Faculty of Computing and Information Technology (FCIT)
Database Systems - ASSIGNMENT#2
BS(CS/IT) - Fall 2023 (Semester Fall 2024)

Total Marks: 100   Deadline: December 16, 24

Instructions

- You must complete all tasks individually. Absolutely **NO** collaboration is allowed.
- Any traces of plagiarism/cheating would result in an "F" grade in this course.
- Late submissions will **NOT** be accepted, in any case.
- You are also required to submit pdf file to the respective TAs.
- Name of your file should be **RollNumber_Assignemnt#2(Normalization)**. Use the same title as a subject of your email for submission.
- Clearly mention the key types of each of the relation obtained as a result of normalization.

Quesiton#1

The set of functional dependencies for relation R(A,B,C,D,E) is:

F = {AB → C, AB → D, D → A, BC → D, BC → E}

a. Find all candidate keys.
b. Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF).
c. Is this relation in BCNF? If not, show all dependencies that violate it.
d. Is this relation in 3NF? If not, show all dependencies that violate it.

Quesiton#2

The set of functional dependencies for a relation R(A,B,C,D,E,F,G) is:

F = {AD → BF, CD → EGC, BD → F, E → D, F → C, D → F}

a. Find all candidate keys.
b. Find F-closure.
c. Find the minimal cover for the above set of functional dependencies.

QUESTION#3

A leasing company lease flats to its clients. The sample data of the company is given in the table below. A place number (placeNo) uniquely identifies each single room in all flats and is used when leasing a room to a student. Perform the following tasks:

- Identify the functional dependencies that exist between the columns of the table and identify the primary key and any alternate keys (if present) for the table.
- Describe why the table is not in 3NF.
- The table is susceptible to update anomalies. Provide examples of how insertion, deletion, and modification anomalies could occur on this table.

leaseNo	bannerID	placeNo	fName	lName	startDate	finishDate	flatNo	flatAddress
10003	B017706	78	Jane	Watt	01/09/2010	30/06/2011	F56	34 High Street, Paisley
10259	B017706	88	Jane	Watt	01/09/2011	30/06/2012	F78	111 Storrie Road, Paisley
10364	B013399	89	Tom	Jones	01/09/2011	30/06/2012	F78	111 Storrie Road, Paisley
10566	B012124	102	Karen	Black	01/09/2011	30/06/2012	F79	120 Lady Lane, Paisley
11067	B034511	88	Steven	Smith	01/09/2012	30/06/2013	F78	111 Storrie Road, Paisley
11169	B013399	78	Tom	Jones	01/09/2012	30/06/2013	F56	34 High Street, Paisley

Question# 4:

A start-up is working on an online personal financial management system. The goal of the system is to provide the users an opportunity to obtain item-level purchase data from as many sources as possible in order to improve the accuracy of budget management and control activities (instead of only at the level of the total of each purchase). For example, let's assume a customer purchases three books from a major online bookseller. For most financial management software systems, the system only receives the total of the purchase from a bank or other financial institution. In the case of this start-up, the intent is to create a link between the financial transaction and the vendor's system data so that the financial management system retrieves product details from the vendor. Now it will be easy for the customer to classify one book as self-help, the other one as a business expense, and the third one as entertainment without having to resort to an analysis of receipts. To provide this capability, the system maintains the following data regarding the transactions:

TransactionID, CustomerID, CustomerName, CustomerEmail, TransactionDate, TransactionTime, TransactionTotalAmount, TransactionTax, ProductID, ProductDescription, ProductCategory, ProductManufacturerID, ManufacturerName, ProductListPrice, ProductPurchasePrice, ProductQuantity, TransactionProductTotal

Sample data for this set of attributes is as follows:
823434434582, 2434254, Silver Patrick, psilver@mail. net, 9/2/2015, 10.28.34, $167.23, $10.37, {78234, "Achieving One's Fullest Potential," self-help, 145432, Brown and Gray, $29.95, $24.75, 1, $24.75 | 4782349, "Programming Server-side Solutions with Python," Programming, 63453632, Green & Yellow, $47.95, $39.99, 2, $79.98 | 2342343, "Murder at Eleven," fiction, 145432, Brown and Gray, $14.95, $12.50, 5, $62.50}. Note that information regarding specific products is repeated multiple times in the sample data set and each repeated set is separated by the "|" symbol

Tasks:
1. Identify the functional dependencies between the attributes.
2. Identify the reasons why this set of data is not in 3NF and indicate the normal form (if any) it is in.
3. Including all intermediate stages, organize the attributes into a set of 3NF relations.

Question#5

A music streaming platform tracks its users, playlists, songs, and playback history in a single table. Each time a user listens to a song, a record is created in the system. Over time, this table has grown large, and the platform management wants to organize the data to reduce redundancy, improve data integrity, and scale effectively for millions of users and songs. Normalize the given table up to 3NF while fulfilling the following requirements:

PlaybackID	UserID	UserName	UserEmail	PlaylistID	PlaylistName	SongID	SongTitle	Artist	Album	PlayDateTime
1	U001	John Doe	john.doe@mail.com	PL001	Chill Vibes	S001	Ocean Eyes	Billie Eilish	Don't Smile At Me	2024-12-01 10:30:00
2	U001	John Doe	john.doe@mail.com	PL001	Chill Vibes	S002	Blinding Lights	The Weeknd	After Hours	2024-12-01 11:00:00
3	U002	Jane Smith	jane.smith@mail.com	PL002	Workout Hits	S002	Blinding Lights	The Weeknd	After Hours	2024-12-02 09:00:00
4	U003	Sarah Johnson	sarah.johnson@mail.com	PL003	Road Trip	S003	Shape of You	Ed Sheeran	Divide	2024-12-03 08:00:00
5	U002	Jane Smith	jane.smith@mail.com	PL002	Workout Hits	S004	Bad Guy	Billie Eilish	When We All Fall	2024-12-02 10:00:00
6	U001	John Doe	john.doe@mail.com	PL004	Party Time	S005	Uptown Funk	Bruno Mars	Uptown Funk	2024-12-01 12:00:00
7	U003	Sarah Johnson	sarah.johnson@mail.com	PL003	Road Trip	S006	Perfect	Ed Sheeran	Divide	2024-12-03 09:00:00
8	U002	Jane Smith	jane.smith@mail.com	PL002	Workout Hits	S007	Rockstar	Post Malone	Beerbongs & Bentleys	2024-12-02 11:00:00
9	U001	John Doe	john.doe@mail.com	PL001	Chill Vibes	S001	Ocean Eyes	Billie Eilish	Don't Smile At Me	2024-12-01 10:30:00
10	U003	Sarah Johnson	sarah.johnson@mail.com	PL003	Road Trip	S008	Happier	Marshmello	Joytime II	2024-12-03 10:00:00

Question# 6:

The University Enrollment System tracks information about students, courses, instructors, departments, and enrollment details in a single table. This table is poorly designed and contains redundant and repeated data. Each time a student enrolls in a course, all related information about the student, instructor, course, and department is repeated in a new row. This leads to issues such as redundancy, anomalies (update, insertion, and deletion), and a lack of scalability.

Normalize the table given below by organizing the data into multiple related tables to eliminate redundancy, maintain data integrity, and scale efficiently for thousands of students, courses, and departments.

Attributes:
StudentID, StudentName, StudentEmail, StudentPhone, CourseID, CourseName, InstructorID, InstructorName, InstructorPhone, InstructorEmail, DepartmentID, DepartmentName, DepartmentHead, EnrollmentDate, Grade, ClassRoom, Schedule

Sample Data

StudentID	StudentName	StudentEmail	StudentPhone	CourseID	CourseName	InstructorID	InstructorName	InstructorPhone	Instructor Email	DepartmentID	DepartmentName	DepartmentHead	EnrollmentDate	Grade	ClassRoom	Schedule
1	Alice Smith	alice.smith@mail.com	12345 67890	101	Data Structures	501	Dr. John Brown	987654 3210	john.brown@mail.com	301	Computer Science	Dr. Emily White	01/15/2024	A	Room 101	Mon/Wed 9:00-10:30
2	Bob Johnson	bob.johnson@mail.com	23456 78901	102	Operating Systems	502	Dr. Susan Green	876543 2109	susan.green@mail.com	301	Computer Science	Dr. Emily White	01/16/2024	B+	Room 102	Tue/Thu 11:00-12:30
1	Alice Smith	alice.smith@mail.com	12345 67890	102	Operating Systems	502	Dr. Susan Green	876543 2109	susan.green@mail.com	301	Computer Science	Dr. Emily White	01/16/2024	A-	Room 102	Tue/Thu 11:00-12:30

Question# 7:

The manager of a consulting firm has asked you to evaluate a database that contains the table structure shown in below. It was created to enable the manager to match clients with consultants. The objective is to match a client within a given region with a consultant in that region and to make sure that the client's need for specific consulting services is properly matched to the consultant's expertise. For example, if the client needs help with database design and is located in the Southeast, the objective is to make a match with a consultant who is located in the Southeast and whose expertise is in database design. (Although the consulting company manager tries to match consultant and client locations to minimize travel expense, it is not always possible to do so.) The following basic business rules are maintained:

- Each **client** is located in one **region**.
- A **region** can contain many **clients**.
- Each **consultant** can work on many **contracts**.
- Each **contract** might require the services of many **consultants**.
- A **client** can sign more than one **contract**, but each **contract** is signed by only one **client**.
- Each **contract** might cover multiple **consulting classifications** (e.g., database design, networking).
- Each **consultant** is located in one **region**.
- A **region** can contain many **consultants**.
- Each **consultant** has one or more **areas of expertise** (e.g., database design, networking).

- Each **area of expertise** (class) can have many **consultants**.

ATTRIBUTE NAME	SAMPLE VALUE	SAMPLE VALUE	SAMPLE VALUE
CLIENT_NUM	298	289	289
CLIENT_NAME	Marianne R. Brown	James D. Smith	James D. Smith
CLIENT_REGION	Midwest	Southeast	Southeast
CONTRACT_DATE	10-Feb-2018	15-Feb-2018	12-Mar-2018
CONTRACT_NUMBER	5841	5842	5843
CONTRACT_AMOUNT	$2,985,000.00	$670,300.00	$1,250,000.00
CONSULT_CLASS_1	Database Administration	Internet Services	Database Design
CONSULT_CLASS_2	Web Applications		Database Administration
CONSULT_CLASS_3			Network Installation
CONSULT_CLASS_4			
CONSULTANT_NUM_1	29	34	25
CONSULTANT_NAME_1	Rachel G. Carson	Gerald K. Ricardo	Angela M. Jamison
CONSULTANT_REGION_1	Midwest	Southeast	Southeast
CONSULTANT_NUM_2	56	38	34
CONSULTANT_NAME_2	Karl M. Spenser	Anne T. Dimarco	Gerald K. Ricardo
CONSULTANT_REGION_2	Midwest	Southeast	Southeast
CONSULTANT_NUM_3	22	45	
CONSULTANT_NAME_3	Julian H. Donatello	Geraldo J. Rivera	
CONSULTANT_REGION_3	Midwest	Southeast	
CONSULTANT_NUM_4		18	
CONSULTANT_NAME_4		Donald Chen	
CONSULTANT_REGION_4		West	

Question# 8:

An online movie rental platform keeps track of movie rentals using a single table. The platform provides details about movies, customers, and rental transactions.

Business Rules

1. Each movie has a unique MovieID, title, and genre.
2. Customers are identified by a unique CustomerID and have details like name and email.
3. A customer can rent multiple movies in a single transaction, leading to multiple MovieIDs being recorded in the same column as a comma-separated list.
4. The rental record includes the rental date and the due date for returning the movies.

Perform the following tasks on the table given below:

- Identify all problems caused by the current table structure, including hints of repeating groups and dependencies.
- Break down the table into separate tables and normalize up to **3NF** to eliminate redundancy, ensure data integrity, and improve scalability.

RentalID	CustomerID	CustomerName	CustomerEmail	MovieIDs	MovieTitles	Genre	RentalDate	DueDate
1	C001	John Doe	john.doe@mail.com	M001, M002	Inception, Interstellar	Sci-Fi, Sci-Fi	2024-12-01	2024-12-05
2	C002	Jane Smith	jane.smith@mail.com	M003	The Godfather	Crime	2024-12-02	2024-12-06
3	C003	Sarah Johnson	sarah.johnson@mail.com	M001, M004, M005	Inception, Titanic, Avatar	Sci-Fi, Romance, Sci-Fi	2024-12-03	2024-12-07
4	C001	John Doe	john.doe@mail.com	M006	The Dark Knight	Action	2024-12-04	2024-12-08
5	C002	Jane Smith	jane.smith@mail.com	M007, M008	La La Land, Whiplash	Musical, Musical	2024-12-05	2024-12-09
6	C003	Sarah Johnson	sarah.johnson@mail.com	M009	Avengers: Endgame	Action	2024-12-06	2024-12-10
7	C004	Robert Brown	robert.brown@mail.com	M001, M002	Inception, Interstellar	Sci-Fi, Sci-Fi	2024-12-07	2024-12-11
8	C002	Jane Smith	jane.smith@mail.com	M003, M010	The Godfather, Pulp Fiction	Crime, Crime	2024-12-08	2024-12-12
9	C005	Alice Taylor	alice.taylor@mail.com	M011	Parasite	Drama	2024-12-09	2024-12-13
10	C003	Sarah Johnson	sarah.johnson@mail.com	M012	Frozen	Animation	2024-12-10	2024-12-14