

Assignment 3: Data Streams

Course: Data Management and Business Intelligence

Professors: D.Chatziantoniou, S.Safras

Students: Ch.Petrakogiannis (F2822112), I.Dimos (F2822102)

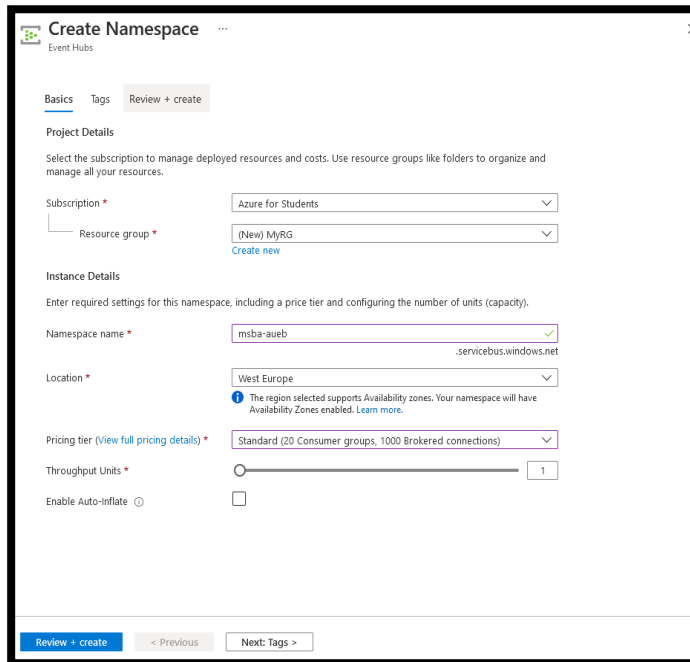
TABLE OF CONTENTS

INTRODUCTION	2
EVENT HUB	2
STORAGE ACCOUNT	4
STREAM ANALYTICS JOB	5
BLOB STORAGE OUTPUT.....	8
QUERIES.....	9

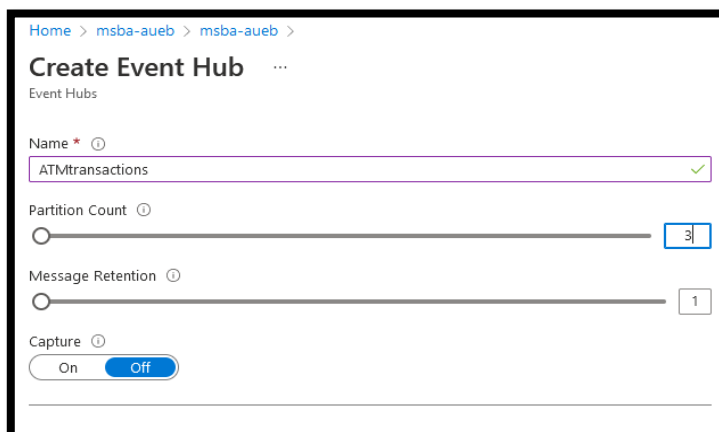
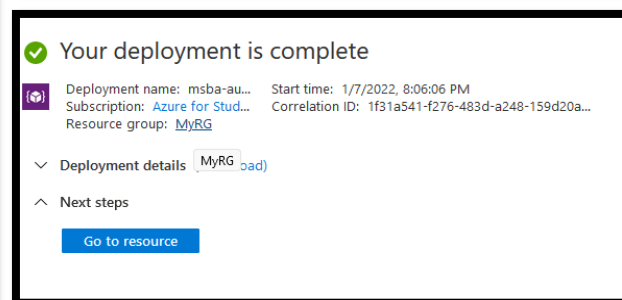
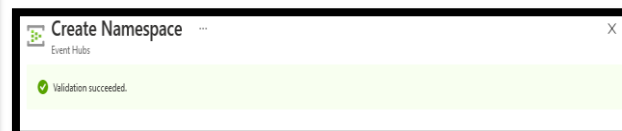
INTRODUCTION

In this specific assignment we aim to work using Azure Stream Analytics in order for us to process a data stream of ATM transactions and give answers to some specified stream queries. For this reason, firstly we had to create a student account on Azure platform and then proceed with the rest of the appropriate steps, that we are going to present in the sections below, in order to succeed our goal.

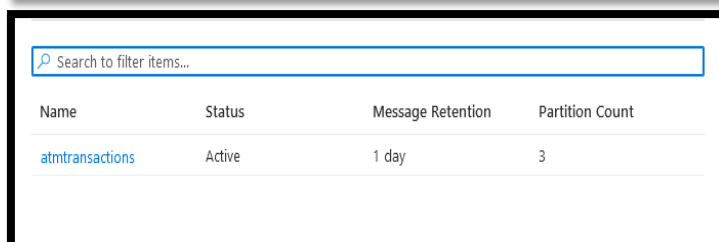
EVENT HUB



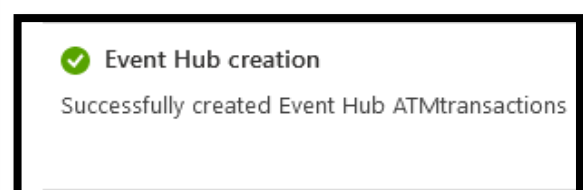
Comment: In first place we had to deploy the event hub feature using a new resource group through our students account.

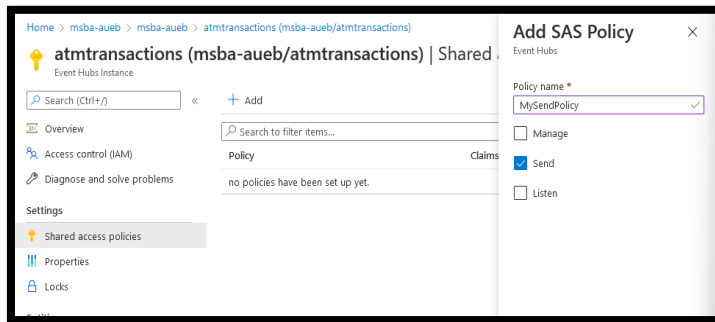


Comment: Then we had to proceed on the creation of a new event hub.



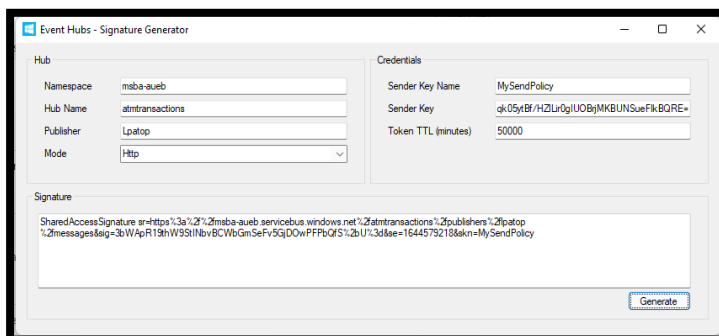
Name	Status	Message Retention	Partition Count
atmtransactions	Active	1 day	3



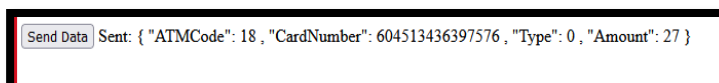


Comment: Later on, we had to set up a send policy on this new event hub that confers the right to send messages to an entity.

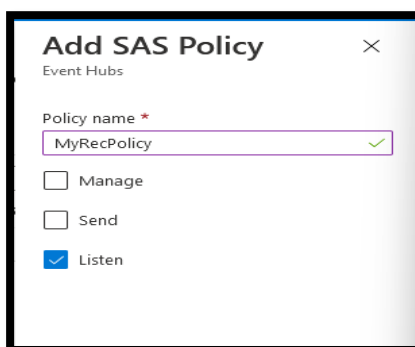
✓ Creating SAS Policy: MySendPolicy
Successfully created SAS policy 'MySendPolicy'.



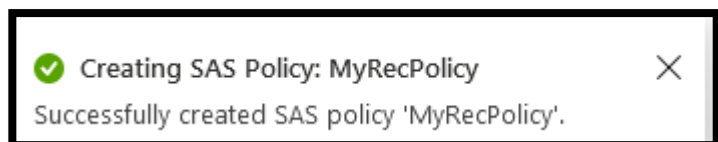
Comment: The next step had to do with the generator in which we updated the CONFIG variables and then we fed the event hub with the use of Generator.html in order to start the data stream through a browser.



Comment: The stream has worked successfully!



Comment: Next in queue was the creation of a receive policy in order for our event hub to obtain the data from the stream.



STORAGE ACCOUNT

Create a storage account ...

Basics Advanced Networking Data protection Tags Review + create

Select the subscription in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription * Azure for Students

Resource group * MyRG [Create new](#)

Instance details

If you need to create a legacy storage account type, please click [here](#).

Storage account name * dimpetsa

Region * (Europe) West Europe

Performance *
 ☒ Standard: Recommended for most scenarios (general-purpose v2 account)
 ☐ Premium: Recommended for scenarios that require low latency.

Redundancy * Geo-redundant storage (GRS)
 ☒ Make read access to data available in the event of regional unavailability.

[Review + create](#) < Previous Next : Advanced >

Comment: After the successful creation of the event hub and the data stream process, we had to create a storage account and use its container for uploading the reference data json files.



Dashboard > dimpetsa_1641580726646 | Overview

Deployment

Search (Ctrl+/) Delete Cancel Redeploy Refresh

Go to resource Pin to dashboard

We'd love your feedback! →

✓ Your deployment is complete

Deployment name: dimpetsa_1641580726... Start time: 1/7/2022, 8:38:54 PM
Subscription: Azure for Students Correlation ID: 810ad4c5-a095-4001-be99-00a6667c...
Resource group: MyRG

Deployment details (Download)

Next steps

Go to resource

Dashboard > dimpetsa

Storage account

dimpetsa | Containers

Search (Ctrl+/) + Container Change access level Restore container

Search containers by prefix Show deleted containers

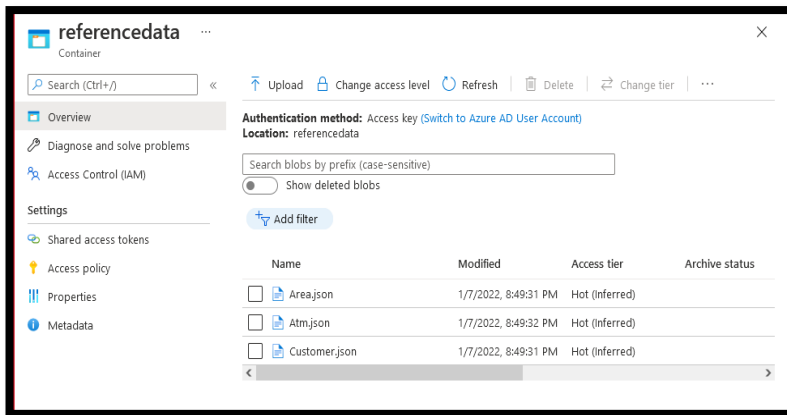
Name	Last modified	Public access level
logs	1/7/2022, 8:39:36 PM	Private

New container

Name * referencedata

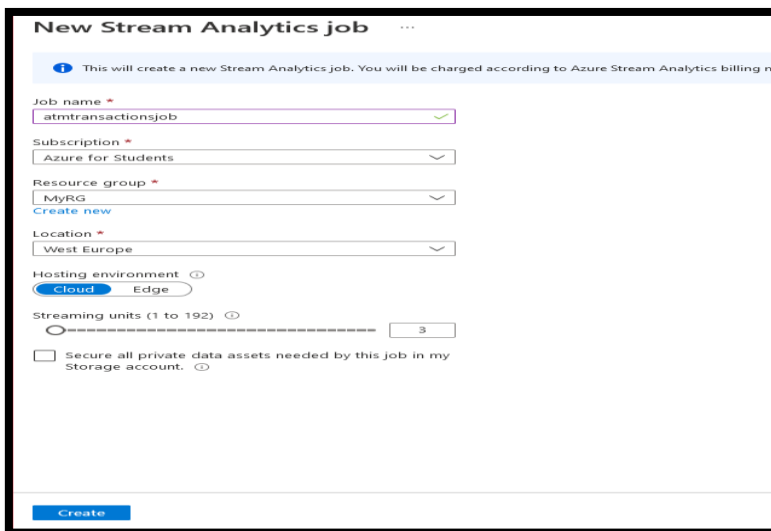
Public access level * Private (no anonymous access)

Advanced

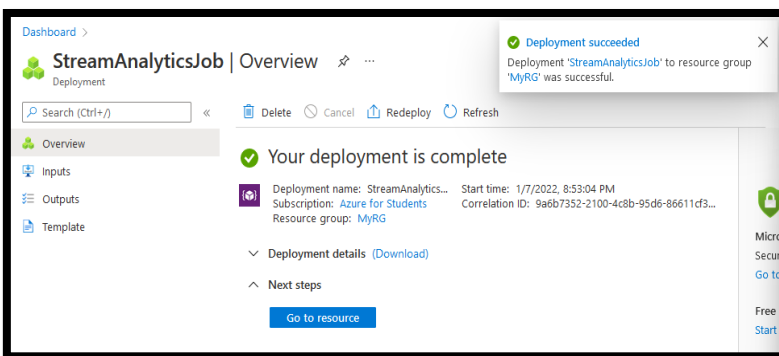


Comment: We have uploaded the files successfully!

STREAM ANALYTICS JOB



Comment: The next step for us in order to come closer to the end, was to set up a new stream analytics job and use the Event Hub + Reference Data Files as input.



Event Hub

New input

Input alias * ✓

☐ Provide Event Hub settings manually
☒ Select Event Hub from your subscriptions

Subscription

Event Hub namespace * ⓘ

Event Hub name * ⓘ
☐ Create new ☒ Use existing

Event Hub consumer group * ⓘ
☒ Create new ☐ Use existing

Authentication mode

The Azure Event Hubs Data Owner role will be granted to the Managed Identity for this Stream Analytics job when you click Save. If grant fails follow the manual grant steps [here](#).

Partition key ⓘ

Event serialization format * ⓘ

Encoding ⓘ

[Save](#)

Successful connection test

Connection to input 'streaminput' succeeded.

Comment: After the Event Hub input we had to upload the reference data and connect them with the stream analytics job.

Home > atmtransactionsjob >

Inputs

[+ Add stream input](#)
[+ Add reference input](#)
[Refresh](#)

Name	Source type	Source
ReferenceDataInputAtm	Reference	Blob storage
ReferenceDataInputCustomer	Reference	Blob storage
referencedataareainput	Reference	Blob storage
streaminput	Stream	Event Hub

Input details

ReferenceDataInputAtm

Test Delete

Input alias
ReferenceDataInputAtm

☒ Provide Blob storage/ADLS Gen2 settings manually
☐ Select Blob storage/ADLS Gen2 from your subscriptions

Subscription
Subscription information not needed

Storage account * ⓘ
dimpetsa

Container
☐ Create new ☒ Use existing
 referencedata

Authentication mode
 Managed Identity: System assigned
 Managed Identity authentication for Blob storage/ADLS Gen2 may need additional steps to work. [Learn more](#)

Path pattern * ⓘ
Atm.json

Date format
YYYY/MM/DD

Input details

ReferenceDataInputCustomer

Test Delete

Input alias
ReferenceDataInputCustomer

☒ Provide Blob storage/ADLS Gen2 settings manually
☐ Select Blob storage/ADLS Gen2 from your subscriptions

Subscription
Subscription information not needed

Storage account * ⓘ
dimpetsa

Container
☐ Create new ☒ Use existing
 referencedata

Authentication mode
 Managed Identity: System assigned
 Managed Identity authentication for Blob storage/ADLS Gen2 may need additional steps to work. [Learn more](#)

Path pattern * ⓘ
Customer.json

Date format
YYYY/MM/DD

Input details

referencedataareainput

Test Delete

Input alias
referencedataareainput

☒ Provide Blob storage/ADLS Gen2 settings manually
☐ Select Blob storage/ADLS Gen2 from your subscriptions

Subscription
Subscription information not needed

Storage account * ⓘ
dimpetsa

Container
☐ Create new ☒ Use existing
 referencedata

Authentication mode
 Managed Identity: System assigned
 Managed Identity authentication for Blob storage/ADLS Gen2 may need additional steps to work. [Learn more](#)

Path pattern * ⓘ
Area.json

Date format
YYYY/MM/DD

Inputs

4 [See more](#)

ReferenceDataInputAtm Blob storage

ReferenceDataInputCust Blob storage

Outputs
0

Start time * ⓘ
 01/07/2022 9:33:33 PM
 Local time (UTC+02:00)

Duration * ⓘ

Days	Hours	Minutes	Seconds
0	0	2 ✓	0

Sample input 'streaminput' succeeded for Stream Analytics job 'atmtransactionsjob'

Sampled events are available to be downloaded.
[Click to download.](#)

a few seconds ago

Comment: The 4 inputs have been loaded and we are able to download the sample!

BLOB STORAGE OUTPUT

Blob storage/ADLS Gen2

New output

Output alias *

StorageOutput ✓

☐ Provide Blob storage/ADLS Gen2 settings manually

☒ Select Blob storage/ADLS Gen2 from your subscriptions

Subscription

Azure for Students

Storage account *

dimpetsa

Container * ⓘ

☐ Create new ☒ Use existing

referencedata

Authentication mode

Managed Identity: System assigned

Path pattern ⓘ

Date format

YYYY/MM/DD

Time format

HH

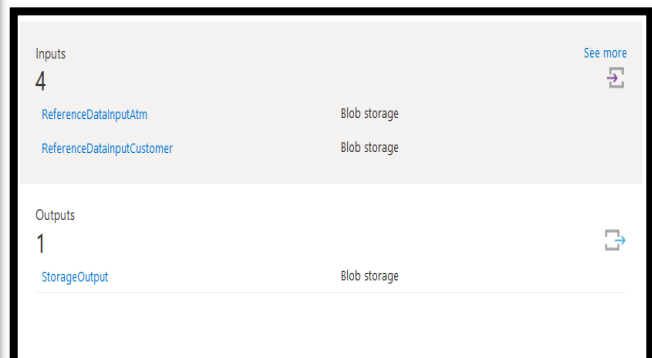
Event serialization format * ⓘ

JSON

Format ⓘ

Save

Comment: Last but not least, we had to create a blob storage output in which we would be able to save the results of our stream queries.



Successful connection test

Connection to output 'StorageOutput' succeeded.

QUERIES

Query 1: Show the total “Amount” of “Type = 0” transactions at “ATM Code = 21” of the last 10 minutes. Repeat as new events keep flowing in (use a sliding window).

Solution:

The screenshot shows the Azure Data Explorer query editor. The query is as follows:

```
1 SELECT sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount, System.Timestamp as Time
2 into StorageOutput
3 from streaminput
4 where CAST([streaminput].[Type] AS BIGINT) = 0 and CAST([streaminput].[ATMCode] AS BIGINT) = 21
5 group by SlidingWindow(minute,10)
```

The results pane shows a single row with the following data:

TotalAmount	Time
81	"2022-01-12T21:08:16.3914887Z"



The screenshot shows the Azure Data Explorer blob storage interface. The left pane shows the 'referencedata' container. The right pane shows a list of blobs, including '0_e4117f86598b4f8d9a8095e96af6c58e_1.json'. The blob details pane shows the blob's metadata, including its name, size, and creation time.

Comment: As we can see from the images above, the query has worked successfully ending up with results from both the sample file and the live data stream. In this specific query, but also in the next ones, we use the “CAST...AS BIGINT” function on a variety of columns that contain numeric data that can become extremely large. Doing that we ensure that the whole number will be visible to us.

Query 2: Show the total “Amount” of “Type = 1” transactions at “ATM Code = 21” of the last hour. Repeat once every hour (use a tumbling window).

Solution:



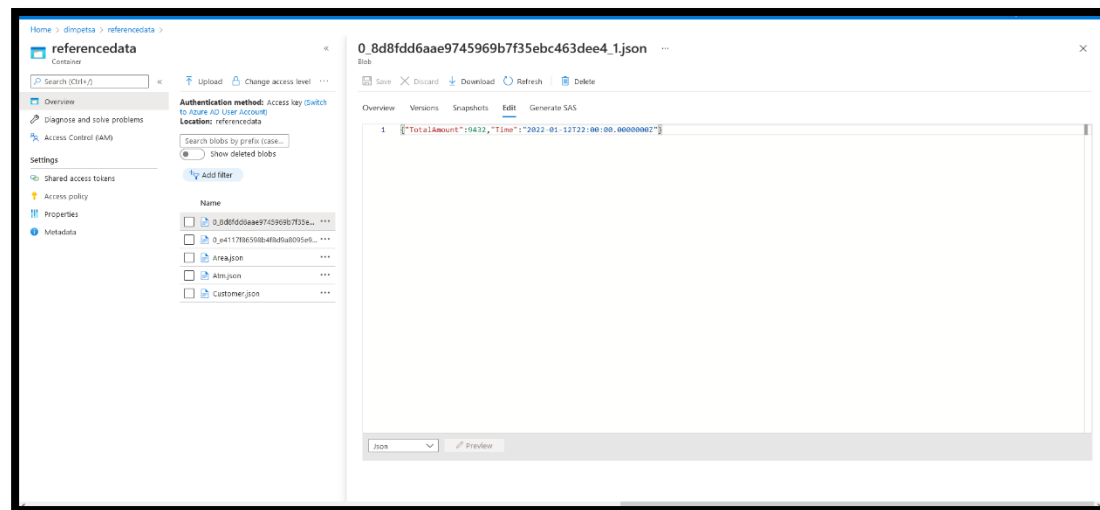
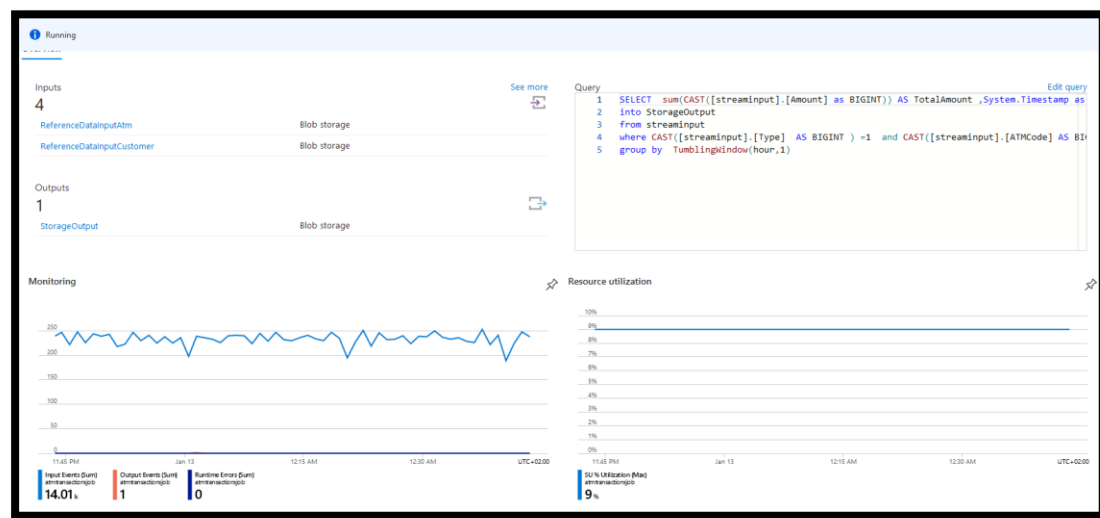
The screenshot shows the Azure Data Studio interface. At the top, a query is written in a text editor:

```
1 SELECT sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount, System.Timestamp as Time
2 into StorageOutput
3 from streaminput
4 where CAST([streaminput].[Type] AS BIGINT) = 1 and CAST([streaminput].[ATMCode] AS BIGINT) = 21
5 group by TumblingWindow(hour,1)
```

Below the query editor, the 'Test results' tab is active, showing a table with two columns: 'TotalAmount' and 'Time'. The table contains one row of data:

TotalAmount	Time
128	"2022-01-12T23:00:00.0000000Z"

At the bottom left, a green status bar indicates 'Success'.



The screenshot shows the Azure Storage Explorer interface. The left pane displays the 'referencedata' container, which contains several blobs. The right pane shows the details of a specific blob named '0.8d8fdd6aae9745969b7f35ebc463dee4_1.json'.

referencedata (Container)

- 0.8d8fdd6aae9745969b7f35ebc463dee4_1.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_2.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_3.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_4.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_5.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_6.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_7.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_8.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_9.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_10.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_11.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_12.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_13.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_14.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_15.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_16.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_17.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_18.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_19.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_20.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_21.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_22.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_23.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_24.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_25.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_26.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_27.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_28.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_29.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_30.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_31.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_32.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_33.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_34.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_35.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_36.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_37.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_38.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_39.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_40.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_41.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_42.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_43.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_44.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_45.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_46.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_47.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_48.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_49.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_50.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_51.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_52.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_53.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_54.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_55.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_56.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_57.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_58.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_59.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_60.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_61.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_62.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_63.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_64.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_65.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_66.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_67.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_68.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_69.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_70.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_71.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_72.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_73.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_74.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_75.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_76.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_77.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_78.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_79.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_80.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_81.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_82.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_83.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_84.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_85.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_86.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_87.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_88.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_89.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_90.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_91.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_92.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_93.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_94.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_95.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_96.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_97.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_98.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_99.json
- 0.8d8fdd6aae9745969b7f35ebc463dee4_100.json

0.8d8fdd6aae9745969b7f35ebc463dee4_1.json (Blob)

Overview Versions Snapshots Edit Generate SAS

1 [{"TotalAmount":9432,"Time":"2022-01-12T23:00:00.0000000Z"}]

Query 3: Show the total “Amount” of “Type = 1” transactions at “ATM Code = 21” of the last hour. Repeat once every 30 minutes (use a hopping window).

Solution:

Test querySave queryDiscard changes

```
1 SELECT sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount ,System.Timestamp as Time
2 into StorageOutput
3 from streaminput
4 where CAST([streaminput].[Type] AS BIGINT ) =1 and CAST([streaminput].[ATMCode] AS BIGINT ) =21
5 group by HoppingWindow(minute,60,30)
6
```

Input previewTest results

Showing 2 rows from 'storageoutput'.

Download results

TotalAmount	Time
53	"2022-01-13T23:30:00.000000002"
53	"2022-01-14T00:00:00.000000002"

SuccessLn 1, Col 1

Running

Inputs4

ReferenceOutputAtmBlob storage

ReferenceDataInputCustomerBlob storage

Outputs1

StorageOutputBlob storage

Monitoring

Resource utilization

Query

1 SELECT sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount ,System.Timestamp as Time
2 into StorageOutput
3 from streaminput
4 where CAST([streaminput].[Type] AS BIGINT) =1 and CAST([streaminput].[ATMCode] AS BIGINT) =21
5 group by HoppingWindow(minute,60,30)
6

Home > dataportal > referencedata

referencedataContainer

Authentication method: Access key Switch to Azure AD User Account

Location: referencedata

Search blobs by prefix (case...)

Show deleted blobs

Add filter

Name

☐ 0_0c710367b2374202b3e5688f5b7d9203_1.json ***

☐ 0_bdd8fdd5a9e745969b7f35e... ***

☐ 0_e4117f86398b4ff5d9a009e5... ***

☐ Area.json ***

☐ Atm.json ***

☐ Customer.json ***

0_0c710367b2374202b3e5688f5b7d9203_1.json

SaveDiscardDownloadRefreshDelete

OverviewVersionsSnapshotsEditGenerate SAS

```
1 [{"TotalAmount":18057,"Time":"2022-01-12T23:00:00.000000002"}]
2 [{"TotalAmount":18385,"Time":"2022-01-12T23:30:00.000000002"}]
```

jsonPreview

Query 4: Show the total “Amount” of “Type = 1” transactions per “ATM Code” of the last one hour (use a sliding window).

Solution:

Test querySave queryDiscard changes

```
1 SELECT CAST([streaminput].[ATMCode] as bigint ) as ATMCode, sum(cast([streaminput].[Amount] as bigint)) as TotalAmount, System.Timestamp as time
2 into StorageOutput
3 from streaminput
4 where cast([streaminput].[type] as bigint) = 1
5 group by cast([streaminput].[ATMCode] as bigint), SlidingWindow(hour,1)
```

Input previewTest results

Showing 11 rows from 'StorageOutput'.

Download results

ATMCode	TotalAmount	time
13	20	"2022-01-13T00:20:02.1761823Z"
20	100	"2022-01-13T00:20:02.1911859Z"
15	205	"2022-01-13T00:20:02.1911859Z"
13	92	"2022-01-13T00:20:02.1911859Z"
18	127	"2022-01-13T00:20:02.1911859Z"
27	27	"2022-01-13T00:20:02.1911859Z"
19	178	"2022-01-13T00:20:02.1911859Z"
10	67	"2022-01-13T00:20:02.1911859Z"
21	43	"2022-01-13T00:20:02.1911859Z"
12	48	"2022-01-13T00:20:02.1911859Z"
13	72	"2022-01-13T01:20:02.1761823Z"

SuccessLn 1, Col 5



Home > dmpetita > referencedata > Container

referencedata

0_225ae86491134793a5bf0ee8b4d8e38c_1.json

Search (Ctrl+F)

UploadChange access level

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Access policy

Metadata

Authentication method: Access key (Switch to Azure AD User Account)

Location: referencedata

Search blobs by prefix (case...)

Show deleted blobs

Add filter

Name
0_0c710167b2374202b3e568...
0_225ae86491134793a5bf0ee...
0_8d8f5d5a9e745949b7f33e...
0_e4117f6039b4f5d9a005e...
Area.json
Atm.json
Customer.json

Overview

Versions

Snapshots

Edit

Generate SAS

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

Icon

Preview

Query 5: Show the total “Amount” of “Type = 1” transactions per “Area Code” of the last hour. Repeat once every hour (use a tumbling window).

Solution:

Test query Save query Discard changes

```

1 SELECT CAST([ReferenceDataInputAtm].[area_code] AS BIGINT) AS AreaCode , sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount ,System.Timestamp as Time
2 into [StorageOutput]
3 from streaminput
4 INNER JOIN [ReferenceDataInputAtm]
5 ON CAST([ReferenceDataInputAtm].[atm_code] AS BIGINT) = CAST([streaminput].[ATMCode] as BIGINT)
6 where CAST([streaminput].[Type] AS BIGINT) = 1
7 group by CAST([ReferenceDataInputAtm].[area_code] AS BIGINT) , TumblingWindow(hour,1)
8

```

Input preview Test results

Showing 10 rows from 'storageoutput'.

AreaCode	TotalAmount	Time
5	142	"2022-01-15T01:00:00.0000000Z"
19	15	"2022-01-15T01:00:00.0000000Z"
1	34	"2022-01-15T01:00:00.0000000Z"
3	26	"2022-01-15T01:00:00.0000000Z"
2	124	"2022-01-15T01:00:00.0000000Z"
7	113	"2022-01-15T01:00:00.0000000Z"
11	133	"2022-01-15T01:00:00.0000000Z"
4	53	"2022-01-15T01:00:00.0000000Z"
10	88	"2022-01-15T01:00:00.0000000Z"
9	90	"2022-01-15T01:00:00.0000000Z"

Success Ln 1, Col 1

Running

Inputs

4

ReferenceDataInputAtm Blob storage

ReferenceDataInputCustomer Blob storage

Outputs

1

StorageOutput Blob storage

Query

```

1 SELECT CAST([ReferenceDataInputAtm].[area_code] AS BIGINT) AS AreaCode , sum(CAST([str
2 into [StorageOutput]
3 from streaminput
4 INNER JOIN [ReferenceDataInputAtm]
5 ON CAST([ReferenceDataInputAtm].[atm_code] AS BIGINT) = CAST([streaminput].[ATMCode] as
6 where CAST([streaminput].[Type] AS BIGINT) = 1
7 group by CAST([ReferenceDataInputAtm].[area_code] AS BIGINT) , TumblingWindow(hour,1)
8

```

Monitoring

Resource utilization

Input Events (Sum) 5.7k

Output Events (Sum) 12

Runtime Errors (Sum) 0

50 % Utilization (Max) 8%

Home > dimpetta > referencedata

referencedata

Container

Search (Ctrl+F)

Authentication method: Access key (Switch to Azure AD User Account)

Location: referencedata

Search blobs by prefix (Case-)

Show deleted blobs

Add filter

Name

- 0_0c710387b374202b3e5888...
- 0_223ae86491134793a35f0ee...
- 0_3b077902d2848f8ab724a6e...
- 0_3ddfd5d5a974590b7f35e...
- 0_3ad9041c3804a6d6d920df...
- 0_3a2e4232c488403695c77a8...
- 0_cdf1e230c34b70be0d423...
- 0_d731079f3ad44962b1cdead...
- 0_e4117f8a39b84f8da095a9...
- Area.json
- Atm.json
- Customer.json

0_d731079f3ad44962b1cdead76234a69b_1.json

Blob

Save Discard Download Refresh Delete

Overview Versions Snapshots Edit Generate SAS

```

1 [{"AreaCode":3,"TotalAmount":164,"Time":"2022-01-15T00:00:00.0000000Z"}]
2 [{"AreaCode":11,"TotalAmount":1081,"Time":"2022-01-15T00:00:00.0000000Z"}]
3 [{"AreaCode":5,"TotalAmount":1823,"Time":"2022-01-15T00:00:00.0000000Z"}]
4 [{"AreaCode":19,"TotalAmount":15,"Time":"2022-01-15T00:00:00.0000000Z"}]
5 [{"AreaCode":19,"TotalAmount":1496,"Time":"2022-01-15T00:00:00.0000000Z"}]
6 [{"AreaCode":12,"TotalAmount":1448,"Time":"2022-01-15T00:00:00.0000000Z"}]
7 [{"AreaCode":7,"TotalAmount":1573,"Time":"2022-01-15T00:00:00.0000000Z"}]
8 [{"AreaCode":20,"TotalAmount":41,"Time":"2022-01-15T00:00:00.0000000Z"}]
9 [{"AreaCode":11,"TotalAmount":1377,"Time":"2022-01-15T00:00:00.0000000Z"}]
10 [{"AreaCode":17,"TotalAmount":54,"Time":"2022-01-15T00:00:00.0000000Z"}]
11 [{"AreaCode":19,"TotalAmount":1299,"Time":"2022-01-15T00:00:00.0000000Z"}]
12 [{"AreaCode":4,"TotalAmount":1166,"Time":"2022-01-15T00:00:00.0000000Z"}]

```

Json Preview

Query 6: Show the total “Amount” per ATM’s “City” and Customer’s “Gender” of the last hour. Repeat once every hour (use a tumbling window).

Solution:

Test query Save query Discard changes

```

1 SELECT [referencedataareainput].[area_city] AS ATMcity ,
2 [ReferenceDataInputCustomer].[gender] AS CustomersGender,
3 sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount ,
4 System.Timestamp as Time
5 Into StorageOutput
6 from streaminput
7 INNER JOIN [ReferenceDataInputCustomer]
8 ON CAST([streaminput].[CardNumber] as BIGINT) = CAST([ReferenceDataInputCustomer].[card_number] as BIGINT)
9 INNER JOIN [ReferenceDataInputAtm]
10 ON CAST([ReferenceDataInputAtm].[atm_code] as BIGINT) = CAST([streaminput].[ATMCode] as BIGINT)
11 INNER JOIN [referencedataareainput]
12 ON CAST([ReferenceDataInputAtm].[area_code] as BIGINT) = CAST([referencedataareainput].[area_code] as BIGINT)
13 group by [ReferenceDataInputCustomer].[gender] , [referencedataareainput].[area_city] , TumblingWindow(hour,1)
14

```

Input preview Test results

Showing 12 rows from 'storageoutput'.

ATMcity	CustomersGender	TotalAmount	Time
"Baltimore"	"Male"	57	"2022-01-13T19:00:00.0000000Z"
"Omaha"	"Male"	61	"2022-01-13T19:00:00.0000000Z"
"Schaumburg"	"Male"	44	"2022-01-13T19:00:00.0000000Z"
"Schaumburg"	"Female"	356	"2022-01-13T19:00:00.0000000Z"
"Memphis"	"Female"	66	"2022-01-13T19:00:00.0000000Z"
"Tacoma"	"Female"	65	"2022-01-13T19:00:00.0000000Z"
"Springfield"	"Male"	234	"2022-01-13T19:00:00.0000000Z"
"Vancouver"	"Male"	17	"2022-01-13T19:00:00.0000000Z"
"Tacoma"	"Male"	30	"2022-01-13T19:00:00.0000000Z"

Success Ln 1, Col 1

Running

Inputs: 4
ReferenceDataInputAtm Blob storage
ReferenceDataInputCustomer Blob storage

Outputs: 1
StorageOutput Blob storage

Query

```

1 SELECT [referencedataareainput].[area_city] AS ATMcity ,
2 [ReferenceDataInputCustomer].[gender] AS CustomersGender,
3 sum(CAST([streaminput].[Amount] as BIGINT)) AS TotalAmount ,
4 System.Timestamp as Time
5 Into StorageOutput
6 from streaminput
7 INNER JOIN [ReferenceDataInputCustomer]
8 ON CAST([streaminput].[CardNumber] as BIGINT) = CAST([ReferenceDataInputCustomer].[ca
9 INNER JOIN [ReferenceDataInputAtm]
10 ON CAST([ReferenceDataInputAtm].[atm_code] as BIGINT) = CAST([streaminput].[ATMCode] as
11 INNER JOIN [referencedataareainput]
12 ON CAST([ReferenceDataInputAtm].[area_code] as BIGINT) = CAST([referencedataareainput]

```

Monitoring

Resource utilization

Input Events (Burst) 26.92
Output Events (Burst) 20
Running Errors (Burst) 0

9% Utilization (Max) 9%

Home > datapeta > referencedata

reference data

Authentication method: Access key [Switch to Azure AD User Account](#)

Location: referencedata

Search blobs by prefix (case...)

Show deleted blobs

Add filter

Properties

Metadata

0_9ad9041c5f804ede8d920dff2c74c12b_1.json

Overview Versions Snapshots Edit Generate SAS

```

1 [{"ATMcity":"Canton","CustomersGender":"Female","TotalAmount":141,"Time":"2022-01-13T18:00:00.0000000Z"}]
2 [{"ATMcity":"Tacoma","CustomersGender":"Male","TotalAmount":288,"Time":"2022-01-13T18:00:00.0000000Z"}]
3 [{"ATMcity":"Schaumburg","CustomersGender":"Female","TotalAmount":59126,"Time":"2022-01-13T18:00:00.0000000Z"}]
4 [{"ATMcity":"Omaha","CustomersGender":"Male","TotalAmount":61438,"Time":"2022-01-13T18:00:00.0000000Z"}]
5 [{"ATMcity":"Vancouver","CustomersGender":"Male","TotalAmount":28978,"Time":"2022-01-13T18:00:00.0000000Z"}]
6 [{"ATMcity":"Springfield","CustomersGender":"Female","TotalAmount":19279,"Time":"2022-01-13T18:00:00.0000000Z"}]
7 [{"ATMcity":"Memphis","CustomersGender":"Male","TotalAmount":39579,"Time":"2022-01-13T18:00:00.0000000Z"}]
8 [{"ATMcity":"Memphis","CustomersGender":"Female","TotalAmount":22312,"Time":"2022-01-13T18:00:00.0000000Z"}]
9 [{"ATMcity":"Schaumburg","CustomersGender":"Male","TotalAmount":18150,"Time":"2022-01-13T18:00:00.0000000Z"}]
10 [{"ATMcity":"Greely","CustomersGender":"Male","TotalAmount":1213,"Time":"2022-01-13T18:00:00.0000000Z"}]
11 [{"ATMcity":"Greely","CustomersGender":"Female","TotalAmount":1213,"Time":"2022-01-13T18:00:00.0000000Z"}]
12 [{"ATMcity":"Baltimore","CustomersGender":"Female","TotalAmount":18379,"Time":"2022-01-13T18:00:00.0000000Z"}]
13 [{"ATMcity":"Tacoma","CustomersGender":"Female","TotalAmount":21893,"Time":"2022-01-13T18:00:00.0000000Z"}]
14 [{"ATMcity":"Dayton","CustomersGender":"Female","TotalAmount":342,"Time":"2022-01-13T18:00:00.0000000Z"}]
15 [{"ATMcity":"Dayton","CustomersGender":"Male","TotalAmount":691,"Time":"2022-01-13T18:00:00.0000000Z"}]
16 [{"ATMcity":"Canton","CustomersGender":"Male","TotalAmount":18853,"Time":"2022-01-13T18:00:00.0000000Z"}]
17 [{"ATMcity":"Springfield","CustomersGender":"Male","TotalAmount":49946,"Time":"2022-01-13T18:00:00.0000000Z"}]
18 [{"ATMcity":"Baltimore","CustomersGender":"Male","TotalAmount":18115,"Time":"2022-01-13T18:00:00.0000000Z"}]
19 [{"ATMcity":"Omaha","CustomersGender":"Female","TotalAmount":16791,"Time":"2022-01-13T18:00:00.0000000Z"}]
20 [{"ATMcity":"Vancouver","CustomersGender":"Female","TotalAmount":282,"Time":"2022-01-13T18:00:00.0000000Z"}]

```

json Preview

Query 7: Alert (Do a simple SELECT “1”) if a Customer has performed two transactions of “Type = 1” in a window of an hour (use a sliding window).

Solution: << The columns that contain information about the customer are not necessary but they are helpful for the viewer >>

Test query ☐ Save query ☒ Discard changes

```
1 select 1 AS ALERT ,
2 CAST([streaminput].[CardNumber] as BIGINT) as CardNumber,
3 [ReferenceDataInputCustomer].[last_name] as LastName,
4 [ReferenceDataInputCustomer].[first_name] as FirstName,
5 count(*) as NumberOfTransact,
6 System.Timestamp as Time
7 into StorageOutput
8 from streaminput
9 inner join [ReferenceDataInputCustomer]
10 on CAST([ReferenceDataInputCustomer].[card_number] as BIGINT) = CAST([streaminput].[CardNumber] as BIGINT)
11 where CAST([streaminput].[Type] AS BIGINT) = 1
12 group by [ReferenceDataInputCustomer].[last_name] ,
13 [ReferenceDataInputCustomer].[first_name] ,
14 CAST([streaminput].[CardNumber] as BIGINT),slidingwindow(hour,1)
15 having NumberOfTransact = 2
```

Input preview Test results

Showing 10 rows from 'storageoutput'.

ALERT	CardNumber	LastName	FirstName	NumberOfTransact	Time
1	560222217915598000	"Carroll"	"Brenda"	2	"2022-01-14T10:04:53.2430494Z"
1	3583257214000023	"Sims"	"Ruth"	2	"2022-01-14T10:04:53.2430494Z"
1	353463361736454	"Moreno"	"Angela"	2	"2022-01-14T10:04:53.2430494Z"
1	5893112367133403000	"Cooper"	"Arthur"	2	"2022-01-14T10:04:53.2430494Z"
1	560224675568900	"Morrison"	"Bruce"	2	"2022-01-14T10:04:53.2430494Z"
1	4026567514157759	"Perry"	"Emily"	2	"2022-01-14T10:04:53.2430494Z"
1	353576637597043	"Day"	"Martha"	2	"2022-01-14T10:04:53.2430494Z"
1	3554025590595485	"Stone"	"Walter"	2	"2022-01-14T10:04:53.2430494Z"

Success

Running

Inputs

4

ReferenceDataInputAtm Blob storage

ReferenceDataInputCustomer Blob storage

Outputs

1

StorageOutput Blob storage

Query

```
1 select 1 AS ALERT ,
2 CAST([streaminput].[CardNumber] as BIGINT) as CardNumber,
3 [ReferenceDataInputCustomer].[last_name] as LastName,
4 [ReferenceDataInputCustomer].[first_name] as FirstName,
5 count(*) as NumberOfTransact,
6 System.Timestamp as Time
7 into StorageOutput
8 from streaminput
9 inner join [ReferenceDataInputCustomer]
10 on CAST([ReferenceDataInputCustomer].[card_number] as BIGINT) = CAST([streaminput].[CardNumber] as BIGINT)
11 where CAST([streaminput].[Type] AS BIGINT) = 1
12 group by [ReferenceDataInputCustomer].[last_name] ,
13 [ReferenceDataInputCustomer].[first_name] ,
14 CAST([streaminput].[CardNumber] as BIGINT),slidingwindow(hour,1)
15 having NumberOfTransact = 2
```

Monitoring

Resource utilization

Input Events (Burst) 13.63%

Output Events (Burst) 0%

Runtime Errors (Burst) 0%

10% Utilization (Max) 10%

Home > dmpgta > referencedata

referencedata

Container

Search (Ctrl-F)

Upload Change access level

Authentication method: Access key Switch to Azure AD User Account

Location: referencedata

Search blobs by prefix (case...)

Show deleted blobs

Add filter

Shared access tokens

Access policy

Properties

Metadata

0_5bcbff7902d2848f3ab724a66dab7bd56_1.json

Overview Versions Snapshots Edit Generate SAS

```
1 [{"ALERT":1,"CardNumber":3554825590595485,"LastName":"Stone","FirstName":"Walter","NumberOfTransact":2,"Time":"2022-01-14T07:13:53.9270000Z"}]
2 [{"ALERT":1,"CardNumber":4026567514157759,"LastName":"Perry","FirstName":"Emily","NumberOfTransact":2,"Time":"2022-01-14T07:13:58.9590000Z"}]
3 [{"ALERT":1,"CardNumber":3549824987623749,"LastName":"Bradley","FirstName":"Jesse","NumberOfTransact":2,"Time":"2022-01-14T07:13:58.9590000Z"}]
4 [{"ALERT":1,"CardNumber":353463361736454,"LastName":"Moreno","FirstName":"Angela","NumberOfTransact":2,"Time":"2022-01-14T07:14:05.9610000Z"}]
5 [{"ALERT":1,"CardNumber":3549870931669297,"LastName":"Snyder","FirstName":"Jose","NumberOfTransact":2,"Time":"2022-01-14T07:14:05.9620000Z"}]
```

json Preview

Query 8: Alert (Do a simple SELECT “1”) if the “Area Code” of the ATM of the transaction is not the same as the “Area Code” of the “Card Number” (Customer’s Area Code) - (use a sliding window)

Solution: << The columns that contain information about the customer are not necessary but they are helpful for the viewer>>

Test query Save query Discard changes

```

1 SELECT 1 AS ALERT,
2 count(*) as Number_of_Different_CustArea_and_AtmArea,
3 cast([ReferenceDataInputCustomer].[area_code] as bigint ) as CustomerArea,
4 cast([ReferenceDataInputAtm].[area_code] as bigint ) as AtmArea, System.Timestamp as Time
5 into [StorageOutput]
6 from streaminput
7 inner join [ReferenceDataInputCustomer]
8 on cast([streaminput].[CardNumber] as bigint) = cast([ReferenceDataInputCustomer].[card_number] as bigint)
9 inner join [ReferenceDataInputAtm]
10 on cast([streaminput].[ATMCode] as bigint) = cast([ReferenceDataInputAtm].[atm_code] as bigint)
11 where cast([ReferenceDataInputAtm].[area_code] as bigint) <> cast([ReferenceDataInputCustomer].[area_code] as bigint)
12 group by cast([ReferenceDataInputCustomer].[area_code] as bigint), cast([ReferenceDataInputAtm].[area_code] as bigint), SlidingWindow(hour,1)

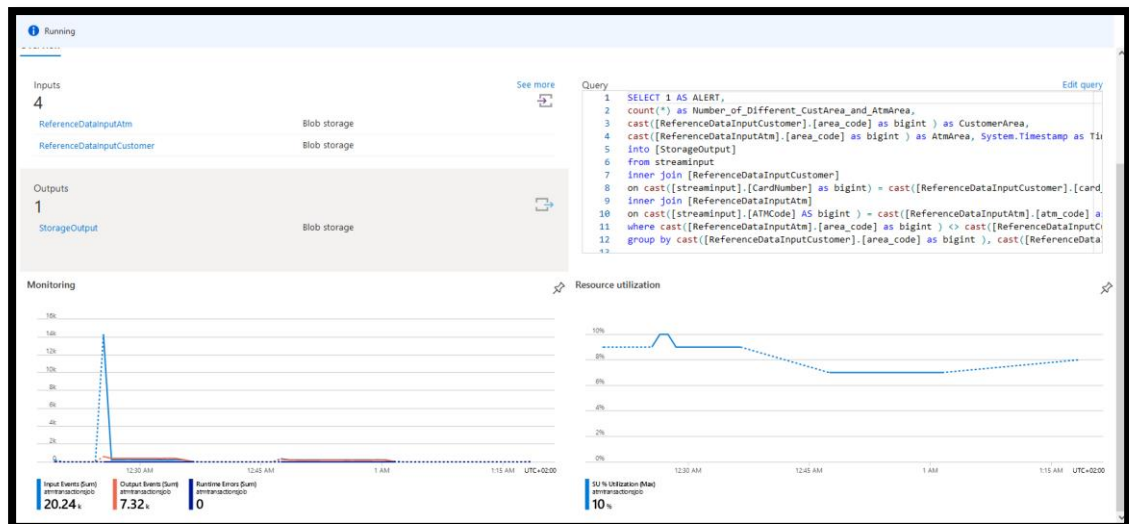
```

input preview Test results

Showing 11 rows from 'storageoutput'.

ALERT	Number_of_Different_CustArea_and_AtmArea	CustomerArea	AtmArea	Time
1	1	2	4	"2022-01-13T23:21:11.8102455Z"
1	11	7	5	"2022-01-13T23:21:11.8302056Z"
1	3	3	7	"2022-01-13T23:21:11.8302056Z"
1	7	2	4	"2022-01-13T23:21:11.8302056Z"
1	6	1	2	"2022-01-13T23:21:11.8302056Z"
1	7	8	11	"2022-01-13T23:21:11.8302056Z"
1	1	4	3	"2022-01-13T23:21:11.8302056Z"
1	3	6	1	"2022-01-13T23:21:11.8302056Z"
1	1	1	2	"2022-01-13T23:21:11.8302056Z"
1	1	10	9	"2022-01-13T23:21:11.8302056Z"

Success



Home > dimpetta > referencedata

referencedata

Container

Search (Ctrl+F)

Upload Change access level

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Access policy

Properties

Metadata

Authentication method: Access key [Switch to Azure AD User Account](#)

Location: referencedata

Show deleted blobs

Add filter

Name
0_0c710367b2374202b3e5688...
0_225a0649113479a3a5f5e...
0_6d8f56aaw974599b7f35e...
0_3a29041c9504e6d8d20df...
0_3b244282488483695c77a...
0_cdf1e8230034b78be0423...
0_e411786858b48da805a5e...
Area.json
Atm.json
Customer.json

0_b82e4282e488483695c77a94f872451_1.json

Blob

Save Discard Download Refresh Delete

Overview Versions Snapshots Edit Generate SAS

```

1 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":690,"CustomerArea":16,"AtmArea":1,"Time":"2022-01-13T23:17:04.7300000Z"]
2 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2562,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:04.7510000Z"]
3 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":647,"CustomerArea":10,"AtmArea":9,"Time":"2022-01-13T23:17:04.7510000Z"]
4 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1937,"CustomerArea":12,"AtmArea":4,"Time":"2022-01-13T23:17:04.7500000Z"]
5 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2563,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:05.0540000Z"]
6 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1877,"CustomerArea":11,"AtmArea":12,"Time":"2022-01-13T23:17:05.0540000Z"]
7 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":679,"CustomerArea":14,"AtmArea":13,"Time":"2022-01-13T23:17:05.0500000Z"]
8 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2562,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:05.7200000Z"]
9 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2831,"CustomerArea":18,"AtmArea":11,"Time":"2022-01-13T23:17:05.7670000Z"]
10 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2563,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:05.7630000Z"]
11 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2832,"CustomerArea":18,"AtmArea":11,"Time":"2022-01-13T23:17:05.6540000Z"]
12 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1928,"CustomerArea":12,"AtmArea":4,"Time":"2022-01-13T23:17:05.0700000Z"]
13 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2562,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:05.0500000Z"]
14 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":16,"CustomerArea":11,"AtmArea":8,"Time":"2022-01-13T23:17:06.7670000Z"]
15 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":669,"CustomerArea":14,"AtmArea":13,"Time":"2022-01-13T23:17:06.7900000Z"]
16 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":566,"CustomerArea":10,"AtmArea":9,"Time":"2022-01-13T23:17:06.7900000Z"]
17 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1876,"CustomerArea":11,"AtmArea":12,"Time":"2022-01-13T23:17:06.8140000Z"]
18 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2563,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:07.0540000Z"]
19 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1887,"CustomerArea":11,"AtmArea":12,"Time":"2022-01-13T23:17:07.0500000Z"]
20 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2564,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:07.1010000Z"]
21 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":691,"CustomerArea":16,"AtmArea":1,"Time":"2022-01-13T23:17:07.1320000Z"]
22 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2563,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:07.7520000Z"]
23 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2831,"CustomerArea":18,"AtmArea":11,"Time":"2022-01-13T23:17:07.7630000Z"]
24 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2830,"CustomerArea":18,"AtmArea":11,"Time":"2022-01-13T23:17:07.7990000Z"]
25 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1875,"CustomerArea":11,"AtmArea":12,"Time":"2022-01-13T23:17:07.7990000Z"]
26 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":2564,"CustomerArea":17,"AtmArea":5,"Time":"2022-01-13T23:17:08.0700000Z"]
27 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":700,"CustomerArea":16,"AtmArea":10,"Time":"2022-01-13T23:17:08.1010000Z"]
28 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":692,"CustomerArea":16,"AtmArea":11,"Time":"2022-01-13T23:17:08.1320000Z"]
29 ["ALERT":1,"Number_of_Different_CustArea_and_AtmArea":1876,"CustomerArea":11,"AtmArea":12,"Time":"2022-01-13T23:17:08.1320000Z"]

```


The End