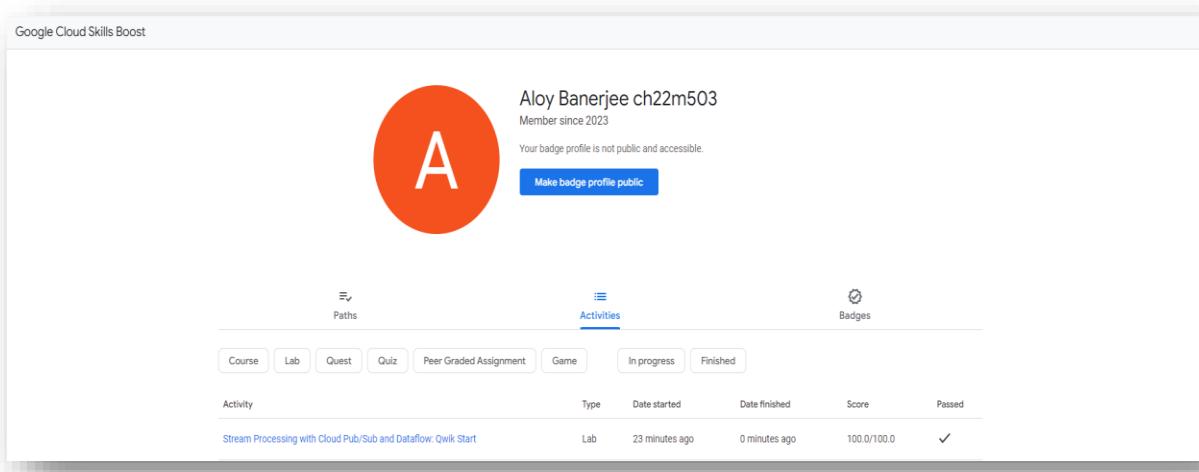


AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

Stream Processing with Cloud Pub/Sub and Dataflow:

Qwik Start: -

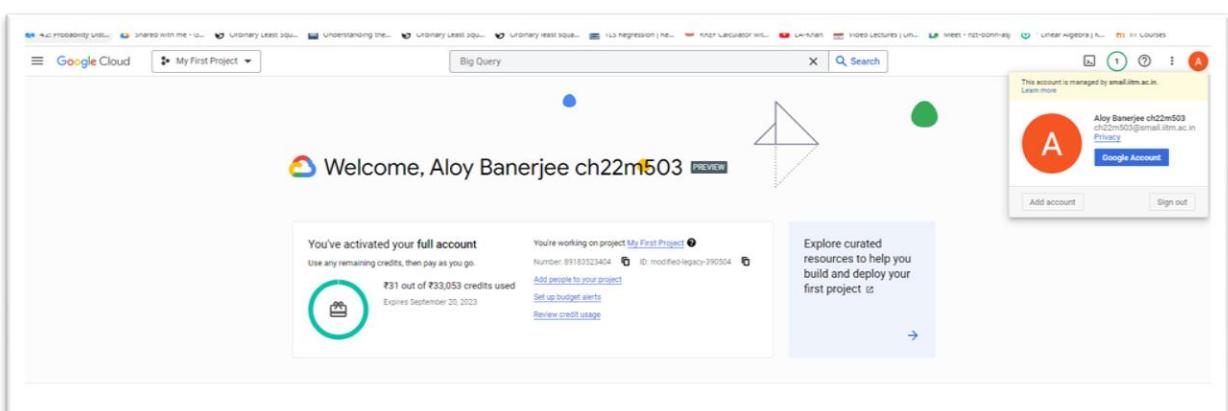
Performing the task in cloud skill boost account:



The screenshot shows a user profile for Aloy Banerjee (ch22m503) on Google Cloud Skills Boost. The profile includes a large orange circular badge with a white letter 'A'. Below it, the name 'Aloy Banerjee ch22m503' and the text 'Member since 2023' are displayed. A button 'Make badge profile public' is visible. The 'Activities' tab is selected, showing a list of completed tasks. One task, 'Stream Processing with Cloud Pub/Sub and Dataflow: Qwik Start', is listed with details: Type: Lab, Date started: 23 minutes ago, Date finished: 0 minutes ago, Score: 100.0/100.0, and Passed status indicated by a checkmark.

Performing the task in my own GCP account:

Showing the credentials:



The screenshot shows the Google Cloud Platform dashboard for the user 'Aloy Banerjee ch22m503'. The dashboard features a 'Welcome' message and a summary of account activation. It also displays the user's current project, 'My First Project', which is active with full credits. On the right side, there is a sidebar with account management options, including a link to 'Google Account' and a sign-out button.

Go to

https://www.cloudskillsboost.google/focuses/18457?catalog_rank=%7B%22rank%22%3A5%2C%22n

AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

um_filters%22%3A1%2C%22has_search%22%3Atrue%7D&parent=catalog&search_id=23687651 for following the steps to be performed for creating the google cloud pub-sub services.

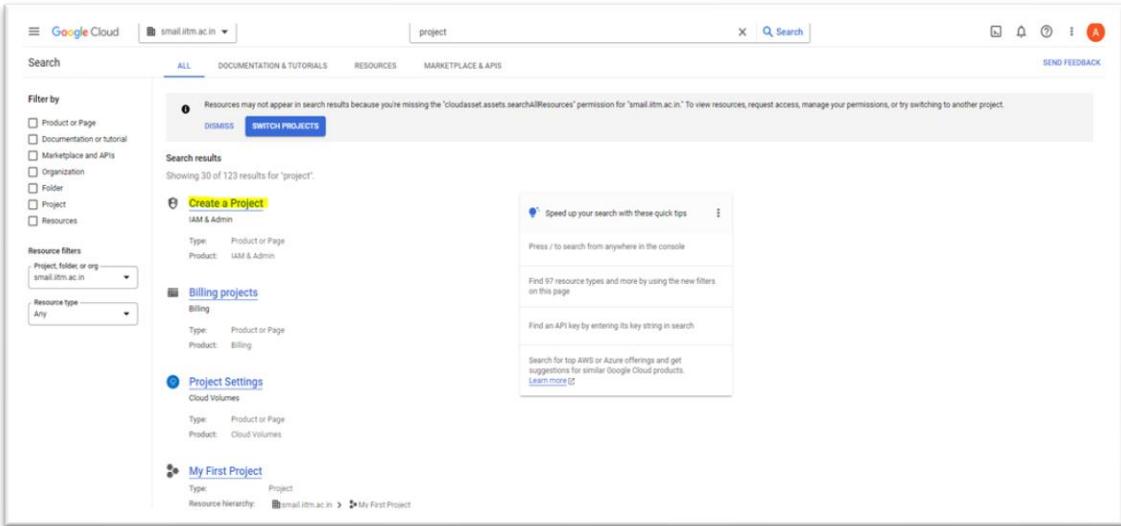
Activities to be performed as part of this assignments are as follows,

- ❖ Read messages published to a Pub/Subtopic
- ❖ Window (or group) the messages by timestamp
- ❖ Write the messages to Cloud Storage

Pre-requisite steps before starting the actual Task,

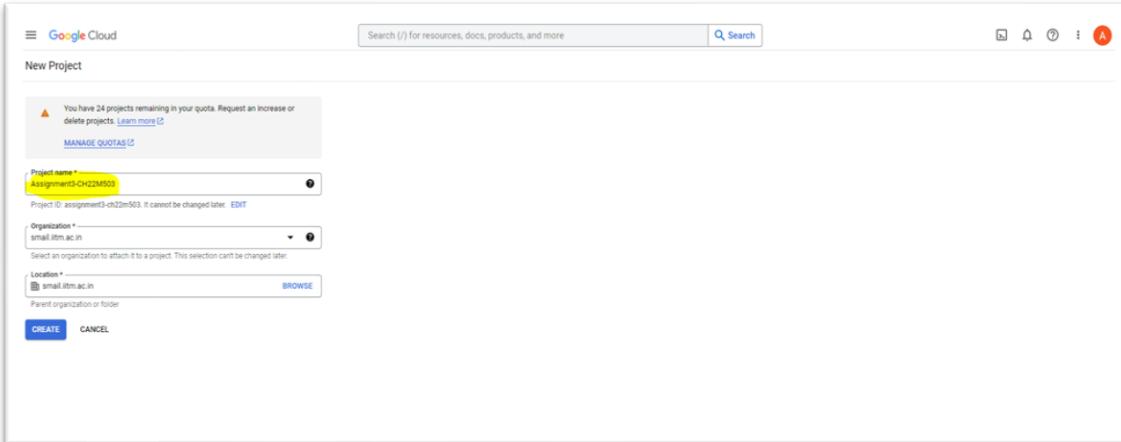
Create a new project under which will perform this task,

1. Search with Project in search bar and click on ‘**FIND**’ button below screen will appear,



The screenshot shows the Google Cloud search interface. The search bar at the top contains the query 'project'. Below the search bar, there are filters on the left for 'Product or Page', 'Documentation or tutorial', 'Marketplace and APIs', 'Organization', 'Folder', 'Project', and 'Resources'. A 'Resource filters' section is also present. The main search results area displays three items: 'Create a Project', 'Billing projects', and 'My First Project'. Each item has a brief description and a 'Type' field. To the right of the search results, there are several quick tips and search-related links.

2. Click on ‘**Create a Project**’ button and fill up the form and click on the ‘**CREATE**’



The screenshot shows the 'New Project' creation form. At the top, it says 'New Project'. Below that, there is a warning message: 'You have 24 projects remaining in your quota. Request an increase or delete projects [learn more](#)'. There is a 'MANAGE QUOTAS' link. The form fields include:

- Project name***: Assignment3-Ch22M503 (highlighted with a yellow box)
- Organization***: small.itm.ac.in
- Location***: small.itm.ac.in

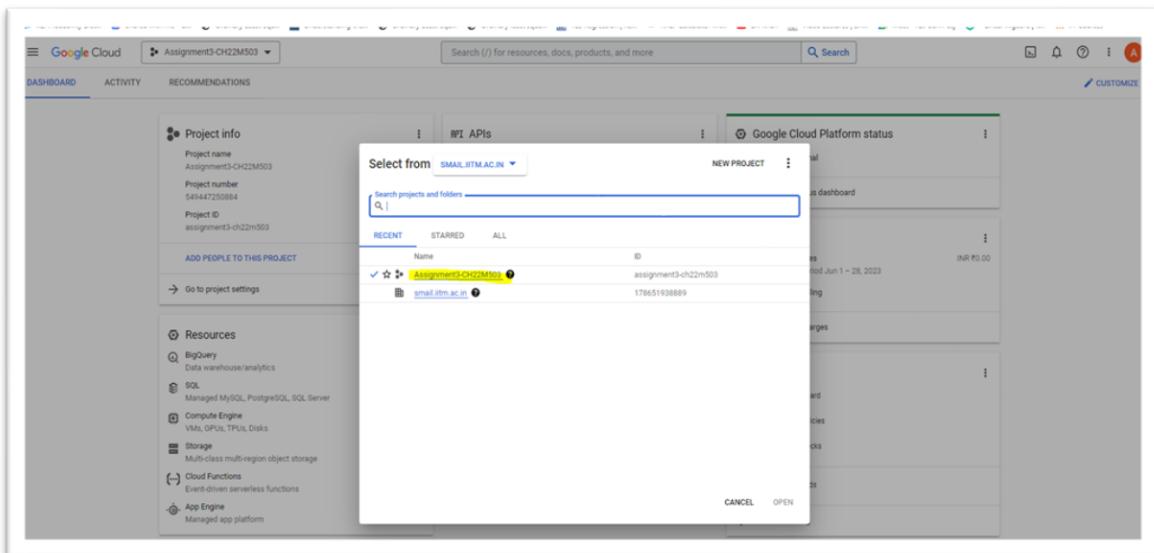
At the bottom of the form are two buttons: 'CREATE' and 'CANCEL'.

3. New project will get created and will be shown in GCP dashboard,

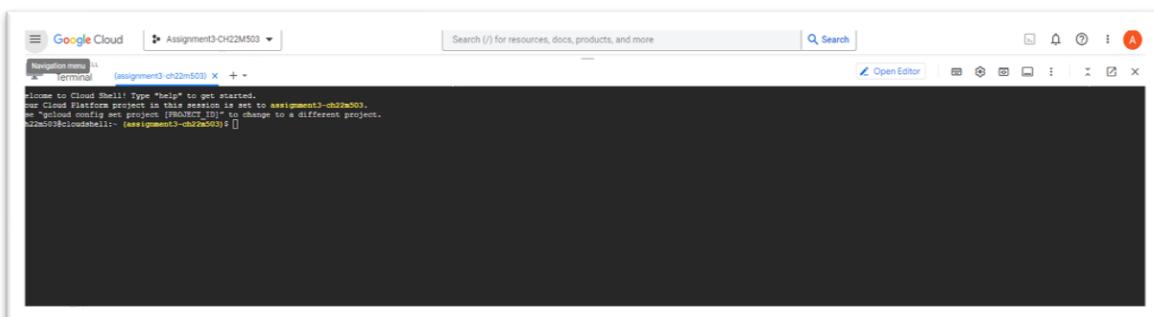
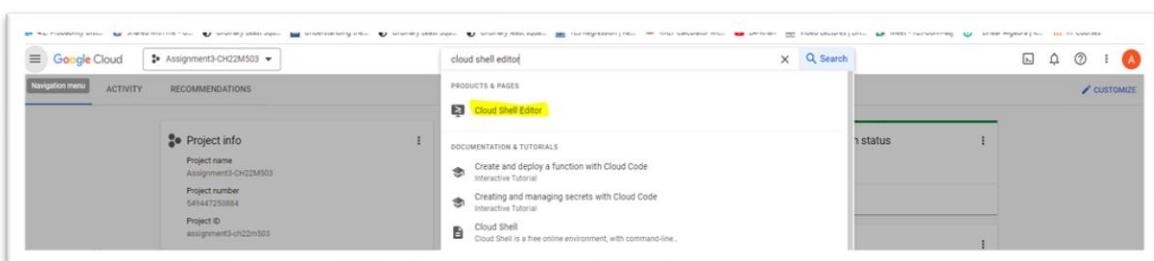
AI Lab 2 (Assignment – 3)

Aloy Banerjee (CH22M503)

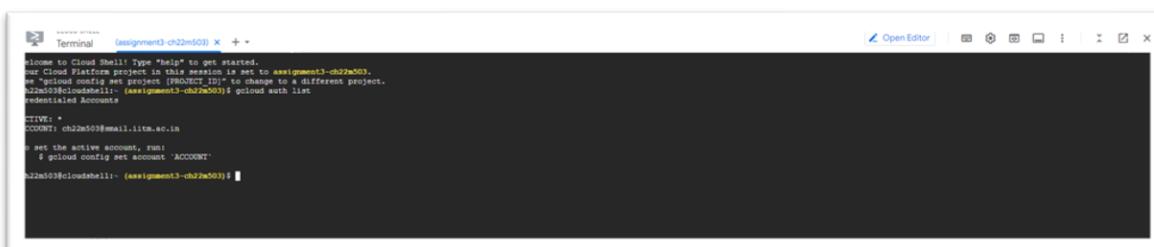
Date of Submission: 04-07-2023



4. Project ID is 'assignment3-ch22m503' in our case for current assignment.
5. Open cloud shell editor for writing the cloud shell command for further operation.



6. Run command 'gcloud auth list' and click on 'Authorize'



7. Run command 'gcloud config list project' it will list out the projects

AI Lab 2 (Assignment – 3)

Aloy Banerjee (CH22M503)

Date of Submission: 04-07-2023



```
Terminal (assignment3-ch22m503) + 
Welcome to Cloud Shell! Type "help" to get started.
our Cloud Platform project in this session is set to assignment3-ch22m503.
or "gcloud config set project [PROJECT_ID]" to change to a different project.
ch22m503@cloudshell:~ (assignment3-ch22m503)$ gcloud auth list
Authenticated accounts:
  * CH22M503 (ch22m503@mail.iitm.ac.in)

Set the active account, run:
  $ gcloud config set account [ACCOUNT]
ch22m503@cloudshell:~ (assignment3-ch22m503)$ gcloud config list project
core
  project = assignment3-ch22m503
our active configuration is: [cloudshell-8632]
ch22m503@cloudshell:~ (assignment3-ch22m503)$
```

- Run command `gcloud config set compute/region "REGION"` to set the region. Use **us-west1** as part of **REGION**.
`gcloud config set compute/region "us-west1"`



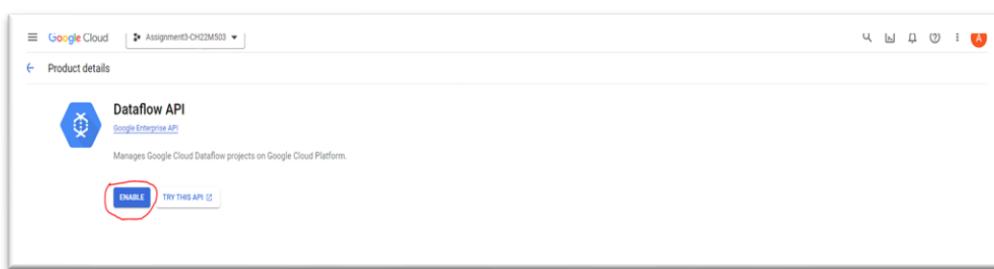
```
Terminal (assignment3-ch22m503) + 
ch22m503@cloudshell:~ (assignment3-ch22m503)$ gcloud config set compute/region "us-west1"
Updated property [compute/region].
ch22m503@cloudshell:~ (assignment3-ch22m503)$
```

Ensure that the Dataflow API is successfully enabled.

- Search with 'DataFlow API' in search bar,

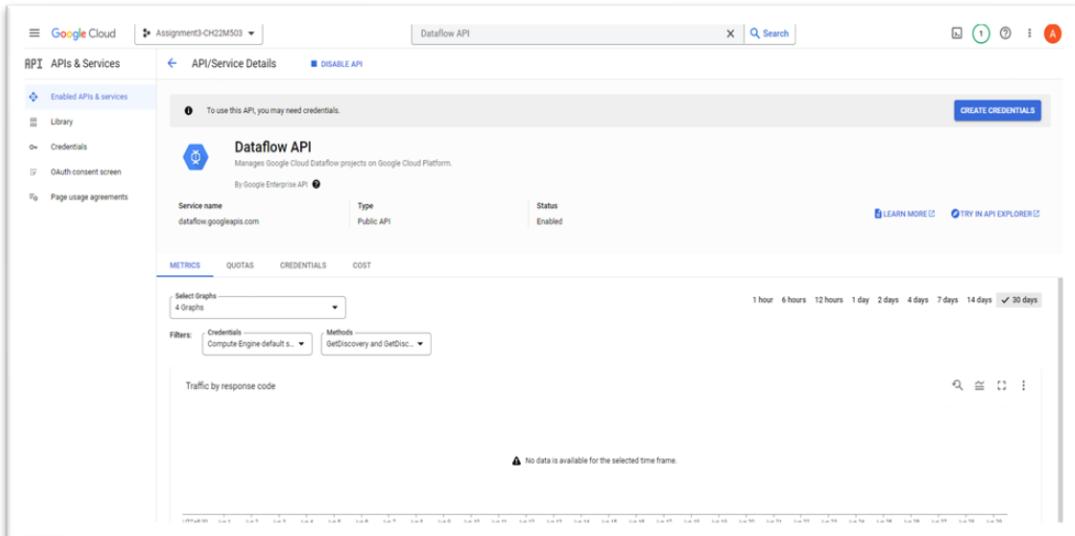


Click on enable to make it enable,



After enabling the same dashboard will open for DataFlow API

AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023



The screenshot shows the Google Cloud API Service Details page for the Dataflow API. The page header includes the project name "Assignment3-CH22M503" and the API name "Dataflow API". A note says "To use this API, you may need credentials." and a "CREATE CREDENTIALS" button is available. The service details show the service name "dataflow.googleapis.com", type "Public API", and status "Enabled". Below this are tabs for "METRICS", "QUOTAS", "CREDENTIALS", and "COST". The "METRICS" tab is selected, showing a chart area with a note "No data is available for the selected time frame." and a legend for traffic by response code. There are also filters for "Credentials" and "Methods".

Task Started Now onwards,

Task 1. Create project resources

1. In Cloud Shell, create variables for your bucket, project, and region.
Run command

```
PROJECT_ID=$(gcloud config get-value project)
BUCKET_NAME="${PROJECT_ID}-bucket"
TOPIC_ID=my-id
REGION="filled in at lab start"
```

```
PROJECT_ID=$(gcloud config get-value project)
BUCKET_NAME="${PROJECT_ID}-bucket"
TOPIC_ID=my-id
REGION=us-west1
```

```
ch22m503@cloudshell:~ (assignment3-ch22m503)$ PROJECT_ID=$(gcloud config get-value project)
BUCKET_NAME="${PROJECT_ID}-bucket"
TOPIC_ID=my-id
REGION=us-west1
Your active configuration is: [cloudshell-18740]
ch22m503@cloudshell:~ (assignment3-ch22m503)$
```

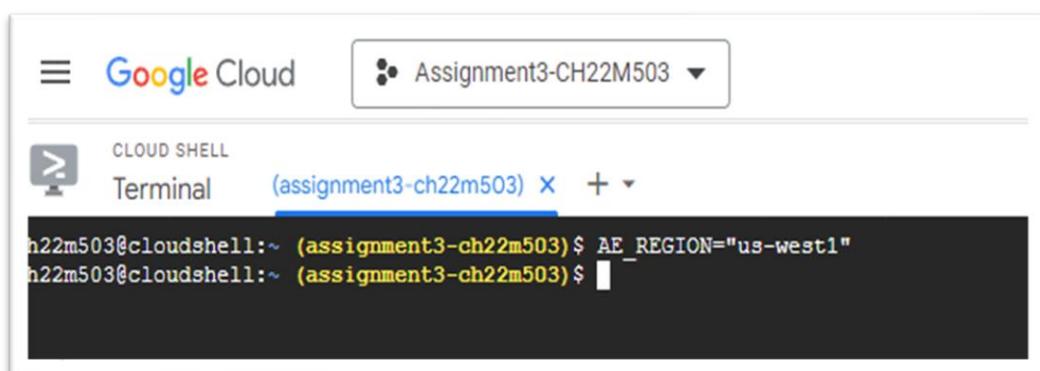
AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

2. Set the App Engine Location by running below command,

```
AE_REGION=
```



➤ AE_REGION="us-west1"



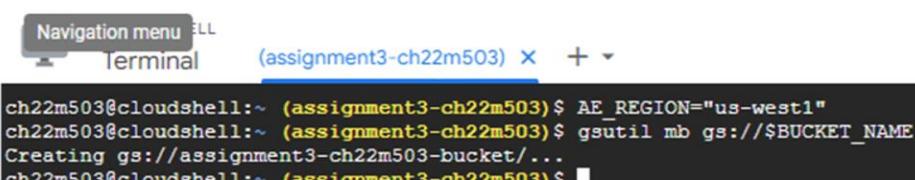
Google Cloud Assignment3-CH22M503

CLOUD SHELL Terminal (assignment3-ch22m503) +

```
h22m503@cloudshell:~ (assignment3-ch22m503)$ AE_REGION="us-west1"
h22m503@cloudshell:~ (assignment3-ch22m503)$
```

3. Create a Cloud Storage bucket owned by this project,

```
gsutil mb gs://$BUCKET_NAME
```



Navigation menu ELL

Terminal (assignment3-ch22m503) +

```
ch22m503@cloudshell:~ (assignment3-ch22m503)$ AE_REGION="us-west1"
ch22m503@cloudshell:~ (assignment3-ch22m503)$ gsutil mb gs://$BUCKET_NAME
Creating gs://assignment3-ch22m503-bucket/...
ch22m503@cloudshell:~ (assignment3-ch22m503)$
```

4. Follow below steps one by one for performing the next set of tasks,

- Create a Pub/Subtopic in this project
- Create a Cloud Scheduler job in this project. The job publishes a message to a Pub/Subtopic at one-minute intervals
- If prompted to enable the Cloud Scheduler API, press **y** and enter
- Start the job

AI Lab 2 (Assignment – 3)

Aloy Banerjee (CH22M503)

Date of Submission: 04-07-2023

```
Cloud Shell Terminal (assignment3-ch22m503) x + v

$ gcloud topics create $TOPIC_ID --region=$REGION
$ gcloud app create --region=$REGION
$ gcloud scheduler jobs create publisher-job --schedule="*/ * * * *"
$ gcloud scheduler jobs run publisher-job

$ gcloud services enable cloudscheduler.googleapis.com
```

Rerun after updating the topic id.

Visit below url to get the information of the running Topics,

[https://console.cloud.google.com/cloudpubsub/topic/list?referrer=search&project=assignment3-ch22m503&pageState=\(%22duration%22:\(%22groupValue%22:%22PT1H%22,%22customValue%22:null\)\)](https://console.cloud.google.com/cloudpubsub/topic/list?referrer=search&project=assignment3-ch22m503&pageState=(%22duration%22:(%22groupValue%22:%22PT1H%22,%22customValue%22:null)))

Filter topics			
Topic ID	Encryption key	Topic name	
CH22M503	Google-managed	projects/assignment3-ch22m503/topics/CH22M503	

- Now to run the code we are doing the cloning of the repository using docker command as below,

AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

Java Python

```
docker run -it -e DEVSHELL_PROJECT_ID=$DEVSHELL_PROJECT_ID python:3.7
git clone https://github.com/GoogleCloudPlatform/python-docs-samples
cd python-docs-samples/pubsub/streaming-analytics
pip install -U -r requirements.txt # Install Apache Beam dependencies
```

Command line output after running the above command,

CLOUD SHELL Terminal (assignment3-ch22m503) +

```
Welcome to Cloud Shell! Type "help" to get started.
Your Cloud Platform project in this session is set to assignment3-ch22m503.
Use "gcloud config set project [PROJECT_ID]" to change to a different project.
ch22m503@cloudshell:~ (assignment3-ch22m503)$ docker run -it -e DEVSHELL_PROJECT_ID=$DEVSHELL_PROJECT_ID python:3.7 /bin/bash
root@c8c9c2eebc989:/# git clone https://github.com/GoogleCloudPlatform/python-docs-samples.git
Cloning into 'python-docs-samples'...
remote: Enumerating objects: 104086, done.
remote: Counting objects: 100% (99/99), done.
remote: Compressing objects: 100% (81/81), done.
remote: Total 104086 (delta 21), reused 79 (delta 11), pack-reused 103987
Receiving objects: 100% (104086/104086), 212.44 MiB | 12.42 MiB/s, done.
Resolving deltas: 100% (62587/62587), done.
Updating files: 100% (4477/4477), done.
root@c8c9c2eebc989:/# cd python-docs-samples/pubsub/streaming-analytics
root@c8c9c2eebc989:/# pip install -U -r requirements.txt # Install Apache Beam dependencies
Collecting apache-beam[gcp,test]==2.42.0
  Downloading apache_beam-2.42.0-cp37-cp37-manylinux2010_x86_64.whl (11.0 MB)
    11.0/11.0 kB 51.6 MB/s eta 0:00:00
Collecting pymongo<4.0.0,>=3.8.0
  Downloading pymongo-3.13.0-cp37-cp37-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (506 kB)
    506.0/506.0 kB 52.2 MB/s eta 0:00:00
Collecting proto-plus<2,>=1.7.1
  Downloading proto_plus-1.22.3-py3-none-any.whl (48 kB)
    48.1/48.1 kB 6.9 MB/s eta 0:00:00
Collecting regex>=2020.6.8
  Downloading regex-2023.6.3-cp37-cp37-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (755 kB)
    755.7/755.7 kB 43.5 MB/s eta 0:00:00
Collecting pyarrow<8.0.0,>=0.15.1
  Downloading pyarrow-7.0.0-cp37-cp37-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (26.7 MB)
    26.7/26.7 kB 46.0 MB/s eta 0:00:00
Collecting protobuf<4,>=3.12.2
  Downloading protobuf-3.20.3-cp37-cp37-manylinux_2_5_x86_64.manylinux1_x86_64.whl (1.0 MB)
    1.0/1.0 kB 47.1 MB/s eta 0:00:00
Collecting typing-extensions>=3.7.0
  Downloading typing_extensions-4.7.0-py3-none-any.whl (33 kB)
Collecting requests<3.0.0,>=2.24.0
  Downloading requests-2.31.0-py3-none-any.whl (62 kB)
    62.6/62.6 kB 9.5 MB/s eta 0:00:00
Collecting crcmod<2.0,>=1.7
  Downloading crcmod-1.7.tar.gz (89 kB)
    89.7/89.7 kB 12.8 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Collecting dill<0.3.2,>=0.3.1.1
  Downloading dill-0.3.1.1.tar.gz (151 kB)
    152.0/152.0 kB 21.2 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Collecting zstandard<1.18.0
  Downloading zstandard-0.21.0-cp37-cp37-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.7 MB)
    2.7/2.7 kB 0.1 MB/s eta 0:00:00
Collecting numpy<1.23.0,>=1.14.3
  Downloading numpy-1.21.6-cp37-cp37-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (15.7 MB)
    15.7/15.7 kB 40.0 MB/s eta 0:00:00
Collecting python-dateutil<3,>=2.8.0
  Downloading python_dateutil-2.8.2-py2.py3-none-any.whl (247 kB)
```

AI Lab 2 (Assignment – 3)

Aloy Banerjee (CH22M503)

Date of Submission: 04-07-2023

```
Navigation menu [L] Terminal (assignment3-ch22m03) x + ▾

Downloading python_dateutil-2.8.2-py2.py3-none-any.whl (247 kB) 267.7/267.7 kB 26.7 MB/s eta 0:00:00
Collecting grpcio!=1.48.0,<2,>=1.33.1
  Downloading grpcio-1.56.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (5.2 MB) 5.2/5.2 kB 69.0 MB/s eta 0:00:00
Collecting pydot<2,>=1.2.0
  Downloading pydot-1.4.2-py2.py3-none-any.whl (21 kB)
Collecting fastavro<2,>=0.23.6
  Downloading fastavro-1.7.4-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.4 MB) 2.4/2.4 kB 60.5 MB/s eta 0:00:00
Collecting httplib2<0.21.0,>=0.8
  Downloading httplib2-0.20.4-py3-none-any.whl (96 kB) 96.0/96.6 kB 13.1 MB/s eta 0:00:00
Collecting orjson<4.0
  Downloading orjson-3.9.1-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (136 kB) 136.9/136.9 kB 18.0 MB/s eta 0:00:00
Collecting cloudpickle==2.1.0
  Downloading cloudpickle-2.1.0-py3-none-any.whl (25 kB)
Collecting hdfs<3.0.0,>=2.1.0
  Downloading hdfs-2.7.0-py3-none-any.whl (34 kB)
Collecting pytz>=2018.3
  Downloading pytz-2023.3-py2.py3-none-any.whl (502 kB) 502.3/502.3 kB 36.5 MB/s eta 0:00:00
Collecting pytest<8.0,>=7.1.2
  Downloading pytest-7.4.0-py3-none-any.whl (323 kB) 323.6/323.6 kB 34.1 MB/s eta 0:00:00
Collecting parameterized<0.9.0,>=0.7.1
  Downloading parameterized-0.8.1-py2.py3-none-any.whl (26 kB)
Collecting pytest-timeout<3,>=2.1.0
  Downloading pytest_timeout-2.1.0-py3-none-any.whl (12 kB)
Collecting sqlalchemy<2.0,>=1.3
  Downloading SQLAlchemy-1.4.48-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2014_x86_64.whl (1.6 MB) 1.6/1.6 kB 55.6 MB/s eta 0:00:00
Collecting pyhamcrest!=1.10.0,<2.0.0,>=1.9
  Downloading PyHamcrest-1.10.1-py3-none-any.whl (48 kB) 48.9/48.9 kB 6.9 MB/s eta 0:00:00
Collecting testcontainers[mysql]<4.0.0,>=3.0.3
  Downloading testcontainers-3.7.1-py2.py3-none-any.whl (45 kB) 45.3/45.3 kB 6.0 MB/s eta 0:00:00
Collecting cryptography<36.0.0
  Downloading cryptography-41.0.1-cp37-manylinux_2_28_x86_64.whl (4.3 MB) 4.3/4.3 kB 77.5 MB/s eta 0:00:00
Collecting scikit-learn>=0.20.0
  Downloading scikit_learn-1.0.2-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (24.8 MB) 24.8/24.8 kB 52.2 MB/s eta 0:00:00
Collecting pytest-xdist<3,>=2.5.0
  Downloading pytest_xdist-2.5.0-py3-none-any.whl (41 kB) 41.7/41.7 kB 5.1 MB/s eta 0:00:00
Collecting pandas<2.0.0
  Downloading pandas-1.3.5-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (11.3 MB) 11.3/11.3 kB 73.7 MB/s eta 0:00:00
Collecting pyyaml<7.0.0,>=3.12
  Downloading PyYAML-6.0-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2010_x86_64.whl (596 kB) 596.3/596.3 kB 41.9 MB/s eta 0:00:00
```

Task 2. Review code to stream messages from Pub/Sub to Cloud Storage: No specific task is needed to perform only we are reviewing and understanding the code written in python.

AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

The given script is a Python programme that reads messages from a Google Cloud Pub/Subtopic, organises them into windows based on the time they were published, and writes them to Google Cloud Storage (GCS). Each section is explained below:

- ✓ **Imports:** The argparse module is used in this script to handle command-line arguments, datetime and random for various operations, logging to log information, and multiple Apache Beam functions for the data processing pipeline.
- ✓ **GroupMessagesByFixedWindows class:** This is a custom PTransform (a transform that may be applied to a pipeline to process data). It organises Pub/Sub messages into fixed-size windows and processes each message by extracting the message's content and publish time. The window_size argument determines the size of the windows.
- ✓ **AddTimestamp class:** This is a DoFn (a function that can be applied to each element in a PCollection). Each element is processed by decoding it from UTF-8 format and retrieving its publication timestamp.
- ✓ **WriteToGCS class:** This is yet another DoFn class. It sends messages in batches (each batch corresponds to a window) to Google Cloud Storage. Each output file is named using the start and end timings of the window as well as a shard ID, and each line in the file comprises a message and its publish time.
- ✓ **run function:** The Beam pipeline is built and executed by this function. It reads from the Pub/Subtopic, groups the messages into windows using the groupMessagesByFixedWindows transform, and then uses the WriteToGCS transform to write the message windows to GCS.
- ✓ **main block:** The script begins to run at this point. It first configures logging before parsing command-line parameters with argparse. These options indicate the GCS directory to send output files to, the window size for grouping messages, the number of shards to utilise when writing to GCS, the Pub/Subtopic to read from, and more. The pipeline is then called by calling the run function with these arguments.

Task 3. Start the pipeline

Following steps are followed,

- ❖ To start the pipeline, run the following command,

AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

Java **Python**

```
python PubSubToGCS.py \
    --project=project_id \
    --region=region \
    --input_topic=projects/project_id/topics/my-id \
    --output_path=gs://bucket_name/samples/output \
    --runner=DataflowRunner \
    --window_size=2 \
    --num_shards=2 \
    --temp_location=gs://bucket_name/temp
```

Updated command as per my data,

```
python PubSubToGCS.py \
    --project=assignment3-ch22m503 \
    --region=us-west1 \
    --input_topic=projects/assignment3-
ch22m503/topics/CH22M503 \
    --output_path=gs://assignment3-ch22m503-
bucket/samples/output \
    --runner=DataflowRunner \
    --window_size=2 \
    --num_shards=2 \
    --temp_location=gs://assignment3-ch22m503-bucket/temp
```

Task 4. Observe job and pipeline progress

Once the command completed, navigated to ‘**DataFlow Console**’ using this url,
<https://console.cloud.google.com/dataflow/jobs?project=assignment3-ch22m503>
To observe the job’s progress,

AI Lab 2 (Assignment – 3)

Aloy Banerjee (CH22M503)

Date of Submission: 04-07-2023

The screenshot shows the Google Cloud Dataflow interface. In the left sidebar, 'Jobs' is selected. The main table lists one job: 'beamapp-root-0629072635-1466104pkln6fc'. The job is 'Running', type 'Streaming', started at 'Jun 29, 2023, 12:56:56 PM', and has an 'ID' of '2.42.0 2023-06-29_00_26_55-11177100026370372116'. It is located in the 'us-west1' region.

Now as a next steps, Open the job details view to see:

- ❖ Job structure
- ❖ Job logs
- ❖ Stage metrics

This screenshot is similar to the previous one, showing the same job details for 'beamapp-root-0629075422-029655-j5olu4d3'. The difference is the presence of a yellow 'INSIGHT' badge in the top right corner of the job row.

This screenshot shows the 'beamapp-root...' job details page. The 'JOB GRAPH' tab is selected, displaying a flowchart of three stages: 'Read from Pub', 'Window into', and 'Write to GCS'. The 'Job info' panel on the right provides detailed configuration information for the job.

Job name	beamapp-root-0629074422-029655-j5olu4d3
Job ID	2023-06-29_00_54_42-15232452723355112928
Job type	Streaming
Job status	Running
SDK version	Apache Beam Python 3.7 SDK 3.42.0
Job region	us-west1
Worker location	us-west1
Current workers	0
Latest worker status	Starting a pool of 1 workers.
Start time	June 29, 2023 at 1:24:43 PM GMT+5
Elapsed time	1 min 53 sec
Encryption type	Google-managed key
Dataflow Prime	Disabled
Runner v2	Enabled
Streaming Engine	Enabled
Vertical Autoscaling	Disabled

Next stage is to view the 'Cloud Storage' details,
Search with Cloud Storage in the search bar and click on 'Search',

AI Lab 2 (Assignment – 3)

Aloy Banerjee (CH22M503)

Date of Submission: 04-07-2023

The screenshot shows the Google Cloud Dataflow interface. On the left, there's a sidebar with options like Overview, Jobs, Pipelines, Workbench, Snapshots, and SQL Workspace. The main area is titled 'Jobs' and shows a list of running jobs. A search bar at the top right contains the text 'cloud storage'. Below it, a dropdown menu lists 'cloud scheduler' and 'cloud storage'. Under 'PRODUCTS & PAGES', there are sections for Cloud Storage (Enterprise-ready object storage), Storage Pools (Cloud Volumes), Settings (Cloud Storage), and Buckets (Cloud Storage). To the right, there's a table showing insights for regions: 42-(2928) and 55-72116, both in us-west1.

It will land into cloud storage dashboard page with the bucket details,

The screenshot shows the Google Cloud Storage dashboard. The left sidebar has 'Navigation menu' with 'Storage' selected, and 'Buckets' is the active tab. It features a 'TRY THE NEW CLOUD STORAGE MONITORING DASHBOARD' section with a 'TRY NOW' button. Below it is a 'View security recommendations' section with a 'VIEW IN TABLE' and 'LEARN MORE' button. The main area displays a table of buckets:

Name	Created	Location type	Location	Default storage class	Last modified	Public access	Access
assignment3-ch22m503-bucket	Jun 29, 2023, 12:10:52 PM	Multi-region	us	Standard	Jun 29, 2023, 12:10:52 PM	Subject to object ACLs	Fine-grained
assignment3-ch22m503.appspot.com	Jun 29, 2023, 12:15:38 PM	Region	us-west1	Standard	Jun 29, 2023, 12:15:38 PM	Subject to object ACLs	Fine-grained
staging assignment3-ch22m503.appspot.com	Jun 29, 2023, 12:15:38 PM	Region	us-west1	Standard	Jun 29, 2023, 12:15:38 PM	Subject to object ACLs	Fine-grained

The screenshot shows the detailed view of a Google Cloud Storage bucket. At the top, there are buttons for UPLOAD FILES, UPLOAD FOLDER, CREATE FOLDER, TRANSFER DATA, MANAGE HOLDS, DOWNLOAD, and DELETE. Below that is a filter section with 'Filter by name prefix only' and a general 'Filter objects and folders' search bar. The main table lists objects:

Name	Size	Type	Created	Storage class	Last modified	Public access	Version
samples/_	—	Folder	—	—	—	—	—
temp/_	—	Folder	—	—	—	—	—

AI Lab 2 (Assignment – 3)
Aloy Banerjee (CH22M503)
Date of Submission: 04-07-2023

Cloud Storage Bucket Overview						
Actions	Name	Size	Type	Created	Storage class	Last modified
<input type="checkbox"/>	output-14:54-14:56-0	34 B	application/octet-stream	Jun 29, 2023, 8:28:15 PM	Standard	Jun 29, 2023, 8:28:15 PM
<input type="checkbox"/>	output-14:56-14:58-0	34 B	application/octet-stream	Jun 29, 2023, 8:28:36 PM	Standard	Jun 29, 2023, 8:28:36 PM
<input type="checkbox"/>	output-14:56-14:58-1	34 B	application/octet-stream	Jun 29, 2023, 8:28:35 PM	Standard	Jun 29, 2023, 8:28:35 PM
<input type="checkbox"/>	output-14:58-15:00-0	34 B	application/octet-stream	Jun 29, 2023, 8:30:35 PM	Standard	Jun 29, 2023, 8:30:35 PM
<input type="checkbox"/>	output-14:58-15:00-1	34 B	application/octet-stream	Jun 29, 2023, 8:30:35 PM	Standard	Jun 29, 2023, 8:30:35 PM

Follow the clean-up steps,

- gcloud scheduler jobs delete publisher-job
- gcloud pubsub topics delete \$TOPIC_ID
- gsutil -m rm -rf "gs://\${BUCKET_NAME}/samples/output*"
- gsutil -m rm -rf "gs://\${BUCKET_NAME}/temp/*"
- gsutil rb gs://\${BUCKET_NAME}

[End of Introduction to Stream Processing with Cloud Pub/Sub and Dataflow: Qwik Start](#)
[hands on](#)