

Big Data Technologies (CSP 554)

Assignment-1

Ques. 1 What was the problem with the Google flu detection algorithm?

The problem with Google Flu Detection Algorithm was that it predicts more than twice the rate of visits to doctors for influenza-like illnesses than the Centers for Disease Control and Prevention (CDC) based on its estimated reports on monitoring from laboratories across the United States. The model was making predictions about flu trends based on search queries.

Two issues that contributed to the problem were: -

- Big Data Hubris
- Algorithm Dynamics

Ques.2 What is big data hubris?

Big Data Hubris is a commonly tacit assumption that big data is a replacement rather than a supplement, to traditional data collection and analysis. However, the quantity of data does not mean that one can ignore fundamental measurement issues and build validity, reliability, and dependencies between data.

Ques.3 What approach could have been used to improve the Google flu detection algorithm?

Google Flu detection could be improved over time if the developer had changed the detection method to “why are these flu-related keywords searched for” instead of generating the data in numbers of trending keywords related to the flu. Also, the concept of algorithm dynamics must be very precise as it is one of the main methods that make the algorithm display erroneous data that results in a high percentage of error.

Ques.4 What is “algorithm dynamics?”

Algorithm dynamics is defined as the changes made by engineers to improve the commercial service and by the consumers while using that service. Algorithm dynamism is important because it helps the algorithm adapt to new changes based on the changes occurring over a certain period in society.

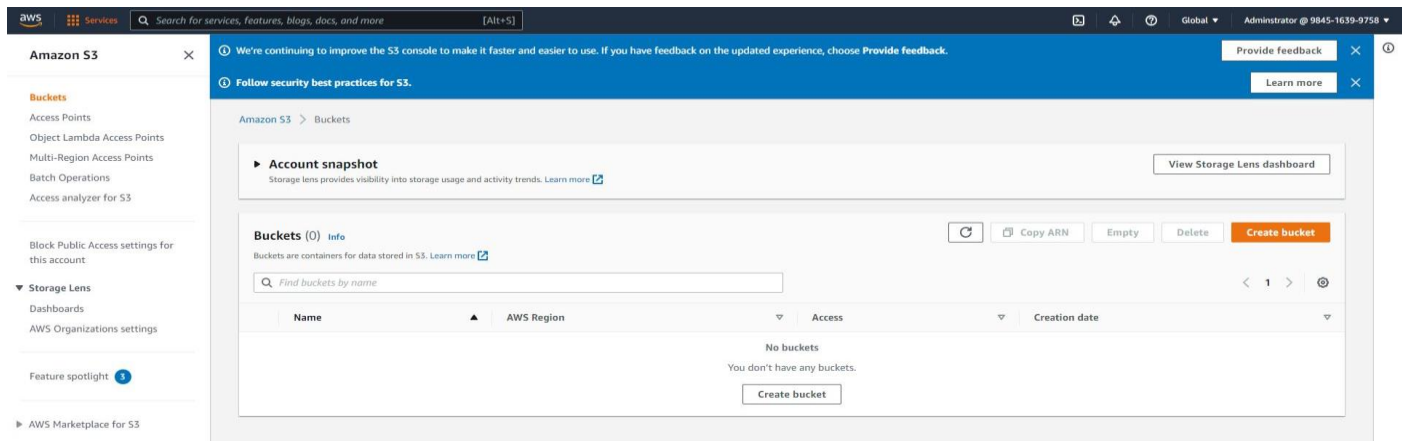
Ques.5 What aspect of algorithm dynamics impacted the Google flu detection algorithm?

All empirical studies are based on measurement. Algorithm dynamics impact the Google flu detection algorithm as GFT acts abnormally by showing unstable reflection about the prevalence of the flu. GFT uses the relative popularity of the search terms in its model. Improvements to the algorithm can negatively impact the GFT estimates.

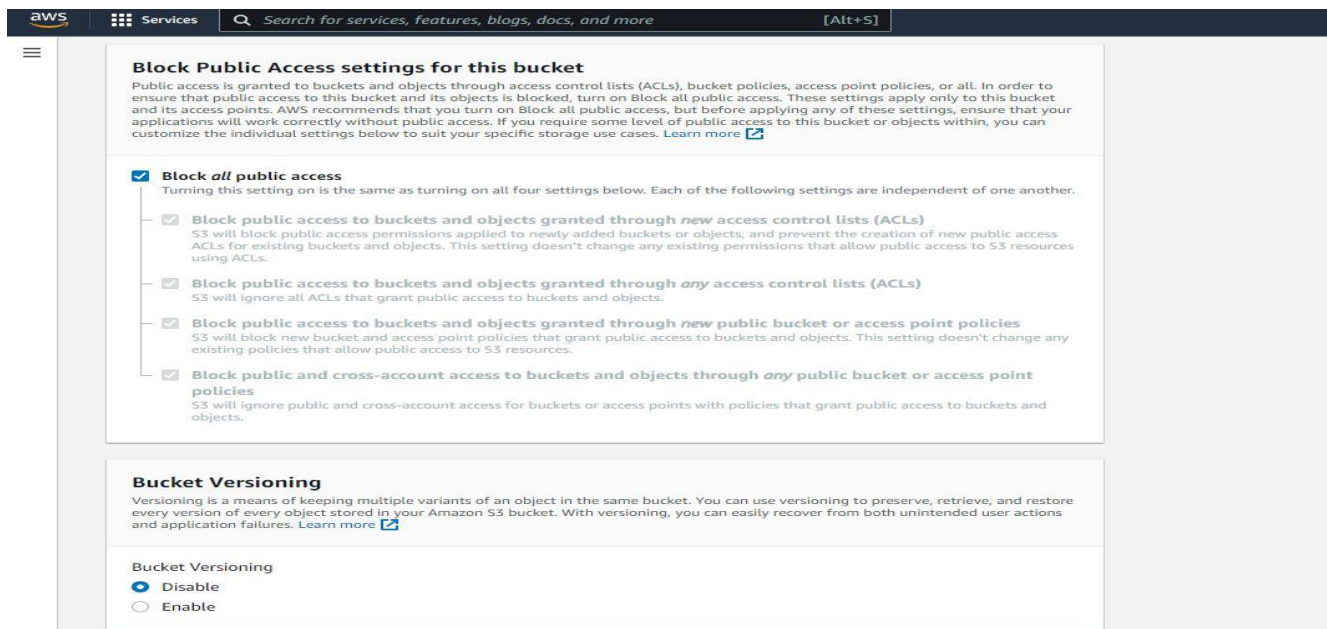
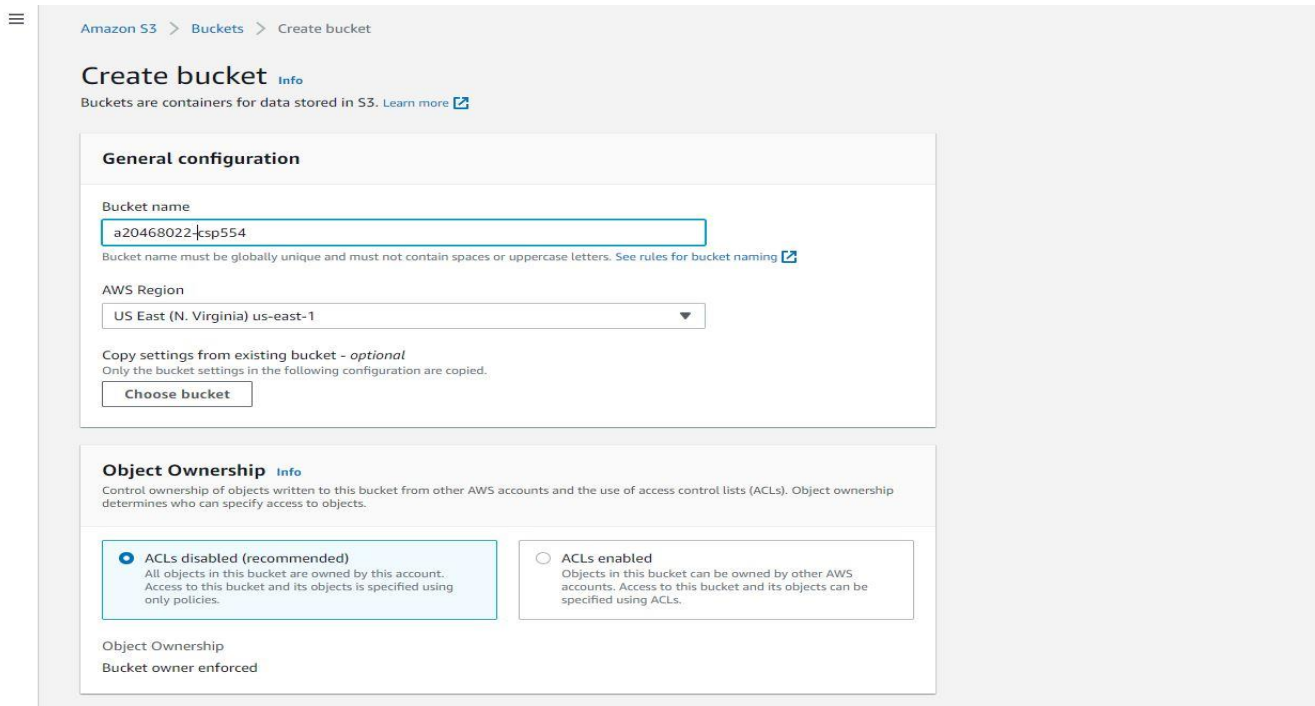
Ques.6

(a) Amazon S3 Bucket

Below is the screenshot of the Amazon S3



(b) Creating an S3 Bucket with the name a20468022-csp554. The series of screenshots below represent the complete process of creating a bucket.



Tags (0) - optional

Track storage cost or other criteria by tagging your bucket. [Learn more](#)

No tags associated with this bucket.

Add tag

Default encryption

Automatically encrypt new objects stored in this bucket. [Learn more](#)

Server-side encryption

☒ Disable

☐ Enable

► Advanced settings

After creating the bucket you can upload files and folders to the bucket, and configure additional bucket settings.

Cancel

Create bucket

Amazon S3

Buckets

Access Points

Object Lambda Access Points

Multi-Region Access Points

Batch Operations

Access analyzer for S3

Block Public Access settings for this account

Storage Lens

Dashboards

AWS Organizations settings

Successfully created bucket "a20468022-csp554"

To upload files and folders, or to configure additional bucket settings choose [View details](#).

Follow security best practices for S3.

Learn more

Amazon S3 > Buckets

Account snapshot

Storage Lens provides visibility into storage usage and activity trends. [Learn more](#)

View Storage Lens dashboard

Buckets (1) Info

[Refresh](#) [Copy ARN](#) [Empty](#) [Delete](#) [Create bucket](#)

< 1 >

[Info](#)

	Name	AWS Region	Access	Creation date
<input type="radio"/>	a20468022-csp554	US East (N. Virginia) us-east-1	Bucket and objects not public	August 30, 2022, 20:33:36 (UTC-05:00)

(c) Creating and Uploading an Object in the Bucket

Amazon S3 > Buckets > a20468022-csp554 > Upload

Upload Info

Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDK or Amazon S3 REST API. [Learn more](#)

Drag and drop files and folders you want to upload here, or choose [Add files](#), or [Add folders](#).

Files and folders (1 Total, 33.3 KB)

[Remove](#) [Add files](#) [Add folder](#)

All files and folders in this table will be uploaded.

< 1 >

<input type="checkbox"/>	Name	Folder	Type	Size
<input type="checkbox"/>	Big Data Course Information.docx	-	application/vnd.openxmlformats-officedocument.wordprocessingml.document	33.3 KB

Destination

Destination

s3://a20468022-csp554

► Destination details

Bucket settings that impact new objects stored in the specified destination.

Services

Search for services, features, blogs, docs, and more

[Alt+S]

Global

Administrator @ 9845-1639-9758

Upload succeeded

View details below.

Upload: status

Close

The information below will no longer be available after you navigate away from this page.

Summary

Destination

s3://a20468022-csp554

Succeeded

1 file, 33.3 KB (100.00%)

Failed

0 files, 0 B (0%)

Files and folders

Configuration

Files and folders (1 Total, 33.3 KB)

Find by name

< 1 >

Name

Folder

Type

Size

Status

Error

Big Data Course Information.docx

-

application/vnd.openxmlformats-officedocument.wordprocessingml.document

33.3 KB

Succeeded

-

(d) Deleting a bucket and its contents

We are deleting a bucket to avoid extra charges. But before deleting a bucket we need to empty it and make sure that there is no object inside it. Below are the screenshots for the same.

Services

Search for services, features, blogs, docs, and more

[Alt+S]

Global

Administrator @ 9845-1639-9758

Amazon S3

Buckets

a20468022-csp554

Empty bucket

Empty bucket

Info

Emptying the bucket deletes all objects in the bucket and cannot be undone.

Objects added to the bucket while the empty bucket action is in progress might be deleted.

To prevent new objects from being added to this bucket while the empty bucket action is in progress, you might need to update your bucket policy to stop objects from being added to the bucket.

Learn more

If your bucket contains a large number of objects, creating a lifecycle rule to delete all objects in the bucket might be a more efficient way of emptying your bucket.

Learn more

Go to lifecycle rule configuration

Permanently delete all objects in bucket "a20468022-csp554"?

To confirm deletion, type *permanently delete* in the text input field.

permanently delete

Cancel

Empty

Services

Search for services, features, blogs, docs, and more

[Alt+S]

Global

Administrator @ 9845-1639-9758

Successfully emptied bucket "a20468022-csp554"

Success details below. If you want to delete this bucket, use the delete bucket configuration.

Empty bucket: status

Cancel

Exit

The details below are no longer available after you navigate away from this page.

Summary

Source

s3://a20468022-csp554

Successfully deleted

1 object, 33.3 KB

Failed to delete

0 objects

Failed to delete (0)

Find objects by name

< 1 >

Name

Prefix

Version ID

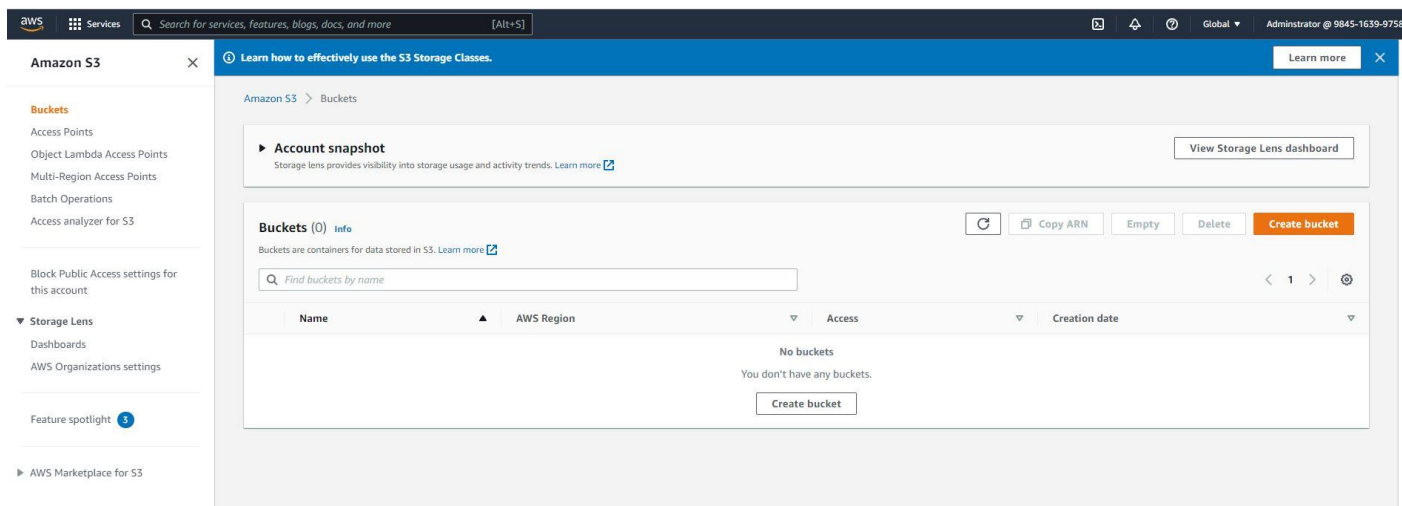
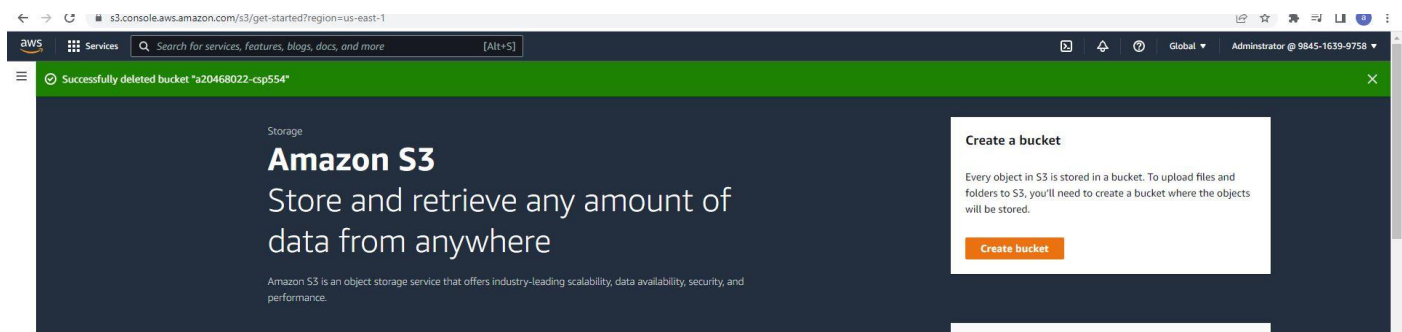
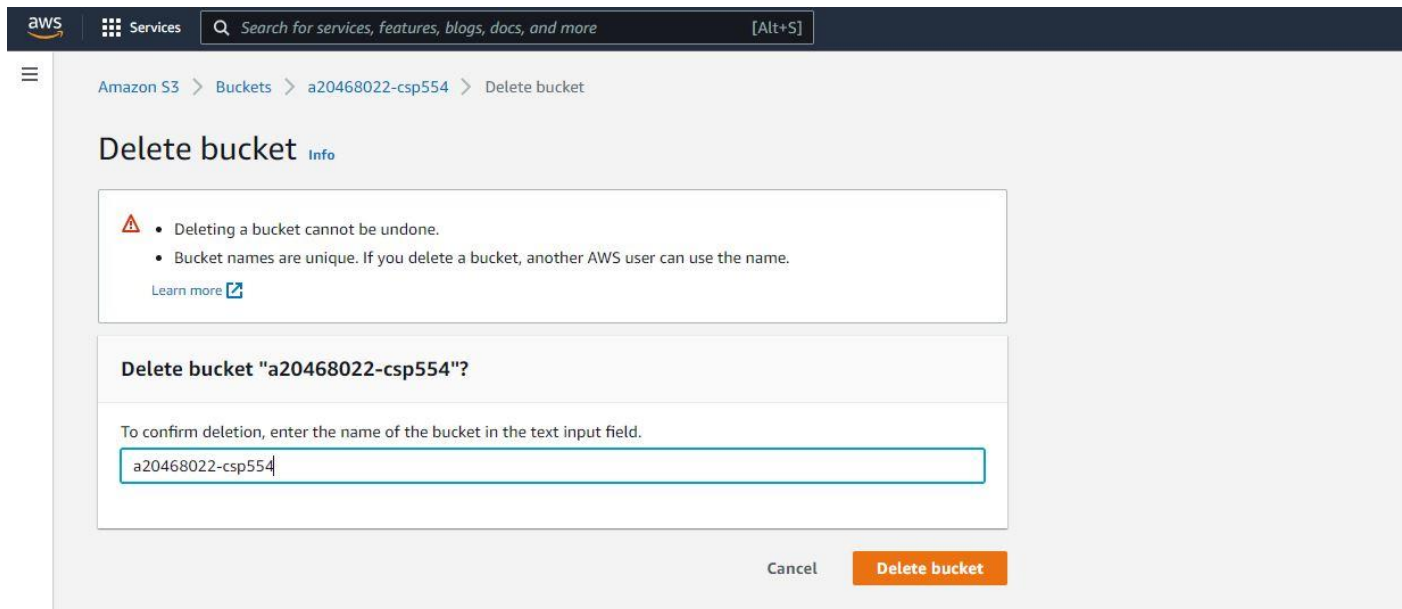
Type

Last modified

Size

Error

No failed object deletions



Submitted By: -

Aastha Dhir

CWID- A20468022

Email id- adhir2@hawk.iit.edu