvolením Cancel.

Odečítání (viz obrázek 15): Funkce *subtract* umožňuje odečítat hodnoty jedné proměnné od hodnot proměnné druhé. Umožňuje také odečíst hodnoty zvolené proměnné od konstanty, nebo naopak odečítání konstanty od proměnné.

Násobení (viz obrázek 16): Funkce *multiply* umožňuje násobení proměnných

Obrázek 21. Rekódování proměnné

Compute

Recode

Old value

Value []

Range: [] through []

Range: Lowest through []

Range: [] through highest

New value

Value []

Label []

As missing

Add Delete

All other values

Keep as original values

Recode to Sysmis

Recode to:

Value []

Label []

As missing

New variable

Name []

Label []

OK Cancel

navzájem, nebo násobení zvolenou konstantou.

Dělení (viz obrázek 17): Funkce *divide* umožňuje dělení proměnné jinou proměnnou, dělení konstanty proměnnou nebo dělení proměnné konstantou.

Procenta (viz obrázek 18): Funkce *percentages* porovnává hodnoty jedné proměnné jako procentní vyjádření vzhledem k druhé proměnné (*As % of:*). Stejně jako v předchozích operacích lze použít místo proměnných i zvolenou konstantu.

Průměr (viz obrázek 19): Funkce *average* vypočítá průměr jedné nebo více proměnných.

Procentní růst (viz obrázek 20): Funkce *percent growth* vypočítá procentní vyjádření hodnot jedné proměnné (*add as first variable*) vzhledem k hodnotám druhé proměnné (*add as second variable*).

Obrázek 21 ukazuje dialog funkce rekódování proměnných, která umožňuje většinu běžných změn uspořádání hodnot proměnných potřebných při analýze dat. Realizace složitějších nebo rozsáhlých transformací je ovšem pomocí této funkce poměrně náročná a někdy neproveditelná.

Systém umožňuje stažení datového souboru do většiny používaných statistických



Download dataset (survey data only)

kých formátů (SPSS, STATA, Statistica, SAS,...). Stahování kompletního datového souboru je podmíněno registrací a podáním žádosti archivu.

Uživatelé se do nabídky stažitelných výzkumů dostanou z hlavní webové stránky archivu pomocí volby „Objednání dat – generování objednávkového formuláře“ (viz obrázek 22). Po zadání uživatelského jména a hesla se otevře okno s možností volit ze seznamu uložených výzkumů, po potvrzení je automaticky vygenerován objednávkový formulář, ten je možné uložit nebo ihned vytisknout. Formulář je nutno poslat na adresu archivu poštou nebo faxem. Archiv po jeho obdržení zpřístupní možnost stažení dat v rámci systému NESSTAR.

Obrázek 22. Objednání dat v SDA

Zdroj: Sociologický datový archiv SOÚ AV ČR.



Export table to spreadsheet

Umožňuje exportovat tabulku, graf nebo výsledek analýzy do formátu Microsoft Excel.



Export pdf

Umožňuje exportovat tabulku, graf nebo výsledek analýzy do formátu PDF.



Print preview (and print)

Ukáže náhled pracovního okna a otevře dialogové okno tisku.



Create bookmark

Umožňuje registrovaným uživatelům vytvářet záložky jak v rámci samotného NESSTARu (Server bookmark), tak záložky do internetového prohlížeče (Browser bookmark).

Poznámky

- 1 O systému NESSTAR vyšly v časopise již dva články ([Lach 2002] a [Kalvas 2006]), na které je také možno odkázat zájemce o podrobnější informace. Kdo by se chtěl dozvědět více o technických vlastnostech NESSTARu, tomu lze doporučit například [Assiny 2002]. Zde pouze podáváme úvod do celé problematiky.
- 2 S touto komercializací je spojena i nepříjemná skutečnost, že NESSTAR již není pro uživatele zadarmo, ale je nutno za licenci k jeho používání platit. Cena je ovšem odstupňovaná podle velikosti datového archivu a není rozhodně nijak zničující.
- 3 Je samozřejmé, že i při rozšíření NESSTARu by data byla stěžejně dostupná bezplatně. Ovšem alespoň informace o těchto datech by dostupné byly. V soukromém sektoru vzniká celá řada potenciálně využitelných dat, jejichž zakoupení by vyšlo levněji než vlastní pořízení.
- 4 DDI – Data Documentation Initiative je sdružení, které vyvíjí standardy datové dokumentace výzkumů ze společenských věd (viz článek v tomto čísle časopisu [Vávra 2007]).

Literatura

- Assiny, P. 2002. *NESSTAR: A Semantic Web Application for Statistical Data and Metadata*. Bergen: NESSTAR. <http://www.nesstar.org/Release-free>
- ICPSR 2005. *Guide to Social Science Data Preparation and Archiving. Best Practice Throughout the Data Life Cycle*. Ann Arbor: Inter-university Consortium for Political and Social Research, University of Michigan. <http://www.icpsr.umich.edu/access/dataprep.pdf>
- Lach, J. 2002. „NESSTAR: operace s daty přes internet“. *SDA Info* 7 (1).
- Kavas, F. 2006. „NESSTAR a DDI pro uživatele datových služeb“. *SDA Info* 4 (1–2).
- Krejčí, J. 2007. „Evropský systém datových služeb“. *Data a výzkum – SDA Info* 1 (1).
- Vávra, M. 2007. „Archivace sociologických dat“. *Data a výzkum – SDA Info* 1 (1).
- NESSTAR. (nedatováno). *Nesstar WebView User Guide, version 3.5*. Bergen: NESSTAR. http://www.nesstar.org/3.5/UserGuides/WebView_UserGuide_v3.5.pdf