AI Assistant Usage Behavior Analysis

🧑‍💻 Project by: Omar Medhat

🎯 Problem Statement

The company owner gave me access to a dataset containing more than **10,000 recorded sessions** of students using an AI assistant for academic purposes. My mission was to analyze how students interacted with the AI and answer specific business-related questions using SQL.

The aim was to:

- Discover which students benefit the most from the AI assistant
- Understand typical usage behavior
- Identify areas for improvement

📊 Dataset Description

The dataset columns are as follows:

Column Name	Description
StudentLevel	Level of student (Undergraduate, Graduate, or High School)
Discipline	Student's academic field (e.g., Computer Science, Psychology)
TaskType	Task category (e.g., Writing, Coding, Brainstorming)
SessionLengthMin	Duration of the session in minutes
TotalPrompts	Number of prompts used during the session
AI_AssistanceLevel	User rating of AI helpfulness (1 to 5)
FinalOutcome	Session result (e.g., Assignment Completed, Idea Drafted)
SatisfactionRating	How satisfied the user was (1 to 5)
UsedAgain	Did the user reuse the tool? (1 = Yes, 0 = No)

🧠 SQL-Based Analysis & Insights

1 - What types of tasks are most common?

```
SELECT TaskType, COUNT(*) AS frequency
FROM ai_usage
GROUP BY TaskType
ORDER BY frequency DESC
```

> Writing and Studying were the most frequently performed tasks.

2 - Who are the most engaged users?

```
SELECT Discipline, StudentLevel, AVG(SessionLengthMin) AS avgSession,
AVG(SatisfactionRating) AS avgSatisfaction
FROM ai_usage
GROUP BY Discipline, StudentLevel
ORDER BY avgSession DESC, avgSatisfaction DESC
```

> Graduate students in Computer Science consistently had longer and more satisfying sessions.

3 - Do longer sessions result in more productive outcomes?

```
SELECT FinalOutcome, COUNT(*) AS count, AVG(SessionLengthMin) AS avgDuration
FROM ai_usage
GROUP BY FinalOutcome
ORDER BY avgDuration DESC
```

> "Assignment Completed" was most common in long sessions, suggesting deeper academic engagement.

4 - How helpful was the AI across disciplines?

```
SELECT Discipline, COUNT(*) AS helpfulSessions
FROM ai_usage
WHERE AI_AssistanceLevel > 4
GROUP BY Discipline
ORDER BY helpfulSessions DESC
```

> History and Biology disciplines reported the highest satisfaction with AI support.

5 - Which tasks often result in "Idea Drafted" outcomes?

```
SELECT TaskType, COUNT(*) AS totalSessions, AVG(TotalPrompts) AS avgPrompts
FROM ai_usage
WHERE FinalOutcome = 'Idea Drafted'
```

```
GROUP BY TaskType
ORDER BY totalSessions DESC
```

> Brainstorming and Writing had the highest idea drafting rate.

6 - Which student level uses the AI for the longest average time?

```
SELECT StudentLevel, AVG(SessionLengthMin) AS avgTime
FROM ai_usage
GROUP BY StudentLevel
```

> Graduate students tend to engage in longer sessions.

7 - What percentage of users reused the tool?

```
SELECT COUNT(*) * 100.0 / (SELECT COUNT(*) FROM ai_usage) AS reuseRate
FROM ai_usage
WHERE UsedAgain = 1
```

> A high reuse rate (~70%) reflects overall trust in the tool.

8 - What is the satisfaction range?

```
SELECT MAX(SatisfactionRating) AS maxRating,
       MIN(SatisfactionRating) AS minRating
FROM ai_usage
```

> Ratings varied from 1 (very unsatisfied) to 5 (very satisfied).

9 - Which academic fields are most active?

```
SELECT Discipline, COUNT(*) AS sessionCount, AVG(SatisfactionRating) AS avgSatisfaction
FROM ai_usage
GROUP BY Discipline
ORDER BY sessionCount DESC
```

> Computer Science and Biology had the highest session counts and engagement.

10 - Which task types involve more back-and-forth prompts?

```
SELECT TaskType, AVG(TotalPrompts) AS avgPrompts
FROM ai_usage
GROUP BY TaskType
ORDER BY avgPrompts DESC
```

> Coding and Brainstorming tasks involved higher prompt exchanges.

11 - Are there weak points? (Low satisfaction, no return)

```
SELECT *
FROM ai_usage
WHERE SatisfactionRating < 2 AND UsedAgain = 0
```

> Identified users who were dissatisfied and never returned ~ (3.43%)

12 - Assignment Completion Rate

```
SELECT COUNT(*) * 100.0 / (SELECT COUNT(*) FROM ai_usage) AS completionRate
FROM ai_usage
WHERE FinalOutcome = 'Assignment Completed'
```

> Around 48% of all sessions led to a completed assignment, which means that AI is mostly used as a helping tool rather than being totally depended on

13 - Top Disciplines by AI Support Rating

```
SELECT Discipline, AVG(AI_AssistanceLevel) AS avgAiAssistance
FROM ai_usage
GROUP BY Discipline
ORDER BY avgAiAssistance DESC
```

> Psychology students consistently rated AI assistance the highest.

✅ What I Learned from This Project

- I improved my SQL skills with real-life data challenges.

- I learned how to investigate user behavior through structured analysis.
- I practiced transforming data into simple business insights.
- I understood how to detect product strengths and weaknesses.

📌 **Conclusion:**

This project provided a complete SQL-driven journey from problem definition to business impact, and confirmed how AI can be valuable for students when used effectively.