

# Research Tools & Methodologies

*Design and Creative Technologies*

*Torrens University, Australia*

**Student:** Luis Guilherme de Barros Andrade Faria - A00187785

**Subject Code:** REM 502

**Subject Name:** Research Methodologies

**Assessment No.:** 2

**Title of Assessment:** Research Tools and Methodologies

**Lecturer:** Dr. Bushra Naeem

**Date:** Nov 2025

Copyright © 2025 by Luis G B A Faria

Permission is hereby granted to make and distribute verbatim copies of this document provided the copyright notice and this permission notice are preserved on all copies.

Table of Contents

1. Introduction .....3

2. Research Questions, Aim and Objectives.....3

3. Comparative Analysis of Research Methodologies .....5

4. Proposed Methodology and Research Methods.....6

    4.1. Design Paradigm.....6

    4.2. Data Collection.....6

    4.3. Data Processing Timeline.....7

    4.4. Quantitative Analysis .....8

5. Rationale for Method Choice.....9

6. Ethical Considerations ..... 10

7. Data Analysis Strategies and Tools ..... 10

8. Limitations and Delimitations..... 11

9. Conclusion..... 11

10. Appendices ..... 13

    10.1. Appendix A – Dataset Overview ..... 13

    10.2. Appendix B – Data Preparation Pipeline ..... 13

    10.3. Appendix C – Descriptive Statistics..... 14

    10.4. Appendix D – Correlation Preview..... 14

11. References ..... 18

# Exploring the Relationship between Net Promoter Score and Revenue Growth in Healthcare Clinics.

## 1. Introduction

The Assessment 2 project continues my study journey on Net Promoter Score, investigating whether this common measure of patient satisfaction, correlates with revenue performance in healthcare clinics. Understanding this relationship can inform evidence-based managerial decisions and guide future AI-enabled feedback systems. In the healthcare sector, however, the bridge between patient sentiment and measurable business outcomes such as revenue, retention, or referrals remains under-explored.

This project proposes an ICT-driven research framework that leverages data analytics and business-intelligence methods to examine whether NPS trends align with financial performance in clinical environments. The study builds directly on the literature gaps identified in Assessment 1 and adopts a quantitative correlational design to ensure statistical accuracy and business relevance.

## 2. Research Questions, Aim and Objectives

**Aim:** To investigate the relationship between Net Promoter Score (NPS) and business performance indicators, specifically revenue growth, in healthcare clinics.

**Research Questions:**

- RQ1: To what extent is NPS correlated with monthly revenue growth in healthcare clinics?
- RQ2: Can NPS trends over time predict short-term revenue fluctuations?

**Hypothesis:**

- H0: There is no statistically significant correlation between NPS scores and monthly revenue.
- H1: There is a positive correlation between NPS scores and monthly revenue.

**Objectives:**

- Collect and prepare Net Promoter Score (NPS) and monthly revenue data from Pro-Corpo's clinics to ensure accuracy and comparability across time periods of the available data.
- Conduct quantitative analysis – including descriptive statistics, correlation, and regression – to examine the relationship between NPS scores and revenue performance.
- Perform correlation analysis (Pearson or Spearman) to quantify the strength and direction of the NPS to revenue relationship).
- Develop a practical ICT framework that demonstrates how NPS-based metrics can inform strategic decision-making and performance evaluation in healthcare.

### 3. Comparative Analysis of Research Methodologies

Healthcare business research operates at the intersection of human experience and organizational performance, requiring methodological approaches that balance measurement rigor with contextual understanding. The choice of methodology fundamentally shapes what can be known: quantitative methods enable hypothesis testing and generalization across populations, while qualitative methods reveal mechanisms and meanings that numbers alone can’t capture (Creswell & Plano Clark, 2023). In ICT research and development (R&D), this tension is particularly salient when translating patient feedback, a qualitative phenomenon, into business intelligence metrics.

This study adopts a **pragmatic paradigm** (Morgan, 2014), prioritizing practical problem-solving over skepticism. Three methodological approaches were evaluated:

Methodology	Description	Strengths	Weakness
Qualitative	Explores human meaning through interviews or thematic coding	Rich context and interpretive depth.	Limited generalizability; prone to researcher bias.
Quantitative	Employs numerical measurement, hypothesis testing, and statistical inference.	Objectivity, replicability, scalability.	May overlook cultural tone.
Mixed Methods	Integrate both qualitative and quantitative strands.	Triangulation improves validity; merges AI outputs with human interpretation.	Requires time and data integration skills.

Given that this study seeks to establish a baseline relationship between existing business metrics (NPS and revenue), a quantitative approach has been chosen as the most appropriate. The availability of large-scale data (27,000 records across 36 months) makes statistical analysis both feasible and methodologically sound (Pallant, 2020). Qualitative methods, while valuable for understanding why correlations exist, require resources (patient interviews, thematic coding) beyond this study's scope. Mixed or qualitative methods may be relevant for future research that integrates NLP-derived sentiment with financial outcomes, but such extensions depend on the foundational quantitative correlation established in this study.

This methodological choice reflects ICT R&D principles: iterative development from simple to complex systems. By demonstrating statistical relationships first, subsequent research can build AI-enabled sentiment analysis with confidence that the underlying NPS-revenue connection merits computational investment.

## 4. Proposed Methodology and Research Methods

### 4.1. Design Paradigm

Given that the research questions seek to quantify relationships and test predictions, a quantitative correlational design is most suitable to establish statistical validity and support evidence-based conclusions (Field, 2018).

### 4.2. Data Collection

- **Primary Source:** Anonymized operational dataset provided by Pro-Corpo Estetica (<https://procorpoestetica.com.br/>), comprising  $\approx 27\,000$  records (2022–2025).
- **Variables:** textual feedback, NPS scores, month/year, clinic ID, and monthly revenue.

- **Data Security:** stored on encrypted drives compliant with the Australian Privacy Act (1988) and Brazilian Data Protection Law.
- **Authorization:** formal company consent letter ensuring confidentiality and academic use only.

The dataset originates from Pro-Corpo’s post-service Net Promoter Score (NPS) program, which automatically invites clients to provide feedback within 24 hours of receiving treatment. Respondents can identify themselves or remain anonymous and answer four brief questions: (1) a 1-to-10 satisfaction rating, (2) optional comments, (3) confirmation of the store visited, and (4) optional mention of staff members for praise or concern. This process has generated approximately 27 000 records collected between 2022 and 2025, providing a rich source of structured (scores, store, month) and unstructured (text feedback) data. Monthly revenue data for each store are also available, enabling correlation between customer sentiment, NPS, and financial performance.

### 4.3. Data Processing Timeline

The plan is to develop the work in 4 phases detailed below:

Milestone	Description
Data extraction	<ul style="list-style-type: none"> <li>• Extract NPS survey responses (score, date, clinic ID, optional text comments)</li> <li>• Extract monthly revenue records (clinic ID, month, total revenue)</li> <li>• Timeframe: January 2022 – December 2024 (36 months)</li> </ul>
Data cleaning	<ul style="list-style-type: none"> <li>• Remove duplicates (based on timestamp + clinic ID)</li> <li>• Handle missing values: <ul style="list-style-type: none"> <li>○ NPS missing: exclude record (cannot impute satisfaction scores)</li> <li>○ Revenue missing: interpolate if isolated gap (linear interpolation); exclude clinic-month if systematic missingness</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>Outlier detection: flag revenue values &gt;3 standard deviations from clinic mean for review</li> </ul>
Data transformation	<ul style="list-style-type: none"> <li>Aggregate NPS to clinic-month level (mean, median, standard deviation, response count)</li> <li>Normalize revenue for clinic size: Revenue per appointment (if appointment data available)</li> <li>Create derived variables such as <b>NPS category</b>: Detractors (0-6), Passives (7-8), Promoters (9-10), <b>Revenue growth</b>: Month-over-month percentage change and <b>Lagged NPS</b>: NPS from 1 and 2 months prior (for temporal analysis)</li> </ul>
Integration	<ul style="list-style-type: none"> <li>Merge datasets on: "clinic_id + year + month"</li> <li>Validate temporal alignment (ensure NPS survey dates precede revenue measurement)</li> <li>Final dataset structure: Each row = one clinic-month observation</li> </ul>

#### 4.4.Quantitative Analysis

- **Step 1: Check Assumptions:** Before running statistical tests, the data must meet certain conditions (Field, 2018): Normality, Linearity, Outliers.
- **Step 2: Describe the Data :** Calculate basic statistics for each clinic and time period like Average NPS score and revenue, show distribution patterns.
- **Step 3: Test Correlation (RQ1):** Measure how strongly NPS and revenue move together:
- **Step 4: Test Prediction (RQ2):** Use regression to see if NPS can predict future revenue:
- **Step 5: Check Robustness:** Verify results are reliable by Re-running tests excluding, testing different time periods separately (2022 vs 2023 vs 2024)



## 5. Rationale for Method Choice

This study adopts a purely **quantitative design** for the methodological reasons below:

- **Nature of Research Questions:** Both RQ1 and RQ2 seek to quantify relationships and predictive power—questions best answered through statistical analysis rather than interpretive methods (Bryman, 2016). The hypothesis-testing framework (H0 vs. H1) requires numerical measurement and inferential statistics to establish evidence-based conclusions.
- **Data Characteristics:** The dataset comprises structured numerical variables (NPS scores 0-10, monthly revenue in currency units) collected systematically across 27,000 observations. This scale and structure align with quantitative methods' strengths: detecting patterns across large samples with statistical confidence (Field, 2018).
- **ICT R&D Context:** In software engineering and IT research, quantitative validation provides the empirical foundation for system design decisions (Wohlin et al., 2012). Establishing whether NPS correlates with revenue — and at what magnitude — determines whether investing in automated NPS analytics systems is justified. Qualitative understanding of **why** correlations exist can follow, but the **whether** question requires quantitative evidence first.

This study focuses solely on quantifiable relationships between NPS and revenue. The optional text comments in NPS surveys could enable qualitative analysis.

This quantitative foundation positions Assessment 3 to propose mixed-methods extensions: "Having established correlation (Assessment 2), future research will employ NLP to identify **which** aspects of patient experience drive revenue outcomes."

## 6. Ethical Considerations

- **Anonymity & Consent:** No personally identifiable data will be used. Pro-Corpo's written authorization ensures institutional consent.
- **Data Governance:** Compliance with the Australian Privacy Act (1988), GDPR, and HIPAA standards.
- **Responsible AI Design:** Any outputs will be advisory, not decision-making.
- **Transparency:** All code and analysis scripts are documented for audit and open-sourced for learning purposes.

## 7. Data Analysis Strategies and Tools

Purpose	Tool / Technique	Outputs	Justification
Data preparation	Python (Pandas, NumPy)	Cleaned dataset	Ensures data quality before analysis
Assumption testing	Excel, Python	Summary stats, averages	Validates distribution assumptions for parametric tests
Descriptive statistics	Exploratory data visualization (scatter plots, line trends)	Validated data integrity and correlation outputs.	Summarizes central tendency and spread for each clinic
Correlation analysis	Streamlit, Matplotlib	Trend plots and correlation heatmaps	Quantifies relationships between NPS and revenue
Regression modeling	Combine AI, NPS and revenue metrics	Composite predictive model	Tests predictive strength of NPS for revenue
Visualization	Streamlit, Matplotlib	Scatterplots, heatmaps, and trend dashboards	Aids interpretation and communication of findings

## 8. Limitations and Delimitations

- **Scope:** Restricted to one clinic group (Pro-Corpo), limiting cross-industry generalizability.
- **Data Bias:** Feedback may be skewed toward extreme experiences.
- **Time Frame:** Analysis limited to 2022–2025 data.
- **Delimitation:** Study focuses on correlational evidence, not causal inference.

While data are anonymized, reflexivity remains essential: as an ex-internal collaborator, the researcher must remain aware of interpretive bias when analyzing familiar organizational contexts.

## 9. Conclusion

This research framework investigates the correlation between Net Promoter Score and revenue growth through a quantitative lens, establishing a foundation for understanding customer experience as a business performance driver in healthcare. By employing descriptive statistics, correlation analysis, and regression modeling, the study provides measurable evidence — or lack thereof — regarding NPS's predictive validity for financial outcomes.

This quantitative approach aligns with ICT R&D principles of evidence-based system design. By first testing the empirical strength of the NPS–revenue relationship, this study provides a statistical foundation that future research can build upon, for example, integrating predictive modelling or AI-enhanced feedback systems, but the fundamental questions must first be answered with statistical rigor: Does NPS correlate with revenue? If so, how strongly? Can past satisfaction scores predict future financial performance? These are prerequisite questions that computational solutions ultimately depend upon.

By establishing baseline correlations in Assessment 2, Assessment 3 can propose extensions: if NPS shows strong predictive power, simpler tracking systems suffice; if correlations are weak, more sophisticated NLP-enabled sentiment analysis becomes an option to capture the nuances traditional metrics miss. This progression from foundational quantitative research to advanced mixed-methods proposals reflects iterative R&D practice — validate core assumptions before building complex systems.

## 10. Appendices

### 10.1. Appendix A – Dataset Overview

Column	Description	Example
store_name	Clinic identifier	“LAPA”
month	Reference month of transaction	2025-04
nps_score	Net Promoter Score from post-service	9
revenue	Monthly clinic revenue (BRL)	85.200
responses	Number of NPS responses in that period	112

Data are anonymized and aggregated by store and month, ensuring no personally identifiable information is retained.

### 10.2. Appendix B – Data Preparation Pipeline

```

✓ # Merging by store and month
df = pd.merge(nps, revenue, on=["Store", "month"], how="inner")

# Computing monthly averages
df["avg_nps"] = df.groupby("Store")["NPS"].transform("mean")

# Dataframe preview
df.head()
] ✓ 0.0s

```

	month	Store	NPS	count_of_responses	revenue	count_of_orders	avg_nps
0	2024-01	BELO HORIZONTE	65.000000	40	111643.77	182	65.000000
1	2024-01	CAMPINAS	91.044776	67	114131.31	246	75.986733
2	2024-01	COPACABANA	80.281690	71	225878.13	250	74.070324
3	2024-01	IPIRANGA	69.230769	65	234789.84	305	69.230769
4	2024-01	ITAIM	30.769231	13	99281.93	128	55.494505

Figure 1: Code example of pandas' dataframe merging

The dataset was merged using store and month as common keys. Only complete entries were retained to maintain consistency in correlation analysis.

### 10.3. Appendix C – Descriptive Statistics

```
[34] df.describe()[["NPS", "revenue"]]
```

	NPS	revenue
count	67.000000	67.000000
mean	75.820018	179285.282836
std	28.704604	112676.584657
min	-100.000000	48331.610000
25%	66.666667	95715.870000
50%	81.818182	158653.140000
75%	92.397388	222530.025000
max	100.000000	601662.020000

Figure 2: Descriptive statistics of Net Promoter Score and revenue across all clinics.

### 10.4. Appendix D – Correlation Preview

The analysis has not been finished yet, but still it is possible to see the beginning of the studies of correlation between NPS score and Revenue income generated on the stores. More to come in the next analysis.

```
corr = df[["NPS", "revenue"]].corr(method="pearson")
corr
```

	NPS	revenue
NPS	1.000000	0.010176
revenue	0.010176	1.000000

Figure 3: Preliminary Pearson correlation matrix..

Heatmap:

```
# Add correlation heatmap between revenue and NPS
import seaborn as sns
import matplotlib.pyplot as plt

# Create correlation heatmap
plt.figure(figsize=(5, 4))
correlation_matrix = df[['revenue', 'NPS']].corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', center=0,
            square=True, linewidths=0.5)
plt.title('Correlation between Revenue and NPS')
plt.tight_layout()
plt.show()

# Print correlation coefficient
correlation = df['revenue'].corr(df['NPS'])
print(f"Correlation coefficient: {correlation:.3f}")
```

✓ 0.0s

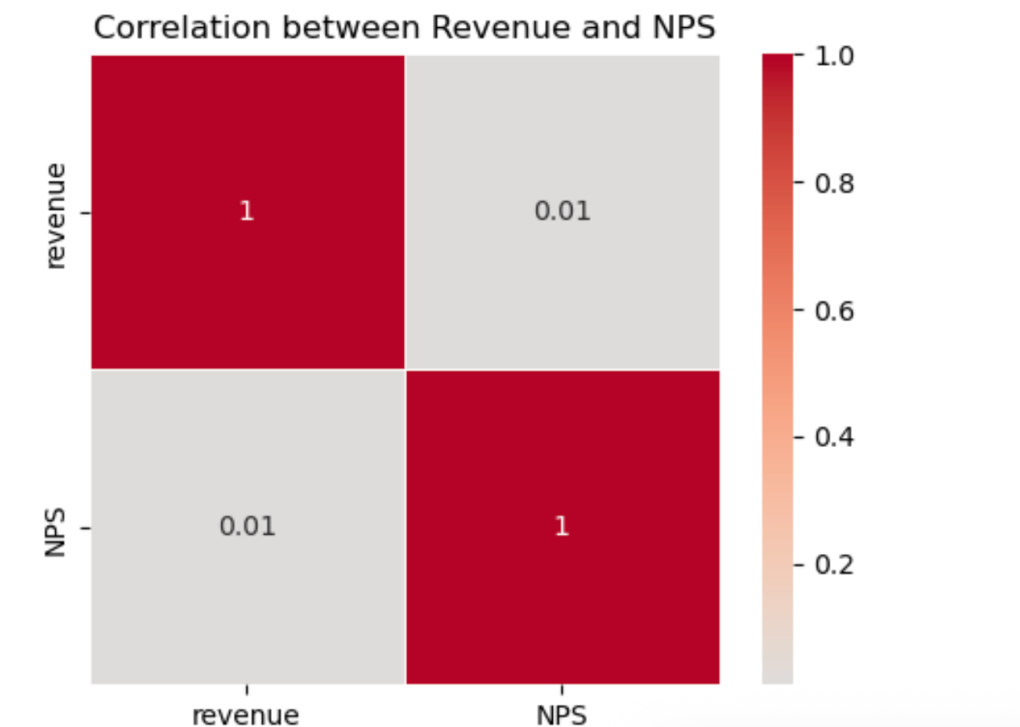


Figure 4: Preliminary Correlation Heatmap between Revenue and NPS score..

## Scatter Plot

```
# Add scatter_nps_vs_revenue between NPS and revenue
def scatter_nps_vs_revenue(df):
    """Create scatter plot of NPS vs revenue"""
    plt.figure(figsize=(5, 4))
    plt.scatter(df['NPS'], df['revenue'], alpha=0.6)
    plt.xlabel('NPS')
    plt.ylabel('revenue')
    plt.title('NPS vs revenue')
    plt.grid(True, alpha=0.3)
    plt.show()

# Call the function
scatter_nps_vs_revenue(df)
```

✓ 0.0s

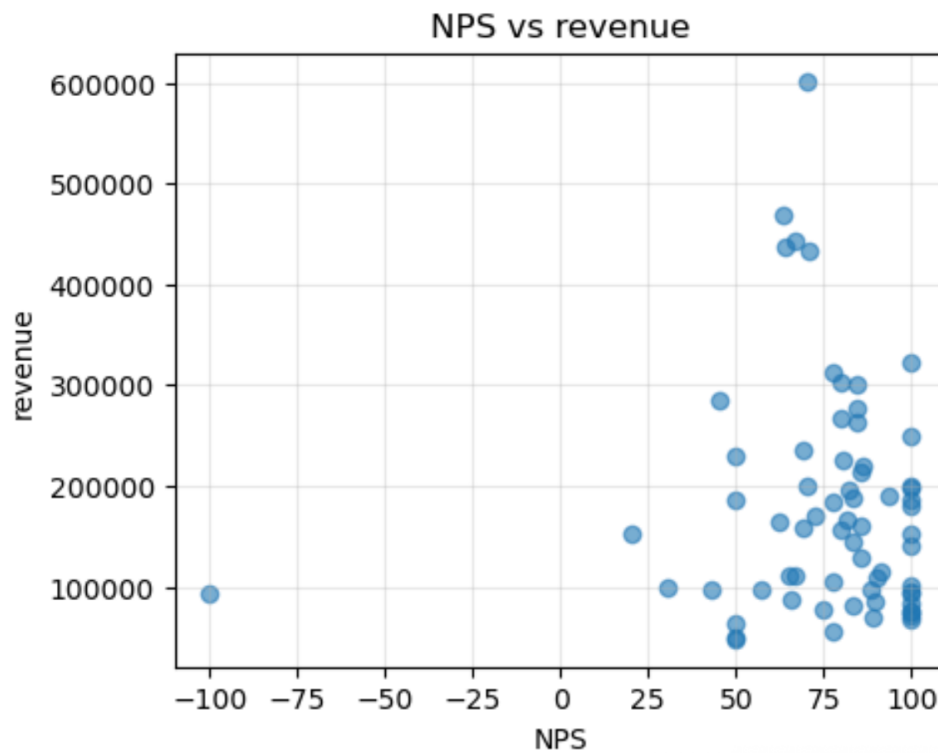


Figure 5: Preliminary Scatter Plot Graphic between Revenue and NPS score.

End of Appendix Section



## **Statement of Acknowledgment**

I acknowledge that I have used the following AI tool(s) in the creation of this report:

- OpenAI ChatGPT (GPT-5): Used to assist with outlining, refining structure, improving clarity of academic language, and supporting APA 7th referencing conventions.

I confirm that the use of the AI tool has been in accordance with the Torrens University Australia Academic Integrity Policy and TUA, Think and MDS's Position Paper on the Use of AI. I confirm that the final output is authored by me and represents my own critical thinking, analysis, and synthesis of sources. I take full responsibility for the final content of this report.

## 11. References

- Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
- Creswell, J. W., & Plano Clark, V. L. (2023). *Designing and conducting mixed methods research* (4th ed.). SAGE Publications.
- Dawes, J. G. (2024). *The net promoter score: What should managers know?* *International Journal of Market Research*, 66(2–3), 182–198.  
<https://doi.org/10.1177/14707853231195003>
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications.
- Godovykh, M., & Pizam, A. (2023). *Measuring patient experience in healthcare*. *International Journal of Hospitality Management*, 112, 103405.  
<https://doi.org/10.1016/j.ijhm.2022.103405>
- Morgan, D. L. (2014). *Pragmatism as a paradigm for social research*. *Qualitative Inquiry*, 20(8), 1045–1053. <https://doi.org/10.1177/1077800413513733>
- Pallant, J. (2020). *SPSS survival manual* (7th ed.). Routledge.
- Polonsky, M. J., & Waller, D. S. (2019). Quantitative data analysis. In *Designing and managing a research project* (4th ed., pp. 222-254). SAGE Publications.  
<https://doi.org/10.4135/9781544316499>
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in software engineering*. Springer.
- Wohlin, C., & Runeson, P. (2021). Guiding the selection of research methodology in industry–academia collaboration in software engineering. *Information and Software Technology*, 140, 106678. <https://doi.org/10.1016/j.infsof.2021.106678>