Dimensional modeling process

The dimensional data model is built based on star schema with a fact table at the center surrounded by a number of dimension tables. The following four-step process is commonly used in dimensional modeling design:

1. Select the business process
2. Declare the grain
3. Identify the dimensions
4. Identify the Fact

Let's examine each step in the modeling process in greater detail.

1. Select the business process to model – business process is daily activities performed in your company supported by an online transaction system (OLTP) or source system. In this step, we have to gather the requirements from business users to select the business process or source of measurement to model. Good examples of business processes are order processing, shipments, materials purchasing, GL, etc.
2. Declare the grain – after having a business process to model, we need to declare the grain of a business process. Declaring grain means describing exactly what a record in a fact table represents. The grains express the level of detail associating with facts in the fact table.
3. Identify the dimensions – in this step, we add a number of dimensions that represent all possible descriptions that take on single values in the context of each fact in the fact table. Date, time, product, customer, store, etc., are several good examples of common dimensions.
4. Identify the facts – in the last step, we select the numeric facts that will be loaded into the fact table. To identify the facts, we need to find the KPIs of the business process or find out what we are trying to measure.

EXERCISE 1

Scenario
Industry: Automobile manufacturing
Company: Millennium Motors
Value chain focus: Sales

Sample business questions:
What are the top 10 selling car models this month?

How do this month's top 10 selling models compare to the top 10 over the last six months?

Show me dealer sales by region by model by day

What is the total number of cars sold by month by dealer by state?

List facts and dimensions

EXERCISE 1 SOLUTION

Facts
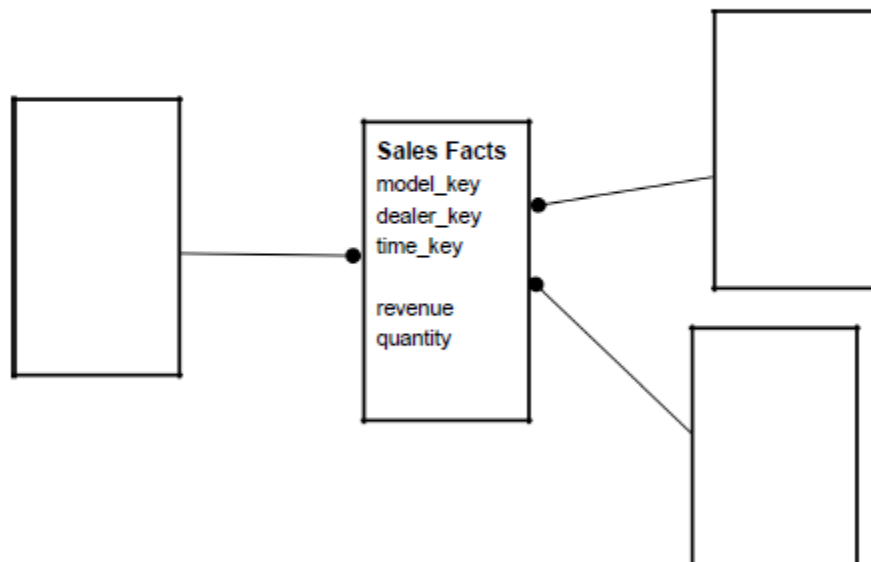Sales revenue
Quantity sold

Dimensions
Model name
Month
Dealer name
Region
State
Date

Example Fact Table

Sales Facts
model_key
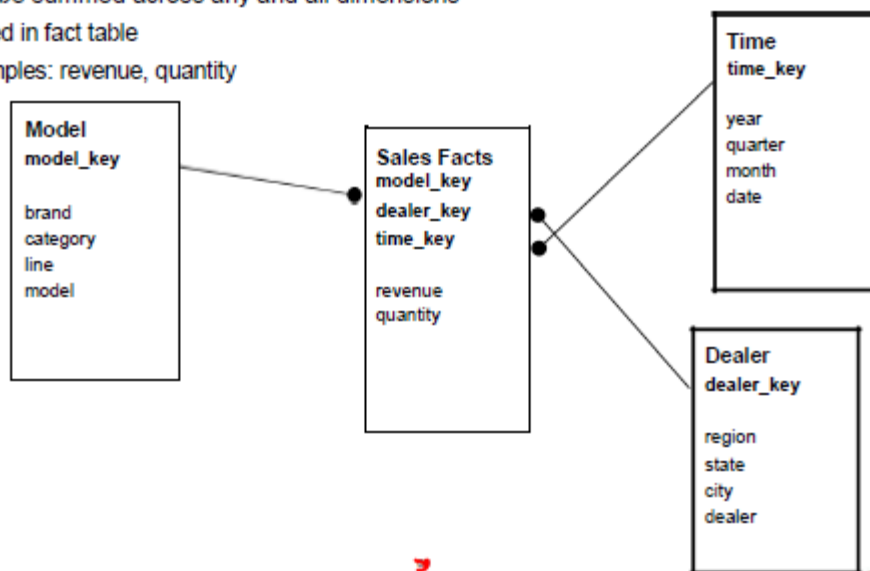dealer_key
time_key

revenue
quantity

Example Fact Table Records

Sales Facts

time_key	model_key	dealer_key	revenue	quantity
1	1	1	75840.27	2
1	2	1	152260.37	3
1	3	1	28360.15	1
1	4	1	132675.22	4
1	5	1	43789.45	1
1	1	2	35678.98	1
1	3	2	57864.78	2
1	5	2	92876.67	2

Primary Key   Facts

Facts

Fully additive

- Can be summed across any and all dimensions
- Stored in fact table
- Examples: revenue, quantity

Time
time_key

year
quarter
month
date

Model
model_key

brand
category
line
model

Sales Facts
model_key
dealer_key
time_key

revenue
quantity

Dealer
dealer_key

region
state
city
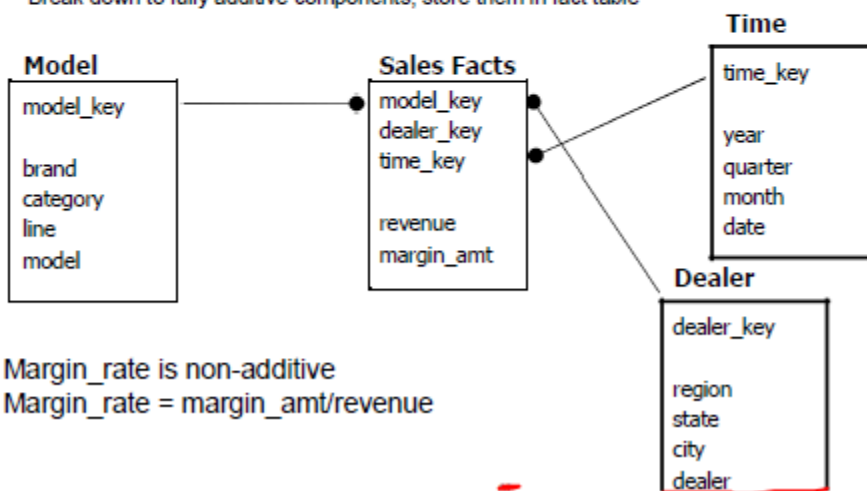dealer

Facts

Semi-additive

- Can be summed across most dimensions but not all
- Examples: Inventory quantities, account balances, or personnel counts
- Anything that measures a —level‖
- Must be careful with ad-hoc reporting
- Often aggregated across the —forbidden dimension‖ by averaging

Model
model_key

brand
category
line
model

Sales Facts
model_key
dealer_key
time_key

inventory

Time
time_key

year
quarter
month
date

Dealer
dealer_key

region
state
city
dealer
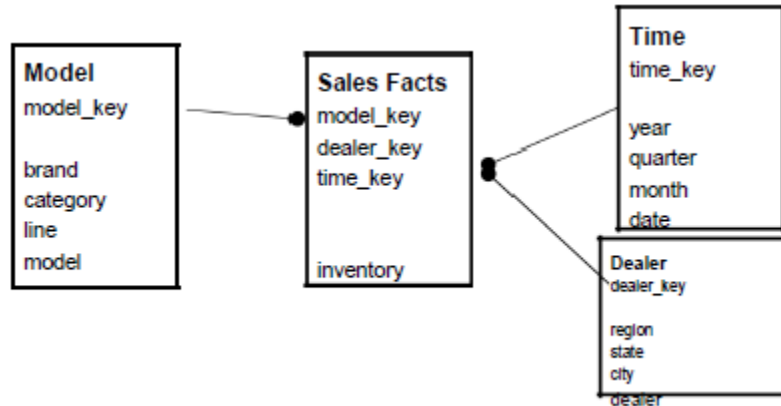
Facts

Non-Additive

- Cannot be summed across any dimension
- All ratios are non-additive
- Break down to fully additive components, store them in fact table

Model
model_key

brand
category
line
model

Sales Facts
model_key
dealer_key
time_key

revenue
margin_amt

Time
time_key

year
quarter
month
date

Dealer
dealer_key

region
state
city
dealer

Margin_rate is non-additive
Margin_rate = margin_amt/revenue

Unit Amounts

Unit price, Unit cost, etc.
- Are numeric, but not measures
- Store the extended amounts which are additive
- Unit amounts may be useful as dimensions for —price point analysis‖
- May store unit values to save space

Fact less Fact Table
- A fact table with no measures in it
- Nothing to measure...
- Except the convergence of dimensional attributes
- Sometimes store a —1‖ for convenience
- Examples: Attendance, Customer Assignments, Coverage

Example Dimension Table Records

Dealer Dimension

dealer_key	region	state	city	dealer
1	Northeast	Massachusetts	Boston	Honest Ted's
2	Northeast	Massachusetts	Boston	Stoller Co.
3	Southwest	Arizona	Tucson	Wright Motors
12	Southwest	California	San Diego	American
245	Central	Illinois	Chicago	Lugwig Motors

Attributes