

WQD7007 Big Data Management

Lab Test (HBase)

Instructions: Answer the following question by using command line interface or Ambari interface. Explain adequately how you get the answers that can include codes used in command line interfaces, codes used in Ambari, a number of print screens and related files. You can create a readme file (either in .txt or .docx) to explain your answer. Zip all the files and submit to the spectrum “Lab test” submission page.

Part 1: (5 marks)

1. Create a file named: “lab_(YourInitial).txt”. E.g. lab_hwl.csv
2. Insert the following content into the file created.

This is a dummy text file for WQD7007.
This file contains 4 lines.
Please remember to save this file.

3. Upload the file to HDFS in directory “/user/hdfs/lab_test/”.

Part 2: (10 marks)

1. Create a HBase table using the following data. Insert the following data to the table created. Based on your judgment, group the attributes of a car to at least two different column family. Display the table after all the data is inserted.

Chevrolet Chevelle Malibu;18.0;8;307.0;130.0;3504.;12.0;70;US
Buick Skylark 320;15.0;8;350.0;165.0;3693.;11.5;70;US
Plymouth Satellite;18.0;8;318.0;150.0;3436.;11.0;70;US
AMC Rebel SST;16.0;8;304.0;150.0;3433.;12.0;70;US
Ford Torino;17.0;8;302.0;140.0;3449.;10.5;70;UK

Attributes used:

Car (String); Miles Per Gallon, MPG (double), Cylinders (int), Displacement (double);
Horsepower (double); Weight (double); Acceleration (double) ;Model (int); Origin (String).

2. Update the record that have horsepower = 150 to 120.
3. Delete the record that have weight = 3433.
4. Display the record that origin from US.