

שיטות להרצת רגרסיה רבת משתנים

קיימות 3 שיטות להרצת הנתונים ברגרסיה מרובה, הנבדלות ביניהן בסדר הכנסת המשתנים הב"ת:

שיטת ENTER

רגרסיה בצעדים חכמים (STEPWISE REGRESSION)

רגרסיה היררכית

שיטת ENTER

מכניסים את כל המשתנים הב"ת למשוואה בעת ובעונה אחת.

דוגמה 2: ניבוי היקף פעולות סניף הבנק ביצוא באמצעות 5 משתנים- היקף פעולות בבורסה, יבוא, חסכונות, עמלות והמחאות נוסעים (כל המשתנים באלפי ש"ח ליום).

ניתוח הנתונים לפי שיטת Enter:

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	X5 Traveler Checks, X1 Stock Market, X4 Commission, X2 Import, X3 Saving	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y Export

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.506 ^a	.256	.221	64.75156

a. Predictors: (Constant), X5 Traveler Checks, X1 Stock Market, X4 Commission, X2 Import, X3 Saving

b. Dependent Variable: Y Export

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	155428.3	5	31085.659	7.414	.000 ^a
	Residual	452818.5	108	4192.764		
	Total	608246.8	113			

a. Predictors: (Constant), X5 Traveler Checks, X1 Stock Market, X4 Commission, X2 Import, X3 Saving

b. Dependent Variable: Y Export

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-259.207	71.279		-3.637	.000
	X1 Stock Market	-2.00E-02	.089	-.020	-.225	.823
	X2 Import	.257	.103	.221	2.490	.014
	X3 Saving	.312	.128	.219	2.447	.016
	X4 Commission	-2.57E-02	.107	-.021	-.240	.811
	X5 Traveler Checks	103.806	29.454	.315	3.524	.001

a. Dependent Variable: Y Export

? על סמך הפלטים הנ"ל, ענה ללא חישוב:

1. האם משוואת הרגרסיה לניבוי פעילות סניף הבנק ביצוא על ידי חמש המשתנים הב"ת מובהקת באוכלוסיה ברמת מובהקות של 0.05?
2. לאיזה מהמשתנים הב"ת תרומה ייחודית מובהקת לניבוי פעילות סניף הבנק ביצוא, ברמת מובהקות של 0.05?
3. מי מבין המשתנים הב"ת תורם את התרומה היחסית הגבוהה ביותר לניבוי פעילות סניף הבנק ביצוא?
4. מהי פרופורציית השונות המוסברת של פעילות סניף הבנק ביצוא ע"י חמשת המשתנים הב"ת?
5. מהו האומדן לשונות הטעויות בניבוי פעילות סניף הבנק ביצוא?

רגרסיה בצעדים חכמים (stepwise)

מטרתה לבנות את משוואת הרגרסיה האופטימאלית. כלומר לבחור את המשתנים הב"ת אשר תרומתם הייחודית לניבוי המשתנה התלוי היא הגבוהה ביותר ומובהקת.

בכל צעד/שלב יכנס משתנה ב"ת אחד בלבד.

בשלב הראשון: יכנס המשתנה הב"ת שיש לו את המתאם הפשוט הגבוה ביותר עם Y ומובהק.

בשלב השני: יכנס המשתנה הב"ת שיש לו את המתאם החלקי הגבוה ביותר עם Y (בניכוי המשתנה הב"ת שנכנס בשלב הראשון למשוואה) ומובהק.

בשלב השלישי: יכנס המשתנה הב"ת שיש לו את המתאם החלקי הגבוה ביותר עם Y (בניכוי המשתנים שנכנסו בשלב הראשון והשני למשוואת הרגרסיה) ומובהק.

וכך הלאה... במידה ובכל שלב, המשתנה בעל המתאם הגבוה ביותר איננו מובהק עוצרים ולא ממשיכים להכניס משתנים.

הדגמה על פי נתוני דוגמא 2 : בניית משוואת הרגרסיה האופטימאלית לניבוי פעילות הבנק ביצוא בשיטת stepwise :

שלב ראשון -בחירת המשתנה הב"ת תעשה מתוך טבלת המתאמים:

Correlations^a

		Y Export	X1 Stock Market	X2 Import	X3 Saving	X4 Commission	X5 Traveler Checks
Y Export	Pearson Correlation Sig. (2-tailed)	1 .	.126 .182	.337** .000	.283** .002	.137 .147	.395** .000
X1 Stock Market	Pearson Correlation Sig. (2-tailed)	.126 .182	1 .	.189* .044	.342** .000	.019 .841	.094 .318
X2 Import	Pearson Correlation Sig. (2-tailed)	.337** .000	.189* .044	1 .	.161 .086	.207* .027	.282** .002
X3 Saving	Pearson Correlation Sig. (2-tailed)	.283** .002	.342** .000	.161 .086	1 .	.109 .249	.119 .206
X4 Commission	Pearson Correlation Sig. (2-tailed)	.137 .147	.019 .841	.207* .027	.109 .249	1 .	.281** .002
X5 Traveler Checks	Pearson Correlation Sig. (2-tailed)	.395** .000	.094 .318	.282** .002	.119 .206	.281** .002	1 .

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

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

a. Listwise N=114

ניתן לראות כי למשתנה "המחאות נוסעים" יש את המתאם הגבוה ביותר עם פעילות הבנק ביבוא (ומובהק)-

$$r_{y5} = 0.395$$

שימו לב כי טבלת המקדמים, המתארת את המתאמים ה"פשוטים" בין Y לבין ה-Xים, טובה רק לשלב הראשון. בשלבים הבאים נבחרים ה-Xים בעלי המתאם החלקי הגבוה ביותר עם Y. את המתאמים החלקיים ניכל לראות בפלט המכונה-excluded variables:

Excluded Variables^d

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	X1 Stock Market	.089 ^a	1.025	.308	.097	.991
	X2 Import	.245 ^a	2.793	.006	.256	.920
	X3 Saving	.239 ^a	2.816	.006	.258	.986
	X4 Commission	.028 ^a	.305	.761	.029	.921
2	X1 Stock Market	.011 ^b	.118	.906	.011	.880
	X2 Import	.216 ^b	2.496	.014	.232	.904
	X4 Commission	.008 ^b	.093	.926	.009	.915
3	X1 Stock Market	-.019 ^c	-.213	.832	-.020	.864
	X4 Commission	-.020 ^c	-.229	.820	-.022	.900

a. Predictors in the Model: (Constant), X5 Traveler Checks

b. Predictors in the Model: (Constant), X5 Traveler Checks, X3 Saving

c. Predictors in the Model: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

d. Dependent Variable: Y Export

ניתן לראות לפי ה-Partial Correlation כי למשתנה "חיסכון" יש את המתאם החלקי הגבוה ביותר עם פעילות הבנק ביצוא (0.258) ומובהק (sig=0.006) ולכן הוא יכנס בשלב השני לתוך מודל הרגרסיה.

באותו אופן ניתן לראות לפי ה-Partial Correlation כי למשתנה יבוא יש את המתאם החלקי הגבוה ביותר עם פעילות הבנק ביצוא (0.232) ומובהק (sig=0.014) ולכן הוא יכנס בשלב השלישי לתוך משוואת הרגרסיה.

בשלב הבא המשתנה בעל המתאם החלקי הגבוה ביותר עם פעילות הבנק ביצוא-"עמלות" (-0.022) כבר לא מובהק (sig=0.820) ולכן לא ייכנס לתוך משוואת הרגרסיה והפרוצדורה תיעצר.

הפלט הבא מציג את המשתנים שנכנסו למשוואת הרגרסיה בכל שלב בשיטת ה-stepwise:

Variables Entered/Removed^d

Model	Variables Entered	Variables Removed	Method
1	X5 Traveler Checks	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	X3 Saving	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	X2 Import	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Y Export

? האם ניתן לדעת, ללא חישוב, האם המשוואה לניבוי פעילות הבנק ביצוא המתקבלת בשלב הסופי על פי שיטת ה-stepwise מובהקת?

? על פי פלט ה-model summary מה יהיה ערכו של הסטטיסטי F לבדיקת מובהקות המשוואה הסופית?

רגרסיה היררכית

החוקר הוא זה שמחליט אילו משתנים יכנסו לתוך משוואת הרגרסיה בכל שלב במטרה לבדוק את השערותיו. בכל שלב יכולים להיכנס כמה משתנים ב"ת ביחד וממילא המשוואה הסופית או מקדמי השיפוע לא חייבים להיות מובהקים (בניגוד לשיטת ה-stepwise).

נדגים על סמך נתוני דוגמא מס' 2. החוקר משער כי המשתנים חיסכון, עמלות ופעילות בבורסה יתרמו תרומה מובהקת לניבוי פעילות הבנק ביצוא מעבר לתרומת המשתנים המחאות נוסעים ויבוא.

בכדי לבדוק את השערותו צריך החוקר לבנות משוואת רגרסיה בשני צעדים/שלבים. בשלב הראשון להכניס את המשתנים - המחאות נוסעים ויבוא ואילו בשלב השני להכניס את שלושת המשתנים הנותרים:

Variables Entered/Removed^d

Model	Variables Entered	Variables Removed	Method
1	X5 Traveler Checks, X2 Import	.	Enter
2	X3 Saving, X4 Commission, X1 Stock Market	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y Export

לפי טבלת המקדמים ניתן לראות כי בניגוד לשיטת ה-stepwise לא כל שיפועי המשתנים הב"ת בשלב האחרון מובהקים:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-231.425	70.532		-3.281	.001
	X2 Import	.285	.102	.245	2.793	.006
	X5 Traveler Checks	107.619	28.975	.326	3.714	.000
2	(Constant)	-259.207	71.279		-3.637	.000
	X2 Import	.257	.103	.221	2.490	.014
	X5 Traveler Checks	103.806	29.454	.315	3.524	.001
	X1 Stock Market	-2.00E-02	.089	-.020	-.225	.823
	X3 Saving	.312	.128	.219	2.447	.016
	X4 Commission	-2.57E-02	.107	-.021	-.240	.811

a. Dependent Variable: Y Export

השוואה בין שיטת stepwise והיררכית

לעיתים אנו מתבקשים לבחון האם הנתונים עובדו בשיטת stepwise או בשיטה ההיררכית.

קיימים כמה אינדיקציות העוזרות להבחין בין השיטות:

1. ב-stepwise נכנס משתנה ב"ת אחד בכל שלב ואילו בהיררכית לא בהכרח.
2. ב-stepwise המשתנה הראשון שנכנס צריך להיות בעל המתאם הפשוט הגבוה ביותר עם המשתנה התלוי ואילו בהיררכית לא בהכרח.
3. ב-stepwise התוספת של כל משתנה ב"ת שנכנס חייבת להיות מובהקת מעבר למה שכבר מצוי במשוואה ואילו בהיררכית לא בהכרח.
4. ב-stepwise כל מקדמי השיפוע במשוואה הסופית חייבים להיות מובהקים ואילו בהיררכית לא בהכרח.
5. ב-stepwise ה-MSRES צריך לרדת משלב לשלב ואילו ה- $AdjR^2$ צריך לעלות משלב לשלב..

? על סמך הפלטים הבאים הערך באיזו שיטה נותחו הנתונים, stepwise או היררכית? תן לפחות 3 אינדיקציות התומכות בערכתך.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.395 ^a	.156	.149	67.68720	.156	20.760	1	112	.000
2	.461 ^b	.213	.198	65.68529	.056	7.931	1	111	.006
3	.505 ^c	.255	.235	64.19061	.042	6.229	1	110	.014

- a. Predictors: (Constant), X5 Traveler Checks
- b. Predictors: (Constant), X5 Traveler Checks, X3 Saving
- c. Predictors: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

ANOVA^d

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	95112.382	1	95112.382	20.760	.000 ^a
	Residual	513134.4	112	4581.557		
	Total	608246.8	113			
2	Regression	129330.9	2	64665.464	14.988	.000 ^b
	Residual	478915.9	111	4314.557		
	Total	608246.8	113			
3	Regression	154999.0	3	51666.334	12.539	.000 ^c
	Residual	453247.8	110	4120.435		
	Total	608246.8	113			

- a. Predictors: (Constant), X5 Traveler Checks
- b. Predictors: (Constant), X5 Traveler Checks, X3 Saving
- c. Predictors: (Constant), X5 Traveler Checks, X3 Saving, X2 Import
- d. Dependent Variable: Y Export

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-221.003	72.540		-3.047	.003
	X5 Traveler Checks	130.447	28.630	.395		
2	(Constant)	-254.426	71.388		-3.564	.001
	X5 Traveler Checks	121.042	27.983	.367		
	X3 Saving	.341	.121	.239		
3	(Constant)	-259.691	69.795		-3.721	.000
	X5 Traveler Checks	102.038	28.387	.309		
	X3 Saving	.301	.119	.211		
	X2 Import	.251	.100	.216		

- a. Dependent Variable: Y Export

Excluded Variables^d

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	X1 Stock Market	.089 ^a	1.025	.308	.097	.991
	X2 Import	.245 ^a	2.793	.006	.256	.920
	X3 Saving	.239 ^a	2.816	.006	.258	.986
	X4 Commission	.028 ^a	.305	.761	.029	.921
2	X1 Stock Market	.011 ^b	.118	.906	.011	.880
	X2 Import	.216 ^b	2.496	.014	.232	.904
	X4 Commission	.008 ^b	.093	.926	.009	.915
3	X1 Stock Market	-.019 ^c	-.213	.832	-.020	.864
	X4 Commission	-.020 ^c	-.229	.820	-.022	.900

a. Predictors in the Model: (Constant), X5 Traveler Checks

b. Predictors in the Model: (Constant), X5 Traveler Checks, X3 Saving

c. Predictors in the Model: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

d. Dependent Variable: Y Export

Correlations^a

		Y Export	X1 Stock Market	X2 Import	X3 Saving	X4 Commission	X5 Traveler Checks
Y Export	Pearson Correlation	1	.126	.337**	.283**	.137	.395**
	Sig. (2-tailed)	.	.182	.000	.002	.147	.000
X1 Stock Market	Pearson Correlation	.126	1	.189*	.342**	.019	.094
	Sig. (2-tailed)	.182	.	.044	.000	.841	.318
X2 Import	Pearson Correlation	.337**	.189*	1	.161	.207*	.282**
	Sig. (2-tailed)	.000	.044	.	.086	.027	.002
X3 Saving	Pearson Correlation	.283**	.342**	.161	1	.109	.119
	Sig. (2-tailed)	.002	.000	.086	.	.249	.206
X4 Commission	Pearson Correlation	.137	.019	.207*	.109	1	.281**
	Sig. (2-tailed)	.147	.841	.027	.249	.	.002
X5 Traveler Checks	Pearson Correlation	.395**	.094	.282**	.119	.281**	1
	Sig. (2-tailed)	.000	.318	.002	.206	.002	.

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

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

a. Listwise N=114

****הערה:** שימו לב כי ניתן לשלב בין השיטות, כלומר שצעדים מסוימים יתבצעו בשיטת stepwise ואילו

אחרים בהיררכית. כך שלמעשה יש לבחון כל שלב ושלב ולראות באיזה שיטה נעשה.

התוספת לניבוי ומובהקותה

כאשר מתבצעת רגרסיה בצעדים (בין אם בשיטה ההיררכית או ב-stepwise) ניתן לחשב את התוספת לניבוי של משתנה/משתנים ב"ת שנכנסו בשלב מסוים אל תוך משוואת הרגרסיה על פני משתנה/משתנים ב"ת אחרים שכבר היו במשוואה. ניתן בנוסף לבדוק האם תוספת זו לניבוי של המשתנה התלוי מובהקת באוכלוסייה.

חישוב התוספת לניבוי

יתבצע ע"י הנוסחה:

$$R_{all}^2 - R_{part}^2 = \text{התוספת המבוקשת}$$

↙
↘

כל המשתנים הב"ת משתנים שהיו כבר במשוואה

לדוגמא:

? על סמך נתוני דוגמא מס' 2:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.395 ^a	.156	.149	67.68720	.156	20.760	1	112	.000
2	.461 ^b	.213	.198	65.68529	.056	7.931	1	111	.006
3	.505 ^c	.255	.235	64.19061	.042	6.229	1	110	.014

- a. Predictors: (Constant), X5 Traveler Checks
- b. Predictors: (Constant), X5 Traveler Checks, X3 Saving
- c. Predictors: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

א. חשב את התוספת לניבוי פעילות הבנק בייצוא של יבוא על פני פעילות בחשבונות חיסכון והמחאות נוסעים.

ב. חשב את התוספת לניבוי פעילות הבנק ביצוא של פעילות בחשבונות חיסכון ויבוא על פני המחאות נוסעים.

מבחן מובהקות התוספת

מטרתנו לבדוק האם התוספת לניבוי של משתנה/משתנים ב"ת מסוימים על פני אחרים מובהקת באוכלוסיה.

נדגים את שלבי המבחן על בדיקת מובהקות התוספות שחושבו בסעיפים א' ו-ב' של השאלה הקודמת.

השערות:

רשומות במונחי הבטה/בטות של התוספת, כאשר בהערה מצוינים המשתנים המצויים במשוואה (part).

למשל, עבור סעיף א' של השאלה הקודמת (תוספת של יבוא על פני חיטון והמחאות נוסעים):

$$\begin{aligned} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{aligned}$$

כאשר המשתנים חיטון והמחאות נוסעים מצויים במשוואה.

עבור סעיף ב' (תוספת של חיטון ויבוא על פני המחאות נוסעים):

$$\begin{aligned} H_0: \beta_2 = \beta_3 = 0 \\ H_1: otherwise \end{aligned}$$

כאשר המשתנה המחאות נוסעים מצוי במשוואה.

הנחות:

אותן שלושת הנחות של מודל רגרסיה מרובה. ההנחות מתייחסות לכל המשתנים הב"ת (ל-all).

לדוגמא, ההנחות של בדיקת מובהקות התוספת של סעיף א' וב' של השאלה הקודמת:

1. דגימה מקרית של לקוחות בנק/טעויות ב"ת.

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0 \quad i \neq j$$

2. $\varepsilon_i \approx N(0, \sigma_\varepsilon^2)$ עבור כל קומבינציה ליניארית קבועה של המחאות נוסעים, יבוא ו-פעילות בחשבונות חיטון.

3. אין מתאם מלא בין המחאות נוסעים, יבוא ו-פעילות בחשבונות חיטון.

סטטיסטי המבחן:

$$F = \frac{(R_{all}^2 - R_{part}^2) / (K_{all} - K_{part})}{(1 - R_{all}^2) / (n - K_{all} - 1)} \stackrel{H_0}{\sim} F_{k_{all} - k_{part}, n - k_{all} - 1}$$

כלל הכרעה ומסקנה:

אם $F < F_{k_{all}-k_{part}, n-k_{all}-1}$, דוחים את H_0 ויש עדות למובהקות התוספת באוכלוסיה.

? צור משוואת רגרסיה אופטימלית לניבוי פעילות הבנק ביצוא באמצעות המשתנים הבאים: פעילות בחשבונות חיסכון, יבוא והמחאות נוסעים. איזה מהמשתנים יכנס בשלב הראשון? איזה מהמשתנים יכנס בשלב השני? פטור על פי טבלת המקדמים בלבד:

Correlations^a

		Y Export	X1 Stock Market	X2 Import	X3 Saving	X4 Commission	X5 Traveler Checks
Y Export	Pearson Correlation Sig. (2-tailed)	1 .	.126 .182	.337** .000	.283** .002	.137 .147	.395** .000
X1 Stock Market	Pearson Correlation Sig. (2-tailed)	.126 .182	1 .	.189* .044	.342** .000	.019 .841	.094 .318
X2 Import	Pearson Correlation Sig. (2-tailed)	.337** .000	.189* .044	1 .	.161 .086	.207* .027	.282** .002
X3 Saving	Pearson Correlation Sig. (2-tailed)	.283** .002	.342** .000	.161 .086	1 .	.109 .249	.119 .206
X4 Commission	Pearson Correlation Sig. (2-tailed)	.137 .147	.019 .841	.207* .027	.109 .249	1 .	.281** .002
X5 Traveler Checks	Pearson Correlation Sig. (2-tailed)	.395** .000	.094 .318	.282** .002	.119 .206	.281** .002	1 .

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

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

a. Listwise N=114

הפלט המתייחס לתוספת ולמובהקות

חלק ה- change statistic בפלט ה-model summary מתייחס לתוספת ולמובהקות.

לדוגמא:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.395 ^a	.156	.149	67.68720	.156	20.760	1	112	.000
2	.461 ^b	.213	.198	65.68529	.056	7.931	1	111	.006
3	.505 ^c	.255	.235	64.19061	.042	6.229	1	110	.014

- a. Predictors: (Constant), X5 Traveler Checks
- b. Predictors: (Constant), X5 Traveler Checks, X3 Saving
- c. Predictors: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

? על פי ה- R square change בלבד חשב את פרופורציית השונות המוסברת של פעילות הבנק ביצוא על ידי המשתנים הב"ת הבאים:

- א. יבוא, המחאות נוסעים ופעילות בחשבונות חיסכון.
- ב. פרופורציית השונות המוסברת על ידי המשתנים המחאות נוסעים ופעילות בחשבונות חיסכון בלבד.
- ג. פרופורציית השונות המוסברת על ידי המשתנים המחאות נוסעים ויבוא בלבד.

הקשר בין מובהקות השיפוע ומובהקות התוספת לניבוי

קיים קשר בין מבחן T למובהקות השיפוע של משתנה ב"ת מסוים כאשר משתנים אחרים מצויים במשוואה (מובהקות התרומה הייחודית שלו) לבין מבחן F למובהקות התוספת של אותו משתנה ב"ת מעבר לאותם משתנים אחרים:

$$t^2\left(\frac{\alpha}{2}\right)(n-k-1) = F(\alpha)(1, n-k-1)$$

sig = sig

נדגים זאת על נתוני דוגמא מס' 2 :

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.395 ^a	.156	.149	67.68720	.156	20.760	1	112	.000
2	.461 ^b	.213	.198	65.68529	.056	7.931	1	111	.006
3	.505 ^c	.255	.235	64.19061	.042	6.229	1	110	.014

- a. Predictors: (Constant), X5 Traveler Checks
- b. Predictors: (Constant), X5 Traveler Checks, X3 Saving
- c. Predictors: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-221.003	72.540		-3.047	.003
	X5 Traveler Checks	130.447	28.630	.395	4.556	.000
2	(Constant)	-254.426	71.388		-3.564	.001
	X5 Traveler Checks	121.042	27.983	.367	4.326	.000
	X3 Saving	.341	.121	.239	2.816	.006
3	(Constant)	-259.691	69.795		-3.721	.000
	X5 Traveler Checks	102.038	28.387	.309	3.595	.000
	X3 Saving	.301	.119	.211	2.521	.013
	X2 Import	.251	.100	.216	2.496	.014

a. Dependent Variable: Y Export

ניתן לראות למשל את הקשר בין מובהקות השיפוע של יבוא כאשר חיסכון והמחאות נוסעים במשוואה (פלט המקדמים) לבין מובהקות התוספת של יבוא מעבר לאותם שני משתנים (פלט ה-model summary):

$$t^2 = 2.496^2 = F = 6.229$$

$$sig = 0.014$$

?

1. על סמך פלט ה- model summary בלבד-מה יהיה סטטיסטי T של מבחן מובהקות התרומה

הייחודית של חשבונות חיסכון כאשר המחאות נוסעים מצוי במשוואה?

2. על סמך פלט המקדמים בלבד- מה יהיה ערך ה- sig של מובהקות התוספת של חשבונות חיסכון

כאשר המחאות נוסעים מצוי במשוואה?

בדיקת קו ליניאריות

קו-ליניאריות משמעה מתאם בין המשתנים הב"ת.

אנו שואפים שהקו-ליניאריות במודל (מתאמים בין המשתנים הב"ת) תהיה נמוכה ככל שניתן כי כידוע מתאם גבוה בין הב"ת פוגם בתרומה הייחודית של המשתנים לניבוי המשתנה התלוי.

לעיתים נתבקש להעריך האם קיימת בעיה של קו-ליניאריות במודל הכולל שני משתנים ב"ת. במקרה זה נצטרך למעשה להעריך האם קיים מתאם גבוה בין שני המשתנים הב"ת שמצויים במודל.

אנו יודעים כי כאשר אין מתאם כלל בין הב"ת- $r_{12} = 0$ מתקיים:

$$(1) \quad \hat{\beta}_1 = \frac{\hat{\beta}_1}{z\hat{\beta}_1} \text{ ברגרסיה פשוטה (כאשר } x_2 \text{ מצוי גם בתוך המשוואה).}$$

$$(2) \quad \begin{aligned} z\hat{\beta}_1 &= r_{y1} \\ z\hat{\beta}_2 &= r_{y2} \end{aligned} \text{ ברגרסיה מרובה (כאשר שני המשתנים הב"ת בתוך המשוואה).}$$

$$(3) \quad \hat{\beta}_1 > \hat{\beta}_1 \text{ במרובה (כאשר } x_2 \text{ מצוי גם הוא במשוואה) בפשוטה}$$

$$(4) \quad R_{y12}^2 = r_{y1}^2 + r_{y2}^2$$

$$(5) \quad \text{Tolerance=1 (בפלט ה- excluded variables).}$$

הסבר: מדד ה-Tolerance נותן את פרופורציית השונות הלא מוסברת של משתנה ב"ת מסוים ע"י משתנים ב"ת אחרים המצויים במשוואה. ככל שפרופורציה זו גבוהה יותר כך קיים מתאם נמוך יותר בין אותו משתנה ב"ת למשתנים ב"ת האחרים המצויים במשוואה.

Tolerance=1 מעיד על חוסר מתאם בין הב"ת ואילו Tolerance=0 מעיד על מתאם מושלם בין המשתנה הב"ת למשתנים הב"ת האחרים המצויים במשוואה.

נדגים באמצעות נתוני דוגמא מס' 2 :

Excluded Variables^d

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	X1 Stock Market	.089 ^a	1.025	.308	.097	.991
	X2 Import	.245 ^a	2.793	.006	.256	.920
	X3 Saving	.239 ^a	2.816	.006	.258	.986
	X4 Commission	.028 ^a	.305	.761	.029	.921
2	X1 Stock Market	.011 ^b	.118	.906	.011	.880
	X2 Import	.216 ^b	2.496	.014	.232	.904
	X4 Commission	.008 ^b	.093	.926	.009	.915
3	X1 Stock Market	-.019 ^c	-.213	.832	-.020	.864
	X4 Commission	-.020 ^c	-.229	.820	-.022	.900

a. Predictors in the Model: (Constant), X5 Traveler Checks

b. Predictors in the Model: (Constant), X5 Traveler Checks, X3 Saving

c. Predictors in the Model: (Constant), X5 Traveler Checks, X3 Saving, X2 Import

d. Dependent Variable: Y Export

? על סמך נתוני דוגמא 2 הערך את גודל המתאם בין המשתנים הב"ת: המחאות נוסעים' (x_5) ו-'פעילות בחשבונות חיסכון' (x_3).

שאלות מסכמות

סקר מטעם הלשכה המרכזית לסטטיסטיקה בחן את הגורמים המשפיעים על רווחי חברות במשק. לשם כך נדגמו 117 חברות במשק ונתקבלו לגביהן הנתונים על המשתנים הבאים:

רווחי החברה

גיל המנהל

רמת השכלתו

הוותק בתפקיד הנוכחי

מס' העובדים בחברה

נתקבלו הממצאים הבאים:

Descriptive statistic

	Mean	Std.deviation	N
Profit	5608.08		117
Time Job seniority	80.02		117
AGE of manager	13.97		117
EDLEVEL	6.888		117
Workers No	3.17		117

Correlations^a

	PROFIT Profit for the firm in \$	TIME Job seniority (months)	AGE Age of manager	EDLEVEL Educational level (years)	EXP Work experience (years)	WORKERS No. of workers
PROFIT Profit for the firm in \$	Pearson Correlation Sig. (2-tailed)	1 .227* .014	-.115 .219	.784** .000	-.093 .318	.575** .000
TIME Job seniority (months)	Pearson Correlation Sig. (2-tailed)	.227* .014	1 .041 .661	.162 .080	.001 .993	.183* .048
AGE Age of manager	Pearson Correlation Sig. (2-tailed)	-.115 .219	.041 .661	1 .270** .003	.798** .000	-.230* .013
EDLEVEL Educational level (years)	Pearson Correlation Sig. (2-tailed)	.784** .000	.162 .080	1 .270** .003	-.171 .065	.824** .000
EXP Work experience (years)	Pearson Correlation Sig. (2-tailed)	-.093 .318	.001 .993	.798** .000	1 .065	-.146 .116
WORKERS No. of workers	Pearson Correlation Sig. (2-tailed)	.575** .000	.183* .048	-.230* .013	.824** .000	1 .116

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

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

a. Listwise N=117

Model summary

Model	R	R square	Adjusted R square	Std. Error of the Estimate
1	a			
2	b			653.959
3	c			645.613

- a. Predictors: (Constant), EDLEVEL.
- b. Predictors: (Constant), EDLEVEL, Workers No.
- c. Predictors: (Constant), EDLEVEL, Workers No., Time Job seniority.

Excluded Variables^d

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	TIME Job seniority (months)	.102 ^a	1.752	.082	.162	.974
	AGE Age of manager	.104 ^a	1.751	.083	.162	.927
	EXP Work experience (years)	.042 ^a	.716	.475	.067	.971
	WORKERS No. of workers	-.220 ^a	-2.188	.031	-.201	.321
2	TIME Job seniority (months)	.114 ^b	1.992	.049	.184	.966
	AGE Age of manager	.103 ^b	1.750	.083	.162	.927
	EXP Work experience (years)	.041 ^b	.707	.481	.066	.971
3	AGE Age of manager	.093 ^c	1.593	.114	.149	.919
	EXP Work experience (years)	.038 ^c	.656	.513	.062	.970

- a. Predictors in the Model: (Constant), EDLEVEL Educational level (years)
- b. Predictors in the Model: (Constant), EDLEVEL Educational level (years), WORKERS No. of workers
- c. Predictors in the Model: (Constant), EDLEVEL Educational level (years), WORKERS No. of workers, TIME Job seniority (months)
- d. Dependent Variable: PROFIT Profit for the firm in \$

א. האם המקדם החלקי של רמת השכלה מובהק כאשר מס' עובדים מצוי במשוואה? בדוק ברמת מובהקות של 0.05 (רשום הנחות, הצב השערות, חשב והסק מסקנה מילולית).

ב. בנה טבלת מקור שונות לניבוי הרווחים על ידי רמת ההשכלה ומס' העובדים בלבד.

ג. מהו סכום הסטיות הריבועיות של הרווחים על פי טבלת מקור השונות שבנית בסעיף הקודם? כיצד ישתנה אם נוסיף את המשתנה הב"ת של וותק המנהל? הסבר.

ד. חשב את גודל התוספת במונחי R^2 של מס' העובדים והוותק על פני רמת ההשכלה. האם תוספת זו מובהקת? בדוק ברמת מובהקות של 0.05 (רשום הנחות, הצב השערות, חשב והסק מסקנה מילולית).

ה. על פי איזה שיטה נותחו הנתונים (stepwise או היררכית)? נמק תשובתך ותן לה סימוכין מתוך הנתונים.

בהמשך הוחלט לנתח את אותם הנתונים בדרך נוספת והתקבלו הממצאים הבאים:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.790 ^a	.624	.618	658.748	.624	94.654	2	114	.000
2	.797 ^b	.636	.623	654.240	.012	1.788	2	112	.172
3	.808 ^c	.653	.637	641.338	.017	5.552	1	111	.020

a. Predictors: (Constant), EDLEVEL Educational level (years), AGE Age of manager

b. Predictors: (Constant), EDLEVEL Educational level (years), AGE Age of manager, TIME Job seniority (months), EXP Work experience (years)

c. Predictors: (Constant), EDLEVEL Educational level (years), AGE Age of manager, TIME Job seniority (months), EXP Work experience (years), WORKERS No. of workers

$$\hat{y}_i = 400.612 - 369.364Branch1 - 245.308Branch2 + 0.806x_i$$

ו. האם המשוואה לניבוי הרווחים על ידי רמת ההשכלה וגיל המנהל מובהקת באוכלוסייה? בדוק ברמת מובהקות של 0.05 (רשום הנחות, הצב השערות והסק מסקנה מילולית בלבד, ללא חישוב).

ז. מה יהיה סטטיסטי T למובהקות מקדם השיפוע של המשתנה – מספר עובדים' כאשר יתר המשתנים הב"ת האחרים מצויים בתוך המשוואה?

ח. איזה משתנים ב"ת יופיעו בפלט ה-excluded variables בכל שלב?