ZOLAXTECH
zolaxtech.com.et

SPSS DATA ANALYSIS GUIDE

FOR CHAPTER FOUR DATA ANALYSIS & INTERPRETATION

From data collection and pre-processing to reliability testing, frequencies, descriptive statistics, correlation, and regression in SPSS

(Quantitative, questionnaire-based research)

A companion to the Final Research Paper Template Guide
www.zolaxtech.com.et

ABOUT THIS GUIDE

This guide walks through the complete practical workflow used to produce a Chapter Four ("Data Analysis and Interpretation") in SPSS starting before you ever open the software, at the moment you collect your questionnaires, and ending with a fully interpreted regression output ready to be written up. It is based on the same analysis pipeline used in an approved MBA thesis studying the effect of leadership style on organizational performance, and it uses SPSS version 26/later menu paths throughout.

Follow the steps in order each step in SPSS builds on the one before it. Skipping data cleaning or variable transformation before running correlation/regression is the single most common reason students get incorrect or inconsistent output.

The Workflow at a Glance

1. Collect the data (questionnaires / Likert-scale instrument)
2. Pre-process the data by hand code, clean, and prepare it before entry
3. Set up SPSS build the Variable View (names, labels, value labels, measurement level)
4. Enter or import the data into the Data View
5. Clean the data inside SPSS (check for out-of-range values and missing data)
6. Transform the data compute composite/mean variables for each construct
7. Run the Reliability Test (Cronbach's Alpha) for each construct
8. Run Frequencies for the demographic questions
9. Run Descriptive Statistics for the Likert-scale (construct) questions
10. Check regression assumptions (normality, linearity, multicollinearity)
11. Run Correlation Analysis to test relationships
12. Run Multiple Linear Regression to test effects
13. Interpret the regression output and test each hypothesis
14. Export tables into your Chapter Four document

PART A BEFORE YOU OPEN SPSS

STEP 1. Collect the Data

Before any SPSS work begins, the questionnaire itself must be well designed, because the quality of your Chapter Four depends entirely on the quality of the data collected here.

What your instrument should contain

- Part I Demographic questions (gender, age, marital status, education, position, tenure) using closed categories
- Part II Construct questions, grouped by variable (one block per independent variable, one block for the dependent variable), each measured on the same 5-point Likert scale (1 = Strongly Disagree … 5 = Strongly Agree)

HOW TO READ THE OUTPUT?

- Keep the same scale direction across all items where possible mixing positively- and negatively-worded items without a plan makes later transformation (reverse-coding) more error-prone.
- Give every questionnaire a unique respondent/serial number before distribution you will need this later to track returned vs. missing forms and to enter data accurately.

Collecting the completed forms

- Distribute according to your sampling plan (e.g., stratified random sampling across departments/branches)
- Log the number distributed vs. returned per stratum as you collect them you will report this as your response rate in Chapter 4.1

Example (from the reference thesis): *343 questionnaires were distributed and 318 were returned and usable a 92.7% response rate, reported at the very start of Section 4.1.*

⚠ ***Common mistake:*** *Don't discard incomplete or unusual responses at collection time decide and document your rule for handling them during pre-processing (Step 2), not on the spot.*

STEP 2. Pre-process the Data (Before Entry)

Pre-processing happens on paper/Excel, before you touch SPSS's Data View. This is where you turn stacks of filled questionnaires into a clean, cod able dataset.

2.1 Screen and sort the returned questionnaires

- Separate fully completed, partially completed, and unusable/blank questionnaires
- Decide and document a consistent rule for partial responses, e.g., "keep if less than 10% of items are missing; treat missing items as system-missing in SPSS"

2.2 Build a codebook

- List every question as a variable with a short SPSS-safe variable name (max 8 characters is safest), a full label, and a coding scheme for every response option

HOW TO READ THE OUTPUT?

- A codebook is simply a table: Variable Name | Question / Label | Values and Meaning | Measurement Level. Build it in Excel before opening SPSS it becomes your direct reference when building the Variable View in Step 3.

Example (from the reference thesis): *Gender → var name "gender": 1 = Male, 2 = Female. Age → var name "age": 1 = 18-25, 2 = 26-35, 3 = 36-45, 4 = 46-55. Leadership items → var names "tls1"-"tls5", "als1"-"als5", "dls1"-"dls5", "els1"-"els5", each coded 1-5 for Strongly Disagree-Strongly Agree. Organizational performance items → "op1"-"op10".*

2.3 Assign a unique ID to every questionnaire

- Number each questionnaire (case ID) in the order it will be entered this lets you trace any suspicious data point in SPSS back to the physical/original form

2.4 Plan for reverse-coding (if applicable)

- Flag any negatively worded item in the codebook now, so you remember to reverse it during Step 6 (Transform) rather than discovering it during analysis

PART B BUILDING THE DATASET IN SPSS

STEP 3. Set Up SPSS (Variable View)

Open SPSS, click the "Variable View" tab at the bottom-left of the Data Editor, and build one row per question/variable using your codebook from Step 2.

IN SPSS DO THIS

1. Open SPSS → click the Variable View tab (bottom left).

2. For each variable, fill in: Name (e.g., tls1), Type (Numeric), Width/Decimals (leave default), Label (the full question wording), Values (define each code, e.g., 1="Strongly Disagree" … 5="Strongly Agree"), Missing (define any missing-value code if used), Measure (Scale for Likert items, Nominal for gender/marital status, Ordinal for age brackets/education/tenure).

3. Repeat for all demographic variables and all construct items (tls1-tls5, als1-als5, dls1-dls5, els1-els5, op1-op8).

HOW TO READ THE OUTPUT?

- Setting the correct "Measure" (Scale/Ordinal/Nominal) matters because SPSS uses it to decide which charts and tests are offered later Likert items are conventionally treated as Scale for the purposes of computing means and running correlation/regression.
- Defining Value Labels now means SPSS will display "Male/Female" or "Strongly Agree" instead of raw numbers in your output tables later this saves you from manually relabeling every table before pasting it into Chapter Four.

⚠ Common mistake: *Typing data first and defining variables later is possible, but it multiplies the chance of coding errors. Always build the Variable View from your codebook before entering a single case.*

STEP 4. Enter or Import the Data

Option A Typing data directly into SPSS

IN SPSS DO THIS

1. Click the Data View tab.
2. Each row = one respondent (case); each column = one variable, in the same order as your Variable View / codebook.
3. Type the coded numeric value for every answer, using the respondent's questionnaire ID as your first column for traceability.

Option B Entering in Excel, then importing (recommended for large samples)

IN SPSS DO THIS

1. Build a spreadsheet in Excel with one column per variable name (matching your SPSS variable names exactly) and one row per respondent.
2. In SPSS: File → Import Data → Excel…, browse to the file, confirm the first row contains variable names, and click OK.
3. Re-open Variable View afterward and confirm labels/measurement levels imported correctly Excel import brings in raw numbers, not your value labels, so you may need to re-apply Value Labels from your codebook.

HOW TO READ THE OUTPUT?

- Whichever method you use, spot-check at least 10% of cases against the original questionnaires after entry this is your basic data-entry quality control step before any analysis begins.

STEP 5. Clean the Data Inside SPSS

Before computing anything, confirm there are no impossible values (e.g., a "7" on a 1-5 scale) and understand your missing-data pattern.

IN SPSS DO THIS

1. Analyze → Descriptive Statistics → Frequencies.
2. Move all variables into the "Variable(s)" box and click OK.
3. Scan the Minimum/Maximum (or the frequency table) for each variable for any value outside the expected range (e.g., outside 1-5), and check the "Missing" count for each.

HOW TO READ THE OUTPUT?

- Any out-of-range value (e.g., a stray "0" or "9") means a data-entry error go back to the original questionnaire using the case's ID number and correct it.
- A small number of "System Missing" cases per item is normal for real survey data (this is exactly why the demographic tables in the thesis example show totals of 315-318 instead of a flat 318 for every item) you do not need to discard these cases, SPSS will exclude them list wise or pairwise depending on the test.

Example (from the reference thesis): *Several construct items in the reference thesis show N = 315 or 316 instead of 318 this simply reflects a few respondents skipping specific questions, which SPSS automatically excludes from that item's statistics.*

STEP 6. Transform the Data (Compute Composite Variables)

Each construct (Transformational Leadership, Authoritarian Leadership, Delegative Leadership, Ethical Leadership, Organizational Performance) is measured with several items. Before you can run correlation or regression at the variable level, you must combine each construct's items into one composite score per respondent.

6.1 Reverse-code any negatively worded items first (if flagged in Step 2.4)

IN SPSS DO THIS

1. Transform → Recode into Different Variables.
2. Move the item to recode, name the output variable (e.g., tls3_r), and click "Old and New Values".
3. Map 1→5, 2→4, 3→3, 4→2, 5→1, then click Change → Continue → OK.

6.2 Compute a mean/sum score for each construct

IN SPSS DO THIS

1. Transform → Compute Variable.
2. Target Variable: type a short name for the new composite, e.g., TLSNEW1.
3. Numeric Expression: use the Mean function, e.g., MEAN (tls1, tls2, tls3, tls4, tls5).
4. Click OK. Repeat for each construct: ALSNEW1, DLSNEW1, ELSNEW1, and the dependent variable OPNEW1.

HOW TO READ THE OUTPUT?

- Using MEAN () (rather than SUM ()) keeps the composite score on the same 1-5 scale as the original items, which makes it directly interpretable against your mean-interpretation scale (Step 8) and easy to explain in the write-up.
- Go back to Variable View and add a Label and set Measure = Scale for each new composite variable you will be using these five new variables (TLSNEW1, ALSNEW1, DLSNEW1, ELSNEW1, OPNEW1) for every remaining step: reliability, descriptive stats, correlation, and regression.

⚠ ***Common mistake:*** *Running correlation or regression directly on individual items (tls1, tls2 ...) instead of the composite construct variable (TLSNEW1) is a common error always analyse relationships at the construct (composite) level, not the item level.*

PART C RUNNING THE CHAPTER FOUR ANALYSES

STEP 7. Run the Reliability Test (Cronbach's Alpha)

Reliability is tested per construct, using the original items (not the composite), to confirm the items in each construct consistently measure the same underlying idea. This produces the reliability table that belongs in Chapter Three (Section 3.9), and is normally the very first test you run once data entry is complete.

IN SPSS DO THIS

1. Analyze → Scale → Reliability Analysis.

2. Move the items for ONE construct only into the "Items" box (e.g., tls1-tls5 for Transformational Leadership).

3. Model: Alpha (default). Click Statistics → tick "Scale if item deleted" (optional but useful) → Continue → OK.

4. Repeat separately for each construct (Authoritarian, Delegative, Ethical, Organizational Performance) and once more with ALL items together for the "Total reliability" row.

HOW TO READ THE OUTPUT?

- Look at "Cronbach's Alpha" in the Reliability Statistics table a value of 0.70 or higher is generally considered acceptable; 0.60-0.70 is marginally acceptable; below 0.60 signals the items may not be measuring the same construct.
- If alpha is low, check the "Cronbach's Alpha if Item Deleted" column an item that, if removed, would raise the overall alpha significantly may be a candidate to drop or reword in future studies.
- Report the Alpha value and number of items for each construct in a single reliability table (see the Chapter Three template guide).

Example (from the reference thesis): *Transformational leadership: α = .879 (5 items). Authoritarian leadership: α = .902 (5 items). Delegative leadership: α = .915 (5 items). Ethical leadership: α = .959 (5 items). Organizational performance: α = .812 (8 items). Overall: α = .938 (28 items) all comfortably above the 0.70 threshold.*

STEP 8. Run Frequencies for the Demographic Questions

This produces the Chapter 4.1 "Demographic Characteristics of Respondents" tables and charts.

IN SPSS DO THIS

1. Analyze → Descriptive Statistics → Frequencies.

2. Move all demographic variables (gender, age, marital status, education, position, tenure) into the "Variable(s)" box.

3. Click Charts → choose Pie Charts (for gender/marital status) or Bar Charts (for age/position/tenure) → Continue.

4. Click OK.

HOW TO READ THE OUTPUT?

- Each output table gives Frequency, Percent, Valid Percent, and Cumulative Percent use "Valid Percent" when writing up results if there is missing data, since it excludes missing cases from the denominator.
- Copy each chart directly from the SPSS Output Viewer (right-click → Copy, or Export) into your Chapter 4.1 figures, and describe the pattern in one sentence per table (e.g., "the majority of respondents, 61.3%, fall in the 26-35 age range").

Example (from the reference thesis): *Frequencies showed 62.3% male vs. 35.5% female respondents, with the 26-35 age bracket the largest group at 61.3% exactly the pattern reported in Table 4.1 and Figures 4.1-4.3.*

STEP 9. Run Descriptive Statistics for the Likert-Scale (Construct) Questions

This produces the item-by-item descriptive tables in Chapter 4.2 (one table per construct) showing how respondents rated each statement.

IN SPSS DO THIS

1. Analyze → Descriptive Statistics → Descriptives.
2. Move the items of ONE construct (e.g., tls1-tls5) plus its composite variable (TLSNEW1) into the "Variable(s)" box.
3. Click Options → tick Mean, Std. Deviation, Minimum, Maximum → Continue → OK.
4. Repeat for each construct separately (Authoritarian, Delegative, Ethical, Organizational Performance) so each gets its own clean table.

HOW TO READ THE OUTPUT?

- Interpret every item's mean using a defined mean-interpretation scale, e.g.: 1.00-1.80 = strongly disagree, 1.81-2.60 = disagree, 2.61-3.40 = neutral, 3.41-4.20 = agree, 4.21-5.00 = strongly agree.
- Write one interpretive sentence per item ("a mean of 4.05 shows respondents agree that …"), then close with the composite variable's overall mean and standard deviation as the summary line for that construct.
- A low standard deviation means respondents answered fairly similarly; a high standard deviation (roughly above 1.0 on a 5-point scale) means opinions were more spread out or divided.

Example (from the reference thesis): *Ethical leadership's overall mean was 3.26 (neutral zone), while delegative leadership's overall mean was 4.10 (agree zone) directly setting up the Discussion's point that ethical leadership was the least clearly present style.*

STEP 10. Check Regression Assumptions

Before running (or when reporting) the regression itself, SPSS output must show that its underlying assumptions are reasonably met. These checks are usually generated together with the regression in Step 12, but are reported first in Chapter 4.2.2.1, ahead of the model results.

10.1 Normality (Skewness and Kurtosis)

IN SPSS DO THIS

1. Analyze → Descriptive Statistics → Descriptives.
2. Move the five composite variables (TLSNEW1, ALSNEW1, DLSNEW1, ELSNEW1, OPNEW1) into the box.
3. Click Options → tick Skewness and Kurtosis → Continue → OK.

HOW TO READ THE OUTPUT?

- An absolute skewness value between -2 and +2, and an absolute kurtosis value between -7 and +7, is generally treated as acceptably close to a normal distribution for parametric tests such as regression.

Example (from the reference thesis): *All five composite variables showed skewness between -1.48 and -0.51 and kurtosis between -1.24 and 2.29 within the accepted range, supporting the use of regression.*

10.2 Linearity

IN SPSS DO THIS

1. This plot is generated automatically as part of the regression procedure (see Step 12): tick "Normal probability plot" under Plots.

2. Alternatively: Graphs → Legacy Dialogs → P-P Plot, select the standardized residual of your regression.

HOW TO READ THE OUTPUT?

- If the plotted points fall reasonably close to the diagonal reference line, the linearity assumption is considered satisfied.

10.3 Multicollinearity

IN SPSS DO THIS

1. This is requested inside the regression dialog (Step 12): Analyze → Regression → Linear → Statistics → tick "Collinearity diagnostics".

HOW TO READ THE OUTPUT?

- Check the Tolerance and VIF (Variance Inflation Factor) values in the Coefficients table: Tolerance above 0.10 and VIF below 10 indicate multicollinearity is not a serious problem among your independent variables.

STEP 11. Run Correlation Analysis (Test Relationships)

Correlation is run on the composite (construct-level) variables to test whether and how strongly each independent variable relates to organizational performance, answering the "is there a relationship" part of your research questions.

IN SPSS DO THIS

1. Analyze → Correlate → Bivariate.

2. Move all five composite variables (TLSNEW1, ALSNEW1, DLSNEW1, ELSNEW1, OPNEW1) into the "Variables" box.

3. Correlation Coefficients: tick Pearson. Test of Significance: Two-tailed. Tick "Flag significant correlations".

4. Click OK.

HOW TO READ THE OUTPUT?

- Read the r-value (Pearson Correlation) and Sig. (2-tailed) value for each independent-variable/dependent-variable pair.
- Use Cohen's (1988) guideline to describe strength: r = .10-.29 small, .30-.49 medium, .50-1.0 large (same logic applies to negative values).
- A relationship is statistically significant when Sig. is below 0.05 (marked with one asterisk*) or below 0.01 (two asterisks**) in the SPSS output.

Example (from the reference thesis): *TLS↔OP: r = .686, p = .000 (large, significant). ALS↔OP: r = .726, p = .000 (large, significant). DLS↔OP: r = .691, p = .000 (large, significant). ELS↔OP: r = .138, p = .014 (small, but still significant) showing every independent variable relates to organizational performance, though ethical leadership's relationship is comparatively weak.*

STEP 12. Run Multiple Linear Regression (Test Effects)

Regression goes a step further than correlation: it tests how much each independent variable actually predicts/explains organizational performance, while controlling for the other independent variables this is the core statistical test behind your hypotheses.

IN SPSS DO THIS

1. Analyze → Regression → Linear.
2. Move your dependent variable (OPNEW1) into the "Dependent" box.
3. Move all independent variables (TLSNEW1, ALSNEW1, DLSNEW1, ELSNEW1) into the "Independent(s)" box. Method: Enter.
4. Click Statistics → tick Estimates, Model fit, R squared change, Collinearity diagnostics → Continue.
5. Click Plots → move *ZRESID to Y and *ZPRED to X (for homoscedasticity) → tick "Normal probability plot" → Continue.
6. Click OK.

12.1 Model Summary table

HOW TO READ THE OUTPUT?

- R = the overall correlation between all independent variables combined and the dependent variable.
- Adjusted R Square = the percentage of variance in the dependent variable explained by the independent variables together report this as "X% of the variance in organizational performance is explained by the leadership style variables".

Example (from the reference thesis): *R = .760, Adjusted R^2 = .572 → the four leadership styles together explain 57.2% of the variance in organizational performance.*

12.2 ANOVA table

HOW TO READ THE OUTPUT?

- Check the Sig. value for the model's F-statistic: if Sig. < 0.05, the regression model as a whole significantly predicts the dependent variable.

Example (from the reference thesis): *F = 105.905, Sig. = .000 the overall model is statistically significant.*

12.3 Coefficients table

HOW TO READ THE OUTPUT?

- For each independent variable, read the Sig. value: below 0.05 means that variable has a statistically significant effect on the dependent variable, holding the others constant.

- Compare the Standardized Beta values across variables (not the unstandardized B) to judge which independent variable has the strongest relative effect the largest absolute Beta is the strongest predictor.
- The sign of B/Beta shows direction: positive means the independent variable increases organizational performance; negative means it decreases it.

Example (from the reference thesis): *Authoritarian leadership: B = .573, Beta = .403, Sig. = .000 (significant, strongest effect). Transformational leadership: B = .331, Beta = .219, Sig. = .001 (significant). Delegative leadership: B = .299, Beta = .195, Sig. = .011 (significant). Ethical leadership: B = -.009, Beta = -.009, Sig. = .811 (not significant).*

STEP 13. Interpret the Output and Test Each Hypothesis

With the Coefficients table in hand, go back to each null hypothesis from Chapter One and apply a single decision rule.

HOW TO READ THE OUTPUT?

- Decision rule: if the Sig. value for that variable in the Coefficients table is less than 0.05, reject the null hypothesis (the effect is significant); if Sig. is 0.05 or greater, fail to reject the null hypothesis (the effect is not significant).
- Build a one-row-per-hypothesis summary table: Hypothesis | Sig. value | Decision (Reject/Accept H0) this becomes your Chapter 4.3/4.4 hypothesis summary table.

Example (from the reference thesis): *H01 (Transformational leadership): Sig. = .001 < .05 → reject H0, the effect is significant. H04 (Ethical leadership): Sig. = .811 > .05 → fail to reject H0, the effect is not significant.*

STEP 14. Export Tables into Your Chapter Four Document

IN SPSS DO THIS

1. In the Output Viewer, right-click any table → Copy (for pasting as an image) or Copy Special → choose "RTF" or "HTML" to paste as an editable Word table.

2. For a full clean export: File → Export… choose Document type Word/RTF, tick "Output Document", select the objects to include, and Export.

3. For charts: right-click the chart → Copy, then paste directly into your Chapter Four figure placeholder.

HOW TO READ THE OUTPUT?

- After pasting each table into Word, reformat it to match your thesis's table style (borders, font, caption numbering) rather than leaving SPSS's default grey shading.
- Caption every table and figure immediately after pasting it (Table 4.x / Figure 4.x) so numbering stays correct as you continue writing.

CHAPTER FOUR SPSS CHECKLIST

Before you consider Chapter Four's analysis complete, confirm:

- Data was cleaned (no out-of-range values) before any test was run
- Reverse-coding (if needed) was applied before computing composite variables
- Composite (mean) variables were created for every construct and used for correlation/regression not the raw items
- Reliability (Cronbach's Alpha) was reported for every construct and the overall scale
- Frequencies (with charts) were run for every demographic variable, and the response rate is reported
- Descriptive statistics (mean, SD) were run and interpreted for every construct item and composite
- Normality, linearity, and multicollinearity were checked and reported before the regression results
- Correlation was run and interpreted (strength + significance) for every independent variable against the dependent variable
- Regression was run with Model Summary, ANOVA, and Coefficients tables, and each is interpreted in the text
- Every hypothesis from Chapter One has a matching decision (reject/fail to reject) drawn directly from the Coefficients Table's Sig. values
- All SPSS tables/charts pasted into the document are captioned and formatted consistently

zolaxtech.com.et