

Статистички софтвер 4

Први час

Марија Радичевић

Математички факултет, Београд

2015.

Садржај

1 Увод

- Шта је IBM SPSS Statistics?
- Примене IBM SPSS Statistics
- Статистике укључене у софтвер

2 Програмско окружење

- Едитор података
- Прозор излазних резултата
- Синтаксни едитор
- Едитор дијаграма
- Прозор помоћи
- Оквир за дијалог

Шта је IBM SPSS Statistics?

- SPSS (Software Package for Social Studies) - статистички пакет за друштвене науке

Шта је IBM SPSS Statistics?

- SPSS (Software Package for Social Studies) - статистички пакет за друштвене науке
- сложени софтверски пакет за статистичку анализу

Шта је IBM SPSS Statistics?

- SPSS (Software Package for Social Studies) - статистички пакет за друштвене науке
- сложени софтверски пакет за статистичку анализу
- садржи низ процедура почев од планирања истраживања и прикупљања података, уношења података у програм, анализирања података, па све до прављења извештаја и визуелног представљања излазних резултата анализе

Примене IBM SPSS Statistics

- статистичка анализа података

Примене IBM SPSS Statistics

- статистичка анализа података
- истраживање тржишта

Примене IBM SPSS Statistics

- статистичка анализа података
- истраживање тржишта
- организација маркетинга

Примене IBM SPSS Statistics

- статистичка анализа података
- истраживање тржишта
- организација маркетинга
- истраживање образовног система

Примене IBM SPSS Statistics

- статистичка анализа података
- истраживање тржишта
- организација маркетинга
- истраживање образовног система
- анкетирање

Примене IBM SPSS Statistics

- статистичка анализа података
- истраживање тржишта
- организација маркетинга
- истраживање образовног система
- анкетирање
- обрада података
-

Статистике укључене у софтвер

- Дескриптивне статистике
- Дводимензионе статистике
- Статистике за предвиђање нумеричких исхода
- Статистике за предвиђање за идентификацију групе

Статистике укључене у софтвер

- Дескриптивне статистике
 - унакрсно табелирање (табеле контингенције)
 - *PP – plot*, *QQ – plot*
- Дводимензионе статистике
- Статистике за предвиђање нумеричких исхода
- Статистике за предвиђање за идентификацију групе

Статистике укључене у софтвер

- Дескриптивне статистике
 - унакрсно табелирање (табеле контингенције)
 - *PP – plot*, *QQ – plot*
- Дводимензионе статистике
 - *t*-тест
 - корелација
 - непараметарски тестови
 - *ANOVA*
- Статистике за предвиђање нумеричких исхода
- Статистике за предвиђање за идентификацију групе

Статистике укључене у софтвер

- Дескриптивне статистике
 - унакрсно табелирање (табеле контингенције)
 - *PP – plot*, *QQ – plot*
- Дводимензионе статистике
 - *t*-тест
 - корелација
 - непараметарски тестови
 - *ANOVA*
- Статистике за предвиђање нумеричких исхода
 - линеарна регресија
- Статистике за предвиђање за идентификацију групе

Статистике укључене у софтвер

- Дескриптивне статистике
 - унакрсно табелирање (табеле контингенције)
 - *PP – plot*, *QQ – plot*
- Дводимензионе статистике
 - *t*-тест
 - корелација
 - непараметарски тестови
 - *ANOVA*
- Статистике за предвиђање нумеричких исхода
 - линеарна регресија
- Статистике за предвиђање за идентификацију групе
 - факторска анализа
 - кластер анализа
 - дискриминанте

Садржај

1 Увод

- Шта је IBM SPSS Statistics?
- Примене IBM SPSS Statistics
- Статистике укључене у софтвер

2 Програмско окружење

- Едитор података
- Прозор излазних резултата
- Синтаксни едитор
- Едитор дијаграма
- Прозор помоћи
- Оквир за дијалог

Програмско окружење

- Прозор едитор података (Data Editor)

Програмско окружење

- Прозор едитор података (Data Editor)
- Прозор излазних резултата (Viewer)

Програмско окружење

- Прозор едитор података (Data Editor)
- Прозор излазних резултата (Viewer)
- Прозор синтаксни едитор (Syntax Window)

Програмско окружење

- Прозор едитор података (Data Editor)
- Прозор излазних резултата (Viewer)
- Прозор синтаксни едитор (Syntax Window)
- Прозор едитор дијаграма (Chart Editor)

Програмско окружење

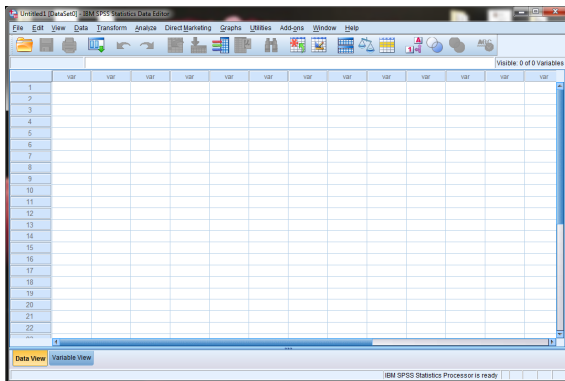
- Прозор едитор података (Data Editor)
- Прозор излазних резултата (Viewer)
- Прозор синтаксни едитор (Syntax Window)
- Прозор едитор дијаграма (Chart Editor)
- Прозор помоћи (Help window)

Програмско окружење

- Прозор едитор података (Data Editor)
- Прозор излазних резултата (Viewer)
- Прозор синтаксни едитор (Syntax Window)
- Прозор едитор дијаграма (Chart Editor)
- Прозор помоћи (Help window)
- Оквири за дијалог (Dialog box)

Облици приказа у едитору података

- Приказ података (Data View)
- Приказ променљивих (Variable View)



Поступци и процедуре у едитор података

- уношење података
- кориговање података
- ажурирање података
- брисање података
- дефинисање променљивих
- промена атрибута променљивих
- трансформација података
-

Уношење података

- Начини уношења:
 - 1 Директно уношење у едитору података
 - 2 Учитавање података из већ постојеће датотеке података (*File* $\Rightarrow$ *Open* $\Rightarrow$ *Data*)

Уношење података

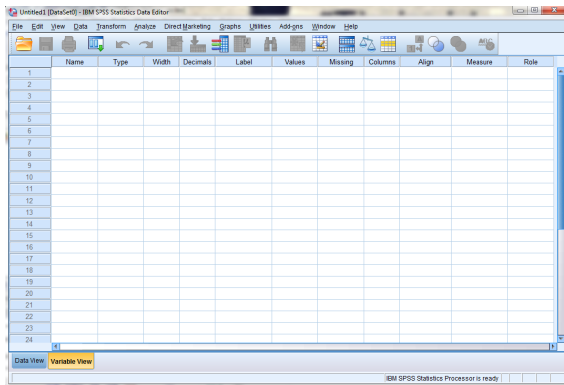
- Начини уношења:
 - ❶ Директно уношење у едитору података
 - ❷ Учитавање података из већ постојеће датотеке података (*File* $\Rightarrow$ *Open* $\Rightarrow$ *Data*)
- Подаци се уносе у ћелије табеле

Уношење података

- Начини уношења:
 - ❶ Директно уношење у едитору података
 - ❷ Учитавање података из већ постојеће датотеке података (*File* $\Rightarrow$ *Open* $\Rightarrow$ *Data*)
- Подаци се уносе у ћелије табеле
- Димензије табеле су одређене бројем променљивих и бројем опсервација

Променљиве

- Карактеристике (атрибути) променљивих дефинисани су у *Variable View*



Основни атрибути сваке променљиве:

- Име (*Name*)
- Тип (*Type*)
- Ширина (*Width*)
- Број децималних места (*Decimals*)
- Ознака (*Label*)
- Вредности (*Values*)
- Недостајући подаци (*Missing*)
- Ширина колоне у едитору података (*Columns*)
- Начин поравњања садржаја у ћелији (*Align*)
- Начин мерења променљивих (*Measure*)
- Улога променљиве (*Role*)

Име променљиве

- јединствено

Име променљиве

- јединствено
- највише 64 знака

Име променљиве

- јединствено
- највише 64 знака
- дозвољени знаци:
 - a* сва слова
 - 1* цифре
 - .* тачка
 - _* знак за подвучено
 - zпaси* \$, #

Име променљиве

- јединствено
- највише 64 знака
- дозвољени знаци:
 - a* сва слова
 - 1* цифре
 - .* тачка
 - _* знак за подвучено
 - znaci* \$, #
- мора почети словом

Име променљиве

- јединствено
- највише 64 знака
- дозвољени знаци:
 - a* сва слова
 - 1* цифре
 - .* тачка
 - _* знак за подвучено
 - zpači* \$, #
- мора почети словом
- не сме да садржи знак за бланко, као и специјалне знаке
!, ?, " , ' , *

Име променљиве

- јединствено
- највише 64 знака
- дозвољени знаци:
 - a* сва слова
 - 1* цифре
 - .* тачка
 - _* знак за подвучено
 - znaci* \$, #
- мора почети словом
- не сме да садржи знак за бланко, као и специјалне знаке
!, ?, ", ', *
- не би требало да садржи тачку или знак за подвучено
као последњи карактер

Име променљиве

- јединствено
- највише 64 знака
- дозвољени знаци:
 - a* сва слова
 - 1* цифре
 - .* тачка
 - _* знак за подвучено
 - znaci* \$, #
- мора почети словом
- не сме да садржи знак за бланко, као и специјалне знаке
!, ?, " , ' , *
- не би требало да садржи тачку или знак за подвучено
као последњи карактер
- не сме да садржи резервисане речи
(ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH)

Име променљиве

- јединствено
- највише 64 знака
- дозвољени знаци:
 - a* сва слова
 - 1* цифре
 - .* тачка
 - _* знак за подвучено
 - znaci* \$, #
- мора почети словом
- не сме да садржи знак за бланко, као и специјалне знаке
!, ?, " , ' , *
- не би требало да садржи тачку или знак за подвучено
као последњи карактер
- не сме да садржи резервисане речи
(ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH)
- програм није осетљив на велика и мала слова

Тип променљиве

- ❶ нумерички тип (*Numeric*)
- ❷ тип зареза (*Comma*)
- ❸ тип тачка (*Dot*)
- ❹ научна нотација (*Scientific notation*)
- ❺ датумски тип (*Data*)
- ❻ тип долара (*Dollar*)
- ❼ наменски тип (*Custom Currency*)
- ❽ алфанумерички тип (*String*)
- ❾ цео број са водећим нулама (*Restricted numeric*)

Напомена

Подразумевани тип је нумерички са укупно 8 карактера (6 испред и 2 иза децималне тачке).

Карактеристике појединих типова променљивих

1. нумерички тип

- све цифре, водећи знак (+ или −) и индикатор за децимално место
- максимално 40 карактера (16 децималног дела)

Карактеристике појединих типова променљивих

1. нумерички тип

- све цифре, водећи знак (+ или −) и индикатор за децимално место
- максимално 40 карактера (16 децималног дела)

2. тип зареза

- исто као код нумеричког, с тим да је укључен зарез за одвајање хиљада

Карактеристике појединих типова променљивих

1. нумерички тип

- све цифре, водећи знак (+ или −) и индикатор за децимално место
- максимално 40 карактера (16 децималног дела)

2. тип зареза

- исто као код нумеричког, с тим да је укључен зарез за одвајање хиљада

3. тип тачка

- све цифре, водећи знак (+ или −), зарез као индикатор одвајања децималних места и тачка за одвајање хиљада

Карактеристике појединих типова променљивих

1. нумерички тип

- све цифре, водећи знак (+ или −) и индикатор за децимално место
- максимално 40 карактера (16 децималног дела)

2. тип зареза

- исто као код нумеричког, с тим да је укључен зарез за одвајање хиљада

3. тип тачка

- све цифре, водећи знак (+ или −), зарез као индикатор одвајања децималних места и тачка за одвајање хиљада

4. научна нотација

- цифре, +, −, E, D

5. датумски тип

- дозвољене су вредности за датум и време

5. датумски тип

- дозвољене су вредности за датум и време

6. тип долара

- нумеричке променљиве код којих је \$ водећи знак
- зарез се користи као индикатор за раздвајање хиљада
- тачка је децимални сепаратор

5. датумски тип

- дозвољене су вредности за датум и време

6. тип долара

- нумеричке променљиве код којих је \$ водећи знак
- зарез се користи као индикатор за раздвајање хиљада
- тачка је децимални сепаратор

7. наменски тип

- нумеричке променљиве
- мора се претходно креирати у *Edit* ⇒ *Options* ⇒ *Currency*

5. датумски тип

- дозвољене су вредности за датум и време

6. тип долара

- нумеричке променљиве код којих је \$ водећи знак
- зарез се користи као индикатор за раздвајање хиљада
- тачка је децимални сепаратор

7. наменски тип

- нумеричке променљиве
- мора се претходно креирати у *Edit* ⇒ *Options* ⇒ *Currency*

8. алфанумерички (string) тип

- било који дозвољени знак до задате дужине
- не може се користити у операцијама нумеричког рачунања

5. датумски тип

- дозвољене су вредности за датум и време

6. тип долара

- нумеричке променљиве код којих је \$ водећи знак
- зарез се користи као индикатор за раздвајање хиљада
- тачка је децимални сепаратор

7. наменски тип

- нумеричке променљиве
- мора се претходно креирати у *Edit* ⇒ *Options* ⇒ *Currency*

8. алфанумерички (string) тип

- било који дозвољени знак до задате дужине
- не може се користити у операцијама нумеричког рачунања

9. цео број са водећим нулама

- слично као нумерички тип, само без разломљеног дела и са водећим нулама

Ширина променљиве

- ширина нумеричке променљиве је 8 места, од којих су два за децимални део броја

Ширина променљиве

- ширина нумеричке променљиве је 8 места, од којих су два за децимални део броја
- подразумеване вредности се мењају у картици *Data*, опције *Edit* $\Rightarrow$ *Options*

Ширина променљиве

- ширина нумеричке променљиве је 8 места, од којих су два за децимални део броја
- подразумеване вредности се мењају у картици *Data*, опције *Edit* $\Rightarrow$ *Options*
- ширина алфанумеричке променљиве је одређена ширином првог унетог податка, али се може променити

Ознака променљиве

- детаљно описује садржај променљиве

Ознака променљиве

- детаљно описује садржај променљиве
- садржи до 256 карактера, при чему су дозвољени и бланко карактери, као и резервисане речи

Вредност променљиве

- свакој вредности променљиве може се доделити одговарајући опис дужине 120 карактера

Вредност променљиве

- свакој вредности променљиве може се доделити одговарајући опис дужине 120 карактера
- значајан за категоричке променљиве

Недостајуће вредности

- 1 системски недостајуће вредности
(*system – missing values*)
- 2 кориснички дефинисане недостајуће вредности
(*user – missing values*)

Недостајуће вредности

- ❶ системски недостајуће вредности
(*system – missing values*)
 - за нумаричке променљиве - тачке у ћелијама
 - за алфанумеричке променљиве - ћелије остају празне
- ❷ кориснички дефинисане недостајуће вредности
(*user – missing values*)

Недостајуће вредности

- ❶ системски недостајуће вредности
(*system – missing values*)
 - за нумаричке променљиве - тачке у ћелијама
 - за алфанумеричке променљиве - ћелије остају празне
- ❷ кориснички дефинисане недостајуће вредности
(*user – missing values*)
 - нема недостајућих вредности
 - 3 дискретне вредности за ознаку недостајућих вредности
 - област, или област и једна дискретна вредност, за ознаку недостајућих вредности

Ширина колоне и начин приказивања садржаја ћелије

- стандардна ширина колоне је одређена изабраном ширином променљиве, али се може променити

Ширина колоне и начин приказивања садржаја ћелије

- стандардна ширина колоне је одређена изабраном ширином променљиве, али се може променити
- поравњање може бити уз леву ивицу ћелије, уз десну ивицу ћелије или центрирано
 - за нумеричке променљиве подразумевано поравњање је уз леву ивицу ћелије
 - за алфанумеричке променљиву подразумевано поравњање уз десну ивицу ћелије

Начин мерења

1 нумеричка скала

Начин мерења

- 1 нумеричка скала
- 2 ординална скала

Начин мерења

- 1 нумеричка скала
- 2 ординална скала
- 3 номинална скала

Улога променљиве

- предиктор (независна) променљива (*input*)
- циљна (зависна) променљива (*output*)
- обе улоге (*input* и *output*)
- није дефинисана улога (*none*)
- променљива за поделу података на узорке за формирање модела, тестирање и оцену (*partition*)
- променљива за поделу код *PASW Modeler* — *a* (*split*)

Улога променљиве

- предиктор (независна) променљива (*input*)
- циљна (зависна) променљива (*output*)
- обе улоге (*input* и *output*)
- није дефинисана улога (*none*)
- променљива за поделу података на узорке за формирање модела, тестирање и оцену (*partition*)
- променљива за поделу код *PASW Modeler* – *a* (*split*)

Напомена

Подразумевана вредност за сваку променљиву је *input* променљива.

Едитор података

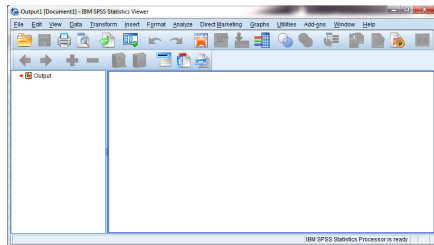
- Датотека се чува са екстензијом **.sav**

Напомена

Све атрибуте променљивих програм може аутоматски да додели (подразумеване вредности).

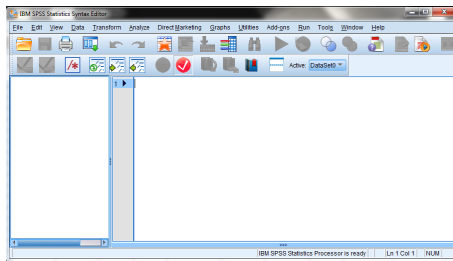
Прозор излазних резултата

- приказује синтаксу команди процедуре и резултате процедуре
- са леве стране структурирано стабло са списком свих обављених процедура
- са десне стране конкретни резултати
- чува се са екстензијом *.spv*



Синтаксни едитор

- команде језика за извршавање процедура
- додатне могућности програма
- покретање: *File* $\Rightarrow$ *New* $\Rightarrow$ *Syntax* или
File $\Rightarrow$ *Open* $\Rightarrow$ *Syntax*
- чува се са екстензијом *.sps*



Едитор дијаграма

- отвара се двоструким кликом на дијаграм
- служи за промену врсте графичког приказивања, изглед и формат дијаграма, текст, фонтови, текстура позадине, итд.

Напомена

Слично се ради са динамичким табелама, које се добијају процедуром *Analyze* $\Rightarrow$ *Reports* $\Rightarrow$ *OLAP Cubes*, а двоструким кликом се прелази у прозор Пивот табеле. Овај прозор омогућава динамички рад са табелама, тј. промену форме табеле, редова, колона, транспоновање редова и колона, итд.

Прозор помоћи

- отвара се активирањем команде *Help* из главног менија
- сваки комуникациони прозор има свој прозор помоћи (*context sensitive help*)

Оквир за дијалог

- привремени прозори који се појављују током процесуирања неке од **SPSS** команди
- у њих се уносе неопходни елементи везани за изабрану статистичку процедуру
- од расположивих променљивих из едитора података бирају се променљиве које учествују у статистичкој процедури и преносе се у један или више празних поља
- дугмад у комуникационим прозорима:

Ok прихватају се све наведене опције и позива се њихово извршавање

Paste све изабране опције замењују се одговарајућом синтаксом **SPSS** команди која се преноси у синтаксни едитор

Reset поништавају се све изабране опције и омогућава се уношење нових

Cancel одустаје се од извршавања опција активног прозора и враћа се на претходни прозор

Help добија се информација о процедури

Хвала на пажњи!