1.0 Introduction to Business Analytics

Eugene Rex L. Jalao, Ph.D.

Associate Professor

Department Industrial Engineering and Operations Research

University of the Philippines Diliman

@thephdataminer

Module 1 of the Business Intelligence and Analytics Track of UP NEC and the UP Center of Business Intelligence

Outline for This Session

- Introductions

- What is Business Analytics?

- The Business Analytics Framework

- Tools of Business Analytics

- Case Study

Introductions: About Me

- PhD, Industrial Engineering, Arizona State University, (2013)
  - Data Mining Specialization
- Certificate in Data Warehousing, University of California Irvine (2013)
- MS Industrial Engineering, UP Diliman, (2009)
  - Optimization Specialization
- BS Industrial Engineering, UP Diliman, (2007)
- Courses Taught
  - Management Science, Simulation, Information Systems, Systems and Procedures, SAP, Data Mining
- Research/Consultancy Interests
  - VB.NET Database System Development, Data Extraction/Data Mining, Advanced Report Generation, Supply Chain Problems, Agent-Based Simulation and Evolutionary Optimization

Introductions: About You

- Name

- Nickname

- Educational Background

- Job Position and Current Duties

- Where did you hear about this course?

UP NEC BA Modules

- Analyst Level

  1. **Introduction to Business Analytics**
  2. Data Warehousing
  3. Data Mining

- Professional Level

  4. Time Series Analysis and Forecasting
  5. Optimization Analysis
  6. Advanced Data Mining

Module 1 Outline

1. **Intro to Business Intelligence**
   – **Case Study on Selecting BI Projects**

2. Data Warehousing
   – Case Study on Data Extraction and Report Generation

3. Descriptive Analytics
   – Case Study on Data Analysis

4. Visualization
   – Case Study on Dashboard Design

5. Classification Analysis
   – Case Study on Classification Analysis

6. Regression and Time Series Analysis
   – Case Study on Regression and Time Series Analysis

7. Unsupervised Learning and Modern Data Mining
   – Case Study on Text Mining

8. Optimization for BI

Opening Vignette

Source: www.target.com

Outline for This Session

- Introductions
- **What is Business Analytics?**
- The Business Analytics Framework
- Tools of Business Analytics
- Case Study

What is Business Analytics?

- Timely

- Accurate

- High-Value

- Actionable

DECISIONS

Via organizational (and sometimes external) data

What is Business Analytics?

Definition 1.1: Business Analytics

- BA is an **umbrella term** that combines architectures, tools, databases, analytical tools, applications, and methodologies

- BA means **different things to different people**

- BA's major objective is to enable **easy access** to data (and models) to provide managers with the ability to conduct analysis

- BA helps *transform* data to information (and knowledge) to **decisions** and finally to **action**

What is Business Analytics?

- Is:
  - An **entire discipline** that encompasses:
    - technology
    - work processes
    - human/ organizational factors
  - **Evolutionary**

- Is **not**:
  - A product or tool
  - A project
  - A collection of reports or dashboards or visualizations…

What is Business Analytics?
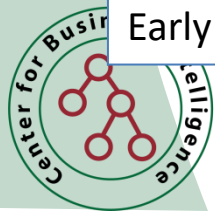
Definition 1.2: Business Analytics

- "Use **sophisticated data-collection technology and analysis** to wring every last drop of value from all your **business processes**."

Source: Davenport, Thomas H. "Competing on analytics." *Harvard Business Review* 84.1 (2006): 98.

What is Business Analytics?

Table 1.1: BA History Timeline

Timeframe	User-Facing	Data/Infrastructure
1970s and early/mid-1980s	Mostly paper reports	File systems; early databases
Late 1980s	Paper reports; start of terminal and PC reports; specialized DSS/EIS;	Databases and extract files; attempts at distributed databases
Early 1990s	Early OLAP	Early data warehousing, special databases
Rest of 1990s	Enhanced OLAP; early dashboards	More data warehousing; data marts; begin terabyte+ DW
Most of 2000s	Next generation OLAP; integrated data mining; dashboards	Inclusion of Data Mining, Enhanced data warehousing including "data warehousing appliances"
Early 2010s	Visualizations; mobile BI	Big data; cloud BA

What is Business Analytics?

- Gartner's Top 10 Strategic Tech Trends for 2015

    1. Computing Everywhere
    2. The Internet of Things
    3. 3-D Printing
    4. **Analytics**
    5. Context-Rich Systems
    6. Smart Machines
    7. Cloud Computing
    8. Software-Defined Applications and Infrastructure
    9. Web-Scale IT
    10. Security

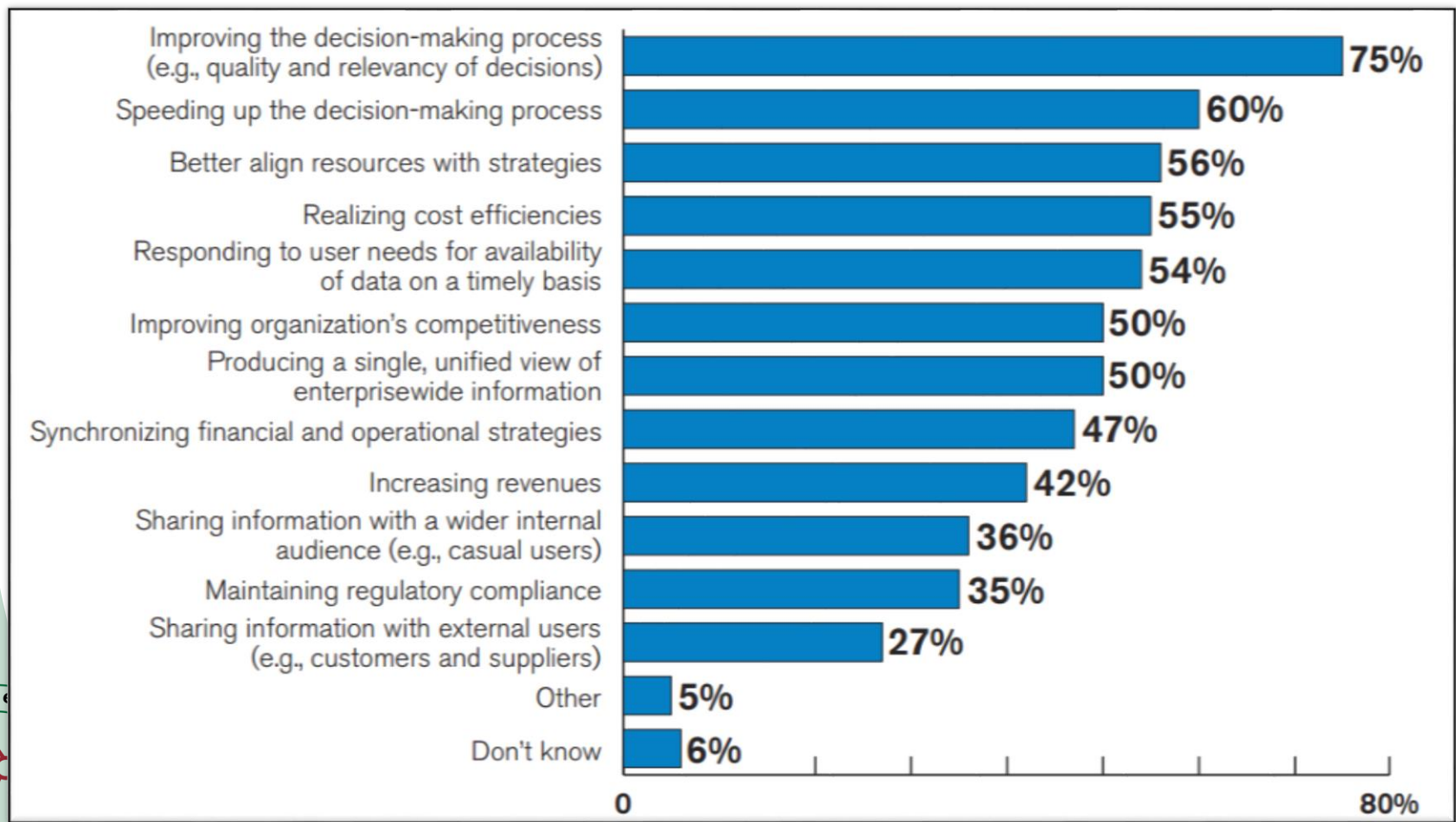http://www.information-management.com/gallery/gartners-top-10-strategic-tech-trends-for-2015-10026168-1.html

What is Business Analytics?

Table 1.2: Investments in BA

Rank	Investment priority	2014	2015
1	BI/analytics	41%	50%
2	Infrastructure and data center	31%	37%
3	Cloud	27%	32%
4	ERP	26%	34%
5	Moblie	24%	36%
6	Digitalization/digital marketing	17%	11%
7	Security	13%	11%
8	Networking, voice and data comms	12%	12%
9	Customer relationship/experience	11%	8%
10	Industry-specific applications	9%	10%
11	Legacy modernization	7%	7%
12	Enterprise applications	6%	2%

What is Business Analytics?

Table 1.3: Benefits of BA

Benefit	Percentage
Improving the decision-making process (e.g., quality and relevancy of decisions)	75%
Speeding up the decision-making process	60%
Better align resources with strategies	56%
Realizing cost efficiencies	55%
Responding to user needs for availability of data on a timely basis	54%
Improving organization's competitiveness	50%
Producing a single, unified view of enterprisewide information	50%
Synchronizing financial and operational strategies	47%
Increasing revenues	42%
Sharing information with a wider internal audience (e.g., casual users)	36%
Maintaining regulatory compliance	35%
Sharing information with external users (e.g., customers and suppliers)	27%
Other	5%
Don't know	6%

What is Business Analytics?

Table 1.4: Some Challenges of BA
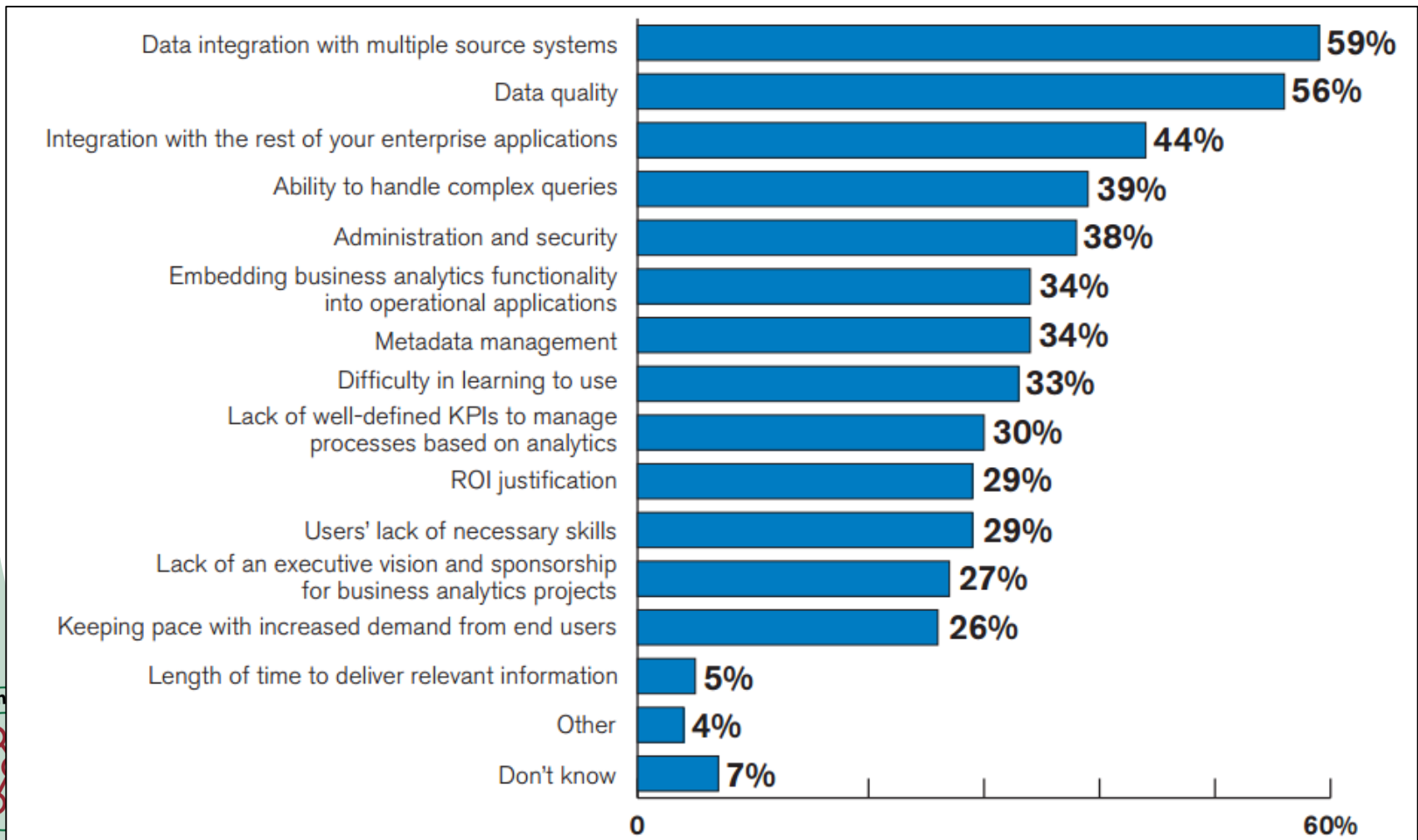
Outline for This Session

- Introductions
- What is Business Analytics?
- **The Business Analytics Framework**
- Tools of Business Analytics
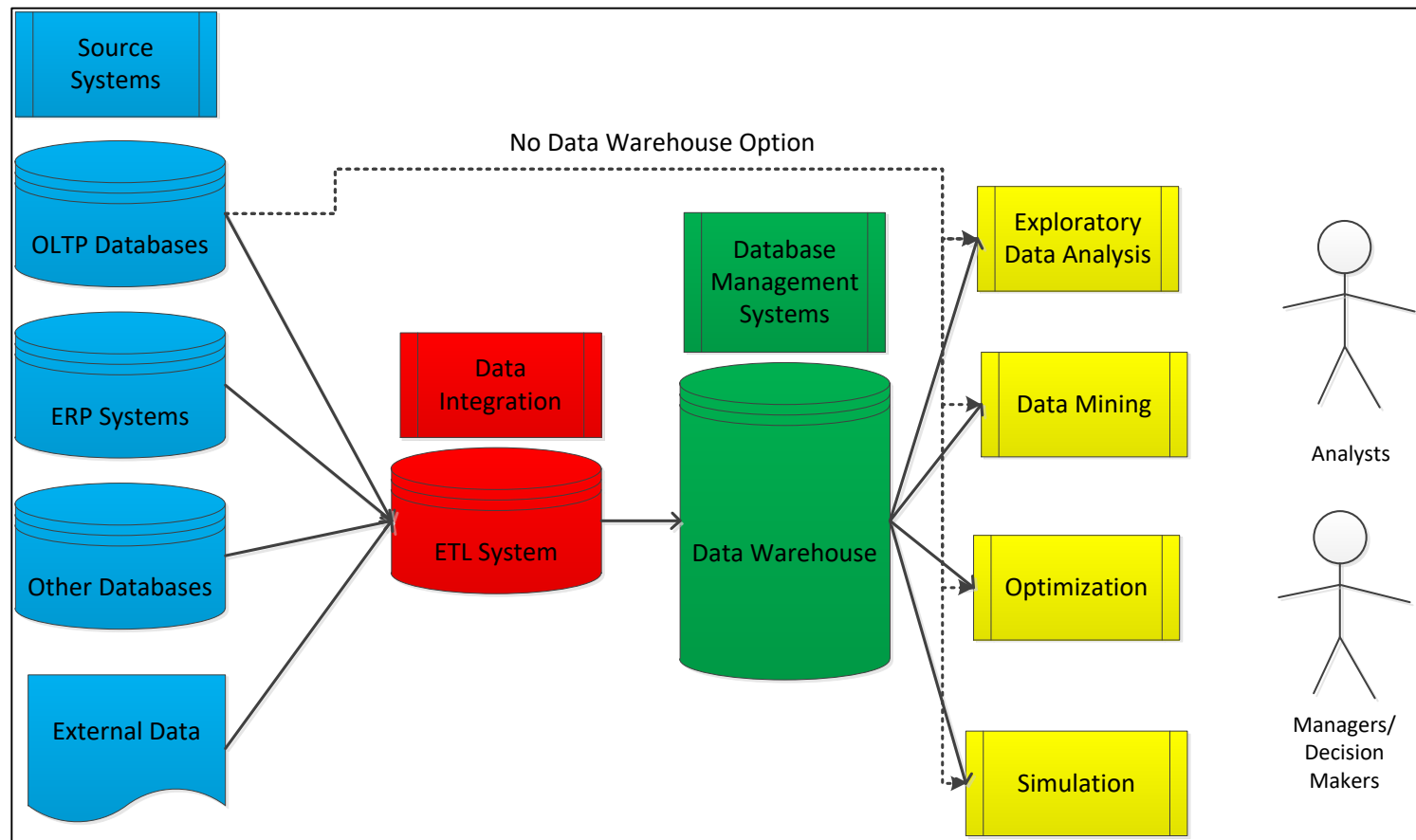- Case Study

The Business Analytics Framework

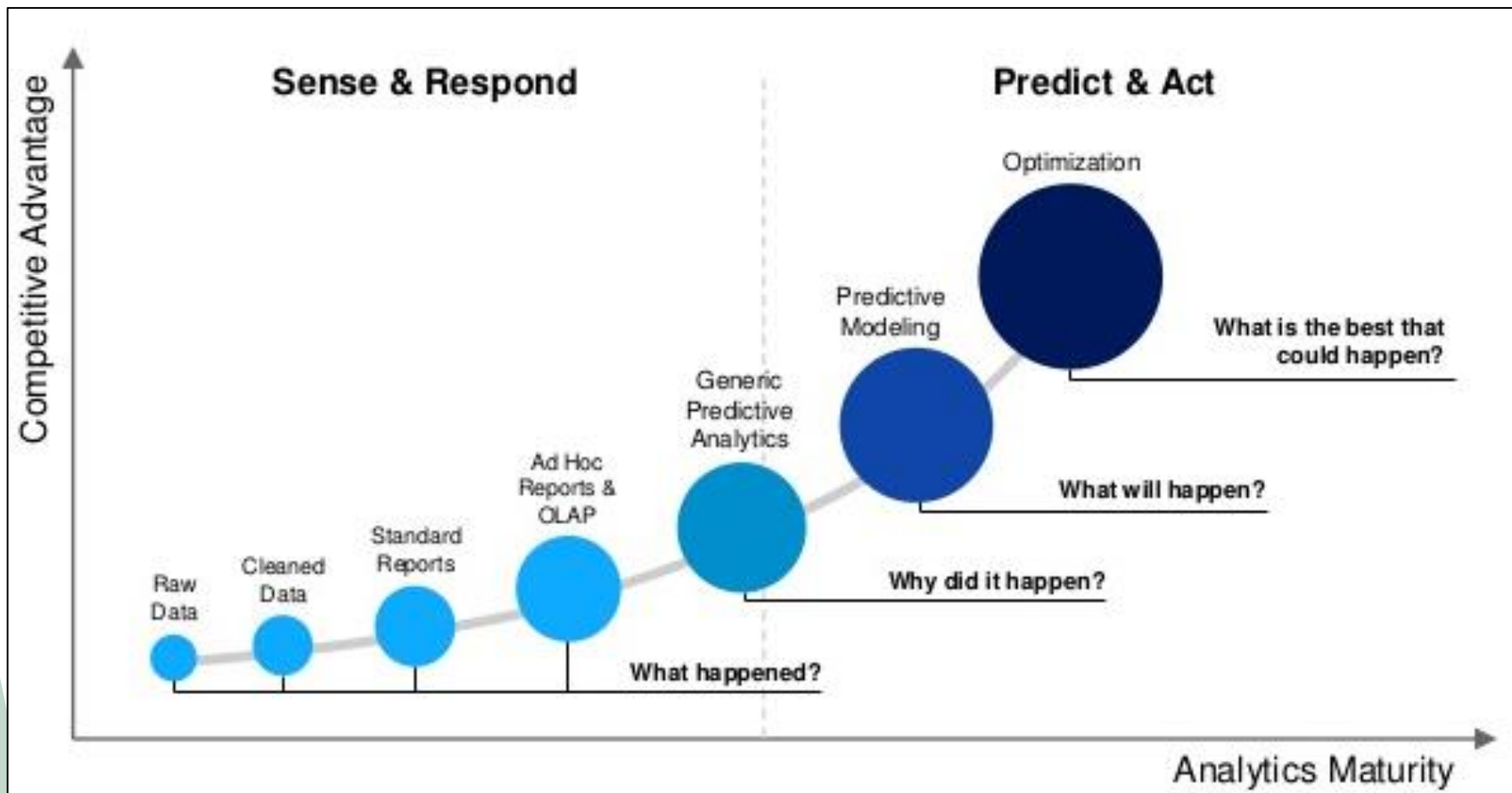*Figure 1.1: BA Framework*

The Business Analytics Framework

Figure 1.2: Types of BA According to Sophistication

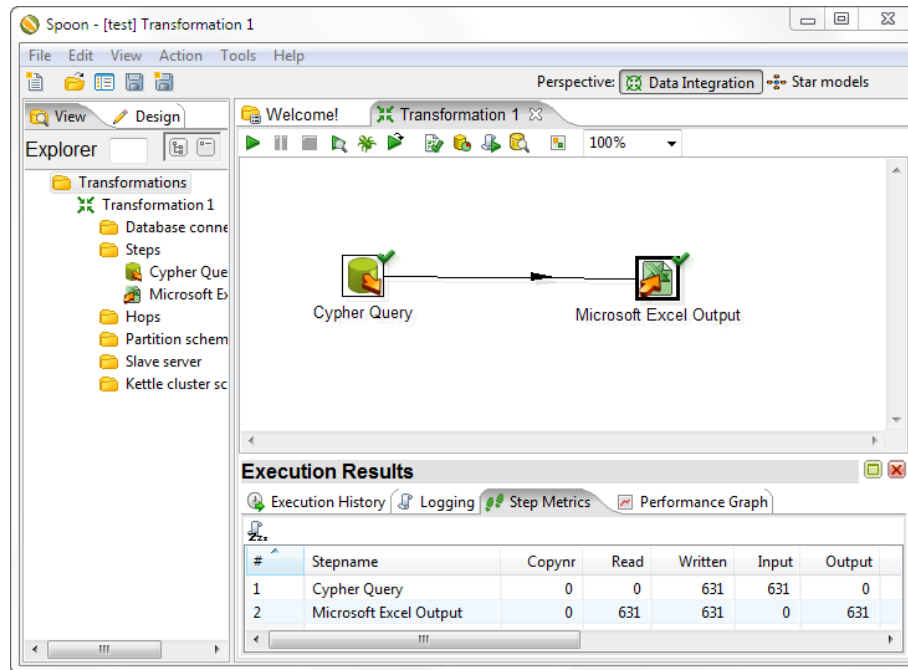# The Business Analytics Framework

Definition 1.3: Source Systems

- Source Systems are **sources of data** that drive the Business Analytics framework.

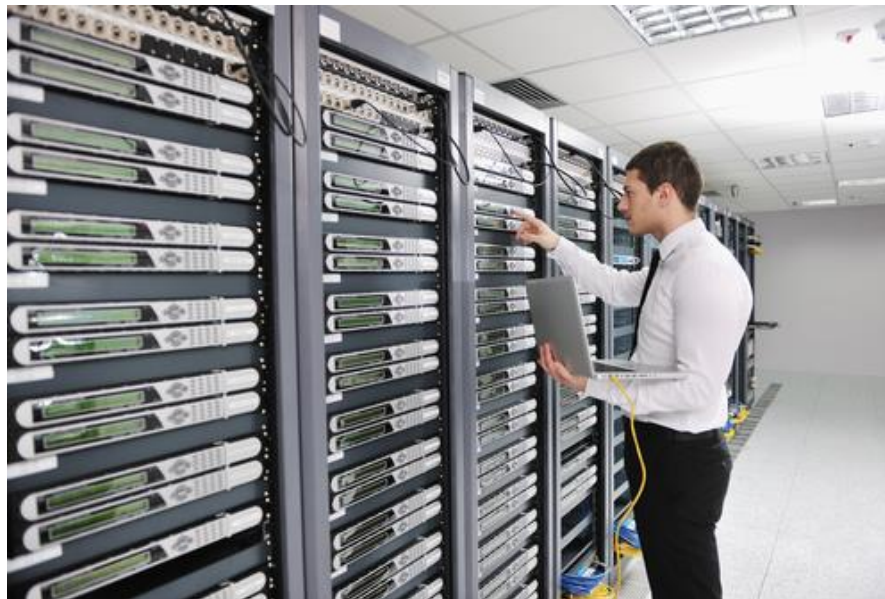# The Business Analytics Framework

- The Data Integration component of BA is a tool that **extracts data** from all Source Systems, transforms them into a unified template and loads it into a Data Warehouse for storage.

The Business Analytics Framework

Definition 1.5: Data Warehouse

- A Data Warehouse is a large database that **stores** transformed and cleaned data as input to various Business Analytics applications.

The Business Analytics Framework

- Types of BA According to Purpose:

  – **Exploratory Data Analysis** (Descriptive Analytics)

    - *Tell Me What has Happened and Why*
    - *Tell Me What is Happening Right Now*

  – **Data Mining** (Predictive Analytics)

    - *Tell Me What is Likely to Happen*
    - *Tell Me Something Interesting Without Me Asking*

  – **Optimization/Simulation** (Prescriptive Analytics)

    - *Tell Me What Might Have Happened*
    - *Tell Me What is the Best Solution*

The Business Analytics Framework

Definition 1.6: Exploratory Data Analysis

- Descriptive Analytics or Exploratory Data Analysis is a type of Business Analytics application where data is **described and summarized** using basic statistical tools and graphs to produce reports and dashboards for decision making.

The Business Analytics Framework

Figure 1.3: Dashboard

The Business Analytics Framework

Figure 1.4: Heat Map/Tree Map
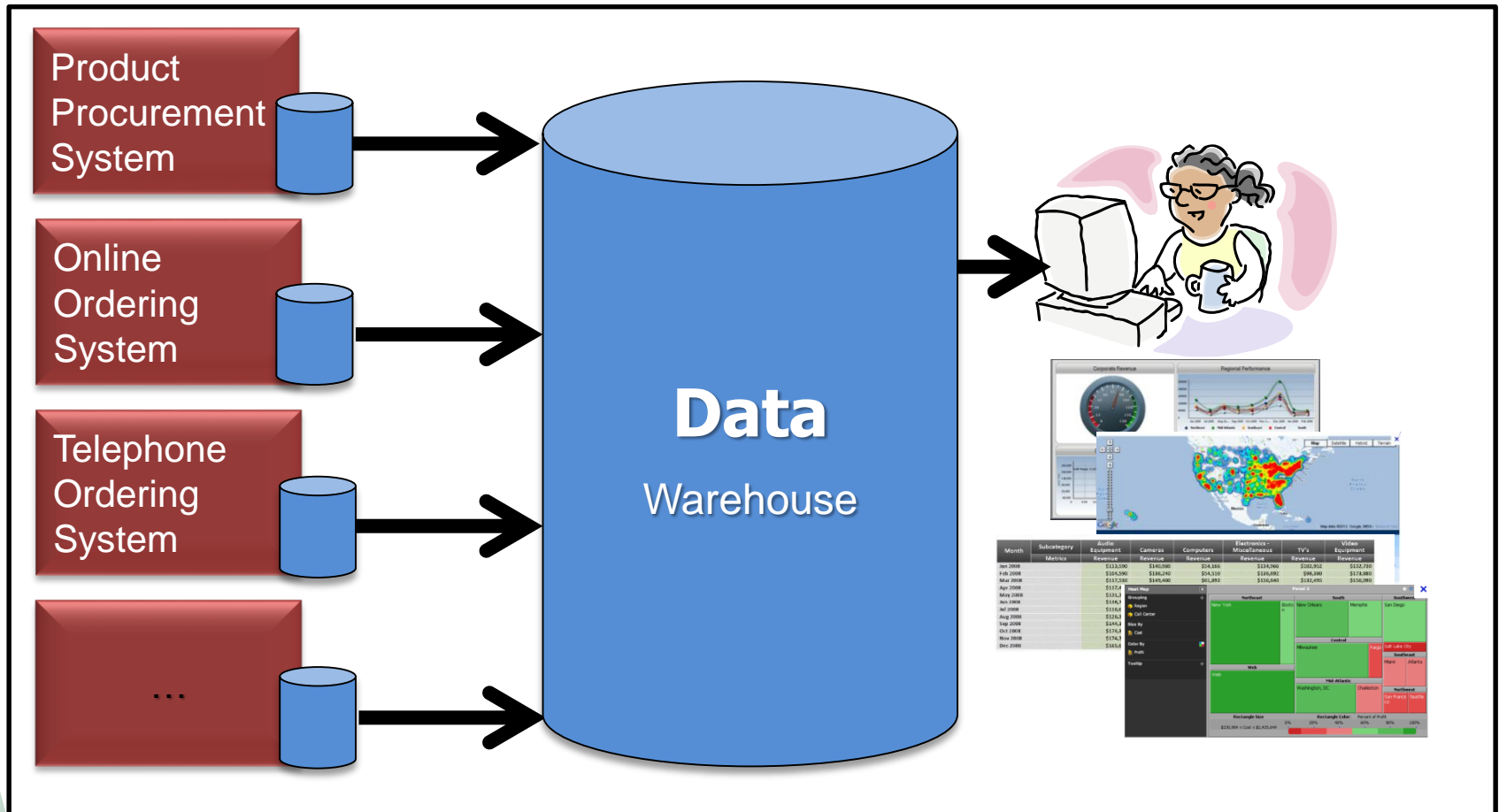
The Business Analytics Framework

Figure 1.5: Framework for EDA

The Business Analytics Framework

- Management would like to create a **daily report** that lists down the items to be ordered today to avoid **stock-outs**.
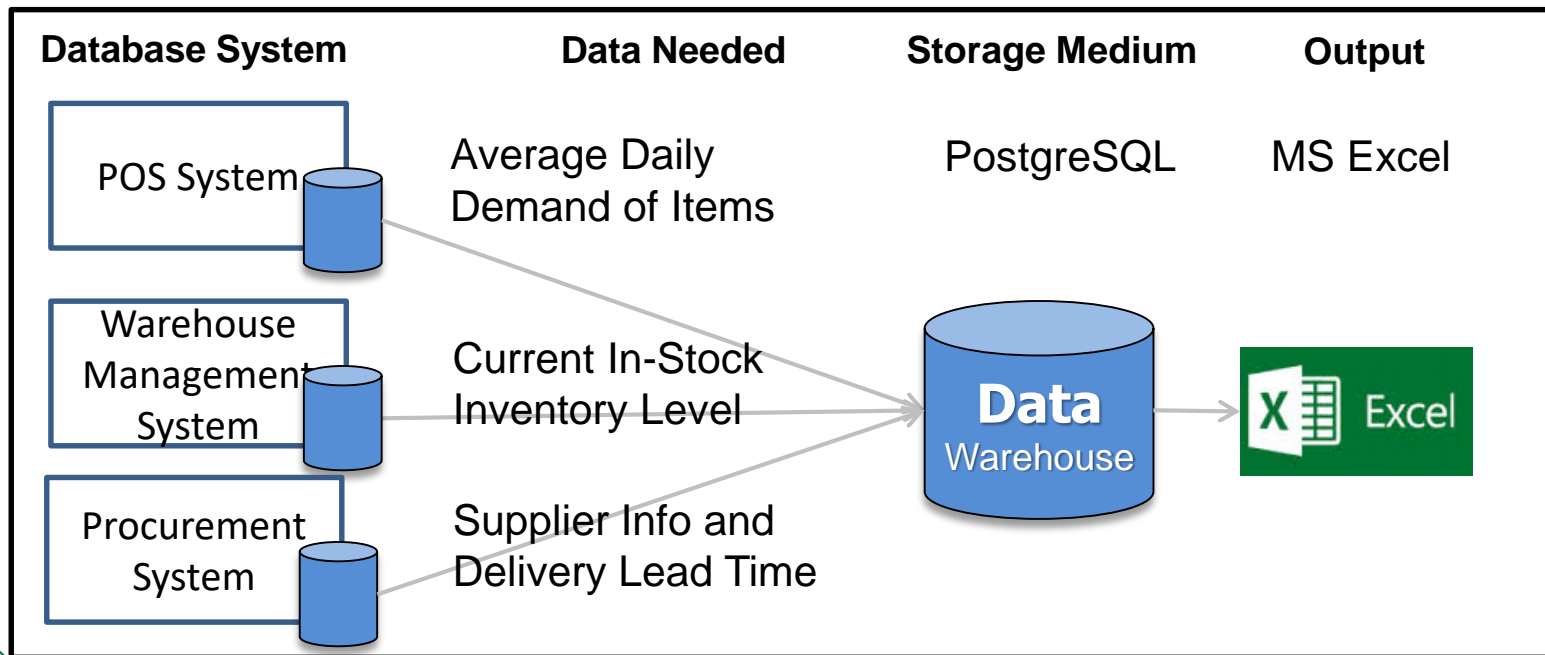

Database System	Data Needed	Storage Medium	Output
POS System	Average Daily Demand of Items	PostgreSQL	MS Excel
Warehouse Management System	Current In-Stock Inventory Level		
Procurement System	Supplier Info and Delivery Lead Time		

Figure 1.6: Descriptive Analytics Framework

29

The Business Analytics Framework

- Types of BA According to Purpose:
  - **Exploratory Data Analysis** (Descriptive Analytics)
    - *Tell Me What has Happened and Why*
    - *Tell Me What is Happening Right Now*
  - **Data Mining** (Predictive Analytics)
    - *Tell Me What is Likely to Happen*
    - *Tell Me Something Interesting Without Me Asking*
  - **Optimization/Simulation** (Prescriptive Analytics)
    - *Tell Me What Might Have Happened*
    - *Tell Me What is the Best Solution*

The Business Analytics Framework

Definition 1.7: Data Mining

- Predictive Analytics or Data Mining is a Business Analytics application that seeks to **find patterns and trends** within historical data to provide useful information that can aid decision makers.
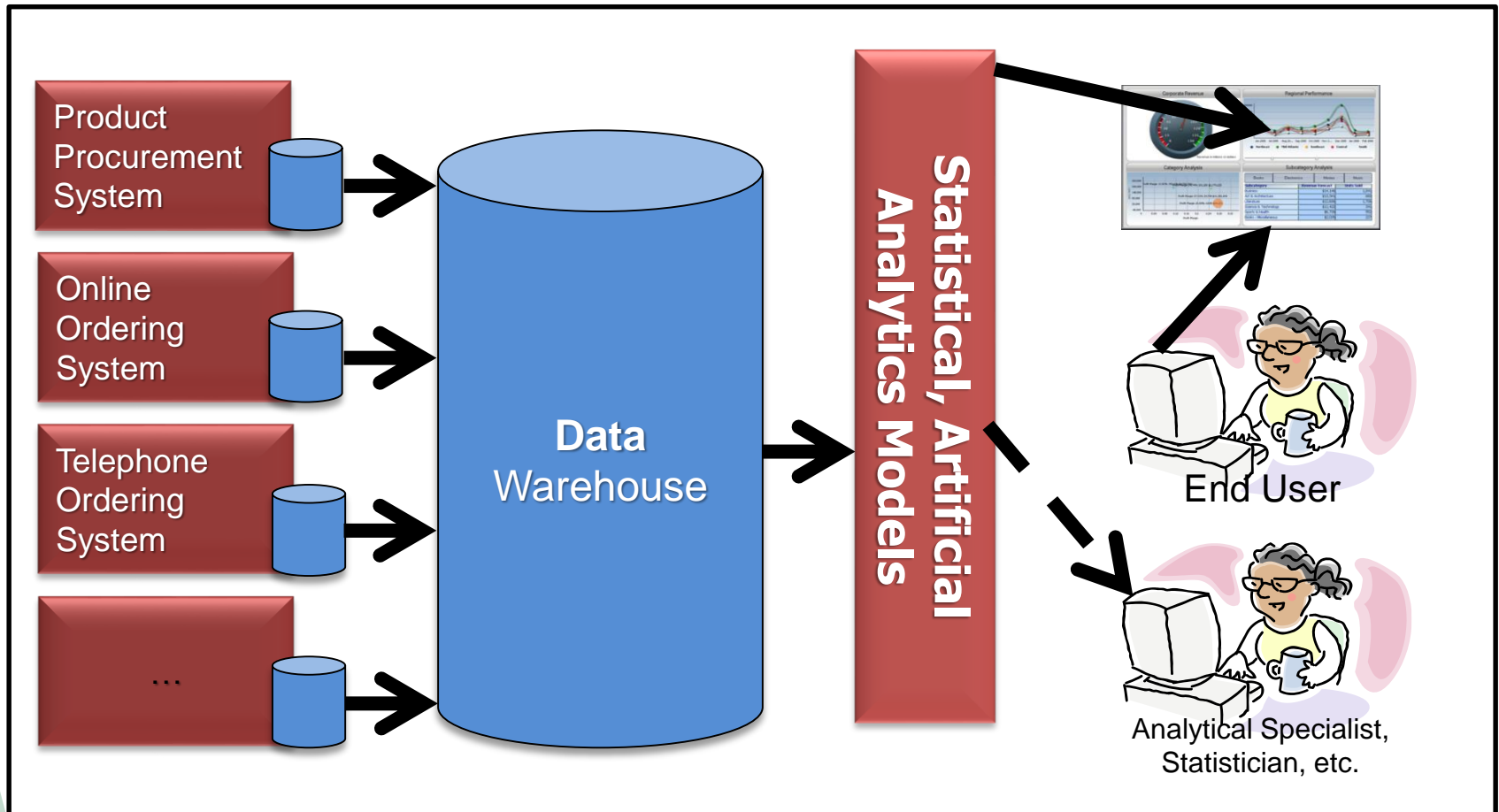
The Business Analytics Framework

Figure 1.7: Data Mining Framework

The Business Analytics Framework

- Types of Data Mining Algorithms
  - Supervised Learning
    - Classification
    - Regression
    - Time Series Analysis
  - Unsupervised Learning
    - Association Analysis
    - Sequential Pattern Analysis
    - Clustering
    - Text Mining/Social Media Sentiment Analysis

The Business Analytics Framework

Definition 1.8: Classification

- Classification is a data mining task of **predicting** the value of a categorical variable by **building a model** based on one or more numerical and/or categorical variables.
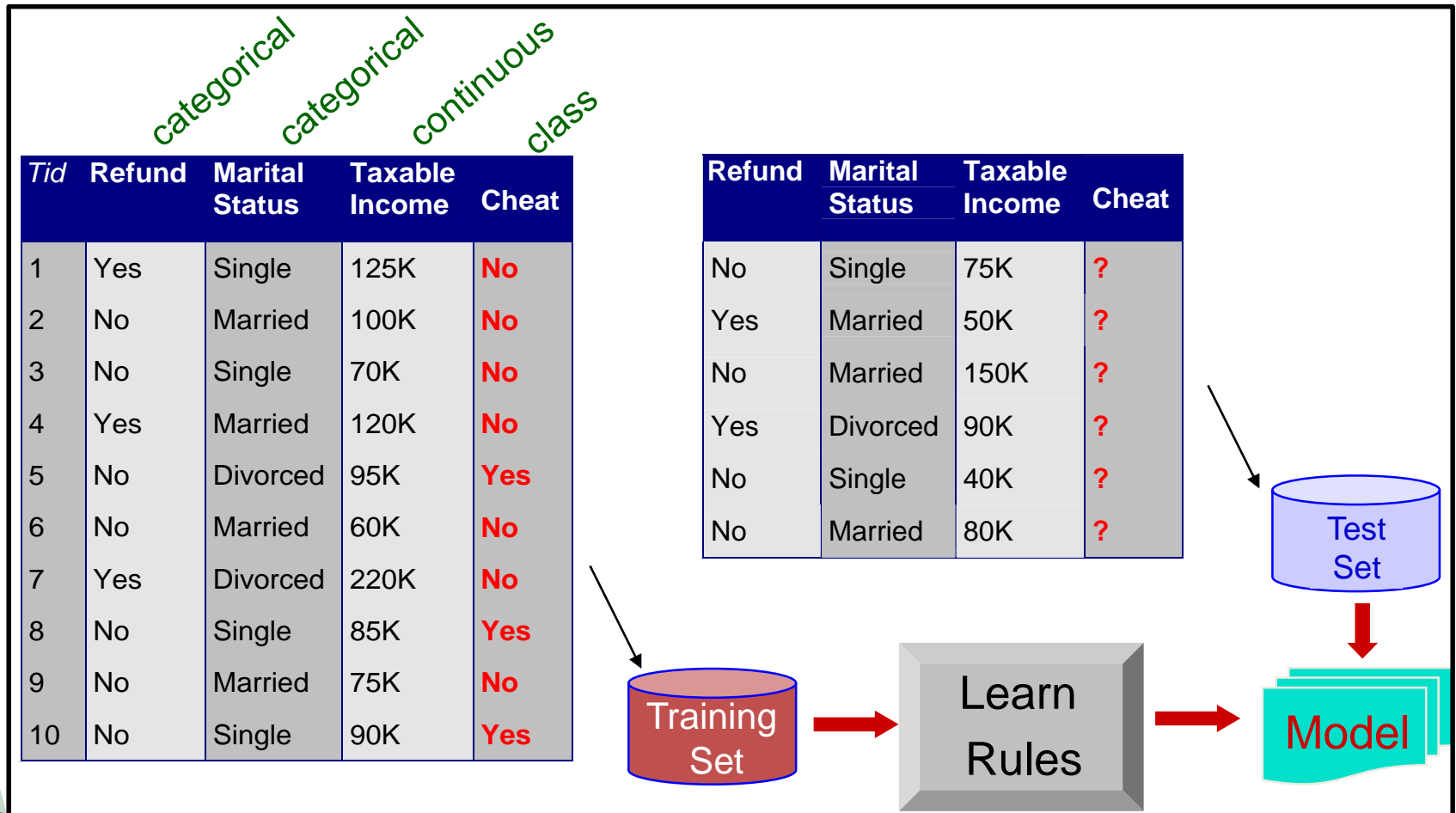
The Business Analytics Framework

categorical categorical continuous class

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	**No**
2	No	Married	100K	**No**
3	No	Single	70K	**No**
4	Yes	Married	120K	**No**
5	No	Divorced	95K	**Yes**
6	No	Married	60K	**No**
7	Yes	Divorced	220K	**No**
8	No	Single	85K	**Yes**
9	No	Married	75K	**No**
10	No	Single	90K	**Yes**

Refund	Marital Status	Taxable Income	Cheat
No	Single	75K	**?**
Yes	Married	50K	**?**
No	Married	150K	**?**
Yes	Divorced	90K	**?**
No	Single	40K	**?**
No	Married	80K	**?**

Test Set

Training Set → Learn Rules → Model

Figure 1.8: Classification Example

The Business Analytics Framework

Example 1.2: Direct Marketing

- Goal: Reduce cost of mailing by *targeting* a set of consumers likely to buy a new cellphone.

- Approach:
  - Use the data for a similar product introduced before.
  - We know which customers decided to buy and which decided otherwise. This *{buy, don't buy}* decision forms the *class attribute*.
  - Collect various demographic, lifestyle, and company-interaction related information about all such customers.
    - Type of business, where they stay, how much they earn, etc.
  - Use this information as input attributes to learn a classifier model.

The Business Analytics Framework

- Sample model framework for predicting probability of churn of subscribers
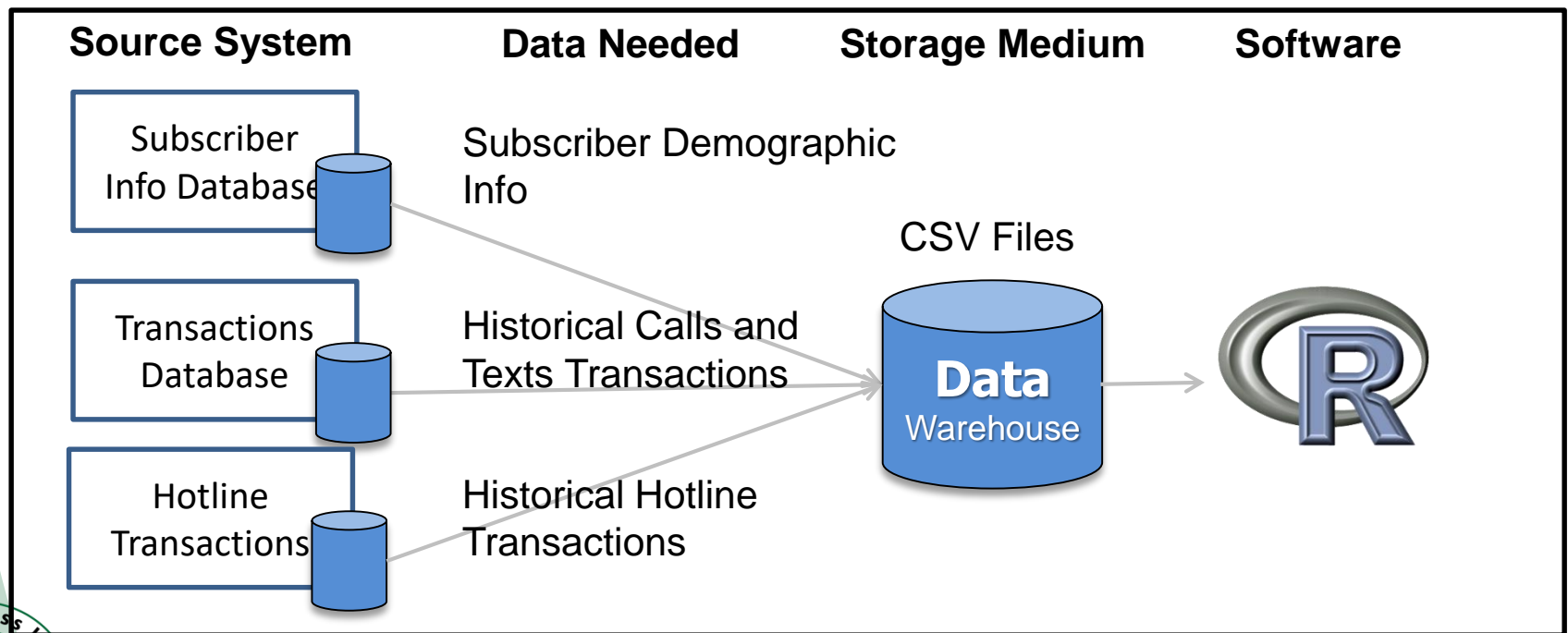
Figure 1.9: Churn Analysis Framework

The Business Analytics Framework

Definition 1.9: Regression

- Regression is a data mining task of **predicting** the value of target (numerical variable) by **building a model** based on one or more predictors (numerical and categorical variables).

- Greatly studied in **statistics, neural network** fields.

- Examples:
  - Predicting sales amounts of new product based on advertising expenditure.
  - Predicting wind velocities as a function of temperature, humidity, air pressure, etc.

The Business Analytics Framework

- Create a regression model to predict the headcount of the merchandisers of a supermarket
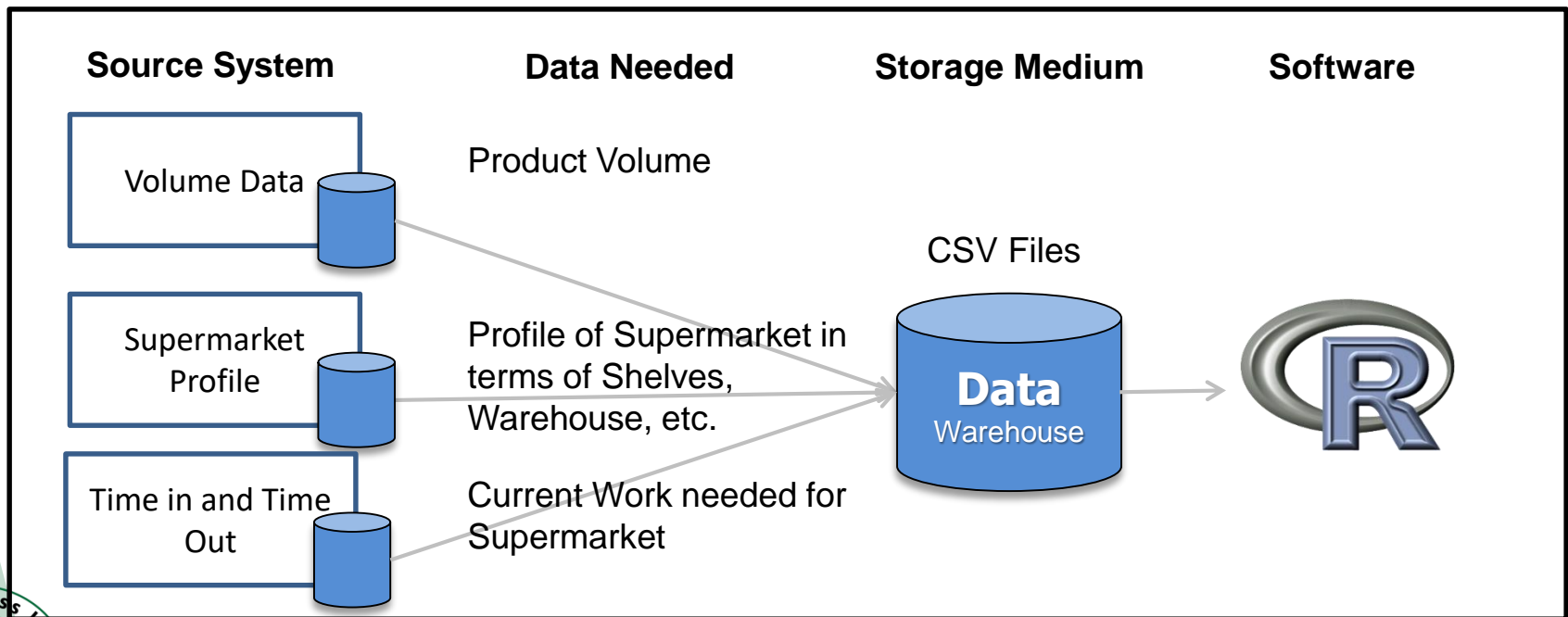
Figure 1.10: Regression Model Framework

The Business Analytics Framework

Definition 1.10: Time Series Analysis

- Time Series Analysis is a data mining task of **predicting** the value of target (numerical variable) by **building a model** based on one predictor (numerical and categorical variables).

- Examples:
  - Forecasting sales
  - Time series prediction of stock market indices.

The Business Analytics Framework

Definition 1.11: Clustering

- Clustering is the process of **dividing** a dataset into **groups** such that the members of each group are as similar (close) as possible to one another, and different groups are as dissimilar (far) as possible from one another.

The Business Analytics Framework

Example 1.5: Market Segmentation

- Goal: **subdivide** a market into distinct subsets of customers where any subset may conceivably be selected as a market target to be reached with a distinct marketing mix.

- Approach:
  - Collect different attributes of customers based on their geographical and lifestyle related information.
  - Find clusters of similar customers.
  - Measure the clustering quality by observing buying patterns of customers in same cluster vs. those from different clusters.

The Business Analytics Framework

Figure 1.9: Clustering Diagram

The Business Analytics Framework

Definition 1.12: Association Rule Analysis

- Definition: Is a data mining task used to identify strong rules that associate elements **together** in datasets using different measures of **interestingness**

The Business Analytics Framework

Example 1.6: Supermarket Basket Analysis

- Given a set of records each of which contain some number of items from a given collection;

- Produce **dependency rules** which will predict occurrence of an item based on occurrences of other items.

TID	Items
1	Bread, Coke, Milk
2	Beer, Bread
3	Beer, Coke, Diaper, Milk
4	Beer, Bread, Diaper, Milk
5	Coke, Diaper, Milk

Rules Discovered:
{Milk} --> {Coke}
{Diaper, Milk} --> {Beer}

Figure 1.10: Basket Transactions

The Business Analytics Framework

- Let the rule discovered be
$$\{Diapers\} \rightarrow \{Beer\}$$

- Beer as consequent: Can be used to determine what should be done to **boost its sales**.

- Diapers in the antecedent: Can be used to see which products **would be affected** if the store discontinues selling Diapers.

- Diapers in antecedent and Beer in consequent: Can be used to see what products should be sold with Diapers to **promote sale** of Beer

The Business Analytics Framework

Example 1.8: Supermarket Shelf Management

- Goal: To identify items that are **bought together by sufficiently many customers**.

- Approach: Process the point-of-sale data collected with barcode scanners to find dependencies among items.

- A classic rule --
  - If a customer buys diaper and milk, then he is very likely to buy beer.

The Business Analytics Framework

- Identify which Menu items are ordered **frequently with each other** such that a promo meal can be launched.
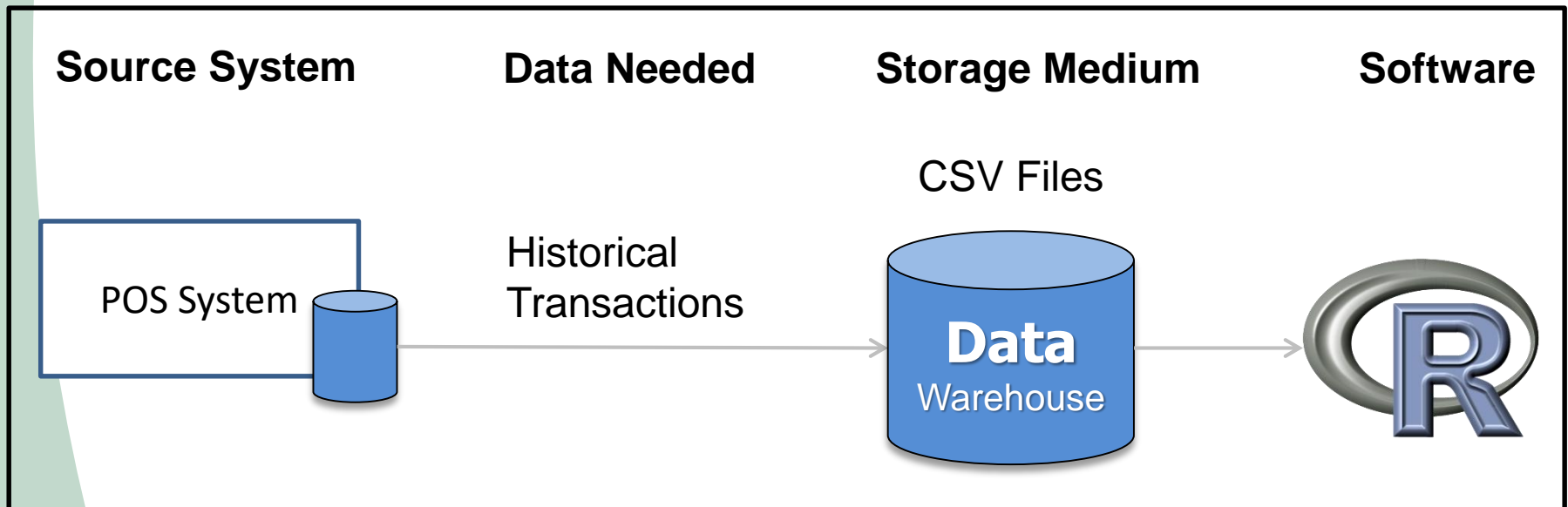
Figure 1.11: Restaurant Promo Pairings

The Business Analytics Framework

- Given a set of objects, with each object associated with its own timeline of events, find rules that predict strong **sequential dependencies among different** events.

$$(A\ \ B)\ \ \ \ (C)\ \ \ \ \ \ (D\ \ E)$$

Figure 1.12: Basket Transactions

- Rules are formed by first discovering patterns. Event occurrences in the patterns are governed by timing constraints.

The Business Analytics Framework

```
21  <{"DEVICE CONFIGURATION"},{"SUCCESSFUL NOT INTERESTED"}>
22     <{"DEVICE CONFIGURATION"},{"SUCCESSFUL INTERESTED"}>
23      <{"MECHANICS PROCEDURE"},{"SUCCESSFUL INTERESTED"}>
24                          <{"SHORT CALL"},{"SHORT CALL"}>
25       <{"MECHANICS PROCEDURE"},{"MECHANICS PROCEDURE"}>
26     <{"DEVICE CONFIGURATION"},{"DEVICE CONFIGURATION"}>
27    <{"SUCCESSFUL INTERESTED"},{"DEVICE CONFIGURATION"}>
28        <{"UNCOMPLETED CALL"},{"DEVICE CONFIGURATION"}>
29             <{"ACCOUNT DETAILS"},{"BILLING INQUIRY"}>
30             <{"BILLING INQUIRY"},{"BILLING INQUIRY"}>
31        <{"AFTERSALES REQUEST"},{"AFTERSALES REQUEST"}>
32         <{"BILLING INQUIRY"},{"AFTERSALES REQUEST"}>
33         <{"BILLING INQUIRY"},{"ACCOUNT DETAILS"}>
```
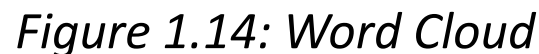
Figure 1.13: Hotline Sequence Transactions

The Business Analytics Framework

Definition 1.14: Text Mining

- Finding **frequently** occurring words from unstructured data, e.g. word files, reviews, journals, articles.

The Business Analytics Framework

Example 1.11: Word Cloud

Figure 1.14: Word Cloud

The Business Analytics Framework

Definition 1.15: Social Media Sentiment Analysis

- Identifying **sentiment** of a customer on a specific product using social media or text mining

The Business Analytics Framework

Example 1.12: Sentiment Analysis Map
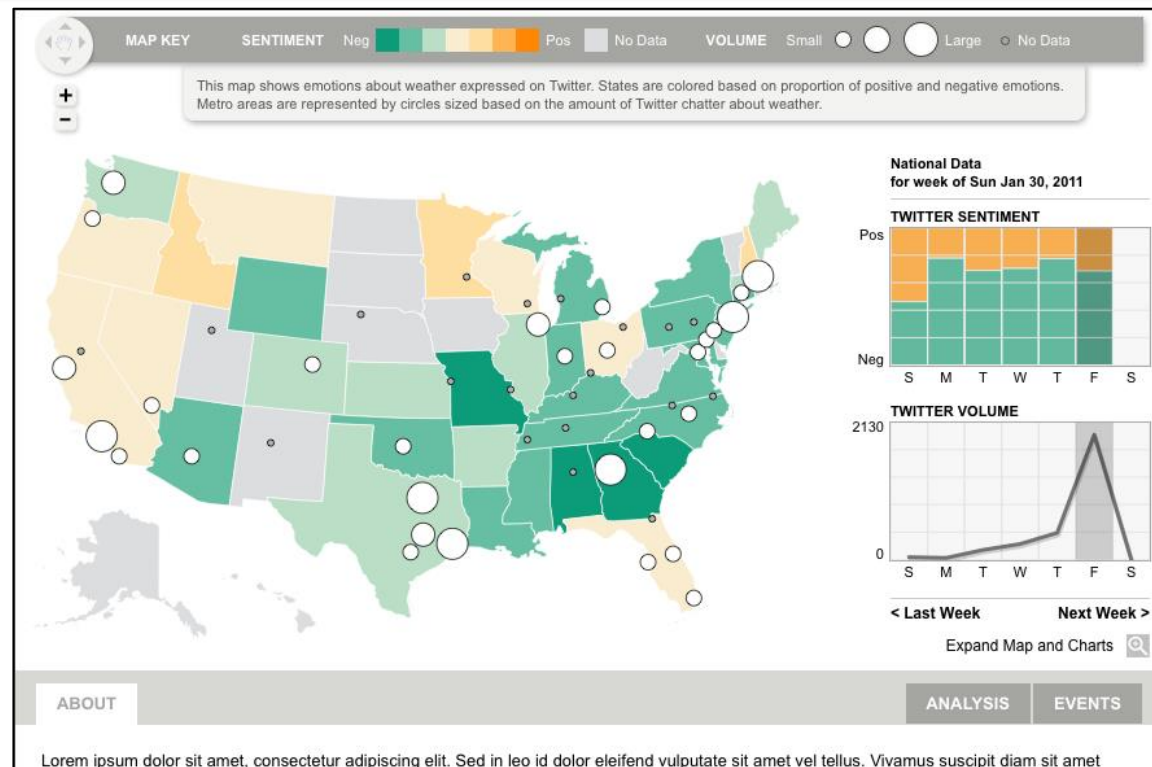
Figure 1.15: Sentiment Analysis

The Business Analytics Framework

- Types of BA According to Purpose:
  - **Exploratory Data Analysis** (Descriptive Analytics)
    - *Tell Me What has Happened and Why*
    - *Tell Me What is Happening Right Now*
  - **Data Mining** (Predictive Analytics)
    - *Tell Me What is Likely to Happen*
    - *Tell Me Something Interesting Without Me Asking*
  - **Optimization/Simulation** (Prescriptive Analytics)
    - *Tell Me What Might Have Happened*
    - *Tell Me What is the Best Solution*

The Business Analytics Framework

Definition 1.16: Optimization

- Prescriptive Analytics or Optimization is an application of Business Analytics that recommends the **optimal solution** to a problem given constraints.

- This application also seeks to find the **best solution** given multiple what-if scenarios
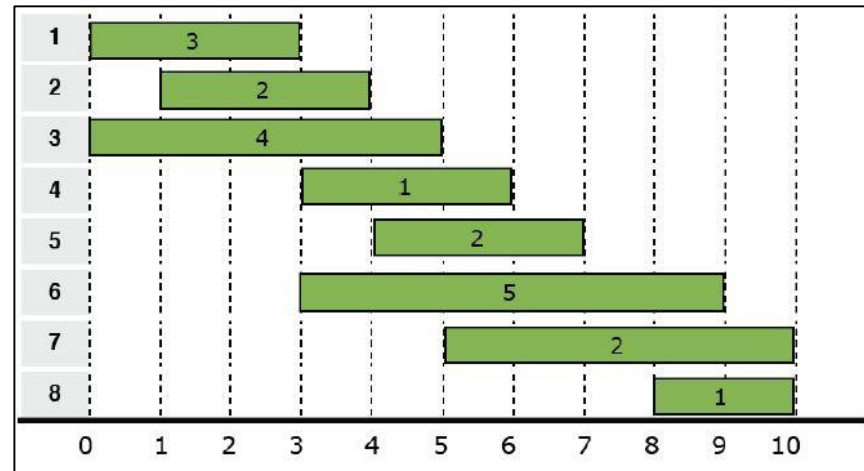
The Business Analytics Framework

Example 1.13: Logistics Optimization

- Find the best route for the delivery of bottled water at 2 PhP per bottle maximum logistics cost
  - Gather truck variable and fixed costs
  - Gather alternative routes
  - Use optimization to find best route given constraints

The Business Analytics Framework

- Determine optimal staffing levels for nurses to man a floor at a hospital
  - Gather historical number of patients per room per floor
  - Estimate needed manpower per shift
  - Identify scheduling constraints
  - Determine shifts and number of nurses to man the floor using optimization

Figure 1.16: Scheduling Chart

58

The Business Analytics Framework

Example 1.15: Ship Optimization for a Cement Company

- Design **optimal assignment** of time-chartered ships to deliver cement bags from plant port in Iligan to different ports in the Philippines.
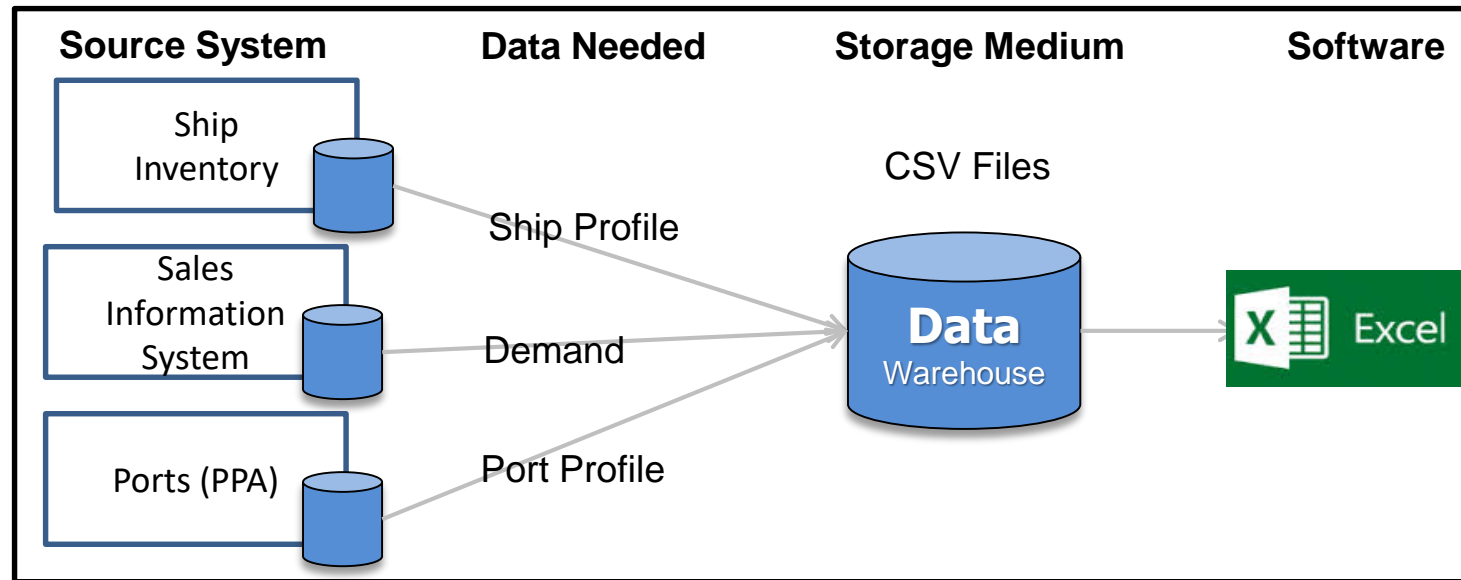

Source System	Data Needed	Storage Medium	Software
Ship Inventory	Ship Profile	CSV Files	Excel
Sales Information System	Demand	Data Warehouse	
Ports (PPA)	Port Profile		

Figure 1.17: Ship Optimization Framework

The Business Analytics Framework

Definition 1.17: Simulation

- It is an **imitation** of the operation of a real world process or system over time.

- Involves the generation of an artificial alternate of a system and **generating inferences** of the real system
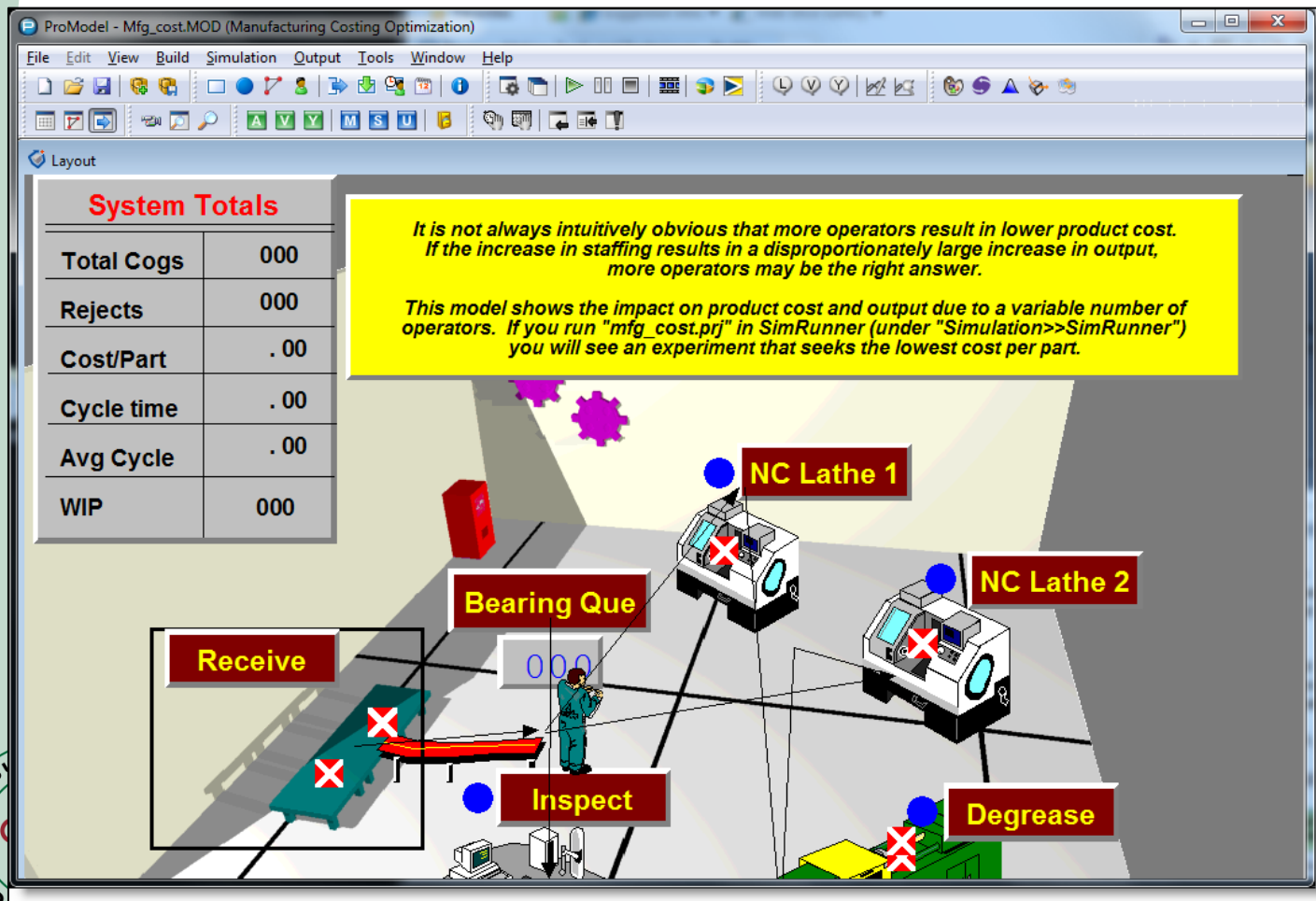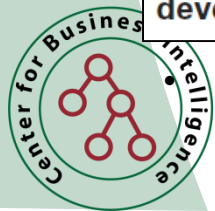
The Business Analytics Framework

Figure 1.18: Simulation Example

The Business Analytics Framework

Table 1.3: Types of Analytics in Different Business Functions

FUNCTION	DESCRIPTION	EXEMPLARS
Supply chain	Simulate and optimize supply chain flows; reduce inventory and stock-outs.	Dell, Wal-Mart, Amazon
Customer selection, loyalty, and service	Identify customers with the greatest profit potential; increase likelihood that they will want the product or service offering; retain their loyalty.	Harrah's, Capital One, Barclays
Pricing	Identify the price that will maximize yield, or profit.	Progressive, Marriott
Human capital	Select the best employees for particular tasks or jobs, at particular compensation levels.	New England Patriots, Oakland A's, Boston Red Sox
Product and service quality	Detect quality problems early and minimize them.	Honda, Intel
Financial performance	Better understand the drivers of financial performance and the effects of nonfinancial factors.	MCI, Verizon
Research and development	Improve quality, efficacy, and, where applicable, safety of products and services.	Novartis, Amazon, Yahoo

Source: Davenport, Thomas H. "Competing on analytics." *Harvard Business Review* 84.1 (2006): 98.

Outline for This Session

- Introductions
- What is Business Analytics?
- The Business Analytics Framework
- **Tools of Business Analytics**
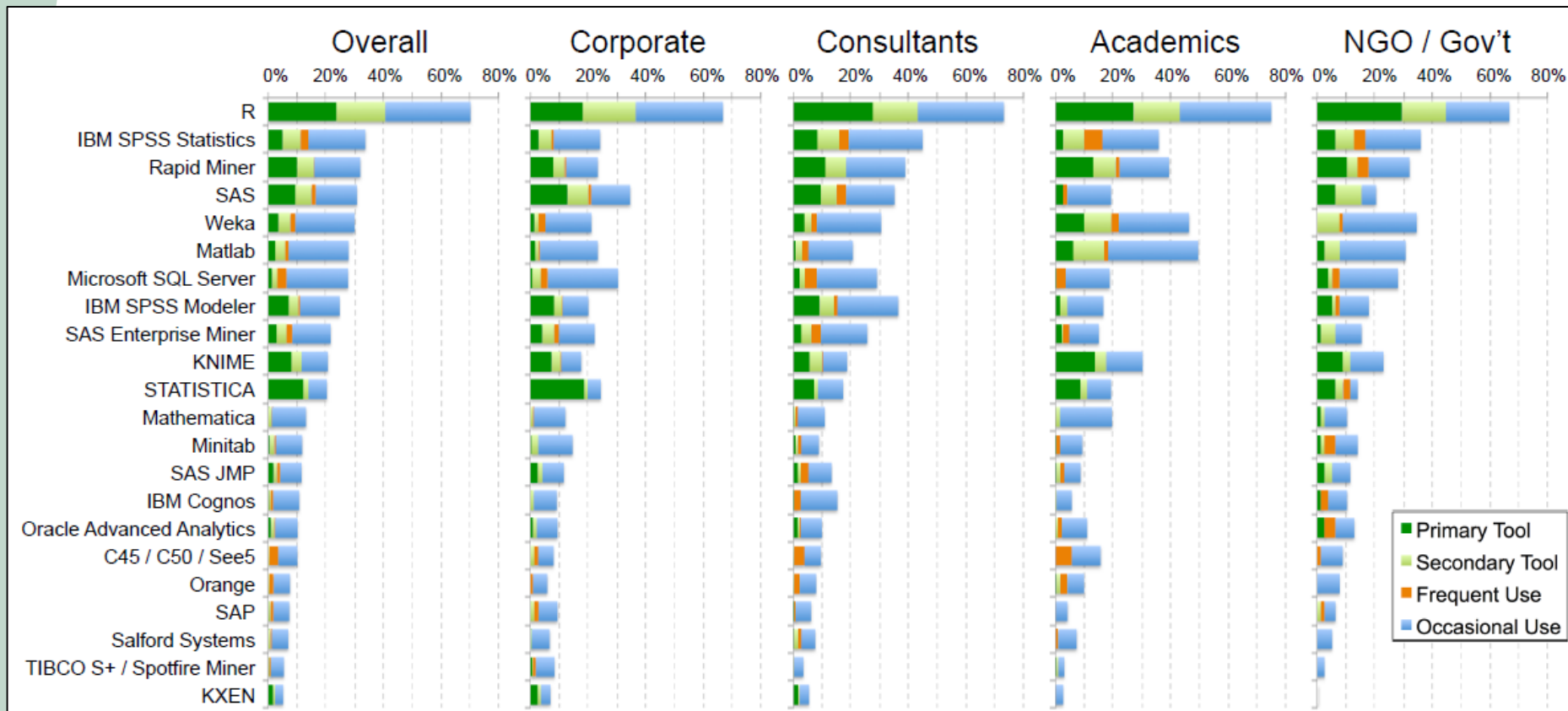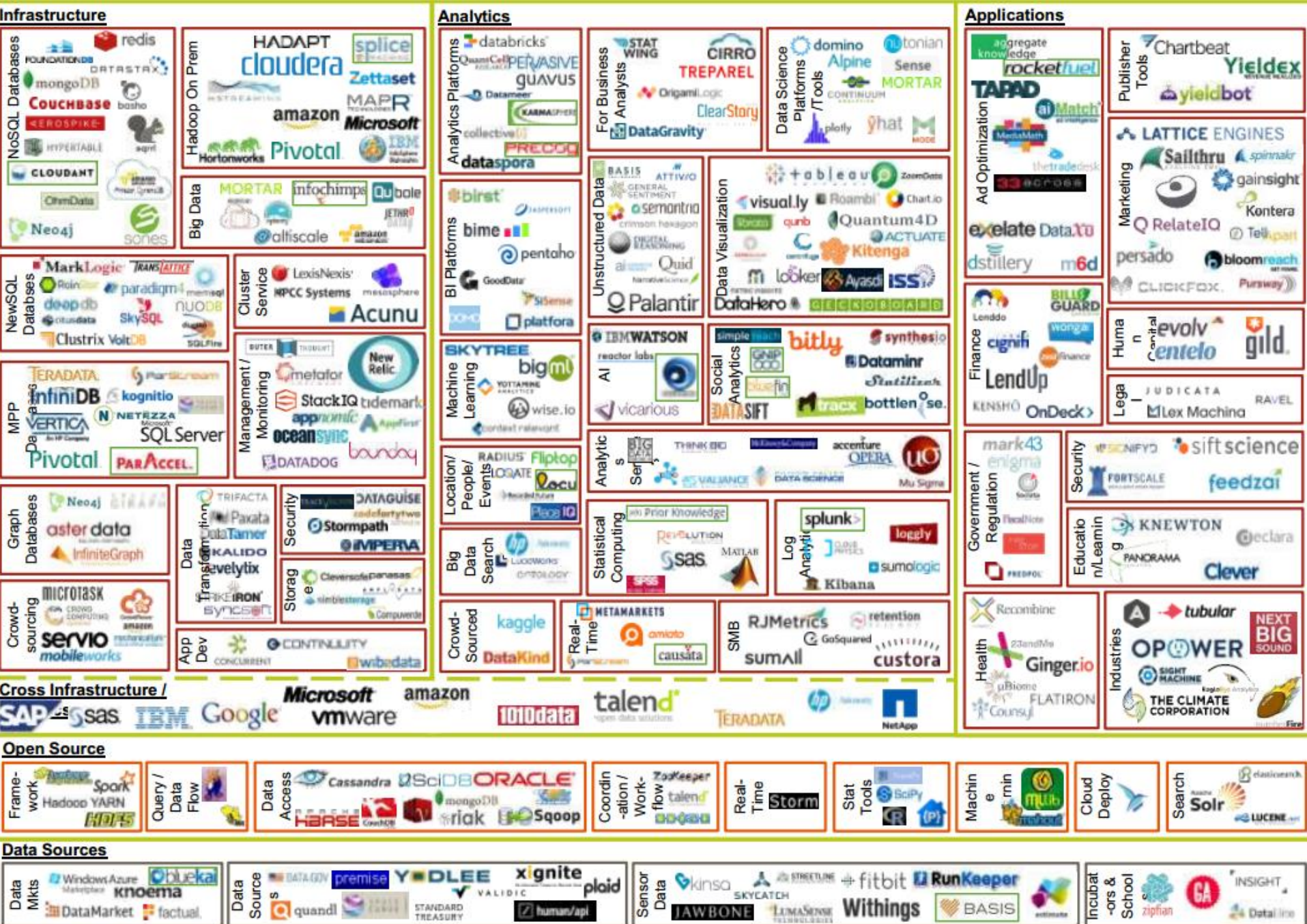- Case Study

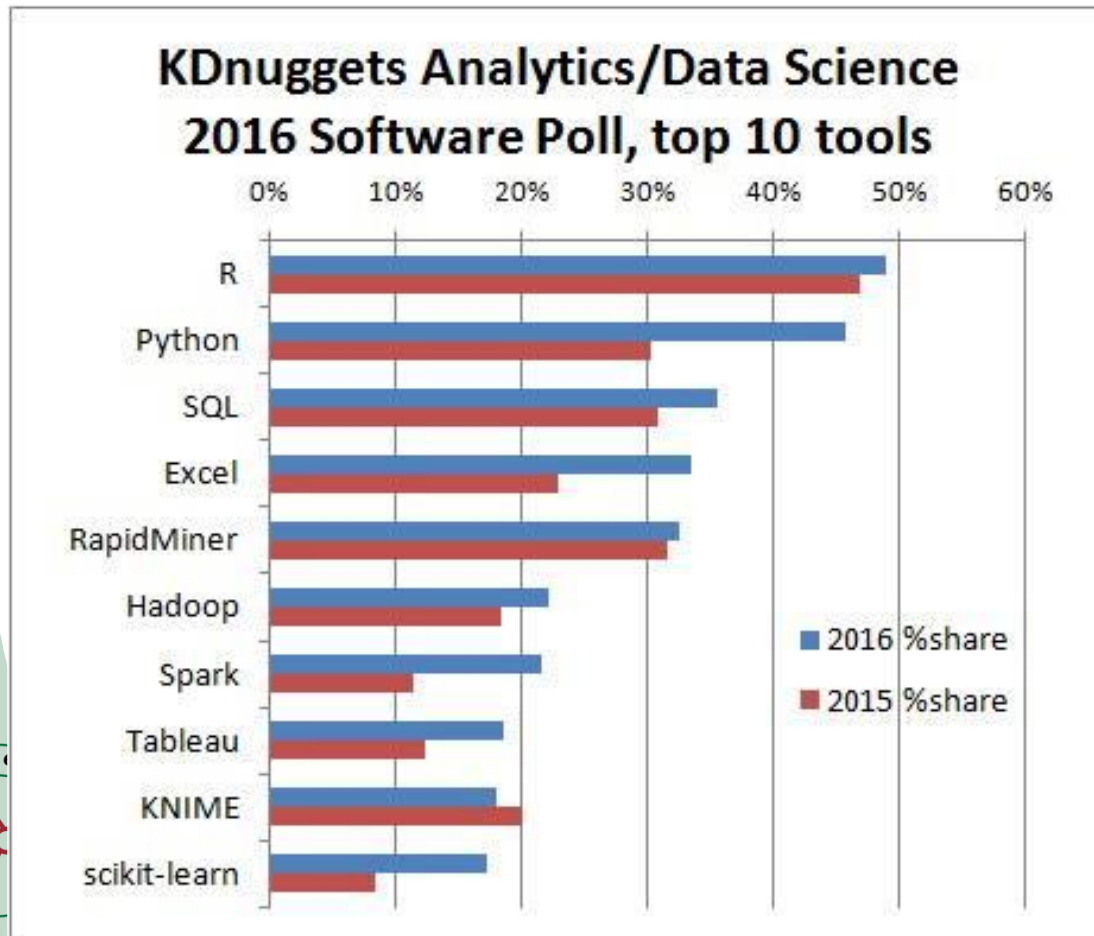# Tools for Business Analytics

Figure 1.19: Tools of Business Analytics

http://www.rexeranalytics.com/Data-Miner-Survey-2013-Intro.html

https://practicalanalytics.files.wordpress.com/2014/09/bigdatauniverse.png

Tools for Business Analytics

- KDNuggets Survey 2016: The top 10 tools **by share of users** were:

http://www.kdnuggets.com/2015/05/poll-r-rapidminer-python-big-data-spark.html

Tools for Business Analytics

- R is a programming language for **computing and analysis**
- RStudio IDE is a powerful and productive user interface for R. It's free and open source, and works great on Windows, Mac, and Linux.
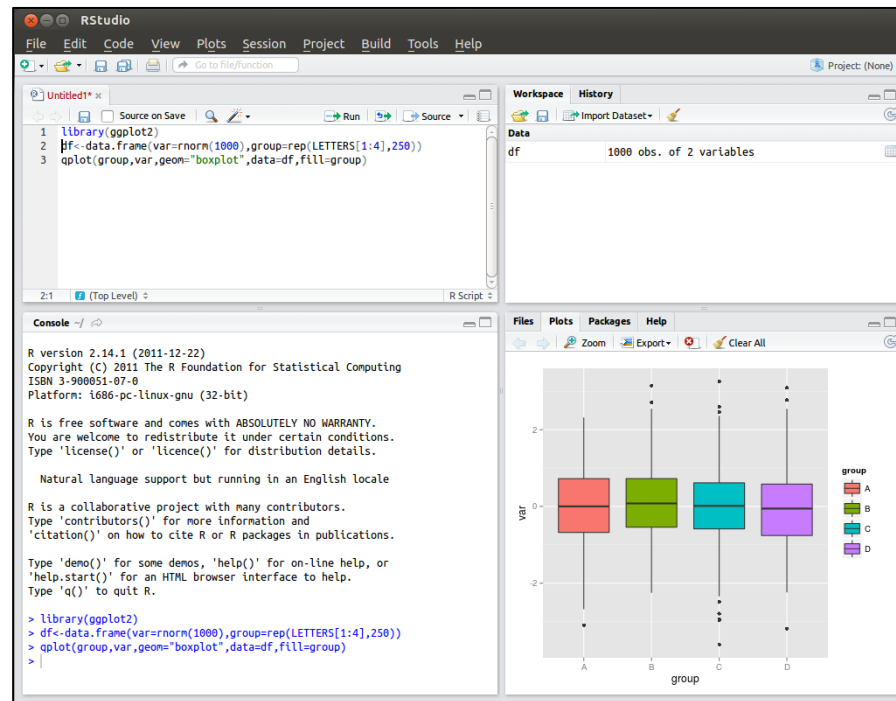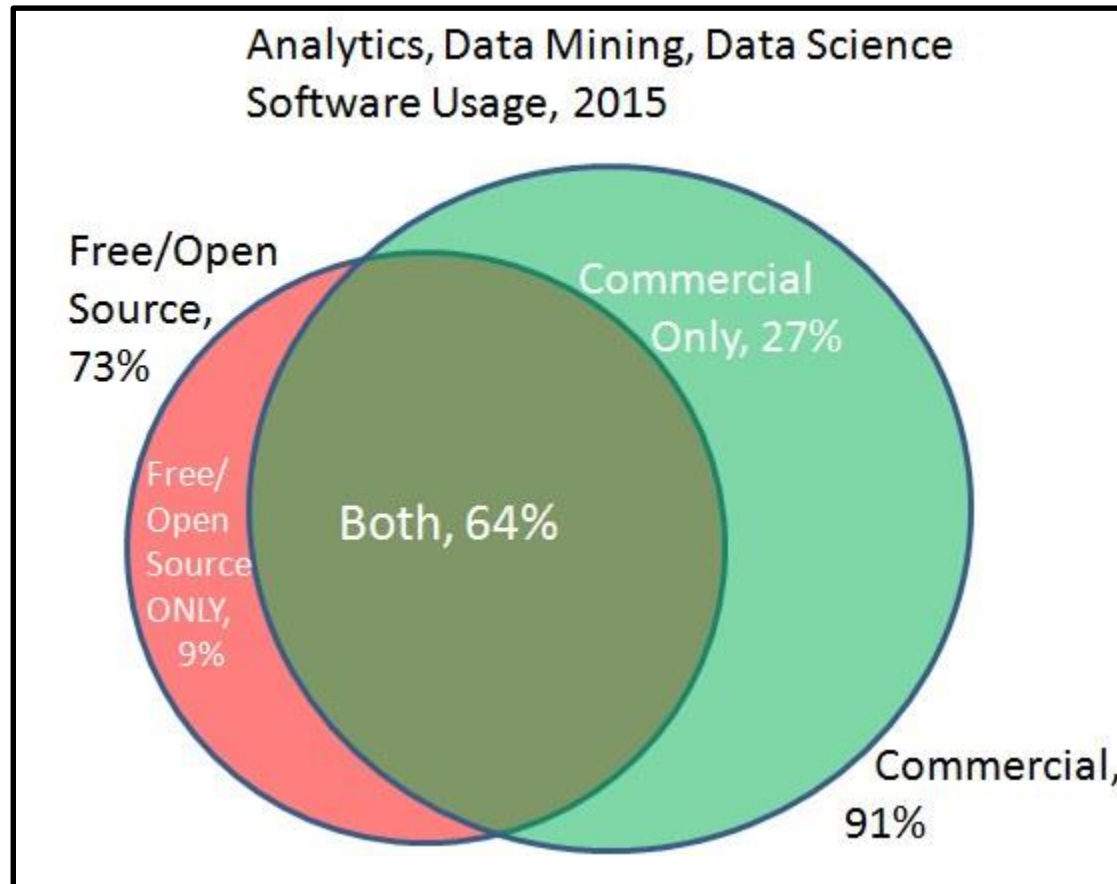

Figure 1.20: R Studio GUI

Tools for Business Analytics

Analytics, Data Mining, Data Science Software Usage, 2015

Free/Open Source, 73%

Commercial Only, 27%

Free/Open Source ONLY, 9%

Both, 64%

Commercial, 91%

Figure 1.21: Distribution of open and paid software

http://www.kdnuggets.com/2015/05/poll-r-rapidminer-python-big-data-spark.html

Outline for This Session

- Introductions
- What is Business Analytics?
- The Business Analytics Framework
- Tools of Business Analytics
- **Case Study**

Case Study

- **Norfolk Southern**
  - **Question 1**: What were the top 3 "big picture" business motivators that drove that company to begin or expand their BA efforts presented in the case?
  - **Question 2**: What were the 4 most valuable <u>specific business functions</u> delivered by the business intelligence initiative?
  - **Question 3**: Consider our three BA types. Which of the three were done <u>according to the case write-up</u>?
  - **Question 4**: Please give examples of what that company might do in a future version in <u>each</u> of the types of BA
- Discuss as a group and create a PPT presentation to identify potential BA initiatives that management can pursue.

Outline for This Session

- Introductions

- What is Business Analytics?

- The Business Analytics Framework

- Tools of Business Analytics

- Case Study

References

- Simon, Alan. CIS 391 PPT Slides
- Tan et al. Intro to Data Mining Notes
- Runger, G. IEE 520 notes
- http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/
- http://www.gartner.com/imagesrv/cio/pdf/cio_agenda_insights2015.pdf
- http://www.umsl.edu/~sauterv/DSS4BI/links/sas_defining_business_analytics_wp.pdf
- http://www.umsl.edu/~sauterv/DSS4BI/links/sas_defining_business_analytics_wp.pdf
- http://www.datasciencecentral.com/m/discussion?id=6448529%3ATopic%3A249904