

## Practical Question 2 — Real-Time Event Ingestion & Analytics

### Objective

To build a real-time analytics pipeline using **Google Cloud Pub/Sub**, **Cloud Functions**, and **BigQuery** to ingest user activity events, process them in real time, store valid records in BigQuery, and route invalid records to a Dead Letter Queue (DLQ).

### Architecture Overview

1. **Pub/Sub Topic** receives user activity events in JSON format.
2. **Cloud Function (Gen2)** is triggered by the Pub/Sub topic.
3. The function:
  - Validates incoming JSON messages
  - Inserts valid events into BigQuery
  - Logs errors and publishes invalid messages to a DLQ topic
4. **BigQuery** stores processed events for analytics.

### Step 1: Create Pub/Sub Topics and Create Pub/Sub Topics

```
gcloud pubsub topics create user-activity
```

```
gcloud pubsub topics create user-activity-dlq-sujatha
```

### Step 2: Create BigQuery Dataset and Table

```
3. bq mk realtime_sujatha
```

And creating table with required schema

```
4. bq mk \ --table \ realtime_sujatha.user_events_sujatha \  
event:STRING,user:STRING,ts:TIMESTAMP
```

### Step 3: Develop Cloud Function

#### Function Responsibilities

- Subscribe to user-activity topic
- Parse JSON messages
- Validate required fields: event, user, ts
- Insert valid records into BigQuery
- Send invalid messages to DLQ and log errors

### Step 4: Deploy Cloud Function

```
gcloud functions deploy ingest_user_activity --gen2 --runtime=python311 --region=us-central1 --  
source=. --entry-point=ingest_user_activity --trigger-topic=user-activity --allow-unauthenticated
```

## Step 5: Publish Test Messages

Publish a valid message

```
gcloud pubsub topics publish user-activity --message "{\"event\":\"LOGIN\\\", \"user\":\"sujatha\\\", \"ts\":\"2025-01-01T10:00:00Z\"}"
```

Publish an invalid message (missing user field)

```
gcloud pubsub topics publish user-activity --message "{\"event\":\"LOGIN\\\", \"ts\":\"2025-01-01T10:00:00Z\"}"
```

## Step 6: Verify BigQuery Data

### Step 7: Verify Dead Letter Queue (DLQ)

```
gcloud pubsub subscriptions pull user-activity-dlq-sub --auto-ack --limit=5
```

## Step 8: View Logs

```
gcloud functions logs read ingest_user_activity --region=us-central1 --limit=10
```

```
Gemini CLI is available in Cloud Shell terminal. type gemini to try it. Learn more
Welcome to Cloud Shell! Type "help" to get started, or type "gemini" to try prompting with Gemini CLI.
Your Cloud Platform project in this session is set to pavanproject-481304.
Use "gcloud config set project [PROJECT_ID]" to change to a different project.
honnalesujatha@cloudshell:~ (pavanproject-481304)$ gcloud services enable cloudfunctions.googleapis.com eventarc.googleapis.com run.googleapis.com firestore.googleapis.com storage.googleapis.com logging.googleapis.com
Operation "operations/acat.p2-49438910600-af25e378-2d49-40f1-bdcf-c3500ed9bdf3" finished successfully.
honnalesujatha@cloudshell:~ (pavanproject-481304)$ gcloud firestore databases create --region=asia-south1
ERROR: (gcloud.firestore.databases.create) unrecognized arguments: --region=asia-south1 (did you mean "--location"?)
```

To search the help text of gcloud commands, run:

```
gcloud help --search TERMS
honnalesujatha@cloudshell:~ (pavanproject-481304)$ ^C
honnalesujatha@cloudshell:~ (pavanproject-481304)$ gcloud firestore databases create --location=asia-south1
metadata:
  'type': type.googleapis.com/google.firestore.admin.v1.CreateDatabaseMetadata
name: projects/pavanproject-481304/databases/(default)/operations/Aa4sD_0156aH8RxyYJoghAqWh0dW9z1MFpc2ELI9QW6e3yBAGyz29ggglCi4a
response:
  'type': type.googleapis.com/google.firestore.admin.v1.Database
apiEngineIntegrationMode: DISABLED
concurrencyMode: PESSIMISTIC
createTime: '2025-12-30T05:09:22.107597Z'
databaseEdition: STANDARD
deleteProtectionState: DELETE_PROTECTION_DISABLED
earliestVersionTime: '2025-12-30T05:09:22.107597Z'
etag: ICKL0+bF5-JEDM2R0+bF5JED
freeTier: true
keyPrefix: j
locationId: asia-south1
name: projects/pavanproject-481304/databases/(default)
pointInTimeRecoveryEnablement: POINT_IN_TIME_RECOVERY_DISABLED
realtimeUpdateMode: REALTIME_UPDATED_MODE_ENABLED
type: FIRESTORE_NATIVE
```

Pub/Sub

Topics

Subscriptions

Snapshots

Schemas

Pub/Sub Lite

Lite reservations

Lite topics

Lite Subscriptions

Release notes

Create topic

Delete

Tags

Show info panel

Learn

Pub/Sub offers SMTs to simplify data transformations for streaming pipelines

Perform single message transforms such as casting, redaction, masking and complex filtering to message attributes and/or the data directly within Pub/Sub using new JavaScript user-defined functions (UDFs).

Learn more

Simplify data lake pipelines with new Pub/Sub Cloud Storage subscriptions

You can now streamline your data ingestion pipelines with Cloud Storage subscriptions, enabling you to write raw streaming data into Cloud Storage without any transformations in between. To get started, create a new Cloud Storage subscription for a Pub/Sub topic.

Create Cloud Storage subscription

Learn more

	Topic	Owner	Project	Subscription	Message Count	Created	Updated	Deleted
<input checked="" type="checkbox"/>	<a href="#">eventarc-asia-south2-sujatha-trigger1-242</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/eventarc-asia-south2-sujatha-trigger1-242	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">eventarc-asia-south2-sujatha-trigger2-264</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/eventarc-asia-south2-sujatha-trigger2-264	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">eventarc-asia-south2-trigger-c9ujlmj-933</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/eventarc-asia-south2-trigger-c9ujlmj-933	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">sainath-user-events</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/sainath-user-events	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">user-activity</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/user-activity	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">user-activity-dlq-sujatha</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/user-activity-dlq-sujatha	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">user-activity-sujatha</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/user-activity-sujatha	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z
<input type="checkbox"/>	<a href="#">user-events</a>	Google-managed	projects/pavanproject-481304	projects/pavanproject-481304/topics/user-events	0	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z	2025-12-30T05:09:22.107597Z

CLOUD SHELL

Terminal

(pavanproject-481304)

Open editor

Google CloudPavanProject

topics

Search

Pub/Sub / Topics / Topic: user-activity-dlq-sujatha

Pub/Sub

Topics

Subscriptions

Snapshots

Schemas

Pub/Sub Lite

Lite reservations

Lite topics

Lite Subscriptions

user-activity-dlq-sujatha

Edit

+ Trigger Cloud Run function

Import

Delete

Learn

Show info panel

Topic name

projects/pavanproject-481304/topics/user-activity-dlq-sujatha

Tags

Export to BigQuery

Export data to a BigQuery table.

Export data

Export to Cloud Storage

Export data to a text or Avro file in Cloud Storage.

Export data

Subscriptions

Snapshots

Metrics

Details

Messages

Only subscriptions attached to this topic are displayed. A subscription captures the stream of messages published to a given topic. You can also stream messages to BigQuery or Cloud Storage by creating a subscription from a Cloud Dataflow job. [Learn more](#)

Create subscription

Export

Filter

Filter subscriptions

Subscription ID

Subscription name

Project

user-activity-dlq-sub

projects/pavanproject-481304/subscriptions/user-activity-dlq-sub

pavanproject-481304

CLOUD SHELL

Terminal

pavanproject-481304

Open editor

Google CloudPavanProject

Search (/) for resources, docs, products and more

Search

Pub/Sub / Topics / Topic: user-activity-sujatha

Pub/Sub

Topics

Subscriptions

Snapshots

Schemas

Pub/Sub Lite

Lite reservations

Lite topics

Lite Subscriptions

user-activity-sujatha

Edit

+ Trigger Cloud Run function

Import

Delete

Learn

Show info panel

Topic name

projects/pavanproject-481304/topics/user-activity-sujatha

Tags

Export to BigQuery

Export data to a BigQuery table.

Export data

Export to Cloud Storage

Export data to a text or Avro file in Cloud Storage.

Export data

Subscriptions

Snapshots

Metrics

Details

Messages

Only subscriptions attached to this topic are displayed. A subscription captures the stream of messages published to a given topic. You can also stream messages to BigQuery or Cloud Storage by creating a subscription from a Cloud Dataflow job. [Learn more](#)

Create subscription

Export

Filter

Filter subscriptions

Subscription ID

Subscription name

Project

user-events-sujatha

projects/pavanproject-481304/subscriptions/user-events-sujatha

pavanproject-481304

Search BigQuery resources

Show starred only

pavanproject-481304

Repositories

Queries

Notebooks

Data canvases

Data preparations

Pipelines

Connections

analytics\_ds

dataflow\_ds

realtime\_sujatha

user\_events\_sujatha

pavanproject-481304 / Datasets / realtime\_sujatha / Tables / user\_events\_sujatha

user\_events\_sujatha

Query

Open in

Share

Copy

Snapshot

Delete

Export

Refresh

Schema

Details

Preview

Table explorer

Preview

Insights

Lineage

Data profile

Data Quality

Filter

Enter property name or value

Field name

Type

Mode

Description

Key

Collation

Default value

Policy tags

Data policies

event

STRING

NULLABLE

-

-

-

-

-

-

user

STRING

NULLABLE

-

-

-

-

-

-

ts

TIMESTAMP

NULLABLE

-

-

-

-

-

-

Edit schema

View row access policies

Search BigQuery resources

Show starred only

pavanproject-481304

Repositories

Queries

Notebooks

Data canvases

Data preparations

Pipelines

Connections

analytics\_ds

dataflow\_ds

realtime\_sujatha

user\_events\_sujatha

pavanproject-481304 / Datasets / realtime\_sujatha / Tables / user\_events\_sujatha

user\_events\_sujatha

QueryOpen inShareCopySnapshotDeleteExportRefresh

SchemaDetailsPreviewTable explorerPreviewInsightsLineageData profileData Quality

Preview may not show all data being streamed to BigQuery. If you want to check for such data, run a SELECT statement over the table.

Row	event	user	ts
1	LOGIN	sujatha	2025-01-01 10:00:00 UTC
2	LOGOUT	sujatha	2025-01-01 13:00:00 UTC

Cloud Run

ServicesDeploy containerConnect repoWrite a functionRefresh

Overview

Services

Jobs

Worker pools

Domain mappings

A service exposes a unique endpoint and automatically scales the underlying infrastructure to handle incoming requests. Deploy a container image, source code or a function to create a service.

Filter services

	Name	Deployment type	Req/sec	Region	Authentication	Ingress	Last deployed	Deployed by
<input type="checkbox"/>	ingest-user-activity	{-} Function	0	us-central1	Public access	All	20 minutes ago	Cloud Run funct
<input type="checkbox"/>	sainath	{-} Function	0	asia-south2	Require authentication	All	13 minutes ago	reddysainath741
<input type="checkbox"/>	sujatha	{-} Function	0.05	asia-south2	Public access	All	35 minutes ago	pavanpanni6@g

FileEditSelectionViewGoRunTerminalHelp

Finals Assement

EXPLORER

main.py

pyenv.cfgvenvrequirements.txt

FINALs ASSEMENT

venv

main.py

requirements.txt

PROBLEMS

OUTPUT

DEBUG CONSOLE

TERMINAL

PORTS

AZURE

availableMemory: 256M

environmentVariables:

LOG\_EXECUTION\_ID: "true"

ingressSettings: ALLOW\_ALL

maxInstanceConcurrency: 1

revision: ingest-user-activity-00001.web

service: projects/pavanproject-481304/locations/us-central1/services/ingest-user-activity

serviceAccountEmail: 49438910600-compute@developer.gserviceaccount.com

timeoutSeconds: 60

uri: https://ingest-user-activity-7ve7vosmua-uc.a.run.app

state: ACTIVE

updateTime: "2025-12-30T05:45:11.818458847Z"

url: https://us-central1-pavanproject-481304.cloudfunctions.net/ingest\_user\_activity

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>cloud pubsub topics publish user-activity --message '{"event":"LOGIN","user":"sujatha","ts":"2025-01-01T10:00:00Z"}'

messageIds:

- '17562595112241815'

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>bq query --use\_legacy\_sql=false "SELECT \* FROM realtime\_sujatha.user\_events\_sujatha ORDER BY ts DESC LIMIT 5"

| event | user | ts |

|-----|-----|-----|

| LOGIN | sujatha | 2025-01-01 10:00:00 |

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>cloud functions logs read ingest\_user\_activity --region=us-central1 --limit=10

LEVEL NAME EXECUTION\_ID TIME\_UTC LOG

I ingest-user-activity 2025-12-30 05:58:04.443

I ingest-user-activity 2025-12-30 05:58:03.973

I ingest-user-activity 2025-12-30 05:58:03.083

I ingest-user-activity 2025-12-30 05:46:15.062

I ingest-user-activity 2025-12-30 05:45:04.334 Default STARTUP TCP probe succeeded after 1 attempt for container "worker" on port 8080.

I ingest-user-activity 2025-12-30 05:45:01.280 Starting new instance. Reason: DEPLOYMENT\_ROLLOUT

- Instance started due to traffic shifting between revisions due to deployment, traffic split adjustment, or deployment

Build with Agent

AI responses may be inaccurate.

Generate Agent instructions to onboard AI onto your codebase.

main.py +

Describe what to build next

Agent Auto

TIMELINE

0 0 0

The VS Code terminal displays the configuration for the `ingest-user-activity` function. The configuration includes labels, deployment tool, name, service config, and environment variables. A message is being ingested from a Pub/Sub topic, and the terminal shows the message details and the function's response.

```

triggerRegion: us-central1
labels:
  deployment-tool: cli-gcloud
name: projects/pavanproject-481304/locations/us-central1/functions/ingest_user_activity
satisfiesPz1: true
serviceConfig:
  allTrafficOnLatestRevision: true
  revision: ingest-user-activity-00002-hoc
  availableMemory: 256M
  environmentVariables:
    LOG_EXECUTION_ID: 'true'
  ingressSettings: ALLOW_ALL
  maxInstanceCount: 12
  maxInstanceRequestConcurrency: 1
  timeoutSeconds: 60
state: ACTIVE
updateTime: '2025-12-30T06:49:57.458963842Z'
url: https://us-central1-pavanproject-481304.cloudfunctions.net/ingest_user_activity

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>gcloud pubsub topics publish user-activity --message '{"event":"LOGIN","ts":"2025-01-01T10:00:00Z"}'
messageIds:
- '17553284450460227'

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>gcloud pubsub subscriptions pull user-activity-dlq-sub --auto-ack --limit=5
+-----+-----+-----+-----+-----+-----+-----+-----+
|          DATA          | MESSAGE_ID | ORDERING_KEY | ATTRIBUTES | DELIVERY_ATTEMPT | ACK_STATUS |
+-----+-----+-----+-----+-----+-----+-----+-----+
| [{"event":"LOGIN","ts":"2025-01-01T10:00:00Z"}] | 17553284450460227 |          |          |          | SUCCESS    |
+-----+-----+-----+-----+-----+-----+-----+

```

The VS Code terminal shows a SQL query being executed in a BigQuery shell. The query selects user events from the `realtime_sujatha` table. The results are displayed in a table format.

```

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>bq query --use_legacy_sql=false "SELECT * FROM realtime_sujatha.user_events_sujatha ORDER BY ts DESC LIMIT 5"
+-----+-----+-----+-----+
| event | user | ts |
+-----+-----+-----+-----+
| LOGIN | sujatha | 2025-01-01 13:00:00 |
| LOGIN | sujatha | 2025-01-01 10:00:00 |
+-----+-----+-----+-----+

(venv) C:\Users\SUJATHA\Downloads\Finals Assement>

```

The Google Cloud console displays the details of the `ingest-user-activity` function. The function is in the `us-central1` region and is configured with a scaling of `Auto (min: 0, max: 12)`. The `Logs` tab is selected, showing a list of log entries with their severity, timestamp, and summary.

Severity	Timestamp	Summary
INFO	2025-12-30 11:28:04.443 IST	GET 200 292 B 3 ms Chrome 101 https://ingest-user-activity-7ve7vosmua-uc.a.run.app/robots.txt
INFO	2025-12-30 11:34:42.797 IST	POST 200 130 B 86 ms APIs-Google: (+https://developers.google.com/cloud/functions/docs/concepts/https-overview) https://ingest-user-activity-7ve7vosmua-uc.a.run...
ERROR	2025-12-30 11:34:42.810 IST	Invalid message: Expecting value: line 1 column 1 (char 0)
INFO	2025-12-30 12:18:56.543 IST	Cloud Run ReplaceInternalService ingest-user-activity service-494389106009cf-admin-robot.i... (@type: type-goog...
INFO	2025-12-30 12:19:31.448 IST	Cloud Run ReplaceInternalService ingest-user-activity service-494389106009cf-admin-robot.i... (@type: type-goog...
INFO	2025-12-30 12:19:37.678 IST	Starting new instance. Reason: DEPLOYMENT_ROLLOUT - Instance started due to traffic shifting between revisions due ...
INFO	2025-12-30 12:19:39.988 IST	Default STARTUP TCP probe succeeded after 1 attempt for container 'worker' on port 8080.
INFO	2025-12-30 12:19:40.120 IST	Cloud Run ReplaceInternalService ingest-user-activity-00002-hoc (@type: type.googleapis.com/google.cloud.audit...
INFO	2025-12-30 12:19:56.262 IST	Cloud Run ReplaceInternalService ingest-user-activity (@type: type.googleapis.com/google.cloud.audit.AuditLog...
INFO	2025-12-30 12:21:41.779 IST	POST 200 130 B 104 ms APIs-Google: (+https://developers.google.com/cloud/functions/docs/concepts/https-overview) https://ingest-user-activity-7ve7vosmua-uc.a.run...
ERROR	2025-12-30 12:21:41.795 IST	Invalid message: Missing required fields
INFO	2025-12-30 12:22:38.242 IST	POST 200 130 B 988 ms APIs-Google: (+https://developers.google.com/cloud/functions/docs/concepts/https-overview) https://ingest-user-activity-7ve7vosmua-uc.a.run...

## Practical Question 3 — Event-Driven Data Ingestion

### Objective

To implement a real-time, event-driven data ingestion pipeline on Google Cloud Platform that captures user activity events using Pub/Sub, processes them with Cloud Functions (Gen 2), and stores the data in BigQuery for analytics.

## Architecture

Application → Pub/Sub → Cloud Functions (Gen 2) → BigQuery

### Step 1: Create Pub/Sub Topic

A Pub/Sub topic is created to receive real-time user activity events.

```
gcloud pubsub topics create honnalesujatha_user_events
```

### Step 2: Create BigQuery Dataset

A BigQuery dataset is created to store analytics data.

```
bq mk honnalesujatha_analytics_ds
```

### Step 3: Create BigQuery Table

A table is created inside the dataset to store user activity events.

#### Schema:

- event → STRING
- user → STRING
- ts → TIMESTAMP

```
bq mk --table honnalesujatha_analytics_ds.honnalesujatha_user_activity \
event:STRING,user:STRING,ts:TIMESTAMP
```

### Step 4: Develop Cloud Function (Gen 2)

A Cloud Function is written in Python to:

- Receive messages from Pub/Sub
- Decode Base64 data
- Parse JSON message
- Insert records into BigQuery

### Step 5: Develop Cloud Function (Gen 2)

The Cloud Function is deployed with Pub/Sub trigger.

```
gcloud functions deploy honnalesujatha_user_function \
--gen2 \
--runtime=python311 \
--region=us-central1 \
--trigger-topic=honnalesujatha_user_events \
```

```
--entry-point=pubsub_to_bigquery
```

## Step:6 Publish Test Message

```
gcloud pubsub topics publish honnalesujatha_user_events \
```

```
--message='{ "event": "LOGIN", "user": "honnalesujatha", "ts": "2025-01-01T10:00:00Z" }'
```

## Step 7: Verify Data in BigQuery

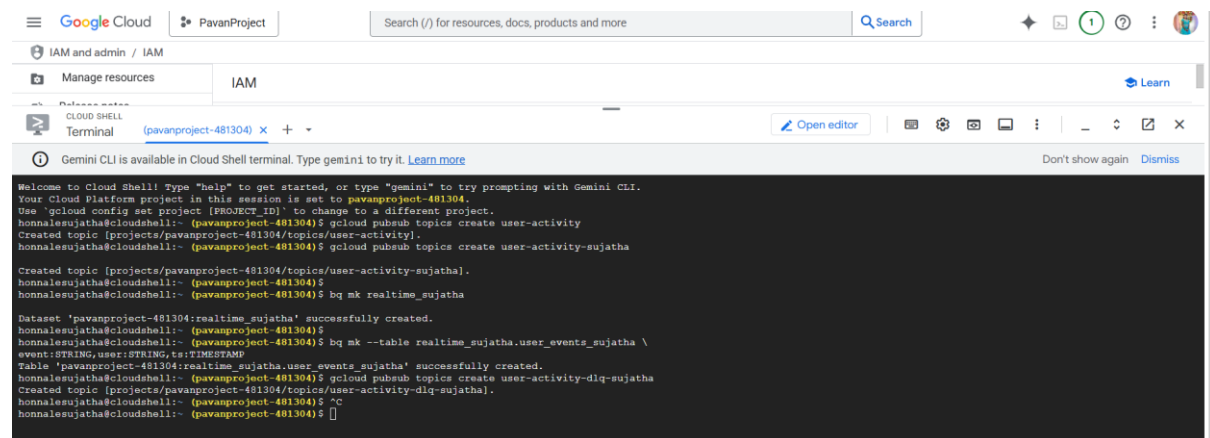
```
bq query --use_legacy_sql=false \
```

```
"SELECT * FROM `pavanproject-481304.honnalesujatha_analytics_ds.honnalesujatha_user_activity`
```

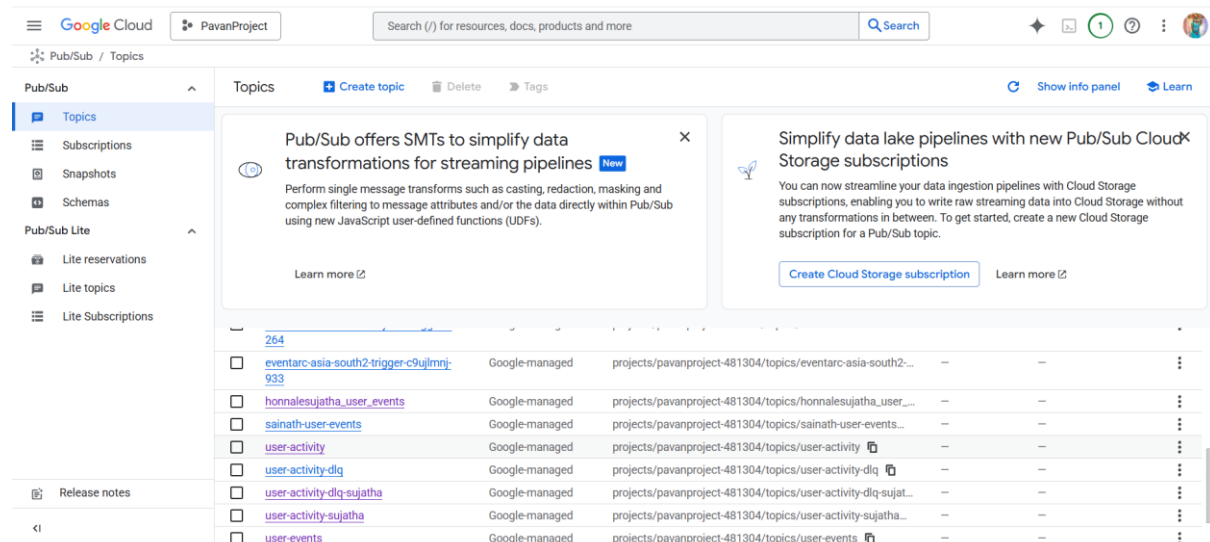
```
ORDER BY ts DESC LIMIT 1"
```

## Step 8: Verify Cloud Function Logs

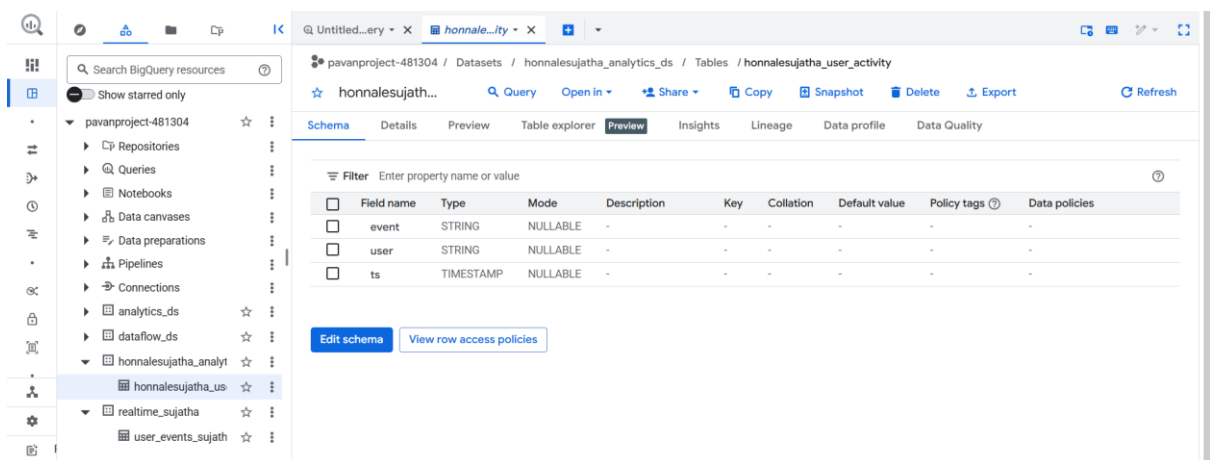
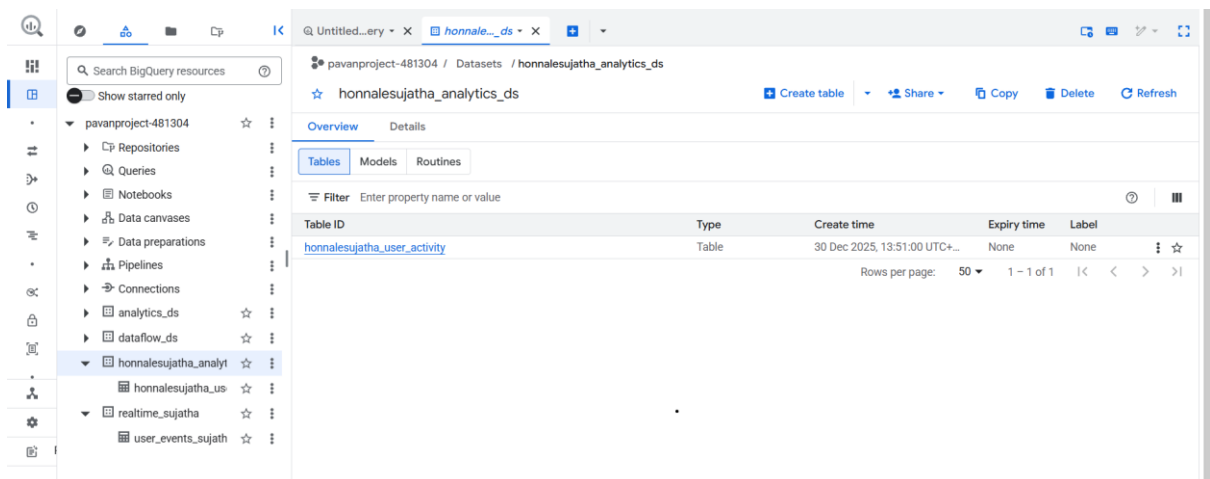
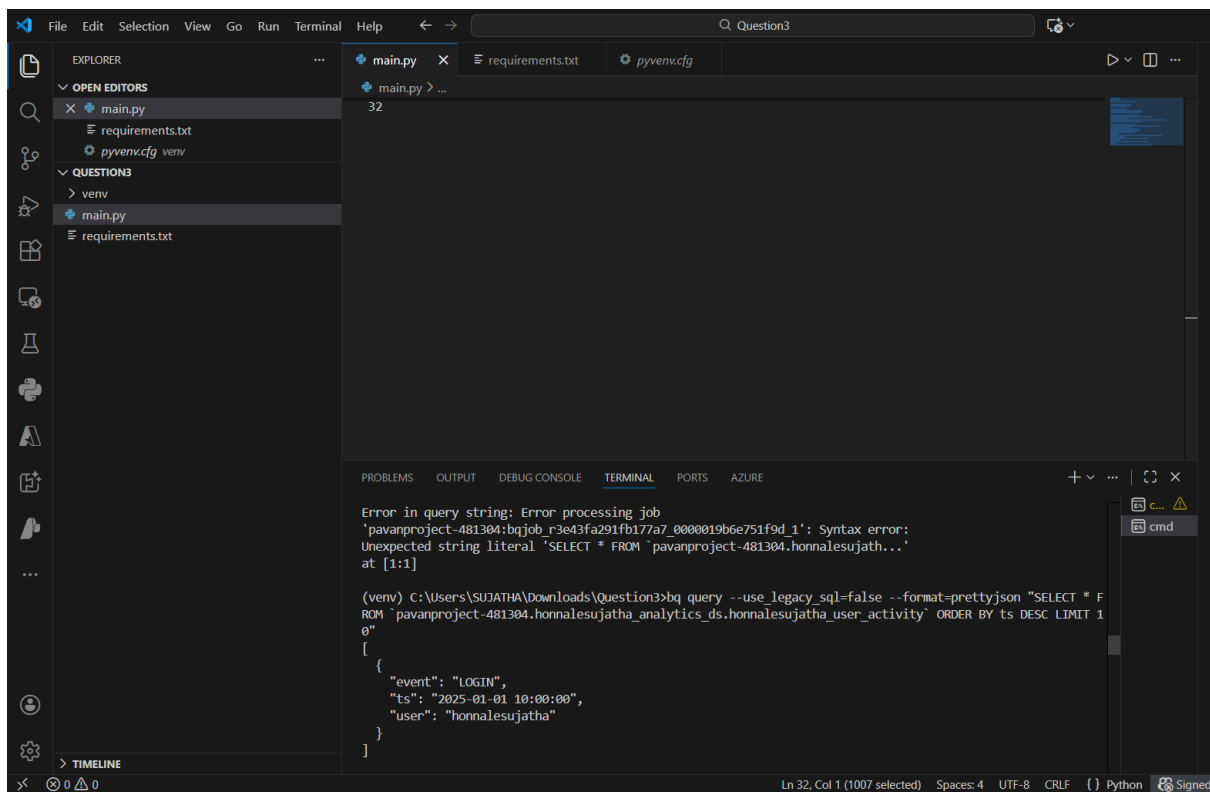
```
gcloud functions logs read honnalesujatha_user_function --gen2 --region=us-central1 --limit=20
```



```
Google Cloud PavanProject Search (/) for resources, docs, products and more
IAM and admin / IAM
Manage resources IAM
Cloud Shell Terminal (pavanproject-481304)
Gemini CLI is available in Cloud Shell terminal. Type gemini to try it.
Welcome to Cloud Shell! Type "help" to get started, or type "gemini" to try prompting with Gemini CLI.
Your Cloud Platform project in this session is set to pavanproject-481304.
Use 'gcloud config set project [PROJECT_ID]' to change to a different project.
honnalesujatha@cloudshell:~ (pavanproject-481304) $ gcloud pubsub topics create user-activity
Created topic [projects/pavanproject-481304/topics/user-activity].
honnalesujatha@cloudshell:~ (pavanproject-481304) $ gcloud pubsub topics create user-activity-sujatha
Created topic [projects/pavanproject-481304/topics/user-activity-sujatha].
honnalesujatha@cloudshell:~ (pavanproject-481304) $ bq mk realtime_sujatha
Dataset 'pavanproject-481304:realtime_sujatha' successfully created.
honnalesujatha@cloudshell:~ (pavanproject-481304) $ bq mk --table realtime_sujatha.user_events_sujatha \
event:STRING,user:STRING,ts:TIMESTAMP
Table 'pavanproject-481304:realtime_sujatha.user_events_sujatha' successfully created.
honnalesujatha@cloudshell:~ (pavanproject-481304) $ gcloud pubsub topics create user-activity-diq-sujatha
Created topic [projects/pavanproject-481304/topics/user-activity-diq-sujatha].
honnalesujatha@cloudshell:~ (pavanproject-481304) $ ^C
honnalesujatha@cloudshell:~ (pavanproject-481304) $
```



Pub/Sub	Topics	Create topic	Delete	Tags	Show info panel	Learn
Topics	Pub/Sub offers SMTs to simplify data transformations for streaming pipelines. Learn more					
Subscriptions	Simplify data lake pipelines with new Pub/Sub Cloud Storage subscriptions. Learn more					
Snapshots						
Schemas						
Pub/Sub Lite						
Lite reservations						
Lite topics						
Lite Subscriptions						
	264					
	eventarc-asia-south2-trigger-c9ujlmnj-933	Google-managed	projects/pavanproject-481304/topics/eventarc-asia-south2-...	--	--	
	honnalesujatha_user_events	Google-managed	projects/pavanproject-481304/topics/honnalesujatha_user_...	--	--	
	sainath-user-events	Google-managed	projects/pavanproject-481304/topics/sainath-user-events...	--	--	
	user-activity	Google-managed	projects/pavanproject-481304/topics/user-activity	--	--	
	user-activity-diq	Google-managed	projects/pavanproject-481304/topics/user-activity-diq	--	--	
	user-activity-diq-sujatha	Google-managed	projects/pavanproject-481304/topics/user-activity-diq-sujat...	--	--	
	user-activity-sujatha	Google-managed	projects/pavanproject-481304/topics/user-activity-sujatha...	--	--	
	user-events	Google-managed	projects/pavanproject-481304/topics/user-events	--	--	





Google Cloud BigQuery interface showing a table named `honnalesujatha_user_activity`. The table has the following schema:

Row	event	user	ts
1	LOGIN	honnalesujatha	2025-01-01 10:00:00 UTC

Google Cloud Cloud Run interface showing a list of services. The service `honnalesujatha-user-function` is highlighted.

Name	Deployment type	Req/sec	Region	Authentication	Ingress	Last deployed	Deployer
<code>honnalesujatha-user-function</code>	Function	0	us-central1	Require authentication	All	1 hour ago	Cloud
<code>ingest-user-activity</code>	Function	0	us-central1	Public access	All	3 hours ago	Cloud
<code>sainath</code>	Function	0	asia-south2	Require authentication	All	2 hours ago	reddy
<code>sainathmyapp</code>	Function	0	asia-south2	Public access	All	1 hour ago	reddy
<code>sujatha</code>	Function	0.04	asia-south2	Public access	All	3 hours ago	pavan

Google Cloud Cloud Run interface showing the logs for the service `honnalesujatha-user-function`. The logs show a successful startup and a deployment rollout.

Severity	Timestamp	Summary
INFO	2025-12-30 13:57:23.667 IST	Default STARTUP TCP probe succeeded after 1 attempt for container "worker" on port 8080.
INFO	2025-12-30 13:57:23.742 IST	Cloud Run ReplaceInternalService honnalesujatha-user-function-00002-dek
INFO	2025-12-30 13:57:25.096 IST	Cloud Run ReplaceInternalService honnalesujatha-user-function
INFO	2025-12-30 14:03:26.308 IST	POST 200 331 B 27 ms APIs-Google; (+https://developers.google.com/api-client-library/python/lib/python3.11/site-packages/googleapiclient/http.py:166) https://honnalesujatha-user-function-7e7vosmua-u...
ERROR	2025-12-30 14:03:26.347 IST	Exception on / [POST] Traceback (most recent call last): File "/layers/google.python.pip/pip/python3.11/site-packages/googleapiclient/http.py:166, in _do_request: HTTPError 404: Not Found
INFO	2025-12-30 14:18:57.476 IST	Cloud Run ReplaceInternalService honnalesujatha-user-function service-49438910600gcf-admin-robot.i...
INFO	2025-12-30 14:19:37.015 IST	Cloud Run ReplaceInternalService honnalesujatha-user-function service-49438910600gcf-admin-robot.i...
INFO	2025-12-30 14:19:56.973 IST	Starting new instance. Reason: DEPLOYMENT_ROLLOUT - Instance started due to traffic shifting between revisions due to...
INFO	2025-12-30 14:19:59.433 IST	Default STARTUP TCP probe succeeded after 1 attempt for container "worker" on port 8080.
INFO	2025-12-30 14:19:59.523 IST	Cloud Run ReplaceInternalService honnalesujatha-user-function-00003-sip
INFO	2025-12-30 14:20:00.886 IST	Cloud Run ReplaceInternalService honnalesujatha-user-function
INFO	2025-12-30 14:21:43.188 IST	POST 200 130 B 1.1 s APIs-Google; (+https://developers.google.com/api-client-library/python/lib/python3.11/site-packages/googleapiclient/http.py:166) https://honnalesujatha-user-function-7e7vosmua-u...
INFO	2025-12-30 14:21:44.272 IST	Inserted: [{"event": "LOGIN", "user": "honnalesujatha", "ts": "2025-01-01T10:00:00Z"}]

## Conclusion

The event-driven data ingestion pipeline was successfully implemented using Google Cloud managed services. User activity events were ingested in real time using Pub/Sub, processed using Cloud Functions (Gen 2), and stored in BigQuery for analytics and querying.

## Practical Question 1 — Event-Driven File Processing

### Objective

To implement an automated event-driven system that processes files uploaded to Google Cloud Storage and stores file metadata in a database.

### Architecture Used

- **Cloud Storage** – File upload trigger
- **Cloud Functions (Gen 2)** – Event processing
- **Cloud Firestore (Native Mode)** – Metadata storage
- **Cloud Logging** – Execution logs

### Step 1: Create Cloud Storage Bucket

```
gsutil mb -l us-central1 gs://file-upload-lab
```

### Step 2: Enable Required APIs

```
gcloud services enable cloudfunctions.googleapis.com eventarc.googleapis.com  
cloudbuild.googleapis.com firestore.googleapis.com
```

### Step 3: Create Firestore Database

```
gcloud firestore databases create --region=us-central
```

### Step 4: Deploy Cloud Function (Gen 2)

```
gcloud functions deploy file_upload_trigger `
--gen2 `
--runtime=python311 `
--region=us-central1 `
--source=. `
--entry-point=file_upload_trigger `
--trigger-event-filters="type=google.cloud.storage.object.v1.finalized" `
--trigger-event-filters="bucket=file-upload-lab"
```

### Step 5: Upload File to Trigger Function

```
"Test file upload" | Out-File test.txt
```

```
gsutil cp test.txt gs://file-upload-lab/
```

Step 6: Verify Cloud Logging

gcloud functions logs read file\_upload\_trigger --region=us-central1

Cloud Storage

Overview

Buckets

Monitoring

Settings

Storage Intelligence

Insights datasets

Configuration

Marketplace

Release notes

Bucket details

sujatha-file-upload-lab

Location

asia-south2 (Delhi)

Storage class

Standard

Public access

Not public

Protection

Soft delete

Hierarchical namespace

Enabled

Objects

ConfigurationPermissionsProtectionLifecycleObservabilityInventory ReportsOperations

Folder browser

sujatha-file-upload-lab

Bucket details

sujatha-file-upload-lab

Create folderUploadTransfer dataOther services

Filter by name prefix onlyFilterFilter objects and foldersShowLive objects only

	Name	Size	Type	Created	Storage class	Last m
	actor.csv	547 B	text/csv	30 Dec 2025, 11:34:00	Standard	30 Dec
	banglore.csv	938 KB	text/csv	30 Dec 2025, 11:38:35	Standard	30 Dec

All databasesDatabase sujatha

Panel view

Query builder

/ > uploaded\_files > esf9y5rvvgeH2mE4tILL

sujatha

+ Start collection

uploaded\_files

uploaded\_files

+ Add document

esf9y5rvvgeH2mE4tILL

esf9y5rvvgeH2mE4tILL

+ Start collection

+ Add field

bucket: "sujatha-file-upload-lab"

file\_name: "actor.csv"

file\_size: 547

Cloud Run

Overview

Services

Jobs

Worker pools

Domain mappings

Service details

Edit and deploy new revision

Connect to repo

Test

Learn

Refresh

Deploying revision

Hide status

Building source (see logs)

Completed

Updating service

Completed

Creating revision

Completed

Routing traffic

Completed

✓

sujatha

Region: asia-south2

URL: <https://sujatha-49438910600.asia-south2.run.app>

Scaling: Auto (min: 0, max: 3)

Observability

Revisions

Source

Triggers

Networking

Security

YAML

Metrics

Logs

SLOs

Errors

Severity

Default

Filter

Search all fields and values

Severity	Timestamp	Summary
✓	2025-12-30 11:33:51.516 IST	Cloud Run deployment succeeded
⚠	2025-12-30 11:33:51.516 IST	POST 500 331 B 19 ms APIs-Google; (+https://developers.goo... https://sujatha-7ve7vosmua-em.a.run.app/7..._GCP_Cloud...
✖	2025-12-30 11:33:51.535 IST	[2025-12-30 06:03:51.529] ERROR in app: Exception on / [POST]
⚠	2025-12-30 11:33:51.535 IST	Traceback (most recent call last): File "/layers/google.python.pip/pip/lib/python3.10/site-packages/flask/app.py", li...
⚠	2025-12-30 11:34:00.919 IST	POST 200 130 B 3 ms APIs-Google; (+https://developers.goo... https://sujatha-7ve7vosmua-em.a.run.app/7..._GCP_Cloud...
✖	2025-12-30 11:34:00.922 IST	Event ID: 17573647956168046
✖	2025-12-30 11:34:00.922 IST	Event type: google.cloud.storage.object.v1.finalized
✖	2025-12-30 11:34:00.922 IST	Bucket: sujatha-file-upload-lab
✖	2025-12-30 11:34:00.923 IST	File: actor.csv
✖	2025-12-30 11:34:00.923 IST	Metageneration: 1
✖	2025-12-30 11:34:00.923 IST	Created: 2025-12-30T06:04:00.155Z
✖	2025-12-30 11:34:00.923 IST	Updated: 2025-12-30T06:04:00.155Z

Release notes

◀

## Conclusion

This practical demonstrates a complete **event-driven file processing pipeline** in Google Cloud using **Cloud Storage, Cloud Functions (Gen 2), Firestore, and Cloud Logging**. The system automatically processes uploaded files and stores metadata reliably.

GitHub link

<https://github.com/HonnaleSujatha/GCP-Assessment>