

Exam : C1000-154

**Title : IBM Watson Data Scientist
v1**

Version : DEMO

1. When anticipating additional data sources that might be relevant, what is a crucial factor to consider?

- A. The color scheme of the data visualization
- B. The data source's popularity on social media
- C. The relevance of the data source to the business problem
- D. The graphical interface of the data source

Answer: C

2. A virtual assistant has been developed and deployed based on the Watson Assistant service. The assistant will support customers by answering FAQs (Frequent Answered Questions).

Which metric is a good indicator of the performance of the virtual assistant?

- A. The Area Under the Curve (AUC)
- B. Measure escalated calls using A/B testing
- C. The Root Mean Squared Error (RMSE) of words
- D. The F1 score of predicted intents in the Analytics tab

Answer: B

3. Which of the following is a critical first step in understanding a business problem for data science projects?

- A. Selecting the machine learning algorithm
- B. Defining the project scope
- C. Choosing the visualization tools
- D. Deploying the model

Answer: B

4. How can data splits be made reproducible in a machine learning experiment?

- A. By using a different random seed each time the data is split
- B. By partitioning the data manually
- C. By using a consistent random seed when splitting the data
- D. By splitting the data in a sequential manner without randomization

Answer: C

5. What is the key difference between batch processing and streaming in data processing?

- A. Batch processing involves real-time data processing, whereas streaming does not process data
- B. Streaming is suitable for large, historical datasets, whereas batch processing is for real-time data analysis
- C. Batch processing processes data in large blocks at a time, whereas streaming processes data in real-time as it arrives
- D. Batch processing processes data in large blocks at a time, whereas streaming processes data in real-time as it arrives

Answer: C