

**DAU™ Certified Data & Analytics Tester (CDAT)
Exam Questions Website**

Released
Version 2020 Syllabus

Data & Analytics United



Copyright © 2020 Data & Analytics United (hereinafter called DAU). All rights reserved.

Purpose of this document

This document contains exam questions for the website for DAU Certified Data & Analytics Tester (CDAT) in the English language.

The sample questions, answer sets and associated justifications in this document have been created by a team of subject matter experts and experienced question writers with the aim of assisting people who are planning to take the DAU Certified Data & Analytics Tester (CDAT) examination.

None of these questions are used in the official DAU Certified Data & Analytics Tester (CDAT) examination, but they are written to the same level of difficulty as the official certification exam. Data & Analytics United recognized training providers are able to use these questions as part of their training to prepare participants for the examination.

Instructions

The question and answer sets are organized in the following way:

- Learning Objectives / Chapters
- Question - including any scenario followed by the question stem
- Answer Set

General Information on the mock exam paper:

- Number of Questions: 10
- Time allowed: 15 Minutes (when English is not your native language, then you will have an extra 4 minutes)
- Total number of points: 10
- 65% required to pass (7 questions in this case)
- Please only choose one answer per question.

List of Chapters

- Chapter 1 - Introduction to Business Intelligence and Data & Analytics
- Chapter 2 - Data & Analytics Testing Strategy
- Chapter 3 - Test Techniques
- Chapter 4 - BI Testing
- Chapter 5 - Data Quality
- Chapter 6 - Environmental Needs

Question 1*(Correct answer is worth 1 point)*

When focusing on Predictive Advanced Analytics, what does one generally focus on?

- (a) What is the best that can happen.
- (b) What has happened.
- (c) Why is something happening.
- (d) What exactly is the problem.

Question 2*(Correct answer is worth 1 point)*

When discussing the V's for Big Data, what is meant by "V for Variety"?

- (a) "Variety" is for increasing amount of data types and sources.
- (b) "Variety" is for increasing range of data types and sources.
- (c) "Variety" is for increasing speed of changing data types.
- (d) "Variety" is for increasing quality of data types and sources.

Question 3*(Correct answer is worth 1 point)*

Which **ONE** of the following options **BEST** describes semantic testing?

- (a) Testing the validity of the input data and then in particular the relations between data.
- (b) Checking validity of the defined columns / fields.
- (c) Testing the dependencies of data columns / fields.
- (d) Checking the variety of the defined columns / fields.

Question 4*(Correct answer is worth 1 point)*

Which **ONE** of the following options is a layer in a data warehouse system?

- (a) Data Connection Layer.
- (b) Data Extraction Layer.
- (c) Staging Rules Engine.
- (d) Data Testing Layer.

Question 5*(Correct answer is worth 1 point)*

At a given company there are 50 employees available, but there are 60 registered in the HR system. Utilizing this above information what would be the **CORRECT** percentage of 'Uniqueness'?

- (a) 0,8%
- (b) 75%
- (c) 83,3%
- (d) 120%

Question 6*(Correct answer is worth 1 point)*

Driving Insurance Scenario:

The driver's age must be at least 18 years. Drivers are eligible for a driving insurance discount of 40% if they didn't have any accident for the last 10 years. Drivers are entitled to a 30% discount if they didn't have an accident in the last 5 to 9 years. Drivers are entitled to a 10% discount if they didn't have an accident in the last 3 to 4 years. Drivers who had an accident the last 2 years are not entitled for a discount.

Using the above scenario, what are the equivalence partitions?

- (a) Age is $0 \leq 18$ (invalid partition). No accidents: for ≥ 10 years (Valid Partition), ≥ 5 to < 10 (Valid Partition), ≥ 3 to ≤ 4 years (Valid Partition), Lower than 3 years (Valid Partition).
- (b) Age is $0 < 18$ (invalid partition). No accidents: for > 10 years (Valid Partition), > 5 to < 10 (Valid Partition), ≥ 3 to ≤ 4 years (Valid Partition), Lower than 3 years (Invalid Partition).
- (c) Age is $0 < 18$ (invalid partition). No accidents: for ≥ 10 years (Invalid Partition), ≥ 5 to < 10 (Valid Partition), > 3 to < 5 years (Valid Partition), Lower than 3 years (Valid Partition).
- (d) Age is $0 < 18$ (invalid partition). No accidents: for ≥ 10 years (Valid Partition), ≥ 5 to < 10 (Valid Partition), ≥ 3 to ≤ 4 years (Valid Partition), Lower than 3 years (Valid Partition).

Question 7*(Correct answer is worth 1 point)*

What does the abbreviation OLTP stand for in the context of data and analytics testing?

- (a) Offline Transaction Process.
- (b) Online Test Processing.
- (c) Online Transactional Processing.
- (d) Offline Transformation Processing.

Question 8*(Correct answer is worth 1 point)*

According to Deming, which **ONE** of the following options is the definition of Quality?

- (a) Quality is defined from the customer's point of view as anything that enhances their satisfaction.
- (b) Fitness for use. Those product features which meet the needs of customers and thereby provide product satisfaction. Freedom from deficiencies.
- (c) The totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs.
- (d) Quality is defined as knowledge of agents that would enable them to provide accurate and consistent solution to the customer at the very first attempt.

Question 9*(Correct answer is worth 1 point)*

Which **ONE** of the following options is the **BEST** definition of Timeliness?

- (a) The total collected amount of data from the required point in time.
- (b) The degree in timeline comparing two or more representations of data against a definition.
- (c) The degree to which data represent reality from the required point in time.
- (d) The degree of exact and precise usage of time and date, when comparing the "real world" with the test data.

Question 10*(Correct answer is worth 1 point)*

What is the reason for testers to execute Data Profiling?

- I. Facilitate better Modeling
 - II. Improve Coverage Levels of Tests
 - III. Enable testing of Error Handling and Audit Trails
 - IV. Facilitate better Data Integration
-
- (a) Options I & IV are correct.
 - (b) Option II & III are correct.
 - (c) Option I, II & IV are correct.
 - (d) All of above reasons (options 1, 2, 3 & 4) are correct.

Answer Key:

Question 1: Answer C

Question 2: Answer B

Question 3: Answer A

Question 4: Answer B

Question 5: Answer C

Question 6: Answer D

Question 7: Answer C

Question 8: Answer A

Question 9: Answer C

Question 10: Answer D