**Using the below given two datasets you need to give a demo on the below joins in hive.**

**1.Bucket Map join**

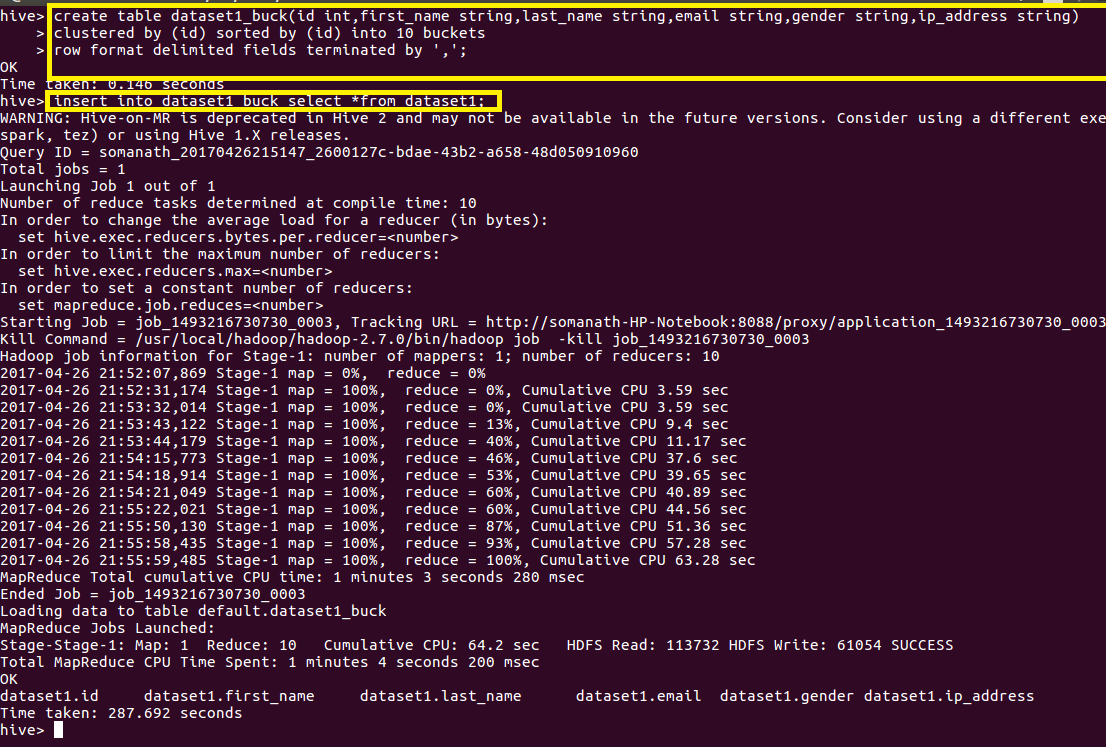
**2.Sort-Merge Bucket join**

**3.Sort-Merge Bucket Map join**

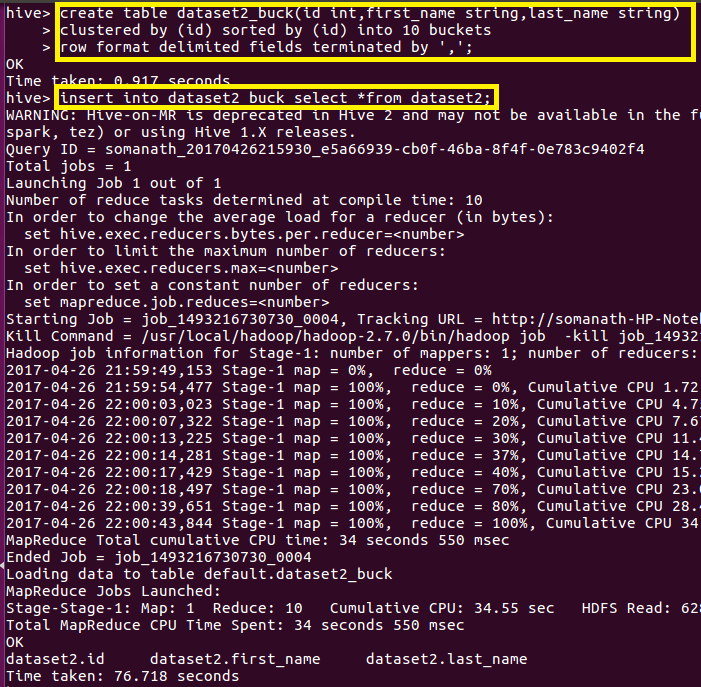
**4.Left semi join**

**Bucket Map join**

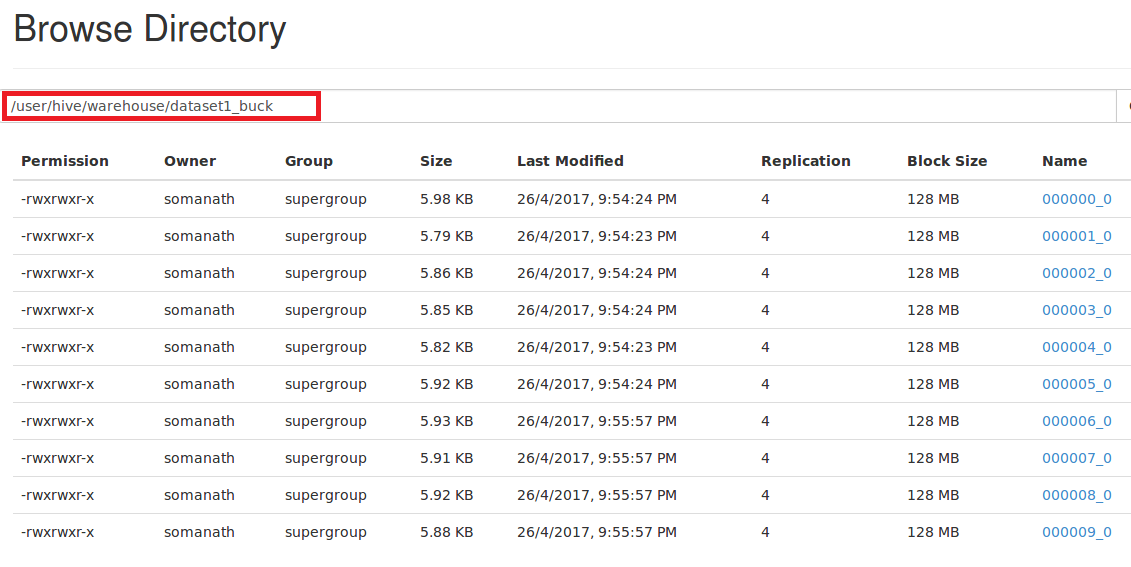
**Step1 Creating Bucketed Table for dataset 1 and inserting the data**

****

**Step2 Creating Bucketed Table for Dataset 2 and inserting the data**

****

**Ensuring Buckets creation**

****

**1.Bucket Map join**

**In normal join, mappers read data of tables on which join needs to be performed and emit key as join key or column on which is expected to be performed . Thus MapReduce framework distributes join key data to single reducer**

**In normal Join in hive joining will take place is reducer but since in this joins joins will be made through buffer memory (which contains small dataset) and large dataset will be streamed in reducer which will actually takes long time**

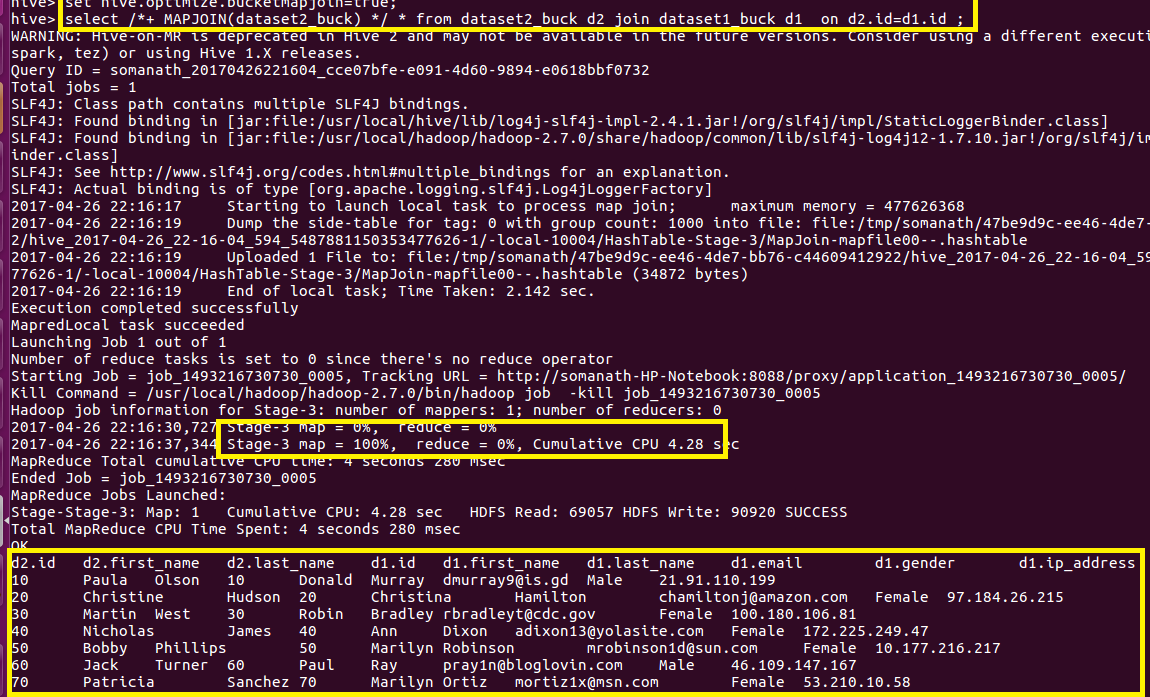
**Instead of this we can use Bucket Map Side Join where joining will be take place in mapper itself**

**Which will results in faster execution of query as this takes place in mapper through distributed cache**

**In order to do Bucket map side join set.hive.auto.convert=true;**

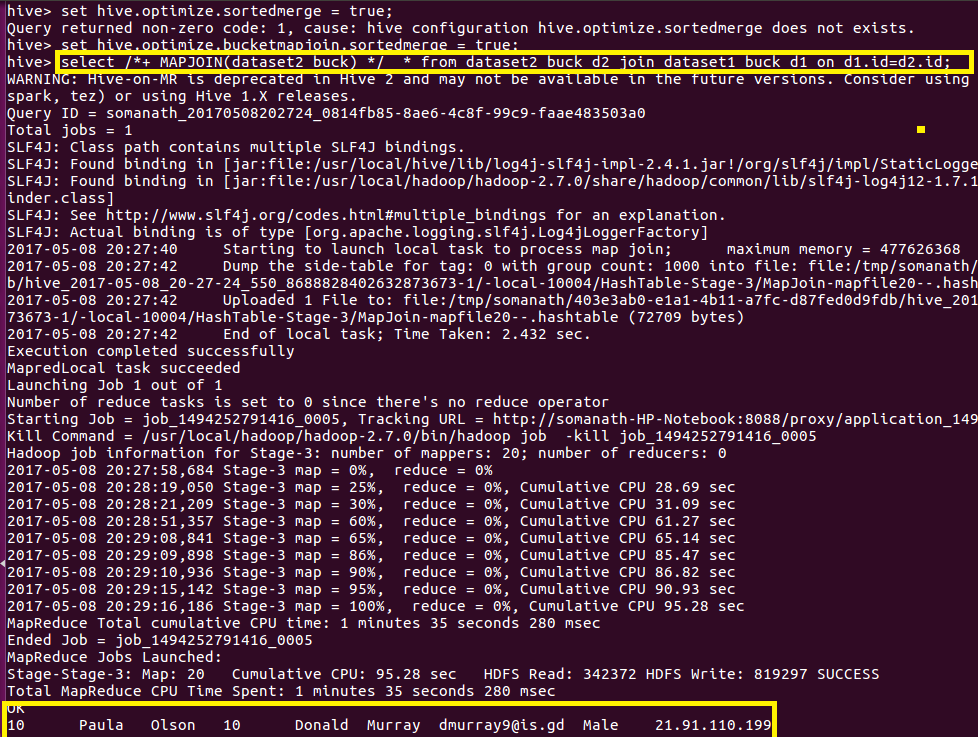
**set.hive.nonconditional.task=true;**

**Set hive.optimise.bucketmapJoin=true**

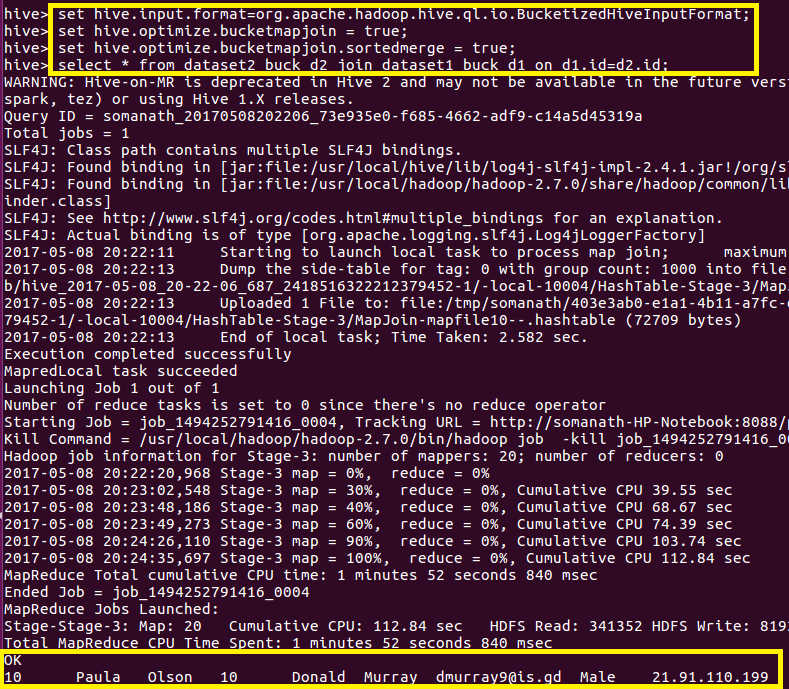
****

**2.Sort-Merge Bucket Map join**

**In this since the data is already sorted and bucketed based on key on which joining should be performed it will be easier for hive to perform join and it results in high optimization.It is a improvisation over bucket map join.**

****

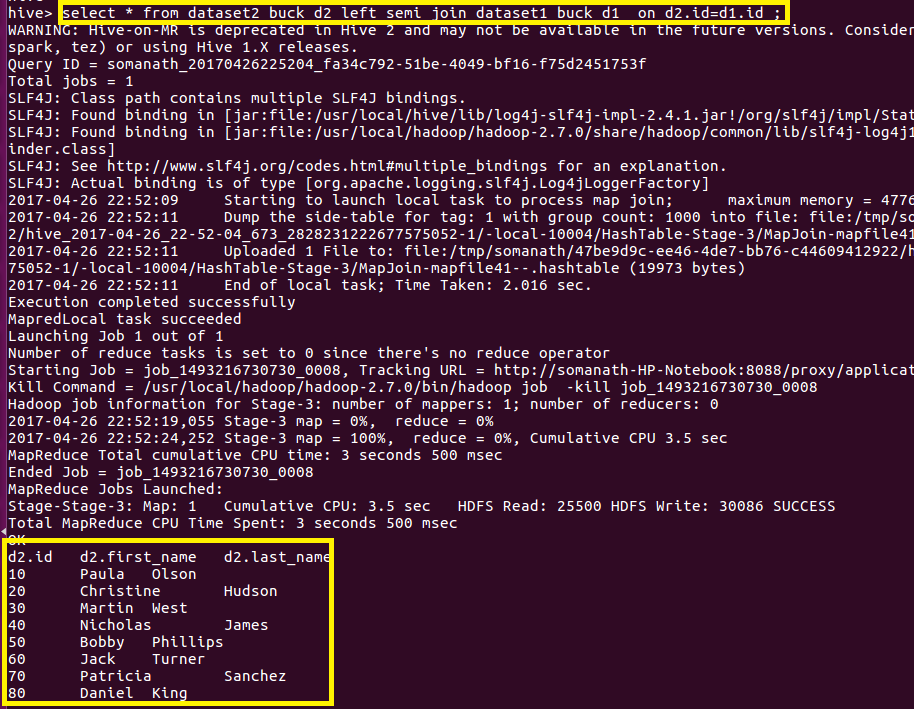
**3.Sort Merge Bucket Join**

****

**4.Left Semi Join**

**It is similar to if in /exists query in efficient way (ie) I t will display all the details of the Table on left side of Join command If a matching value for the join column exists in right table**

**For example In this case all the details of Dataset2 is displayed for every matchingId in right column**

****

**Complete Outpiut ATTACHED**