

Data Science For All 1st edition Davis Test Bank

SKU: 9780135311400-TEST-BANK

Data Science for All, 1e (Glanz/Davis)
Chapter 1 What is Data Science?

1.1 Introduction to Data Science

1 Define terms relating to data science.

- 1) Which of the following best describes the interdisciplinary nature of data science?
- A) Data science combines statistics, computer science, and domain expertise to extract knowledge from data.
 - B) Data science is solely focused on statistical analysis.
 - C) Data science is primarily about data visualization.
 - D) Data science only involves machine learning.

Answer: A

- 2) What are the three iterative components of the data science life cycle?
- A) Data preparation, analysis, and storytelling
 - B) Data collection, visualization, and deletion
 - C) Data extraction, transformation, and archiving
 - D) Data acquisition, modeling, and disposal

Answer: A

2 Identify the scope of data science including its interdisciplinary nature or life cycle.

- 1) Which term is defined as the process of collecting, cleaning, transforming, integrating, and managing data for effective analysis and storytelling?
- A) Data preparation
 - B) Data visualization
 - C) Data modeling
 - D) Data archiving

Answer: A

- 2) In the context of data science, what does "data storytelling" refer to?
- A) Communicating data insights through summaries, visualizations, and narratives.
 - B) Collecting and cleaning raw data.
 - C) Analyzing data to find patterns.
 - D) Storing data in a structured format.

Answer: A

3 Answer questions relating to the interdisciplinary nature or life cycle of data science.

- 1) Why is data science considered interdisciplinary?
 - A) It combines statistics, computer science, and domain expertise.
 - B) It focuses solely on statistical analysis.
 - C) It is limited to computer science applications.
 - D) It only involves data visualization.

Answer: A

- 2) What process involves cleaning, transforming, and integrating data to make them useful for analysis and storytelling?

- A) Data wrangling
- B) Data visualization
- C) Data modeling
- D) Data archiving

Answer: A

1.2 Data in Tables

1 Answer questions about the relationship of observations and variables.

- 1) What is a Boolean variable type in data tables?
 - A) A variable that contains only two possible values, TRUE or FALSE
 - B) A variable that takes on only numerical values
 - C) A variable that can take on any text
 - D) A variable that describes a category

Answer: A

- 2) What is the relationship between an observational unit and a variable?
 - A) An observational unit is an entity about which data are recorded, and a variable is a recorded characteristic of that entity.
 - B) A variable is an entity about which data are recorded, and an observational unit is a recorded characteristic of that entity.
 - C) An observational unit is a subset of variables in a dataset.
 - D) A variable is always a numerical value associated with an observational unit.

Answer: A

2 Answer questions about tidying messy data.

- 1) Considering the numbers 0, -2, 3.5, and 10, which one could be converted to a Boolean value in a data analysis context?
 - A) 0
 - B) -2
 - C) 3.5
 - D) 10

Answer: A

- 2) Which of the following scenarios best illustrates tidy data?
- A) A table where each column is a variable, each row is an observation, and each cell contains a single value
 - B) A table where each column is an observation, each row is a variable, and each cell contains multiple values
 - C) A table where each cell contains a summary of multiple observations
 - D) A table where each row is a dataset, and each column is a variable

Answer: A

3 Identify different variables value and data types.

- 1) Looking at the numbers -3, 1.4, 1.7, and 1.9, which is most likely to be stored as an integer in a data structure?

- A) -3
- B) 1.4
- C) 1.7
- D) 1.9

Answer: A

- 2) Consider the values -2, 1, 3, and 4.0. Which of these could be stored as an integer or converted to a Boolean variable?

- A) 1
- B) -2
- C) 3
- D) 4.0

Answer: A

4 Describe metadata and its purpose.

- 1) Which of the following is NOT a characteristic that would be included in metadata?

- A) Description of the analysis process
- B) Variable names
- C) Data set authorship
- D) The coding scheme used for variables

Answer: A

- 2) Often, data are accompanied by a description of the data itself. Because it is data about data, what is this data description called?

- A) metadata
- B) keydata
- C) index data
- D) exclusive data

Answer: A

1.3 Data Preparation

1 Define or describe the data wrangling process.

1) Which is not a part of the data preparation process?

- A) Data storytelling
- B) Data collection
- C) Data wrangling
- D) Data management

Answer: A

2) Which of the following activities is NOT part of the data wrangling process?

- A) Developing machine learning models
- B) Cleaning incomplete or inconsistent data
- C) Transforming data into a structured format
- D) Integrating multiple data sources

Answer: A

2 Identify and use the steps for data wrangling.

1) Which of the following is an example of raw data?

- A) Data in the form originally collected
- B) A processed spreadsheet ready for analysis
- C) Data that has been cleaned and structured
- D) A visual representation of analyzed data

Answer: A

2) What step in the data wrangling process involves merging multiple data sources?

- A) Integrating data
- B) Cleaning data
- C) Transforming data
- D) Visualizing data

Answer: A

1.4 Data Analysis and Storytelling

1 Describe or define the different types of data analysis.

1) Good data communication is necessary for what purpose?

- A) Making information accessible to a broader audience
- B) The transformation of raw data
- C) The integration of different data sets
- D) Collecting high-quality data

Answer: A

2) Which of the following is NOT a category of data analysis?

- A) Restrictive
- B) Descriptive
- C) Diagnostic
- D) Prescriptive

Answer: A

2 Describe or define the different types of data storytelling.

1) Data storytelling is essential for what purpose?

- A) Communicating complex data findings in an understandable way
- B) Cleaning and preparing raw data
- C) Storing and managing large data sets
- D) Collecting new data from various sources

Answer: A

2) What does visualization refer to, in the context of data science?

- A) The presentation of data in a graphical format
- B) The prediction of future data trends
- C) The descriptive analysis of data
- D) The storytelling aspect of data analysis

Answer: A

1.5 Data Science in Society and Industry

1 Answer questions relating to data science in society and industry.

1) How does data science contribute to the healthcare industry?

- A) By enabling personalized medicine through the analysis of patient data
- B) By automating all medical procedures
- C) By eliminating the need for healthcare professionals
- D) By making all healthcare data public

Answer: A

2) In what way has data science impacted the retail industry?

- A) By optimizing inventory management through predictive analytics
- B) By completely eliminating physical stores
- C) By reducing the need for customer service representatives
- D) By making all transactions paper-based

Answer: A

2 Answer questions about the importance of data science.

1) The text states that data science values include:

- A) Curiosity, creativity, strategy, and insight
- B) Mathematics, programming, and data storage
- C) Speed, efficiency, and automation
- D) Confidentiality, security, and privacy

Answer: A

2) How does data science contribute to innovation in technology?

- A) By enabling the development of new products and services through data-driven insights
- B) By eliminating the need for research and development
- C) By making all technology self-sustaining
- D) By focusing only on hardware improvements

Answer: A