**Milestone 1: Data Cleaning and Staging**

**Original datasets in S3 bucket i.e. etlemrbucket**

A screenshot of a computer

Description automatically generated

**Cleaned datasets in S3 bucket under folder named cleandata**

A screenshot of a computer

Description automatically generated

**Copying the datasets from “cleandata” folder into respective folders for easy handling**

A screenshot of a computer

Description automatically generated

**Milestone 2: Glue Table Creation**

**Creating crawler i.e. covid19-crawler**

**A screenshot of a computer

Description automatically generated**

**Tables creation after running crawler**

**A screenshot of a computer

Description automatically generated**

**Showing schema of some datasets**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**Milestone 3: Athena Preview and Validation**

**1. c-enigma-jhu.csv**

**Previewing data:**

A screenshot of a computer

Description automatically generated

**Validation:**

**Counting rows**

A screenshot of a computer

Description automatically generated

**Checking null values (column: country\_region)**

A screenshot of a computer

Description automatically generated

**Statistical summary (column: deaths)**

A screenshot of a computer

Description automatically generated

**2. c-states\_daily.csv**

**Previewing data**

A screenshot of a computer

Description automatically generated

**Validation:**

**Counting rows**

A screenshot of a computer

Description automatically generated

**Checking null values (column: negative)**

A screenshot of a computer

Description automatically generated

**Statistical summary (column: positive)**

A screenshot of a computer

Description automatically generated

**3. c-us\_counties.csv**

**Previewing data**

A screenshot of a computer

Description automatically generated

**Counting rows**

A screenshot of a computer

Description automatically generated

**Checking null values (column: county)**

A screenshot of a computer

Description automatically generated

**Statistical summary (column: fips)**

A screenshot of a computer

Description automatically generated

**4. c-us\_daily.csv**

**Previewing data**

A screenshot of a computer

Description automatically generated

**Counting rows**

A screenshot of a computer

Description automatically generated

**Checking null values (column: deathIncrease)**

A screenshot of a computer

Description automatically generated

**Statistical summary (column: deathIncrease)**

A screenshot of a computer

Description automatically generated

**5. c-us\_states.csv**

**Previewing data**

A screenshot of a computer

Description automatically generated

**Validation**

**Counting rows**

A screenshot of a computer

Description automatically generated

**Checking null values (column: cases)**

A screenshot of a computer

Description automatically generated

**Statistical summary (column: cases)**

A screenshot of a computer

Description automatically generated

**6. c-us-total.csv**

**Previewing data**

A screenshot of a computer

Description automatically generated

**Validation**

**Counting rows**

A screenshot of a computer

Description automatically generated

**Checking null values (column: recovered)**

A screenshot of a computer

Description automatically generated

**Statistical summary (column: recovered)**

A screenshot of a computer

Description automatically generated