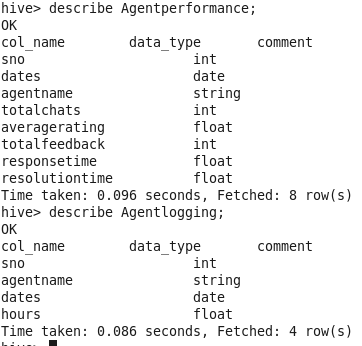
**1. Create a schema based on the given dataset**



**2. List of all agents' names**

select distinct(agentname) from Agentperformance;

**3. Find out agent average rating.**

select agentname, avg(averagerating) as AverageRating from Agentperformance group by agentname;

**4. Total working days for each agents**

select count(dates) as TotalWorkingDays, agentname from Agentperformance group by agentname;

**5. Total query that each agent have taken**

select agentname,sum(totalchats) as TotalQueriesTooked from Agentperformance group by agentname;

**6. Total Feedback that each agent have received**

select agentname,sum(totalfeedback) as TotalFeedback from Agentperformance group by agentname;

**8. Agent name who have average rating between 3.5 to 4**

**8- Agent name who have rating less than 3.5**

select agentname from Agentperformance where averagerating < 3.5 group by agentname;

**9-Agent name who have rating more than 4.5**

select agentname from Agentperformance where averagerating > 4.5 group by agentname;

**10. Agent name who have rating more than 4.5**

select agentname from Agentperformance where averagerating >4.5 group by agentname;

**11. How many feedback agents have received more than 4.5 average**

select sum(totalfeedback) from Agentperformance where averagerating>4.5;

**Perform inner join**

select ap.agentname, ap.averagerating ,al.hours from Agentperformance ap inner join Agentlogging al on (ap.agentname=al.agentname);

INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/Desktop/exportedfiles/data.csv' ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' select ap.agentname, ap.averagerating ,al.hours from Agentperformance ap inner join Agentlogging al on (ap.agentname=al.agentname);

**Perform Left Outer JOIN:**

INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/Desktop/exportedfiles/data.csv' ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' select ap.agentname, ap.averagerating ,al.hours from Agentperformance ap left outer join Agentlogging al on ap.agentname=al.agentname;

**RIGHT OUTER JOIN**

INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/Desktop/exportedfiles/data.csv' ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' select ap.agentname, ap.averagerating ,al.hours from Agentperformance ap right outer join Agentlogging al on ap.agentname=al.agentname;

**CREATE TABLE AGENT PERFORMANCE WITH PARTITIONING AND BUCKETING**

create table Agentperformance\_DPB

> (

> srno int,

> Tareekh date,

> totalchats int,averagerating float,totalfeedback int,responsetime float,resolutiontime float)partitioned by (agentname string)clustered by (srno) into 10 buckets;

set hive.exec.dynamic.partition.mode=nonstrict;

set hive.enforce.bucketing=true;

insert overwrite table Agentperformance\_DPB partition(agentname) select sno,dates,totalchats,averagerating,totalfeedback,responsetime,resolutiontime,agentname from Agentperformance;

**CREATE TABLE AGENT LOGGING WITH PARTITIONING AND BUCKETING**

create table Agentlogging\_DPB

> (

> srno int,

> tareekh date,

> hours float

> )

> partitioned by (agentname string)clustered by (srno) into 10 buckets;

insert overwrite table Agentlogging\_DPB partition(agentname) select sno,dates,hours,agentname from Agentlogging;