README — BNPL Adoption Behavioral Analysis

Replication Guide for Code and Dataset

Dataset Description

Dataset Name: Raw-Data_Paylater-and-Non-Paylater-User.xlsx
Source: Primary structured survey
Sample Type: BNPL and Non-BNPL users
Scale Used: Likert scale 1–5 (Strongly Disagree → Strongly Agree)

Key Variable Categories:

Category	Examples
Demographics	Age, Gender, Education, Income
Financial Behaviour	Spending habits, repayment behavior
Technology Adoption	PU – Perceived Usefulness, PEOU – Ease of Use
Psychological Factors	Trust, Perceived Risk, Security
Behavioral Outcome	Adoption Intention + BNPL Actual Usage

Data includes classification into two groups:

- **BNPL Users**

- **Non-BNPL Users**

Variables and Constructs Table

Construct	Indicator Variables	Purpose
PU (Usefulness)	PU1, PU2, PU3	Benefits of BNPL
PEOU (Ease of Use)	PE1, PE2, PE3	Convenience to adopt
Trust	TR1, TR2, TR3	Confidence in service
Perceived Risk	PR1, PR2, PR3	Fear of financial loss
Behavioral Intention	BI1, BI2, BI3	Likelihood of future usage
Actual Usage	Use Frequency Score	Adoption outcome

Software Requirements

Library	Use
pandas, numpy	Data preprocessing
statsmodels	OLS Regression
semopy (or Lavaan/Mplus)	CFA, SEM
scikit-learn	K-Means Clustering
seaborn, matplotlib	Visual analytics

Install:
pip install pandas numpy matplotlib seaborn statsmodels scikit-learn semopy

Replication Instructions (Run Order)

Step	Analysis	Output
1	Import dataset & clean	Final dataframe
2	Reliability → Cronbach α	α ≥ 0.70 acceptable
3	OLS Regression (Predictive Check)	Screening of significant predictors
4	CFA (Measurement Model)	Validity confirmation

5	SEM (Structural Model)	Mediation + causal relationships
6	Multi-Group SEM	BNPL vs Non-Users comparison
7	Predictive Power Test	RMSE, R² scores
8	Clustering	Segmented consumer groups

Methodology (What Each Model Does)

OLS Regression

Used as preliminary check:

$$BI = \beta_0 + \beta_1 \cdot PU + \beta_2 \cdot PEOU + \beta_3 \cdot Trust - \beta_4 \cdot Risk + \varepsilon$$

Purpose → identify strongest behavioral predictors.

CFA — Measurement Validation

Model fit indices checked:

Metric	Acceptance
CFI > 0.90	Good Fit
TLI > 0.90	Construct fit
RMSEA < 0.08	Error margin
SRMR < 0.08	Residuals
AVE > 0.50	Convergent validity
CR > 0.70	Reliability

Multi-Group SEM — Group Differences

Comparing **BNPL Users vs Non-Users**:

Path	Stronger In
PU → BI	BNPL Users

Risk → BI   Non-BNPL Users

Trust → BI   BNPL Users

Interpretation:
1.Users rely more on *benefits*
2. Non-users fear *financial risks*

Predictive Power Evaluation

Held-out dataset scoring:

- Good generalization

- Behaviors predict usage correctly

Metrics reported: R^2, MAE, RMSE

Clustering Results (Behavioral Segmentation)

K-Means → 3 or 4 clusters typical
Segments identified:

Cluster	Profile	Interpretation
1	Tech-savvy & Trusting	Strong adoption drivers
2	Risk-averse & skeptical	Requires assurance
3	Neutral / Price-sensitive	Motivated by discounts

Marketing Insight:
Personalized targeting improves BNPL conversion.

Outputs Generated from Code

Distribution plots
OLS summary tables
CFA factor loadings
SEM path diagrams
Group-wise significance comparison

Cluster scatterplots & centroid plots
Heatmaps and correlation matrices

All results visible directly in the notebook.

Reproducibility Requirements

Do not change dataset filename
Maintain data types of Likert responses
Full execution without skipping cells

Key Findings Summary

- **Trust**, **Usefulness**, and **Ease of Use** → drive adoption

- **Risk concerns** deter non-users

- Clustering → distinct **behavioral personas** for targeted marketing

- Data strongly supports BNPL expansion strategy in young digital users