Below are the step by screenshots of the two jobs that I created and the following updates in the tables of the landing dataset which was in the S3 bucket. I will not go through the detailed steps as it is already present in the document containing detailed lab steps, but just give an overview.

**Step 1:** Creating database “gluestudio”

Graphical user interface, text, application, Word

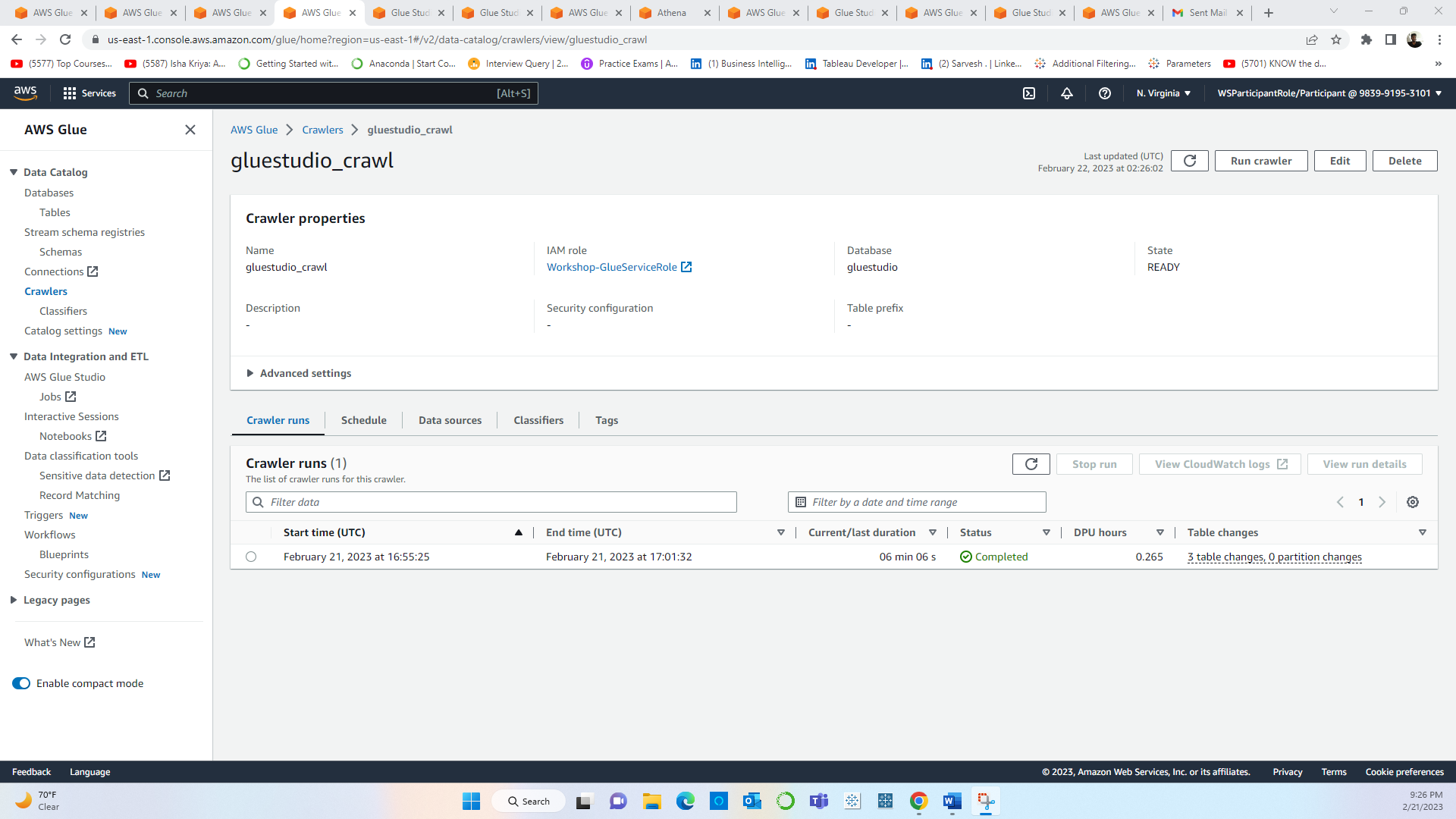
Description automatically generated

**Step 2:** Creating the crawler “gluestudio\_crawl”

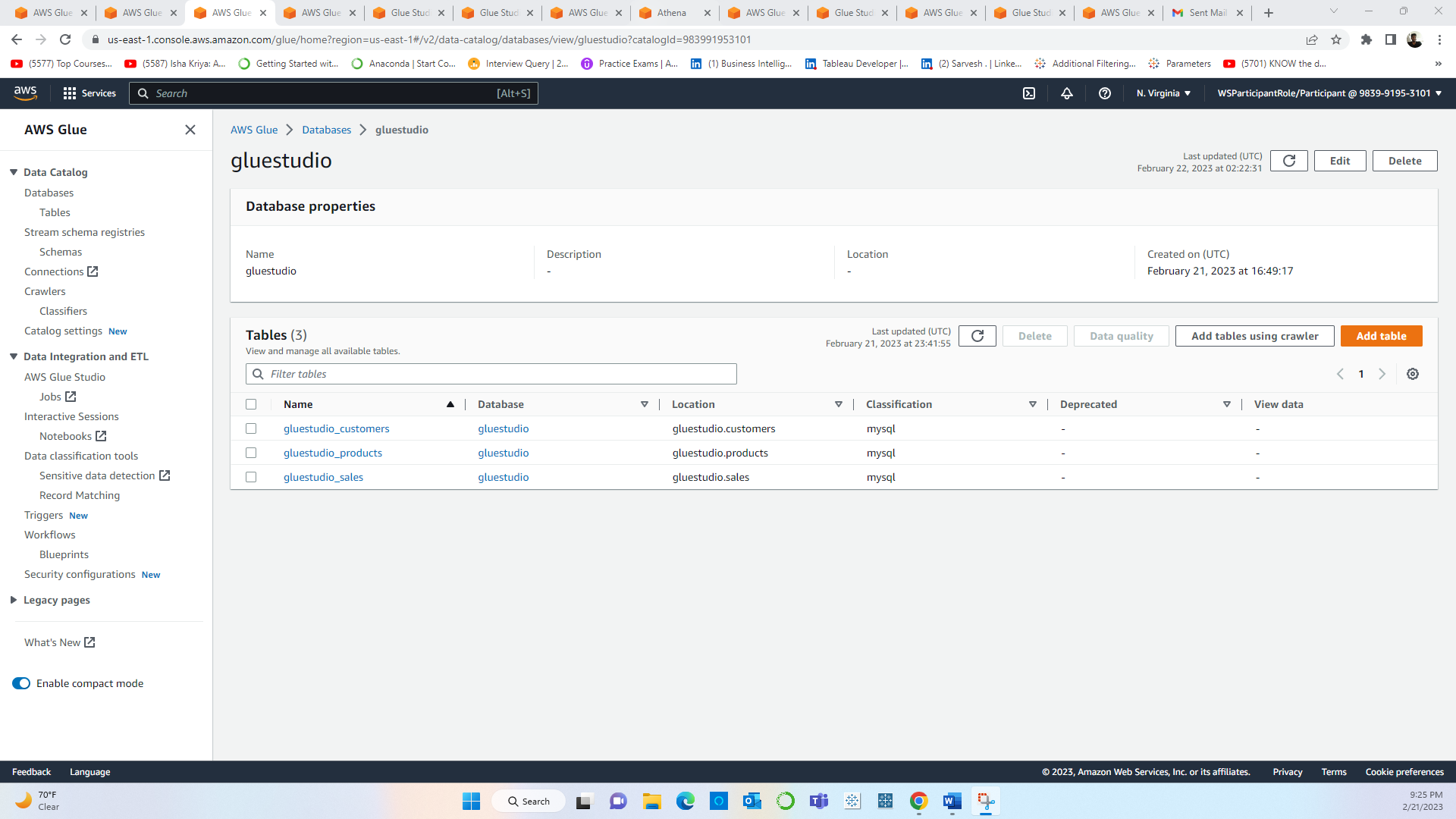
Graphical user interface, text, application, Word

Description automatically generated

**Step 3:** Crawler run successfully completed.



**Step 4:** After the crawler is executed successfully all the tables can now be seen in the technical Data Catalog of Glue



**Step 5:** Job 1 - Extract from a MySql DB and **transform** the data visually performing multiple actions of changing data types, selecting specific columns, joining tables, filtering data, and performing aggregations.

Graphical user interface

Description automatically generated

Graphical user interface, text, application, table

Description automatically generated

**Step 6**: Job 1 successfully Run.

Graphical user interface, application

Description automatically generated

**Step 7**: “sales-aggregation-report table” successfully created.

Graphical user interface, application, table, Word

Description automatically generated

**Step 8:** Querying the table “sales-aggregation-report table” in **Athena** to verify if the data has been correctly populated.

A screenshot of a computer

Description automatically generated

**Step 9:** Creating the custom Regex pattern to mask date of birth.

Graphical user interface, application, Word

Description automatically generated

**Step 10**: Creating the second ETL Job - Redact sensitive data from dataset using Glue built-in PII feature without writing code.

