




Korelačná Analýza 2

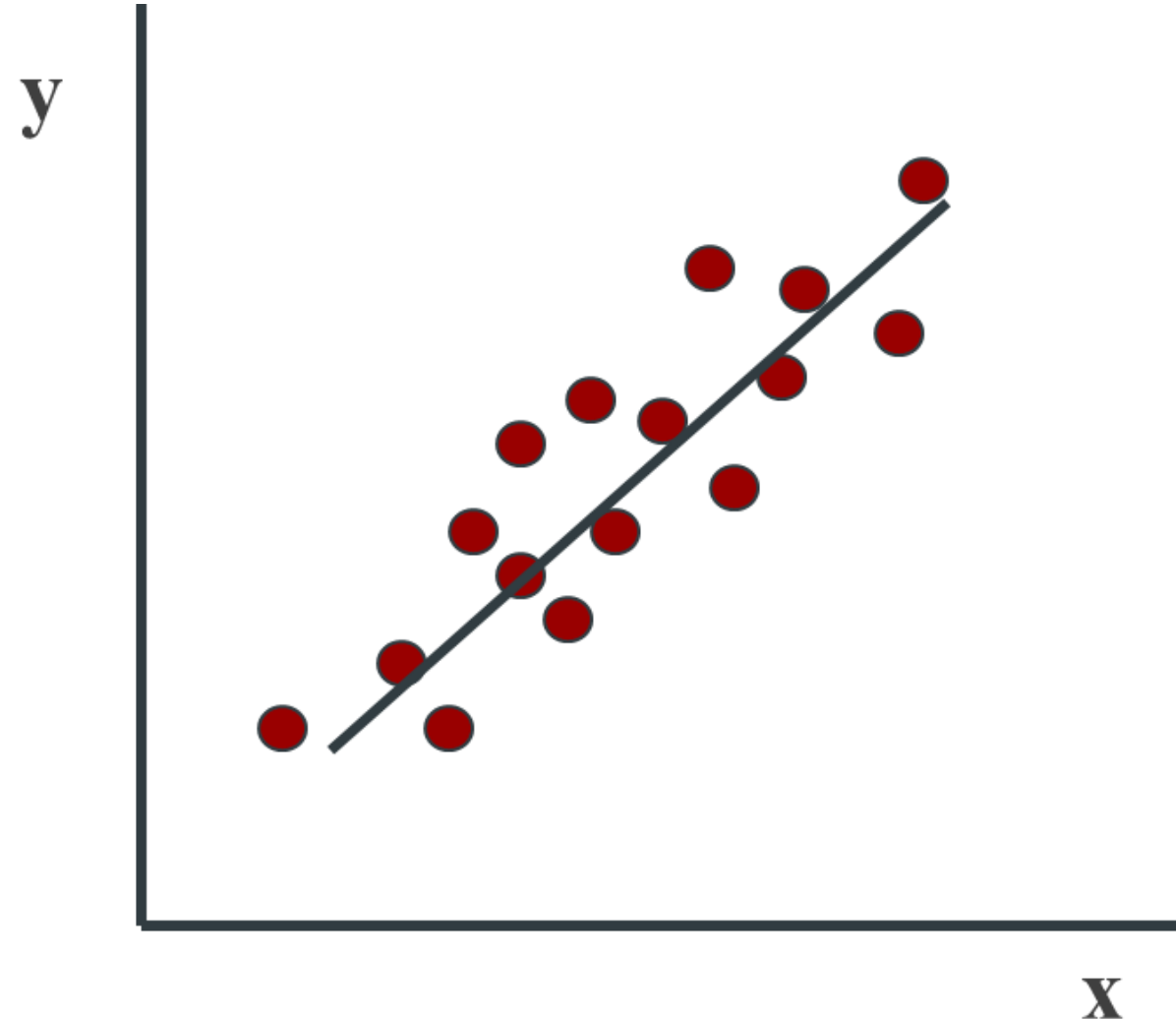


Použitie Korel. Koeficientov

Korelácia	Typ vzťahu	Použitie
 Pearson (r)	Lineárny vzťah	Výkonnosť serverov vs. doba odozvy, reklamné výdavky vs. predaj, mzdy vs. spokojnosť zamestnancov, počet chýb v kóde vs. skúsenosti programátora, dĺžka pracovného času vs. produktivita, čas strávený na stránke vs. konverzný pomer
 Spearman (ρ)	Monotónny vzťah	Poradie stránok vo výsledkoch vyhľadávania vs. návštevnosť, pozícia produktu v rebríčku predajnosti vs. predajnosť, hodnotenie influencerov vs. úspešnosť kampaní, šanca na povýšenie vs. pracovný výkon, preferované benefity vs. spokojnosť zamestnancov, počet testovacích prípadov vs. stabilita softvéru
 Kendall (τ)	Poradie hodnôt	Poradie kandidátov HR vs. ich prijatie, ranking produktov podľa recenzií vs. predaj, firemné oddelenia podľa spokojnosti vs. fluktuácia, najpopulárnejšie programovacie jazyky vs. počet nových vývojárov, poradie bugov podľa závažnosti vs. počet nahlásení, poradie sociálnych médií podľa angažovanosti vs. počet reklám

Korelačná Analýza

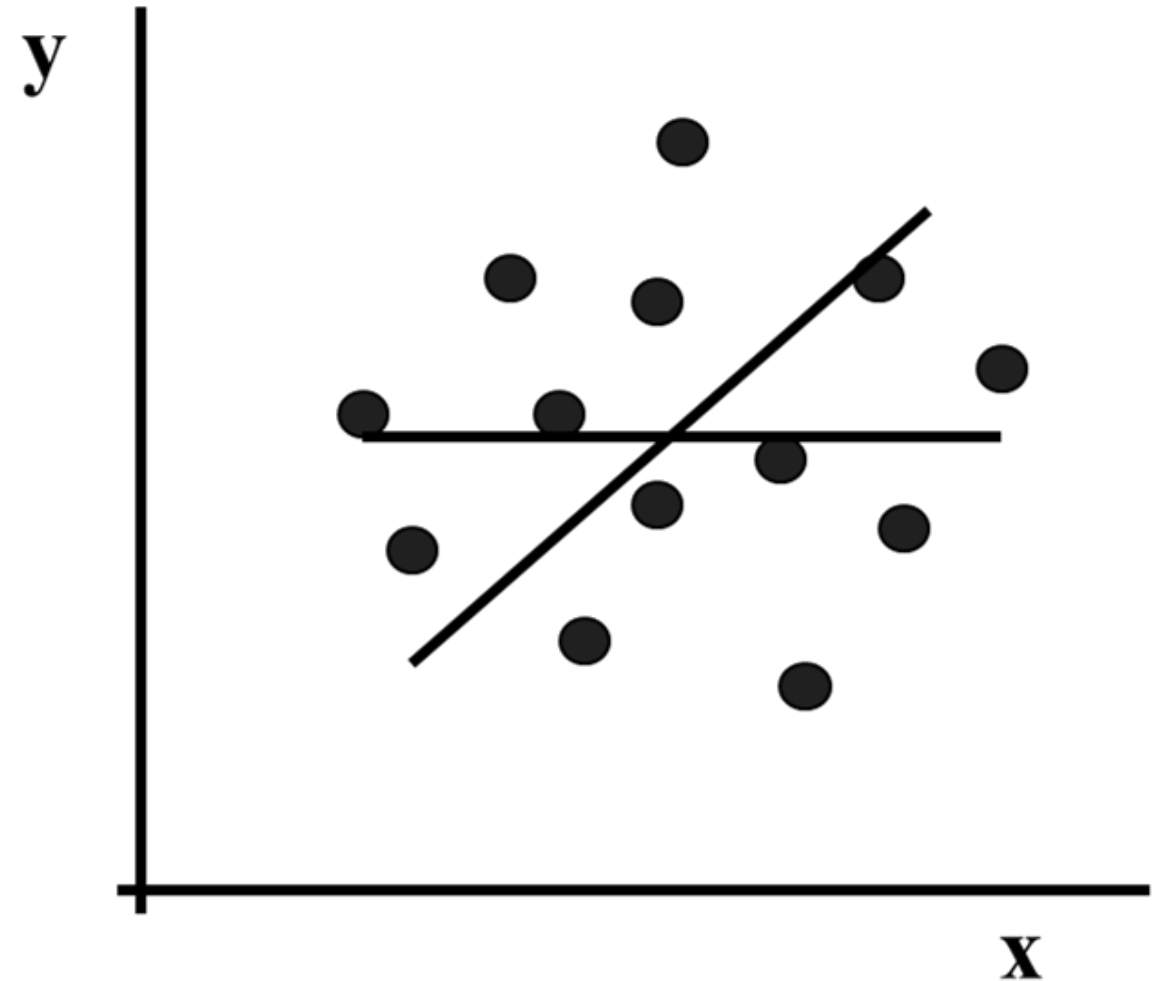
- Overenie **vypovedacej schopnosti kvantifikovaných regresných modelov** ako **celku**, aj **jeho častí**
- Výpočet **číselných charakteristík**, ktoré v koncentrovanej forme popisujú **kvalitu vypočítaných modelov**
- Požadujeme od nich, aby sa **pohybovali v pevne ohraničenom intervale**
- V rámci **intervalu rástli s vyššou silou závislosti**
- Porovnanie **2 prípadov závislosti**



Korelačná Analýza

Miery tesnosti štatistickej závislosti:

- 1. Kovariancia** – COV_{yx}
 - len pre lineárnu závislosť
- 2. Koeficient Korelácie** r_{yx}
 - len pre lineárnu závislosť
- 3. Koeficient Determinácie** r_{yx}^2
 - len pre lineárnu závislosť
4. index korelácie i_{yx}
5. index determinácie i_{yx}^2

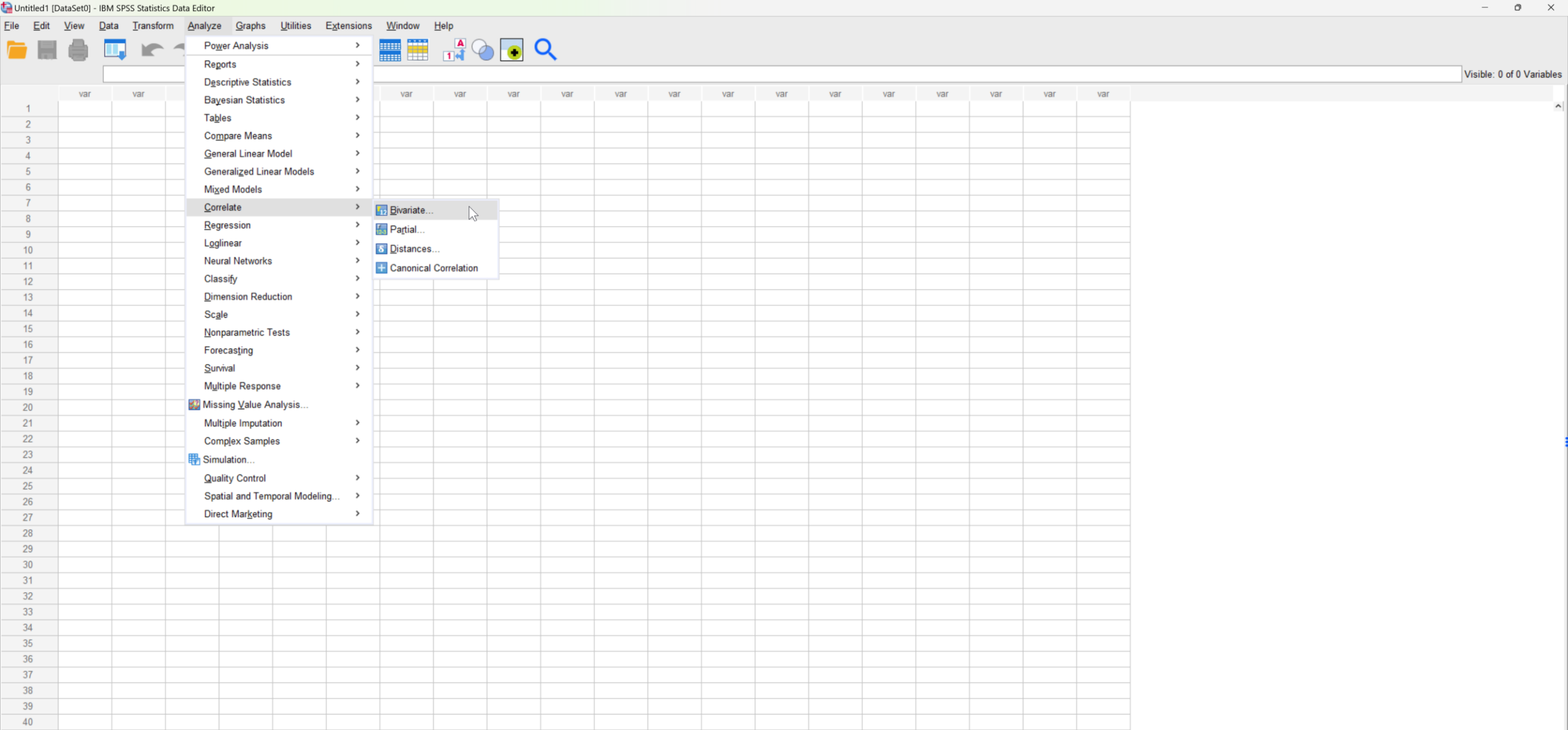


Kovariancia

- **Štatistická miera**, ktorá sa **používa**, aby sme **potvrdili** alebo **vyvrátili existenciu lineárnej závislosti (korelácie)**
- Zo spôsobu výpočtu možno odvodiť, kedy **potvrdzuje existenciu pozitívnej, negatívnej korelácie** a kedy **nekorelovanosti**
- Ak kovariancia, **potvrdí neexistenciu lineárneho vzťahu**
- **premennými môže existovať nelineárny vzťah** medzi
- Ak **kovariancia potvrdí existenciu lineárneho vzťahu**, môžeme merať jeho intenzitu

$$\text{cov } xy = \frac{1}{n} \sum \left(x_i - \bar{x} \right) \left(y_i - \bar{y} \right)$$

$\text{cov } xy = 0$, medzi prem. **nie je** lineárny vzťah
 $\text{cov } xy > 0$, medzi prem. **je priamy** lineárny vzťah
 $\text{cov } xy < 0$, medzi prem. **je nepriamy** lineárny vzťah



Analyze > Correlate

Prehľad Metód Korelácie IBM SPSS

Metóda	Kedy ju použiť?	Postup v SPSS
Bivariate correlation (Dvojrozmerná korelácia)	Keď chceš zistiť vzťah medzi dvoma premennými (napr. výška a hmotnosť).	<i>Analyze</i> → <i>Correlate</i> → <i>Bivariate...</i>
Partial correlation (Čiastočná korelácia)	Keď chceš zistiť vzťah medzi dvoma premennými po odstránení vplyvu tretej premennej (napr. korelácia výšky a hmotnosti po eliminácii vplyvu veku).	<i>Analyze</i> → <i>Correlate</i> → <i>Partial...</i>
Multiple correlation (Mnohonásobná korelácia)	Keď chceš analyzovať vzťah viacerých nezávislých premenných k jednej závislej premennej (napr. vplyv veku, výšky a stravy na hmotnosť).	<i>Analyze</i> → <i>Regression</i> → <i>Linear...</i>
Canonical correlation (Kanonická korelácia)	Keď chceš skúmať vzťah medzi dvoma skupinami premenných (napr. ako akademické výsledky (testy, GPA) súvisia so psychologickými faktormi (motivácia, stres)).	<i>Analyze</i> → <i>Regression</i> → <i>Canonical Correlation</i> (iba v rozšírených moduloch SPSS)
Spearmanova a Kendallova korelácia	Keď premenné nemajú normálne rozdelenie alebo sú ordinálne (poradové údaje) .	<i>Analyze</i> → <i>Correlate</i> → <i>Bivariate...</i> (zaškrtnúť Spearman alebo Kendall)
Point-biserial correlation (Bodovo-biseriálna korelácia)	Keď jedna premenná je dichotomická (0/1, áno/nie) a druhá je spojitá (napr. vzťah medzi pohlavím a výškou).	<i>Analyze</i> → <i>Correlate</i> → <i>Bivariate...</i> (Použiť Pearsonov koeficient)
Phi coefficient (Φ koeficient)	Keď obe premenné sú dichotomické (0/1, muž/žena, prešiel/neprešiel) .	<i>Analyze</i> → <i>Descriptive Statistics</i> → <i>Crosstabs...</i> (zaškrtnúť <i>Phi and Cramér's V</i>)
Factor analysis (Faktorová analýza)	Keď chceš zistiť, či viaceré premenné možno zoskupiť do menšieho počtu faktorov (napr. otázky v dotazníku mapujú rôzne dimenzie osobnosti).	<i>Analyze</i> → <i>Dimension Reduction</i> → <i>Factor...</i>
Cluster analysis (Klastrová analýza)	Keď chceš rozdeliť súbor údajov na homogénne skupiny na základe podobnosti (napr. segmentácia zákazníkov na základe správania).	<i>Analyze</i> → <i>Classify</i> → <i>Cluster...</i>
Structural Equation Modeling (SEM - Štrukturálne modelovanie rovníc)	Keď chceš modelovať komplexné vzťahy medzi viacerými premennými naraz (napr. ako socioekonomický status ovplyvňuje vzdelanie a príjem).	Používa sa v rozšírených moduloch IBM SPSS AMOS

C:\Users\miros\Desktop\testovacie_data.xlsx

Worksheet: Testovacie Data [A1:E31]

Range:

☒ Read variable names from first row of data☒ Percentage of values that determine data type: 95☒ Ignore hidden rows and columns☐ Remove leading spaces from string values☐ Remove trailing spaces from string values

Preview

	ID	Hodiny_studia	Skore_testu	Vek_studentov	Spokojnost_studentov
1	1	7	88	29	1
2	2	15	67	19	8
3	3	11	53	27	3
4	4	8	74	21	3
5	5	7	63	25	1
6	6	19	99	24	5
7	7	11	58	29	7
8	8	11	75	26	9
9	9	4	51	25	7
10	10	8	69	22	9
11	11	3	77	19	8
12	12	2	96	22	2



Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500 columns.

OK






Paste

Reset

Cancel

Help

Variables:

 ID
 Hodiny_studia
 Skore_testu
 Vek_studentov
 Spokojnost_studentov

[Options...](#)[Style...](#)[Bootstrap...](#)[Confidence interval...](#)

Correlation Coefficients

☒ Pearson ☐ Kendall's tau-b ☐ Spearman

Test of Significance

☒ Two-tailed ☐ One-tailed

☒ Flag significant correlations ☐ Show only the lower triangle ☒ Show diagonal

OK

[Paste](#)[Reset](#)[Cancel](#)[Help](#)

➔ Correlations

Correlations

		ID	Hodiny_studia	Skore_testu	Vek_studentov	Spokojnost_studentov
ID	Pearson Correlation	1	,134	-,024	,003	,052
	Sig. (2-tailed)		,481	,899	,988	,783
	N	30	30	30	30	30
Hodiny_studia	Pearson Correlation	,134	1	-,088	-,062	-,051
	Sig. (2-tailed)	,481		,644	,745	,791
	N	30	30	30	30	30
Skore_testu	Pearson Correlation	-,024	-,088	1	,046	-,068
	Sig. (2-tailed)	,899	,644		,810	,719
	N	30	30	30	30	30
Vek_studentov	Pearson Correlation	,003	-,062	,046	1	-,175
	Sig. (2-tailed)	,988	,745	,810		,355
	N	30	30	30	30	30
Spokojnost_studentov	Pearson Correlation	,052	-,051	-,068	-,175	1
	Sig. (2-tailed)	,783	,791	,719	,355	
	N	30	30	30	30	30

Descriptive Statistics

	Mean	Std. Deviation	N
ID	15,50	8,803	30
Hodiny_studia	9,93	5,388	30
Skore_testu	72,70	16,110	30
Vek_studentov	24,77	3,461	30
Spokojnost_studentov	5,07	2,766	30

Correlations

		ID	Hodiny_studia	Skore_testu	Vek_studentov	Spokojnost_studentov
ID	Pearson Correlation	--				
	Sum of Squares and Cross-products	2247,500				
	Covariance	77,500				
	N	30				
Hodiny_studia	Pearson Correlation	,134	--			
	Sig. (2-tailed)	,481				
	Sum of Squares and Cross-products	184,000	841,867			
	Covariance	6,345	29,030			
Skore_testu	Pearson Correlation	-,024	-,088	--		
	Sig. (2-tailed)	,899	,644			
	Sum of Squares and Cross-products	-99,500	-221,600	7526,300		
	Covariance	-3,431	-7,641	259,528		
Vek_studentov	Pearson Correlation	,003	-,062	,046	--	
	Sig. (2-tailed)	,988	,745	,810		
	Sum of Squares and Cross-products	2,500	-33,467	73,900	347,367	
	Covariance	,086	-1,154	2,548	11,978	
Spokojnost_studentov	Pearson Correlation	,052	-,051	-,068	-,175	--
	Sig. (2-tailed)	,783	,791	,719	,355	
	Sum of Squares and Cross-products	37,000	-21,867	-88,400	-48,533	221,867
	Covariance	1,276	-,754	-3,048	-1,674	7,651
	N	30	30	30	30	30



Bivariate Correlations: Options



Statistics



Means and standard deviations



Cross-product deviations and covariances

Missing Values



Exclude cases pairwise



Exclude cases listwise

Continue

Cancel

Help



Confidence interval

☐ Estimate confidence interval of bivariate correlation parameter

Confidence interval (%)

Pearson Correlation

☐ Apply the bias adjustment

Spearman Correlation

Estimate the variance by the method proposed by:

- ☒ Fieller, Hartley and Pearson
- ☐ Bonett and Wright
- ☐ Caruso and Cliff

[Continue](#)

[Cancel](#)

[Help](#)



☐ Perform bootstrapping

Number of samples:

☐ Set seed for Mersenne Twister

Seed:

Confidence Intervals

Level(%):

- ☒ Percentile
- ☐ Bias corrected accelerated (BCa)

Sampling

- ☒ Simple
- ☐ Stratified

Variables:

☐ ID
☐ Hodiny_studia
☐ Skore_testu
☐ Vek_studentov
☐ Spokojnost_stude...



Strata Variables:

[Continue](#)

[Cancel](#)

[Help](#)

Po úvodnom grafickom preskúmaní nastupuje fáza hľadania presných štatistík, ktoré potvrdia odhady z grafov. Pre tieto účely používame **štatistické nástroje** korelačnej analýzy.

Korelačné štatistiky zisťujú **či medzi premennými existuje korelácia aká je sila korelácie. Koreláciou nazývame vzájomný lineárny vzťah – závislosť dvoch premenných.** Tento vzťah môže byť:

priamy – s rastúcimi hodnotami jednej premennej rastú aj hodnoty druhej premennej

nepriamy – s rastúcimi hodnotami jednej premennej klesajú hodnoty druhej premennej

Ak medzi hodnotami dvoch premenných neexistuje ani priama ani nepriama lineárna závislosť, hovoríme, že sú **nekorelované**.

