Some questions covering different aspects of **Data Science Methodology unit-3**:

**1. Business Understanding**

1. Explain the significance of the **Business Understanding** phase in the data science lifecycle. How does it influence the success of a data science project?
2. Using a real-world example, Describe how business objectives are translated into data science problems.
3. Discuss the challenges faced in defining business problems accurately in data science projects and how they can be mitigated.

**2. Analytic Approach**

1. Explain the difference between a **descriptive, predictive, and prescriptive** analytic approach in data science. Provide examples for each.
2. How do data scientists decide whether to use a **machine learning-based** or a **rule-based** approach to solve a problem?
3. Discuss the importance of **choosing the right analytic approach** and how incorrect selection can impact a project.

**3. Data Requirements & Data Collection**

1. What are the key factors to consider while defining **data requirements** for a data science project? Explain with examples.
2. Describe various **data sources** used in data science (structured, unstructured, internal, external). How do they impact model performance?
3. Discuss the ethical and legal challenges associated with **data collection and usage** in data science projects.

**4. Data Understanding & Data Preparation**

1. Explain the role of **Exploratory Data Analysis (EDA)** in the Data Understanding phase. What techniques are commonly used?
2. Describe the steps involved in **data cleaning, transformation, and feature engineering** with suitable examples.
3. Why is **data preprocessing** important in a data science project? Discuss the impact of missing values, outliers, and data scaling.

**5. Modeling & Evaluation**

1. Compare and contrast **supervised and unsupervised learning models**. In what scenarios are they used in data science?
2. Explain different **model evaluation techniques** (accuracy, precision, recall, F1-score, ROC curve) and their significance.
3. Discuss the importance of **model deployment and feedback loops** in real-world applications. How does continuous evaluation improve model performance?