

AWS Big Data: Module 4 Assignment

Problem Statement:

You work as a Big Data Engineer at GrapeVine Pvt. Ltd. Your company is currently working on data analysis. You have been assigned certain tasks for the fulfillment of this activity. Your organization uses AWS services for its Big Data needs.

Tasks:

- Create a Redshift cluster with two nodes (dc2.large)
 - Name the database as 'maindata'
 - Create a table 'sales' according to the data in the 'videogamesales.csv' file
 - Upload the CSV file to S3 and use the COPY command to upload the file's data into the 'sales' table
 - Query all the games released in the year 2007 and make a CSV file named 'games2007' out of the result
- Create a manual backup of the Redshift cluster and delete the cluster
- Restore the cluster using the created backup