

DAMG 7275 Project Team 8

Team Member

1. Jiayue Han - NUID: 002965004
2. Zhiqing Wang - NUID: 002916484
3. Kewei Shi - NUID: 001359675
4. Xinan Wang - NUID:002916472

Topic

US Used Car Data Management System

Data Refresh Implementation

Since the dataset for our project is from Kaggle.com and it is not likely that data will be updated regularly, we decided to implement data refresh and pipeline automation using trigger in azure data factory.



Our plan is to update data by manually uploading new CSV documents to the blob storage account, and the event trigger would be triggered when new file is uploaded, or existing file is deleted. Once an event happens, the whole pipeline will be executed and the CosmosDB items will be updated accordingly. Blob path is also set up so the trigger only function when file with specific name and format is uploaded to the blob storage account.

Name *

Trigger_Event

Description

Type *

Storage events

Rest of the trigger setting is shown in next page.

Account selection method * ⓘ

☒ From Azure subscription ☐ Enter manually

Azure subscription ⓘ

Pay-As-You-Go (004e0cb3-cffd-4632-81bc-5934d1a41dec) ▼

Storage account name * ⓘ

team8project ▼



Container name * ⓘ

input ▼

Blob path begins with ⓘ

used

Blob path ends with ⓘ

.csv

Event * ⓘ



Blob created



Blob deleted

Ignore empty blobs * ⓘ



Yes



No