Lecture 1

1. The content of employee emails is an example of a non-structured data source.
    a. True
    b. False

2. The increase in the amount of data available due to automatic data capture technologies is an example of Big Data velocity.
    a. True
    b. False

3. The increase in the rate at which data flow into an organization due to online sales, mobile computing, smartphones, and online social networks is an example of Big Data velocity.
    a. True
    b. False

4. The increase in the number of sources and data formats due to improved data capture technologies is an example of Big Data variety.
    a. True
    b. False

5. Big Data demands that data scientists and managers implement new techniques and, above all, a new way of thinking.

a. True

b. False

6. Databases and data warehouse technologies are an important part of business analytics.

   a. True

   b. False

7. Histograms are a common tool used in predictive analytics.

   a. True

   b. False

8. Predictive analytics uses statistical modeling to draw conclusions and predict behavior based on the assumption that what has happened in the past will continue to happen.

   a. True

   b. False

9. The examples of Target, LinkedIn, First Tennessee, and many other companies that have implemented data analytics initiatives indicate that business analytics programs are more successful in large corporations.

   a. True

   b. False

10. Spreadsheet modeling offers a high degree of data integrity and, as such, is an excellent tool for prescriptive analytics.
    a. True
    b. False

11. Big Data models can sometimes lead to bad business results, conclusions, and recommendations due to inaccurate information.
    a. True
    b. False

12. Mathematical models are often created through the process of abstraction and, as a result, can produce inaccurate results.
    a. True
    b. False

13. Big Data includes both structured and non-structured data sources. Which of the following is usually considered to be a structured data source?
    a. The content of emails
    b. The content of web pages
    c. The content of employee information
    d. The content of daily memos
    e. All of those

14. Big Data is usually characterized by three Vs: volume, velocity, and variety. Which of the following statements best describes the velocity of data?
    a. The amount of data available is significantly increased due to automatic data capture technologies.
    b. The rate at which data flow into an organization has increased due to online sales, mobile computing, smartphones, and online social networks.
    c. The number of sources and data formats has increased due to the increase of data capture technologies such as text and images from online social networks, images and voice from a camera, and voice recognition technologies, among others.
    d. All of those describe the velocity of Big Data.

15. Big Data is usually characterized by three Vs: volume, velocity, and variety. Which of the following statements best describes the variety of data?
    a. The amount of data available is significantly increased due to automatic data capture technologies.
    b. The rate at which data flow into an organization has increased due to online sales, mobile computing, smartphones, and online social networks.
    c. The number of sources and data formats has increased due to the increase of data capture technologies such as text and images from online social networks, images and voice from a camera, and voice recognition technologies, among others.
    d. All of those describe the volume of Big Data.

16. Big Data demands that data scientists and managers implement new techniques and, above all, a new way of thinking. Specifically, Big Data requires that:
    a. Managers think more anecdotally than probabilistically.
    b. Managers learn to focus more on the signals than on the noise.
    c. Managers apply business analytics more at the operational level than at the strategic level.
    d. All of those are required in the Big Data paradigm.

17. What is the difference between analytics and statistics?
    a. Analytics includes statistics and other components such as databases, data warehouses, and optimization techniques.
    b. Statistics includes analytics and other components such as databases, data warehouses, and optimization techniques.
    c. Analytics and statistics are two completely different fields of study.
    d. There is no difference; these concepts can be used interchangeably.

18. Which of the following is a major field of study included in business analytics?
    a. Descriptive analytics
    b. Predictive analytics
    c. Prescriptive analytics
    d. All of those

19. Which of the following fields of business analytics can be used to effectively store and efficiently retrieve information?
    a. Databases and data warehouses
    b. Descriptive analytics
    c. Predictive analytics
    d. Prescriptive analytics

20. Which of the following fields of business analytics can be used to report what has happened in the past?
    a. Databases and data warehouses
    b. Descriptive analytics
    c. Predictive analytics
    d. Prescriptive analytics

21. Which of the following fields of business analytics uses past data to create models that predict the future?
    a. Databases and data warehouses
    b. Descriptive analytics
    c. Predictive analytics
    d. Prescriptive analytics

22. Which of the following fields of business analytics aims to summarize a sample instead of using data to learn about the population that the sample of data is thought to represent?

a. Databases and data warehouses

b. Descriptive analytics

c. Predictive analytics

d. Prescriptive analytics

23. Which of the following is a tool commonly used in descriptive statistics?

a. Mean

b. Standard deviation

c. Variance

d. Stem and leaf diagrams

e. All of those

24. Which of the following is not a tool commonly used in descriptive statistics?

a. Histograms

b. Quartiles

c. Frequency distributions

d. Neural networks

e. All of those

25. Which of the following describes the change in the nature of management science to accommodate the need to process large volumes of data?

a. Reliance on spreadsheet modeling

b. Connection of optimization models with live operational databases

c. The use of distributed file systems technologies such as Hadoop to process a variety of data sources

d. All of those are used by management scientists to process large volumes of data.

26. Which of the following describes the change in the nature of management science to accommodate the need to process high-velocity data?

a. Reliance on spreadsheet modeling

b. Connection of optimization models with live operational databases

c. The use of distributed file systems technologies such as Hadoop to process a variety of data sources

d. All of those

27. Which of the following describes the change in the nature of management science to accommodate the need to process high variety of data?

a. Reliance on spreadsheet modeling

b. Connection of optimization models with live operational databases

c. The use of distributed file systems technologies such as Hadoop to process a variety of data sources

d. All of those

28. The examples of Target, LinkedIn, First Tennessee, and many other companies that have implemented data analytics initiatives indicate that:

a. Business analytics programs are more successful in large corporations.

b. Business analytics programs are more successful in medium-sized or small corporations.
c. Corporations must have a clear, step-by-step plan to implement any business analytics initiative.
d. All of those are lessons learned by organizations that have successfully implemented business analytics.

29. Which of the following statements does not reflect LinkedIn's approach to its "People You May Know" campaign?
    a. LinkedIn uses information about its members that is housed in operational databases.
    b. LinkedIn organizes information about its members in a data warehouse.
    c. LinkedIn uses prescriptive analytics to generate reports and discover patterns.
    d. LinkedIn uses predictive analytics to discover that speed is very important in receiving positive responses.
    e. All of those accurately describe LinkedIn's approach to its "People You May Know" campaign.

30. Assume that you are a regular customer of a restaurant. Which of the following best describes the information management component of business analytics used by the restaurant to make you a more valuable customer?
    a. Storing information about you and loading it into data warehouses
    b. Investigating correlations, sampling customer data, and summarizing data
    c. Determining the menu item that you will most likely order on your next visit

d. Offering coupons or discounts to maximize sales or profits

31. Assume that you are a regular customer of a grocery store. Which of the following best describes the descriptive analytics component used by the store to make you a more valuable customer?
    a. Storing information about you and loading it into data warehouses
    b. Investigating correlations, sampling customer data, and summarizing data
    c. Determining the grocery item that you will most likely purchase on your next visit
    d. Offering coupons or discounts to maximize sales or profits

32. Assume that you are a regular customer of a clothing store. Which of the following best describes the predictive analytics component used by the store to make you a more valuable customer?
    a. Storing information about you and loading it into data warehouses
    b. Investigating correlations, sampling customer data, and summarizing data
    c. Determining the clothing item that you will most likely order on your next visit
    d. Offering coupons or discounts to maximize sales or profits

33. Which of the following supports the assessment that descriptive and predictive analytics benefit from a high volume of data?
    a. Statistical analysis and reliability of predictions are better when the population size increases.

b. A forecasting method with many input factors can predict better than the one with only a few input factors.
c. More data beats better modeling.
d. All of those

34. Data variety is defined as a mix of different data sources in different formats. Which of the following does not generate high data variety?
    a. An automated toll system that captures photos, license plates, and driver information
    b. A sales transaction system that captures customer and product information
    c. A weather station that captures satellite inputs, temperature sources, and live video streaming of weather events
    d. All of those generate data variety.

35. Which of the following is considered to be a challenge that organizations face when analyzing Big Data?
    a. Privacy invasion
    b. Financial exposure
    c. Mistaking noise for the signal
    d. Poorly defined business problems
    e. All of those

36. Which of the following is an advantage of using Excel modeling for prescriptive analytics?

a. Spreadsheets are easy to use and available at universities and other organizations.
b. Spreadsheet modeling offers a high degree of data integrity.
c. Spreadsheet models can be more accurate than specialized optimization software programs.
d. All of those

37. Which of the following explains why Big Data models can sometimes lead to poor business results, conclusions, and recommendations?
a. Big Data is sometimes associated with inaccurate information.
b. Mathematical models are often created through the process of abstraction.
c. Mathematical models often allow relevant information to enter the model.
d. All of those