

Статистички софтвер 4

Други час

Марија Радичевић

Математички факултет, Београд

2015.

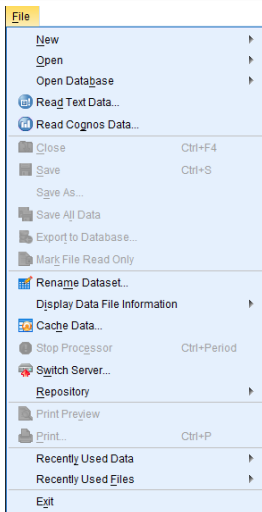
Садржај

- 1 Главне команде
 - File
 - Edit
 - View
 - Data
 - Transform
- 2 Команде трансформације SPSS командног језика
- 3 Преузимање података из других програмских окружења
- 4 Примери

Главни мени у прозору едитора података

- 1 File
- 2 Edit
- 3 View
- 4 Data
- 5 Transform
- 6 Analyze
- 7 Direct marketing
- 8 Graphics
- 9 Utilities
- 10 Add-ons
- 11 Window
- 12 Help

File



File омогућава рад са датотекама података.

Open database учитавање података из система за рад са базама података

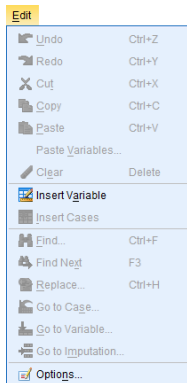
Read Text data учитавање података из текстуалне датотеке

Export to Database пренос података из SPSS датотеке у изабрану базу података

Display Data File Info приказ података о посебно наведеној датотеци података

- Working File
- External File

Edit



Edit омогућава премештање, копирање, преношење текста и дијаграма из активног прозора у програм за привременено чување садржаја, као и преузимање садржаја из њега, брисање, претраживање, итд.

Insert Variable отвара се нова колона за унос нове променљиве

Insert Cases убацује се нови ред

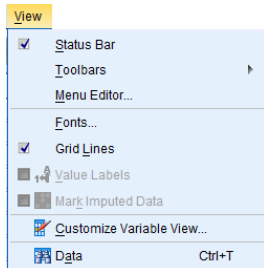
Go to Variable прелази се на колону (променљиву) са наведеним редним бројем

Go to Case прелази се на ред (опсервацију) са наведеним редним бројем

Go to Imputation активна команда само у случају рада са недостајућим подацима

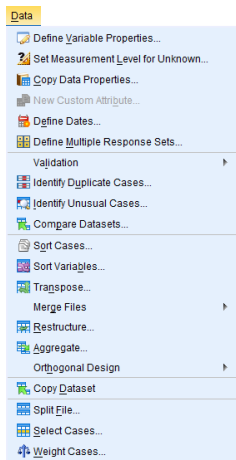
Options подешавања SPSS програмског окружења

View



View се користи за подешавање радног окружења.

Data



Data садржи скуп команди и њихов подсистем које пружају разноврсне могућности за рад са подацима.

Define Variable Properties отвара се комуникациони прозор за одређивање атрибута променљиве

Copy Data Properties копирање свих атрибута једне променљиве на другу променљиву

Define Dates генерисање датумске променљиве

Define Multiple Response Sets формира се низ за вишеструке одговоре, тј. више променљивих

Validation контрола исправности података у складу са постављеним правилима

- 1 претходно дефинисана правила
- 2 дефинисање правила
 - *Single – Variable Rules*
 - *Cross – Variable Rules*
- 3 спровођење контроле

Identify Duplicate Cases проверава за задату променљиву да ли постоје опсервације са истом вредношћу

Identify Unusual Cases открива "необичне" опсервације које се значајно разликују у односу на друге опсервације

Sort Cases сортира редове (опсервације) активне датотеке података према вредностима изабране једне или више променљивих

Sort Variables сортирање колона (променљивих) и то по било ком атрибуту променљивих

Transpose креирање новог едитора података са транспонованим редовима и колонама

Restructure формирање нове структуре података на бази већ постојеће табеле података

- превођење изабране опсервације у преоменљиве
- превођење изабране променљиве у опсервације
- транспоновање свих података

Merge Files удруживање података из две датотеке: активне датотеке података и датотеке података сачуване у **SPSS** формату

- удруживање датотека података које садрже исту променљиву а различите опсервације
- удруживање датотека података које садрже исте опсервације а различите променљиве

Aggregate формирање нове датотеке агрегираних података тако што се врши агрегирање свих опсервација изабраних променљивих у активној датотеци података, а према вредности променљиве која је изабрана за променљиву прелома

Split File подела активне датотеке података на више делова

Select Cases издвајање опсервација активне датотеке података које ће се даље процесуирати

- дефинисање логичког услова
- избор случајног узорка
- одређена област опсервација
- филтер променљива

Weight Cases додељује пондере опсервацијама, да би тако пондерисане вредности опсервација ушле у следећи поступак обраде и анализе

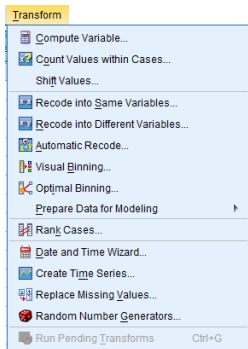
Transform

Transform садржи систем команди које пружају различите могућности за трансформисање променљивих и података.

Compute рачуна вредности нове или већ постојеће променљиве применом нумеричких израза у којима се могу појавити променљиве из активне датотеке података, неке од уграђених функција, аритметички и логички оператори, као и константе

Count Values Within Cases формирање нове променљиве у активној датотеци података која броји појављивање одређене вредности за изабране променљиве посебно за сваку опсервацију

Shift Values формирање нове променљиве која садржи вредности већ постојеће променљиве, али померене у односу на првобитну променљиву



Recode into Same Variables промена вредности (прекодирање) изабране(их) променљиве(их) према наведеној спецификацији

Record into Different Variables слично као претходна команда, с том разликом што се вредности изабране променљиве прекодирају и додељују новој променљивој

Automatic Recode активирање процедуре којом се за изабрану променљиву (променљиве) формира нова променљива (променљиве) чије су вредности цели бројеви

Visual Binning формирање категоричке променљиве на бази изабране нумеричке променљиве

Optimal Binning подела (једне или више) непрекидне променљиве на групу

Prepere Data for Modeling извршавање алгоритма аутоматизоване припреме података

- Interactive
- Automatic
- Backtransform Score

Rank Cases активирање процедуре која за наведену(е) нумеричку(е) променљиву(е) формира нову променљиву која садржи ранг, нормалне скорове или неке друге ранг величине

- Rank
- Savage score
- Fractional rank
- Sum of case weights
- Ntiles $g_i = \text{truncate}(1 + \frac{R^n}{W+1})$
- Proportional estimates
- Normal Score

Data and Time Wizard спровођење операција које се најчешће користе у раду са датумским и временским величинама

Create Time Series покретање процедуре за креирање нових променљивих од изабраних променљивих применом неке функције временске серије

Replace Missing Values формирање нове променљиве (променљивих) код које је извршена замена недостајућих података за изабрану променљиву (променљиве)

Random Number Generators избор једног од два понуђена генератора случајних бројева

Садржај

- 1 Главне команде
 - File
 - Edit
 - View
 - Data
 - Transform
- 2 Команде трансформације SPSS командног језика
- 3 Преузимање података из других програмских окружења
- 4 Примери

Команде трансформације

IF: команда за условну трансформацију

IF [логички израз] циљна променљива=израз.

DO IF-END IF: омогућава спровођење више трансформација

на групи опсервација

DO IF[логички израз].

команде трансформације.

ELSE IF [логички израз].

команде трансформације.

...

ELSE.

команде трансформације.

END IF.

LOOP-END LOOP: трансформације се понављају на истој

опсервацији док се не постигне услов за излазак

из структуре

LOOP.

команде трансформације.

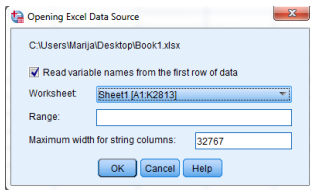
END LOOP.

Садржај

- 1 Главне команде
 - File
 - Edit
 - View
 - Data
 - Transform
- 2 Команде трансформације SPSS командног језика
- 3 Преузимање података из других програмских окружења
- 4 Примери

Преузимање података из:

- Excel: File⇒Open⇒Data



- Текстуални документ File⇒Read Text File
- програма за управљање базама података File⇒Open Database

Преузимање података из текстуалног документа

The screenshots illustrate the following steps in the Text Import Wizard:

- Step 1 of 6:** Selecting the file to import and confirming the encoding (UTF-8).
- Step 2 of 6:** Choosing how to parse the data (Delimited or Fixed width) and which variables to import.
- Step 3 of 6:** Selecting the starting row and the number of records to import.
- Step 4 of 6:** Choosing which delimiters to use.
- Step 5 of 6:** Specifying the date format.
- Step 6 of 6:** Reviewing the data preview and the final import options.

Садржај

- 1 Главне команде
 - File
 - Edit
 - View
 - Data
 - Transform
- 2 Команде трансформације SPSS командног језика
- 3 Преузимање података из других програмских окружења
- 4 Примери

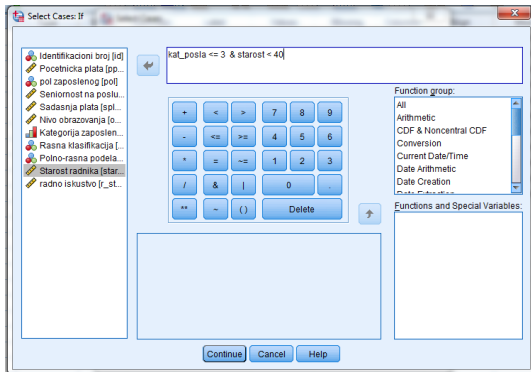
Пример 1

Изабрати банкарске службенике из прве три категорије посла који су млађи од 40 година.

Пример 1

Изабрати банкарске службенике из прве три категорије посла који су млађи од 40 година.

- *Data* ⇒ *Select Cases*.

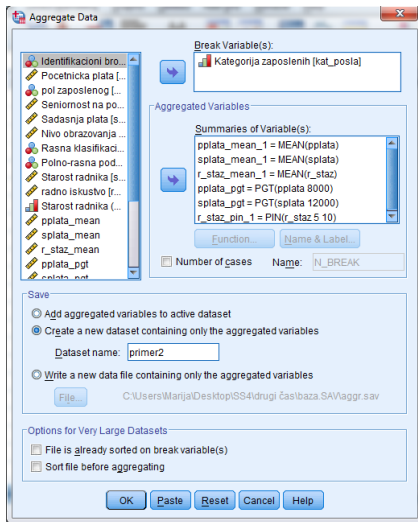


Пример 2

За сваку категорију посла запослених приказати следеће:

- Просечну вредност почетне плате
- Просечну вредност садашње плате
- Просечну вредност радног искуства
- Процент запослених чија је садашња плата већа од 12000
- Процент запослених чија је почетна плата већа од 8000
- Процент запослених са радним искуством између 5 и 10 година
- Број запослених у свакој категорији

Data ⇒ Aggregate



Пример 3

Колика је вероватноћа да случајна променљива која има $N(3, 1)$ расподелу узме вредност ≤ 4 .

Пример 3

Колика је вероватноћа да случајна променљива која има $N(3, 1)$ расподелу узме вредност ≤ 4 .

- *Transform* \Rightarrow *Compute Variables* и функција $CDF.NORMAL(4, 3, 1)$. Решење је 0.84.

Пример 4

Колика је вероватноћа да случајна променљива која има $U(0, 5)$ расподелу узме вредност > 2 .

Пример 4

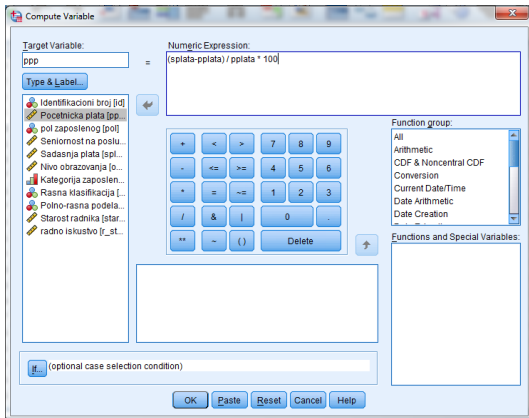
Колика је вероватноћа да случајна променљива која има $U(0, 5)$ расподелу узме вредност > 2 .

- *Transform* \Rightarrow *Compute Variables* и функција $1 - CDF.UNIFORM(2, 0, 5)$. Решење је 0.6.

Пример 5

У датотеци *baza.sav* формирати променљиву *ppp* (процент пораста плате) која показује колико је процената порасла садашња у односу на почетну плату за сваког запосленог. Затим формирати категоријску променљиву *kpp* (категорија пораста плате) којом се службеници групишу у 4 категорије у зависности од процента пораста плате (до 50%, од 50.01 – 100%, од 100.01 – 150%, преко 150%)

- ppp: Transform \Rightarrow Compute Variable



- kpp:Transform⇒Visual Binning

Visual Binning

Scanned Variable List: ppp

Name: Current Variable: ppp Binned Variable: kpp Label: ppp (Binned)

Minimum: 24,76 Nonmissing Values Maximum: 407,94

Enter interval cutpoints or click Make Cutpoints for automatic intervals. A cutpoint value of 10, for example, defines an interval starting above the previous interval and ending at 10.

Grid:	Value	Label
1	50,000	<= 50,00
2	100,000	50,01 - 100,00
3	150,000	100,01 - 150,00
4		HIGH 150,01+
5		

Cases Scanned: 474

Missing Values: 0

Copy Bins: From Another Variable, To Other Variables...

Upper Endpoints: Included (<=), Excluded (<)

Make Cutpoints, Make Labels, Reverse scale

OK, Paste, Reset, Cancel, Help

Хвала на пажњи!