# Weekly Activity#2

Progress Report: AWS EC2 and S3 Implementation

Date: April 29, 2025

To: Project Manager

From: Sahil Sharma

Subject: Status Update on AWS EC2 Virtual Server and S3 Data Storage Implementation

# **Project Overview:**

As a part of the initiative emigrate academic statistics processing and garage to the cloud, this undertaking makes a speciality of deploying a digital server the use of AWS EC2 and setting up a dependent records repository in AWS S3. The goal is to make sure scalable, steady, and green managing of datasets while adhering to quality practices in cloud architecture.

#### **Key Accomplishments:**

- EC2 Instance Deployment
- A virtual server (Instance ID: i-0588c70d695df587a) was successfully launched in the US East (N. Virginia) region.

#### **Specifications:**

- Instance Type: t2.micro (cost-effective for initial testing).
- Network Configuration: Assigned a public IPv4 address (3.238.108.10) and private IP (172.31.8.8) for secure internal communication.
- VPC Setup: A dedicated Virtual Private Cloud (Academic-Vol-SahilSharma, VPC ID: vpc-0606z3b2cf24aef16) was configured with a CIDR block of 172.31.0.0/16, ensuring network isolation.

#### **Security and Compliance:**

• Security Group: Academic-SG-SahilSharma (ID: sg-0bf72259a5c832807) was created to manage inbound/outbound traffic. This restricts unauthorized access while allowing necessary protocols for data transfer.

• IAM Access Analyzer: Enabled for S3 to monitor resource permissions and ensure compliance with least-privilege principles.

## S3 Bucket Configuration:

- A general-purpose bucket named academic-raw-mah was created in us-east-1 to store raw datasets.
- Storage Monitoring: AWS Storage Lens was activated to track storage usage trends and optimize costs.

#### **Project Documentation & Collaboration:**

- A structured directory (Group 8 Academic > Week #2 > Academic Data Team > Design) was established on AWS for team collaboration.
- The Academic Data Lake Design document was uploaded and updated by Saha Sharma, outlining the architecture for data ingestion and processing.
- Supplementary diagrams (fishbone and Draw.io) were created to visualize workflows and dependencies, ensuring alignment across stakeholders.

# **Challenges & Mitigation:**

- Initial VPC Configuration Complexity: The custom VPC setup required careful subnet and routing table configurations. This was resolved by leveraging AWS's default DHCP and route table settings for simplicity.
- Security Group Rules: Initially, overly restrictive rules blocked SSH access. Adjustments were made to permit SSH (port 22) from trusted IP ranges.

## **Next Steps:**

- Data Migration: Begin transferring academic datasets to the academic-raw-mah bucket using AWS CLI/SDKs.
- Automation: Implement lifecycle policies in S3 to archive older data to Glacier, reducing storage costs.
- EC2 Scaling: Monitor instance performance and upgrade to a larger instance type if CPU/memory usage exceeds thresholds.
- Access Controls: Finalize IAM roles and bucket policies to grant granular access to team members.

#### **Conclusion:**

The foundational AWS infrastructure is now operational, with EC2 and S3 components securely configured. The team is positioned to proceed with data processing tasks and further optimizations. Regular audits via Storage Lens and IAM Analyzer will ensure ongoing compliance and cost efficiency.

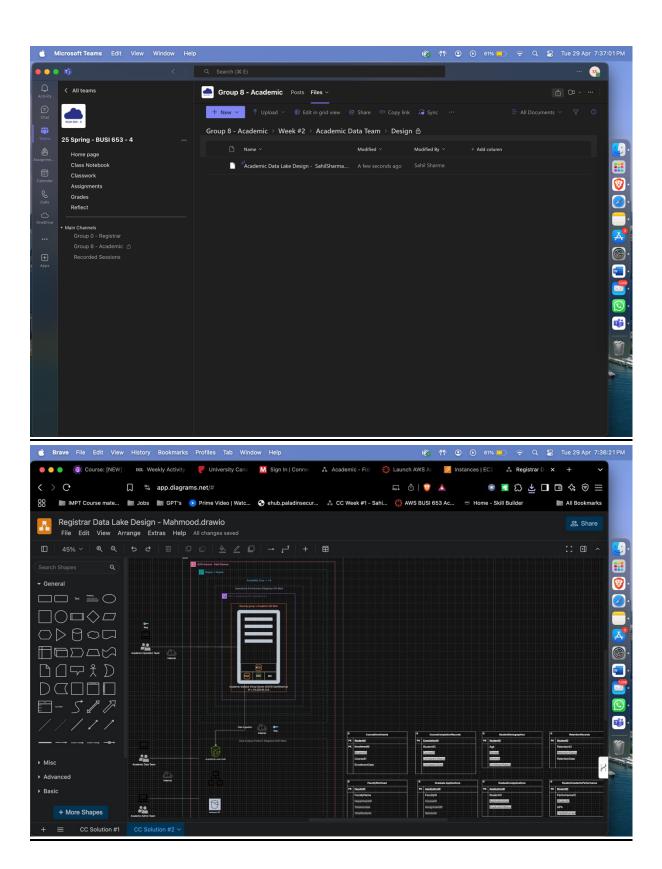
Attachments: Screenshots of EC2, S3, VPC, and Security Group configurations.

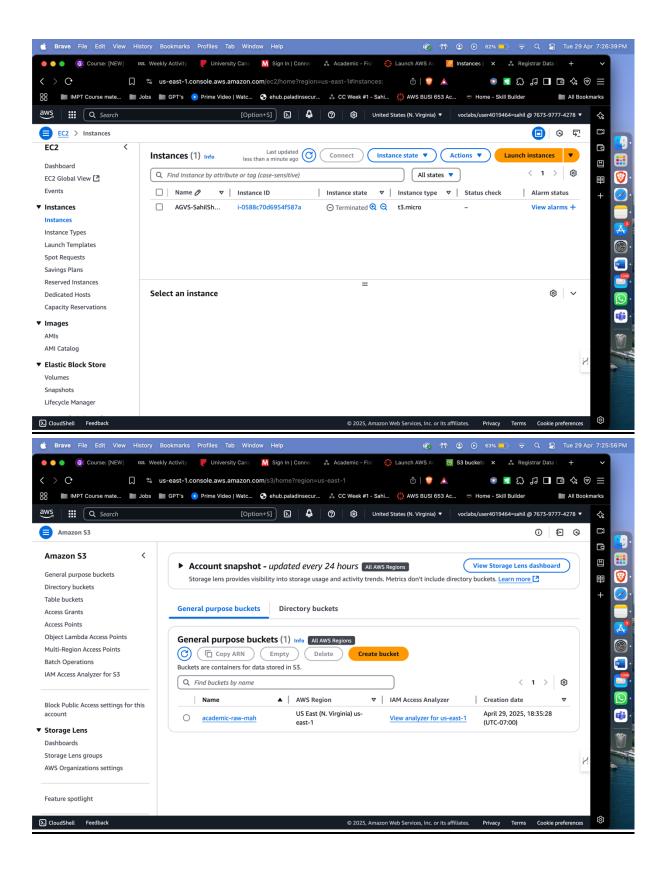
Respectfully submitted,

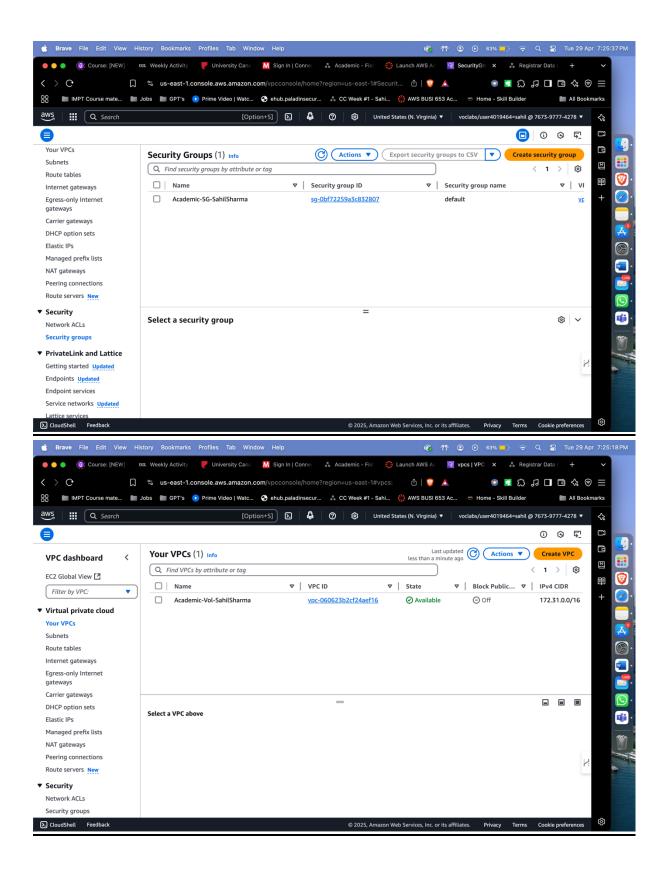
[Your Name]

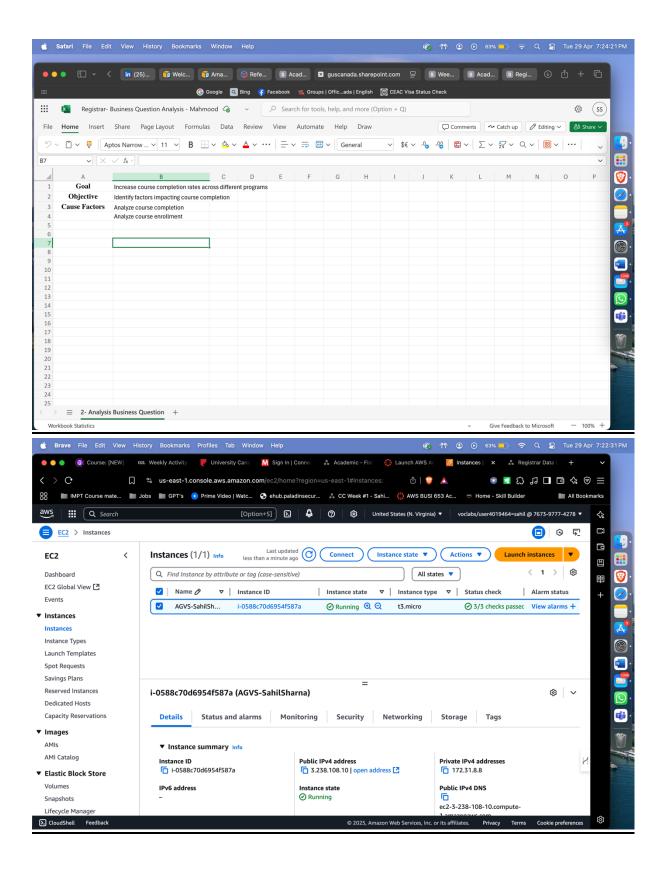
Cloud Infrastructure Team

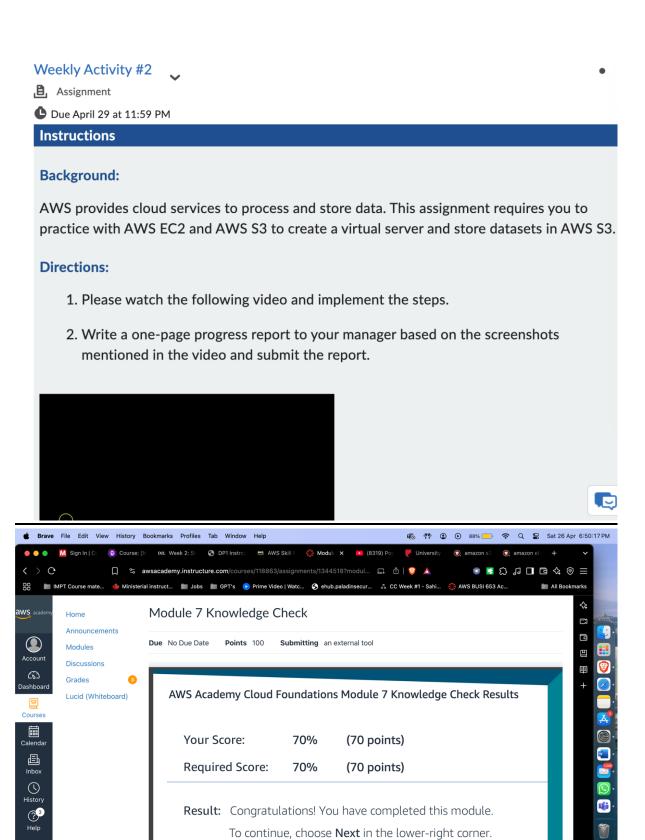
# **Screenshots:**











Next ▶

◆ Previous

