



Renáta

Benda Prokeiová

štatistika v **SAS**-e

SAS v štatistike



© 2014 Renáta Benda -Prokeinová

Obsah tejto publikácie je chránený v zmysle Autorského zákona č. 383/1997 Zb. v znení neskorších predpisov a príslušnými platnými právnymi predpismi Slovenskej republiky. Publikovanie resp. ďalšie šírenie časti alebo celého obsahu publikácie akýmkoľvek spôsobom bez predchádzajúceho písomného súhlasu autorky je zakázané bez ohľadu na uvedenie či neuvedenie zdroja.

Recenzenti:

prof. Dr.Ing. Elena Horská

Mgr. Katarína Frühwaldová, PhD.

Doc. Ing. Peter Obtulovič, CSc.

ISBN 978-80-972772-6-0

EAN 9788097277260

Obsah

Zoznam ilustrácií.....	4
Zoznam tabuliek.....	7
Úvod.....	9
1 Import údajov v SAS Enterprise Guide.....	10
2 Zobrazenie údajov.....	17
3 Sumárne štatistiky.....	29
4 Sumárne tabuľky.....	50
5 Distribučná analýza.....	59
6 Jednostupňové triedenie početností.....	88
7 Tabuľkové analýzy.....	99
7.1 Prípadová štúdia – ukážka využitia systému SAS pri marketingovom prieskume.....	111
8 Testovanie štatistických hypotéz.....	115
8.1 Test zhody strednej hodnoty s konštantou.....	116
8.2 Test zhody dvoch stredných hodnôt pre závislé súbory.....	125
8.3 Test zhody dvoch stredných hodnôt pre nezávislé súbory.....	130
9 ANOVA.....	137
10 Regresia.....	153
10.1 Lineárny regresný model.....	154
10.2 Viacnásobný lineárny regresný model.....	162
11 Korelačná analýza.....	174
11.1 Korelačná analýza v SAS Enterprise Guide 4.2.....	176
11.2 Prípadová štúdia Korelačná analýza aplikovaná na výskume o vnímaní imidžu krajiny pôvodu produktu realizovanom na Slovensku.....	184
11.3 Prípadová štúdia Aplikácia cronbach alfa koeficientu v marketingovom výskume.....	187
12 Zhuková analýza.....	190
12.1 Zhuková analýza v SAS Enterprise Guide 4.2.....	191
Vnímanie aspektov kvality potravinárskych produktov na Slovensku – aplikácia zhukovej analýzy.....	200
12.2 Prípadová štúdia.....	205
Komparácia viacerých metód zhukovej analýzy.....	205
Summary.....	212
Zoznam literatúry.....	213

Zoznam ilustrácií

Obrázok 1 Uvítacie dialógové okno	10
Obrázok 2 Import údajov	11
Obrázok 3 Otvorenie importovaného súboru	12
Obrázok 4 Import súboru krok 1	13
Obrázok 5 Import súboru krok 2	13
Obrázok 6 Import súboru krok 3	14
Obrázok 7 Import súboru krok 4	15
Obrázok 8 Zobrazenie ikoniek importovania na pracovnej ploche softvéru	15
Obrázok 9 Zobrazenie naimportovanej databázy	16
Obrázok 10 Zobrazenie naimportovanej databázy	17
Obrázok 11 Zobrazenie položky Data	18
Obrázok 12 Zoznam premenných vypísaných na obrazovke	19
Obrázok 13 Priraďovanie úloh premenným podľa skupín	19
Obrázok 14 Zobrazenie premenných na obrazovke podľa skupín	20
Obrázok 15 Zobrazenie premenných na obrazovke podľa strany	20
Obrázok 16 Zobrazenie premenných na obrazovke so sumami	21
Obrázok 17 Zobrazenie premenných na obrazovke s medzisúčtami	22
Obrázok 18 Zobrazenie premenných na obrazovke bez poradového čísla	22
Obrázok 19 Možnosti zobrazenia premenných na obrazovke	23
Obrázok 20 Hlavička a päta zobrazenia premenných na obrazovke	25
Obrázok 21 Vlastnosti zobrazenia premenných na obrazovke	26
Obrázok 22 Editácia vlastností zobrazenia premenných na obrazovke	26
Obrázok 23 Riešenie úlohy 2.1	27
Obrázok 24 Zobrazenie úlohy Summary Statistics	29
Obrázok 25 Zobrazenie položky Data	30
Obrázok 26 Zobrazenie položky Data a možnosti klasifikačnej premennej	31
Obrázok 27 Zobrazenie položky Základné štatistiky	33
Obrázok 28 Zobrazenie položky Percentily	35
Obrázok 29 Zobrazenie položky Dodatočné štatistiky	37
Obrázok 30 Zobrazenie položky Grafy	38
Obrázok 31 Škatuľový graf	39
Obrázok 32 Riešenie úlohy 3.1	40
Obrázok 33 Riešenie úlohy 3.1	41
Obrázok 34 Riešenie úlohy 3.2.1	42
Obrázok 35 Riešenie úlohy 3.2.2	44
Obrázok 36 Riešenie úlohy 3.3.1	45
Obrázok 37 Riešenie úlohy 3.3.2	47
Obrázok 38 Zobrazenie úlohy Sumárne štatistiky	50
Obrázok 39 Zobrazenie položky Data v úlohe Sumárne štatistiky	51
Obrázok 40 Zobrazenie položky Sumárne tabuľky v úlohe Sumárne štatistiky	52
Obrázok 41 Zobrazenie viacerých štatistik v úlohe Sumárne štatistiky	53
Obrázok 42 Vytváranie štruktúrovaných sumárnych tabuliek	54
Obrázok 43 Formátovanie hodnôt v sumárnych tabuľkách	54
Obrázok 44 Vytváranie súčtov a medzisúčtov v sumárnych tabuľkách	55
Obrázok 45 Riešenie úlohy 4.1	55
Obrázok 46 Riešenie úlohy 4.2	57
Obrázok 47 Zobrazenie úlohy Distribučná analýza	59
Obrázok 48 Zobrazenie položky Data	60
Obrázok 49 Gaussova krivka normálneho rozdelenia s pravidlom 3 sigma	63
Obrázok 50 Typy rozdelení a možnosti normálneho rozdelenia	64

Obrázok 51 Typy grafov	65
Obrázok 52 Možnosti v položke Tabuľky.....	67
Obrázok 53 Riešenie úlohy 5.1	77
Obrázok 54 Riešenie úlohy 5.2	80
Obrázok 55 Výstup úlohy 5.2.....	81
Obrázok 56 Riešenie úlohy 5.3	82
Obrázok 57 Riešenie úlohy 5.4	84
Obrázok 58 Riešenie úlohy 5.5	87
Obrázok 59 Výstup úlohy 5.5.....	87
Obrázok 60 Zobrazenie úlohy Jednostupňového triedenia.....	88
Obrázok 61 Zobrazenie položky Data	89
Obrázok 62 Štatistiky jednostupňovej tabuľky	90
Obrázok 63 Typy grafov jednostupňovej tabuľky	94
Obrázok 64 Vertikálny graf jednostupňovej tabuľky – histogram.....	94
Obrázok 65 Možnosti výsledkov jednostupňovej tabuľky.....	95
Obrázok 66 Riešenie úlohy 6.1	96
Obrázok 67 Zobrazenie úlohy Tabuľková analýza	99
Obrázok 68 Zobrazenie položky Data.....	100
Obrázok 69 Zobrazenie položky Data.....	100
Obrázok 70 Zobrazenie položky Tabuľky.....	101
Obrázok 71 Zobrazenie položky Tabuľky – vytvorenie kontingenčnej tabuľky.....	101
Obrázok 72 Zobrazenie položky Štatistiky v bunkách.....	102
Obrázok 73 Zobrazenie položky Tabuľkové štatistiky – Asociácie.....	107
Obrázok 74 Zobrazenie položky Tabuľkové štatistiky – Možnosti výpočtu.....	109
Obrázok 75 Zobrazenie databázy k Testu zhody strednej hodnoty s konštantnou.....	117
Obrázok 76 Zobrazenie úlohy t-test.....	117
Obrázok 77 Výber Testu zhody strednej hodnoty s konštantou.....	118
Obrázok 78 Výber položky Data.....	119
Obrázok 79 Stanovenie nulovej hypotézy.....	119
Obrázok 80 Typy grafov	120
Obrázok 81 Sumárny graf	121
Obrázok 82 Histogram	121
Obrázok 83 Škatuľkový graf.....	122
Obrázok 84 Graf intervalu spoľahlivosti pre strednú hodnotu	122
Obrázok 85 Q-Q graf.....	123
Obrázok 86 Vlastnosti a nadpisy reportu úlohy t-test	123
Obrázok 87 Riešenie úlohy 8.1	124
Obrázok 88 Vstupná databáza – Test zhody dvoch stredných hodnôt pre závislé súbory	126
Obrázok 89 Výber testu – Test zhody dvoch stredných hodnôt pre závislé súbory.....	127
Obrázok 90 Položka Data v úlohe t-test.....	128
Obrázok 91 Stanovenie hypotézy.....	128
Obrázok 92 Riešenie úlohy 8.2	130
Obrázok 93 Zobrazenie vstupnej databázy – Test zhody dvoch stredných hodnôt pre nezávislé súbory	131
Obrázok 94 Výber testu – Test zhody dvoch stredných hodnôt pre nezávislé súbory	132
Obrázok 95 Položka Data úlohy t-test.....	132
Obrázok 96 Stanovenie hypotézy.....	133
Obrázok 97 Riešenie úlohy 8.3	134
Obrázok 98 Zobrazenie vstupnej databázy.....	138
Obrázok 99 Zobrazenie úlohy ANOVA.....	139
Obrázok 100 Zobrazenie položky Data	140
Obrázok 101 Zobrazenie položky Testy.....	141

Obrázok 102	Zobrazenie položky Testy mnohonásobného porovnania	142
Obrázok 103	Zobrazenie položky Základné popisné charakteristiky	145
Obrázok 104	Zobrazenie typov grafov v úlohe ANOVA.....	146
Obrázok 105	Nastavenie výsledkov a nadpisov v úlohe ANOVA.....	147
Obrázok 106	Zobrazenie vlastností v úlohe ANOVA.....	148
Obrázok 107	Riešenie úlohy 9.1	149
Obrázok 108	Riešenie úlohy 9.2	151
Obrázok 109	Zobrazenie úlohy Lineárna regresia	155
Obrázok 110	Zobrazenie položky Data.....	155
Obrázok 111	Zobrazenie položky Model.....	156
Obrázok 112	Zobrazenie položky Štatistiky	157
Obrázok 113	Zobrazenie položky Grafy	159
Obrázok 114	Riešenie úlohy 10.1	160
Obrázok 115	Zobrazenie databázy Viacnásobnej lineárnej regresie.....	163
Obrázok 116	Zobrazenie úlohy Viacnásobnej lineárnej regresie.....	164
Obrázok 117	Zobrazenie položky Data.....	165
Obrázok 118	Zobrazenie položky Model.....	165
Obrázok 119	Zobrazenie položky Grafy, Vlastnosti, Nadpisy.....	168
Obrázok 120	Riešenie úlohy 10.2.1	169
Obrázok 121	Riešenie úlohy 10.2.2	171
Obrázok 122	Príklady rôznych korelácií.....	175
Obrázok 123	Zobrazenie úlohy Korelácia.....	176
Obrázok 124	Zobrazenie položky Data v úlohe Korelácia.....	177
Obrázok 125	Zobrazenie položky Option v úlohe Korelácia.....	178
Obrázok 126	Zobrazenie položky Result v úlohe Korelácia.....	180
Obrázok 127	Zobrazenie položky Output a Tittles v úlohe Korelácia.....	181
Obrázok 128	Zobrazenie položky Properties v úlohe Korelácia.....	181
Obrázok 129	Zobrazenie procesu výpočtu korelácie	182
Obrázok 130	Zobrazenie výstupu výpočtu korelácie	183
Obrázok 131	Typ otázky vhodný na zhlukovú analýzu	191
Obrázok 132	Upravená databáza vhodná na výpočet zhlukovej analýzy	192
Obrázok 132	Zobrazenie úlohy Cluster analysis.....	193
Obrázok 132	Zobrazenie položky Data v úlohe Cluster analysis.....	194
Obrázok 133	Zobrazenie položky Cluster v úlohe Cluster analysis.....	195
Obrázok 134	Zobrazenie položky Plots v úlohe Cluster analysis	196
Obrázok 135	Zobrazenie položky Results a Tittles v úlohe Cluster analysis	197
Obrázok 136	Zobrazenie položky Properties v úlohe Cluster analysis.....	198
Obrázok 137	Proces a výstup úlohy Cluster analysis.....	199
Obrázok 138	Dendrogram pre Slovensko	203
Obrázok 139	Zhluk Single metódou	207
Obrázok 140	Zhluk Ward metódou.....	208
Obrázok 141	Zhluk Median metódou	210

Zoznam tabuliek

Tabuľka 1 Výstup úlohy 2.1.....	27
Tabuľka 2 Výstup úlohy 3.1.....	40
Tabuľka 3 Výstup úlohy 3.2.1.....	42
Tabuľka 4 Výstup úlohy 3.2.2.....	44
Tabuľka 5 Výstup úlohy 3.3.1.....	45
Tabuľka 6 Výstup úlohy 3.3.2.....	47
Tabuľka 7 Výstup úlohy 4.1.....	56
Tabuľka 8 Výstup úlohy 4.2.....	57
Tabuľka 9 Intervaly spoľahlivosti.....	67
Tabuľka 10 Základné popisné charakteristiky.....	68
Tabuľka 11 Test zhody priemeru.....	69
Tabuľka 12 Extrémne riadky.....	70
Tabuľka 13 Extrémne hodnoty.....	71
Tabuľka 14 Početnosti.....	71
Tabuľka 15 Modusy.....	72
Tabuľka 16 Momenty.....	72
Tabuľka 17 Kvantily.....	74
Tabuľka 18 Robustné miery škál.....	74
Tabuľka 19 Testy normality.....	75
Tabuľka 20 Orezané priemery.....	76
Tabuľka 21 Nahradzované priemery.....	76
Tabuľka 22 Výstup úlohy 5.1.....	77
Tabuľka 23 Výstup úlohy 5.2.....	80
Tabuľka 24 Výstup úlohy 5.3.....	82
Tabuľka 25 Výstup úlohy 5.4.....	84
Tabuľka 26 Absolútne a relatívne početnosti a kumulatívne početnosti.....	90
Tabuľka 27 Absolútne a kumulatívne absolútne početnosti.....	91
Tabuľka 28 Absolútne a relatívne početnosti.....	92
Tabuľka 29 Absolútne početnosti.....	92
Tabuľka 30 Výstup úlohy 6.1.....	96
Tabuľka 31 Kumulatívne percento stĺpca.....	103
Tabuľka 32 Percento riadka.....	103
Tabuľka 33 Percento stĺpca.....	104
Tabuľka 34 Absolútne početnosti v bunkách.....	104
Tabuľka 35 Relatívne početnosti.....	105
Tabuľka 36 Pearsonove chí kvadrát štatistiky.....	105
Tabuľka 37 Absolútne početnosti odchýlok od očakávaných hodnôt v bunkách.....	106
Tabuľka 38 Očakávané (teoretické) početnosti v bunkách.....	106
Tabuľka 39 Chí kvadrát testy a korelačné koeficienty.....	112
Tabuľka 40 Výstup úlohy 8.1.....	125
Tabuľka 41 Výstup úlohy 8.2.....	130
Tabuľka 42 Výstup úlohy 8.3.....	134
Tabuľka 43 Hodnotenie najčastejšie používaných testov.....	144
Tabuľka 44 Výstup úlohy 9.1.....	149
Tabuľka 45 Výstup úlohy 9.2.....	151
Tabuľka 46 Výstup úlohy 10.1.....	160
Tabuľka 47 Výstup úlohy 10.2.1.....	170

Tabuľka 48 Výstup úlohy 10.2.2.....	171
Tabuľka 49 Výstup úlohy korelačnej analýzy.....	185
Tabuľka 50 Výstup Cronbach alpha analýzy	188
Tabuľka 51 Kovariančná matica pre Slovensko	200
Tabuľka 52 Celková suma štvorcov odchýlok.....	201
Tabuľka 53 História zhlukovania pre Slovensko	201
Tabuľka 54 Zhluky podobnosti faktorov imidžu krajiny pôvodu produktu vnímané slovenskými respondentmi	203
Tabuľka 55 Zhluky podľa metódy Single	207
Tabuľka 56 Zhluky podľa Wardovej metódy.....	209
Tabuľka 57 Zhluky podľa Mediánovej metódy	210

Úvod

Skriptá štatistika v SAS-e, SAS v štatistike vznikli z dôvodu absencie študijnej literatúry k predmetu SAS ako nástroj kvantitatívnych analýz na Fakulte ekonomiky a manažmentu Slovenskej poľnohospodárskej univerzity v Nitre.

Kniha je určená nielen študentom, ale aj vedeckým pracovníkom, ktorí využívajú softvér SAS na svoje výskumy.

Zvládnutie základov štatistiky je nevyhnutné pre pochopenie ďalších sofistikovanejších metód, akými sú ekonometria, metódy prognózovania či marketingový výskum. Základy štatistiky dokážu študenti počítať aj v MS Exceli, no pri komplikovanejších vzorcoch sa často mýlia a výsledok nie je relevantný. Štatistický softvér SAS Enterprise Guide 4.2 odbúrava možnosť pomýlenia sa vo vzorci, pretože vzorec je už naprogramovaný priamo v softvéri. Študent musí presunúť pozornosť na kvalitu vstupnej databázy a možné chyby v databáze, ktoré môžu skresliť výsledok testu.

Cieľom predkladanej publikácie je poskytnúť návody na výpočty základných štatistických metód, počnúc importom databázy cez základné popisné charakteristiky, štatistickú indukciu až po ANOVU a regresiu. Zaujímavosťou publikácie je to, že kapitoly neobsahujú vzorce analyzovaných metód. Druhé vydanie je doplnené o kapitoly z korelačnej zhlukovej analýzy. Po každej kapitole nasledujú príklady, ktoré sú následne vyriešené, a zinterpretované výsledky. Kapitoly tvoria samostatné celky rozdelené podľa štruktúry v softvéri. Každá kapitola má rovnakú štruktúru: krátky popis úlohy v softvéri, stručná definícia metódy, postup a vysvetlenie jednotlivých položiek pri výpočte a následné úlohy s riešením a interpretáciami.

Pevne verím, že kniha bude prínosom nielen pre študentov, ale aj pre vedeckých pracovníkov. Zároveň uvítam všetky pripomienky čitateľov k skvalitneniu obsahu učebnice.

Na záver by som sa chcela poďakovať recenzentom: prof. Dr. Ing. Elene Horskej z SPU v Nitre, doc. Ing. Petrovi Obtulovičovi, CSc. z SPU v Nitre a Mgr. Kataríne Frühwaldovej, PhD. zo spoločnosti SAS Slovakia, s. r. o., ktorí prispeli k zlepšeniu obsahu knihy.

Osobitne vyjadrujem poďakovanie spoločnosti SAS Slovakia, s. r. o., za odbornú a finančnú podporu, bez ktorej by nemohla byť učebnica vytvorená a vydaná.

Autorka

Import údajov v SAS Enterprise Guide

1 Import údajov v SAS Enterprise Guide

Kapitola je zameraná na prácu so softvérom SAS Enterprise Guide 4.2 (ďalej len SAS EG), konkrétne na vkladanie údajov do softvéru. Je nevyhnutné vysvetliť niektoré postupy prostredníctvom screenshotov softvéru. Užívateľ, samozrejme, môže používať systém SAS HELP, ktorý je intuitívny a vhodný pre bežného užívateľa.

Po odkliknutí ikony SAS EG na ploche alebo v ponuke Štart sa otvorí okno softvéru, kde sa nachádza prázdna pracovná plocha. Vzápätí sa spustí dialógové okno, ktoré nám ponúka niekoľko možností.



Obrázok 1 Uvítacie dialógové okno

Zdroj: softvér SAS

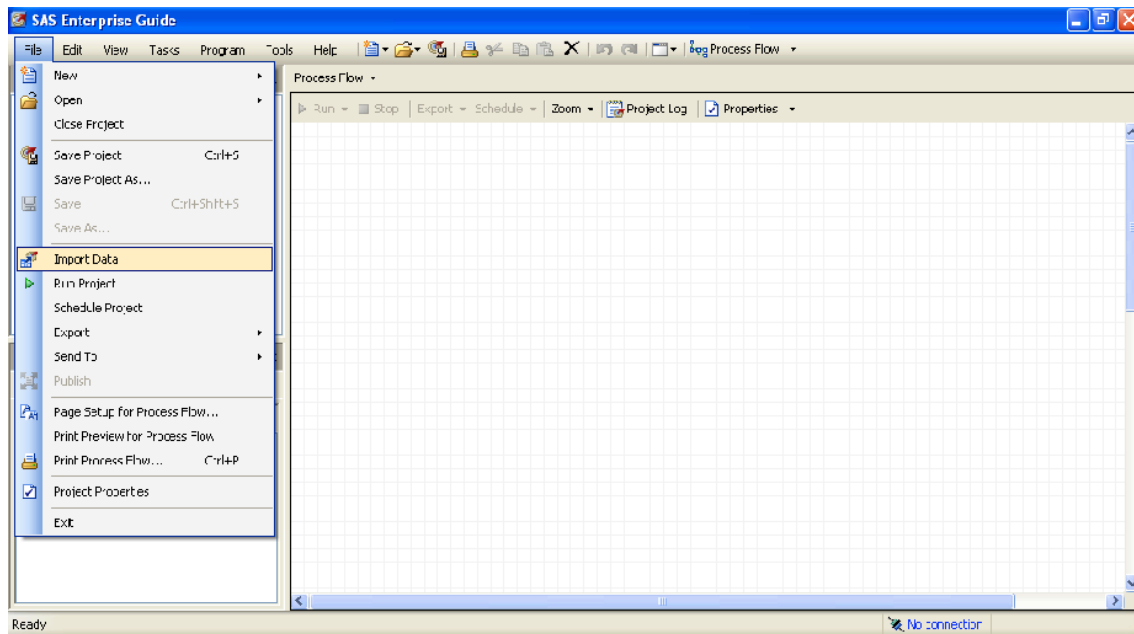
Jednotlivé výpočty v SAS EG pracujú na báze projektov. Project.egp je základná štruktúra, s ktorou užívateľ pracuje. Úvodné dialógové okno umožňuje:

- vybrať zo zoznamu existujúcich projektov,
- vytvoriť nový projekt,
- vytvoriť nový program,
- vytvoriť nový dátový súbor,
- samozrejme, neustále ponúka tutoriál pre prácu s EG.

Do EG sa priebežne ukladajú všetky analýzy vykonávané v prostredí EG. V našom prípade sme si založili nový projekt, do ktorého chceme importovať údaje z excelového súboru.

Import údajov v SAS Enterprise Guide

Údaje musia byť vhodne upravené. V hlavičke databázy (*label*) je vhodnejšie používať text bez interpunkcie, medzier a špeciálnych znakov, teda môže obsahovať len písmená, podčiarkovníky a čísla.

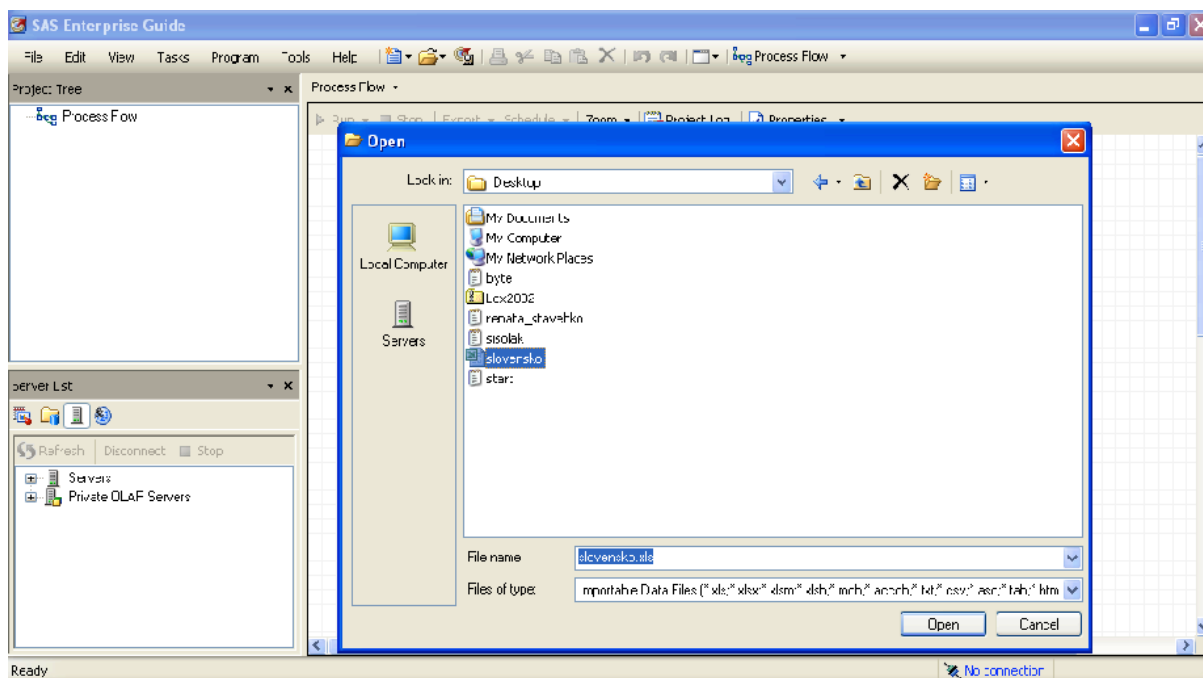


Obrázok 2 Import údajov

Zdroj: softvér SAS

Import údajov spustíme cez úlohu nachádzajúcu sa na lište nástrojov v ponuke „File“ → „Import data“. Uvedená možnosť nám automaticky otvorí dialógové okno obsahujúce dve dôležité položky *Local Computer* a *Servers*. Keďže sa súbor nachádzal v *Local Computer* na pracovnej ploche, označili sme súbor, ktorý chceme importovať, a potvrdili možnosť „Open“.

Import údajov v SAS Enterprise Guide

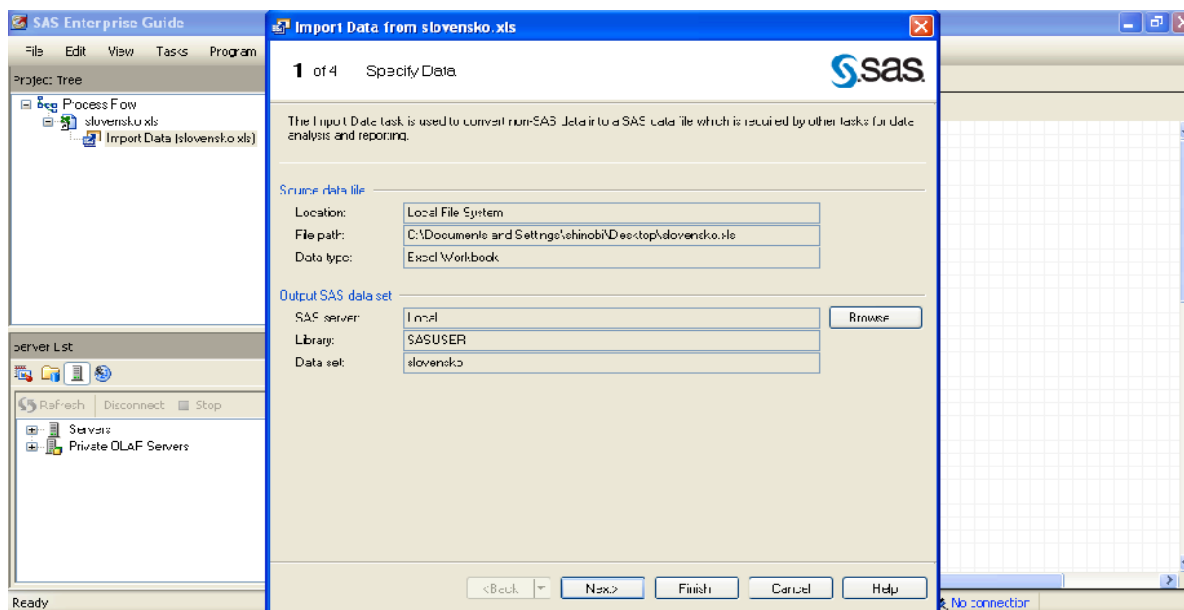


Obrázok 3 Otvorenie importovaného súboru

Zdroj: softvér SAS

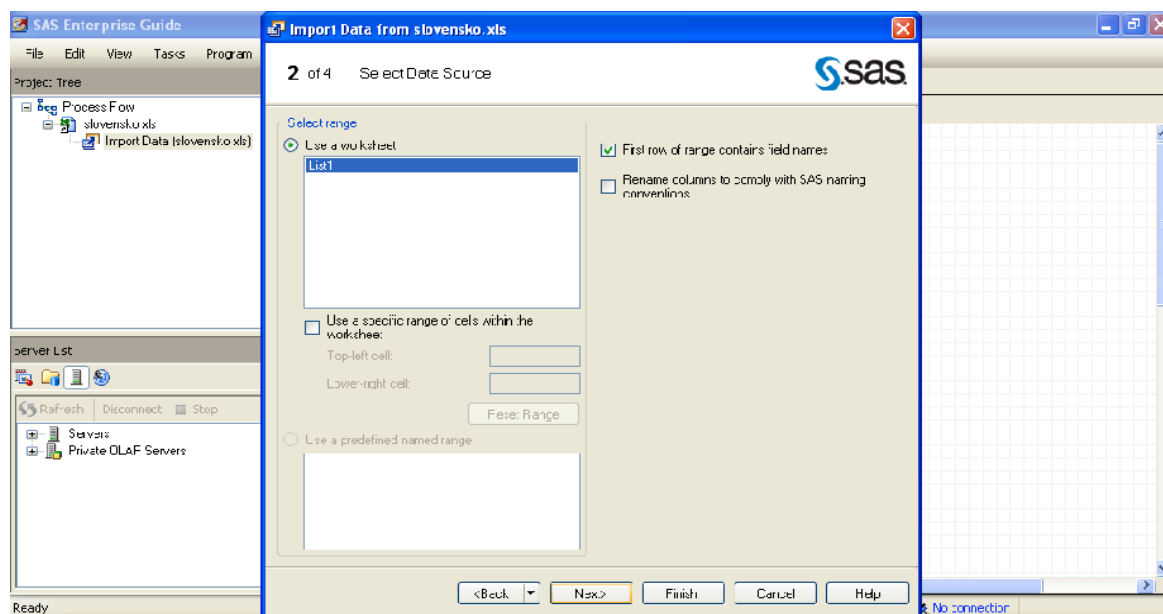
Po spustení súboru sa nám otvorí dialógové okno importovania, pozostávajúce zo 4 krokov, ktoré by bolo vhodné detailnejšie priblížiť. V prvom kroku nemusíme meniť žiadne nastavenia. Môžeme skonštatovať, že prvý krok má informatívny charakter, ktorý nám poskytuje informáciu o pozícii na disku, odkiaľ sme súbor prevzali, a následne popisuje lokalizáciu, kam databázu uložil. SAS používa spravidla tri typy knižníc – **SAS help**, **SAS user** a **Work**. **SAS help** obsahuje pomocné dátové súbory, **SAS user** predstavuje trvalú knižnicu (po ukončení práce so softvérom sa údaje nestratia) a **Work** predstavuje dočasnú knižnicu, v ktorej sa údaje neukladajú po ukončení práce so softvérom. Výhodou EG je skutočnosť, že dáta sú automaticky vkladané do knižnice *SAS user*.

Import údajov v SAS Enterprise Guide



Obrázok 4 Import súboru krok 1
Zdroj: softvér SAS

V druhom kroku importovania už EG vyžaduje aktivitu užívateľa. V danom kroku si SAS žiada informáciu, z ktorého hárku v Exceli má prevziať údaje, pretože v našom súbore sa nachádzal iba jeden hárk.

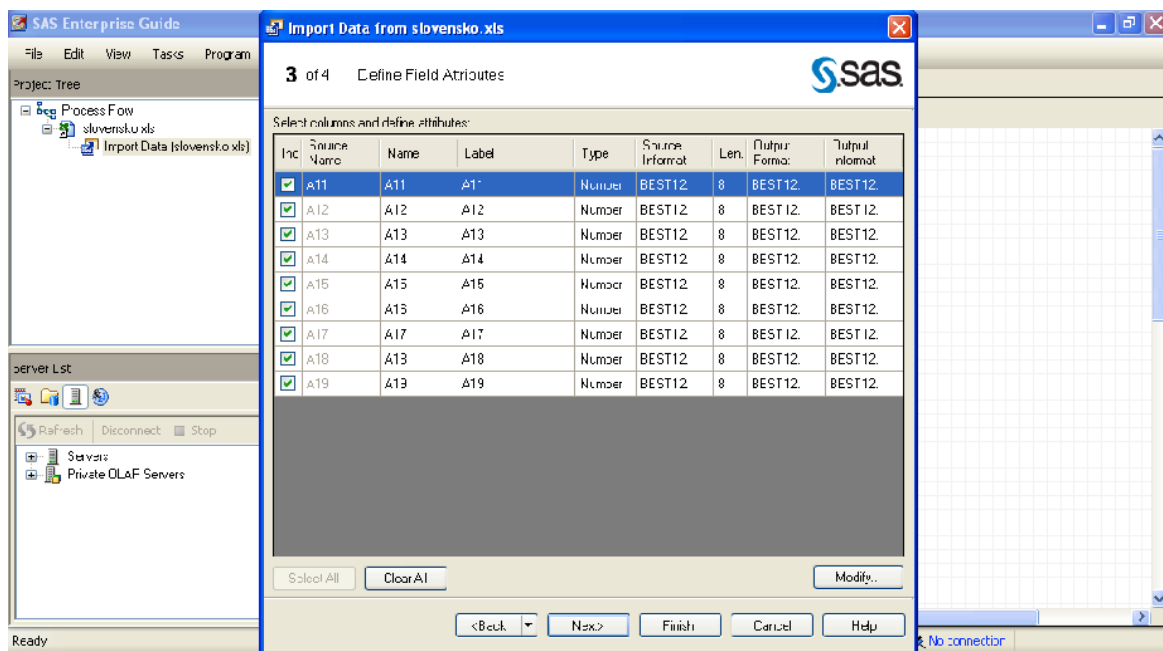


Obrázok 5 Import súboru krok 2
Zdroj: softvér SAS

Štandardne je zaškrtnuté políčko, ktoré informuje o tom, že SAS eviduje prvú hodnotu v stĺpci ako popis stĺpca.

Import údajov v SAS Enterprise Guide

V kroku 2 je dokonca možné premenovať názvy stĺpcov na SAS mená podľa SAS konvencií, v prípade, ak popisy stĺpcov neboli upravené pred importovaním. Názov stĺpca môže mať maximálne 32 znakov a úprava pri importovaní spočíva v tom, že v názvoch budú odstránené medzery a nahradí ich podčiarkovník, zároveň bude odstránená interpunkcia. Uvedenú skutočnosť dosiahneme zaškrtnutím druhého políčka.



Obrázok 6 Import súboru krok 3

Zdroj: softvér SAS

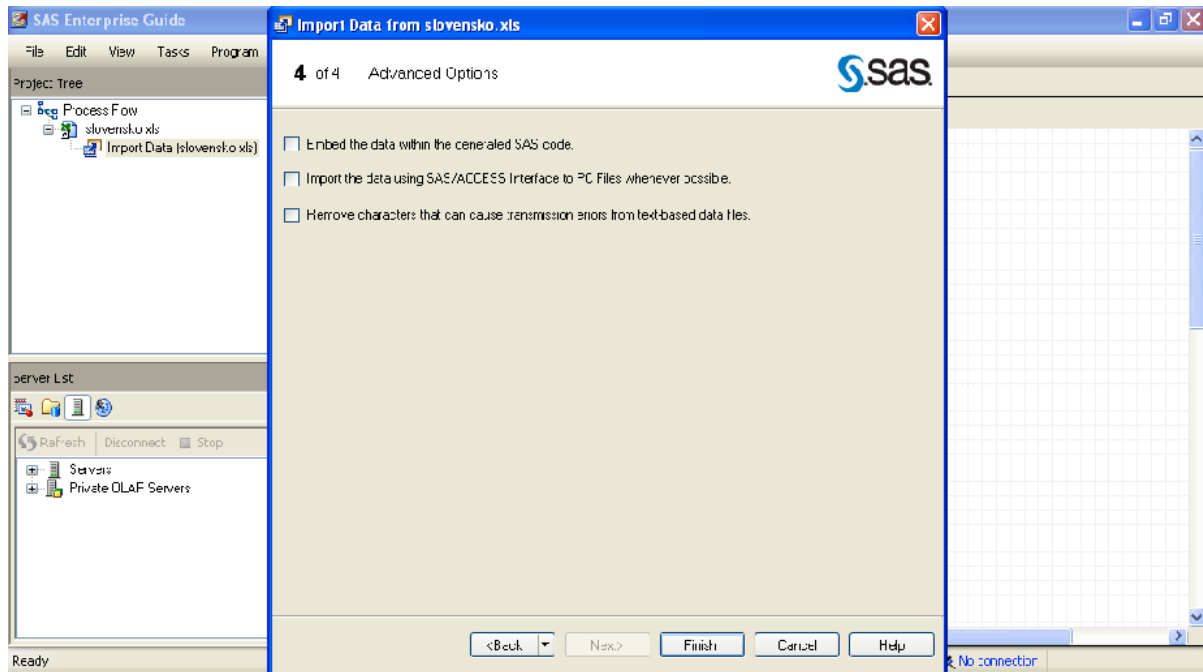
V treťom kroku vidíme formát údajov, teda ako ich zhodnotil SAS EG. Naše údaje majú povahu numerických premenných, teda v stĺpci *TYPE* vidíme definíciu *Number*. V prípade, ak by SAS EG údaje prečítal a ukázal by neadekvátny typ dát, je vhodné upraviť údaje v Exceli a opäť ich nainportovať. Priamo v SAS EG je možné tiež meniť charakter údajov. Ak je v databáze významnejšia chyba, bude pracovať aj s tou chybou a výsledky môžu byť skreslené.

V tomto kroku je nevyhnutné si zadať typ atribútov, s ktorými sa môžeme pri databázach stretnúť:

- *Number* – numerický atribút, numerická hodnota znaku,
- *String* – textový atribút, textová premenná, môže to byť slovo alebo písmeno,
- *Currency* – atribút meny,
- *Date* – atribút dátumu,
- *Time* – atribút času,
- *Date/Time* – atribút dátumu a času.

Import údajov v SAS Enterprise Guide

Typická dĺžka stĺpca v databáze má 8 znakov, môže byť aj dlhšia, aj kratšia, v závislosti od obsahu hlavičky stĺpca.

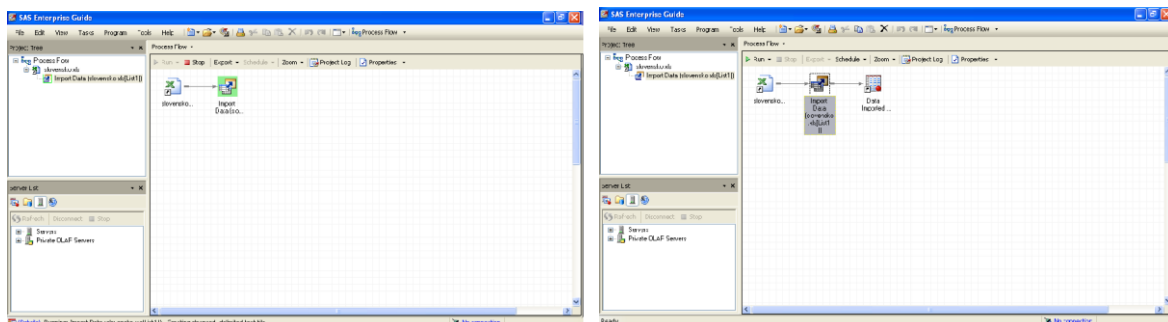


Obrázok 7 Import súboru krok 4

Zdroj: softvér SAS

V štvrtom kroku ponúka rozšírené možnosti. Ak už v predchádzajúcom kroku bola databáza v poriadku, v tomto kroku nemusíme nič zaškrtnávať a zadáme možnosť *FINISH*. V prípade, že si nie sme istí úpravami svojej databázy, môžeme zaškrtnúť niekoľko možností:

- skrátiť dĺžku údajov (týka sa najmä desatinných miest a dlhých textových premenných),
- importovať údaje použitím modulu *SAS/ACCESS* Interface do súborov v počítači, ak je to možné,
- odstrániť znaky, ktoré môžu zapríčiniť chyby v údajoch.



Obrázok 8 Zobrazenie ikoniek importovania na pracovnej ploche softvéru

Zdroj: softvér SAS

Import údajov v SAS Enterprise Guide

	Product_ID	Product_Line	Order_ID	Product_Category	Product_Name	Supplier_Name	Supplier_ID
1	220101300017	Clothes & Shoes	1230058123	Clothes	Toncot Beely-T E...	A Team Sports	32
2	230100500026	Outdoors	1230080101	Outdoors	Trekking Tent	Prime Sports Ltd	3
3	240600100090	Sports	1230106893	Swim Sports	Sharky Swimming...	Dolphin Sportswea...	16
4	240600100010	Sports	1230147441	Swim Sports	Goggles, Assorte...	Nautlius SportsWe...	61
5	240200200039	Sports	1230315085	Golf	Eagle Plain Cap	HighPoint Trading	10
6	240100400005	Sports	1230333319	Assorted Sports Articl...	Cool Fit Men's Rol...	Twain Inc	13
7	240800200062	Sports	1230338566	Winter Sports	Top Equipe 07 Gr...	Twain Inc	13
8	240800200063	Sports	1230338566	Winter Sports	Top Equipe 99 Bla...	Twain Inc	13
9	240500100004	Sports	1230371142	Running - Jogging	Pants N	Luna sastreria S.A.	47
10	240500200003	Sports	1230371142	Running - Jogging	Men's Singlet	Force Sports	58
11	220200300002	Clothes & Shoes	1230404278	Shoes	Pytossage Bathin...	3Top Sports	28
12	230100600005	Outdoors	1230440481	Outdoors	Basic 10, Left ,Ye...	Top Sports	7
13	230100600018	Outdoors	1230450371	Outdoors	Expedition Zero.S...	Top Sports	7
14	240500200056	Sports	1230453723	Running - Jogging	Breath-brief Long...	Force Sports	58
15	240200100233	Sports	1230455630	Golf	Tee18 Troon 7 BL	Van Dammeren Int...	28
16	230100100025	Outdoors	1230478006	Outdoors	Women's Shorts	Scandinavian Clot...	16
17	230100300006	Outdoors	1230498538	Outdoors	Gloves Le Fly Mitt...	Magnifico Sports	16
18	240200100131	Sports	1230500669	Golf	HrHy Tour Advanc...	Van Dammeren Int...	28
19	2201014000310	Clothes & Shoes	1230503155	Clothes	Short Women's TL	3Top Sports	28
20	220200200024	Clothes & Shoes	1230591673	Shoes	Pro Fit Gel Gt 203...	Pro Sportswear Inc	17
21	240500100039	Sports	1230591675	Running - Jogging	Sweatshirt W'ome...	A Team Sports	32
22	240200100076	Sports	1230591684	Golf	Expert Men's Fire...	Twain Inc	13
23	220200100092	Clothes & Shoes	1230619748	Shoes	Big Guy Men's Air...	Eclipse Inc	13
24	220200300005	Clothes & Shoes	1230619748	Shoes	Liga Football Boot	3Top Sports	28
25	230100500082	Outdoors	1230642273	Outdoors	Lucky Tech Inter...	AllSeasons Outdo...	7
26	240100100646	Sports	1230657844	Assorted Sports Articl...	Lyon Men's Jacket	Luna sastreria S.A.	47
27	240200100043	Sports	1230690733	Golf	Bretagne Perform...	Van Dammeren Int...	28
28	220100100044	Clothes & Shoes	1230699509	Clothes	Sports glasses Sa...	Eclipse Inc	13
29	220100400005	Clothes & Shoes	1230699509	Clothes	Big Guy Men's Air...	Eclipse Inc	13
30	230100100006	Outdoors	1230700421	Outdoors	Jacket Nome	Luna sastreria S.A.	47
31	230100500045	Outdoors	1230700421	Outdoors	Cup Picnic Mug 2...	Prime Sports Ltd	3
32	230100500068	Outdoors	1230700421	Outdoors	Plate Picnic Deep	Prime Sports Ltd	3
33	230100600026	Outdoors	1230738723	Outdoors	Expedition 20.Lar...	Top Sports	7
34	220100700024	Clothes & Shoes	1230744524	Clothes	Armadillo Road D...	Fuller Trading Co.	16

Obrázok 9 Zobrazenie naimportovanej databázy
Zdroj: softvér SAS

Následne uvádzame screenshoty hlavnej pracovnej plochy, kde vidíme ikony jednotlivých aktivít, ktoré sme práve vykonali. V závere kapitoly môžeme vidieť výsledok importovania.

Prístup k úlohám, ktoré budú popísané v tejto učebnici, môžete nájsť na základnej hornej lište v záložke *Task* alebo cez položku *Describe* nachádzajúcu sa nad importovanou databázou.

SAS Enterprise Guide poskytuje pre niektoré úlohy dve možnosti nastavenia úloh, buď pomocou *Wizards* (tzv. help, ktorý ponúka jednotlivé kroky), alebo pomocou štandardného nastavenia úlohy.

Zobrazenie údajov

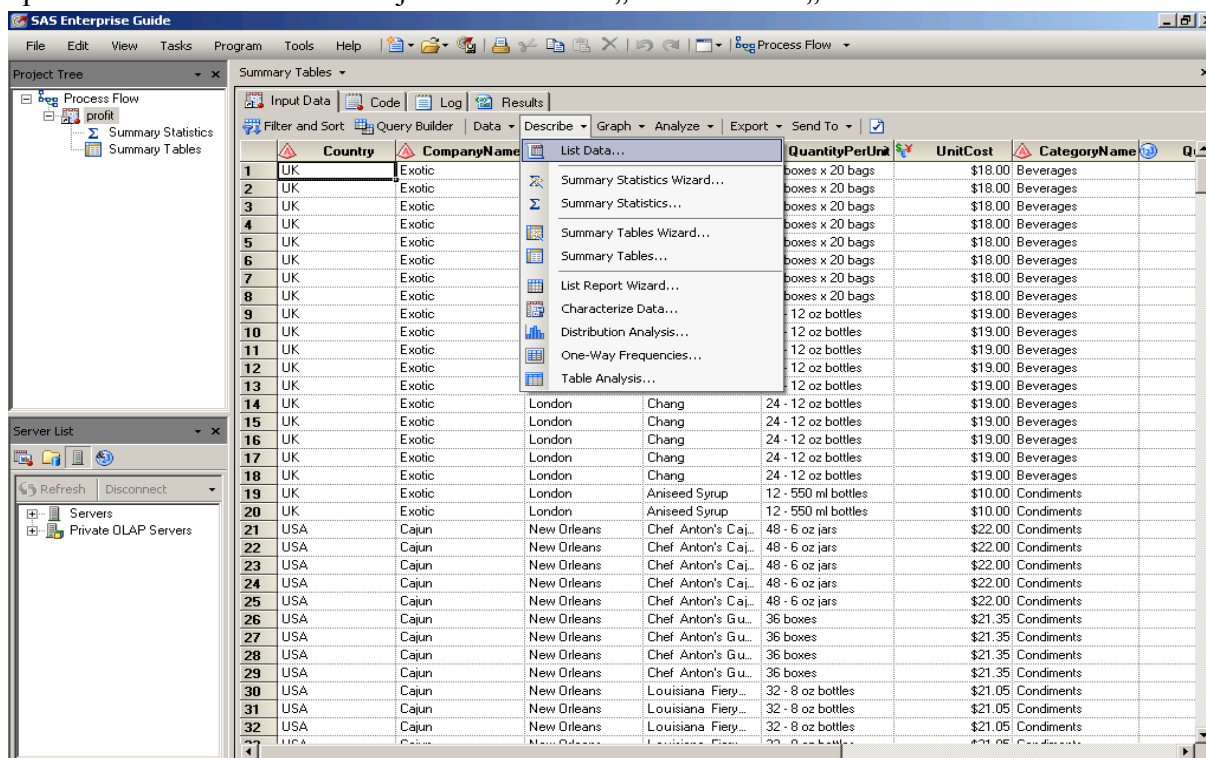
2 Zobrazenie údajov

Po naimportovaní údajov do SAS-u môžeme prísť k samotným analýzám. Z dôvodu praktického využitia dávame do pozornosti úlohu *List Data*.

List Data zobrazí na obrazovke pozorovania databázy v SAS-e. Na výstup vo forme reportu môžeme využívať všetky alebo len niektoré z premenných. Je možné vytvoriť rozličné reporty, od jednoduchého zoznamu po striktno upravovaný report, v ktorom sú údaje zoskupované a prepočítavané cez súčty a medzisúčty numerických premenných.

Napríklad, môžeme použiť *List Data* na vytvorenie reportu, ktorý spočíta sumy nákladov a tržieb pre každý región, a porovnáme tieto sumy s hodnotami celkových nákladov a tržieb pre každú spoločnosť.

Spomínanú úlohu *List Data* nájdeme v záložke „Describe“ → „List data“.

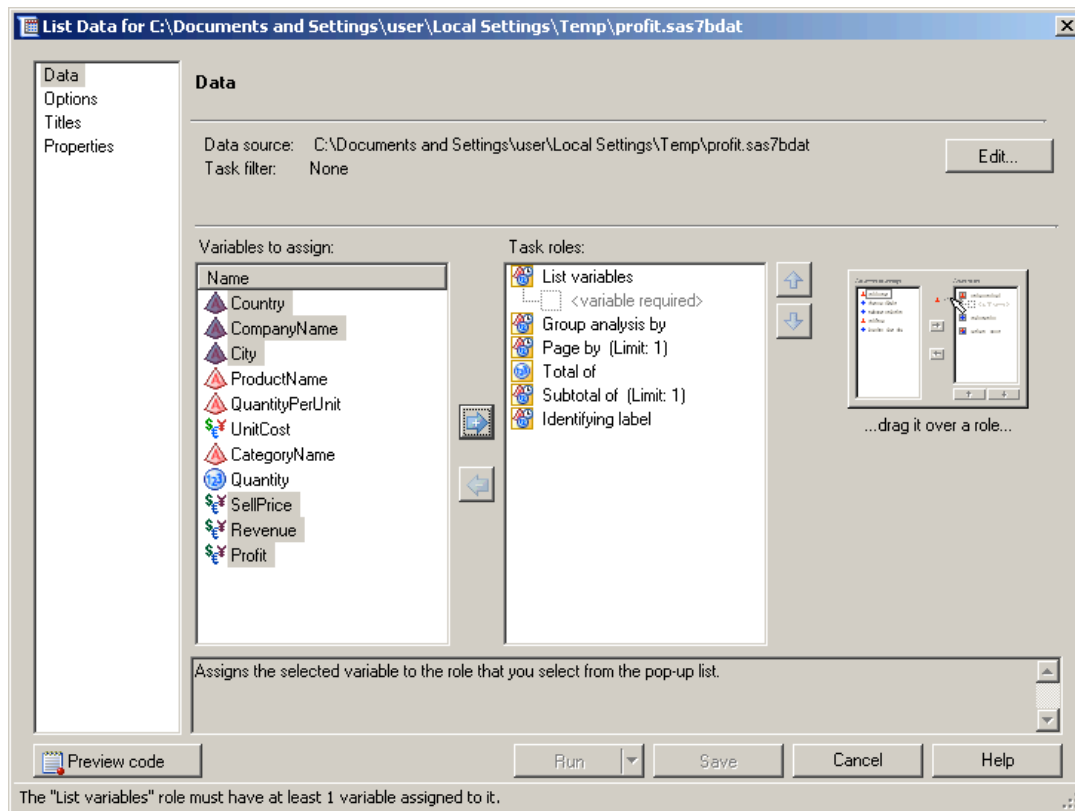


Obrázok 10 Zobrazenie naimportovanej databázy

Zdroj: softvér SAS

Po odkliknutí sa nám zobrazí dialógové okno, ktoré obsahuje niekoľko dôležitých položiek:

Zobrazenie údajov



Obrázok 11 Zobrazenie položky Data
Zdroj: softvér SAS

Data – identifikovanie údajov, ktoré chceme zobrazit'. V položke *Data* sa nachádzajú dve okná *Variable to assign* a *Task roles*.

V okne *Variables to assign* vidíme všetky premenné, ktoré sa nachádzajú v databáze. Z uvedeného okna presúvame premenné podľa potreby do druhého okna *Task roles*. Toto presúvanie nazývame **priradovanie úloh premenným**.

V okne *Task roles* sa nachádza niekoľko položiek:

Zoznam premenných (*List Variables*)

Zobrazí na obrazovke premenné v poradí, v akom sú uvedené. Musíme priradiť aspoň jednu premennú k tejto úlohe.

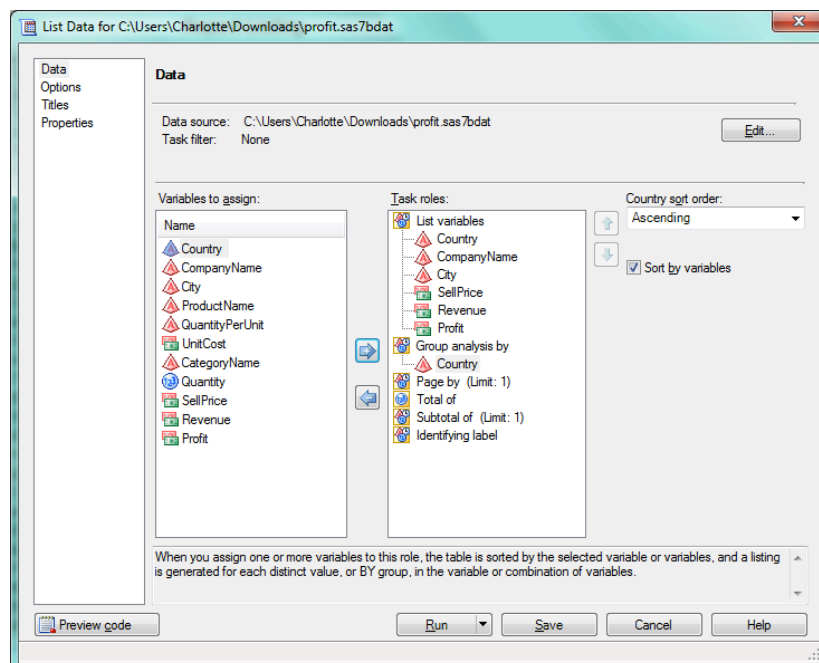
Zobrazenie údajov

Row number	Country	CompanyName	City	Sell Price	Revenue	Profit
1	UK	Exotic	London	\$22.50	\$1,012.50	\$202.50
2	UK	Exotic	London	\$22.50	\$405.00	\$81.00
3	UK	Exotic	London	\$22.50	\$450.00	\$90.00
4	UK	Exotic	London	\$22.50	\$337.50	\$67.50
5	UK	Exotic	London	\$22.50	\$270.00	\$54.00
6	UK	Exotic	London	\$22.50	\$337.50	\$67.50
7	UK	Exotic	London	\$22.50	\$225.00	\$45.00
8	UK	Exotic	London	\$22.50	\$540.00	\$108.00
9	UK	Exotic	London	\$23.80	\$476.00	\$96.00
10	UK	Exotic	London	\$23.80	\$1,190.00	\$240.00
11	UK	Exotic	London	\$23.80	\$633.00	\$126.00
12	UK	Exotic	London	\$23.80	\$992.00	\$192.00
13	UK	Exotic	London	\$23.80	\$595.00	\$120.00
14	UK	Exotic	London	\$23.80	\$166.60	\$33.60
15	UK	Exotic	London	\$23.80	\$571.20	\$115.20
16	UK	Exotic	London	\$23.80	\$595.00	\$120.00
17	UK	Exotic	London	\$23.80	\$1,428.00	\$288.00
18	UK	Exotic	London	\$23.80	\$238.00	\$48.00
19	UK	Exotic	London	\$11.20	\$336.00	\$36.00
20	UK	Exotic	London	\$11.20	\$560.00	\$60.00
21	USA	Cajun	New Orleans	\$24.60	\$492.00	\$52.00
22	USA	Cajun	New Orleans	\$24.60	\$590.40	\$62.40
23	USA	Cajun	New Orleans	\$24.60	\$442.80	\$46.80
24	USA	Cajun	New Orleans	\$24.60	\$246.00	\$26.00
25	USA	Cajun	New Orleans	\$24.60	\$661.00	\$31.00
26	USA	Cajun	New Orleans	\$23.90	\$1,553.50	\$165.75
27	USA	Cajun	New Orleans	\$23.90	\$286.80	\$30.60
28	USA	Cajun	New Orleans	\$23.90	\$478.00	\$51.00
29	USA	Cajun	New Orleans	\$23.90	\$764.80	\$81.60
30	USA	Cajun	New Orleans	\$23.60	\$354.00	\$38.25

Obrázok 12 Zoznam premenných vypísaných na obrazovke
Zdroj: softvér SAS

Analýza podľa skupín (*Group Analysis by*)

Po priradení jednej alebo viacerých premenných k tejto úlohe je tabuľka zoradená podľa vybranej premennej alebo premenných a výpis je generovaný pre každú jednotlivú hodnotu, alebo podľa skupiny, do premennej alebo kombinácií premenných.



Obrázok 13 Priradovanie úloh premenným podľa skupín
Zdroj: softvér SAS