



Task 6: Data Analysis

Prepared by

ALI MMADI, PhD student-Midocean University



Research Issues (2)

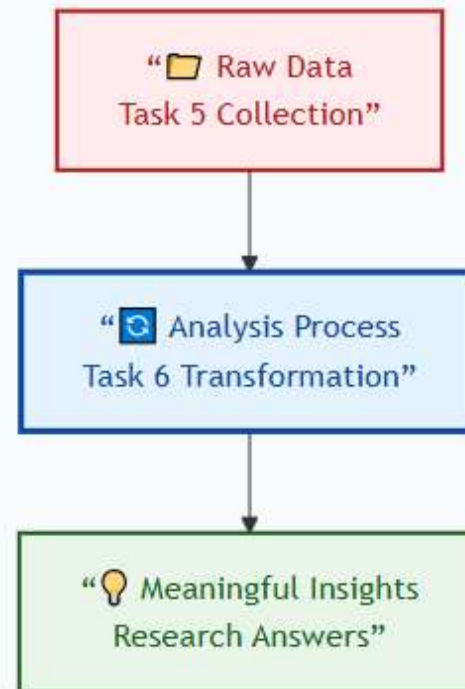
Task 6

Introduction and link with the previous task

Assalam Anlaikum,
We are now moving on to task 6, which focuses on data analysis. This step is crucial because it transforms the raw data collected during the previous task into meaningful and interpretable results.

If data collection provided us with the raw material, analysis is the process that will give it meaning and value to answer our research questions.

I will present the analysis methods I will use, explaining their relevance and the clear steps for their implementation.



Relevance of analytical methods

Let's start with the relevance of the methods. In line with our mixed approach, I will use three types of analysis:

Data Type	Analysis Method	Relevance
Quantitative Data (detection rate, performance)	Descriptive and inferential statistical analysis	Allows for objectively measuring effectiveness and generalizing the results
Qualitative Data (logs, alerts, observations)	Thematic and comparative analysis	Understanding the underlying patterns and reasons for performance
Coverage Data (ATT&CK Matrix)	Analysis by gap analysis	Visually identify the strengths and weaknesses of solutions patterns

This combination is relevant because it allows for data triangulation, strengthening the validity of the conclusions.

Steps of quantitative analysis

Let's now move on to the clarity of the steps, starting with the quantitative analysis:

Step 1: Data Cleaning and Preparation

- Checking the integrity of collected data
- Eliminating outliers and standardizing formats

Step 2: Descriptive Statistical Analysis

- Calculating means, medians, and standard deviations for detection rates
- Creating visualizations (box plots, histograms)

Step 3: Inferential statistical tests

- ANOVA** to compare performance between the 3 EDR solutions
- Student's t-test** for pairwise comparisons

-**Significance threshold:** $p < 0.05$

Step 4: Interpretation of results

- Identification of statistically significant differences
- Calculation of effect sizes to assess practical significance

Each test will be repeated 10 times to ensure the reliability of the quantitative data.

Steps of qualitative analysis(Thematic analysis process)

Let's now move on to the clarity of the steps, starting with the quantitative analysis:

Step 1: Coding qualitative data

- Thorough reading of logs and alert reports
- Identification of initial codes (e.g., "late detection", "false positive")

Step 2: Identifying themes

- Grouping codes into emerging themes
- Examples of themes: "effectiveness against fileless techniques", "impact of configurations"

Step 3: Comparative Analysis

- Cross-comparison between the three EDR solutions
- Identification of common patterns and notable differences

Step 4: Triangulation

- Comparison of qualitative results with quantitative data
- Search for convergences and explanation of divergences

Tools and validation

To carry out these analyses, I will use specialized tools:

Type of Analysis	Outils	Validation
Quantitative	Python (pandas, scipy), R	Reproducibility of scripts, peer validation
Qualitative	NVivo, manual analysis	Audit trail, inter-coder agreement
Visualization	Matplotlib, Power BI	Clarity and accuracy of representations

Validation will be ensured by:

- Review by a second researcher for the qualitative analysis
- Random recalculation of 10% of the quantitative analyses
- Analysis journal documenting all methodological decisions

Summary and conclusion

In conclusion, this analytical approach:

- Is perfectly relevant for handling the different types of data collected
- Follows clear and systematic steps ensuring scientific rigor
- Fits within the logical continuity of data collection
- Lays the groundwork for the interpretation and discussion of results

This robust analytical methodology will allow us to draw reliable and actionable conclusions from our experimental data.

Thank you for your attention.

Bibliographic References

1) Statistical Analysis & Quantitative Methods

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). SAGE Publications.

2) Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.

3) McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (2nd ed.). O'Reilly Media.

4) Braun, V., & Clarke, V. (2022). *Thematic Analysis: A Practical Guide*. SAGE Publications.

5) Bazeley, P. (2013). *Qualitative Data Analysis: Practical Strategies*. SAGE Publications.

6) Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE Publications.

7) Guest, G. (2013). *Describing Mixed Methods Research: An Alternative to Typologies*. *Journal of Mixed Methods Research*, 7(2), 141–151.

8) **Heale, R., & Forbes, D. (2013). *Understanding triangulation in research. *Evidence-Based Nursing*, 16(4), 98.**

9) *O'Connor, C., & Joffe, H. (2020). *Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines*. *International Journal of Qualitative Methods*, 19.



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY



Thank You

