

The first 49 questions are based on this dataset downloaded from Kaggle.

<https://www.kaggle.com/code/parthdande/amazon-book-dataset-analysis>

Questions from 50 to 108 are based on artificially generated data for an educational website. Here is the detailed explanation for each table.

### 1. Users Table

Description: This table contains information about all users who have signed up on the platform. Users can have roles such as students, instructors, or administrators.

Columns:

user\_id: Unique 6-digit identifier for each user.

name: Name of the user

date\_of\_birth: Date of birth of the user (dd-mm-yyyy format).

email: Unique email address for each user.

gender: Gender of the user (male/female).

category: User category (working, student, homemaker, retired).

registration\_date: The date the user registered on the platform.

country: The country of the user

last\_login\_date: The most recent date the user logged in.

user\_type: Type of user (student, instructor, admin).

### 2. Courses Table

Description: This table stores details of all courses offered on the platform. Courses are divided into free and paid categories.

Columns:

course\_id: Unique 5-digit identifier for each course.

title: Name of the course

category: Broad category of the course (e.g., Technology, Arts, Business).

language: Language of the course

subtitles: Indicates if subtitles are available (yes/no).

price: Cost of the course. Free courses have a price of 0; others have varying prices.

instructor\_id: Reference to the user ID of the instructor.

length: Duration of the course in hours and minutes (e.g., "10:30").

difficulty\_level: Difficulty level of the course (Beginner, Intermediate, Advanced).

created\_on: Date the course was created.

last\_updated\_on: Date the course was last updated.

number\_of\_videos: Total number of videos in the course.

number\_of\_quizzes: Total number of quizzes in the course.

view\_count: Number of views for the course.

average\_watch\_duration: Average watch time for the course (in hours and minutes).

certification\_available: Whether a certificate is granted upon completion (yes/no).

completion\_rate: Percentage of enrolled users who completed the course.

bookmark\_count: Number of users who bookmarked the course.

### 3. Ratings Table

Description: This table contains user reviews and ratings for various courses.

Columns:

review\_id: Unique identifier for each review.

course\_id: Reference to the course being reviewed.

student\_id: Reference to the user who submitted the review.

rating: Numeric rating given by the student (1 to 5).

review\_text: Text of the review (optional, some reviews have no text).

first\_reviewed\_on: Date the review was first submitted.

last\_updated\_on: Date the review was last updated (if applicable).

helpful\_count: Number of times the review was marked as helpful by others.

reported\_count: Number of times the review was flagged as inappropriate.

#### 4. Payments Table

Description: This table records payment transactions for paid courses. It includes details like discount percentages, payment methods, and payment statuses.

Columns:

payment\_id: Unique identifier for each payment transaction.

enrollment\_id: Reference to the enrollment the payment is associated with.

discount\_percent: Discount applied (if any).

amount\_paid: The final amount paid after applying discounts.

coupon\_code: Coupon code applied, if any (e.g., "UTUBE23").

payment\_method: Payment method used (e.g., card, PayPal, UPI).

transaction\_date: Date and time of the transaction.

transaction\_id: Unique identifier generated by the payment gateway.

payment\_status: Status of the payment (success, failed, pending).

#### 5. Enrollments Table

Description: This table tracks enrollments for both free and paid courses. Only successful payments are linked here.

Columns:

enrollment\_id: Unique identifier for each enrollment.

student\_id: Reference to the user who enrolled.

course\_id: Reference to the course enrolled in.

enrollment\_date: Date the enrollment occurred.

#### 6. Refunds Table

Description: This table contains details about refunds issued for specific enrollments.

Columns:

refund\_id: Unique identifier for each refund record.

enrollment\_id: Reference to the enrollment for which the refund was issued.

refund\_date: Date the refund was processed.

refund\_reason: Reason for the refund (optional).

refund\_amount: The amount refunded to the user.