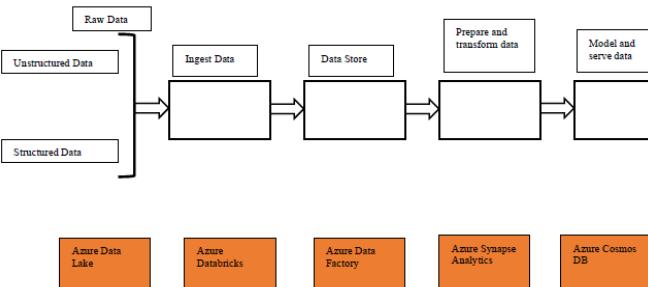


**1. [Marks: 5] Explain below the 5 components shown in orange boxes. Explain which Azure components you will use where in this big data architecture and why.**



*Azure Data Lake:*

Scalable raw data storage used to store large volumes of structured and unstructured raw data and act as the first landing zone from a variety of resources.

*Azure Data Factory:*

Cloud-based data integration service for orchestrating and automating data movement and workflows.

*Azure Synapse Analytics:*

Data platform act as analytical data warehouse and used for storing transformed data and supporting large-scale querying and reporting

*Azure Databricks:*

Apache spark-based analytics platform optimized for Microsoft Azure provided collaborative environment for sorting, cleaning, joining and outputting the data.

*Azure Cosmos DB:*

Globally distributed, multi-model database service for real-time application and aim to serve processed data to end-user applications at low latency

In this big data architecture:

**Ingest Data → Azure Data Factory:** Ingest the unstructured and structured data for preparation of data storing

**Data Store → Azure Data Lake:** Store the data collected from Azure Data Factory

**Prepare and Transform Data → Azure Databricks:** From Azure Data Factory, process and transform the data for data analysis

**Model and Serve Data → Azure Synapse Analytics:** Provide limitless analytics service for large-scale dataset and combine with data warehousing as big data analysis and modeling service

**2. [Marks: 5] Explain how Stream Analytics works in Azure. Mention at least two common use cases or applications for this service.**

In Microsoft Azure, Azure Stream Analytics works as a real-time analytics service for processing data streams from various sources like real-time IoT device or applications. It digests high-velocity data from the data source and uses a SQL-like language to filter, aggregate, and detect patterns within the data stream. It can output the data to multiple data sinks like data lakes or Azure SQL database and enabling real-time insight and actions.

**Two common use case for Azure Stream Analytics:**

1. Real-time IoT device monitoring:  
process real-time data from IoT devices or sensors. It can be used to tracking the temperature of sensors or set up alerts when any data is monitored
2. Social media sentiment monitoring and analysis:  
It can monitor ingest real-time Twitter(X) or Instagram feeds for trending topics or public sentiment and provide visualizations

**3. [Marks: 10] Deploy all the resources in Azure Portal. Implement a Stream Analytics job by using the Azure portal. See this for reference - <https://learn.microsoft.com/en-us/azure/streamanalytics/stream-analytics-quick-create-portal>**

For query use below:

**SELECT \***

**INTO BlobOutput**

**FROM IoTHubInput**

**HAVING Temperature > 25**

See the below screenshot and show the top 30 results for your output.

The screenshot shows the Azure Storage Blobs interface. A file named '0\_ee2fc9a55474864a45dc288edc6231.json' is selected. The file content is displayed as a JSON array of 30 objects, each representing a message from a Raspberry Pi web client. The properties include 'deviceId', 'temperature', and 'humidity'. The temperature values are consistently above 25, as specified in the query.

```

[{"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 27.5628055546589, "humidity": 77}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 31.7608318991918, "humidity": 64}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 29.7808318991918, "humidity": 64}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 29.7808318991918, "humidity": 70}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 29.537538745134363, "humidity": 73}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 29.537538745134363, "humidity": 73}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 29.537538745134363, "humidity": 73}, {"messageId": "776", "deviceId": "Raspberry Pi web Client", "temperature": 27.2370885128513, "humidity": 63}, {"messageId": "780", "deviceId": "Raspberry Pi web Client", "temperature": 30.5439280849564, "humidity": 60}, {"messageId": "780", "deviceId": "Raspberry Pi web Client", "temperature": 29.7808318991918, "humidity": 64}, {"messageId": "791", "deviceId": "Raspberry Pi web Client", "temperature": 23.68798512762802, "humidity": 75}, {"messageId": "791", "deviceId": "Raspberry Pi web Client", "temperature": 23.68798512762802, "humidity": 75}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 28.8599218243889, "humidity": 69}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 30.7820961384389, "humidity": 71}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 30.7820961384389, "humidity": 71}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 30.7820961384389, "humidity": 71}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 33.1118801842066, "humidity": 64}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 27.1360801528064, "humidity": 64}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 27.1360801528064, "humidity": 64}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 33.0122804325575, "humidity": 68}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 33.0122804325575, "humidity": 68}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 33.0122804325575, "humidity": 68}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 33.7359424979905, "humidity": 68}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 27.80682207890739, "humidity": 60}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 27.77980807373895, "humidity": 68}, {"messageId": "794", "deviceId": "Raspberry Pi web Client", "temperature": 27.77980807373895, "humidity": 68}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 30.3662234502802, "humidity": 64}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 29.412532075731782, "humidity": 69}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 29.412532075731782, "humidity": 69}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 29.412532075731782, "humidity": 69}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 30.88946403523278, "humidity": 64}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 30.88946403523278, "humidity": 64}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 30.88946403523278, "humidity": 64}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 33.1758030772866, "humidity": 64}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 33.1758030772866, "humidity": 64}, {"messageId": "806", "deviceId": "Raspberry Pi web Client", "temperature": 33.74810189795484, "humidity": 64}

```

Create stream analytics jobs, resource group, storage account, IOT Hub etc. as below

The screenshot shows the Azure Resource Groups blade. It displays a list of resources across different categories. The resources listed include:

- Stream Analytics job: 1628a5
- Resource group: DefaultResourceGroup-eastus2
- Storage account: 1628a5
- IoT Hub: 1628a5
- Machine learning online endpoint: mle1628a3-yolnr
- Subscription: Azure subscription 1

## Query as below:

The screenshot shows the Azure Stream Analytics job configuration interface. The left sidebar lists the job's overview, activity log, access control (IAM), tags, and diagnostic settings. The main area shows the inputs and outputs configuration. Under 'Inputs' (2), there is an IoT Hub input named '1628a5-1'. Under 'Outputs' (2), there is a Blob Output named '1628a5-1'. The 'Query' tab is selected, displaying the following Stream Analytics Query Language (SAL) code:

```

1  2  Here are links to help you get started with Stream Analytics Query Language:
2  3  Common query patterns - https://go.microsoft.com/fwlink/?linkid=2119153
3  4  Query language - https://docs.microsoft.com/stream-analytics-query/query-language-elements-azure-stream-analytics
4  5  */
5  6  SELECT *
6  7  INTO BlobOutput
7  8  FROM IoTHubInput
8  9  HAVING Temperature > 25

```

Below the query editor, the 'Input preview' tab is active, showing the message 'While sampling data, no data was received from 2 partitions.'

The screenshot shows the Raspberry Pi Azure IoT Online Simulator interface. On the left, a Fritzing schematic diagram of a Raspberry Pi connected to a DHT22 sensor and a relay module. On the right, a code editor window titled 'Raspberry Pi Azure IoT Online Simulator' displays C# code for a Raspberry Pi device. The code handles sensor data, connects to an IoT hub, and manages a relay. A red box highlights the 'Run' button at the bottom of the code editor.

sent query:

The screenshot shows the Raspberry Pi Azure IoT Online Simulator interface. The Fritzing schematic and code editor are identical to the previous screenshot. A red box highlights the 'Output' terminal window at the bottom right, which displays the results of the executed C# code. The output shows messages being sent to an IoT hub, including sensor data and relay status updates.

## Top 30 Result similar to the provided screenshot:

The screenshot shows the Microsoft Azure Storage Explorer interface. A container named '1628a5' is selected. Inside, a blob named '0\_b4f317c51fd46d089b558c1266545d2\_1.json' is displayed. The JSON file contains 32 lines of data, each representing a message from a Raspberry Pi Web Client. The data includes fields for deviceId, deviceName, temperature, and humidity. The JSON structure is as follows:

```
1  {"deviceId": "37", "deviceName": "Raspberry Pi Web Client", "temperature": 28.29419891976918, "humidity": 74.99730805649862, "EventProcessedU  
2  {"deviceId": "38", "deviceName": "Raspberry Pi Web Client", "temperature": 25.928187295335203, "humidity": 72.05698104844154, "EventProcessedU  
3  {"deviceId": "41", "deviceName": "Raspberry Pi Web Client", "temperature": 28.056476120846604, "humidity": 68.44382094693071, "EventProcessedU  
4  {"deviceId": "43", "deviceName": "Raspberry Pi Web Client", "temperature": 25.5979331363464, "humidity": 70.9647078872492, "EventProcessedU  
5  {"deviceId": "44", "deviceName": "Raspberry Pi Web Client", "temperature": 27.668318701841113, "humidity": 77.20033910648651, "EventProcessedU  
6  {"deviceId": "45", "deviceName": "Raspberry Pi Web Client", "temperature": 30.4148008815595, "humidity": 74.72594873717682, "EventProcessedU  
7  {"deviceId": "46", "deviceName": "Raspberry Pi Web Client", "temperature": 29.455991171872704, "humidity": 70.43151579087306, "EventProcessedU  
8  {"deviceId": "47", "deviceName": "Raspberry Pi Web Client", "temperature": 29.4348131439711, "humidity": 63.56514635658544, "EventProcessedU  
9  {"deviceId": "48", "deviceName": "Raspberry Pi Web Client", "temperature": 29.42836702570044, "humidity": 76.18019636508481, "EventProcessedU  
10 {"deviceId": "49", "deviceName": "Raspberry Pi Web Client", "temperature": 29.4372578562488, "humidity": 79.6372578562488, "EventProcessedU  
11 {"deviceId": "50", "deviceName": "Raspberry Pi Web Client", "temperature": 25.75423840173954, "humidity": 67.791545169828424, "EventProcessedU  
12 {"deviceId": "51", "deviceName": "Raspberry Pi Web Client", "temperature": 27.794169287744975, "humidity": 70.92995727086105, "EventProcessedU  
13 {"deviceId": "52", "deviceName": "Raspberry Pi Web Client", "temperature": 28.243807062254252, "humidity": 65.73780074751548, "EventProcessedU  
14 {"deviceId": "54", "deviceName": "Raspberry Pi Web Client", "temperature": 30.92193863758927, "humidity": 65.51378791578657, "EventProcessedU  
15 {"deviceId": "55", "deviceName": "Raspberry Pi Web Client", "temperature": 30.96967724544976, "humidity": 69.5991652695911, "EventProcessedU  
16 {"deviceId": "56", "deviceName": "Raspberry Pi Web Client", "temperature": 31.46297753746157, "humidity": 62.76487901172665, "EventProcessedU  
17 {"deviceId": "57", "deviceName": "Raspberry Pi Web Client", "temperature": 27.660789731151297, "humidity": 69.19458268083938, "EventProcessedU  
18 {"deviceId": "59", "deviceName": "Raspberry Pi Web Client", "temperature": 28.639216460600394, "humidity": 71.02459328116849, "EventProcessedU  
19 {"deviceId": "60", "deviceName": "Raspberry Pi Web Client", "temperature": 25.92898722724147, "humidity": 77.22870626583418, "EventProcessedU  
20 {"deviceId": "63", "deviceName": "Raspberry Pi Web Client", "temperature": 25.842716380900413, "humidity": 64.28618167228804, "EventProcessedU  
21 {"deviceId": "64", "deviceName": "Raspberry Pi Web Client", "temperature": 29.596278491354678, "humidity": 72.862936128122, "EventProcessedU  
22 {"deviceId": "65", "deviceName": "Raspberry Pi Web Client", "temperature": 28.489417244337243, "humidity": 61.91562999876396, "EventProcessedU  
23 {"deviceId": "67", "deviceName": "Raspberry Pi Web Client", "temperature": 28.018849308844772, "humidity": 68.557999519584, "EventProcessedU  
24 {"deviceId": "68", "deviceName": "Raspberry Pi Web Client", "temperature": 29.90156931660552, "humidity": 66.758638858001454, "EventProcessedU  
25 {"deviceId": "70", "deviceName": "Raspberry Pi Web Client", "temperature": 30.38924193054517, "humidity": 69.21235188080895, "EventProcessedU  
26 {"deviceId": "71", "deviceName": "Raspberry Pi Web Client", "temperature": 28.873844241538437, "humidity": 74.9891044634951, "EventProcessedU  
27 {"deviceId": "76", "deviceName": "Raspberry Pi Web Client", "temperature": 28.146925266335245, "humidity": 74.28004282770246, "EventProcessedU  
28 {"deviceId": "77", "deviceName": "Raspberry Pi Web Client", "temperature": 26.37030371774628, "humidity": 76.9538268472931, "EventProcessedU  
29 {"deviceId": "78", "deviceName": "Raspberry Pi Web Client", "temperature": 31.0584362488063, "humidity": 63.715362688130612, "EventProcessedU  
30 {"deviceId": "79", "deviceName": "Raspberry Pi Web Client", "temperature": 27.300988514034355, "humidity": 66.640213319819324, "EventProcessedU  
31 {"deviceId": "82", "deviceName": "Raspberry Pi Web Client", "temperature": 31.231224857089956, "humidity": 74.6445881121089, "EventProcessedU  
32 {"deviceId": "85", "deviceName": "Raspberry Pi Web Client", "temperature": 29.67245090427714523, "humidity": 75.445065916727632, "EventProcessedU
```