D Y PATIL INTERNATIONAL UNIVERSITY
AKURDI PUNE

School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

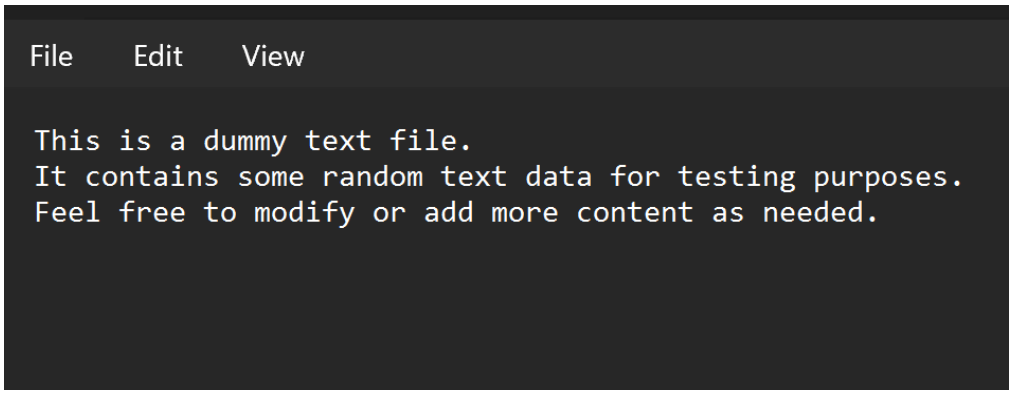
Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.

Step 1: Create a S3 Bucket and Upload some sample csv files in it.

- Sign in to the AWS Management Console.
- Navigate to Amazon S3
- Create a bucket (if you don't have one already)
- In the S3 Bucket, upload four files:
  - One .txt file - dummy.txt – contains random text
  - Three .csv files - Sample1.csv (Contains Employee IDs).

    Sample2.csv (Contains PAN Card numbers).

    Sample3.csv (Contains non-sensitive information).
- Click on Upload.



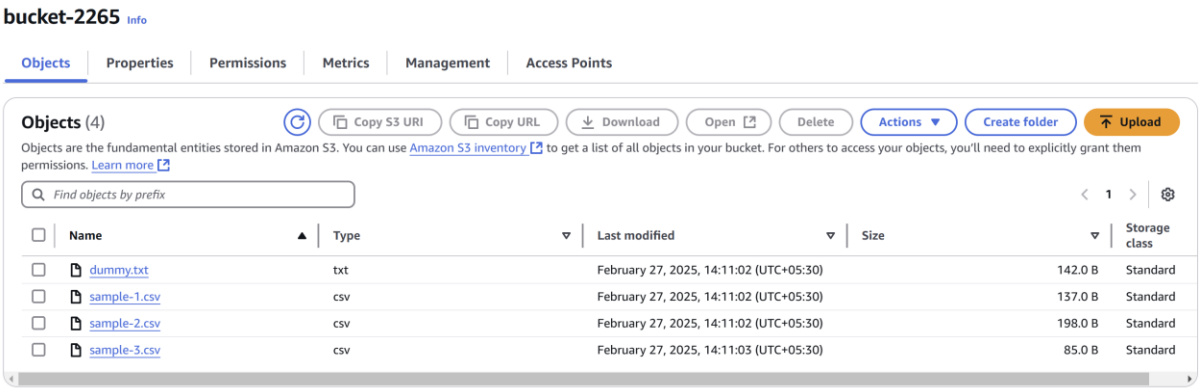
EmpId	Name	State
MH-01	Amit Sharma	Maharashtra
MH-02	Priya Singh	Maharashtra
MH-03	Rohit Verma	Maharashtra
MH-04	Sneha Joshi	Maharashtra

EmpId	Name	State	Pancard
MH-01-JP	Amit Sharma	Maharashtra	ABCDE1234F
MH-02-KR	Priya Singh	Maharashtra	WXYZA5678G
MH-03-LM	Rohit Verma	Maharashtra	LMNOS9012H
MH-04-NP	Sneha Joshi	Maharashtra	PQRSW3456J

Name of the Student: Sahil S. Mandawgade **PRN:** 20220802265

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.

City	State
Mumbai	Maharashtra
Pune	Maharashtra
Nagpur	Maharashtra
Nashik	Maharashtra



Step 2: Create Two Macie Jobs.

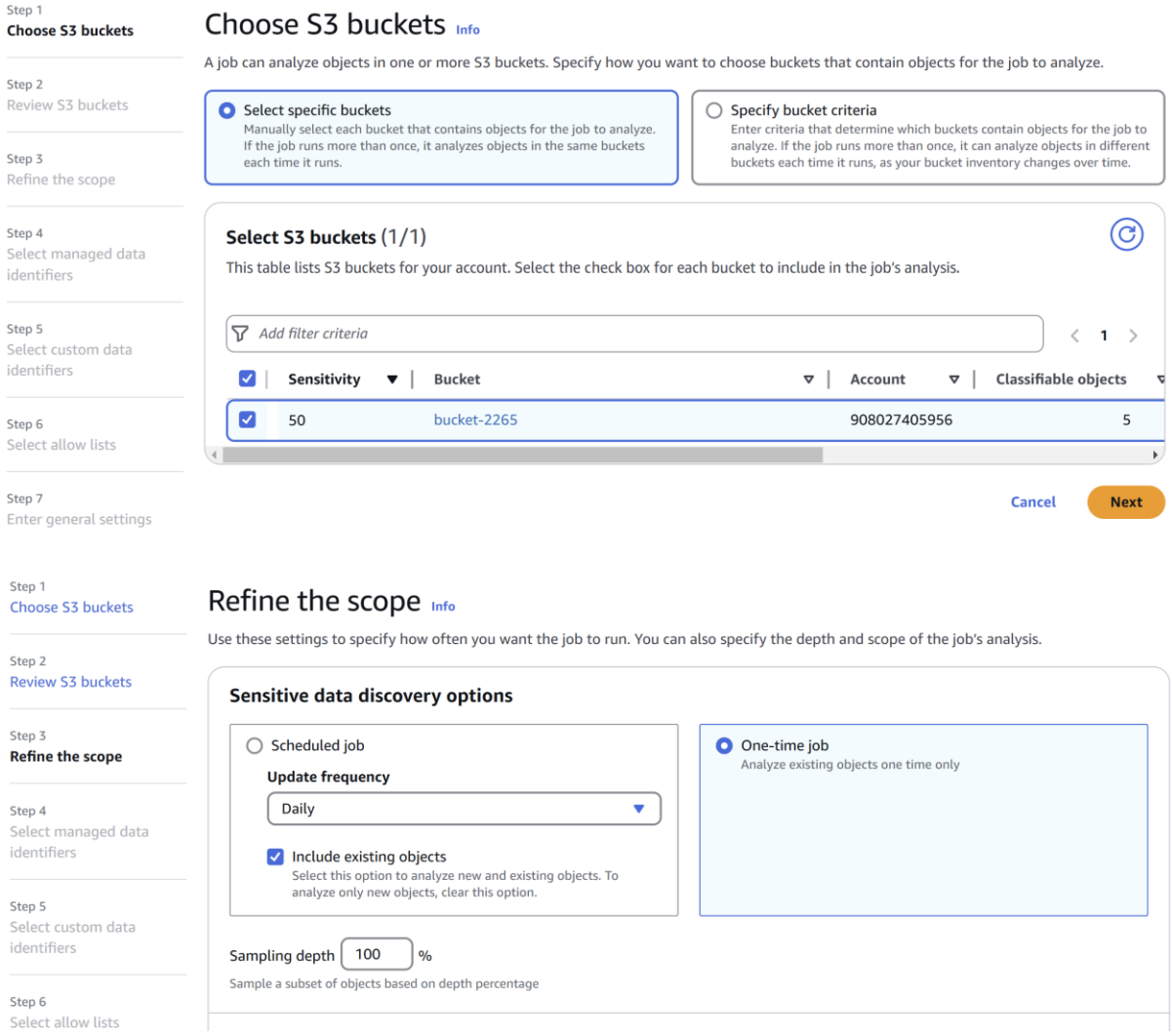
- Open Amazon Macie in the AWS Console.
- Enable Macie.
- Go to “Jobs” and click "Create Job".
- In Step 1, Choose your S3 Bucket - (bucket-2265) and click on next.
- In Step 2, review and click next.
- In Step 3, Select ‘One Time Job’.
- Under ‘Additional settings’, in ‘Object criteria’, type ‘csv’ in the box and click on ‘Include’ and the click on next.
- In Step 4, keep everything as is (recommended) and click on next.

D Y PATIL INTERNATIONAL UNIVERSITY
AKURDI PUNE

School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

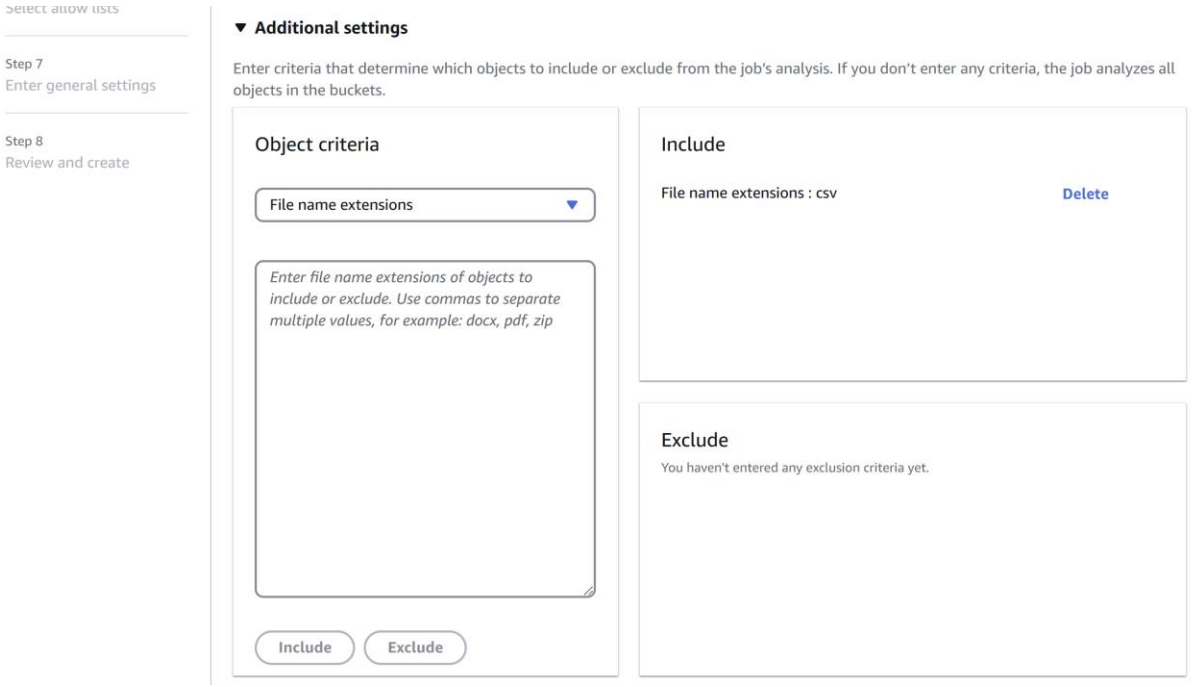
Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.



Name of the Student: Sahil S. Mandawgade **PRN:** 20220802265

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.

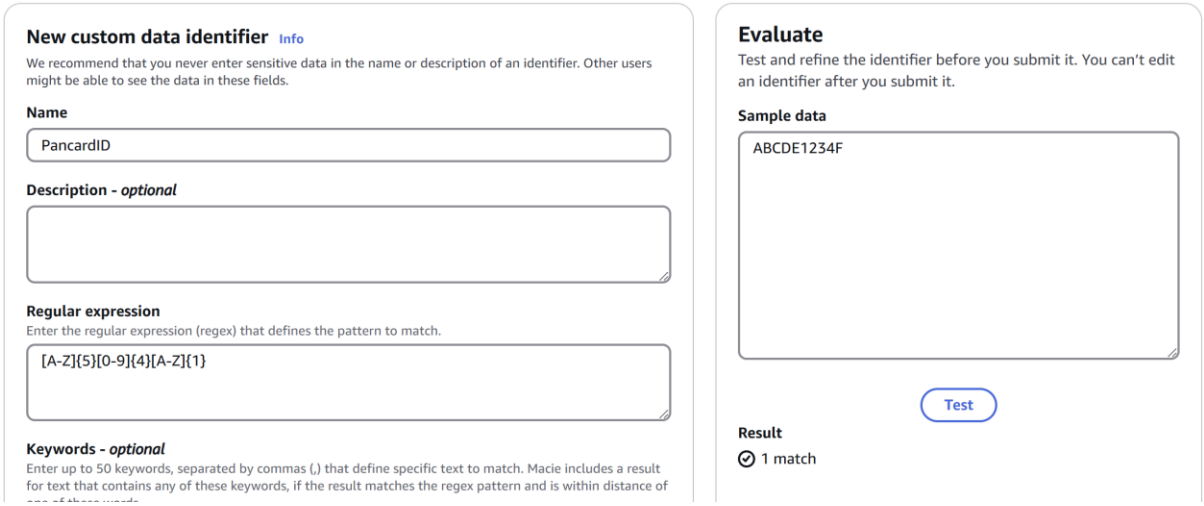


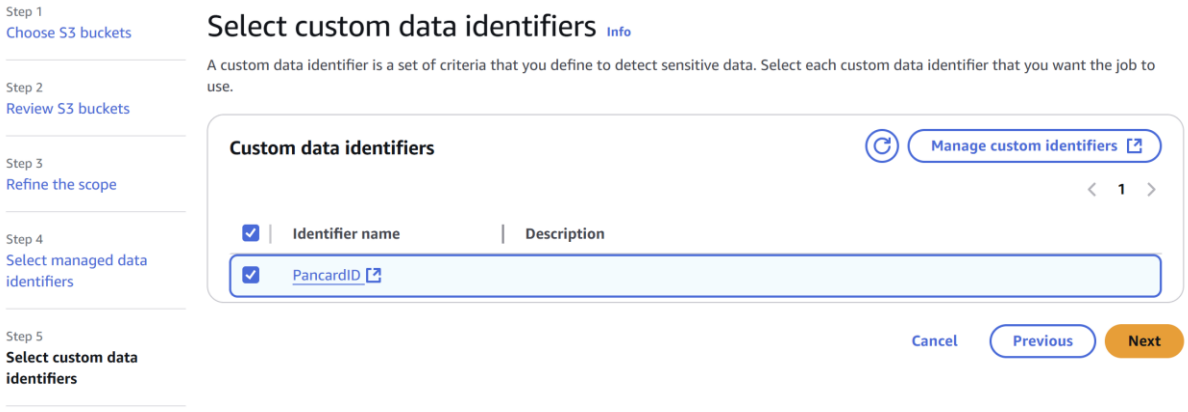
- In Step 5, click on 'Manage custom identifiers' to create a custom identifier.
  - Set name – 'PancardID'.
  - Add Regular Expression – '**[A-Z]{5}[0-9]{4}[A-Z]{1}**'
  - Keep Severity settings as 'Use Medium severity'.
  - Click on submit.
  - Note: You can check whether the regex is correct or not by testing sample data from csv files in 'Evaluate' box.
- Go back to Step 5 and refresh to find newly created identifier.
- Click on next.

School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.



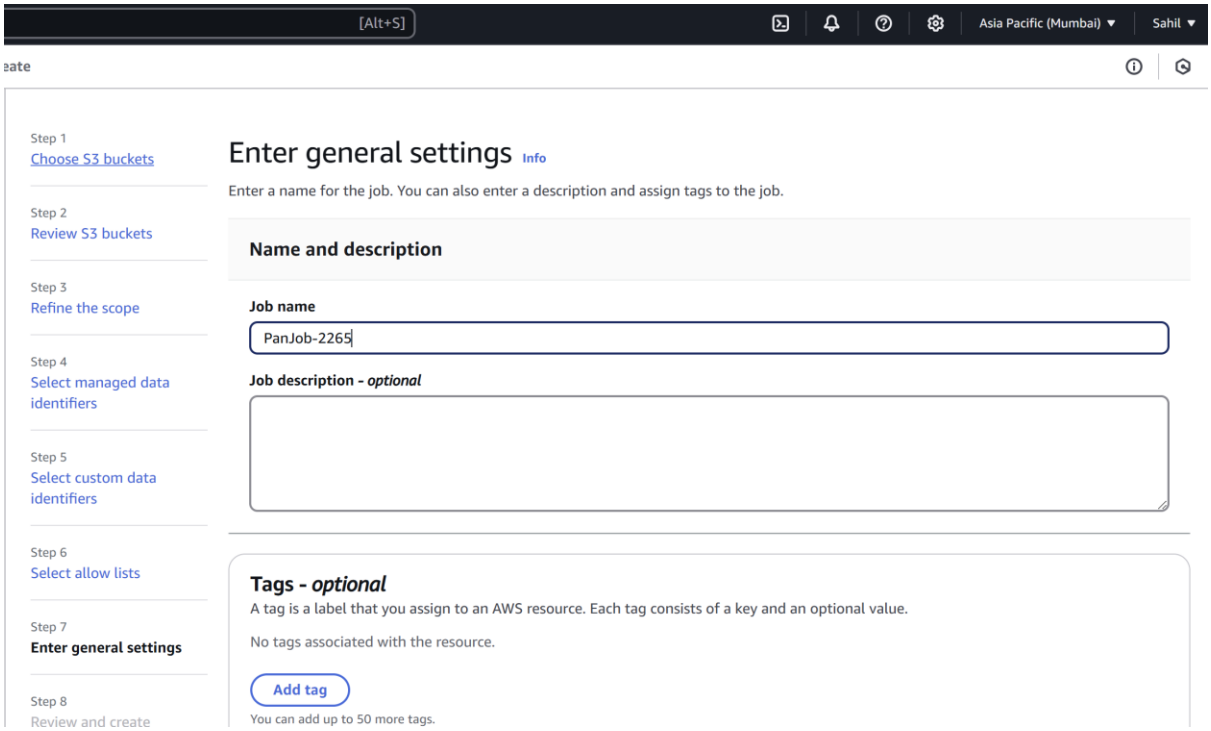


- In Step 6, click on next.
- In Step 7, set job name – 'PanJob-2265' and click on next.
- In Step 8, click on submit.
- The job will take up to a few minutes to complete.

School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.

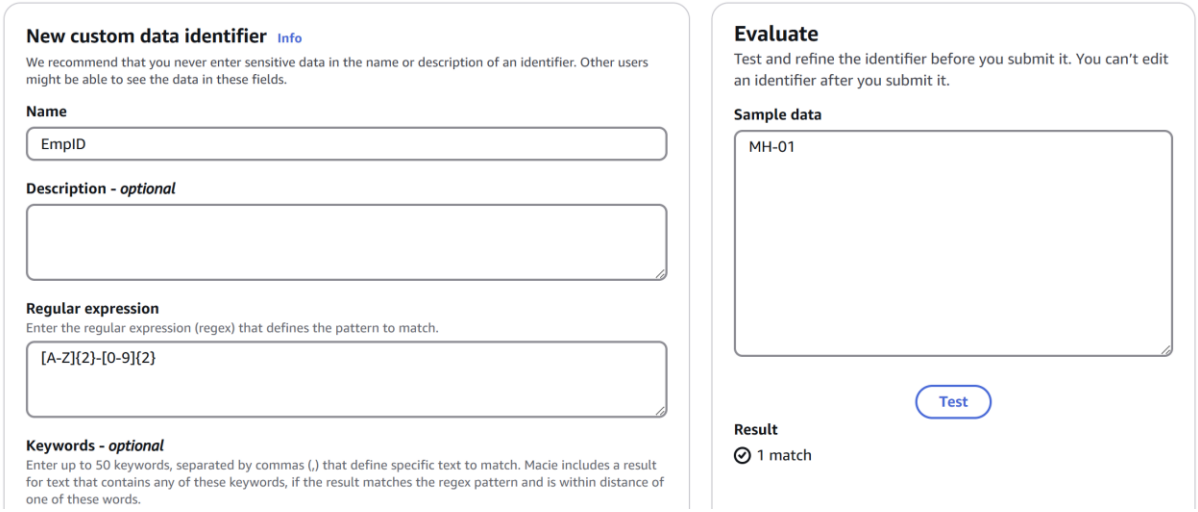


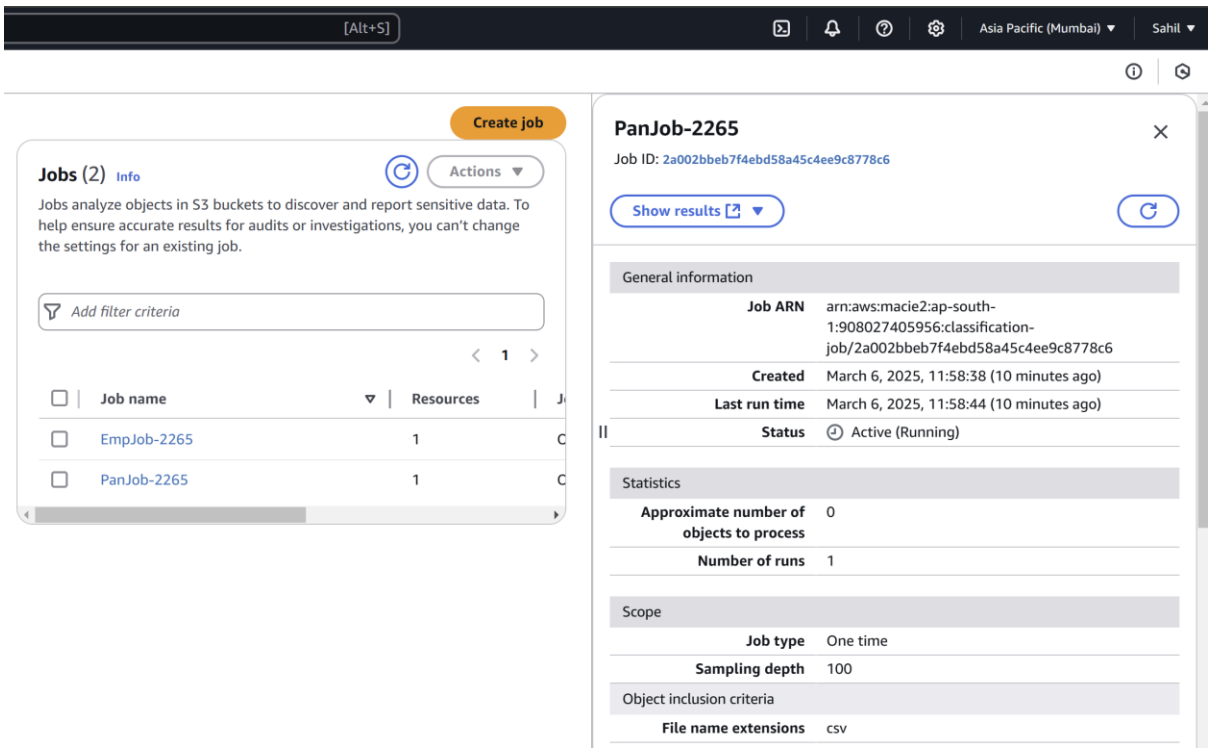
- Similarly, create another Job.
- Choose the same S3 Bucket.
- Select 'One-time job' under 'Sensitive data discovery options'.
- Create another data identifier named – 'EmpID'.
- Add Regular Expression – '**[A-Z]{2}-[0-9]{2}**'.
- Keep Severity as – Medium severity.
- Test sample data from csv file.
- Go back to creating job to select the Custom data identifier.
- Name the 2nd job as – 'EmpJob-2265'.

School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.





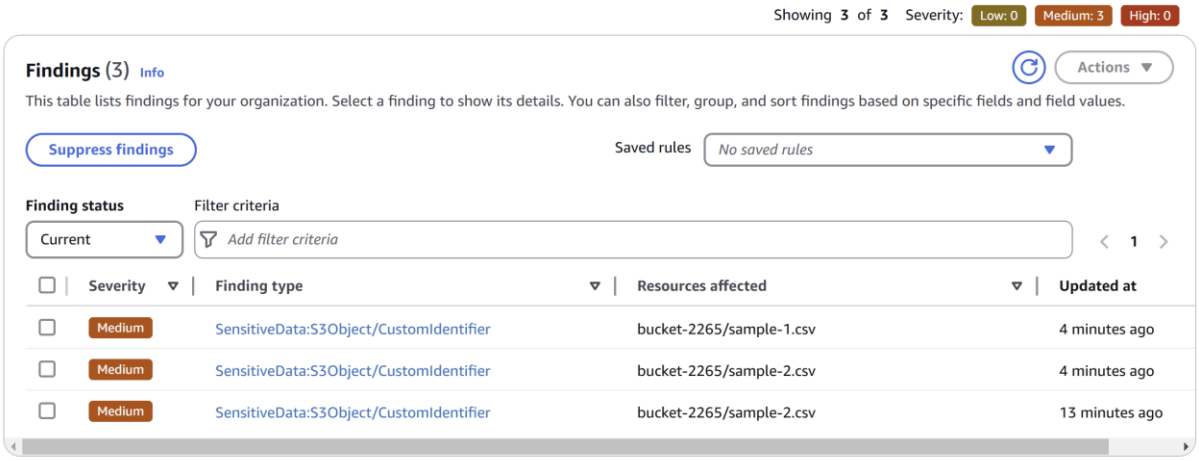
- As we can see, 2 Jobs have been created.

Name of the Student: Sahil S. Mandawgade **PRN:** 20220802265

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.

Step 3: View Findings and Extract JSON Output.

- Click on the Job ID to open details.
- Click on "Show Findings".
- Select a finding (related to PAN or Employee ID).
- View and copy the JSON output.



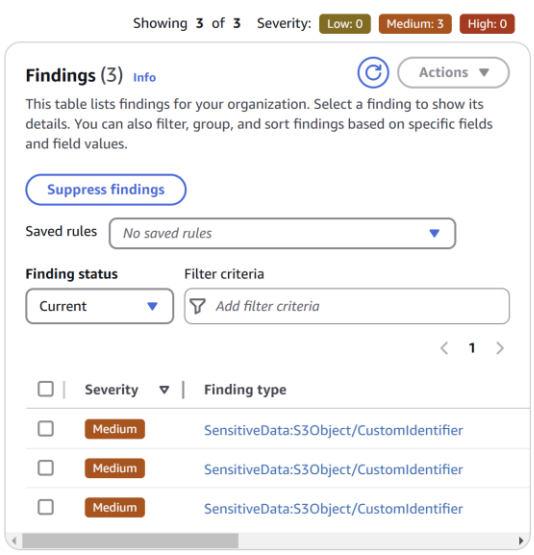
- Wee can see Sensitive Data count as 4.

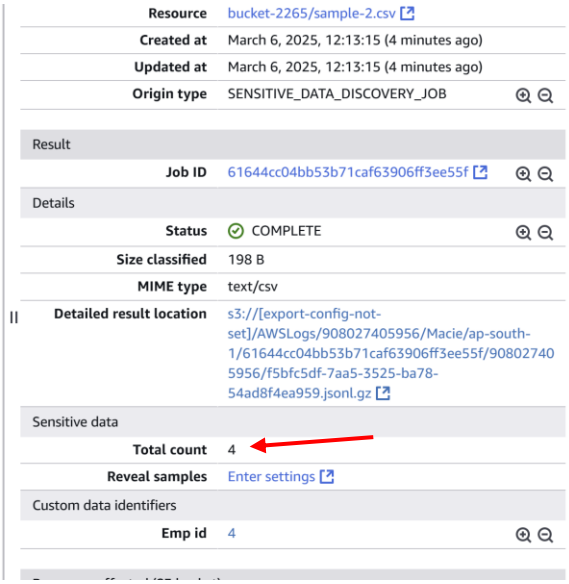
D Y PATIL INTERNATIONAL UNIVERSITY
AKURDI PUNE

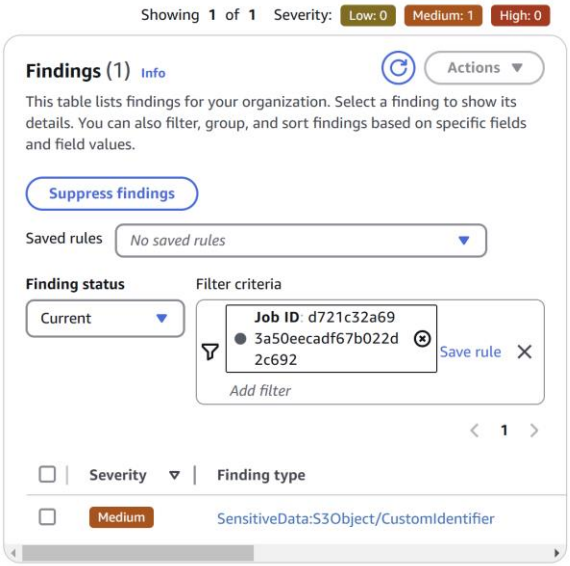
School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.







School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN:** 20220802265

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.

School of Computer Science, Engineering and Applications (SCSEA)
B. Tech TY (CCSA)
Subject: Cloud Architecture And Protocol

Name of the Student: Sahil S. Mandawgade **PRN: 20220802265**

Title of Practical: 7. AWS Macie: Recognizing and Assessing Sensitive Data with PAN Card Details in Stored Documents.