**Data Management**

**September 2024 Examination**

**Q1. Discuss the concept of normalization in relational databases and discuss why it is important in database design. Describe the different types of normal forms used in the normalization process with appropriate example. (10 Marks)**

**Ans 1.**

**Introduction**

Normalization is a fundamental concept in relational database design, essential for optimizing database structure to reduce redundancy and improve data integrity. By distributing fields among multiple tables and establishing relationships through foreign keys, normalization ensures that the database is efficient and maintains consistency. This process not only helps in organizing data efficiently but also in achieving the scalability and flexibility required in managing large databases. This discussion will delve into the concept of normalization, explaining its significance and detailing the various normal forms with practical examples. Through this

It is only half solved

Buy Complete from our online store

<https://nmimsassignment.com/online-buy-2/>

NMIMS Fully solved assignment available for**session SEP 2024,**

your**last date is 29th August 2024**.

Lowest price guarantee with quality.

Charges**INR 350 only per assignment.**For more information you can get via mail or Whats app also

Mail id is [aapkieducation@gmail.com](mailto:aapkieducation@gmail.com)

Our website [www.aapkieducation.com](http://www.aapkieducation.com/)

After mail, we will reply you instant or maximum

1 hour.

Otherwise you can also contact on our

Whatsapp no OR Contact no is +91 8755555879

**Q2. Imagine you are working as a data analyst at a company that specializes in customer insights. Your team is tasked with analysing a large dataset containing information on customer purchases, demographics, and product reviews.**

**Given this scenario, how would you use R programming to effectively manage and analyse this dataset? Discuss which data structures in R you would use to store and manipulate the data, such as vectors, lists, data frames, and matrices. Explain why you would choose each of these structures, and describe specific R functions or techniques that would be helpful in this context.**

**Additionally, consider a situation where you need to create visualizations and summary statistics from this dataset. Which R packages and tools would you recommend, and how would they assist in generating insights for your team's analysis? (10 Marks)**

**Ans 2.**

**Introduction**

As a data analyst working with a company focused on customer insights, the ability to efficiently manage and analyze large datasets is crucial. R programming offers a robust environment for handling and analyzing such data, thanks to its extensive range of data structures and analytical tools. This discussion will explore how to use R to store, manipulate, and analyze a large dataset containing customer purchases, demographics, and product reviews. We will examine the specific data structures in R—such as vectors, lists, data frames, and matrices—that are best suited for different types of data. Additionally, we will discuss the R packages and

**Q3. Given a database schema with four tables: emp, proj, dept, and workson, each having specific attributes.**

**emp (eno , ename, bdate, title, salary, dno)**

**proj (pno , pname, budget, dno)**

**dept (dno , dname, mgreno) workson (eno , pno Questions: , resp, hours)**

**[Please Note :**

**emp: Contains employee data, including employee number (eno), employee name (ename), birthdate (bdate), title (title), salary (salary), and department number (dno).**

**proj: Contains project data, including project number (pno), project name (pname), budget**

**(budget), and department number (dno).**

**dept: Contains department data, including department number (dno), department name**

**(dname), and manager employee number (mgreno).**

**workson: Contains information about employees working on projects, including employee number (eno), project number (pno), responsibility (resp), and hours worked (hours). ]**

**a) Create emp, proj, dept, and workson tables as per above schema in MSQL and add any 5 dummy records using Insert into command. (5 Marks)**

**Ans 3a.**

**Introduction**

In a relational database management system like MySQL, creating tables according to a specific schema is a foundational task that sets the stage for storing and managing data. This exercise involves creating four tables—emp, proj, dept, and workson—based on the provided schema. These tables will be structured to store detailed information about employees, projects, departments, and the specific responsibilities of employees working on various projects. After creating the tables, we will insert five dummy records into each table to populate them with

**b) Perform following SQL queries on above tables. (5 Marks)**

**i. Find the project number and name for projects with a budget greater than 100,000.**

**ii. Return all records from workson where the hours worked are less than 10 and the responsibility is 'Manager'.**

**Ans 3b.**

**Introduction**

Executing SQL queries on a well-structured database is a crucial task for retrieving specific information based on defined criteria. In this exercise, we will perform two SQL queries on the previously created tables (emp, proj, dept, and workson). The first query will focus on extracting project details where the budget exceeds 100,000, while the second query will filter records from the workson table based on the number of hours worked and the specified responsibility. These queries demonstrate how SQL can be used to extract meaningful insights from a relational