

# Meaning of Dataset Columns

## 1. features.csv (Features Table)

- **feature\_id**: Primary key for this table
- **feature\_name**: The common name used to refer to the feature (e.g., "AI Assistant", "Task Reminders")
- **category**: Helps group features (e.g., "Core", "Collaboration", "Productivity", "Admin", "New", "UI/UX"). "Unknown" if not categorized
- **launch\_date**: For new features (Task Reminders, Voice Assistant, Custom Themes), this is February 20, 2025. For pre-existing features, it defaults to the overall user data start date (user\_start\_date)
- **available\_plans**: Indicates target user tiers (e.g., "All", "Basic, Pro, Enterprise", "Pro, Enterprise", "Enterprise")

## 2. users.csv (Users Table)

- **user\_id**: Primary key for identifying a user
- **name**: The user's provided name
- **email**: User's contact email
- **location**: User's geographical location
- **age**: User's age
- **plan\_type**: The plan the user is associated with
- **sign\_up\_date**: Marks the beginning of the user's lifecycle with the product.
- **is\_active**: Indicates current user status.
- **churn\_date**: If is\_active is False, this is the date they became inactive. NULL if the user is active
- **last\_login\_date**: Indicates the most recent point of known engagement or status change

## 3. subscriptions.csv (Subscriptions Table)

- **subscription\_id**: Primary key for this table
- **user\_id**: Foreign key linking to the users table
- **plan**: The plan active during this segment
- **duration**: Indicates if the plan is "monthly" or "yearly"
- **start\_date**: When this particular plan period began for the user

- **end\_date:** When this plan period ended, due to churn, plan change
- **status:** Describes the state, such as "active", "cancelled" (due to churn), or "changed\_plan"

#### • **4. billing.csv (Billing Table)**

- **billing\_id:** Primary key for this table
- **user\_id:** Foreign key linking to the users table
- **billing\_date:** The transaction date for the billing
- **plan\_type:** The plan associated with this specific charge
- **amount:** The monetary value of the charge
- **currency:** The currency code for the transaction
- **status:** Typically "paid" in this dataset
- **payment\_method:** The financial instrument used (e.g., "credit\_card", "paypal", "bank\_transfer")

#### **5. sessions.csv (Sessions Table)**

- **session\_id:** Conceptually the primary key for a unique session instance
- **user\_id:** Foreign key linking to the users table
- **login\_time:** Marks the beginning of a user's interaction period
- **logout\_time:** Marks the end of a user's interaction period
- **device\_type:** The platform from which the user accessed the service

#### • **6. activity\_log.csv (Activity Log Table)**

- **activity\_id:** Primary key for this table
- **session\_id:** Foreign key linking to the sessions table
- **user\_id:** Foreign key linking to the users table
- **activity\_type:** A string describing the action (e.g., "login", "create\_task", "used\_ai\_assistant", "task\_reminder")
- **timestamp:** The precise moment of the action

#### **7. feedback.csv (Feedback Table)**

- **feedback\_id:** Primary key for this table
- **user\_id:** Foreign key linking to the users table
- **session\_id:** Foreign key linking to the sessions table

- **submission\_timestamp**: When the feedback was recorded
- **rating**: A quantitative measure of satisfaction
- **comment\_type**: Helps classify feedback (e.g., "Positive", "Negative", "Neutral/Suggestion", or "Unknown" if the rating was invalid)
- **comment**: Qualitative insights from the user
- **feature\_area**: Helps associate feedback with product components

## 8. support\_tickets.csv (Support Tickets Table)

- **ticket\_id**: Primary key for this table
- **user\_id**: Foreign key linking to the users table
- **feature**: The area of the product the user had an issue with (e.g., "AI Assistant", "Billing", "Task Manager", or "General Inquiry"). Conceptually links to features.feature\_name
- **submitted\_at**: When the user reported the issue
- **priority**: Indicates urgency (e.g., "low", "medium", "high")
- **resolved**: Current resolution status
- **resolved\_at**: When the issue was addressed. NULL if the ticket is not yet resolved
- **status**: The stage in the support workflow (e.g., "Open", "In Progress", "Resolved")

## 9. system\_metrics.csv (System Metrics Table)

- **timestamp**: Primary key (conceptually, as each hour should be unique)
- **active\_users**: A measure of concurrent user engagement
- **request\_count**: An indicator of overall system load
- **error\_count**: A measure of system instability or issues
- **error\_rate**: Normalized measure of error frequency
- **cpu\_usage**: An indicator of processing load
- **memory\_usage**: An indicator of memory pressure
- **response\_time**: A key performance indicator for user experience