MINISTERUL EDUCAȚIEI ȘI CERCETĂRII AL REPUBLICII MOLDOVA
Universitatea Pedagogică de Stat „Ion Creangă” din Chișinău
Facultatea de Psihologie și Psihopedagogie Specială
Catedra Psihologie
Program de studii: **Psihologie**

LARISA CHIREV

ANALIZA STATISTICĂ COMPUTERIZATĂ / КОМПЬЮТЕРИЗИРОВАННЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ / COMPUTERIZED STATISTICAL ANALYSIS

Ghid metodic

Chișinău, 2026

Aprobat pentru editare de Senatul Universității Pedagogice de Stat „Ion Creangă” din Chișinău în ședința din 17.12.2025, proces verbal nr. 5

Recenzenți:

ADĂSCĂLIȚĂ VIORICA, doctor în psihologie, conferențiar universitar, Catedra de Psihologie, Universitatea Pedagogică de Stat „Ion Creangă”

DANILIUC Natalia, doctor în psihologie, conferențiar universitar, Catedra de Sănătate mintală, psihologie medicală și psihoterapie, Universitatea de Stat de Medicină și Farmacie „Nicolae Testemițanu”

DESCRIEREA CIP A CAMEREI NAŢIONALE A CĂRŢII DIN REPUBLICA MOLDOVA

Chirev, Larisa.

Analiza statistică computerizată = Компьютеризированный статистический анализ = Computerized statistical analysis / Larisa Chirev : Ghid metodic / Larisa Chirev ; Ministerul Educaţiei şi Cercetării al Republicii Moldova, Universitatea Pedagogică de Stat Ion Creangă din Chişinău, Facultatea de Psihologie şi Psihopedagogie Specială, Catedra Psihologie, Program de studii: Psihologie. – Chişinău : [S. n.], 2026 (CEP UPSC). – 82 p. : tab.

Text paral.: lb. rom., engl., rusă. – Referinţe bibliogr. la sfârşitul temelor. – [100] ex.

ISBN 978-9975-48-366-7.

159.9.07:311.2(075.8)=135.1=111=161.1

C 43

Tipărit la Centrul Editorial-Poligrafic al Universității Pedagogice de Stat “Ion Creangă” din Chișinău, str. Ion Creangă, nr. 1, MD-2069

© Chirev Larisa

© CEP UPSC

CUPRINS

INTRODUCERE 4
Partea I. STATISTICA DESCRIPTIVĂ 6
Tema 1. Concepte utilizate în prelucrarea statistică 6
Tema 2. Introducerea, editarea și crearea datelor în SPSS 11
Tema 3. Analiza descriptivă a datelor 13
Partea II. STATISTICA INFERENȚIALĂ 15
Tema 4. Corelații sau asocieri între variabile 15
Tema 5. Diferențe între 2 grupuri 20
Tema 6. Diferențe între 3 și mai multe grupuri 26

ВВЕДЕНИЕ 30
Часть I. ОПИСАТЕЛЬНАЯ СТАТИСТИКА 32
Тема 1. Понятия, используемые в статистической обработке данных 32
Тема 2. Ввод, редактирование и создание новых переменных в SPSS 37
Тема 3. Описательный анализ данных 39
Часть II. ИНДУКТИВНАЯ СТАТИСТИКА 41
Тема 4. Корреляции или ассоциации (зависимость) между переменными 47
Тема 5. Различия между двумя группами 53
Тема 6. Различия между 3-мя и более группами .....

INTRODUCTION 57
Part I. DESCRIPTIVE STATISTICS 59
Topic 1. Terms and concepts used in describing a distribution 59
Topic 2. Data entry, editing and creation in SPSS 64
Topic 3. Descriptive data analysis 66
Part II. INFERENTIAL STATISTICS 68
Topic 4. Correlations or associations between variables 68
Topic 5. Differences between two groups 73
Topic 6. Differences among three or more groups 79

INTRODUCERE

Acest ghid metodic este destinat studenților din anul terminal al programului de studii PSIHOLOGIE pentru a facilita activitatea lor de învățare în cadrul disciplinei **Analiza statistică computerizată** prevăzută, în conformitate cu planului de studii, pentru semestrul VI.

Unitatea de curs de specialitate are ca **obiectiv major** aprofundarea cunoștințelor teoretice despre utilizarea statisticii în domeniul psihologiei și formarea competențelor necesare pentru utilizarea programului SPSS. Aceste competențe sunt absolut necesare pentru activitatea de cercetare, dar și pentru înțelegerea corectă a rezultatelor unui studiu științific într-o documentare de rutină. Rezultatele învățării vor fi aplicate în planificarea cercetării experimentale, analiza descriptivă și inferențială a datelor în lucrarea de licență.

Competențele dezvoltate în cadrul cursului:

CG 2. Analiza datelor referitoare la situațiile concrete de natură psihologică.

CP 6. Adaptarea metodelor / tehnicilor de asistență psihologică la diverse contexte.

Rezultatele învățării:

La finalizarea cursului, studentul va fi capabil:

F-1. Să examineze / cerceteze obiectiv datele, informațiile, circumstanțele problemei / situației concrete de natură psihologică, evaluând calitatea datelor și a informațiilor acumulate pentru rezolvarea problematicii / situației.

F-2. Să aplice metode eficiente de evaluare psihologică în funcție de context, analizând rezultatele obținute.

Conținutul unităților de curs:

Tema 1. Concepte utilizate în prelucrarea statistică

Tipuri de variabile, indicatorii tendinței centrale, indicatorii de împrăștiere, indicatori ai tipului de distribuție, tipuri de distribuții, caracteristicile distribuției normale.

Tema 2. Introducerea, editarea și crearea datelor în SPSS

Copierea datelor dintr-un editor tabelar în fereastra Data View, editarea variabilelor în fereastra Variable View. Salvarea documentului construit în SPSS. Transformarea scorurilor și crearea de noi variabile.

Tema 3. Analiza descriptivă a datelor

Determinarea frecvențelor variabilelor nominale și ordinale și reprezentarea grafică a lor. Analiza descriptivă a variabilelor de tip scală. Testul de normalitate.

Lucrare de laborator nr. 1: Crearea unei baze de date în SPSS. Identificarea normalității distribuției pentru variabilele continue. Determinarea frecvențelor pentru variabilele dependente în raport cu o variabilă independentă.

Tema 4. Corelații sau asocieri între variabile

Selectarea testului statistic în funcție de caracteristicile distribuțiilor analizate. Utilizarea metodelor statistice parametrice (testul de corelație Pearson) și neparametrice (testul de asociere Kendall, testul de asociere Spearman, testul Chi-pătrat și testul exact Fisher). Interpretarea datelor.

Lucrare de laborator nr. 2: Corelarea / asocierea datelor și interpretarea rezultatelor

Tema 5. Diferențe între 2 grupuri

Selectarea testului statistic în funcție de caracteristicile distribuțiilor analizate. Compararea a 2 eșantioane independente (testul t pentru eșantioane independente și testul U Mann-Whitney) și a 2 eșantioane dependente (testul t pentru eșantioane dependente și testul Wilcoxon). Interpretarea datelor.

Tema 6. Diferențe între 3 și mai multe grupuri

Compararea a 3 și mai multe eșantioane independente (ANOVA unifactorială și Kruscal-Wallis). Interpretarea datelor.

Lucrare de laborator nr. 3: Compararea datelor a două și mai multe eșantioane și interpretarea rezultatelor.

Partea I. STATISTICA DESCRIPTIVĂ

Tema 1. CONCEPTE UTILIZATE ÎN PRELUCRAREA STATISTICĂ

A. Tipuri de variabile

1. ***Variabilele nominale***

- clasificarea diferitelor cazuri în categoriile prestabilite ale unei variabile
- Ex: gen, gtup de vârstă (copii și tineri), religie, cu / fără cafea
- Statistici: frecvența fiecărei categorii, modul

2. ***Variabile ordinale***

- categorii ordonate de la „inferior" la „superior", în funcţie de gradul în care acestea posedă trăsătura măsurată
- Ex: nivel educațional, nivelul intensității unei trăsături (redus / mediu / ridicat)
- distanțele dintre categorii nu sunt neapărat egale
- statistici: frecvența, modul, poziția în ierarhie a fiecărui subiect = rang

3. ***Variabile scală de tip interval***

- Subiecții pot lua scoruri aflate la intervale egale
- Răspund la întrebările: Care este diferența dintre? Cu cât este mai...?
- Ex: anul nașterii, scorul la un instrument de evaluare psiho-socială (IE, IQ, anxietate, agresivitate etc.)
- Zero stabilit arbitrar, nu absolut (ex: temperatura Celsius / Kelvin)
- De aceea nu permite calcularea proporțiilor
- Permite calcularea mediei și variației – teste statistice parametrice

4. ***Variabile scală de tip raport***

- Are punct zero natural – reflectă absența acelei caracteristici
- Răspund la întrebarea De câte ori este mai ...?
- Ex: număr de ani de școală, vârsta în ani, timp de reacție, nivelul adrenalinei
- Permite stabilirea proporțiilor – comparații între elemente

 **Exersare:**	**Exercițiul 1.** Clasificaţi datele oferite de răspunsurile la întrebările de mai jos în cantitative și calitative și indicaţi scala de măsurare pentru fiecare dintre ele. a) Care este vârsta dumneavoastră? b) Sunteţi persoana care face de obicei cumpărături pentru familia dvs.? c) Cate persoane sunt în familia dv.? d) Cum notaţi, pe o scală de la 1 la 10, gustul produsului, dacă 1 este cel mai puţin gustos? e) Veţi cumpăra acest produs dacă va fi disponibil în magazine? f) Dacă răspunsul la e) este "Da", cât de des veţi cumpăra produsul?

B. Concepte utilizate în descrierea unei distribuții

1) Indicatorii tendinței centrale

- ***Media („m" sau „M")*** – media aritmetică a unui grup, se calculează prin sumarea scorurilor individuale ale tuturor subiecților din grup și împărțirea sumei obținute la nr de subiecți din grup

  Ex: 2, 3, 5, 6, 7 M=(2+3+5+6+7):5= 23:5 = 4,6

- ***Mediana („Me")*** – este scorul central, punctul de mijloc al mulțimii de scoruri pe care le împarte în două părți egale

  - pentru număr impar de scoruri – e nr întreg *(ex: 2, 3, **5**, 6, 7 Me=5)*

  - pentru număr par de scoruri – se află între valori

  Ex: 2, 3, 4 /5, 6, 7 Me=4,5 sau Ex: 2, 3 /5, 6 Me=4

- ***Modul („Mo")*** – scorul cu cea mai ridicată frecvență

  Ex. 2, 3, 5, 6, 6, 7 Mo=6 (distribuție unimodală)

  Ex. 2, 3, 3, 5, 6, 6, 7 Mo=3 și Mo=6 (distribuție bimodală)

  Ex. 2, 3, 3, 3, 5, 6, 6, 6, 7, 8, 9, 9, 9 Mo=3, Mo=6 și Mo=9 și (distribuție multimodală)

2) Indicatori de împrăștiere

- ***Amplitudinea variației („A")*** – diferența dintre scorul maxim și scorul minim al distribuției: A = V max – V min

  Ex: 2, 3, 5, 6, 7 A=7-2=5

- ***Abaterea medie (d)*** – modul în care se abat rezultatele de la medie: **d ind= V– M**. Suma abaterilor medii întotdeauna este egală cu 0.

  Ex: 2, 3, 5, 6, 7 M=4,6;

Abaterea medie a scorurilor individuale: d1=2-4,6= - 2.6

d2=3-4,6= - 1.6

d3=5-4,6= 0.4

d4=6-4,6= 1.4

d5=7-4,6= 2.4

Abaterea medie a unei serii se calculează: $\mathbf{d = \sum |V - M| / n}$. Se adună toate abaterile fără a ține cont de valență și se împart la n: 8,4/5=1,68 (abaterea medie a seriei de date).

- ***Dispersia / varianța („s²")*** – media pătratică a diferențelor scorurilor individuale față de media aritmetică (media abaterilor medii): $\boldsymbol{s^2 = \sum d^2/n}$

  Ex: $s^2 = ((-2,6)^2 + (-1,6)^2 + (0,4)^2 + (1,4)^2 + (2,4)^2)/5 = (6,76+2,56+0,16+1,96+5,76) / 5 = 17,2 / 5 = 3,44$

- ***Abaterea standard („s")*** **–** rădăcina pătrată din valoarea dispersiei: $\mathbf{s=\sqrt{s^2}}$

  Ex: $s=\sqrt{3,44} = 1,85$

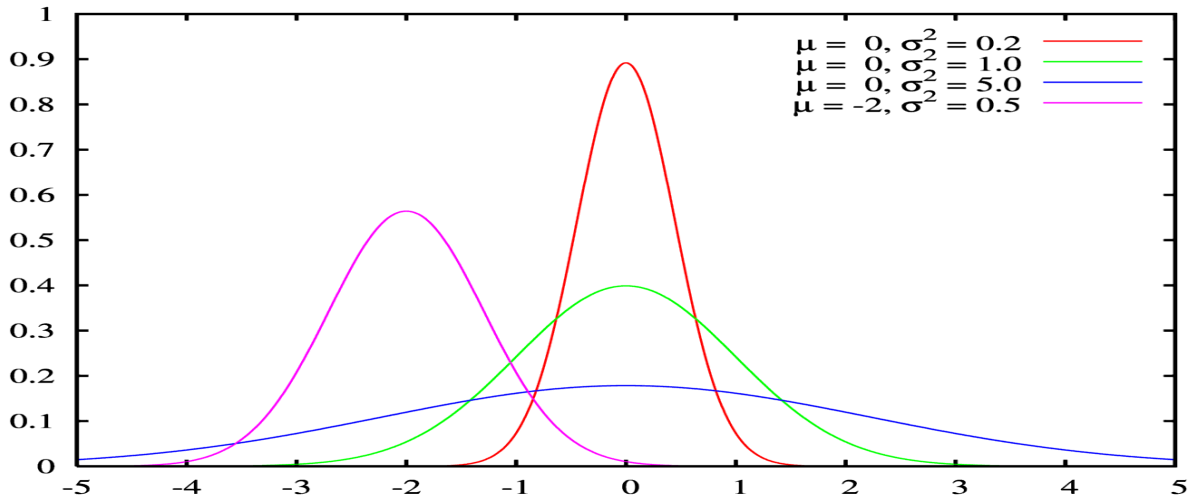
  Abaterea standard este un indicator care ne oferă informaţii privind modul de împrăştiere a termenilor unei serii cu tendinţă de normalitate. Astfel, o regulă empirică spune:

  - 68,37% din valorile unei serii se situează în intervalul (M – s , M + s);
  - 98,45% din valorile unei serii se situează în intervalul (M – 2s , M + 2s);
  - 99,73% din termenii unei serii se situează în intervalul (M – 3s , M + 3s);
  - 99,94% din termeni se găsesc în intervalul (M – 4s , M + 4s).

- ***Coeficientul de variație („CV")*** **–** gradul de împrăștierea a scorurilor în jurul mediei, se calculează împărțind abaterea standard la medie și înmulțind cu 100%: CV=s/M·100% *(ex: 1,85 / 4,6 x 100% =40,2%)*
  - Valorile sub 10% – indică o populație omogenă
  - Valorile între 10%-20% – indică o populație relativ omogenă
  - Valorile între 20%-30% – indică o populație relativ eterogenă
  - Valorile peste 30% – indică o populație eterogenă.

3) Indicatori ai tipului de distribuție

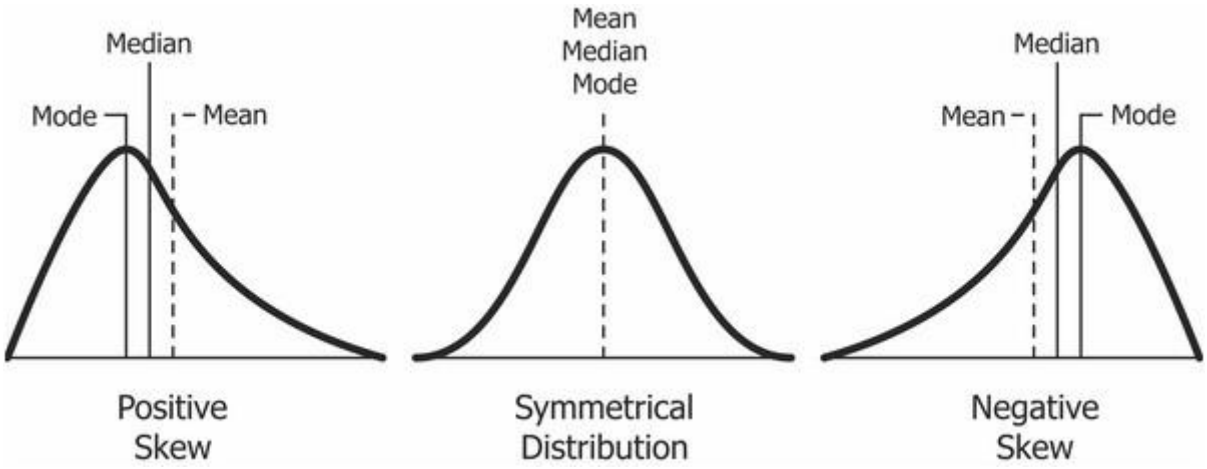
- ***Boltirea (kurtosis)*** **–** arată cât de ascuțită sau aplatizată este o distribuție:
  - distribuție normală: boltirea=3, excesul de boltire=0
  - distribuție leptocurtică (ascuțită): boltirea>3, excesul de boltire>0 *(roșie)*
  - distribuție platicurtică (aplatizată): boltirea<3, excesul de boltire<0 *(albastră)*



- ***Simetria (skewness)***
    - distribuție simetrică – de la – ½ până la + ½
    - distribuție izomoderat asimetrică – între -1 și -½ sau între ½ și 1
    - distribuție puternic asimetrică – mai mică ca -1 sau mai mare ca 1

Asimetrie de stânga sau pozitivă – predomină valori mai mici decât media aritmetică

Asimetrie de dreapta sau negativă – predomină valori mai mari decât media aritmetică



Caracteristicile distribuției normale:

- Este unimodală
- Simetrică față de medie
- Boltirea normală
- Limitele stânga – dreapta tind către zero
- Este valabilă doar pentru variabile continue
- Indicatorii tendinței centrale au aceeași valoare: M=Me=Mo

<table>
<tr>
<td rowspan="2">
Exersare:</td>
<td>Exercițiul 2.
Determinați media, mediana, modul, amplitudinea variației, abaterea medie, dispersia, abaterea standard și coeficientul de variație a celor 2 mulțimi de scoruri:
I mulțime: 1, 4, 5, 2, 2, 3, 4
M=
Me=
Mo=
A=
d=
s^2=
s=
Cv=
II mulțime: 1, 4, 5, 2, 2, 3, 4, 5, 5
M=
Me=
Mo=
A=
d=
s^2=
s=
Cv=</td>
</tr>
<tr>
<td>Exercițiul 3.
Reieșind din datele obținute în exercițiul 2 pentru cele două mulțimi de date și caracteristicile unei distribuții normale, determinați tipul distribuției (normală sau eparametrică) a scorurilor acestor mulțimi. Argumentați-vă opinia.</td>
</tr>
</table>

Surse recomandate:

1. CLOCOTICI, V., STAN, A. Statistică aplicată în psihologie. Iași: Polirom, 2000. ISBN 973-683-577-4
2. GHEORGIU, D. Statistică pentru psihologi. Note de curs 1. Disponibil: https://www.studocu.com/ro/document/universitatea-titu-maiorescu/statistica-psihologica/statistica-i-note-de-curs-1/10155971
3. RACU, Ig. Psihodiagnostic și statistică psihologică. Chișinău: CEP UPS, 2017. ISBN 978-9975-46-321-8.
4. VASILIU, D. Momento de cercetare în psihologie. București: Editura Didactică și pedagogică, 2018. ISBN 978-606-31-0501-2
5. http://www.phys.ubbcluj.ro/~dana.maniu/BIOSTAT/C4.pdf

Tema 2. INTRODUCEREA, EDITAREA ȘI CREAREA DATELOR ÎN SPSS

1. Copiați datele din Excel în SPSS.

2. În editorul VARIABLE VIEW completați datele:

- ***Name*** – scrieți numele scurt al variabilei fără spații (indica numele variabilei în editorul DATA VIEW)
- ***Type*** – numeric
- ***Decimals*** – scrieți 0 dacă aveți numere întregi sau 2 dacă aveți numere cu sutimi
- ***Label*** – scrieți numele variabilei
- ***Values*** – introduceți codurile doar la variabilele nominale sau ordinale
- ***Align*** – pe centru
- ***Measure*** – tipul variabilei (nominal, ordinal, scale (scală de interval sau de raport)

3. Determinați indicatorii tendinței centrale, de împrăștiere și a tipului de distribuție pentru variabilele de interval / de raport:

Analyse – Descriptive Statistics – Frequencies

Statistics:

Central Tehndency: ***Mean, Median, Mode***

Dispersion: ***Std. deviation, Variance, Range, Minimum, Maximum***

Distribution: ***Skewness, Kurtosis***

4. Transformați veniturile lunare în categorii „sub medie” și „peste medie”:

Transform – Recode into Different Variables – duceți variabila necesară în fereastră.

Name: nivVL08, **Label**: nivel venit lunar 08

Old an New Values – **Range** (înscrieți valoarea minimă), **through** (înscrieți valoarea mediei)

New Value – înscrieți „1”, tastați butonul **Add**

Old an New Values – Range (înscrieți valoarea mediei+1), **through** (înscrieți valoarea maximă)

New Value – înscrieți „2”, tastați butonul **Add**, apoi butonul **Continue** de mai jos.

Pe pagina **Data wiew** editați **Label** pentru variabila nou creată**.**

5. Transformați (pașii din p. 4) scorurile de la variabilele „intergrare în colectiv” în nivele în corespundere cu următoarele intervale:

1-4 – nivel scăzut

5-7 – nivel mediu

8-10 – nivel înalt

Denumiți noile variabile create în felul următor: ***Name:*** NivInt08, ***Label:*** nivel integrare 08.

6. Creați variabila MEDIA INTEGRĂRII PE 3 ANI:

   Transform – Compute Variable

   Target Variable: MedInt

   Numeric Expression: (integr 08+ integr 09 + integr 10)/3. Tastați butonul **OK**

7. Transformați „media integrare” în nivele (nivelele din p.5, pașii din p. 4)

 **Exersare:**	**Exercițiul 4.** Creați un nou document **.sav**, utilizând o nouă bază de date în excel. Editați variabilele.
	Exercițiul 5. Creați noi variabile transformând variabilele continue în variabile ordinale. Editați noile variabile.
	Exercițiul 6. Determinați pentru toate variabilele continue următorii indicatori: media, mediana, modul, amplitudinea variației, dispersia, abaterea standard, coeficientul de variație, valoarea minimă și maximă.

Surse recomandate:

1. Introducere în SPSS. Partea I - Componente SPSS. Disponibil: https://www.youtube.com/watch?v=mtKHtsMB7s0
2. Introducere în SPSS. Partea II – Intrare variabile în SPSS. Disponibil: https://www.youtube.com/watch?v=sszSeqOKiic
3. VASILIU, D. Momento de cercetare în psihologie. București: Editura Didactică și pedagogică, 2018. ISBN 978-606-31-0501-2

Tema 3. ANALIZA DESCRIPTIVĂ A DATELOR

1. Utilizând baza de date 1 (creată la tema 2), determinați frecvențele în funcție de gen, nivelul venitului lunar pe ani, prezența sancțiunilor pe ani, atitudinea față de management, nivelul autonomiei percepute și nivelul integrării în colectiv pe ani:

   Analyse – Descriptive Statistics – Frequencies

2. Determinați frecvențele ***în funcție de gen*** a următoarelor variabile:
   - nivelul venitului lunar pe ani,
   - prezența sancțiunilor pe ani,
   - atitudinea față de management,
   - nivelul autonomiei percepute
   - nivelul integrării în colectiv pe ani,

   Analyse – Descriptive Statistics – Crosstabs – Rows (variabilele care ne interesează) / **Columns** (variabila independentă – ex. gen)
   - **Cells – Percentages: *Row*, *Column, Total***

     Tastați butonul **Continue**, apoi **OK**

3. Determinați frecvențele ***în funcție sancțiuni 08*** a următoarelor variabile (pașii din p. 9):
   - nivelul venitului lunar 08,
   - atitudinea față de management,
   - nivelul autonomiei percepute 08.

4. Verificați normalitatea distribuției scorurilor pentru „venit lunar 08” și „venit lunar 09” cu ajutorul testului Kolmogorov-Smirnov:

   Analyse – Descriptive Statistics – Explore

   Plots – bifați ***Normality plots with tests***

   Tastați **Continue**, apoi **OK**

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IQ	.145	40	.034	.920	40	.008
EQ	.087	40	.200*	.977	40	.568

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Testul Kolmogorov-Smirnov compară distribuția empirică cu cea teoretică (normală).

Privim **Sig.** Dacă valoarea este mai mare de 0,05, înseamnă că distribuția se apropie de cea normală, respectiv putem utiliza teste parametrice de analiză inferențială.

Dacă e mai mică de 0,05 – există diferențe între distribuția empirică și cea normală.

În exemplul dat la variabila IQ p=0,034, deci distribuția scorurilor individuale la variabila IQ ale celor 40 de subiecți diferă de distribuția normală (curba Gauss).

Pentru variabila EQ p=0,2, distribuția scorurilor la această variabilă nu diferă semnificativ de distribuția normală.

 **Exersare:**	**Exercițiul 7.** Determinați frecvențele (nr și %) scorurilor variabilelor continue.
	Exercițiul 8. Utilizând opțiunea ***Crosstabs***, generați tabele cu frecvențe (nr și %) prin încrucișarea a doua variabile categoriale.
	Exercițiul 9. Utilizând testul de normalitate, determinați tipul distribuției datelor la toate variabilele continue.

Surse recomandate:

1. Introducere în SPSS. Partea III – Statistici descriptive: Media. Disponibil: https://www.youtube.com/watch?v=z6f__Nz6Imc
2. Introducere în SPSS. Partea III – Statistici descriptive: Modul. Disponibil: https://www.youtube.com/watch?v=m40dYChK7pk
3. Introducere în SPSS. Partea III – Statistici descriptive: Modul. Disponibil: https://www.youtube.com/watch?v=WeuaQ4IDPTY
4. Introducere în SPSS - Partea IV - Amplitudinea, Varianța și Abaterea Standard. Disponibil: https://www.youtube.com/watch?v=DjXTEZq59Zk
5. VASILIU, D. Momento de cercetare în psihologie. București: Editura Didactică și pedagogică, 2018. ISBN 978-606-31-0501-2

Partea II. STATISTICA INFERENȚIALĂ

Tema 4. CORELAȚII SAU ASOCIERI ÎNTRE VARIABILE

Decizia cu referire la testul statistic care va fi utilizat se va lua în funcție de tipul variabilelor și a caracteristicilor distribuției, conform tabelului:

Variabila 1	*Variabila 2*	*Testul statistic*
Interval / raport	Interval / raport	**r Pearson** (pentru distribuții parametrice)
		Kt Kendall-tau (pentru distribuții neparametrice)
Ordinală	Ordinală	**rs Spearman**
Categorială (nominală sau ordinală)	Categorială (nominală sau ordinală)	**χ^2 Chi-pătrat** (dacă măcar o variabilă are 3 și mai multe categorii) **Testul exact Fisher** (doar daca ambele variabile sunt dihotomice)

Aprecierea **tăriei corelației** după valoarea coeficientului de corelație, conform lui Davis:

0,01 – 0,09	Corelație neglijabilă
0,10 – 0,29	Corelație scăzută / slabă
0,30 – 0,49	Corelație moderată
0,50 – 0,69	Corelație substanțială
0,70 – 1	Corelație foarte puternică

1. Corelație Pearson

Analyze – Correlate – Bivariate

Se bifează **Pearson**, **Two-tailed**, **Flag significant correlations**

Tastați **OK**

Correlations

		IQ	EQ
IQ	Pearson Correlation	1	-.081
	Sig. (2-tailed)		.620
	N	40	40
EQ	Pearson Correlation	-.081	1
	Sig. (2-tailed)	.620	
	N	40	40

Coeficientul de corelație r= -0,081 (este o corelație negativă), iar pragul de semnificație p=0,62. Deci, IQ nu corelează cu EQ.

Dacă ar fi r= 0,75, p≤ 0,05, atunci am fi afirmat: *EQ corelează foarte puternic cu IQ, această corelație fiind una pozitivă / direct proporțională: r= 0,75, p≤ 0,05.*

2. Asociere Kendall

Analyze – Correlate – Bivariate

Se bifează **Kendall's tau-b**, **Two-tailed**, **Flag significant correlations**

Tastați **OK**

Correlations

			IQ	EQ
Kendall's tau_b	IQ	Correlation Coefficient	1.000	-.049
		Sig. (2-tailed)	.	.664
		N	40	40
	EQ	Correlation Coefficient	-.049	1.000
		Sig. (2-tailed)	.664	.
		N	40	40

Datele obținute la variabila EQ nu au relevat o asociere semnificativă între ranguri cu cele obținute la variabila IQ: Kt = -0,049, p= 0,664.

În cercetări se prezintă doar valorile coeficienților care indică corelație sau asociere semnificativă din punct de vedere statistic, adică sunt la p≤0,05.

3. Asociere Spearman

Analyze – Correlate – Bivariate

Se bifează **Spearman, Two-tailed, Flag significant correlations**

Tastați **OK**

Correlations

			IQ	EQ
Spearman's rho	IQ	Correlation Coefficient	1.000	-.063
		Sig. (2-tailed)	.	.698
		N	40	40
	EQ	Correlation Coefficient	-.063	1.000
		Sig. (2-tailed)	.698	.
		N	40	40

Intre rangurile variabilei „IQ" și cele ale variabilei „EQ" există o asociere nesemnificativă / nu există o asociere semnificativă (p=0,698).

4. Testul Chi-pătrat

Pentru variabile categoriale: nominale sau ordinale. ***Chi-pătrat se utilizează doar atunci când cel puțin o variabilă are 3 și mai multe categorii*** **(**de ex. variabila sex are 2 categorii – fete și băieți; variabila nivel de studii are 3 categorii: liceu, colegiu, universitate). De exemplu, vrem să vedem dacă este vreo asociere între gen și nivelul studiilor, cu alte cuvinte dacă diferă semnificativ numărul fetelor

și băieților din liceu, colegiu și universitate. Sau dorim să vedem dacă există o asociere între nivelul de studii și nivelul IQ.

Analyse – Descriptive Statistics – Crosstabs – Rows (o variabilă), **Columns** (altă variabilă). Dacă doriți histogramele frecvențelor, bifați **Display clustered bar charts**

- **Statistics: Chi pătrat; Phi and Cramers V**. Tastați **Continue**
- **Cells: Percentages – Row, Column.** Tastați butonul **Continue**, apoi **OK**

genul * studii Crosstabulation

			studii			Total
			liceu	colegiu	universitate	
genul	feminin	Count	7	10	3	20
		% within genul	35.0%	50.0%	15.0%	100.0%
		% within studii	100.0%	50.0%	23.1%	50.0%
	masculin	Count	0	10	10	20
		% within genul	0.0%	50.0%	50.0%	100.0%
		% within studii	0.0%	50.0%	76.9%	50.0%
Total		Count	7	20	13	40
		% within genul	17.5%	50.0%	32.5%	100.0%
		% within studii	100.0%	100.0%	100.0%	100.0%

Cum se citesc datele din tabel:

Pe rând: Din totalul de 20 de fete 7 (35%) au absolvit liceul, 10 (50%) – colegiul și doar 3 (15%) au făcut facultate.

Pe coloană: Din cei 7 subiecți cu liceu, toate sunt fete (100%), din cei 20 subiecți ce au absolvit colegiu 50% sunt fete (10), 50% - băieți (10), din cei 13 subiecți cu facultate 23,1% sunt fete (3) și 76,9 (10) sunt băieți.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.769[a]	2	.005
Likelihood Ratio	13.681	2	.001
Linear-by-Linear Association	10.005	1	.002
N of Valid Cases	40		

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is 3.50.

Cum se citesc datele din tabel:

Din tabelul de mai sus ne interesează valoarea coeficientului Chi-pătrat și pragului de semnificație.

Cercetarea făcută pe 40 de subiecți confirmă o frecvență mai mare a băieților cu facultate în comparație cu fetele (χ^2 = 10, 697, df=2, p=0,005≤0,05), ceea ce înseamnă că este o asociere semnificativă a nivelului studiilor și genului (frecvența mai mare a fetelor cu liceu spre deosebire de frecvența mai mare a băieților cu facultate).

Dat fiind faptul că am obținut un rezultat statistic semnificativ, este necesar un indicator suplimentar care să indice intensitatea legăturii – acesta e coeficientul phi Cramer din tabelul de mai jos.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.519	.005
	Cramer's V	.519	.005
N of Valid Cases		40	

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

Cum se citesc datele din tabel:

Diferențele de frecvențe ale fetelor și băieților în raport cu nivelul studiilor sunt de intensitate mare: phi=0,519.

Valoarea lui Phi se raportează la următorul tabel:

Phi / φ	0, 01 – 0,24	Efect mic
	0,25 – 0,39	Efect mediu
	Peste 0,40	Efect mare

5. Testul exact Fisher

Pentru variabile categoriale: nominale sau ordinale. ***Testul se aplică când ambele variabile sunt dihotomice*** (ex: variabila sex cu 2 categorii – fete și băieți, variabila nivel EQ cu 2 categorii – sub medie și peste medie).

De exemplu, vrem să vedem dacă este există o asociere semnificativă între gen și nivelul EQ, cu alte cuvinte, diferă semnificativ numărul fetelor și băieților cu EQ sub medie de numărul fetelor și băieților cu EQ peste medie. Testul lucrează cu frecvențele, nu cu scorurile individuale.

Analyse – Descriptive Statistics – Crosstabs – Rows (o variabilă), **Columns** (altă variabilă).

- Exact: bifați **Exact.** Tastați **Continue**
- Statistics: **Chi pătrat**. Tastați **Continue**

Tastați butonul **OK**

genul * nivel EQ Crosstabulation

			nivel EQ		Total
			sub medie	peste medie	
genul	feminin	Count	9	11	20
		% within genul	45.0%	55.0%	100.0%
		% within nivel EQ	47.4%	52.4%	50.0%
	masculin	Count	10	10	20
		% within genul	50.0%	50.0%	100.0%
		% within nivel EQ	52.6%	47.6%	50.0%
Total		Count	19	21	40
		% within genul	47.5%	52.5%	100.0%
		% within nivel EQ	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.100[a]	1	.752	1.000	.500	
Continuity Correction[b]	.000	1	1.000			
Likelihood Ratio	.100	1	.751	1.000	.500	
Fisher's Exact Test				1.000	.500	
Linear-by-Linear Association	.098[c]	1	.755	1.000	.500	.236
N of Valid Cases	40					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 9.50.
b. Computed only for a 2x2 table
c. The standardized statistic is -.313.

Pragul de semnificație bidirectional este 1, pragul de semnificație unidirectional este 0,5, ambele fiind mai mari de 0,05 – pragul de semnificație admis.

Cum se citesc datele din tabel:

Între participanții de gen feminin și cei de gen masculin nu există diferențe privind nivelul EQ. Raportul fete – băieți cu EQ sub medie nu diferă de raportul fete ă băieți cu EQ peste medie.

 **Exersare:**	**Exercițiul 10.** Corelați 3 variabile continue selectând testul potrivit în funcție de tipul de distribuție a datelor.
	Exercițiul 11. Asociați 2 variabile categoriale selectând testul potrivit în funcție de nr de categorii al variabilelor.

Surse recomandate:

1. Analiza corelației dintre două variabile. Disponibil: https://www.youtube.com/watch?v=r_w-WYMx4Ug
2. CLOCOTICI, V., STAN, A. Statistică aplicată în psihologie. Iași: Polirom, 2000. ISBN 973-683-577-4
3. Introducere în SPSS - Partea V – Corelația. Disponibil: https://www.youtube.com/watch?v=c7oNwdKHWA8
4. VASILIU, D. Momento de cercetare în psihologie. București: Editura Didactică și pedagogică, 2018. ISBN 978-606-31-0501-2

Tema 5. DIFERENȚE ÎNTRE DOUĂ GRUPURI

Decizia cu privire la testul utilizat se va lua conform indicațiilor din tabel:

Variabila independentă			*Variabila dependentă*	*Testul statistic*
Categorială (nominală sau ordinală)	Două categorii	Grupuri independente	Interval / raport (parametrică) Distribuția este una normală	**t** pentru eșantioane independente
			Ordinală (neparametrică) Interval / raport (distribuția NU este normală)	**U Mann-Whitney**
		Grupuri dependente	Interval / raport (parametrică)	**t** pentru eșantioane dependente
			Ordinală (neparametrică) sau Interval / raport (distribuția NU este normală)	**Wilcoxon**
	Trei și mai multe categorii	Grupuri independente	Interval / raport (parametrică)	**ANOVA unifactorială**
			Ordinală (neparametrică) sau Interval / raport (distribuția NU este normală)	**Kruskal-Wallis**

1. Testul t pentru eșantioane independente

Testul se folosește pentru a **compara rezultatele a 2 grupuri obținuite la variabilele continue, distribuțiile cărora sunt normale**.

De exemplu, ne interesează dacă există diferențe între scorurile subiecților obținute la testele de IQ și EQ în funcție de gen (2 categorii – fete și băieți) și distribuțiile datelor la aceste variabile sunt normale.

Analyze – Compare Means – Independent-Simple T Test

În **Test Variables** se transferă variabilele dependente (în cazul nostru ***EQ și IQ***)

În **Grouping Variable** se transferă variabila independentă (în cazul nostru sex), apoi definim variabila independentă tastând butonul ***Define Group***: Group 1 și Group 2. Contează cum au fost codificate categoriile în baza de date SPSS pe pagina **Label View** în coloana **Values**.

Tastați butonul **Continue**, apoi **OK**

Group Statistics

	genul	N	Mean	Std. Deviation	Std. Error Mean
IQ	feminin	20	85.7000	18.22404	4.07502
	masculin	20	116.6000	8.97013	2.00578
EQ	feminin	20	46.6500	10.20462	2.28182
	masculin	20	45.4500	8.70859	1.94730

La variabila „IQ” media în grupul de fete este 85,7, iar media în grupul de băieți este 116,6. Se pare că la băieți este mai dezvoltat IQ. În schimb la fete este mai dezvoltat EQ (media 46,65) decât la băieți (media 45,45). La întrebarea *Sunt statistic semnificative aceste diferențe?* ne răspund datele din tabelul de mai jos:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
IQ	Equal variances assumed	7.841	.008	-6.803	38	.000	-30.90000	4.54191	-40.09462	-21.70538
	Equal variances not assumed			-6.803	27.696	.000	-30.90000	4.54191	-40.20828	-21.59172
EQ	Equal variances assumed	.008	.931	.400	38	.691	1.20000	2.99978	-4.87274	7.27274
	Equal variances not assumed			.400	37.083	.691	1.20000	2.99978	-4.87767	7.27767

La fiecare variabilă sunt câte 2 valori ale testului t, de ex. la IQ e -6,803, și 2 valori ale lui p, de ex. ≤0,01. În cazul dat ele sunt egale, dar se întâmplă ca ele să fie diferite și atunci urmează de decis pe care să le luăm?

Pentru a decide, trebuie luat în considerație **p** din coloana a treia: testul Levene indică dacă variațele celor 2 distribuții sunt egale sau nu, în cazul nostru dacă distribuția datelor în grupul de fete este egală cu distribuția datelor în grupul de băieți. Dacă p≤0,05 înseamnă că sunt diferențe semnificative în variațele celor 2 distribuții și luam valoarea lui t din rândul de jos (Equal variances not assumed). În exemplul nostru, pentru IQ se va lua valoarea lui din rândul de jos (Equal variances not assumed), iar în cazul EQ – valoarea lui t din rândul de sus (Equal variances assumed).

Cum se citesc datele din tabel:

Între media scorurilor obținute de fete (M=85,7) și media celor obținute de către băieți (M=116,6) sunt diferențe statistic semnificative (t=-6,803, p≤0,01).

Scorurile obținute de fete (M=85,7) la testarea IQ sunt semnificativ mai mici (t=-6,803, p≤0,01) decât cele obținute de către băieți (M=116,6).

Media scorurilor obținute la variabila „IQ” este semnificativ mai mare la băieți (media 116,6) decât la fete (media 85,7): t=-6,803, p≤0,01.

La variabila EQ nu sunt diferențe de gen.

2. Testul U Mann-Whitney

Testul se folosește pentru a **compara rezultatele a 2 grupuri obținute la variabilele ordinale sau continue,** ale căror **distribuții sunt neparametrice** (diferă de distribuția normală).

De exemplu, ne interesează dacă există diferențe între scorurile subiecților obținute la testele de IQ și EQ în dependență de gen (2 categorii – fete și băieți), iar distribuțiile datelor la aceste variabile sunt neparametrice (diferă de distribuția normală).

Analyze – Nonparametric Tests – Legacy Dialogs – 2 Independent Simples

În **Test Variables Lists** se transferă variabilele dependente (în cazul nostru ***IQ și EQ***)

În **Grouping Variable** se transferă variabila independentă (în cazul nostru sex) – apoi definim variabila independentă tastând butonul ***Define Group***. Contează cum au fost codificate categoriile în baza de date SPSS pe pagina **Label View** în coloana **Values**

Bifați **U Mann-Whitney,** apoi tastați butonul **Continue**, apoi **OK**

Ranks

	genul	N	Mean Rank	Sum of Ranks
IQ	feminin	20	10.50	210.00
	masculin	20	30.50	610.00
	Total	40		
EQ	feminin	20	21.45	429.00
	masculin	20	19.55	391.00
	Total	40		

La variabila „IQ” în grupul de fete media rangurilor este 10,5, în grupul de băieți media rangurilor este 30,5. De aici rezultă că scorurile băieților sunt mai mari decât ale fetelor.

Test Statistics[a]

	IQ	EQ
Mann-Whitney U	.000	181.000
Wilcoxon W	210.000	391.000
Z	-5.434	-.515
Asymp. Sig. (2-tailed)	.000	.607

a. Grouping Variable: genul

Cum se citesc datele din tabel:

Scorurile la variabila „IQ” ale băieților sunt semnificativ mai mari (media rangurilor 30,5) decât cele ale fetelor (media rangurilor 10,5): U=0, N_1=20, N_2=20, p≤0,01. La EQ nu au fost constatate diferențe de gen semnificative din punct de vedere statistic.

3. Testul t pentru eșantioane dependente/perechi

Testul se folosește când **au fost efectuate 2 măsurători cu aceeași probă asupra aceluiași eșantion în momente diferite** (test și retest), iar **distribuțiile scorurilor sunt normale**.

Analyze – Compare Means – Paired-Simple T Test

În **Paired Variables** se transferă una din variabile în celula ***Variable 1*** (de ex. EQ 1) și pe cealaltă în ***Variable 2*** (de ex. EQ 2)

Tastați butonul **OK**

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	EQ 1	46.0500	40	9.38343	1.48365
	EQ 2	47.9000	40	7.93499	1.25463

La variabila „EQ 1" media test a fost 46,05, media retest este 47,9. De obicei, retestarea se face după o careva intervenție. Vedem că a crescut media după intervenție, întrebarea e „Este oare semnificativă din punct de vedere statistic această creștere?". Răspunsul stă în tabelul următor:

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	EQ 1 - EQ 2	-1.85000	2.25945	.35725	-2.57261	-1.12739	-5.178	39	.000

Cum se citesc datele din tabel:

Scorurile la testarea empatiei după realizarea programului de intervenție (m=47,9, s=7,935) diferă semnificativ (t= - 5,178, df=39, p≤0,05) de cele de la testarea inițială (m=46,05, s=9,383).

df – numărul gradelor de libertate

4. Testul Wilcoxon

Testul se folosește când **au fost efectuate 2 măsurători cu aceeași probă asupra aceluiași eșantion în momente diferite** (test și retest) și avem: 1) **variabile ordinale** sau 2) **variabile de raport sau interval, iar distribuțiile scorurilor NU sunt parametrice**.

Analyze – Nonparametric Tests – Legacy Dialogs – 2 Related Simples

În **Test Pair** se transferă una din variabile în celula ***Variable 1*** (de ex. EQ1) și pe cealaltă în ***Variable 2*** (de ex. EQ2)

În **Test Type** bifați **Wilcoxon,** apoi tastați butonul **OK**

Ranks

		N	Mean Rank	Sum of Ranks
EQ 2 - EQ 1	Negative Ranks	0[a]	.00	.00
	Positive Ranks	22[b]	11.50	253.00
	Ties	18[c]		
	Total	40		

a. EQ 2 < EQ 1
b. EQ 2 > EQ 1
c. EQ 2 = EQ 1

La 22 subiecți scorurile au crescut după intervenție, la 18 – au rămas aceleași.

Test Statistics[a]

	EQ 2 - EQ 1
Z	-4.124[b]
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

Cum se citesc datele din tabel:

Intervenția a avut efect (Z/W= -4,124, p≤0,01).

Efectul este prezent și la fete, și la băieți?

Data – Split file – bifăm ***Organize output by groups*** și transferăm în câmpul **Groups Based on** variabila ***gen***

Tastăm **OK**

Ranks[a]

		N	Mean Rank	Sum of Ranks
EQ 2 - EQ 1	Negative Ranks	0[b]	.00	.00
	Positive Ranks	10[c]	5.50	55.00
	Ties	10[d]		
	Total	20		

a. genul = feminin
b. EQ 2 < EQ 1
c. EQ 2 > EQ 1
d. EQ 2 = EQ 1

Scorul a crescut la 10 fete.

Test Statistics[a,b]

	EQ 2 - EQ 1
Z	-2.829[c]
Asymp. Sig. (2-tailed)	.005

a. genul = feminin
b. Wilcoxon Signed Ranks Test
c. Based on negative ranks.

Intervenția a avut efect semnificativ la fete (Z/W= -2,829, p≤0,01).

Ranks[a]

		N	Mean Rank	Sum of Ranks
EQ 2 - EQ 1	Negative Ranks	0[b]	.00	.00
	Positive Ranks	12[c]	6.50	78.00
	Ties	8[d]		
	Total	20		

a. genul = masculin
b. EQ 2 < EQ 1
c. EQ 2 > EQ 1
d. EQ 2 = EQ 1

Scorul a crescut la 12 băieți.

Test Statistics[a,b]

	EQ 2 - EQ 1
Z	-3.086[c]
Asymp. Sig. (2-tailed)	.002

a. genul = masculin
b. Wilcoxon Signed Ranks Test
c. Based on negative ranks.

Intervenția a avut efect semnificativ la băieți (Z/W= -3,086, p≤0,01).

 **Exersare:**	**Exercițiul 12.** Comparați scorurile obținute la 3 variabile continue a 2 grupe independente selectând testul potrivit în funcție de tipul de distribuție a datelor.
	Exercițiul 13. Comparați scorurile obținute la o variabilă continuă a 2 grupe dependente selectând testul potrivit în funcție de tipul de distribuție a datelor.

Surse recomandate:

1. CLOCOTICI, V., STAN, A. ***Statistică aplicată în psihologie***. Iași: Polirom, 2000. ISBN 973-683-577-4
2. Compară mediile a două grupuri în SPSS. Disponibil: https://www.youtube.com/watch?v=evkmFrssZ74
3. Interpretarea Testului T pentru eșantioane independente. Disponibil: https://www.youtube.com/watch?v=6WhPpMOZDls
4. OPARIUC-DAN, C. ***Statistica aplicată în științele socio-umane. Analiza asocierilor și diferențelor statistice***. Constanța: Universitatea Ovidius, 2011
5. RACU, Ig. ***Psihodiagnostic și statistică psihologică***. Chișinău: CEP UPS, 2017. ISBN 978-9975-46-321-8.
6. Testul T pentru eșantioane independente. Disponibil: https://www.youtube.com/watch?v=cRt4rE_3eQk&list=PLR3mG_n9dbTVtPr5wbmoU54Huk0WnwUlr&index=2
7. Testul U Mann Whitney. Disponibil: https://www.youtube.com/watch?v=JStVynV6Muk&list=PLR3mG_n9dbTVtPr5wbmoU54Huk0WnwUlr&index=4
8. Testul Wilcoxon. Disponibil: https://www.youtube.com/watch?v=Mqn4TesqDHA&list=PLR3mG_n9dbTVtPr5wbmoU54Huk0WnwUlr&index=5
9. VASILIU, D. ***Momento de cercetare în psihologie: un ghid prietenos pentru studenți și nu numai…*** București: Editura Didactică și pedagogică, 2018. ISBN 978-606-31-0501-2

Tema 6. DIFERENȚE ÎNTRE 3 ȘI MAI MULTE GRUPURI

1. ANOVA unifactorială

Se utilizează atunci când vrem să **comparăm 3 și mai multe grupuri independente**, iar **variabila comparată este la scală de interval / de raport cu o distribuție normală a scorurilor**. De ex. dorim să determinăm dacă diferă EQ la subecții cu diferit nivel al IQ.

Analyze – Compare Means – One-Way ANOVA

În **Dependent List** se transferă variabila dependentă, iar în câmpul **Factor** – se transferă variabila categorială cu mai mult de 2 categorii.

Se deschide caseta **Post Hoc**:

- în **Equal Variances Assumed** se bifează opțiunea **Bonferroni**
- în **Equal Variances Not Assumed** se bifează opțiunea **Tamhane's T2**. Tastați **Continue**

Se deschide caseta **Options**:

- în **Statistics** se bifează **Descriptives**, **Homogeneity-of-variance test, Welch**. Tastați **Continue,** apoi **OK**.

Descriptives

EQ 1

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
sub medie	6	46.33	3.327	1.358	42.84	49.82	43	52
medie	10	44.50	13.134	4.153	35.10	53.90	25	64
peste medie	14	48.36	8.608	2.301	43.39	53.33	31	64
exceptionala	10	44.20	9.004	2.847	37.76	50.64	31	57
Total	40	46.05	9.383	1.484	43.05	49.05	25	64

Test of Homogeneity of Variances

EQ 1

Levene Statistic	df1	df2	Sig.
2.920	3	36	.047

Testul Levene indică dacă dispersiile în interiorul grupurilor sunt sau nu omogene. Odată ce avem p=0,047≤0,05, dispersiile în cele 4 grupuri diferă. Dacă p>0,05, atunci dispersiile în interiorul grupurilor sunt omogene.

ANOVA

EQ 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	133.252	3	44.417	.484	.695
Within Groups	3300.648	36	91.685		
Total	3433.900	39			

Dacă p>0,05, atunci variabila dependentă (în cazul nostru – EQ) nu diferă semnificativ în funcție de categoriile variabilei independente (în cazul nostru – nivelele IQ). Diferențele semnificative sunt doar dacă p≤0,05.

Robust Tests of Equality of Means

EQ 1

	Statistic[a]	df1	df2	Sig.
Welch	.477	3	19.096	.702

a. Asymptotically F distributed.

Testul Welch confirmă inegalitatea (dacă p≤0,05) sau egalitatea (dacă p>0,05) mediilor între grupuri.

Dacă p>0,05, vom consulta rezultatele din dreptul testului Bonferroni din tabelul de mai jos.

Dacă p≤0,05, vom consulta rezultatele din dreptul testului Tamhane din tabelul de mai jos.

Multiple Comparisons

Dependent Variable: EQ 1

	(I) nivel inteligenta	(J) nivel inteligenta	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Bonferroni	sub medie	medie	1.833	4.945	1.000	-11.97	15.64
		peste medie	-2.024	4.672	1.000	-15.07	11.02
		exceptionala	2.133	4.945	1.000	-11.67	15.94
	medie	sub medie	-1.833	4.945	1.000	-15.64	11.97
		peste medie	-3.857	3.965	1.000	-14.93	7.21
		exceptionala	.300	4.282	1.000	-11.66	12.26
	peste medie	sub medie	2.024	4.672	1.000	-11.02	15.07
		medie	3.857	3.965	1.000	-7.21	14.93
		exceptionala	4.157	3.965	1.000	-6.91	15.23
	exceptionala	sub medie	-2.133	4.945	1.000	-15.94	11.67
		medie	-.300	4.282	1.000	-12.26	11.66
		peste medie	-4.157	3.965	1.000	-15.23	6.91
Tamhane	sub medie	medie	1.833	4.370	.999	-12.19	15.85
		peste medie	-2.024	2.671	.975	-9.91	5.87
		exceptionala	2.133	3.155	.986	-7.71	11.98
	medie	sub medie	-1.833	4.370	.999	-15.85	12.19
		peste medie	-3.857	4.748	.966	-18.31	10.59
		exceptionala	.300	5.036	1.000	-14.81	15.41
	peste medie	sub medie	2.024	2.671	.975	-5.87	9.91
		medie	3.857	4.748	.966	-10.59	18.31
		exceptionala	4.157	3.660	.849	-6.58	14.90
	exceptionala	sub medie	-2.133	3.155	.986	-11.98	7.71
		medie	-.300	5.036	1.000	-15.41	14.81
		peste medie	-4.157	3.660	.849	-14.90	6.58

Dacă p≤0,05, există diferențe semnificative între mediile grupurilor, dacă p>0,05, nu există diferențe semnificative între grupuri. În cazul nostru, *rezultatele la obținute la testul de EQ nu diferă semnificativ în dependență de nivelul IQ.*

5. Testul Kruskal-Wallis

Se utilizează atunci când vrem să **comparăm 3 și mai multe grupuri independente, iar variabila comparată este la scală ordinală sau de interval / de raport cu o distribuție ce diferă de cea normală**. De ex. dorim să determinăm dacă diferă IQ la participanții cu diferit nivel al studiilor.

Analyze – Nonparametric Tests – Legacy Dialogs – K-Independent Simples

În **Test Variables Lists** se transferă variabila dependentă (în cazul nostru ***IQ***)

În **Grouping Variable** se transferă variabila independentă (în cazul nostru nivelul studiilor)

În ***Define Range*** scriem valoarea minima și maximă ale variabilei corespunzătoare (în cazul nostru va fi 1 și 3).

Bifăm testul **Kruscal-Wallis**, tasăm **OK.**

Ranks

	studii	N	Mean Rank
IQ	liceu	7	7.86
	colegiu	20	19.48
	universitate	13	28.88
	Total	40	

La variabila IQ media rangurilor subiecților ce au studii de liceu este 7,86, a cu colegiu este 19,48 și a celor cu universitate este de 28,88. Constatăm că cele mai mari scoruri le au cei cu universitate, iar cele mai mici – cei cu liceu. *Sunt însă oare semnificative aceste diferențe?Răspunsul îl obținem din datele următorului tabel.*

Test Statistics[a,b]

	IQ
Chi-Square	15.160
df	2
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: studii

Diferențele semnificative constatăm doar dacă **p**≤0,05. În cazul nostru există diferențe semnificative privind IQ subiecților cu nivel diferit al studiilor: χ^2 =15,16, df=2, p≤0,01.

 **Exersare:**	**Exercițiul 14.** Comparați scorurile obținute la 3 variabile continue a 3 și mai multe grupe independente selectând testul potrivit în funcție de tipul de distribuție a datelor

Surse recomandate:

1. Analiza variației ANOVA. Disponibil: https://www.youtube.com/watch?v=3tGl9CvPJpc
2. CLOCOTICI, V., STAN, A. ***Statistică aplicată în psihologie***. Iași: Polirom, 2000. ISBN 973-683-577-4
3. OPARIUC-DAN, C. ***Statistica aplicată în științele socio-umane. Analiza asocierilor și diferențelor statistice***. Constanța: Universitatea Ovidius, 2011
4. RACU, Ig. ***Psihodiagnostic și statistică psihologică***. Chișinău: CEP UPS, 2017. ISBN 978-9975-46-321-8.
5. Testul One-Way ANOVA. Disponibil: https://www.youtube.com/watch?v=-gMZac8VAsc&list=PLR3mG_n9dbTVtPr5wbmoU54Huk0WnwUlr&index=6
6. V16.6 - Kruskal-Wallis in SPSS. Disponibil: https://www.youtube.com/watch?v=PV-kEFgXrIQ
7. VASILIU, D. ***Momento de cercetare în psihologie: un ghid prietenos pentru studenți și nu numai…*** București: Editura Didactică și pedagogică, 2018. ISBN 978-606-31-0501-2

ВВЕДЕНИЕ

Методическое пособие предназначено для студентов третьего курса, специальность «ПСИХОЛОГИЯ», для организации учебной деятельности в рамках курса «Компьютеризированный статистический анализ», предусмотренный в соответствии с учебным планом на VI семестр обучения.

Специализированный учебный курс направлен на углубление теоретических знаний в области использования статистики в психологии и развитие навыков работы с программой SPSS. Эти навыки необходимы как для исследовательской деятельности, так и для правильного понимания результатов научных психологических исследований. Полученные результаты обучения будут использованы в планировании экспериментальных исследований, в описательном и индуктивном анализе данных в рамках дипломной работы.

Компетенции, формируемые в ходе курса:

CG 2. Анализ данных, касающихся конкретных психологических ситуаций.

CP 6. Адаптация методов/техник психологической помощи к различным контекстам.

По окончании курса, студент сможет:

F-1. Объективно исследовать/изучать данные, информацию, обстоятельства конкретной психологической проблемы/ ситуации, оценивая качество собранных данных и информации для решения проблемы/ситуации.

F-2. Применять эффективные методы психологической оценки в зависимости от контекста, анализируя полученные результаты.

Содержание курса:

Тема 1. Понятия, используемые в статистической обработке данных

Типы переменных, показатели центральной тенденции, показатели дисперсии, показатели типа распределения, типы распределений, характеристики нормального распределения.

Тема 2. Ввод, редактирование и создание новых переменных в SPSS

Копирование данных из табличного редактора в окно «Просмотр данных», редактирование переменных в окне «Просмотр переменных». Сохранение документа, созданного в SPSS. Преобразование балов и создание новых переменных.

Тема 3. Описательный анализ данных

Определение частот номинальных и порядковых переменных и их графическое представление. Описательный анализ количественных переменных. Проверка на нормальность распределения.

Лабораторная работа № 1: Создание базы данных в SPSS. Определение нормальности распределения непрерывных величин. Определение частот зависимых переменных относительно независимой переменной.

Тема 4. Корреляции или ассоциации между переменными

Выбор статистического теста в соответствии с характеристиками анализируемых распределений. Использование параметрических (критерий корреляции Пирсона) и непараметрических статистических методов (критерий ассоциации Кендалла, критерий ассоциации Спирмена, критерий хи-квадрат и точный критерий Фишера). Интерпретация результатов.

Лабораторная работа №2: Корреляционный анализ данных и интерпретация результатов.

Тема 5. Различия между двумя группами

Выбор статистического теста в соответствии с характеристиками анализируемых распределений. Сравнение двух независимых выборок (t-критерий для независимых выборок и U-критерий Манна-Уитни) и двух зависимых выборок (t-критерий для зависимых выборок и критерий Вилкоксона). Интерпретация результатов.

Тема 6. Различия между 3 и более группами

Сравнение 3 и более независимых выборок (Однофакторный дисперсионный анализ и критерий Крускала-Уоллиса). Интерпретация результатов.

Лабораторная работа № 3: Сравнение данных двух или более групп и интерпретация результатов.

Часть I. ОПИСАТЕЛЬНАЯ СТАТИСТИКА

Тема 1. ПОНЯТИЯ, ИСПОЛЬЗУЕМЫЕ В СТАТИСТИЧЕСКОЙ ОБРАБОТКЕ ДАННЫХ

А. Типы шкал:

1. ***Шкала наименований (номинальная, классификационная)*** - наименование класса эквивалентности, к которому принадлежит рассматриваемый объект.

- Пример: пол, этнос, с / без кофе
- Статистика: частота, мода

2. ***Порядковая шкала (или ранговая)*** **–** включает отношения тождества и порядка. Объекты в данной шкале ранжированы. *Биполярные шкалы* - противоположные значения свойств объектов

    - Пример: возрастной период, уровень тревожности...
    - Статистика: частота, медиана, мода

3. ***Интервальная шкала (Шкала разностей)*** – разность (интервал) между двумя значениями имеет эмпирическую значимость.

- В шкале разностей действует отношение «на столько-то больше».
    - Пример: IQ , психо-социальные свойства
    - Статистика: среднее значение, вариация

4. ***Абсолютная шкала (Шкала отношений)*** **–** это интервальная шкала, в которой присутствует дополнительное свойство — естественное и однозначное присутствие нулевой точки.

    - В шкале отношений действует отношение «во столько-то раз больше».
    - Пример: число людей в аудитории, масса, длина, сила
    - Статистика: среднее значение, вариация

 **Тренировка:**	**Упражнение 1.** Классифицируйте на количественные и качественные данные, исходя из ответов полученные на вопросы ниже, и укажите шкалу измерения для каждого из них. а) Сколько вам лет? б) Вы обычно ходите за покупками для своей семьи? в) Сколько человек в вашей семье? г) Как бы вы оценили вкус продукта по шкале от 1 до 10, если 1 соответствует наименее вкусному? д) Купили бы вы этот продукт, если бы он был доступен в магазинах? е) Если ответ на вопрос д) «Да», как часто вы бы покупали этот продукт?

Б. Понятия, используемые при описании распределения

1) **Параметры / показатели центральной тенденции**

- ***Среднее значение признака/ среднее арифметическое (M)*** – сумма всех наблюдаемых значений признака делится на количество наблюдений.

Пример: 2, 3, 5, 6, 7 M=(2+3+5+6+7):5= 23:5 = 4,6

- ***Медиана (Me)*** – делит выборку испытуемых по интересующему нас параметру ровно пополам.

Примеры: а) 2, 3, 5, 6, 7 б) 2, 3, 4 | 5, 6, 7 в) 2, 3 | 5, 6

Me=5 Me=4,5 Me=4

- ***Мода (Mo)*** – наиболее часто встречающееся значение интересующей нас переменной.

Пример: 2, 3, 5, 6, 6, 7 Mo=6 (унимодальное распределение)

Пример: 2, 3, 3, 5, 6, 6, 7 Mo=3 şi Mo=6 (бимодальное распределение)

Пример: 2, 3, 3, 3, 5, 6, 6, 6, 7, 8, 9, 9, 9 Mo=3, Mo=6 şi Mo=9 şi (мультимодальное распределение)

2) Параметры / показатели распределения

- ***Амплитуда ряда / вариации (A)*** – разница между максимальной и минимальной величинами ряда: **A = V max – V min**

Пример: 2, 3, 5, 6, 7 A=7-2=5

- ***Отклонение от среднего значения (d)*** – разница между средним значением и вариантой (индивидуальный результат): **d = V – M.** Сумма всех отклонений от среднего значения равна 0.

Пример: 2, 3, 5, 6, 7 M=4,6;

Отклонения от среднего значения: d_1=2-4,6= - 2.6 d_2=3-4,6= - 1.6

d_3=5-4,6= 0.4 d_4=6-4,6= 1.4 d_5=7-4,6= 2.4

- ***Дисперсия (s^2)*** – среднее значение квадратов отклонения величин от среднего значения: $\boldsymbol{s^2 = \sum d^2/n}$

Пример: s^2 = ((-2,6)²+ (-1,6)²+ (0,4)²+ (1,4)²+ (2,4)²)/5 = (6,76+2,56+0,16+1,96+5,76) / 5= 17,2 / 5= 3,44

- ***Стандартное отклонение (s)*** - квадратный корень из значения дисперсии: $\mathbf{s=\sqrt{s^2}}$

Пример: s=√3,44 = 1,85

Стандартное отклонение — это индикатор, который дает нам информацию о том, насколько ряд данных близок к нормальному распределению. Таким образом, эмпирическое правило гласит:

- 68,37% значений ряда расположены в интервале (M – s , M + s)
- 98,45% значений ряда расположены в интервале (M – 2s , M + 2s)

- 99,73% значений ряда расположены в интервале (M – 3s , M + 3s)
- 99,94% значений ряда расположены в интервале (M – 4s , M + 4s).

- ***Коэффициент вариации (Cv)*** - это процентное отношение стандартного отклонения к среднеарифметической величине: **Cv = s / M x 100%.**

Применяется коэффициент вариации для оценки разнообразия каждого конкретного вариационного ряда и, соответственно, суждения о типичности отдельной средней (т.е. ее способности быть полноценной обобщающей характеристикой данного ряда).

<10% – слабое разнообразие ряда, группа гомогенная

10%-20% – относительно гомогенная группа

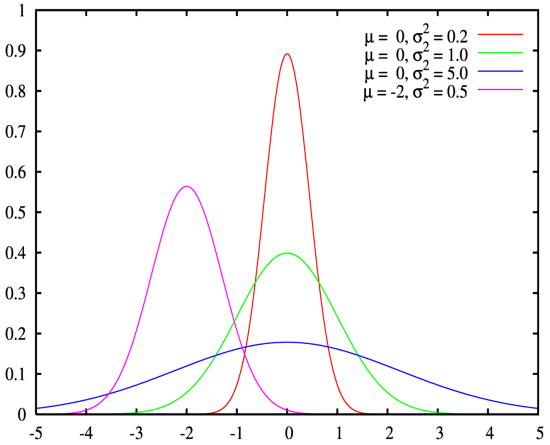
20%-30% – относительно гетерогенная группа

>30% – сильное разнообразие ряда, группа гетерогенная

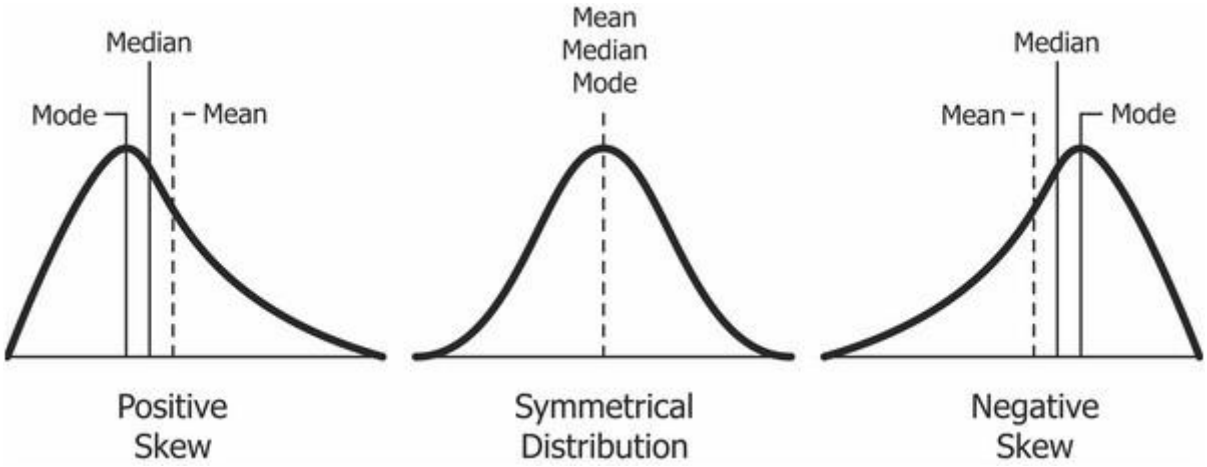
Сильное разнообразие ряда свидетельствует о малой представительности (типичности) соответствующей средней величины и, следовательно, о нецелесообразности ее использования в практических целях.

3) Параметры / показатели вида распределения:

- ***Коэффициент эксцесса (kurtosis)*** — числовая характеризующая степени остроты пика распределения случайной величины:
  - нормальное распределение ≈ 3
  - островершинное распределение >3
  - плосковершинное распределение <3



- ***Коэффициент асимметрии (skewness)***
  - симметричное распределение – от -½ до + ½
  - умеренно асимметричное распределение (скошенное) – от -1 до -½ или от ½ до 1
  - крайне асимметричное распределение – меньше -1 или больше 1.



Для нормального распределения характерно следующее**:**

- оно унимодальное
- кривая распределения частот и значений совершенно симметрична по отношению к средней, т. е. слева и справа от нее лежит 50% вариантов
- нормальная острота пика
- крайние значения (наибольшие и наименьшие) появляются редко, но чем ближе значение признака к центру (к средней арифметической), тем оно чаще встречается
- Это действительно только для непрерывных переменных (для интервальных и абсолютных шкал)
- все три средние совпадают: M=Me=Mo

 **Тренировка:**	**Упражнение 2.** Определите среднее значение, медиану, моду, размах, стандартное отклонение, дисперсию, стандартное отклонениеб коэффициент вариацииб минимальный и максимальный результаты для двух серий данных: **I серия: 1, 4, 5, 2, 2, 3, 4** **M=** **Me=** **Mo=** **A=** **d=** s^2**=** **s=** **Cv=** **I серия: 1, 4, 5, 2, 2, 3, 4, 5, 5** **M=** **Me=** **Mo=** **A=** **d=** s^2**=** **s=** **Cv=**

	Упражнение 3. На основании данных, полученных при выполнении 2-го упражнения для двух серий данных, и характеристик нормального распределения определите тип распределения (нормальное или нет) этих серий. Обоснуйте своё мнение.

Рекомендуемые источники:

1. БАВРИНА, А.П. Основные понятия статистики. Доступно: https://cyberleninka.ru/article/n/osnovnye-ponyatiya-statistiki/viewer
2. КРЕЧЕТОВА, М.А. *Статистика*: Учебное пособие / НФИ КемГУ. Новокузнецк, 2003. Доступно: https://studfile.net/preview/3267558/
3. Описательная статистика (часть 1): ключевые определения за 15 минут. Доступно: https://www.youtube.com/watch?v=4SEX9Bozktw
4. 01-06 Описательные статистики. Доступно: https://www.youtube.com/watch?v=M3JPIAq6Vwg
5. Описательная статистика. Доступно: https://www.youtube.com/watch?v=MdoJ4laEMcA
6. Описательная статистика как базовый навык для написания научной публикации. Коротко о главном. Доступно: https://www.youtube.com/watch?v=Zm8sBDSmU3M&t=1132s
7. Элементарные понятия статистики. Доступно: https://masters.donntu.ru/2001/kita/kaluzhny/lib/esc.html

Тема 2. ВВОД, РЕДАКТИРОВАНИЕ И СОЗДАНИЕ НОВЫХ ПЕРЕМЕННЫХ В SPSS

1. Скопируйте данные из EXCEL в SPSS

2. В редакторе ***VARIABLE VIEW*** заполните данные:

- ***Name*** - введите «пол» (указывает имя переменной в редакторе DATA VIEW)
- ***Type*** - numeric
- ***Decimals*** – введите 0, если нет десятичных
- ***Labe***l - напишите «пол»
- ***Value***s - напишите «1 женский» и «2 мужской»
- ***Measure*** - тип шкалы (nominal - номинальная, ordinal - порядковая, scale – шкала: интервальная или абсолютная)

3. Определите показатели центральной тенденции, показатели распределения и показатели вида распределения для количественных шкал

Analyse – Descriptive Statistics – Frequencies

Central Tehndency: Mean, Median, Mode

Dispersion: Std. deviation, Variance

Distribution: Skewness, Kurtosis

4. Переведите месячные зарплаты в категории «ниже средней» и «выше средней»:

Transform – Recode into Different Variables – введите нужное переменное в окно.

Name: зарплата08, ***Label:*** уровень месячного дохода 08. Нажмите ***Change***

Old an New Vaues – ***Range*** (впишите наименьшее значение), ***through*** (впишите среднее значение)

New Value – впишите „1”, нажмите ***Add***

Old an New Values – ***Range*** (впишите среднее значение+1), ***through*** (впишите наибольшее значение)

New Value – впишите „2”, нажмите ***Add***, после ***Continue*** снизу.

5. Данные переменных „интеграция в группе 08”, „интеграция в группе 09” и „интеграция в группе 10” переведите в уровни в соответствии с шагами, описанными в пункте 4.

Name: УрИнтег08, ***Label:*** уровень интеграции в группе 08.

Уровни: от 1 до 4 - низкий уровень, 5-7 - средний уровень, 8-10 - высокий уровень

6. Создайте переменную СРЕДНЕЕ ЗНАЧЕНИЕ ИНТЕГРАЦИИ:

   Transform – Compute Variable

   Target Variable: СрИнтегр

   Numeric Expression: (интегр08+ интегр09+ интегр10)/3. Нажмите ***OK***

7. Переведите среднее значение интеграции в коллектив в уровни (уровни в п. 5, шаги - в п. 4)

 **Тренировка:**	**Упражнение 4.** Создайте новый .**sav**-документ, используя новую базу данных Excel. Отредактируйте переменные.
	Упражнение 5. Создайте новые переменные, преобразовав непрерывные переменные в порядковые. Отредактируйте новые переменные.
	Упражнение 6. Определите для всех непрерывных переменных следующие показатели: среднее значение, медиана, мода, амплитуду ряда, среднее отклонение, дисперсия, стандартное отклонение, коэффициент вариации, минимальное и максимальное значения.

Рекомендуемые источники:

1. Анализ средних за 2 минуты в SPSS. Доступно: https://www.youtube.com/watch?v=ztWMBDIz7Z4
2. Знакомство со статистической программой IBM SPSS. Доступно: https://www.youtube.com/watch?v=swtTUP0vPPc
3. Как правильно вбить данные в СПСС. Доступно: https://www.youtube.com/watch?v=8Ba1ny5WSR0
4. Подготовка данных к работе в SPSS. Доступно: https://www.youtube.com/watch?v=In9XfFKqAcc
5. Работа с SPSS. Доступно: https://www.youtube.com/watch?v=3e8w16Y1j34&t=3411s

Тема 3. ОПИСАТЕЛЬНЫЙ АНАЛИЗ ДАННЫХ

1. Опишите выборку, определяя частоты по следующим признакам: полу, уровню ежемесячного дохода по годам, наличие санкций по годам, тип отношения к управлению организацией, уровню воспринимаемой автономии по годам, уровню интеграции в коллективе по годам и по уровню среднего значения интеграции:

   Analyse – Descriptive Statistics – Frequencies

2. Определите частоты следующих порядковых переменных ***в зависимости от пола***:
   - уровень ежемесячного дохода по годам,
   - наличие санкций по годам,
   - тип отношения к управлению организацией,
   - уровень воспринимаемой автономии по годам
   - уровень интеграции в коллективе по годам
   - уровень среднего значения интеграции:

   Analyse – Crosstabs – Rows (интересующие переменные) / ***Columns*** – пол

   Cells – Percentages – Row

   Нажмите ***Continue***, после – ***OK***

3. Определите частоты следующих переменных ***в зависимости от наличия санкций в 2008 году***:
   - уровень ежемесячного дохода по годам,
   - тип отношения к управлению организацией,
   - уровень воспринимаемой автономии по годам
   - уровень интеграции в коллективе по годам
   - уровень среднего значения интеграции:

4. Определите частоты переменных ***в зависимости от типа отношения к управлению организацией***

5. Определите тип распределения ежемесячной заработной платы за 2008 и 2009 годы

   Analyse – Descriptive Statistics – Explore

   Plots – вставте галочку ***Normality plots with tests***

   Нажмите ***Continue***, после ***OK***

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IQ	.145	40	.034	.920	40	.008
EQ	.087	40	.200*	.977	40	.568

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Тест Колмогорова-Смирнова сравнивает эмпирическое распределение с теоретическим (нормальным) распределением.

Смотрим на **Sig** (порог значимости). Если значение больше 0,05, это означает, что распределение близко к нормальному, следовательно можем использовать параметрические методы индуктивного анализа.

Если оно меньше 0,05 – имеются достоверные различия между эмпирическим и нормальным распределениями.

В примере, приведенном для переменной IQ, p=0,034, таким образом, распределение индивидуальных баллов для переменной IQ у 40 испытуемых отличается от нормального распределения (кривая Гаусса).

Для переменной EQ p=0,2, распределение баллов по этой переменной существенно не отличается от нормального распределения, следовательно распределение параметрическое.

 **Тренировка:**	**Упражнение 7.** Определите частоты (количество и %) результатов для непрерывных переменных.
	Упражнение 8. Используя опцию Crosstabs, создайте таблицы частот (число и %) путем пересечения двух категориальных переменных.
	Упражнение 9. Используя тест Колмогорова—Смирнова для определения нормальности распределения, определите тип распределения данных для всех непрерывных переменных.

Рекомендуемые источники:

1. Начало работы с SPSS: описательные статистики. Доступно: https://www.youtube.com/watch?v=TqDzZBpzLQc&t=205s https://www.youtube.com/watch?v=TqDzZBpzLQc
2. Описательные статистики в SPSS ч 2 Descriptive Statistics in SPSS. Доступно: https://www.youtube.com/watch?v=HOk4bWxHriU
3. 14. SPSS. Проверка нормальности распределения. Тест Колмогорова-Смирнова. Расчет. Вариант 1. Доступно: https://www.youtube.com/watch?v=Sur_du0ELUA
4. 15. SPSS. Проверка нормальности распределения. Тест Колмогорова-Смирнова. Расчет. Вариант 2. Доступно: https://www.youtube.com/watch?v=gcX8SNUvJ9U
5. 16. SPSS. Проверка нормальности распределения. Тест Колмогорова-Смирнова. Интерпретация. Доступно: https://www.youtube.com/watch?v=GaLbtrF5lAs

Часть II. ИНДУКТИВНАЯ СТАТИСТИКА

Тема 4. КОРРЕЛЯЦИИ ИЛИ АССОЦИАЦИИ (ЗАВИСИМОСТЬ) МЕЖДУ ПЕРЕМЕННЫМИ

Решение о том, какой статистический тест будет использоваться, будет принято в зависимости от типа переменных и характеристик распределения в соответствии с таблицей:

Переменная 1	*Переменная 2*	*Статистический тест*
Количественная	**Количественная**	**r Пирсона** (для нормальных распределений)
		Коэффициент корреляции Кендалла (распределение отличается от нормального)
Порядковая	**Порядковая**	**Коэффициент корреляции Спирмена**
Категориальная (номинальная или порядковая)	**Категориальная (номинальная или порядковая)**	**Хи-квадрат (критерий согласия Пирсона) -** если хотя бы одна переменная имеет 3 или более категорий **Точный тест Фишера** (только если обе переменные дихотомические)

Оценка силы корреляции согласно значению коэффициента корреляции, согласно **Дэвису**:

0,10 – 0,29	**Слабая зависимость / корреляция**
0,30 – 0,49	**Средняя зависимость / корреляция**
0,50 – 0,69	**Высокая зависимость / корреляция**
0,70 – 1	**Очень высокая зависимость / корреляция**

1. Корреляция Пирсона

Analyze – Correlate – Bivariate

Отметить ***Pearson , Two-tailed, Flag significant correlations***

OK

Correlations

		IQ	EQ
IQ	Pearson Correlation	1	-.081
	Sig. (2-tailed)		.620
	N	40	40
EQ	Pearson Correlation	-.081	1
	Sig. (2-tailed)	.620	
	N	40	40

Коэффициент корреляции r = -0,081 (это отрицательная корреляция), а порог значимости p = 0,62. Таким образом, данные, полученные для переменной IQ, не коррелируют с данными, полученными для переменной EQ.

2. Коэффициент корреляции Кендалла

Analyze – Correlate – Bivariate

Отметить ***Kendall's tau-b, Two-tailed, Flag significant correlations***

OK

Correlations

			IQ	EQ
Kendall's tau_b	IQ	Correlation Coefficient	1.000	-.049
		Sig. (2-tailed)	.	.664
		N	40	40
	EQ	Correlation Coefficient	-.049	1.000
		Sig. (2-tailed)	.664	.
		N	40	40

Как читать данные в таблице:

Не выявили значимой связи между рангами данных, полученные для переменной EQ, и рангами, полученными для переменной IQ: Kt = -0,049, p = 0,664.

В исследовании указываем только значения коэффициентов, которые соответствуют статистически значимым корреляциям, т.е. только те у которых p≤0,05.

3. Коэффициент корреляции Спирмена

Analyze – Correlate – Bivariate

Отметить ***Spearman, Two-tailed, Flag significant correlations***

OK

Correlations

			IQ	EQ
Spearman's rho	IQ	Correlation Coefficient	1.000	-.063
		Sig. (2-tailed)	.	.698
		N	40	40
	EQ	Correlation Coefficient	-.063	1.000
		Sig. (2-tailed)	.698	.
		N	40	40

Как читать данные в таблице:

Существует незначительная связь между рангами переменной «IQ» и рангами переменной «EQ» / нет значимой связи (p = 0,698).

4. Хи-квадрат (критерий согласия Пирсона)

Для категориальных переменных: номинальные или порядковые. Например, мы хотим узнать, сильно ли различается количество девочек и мальчиков в зависимости от уровня образования. Тест работает с частотами, а не с индивидуальными баллами. **Хи-квадрат используется когда хотя бы одна переменная имеет 3 или более категорий** (например, переменная «пол» имеет 2 категории - девочки и мальчики; переменная «класс» имеет 3 категории: лицей, колледж и университет).

Analyse – Descriptive Statistics – Crosstabs: Rows (одну переменную) / ***Columns*** (вторую переменную)

– ***Statistics*: *Chi square, Phi and Cramers V.*** Нажмите Continue

– ***Cells*: *Percentages – Row, Columns, Total***

Нажмите ***Continue***, потом **OK**

пол * уровень образования Crosstabulation

			Уровень образования			Всего
			лицей	колледж	университет	
пол	женский	Count	7	10	3	20
		% within genul	35.0%	50.0%	15.0%	100.0%
		% within studii	100.0%	50.0%	23.1%	50.0%
	мужской	Count	0	10	10	20
		% within genul	0.0%	50.0%	50.0%	100.0%
		% within studii	0.0%	50.0%	76.9%	50.0%
Всего		Count	7	20	13	40
		% within genul	17.5%	50.0%	32.5%	100.0%
		% within studii	100.0%	100.0%	100.0%	100.0%

Как читать данные в таблице:

Из 20 девушек 7 (35%) окончили среднюю школу, 10 (50%) - колледж и только 3 (15%) учились в университете.

Из 7 учеников средней школы все девушки (100%), из 20 учеников, окончивших колледж, 50% (10) - девушки, 50% (10) - юноши, из 13 учеников с высшим образованием 23,1% (3) - девушки и 76,9% (10) - юноши.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.769[a]	2	.005
Likelihood Ratio	13.681	2	.001
Linear-by-Linear Association	10.005	1	.002
N of Valid Cases	40		

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is 3.50.

Из приведенной выше таблицы нас интересуют значение коэффициента хи-квадрат и порог значимости.

Как читать данные в таблице:

Исследование, проведенное на 40 испытуемых, подтверждает более высокую частоту юношей в колледже по сравнению с девушками (χ^2 = 10,769, df = 2, p = 0,005≤0,05), что указывает на значимую связь между уровнем образования и полом (более высокая частота девушек в средней школе по сравнению с более высокой частотой юношей в колледже).

Учитывая, что мы получили статистически значимый результат, необходим дополнительный индикатор для обозначения интенсивности связи – это коэффициент Фи Крамера в таблице ниже.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.519	.005
	Cramer's V	.519	.005
N of Valid Cases		40	

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

Как читать данные в таблице: *Частотные различия девочек и мальчиков по уровню образования сильно выражены: φ = 0,519.*

Значения Фи Крамера в следующей таблице:

Phi / φ	0, 01 – 0,24	***низкая надежность***
	0,25 – 0,39	***средняя надежность***
	Больше 0,40	***высокая надежность***

5. Точный тест Фишера

Для категориальных переменных: номинальные или порядковые. Тест применяется, когда **обе переменные являются дихотомическими** (например: переменная «пол» с двумя категориями - девочки и мальчики, переменная «среда обитания» с двумя категориями - городская и сельская). Тест работает с частотами, а не с индивидуальными оценками.

Analyse – Descriptive Statistics – Crosstabs : Rows (одну переменную) / ***Columns*** (вторую переменную)

– ***Exact***: Отметить ***Exact***. Нажмите ***Continue***

– ***Statistics***: Chi square.

Нажмите ***Continue***, потом **OK**

genul * nivel EQ Crosstabulation

			уровень EQ		Всего
			Ниже среднего	Выше среднего	
пол	женский	Count	9	11	20
		% within genul	45.0%	55.0%	100.0%
		% within nivel EQ	47.4%	52.4%	50.0%
	мужской	Count	10	10	20
		% within genul	50.0%	50.0%	100.0%
		% within nivel EQ	52.6%	47.6%	50.0%
Всего		Count	19	21	40
		% within genul	47.5%	52.5%	100.0%
		% within nivel EQ	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.100[a]	1	.752	1.000	.500	
Continuity Correction[b]	.000	1	1.000			
Likelihood Ratio	.100	1	.751	1.000	.500	
Fisher's Exact Test				1.000	.500	
Linear-by-Linear Association	.098[c]	1	.755	1.000	.500	.236
N of Valid Cases	40					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 9.50.
b. Computed only for a 2x2 table
c. The standardized statistic is -.313.

Порог двунаправленной значимости равен 1, порог однонаправленной значимости - 0,5, причем оба значения выше 0,05 - допустимого порога значимости.

Как читать данные в таблице:

Нет различий в уровнях EQ между участниками женского и мужского пола.

Соотношение девочек и мальчиков с EQ ниже среднего не отличается от соотношения девочек и мальчиков с EQ выше среднего.

 **Тренировка:**	**Упражнение 10.** Коррелируйте 3 непрерывные переменные, выбрав соответствующий тест в соответствии с типом распределения данных.
	Упражнение 11. Выявите созависимость 2 категориальных переменных, выбрав соответствующий тест исходя из количества категорий для каждой переменной.

Рекомендуемые источники:

1. Как посчитать коэффициенты корреляции в SPSS. Доступно: https://www.youtube.com/watch?v=bmvi9kp7nQs
2. Корреляционный анализ в SPSS. Доступно: https://www.youtube.com/watch?v=mfD_HRyc4D0
3. Расчет коэффициента корреляции Кендалла в SPSS. Доступно: https://www.youtube.com/watch?v=jSvGOLFz6YU
4. Расчет корреляции Пирсона в SPSS. Доступно: https://www.youtube.com/watch?v=yVPUw8d01Yc
5. Результаты расчета коэффициента корреляции Пирсона в SPSS. Доступно: https://www.youtube.com/watch?v=j622ZpLLbJY
6. Расчет корреляции Спирмена в SPSS. Доступно: https://www.youtube.com/watch?v=Rh9J-wvyw8o
7. Результаты расчета коэффициента корреляции Спирмена в SPSS. Доступно: https://www.youtube.com/watch?v=okU73oCmF-Y
8. 22. SPSS. Равенство (однородность) двух выборок. Критерий Хи-квадрат. Вариант 1. Доступно: https://www.youtube.com/watch?v=wxmMIvWayb8
9. 24. SPSS. Равенство (однородность) двух выборок. Критерий Хи-квадрат. Интерпретация. Доступно: https://www.youtube.com/watch?v=b-YhG_erR38

Тема 5. РАЗЛИЧИЯ МЕЖДУ ДВУМЯ ГРУППАМИ

Решение относительно используемого статистического критерия будет принято в соответствии с указаниями в таблице:

<table>
<tr><th colspan="3">Независимая переменная</th><th>Зависимая переменная</th><th>Статистический критерий</th></tr>
<tr><td rowspan="5">Категориальная (номинальная или порядочная)</td><td rowspan="3">Две категории</td><td rowspan="2">Независимые группы</td><td>Количественная (распределение ближе к нормальному – параметрическое распределение)</td><td>Т-критерий для независимых выборок</td></tr>
<tr><td>Порядковая или количественная (распределение отличается от нормального – непараметрическое распределение)</td><td>U-критерий Манна-Уитни</td></tr>
<tr><td rowspan="2">Зависимые группы</td><td>Количественная (распределение ближе к нормальному – параметрическое распределение)</td><td>Т-критерий для парных выборок</td></tr>
<tr><td>Порядковая или количественная (распределение отличается от нормального – непараметрическое распределение)</td><td>Критерий Вилкоксона</td></tr>
<tr><td rowspan="2">Три и более категорий</td><td rowspan="2">Независимые группы</td><td>Количественная (распределение ближе к нормальному – параметрическое распределение)</td><td>Однофакторный дисперсионный анализ</td></tr>
<tr><td>Порядковая или количественная (распределение отличается от нормального – непараметрическое распределение)</td><td>Критерий Н Крускалла-Уолеса</td></tr>
</table>

1. Т-критерий для независимых выборок

Тест используется для **сравнения результатов двух групп**, полученных по **непрерывным переменным, распределение которых является нормальным**. Например, нас интересует, есть ли различия между баллами учащихся, полученными для составляющих эмпатии, в зависимости от: пола (2 категории - девочки и мальчики), среды проживания (2 категории – городские и сельские).

Analyze – ***Compare Means – Independent-Simple T Test***

В ***Test Variables*** – вводим зависимые переменные

В ***Grouping Variable*** – вводим независимую переменную, после определяем группы нажав на ***Define Group*** и заполнив ***Group 1*** и ***Group 2***

Нажмите ***Continue***, после – **OK**

Group Statistics

	пол	N	Mean	Std. Deviation	Std. Error Mean
Рацион канал	жен	77	3,19	1,113	,127
	муж	63	4,08	1,052	,133
Эмоцион канал	жен	77	3,65	1,355	,154
	муж	63	2,97	1,204	,152
Интуитивн канал	жен	77	3,10	1,294	,147
	муж	63	2,97	1,332	,168
Эмпатия	жен	77	47,58	8,881	1,012
	муж	63	44,13	9,871	1,244

Среднее значение переменной «Рациональный канал эмпатии» в группе девочек составляет 3,19, а в группе мальчиков - 4,08. Похоже, у мальчиков этот канал эмпатии более развит. В то время как у девочек более развит эмоциональный канал эмпатии (среднее значение 3,65), чем у мальчиков (среднее значение 2,97). Ответ на вопрос *Являются ли эти различия статистически значимыми?* ищем в следующей таблице:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Рацион канал	Equal variances assumed	,126	,723	-4,796	138	,000	-,885	,184	-1,249	-,520
	Equal variances not assumed			-4,823	135,104	,000	-,885	,183	-1,247	-,522
Эмоцион канал	Equal variances assumed	1,698	,195	3,109	138	,002	,681	,219	,248	1,114
	Equal variances not assumed			3,146	137,023	,002	,681	,216	,253	1,109
Интуитивн канал	Equal variances assumed	,158	,691	,609	138	,543	,136	,223	-,305	,576
	Equal variances not assumed			,607	130,995	,545	,136	,223	-,306	,577
Эмпатия	Equal variances assumed	,817	,368	2,179	138	,031	3,457	1,587	,320	6,594
	Equal variances not assumed			2,156	126,174	,033	3,457	1,603	,284	6,631

Для каждой переменной есть два значения t-критерия, например, для переменной рациональный канал -4,796 и -4,823. Какое значение указываем?

Чтобы принять решение, анализируем **p** из третьего столбца – критерий Ливиня указывает, являются ли дисперсии двух распределений равными или нет (в нашем случае, если распределение данных в группе девочек равно распределению данных в группе мальчиков).

Если p≤0,05 означает, что имеются существенные различия в дисперсиях 2 распределений, тогда берем значение t из строки 2, где написано: ***Equal variances not***

assumed / ***Равенство дисперсий не предполагается*** (для рационального канала это будет t = -4,823).

Как читать данные в таблице:

Были выявлены существенные гендерные различия по следующим переменным:

- рациональный канал (t = -4,796, p≤0,01), этот канал эмпатии более развит у мальчиков (среднее значение 4,08), чем у девочек (среднее значение 3,19).

- эмоциональный канал …

или …

- Средний балл, полученный по переменной «рациональный канал эмпатии», значительно выше у мальчиков (среднее значение 4,08), чем у девочек (среднее значение 3,19): t = -4,796, p≤0,01

2. U-критерий Манна-Уитни

Тест используется для **сравнения результатов двух групп, полученных по порядковым или непрерывным переменным, распределения которых являются непараметрическими** (отличаются от нормального распределения).

Analyze – ***Nonparametric Tests*** – ***Legacy Dialogs*** – ***2 Independent Simples***

В ***Test Variables Lists*** – вводим зависимые переменные

В ***Grouping Variable*** – вводим независимую переменную,

после определяем группы нажав на ***Define Group*** и заполнив ***Group 1*** и ***Group 2***

Ставим галочку перед ***U Mann-Whitney*** и нажимаем на ***Continue***, после – **OK**

Ranks

	пол	N	Mean Rank	Sum of Ranks
Рацион канал	ж	77	56,60	4358,50
	м	63	87,48	5511,50
	Всего	140		
Эмоцион канал	ж	77	81,10	6244,50
	м	63	57,55	3625,50
	Всего	140		
Интуитивн канал	ж	77	72,34	5570,50
	м	63	68,25	4299,50
	Всего	140		

Средний ранг переменной «Рациональный канал эмпатии» в группе девочек составляет 56,6, а в группе мальчиков - 87,48. Следовательно, этот канал эмпатии более развит у мальчиков. *Значимы/достоверны ли эти различия?*

Test Statistics[a]

	C rational	C emotional	C intuitiv
Mann-Whitney U	1355,500	1609,500	2283,500
Wilcoxon W	4358,500	3625,500	4299,500
Z	-4,625	-3,509	-,610
Asymp. Sig. (2-tailed)	,000	,000	,542

a. Grouping Variable: пол

Как читать данные в таблице:

Мальчики (средний ранг 87,48) превосходят девочек (средний ранг 56,6) по уровню «рационального канала эмпатии»: U=1355,5, N1=63, N2=77, p≤0,01.

3. Т-критерий для зависимых (парных) выборок

Критерий используется для выявления различий результатов по одной и той же методике, проведенной на той же выборке (тестирование и повторное тестирование). Только для количественных переменных с параметрическими распределениями!!!

Analyze – Compare Means – Paired-Simple T Test

В ***Paired Variables:*** вводим результат тестирования в ***Variable 1*** (например, EQ 1), а результат повторного тестирования вводим в ***Variable 2*** (например, EQ 2).

Нажимаем **OK.**

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Эмпатия 1	46,03	140	9,464	,800
	Эмпатия 2	46,57	140	8,981	,759

Среднее значение результатов тестирования эмпатии 46,03, среднее значение результатов повторного тестирования эмпатии – 46,57.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Эмпатия 1 - Эмпатия 2	-,543	1,723	,146	-,831	-,255	-3,727	139	,000

Как читать данные в таблице:

У участвующих в тренинге по развитию эмпатии результаты повторного тестирования (m = 46,57, s=8,981) значительно выше (t= - 3,727, df=139, p≤0,01) результатов начального тестирования (m = 46,03, s=9,464).

df – степень свободы

4. Критерий Вилкоксона

Критерий используется для выявления различий результатов по одной и той же методике, проведенной на той же выборке – тестирование и повторное тестирование. Для порядковых переменных или количественных переменных с непараметрическими распределениями!!!

Analyze – ***Nonparametric Tests*** – ***Legacy Dialogs*** – ***2 Related Simples***

В ***Test Pair***: вводим результат тестирования в ***Variable 1*** (например, EQ 1), а результат повторного тестирования вводим в ***Variable 2*** (например, EQ 2).

В ***Test Type*** ставим галочку перед ***Wilcoxon***.

Нажимаем OK.

Ranks

		N	Mean Rank	Sum of Ranks
Эмпатия 2 - Эмпатия 1	Negative Ranks	7[a]	11,50	80,50
	Positive Ranks	26[b]	18,48	480,50
	Ties	107[c]		
	Total	140		

a. Эмпатия 2 < Эмпатия 1

b. Эмпатия 2 > Эмпатия 1

c. Эмпатия 2 = Эмпатия 1

У 7 испытуемых показатели снизились после тренинга, у 26 - увеличились, у 107 субъектов показатели эмпатии остались прежними.

Test Statistics[a]

	Empatie 2 - Empatie 1
Z	-3,603[b]
Asymp. Sig. (2-tailed)	,000

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Как читать данные в таблице:

Тренинг способствовал развитию эмпатии (Z/W= -3,603, p≤0,01).

 **Тренировка:**	**Упражнение 12.** Сравните результаты двух независимых групп, полученные по трем непрерывным переменным, выбрав соответствующий тест в зависимости от типа распределения данных.

	Упражнение 13. Сравните результаты двух зависимых групп, полученные по непрерывной переменной, выбрав соответствующий тест в зависимости от типа распределения данных.

Рекомендуемые источники:

1. Расчет критерия Т Стьюдента для независимых выборок в SPSS. Доступно: https://www.youtube.com/watch?v=bQ4ER_KWZ8k
2. Результаты расчета критерия Т Стьюдента для независимых выборок в SPSS. Доступно: https://www.youtube.com/watch?v=NKJ9-xKG7ms
3. 30. SPSS. Значимость различий. U- Критерий Манна-Уитни. Расчет. Доступно: https://www.youtube.com/watch?v=6mAhuzFEssQ
4. 32. SPSS. Значимость различий. U- Критерий Манна-Уитни. Интерпретация. Доступно: https://www.youtube.com/watch?v=jmNghwb6Cz0
5. SPSS. Критерий Манна-Уитни на констатирующем этапе. Две независимые выборки. Доступно: https://www.youtube.com/watch?v=p4y6PzobWRQ
6. Расчет критерия Т Стьюдента для зависимых выборок в SPSS. Доступно: https://www.youtube.com/watch?v=vDxdslY3850
7. Расчет критерия Т Вилкоксона в SPSS. Доступно: https://www.youtube.com/watch?v=2OPE3OEZEgs
8. Результаты расчета критерия Т Вилкоксона в SPSS. Доступно: https://www.youtube.com/watch?v=6rCnWDEhsGo

Тема 6. РАЗЛИЧИЯ МЕЖДУ 3-мя И БОЛЕЕ ГРУППАМИ

1. Однофакторный дисперсионный анализ (ANOVA)

Используется для сравнения результатов 3-х или более групп, только если распределения количественных переменных являются параметрическими.

Analyze – Compare Means – One-Way ANOVA

В ***Dependent List*** вводим зависимую переменную, в ***Factor*** – вводим независимую переменную с 3 (и более) категориями.

В ***Post Hoc***:

- в ***Equal Variances Assumed*** отмечаем **Bonferroni**
- в ***Equal Variances Not Assumed*** отмечаем **Tamhane's T2**.

Нажимаем ***Continue***

В ***Options***:

- в ***Statistics*** отмечаем ***Descriptives***, ***Homogeneity-of-variance test***, ***Welch***,
- ***Means Plots***ю

Нажимаем ***Continue***, после – ***OK.***

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Рацион канал	1,035	2	137	,358
Эмоцион канал	1,419	2	137	,245
Интуитивн канал	5,191	2	137	,007
Эмпатия	,218	2	137	,805

В этой таблице представлены результаты теста Ливиня, сравнивающего дисперсии в группах (в нашем случае в 3 классах). Если p>0,05, дисперсии внутри групп являются однородными. Если p≤0,05, тогда дисперсии внутри групп не являются однородными.

Таким образом, дисперсии переменных ***рациональный канал эмпатии***, ***эмоциональный канал эмпатии*** и ***эмпатия 1*** являются однородными во всех 3х группах, но дисперсии переменной интуитивный канал различаются между группами (p=0,007).

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Рацион канал	Between Groups	4,978	2	2,489	1,845	,162
	Within Groups	184,815	137	1,349		
	Total	189,793	139			
Эмоцион канал	Between Groups	6,743	2	3,371	1,934	,148
	Within Groups	238,800	137	1,743		
	Total	245,543	139			
Интуитивн канал	Between Groups	2,143	2	1,071	,623	,538
	Within Groups	235,600	137	1,720		
	Total	237,743	139			
Эмпатия	Between Groups	70,406	2	35,203	,390	,678
	Within Groups	12379,480	137	90,361		
	Total	12449,886	139			

Если p>0,05, нет различий между группами. Статистические различия есть только при p≤0,05.

Как читать данные в таблице:

Нет существенных различий между классами (9, 10 и 12) по эмпатии и ее составляющим.

Robust Tests of Equality of Means

		Statistic[a]	df1	df2	Sig.
Рацион канал	Welch	1,679	2	86,787	,193
Эмоцион канал	Welch	2,145	2	90,020	,123
Интуитивн канал	Welch	,772	2	83,993	,466
Эмпатия	Welch	,381	2	88,508	,684

a. Asymptotically F distributed.

Тест Велча подтверждает неравенство (если p≤0,05) или равенство (если p>0,05) средних значений в группах.

При p>0,05, смотрим различия по критерию Бонферрони, при p≤0,05 смотрим различия по критерию Тамхейн.

Multiple Comparisons

Dependent Variable		(I) класс	J) класс	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Рацион канал	Bonferroni	9	10	,465	,246	,184	-,13	1,06
			12	,185	,246	1,000	-,41	,78
		10	9	-,465	,246	,184	-1,06	,13
			12	-,280	,232	,690	-,84	,28
		12	9	-,185	,246	1,000	-,78	,41
			10	,280	,232	,690	-,28	,84
	Тамхейн	9	10	,465	,262	,219	-,17	1,10
			12	,185	,250	,845	-,43	,80
		10	9	-,465	,262	,219	-1,10	,17
			12	-,280	,222	,508	-,82	,26
		12	9	-,185	,250	,845	-,80	,43
			10	,280	,222	,508	-,26	,82
Эмоцион канал	Bonferroni	9	10	-,520	,280	,197	-1,20	,16
			12	-,440	,280	,355	-1,12	,24
		10	9	,520	,280	,197	-,16	1,20
			12	,080	,264	1,000	-,56	,72
		12	9	,440	,280	,355	-,24	1,12
			10	-,080	,264	1,000	-,72	,56
	Tamhane	9	10	-,520	,265	,150	-1,16	,12
			12	-,440	,280	,316	-1,12	,24
		10	9	,520	,265	,150	-,12	1,16
			12	,080	,272	,988	-,58	,74
		12	9	,440	,280	,316	-,24	1,12
			10	-,080	,272	,988	-,74	,58
Интуитивн канал	Bonferroni	9	10	,220	,278	1,000	-,45	,89
			12	-,060	,278	1,000	-,73	,61
		10	9	-,220	,278	1,000	-,89	,45
			12	-,280	,262	,863	-,92	,36
		12	9	,060	,278	1,000	-,61	,73
			10	,280	,262	,863	-,36	,92
	Tamhane	9	10	,220	,307	,856	-,53	,97
			12	-,060	,307	,996	-,81	,69

		10	9	-,220	,307	,856	-,97	,53
			12	-,280	,230	,538	-,84	,28
		12	9	,060	,307	,996	-,69	,81
			10	,280	,230	,538	-,28	,84
Эмпатия	Bonferroni	9	10	-1,480	2,016	1,000	-6,37	3,41
			12	,000	2,016	1,000	-4,89	4,89
		10	9	1,480	2,016	1,000	-3,41	6,37
			12	1,480	1,901	1,000	-3,13	6,09
		12	9	,000	2,016	1,000	-4,89	4,89
			10	-1,480	1,901	1,000	-6,09	3,13
	Tamhane	9	10	-1,480	2,061	,855	-6,50	3,54
			12	,000	2,011	1,000	-4,90	4,90
		10	9	1,480	2,061	,855	-3,54	6,50
			12	1,480	1,882	,818	-3,09	6,05
		12	9	,000	2,011	1,000	-4,90	4,90
			10	-1,480	1,882	,818	-6,05	3,09

Если p≤0,05, средние значения групп статистические различаются, если p>0,05, нет достоверных различий.

2. Критерий Н Крускалла-Уолеса

Используется для сравнения результатов 3-х или более групп для порядковых переменных или количественных переменных, если распределения являются непараметрическими.

Analyze – ***Nonparametric Tests*** – ***Legacy Dialogs*** – ***K-Independent Simples***

В ***Test Variables Lists*** вводим зависимые переменные.

В ***Grouping Variable*** вводим независимую переменную.

В ***Define Range*** вставляем минимальное и максимальное значения. Нажимаем ***Continue***,

Отмечаем ***Kruscal-Wallis.*** Нажимаем ***OK.***

Ranks

	класс	N	Mean Rank
Рацион канал	9	40	78,21
	10	50	63,42
	12	50	71,41
	Всего	140	
Эмоцион канал	9	40	60,24
	10	50	75,69
	12	50	73,52
	Всего	140	
Интуитивн канал	9	40	70,08
	10	50	65,70
	12	50	75,64
	Всего	140	
Эмпатия	9	40	68,78
	10	50	73,55
	12	50	68,83
	Всего	140	

Test Statistics[a,b]

	Рацион канал	Эмоцион канал	Интуитивн канал	Установки
Chi-Square	3,189	3,856	1,588	6,714
df	2	2	2	2
Asymp. Sig.	,203	,145	,452	,035

a. Kruskal Wallis Test
b. Grouping Variable: класс

Только при $p \leq 0{,}05$ есть достоверные различия.

Как читать данные в таблице:

Есть достоверные различия между классами только по переменной «установки способствующие эмпатии»: $\chi^2 = 6{,}714$, $df = 2$, $p = 0{,}035$.

 **Тренировка:**	**Упражнение 14.** Сравните результаты 3-х и более независимых групп, полученные по 3 непрерывным переменным, выбрав соответствующий тест в зависимости от типа распределения данных.

Рекомендуемые источники:

1. Дисперсионный анализ в SPSS. Доступно: https://www.youtube.com/watch?v=jL7DzzhlZV8
2. 13. Факторный анализ в программе SPSS. Интерпретация результатов. Доступно: https://www.youtube.com/watch?v=i2sKVcEYEyw
3. Расчет критерия Н Крускалла Уоллеса в SPSS. Доступно: https://www.youtube.com/watch?v=t61zx9nu7AI
4. Результаты расчета критерия Н Крускалла Уоллеса в SPSS. Доступно: https://www.youtube.com/watch?v=f6GLwfwBVfs

INTRODUCTION

This methodological guide is intended for students in the final year of the PSYCHOLOGY study program to facilitate their learning activity within the Computerized Statistical Analysis course provided for the VI semester in accordance with the study plan.

The specialized course has as its primary objective the deepening of theoretical knowledge regarding the use of statistics in the field of psychology and the development of the skills necessary for operating the SPSS software. These competencies are essential both for conducting research activities and for the accurate interpretation of the results of scientific studies during routine documentation. The learning outcomes will be applied in the planning of experimental research, in the descriptive and inferential analysis of the bachelor's thesis data.

Skills developed during the course:

CG 2. Analysis of data pertaining to concrete situations of a psychological nature.

CP 6. Adaptation of psychological assistance methods and techniques to diverse contexts.

Study results:

Upon completion of the course, the student will be able to:

F-1. Objectively examine and investigate the data, information, and circumstances related to concrete problems or situations of a psychological nature, assessing the quality of the data and information gathered for addressing the issue or situation.

F-2. Apply effective methods of psychological assessment according to the specific context, analyzing the results obtained.

The content of the course units

Topic 1. Concepts used in statistical processing

Types of variables, indicators of central tendency, indicators of dispersion, indicators of distribution type, types of distributions, characteristics of the normal distribution.

Topic 2. Data entry, editing, and creation in SPSS

Copying data from a spreadsheet editor into the *Data View* window, editing variables in the *Variable View* window. Saving the document created in SPSS. Transforming scores and creating new variables.

Topic 3. Descriptive data analysis

Determining the frequencies of nominal and ordinal variables and their graphical representation. Descriptive analysis of scale-type variables. Normality test.

Laboratory Work No. 1: Identifying the normality of the distribution. Determining the frequencies of dependent variables in relation to an independent variable.

Topic 4. Correlations or associations between variables

Selecting the statistical test according to the characteristics of the analyzed distributions. Using parametric statistical methods (Pearson correlation test) and nonparametric methods (Kendall's tau test, Spearman's rank correlation test, Chi-square test, and Fisher's exact test). Interpreting the data.

Laboratory Work No. 2: Correlating/associating data and interpreting the results.

Topic 5. Differences between two groups

Selecting the statistical test according to the characteristics of the analyzed distributions. Comparing two independent samples (independent samples t-test and Mann-Whitney U test) and two dependent samples (paired samples t-test and Wilcoxon signed-rank test). Interpreting the data.

Topic 6. Differences among three or more groups

Comparing three or more independent groups (one-way ANOVA and Kruskal-Wallis test). Interpreting the data.

Laboratory Work No. 3: Comparing data from two or more groups and interpreting the results.

Part I. DESCRIPTIVE STATISTICS

Topic 1. TERMS AND CONCEPTS USED IN DESCRIBING A DISTRIBUTION

A. Types of variables

1. Nominal variables

- Classify different cases into the predefined categories of a variable
- Example: gender, age, religion, with/without coffee
- Statistics: frequency of each category, mode

2. Ordinal variables

- Categories ordered from "lower" to "higher", according to the degree to which they possess the measured trait
- Example: educational level, level of intensity of a trait (low / medium / high)
- Distances between categories are not necessarily equal
- Statistics: frequency, mode
- The position in the hierarchy for each subject = rank

3. Interval scale variables

- Subjects can take scores that are equally spaced
- Answer questions like: What is the difference between...? How much more...?
- Example: year of birth, score on a psycho-social assessment tool (EQ, IQ, anxiety, aggression, etc.)
- Zero is set arbitrarily, not absolute (e.g., Celsius/Kelvin temperature)
- Therefore, it does not allow calculation of ratios
- Allows calculation of mean and variance — parametric statistical tests

4. Ratio scale variables

- Have a natural zero point — reflects the absence of that characteristic
- Answer the question: How many times more...?
- Example: number of years of schooling, reaction time, adrenaline level
- Allows calculation of ratios – comparisons between elements

 **Training:**	**Exercise 1** Classify the data provided by the answers as quantitative or qualitative and indicate the measurement scale for each one. a) What is your age? b) Are you the person who usually does the shopping for your family? c) How many people are in your family? d) How would you rate the taste of the product on a scale from 1 to 10, where 1 means least tasty? e) Will you buy this product if it is available in stores? f) If your answer to (e) is "Yes", how often will you buy the product?

B. Concepts used in describing a distribution

1) Indicators of central tendency

- ***Mean ("m" or "M")*** **–** the arithmetic mean of a group, calculated by summing the individual scores of all subjects in the group and dividing the total by the number of subjects in the group.

  Example: 2, 3, 5, 6, 7 M=(2+3+5+6+7):5= 23:5 = 4,6

- ***Median ("Me")*** **–** the central score, the midpoint of the set of scores that divides it into two equal parts.

  – For an odd number of scores **–** it is a whole number (e.g., 2, 3, **5**, 6, 7 → Me = 5)

  – For an even number of scores **–** it falls between two values.

  Example: 2, 3, 4 /5, 6, 7 Me=4,5 or *Ex: 2, 3 /5, 6 Me=4*

- ***Mode ("Mo")*** **–** the score with the highest frequency.

  Example: 2, 3, 5, 6, 6, 7 → Mo = 6 (unimodal distribution)

  Example: 2, 3, 3, 5, 6, 6, 7 → Mo = 3 and Mo = 6 (bimodal distribution)

  Example: 2, 3, 3, 3, 5, 6, 6, 6, 7, 8, 9, 9, 9 → Mo = 3, Mo = 6, and Mo = 9 (multimodal distribution)

2) Indicators of dispersion

- ***Range ("A")*** **–** the difference between the maximum and minimum scores in the distribution: **A = V max – V min**

  Example: 2, 3, 5, 6, 7 → A = 7 − 2 = 5

- ***Mean deviation ("d") –*** shows how much the individual scores deviate from the mean: **d ind= V– M**. The sum of the mean deviations is always equal to zero.

  Example: 2, 3, 5, 6, 7 → M = 4.6

  Individual mean deviations:

  $d_1 = 2 - 4.6 = -2.6$

  $d_2 = 3 - 4.6 = -1.6$

  $d_3 = 5 - 4.6 = 0.4$

  $d_4 = 6 - 4.6 = 1.4$

  $d_5 = 7 - 4.6 = 2.4$

The mean deviation of a series is calculated as: $\mathbf{d = \sum |V - M| / n}$.

Add up all deviations disregarding sign and divide by n: 8.4 / 5 = 1.68 (mean deviation of the dataset).

- ***Variance ("s²") –*** the mean of the squared differences between individual scores and the arithmetic mean (mean of squared deviations): $s^2 = \sum d^2 / n$

  Example: $s^2 = ((-2,6)^2 + (-1,6)^2 + (0,4)^2 + (1,4)^2 + (2,4)^2)/5 = (6,76+2,56+0,16+1,96+5,76) / 5 = 17,2 / 5 = 3,44$

- ***Standard deviation ("s") –*** the square root of the variance: $\mathbf{s=\sqrt{s^2}}$

  Example: $s=\sqrt{3,44} = 1,85$

The standard deviation shows how much the values of a series tend to spread around the mean. An empirical rule states that:

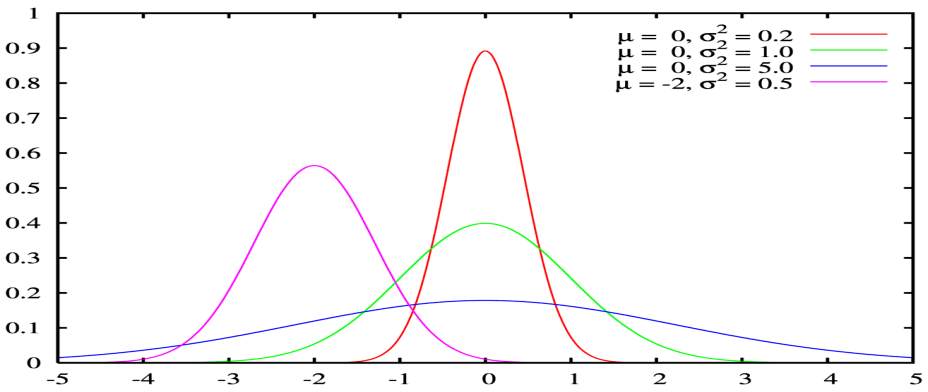
- 68.37% of the values lie within the interval (M − s , M + s)
- 98.45% of the values lie within the interval (M − 2s , M + 2s)
- 99.73% of the values lie within the interval (M − 3s , M + 3s)
- 99.94% of the values lie within the interval (M − 4s , M + 4s)

- ***Coefficient of variation ("CV") –*** shows the degree of dispersion of scores around the mean, calculated by dividing the standard deviation by the mean and multiplying by 100%: $CV=s/M \cdot 100\%$ *(ex: 1,85 / 4,6 x 100% =40,2%)*
  - Values below 10% indicate a homogeneous population
  - Values between 10%–20% indicate a relatively homogeneous population
  - Values between 20%–30% indicate a relatively heterogeneous population
  - Values above 30% indicate a heterogeneous population.

3) Indicators of distribution shape

- ***Kurtosis –*** indicates how peaked or flat a distribution is:
  - Normal distribution: kurtosis = 3, excess kurtosis = 0

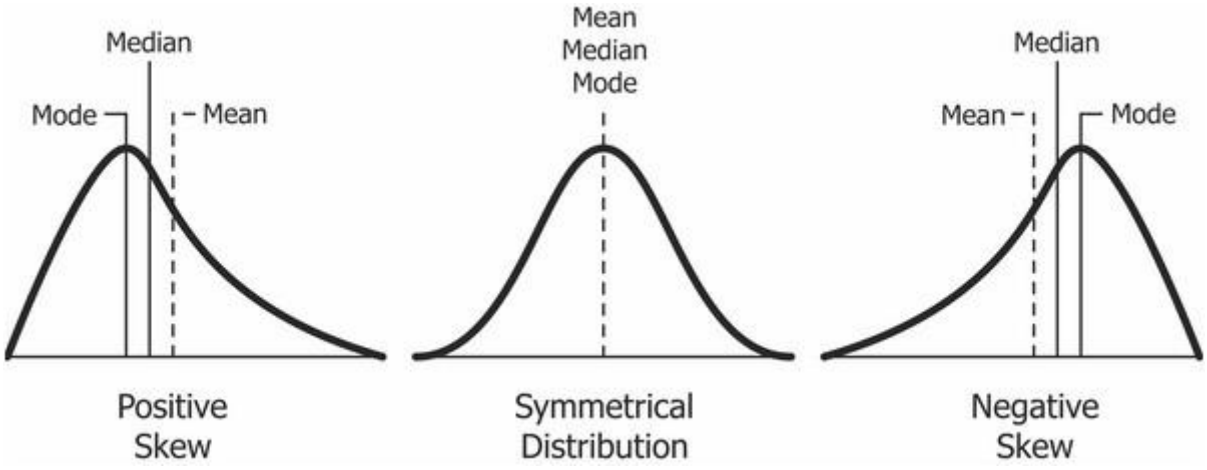
- Leptokurtic distribution (peaked): kurtosis > 3, excess kurtosis > 0 (red)
- Platykurtic distribution (flat): kurtosis < 3, excess kurtosis < 0 (blue)



- ***Skewness***
  - Symmetric distribution **–** from −½ to +½
  - Moderately asymmetric distribution **–** between −1 and −½ or between +½ and +1
  - Strongly asymmetric distribution **–** less than −1 or greater than +1

Left (positive) skewness **–** lower values than the mean predominate

Right (negative) skewness **–** higher values than the mean predominate



Characteristics of the normal distribution:

– It is unimodal

– Symmetric around the mean

– Normal kurtosis

– The left and right tails approach zero

– It applies only to continuous variables

– The measures of central tendency have the same value: M = Me = Mo



Training:

Exercise 2.

Determine the mean, median, mode, range, mean deviation, variance, standard deviation, and coefficient of variation for the given sets of scores:

Set I: 1, 4, 5, 2, 2, 3, 4	**Set II: 1, 4, 5, 2, 2, 3, 4, 5, 5**
M=	**M=**
Me=	**Me=**
Mo=	**Mo=**
A=	**A=**
d=	**d=**
s^2**=**	s^2**=**
s=	**s=**
Cv=	**Cv=**

Exercise 3.

Based on the data obtained in Exercise 1 for the two sets of scores and the characteristics of a normal distribution, determine the type of distribution (normal or not) for each set of scores. Justify your answer.

Recommended sources:

1. Descriptive Statistics: Definition, Overview, Types, and Examples. Available: https://www.investopedia.com/terms/d/descriptive_statistics.asp
2. Descriptive Statistics: FULL Tutorial - Mean, Median, Mode, Variance & SD (With Examples). Available: https://www.youtube.com/watch?v=SplCk-t1BeA
3. Descriptive Statistics [Simply explained]. Available: https://www.youtube.com/watch?v=FzujIYo9GYo
4. Introduction to statistics. Available: https://uniskills.library.curtin.edu.au/numeracy/statistics/introduction/
5. Normal Distribution | Examples, Formulas, & Uses. Available: https://www.scribbr.com/statistics/normal-distribution/
6. Normal Distributions Explained – With Real-World Examples. Available: https://www.youtube.com/watch?v=xlxaa9YhT6A
7. The Normal Distribution, Clearly Explained!!!. Available: https://www.youtube.com/watch?v=rzFX5NWojp0
8. OTT, R.L., LONGNECKER, M. *An Introduction to Statistical Methods and Data Analysis*. Texas A&M University. 5-th ed. ISBN 0-543-25122-6. Available: https://eclass.uoa.gr/modules/document/file.php/ECD433/%CE%91%CE%BD%CE%AC%CE%BB%CF%85%CF%83%CE%B7%20%CE%B4%CE%B5%CE%B4%CE%BF%CE%BC%CE%AD%CE%BD%CF%89%CE%BD%20%CE%BC%CE%B5%20%CF%84%CE%BF%20SPSS/%CE%92%CE%B9%CE%B2%CE%BB%CE%B9%CE%BF%CE%B3%CF%81%CE%B1%CF%86%CE%AF%CE%B1/Statistical%20Methods%20and%20Data%20Analysis%205Ed.pdf
9. What is Descriptive Statistics? A Beginner's Guide to Descriptive Statistics! Available: https://www.youtube.com/watch?v=jfa8hDU6Mz8

Topic 2. DATA ENTRY, EDITING, AND CREATION IN SPSS

1. Copy the data from Excel into SPSS.
2. In the VARIABLE VIEW editor, fill in the following information:
   - **Name –** enter a short name for the variable without spaces (this will display as the variable name in DATA VIEW)
   - **Type –** numeric
   - **Decimals –** enter 0 if you have whole numbers or 2 if you have numbers with decimals
   - **Label –** enter the name of the variable
   - **Values –** enter codes only for nominal variables
   - **Align –** center
   - **Measure –** the type of variable (nominal, ordinal, scale — either interval or ratio)
3. Determine the indicators of central tendency, dispersion, and distribution shape for the interval or ratio scale variables:

   Analyse – Descriptive Statistics – Frequencies

   Statistics:

   Central Tehndency: ***Mean, Median, Mode***

   Dispersion: ***Std. deviation, Variance***

   Distribution: ***Skewness, Kurtosis***
4. Transform the monthly income into the categories "below average" and "above average":

   Transform → Recode into Different Variables (move the required variable into the window).

   Name: monthly income 08, **Label**: monthly income level 08

   Old and New Values – Range (enter the minimum value) **through** (enter the mean value)

   New Value — enter “1”, then click **Add**

   Old and New Values – Range (enter mean value + 1) **through** (enter maximum value)

   New Value — enter “2”, then click **Add**, and then **Continue** below.

   On the **Data View** page, edit the **Label** for the newly created variable.
5. Transform (following the steps from point 4) the scores from the variable “integration into the group” into levels according to the following intervals:

   1–4 — low level

   5–7 — medium level

   8–10 — high level

   Name the new variables as follows: **Name**: NivInt08, **Label**: integration level 08.
6. Create the variable AVERAGE INTEGRATION OVER 3 YEARS:

   Transform – Compute Variable

   Target Variable: MedInt

Numeric Expression: (integr 08+ integr 09 + integr 10)/3. Click the button **OK**

7. Transform the "average integration" into levels (use the levels from point 5 and the steps from point 4).

 **Training:**	**Exercise 4.** Create a new **.sav** document, using a new database in excel. Edit the variables.
	Exercise 5. Create new variables by converting continuous variables to ordinal variables. Edit the new variables.
	Exercise 6. Determine for all continuous variables the following indicators: mean, median, mode, range, mean deviation, dispersion, standard deviation, coefficient of variation, minimum and maximum value.

Recommended sources:

1. How to Enter Data in SPSS for Beginners | Quick Tutorial. Available: https://www.youtube.com/watch?v=oGZmaDstCLM
2. 01 How to Use SPSS - An Introduction to SPSS for Beginners. Available: https://www.youtube.com/watch?v=_zFBUfZEBWQ&list=PLVI_iGT5ZuRkk2d-kePUmlIHD5pmOuquN
3. Learn basics of SPSS in 10 minutes | Data Entry in SPSS | Scale, ordinal, Nominal Variables |. Available: https://www.youtube.com/watch?v=Wp2eUfLlusk

Topic 3. DESCRIPTIVE DATA ANALYSIS

1. Determine the frequencies according to gender, monthly income level by year, presence of sanctions by year, attitude toward management, level of perceived autonomy, and level of group integration by year:

   Analyse – Descriptive Statistics – Frequencies

2. Determine the frequencies by gender for the following variables:
   - monthly income level by year,
   - presence of sanctions by year,
   - attitude toward management,
   - level of perceived autonomy,
   - level of group integration by year.

   Analyse – Descriptive Statistics – Crosstabs – Rows (the variables we are interested in) / **Columns** (Independent variable – e.g., gender)

   – **Cells – Percentages: *Row*, *Column, Total***

   Click the button **Continue**, then **OK**

3. Determine the frequencies according to **sanctions 08** for the following variables (steps from p. 2):
   - monthly income level by year
   - attitude toward management
   - level of perceived autonomy
   - level of group integration by year

4. Check the normality of the score distributions for "monthly income 08" and "monthly income 09" using the Kolmogorov-Smirnov test:

   Analyse – Descriptive Statistics – Explore

   Plots – select Normality plots with tests

   Click the button **Continue**, then **OK**

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IQ	.145	40	.034	.920	40	.008
EQ	.087	40	.200*	.977	40	.568

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

***The Kolmogorov-Smirnov test* compares the empirical distribution with the theoretical (normal) distribution.**

We look at the ***Sig.*** (significance) value. If the value is greater than 0.05, it means the distribution approximates normality, and therefore parametric inferential tests can be used.
If it is less than 0.05, there are differences between the empirical distribution and the normal distribution.

In the given example *for the IQ variable, p = 0.034, so the distribution of individual IQ scores for the 40 subjects differs from the normal distribution (Gaussian curve). For the EQ variable, p = 0.2, so the distribution of scores for this variable does not significantly differ from the normal distribution.*

 **Training:**	**Exercise 7.** Determine the frequencies (nr and %) of the scores of continuous variables.
	Exercise 8. Using the Crosstabs option, generate frequency tables (number and %) by crossing two categorical variables.
	Exercise 9. Using the normality test, determine the type of data distribution for all continuous variables.

Recommended sources:

1. 02 Descriptive Statistics and Frequencies in SPSS – SPSS for Beginners. Available: https://www.youtube.com/watch?v=bapuGcjwiLQ&list=PLVI_iGT5ZuRkk2d-kePUmlIHD5pmOuquN&index=2
2. 03 Descriptive Statistics and z Scores in SPSS – SPSS for Beginners. Available: https://www.youtube.com/watch?v=99fGYHGyO5U&list=PLVI_iGT5ZuRkk2d-kePUmlIHD5pmOuquN&index=3
3. SPSS course for beginners: Stats basics, creating variables, data entry, descriptive stats. Available: https://www.youtube.com/watch?v=-UF2k0PTw5w
4. Data cleaning, Sifting, data outlier, Cross tab, Frequencies in SPSS (Lesson 5). Available: https://www.youtube.com/watch?v=_1E5cFCgx6o
5. Normality Tests in SPSS. Available: https://www.youtube.com/watch?v=2GRZ_d4ftoo&t=298s
6. Normality Tests in SPSS. Available: https://www.youtube.com/watch?v=2GRZ_d4ftoo

Part II. INFERENTIAL STATISTICS

Topic 4. CORRELATIONS OR ASSOCIATIONS BETWEEN VARIABLES

The decision regarding which statistical test to use depends on the type of variables and distribution characteristics, according to the table:

Variable 1	*Variable 2*	*Statistical test*
Interval / ratio	Interval / ratio	**r Pearson** (for parametric distributions)
		Kt Kendall-tau (for non-parametric distributions)
Ordinal	Ordinal	**rs Spearman**
Categorical (nominal or ordinal)	Categorical (nominal or ordinal)	**χ^2 Chi-Square** (if at least one variable has 3 or more categories) **Fisher's Exact Test** (only if both variables are dichotomous)

Assessment of correlation strength based on the correlation coefficient value, according to Davis:

Correlation coefficient	**Strength of correlation**
0,01 – 0,09	Negligible correlation
0,10 – 0,29	Low / weak correlation
0,30 – 0,49	Moderate correlation
0,50 – 0,69	Substantial correlation
0,70 – 1	Very strong correlation

1. Pearson Correlation Test

Analyze – Correlate – Bivariate

Select ***Pearson***, ***Two-tailed***, ***Flag significant correlations***

Click ***OK***

Correlations

		IQ	EQ
IQ	Pearson Correlation	1	-.081
	Sig. (2-tailed)		.620
	N	40	40
EQ	Pearson Correlation	-.081	1
	Sig. (2-tailed)	.620	
	N	40	40

The correlation coefficient r = −0.081 (a negative correlation), and the significance level p = 0,62. Therefore, IQ does not correlate with EQ.

If r were 0,75(r= 0,75), p ≤ 0.05, then we would say: *EQ correlates very strongly with IQ, and this correlation is positive / directly proportional: r = 0.75, p ≤ 0.05*.

2. Kendall's Tau Association Test

Analyze – Correlate – Bivariate

Select ***Kendall's tau-b, Two-tailed, Flag significant correlations***

Click ***OK***

Correlations

			IQ	EQ
Kendall's tau_b	IQ	Correlation Coefficient	1.000	-.049
		Sig. (2-tailed)	.	.664
		N	40	40
	EQ	Correlation Coefficient	-.049	1.000
		Sig. (2-tailed)	.664	.
		N	40	40

The data obtained for the EQ variable did not show a significant association between the ranks and those obtained for the IQ variable: Kt = −0.049, p = 0.664.

In research, only the coefficient values that indicate a statistically significant correlation or association (i.e., with p ≤ 0.05) are reported.

3. Spearman Association Test

Analyze – Correlate – Bivariate

Select ***Spearman, Two-tailed, Flag significant correlations***

Click ***OK***

Correlations

			IQ	EQ
Spearman's rho	IQ	Correlation Coefficient	1.000	-.063
		Sig. (2-tailed)	.	.698
		N	40	40
	EQ	Correlation Coefficient	-.063	1.000
		Sig. (2-tailed)	.698	.
		N	40	40

There is an insignificant association / no significant association between the ranks of the "IQ" variable and those of the "EQ" variable (p = 0.698).

4. Chi square Test

It is only for categorical variables: nominal or ordinal. ***The Chi-square test is used only when at least one variable has three or more categories*** (for example, the variable "sex" has 2 categories – girls and boys; the variable "level of education" has 3 categories: high school, college, university). For instance, we may want to see whether there is an association between gender and level of education – in other words, whether the number of girls and boys differs significantly across high

school, college, and university. Or we may want to see if there is an association between level of education and IQ level.

Analyse – Descriptive Statistics – Crosstabs – Rows (one variable), ***Columns*** (another variable).

Statistics: Chi square; Phi and Cramers V. Click ***Continue***

- ***Cells: Percentages – Row, Column, Total.*** Click ***Continue***, after ***OK***

sex* studii Crosstabulation

			level of education			Total
			high school	college	university	
sex	female	Count	7	10	3	20
		% within sex	35.0%	50.0%	15.0%	100.0%
		% within level of education	100.0%	50.0%	23.1%	50.0%
	male	Count	0	10	10	20
		% within sex	0.0%	50.0%	50.0%	100.0%
		% within level of education	0.0%	50.0%	76.9%	50.0%
Total		Count	7	20	13	40
		% within sex	17.5%	50.0%	32.5%	100.0%
		% within level of education	100.0%	100.0%	100.0%	100.0%

How to read the data from the table:

By row: Out of a total of 20 girls, 7 (35%) graduated from high school, 10 (50%) from college, and only 3 (15%) attended university.

By column: Of the 7 subjects with a high school education, all are girls (100%); of the 20 subjects who graduated from college, 50% are girls (10) and 50% are boys (10); of the 13 subjects with a university degree, 23.1% are girls (3) and 76.9% (10) are boys.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.769[a]	2	.005
Likelihood Ratio	13.681	2	.001
Linear-by-Linear Association	10.005	1	.002
N of Valid Cases	40		

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is 3.50.

How to read the data from the table:

From the table above, we are interested in the Chi-square coefficient and the significance level.

The study conducted on 40 subjects confirms a higher frequency of boys with a university degree compared to girls (χ^2 = 10,697, df = 2, p = 0,005 ≤ 0,05), which means there is a significant association between level of education and gender (a higher frequency of girls with only a high school education compared to a higher frequency of boys with a university degree).

Since we obtained a statistically significant result, an additional indicator is needed to show the strength of the association — this is the Cramér's V coefficient shown in the table below.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.519	.005
	Cramer's V	.519	.005
N of Valid Cases		40	

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

How to read the data from the table:

The frequency differences between girls and boys in relation to the level of education show a strong association: phi = 0.519.

The Phi value is interpreted according to the following table:

Phi / φ	0, 01 – 0,24	Small effect
	0,25 – 0,39	Medium effect
	Peste 0,40	Large effect

5. Fisher's Exact Test

It is for categorical variables: nominal or ordinal. ***The test is applied when both variables are dichotomous*** (e.g., the variable "gender" with 2 categories – girls and boys; the variable "EQ level" with 2 categories – below average and above average).

For example, we want to see if there is a significant association between gender and EQ level; in other words, whether the number of girls and boys with below-average EQ differs significantly from the number of girls and boys with above-average EQ. The test works with frequencies, not with individual scores.

Analyse – Descriptive Statistics – Crosstabs – Rows (one variable), ***Columns*** (another variable).

- ***Exact***: select ***Exact.*** Click ***Continue***
- ***Statistics***: ***Chi Square***. Click ***Continue***

Click ***OK***

sex * level EQ Crosstabulation

			levelEQ		Total
			below average	above average	
sex	female	Count	9	11	20
		% within genul	45.0%	55.0%	100.0%
		% within nivel EQ	47.4%	52.4%	50.0%
	male	Count	10	10	20
		% within genul	50.0%	50.0%	100.0%
		% within nivel EQ	52.6%	47.6%	50.0%
Total		Count	19	21	40
		% within genul	47.5%	52.5%	100.0%
		% within nivel EQ	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.100[a]	1	.752	1.000	.500	
Continuity Correction[b]	.000	1	1.000			
Likelihood Ratio	.100	1	.751	1.000	.500	
Fisher's Exact Test				1.000	.500	
Linear-by-Linear Association	.098[c]	1	.755	1.000	.500	.236
N of Valid Cases	40					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 9.50.
b. Computed only for a 2x2 table
c. The standardized statistic is -.313.

The two-tailed significance level is 1, the one-tailed significance level is 0.5, both greater than 0.05 – the accepted significance threshold.

How to read the data from the table:

There are no differences between female and male participants regarding EQ level. / The ratio of girls to boys with below-average EQ does not differ from the ratio of girls to boys with above-average EQ.

Recommended sources:

1. Correlation Analysis in SPSS. Available: https://www.youtube.com/watch?v=k58PC9si6r4
2. 04 Correlation in SPSS – SPSS for Beginners. Available: https://www.youtube.com/watch?v=6EH5DSaCF_8&list=PLVI_iGT5ZuRkk2d-kePUmlIHD5pmOuquN&index=5
3. Pearson Correlation – SPSS. Available: https://www.youtube.com/watch?v=VOI5IlHfZVE
4. Interpret SPSS output for correlations: Pearson's r. Available: https://www.youtube.com/watch?v=pgQ9T_SdzoQ
5. SPSS Pearson Chi square independence. Available: https://www.youtube.com/watch?v=d4O_gflyj9o
6. How to Run CHi-Square and Fishers Exact | SPSS. Available: https://www.youtube.com/watch?v=x8vUQSGPH1U
7. SPSS - Fisher exact test. Available: https://www.youtube.com/watch?v=KomStQDeraY
8. How to Use SPSS-Fisher's Exact Test. Available: https://www.youtube.com/watch?v=LynasIsG0xI

Topic 5. DIFFERENCES BETWEEN TWO GROUPS

The decision regarding which test to use will be made according to the instructions in the table:

Independent Variable			*Dependent Variable*	*Statistical Test*
Categorical (nominal or ordinal)	Two categories	Independent groups	Interval / ratio (parametric distribution)	**Independent samples t-test**
			Ordinal (non-parametric) or Interval / ratio (distribution is NOT normal)	**Mann–Whitney U test**
		Dependent groups	Interval / ratio (parametric distribution)	**Paired samples t-test**
			Ordinal (non-parametric) or Interval / ratio (distribution is NOT normal)	**Wilcoxon signed-rank test**
	Three or more categories	Independent groups	Interval / ratio (parametric distribution)	**One-way ANOVA**
			Ordinal (non-parametric) or Interval / ratio (distribution is NOT normal)	**Kruskal-Wallis test**

1. Independent Samples t-test

For example, we are interested in whether there are differences in the scores obtained by subjects on the IQ and EQ tests depending on gender (2 categories — girls and boys), and the data distributions for these variables are normal.

Analyze → Compare Means → Independent-Samples T Test

In ***Test Variables,*** transfer the dependent variables (in this case, *EQ* and *IQ*).

In ***Grouping Variable,*** transfer the independent variable (in this case, gender), then define the independent variable by clicking the ***Define Group*** button: enter *Group 1* and *Group 2*. How the categories were coded in the SPSS database matters — check this in the ***Label View*** under the ***Values*** column.

Click ***Continue,*** then ***OK.***

Group Statistics

	sex	N	Mean	Std. Deviation	Std. Error Mean
IQ	female	20	85.7000	18.22404	4.07502
	male	20	116.6000	8.97013	2.00578
EQ	female	20	46.6500	10.20462	2.28182
	male	20	45.4500	8.70859	1.94730

For the “IQ” variable, the mean in the girls’ group is 85,7, while the mean in the boys’ group is 116,6. It appears that IQ is higher among boys. On the other hand, EQ is higher among girls (mean 46,65) than among boys (mean 45,45). The question “Are these differences statistically significant?” is answered by the data in the table below:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
IQ	Equal variances assumed	7.841	.008	-6.803	38	.000	-30.90000	4.54191	-40.09462	-21.70538
	Equal variances not assumed			-6.803	27.696	.000	-30.90000	4.54191	-40.20828	-21.59172
EQ	Equal variances assumed	.008	.931	.400	38	.691	1.20000	2.99978	-4.87274	7.27274
	Equal variances not assumed			.400	37.083	.691	1.20000	2.99978	-4.87767	7.27767

For each variable, there are two t-test values — for example, for IQ it is −6.803 — and two p-values, for example, ≤ 0.01. In this case, they are equal, but sometimes they differ, and then you need to decide which one to use.

To decide, you need to consider the ***p-value*** in the third column: ***Levene's test*** indicates whether the variances of the two distributions are equal or not — in our case, whether the distribution of data in the girls' group is equal to the distribution in the boys' group. If $p \leq 0.05$, it means there are significant differences in the variances of the two distributions, and you should use the t-value from the bottom row (***Equal variances not assumed***). In our example, for IQ you should take the value from the bottom row (***Equal variances not assumed***), and for EQ — the t-value from the top row (***Equal variances assumed***).

How to read the data from the table:

There are statistically significant differences between the mean scores of girls (M = 85.7) and the mean scores of boys (M = 116.6) (t = −6.803, p ≤ 0.01). / The scores obtained by girls (M = 85.7) on the IQ test are significantly lower (t = −6.803, p ≤ 0.01) than those obtained by boys (M = 116.6). / The mean score for the "IQ" variable is significantly higher among boys (mean 116.6) than among girls (mean 85.7): t = −6.803, p ≤ 0.01.

There are no gender differences for the EQ variable.

2. Mann-Whitney U test

For example, we are interested in whether there are differences in the scores obtained by subjects on the IQ and EQ tests depending on gender (2 categories — girls and boys) and the distributions of the data for these variables differ from normal.

Analyze – Nonparametric Tests – Legacy Dialogs – 2 Independent Simples

In ***Test Variables List***, transfer the dependent variables (in this case, IQ and EQ).

In ***Grouping Variable***, transfer the independent variable (in this case, sex) — then define the independent variable by clicking the ***Define Group*** button. How the categories were coded in the SPSS database matters — check this in the ***Label View*** under the ***Values*** column. Select ***Mann–Whitney U***, then click ***Continue***, then ***OK***.

Ranks

	sex	N	Mean Rank	Sum of Ranks
IQ	female	20	10.50	210.00
	male	20	30.50	610.00
	Total	40		
EQ	female	20	21.45	429.00
	male	20	19.55	391.00
	Total	40		

For the "IQ" variable, the mean rank in the girls' group is 10.5, while the mean rank in the boys' group is 30.5. This shows that boys' scores are higher than girls' scores.

Test Statistics[a]

	IQ	EQ
Mann-Whitney U	.000	181.000
Wilcoxon W	210.000	391.000
Z	-5.434	-.515
Asymp. Sig. (2-tailed)	.000	.607
Exact Sig. [2*(1-tailed Sig.)]	.000[b]	.620[b]

a. Grouping Variable: genul
b. Not corrected for ties.

How to read the data from the table:

The scores on the "IQ" variable for boys are significantly higher (mean rank 30.5) than those for girls (mean rank 10.5): $U = 0$, $N_1 = 20$, $N_2 = 20$, $p \leq 0.01$. For EQ, no statistically significant gender differences were found.

3. Paired Samples t-test

This test is used when two measurements have been taken on the same sample at different times (test and retest) and the score distributions are normal.

Analyze – Compare Means – Paired-Simple T Test

In ***Paired Variables***, transfer one of the variables to ***Variable 1*** (e.g., EQ 1) and the other to ***Variable 2*** (e.g., EQ 2)

Click ***OK***.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	EQ 1	46.0500	40	9.38343	1.48365
	EQ 2	47.9000	40	7.93499	1.25463

For the "EQ 1" variable, the mean for the test was 46.05, and the mean for the retest is 47.9. Usually, a retest is done after some form of intervention. We can see that the mean increased after the intervention, but the question is: "Is this increase statistically significant?"

The answer is shown in the table below:

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	EQ 1 - EQ 2	-1.85000	2.25945	.35725	-2.57261	-1.12739	-5.178	39	.000

How to read the data from the table:

The empathy test scores after the intervention program (M = 47.9, SD = 7.935) differ significantly (t = –5.178, df = 39, p ≤ 0.05) from the initial test scores (M = 46.05, SD = 9.383).

df — degrees of freedom

4. Wilcoxon Test

This test is used when two measurements have been taken on the same sample at different times (test and retest) and we have: 1) **ordinal variables**, or 2) **ratio or interval variables, but the score distributions are NOT parametric.**

Analyze – Nonparametric Tests – Legacy Dialogs – 2 Related Simples

In ***Test Pair***, transfer one variable to ***Variable 1*** (e.g., EQ1) and the other to ***Variable 2*** (e.g., EQ2).

In ***Test Type***, select ***Wilcoxon***, then click ***OK***.

Ranks

		N	Mean Rank	Sum of Ranks
EQ 2 - EQ 1	Negative Ranks	0[a]	.00	.00
	Positive Ranks	22[b]	11.50	253.00
	Ties	18[c]		
	Total	40		

a. EQ 2 < EQ 1
b. EQ 2 > EQ 1
c. EQ 2 = EQ 1

For 22 subjects, the scores increased after the intervention, while for 18 subjects, they remained the same.

Test Statistics[a]

	EQ 2 - EQ 1
Z	-4.124[b]
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

How to read the data from the table:

The intervention had an effect (Z/W = −4.124, p ≤ 0.01).

Is the effect present in both girls and boys?

Go to ***Data → Split File*** — check ***Organize output by groups*** and transfer the ***gender*** variable to the ***Groups Based on*** field.

Click ***OK***.

Ranks[a]

		N	Mean Rank	Sum of Ranks
EQ 2 - EQ 1	Negative Ranks	0[b]	.00	.00
	Positive Ranks	10[c]	5.50	55.00
	Ties	10[d]		
	Total	20		

a. genul = feminin
b. EQ 2 < EQ 1
c. EQ 2 > EQ 1
d. EQ 2 = EQ 1

Ranks[a]

		N	Mean Rank	Sum of Ranks
EQ 2 - EQ 1	Negative Ranks	0[b]	.00	.00
	Positive Ranks	12[c]	6.50	78.00
	Ties	8[d]		
	Total	20		

a. genul = masculin
b. EQ 2 < EQ 1
c. EQ 2 > EQ 1
d. EQ 2 = EQ 1

The score increased for 10 girls.

The score increased for 12 boys.

Test Statistics[a,b]

	EQ 2 - EQ 1
Z	-2.829[c]
Asymp. Sig. (2-tailed)	.005

a. genul = feminin
b. Wilcoxon Signed Ranks Test
c. Based on negative ranks.

Test Statistics[a,b]

	EQ 2 - EQ 1
Z	-3.086[c]
Asymp. Sig. (2-tailed)	.002

a. genul = masculin
b. Wilcoxon Signed Ranks Test
c. Based on negative ranks.

The intervention had a significant effect on girls (Z = −2.829, p ≤ 0.01).

The intervention had a significant effect on boys (Z = −3.086, p ≤ 0.01).



Training:

Exercise 12.

Compare the scores obtained on 3 continuous variables of 2 independent groups by selecting the appropriate test depending on the type of data distribution.

Exercise 13.

Compare the scores obtained on a continuous variable of 2 dependent groups by selecting the appropriate test depending on the type of data distribution.

Recommended sources:

1. 07 Independent Samples t-Tests in SPSS – SPSS for Beginners. Available: https://www.youtube.com/watch?v=-qGFZFOQx7Q

2. SPSS - Independent Samples t-Test. Available: https://www.youtube.com/watch?v=VUUfBWlziIU
3. Interpret independent t-test output from SPSS. Available: https://www.youtube.com/watch?v=I4mX_MqpDhU
4. How to run a Mann Whitney U Test | SPSS. Available: https://www.youtube.com/watch?v=OBw-8g5qjF8
5. How to run Mann-Whitney U test in SPSS by Samuel. Available: https://www.youtube.com/watch?v=F-T2tK_KB8U
6. Paired Sample t-test. Available: https://www.youtube.com/watch?v=Qt-TlsY_b3k
7. Interpret SPSS output for a paired t-test. Available: https://www.youtube.com/watch?v=siFNMaXobBE

Topic 6. DIFFERENCES AMONG THREE OR MORE GROUPS

1. One-Way ANOVA

It is used when we want **to compare 3 or more independent groups**, and the variable being compared is on an **interval or ratio scale with a normal distribution of scores**. For example, we want to determine whether EQ differs among subjects with different levels of IQ.

Analyze – Compare Means – One-Way ANOVA

In the ***Dependent List***, transfer the dependent variable, and in the ***Factor*** field, transfer the categorical variable with more than 2 categories.

Open the ***Post Hoc*** dialog box:

- In ***Equal Variances Assumed***, check the ***Bonferroni*** option.
- In ***Equal Variances Not Assumed***, check the ***Tamhane's T2*** option. Click ***Continue***.

Open the ***Options*** dialog box:

- In ***Statistics***, check ***Descriptives***, ***Homogeneity-of-variance test***, and ***Welch***. Click ***Continue***, then ***OK***.

Descriptives

EQ 1

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
below average	6	46.33	3.327	1.358	42.84	49.82	43	52
average	10	44.50	13.134	4.153	35.10	53.90	25	64
above average	14	48.36	8.608	2.301	43.39	53.33	31	64
exceptional	10	44.20	9.004	2.847	37.76	50.64	31	57
Total	40	46.05	9.383	1.484	43.05	49.05	25	64

Test of Homogeneity of Variances

EQ 1

Levene Statistic	df1	df2	Sig.
2.920	3	36	.047

Levene's test indicates whether the within-group variances are homogeneous or not. Since we have $p = 0.047 \leq 0.05$, the variances in the four groups differ. If $p > 0.05$, then the within-group variances are homogeneous.

ANOVA

EQ 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	133.252	3	44.417	.484	.695
Within Groups	3300.648	36	91.685		
Total	3433.900	39			

If p > 0.05, then the dependent variable (in our case – EQ) does not differ significantly according to the categories of the independent variable (in our case – IQ levels). Significant differences exist only if $p \leq 0.05$.

Robust Tests of Equality of Means

EQ 1

	Statistic[a]	df1	df2	Sig.
Welch	.477	3	19.096	.702

a. Asymptotically F distributed.

The Welch test confirms the inequality (if $p \leq 0.05$) or equality (if $p > 0.05$) of means between groups.

If p > 0.05, we refer to the results under the Bonferroni test in the table below.

If p ≤ 0.05, we refer to the results under the Tamhane test in the table below.

Multiple Comparisons

Dependent Variable: EQ 1

	(I) nivel inteligenta	(J) nivel inteligenta	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Bonferroni	below average	average	1.833	4.945	1.000	-11.97	15.64
		above average	-2.024	4.672	1.000	-15.07	11.02
		exceptional	2.133	4.945	1.000	-11.67	15.94
	average	below average	-1.833	4.945	1.000	-15.64	11.97
		above average	-3.857	3.965	1.000	-14.93	7.21
		exceptional	.300	4.282	1.000	-11.66	12.26
	above average	below average	2.024	4.672	1.000	-11.02	15.07
		average	3.857	3.965	1.000	-7.21	14.93
		exceptional	4.157	3.965	1.000	-6.91	15.23
	exceptional	below average	-2.133	4.945	1.000	-15.94	11.67
		average	-.300	4.282	1.000	-12.26	11.66
		above average	-4.157	3.965	1.000	-15.23	6.91
Tamhane	below average	average	1.833	4.370	.999	-12.19	15.85
		above average	-2.024	2.671	.975	-9.91	5.87
		exceptional	2.133	3.155	.986	-7.71	11.98
	average	below average	-1.833	4.370	.999	-15.85	12.19
		above average	-3.857	4.748	.966	-18.31	10.59
		exceptional	.300	5.036	1.000	-14.81	15.41
	above average	below average	2.024	2.671	.975	-5.87	9.91
		average	3.857	4.748	.966	-10.59	18.31
		exceptional	4.157	3.660	.849	-6.58	14.90
	exceptional	below average	-2.133	3.155	.986	-11.98	7.71
		average	-.300	5.036	1.000	-15.41	14.81
		above average	-4.157	3.660	.849	-14.90	6.58

If $p \leq 0.05$, there are significant differences between the group means; if $p > 0.05$, there are no significant differences between the groups.

In our case, *the results obtained on the EQ test do not differ significantly depending on the IQ level.*

2. Kruskal-Wallis test

It is used when we want **to compare 3 or more independent groups**, and the variable being compared is on **an ordinal or interval/ratio scale with a distribution that differs from normal**.

For example, we want to determine whether IQ differs among subjects with different education levels.

Analyze – Nonparametric Tests – Legacy Dialogs – K Independent Samples

In ***Test Variable List,*** transfer the dependent variable (in our case, ***IQ***).

In ***Grouping Variable,*** transfer the independent variable (in our case, education level).

In ***Define Range,*** enter the minimum and maximum values of the corresponding variable (in our case, it will be 1 and 3).

Check the ***Kruskal-Wallis,*** then click ***OK.***

Ranks

	studii	N	Mean Rank
IQ	liceu	7	7.86
	colegiu	20	19.48
	universitate	13	28.88
	Total	40	

For the IQ variable, the mean rank of subjects with high school education is 7.86, for those with college education it is 19.48, and for those with university education it is 28.88.

We observe that the highest scores are found in the group with university education, while the lowest are in the group with only high school education. *But are these differences significant?* The answer can be found in the data from the following table.

Test Statistics[a,b]

	IQ
Chi-Square	15.160
df	2
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: studii

Significant differences are observed only if $p \leq 0.05$. In our case, *there are significant differences in IQ among subjects with different levels of education:* $\chi^2 = 15.16$, $df = 2$, $p \leq 0.05$.



Training:

Exercise 14.

Compare scores obtained on 3 continuous variables of 3 and more independent groups by selecting the appropriate test depending on the type of data distribution.

Recommended sources:

1. One-way ANOVA and Post Hoc Test Using SPSS. Available: https://www.youtube.com/watch?v=ehOWuOcyejI
2. How to do a One-Way ANOVA in SPSS (12-6). Available: https://www.youtube.com/watch?v=rS3k8ONVN-o
3. Oneway ANOVA - SPSS (part 1). Available: https://www.youtube.com/watch?v=_btBuD3LIsM
4. Oneway ANOVA - SPSS (part 2). Available: https://www.youtube.com/watch?v=ZxLfDYCTlhE
5. Kruskal-Wallis H Test in SPSS. Available: https://www.youtube.com/watch?v=GK2PKB_veQc
6. Non parametric Kruskal Wallis test + POST HOC TEST in SPSS TUTORIAL. Available: https://www.youtube.com/watch?v=uXrLDq_aIas