UNIVERSITY OF PÉCS
FACULTY OF HEALTH SCIENCES

PONGRÁC ÁCS

RESEARCH METHODOLOGY IN SPORT SCIENCES - WORKBOOK

University of Pécs, Faculty of Health Sciences

Institute of Physiotherapy and Sport Science

RESEARCH METHODOLOGY IN SPORT SCIENCES

Workbook

Pongrác ÁCS

Pécs, 2015

PÉCSI TUDOMÁNYEGYETEM
UNIVERSITY OF PÉCS

RESEARCH METHODOLOGY IN SPORT SCIENCES

Workbook

Editor: Dr. Pongrác Ács

PUBLISHED BY

UNIVERSITY OF PÉCS

FACULTY OF HEALTH SCIENCES

Reviewed by

Dr. habil Ferenc Ihász

Cover design and technical support by

Gábor Varga

Translated by

Kata Füge

ISBN 978-963-642-972-0

The manual has been produced in the framework of a project registered as TÁMOP-4.1.2. E-15/1/KONV-2015-0003

SZÉCHENYI 2020

MAGYARORSZÁG KORMÁNYA

Európai Unió
Európai Szociális Alap

BEFEKTETÉS A JÖVŐBE

CONTENTS

Foreword

It became apparent during the times since the first edition of our coursebook that the examples of the book – also discussed during class – are too few in number to serve as a basis for solid practical statistical knowledge . Furthermore, students' feedback made it clear that there is a need for an accompanying workbook, containing further examples.

In light of all these experiences and needs we prepared our electronic workbook, containing example exercises and databases to support acquiring knowledge on statistical theories explained by the coursebook. Apart from the exercises, the workbook also contains an answer key with results, thus calculations are easy to check. Descriptions of answers differ in case of multivariate exercises (factor- and cluster analysis) as we provide only one solution in detail, while – due to experts' heterogeneity – there might be several good answers as well. The numbering of exercises matches the tab numbers of the workbook's companion Excel file, thus databases may easily be found.

Kozármisleny, 8th August 2015

Pongrác Ács
author

EXERCISE 1.

The following table shows the number of visitors at Hungarian soccer matches.

Season	Number of Visitors' Mean	2010/2011= 100%	Previous Year = 100 %
2010/2011			
2011/2012	3812	138,17%	138,17%
2012/2013			72,46%
2013/2014	2969		
2014/2015		102,10%	

Source: www.hlsz.hu

Exercise (Excel):

- Count the missing data of the table.
- Present graphically the tendency of base-relatives.
- Present graphically and explain the tendency of chain-relatives.

EXERCISE 2.

The following table shows the average number of visitors at home matches of Hungarian premiere league teams.

Name of Team	Number of Visitors 2014/2015	Number of Visitors 2013/2014	Change Compared to Previous Year	Change in Number of Visitors (persons) Compared to Previous Year
Ferencváros	9000		99,87%	
Diósgyőr		4782		214
Újpest	4686	2561		
Győr	4079			1152
Debrecen	3468		46,86%	
Videoton	3143	3336		-193
Nyíregyháza	2671	No data	No data	No data
Haladás		2841	85,99%	
Paks		1273		510
Kecskemét	1751	1863		
Pápa	1516	1580		
Pécs	1253			-474
Honvéd	999		66,87%	
Puskás Akadémia	956	1045		-89
MTK		1347	54,34%	
Dunaújváros	443	No data	No data	No data

Source: www.hlsz.hu

Exercise (Excel):

- Count the missing data of the table.
- Present graphically the average number of visitors of Hungarian premiere league teams in the 2014/2015 season using a bar chart.

EXERCISE 3.

The following table presents the number of members in a sport association according to age, in four years' time.

Age	2010	2011	2012	2013	2014
-7	45	49	52	51	55
7,1-13,99	92	97	99	101	110
14,00-17,99	110	99	116	120	121
18,00-20,99	182	174	177	180	188
21,00-22,99	201	195	202	210	207
23-	211	201	222	198	211

Exercise (Excel):

- In which year was the average age the lowest?
- In which year was the average age in the sport association the highest?

EXERCISE 4.

The following table shows the distribution of workers at a sports equipment store according their wage.

Income (HUF/person)	Number of Employees (persons)
-60 000	1
60 001 - 80 000	3
80 001 - 100 000	10
100 001 - 120 000	24
120 001 - 140 000	35
140 001 - 160 000	18
160 001 -	9
Total	100

Exercise (Excel):

- Count the average wage of workers at the store.
- Describe mean values of wage according to their place in the line of data (median, mode).
- Describe wage using standard deviation and relative standard deviation.

EXERCISE 5.

Prepare the summarizing descriptive statistics based on the frequency row of the 'weight' variable (source: workbook.xlsx).

Exercise (Excel):

- Count the average weight of university students.
- How big is the total weight of the examined university students?
- Analyse the measures of skewness.

EXERCISE 6.

Examine the database of basketball players and analyse if there is a significant difference in the number of turnovers according to the player's role (point guard, defender, forward, power forward, center). Use the NBA database.

Exercise (SPSS):

- Calculate the average number and standard deviation of turnovers according to players' roles.
- Test and analyse homogeneity of standard deviations.
- Evaluate results (F; p).

EXERCISE 7.

Several scientific examinations concluded that there is a significant difference between men and women of the young adult age group (18-23 years) in body fat % category. Analyse whether this result is generalizable based on the examination that was carried out with 57 university students, using the primary database (source: fittségi57_adatbázis_alap_bmikat.sav).

Exercise (SPSS):

- Calculate the mean body fat % of men and women.
- Examine the equivalence of variances.
- Evaluate and explain the results.

EXERCISE 8.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (workbook_exercise8). Estimate values of standing long jump of girls and boys by 95% confidence.

Exercise (Excel):

- Calculate the mean values of boys and girls at standing long jump.
- Calculate the confidence intervals matching the mean values of boys and girls at standing long jump.
- Evaluate and explain the results.
- Show results on graphs.

EXERCISE 9.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: fittségi57_adatbázis_alap_bmikat.sav). Estimate values of standing long jump of girls and boys by 95% confidence.

Exercise (SPSS):

- Calculate the mean values of boys and girls at standing long jump.
- Calculate the confidence intervals matching the mean values of boys and girls at standing long jump.
- Evaluate and explain the results.
- Show results on graphs.

EXERCISE 10.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: fittségi57_adatbázis_alap_bmikat.sav). Examine whether drill run results of female university students are homogenous according to body fat % categories.

Exercise (SPSS):

- Prepare body fat % categories for females according to the intervals bellow, and label them as well.

Boys	Body fat % categories			
	underweight	healthy zone	exercise is required	increased exercise is required
	<16,4	16,5-31,3	31,4 -38,5,0	38,6<

- Give mean numerical values for the drill runs in each category.
- Examine whether there is a significant difference between the drill run values of various categories.
- Between which categories can you find differences (if any)? Describe numerically.

EXERCISE 11.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: workbook.xlsx). Analyse the relationship between height and long jump.

Exercise (Excel):

- Calculate and interpret the association coefficient of height and long jump.
- In what % does height determine the value of long jump?
- Calculate and interpret the regression coefficient.
- Give the formula for and interpret the regression line.
- Estimate the expected value of long jump of a student whose height is 190 cm.

EXERCISE 12.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: fittségi57_adatbázis_alap_bmikat.sav). Analyse the relationship between height and long jump.

Exercise (SPSS):

- Calculate and interpret the association coefficient of height and long jump.
- In what % does height determine the value of long jump?
- Calculate and interpret the regression coefficient.
- Give the formula for and interpret the regression line.
- Estimate the expected value of long jump of a student whose height is 190 cm.

EXERCISE 13.

We built a database of 53 motorbikes using a motorbike catalogue. We defined categories according to top speed and price, based on which we organised motorbikes into a contingency table (source: workbook. xlsx).

Exercise (Excel):

- Create the contingency table.
- Create the table for the frequencies under the assumption of independence.
- Describe numerically the strength of association using the Cramer coefficient.
- In what % does top speed determine the price?

EXERCISE 14.

We built a database of 53 motorbikes using a motorbike catalogue. Let's create categories for speed and price based on top speed and price, according to the following intervals:

Price	**Cheap**	**Medium**	**Expensive**
	<2 500 000	2 500 001- 4 000 000	>4 000 001
Speed	**Slow**	**Medium**	**Fast**
	<180	181-260	>261

Exercise (SPSS):

- Create the new variables containing the categories.
- Create the combinational table.
- Create the table for the frequencies under the assumption of independence.
- Describe numerically the strength of association using the Cramer coefficient.
- In what % does top speed determine the price?
- Examine between which categories is there an association definitely.

EXERCISE 15.

According to our hypothesis there is a gender-related difference in rhythmical sit-up results. Examine if there is significant difference between the rhythmical sit-up results of males and females (source: fittségi_57fő_alap_bmikat.xlsx or workbook.xlsx/ exercise 15.).

Exercise (Excel):

- Calculate mean values of males' and females' rhythmical sit-ups.
- Examine the equivalence of variances.
- Evaluate and explain the results (t, p).

EXERCISE 16.

In our examination we aimed to increase physical activity of patients living with diabetes, with group training carried out three times per week. Parallel to the increase of physical activity we anticipated the improvement in patients' blood glucose results as well. Examine whether there was any difference between the first and last measurement of blood glucose parameters by the group of 10 patients (source: workbook.xlsx/ exercise 16.).

Exercise (Excel):

- Calculate the mean values and variance of blood glucose levels.
- Calculate if there is any difference between the first and the least measurement.
- Evaluate the results.

EXERCISE 17.

One research examined the physical activity of respondents, sorting them into three groups: low-, medium-, and high activity groups. The researchers assumed that there was a difference between the weights of the members of the three groups. Examine if there is indeed a significant difference between the weight of the three activity groups (source: workbook.xlsx/ exercise 17.).

Exercise (Excel):

- Calculate the mean weight of the three groups and the matching variances.
- Examine if there is a difference between the mean weights of the groups, and interpret the results (F, p).

EXERCISE 18.

One research examined the nutritional behaviour and physical activity of respondents less than 18 years old. According to one of the hypotheses, there will be an association between gender and body shape. Examine if there is an association between gender and body shape (source: workbook.xlsx/ exercise 17.).

Exercise (Excel):

- Prepare a report and evaluate results according to the steps used in earlier exercises.

EXERCISE 19.

We built a secondary database using 2014/2015 regular season results of NBA players. The database contains the following variables (33 pieces) (source: NBA6.sav):

Number of Variable	Name of Variable	Meaning of Variable	Number of Variable	Name of Variable	Meaning of Variable
1	Number	Number of player in the database	21	Effectiveness_free _throw	Effectiveness of free-throws %
2	Name	Name of player	22	Offensive_rebounds	Number of offensive rebounds
3	Role	Role of player (PG: 1, SG: 2, SF: 3, PF: 4, C: 5)	23	Defesive_rebounds	Number of defensive rebounds
4	Age	Age of player	24	Total_rebounds	Number of total rebounds
5	Team	Team of player	25	Assists	Number of assists
6	Reagular season_matches	Number of completed matches in regular season (when the player played, maximum 82)	26	Interceptions	Number of interceptions
7	Regular season_starting _lineup	Number of completed matches in regular season as member of the starting lineup (maximum 82)	27	Blocks	Number of blocks
8	Total_time_regular_season	Total time on court at regular season matches	28	Personal_fouls	Number of personal fouls
9	Successful_field_goal_attempts	Number of successful field goal attempts (three-points included)	29	Turnovers	Number of turnovers
10	Field_goal_attempts	Number of field goal attempts (three-points included)	30	Points	Total scored points
11	Effectiveness_field_goals	Effectiveness of field goals %	31	Height	Body height
12	Goals_2p	Number of two-point goals	32	Weight	Weight
13	Attempts_3p	Number of three-point attempts	33	Wage	Wage (2014-15)
14	Effectiveness_3p	Effectiveness of three-points			
15	Successful_2p	Number of successful two-points			
16	Attempts_2p	Number of two-point attempts			
17	Effectiveness_2p	Effectiveness of two-points %			
18	Eff_field goal_effectiveness	Effective field goal effectiveness (FG + 0.5 * 3P)			
19	Successful_free_throws	Number of successful free throws			
20	Free_throw_attempts	Number of free_throw attempts			

Exercise (SPSS):

- Building on your professional opinion group the variables into four factors using factor analysis.
- Explain and interpret the four factors.

ANSWERS

ANSWER – EXERCISE 1.

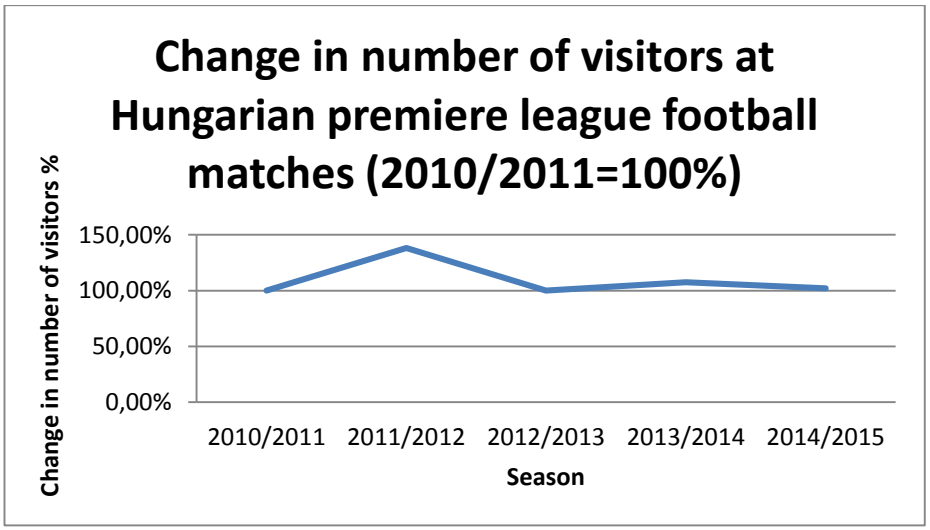
The following table shows the number of visitors at Hungarian soccer matches (source: workbook_excel.xlsx).

Exercise, answer (Excel):

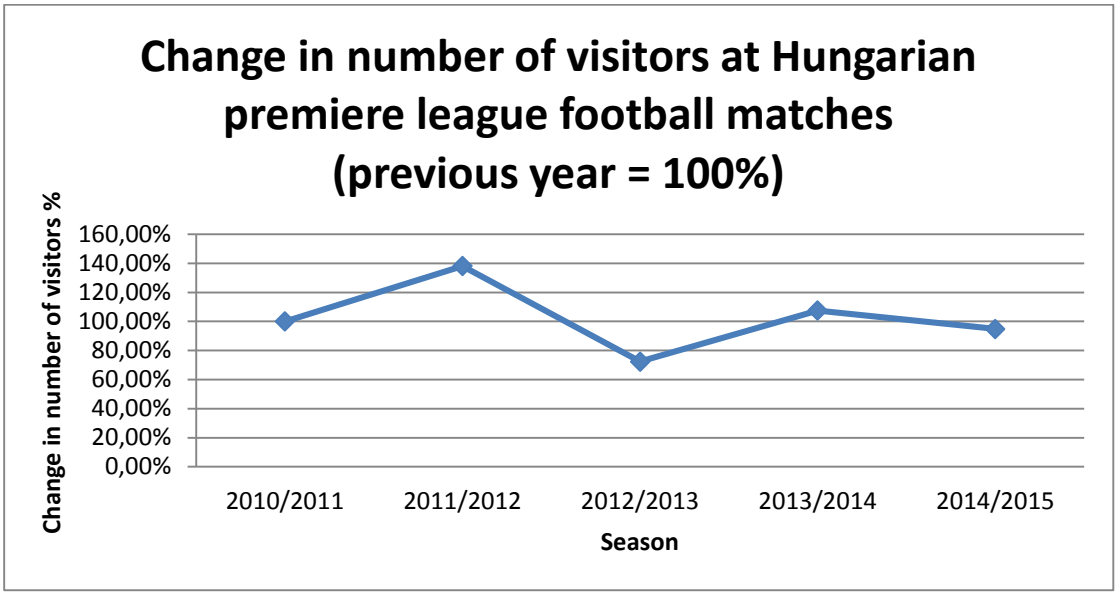
- Count the missing data of the table.

Season	Mean number of visitors	2010/2011= 100%	Previous year= 100 %
2010/2011	2759	100,00%	100,00%
2011/2012	3812	138,17%	138,17%
2012/2013	2762	100,11%	72,46%
2013/2014	2969	107,61%	107,49%
2014/2015	2817	102,10%	94,88%

- Present graphically the tendency of base-relatives.



- Present graphically and explain the tendency of chain-relatives.



ANSWER – EXERCISE 2.

The following table shows the average number of visitors at home matches of Hungarian premiere league teams (source: workbook_excel.xlsx).

Name of Team	Number of Visitors 2014/2015	Number of Visitors 2013/2014	Change Compared to Previous Year	Change in Number of Visitors (persons) Compared to Previous Year
Ferencváros	9000		99,87%	
Diósgyőr		4782		214
Újpest	4686	2561		
Győr	4079			1152
Debrecen	3468		46,86%	
Videoton	3143	3336		-193
Nyíregyháza	2671	No data	No data	No data
Haladás		2841	85,99%	
Paks		1273		510
Kecskemét	1751	1863		
Pápa	1516	1580		
Pécs	1253			-474
Honvéd	999		66,87%	
Puskás Akadémia	956	1045		-89
MTK		1347	54,34%	
Dunaújváros	443	No data	No data	No data

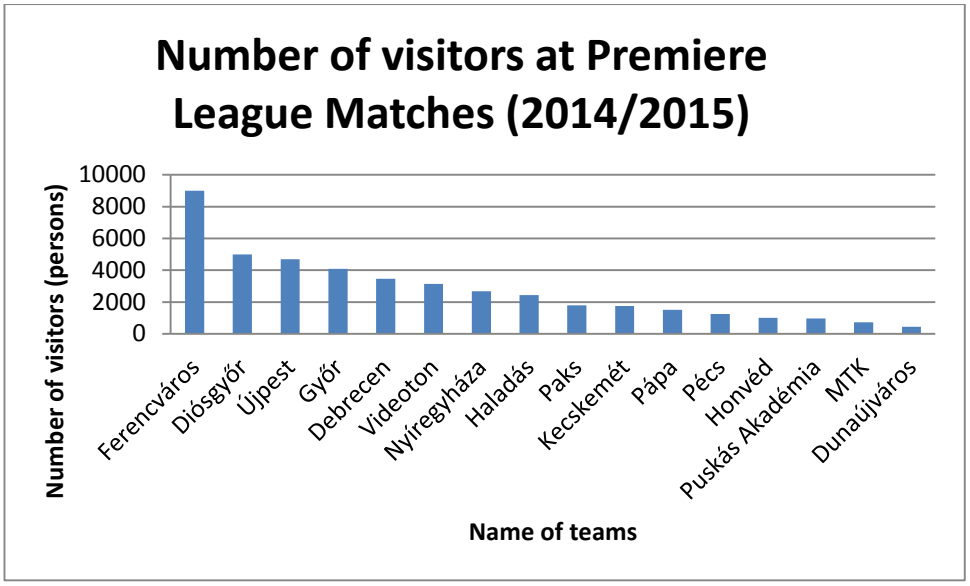
Source: www.hlsz.hu

Exercise (Excel):

- Count the missing data of the table.

Name of Team	Number of Visitors 2014/2015	Number of Visitors 2013/2014	Change Compared to Previous Year	Change in Number of Visitors (persons) Compared to Previous Year
Ferencváros	9000	9012	99,9%	-12
Diósgyőr	4996	4782	104,5%	214
Újpest	4686	2561	183,0%	2125
Győr	4079	2927	139,4%	1152
Debrecen	3468	7400	46,9%	-3932
Videoton	3143	3336	94,2%	-193
Nyíregyháza	2671			2671
Haladás	2443	2841	86,0%	-398
Paks	1783	1273	140,1%	510
Kecskemét	1751	1863	94,0%	-112
Pápa	1516	1580	95,9%	-64
Pécs	1253	1727	72,6%	-474
Honvéd	999	1494	66,9%	-495
Puskás Akadémia	956	1045	91,5%	-89
MTK	732	1347	54,3%	-615
Dunaújváros	443			

- Present graphically the average number of visitors of Hungarian premiere league teams in the 2014/2015 season using a bar chart.



ANSWER – EXERCISE 3.

The following table presents the number of members in a sport association according to age, in four years' time (source: workbook_excel.xlsx).

Age	2010	2011	2012	2013	2014
-7	45	49	52	51	55
7,1-13,99	92	97	99	101	110
14,00-17,99	110	99	116	120	121
18,00-20,99	182	174	177	180	188
21,00-22,99	201	195	202	210	207
23-	211	201	222	198	211

Exercise, answer (Excel):

mean	**2010**	**18,93**
mean	**2011**	**18,75**
mean	**2012**	**18,78**
mean	**2013**	**18,65**
mean	**2014**	**18,57**

- In which year was the average age the lowest?

The average age the lowest in 2014.

- In which year was the average age in the sport association the highest?

The average age in the sport association was the highest in 2010.

ANSWER – EXERCISE 4.

The following table shows the distribution of workers at a sports equipment store according their wage (source: workbook_excel. xlsx):

Wage (HUF/person)	Number of Employees (persons)
-60 000	1
60 001 - 80 000	3
80 001 - 100 000	10
100 001 - 120 000	24
120 001 - 140 000	35
140 001 - 160 000	18
160 001 -	9
Total	100

Exercise, answer (Excel):

Wage		**Persons (f)**	**x**	**fx**	**f'**
40 000	60 000	1	50 000	50 000	1
60 000	80 000	3	70 000	210 000	4
80 000	100 000	10	90 000	900 000	14
100 000	120 000	24	110 000	2 640 000	38
120 000	140 000	35	130 000	4 550 000	73
140 000	160 000	18	150 000	2 700 000	91
160 000	180 000	9	170 000	1 530 000	100

- Count the average wage of workers at the store.

The average wage of workers is 125 800 Ft.

- Describe mean values of wage according to their place in the line of data (median, mode).

Median wage of workers: 127 857 Ft.

Mode wage of workers: 126 857 Ft.

ANSWER – EXERCISE 5.

Prepare the summarizing descriptive statistics based on the frequency row of the 'weight' variable (source: workbook.xlsx).

Exercise, answer (Excel):

Weight	
Expected value	70,6245614
Standard error	1,626640835
Median	67,5
Mode	62
Variance	12,28086899
Variance of the sample	150,8197431
Kurtosis	7,70283507
Skewness	2,124865603
Range	74,7
Minimum	53,6
Maximum	128,3
Total	4025,6
Pieces	57
Confidence level (95,0%)	3,258553155

- Count the average weight of university students.

The average weight of university students is 70.63 kg.

- How big is the total weight of the examined university students?

The total weight of the examined university students is 4025.6 kg.

- Analyse the measures of skewness.

The positive value of the skewness measure indicates right-sided asymmetry.

ANSWER – EXERCISE 6.

Examine the database of basketball players and analyse if there is a significant difference in the number of turnovers according to the player's role (point guard, defender, forward, power forward, center). Use the database NBA.

Exercise (SPSS):

- Calculate the average number and standard deviation of turnovers according to players' roles.

Descriptives

Number of Turnovers

	N	Mean	Variance	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Point Guard	**19**	**119,11**	**62,724**	14,390	88,87	149,34	3	251
Guard	**24**	**90,75**	**66,490**	13,572	62,67	118,83	4	254
Forward	**20**	**107,60**	**67,906**	15,184	75,82	139,38	3	240
Power Forward, Forward	**21**	**93,71**	**54,514**	11,896	68,90	118,53	0	184
Center	**16**	**110,06**	**67,235**	16,809	74,24	145,89	0	222
Total	100	103,22	63,478	6,348	90,62	115,82	0	254

- Test and analyse homogeneity of standard deviations.

Test of Homogeneity of Variances

number of turnovers

Levene Statistic	df1	df2	Sig.
,347	4	95	**,845**

Variances are equal.

- Evaluate results (F; p).

ANOVA

number of turnovers

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	11556,847	4	2889,212	**,709**	**,588**
Within Groups	387356,313	95	4077,435		
Total	398913,160	99			

Based on the p (0.59) matching the calculated F value (0.7) we may state that there is no significant difference between players' roles and turnovers.

ANSWER – EXERCISE 7.

Several scientific examinations concluded that there is a significant difference between men and women of the young adult age group (18-23 years) in body fat % category. Analyse whether this result is generalisable based on the examination that was carried out with 57 university students, using the primary database (source: fittségi57_adatbázis_alap_bmikat.sav).

Exercise, answer (SPSS):

- Calculate the man body fat % of men and women.

Group Statistics

	Gender	Number of participants	Mean	Variance
Body fat (%)	male	28	17,5321	3,95615
	female	29	30,4759	7,77646

- Examine the equivalence of variances.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Body fat (%)	Equal variances assumed	2,499	,120	-7,877	55	,000	-12,94372	1,64331	-16,23698	-9,65046
	Equal variances not assumed			-7,960	41,901	,000	-12,94372	1,62612	-16,22559	-9,66185

- Evaluate and explain the results.

Based on the examination we may state that the difference between the body fat % of male and female respondents is significant, thus it can not only be due to chance.

ANSWER – EXERCISE 8.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (workbook_exercise8). Estimate values of standing long jump of girls and boys by 95% confidence.

Exercise, answer (Excel):

	Male	Female
Mean	234,25	182,48
Variance	17,97039747	21,51016
Number of participants	28	29
Standard error	3,396085905	3,994337
Error limit	6,656328373	7,828901
Lower error limit	227,59	174,65
Upper error limit	240,91	190,31

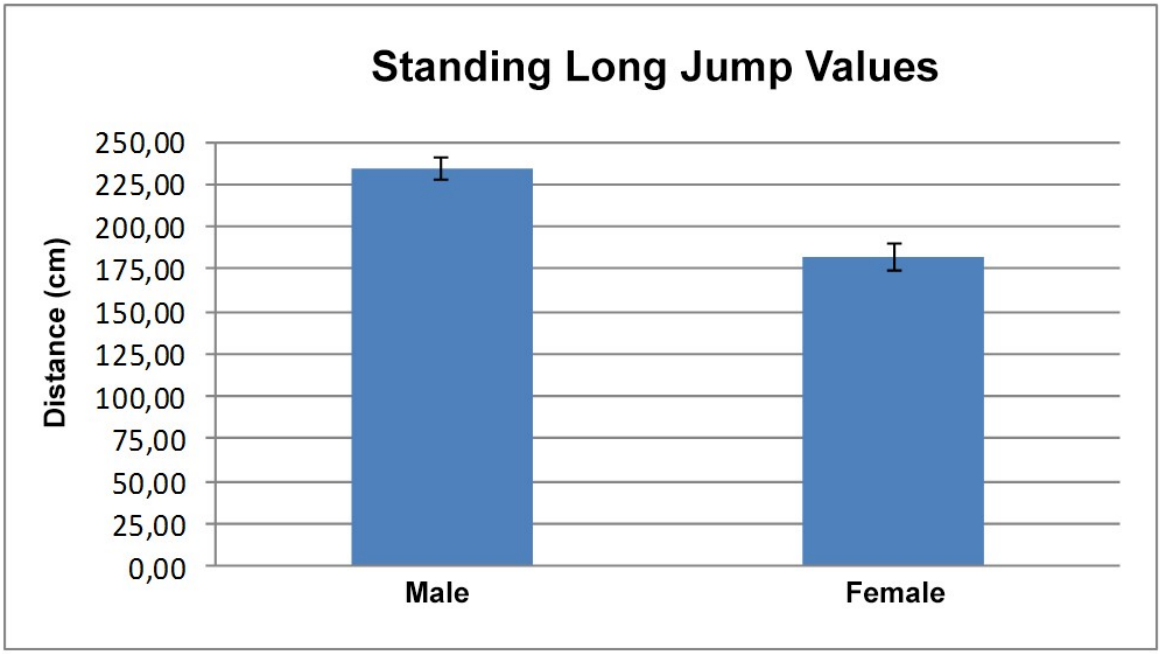
- Calculate the mean values of boys and girls at standing long jump.

Mean value of boys at standing long jump is 234.25 cm; that of girls is 182. 48 cm.

- Calculate the confidence intervals matching the mean values of boys and girls at standing long jump.

Confidence interval matching the mean values of boys is 6.65 cm,; that of girls is 7.83 cm.

- Evaluate and explain the results.
- Show results on graphs.


Standing Long Jump Values
Distance (cm)
250,00
225,00
200,00
175,00
150,00
125,00
100,00
75,00
50,00
25,00
0,00
Male
Female

ANSWER – EXERCISE 9.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: fittségi57_adatbázis_alap_bmikat.sav). Estimate values of standing long jump of girls and boys by 95% confidence.

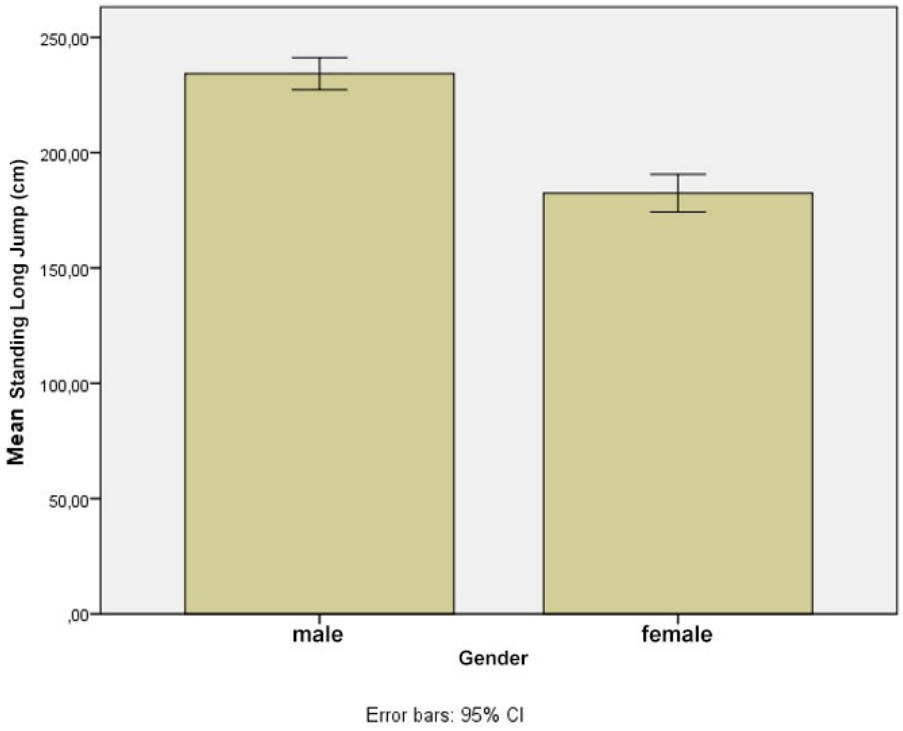
Exercise, answer (SPSS):

- Calculate the mean values of boys and girls at standing long jump.
- Calculate the confidence intervals matching the mean values of boys and girls at standing long jump.
- Evaluate and explain the results.

Descriptives

	Gender			Statistic	Std.
Standing long jump (cm)	Male	Mean		234,2500	3,39609
		95% Confidence Interval for Mean	Lower Bound	227,2818	
			Upper Bound	241,2182	
		5% Trimmed Mean		234,2937	
		Median		237,5000	
		Variance		322,935	
		Std. Deviation		17,97040	
		Minimum		203,00	
		Maximum		265,00	
		Range		62,00	
		Interquartile Range		24,50	
		Skewness		-,186	,441
		Kurtosis		-,787	,858
	Female	Mean		182,4828	3,99434
		95% Confidence Interval for Mean	Lower Bound	174,3007	
			Upper Bound	190,6648	
		5% Trimmed Mean		184,2586	
		Median		184,0000	
		Variance		462,687	
		Std. Deviation		21,51016	
		Minimum		120,00	
		Maximum		213,00	
		Range		93,00	
		Interquartile Range		19,00	
		Skewness		-1,542	,434
		Kurtosis		3,540	,845

- Show results on graphs.



ANSWER – EXERCISE 10.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: fittségi57_adatbázis_alap_bmikat.sav). Examine whether drill run results of female university students are homogenous according to body fat % categories.

Exercise, answer (SPSS):

- Prepare body fat % categories for females according to the intervals bellow, and label them as well.

Boys	Body fat % categories			
	underweight	healthy zone	exercise is required	increased exercise is required
	<16,4	16,5-31,3	31,4 -38,5,0	38,6<

- Give mean numerical values for the drill runs in each category.

Descriptives

drill runs (number of 20 meter full runs)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
healthy zone	20	60,7500	13,44727	3,00690	54,4565	67,0435	34,00	80,00
exercise is required	6	44,6667	15,09525	6,16261	28,8252	60,5082	24,00	61,00
increased exercise is required	3	16,3333	4,04145	2,33333	6,2938	26,3729	12,00	20,00
Total	29	52,8276	19,14059	3,55432	45,5469	60,1083	12,00	80,00

- Examine whether there is a significant difference between the drill run values of various categories.

Test of Homogeneity of Variances

drill runs (number of 20 meter full runs)

Levene Statistic	df1	df2	Sig.
1,776	2	26	,189

ANOVA

drill runs (number of 20 meter full runs)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5650,388	2	2825,194	15,942	,000
Within Groups	4607,750	26	177,221		
Total	10258,138	28			

- Between which categories can you find differences (if any)? Describe numerically.

Multiple Comparisons

Dependent Variable: drill runs (number of 20 meter full runs)

						95% Confidence Interval	
	(I) bodyfat_cat	(J) bodyfat_cat	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Scheffe	healthy zone	exercise is required	16,08333	6,19661	,050	-,0016	32,1683
		increased exercise is required	44,41667*	8,24226	,000	23,0216	65,8117
	exercise is required	healthy zone	-16,08333	6,19661	,050	-32,1683	,0016
		increased exercise is required	28,33333*	9,41332	,021	3,8985	52,7682
	increased exercise is required	healthy zone	-44,41667*	8,24226	,000	-65,8117	-23,0216
		exercise is required	-28,33333*	9,41332	,021	-52,7682	-3,8985

Bonferroni	healthy zone	exercise is required	16,08333*	6,19661	,046	,2266	31,9401
		increased exercise is required	44,41667*	8,24226	,000	23,3252	65,5081
	exercise is required	healthy zone	-16,08333*	6,19661	,046	-31,9401	-,2266
		increased exercise is required	28,33333*	9,41332	,017	4,2452	52,4215
	increased exercise is required	healthy zone	-44,41667*	8,24226	,000	-65,5081	-23,3252
		exercise is required	-28,33333*	9,41332	,017	-52,4215	-4,2452

*. The mean difference is significant at the 0.05 level.

ANSWER – EXERCISE 11.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: workbook.xlsx). Analyse the relationship between height and long jump.

Exercise, answer (Excel):

- Calculate and interpret the association coefficient of height and long jump.

The association has medium strength (R=0.48).

- In what % does height determine the value of long jump?

Height determines the value of long jump in 23.21%.

SUMMARY TABLE	
REGRESSION STATISTICS	
R VALUE	0,481815
R-SQUARE	0,232146
CORRECTED R-SQUARE	0,218185
STANDARD ERROR	28,90627
NUMBER OF SURVEYS	57

- Calculate and interpret the regression coefficient.

b1=1.96, which means that 1 cm growth in height will increase the long jump value by 1.96 cm.

	COEFFICIENT	STANDARD ERROR	T VALUE	P VALUE	LOWER 95%	UPPER 95%	LOWER 95,0%	UPPER 95,0%
INTERSECTION	-133,105	83,71602	-1,58996	0,117576	-300,876	34,66559	-300,876	34,66559
HEIGHT	1,956812	0,479873	4,077769	0,000148	0,995124	2,918499	0,995124	2,918499

- Give the formula for and interpret the regression line.

y=-133.1+1.96x

- Estimate the expected value of long jump of a student whose height is 190 cm.

239.3=-133.1+1.96*190

ANSWER – EXERCISE 12.

We examined the fitness of 57 university students using the tests defined by the Hungarian National Student Fitness Test (source: fittségi57_adatbázis_alap_bmikat.sav). Analyse the relationship between height and long jump.

Exercise, answer (SPSS):

- Calculate and interpret the association coefficient of height and long jump

Correlations

		Standing long jump (cm)	Height (cm)
Pearson Correlation	Standing long jump (cm)	1,000	**,482**
	Height (cm)	,482	1,000
Sig. (1-tailed)	Standing long jump (cm)	.	,000
	Height (cm)	,000	.
N	Standing long jump (cm)	57	57
	Height (cm)	57	57

- In what % does height determine the value of long jump?

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	**,482**[a]	**,232**	,218	28,90627

a. Predictors: (Constant), Height (cm)

- Calculate and interpret the regression coefficient.
- Give the formula for and interpret the regression line.

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	**-133,105**	83,716		-1,590	,118
Height (cm)	**1,957**	,480	,482	4,078	,000

a. Dependent Variable: Standing long jump (cm)

- Estimate the expected value of long jump of a student whose height is 190 cm.

239.3=-133.1+1.96*190

ANSWER – EXERCISE 13.

We built a database of 53 motorbikes using a motorbike catalogue. We defined categories according to top speed and price, based on which we organised motorbikes into a contingency table (source: workbook. xlsx).

Exercise, answer (Excel):

- Create the contingency table.

	Expensive	**Medium**	**Cheap**	**Total price**
fast	2	10		12
medium	5	12	8	25
slow	4	3	6	13
Total price	**11**	**25**	**14**	**50**

- Create the table for the frequencies under the assumption of independence.

	Expensive	**Medium**	**Cheap**	**Total price**
fast	2,64	6	3,36	12
medium	5,5	12,5	7	25
slow	2,86	6,5	3,64	13
Total price	**11**	**25**	**14**	**50**

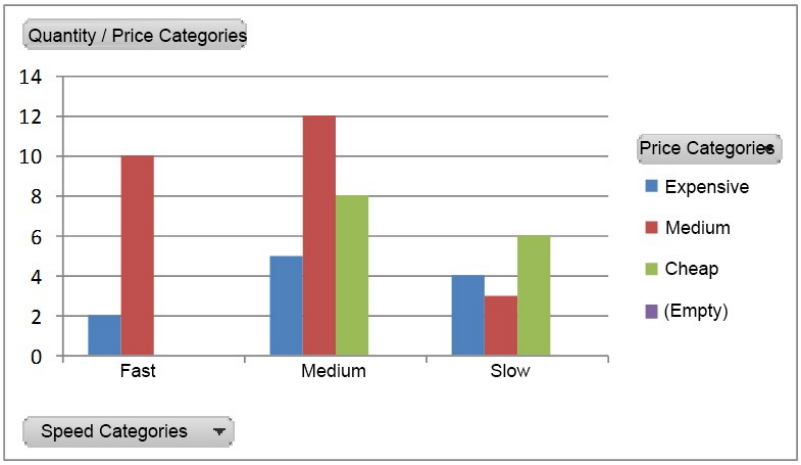
- Describe numerically the strength of association using the Cramer coefficient.

	Expensive	**Medium**	**Cheap**	**Total price**
fast	0,16	2,67	3,36	12
medium	0,05	0,02	0,14	25
slow	0,45	1,88	1,53	13
Total price	**11**	**25**	**14**	**50**

khi	10,26
cramer	0,32

- In what % does top speed determine the price?

In 10.24 %.



ANSWER – EXERCISE 14.

We built a database of 53 motorbikes using a motorbike catalogue. Let's create categories for speed and price based on top speed and price, according to the following intervals:

Price	Cheap	Medium	Expensive
	<2 500 000	2 500 001- 4 000 000	>4 000 001
Speed	Slow	Medium	Fast
	<180	181-260	>261

Exercise, answer (SPSS):

- Create the new variables containing the categories.
- Create the combinational table.
- Create the table for the frequencies under the assumption of independence.

Speed Category * Price Category Crosstabulation

			Price Categories			
			Cheap	Medium	Expensive	Total
Speed Cat.	Slow	Count	6	3	4	13
		Expected Count	3,6	6,5	2,9	13,0
		% within speed cat.	46,2%	23,1%	30,8%	100,0%
		% within proce cat.	42,9%	12,0%	36,4%	26,0%
		% of Total	12,0%	6,0%	8,0%	26,0%
		Adjusted Residual	1,7	-2,3	,9	
		Count	8	12	5	25
		Expected Count	7,0	12,5	5,5	25,0
		% within speed cat.	32,0%	48,0%	20,0%	100,0%
		% within proce cat.	57,1%	48,0%	45,5%	50,0%
		% of Total	16,0%	24,0%	10,0%	50,0%
		Adjusted Residual	,6	-,3	-,3	
		Count	0	10	2	12
		Expected Count	3,4	6,0	2,6	12,0
		% within speed cat.	0,0%	83,3%	16,7%	100,0%
		% within proce cat.	0,0%	40,0%	18,2%	24,0%
		% of Total	0,0%	20,0%	4,0%	24,0%
		Adjusted Residual	-2,5	2,6	-,5	
Total		Count	14	25	11	50
		Expected Count	14,0	25,0	11,0	50,0
		% within speed cat.	28,0%	50,0%	22,0%	100,0%
		% within proce cat.	100,0%	100,0%	100,0%	100,0%
		% of Total	28,0%	50,0%	22,0%	100,0%

- Describe numerically the strength of association using the Cramer coefficient.

The Cramer coefficient shows that the association between the two variable is of medium strength (c=0.32).

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	,453	,036
	Cramer's V	,320	,036
N of Valid Cases		50	

- In what % does top speed determine the price?

In 10.24 %.

- Examine between which categories is there an association definitely.

Between those indicated by blue colour.

ANSWER – EXERCISE 15.

According to our hypothesis there is a gender-related difference in rhythmical sit-up results. Examine if there is significant difference between the rhythmical sit-up results of males and females (source: fittségi_57fő_alap_bmikat.xlsx or workbook.xlsx/ exercise 15.).

Feladat, megoldás (Excel):

- Calculate mean values of males' and females' rhythmical sit-ups.

Mean values of males' rhythmical sit-ups is 76(±11.89); that of girls' is 60(±20.03).

- Examine the equivalence of variances.

According to the result of the F test, variances are not equal.

- Evaluate and explain the results (t, p).

Based on the two-sample t-test and the alienation examination there is significant difference between the rhythmical sit-up values of males and females (p<0.001).

ANSWER – EXERCISE 16.

In our examination we aimed to increase the level of physical activity of patients living with diabetes, with group training carried out three times per week. Parallel to the increase of physical activity we anticipated the improvement in patients' blood glucose results as well. Examine whether there was any difference between the first and last measurement of blood glucose parameters by the group of 10 patients (source: workbook.xlsx/ exercise 16.).

Exercise, answer (Excel):

- Calculate the mean values and variance of blood glucose levels.

The mean blood glucose level of the group was 8.84 (±2.02) at the first measurement, and 7.02(±1.45) at the last measurement.

- Calculate if there is any difference between the first and the least measurement.

Based on the result of the paired t-test, there was a significant improvement in the blood glucose parameters of the group ($p<0.001$).

ANSWER – EXERCISE 17.

One research examined the physical activity of respondents, sorting them into three groups: low-, medium-, and high activity groups. The researchers assumed that there was a difference between the weights of members of the three groups. Examine if there is indeed a significant difference between the weight of the three activity groups (source: workbook.xlsx/ exercise 17.).

Exercise, answer (Excel):

- Calculate the mean weight of the three groups and the matching variances.

Mean weight of the group with low physical activity is 74.93(±13.46); that of medium physical activity is 67.25(±19.89); that of high physical activity is 63.19(±10.61) kg.

- Examine if there is a difference between the mean weight of the groups, and interpret the results (F, p).

Based on the result of the variance analysis, there is significant difference between the weight of the three groups ($F=11.06$, $p<0.001$).

ANSWER – EXERCISE 18.

One research examined the nutritional behaviour and physical activity of respondents less than 18 years old. According to one of the hypotheses, there will be an association between gender and body shape. Examine if there is an association between gender and body shape (source: workbook.xlsx/ exercise 17.).

Exercise, answer (Excel):

- Prepare a report and evaluate results according to the steps used in earlier exercises.

According to the result of the chi-square test, there is no significant difference between gender and body shape ($p>0.05$).

ANSWER – EXERCISE 19.

We built a secondary database using 2014/2015 regular season results of NBA players. The database contains the following variables (33 pieces) (source: NBA6.sav):

Number of Variable	Name of Variable	Meaning of Variable	Number of Variable	Name of Variable	Meaning of Variable
1	Number	Number of player in the database	21	Effectiveness_free_throw	Effectiveness of free-throws %
2	Name	Name of player	22	Offensive_rebounds	Number of offensive rebounds
3	Role	Role of player (PG: 1, SG: 2, SF: 3, PF: 4, C: 5)	23	Defesive_rebounds	Number of defensive rebounds
4	Age	Age of player	24	Total_rebounds	Number of total rebounds
5	Team	Team of player	25	Assists	Number of assists
6	Reagular season_matches	Number of completed matches in regular season (when the player played, maximum 82)	26	Interceptions	Number of interceptions
7	Regular season_starting_lineup	Number of completed matches in regular season as member of the starting lineup (maximum 82)	27	Blocks	Number of blocks
8	Total_time_regular_season	Total time on court at regular season matches	28	Personal_fouls	Number of personal fouls
9	Successful_field_goal_attempts	Number of successful field goal attempts (three-points included)	29	Turnovers	Number of turnovers
10	Field_goal_attempts	Number of field goal attempts (three-points included)	30	Points	Total scored points
11	Effectiveness_field_goals	Effectiveness of field goals %	31	Height	Body height
12	Goals_2p	Number of two-point goals	32	Weight	Weight
13	Attempts_3p	Number of three-point attempts	33	Wage	Wage (2014-15)
14	Effectiveness_3p	Effectiveness of three-points			
15	Successful_2p	Number of successful two-points			
16	Attempts_2p	Number of two-point attempts			
17	Effectiveness_2p	Effectiveness of two-points %			
18	Eff_field goal_effectiveness	Effective field goal effectiveness (FG + 0.5 * 3P)			
19	Successful_free_throws	Number of successful free throws			
20	Free_throw_attempts	Number of free_throw attempts			

Exercise, answer (SPSS):

- Building on your professional opinion group the variables into four factors using factor analysis.

Descriptive Statistics

	Mean	Variance	Number of Participants
Age	27,02	4,359	97
Wage	4186655,6701	4925604,87059	97
Weight	98,2472	13,61402	97
Points (pts)	497,98	370,261	97
Number of personal fouls	68,74	52,804	97
Total time on court at regular season matches	1271,18	798,860	97
Number of assists	107,42	105,363	97
Number of interceptions	40,39	33,249	97
Number of turnovers	105,07	63,103	97
Number of completed matches in regular season	54,66	24,525	97
Number of defensive rebounds	166,18	124,226	97
Effective field goal effectiveness	,4816	,06932	97
Effectiveness of field goals %	,4348	,07319	97
Effectiveness of two-points %	,4628	,07321	97
Effectiveness of free-throws %	,6914	,19935	97
Effectiveness of three-points %	,2701	,16068	97
Number of blocks	21,59	23,769	97
Number of offensive rebounds	54,58	54,928	97

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,825
Bartlett's Test of Sphericity	Approx. Chi-Square	1910,785
	df	153
	Sig.	,000

Communalities

	Initial	Extraction
Age	1,000	,675
Wage	1,000	,696
Weight	1,000	,648
Points (pts)	1,000	,854
Number of personal fouls	1,000	,888
Total time on court at regular season matches	1,000	,940
Number of assists	1,000	,765
Number of interceptions	1,000	,778
Number of turnovers	1,000	,801
Number of completed matches in regular season	1,000	,791
Number of defensive rebounds	1,000	,889
Effective field goal effectiveness	1,000	,881
Effectiveness of field goals %	1,000	,899
Effectiveness of two-points %	1,000	,824
Effectiveness of free-throws %	1,000	,553
Effectiveness of three-points %	1,000	,638
Number of blocks	1,000	,728
Number of offensive rebounds	1,000	,810

Extraction Method: Principal Component Analysis.

Rotated Component Matrixa

	Component			
	1	2	3	4
Number of personal fouls	,939			
Total time on court at regular season matches	,921			
Points (pts)	,884			
Number of interceptions	,874			
Number of assists	,848			
Number of turnovers	,778	,389		
Number of completed matches in regular season	,741	,482		
Number of defensive rebounds	,695	,368	,470	
Effectiveness of free-throws %	,457		-,415	,335
Effective field goal effectiveness		,926		
Effectiveness of field goals %		,879	,347	
Effectiveness of two-points %		,814	,314	
Effectiveness of three-points %			-,767	
Number of blocks	,325	,384	,684	
Number of offensive rebounds	,372	,454	,669	
Weight	-,423		,625	
Age				,804
Wage	,362			,717

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

- Explain and interpret the four factors.

Factor 1	**Factor 2**	**Factor 3**	**Factor 4**
Number of personal fouls	Effective field goal effectiveness	Effectiveness of free-throws %	Age
Total time on court at regular season matches	Effectiveness of field goals %	Effectiveness of three-points %	Wage
Points (pts)	Effectiveness of two-points %	Number of blocks	
Number of interceptions		Number of offensive rebounds	
Number of assists		Weight	
Number of turnovers			
Number of completed matches in regular season			
Number of defensive rebounds			
Player utility indicator	**Field effectiveness feature**	**Extra added value indicator**	**Maturing indicator**

THE MANUAL HAS BEEN PRODUCED IN THE FRAMEWORK OF A PROJECT REGISTERED AS
TÁMOP-4.1.2. E-15/1/KONV-2015-0003

SZÉCHENYI 2020
MAGYARORSZÁG KORMÁNYA
Európai Unió
Európai Szociális Alap
BEFEKTETÉS A JÖVŐBE