Roll No&Name. ___Ehtisham Ali___11636_____

NATIONAL UNIVERSITY OF MODERN LANGUAGES ISLAMABAD

DEPARTMENT OF SOFTWARE ENGINEERING

Mid-Term Examination – Spring 2021 – BSSE 8 A & B (Morning)

Subject: ISDP - II **Instructor: Dr. Raheel Zafar/Ms.Benish Jamil**

Allowed Time: 1 hr 45 Minutes **Total Marks: 20**

Subjective Paper

Please attempt all questions.

Q 1): Please create Star Schema (Diagram) from the below mentioned tables. **(4 Marks)**

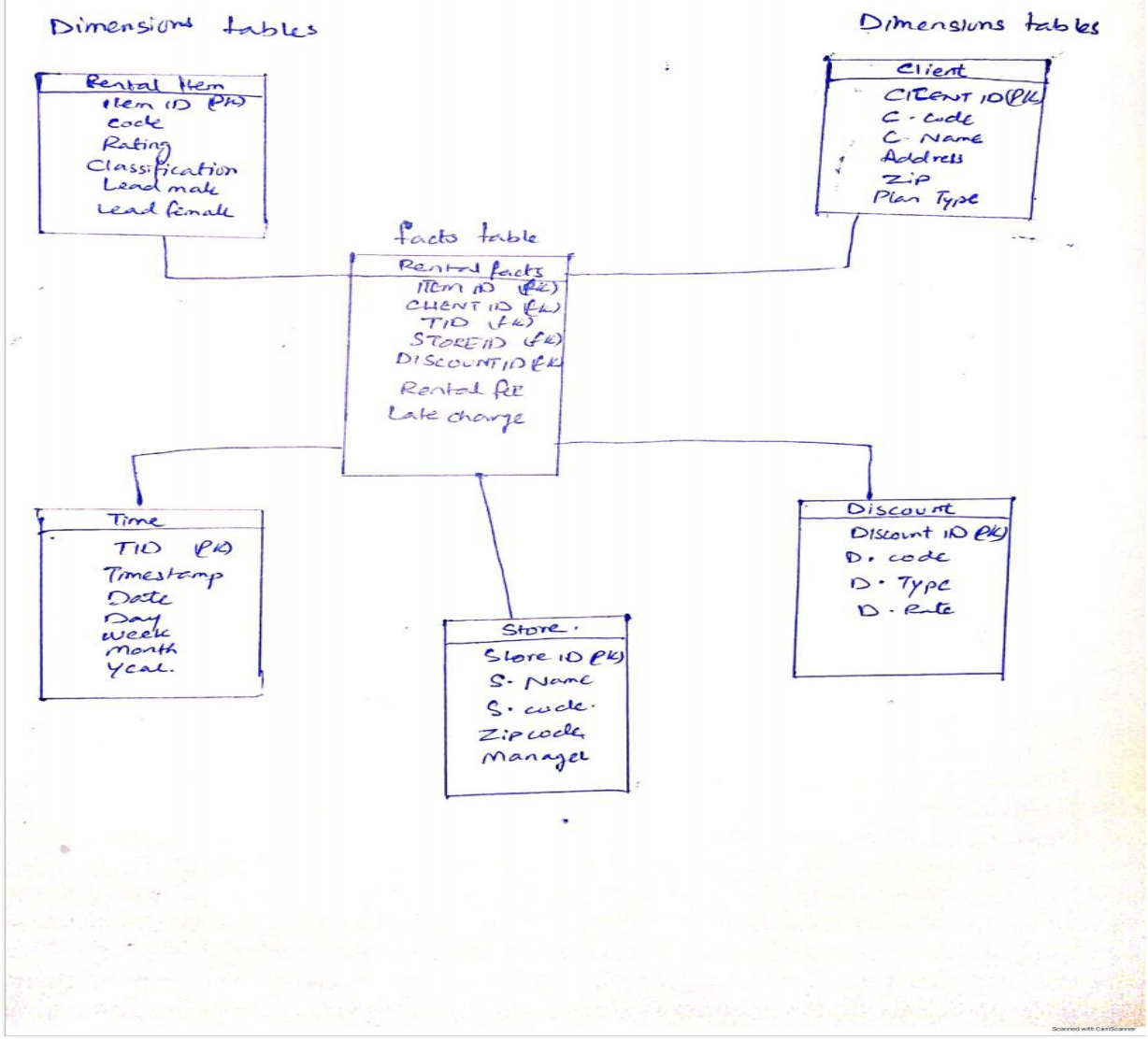
- Identify primary and foreign keys
- Create star schema properly showing linkages between dimension and fact tables.

STAR Schema: Video Rental

Table Name: RENTAL ITEM	**Table Name: CLIENT**	**Table Name: DISCOUNT**
1. Item ID 2. Item Code 3. Rating 4. Classification 5. Sub-classification 6. Director 7. Lead Male 8. Lead Female	1. Client ID 2. Client Code 3. Client Name 4. Address 5. State 6. Zip 7. Rental Plan Type	1. Discount ID 2. Discount Code 3. Discount Type 4. Discount Rate
Table Name: TIME	**Table Name: STORE**	**Table Name: RENTAL FACTS**
1. TID 2. Timestamp 3. Date 4. Day of Month 5. Week 6. Month 7. Quarter 8. Year	1. Store ID 2. Store Name 3. Store Code 4. Zip Code 5. Manager	1. Rental Fee 2. Late Charge

11636

Star schema Video Rental.



Q 2): How does a snowflake schema differ from a STAR schema? (2 Marks)

A. In snowflake schema data is normalized. They uses less space to store data but are more complex then star schema.

Q 3): What are aggregate fact tables? Why are they needed? Explain three ways of creating aggregate fact tables. (4 Marks)

A. Aggregate fact tables contain pre-calculated summaries derived from the most detailed fact table.

Uses: They are used to reduce the runtime processing.

They are created as a specific summarization across any number of dimensions.

Ways to create aggregate fact tables:

- **Transaction Fact Table**
  A transactional fact table is a fact table where each event is stored in the fact table only once. It has a date column indicating when the event occurred.
- **Periodic snapshot tables**
  A snapshot table holds the same raw, transactional data as its source in the transactional system, with additional fields for tracking the snapshot date.
- **Accumulating snapshot tables**
  Accumulating snapshot fact table summarizes the measurement events occurring at predictable steps between the beginning and the end of a process.

Q 4): Describe slowly changing dimensions. What are the three types? Explain each type briefly.

(4 Marks)

A. A Slowly changing dimension (SCD) is a dimension which stores and manages both current and historical data over time in a data warehouse.
TYPES:
There are 3 types of SCD.
SCD Type 1:

- SCD type 1 is used when there is no need to store historical data in the dimension table. (stores only current value).
- This method overwrites the old data in the dimension table with the new data.
- It is used to correct data errors in the dimension.

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name
12345	ABC922-Z	IntelliKidz	Education

Updated row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name
12345	ABC922-Z	IntelliKidz	Strategy

SCD Type 2:

- SCD type 2 stores the entire history the data in the dimension table.
- With type 2 we can store unlimited history in the dimension table.

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name	...	Row Effective Date	Row Expiration Date	Current Row Indicator
12345	ABC922-Z	IntelliKidz	Education	...	2012-01-01	9999-12-31	Current

Rows in Product dimension following department reassignment:

Product Key	SKU (NK)	Product Description	Department Name	...	Row Effective Date	Row Expiration Date	Current Row Indicator
12345	ABC922-Z	IntelliKidz	Education	...	2012-01-01	2013-01-31	Expired
25984	ABC922-Z	IntelliKidz	Strategy	...	2013-02-01	9999-12-31	Current

SCD Type 3:

- In type 3 method, only the current status and previous status of the row is maintained in the table.
- To track these changes two separate columns are created in the table.
- The type 3 method will have limited history and it depends on the number of columns you create.

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name
12345	ABC922-Z	IntelliKidz	Education

Updated row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name	Prior Department Name
12345	ABC922-Z	IntelliKidz	Strategy	Education

Q 5): How OLAP can be used for following applications?

i) Forecasting ii) Trend analysis iii) Tax planning (3 Marks)

A i) forecasting

We can create forecasts in Analytic Workspace Manager by defining a forecast step in a Calculation Plan. These are the steps for creating a forecast.

1. Creating the Forecast Time Periods
2. Creating a Forecast Measure
3. Selecting the Historical Data
4. Identifying the Levels for the Forecast
5. Creating a Forecast Step
6. Generating the Forecast Data
7. Evaluating the Forecast Results

ii) Trend analysis

OLAP in trend analysis is used to project both contemporary and future movement of events through use of time collection facts analysis which includes assessment of data over a sequential period of time to spot a pattern or trend.

iii) Tax palnning

OLAP takes a certain tax planning system, and checks the crucial frame, modules and technical difficulties or problems of the system.

Q 6): Explain any 5 tasks to be performed during Data Transformation in ETL process? (3 Marks)

A. Five tasks performed in data transformation are:

1. **Format changes:**
2. **Deduplication**
3. **Splitting up fields**
4. **Replacement of codes.**
5. **Aggregates**

1. **Format changes:**
   This task transforms the data format from one particular format or structure to another.
2. **Deduplication:**

This task transforms or basically remove the redundant or duplicate data.

3. **Splitting up fields:**
   This task splits up the field in multiple sperate fields. Like if a person has filed with first and last name then it will split first name field in another field and last name in another.

4. **Replacement of codes:**
   In data transformation code replacement is performed to change the code that the code generator produces for functions and operators to meet application code requirements.
5. **Aggregates:**
   This task in data transformation summaries the data by combining things so we can refer to them collectively. It reduces runtime and helps in achieve maximum business objectives.