Ch13 Business Intelligence Exercises

As usual, please ignore question numbers and other copy/paste anomalies ☺

FIGURE 13.1 BUSINESS INTELLIGENCE FRAMEWORK

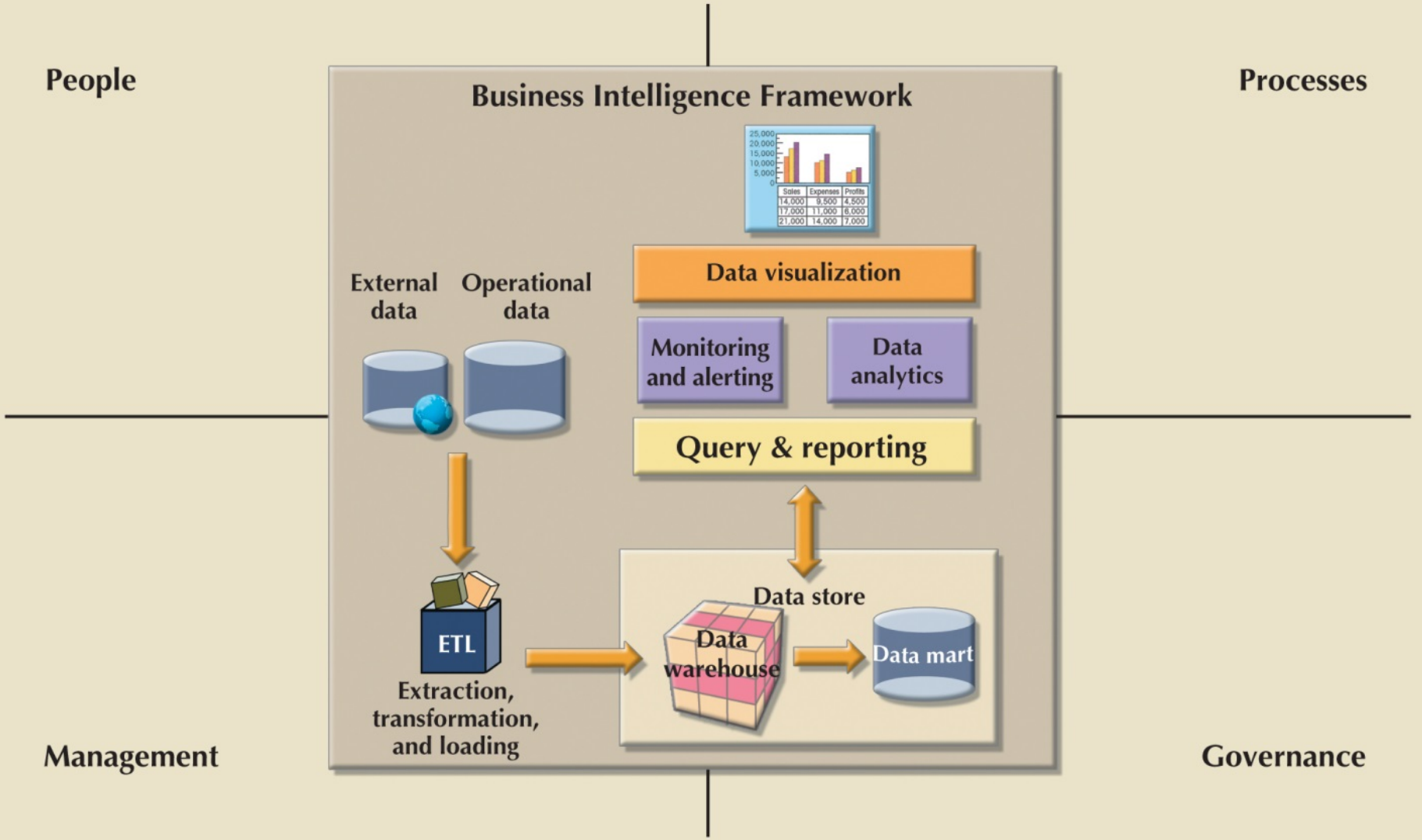


FIGURE 13.5 THE ETL PROCESS

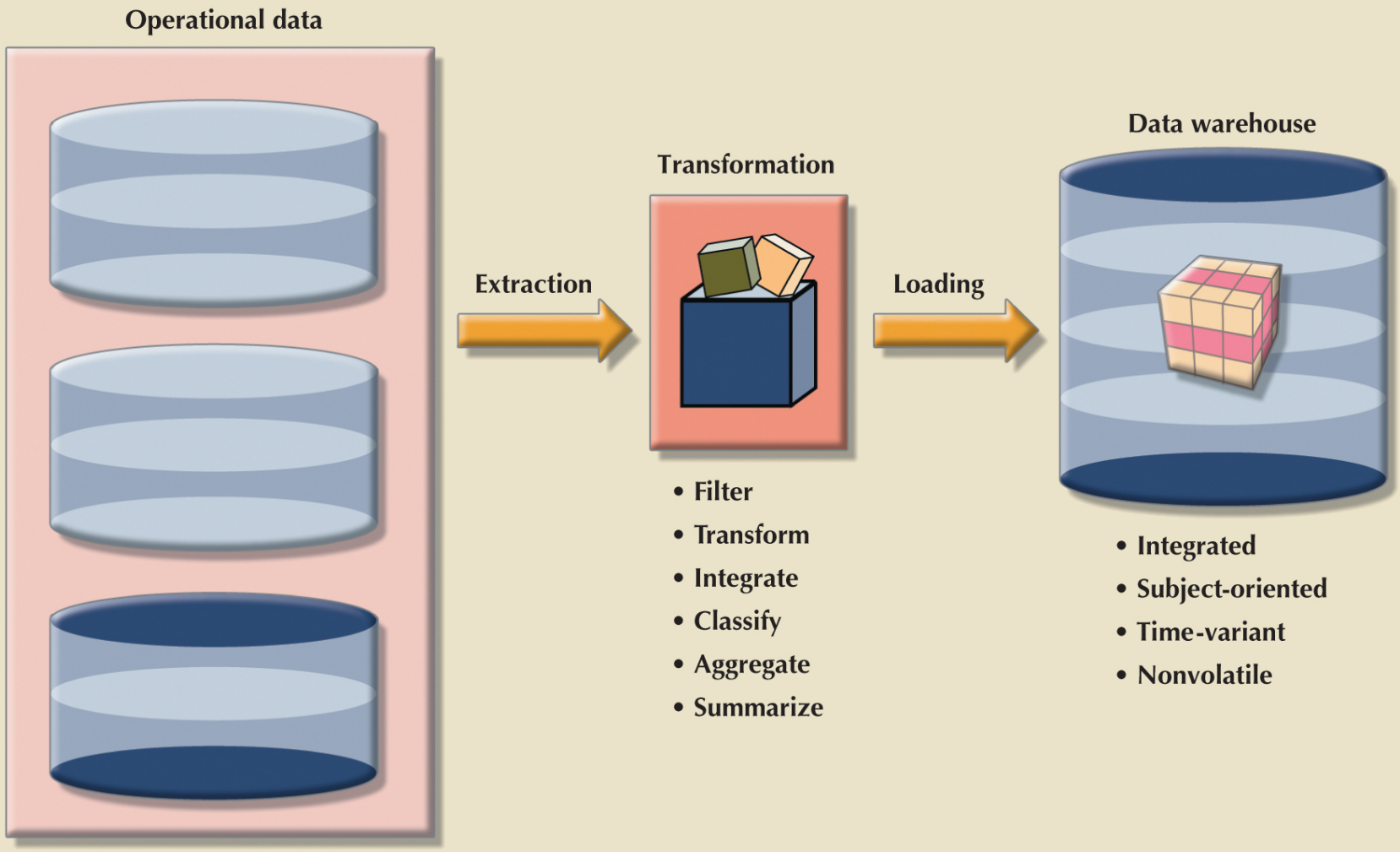


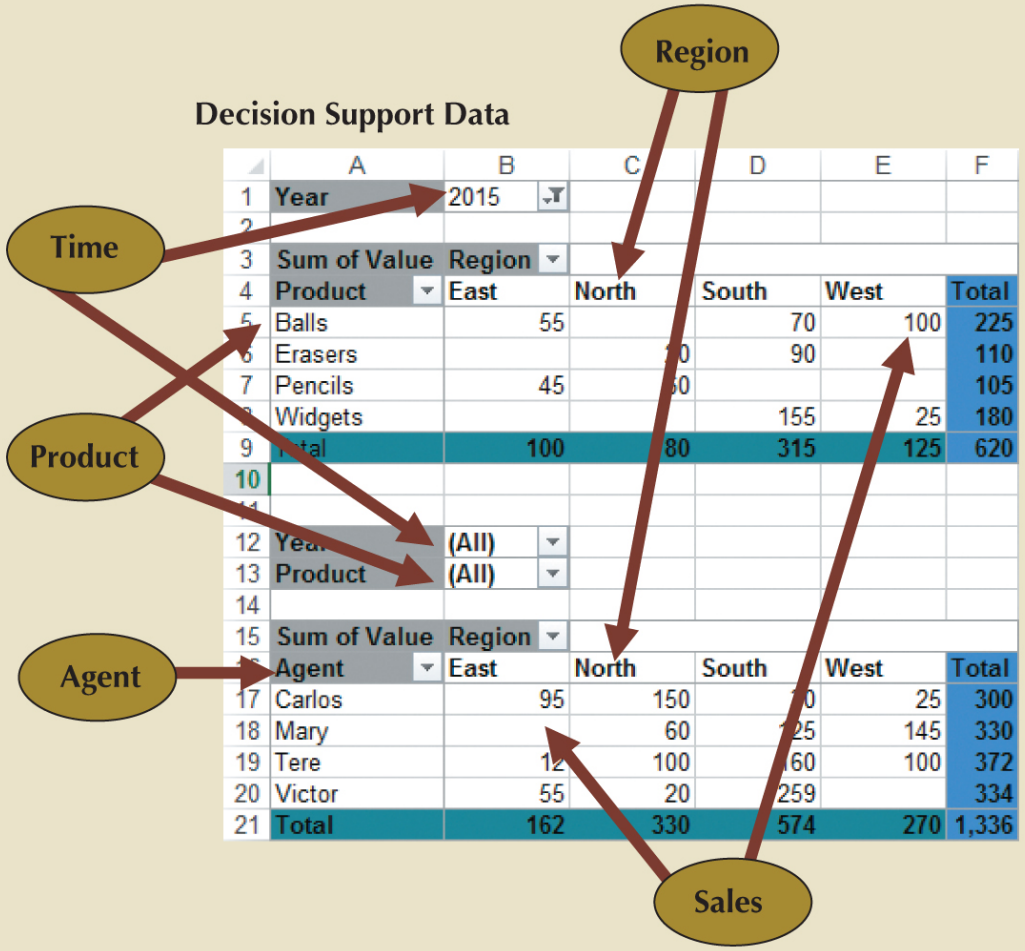
FIGURE 13.4 TRANSFORMING OPERATIONAL DATA INTO DECISION SUPPORT DATA

Operational Data

	A	B	C	D	E
1	Year	Region	Agent	Product	Value
2	2014	East	Carlos	Erasers	50
3	2014	East	Tere	Erasers	12
4	2014	North	Carlos	Widgets	120
5	2014	North	Tere	Widgets	100
6	2014	North	Carlos	Widgets	30
7	2014	South	Victor	Balls	145
8	2014	South	Victor	Balls	34
9	2014	South	Victor	Balls	80
10	2014	West	Mary	Pencils	89
11	2014	West	Mary	Pencils	56
12	2015	East	Carlos	Pencils	45
13	2015	East	Victor	Balls	55
14	2015	North	Mary	Pencils	60
15	2015	North	Victor	Erasers	20
16	2015	South	Carlos	Widgets	30
17	2015	South	Mary	Widgets	75
18	2015	South	Mary	Widgets	50
19	2015	South	Tere	Balls	70
20	2015	South	Tere	Erasers	90
21	2015	West	Carlos	Widgets	25
22	2015	West	Tere	Balls	100

Operational data has a narrow time span, low granularity, and single focus. Such data is usually represented in tabular format, in which each row represents a single transaction. This format often makes it difficult to derive useful information.

Decision Support Data



	A	B	C	D	E	F
1	Year	2015				
2						
3	Sum of Value	Region				
4	Product	East	North	South	West	Total
5	Balls	55		70	100	225
6	Erasers			90		110
7	Pencils	45				105
8	Widgets			155	25	180
9	Total	100		315	125	620
10						
11						
12	Year	(All)				
13	Product	(All)				
14						
15	Sum of Value	Region				
16	Agent	East	North	South	West	Total
17	Carlos	95	150		25	300
18	Mary		60		145	330
19	Tere		100		100	372
20	Victor	55	20	259		334
21	Total	162	330	574	270	1,336

Decision support system (DSS) data focuses on a broader time span, tends to have high levels of granularity, and can be examined in multiple dimensions. For example, note these possible aggregations:

- **Sales by product, region, agent, and so on**
- **Sales for all years or only a few selected years**
- **Sales for all products or only a few selected products**

FIGURE 13.6 SIMPLE STAR SCHEMA

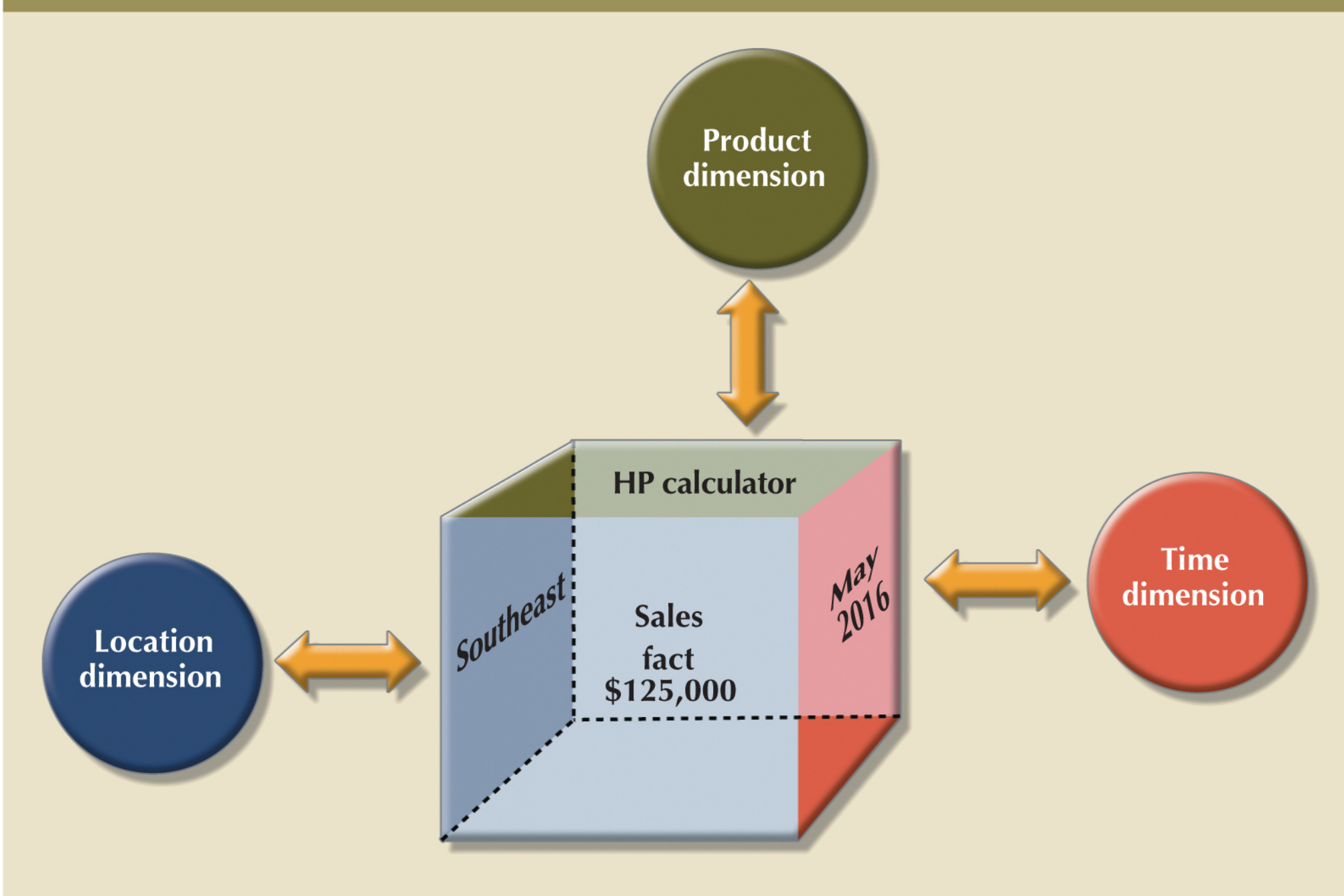


FIGURE 13.10 ATTRIBUTE HIERARCHIES IN MULTIDIMENSIONAL ANALYSIS

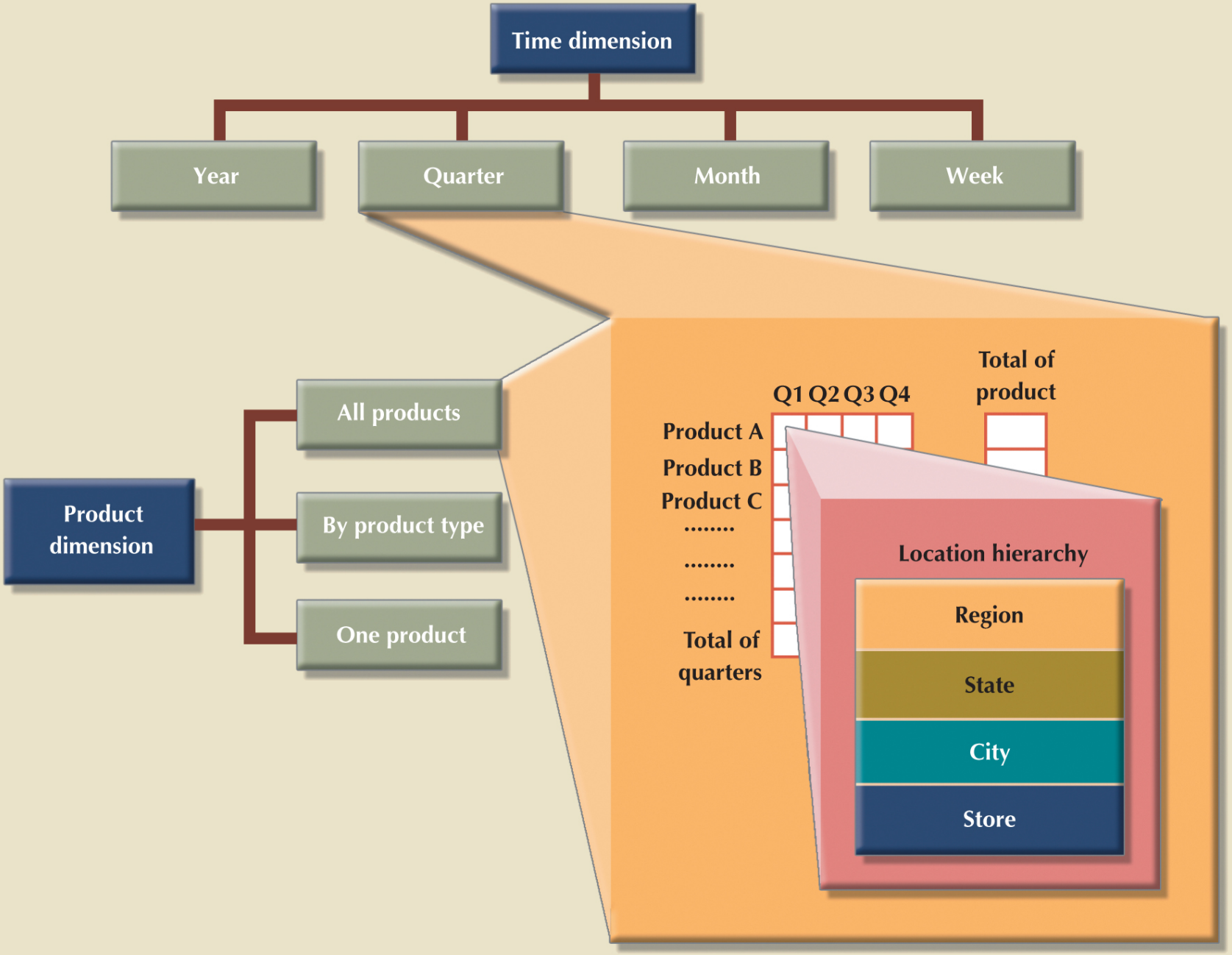


FIGURE 13.11 STAR SCHEMA FOR SALES

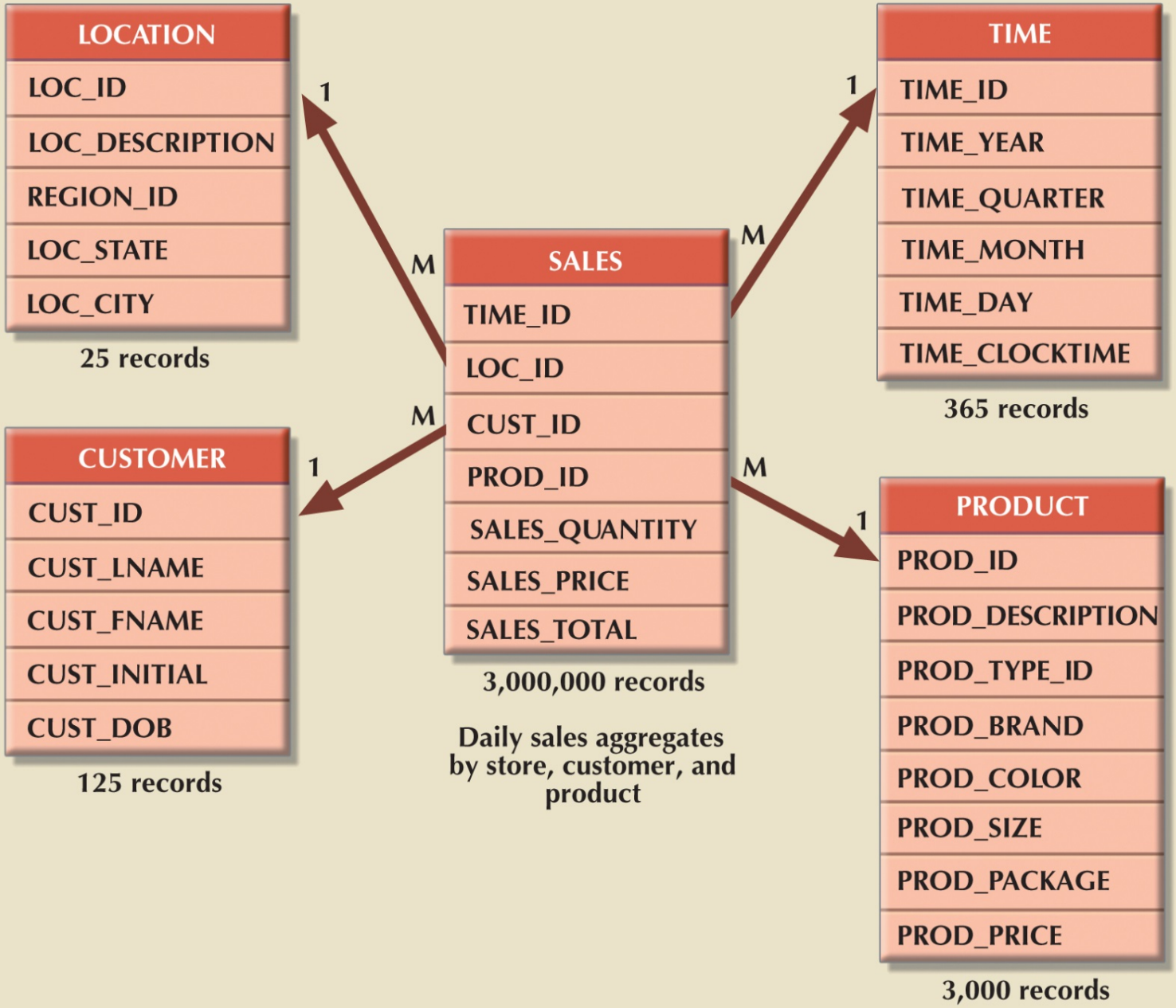


FIGURE 13.13 NORMALIZED DIMENSION TABLES

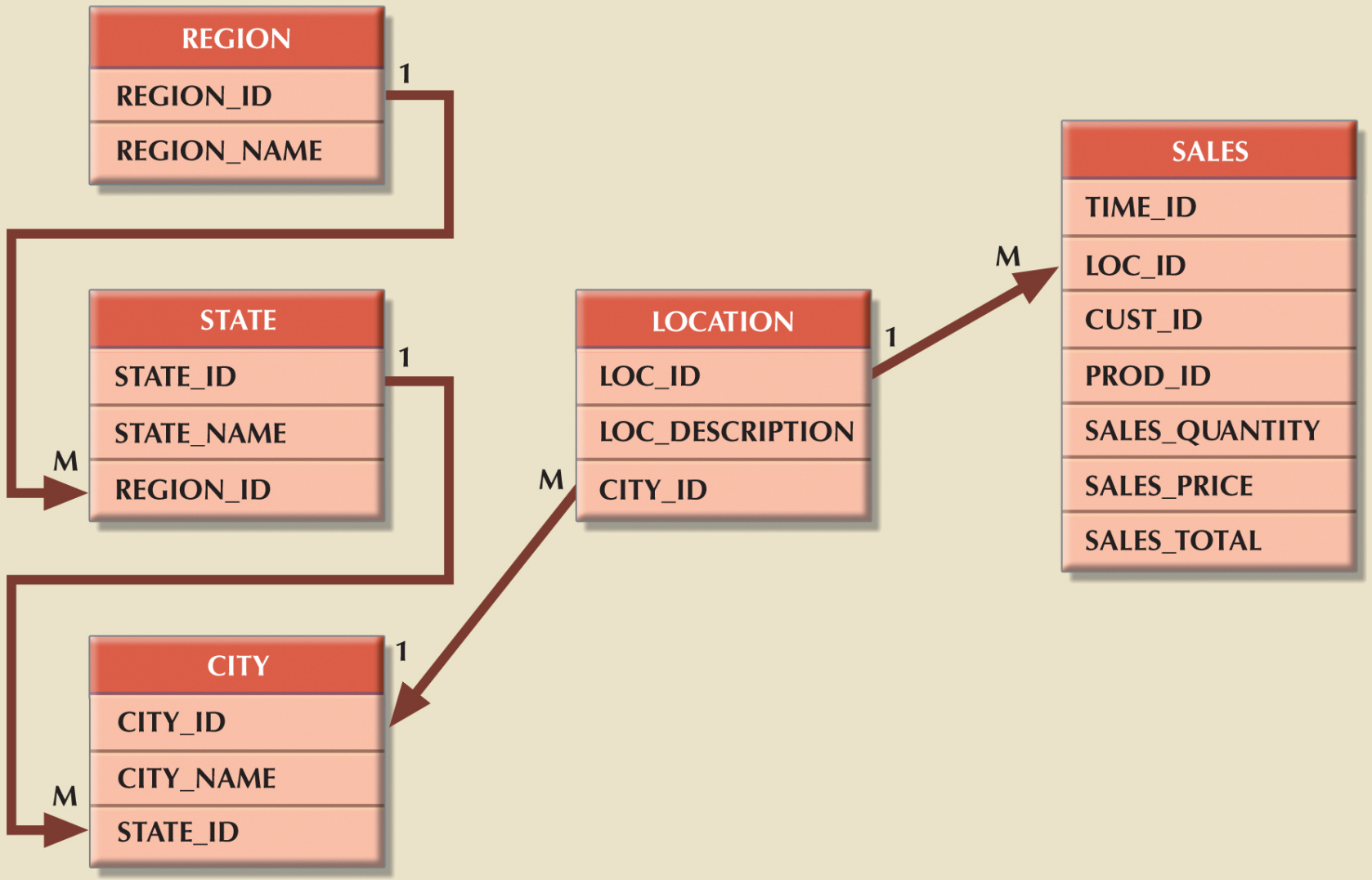


FIGURE 13.14 MULTIPLE FACT TABLES

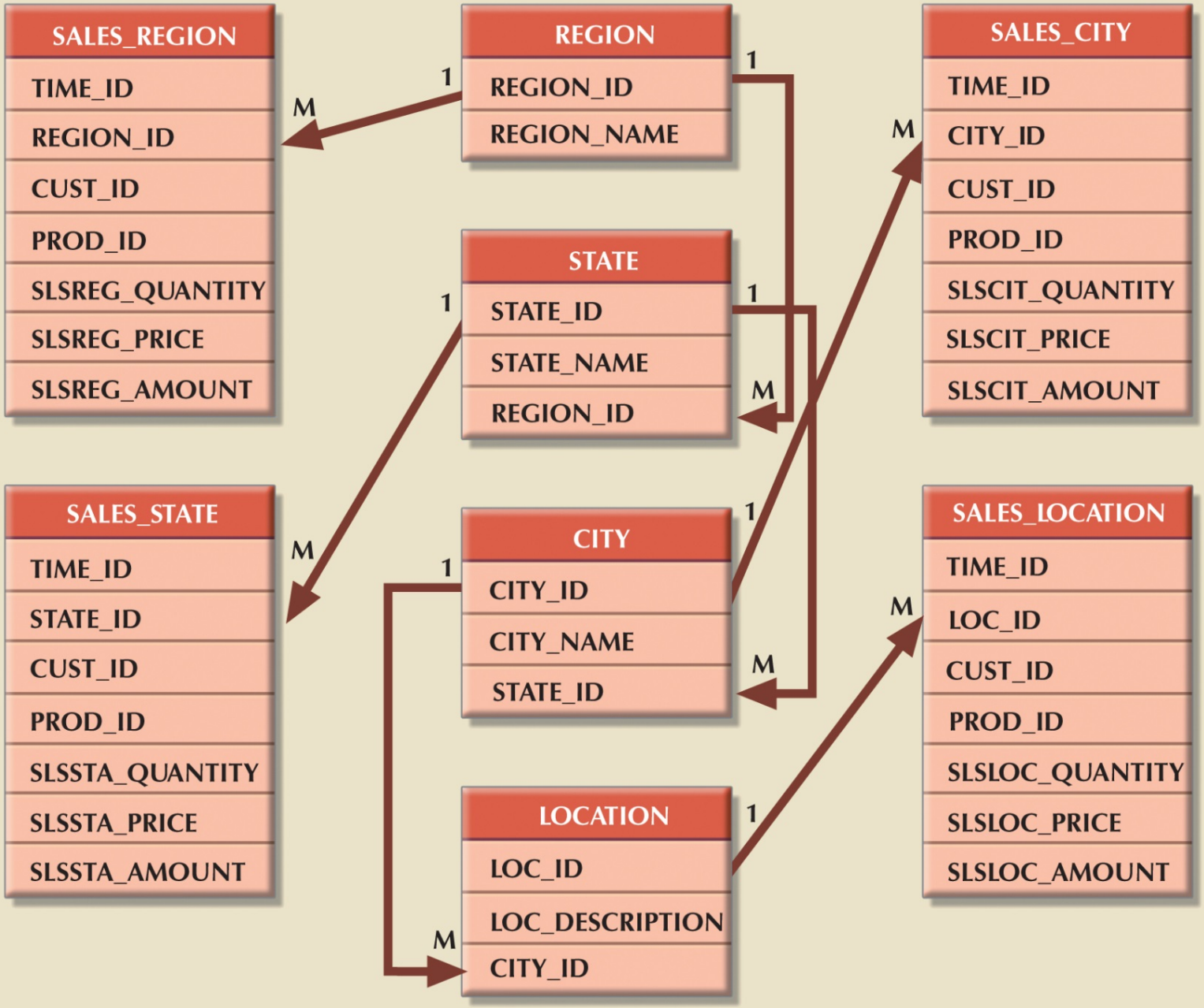


FIGURE 13.15 OPERATIONAL VS. MULTIDIMENSIONAL VIEW OF SALES

Table name: DW_INVOICE

INV_NUM	INV_DATE	CUS_NAME	INV_TOTAL
2034	15-May-16	Dartonik	1400.00
2035	15-May-16	Summer Lake	1200.00
2036	16-May-16	Dartonik	1350.00
2037	16-May-16	Summer lake	3100.00
2038	16-May-16	Trydon	400.00

Operational Data

Table name: DW_LINE

INV_NUM	LINE_NUM	PROD_DESCRIPTION	LINE_PRICE	LINE_QUANTITY	LINE_AMOUNT
2034	1	Optical Mouse	45.00	20	900.00
2034	2	Wireless RF remote and laser pointer	50.00	10	500.00
2035	1	Everlast Hard Drive, 60 GB	200.00	6	1200.00
2036	1	Optical Mouse	45.00	30	1350.00
2037	1	Optical Mouse	45.00	10	450.00
2037	2	Roadster 56KB Ext. Modem	120.00	5	600.00
2037	3	Everlast Hard Drive, 60 GB	205.00	10	2050.00
2038	1	NoTech Speaker Set	50.00	8	400.00

Multidimensional View of Sales
(using MS Excel PivotTable)

Time Dimension

Multidimensional View of SALES	INV_DATE		
CUS_NAME	15-May-16	16-May-16	Grand Total
Dartonik	$ 1,400.00	$1,350.00	$ 2,750.00
Summer Lake	$ 1,200.00	$3,100.00	$ 4,300.00
Trydon		$ 400.00	$ 400.00
Grand Total	$ 2,600.00	$4,850.00	$ 7,450.00

Customer Dimension

Sales are located in the intersection of a customer row and date (time) column

Aggregations (grand total sales) are provided for both dimensions (time and customer)

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA

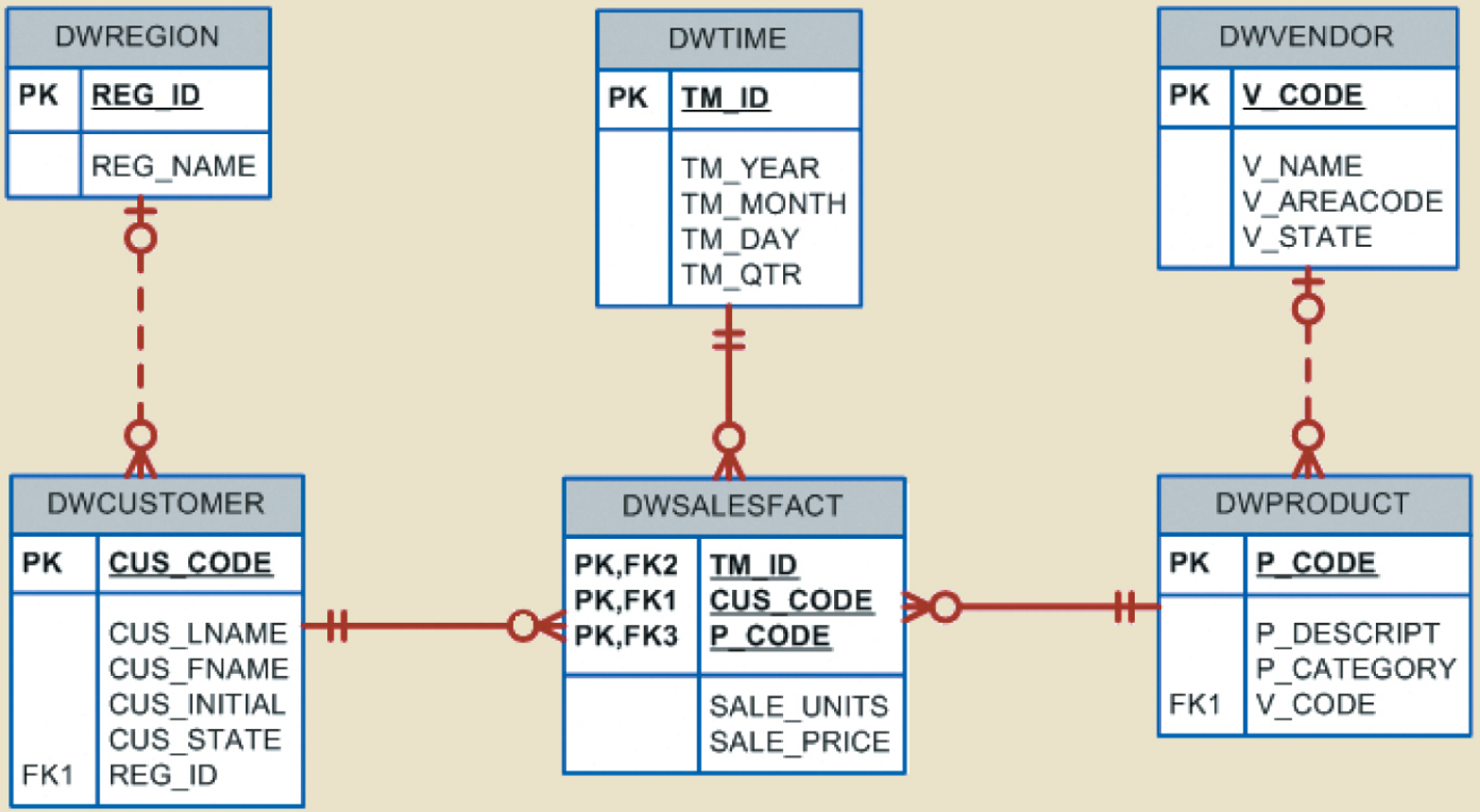


FIGURE 13.20 ROLLUP EXTENSION

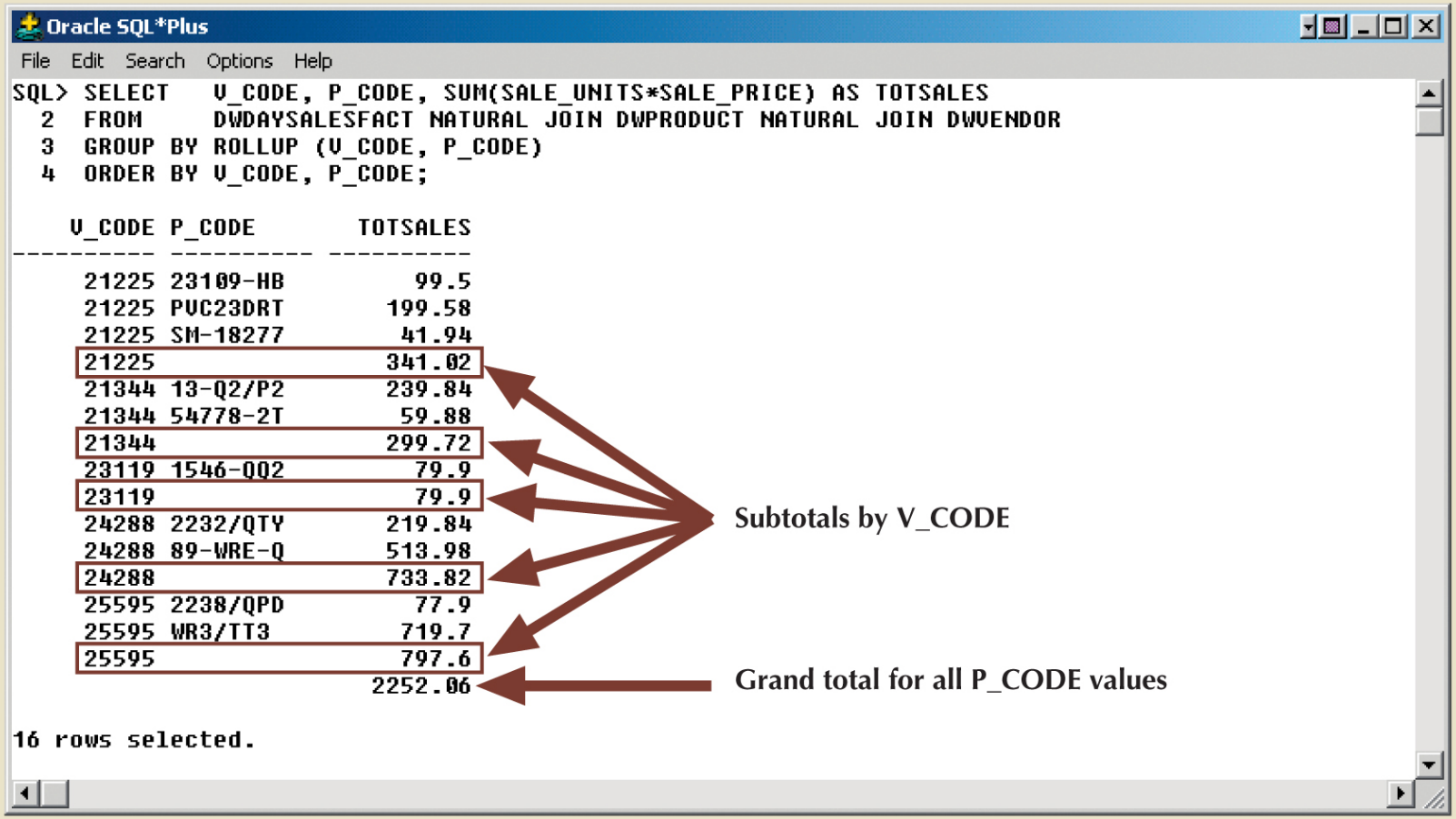


FIGURE 13.21 CUBE EXTENSION

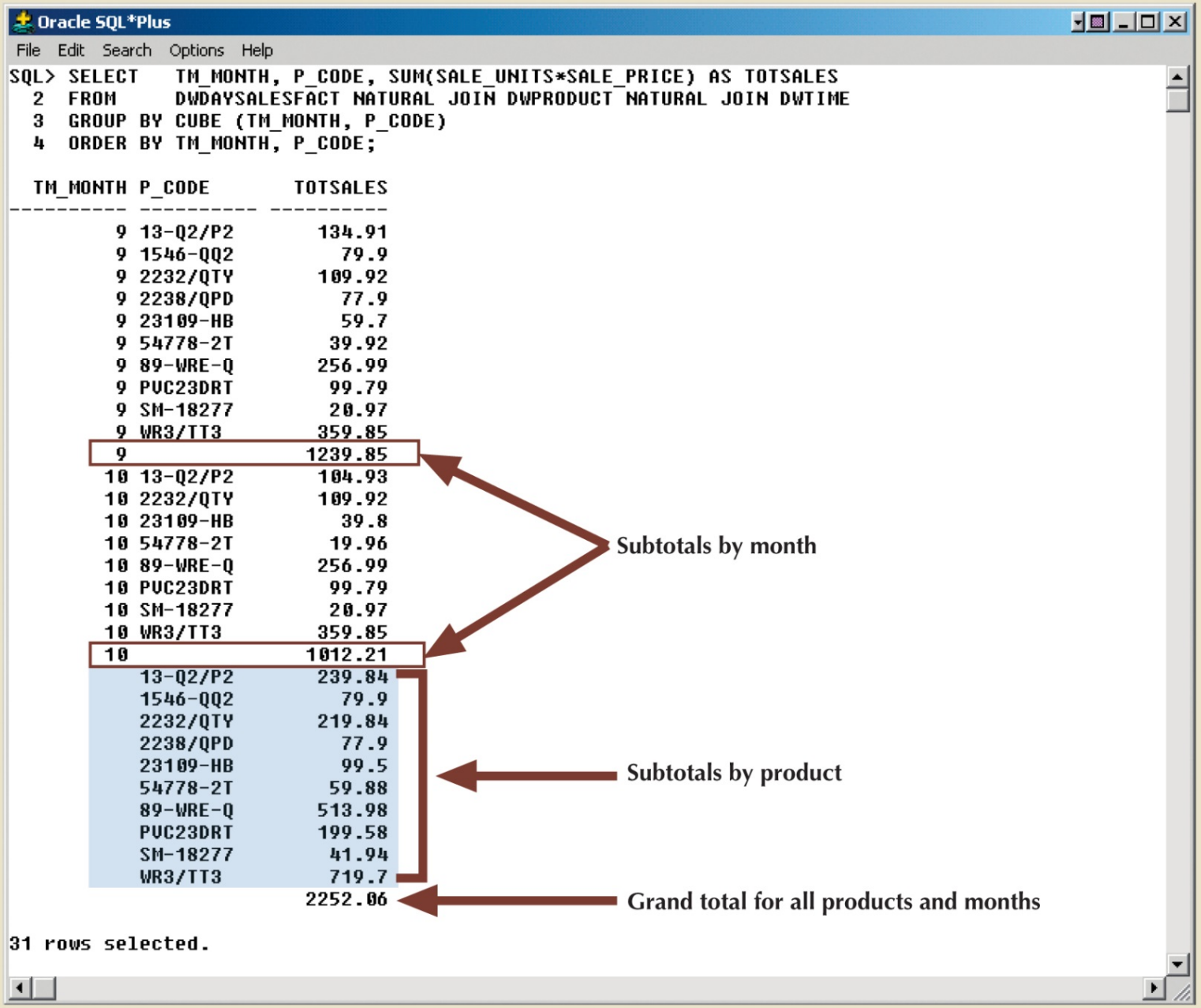
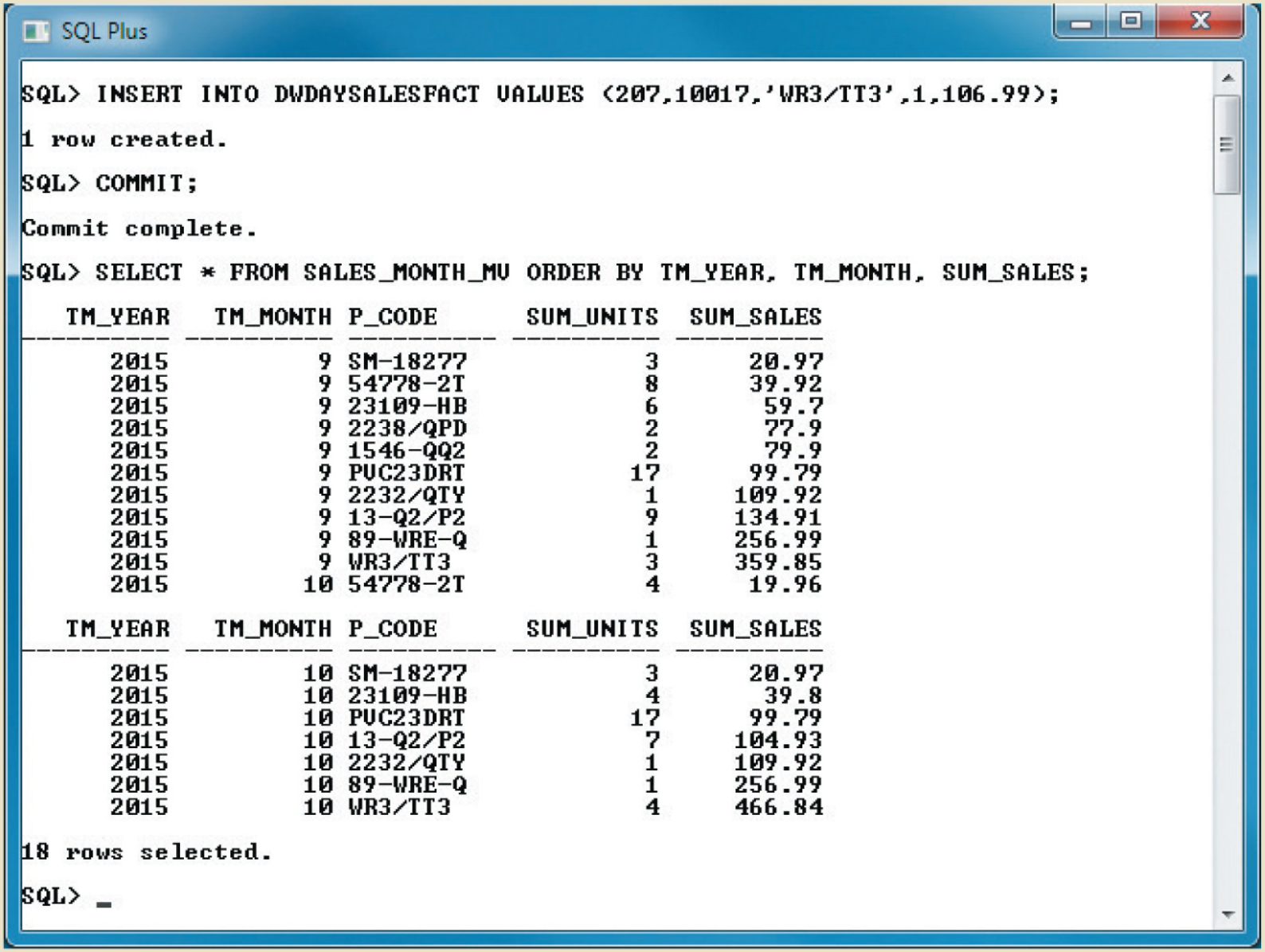


FIGURE 13.22 CREATING A MATERIALIZED VIEW

SQL Plus

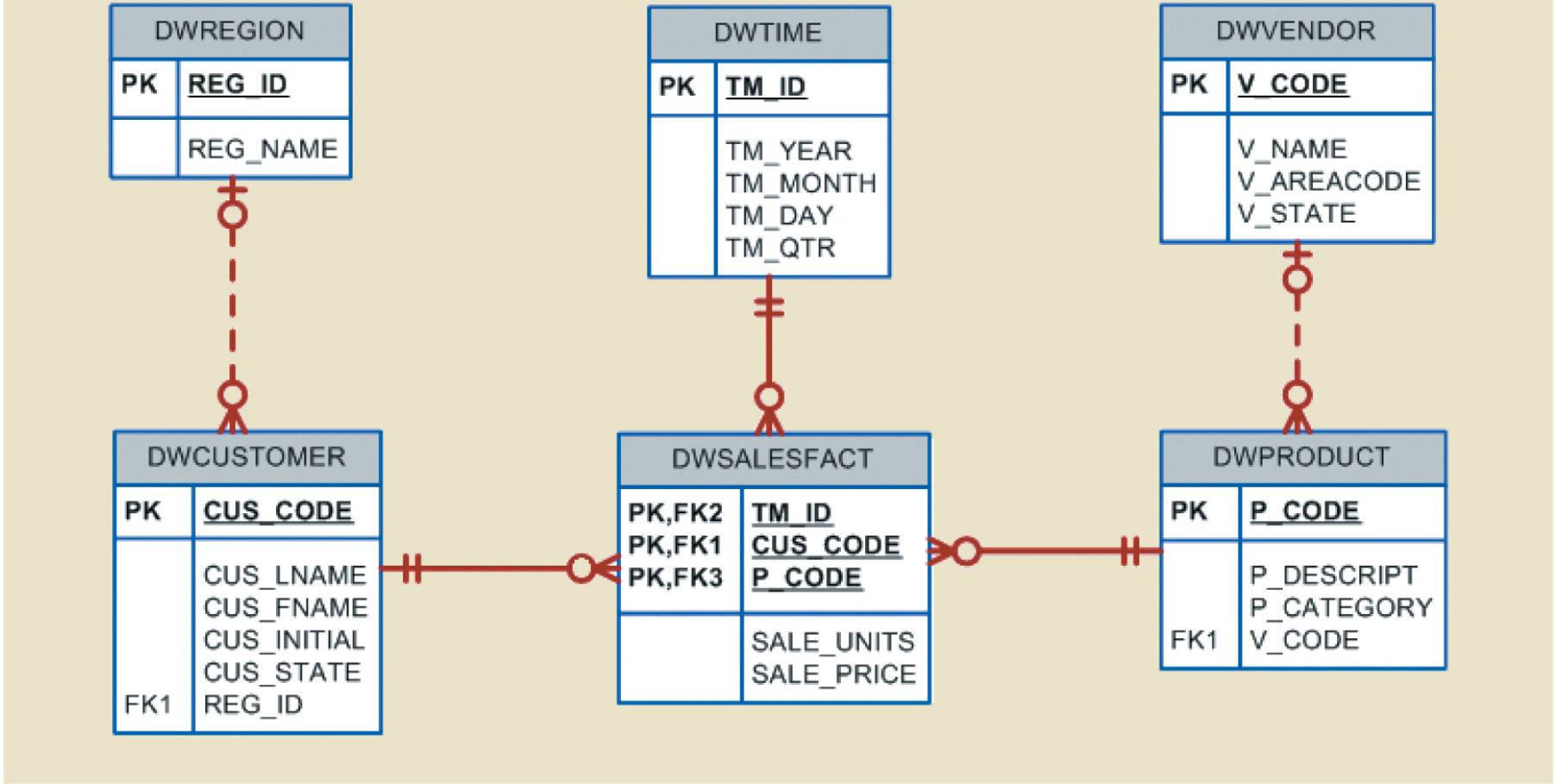
```
SQL> CREATE MATERIALIZED VIEW LOG ON DWTIME
  2  WITH ROWID, SEQUENCE INCLUDING NEW VALUES;

Materialized view log created.

SQL> CREATE MATERIALIZED VIEW LOG ON DWDAYSALESFACT
  2  WITH ROWID, SEQUENCE INCLUDING NEW VALUES;

Materialized view log created.

SQL> CREATE MATERIALIZED VIEW SALES_MONTH_MV
  2     BUILD IMMEDIATE
  3     REFRESH FORCE ON COMMIT
  4     ENABLE QUERY REWRITE
  5     AS SELECT     TM_YEAR, TM_MONTH, P_CODE,
  6                   SUM(SALE_UNITS) AS SUM_UNITS,
  7                   SUM(SALE_PRICE*SALE_UNITS) AS SUM_SALES
  8        FROM       DWTIME T, DWDAYSALESFACT S
  9        WHERE      S.TM_ID = T.TM_ID
 10        GROUP BY      TM_YEAR, TM_MONTH, P_CODE;

Materialized view created.

SQL> SELECT * FROM SALES_MONTH_MV ORDER BY TM_YEAR, TM_MONTH, SUM_SALES;

   TM_YEAR   TM_MONTH P_CODE      SUM_UNITS  SUM_SALES
---------- ---------- ---------- ---------- ----------
      2015          9 SM-18277            3      20.97
      2015          9 54778-2T            8      39.92
      2015          9 23109-HB            6       59.7
      2015          9 2238/QPD            2       77.9
      2015          9 1546-QQ2            2       79.9
      2015          9 PVC23DRT           17      99.79
      2015          9 2232/QTY            1     109.92
      2015          9 13-Q2/P2            9     134.91
      2015          9 89-WRE-Q            1     256.99
      2015          9 WR3/TT3             3     359.85
      2015         10 54778-2T            4      19.96

   TM_YEAR   TM_MONTH P_CODE      SUM_UNITS  SUM_SALES
---------- ---------- ---------- ---------- ----------
      2015         10 SM-18277            3      20.97
      2015         10 23109-HB            4       39.8
      2015         10 PVC23DRT           17      99.79
      2015         10 13-Q2/P2            7     104.93
      2015         10 2232/QTY            1     109.92
      2015         10 89-WRE-Q            1     256.99
      2015         10 WR3/TT3             3     359.85

18 rows selected.

SQL> COMMIT;

Commit complete.

SQL> _
```

FIGURE 13.23 REFRESHING A MATERIALIZED VIEW



What is the SQL command to list the total sales by customer and by product, with subtotals by customer and a grand total for all product sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the total sales by customer and by product, with subtotals by customer and a grand total for all product sales?

Oracle:

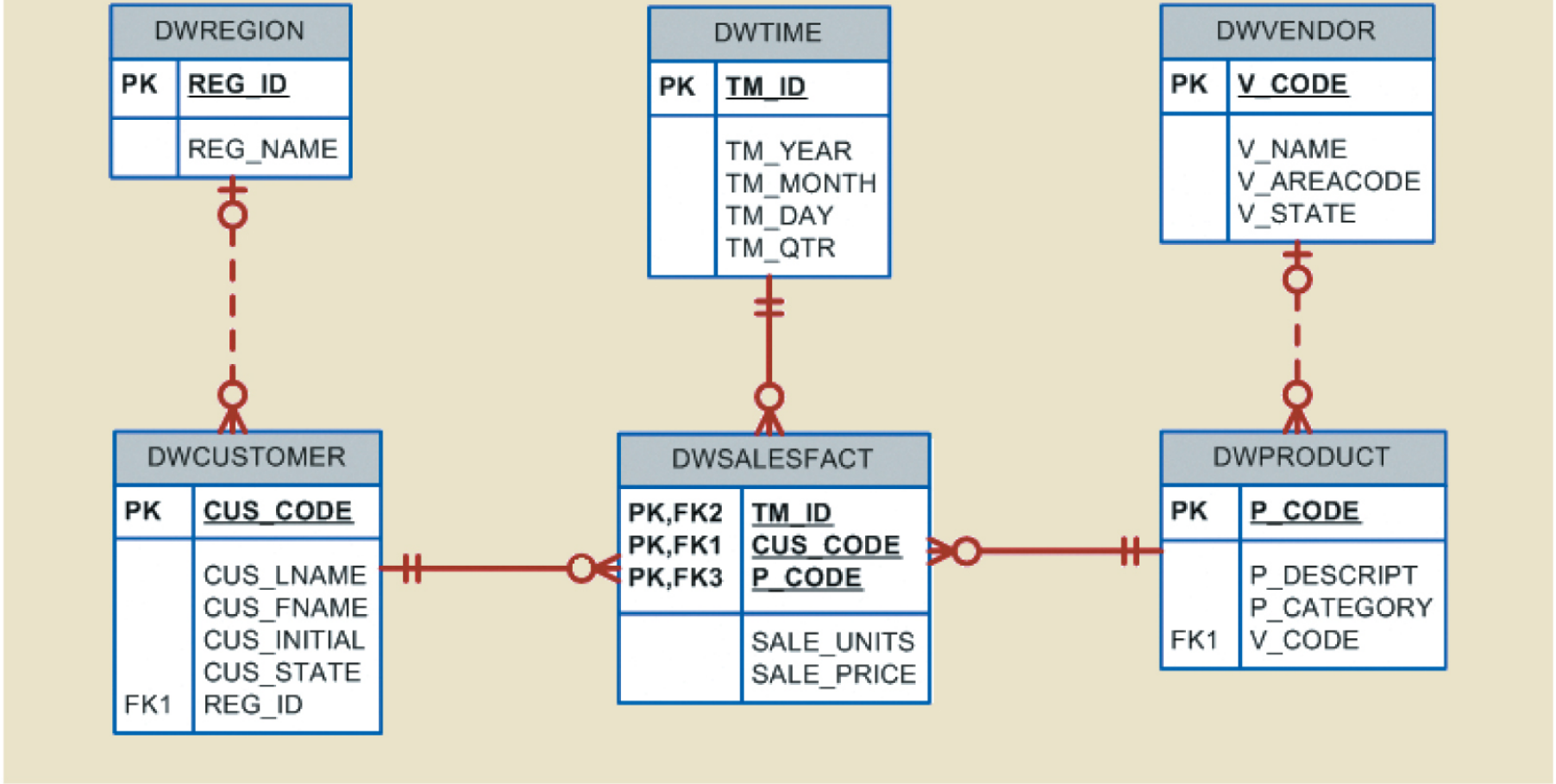
```
SELECT      CUS_CODE, P_CODE, SUM(SALE_UNITS*SALE_PRICE)  AS TOTSALES
FROM        DWDAYSALESFACT
GROUP BY    ROLLUP (CUS_CODE, P_CODE);
```

SQL Server and MySQL:

```
SELECT      CUS_CODE, P_CODE, SUM(SALE_UNITS*SALE_PRICE)  AS TOTSALES
FROM        DWDAYSALESFACT
GROUP BY    CUS_CODE, P_CODE WITH ROLLUP;
```

What is the SQL command to list the total sales by customer, month and product, with subtotals by customer and by month and a grand total for all product sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the total sales by customer, month and product, with subtotals by customer and by month and a grand total for all product sales?

Oracle:

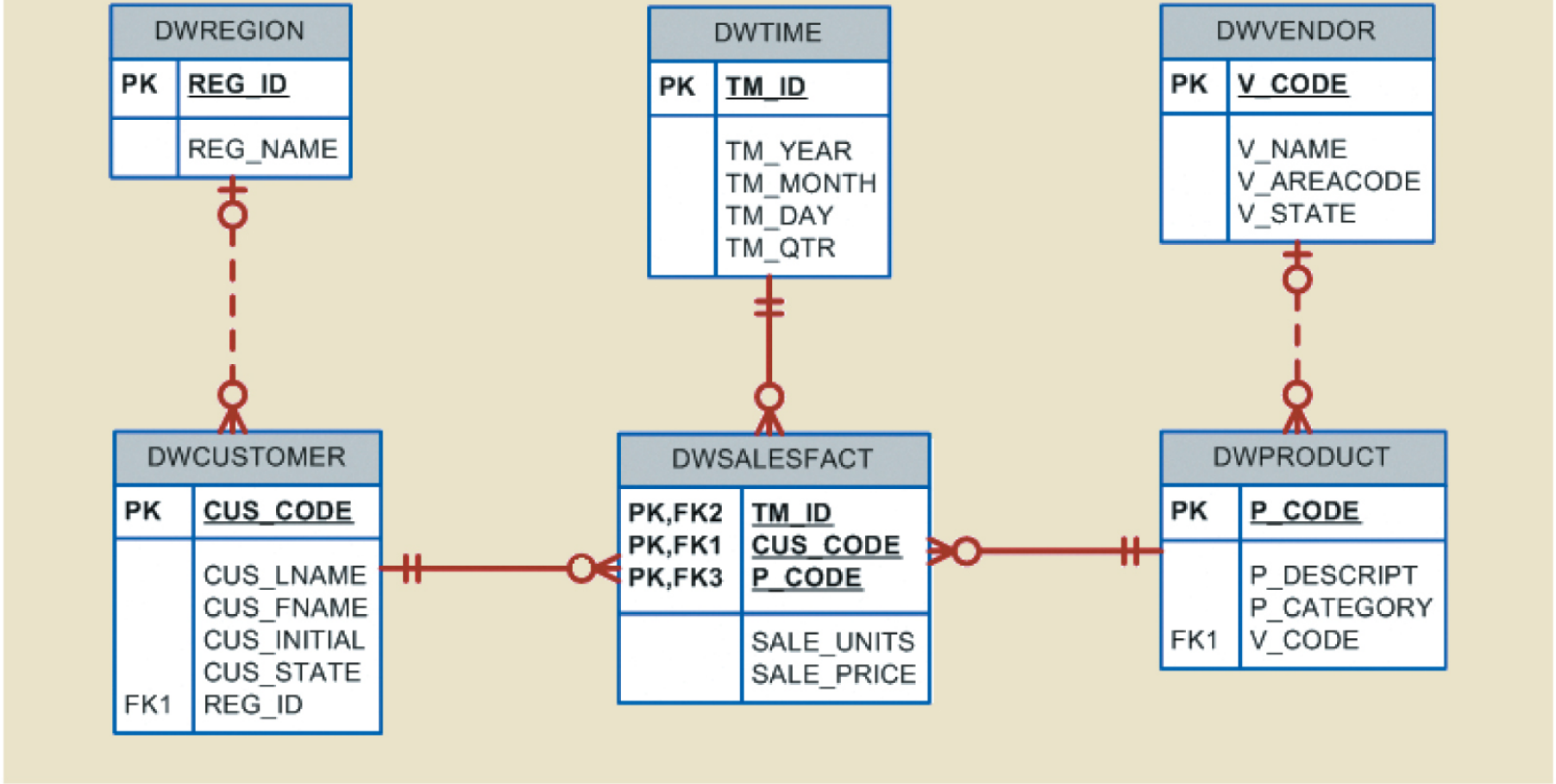
```
SELECT      CUS_CODE, TM_MONTH, P_CODE, SUM(SALE_UNITS*SALE_PRICE)
            AS TOTSALES
FROM        DWDAYSALESFACT  S JOIN DWTIME  T ON S.TM_ID = T.TM_ID
GROUP BY    ROLLUP (CUS_CODE, TM_MONTH, P_CODE);
```

SQL Server and MySQL:

```
SELECT      CUS_CODE, TM_MONTH, P_CODE, SUM(SALE_UNITS*SALE_PRICE)
            AS TOTSALES
FROM        DWDAYSALESFACT  S JOIN DWTIME  T ON S.TM_ID = T.TM_ID
GROUP BY    CUS_CODE, TM_MONTH, P_CODE WITH ROLLUP;
```

What is the SQL command to list the total sales by region and customer, with subtotals by region and a grand total for all sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the total sales by region and customer, with subtotals by region and a grand total for all sales?

Oracle:

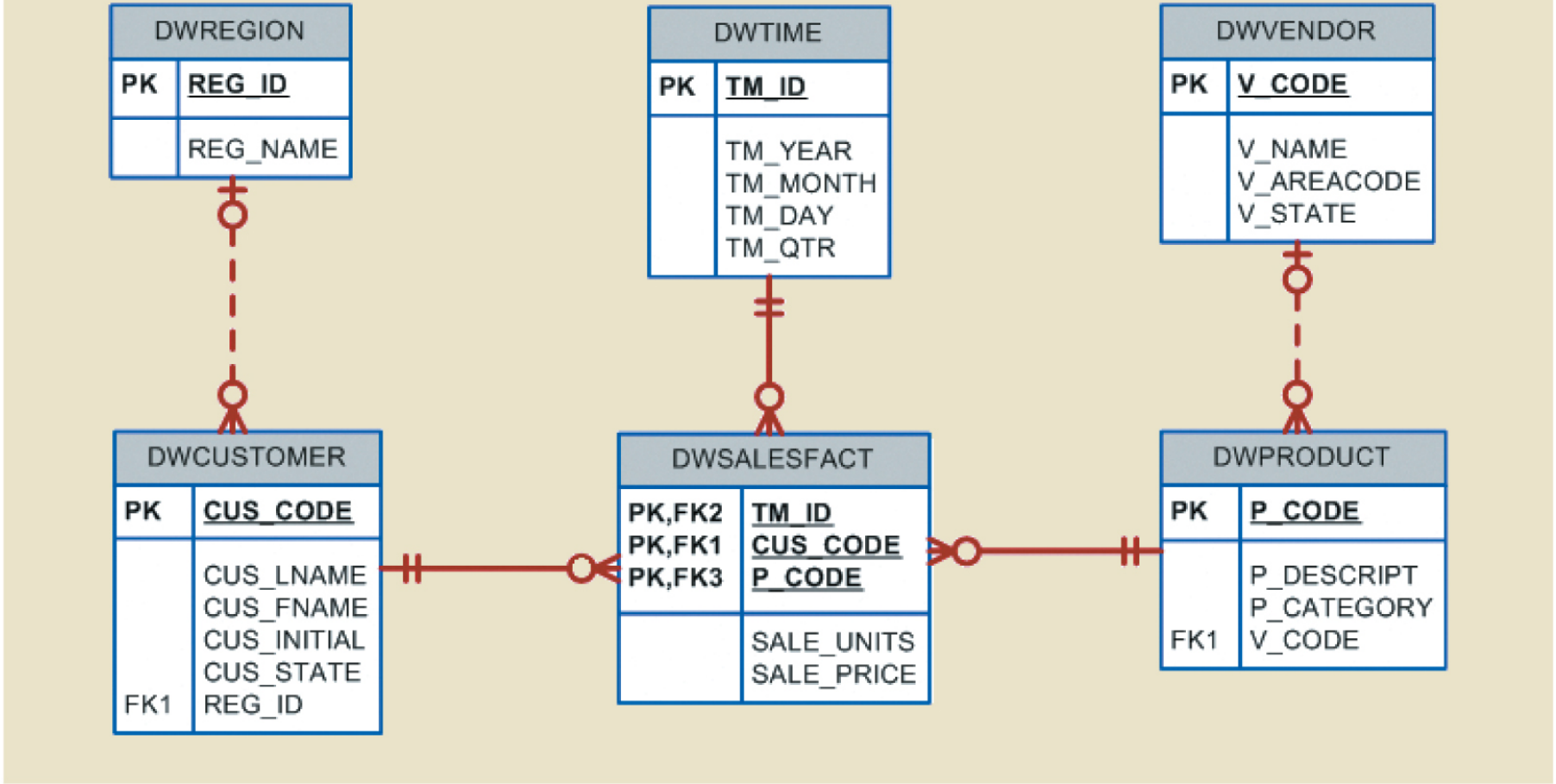
```
SELECT      REG_ID, CUS_CODE, SUM(SALE_UNITS*SALE_PRICE)  AS TOTSALES
FROM        DWDAYSALESFACT  S JOIN DWCUSTOMER  C
            ON S.CUS_CODE = C.CUS_CODE
GROUP BY    ROLLUP (REG_ID, CUS_CODE);
```

SQL Server and MySQL:

```
SELECT      REG_ID, CUS_CODE, SUM(SALE_UNITS*SALE_PRICE)  AS TOTSALES
FROM        DWDAYSALESFACT  S JOIN DWCUSTOMER  C
            ON S.CUS_CODE = C.CUS_CODE
GROUP BY    REG_ID, CUS_CODE WITH ROLLUP;
```

What is the SQL command to list the total sales by month and product category, with subtotals by month and a grand total for all sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the total sales by month and product category, with subtotals by month and a grand total for all sales?

Oracle:

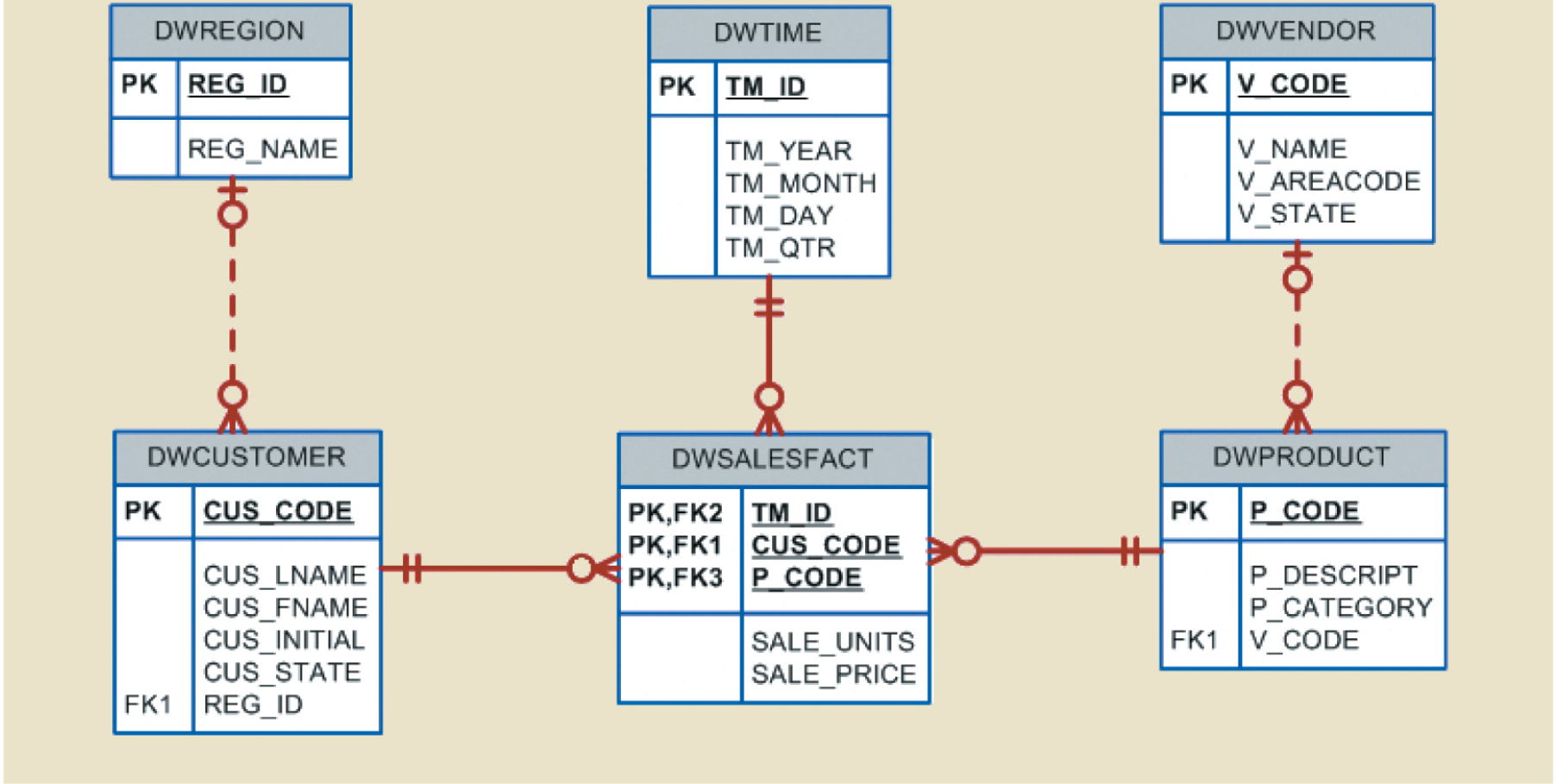
```
SELECT      TM_MONTH, P_CATEGORY, SUM(SALE_UNITS*SALE_PRICE)
            AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
            JOIN DWTIME T ON S.TM_ID = T.TM_ID
GROUP BY    ROLLUP (TM_MONTH, P_CATEGORY);
```

SQL Server and MySQL:

```
SELECT      TM_MONTH, P_CATEGORY, SUM(SALE_UNITS*SALE_PRICE)
            AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
            JOIN DWTIME T ON S.TM_ID = T.TM_ID
GROUP BY    TM_MONTH, P_CATEGORY WITH ROLLUP;
```

What is the SQL command to list the number of product sales (number of rows) and total sales by month, with subtotals by month and a grand total for all sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the number of product sales (number of rows) and total sales by month, with subtotals by month and a grand total for all sales?

Oracle:

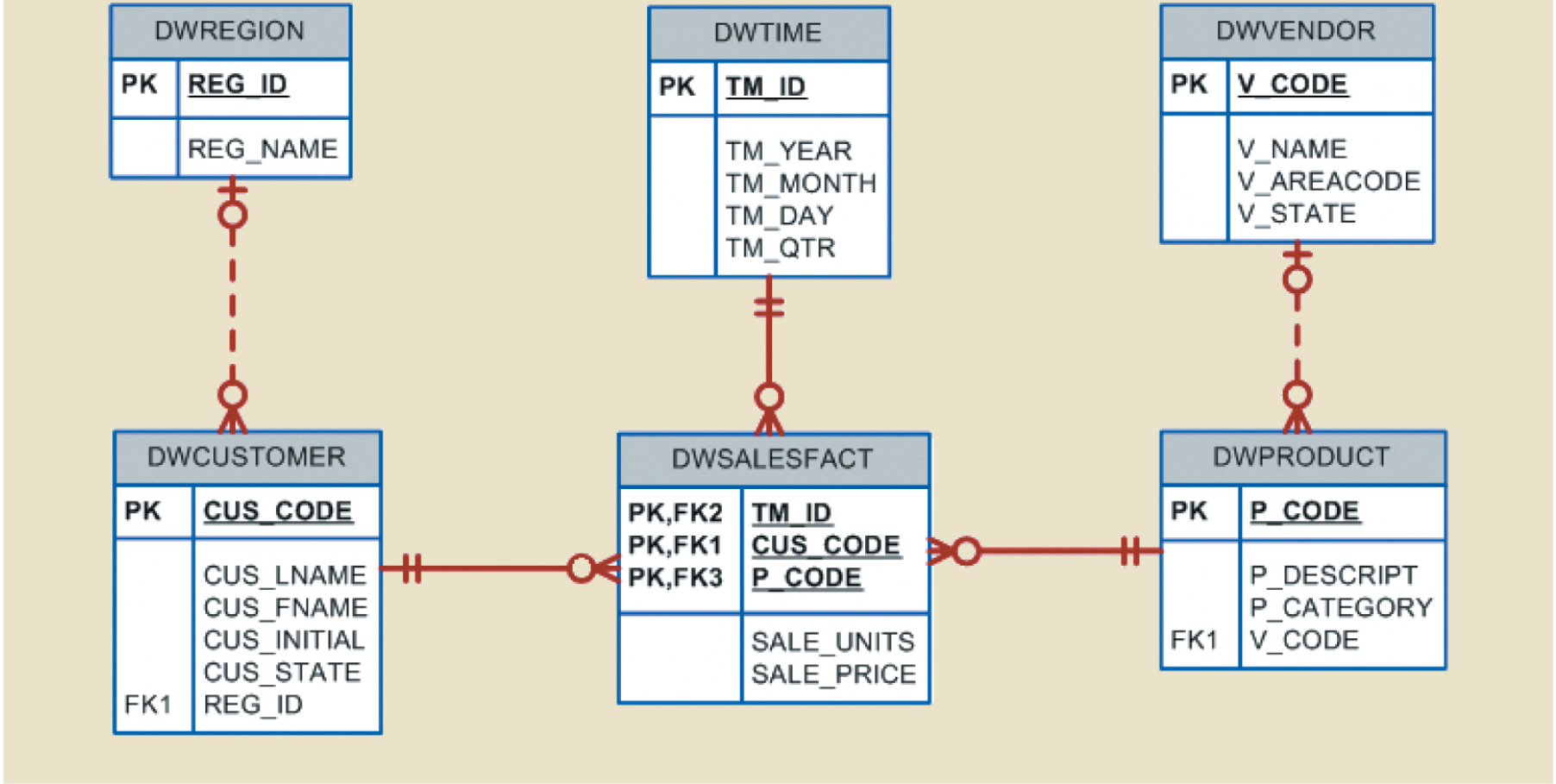
```
SELECT      TM_MONTH, COUNT(*) AS NUMPROD, SUM(SALE_UNITS*SALE_PRICE)
    AS TOTSALES
FROM        DWDAYSALESFACT  S JOIN DWTIME  T ON S.TM_ID = T.TM_ID
GROUP BY    ROLLUP (TM_MONTH);
```

SQL Server and MySQL:

```
SELECT      TM_MONTH, COUNT(*) AS NUMPROD, SUM(SALE_UNITS*SALE_PRICE)
    AS TOTSALES
FROM        DWDAYSALESFACT  S JOIN DWTIME  T ON S.TM_ID = T.TM_ID
GROUP BY    TM_MONTH WITH ROLLUP;
```

What is the SQL command to list the number of product sales (number of rows) and total sales by month and product category with subtotals by month and product category and a grand total for all sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the number of product sales (number of rows) and total sales by month and product category with subtotals by month and product category and a grand total for all sales?

Oracle:

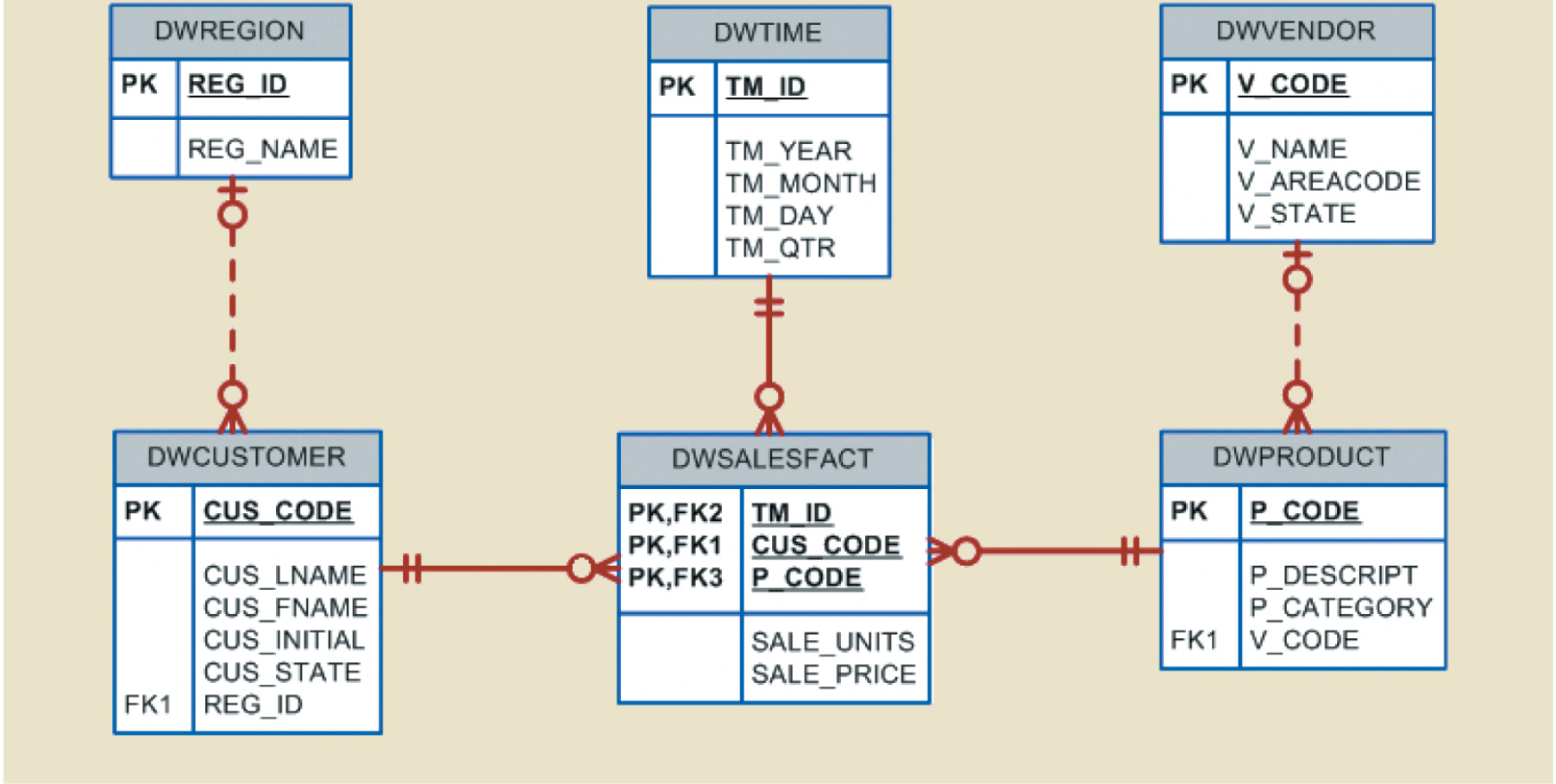
```
SELECT      TM_MONTH, P_CATEGORY, COUNT(*) AS NUMPROD,
            SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
            JOIN DWTIME T ON S.TM_ID = T.TM_ID
GROUP BY    ROLLUP (TM_MONTH, P_CATEGORY);
```

SQL Server and MySQL:

```
SELECT      TM_MONTH, P_CATEGORY, COUNT(*) AS NUMPROD,
            SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
            JOIN DWTIME T ON S.TM_ID = T.TM_ID
GROUP BY    TM_MONTH, P_CATEGORY WITH ROLLUP;
```

What is the SQL command to list the number of product sales (number of rows) and total sales by month, product category and product, with subtotals by month and product category and a grand total for all sales?

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



What is the SQL command to list the number of product sales (number of rows) and total sales by month, product category and product, with subtotals by month and product category and a grand total for all sales?

Oracle:

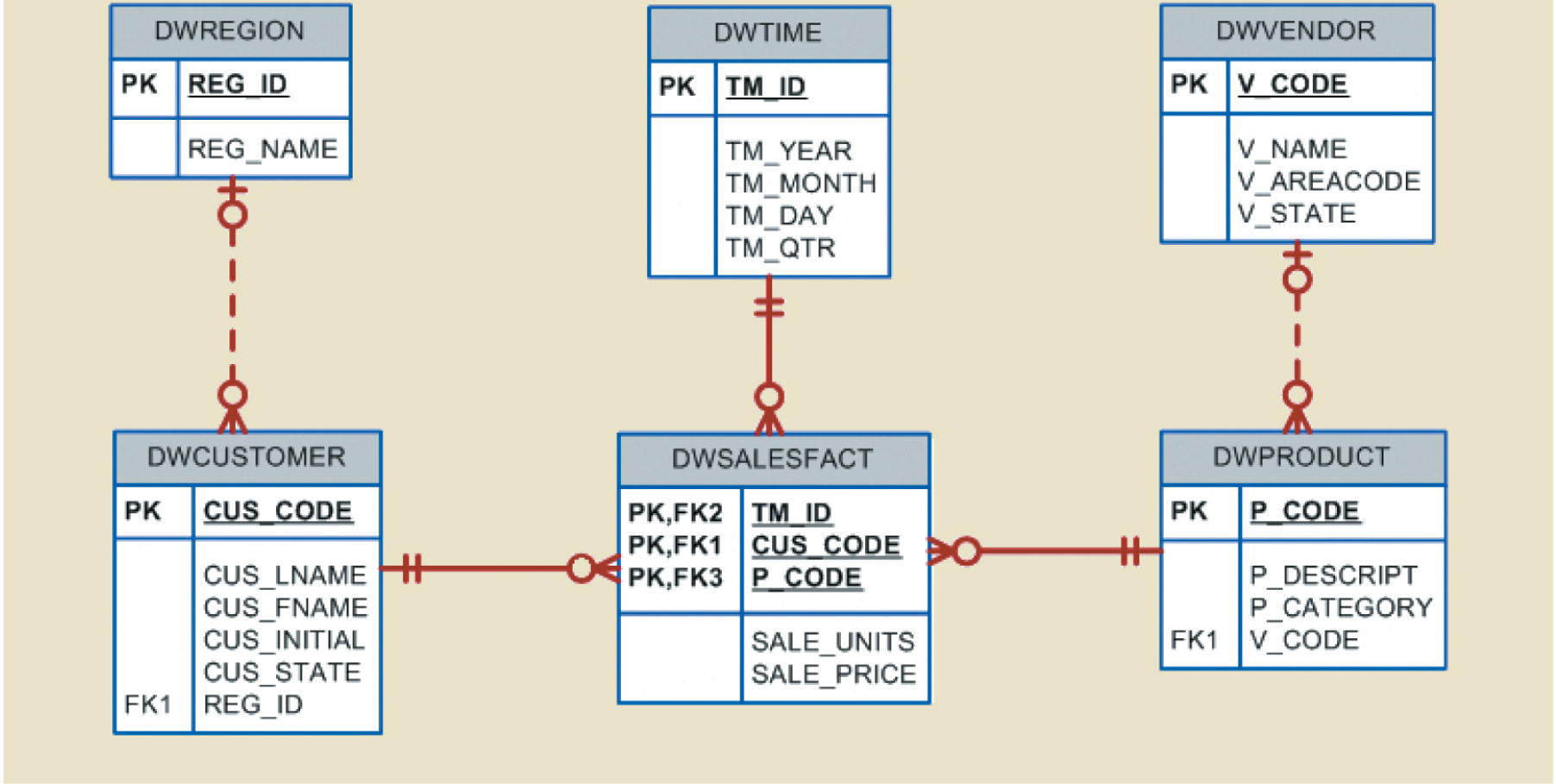
```
SELECT      TM_MONTH, P_CATEGORY, P_CODE, COUNT(*) AS NUMPROD,
            SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWTIME T ON S.TM_ID = T.TM_ID
            JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
GROUP BY    ROLLUP (TM_MONTH, P_CATEGORY, P_CODE);
```

SQL Server and MySQL:

```
SELECT      TM_MONTH, P_CATEGORY, P_CODE, COUNT(*) AS NUMPROD,
            SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWTIME T ON S.TM_ID = T.TM_ID
            JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
GROUP BY    TM_MONTH, P_CATEGORY, P_CODE WITH ROLLUP;
```

Using the answer to Problem 10 as your base, what command would you need to generate the same output but with subtotals in all columns? (*Hint*: Use the CUBE command).

FIGURE 13.19 SALECO SNOWFLAKE SCHEMA



Using the answer to Problem 10 as your base, what command would you need to generate the same output but with subtotals in all columns? (*Hint*: Use the CUBE command).

Oracle:

```
SELECT      TM_MONTH, P_CATEGORY, COUNT(*) AS NUMPROD,
            SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
            JOIN DWTIME T ON S.TM_ID = T.TM_ID
GROUP BY    CUBE (TM_MONTH, P_CATEGORY);
```

SQL Server:

```
SELECT      TM_MONTH, P_CATEGORY, COUNT(*) AS NUMPROD,
            SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
FROM        DWDAYSALESFACT S JOIN DWPRODUCT P ON S.P_CODE = P.P_CODE
            JOIN DWTIME T ON S.TM_ID = T.TM_ID
GROUP BY    TM_MONTH, P_CATEGORY WITH CUBE;
```

1. **Victoria Ephanor manages a small product distribution company. Because the business is growing fast, Ephanor recognizes that it is time to manage the vast information pool to help guide the accelerating growth. Ephanor, who is familiar with spreadsheet software, currently employs a sales force of four people. She asks you to develop a data warehouse application prototype that will enable her to study sales figures by year, region, salesperson, and product. (This prototype is to be used as the basis for a future data warehouse database.)**

   Using the data supplied in the Ch13_P2.xls file, complete the following seven problems:

   a. Identify the appropriate fact table components.

Year	Region	Agent	Product	Value
2012	East	Carlos	Erasers	50
2012	East	Tere	Erasers	12
2012	North	Carlos	Widgets	120
2012	North	Tere	Widgets	100
2012	North	Carlos	Widgets	30
2012	South	Victor	Balls	145
2012	South	Victor	Balls	34
2012	South	Victor	Balls	80
2012	West	Mary	Pencils	89
2012	West	Mary	Pencils	56
2013	East	Carlos	Pencils	45
2013	East	Victor	Balls	55
2013	North	Mary	Pencils	60
2013	North	Victor	Erasers	20
2013	South	Carlos	Widgets	30
2013	South	Mary	Widgets	75
2013	South	Mary	Widgets	50
2013	South	Tere	Balls	70
2013	South	Tere	Erasers	90
2013	West	Carlos	Widgets	25
2013	West	Tere	Balls	100

1. **Victoria Ephanor manages a small product distribution company. Because the business is growing fast, Ephanor recognizes that it is time to manage the vast information pool to help guide the accelerating growth. Ephanor, who is familiar with spreadsheet software, currently employs a sales force of four people. She asks you to develop a data warehouse application prototype that will enable her to study sales figures by year, region, salesperson, and product. (This prototype is to be used as the basis for a future data warehouse database.)**

   Using the data supplied in the Ch13_P2.xls file, complete the following seven problems:

   a. Identify the appropriate fact table components.

   The dimensions for this star schema are: Year, Region, Agent, and Product. (These are shown in Figure P13.2c.)

Year	Region	Agent	Product	Value
2012	East	Carlos	Erasers	50
2012	East	Tere	Erasers	12
2012	North	Carlos	Widgets	120
2012	North	Tere	Widgets	100
2012	North	Carlos	Widgets	30
2012	South	Victor	Balls	145
2012	South	Victor	Balls	34
2012	South	Victor	Balls	80
2012	West	Mary	Pencils	89
2012	West	Mary	Pencils	56
2013	East	Carlos	Pencils	45
2013	East	Victor	Balls	55
2013	North	Mary	Pencils	60
2013	North	Victor	Erasers	20
2013	South	Carlos	Widgets	30
2013	South	Mary	Widgets	75
2013	South	Mary	Widgets	50
2013	South	Tere	Balls	70
2013	South	Tere	Erasers	90
2013	West	Carlos	Widgets	25
2013	West	Tere	Balls	100

1. **Victoria Ephanor manages a small product distribution company. Because the business is growing fast, Ephanor recognizes that it is time to manage the vast information pool to help guide the accelerating growth. Ephanor, who is familiar with spreadsheet software, currently employs a sales force of four people. She asks you to develop a data warehouse application prototype that will enable her to study sales figures by year, region, salesperson, and product. (This prototype is to be used as the basis for a future data warehouse database.)**

   Using the data supplied in the Ch13_P2.xls file, complete the following seven problems:

   a. Identify the appropriate fact table components.

   The dimensions for this star schema are: Year, Region, Agent, and Product. (These are shown in Figure P13.2c.)

   b. Identify the appropriate dimension tables.

   (These are shown in Figure P13.2c.)

   c. Draw a star schema diagram for this data warehouse.

1. **Victoria Ephanor manages a small product distribution company. Because the business is growing fast, Ephanor recognizes that it is time to manage the vast information pool to help guide the accelerating growth. Ephanor, who is familiar with spreadsheet software, currently employs a sales force of four people. She asks you to develop a data warehouse application prototype that will enable her to study sales figures by year, region, salesperson, and product. (This prototype is to be used as the basis for a future data warehouse database.)**

Using the data supplied in the Ch13_P2.xls file, complete the following seven problems:

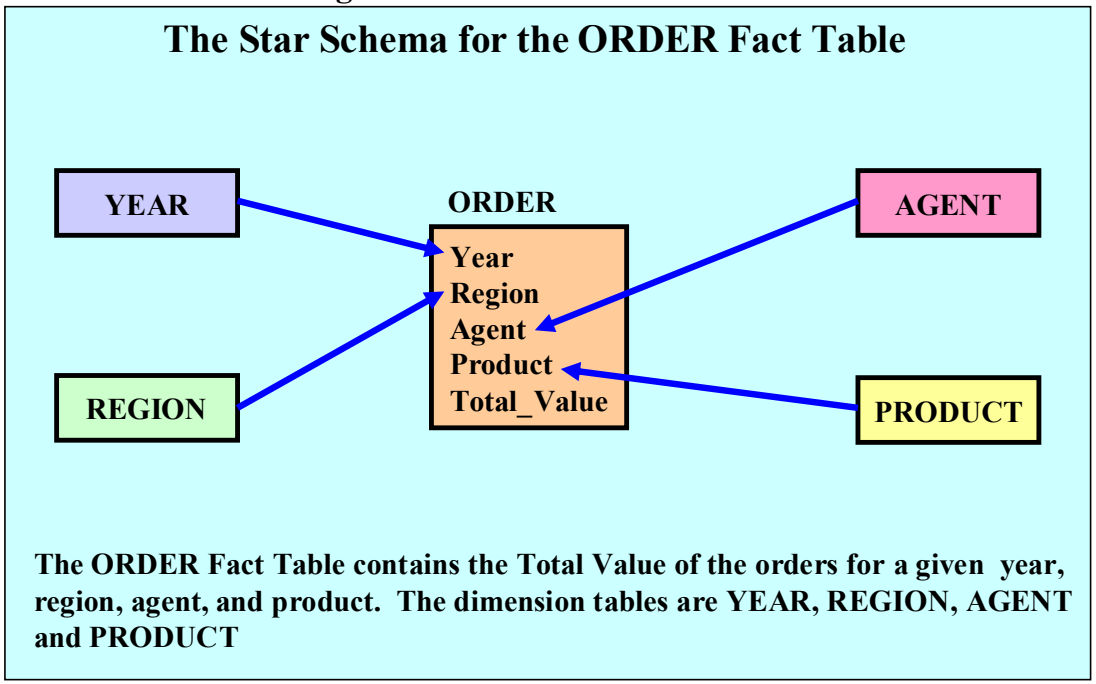
a. Identify the appropriate fact table components.

The dimensions for this star schema are: Year, Region, Agent, and Product. (These are shown in Figure P13.2c.)

b. Identify the appropriate dimension tables.

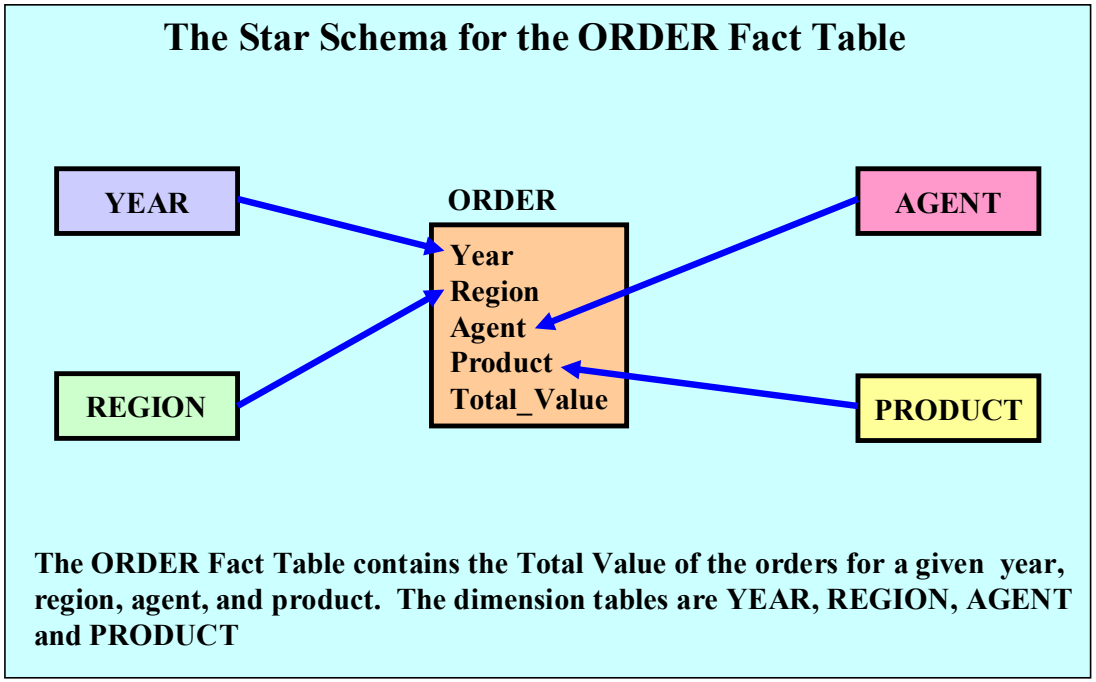
(These are shown in Figure P13.2c.)

c. Draw a star schema diagram for this data warehouse.



a. Identify the attributes for the dimension tables that will be required to solve this problem.

a. Identify the attributes for the dimension tables that will be required to solve this problem.



Using a Microsoft Excel or any other spreadsheet capable of producing pivot tables, generate a pivot table to show the sales by product and by region. The end user must be able to specify the display of sales for any given year. The sample output is shown in the first pivot table in Figure P13.2E.

	A	B	C	D	E	F
1	**Year**	2011				
2						
3	**Sum of Value**	**Region**				
4	**Product**	**East**	**North**	**South**	**West**	**Total**
5	Balls	55		70	100	**225**
6	Erasers		20	90		**110**
7	Pencils	45	60			**105**
8	Widgets			155	25	**180**
9	**Total**	**100**	**80**	**315**	**125**	**620**
10						
11						
12	**Year**	(All)				
13	**Product**	(All)				
14						
15	**Sum of Value**	**Region**				
16	**Agent**	**East**	**North**	**South**	**West**	**Total**
17	Carlos	95	150	30	25	**300**
18	Mary		60	125	145	**330**
19	Tere	12	100	160	100	**372**
20	Victor	55	20	259		**334**
21	**Total**	**162**	**330**	**574**	**270**	**1,336**

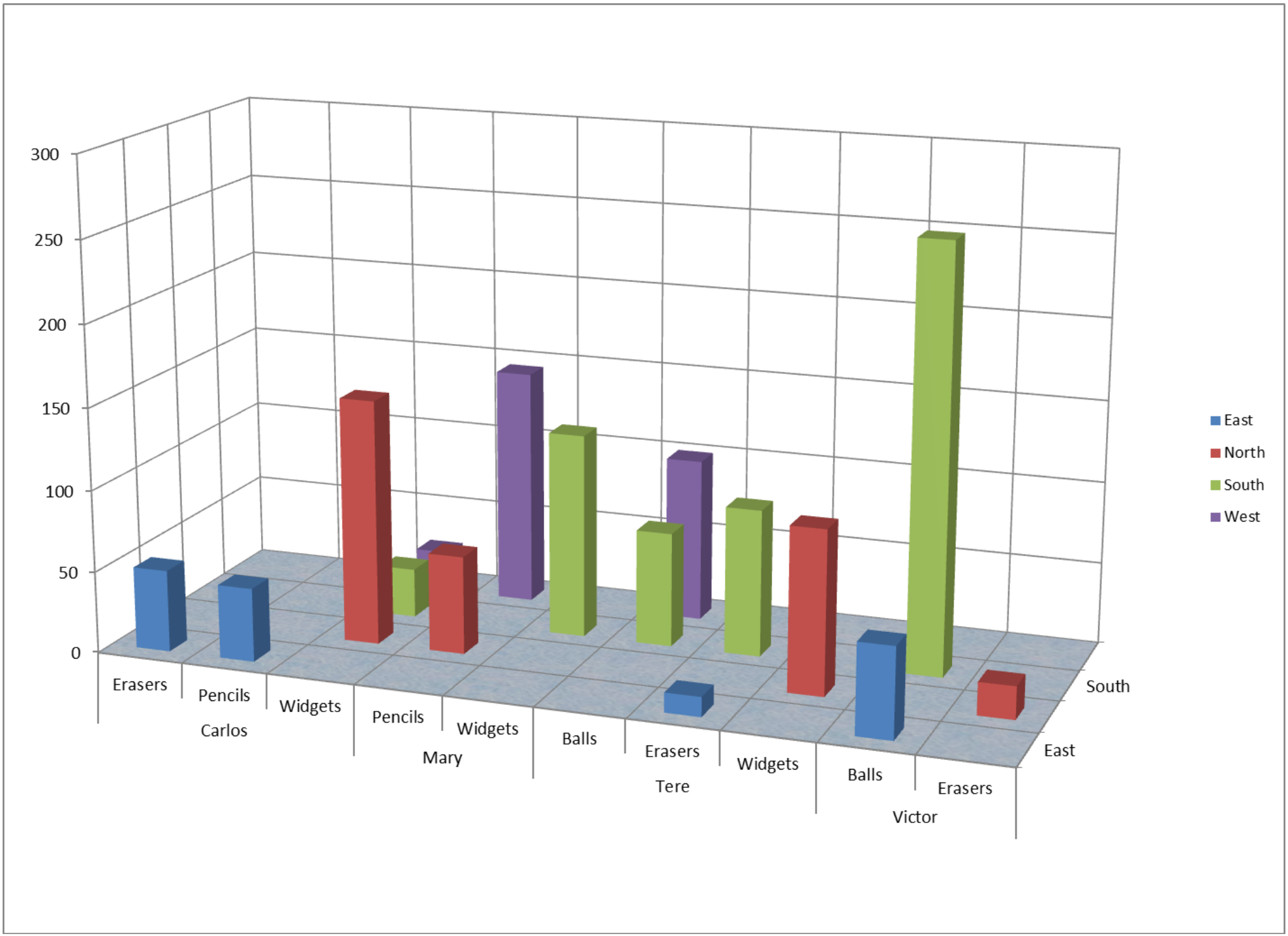
a. Using Problem 2e as your base, add a second pivot table (see Figure P13.2E) to show the sales by salesperson and by region. The end user must be able to specify sales for a given year or for all years, and for a given product or for all products.

a. Using Problem 2e as your base, add a second pivot table (see Figure P13.2E) to show the sales by salesperson and by region. The end user must be able to specify sales for a given year or for all years, and for a given product or for all products.

Sum of Value		Region				
Agent	Product	East	North	South	West	Grand Total
Carlos	Erasers	50				50
	Pencils	45				45
	Widgets		150	30	25	205
Carlos Total		95	150	30	25	300
Mary	Pencils		60		145	205
	Widgets			125		125
Mary Total			60	125	145	330
Tere	Balls			70	100	170
	Erasers	12		90		102
	Widgets		100			100
Tere Total		12	100	160	100	372
Victor	Balls	55		259		314
	Erasers		20			20
Victor Total		55	20	259		334
Grand Total		162	330	574	270	1336

Create a 3-D bar graph to show sales by salesperson, by product, and by region. (See the sample output in Figure P13.2G.)

Create a 3-D bar graph to show sales by salesperson, by product, and by region. (See the sample output in Figure P13.2G.)



1. **David Suker, the inventory manager for a marketing research company, wants to study the use of supplies within the different company departments. Suker has heard that his friend, Ephanor, has developed a spreadsheet-based data warehouse model that she uses in her analysis of sales data (See Problem 2). Suker is interested in developing a data warehouse model like Ephanor's so he can analyze orders by department and by product. He will use Microsoft Access as the Data Warehouse DBMS and Microsoft Excel as the analysis tool.**

   a. **Develop the order star schema.**

ORDERS

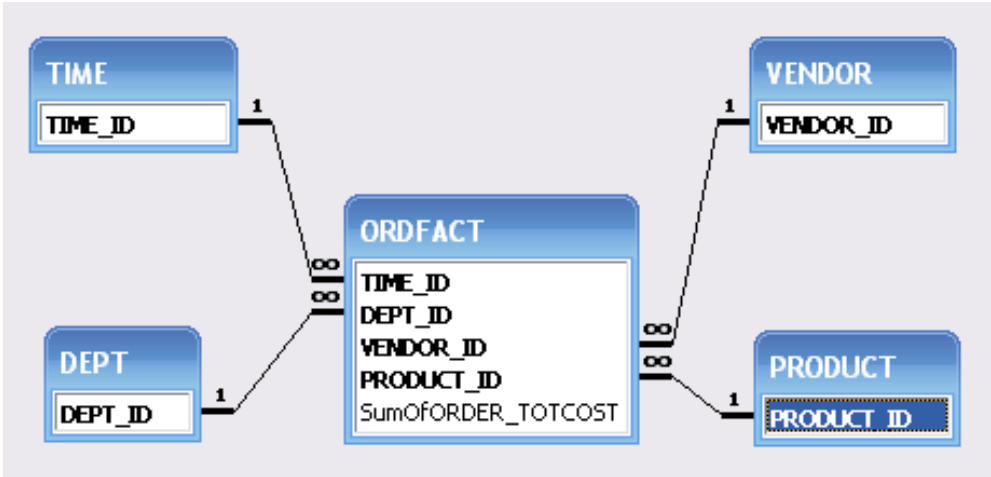
ORDER	ORDER_DAT	DEPT_I	VENDOR_	PRODUCT_ID	ORDER_TOTCOS
1	28-Aug-11	PRODUC	ADM	Pencil	33.00
2	29-Aug-11	PRODUC	GLOBAL	Paper	216.87
3	07-Sep-11	ACCTNG	GLOBAL	Ribbons	178.49
4	07-Sep-11	ACCTNG	GLOBAL	Disks	116.95
5	02-Apr-11	ACCTNG	GLOBAL	Toners	110.47
6	11-Apr-11	ACCTNG	GLOBAL	Labels	136.50
7	17-Apr-11	ACCTNG	ADM	Paper	78.00
8	27-Feb-11	ACCTNG	GLOBAL	Paper	1335.00
9	19-Sep-11	ACCTNG	XTT	Labels	189.00
10	09-Oct-11	ACCTNG	GLOBAL	Labels	201.80
11	06-Oct-11	ACCTNG	MCC	Labels	252.00
12	13-Oct-11	ACCTNG	XTT	Pencil	52.35
13	24-Oct-11	ACCTNG	MCC	Paper	63.00
14	25-Oct-11	ACCTNG	GLOBAL	Labels	169.19
15	31-Oct-11	ACCTNG	MCC	Labels	55.00
16	06-Nov-11	DESIGN	XTT	Paper	210.00
17	13-Nov-11	ACCTNG	XTT	Labels	297.00
18	14-Dec-11	ACCTNG	GLOBAL	Labels	287.79
19	25-Dec-11	ACCTNG	GLOBAL	Labels	61.02
20	10-Jan-12	ACCTNG	GLOBAL	Ribbons	288.29
21	18-Jan-12	ACCTNG	GLOBAL	Labels	193.57
22	23-Jan-12	PRODUC	GLOBAL	Labels	92.19
23	24-Jan-12	PRODUC	MCC	Ribbons	299.00
24	26-Jan-12	PRODUC	BCC	Labels	146.50
25	26-Jan-12	PRODUC	GLOBAL	Ribbons	289.88
26	30-Jan-12	PRODUC	XTT	Toners	300.00

Record: 1 of 330 Unfiltered Search

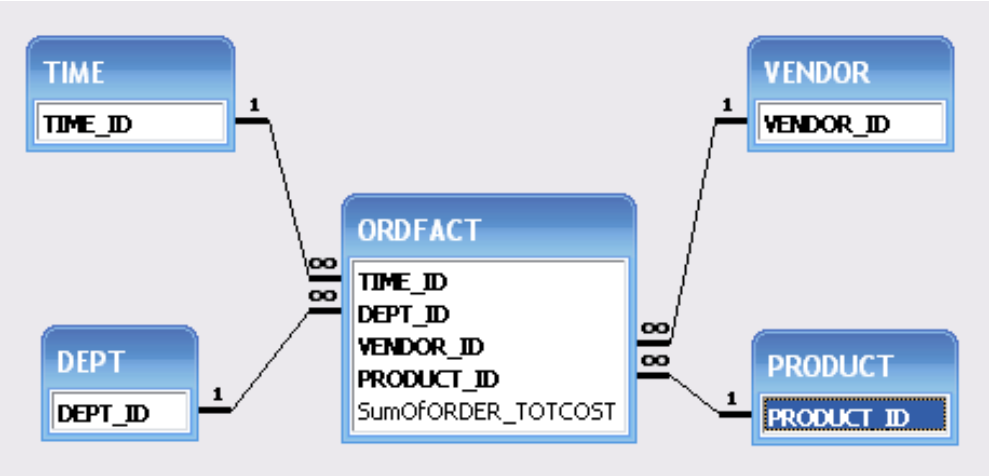
1. **David Suker, the inventory manager for a marketing research company, wants to study the use of supplies within the different company departments. Suker has heard that his friend, Ephanor, has developed a spreadsheet-based data warehouse model that she uses in her analysis of sales data (See Problem 2). Suker is interested in developing a data warehouse model like Ephanor's so he can analyze orders by department and by product. He will use Microsoft Access as the Data Warehouse DBMS and Microsoft Excel as the analysis tool.**

   a. **Develop the order star schema.**

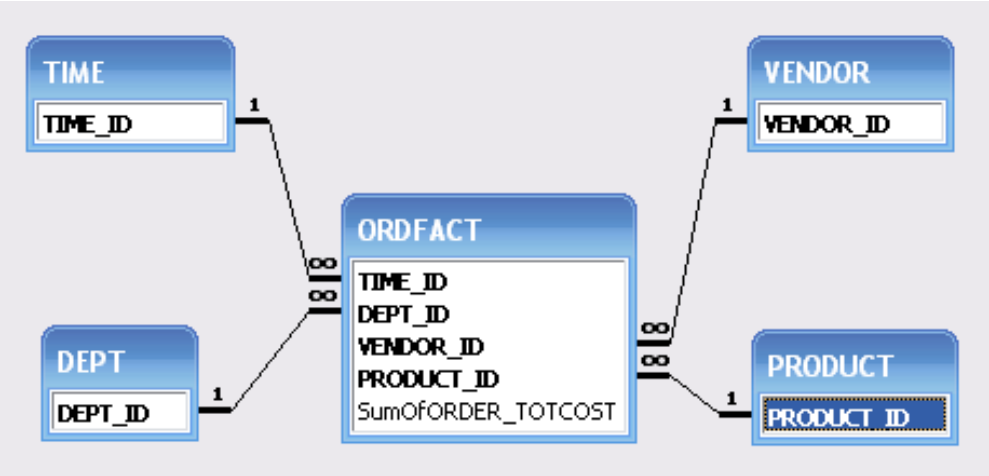
      Figure P13-3A's MS-Access relational diagram reflects the star schema and its relationships. Note that the students are given only the ORDERS table. The student must study the data set and make the queries necessary to create the dimension tables (TIME, DEPT, VENDOR and PRODUCT) and the ORDFACT fact table.



a. Identify the appropriate dimension attributes.

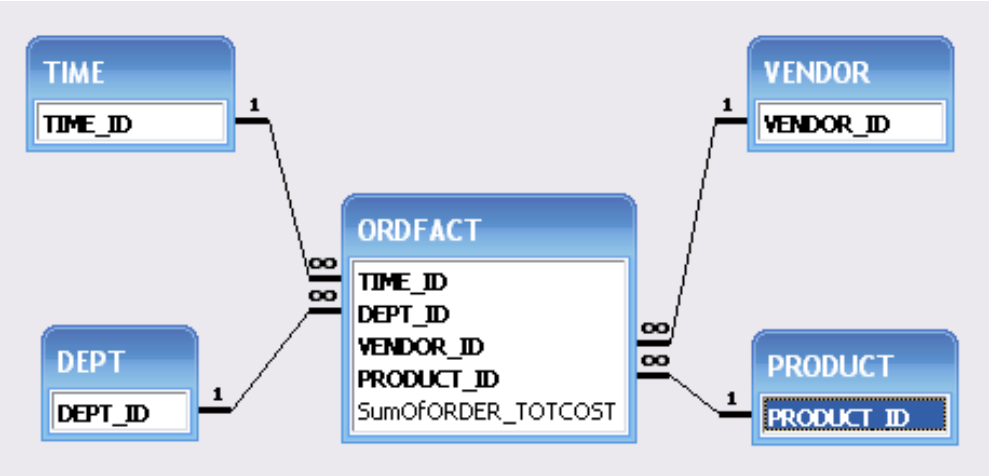


a. Identify the appropriate dimension attributes.



The dimensions are: TIME, DEPT, VENDOR, and PRODUCT. (See Figure P13.3A.)

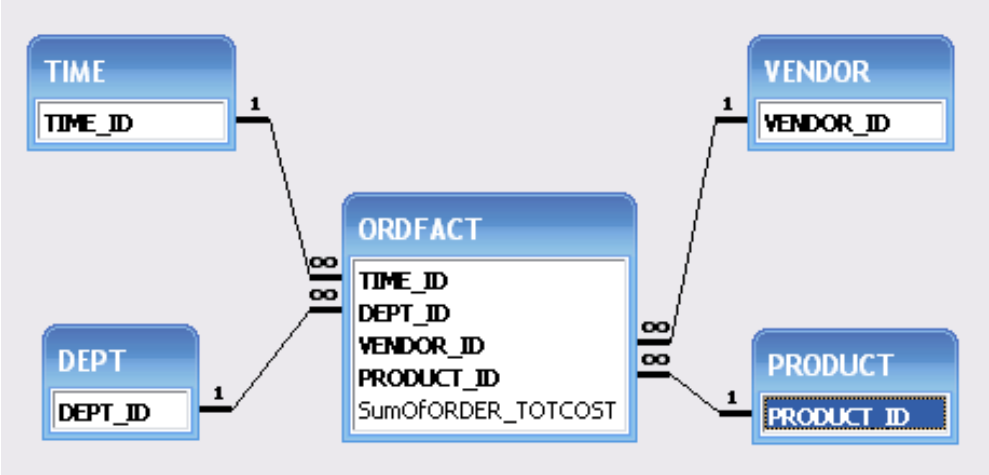
a. Identify the attribute hierarchies required to support the model.



a. Identify the attribute hierarchies required to support the model.

The main hierarchy used for data drilling purposes is represented by TIME-DEPT-VENDOR-PRODUCT sequence. (See Figure P13.3A.) Within this hierarchy, the user can analyze data at different aggregation levels.

Additional hierarchies can be constructed in the TIME dimension to account for quarters or, if necessary, by daily aggregates. The VENDOR dimension could also be expanded to include geographic information that could be used for drill-down purposes.



a. Develop a crosstab report (in Microsoft Access), using a 3-D bar graph to show sales by product and by department. (The sample output is shown in Figure P13.3.)

FIGURE P13.3 A Crosstab Report: Sales by Product and Department

Orders by Department and Product

Friday, January 31, 2014 1/31/2014 9:50:16 AM

Product	Accounting	Design	Marketing	Production	Row Total
Disks	$1,429.79	$17,268.80	$13,510.09	$6,312.24	**$38,520.92**
Envelopes	$329.16	$1,165.79	$17,074.33	$4,517.91	**$23,087.19**
Labels	$3,651.76	$1,514.15	$2,356.72	$8,464.79	**$15,987.42**
Paper	$1,761.90	$5,246.74	$14,222.35	$3,928.99	**$25,159.98**
Pencil	$741.83	$1,585.21	$2,014.56	$1,370.30	**$5,711.90**
Ribbons	$1,916.92	$525.00	$1,873.21	$3,203.82	**$7,518.95**
Toners	$110.47	$448.55	$358.25	$1,589.57	**$2,506.84**
	$9,941.83	**$27,754.24**	**$51,409.51**	**$29,387.62**	**$118,493.20**