**Assignment: Real-Time Data Pipeline with Azure**

1. **Set Up Azure Resources**
2. Resource Name: **Resources23**
3. Event Hub Name: **myevent-hub**
4. Storage Accounts: **microstoragedb**

Containers: **raw**

**processed**

1. Stream Analytics Job

Job1: **streamananlytics1**

Job 2: **streamananlytics2**

A screenshot of a computer

Description automatically generated

1. **Configure Event Hub and Generate Data**
2. Event Hub Instance Name: **myinstance**
3. Data Generation:
4. Go to Data Explorer (preview)
5. Select your hub (**myinstance**) 🡪 Send Event
6. Select Dataset as: **Weather data**

**Code:** {

"country": "USA",

"city": "New York",

"date": "2023-05-11",

"temperature": 20,

"humidity": 60,

"windSpeed": 10,

"windDirection": "NW",

"precipitation": 0,

"cloudCover": 20,

"visibility": 10,

"pressure": 1013,

"dewPoint": 10,

"uvIndex": 5,

"sunrise": "05:30",

"sunset": "20:15",

"moonrise": "22:00",

"moonset": "08:00",

"moonPhase": "Waning Gibbous",

"conditions": "Partly Cloudy",

"icon": "partly-cloudy-day"

}

A screenshot of a computer

Description automatically generated

1. **Stream Analytics Job 1: Event Hub to Raw Container**

Go to Job Topology

Add input 🡪 Select Event Hub

1. **Input:** myinstance

Event Hub Name: myevent-hub

Save and Test

1. **Output:** raw

**Event:** json

**Path Pattern:** raw/{date}/{hour}/events.json

Save and Test

1. **Query:**

Code: SELECT \* INTO [raw]FROM [myinstance]

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. **Stream Analytics Job 2: Raw to Processed Container**

Go to Job Topology

Add input 🡪 Select Blob Storage

1. **Input:** raw

Save and Test

Go to Job Topology

Add output 🡪 Select Blob Storage

1. **Output:** processed

Path Pattern:raw/{date}/{hour}/events.json

Save and Test

1. **Query:**

Code:

SELECT

     sensorId,

      AVG(temperature) AS AvgTemperature,

     System.Timestamp AS ProcessedTime

INTO [processed]

FROM [raw]

GROUP BY

     sensorId,

     TumblingWindow(minute, 5)

1. **Verification and Testing**
2. Raw:
3. Processed:
4. **Documentation and Submission**