

CSP 554 BIG DATA TECHNOLOGIES
MOHAMMED WASIM R D(A20497053)

ASSIGNMENT 1

1. Answer each of the following questions about the article in just one to three sentences each:

- **What was the problem with the Google flu detection algorithm?**

The goal behind Google Flu Trends was to evaluate the vast number of Google search queries obtained to see if there was a presence of flu-like illness in a population by monitoring millions of user's health tracking activities online.

- **What is big data hubris?**

Big data hubris is a belief that big data should be used as a replacement for rather than an addition to traditional data collecting and analysis

- **What approach could have been used to improve the Google flu detection algorithm?**

Larger value can be obtained by combining GFT with other real time health data

- a) Combining GFT and lagged CDC data
- b) Dynamically recalibrating GFT

- **What is “algorithm dynamics?”**

Algorithm dynamics are the changes made by engineers to improve the commercial service and by consumers in using that service. Several changes in Google’s search algorithm and user behavior likely affected GFT’s tracking.

- **What aspect of algorithm dynamics impacted the Google flu detection algorithm?**

The most common explanation for GFT’s error is a media-stoked panic last flu season (1, 15) . Although this may have been a factor, it cannot explain why GFT has been missing high by wide margins for more than 2 years . Moreover, the data sources for GFT were not appropriate to produce valid and reliable results compatible for scientific research

2. Setting up AWS account and uploading files to S3

a) Sign up to AWS Account as root user



Sign in

☒ **Root user**

Account owner that performs tasks requiring unrestricted access. [Learn more](#)

☐ **IAM user**

User within an account that performs daily tasks. [Learn more](#)

Root user email address

mrafeeqahameddilshad@hawk.iit.edu|

Next

By continuing, you agree to the [AWS Customer Agreement](#) or other agreement for AWS services, and the [Privacy Notice](#). This site uses essential cookies. See our [Cookie Notice](#) for more information.

New to AWS?

Create a new AWS account



b) Creating bucket with id and course code

The screenshot shows the AWS S3 console interface. At the top, there's a navigation bar with the AWS logo, 'Services' link, a search bar, and user information. A left-hand navigation pane lists various S3 features like Buckets, Access Points, and Storage Lens. The main content area displays a success message: 'Successfully created bucket "a20497053-csp554"'. Below this, there's a table showing the details of the created bucket.

Name	AWS Region	Access	Creation date
a20497053-csp554	US East (N. Virginia) us-east-1	Bucket and objects not public	September 1, 2022, 14:48:42 (UTC-05:00)

c) Creating folder in the bucket

The screenshot shows the AWS S3 console interface. On the left, the 'Amazon S3' sidebar is visible with options like Buckets, Access Points, and Storage Lens. The main content area displays a notification bar at the top with a green message: 'Successfully created folder "Assignment 1". Operation successfully completed.' Below this, there's a blue bar with a message about replicating data to any storage class. The bucket name 'a20497053-csp554' is shown. The 'Objects' tab is selected, showing a list of objects. The 'Assignment 1/' folder is listed as a 'Folder' type. The 'Create folder' button is highlighted in orange.

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

Feature spotlight 3

► AWS Marketplace for S3

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose [Provide feedback](#).

Successfully created folder "Assignment 1".
Operation successfully completed.

Replicate your data to any storage class to save on storage costs. [Get started](#)

a20497053-csp554 [Info](#)

Objects | Properties | Permissions | Metrics | Management | Access Points

Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

[Refresh](#) [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#)

[Create folder](#) [Upload](#)

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	Assignment 1/	Folder	-	-	-

d) Uploading files in folder which are called objects

The screenshot shows the 'Upload: status' page in the AWS S3 console. It displays a summary of the upload operation, indicating that 1 file (199.9 KB) was successfully uploaded to the 'Assignment 1' folder. Below the summary, the 'Files and folders' tab is selected, showing a table with the uploaded file details.

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose [Provide feedback](#).

Upload succeeded
View details below.

Upload: status

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

Summary

Destination	Succeeded	Failed
s3://a20497053-csp554/Assignment 1/	1 file, 199.9 KB (100.00%)	0 files, 0 B (0%)

Files and folders | Configuration

Files and folders (1 Total, 199.9 KB)

Name	Folder	Type	Size	Status	Error
AWS01 v2.pdf	-	application/pdf	199.9 KB	Succeeded	-

e) Overview of uploaded file

The screenshot shows the AWS S3 console interface. The left sidebar contains navigation options like Buckets, Access Points, and Storage Lens. The main content area displays the 'Object overview' for 'AWS01 v2.pdf'. The overview is divided into two columns: 'Properties' and 'S3 URI'. The 'Properties' column lists details such as Owner (mrafeeqahameddilshad), AWS Region (US East (N. Virginia) us-east-1), Last modified (September 1, 2022, 14:51:12 (UTC-05:00)), Size (199.9 KB), Type (pdf), and Key (Assignment 1/AWS01 v2.pdf). The 'S3 URI' column lists details such as S3 URI (s3://a20497053-csp554/Assignment 1/AWS01 v2.pdf), Amazon Resource Name (ARN) (arn:aws:s3::a20497053-csp554/Assignment 1/AWS01 v2.pdf), Entity tag (Etag) (07fbfc4a3c52bd24602d2b109b271b4b), and Object URL (https://a20497053-csp554.s3.amazonaws.com/Assignment+1/AWS01+v2.pdf).

Amazon S3 > Buckets > a20497053-csp554 > Assignment 1/ > AWS01 v2.pdf

AWS01 v2.pdf [Info](#)

[Copy S3 URI](#) [Download](#) [Open](#) [Object actions](#)

Properties | Permissions | Versions

Object overview

Properties	S3 URI
Owner mrafeeqahameddilshad	S3 URI s3://a20497053-csp554/Assignment 1/AWS01 v2.pdf
AWS Region US East (N. Virginia) us-east-1	Amazon Resource Name (ARN) arn:aws:s3::a20497053-csp554/Assignment 1/AWS01 v2.pdf
Last modified September 1, 2022, 14:51:12 (UTC-05:00)	Entity tag (Etag) 07fbfc4a3c52bd24602d2b109b271b4b
Size 199.9 KB	Object URL https://a20497053-csp554.s3.amazonaws.com/Assignment+1/AWS01+v2.pdf
Type pdf	
Key Assignment 1/AWS01 v2.pdf	

f) Deleting folder and objects

The screenshot shows the AWS S3 console interface after a successful deletion. A green banner at the top indicates 'Successfully deleted objects'. Below this, a summary table shows the deletion results. The 'Failed to delete' section is currently empty, showing '0' failed objects.

① We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose [Provide feedback](#).

② Successfully deleted objects
[View details below.](#)

③ Replicate your data to any storage class to save on storage costs. [Get started](#)

④ The information below will no longer be available after you navigate away from this page.

Summary

Source	Successfully deleted	Failed to delete
s3://a20497053-csp554	✔ 2 objects, 199.9 KB	0 objects

Failed to delete | Configuration

⊗ Failed to delete (0)

< 1 >

Name	Folder	Type	Last modified	Size	Error
No objects failed to delete.					

g) Deleting Bucket

aws

Services

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

[Option+S]

Global

Mohammed Wasim

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose [Provide feedback](#).

Provide feedback

Successfully emptied bucket "a20497053-csp554"

View 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	Successfully deleted	Failed to delete
s3://a20497053-csp554	✔ 4 objects, 199.9 KB	0 objects

Failed to delete (0)

Find objects by name

< 1 >

Name	Prefix	Version ID	Type	Last modified	Size	Error
No failed object deletions						