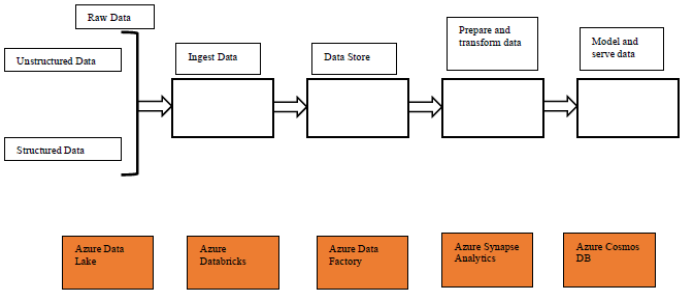
1. [Marks: 5] Explain below the 5 components shown in orange boxes. Explain which Azure components you will use where in this big data architecture and why.



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

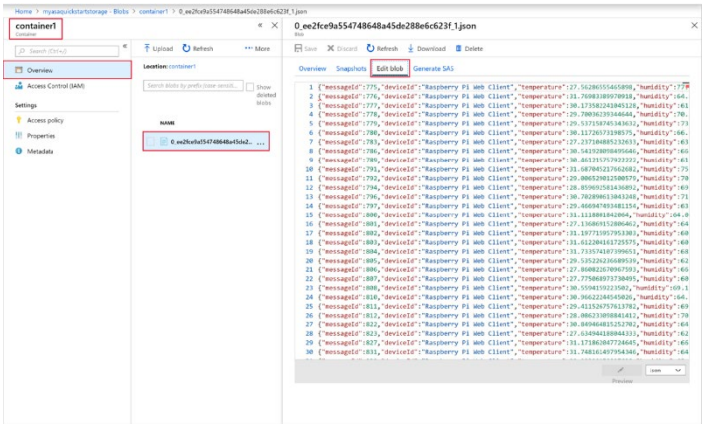
2. [Marks: 5] Explain how Stream Analytics works in Azure. Mention at least two common use cases or applications for this service.

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.

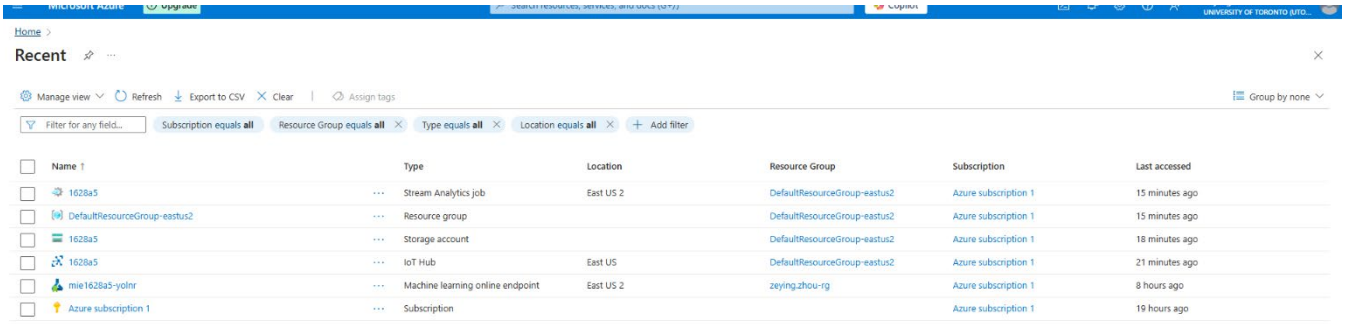
Two common use case for Azure Stream Analytics:

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

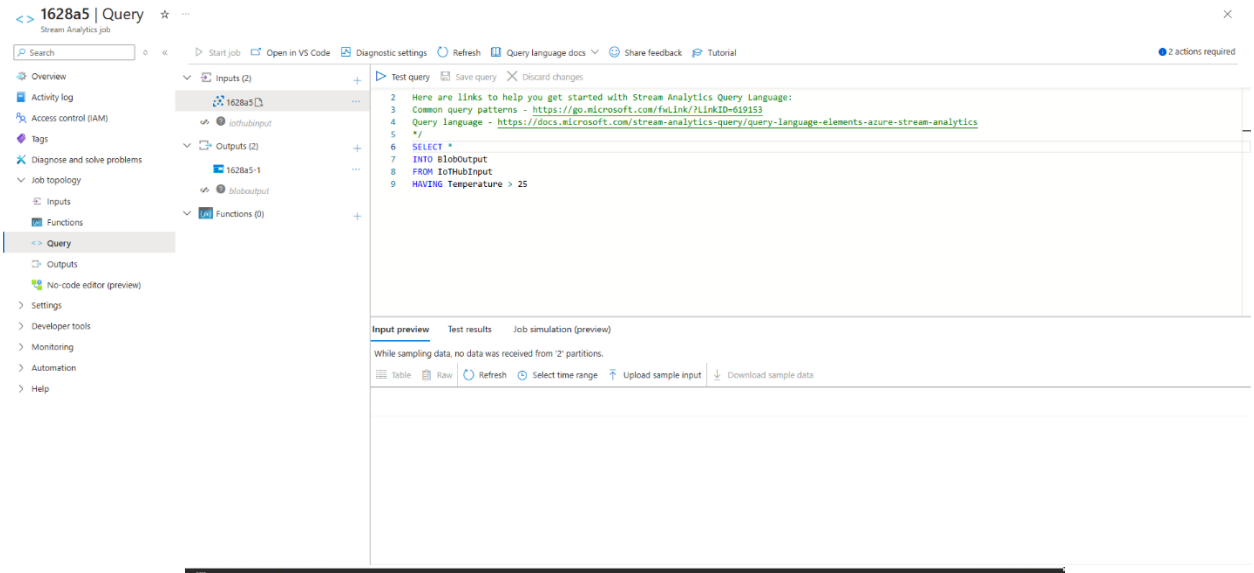
3. [Marks: 10] Deploy all the resources in Azure Portal. Implement a Stream Analytics job by using the Azure portal. See this for reference - https://learn.microsoft.com/en-us/azure/streamanalytics/stream-analytics-quick-create-portal
For query use below:
SELECT *
INTO BlobOutput
FROM IoTHubInput
HAVING Temperature > 25
See the below screenshot and show the top 30 results for your output.

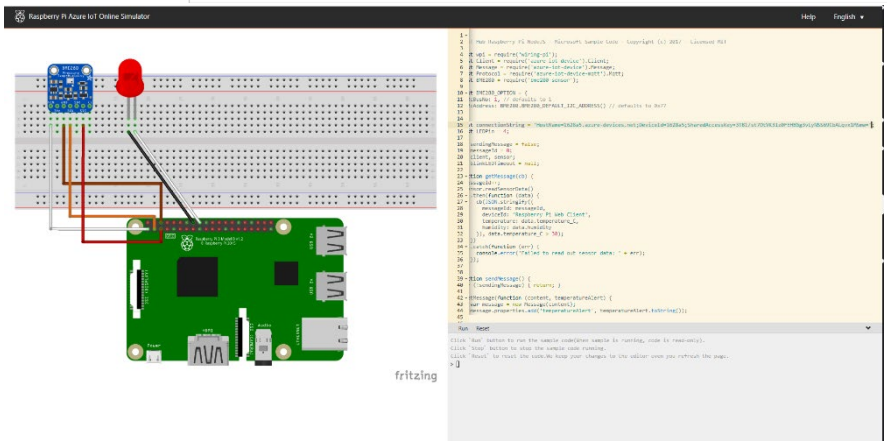


Create stream analytics jobs, resource group, storage account, IOT Hub etc. as below

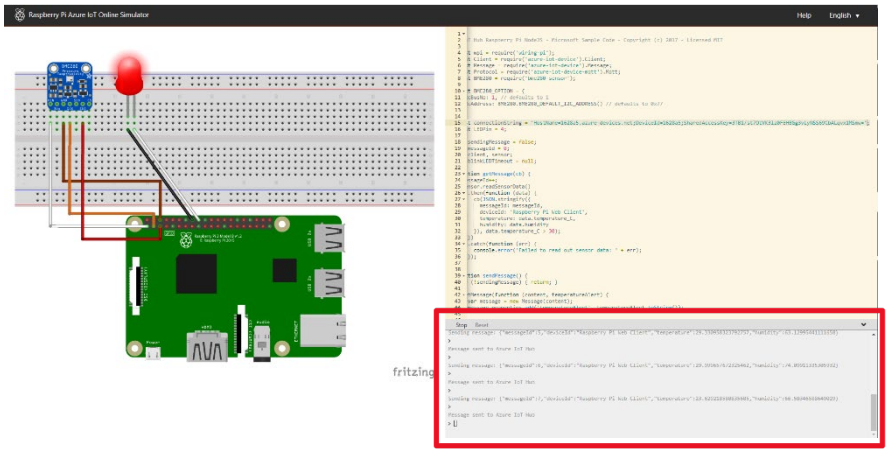


Query as below:





sent query:



Top 30 Result similar to the provided screenshot: