AWS Glue

1. Setting up IAM permissions.

A screenshot of a computer

Description automatically generated

1. Created database named “yyz-tickets” and a table “tickets count”.

A screenshot of a computer

Description automatically generated

1. Created stack, by doing so it creates the necessary resources to do the job.

A screenshot of a computer

Description automatically generated

1. Adding AWS Glue tables as data sources from S3.

A screenshot of a computer

Description automatically generated

1. Transform: Changed the data type of both tables and removed column from ‘Tickets’ table.

A screenshot of a computer

Description automatically generated

1. Join: Merging two tables together by ‘ticket\_number = parking\_ticket\_number’.

A screenshot of a computer

Description automatically generated

1. Aggregate: Aggregating the data by ‘location’, ‘description’ and ‘count’.

A screenshot of a computer

Description automatically generated

1. Created new node named ‘select\_aggregated\_data’ with the result data from the aggregation.

A screenshot of a computer

Description automatically generated

1. Storing: After the customization of the job, storing the result set back in the S3 bucket by using the ‘collection transform’ and saving it in the destination.

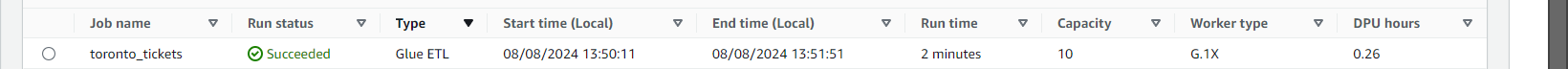
A screenshot of a computer

Description automatically generated

1. Job configuration and running the job.

A screenshot of a computer

Description automatically generated



1. Querying the created table ‘tickets\_count’.

A screenshot of a computer

Description automatically generated

1. Output of the query.

A screenshot of a computer

Description automatically generated