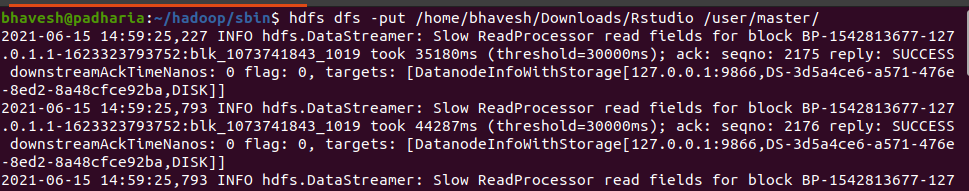
Name : Bhavesh Padharia

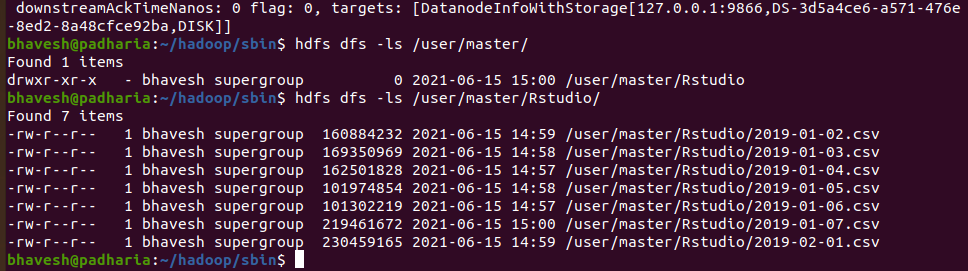
HDFS-Hive Task 1.

1. Import RStudio Log Files from one week in Jun 2021 into HDFS

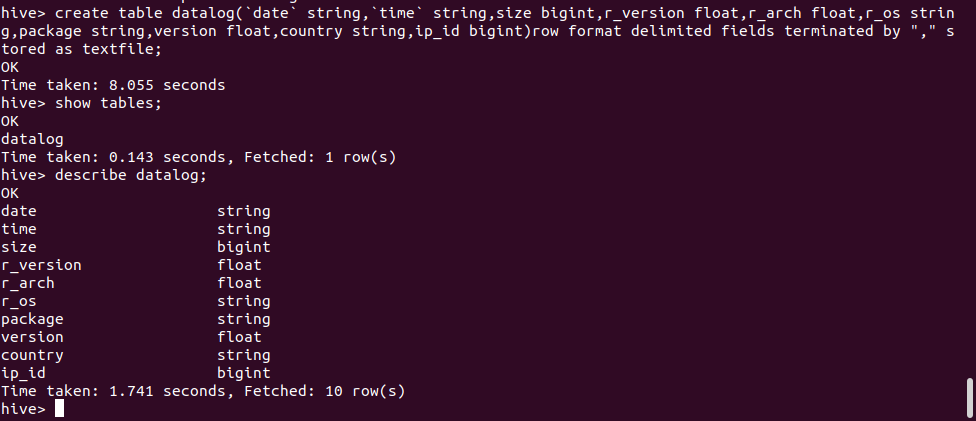
1: Download one week Logs from <http://cran-logs.rstudio.com/>

2: Store RlogFiles in HDFS directory using -put command

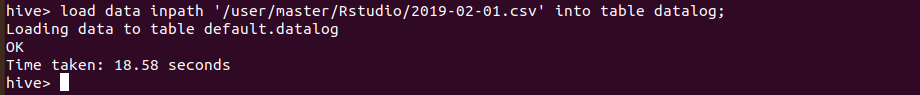




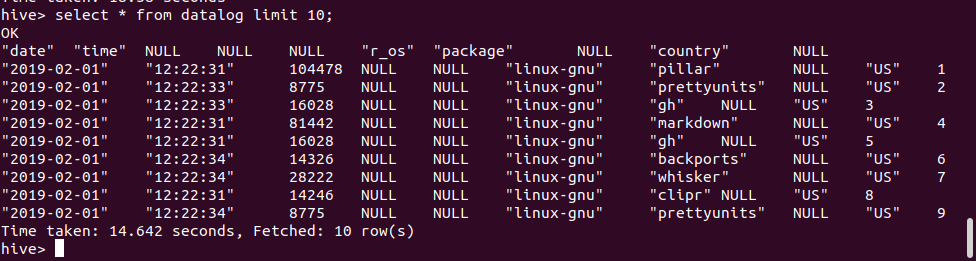
3: Create table in hive to check log files.



4: Load csv file into table datalog in hive

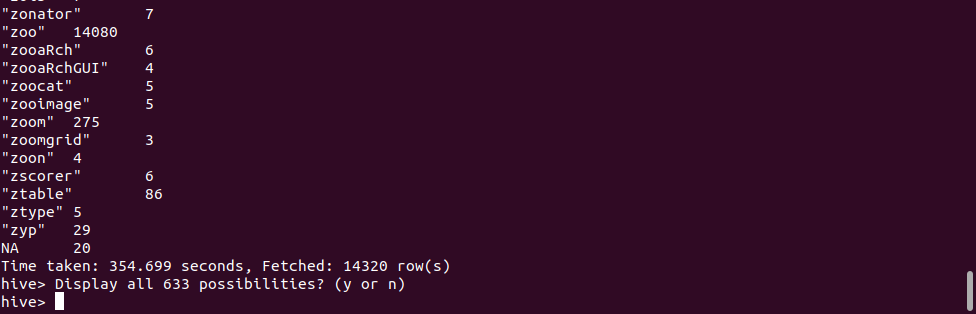


5: Check the first 10 entries of datalog table



6: Count number of occurrences of different packages

Select package, count(\*) from datalog group by package;



7 Count number of Occurrences of different packages by operating system

