

Data Analytics for Accounting, 1e (Richardson)
Chapter 2 Data Preparation and Cleaning

- 1) Mastering the data requires a firm understanding of what data is available to you and where it is stored, as well as being skilled in the process of extracting, transforming, and loading (ETL).
- 2) A flat file is a means of storing data in one place, such as in an Excel spreadsheet, as opposed to storing the data in multiple tables, such as in a relational database.
- 3) A foreign key is an attribute that is required to exist in each table of a relational database and serves as the unique identifier for each record in a table.
- 4) A primary key is an attribute that is required to exist in each table of a relational database and serves as the unique identifier for each record in a table.
- 5) A foreign key is an attribute that exists in relational databases in order to carry out the relationship between two tables.
- 6) A composite primary key is made up of the three or more primary keys in the tables that it is linking.
- 7) Descriptive attributes are attributes that exist in relational databases that are neither primary nor foreign keys.
- 8) Once you have extracted the data of interest, it will need to be validated for completeness and existence.
- 9) The E in IMPACT Cycle represents existence.
- 10) The T in IMPACT Cycle represents transfer.
- 11) The L in IMPACT Cycle represents loading.
- 12) In order to obtain the right data, it is important to have a firm grasp of what data is available and how it is stored.
- 13) Data normalization can reduce data redundancy and improve data integrity.
- 14) Much like the IMPACT cycle, requesting data is often an iterative process.
- 15) Unlike the IMPACT cycle, requesting data is not an iterative process.
- 16) If the extraction and transformation steps have been done correctly, the loading part of the ETL process should be the simplest step.
- 17) After obtaining the data and determining the purpose and scope of the data request, the next

step is to validate the data.

18) Comparing the number of records that were extracted to the number of records in the source database is an example of validating the data for integrity.

19) Formatting negative numbers is an example of cleaning the data.

20) A template can make communication easier between data requestor and provider.

21) Mastering the data can also be described via the ETL process. The ETL process stands for:

- A) Extract, total, and load data.
- B) Extract, transform, and load data.
- C) Enter, transform, and load data.
- D) Enter, total, and load data.

22) All of the following are Audit Data Standards (ADS) developed by the American Institute of Certified Accountants *except*:

- A) Investments subledger standards
- B) General Ledger standards
- C) Procure-to-Pay subledger standards
- D) Order-to-Cash subledger standards

23) When using [EmployeeID] as the unique identifier of the Employee table, [EmployeeID] is an example of which of the following:

- A) Foreign key
- B) Composite key
- C) Primary key
- D) Key attribute

24) The purpose of extracting data is:

- A) To validate the data for completeness and integrity
- B) To load the data into the appropriate tool for analysis
- C) To identify and obtain the data from the appropriate source
- D) To identify which approach to data analytics should be used

25) The purpose of transforming data is:

- A) To validate the data for completeness and integrity
- B) To load the data into the appropriate tool for analysis
- C) To identify and obtain the data from the appropriate source
- D) To identify which approach to data analytics should be used

- 26) The purpose of loading data is:
- A) To validate the data for completeness and integrity
 - B) To put the data into the appropriate tool for analysis
 - C) To identify and obtain the data from the appropriate source
 - D) To identify which approach to data analytics should be used
- 27) All of the following are included in the five steps of the ETL process *except*:
- A) Determine the purpose and scope of the data request
 - B) Obtain the data
 - C) Validate the data for completeness and integrity
 - D) Scrub the data
- 28) Which of the following best exemplifies a way that data will need to be cleaned after extraction and validation?
- A) Remove headings and subtotals
 - B) Validate date/time fields
 - C) Remove trailing zeroes
 - D) Compare string limits for text fields
- 29) _____ is the metadata that describes each attribute in a database.
- A) Relational database
 - B) Data dictionary
 - C) Descriptive attributes
 - D) Flat file
- 30) Removing headings or subtotals from data is an example of which of the following?
- A) Validating the data for completeness
 - B) Validating the data for integrity
 - C) Cleaning the data
 - D) Obtaining the data
- 31) Correcting inconsistencies across data is an example of which of the following?
- A) Validating the data for completeness
 - B) Validating the data for integrity
 - C) Cleaning the data
 - D) Obtaining the data
- 32) Formatting negative numbers in the data is an example of which of the following?
- A) Validating the Data for Completeness
 - B) Validating the Data for Integrity
 - C) Cleaning the Data
 - D) Obtaining the Data

33) Removing leading zeroes and non-printable characters from the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

34) Comparing descriptive statistics for numeric fields within the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

35) Comparing the number of records within the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

36) Validating date/time fields within the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

37) Which of the following best describes the purpose of a primary key?

- A) To ensure that each row in the table is unique
- B) To create the relationship between two tables
- C) To provide business information, but are not required to build a database
- D) To support business processes across the organization

38) Which of the following best describes the purpose of a non-key attribute?

- A) To ensure that each row in the table is unique
- B) To create the relationship between two tables
- C) To provide business information
- D) To support business processes across the organization

39) Which of the following best describes the purpose of relational databases?

- A) To ensure that business rules are enforced
- B) To increase information redundancy in the organization
- C) To provide business information to data analysts
- D) To support business processes across the organization

- 40) Which of the following best describes the purpose of a foreign key?
- A) To ensure that each row in the table is unique
 - B) To create the relationship between two tables
 - C) To provide business information
 - D) To support business processes across the organization
- 41) A data dictionary is paramount in helping data analysts do which of the following?
- A) Maintain databases.
 - B) Identify the data they need to use.
 - C) Communicating insights.
 - D) Track outcomes.
- 42) A data dictionary is paramount in helping database administrators do which of the following?
- A) Maintain databases.
 - B) Identify the data they need to use.
 - C) Communicating insights.
 - D) Track outcomes.
- 43) At which step of the ETL process should you try to answer the question "What tools will be used to perform data analytic tests or procedures and why?"
- A) Step 1: Determine the purpose and scope of the data request.
 - B) Step 2: Obtain the data.
 - C) Step 3 or 4: Transformation.
 - D) Step 5: Loading the data for data analysis.
- 44) At which step of the ETL process should you try to answer the question "What other information will impact the nature, timing and extent of the data analysis?"
- A) Step 1: Determine the purpose and scope of the data request.
 - B) Step 2: Obtain the data.
 - C) Step 3 or 4: Transformation.
 - D) Step 5: Loading the data for data analysis.
- 45) At which step of the ETL process should you try to answer the question "What business problem will the data address?"
- A) Step 1: Determine the purpose and scope of the data request.
 - B) Step 2: Obtain the data.
 - C) Step 3 or 4: Transformation.
 - D) Step 5: Loading the data for data analysis.
- 46) At which step of the ETL process should you try to answer the question "Where are the data located in the financial or other related systems?"
- A) Step 1: Determine the purpose and scope of the data request.
 - B) Step 2: Obtain the data.
 - C) Step 3 or 4: Transformation.
 - D) Step 5: Loading the data for data analysis.

47) When obtaining the data yourself, you should do all of the following before you begin *except*:

- A) Identify the tables that contain the information you need.
- B) Identify which attributes specifically hold the information you need in each table.
- C) Identify how those tables are related to each other.
- D) Identify any errors or issues from the extraction.

48) There are a variety of methods that you could take to retrieve the data, including SQL. What does SQL stand for?

- A) Systems Query Language.
- B) Systems Question Language.
- C) Structured Question Language.
- D) Structured Query Language.

49) All of the following are benefits of using a normalized relational database *except*:

- A) Completeness.
- B) No redundancy.
- C) Business rules are enforced.
- D) Data is stored in one place.

50) Which of the following is most likely to be the primary key in an Employee table?

- A) Employee ID
- B) Employee Social Security Number
- C) Employee Name
- D) Employee Type

51) Chapter 2 describes, the various ways in which data can be stored for differing purposes. Describe the two ways data can be organized and the purpose for each organizational structure.

52) Taylor is a new staff accountant for a fortune 100 company. After hearing that she just successfully completed an Accounting Data Analytics course, her boss said, "Get me a listing of all our deadbeat customer, so I can cut off their credit." After asking clarifying questions, Taylor was able to determine that root request was "Which customer, with a credit limit over \$10,000, have more than \$5,000 outstanding for more than 90 days at the prior quarter's end?" Use the ETL Techniques briefly describe the process Taylor with have to complete to answer her boss' question. (Assume that Taylor does not have direct access the data, data will export into an Excel file, and she will complete the analysis with an Excel Pivot Table.)

53) Define and compare a primary key, a foreign key, and non-key attribute.

54) Assume that you will be up for a promotion next month and you'd like to analyze a recently acquired database to show off your data analytic skills. After identifying the goal of your data analysis, using the first step of the IMPACT cycle, what steps would you take if you have direct access to the database?

55) Assume that you have just completed the extraction process on a data set. As you begin validating the data for completeness and integrity, you notice an error. Describe the steps you might take to determine the source of an error.

Data Analytics for Accounting, 1e (Richardson)
Chapter 2 Data Preparation and Cleaning

1) Mastering the data requires a firm understanding of what data is available to you and where it is stored, as well as being skilled in the process of extracting, transforming, and loading (ETL).

Answer: TRUE

Difficulty: 1 Easy

Topic: How Data Are Used and Stored in The Accounting Cycle

Learning Objective: 02-01 Understand how data are organized in an accounting information system.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

2) A flat file is a means of storing data in one place, such as in an Excel spreadsheet, as opposed to storing the data in multiple tables, such as in a relational database.

Answer: TRUE

Difficulty: 1 Easy

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology

3) A foreign key is an attribute that is required to exist in each table of a relational database and serves as the unique identifier for each record in a table.

Answer: FALSE

Difficulty: 1 Easy

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology

4) A primary key is an attribute that is required to exist in each table of a relational database and serves as the unique identifier for each record in a table.

Answer: TRUE

Difficulty: 1 Easy

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology

5) A foreign key is an attribute that exists in relational databases in order to carry out the relationship between two tables.

Answer: TRUE

Difficulty: 1 Easy

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

6) A composite primary key is made up of the three or more primary keys in the tables that it is linking.

Answer: FALSE

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

7) Descriptive attributes are attributes that exist in relational databases that are neither primary nor foreign keys.

Answer: TRUE

Difficulty: 1 Easy

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

8) Once you have extracted the data of interest, it will need to be validated for completeness and existence.

Answer: FALSE

Difficulty: 2 Medium

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

9) The E in IMPACT Cycle represents existence.

Answer: FALSE

Difficulty: 1 Easy

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

10) The T in IMPACT Cycle represents transfer.

Answer: FALSE

Difficulty: 1 Easy

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

11) The L in IMPACT Cycle represents loading.

Answer: TRUE

Difficulty: 1 Easy

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

12) In order to obtain the right data, it is important to have a firm grasp of what data is available and how it is stored.

Answer: TRUE

Difficulty: 2 Medium

Topic: How Data Are Used and Stored in The Accounting Cycle

Learning Objective: 02-01 Understand how data are organized in an accounting information system.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

13) Data normalization can reduce data redundancy and improve data integrity.

Answer: TRUE

Difficulty: 1 Easy

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

14) Much like the IMPACT cycle, requesting data is often an iterative process.

Answer: TRUE

Difficulty: 2 Medium

Topic: The Data Analytics Process Using the Impact Cycle

Learning Objective: 02-04 Describe the Data Analytics Process Using the IMPACT Cycle.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

15) Unlike the IMPACT cycle, requesting data is not an iterative process.

Answer: FALSE

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

16) If the extraction and transformation steps have been done correctly, the loading part of the ETL process should be the simplest step.

Answer: TRUE

Difficulty: 2 Medium

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

17) After obtaining the data and determining the purpose and scope of the data request, the next step is to validate the data.

Answer: TRUE

Difficulty: 1 Easy

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

18) Comparing the number of records that were extracted to the number of records in the source database is an example of validating the data for integrity.

Answer: FALSE

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

19) Formatting negative numbers is an example of cleaning the data.

Answer: TRUE

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

20) A template can make communication easier between data requestor and provider.

Answer: TRUE

Difficulty: 1 Easy

Topic: Obtain the Data

Learning Objective: 02-01 Understand how data are organized in an accounting information system.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; BB Industry

21) Mastering the data can also be described via the ETL process. The ETL process stands for:

A) Extract, total, and load data.

B) Extract, transform, and load data.

C) Enter, transform, and load data.

D) Enter, total, and load data.

Answer: B

Difficulty: 1 Easy

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

22) All of the following are Audit Data Standards (ADS) developed by the American Institute of Certified Accountants *except*:

A) Investments subledger standards

B) General Ledger standards

C) Procure-to-Pay subledger standards

D) Order-to-Cash subledger standards

Answer: A

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

23) When using [EmployeeID] as the unique identifier of the Employee table, [EmployeeID] is an example of which of the following:

- A) Foreign key
- B) Composite key
- C) Primary key
- D) Key attribute

Answer: C

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

24) The purpose of extracting data is:

- A) To validate the data for completeness and integrity
- B) To load the data into the appropriate tool for analysis
- C) To identify and obtain the data from the appropriate source
- D) To identify which approach to data analytics should be used

Answer: C

Difficulty: 1 Easy

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

25) The purpose of transforming data is:

- A) To validate the data for completeness and integrity
- B) To load the data into the appropriate tool for analysis
- C) To identify and obtain the data from the appropriate source
- D) To identify which approach to data analytics should be used

Answer: A

Difficulty: 1 Easy

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

26) The purpose of loading data is:

- A) To validate the data for completeness and integrity
- B) To load the data into the appropriate tool for analysis
- C) To identify and obtain the data from the appropriate source
- D) To identify which approach to data analytics should be used

Answer: B

Difficulty: 1 Easy

Topic: Loading

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

27) All of the following are included in the five steps of the ETL process *except*:

- A) Determine the purpose and scope of the data request
- B) Obtain the data
- C) Validate the data for completeness and integrity
- D) Scrub the data

Answer: D

Difficulty: 1 Easy

Topic: How Data Are Used and Stored in The Accounting Cycle

Learning Objective: 02-01 Understand how data are organized in an accounting information system.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

28) Which of the following best exemplifies a way that data will need to be cleaned after extraction and validation?

- A) Remove headings and subtotals
- B) Validate date/time fields
- C) Remove trailing zeroes
- D) Compare string limits for text fields

Answer: A

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

29) _____ is the metadata that describes each attribute in a database.

- A) Relational database
- B) Data dictionary
- C) Descriptive attributes
- D) Flat file

Answer: B

Difficulty: 2 Medium

Topic: Data Dictionaries

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

30) Removing headings or subtotals from data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

Answer: C

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

31) Correcting inconsistencies across data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

Answer: C

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

32) Formatting negative numbers in the data is an example of which of the following?

- A) Validating the Data for Completeness
- B) Validating the Data for Integrity
- C) Cleaning the Data
- D) Obtaining the Data

Answer: C

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

33) Removing leading zeroes and non-printable characters from the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

Answer: C

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

34) Comparing descriptive statistics for numeric fields within the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

Answer: A

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

35) Comparing the number of records within the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

Answer: A

Difficulty: 2 Medium

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

36) Validating date/time fields within the data is an example of which of the following?

- A) Validating the data for completeness
- B) Validating the data for integrity
- C) Cleaning the data
- D) Obtaining the data

Answer: B

Difficulty: 1 Easy

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

37) Which of the following best describes the purpose of a primary key?

- A) To ensure that each row in the table is unique
- B) To create the relationship between two tables
- C) To provide business information, but are not required to build a database
- D) To support business processes across the organization

Answer: A

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

38) Which of the following best describes the purpose of a non-key attribute?

- A) To ensure that each row in the table is unique
- B) To create the relationship between two tables
- C) To provide business information
- D) To support business processes across the organization

Answer: C

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

39) Which of the following best describes the purpose of relational databases?

- A) To ensure that business rules are enforced
- B) To increase information redundancy in the organization
- C) To provide business information to data analysts
- D) To support business processes across the organization

Answer: D

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

40) Which of the following best describes the purpose of a foreign key?

- A) To ensure that each row in the table is unique
- B) To create the relationship between two tables
- C) To provide business information
- D) To support business processes across the organization

Answer: B

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

41) A data dictionary is paramount in helping data analysts do which of the following?

- A) Maintain databases.
- B) Identify the data they need to use.
- C) Communicating insights.
- D) Track outcomes.

Answer: B

Difficulty: 2 Medium

Topic: Data Dictionaries

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

42) A data dictionary is paramount in helping database administrators do which of the following?

- A) Maintain databases.
- B) Identify the data they need to use.
- C) Communicating insights.
- D) Track outcomes.

Answer: A

Difficulty: 2 Medium

Topic: Data Dictionaries

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

43) At which step of the ETL process should you try to answer the question "What tools will be used to perform data analytic tests or procedures and why?"

- A) Step 1: Determine the purpose and scope of the data request.
- B) Step 2: Obtain the data.
- C) Step 3 or 4: Transformation.
- D) Step 5: Loading the data for data analysis.

Answer: B

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

44) At which step of the ETL process should you try to answer the question "What other information will impact the nature, timing and extent of the data analysis?"

- A) Step 1: Determine the purpose and scope of the data request.
- B) Step 2: Obtain the data.
- C) Step 3 or 4: Transformation.
- D) Step 5: Loading the data for data analysis.

Answer: A

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

45) At which step of the ETL process should you try to answer the question "What business problem will the data address?"

- A) Step 1: Determine the purpose and scope of the data request.
- B) Step 2: Obtain the data.
- C) Step 3 or 4: Transformation.
- D) Step 5: Loading the data for data analysis.

Answer: A

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

46) At which step of the ETL process should you try to answer the question "Where are the data located in the financial or other related systems?"

- A) Step 1: Determine the purpose and scope of the data request.
- B) Step 2: Obtain the data.
- C) Step 3 or 4: Transformation.
- D) Step 5: Loading the data for data analysis.

Answer: B

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

47) When obtaining the data yourself, you should do all of the following before you begin *except*:

- A) Identify the tables that contain the information you need.
- B) Identify which attributes specifically hold the information you need in each table.
- C) Identify how those tables are related to each other.
- D) Identify any errors or issues from the extraction.

Answer: D

Difficulty: 2 Medium

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Understand

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

48) There are a variety of methods that you could take to retrieve the data, including SQL. What does SQL stand for?

- A) Systems Query Language.
- B) Systems Question Language.
- C) Structured Question Language.
- D) Structured Query Language.

Answer: D

Difficulty: 1 Easy

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

49) All of the following are benefits of using a normalized relational database *except*:

- A) Completeness.
- B) No redundancy.
- C) Business rules are enforced.
- D) Data is stored in one place.

Answer: D

Difficulty: 2 Medium

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Remember

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

50) Which of the following is most likely to be the primary key in an Employee table?

- A) Employee ID
- B) Employee Social Security Number
- C) Employee Name
- D) Employee Type

Answer: A

Difficulty: 2 Medium

Topic: Relational Database Essay and Computational Questions

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Apply

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making

51) Chapter 2 describes, the various ways in which data can be stored for differing purposes. Describe the two ways data can be organized and the purpose for each organizational structure.

Answer: Answers may vary slightly!

A **relational database**, is the type of database you are most likely to come across when extracting and using accounting and financial data. While it is often preferred to analyze data from a **flat file** (e.g., in an Excel spreadsheet, in which all the data are stored in one place), when it comes to storing data and maintaining data integrity, a relational database is preferred because of its ability to maintain "one version of the truth" across multiple data elements.

Difficulty: 3 Hard

Topic: How Data Are Used and Stored in The Accounting Cycle

Learning Objective: 02-01 Understand how data are organized in an accounting information system.

Bloom's: Evaluate

AACSB: Reflective Thinking; Analytical Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; BB Industry

52) Taylor is a new staff accountant for a fortune 100 company. After hearing that she just successfully completed an Accounting Data Analytics course, her boss said, "Get me a listing of all our deadbeat customer, so I can cut off their credit." After asking clarifying questions, Taylor was able to determine that root request was "Which customer, with a credit limit over \$10,000, have more than \$5,000 outstanding for more than 90 days at the prior quarter's end?" Use the ETL Techniques briefly describe the process Taylor will have to complete to answer her boss' question. (Assume that Taylor does not have direct access the data, data will export into an Excel file, and she will complete the analysis with an Excel Pivot Table.)

Answer: Answers will vary but should include some of these items.

- **EXTRACTION:** Taylor completed step 1 by asking her boss clarifying question to determine purpose and scope of the data request. To complete step 2 she should complete a template to request the data from the system administrator.
- **TRANSFORMATION:** For step 3, Taylor will need to validate the data for completeness and integrity. She can compare the total AR balance of the data extracted to the gross AR amount on the financial statement from the prior quarter's end. For step 4, Taylor should clean the data by removing headings or subtotals, ensuring formatting is consistent, etc.
- **LOADING:** No additional loading is necessary as the analysis will be run in Excel.

Difficulty: 3 Hard

Topic: Extraction, Transformation, and Loading (ETL) of Data

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Apply; Analyze

AACSB: Reflective Thinking; Knowledge Application

AICPA: BB Industry; FN Decision Making; BB Leveraging Technology; FN Leveraging Technology

53) Define and compare a primary key, a foreign key, and non-key attribute.

Answer: Answers will vary!

- The purpose of the primary key is to ensure that each row in the table is unique, so it is often referred to as a "unique identifier."
 - a. Each table must have a primary key. The primary key is typically made up of one column, but it can occasionally be made up of a combination of columns. It is rarely truly descriptive; instead, a collection of letters or simply sequential numbers are often used.
- The purpose of the foreign key is to create the relationship between two tables. The relationship is created by placing a foreign key in one of the two tables that are related.
 - a. Whenever two tables are related, one of those tables must contain a foreign key to create the relationship. The foreign key is special type of attribute as it must be the primary key in a related table.
- The other columns in a table are descriptive attributes.
 - a. Primary and foreign keys facilitate the structure of a relational database, and the descriptive attributes provide actual business information.

Difficulty: 3 Hard

Topic: Relational Database

Learning Objective: 02-02 Understand how data are stored in a relational database.

Bloom's: Analyze

AACSB: Reflective Thinking

AICPA: BB Industry; BB Leveraging Technology; FN Decision Making

54) Assume that you will be up for a promotion next month and you'd like to analyze a recently acquired database to show off your data analytic skills. After identifying the goal of your data analysis, using the first step of the IMPACT cycle, what steps would you take if you have direct access to the database?

Answer: Answers may vary!

1. Identify the tables that contain the information you need. You can do this by looking through the data dictionary or the relationship model.
2. Identify which attributes, specifically, hold the information you need in each table.
3. Identify how those tables are related to each other.

Difficulty: 3 Hard

Topic: Extraction

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Analyze

AACSB: Reflective Thinking

AICPA: BB Industry; BB Leveraging Technology; FN Decision Making

55) Assume that you have just completed the extraction process on a data set. As you begin validating the data for completeness and integrity, you notice an error. Describe the steps you might take to determine the source of an error.

Answer: Answers will vary! A potential answer might include:

If an error is found, depending on the size of the dataset, you may be able to easily find the missing or erroneous data by scanning it with your eyes. However, if the dataset is large, or if the error is difficult to find, it may be easiest to go back to the extraction and examine how the data was extracted, fix any errors in the SQL code, and re-run the extraction.

Difficulty: 3 Hard

Topic: Transformation

Learning Objective: 02-03 Explain and apply extraction, transformation, and loading (ETL) techniques.

Bloom's: Apply

AACSB: Reflective Thinking

AICPA: BB Leveraging Technology; FN Leveraging Technology; FN Decision Making