Chapter 13

Business Intelligence and Data Warehouses

Learning Objectives

- In this chapter, you will learn:
  - How business intelligence provides a comprehensive business decision support framework
  - About business intelligence architecture, its evolution, and reporting styles
  - About the relationship and differences between operational data and decision support data
  - What a data warehouse is and how to prepare data for one

Learning Objectives

- In this chapter, you will learn:
  - What star schemas are and how they are constructed
  - About data analytics
  - About online analytical processing (OLAP)
  - How SQL extensions are used to support OLAP-type data manipulations

Business Intelligence (BI)

- Comprehensive, cohesive, integrated set of tools and processes
  - Captures, collects, integrates, stores, and analyzes data
- Generates and presents information to support business decision making
- Allows a business to transform:
  - Data into information
  - Information into knowledge
  - Knowledge into wisdom

Business Intelligence (BI)

- Concepts, practices, tools and techniques to help business
  - Understand its core capabilities
  - Provide snapshots of the company situation
  - Identify key opportunities to create a competitive advantage
- Provides a framework for
  - Collecting and storing operational data and aggregating it into decision support data
  - Analyzing decision support data and presenting generated information to end users to support business decisions
  - Making business decision which generates more data
  - Monitoring results to evaluate outcomes and predicting future outcomes with a high degree of accuracy

Figure 13.1 - Business Intelligence Framework

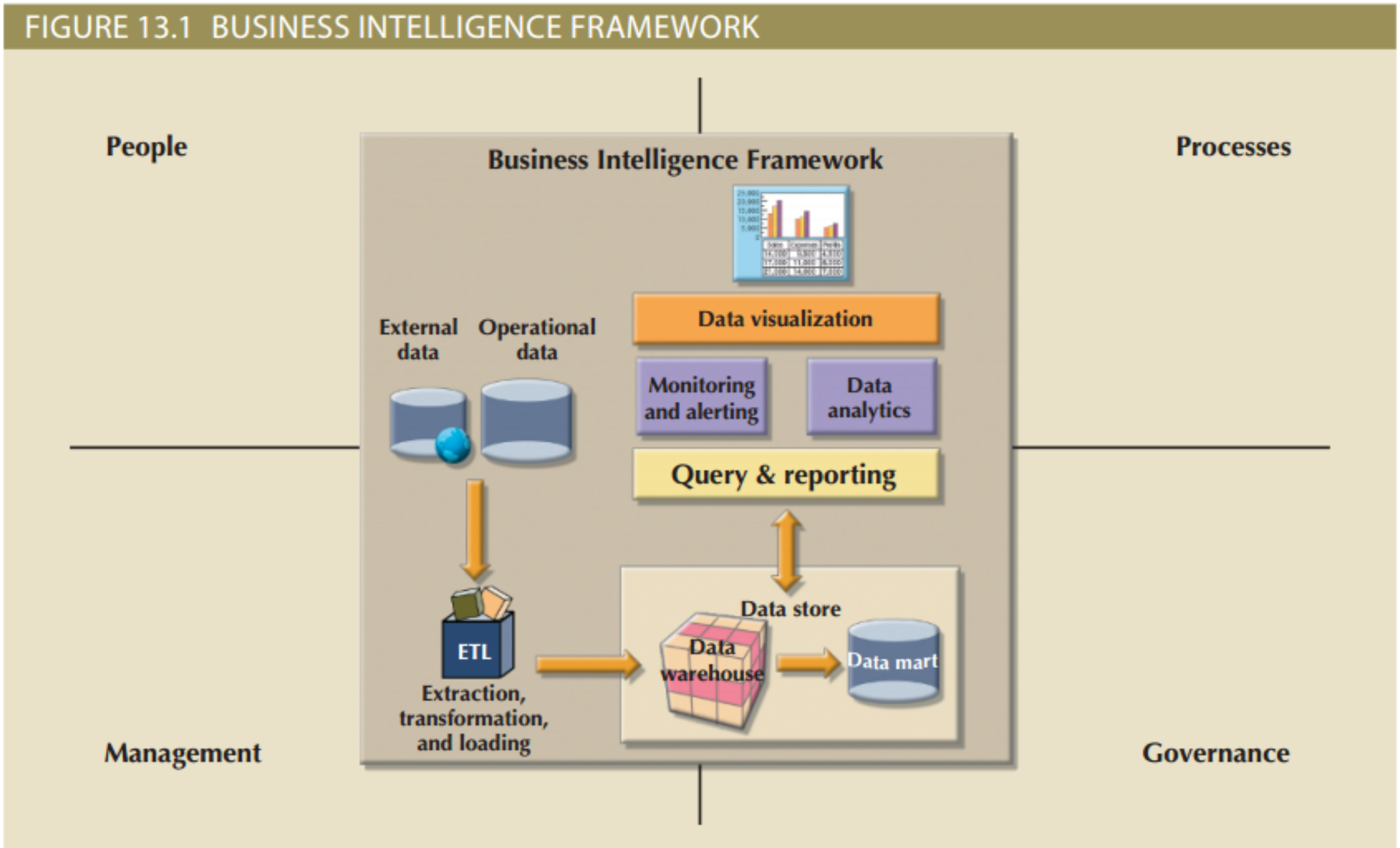


Table 13.3 – Sample of Business Intelligence Tools

TABLE 13.3

SAMPLE OF BUSINESS INTELLIGENCE TOOLS

TOOL	DESCRIPTION	SAMPLE VENDORS
Dashboards and business activity monitoring	**Dashboards** use web-based technologies to present key business performance indicators or information in a single integrated view, generally using graphics that are clear, concise, and easy to understand.	Salesforce IBM/Cognos BusinessObjects Information Builders iDashboards
Portals	**Portals** provide a unified, single point of entry for information distribution. Portals are a web-based technology that use a web browser to integrate data from multiple sources into a single webpage. Many different types of BI functionality can be accessed through a portal.	Oracle Portal Actuate Microsoft SAP
Data analysis and reporting tools	These advanced tools are used to query multiple and diverse data sources to create integrated reports.	Microsoft Reporting Services MicroStrategy SAS WebReportStudio
Data-mining tools	These tools provide advanced statistical analysis to uncover problems and opportunities hidden within business data. Chapter 14 covers data mining in more detail.	SAP Teradata MicroStrategy MS Analytics Services

Table 13.3 – Sample of Business Intelligence Tools (cont'd)

TABLE 13.3		
SAMPLE OF BUSINESS INTELLIGENCE TOOLS		
TOOL	**DESCRIPTION**	**SAMPLE VENDORS**
Data warehouses (DW)	The data warehouse is the foundation of a BI infrastructure. Data is captured from the production system and placed in the DW on a near real-time basis. BI provides company-wide integration of data and the capability to respond to business issues in a timely manner.	Microsoft Oracle IBM/Cognos Teradata
OLAP tools	Online analytical processing provides multidimensional data analysis.	IBM/Cognos BusinessObjects Oracle Microsoft
Data visualization	These tools provide advanced visual analysis and techniques to enhance understanding and create additional insight of business data and its true meaning.	Dundas Tableau QlikView Actuate

Practices to Manage Data

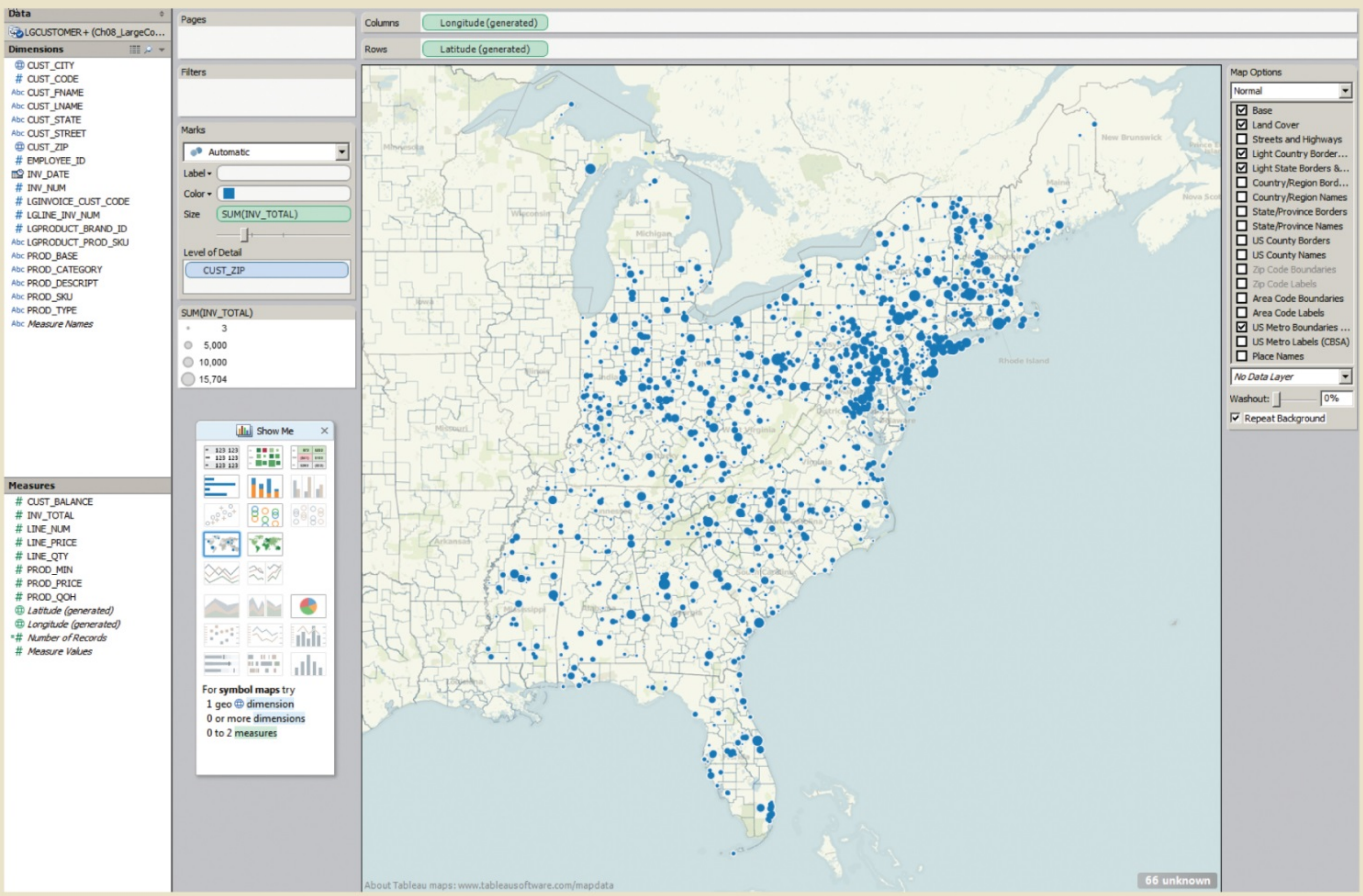
- **Master data management (MDM)**: Collection of concepts, techniques, and processes for identification, definition, and management of data elements
- **Governance**: Method of government for controlling business health and for consistent decision making
- **Key performance indicators (KPI)**: Numeric or scale-based measurements that assess company's effectiveness in reaching its goals

Practices to Manage Data

- **Data visualization**: Abstracting data to provide information in a visual format
  - Enhances the user's ability to efficiently comprehend the meaning of the data
  - Techniques:
    - Pie charts and bar charts
    - Line graphs
    - Scatter plots
    - Gantt charts
    - Heat maps

Figure 13.2 – Visualizing Sales Total by ZIP Code

FIGURE 13.2 VISUALIZING SALES TOTAL BY zip CODE



Courtesy of Tableau

Reporting Styles of a Modern BI System

Advanced reporting

Monitoring and alerting

Advanced data analytics

Business Intelligence Benefits

- Improved decision making
- Integrating architecture
- Common user interface for data reporting and analysis
- Common data repository fosters single version of company data
- Improved organizational performance

Table 13.4 - Business Intelligence Evolution

TABLE 13.4

BUSINESS INTELLIGENCE EVOLUTION

SYSTEM TYPE	DATA SOURCE	DATA EXTRACTION/ INTEGRATION PROCESS	DATA STORE	END-USER QUERY TOOL	END USER PRESENTATION TOOL
Traditional mainframe-based online transaction processing (OLTP)	Operational data	None Reports read and summarized data directly from operational data	None Temporary files used for reporting purposes	Very basic Predefined reporting formats Basic sorting, totaling, and averaging	Very basic Menu-driven, predefined reports, text and numbers only
Managerial information system (MIS)	Operational data	Basic extraction and aggregation Read, filter, and summarize operational data into intermediate data store	Lightly aggregated data in RDBMS	Same as above, in addition to some ad hoc reporting using SQL	Same as above, in addition to some ad hoc columnar report definitions
First-generation departmental decision support system (DSS)	Operational data External data	Data extraction and integration process populates DSS data store Run periodically	First DSS database generation Usually RDBMS	Query tool with some analytical capabilities and predefined reports	Spreadsheet style Advanced presentation tools with plotting and graphics capabilities

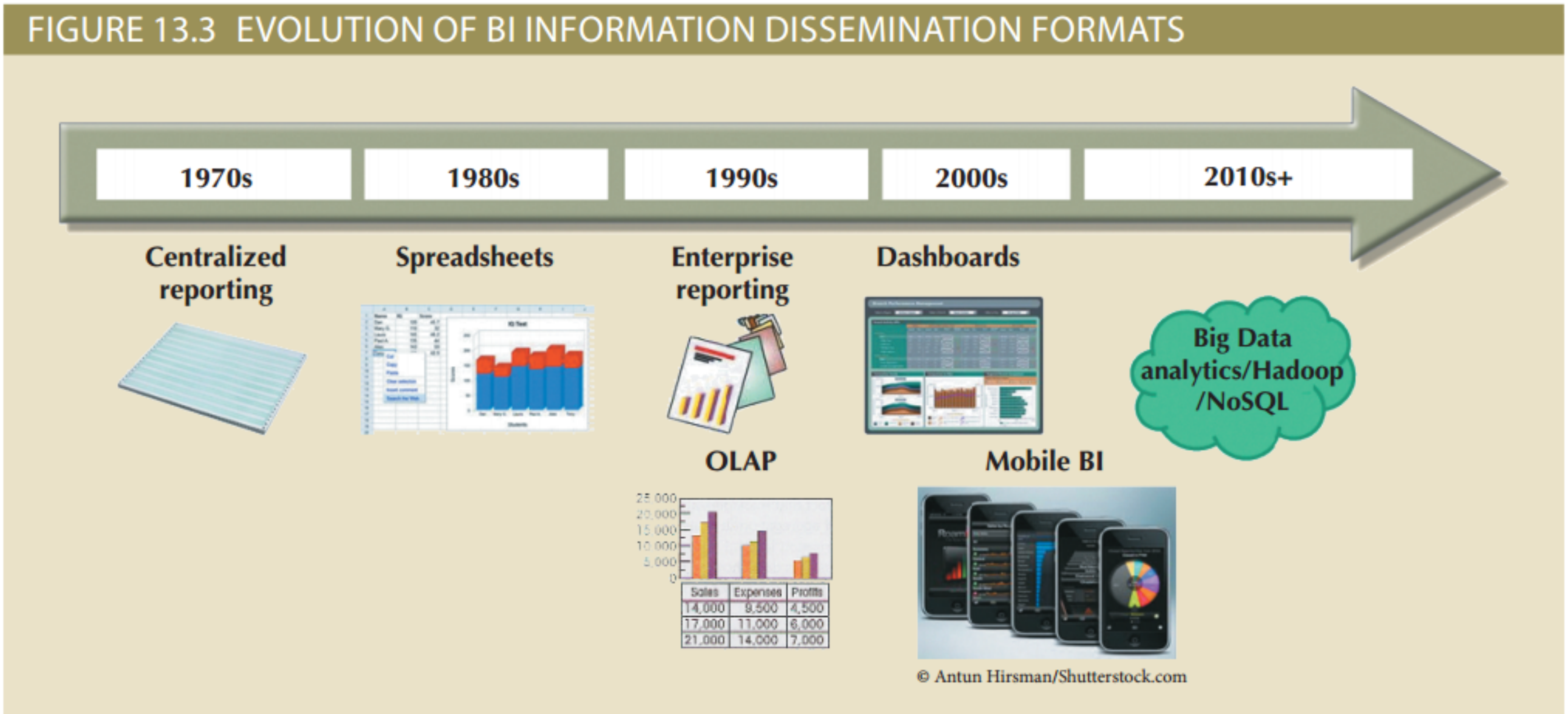
Table 13.4 - Business Intelligence Evolution (cont'd)

TABLE 13.4

BUSINESS INTELLIGENCE EVOLUTION

SYSTEM TYPE	DATA SOURCE	DATA EXTRACTION/ INTEGRATION PROCESS	DATA STORE	END-USER QUERY TOOL	END USER PRESENTATION TOOL
First-generation BI	Operational data External data	Advanced data extraction and integration Access diverse data sources, filters, aggregations, classifications, scheduling, and conflict resolution	Data warehouse RDBMS technology Optimized for query purposes Star schema model	Same as above	Same as above, in addition to multidimensional presentation tools with drill-down capabilities
Second-generation BI Online analytical processing (OLAP)	Same as above	Same as above	Data warehouse stores data in MDBMS Cubes with multiple dimensions	Adds support for end-user-based data analytics	Same as above, but uses cubes and multidimensional matrixes; limited by cube size Dashboards Scorecards Portals
Third-generation Mobile, cloud-based, and Big Data	Same as above Includes social media and machine-generated data	Same as above Cloud-based	Same as above Cloud-based Hadoop and NoSQL databases	Advanced analytics Limited ad hoc interactions	Mobile devices: smartphones and tablets

Figure 13.3 - Evolution of BI Information Dissemination Formats



Business Intelligence Technology Trends

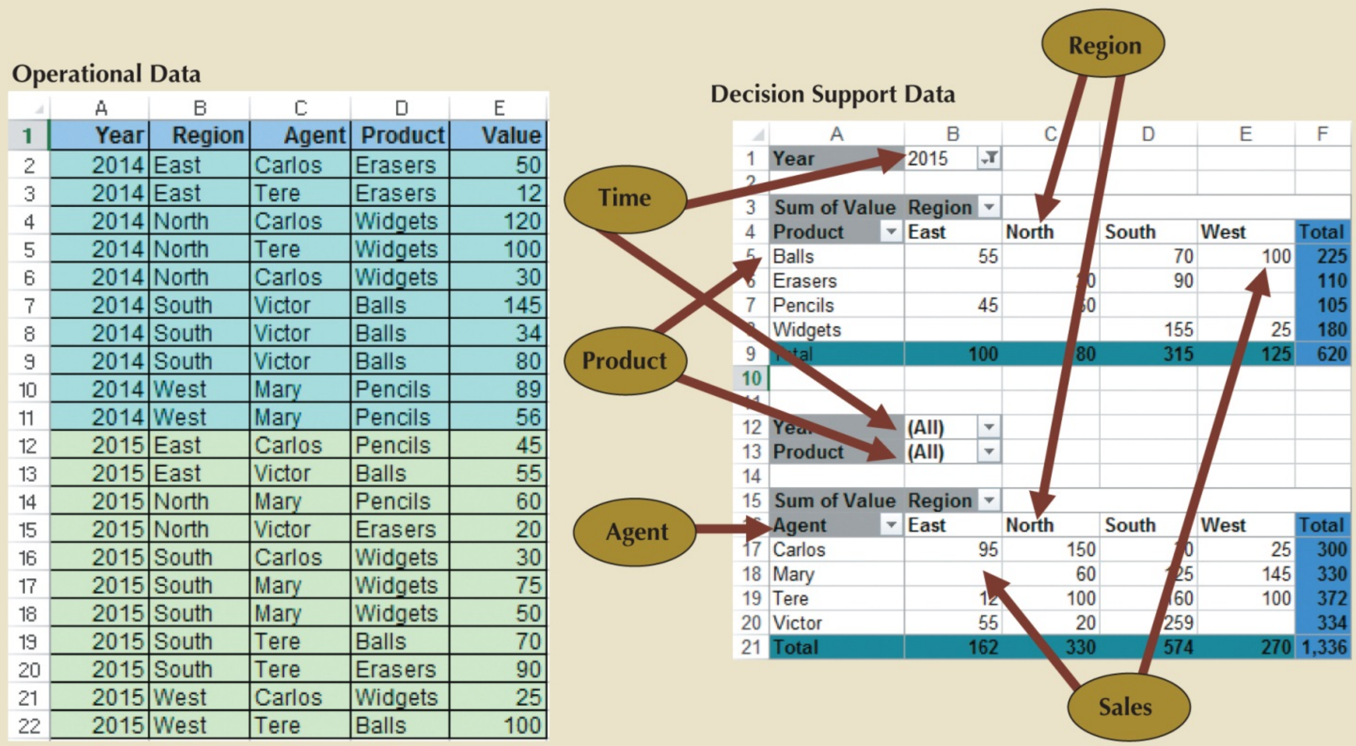
- Data storage improvements
- Business intelligence appliances
- Business intelligence as a service
- Big Data analytics
- Personal analytics

Decision Support Data

- Effectiveness of BI depends on quality of data gathered at operational level
- Operational data
  - Seldom well-suited for decision support tasks
  - Stored in relational database with highly normalized structures
  - Optimized to support transactions representing daily operations

Figure 13.4 –Transforming Operational Data into Decision Support Data

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

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		[illegible]	90		110
7	Pencils	45	[illegible]			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	[illegible]	25	300
18	Mary		60	[illegible]	145	330
19	Tere	12	100	[illegible]	100	372
20	Victor	55	20	259		334
21	Total	162	330	574	270	1,336

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 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

Decision Support Data

- Differ from operational data in:
  - Time span
  - Granularity
    - **Drill down**: Decomposing a data to a lower level
    - **Roll up**: Aggregating a data into a higher level
  - Dimensionality

Table 13.5 - Contrasting Operational and Decision Support Data Characteristics

TABLE 13.5

CONTRASTING OPERATIONAL AND DECISION SUPPORT DATA CHARACTERISTICS

CHARACTERISTIC	OPERATIONAL DATA	DECISION SUPPORT DATA
Data currency	Current operations Real-time data	Historic data Snapshot of company data Time component (week/month/year)
Granularity	Atomic-detailed data	Summarized data
Summarization level	Low; some aggregate yields	High; many aggregation levels
Data model	Highly normalized Mostly relational DBMSs	Non-normalized Complex structures Some relational, but mostly multidimensional DBMSs
Transaction type	Mostly updates	Mostly query
Transaction volumes	High-update volumes	Periodic loads and summary calculations
Transaction speed	Updates are critical	Retrievals are critical
Query activity	Low to medium	High
Query scope	Narrow range	Broad range
Query complexity	Simple to medium	Very complex
Data volumes	Hundreds of gigabytes	Terabytes to petabytes

Decision Support Database Requirements

- Database schema
  - Must support complex, non-normalized data representations
  - Data must be aggregated and summarized
  - Queries must be able to extract multidimensional time slices

Decision Support Database Requirements

- Data extraction and loading
  - Allow batch and scheduled data extraction
  - Support different data sources and check for inconsistent data or data validation rules
  - Support advanced integration, aggregation, and classification
- Database size should support
  - **Very large databases (VLDBs)**
  - Advanced storage technologies
  - Multiple-processor technologies

Table 13.6 – Ten-Year Sales History for a Single Department, in Millions of Dollars

TABLE 13.6

TEN-YEAR SALES HISTORY FOR A SINGLE DEPARTMENT, IN MILLIONS OF DOLLARS

YEAR	SALES
2006	8,227
2007	9,109
2008	10,104
2009	11,553
2010	10,018
2011	11,875
2012	12,699
2013	14,875
2014	16,301
2015	19,986

Table 13.8 - Characteristics of Data Warehouse Data and Operational Database Data

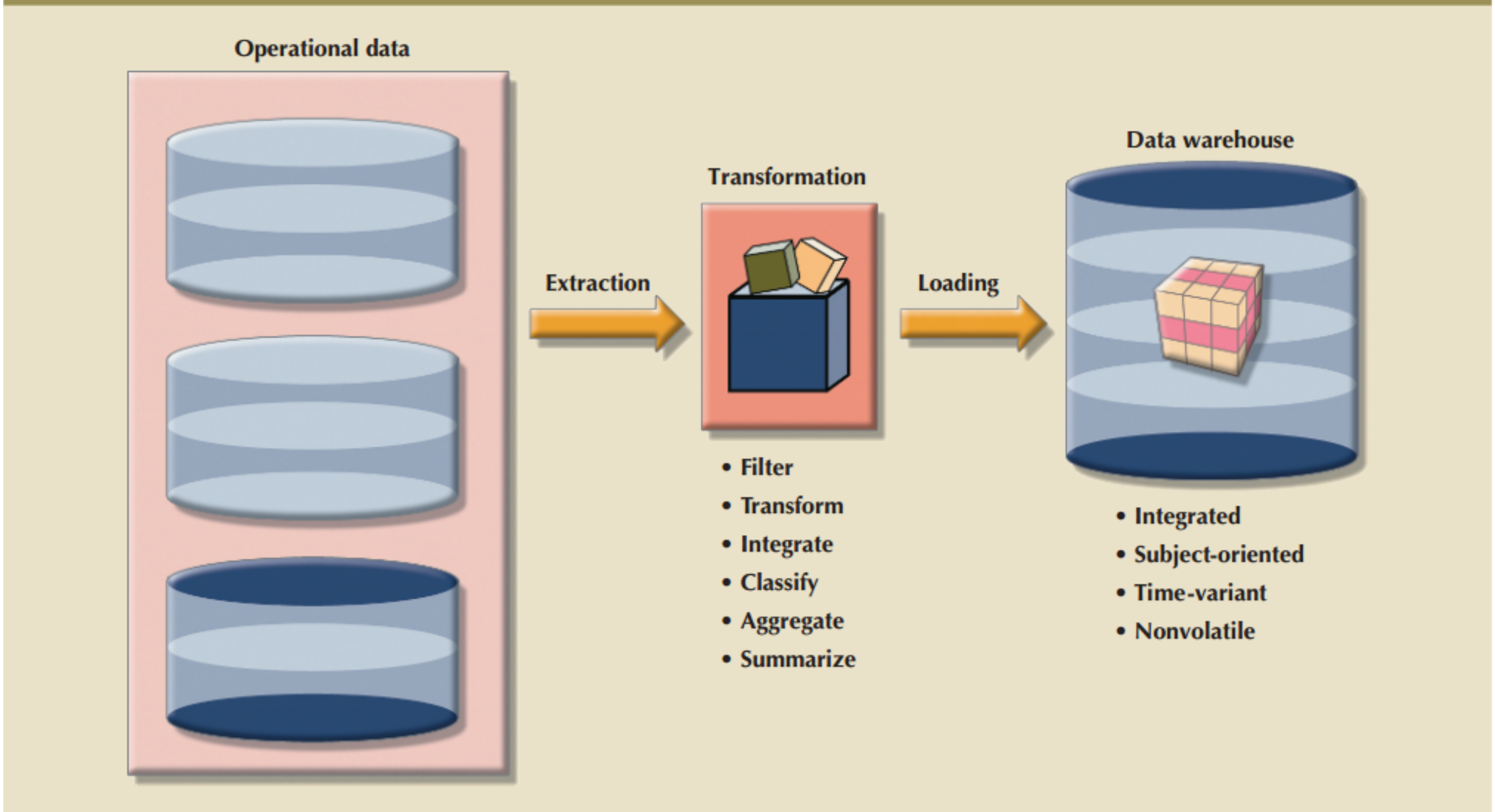
TABLE 13.8

CHARACTERISTICS OF DATA WAREHOUSE DATA AND OPERATIONAL DATABASE DATA

CHARACTERISTIC	OPERATIONAL DATABASE DATA	DATA WAREHOUSE DATA
Integrated	Similar data can have different representations or meanings. For example, Social Security numbers may be stored as ###-##-#### or as #########, and a given condition may be labeled as T/F or 0/1 or Y/N. A sales value may be shown in thousands or in millions.	Provide a unified view of all data elements with a common definition and representation for all business units.
Subject-oriented	Data is stored with a functional, or process, orientation. For example, data may be stored for invoices, payments, and credit amounts.	Data is stored with a subject orientation that facilitates multiple views of the data and decision making. For example, sales may be recorded by product, division, manager, or region.
Time-variant	Data is recorded as current transactions. For example, the sales data may be the sale of a product on a given date, such as $342.78 on 12-MAY-2016.	Data is recorded with a historical perspective in mind. Therefore, a time dimension is added to facilitate data analysis and various time comparisons.
Nonvolatile	Data updates are frequent and common. For example, an inventory amount changes with each sale. Therefore, the data environment is fluid.	Data cannot be changed. Data is added only periodically from historical systems. Once the data is properly stored, no changes are allowed. Therefore, the data environment is relatively static.

Figure 13.5 - The ETL Process

FIGURE 13.5 THE ETL PROCESS



Cengage Learning © 2015

Data Marts

- Small, single-subject data warehouse subset
- Provide decision support to a small group of people
- Benefits over data warehouses
  - Lower cost and shorter implementation time
  - Technologically advanced
  - Inevitable people issues

Table 13.9 - Twelve Rules for a Data Warehouse

TABLE 13.9

TWELVE RULES FOR A DATA WAREHOUSE

RULE NO.	DESCRIPTION
1	The data warehouse and operational environments are separated.
2	The data warehouse data is integrated.
3	The data warehouse contains historical data over a long time.
4	The data warehouse data is snapshot data captured at a given point in time.
5	The data warehouse data is subject oriented.
6	The data warehouse data is mainly read-only with periodic batch updates from operational data. No online updates are allowed.
7	The data warehouse development life cycle differs from classical systems development. Data warehouse development is data-driven; the classical approach is process-driven.
8	The data warehouse contains data with several levels of detail: current detail data, old detail data, lightly summarized data, and highly summarized data.
9	The data warehouse environment is characterized by read-only transactions to very large data sets. The operational environment is characterized by numerous update transactions to a few data entities at a time.
10	The data warehouse environment has a system that traces data sources, transformations, and storage.
11	The data warehouse's metadata is a critical component of this environment. The metadata identifies and defines all data elements. The metadata provides the source, transformation, integration, storage, usage, relationships, and history of each data element.
12	The data warehouse contains a chargeback mechanism for resource usage that enforces optimal use of the data by end users.

Star Schema

- Data-modeling technique
- Maps multidimensional decision support data into a relational database
- Creates the near equivalent of multidimensional database schema from existing relational database
- Yields an easily implemented model for multidimensional data analysis

Figure 13.6 – Simple Star Schema

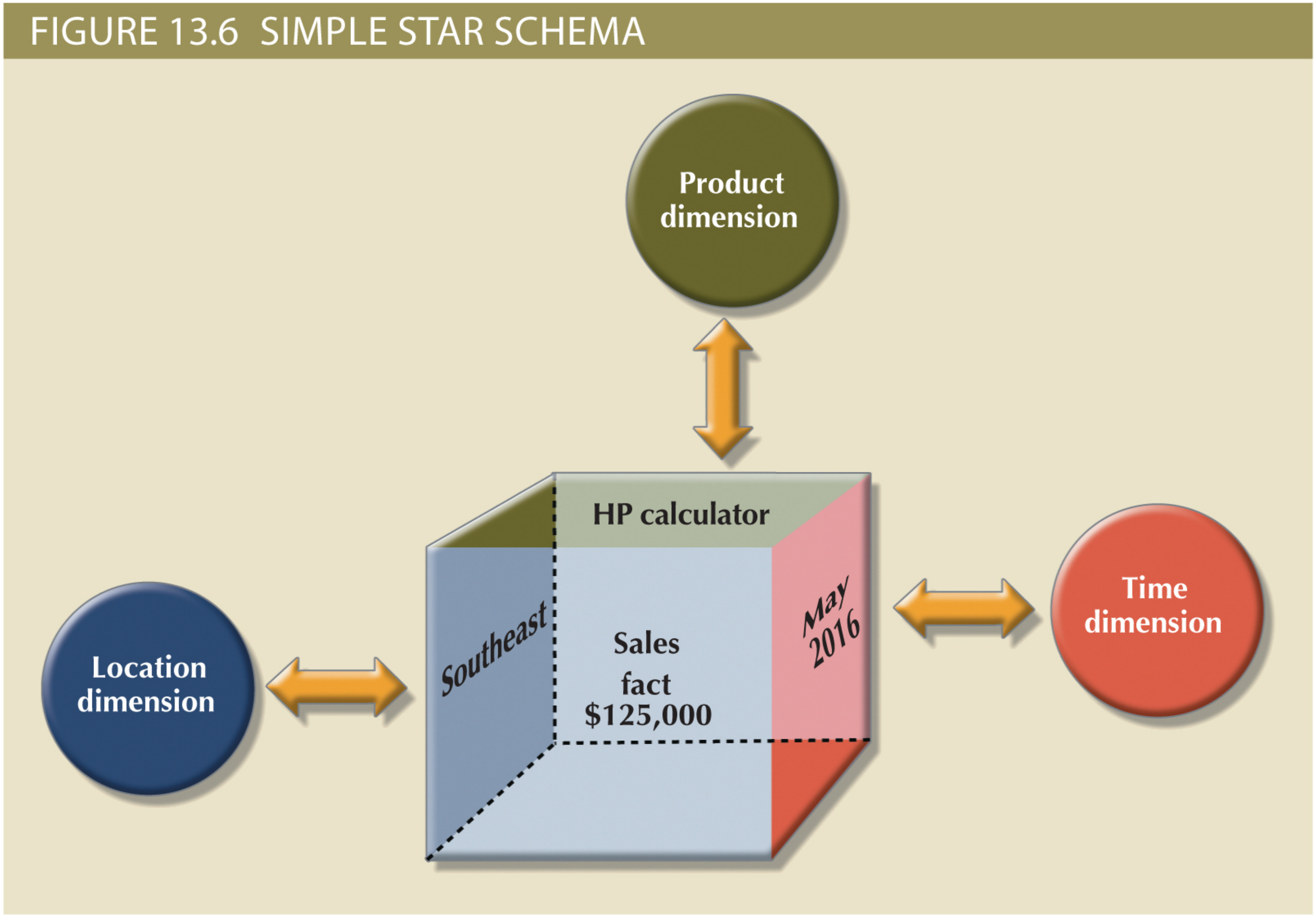


Figure 13.7 – Three-Dimensional View of Sales

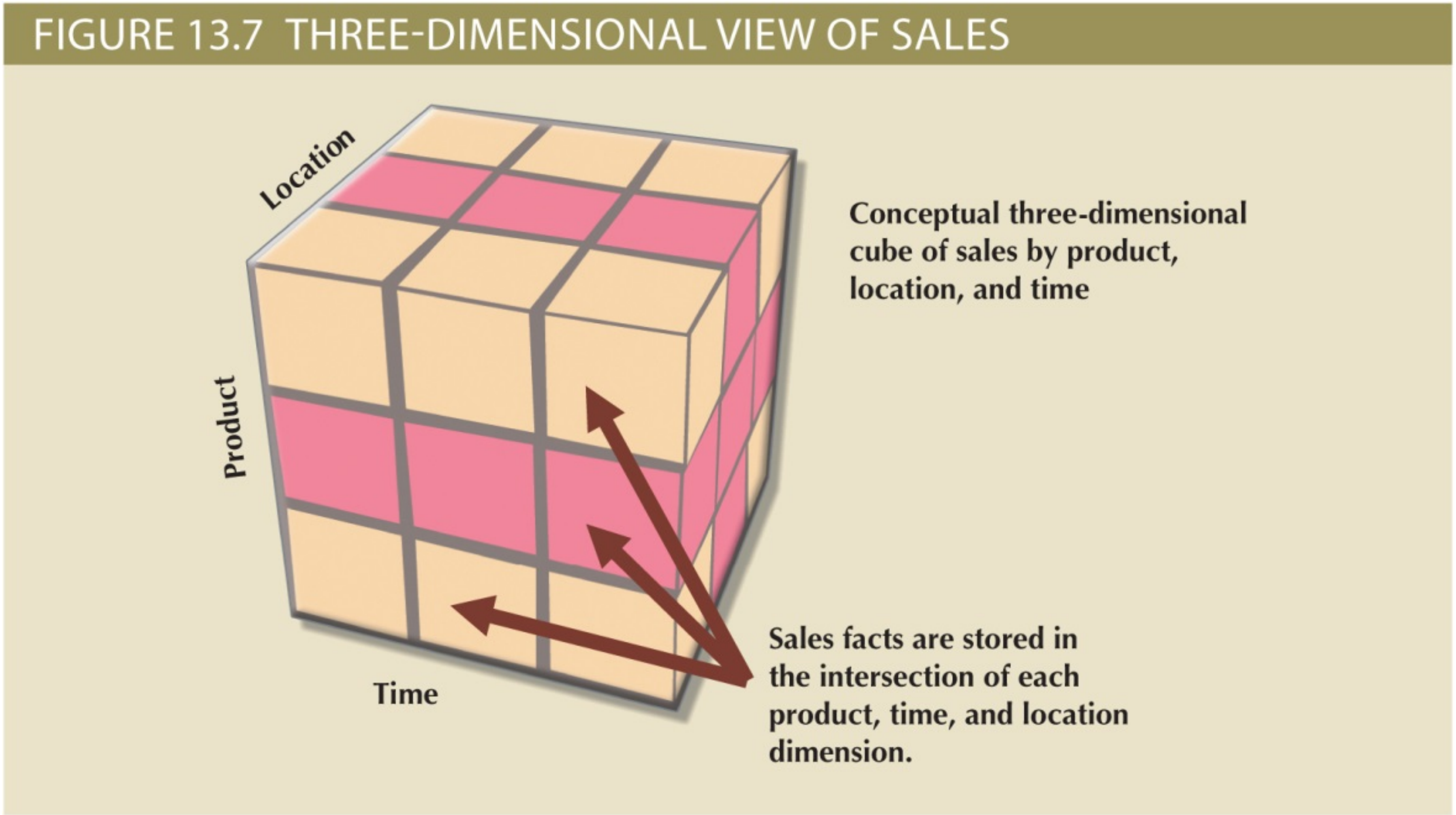
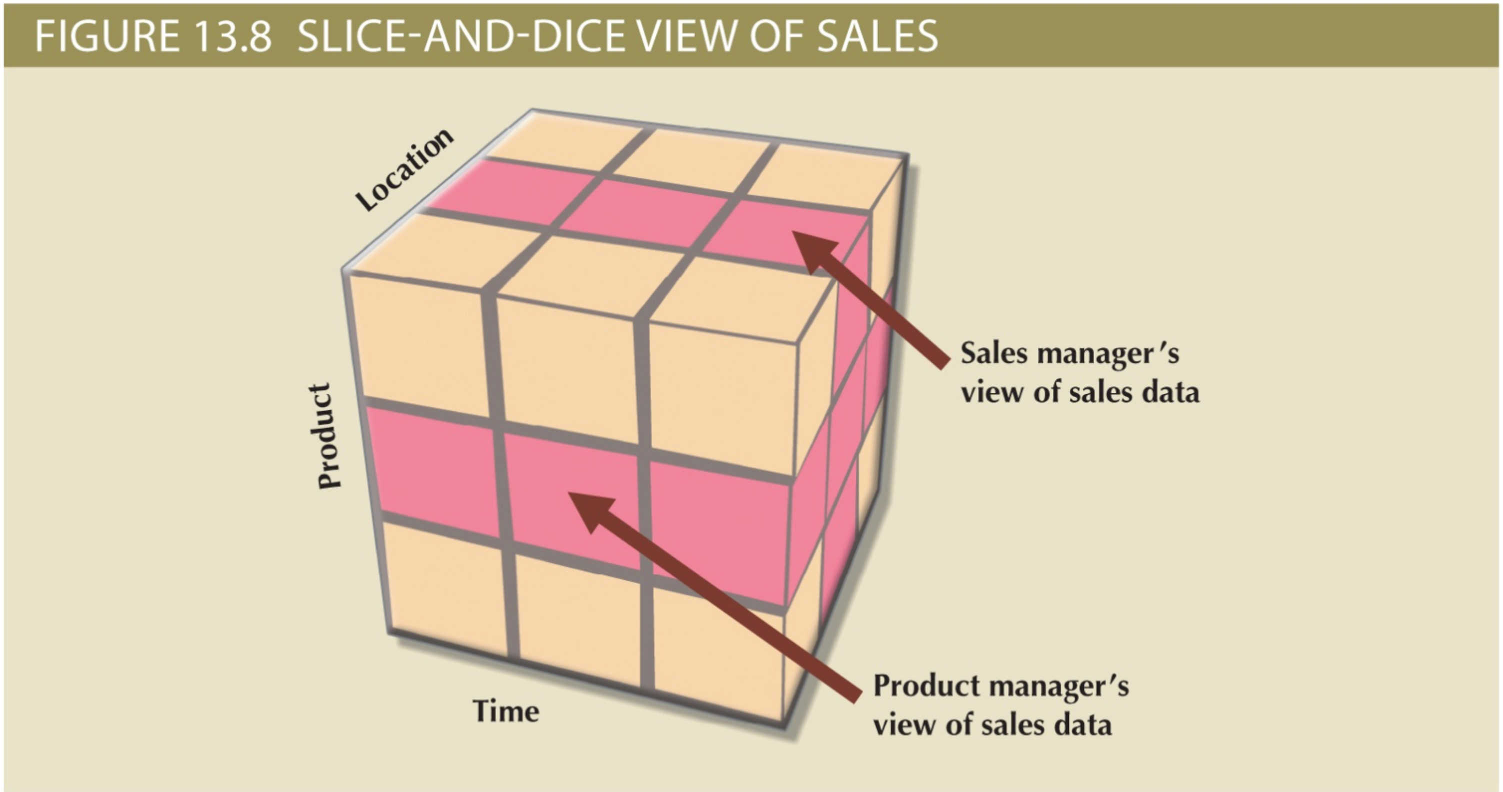


Figure 13.8 – Slice-and-Dice View of Sales



Components of Star Schemas

Facts

- Numeric values that represent a specific business aspect

Dimensions

- Qualifying characteristics that provide additional perspectives to a given fact

Attributes

- Used to search, filter, and classify facts
- **Slice and dice**: Ability to focus on slices of the data cube for more detailed analysis

Attribute hierarchies

- Provides a top-down data organization

Figure 13.9 – Location Attribute Hierarchy

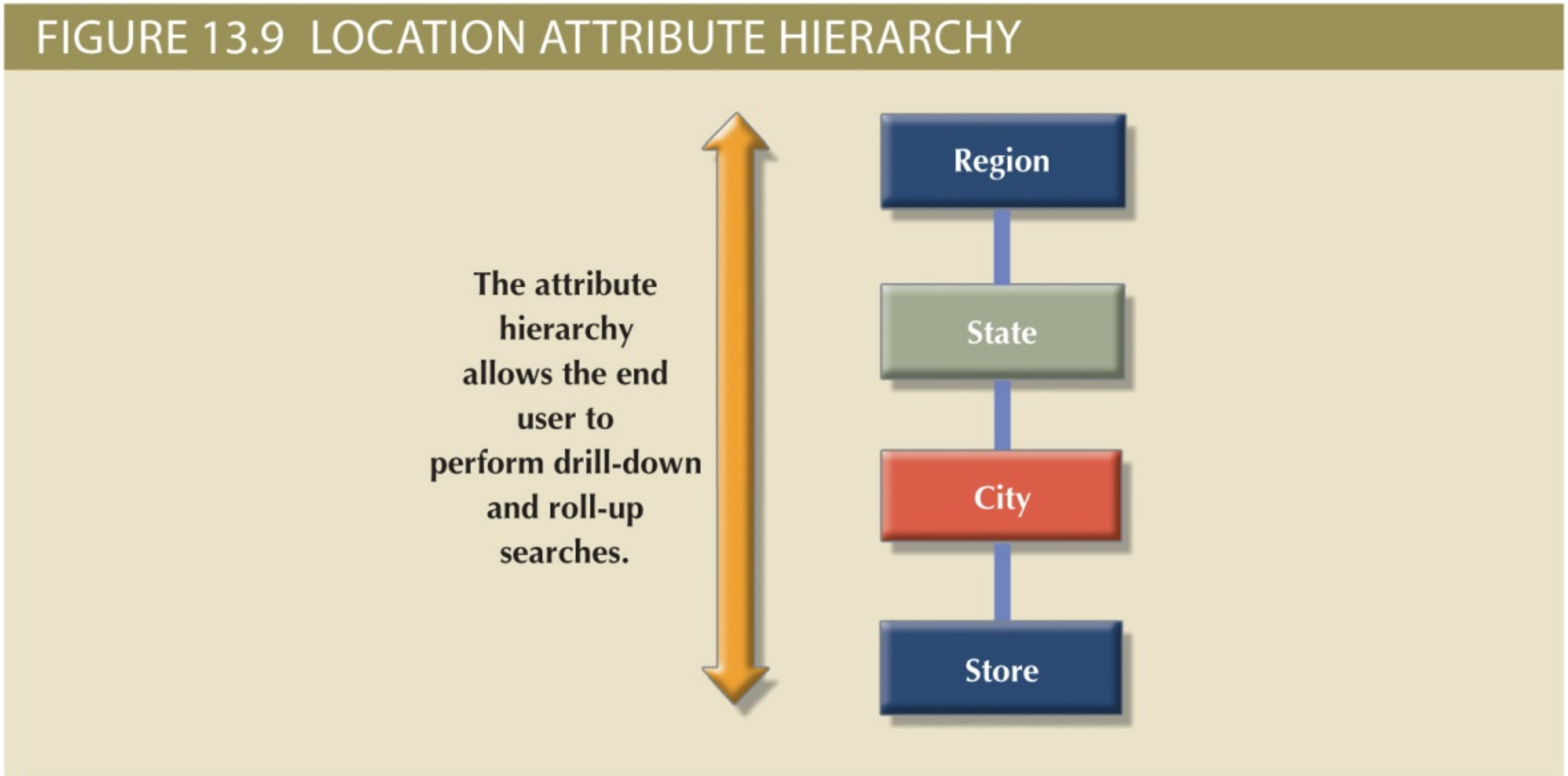
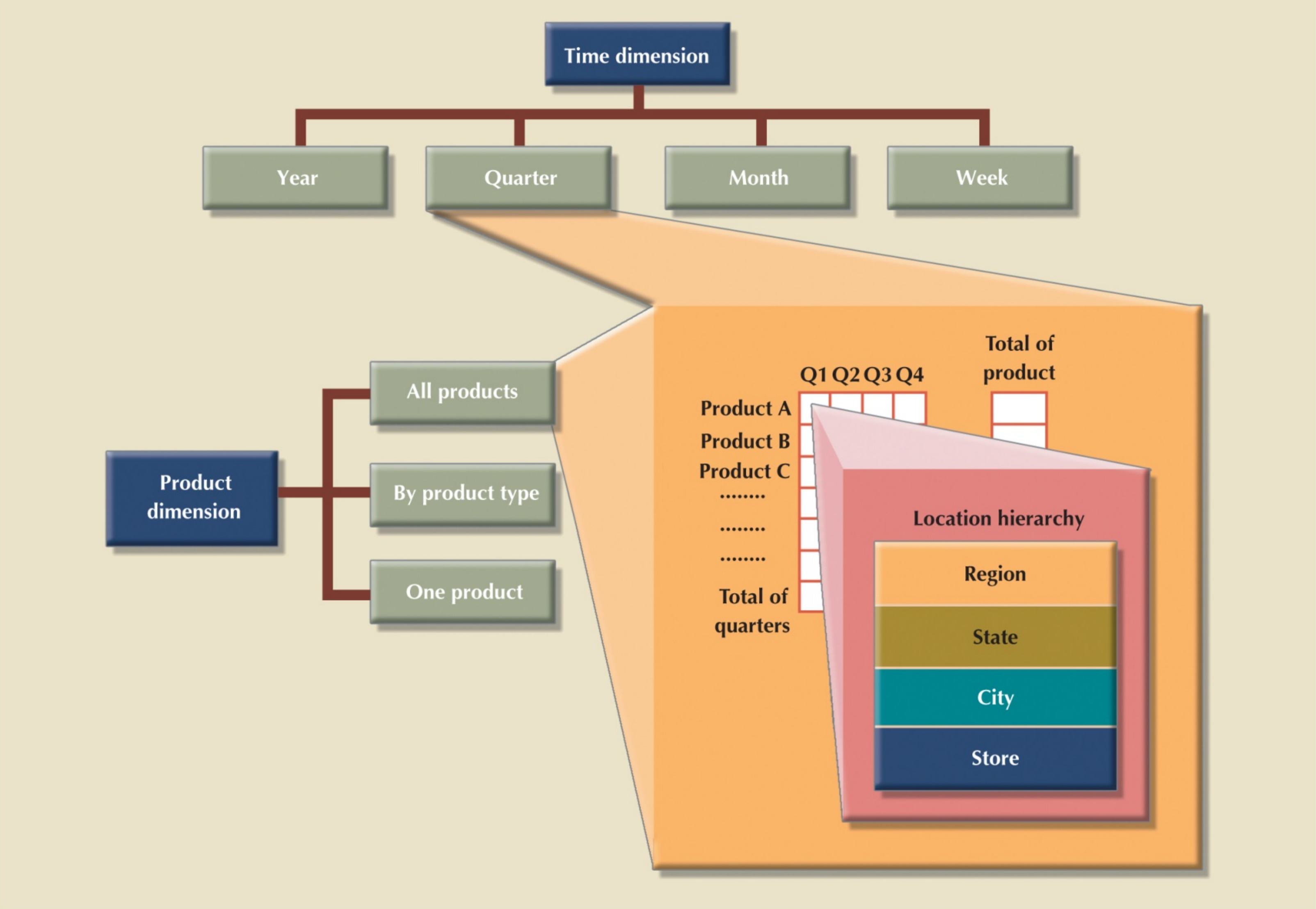


Figure 13.10 – Attribute Hierarchies in Multidimensional Analysis

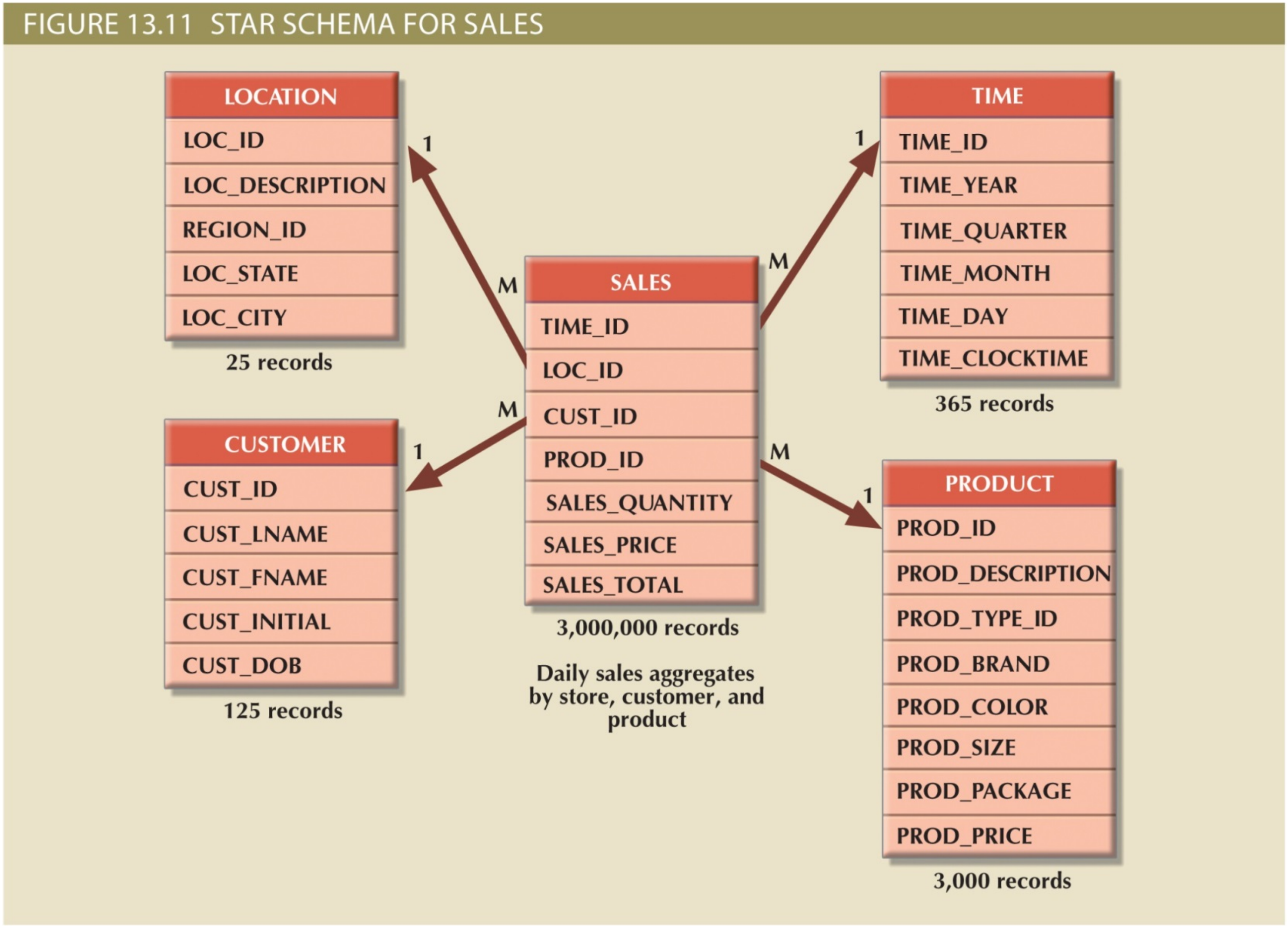
FIGURE 13.10 ATTRIBUTE HIERARCHIES IN MULTIDIMENSIONAL ANALYSIS



Star Schema Representation

- Facts and dimensions represented by physical tables in data warehouse database
- Many-to-one (M:1) relationship between fact table and each dimension table
- Fact and dimension tables
  - Related by foreign keys
  - Subject to primary and foreign key constraints
- Primary key of a fact table
  - Is a composite primary key because the fact table is related to many dimension tables
  - Always formed by combining the foreign keys pointing to the related dimension tables

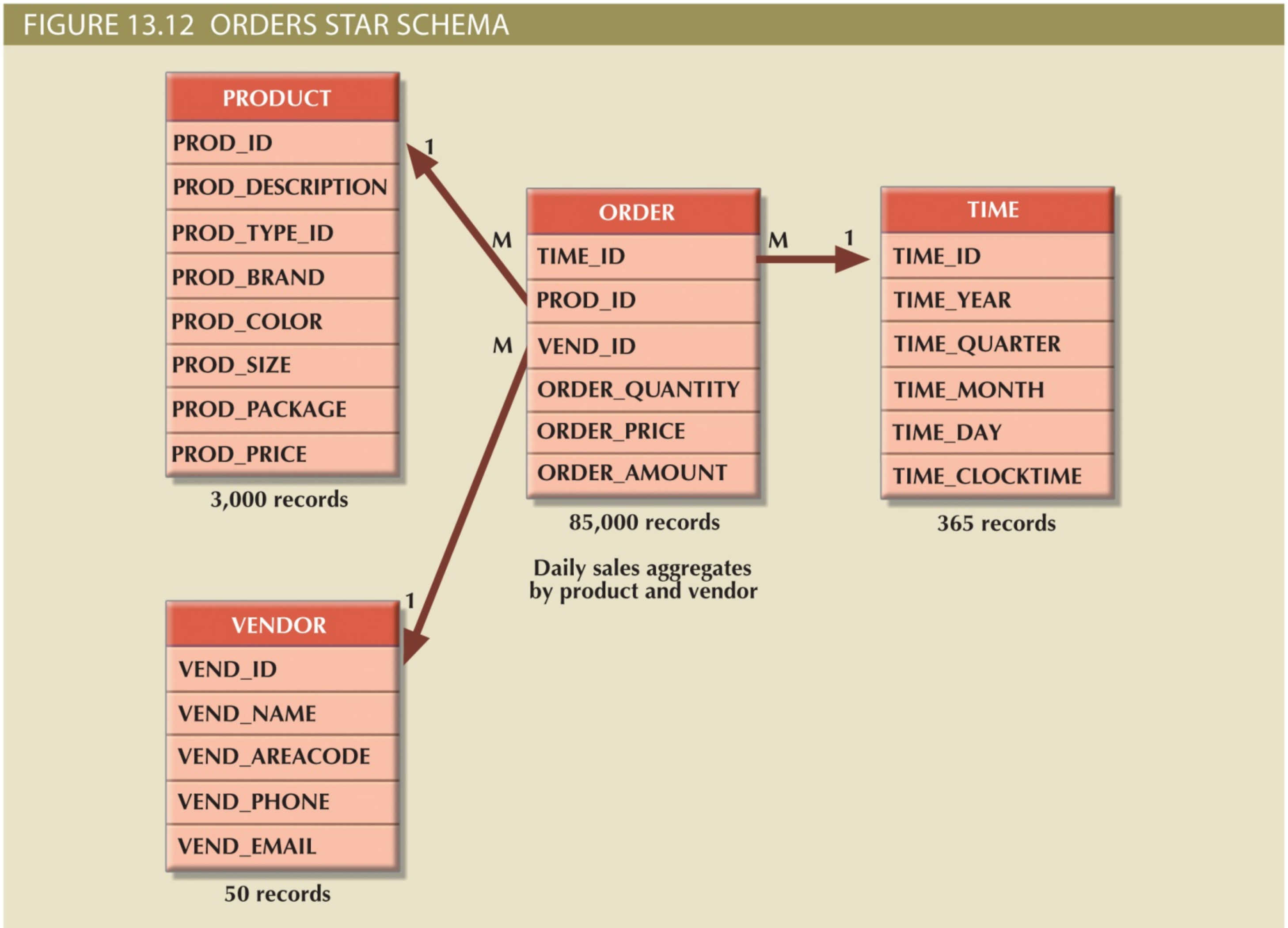
Figure 13.11 – Star Schema for Sales



Performance-Improving Techniques for the Star Schema

- Normalizing dimensional tables
  - **Snowflake schema**: Dimension tables can have their own dimension tables
- Maintaining multiple fact tables to represent different aggregation levels
- Denormalizing fact tables

Figure 13.12 – Orders Star Schema



Performance-Improving Techniques for the Star Schema

- Partitioning and replicating tables
  - **Partitioning**: Splits tables into subsets of rows or columns and places them close to customer location
  - **Replication**: Makes copy of table and places it in a different location
  - **Periodicity**: Provides information about the time span of the data stored in the table

Figure 13.13 – Normalized Dimension Tables

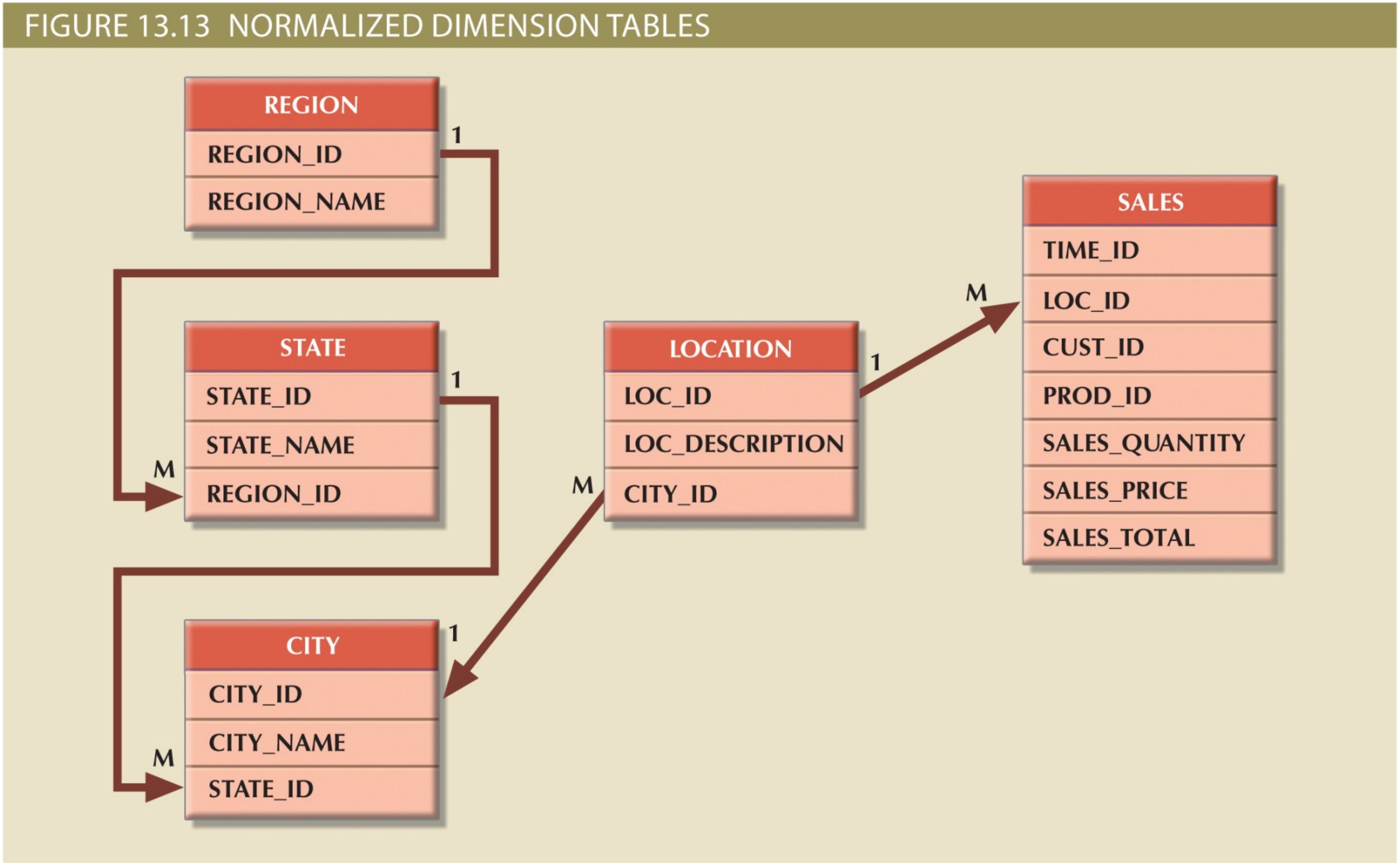
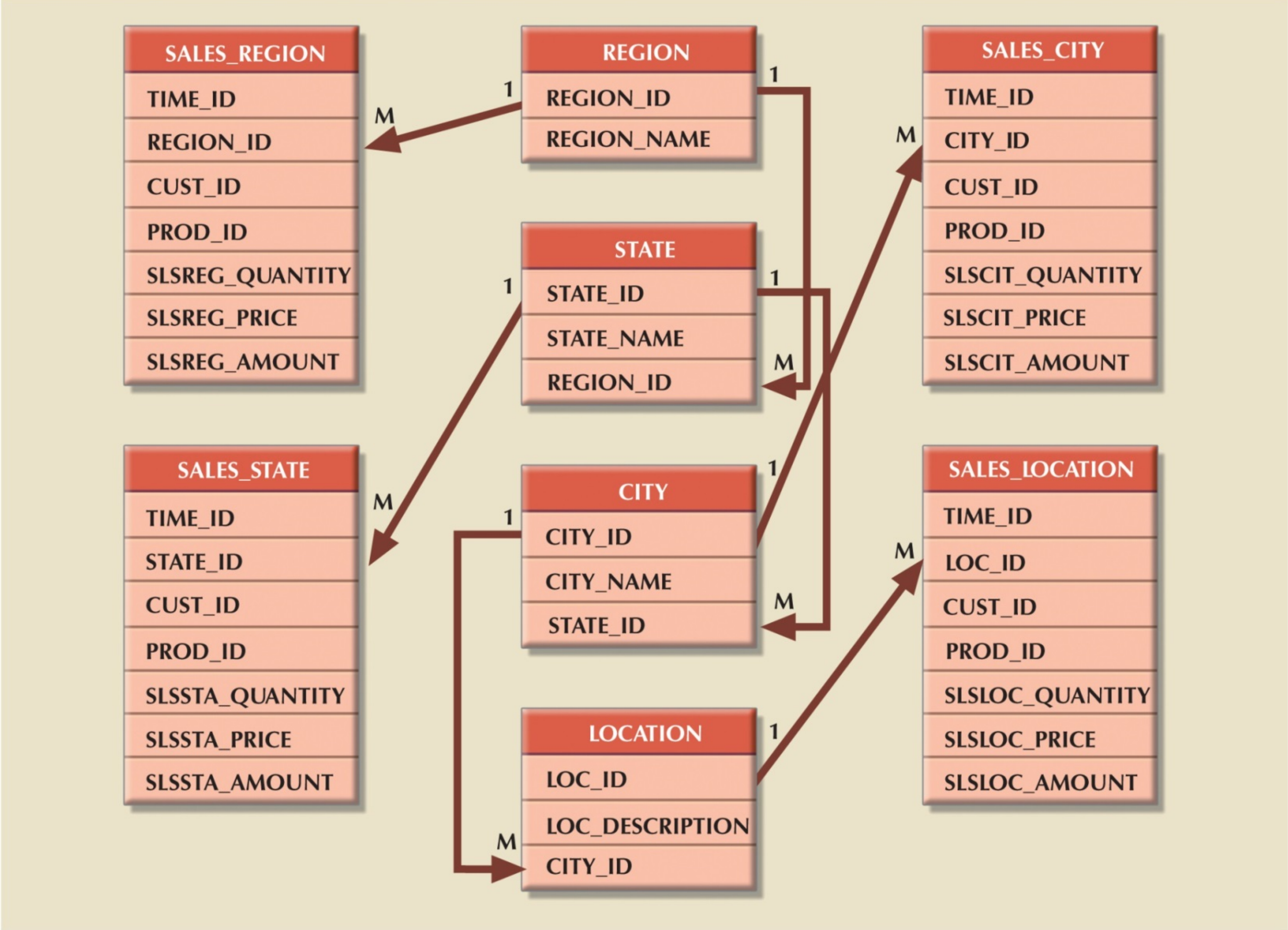


Figure 13.14 – Multiple Fact Tables

FIGURE 13.14 MULTIPLE FACT TABLES



Online Analytical Processing (OLAP)

- Advanced data analysis environment that supports decision making, business modeling, and operations research
- Characteristics:
  - Multidimensional data analysis techniques
  - Advanced database support
  - Easy-to-use end-user interfaces

Multidimensional Data Analysis Techniques

- Data are processed and viewed as part of a multidimensional structure
- Augmented by the following functions:
  - Advanced data presentation functions
  - Advanced data aggregation, consolidation, and classification functions
  - Advanced computational functions
  - Advanced data-modeling functions

Figure 13.15 – Operational vs. Multidimensional View of Sales

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)

Advanced Database Support

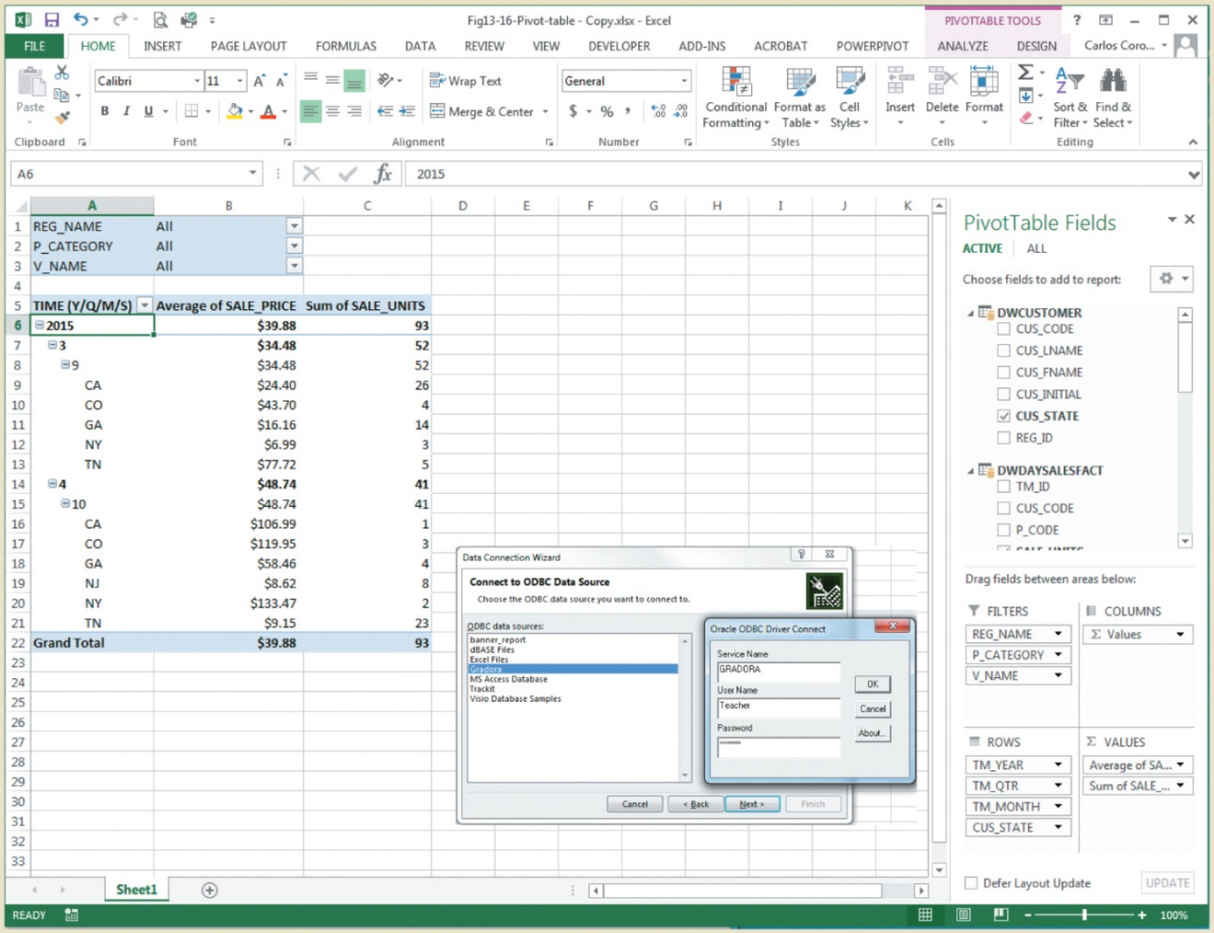
- OLAP tools must have the following features to deliver efficient decision support:
  - Access to many different kinds of DBMSs, flat files, and internal and external data sources
  - Access to aggregated data warehouse data and operational database detail data
  - Advanced data navigation features
  - Rapid and consistent query response times
  - Ability to map end-user requests
  - Support for very large databases

Easy-to-Use End-User Interface

- Proper implementation leads to simple navigation and accelerated decision making or data analysis
- Advanced OLAP features are more useful when access is kept simple
- Many interface features are borrowed from previous generations of data analysis tools

Figure 13.16 – Integration of OLAP with Spreadsheet Program

FIGURE 13.16 INTEGRATION OF OLAP WITH A SPREADSHEET PROGRAM



Source: Microsoft LLC (Excel screenshot); Oracle OCBC (Oracle windows)

Figure 13.17 - OLAP Architecture

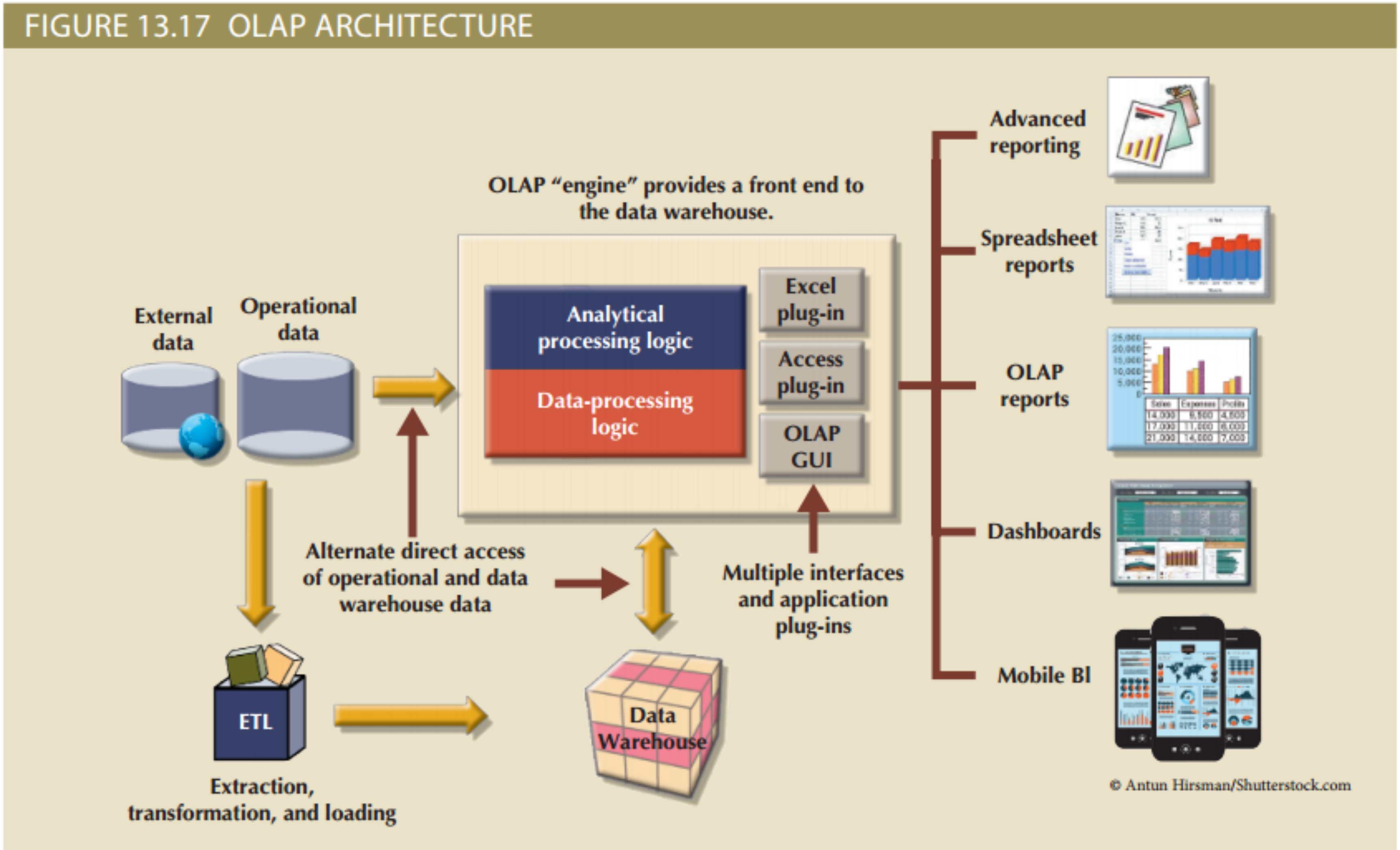
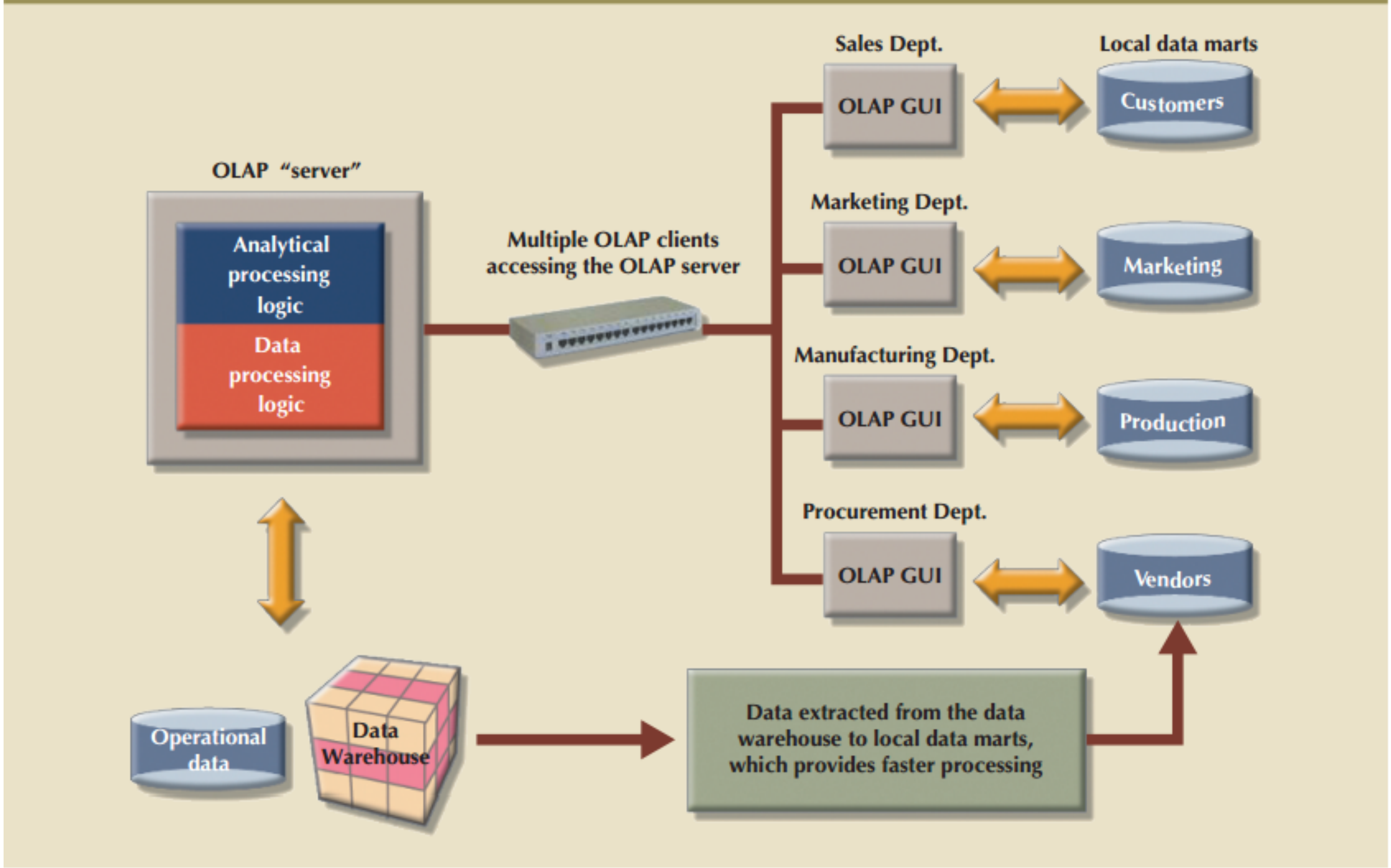


Figure 13.18 - OLAP Server with Local Mini Data Marts

FIGURE 13.18 OLAP SERVER WITH LOCAL MINI DATA MARTS



Relational Online Analytical Processing (ROLAP)

- Provides OLAP functionality using relational databases and familiar relational tools to store and analyze multidimensional data
- Extensions added to traditional RDBMS technology
  - Multidimensional data schema support within the RDBMS
  - Data access language and query performance optimized for multidimensional data
  - Support for very large databases (VLDBs)

Multidimensional Online Analytical Processing (MOLAP)

- Extends OLAP functionality to multidimensional database management systems (MDBMSs)
  - **MDBMS**: Uses proprietary techniques store data in matrix-like n-dimensional arrays
  - End users visualize stored data as a 3D **data cube**
    - Grow to n dimensions, becoming hypercubes
    - Held in memory in a **cube cache** to speed access
- **Sparsity**: Measures the density of the data held in the data cube

Table 13.12 - Relational vs. Multidimensional OLAP

TABLE 13.12

RELATIONAL VS. MULTIDIMENSIONAL OLAP

CHARACTERISTIC	ROLAP	MOLAP
Schema	Uses star schema Additional dimensions can be added dynamically	Uses data cubes Multidimensional arrays, row stores, column stores Additional dimensions require re-creation of the data cube
Database size	Medium to large	Large
Architecture	Client/server Standards-based	Client/server Open or proprietary, depending on vendor
Access	Supports ad hoc requests Unlimited dimensions	Limited to predefined dimensions Proprietary access languages
Speed	Good with small data sets; average for medium-sized to large data sets	Faster for large data sets with predefined dimensions

SQL Extensions for OLAP

The ROLLUP extension

- Used with GROUP BY clause to generate aggregates by different dimensions
- Enables subtotal for each column listed except for the last one, which gets a grand total
- Order of column list important

The CUBE extension

- Used with GROUP BY clause to generate aggregates by the listed columns
- Includes the last column

Figure 13.19 – Saleco Snowflake Schema

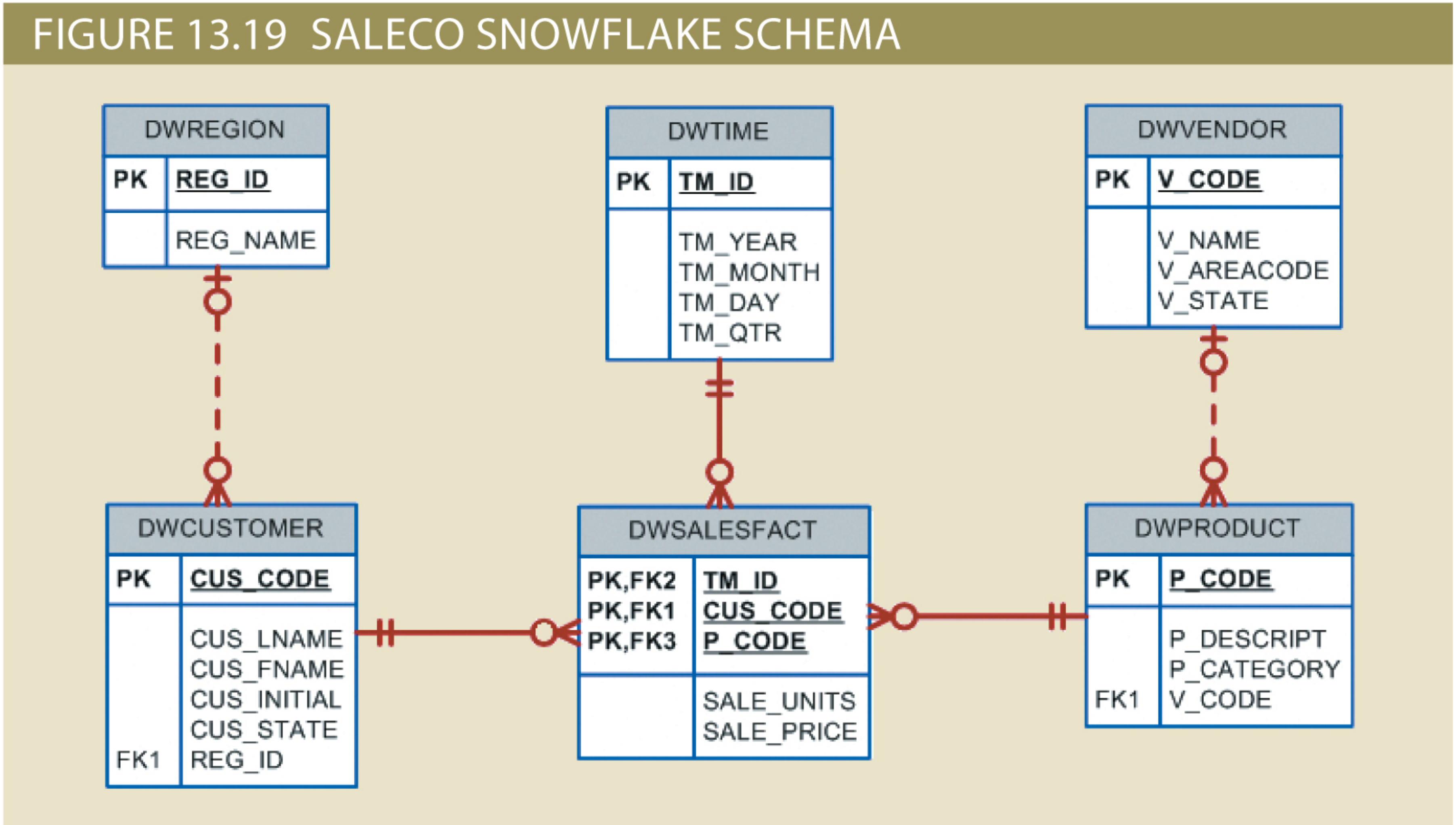


Figure 13.19 – Saleco Snowflake Schema

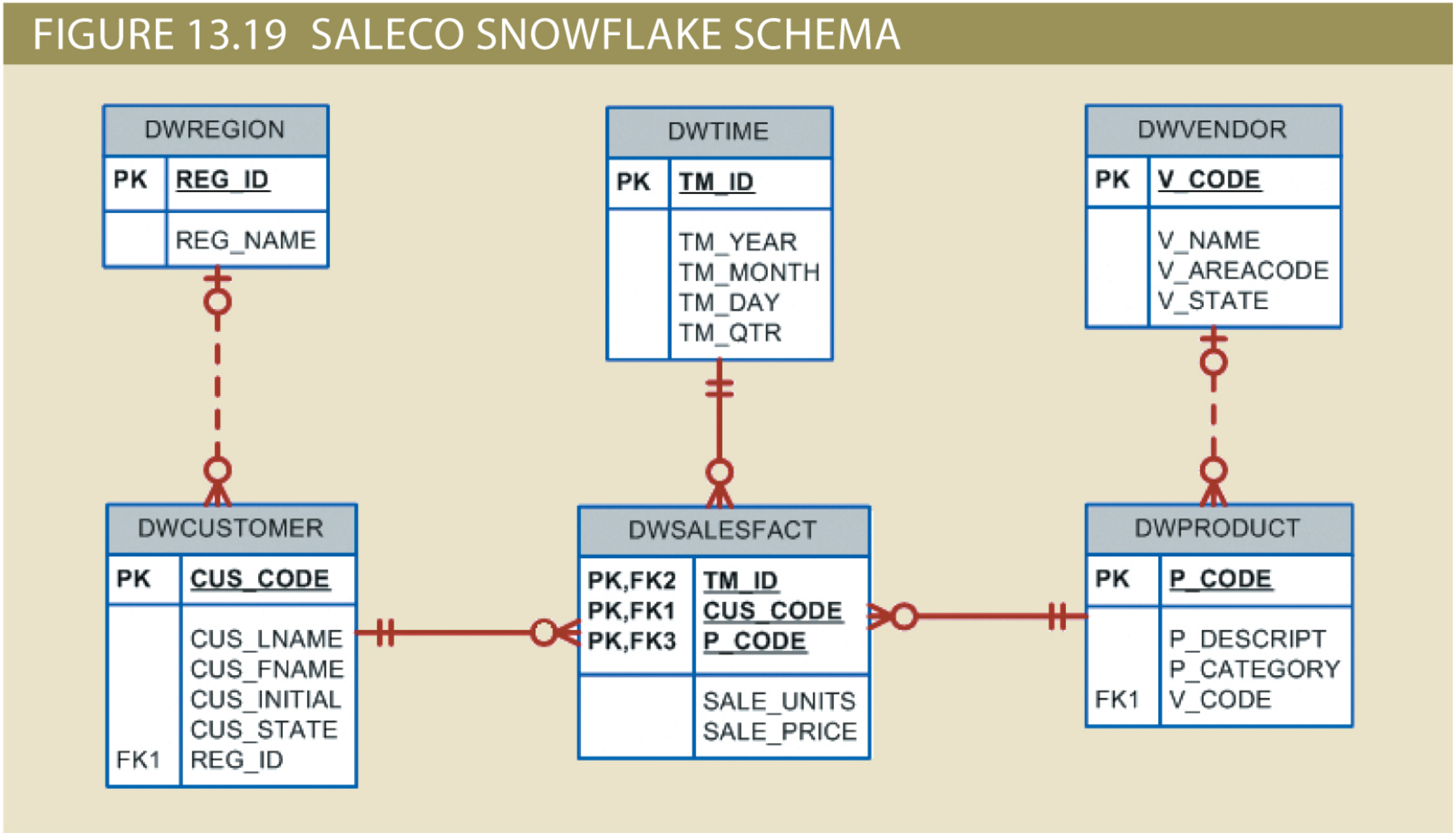


Figure 13.20 – Rollup Extension

FIGURE 13.20 ROLLUP EXTENSION

Oracle SQL*Plus

File Edit Search Options Help

```
SQL> SELECT   V_CODE, P_CODE, SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
  2  FROM     DWDAYSALESFACT NATURAL JOIN DWPRODUCT NATURAL JOIN DWVENDOR
  3  GROUP BY ROLLUP (V_CODE, P_CODE)
  4  ORDER BY V_CODE, P_CODE;

    V_CODE P_CODE       TOTSALES
---------- ---------- ----------
     21225 23109-HB         99.5
     21225 PVC23DRT       199.58
     21225 SM-18277        41.94
     21225                341.02
     21344 13-Q2/P2       239.84
     21344 54778-2T        59.88
     21344                299.72
     23119 1546-QQ2         79.9
     23119                  79.9
     24288 2232/QTY       219.84
     24288 89-WRE-Q       513.98
     24288                733.82
     25595 2238/QPD         77.9
     25595 WR3/TT3         719.7
     25595                797.6
                         2252.06

16 rows selected.
```

Subtotals by V_CODE

Grand total for all P_CODE values

Figure 13.21 – Cube Extension

FIGURE 13.21 CUBE EXTENSION

Oracle SQL*Plus

File Edit Search Options Help

```
SQL> SELECT   TM_MONTH, P_CODE, SUM(SALE_UNITS*SALE_PRICE) AS TOTSALES
  2  FROM     DWDAYSALESFACT NATURAL JOIN DWPRODUCT NATURAL JOIN DWTIME
  3  GROUP BY CUBE (TM_MONTH, P_CODE)
  4  ORDER BY TM_MONTH, P_CODE;

  TM_MONTH P_CODE       TOTSALES
---------- ---------- ----------
         9 13-Q2/P2       134.91
         9 1546-QQ2         79.9
         9 2232/QTY       109.92
         9 2238/QPD         77.9
         9 23109-HB         59.7
         9 54778-2T        39.92
         9 89-WRE-Q       256.99
         9 PVC23DRT        99.79
         9 SM-18277        20.97
         9 WR3/TT3        359.85
         9                1239.85
        10 13-Q2/P2       104.93
        10 2232/QTY       109.92
        10 23109-HB         39.8
        10 54778-2T        19.96
        10 89-WRE-Q       256.99
        10 PVC23DRT        99.79
        10 SM-18277        20.97
        10 WR3/TT3        359.85
        10               1012.21
           13-Q2/P2       239.84
           1546-QQ2         79.9
           2232/QTY       219.84
           2238/QPD         77.9
           23109-HB         99.5
           54778-2T        59.88
           89-WRE-Q       513.98
           PVC23DRT       199.58
           SM-18277        41.94
           WR3/TT3         719.7
                         2252.06

31 rows selected.
```

Subtotals by month

Subtotals by product

Grand total for all products and months

Materialized View

- Dynamic table that contains SQL query command to generate rows and stores the actual rows
- Created the first time query is run
  - Summary rows are stored in the table
- Automatically updated when base tables are updated
- Requires specified privileges

Figure 13.22 – Creating Materialized View

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 Materialized View

FIGURE 13.23 REFRESHING A MATERIALIZED VIEW

SQL Plus

```
SQL> INSERT INTO DWDAYSALESFACT VALUES (207,10017,'WR3/TT3',1,106.99);

1 row created.

SQL> COMMIT;

Commit complete.

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             4     466.84

18 rows selected.

SQL> _
```

Figure 13.24 –Sample OLAP Applications

FIGURE 13.24 SAMPLE OLAP APPLICATIONS