Homework 01: Data Modeling

CS 355 Database Systems
Habib University
Fall 2024

1 Instructions

- The deadline for submitting this Homework is 29 September 2024, 11:59 PM.
- This homework must be submitted online via CANVAS.
- You are required to submit a **single pdf** which contains solutions of all questions.
- This homework will be done in pairs. One grade will be allocated to the entire team, and there will not be individual grades for each member.
- 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 Question 1
- 30 Marks are for Question 2
- 40 Marks are for Question 3

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 group/student for Viva to explain their submission.

2 Objective

This assignment enables students to analyze real-world business scenarios, identify data to be stored, and build a normalized relational model. The assignment also assesses their understanding of different stages of normalization.

3 Exercise

Question 01 - Employee Management System

We are building an enterprise application for an organization to assist with employee appraisals. The system will store information about employees, their work histories at the organization, and the projects they are working on.

The company operates in multiple cities across Pakistan, with its headquarters currently located in Islamabad. However, the system should allow for the headquarters to be changed in the future if necessary. The company has over 10 departments, each with more than 50 employees. An employee can belong to only one department, but a manager may oversee multiple departments.

There is no physical department concept within the company, meaning it does not explicitly track which department exists in which city. Instead, this information can be inferred from employee data; if an employee from a specific department is located in a particular city, it indicates that the department has a presence in that city.

Each employee is assigned to a project, and each project has a start date. If a project is still ongoing, it will not yet have an end date. An employee may be assigned to a project but not yet have any tasks assigned. While some projects may be managed by a single department, most involve multiple departments. One of the employees assigned to a project can be designated as the project lead. The project lead creates different tasks for the project, specifying each task's start date, expected end date, and completion date, and assigns these tasks to employees. The completion date of each task is updated by the project lead once the task is completed. If no completion date is assigned, it means the task is incomplete. Each task may also be assigned a priority of "Top," "Medium," or "Low."

The top management is intended to receive the following reports after the implementation of the system.

- List the number of tasks completed by each employee.
- List employees who have missed deadlines, including the number of deadlines missed.
- Rank employees based on the number of tasks completed and their average task completion time.
- List all ongoing projects, including start dates, current progress, and

departments involved.

- List project leads along with their assigned projects, including task completion rates and the number of deadlines met and missed.
- List all projects with their start date, expected end date, and completion status.
- List all departments with the number of projects in progress over a specified period.
- List cities with the number of employees and departments in each city.
- List all tasks assigned to employees, including start dates, expected end dates, and priorities (Top, Medium, Low).
- List all incomplete tasks, including assigned employees, start dates, and priorities.
- Provide a summary of the number of tasks remaining, categorized by priority.
- List the projects an employee has worked on since joining the company.
- List all tasks assigned to an employee since joining the company, including whether the deadlines were met or not.

Entities correctly identified	25%
All Attributes mentioned	15%
Relationships drawn and resolved correctly	20%
Correct Cardinalities	15%
PK, FK identified and FK is placed in appropriate Entity table	25%

Table 1: Rubric for Question 01

Question 2: Cricket Equipment Shop

Assume that in your Database course, you are working on a group project to develop a software application for managing a cricket equipment shop that is currently being managed through some MS Excel sheets in the owner's computer. During your interview with the shop owner, he provided you with the snapshots (Table 3 & Table 4 on the next page) from Excel sheets that he currently uses.

The shop owner has also informed you about the following facts regarding his business:

- For each ItemID, there is only one Supplier.
- One PurchaseOrder can only include purchases from only one supplier.
- One SalesOrder can only include sales to only one customer.

Based on this information, your team member has sent you the following proposed relational schema for your software application:

> SalesOrders (SalesOrderID, ItemID, ItemName, Quantity, UnitPrice, CustomerID, CustomerName, CustomerPhone)
>
> PurchaseOrders (PurchaseOrderID, ItemID, ItemName, Quantity, UnitPrice, SupplierID, SupplierName, SupplierPhone)

Is the relational schema proposed by your team member in 1NF? If not, convert it into 1NF. Then, check if the converted schema is in 2NF. If not, convert it into 2NF. Then, check if your new schema is in 3NF. If not, convert into 3NF. Please provide the relevant functional dependencies that guided your conversion process at each step.

Rubric

1NF is correct	15%
2NF is correct	30%
3NF is correct	35%
Functional dependencies & transitive dependencies correctly identified	20%

Table 2: Rubric for Question 02

SalesOrderID	ItemID	ItemName	Quantity	UnitPrice	CustomerID	CustomerName	CustomerPhone
1	2	CA Bat	1	55000	1	Amir	1234567
	4	Duke Ball	2	5000			
2	3	Kookaburra Ball	5	4000	3	Shoaib	1234569
3	1	CA Gloves	2	10000	4	Rizwan	1234568
	5	Gray-Nicolls Bat	2	75000			
	6	Gray-Nicolls Gloves	2	15000			
4	3	Kookaburra Ball	3	4100	6	Waqar	1234561
5	2	CA Bat	1	52000	7	Inzamam	1234560

Table 3: Sales Order Details

PurchaseOrderID	ItemID	ItemName	Quantity	UnitPrice	SupplierID	SupplierName	SupplierPhone
1	2	CA Bat	10	50000	1	CA	2234567
	1	CA Gloves	20	9000			
2	3	Kookaburra Ball	50	3500	3	Kookaburra	2234569
3	6	Gray-Nicolls Gloves	10	12500	2	Gray-Nicolls	2234561
	5	Gray-Nicolls Bat	10	70000			
4	4	Duke Ball	50	4500	5	Duke	2234560
5	2	CA Bat	5	50000	1	CA	2234567

Table 4: Purchase Order Details

Question 03 - Khadi Online Store

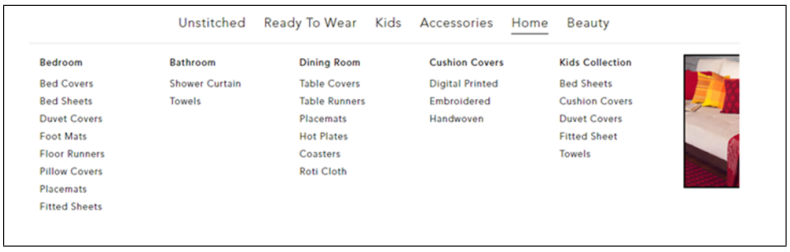
Khaadi is one of the most popular fashion brands in Pakistan. You have been assigned the task of designing a data model for their online product catalog. The business details are provided below. Your task is to construct a normalized data model that includes:

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

Business Details

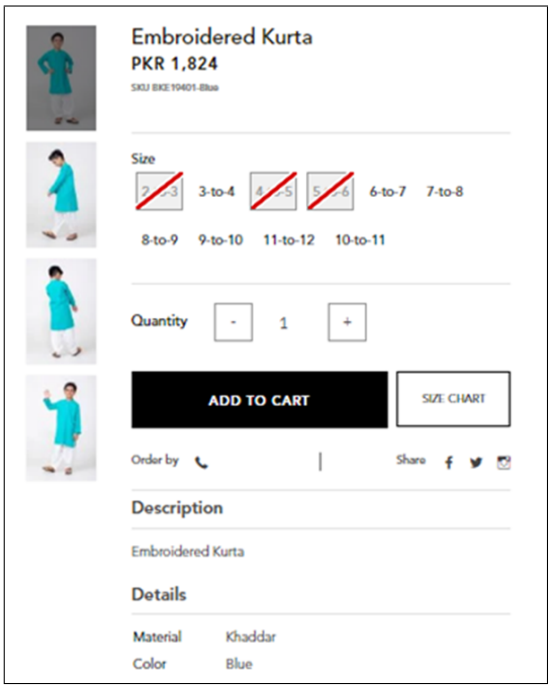
Khaadi's catalog is organized into various departments as is evident from their website:

Each department contains multiple categories and sub-categories.



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.



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.

Rubric

Entities correctly identified	20%
All Attributes mentioned	15%
Relationships drawn and resolved correctly	20%
Correct Cardinalities	10%
PK, FK identified and FK is placed in appropriate Entity table	10%
ERD is 3NF	25%

Table 5: Rubric for Question 03