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

Ans: -Start the hive shell and then create the database project.

Command = create database project;

Use project;

Create table agent\_login\_report\_csv and agent\_performance\_csv

1. **Dump the data inside the hdfs in the given schema location?**

Ans: -Move the data from local to HDFS by using filezila

Then dump the data into the tables which we created using command

Command= load data local inpath

‘file:/path of the file in local’ into table Agent\_Login\_Report\_csv and

Agent\_performance\_csv.

1. **List of all agents' names?**

Ans: -Select agentname fromAgent\_performance\_csv.

1. **Find out agent average rating?**

Ans: - command= select AVG (averagetrating) from agent\_performance\_csv;

1. **Total working days for each agent?**

Ans: - select agent , count(date) from agent\_loging\_report\_csv group\_by agent;

6. Total query that each agent have taken

**7. Total Feedback that each agent have received?**

Ans: - select agentname , count(totalfeedback) from agent\_performance\_data\_csv group\_by agentname;

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

Ans: - select agentname from agent\_performance\_data\_csv where agentrating between 3.5 AND 4;

**9. Agent name who have rating less than 3.5?**

Ans: - select agentname from agent\_performance\_data\_csv where averagerating<3.5;

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

Ans: - select agentname from agent\_performance\_data\_csv where averagerating<3.5;

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

Ans: - select agentname from agent\_performance\_report\_csv where totalfeedback>4.5;

**12. average weekly response time for each agent?**

Ans: - select avg(averageresponsetime) , over(partition by date group by name ) from agent\_performance\_data\_csv;

**13. average weekly resolution time for each agents ?**

Ans: - select avg(averageresolutiontime) , over(partition by date group by name ) from agent\_performance\_data\_csv;

**14. Find the number of chat on which they have received a feedback?**

Ans: - select totalchats from agent\_performance\_data\_csv where totalfeedback>0;

15. Total contribution hour for each and every agents weekly basis

**16. Perform inner join, left join and right join based on the agent column and after joining the table export that data into your local system?**

Ans: - inner join command = select agent\_loging\_report\_csv.SLNO, agent\_performance\_data\_csv.agentname

from agent\_loging\_report\_csv

INNER JOIN agent\_performance\_data\_csv on agent\_loging\_report.SLNO = agent\_performance\_report\_csv.SLNO

Left join command = select agentame. agent\_perfromance\_data\_csv , SLNO.agent\_loging\_report\_csv

FROM agent\_performance\_data\_csv

LEFT JOIN on agent\_loging\_report\_csv.SLNO = agent\_performance\_data\_csv.SLNO

ORDER BY agent\_perfromance\_data\_csv.agentname;

Right join command =select agent\_loging\_report\_csv.SLNO , agent\_performance\_data\_csv.agentname

FROM agent\_loging\_report\_csv RIGHT JOIN

agent\_performance\_data\_csv ON agent\_loging\_report\_csv.agent = agent\_perfromance\_data\_csv.agentname order by agent\_loging\_report\_csv.SLNO;

**17. Perform partitioning on top of the agent column and then on top of that perform bucketing for each partitioning?**

Ans: - create table agent\_loging\_report1\_csv(

SLno int,

Agent string,

Date string,

Login time string,

Logout time string,

Duration string)

Partition by hash (date)

Partitions 21**;**