Consider the following dataset:.

https://drive.google.com/file/d/10gQdcl-eO72eYGZMKRQ383h_-XkcdGFX/view?usp=share_link

Complete the steps given below:

    1. Add this dataset on HDFS

   => Make a new directory on HDFS

   => Send this data from local file system to HDFS

   2. Load the data in Pig and filter for Device Type SmartWatch.

   => Start Pig

   => Load the data into pig

   => Filter command to get the device type smartwatch

   3. Store the result in HDFS

   => Store the result that you get after filtering

   4. Load the output of Pig in Hive. Display the data sorted by Rating.

   => Start the Hive shell

   => Create a database

   => Select the database

   => Create a table with the correct schema

   => Load the output of Pig that you got after filtering into that table

   => Write a select query to sort the data by Rating (order by clause)

   5. Store that results in HDFS.

   => Create an external table

Create external table <tablename>(<schema>) row format delimited fields terminated by ','
location '<path in HDFS>';

=> Insert the result of the select query to sort the data by rating into that external table

Insert into table <name of external table> <select query to sort the data by rating>;

Task	Grade

5 out of 5 steps done correctly	A+
4 out of 5 steps done correctly	A
3 out of 5 steps done correctly	B+
2 out of 5 steps done correctly	B
1 out of 5 steps done correctly	C
0 out of 5 steps done correctly	D

After completing, convert the project to PDF format and send it as an attachment to a mail to arshdeep@itvedant.com. Please mention your full name in the email message. Images of outputs are NOT compulsory.