Statistical Methods in Educational & Humanities Research

Assoc. Prof. Katerina Zlatkova-Doncheva, PhD

St. Cyril and St. Methodius University of Veliko Tarnovo

OpenEU International Summer School

“Learning and Innovation for Resilience in Sustainability”

28.-30. August 2026 | Den Haag, Netherlands

AGENDA

- **Research Methods**
- **Criteria and Indicators**
- **Instruments**
- **Variables**
- **Parametric and Nonparametric Tests**
- **Distribution**
- **Data Transformation and Presentation**
- **Mean Values**
- **Reliability**
- **Correlation**
- **Analysis**
- **Hypothesis Testing and T-test**
- **Factor Analysis and ANOVA**
- **Data Presentation (APA Style)**

Research Methods

Review

- Document Analysis
- Literature Review
- Study of Research Results (materials, tools, etc.)

Research

- Observation
- Survey
- Interview
- Expert Evaluation
- Test
- Focus Group
- Rating Scales

Experiment

- Intervention
- Observation
- Test
- Interview
- Expert Evaluations
- Focus Groups
- Rating Scales

Criteria and Indicators

Criteria	Indicator	Observed behavior
Aggression	*Physical Aggression*	Engages in physical fights at least once
Aggression	*Verbal Aggression*	Uses offensive language and verbally insults other children
Aggression	*Hostility*	Disrupts or interferes with the work of other children when attention is not received
Motivation	*Attitude toward the school as an institution*	Attend school willingly
Motivation	*Attitude toward academic work*	Regularly complete their homework
Motivation	*Attitude toward the teachers*	Show respect for their teachers.

Criteria	Indicator	Observed Behavior
Cognitive and Functional Skills	Sequencing of Actions	Follows instructions in the correct sequence
Cognitive and Functional Skills	Following Multi-Step Instructions	Is able to follow instructions involving more than one step
Engagement and Participation	Initiative	Initiates and begins activities
Engagement and Participation	More Active Participation in Learning and Functional Activities	Actively engages in activities
Bulgarian Language and Literature (Adapted Literacy Tasks)	Symbol-Supported Reading	Is able to complete tasks involving reading and/or constructing sentences using symbols.

Instruments/Measures

Survey

	Please rate the students' attitude toward the learning process.	No	Rather No	Rather Yes	Yes
1	**Attend school willingly**	1	2	3	4
2	**Regularly complete their homework**	1	2	3	4
3	**Show respect for their teachers.**	**1**	**2**	**3**	**4**

Observation

	Indicator (behavioral responses)	Variables for tables	Yes	No
A1	Engages in physical fights at least once	Physical Aggression	1	0
A2	Uses offensive language and verbally insults other children	Verbal Aggression	1	0
A3	Disrupts or interferes with the work of other children when attention is not received	Hostility	1	0

Instruments/Measures

Expert evaluation

Question	Indicator	Yes, often	Yes, sometimes	Rarely	No, never
Aggression					
Engages in physical fights	Physical Aggression	4	3	2	1
Uses offensive language	Verbal Aggression	4	3	2	1
Disrupts other children	Hostility	4	3	2	1

Interview

Question	Yes, often	Yes, sometimes	Rarely	No, never
Misbehaves in class	4	3	2	1
Frequently skips school	4	3	2	1
Falls behind in school	4	3	2	1

Variables

Dependent variables

- The variables that will be measured during the study
- Those that will be directly influenced (manipulated)
- The determination of these variables is guided by theoretical frameworks
- **All criteria and indicators:** aggression, motivation, reading literacy, height, participation, activity, etc.

Independent variables

- Control variables refer to factors that may directly or indirectly influence the results of the experiment or study
- Such as **gender, age, social status, school, class, height, weight, diagnosis, position, workplace, professional experience**, and others.

Operationalization of Independent variables

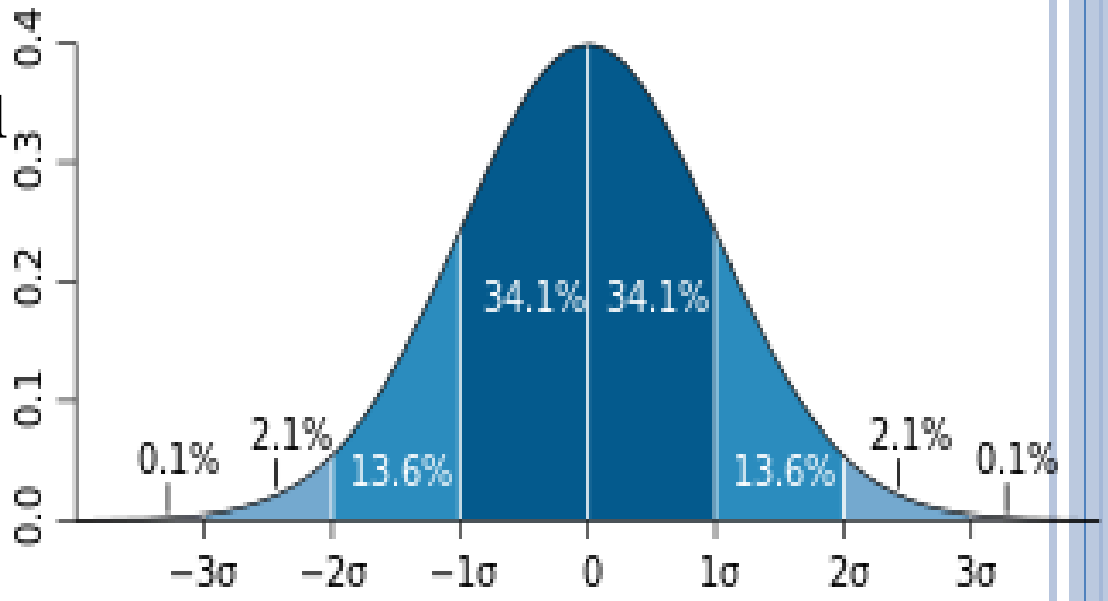
- Defined through measurable indicators.
- For example, ***participation*** is operationalized based on frequency of involvement in the experimental activities.
- The levels of participation are categorized as follows:

- low level of participation (0–20 times) - 1
- medium level of participation (21–60 times) - 2
- high level of participation (over 61 times) - 3

Distribution

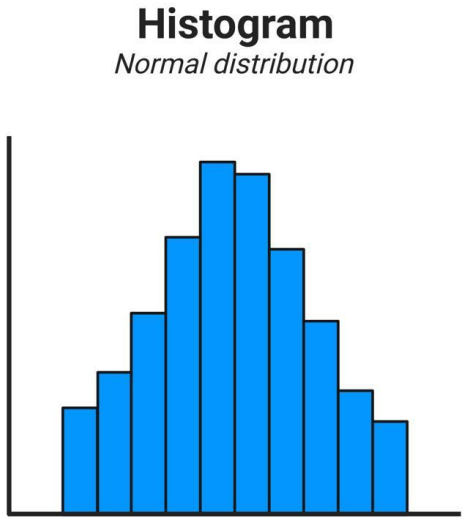
Normal

- It describes how values of a variable are distributed around the mean.
- Most observations cluster near the central value and fewer occur at the extremes.
- The mean, median, and mode are equal
- The distribution is symmetric around the mean
- The **histogram** is bell-shaped



Normal Distribution

- **For example, imagine we ask 10 students to rate a course from 1 to 5:**
- 2, 3, 3, 3, 4, 4, 4, 4, 5, 5
- The distribution tells us that:
- 2 occurs once
- 3 occurs three times
- 4 occurs four times
- 5 occurs twice
- So, rather than looking only at the individual answers, the distribution gives us a picture of the whole dataset.



Normal Distribution

.

- **Mean** value:

2 + 4 + 6 + 8 = 20

Divide by how many numbers there are (4):20 ÷ 4 = **5**

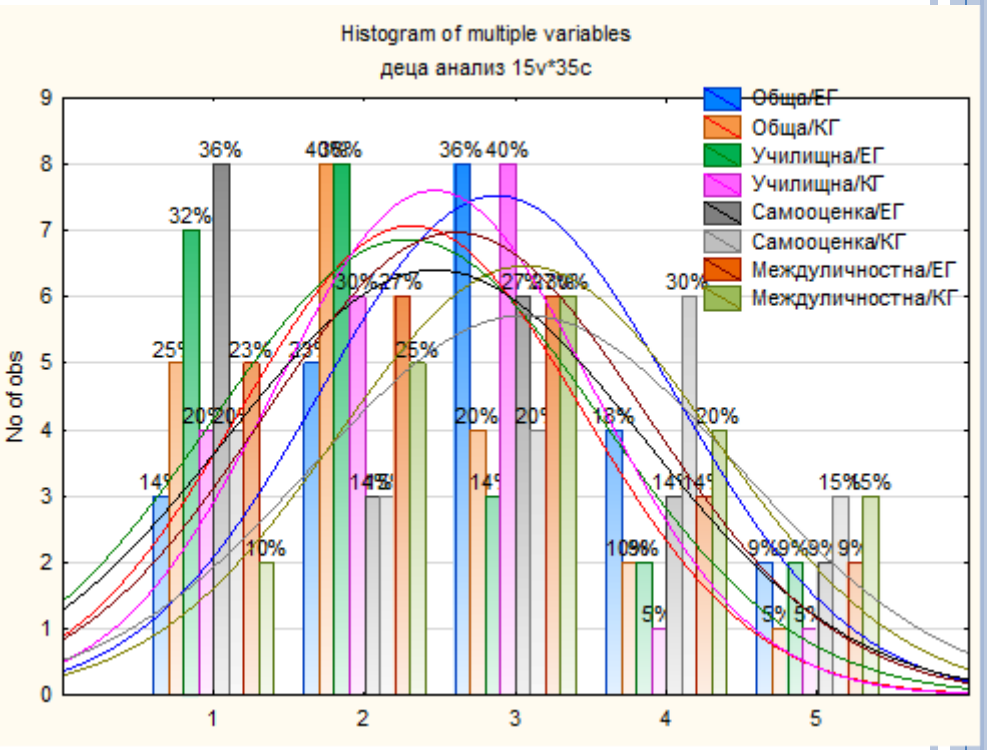
- **Mode** –the value that appears most frequently in a dataset.

In: 2, 3, 3, 5, 7 → mode = 3 (appears most often)

- **Median** –the middle value in a dataset when the numbers are arranged in order (from smallest to largest). Center of data

1, 3, 5 → median = 3

2, 4, 6, 8 → (4 + 6) ÷ 2 = 5



Mean → average value

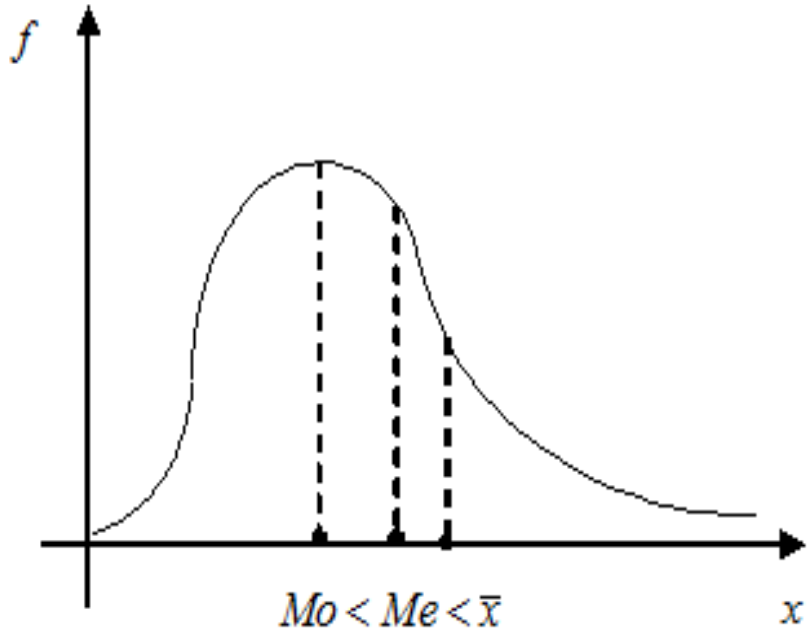
Median → middle value

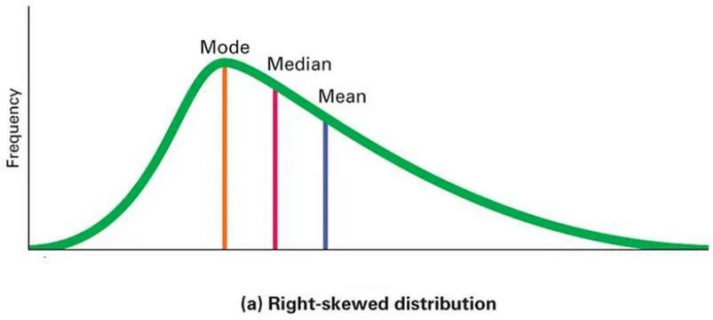
Mode → most frequent value

Asymmetric Distribution(Skewed Distribution)

Positive skew (right-skewed)

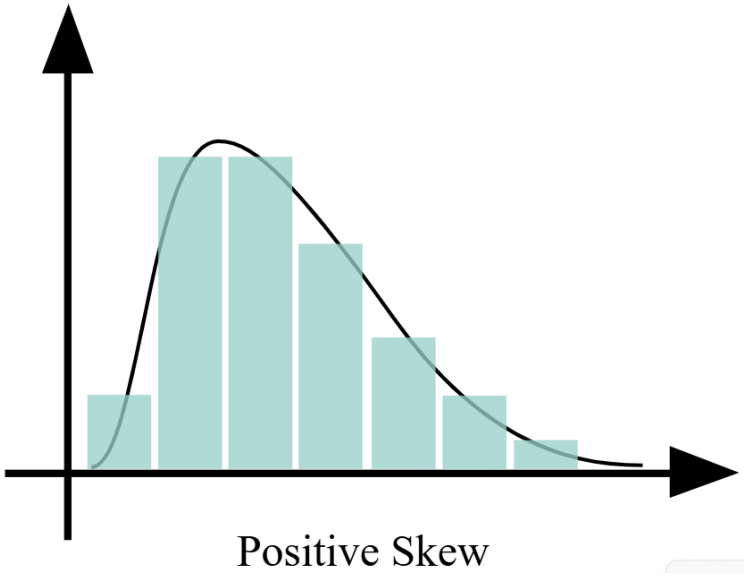
- Tail extends to the right
- Most values are on the left (smaller numbers with high values)
- Mean > Median > Mode
- Differences between mean, median, and mode
- Income distribution → usually positively skewed (few very high incomes)





Example

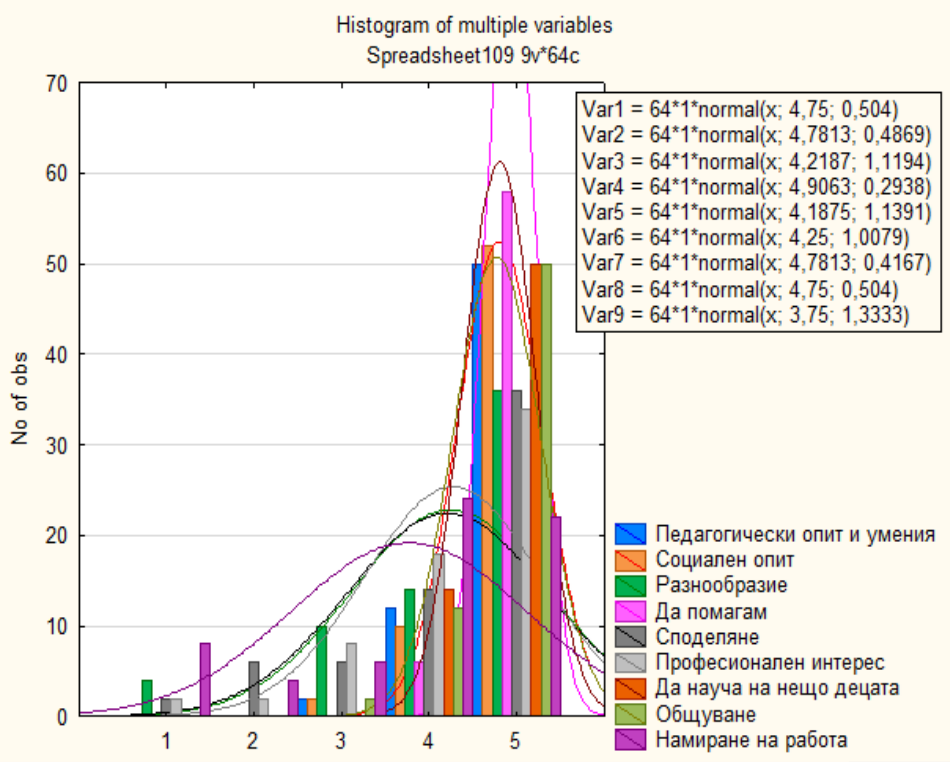
- Students' test scores
- Scores: 40, 50, 55, 60, 65, 70, 95
- Most students scored between 40–70
- One high score (95) pulls the distribution to the right



Asymmetric Distribution(Skewed Distribution)

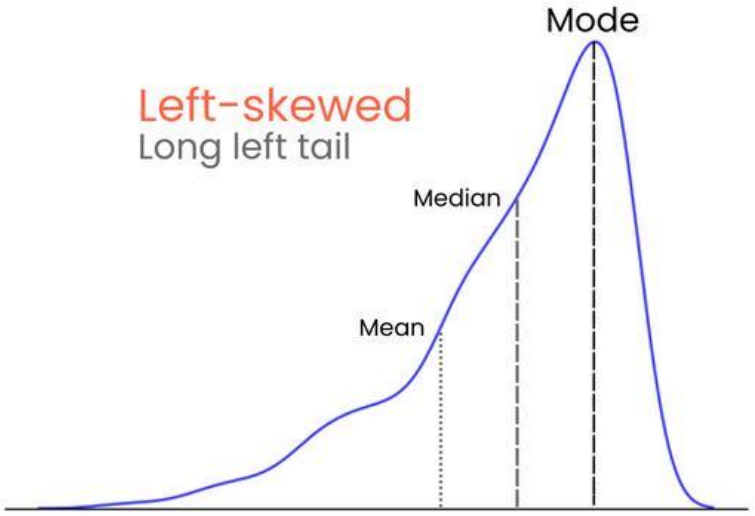
Negative skew (left-skewed)

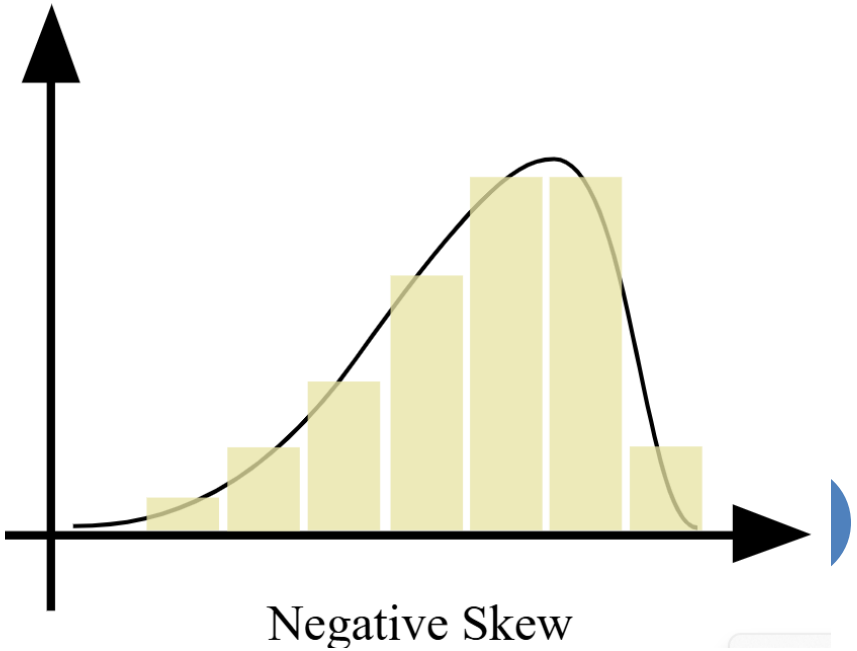
- Tail extends to the left
- Most values are on the right – smaller number of low rates
- Mean < Median < Mode
- Differences between mean, median, and mode

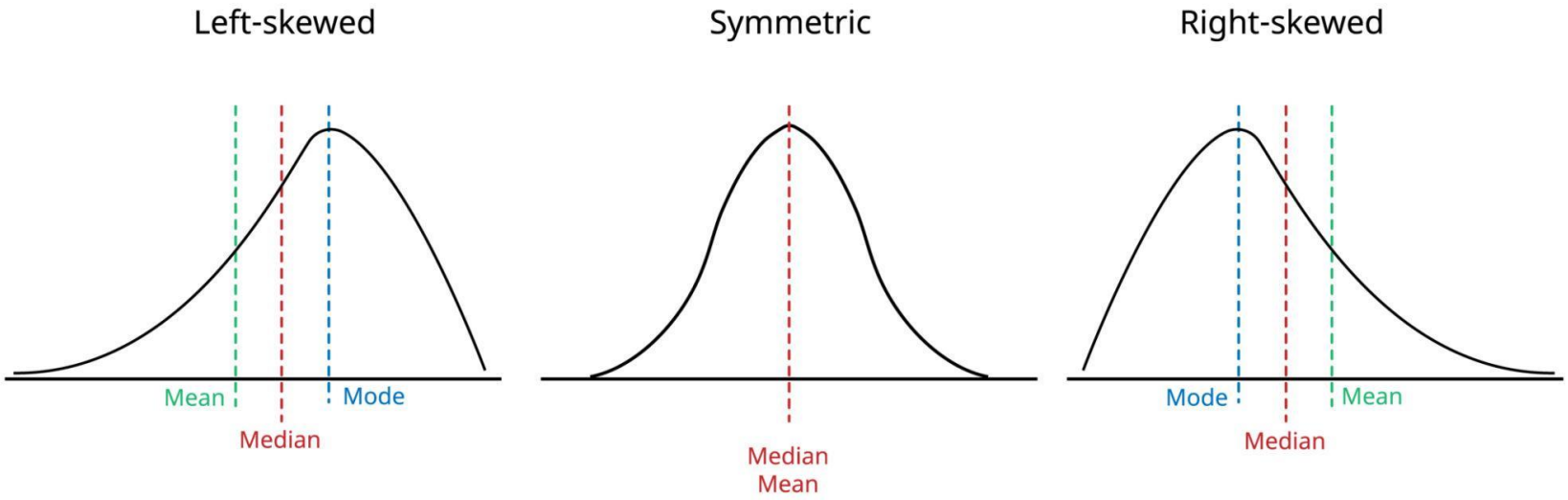


Example

- Students' test scores
- Scores: 30, 60, 65, 70, 75, 80, 85
- Most students scored high (60–85)
- One low score (30) pulls the distribution to the left







Skewness = 0.15 → almost symmetrical

Skewness = +1.40 → clearly right-skewed

Skewness = −1.30 → clearly left-skewed

Kurtosis = 0.20 → close to normal

Kurtosis = +2.50 → relatively heavy-tailed/peaked

Kurtosis = −1.20 → relatively flat/light-tailed

Normality tests SPSS

Analyze → Descriptive Statistics → Explore

Descriptives

			Statistic	Std. Error
Krit1	Mean		2.8261	.02243
	95% Confidence Interval for Mean	Lower Bound	2.7821	
		Upper Bound	2.8702	
	5% Trimmed Mean		2.8334	
	Median		2.7778	
	Variance		.202	
	Std. Deviation		.44971	
	Minimum		1.56	
	Maximum		4.00	
	Range		2.44	
	Interquartile Range		.56	
	Skewness		-.212	.122
	Kurtosis		.353	.243

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Krit1	.080	402	<.001	.986	402	<.001

a. Lilliefors Significance Correction

−1 to +1 as a simple rule

Test	Sig. (p)	Interpretation
Shapiro–Wilk	**p > .05**	Normality
Shapiro–Wilk	**p < .05**	No normality

This is why we should never judge normality from only one statistic.

Scales (weak)

Nominal

- Used to name or label categories (categorical variables)
- It does not show order or quantity — just different groups.
- Eye color: blue, brown, green
- Dichotomous nominal scale with only two categories. – gender (male/female;yes/no) 0/1

Ordinal

- Used to rank or order categories.
- The order matters, but the distance between values is not equal or known
- ✓ *Satisfaction: low, medium, high*
- ✓ *Education level: primary, secondary, higher*
- ✓ *Ranking: 1st, 2nd, 3rd*
- ✓ *Frequency: never, sometimes, often*

Scales (strong)

Interval

- Values are ordered
- The distance between values is equal
  - *Temperature in Celsius or Fahrenheit*
  - *IQ scores*
  - *Calendar years*
- Differences between individual classes can be assessed

Ratio Scale

- Order
- Numbers with equal intervals
- *Height (0 cm = no height)*
- *Weight (0 kg = no weight)*
- *Age*
- *Income*
- *Number of students*

Feature	Nominal Scale	Ordinal Scale	Interval Scale	Ratio Scale
Meaning	Categories (labels)	Ordered categories	Ordered with equal intervals	Ordered with equal intervals + true zero
Order	❌ No	✅ Yes	✅ Yes	✅ Yes
Distance between values	❌ Not defined	❌ Not equal/unknown	✅ Equal	✅ Equal
True zero	❌ No	❌ No	❌ No	✅ Yes
Can compare (> <)	❌ No	✅ Yes	✅ Yes	✅ Yes
Can add/subtract	❌ No	❌ No	✅ Yes	✅ Yes
Can multiply/divide	❌ No	❌ No	❌ No	✅ Yes

Nominal and Scale Variables

Feature	Nominal Variables	Scale Variables
Meaning	Categories (labels)	Numerical measurements
Type of data	Categorical	Quantitative
Order	❌ No order	✅ Ordered
Calculations	❌ Cannot calculate mean, median	✅ Can calculate mean, median, etc.
Examples	Gender, nationality, eye color	Height, weight, age, income
SPSS term	Nominal	Scale (includes ordinal, interval & ratio)

Parametric and non-parametric methods

Parametric

- Normal distribution.
- Data is numeric (interval/ratio)
- Large samples with large data sets (N>30).
- Example: t-test
- They are précise but sensitive to deviation

Non-parametric

- Skewed distribution.
- Data is categorical or ordinal
- Small samples (N<30).
- More flexible, accommodating different data types, including ordinal and nominal.
- Example: Mann–Whitney test
- Resistant to deviations

Parametric vs Non-Parametric

Feature	Parametric Methods	Non-Parametric Methods
Assumptions	Require assumptions (e.g. normal distribution)	Fewer or no strict assumptions
Data type	Interval or ratio data	Nominal, ordinal, or non-normal data
Distribution	Usually assumes **normal distribution**	No need for normal distribution
Sensitivity	More powerful if assumptions are met	More flexible but less powerful
Examples of tests	t-test, ANOVA, Pearson correlation	Mann–Whitney, Wilcoxon, Chi-square, Spearman

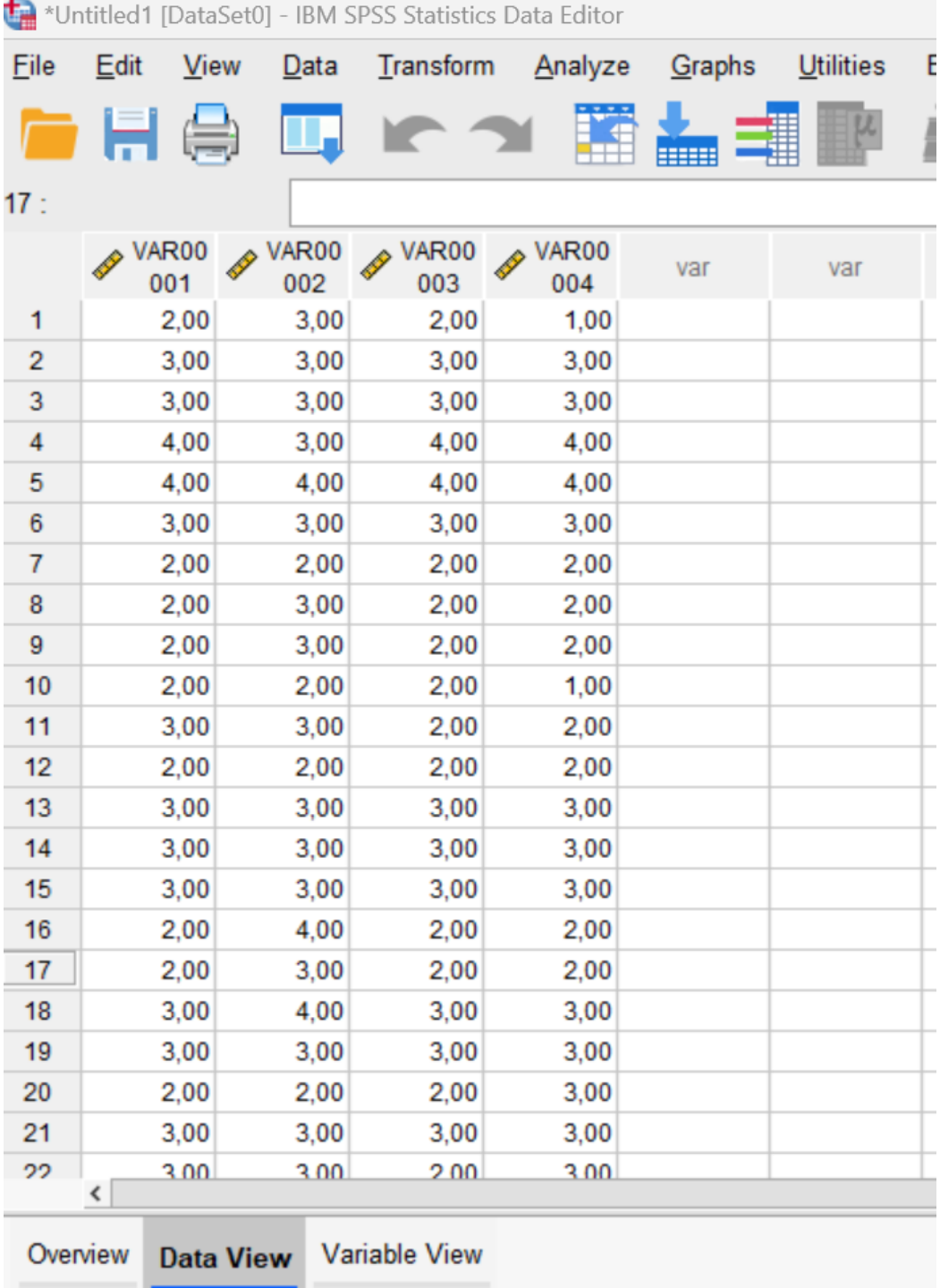
Data arrangement and transformation

- A dataset should always be organized in a specific way: the rows represent the units (e.g., individuals or observed objects) for which data are collected, while the columns represent the variables.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
ID	Q1-1	Q1-2	Q1-3	Q1-4	Q2-1	Q2-2	Q2-3	Q2-4	Q2-5	Q2-6	Q3-1	Q3-2	Q3-3	Q3-4
1	2	3	2	1	2	1	2	3	2	1	2	2	1	1
2	3	3	3	3	3	3	3	3	3	3	3	2	3	3
3	3	3	3	3	2	1	1	2	2	1	1	2	2	1
4	4	3	4	4	4	4	4	4	4	4	3	3	3	4
5	4	4	4	4	3	3	3	3	3	3	3	2	4	3
6	3	3	3	3	4	4	4	3	3	3	3	2	4	3
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	3	2	2	3	2	2	3	2	3	2	2	3	2
9	2	3	2	2	3	2	2	3	2	3	2	2	3	2
10	2	2	2	1	3	2	2	3	2	2	2	3	2	2
11	3	3	2	2	3	2	3	3	2	3	2	2	3	3
12	2	2	2	2	3	4	3	4	3	2	3	3	2	3
13	3	3	3	3	3	3	3	4	2	4	2	3	3	2
14	3	3	3	3	3	3	3	3	2	4	2	3	3	2
15	3	3	3	3	3	2	3	3	4	4	2	3	2	3
16	2	4	2	2	3	3	2	3	2	3	1	3	1	1
17	2	3	2	2	3	2	3	3	2	3	1	3	1	1
18	3	4	3	3	3	3	3	4	3	3	4	3	4	3
19	3	3	3	3	3	3	3	3	3	3	3	3	2	2

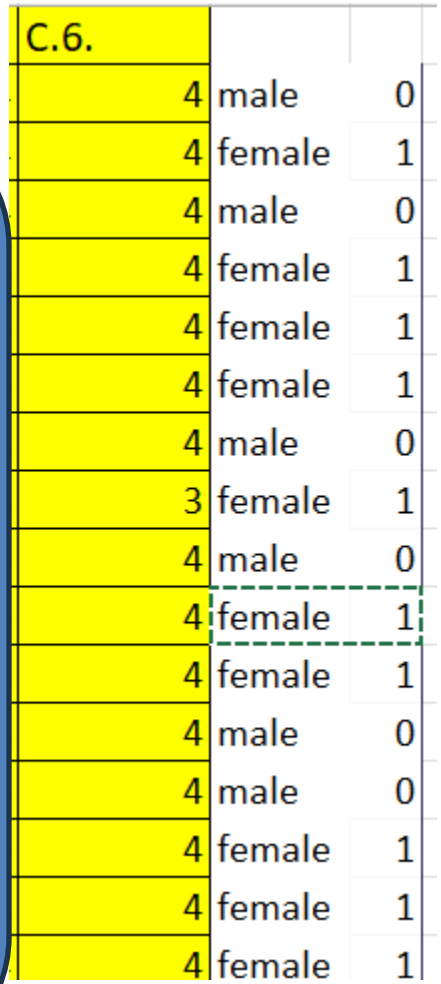
record_id	gender_caregiver	age_mother	edu_mother	work_mother	age_father	edu_father	work_father	famincome
11	1	47	4	4	48	4	4	5
26	1	42	3	0	42	3	5	5
43	1	44	3	4	46	1	5	5
48	1	39	2	3	40	1	0	3
5	1	44	4	5	45	4	3	3
6	1	39	2	0	40	3	4	3
15	2	46	3	0	56	2	5	5
16	1	46	3	0	56	2	5	5
19	1	38	1	1		2	4	1
36	0	41	3	3	43	4	4	5
37	1	46	3	5	46	1	0	2
39	1	45	4	5	44	3	5	5
1	1	40	4	5	43	4	4	5
8	1	33	2	5	34	2	4	3
18	1	40	2	0		.	.	0
29	1	38	4	5	39	1	4	4
32	1	46	4	5	54	2	3	5
38	1	46	4	2	45	3	5	5
41	1	44	4	0	44	3	0	0
4	1	37	0	3	53	0	4	1
25	1	28	1	3	34	1	5	3
46	1	54	4	0		4	.	1

A	B	C	D	E	F	G	H	I	J	K
ID	Prop bus 1	Prop bus 2	Prop bus 1	Prop bus 2	Prop hall 1	Prop hall 2	Prop hall 1	Prop hall 2	Prop lunch 1	Prop lunch 2
11	0	0	0	0	0	0	0	0	0	0
26	1	1	2	2	0	0	0	0	1	1
43	1	0	2	0	0	0	2	0	1	1
48	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
15	1	1	2	2	1	0	2	0	0	0
16	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0
36	1	0	2	0	0	0	1	0	0	1
37	1	1	2	2	1	1	2	2	1	1
39	0	0	1	0	0	0	0	0	0	0
1	1	1	2	2	0	0	2	0	0	0
8	0	0	0	1	0	0	0	0	0	0
18	1	1	1	2	0	0	0	0	0	0
29	1	0	2	0	1	1	0	2	1	1
32	0	1	0	2	0	0	0	0	0	0
38	0	1	0	1	0	0	0	0	0	0
41	0	0	1	0	1	0	1	0	0	0

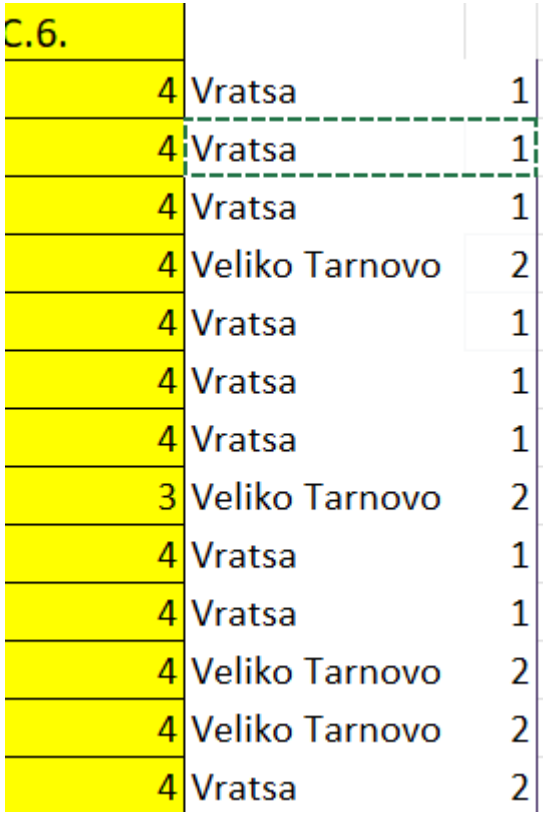

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities
17 :
VAR00001 VAR00002 VAR00003 VAR00004 var var
1 2,00 3,00 2,00 1,00
2 3,00 3,00 3,00 3,00
3 3,00 3,00 3,00 3,00
4 4,00 3,00 4,00 4,00
5 4,00 4,00 4,00 4,00
6 3,00 3,00 3,00 3,00
7 2,00 2,00 2,00 2,00
8 2,00 3,00 2,00 2,00
9 2,00 3,00 2,00 2,00
10 2,00 2,00 2,00 1,00
11 3,00 3,00 2,00 2,00
12 2,00 2,00 2,00 2,00
13 3,00 3,00 3,00 3,00
14 3,00 3,00 3,00 3,00
15 3,00 3,00 3,00 3,00
16 2,00 4,00 2,00 2,00
17 2,00 3,00 2,00 2,00
18 3,00 4,00 3,00 3,00
19 3,00 3,00 3,00 3,00
20 2,00 2,00 2,00 3,00
21 3,00 3,00 3,00 3,00
22 3,00 3,00 2,00 3,00
Overview Data View Variable View

Recoding

Recoding is a procedure in which the entered values of a variable are transformed according to a scheme selected by the researcher.



C.6.		
4	male	0
4	female	1
4	male	0
4	female	1
4	female	1
4	female	1
4	male	0
3	female	1
4	male	0
4	female	1
4	female	1
4	male	0
4	male	0
4	female	1
4	female	1
4	female	1



C.6.		
4	Vratsa	1
4	Vratsa	1
4	Vratsa	1
4	Veliko Tarnovo	2
4	Vratsa	1
4	Vratsa	1
4	Vratsa	1
3	Veliko Tarnovo	2
4	Vratsa	1
4	Vratsa	1
4	Veliko Tarnovo	2
4	Veliko Tarnovo	2
4	Vratsa	2

Descriptive statistics: frequencies

Frequencies show how often each value or category appears in your data.

Statistics

		Височина	Пол	Височина_следд
N	Valid	10	12	10
	Missing	2	0	2

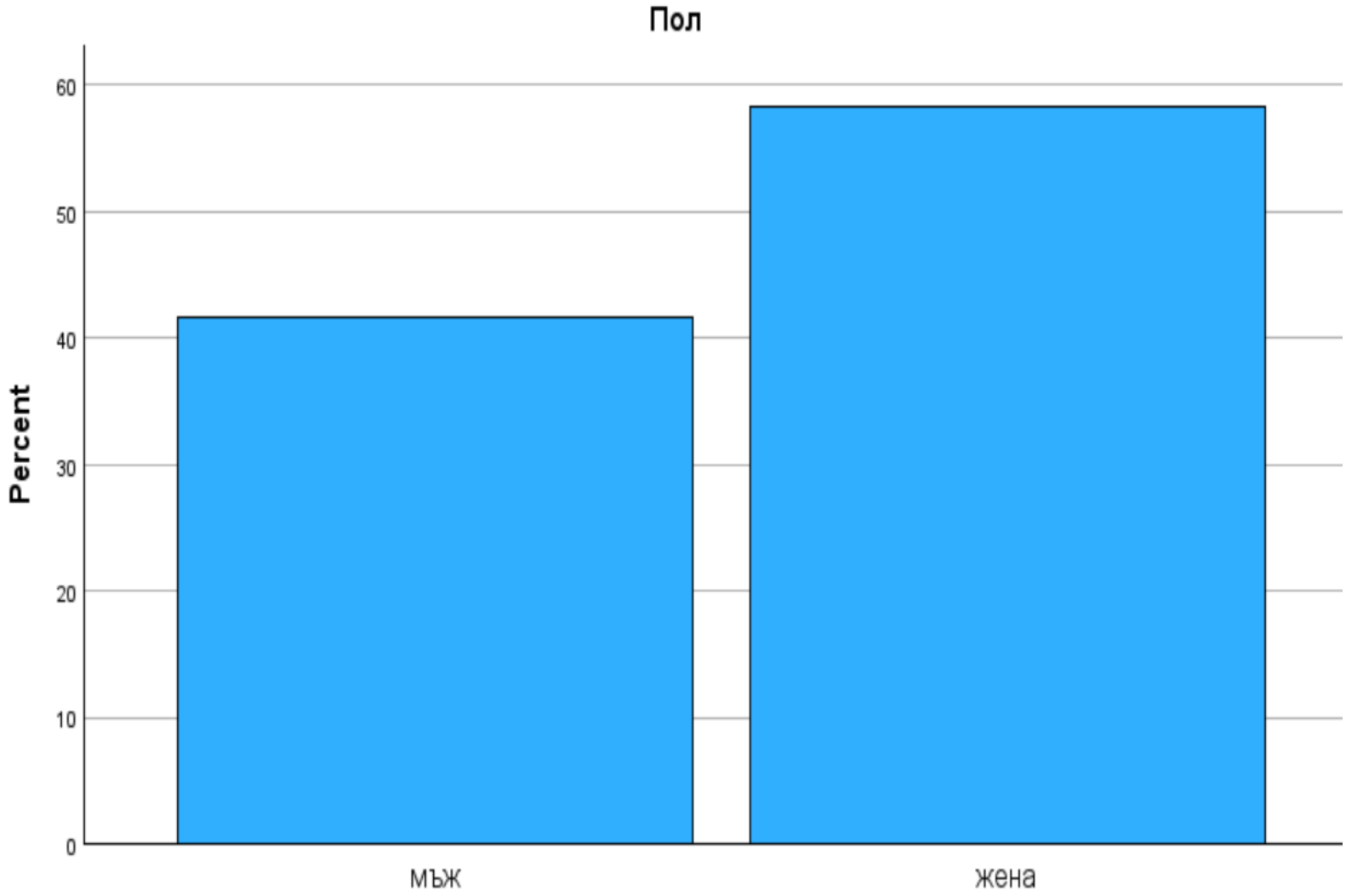
Пол

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	мъж	5	41,7	41,7	41,7
	жена	7	58,3	58,3	100,0
	Total	12	100,0	100,0	

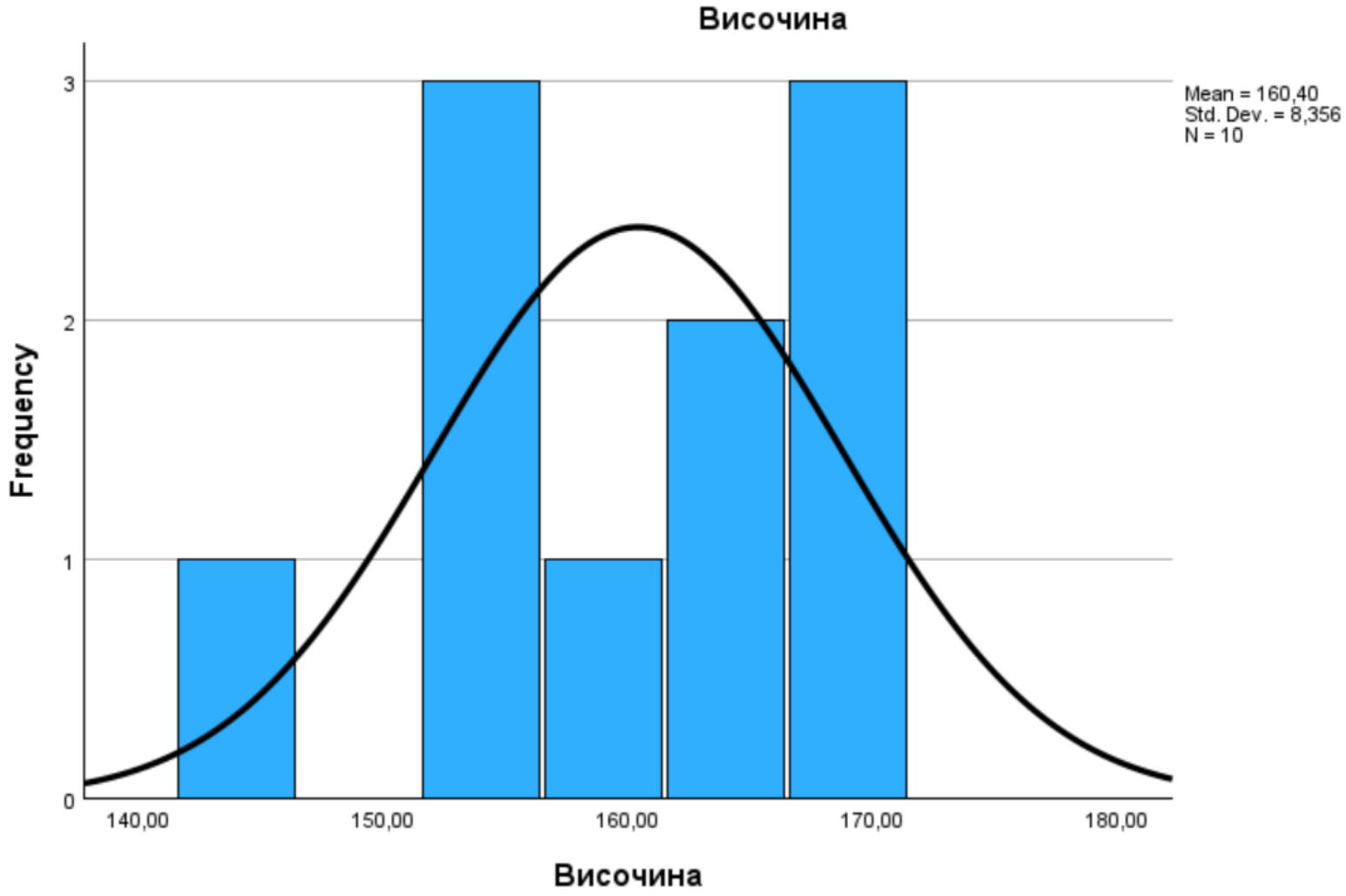
Височина

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	144,00	1	8,3	10,0	10,0
	155,00	3	25,0	30,0	40,0
	160,00	1	8,3	10,0	50,0
	162,00	1	8,3	10,0	60,0
	165,00	1	8,3	10,0	70,0
	167,00	1	8,3	10,0	80,0
	170,00	1	8,3	10,0	90,0
	171,00	1	8,3	10,0	100,0
	Total	10	83,3	100,0	
Missing	System	2	16,7		
Total		12	100,0		

Nominal variable – frequencies and bar chart



Scale variables – mean, SD, histogram



MEAN VALUE *(M)*

- The mean value (or average) is the result you get when you: add all values and divide by the number of values

$$\bar{x} = \frac{\sum x}{n}$$

- $\bar{x}$ = mean (average)
- $\sum x$ = sum of all values
- n = number of values
- Example: Mean of 4, 7 , 1 = (4+7+1)=12/3=**4**

Standard Deviation (*SD*)

- Standard deviation shows how spread out the data is around the mean. It tells you if values are close to the average or far from it.
- Small standard deviation → values are close to the mean
- Large standard deviation → values are spread out
- Data: 5, 6, 7 ; Mean = 6; Values are close → small standard deviation
- Data: 2, 6, 10; Mean = 6; Values are far → large standard deviation

SD

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

Where:

- s = standard deviation
- x = each value
- $\bar{x}$ = mean
- n = number of values

Four friends compare their exam results.

- Step 1 1: Find the mean value.

$$\mu = \frac{6+2+3+1}{4} = \frac{12}{4} = 3$$

- Step 2: Get the mean value out of each score

Резултат: x_i	Отклонение: $(x_i - \mu)$
6	$6 - 3 = 3$
2	$2 - 3 = -1$
3	$3 - 3 = 0$
1	$1 - 3 = -2$

- Стъпка 4: Add up the squared deviations

Резултат: x_i	Отклонение: $(x_i - \mu)$	Отклонение на квадрат: $(x_i - \mu)^2$
6	$6 - 3 = 3$	$(3)^2 = 9$
2	$2 - 3 = -1$	$(-1)^2 = 1$
3	$3 - 3 = 0$	$(0)^2 = 0$
1	$1 - 3 = -2$	$(-2)^2 = 4$

9+1+0+4=14

- Step 5: Divide the sum by the number of

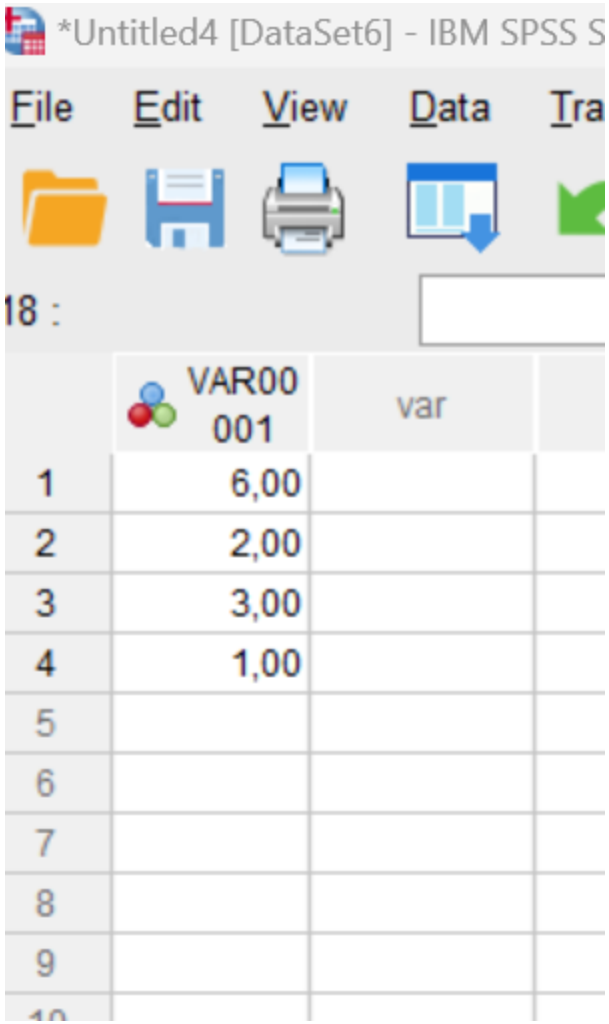
-for the sample 14/4-1=4,66

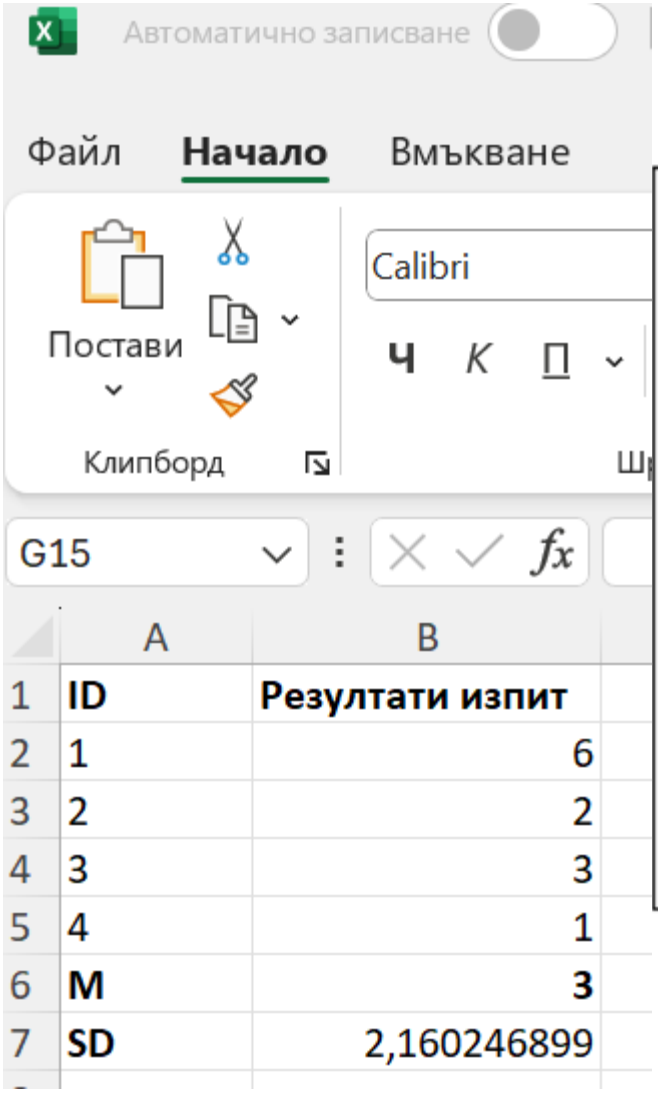
$$\frac{14}{4} = 3{,}5$$

- Step 6: Find the square root of the result from Step 5.

$$\sqrt{3{,}5} \approx 1{,}87$$

- **SD = 1,87 for the population**
- **SD = 2,15 for the sample**





Statistics

VAR00001

N	Valid	4
	Missing	0
Mean		3,0000
Median		2,5000
Mode		1,00[a]
Std. Deviation		2,16025
Variance		4,667

a. Multiple modes exist. The smallest value is shown

Means

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Височина * Пол	10	83,3%	2	16,7%	12	100,0%

Report

Височина

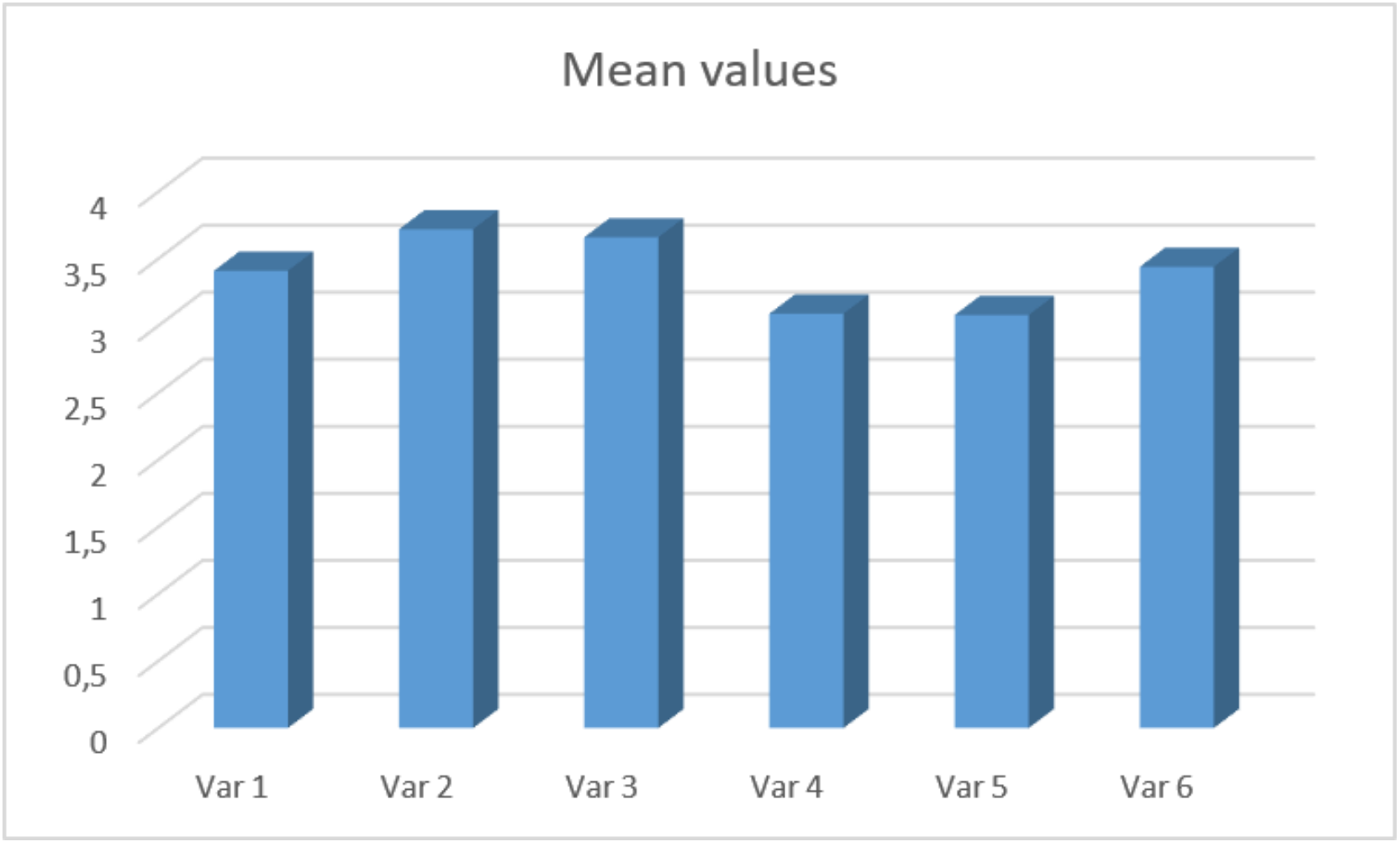
Пол	Mean	N	Std. Deviation
мъж	164,6000	5	6,87750
жена	156,2000	5	8,10555
Total	160,4000	10	8,35597

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic
Използване на ДАК средства	402	2	4	3.79	.492
Използване на Cboard "	402	1	4	2.08	1.074
Мимики и жестове	402	1	4	3.39	.838
Написан текст	402	1	4	3.09	.910
Картинки, снимки и символи"	402	1	4	3.79	.504
Реални предмети"	402	1	4	3.65	.635
Устройства за синтезиран говор с бутони	402	1	4	1.82	.892
Специализиран софтуер"	402	1	4	1.90	.950
PECS, Макатон или др.	402	1	4	1.92	1.003
Valid N (listwise)	402				

Mean values for text

Показател	N	M	SD	Min	Max
Използване на ДАК средства	402	3.79	.492	2	4
Използване на Cboard "	402	2.08	1.074	1	4
Мимики и жестове	402	3.39	.838	1	4
Написан текст	402	3.09	.910	1	4
Картинки, снимки и символи"	402	3.79	.504	1	4
Реални предмети"	402	3.65	.635	1	4
Устройства за синтезиран говор с бутони	402	1.82	.892	1	4
Специализиран софтуер"	402	1.90	.950	1	4
PECS, Макатон или др.	402	1.92	1.003	1	4


Mean values
4
3,5
3
2,5
2
1,5
1
0,5
0
Var 1
Var 2
Var 3
Var 4
Var 5
Var 6

Mean for text needed

Scale	Mean	SD	N
Leadership & Org Culture	3.78	0.10	226
Planning	3.79	0.11	226
Motivation	3.77	0.11	226
Control	3.78	0.20	226
Overall	3.78	0.13	226

Table 1. Bulgarian students: descriptives by scale (1–4).

Reliability test

- **Cronbach's alpha (α)** is used to measure the reliability (internal consistency) of a test or questionnaire.
- If you have a questionnaire about motivation with 10 questions: Cronbach's alpha checks if all 10 questions are consistent with each other(i.e. they measure motivation in a similar way).
- High alpha = items are consistent
- Low alpha = items may not measure the same thing

Alpha (α)	Interpretation
≥ 0.9	Excellent
≥ 0.8	Good
≥ 0.7	Acceptable
≥ 0.6	Questionable
< 0.6	Poor

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.889	.873	8

In SPSS

- Analyze → Scale → Reliability Analysis

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.567	.576	5

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Готови помощни материали и работни листи с картинки	13.55	3.605	.316	.146	.519
Работни листи с конкретни образователни задачи	13.23	4.088	.391	.211	.497
Картинки	13.30	4.255	.204	.126	.571
Методически указания	13.87	3.511	.315	.122	.521
Готови адаптирани уроци с обяснения и упражнения"	13.87	2.999	.449	.252	.429

Reliability in text Cronbach alpha

Scale	Items (n)	Bulgaria α	U.S. α
Leadership & Organizational Culture	7	0.745	0.883
Planning	4	0.557	0.825
Motivation	7	0.872	.741
Control	6	0.630	.750

Correlation

Measures the relationship between two variables– between -1 and +1

1. Positive correlation

- As one variable increases → the other also increases
- Example: study time ↑ → test scores ↑

2. Negative correlation

- As one variable increases → the other decreases
- Example: stress ↑ → performance ↓

3. No correlation

- No relationship between variables

Correlation analysis

- **Correlation coefficient (r):** $r = -1$ *to* $+1$
- **+1** → perfect positive correlation
- **0** → no correlation
- **-1** → perfect negative correlation
- ***p*< .05**
- **The p-value (probability value) significance**
- **It tells you how likely your results are if there is**

➔ **Correlations**

Correlations

		Височина	Височина_след
Височина	Pearson Correlation	1	,875**
	Sig. (2-tailed)		<,001
	N	10	10
Височина_след	Pearson Correlation	,875**	1
	Sig. (2-tailed)	<,001	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

- **Pearson correlation (r)**→ for numerical, normally distributed data
- **Spearman correlation (r)**→ for ordinal or non-normal data

❗ Correlation does not mean causation(just because two things are related does not mean one causes the other)

Interpreting results:

- r > 0.7 – strong correlation
- r < . 6 r > . 04 – moderate correlation
- r < .40 – weak correlation

- **Correlations**

Correlations

		Височина	Височина_след	Пол
Височина	Pearson Correlation	1	,875**	-,530
	Sig. (2-tailed)		<,001	,115
	N	10	10	10
Височина_след	Pearson Correlation	,875**	1	-,725*
	Sig. (2-tailed)	<,001		,018
	N	10	10	10
Пол	Pearson Correlation	-,530	-,725*	1
	Sig. (2-tailed)	,115	,018	
	N	10	10	12

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		Височина	Пол	Тегло
Височина	Pearson Correlation	1	-,530	,011
	Sig. (2-tailed)		,115	,977
	N	10	10	10
Пол	Pearson Correlation	-,530	1	,303
	Sig. (2-tailed)	,115		,394
	N	10	12	10
Тегло	Pearson Correlation	,011	,303	1
	Sig. (2-tailed)	,977	,394	
	N	10	10	10

Correlations

		Височина	Тегло
Височина	Pearson Correlation	1	,011
	Sig. (2-tailed)		,977
	N	10	10
Тегло	Pearson Correlation	,011	1
	Sig. (2-tailed)	,977	
	N	10	10

KENDALL TAU – FOR SMALL SAMPLES

Nonparametric Correlations

Correlations

			Височина	Височина_след
Kendall's tau_b	Височина	Correlation Coefficient	1,000	,706**
		Sig. (2-tailed)	.	,006
		N	10	10
	Височина_след	Correlation Coefficient	,706**	1,000
		Sig. (2-tailed)	,006	.
		N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

How do you present

VARIABLES	1	2	3	4	5	6	7	8	9
VARIABLE 1	—	ρ=.183; p<.001	ρ=.118; p=.018	ρ=.202; p<.001	ρ=.240; p<.001	ρ=.244; p<.001	ρ=.302; p<.001	ρ=.328; p<.001	ρ=-.188; p<.001
VARIABLE 2		—	ρ=-.121; p=.015	ρ=.213; p<.001	ρ=.171; p<.001	ρ=.215; p<.001	ρ=-.195; p<.001	ρ=.135; p=.007	ρ=-.235; p<.001
VARIABLE 3			—	ρ=-.119; p=.017	ρ=-.077; p=.121	ρ=-.040; p=.425	ρ=.295; p<.001	ρ=-.141; p=.005	ρ=-.125; p=.012
VARIABLE 4				—	ρ=.123; p=.013	ρ=.187; p<.001	ρ=.052; p=.302	ρ=.262; p<.001	ρ=-.004; p=.929
VARIABLE 5					—	ρ=.305; p<.001	ρ=.013; p=.798	ρ=.156; p=.002	ρ=.083; p=.096
VARIABLE 6						—	ρ=.035; p=.484	ρ=.157; p=.002	ρ=.020; p=.683
VARIABLE 7							—	ρ=.290; p<.001	ρ=-.290; p<.001
VARIABLE 8								—	ρ=-.133; p=.007
VARIABLE 9									—

Hypothesis

H_1: Alternative Hypothesis

- There is a difference between the two variances.
- The intervention is working
- There is an effect
- Something has changed
- H_1: There is a difference between two teaching methods

H_0: null hypothesis

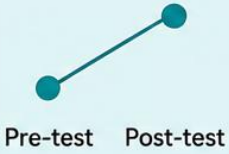
- There is no difference between the two variances.
- The intervention is not working
- There is no effect
- Nothing has changed
- **It is the default assumption**
- H_0: There is no difference between two teaching methods

Decision rule (using p-value):

If $p < 0.05 \rightarrow$ reject $H_0 \rightarrow$ support H_1 ✅

If $p \geq 0.05 \rightarrow$ do not reject H_0 ❌

T-TESTS

TEST	WHAT ARE WE COMPARING?	SIMPLE EXAMPLE
One-Sample t-test	One group vs. a known/reference value	 Is the average student score different from 50?
Paired-Samples t-test	The same group measured twice	 Did students improve from pre-test to post-test?
Independent-Samples t-test	Two different groups	 Do male and female students have different mean scores?

One sample t-test

- One-sample t-test = compare your group's average to a fixed value
- Used to check whether the mean of one group is different from a known or expected value.
- SPSS - Analyze – Compare means – One-sample T-test

p < 0.05

t-value and critical value

Confidence interval crosses 0

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Височина	10	160,4000	8,35597	2,64239

One-Sample Test

Test Value = 165

	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper
Височина	-1,741	9	,058	,116	-4,60000	-10,5775	1,3775

One-Sample Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Височина	Cohen's d	8,35597	-,551	-1,206	,131
	Hedges' correction	9,14345	-,503	-1,102	,120

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

Paired sample T-test

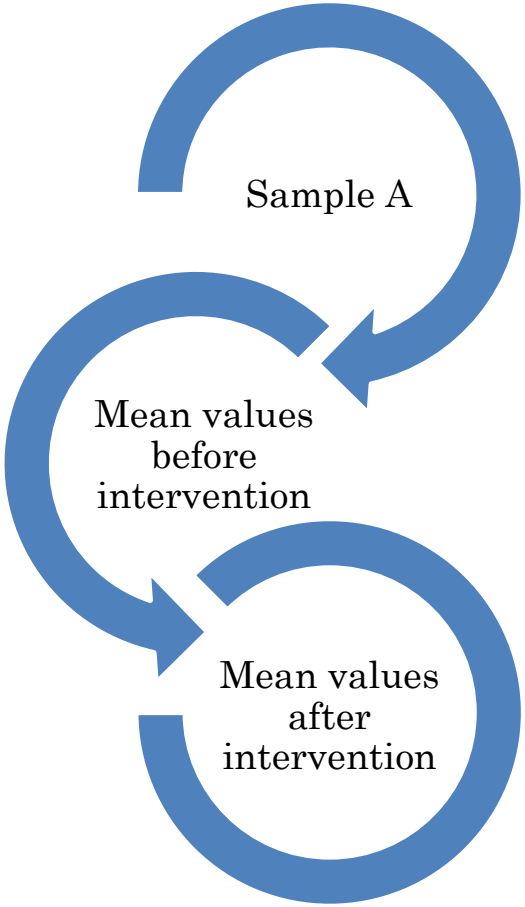
- Repeated measure design
- Compares the means of the same group measured twice
- Pre test и post test
- Students' scores before training and after training

Analyze – Compare means – Paired-sample T-test

p < 0.05

t-value and critical value

Confidence interval crosses 0 (confidence interval – probability that a repeated test will have identical results)



Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Височина	160,4000	10	8,35597	2,64239
	Височина_след	172,0000	10	6,11010	1,93218

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	Височина & Височина_след	10	,875	<,001	<,001

Paired Samples Test

		Paired Differences					t	df	Significance One-Sided p	Significance Two-Sided p
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper				
Pair 1	Височина - Височина_след	-11,60000	4,22164	1,33500	-14,61998	-8,58002	-8,689	9	<,001	<,001

Paired Samples Correlations

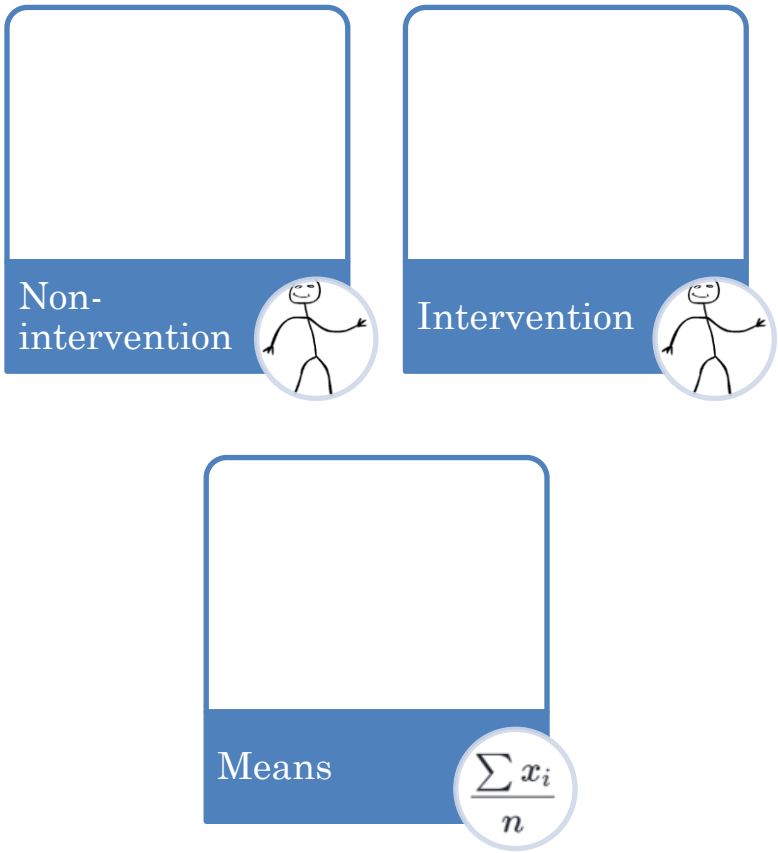
		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	Височина & Височина_след	10	,875	<,001	<,001
Pair 2	Височина_след & Височина	10	,875	<,001	<,001

Paired Samples Test

		Paired Differences					t	df	Significance One-Sided p	Significance Two-Sided p
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper				
Pair 1	Височина - Височина_след	-11,60000	4,22164	1,33500	-14,61998	-8,58002	-8,689	9	<,001	<,001
Pair 2	Височина_след - Височина	11,60000	4,22164	1,33500	8,58002	14,61998	8,689	9	<,001	<,001

Independent sample t-test

- Compares the means of two different groups
- Is there a statistically significant difference between 2 groups? (male-female, EG/CG)
- SPSS - Analyze – Compare means – Independent sample T-test
- **p < 0.05**
- **t-value and critical value**
- **Confidence interval crosses 0**



Independent vs Paired T-Test

FEATURE	INDEPENDENT T-TEST	PAIRED T-TEST
GROUPS	Two different groups	One group (twice measured)
DATA TYPE	Independent	Related / paired
EXAMPLE	Male vs female	Before vs after

- **Example (one-sample t-test):**
- A one-sample t-test showed that the mean score ($M = 73.8$, $SD = 3.0$) was higher than the expected value of 70, $t(4) = 2.85$, $p = .046$.
- **Example (independent t-test):** There was a significant difference between groups,

$t(38) = 2.10$, $p = .042$.

- **Example (paired t-test):** Scores after the intervention ($M = 80.2$) were significantly higher than before ($M = 70.5$), $t(19) = 3.25$, $p = .004$.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
						Significance			
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference
Височина	Equal variances assumed	,004	,950	1,767	8	,058	,115	8,40000	4,75395
	Equal variances not assumed			1,767	7,793	,058	,116	8,40000	4,75395
Височина_след	Equal variances assumed	,984	,350	2,974	8	,009	,018	8,40000	2,82489
	Equal variances not assumed			2,974	7,967	,010	,020	8,40000	2,82489

- Variable: Височина (Height)
  - t(8) = 1.767
  - p (two-sided) = .115 p > 0.05 ✗ Not statistically sig. Variable: Височина_след (Height_after)
  - t(8) = 2.974
  - p (two-sided) = .018 p < 0.05 ✓ Statistically sig.

Height (before): There was no significant difference between groups,

t(8) = 1.77, p = .115.

Height (after): There was a significant difference between groups,

t(8) = 2.97, p = .018.

At the start → groups are similar; After → groups are different

ANOVA (Analysis of variance)

- ANOVA is a statistical test used to compare the means of three or more groups.
- It tells you if there is a significant difference between groups. **Are these differences large enough to be statistically significant, or could they simply be due to chance?**
- When you want to compare their average values

Test result (F-value):ANOVA gives you:

F statistic; p-value

If $p < 0.05$ → significant difference ✅

If $p \geq 0.05$ → no difference ❌

One-way ANOVA → one factor (e.g. teaching method)

Two-way ANOVA → two factors (e.g. method + gender)

- This is basically an extension of the Independent-Samples
- t-test:t-test → 2 groups
- ANOVA → 3 or more groups
- **Levene's test**
- Before interpreting the standard ANOVA, we check one important assumption: homogeneity of variances. **Are the groups approximately equally variable?**

Test	What does it test?	If $p > .05$	If $p < .05$	What to do?
One-Way ANOVA	Are the means of 3+ groups different?	No significant difference	Significant difference between at least two groups	If significant → Post-hoc test
Levene's Test	Are the variances equal across groups?	Variances are similar (homogeneity assumed)	Variances are significantly different	$p > .05$ → Standard ANOVA; $p < .05$ → Welch ANOVA
Welch ANOVA	Are the means of 3+ groups different when variances are unequal?	No significant difference	Significant difference between at least two groups	If significant → appropriate Post-hoc test

ANOVA TESTS – QUICK GUIDE

TEST	WHAT DOES IT TEST?	IF p > .05 (not significant)	IF p < .05 (significant)	WHAT TO DO?
One-Way ANOVA	Are the means of 3+ groups different?	No significant difference	Significant difference between at least two groups	If significant → Post-hoc test
Levene's Test	Are the variances equal across groups?	Variances are similar (homogeneity assumed)	Variances are significantly different	p > .05 → **Standard ANOVA** p < .05 → **Welch ANOVA**
Welch ANOVA	Are the means of 3+ groups different when variances are unequal?	No significant difference	Significant difference between at least two groups	If significant → appropriate Post-hoc test

Note: Always start with Levene's Test to check if variances are equal. Then choose the appropriate ANOVA test.

ANOVA

Височина_след

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	176,400	1	176,400	8,842	,018
Within Groups	159,600	8	19,950		
Total	336,000	9			

- $F(1, 8) = 8.842$
- p (Sig.) $= .018 < 0.05$
- The result is statistically significant
- There is a significant difference between the groups for Height after. This suggests the **intervention or group effect had an impact**
- The groups were different after the experiment
- This supports the idea that your independent variable had an effect

Data representing

ANOVA

Височина_след

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	176,400	1	176,400	8,842	,018
Within Groups	159,600	8	19,950		
Total	336,000	9			

$F(1, 8) = 8{,}842,\ p > .001$

- Report in text: There was a significant effect of group on height after the intervention,
- F(1, 8) = 8.842, p = .018.

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound
1	38	3.5175	.39094	.06342	3.3890	3.6460
2	94	3.4433	.45761	.04720	3.3495	3.5370
3	164	3.3201	.46485	.03630	3.2484	3.3918
4	90	3.4185	.41676	.04393	3.3312	3.5058
5	16	3.5417	.36260	.09065	3.3484	3.7349
Total	402	3.3984	.44632	.02226	3.3547	3.4422

Robust Tests of Equality of Means

Krit1

	Statistic[b]	df1	df2	Sig.[a]
Welch	.275	4	78.047	.893

a. Confidence Interval: 95%

b. Asymptotically F distributed.

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.[a]
Krit1	Based on Mean	3.799	4	397	.005
	Based on Median	3.006	4	397	.018
	Based on Median and with adjusted df	3.006	4	374.136	.018
	Based on trimmed mean	3.611	4	397	.007

a. Confidence Interval: 0%

ANOVA

Krit1

	Sum of Squares	df	Mean Square	F	Sig.[a]
Between Groups	.423	4	.106	.520	.721

Post Hoc Test

Multiple Comparisons

Dependent Variable: Krit2
Tukey HSD

(I) Възраст	(J) Възраст	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	.07428	.08509	.907	-.1589	.3075
	3	.19742	.07969	.098	-.0210	.4158
	4	.09903	.08563	.776	-.1356	.3337
	5	-.02412	.13191	1.000	-.3856	.3374
2	1	-.07428	.08509	.907	-.3075	.1589
	3	.12314	.05726	.201	-.0338	.2801
	4	.02474	.06528	.996	-.1541	.2036
	5	-.09840	.11971	.924	-.4264	.2296
3	1	-.19742	.07969	.098	-.4158	.0210
	2	-.12314	.05726	.201	-.2801	.0338
	4	-.09840	.05807	.438	-.2575	.0607
	5	-.22154	.11593	.313	-.5392	.0961
4	1	-.09903	.08563	.776	-.3337	.1356
	2	-.02474	.06528	.996	-.2036	.1541
	3	.09840	.05807	.438	-.0607	.2575
	5	-.12315	.12009	.844	-.4522	.2059
5	1	.02412	.13191	1.000	-.3374	.3856
	2	.09840	.11971	.924	-.2296	.4264
	3	.22154	.11593	.313	-.0961	.5392
	4	.12315	.12009	.844	-.2059	.4522

Tukey HSD	1	2	.232	.142	.476	-.16	.62
		3	.253	.133	.315	-.11	.62
		4	.599*	.143	<.001	.21	.99
		5	.461	.220	.224	-.14	1.06
	2	1	-.232	.142	.476	-.62	.16
		3	.021	.095	.999	-.24	.28
		4	.368*	.109	.007	.07	.67
		5	.229	.199	.781	-.32	.78
	3	1	-.253	.133	.315	-.62	.11
		2	-.021	.095	.999	-.28	.24
		4	.346*	.097	.004	.08	.61
		5	.207	.193	.820	-.32	.74
	4	1	-.599*	.143	<.001	-.99	-.21
		2	-.368*	.109	.007	-.67	-.07
		3	-.346*	.097	.004	-.61	-.08
		5	-.139	.200	.958	-.69	.41
	5	1	-.461	.220	.224	-1.06	.14
		2	-.229	.199	.781	-.78	.32
		3	-.207	.193	.820	-.74	.32
		4	.139	.200	.958	-.41	.69
Games-Howell	1	2	.232	.112	.241	-.08	.54
		3	.253	.106	.131	-.04	.55
		4	.599*	.139	<.001	.21	.98
		5	.461	.194	.156	-.11	1.03
	2	1	-.232	.112	.241	-.54	.08
		3	.021	.083	.999	-.21	.25
		4	.368*	.122	.025	.03	.70
		5	.229	.182	.720	-.32	.78
	3	1	-.253	.106	.131	-.55	.04
		2	-.021	.083	.999	-.25	.21
		4	.346*	.117	.029	.02	.67
		5	.207	.179	.773	-.33	.75
	4	1	-.599*	.139	<.001	-.98	-.21

Levene p > .05 → ANOVA → Tukey

Levene p < .05 → Welch ANOVA → Games–Howell

How we present

Variable	Levene p	Test used	Statistics	p (2-tailed)	Significance	Post Hoc
Variable 1	.005	Welch	F(4, 78.047) = .275	.893	No	-
Variable 2	.784	ANOVA	F(4, 397) = 2.677	.032	Yes	No significant pairs (Tukey)
Variable 3	<.001	Welch	F(4, 81.808) = 4.942	.001	да	До 30 г. > 51-60 г. (p < .001); 31–40 г. > 51-60 г. (p = .025); 41-50 г. > 51 – 60 г (p= .029) (Games-Howell)

Variable	**Levene p**	**Test used**	**Statistics**	**p** (2-tailed)	**Significance**	**Group with higher value (M)**
Variable 1	.030	Welch	F(1, 34.67) = 9.73	.004	Significant	Females (M=2.85)
Variable 2	.067	t-test	t(400) = 3.81	< .001	Significant	Females (M=3.42)
Variable 3	.511	t-test	t(400) = .687	.493	Non signifigant	Females (M=3.41)

APA format

- (r= .80; p < .001) – correlations parametric
- (r_T = .80; p <.05) - correlations non-parametric

- *(M = 4.43, SD = 0.90) – mean values*

- $F(1, 8) = 8{,}842$, $p > .001$

- ANOVA

Scale	Mean	SD	N
Leadership & Org Culture	3.78	0.10	226
Planning	3.79	0.11	226
Motivation	3.77	0.11	226
Control	3.78	0.20	226
Overall	3.78	0.13	226

ST SYRIL AND ST METHODIUS UNIVERSITY OF VELIKO TARNOVO FACULTY OF EDUCATION

THANK YOU FOR YOUR ATTENTION!

E-mail: kzlatkova@gmail.com

katerina@ts.uni-vt.bg