

# Real-time Analytics with Azure Stream Analytics

Michael Johnson  
[@MikeJohnsonZA](#)



# What I'd like to share with you today

Introduction to streaming data

Overview of Azure Stream Analytics

Demonstrate a simple streaming data solution

# What is streaming data

# What is streaming data?

Batch Data





# What is streaming data?

Streaming data



# What is streaming data?

Example



# What is streaming data?

## Sources of streaming data



# Azure Stream Analytics



# Azure Stream Analytics

What is Azure  
Stream  
Analytics?

Real-time event processing engine



# Azure Stream Analytics

Easy to use

Scalable

What  
Microsoft like  
you to know.

Reliable

Low Cost

# Architecture

# Architecture





# Architecture

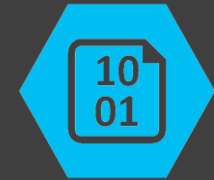
Inputs



Azure Event hub



Azure IOT hub



Azure Blob storage

# Architecture



Azure SQL database



Azure document DB



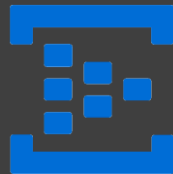
Azure Blob storage



Azure data lake



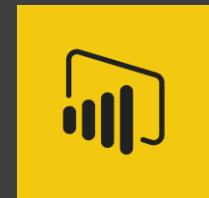
Azure Service bus



Azure Event hub



Azure table storage



Power BI

Outputs

Query language

# Query language

What does a  
query look like

SELECT

Make,  
System.TimeStamp as Time,  
count(\*) [NumOfCars]

INTO

[Output]

FROM

[Input] TIMESTAMP BY Time

GROUP BY

Make,  
TumblingWindow(Second, 10)

HAVING

[Count] >= 3



# Query language

## Query language overview

### DML Statements

- SELECT
- FROM
- WHERE
- GROUP BY
- HAVING
- CASE
- JOIN
- UNION

### Date and time

- DATEPART
- DAY
- MONTH
- YEAR
- DATEDIFF
- DATEADD

### Aggregation functions

- SUM
- COUNT
- AVG
- MIN
- MAX

### String functions

- LEN
- CONCAT
- CHARINDEX

### Scaling functions

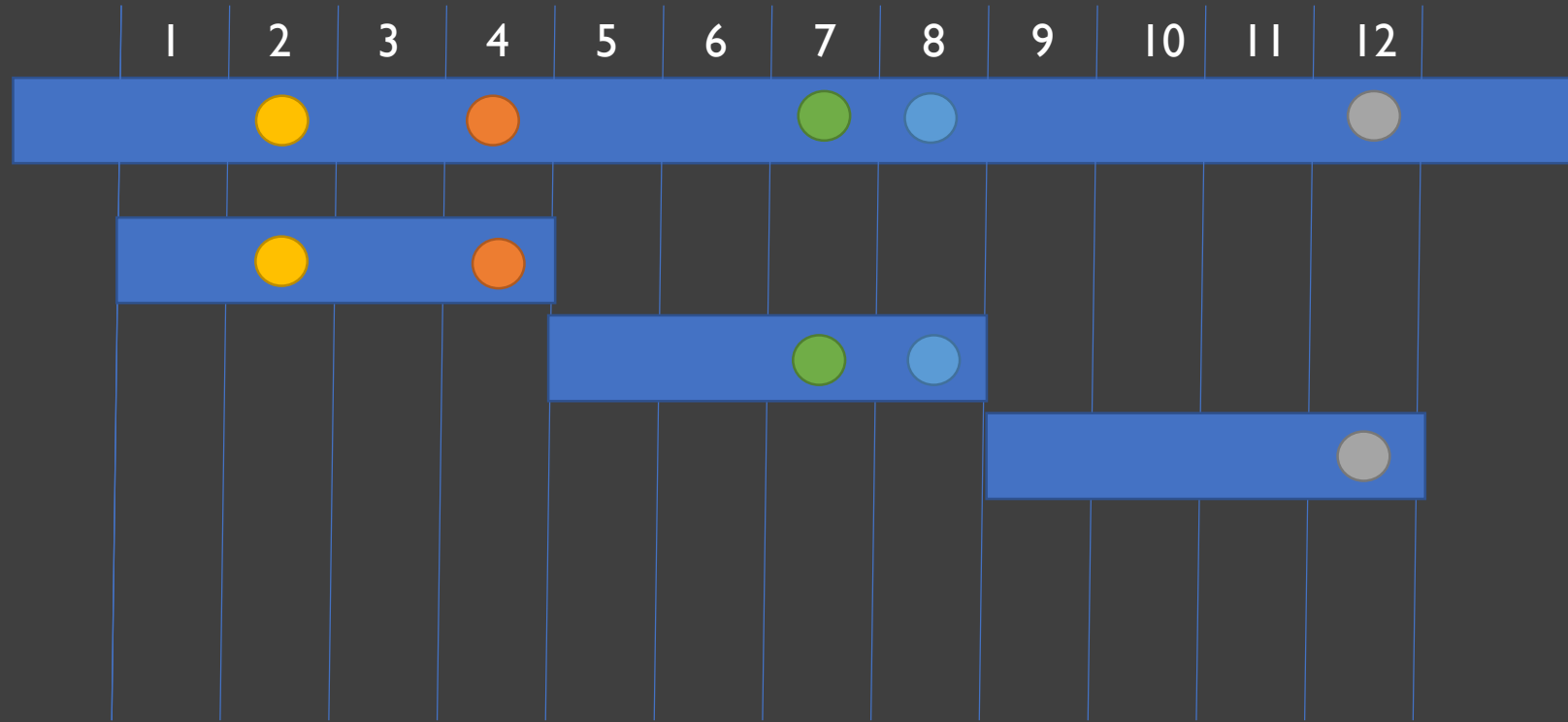
- WITH
- PARTITION BY

### Windowing functions

- TumblingWindow
- HoppingWindow
- SlidingWindow

# Query language

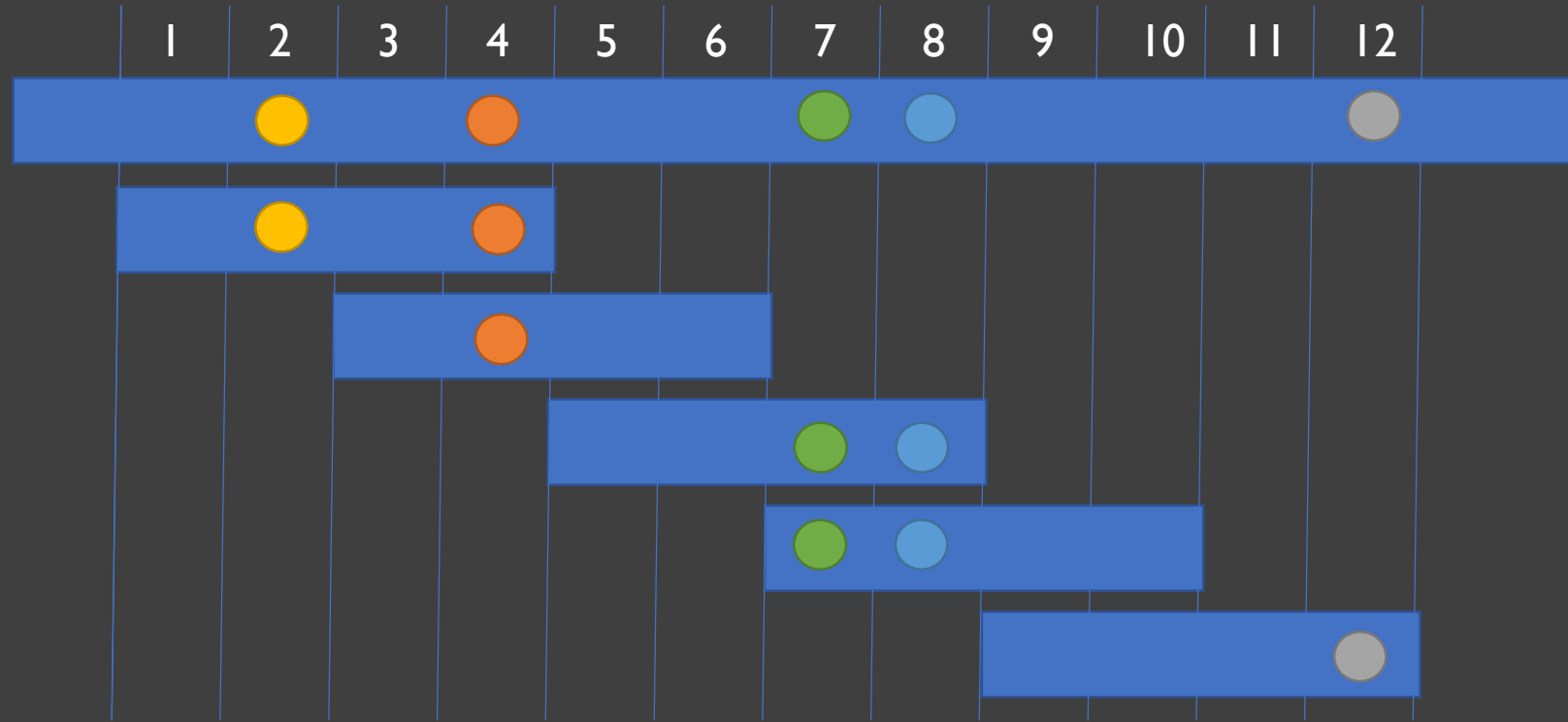
Tumbling  
Window



```
TUMBLINGWINDOW ( timeunit , windowsize,[offsetsize] )  
TUMBLINGWINDOW ( SECOND , 4)
```

# Query language

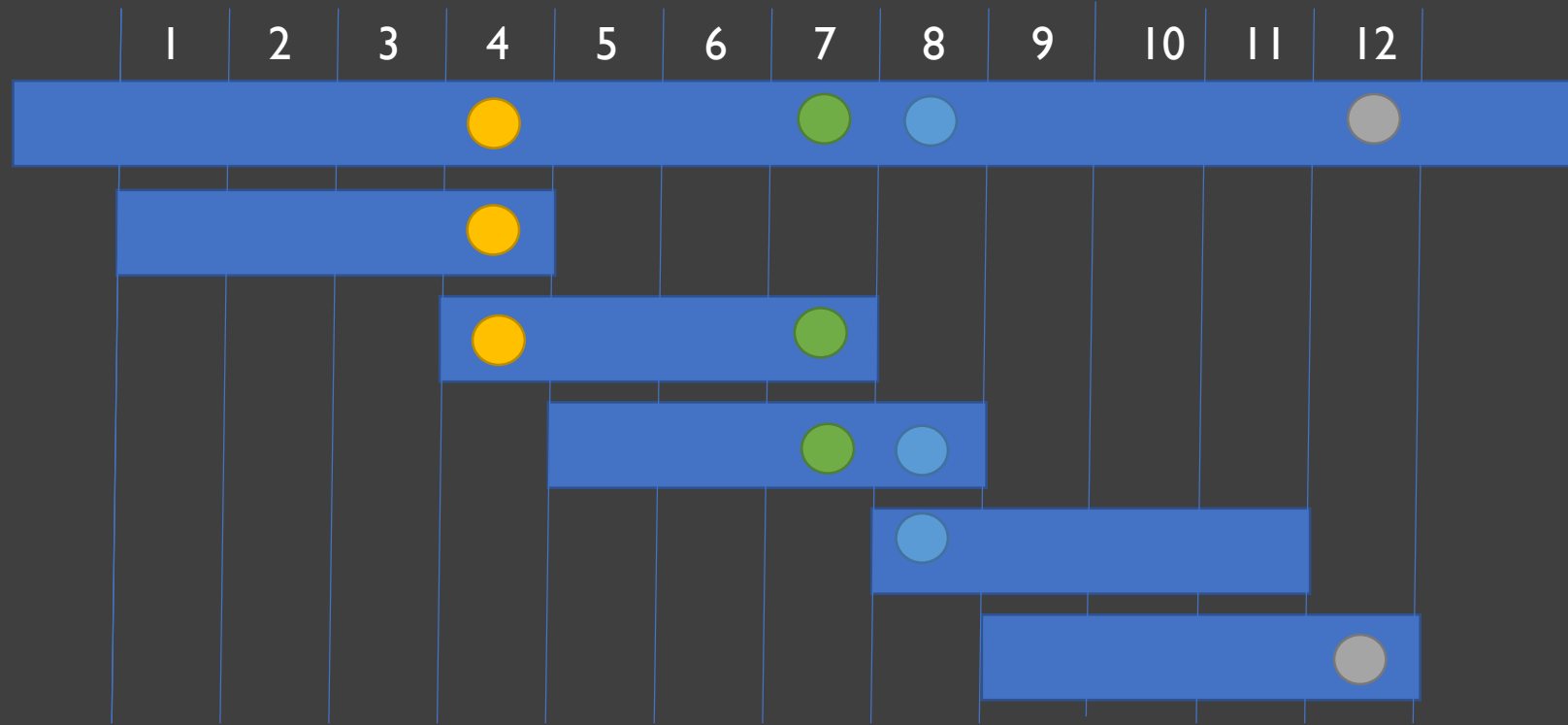
Hopping  
Window



```
HOPPINGWINDOW ( timeunit , windowsize , hopsize, [offsetsize] )  
HOPPINGWINDOW ( SECOND , 4 , 2 )
```

# Query language

Sliding  
Window



```
SLIDINGWINDOW ( timeunit , windowsize )  
SLIDINGWINDOW (SECOND,4)
```



# Query language

- Limitations

- Stateless
- Compute only
- No external connectivity

JavaScript UDF  
integration

- Why

- JavaScript libraries
  - REGEX
  - Math

Code

```
// Convert Hex value to integer.  
function main(hexValue) {  
    return parseInt(hexValue, 16);  
}
```

Code

```
SELECT  
    time,  
    UDF.hex2Int(offset) AS IntOffset  
INTO  
    output  
FROM  
    InputStream
```

# Query language

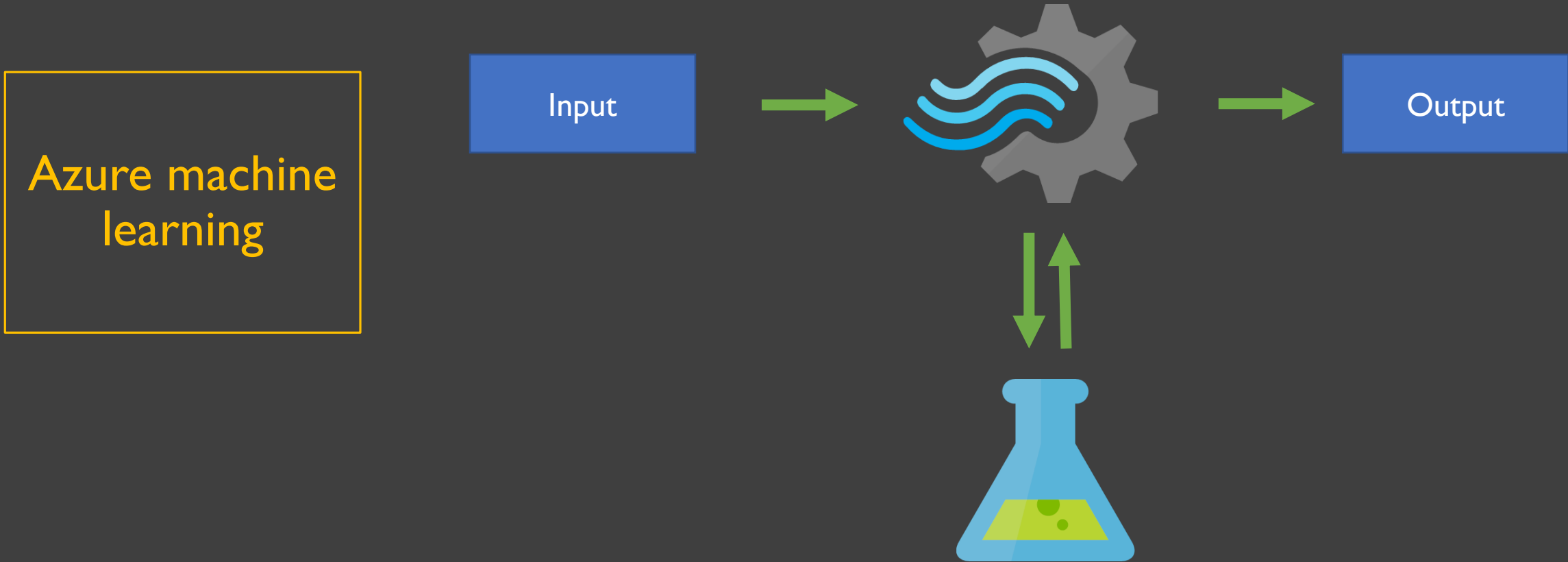
## Geospatial queries

- Data types
  - Point
  - Line
  - Polygon
- Functions
  - ST\_DISTANCE
  - ST\_OVERLAP
  - ST\_INTERSECTS
  - ST\_WITHIN



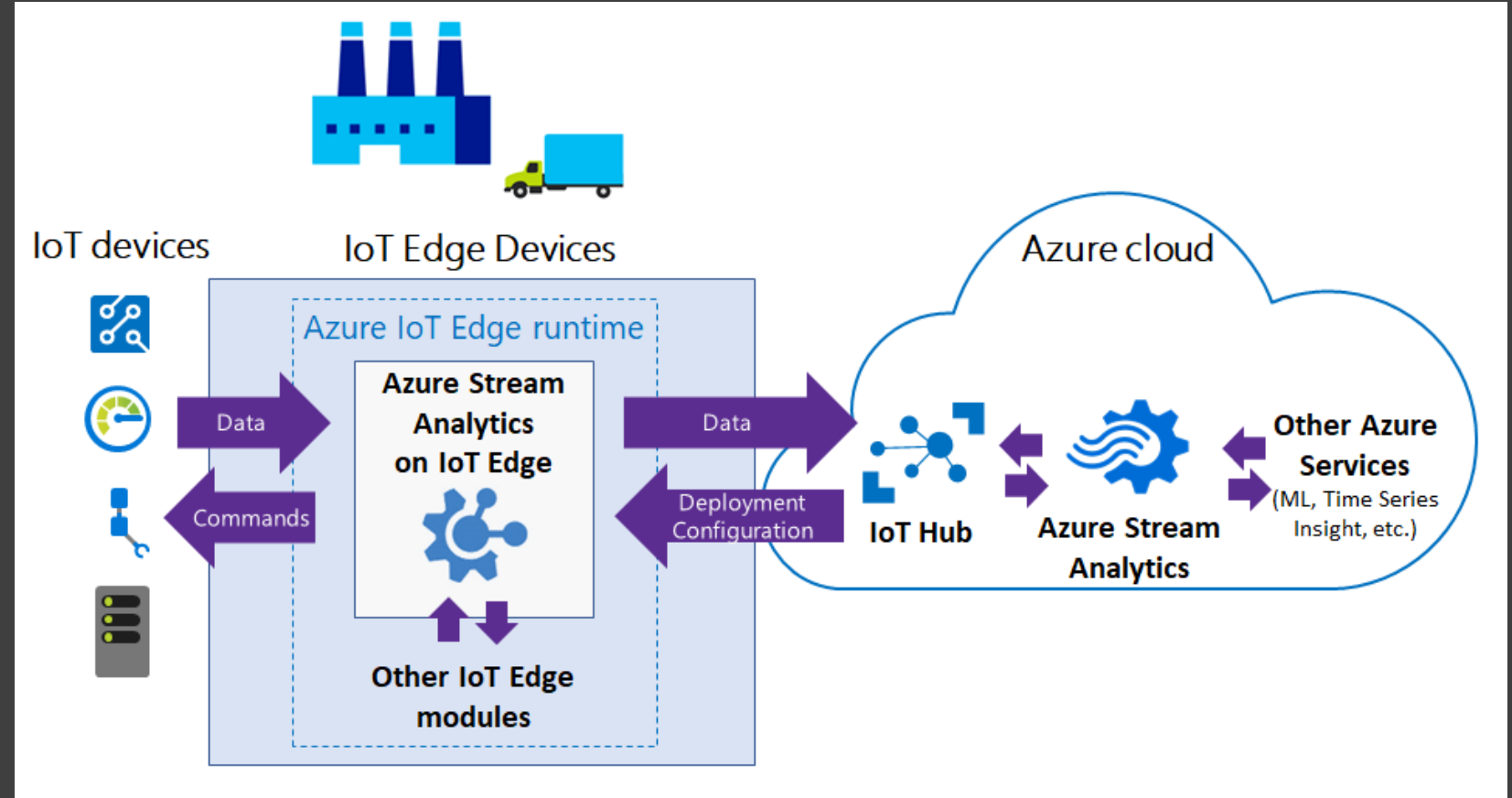
```
SELECT  c.Location, s.Location
FROM Cars c
JOIN Station s
ON ST_DISTANCE(c.location,s.location) < 10000
```

# Query language



# Azure Stream analytics on Azure IoT Edge

Preview





# Anomaly Detection

Preview

```
ANOMALYDETECTION(<scalar_expression>)  
OVER ([PARTITION BY <partition key>]  
LIMIT DURATION(<unit>, <length>) [WHEN  
boolean_expression])
```

# Development and deployment options

# Development and deployment options

## Azure Portal

The screenshot displays the Microsoft Azure Portal interface for a Stream Analytics job named 'BAVCDemo1'. The top navigation bar includes the 'Microsoft Azure' logo, the job name 'BAVCDemo1', and a user profile for 'MichaelAdrianJohnso... MIKE JOHNSON'. The left sidebar contains a search bar and a list of navigation options: Overview (selected), Activity log, Access control (IAM), Tags, Diagnose and solve problems, SETTINGS (Locks), JOB TOPOLOGY (Inputs, Functions, Query, Outputs), and a bottom section with icons for various Azure services. The main content area shows the 'Created' status of the job. Under the 'Essentials' section, details for the resource group 'BAVCDemo' are listed, including its status, location (West Europe), subscription name (Visual Studio Enterprise), and subscription ID. The 'Job Topology' section displays three panels: 'Inputs' with 0 results, 'Query' with a code editor icon, and 'Outputs' with 0 results. The 'Query' panel is currently active, showing a code editor with a '<>' icon.

Microsoft Azure BAVCDemo1

BAVCDemo1  
Stream Analytics job

Search (Ctrl+/)

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

SETTINGS

Locks

JOB TOPOLOGY

Inputs

Functions

Query

Outputs

Settings Start Stop Delete

Created

Essentials

Resource group (change)  
BAVCDemo

Status  
Created

Location  
West Europe

Subscription name (change)  
Visual Studio Enterprise

Subscription ID  
f5f2686c-a11d-4e97-96f5-578881d4da02

Send feedback  
UserVoice

Created  
Thursday, 11 May 2017, 13:59:17

Started  
-

Last output  
-

Job Topology

Inputs

0

No results.

Query

<>

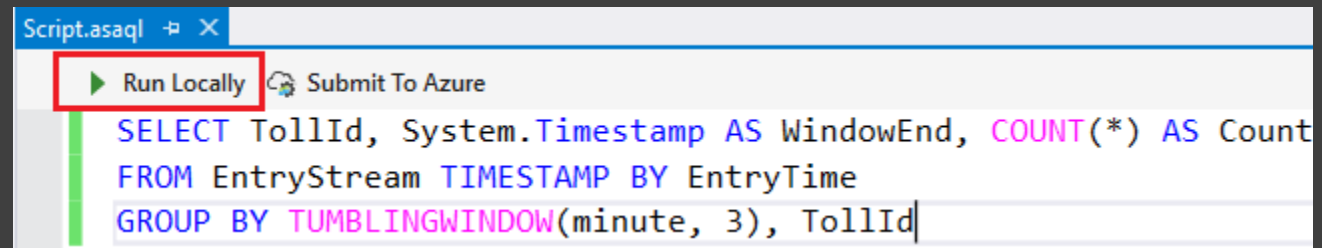
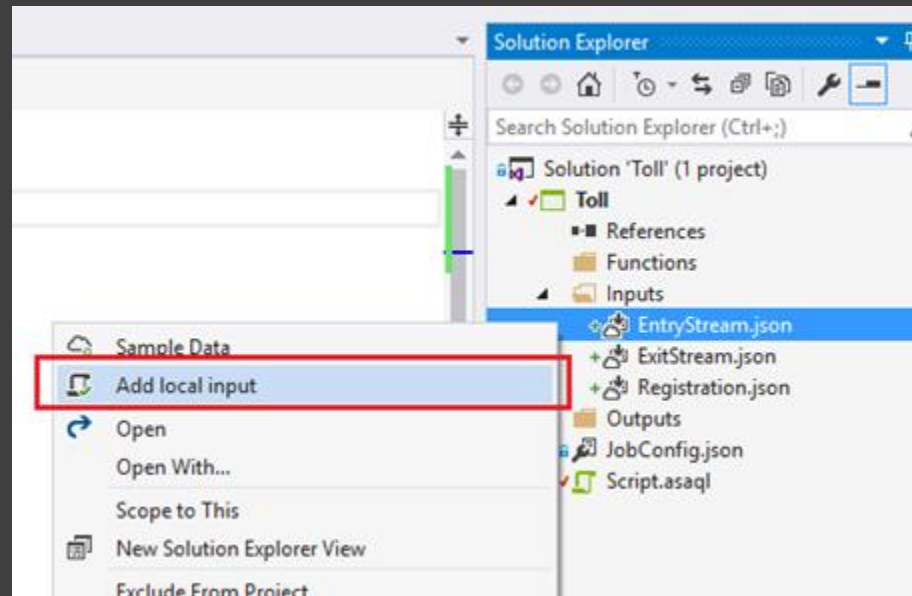
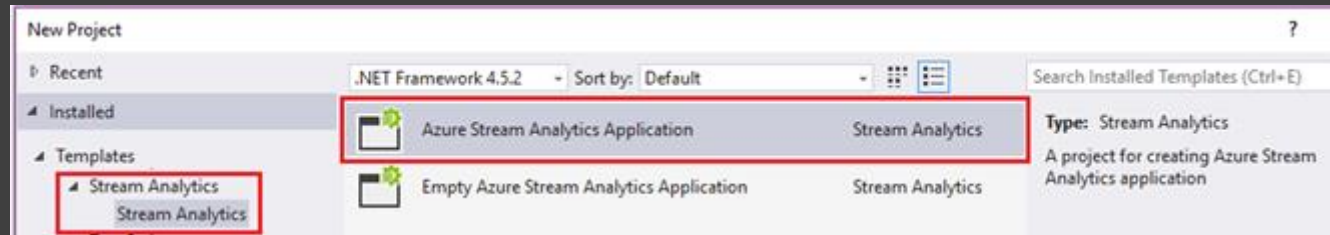
Outputs

0

No results.

# Development and deployment options

## Visual Studio



# Development and deployment options

## PowerShell

```
Administrator: Microsoft Azure PowerShell
PS C:\> cd .\TollApp\TollApp
PS C:\TollApp\TollApp> .\Setup.ps1
You are signed-in with dipanjan.banik@live.com

Your Azure Subscriptions:

RowNumber SubscriptionId SubscriptionName
-----
1 d115f887-f937-4c61-9e59- Visual Studio Ultimate with MSDN
2 7a74e4df-81b1-4817-9616-

Enter the row number (1 - 2) of a subscription: 1
Identifying Location for your lab East US
Creating/Validating resources for Toll App
Creating the Service Bus Namespace I TollData104094364 I.....created.
Creating EventHub I entry I.....created.
Creating EventHub I exit I.....created.
Creating SQL Server ..... [svr name: bysttpgt8d I....created.
Creating SQL DB I TollDataDB I.....created.
Creating required sqlTables.....created.
Creating AzureStorageAccount I tolldata104094364 I....
Write-Host AzureStorageAccount I tolldata104094364 I created.
Creating Container I tolldata I....created
Uploading reference data to container....Completed

All Resource Names

You are signed-in with dipanjan.banik@live.com
Subscription Id: d115f887-f937-4c61-9e59-
Subscription Name: TollApp

Service Bus:
Namespace: TollData104094364
SharedAccessKeyName: RootManageSharedAccessKey
SharedAccessKey: xg6kkULCxouRoP4BW4gZESuJ9RLQEv2IIa6/wESQ2o=

Sql Server:
Server: bysttpgt8d + .database.windows.net
SqlLogin: tolladmin
Password: 123toll!
DatabaseName: TollDataDB

Storage Account:
AccountName: tolldata104094364
AccountKey: 1sWv00ZQIIMqo7vqAVDlywTQuqGESNTMIr6CS9Xoalz0aY9dmYRo2T+77WK/1y66412ThAJ/Eo+yMUAh5upUgQ==

Location: East US

OperationDescription OperationId OperationStatus
-----
New-AzureStorageAccount 8b50b564-a275-6e30-a4bf-d28b0774a236 Succeeded

PS C:\TollApp\TollApp>
```

# Capacity and Pricing

## Streaming unit

- Streaming unit
  - 1 streaming unit = 1MB/Second
    - $1 \times 60 \times 60 \times 24 = 84.375$  GB/day
  - Data volume + streaming unit cost
    - Data volume = \$ 0.001 per GB
    - Streaming unit = \$ 0.12 per hour
  - If run at max through put for a month
    - Total data ingested = 2.47 TB
    - Data = \$ 2.50
    - Streaming Unit = \$ 86
    - Total cost = \$ 88.50

DEMO

Create an Event Hub

Create event generator

Connect Event Hub to Stream Analytics

Persist output events to a database

Create a Power BI output for real-time monitoring

# Questions



Just like Jimi Hendrix ...

We love to get feedback

Please complete the session feedback forms

# It's all about the community...

SQL BITS

Please visit Community Corner, we are trying this year to get more people to learn about the SQL Community, equally if you would be happy to visit the community corner we'd really appreciate it.

Thank you