Exploratory Data Analysis Report

**Dataset:** Initial Survey.xlsx **Generated on:** 2025-10-01 15:51:54 **Source of data:** Uploaded Excel file (local dataset). **Methodology:** Automated exploratory analysis was performed across all sheets. Variable types were inferred using data types and cardinality heuristics; numeric variables were summarized with distributional statistics and bivariate correlations; categorical variables were summarized with frequency tables; datetime-like variables were aggregated by day. Missingness was quantified per variable. Visualizations were produced with matplotlib.

# 1. Dataset Overview

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sheet | Rows | Columns | Numeric columns | Categorical columns | Datetime-like columns | Missing cells (%) |
| in | 32 | 15 | 2 | 8 | 0 | 16.88 |

# 2. in

**Dimensions:** 32 rows × 15 columns.

**Inferred variable types.** Numeric: 2; Categorical: 8; Datetime-like: 0.

## in: Missingness by variable

|  |  |
| --- | --- |
| Variable | Missing\_Rate(%) |
| Why do you not use AI for studying? | 93.75 |
| Will you use AI if it solves the above stated problems? | 93.75 |
| What problems do you face while using these tools? (eg, inaccurate answers, slow, etc) | 18.75 |
| Why do you use AI for studying? | 15.62 |
| Which AI tools do you use for studying? | 6.25 |
| How often do you use AI in your studying? | 6.25 |
| In what ways do you use AI for studying? | 6.25 |
| On a scale of 1-5, how effective do you find AI compared to traditional methods like lecture notes? | 6.25 |
| Has using AI improved your study outcomes? | 6.25 |
| Timestamp | 0.0 |
| What is your current level of education? | 0.0 |
| What is your age group? | 0.0 |
| What is your primary major or field of study? | 0.0 |
| On a scale of 1-5, how would you rate your overall academic performance? | 0.0 |
| Do you use AI tools for studying? | 0.0 |

## in: Descriptive statistics (numeric)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | count | mean | std | min | 25% | median | 75% | max | missing\_% |
| On a scale of 1-5, how would you rate your overall academic performance? | 32.0 | 3.59375 | 0.7560241931612779 | 1.0 | 3.0 | 4.0 | 4.0 | 5.0 | 0.0 |
| On a scale of 1-5, how effective do you find AI compared to traditional methods like lecture notes? | 30.0 | 3.9 | 0.8448627719625691 | 2.0 | 3.0 | 4.0 | 4.75 | 5.0 | 6.25 |

## in: Correlation matrix (numeric)

|  |  |  |
| --- | --- | --- |
| Variable | On a scale of 1-5, how would you rate your overall academic performance? | On a scale of 1-5, how effective do you find AI compared to traditional methods like lecture notes? |
| On a scale of 1-5, how would you rate your overall academic performance? | 1.0 | 0.07533363093289291 |
| On a scale of 1-5, how effective do you find AI compared to traditional methods like lecture notes? | 0.07533363093289291 | 1.0 |

## in: Frequency – Do you use AI tools for studying? (top 10)

|  |  |
| --- | --- |
| Do you use AI tools for studying? | Count |
| Yes | 19 |
| Sometimes | 11 |
| No | 2 |

## in: Frequency – Why do you not use AI for studying? (top 10)

|  |  |
| --- | --- |
| Why do you not use AI for studying? | Count |
| nan | 30 |
| Unreliable accuracy;Prefer traditional methods;Ethical concerns | 1 |
| Prefer traditional methods | 1 |

## in: Frequency – Will you use AI if it solves the above stated problems? (top 10)

|  |  |
| --- | --- |
| Will you use AI if it solves the above stated problems? | Count |
| nan | 30 |
| Yes | 2 |

## in: Frequency – What is your age group? (top 10)

|  |  |
| --- | --- |
| What is your age group? | Count |
| 18-24 | 23 |
| 25-34 | 8 |
| 35+ | 1 |

## in: Frequency – What is your current level of education? (top 10)

|  |  |
| --- | --- |
| What is your current level of education? | Count |
| Post Graduate | 12 |
| Undergraduate | 11 |
| Graduate | 9 |

## in: Frequency – How often do you use AI in your studying? (top 10)

|  |  |
| --- | --- |
| How often do you use AI in your studying? | Count |
| Weekly | 15 |
| Daily | 7 |
| Rarely | 6 |
| nan | 2 |
| Monthly | 2 |

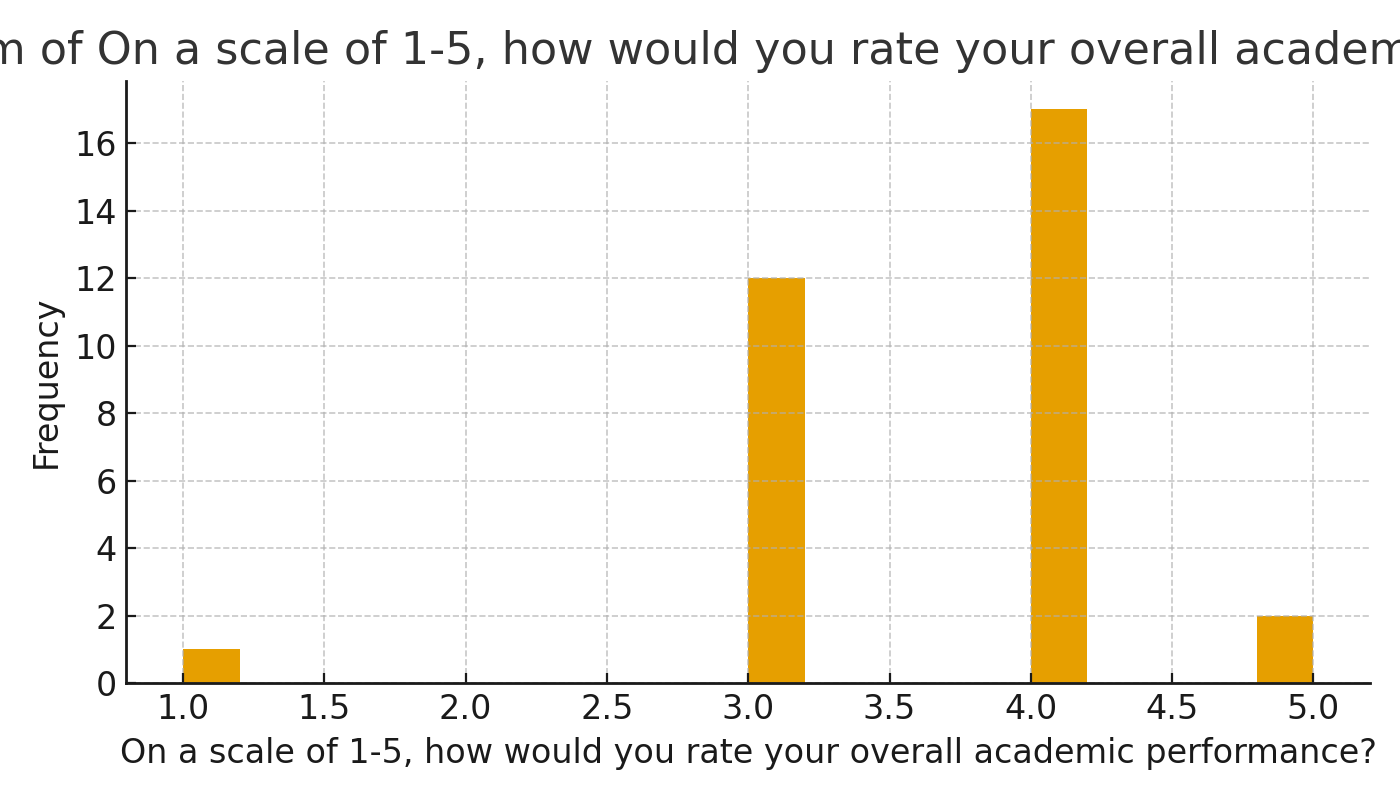
## in: Frequency – Has using AI improved your study outcomes? (top 10)

|  |  |
| --- | --- |
| Has using AI improved your study outcomes? | Count |
| Yes somewhat | 18 |
| Yes significantly | 12 |
| nan | 2 |

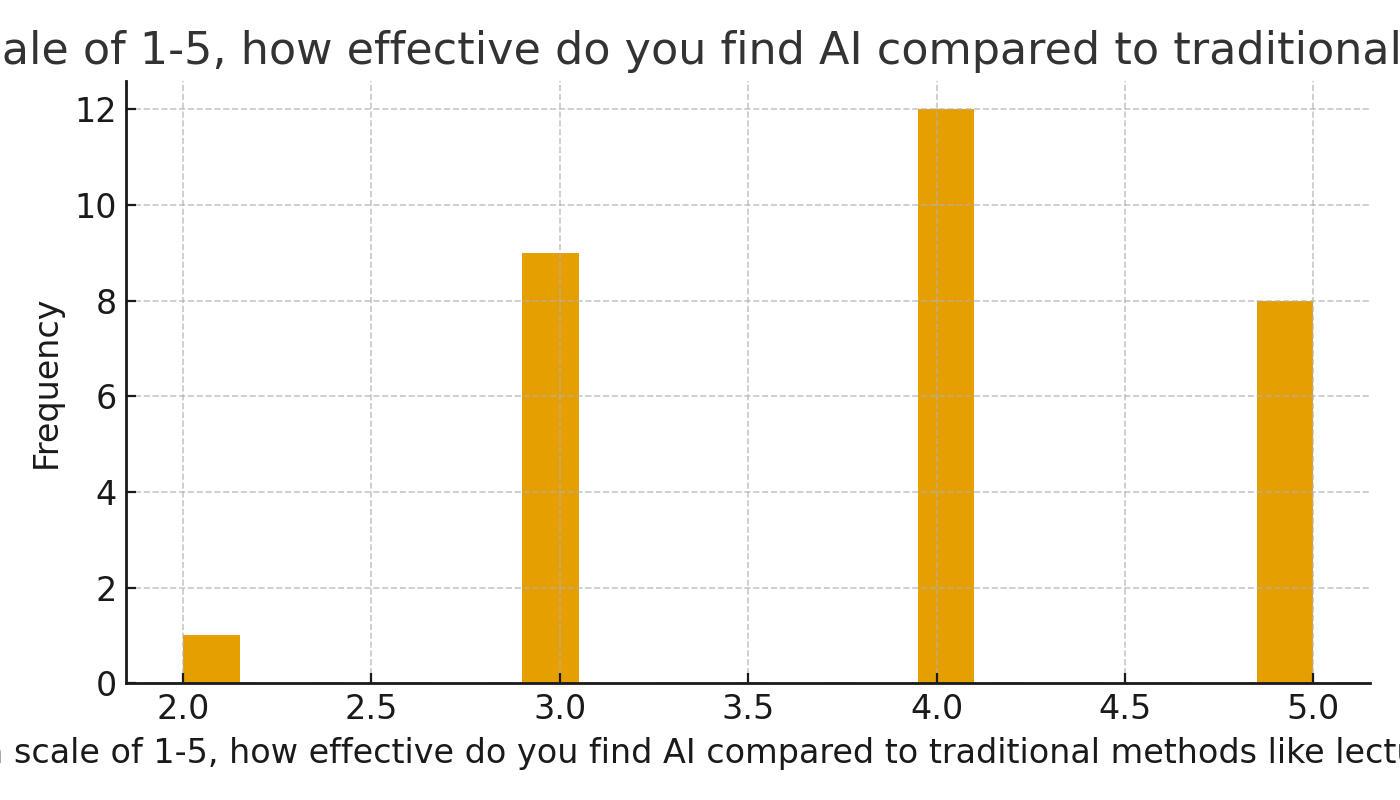
## in: Frequency – Which AI tools do you use for studying? (top 10)

|  |  |
| --- | --- |
| Which AI tools do you use for studying? | Count |
| ChatGPt | 17 |
| ChatGPt;Gemini | 7 |
| nan | 2 |
| ChatGPt;Grok;Adobe AI | 1 |
| ChatGPt; | 1 |
| ChatGPt;Notion AI | 1 |
| ChatGPt;Claude | 1 |
| Snapchat Ai | 1 |
| ChatGPt;Gemini;Claude, Deepseek | 1 |

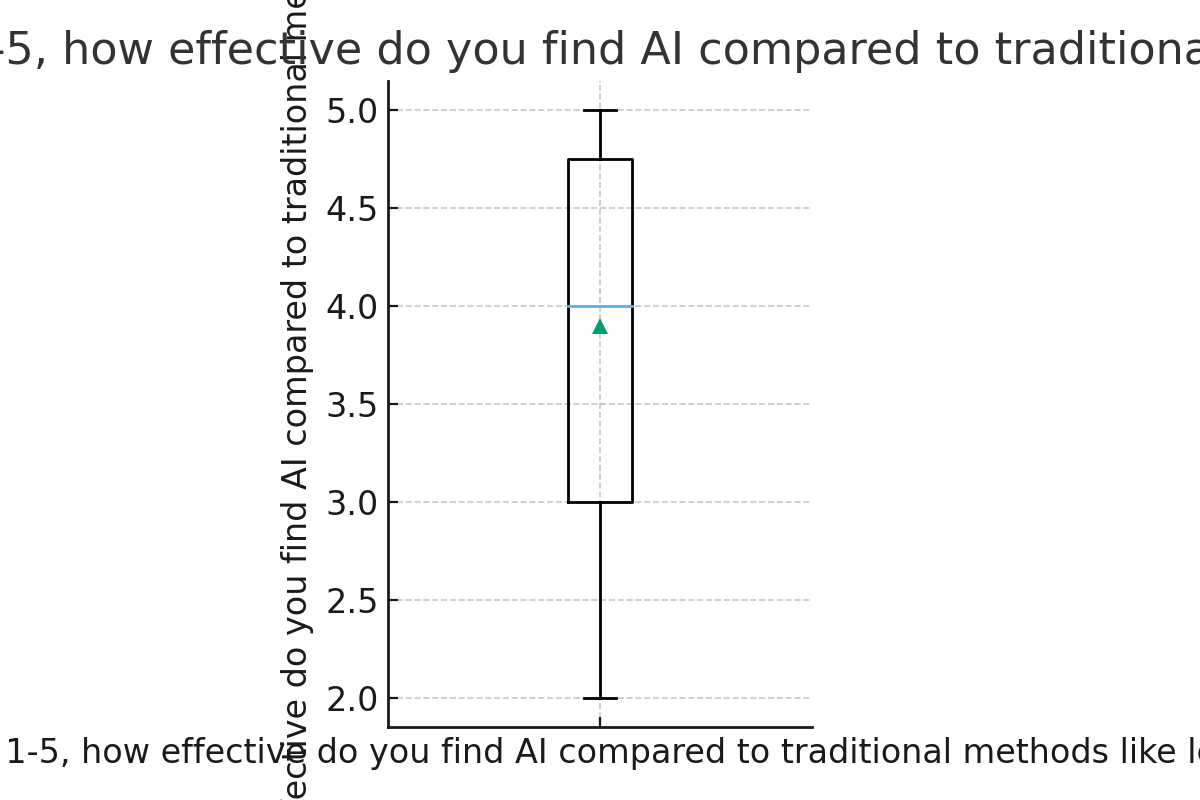
*Figure: Histogram of On a scale of 1-5, how would you rate your overall academic performance? (in)*



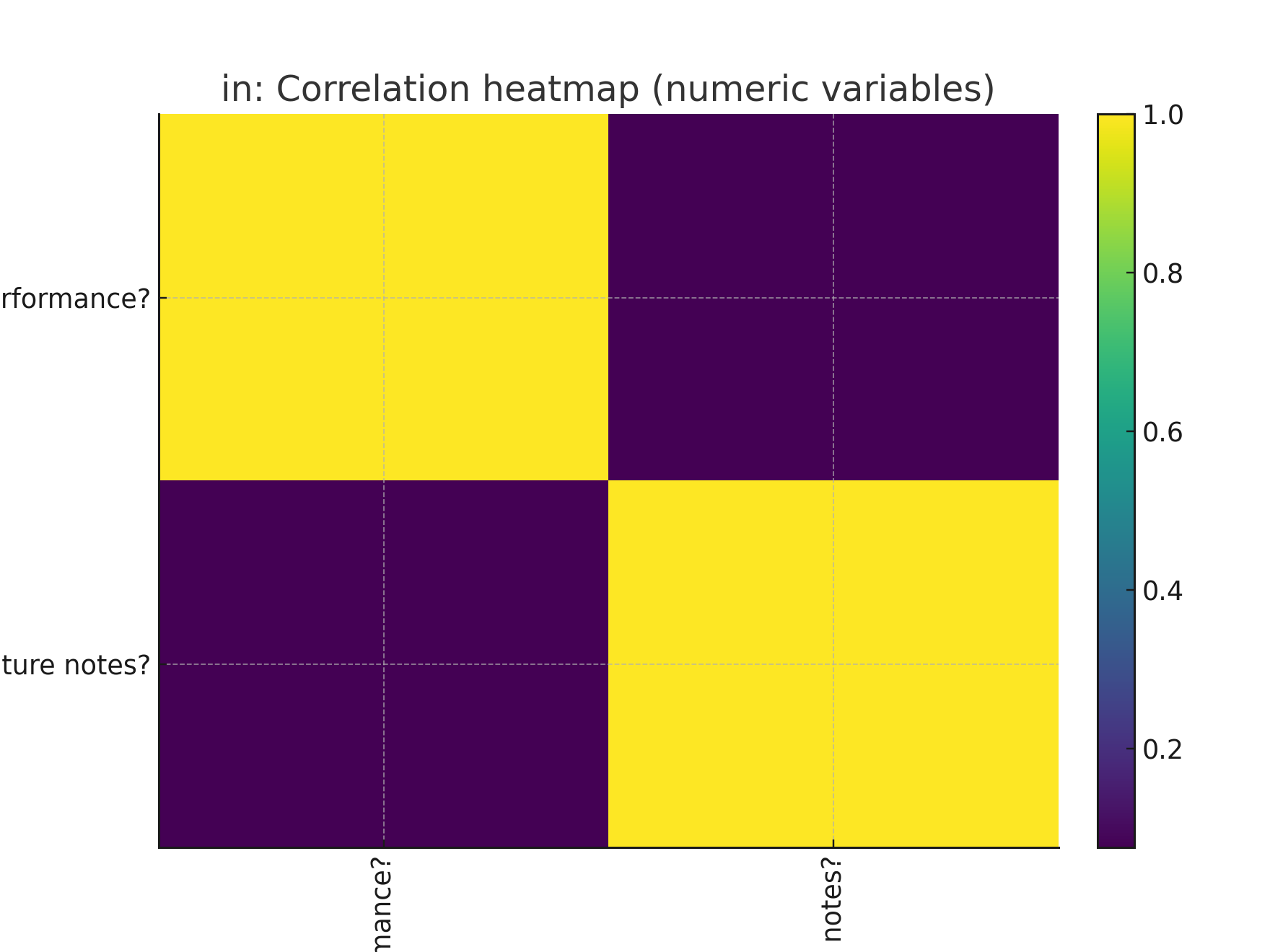
*Figure: Histogram of On a scale of 1-5, how effective do you find AI compared to traditional methods like lecture notes? (in)*



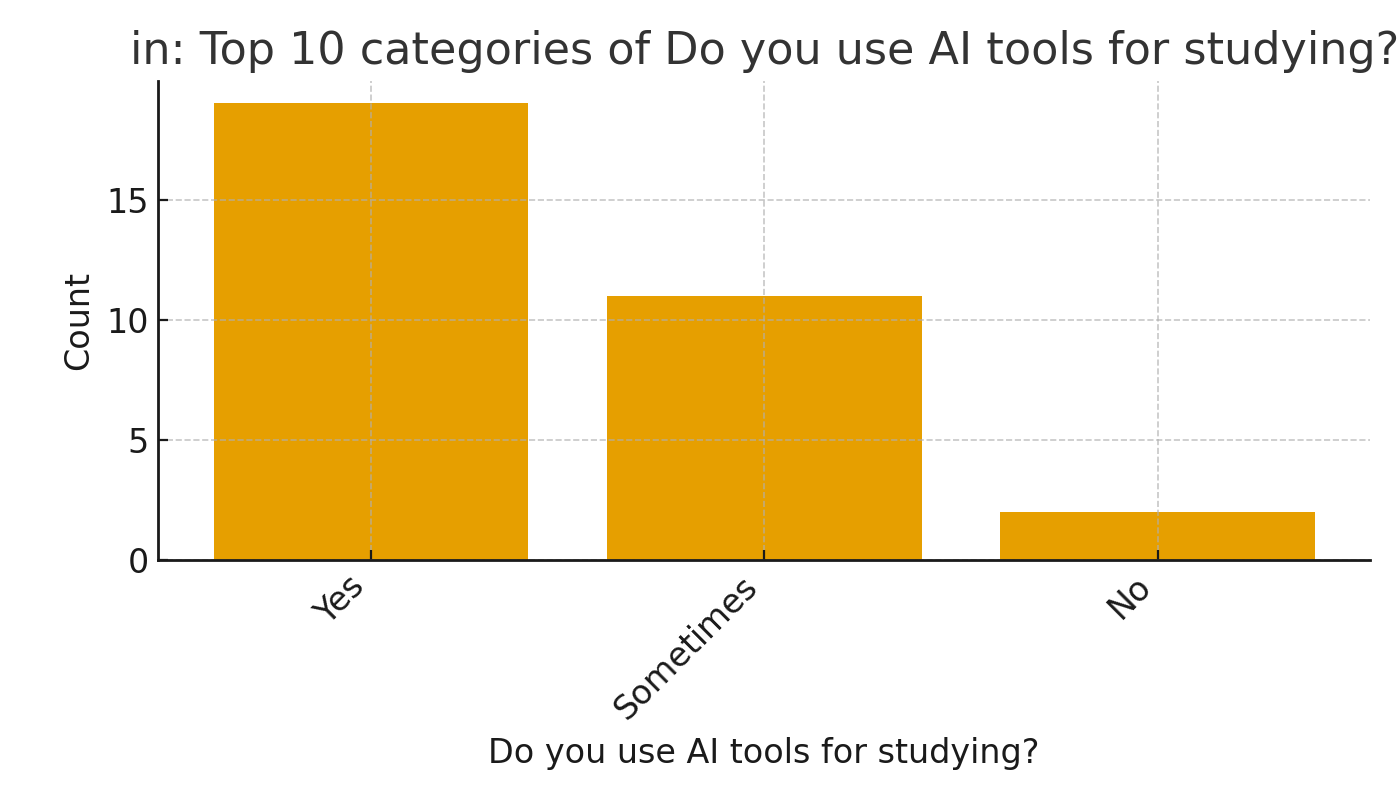
*Figure: Boxplot of on a scale of 1-5, how effective do you find AI compared to traditional methods like lecture notes? (in)*



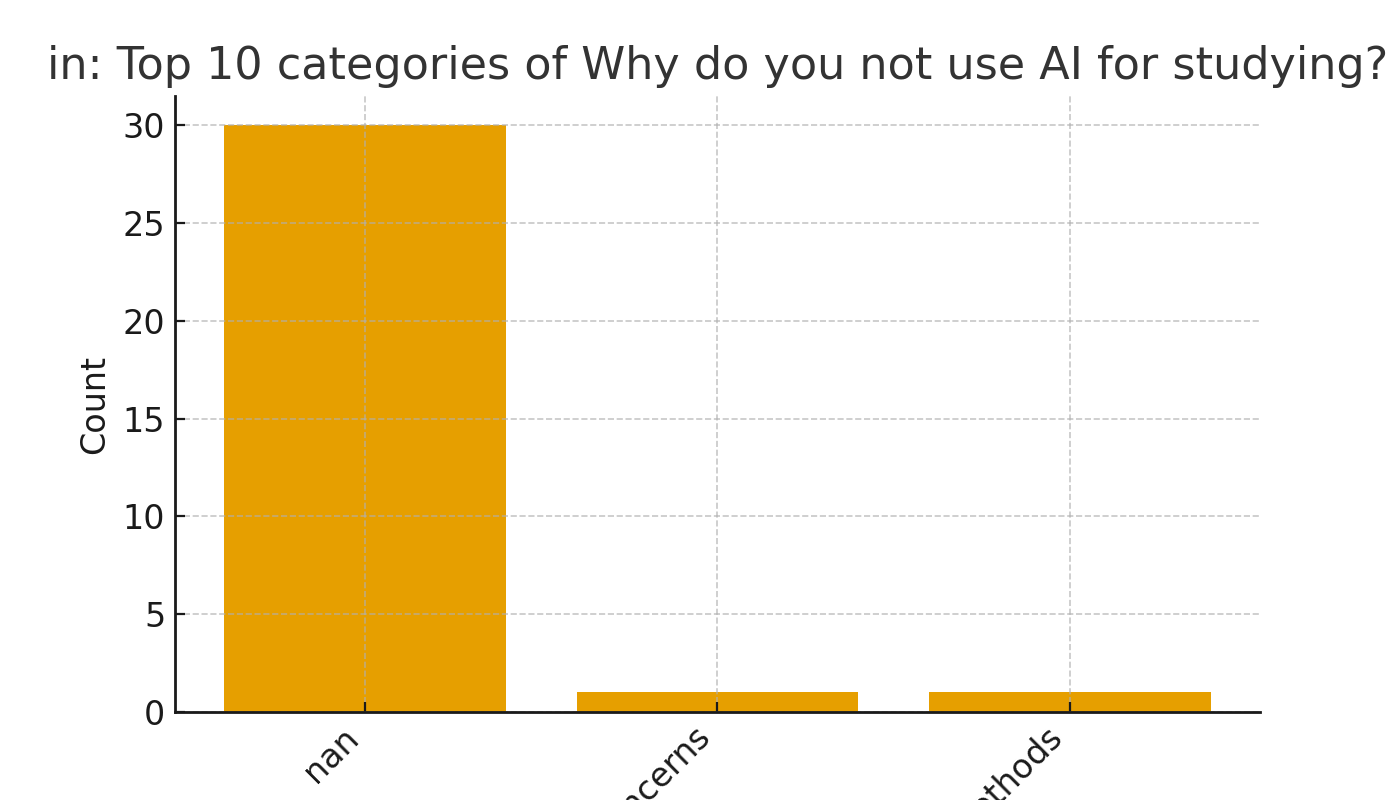
*Figure: Correlation heatmap (in)*



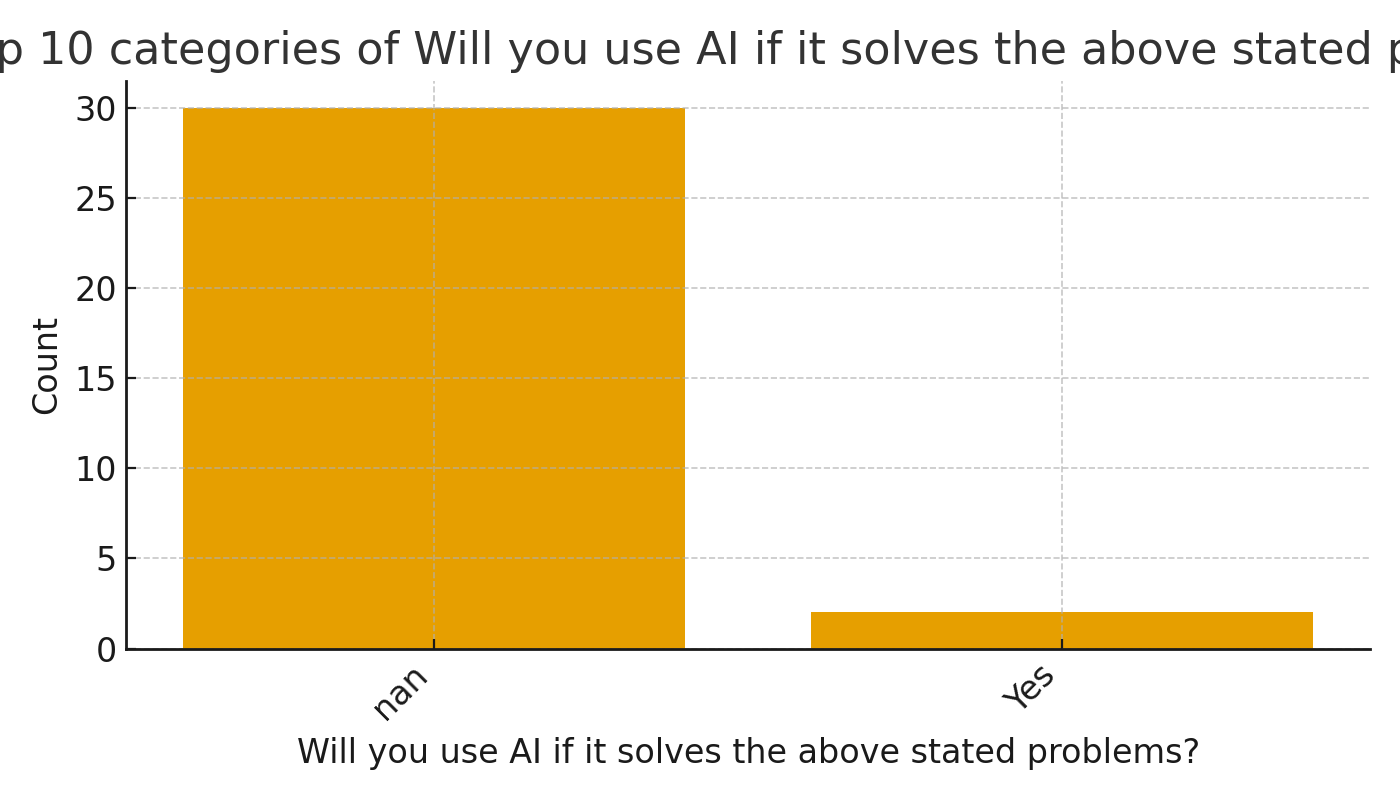
*Figure: Top 10 categories – Do you use AI tools for studying? (in)*



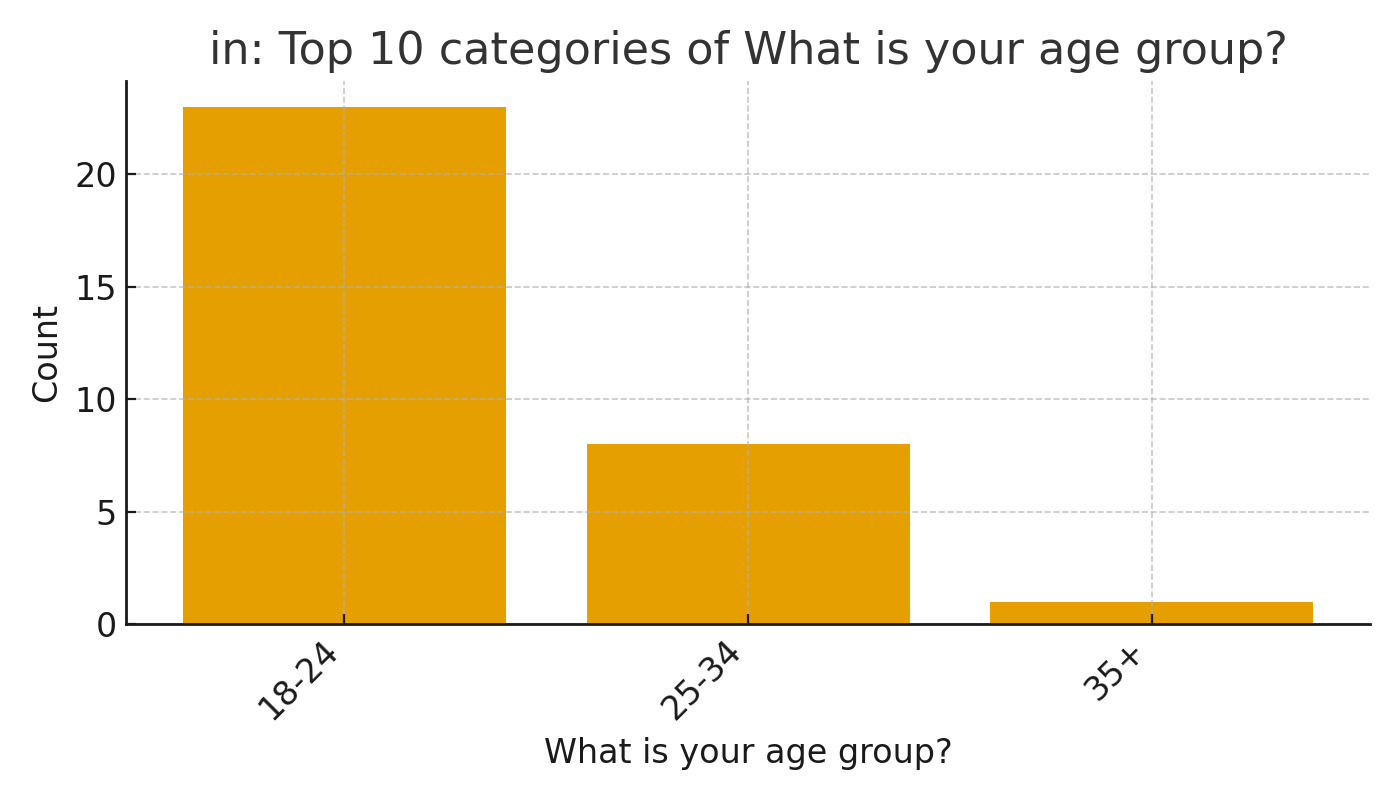
*Figure: Top 10 categories – Why do you not use AI for studying? (in)*



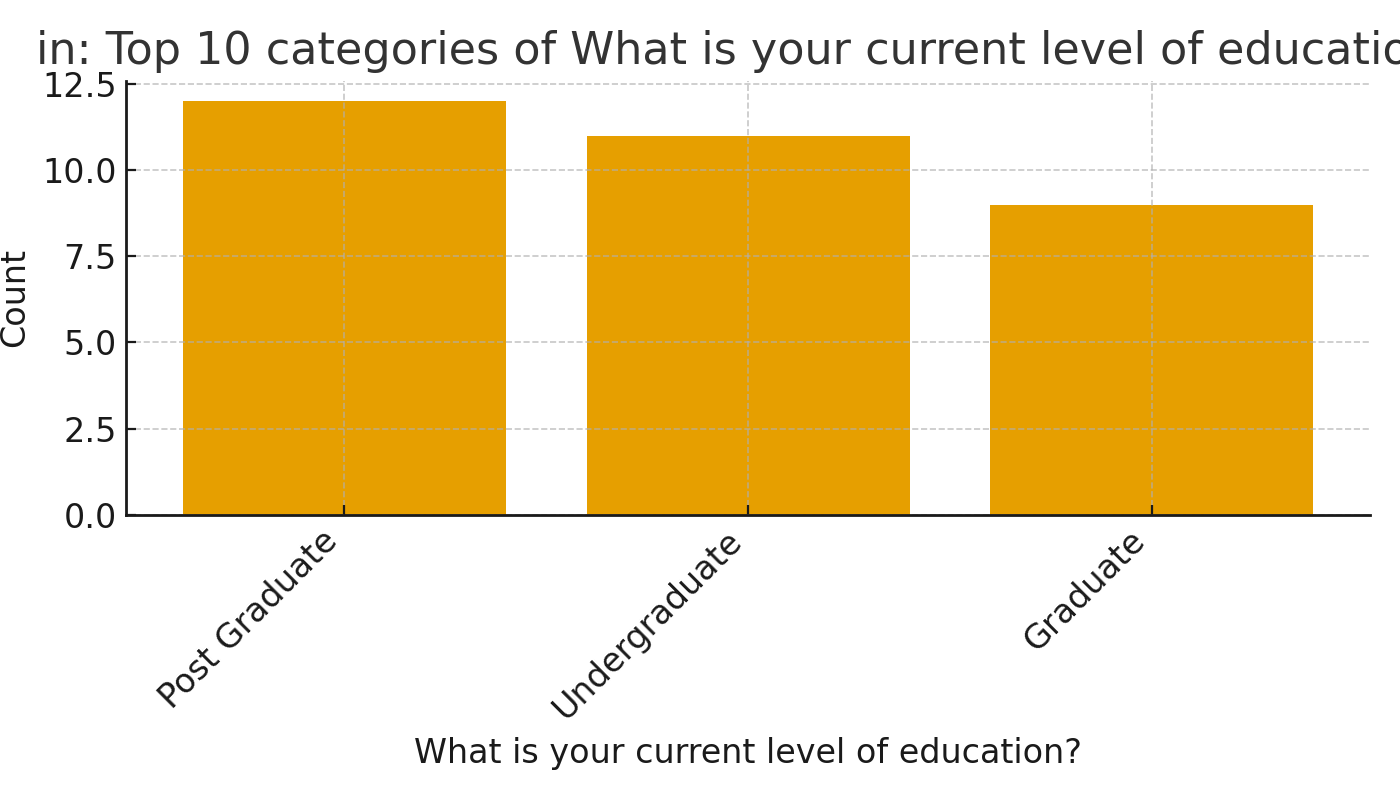
*Figure: Top 10 categories – Will you use AI if it solves the above stated problems? (in)*



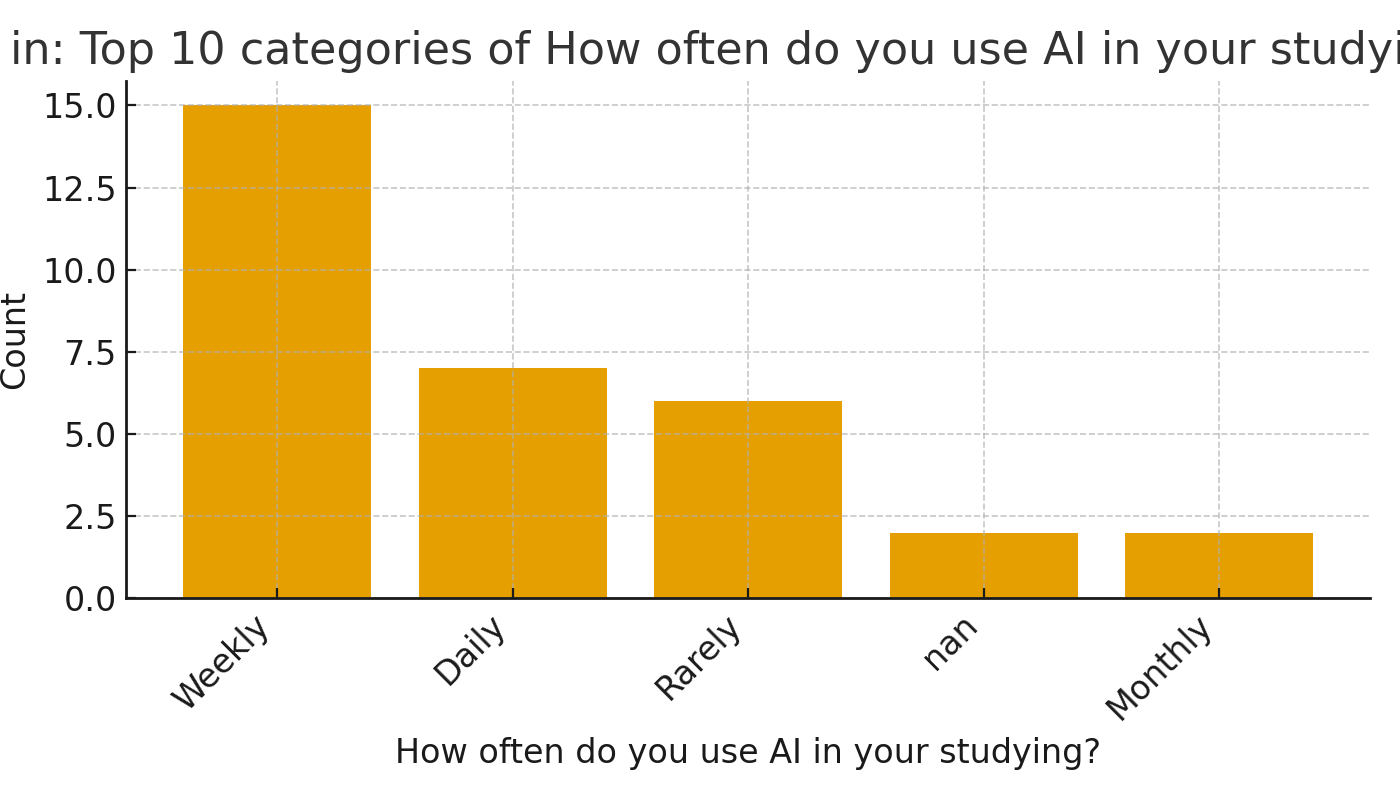
*Figure: Top 10 categories – What is your age group? (in)*



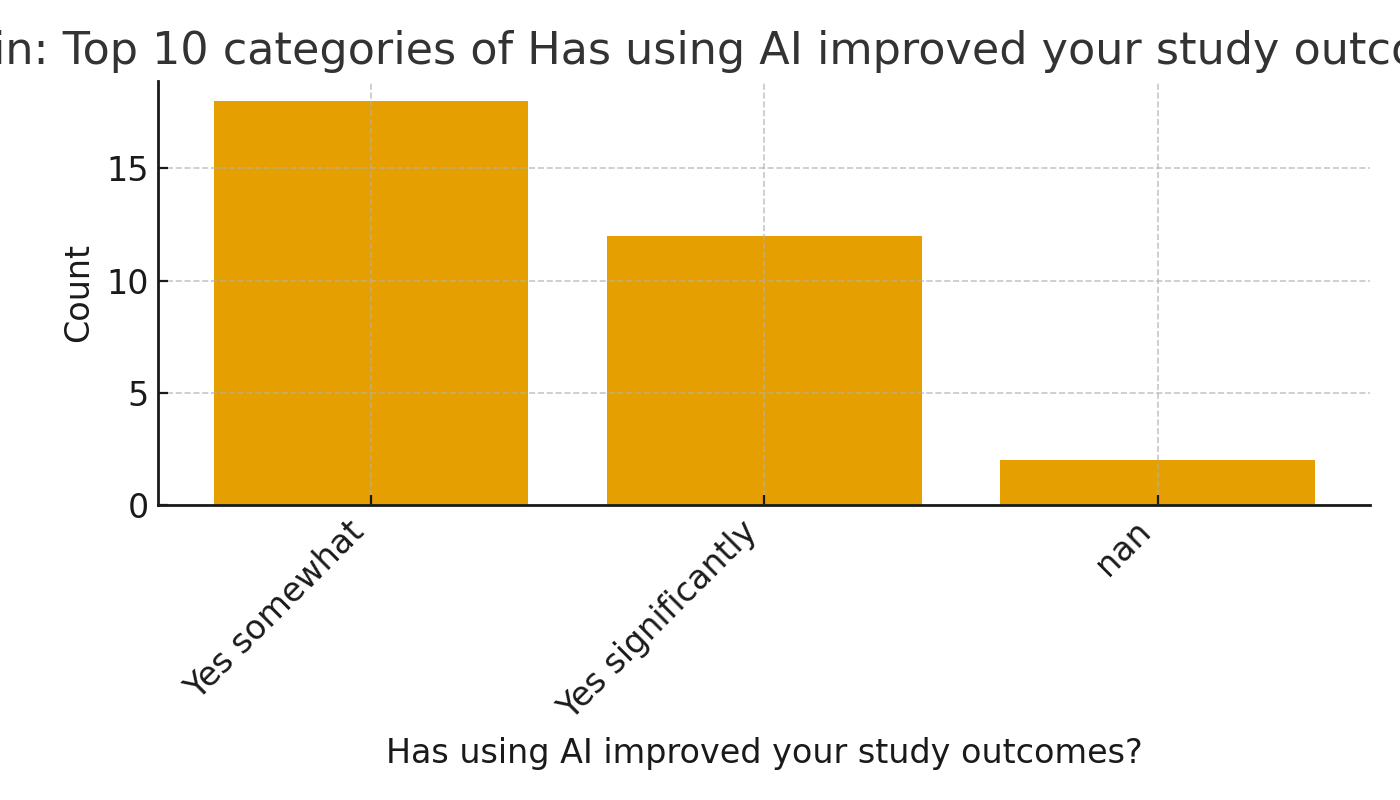
*Figure: Top 10 categories – What is your current level of education? (in)*



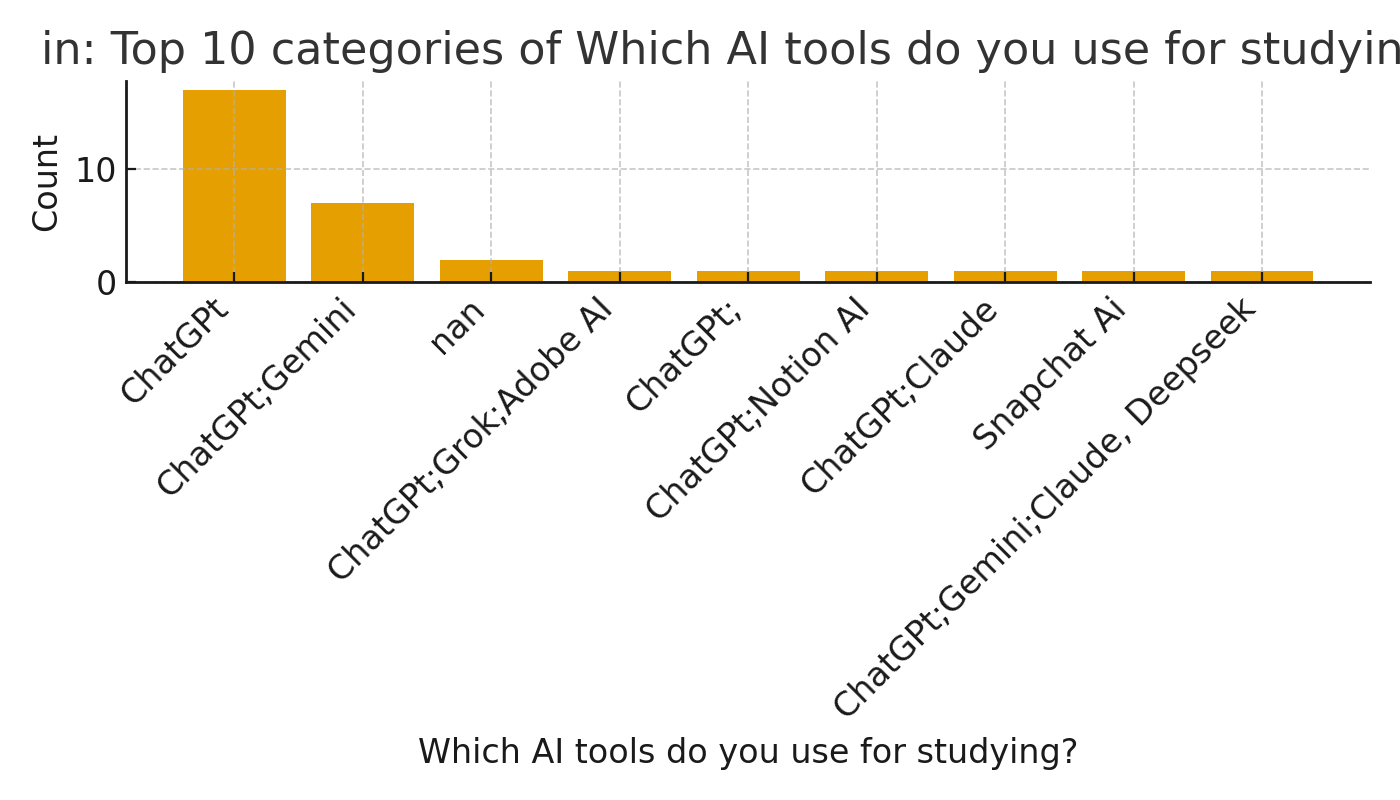
*Figure: Top 10 categories – How often do you use AI in your studying? (in)*



*Figure: Top 10 categories – Has using AI improved your study outcomes? (in)*



*Figure: Top 10 categories – Which AI tools do you use for studying? (in)*



**Interpretation Notes:** Distributions were examined for skewness and outliers using histograms and boxplots. Where applicable, bivariate correlation structure was inspected to identify potential multicollinearity. Categorical distributions reveal the concentration of responses across levels; temporal plots indicate seasonality or structural breaks if present. Observed patterns should be interpreted in light of the data-generating process and any sampling limitations.

# 3. Limitations

This report is fully descriptive and relies solely on the uploaded dataset without external validation. Automated type inference may misclassify variables with atypical encodings. Missing data mechanisms were not modelled; figures and tables reflect available observations only. No causal claims are made.

# 4. Reproducibility

All computations were executed within a controlled environment using Python (pandas, numpy, matplotlib, python-docx). Charts were generated with default matplotlib settings, one figure per chart. Analytical steps can be replicated by re-running this notebook block against the same input file.