ITMD 523	Advanced Topics in Data Management	HW 3

Student Name	**Sivaranjani Prabasankar**	Section
Instructor	***Luke Papademas***	Due Date 3/2/2019

Part	1	2	3	4	TOTAL	Score
Maximum Points	25 points	25 points	25 points	25 points	**100** points	

Textbook Reading Assignment   Thoroughly read Week 1 - 4 course lecture notes.

Part 1	**Concept Check - Advanced Topics in Data Management**

(1)   (Data Warehousing: OLAP versus OLTP)

An information system category that prioritizes transaction processing, which deal with operational data is Online Transaction Processing (OLTP) . Another type of information system category is Online Analytical Processing (OLAP) .

OLAP concentrates on performing analytical processing and OLTP focuses on providing transactional processing.

View the video production at the link shown below (or similar video) and list <u>five</u> facts that you learned concerning these two information systems.

https://www.youtube.com/watch?v=q5f1xOA9IQQ

S. No	OLAP	OLTP
Abbreviation	OnLine Analytical Process	OnLine Transaction Process
Data Type	Historical or Archived Data	Real Time Data
Volume of Data	High	Low
Usage	Involves complex calculation	Involves short online transaction
Inserts and Updates	Batch run process	Can be done by end users
Queries	Complex involving aggregations	Simple
Database Design	Less joins and de-normalization used	More joins and Normalization used
Effectiveness Measure	Response time	Fast query process
Usage	Used in Data mining Techniques	To store transactional databases is the entity mode
Data Type	Analytical Data	Operational Data
Performance Measure	Query throughput	Transaction throughput
Optimization	Many indexes for performance optimization	Few indexes for performance optimization

(2)   (Data Science and Business Intelligence: OLAP)

OLAP focuses on multi - dimensional views and analysis of data queries. For example, we can query data from a standpoint of examining company sales by region, year, sales amount and salesperson.

The typical OLAP operations will include: (1) drill - down, (2) dice and slice, (3) rotate or pivot and (4) roll - up or consolidation

Student Name **Sivaranjani Prabasankar**   Section _____

Using one or more of the business or organizational entities listed below discuss a multi - dimensional analysis that can be performed on the entities with any or all of the above operations.

Business / Organizational Entities

- a global parcel shipping company
- a nationwide college or university
- a talent agency
- a sports memorabilia firm
- an office supply superstore
- a vitamin and health shop

You can review the information at the link below, which summaries the OLAP operations.

http://www.cs.ccsu.edu/~markov/ccsu_courses/DataMining-2.html

---

OLAP in Business Intelligence

OLAP – Online Analytical Processing is a section of database processing that has its implications on business intelligence. It enables analysts, Managers and other executives with the processed information which is required to make strategic and calculated decisions about their organization. Multi-dimensional OLAP (MOLAP) uses array-based multidimensional storage engines for multidimensional views of data. With multidimensional data stores, the storage utilization may be low if the data set is sparse.

Let us consider **a Nationwide college or university** for our discussion.

1) Drill Down

Drill-down is disperse or fraction of a data cube. It is performed by either of the following ways

- By stepping down a concept hierarchy for a dimension
- By introducing a new dimension.

Example: Drill down college admission to get insights about admission in each Department.

2) Roll Up

Roll-up or consolidation performs aggregation on a data cube in any of the following ways −

- By climbing up a concept hierarchy for a dimension
- By dimension reduction

Example: Roll up no. of. students graduated in each department to get insights about overall graduation percentage.

3) Slice and Dice

The slice operation selects one dimension from a given cube and provides a new sub-cube.
Dice selects two or more dimensions from a given cube and provides a new sub-cube.

Example: Slice and Dice performance of students in each department for various semester to get insights about overall performance.

Student Name   **Sivaranjani Prabasankar**   Section

> **4)   Pivot**
> The pivot operation is also known as rotation. It rotates the data axes in view to provide an alternative presentation of data.
> **Example:** **Pivot of student financial aid for every semester combining with course details to get finalize fee.**

(3)   (Data Maintenance: Scrubbing Data in an MS Access Database Table)

You have been assigned the task of scrubbing an MS Access database table, which is used for data maintenance purposes. Explain how you would use MS Access to perform each of these scrubbing tasks.

(a)   Locate any fields that contain two spaces and replace the two spaces with one space.

(b)   Locate any duplicate records.

(c)   Eliminate any records that have a particular field having a number value.

(d)   Eliminate any records that have a field having an alpha - numeric value.

(e)   Update any records that have any non - printing characters.

You can create your own sample table and then show explain or how the table is scrubbed.

Initial Table Before Querying

(A) Locate any fields that contain two spaces and replace the two spaces with one space.

SQL Query

SELECT TY2018.ID, TY2018.State, TY2018.Revenue, TY2018.[Street Name],
REPLACE(TY2018.[Street Name],' ',' ') AS Street
FROM TY2018;

Screenshots

(B) Locate any duplicate records

SQL Query

SELECT First(TY2018.[ID]) AS [ID Field], First(TY2018.[State]) AS [State Field], First(TY2018.[Revenue]) AS [Revenue Field], First(TY2018.[Street Name]) AS [Street Name Field], Count(TY2018.[ID]) AS NumberOfDups
FROM TY2018
GROUP BY TY2018.[Street Name]
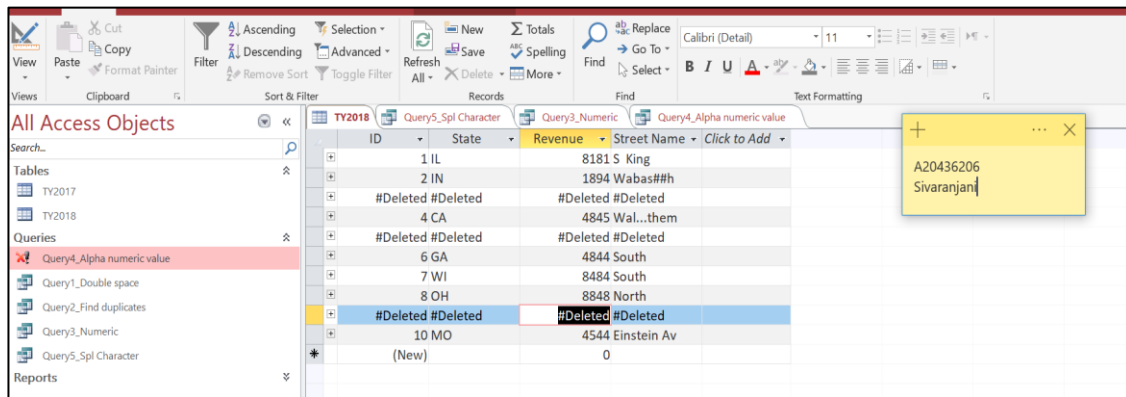HAVING ((Count(TY2018.[Street Name]))>1);

Screenshots

**(C) Eliminate any records that have a particular field having a number value**

SQL Query

DELETE TY2018.ID, TY2018.State, TY2018.Revenue, TY2018.[Street Name]
FROM TY2018
WHERE (((TY2018.[Revenue]) Like "4814"));

Student Name   **Sivaranjani Prabasankar**   Section

Screenshots

(D) Eliminate any records that have a field having an alpha - numeric value

SQL Query

DELETE TY2018.ID, TY2018.State, TY2018.Revenue, TY2018.[Street Name]
FROM TY2018
WHERE (((TY2018.[Street Name]) Like "*2*"));

Screenshots
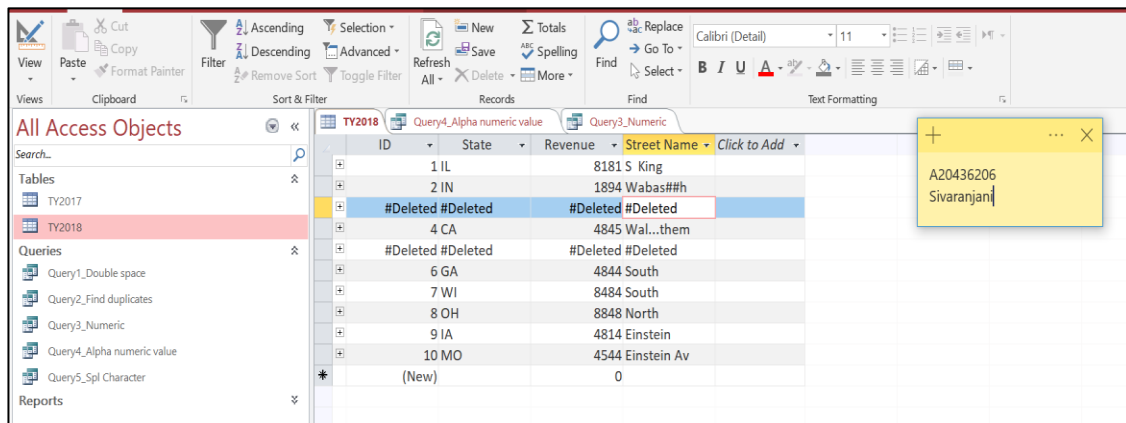
(E) Update any records that have any non - printing characters

SQL Query

UPDATE TY2018 SET TY2018.[Street Name]='UPDATED'
WHERE (((TY2018.[Street Name]) Like "*[^!-~.]*"));

Student Name   **Sivaranjani Prabasankar**   Section

Screenshots
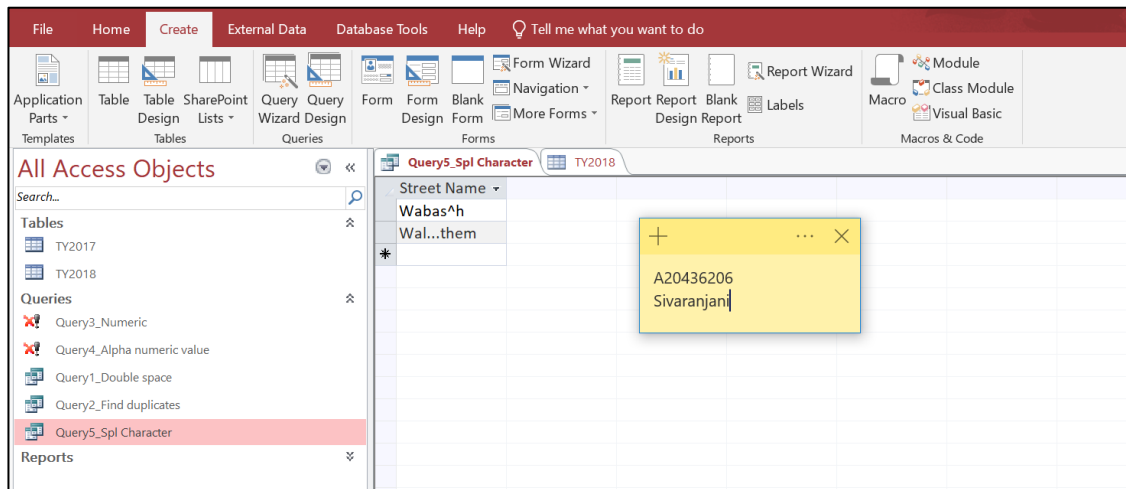
Final Table After Querying
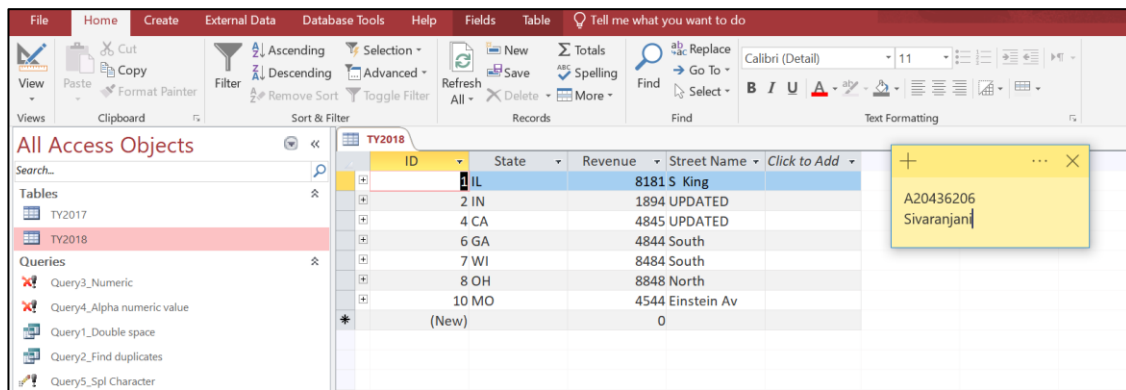
Part 2   DBMS Concepts - Advanced Topics in Data Management

(1)   (Entity Relationship Diagrams - ERDs)

An entity relationship diagram (ERD) depicts the conceptual database as viewed by end user. An ERD also illustrates the database's main components: Entities, Attributes, Relationships

eir attributes and their relationships.

Orders {OrderID, OrderDate, CustomerID, ShippingDate, OrderStatus}

OrderDetails {OrderID, ProductID, Quantity, LineNumber}

Products {ProductID, Name, Description, Quantity, UnitPrice}

Student Name   **Sivaranjani Prabasankar**   Section _____

Without writing any SQL code statements, discuss <u>five</u> separate queries that would be useful to ascertain information regarding this database schema.

Attributes and Relationship
Table 1: Order: To hold the details of overall orders placed
  Attributes:
    Order ID, Order Date, Customer ID, Shipping Date, Order Status
  Relationship:
    Order table has 1 to M relationship with Order details table
    Order table has 1 to M relationship with Product table

Table 2: Order Details: To hold the details of individual order
  Attributes: Order ID, Product ID, Quantity, Line Number
  Relationship:
    Order details table has 1 to M relationship with Order table
    Order details table has 1 to M relationship with Product table

Table 3: Products To hold the details of individual products
  Attributes: Product ID, Name, Description, Quantity, Unit Price
  Relationship:
    Product table has 1 to M relationship with Order details table
    Product table has 1 to M relationship with Order table

Using this schema, we can query for following information:
⇒ The best-selling product for a date range (Using Order Date).
⇒ Total preparation time taken by products owners (Using Order Date and Shipping Date).
⇒ The products that are ordered by a customer (Using Customer ID).
⇒ The total amount of an order made by customer (Counting Order id for each Customer ID)
⇒ Number of orders cancelled in a month (Using Order Status).

---

(2)   (Types of Relationships)

When designing an ERD , what is meant by a weak relationship? Provide an example of such a relationship.

Weak Entity
  ➢ A weak entity is simply an entity where its existence depends on another entity.
  ➢ Weak entity doesn't have the primary key instead it has a partial key that uniquely discriminates the weak entities.
  ➢ It does not have enough attributes to form a primary key and depends on another identifying entity or owner entity set to become meaningful.
  ➢ The primary key of a weak entity is a composite key formed from the primary key of the strong entity and partial key of the weak entity.

Weak Relationship
  ➢ Weak Entity Relationship is a relationship between a weak entity and another entity, it's sketched using double lines.
    Example: Dependents details for Employee
      ⇒ You can't logically have dependent (son, daughter etc.) with the absence of the employee table.

Student Name **Sivaranjani Prabasankar**   Section _____

| Part 3 | Data Models / Analytics - Advanced Topics in Data Management |

(1)   (Data Modeling)

Explain multivalued attributes with the help of examples. How are multivalued attributes indicated in the Chen Entity Relationship model?

Multi Valued Attributes

   A multivalued attribute can have more than one value at a time for an attribute. Or in other words an attribute is considered multi-valued if there can be many values associated with it at any one point in time.
   ⇒ An instance of a multivalued attribute can contain either a single value or multiple values.
   ⇒ Multivalued attributes often mask lists and collections.
   ⇒ Multivalued attribute is depicted by a dual oval in Chen ER Model

For ex.,
   1.   Degrees earned by employees in a business firm.
   2.   skills of a surgeon is a multivalued attribute since a surgeon can have more than one skill.
   3.   Addresses that a University might keep track of for a student: mailing address, home address, etc.
   Suppose someone proposes to track the university degrees earned by employees with an attribute named empDegree, then empDegree attribute could have none, one, or several values, we say the attribute is multi-valued. If you need to keep track of all the degrees that each person has obtained, then there would be multiple values to store.

---

(2)   (Data Analytics / Predictive Analytics: Aggregate Functions)

In SQL database systems such as Oracle and DB2 , aggregate functions, such as averages and other statistical measures, can be used for analytical purposes.
The following data illustrates company sales for the second half of the year.
List <u>five</u> generic aggregate functions that can be used with the following data.
Comment on their usefulness in describing the data.

Table: tblSales

index	month	sales
1	July	$16,000
2	August	$17,280
3	September	$23,660
4	October	$17,493
5	November	$28,805
6	December	$32,470

Student Name   **Sivaranjani Prabasankar**   Section _____

These data represent company sales for 6 months of a year. The following functions could be used to retrieve information.

- AVG (sales) – calculates the average of sales for 6 months.
- COUNT (sales) – counts rows with sales entry.
- MIN (sales)– gets the minimum value of sales in 6 months.
- MAX (sales)– gets the maximum value of sales in 6 months.
- SUM (sales)– calculates the sum of sales for 6 months.

Part 4   Data Design Concepts - Advanced Topics in Data Management

(1)   (Database Models)

What is a ternary relationship? Provide some business rules examples that specify the need for a ternary or higher - order relationship.

A **ternary relationship** is when three entities participate in the relationship. It is the is a relationship of degree three.

For Example: Consider a Mobile manufacture company. Three different entities involved are:

1. Car - Manufactured by company.
2. Part - Mobile Part which company get from Supplier.
3. Supplier - Supplier supplies Mobile parts to Company.

Car, Part and Supplier will participate simultaneously in a relationship.
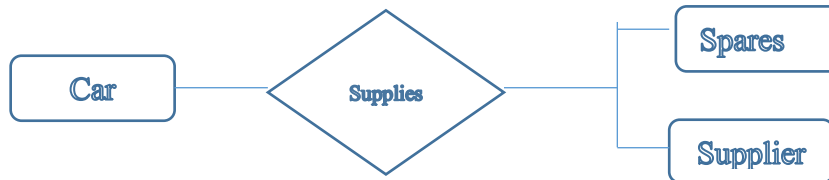
For Example: Consider a Medication example. Three different entities involved are:

1. A DOCTOR writes one or more PRESCRIPTIONs.
2. A PATIENT may receive one or more PRESCRIPTIONs.
3. A DRUG may appear in one or more PRESCRIPTIONs.

It is possible to replace a ternary or n-ary or higher order relationship by a collection of binary relationship connecting pairs of the original entities. But, in many cases it is hard to replace ternary relationship with two or more binary relationships because some information could be lost. This could lead to data inconsistency.

(2)   (Data Models and eCommerce Web Databases)

Visit the home page of the Barnes and Noble Web site and examine especially the links towards the bottom portion of the page.

Student Name   **Sivaranjani Prabasankar**   Section _____

https://www.barnesandnoble.com

B&N Services	Shipping & Delivery	NOOK	About Us	Quick Help
Advertise	About Free Shipping	NOOK Tablets & eReaders	About B&N	Customer Service
Affiliate Program	About Shipping	NOOK Mobile Apps	Investor Relations	Order Status
Publisher & Author Guidelines	Shipping Rates		Barnes & Noble, Inc.	Easy Returns
Bulk Order Discounts			Careers at B&N	Product Recalls
B&N Membership				
B&N Mastercard				
B&N Kids' Club				
B&N Educators				
B&N Bookfairs				

Then choose one of these categories, such as Shipping Rates or B&N Membership, and discuss how database management plays a role in connecting the Web site with the customers of the company.

Barnes & Noble website is an online book store.

At first, the store needs to maintain all kinds of data about a customer. The type of data includes the customer personal and payment details, books or ordered, subscription associated, membership details, offers allowed to the customer, etc.

On clicking **Membership section**, we can perform following operations

o   Sign up
   ▪   **Database Usage:**
      Create customer record with User id, Password in user table
      Update personal details in Customer table
      Update payment details in payment table
      Update account details in transaction table

o   Sign In
   ▪   **Database Usage:**
      Validate user id against password from user table
      Update password if required in user table

o   Renew Membership
   ▪   **Database Usage:**
      Update customer record in Customer table
      Update membership details with expiration date in Membership table
      Update payment details in payment table

o   Avail Discount
   ▪   **Database Usage:**
      Update offer details in offer table

Student Name   **Sivaranjani Prabasankar**   Section

- o   Exclusive offers
    - ▪ **Database Usage:**
      Display offer details from offer table

- o   Manage your profile
    - ▪ **Database Usage:**
      Display customer details, payment details, subscription detail from various table
      Update personal details in customer table
      Update payment details

- o   FAQ's
    - ▪ **Database Usage:**
      Update customer queries
      Update response time