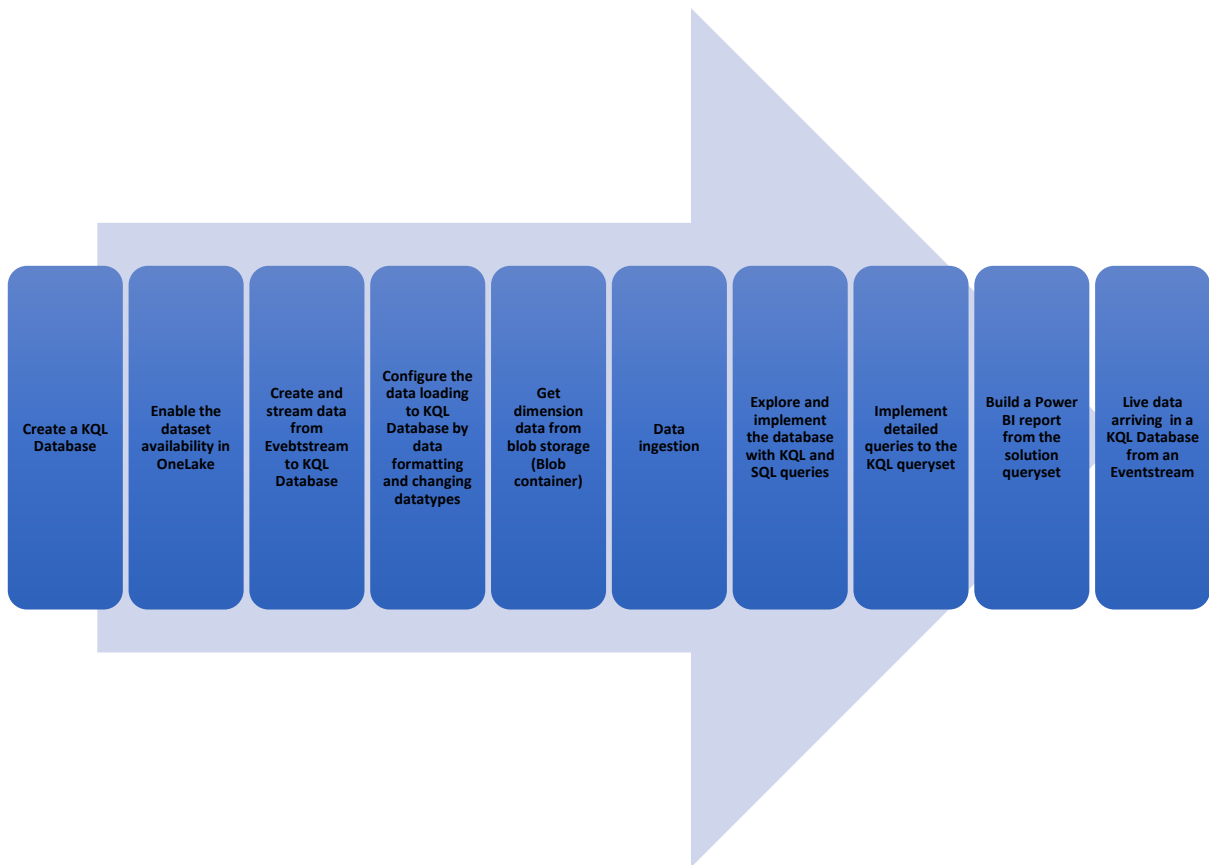


REAL-TIME ANALYTICS: NYC TAXI TRIPS

Using Microsoft Fabric- an end-to-end analytics solution with full-service capabilities including data movement, data lakes, data engineering, data integration, data science, real-time analytics, and business intelligence.



The screenshot displays the Microsoft Fabric interface for the 'NycTaxiDB' database. The left sidebar shows the 'Data tree' with a search bar and a list of items: 'NycTaxiDB' (selected), 'Tables', 'Shortcuts', 'Materialized views', 'Functions', and 'Data stream'. The main panel shows the 'Database: NycTaxiDB' details, including a table of database information and a 'Size' section.

Database details	
Created by	Vidya Kailasam Palaninatharaja
Region	northcentralus
Created on	Today, 4 hr ago
Last ingestion	Today, 1 min ago
Query URI	Copy URI
Ingestion URI	Copy URI
OneLake folder	Copy path Active

Size

Original size	Compressed size	Compression ratio
8.09 MB	963.18 KB	8.6

Top tables | Most active users

Create

Workspaces

My workspace

NycTaxiDB

stockmarket

qs

FabricDB

...

Real-Time Analytics

Data tree

Database

Search

NycTaxiDB

Tables

Locations

nyc-taxitrips

Shortcuts

No results

Materialized views

No results

Functions

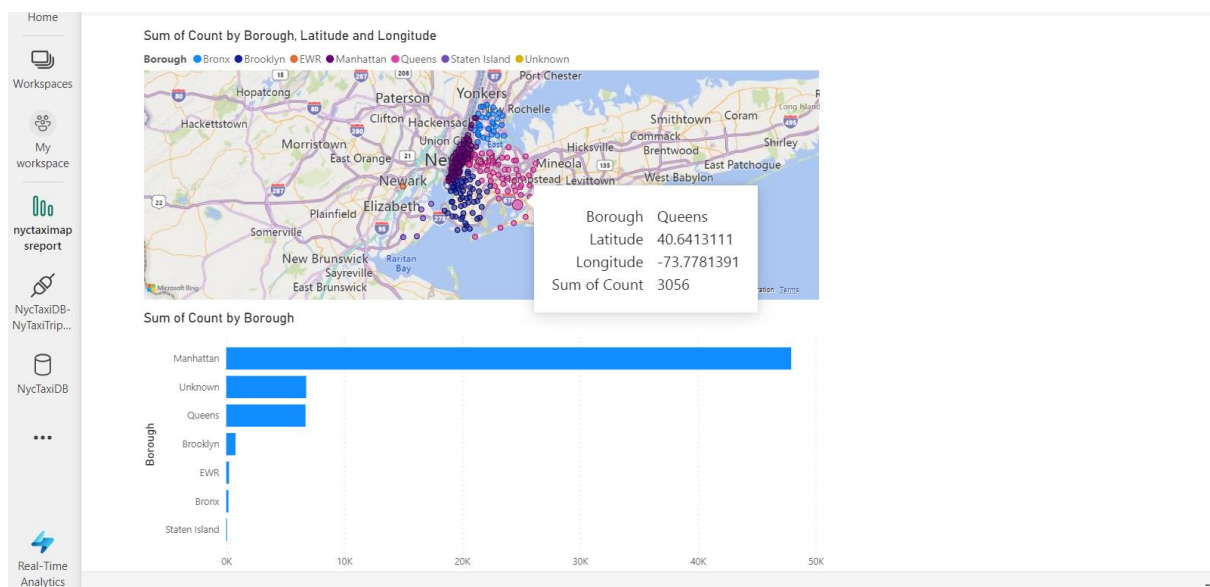
No results

Schema

Find by keyword

Filter

Column name	Type
VendorID	# Integer
tpep_pickup_datetime	🕒 Date/Time
tpep_dropoff_datetime	🕒 Date/Time
passenger_count	# Long
trip_distance	# Real
RatecodeID	# Real
store_and_fwd_flag	T String



nyc-taxitrips

```
| summarize Count=count() by PULocationID
| top 10 by Count
```

nyc-taxitrips

```
| lookup (Locations) on $left.PULocationID == $right.LocationID
| summarize Count=count() by Zone
| top 10 by Count
| render columnchart
```

nyc-taxitrips

```
| lookup (Locations) on $left.PULocationID==$right.LocationID
| where Borough == "Manhattan"
| make-series s1 = avg(tip_amount) on tpep_pickup_datetime from datetime(2022-06-01) to datetime(2022-06-04) step 1h
| extend anomalies = series_decompose_anomalies(s1)
| render anomalychart with (anomalycolumns=anomalies)
```

nyctaxitrips

```
| lookup (Locations) on $left.PULocationID==$right.LocationID
| where Borough == "Manhattan"
| make-series s1 = count() on tpep_pickup_datetime from datetime(2022-06-01)
to datetime(2022-06-08)+3d step 1h by PULocationID
| extend forecast = series_decompose_forecast(s1, 24*3)
| render timechart
```

nyctaxitrips

```
| where PULocationID == DOLocationID
| lookup (Locations) on $left.PULocationID==$right.LocationID
| summarize Count=count() by Borough, Zone, Latitude, Longitude
```

The screenshot shows a Power BI workspace with a database named 'NycTaxiDB'. The 'Database' pane on the left lists tables: 'Locations' (with columns LocationID, Borough, Zone, service_zone, Latitude, Longitude) and 'nyctaxitrips' (with columns VendorID, PULocationID, DOLocationID, tpep_pickup_datetime). The main query editor shows a SQL query that joins 'Locations' with 'nyctaxitrips' on 'PULocationID' and 'DOLocationID', filters for 'Borough = Manhattan', and summarizes the count by 'Borough, Zone, Latitude, Longitude'. The results pane shows a table with 218 records.

Borough	Zone	Latitude	Longitude	Count
Manhattan	Murray Hill	40.7479	-73.9757	1,451
Unknown	NV			6,080
Manhattan	Lincoln Square West	40.7734551	-73.985679	798
Manhattan	Midtown East	40.7553	-73.9693	1,284
Manhattan	Lincoln Square East	40.7734551	-73.985679	1,670
Manhattan	Midtown Center	40.7577	-73.9857	1,934
Manhattan	Upper West Side South	40.7829	-73.9765	2,200
Manhattan	Garment District	40.753995	-73.988947	385
Manhattan	Upper West Side North	40.7958	-73.9719	1,522
Manhattan	Morningside Heights	40.8105	-73.9624	433