BANK DOMAIN PROJECT

Project Overview:

This project focuses on building a cloud-based data ingestion pipeline for banking datasets using Azure serverless components. The goal is to ingest raw banking data into a centralized system where it can later be transformed, validated, and analyzed for fraud detection, customer insights, and reporting needs.

Objectives Completed on Day1 :

1. Storage Account (ADLS Gen2)

Storage Account Purpose: Stores raw incoming banking files.

Name: azurebankstore

Folder Structure: `raw/transactions`

`raw/customers/`

`raw/atm/`

`raw/upi/`

2. Function App

Executes automatically on arrival of events/messages.

Name: Azurebankfunc
Runtime: Python 3.11
Hosting: Consumption plan

3. Azure Service Bus Namespace & Queue

Holds metadata messages sent from EventGrid-based triggers

Name: ingestion-queue

4. Configure Event Subscriptions

Enable Event Grid on Storage Account

Storage Account → Events
Created a System Topic automatically:
Name: Event-Trigger-New-File-Topic
Topic Type: Microsoft.Storage.StorageAccounts

5. Azure Functions Implemented

5.1 Function-1: eventgrid_ingest_blob

Trigger Type: Event Grid
Triggered When: A file is uploaded to storage

Purpose: Detect file upload and Extract metadata then Send this metadata into queue to be processed later

5.2 Function-2: process_event_message (Created but not implemented yet)

Trigger Type: Service Bus Queue Trigger

Further Usecase :

Will run when a queue message arrives
Will download actual CSV
Will extract data
Will validate & transform
Will later insert into Cosmos DB

➞ I have created the function trigger the Actual processing logic will be implemented later

5.3 Function-3: eventgrid_high_value_txn

Trigger Type: Event Grid

Future Purpose:

Handle high-value alerts
Save alerts into FraudAlerts container

This will act as fraud detection component later.

5.4 Function-4: Timer Trigger

Trigger Type: Schedule trigger

Future Role:

This is useful for daily batch activities.

6. Local Execution Validation

Using: func start

I have verified:

All functions load correctly
Triggers are registered
No runtime errors
Environment variables working

Still these functions has not Deployed. The Deployement will be done in Day 2

Current Workflow:

File Upload → Event Grid Triggered → Function receives metadata → Message prepared

Today only the detection and trigger phase has been Compeleted.

Day -2 Objectivites to Complete

DEPLOYMENT OF THE ALL THE FUNCTION CREATED ON DAY1

COMPLETE SERVICE BUS PROCESSING FUNCTION

CONFIGURE EVENT GRID IN PORTAL

INSERT INTO COSMOS DB

IMPLEMENT FRAUD LOGIC

RUN TIMER BASED DAILY JOBS