Difference between static and dynamic partitioning

Partition manages the underlying structures of the table’s data directory. In case of partitioned tables, subdirectories are created under the table’s data directory for each unique value of a partition column. When a partition table is queried with one or both partition columns in criteria or in the where clause, what hive effectively does is partition elimination by scanning only those data directories that are needed.

Types of partition

Static partitioning

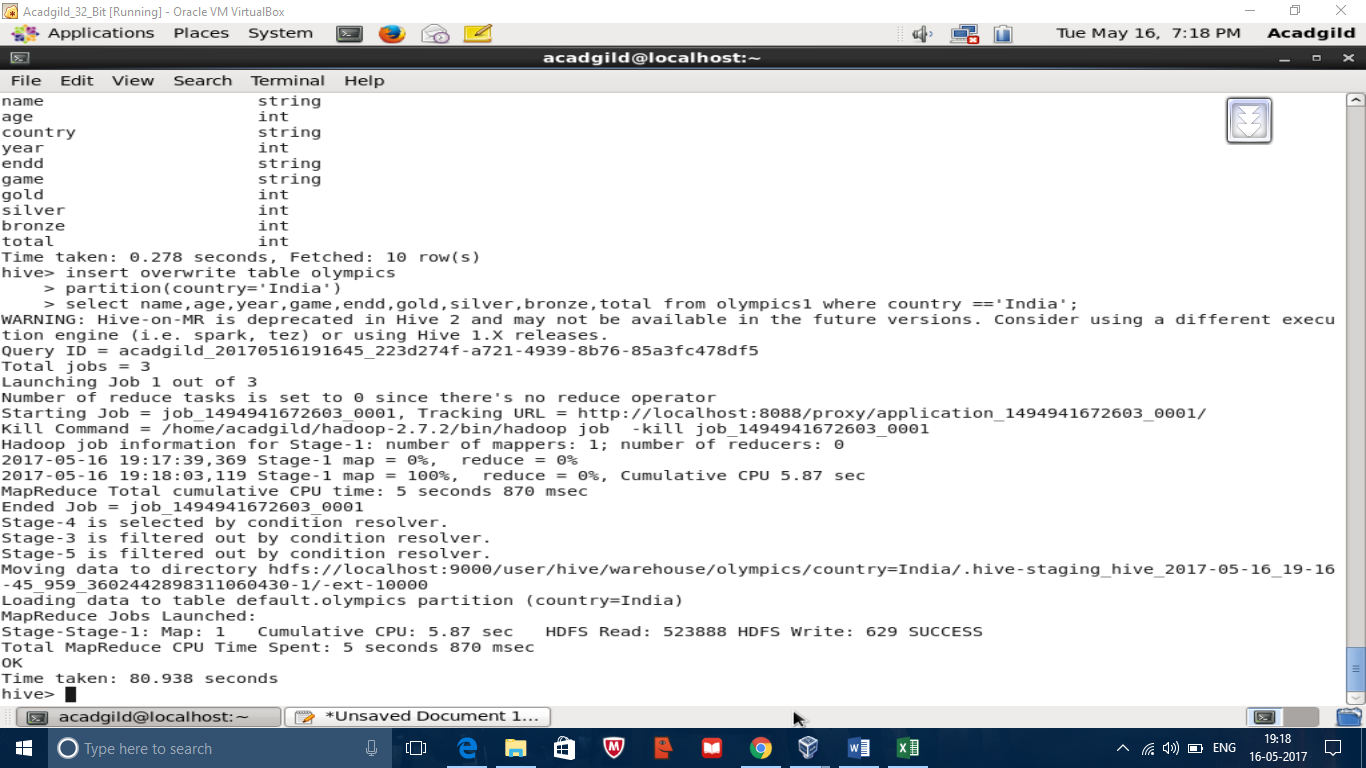
Dynamic partitioning

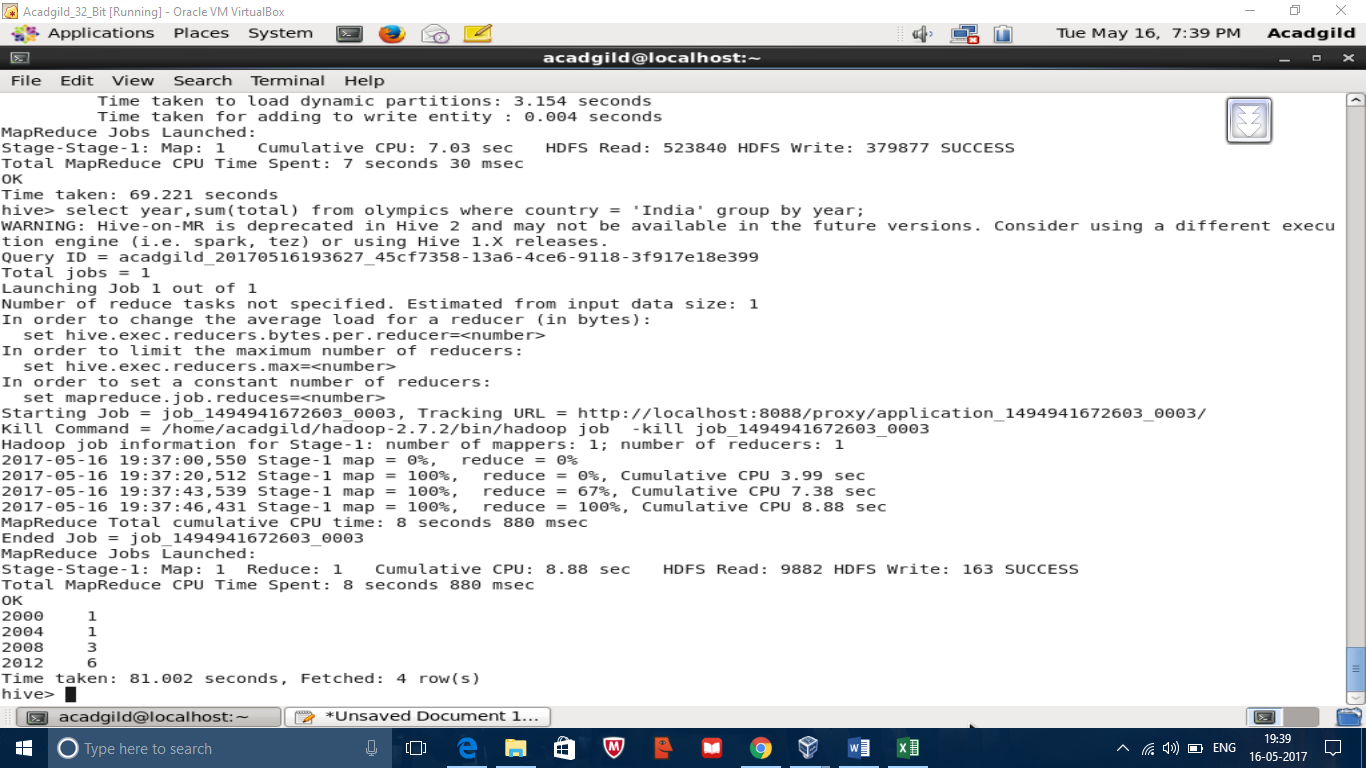
Static partitioning needs to be applied when we know data belongs to which partition

In static partitioning, every partition needs to be backed with individual hive statement which is not feasible for large number of partitions as it will require writing of lot of hive statements. In that scenario dynamic partitioning is suggested s we can create as many number of partitions with single hive statement.

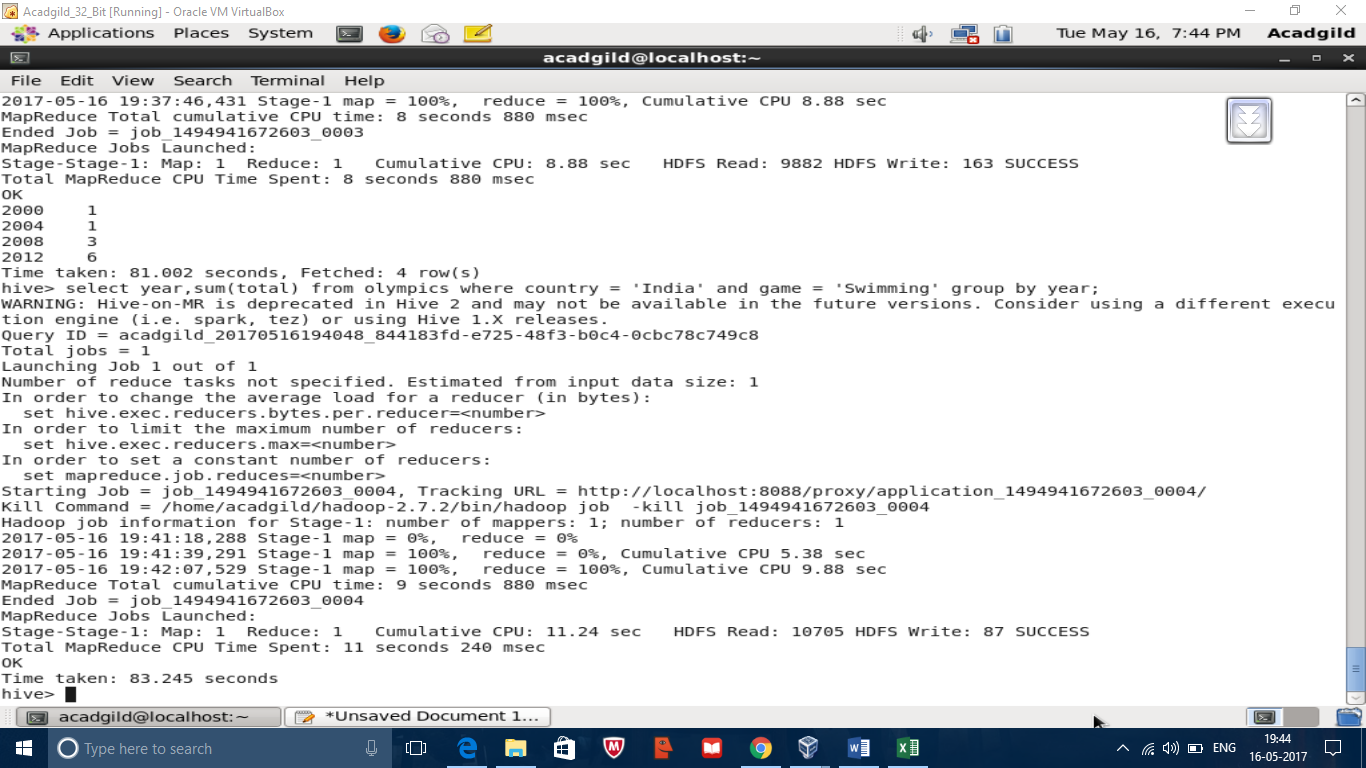
Problem 1.1

Loading using static partitioning

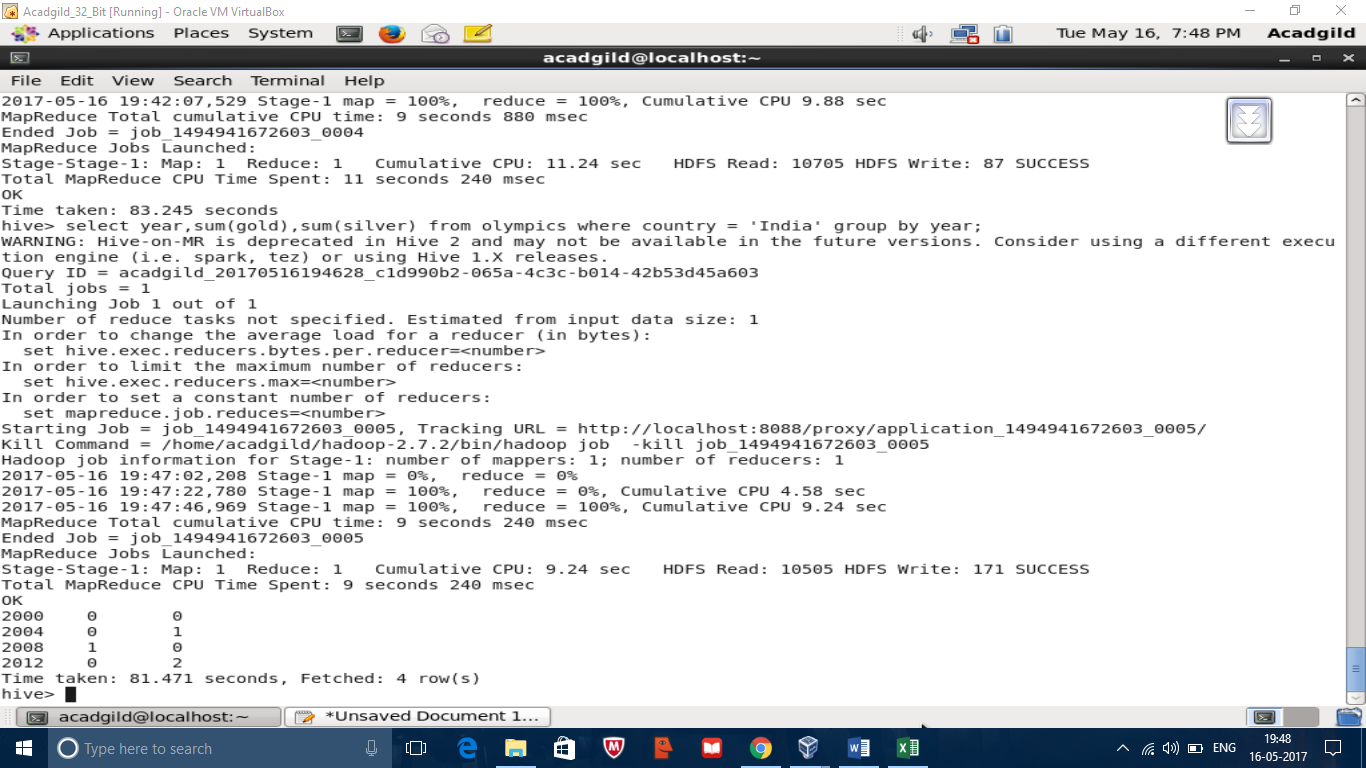




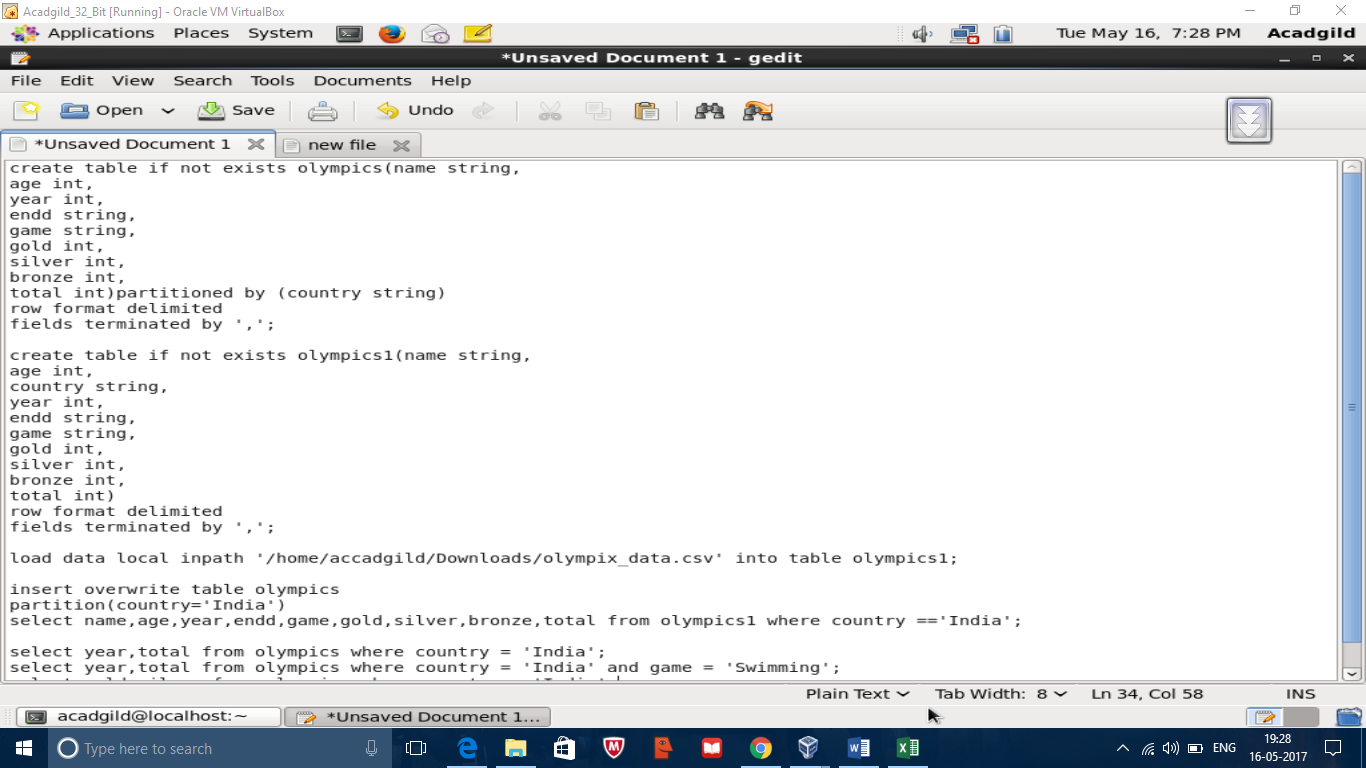
Problem 1.2

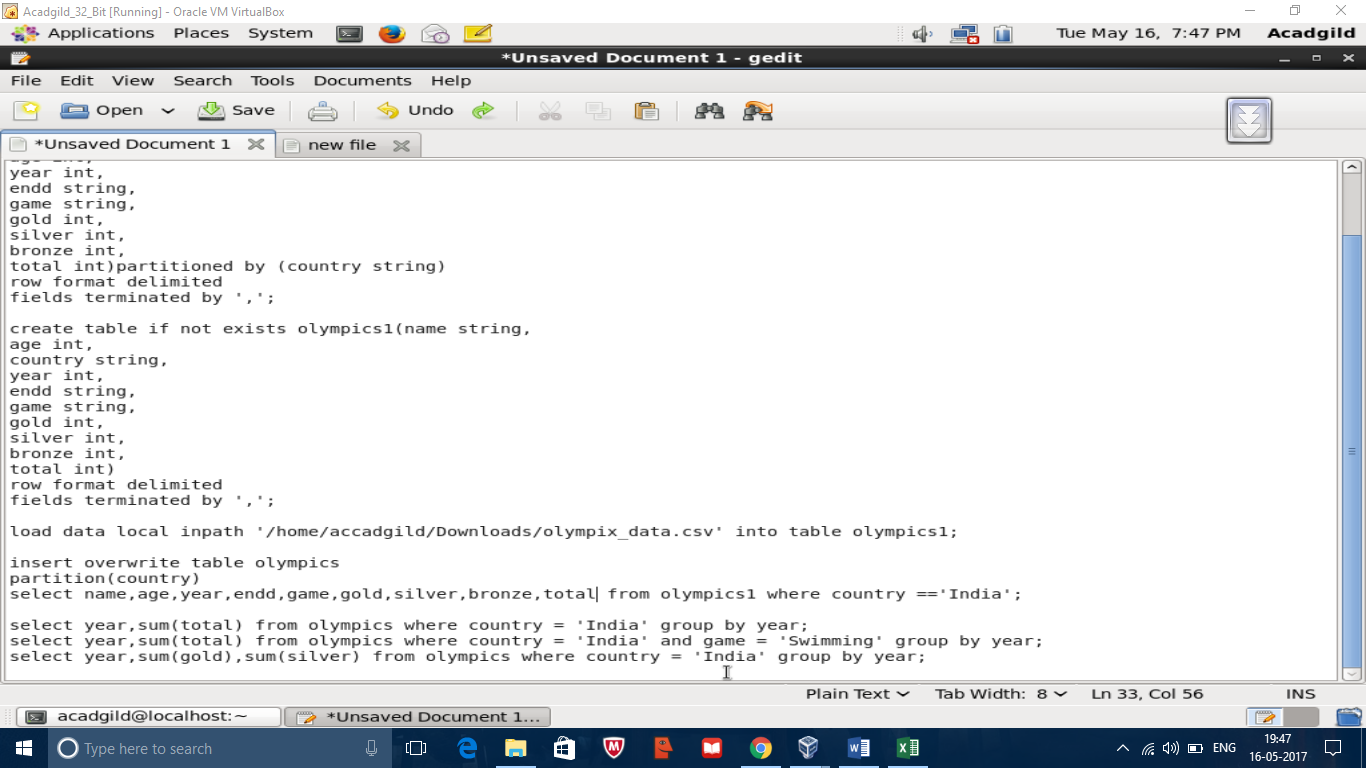


Problem 1.3

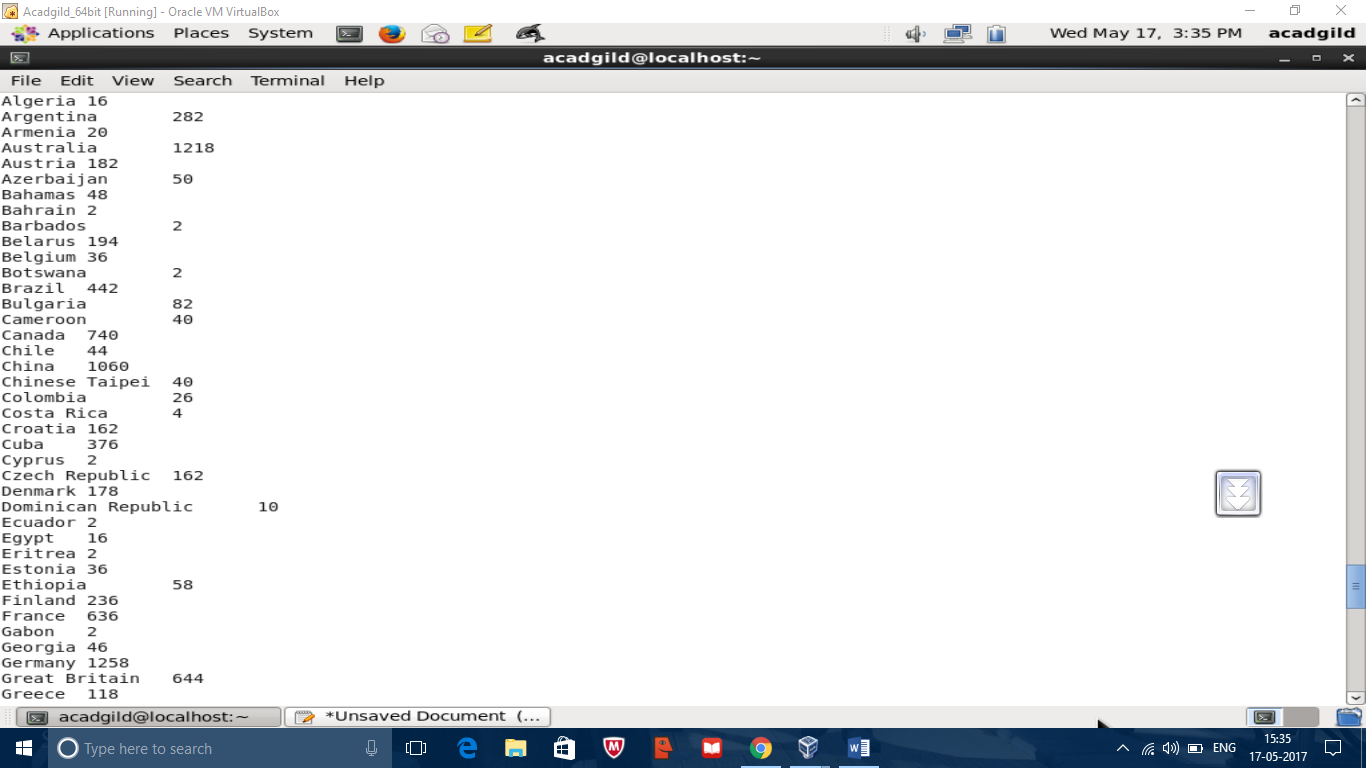


Codes

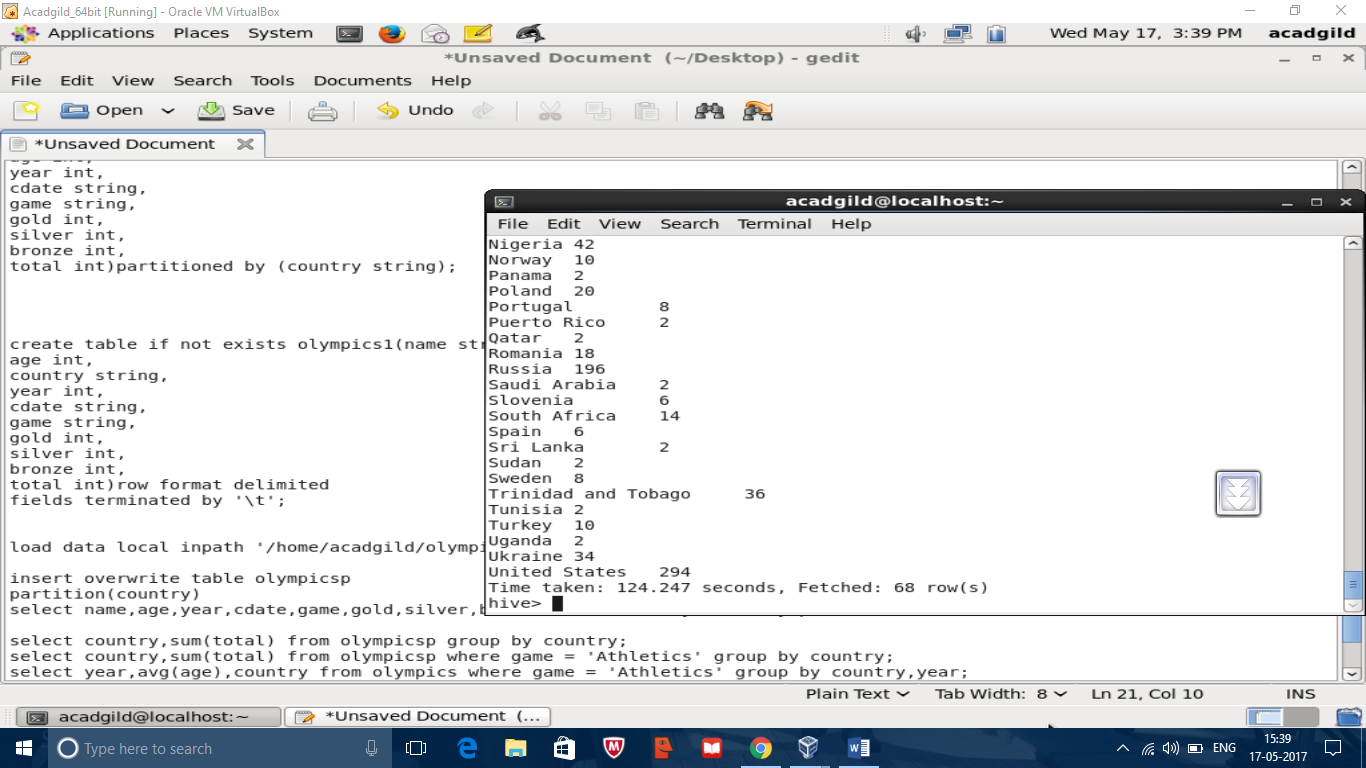




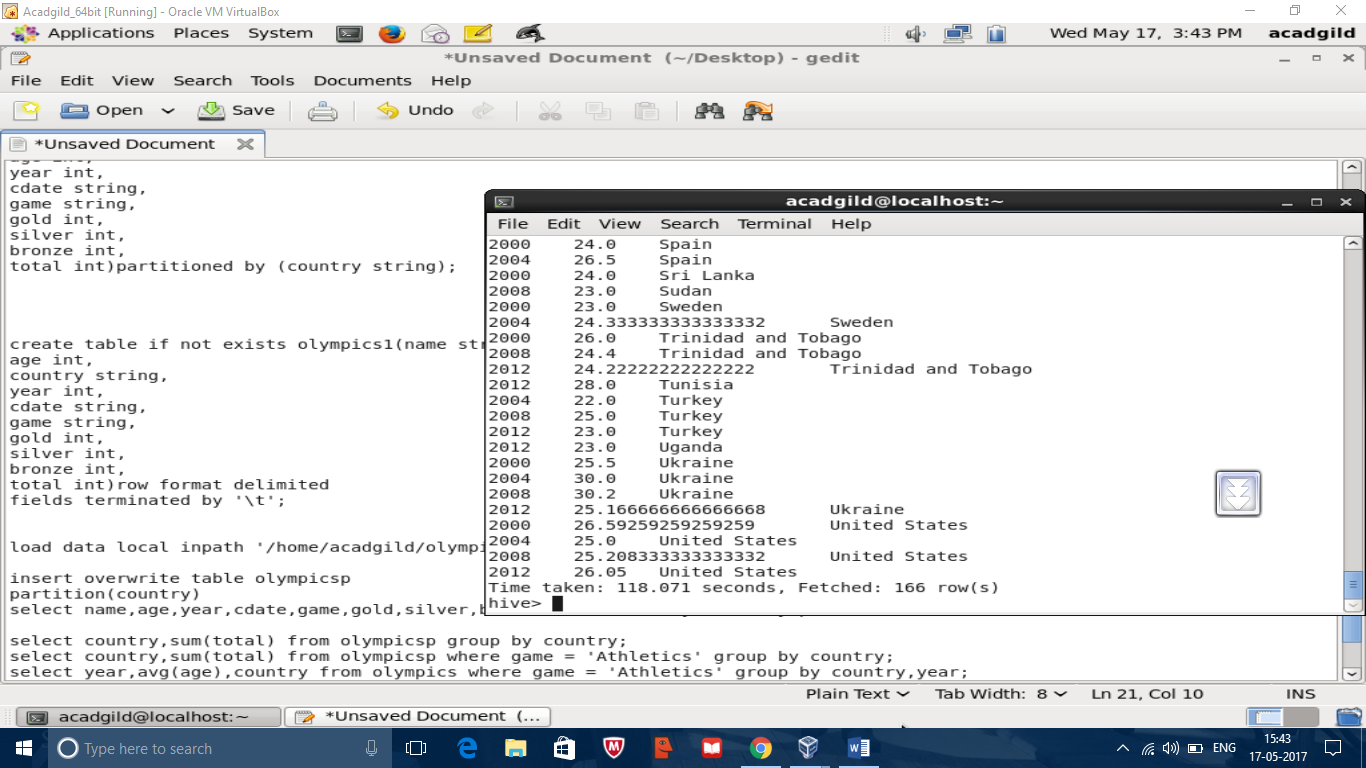
Problem 2.1



Problem 2.2



Problem 2.3



Code

