Module 6: HIVE

Assignment

Version 2.1

Module 6: HIVE

Assignment: Calculating a Stock's Covariance

Table of Contents

Introduction

This assignment will test your skills on HiveQL commands and scripting concepts learnt in the Live Class. You need to use the HiveQL commands to analyse the stocks data from a **'New York Stock Exchange'** dataset and calculate the covariance for a stock.

Dataset:
This dataset is s a comma separated file (CSV) named **'NYSE_daily_prices_Q.csv'** that contains the stock information such as daily quotes etc. at **New York Stock Exchange**.

Covariance: This finance term represents the degree or amount that two stocks or financial instruments move together or apart from each other. With covariance, investors have the opportunity to seek out different investment options based upon their respective risk profile. It is a statistical measure of how one investment moves in relation to another.

- A positive covariance means that asset returns moved together. If investment instruments or stocks tend to be up or down during the same time periods, they have positive covariance.
- A negative covariance means returns move inversely. If one investment tends to be up while the other is down, they have negative covariance.

1. Problem statement

Use HiveQL to analyse the stock exchange dataset and calculate the covariance between the stocks for each month. This will help a stock broker in recommending the stocks to his customers.
After completing the task, you can review the solution in your LMS.