2.1 What is Integration Runtime in Azure Data Factory, and what are its types?  
which one you used in your projects?  
2.2 What are the different types of pipeline triggers available in Azure Data Factory?  
2.3 Can you explain each type of trigger in ADF and their use cases?  
2.4 When would you use a Tumbling Window Trigger in ADF?  
2.5 What is the difference between Tumbling Window Trigger and Schedule Trigger in ADF?  
2.6 If a pipeline is taking longer than expected to run, what steps would you follow to troubleshoot and resolve the issue?  
2.7 What are the key activities you have used while creating pipelines in your project?  
2.8 What is the difference between the Get Metadata activity and the Lookup activity in Azure Data Factory?  
2.9 You have designed a critical pipeline with multiple activities. If it fails during execution, what steps would you follow to identify and resolve the issue?  
2.10 What are some of the key challenges you have faced in your project, and how did you overcome them?  
2.11 Can you explain the difference between variables and parameters in Azure Data Factory?  
2.12 Can you describe the incremental load approach you implemented while creating your pipeline?  
2.13 Why would you use PySpark in Databricks for data transformation when similar transformations can be done using Mapping Data Flows in Azure Data Factory?  
  
Apache Spark :  
  
3.1 What optimization techniques have you implemented while working with Apache Spark?  
3.2 Can you write a PySpark code snippet to demonstrate the salting technique used to handle data skew?  
3.3 Can you write a PySpark code snippet to replace null values in a DataFrame with the column’s mean, median, or mode?  
3.4 What are variables in Apache Spark, and can you explain their usage briefly?  
3.5 Can you write a PySpark code snippet to drop rows from a DataFrame where any column has a null value?  
  
SQL Question:  
  
4.1 There is a queue of people waiting to board a bus. However, the bus has a weight limit of 1000 kilograms, so there may be some people who cannot board.  
Write a solution to find the person\_name of the last person that can fit on the bus without exceeding the weight limit. The test cases are generated such that the first person does not exceed the weight limit.  
CREATE TABLE queue (  
 id INT PRIMARY KEY,  
 name VARCHAR(100),  
 weight INT,  
 turn INT  
);  
INSERT INTO queue (id, name, weight, turn) VALUES  
(5, 'Raftar', 250, 1),  
(4, 'HHH', 175, 5),  
(3, 'Virat', 400, 2),  
(6, 'Messi', 400, 3),  
(1, 'Sania', 500, 6),  
(2, 'Bolt', 200, 4);