

DataEng: Data Maintenance In-class Assignment

A. Discussion Questions

Response:

We can use compression when we need to send a large amount of data over the network. Also we can do compression when the storage is limited. We can do encryption when we have some PII information or confidential information. We can reduce the data when there is a lot of similar data and when we would not lose much data when we reduce it. For example we have breadcrumb data for 5 seconds but if we reduce the data to 30 second breadcrumb data we would not lose much data.

We can combine the transformations based on our requirements. For example, We can combine encryption and compression when we have limited space and need to store a file that has PII information. For the breadcrumb data we can reduce the data and then compress the data.

C. Store Data to GCP Storage Bucket

How much bucket space (in KiBs) does it take to store 1 day of breadcrumbs for 100 vehicles?

Response:

To store one days data for 100 vehicles it takes about 54.1 MB or 55362.4 KiB

archbuckets

Location	Storage class	Public access	Protection
us-west1 (Oregon)	Nearline	Not public	Soft Delete

OBJECTS

CONFIGURATION

PERMISSIONS

PROTECTION

LIFECYCLE

OBSERVABILITY

INVENTORY REPORTS

OPERATIONS

Folder browser

Buckets > archbuckets

UPLOAD FILES UPLOAD FOLDER CREATE FOLDER TRANSFER DATA MANAGE HOLDS EDIT RETENTION DOWNLOAD

DELETE

Filter by name prefix only Filter Filter objects and folders Show Live objects only

<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified	
<input type="checkbox"/>	archive_2024-05-22.json	54.1 MB	application/json	May 21, 2024, 8:08:05 PM	Nearline	May 21, 2024, 8:08:05 PM	Download

D. Compress

How much bucket space (in KiBs) does it take to store the compressed data?

Response:

To store one days compressed data for 100 vehicles it takes about 3.9 MB or 3808.59 KiB

archbuckets

Location

us-west1 (Oregon)

Storage class

Nearline

Public access

Not public

Protection

Soft Delete

OBJECTS

CONFIGURATION

PERMISSIONS

PROTECTION

LIFECYCLE

OBSERVABILITY

INVENTORY REPORTS

OPERATIONS

Folder browser

archbuckets

Buckets > archbuckets

UPLOAD FILES

UPLOAD FOLDER

CREATE FOLDER

TRANSFER DATA

MANAGE HOLDS

EDIT RETENTION







DOWNLOAD

DELETE

Filter by name prefix only

Filter objects and folders

Show Live objects only

<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified	
<input type="checkbox"/>	 archive_2024-05-22.json	54.1 MB	application/json	May 21, 2024, 8:24:37 PM	Nearline	May 21, 2024, 8:24:37 PM	 
<input type="checkbox"/>	 archivecompress_2024-05-22.json.gz	3.9 MB	application/x-gzip	May 21, 2024, 8:32:15 PM	Nearline	May 21, 2024, 8:32:15 PM	 

E. Encrypt

How much bucket space (in KiBs) does it take to store the encrypted, compressed data?

Response:

To store one days compressed and encrypted data for 100 vehicles it takes about 6.8 MB or 6640.62 KiB

archbuckets

Location

Storage class

Public access

Protection

us-west1 (Oregon)

Nearline

Not public

Soft Delete

OBJECTS

CONFIGURATION

PERMISSIONS

PROTECTION

LIFECYCLE

OBSERVABILITY

INVENTORY REPORTS

OPERATIONS

Folder browser

archbuckets

Buckets > archbuckets

UPLOAD FILES

UPLOAD FOLDER

CREATE FOLDER

TRANSFER DATA

MANAGE HOLDS

EDIT RETENTION

DOWNLOAD

DELETE

Filter by name prefix only

Filter

Filter objects and folders

Show

Live objects only

<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified	
<input type="checkbox"/>	<div><div></div><div>archive_2024-05-22.json</div></div>	54.1 MB	application/json	May 21, 2024, 8:24:37 PM	Nearline	May 21, 2024, 8:24:37 PM	<div><div></div><div></div></div>
<input type="checkbox"/>	<div><div></div><div>archivecompress_2024-05-22.json</div></div>	3.9 MB	application/x-gzip	May 21, 2024, 8:32:15 PM	Nearline	May 21, 2024, 8:32:15 PM	<div><div></div><div></div></div>
<input type="checkbox"/>	<div><div></div><div>archivecompressencrypt_2024-05-22.json</div></div>	6.8 MB	application/octet-stream	May 23, 2024, 8:18:28 PM	Nearline	May 23, 2024, 8:18:28 PM	<div><div></div><div></div></div>