Online Retail Application using HBase and HiveQL

i. Create and Load table with Online Retail data in Hive

```
CREATE TABLE online_retail (
    InvoiceNo STRING,
    StockCode STRING,
    Description STRING,
    Quantity INT,
    InvoiceDate STRING,
    UnitPrice DOUBLE,
    CustomerID STRING,
    Country STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;

LOAD DATA INPATH '/user/hive/warehouse/online_retail.csv' INTO TABLE online_retail;
```

j. Create Index on Online Retail Table in Hive

```
CREATE INDEX idx_customer_id ON TABLE online_retail (CustomerID)
AS 'COMPACT'
WITH DEFERRED REBUILD;

ALTER INDEX idx_customer_id ON online_retail REBUILD;
```

k. Find the Total and Average Sales in Hive

```
SELECT
    SUM(Quantity * UnitPrice) AS total_sales,
    AVG(Quantity * UnitPrice) AS average_sales
FROM online_retail;
```

l. Find Order Details with Maximum Cost

```
SELECT InvoiceNo, SUM(Quantity * UnitPrice) AS order_total
FROM online_retail
GROUP BY InvoiceNo
ORDER BY order_total DESC
LIMIT 1;
```

m. Find Customer Details with Maximum Order Total

```
SELECT CustomerID, SUM(Quantity * UnitPrice) AS customer_total
FROM online_retail
GROUP BY CustomerID
ORDER BY customer_total DESC
LIMIT 1;
```

n. Find Country with Maximum and Minimum Sale

```
-- Maximum
SELECT Country, SUM(Quantity * UnitPrice) AS total_sales
FROM online_retail
GROUP BY Country
ORDER BY total_sales DESC
LIMIT 1;

-- Minimum
SELECT Country, SUM(Quantity * UnitPrice) AS total_sales
FROM online_retail
GROUP BY Country
HAVING total_sales > 0
ORDER BY total_sales ASC
LIMIT 1;
```

o. Creating an External Hive Table to Connect to the HBase for OnlineRetail

```
-- In HBase shell:
create 'online_retail_hbase', 'meta', 'details'

-- In Hive:
CREATE EXTERNAL TABLE online_retail_hive (
    rowkey STRING,
    InvoiceNo STRING,
    StockCode STRING,
    Description STRING,
    Quantity INT,
    InvoiceDate STRING,
    UnitPrice DOUBLE,
    CustomerID STRING,
    Country STRING
)
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
WITH SERDEPROPERTIES (
                                              "hbase.columns.mapping"                          =
":key,meta:InvoiceNo,meta:StockCode,meta:Description,details:Quantity,details:InvoiceDate,details:UnitPrice,details:Cu
stomerID,details:Country"
)
TBLPROPERTIES (
```

Online Retail Application using HBase and HiveQL

```
  "hbase.table.name" = "online_retail_hbase"
);
```

p. Display Records of OnlineRetail Table in HBase

```
-- From HBase shell:
scan 'online_retail_hbase', {LIMIT => 10}
```