Group A

LAB 2b

Hive Part1

Participants:

1. Aadarsha chapagain

2.Jyoti shukla

3.Rishi Phaneendra Varma

4.Priti Bhale

5.Sreya Treesa Johny

6.Piyush Bhatia

**Objective:** *The objective is to find the number of male staff who has been serving more than 15 years and are still active*.

1. The public data set was uploaded with winscp and copied to hdfs file system and permission was changed.

**Graphical user interface, text, application

Description automatically generated**

1. **Managed table without partition**

**Graphical user interface, text, application

Description automatically generated**

**Analysis:**

To achieve the objective the following query was run.

*select count(\*) from staff\_man where gender\_full='Male'and length\_of\_service > 15 and status='ACTIVE'*;

Timeline

Description automatically generated

Text

Description automatically generated

1. **Managed table with partition**

*Create table staff\_man\_part(age int, length\_of\_service int, city\_name string,*

*department\_name string, job\_title string, STATUS string, BUSINESS\_UNIT string)*

*Partitioned by(gender\_full string) row format delimited fields terminated by ',' stored as textfile tblproperties("skip.header.line.count"="1")*

*Load data inpath '/user/aadarsha/externaldata/sample\_data.csv' into table staff\_man\_part*

*partition(gender\_full='Male')*

Text

Description automatically generated

1. **External table without partition**

Keyword external is used to create a external table;

**Graphical user interface, text, application

Description automatically generated**

**Analysis**

*select count(\*) from staff\_ext where gender\_full='Male' and length\_of\_service >15 and status ='ACTIVE';*

A picture containing timeline

Description automatically generated

Graphical user interface, text, application

Description automatically generated

1. **External table with partition**

Graphical user interface, text, application

Description automatically generated

1. **External table with dynamic partition**

Graphical user interface, text, application

Description automatically generated

*Here it is seen that for dynamic partition insert command is used instead of load command and the field by which partition is done is mentioned at last***.**

*select count(\*) from staff\_dynamic\_part where gender\_full='Male' and status='ACTIVE' and length\_of\_service>15;*

Graphical user interface, text, application

Description automatically generated

And as per our objective:

Text

Description automatically generated