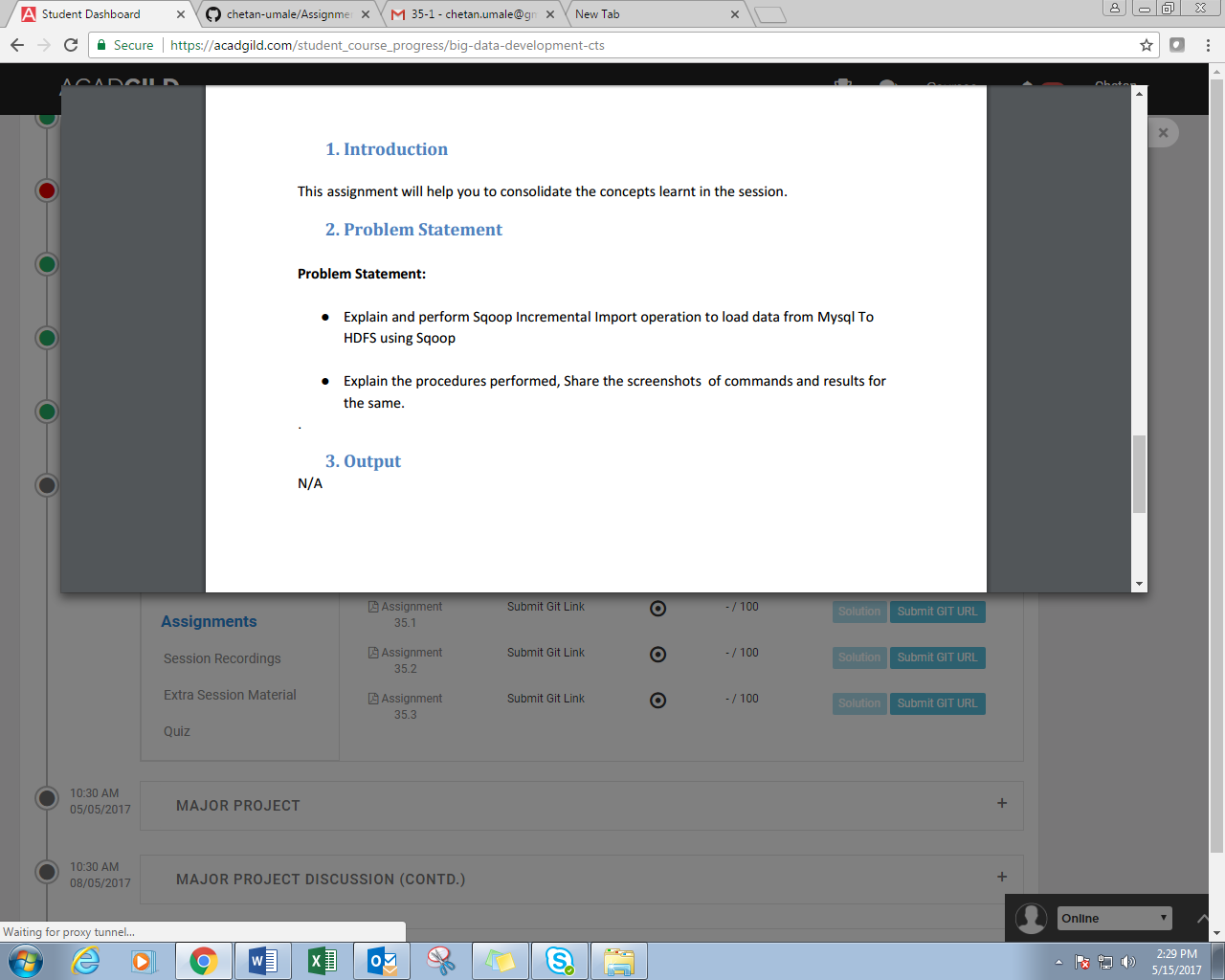
**ASSIGNMENT 35.2**

**PROBLEM STATEMENT:**



**SOLUTION:**

1. Sqoop provides an incremental import mode which can be used to retrieve only rows newer than some previously-imported set of rows.

2. Sqoop supports two types of incremental imports: append and lastmodified.

3. You can use the --incremental argument to specify the type of incremental import to perform.

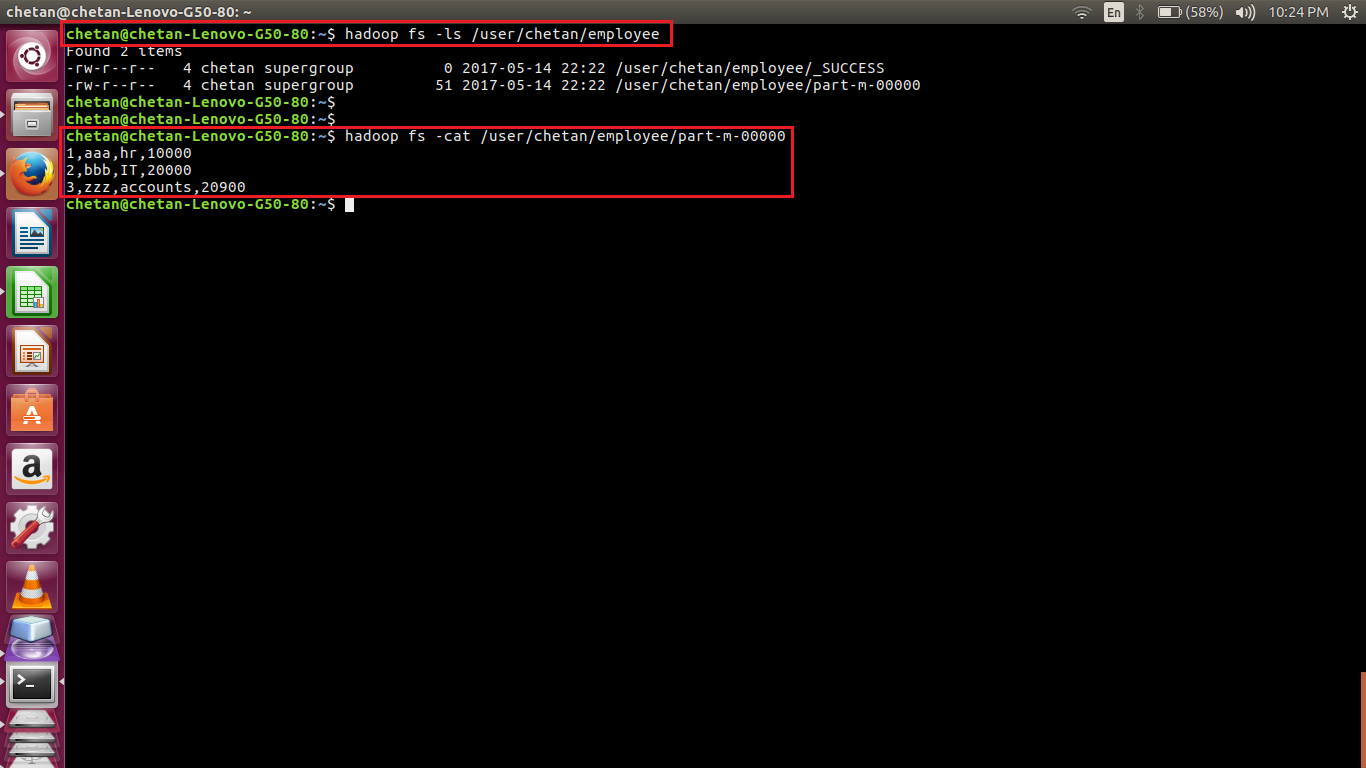
4. You should specify append mode when importing a table where new rows are continually being added with increasing row id values.

5. You specify the column containing the row’s id with --check-column. Sqoop imports rows where the check column has a value greater than the one specified with --last-value.

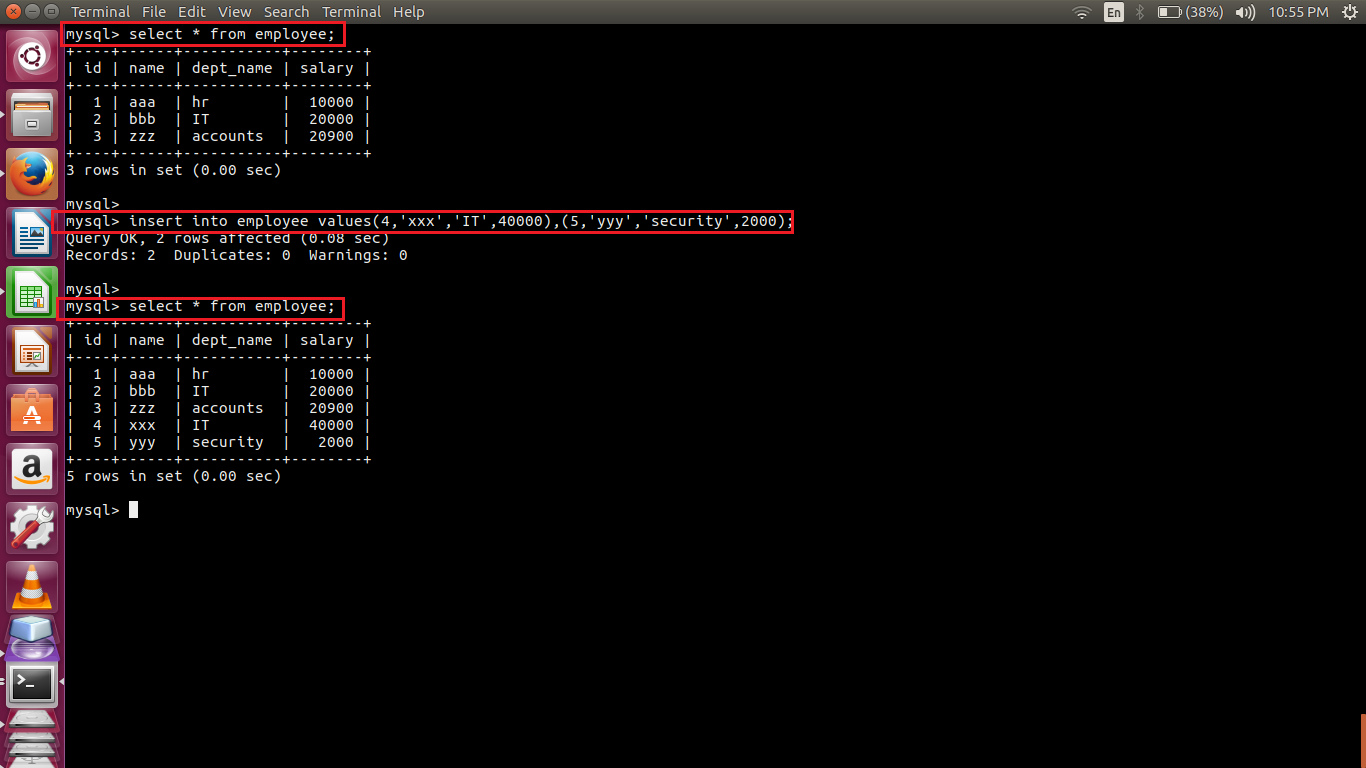
6. An alternate table update strategy supported by Sqoop is called lastmodified mode. You should use this when rows of the source table may be updated, and each such update will set the value of a last-modified column to the current timestamp. Rows where the check column holds a timestamp more recent than the timestamp specified with --last-value are imported.

7. At the end of an incremental import, the value which should be specified as --last-value for a subsequent import is printed to the screen. When running a subsequent import, you should specify --last-value in this way to ensure you import only the new or updated data. This is handled automatically by creating an incremental import as a saved job, which is the preferred mechanism for performing a recurring incremental import.

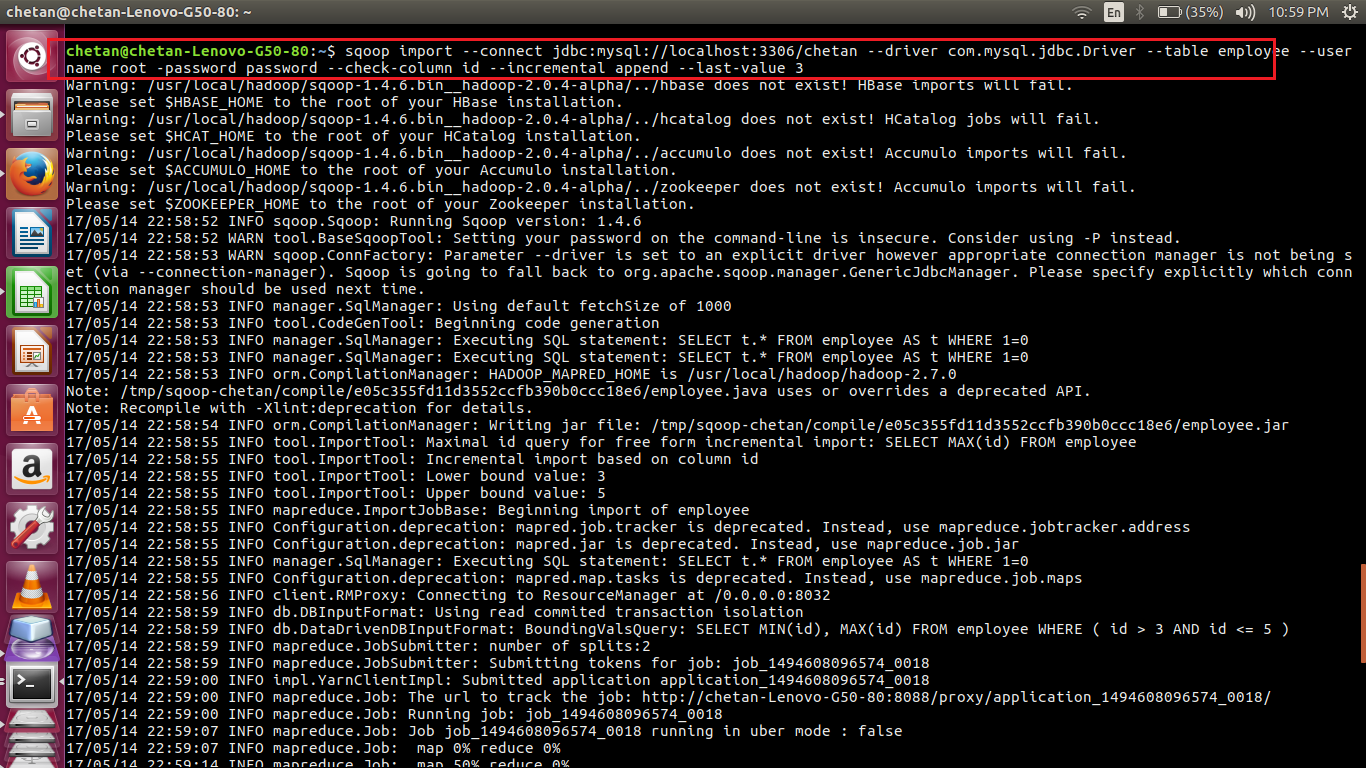
//displaying the contents of the file already present in hdfs.



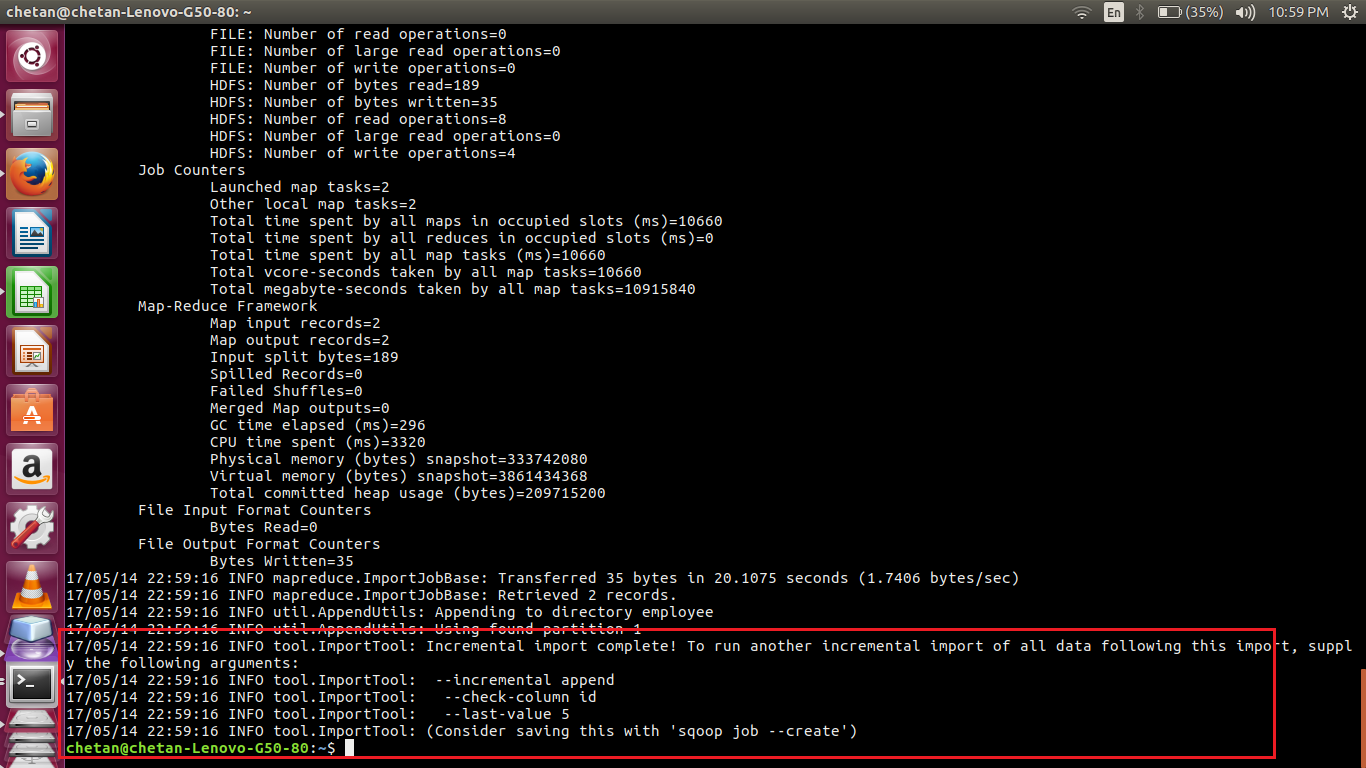
//displaying the contents of table. Adding new records and displaying new contents of table



//sqoop incremental import command



//incremental import complete



//2 new files are created as 2 new records are imported. The contents of these files are displayed

