

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
DEPARTMENT OF MANAGEMENT, SCIENCE & TECHNOLOGY
MSc BUSINESS ANALYTICS

“Data Streams”

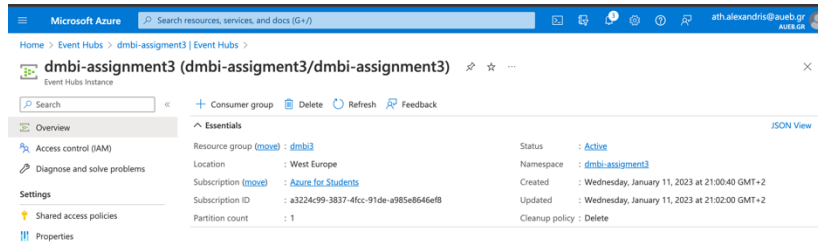
Full Name: ATHANASIOS ALEXANDRIS
Register Number: p2822202
&
Full Name: KRISTJANA KOLAJ
Register Number: p2822213

Table of Contents

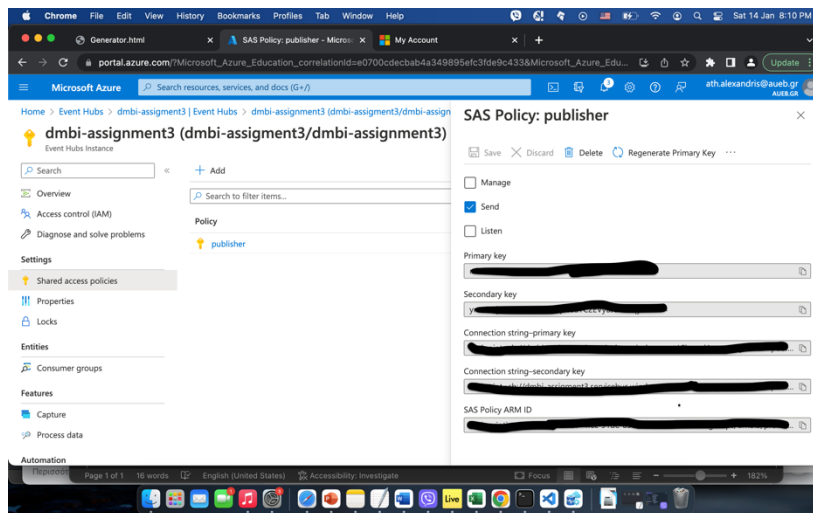
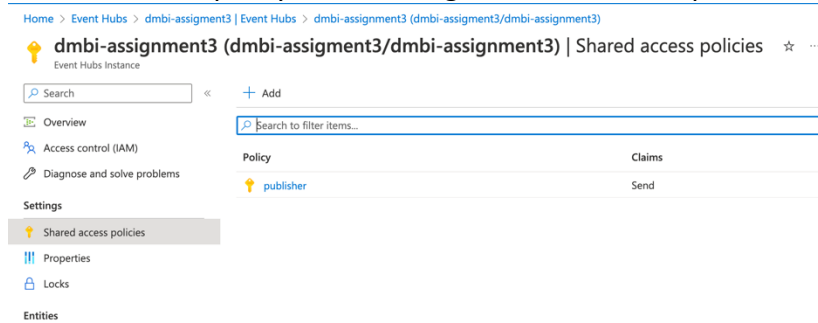
1. Setting up azure account process.....	page 3
2. Queries and outputs.....	page 8

1.Setting up azure account process

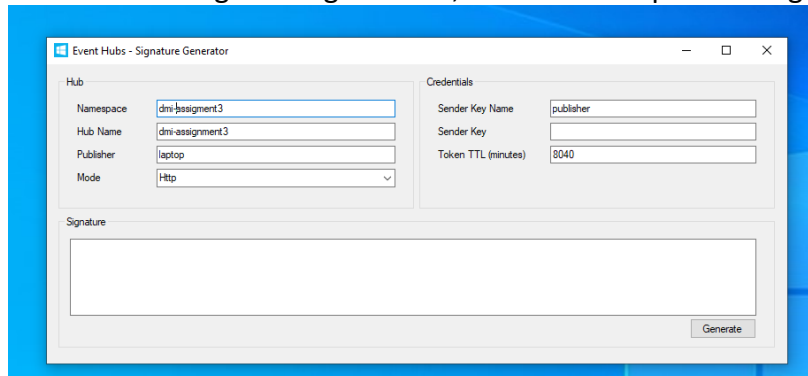
1. We created an event on 'Event Hubs'.



2.We created a policy from Settings→Shared access policies



3. We use the signature generator, in order to acquire the signature for the 'Generator.html'



4. We updated the 'Generator.html' with our credentials and keys.

Generator.html x | Area.json | Atm.json | Customer.json

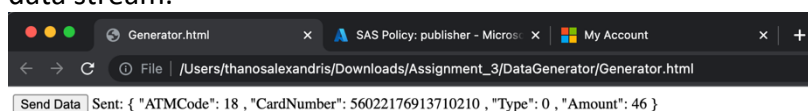
```

1 <html>
2 <head>
3   <script src="js/lodash.js"></script>
4 </head>
5 <body>
6   <input type="button" value="Send Data" onclick="sendDummyData()" />
7   <div id="status" style="display: inline-block;"></div>
8   <script type="text/javascript">
9     function sendDummyData() {
10
11       /*****
12       /*** CONFIG ****/
13       /*****/
14
15       //Use the signature generator: https://github.com/sandrinodimattia/RedDog/releases
16       var sas = '
17
18       var serviceNamespace = "dmbi-assignment3";
19       var hubName = "dmbi-assignment3";
20       var deviceName = "laptop";
21
22       /*****
23       /*** GENERATOR ****/
24       /*****/
25       var atms = [{ "atm_code":1,"area_code":20}, {"atm_code":2,"area_code":17}, {"atm_code":3,"area_code":18},
26         {"atm_code":4,"area_code":19}, {"atm_code":5,"atm_code":16}, {"atm_code":16,"area_code":12}, {"atm_code":7,
27         "area_code":14}, {"atm_code":8,"area_code":13}, {"atm_code":9,"area_code":15}, {"atm_code":10,
28         "area_code":11}, {"atm_code":11,"area_code":11, "area_code":12, "atm_code":12, "area_code":9, "atm_code":13,
29         "area_code":10}, {"atm_code":14,"area_code":16}, {"atm_code":15,"area_code":5}, {"atm_code":16,
30         "area_code":17}, {"atm_code":17,"area_code":3}, {"atm_code":18,"area_code":4}, {"atm_code":19,"area_code":2},
31         {"atm_code":20,"area_code":1}, {"atm_code":21,"area_code":11}, {"atm_code":22,"area_code":6}];
32
33       var customers = [{"card_number":5446210381593272, "first_name":"Eugene", "last_name":"Mason", "age":67,
34         "deviceName": 'laptop' }];
35     }
36   }
37 </script>
38 </body>
39 </html>

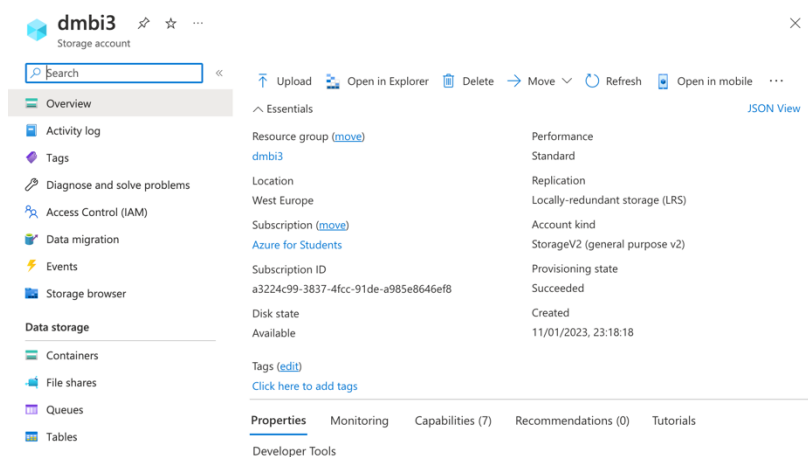
```

Ln 18, Col 29 Spaces: 2 UTF-8 CRLF HTML

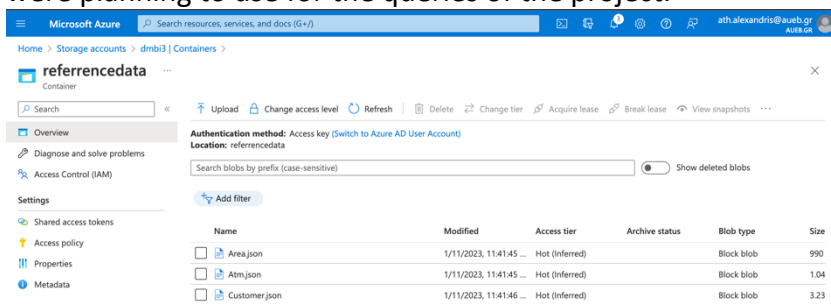
5. We opened the Generator.html, and we pressed send data to begin the process of the data stream.



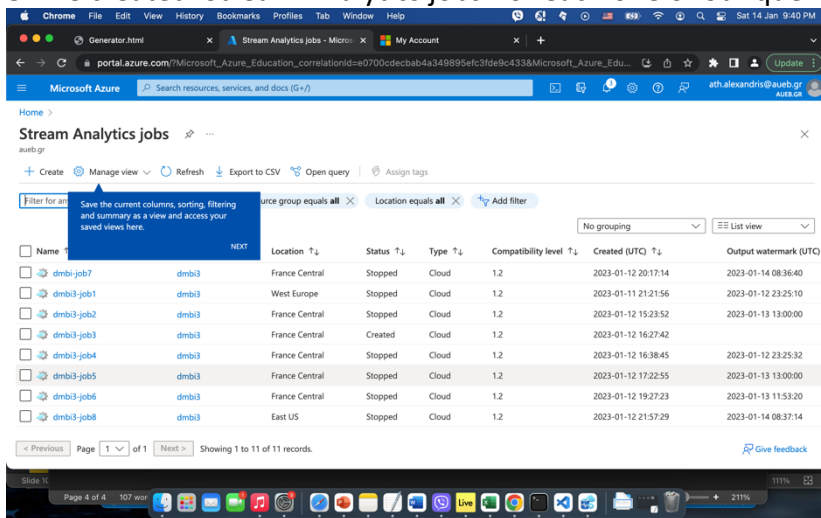
6. We created a storage account, where our stream outputs will be placed.



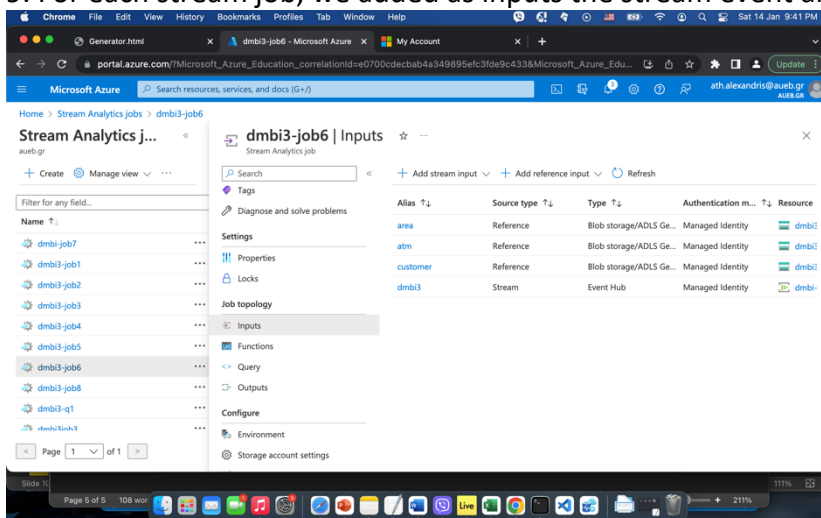
7. We created a container named 'referencedata', where we stored the .json files that we were planning to use for the queries of the project.



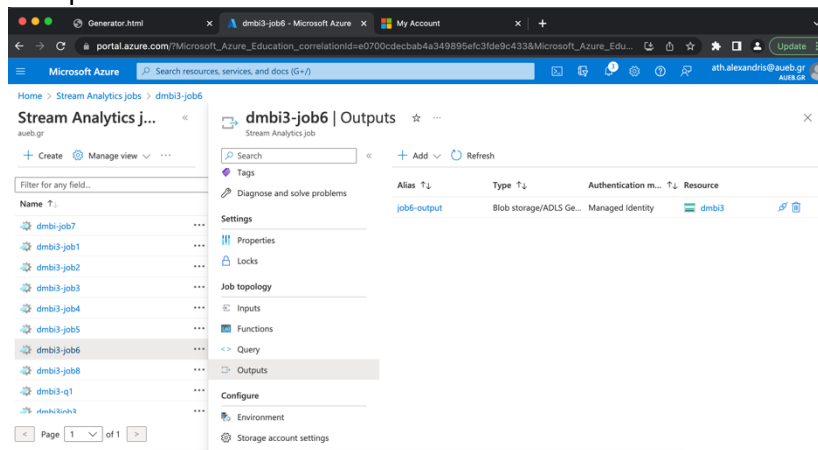
8. We created 'Stream Analytics jobs' for each one of our queries.



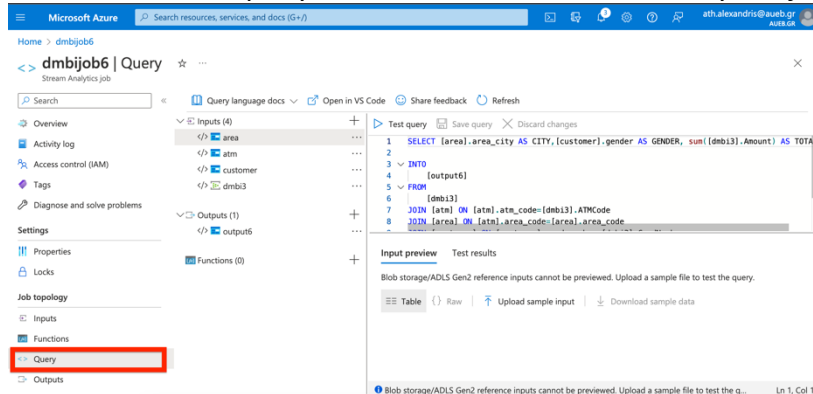
9. For each stream job, we added as inputs the stream event and the reference files



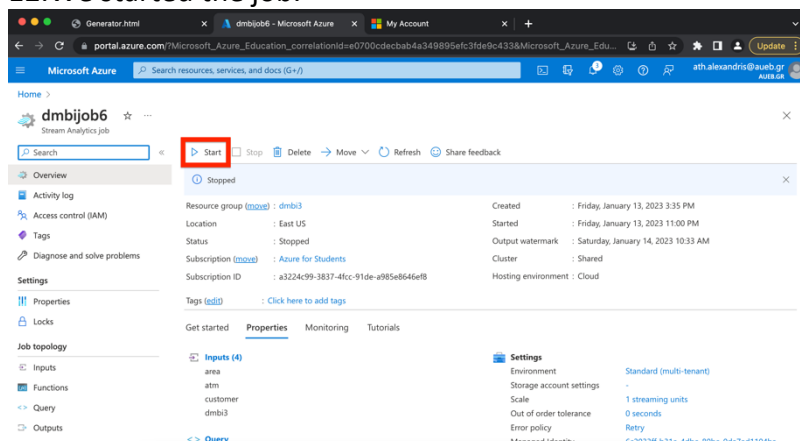
10. For each stream job, we created blob outputs that were stored in our storage accounts in separate containers.



11. We wrote the query that we want to be executed by the job.



12. We started the job.



13. The outcomes of each job are stored as .json files, in our storage account in separate containers

The screenshot shows the Microsoft Azure portal interface. The breadcrumb navigation at the top left is 'Home > dmbi3 | Containers', which is highlighted with a red box. Below this, the 'job3output' container is selected. The left sidebar contains navigation links: Overview, Diagnose and solve problems, Access Control (IAM), Settings, Shared access tokens, Access policy, Properties, and Metadata. The main content area shows the container's overview, including the authentication method (Access key), location (job3output), and a search bar. A table of blobs is displayed, with one blob highlighted by a red box:

Name	Modified	Access tier	Archive status	Blob type	Size
0_86808a1e7ee64b7b88ae38352c46c789_1.json	1/13/2023, 2:00:01 PM	Hot (inferred)		Block blob	508

2.Queries and outputs:

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).

Query

Inputs (1)

</> input3

Outputs (1)

</> job1output

Functions (0)

Test query

Save query

Discard changes

```
1 SELECT
2     SUM([Amount]) AS TOTAL
3 INTO
4     [job1output]
5 FROM
6     [input3]
7 WHERE [ATMCode]=21 AND [Type]=0
8 GROUP BY SlidingWindow(minute,10)
```

Output

job1.json 1 x

Users > thanosalexandris > Downloads > azure dmbi3 job outputs > job1.json > ...

```
1 {"TOTAL":609.0}
2 {"TOTAL":566.0}
3 {"TOTAL":584.0}
4 {"TOTAL":567.0}
5 {"TOTAL":524.0}
6 {"TOTAL":478.0}
7 {"TOTAL":457.0}
8 {"TOTAL":441.0}
9 {"TOTAL":422.0}
10 {"TOTAL":409.0}
11 {"TOTAL":377.0}
12 {"TOTAL":350.0}
13 {"TOTAL":315.0}
14 {"TOTAL":301.0}
15 {"TOTAL":253.0}
16 {"TOTAL":234.0}
17 {"TOTAL":218.0}
18 {"TOTAL":196.0}
19 {"TOTAL":146.0}
20 {"TOTAL":98.0}
21 {"TOTAL":82.0}
22 {"TOTAL":52.0}
23 {"TOTAL":18.0}
24 {"TOTAL":34.0}
25 {"TOTAL":76.0}
26 {"TOTAL":118.0}
27 {"TOTAL":157.0}
28 {"TOTAL":177.0}
29 {"TOTAL":214.0}
30 {"TOTAL":240.0}
31 {"TOTAL":251.0}
```

Ln 1, Col 1 Spaces: 4 UTF-8 CRLF JSON Change SQL language provider

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).

Query

Inputs (1)

</> dmbi3

Outputs (1)

</> job2output

Functions (0)

Test query

Save query

Discard changes

```
1 SELECT sum([Amount]) AS TOTAL
2
3 INTO
4     [job2output]
5 FROM
6     [dmbi3]
7 WHERE [ATMCode]=21 AND [Type]=1
8 GROUP BY TumblingWindow(hour,1)
```


Output

```

1 {"TOTAL":582.0}
2 {"TOTAL":54.0}
3 {"TOTAL":248.0}
4 {"TOTAL":407.0}
5 {"TOTAL":58.0}
6 {"TOTAL":1416.0}
7 {"TOTAL":270.0}
8 {"TOTAL":11.0}
9 {"TOTAL":107.0}
10 {"TOTAL":80.0}
11 {"TOTAL":26.0}
12 {"TOTAL":70.0}
13 {"TOTAL":21.0}
14 {"TOTAL":16.0}
15 {"TOTAL":34.0}
16 {"TOTAL":83.0}

```

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).

Query

```

1 SELECT sum([Amount]) AS TOTAL
2
3 INTO
4 [job3output]
5 FROM
6 [dmbi3]
7 WHERE [ATMCode]=21 AND [Type]=1
8 GROUP BY HoppingWindow(minute,60,30)

```

Output

```

1 {"TOTAL":15.0}
2 {"TOTAL":248.0}
3 {"TOTAL":248.0}
4 {"TOTAL":407.0}
5 {"TOTAL":421.0}
6 {"TOTAL":58.0}
7 {"TOTAL":756.0}
8 {"TOTAL":1416.0}
9 {"TOTAL":704.0}
10 {"TOTAL":270.0}
11 {"TOTAL":270.0}
12 {"TOTAL":11.0}
13 {"TOTAL":11.0}
14 {"TOTAL":50.0}
15 {"TOTAL":107.0}
16 {"TOTAL":57.0}
17 {"TOTAL":80.0}
18 {"TOTAL":106.0}
19 {"TOTAL":26.0}
20 {"TOTAL":70.0}
21 {"TOTAL":70.0}
22 {"TOTAL":21.0}
23 {"TOTAL":21.0}
24 {"TOTAL":16.0}
25 {"TOTAL":16.0}
26 {"TOTAL":10.0}
27 {"TOTAL":34.0}
28 {"TOTAL":24.0}
29 {"TOTAL":83.0}

```

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

Query

Inputs (1)

</> dmbi3

...

Outputs (1)

</> job4output

...

Functions (0)

+

Test query Save query Discard changes

```

1 SELECT [ATMCode],SUM([Amount]) AS TOTAL
2
3 INTO
4 [job4output]
5 FROM
6 [dmbi3]
7 WHERE [Type]=1
8 GROUP BY [ATMCode], SlidingWindow(hour,1)

```

Output

job4.json

Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More

job1.json 1 job2.json 1 job3.json 1 job4.json 1

...

Users > thanosalexandris > Downloads > azure dmbi3 job outputs > job4.json > ...

```

1 {"ATMCode":10,"TOTAL":204.0}
2 {"ATMCode":13,"TOTAL":223.0}
3 {"ATMCode":18,"TOTAL":210.0}
4 {"ATMCode":19,"TOTAL":263.0}
5 {"ATMCode":18,"TOTAL":183.0}
6 {"ATMCode":18,"TOTAL":211.0}
7 {"ATMCode":20,"TOTAL":114.0}
8 {"ATMCode":12,"TOTAL":85.0}
9 {"ATMCode":18,"TOTAL":174.0}
10 {"ATMCode":12,"TOTAL":69.0}
11 {"ATMCode":19,"TOTAL":218.0}
12 {"ATMCode":16,"TOTAL":33.0}
13 {"ATMCode":18,"TOTAL":156.0}
14 {"ATMCode":15,"TOTAL":124.0}
15 {"ATMCode":19,"TOTAL":176.0}
16 {"ATMCode":10,"TOTAL":243.0}
17 {"ATMCode":15,"TOTAL":165.0}
18 {"ATMCode":13,"TOTAL":178.0}
19 {"ATMCode":21,"TOTAL":15.0}
20 {"ATMCode":10,"TOTAL":217.0}
21 {"ATMCode":20,"TOTAL":69.0}
22 {"ATMCode":18,"TOTAL":198.0}
23 {"ATMCode":13,"TOTAL":146.0}
24 {"ATMCode":19,"TOTAL":221.0}
25 {"ATMCode":16,"TOTAL":69.0}
26 {"ATMCode":13,"TOTAL":112.0}
27 {"ATMCode":20,"TOTAL":40.0}
28 {"ATMCode":18,"TOTAL":217.0}
29 {"ATMCode":19,"TOTAL":183.0}
30 {"ATMCode":19,"TOTAL":197.0}
31 {"ATMCode":19,"TOTAL":230.0}

```

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).

Query

Inputs (3)

</> dmbi3

reference-area

</> reference-atm

Outputs (1)

</> job5-output

...

Test query Save query Discard changes

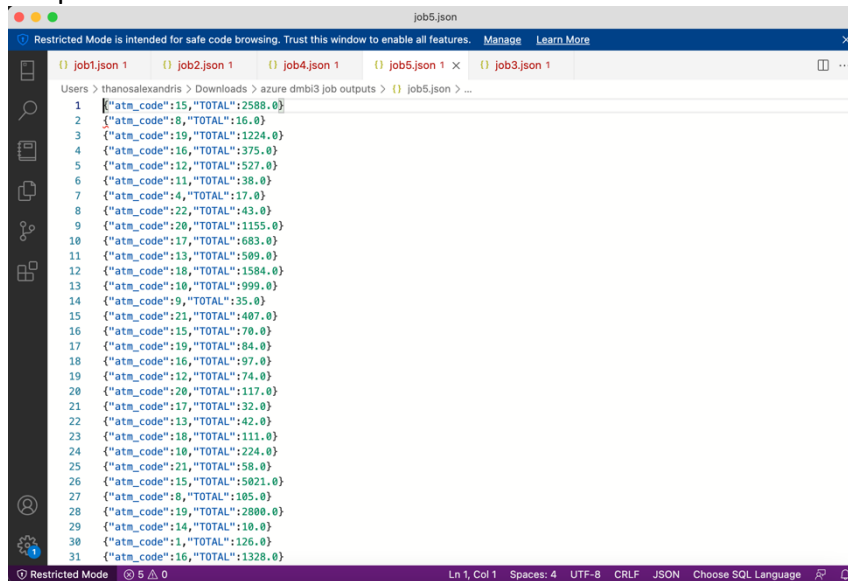
```

1 SELECT [reference-atm].atm_code, sum([dmbi3].Amount) AS TOTAL
2
3 INTO
4 [job5-output]
5 FROM
6 [dmbi3]
7 JOIN [reference-atm] ON [reference-atm].atm_code=[dmbi3].ATMCode
8 WHERE [dmbi3].Type=1
9 GROUP BY [reference-atm].atm_code, TumblingWindow(hour,1)

```

10

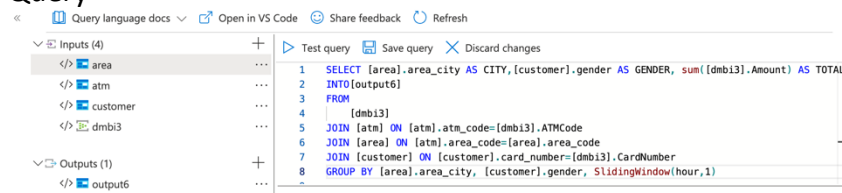
Output



```
1 [{"atm_code":15,"TOTAL":2588.0}]
2 [{"atm_code":8,"TOTAL":16.0}]
3 [{"atm_code":19,"TOTAL":1224.0}]
4 [{"atm_code":16,"TOTAL":375.0}]
5 [{"atm_code":12,"TOTAL":527.0}]
6 [{"atm_code":11,"TOTAL":38.0}]
7 [{"atm_code":4,"TOTAL":17.0}]
8 [{"atm_code":22,"TOTAL":43.0}]
9 [{"atm_code":20,"TOTAL":1155.0}]
10 [{"atm_code":17,"TOTAL":683.0}]
11 [{"atm_code":13,"TOTAL":589.0}]
12 [{"atm_code":16,"TOTAL":1584.0}]
13 [{"atm_code":10,"TOTAL":999.0}]
14 [{"atm_code":9,"TOTAL":35.0}]
15 [{"atm_code":21,"TOTAL":407.0}]
16 [{"atm_code":15,"TOTAL":70.0}]
17 [{"atm_code":19,"TOTAL":84.0}]
18 [{"atm_code":16,"TOTAL":97.0}]
19 [{"atm_code":12,"TOTAL":74.0}]
20 [{"atm_code":20,"TOTAL":117.0}]
21 [{"atm_code":17,"TOTAL":32.0}]
22 [{"atm_code":13,"TOTAL":42.0}]
23 [{"atm_code":18,"TOTAL":111.0}]
24 [{"atm_code":10,"TOTAL":224.0}]
25 [{"atm_code":21,"TOTAL":58.0}]
26 [{"atm_code":15,"TOTAL":5021.0}]
27 [{"atm_code":8,"TOTAL":105.0}]
28 [{"atm_code":19,"TOTAL":2800.0}]
29 [{"atm_code":14,"TOTAL":10.0}]
30 [{"atm_code":11,"TOTAL":126.0}]
31 [{"atm_code":16,"TOTAL":1328.0}]
```

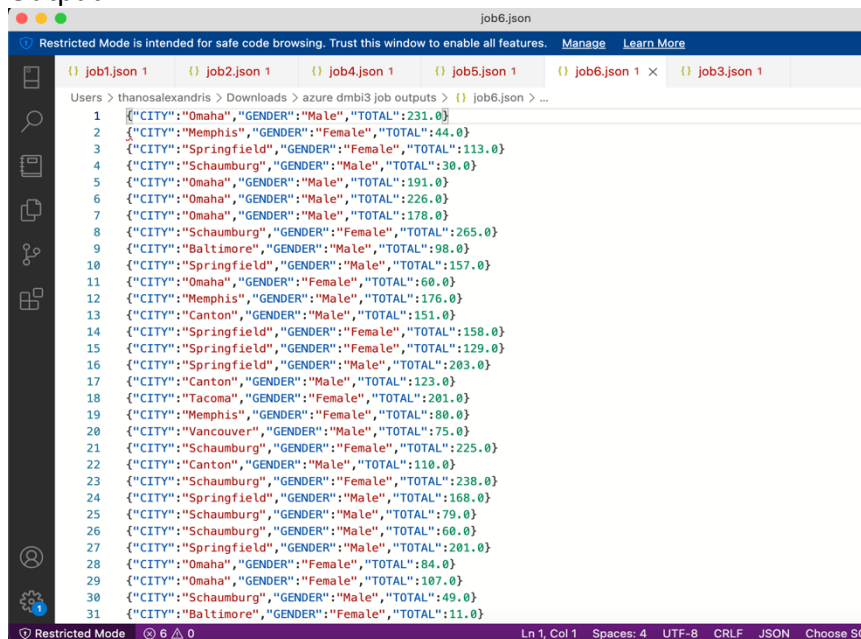
Query 6: Show the total “Amount” per ATM’s “City” and Customer’s “Gender” of the last hour. Repeat once every hour (use a tumbling window).

Query



```
1 SELECT [area].area_city AS CITY, [customer].gender AS GENDER, sum([dmbi3].Amount) AS TOTAL
2 INTO [output6]
3 FROM
4 [dmbi3]
5 JOIN [atm] ON [atm].atm_code=[dmbi3].ATMCode
6 JOIN [area] ON [area].area_code=[atm].area_code
7 JOIN [customer] ON [customer].card_number=[dmbi3].CardNumber
8 GROUP BY [area].area_city, [customer].gender, SlidingWindow(hour,1)
```

Output



```
1 [{"CITY":"Omaha","GENDER":"Male","TOTAL":231.0}]
2 [{"CITY":"Memphis","GENDER":"Female","TOTAL":44.0}]
3 [{"CITY":"Springfield","GENDER":"Female","TOTAL":113.0}]
4 [{"CITY":"Schaumburg","GENDER":"Male","TOTAL":30.0}]
5 [{"CITY":"Omaha","GENDER":"Male","TOTAL":191.0}]
6 [{"CITY":"Omaha","GENDER":"Male","TOTAL":226.0}]
7 [{"CITY":"Omaha","GENDER":"Male","TOTAL":178.0}]
8 [{"CITY":"Schaumburg","GENDER":"Female","TOTAL":265.0}]
9 [{"CITY":"Baltimore","GENDER":"Male","TOTAL":98.0}]
10 [{"CITY":"Springfield","GENDER":"Male","TOTAL":157.0}]
11 [{"CITY":"Omaha","GENDER":"Female","TOTAL":60.0}]
12 [{"CITY":"Memphis","GENDER":"Male","TOTAL":176.0}]
13 [{"CITY":"Canton","GENDER":"Male","TOTAL":151.0}]
14 [{"CITY":"Springfield","GENDER":"Female","TOTAL":158.0}]
15 [{"CITY":"Springfield","GENDER":"Female","TOTAL":129.0}]
16 [{"CITY":"Springfield","GENDER":"Male","TOTAL":203.0}]
17 [{"CITY":"Canton","GENDER":"Male","TOTAL":123.0}]
18 [{"CITY":"Tacoma","GENDER":"Female","TOTAL":201.0}]
19 [{"CITY":"Memphis","GENDER":"Female","TOTAL":80.0}]
20 [{"CITY":"Vancouver","GENDER":"Male","TOTAL":75.0}]
21 [{"CITY":"Schaumburg","GENDER":"Female","TOTAL":225.0}]
22 [{"CITY":"Canton","GENDER":"Male","TOTAL":110.0}]
23 [{"CITY":"Schaumburg","GENDER":"Female","TOTAL":238.0}]
24 [{"CITY":"Springfield","GENDER":"Male","TOTAL":168.0}]
25 [{"CITY":"Schaumburg","GENDER":"Male","TOTAL":79.0}]
26 [{"CITY":"Schaumburg","GENDER":"Male","TOTAL":60.0}]
27 [{"CITY":"Springfield","GENDER":"Male","TOTAL":201.0}]
28 [{"CITY":"Omaha","GENDER":"Female","TOTAL":84.0}]
29 [{"CITY":"Omaha","GENDER":"Female","TOTAL":107.0}]
30 [{"CITY":"Schaumburg","GENDER":"Male","TOTAL":49.0}]
31 [{"CITY":"Baltimore","GENDER":"Female","TOTAL":11.0}]
```

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).

Query

Query language docs Open in VS Code Share feedback Refresh

Inputs (4)

- area
- Atm
- customer
- dmbi3

Outputs (1)

- job7output

Test query Save query Discard changes

```

1 SELECT 1 AS ALERT
2 INTO [job7output]
3 FROM
4     [dmbi3]
5 JOIN customer ON [customer].card_number=[dmbi3].CardNumber
6 WHERE [dmbi3].Type=1
7 GROUP BY SlidingWindow(hour,1)
8 HAVING count(*)>=2

```

Output

job7.json

Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More

job1.json 1 job2.json 1 job4.json 1 job5.json 1 job6.json 1 job7.json 1 job3.json 1

Users > thanosalexandris > Downloads > azure dmbi3 job outputs > job7.json > ...

```

1 {"ALERT":1}
2 {"ALERT":1}
3 {"ALERT":1}
4 {"ALERT":1}
5 {"ALERT":1}
6 {"ALERT":1}
7 {"ALERT":1}
8 {"ALERT":1}
9 {"ALERT":1}
10 {"ALERT":1}
11 {"ALERT":1}
12 {"ALERT":1}
13 {"ALERT":1}
14 {"ALERT":1}
15 {"ALERT":1}
16 {"ALERT":1}
17 {"ALERT":1}
18 {"ALERT":1}
19 {"ALERT":1}
20 {"ALERT":1}
21 {"ALERT":1}
22 {"ALERT":1}
23 {"ALERT":1}
24 {"ALERT":1}
25 {"ALERT":1}
26 {"ALERT":1}
27 {"ALERT":1}
28 {"ALERT":1}
29 {"ALERT":1}
30 {"ALERT":1}
31 {"ALERT":1}

```

Ln 1, Col 1 Spaces: 4 UTF-8 CRLF JSON Choose SQL Language

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)

Query

Test query Save query Discard changes

Inputs (4)

- Area
- Atm
- Customer
- dmbi3

Outputs (1)

- inh8output

```

1 SELECT 1 AS ALERT
2 INTO
3     [job8output]
4 FROM
5     [dmbi3]
6 JOIN [Customer] ON [Customer].card_number=[dmbi3].CardNumber
7 JOIN [Atm] ON [Atm].atm_code=[dmbi3].ATMCode
8 WHERE [Atm].area_code!= [Customer].area_code

```

Output

```
1 {"ALERT":1}
2 {"ALERT":1}
3 {"ALERT":1}
4 {"ALERT":1}
5 {"ALERT":1}
6 {"ALERT":1}
7 {"ALERT":1}
8 {"ALERT":1}
9 {"ALERT":1}
10 {"ALERT":1}
11 {"ALERT":1}
12 {"ALERT":1}
13 {"ALERT":1}
14 {"ALERT":1}
15 {"ALERT":1}
16 {"ALERT":1}
17 {"ALERT":1}
18 {"ALERT":1}
19 {"ALERT":1}
20 {"ALERT":1}
21 {"ALERT":1}
22 {"ALERT":1}
23 {"ALERT":1}
24 {"ALERT":1}
25 {"ALERT":1}
26 {"ALERT":1}
27 {"ALERT":1}
28 {"ALERT":1}
29 {"ALERT":1}
30 {"ALERT":1}
31 {"ALERT":1}
```