



## **PROOF OF CONCEPT (POC)**

### **REPORT ON BIG DATA**

**Topic: COVID-19 and its Impact on Students**

Submitted by:  
Aryan Sengar  
2020BTechCSE013

**CS1312: Big Data Engineering**

Under the Guidance of:  
Dr. Alok Kumar

**Institute of Engineering & Technology  
JK Lakshmi Pat University, Jaipur**

**April 2023**

## **DECLARATION**

I hereby declare that I have learnt the concepts of Big Data under the course CS1312: Big Data Engineering under the guidance of (Dr. Alok Kumar). I hereby undertake that the project undertaken by me is the genuine work of mine.

Name of Student: Aryan Sengar  
Registration no: 2020BTechCSE013

Date:13/04/2023

## **ACKNOWLEDGEMENT**

I would like to express my thanks and gratitude to my teachers (**Dr. Alok Kumar**) who have equipped me and motivated to do this project on the field of Data analysis (COVID-19 and its Impact on Students), I have learned some terms of big data explored the data with new tools and techniques that I have learned and I am really thankful to him.

- Aryan Sengar

## **Table of Contents**

<b>S.NO.</b>	<b>Topic</b>
<b>1.</b>	Introduction
	1.1 Purpose
	1.2 Project Scope
	1.3 References
	1.4 About Data Set
	1.5 Details of Data Set
	1.6 Screenshots
	1.7 Attribute Information
<b>2.</b>	Tools and Working Essentials
	2.1 Assumption and Dependencies
	2.2 Installation
<b>3.</b>	Coding on Tools
	3.1 HDFS
	3.2 Sqoop
	3.3 HIVE
	3.4 PIG

# **1. INTRODUCTION**

## **1.1. Purpose**

The aim of this project is to analyse the Covid-19 student survey data. The data consists of data for Covid-19 student responses about impact on studies due to the lockdown and online mode education.

## **1.2. Project Scope**

1. It could be used to understand the effectiveness of online based learning.
2. It could be used to check impact of social media on education.
3. The insights from the data can be used to make better education transformation and how the quality of online classes can be improved.

## **1.3. References**

<https://www.kaggle.com/datasets/kunal28chaturvedi/covid19-and-its-impact-on-students>

## **1.4 About Data Set**

In this data, a cross-sectional survey is conducted with a sample size of 1182 students of different age groups from different educational institutions in Delhi National Capital Region (NCR).

## **1.5 The details of the data set are given below:**

The data set is in csv (comma delimited) format. The data is hosted on Kaggle. This data contains student responses on the impact on their studies due to Covid-19.

## 1.6 Screenshots

ID	Region of residence	Age of Subject	Time spent on Online Class	Rating of Online Class experience	Medium for online class	Time spent on self-study	Preferred : Time spent on self-study	Number of hours spent on self-study	Change in Health issues	Stress bus	Time utili:	Do you fir	What you miss the most	
R1	Delhi-NCR	21	2 Good	Laptop/Desktop	4	0	7	3 LinkedIn	1	4 Increased	NO	Cooking YES	YES	School/college
R2	Delhi-NCR	21	0 Excellent	Smartphone	0	2	10	3 YouTube	0	3 Decrease	NO	Scrolling t YES	NO	Roaming around freely
R3	Delhi-NCR	20	7 Very poor	Laptop/Desktop	3	0	6	2 LinkedIn	0	3 Remain	C NO	Listening t NO	YES	Travelling
R4	Delhi-NCR	20	3 Very poor	Smartphone	2	1	6	5 Instagram	0	3 Decrease	NO	Watching NO	NO	Friends , relatives
R5	Delhi-NCR	21	3 Good	Laptop/Desktop	3	1	8	3 Instagram	1	4 Remain	C NO	Social Me NO	NO	Travelling
R6	Delhi-NCR	21	0 Very poor	Smartphone	6	0	5	1 YouTube	0	1 Decrease	YES	Coding an NO	YES	School/college
R7	Delhi-NCR	19	2 Very poor	Smartphone	2	1	5	4 Instagram	0	3 Increased	NO	Watching NO	YES	Friends , relatives
R8	Outside D	19	2 Very poor	Tablet	1	1	10	5 Instagram	0	3 Increased	YES	Scrolling t NO	YES	Eating outside
R9	Delhi-NCR	21	3 Very poor	Laptop/Desktop	4	1	8	2 WhatsApp	1	3 Increased	NO	Online sui NO	NO	Friends , relatives
R10	Outside D	20	0 Very poor	Laptop/Desktop	1	0.5	8	5 Instagram	3	3 Decrease	YES	live stree NO	NO	School/college
R11	Delhi-NCR	21	3 Good	Laptop/Desktop	3	1	8	3 Instagram	1	4 Remain	C NO	Social Me NO	NO	Colleagues
R12	Delhi-NCR	21	1 Very poor	Laptop/Desktop	0	0	7	3 Instagram	1	2 Increased	YES	Watching YES	YES	Eating outside
R13	Delhi-NCR	21	3 Average	Laptop/Desktop	0	0	8	3 Instagram	0	3 Increased	YES	Listening t NO	NO	School/college
R14	Outside D	22	1 Good	Laptop/Desktop	2	0	7	0 None	0.5	2 Increased	NO	Reading YES	YES	Travelling
R15	Delhi-NCR	20	5 Very poor	Laptop/Desktop	1	0	8	3 Instagram	0	3 Remain	C NO	Scrolling t NO	NO	Eating outside
R16	Delhi-NCR	22	3 Average	Smartphone	3	1	6	2 Instagram	1	3 Increased	NO	Online gai YES	YES	Eating outside
R17	Outside D	20	0 Good	Smartphone	0	0	8	2 Reddit	1	3 Increased	NO	Online gai NO	YES	Colleagues
R18	Delhi-NCR	20	1 Very poor	Smartphone	2	1	8	1 YouTube	0	3 Remain	C NO	Listening t NO	NO	Roaming around freely
R19	Delhi-NCR	21	0 Very poor	Laptop/Desktop	4	0	7	7 YouTube	1	2 Remain	C NO	Reading b NO	NO	Friends , relatives
R20	Delhi-NCR	22	5 Very poor	Smartphone	2	1	8	4 Instagram	0	3 Increased	NO	Dancing YES	YES	Friends , relatives
R21	Delhi-NCR	20	4 Excellent	Laptop/Desktop	5	0	6	2 LinkedIn	1	2 Decrease	NO	Listening t YES	YES	Roaming around freely
R22	Delhi-NCR	21	5 Average	Laptop/Desktop	2	1	10	5 Instagram	2	3 Remain	C NO	Talking w/ YES	NO	School/college
R23	Delhi-NCR	21	4 Very poor	Smartphone	1	0	8	6 Instagram	1	4 Decrease	NO	Watching NO	YES	School/college
R24	Delhi-NCR	21	4 Average	Laptop/Desktop	2	1	7	1 LinkedIn	0	3 Decrease	NO	Dancing YES	YES	Travelling
R25	Delhi-NCR	20	3 Very poor	Smartphone	4	2	9	2 Instagram	1	3 Decrease	NO	Reading b YES	NO	School/college
R26	Delhi-NCR	20	5 Very poor	Smartphone	2	0	9	5 LinkedIn	3	2 Increased	NO	Talking to YES	YES	Eating outside
R27	Delhi-NCR	21	4 Very poor	Laptop/Desktop	5	0	8	3 Instagram	0	2 Increased	YES	Listening t NO	YES	School/college
R28	Delhi-NCR	20	3 Very poor	Laptop/Desktop	6	0	9	2 WhatsApp	0	2 Remain	C NO	Listening t NO	NO	Roaming around freely

## 1.7 Attribute Information

Attribute	Description
ID	Unique id of each student
Region of residence	Name of the region
Age of Subject	Age of subjects
Time spent on Online Class	Time spent daily on online classes
Rating of Online Class experience	Rating given by all students about their experience about online classes.
Medium for online class	This contains the medium used to attend online class.
Time spent on self-study	Time spent on self-study by subject per day.

Time spent on fitness	Time spent on fitness by subject per day.
-----------------------	---

Time spent on sleep	Time spent on sleep by subject per day.
Time spent on social media	Time spent on social media by subject per day.
Preferred social media platform	Preferred social media platform of each subject
Time spent on TV	Time spent on television by subject per day.
Number of meals per day	Number of meals per day of each student
Change in your weight	Does the weight of subject change during lockdown
Health issue during lockdown	Does subject have any health issue during lockdown
Stress busters	Stress buster of each subject
Time utilized	Does the time utilize by the student in this period
Do you find yourself more connected with your family, close friends, relatives?	This contains information if student find themselves more connected to their close people.
What you miss the most	This contains information what the subject missed the most during this interval

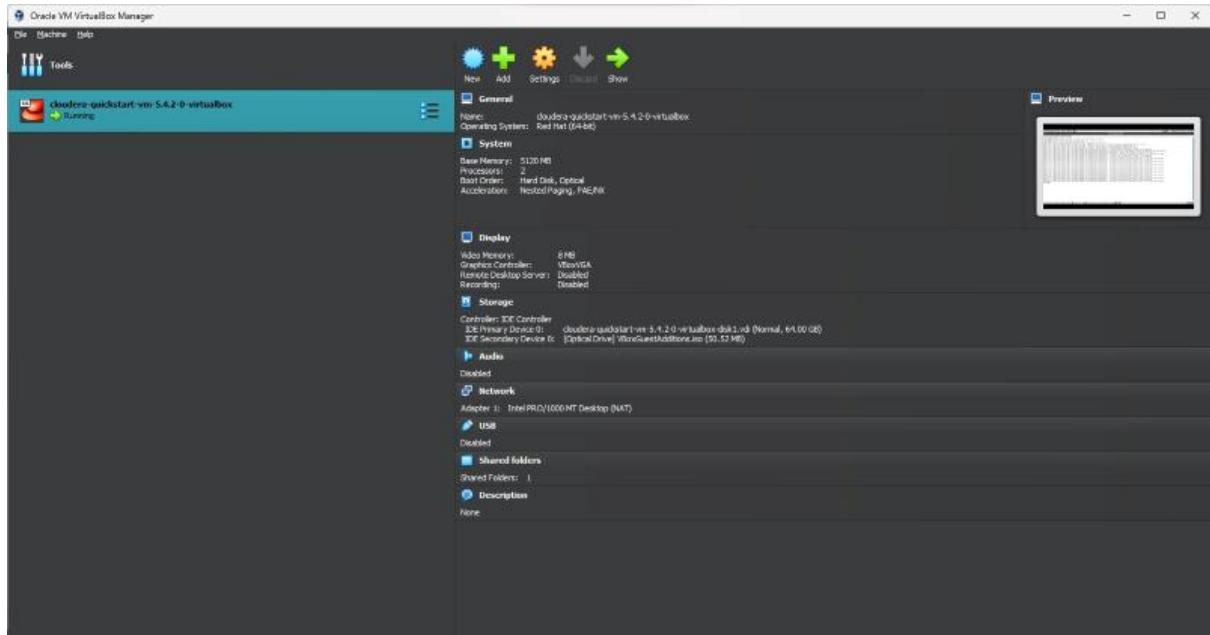
## 2. TOOLS AND WORKING ENVIRONMENT:

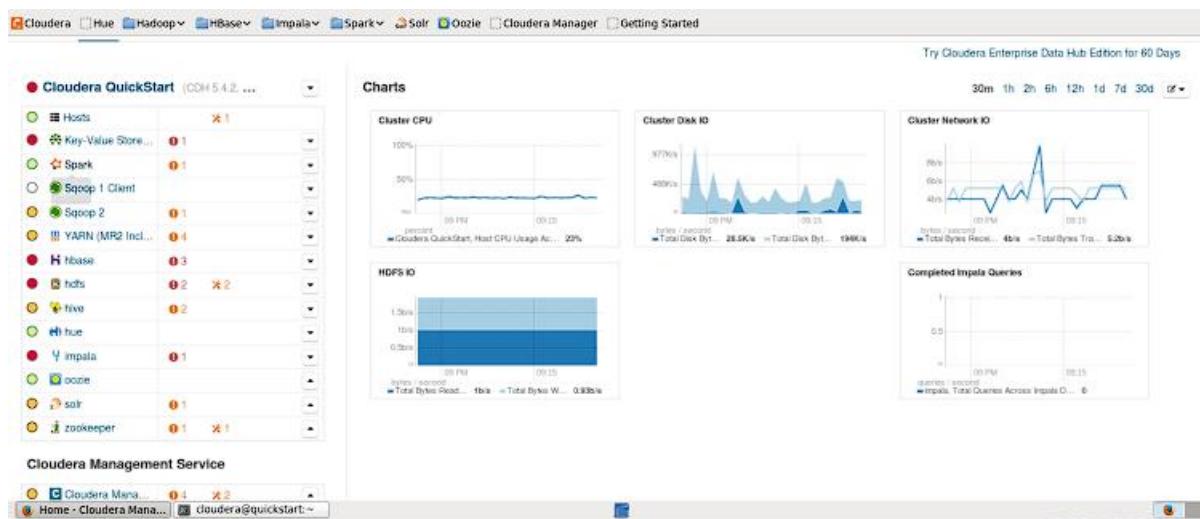
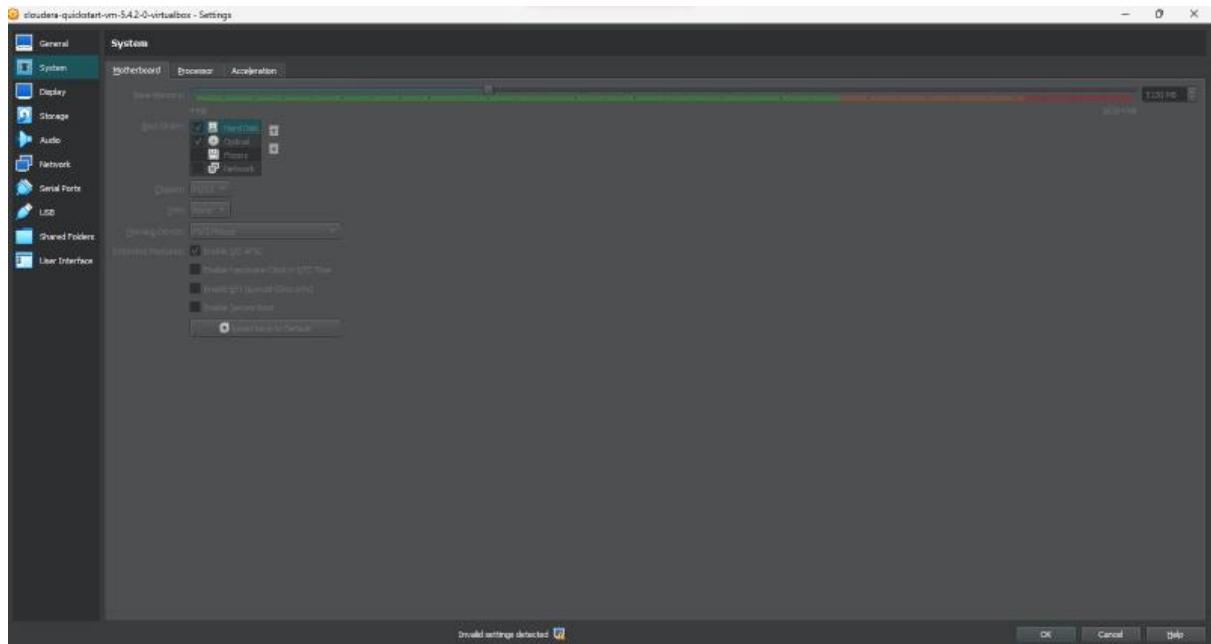
1. HDFS (Hadoop Distributed File System)
2. Sqoop
3. Hive
4. Pig

### 2.1 Assumptions and Dependencies

1. VM Ware - This software is needed for accessing a cloud and work on data using the tools HDFS, Sqoop, Hive and Pig.
2. Cloudera - This is a cloud data platform with essential tools for data analysis.

### 2.2 Installation:





## INSTALLATION OF FILE TRANSFER FROM LOCAL TO CLOUDERA:

```
VirtualBox Guest Additions installation
File Edit View Search Terminal Help
Verifying archive integrity... 100% MD5 checksums are OK. All good.
Uncompressing VirtualBox 7.0.4 Guest Additions for Linux 100%
VirtualBox Guest Additions installer
Removing installed version 4.3.10 of VirtualBox Guest Additions...
VirtualBox Guest Additions: Starting.
VirtualBox Guest Additions: Setting up modules
VirtualBox Guest Additions: Building the VirtualBox Guest Additions kernel
modules. This may take a while.
VirtualBox Guest Additions: To build modules for other installed kernels, run
VirtualBox Guest Additions: /sbin/recvboxadd quicksetup <version>
VirtualBox Guest Additions: or
VirtualBox Guest Additions: /sbin/recvboxadd quicksetup all
VirtualBox Guest Additions: Building the modules for kernel
2.6.32-358.el6.x86_64.
VirtualBox Guest Additions: Running kernel modules will not be replaced until
the system is restarted
Setting up modules
VirtualBox Guest Additions kernel modules 7.0.4 r154605 are already available
for kernel 2.6.32-358.el6.x86_64 and do not require to be rebuilt.
Running kernel modules will not be replaced until the system is restarted
vboxadd-service.sh: Starting VirtualBox Guest Addition service.
Press Return to close this window...
```

### 3. Coding On Tools:

#### 3.1. HDFS

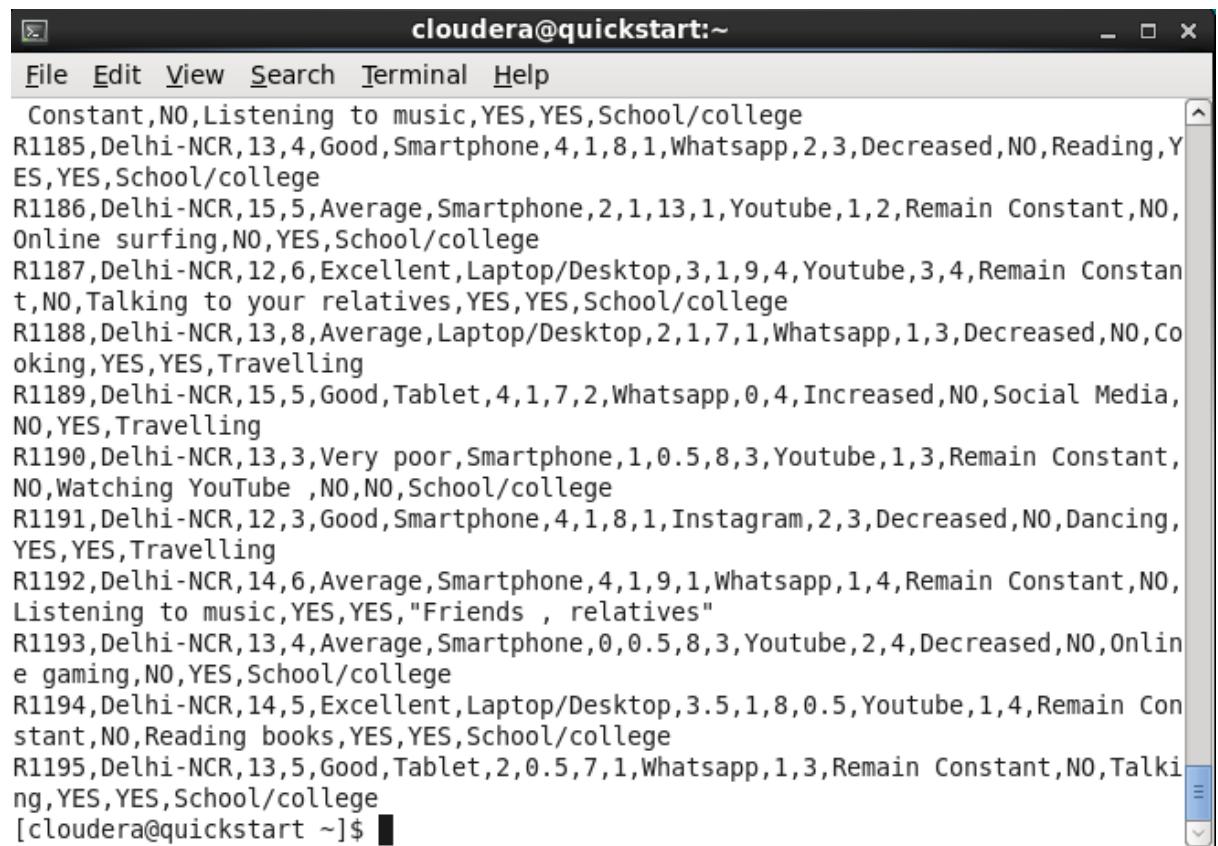
HDFS stands for Hadoop Distributed File System. By using HDFS we will run commands to move the dataset to Hadoop. So, the commands snapshots: -

1. Loading data into hdfs:

In directory name BDPOC Survey.csv file is imported using hdfs commands

```
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -mkdir BDPOC;
[cloudera@quickstart ~]$ hadoop fs -put '/home/cloudera/Desktop/Survey.csv' '/user/cloudera/BDPOC/'
[cloudera@quickstart ~]$ hadoop fs -cat '/user/cloudera/BDPOC/Survey.csv'
```

2. Data loaded successfully in hdfs here is the output from the above command in the terminal window here all attributed can be viewed in csv format.



```
cloudera@quickstart:~$ File Edit View Search Terminal Help
Constant,NO,Listening to music,YES,YES,School/college
R1185,Delhi-NCR,13,4,Good,Smartphone,4,1,8,1,Whatsapp,2,3,Decreased,NO,Reading,YES,YES,School/college
R1186,Delhi-NCR,15,5,Average,Smartphone,2,1,13,1,Youtube,1,2,Remain Constant,NO,Online surfing,NO,YES,School/college
R1187,Delhi-NCR,12,6,Excellent,Laptop/Desktop,3,1,9,4,Youtube,3,4,Remain Constant,NO,Talking to your relatives,YES,YES,School/college
R1188,Delhi-NCR,13,8,Average,Laptop/Desktop,2,1,7,1,Whatsapp,1,3,Decreased,NO,Cooking,YES,YES,Travelling
R1189,Delhi-NCR,15,5,Good,Tablet,4,1,7,2,Whatsapp,0,4,Increased,NO,Social Media,NO,YES,Travelling
R1190,Delhi-NCR,13,3,Very poor,Smartphone,1,0.5,8,3,Youtube,1,3,Remain Constant,NO,Watching YouTube ,NO,NO,School/college
R1191,Delhi-NCR,12,3,Good,Smartphone,4,1,8,1,Instagram,2,3,Decreased,NO,Dancing,YES,YES,Travelling
R1192,Delhi-NCR,14,6,Average,Smartphone,4,1,9,1,Whatsapp,1,4,Remain Constant,NO,Listening to music,YES,YES,"Friends , relatives"
R1193,Delhi-NCR,13,4,Average,Smartphone,0,0.5,8,3,Youtube,2,4,Decreased,NO,Online gaming,NO,YES,School/college
R1194,Delhi-NCR,14,5,Excellent,Laptop/Desktop,3.5,1,8,0.5,Youtube,1,4,Remain Constant,NO,Reading books,YES,YES,School/college
R1195,Delhi-NCR,13,5,Good,Tablet,2,0.5,7,1,Whatsapp,1,3,Remain Constant,NO,Talking,YES,YES,School/college
[cloudera@quickstart ~]$
```

3. LOADED FILE CAN BE VIEWED IN DIRECTORY AND TERMINAL:

```
[cloudera@quickstart ~]$ hdfs dfs -ls
Found 3 items
drwxr-xr-x  - cloudera cloudera          0 2023-03-16 01:44 13ar
drwxr-xr-x  - cloudera cloudera          0 2023-03-16 01:46 BDPOC
drwxr-xr-x  - cloudera cloudera          0 2023-03-16 01:39 POCBD
[cloudera@quickstart ~]$
```

### Browse Directory

Browse Directory						
/user/cloudera/BDPOC						Go!
Permission	Owner	Group	Size	Replication	Block Size	Name
-rw-r--r--	cloudera	cloudera	141.39 KB	1	128 MB	Survey.csv

Hadoop, 2014.

## 3.2. Sqoop:

### To Load Data into Sqoop.

1. Open the terminal.
2. Write the Sqoop command on the terminal.
3. This will open the shell.
4. Now you need to create the table in SQL and insert data into it from Sqoop.

1. Database name BDPOC is created in sqoop and can be viewed:

```
[cloudera@quickstart ~]$ mysql -uroot -pcloudera
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 62
Server version: 5.1.66 Source distribution

Copyright (c) 2000, 2012, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> create database BDPOC;
Query OK, 1 row affected (0.01 sec)

mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| BDPOC |
| firehose |
| hue |
| metastore |
| mysql |
| nav |
| navms |
| oozie |
| retail_db |
| rman |
| sentry |
+-----+
13 rows in set (0.01 sec)

mysql> █
```

2. Then table is created in the database and values are inserted in the table

```
mysql> use inpod;
Database changed
mysql> create table table(name varchar(35), age int, class varchar(35));
ERROR 1064 (42000): You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near 'table(name varchar(35), age int, class var
char(35))' at line 1
mysql> create table student(name varchar(35),age int,enrolledbatch varchar(35));
Query OK, 0 rows affected (0.14 sec)

mysql> insert into student("Aryan","20","JKLU20");
ERROR 1064 (42000): You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near ''"Aryan","20","JKLU20"' at line 1
mysql> insert into student("Aryan","20","JKLU20");
ERROR 1064 (42000): You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near ''"Aryan","20","JKLU20"' at line 1
Query OK, 1 row affected (0.02 sec)

mysql> insert into student values("Piyush","19","JKLU20");
Query OK, 1 row affected (0.05 sec)

mysql> insert into student values("Dhawal","21","JKLU20");
Query OK, 1 row affected (0.06 sec)
```

3. Description of the table:

```
mysql> desc student;
+-----+-----+-----+-----+-----+
| Field | Type   | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| name  | varchar(35) | YES |   | NULL    |       |
| age   | int(11)    | YES |   | NULL    |       |
| enrolledbatch | varchar(35) | YES |   | NULL    |       |
+-----+-----+-----+-----+-----+
3 rows in set (0.05 sec)
```

**4. After creating the database and table we will first connect with Cloudera sqoop with our local directory:**

```
[cloudera@quickstart ~]$ hostname -f
quickstart.cloudera
[cloudera@quickstart ~]$ sqoop list-databases --connect jdbc:mysql://localhost -username root --password cloudera;
Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
No such sqoop tool: list-databases. See 'sqoop help'.
[cloudera@quickstart ~]$ sqoop list-databases --connect jdbc:mysql://localhost -username root --password cloudera;
Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
23/03/19 22:03:52 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.2
23/03/19 22:03:52 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
23/03/19 22:03:53 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
information_schema
BDPOC
cm
firehose
hue
inpoc
metastore
mysql
nav
navms
oozie
retail_db
rman
sentry
[cloudera@quickstart ~]$
```

**5. Listing table in sqoop:**

```
[cloudera@quickstart ~]$ sqoop list-tables --connect jdbc:mysql://localhost/BDPOC -username root --password cloudera;
Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
23/03/19 21:10:39 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.2
23/03/19 21:10:39 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P .
23/03/19 21:10:40 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
student
```

**6. After this, we will load the data by importing the table into the Sqoop cluster.**

```
[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://localhost/BDPOC -username root --password cloudera --table student --m 1;
Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
23/03/19 21:13:47 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.2
23/03/19 21:13:47 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
23/03/19 21:13:51 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
23/03/19 21:13:51 INFO tool.CodeGenTool: Beginning code generation
23/03/19 21:13:52 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `student` AS t LIMIT 1
23/03/19 21:13:52 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `student` AS t LIMIT 1
23/03/19 21:13:52 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce
Note: /tmp/sqoop-cloudera/compile/498f7a583dcc2059be2b158ca92a11cd/student.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
23/03/19 21:14:16 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/498f7a583dcc2059be2b158ca92a11cd/student.jar
23/03/19 21:14:16 WARN manager.MySQLManager: It looks like you are importing from mysql.
23/03/19 21:14:16 WARN manager.MySQLManager: This transfer can be faster! Use the -direct
23/03/19 21:14:16 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
23/03/19 21:14:16 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
23/03/19 21:14:16 INFO mapreduce.ImportJobBase: Beginning import of student
23/03/19 21:14:23 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
23/03/19 21:14:29 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
23/03/19 21:14:31 INFO client.RMProxy: Connecting to ResourceManager at quickstart.cloudera/10.0.2.15:8032
23/03/19 21:14:45 INFO db.DBInputFormat: Using read committed transaction isolation
23/03/19 21:14:46 INFO mapreduce.JobSubmitter: number of splits:1
23/03/19 21:14:47 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1679285152671_0001
23/03/19 21:14:55 INFO impl.YarnClientImpl: Submitted application application_1679285152671_0001
23/03/19 21:14:55 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application_1679285152671_0001/
23/03/19 21:14:55 INFO mapreduce.Job: Running job: job_1679285152671_0001
23/03/19 21:15:49 INFO mapreduce.Job: Job job_1679285152671_0001 running in uber mode : false
23/03/19 21:15:49 INFO mapreduce.Job: map 0% reduce 0%
23/03/19 21:16:24 INFO mapreduce.Job: map 100% reduce 0%
23/03/19 21:16:25 INFO mapreduce.Job: Job job_1679285152671_0001 completed successfully
23/03/19 21:16:25 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=135428
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
```

```

23/03/19 21:16:25 INFO mapreduce.Job: Job job_1679285152671_0001 completed successfully
23/03/19 21:16:25 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=135428
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=87
    HDFS: Number of bytes written=50
    HDFS: Number of read operations=4
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
  Job Counters
    Launched map tasks=1
    Other local map tasks=1
    Total time spent by all maps in occupied slots (ms)=3081728
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=24076
    Total vcore-seconds taken by all map tasks=24076
    Total megabyte-seconds taken by all map tasks=3081728
  Map-Reduce Framework
    Map input records=3
    Map output records=3
    Input split bytes=87
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=1669
    CPU time spent (ms)=870
    Physical memory (bytes) snapshot=103182336
    Virtual memory (bytes) snapshot=718426112
    Total committed heap usage (bytes)=47710208
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=50
23/03/19 21:16:25 INFO mapreduce.ImportJobBase: Transferred 50 bytes in 116.388 seconds (0.4296 bytes/sec)
23/03/19 21:16:25 INFO mapreduce.ImportJobBase: Retrieved 3 records.

```

7. This is the table that we loaded into Sqoop cluster:

```

[cloudera@quickstart ~]$ hadoop fs -ls
Found 7 items
drwx-----  - cloudera cloudera      0 2023-03-19 21:16 .staging
drwxr-xr-x  - cloudera cloudera      0 2023-03-16 01:44 13ar
drwxr-xr-x  - cloudera cloudera      0 2023-03-16 01:46 BDPOC
drwxr-xr-x  - cloudera cloudera      0 2023-03-16 01:39 POCBD
drwxr-xr-x  - cloudera cloudera      0 2023-03-17 02:02 bhavishidata
drwxr-xr-x  - cloudera cloudera      0 2023-03-16 02:13 inpoc
drwxr-xr-x  - cloudera cloudera      0 2023-03-19 21:16 student

```

8. These are the values that we created in our local storage and now it is loaded into the sqoop cluster.

```

[cloudera@quickstart ~]$ hadoop fs -ls student
Found 2 items
-rw-r--r--  1 cloudera cloudera      0 2023-03-19 21:16 student/_SUCCESS
-rw-r--r--  1 cloudera cloudera     50 2023-03-19 21:16 student/part-m-000
[cloudera@quickstart ~]$ hadoop fs -cat /user/cloudera/student/part*/
Aryan,20,JKLU20
Piyush,19,JKLU20
Dhawal,21,JKLU20
[cloudera@quickstart ~]$ █

```

### 3.3. HIVE:

Load Data into hive environments.

- Open the terminal.
- Write the Hive command on the terminal.
- This will open the hive shell:-

```
[cloudera@quickstart ~]$ hive
Logging initialized using configuration in jar:file:/usr/jars/hive-common-1.1.0-cdh5.4.2.jar!/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> █
```

#### 1. Create table in hive:

```
hive> create table survey(id string ,Region_of_residence string ,Age_of_subject int ,Time_spent_on_Online_Class float ,Rating_of_Online_Class_experience string ,Medium_for_online_class string ,Time_spent_on_self_study float ,Time_spent_on_fitness float ,Time_spent_on_sleep float ,Time_spent_on_social_media float ,Preferred_social_media_platform string ,Time_spent_on_TV float ,Number_of_meals_per_day int ,Change_in_your_weight string ,Health_issue_during_lockdown string ,Stress_busters string ,Time_utilized boolean ,Do_you_find_yourself_more_connected_with_your_family_close_friend_relatives boolean ,What_you_miss_the_most string)row format delimited fields terminated by ',' stored as textfile;
OK
Time taken: 3.792 seconds
hive> █
```

#### 2. Description of the table:

```
hive> desc survey;
OK
id                      string
region_of_residence      string
age_of_subject           int
time_spent_on_online_class    float
rating_of_online_class_experience   string
medium_for_online_class string
time_spent_on_self_study     float
time_spent_on_fitness      float
time_spent_on_sleep        float
time_spent_on_social_media  float
preferred_social_media_platform string
time_spent_on_tv            float
number_of_meals_per_day    int
change_in_your_weight      string
health_issue_during_lockdown string
stress_busters             string
time_utilized              boolean
do_you_find_yourself_more_connected_with_your_family_close_friend_relatives   boolean
what_you_miss_the_most     string
Time taken: 0.379 seconds, Fetched: 19 row(s)
hive> █
```

#### 3. Now importing Survey table into hive table:

```
hive> load data local inpath '/home/cloudera/Desktop/Survey.csv' overwrite into table survey;
Loading data to table default.survey
Table default.survey stats: [numFiles=1, numRows=0, totalSize=144782, rawDataSize=0]
OK
Time taken: 0.709 seconds
hive> █
```

#### 4. Table is imported into database and can be viewed in terminal:

```

hive> select * from survey limit 10;
OK
ID      Region of residence    NULL    NULL    Rating of Online Class experience
e in your weight          Health issue during lockdown    Stress busters    NULL    NULL
R1      Delhi-NCR           21     2.0     Good    Laptop/Desktop 4.0   0.0    7.0    3.0    LinkedIn 1.0   4    Increased    NO    Cooking    NULL    NULL    School/college    NULL    NULL    Chang
R2      Delhi-NCR           21     0.0     Excellent  Smartphone  0.0    2.0    10.0   3.0    Youtube 0.0  3    Decreased    NO    Scrolling through social media    NULL    NULL    NULL
R3      Delhi-NCR           20     7.0     Very poor  Laptop/Desktop 3.0   0.0    6.0    2.0    LinkedIn 0.0   3    Remain Constant    NO    Listening to music    NULL    NULL    NULL
R4      Travelling Friends  Delhi-NCR           20     3.0     Very poor  Smartphone  2.0    1.0    6.0    5.0    Instagram 0.0   3    Decreased    NO    Watching web series    NULL    NULL    NULL
R5      Delhi-NCR           21     3.0     Good    Laptop/Desktop 3.0   1.0    8.0    3.0    Instagram 1.0   4    Remain Constant    NO    Social Media    NULL    NULL    Travelling
R6      Delhi-NCR           21     0.0     Very poor  Smartphone  6.0    0.0    5.0    1.0    Youtube 0.0  1    Decreased    YES    Coding and studying for exams    NULL    NULL    NULL
R7      Delhi-NCR           19     2.0     Very poor  Smartphone  2.0    1.0    5.0    4.0    Instagram 0.0   3    Increased    NO    Watching web series    NULL    NULL    NULL
R8      Outside Delhi-NCR  19     2.0     Very poor  Tablet    1.0    1.0    10.0   5.0    Instagram 0.0   3    Increased    YES    Scrolling through social media    NULL    NULL    NULL
R9      Eating outside      Delhi-NCR           21     3.0     Very poor  Laptop/Desktop 4.0   1.0    8.0    2.0    Whatsapp 1.0   3    Increased    NO    Online surfing    NULL    NULL    "Frie
Time taken: 0.437 seconds, Fetched: 10 row(s)
hive> ■

```

## QUERIES ON HIVE:

### 1. Retrieve all students which have health issue during the lockdown period:

```

hive> SELECT ID,Health_issue_during_lockdown FROM survey WHERE Health_issue_during_lockdown IS NOT NULL and Health_issue_during_lockdown="YES";
OK
R6      YES
R8      YES
R10     YES
R12     YES
R13     YES
R27     YES
R31     YES
R34     YES
R43     YES
R48     YES
R51     YES
R56     YES
R59     YES
R69     YES
R76     YES
R81     YES
R87     YES
R89     YES
R91     YES
R93     YES
R99     YES
R103    YES
R112    YES
R125    YES
R141    YES
R147    YES
R149    YES
R155    YES
R156    YES
R160    YES
R162    YES
R164    YES
R165    YES
R171    YES
R172    YES
R176    YES
R183    YES

```

```
R826    YES
R829    YES
R830    YES
R837    YES
R853    YES
R865    YES
R872    YES
R878    YES
R890    YES
R891    YES
R895    YES
R904    YES
R907    YES
R934    YES
R937    YES
R938    YES
R947    YES
R950    YES
R963    YES
R970    YES
R971    YES
R974    YES
R992    YES
R995    YES
R998    YES
R1009   YES
R1012   YES
R1020   YES
R1023   YES
R1028   YES
R1035   YES
R1041   YES
R1047   YES
R1051   YES
R1060   YES
R1082   YES
R1126   YES
R1151   YES
Time taken: 0.186 seconds, Fetched: 161 row(s)
hive> █
```

2. Retrieve count of each social media platform preferred during the lockdown period:

```

hive> SELECT Prefered_social_media_platform, COUNT(*) AS count FROM survey GROUP BY Prefered_social_media_platform ORDER BY count DESC;
Query ID = cloudera_20230320021313_6035bb3f-b1fb-4e58-809f-75664e961b0c
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-03-20 02:13:45,954 Stage-1 map = 0%,  reduce = 0%
2023-03-20 02:14:04,633 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.02 sec
2023-03-20 02:14:13,967 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 1.99 sec
MapReduce Total cumulative CPU time: 1 seconds 990 msec
Ended Job = job_1679285152671_0002
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0003
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2023-03-20 02:14:23,797 Stage-2 map = 0%,  reduce = 0%
2023-03-20 02:14:29,141 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 0.63 sec
2023-03-20 02:14:36,406 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 1.63 sec
MapReduce Total cumulative CPU time: 1 seconds 630 msec
Ended Job = job_1679285152671_0003
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 1.99 sec  HDFS Read: 154770 HDFS Write: 567 SUCCESS
Stage-Stage-2: Map: 1  Reduce: 1  Cumulative CPU: 1.63 sec  HDFS Read: 4959 HDFS Write: 203 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 620 msec

```

## RESULT:

OK	
Instagram	352
Whatsapp	336
Youtube	314
Linkedin	61
Facebook	52
Twitter	28
None	17
Snapchat	8
Reddit	5
Telegram	3
Quora	1
WhatsApp	1
Talklife	1
Prefered social media platform	1
Omegle	1
None	1
Elyment	1

Time taken: 77.375 seconds, Fetched: 17 row(s)

hive> █

3. Retrieve count of each stress buster activity preferred during the lockdown period:

```

time taken: 27.777 seconds, Fetched: 1 row(s)
hive> SELECT Stress_busters, COUNT(*) AS count FROM survey GROUP BY Stress_busters;
Query ID = cloudera_20230320022525_e3057c90-14e2-4907-8a89-fe901ba86407
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0006/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0006
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-03-20 02:25:41,223 Stage-1 map = 0%, reduce = 0%
2023-03-20 02:25:47,532 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 0.84 sec
2023-03-20 02:25:56,015 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.02 sec
MapReduce Total cumulative CPU time: 2 seconds 20 msec
Ended Job = job_1679285152671_0006
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.02 sec HDFS Read: 155523 HDFS Write: 1647 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 20 msec
OK
" listening music      1
"Drawing      1
"Netflix      1
"Online gaming  1
"Painting      1
"Poetry       1
"Reading books 1
"Sleeping      1
"listening to music    1
"sketching      1
"watching movies    1
All reading books watching web series listening to music and talking to friends 1
Anime Manga     1

```

```

Reading 19
Reading books    77
Running 1
Scrolling through social media  74
Sketching      1
Sketching and writing   1
Sleep      34
Sleeping      71
Social Media    26
Sports       1
Stress busters  1
Taking with parents  1
Talk with childhood friends.  1
Talking 19
Talking to friends    1
Talking to your relatives  37
Talking with friends   1
Watching YouTube    1
Watching orgasm releasing videos    1
Watching ted talks and music and books  1
Watching web series    102
Web Series     21
Whatever want    1
With a friend    1
Work       1
Workout      1
Workout and listening music    1
Writing 2
Writing my own Comics & novels  1
Youtube 2
drawing      1
gardening cartoon    1
live stream watching  1
no stress      1
pubg        1
singing 1
working out and some physical activity  1
workout 1
Time taken: 22.048 seconds, Fetched: 87 row(s)
hive> █

```

#### 4. Retrieve the preferred social media platform with the highest usage time:

```

hive> SELECT Preferred_social_media_platform, SUM(Time_spent_on_social_media) AS total_time_social_media
    > from survey
    > GROUP BY Preferred_social_media_platform
    > ORDER BY total_time_social_media DESC
    > LIMIT 1;
Query ID = cloudera_20230410030606_640529ec-ddbd-4d30-a90e-902e2ce91ec0
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-10 03:06:51,936 Stage-1 map = 0%,  reduce = 0%
2023-04-10 03:07:15,649 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.96 sec
2023-04-10 03:07:24,020 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.07 sec
MapReduce Total cumulative CPU time: 2 seconds 70 msec
Ended Job = job_1679285152671_0008
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0009, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0009/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0009
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2023-04-10 03:07:33,019 Stage-2 map = 0%,  reduce = 0%
2023-04-10 03:08:05,166 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.25 sec
2023-04-10 03:08:13,422 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 2.21 sec
MapReduce Total cumulative CPU time: 2 seconds 210 msec
Ended Job = job_1679285152671_0009

```



```

Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-10 03:06:51,936 Stage-1 map = 0%,  reduce = 0%
2023-04-10 03:07:15,649 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.96 sec
2023-04-10 03:07:24,020 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.07 sec
MapReduce Total cumulative CPU time: 2 seconds 70 msec
Ended Job = job_1679285152671_0008
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0009, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0009/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0009
Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1
2023-04-10 03:07:33,019 Stage-2 map = 0%,  reduce = 0%
2023-04-10 03:08:05,166 Stage-2 map = 100%,  reduce = 0%, Cumulative CPU 1.25 sec
2023-04-10 03:08:13,422 Stage-2 map = 100%,  reduce = 100%, Cumulative CPU 2.21 sec
MapReduce Total cumulative CPU time: 2 seconds 210 msec
Ended Job = job_1679285152671_0009
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 2.07 sec  HDFS Read: 154795 HDFS Write: 672 SUCCESS
Stage-Stage-2: Map: 1  Reduce: 1  Cumulative CPU: 2.21 sec  HDFS Read: 5202 HDFS Write: 29 SUCCESS
Total MapReduce Time Spent: 4 seconds 280 msec
OK
Instagram      1025.0999999716878
Time taken: 115.779 seconds, Fetched: 1 row(s)

```

5. Retrieve the number of students who reported a change in weight:

```

hive> SELECT COUNT(*) AS num_stud
    > FROM survey
    > WHERE Change_in_your_weight IS NOT NULL;
Query ID = cloudera_20230410031919_e5b6f971-cfc9-4c86-a02a-7804e889d9c1
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0010
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-10 03:20:04,043 Stage-1 map = 0%, reduce = 0%
2023-04-10 03:20:30,189 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.45 sec
2023-04-10 03:20:37,366 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.48 sec
MapReduce Total cumulative CPU time: 2 seconds 480 msec
Ended Job = job_1679285152671_0010
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.48 sec HDFS Read: 155840 HDFS Write: 5 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 480 msec
OK
1183
Time taken: 45.053 seconds, Fetched: 1 row(s)
hive> ■

```

---

6. Retrieve the total time spent on fitness activities for students who rated online class experience as "Excellent":

```

hive> SELECT SUM(time_spent_on_fitness) AS total_TT
    > FROM survey
    > WHERE rating_of_online_class_experience= 'Excellent';
Query ID = cloudera_20230410032626_70e6663c-b20f-4522-ac00-8ee49bbb8b66
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0011, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0011/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0011
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-10 03:26:10,747 Stage-1 map = 0%, reduce = 0%
2023-04-10 03:26:39,805 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.61 sec
2023-04-10 03:26:48,317 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.78 sec
MapReduce Total cumulative CPU time: 2 seconds 780 msec
Ended Job = job_1679285152671_0011
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.78 sec HDFS Read: 156014 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 780 msec
OK
97.85000002384186
Time taken: 46.778 seconds, Fetched: 1 row(s)

```

7. Retrieve the average time spent on self-study for students below the age of 18:

```
hive> SELECT AVG(time_spent_on_self_study) AS avg_time
    > FROM survey
    > WHERE age_of_subject < 18;
Query ID = cloudera_20230410033030_1307f644-b150-4023-8d61-d52e7c6f25a9
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0012, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0012/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0012
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-10 03:30:54,499 Stage-1 map = 0%, reduce = 0%
2023-04-10 03:31:10,019 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.45 sec
2023-04-10 03:31:19,418 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.71 sec
MapReduce Total cumulative CPU time: 2 seconds 710 msec
Ended Job = job_1679285152671_0012
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.71 sec HDFS Read: 156413 HDFS Write: 19 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 710 msec
OK
2.8294701986754967
Time taken: 34.335 seconds, Fetched: 1 row(s)
```

8. Retrieve the number of students who spent more than 3 hours on TV and reported a health issue:

```
hive> SELECT COUNT(*) AS num_stud_tvhealth
    > FROM survey
    > WHERE time_spent_on_tv > 3 AND health_issue_during_lockdown="YES";
Query ID = cloudera_20230410035353_3893d13b-dc4f-4dd3-8b24-fe6274451479
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0014, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0014/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0014
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-10 03:53:30,730 Stage-1 map = 0%, reduce = 0%
2023-04-10 03:53:39,230 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.28 sec
2023-04-10 03:53:46,483 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.46 sec
MapReduce Total cumulative CPU time: 2 seconds 460 msec
Ended Job = job_1679285152671_0014
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 2.46 sec HDFS Read: 156398 HDFS Write: 2 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 460 msec
OK
6
Time taken: 24.517 seconds, Fetched: 1 row(s)
```

9. Retrieve the total time utilized by students, calculated as the sum of time spent on online classes, self-study, fitness activities, and sleep:

```

hive> SELECT (SUM(time_spent_on_self_study)/SUM(time_spent_on_online_class + time_spent_on_fitness
s + time_spent_on_social_media + time_spent_on_tv)) * 100 as studpercnt
    > FROM survey;
Query ID = cloudera_20230411230606_4d6a3b9e-bac9-4b1f-9ee7-90386c2a3a8a
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0015, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0015
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0015
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-11 23:06:41,551 Stage-1 map = 0%,  reduce = 0%
2023-04-11 23:07:22,224 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 1.68 sec
2023-04-11 23:07:31,456 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 3.38 sec
MapReduce Total cumulative CPU time: 3 seconds 380 msec
Ended Job = job_1679285152671_0015
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 3.38 sec  HDFS Read: 157511 HDFS Write: 18 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 380 msec
OK
40.01976857677319
Time taken: 67.209 seconds, Fetched: 1 row(s)

```

10. Retrieve the number of students who reported a health issue and spent more time on fitness activities than on online classes:

```

hive> SELECT COUNT(*) AS num_stud
    > FROM survey
    > WHERE health_issue_during_lockdown= 'YES' AND time_spent_on_fitness > time_spent_on_online_
class
    > ;
Query ID = cloudera_20230411231616_f7694562-64a9-4262-beda-f9294fd99bbc
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1679285152671_0016, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1679285152671_0016
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1679285152671_0016
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2023-04-11 23:16:09,514 Stage-1 map = 0%,  reduce = 0%
2023-04-11 23:16:24,486 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 0.9 sec
2023-04-11 23:16:34,256 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 2.15 sec
MapReduce Total cumulative CPU time: 2 seconds 150 msec
Ended Job = job_1679285152671_0016
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1  Cumulative CPU: 2.15 sec  HDFS Read: 155223 HDFS Write: 5 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 150 msec
OK
1183
Time taken: 33.062 seconds, Fetched: 1 row(s)

```

## Queries on PIG:

Load Data into Pig environments.

1. Open the terminal.
2. Write the Pig Command (`pig -x local`) on the terminal.
3. This will open the grunt shell in the local mode.
4. Now you need to create a table and load data into it from the local file system.

Pig in terminal:

```
[cloudera@quickstart ~]$ pig -x local
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
2023-04-12 02:23:54 [main] INFO org.apache.pig.Main - Apache Pig version 0.12.0-cdh5.4.2 (exported) compiled May 19 2015, 17:03:41
2023-04-12 02:23:54,550 [main] INFO org.apache.pig.Main - Logging error messages to: /home/cloudera/pig_1681291434506.log
2023-04-12 02:23:54,638 [main] INFO org.apache.pig.impl.util.Utils - Default bootup file /home/cloudera/.pigbootup not found
2023-04-12 02:23:54,960 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:54,960 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2023-04-12 02:23:54,961 [main] INFO org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop file system at: file:///
2023-04-12 02:23:55,336 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,340 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,454 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,456 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,507 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,509 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,546 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,547 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,594 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,595 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,620 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,621 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,655 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,656 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,689 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,691 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 02:23:55,736 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 02:23:55,747 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
grunt> ■
```

### 1. Create table in pig :

```
grunt> Survey= LOAD 'Desktop/Survey.csv' USING PigStorage(',') AS (id: chararray, region_of_residence: chararray, age_of_subject: int, time_spent_on_online_class: float, rating_of_online_class_experience: chararray, medium_for_online_class: chararray, time_spent_on_study: float, time_spent_on_fitness: float, time_spent_on_sleep: float, time_spent_on_social_media: float, preferred_social_media_platform: chararray, time_spent_on_tv: float, number_of_meals_per_day: int, change_in_your_weight: chararray, health_issue_during_lockdown: chararray, stress_busters: chararray, time_utilized: boolean, do_you_find_yourself_more_connected_to_family: boolean, what_you_miss_the_most: chararray);■
```

### 2. Created table can be viewed in terminal:

```
(R727,Outside Delhi-NCR,18,3.0,Good,Laptop/Desktop,3.0,0.0,10.0,2.0,Instagram,0.0,2,Increased,NO,Listening to music,,,Colleagues)
(R728,Delhi-NCR,20,5.0,Very poor,Laptop/Desktop,2.0,3.0,10.0,1.0,Youtube,0.5,2,Decreased,YES,Writing my own Comics & novels,,,Playing)
(R729,Delhi-NCR,20,2.0,Very poor,Smartphone,6.0,1.0,5.0,4.0,Youtube,0.5,1,Increased,YES,Listening to music,,,Travelling)
(R730,Delhi-NCR,23,0.0,Average,Laptop/Desktop,1.0,1.0,9.0,0.3,LinkedIn,0.0,3,Increased,NO,Listening to music,,,Travelling)
(R731,Outside Delhi-NCR,29,2.0,Average,Laptop/Desktop,2.0,1.0,8.0,2.0,Youtube,0.0,2,Remain Constant,NO,Reading books,,,nothing)
(R732,Delhi-NCR,28,0.0,Very poor,Laptop/Desktop,6.0,0.0,9.0,1.0,Whatsapp,1.0,3,Increased,NO,Sleeping,,,Travelling)
(R733,Outside Delhi-NCR,22,1.0,Very poor,Smartphone,4.0,0.25,10.0,1.0,LinkedIn,0.0,3,Remain Constant,NO,Sleeping,...)
(R734,Delhi-NCR,18,2.0,Average,Laptop/Desktop,3.0,1.0,8.0,2.0,Youtube,3.0,3,Increased,NO,Online surfing,,,School/college)
(R735,Delhi-NCR,15,4.0,Good,Smartphone,4.0,1.0,8.0,3.0,Youtube,2.0,3,Remain Constant,NO,Sleeping,,,Colleagues)
(R736,Delhi-NCR,19,4.0,Average,Smartphone,5.0,1.0,8.0,2.0,Youtube,1.0,3,Remain Constant,NO,Meditation,,,Travelling)
(R737,Delhi-NCR,20,2.0,Very poor,Smartphone,2.0,0.0,9.0,10.0,Youtube,1.0,2,Decreased,NO,Online gaming,,,Nothing)
(R738,Outside Delhi-NCR,21,1.0,Very poor,Smartphone,4.0,1.0,9.0,0.1,Whatsapp,1.0,3,Remain Constant,NO,Watching web series,,,Eating outside)
(R739,Outside Delhi-NCR,21,4.0,Average,Laptop/Desktop,3.0,0.0,8.0,2.0,Youtube,1.0,3,Decreased,NO,Online gaming,,,Roaming around freely)
(R740,Delhi-NCR,20,4.0,Very poor,Smartphone,1.0,1.0,10.0,2.0,Instagram,1.0,3,Increased,YES,Dancing,,,Eating outside)
(R741,Delhi-NCR,20,2.0,Average,Smartphone,1.0,2.0,8.0,4.0,Facebook,2.0,0.4,Decreased,NO,Reading books,,,Travelling)
(R742,Outside Delhi-NCR,19,0.0,Very poor,Smartphone,2.0,1.0,7.0,4.0,Instagram,0.0,2,Remain Constant,NO,Meditation,,,Travelling)
(R743,Outside Delhi-NCR,28,3.0,Good,Laptop/Desktop,1.0,0.0,10.0,1.0,Whatsapp,0.0,2,Increased,NO,Online gaming,,,Eating outside)
(R744,Delhi-NCR,19,6.0,Good,Laptop/Desktop,6.0,0.0,8.0,2.0,Youtube,0.5,2,Increased,NO,Listening to music,,,Travelling)
(R745,Delhi-NCR,21,2.0,Average,Smartphone,2.0,1.0,8.0,1.0,Youtube,0.0,3,Remain Constant,NO,Sleeping,,,Friends )
(R746,Delhi-NCR,23,0.0,Good,Laptop/Desktop,3.0,0.0,8.0,10.0,4.0,Youtube,1.0,3,Increased,NO,Online gaming,,,Eating outside)
(R747,Delhi-NCR,20,3.0,Good,Laptop/Desktop,2.0,1.0,7.0,5.0,Whatsapp,2.0,0.3,Remain Constant,NO,Reading,,,School/college)
(R748,Delhi-NCR,22,4.0,Good,Laptop/Desktop,2.0,1.0,12.0,2.0,Instagram,2.0,3,Remain Constant,NO,Talking,,,School/college)
(R749,Delhi-NCR,20,2.0,Average,Smartphone,4.0,1.0,8.0,2.0,Youtube,0.0,3,Increased,NO,Listening to music,,,School/college)
(R750,Outside Delhi-NCR,22,0.0,NA,NA,0.0,1.0,5.0,0.4,Instagram,0.0,2,Decreased,NO,Watching web series,,,Travelling)
(R751,Outside Delhi-NCR,20,5.0,Good,Laptop/Desktop,4.0,1.0,7.0,3.0,Youtube,1.0,3,Increased,NO,Reading,,,Travelling)
(R752,Outside Delhi-NCR,23,3.0,Good,Laptop/Desktop,2.0,1.0,6.0,1.0,Whatsapp,0.0,2,Remain Constant,NO,Reading,,,Eating outside)
(R753,Outside Delhi-NCR,21,2.0,Very poor,Smartphone,2.0,1.0,8.0,4.0,Whatsapp,1.0,4,Increased,NO,Scrolling through social media,,,Travelling)
(R754,Delhi-NCR,19,3.0,Very poor,Smartphone,2.0,1.0,9.0,5.0,Instagram,1.0,3,Remain Constant,NO,Listening to music,,,Friends )
(R755,Outside Delhi-NCR,23,3.0,Good,Laptop/Desktop,5.0,0.0,6.0,1.0,Youtube,0.0,2,Remain Constant,NO,Web Series,,,School/college)
(R756,Outside Delhi-NCR,20,4.0,Average,Smartphone,2.0,0.0,8.0,2.0,Whatsapp,4.0,0.3,Remain Constant,NO,Sleeping,,,Friends )
(R757,Outside Delhi-NCR,22,2.0,Good,Laptop/Desktop,2.0,1.0,8.0,2.0,Whatsapp,0.0,4,Increased,NO,Web Series,,,Travelling)
(R758,Delhi-NCR,19,2.0,Poor,Smartphone,2.0,2.0,8.0,3.0,Youtube,0.0,4,Decreased,YES,Gym,,,Roaming around freely)
(R759,Outside Delhi-NCR,21,5.0,Very poor,Smartphone,4.0,1.0,7.0,1.0,Youtube,0.0,3,Remain Constant,NO,Meditation,,,Friends )
(R760,Outside Delhi-NCR,20,0.0,Very poor,Smartphone,5.0,0.0,8.0,1.0,Youtube,3,Decreased,YES,I have no problem of stress,,,I have missed nothing )
(R761,Outside Delhi-NCR,21,1.0,Very poor,Laptop/Desktop,5.0,1.0,11.0,2.0,Instagram,0.0,3,Increased,NO,workout,,,Eating outside)
(R762,Delhi-NCR,22,5.0,Good,Laptop/Desktop,3.0,0.0,8.0,7.0,2.0,Whatsapp,0.0,3,Remain Constant,NO,Sleeping,,,Travelling)
(R763,Delhi-NCR,18,4.0,Good,Smartphone,2.0,0.0,5.0,0.2,Whatsapp,0.0,4,Remain Constant,NO,Online gaming,,,School/college)
(R764,Delhi-NCR,21,1.0,Good,Laptop/Desktop,1.0,0.2,7.0,2.0,Youtube,0.0,3,Remain Constant,NO,Listening to music,,,Colleagues)
(R765,Outside Delhi-NCR,22,2.0,Average,Laptop/Desktop,5.0,0.3,8.0,1.0,LinkedIn,1.0,3,Increased,YES,Listening to music,,,Roaming around freely)
(R766,Out Browsing HDFS - Mozilla Firefox 3,Laptop/Desktop,3.0,1.0,9.0,0.3,Instagram,1.0,2,Increased,NO,Meditation,,,Friends )■
```



## PIG QUERY:

### 1. Filter data for subjects who spent more than 3 hours on online class.

```
grunt> filtered data= FILTER Survey BY time_spent_on_online_class > 3.0;
2023-04-12 03:06:27,275 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 2 time(s).
grunt> dump filtered_data;
```

### Answer:

```
(R50,Delhi-NCR,18,4.0,Good,Laptop/Desktop,2.0,0.0,8.0,4.0,Instagram,0.0,3,Remain Constant,NO,Social Media,,,Colleagues)
(R52,Delhi-NCR,20,5.0,Average,Laptop/Desktop,6.0,1.0,7.0,2.0,Whatsapp,0.0,3,Remain Constant,NO,Reading books,,,Roaming around freely)
(R55,Delhi-NCR,21,4.0,Average,Smartphone,4.0,1.0,8.0,2.0,Youtube,2.0,4,Remain Constant,NO,Sleep,,,Gym)
(R56,Delhi-NCR,19,4.0,Very poor,Smartphone,2.0,0.0,7.0,1.0,Whatsapp,0.0,3,Increased,YES,Watching web series,,,Roaming around freely)
(R59,Delhi-NCR,20,5.0,Average,Laptop/Desktop,1.0,1.0,7.0,1.0,Whatsapp,0.0,3,Increased,YES,Listening to music,,,School/college)
(R61,Delhi-NCR,22,6.0,Very poor,Laptop/Desktop,6.0,0.0,6.0,1.0,Instagram,0.0,2,Increased,NO,Watching web series,,,Friends )
(R66,Delhi-NCR,20,4.0,Average,Smartphone,4.0,2.0,7.0,2.0,Instagram,1.0,3,Increased,NO,Listening to music,,,Eating outside)
(R72,Delhi-NCR,22,6.0,Very poor,Smartphone,4.0,0.0,7.0,2.0,Whatsapp,0.0,1,Increased,NO,Online gaming,,,Eating outside)
(R81,Delhi-NCR,21,5.0,Average,Laptop/Desktop,5.0,0.0,8.0,7.0,Instagram,1.0,2,Increased,YES,Scrolling through social media,,,Friends )
(R87,Delhi-NCR,20,5.0,Average,Smartphone,3.0,1.0,7.0,1.5,Instagram,0.0,3,Increased,YES,drawing,,School/college)
(R88,Outside Delhi-NCR,21,4.0,Very poor,Smartphone,3.0,1.0,8.0,1.0,Linkedin,0.0,3,Decreased,NO,Meditation,,,All of them)
(R90,Outside Delhi-NCR,19,4.0,Average,Smartphone,3.0,1.0,7.0,3.0,Instagram,2.0,3,Increased,NO,Listening to music,,,School/college)
(R91,Outside Delhi-NCR,25,4.0,Very poor,Smartphone,3.0,1.0,7.0,1.0,Instagram,0.0,3,Remain Constant,YES,Cricket,,,Eating outside)
(R93,Delhi-NCR,19,4.0,Very poor,Laptop/Desktop,0.0,0.0,8.0,2.0,Instagram,1.5,3,Increased,YES,Anime Manga,,,School/college)
(R94,Delhi-NCR,20,4.0,Good,Smartphone,2.0,2.0,8.0,6.0,Instagram,0.0,3,Decreased,NO,Sleeping,,,School/college)
(R96,Delhi-NCR,19,5.0,Good,Tablet,3.0,1.0,8.0,2.0,Whatsapp,0.0,2,Remain Constant,NO,Reading,,,Travelling)
(R101,Outside Delhi-NCR,16,6.0,Poor,Smartphone,2.0,2.0,7.0,2.0,Youtube,1.0,3,Remain Constant,NO,Scrolling through social media,,,Travelling)
(R103,Outside Delhi-NCR,21,4.0,Very poor,Smartphone,4.0,1.0,7.0,1.0,Whatsapp,0.0,4,Decreased,YES,Listening to music,,,Travelling)
(R106,Outside Delhi-NCR,23,4.0,Excellent,Laptop/Desktop,6.0,1.0,8.0,0.3,Whatsapp,0.0,3,Increased,NO,Talking to your relatives,,,Travelling)
(R114,Outside Delhi-NCR,14,5.0,Average,Laptop/Desktop,4.0,1.0,7.0,4.0,Instagram,1.0,5,Increased,NO,Scrolling through social media,,,Friends )
(R117,Outside Delhi-NCR,24,4.0,Excellent,Laptop/Desktop,8.0,1.0,8.0,2.0,Youtube,2.0,3,Increased,NO,Online surfing,,,Friends )
(R119,Outside Delhi-NCR,21,4.0,Good,Laptop/Desktop,2.0,2.0,7.0,2.0,Instagram,1.0,3,Increased,NO,Dancing,,,Travelling)
(R122,Outside Delhi-NCR,23,6.0,Average,Laptop/Desktop,5.0,1.0,6.0,1.0,Linkedin,0.0,3,Remain Constant,NO,Meditation,,School/college)
(R123,Outside Delhi-NCR,21,4.0,Average,Smartphone,4.0,1.0,8.0,1.0,Youtube,0.0,3,Remain Constant,NO,Meditation,,Friends )
(R124,Delhi-NCR,18,5.0,Very poor,Smartphone,3.0,0,0,9.0,7.0,Whatsapp,0.0,3,Increased,NO,Watching web series,,,Roaming around freely)
(R126,Outside Delhi-NCR,20,5.0,Average,Smartphone,4.0,1.0,7.0,2.0,Whatsapp,1.0,3,Increased,NO,Meditation,,School/college)
(R132,Delhi-NCR,17,6.0,Average,Laptop/Desktop,4.0,1.0,8.0,1.0,Youtube,1.0,2,Remain Constant,NO,Listening to music,,,School/college)
(R135,Delhi-NCR,18,5.0,Average,Smartphone,1.0,1.0,8.0,1.0,Youtube,0.0,4,Increased,NO,Sleep,,,Roaming around freely)
(R138,Delhi-NCR,15,4.0,Average,Smartphone,2.0,0.5,6.0,1.0,Whatsapp,1.0,3,Remain Constant,NO,Dancing,,,Eating outside)
(R146,Outside Delhi-NCR,14,4.0,Average,Smartphone,2.0,1.0,8.0,2.0,Whatsapp,1.0,3,Remain Constant,NO,Listening to music,,,Friends )
(R147,Delhi-NCR,20,6.0,Average,Laptop/Desktop,1.0,1.0,6.0,1.0,Linkedin,0.0,3,Increased,YES,Talking to your relatives,,,School/college)
(R149,Delhi-NCR,19,6.0,Very poor,Smartphone,3.0,0,0,7.0,2.0,Whatsapp,0.5,2,Increased,YES,Online gaming,,,Friends )
(R152,Outside Delhi-NCR,21,4.0,Very poor,Smartphone,4.0,1.0,8.0,2.0,Instagram,0.0,3,Increased,NO,Sleep,,,Roaming around freely)
(R153,Outside Delhi-NCR,15,4.0,Average,Smartphone,7.0,2.0,6.0,1.0,Instagram,1.0,4,Remain Constant,NO,Listening to music,,,Friends )
(R154,Delhi-NCR,20,5.0,Very poor,Tablet,3.0,1.0,8.0,2.0,Whatsapp,3.0,4,Increased,NO,Listening to music,,,Roaming around freely)
(R158,Delhi-NCR,21,5.0,Average,Smartphone,1.0,0,0,5.0,1.0,Whatsapp,0.0,2,Decreased,NO,Online gaming,,,Travelling)
(R159,Delhi-NCR,21,4.0,Excellent,Laptop/Desktop,6.0,2.0,6.0,1.0,Youtube,1.0,3,Remain Constant,NO,Cooking,,,Eating outside)
(R161,Delhi-NCR,20,4.0,Average,Laptop/Desktop,2.0,0,0,7.0,2.0,Whatsapp,4.0,3,Increased,NO,Online gaming,,,Friends )
(R162,Outside Delhi-NCR,20,6.0,Very poor,Smartphone,3.0,0,0,8.0,4.0,Linkedin,2.0,3,Increased,YES,Sleep,,,School/college)
(R164,Outside Delhi-NCR,21,4.0,Very poor,Smartphone,3.0,1.0,8.0,1.0,Youtube,1.0,3,Remain Constant,YES,Online surfing,,,Friends )
(R1151,Delhi-NCR,14,5.0,Average,Smartphone or Laptop/Desktop,2.0,0.0,9.0,0.5,Whatsapp,0.5,4,Increased,YES,Listening to music,,,Travelling)
(R1153,Delhi-NCR,13,4.0,Average,Smartphone,1.0,0,0,10.0,1.0,None,1.0,2,Remain Constant,NO,Online gaming,,,Friends )
(R1154,Delhi-NCR,15,4.0,Very poor,Smartphone,2.0,1,0,9.0,0.5,Instagram,2.0,4,Remain Constant,NO,Online gaming,,,Friends )
(R1155,Outside Delhi-NCR,17,6.0,Very poor,Smartphone,3.0,1.0,7.0,1.0,Instagram,1.0,2,Decreased,NO,Gym,,,School and friends.)
(R1156,Outside Delhi-NCR,22,5.0,Average,Laptop/Desktop,4.0,0,0,8.0,3.0,Linkedin,0.0,2,Increased,NO,Talking to your relatives,,,School/college)
(R1157,Outside Delhi-NCR,13,6.0,Good,Smartphone,2.0,0,0,11.0,1.0,Whatsapp,0.5,3,Remain Constant,NO,Online gaming,,,Colleagues)
(R1158,Outside Delhi-NCR,11,5.0,Average,Smartphone,2.0,1,0,8.0,1.0,Whatsapp,0.5,1,Decreased,NO,Watching web series,,,School/college)
(R1159,Delhi-NCR,13,4.0,Good,Smartphone,1.0,0,5.7,1.0,1,Instagram,2.0,4,Increased,NO,Listening to music,,,School/college)
(R1160,Outside Delhi-NCR,14,5.0,Excellent,Smartphone,2.0,0,0,12.0,2.0,2,Youtube,1.0,3,Remain Constant,NO,Listening to music,,,Friends )
(R1161,Delhi-NCR,12,6.0,Good,Laptop/Desktop,2.0,2,0,8.0,7.0,Whatsapp,2.0,3,Remain Constant,NO,Social Media,,,Eating outside)
(R1162,Delhi-NCR,13,4.0,Excellent,Laptop/Desktop,4.0,1,0,8.0,0.0,Whatsapp,1.0,2,Remain Constant,NO,Listening to music and reading books both . ,ALL)
(R1164,Delhi-NCR,12,4.0,Good,Smartphone,2.0,1,0,8.0,0.0,Whatsapp,3.0,4,Remain Constant,NO,Meditation,,,Colleagues)
(R1165,Delhi-NCR,12,5.0,Excellent,Laptop/Desktop,3.0,1,0,7.0,0.0,None,2,0,2,Remain Constant,NO,Listening to music,,,Eating outside)
(R1166,Delhi-NCR,13,6.0,Excellent,Smartphone,7.0,0,0,5.0,0.0,0,Youtube,0.0,3,Remain Constant,NO,Poetry,,YES)
(R1167,Outside Delhi-NCR,14,6.0,Excellent,Laptop/Desktop,1.0,2,0,7,1.0,Whatsapp,1.0,4,Increased,NO,Sleeping,,,Roaming around freely)
(R1168,Delhi-NCR,14,6.0,Good,Smartphone,2.0,1,0,8,0,1.0,Whatsapp,1.0,1,Decreased,NO,Social Media,,,Colleagues)
(R1169,Delhi-NCR,12,6.0,Excellent,Smartphone,5.0,1,0,8,0,2.0,Youtube,1.0,2,Remain Constant,NO,Listening to music,,,School/college)
(R1171,Delhi-NCR,14,6.0,Average,Smartphone,2.0,0,5,10,0,1,0,Whatsapp,1.0,3,Remain Constant,NO,Sleeping,,,School and my school friends)
(R1173,Delhi-NCR,12,4.0,Excellent,Smartphone,3.0,2,0,10,0,1,0,Whatsapp,1.0,4,Increased,NO,Sleeping,,,Friends )
(R1174,Delhi-NCR,14,4.0,Excellent,Smartphone,2.0,1,0,6,0,1,0,Whatsapp,1.0,2,Remain Constant,NO,Listening to music,,,School/college)
(R1175,Delhi-NCR,13,6,0,Very poor,Laptop/Desktop,3.0,1,0,8,0,2,0,Whatsapp,2,0,3,Decreased,NO,Listening to music,,,Friends )
(R1176,Delhi-NCR,12,5,0,Excellent,Smartphone,2.0,1,0,12,0,2,0,Youtube,1.0,3,Remain Constant,NO,Dancing,,,Friends )
(R1177,Delhi-NCR,13,5,0,Excellent,Laptop/Desktop,4,0,1,0,8,0,1,0,Whatsapp,1,5,3,Decreased,NO,Listening to music,,,School/college)
(R1178,Delhi-NCR,12,4,0,Very poor,Smartphone,1.5,0,0,9,0,5,0,Youtube,1,0,4,Increased,NO,Painting,,,Friends )
(R1179,Delhi-NCR,13,6,0,Excellent,Laptop/Desktop,2,0,0,5,9,0,1,0,Whatsapp,3,0,4,Remain Constant,NO,Online gaming,,,Friends )
(R1180,Delhi-NCR,12,6,0,Average,Smartphone,3,0,1,0,9,0,1,0,Whatsapp,1,0,4,Remain Constant,NO,Listening to music,,,Roaming around freely)
(R1181,Delhi-NCR,12,5,0,Good,Laptop/Desktop,2,0,0,0,8,0,2,0,Youtube,1,0,3,Remain Constant,NO,Listening to music,,,Colleagues)
(R1182,Delhi-NCR,13,6,0,Good,Tablet,2,0,1,0,6,0,1,0,Whatsapp,2,0,3,Decreased,NO,Dancing,,,Colleagues)
(R1183,Delhi-NCR,16,7,5,Good,Laptop/Desktop,4,5,0,0,6,0,0,5,Whatsapp,2,0,4,Remain Constant,NO,Listening to music,,,Colleagues)
(R1184,Outside Delhi-NCR,13,5,0,Excellent,Laptop/Desktop,6,0,1,0,6,0,0,0,Youtube,0,0,4,Remain Constant,NO,Listening to music,,,School/college)
(R1185,Delhi-NCR,13,4,0,Good,Smartphone,4,0,1,0,8,0,1,0,Whatsapp,2,0,3,Decreased,NO,Reading,,,School/college)
(R1186,Delhi-NCR,15,5,0,Average,Smartphone,2,0,1,0,13,0,1,0,Youtube,1,0,2,Remain Constant,NO,Online surfing,,,School/college)
(R1187,Delhi-NCR,12,6,0,Excellent,Laptop/Desktop,3,0,1,0,9,0,4,0,0,Youtube,3,0,4,Remain Constant,NO,Talking to your relatives,,,School/college)
(R1188,Delhi-NCR,13,8,0,Average,Laptop/Desktop,2,0,1,0,7,0,1,0,Whatsapp,1,0,3,Decreased,NO,Cooking,,Travelling)
(R1189,Delhi-NCR,15,5,0,Good,Tablet,4,0,1,0,7,0,2,0,Whatsapp,0,0,4,Increased,NO,Social Media,,Travelling)
(R1192,Delhi-NCR,14,6,0,Average,Smartphone,4,0,1,0,9,0,1,0,Whatsapp,1,0,4,Remain Constant,NO,Listening to music,,,Friends )
(R1193,Delhi-NCR,13,4,0,Average,Smartphone,0,0,0,5,8,0,3,0,Youtube,2,0,4,Decreased,NO,Online gaming,,,School/college)
(R1194,Delhi-NCR,14,5,0,Excellent,Laptop/Desktop,3,5,1,0,8,0,0,5,Whatsapp,1,0,4,Remain Constant,NO,Reading books,,,School/college)
(R1195,Delhi-NCR,13,5,0,Good,Tablet,2,0,0,5,7,0,1,0,Whatsapp,1,0,3,Remain Constant,NO,Talking,,,School/college)
grunt>
```

**2. Filter data for subjects who spent more than 3 hours on online class and have a rating of "Excellent":**

```
grunt> fil_data = FILTER Survey BY time_spent_on_online_class > 3.0 AND rating_of_online_class_experience == 'Excellent';
2023-04-12 03:32:50,053 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> dump fil_data;
```

**MAPREDUCE:**

```
2023-04-12 03:32:57,337 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 1 time(s).
2023-04-12 03:32:57,338 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: FILTER
2023-04-12 03:32:57,339 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES ENABLED=[AddForEach, ColumnMapKeyPrune, DuplicateForEach, El Setter, ImplicitSplitInserter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NewPartitionFilterOptimizer, PushDownForEachFlatten, PushUpFilter, Partitioner], RULES DISABLED=[FilterLogicExpressionsSimplifier, PartitionerOptimizer]}
2023-04-12 03:32:57,422 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2023-04-12 03:32:57,429 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2023-04-12 03:32:57,429 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2023-04-12 03:32:57,430 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 03:32:57,453 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig script settings are added to the job
2023-04-12 03:32:57,472 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - mapred.job.reduce.markreset.buffer.percent
2023-04-12 03:32:57,472 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Setting up single store job
2023-04-12 03:32:57,472 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Key [pig.schematuple] is false, will not generate code.
2023-04-12 03:32:57,472 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Starting process to move generated code to distributed cache
2023-04-12 03:32:57,472 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Distributed cache not supported or needed in local mode. Setting key [pig.schematuple]: /tmp/1681295577479-0
2023-04-12 03:32:57,481 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLauncher - 1 map-reduce job(s) waiting for submission.
2023-04-12 03:32:57,484 [JobControl] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 03:32:57,496 [JobControl] WARN org.apache.hadoop.mapreduce.JobSubmitter - No job jar file set. User classes may not be found. See Job or Job#setJar()
2023-04-12 03:32:57,503 [JobControl] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 03:32:57,503 [JobControl] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
2023-04-12 03:32:57,508 [JobControl] INFO org.apache.hadoop.mapreduce.JobSubmitter - number of splits:1
2023-04-12 cloudera@quickstart:~ [control] INFO org.apache.hadoop.mapreduce.JobSubmitter - Submitting tokens for job: job_local1010530712_0004
```



**RESULT:**

```
2023-04-12 03:32:57,991 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-04-12 03:32:58,076 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 03:32:58,076 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(R21,Delhi-NCR,20,4.0,Excellent,Laptop/Desktop,5.0,0.0,6.0,2.0,LinkedIn,1.0,2,Decreased,NO,Listening to music,,Roaming around freely)
(R32,Delhi-NCR,23,4.0,Excellent,Laptop/Desktop,18.0,1.0,8.0,2.0,Snapchat,0.0,2,Decreased,NO,Reading books,,Travelling)
(R106,Outside Delhi-NCR,23,4.0,Excellent,Laptop/Desktop,6.0,1.0,8.0,0.3,Whatsapp,0.0,3,Increased,NO,Talking to your relatives,,,Travelling)
(R117,Outside Delhi-NCR,24,4.0,Excellent,Laptop/Desktop,8.0,1.0,8.0,2.0,Youtube,2.0,3,Increased,NO,Online surfing,,,"Friends")
(R159,Delhi-NCR,21,4.0,Excellent,Laptop/Desktop,6.0,2.0,6.0,1.0,Youtube,1.0,3,Remain Constant,NO,Cooking,,Eating outside)
(R173,Outside Delhi-NCR,16,4.0,Excellent,Laptop/Desktop,3.0,2.0,10,0.1,0,Youtube,1.0,3,Increased,NO,Online gaming,,,"Friends")
(R290,Outside Delhi-NCR,20,5.0,Excellent,Laptop/Desktop,1.0,1.0,8.0,3.0,Instagram,0.0,3,Remain Constant,NO,Listening to music,,,"Friends")
(R294,Outside Delhi-NCR,16,7.0,Excellent,Tablet,2.0,1.0,6.0,2.0,Whatsapp,0.0,3,Decreased,NO,Online surfing,,School/college)
(R305,Outside Delhi-NCR,19,5.0,Excellent,Smartphone,5.0,1.0,7.0,2.0,Youtube,3.0,3,Increased,NO,Sleep,,Travelling)
(R366,Outside Delhi-NCR,31,4.0,Excellent,Laptop/Desktop,2.0,1.0,8.0,1.0,Facebook,1.0,3,Increased,NO,Sleep,,School/college)
(R367,Delhi-NCR,23,5.0,Excellent,Smartphone,1.0,1.0,9.0,2.0,Whatsapp,0.0,4,Decreased,NO,Online gaming,,Football)
(R407,Delhi-NCR,17,7.0,Excellent,Smartphone,1.0,1.0,4.0,1.0,Youtube,0.0,1,Decreased,NO,Cooking,,School/college)
(R423,Delhi-NCR,19,8.0,Excellent,Laptop/Desktop,10,0,1.0,6.0,1.0,LinkedIn,0.0,4,Remain Constant,NO,Online gaming,,School/college)
(R486,Delhi-NCR,21,6.0,Excellent,Laptop/Desktop,6.0,1.0,6.0,1.0,Instagram,0.0,4,Decreased,NO,Driving,,Roaming around freely)
(R569,Delhi-NCR,20,8.0,Excellent,Smartphone,8.0,1.0,7.0,2.0,Youtube,1.0,4,Increased,NO,Watching web series,,,Roaming around freely)
(R631,Delhi-NCR,13,4.0,Excellent,Smartphone,5.0,1.0,13.0,2.0,Whatsapp,3.0,2,Decreased,YES,Reading books,,,"Friends")
(R649,Delhi-NCR,30,4.0,Excellent,Laptop/Desktop,4.0,0,0,5,0,6,0,Whatsapp,1.0,2,Remain Constant,NO,Watching web series,,,Roaming around freely)
(R724,Delhi-NCR,38,4.0,Excellent,Laptop/Desktop,5.0,1.0,6.0,1.0,Youtube,2.0,3,Increased,NO,Online gaming,,Roaming around freely)
(R772,Outside Delhi-NCR,23,5.5,Excellent,Laptop/Desktop,3.0,1.0,8.0,1.0,Whatsapp,0.0,2,Decreased,NO,Watching web series,,School/college)
(R776,Outside Delhi-NCR,40,5.0,Excellent,Laptop/Desktop,4.0,0,0,5,6,0,1.0,Whatsapp,2.0,2,Remain Constant,NO,Talking to your relatives,,,Travelling)
(R789,Outside Delhi-NCR,24,5.0,Excellent,Laptop/Desktop,3.0,3,0,8.0,2.0,Whatsapp,2.0,3,Remain Constant,NO,By engaging in my work,,,ALL)
(R800,Delhi-NCR,23,5.0,Excellent,Laptop/Desktop,2.0,1.0,8.0,1.0,Whatsapp,4.0,3,Remain Constant,YES,Meditation,,,"Friends")
(R838,Outside Delhi-NCR,40,6.0,Excellent,Laptop/Desktop,2.0,1.0,6,0,2.0,Whatsapp,2.0,3,Decreased,NO,Listening to music,,School/college)
(R923,Outside Delhi-NCR,9,5.0,Excellent,Smartphone or Laptop/Desktop,2.0,0,6,7,0,1.0,Whatsapp,0.0,4,Remain Constant,NO,Online gaming,,Eating outside)
(R940,Outside Delhi-NCR,40,5.0,Excellent,Laptop/Desktop,4.0,1,0,7,0,6,0,LinkedIn,2.0,3,Remain Constant,NO,Listening to music,,School/college)
(R1047,Outside Delhi-NCR,21,4.0,Excellent,Laptop/Desktop,2.0,0,0,8,0,0,5,Whatsapp,4.0,3,Increased,YES,Meditation,,,"Friends")
(R1065,Outside Delhi-NCR,16,4.0,Excellent,Smartphone or Laptop/Desktop,3.0,1,0,8,0,5,0,Youtube,5.0,3,Increased,NO,Listening to music,,Roaming around freely)
(R1087,Delhi-NCR,15,5.0,Excellent,Laptop/Desktop,4,0,0,5,8,0,0,0,None,1,0,3,Remain Constant,NO,Sleeping,,,"Friends")
(R1096,Delhi-NCR,15,9.0,Excellent,Laptop/Desktop,10,0,0,5,6,0,0,0,Whatsapp,2.0,3,Remain Constant,NO,Reading books,,,"Friends")
(R1111,Delhi-NCR,16,5.0,Excellent,Smartphone,1.0,1,0,11,0,0,0,Whatsapp,2,0,3,Remain Constant,NO,Listening to music,,,"Friends")
(R1118,Delhi-NCR,14,4.0,Excellent,Smartphone,2,0,1,0,9,0,1,0,Whatsapp,2,0,3,Remain Constant,NO,Listening to music,,,"Friends")
(R1119,Outside Delhi-NCR,13,4.0,Excellent,Smartphone,3,0,0,25,8,0,1,0,Youtube,0,5,1,Increased,NO,Online gaming,,,"Friends")
(R1122,Delhi-NCR,13,5.0,Excellent,Tablet,4,0,1,0,8,0,0,0,Youtube,0,5,4,Remain Constant,NO,Listening to music,,School/college)
(R1126,Delhi-NCR,12,4.0,Excellent,Smartphone,2,0,1,0,6,0,0,0,Youtube,2,0,2,Remain Constant,YES,Listening to music,,,"Friends")
(R1134,Delhi-NCR,14,4.0,Excellent,Smartphone,2,0,2,0,12,0,1,0,Whatsapp,2,0,3,Decreased,NO,Watching web series,,School/college)
(R1142,Delhi-NCR,11,4.0,Excellent,Smartphone,2,0,0,0,11,0,1,0,Whatsapp,2,0,2,Remain Constant,NO,Meditation,,,"Friends")
(R1143,Browsing HDFS - Mozilla Firefox,plet,2,0,0,5,9,0,1,0,Whatsapp,1,0,3,Increased,NO,Youtube,,,"Friends")
```



**3. Calculate the average age and total time spent on self-study for each region of residence:**

```
details at /tmp/cloudera/pig_1001252105440.log
grunt> group data= GROUP Survey BY region_of_residence;
2023-04-12 03:36:42,597 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> result = FOREACH group data GENERATE group AS region, AVG(Survey.age_of_subject) AS avg_age, SUM(Survey.time_spent_on_self_study) AS total_time;
2023-04-12 03:40:00,721 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> dump result;
```

**MapReduce:**

```
2023-04-12 03:40:08,449 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP BY
2023-04-12 03:40:08,450 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, DuplicateForEachColumnRewrite, GroupByConstParaller, ImplicitSplitInserter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NewPartitionFilterOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter], RULES_DISABLED=[FilterLogicExpressionSimplifier, PartitionFilterOptimizer]}
2023-04-12 03:40:08,475 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2023-04-12 03:40:08,477 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.CombineOptimizer - Choosing to move algebraic foreach to combiner
2023-04-12 03:40:08,507 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2023-04-12 03:40:08,508 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2023-04-12 03:40:08,509 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 03:40:08,512 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig script settings are added to the job
2023-04-12 03:40:08,561 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2023-04-12 03:40:08,562 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Reduc phase detected, estimating # of required reducers.
2023-04-12 03:40:08,566 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using reducer estimator: org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2023-04-12 03:40:08,604 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator - BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=144782
2023-04-12 03:40:08,604 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Setting Parallelism to 1
2023-04-12 03:40:08,610 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Setting up single store job
2023-04-12 03:40:08,611 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Key [pig.schematuple] is false, will not generate code.
2023-04-12 03:40:08,611 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Starting process to move generated code to distributed cache
2023-04-12 03:40:08,611 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Distributed cache not supported or needed in local mode. Setting key [pig.schematuple.local.dir] with code temp direct
2023-04-12 03:40:08,918 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1 map-reduce job(s) waiting for submission.
2023-04-12 03:40:08,922 [JobControl] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 03:40:08,971 [JobControl] WARN org.apache.hadoop.mapreduce.JobSubmitter - No jar job file set. User classes may not be found. See Job or Job#setJar(String).
2023-04-12 03:40:08,977 [JobControl] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 03:40:08,977 [JobControl] INFO org.apache.pig.backend.hadoop.executionengine.util.MapReduceUtil - Total input paths to process : 1
2023-04-12 03:40:08,980 [JobControl] INFO org.apache.hadoop.mapreduce.JobSubmitter - number of splits:1
2023-04-12 03:40:09,018 [JobControl] INFO org.apache.hadoop.mapreduce.JobSubmitter - Submitting tokens for job: job_local154093582_0005
2023-04-12 03:40:09,211 [JobControl] INFO org.apache.hadoop.mapreduce.Job - The url to track the job: http://localhost:8080/
2023-04-12 03:40:18,799 [pool-11-thread-1] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt_local154093582_0005_r_000000_0' to file:/tmp/temp952197486/6/tmp95333657/_temporary/_r/task_local154093582_0005_r_000000_0
2023-04-12 03:40:18,800 [pool-11-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2023-04-12 03:40:18,791 [pool-11-thread-1] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local154093582 0005_r_000000_0' done.
2023-04-12 03:40:18,791 [pool-11-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local154093582_0005_r_000000_0
2023-04-12 03:40:18,791 [Thread-23] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce task executor complete.
2023-04-12 03:40:19,101 [main] WARN org.apache.pig.tools.pigstats.PigStatsUtil - Failed to get RunningJobs for job job_local154093582_0005
2023-04-12 03:40:19,101 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-04-12 03:40:19,101 [main] INFO org.apache.pig.tools.pigstats.SimplePigState - Detected Local mode. Stats reported below may be incomplete
2023-04-12 03:40:19,102 [main] INFO org.apache.pig.tools.pigstats.SimplePigState - Script Statistics:
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.6.0-cdh5.4.2 0.12.0-cdh5.4.2 cloudera 2023-04-12 03:40:08 2023-04-12 03:40:19 GROUP_BY
Success!
Job Stats (time in seconds):
JobId Alias Feature Outputs
job_local154093582_0005 Survey,group_data,result GROUP_BY,COMBINER file:/tmp/temp952197486/tmp69333657,
Input(s):
Successfully read records from: "file:///home/cloudera/Desktop/Survey.csv"
Output(s):
Successfully stored records in: "file:/tmp/temp952197486/tmp69333657"
Job DAG:
job_local154093582_0005

2023-04-12 03:40:19,103 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-04-12 03:40:19,104 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 03:40:19,105 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 03:40:19,107 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-04-12 03:40:19,134 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 03:40:19,134 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
```

**RESULT:**

```
(Delhi-NCR,19.832177531206657,2113.0)
(Outside Delhi-NCR,20.68763557483731,1328.5)
(Region of residence,,)
grunt> █
```

**4. Filter data for subjects who reported a health issue and spent more than 2 hours on fitness:**

```
grunt> fd= FILTER Survey BY health_issue_during_lockdown == 'YES' AND time_spent_on_fitness > 2;
2023-04-12 03:53:44,694 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_FLOAT 1 time(s).
2023-04-12 03:53:44,694 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> dump fd;
```

## MAPREDUCE:

```

2023-04-12 03:53:58,275 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_FLOAT 1 time(s).
2023-04-12 03:53:58,275 [main] INFO org.apache.tools.pigstats.ScriptState - Pig features used in the script: FILTER
2023-04-12 03:53:58,276 [main] INFO org.apache.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, DuplicateForEachColumnRewrite, GroupByConstParallelSetter, ImplicitSplitInserter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NewPartitionFilterOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter], RULES_DISABLED=[FilterLogicExpressionSimplifier, PartitionFilterOptimizer]}
2023-04-12 03:53:58,288 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - map concatenation threshold: 100 optimistic? false
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - MR plan size before optimization: 1
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - MR plan size after optimization: 1
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - MapReduceLauncher - already initialized
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Pig script settings are added to the job
2023-04-12 03:53:58,290 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2023-04-12 03:53:58,312 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Setting up single store job
2023-04-12 03:53:58,317 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Key [pig.schematuple] is false, will not generate code.
2023-04-12 03:53:58,317 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Starting process to move generated code to distributed cache
2023-04-12 03:53:58,317 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Distributed cache not supported or needed in local mode. Setting key [pig.schematuple.local.dir] with code temp directory: /tmp/1681296838314-0
2023-04-12 03:53:58,320 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1 map-reduce job(s) waiting for submission.
2023-04-12 03:53:58,327 [JobControl] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 03:53:58,340 [JobControl] WARN org.apache.hadoop.mapreduce.Lib - No job jar file set. User classes may not be found. See Job or Job#setJar(String).
2023-04-12 03:53:58,349 [JobControl] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 03:53:58,350 [JobControl] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
2023-04-12 03:53:58,350 [JobControl] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Total input paths (combined) to process : 1
2023-04-12 03:53:58,350 [JobControl] INFO org.apache.hadoop.mapreduce.JobSubmitter - Submitting tokens for job: job_local1845654830_0006
2023-04-12 03:53:58,356 [JobControl] INFO org.apache.hadoop.mapreduce.Job - The url to track the job: http://localhost:8080/
2023-04-12 03:53:58,430 [Thread-33] INFO org.apache.hadoop.mapred.LocalJobRunner - OutputCommitter set in config null
2023-04-12 03:53:58,434 [Thread-33] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 03:53:58,434 [Thread-33] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use mapreduce.reduce.markreset.buffer.percent
2023-04-12 03:53:58,434 [Thread-33] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2023-04-12 03:53:58,434 [Thread-33] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 03:53:58,554 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.Task - Task attempt_local1845654830_0006_m_000000_0 done.
2023-04-12 03:53:58,555 [LocalJobRunner Map Task Executor #0] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local1845654830_0006_m_000000_0
2023-04-12 03:53:58,555 [Thread-33] INFO org.apache.hadoop.mapred.LocalJobRunner - map task executor complete.
2023-04-12 03:53:58,827 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - HadoopJobId: job_local1845654830_0006
2023-04-12 03:53:58,827 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Processing aliases Survey,fd
2023-04-12 03:53:58,827 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - detailed locations: M: Survey[1,8],Survey[-1,-1],fd[8,4] C: R:
2023-04-12 03:53:58,835 [main] WARN org.apache.tools.pigstats.PigStatsUtil - Failed to get RunningJob for job job_local1845654830_0006
2023-04-12 03:53:58,835 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-04-12 03:53:58,835 [main] INFO org.apache.pig.tools.pigstats.SimplePigStats - Detected Local mode. Stats reported below may be incomplete
2023-04-12 03:53:58,835 [main] INFO org.apache.pig.tools.pigstats.SimplePigStats - Script Statistics:
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.6.0-cdh5.4.2 0.12.0-cdh5.4.2 cloudera 2023-04-12 03:53:58 2023-04-12 03:53:58 FILTER
Success!
Job Stats (time in seconds):
JobId Alias Feature Outputs
job_local1845654830_0006 Survey,fd MAP_ONLY file:/tmp/temp952197486/tmp2049490070,
Input(s):
Successfully read records from: "file:///home/cloudera/Desktop/Survey.csv"
Output(s):
Successfully stored records in: "file:/tmp/temp952197486/tmp2049490070"
Job DAG:
job_local1845654830_0006

2023-04-12 03:53:58,836 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-04-12 03:53:58,836 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 03:53:58,836 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 03:53:58,836 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-04-12 03:53:58,860 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 03:53:58,861 [main] INFO org.apache.oio.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1

```

## RESULT:

```

(R193,Delhi-NCR,21,3.0,Very poor,Smartphone,8.0,3.0,7.0,3.0,Youtube,0.0,3,Decreased,YES,Talking to your relatives,,,School/college)
(R728,Delhi-NCR,20,5.0,Very poor,Laptop/Desktop,2.0,3.0,10.0,1.0,Youtube,0.5,2,Decreased,YES,Writing my own Comics & novels,,,Playing)
(R826,Outside Delhi-NCR,21,0.0,Very poor,Smartphone,2.0,3.0,9.0,1.0,Whatsapp,1.0,1,Decreased,YES,Online surfing,,,Roaming around freely)
grunt> ■

```

## 5. Calculate the total number of subjects who prefer each social media platform:

```

grunt> gd= GROUP Survey by prefered_social_media_platform;
2023-04-12 04:02:20,225 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_FLOAT 1 time(s).
2023-04-12 04:02:20,225 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> rse= FOREACH gd GENERATE group AS sc, COUNT(Survey) AS tt;
2023-04-12 04:03:23,500 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_FLOAT 1 time(s).
2023-04-12 04:03:23,500 [main] WARN org.apache.pig.PigServer - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> dump rse;

```

## MAPREDUCE:

```

2023-04-12 04:03:36,859 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: GROUP_BY
2023-04-12 04:03:36,859 [main] INFO org.apache.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, DuplicateForEachColumnRewrite, GroupByConstParallelSetter, ImplicitSplitInserter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NewPartitionFilterOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter], RULES_DISABLED=[FilterLogicExpressionSimplifier, PartitionFilterOptimizer]}
2023-04-12 04:03:36,899 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2023-04-12 04:03:36,900 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.CombineOptimizer - Choosing to move algebraic foreach to combiner
2023-04-12 04:03:36,903 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.CombineOptimizer - MR plan size before optimization: 1
2023-04-12 04:03:36,903 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.CombineOptimizer - MR plan size after optimization: 1
2023-04-12 04:03:36,909 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 04:03:36,911 [main] INFO org.apache.tools.pigstats.ScriptState - Pig script settings are added to the job
2023-04-12 04:03:36,920 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2023-04-12 04:03:36,921 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Reduce phase detected, estimating # of required reducers.
2023-04-12 04:03:36,921 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using reducer estimator: org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2023-04-12 04:03:36,921 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator - BytesPerReducer=10000000000 maxReducers=999 totalInputFileSize=144782
2023-04-12 04:03:36,921 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Setting Parallelism to 1
2023-04-12 04:03:36,934 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Setting up single store job
2023-04-12 04:03:36,937 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Key [pig.schematuple] is false, will not generate code.
2023-04-12 04:03:36,937 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Starting process to move generated code to distributed cache
2023-04-12 04:03:36,937 [main] INFO org.apache.pig.data.SchemaTupleFrontend - Distributed cache not supported or needed in local mode. Setting key [pig.schematuple.local.dir] with code temp directory: /tmp/1681297416935-0
2023-04-12 04:03:36,959 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1 map-reduce job(s) waiting for submission.
2023-04-12 04:03:36,961 [JobControl] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Can initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-04-12 04:03:36,961 [JobControl] WARN org.apache.hadoop.mapreduce.Lib - No job jar file set. User classes may not be found. See Job or Job#setJar(String).
2023-04-12 04:03:36,978 [JobControl] INFO org.apache.hadoop.mapreduce.Lib - Input paths to process : 1
2023-04-12 04:03:36,978 [JobControl] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 04:03:36,978 [JobControl] INFO org.apache.hadoop.mapreduce.Job - Reducer number of splits:1
2023-04-12 04:03:36,979 [JobControl] INFO org.apache.hadoop.mapreduce.JobSubmitter - Submitting tokens for job: job_local1512296766_0007
2023-04-12 04:03:40,000 [JobControl] INFO org.apache.hadoop.mapreduce.Job - The url to track the job: http://localhost:8080/
2023-04-12 04:03:40,000 [JobControl] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - HadoopJobId: job_local1512296766_0007
2023-04-12 04:03:40,000 [JobControl] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Processing aliases Survey,gd,rse

```

```
Success!
Job Stats (time in seconds):
JobId Alias Feature Outputs
job_local512296766_0007 Survey,gd,rse GROUP_BY,COMBINER      file:/tmp/temp952197486/tmp1385900907,
Input(s):
Successfully read records from: "file:///home/cloudera/Desktop/Survey.csv"
Output(s):
Successfully stored records in: "file:/tmp/temp952197486/tmp1385900907"

Job DAG:
job_local512296766_0007

2023-04-12 04:03:41,683 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-04-12 04:03:41,683 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2023-04-12 04:03:41,683 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2023-04-12 04:03:41,683 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-04-12 04:03:41,771 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2023-04-12 04:03:41,771 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(None 17)
```

#### RESULT:

```
(None,17)
(None ,1)
(Quora,1)
(Omegle,1)
(Reddit,5)
(Elyment,1)
(Twitter,28)
(Youtube,314)
(Facebook,52)
(Linkedin,61)
(Snapchat,8)
(Talklife,1)
(Telegram,3)
(WhatsApp,1)
(Whatsapp,336)
(Instagram,352)
(Preferred social media platform,1)
grunt> █
```