[EN] AI Assistant Usage in Student Life — Analytical Report

Author: Ekaterina Suvalova — Junior Data Analyst

Date: August 2025

Tools: Excel, Power BI (Power Query, DAX), GitHub, Notion

1) Executive Summary

- This report covers **10,000 sessions** of AI Assistant (hereafter — AI) usage by **students** over the course of **one year**. The main audience consisted of **undergraduates — 60%**, **high school students** and **graduate students — 20% each.**
- In **48%** of sessions, students **successfully completed** the assignment; in **29%**, they **drafted an idea**; in **16%**, they got **confused**; and in only **8%**, they completely **gave up**.
- **71%** of students **reused** the AI tool after their first session.
- The highest **efficiency** was observed in sessions lasting **5–10 minutes**; low-efficiency sessions involved **2–3× more prompts**.
- Peaks of **activity** occurred on **Fridays** and **Sundays**, with the busiest months being **August**, **June**, and **January**.
- The most common **Task Types** were **Writing** (3.1K) and **Studying/Homework Help** (2.0K). **Coding** showed the highest share of **successfully completed sessions — 60%**.

2) Project Objective

Objective: To demonstrate a complete analytical process using an educational dataset: from data preparation and metric calculation to visualization and interpretation of results. The project showcases proficiency in Excel, Power BI (Power Query, DAX), as well as the ability to draw meaningful conclusions and articulate the limitations of analysis when working with synthetic data.

Key Questions:

1. Which groups of students use AI more actively?
2. How are sessions distributed across disciplines and task types?
3. How are sessions distributed across efficiency levels, and how do duration and number of prompts influence efficiency?
4. In which months and days of the week do students use AI most frequently?

5. What are the final session outcomes by day of the week and by task type?
6. How consistently do students return to using AI?

---

3) Dataset Overview

- **Source: AI Assistant Usage in Student Life** (Kaggle).
- **Volume and period: 10,000 sessions** recorded over a **full calendar year** (June 24, 2024 – June 24, 2025).
- **Purpose:** This dataset was created **to simulate realistic student interactions with AI tools** (such as ChatGPT) for academic purposes. It is fully **synthetic** but designed to reflect real-world logic and behavior, making it suitable for exploratory data analysis, machine learning experiments, feature engineering, and visualization practice.
- **Main Fields:**

StudentLevel	High School, Undergraduate, Graduate
Discipline	e.g., CS, Psychology, etc.
SessionDate	session timestamp
SessionLengthMin	session duration in minutes
TotalPrompts	number of prompts in the session
TaskType	e.g., Writing, Coding, Research
AI_AssistanceLevel	1–5 rating of helpfulness
FinalOutcome	Assignment Completed, Idea Drafted, Confused, Gave Up
UsedAgain	UsedAgain
SatisfactionRating	1–5 satisfaction score

- **Additional Derived Metrics:**

Prompt density 10min	frequency of prompts per 10 minutes of a session	Formula: Total prompts / (Session length min / 10)
Efficiency score	efficiency index of AI usage: the ratio of AI Assistance Level to Prompt density **10min**. The fewer prompts are needed for a high assistance level, the higher the efficiency	Formula: AI Assistance Level / Prompt density 10min

---

4) Tools Used

Excel:

- initial data cleaning
- data type validation, date/time formatting, anomaly detection
- calculation of derived metrics

- creation of pivot tables

Power BI:

- data analysis and visualization
- creating measures and calculated columns in DAX, adding custom columns in Power Query
- development and design of interactive dashboards with tooltips

GitHub/Notion: project publication and presentation (RU/EN)

---

5) Methodology

1. **Data import and initial cleaning in Excel**: removal of duplicates, handling of missing values, standardizing date and time formats.
2. **Structure validation**: analysis of distributions, identification of logical relationships between fields.
3. **Data preparation in Power Query**: formatting and unit adjustments, renaming fields for clarity, creation of conditional columns, building sorting keys, grouping metrics into categories (binning Efficiency score, SessionLengthMin, Prompt density 10min).
4. **Creation of DAX measures**: calculation of additional metrics (DayOfWeek, High Efficiency Share %, Outcome %, etc.; see section 6).
5. **Data analysis and interpretation**: identification of patterns and key differences by session duration, number of prompts, task types, efficiency scores, etc.
6. **Dashboard design and development**: creating informative visualizations, adding tooltips with concise insights, setting up a consistent color palette and design style.
7. **Formulation of conclusions**: summarizing analysis results and visualizations, preparing a concise written summary for project presentation.
8. **Preparation of project presentation:** compiling visuals (screenshots, GIFs) and structuring content in Notion for clear and engaging project delivery to audience.

---

6) DAX: Calculated Columns and Measures

Alongside the metrics calculated in Excel and conditional columns in Power Query, I created additional measures and calculated columns in DAX.

DayOfWeek

Calculated column converting the **session date into the day-of-week name** (Monday, Tuesday, etc.).
Used in the chart Number of Sessions by Day of Week on the Time & Behavior dashboard.

```
DayOfWeek = FORMAT([Session date], "dddd")
```

DayOfWeekNum

Auxiliary column assigning a sequential number (1–7) to each day of the week.
Used as a **sorting key for the DayOfWeek** to ensure calendar order.

```
DayOfWeekNum = WEEKDAY([Session date], 2)
```

High Efficiency Share %

Measure that calculates **the share of high-efficiency sessions among all sessions.**
Used in a KPI card on the Efficiency & Prompt Analysis dashboard.

```
High Efficiency Share % =
DIVIDE(
    CALCULATE(DISTINCTCOUNT('ai_assistant_usage_student_life'[Sessions]), 'ai_assi
stant_usage_student_life'[EfficiencyScoreGroup] = "High (> 2.2)"),
    CALCULATE(DISTINCTCOUNT('ai_assistant_usage_student_life'[Sessions]), REMOV
EFILTERS('ai_assistant_usage_student_life'[EfficiencyScoreGroup]))
)
```

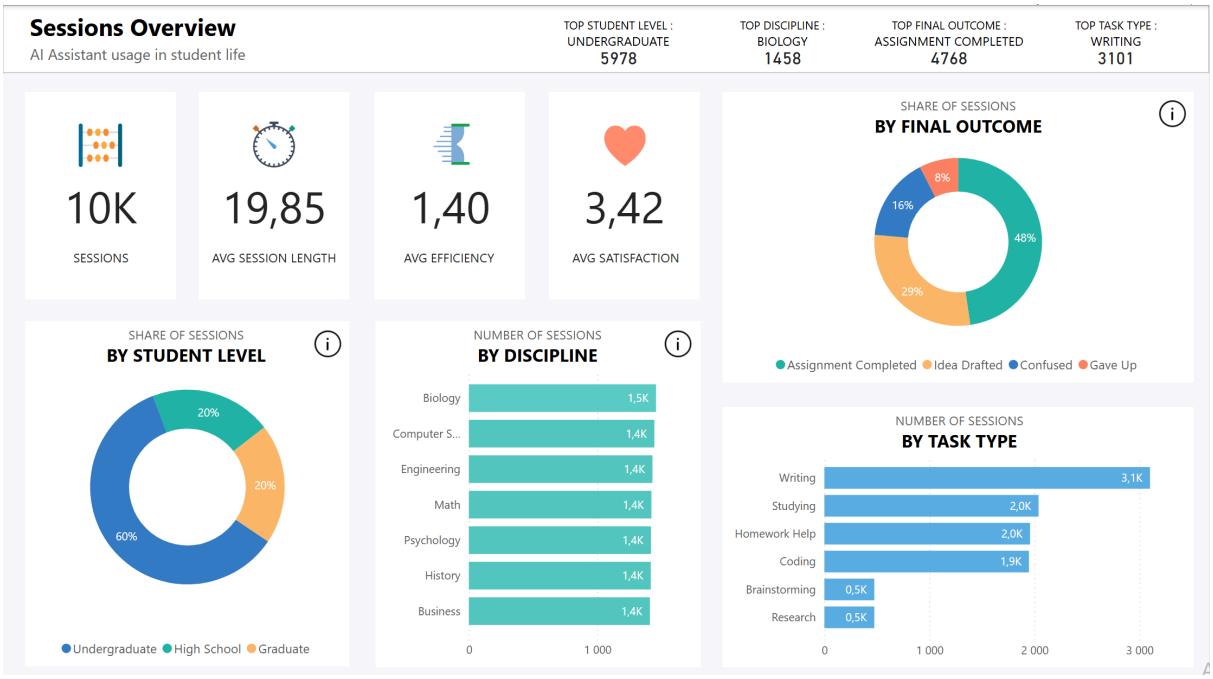
Outcome %

Measure that calculates the **percentage distribution of outcomes** within each Task Type.
Used in the chart Share of Tasks by Final Outcome on the Outcome Analysis dashboard.

```
Outcome % =
VAR Numerator =
    [Sessions (distinct)]
VAR Denominator =
    CALCULATE (
        [Sessions (distinct)],
        ALLEXCEPT (
            'ai_assistant_usage_student_life',
            'ai_assistant_usage_student_life'[Task type]
        )
    )
RETURN
    DIVIDE ( Numerator, Denominator )
```

7) Dashboard Insights

7.1 Overview



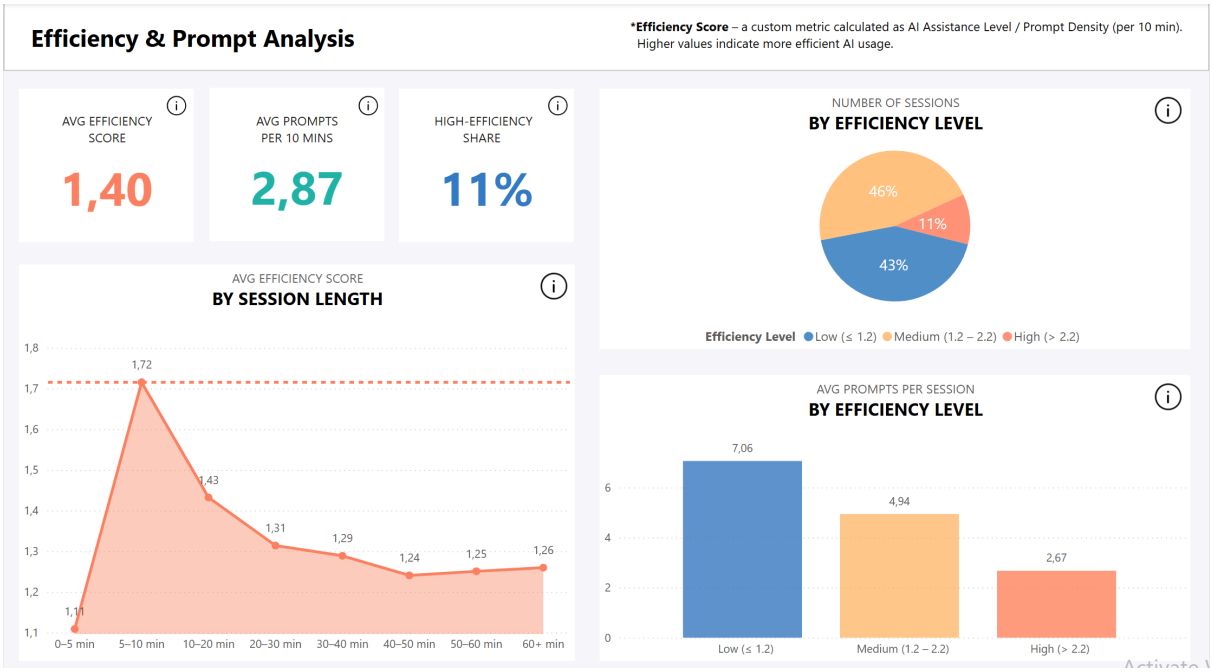
- **Share of Sessions by Final Outcome**
  Almost half of all sessions (48%) **result in successfully completing an assignment**, indicating high efficiency of AI usage. Another **29%** of users employ the system for **drafting ideas**, turning it into a source of inspiration. Together, these two categories account for more than three-quarters of all outcomes, demonstrating the significant educational potential of AI. At the same time, in **16%** of cases students feel **confused** — primarily in tasks related to **Studying** and **Homework Help**. This may point to AI limitations: it is not always able to clarify complex concepts, or its explanations may not be clear enough for learners. Only **8%** of users **gave up**, which highlights the rarity of complete disengagement and shows that most students remain engaged — especially considering that 71% of students return to the AI after their first session.
- **Share of Sessions by Student Level**
  Undergraduates account for **60%** of all activity, making them the primary group of AI users. **High school** and **graduate** students each contribute **20%**, while using AI noticeably less often. This may indicate that high school students currently face fewer tasks where AI feels relevant, while graduate students tend to rely more on their own research skills. **Undergraduates are therefore the most active users**, though AI shows meaningful potential among the other groups as well.
- **Number of Sessions by Discipline**
  Biology slightly leads other disciplines in session count (≈1.5k), but overall the **distribution remains fairly balanced**: Computer Science, Engineering, Maths, Psychology, History, and Business show substantial activity as well (≈1.4k each). This shows that **AI is being applied across a wide range of disciplines** — from the humanities

to the sciences and applied fields. Such balance highlights its versatility and value for students from diverse academic backgrounds.

- **Number of Sessions by Task Type**
  Writing tasks clearly dominate — with over 3.1k sessions, they account for nearly **one-third of all activity**. Which is expected, since working with texts is a central part of academic life. Next come **Studying** and **Homework Help** (≈2.0k each), as well as **Coding** (1.9k), underscoring the demand for AI both in preparation and in hands-on assignments. **Brainstorming** and **Research are used far less often** (≈0.5k each), suggesting that students view AI more as a **tool for tackling specific tasks than as a resource for idea generation or conducting research.**

What this means:
AI is gradually becoming an integral part of the learning process, especially **popular among undergraduates and in writing** tasks. It demonstrates both **strong efficiency and broad applicability**, being used across humanities, sciences, and applied disciplines. While some students (16%) end up Confused, the **low share of students who Gave Up** entirely (8%) and the **high level of repeat usage** indicate that **AI is perceived as a reliable and valuable study tool.**

7.2 Efficiency & Prompt Analysis



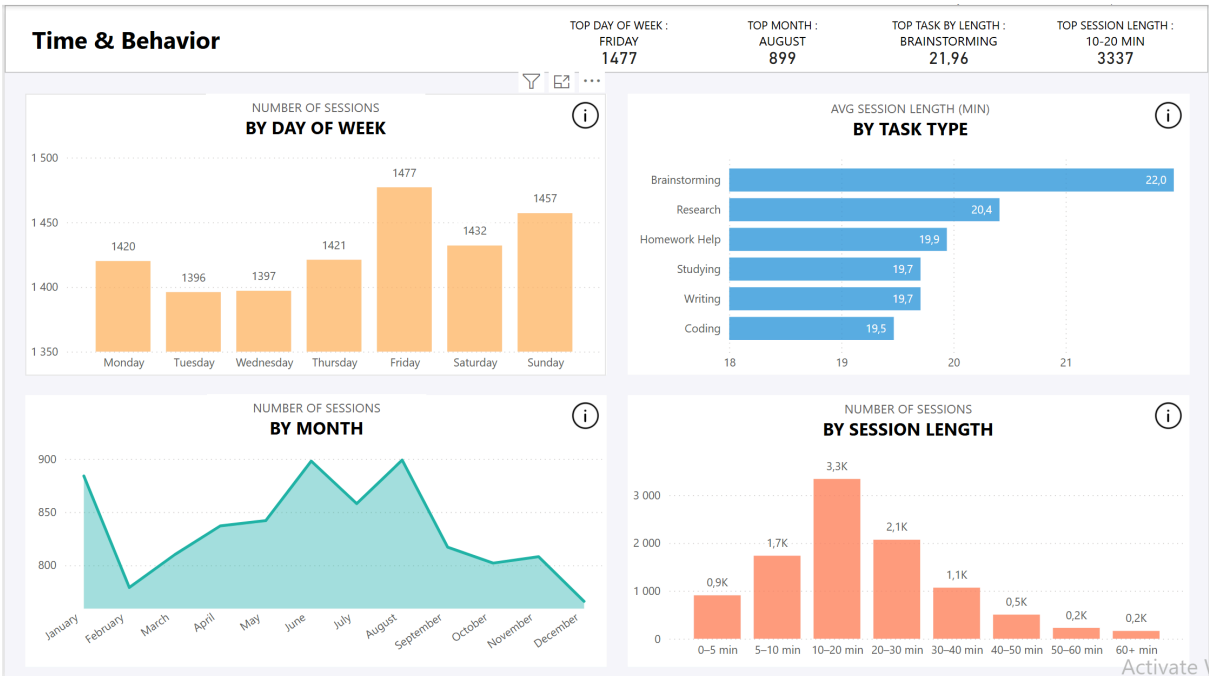
- **Avg Efficiency Score by Session Length**
  Short sessions (5–10 minutes) deliver the highest efficiency (1.72), suggesting that quick and focused interactions with AI lead to better results. **In sessions longer than 10 minutes, efficiency drops and remains relatively stable across all longer time ranges** (≈1.2–1.4), possibly due to user fatigue or less targeted use. **The shortest sessions (0–5 minutes) tend to be the least efficient** (1.11), likely because they often consist of test runs, incomplete attempts, or quick checks without full engagement in the task.

It is worth noting that efficiency scores are also influenced by task type. For instance, research and brainstorming naturally take more time and require more prompts, which may lead to lower efficiency compared to more straightforward, applied tasks.
For clarity, the minimum value on the chart was set to 1.1.

- **Number of Sessions by Efficiency Level**
  The Efficiency Score is divided into three terciles: low (<1.2), medium (1.2–2.2), and high (>2.2). **The majority of sessions fall into the low (43%) and medium (46%) categories.** Nearly half of all cases are in the medium range, suggesting that most students use AI in a fairly balanced way. **High efficiency is observed in only 11% of sessions**, which may reflect both the specifics of the Efficiency Score metric and the nature of academic tasks themselves: in most cases, several follow-up prompts are needed, which naturally keeps students in the "medium" zone.
- **Avg Prompts per Session by Efficiency Level**
  The chart visually represents the logic behind the Efficiency Score: **the fewer prompts it takes to reach a result, the higher the efficiency.** Students in the **low-efficiency group average 7.06** prompts per session, whereas those in the **high-efficiency group need only 2.67 prompts.**

What this means:
The data suggest that **efficiency in working with AI depends on combining a small number of well-thought-out prompts with focused interaction. The most productive sessions are short** (5–10 minutes) and built around a limited set of clear queries. At the same time, **high efficiency is relatively rare** (11%), reflecting both the specifics of the Efficiency Score metric and the nature of most academic tasks.

7.3 Time & Behavior



- **Number of Sessions by Day of Week**
  The highest activity happens on Fridays (1,477 sessions), **Sundays** (1,457), **and Saturdays** (1,432). **The fewest sessions are on Tuesdays** (1,396) **and Wednesdays** (1,397). This dynamic may reflect students' study rhythm: in the middle of the week, their academic schedules are denser, leaving less time for independent study. **Toward the end of the week, activity increases** — likely because homework deadlines peak on Fridays and weekends. Interestingly, coding tasks are more popular on Fridays, whereas studying and homework help dominate on weekends.
  For clarity, the minimum value on the chart was set to 1,350.

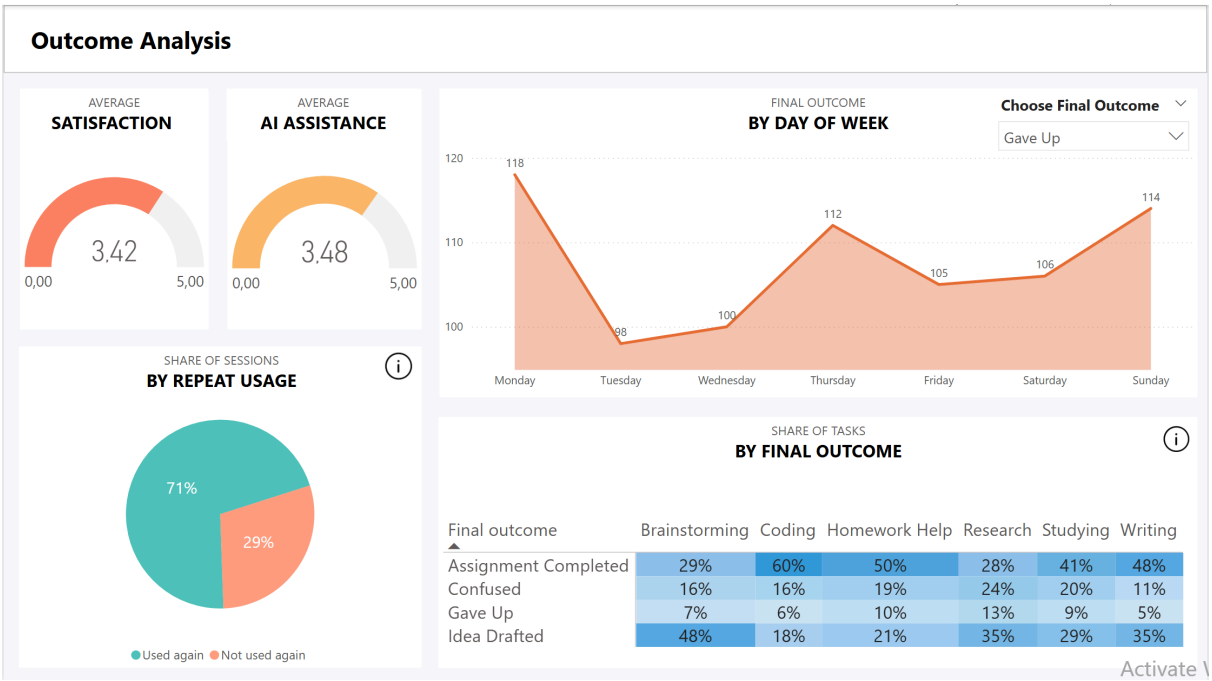
- **Avg Session Length (in minutes) by Task Type**
  The longest sessions are common in Brainstorming (22 minutes) **and Research** (20.4 minutes) — tasks that require idea generation and information analysis. More applied and structured activities — **Homework Help** (19.9), **Studying** (19.7), **Writing** (19.7), and **Coding** (19.5) — **average slightly under 20 minutes**. This reflects the difference between creative, open-ended tasks and more concrete, practical ones.
  For clarity, the minimum value on the chart was set to 18.

- **Number of Sessions by Month**
  Activity peaks in June (898)**, August** (899)**, and January** (884), coinciding with exam periods and university admissions. The spike in August may also be linked to the IT hiring season, as reflected in the noticeable increase in Coding sessions. **The lowest level is seen in February** (779), likely due to winter holidays. Overall, the trend shows that **students' activity grows during periods of academic and career intensity and declines during holiday breaks.**
  For clarity, the minimum value on the chart was set to 750.

- **Number of Sessions by Session Length**
  The majority of sessions last 10–20 minutes (3.3k). **Sessions lasting 20–30 minutes** (2.1k) **and 5–10 minutes** (1.7k) **are also common.** Longer interactions **(30+ minutes) are much less frequent**, while the shortest ones **(<5 minutes) account for only 0.9k cases.** Overall, **most sessions fall within the 5–30 minute range**, reflecting the nature of student tasks — from quick clarifications to working on short portions of study material.

What this means:
Students **most often use AI for short to medium-length sessions (5–30 minutes)**, which aligns with the typical study format: working through assignments or generating ideas. **The longest sessions occur in Research and Brainstorming**, while **applied assignments generally fit into more compact time frames.** Weekly activity is relatively even, with **rises on Fridays and weekends.** Seasonal patterns show **spikes in AI usage during summer and winter exam periods.** Taken together, this indicates that AI serves as a flexible study companion, supporting students both in their daily studies and in high-pressure periods like exams and career preparation.

7.4 Outcome Analysis



- **Share of Sessions by Repeat Usage**
  71% of students return to AI after their first session, indicating that they generally perceive it as a useful tool that meets their expectations. Only **29% of students did not return to AI**, a relatively small share that may be explained by the fact that some students used it for a one-off task or did not gain sufficient benefit from the interaction.
- **Final Outcome by Day of Week**
  The distribution of session outcomes by day of week shows noticeable variations. **The number of sessions that ended with a completed assignment peaks on Fridays**, which aligns with the overall peak of sessions that day and indicates students' tendency to finish tasks by the end of the week. **Sessions resulting in an idea draft are most frequent on Sundays**, which may reflect a preference for using weekends for brainstorming and idea generation. **The proportion of sessions where students get confused rises significantly on Fridays and Saturdays**, coinciding with the overall peak in total activity (see Number of Sessions by Day of Week). **On Sundays**, however, despite the second-highest number of sessions, **the share of confused outcomes declines**, suggesting a more focused and intentional use of AI. **The number of sessions ending with students giving up stays relatively stable across the week**, with slight increases on Mondays, Thursdays, and Sundays, reflecting natural fluctuations in engagement. Overall, the chart highlights the cyclic nature of productivity and challenges depending on the day of the week.
- **Share of Tasks by Final Outcome**
  A comparison of task types by final outcomes shows clear differences in efficiency. **The highest share of completed assignments is found in Coding** (60%), **Homework Help** (50%), and **Writing** (48%), pointing to high efficiency of AI in these task types. **Brainstorming most often results in an idea being drafted** (48%), confirming its value as a preparatory step. **Research stands out with the largest share of sessions ending in**

giving up, along with a more even distribution across other outcomes, reflecting the complexity and uncertainty of this task type. **Studying is largely associated with successful completion but also frequently leads to confusion or idea drafting**, showing that the learning process does not always result in a direct outcome but may include stages of exploration and overcoming difficulties.

What this means:
The analysis of session outcomes shows that **students generally perceive AI as a helpful tool and tend to return to it regularly.** However, **the results of interactions depend on context:** tasks involving practical work (Coding, Homework Help, Writing) most often end in successful completion, whereas more open-ended formats (Research, Brainstorming) are more likely to result in idea generation or giving up. The weekly breakdown highlights recurring patterns: **on Fridays students complete more assignments successfully, while on weekends they turn to AI more often for generating ideas. The number of sessions where students get confused or give up fluctuates** but remains within the bounds of the natural academic workload. Overall, the data highlight that **AI has become embedded in the learning process as a tool for completing specific tasks, generating ideas, and providing support in moments of difficulty, with its perception and efficiency depending both on task type and time of use.**

---

8) Project Conclusions

AI has become integrated into students' learning practices as a **flexible assistant**: most often, they use it in the form of **short to medium-length sessions (5–30 minutes)** during periods of **academic and career intensity** — **at the end of the week and in exam months**. **Efficiency is higher in shorter, more focused sessions** with fewer but well-formulated prompts. The final outcome largely depends on the type of activity: **applied tasks** such as Coding, Homework Help, or Writing **most often end in successful completion**, whereas **creative and research-oriented tasks** such as Brainstorming and Research **often lead to idea generation or giving up**. At the same time, **most interactions result in a positive outcome**: many sessions result in assignment completion, and a significant share in idea drafting. **The proportion of sessions ending in Gave Up remains low** — which aligns with the **71% repeat usage rate**.

Overall, AI has become a **familiar working tool** for students: they turn to it when they need to **quickly complete an assignment, generate an idea, or make sense of a complex topic**. Even if the result is not always perfect, **the experience of using AI is most often perceived as useful** — which explains why the majority of students come back to it again.

---

9) Limitations of the Analysis

- **The Efficiency Score is a Conditional Metric**
  It reflects the ratio between the number of prompts and the achieved outcome, but its interpretation is valid only within this framework (fewer prompts = higher efficiency). In reality, the productivity of interaction with AI may also be influenced by other factors, such as the complexity of the task or the student's individual working style.

- **Contextual Limitations**
  The dataset does not include information about students' geography, level of preparation, or motivation, nor does it capture their academic performance or the technical parameters of interaction. External circumstances such as class schedules, exam periods, and overall academic workload are also not taken into account, even though they may influence students' activity when interacting with AI. As a result, the analysis reflects only the patterns of the sessions themselves rather than students' actual academic success.
- **Synthetic Nature of the Data**
  The dataset was artificially created to model educational scenarios. This means it does not necessarily reflect real student behavior patterns, but rather imitates them based on predefined rules. Therefore, the results of the analysis should be viewed more as a scenario simulation than as a study of actual student experience.
- **Simplification of Outcome Categories**
  All results are grouped into only four categories (Assignment Completed, Idea Drafted, Confused, Gave Up). While this provides a clear and consistent framework, it does not capture more nuanced scenarios — for example, cases of partial success, incomplete understanding, or receiving an incorrect answer.

As a result, the analysis highlights **potential patterns within synthetically modeled data**, but its main value lies in **demonstrating the analytical approach and the use of tools** rather than providing a precise interpretation of real student behavior.

---

10) How to Reproduce

1. **Download** the files .pbix and .xlsx from GitHub.
2. **Open** the .pbix file in Power BI Desktop (only users with Power BI Desktop installed will be able to run the project).
3. **Navigate between report pages**; when hovering over the visualizations, tooltips with brief explanations will appear.
4. If Power BI is not available, **refer to the Notion reports** (RU/EN) — they contain screenshots and GIF animations of the dashboards.

---

11) References and Files

GitHub Repository: ai-assistant-student-life

Excel: ai_assistant_usage_student_life.xlsx

Power BI: ai_assistant_usage_student_life.pbix

Notion presentation (RU): [RU] AI Assistant Usage in Student Life — Аналитический отчет

Notion presentation (EN): [EN] AI Assistant Usage in Student Life — Analytical Report

Data Source: Kaggle