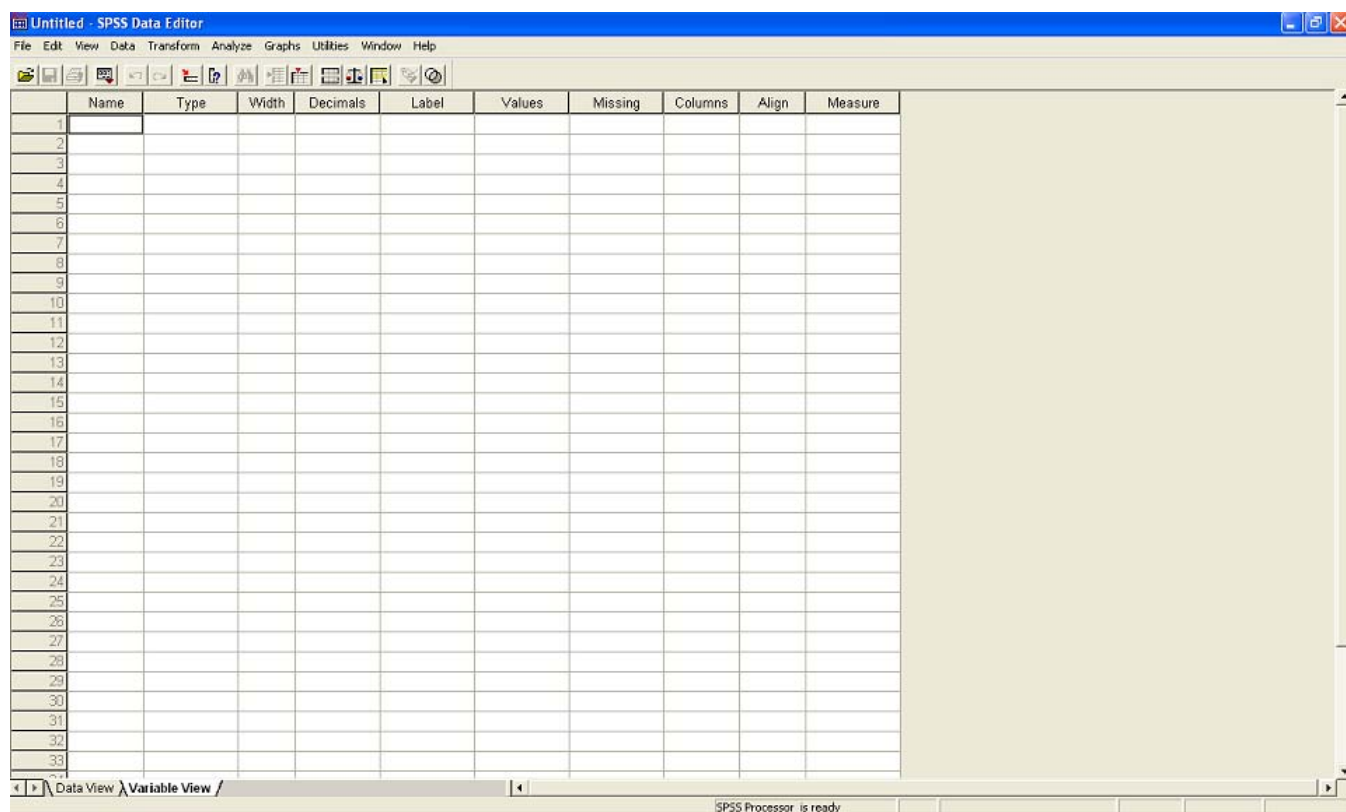


Prvé kroky v SPSS

Statistical Package for Social Science (SPSS) je počítačový program pracujúci pod operačným systémom Windows a špecializovaný na klasifikáciu, spracovanie a analýzu dát. Ako naznačuje už jeho názov, dá sa použiť najmä na analýzu v oblasti sociologického a marketingového výskumu. Je však vhodný aj na aplikácie v inžinierskych špecializáciách, v matematike a iných prírodných vedách.

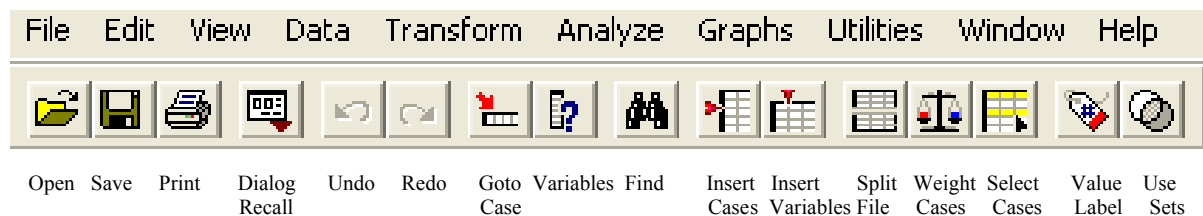
Program SPSS naštartujeme v operačnom systéme Windows nasledovne: Start → Programs → SPSS for Windows. Otvorí sa dátové okno SPSS Data Edition window, ktoré je zobrazené na obrázku:



Obrázok 1

SPSS pracuje s dvoma oknami: **Data Editor** – pre vstup a editovanie dát a okno **Output Viewer** – pre tlač výsledkov štatistického vyhodnocovania – tabuľky a grafy. Prepínanie medzi oknami zabezpečuje tabulátor v dolnej lište obrazovky. V dátovom okne (SPSS Data Editor) sa vkladajú, definujú a editujú štatistické dáta. Zobrazovacie okno sa aktivuje automaticky vždy po ukončení štatistického spracovania dát. Zobrazia sa v ňom všetky vypočítané štatistické indexy a grafické údaje.

Menu je v oboch oknách rovnaké, ale ikony sú odlišné. Menu programu SPSS obsahuje niektoré zo základných sub-menu programov pre Windows, ako aj niektoré špeciálne funkcie. V editačnom okne Data Editor menu vyzerá nasledovne:



Obrázok 2

Príklad: V tabuľke 1 sú zobrazené základné indexy banky UniCredit Bulbank. Každý rok sa považuje za samostatnú jednotku, a indexy za samostatné kvantivy. Matica dát, ktorá je zobrazená v nasledujúcej tabuľke, má rozmery 7×5 .

Tabuľka 1

Základné indexy (v miliónoch) banky UniCredit Bulbank za obdobie rokov 2000-2006

№	Year	Net profit	Own funds	Assets	Client deposits	Credits
		X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}
1	2000	160,065	602,776	2 559,476	1 692,270	316,380
2	2001	68,912	490,479	2 731,686	2 021,634	362,353
3	2002	79,130	513,249	2 721,980	2 049,957	563,935
4	2003	89,752	550,026	2 825,439	2 177,781	916,634
5	2004	86,279	574,112	3 614,696	2 917,169	1 393,968
6	2005	96,116	609,609	3 474,829	2 618,771	1 706,858
7	2006	120,654	630,781	4 346,594	3 336,875	2 131,577

Ak sú rôzne premenné uvádzané v rôznych jednotkách a mierkach, je potrebné zabezpečiť ich štandardizáciu. Ak sa majú napríklad premenné rôznych typov – interval a nominál, používať v klastrovej analýze, musíme ich štandardizovať na jeden typ – interval. Štandardizácia sa dá v programe SPSS uskutočniť automaticky, pomocou rôznych metód. Najznámejšou je metóda Z scores, pomocou ktorej sa hodnoty všetkých premenných transformujú na ten istý rozsah od -3 do +3 a potom sú premenené na interval. Pomocou Z-metódy budú mať získané štandardizované hodnoty strednú hodnotu nula a štandardnú odchýlku 1. Iné metódy pracujú v rozsahu intervalu od -1 do

1 alebo od 0 do 1, alebo v štandardizácii vzhľadom na maximálnu hodnotu, strednú hodnotu, či štandardnú odchýlku. Ak sú premenné uvedené v tých istých jednotkách a mierkach, ich štandardizácia sa neodporúča. V predchádzajúcom príklade boli pre rôzne pozorované údaje použité iba symboly X_{ij} (prvky matice X), pričom však musíme pamätať na to, že tieto symboly sa môžu vzťahovať na neštandardizované alebo na štandardizované premenné, v súvislosti s predchádzajúcimi úvahami.

V štatistike platí, že premenné (indikátory), ktorých hodnoty sú udané slovné, sa nazývajú nečíselné (**character -znak**). Väčšina údajov v anketách je práve tohto typu, pretože sú to otázky, na ktoré je potrebné získať verbálnu odpoveď. Ostatné premenné sú numerického typu, ich hodnoty sú vyjadrené číslami.

Prednostne sa používajú číselné premenné, pretože to uľahčuje štatistické spracovanie a analýzu. Z tohto dôvodu sa tiež nečíselné indikátory umelo konvertujú na numerické. Napríklad premenná „rod“, ktorá je nečíselná, je umelo konvertovaná na numerickú premennú zakódovaním mužov hodnotou 1 a žien hodnotou 2. Takto sa dotazníkové ankety spracúvajú a pripravujú na štatistické spracovanie a analýzu.

Vstup dát sa v programe SPSS uskutočňuje postupne, ako prvé sa ukladajú názvy premenných. Nachádzajú sa v prvom riadku tabuľky, v okne **Data View**, kde je na začiatku v každej bunke napísané **var**. Po prepnutí do okna **Variable View**, kde je potrebné zapísať názov, typ a opis (vysvetlenie) každej premennej.

	net_profit	own_funds	assets	client_deposits	credits	var
1	160,065	602,776	2559,476	1692,270	316,380	
2	68,912	490,479	2731,686	2021,634	362,353	
3	79,130	513,249	2721,980	2049,957	563,935	
4	89,752	550,026	2825,439	2177,781	916,634	
5	86,279	574,112	3614,696	2917,169	1393,968	
6	96,116	609,609	3474,829	2618,771	1706,858	
7	120,654	630,781	4346,594	3336,875	2131,577	
8						
9						
10						

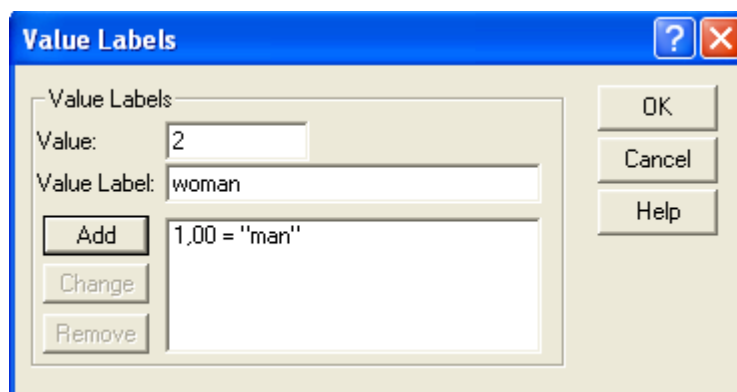
Obrázok 3

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	net_profit	Numeric	8	3	Net profit	None	None	15	Right	Scale
2	own_funds	Numeric	8	3	Own funds	None	None	15	Right	Scale
3	assets	Numeric	8	3	Assets	None	None	15	Right	Scale
4	client_deposits	Numeric	8	3	Client deposits	None	None	15	Right	Scale
5	credits	Numeric	8	3	Credits	None	None	15	Right	Scale
6										
7										

Obrázok 4


Do stĺpca **Name** zapíšeme krátky názov premennej, ktorý musí byť jedinečný. To znamená, že dve rôzne premenné nemôžu mať ten istý názov. Potom sa pomocou kurzora ovládaného myšou presunieme do stĺpca **Type** a vyberieme typ premennej. V našom prípade je to číselná premenná **Numeric**. Šírku stĺpca môžeme zmeniť pomocou hodnoty v stĺpci **Width** a počet desatinných miest hodnotou v stĺpci **Decimal Places**. Pole s názvom **Label** sa používa na opis premennej, prípadne jej dlhší názov, a pole **Values** na opis jej hodnôt.

Pri konverzii číselnej premennej na nečíselnú, sa kódy a opisy premenných vkladajú do stĺpca **Values**. Po ťuknutí na zelené tlačidlo, nachádzajúce sa v pravej časti stĺpca, sa objaví dialógové okno **Value Labels**. Do poľa Value sa zapíše kód, ktorým bude zakódovaná príslušná kategória indikátora, napríklad 1, a do stĺpca Value Label zapíšeme zodpovedajúci slovný opis muž - „man”. Ťuknutím na Add sa údaj 1=„man” zapíše do príslušného poľa. To isté zopakujeme pre ženy – do stĺpca Value vložíme 2 a do stĺpca Value Label hodnotu „woman” a potvrdíme stlačením Add. Rovnakú procedúru zopakujeme pre všetky kategórie indikátorov a potvrdíme tlačidlom OK.



Obrázok 5

Pri vkladani hodnot číselných premenných zapíšeme do každej bunky príslušný číselný údaj, pri vkladani hodnot nečíselných premenných zodpovedajúci kód. Napríklad pre nečíselnú premennú rod uvedieme kód 2, ak je prvou respondentkou ankety žena.

Opísané kódovanie sa dá zabezpečiť aj jednoduchšie – aktiváciou okna označeného  z menu. Prejdeme na príslušnú bunku nečíselného indikátora a vyberieme z ponuky v zobrazenom zozname možných definícií indikátora – v tomto prípade „man” alebo „woman”. Všetko zopakujeme pre každú nečíselnú premennú. Opisy nedefinujeme pre číselné premenné, pretože tieto nemajú kódy.

Pomocou **Missing Values** môžeme definovať niektoré hodnoty ako chýbajúce, zaradiť ich medzi ostatné a analyzovať ich oddelene. Takto môžeme kódovať otázky s odpoveďami typu "Neviem", ktoré by sa nemali stratiť, ale nie sú nevyhnutné pre analýzu. Šírka stĺpca sa dá definovať pomocou **Columns**, zarovnanie v **Align**, a pomocou hodnoty v stĺpci **Measure** môžeme definovať rozsah premennej. Číselné hodnoty sú interval/mierka, a symbolické - nominály. Môžeme definovať ľubovoľné číselné premenné, kódované ako nominály, a iné, pri ktorých je podstatné poradie dôležitosti, ako radoví premenné - ordinály. Premenná sa hneď potom, ako ju definujeme, objaví v prvom stĺpci tabuľky. Takto postupne definujeme všetky premenné.