

# FBLA Data Science & AI Practice Test (Sample)

## Study Guide



**Everything you need from our exam experts!**

**Copyright © 2026 by Examzify - A Kaluba Technologies Inc. product.**

**ALL RIGHTS RESERVED.**

**No part of this book may be reproduced or transferred in any form or by any means, graphic, electronic, or mechanical, including photocopying, recording, web distribution, taping, or by any information storage retrieval system, without the written permission of the author.**

**Notice: Examzify makes every reasonable effort to obtain accurate, complete, and timely information about this product from reliable sources.**

**SAMPLE**

# Table of Contents

<b>Copyright</b> .....	<b>1</b>
<b>Table of Contents</b> .....	<b>2</b>
<b>Introduction</b> .....	<b>3</b>
<b>How to Use This Guide</b> .....	<b>4</b>
<b>Questions</b> .....	<b>5</b>
<b>Answers</b> .....	<b>8</b>
<b>Explanations</b> .....	<b>10</b>
<b>Next Steps</b> .....	<b>16</b>

SAMPLE

# Introduction

Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.

This guide was designed to help you do exactly that.

Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.

The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:

- Practice answering questions under realistic conditions,
- Improve accuracy and speed,
- Review explanations to strengthen weak areas, and
- Approach the exam with greater confidence.

We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.

**Remember:** successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.

Let's begin.

# How to Use This Guide

**This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here's how to get the most out of your Examzify study guide:**

## **1. Start with a Diagnostic Review**

**Skim through the questions to get a sense of what you know and what you need to focus on. Your goal is to identify knowledge gaps early.**

## **2. Study in Short, Focused Sessions**

**Break your study time into manageable blocks (e.g. 30 - 45 minutes). Review a handful of questions, reflect on the explanations.**

## **3. Learn from the Explanations**

**After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.**

## **4. Track Your Progress**

**Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.**

## **5. Simulate the Real Exam**

**Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.**

## **6. Repeat and Review**

**Don't just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning. Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.**

**There's no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly, adapt the tips above to fit your pace and learning style. You've got this!**

## Questions

SAMPLE

- 1. What is defined as the value that appears most frequently in a dataset?**
  - A. Mean**
  - B. Mode**
  - C. Median**
  - D. Variance**
  
- 2. What data type is suitable for data that has a fixed format and can be easily organized into tables?**
  - A. Unstructured Data**
  - B. Categorical Data**
  - C. Numeric Data**
  - D. Structured Data**
  
- 3. What term refers to AI systems that are designed to monitor individuals and raise privacy concerns?**
  - A. Logic-Based Reasoning**
  - B. AI Surveillance**
  - C. Probability-Based Reasoning**
  - D. LLM Credibility Issues**
  
- 4. When is the median used instead of the mean?**
  - A. With normally distributed data**
  - B. With skewed data distributions**
  - C. When calculating mode**
  - D. With categorical data**
  
- 5. Which method is commonly used to address missing data?**
  - A. Normalization**
  - B. Imputation**
  - C. Segmentation**
  - D. Encoding**

- 6. What method involves solving problems using established rules and logic?**
- A. Symbolic Reasoning**
  - B. Inference**
  - C. Ethical AI**
  - D. Algorithmic Bias**
- 7. Which architecture is typically associated with deep learning models?**
- A. Decision Trees**
  - B. Support Vector Machines**
  - C. Neural Networks**
  - D. Linear Regression**
- 8. What term refers to all possible outcomes in a scenario?**
- A. Event**
  - B. Sample Space**
  - C. Experiment**
  - D. Outcome Set**
- 9. Which of the following is NOT a role of data analysts?**
- A. Building predictive algorithms**
  - B. Creating detailed reports**
  - C. Performing data storage and maintenance**
  - D. Conducting qualitative research**
- 10. What is termed as a variable with countable values, such as the number of students?**
- A. Continuous Variable**
  - B. Discrete Variable**
  - C. Nominal Variable**
  - D. Interval Variable**

## Answers

SAMPLE

1. B
2. D
3. B
4. B
5. B
6. A
7. C
8. B
9. D
10. B

SAMPLE

## **Explanations**

SAMPLE

**1. What is defined as the value that appears most frequently in a dataset?**

- A. Mean**
- B. Mode**
- C. Median**
- D. Variance**

The value that appears most frequently in a dataset is known as the mode. In statistical terms, the mode is a measure of central tendency, much like the mean and median, but it specifically focuses on frequency rather than the sum or middle value of the data points. This makes the mode particularly useful for categorical data where we want to identify the most common category or value present. Understanding the mode is important because it can provide insights into the data distribution, especially when analyzing data sets with multiple modes (bi-modal or multi-modal distributions), which indicate that there might be different groups or categories present within the data. Unlike the mean, which can be affected by extreme values, the mode is robust to outliers, making it a valuable measure in many practical data analysis scenarios.

**2. What data type is suitable for data that has a fixed format and can be easily organized into tables?**

- A. Unstructured Data**
- B. Categorical Data**
- C. Numeric Data**
- D. Structured Data**

Structured data refers to data that is organized in a predefined manner, typically in tables with rows and columns. This type of data is easily searchable and can be efficiently processed by algorithms and databases due to its clear format. Examples of structured data include data stored in relational databases, where each field has a defined type (such as integers, dates, or strings) and the relationships between different pieces of data are well-defined. In contrast, unstructured data lacks such a format and does not fit neatly into tables, making it more challenging to analyze. Categorical data consists of labels or categories that can be used to organize data but may not have a fixed structure. Numeric data refers specifically to data that is quantitative, which can be structured but does not encompass the broader definition necessary for the question. Thus, structured data is the most appropriate answer because of its fixed format and compatibility with tabular organization.

**3. What term refers to AI systems that are designed to monitor individuals and raise privacy concerns?**

- A. Logic-Based Reasoning**
- B. AI Surveillance**
- C. Probability-Based Reasoning**
- D. LLM Credibility Issues**

AI Surveillance refers to artificial intelligence systems that are developed specifically to monitor individuals, often through the analysis of data, visual recognition, and behavior tracking. This can include technologies such as facial recognition, tracking online activities, and analyzing data from various sources to create profiles of individuals or groups. The use of such systems raises significant privacy concerns because they can infringe on individuals' rights to privacy and potentially lead to misuse of data, discrimination, and unauthorized surveillance. The context surrounding AI surveillance has sparked debates on ethical guidelines, regulations, and the need for transparency in how such technologies are implemented and used. Given the capacity of these systems to collect and analyze vast amounts of personal information, they have become focal points in discussions about the balance between security, public safety, and individual privacy rights.

**4. When is the median used instead of the mean?**

- A. With normally distributed data**
- B. With skewed data distributions**
- C. When calculating mode**
- D. With categorical data**

The median is often used instead of the mean particularly in the context of skewed data distributions because it provides a better measure of central tendency in such cases. When data is skewed, one or more extreme values can significantly affect the mean, pulling it in the direction of the skew. This can result in a mean that does not accurately represent the center of the data distribution. For example, consider a dataset of home prices in a neighborhood where most homes are valued around \$300,000 but a few luxurious homes are valued at \$1 million or more. The mean price will be elevated due to these high values, making it seem like homes are more expensive on average than they truly are for the majority of homeowners. The median, which is the middle value when the data is ordered, remains unaffected by these outliers and provides a clearer picture of a typical home price in that neighborhood. In contrast, while normally distributed data may have a mean and median that are similar, using the median in skewed distributions can present a more representative value for the center of the data. Other contexts mentioned, like calculating the mode or dealing with categorical data, do not directly relate to the advantages of using the median over the mean. Thus, the median is especially valuable

## 5. Which method is commonly used to address missing data?

- A. Normalization
- B. Imputation**
- C. Segmentation
- D. Encoding

Imputation is a commonly used method for addressing missing data in datasets. This technique involves replacing the missing values with substituted values, allowing the dataset to remain intact and usable for analysis. Imputation can take various forms, such as filling in missing values with the mean, median, mode of the data, or using more complex statistical approaches according to the underlying relationships in the data. Commonly, missing data can introduce bias or lead to inaccurate conclusions if not properly managed. By employing imputation, analysts can maintain the structure and integrity of the dataset, thereby enabling more robust statistical analysis or machine learning models. This method helps in generating complete datasets, which is crucial for valid results in data science projects. In contrast, normalization focuses on scaling the values of features to fall within a certain range, which does not directly address gaps in the data. Segmentation involves dividing a dataset into distinct groups or segments based on certain characteristics, and encoding refers to transforming categorical variables into numeric formats for analysis. While these methods are essential in various stages of data preprocessing, they do not specifically target the issue of missing data like imputation does.

## 6. What method involves solving problems using established rules and logic?

- A. Symbolic Reasoning**
- B. Inference
- C. Ethical AI
- D. Algorithmic Bias

Symbolic reasoning is a method that specifically involves applying established rules and logic to solve problems. It focuses on the manipulation of symbols and the relationships between them to derive conclusions, making it particularly useful in fields like mathematics, computer science, and artificial intelligence. By establishing clear rules and definitions, symbolic reasoning allows for systematic problem-solving and the formulation of logical arguments. In contrast, inference refers to the process of drawing logical conclusions from premises known or assumed to be true. While it does involve some degree of logic, it doesn't solely rely on a set of established rules in the way that symbolic reasoning does. Ethical AI pertains to the development and deployment of artificial intelligence systems in a manner that aligns with ethical guidelines, ensuring fairness, transparency, and accountability. Algorithmic bias addresses the biases that may occur in algorithms, impacting the outputs of AI systems, rather than serving as a structured problem-solving approach. Thus, the aspect of utilizing established rules and logic distinctly aligns with symbolic reasoning.

**7. Which architecture is typically associated with deep learning models?**

- A. Decision Trees**
- B. Support Vector Machines**
- C. Neural Networks**
- D. Linear Regression**

Deep learning models are primarily based on neural networks, which are designed to simulate the way human brains operate. These networks are composed of layers of interconnected nodes or "neurons," allowing them to learn complex patterns from large amounts of data through processes such as backpropagation and gradient descent. Neural networks can capture intricate relationships in data due to their multi-layered structure, enabling them to perform tasks such as image recognition, natural language processing, and more, which are characteristic of deep learning applications. The depth of these networks—reflected in the number of layers—allows for hierarchical feature extraction and represents a significant advancement over traditional machine learning techniques. In contrast, the other architectures listed, such as decision trees, support vector machines, and linear regression, do not embody the multi-layered, interconnected structure that defines deep learning. They are typically more straightforward models with limited capacity for handling high-dimensional data and complex data representations.

**8. What term refers to all possible outcomes in a scenario?**

- A. Event**
- B. Sample Space**
- C. Experiment**
- D. Outcome Set**

The term that refers to all possible outcomes in a scenario is known as the sample space. In probability and statistics, the sample space is the complete set of all outcomes that can occur from a given random experiment. For instance, if you are rolling a six-sided die, the sample space would include the numbers 1 through 6, representing each possible result of that roll. Understanding the concept of sample space is crucial in probability because it provides a foundational framework for calculating probabilities, determining likelihoods, and analyzing events. The remaining terms do not encapsulate the complete set of outcomes. An event refers to a specific set of outcomes, typically of interest within the context of the sample space. An experiment is the procedure or action that generates outcomes, and an outcome set is not a standard term in probability; it may mistakenly imply a collection of events rather than the entirety of possible outcomes. Thus, the sample space is the precise term that defines all possible outcomes in any probabilistic scenario.

**9. Which of the following is NOT a role of data analysts?**

- A. Building predictive algorithms**
- B. Creating detailed reports**
- C. Performing data storage and maintenance**
- D. Conducting qualitative research**

Data analysts primarily focus on examining data to extract insights, support decision-making, and enhance business outcomes. They excel in creating detailed reports that summarize their findings, handling data storage and maintenance to ensure data integrity and accessibility, and building predictive algorithms that leverage data to forecast trends and behaviors. Conducting qualitative research, while it may include some data analysis, typically falls under the purview of qualitative researchers or social scientists. This role emphasizes understanding phenomena through non-numerical data, such as interviews and observations, rather than quantifiable data analysis. It involves interpreting subjective experiences rather than the quantitative metrics that data analysts deal with. Thus, it is not considered a core responsibility of data analysts.

**10. What is termed as a variable with countable values, such as the number of students?**

- A. Continuous Variable**
- B. Discrete Variable**
- C. Nominal Variable**
- D. Interval Variable**

A discrete variable is characterized by having countable values, meaning it can take specific values that can be counted individually. For example, the number of students in a classroom is a discrete variable because you can count students as whole numbers—1, 2, 3, and so forth—without any fractional values. This type of variable is often associated with counts of items or occurrences where the values cannot be subdivided into finer increments. Continuous variables, on the other hand, can take on any value within a given range, including fractions or decimals, making them suitable for measurements rather than counts. Nominal variables classify data without a specific numerical order, such as categories like colors or names. Interval variables involve numerical data where the difference between values is meaningful, but lack a true zero point, such as temperature in Celsius. These distinctions clarify why a discrete variable is the correct classification for scenarios involving countable quantities.

## Next Steps

**Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.**

**As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.**

**If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at [hello@examzify.com](mailto:hello@examzify.com).**

**Or visit your dedicated course page for more study tools and resources:**

**<https://fbladatascienceai.examzify.com>**

**We wish you the very best on your exam journey. You've got this!**

SAMPLE