Lecture 1: Introduction to Business Intelligent Analytics

What are Data Analytics?

- Analytic is the use of:
  - Data,
  - Information technology,
  - Statistical analysis,
  - Quantitative methods,
  - Mathematical or computer-based models.
- To help managers gain improved insight about their business operations and make better, fact-based decisions.
- Business Analytics (BA) is a subset of Data Analytics.

Business Analytics applications:

- Management of customer relationships.
- Financial and marketing activities.
- Supply chain management.
- Human resource planning.
- Pricing decisions.
- Sport team game strategies.

Importance of Business Analytics:

- There is a strong of relationship of BA with:
  - Profitability of businesses.
  - Revenue of businesses.
  - Shareholder return.
- BA enhances understanding of data.
- BA is vital for businesses to remain competitive.
- BA enables creation of informative reports.

Scope of Business Analytics

- Descriptive analytics:
  - Uses data to understand past and present.
- Predictive analytics:
  - Analyzes past performance.
- Prescriptive analytics:
  - Uses optimization techniques.

Scope of Business Analytics - Retail Markdown Decisions

- Most department stores clear seasonal inventory by reducing prices.
- The question is:
  - When to reduce the price and how much?

- Descriptive analytics: examine historical data for similar products (prices, unit sold, advertising, etc.)
- Predictive analytics: predict sales based on price.
- Prescriptive analytics: find the best sets of pricing and advertising to maximize sales revenues.

Data for Business Analytics

- DATA:
  - Collected facts and figures.
- DATABASE:
  - Collection of computer files containing data.
- INFORMATION:
  - Comes from analyzing data.

Data for Business Analytics

- Metrics are used to quantify performance.
- Measures are numerical values of metrics.
- Discrete metrics involve counting.
  - On time or not on time.
  - Number or proportion of on time deliveries.
- Continuous metrics are measured on a continuum.
  - Delivery time.
  - Package weight.
  - Purchase price.

Data for Business Analytics

- A Sales Transaction Database File:

	A	B	C	D	E	F	G	H
1	**Sales Transactions: July 14**							
2								
3	**Cust ID**	**Region**	**Payment**	**Transaction Code**	**Source**	**Amount**	**Product**	**Time Of Day**
4	10001	East	Paypal	93816545	Web	$20.19	DVD	22:19
5	10002	West	Credit	74083490	Web	$17.85	DVD	13:27
6	10003	North	Credit	64942368	Web	$23.98	DVD	14:27
7	10004	West	Paypal	70560957	Email	$23.51	Book	15:38
8	10005	South	Credit	35208817	Web	$15.33	Book	15:21
9	10006	West	Paypal	20978903	Email	$17.30	DVD	13:11
10	10007	East	Credit	80103311	Web	$177.72	Book	21:59
11	10008	West	Credit	14132683	Web	$21.76	Book	4:04
12	10009	West	Paypal	40128225	Web	$15.92	DVD	19:35
13	10010	South	Paypal	49073721	Web	$23.39	DVD	13:26

Records

Entities

Fields or Attributes

Figure 1.1

What is Big Data?

- Information from multiple internal and external sources:
  - Transactions.
  - Social media.
  - Enterprise content.

    - Sensors.
    - Mobile devices.
- Companies leverage data to adapt products and services to:
    - Meet customer needs.
    - Optimize operations.
    - Optimize infrastructure.
    - Find new resource of revenue.
    - Can reveal more patterns and anomalies.
- IBM estimates that by 2015 4.4 million jobs will be created globally to support Big Data.
    - 1.9 million of these jobs will be in United States.

Types of Data

- When collecting or gathering data we collect data from individuals' cases on variables.
- A variable is a unit of data collection whose value can vary.
- Variables can be defined into types according to the level of mathematical scaling that can carried out on the data.
- There are four types of data or levels of measurement:

1. Categorical (Nominal)	**2. Ordinal**
3. Interval	**4. Ratio**

Data for Business Analytics

- Classifying Data Elements in a Purchasing Database

	A	B	C	D	E	F	G	H	I	J
1	**Purchase Orders**									
2										
3	**Supplier**	**Order No**	**Item No.**	**Item Description**	**Item Cost**	**Quantity**	**Cost per order**	**A/P Terms (Month**	**Order Date**	**Arrival Date**
4	Spacetime Technologies	A0111	6489	O-Ring	$ 3.00	900	$ 2,700.00	25	10/10/11	10/18/11
5	Steelpin Inc.	A0115	5319	Shielded Cable/ft.	$ 1.10	17,500	$ 19,250.00	30	08/20/11	08/31/11
6	Steelpin Inc.	A0123	4312	Bolt-nut package	$ 3.75	4,250	$ 15,937.50	30	08/25/11	09/01/11
7	Steelpin Inc.	A0204	5319	Shielded Cable/ft.	$ 1.10	16,500	$ 18,150.00	30	09/15/11	10/05/11
8	Steelpin Inc.	A0205	5677	Side Panel	$195.00	120	$ 23,400.00	30	11/02/11	11/13/11
9	Steelpin Inc.	A0207	4312	Bolt-nut package	$ 3.75	4,200	$ 15,750.00	30	09/01/11	09/10/11
10	Alum Sheeting	A0223	4224	Bolt-nut package	$ 3.95	4,500	$ 17,775.00	30	10/15/11	10/20/11
11	Alum Sheeting	A0433	5417	Control Panel	$255.00	500	$ 127,500.00	30	10/20/11	10/27/11
12	Alum Sheeting	A0443	1243	Airframe fasteners	$ 4.25	10,000	$ 42,500.00	30	08/08/11	08/14/11
13	Alum Sheeting	A0446	5417	Control Panel	$255.00	406	$ 103,530.00	30	09/01/11	09/10/11
14	Spacetime Technologies	A0533	9752	Gasket	$ 4.05	1,500	$ 6,075.00	25	09/20/11	09/25/11
15	Spacetime Technologies	A0555	6489	O-Ring	$ 3.00	1,100	$ 3,300.00	25	10/05/11	10/10/11

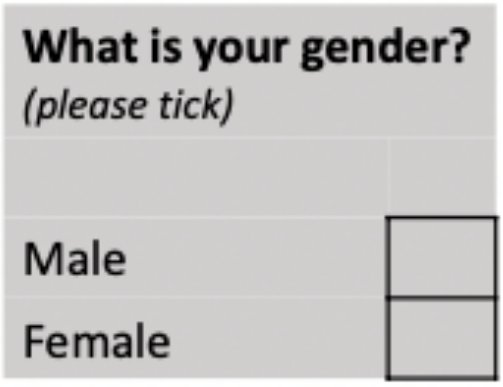
Figure 1.2

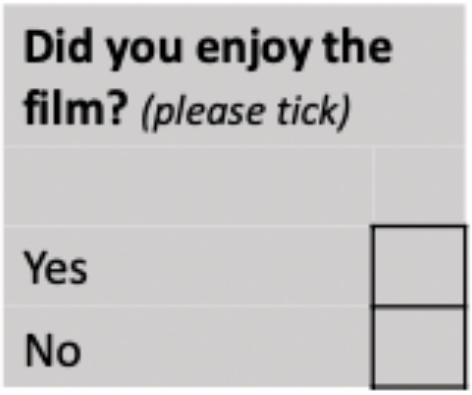
Categorical Categorical Categorical Categorical Ratio Ratio Ratio Ratio Interval Interval

Categorical (Nominal) data

- Nominal or categorical data is data that comprises of categories that cannot be rank ordered – each category is just different.
- The categories available cannot be placed in any order and no judgement can be made about the relative size or distance from one category to another.
    - Categories bear no quantitative relationship to one another.

    - Examples:
        - Customer's location (American, Europe, Asia).
        - Employee classification (manager, supervisor, associate).
- What does this mean? No mathematical operations can be performed on the data relative to each other.
- Therefore, nominal data reflect qualitative differences rather than quantitative.





- Systems for measuring nominal data must ensure that each category is mutually exclusive, and the system of measurement needs to be exhaustive.
- Variables that have only 2 responses, i.e., Yes or No, are known as dichotomies.

Ordinal data

- Ordinal data is data that comprises of categories that can be ranked ordered.
- Similarly with nominal data, the distance between each category cannot be calculated but the categories can be ranked above or below each other.
    - No fixed units of measurement.
    - Examples:
        - College football rankings.
        - Survey responses (poor, average, good, very good, excellent).
- What does this mean? Can make statical judgements and perform limited math.

How satisfied are you with the level of service you have received? *(please tick)*	
Very satisfied	
Somewhat satisfied	
Neutral	
Somewhat dissatisfied	
Very dissatisfied	

Interval and ratio data

- Both interval and ratio are examples of scale data.
- Scale data:
  - Data is in numeric format ($50, $100, $200).
  - Data that can be measured on a continuous scale.
  - The distance between each can be observed and as a result measured.
  - The data can be placed in rank order.

Interval data

- Ordinal data but with constant differences between observations.
- Ratios are not meaningful.
- Examples:
  - Time – moves along a continuous measure of seconds, minutes, and so on, and is without a zero point of time.
  - Temperature – moves along a continuous measure of degrees and is without a true zero.
  - SAT scores.

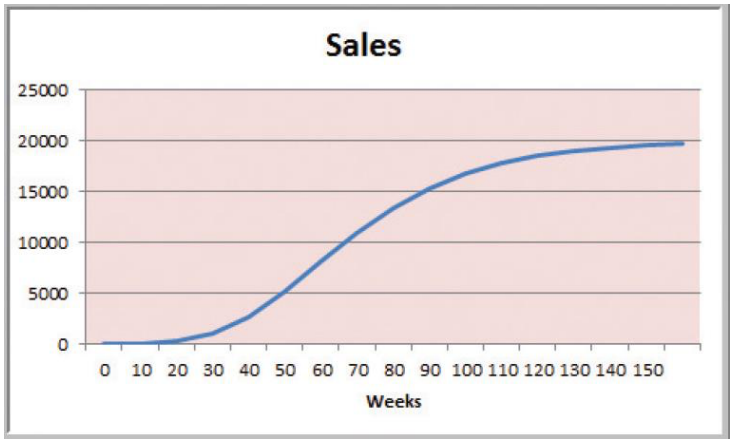
Ratio data

- Ratio data measured on a continuous scale and does have a natural zero point.
  - Ratios are meaningful.
  - Examples:
    - Monthly sales.
    - Delivery times.
    - Weight.
    - Height.

        - Age.

Types of analytics – Decision Models

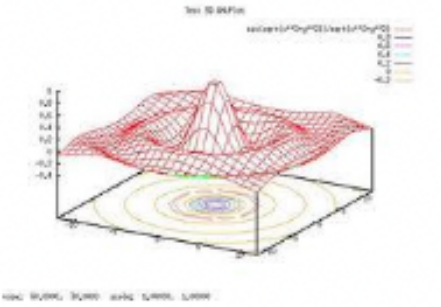
- Model:
    - An abstraction or representation of a real system, idea, or object.
    - Captures the most important features.
    - Can be:
        - written or verbal description,
        - a visual display,
        - a mathematical formula,
        - a spreadsheet representation.
- Decision Models:



    - Is a model used to understand, analyze, or facilitate decision making.
    - Types of model input:
        - Data.
        - Uncontrollable variables.
        - Decision variables (controllable).
- Descriptive Decision Models:
    - Simply tell “what is” and describe relationships.
    - Do not tell managers what to do.

Descriptive Analytics

- Description analytics, such as reporting/OLAP, dashboards, and data visualization, have been widely used for some time.
- They are the core of traditional BI.

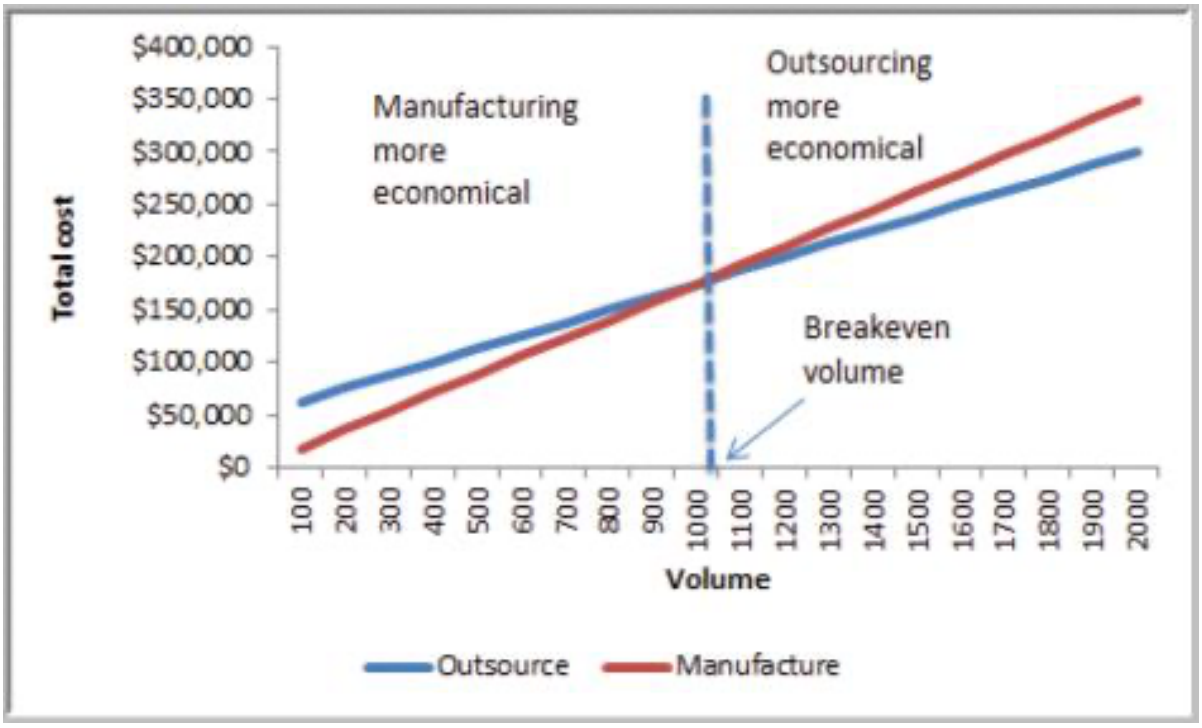


What has occurred?

Descriptive analytics, such as data visualization, is important in helping users interpret the output from predictive and predictive analytics.

Decision Models

- A Break-even Decision Model:
  - $TC(manufacturing) = \$50.000 + \$125 * Q$
  - $TC(outsourcing) = \$175 * Q$
- Break-even Point:
  - Set $TC(manufacturing) = TC(outsourcing)$.
  - Solve for Q = 1000 units.

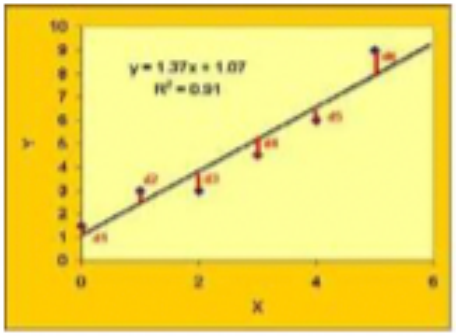


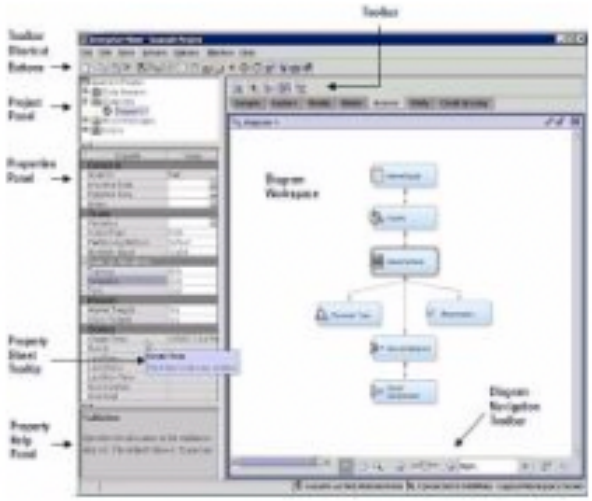
Predictive Decision Models

- Predictive Decision Models often incorporate uncertainty to help managers analyze risk.
- Aim to predict what will happen in the future.
- Uncertainty is imperfect knowledge of what will happen in the future.
- Risk is associated with the consequences of what happens.

Predictive Analytics

- Algorithms for predictive analytics, such as regression analysis, machine learning, and neural networks, have also been around for some time.
- Prescriptive analytics are often referred to as advanced analytics.





What will occur?

- Marketing is the target for many predictive analytics applications.
- Descriptive analytics, such as data visualization, is important in helping users interpret the output from predictive and prescriptive analytics.

Decision Models

- A Linear Demand Prediction Model:
    - As price increases, demand falls.

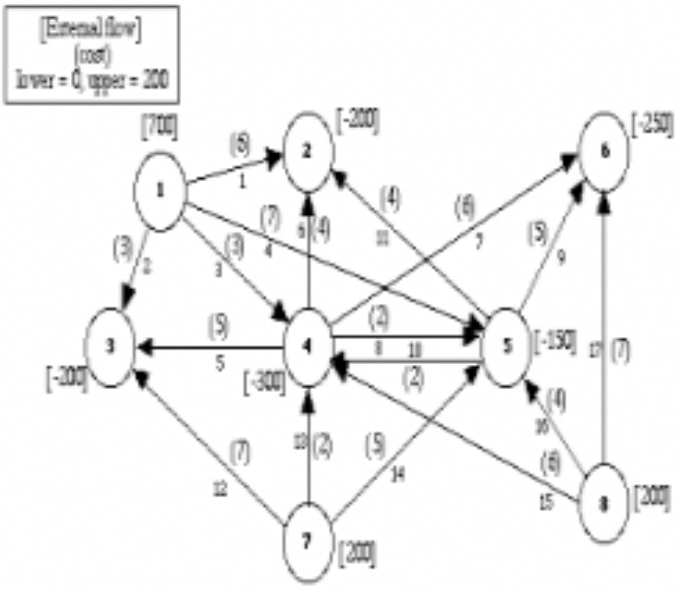
- A Non-Linear Demand Prediction Model:

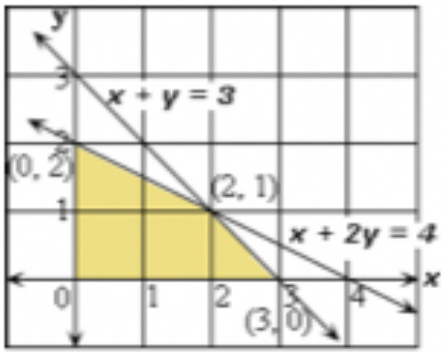
- Assumes price elasticity (constraint ratio of % change in demand to % change in price).

- Prescriptive Decision Models help decision makers identify the best solution.
  - Optimization – finding values for decision variables that minimize (or maximize) something such as cost (or profit).
  - Objective function – the equation that minimizes (or maximizes) the quantity of interest.
  - Constraints – limitation or restrictions.
  - Optimal solution – values of decision variables at the minimum (or maximum) point.

Prescriptive Analytics

- Prescriptive analytics are often referred to as advanced analytics.
- Regression analysis, machine learning and neural networks.
- Often for the allocation of scarce resources.





What should occur?

- For example, the use of mathematical programming for revenue management is common for organizations that have "perishable" good (e.g., rental cars, hotel rooms, airline seats).
- Harrah's has been using revenue management for hotel room pricing for some time.

Organizational Transformation

- Brought about by opportunity or necessity.
- The firm adopts a new business model enabled by analytics.
- Analytics are competitive requirements.

2013 Academic Research

- A 2011 TDWI report on Big Data Analytics found that 85% of respondents indicated that their firms would be using advanced analytics within 3 years.
- A 2011 IBM/MIT Sloan Management Review research study found that top performing companies in their industry are much more likely to use analytics rather than intuition across the widest range of possible decisions.

Conditions that Lead to Analytics-based Organizations

- The nature of the industry.
- Seizing an opportunity.
- Responding to a problem.

Complex Systems

- Tackle complex problems and provide individualized solutions.
- Products and services are organized around the needs of individual customers.
- Dollar value of interactions with each customer is high.
- There is considerable interaction with each customer.
- Examples: IBM, World Bank, Halliburton.

Volume Operations

- Serves high-volume markets through standardized products and services.
- Each customer interaction has a low dollar value.
- Customer interactions are generally conducted through technology rather than person-to-person.
- Are likely to be analytics-based.

- Examples: Amazon, eBay, Hertz.

The Nature of the Industry: Online Retailers

- BI Applications
    - Analysis of clickstream data.
    - Customer profitability analysis.
    - Customer segmentation analysis.
    - Product recommendations.
    - Campaign management.
    - Pricing.
    - Forecasting.
    - Dashboards.

The Nature of the Industry

- Online retailers like Amazon.com and Overstock.com are high volume operations who rely on analytics to complete.
- When you enter their sites, a cookie is placed on your PC and all clicks are recorded.
- Based on your blocks and any search terms, recommendation engines decide what products to display.
- After you purchase an item, they have additional information that is used in marketing campaigns.
- Customer segmentation analysis is used in deciding what promotions to send you.
- How profitable you are influencing how the customer care center treats you.
- A pricing team helps set prices and decides what prices are needed to clear out merchandise.
- Forecasting models are used to decide how many items to order for inventory.
- Dashboards monitor all aspects of organizational performance.

Analytics Help the Cincinnati Zoo Know Its Customers

- What management, organization, and technology factors were behind the Cincinnati Zoo losing opportunities to increase revenue?
- Why was replacing legacy point-of-sale systems and implementing a data warehouse essential to an information system solution?
- How did the Zoo benefit from business intelligence? How did it enhance operational performance and decision making? What role was played by predictive analytics?
- Visit the IBM Cognos Web site and describe the business intelligence tools that would be the most useful for the Zoo.

Table of Contents