Write an application using HBase and HiveQL for Customer information system which will include

a. Creation of –Cutomer_info(Cust-ID,Cust-Name,orderID), order_info(OrderID,ItemID,Quantity), item_info(Item-ID,Item-Name,ItemPrice) tables in Hive

```
CREATE TABLE customer_info (
  cust_id INT,
  cust_name STRING,
  order_id INT
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';

CREATE TABLE order_info (
  order_id INT,
  item_id INT,
  quantity INT
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';

CREATE TABLE item_info (
  item_id INT,
  item_name STRING,
  item_price FLOAT
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';
```

b. Load table with data from local storage in Hive.

Open text editor, enter values for each table separated by commas ‘,’

Save as .csv file

customer.csv

1,Alice,101

2,Bob,102

3,Charlie,103

4,Diana,104

5,Edward,105

order.csv

101,201,2

102,202,1

103,203,5

104,204,3

105,201,1

item.csv

201,Laptop,50000

202,Mouse,500

203,Keyboard,1200

204,Monitor,15000

205,USB Cable,150

Hive>

```
LOAD DATA LOCAL INPATH '/home/cloudera/customer.csv' INTO TABLE customer_info;
LOAD DATA LOCAL INPATH '/home/cloudera/order.csv' INTO TABLE order_info;
LOAD DATA LOCAL INPATH '/home/cloudera/item.csv' INTO TABLE item_info;
```

c. Perform Join tables with Hive

```
SELECT
  ci.cust_id,
  ci.cust_name,
  oi.order_id,
  ii.item_name,
  ii.item_price,
  oi.quantity,
  (ii.item_price * oi.quantity) AS total_cost
FROM customer_info ci
JOIN order_info oi ON ci.order_id = oi.order_id
JOIN item_info ii ON oi.item_id = ii.item_id;
```

d. Create Index on Customer information system in Hive.

```
CREATE INDEX idx_cust_id
ON TABLE customer_info (cust_id)
AS 'COMPACT'
WITH DEFERRED REBUILD;

ALTER INDEX idx_cust_id ON customer_info REBUILD;
```

e. Find the total, average sales in Hive

```
SELECT
   SUM(ii.item_price * oi.quantity) AS total_sales,
   AVG(ii.item_price * oi.quantity) AS avg_sales
FROM order_info oi
JOIN item_info ii ON oi.item_id = ii.item_id;
```

f. Find Order details with maximum cost.

```
SELECT
   ci.cust_id,
   ci.cust_name,
   oi.order_id,
   ii.item_name,
   (ii.item_price * oi.quantity) AS total_cost
FROM customer_info ci
JOIN order_info oi ON ci.order_id = oi.order_id
JOIN item_info ii ON oi.item_id = ii.item_id
ORDER BY total_cost DESC
LIMIT 1;
```

g. Creating an external Hive table to connect to the HBase for Customer Information System.

hbase shell

```
create 'cust_hbase', 'info'
put 'cust_hbase', '1', 'info:cust_name', 'Alice'
put 'cust_hbase', '1', 'info:order_id', '101'

CREATE EXTERNAL TABLE hbase_customer (
   key STRING,
   cust_name STRING,
   order_id STRING
)
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
WITH SERDEPROPERTIES (
  "hbase.columns.mapping" = ":key,info:cust_name,info:order_id"
)
TBLPROPERTIES("hbase.table.name" = "cust_hbase");
```

h. Display records of Customer Information Table in Hbase.

hbase shell

```
scan 'cust_hbase'
```