**Assignment-10.1**

**1.Explain in brief the need of hive.**

* Apache Hive supports analysis of large datasets stored in Hadoop's [HDFS](https://en.wikipedia.org/wiki/HDFS) and compatible file systems such as [Amazon S3](https://en.wikipedia.org/wiki/Amazon_S3) filesystem.
* Supports SQL-like queries (HiveQL), which are implicitly converted into MapReduce or Tez, or Spark jobs. Operating on compressed data stored into the Hadoop ecosystem using algorithms including [DEFLATE](https://en.wikipedia.org/wiki/DEFLATE), [BWT](https://en.wikipedia.org/wiki/BWT), [snappy](https://en.wikipedia.org/wiki/Snappy_(compression)), etc.
* Supports different storage types such as plain text, [RCFile](https://en.wikipedia.org/wiki/RCFile), [HBase](https://en.wikipedia.org/wiki/HBase), ORC, and others.

**2.Is Hive suitable to be used for OLTP systems? Explain in brief.**

No, since Hive does not provide insert and update at row level, it is not suitable for OLTP system.

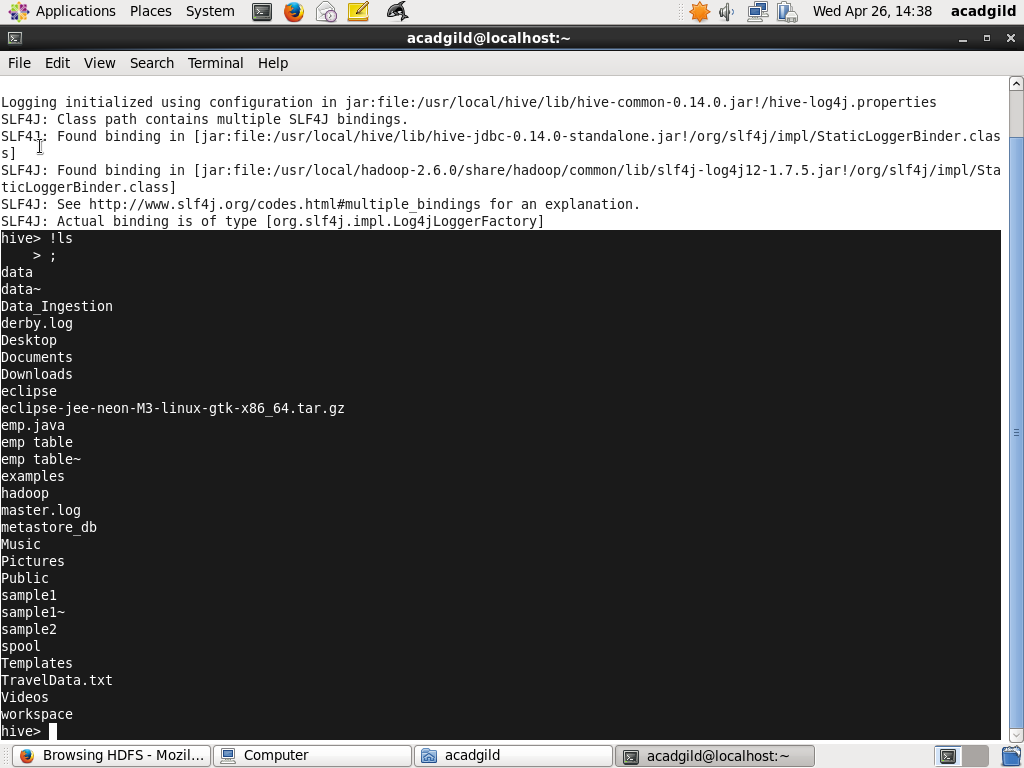
**3.What is a metastore in Hive?**

* Hive metastore is a component that stores all the structure information of objects like tables and partitions in the warehouse including column and column type information, the serializers and deserializers necessary to read and write data and the corresponding HDFS files where the data is stored.
* Any JDBC compliant database can be used as the metastore for Hive, Derby database is the default metastore which supports only one user.

**4.Can we run unix shell commands from hive? Give 5 examples and share the screenshot**.

Yes we can,Hive shell just append the unix shell command by !

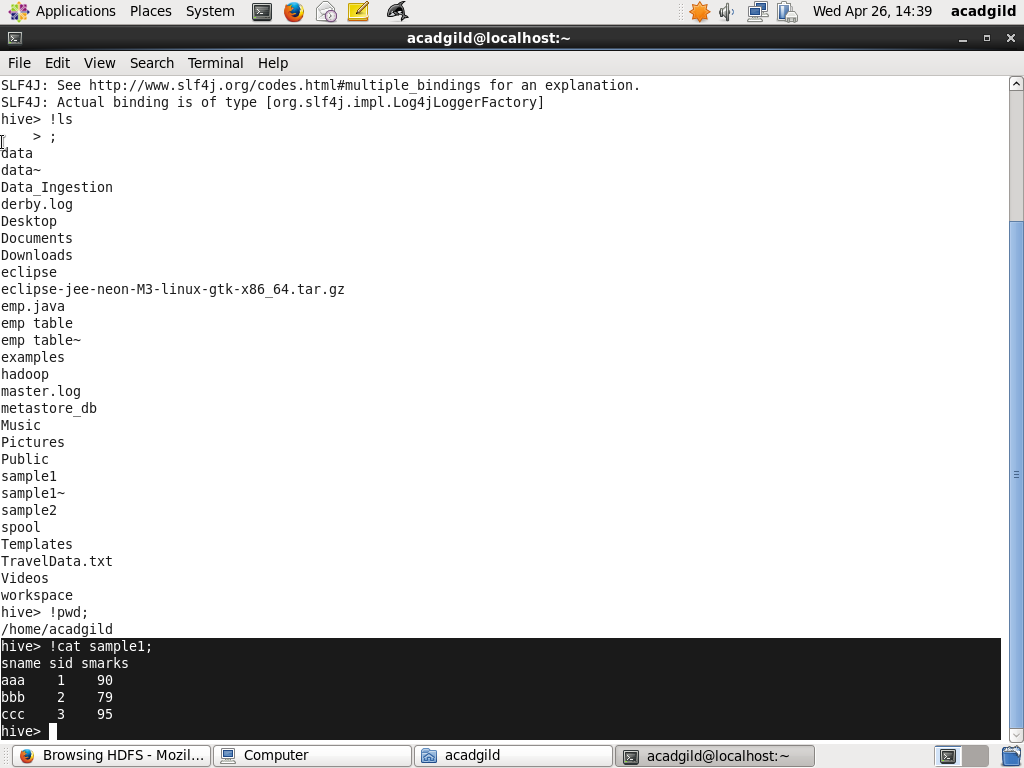
* Ls Command



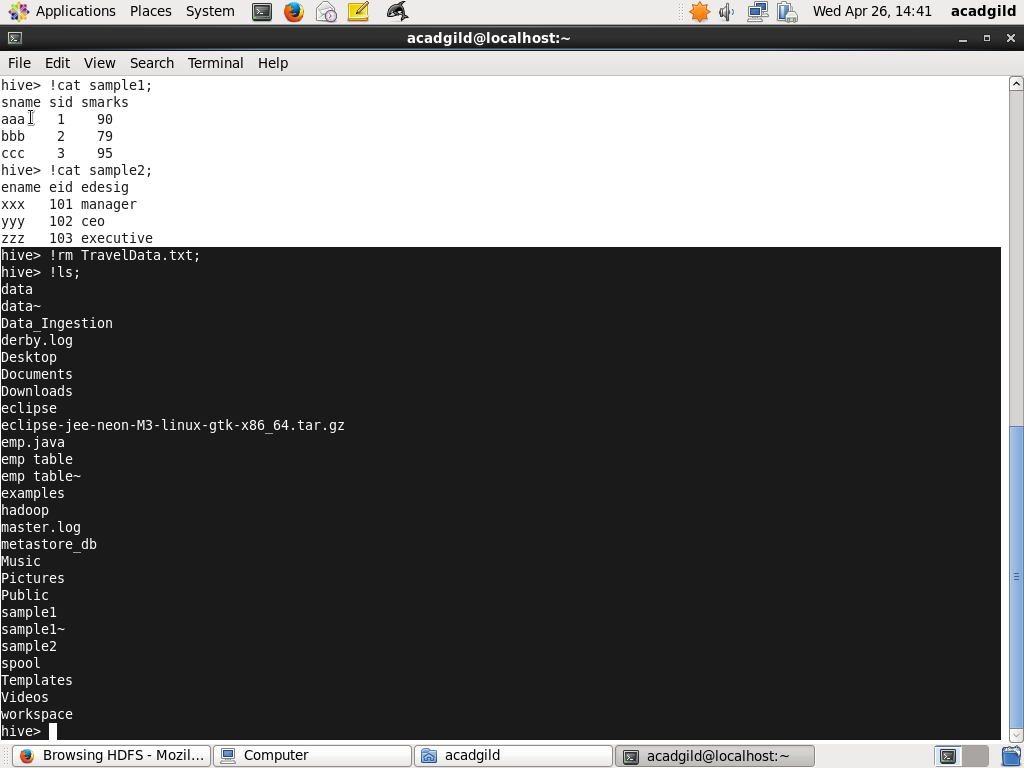
* Pwd Command



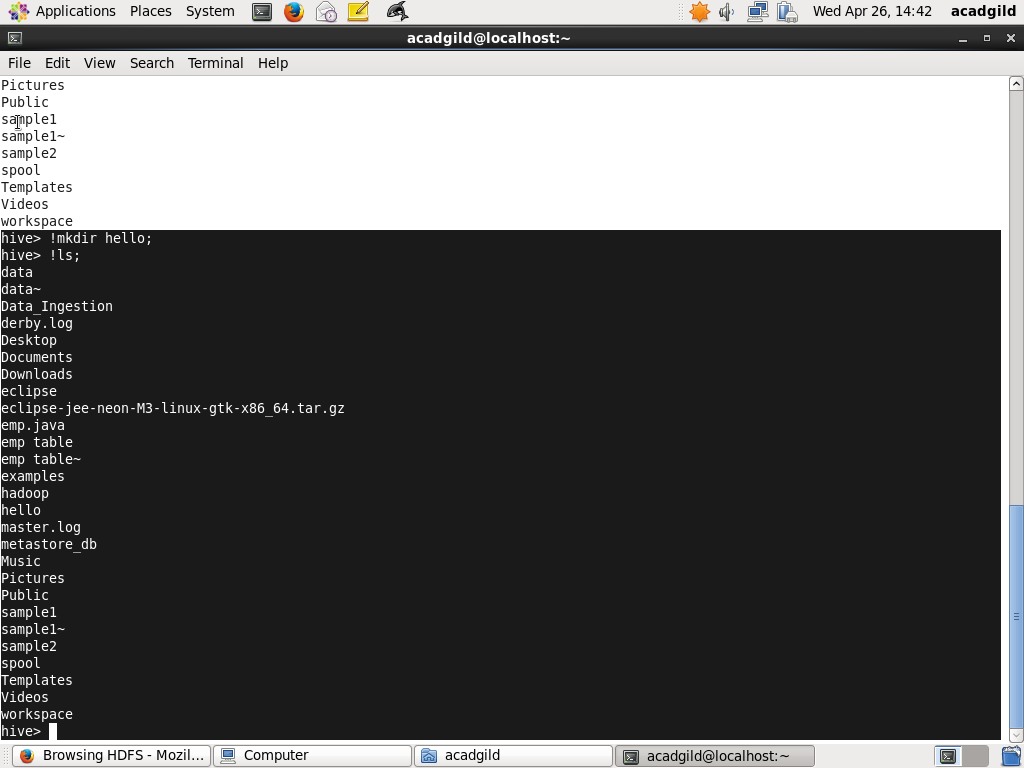
* Cat Command



* Rm Command



* Mkdir Command



**5.Hive can process any type of data formats? Explain in brief.**

* Yes, as Hive Supports any type of data formats such as plain text, [RCFile](https://en.wikipedia.org/wiki/RCFile), [HBase](https://en.wikipedia.org/wiki/HBase), ORC, and others.
* Hive uses the SerDe interface for IO operations. Different SerDe interfaces can read and write any type of data.
* Normally process the data where as different type of data in hadoop, Hive uses different SerDe interface to process such data.