# Ex no: 6 210701511

ImportaJSONfilefromthecommandlineandapplyactions withthedata presentintheJSONfile

# Aim:

ToimportaJSONfilefromthecommandlineandapplythefollowingactionswiththe

datapresentintheJSONfilewhere,projection,aggregation,remove,count,limit,skipandsort.

# Procedure:

**HiveDownload andinstallation:**

1. StartingHadoopServices

OpenPowerShellasadministratorandgotoHadoopsbindirectoryandstarthadoopservices using thefollowing commands:Start-all.cmd



1. Createa.json filewiththebelowcontent:

{"id":1,"name":"JohnDoe","age": 30,"salary":50000}

{"id":2,"name":"JaneSmith","age":25,"salary":60000}

{"id":3,"name":"AliceJohnson","age":28,"salary":55000}

{"id":4,"name":"BobBrown","age": 35,"salary":70000}

{"id":5,"name":"CharlieDavis","age":40,"salary":80000}

{"id":6,"name":"EveWhite","age":22,"salary":48000}

{"id":7,"name":"FrankBlack","age":32,"salary": 65000}

{"id":8,"name":"GraceGreen","age":27,"salary":52000}

{"id":9,"name":"HenryGold","age":29,"salary":59000}

{"id":10,"name":"IsabelBlue","age":33,"salary":73000}DerbyNetworkServer:

Run the following command to open Derby:StartNetworkServer-h 0.0.0.0



GotofirstPowerShellwindowandcheckwhether NetworkServerControlisrunning.



1. StartingApacheHive:

GotoApacheHive’sbinlocationwithcdcommandandrunthefollowingcommand:hive--serviceschematool-dbTypederby–initSchema

8. Open Hive shell by typing:hive

# CreateaDatabase:

Startbycreating adatabase.Open theHiveCLIandfollowthestepsbelow:

1. Usethe**CREATEDATABASE**statementtocreateanewdatabase:CREATEDATABASEIFNOTEXISTSemp\_json;
2. Verifythedatabaseispresent:

SHOWDATABASES;



# CreateaTableinHive:

CREATETABLEemployees\_table(idINT,

name STRING,age INT, salaryDOUBLE

)

ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'STOREDASTEXTFILE

LOCATION'/ser/hive/warehouse/emp\_json/';



# AddDatatotheTABLE:

Run the **LOAD DATAINPATH**command:

LOADDATAINPATH'/user/hive/warehouse/emp\_json/employee.json' INTOTABLEemployees\_table;



# ListHiveTablesandData:

Toshowalltablesinaselecteddatabase,usethefollowingstatement:

SHOWTABLES;



Toshowtablecolumnnamesanddatatypes,run:DESC employees\_table;

Todisplaytabledata,usea**SELECT**statement.Forexample,toselecteverythinginatable,run:

SELECT\*FROMemployees\_table;



**PerformVariousOperationsontheDatainthetable:WHERE:**

SELECTid,name,age,salaryFROMemployees\_tableWHEREsalary>60000;



# PROJECTION:(SelectingSpecificColumns)

SELECTid,nameFROMemployees\_table;





# AGGREGATION:(e.g.,Summing SalariesbyAgeGroup)

SELECTage,MAX(salary)ASmax\_salaryFROM employees\_tableGROUPBYage;



# REMOVE:(RemoveSpecificRecords)

SELECT\*FROMemployees\_tableWHEREsalary>70000;



# COUNT:(CountingtheNumberofRecords)

SELECTCOUNT(\*)FROMemployees\_table;



# LIMIT:(RestricttheNumberofRowsReturned)

SELECT\*FROMemployees\_tableLIMIT5;



# SKIP:(SkippingtheFirstNRows,usingRowNumber)

SELECT\*FROM(SELECT\*, ROW\_NUMBER()OVER ()ASrow\_numFROM

employees\_table)tempWHERErow\_num>3;



# SORT:(OrdertheDatabySalary)

SELECT\*FROMemployees\_tableORDERBYsalaryDESC;



# Result:

Thus,toimport aJSONfilefrom thecommand lineandapply thefollowingactionswiththe data present in the JSON file where, projection, aggregation, remove, count, limit, skip andsortwas completed successfully.