Using the earnings CSV as a base, prepare a new data file with employees' office locations. Make sure there are 5-6 distinct locations that are shared between

employees.

Create a Glue job that aggregates the data based on the office location to calculate average salaries and raise percentages for these locations.

Created Bucket with Name khan –glue

Added three folders with all relevant data
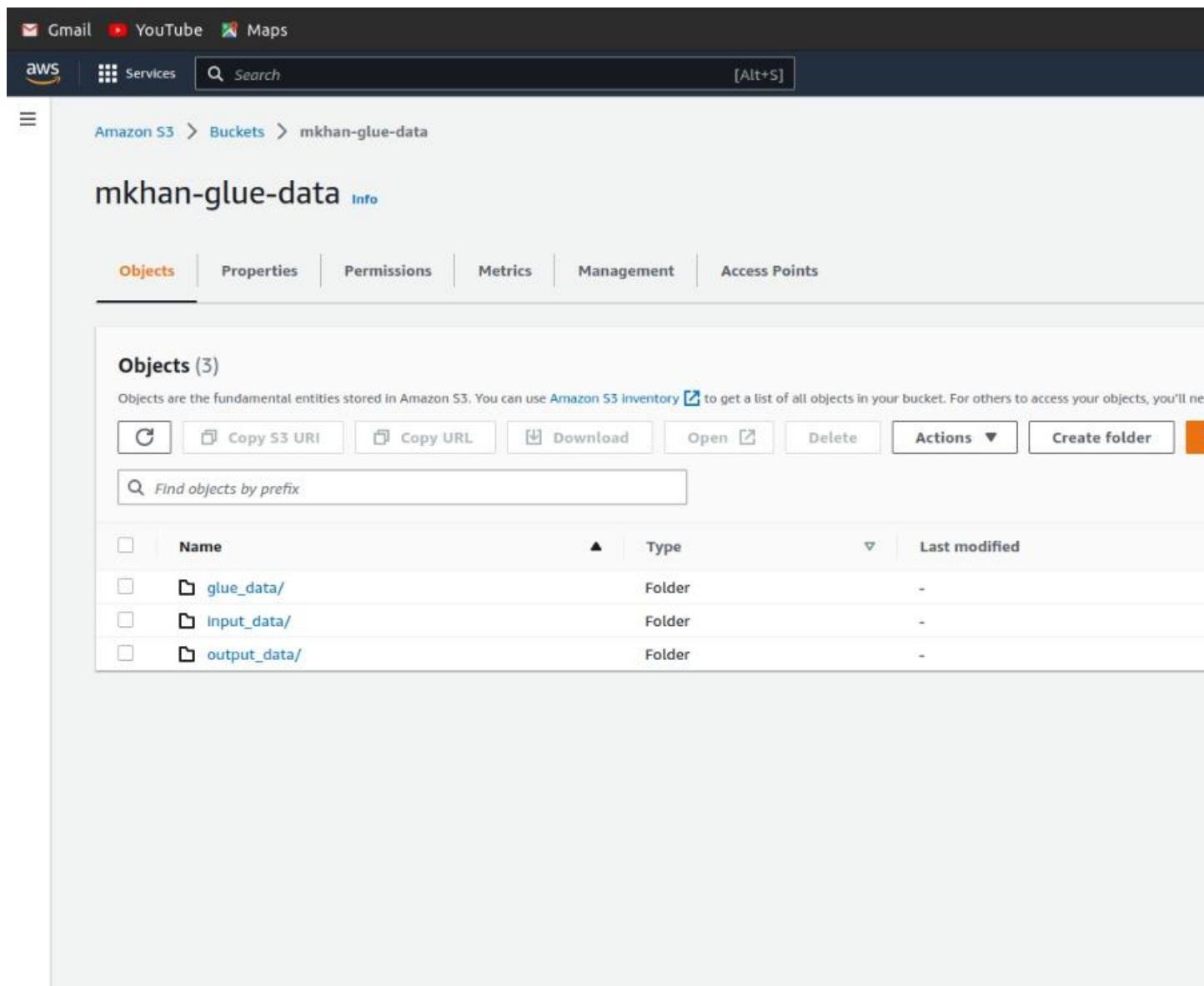
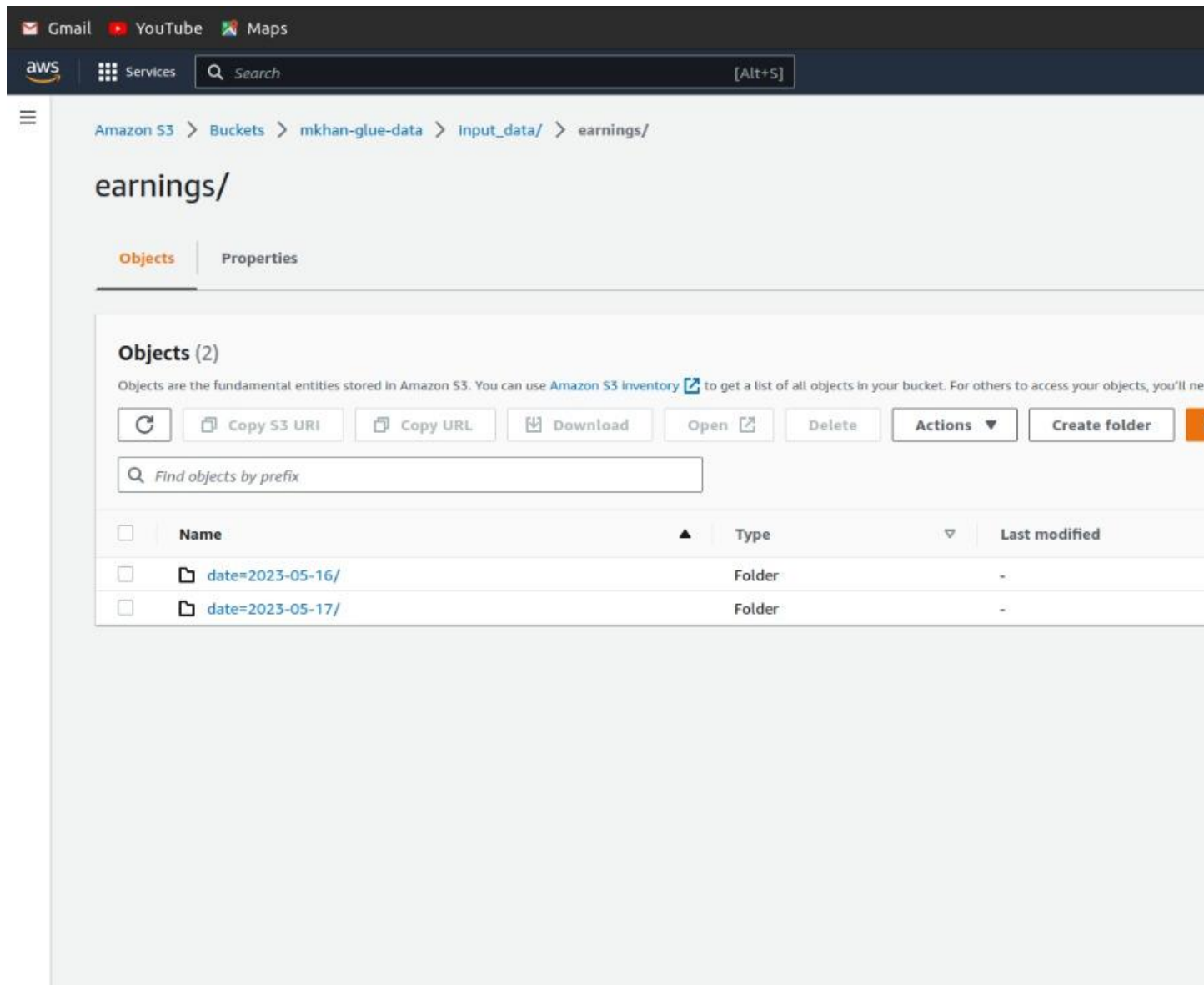**Inside input folder**

Created two other folders namely

Date = 2023-06-16

Date = 2023-06-17

Created DBS database Named "mkhan-emlopyees-db"

DB details
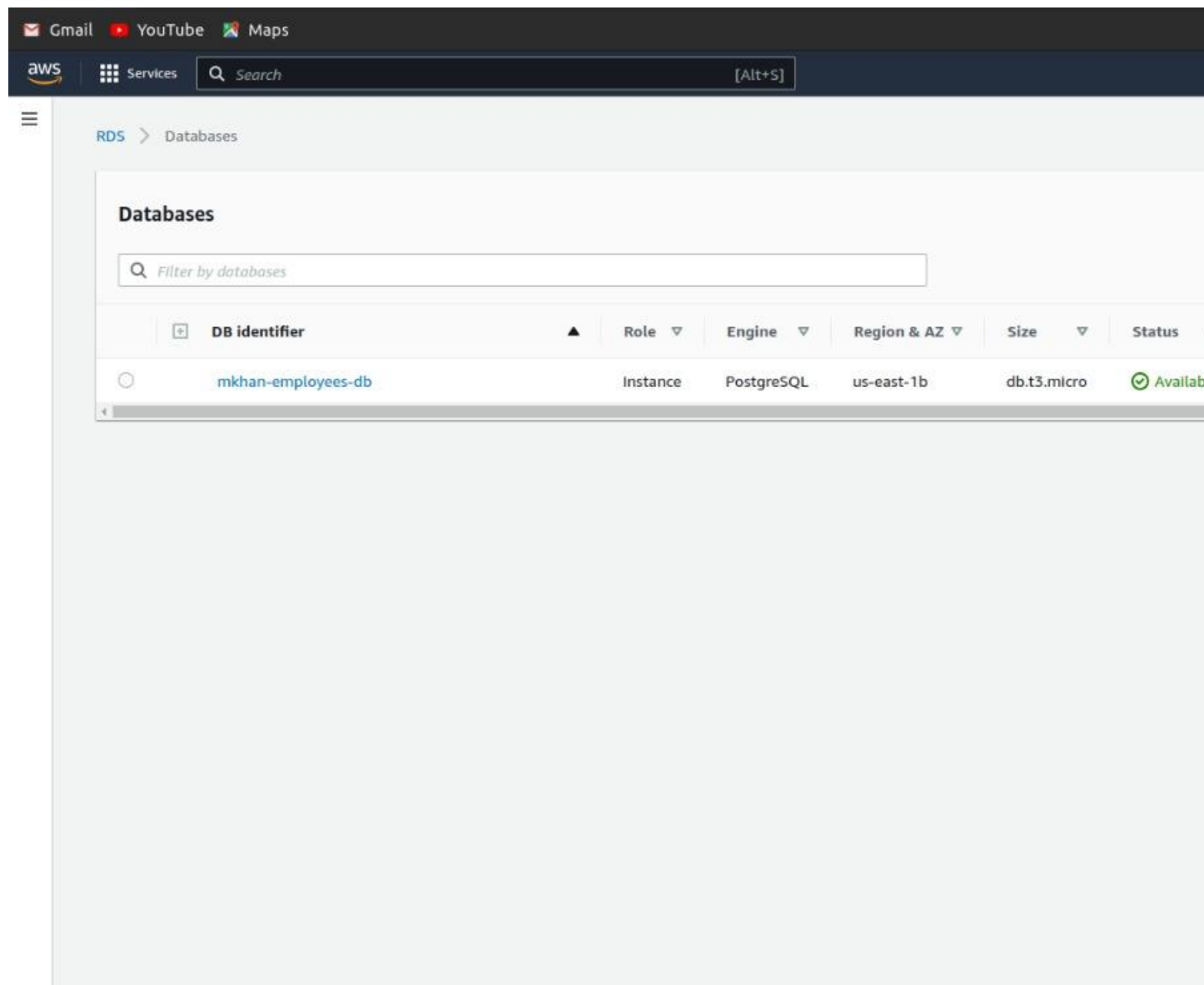
Created IAM role rules for S3 and glue access

aws   Services   Q Search   [Alt+S]

IAM > Roles > mkhan-glue-role

mkhan-glue-role

Allows Glue to call AWS services on your behalf.

Summary

Creation date
May 16, 2023, 14:43 (UTC+05:00)

ARN
arn:aws:iam::5

Last activity
✓ 17 minutes ago

Maximum session
1 hour

| Permissions | Trust relationships | Tags | Access Advisor | Revoke sessions |

Permissions policies (2) Info
You can attach up to 10 managed policies.

Q Filter policies by property or policy name and press enter.

	Policy name ⬈	Type
☐	⊞ 🗊 AmazonS3FullAccess	AWS managed
☐	⊞ 🗊 AWSGlueServiceRole	AWS managed

Permissions boundary - (not set) Info
Set a permissions boundary to control the maximum permissions this role can have. This is not a common setting but can be used to delegate permission management to others.

VPC Created

VPC Configuration

Created crawlers and they started running successfully

Raise percentage

Avg salary

aws   Services   Q Search   [Alt+S]

employe-salaries-etl-job

⚠ **[avg_salaries] input updated**
Your recent changes may have affected the input of your SQL node. Please verify its output is still valid with the updated definition.

Visual   Script   Job details   Runs   Data quality New   Schedules   Version Control

Source ▾   Action ▾   Target ▾   Undo   Redo   Remove   ⊕   ⊖   ⊡

Transform

Data previe

Q Filter samp

location

B

C

A

D

E

Glue studio Running successfully

Final Output

aws Services Q Search [Alt+S]

Amazon S3 ✕

Buckets

Access Points

Object Lambda Access Points

Multi-Region Access Points

Batch Operations

IAM Access Analyzer for S3

Block Public Access settings for this account

▼ Storage Lens

Dashboards

AWS Organizations settings

Feature spotlight 3

▶ AWS Marketplace for S3

Amazon S3 > Buckets > mkhan-glue-data > output_data/

output_data/

Objects   Properties

Objects (5)

Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory ↗ to get a list of all objects in your bucke

| ⟳ | Copy S3 URI | Copy URL | ⤓ Download | Open ↗ | Delete | Ac |

Q Find objects by prefix

☐	Name	▲	Type
☐	🗋 run-1684324733853-part-block-0-r-00002-snappy.parquet		parqu
☐	🗋 run-1684324733853-part-block-0-r-00014-snappy.parquet		parqu
☐	🗋 run-1684324733853-part-block-0-r-00021-snappy.parquet		parqu
☐	🗋 run-1684324733853-part-block-0-r-00025-snappy.parquet		parqu
☐	🗋 run-1684324733853-part-block-0-r-00031-snappy.parquet		parqu