**44-560 Adv Topics in DB Systems Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Exam 01 (100 points) KEY** *please print*

**Multiple choice (78 points – 2 points each).**  Select the **BEST** correct answer for each of the following. ***Write the letter corresponding to the BEST correct answer on your answer sheet.***

***SELECT ONLY ONE ANSWER FOR EACH QUESTION. IF YOU SELECT MORE THAN ONE ANSWER, THE ENTIRE QUESTION WILL BE COUNTED AS WRONG.***

1. Modeling a many-to-many relationship requires the introduction of an additional entity called a(n)
   1. linking entity
   2. join entity
   3. associative entity
   4. assistive entity
2. A \_\_\_\_\_ schema is usually the result of normalizing dimension tables.
   1. snowflake
   2. starflake
   3. mesh
   4. matrix
3. Consider the following table:

Movie (movieId, movieTitle, length, directorId, dirLName, dirFName)

Which of the following statements is true?

* 1. This table is in 3NF.
  2. This table is not in 3NF because it has a partial dependency.
  3. This table is not in 3NF because it has a transitive dependency.

1. In a dimensional model, dimension tables usually contain descriptive data.
   1. true
   2. false
2. \_\_\_\_\_ means that a foreign key matches the primary key value of some row in the referenced table, or is null.
   1. referential integrity
   2. entity integrity
   3. functional dependence
   4. relational integrity
3. When data is moved into a data warehouse, consistent naming conventions and formats must be used. This is referred to as data \_\_\_\_\_.
   1. volatility
   2. time-variation
   3. subject-orientation
   4. integration
4. Primary keys for dimension tables are usually composite keys, with each component of the primary key also being a foreign key into the fact table.
   1. true
   2. false
5. If there is a one-to-many relationship between two entities, it is necessary to create an additional entity in order to model the one-to-many relationship.
   1. true
   2. false
6. The level of detail in the fact table is referred to as the \_\_\_.
   1. dimension
   2. grain
   3. depth
   4. size
7. The conflicts between design efficiency, information requirements, and processing speed are often resolved through \_\_\_\_\_.
   1. conversion from 1NF to 2NF
   2. conversion from 2NF to 3NF
   3. compromises that include denormalization
   4. conversion from 3NF to 4NF
8. Data warehouse design uses the dimensional model, also referred to as a(n)
   1. ER model
   2. snowflake
   3. star schema
   4. conceptual model
9. The \_\_\_\_\_ of a relationship is the number of entities involved in the relationship.
   1. cardinality
   2. relationship number
   3. depth
   4. degree
10. The relational data model was developed in the \_\_\_\_\_.
    1. 1960s
    2. 1970s
    3. 1980s
    4. 1990s
11. Within a specialization hierarchy, every subtype can have \_\_\_\_ supertype(s) to which it is directly related.
    1. zero
    2. only one
    3. one or two
    4. many
12. The only normal forms are 1NF, 2NF, and 3NF.
    1. true
    2. false
13. In the grocery store model that we studied in class, the transaction dimension had no attributes to be stored except for the transaction number. Therefore, we stored the transaction number in the fact table, without linking it to any dimension table. Thus, we were treating the transaction dimension as a(n) \_\_\_\_\_\_\_ dimension.
    1. non-existent
    2. unnormalized
    3. empty
    4. degenerate
14. In a database context, a(n) \_\_\_\_\_ indicates the use of different names to describe the same attribute.
    1. entity
    2. duplicate
    3. synonym
    4. homonym
15. A relation with no repeating groups is said to be in \_\_\_\_\_.
    1. 1NF
    2. 2NF
    3. 3NF
    4. 4NF
16. \_\_\_\_ is an example of a non-additive fact
    1. bank balance
    2. daily high temperature
    3. dollar sales amount
    4. quantity sold
17. A \_\_\_\_\_ entity has a primary key that is partially derived from the parent entity in the relationship.
    1. strong
    2. weak
    3. business
    4. relationship
18. Which of the following is an example of structured data?
    1. a web page
    2. an e-mail
    3. a memo
    4. a spreadsheet
19. A table that is in 1NF and includes no partial dependencies is said to be in \_\_\_\_\_.
    1. BCNF
    2. 2NF
    3. 3NF
    4. 4NF
20. Assume we have the following dimensional model used for a data warehouse to store sales information for the DogsAndMore chain of stores:



Assume we are storing data for five years (5 \* 365 = 1,825 days). We initially have 100,000 customers and 50,000 items. The chain expects 10,000 new customers each year and 5,000 new items each year. (So after the first year, there will be 110,000 customers; after the second year, 120,000 customers, and so forth.) Each day, the company receives orders from approximately 1,000 customers, with each customer ordering, on average, two items.

A sample row in the fact table will look like this:

(“C1”, “I3”, “D5”, 35, 1035.33)

A row like this tells us that the customer with customerKey of C1 ordered 35 of the item with itemKey of I3 on the date with dateKey D5. Dollar amount for this sale was $1,035.33.

How many rows will the fact table contain after five years?

5 years \* 365 days / year \* 1,000 customer orders / day \* 2 items / customer order =

1825 \* 1,000 \* 2 items= 3,650,000 items: 3,650,000 rows in the fact table.

* 1. (100,000 + 50,000) \* 2 \* 1,000 \* 1,825
  2. 2 \* 1,000 \* 1,825 \* 100,000 \* 50,000
  3. 2 \* 1,000 \* 1,825
  4. 2 \* 1,825

1. Given the same scenario as in the previous problem, how large, in bytes, will the customer dimension be after five years? Assume each field is approximately 10 bytes in length.

100,000 customers initially.

10,000 new customers each year, so 50,000 additional customers after five years for a total of 150,000.

150,000 rows \* 5 fields / row \* 10 bytes / field = 7,500,000 bytes

* 1. 100,000
  2. 6,000,000
  3. 7,500,000
  4. 15,000,000

1. Which of the following is true for data warehouses?
   1. uses periodic data
   2. uses transient data
   3. is always fully normalized
   4. none of the above is true
2. Operational databases are designed for decision support.
   1. true
   2. false
3. In a relational table, each column has a specific range of values known as its \_\_\_\_\_.
   1. domain
   2. type
   3. universal set
   4. key
4. An attribute that is made up of two or more simpler attributes is called a
   1. complex attribute
   2. composite attribute
   3. constant attribute
   4. unified attribute
5. The attribute B \_\_\_\_\_ the attribute A if each value in column A determines one and only one value in column B.
   1. is logically dependent on
   2. is owned by
   3. functionally determines
   4. is functionally dependent on
6. Suppose we have two entities, A and B, with a 1:M relationship. For each instance of A, there are many instances of B associated with it. For each instance of B, there is only one instance of A associated with it. To represent this relationship, which of the following is true?
   1. You must introduce a third entity.
   2. A must contain a foreign key referencing B.
   3. B must contain a foreign key referencing A.
   4. Both A and B will contain a foreign key, referencing each other.
7. A Customer table’s primary key is custCode. The Customer primary key column has no null entries, and all entries are unique. This is an example of
   1. referential integrity
   2. entity integrity
   3. functional dependence
   4. relational integrity
8. \_\_\_\_ is an example of a semi-additive fact.
   1. bank balance
   2. daily high temperature
   3. dollar sales amount
   4. quantity sold
9. Fact and dimension tables are related by \_\_\_\_\_ keys.
   1. shared
   2. primary
   3. foreign
   4. linked
10. Attributes are types of entities.
    1. true
    2. false
11. Consider the following table:

Book (ISBN, bookTitle, publisherID, publisherName)

Which of the following statements is true?

* 1. This table is in 3NF.
  2. This table is not in 3NF because it has a partial dependency.
  3. This table is not in 3NF because it has a transitive dependency.

1. In a dimensional model, fact tables are normally 3NF.
   1. true
   2. false
2. Most operational databases use \_\_\_\_\_ data in which existing records are overwritten with new data.
   1. transient
   2. temporary
   3. periodic
   4. permanent
3. Data warehouses are built from operational databases using a set of processes known as
   1. CRN
   2. ETL
   3. LET
   4. SAP
4. \_\_\_\_\_ attributes can have several values
   1. Composite
   2. Simple
   3. Single-valued
   4. Multi-valued

***\*\*\*\* END OF MULTIPLE CHOICE QUESTIONS\*\*\****

1. (10 pts) Draw a Visio-style ER model for the following scenario. You must use Crow’s foot notation, and your model, through drawn by hand, must look like a Visio model, with the following exceptions:

* You do not have to indicate required attributes – so no bolding is necessary.
* You do not have to use dotted lines to indicate non-identifying relationships.

Note that you *must* indicate primary and foreign keys in the same way as in Visio.

**Scenario:** This database stores information about books, publishers, authors, and patrons who have checked out books from a library that has several branches.

**Entities and attributes:**

Book: For each book, store the ISBN number of the book (a unique identifier). In addition, store the title of the book and the date of publication.

Publisher: This entity represents a publishing company. For each publisher, store the publisher id (a unique identifier), the name of the publisher, and the location (city and state) of the publisher.

Author: This entity represents people who write books. For each author, store the author id (a unique identifier), and the name of the author.

Patron: This entity represents people who borrow books from a library branch. For each patron, store the patron id (a unique identifier) and the name of the patron.

Branch: For each library branch, store the branch id (a unique identifier) and the name of the branch.

**Relationships:**

Each book is written by at least one author. Each author has written at least one book. For each book and author combination, we store the royalty percentage (the percent of the sale price that the author will receive for each book sold).

Each book is published by exactly one publisher. Most publishers have published many books, but a new publisher may not have published any books.

A book may be checked out by many patrons from many different branches, but a few books may never be checked out by any one from any branch. At each branch, many books are checked out by many patrons, although a new branch many not have had any books checked out by any patron. Each patron has checked out at least one book from at least one branch. When a patron checks out a book from a branch, the date the book was checked out is stored, and the due date is stored.



1. (12 pts) Suppose we have this ER model for an operational database. The current data in each table is displayed below the model.



**Current data stored in the tables is shown here.**

| **customer** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **custNumber** | **custName** | **address** | **currBal** | **credLimit** | **repNum** |
| 124 | Sally Adams | 481 Oak, Lansing, MI | $418.75 | $500.00 | 3 |
| 311 | Don Charles | 48 College, Ira, MI | $200.10 | $300.00 | 12 |
| 522 | Mary Nelson | 108 Pine, Ada, MI | $49.50 | $800.00 | 12 |

| **parts** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **partNum** | **partDesc** | **category** | **unitsOnHand** | **warehouseNum** | **unitPrice** |
| ax12 | iron | appliance | 104 | 3 | $10.00 |
| az52 | skates | sporting | 20 | 2 | $25.00 |
| ba74 | baseball | sporting | 40 | 1 | $5.00 |
| bt04 | stove | appliance | 11 | 2 | $500.00 |
| bz66 | washer | appliance | 52 | 3 | $400.00 |
| ca14 | skillet | houseware | 2 | 3 | $5.00 |
| cb03 | bike | sporting | 44 | 1 | $100.00 |
| cx11 | mixer | houseware | 112 | 3 | $35.00 |
| cz81 | weights | sporting | 208 | 2 | $250.00 |

| **slsrep** | | | | |
| --- | --- | --- | --- | --- |
| **repNum** | **repName** | **repAddress** | **totComm** | **commRate** |
| 3 | Mary Jones | 123 Mina, Grant, MI | $2,150.00 | 0.05 |
| 6 | William Smith | 102 Raymond, Ada, MI | $4,912.50 | 0.07 |
| 12 | Sam Brown | 41 Harper, Ada, MI | $2,150.00 | 0.05 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **orderline**   | **orders** | | | | --- | --- | --- | | **orderNum** | **orderDate** | **custNum** | | 12489 | 9/2/1994 | 124 | | 12491 | 9/2/1994 | 311 | | 12498 | 9/5/1994 | 522 | | 12500 | 9/5/1994 | 124 | | 12504 | 9/5/1994 | 522 | | | |
| **orderNum** | **partNum** | **numOrdered** |
| 12489 | ax12 | 1 |
| 12489 | bt04 | 1 |
| 12489 | bz66 | 1 |
| 12489 | cb03 | 2 |
| 12489 | cx11 | 1 |
| 12489 | cz81 | 1 |
| 12491 | bt04 | 1 |
| 12491 | bz66 | 1 |
| 12498 | az52 | 2 |
| 12498 | ba74 | 2 |
| 12500 | bt04 | 1 |
| 12504 | cz81 | 1 |

A data warehouse has been constructed using the model shown below.



Data has been moved from the operational database to the data warehouse. Data for the dimension tables is shown below:

**Date Dimension**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dateKey | day | month | quarter | year |
| 1 | 02 | September | 3 | 1994 |
| 2 | 05 | September | 3 | 1994 |

**Customer Dimension**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| customerKey | custNumber | custLastName | custFirstName | Street | City | State | Zip |
| 1 | 124 | Adams | Sally | 481 Oak | Lansing | Michigan | 62222 |
| 2 | 311 | Charles | Don | 48 College | Ira | Michigan | 62223 |
| 3 | 522 | Nelson | Mary | 108 Pine | Ada | Michigan | 62222 |

**Part Dimension**

|  |  |  |  |
| --- | --- | --- | --- |
| partKey | partNum | partDesc | category |
| 1 | ax12 | iron | appliance |
| 2 | az52 | skates | sporting |
| 3 | ba74 | baseball | sporting |
| 4 | bt04 | stove | appliance |
| 5 | bz66 | washer | appliance |
| 6 | ca14 | skillet | houseware |
| 7 | cb03 | bike | sporting |
| 8 | cx11 | mixer | houseware |
| 9 | cz81 | weights | sporting |

**Your task:** Fill in the fact table on your answer sheet. ***Fill in only those rows for which the partKey has one of these values: 5, 7, 8, 9.***

**Sales Fact Table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **dateKey** | **customerKey** | **partKey** | **amount** | **quantity** |
| **1** | **1** | **5** | **400.00** | **1** |
| **1** | **1** | **7** | **200.00** | **2** |
| **1** | **1** | **8** | **35.00** | **1** |
| **1** | **1** | **9** | **250.00** | **1** |
| **1** | **2** | **5** | **400.00** | **1** |
| **2** | **3** | **9** | **250.00** | **1** |