

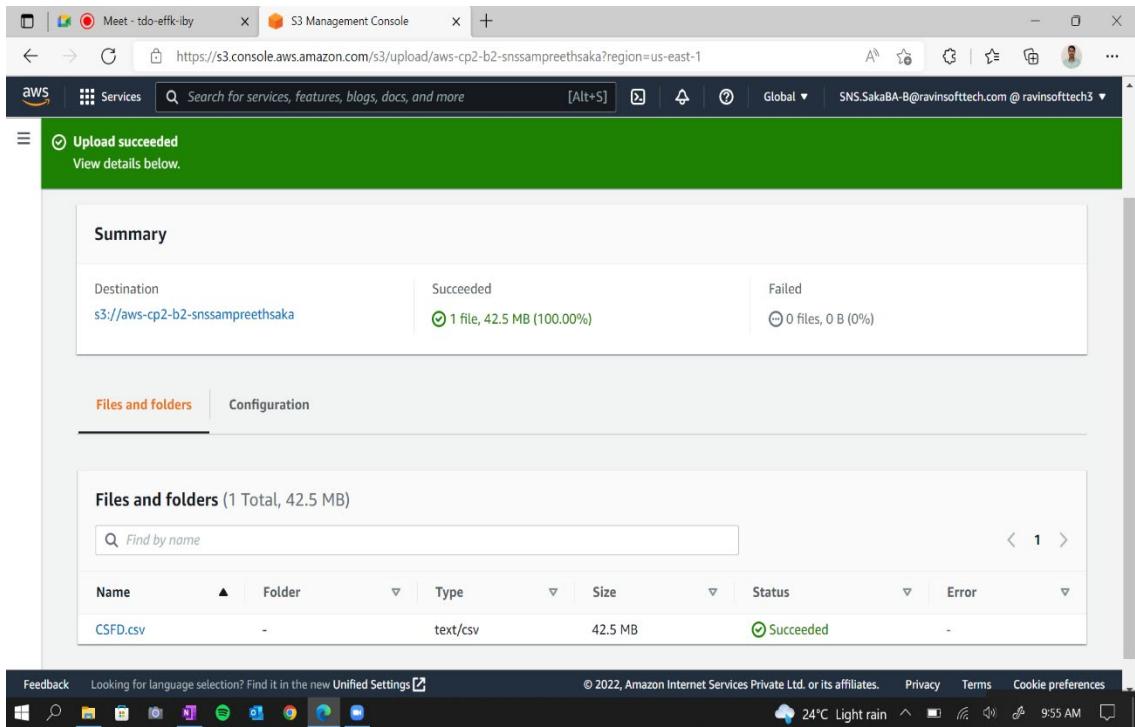
AWS Cloud Checkpoint-3 Batch 2

B2 Team Members

SAGAR KUMAR
SNS SAMPREETH SAKA
ADITHYA SINGH
K PREM KUMAR
PRAGYA SRIVASTAVA

1. Redshift to PowerBI Connectivity

1. Move the Datasets to AWS s3



The screenshot shows a browser window with the AWS S3 Management Console. The URL in the address bar is <https://s3.console.aws.amazon.com/s3/upload/aws-cp2-b2-snssampreethsaka?region=us-east-1>. The main content area displays a green success message: "Upload succeeded" with a link to "View details below". Below this, there's a "Summary" section showing the destination as "s3://aws-cp2-b2-snssampreethsaka" and the results: "Succeeded" (1 file, 42.5 MB (100.00%)) and "Failed" (0 files, 0 B (0%)). At the bottom, there are tabs for "Files and folders" (which is selected) and "Configuration". Under "Files and folders", there's a table titled "Files and folders (1 Total, 42.5 MB)". The table has columns: Name, Folder, Type, Size, Status. One row is shown: "CSFD.csv", "-", "text/csv", "42.5 MB", "Succeeded". The status column shows a green checkmark icon. The bottom of the page includes standard browser navigation and search bars, along with footer links for Feedback, Unified Settings, © 2022 Amazon, Privacy, Terms, and Cookie preferences.

The screenshot shows the AWS S3 console interface. At the top, there are two tabs: 'Meet - tdo-eflk-iby' and 'aws-cp2-b2-snssampreethsaka'. The second tab is active. Below the tabs, the URL is https://s3.console.aws.amazon.com/s3/buckets/aws-cp2-b2-snssampreethsaka?region=us-east-1&tab=objects. The main header includes the AWS logo, 'Services', a search bar, and user information: SNS.SakaBA-B@ravinsofttech.com @ravinsofttech3. The navigation bar shows 'Amazon S3 > Buckets > aws-cp2-b2-snssampreethsaka'. Below the navigation, there are tabs: Objects (highlighted in orange), Properties, Permissions, Metrics, Management, and Access Points. The 'Objects' section displays a table with one item: 'CSFD.csv'. The table columns are Name, Type, Last modified, Size, and Storage class. The file 'CSFD.csv' is a CSV file last modified on August 4, 2022, at 09:55:13 (UTC+05:30), with a size of 42.5 MB and a storage class of Standard. Action buttons include Copy S3 URI, Copy URL, Download, Open, Delete, Actions (dropdown), and Create folder. A search bar 'Find objects by prefix' is also present.

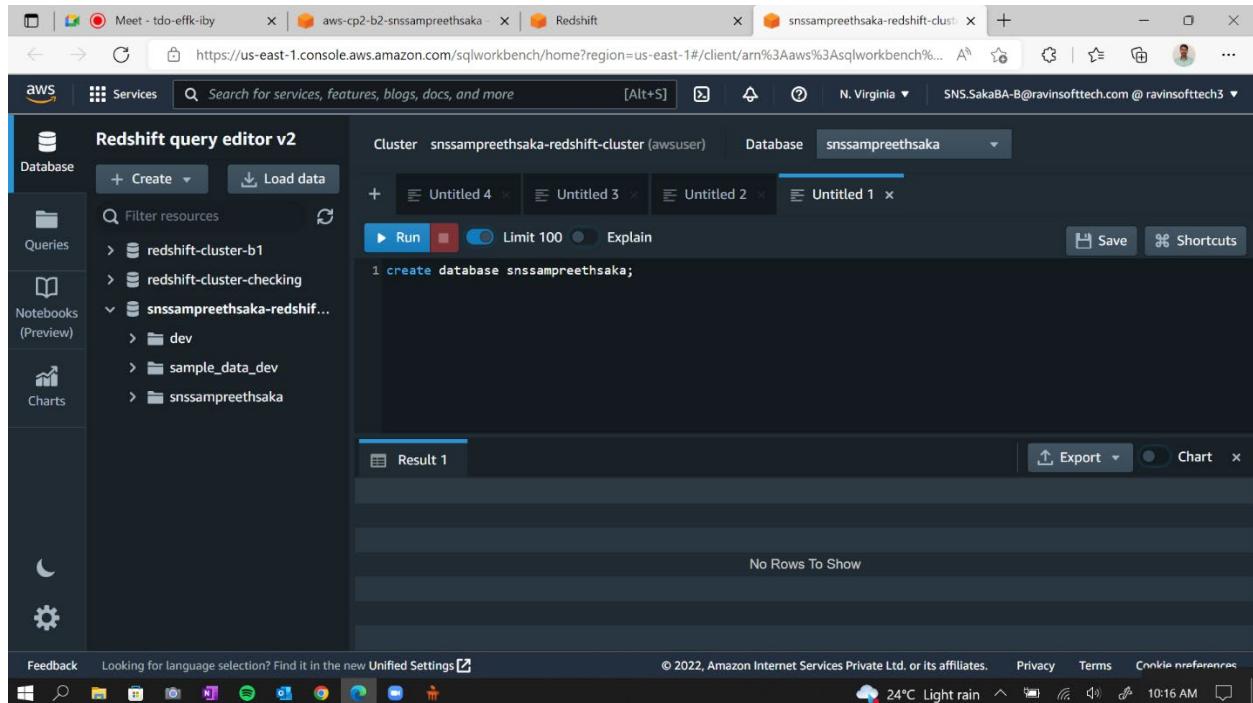
2. Create Redshift Instance

The screenshot shows the AWS Redshift console interface. At the top, there are two tabs: 'Meet - tdo-eflk-iby' and 'Redshift'. The second tab is active. Below the tabs, the URL is https://us-east-1.console.aws.amazon.com/redshiftv2/home?region=us-east-1#cluster-details?cluster=snssampreethsaka-r... The main header includes the AWS logo, 'Services', a search bar, and user information: N. Virginia and SNS.SakaBA-B@ravinsofttech.com @ravinsofttech3. The navigation bar shows 'Amazon Redshift > Clusters > snssampreethsaka-redshift-cluster'. Below the navigation, there are buttons: Actions (dropdown), Edit, Add partner integration, and Query data (dropdown). The 'General information' section displays cluster details in a table:

Cluster identifier	Status	Node type	Endpoint
snssampreethsaka-redshift-cluster	Available	dc2.large	snssampreethsaka-redshift-cluster.c...
Cluster namespace	Date created	Number of nodes	JDBC URL
cb5fa7c9-a90a-4332-bd23-ab57297a70aa	August 04, 2022, 10:03 (UTC+05:30)	1	jdbc:redshift://snssampreethsaka-re...
	Storage used	AQUA	ODBC URL
	-	Not available	Driver={Amazon Redshift (x64)}; Ser...

The cluster was created on August 04, 2022, at 10:03 (UTC+05:30). The screenshot also shows the Windows taskbar at the bottom with various pinned icons and system status indicators like battery level and network connection.

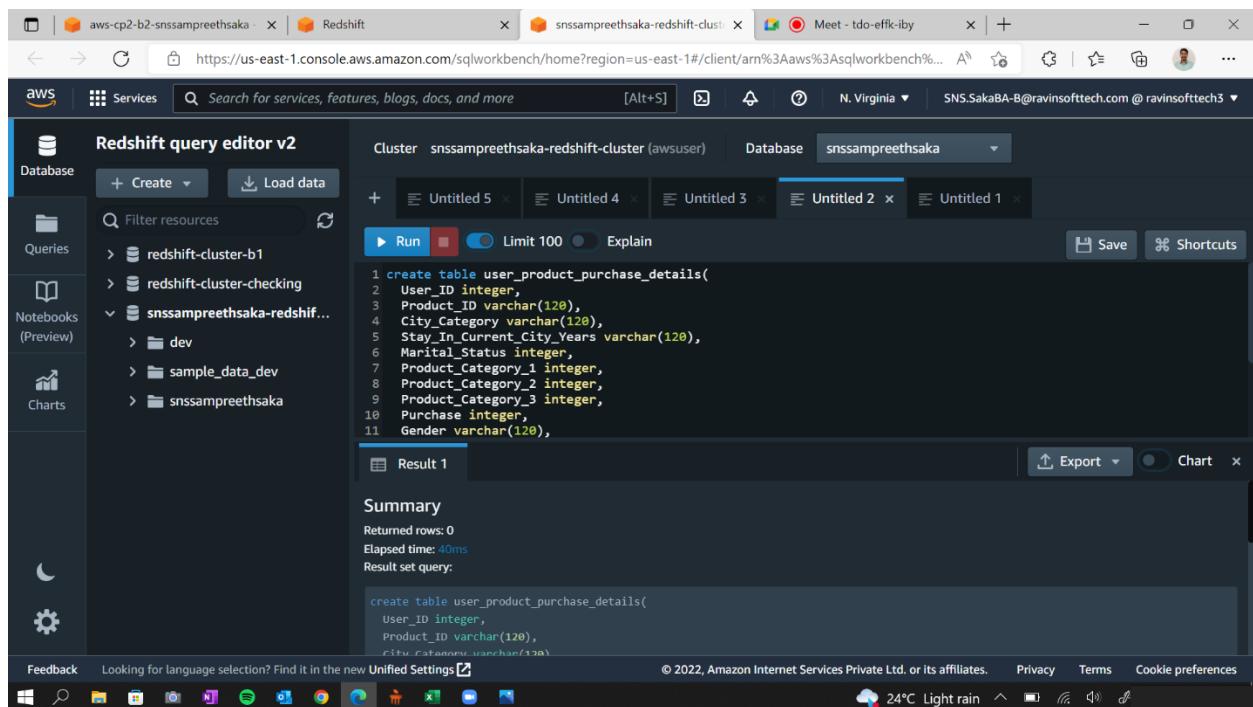
3. Ensure you create required tables in Redshift



The screenshot shows the AWS Redshift Query Editor v2 interface. The left sidebar shows the 'Database' section with 'Queries' selected. The main area displays a query to create a database:

```
1 create database snssampreethsaka;
```

The results pane below shows "No Rows To Show".



The screenshot shows the AWS Redshift Query Editor v2 interface. The left sidebar shows the 'Database' section with 'Queries' selected. The main area displays a query to create a table:

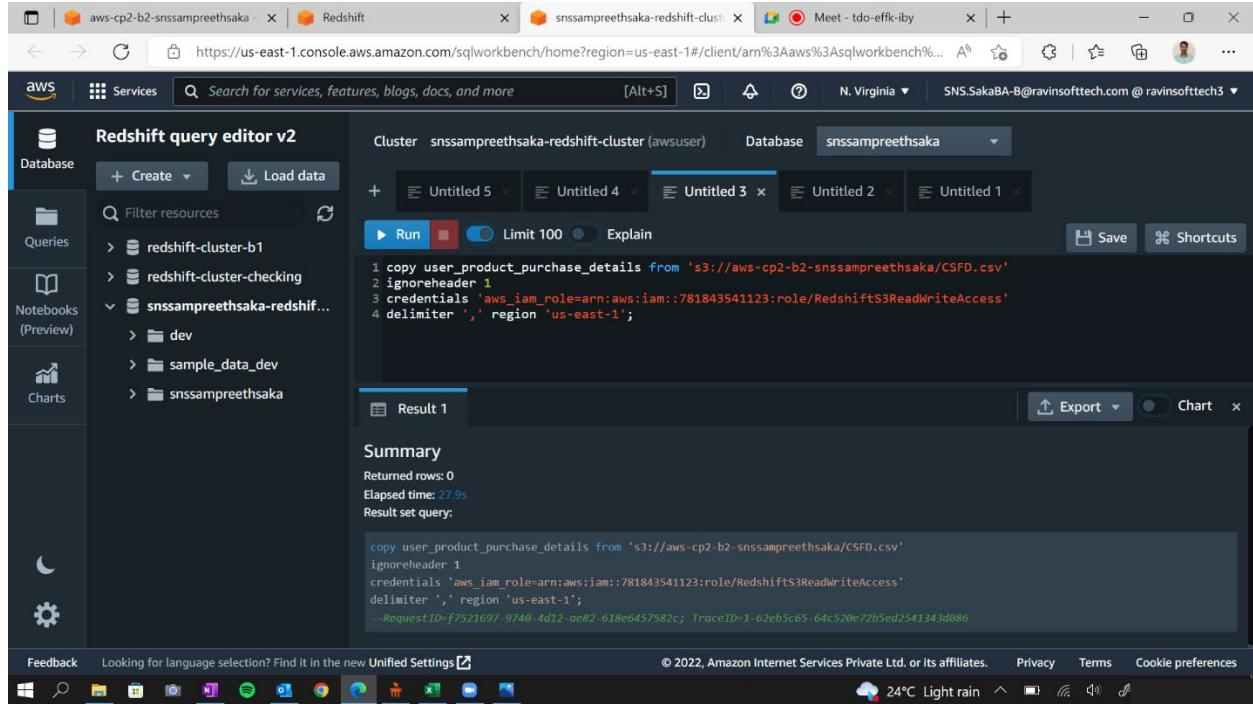
```
1 create table user_product_purchase_details(
2     User_ID integer,
3     Product_ID varchar(120),
4     City_Category varchar(120),
5     Stay_In_Current_City_Years varchar(120),
6     Marital_Status integer,
7     Product_Category_1 integer,
8     Product_Category_2 integer,
9     Product_Category_3 integer,
10    Purchase integer,
11    Gender varchar(120),
```

The results pane below shows a summary:

Summary
Returned rows: 0
Elapsed time: 40ms
Result set query:

```
create table user_product_purchase_details(
    User_ID integer,
    Product_ID varchar(120),
    City_Category varchar(120)
```

4.Create a data pipeline/copy command to move the data from storage to data warehouse(Redshift). You are allowed to use other copy commands as well to move the data from storage to data warehouse.



```

copy user_product_purchase_details from 's3://aws-cp2-b2-snssampreethsaka/CSFD.csv'
ignoreheader 1
credentials 'aws_iam_role=arn:aws:iam::781843541123:role/RedshiftS3ReadWriteAccess'
delimiter ',' region 'us-east-1';

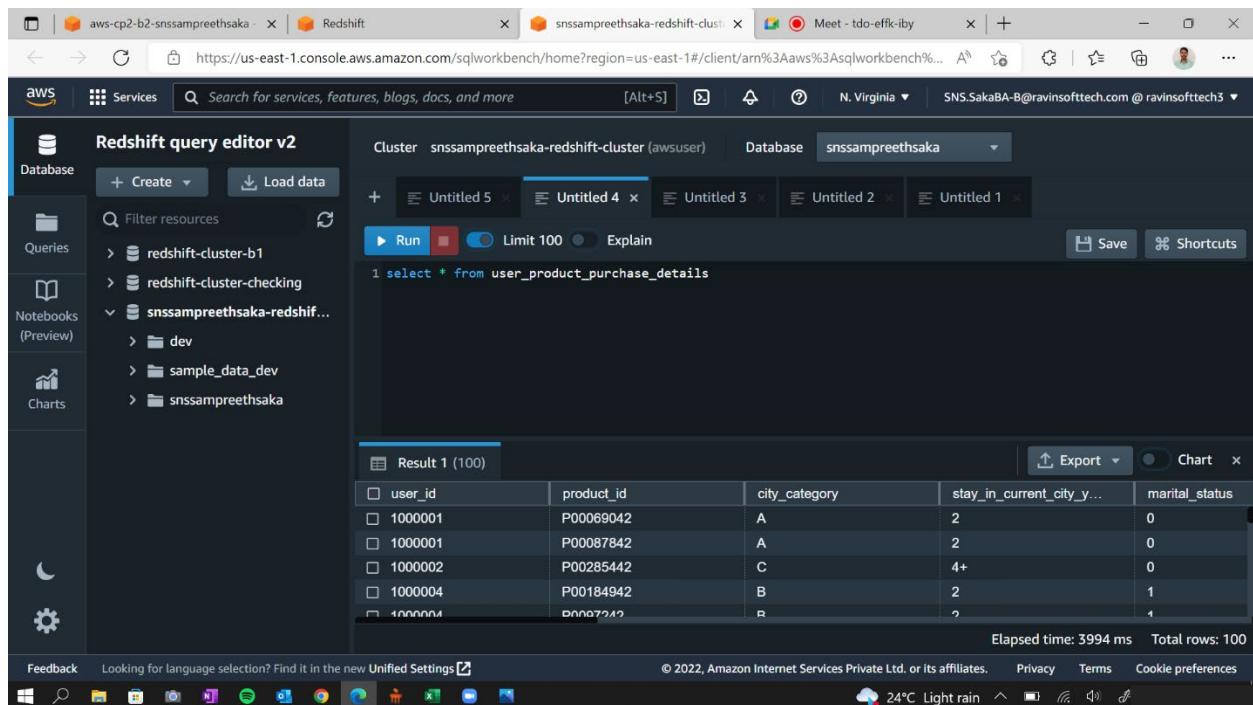
```

Summary
 Returned rows: 0
 Elapsed time: 27.9s
 Result set query:

```

copy user_product_purchase_details from 's3://aws-cp2-b2-snssampreethsaka/CSFD.csv'
ignoreheader 1
credentials 'aws_iam_role=arn:aws:iam::781843541123:role/RedshiftS3ReadWriteAccess'
delimiter ',' region 'us-east-1';

```



```

select * from user_product_purchase_details

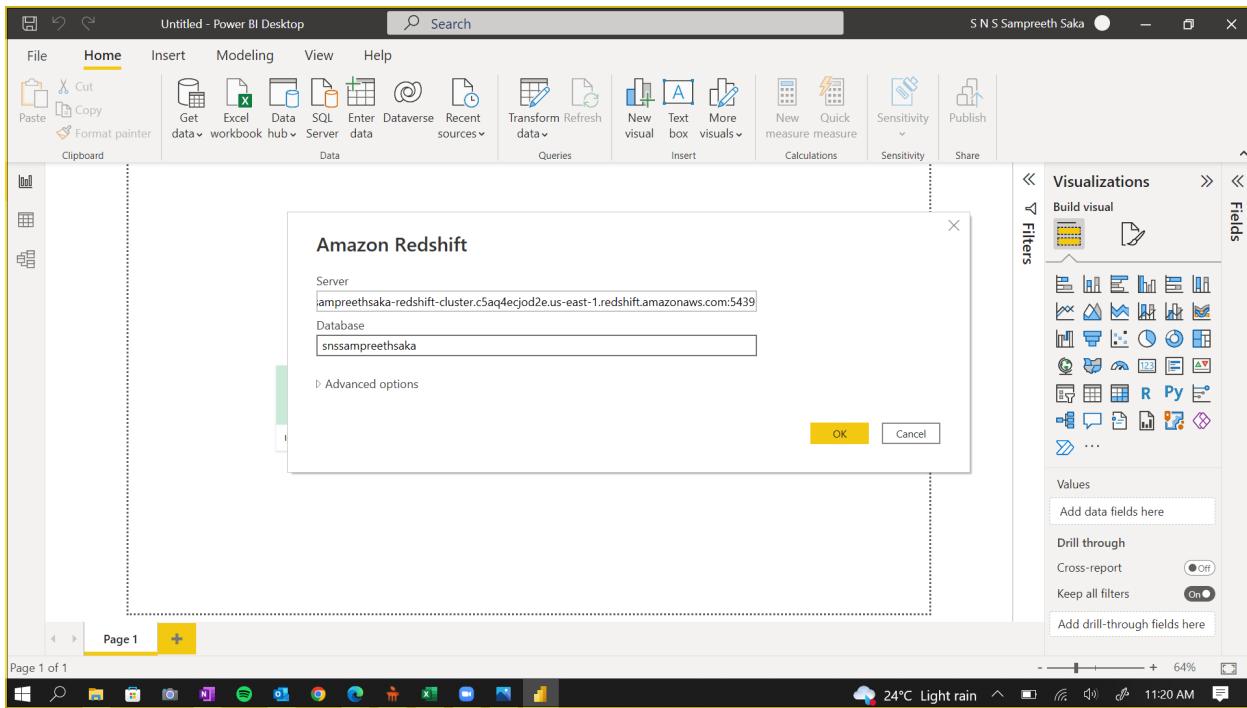
```

Result 1 (100)

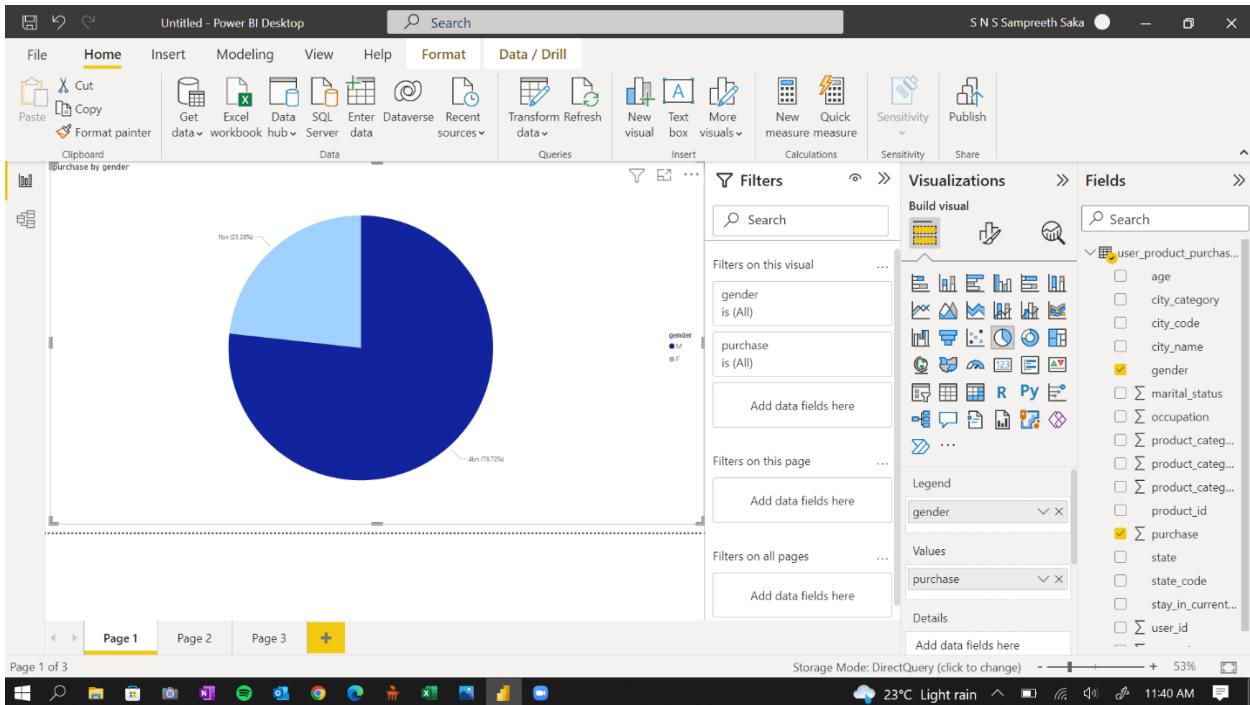
user_id	product_id	city_category	stay_in_current_city_y...	marital_status
1000001	P00069042	A	2	0
1000001	P00087842	A	2	0
1000002	P00285442	C	4+	0
1000004	P00184942	B	2	1
1000004	P0007242	B	2	1

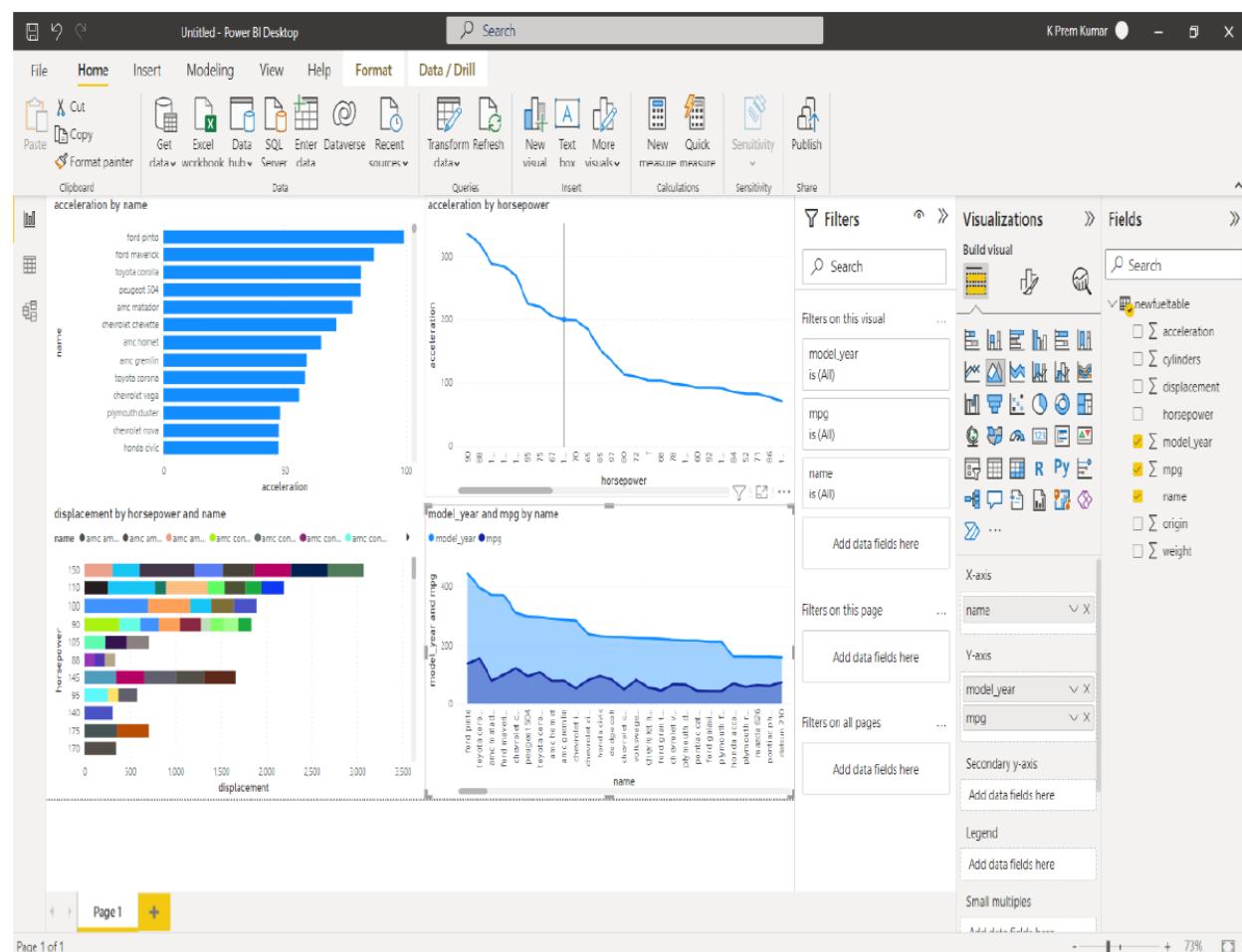
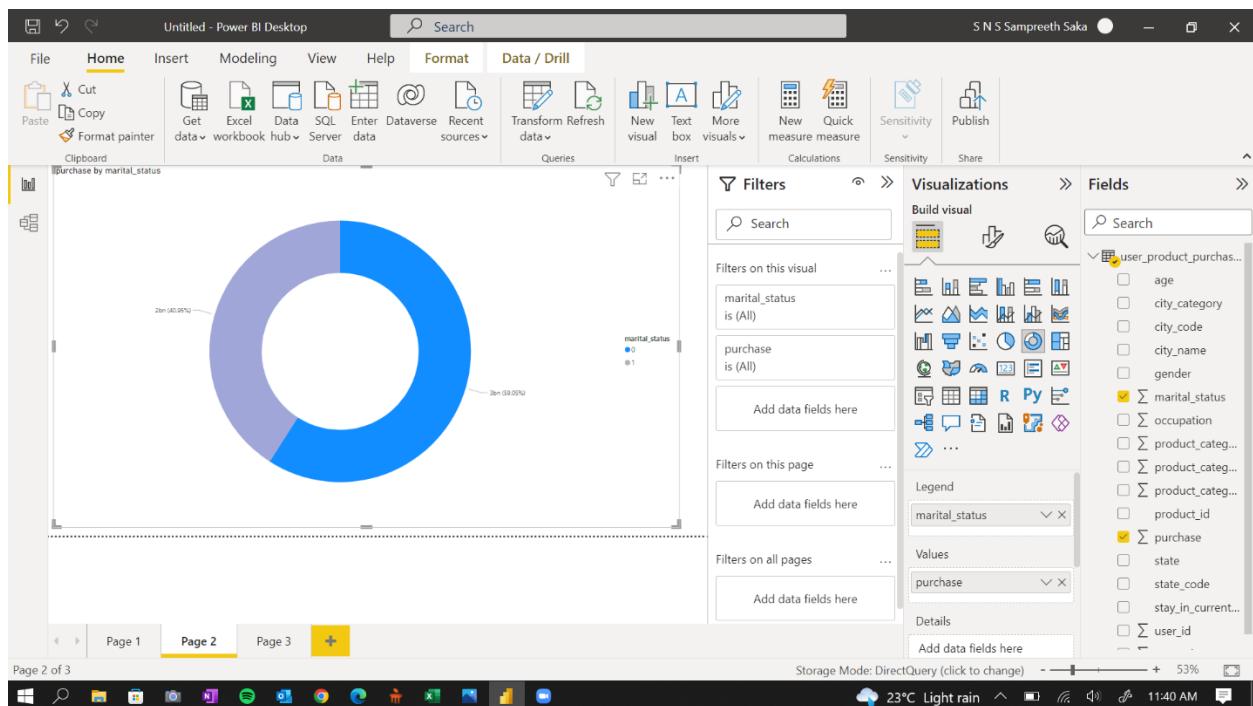
Elapsed time: 3994 ms Total rows: 100

5. Connect the Redshift data to PowerBI



6. Perform the tasks mentioned in Task 2.1(Only 4-5 core reports)





1.Dynamodb to s3 bucket configuire SNS notifications for any new records added in the Dynamodb

The screenshot shows the AWS DynamoDB service page. A green success message at the top states: "The Checkpoint3-B2-DynamodbtoS3 table was created successfully." Below this, the "Tables" section displays three tables: "Checkpoint3-B2-DynamodbtoS3" (Active, User_ID \$), "test" (Active, User_ID \$), and "usedCars-ankur" (Active, Sales_ID N). The left sidebar includes links for Dashboard, Tables (selected), Update settings, Explore items, PartiQL editor (New), Backups, Exports to S3, Reserved capacity, Settings (New), DAX, Clusters, Subnet groups, Parameter groups, and Events.

The screenshot shows the AWS S3 service page. It is on the "Create bucket" step of the wizard. The "General configuration" section is filled out with a bucket name "checkpoint3-b2" and the AWS Region set to "US East (N. Virginia) us-east-1". The "Object Ownership" section has "ACLs disabled (recommended)" selected. The bottom of the screen shows the Windows taskbar with various pinned icons and the system tray indicating "Rain showers" and the date/time as 8/4/2022.

Screenshot of the AWS S3 Management Console showing the creation of a new folder named "log" in the "checkpoint3-b2" bucket.

The "Create folder" dialog is open, displaying the following information:

- Folder name:** log/
- Server-side encryption:** Disable (radio button selected)

A message at the top of the dialog states: "Your bucket policy might block folder creation. If your bucket policy prevents uploading objects without specific tags, metadata, or access control list (ACL) grantees, you will not be able to create a folder using this configuration. Instead, you can use the upload configuration to upload an empty folder and specify the appropriate settings."

The browser status bar shows the URL: https://s3.console.aws.amazon.com/s3/buckets/checkpoint3-b2/object/create_folder?region=us-east-1.

Screenshot of the AWS S3 console showing the creation of a folder named "output".

The browser tabs include: Subject Content | EduNxt, AWS Data Pipeline Management, List tables | Amazon DynamoDB, checkpoint3-b2 - S3 bucket, Meet - tdo-eflk-lby.

The S3 console shows:

- 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.
- AWS Marketplace for S3.

The main view shows the "checkpoint3-b2" bucket with 2 objects:

Name	Type	Last modified	Size	Storage class
logs/	Folder	-	-	-
output/	Folder	-	-	-

Below the table, there is a message from Naga Brahmam Y.

Feedback: Looking for language selection? Find it in the new Unified Settings.

Screenshot of the AWS Data Pipeline Management console showing the creation of a new pipeline.

DynamoDB Tables:

- Tables (3): Checkpoint3-B2-Dynamodbtos3, test, usedCars-ankur

Items returned (6) - Scan/Query items:

User_ID	Age	City_Category	City_Code	City_Name
1000006	51-55	A	C64	Los Angeles
1000005	26-35	A	C150	Fort Lauder...
1000004	46-50	B	C150	Fort Lauder...
1000003	26-35	A	C64	Los Angeles
1000002	55+	C	C259	Henderson
1000001	0-17	A	C259	Henderson

Create Pipeline:

Name: CheckPoint3-B2-Dynamodbtos3

Description (optional): (empty)

Source: Build using a template (selected) - Export DynamoDB table to S3

Parameters:

- Source DynamoDB table name: Checkpoint3-B2-Dynamodbtos3
- Output S3 folder: s3://checkpoint3-b2/output/
- DynamoDB read throughput ratio: 0.25
- Region of the DynamoDB table: us-east-1

Schedule:

You can run your pipeline once or specify a schedule. [More](#)

Run on pipeline activation

Pipeline Configuration

Logging: Enabled

S3 location for logs: `s3://checkpoint3-b2/logs/`

Security/Access

IAM roles: IAM Roles let you control permissions for AWS Data Pipeline and your EC2 applications. [More](#)

Pipeline role: `DataPipelineDefaultRole`

EC2 instance role: `DatapipelineDefaultEc2Role`

Tags

Add up to 10 tags to your pipeline. These tags will be applied to the pipeline as well as any resources created by the pipeline. A tag consists of a case-sensitive key-value pair. [Learn more](#)

Key	Value (Optional)
<code>Add key to create</code>	

This pipeline launches an Amazon EMR cluster in your account on every scheduled execution of the pipeline. [Normal service charges](#) for this resource will apply in addition to charges for other AWS services used by this pipeline.

Activate

Data Pipeline > **List Pipelines** > **Architect: CheckPoint3-B2-Dynamodbtos3 (df-032888074SR1C71230F) [Pending]**

Add **Save** **Activate** **Export** **View/Edit tags**

```

graph TD
    CD[Configuration Default] --> DDD[DynamoDBDataNode  
DDBSourceTable]
    CD --> ECFB[EmrCluster  
EmrClusterForBackup]
    DDD --> EA[EmrActivity  
TableBackupActivity]
    ECFB --> EA
    EA --> SDN[S3DataNode  
S3BackupLocation]
  
```

Parameters

- `myDDTableName`: `Checkpoint3-B2-Dynamodbtos3`
- `myOutputS3Loc`: `s3://checkpoint3-b2/output/`
- `myDDBReadThroughputRatio`: `0.25`
- `myDDBRegion`: `us-east-1`

Errors/Warnings

Object: `EmrClusterForBackup`
WARNING: `'TerminateAfter'` is missing. It is recommended to set this to avoid leaving resource running for long duration.

Screenshot of the AWS Data Pipeline Management console showing a pending pipeline named "Architect: CheckPoint3-B2-DynamodbtoS3 (df-032888074SR1C71230F)".

The pipeline diagram shows a flow from a "Configuration Default" node to two parallel nodes: "DynamoDBDataNode" (with sub-node "DDBSourceTable") and "EmrCluster" (with sub-node "EmrClusterForBackup"). Both of these nodes converge into a single "EmrActivity" node (with sub-node "TableBackupActivity").

The "EmrCluster" configuration details are displayed on the right:

- Name: EmrCluster
- Type: EmrCluster
- Core Instance Count: 1
- Core Instance Type: m3.xlarge
- Release Label: emr-5.23.0
- Master Instance Type: m3.xlarge
- Region: #(myDBRegion)
- Subnet Id: subnet-07bd9493c89d5bcad

Errors/Warnings

Object: EmrClusterForBackup
 WARNING: 'terminateAfter' is missing. It is recommended to set this to avoid leaving resource running for long duration.

Below the pipeline view, the browser navigation bar shows the URL: https://us-east-1.console.aws.amazon.com/sns/v3/home?region=us-east-1#/topic/armawssnsus-east-1:781843541123:testpk

Screenshot of the Amazon SNS Topics page for the topic "testpk".

Details

Name	testpk	Display name	-
ARN	arn:aws:sns:us-east-1:781843541123:testpk	Topic owner	781843541123
Type	Standard		

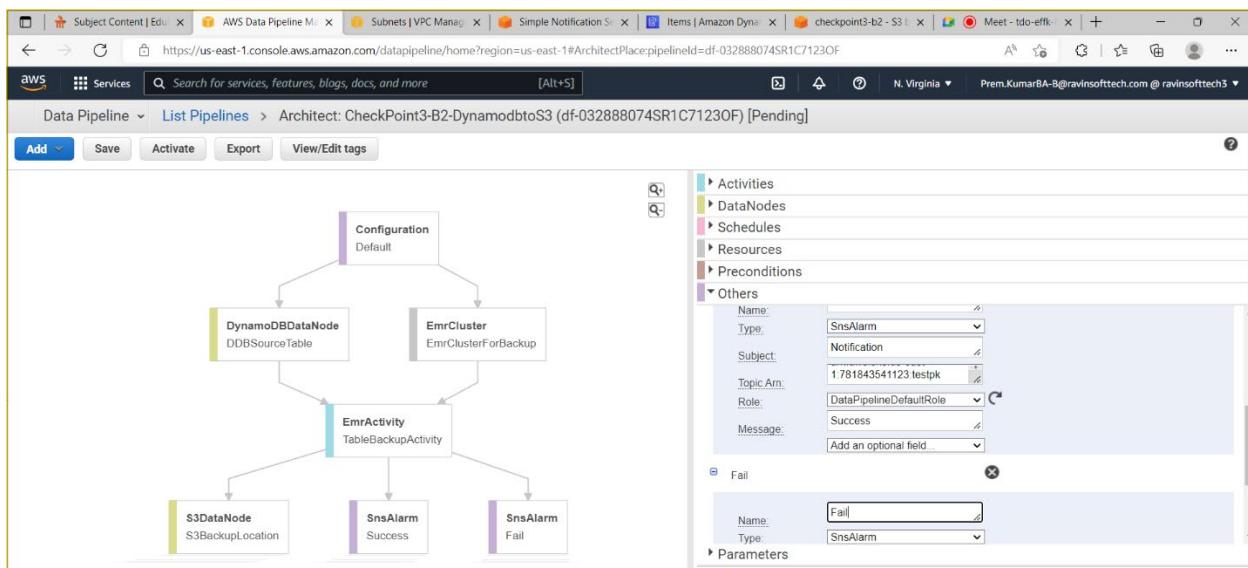
Subscriptions

Subscriptions (1)

ID	Endpoint	Status	Protocol

Actions

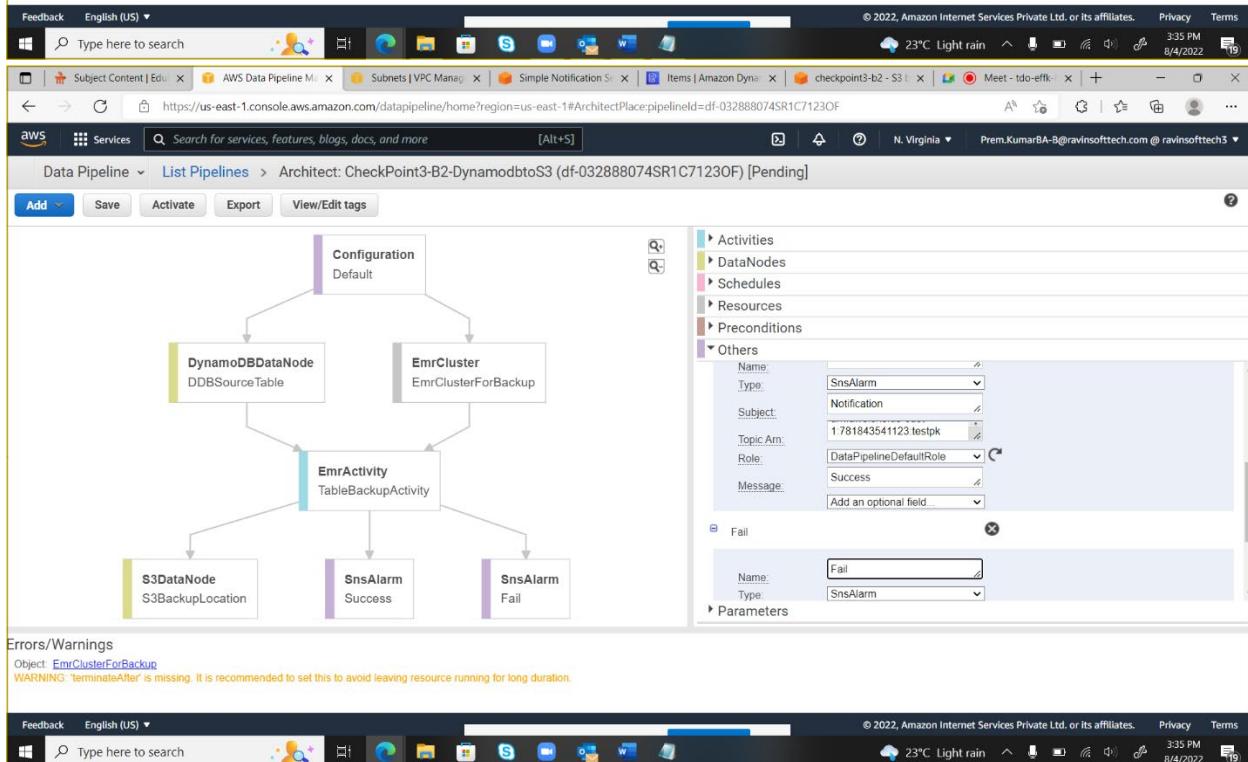
- Edit
- Delete
- Request confirmation
- Confirm subscription
- Create subscription



Errors/Warnings

Object: EmrClusterForBackup

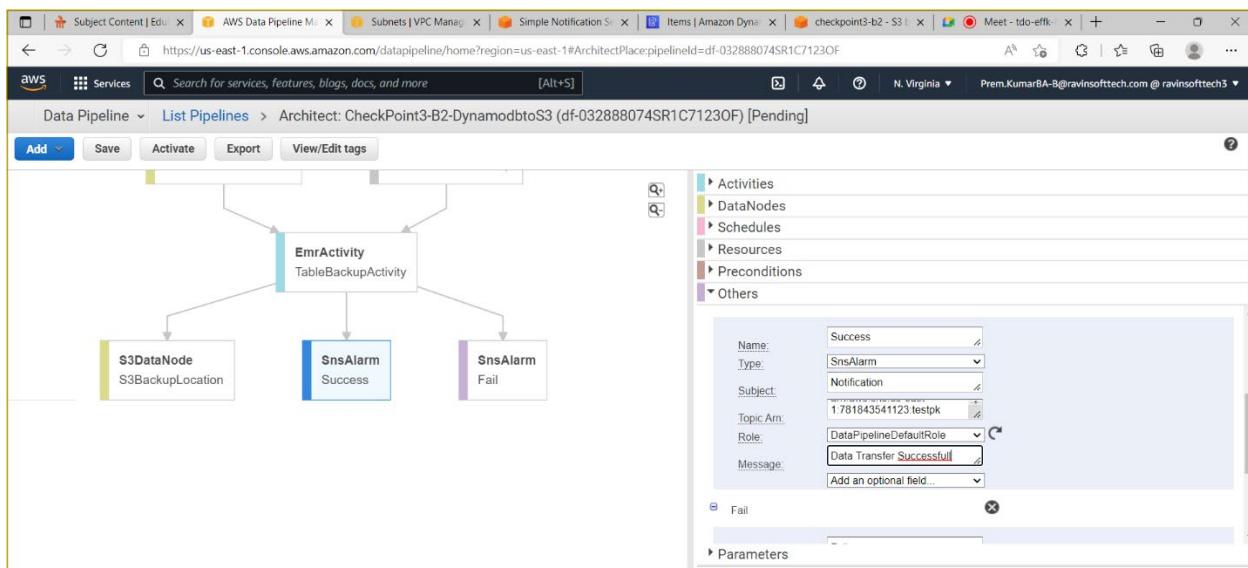
WARNING: 'terminateAfter' is missing. It is recommended to set this to avoid leaving resource running for long duration.



Errors/Warnings

Object: EmrClusterForBackup

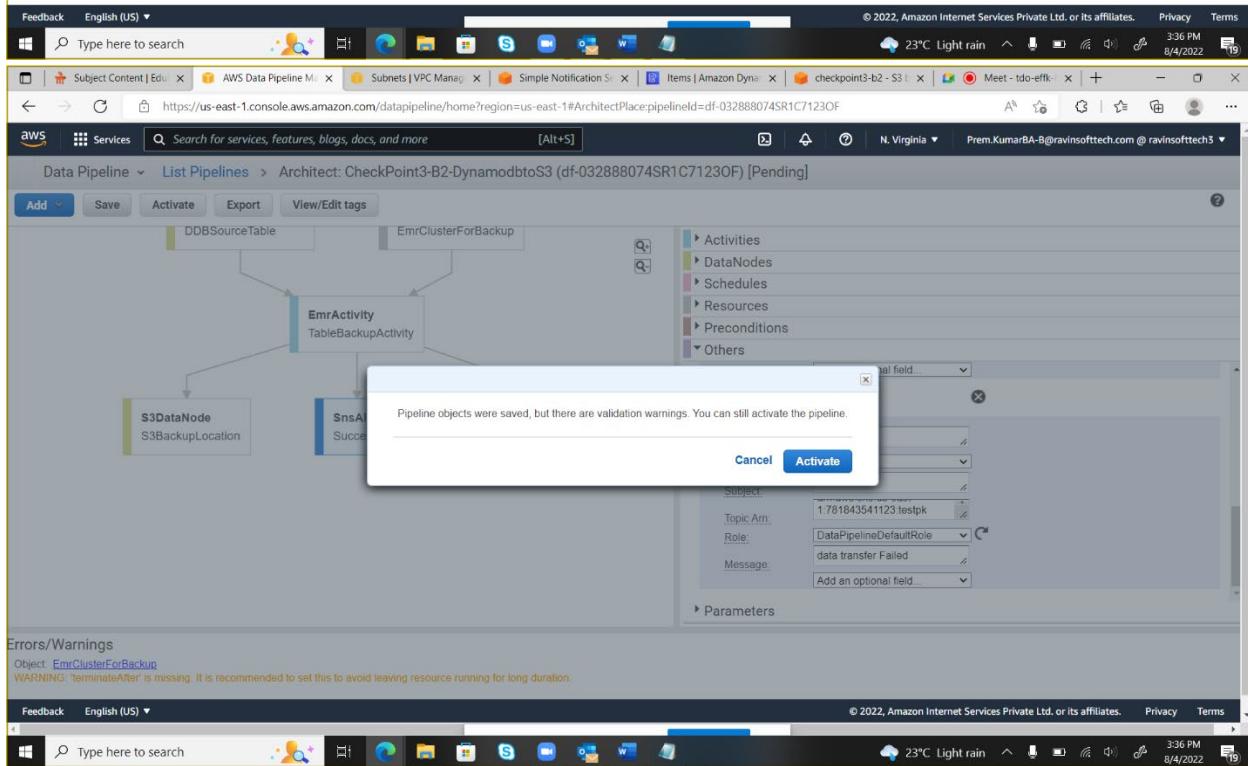
WARNING: 'terminateAfter' is missing. It is recommended to set this to avoid leaving resource running for long duration.



Errors/Warnings

Object: EmrClusterForBackup

WARNING: 'terminateAfter' is missing. It is recommended to set this to avoid leaving resource running for long duration.



Errors/Warnings

Object: EmrClusterForBackup

WARNING: 'terminateAfter' is missing. It is recommended to set this to avoid leaving resource running for long duration.

Screenshot of the AWS Data Pipeline console showing the execution details of a pipeline named "CheckPoint3-B2-DynamodbtoS3".

The pipeline has one component, "TableBackupActivity", which is an EmrActivity. It was scheduled for August 4, 2022, between 10:06:36 and 10:06:38 UTC. The status is "WAITING_FOR_RUNNER", and it started at 10:06:38 UTC on August 4, 2022.

Component Name	Schedule Interval (UTC)	Type	Status	Execution Start (UTC)	Execution End (UTC)	Attempt
TableBackupActivity	2022-08-04 – 2022-08-04 10:06:36 – 10:06:38	EmrActivity	WAITING_FOR_RUNNER	2022-08-04 10:06:38	–	1 of 3

Screenshot of the AWS Data Pipeline console showing the execution details of the same pipeline "CheckPoint3-B2-DynamodbtoS3".

The pipeline component "TableBackupActivity" has completed its execution. The status is now "FINISHED", and the execution ended at 10:13:44 UTC on August 4, 2022.

Component Name	Schedule Interval (UTC)	Type	Status	Execution Start (UTC)	Execution End (UTC)	Attempt
TableBackupActivity	2022-08-04 – 2022-08-04 10:06:36 – 10:06:38	EmrActivity	FINISHED	2022-08-04 10:06:38	2022-08-04 10:13:44	1 of 3



Screenshot of the AWS Data Pipeline console showing the list of pipelines:

Pipeline ID	Name	Schedule State	Health Status	Creation Time (UTC)
df-032888074SR1C7123OF	CheckPoint3-B2-DynamodbToS3	SCHEDULED Runs every 100 years	HEALTHY	2022-08-04 10:00:16
df-06030142LBFRQ25YNPCI	rmcf-dynamodb-to-s3-pipeline	PENDING	Pipeline is not active	2022-08-03 12:04:31
df-01431396WQVQ5XBJN5P	pipeline-dynamoDB-S3-ankur	FINISHED Runs every 100 years	HEALTHY	2022-08-04 05:28:28
df-039797211V96RVUTWFMBM	exportdatafromDynamoDBtoS3	FINISHED Runs every 100 years	HEALTHY	2022-08-03 11:10:10

Screenshot of the AWS S3 console showing the contents of the 'output' folder in the 'checkpoint3-b2' bucket:

Name	Type	Last modified	Size	Storage class
2022-08-04-10-06-36/	Folder	-	-	-

Screenshot of the AWS S3 console showing the contents of the 'output' folder in the 'checkpoint3-b2' bucket. The date selected is 2022-08-04-10-06-36.

The objects listed are:

Name	Type	Last modified	Size	Storage class
_SUCCESS	-	August 4, 2022, 15:43:29 (UTC+05:30)	0 B	Standard
a35bf897-05b6-4084-915f-ad75f3cda72b	-	August 4, 2022, 15:43:20 (UTC+05:30)	2.7 KB	Standard
manifest	-	August 4, 2022, 15:43:29 (UTC+05:30)	165.0 B	Standard

Below the S3 interface, there is a Microsoft Outlook window showing an email from AWS Notifications. The subject is '[EXT] Notification - Message (Plain Text)'. The body of the email contains the following text:

Data Transfer Successfull

--

If you wish to stop receiving notifications from this topic, please click or visit the link below to unsubscribe:
<https://sns.us-east-1.amazonaws.com/unsubscribe.html?SubscriptionArn=arn:aws:sns:us-east-1:781843541123:testpk1b72c4c4-8b49-49d7-929e-3c22f116b326&Endpoint=kprekumar@deloitte.com>

Please do not reply directly to this email. If you have any questions or comments regarding this email, please contact us at <https://aws.amazon.com/support>

3. Transfer the AWS s3 data to AWS Quicksight perform the same analysis doe using powerBI(Any 5 core reports)

The screenshot shows the AWS S3 service page. On the left, there's a sidebar with options like 'Buckets', 'Access Points', 'Storage Lens', and 'Feature spotlight'. The main area displays a table of buckets:

Name	AWS Region	Access	Creation date
aws-cloudtrail-logs-3rdaug	Not found		August 3, 2022, 10:49:10 (UTC+05:30)
aws-cloudtrail-logs-781843541123-b19c28a6	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 10:56:33 (UTC+05:30)
aws-cloudtrail-logs-aakash	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 11:33:38 (UTC+05:30)
aws-cloudtrail-logs-nishanth	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 11:45:41 (UTC+05:30)
aws-cloudtrail-logs3rdaug	US East (N. Virginia)	Bucket and objects	August 3, 2022, 10:55:12

The screenshot shows the 'Create bucket' wizard. The first step, 'General configuration', is displayed. It includes fields for 'Bucket name' (set to 'quicksight-bucket-capstone'), 'AWS Region' (set to 'US East (N. Virginia) us-east-1'), and a section for 'Copy settings from existing bucket - optional'.

The screenshot shows the AWS S3 console with the URL <https://s3.console.aws.amazon.com/s3/bucket/create?region=us-east-1>. The page displays options for blocking public access to buckets and objects. A warning message states: "Turning off block all public access might result in this bucket and the objects within becoming public. AWS recommends that you turn on block all public access, unless public access is required for specific and verified use cases such as static website hosting." A checkbox is checked, acknowledging this risk.

Block all public access

Turning this setting on is the same as turning on all four settings below. Each of the following settings are independent of one another.

- Block public access to buckets and objects granted through new access control lists (ACLs)
- Block public access to buckets and objects granted through any access control lists (ACLs)
- Block public access to buckets and objects granted through new public bucket or access point policies
- Block public and cross-account access to buckets and objects through any public bucket or access point policies

A Turning off block all public access might result in this bucket and the objects within becoming public
AWS recommends that you turn on block all public access, unless public access is required for specific and verified use cases such as static website hosting.

I acknowledge that the current settings might result in this bucket and the objects within becoming public.

The screenshot shows the AWS S3 console with the URL <https://s3.console.aws.amazon.com/s3/buckets?region=us-east-1>. A green success message box is displayed: "Successfully created bucket \"quicksight-bucket-capstone\". To upload files and folders, or to configure additional bucket settings choose View details." The main table lists existing buckets:

Name	AWS Region	Access	Creation date
aws-cloudtrail-logs-nishanth	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 11:45:41 (UTC+05:30)
aws-cloudtrail-logs3rdaug	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 10:55:12 (UTC+05:30)
cloudwatchbucketdemo	US East (N. Virginia) us-east-1	Objects can be public	August 3, 2022, 12:40:31 (UTC+05:30)
quicksight-bucket-capstone	US East (N. Virginia) us-east-1	Objects can be public	August 3, 2022, 14:50:34 (UTC+05:30)
quicksitebucket	US East (N. Virginia) us-east-1	Objects can be public	August 3, 2022, 09:12:38 (UTC+05:30)
sam-cloud-buck	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 11:46:44 (UTC+05:30)
sourav-aws-cloudtrail-logs	US East (N. Virginia) us-east-1	Bucket and objects not public	August 3, 2022, 11:54:51 (UTC+05:30)

[View details](#)

Screenshot of the AWS S3 console showing the 'quicksight-bucket-capstone' bucket.

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

Feedback Looking for language selection? Find it in the new Unified Settings [\[?\]](#)

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

CloudWatch S3 Manager quicksight Home Instances 816026083 ...

https://s3.console.aws.amazon.com/s3/buckets/quicksight-bucket-capstone?region=us-east-1&tab=... [Alt+S] Global Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3 ...

quicksight-bucket-capstone [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 \[?\]](#)

[G](#) [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open \[?\]](#) [Delete](#) [Actions ▾](#)

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

Find objects by prefix

Name	Type	Last modified	Size	Storage class
User_product_purchase_final.csv	csv	August 3, 2022, 14:51:57 (UTC+05:30)	42.5 MB	Standard

Screenshot of the Amazon QuickSight console showing a dataset named 'batch b2 dataset analysis'.

Visualize Dataset SPICE batch b2 dataset

Fields list

- City_Code
- City_Name
- Gender
- Marital_Status
- Occupation
- Product_Category_1
- Product_Category_2
- Product_Category_3
- Product_ID
- Purchase
- State
- State_Code
- Stay_In_Current_City_Years
- User_ID
- Zip_Code

Field wells

Sheet 1 +

Sum of Purchase by Age

The data source or dataset was deleted or became unavailable while Amazon QuickSight was ingesting the data into SPICE. You can import the data to try again.

Show details

Import failed:
Too many rows were skipped while importing data into SPICE. Review your data and try again.
[Retry import](#) [View import summary](#)

CloudWatch quicksight quicksight quicksight batcl Supp New quic redu ...

https://us-east-1.quicksight.aws.amazon.com/sn/analyses/5b96e2ae-b3f7-4fe2-8ffe-cd3f78e2e45f

25°C Rain 252 PM 8/3/2022

Launch Me Section Hover CloudWatch S3 Manager quicksight Home Instances 816026083 ...

[Alt+S] Global Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3 ...

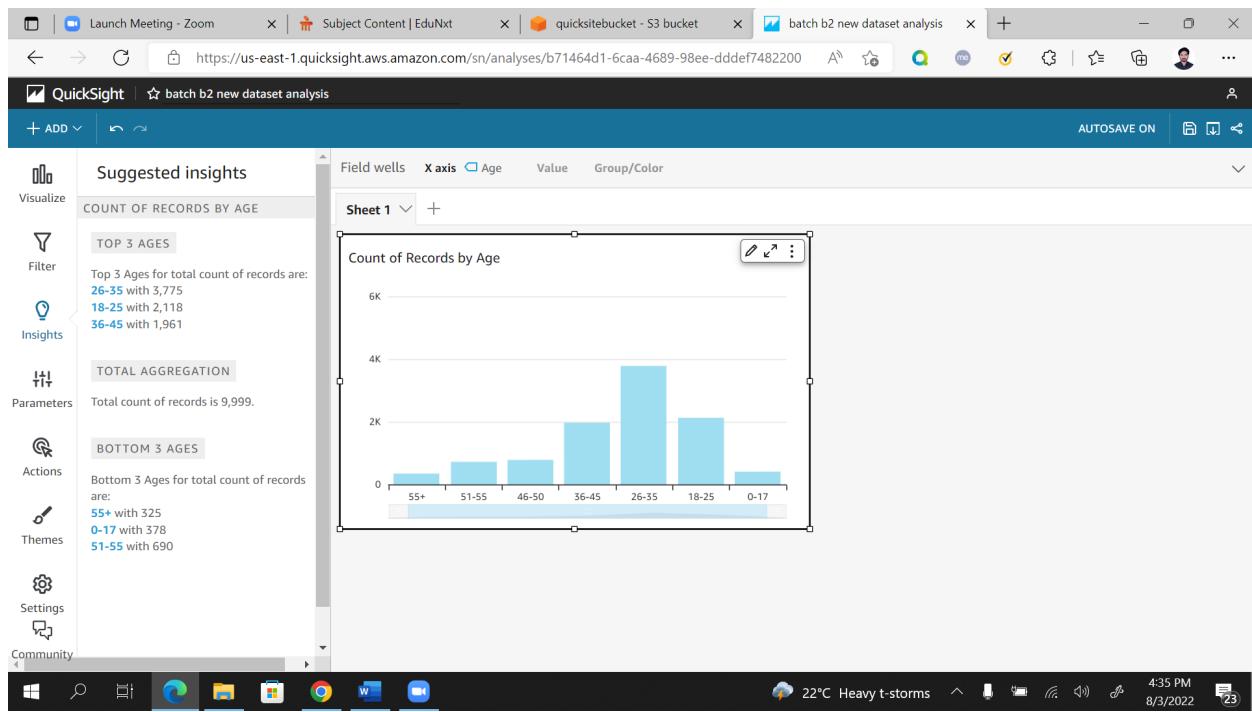
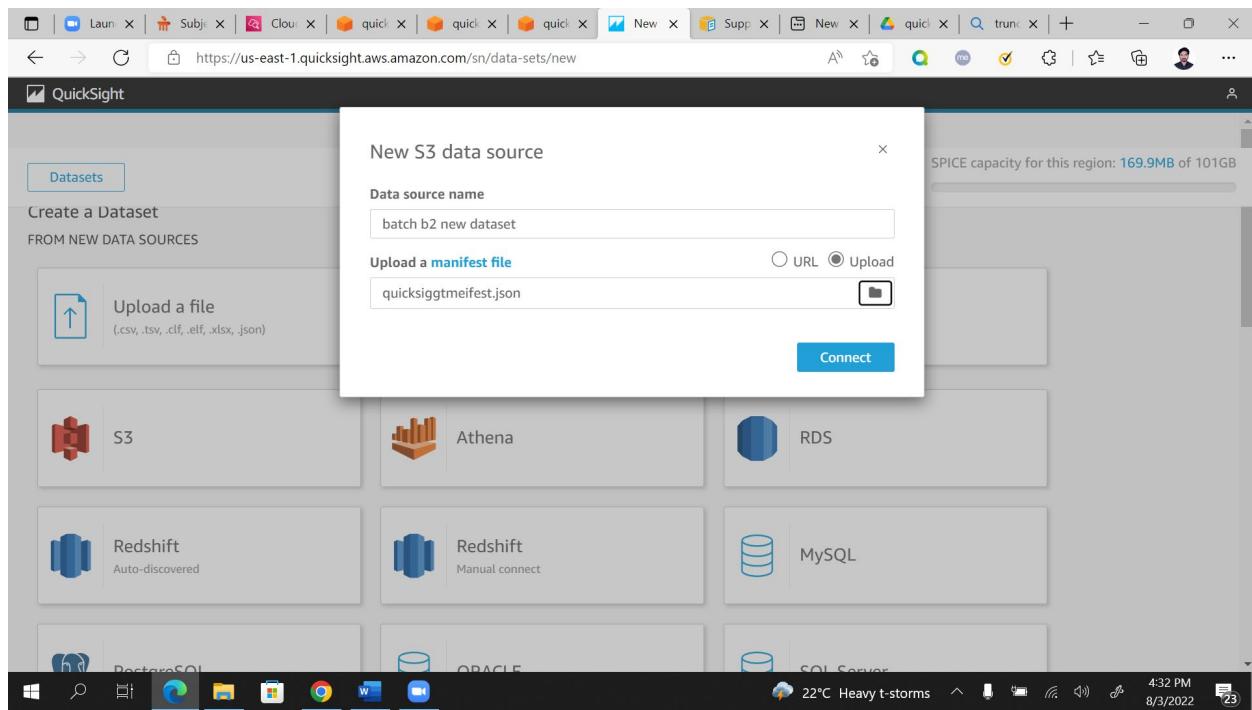
Feedback Looking for language selection? Find it in the new Unified Settings [\[?\]](#)

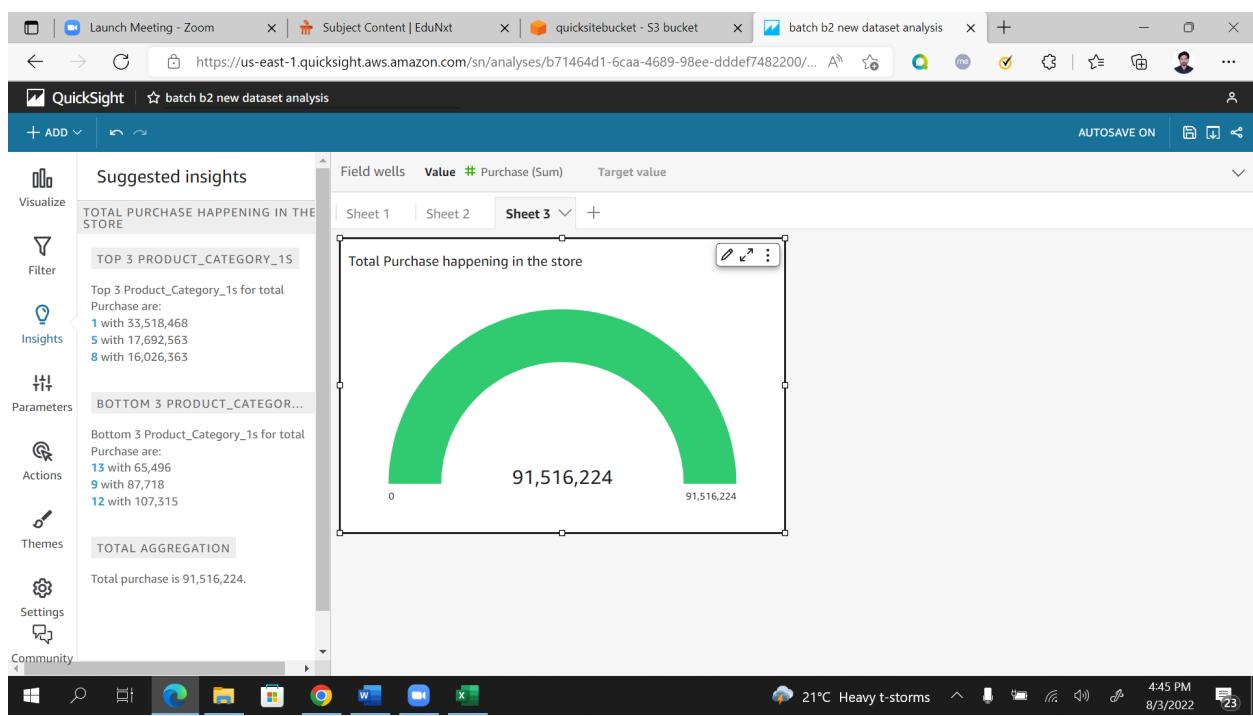
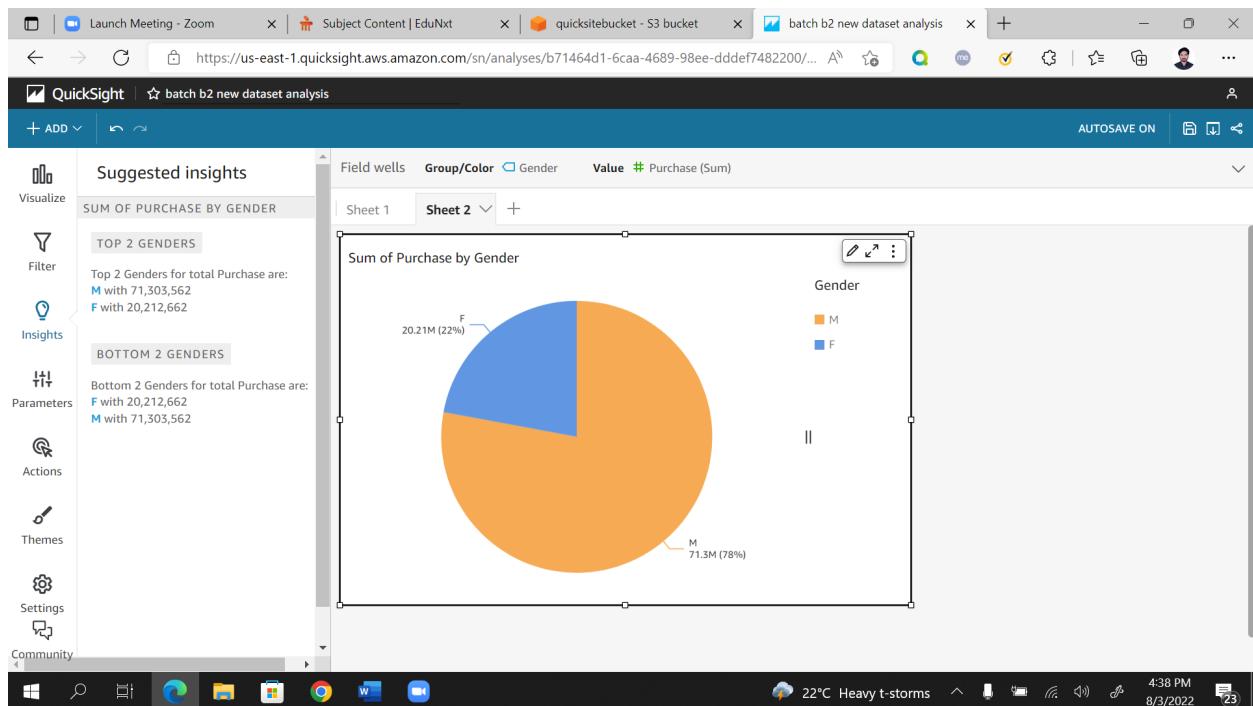
© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

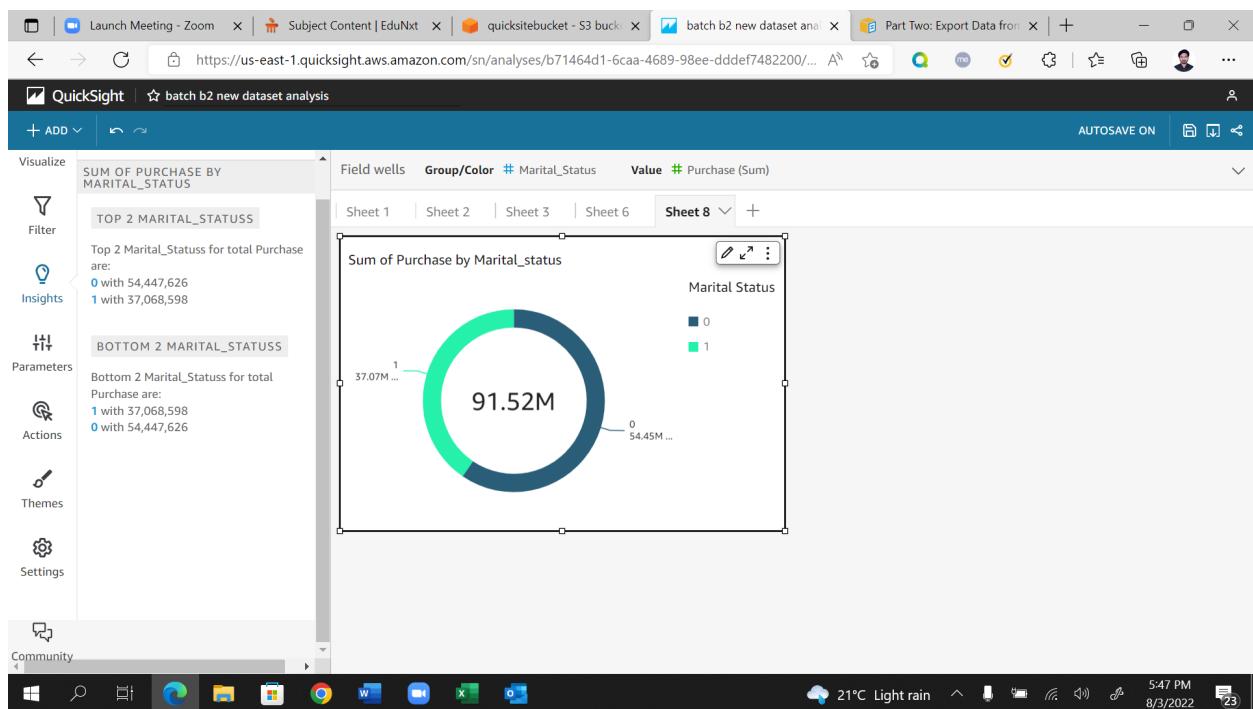
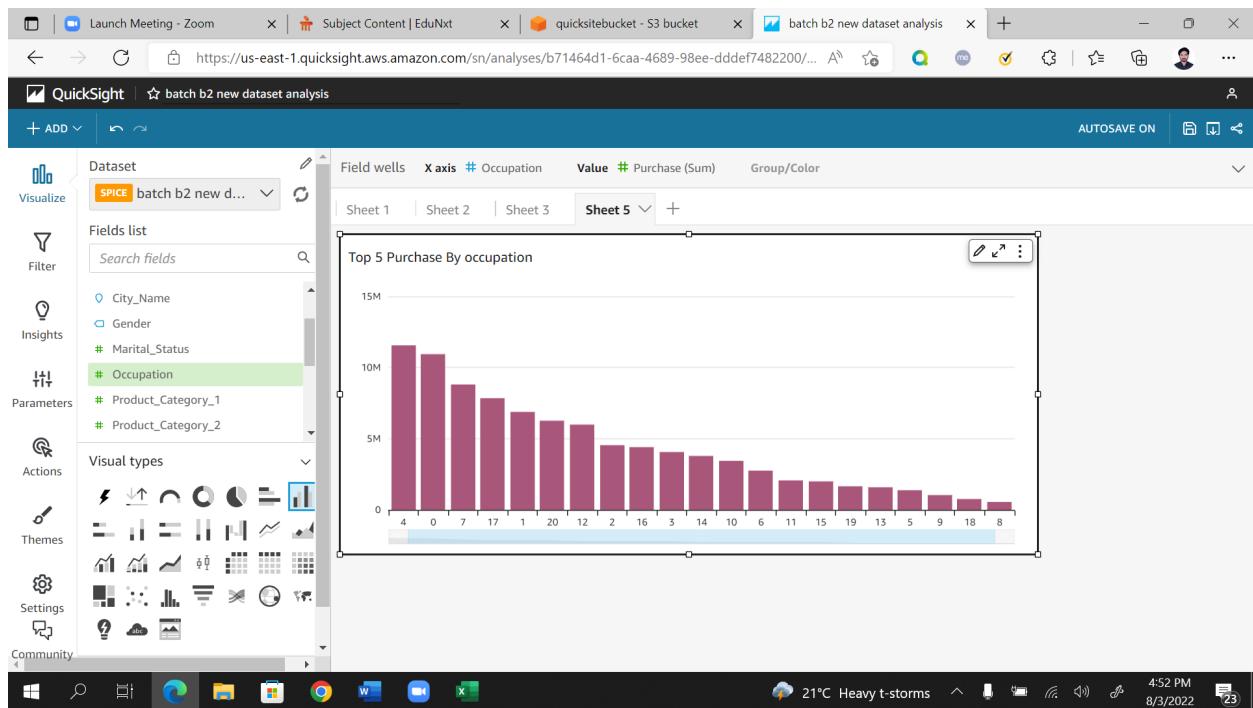
CloudWatch S3 Manager quicksight Home Instances 816026083 ...

https://us-east-1.quicksight.aws.amazon.com/sn/analyses/5b96e2ae-b3f7-4fe2-8ffe-cd3f78e2e45f [Alt+S] Global Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3 ...

422 PM 8/3/2022







4. 50% of storage exceeds then cloud watch has to trigger the alaram

Create bucket Info

Buckets are containers for data stored in S3. [Learn more](#)

General configuration

Bucket name

Bucket name must be globally unique and must not contain spaces or uppercase letters. [See rules for bucket naming](#)

AWS Region

Copy settings from existing bucket - *optional*
Only the bucket settings in the following configuration are copied.

[Choose bucket](#)

Object Ownership Info

Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.

Select metric

18:30 21:30 00:30 03:30 06:30 09:30 12:30 15:30 18:30 21:30 00:30 03:30 06:30

[Browse](#) [Query](#) [Graphed metrics \(1\)](#) [Options](#) [Source](#) [Add math](#) [Add query](#)

Metrics (8)

All > S3 > Storage Metrics

BucketName (8)	StorageType	Metric name
<input type="checkbox"/> muskan0604 ▾	AllStorageTypes ▾	NumberOfObjects ▾
<input checked="" type="checkbox"/> muskan0604 ▾	StandardStorage ▾	BucketSizeBytes ▾

[Graph with SQL](#) [Graph search](#)

[Cancel](#) [Select metric](#)

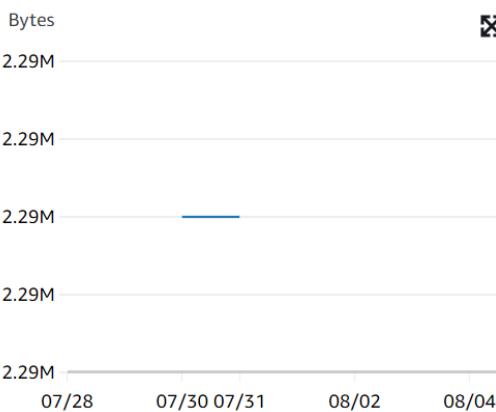
Specify metric and conditions

Metric

[Edit](#)

Graph

This alarm will trigger when the blue line goes above the red line for 1 datapoints within 1 day.



Namespace

AWS/S3

Metric name

BucketSizeBytes

StorageType

StandardStorage

BucketName

Threshold type

 Static

Use a value as a threshold

 Anomaly detection

Use a band as a threshold

Whenever BucketSizeBytes is...

Define the alarm condition.

 Greater

> threshold

 Greater/Equal

>= threshold

 Lower/Equal

<= threshold

 Lower

< threshold

than...

Define the threshold value.

1



Must be a number

► Additional configuration

[Cancel](#)[Next](#)

Notification

Alarm state trigger

Define the alarm state that will trigger this action.

[Remove](#)

In alarm

The metric or expression is outside of the defined threshold.

OK

The metric or expression is within the defined threshold.

Insufficient data

The alarm has just started or not enough data is available.

Send a notification to the following SNS topic

Define the SNS (Simple Notification Service) topic that will receive the notification.

Select an existing SNS topic

Create new topic

Use topic ARN to notify other accounts

Send a notification to...



Only email lists for this account are available.

Email (endpoints)

Topic is configured with email address and a subscription.

Upload the file to the bucket

Files and folders Configuration

Files and folders (1 Total, 5.9 MB)

Name	Folder	Type	Size	Status	Error
datapielinness.zip	-	application/x-zip-compressed	5.9 MB	Succeeded	-

On checking email, we saw this notification

Alarm Details:

- Name: s3
- Description:
- State Change: INSUFFICIENT_DATA -> ALARM
- Reason for State Change: Threshold Crossed: 1 out of the last 1 datapoints [2294504.0 (31/07/22 07:21:00)] was greater than the threshold (1.0) (minimum 1 datapoint for OK -> ALARM transition)
- Timestamp: Wednesday 03 August, 2022 07:21:37 UTC
- AWS Account: 781843541123
- Alarm Arn: arn:aws:cloudwatch:us-east-1:781843541123:alarm:s3

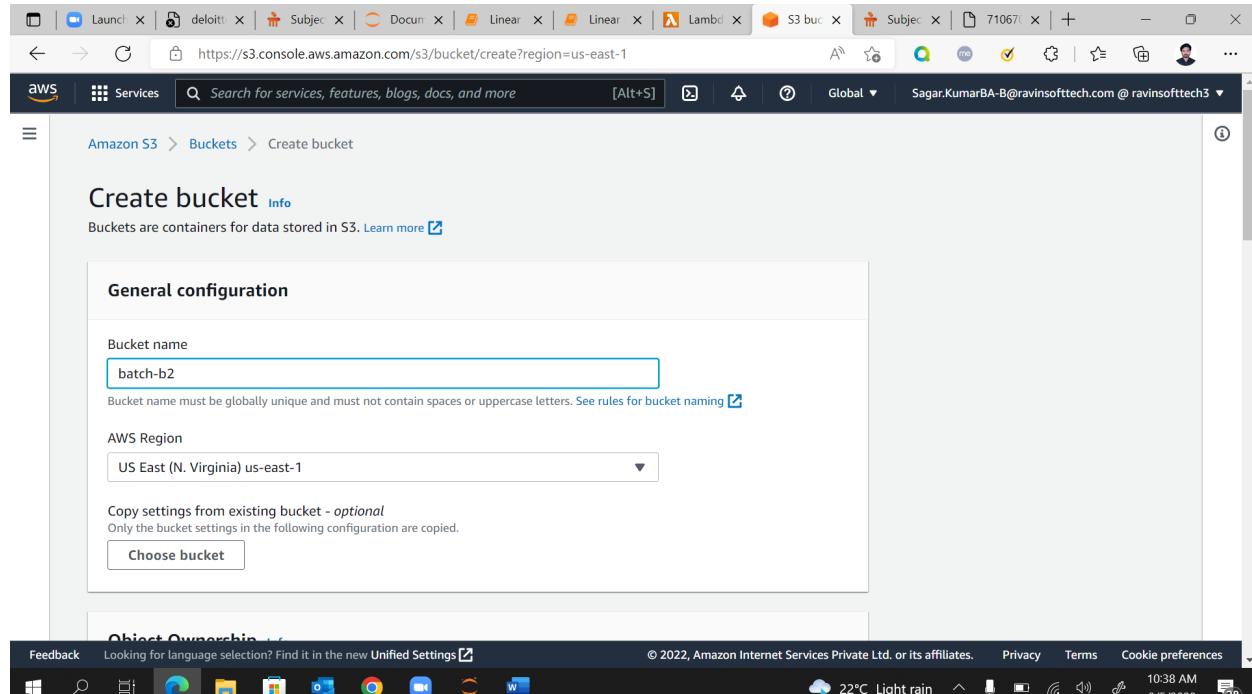
Threshold:

- The alarm is in the ALARM state when the metric is GreaterThanThreshold 1.0 for at least 1 of the last 1 period(s) of 86400 seconds.

Monitored Metric:

- MetricNamespace: AWS/S3
- MetricName: BucketSizeBytes
- Dimensions: [BucketName = muskan0604] [StorageType = StandardStorage]
- Period: 86400 seconds
- Statistic: Average
- Unit: not specified
- TreatMissingData: missing

5. Write a Lambda function which logs in cloud trail about S3 file type and size.



Create bucket Info

Buckets are containers for data stored in S3. [Learn more](#)

General configuration

Bucket name Bucket name must be globally unique and must not contain spaces or uppercase letters. [See rules for bucket naming](#)

AWS Region

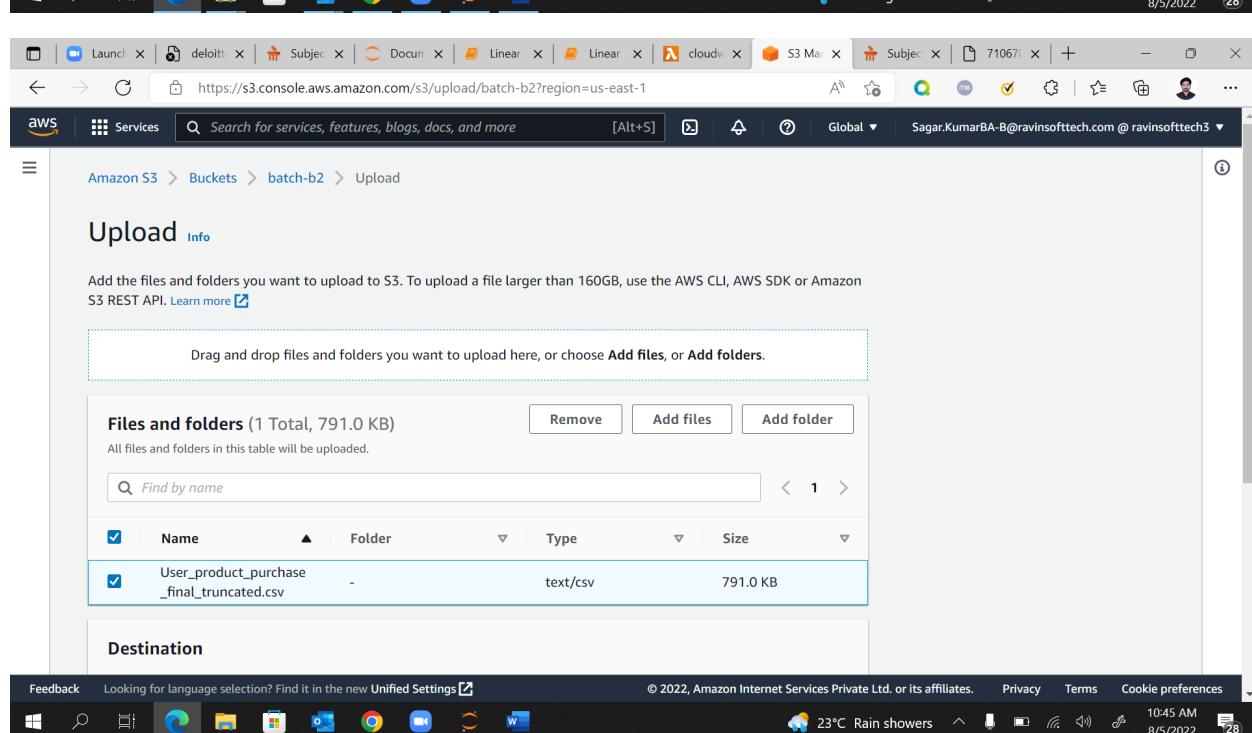
Copy settings from existing bucket - *optional*
Only the bucket settings in the following configuration are copied.
[Choose bucket](#)

Object Ownership

Feedback Looking for language selection? Find it in the new Unified Settings [Learn more](#)

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

Windows 10 taskbar: Launch, deloitte, Subject, Document, Linear, Linear, Lambda, S3 bucket, Subject, 71067, etc.



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, 791.0 KB)

All files and folders in this table will be uploaded.

<input checked="" type="checkbox"/>	Name	Folder	Type	Size
<input checked="" type="checkbox"/>	User_product_purchase_final_truncated.csv	-	text/csv	791.0 KB

Destination

Feedback Looking for language selection? Find it in the new Unified Settings [Learn more](#)

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

Windows 10 taskbar: Launch, deloitte, Subject, Document, Linear, Linear, Lambda, S3 bucket, Subject, 71067, etc.

Launch x deloitte x Subject C x Document x Linear Re x Linear Re x Lambda x Subject C x 7106707 x + – □ ...

aws | Services Search for services, features, blogs, docs, and more [Alt+S] ... N. Virginia Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech ...

Lambda > Functions > Create function > Configure blueprint s3-get-object-python ...

Basic information Info

Function name cloudwatch-storage

Execution role
Choose a role that defines the permissions of your function. To create a custom role, go to the [IAM console](#).

Create a new role with basic Lambda permissions
 Use an existing role
 Create a new role from AWS policy templates

ⓘ Role creation might take a few minutes. Please do not delete the role or edit the trust or permissions policies in this role.

Role name
Enter a name for your new role.
myRoleName

Use only letters, numbers, hyphens, or underscores with no spaces.

Feedback Looking for language selection? Find it in the new [Unified Settings](#) ... © 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

... 10:36 AM 8/5/2022 28

```
2 "Records": [
3 {
4     "eventVersion": "2.0",
5     "eventSource": "aws:s3",
6     "awsRegion": "us-east-1",
7     "eventTime": "1970-01-01T00:00:00.000Z",
8     "eventName": "ObjectCreated:Put",
9     "userIdentity": {
10         "principalId": "EXAMPLE"
11     },
12     "requestParameters": {
13         "sourceIPAddress": "127.0.0.1"
14     },
15     "responseElements": {
16         "x-amz-request-id": "EXAMPLE123456789",
17         "x-amz-id-2": "EXAMPLE123/5678abcdefghijklmabdaisawesome/mnopqrstuvwxyzABCDEFGH"
18     },
19     "s3": {
20         "s3SchemaVersion": "1.0",
21         "configurationId": "testConfigRule",
22         "bucket": {
23             "name": "batch-b2",
24             "ownerIdentity": {
25                 "principalId": "EXAMPLE"
26             },
27             "arn": "arn:aws:s3:::batch-b2"
28         },
29         "object": {
30             "key": "index.html",
31             "size": 1024,
32             "eTag": "0123456789abcdef0123456789abcdef",
33             "versionId": "20220726T123456Z"
34         }
35     }
36 }
```

The test event Cloudtraillogs was successfully saved.

Code source Info

File Edit Find View Go Tools Window Test Deploy

Execution results

Test Event Name Cloudtraillogs

Response "text/csv"

Function Logs

```
START RequestId: dab3354a-5163-49ca-b9a2-75d00fccc401 Version: $LATEST
CONTENT TYPE: text/csv
END RequestId: dab3354a-5163-49ca-b9a2-75d00fccc401
REPORT RequestId: dab3354a-5163-49ca-b9a2-75d00fccc401 Duration: 242.00 ms Billed Duration: 242 ms Memory Size: 128 MB Max Memory
```

Request ID dab3354a-5163-49ca-b9a2-75d00fccc401

Feedback Looking for language selection? Find it in the new Unified Settings © 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

23°C Rain showers 10:53 AM 8/5/2022 28

CloudWatch

Favorites and recents

Dashboard

Alarms △ 2 ○ 6 ⚡ 9

In alarm

All alarms

Billing

Logs

Log groups

Logs Insights

Metrics

All metrics

Explorer

Streams

Log streams (1/2)

Retention Never expire

Creation time 11 minutes ago

Stored bytes -

ARN arn:aws:logs:us-east-1:781843541123:log-group:/aws/lambda/cloudwatch-storage:*

KMS key ID -

Metric filters 0

Subscription filters 0

Contributor Insights rules -

Log stream 2022/08/05[\$LATEST]06b304117fd9495f841776... Last event time 2022-08-05 10:53:41 (UTC+05:30)

Log stream 2022/08/05[\$LATEST]adb7eba8d3f5464eabf57a... Last event time 2022-08-05 10:44:31 (UTC+05:30)

Feedback Looking for language selection? Find it in the new Unified Settings © 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

23°C Rain showers 10:55 AM 8/5/2022 28

The screenshot shows the AWS CloudWatch Logs console for the us-east-1 region. The left sidebar navigation includes 'CloudWatch' (selected), 'Dashboards', 'Alarms' (with 2 alerts, 6 in alarm, 9 total), 'Logs' (selected), 'Log groups' (highlighted in orange), 'Logs Insights', 'Metrics', 'All metrics', 'Explorer', and 'Streams'. The main content area displays log events for a Lambda function. The first event is a 'Loading function' message. Subsequent events show the Lambda starting, processing a request (RequestID: dab3354a-5163-49ca-b9a2-75d00fccd401), and returning a response (RequestID: dab3354a-5163-49ca-b9a2-75d00fccd401). The final event is a 'REPORT' message indicating the function took 242.00 ms and was billed 242 ms. A message at the bottom states 'No newer events at this moment. Auto retry paused. Resume'.

The screenshot shows the AWS CloudWatch Alarms console for the us-east-1 region. The left sidebar navigation includes 'CloudWatch' (selected), 'Dashboards', 'Alarms' (with 0 in alarm, 0 total, 5 total), 'Logs', 'Metrics', 'X-Ray traces', 'Events', 'Application monitoring', 'Insights', and 'Settings'. The main content area displays a table titled 'Alarms (5)'. The table columns are 'Name', 'State', 'Last state update', and 'Conditions'. The alarms listed are:

Name	State	Last state update	Conditions
s3	Insufficient data	2022-08-04 05:31:37	BucketSizeBytes > 1 for 1 datapoints within 1 day
Ec_2_alarm	Insufficient data	2022-08-03 18:20:17	CPUUtilization >= 70 for 1 datapoints within 1 minute
EC2Alarm	Insufficient data	2022-08-03 14:48:01	CPUUtilization >= 70 for 1 datapoints within 1 minute
s3 alarm	Insufficient data	2022-08-03 12:22:34	BucketSizeBytes >= 1 for 1 datapoints within 1 minute
ec2alarm	Insufficient data	2022-08-03 12:07:53	CPUUtilization >= 70 for 1 datapoints within 1 minute

AWS Management Console | S3 Management Console | CloudWatch Management Console | Simple Notification Service | +

https://s3.console.aws.amazon.com/s3/buckets?region=us-east-1

Services | Search for services, features, blogs, docs, and more [Alt+S] | Global | Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3

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

Successfully created bucket "codepipeline-us-east-1-749217320090"
To upload files and folders, or to configure additional bucket settings choose View details.

Name	AWS Region	Access	Creation date
capstoneb2	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 10:38:59 (UTC+05:30)
checkpoint3-b1	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 09:52:25 (UTC+05:30)
checkpoint3-b1-sagemaker	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 14:45:35 (UTC+05:30)
checkpoint3pk	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 13:49:10 (UTC+05:30)
cloudtrail-ankur-usedcars	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 14:24:48 (UTC+05:30)
codepipeline-us-east-1-749217320090	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 15:17:41 (UTC+05:30)
lambda-s3-07	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 08:46:48 (UTC+05:30)
muskan0604	US East (N. Virginia) us-east-1	Objects can be public	August 4, 2022, 14:45:33 (UTC+05:30)

s3.console.aws.amazon.com/s3/buckets/codepipeline-us-east-1-749217320090... settings

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

22°C Rain showers 3:17 PM 8/4/2022 28

AWS Management Console | S3 Management Console | CloudWatch Management Console | Simple Notification Service | +

https://us-east-1.console.aws.amazon.com/cloudwatch/home?region=us-east-1#alarmsV2:create?~(P...

Services | Search for services, features, blogs, docs, and more [Alt+S] | Global | N. Virginia | Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3

Configure actions

Step 3 Add name and description

Step 4 Preview and create

Alarm state trigger
Define the alarm state that will trigger this action.

In alarm
The metric or expression is outside of the defined threshold.

OK
The metric or expression is within the defined threshold.

Insufficient data
The alarm has just started or not enough data is available.

Remove

Send a notification to the following SNS topic
Define the SNS (Simple Notification Service) topic that will receive the notification.

Select an existing SNS topic

Create new topic

Use topic ARN to notify other accounts

Send a notification to...

capstone-alarm

Only email lists for this account are available.

Email (endpoints)
kumarsagar620@gmail.com - View in SNS Console

Add notification

Feedback Looking for language selection? Find it in the new Unified Settings

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

22°C Rain showers 3:19 PM 8/4/2022 28

Screenshot of the AWS CloudWatch Metrics selection interface.

The interface shows a graph at the top with a Y-axis from 0 to 0.5 and an X-axis from 01:45 to 04:30. Below the graph is a navigation bar with tabs: **Browse** (selected), Query, Graphed metrics, Options, and Source. Buttons for Add math and Add query are also present.

The main area displays a grid of metrics:

S3	8	SNS	44
AWS/SecretsManager	1	States	4
SageMaker	40	Usage	470

At the bottom right of the grid is a button labeled "Select a single metric to continue".

Screenshot of the AWS CloudWatch Create alarm step 3 interface.

The left sidebar shows the steps: Step 1 Specify metric and conditions, Step 2 Configure actions, Step 3 Add name and description (selected), and Step 4 Preview and create.

The main area is titled "Add name and description". It contains fields for "Alarm name" (BucketSizeAlarm) and "Alarm description - optional" (with a placeholder "Alarm description" and character limit "Up to 1024 characters (0/1024)").

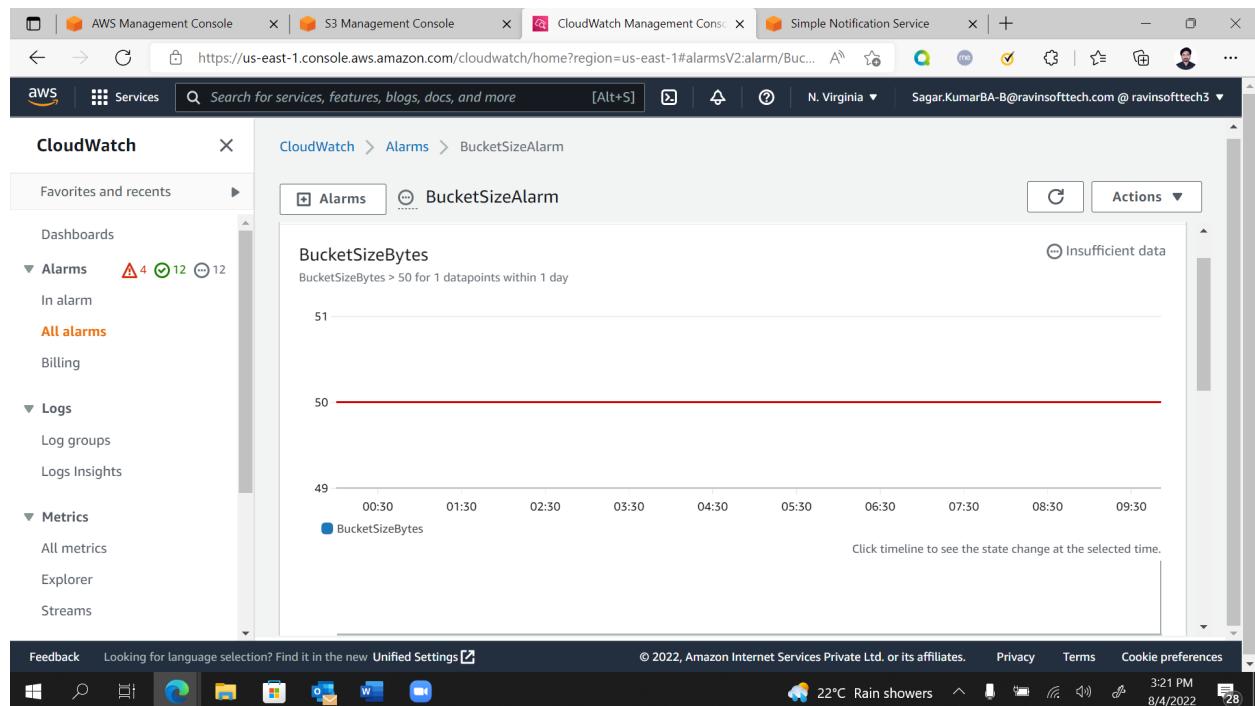
Buttons at the bottom right include Cancel, Previous, and Next (highlighted).

Screenshot of the AWS CloudWatch Metrics console showing the 'Select metric' step. A graph shows a single data series over time from 01:45 to 04:30. Below the graph, there are tabs for 'Browse', 'Query', 'Graphed metrics', 'Options', and 'Source'. A search bar at the top right says 'Search for services, features, blogs, docs, and more'. The main area displays 'Metrics (8)' under 'Storage Metrics'.

Screenshot of the AWS CloudWatch Metrics console showing the 'CloudWatch' dashboard. The left sidebar lists 'Alarms' (4 pending, 12 in alarm), 'Logs', and 'Metrics'. The main pane shows a list of alarms:

Name	State	Last state update	Conditions
BucketSizeAlarm	Insufficient data	2022-08-04 15:20:45	BucketSizeBytes > 50 for 1 datapoints within 1 day
alarm-ankur-usedcars	Insufficient data	2022-08-04 14:57:28	BucketSizeBytes >= 2000000 for 1 datapoints within 1 minute
cloudwatchts3alarm	Insufficient data	2022-08-04 14:46:48	BucketSizeBytes > 1 for 1 datapoints within 1

A notification bar at the top indicates 'Some subscriptions are pending confirmation' and 'Successfully created alarm BucketSizeAlarm'. The status bar at the bottom shows '22°C Rain showers' and the date '8/4/2022'.



6. Build appropriate ML model/s on the data using AWS Sagemaker , Identify the right metric to evaluate the performance of the model and Deploy on AWS Sagemaker

The screenshots show the AWS SageMaker console interface across three different stages of configuration:

- Screenshot 1: Create notebook instance**

This screen shows the "Notebook instance settings" section. The "Notebook instance name" is set to "sagemaker-b2". The "Notebook instance type" is "ml.t3.medium". The "Platform identifier" is "Amazon Linux 2, Jupyter Lab 1". There is also an "Additional configuration" section.
- Screenshot 2: Permissions and encryption**

This screen shows the "Permissions and encryption" section. The "IAM role" is set to "AmazonSageMaker-ExecutionRole-20220802T070555". The "Root access" option is selected ("Enable - Give users root access to the notebook"). The "Encryption key" dropdown is set to "No Custom Encryption".
- Screenshot 3: Network and Git repositories**

This screen shows optional configurations: "Network" and "Git repositories". Both sections have a note indicating they are optional.

Amazon SageMaker

Getting started

Control panel

- Studio
- Studio Lab
- Canvas
- RStudio

SageMaker dashboard

- Images
- Lifecycle configurations
- Search

► Ground Truth

▼ Notebook

Feedback Looking for language selection? Find it in the new Unified Settings

https://us-east-1.console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebook-instances

N. Virginia Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3

We want to hear from you. Take our short 5-10 minute survey and help us make SageMaker better for customers like you.

Success! Your notebook instance is being created.

Open the notebook instance when status is InService and open a template notebook to get started.

View details

Amazon SageMaker Notebook instances

Notebook instances

Name	Instance	Creation time	Status	Actions
sagemaker-b2	ml.t3.medium	Aug 04, 2022 05:15 UTC	InService	Open Jupyter Open JupyterLab
demonotebook	ml.t3.medium	Aug 04, 2022 05:13 UTC	InService	Open Jupyter Open JupyterLab

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

Cloud 25°C Light rain 10:52 AM 8/4/2022 25

jupyter

Files Running Clusters SageMaker Examples Conda

Select items to perform actions on them.

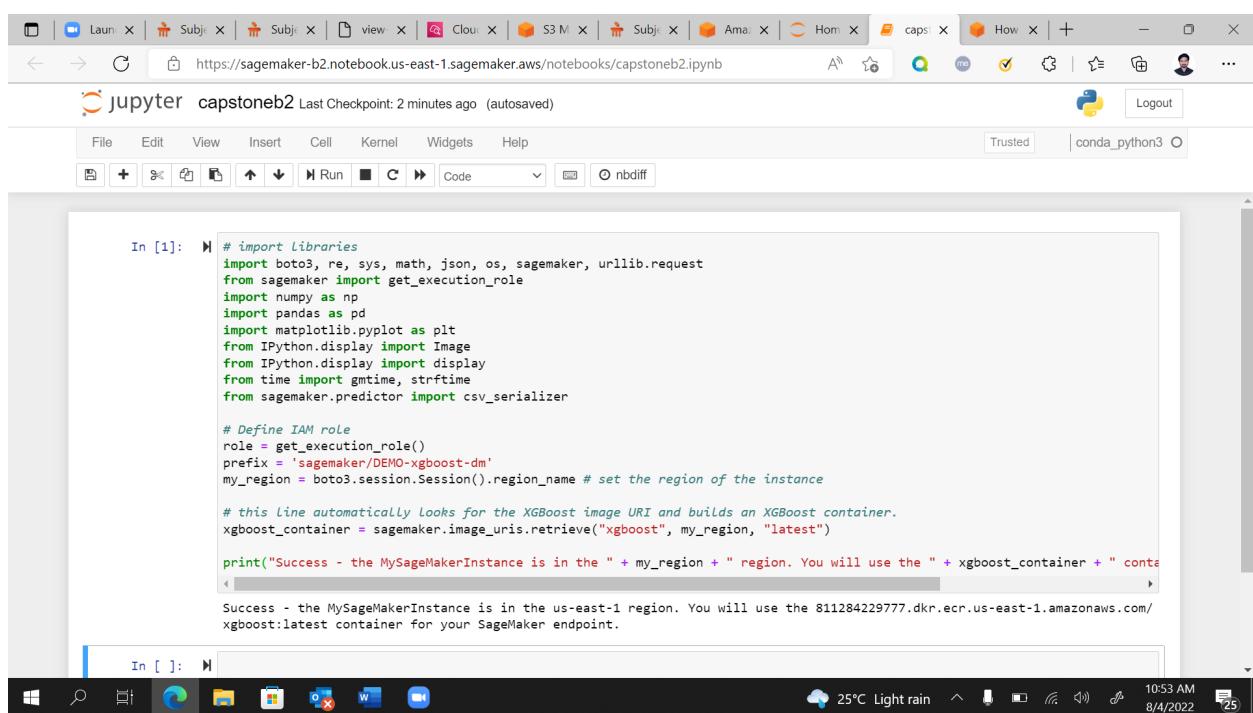
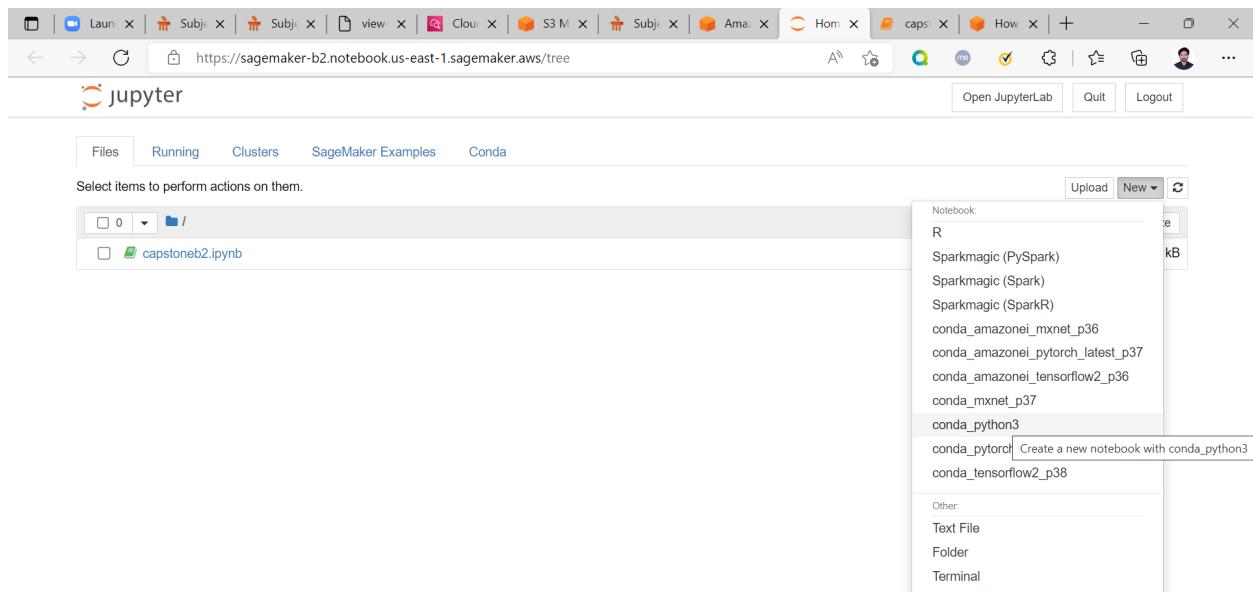
Upload New

0 / Untitled.ipynb

Name Last Modified File size

Running 2 minutes ago 72 B

Cloud 25°C Light rain 10:52 AM 8/4/2022 25



The screenshot shows a Jupyter Notebook interface running on AWS SageMaker. The notebook is titled "capstoneb2" and has a last checkpoint at 4 minutes ago. The code cell contains Python code for creating an S3 bucket:

```
prefix = sagemaker.vpc.xgboost_uris.XGBContainer_uris.sagemaker_session().region_name # set the region of the instance
my_region = boto3.Session().region_name
# this line automatically looks for the XGBoost image URI and builds an XGBoost container.
xgboost_container = sagemaker.image_uris.retrieve("xgboost", my_region, "latest")
print("Success - the MySageMakerInstance is in the " + my_region + " region. You will use the " + xgboost_container + " container")
Success - the MySageMakerInstance is in the us-east-1 region. You will use the 811284229777.dkr.ecr.us-east-1.amazonaws.com/xgboost:latest container for your SageMaker endpoint.
```

In []: In the code cell, the user has typed:

```
bucket_name = 'capstone-b2-sagemaker' # <-- CHANGE THIS VARIABLE TO A UNIQUE NAME FOR YOUR BUCKET
s3 = boto3.resource('s3')
try:
    if my_region == 'us-east-1':
        s3.create_bucket(Bucket=bucket_name)
    else:
        s3.create_bucket(Bucket=bucket_name, CreateBucketConfiguration={'LocationConstraint': my_region})
    print('S3 bucket created successfully')
except Exception as e:
    print('S3 error:', e)
```

The screenshot shows the AWS S3 console. The left sidebar shows the navigation path: Amazon S3 > Buckets > capstone-b2-sagemaker > output/. The main area displays the contents of the 'output/' folder, which is currently empty. The 'Objects' tab is selected, showing one item: linear-learner-2022-08-04-11-58-09-222/. The status bar at the bottom indicates the date and time as 8/4/2022 and 5:40 PM.

The screenshot shows the AWS S3 console interface. The left sidebar has a tree view with 'Amazon S3' selected, under 'Buckets'. The main area shows the 'output/' folder within the 'capstone-b2-sagemaker' bucket. The 'Objects' tab is selected, showing one item: 'linear-learner-2022-08-04-11-58-09-222/'. The 'Upload' button is highlighted in orange. The top navigation bar shows the URL as https://s3.console.aws.amazon.com/s3/buckets/capstone-b2-sagemaker?region=us-east-1&prefix=o... and the user as Sagar.KumarBA-B@ravinsofttech.com @ ravinsofttech3.

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

Feedback Looking for language selection? Find it in the new Unified Settings

© 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

23°C Rain 5:40 PM 8/4/2022 28

Name	Type	Last modified	Size	Storage class
linear-learner-2022-08-04-11-58-09-222/	Folder	-	-	-

