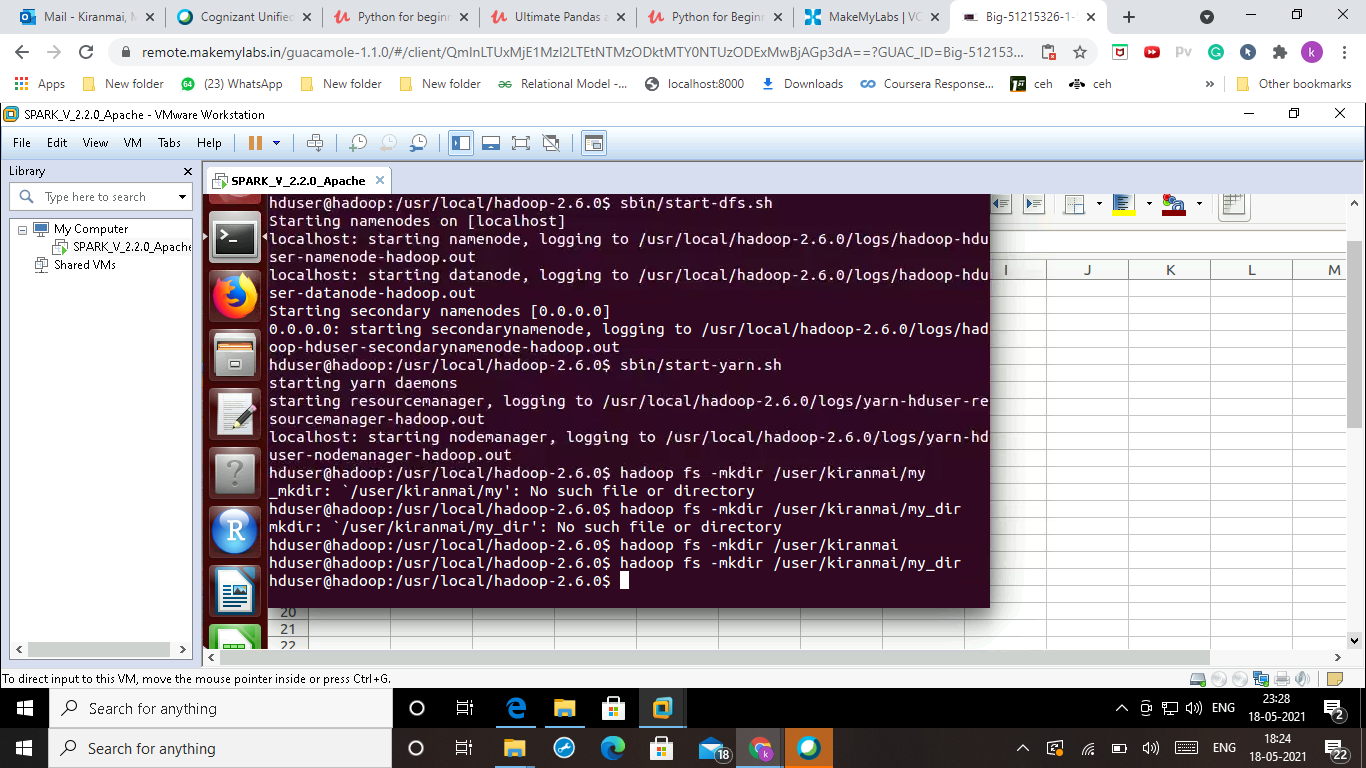
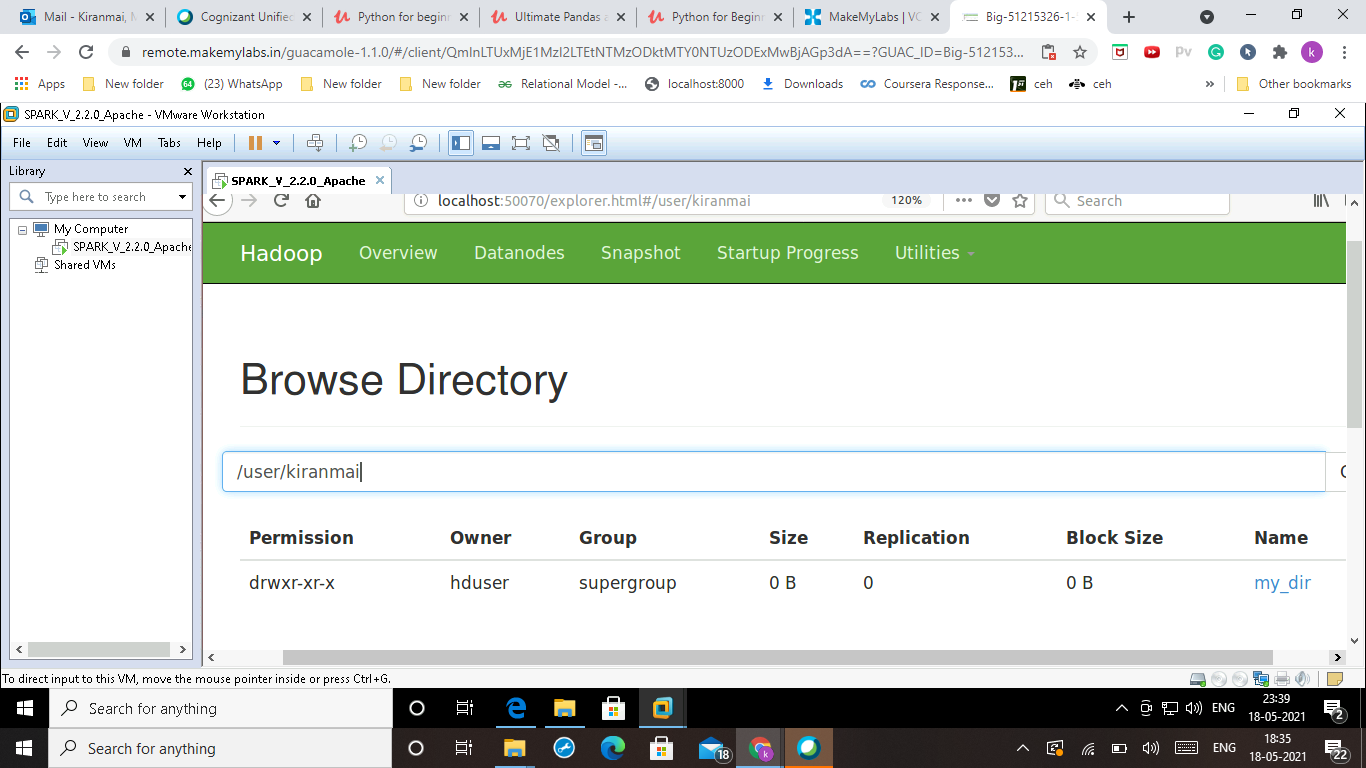
LAB ASSIGNMENTS

HDFS:

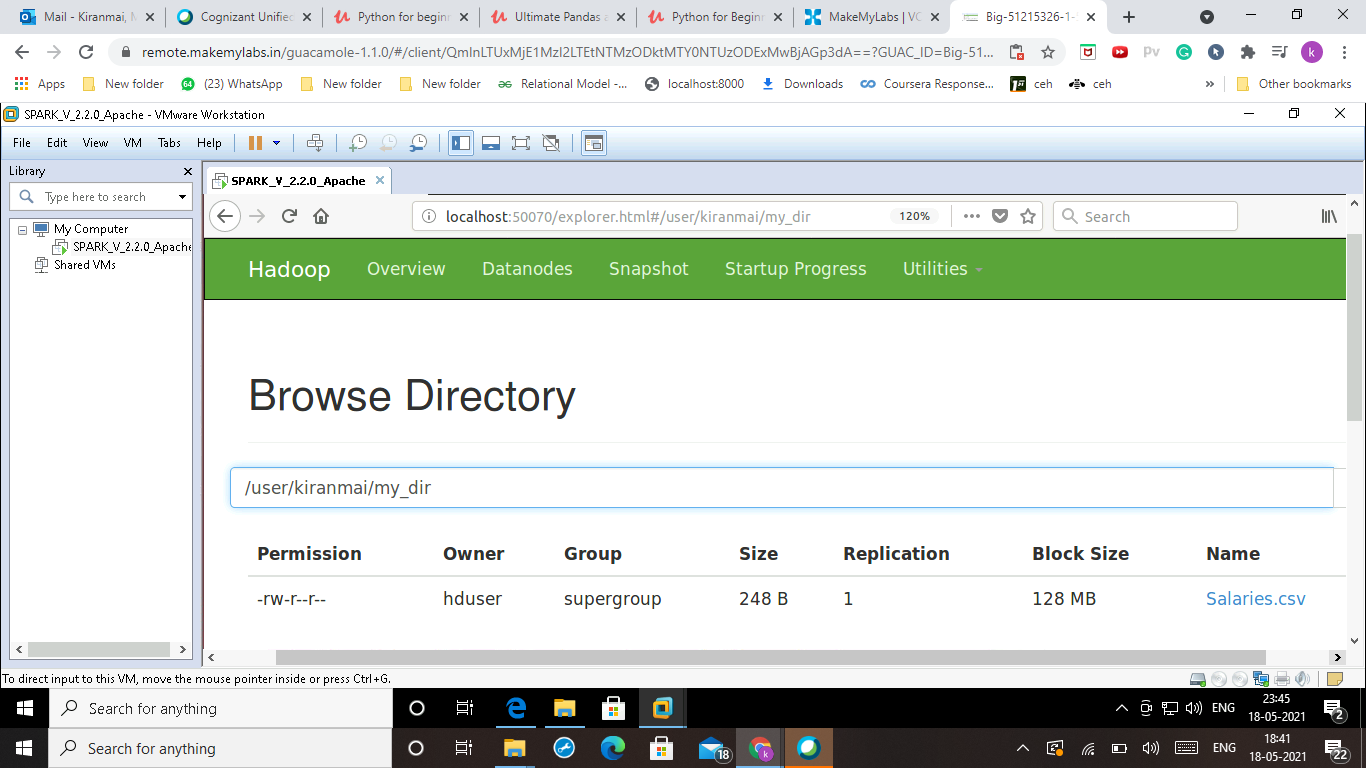
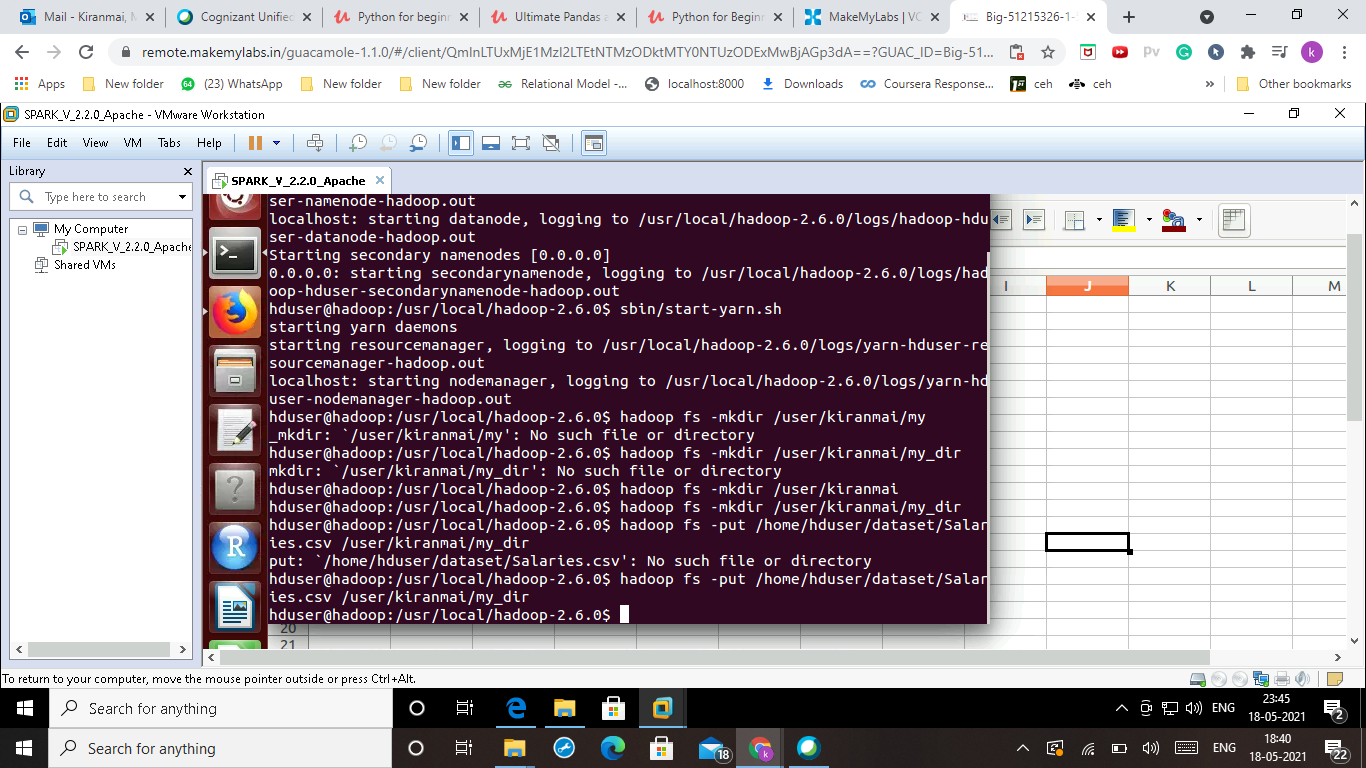
1.Create a directory with your name under/user/<yourname>/hdfs directory in HDFS



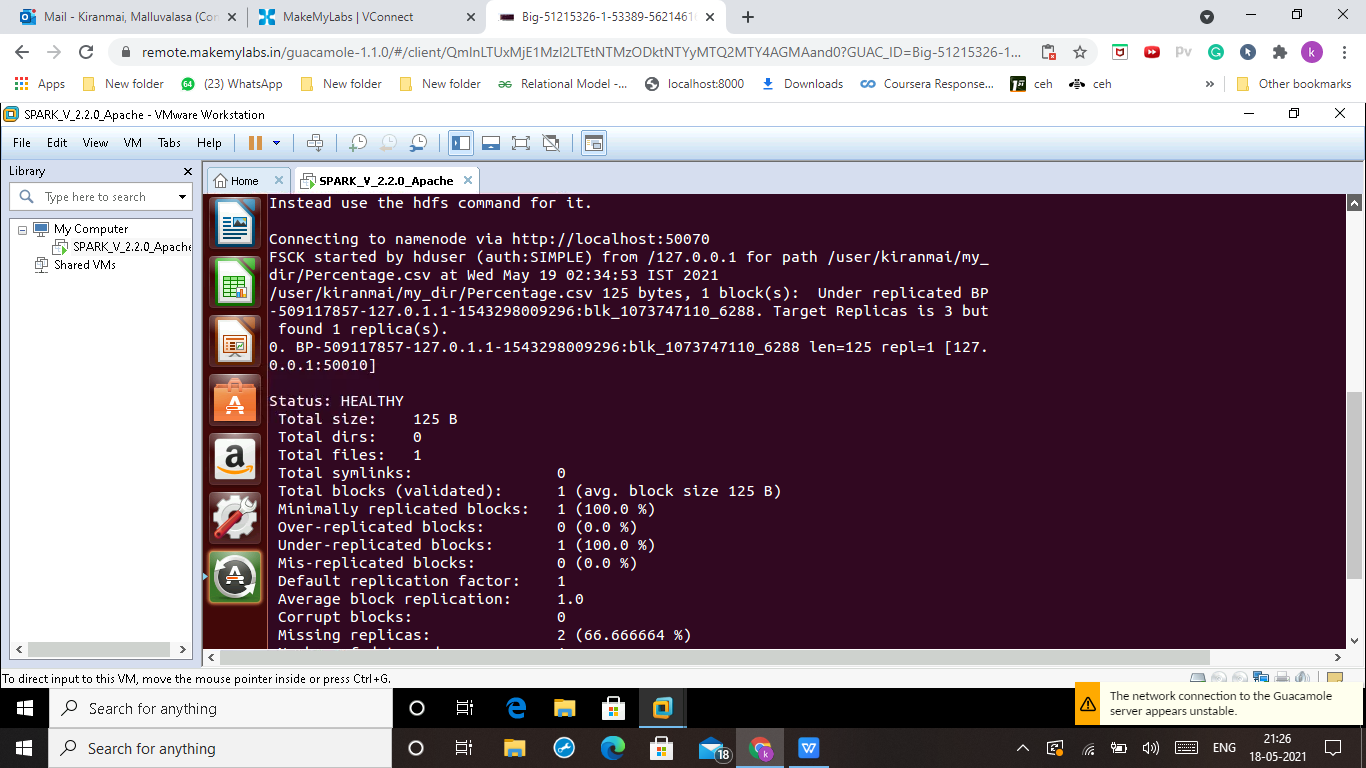
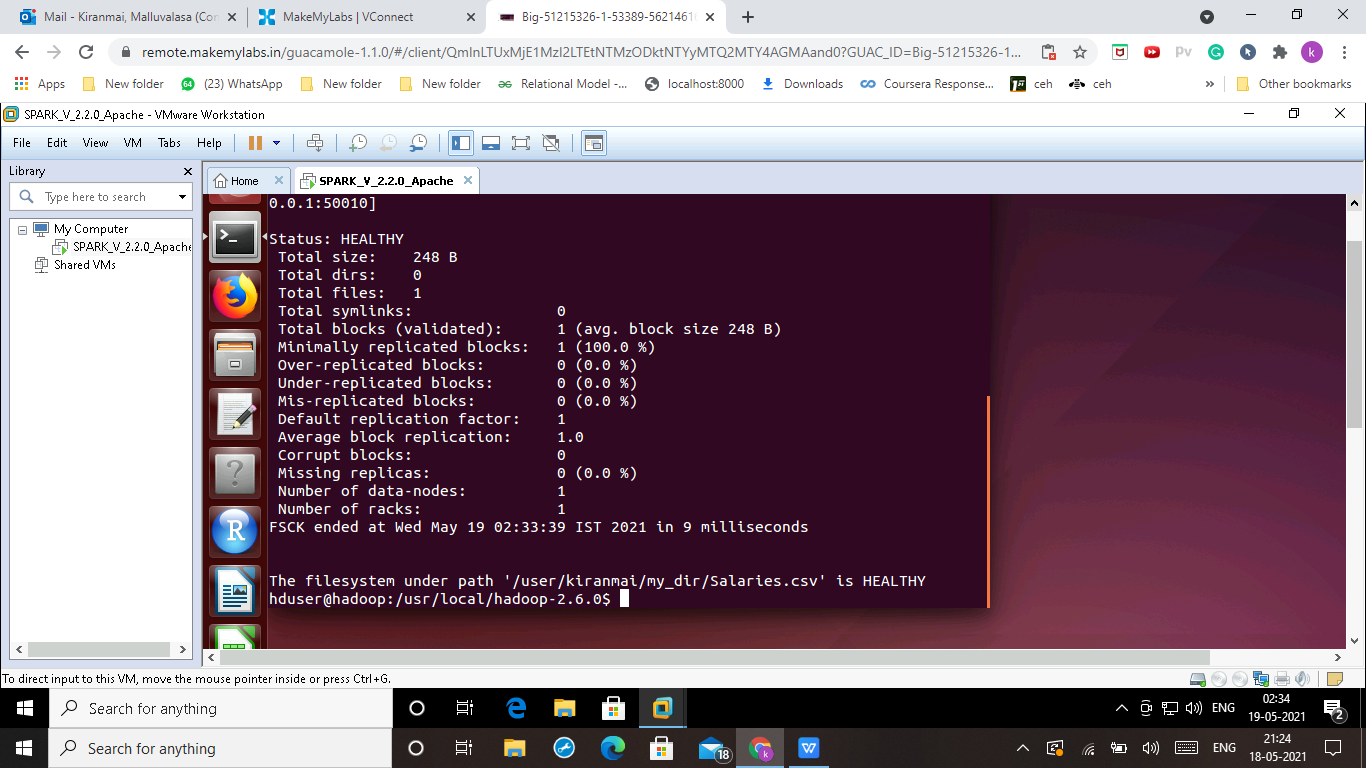


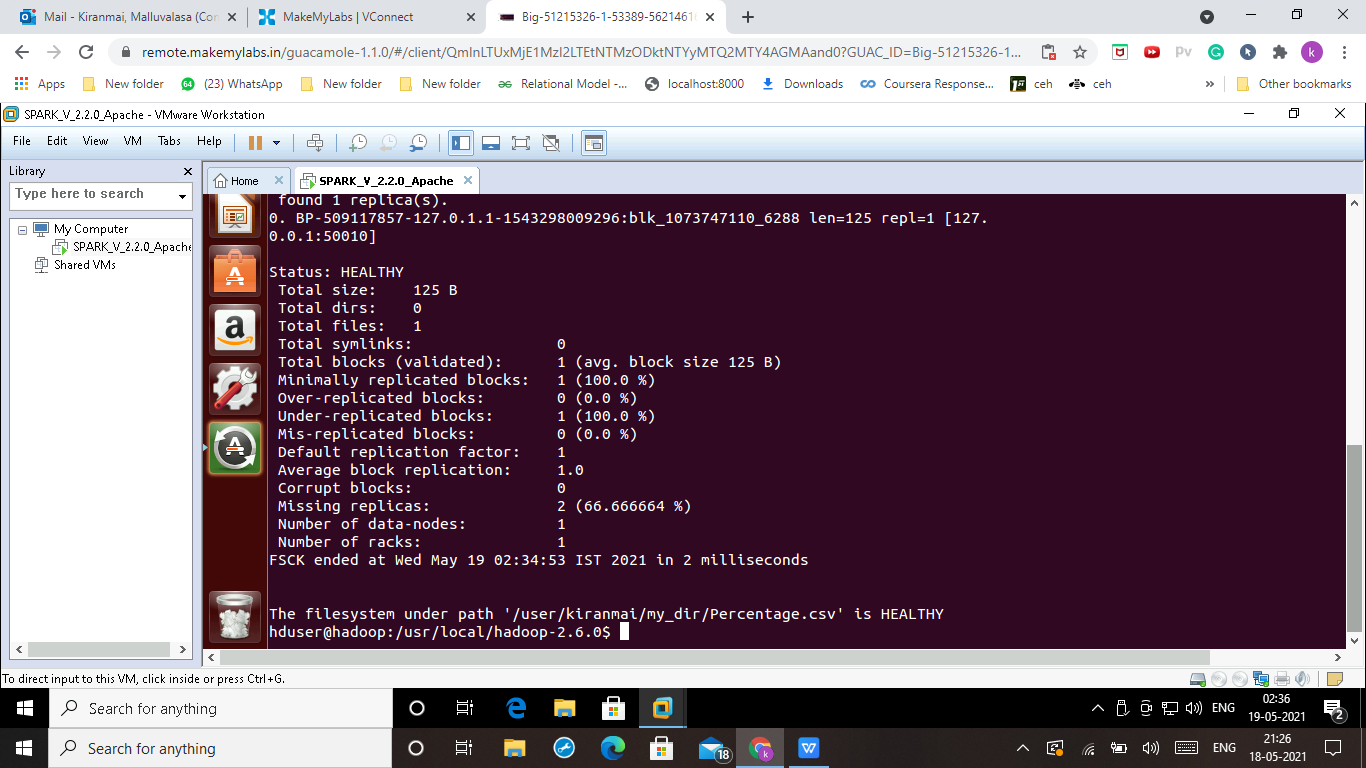
2.Create sample file on you local machine and upload that file in your directory in HDFS



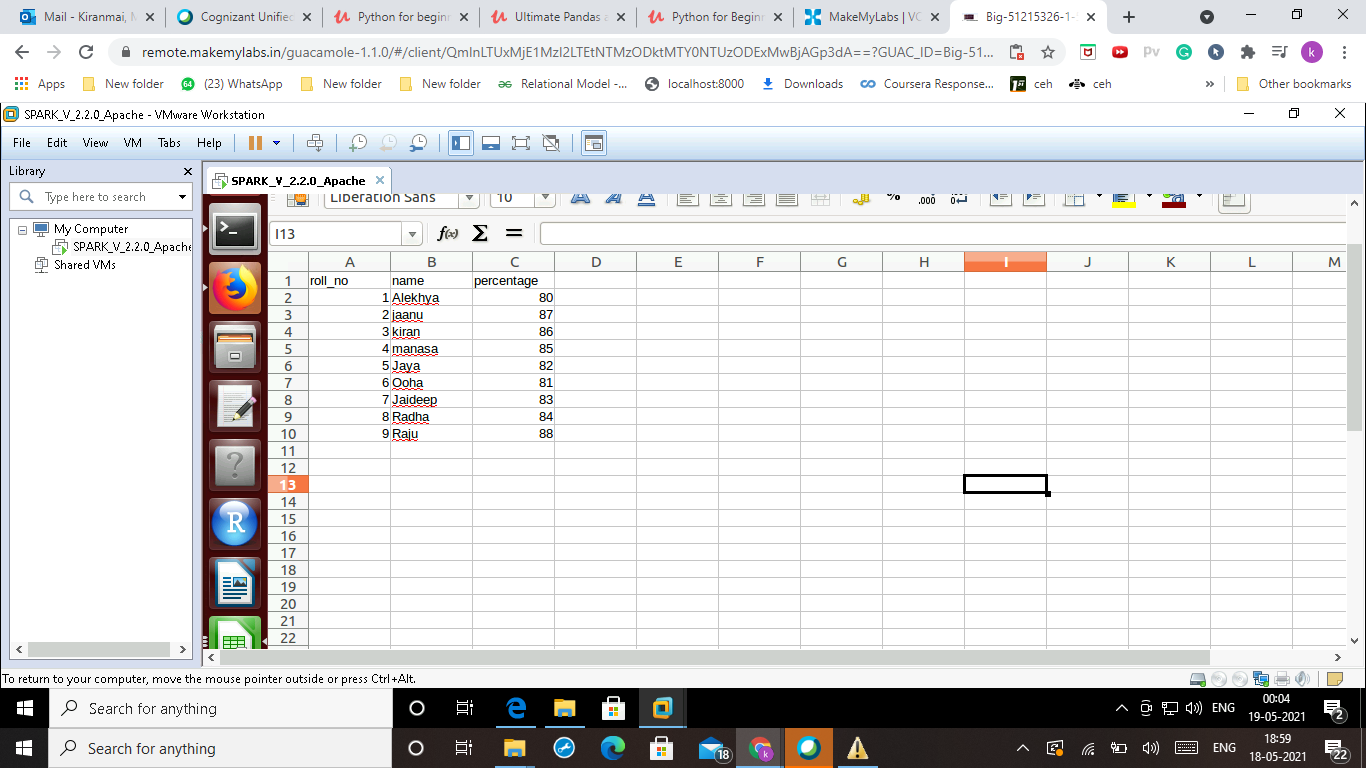


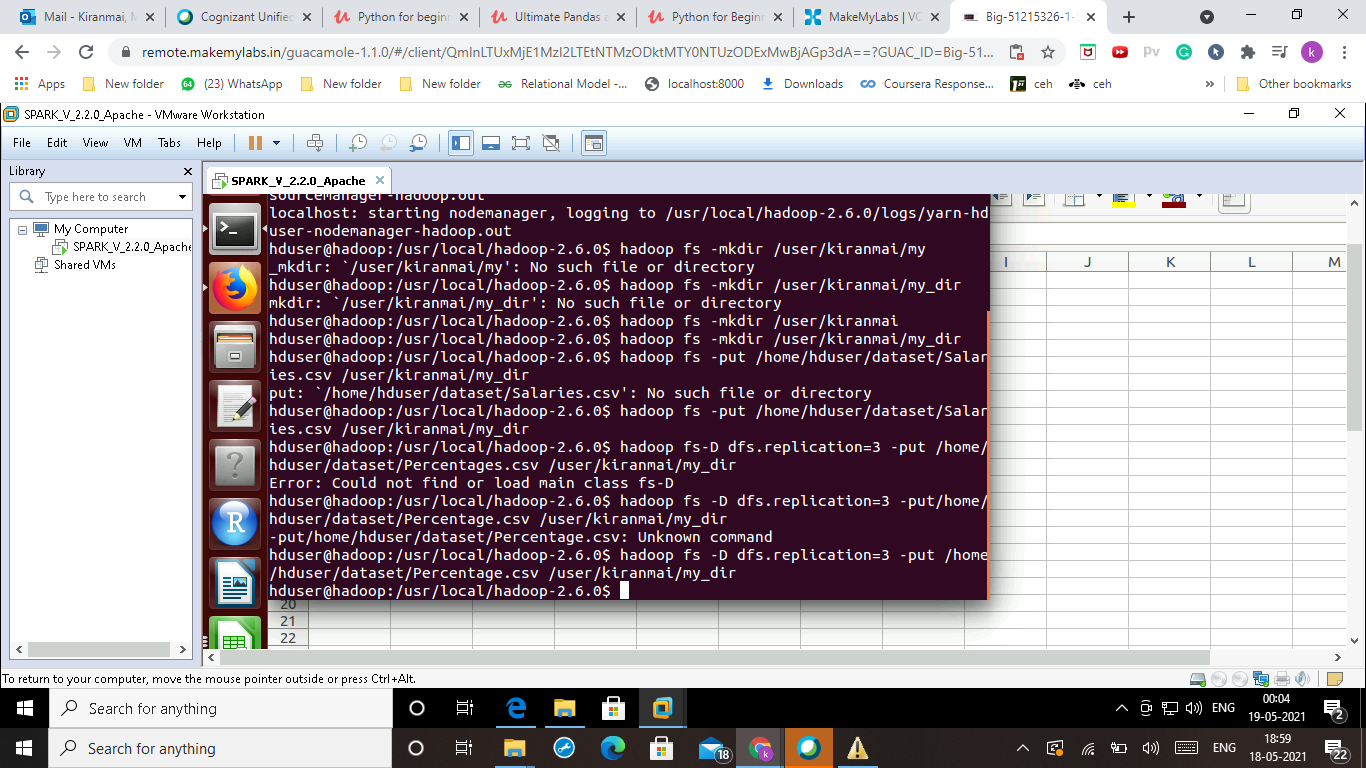
1. Execute HDFS report and find out block names in which your stored on Hdfs

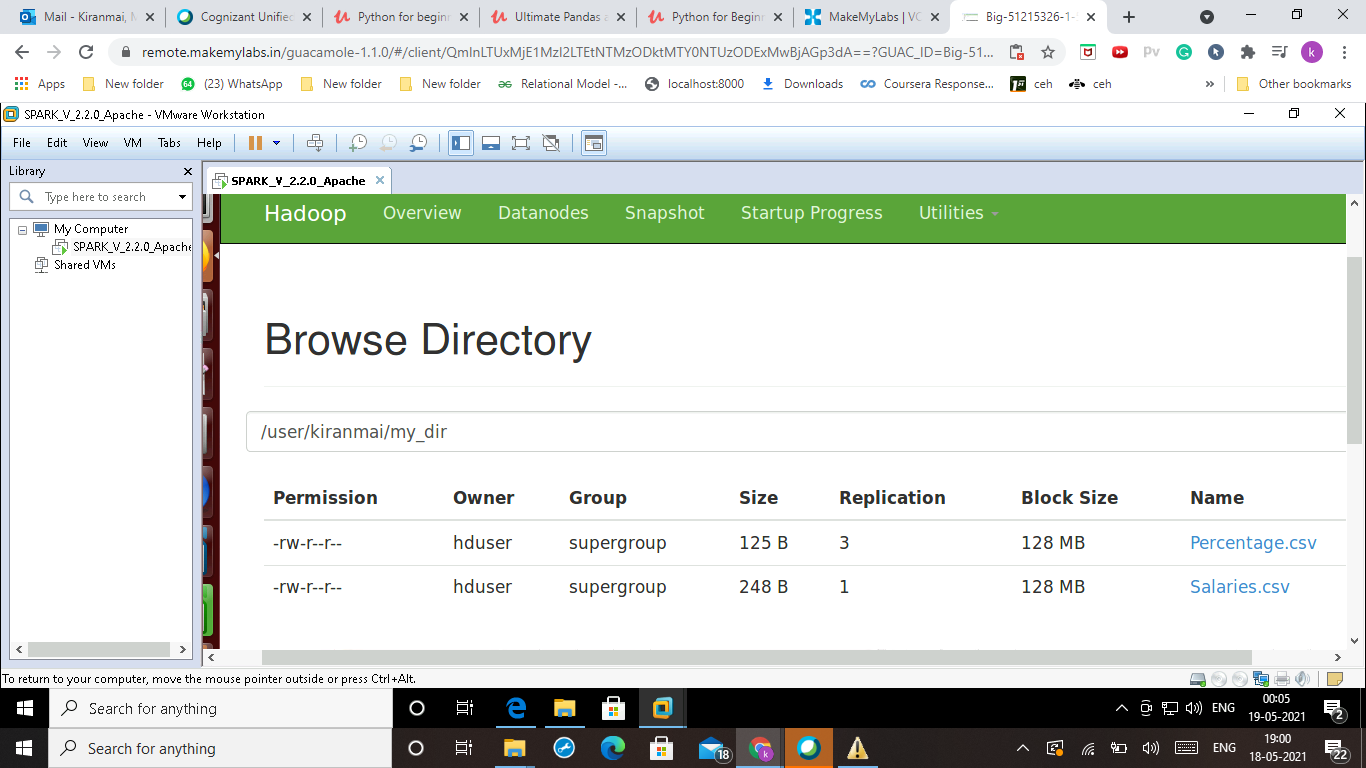




1. Create a new files on your local machine and upload it with replication factor 3 in your HDFS directory



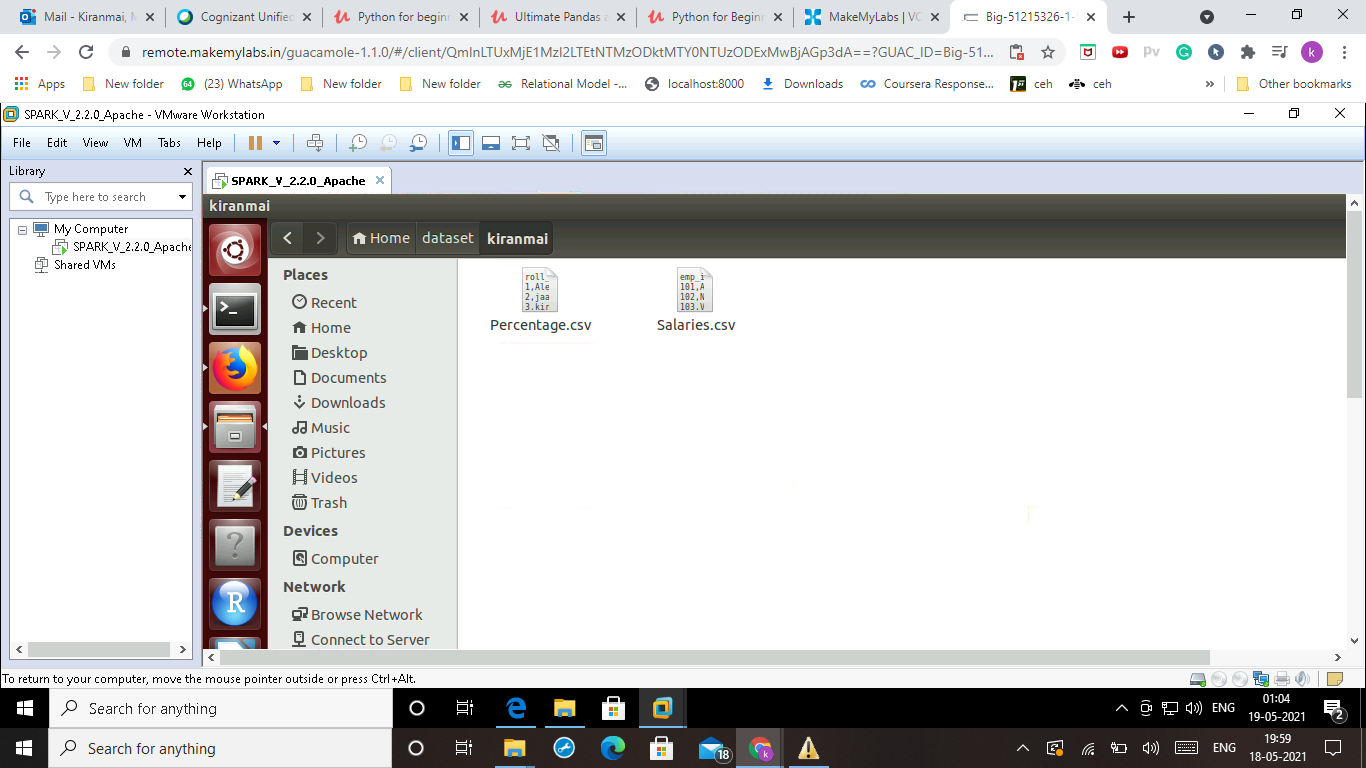




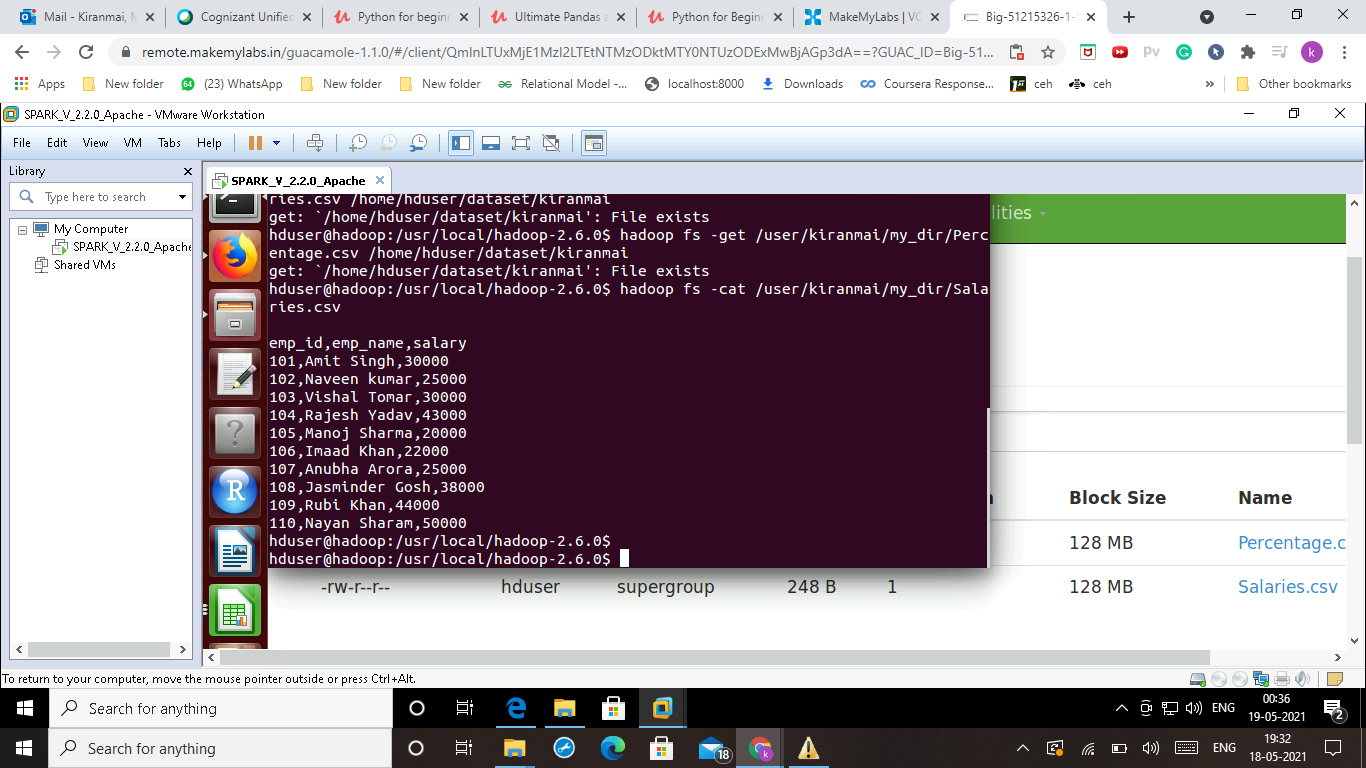
1. Download your both files from your directory on HDFS to your local machine

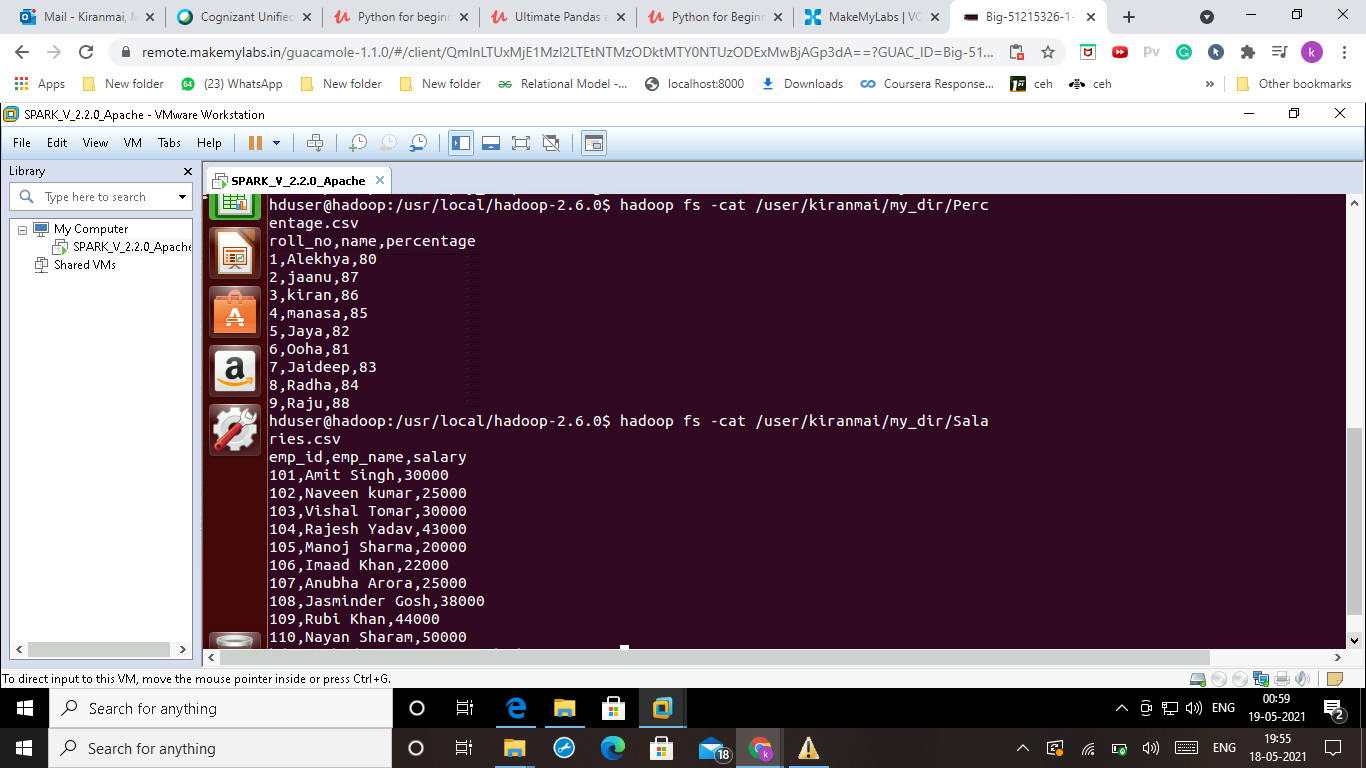
hadoop fs -get /user/kiranmai/my\_dir/Salaries.csv /home/hduser/dataset/kiranmai

hadoop fs -get /user/kiranmai/my\_dir/Percentages.csv /home/hduser/dataset/kiranmai

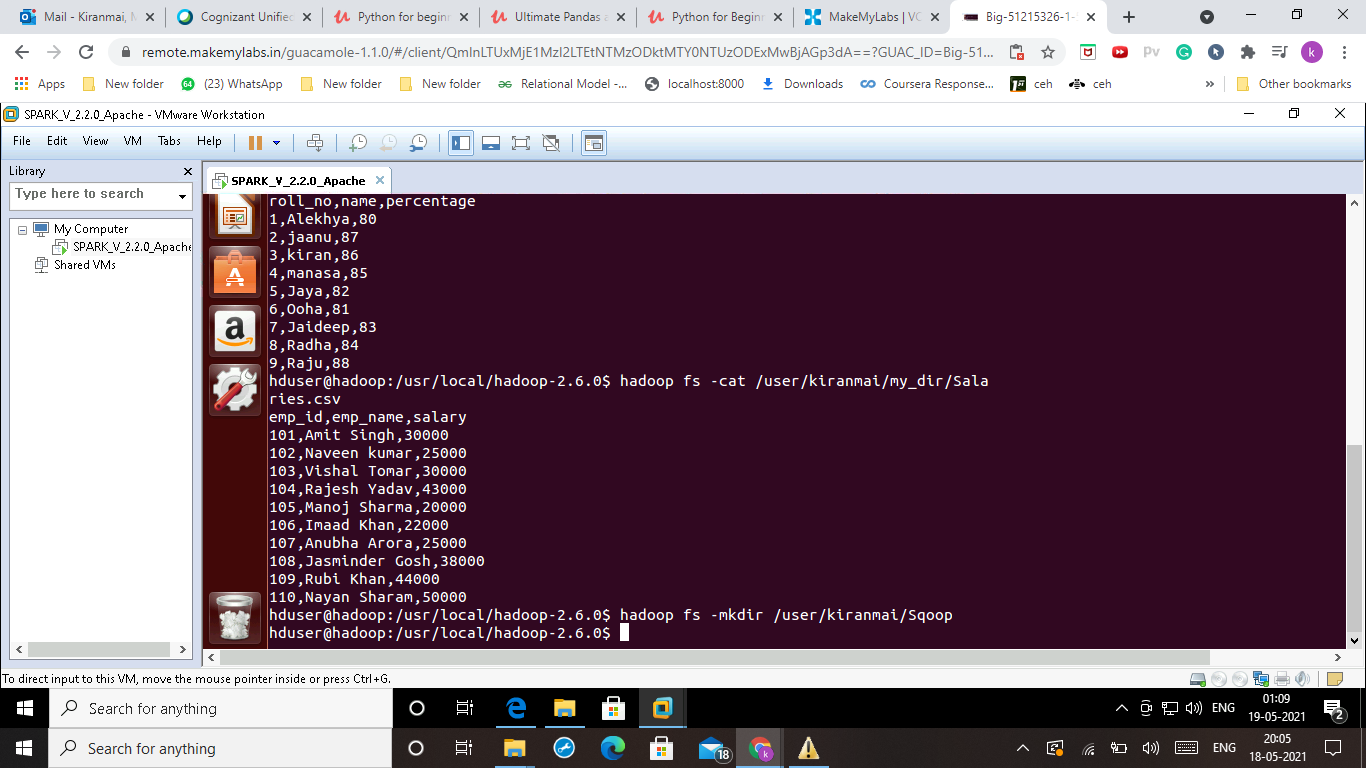


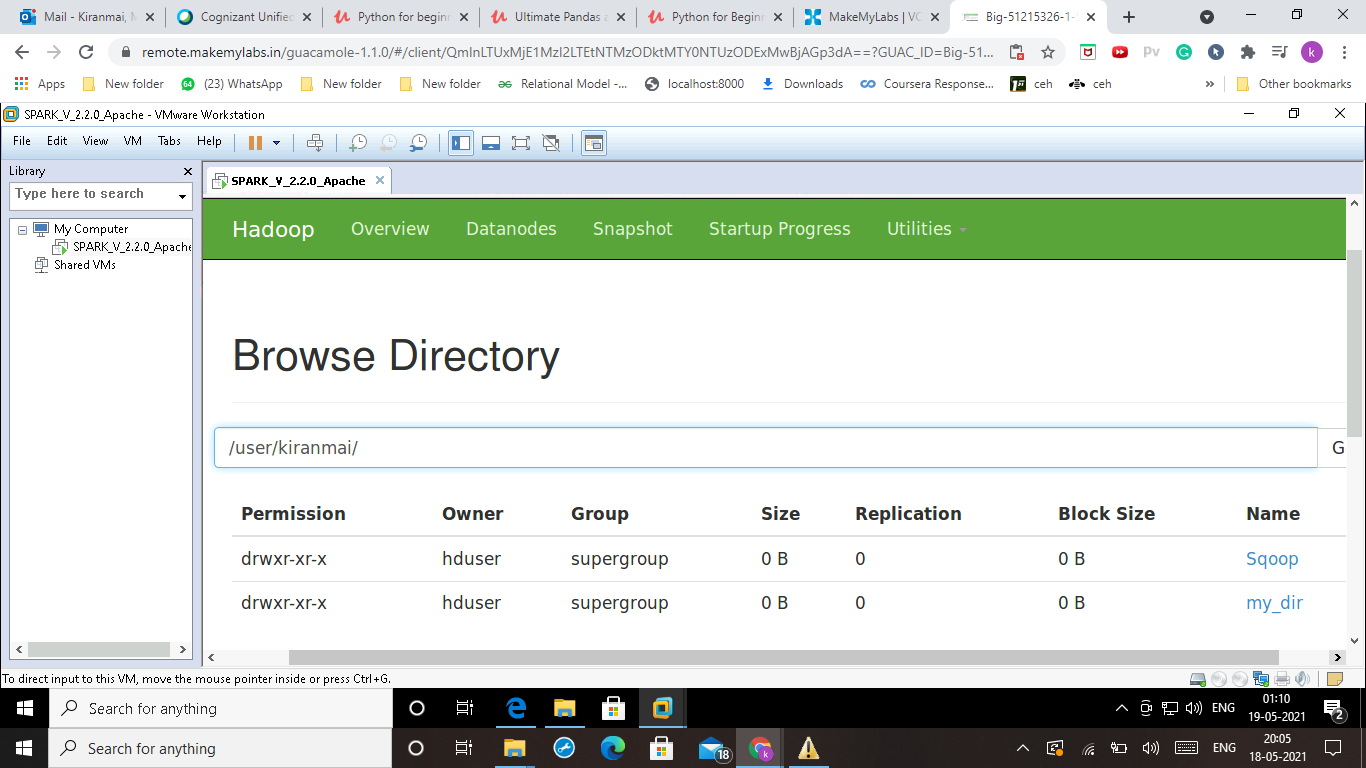
6.Display content of your file from your directory on HDFS to your local machine





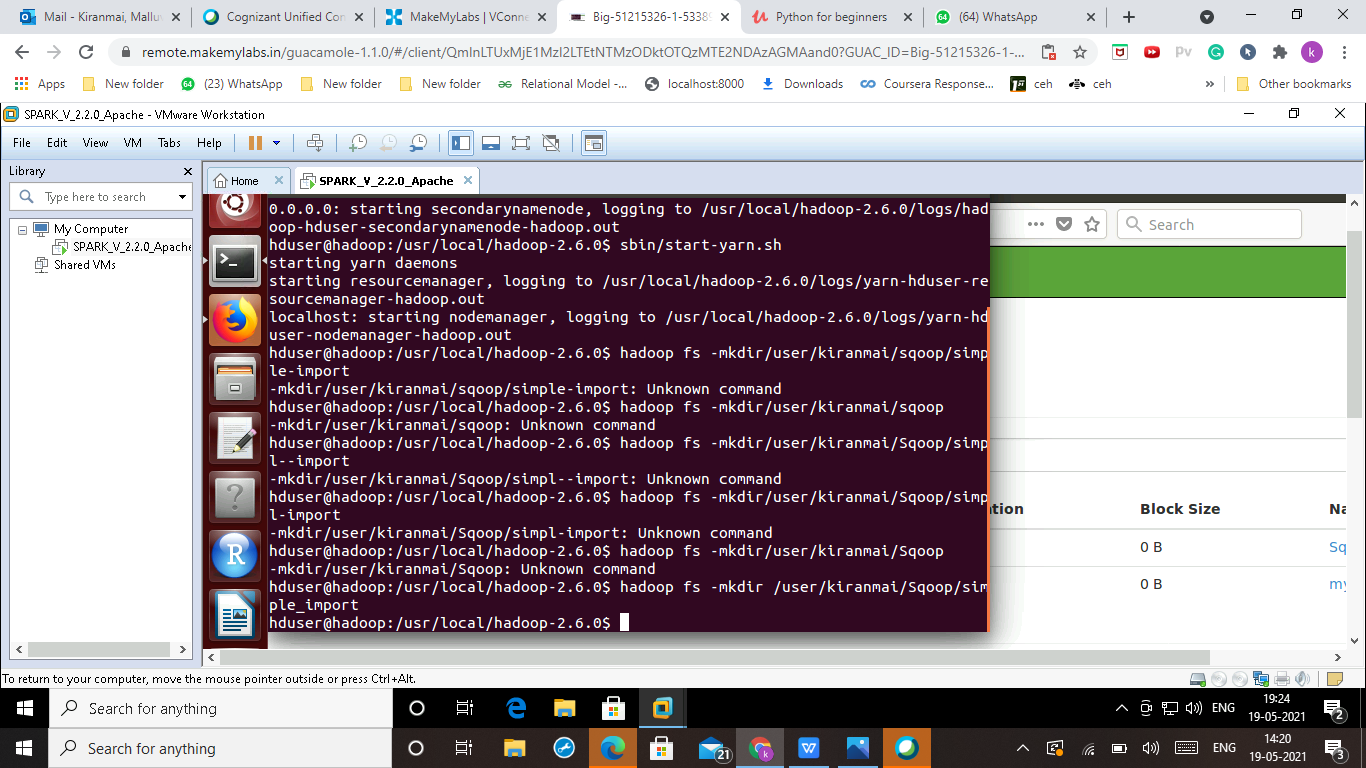
7.Create directory names “sqoop” under your directoty in HDFS

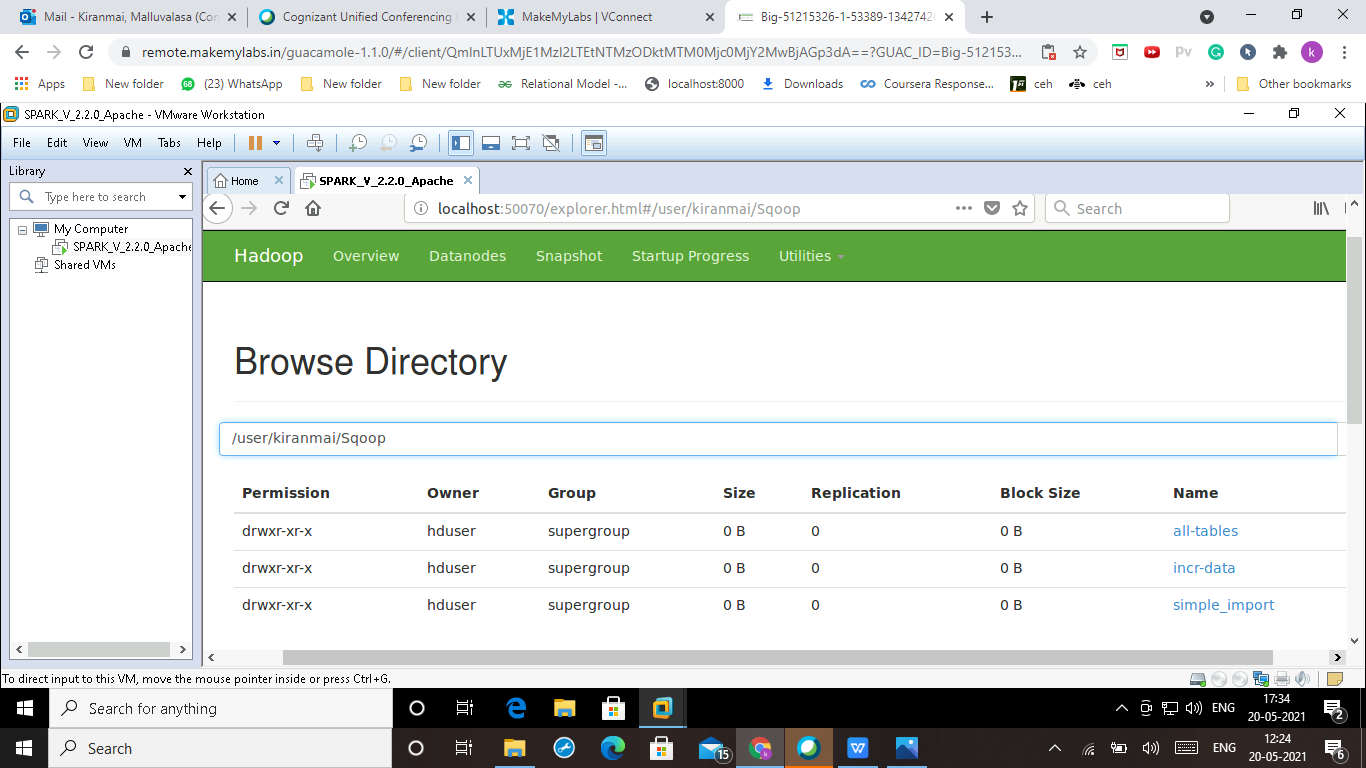




SQOOP:

1. Under sqoop directory on HDFS in your folder, create a new directories called “simple-import” and ”all-tables” and “incr-data”





1. Connect to mysql installed on your local machine and create a database called “retail\_db”. Choose retail\_db in mysql and execute all commands in “retail\_db.sql” file to create tables with data.

🡪Commands:-

mysql -u root -p

password: root

mysql> create database retail\_db;

mysql> use retail\_db;

mysql> CREATE TABLE categories (

category\_id int(11) NOT NULL AUTO\_INCREMENT,

category\_department\_id int(11) NOT NULL,

category\_name varchar(45) NOT NULL,

PRIMARY KEY (category\_id)

);

mysql> insert into categories values ((1, 2, ‘Football’),

(2, 2, ‘Soccer’), (3, 2, ‘Baseball’), (4, 2, ‘Basketball’),

(5, 2, ‘Lacrosse’), ),(6,2,'Tennis & Racquet'),(7,2,'Hockey'),

(8,2,'More Sports'),(9,3,'Cardio Equipment'),(10,3,'Strength Training');

mysql> CREATE TABLE customers(

customer\_id int(11) NOT NULL AUTO\_INCREMENT,

customer\_fname varchar(45) NOT NULL,

customer\_lname varchar(45) NOT NULL,

customer\_email varchar(45) NOT NULL,

customer\_password varchar(45) NOT NULL,

customer\_street varchar(255) NOT NULL,

customer\_city varchar(45) NOT NULL,

customer\_state varchar(45) NOT NULL,

customer\_zipcode varchar(45) NOT NULL,

PRIMARY KEY (customer\_id)

);

mysql> INSERT INTO customers VALUES

(1,'Richard','Hernandez','XXXXXXXXX','XXXXXXXXX','6303 Heather Plaza','Brownsville','TX','78521'),(2,'Mary','Barrett','XXXXXXXXX','XXXXXXXXX','9526 Noble Embers Ridge','Littleton','CO','80126'),(3,'Ann','Smith','XXXXXXXXX','XXXXXXXXX','3422 Blue Pioneer Bend','Caguas','PR','00725'),(4,'Mary','Jones','XXXXXXXXX','XXXXXXXXX','8324 Little Common','San Marcos','CA','92069'),(5,'Robert','Hudson','XXXXXXXXX','XXXXXXXXX','10 Crystal River Mall ','Caguas','PR','00725'),(6,'Mary','Smith','XXXXXXXXX','XXXXXXXXX','3151 Sleepy Quail Promenade','Passaic','NJ','07055'),(7,'Melissa','Wilcox','XXXXXXXXX','XXXXXXXXX','9453 High Concession','Caguas','PR','00725'),(8,'Megan','Smith','XXXXXXXXX','XXXXXXXXX','3047 Foggy Forest Plaza','Lawrence','MA','01841'),(9,'Mary','Perez','XXXXXXXXX','XXXXXXXXX','3616 Quaking Street','Caguas','PR','00725'),(10,'Melissa','Smith','XXXXXXXXX','XXXXXXXXX','8598 Harvest Beacon Plaza','Stafford','VA','22554');

mysql> CREATE TABLE departments (

department\_id int(11) NOT NULL AUTO\_INCREMENT,

department\_name varchar(45) NOT NULL,

PRIMARY KEY (department\_id)

);

mysql> INSERT INTO departments VALUES (2,'Fitness'),(3,'Footwear'),(4,'Apparel'),(5,'Golf'),(6,'Outdoors'),(7,'Fan Shop');

mysql> CREATE TABLE order\_items(

order\_item\_id int(11) NOT NULL AUTO\_INCREMENT,

order\_item\_order\_id int(11) NOT NULL,

order\_item\_product\_id int(11) NOT NULL,

order\_item\_quantity tinyint(4) NOT NULL,

order\_item\_subtotal float NOT NULL,

order\_item\_product\_price float NOT NULL,

PRIMARY KEY (order\_item\_id)

);

mysql> INSERT INTO order\_items VALUES (1,1,957,1,299.98,299.98),(2,2,1073,1,199.99,199.99),(3,2,502,5,250,50),(4,2,403,1,129.99,129.99),(5,4,897,2,49.98,24.99),(6,4,365,5,299.95,59.99),(7,4,502,3,150,50),(8,4,1014,4,199.92,49.98),(9,5,957,1,299.98,299.98),(10,5,365,5,299.95,59.99);

mysql> CREATE TABLE orders (

order\_id int(11) NOT NULL AUTO\_INCREMENT,

order\_date datetime NOT NULL,

order\_customer\_id int(11) NOT NULL,

order\_status varchar(45) NOT NULL,

PRIMARY KEY (order\_id)

);

mysql> INSERT INTO orders VALUES (1,'2013-07-25 00:00:00',11599,'CLOSED'),(2,'2013-07-25 00:00:00',256,'PENDING\_PAYMENT'),(3,'2013-07-25 00:00:00',12111,'COMPLETE'),(4,'2013-07-25 00:00:00',8827,'CLOSED'),(5,'2013-07-25 00:00:00',11318,'COMPLETE'),(6,'2013-07-25 00:00:00',7130,'COMPLETE'),(7,'2013-07-25 00:00:00',4530,'COMPLETE'),(8,'2013-07-25 00:00:00',2911,'PROCESSING'),(9,'2013-07-25 00:00:00',5657,'PENDING\_PAYMENT'),(10,'2013-07-25 00:00:00',5648,'PENDING\_PAYMENT');

mysql> CREATE TABLE products (

product\_id int(11) NOT NULL AUTO\_INCREMENT,

product\_category\_id int(11) NOT NULL,

product\_name varchar(45) NOT NULL,

product\_description varchar(255) NOT NULL,

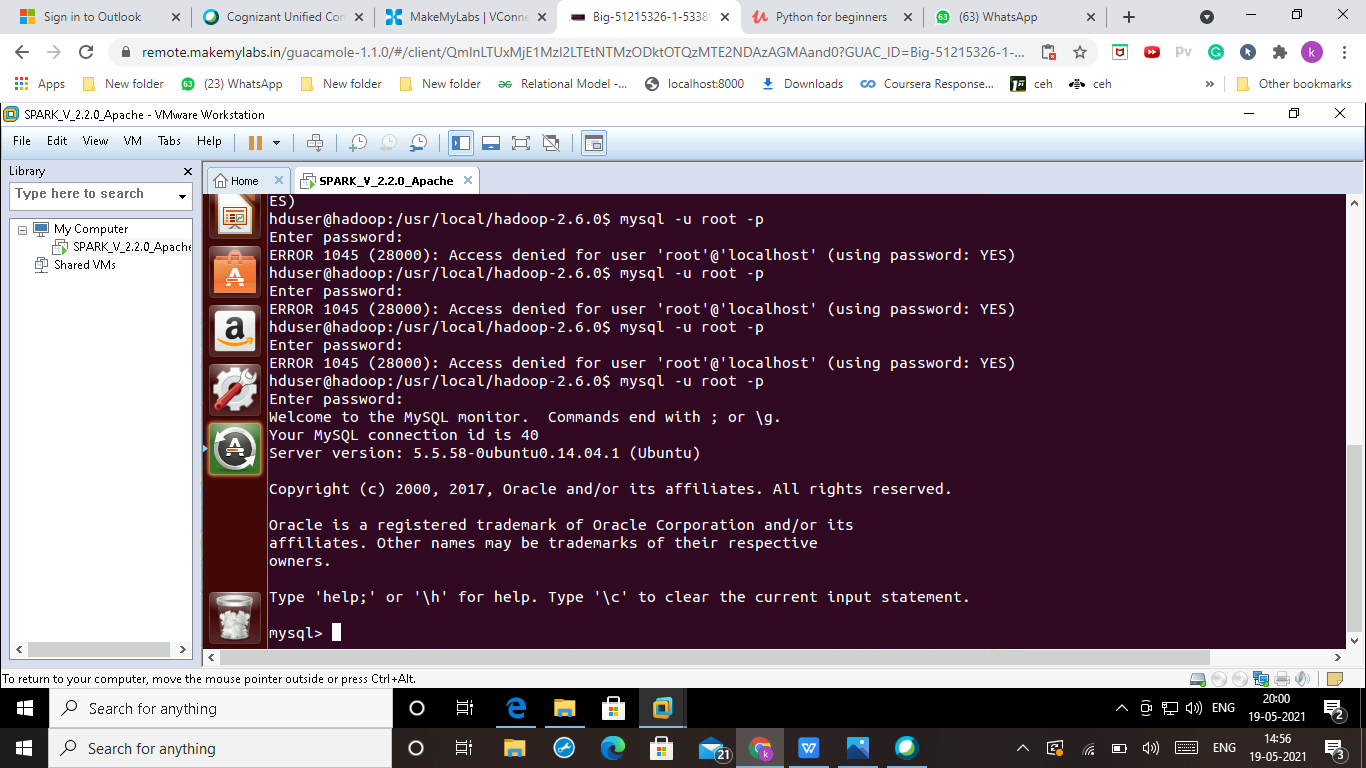
product\_price float NOT NULL,

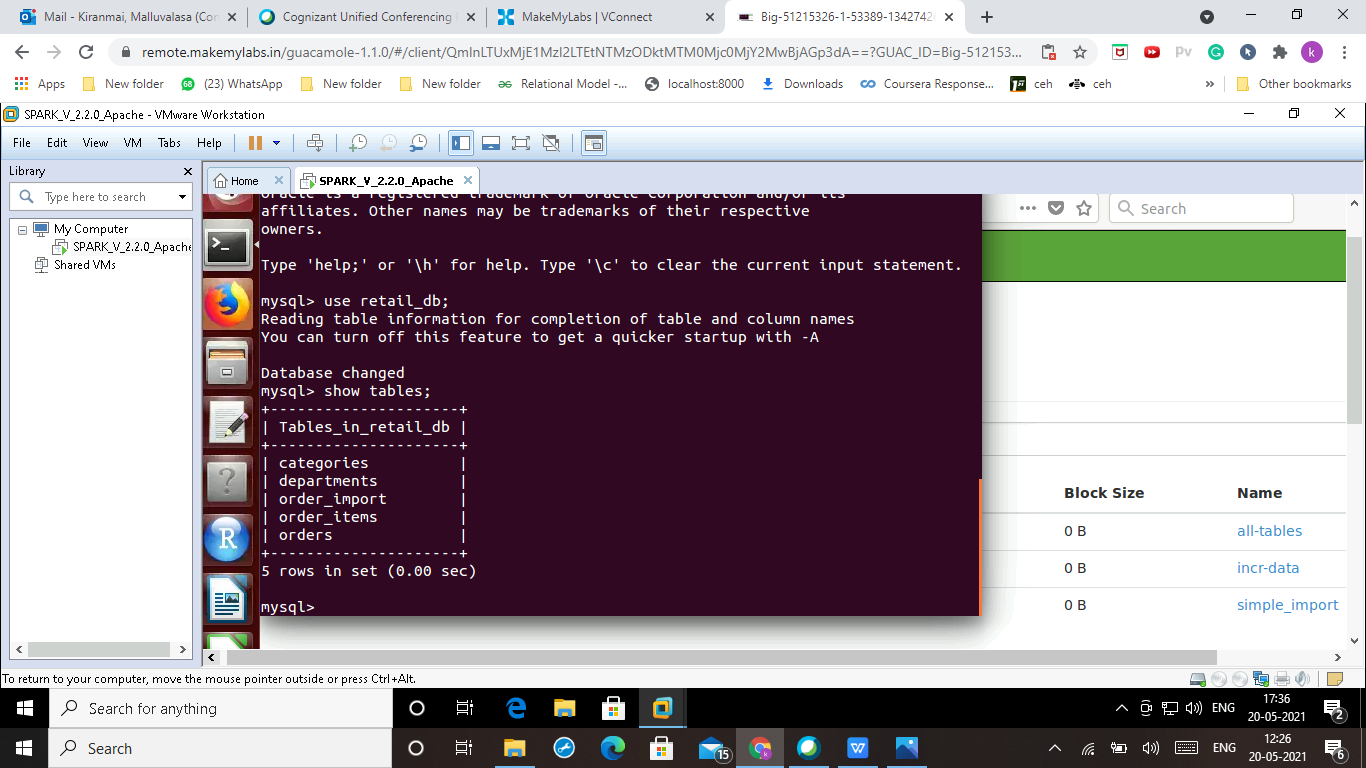
product\_image varchar(255) NOT NULL,

PRIMARY KEY (product\_id)

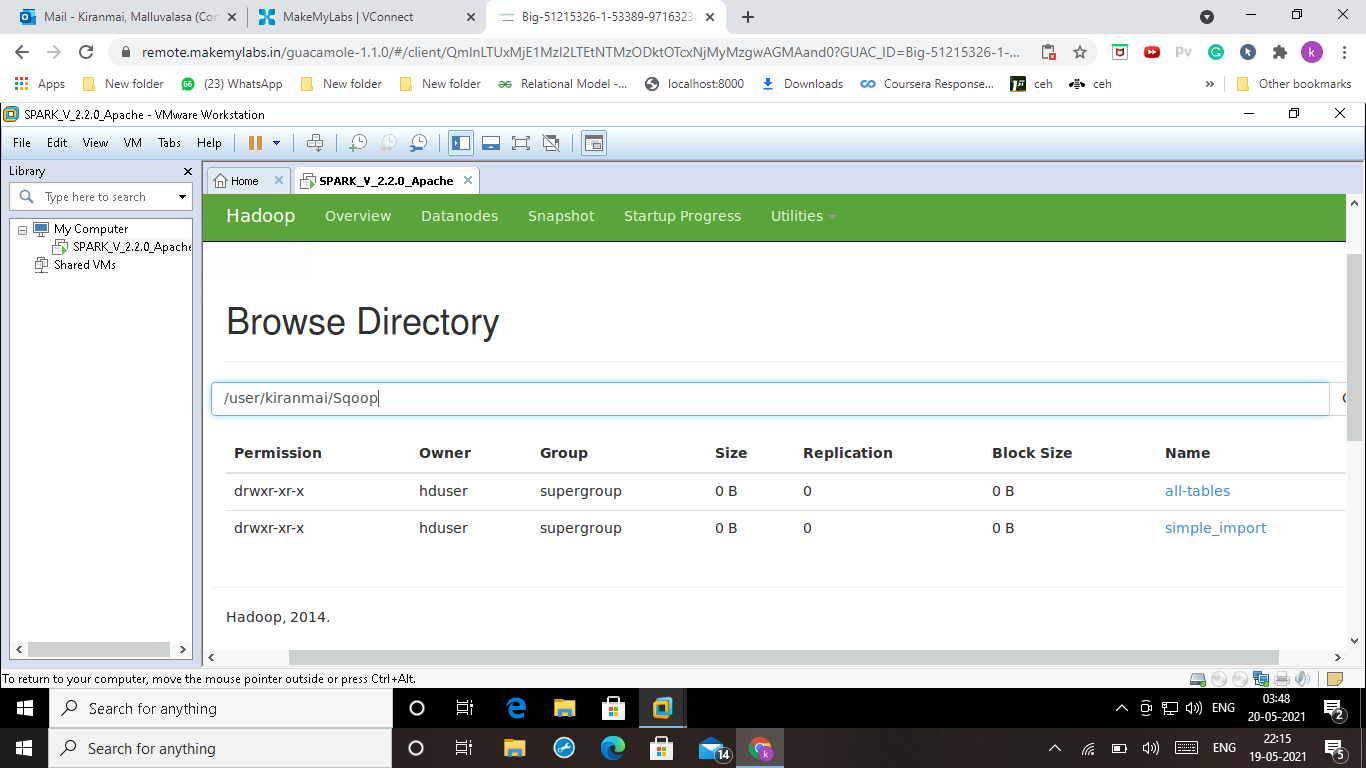
);

mysql> INSERT INTO products VALUES (1,2,'Quest Q64 10 FT. x 10 FT. Slant Leg Instant U','',59.98,'http://images.acmesports.sports/Quest+Q64+10+FT.+x+10+FT.+Slant+Leg+Instant+Up+Canopy'),(2,2,'Under Armour Men\'s Highlight MC Football Clea','',129.99,'http://images.acmesports.sports/Under+Armour+Men%27s+Highlight+MC+Football+Cleat'),(3,2,'Under Armour Men\'s Renegade D Mid Football Cl','',89.99,'http://images.acmesports.sports/Under+Armour+Men%27s+Renegade+D+Mid+Football+Cleat'),(4,2,'Under Armour Men\'s Renegade D Mid Football Cl','',89.99,'http://images.acmesports.sports/Under+Armour+Men%27s+Renegade+D+Mid+Football+Cleat'),(5,2,'Riddell Youth Revolution Speed Custom Footbal','',199.99,'http://images.acmesports.sports/Riddell+Youth+Revolution+Speed+Custom+Football+Helmet');

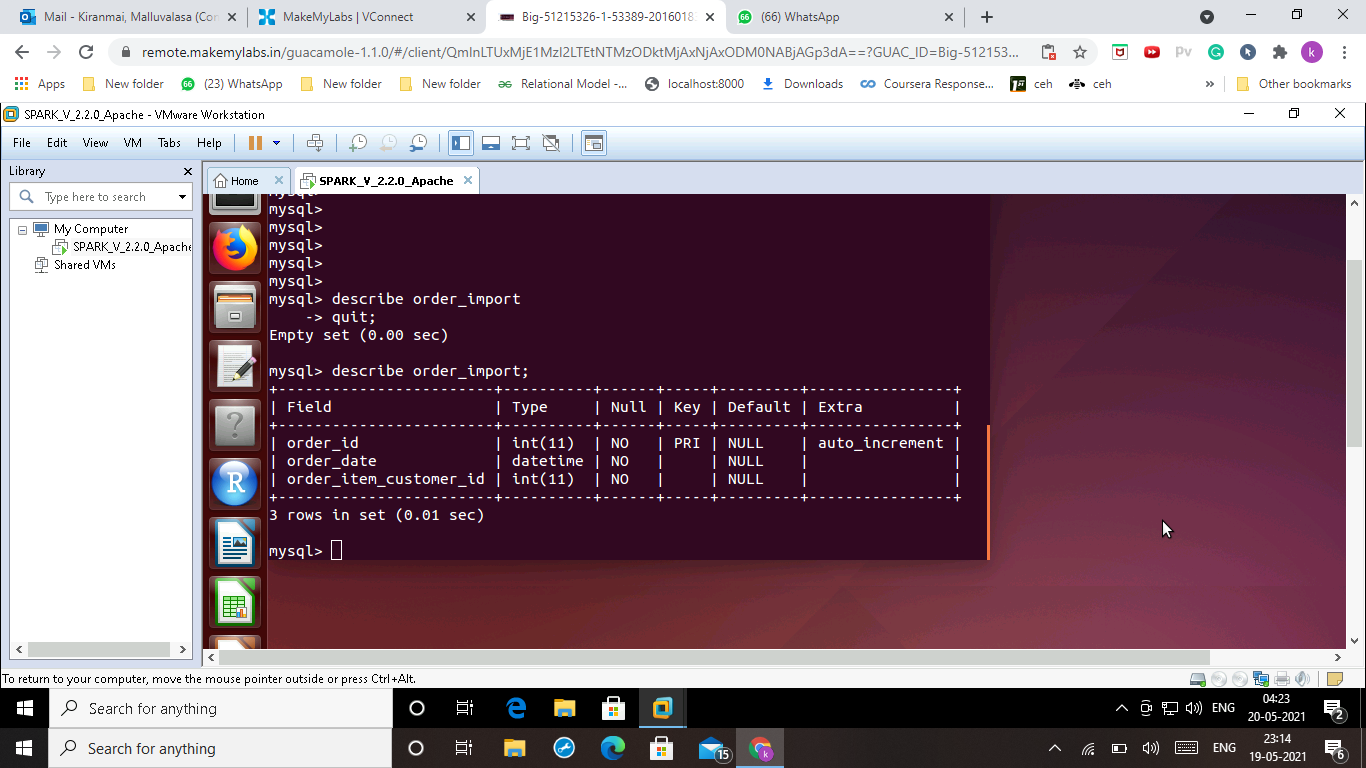




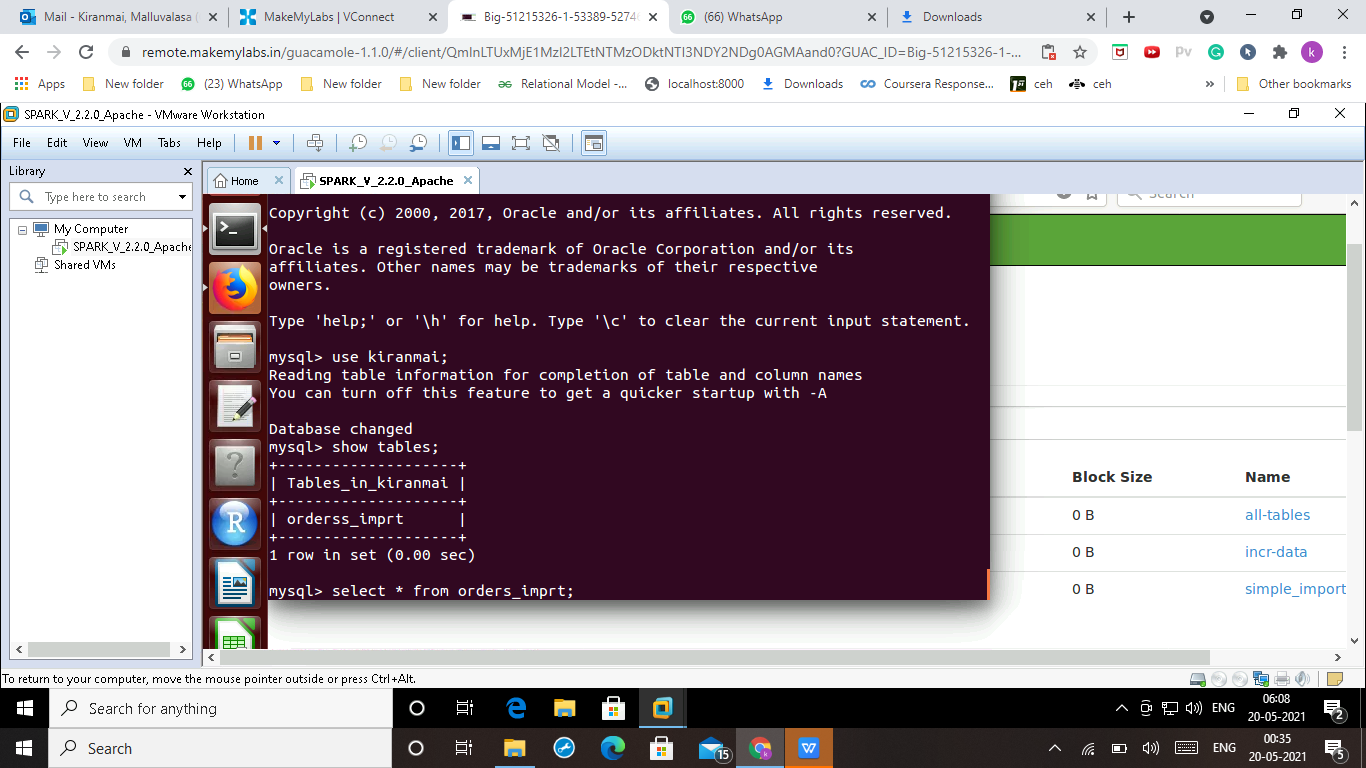
1. Using sqoop import orders table under “simple-import” directory on HDFS.

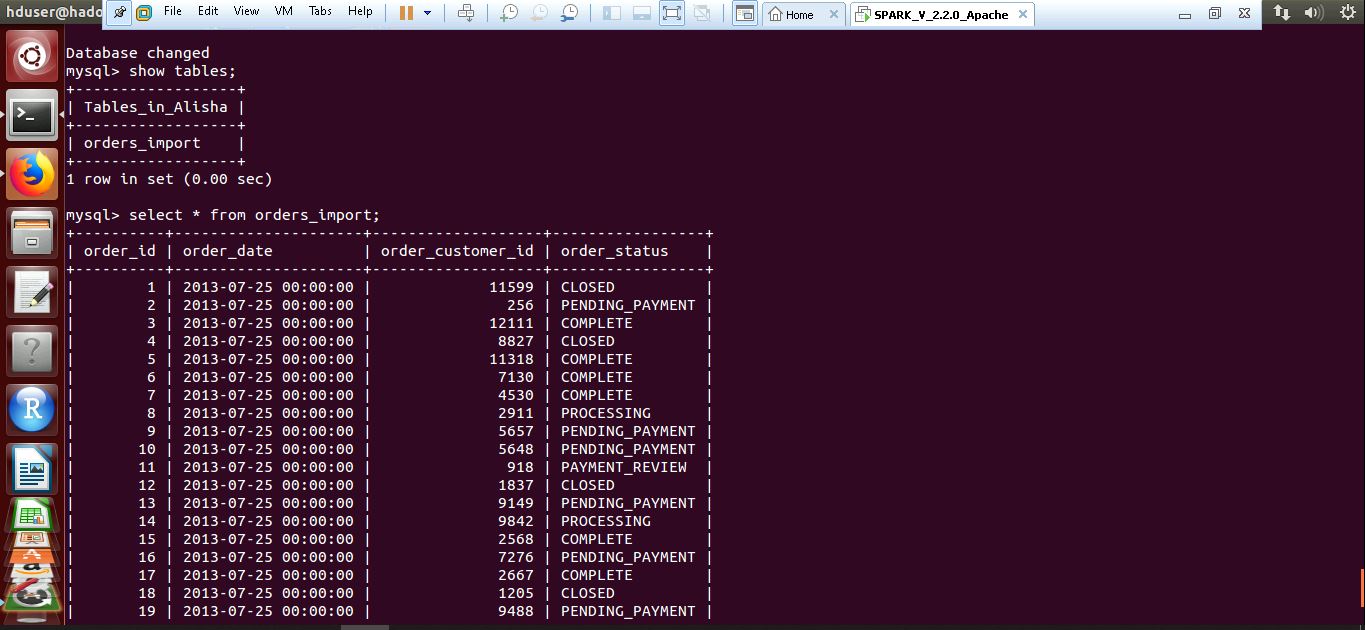


1. Connect to mysql and create a database with your name & create orders\_import table using same DDL as orders table

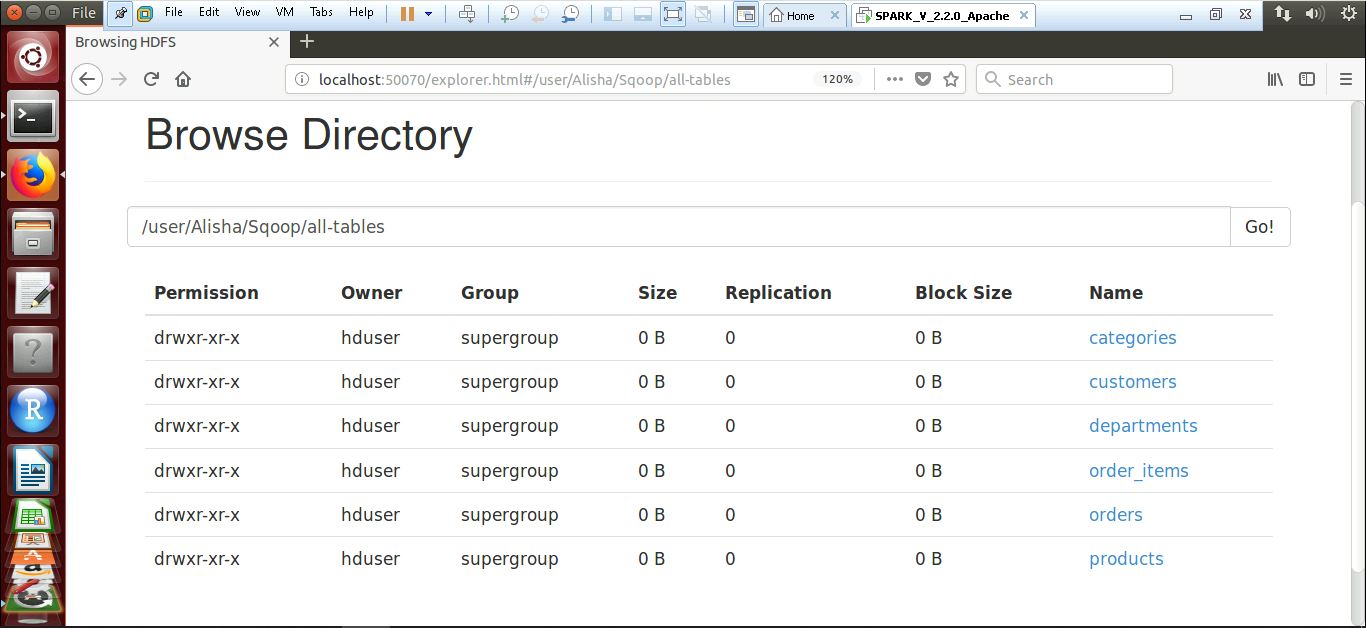


1. Using sqoop , export order data from HDFS under “simple-import” directory into your order\_import tables in your database

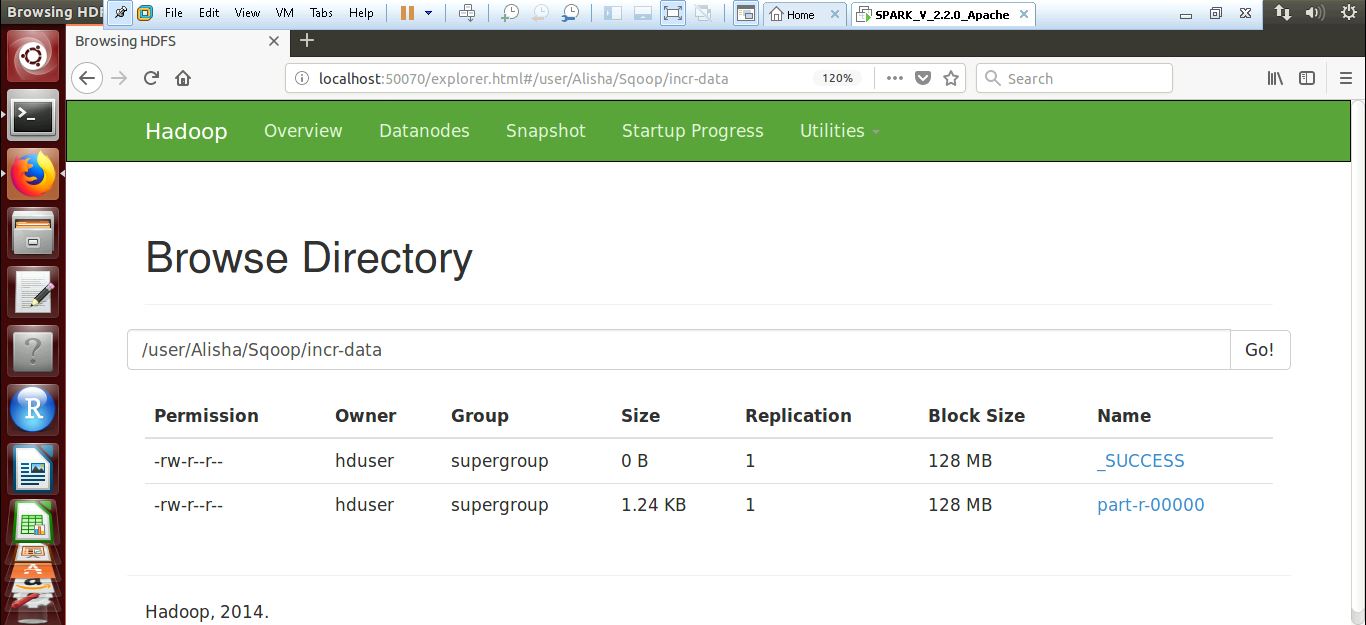




1. Using sqoop import-all-tables , dump retail\_db in mysql in to “all-tables” directory. Verify data on HDFS



1. Create a sqoop job to import incremental data from orders table in to “incr-data” folder on HDFS



1. Execute sqoop job for 3 times to load incremental data. Make sure to add/update records in retail\_db.orders table before running job.

mysql> insert into orders values(32, '2013-08-14', 51, 'CLOSED');

mysql>exit;

sqoop import --connect jdbc:mysql://localhost/retail\_db --username root --password root --table orders --incremental lastmodified --check-column order\_date --last-value "2013-07-25 00:00:00" --merge-key order\_id --target-dir /user/kiranmai/Sqoop/incr-data

mysql> insert into orders values(33, '2013-09-14', 52, 'PROCESSING');

mysql>exit;

sqoop import --connect jdbc:mysql://localhost/retail\_db --username root --password root --table orders --incremental lastmodified --check-column order\_date --last-value "2013-08-14 00:00:00" --merge-key order\_id --target-dir /user/kiranmai/Sqoop/incr-data

mysql> insert into orders values(34, '2013-10-14', 52, 'COMPLETED');

mysql>exit;

sqoop import --connect jdbc:mysql://localhost/retail\_db --username root --password root --table orders --incremental lastmodified --check-column order\_date --last-value "2013-09-14 00:00:00" --merge-key order\_id --target-dir /user/kiranmai/Sqoop/incr-data