


Big Data

GCP Analytics





Cloud Scheduler:





Cloud Scheduler


Jobs


 CREATE JOB


 REFRESH

 EDIT


 COPY

 PAUSE

 RESUME


 DELETE


SCHEDULER JOBS






Filter

Filter jobs





<input type="checkbox"/>	Name 	Status of last execution	Region	State	Description	Frequency	Target	Last run	Next run	Action
<input type="checkbox"/>	flight-scheduler1	 Success	us-central1	Enabled		0 * / 1 * * * (America/Indianapolis)	URL : https://us-central1-driven-binder-366618.cloudfunctions.net/flight-streaming-1	Dec 5, 2022, 3:00:00 PM	Dec 5, 2022, 4:00:00 PM	

Cloud Scheduler is a fully managed enterprise-grade cron job scheduler. Here, I have set it to an interval of 1 hour so that the live data is loaded into the bucket on an hourly basis.

Cloud Scheduler ← flight-scheduler1

Define the schedule

Region

us-central1

Description

Frequency *

0 */1 * * *

Schedules are specified using unix-cron format. E.g. every minute: "* * * * *", every 3 hours: "0 */3 * * *", every Monday at 9:00: "0 9 * * 1". [Learn more](#)

Minute:

◦ Equal to 0

and Hour:

◦ From 0 to 23

Timezone *

Eastern Standard Time (EST)

Cloud Function:

Cloud Functions

Functions

CREATE FUNCTION

REFRESH

RELEASE NOTES

LEARN

Filter

Filter functions

<input type="checkbox"/>	<input checked="" type="radio"/>	Environment	Name ↑	Last deployed	Region	Trigger	Runtime	Memory allocated	Executed function	Authentication ?	Actions
<input type="checkbox"/>	<input checked="" type="radio"/>	1st gen	flight-streaming-1	Dec 3, 2022, 6:07:30 PM	us-central1	HTTP	Python 3.10	128 MB	flightStreaming	Allow unauthenticated	<div></div>

A serverless execution environment for constructing and linking cloud services is called Google Cloud Functions. We can create straightforward, one-purpose functions with Cloud Functions that are linked to events released by our cloud infrastructure and services. When an event being watched fires, the function is called.

Here, I have created a cloud function 'flight-streaming-1'. The properties are shown below:

Cloud Functions

Create function

Environment

1st gen

Function name *

flight-streaming-1

Region

us-central1

Trigger

HTTP

Trigger URL

https://us-central1-driven-binder-366618.cloudfunctions.net/flight-streaming-1

Authentication

Allow unauthenticated invocations

Require HTTPS

The entry point is 'flightStreaming' which will trigger the function.

HTTP

Trigger URL

<https://us-central1-driven-binder-366618.cloudfunctions.net/flight-streaming-1>

HTTPS not required

METRICS

DETAILS

SOURCE

VARIABLES

TRIGGER

PERMISSIONS

LOGS

TESTING

Runtime: Python 3.10

Entry point: flightStreaming

requirements.txt

flight

__init__.py

stream

__init__.py

openSkyParser.py

opensky_api.py

```

1 import functions_framework
2 from flight.stream.openSkyParser import parse
3
4 @functions_framework.http
5 def flightStreaming(request):
6     """HTTP Cloud Function.
7     Args:
8         request (flask.Request): The request object.
9         <https://flask.palletsprojects.com/en/1.1.x/api/#incoming-request-data>
10    Returns:
11        The response text, or any set of values that can be turned into a
12        Response object using `make_response`
13        <https://flask.palletsprojects.com/en/1.1.x/api/#flask.make_response>.
14    """
15    return parse(request)

```

Logs are generated everytime the function is called.

flight-streaming-1

1st gen

Version

Version 2, deployed at Dec 3, 2022, 6:07:30 PM...

METRICS

DETAILS

SOURCE

VARIABLES

TRIGGER

PERMISSIONS

LOGS

TESTING

Logs

150 log entries

Severity

Default

Filter

Search all fields and values

?

SEVERITY	TIMESTAMP	SUMMARY		
	2022-12-04 15:55:00.440 EST	flight-streaming-1	4g1pt6yod28	function execution started
>	2022-12-04 15:55:00.447 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:00.447Z, /workspace/flight/stream/openSkyParser.py:96, DEBUG, openSkyParser, _getMessageJSON...
>	2022-12-04 15:55:00.447 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:00.447Z, /workspace/flight/stream/openSkyParser.py:104, DEBUG, openSkyParser, _getMessageJSON...
>	2022-12-04 15:55:00.447 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:00.447Z, /workspace/flight/stream/openSkyParser.py:424, INFO, openSkyParser, parse: {"log": ...
>	2022-12-04 15:55:00.448 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:00.447Z, /workspace/flight/stream/openSkyParser.py:428, INFO, openSkyParser, parse: {"log": ...
>	2022-12-04 15:55:00.448 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:00.448Z, /workspace/flight/stream/openSkyParser.py:333, DEBUG, openSkyParser, _scavengeRows: ...
>	2022-12-04 15:55:00.448 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:00.448Z, /workspace/flight/stream/openSkyParser.py:308, DEBUG, openSkyParser, _getLatestFlig...
>	2022-12-04 15:55:04.913 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:04.913Z, /workspace/flight/stream/openSkyParser.py:349, DEBUG, openSkyParser, _scavengeRows: ...
>	2022-12-04 15:55:04.913 EST	flight-streaming-1	4g1pt6yod28	2022-12-04 20:55:04.913Z, /workspace/flight/stream/openSkyParser.py:349, DEBUG, openSkyParser, _scavengeRows: ...

Data Metrics for the cloud function:



Cloud Storage:

The batch and streaming data would be loaded into the folder I created in the bucket. Here, ‘scheduled-streaming-data’ folder would be filled with data regularly based on our frequency.

Buckets > 2022_abhi_bigdata

UPLOAD FILES UPLOAD FOLDER CREATE FOLDER TRANSFER DATA MANAGE HOLDS DOWNLOAD DELETE

Filter by name prefix only Filter Filter objects and folders Show deleted data

<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified	Public access	Version history	E
<input type="checkbox"/>	data/	—	Folder	—	—	—	—	—	—
<input type="checkbox"/>	flight-streaming/	—	Folder	—	—	—	—	—	—
<input type="checkbox"/>	scheduled-streaming-data/	—	Folder	—	—	—	—	—	—
<input type="checkbox"/>	testfunction.zip	11.6 KB	application/zip	Dec 3, 2022, 2:24:21 PM	Standard	Dec 3, 2022, 2:24:21 PM	Not public	—	G

Live data gets uploaded into the folder every 1 hour based on our settings.

Buckets > 2022_abhi_bigdata > scheduled-streaming-data

UPLOAD FILES UPLOAD FOLDER CREATE FOLDER TRANSFER DATA MANAGE HOLDS DOWNLOAD DELETE

Filter by name prefix only Filter Filter objects and folders Show deleted data

<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified	Public access	Version history	Enc
<input type="checkbox"/>	2022-12-03_23-24-52_0	449 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_1	424 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_2	446 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_3	446 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_4	433 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_5	428 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_6	424 B	text/plain	Dec 3, 2022, 6:24:54 PM	Standard	Dec 3, 2022, 6:24:54 PM	Not public	—	Goc
<input type="checkbox"/>	2022-12-03_23-24-54_7	449 B	text/plain	Dec 3, 2022, 6:24:55 PM	Standard	Dec 3, 2022, 6:24:55 PM	Not public	—	Goc

Dataflow:

I ran the following batch and streaming jobs to get the data into our folder in the bucket.

Jobs CREATE JOB FROM TEMPLATE ENABLE SORTING REFRESH LEARN

Running Filter Filter jobs

Name	Type	End time	Elapsed time	Start time	Status	SDK version	ID	Region
streaming_GCS_PS3	Streaming		7 hr 40 min	Dec 4, 2022, 9:35:53 AM	Running	2.41.0	2022-12-04_06_35_52-311974886284103074	us-central1
streaming_ps_bq	Streaming		7 hr 58 min	Dec 4, 2022, 9:17:34 AM	Running	2.41.0	2022-12-04_06_17_32-16280172724710077510	us-central1
streaming2_GCS_PS2	Batch	Dec 4, 2022, 9:26:35 AM	4 min 39 sec	Dec 4, 2022, 9:21:56 AM	Succeeded	2.41.0	2022-12-04_06_21_56-5755714518047960403	us-central1
streaming2_GCS_PS	Batch	Dec 4, 2022, 8:53:47 AM	5 min 31 sec	Dec 4, 2022, 8:48:16 AM	Succeeded	2.41.0	2022-12-04_05_48_16-8184400348264625087	us-central1

Batch Dataflow:

To read txt files from Google Cloud Storage to PubSub, I have used a Batch Job in Dataflow using the “Text files on CloudStorage to PubSub” data template.

streaming2_GCS_PS2 CLONE STOP IMPORT AS PIPELINE SHARE

JOB GRAPH EXECUTION DETAILS JOB METRICS RECOMMENDATIONS

Job steps view Graph view CLEAR SELECTION

Logs HIDE 1

JOB LOGS WORKER LOGS DIAGNOSTICS

31 log entries Severity Debug Filter Search all fields and values MAX TIME

SEVERITY	TIMESTAMP	SUMMARY
>	2022-12-04 09:25:56.518 EST	Cleaning up.
>	2022-12-04 09:25:56.563 EST	Starting worker pool teardown.
>	2022-12-04 09:25:56.593 EST	Stopping worker pool...
>	2022-12-04 09:26:30.754 EST	Autoscaling: Resized worker pool from 1 to 0.
>	2022-12-04 09:26:30.801 EST	Worker pool stopped.
>	2022-12-04 09:26:30.829 EST	Tearing down pending resources...

Job info

Job name

streaming2_GCS_PS2

Job ID

2022-12-04_06_21_56-5755714518047960403

Job type

Batch

Job status

Succeeded

SDK version

Apache Beam SDK for Java 2.41.0
A newer version of the SDK family exists and updating is recommended. [Learn more](#)

Job region

us-central1

Worker location

us-central1

Current workers

0

Latest worker status

Worker pool stopped.

Start time

December 4, 2022 at 9:21:56 AM GMT-5

Elapsed time

4 min 40 sec

Encryption type

Google-managed key

Dataflow Prime

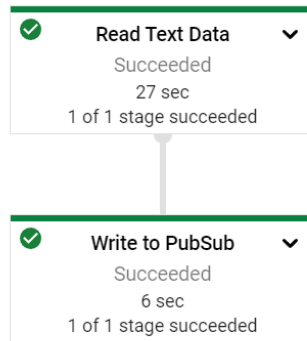
Disabled

Runner v2

Disabled

Dataflow Shuffle

Enabled

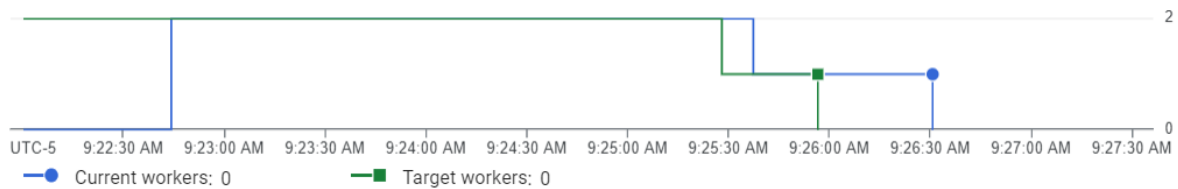


- streaming2_GCS_PS2 [CLONE](#) [STOP](#) [+ IMPORT AS PIPELINE](#) [SHARE](#)

[job list page](#) [EXECUTION DETAILS](#) [JOB METRICS](#) [RECOMMENDATIONS](#)

☐ Auto Refresh [MAX TIME](#) [v](#)

Autoscaling [?](#)



Latest worker status: Worker pool stopped.

[Metrics](#)

[I](#)

Throughput

[EXPORT TO CLOUD DASHBOARD](#)

[OVERVIEW METRICS](#)

[Throughput](#)

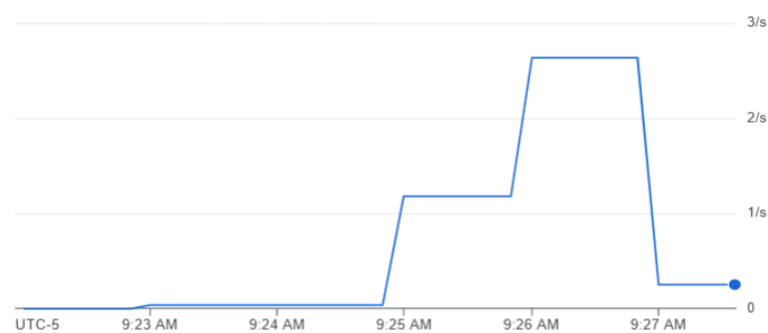
[Errors](#)

[RESOURCE METRICS](#)

[CPU utilization](#)

[Memory utilization](#)

Throughput (elements/sec) [?](#) [v](#)



[Name](#)

[Value](#)

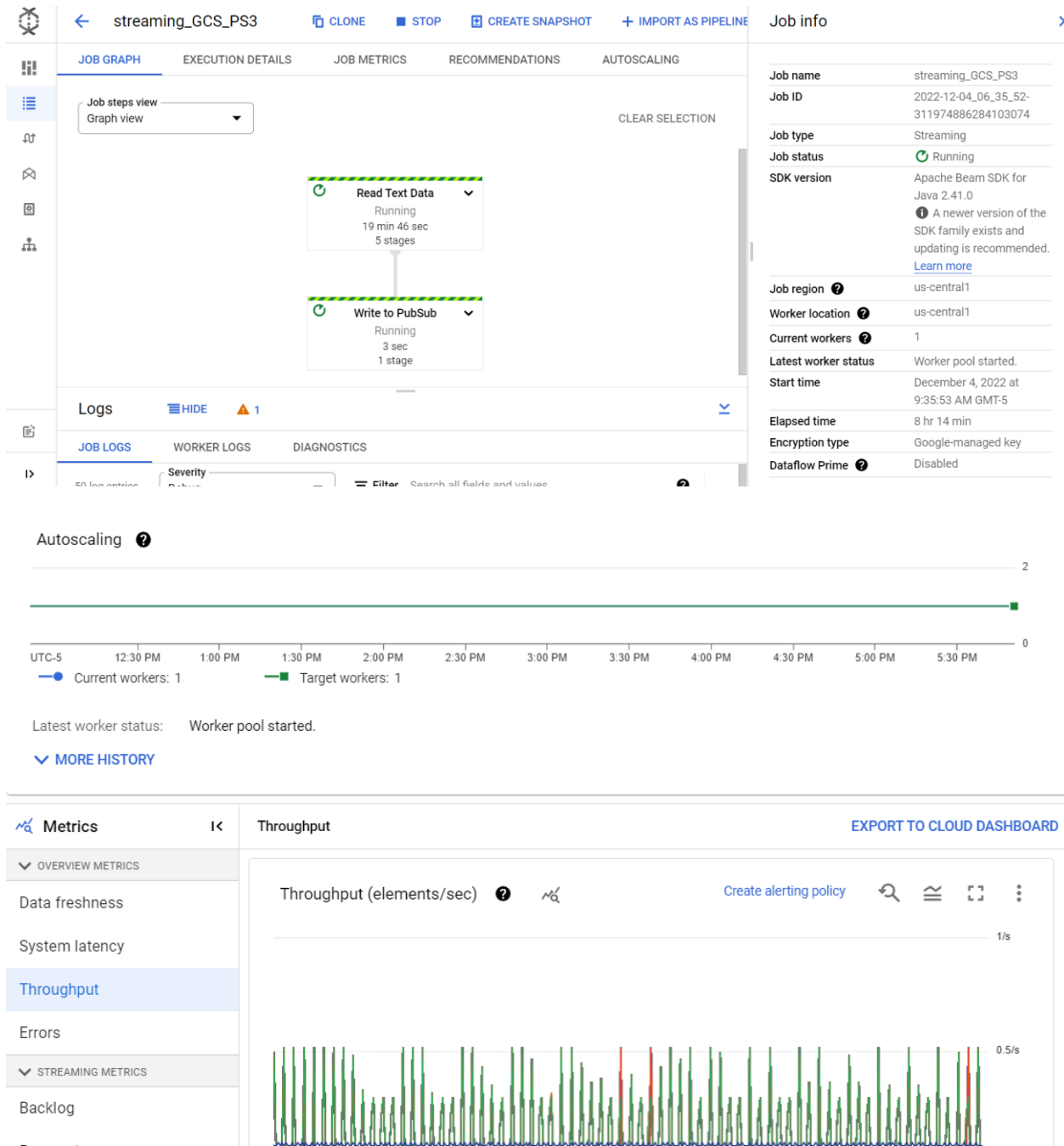
[Read Text Data/Read](#)

0.25/s

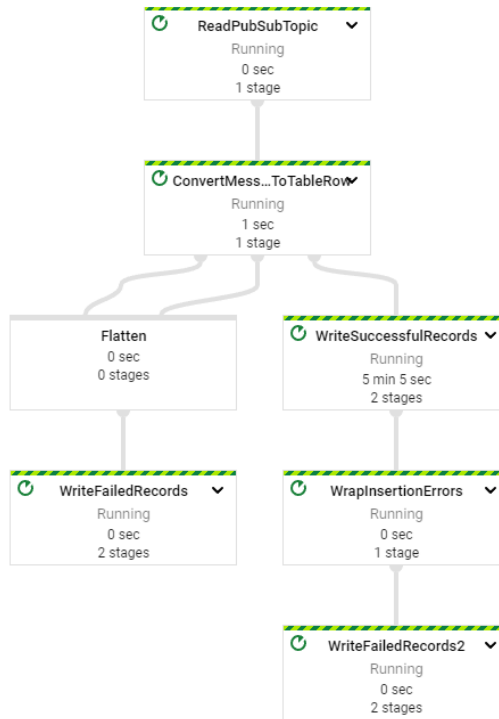
Streaming dataflow:

To ingest messages from a PubSub topic, transform them, and write them to a BigQuery table that I created by using the provided schema I have used a streaming job in Data flow. The template used is “PubSub Topic to BigQuery”.

GCS to Pub/Sub:



Pub/Sub to BigQuery:



streaming_ps_bq

CLONE

STOP

CREATE SNAPSHOT

IMPORT AS PIPELINE

SHARE

Up to Dataflow job list page

EXECUTION DETAILS

JOB METRICS

RECOMMENDATIONS

AUTOSCALING

Graph view

Stage progress

Filter stages

F32, F33, F34, and F35

Auto Refresh

6 HOURS

Logs

HIDE

JOB LOGS

WORKER LOGS

DIAGNOSTICS

50 log entries

Severity

Debug

Filter

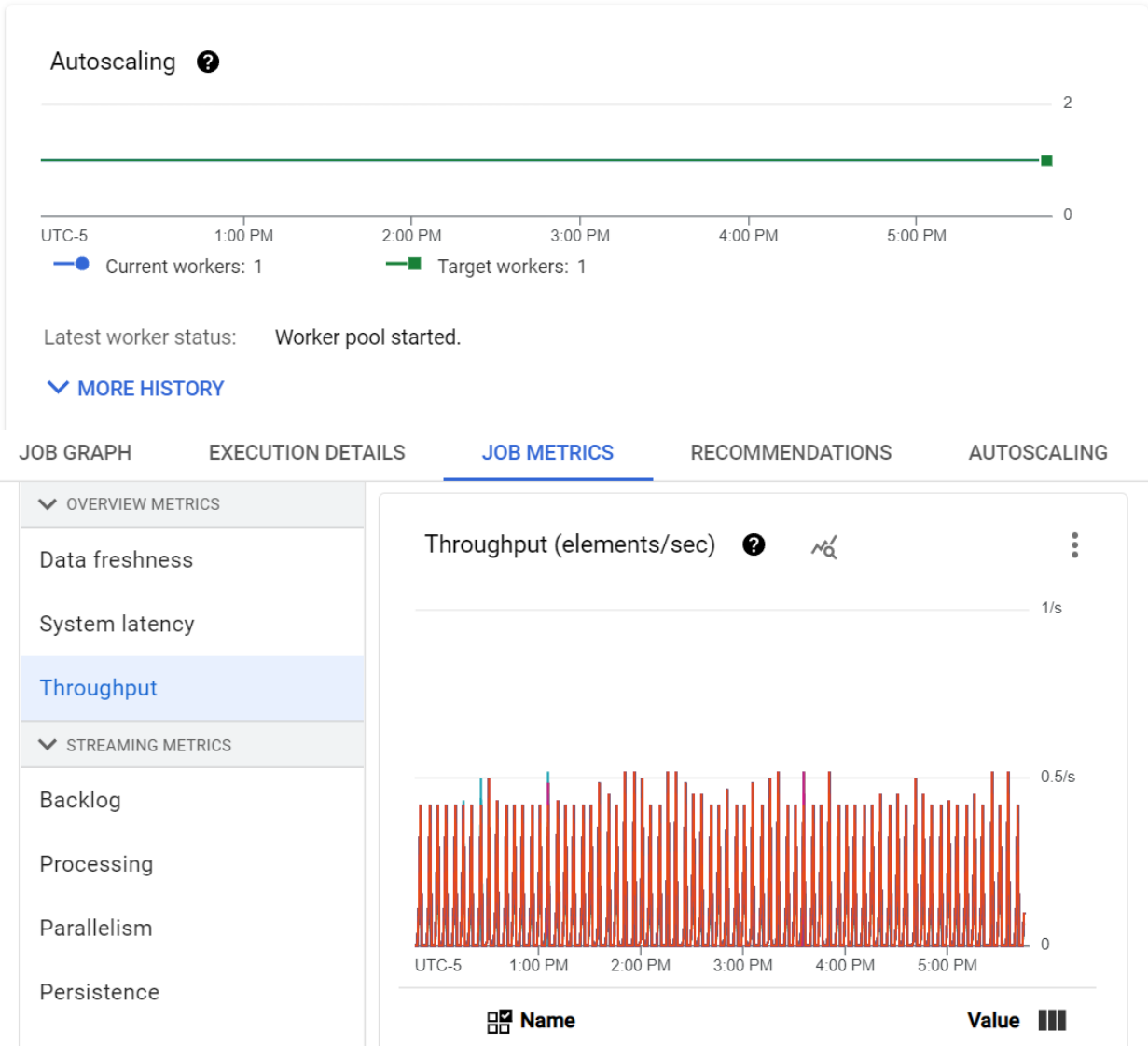
Search all fields and values

6 HOURS

SEVERITY	TIMESTAMP	SUMMARY
Scanned up to 12/4/22, 9:17 AM. Scanned 61.1 KB.		
>	2022-12-04 09:17:51.532 EST	Fusing consumer WriteFailedRecords/WriteFailedRecordsToBigQuery/PrepareWrite/ParDo(An...
>	2022-12-04 09:17:51.567 EST	Fusing consumer ReadPubSubTopic/MapElements/Map into ReadPubSubTopic/PubsubUnboundedS...
>	2022-12-04 09:17:51.600 EST	Fusing consumer ConvertMessageToTableRow/MapToRecord into ReadPubSubTopic/MapElements...
>	2022-12-04 09:17:51.637 EST	Fusing consumer ConvertMessageToTableRow/InvokeUDF/ProcessUdf into ConvertMessageToTa...
>	2022-12-04 09:17:51.670 EST	Fusing consumer ConvertMessageToTableRow/JsonToTableRow/JsonToTableRow into ConvertMe...
>	2022-12-04 09:17:51.706 EST	Fusing consumer WriteSuccessfulRecords/PrepareWrite/ParDo(Anonymous) into ConvertMess...

Job info

Job name	streaming_ps_bq
Job ID	2022-12-04_06_17_32-16280172724710077510
Job type	Streaming
Job status	Running
SDK version	Apache Beam SDK for Java 2.41.0 A newer version of the SDK family exists and updating is recommended. Learn more
Job region	us-central1
Worker location	us-central1
Current workers	1
Latest worker status	Worker pool started.
Start time	December 4, 2022 at 9:17:34 AM GMT-5
Elapsed time	8 hr 28 min
Encryption type	Google-managed key
Dataflow Prime	Disabled
Runner v2	Disabled
Streaming Engine	Disabled
Vertical Autoscaling	Disabled



Pub/Sub:

I created a PubSub topic to which messages were sent from the data flow job. The messages were in JSON and these were then stored as entries in a BigQuery table.

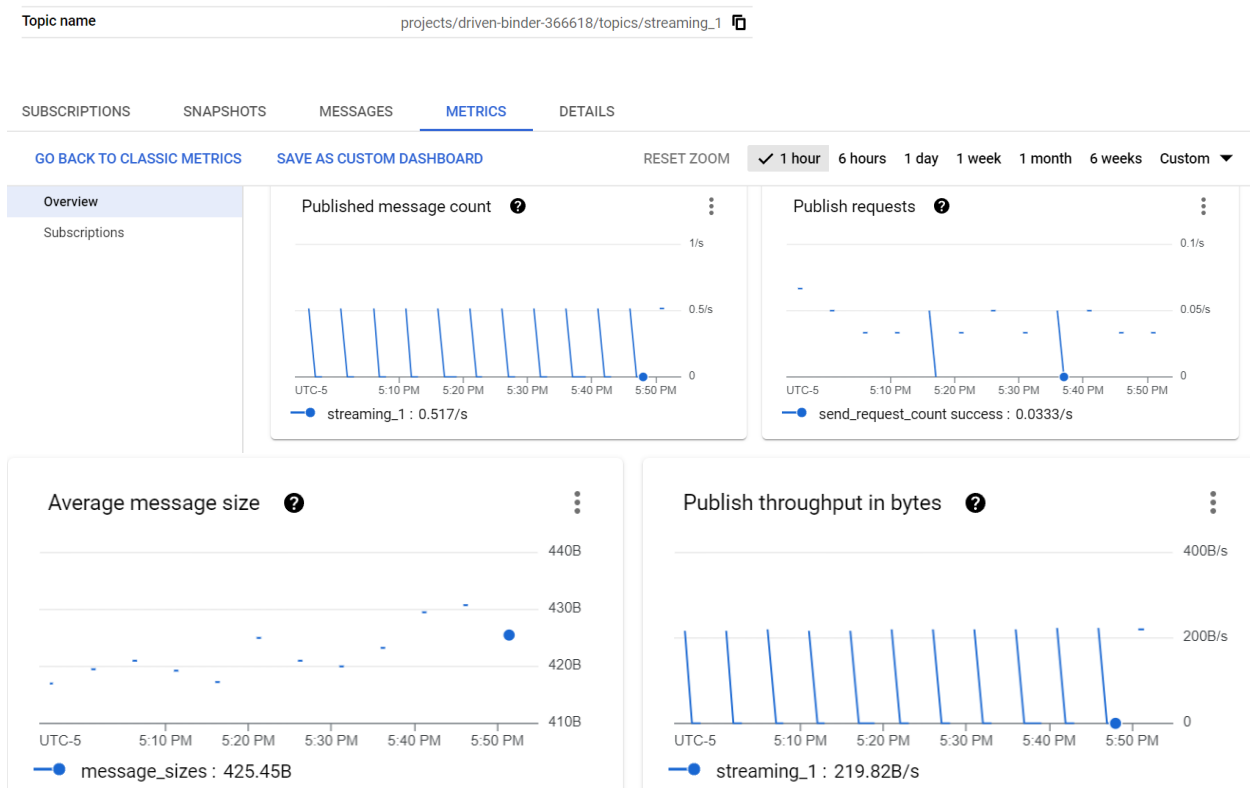
Topic:

Topics [+ CREATE TOPIC](#) [DELETE](#) [SHOW INFO PANE](#)

[LIST](#) METRICS

Filter Filter topics ?

<input type="checkbox"/>	Topic ID ↑	Encryption key	Topic name	Retention	
<input type="checkbox"/>	streaming_1	Google-managed	projects/driven-binder-366618/topics/streaming_1	—	⋮ ▼



Subscription:

Subscriptions [+ CREATE SUBSCRIPTION](#) [DELETE](#) [SHOW INFO PANEL](#)

LIST METRICS

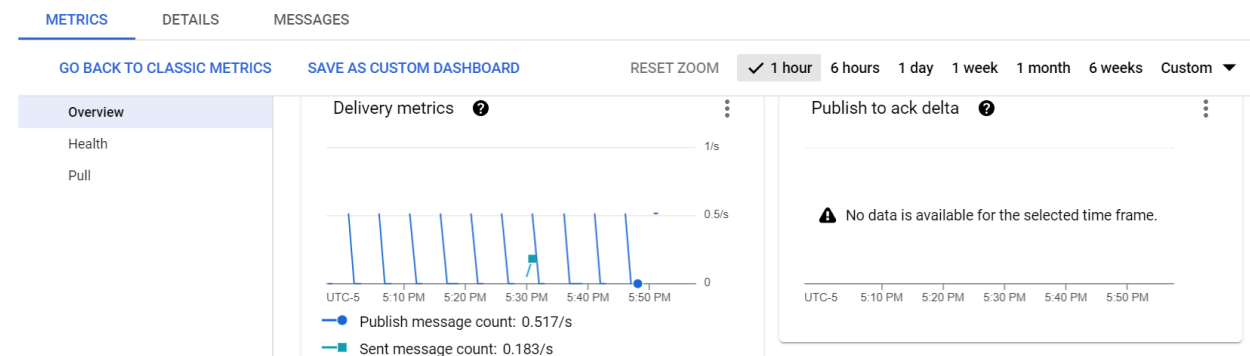
Filter Filter subscriptions

<input type="checkbox"/>	State	Subscription ID ↑	Delivery type	Topic name	Ack deadline	Retention	Message ordering	
<input type="checkbox"/>	✓	streaming_1-sub	Pull	projects/driven-binder-36...	10 seconds	7 days	Disabled	⋮ ▼
<input type="checkbox"/>	✓	streaming_1.subscription-6204211500883520792	Pull	projects/driven-binder-36...	1 minute	7 days	Disabled	⋮ ▼

Subscription name [projects/driven-binder-366618/subscriptions/streaming_1-sub](#)

Subscription state ✓ active

Topic name [projects/driven-binder-366618/topics/streaming_1](#)



METRICS

DETAILS

MESSAGES

PULL

☐ Enable ack messages

Filter

Filter messages

?

⋮

Publish time	Attribute keys	Message body	Ack ↑
Dec 4, 2022, 4:50:11 PM	—	{\"icao24\": \"4b1812\", \"callsign\": \"SWR9RL\", \"origin\": \"Switzerland\", \"time\": 1670190597, \"contact\": 1670190598, \"longitude\": 2.3283, \"latitude\": 41.6565, \"altitude\": 3939.54, \"on_ground\": false, \"velocity\": 135.49, \"heading\": 230.08, \"vertical_rate\": -4.55, \"baro_altitude\": 3878.58, \"squawk\": 3003, \"spi\": false, \"position_source\": 0, \"time_bq\": \"2022-12-04 21:49:57\", \"contact_bq\": \"2022-12-04 21:49:58\", \"query_time_bq\": \"2022-12-04 21:50:00\"}	Deadline exceeded ⬆

BigQuery:

Data is loaded into the table ‘gcs-ps-dest1’ which is under the dataset ‘streaming_dataset1’.

Explorer

+ ADD DATA

IK

Type to search

?

Viewing all resources. [Show starred resources only.](#)

HW1-Q4_2

HW1_BD

test1_class10

test2_class10

test_1511

2022_abhi_dataset

Models (1)

Flights

aircraft

BD_HW1

Flights_class

streaming_dataset1

gcs-ps-dest1

Flights

*Unsaved query 2

HW1-Q2

*Unsaved query 3

gcs-ps-dest1

gcs-ps-dest1

QUERY

SHARE

COPY

SNAPSHOT

DELETE

EXPORT

SCHEMA

DETAILS

PREVIEW

Row	squawk	spi	baro_altitude	vertical_rate	icao24	on_gro	heading	po
1	7006	true	10058.4	0.33	a1d18f	fal...	56.37	po
2	nuli	fal...	3017.52	0.0	4d0107	fal...	236.1	
3	1437	fal...	8351.52	6.18	a00009	fal...	303.97	
4	nuli	fal...	nuli	nuli	4b1804	true	185.62	
5	nuli	fal...	9753.6	0.0	aa56d5	fal...	282.86	
6	5765	fal...	1493.52	-1.95	4b1806	fal...	358.55	
7	7312	fal...	2499.36	-6.5	abd254	fal...	88.78	
8	nuli	fal...	7840.98	13.66	880458	fal...	342.51	
9	nuli	fal...	nuli	nuli	4b1804	fal...	135.0	
10	nuli	fal...	nuli	0.33	a58213	fal...	341.9	
11	nuli	fal...	11285.22	0.33	7c6b30	fal...	19.04	
12	nuli	fal...	6271.26	-8.13	842194	fal...	189.13	
13	2000	fal...	424.24	nuli	4b1805	true	140.05	

Results per page: 50 1 – 50 of 1922

To conduct further analysis and modeling, I wrote queries in BigQuery.

RUN
SAVE
SHARE
SCHEDULE
MORE
Query completed

```
1 SELECT count(*) FROM `driven-binder-366618.streaming_dataset1.gcs-ps-dest1` LIMIT 1000
```

Press Alt+F1 for Accessibility Option

Query results
SAVE RESULTS
EXPLORE DATA

JOB INFORMATION
RESULTS
JSON
EXECUTION DETAILS
EXECUTION GRAPH
PREVIEW

Row	f0_
1	3596

Data gets uploaded into the table I created in the bucket.

1. BQML logistic and linear regression models to predict, for any two-hour time window, for flights leaving ORD:

Create a logistic regression model to predict how many flights are ascending based on baro_altitude, heading, and velocity

```
1 CREATE OR REPLACE MODEL `2022_abhi_dataset.HW2_LogReg1`
2 OPTIONS ( MODEL_TYPE = 'LOGISTIC_REG',
3           INPUT_LABEL_COLS = ['ASCENDING'],
4           OPTIMIZE_STRATEGY = 'AUTO_STRATEGY',
5           MAX_ITERATIONS = 50,
6           LEARN_RATE_STRATEGY = 'CONSTANT',
7           LEARN_RATE = 0.1,
8           EARLY_STOP = FALSE,
9           DATA_SPLIT_METHOD = 'AUTO_SPLIT',
10          AUTO_CLASS_WEIGHTS = TRUE ) AS
11 SELECT
12     heading,
13     velocity,
14     baro_altitude,
15     ASCENDING
16 FROM
17     `2022_abhi_dataset.Flights` bts
18 INNER JOIN
19     `2022_abhi_dataset.gcs_ps_new2` ls
20 ON
21     (ls.IATA_Code_rep = bts.IATA_CODE_Reporting_Airline
22     AND CAST(bts.Flight_Number_Reporting_Airline AS STRING) = ls.reporting_airlined_code)
23 WHERE
24     OriginAirportID = 13930 --Origin:ORD
25     AND time_bq BETWEEN '2022-12-03 23:18:31 UTC' AND '2022-12-04 01:18:31 UTC';
```

Query results

SAVE RESULTS
 EXPLORE DATA

JOB INFORMATION

RESULTS

JSON

EXECUTION DETAILS

EXECUTION GRAPH

PREVIEW

Row	predicted_ASCE	predict... label	predict... prob	heading	velocity	baro_altitude	ASCENDING
Results per page: 50 1 – 50 of 658 < < > >							

HW2_LogReg1

QUERY MODEL
 DELETE MODEL
 EXPORT MODI

DETAILS

TRAINING

EVALUATION

SCHEMA



▶

RUN

📄

SAVE

👤

SHARE

🕒

SCHEDULE

⚙️

MORE

✔️ This query is valid

```
1 CREATE OR REPLACE MODEL `2022_abhi_dataset.HW2_LinReg1`
2 OPTIONS ( MODEL_TYPE = 'LINEAR_REG',
3           INPUT_LABEL_COLS = ['VELOCITY'],
4           OPTIMIZE_STRATEGY = 'AUTO_STRATEGY',
5           MAX_ITERATIONS = 50,
6           LEARN_RATE_STRATEGY = 'CONSTANT',
7           LEARN_RATE = 0.1,
8           EARLY_STOP = FALSE,
9           DATA_SPLIT_METHOD = 'AUTO_SPLIT') AS
10 (
11   SELECT DISTINCT Flight_Number_Reporting_Airline,
12                  longitude,
13                  latitude,
14                  vertical_rate,
15                  baro_altitude,
16                  velocity
17   FROM
18     `2022_abhi_dataset.Flights` bts
19   INNER JOIN
20     `2022_abhi_dataset.gcs_ps_new2` ls
21   ON
22     (ls.IATA_Code_rep = bts.IATA_CODE_Reporting_Airline
23     AND CAST(bts.Flight_Number_Reporting_Airline AS STRING) = ls.reporting_airlined_code)
24   WHERE
25     OriginAirportID= 13930 --Origin:ORD
26     AND time_bq BETWEEN '2022-12-03 23:18:31 UTC' AND '2022-12-04 01:18:31 UTC'
27 )
```

HW2_LinReg1

DETAILS	TRAINING	EVALUATION	SCHEMA
View as			
<input type="radio"/> Graphs			
<input checked="" type="radio"/> Table			
Iteration	Training Data Loss	Duration (seconds)	
0	2,572.8547	2.31	

DETAILS	TRAINING	EVALUATION	SCHEMA
Mean absolute error			
32.8561			
Mean squared error			
2,572.8547			
Mean squared log error			
0.8143			
Median absolute error			
19.3779			
R squared			
0.4184			

```

1 SELECT * FROM ML.PREDICT(MODEL `2022_abhi_dataset.HW2_LinReg1`,
2 (
3   SELECT DISTINCT Flight_Number_Reporting_Airline,
4     longitude,
5     latitude,
6     vertical_rate,
7     baro_altitude,
8     velocity
9   FROM
10    `2022_abhi_dataset.Flights` bts
11   INNER JOIN
12    `2022_abhi_dataset.gcs_ps_new2` ls
13   ON
14    (ls.IATA_Code_rep = bts.IATA_CODE_Reporting_Airline
15     AND CAST(bts.Flight_Number_Reporting_Airline AS STRING) = ls.reporting_airlined_code)
16  WHERE
17    OriginAirportID= 13930 --Origin:ORD
18    AND time_bq BETWEEN '2022-12-03 23:18:31 UTC' AND '2022-12-04 01:18:31 UTC' ))

```

PI

Query results

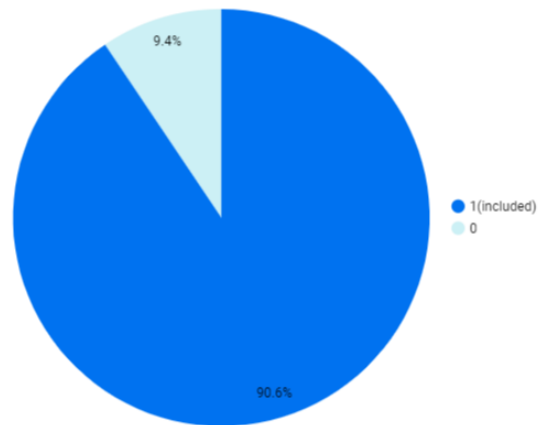
[SAVE RESULTS](#) ▼

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		EXECUTION GRAPH		PREVIEW
row	predicted_VELO	Flight_Number_	longitude	latitude	vertical_rate	baro_altitude	velocity	
1	227.513657...	1286	-112.4781	26.8694	0.0	11582.4	233.62	
2	231.590086...	582	-97.2245	31.7619	-0.33	11871.96	236.16	

2. DataStudio Dashboard to display results



Logistic Regression

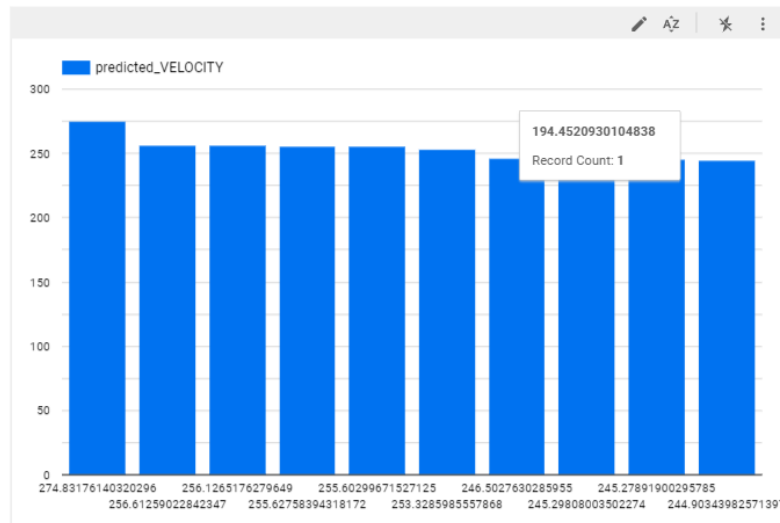


<https://datastudio.google.com/reporting/3405df41-e8ea-4431-93a5-30061585e1aa>

About 90.6% of the prediction is correct for the ascending flights based on the data.



Linear Regression



<https://datastudio.google.com/reporting/69a91202-4175-42aa-bfff-b72182402fdd>