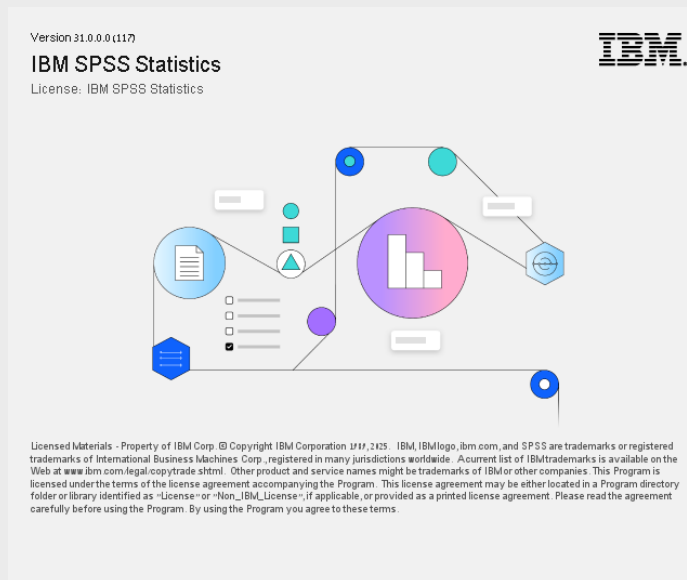


# IBM SPSS Statistics 31.0.1



מני ברגר  
מדען נתונים, ג'ניוס מערכות  
ספטמבר 2025

[Meni@genius.co.il](mailto:Meni@genius.co.il) ■

קבוצת הפייסבוק - SPSS Israel ■

[www.dataninapro.com](http://www.dataninapro.com) ■

<https://github.com/Dataninapro> ■



- קצת על התכונות החדשות של גרסה 31.0.2
- הצצה על העתיד: מנוע ה-AI של SPSS
- סקירה על החידושים בגרסאות קודמות

## תחילתה של מהפכה!

בחר את ה-LLM המועדף עליך ותן לבינה המלאכותית לספק פרשנות, הסברים ותובנות מתוך פלט ה-SPSS שלך.

previous inquiries:

Current AI model: **ibm/granite-3-sb-instruct** Change model...

watsonx.ai account: **Using own account** Edit account.

Objects to inquire about: ☐ All ☐ All visible ☒ Selected

Ask the selected AI model about your output:

Professional Statistical Interpretation

Professional Statistical Interpretation

Executive Summary for Business Stakeholders

Instructional Summary for Students

Diagnostic Review for Data Quality

Enter Prompt Manually

View or Edit AI Prompt

AI prompt text:

I am a statistician analyzing output from IBM SPSS Statistics. Below is the text extracted from an SPSS procedure output (e.g., regression, correlation, or descriptive statistics).

Please help interpret this output and extract meaningful insights such as:

- The interpretation of coefficients, p-values, R-squared, correlations, or test statistics.
- Any trends or notable patterns you notice in the data.
- Issues that might need attention (e.g., multicollinearity, missing values, insignificant predictors).
- Suggestions for additional analysis or improvement.

Treat me as a fellow professional — I'm looking for technically sound insights with explanations.

OK

Cancel

Help

AI Output Assistant

The AI Output Assistant uses IBM AI technology to offer insights into the meaning and significance of your output. From this location you can select the output content and model to use for evaluation. You can also set a custom prompt that will build upon the results of your previous inquiries.

Current AI model: **ibm/granite-3-sb-instruct** Change model...

watsonx.ai account: **Using own account** Edit account.

Objects to inquire about: ☐ All ☐ All visible ☒ Selected

Ask the selected AI model about your output:

Professional Statistical Interpretation Edit/View Prompt.

Execute your inquiry against watsonx.ai using the selected model and your chosen prompt. Preview AI Inquiry

Preview:

4. Variable 312.50010.15012.667(0.009)0.05010.150

Interpretation:

1. **R-squared (0.454) and Adjusted R-squared (0.421):** These values indicate that approximately 45.4% and 42.1% of the variance in the dependent variable can be explained by the independent variables, respectively. The slight difference between R-squared and Adjusted R-squared suggests that the model might be slightly overfitting the data.
2. **Regression Sum of Squares (0.000):** This value represents the total variance explained by the regression model.
3. **Residual Sum of Squares (0.500):** This value represents the total unexplained variance.
4. **F-statistic (50.000) and its significance (0.000):** The F-statistic is significant, indicating that at least one of the predictors in the model is significantly related to the dependent variable.
5. **Coefficients:**
  - **Variable 1:** The unstandardized coefficient (B = 3.500) indicates that for every one-unit increase in Variable 1, the dependent variable is predicted to increase by 3.500 units, holding other variables constant. The standardized coefficient (Beta = 0.350) suggests that Variable 1 has a moderate effect on the dependent variable. The t-value (3.333) and p-value (0.000) indicate that this relationship is statistically significant.
  - **Variable 2:** Similarly, for every one-unit increase in Variable 2, the dependent variable is predicted to increase by 2.000 units. The standardized coefficient (Beta = 0.200) suggests a weak to moderate effect. The t-value (3.571) and p-value (0.001) indicate that this relationship is statistically significant.
  - **Variable 3:** For every one-unit increase in Variable 3, the dependent variable is predicted to increase by 1.500 units. The standardized coefficient (Beta = 0.150) suggests a weak effect. The t-value (2.667) and p-value (0.009) indicate that this relationship is statistically significant.
6. **No multicollinearity issues are apparent from this output, as the Variance Inflation Factor (VIF) for each predictor is not provided. However, if VIF values were high (>5), it would indicate multicollinearity.**
7. **Missing values.** The output does not provide information about missing values. If there were missing values, it would be important to investigate their impact on the analysis.
8. **Additional analysis.** To further understand the relationships between the variables, consider creating scatter plots or partial regression plots. Additionally, you might want to explore interaction terms between the predictors if

\*\*\*\*\*

This content has been generated by an AI system and is provided for informational purposes only. It is not an official IBM SPSS output. The information presented here is a summarization or interpretation by AI and may not be fully accurate, complete, or deterministic. Please validate critical details independently before use.

1 Add to Output Cancel Help



**הגיע הזמן להפסיק להחמיץ קשרים**  
מדדי קשר מסורתיים מצליחים להצביע על קשרים לינאריים  
DC מזהה גם קשרים לא לינאריים בין משתנים.

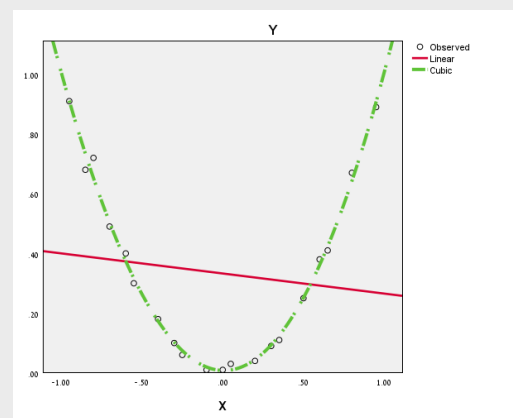
**Distance Correlation Coefficients<sup>a</sup>**

| Variable | dCor Value | Confidence Interval (Permutation, Sig. Level = 0.05) |       | 95.00% Confidence Interval |       | Sig.  |
|----------|------------|------------------------------------------------------|-------|----------------------------|-------|-------|
|          |            | Lower                                                | Upper | Lower                      | Upper |       |
| X vs. Y  | .2660      | .0507                                                | .3010 | .0000                      | .6339 | .0400 |

a. Number of Replicates (R) = 100

**Correlations**

|   |                     | X     | Y     |
|---|---------------------|-------|-------|
| X | Pearson Correlation | 1     | -.131 |
|   | Sig. (2-tailed)     |       | .582  |
|   | N                   | 20    | 20    |
| Y | Pearson Correlation | -.131 | 1     |
|   | Sig. (2-tailed)     | .582  |       |
|   | N                   | 20    | 20    |



**איתור קשר קוואדראטי  
בהעדר קשר לינארי  
(נתוני סימולציה)**

# Conditional Inference Trees

## CHAID

### 3 Levels

| Observed           | Predicted |       | Percent Correct |
|--------------------|-----------|-------|-----------------|
|                    | 0 NO      | 1 YES |                 |
| 0 NO               | 20809     | 3680  | 85.0%           |
| 1 YES              | 4893      | 13770 | 73.8%           |
| Overall Percentage | 59.6%     | 40.4% | 80.1%           |

Growing Method: CHAID

Dependent Variable: Accept\_Cashcard Accept the new cash card?

## CIT

### Multi Level

Confusion Matrix - Accept\_Cashcard

| Observed | Predicted |       | Total |
|----------|-----------|-------|-------|
|          | NO        | YES   |       |
| NO       | 21245     | 3244  | 24489 |
| YES      | 4751      | 13912 | 18663 |
| Total    | 25996     | 17156 | 43152 |

Correct: count: 35157, Percentage: 81.47

### 3 Levels

Confusion Matrix - Accept\_Cashcard

| Observed | Predicted |       | Total |
|----------|-----------|-------|-------|
|          | NO        | YES   |       |
| NO       | 18541     | 5948  | 24489 |
| YES      | 3427      | 15236 | 18663 |
| Total    | 21968     | 21184 | 43152 |

Correct: count: 33777, Percentage: 78.27

### 5 Levels

Confusion Matrix - Accept\_Cashcard

| Observed | Predicted |       | Total |
|----------|-----------|-------|-------|
|          | NO        | YES   |       |
| NO       | 20926     | 3563  | 24489 |
| YES      | 4458      | 14205 | 18663 |
| Total    | 25384     | 17768 | 43152 |

Correct: count: 35131, Percentage: 81.41

## CIT

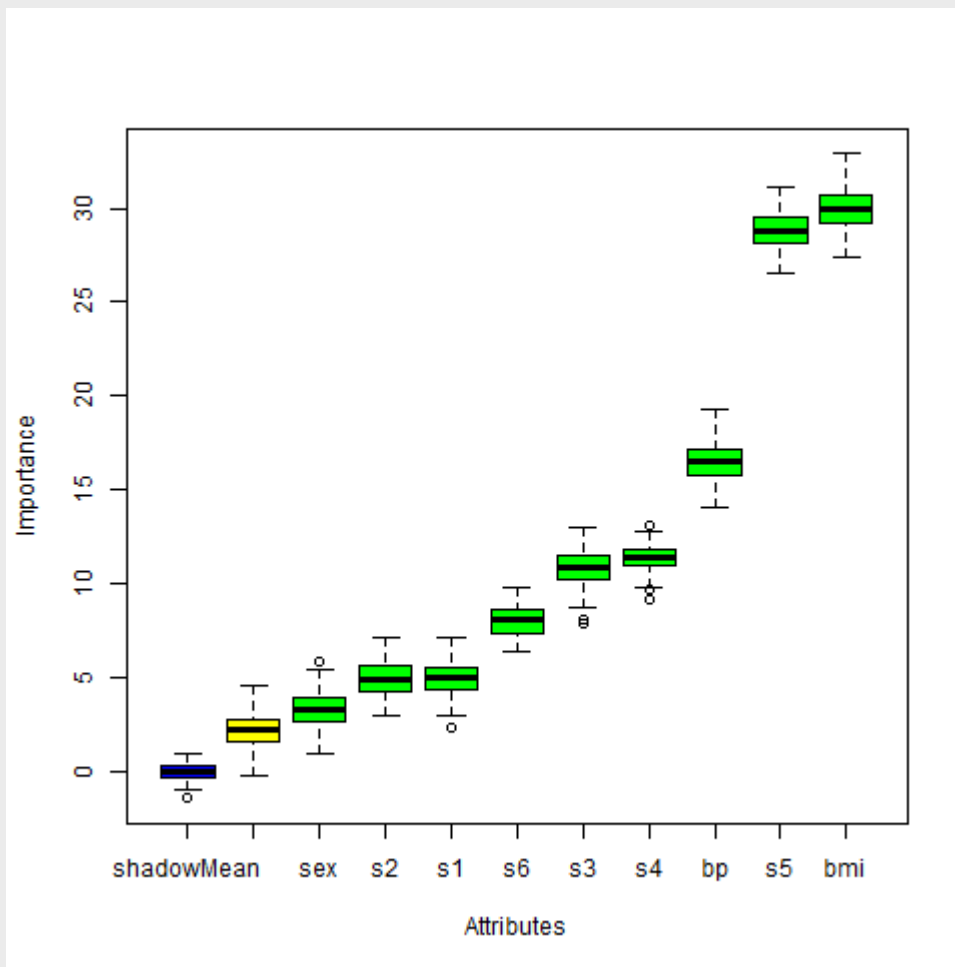
משפחה חדשה של עצי החלטה א-פרמטריים מבוססי p-value כקריטריון לפיצול בנתונים. מסייעים במניעת הטיות בסיווג המבוסס על קריטריון מסורתיים.



## Features

|     | Decision  |
|-----|-----------|
| s6  | Confirmed |
| s5  | Confirmed |
| s4  | Confirmed |
| s3  | Confirmed |
| s2  | Confirmed |
| s1  | Confirmed |
| bp  | Confirmed |
| bmi | Confirmed |
| sex | Confirmed |
| age | Tentative |

Dependent variable:  
target  
Pvalue: 0.01, Bonferroni:  
Yes, Boruta Version:  
9.0.0



## Boruta

**אלגוריתם לבחירת משתנים**  
בנוי מעל-Random Forest.  
המטרה שלו היא לזהות אילו  
משתנים באמת רלוונטיים  
לניבוי.



Correlations

|          |                     | price   | cost    | calories | sodium | alcohol |
|----------|---------------------|---------|---------|----------|--------|---------|
| price    | Pearson Correlation | 1       | .1000** | .217     | -.090  | .196    |
|          | Sig. (2-tailed)     |         | <.001   | .210     | .608   | .259    |
|          | N                   | 35      | 35      | 35       | 35     | 35      |
| cost     | Pearson Correlation | .1000** | 1       | .214     | -.090  | .196    |
|          | Sig. (2-tailed)     | <.001   |         | .216     | .606   | .260    |
|          | N                   | 35      | 35      | 35       | 35     | 35      |
| calories | Pearson Correlation | .217    | .214    | 1        | .285   | .914**  |
|          | Sig. (2-tailed)     | .210    | .216    |          | .097   | <.001   |
|          | N                   | 35      | 35      | 35       | 35     | 35      |
| sodium   | Pearson Correlation | -.090   | -.090   | .285     | 1      | .205    |
|          | Sig. (2-tailed)     | .608    | .606    | .097     |        | .237    |
|          | N                   | 35      | 35      | 35       | 35     | 35      |
| alcohol  | Pearson Correlation | .196    | .196    | .914**   | .205   | 1       |
|          | Sig. (2-tailed)     | .259    | .260    | <.001    | .237   |         |
|          | N                   |         |         |          |        |         |

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

Pearson Correlations

- Highly Positive:** (price <---> cost), (calories <---> alcohol)
- Positive:** (price <---> calories), (price <---> alcohol), (cost <---> calories), (cost <---> alcohol), (calories <---> sodium), (sodium <---> alcohol)
- No Linear Correlation:** (None)
- Negative:** (price <---> sodium), (cost <---> sodium)
- Highly Negative:** (None)

**SPSS מסייע בהפקת התובנות**  
פלט המתאמים מסומן בקוד צבע בהתאם לעוצמת המתאם. פרשנות נוספת מסופקת לפי ההקשר כטקסט בהמשך המטריצה.

**זמין כעת רק ב-  
Correlations וגם  
בניתוחי T**

Paired Samples Test

|        |                    | Paired Differences |                |                 |                                           |          | t       | df  | Significance |             |
|--------|--------------------|--------------------|----------------|-----------------|-------------------------------------------|----------|---------|-----|--------------|-------------|
|        |                    | Mean               | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference |          |         |     | One-Sided p  | Two-Sided p |
|        |                    |                    |                |                 | Lower                                     | Upper    |         |     |              |             |
| Pair 1 | creddebt - othdebt | -1.50198           | 2.59829        | .08912          | -1.67691                                  | -1.32706 | -16.853 | 849 | < .001       | < .001      |

☒ Estimate effect sizes

Paired Samples Test: One-Sided p

- The average difference is not significant for the paired variable(s): (None)
- The average difference is significant for the paired variable(s): (creddebt - othdebt)

Paired Samples Test: Two-Sided p

- The average difference is not significant for the paired variable(s): (None)
- The average difference is significant for the paired variable(s): (creddebt - othdebt)

Note: Curated Help is calculated based on actual cell values, not the formatted values.

Time Series Filters

Filters

☒ Hodrick-Prescott (HP)

☐ Baxter-King (BK)

☐ Christiano-Fitzgerald (CF)

Variables:

Sort: None

- custid
- mortgage
- life\_ins
- cre\_card
- deb\_card
- mob\_bank
- curr\_acc
- internet
- perloan
- savings
- atm\_user
- odrtlim
- isa
- age

HP Variable:

Lambda: 1600

BK and CF Variables:

Log: 6

High: 32

K: 12

☐ Drift

OK Paste Reset Cancel Help

## Time Series Filter

פתרונות סינון לתנודתיות ו"רעשי רקע" בנתוני סדרות עתיות (Time Series). נתונים נקיים ניבוי מדויק יותר.

# Multivariate Adaptive Regression Splines (MARS)

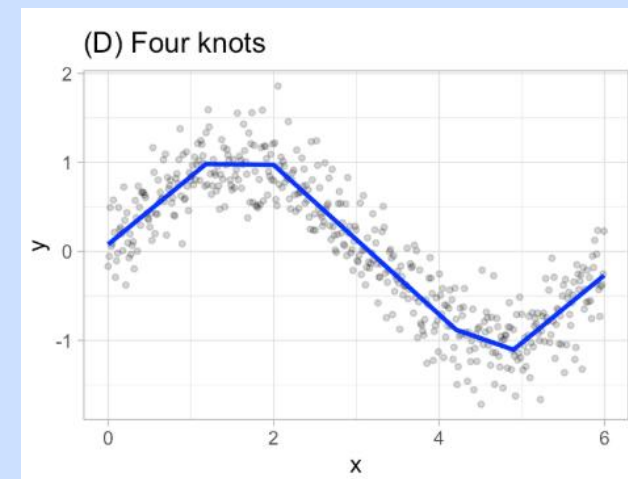
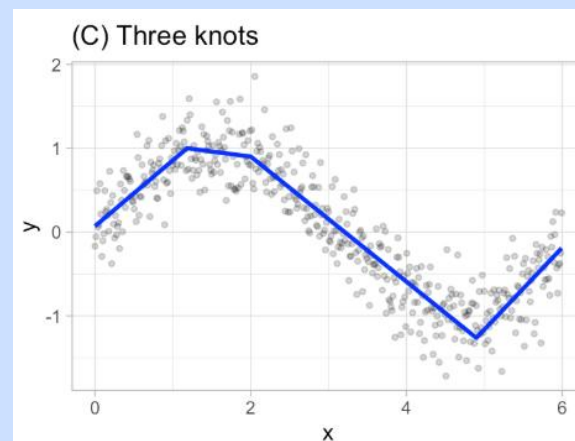
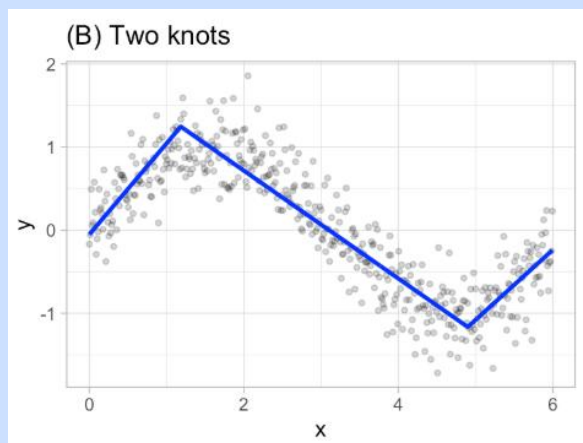
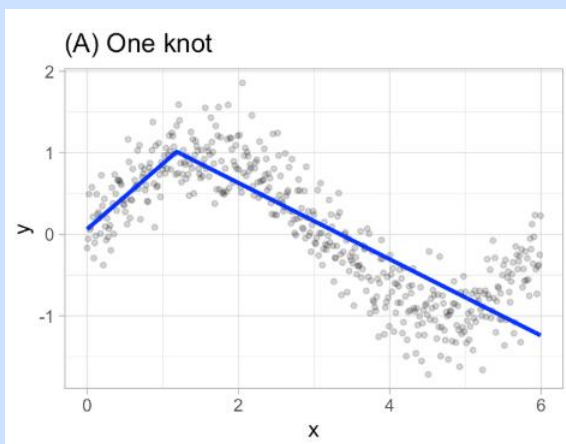
- מודל חכם המבצע חיפוש לאורך הסקאלה של המשתנה כדי למצוא קשרים מועילים.
- במקום להניח קשר ליניארי אחד לכל משתנה, MARS בונה את הפתרון מחלקים קטנים.
- חלקים אלו הם **Hinge Functions** המורכבים מ"שברים" של קווים המחוברים באמצעות **Knots**.

# Multivariate Adaptive Regression Splines (MARS)

שלב 1 **Forward** - הוספת נקודות שבירה ואיברים כדי ללכוד דפוסים לא ליניאריים ואינטראקציות.

שלב 2 **Backward pruning** - "גיזום" האיברים בעלי התרומה הנמוכה ביותר למודל.

תוצר - מודל קומפקטי עם הכללה טובה יותר.



# Multivariate Adaptive Regression Splines (MARS)

## טיפול אוטומטי באי-לינאריות:

אין צורך להגדיר טרנספורמציות, פולינומים, קטגוריזציה ידנית ו- step-functions.

## איתור אינטראקציות אדפטיבי:

למשל כאשר משתנה אחד משפיע רק אחרי סף מסוים ובתנאי שמשתנה אחר גבוה/נמוך.

## סיווג משופר ב"אזורים בעייתיים":

מתגבר על בעיות הערכת יתר/חסר של מודל לוגיסטי הסובל מתת-טווח.

## Interpretability טובה:

מציג בבהירות יחסית אילו משתנים נכללים במודל, נקודות השבירה ואיך כל משתנה משפיע בחלקים השונים של הטווח.

השוואת סיווג בינארי לניבוי תחלואה לבבית-  
מודל רגרסיה לוגיסטית בינארית אל מול מודל MARS

קובץ הנתונים מתוך Kaggle - 70,000 תצפיות על פני 12 משתנים מנבאים.  
המנבאים הם מדדים אנתרופומטריים וקליניים של הנבדקים.  
חלקם רציפים וחלקם קטגוריאליים.  
המשתנה המנובא הוא קיומה של מחלת לב (מקודד 1).

## איכות הסיווג- מודל לוגיסטי בינארי

|                      |             | ניבוי Predicted |            |
|----------------------|-------------|-----------------|------------|
|                      |             | control (0)     | target (1) |
| מצב במציאות observed | control (0) | 26844           | 8177       |
|                      | target (1)  | 11326           | 23653      |

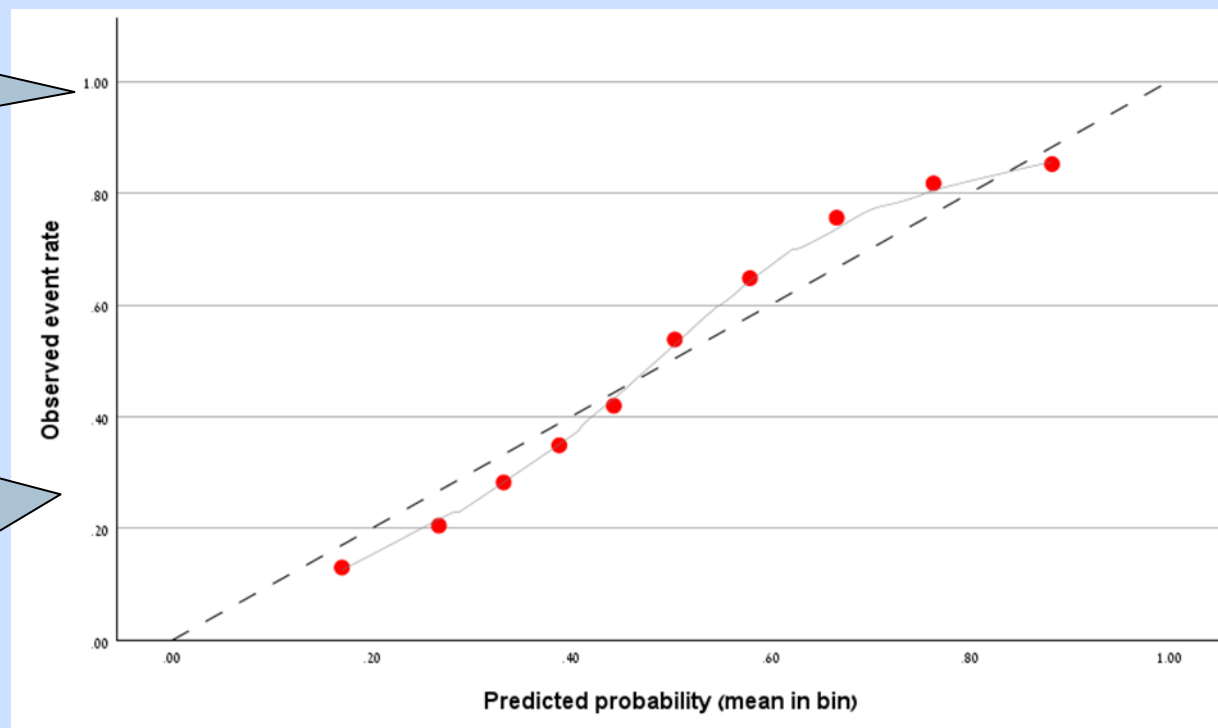
|                                                    |        |
|----------------------------------------------------|--------|
| <b>Sensitivity (recall)- true positive</b>         | 67.6%  |
| <b>Specificity- True negative</b>                  | 76.7%  |
| <b>Precision (PPV) – Positive Predictive Value</b> | 74.3%  |
| <b>Negative Predictive Value (NPV)</b>             | 70.3%  |
| <b>False Positive Rate (FPR)</b>                   | 23.3%  |
| <b>False Negative Rate (FNR)</b>                   | 32.4%  |
| <b>AUC</b>                                         | 0.785* |
| <b>Cohen's kappa</b>                               | 0.42   |

## תרשים קליברציה- מודל לוגיסטי בינארי

נקודות מעל האלכסון- המודל ממעיט בסיכון (בפועל קורה יותר ממה שנחזה).  
נקודות מתחת לאלכסון- המודל מגזים בסיכון (נחזה יותר ממה שקורה בפועל).

שיעור האנשים שחוו  
את האירוע בפועל,  
בכל עשירון.

בטווחי הניבוי הנמוכים-  
בינוניים, המודל מנבא  
יותר מדי סיכון.  
המגמה מתהפכת  
והמודל ממעיט בסיכון  
בטווחים הגבוהים.



ממוצע ההסתברות  
החזויה בכל עשירון.

## איכות הסיווג- מודל MARS (EARTH)

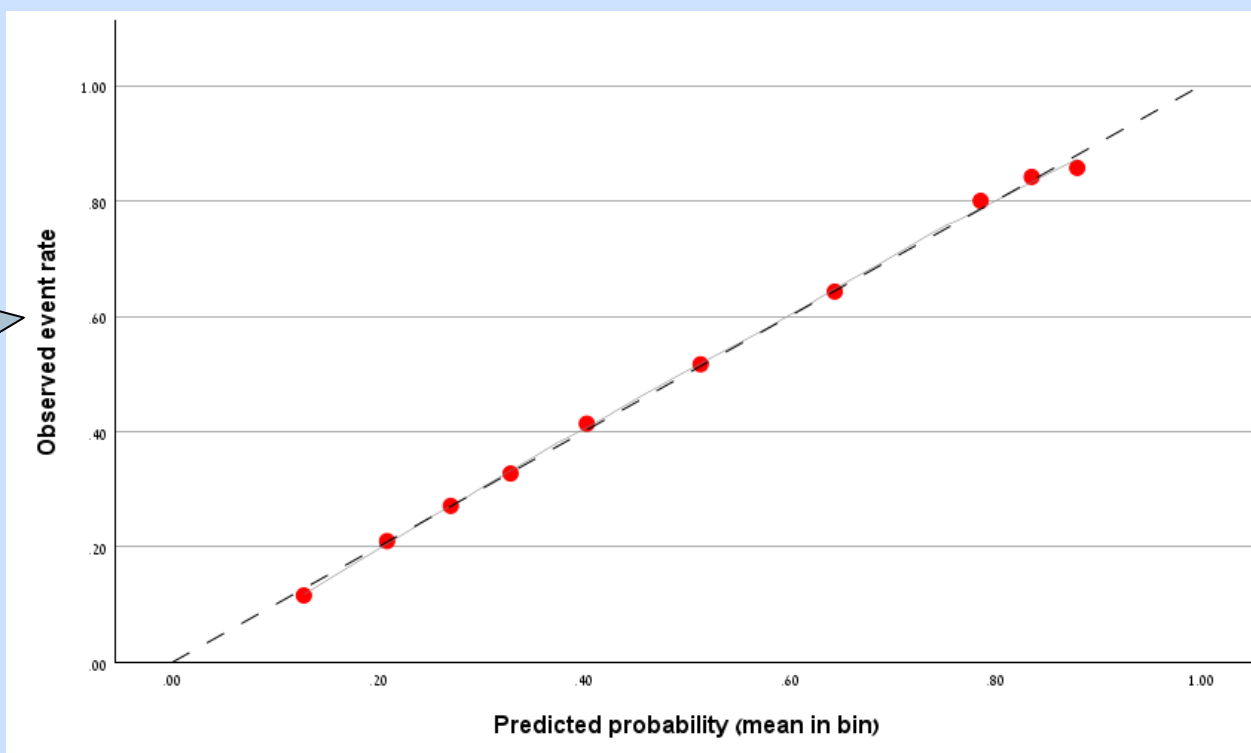
|                                                    |        |
|----------------------------------------------------|--------|
| <b>Sensitivity (recall)- true positive</b>         | 67.6%  |
| <b>Specificity- True negative</b>                  | 76.7%  |
| <b>Precision (PPV) – Positive Predictive Value</b> | 74.3%  |
| <b>Negative Predictive Value (NPV)</b>             | 70.3%  |
| <b>False Positive Rate (FPR)</b>                   | 23.3%  |
| <b>False Negative Rate (FNR)</b>                   | 32.4%  |
| <b>AUC</b>                                         | 0.785* |
| <b>Cohen's kappa</b>                               | 0.42   |

|                                                    |       |
|----------------------------------------------------|-------|
| <b>Sensitivity (recall)- true positive</b>         | 69.5% |
| <b>Specificity- True negative</b>                  | 76.9% |
| <b>Precision (PPV) – Positive Predictive Value</b> | 75.1% |
| <b>Negative Predictive Value (NPV)</b>             | 71.6% |
| <b>False Positive Rate (FPR)</b>                   | 23.1% |
| <b>False Negative Rate (FNR)</b>                   | 30.5% |
| <b>AUC</b>                                         | .798* |
| <b>Cohen's kappa</b>                               | 0.44  |

|                            |             | ניבוי Predicted |            |
|----------------------------|-------------|-----------------|------------|
|                            |             | control (0)     | target (1) |
| מצב<br>observed<br>במציאות | control (0) | 26945           | 8076       |
|                            | target (1)  | 10677           | 24302      |

## תרשים קליברציה- מודל MARS (EARTH)

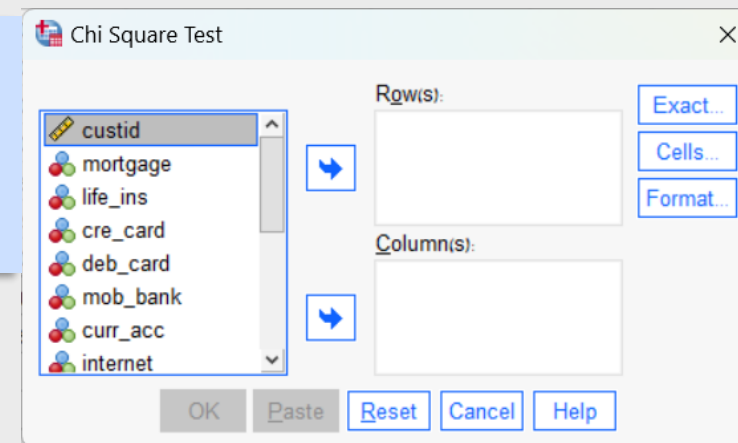
מודל MARS מציג  
שיפור ברור  
בקליברציה לעומת  
המודל הלוגיסטי.



ממצאי מודל MARS  
מציגים הסתברויות  
אמינות יותר!

מבחן חי-בריבוע

לא תלוי ב- Crosstabs

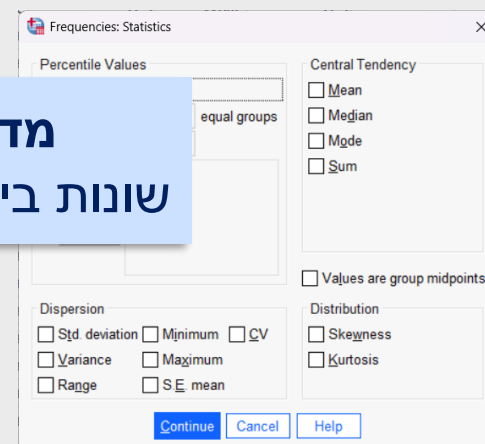


עדכון קבצים

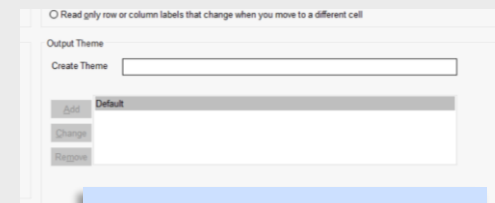
יותר קל למזג קבצים!

מדד CV

שונות ביחס לממוצע

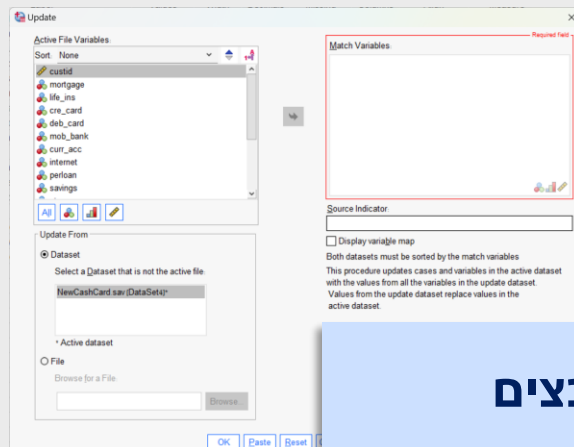


עיצוב הפלט



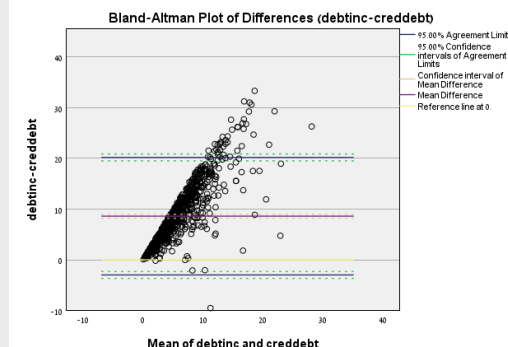
מבחן Bland-Altman

בדיקת הסכמה בין 2 שיטות  
מידעה. שלילת הטיה  
באמצעות בדיקת גבולות  
ההסכמה.



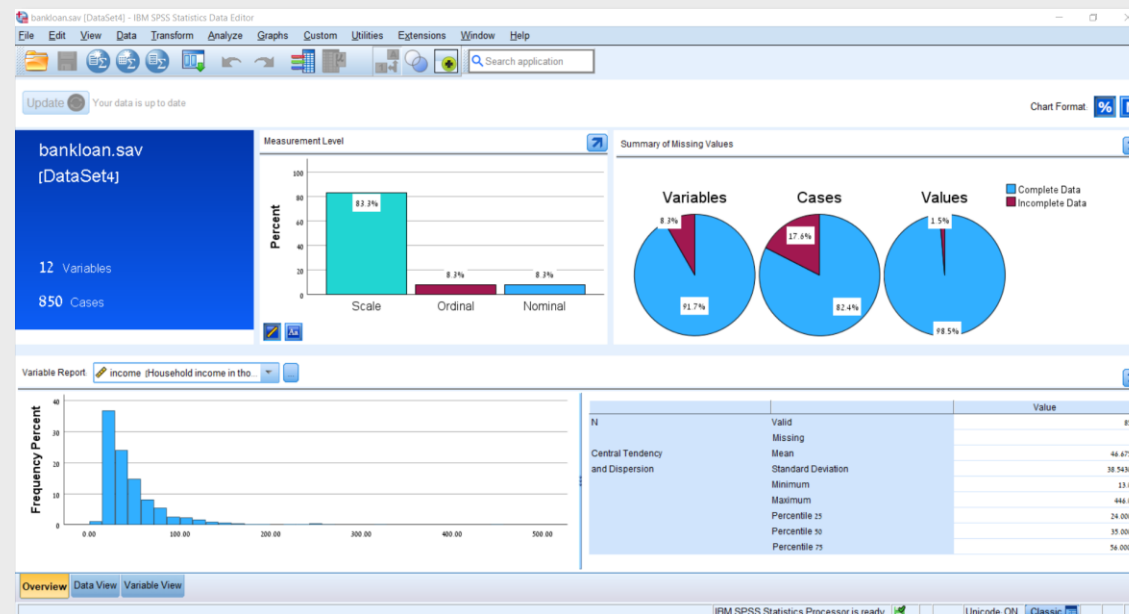
Confidence Intervals of Agreement Limits (debtinc and creddeb)

|                         | 95.00% Agreement Limit Lower | 95.00% Confidence Interval Lower | 95.00% Confidence Interval Upper | 95.00% Agreement Limit Upper | 95.00% Confidence Interval Lower | 95.00% Confidence Interval Upper |
|-------------------------|------------------------------|----------------------------------|----------------------------------|------------------------------|----------------------------------|----------------------------------|
| debtinc - creddeb       | -2.994                       | -3.674                           | -2.313                           | 20.183                       | 19.503                           | 20.864                           |
| (debtinc-creddeb) /Mean | -.510                        | -.625                            | -.394                            | 3.436                        | 3.320                            | 3.552                            |

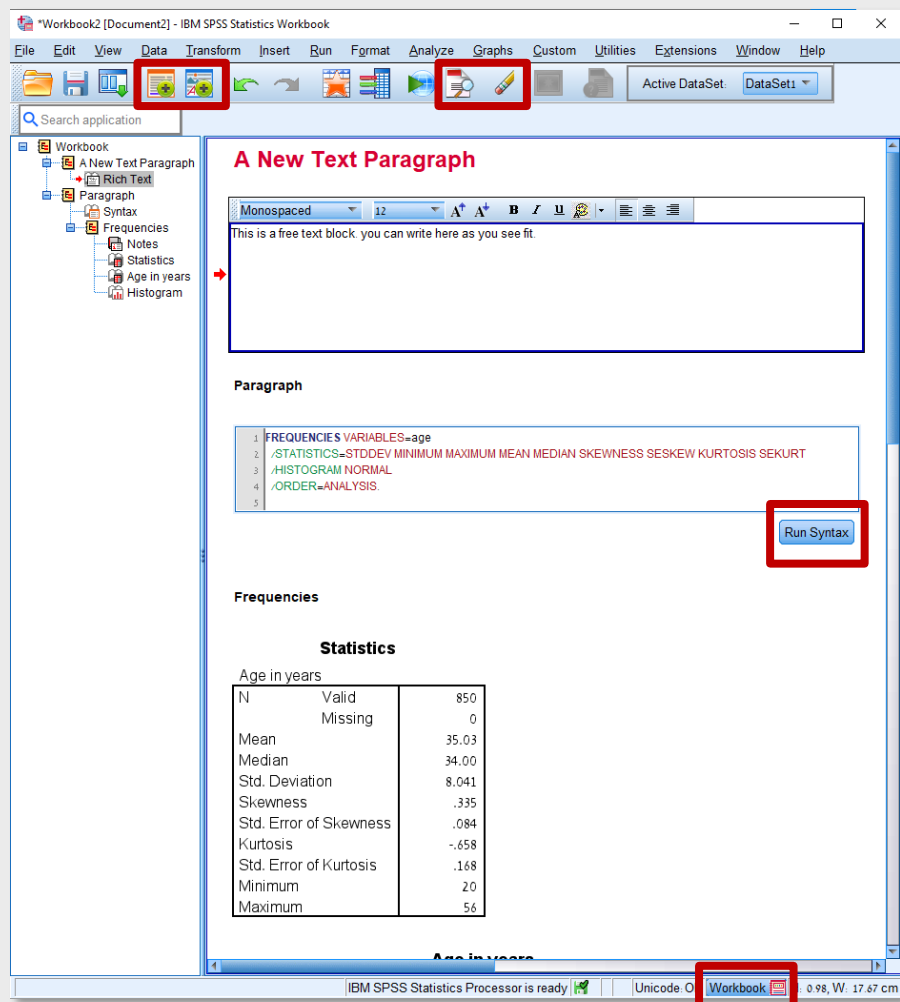


## לשונית ה- Overview

- נכסי מידע מטוייבים ושלמים הכרחיים  
לניתוח ודיווח משמעותי
- ה Overview-היא "תחנה" המאפשרת  
הבנה יסודית של נכסי המידע צעד לפני  
התחלת הניתוח
- באמצעות פאנל ייחודי ניתן לחקור את  
הנתונים בשלמותם ולאחר מכן להתמקד  
במשתנים בודדים



# Application Mode - Workbook



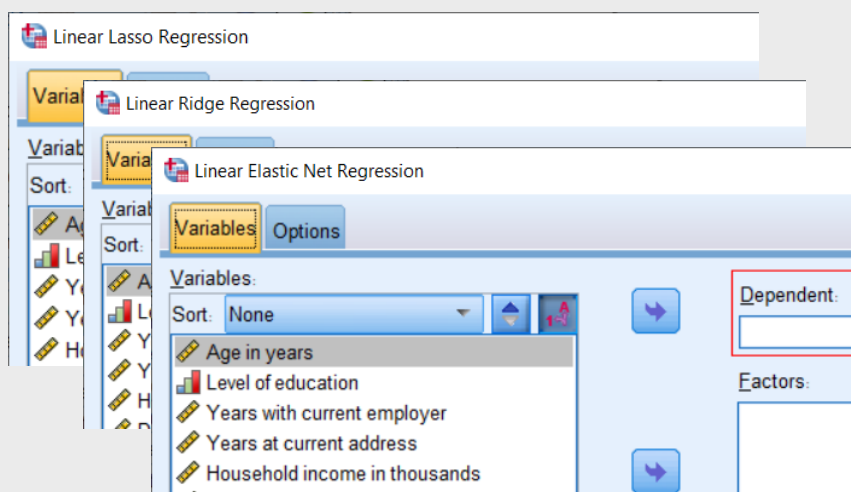
- גישת עבודה חדשה המציגה סינתזה דינמית בין מסמך ה-Syntax ומסמך ה-Output

- בשיטת עבודה זו ניתן לעדכן את ה-Syntax בכל עת והפלט המקושר אליו יתעדכן בהתאם

- ניתן להוסיף פסקאות טקסט כדי ללוות את התהליך האנליטי בהסברים

## אלטרנטיבות למודל רגרסיה OLS – רגרסיה באמצעות רגולריזציה

- Ridge Regression (L2 Regularization) – ניבוי באמצעות הפחתת הרגישות של המנבאים, על-ידי צמצום השאריות של המנבאים. יעיל במיוחד כשיש כמות גדולה של מנבאים התורמים למודל.
- Lasso Regression (L1 Regularization) – בעל דמיון ל-Ridge עם אפשרות לאפס לחלוטין שונות מנבאים בלתי רלוונטיים. יעיל במיוחד בצמצום כמות גדולה של מנבאים שאינם תורמים למודל הניבוי.
- Elastic Net Regression – מודל השופט באופן אוטומטי ומאפשר שילוב היתרונות של OLS, ורגולריזציה בשיטות Ridge ו-Lasso. מנצל את יתרונות 2 השיטות לעיל וגם יעיל במיוחד במצב של מתאם גבוה בין המנבאים.



\*כל הפרוצדורות מכילות אפשרות מובנית ל- Training\Testing

### שימוש עיקרי-

1. הפחתת כמות גבוהה של מנבאים.
2. צמצום קורלציות חזקות בין מנבאים.
3. התאמת מודל חיזוי יציב עם הכללה טובה.
4. "דיוק" המנבאים ובחירת סט קטן יחסית של מנבאים יעילים

**בחינת יכול מודל Elastic Net בצמצום מנבאים והפחתת מולטיקוליניאריות במודל רגרסיה לניבוי חוב כספי.**

**קובץ הנתונים מתוך ספריית קובצי הדוגמה של SPSS Statistics לצורך הדוגמה נבחרו 114 מנבאים ללא בדיקה מוקדמת של הקשר הליניארי שלהם למנובא והתפלגות השאריות. חלק מהמנבאים הינם קטגוריאליים. לא בוצע קידוד דמי.**

Linear Elastic Net Regression

Variables Options

Mode:

☐ Specify individual L1 ratios  
Value(s):

☒ Specify grid of L1 ratio values  
Start:   
End:   
By:

☒ Specify individual alphas  
Value(s):

☒ Specify grid of alpha values  
Start:   
End:   
By:

Alpha metric  
☐ Linear  
☒ Base 10 logarithmic

Criteria  
☒ Include intercept  
☒ Standardize predictors  
Number of crossvalidation folds:   
Python random state:   
Time limit (minutes):

Display  
☐ Best  
☒ Compare models  
☐ Compare models and folds

Plot  
☒ Average crossvalidation mean squared error (MSE) vs. alpha  
☐ Average crossvalidation R Square vs. alpha  
☒ Observed vs. Predicted  
☒ Residuals vs. Predicted

Save  
☐ Predicted values  
Custom variable name:   
☐ Residuals  
Custom variable name:

OK Paste Reset Cancel Help

מצב המודל-  
בחירת השילוב שנותן  
את ביצועי החיזוי הטובים  
ביותר על Test.

גריד של יחסים בין L1  
ל- L2.  
איזון בין בחירת משתנים  
ליציבות.

גריד של עוצמות  
רגולריזציה שונות.  
איזון בין פשטות  
מול דיוק.

מספר חלקי הדאטה  
ל- Train/Test  
המשמש למציאת  
"ממוצע" תוצאות יציב.

# תצוגת תכלית Elastic Net

Best Model Summary<sup>a,b</sup>

| L1 Ratio | Alpha | Number of Crossvalidation Folds | Training R Square | Average Test Subset R Square | Holdout R Square |
|----------|-------|---------------------------------|-------------------|------------------------------|------------------|
| .260     | 6.310 | 5.000                           | .415              | -.260                        | .201             |

a. Dependent Variable: creddebt

b. Model: region, townsize, gender, agecat, edcat, jobcat, union, empcat, retire, inccat, default, jobsat, marital, spousedcat, homeown, hometype, addresscat, cars, carown, cartype, carcatvalue, carbought, carbuy, commute, commutecat, commutecar, commutematerials, commutecarpool, commutebus

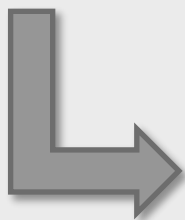
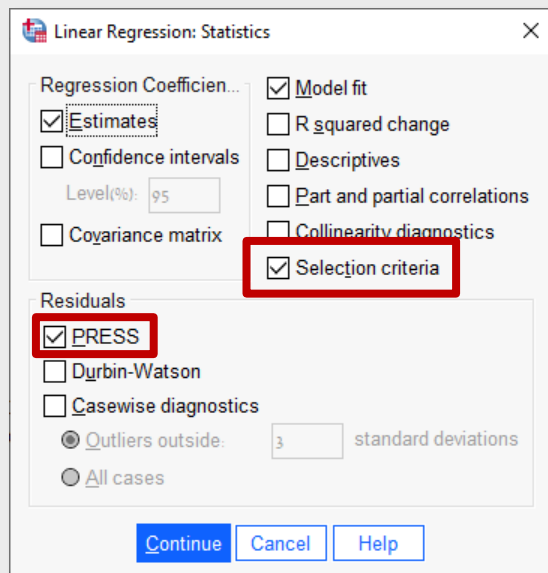
שוונות מנובאת ( $R^2$ )  
המודל הטוב ביותר

טבלת מקדמים  
משמעותיים וערכיהם.  
כל יתר המקדמים  
"מאופסים".

|                        | Mean    | Std. Dev. | Standardized Coefficients | Unstandardized Coefficients |
|------------------------|---------|-----------|---------------------------|-----------------------------|
| Intercept <sup>b</sup> |         |           | 2.825                     | 0.601                       |
| othdebt                | 4.921   | 9.715     | 0.977                     | 0.101                       |
| income                 | 68.200  | 100.814   | 0.870                     | 0.009                       |
| cardspent              | 402.791 | 313.306   | 0.657                     | 0.002                       |
| card2spent             | 193.755 | 183.850   | 0.163                     | 0.001                       |
| debtinc                | 9.795   | 6.501     | 0.137                     | 0.021                       |
| [default=1]            | 0.181   | 0.385     | 0.082                     | 0.213                       |
| [inccat=5]             | 0.103   | 0.304     | 0.068                     | 0.223                       |
| carvalue               | 29.023  | 21.607    | 0.019                     | 0.001                       |
| [cars=4]               | 0.103   | 0.304     | 0.017                     | 0.057                       |

## מדד \*PRESS - מאפשר Cross-validation בין מודלים

### מדדי Selection Criteria לטיב המודל המבוססים על שיטות שונות



| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate | Selection Criteria           |                              |                               |                            | PRESS    |
|-------|-------------------|----------|-------------------|----------------------------|------------------------------|------------------------------|-------------------------------|----------------------------|----------|
|       |                   |          |                   |                            | Akaike Information Criterion | Amemiya Prediction Criterion | Mallows' Prediction Criterion | Schwarz Bayesian Criterion |          |
| 1     | .645 <sup>a</sup> | .416     | .415              | 1.62557                    | 827.960                      | .587                         | 366.617                       | 837.451                    |          |
| 2     | .676 <sup>b</sup> | .457     | .456              | 1.56773                    | 767.369                      | .546                         | 282.533                       | 781.605                    |          |
| 3     | .768 <sup>c</sup> | .590     | .589              | 1.36289                    | 530.320                      | .413                         | 8.365                         | 549.301                    |          |
| 4     | .768 <sup>d</sup> | .590     | .589              | 1.36228                    | 528.564                      | .413                         | 6.609                         | 542.800                    |          |
| 5     | .770 <sup>e</sup> | .593     | .592              | 1.35800                    | 524.218                      | .410                         | 2.282                         | 543.198                    | 1606.669 |

a. Predictors: (Constant), Other debt in thousands

b. Predictors: (Constant), Other debt in thousands, Household income in thousands

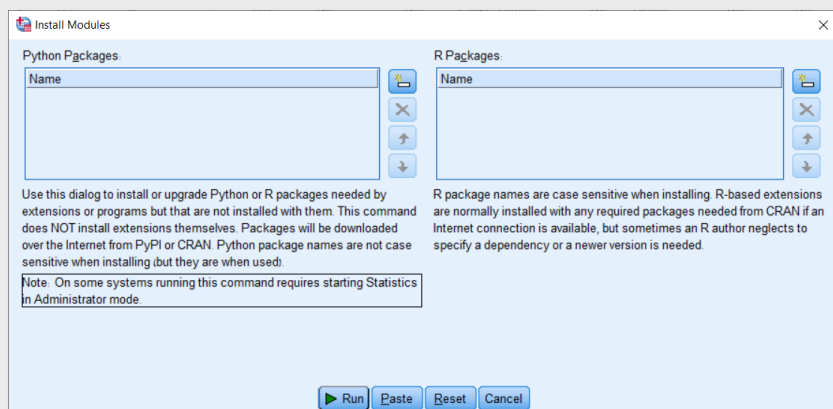
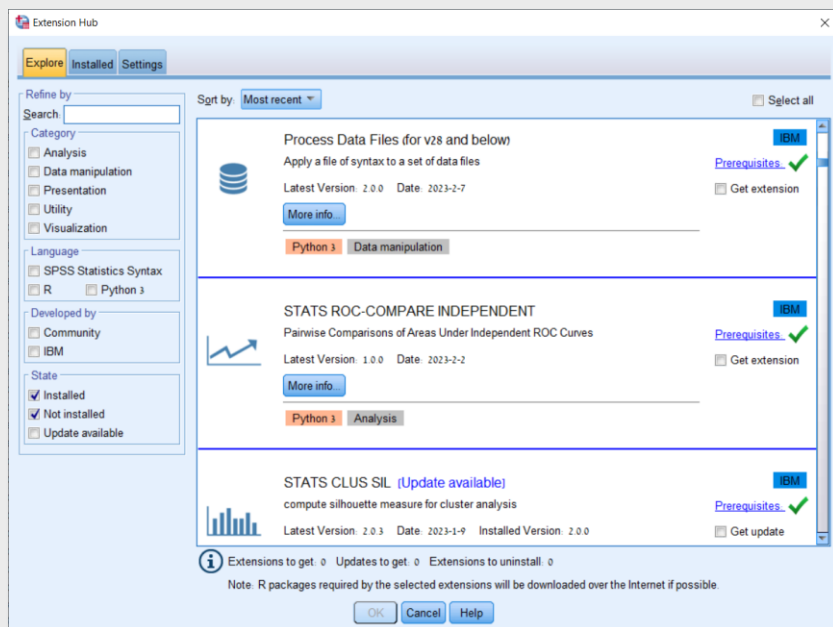
c. Predictors: (Constant), Other debt in thousands, Household income in thousands, Debt to income ratio (x100)

d. Predictors: (Constant), Household income in thousands, Debt to income ratio (x100)

e. Predictors: (Constant), Household income in thousands, Debt to income ratio (x100), Years with current employer

f. Dependent Variable: Credit card debt in thousands

\*predicted residual sum of squares



- Extension Hub - חיבור ישיר ל-Repository ב-GitHub המכיל תוספות לתוכנה במגוון נושאים ושפות קוד. ממשק חיפוש והתקנה יעיל ופשוט.

- מגוון פונקציות וניתוחים סטטיסטיים זמינים להורדה גם ישירות מתוך הפורטל (קישור בצ'אט)

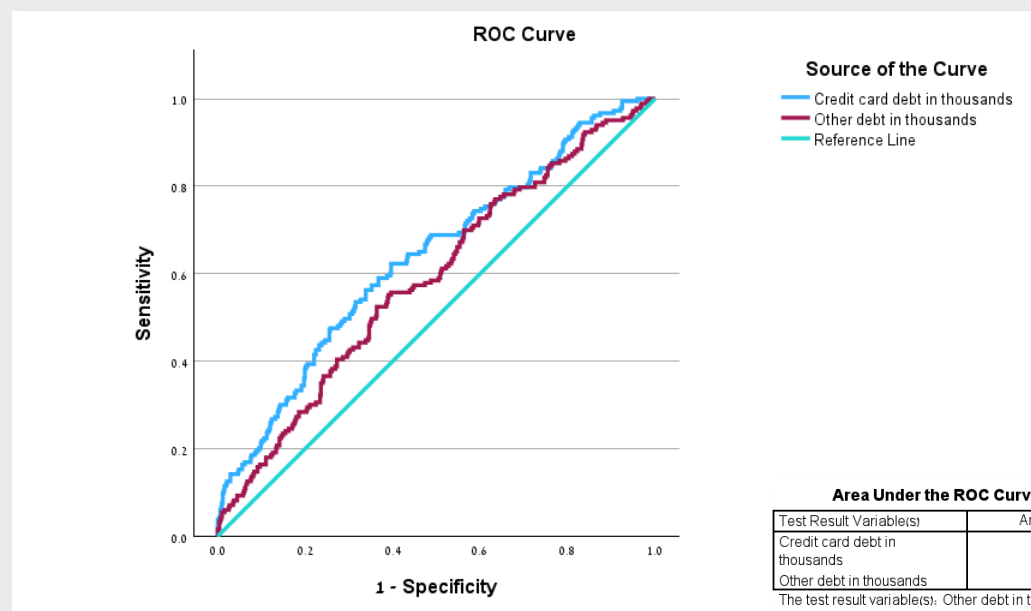
- פקודת Begin Program מאפשרת לכתוב קוד R או Python ישירות בתוך ה-Syntax

- ניתן לייצר תיבות דו-שיח מותאמות אישית באמצעות ממשק Custom Dialog Builder for Extensions כדי להריץ קוד Open Source בנוחות.

- החל מגרסה 29.0.1 ניתן להתקין חבילות R ו-Python ישירות מתוך SPSS ובכך להקל על תהליך הרצת הקוד.

בנוסף להשוואת AUC (מבחן DeLong) התווסף מדד Youden's J :

מאפשר אינדקס המשלב בין Sensitivity לבין Specificity למדד יחיד ופשוט בכל נקודת חיתוך של עקומת ה-ROC



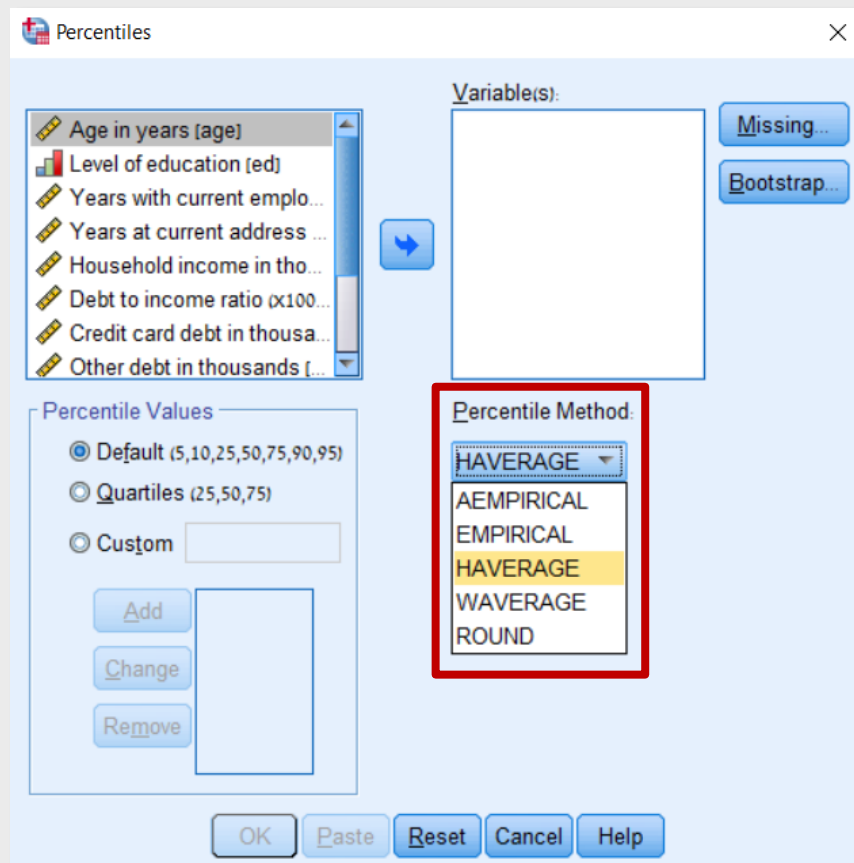
**Paired-Sample Area Difference Under the ROC Curves**

| Test Result Pair(s) | Asymptotic |                            | AUC Difference | Std. Error Difference <sup>b</sup> | Asymptotic 95% Confidence Interval |             |
|---------------------|------------|----------------------------|----------------|------------------------------------|------------------------------------|-------------|
|                     | z          | Sig. (2-tail) <sup>a</sup> |                |                                    | Lower Bound                        | Upper Bound |
| creddebt - othdebt  | 2.270      | .023                       | .049           | .219                               | .007                               | .092        |

a. Null hypothesis: true area difference = 0  
b. Under the nonparametric assumption

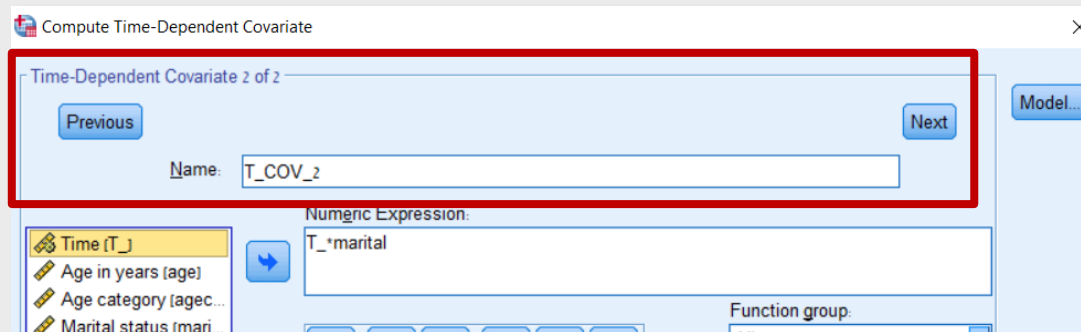
**Coordinates of the ROC Curve**

| Test Result Variable(s)       | Positive if Greater Than or Equal To <sup>a</sup> | Sensitivity | 1 - Specificity | Youden's Index |
|-------------------------------|---------------------------------------------------|-------------|-----------------|----------------|
| Credit card debt in thousands | -.9883                                            | 1.000       | 1.000           | .000           |
|                               | .0133                                             | 1.000       | .998            | .002           |
|                               | .0197                                             | 1.000       | .996            | .004           |
|                               | .0246                                             | 1.000       | .994            | .006           |
|                               | .0248                                             | 1.000       | .992            | .008           |
|                               | .0272                                             | 1.000       | .990            | .010           |
|                               | .0297                                             | 1.000       | .988            | .012           |
|                               | .0301                                             | 1.000       | .986            | .014           |
|                               | .0304                                             | 1.000       | .985            | .015           |
|                               | .0306                                             | 1.000       | .983            | .017           |



- זמין גם דרך תפריט **!Explore**
- מאפשר התמקדות בניתוח אחוזונים
- תפריט **Percentile Method** מאפשר לבחור את שיטת החיתוך של המדגם לאחוזונים
- מאפשר בקרה גבוהה על אופן החלוקה

- רגרסיית COX - ניבוי אירוע קריטי בתרחישים א-פרמטריים.
- תחת הנחות המודל ניתן לנבא תחלואה ותמותה, מאורע התנהגותי כמו נטישת לקוחות או רכישת מוצר/שרות, כשל ברכיבי מכונה ועוד!
- המודל משתמש בנתונים מסוג "זמן לאירוע" (הנחת Proportional Hazards).
- ניתן לבדוק את תלותו של Covariate במשתנה הזמן - Time Dependent Covariate, כלומר, האם חל שינוי באפקט ה-Covariate לאורך הזמן עד לאירוע המנובא
- החל מגרסה 29.0.1 ניתן להוסיף בקלות יותר מ-TDC אחד ובכך ליצור מודלים מדויקים ואמינים יותר.



**מודל הישרדות פרמטרי מאפשר לבחון את זמן ההישרדות אם הוא מתפלג באופן ידוע.**

**גרסה 29.0.1 מציעה 2 שיטות חדשות:**

- **Parametric Accelerated Failure Time:** מאפשר לבנות מודל הישרדות פרמטרי עם משתנים מפקחים הדועכים או מתגברים באופן פרופורציונלי לזמן ההישרדות.
- **Parametric Shared Frailty:** מאפשר לבנות מודל הישרדות המכיל נתוני תצפיות באשכולות עם משתנים לטנטיים ורנדומליים בין קבוצות.

# Output Modify - עיצוב מותנה

- מגוון רחב של התניות לעיצוב כל סוגי הטבלאות
- זמין גם בתצורת Syntax לשימוש חוזר
- נגישת באמצעות לחצן ימני בפלט

|               |             |                              | Level of education |            |
|---------------|-------------|------------------------------|--------------------|------------|
|               |             |                              | Count              | Column N % |
| Social status | Lower class | Did not complete high school | 203                | 45.8%      |
|               |             | High school degree           | 176                | 39.7%      |
|               |             | College degree               | 56                 | 12.6%      |
|               |             | Post-undergraduate degree    | 8                  | 1.8%       |
|               |             | Total                        | 443                | 100.0%     |
| Middle class  |             | Did not complete high school | 152                | 38.5%      |
|               |             | High school degree           | 153                | 38.7%      |
|               |             | College degree               | 77                 | 19.5%      |
|               |             | Post-undergraduate degree    | 13                 | 3.3%       |
|               |             | Total                        | 395                | 100.0%     |
| Upper class   |             | Did not complete high school | 58                 | 35.8%      |
|               |             | High school degree           | 73                 | 45.1%      |
|               |             | College degree               | 24                 | 14.8%      |
|               |             | Post-undergraduate degree    | 7                  | 4.3%       |
|               |             | Total                        | 162                | 100.0%     |
| Total         |             | Did not complete high school | 413                | 41.3%      |
|               |             | High school degree           | 402                | 40.2%      |
|               |             | College degree               | 157                | 15.7%      |
|               |             | Post-undergraduate degree    | 28                 | 2.8%       |
|               |             | Total                        | 1000               | 100.0%     |

Modify Output - Document6

Modify the properties of individual or groups of objects in the Output document. Property settings are applied, in order, top-down. Later changes to property settings may override previous property settings.

These changes cannot be undone. If you want to restore the document to its original state, you should save a copy before applying the changes.

Create a Backup of the Output

Selections and Properties

| Selection               | Type                | Delete | Visible | Properties              |
|-------------------------|---------------------|--------|---------|-------------------------|
| All Tables/Custom Table | Tables/Custom Table | No     | As Is   | Table Title: A New Look |

Add  
Duplicate  
Move Up  
Move Down  
Delete

Create a report of the property changes

| Property            | Value                                             |
|---------------------|---------------------------------------------------|
| Indexing Format     |                                                   |
| Table Title         | A New Look                                        |
| Table Look          | Cobalt                                            |
| Transparencies      |                                                   |
| Conditional Styling | PERCENT, -10.0<=COUNT, -10.0<=COUNT, -10.0<=COUNT |
| Sort                | Column N %                                        |
| Sort Direction      | Ascending                                         |
| Table Comment       |                                                   |

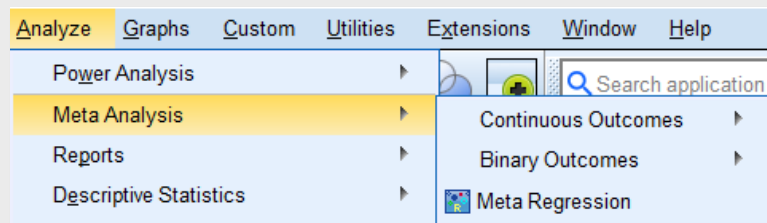
OK | Paste | Reset | Cancel | Help



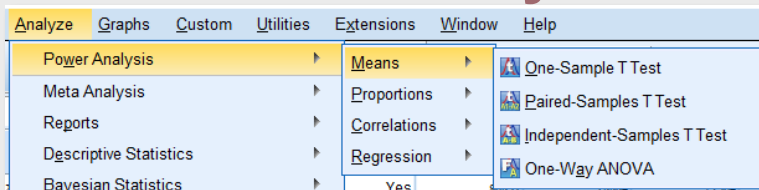
A New Look!

|               |             |                              | Level of education |            |
|---------------|-------------|------------------------------|--------------------|------------|
|               |             |                              | Count              | Column N % |
| Social status | Lower class | Did not complete high school | 203                | 45.8%      |
|               |             | High school degree           | 176                | 39.7%      |
|               |             | College degree               | 56                 | 12.6%      |
|               |             | Post-undergraduate degree    | 8                  | 1.8%       |
|               |             | Total                        | 443                | 100.0%     |
| Middle class  |             | Did not complete high school | 152                | 38.5%      |
|               |             | High school degree           | 153                | 38.7%      |
|               |             | College degree               | 77                 | 19.5%      |
|               |             | Post-undergraduate degree    | 13                 | 3.3%       |
|               |             | Total                        | 395                | 100.0%     |
| Upper class   |             | Did not complete high school | 58                 | 35.8%      |
|               |             | High school degree           | 73                 | 45.1%      |
|               |             | College degree               | 24                 | 14.8%      |
|               |             | Post-undergraduate degree    | 7                  | 4.3%       |
|               |             | Total                        | 162                | 100.0%     |
| Total         |             | Did not complete high school | 413                | 41.3%      |
|               |             | High school degree           | 402                | 40.2%      |
|               |             | College degree               | 157                | 15.7%      |
|               |             | Post-undergraduate degree    | 28                 | 2.8%       |
|               |             | Total                        | 1000               | 100.0%     |

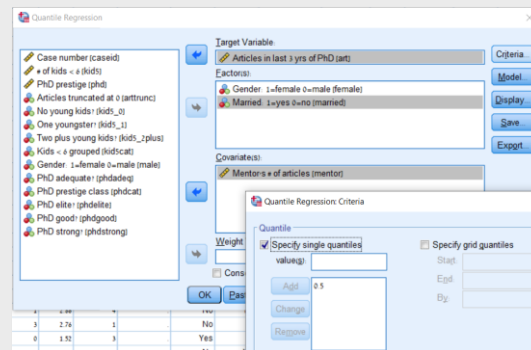
## Meta-Analysis



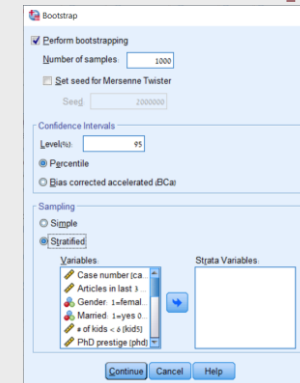
## Power-Analysis



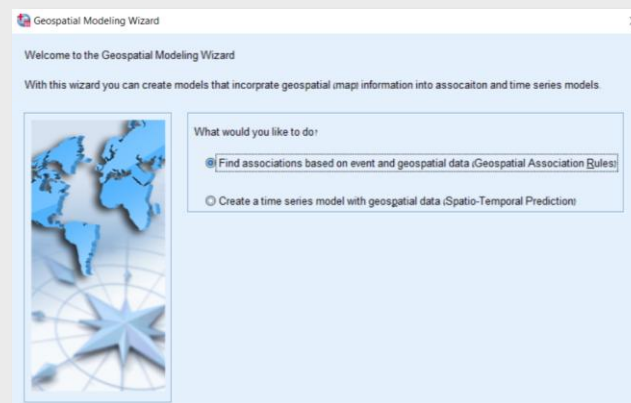
## Quantile Regression



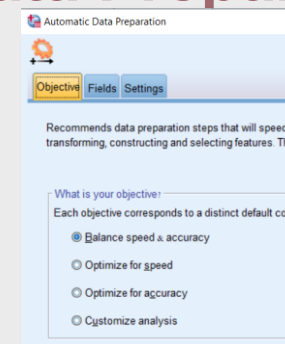
## Bootstrap



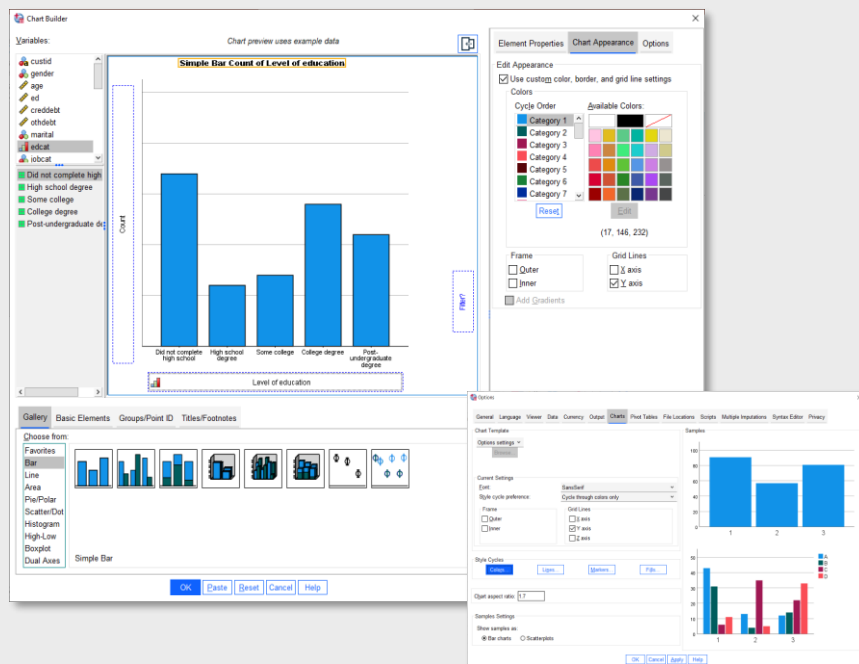
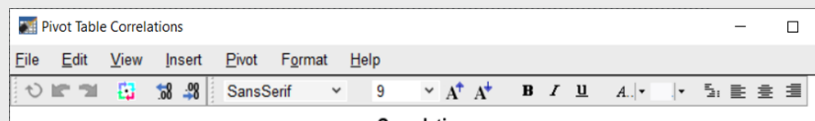
## Geospatial Modeling



## Data Preparation



## עיצוב תרשימים ולוחות



## רב"ס והצגת משולש תחתון - Bivariate Correlation

**Correlations**

|          |                     | creddebt | othdebt | employ | debtinc |
|----------|---------------------|----------|---------|--------|---------|
| creddebt | Pearson Correlation | --       |         |        |         |
|          | N                   | 850      |         |        |         |
| othdebt  | Pearson Correlation | .645**   | --      |        |         |
|          | Sig. (2-tailed)     | <.001    |         |        |         |
|          | N                   | 850      | 850     |        |         |
| employ   | Pearson Correlation | .382**   | .414**  | --     |         |
|          | Sig. (2-tailed)     | <.001    | <.001   |        |         |
|          | N                   | 850      | 850     | 850    |         |
| debtinc  | Pearson Correlation | .515**   | .573**  | -.034  | --      |
|          | Sig. (2-tailed)     | <.001    | <.001   | .327   |         |
|          | N                   | 850      | 850     | 850    | 850     |

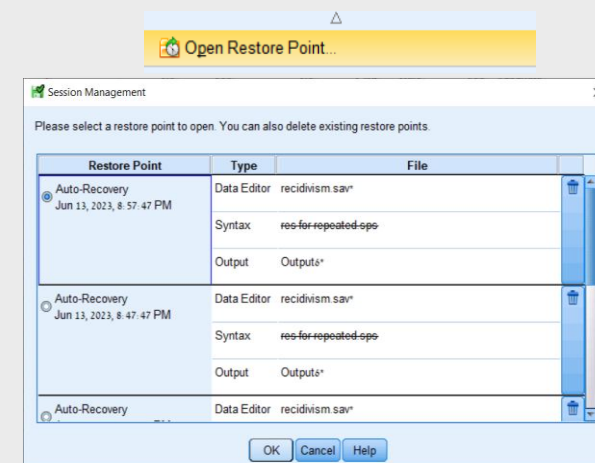
\*\* . Correlation is significant at the 0.01 level (2-tailed).

**Confidence Intervals**

|                    | Pearson Correlation | Sig. (2-tailed) | 95% Confidence Intervals (2-tailed) <sup>a</sup> |       |
|--------------------|---------------------|-----------------|--------------------------------------------------|-------|
|                    |                     |                 | Lower                                            | Upper |
| creddebt - othdebt | .645                | <.001           | .604                                             | .682  |
| creddebt - employ  | .382                | <.001           | .323                                             | .438  |
| creddebt - debtinc | .515                | <.001           | .464                                             | .563  |
| othdebt - employ   | .414                | <.001           | .357                                             | .468  |
| othdebt - debtinc  | .573                | <.001           | .525                                             | .616  |
| employ - debtinc   | -.034               | .327            | -.101                                            | .034  |

a. Estimation is based on Fisher's r-to-z transformation with bias adjustment.

## שחזור אוטומטי וניהול גרסאות



## שורת חיפוש גלובלית בסרגל

| Type         | Content                                                                                |
|--------------|----------------------------------------------------------------------------------------|
| Menu Dialogs | Analyze → Bayesian Statistics → One-way Repeated Measures ANOVA                        |
|              | Analyze → Mixed Models → Linear                                                        |
|              | Analyze → Regression → Quantile                                                        |
|              | Analyze → Missing Value Analysis                                                       |
|              | Analyze → Mixed Models → Generalized Linear                                            |
| Extensions   | Extensions → Extension Hub... (Linear and Integer Programming Problem Solver)          |
| Help Topics  | Core features → Legacy Dialog Charts → Mixed Charts                                    |
|              | Advanced statistics → Linear Mixed Models                                              |
|              | Advanced statistics → Linear Mixed Models → Linear Mixed Models: Subjects and Repeated |