Loading from Data from SQL to Amazon Redshift

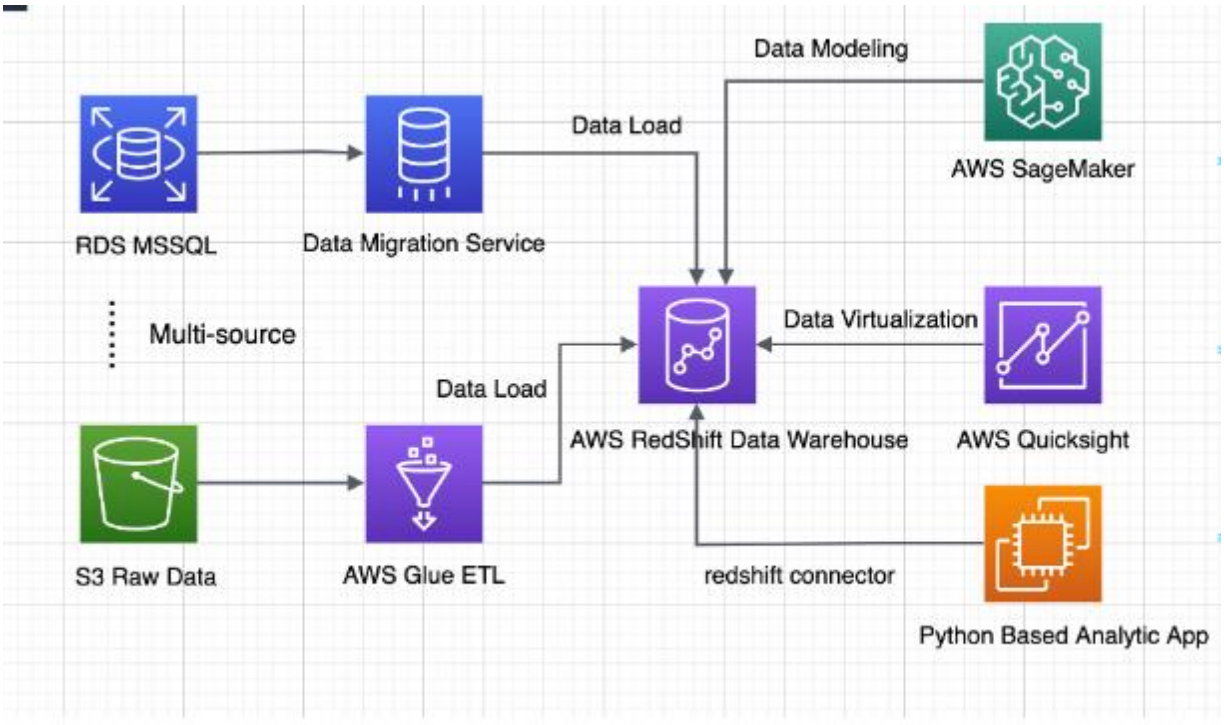
If I want to load data from SQL to Redshift, there are three methods:

1. Using AWS Data Migration Service (DMS)
2. Loading data via an S3 bucket
3. Using AWS Glue

To generate reports , there are three visualization options:

4. Using QuickSight itself
5. Utilizing AWS SageMaker
6. Employing a Python-based analytic app

I have learned these three options as well.



1) Move selected data from multiple sources to the data warehouse
2) Centralized data repository in a tabular format for analytics
3) Data analysis with AWS Redshift cluster and Python-based analytic apps
4) Data BI virtualization with AWS Quicksight
5) Data modeling with AWS Sagemaker

Loading Data from SQL to Redshift by Using DMS

- Set Up Source and Target Endpoints [Create an RDS endpoint for your SQL database | Create an Amazon Redshift endpoint]
- Create a Replication Instance [Launch a DMS replication instance that will perform the data migration]
- Configure Source and Target Endpoints [Specify the SQL database as the source endpoint |Specify the Redshift cluster as the target endpoint]
- Create a Migration Task [Define a migration task that specifies the tables to be migrated | Choose the migration type (full load, CDC, or both)]
- Start the Migration Task [Start the task to begin migrating data from SQL to Redshift]
- Monitor the Migration [Use the DMS console to monitor the progress and ensure the migration completes successfully]

Loading Data from SQL to Redshift by Using S3

- Extract Data from SQL [Export the data from your SQL database to CSV files or another suitable format]
- Upload Data to S3 [Upload the exported data files to an Amazon S3 bucket]
- Set Up Redshift Cluster [Ensure you have an Amazon Redshift cluster set up and accessible]
- Create a Table in Redshift [Define the schema and create a table in Redshift where the data will be loaded]
- Grant Redshift Access to S3 [Create an IAM role with the necessary permissions and associate it with your Redshift cluster]
- Load Data into Redshift [Use the COPY command in Redshift to load data from the S3 bucket into your Redshift table. For example:
  COPY my_table
  FROM 's3://my-bucket/my-data-file.csv'
  IAM_ROLE 'arn:aws:iam::account-id:role/my-redshift-role'
  CSV;]
- Verify Data Load [Check the data in your Redshift table to ensure it has been loaded correctly]

Loading Data from SQL to Redshift by Using AWS Glue

- Set Up AWS Glue [Create an AWS Glue Data Catalog database]
- Create a Connection [In AWS Glue, set up a connection to your SQL database]
- Create a Crawler [Define a crawler that uses the connection to scan the SQL database and populate the Data Catalog with table definitions | Run the crawler to discover the schema and create table metadata in the Data Catalog]
- Create an ETL Job [Define an AWS Glue ETL job to extract data from the SQL database and load it into Amazon Redshift | Choose your SQL database as the source and Redshift as the target | Use the Data Catalog tables for both source and target schemas]
- Configure the Job [Specify the transformation logic if needed, or use the default mappings | Set up the necessary IAM roles for Glue to access both the SQL database and Redshift]
- Run the ETL Job [Execute the Glue job to start the ETL process. This will extract data from SQL, transform it if necessary, and load it into Redshift]
- Verify Data Load [Check the data in your Redshift tables to ensure it has been loaded correctly]

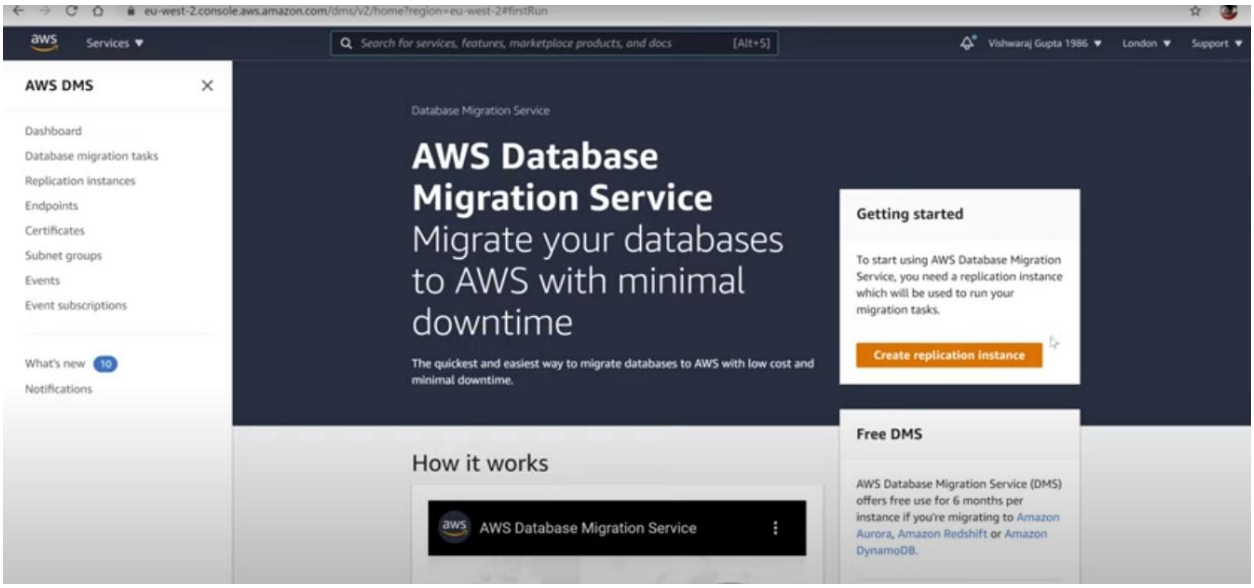
Generate Reports and Visualization in Quicksight from Redshift Database

- Ensure your Amazon Redshift cluster is set up and your data is loaded
- Log in to your Amazon QuickSight account
- From the QuickSight dashboard, click on "Manage data" and then "New dataset".
- Select "Redshift" as your data source
- Enter your Redshift cluster details: [Cluster hostname, Database name, Username and password]
- After connecting, choose the appropriate schema and tables you want to use for your analysis.
- QuickSight provides an option to prepare data. You can join tables, create calculated fields, and apply filters as needed.
- Save and preview your data
- Go back to the QuickSight dashboard and click on "New analysis"
- Select the dataset you just created from Redshift.
- Use the QuickSight interface to create various visualizations such as charts, tables, maps, and more.
- Drag and drop fields onto the canvas to start building your visualizations.
- Customize your visualizations by configuring various settings, applying filters, and adding calculated fields.

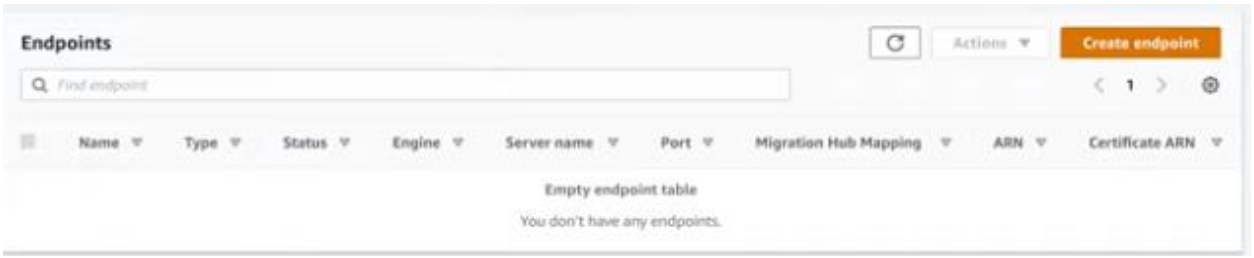
- You can also format and style your visualizations to match your reporting needs.
- Once you have multiple visualizations, you can combine them into a dashboard.
- Click on "Add to dashboard" to create a new dashboard or add to an existing one.
- After creating your dashboard, you can share it with others.

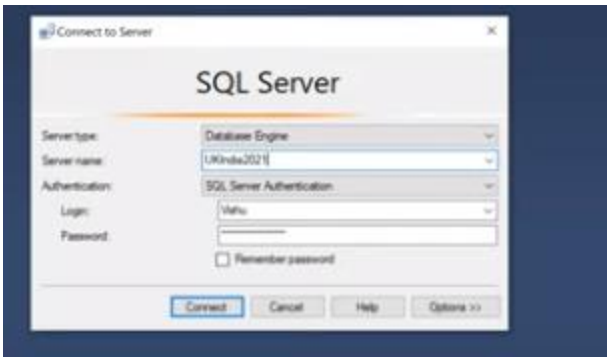
Loading MYSQL Database to Redshift by Using DMS

Once you go to DMS in AWS, click on Data Migration Services and then create a replication instance. It will look like this and ask for several details.

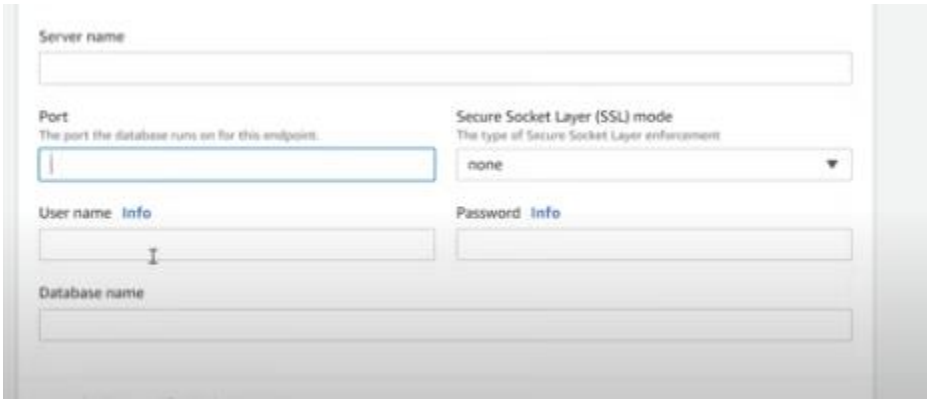


After creating the replication instance, go to the Endpoints section. Endpoints consist of source endpoints and target endpoints.

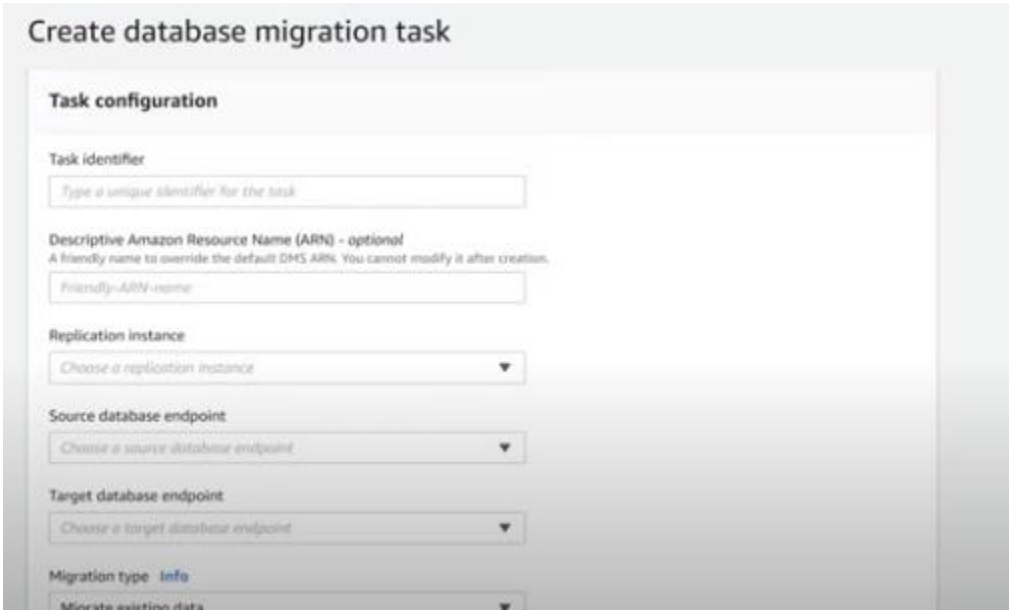


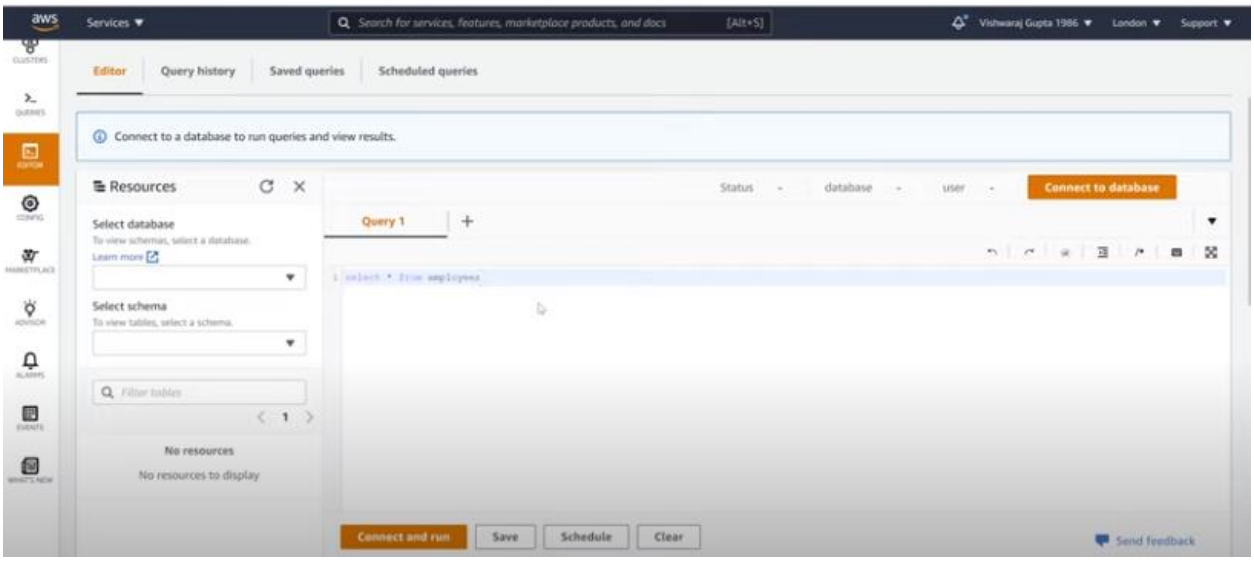

Connect to Server
SQL Server
Server type:
Database Engine
Server name:
UKindia2021
Authentication:
SQL Server Authentication
Login:
Vishu
Password:
Remember password
Connect
Cancel
Help
Options >>


SELECT
CONNECTIONPROPERTY('net_transport') AS net_transport,
CONNECTIONPROPERTY('protocol_type') AS protocol_type,
CONNECTIONPROPERTY('auth_scheme') AS auth_scheme,
CONNECTIONPROPERTY('local_net_address') AS local_net_address,
CONNECTIONPROPERTY('local_tcp_port') AS local_tcp_port,
CONNECTIONPROPERTY('client_net_address') AS client_net_address

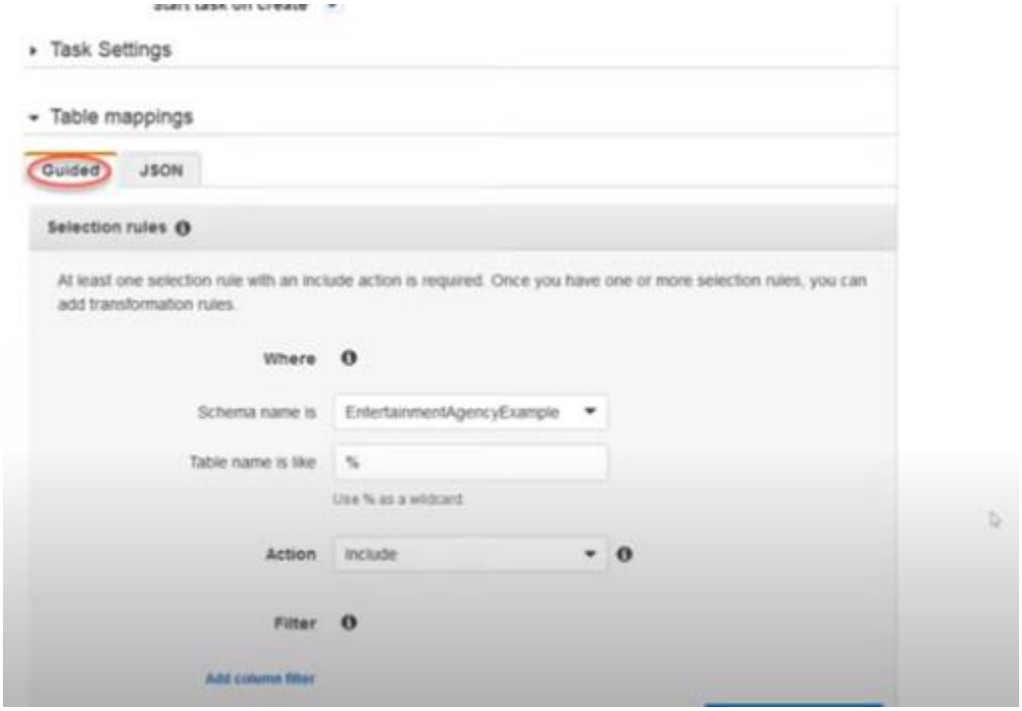

Server name
Port
The port the database runs on for this endpoint.
Secure Socket Layer (SSL) mode
The type of Secure Socket Layer enforcement
none
User name Info
Password Info
Database name

Now, go to the Data Migration Tasks section





Finally, table mapping is a crucial step in the data migration task.



In QuickSight, click on "Analysis" and set up the datasets. And start visualize