

CH 3 DATA LITERACY

Multiple Choice Questions:

1. Which of the following describes the process of transforming raw data into useful information?

- (a) Data encryption
- (c) Data processing
- (b) Data aggregation
- (d) Data storage

2. In the context of the DIKW Model, which stage involves deriving actionable insights from processed

information?

- (a) Wisdom
- (c) Information
- (b) Knowledge (d) Data

3. Which aspect of data security involves confirming the identity of users trying to access data?

- (a) Data encryption
- (c) Authentication
- (b) Access controls
- (d) Datamasking

4. What is one advantage of using a data dashboard for tracking progress?

- (a) It ensures data accuracy
- (c) It requires manual updates
- (b) It saves time by consolidating information (d) It increases data complexity

5. Which type of data is used to specify the taste of sugar?

- (a) Qualitative
- (b) discrete

Ans. 1. (c), 2. (b), 3. (c), 4. (b), 5. (a).

Fill in the blanks

1. One key aspect of data security is _____, which transforms data into a coded format to prevent unauthorized access.

2. The DIKW Model represents the relationship between data, information, _____ and wisdom
3. Excel is primarily known as spreadsheet software, but it is also a powerful tool for _____ data.
4. _____ data can be subdivided into smaller parts according to the measurement precision.
5. A data dashboard helps you see and understand key information quickly by using visuals like _____ and _____

Ans. 1. Encryption, 2. Knowledge, 3. Visualizing, 4. Continuous, 5. Charts, graphs.

State Whether the given statement are True or False

1. Excel is primarily a data visualization tool designed for creating interactive dashboards.
2. A bar chart can be used to show sales trends over a period of time.
3. Discrete data can take any value in a range
4. Data presentation should focus on accuracy and relevance to make the information actionable.
5. Kaggle assigns a usability score to the data sets that are present on the website based on scores given by the users of that data

Ans. 1.False, 2. False, 3. False. 4. True 5. True

Match the following

| | |
|--------------------------|---|
| 1.Data Discovery | [a] creating new data artificially ,often using algorithms or models, to simulate real-world scenarios or to fill gaps in existing datasets |
| 2. Data Augmentation | [b] Data collected first-hand for a specific purpose by the researcher or organization |
| 3. Data Generation | [c] Data collected by someone else for a different purpose but used by another party. |
| 4. Primary Data Source | [d] Process of enhancing existing data by adding new information or modifying it to improve its quality and usefulness. |
| 5. Secondary Data Source | [e] Involves finding and identifying relevant data from various sources to use for analysis. |

Ans. 1.[e], 2.[d], 3.[a], 4.[b], 5.[c].

1. Assertion (A): Data visualization is the representation of data or information in a graph, chart, or other Reason(R): Data visualization does not help to see and understand trends and outliers

(a) Both A and R are true and R is the correct explanation of A

b) Both A and R are true but R is not the correct explanation of A

(c) A is true but R is false.

(d) A is false but R is true.

2. Assertion (A): Data is a powerless tool.

Reason (R): Data is present everywhere, it should be properly reviewed and analyzed accurately.

(a) Both A and R are true and R is the correct explanation of A.

(b) Both A and R are true but R is not the correct explanation of A.

(c) A is true but R is false.

(d) A is false but R is true.

Ans 1 c

2. D

Very Short Answer Type Questions:

1. How can data literacy impact decision-making?

Ans. It helps individuals make informed decisions by understanding and analyzing data effectively.

2. What is a key purpose of using data dashboards?

Ans. To display important information in one place.

3. What are three different ways of data representation?

Ans. Textual, tabular, Graphical

4. Give five examples of Continuous data around you.

Ans. Weight, Height, Time, Wind Speed, Temperature.

5. Which type of data best represents the time displayed on a clock? Explain.

Ans. It is Continuous Data. The continuous data can be broken down into decimal and fractions, so it can be subdivided into smaller parts according to the measurement precision.

Short Answer Type Questions

1. How does data security differ from data privacy?

Ans. Data security focuses on protecting data from unauthorized access, while data privacy concerns the rights and control individuals have over their personal information.

2. What does Tableau Public offer that Tableau Desktop does not?

Ans. Tableau Public allows users to create and share visualizations publicly on the web, whereas Tableau Desktop offers advanced features and private data storage.

3. Why is visual appeal important in data presentation?

Ans. Visual appeal is important because it makes the data engaging and easier to understand, helping the audience grasp key points quickly.

4. Explain Data Processing with example.

Ans. Data processing involves the methods and techniques used to collect, clean, transform, and organize raw data into a structured format that can be analyzed. For Example: If you have a list of customer survey responses, data processing might involve cleaning up the responses by correcting typos, removing incomplete entries and converting the data into a consistent form for analysis.

5. What is Web Scraping?

Ans. Web scraping is the automated process of extracting data from websites, It involves using software tools known as web scrapers, to gather specific information from web pages. This data can then be used for various purposes such as analysis, research, or integration into other applications.

Long Answer Type Questions

1. What are the best practices for acquiring data?

Ans. The best practices for acquiring data involve

(a) Clearly stating the purpose of data collection.

(b) Selecting data from reliable and authentic primary and secondary sources.

(c) Verifying data for accuracy, completeness, consistency, and reliability.

(d) Choosing appropriate data collection methods and implementing standardized process. (e) Obtaining necessary permissions, ensuring privacy, and complying with regulations. (f) Ensuring secure and organized data storage with backup plans.

(g) Performing data cleaning to maintain data quality.

2. What are Data Features? Explain the two primary categories of features in data analysis and machine learning.

Ans. Data features are individual measurable properties or characteristics of a dataset that help in describing and analyzing the data. Each feature represents a variable that can be used for various types of analysis, including statistical analysis, machine learning, and data visualization. For Example To analyze how customer satisfaction trends over time, we use features like rating and review date.

In data analysis and machine learning, features (or variables) are classified into dependent and independent features based on their roles in a model.

Independent Features or predictors are the input variables used to predict or explain changes in the dependent feature. They are the variables you use to build your model.

Dependent Feature or target is the output variable that you want to predict or explain. It depends on the independent features.

3. What are different data collection methods for Quantitative Data Interpretation?

Polls: Surveys with a single question to quickly gather numerical responses. For Example: Asking people to vote on their favorite movie from a list of options.

Observations: Recording numerical data based on observed behavior or events. For Example: Counting the number of people visiting a website over a specific period.

Lists: Compiling rankings or counts for specific categories. For Example: Top 5 Trending Athletes: A list of the most popular athletes based on social media mentions.

Counter: Recording the number of occurrences of a particular event. For Example: Counting the number of students visiting a website.

Interviews: Structured interviews with predefined questions to gather numerical data. For Example: Conducting interviews to collect data on customer satisfaction scores.