6. One of the core tenets of dimensional modeling is that all measurement columns in a fact table must be at the same grain. Suppose that we want to capture two facts in a transaction fact table (or a periodic snapshot fact table). Which one of the following options violates the above core tenet?
   a. Requisition quantity and Dollar value
   b. Sales quantity and Unit price
   c. Quantity on hand and Dollar value at cost
   d. Period end balance amount and Year-to-date amount
   e. Period debit amount and Period credit amount

7. The simplest Inventory Periodic Snapshot Fact table involves only a single fact: quantity on hand. For most inventory analysis, however, quantity on hand alone is not enough. Quantity on hand needs to be used in conjunction with another fact to measure the velocity of inventory movement. What additional fact should be captured in each row of the Inventory Periodic Snapshot Fact table so that inventory analysts can compute the number of turns and number of day's supply?
   a. Minimum reorder quantity
   b. Quantity sold
   c. Unit price
   d. Dollar value at cost
   e. Dollar value at latest selling price

8. Select the best answer. Dimension role playing refers to which one of the following situations?
   a. A single dimension simultaneously appears several times in the same fact table. The underlying dimension exists as a single physical table, but we create multiple views from it. Each of the views represents a different role.
   b. A single dimension simultaneously appears several times in the same fact table. The underlying dimension exists as a single physical table, but we join this single dimension table to the fact table multiple times. Each of the join represents a different role.
   c. A single dimension simultaneously appears two times in the same fact table. We create two physical copies of the same dimension table. We then join each copy to the fact table. Each join represents a role.
   d. A single dimension simultaneously appears two times in the same fact table. We create two physical copies of the same fact table. We then join each copy to the same dimension table. Each join represents a role.
   e. None of the above

9. Which of the following is an appropriate guideline for including numeric values as attributes?
   a. If the number is used in calculations, place it in a fact table.
   b. If the attribute is used for grouping or filtering, create a limited number of categories or bands for the dimension table.
   c. Do not include the numeric quantity in the dimension table if there are many possible values.
   d. All of the above.
   e. A and C only.

10. Which of the following describes a best practice in dimensional modeling, according to the Kimball guidelines?
   a. Use natural keys to join facts and dimension tables
   b. Normalize tables for optimal query performance
   c. Use NULL values for missing attributes in dimension tables
   d. Group quantitative measures into a limited number of bands within fact tables.
   e. None of the above

11. Which of the following is true about factless-fact tables?
   a. The factless-fact table can be used to store all the possibilities of events that might happen. It is used in conjunction with an activity table that contains the events that did happen.
   b. The factless-fact table contains one transaction per primary key.
   c. The factless-fact table enables viewing of assignments between entities, such as sales representatives to clients, even when no transactions took place between them.
   d. All of the above
   e. Options (a) and (c) only

6. One of the core tenets of dimensional modeling is that all measurement columns in a fact table must be at the same grain. Suppose that we want to capture two facts in a transaction fact table (or a periodic snapshot fact table). Which one of the following options violates the above core tenet?
   a. Requisition quantity and Dollar value
   b. Sales quantity and Unit price
   c. Quantity on hand and Dollar value at cost
   d. Period end balance amount and Year-to-date amount
   e. Period debit amount and Period credit amount

7. The simplest Inventory Periodic Snapshot Fact table involves only a single fact: quantity on hand. For most inventory analysis, however, quantity on hand alone is not enough. Quantity on hand needs to be used in conjunction with another fact to measure the velocity of inventory movement. What additional fact should be captured in each row of the Inventory Periodic Snapshot Fact table so that inventory analysts can compute the number of turns and number of day's supply?
   a. Minimum reorder quantity
   b. Quantity sold
   c. Unit price
   d. Dollar value at cost
   e. Dollar value at latest selling price

8. Select the best answer. Dimension role playing refers to which one of the following situations?
   a. A single dimension simultaneously appears several times in the same fact table. The underlying dimension exists as a single physical table, but we create multiple views from it. Each of the views represents a different role.
   b. A single dimension simultaneously appears several times in the same fact table. The underlying dimension exists as a single physical table, but we join this single dimension table to the fact table multiple times. Each of the join represents a different role.
   c. A single dimension simultaneously appears two times in the same fact table. We create two physical copies of the same dimension table. We then join each copy to the fact table. Each join represents a role.
   d. A single dimension simultaneously appears two times in the same fact table. We create two physical copies of the same fact table. We then join each copy to the same dimension table. Each join represents a role.
   e. None of the above

9. Which of the following is an appropriate guideline for including numeric values as attributes?
   a. If the number is used in calculations, place it in a fact table.
   b. If the attribute is used for grouping or filtering, create a limited number of categories or bands for the dimension table.
   c. Do not include the numeric quantity in the dimension table if there are many possible values.
   d. All of the above.
   e. A and C only.

10. Which of the following describes a best practice in dimensional modeling, according to the Kimball guidelines?
   a. Use natural keys to join facts and dimension tables
   b. Normalize tables for optimal query performance
   c. Use NULL values for missing attributes in dimension tables
   d. Group quantitative measures into a limited number of bands within fact tables.
   e. None of the above

11. Which of the following is true about factless-fact tables?
   a. The factless-fact table can be used to store all the possibilities of events that might happen. It is used in conjunction with an activity table that contains the events that did happen.
   b. The factless-fact table contains one transaction per primary key.
   c. The factless-fact table enables viewing of assignments between entities, such as sales representatives to clients, even when no transactions took place between them.
   d. All of the above
   e. Options (a) and (c) only

...questions, circle the best answer.

1. What is the grain of the General Ledger Journal Entry Fact table?
   a. One row per day for every account in the chart of accounts.
   b. One row per week for every account in the chart of accounts.
   c. One row per ledger account
   d. One row per month for every account in the chart of accounts.
   e. One row for every general ledger journal entry transaction.

2. Select the best answer. For an expense budget line item, what is the grain of the Budget Fact table?
   a. Each row in the Budget Fact table represents the net change of the budget line item in an organizational cost center that occurred during the month.
   b. Each row in the Budget Fact table identifies what an organization in the company is allowed to spend for what purpose during a given time frame.
   c. Each row in the Budget Fact table is a snapshot of the current status of each line item in each budget each month.
   d. Each row in the Budget Fact table represents the net change of the budget line item that occurred during the month.
   e. Each row in the Budget Fact table represents the cumulative amount of the budget line item in an organizational cost center.

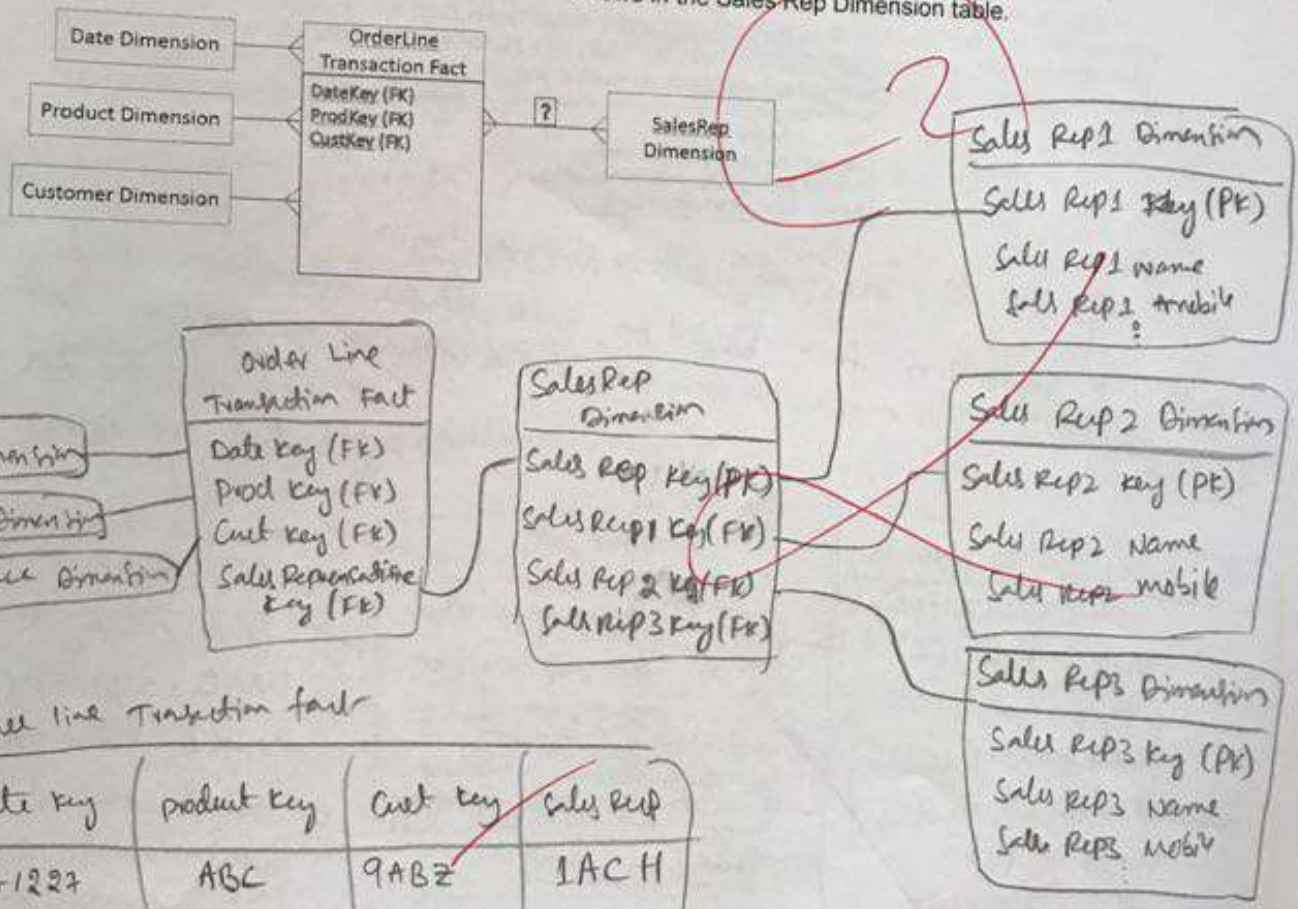
3. How to design an order line transaction fact table for a large multinational company with sale offices around the world? Note that this type of multinational company may be capturing order transactions in more than 15 different currencies.
   a. Include a column in the fact table for each currency.
   b. Include a single standard corporate currency column in the order line transaction fact table.
   c. Express each order line transaction fact amount in both the local transaction currency and the standardized corporate currency, such as U.S. dollars.
   d. Include a single local currency column in the order line transaction fact table.
   e. None of the above

4. Business users often need to know year-to-date (YTD) values in order to perform analysis. Where should the year-to-date values be stored?
   a. Year-to-date values are facts. They should be stored in fact tables
   b. Year-to-date values are description of facts. They should be stored in dimension tables
   c. Year-to-date values are not facts nor descriptive values. They should be stored in bridge tables
   d. Year-to-date values should not be stored in any table. They should be calculated in the Business Intelligence applications or OLAP cubes.
   e. None of the above.

5. Very large companies may have multiple ledgers arranged in an ascending hierarchy. At the lowest level, department ledger entries may be consolidated to roll up to a single division ledger entry. Then the division ledger entries may be consolidated to the enterprise level. How can you model this hierarchy in the General Ledger Snapshot Fact table?
   a. By adding a Hierarchical Ledger dimension
   b. By introducing a Parent Snapshot Key column in the fact table.
   c. By adding an explicit fact table surrogate key, a single column numeric identifier, which is incremented as you add rows to the fact table.
   d. Both options (a) and (b)
   e. Both options (b) and (c)

Essay Questions

1. (12.5%) Most of the time, every row in the fact table is related to a single row in the dimension table. For example, in a transaction fact table, each transaction line is related to exactly one product. Unfortunately, the real world is not always that simple. What happens when a fact may be related to one or more instances of a dimension? For example, more than one sales person can collaborate and work as a group on a single order line, as shown below. Assume that the percentage contribution of each sales person is known in this collaboration. In this case, the fact table cannot hold a foreign key reference to the sales rep dimension table since there maybe more than one sales person. How do you design a dimensional model in this situation?
   a. Show a dimensional model schema that illustrates your design. In your schema, show only the appropriate foreign keys and primary keys.
   b. Based on your new schema, give sample data that illustrate a single row in the OrderLine Transaction Fact table is related to two rows in the Sales Rep Dimension table.



(b)

Order line Transaction fact

Date key	product key	Cust key	Sales Rep
2017/1227	ABC	9ABZ	1ACH

Sales Representation Table

Sales Rep key	Sales Rep1 key	Sales Rep2 key	Sales Rep3 key
1ACH	1001	NA	2002

Sales Rep1 key.

1001	Albert	+1 9112923121
2001	Chris	+1 212 314 678

20. Select the best answer. The following Slowly Changing Dimension types are most relevant if you have been asked to preserve the historically accurate dimension attribute associated with a fact event, while supporting the option to report historical facts according to the current attribute values.
    a. Types 1 and 2
    b. Types 3 and 4
    c. Types 0, 1, and 2
    d. Types 3, 4, and 5
    e. Types 5, 6, and 7

21. Select the best answer. What Slowly Changing Dimension types modify the relevant fact table?
    a. Type 0
    b. Type 1
    c. None of the Slowly Changing Dimension types modify the relevant fact table.
    d. Type 3
    e. Type 4

22. In the order transaction fact table, each line item row includes the order number as a degenerate dimension. Why do we want to include this order number in the fact table?
    a. The order number enables us to group the separate line items on the order and answer questions such as "What is the average number of line items on an order?"
    b. The order number is occasionally used to link the data warehouse back to the operational database.
    c. The order number may play a role in the fact table's primary key.
    d. All of the above
    e. None of the above

23. When modeling the order line transaction fact table and its dimensions, do we want to have an order header dimension that joins to the fact table via the order number column (field)?
    a. Yes, because the order header dimension has customer descriptive data, which would be needed in the analysis of the order line transaction fact table.
    b. Yes, because the order header dimension is a replica of the order header table in the operational database system.
    c. No, because the order header dimension is likely very large.
    d. No, because we do not need an order number column (field) in the order line fact table.
    e. Yes, because this design enables us to group the separate line items on the order and answer questions such as "What is the average number of line items on an order?"

24. On a single order, there may be a shipping charge that applies to the entire items in the order. The dimensional modeler's first response should be to try to distribute this shipping charge to all items in the order. This procedure is broadly referred to as _____.
    a. Cost allocation
    b. Cost streamlining
    c. Cost replication
    d. Activity-based costing
    e. Cost drill down

25. What framework provides a rational approach to decomposing the enterprise Data Warehouse and Business Intelligence system?
    a. ETL
    b. The enterprise data warehouse bus architecture
    c. The independent data mart architecture
    d. The hub-and-spoke architecture
    e. The hybrid architecture

4. (12.5%) Assume that Product Key 12345 (Intellikidz) was initially assigned to Education department; and the original product dimension is as follows:

Original row in Product dimension:

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

Moreover, assume that Intellikidz product was introduced on January 1, 2014. On October 1, 2016, the marketing manager reassigned IntelliKidz to the Strategy department. You have been asked to redesign the Product dimension so that the new one preserves the historically accurate department reassignments for IntelliKidz (or another) product, while supporting the option to report historical facts according to the current IntelliKidz department.

a. What Slowly Changing Dimension Type should be used?
b. Show the new design for the Product dimension table, including two sample rows. The first row records the original IntelliKidz product when it was assigned to the Education department and then it expired on September 30, 2016; and the second row record the IntelliKidz product after it was reassigned to the Strategy department on October 1, 2016; and until now, IntelliKidz is still with the Strategy department.

a) We will use Type 6 SCD Here as per Requirements

b)

Product key	SKU (NK)	product description	Historic Dept name	Current dept name	Row effective	Row Expiration Date	Current row Indicator
1A1236	9A1B-Y	Intellikidz	Education	Strategy	2014-01-01	2016-09-30	Expired
2AB2C7	9A1B-Y	Intellikidz	Strategy	Strategy	2016-10-01	9999-12-31	Current

→ Here, we can see the historic data from History Department name table and can see effective and expired dates to check the given product effective dates. we have added new attributes to represent these.

3. (12.5%) Suppose that you work as a dimensional data modeler for a large nation-wide retailer. You are asked to design a dimensional data model to analyze daily quantity-on-hand inventory levels by product and store. You follow Kimball's four-step dimensional process to design the schema.

a. What is the business process that you are going to model?

b. What is the grain of your fact table?

c. What are the dimensions for this model?

d. What is the fact (numeric measure) that you are going to capture?

e. Draw a sample schema to illustrate your dimensional data model.

f. To enhance the inventory analysis, the retailer wants to store the retail cost and minimum reorder quantity for each product in the data warehouse. Assuming that the cost and minimum reorder quantity varies for a product by store, where would you store the retail cost and minimum reorder quantity data items in your dimensional data model? Draw a new schema that includes the retails cost and minimum order quantity data items.

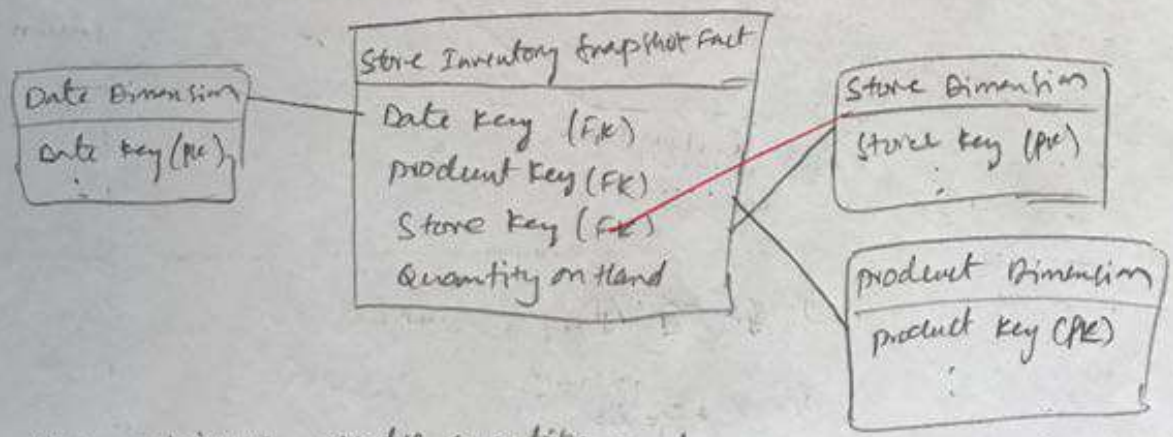
a) Business process:- periodic Snapshot of Retail Store Inventory

b) grain:- Daily Inventory Level of each product Per store

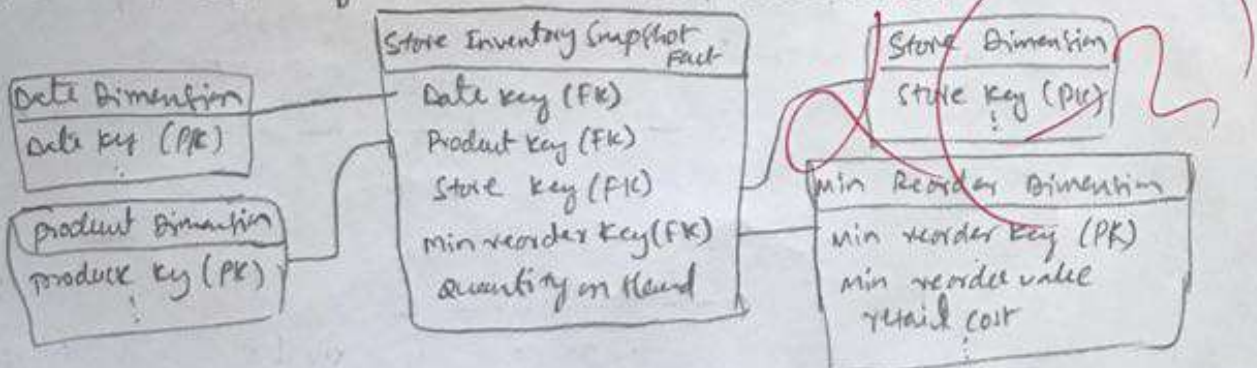
c) Dimensions:- product, store and Date

d) Fact :- quantity on Hand

e)



f) If the minimum reorder quantity and retail cost values for a product by store, then we can include that in any of the dimension table or even in Inventory Fact table. Here, I have added a new dimension table.



12. Which of the following is an appropriate guideline for designing dimensions or fact tables?
    a. Combine items that are highly correlated into the same dimension.
    b. Rather than using natural keys, we should create surrogate keys for dimension tables.
    c. To facilitate partitioning, the primary key of a date dimension can be more meaningful, such as YYYYMMDD instead of a non-meaningful surrogate key.
    d. All of the above
    e. Options (a) and (c) only

13. Which of the following is true of "degenerate dimensions"?
    a. They have no associated dimension table.
    b. They can serve as useful grouping keys for pulling together related transactions.
    c. They are inherently a natural choice for use as primary keys in fact tables.
    d. All of the above.
    e. Options (a) and (b) only.

14. Suppose that we want to add a new dimension, say Frequent Shopper dimension, to an existing schema. As a dimensional modeler, what should you do after you create this new dimension table?
    a. Add two rows in this new dimension table: one row is to represent "Frequent Shopper Not Identified" and another row is to represent "Prior to Frequent Shopper Program".
    b. Add a foreign key in the fact table to relate the fact table to this new dimension table.
    c. On each of the old fact table rows, insert a default foreign key that corresponds to a "Prior to Frequent Shopper Program" dimension row.
    d. All of the above
    e. Options (a) and (b) only

15. Which one of the following transactions is not a procurement transaction?
    a. Purchase requisition
    b. Purchase order
    c. Package product for shipment
    d. Shipping notification
    e. Payment

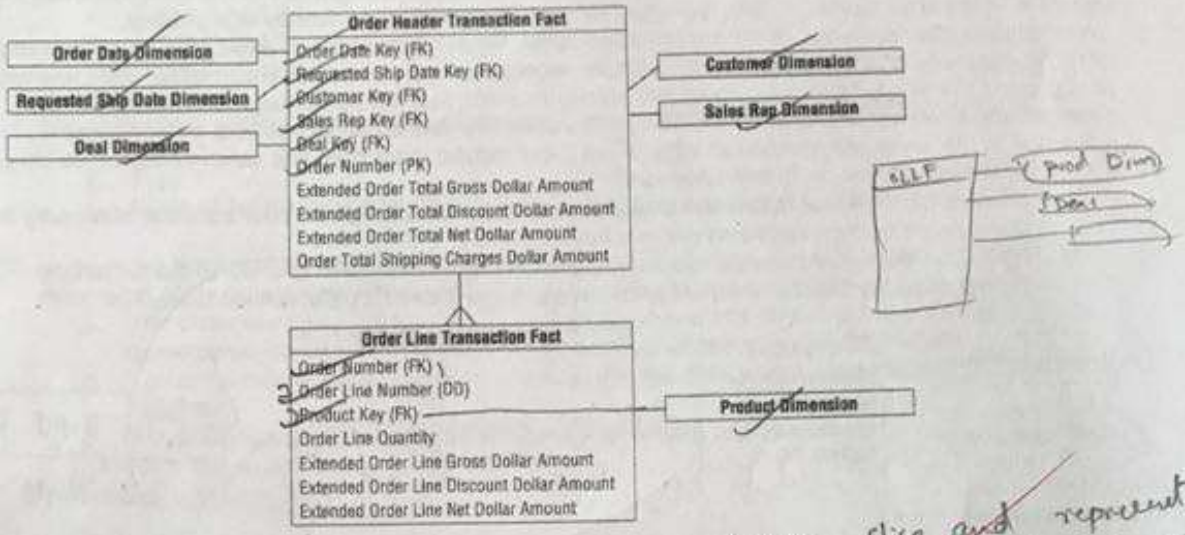
16. In the following list of dimensions, which one is not a dimension in procurement business processes?
    a. Account dimension
    b. Date dimension
    c. Product dimension
    d. Vendor dimension
    e. Employee dimension

17. In the following list, what item is implemented in a multi-dimensional database?
    a. OLAP Cubes
    b. Star-schema
    c. Pivot table
    c. All of the above
    d. Options (a) and (b)

18. If a business firm is interested in monitoring product movement as it proceeds through the procurement pipeline, what type of fact table does it need?
    a. Procurement transaction fact table
    b. Procurement periodic snapshot fact table
    c. Procurement accumulating snapshot fact table
    d. Procurement factless-fact table
    e. Procurement aggregate fact table

19. Select the best answer. In order management dimensional modeling, what is the primary reason to construct factless-fact table for sales rep assignments to customers?
    a. To provide a complete map of the historical assignments of sales reps to customers, even if some of the assignments never resulted in a sale.
    b. To capture the varying relationship between sales reps and customers over time.
    c. To indicate that sales rep and customer dimensions may participate independently in other fact tables.
    d. To signify that corporate business thinks sales rep and customer as separate things.
    e. To implement the many-to-many relationship between sales reps and customers.

2. (12.5%) Consider the following schema. Explain why this schema was not designed properly. Show a better new schema that is equivalent to this schema.



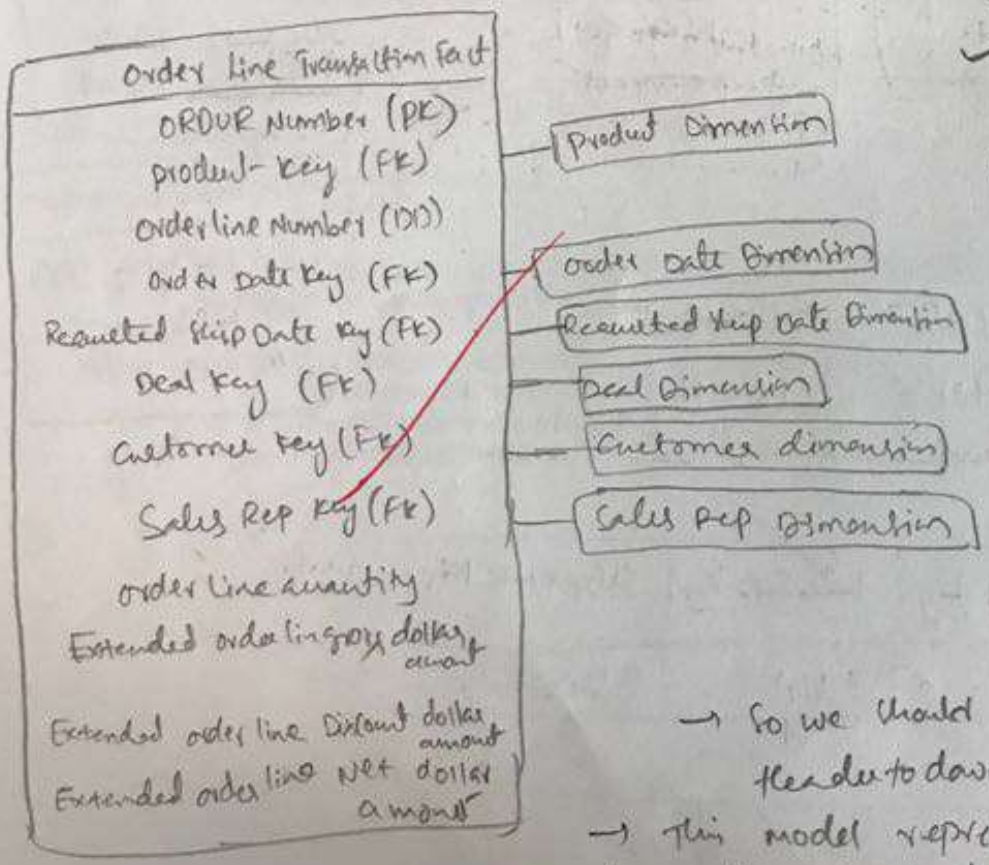
In the given schema, if we has to slice and represent order line transaction fact based on order header transaction fact, then it required to traverse on very large order header transaction fact.

the order header virtually replicated in order line Transaction Dimension

✓ Header dimension Contain all the information from Line F

✓ The order header is lik large, especially relative to the Line Fact table

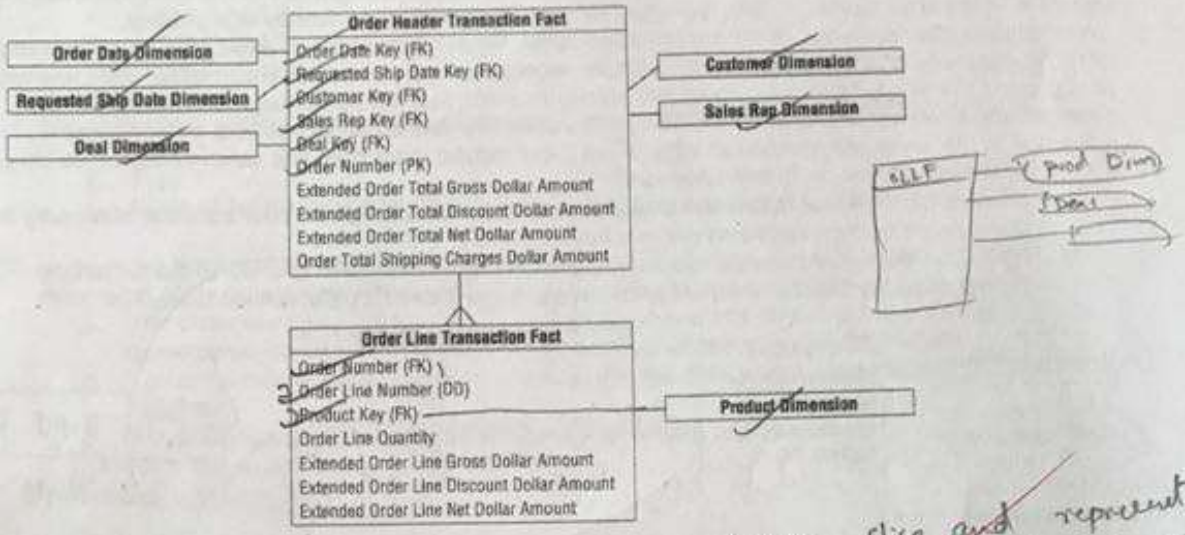
✓ with this design you would add 1 row to the dimension table, and the average of line rows to the fact table for every new order



→ So we should bring the dimensionality of header to down the order line

→ This model represents data ~~from~~ relationship from the order header / source systems. But we abonded the operational metality surrounding header f

2. (12.5%) Consider the following schema. Explain why this schema was not designed properly. Show a better new schema that is equivalent to this schema.



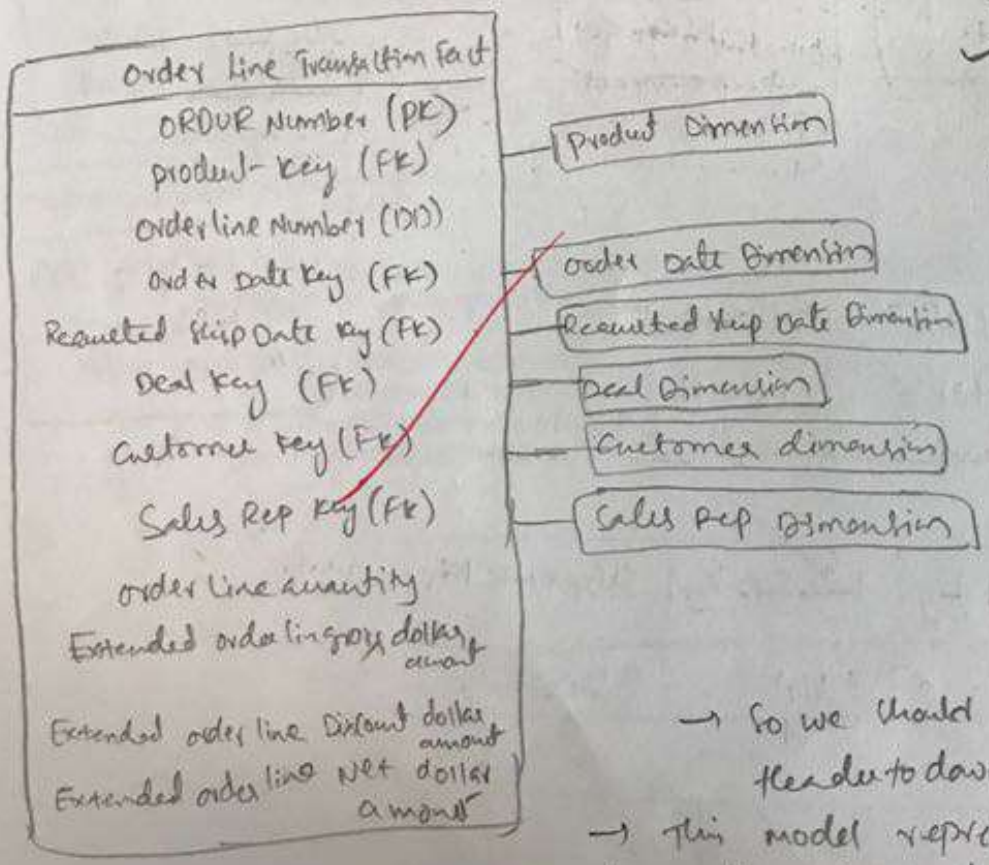
In the given schema, if we has to slice and represent order line transaction fact based on order header transaction fact, then it required to traverse on very large order header transaction fact.

the order header virtually replicated in order line Transaction Dimension

✓ Header dimension Contain all the information from Line F

✓ The order header is lik large, especially relative to the Line Fact table

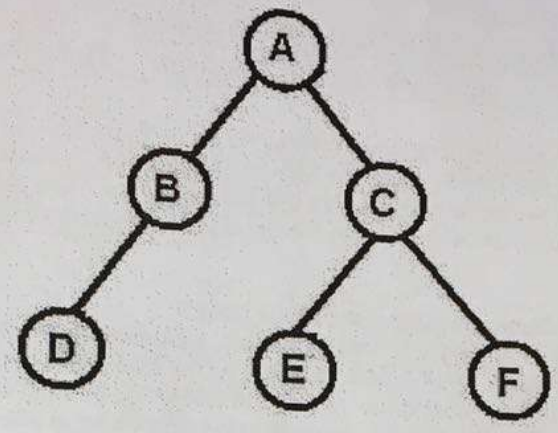
✓ with this design you would add 1 row to the dimension table, and the average of five rows to the fact table for every new order



→ So we should bring the dimensionality of header to down the order line

→ This model represents data ~~from~~ relationship from the order header / source systems. But we abonded the operational metality surrounding header f

28. (12.5%) Suppose that a large organization has the following rollup structure (see the diagram below). Create a map bridge table for this organization that shows sample rows. The grain of this bridge table is each path in the tree from a parent to all the children below that parent. A row must be constructed from each possible parent to each possible child, including a row that connects the parent to itself.



'Bridge table

Parent Organization key	Child Organization key	Depth From parent	Highest Parent flag	Lowest Child fla
A	A	0	TRUE	FALSE
A	B	1	TRUE	FALSE
A	C	1	TRUE	FALSE
A	D	2	TRUE	TRUE
A	E	2	TRUE	TRUE
A	F	2	TRUE	TRUE
B	B	0	FALSE	FALSE
B	D	1	FALSE	TRUE
C	C	0	FALSE	FALSE
C	E	1	FALSE	TRUE
C	F	1	FALSE	TRUE
D	D	0	FALSE	TRUE
E	E	0	FALSE	TRUE
E	F	0	FALSE	TRUE