BIG DATA DEVELOPMENT

ACADGILD

Session 06: Hive Introduction

Assignment 2

Table of Contents

Big Data and Hadoop Development

1. Introduction

In this assignment you need to perform the tasks given.

2. Objective

This assignment will help you to consolidate the concepts learnt in the session.

3. Prerequisites
None

4. Associated Data Files

https://drive.google.com/file/d/0Bxr27gVaXO5sa0JBamZXdkpYUFk/view?usp=sharing

5. Problem Statement

- Fetch date and temperature from temperature_data where zip code is greater than 300000 and less than 399999.

- Calculate maximum temperature corresponding to every year from temperature_data table.

- Calculate maximum temperature from temperature_data table corresponding to those years which have at least 2 entries in the table.

- Create a view on the top of last query, name it temperature_data_vw.

- Export contents from temperature_data_vw to a file in local file system, such that each file is '|' delimited.

6. Expected Output

N/A

7. Approximate Time to Complete Task