

### Question 1

In the context of data science methodology, select the answer that describes what happens during the Data Requirements stage. 1 / 1 point

- Selecting the programming language for analysis
- Organizing data into columns and rows
- Identifying the necessary data content, formats, and sources
- Defining the ingredients needed for a meal

Correct

Correct! The Data Requirements stage in data science involves identifying the necessary data content, formats, and sources for the analysis. This stage is comparable to identifying the ingredients needed for a meal before cooking.

### Question 2

What technique can you apply to the data set during the data collection stage to assess data content, quality, and initial insights? 1 / 1 point

- Data extraction and merging
- Predictive model
- Data visualization and descriptive statistics
- Data preparation

Correct

Correct! You can apply data visualization and descriptive statistics to the data set during the data collection stage to assess content, quality, and initial insights about the data, similar to examining the ingredients on a cutting board.

### Question 3

Imagine you are a data scientist working on a healthcare project. Your task is to define the data requirements for a decision tree classification approach to predict patient readmission rates.

Select the answer that describes the purpose of defining data requirements in the context of the healthcare project. 1 / 1 point

- To select the appropriate patients for the cohort
- To identify the necessary data content, formats, and source
- To identify the best programming language for analysis
- To determine the total budget for the project

Correct

Correct! Defining data requirements involves identifying the necessary data content, formats, and sources for the analysis, similar to identifying the ingredients needed for a meal.

#### Question 4

You are a data scientist tasked with collecting and preparing data for a predictive modelling project related to patient outcomes. You have encountered some data sources, mainly drug information, that are missing. How would you handle this situation?

What is a suitable approach when encountering missing data sources in the data collection process context? 1 / 1 point

Attempt to acquire the missing data after obtaining intermediate results.

Exclude the entire patient cohort from the analysis.

Discard the project and start over with a different methodology.

Proceed with incomplete data and hope for accurate results.

Correct

Correct! In the provided scenario, it is acceptable to defer decisions about unavailable data and attempt to acquire the data later, such as after obtaining intermediate results from predictive modeling.

#### Question 5

Which stage of the methodology involves collaborating with DBAs and programmers to extract and merge data from various sources? 1 / 1 point

Data Preparation

Data Requirements

Data Collection

Data Understanding

Correct

Correct! Data collection involves extracting and merging data from various sources, and collaboration with DBAs and programmers is essential to ensuring data quality and efficiency.