

Основи на SPSS

Уважаеми колеги,

В този текст са представени анотациите на темите по “Основи на SPSS”.

Всяка тема е представена заедно със списък на основната и допълнителната литература. За някои теми няма достатъчно информация в посочената основна и допълнителна литература, затова в края на текста са дадени допълнителни текстове по тях.

Нека преди да започнем с въведението в SPSS да си припомним етапите на провеждане на емпиричното социологическо изследване, с които се запознахте в курса по „Статистически методи в социологията”:

- Първият етап е предварителната подготовка. На този етап се:
 - ◆ Формулират целта и задачите на изследването.
 - ◆ Дефинират съвкупността, съставлящите я единици и изучаваните признаци.
 - ◆ Изготвя формата за регистрация на наблюдаваните значения на признаците.
 - ◆ Провежда се пробно изследване.
- Вторият етап е събирането на първичната информация.
- Третият етап е обработката и анализът на данните. Той включва:
 - ◆ Прехвърлянето на информацията от регистрационната форма на общ (електронен) носител.
 - ◆ Контрол за грешки в първичната информация.
 - ◆ Статистически анализ на данните.

Тези три етапа изчерпват статистическата част на изследването. След тях обаче следва съдържателната интерпретация на получените резултати, която се прави от специалистите по социология.

Работата със SPSS започва след като първичната информация вече е събрана. Основното предназначение на програмата е да автоматизира дейностите, извършвани на третия етап на изследването.

1. Запознаване с SPSS

SPSS е абривиатура на "Statistical package for social science". По същество това е статистически софтуер, насочен към потребители, занимаващи се със социални науки. Като функционални възможности той не отстъпва, но и не превъзхожда другите статистически програми, използвани за анализ на данни, като Statistica, Statgraphics, EViews и др. Голямото предимство на SPSS е в това, че е лесен за употреба, с много добре организиран вход на първичната информация и че готовите резултати се получават във вид, които е ясен и разбираем за всеки нестатистик. За да можете да работите със SPSS са ви необходими четири неща:

- Елементарна компютърна грамотност, за да се справяте с компютъра.
- Добро познание на английски език, за да разбирате позициите от менюто.
- Знания по статистика, за да знаете коя позиция от менюто трябва да изберете.
- Социологическа компетентност, за да можете да интерпретирате получените резултати.

2. Изготвяне на макет на таблица

Първичната информация от емпиричните социологически изследвания обикновено е във вид на анкетни карти. За да може тази информация да се използва, тя трябва да се вкара в компютъра. Вкарването е организирано в специална таблица, която трябва да се подготви предварително. Тази предварителна подготовка се нарича изготвяне на макет на таблицата.

От тази тема ще научите:

- Как се описват анкетните въпроси, които предполагат само един отговор.
- Какви са вариантите за описване на анкетните въпроси, които предполагат повече от един отговор.

Основна литература: [1, стр. 9-15, 20-24 и 29-34]

Допълнителна литература: [2, стр. 33-36 и 38-40]

3. Вкарване на данни

След като макетът на таблицата е готов, можем да започнем вкарването на данните от анкетните карти.

От тази тема ще научите:

- Как се вкарват отговорите на анкетните въпроси, които предполагат само един отговор
- Как се вкарват отговорите на анкетните въпроси, които предполагат повече от един отговор
- Как се вкарват в SPSS данни, вече вкарани с помощта на друг софтуер

Основна литература: [1, стр. 34-35 и 36]

Допълнителна литература: [2, стр. 36-38 и 43-45]

4. Откриване и отстраняване на грешки в първичната информация

След вкарването на данните, **задължително** трябва да се направи контрол за откриване на евентуални грешки в първичната информация.

От тази тема ще научите:

- Как се откриват грешки, допуснати от операторите при вкарването на данните
- Как се идентифицират единиците, при които са допуснати тези грешки
- Как се отстраняват грешките, допуснати от операторите при вкарването на данните
- Как се откриват противоречия в отговорите на анкетираните лица
- Как се идентифицират единиците, при които са допуснати тези противоречия
- Как се отстраняват противоречията в отговорите на анкетираните лица
- Как SPSS може да ни помогне да уловим недобросъвестни анкетьори

Основна литература: [1, стр. 35-36]

5. Едномерни разпределения

Анализът на информацията от емпирични социологически изследвания започва с анализ на едномерни разпределения.

От тази тема ще научите:

- Как се прави едномерно разпределение по отговорите на анкетни въпроси, които предполагат само един отговор
- Как се прави едномерно разпределение по отговорите на анкетни въпроси, които предполагат повече от един отговор
- Как се прави едномерно разпределение по количествен признак (интервална групировка)
- Как се интерпретират получените едномерни разпределения
- Как се описват разпределенията по количествен признак:
 - ◆ От гледна точка на съществени техни черти.
 - ◆ От гледна точка на местоположението на конкретните единици.
- Как се интерпретират получените числови характеристики
- Как едномерните разпределения се илюстрират графично

Основна литература: [1, стр. 41-42, 44 и 47-87]

Допълнителна литература: [2, стр. 49-62, 65-68 и 77-82]

6. Статистическо оценяване

Когато данните са получени от представителна извадка (което е най-честият случай в практиката на емпиричните социологически изследвания), трябва да се построят доверителните интервали на относителните дялове и/или на средните аритметични. В SPSS не е предвидена възможността да се построят доверителни интервали на относителни дялове. Също така построяването на доверителни интервали на средната аритметична не е напълно автоматизирано – необходимо е част от работата да бъде свършена лично.

От тази тема ще научите:

- Как се получават стойностите на средната аритметична и на стандартната грешка на средната аритметична
- Как ръчно се построява доверителният интервал на средната аритметична
- Как се интерпретира построенят доверителен интервал

- Как построеният доверителен интервал се илюстрира графично
Основна литература: [1, стр. 42-43 и 44-47]
Допълнителна литература: [2, стр. 84-91, 95-103, 104-106 и 144-150]

7. Проверка на статистически хипотези (тестове за значимост)

В практиката на емпиричните социологически изследвания, когато информацията е получена от представителни извадки, често възниква необходимост да се проверяват хипотези относно вида или основни числови характеристики на едномерните разпределения. В SPSS не е предвидена възможността да се проверяват хипотези относно относителни дялове, а само относно средни аритметични и относно вида на разпределението на единиците.

От тази тема ще научите:

- Как се прави проверка за равенство на средна аритметична с дадено число
- Как се прави проверка за равенство на две средни аритметични
 - ◆ При независими извадки
 - ◆ При зависими извадки
- Как проверката за равенство на две средни аритметични се илюстрира графично
- Как се прави проверка дали емпиричното разпределение съвпада с някое от познатите теоретични разпределения
- Как се прави проверка дали две емпирични разпределения в две различни генерални съвкупности са еднакви
- Как сравняването на две разпределения се илюстрира графично
- Как се прилагат непараметричните методи за проверка на хипотези
 - ◆ При независими извадки
 - ◆ При зависими извадки

Основна литература: [1, стр. 90-108]

Допълнителна литература: [2, стр. 154-164, 178-189, 212-218, 432-435, 436 и 437-440]

8. Двумерни разпределения (кростаблици)

Едномерните разпределения са полезен, но недостатъчен компонент на социологическия анализ. Това е така, защото те дават отговор на въпросите "Какво?" и "Колко?", но по никакъв начин не могат да отговорят на въпросите "Как?" и "Защо?". Иначе казано, освен да описваме изучаваната съвкупност признак по признак, трябва да можем да откриваме и връзките между признаците. В основата на анализа на връзки стоят двумерните разпределения.

От тази тема ще научите:

- Как се прави двумерно разпределение
- Как се получават относителни дялове по редове, по колони и общо
- Как се интерпретират относителните дялове по редове и по колони от гледна точка на връзката между два признака

Основна литература: [1, стр. 111-120]

Допълнителна литература: [2, стр. 286-296]

9. Анализ на връзки при качествен фактор и качествен резултат

При анализа на връзки между два качествени признака на базата на информация от представителна извадка, трябва да отговорим на въпроса доколко наблюдаваното различие между относителните дялове по редове или по колони в извадката е показателно за различието в генералната съвкупност. Това означава, че е необходим статистически тест за установяване на наличие на връзка. След като се установи връзка, можем да преминем към измерване на посоката и силата ѝ.

От тази тема ще научите:

- Как се получава стойността на Хи-квадрат
- Как от стойността на Хи-квадрат се разбира дали има или няма връзка между признаците
- Как се изследва посоката на връзката
- Как се получават стойностите на Фи и на коефициента на Крамер
- Как от стойностите на Фи и на коефициента на Крамер се разбира каква е силата на връзката между признаците
- Как връзките между два качествени признака се илюстрират графично

Основна литература: [1, стр. 125-133]

Допълнителна литература: [2, стр. 298-308 и 311-312]

10. Анализ на връзки при качествено фактор и количествен резултат

Когато резултатът е количествен признак, можем вместо относителните дялове да използваме средната аритметична. Различията между средните стойности на резултата във всяка група на фактора ще ни покаже дали има или няма връзка. След като установим, че има връзка можем да изследваме посоката и силата ѝ.

От тази тема ще научите:

- Как се получават средните стойности на признака резултат във всяка група на признака фактор
- Как се разбира дали има или няма връзка между признаците при изчерпателни изследвания
- Как се разбира дали има или няма връзка между признаците при извадкови изследвания (дисперсионен анализ)
- Как се изследва посоката на връзката
- Как се получава стойността на коефициента на определеност
- Как се интерпретира коефициентът на определеност
- Как се проверява дали са изпълнени условията за приложение на дисперсионния анализ.
- Как се идентифицират тези двойки групи, формирани по значенията на признака фактор, между които има значимо различие в средните стойности на признака резултат
- Как връзките между качествено фактор и количествен признак резултат се илюстрират графично

Основна литература: [1, стр. 133-146]

Допълнителна литература: [2, стр. 224-234 и 333-334]

11. Анализ на връзки при количествен фактор и количествен резултат

Анализът на връзки между два количествени признака се различава от предходните два случая на анализ на връзки. Особеното е, че връзките винаги се изследват с помощта на даден математически модел.

От тази тема ще научите:

- Как се получават коефициентите на различните модели на връзката
- Как се интерпретират получените коефициенти
- Как получените коефициенти показват наличието и посоката на връзката
- Как се интерпретира коефициентът на определеност
- Как се проверява кои от моделите на връзката е адекватен
- Как се избира най-добрият модел от всички адекватни модели
- Как връзките между два количествени признака се илюстрират графично
- Как се прилага самостоятелно корелационният анализ
- Как се получават ранговите коефициенти на корелация
- Как се интерпретират обикновените и ранговите коефициенти на корелация

Основна литература: [1, стр. 146-162]

Допълнителна литература: [2, стр. 237-256 и 445-450]

12. Графично представяне на информацията

Програмата SPSS дава възможност за графична илюстрация на получените аналитични резултати.

От тази тема ще научите:

- Как се получават и за какви цели се използват:
 - ◆ Правоъгълна диаграма (Bar...)
 - ◆ Линейна диаграма (Line...)
 - ◆ Area...
 - ◆ Кръгова диаграма (Pie...)
 - ◆ "Кутия с мустаци" (Boxplot...)
 - ◆ Error bar...

- ◆ Точкова диаграма (Scater...)
- ◆ Хистограма (Histogram...)
- ◆ P-P...
- ◆ Q-Q...

13. Работа с Output

Резултатите в SPSS се получават в специална форма наречена Output. Екрана на Output-а е разделен на две части. В дясната част са самите обекти (таблицы и/или графики), а в лявата част са наименованията на тези обекти. SPSS позволява обектите в дясна страна на екрана да бъдат маркирани, премествани, копирани, отпечатвани и изтривани. Наименованията в лявата част на екрана служат за бърз достъп до съответните обекти, както и отново за маркиране, преместване, копиране, отпечатване и изтриване.

От тази тема ще научите:

- Как се маркират, преместват, копират, отпечатват и изтриват обектите в дясната част на Output-а
- Как се осъществява бърз достъп до обектите в дясната част на Output-а чрез наименованията им в лявата част
- Как се маркират, преместват, копират, отпечатват и изтриват обектите в дясната част на Output-а посредством наименованията им в лявата част

14. Трансфер на обекти между различни приложения

Наличието на Output е голямо улеснение в работата на социолога. Това е така, защото Output-ът може директно да бъде експортиран във формат Microsoft Word и след това остава само да се напише интерпретацията на получените резултати. От друга страна SPSS дава възможност да се прехвърлят файлове от и към Microsoft Excel. Това също е голямо улеснение, защото позволява данни, вкарани на SPSS да се използват от други приложения и обратно.

От тази тема ще научите:

- Как Output-ът или отделни негови части се експортира във формат Microsoft Word
- Как данните от SPSS се прехвърлят във формат Microsoft Excel

- Как данни от формат Microsoft Excel се прехвърлят в SPSS

* *
*
*
*

Литература:

1. Гоев, В., Статистическа обработка и анализ на информацията от социологически, маркетингови и политически изследвания със SPSS, С, 1996
2. Манов, А., Статистика със SPSS, С, 2001

ДОПЪЛНИТЕЛНИ ТЕКСТОВЕ

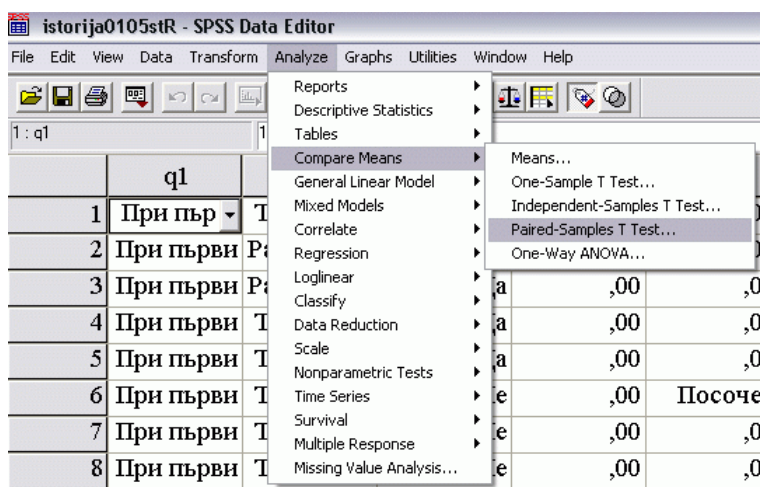
Към темите, за които няма достатъчно информация в посочената към тях основна и допълнителна литература

Технологично допълнение към тема №7

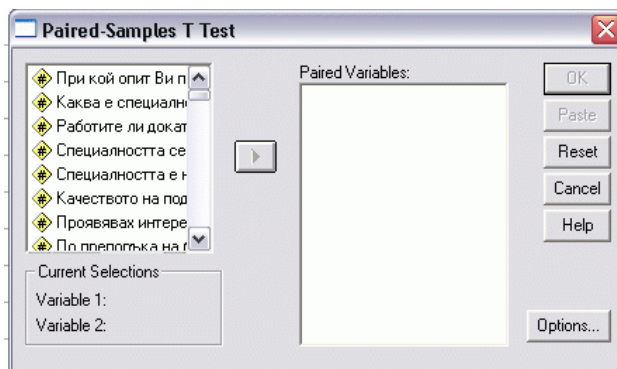
Проверка на статистически хипотези (тестове за значимост)

Проверката на хипотезата за равенство на две средни аритметични при зависими извадки преминава през следната последователност:

- От менюто се избира:
 - ◆ Analyze
 - Compare Means
 - ❖ Paired-Samples T Test...



Появява се следният прозорец:



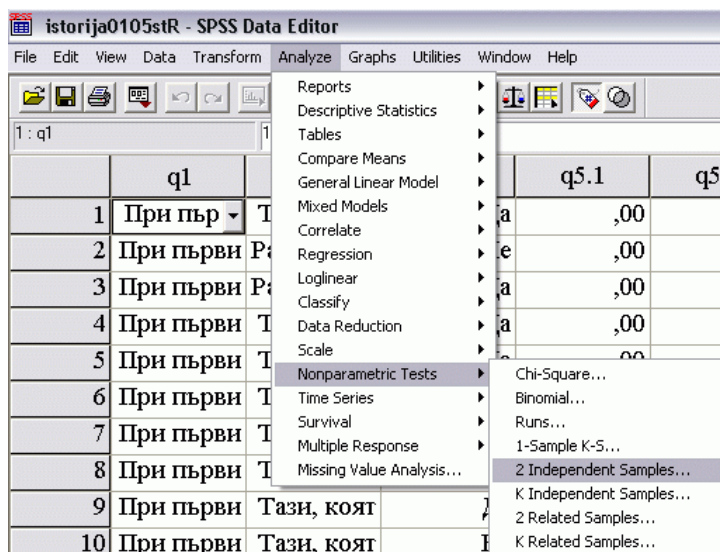
- Формират се двойки от съответните признаци, по един за всяка една от зависимите извадки и се натиска бутона ОК.

Проверката на хипотези за съгласуваност на две емпирични разпределения преминава през следната последователност:

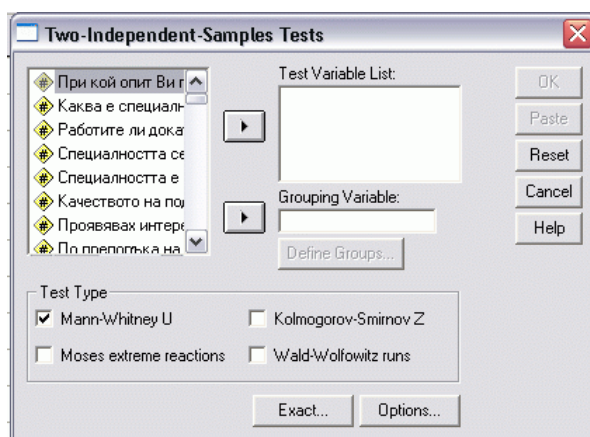
- От менюто се избира:
 - ◆ Analyze

➤ Nonparametric Tests

❖ 2 Independent Samples...

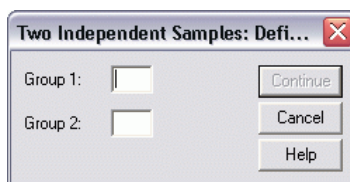


Появява се следният прозорец:



- В позицията Test Variable List се избира(т) признака (признаците), чиито разпределения ще се сравняват.
- В позицията Grouping Variable се избира признака, по който са формираните сравняваните подсъвкупности.
- Натиска се бутона Define Groups.

Появява се следният прозорец:

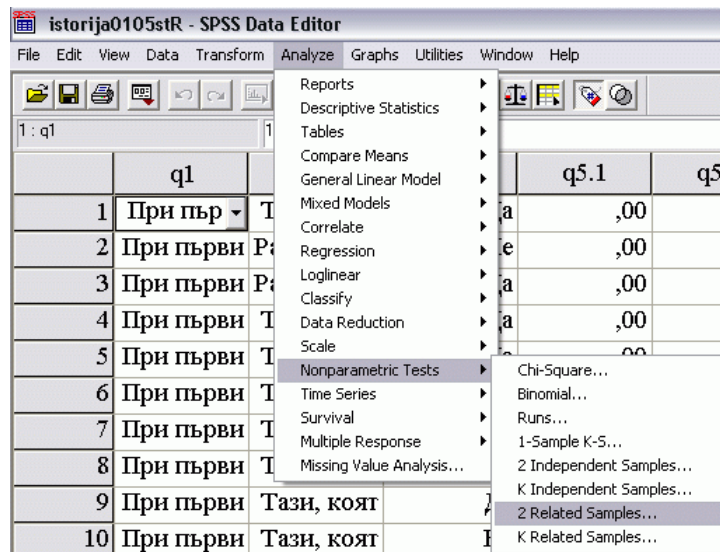


- В полетата Group 1 и Group 2 се избират шифрите на сравняваните подсъвкупности и се натиска бутона Continue.

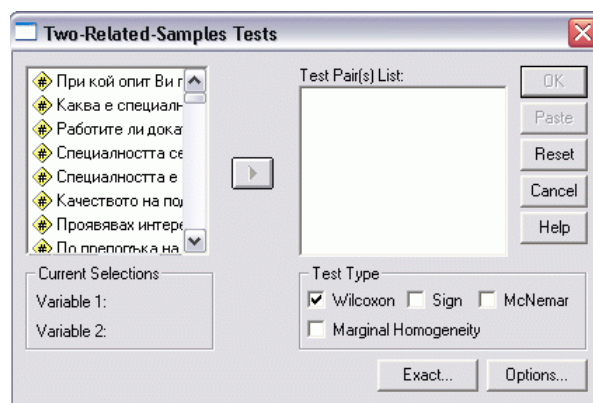
- В правоъгълника Test Type се маркира само Kolmogorov-Smirnov Z и се натиска бутона ОК.

Използване на непараметрични методи за проверка на хипотези при зависими извадки преминава през следната последователност:

- От менюто се избира:
 - ◆ Analyze
 - Nonparametric Tests
 - ❖ 2 Related Samples...



Появява се следният прозорец:



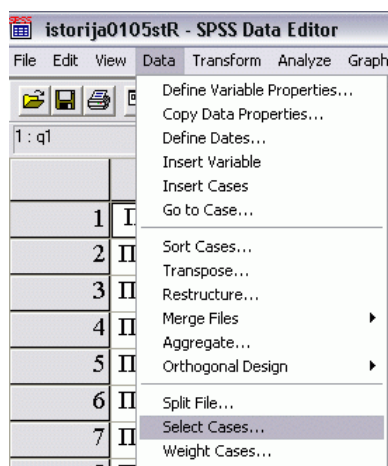
- Формират се двойки от съответните признаци, по един за всяка една от зависимите извадки.
- В правоъгълника Test Type се маркира Wilcoxon и Sign и след това се натиска бутона ОК.

Технологично допълнение към тема №10

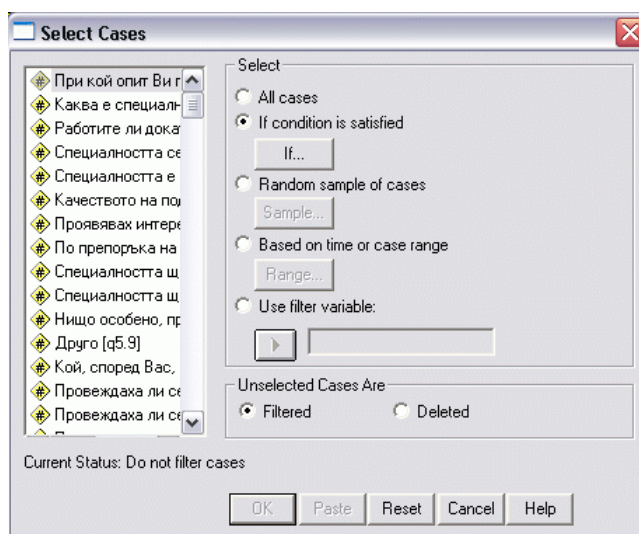
Анализ на връзки при качествен фактор и количествен резултат

Едно от условията за приложение на дисперсионния анализ е условието за нормалност на разпределенията по признака резултат във всяка група, формирана по значенията на признака фактор. Това условие се проверява по начина, разгледан в тема №7. Разликата тук е, че трябва предварително да се обособят групите по значенията на признака фактор. Това става в следната последователност:

- От менюто на файла с данни се избира:
 - ◆ Data
 - Select Cases...

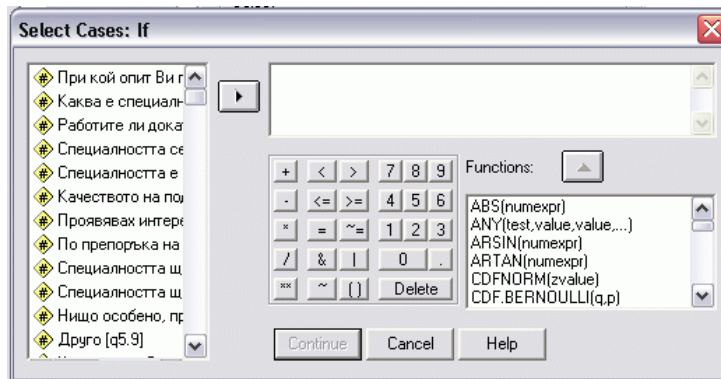


Появява се следният прозорец:



- В полето Select се избира If condition is satisfied и се натиска бутона If...

Появява се следният прозорец:



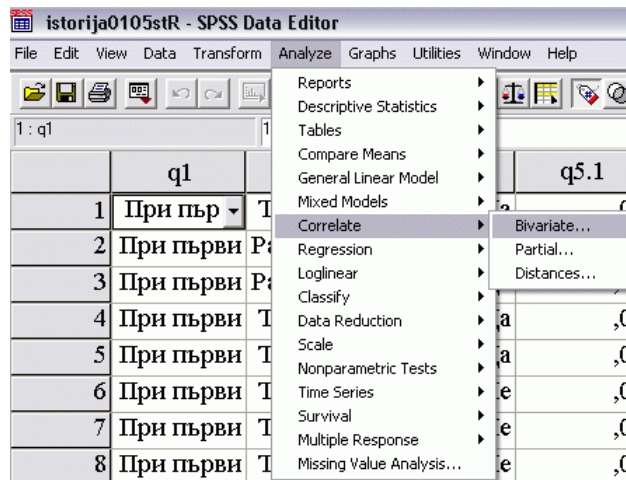
- В празното поле се описва групата, формирана по значението на признака фактор с помощта на името на признака от левия прозорец и на шифъра на конкретното значение.
- Последователно се натискат бутоните Continue и ОК.

Технологично допълнение към тема №11

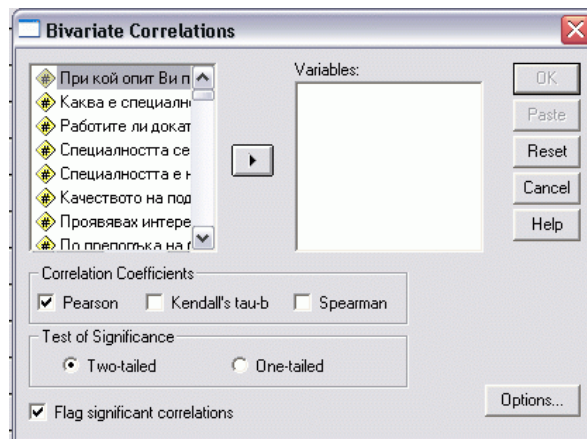
Анализ на връзки при количествен фактор и количествен резултат

Получаването на стойностите на обикновените и на ранговите коефициенти на корелация преминава през следната последователност:

- От менюто се избира:
 - ◆ Analyze
 - Correlate
 - ❖ Bivariate...



Появява се следният прозорец:



- В позицията Variables се избират всички признаци, между които ще се изчисляват коефициенти на корелация.
- В позицията Correlation Coefficients се избират Pearson (за обикновените коефициенти на корелация), Kendall's tau-b (за ранговия коефициент на корелация на Кендал) и Spearman (за ранговия коефициент на корелация на Спирман) и след това се натиска бутона ОК.

В резултат на тази процедура се получава т.нар. корелационна матрица, която показва наличието, посоката и силата на връзката между всеки два признака.