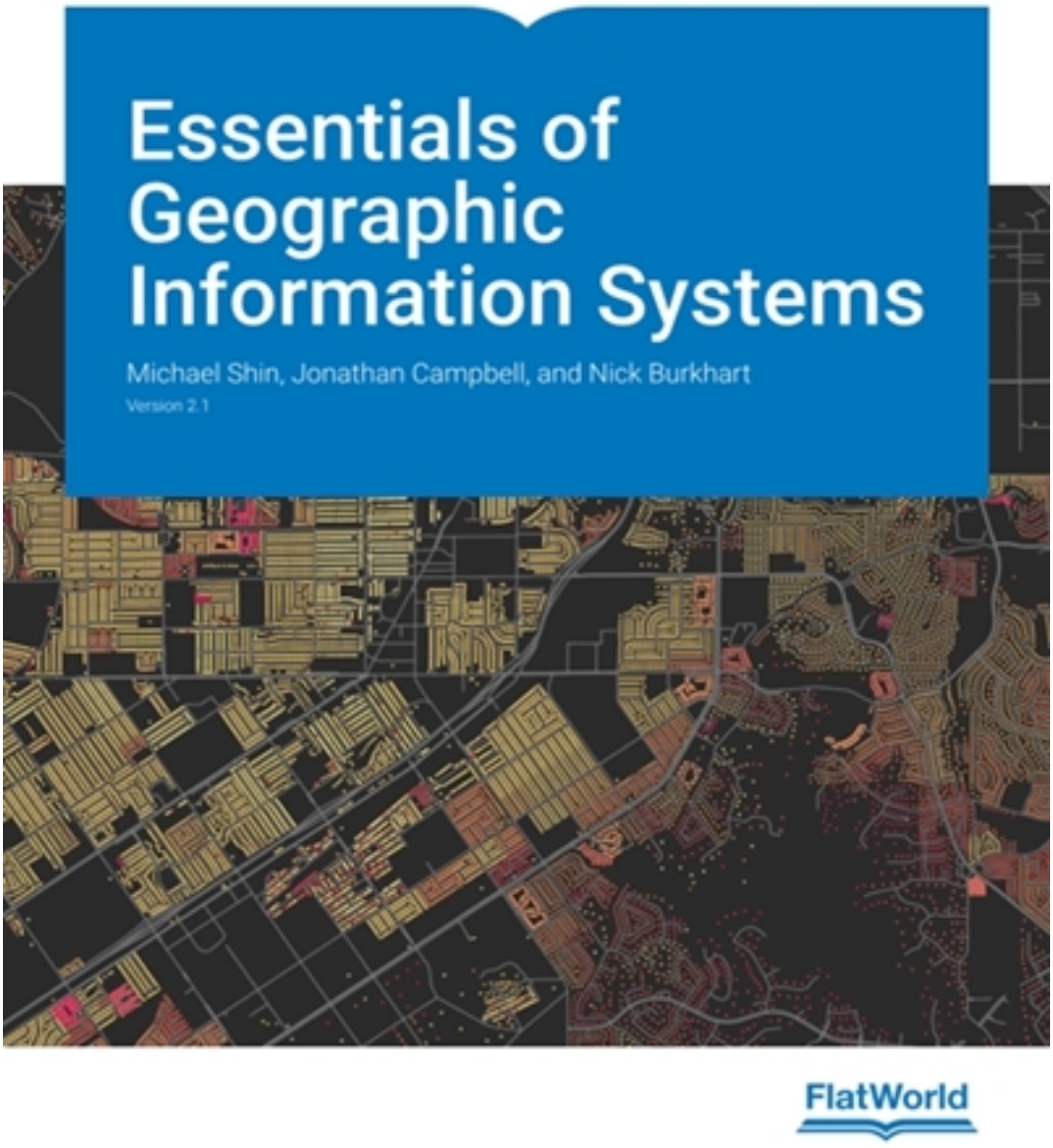
Test Bank for Essentials of Geographic Information Systems Version 2 1 2nd Edition by Shin



Test Bank

Chapter 2
Data, Information, and Where to Find Them

True/False Questions

1. Data refer to facts, measurements, characteristics, or traits of an object of interest.
   True; Easy

2. Information situated within analytical frameworks is known as data.
   False; Easy

3. Information refers to the knowledge of value obtained through the collection, interpretation, and/or analysis of data.
   True; Easy

4. Geographic data is also known as attribute data.
   False; Easy

5. Spatial data refers to data that describe geographic aspects of phenomena.
   True; Easy

6. Street addresses, postal codes, and political boundaries are classified as attribute data.
   False; Moderate

7. Attribute data are concerned with the nongeographic traits and characteristics of an object of interest.
   True; Easy

8. A GIS functions only on geographic data and is independent of attribute data.
   False; Moderate

9. A file extension is used to distinguish one computer file from another.
   True; Easy

10. A file named AnnualBudget.doc is in the Adobe portable document format.
    False; Easy

11. A file named image_file.xml is in the tagged image format.
    False; Easy

12. If Harry wants to store spatial data in a compressed archive, he will have to use a file with a .zip extension.
    True; Moderate

13. Files that end with the extension ".txt" contain alphanumeric text with special formatting.
    False; Easy

14. Each row of a flat file contains one and only one record.

True; Moderate

15. Columns in a flat file are called attributes.
True; Easy

16. A comma cannot be used as a delimiter within a flat file.
False; Easy

17. A grouping of multiple files to collect, organize, and analyze data is called a database.
True; Easy

18. Recording the temperature at noon every day in May in order to obtain the average temperature of the month is an example of primary data collection.
True; Moderate

19. A geologist working in Tanzania with geographic data released by the government of Tanzania is using primary data.
False; Moderate

20. Using primary data provides more time and cost savings than using secondary data.
False; Easy

21. Geospatial metadata refers to a class of metadata that contains information about the attributes of a data set.
False; Easy

22. In large organizations, the use of metadata increases redundancy.
False; Moderate

23. All GIS programs can read and use shapefiles.
False; Easy

24. Data that can be freely distributed are called public data.
True; Easy

25. Proprietary data are free to use.
False; Easy

26. Mean is the measure of central tendency that represents the most frequently occurring value in an array.
False; Easy

27. The mean of an array is also referred to as the average.
True; Easy

28. The measure of distribution of a variable is a summary of the frequency of values over the range of the dataset.
True; Easy

29. Given array: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}. The mean of the given array is 8.
False; Moderate

30. The following is a sorted array of exam scores: {45, 50, 55, 66, 78}. The mean of the given data is 73.5.
False; Moderate

31. The difference between the upper and lower quartile gives the interquartile range.
True; Easy

32. A small standard deviation suggests that the values are scattered widely around the mean.
False; Easy

Multiple Choice Questions

1. Which of the following refers to facts, measurements, and characteristics of something of interest?
   a. information
   b. variable
   c. data
   d. observation
   e. illustration
   c; Easy

2. What is information?
   a. The knowledge and insights acquired through the analysis of data
   b. A collection of data files
   c. The graphical representation of data
   d. A modification of existing data
   e. An inventory of existing, unorganized data
   a; Easy

3. Which of the following is true of data and information?
   a. Information is synonymous with data.
   b. Information is a generic term whereas data is related to the task undertaken.
   c. Information is acquired through the analysis of data.
   d. A GIS can function without data but information is indispensable.
   e. Data is constituted by information situated within analytical frameworks.
   c; Moderate

4. _____ data describes the geographical aspects of phenomena.
   a. Attribute
   b. Specific
   c. Generic
   d. Spatial
   e. Phenomenal
   d; Easy

5. A cartographer determines that the Salto Angel waterfall in Venezuela is located at 05°57'N latitude and 62°30'W longitude. This is an example of _____.
   a. attribute data
   b. spatial data
   c. metadata

d. generic data
e. geospatial metadata
b; Moderate

6. Which of the following statements is true of spatial data?
a. It describes the qualities of a phenomenon.
b. It describes the characteristics of an event that occurred in the distant past.
c. It is also known as attribute data.
d. It enables us to define the location of an object on the earth's surface.
e. It does not include street addresses, postal codes, and political boundaries.
d; Moderate

7. _____ data describes the qualities and characteristics of particular phenomena.
a. Public
b. Generic
c. Continuous
d. Private
e. Attribute
e; Easy

8. Which of the following constitute(s) attribute data?
a. The latitudinal and longitudinal coordinates of a point of interest
b. The postal code of a village in Scotland
c. The approximate distance of Miami from the Greenwich meridian
d. The color and texture of the giant redwood trees in California
e. The point of origin of the hurricane Irene
d; Moderate

9. Which of the following statements is true of spatial and attribute data?
a. Spatial data is geographic data whereas attribute data is nongeographic data.
b. A GIS functions only on geographic data and is independent of attribute data.
c. Attribute data is used by a GIS to find the coordinate location of a point of interest.
d. Unlike spatial data, attribute data can be tabulated in a DBMS.
e. Spatial data is sufficient to get complete information of a phenomenon.
a; Moderate

10. What is the first step in organizing a GIS?
a. Analyzing the spatial data of the point of interest
b. Differentiating between spatial and attribute data
c. Analyzing the attribute data of the point of interest
d. Integrating geographic data and spatial data
e. Analyzing spatial data to determine its sufficiency
b; Easy

11. A .txt extension denotes a(n) _____ file.
a. text-based spreadsheet
b. formatted text
c. simple text
d. spreadsheet
e. extensible markup language
c; Easy

12. A file named important.doc is a(n) _____.
    a. simple text file
    b. Microsoft Word document
    c. Adobe portable document
    d. zipped archive
    e. extensible markup language file
    b; Easy

13. Which of the following extensions is used to denote a compressed image file?
    a. .jpg
    b. .tif
    c. .pdf
    d. .zip
    e. .xml
    a; Easy

14. Kevin's class is working on a project that involves collecting crop-rotation data from farmers across the county. Which of the following file extensions must Kevin use to upload his project as a web page?
    a. .doc
    b. .zip
    c. .pdf
    d. .tif
    e. .html
    e; Moderate

15. The file finaldoc.zip is in the _____ format.
    a. tagged image
    b. extensible markup language
    c. compressed archive
    d. hypertext markup
    e. simple text
    c; Easy

16. Which of the following files is the Adobe portable document format?
    a. sun.doc
    b. sun.tif
    c. sun.xml
    d. sun.pdf
    e. sun.html
    d; Easy

17. Which of the following is true of .txt files?
    a. They are in the Adobe portable document format.
    b. They are used to create web pages.
    c. They are written in extensible markup language.
    d. They cannot contain special formatting.
    e. They are capable of supporting complex graphics with text.
    d; Moderate

18. When the first row of a flat file contains a descriptor, rather than data, it is called a(n) _____.
    a. attribute
    b. header
    c. field
    d. variable
    e. footer
    b; Easy

19. Most operating systems read simple text files with a program called _____.
    a. compiler
    b. text viewer
    c. Adobe reader
    d. pdf reader
    e. text editor
    e; Moderate

20. A text file that contains data structured in some fashion is called a _____.
    a. zipped archive
    b. compressed archive
    c. flat file
    d. portable file
    e. private data file
    c; Easy

21. Which of the following is true about a flat file?
    a. It contains unstructured data.
    b. It is denoted by an .xml extension.
    c. It consists of data depicted as images or complex graphics.
    d. It is organized in a tabular format or line by line.
    e. It is used to create web pages.
    d; Moderate

22. Which of the following is most commonly used as a delimiter in a flat file?
    a. semi-colon
    b. comma
    c. colon
    d. period
    e. hash
    b; Easy

23. Why is a delimiter used in a flat file?
    a. It highlights the most important term in a column.
    b. It separates two columns.
    c. It reduces the overall size of the file.
    d. It denotes the beginning of a footnote.
    e. It specifies the end of a line of data.
    b; Easy

24. In a flat file, rows are reserved for _____.
    a. records
    b. fields

c. variables
d. attributes
e. delimiters
a; Easy

25. In a flat file, columns are reserved for _____.
a. attributes
b. records
c. delimiters
d. comments
e. additional notes
a; Easy

26. A grouping of multiple files to collect, organize, and analyze data is called a _____.
a. compilation
b. program
c. database
d. compressed archive
e. portable document
c; Easy

27. Which of the following terms denotes a system that is used to update, edit, integrate, and share the data within a database?
a. DOS
b. GSM
c. DTMs
d. DBMS
e. DDOS
d; Easy

28. What is primary data?
a. The most significant set of numbers fed into a GIS
b. Data that is collected by a third party
c. Spatial or geographic data of a point of interest
d. Attributes of a point of interest
e. Data that is collected on a firsthand basis
e; Easy

29. Economic data that is collected and distributed by the government is an example of _____ data.
a. spatial
b. generic
c. primary
d. secondary
e. tertiary
d; Moderate

30. Which of the following is true about primary and secondary data?
a. Primary data refers to the data released by a country's government.
b. Secondary data refers to data collected directly by a researcher.
c. Secondary data provides better time and cost savings than primary data.

d. Primary data is more effective than secondary data.
e. A GIS functions using primary data alone; secondary data is discarded by the system.

c; Moderate

31. Data and information that describe data are known as _____.
    a. primary data
    b. secondary data
    c. tertiary data
    d. attribute data
    e. metadata

    e; Easy

32. _____ metadata is a special class of metadata that contains information about the geographical qualities of a data set.
    a. Attribute
    b. Geospatial
    c. Primary
    d. Secondary
    e. Geological

    b; Easy

33. In a table, the header row provides data about subsequent rows of data. This is an example of which of the following?
    a. Primary data
    b. Secondary data
    c. Metadata
    d. Attribute data
    e. Spatial data

    c; Moderate

34. Why are geospatial metadata used?
    a. To manage the data in the DBMS of a GIS
    b. To integrate spatial data with attribute data
    c. To organize the database of existing geographic data
    d. To eliminate the need to collect primary data
    e. To document geographic digital resources

    e; Easy

35. Which of the following refers to a common set of files used by many GIS software programs that contain both attribute and spatial data?
    a. Metadata
    b. Flat file
    c. Shapefile
    d. Compressed archive
    e. Spreadsheet

    c; Easy

36. Data that can be shared and distributed freely is known as _____ data.
    a. proprietary
    b. public
    c. private

d. meta
e. attribute
b; Easy

37. _____ data must be purchased and are subject to certain terms of use.
a. Public
b. Geospatial
c. Meta
d. Attribute
e. Proprietary
e; Easy

38. In the context of descriptive statistics, an individual observation of a variable is known as a(n) _____.
a. value
b. sample
c. population
d. array
e. sorted array
a; Easy

39. In the context of descriptive statistics, the universe of all possible values for a variable is known as a(n) _____.
a. variable
b. value
c. population
d. array
e. sorted array
c; Easy

40. In a geographic information system, an array is also called a _____.
a. field
b. value
c. sample
d. population
e. variable
a; Easy

41. What is a variable?
a. It a symbol used to represent any given value or set of values.
b. It is a subset of the population.
c. It is a sequence of observed measures.
d. It is an ordered, quantitative array.
e. It is an individual observation of a value.
a; Easy

42. What is a value?
a. It is an individual observation of a variable.
b. It is a subset of the population.
c. It is a sequence of observed measures.
d. It is an ordered, quantitative array.

e. It is a type of geodatabase that allows only single-user editing for unique feature datasets within a geodatabase.

a; Easy

43. What is a population?
    a. It is an individual observation of a variable.
    b. It is the universe of all possible values for a variable.
    c. It is a sequence of observed measures.
    d. It is an ordered, quantitative array.
    e. It is a narrow gap formed when the shared boundary of two polygons do not meet exactly.

    b; Easy

44. What is a sample?
    a. It is an individual observation of a variable.
    b. It is a subset of the population.
    c. It is the process of converting raster graphics to vector graphics.
    d. It is a geoprocessing operation that extracts those features from an input point, line, or polygon layer that falls within the spatial extent of a clip layer.
    e. It is a narrow gap formed when the shared boundary of two polygons do not meet exactly.

    b; Easy

45. What is a histogram?
    a. It is a graph showing the relationship between a variable and a constant in the form of a line chart.
    b. It is a graph showing the number of data values within each class range.
    c. It is a circular graph divided into sectors, illustrating proportion.
    d. It is a scatterplot that illustrates the relationship between two variables.
    e. It is a type of chart used to show conditions at which thermodynamically distinct phases occur at equilibrium.

    b; Easy

46. Adding all the values in an array and then dividing that sum by the number of observations gives the _____.
    a. mean
    b. mode
    c. median
    d. standard deviation
    e. variance

    a; Easy

47. If the sum of an array is 240 and there are 20 observations, the mean is _____.
    a. 24
    b. 20
    c. 10
    d. 12
    e. 42

    d; Easy

48. Array of ages of the students of a cooking class: {24, 23, 25, 23, 50, 40, 23, 70, 24, 45, 50, 23, 75, 60, 43, 32, 23, 30, 27, 30, 45, 62}. The mean of the given data is _____.
    a. 38.5
    b. 40
    c. 23
    d. 24.5
    e. 45
    a; Moderate

49. The _____ is the measure of central tendency that represents the most frequently occurring value in an array.
    a. mean
    b. mode
    c. median
    d. standard deviation
    e. variance
    b; Easy

50. In the context of central tendencies, which of the following is true of the mode?
    a. It is the most frequently occurring value in an array.
    b. It is the value in the middle of a sorted array when there are an odd number of observations.
    c. It is the mean of two central values of a sorted array containing an even number of observations.
    d. It is the sum of all the values in an array.
    e. It is the sum of all the values in an array, divided by the sum of the number of observations.
    a; Easy

51. Array of ages of the students of a cooking class: {24, 23, 25, 23, 50, 40, 23, 70, 24, 45, 50, 23, 75, 60, 43, 32, 23, 30, 27, 30, 45, 62}. The mode of the given data is _____.
    a. 38.5
    b. 40
    c. 23
    d. 24.5
    e. 45
    c; Moderate

52. Array of exam scores: {87, 76, 89, 90, 64, 67, 59, 79, 88, 74, 72, 99, 81, 77, 75, 86, 94, 66, 75, 74, 83,100, 92, 75, 73, 70, 60, 80, 85, 57}. The mode of the given data is _____.
    a. 75
    b. 76
    c. 77
    d. 76.5
    e. 78
    a; Moderate

53. The _____ is the value in the middle of a sorted array when there are an odd number of observations.
    a. mean
    b. mode

c. median
d. standard deviation
e. variance
c; Easy

54. Sorted array of exam scores: {57, 59, 60, 64, 66, 67, 70, 72, 73, 74, 74, 75, 75, 75, 76, 77, 79, 80, 81, 83, 85, 86, 87, 88, 89, 90, 92, 93, 94, 99}. The median of the given data is _____.
a. 75.5
b. 77
c. 77.5
d. 76.5
e. 78.5
d; Hard

55. Sorted array of exam scores: {57, 59, 60, 64, 66, 67, 70, 72, 73, 74, 74, 75, 75, 75, 76, 77, 79, 80, 81, 83, 85, 86, 87, 88, 89, 90, 92, 93, 94}. The median of the given data is _____.
a. 75.5
b. 77
c. 77.5
d. 76
e. 78
d; Hard

56. In an array, the range equals the:
a. least variable.
b. value in the middle of the sorted array when there are an odd number of observations.
c. mean of the two central values of the sorted array when there are an even number of observations.
d. sum of all the values.
e. largest value minus the smallest in the dataset.
e; Easy

57. Array of exam scores: {87, 76, 89, 90, 64, 67, 59, 79, 88, 74, 72, 99, 81, 77, 75, 86, 94, 66, 75, 74, 83,100, 92, 75, 73, 70, 60, 80, 85, 57}. Which of the following values is the range of the given data?
a. 43
b. 75
c. 76
d. 45
e. 75.5
a; Hard

58. The formula for variance is _____.
a. $s = \sqrt{\frac{n}{\sum\left(x-\bar{x}\right)^2}}$
b. $s^2 = \frac{n}{\sum\left(x-\bar{x}\right)^2}$

c. $s^2 = \frac{1}{n\left[\sum\left(x-\bar{x}\right)^2\right]}$

d. $s^2 = n\left[\sum\left(x-\bar{x}\right)^2\right]$

e. $s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n}$

e; Easy

59. Which of the following is the formula for standard deviation?

a. $s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n}}$

b. $s^2 = \frac{n}{\sum\left(x-\bar{x}\right)^2}$

c. $s = \sqrt{n\left[\sum\left(x-\bar{x}\right)^2\right]}$

d. $s = \sqrt{\frac{n}{\sum\left(x-\bar{x}\right)^2}}$

e. $s = \sqrt{\frac{1}{n\left[\sum\left(x-\bar{x}\right)^2\right]}}$

a; Easy

Short Answer Questions

1. How can information technology be of use in processing data for a GIS?
   Computers can automate repetitive tasks, store data efficiently in terms of space and cost, and provide a range of tools for analyzing data from spreadsheets to a GIS. They also help in collecting incredible amounts of data by satellites, grocery store product scanners, traffic sensors, temperature gauges, etc.
   Easy

2. Provide an example of spatial data and attribute data.
   The location of a house can be defined in many ways, such as with a street address, the street names of the nearest intersection, or the postal code. This is spatial data. Details such as the number of bedrooms, nature of the floor, air conditioning, and the like, constitute attribute data.
   Moderate

3. Why is it important to understand how spatial data and attribute data differ from and complement each other?

Because a GIS requires and integrates both spatial and attribute data, being able to differentiate between geographic and attribute data is the first step in organizing a GIS. Furthermore, being able to determine which kinds of data are required will ultimately aid in the implementation and use of a GIS.
Moderate

4. What is a file extension? How is it useful?
   The term file extension refers to the letters that follow the period (".") after the name of the file. For example, filename.txt is a text file. A file extension is used to distinguish one file from another.
   Easy

5. What is a flat file?
   A flat file is a text file that contains data organized or structured in some fashion. However, the file extension remains the same, i.e., .txt. Generally, flat files are organized in a tabular format or line by line. Each line or row of the file contains one and only one record.
   Easy

6. What is a delimiter?
   The character that is used to separate columns within a flat file is called the delimiter or separator. Though any character can be used as a delimiter, the most common delimiters are the tab, the comma, and a single space.
   Easy

7. What are databases? How is a DBMS relevant to the study of GIS?
   A grouping of multiple files to collect, organize, and analyze data is called a database. DBMS is a system that allows one to work, update, edit, integrate, share, and display the various data within the database. A GIS works with multiple files of various formats and a DBMS would help in the efficient management of data stored in these files.
   Easy

8. Explain the need for metadata.
   There is an incredibly vast and growing amount of data and information available to us, and most online searches return hundreds—if not thousands—of results. To overcome this data and information overload, it is necessary to obtain data about data. This special kind of data called metadata gives information and data about data.
   Moderate

9. Why is it a good idea to take an inventory of existing data before beginning the process of data collection?
   By taking an inventory of existing data before beginning the process of data collection, a lot of time and effort can be saved. Furthermore, identifying existing resources gives a better idea about the resources required. For instance, though we may already have census data (i.e., attribute data), we may need updated geographic data that contains the boundaries of US states or counties.
   Moderate

10. Distinguish between public and proprietary data.
    Data that is free to use is called public data. This data can be shared and distributed with no charge. Most government data is free. Proprietary data is data that must be purchased and is subject to certain terms of use.

Easy

11. What is a variable?
    A variable is a symbol used to represent any given value or set of values.
    Easy

12. What is a value?
    An individual observation of a variable is called a value.
    Easy

13. How many primary measures of central tendency are there? Name them.
    There are three primary measures of central tendency. They are mean, mode, and median.
    Easy

Fill in the Blanks

1. Once data is put into context, it becomes _____.
   information; Easy

2. _____ data describes the qualities and characteristics of a particular phenomenon.
   Attribute; Easy

3. _____ is the extension used for a compressed image file.
   .jpg; Easy

4. A _____ is a common set of files used by many GIS software programs that contain both spatial and attribute data.
   shapefile; Easy

5. A graph showing the number of data values within each class range is called a(n) _____.
   histogram; Easy

6. _____ are those that maintain values that are not symmetrical around the mean.
   Skewed data; Easy

Essay Questions

1. Distinguish between data and information with examples.
   Data refer to facts, measurements, characteristics, or traits of an object of interest. For example, the length of rainbow trout in a Colorado stream, the number of vegetarians in Alaska, and the diameter of mahogany tree trunks in the Brazilian rainforest are considered data. Once data are put into context, used to answer questions, situated within analytical frameworks, or used to obtain insights, they become information. The term information refers to the knowledge of value obtained through the collection, interpretation, and/or analysis of data.
   Moderate

2. How are geographic data different from attribute data?

Geographic or spatial data refer to geographic facts, measurements, or characteristics of an object that permit us to define its location on the surface of the earth. Such data include but are not restricted to the latitude and longitude coordinates of points of interest, street addresses, postal codes, political boundaries, and even the names of places of interest. Attribute data are concerned with nongeographic traits and characteristics of the point of interest.
Moderate

3. Explain the basic elements of a computer file. What is a flat file?
Files in a computer can contain any number of things from a complex set of instructions (e.g., a computer program) to a list of numbers and letters (e.g., address book). Furthermore, computer files come in all different sizes and types. One of the clues we can use to distinguish one file from another is the file extension. The file extension refers to the letters that follow the period (".") after the name of the file. For instance, a file mydocument.txt refers to a text file. Similarly, a .jpg extension denotes a compressed image file. A flat file is a text file that contains data organized or structured in some fashion, such as a table.
Easy

4. With regard to data sources, explain the terms primary data, secondary data, and metadata.
Primary data refer to data that are collected directly or on a firsthand basis. For example, recording the temperature during afternoons in May to determine the average temperature of the month would form a primary data set. Conversely, secondary data refer to data collected by someone else or some other party. For instance, using census or economic data collected and distributed by the government is secondary data. Metadata is data about data. For instance, a header row in a simple text file is analogous to metadata as it provides data about the subsequent rows of data.
Moderate

5. How is the format of data critically important to a GIS? Why is better to use familiar formats of data than exploring newer, potentially more efficient ones? Though many programs can read many formats of data, some data types can only be read by certain programs. Also, some programs require particular data formats. Understanding what data formats can or cannot be used will aid in the search for data. For instance, one of the most common forms of geographic information system (GIS) data is called the shapefile. Not all GIS programs can read or use shapefiles, but it may be necessary to convert to or from a shapefile or some other format. Hence, familiarity with a large number of data formats is beneficial for the search for data as it provides an understanding of not only what formats can be used but also what format conversions will need to be made, if necessary. **Moderate**

6. What are the key considerations that should be kept in mind when searching for data?
Prior to searching for data, ask yourself the following questions:
   - Why do I need the data?
   - At what time scale do I need the data?
   - At what geographic scale do I want the data?
   - What data already exist?
   - What format do I need the data?

   Moderate

7. Name the three primary measures of central tendency and give a brief description of each.
The three primary measures of central tendency are the mean, mode, and median.

The mean is the sum total of all the values in an array divided by the number of observations.
The mode is the measure of central tendency that represents the most frequently occurring value in the array.
The median is the observation that, when the array is ordered from lowest to highest, falls exactly in the center of the sorted array.
Easy