Problem Statement:

In continuation with previous week's Hive assignment, once We have the following tables available:

1. Two optimized hive tables created based on Covid19_india data and State-wise testing details data,
2. One consolidated table formed in hive by joining these two tables.

Now we want to dump the data in Hbase table for quick searching.

For that a hive managed Hbase table needs to be created. We need to populate the Hbase table using normal hive table. Schema should be same as hive table.

We want the UI team to use this Hive managed Hbase table to quickly search the records.

Hints:

- state + date will be a Rowkey in HBase, as Rowkey needs to be unique we can choose this combination.
- Insert data into HBase table managed by Hive by loading the output of Hive mapside join operation into the Hbase table.
- Search a record in this Hive managed Hbase table. See how much time it takes. See whether the subsequent read of the same record takes lesser time than previous read.
- Search same record in hive and see that it takes more time.
- Conclude that Hbase performs quicker searches than Hive.

NOTE: This Assignment is for self evaluation so no need to submit. Keep this Assignment with you and analyze yourself once the solution is released.