

Homework 01: Data Modeling

CS 355 Database Systems
Habib University
Fall 2025

1 Instructions

- The deadline for submitting this Homework is 1st October 2025, 11:59 PM.
- You may submit it by 2nd October 2025, 11:59 PM with a 20% penalty.
- This homework must be submitted online via CANVAS.
- You are required to submit a **single pdf** which contains the solutions of all 3 questions.
- This homework will be done in pairs. Please ensure that both the members have joined the same team on Canvas. **If your team does not correspond with your group name on Canvas, then your submission will not be graded.**
- The pdf file should be named HW_01_team_XX.pdf where XX will be replaced with your respective team number on Canvas.
- *Files that don't follow the appropriate naming convention will not be graded.*

1.1 Marking scheme

This Homework will be marked out of 100.

- 30 Marks are for Normalization
- 40 Marks are for Orchard ERD
- 30 Marks are for Khaadi ERD

1.2 Submission Guidelines

- Models can be developed in DB Designer or any Database Diagram tool that can generate a schema.
- **No handwritten submissions will be accepted**
- If you make any assumptions regarding the scenario, clearly state them in your pdf document. **Your assumptions should not contradict the actual scenario.**

1.3 Late submission policy

Refer to the course syllabus for the late submission policy.

1.4 Use of AI

Taking help from any AI-based tools such as ChatGPT is strictly prohibited and will be considered plagiarism.

1.5 Viva

Course staff may call any student for Viva to provide explanation for their submission.

2 Objective

The purpose of this homework is to enable students to analyze real-world business specifications and construct their data model. This skill helps them build the data management system of any business entity.

3 Questions

Question 01 - Normalization

Imagine you are employed as a Data Analyst at Habib University responsible for managing students' extracurricular activities information. A student can be a part of multiple extracurricular activities. Each of these activities are scheduled at different timings. During each time slot, there is always an Activity Lead who overlooks all the participating students in that group. A student can attend any of the time slots, as per their availability, so no specific slot will always have the same set of students. The university administration is using a table to store this information. Some of the rows from that table are given in Figure 1 for reference.

- Please analyze the provided extracurricular activities data table, and derive a relation schema for it in 1NF, by specifying the primary keys. Give reasons for choosing the primary keys.
- Identify full, partial, and transitive functional dependencies that exist among its attributes. If there are no dependencies, please state that explicitly.
- Normalize the schema up to the Third Normal Form (3NF) by creating new tables. Provide schemas for 2NF, and 3NF.

StudentID [PK]	StudentName	DOB	Year	Activity	ActivityType	Activity Timings	ActivityLead	ActivityJoiningDate	ActivityLevel	GuardianID	GuardianName	GuardianContact
101	Ammara Javed	13-02-2005	1st	Basketball	Sports	Morning	Mr. X	31-08-2024	Intermediate	1	Javed Jaffrey	555-1234
				Drama Club	Performance	Evening	Mr. Y	15-09-2024	Beginner			
				Chess Club	Hobby	Morning	Mr. B	10-09-2023	Advanced			
102	Bilal Syed	22-08-2004	2nd	Coding	Educational	Afternoon	Mr. C	01-01-2024	Intermediate	2	Mariam Syed	555-5678
				Debate	Performance	Evening	Ms. D					
						Afternoon	Ms. G	15-06-2024	Beginner			
103	Dawood Waqar	10-09-2005	1st	Football	Sports	Morning	Mr. E	10-08-2024	Advanced	3	Sarah Bilal	555-4321
				Robotics	Educational	Afternoon	Mr. H					
						Evening	Mr. I	01-09-2024	Beginner			
104	Erum Pervez	20-01-2002	4th	Drama Club	Performance	Evening	Ms. A	01-09-2021	Advanced	4	Pervez Waheed	555-9876
				Cooking Club	Hobby	Afternoon	Ms. J	19-01-2024	Beginner			
				Debate	Performance	Afternoon	Ms. G	04-12-2023	Beginner			
105	Faheem Haq	23-05-2003	3rd	Basketball	Sports	Morning	Mr. X	22-11-2022	Advanced	5	Nadeem Haq	555-6543
						Afternoon	Mr. Y					
				Coding	Educational	Evening	Mr. C	07-10-2022	Advanced			
							Ms. D					
106	Ghazal Khan	24-07-2004	2nd	Chess Club	Hobby	Morning	Mr. B	26-10-2023	Intermediate	6	Mikaal Khan	555-3456
				Robotics	Educational	Morning	Mr. H	19-06-2024	Beginner			
						Evening	Mr. I					
				Cooking Club	Hobby	Afternoon	Ms. J	29-04-2024	Beginner			
107	Harish Chandra	01-01-2022	4th	Football	Sports	Morning	Mr. E	18-03-2021	Advanced	7	Laxman Chandra	555-7890
						Afternoon	Mr. F					
108	Ibrahim Shah	21-04-2003	3rd	Basketball	Sports	Morning	Mr. X	12-12-2022	Intermediate	8	Adam Shah	555-6789
						Afternoon	Mr. Y					
109	Javeria Merchant	20-09-2004	2nd	Cooking Club	Hobby	Afternoon	Ms. J	17-11-2023	Intermediate	9	Hasham Merchant	555-1239
110	Kiran Zahid	29-03-2005	1st	Coding	Educational	Afternoon	Mr. C	30-08-2024	Beginner	10	Zahid Siddiqui	555-0123
				Drama Club	Performance	Evening	Ms. D					
						Evening	Ms. A	30-08-2024	Beginner			
111	Monis Shahmeer	16-04-2003	3rd	Coding	Educational	Afternoon	Mr. C	15-06-2023	Intermediate	11	Shahmeer Zulfiqar	555-2345
						Evening	Ms. D					
				Football	Sports	Morning	Mr. E	18-12-2022	Advanced			
						Afternoon	Mr. F					
				Football	Sports	Morning	Mr. E	25-08-2021	Advanced			
						Afternoon	Mr. F					
112	Abdul Ahad	13-05-2002	4th	Basketball	Sports	Morning	Mr. X	27-08-2021	Advanced	12	Abdul Samad	555-4567
						Afternoon	Mr. Y					
				Debate	Performance	Afternoon	Ms. G	01-01-2022	Intermediate			

Figure 1: Extracurricular activities information

Question 02 - Orchard ERD

A fruit orchard company manages multiple orchards across different regions of Pakistan. Each orchard specializes in cultivating a variety of fruits (e.g., mango, orange, apple, guava, etc). Each orchard has a location (city, province) and is managed by a designated orchard manager.

Every orchard must hold a government-issued license to operate. Each orchard has exactly one license, and each license belongs to exactly one orchard. For each license, the system records the license ID, issue date, and expiry date.

Every orchard is divided into multiple plots of land. Each plot is used to cultivate exactly one type of fruit, but the same fruit may be grown in different plots (sometimes even within the same orchard). Each plot has information about its soil type, area (in acres), and year it was first cultivated.

Each fruit type has details such as fruit name, season, and average yield per acre. Fruits are harvested multiple times in a year. For each harvest, the system keeps track of the date, the quantity harvested (in kg), and the plot it came from.

The orchard company employs multiple workers. Each worker is assigned to one orchard but may be rotated between different plots within that orchard. Workers can be specialized (e.g., irrigation, pesticide application, harvesting), and the system stores each worker's name, CNIC, contact, and specialization. Workers are supervised by orchard managers, but an orchard manager may supervise multiple orchards.

The company sells fruits in bulk to different buyers (wholesalers, retailers, and/or exporters). For each sale, the system stores the buyer details (name, type, contact), the fruit type sold, the quantity sold, the price per kg, and the date of sale. Sometimes, a single sale can include multiple fruit types.

Management is interested in generating the following reports:

- Annual yield per fruit type per orchard
- Which orchard produces the most fruit of a given type
- Productivity of workers in terms of harvested quantity
- Sales trends of different fruit types across years
- The contribution of each orchard to total revenue

- Expiring orchard licenses in the upcoming year

Based on the above scenario, construct an Entity-Relationship Diagram (ERD) that includes:

- Entities with appropriate attributes (identify primary keys and foreign keys).
- Relationships among entities with appropriate cardinalities.
- Any assumptions you make to clarify ambiguities in the scenario.

Question 03 - Khaadi ERD

Khaadi is one of the most popular fashion brands in Pakistan. You have been assigned the task of designing a data model for their online product catalog. The business details are provided below. Your task is to construct a normalized data model that includes:

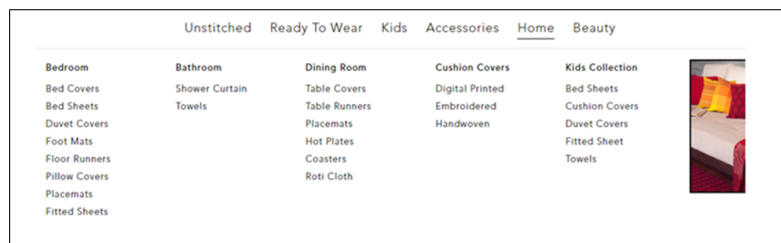
- Entities, attributes, primary keys (PKs), and foreign keys (FKs)
- Relationships and their cardinalities

Business Details

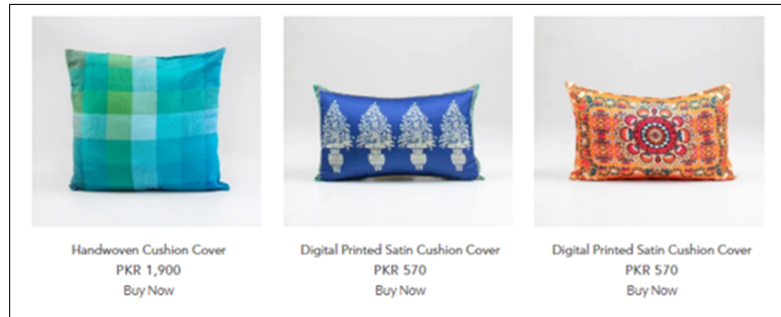
Khaadi's catalog is organized into various departments as is evident from their website:



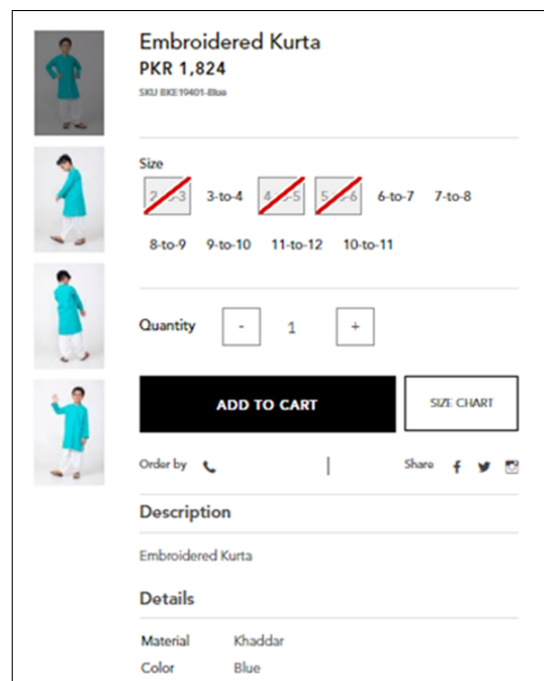
Each department contains multiple categories and sub-categories. A category can belong to multiple departments. A sub-category can belong to multiple categories.



When a category is selected, the products associated with that category are displayed.



Clicking on a product reveals further details, including the product title, SKU, price, description, size, material, color, and quantity in stock. Certain products are available in multiple sizes, and the system tracks the stock level for each size. If a product is available in multiple sizes, they share the same SKU. The price of a product does not change with size.



Throughout the year, different sales events are launched. During these sales, selected products are offered at discounted prices. To set up a sale, the system specifies a timeline (start and end dates), the products included in the sale, and the percentage discount applied to each product. Not all

products are required to be part of a sale, and different products may have different discount percentages.