Challenge in Hive:-
==========

Creating CSV table :-

```
create table sales_order_data_csv
(
ORDERNUMBER int,
QUANTITYORDERED int,
PRICEEACH float,
ORDERLINENUMBER int,
SALES float,
STATUS string,
QTR_ID int,
MONTH_ID int,
YEAR_ID int,
PRODUCTLINE string,
MSRP int,
PRODUCTCODE string,
PHONE string,
CITY string,
STATE string,
POSTALCODE string,
COUNTRY string,
TERRITORY string,
CONTACTLASTNAME string,
CONTACTFIRSTNAME string,
DEALSIZE string
)
row format delimited
fields terminated by ','
tblproperties("skip.header.line.count"="1");
```

Loading data into csv table:-

```
load data local inpath ‘file:///tmp/hive_class_1/vehical_sales_data.txt’ into table
sales_order_data_csv;
```

Creating orc table:-

```
create table vehical_sales_order_data_orc_v1
(
ORDERNUMBER int,
```

QUANTITYORDERED int,
PRICEEACH float,
ORDERLINENUMBER int,
SALES float,
STATUS string,
QTR_ID int,
MONTH_ID int,
YEAR_ID int,
PRODUCTLINE string,
MSRP int,
PRODUCTCODE string,
PHONE string,
CITY string,
STATE string,
POSTALCODE string,
COUNTRY string,
TERRITORY string,
CONTACTLASTNAME string,
CONTACTFIRSTNAME string,
DEALSIZE string
)
stored as orc;

Loading data into orc table using csv table data:-

from sales_order_data_csv insert overwrite table vehical_sales_order_data_orc_v1 select *;

Performing Operations using vehical_sales_order_data_orc_v1 table:-

1.Calculate Total Sales per year:

> select year_id, sum(sales) as total_sales from vehical_sales_order_data_orc_v1 group by year_id;

```
hive> select year_id, sum(sales) as total_sales from vehical_sales_order_data_orc_v1 group by year_id;
```

Output:-

```
OK
year_id total_sales
2003    3516979.547241211
2004    4724162.593383789
2005    1791486.7086791992
Time taken: 22.827 seconds, Fetched: 3 row(s)
```

2.Find a product for which maximum orders were placed.

>select productline, count(quantityordered) as total_ordered from vehical_sales_ordered_data_orc_v1 group by productline order by total_ordered desc limit 1;

```
hive> select productline, count(quantityordered) as total_ordered from vehical_sales_order_data_orc_v1 group by productline order by total_ordered desc limit 1;
```

Output:-

```
OK
Classic Cars    967
Time taken: 41.36 seconds, Fetched: 1 row(s)
hive>
```

3.Calculate the total sales for each quarter.

>select qtr_id,sum(sales) as total_sales from vehical_sales_order_data_orc_v1 group by qtr_id;

```
hive> select qtr_id,sum(sales) as total_sales from vehical_sales_order_data_orc_v1 group by qtr_id ;
```

```
OK
1       2350817.726501465
2       2048120.3029174805
3       1758910.808959961
4       3874780.010925293
Time taken: 16.168 seconds, Fetched: 4 row(s)
```

4.In which quarter sales was minimum

>select qtr_id,sum(sales) as total_sales from vehical_sales_order_data_orc_v1 group by qtr_id order by total_sales limit 1;

```
hive> select qtr_id,sum(sales) as total_sales from vehical_sales_order_data_orc_v1 group by qtr_id order by total_sales limit 1;
```

```
OK
3       1758910.808959961
Time taken: 44.978 seconds, Fetched: 1 row(s)
```

5.In which country sales was maximum and in which country sales was minimum

>select country,max(sales) as sales_data from vechical_sales_order_data_orc_v1 group by country order by sales_data desc limit 1
union all
select country,min(sales) as sales_data from vechical_sales_order_data_orc_v1 group by country order by sales_data limit 1

```
hive> select country,max(sales) as sales_data from vehical_sales_order_data_orc_v1 group by country order by sales_data desc limit 1
    > union all
    > select country,min(sales) as sales_data from vehical_sales_order_data_orc_v1 group by country order by sales_data limit 1;
```

```
OK
France  482.13
USA     14082.8
Time taken: 100.509 seconds, Fetched: 2 row(s)
```

6.Calculate quarterly sales for each city.

>select city,qtr_id,sum(sales) from vehical_sales_order_data_orc_v1 group by city,qtr_id;

```
hive> select city,qtr_id,sum(sales) from vehical_sales_order_data_orc_v1 group by city,qtr_id;
```

```
Pasadena        1       44273.359436035156
Pasadena        3       55776.119873046875
Pasadena        4       4512.47998046875
Philadelphia    1       27398.820434570312
Philadelphia    2       7287.240234375
Philadelphia    4       116503.07043457031
Reggio Emilia   2       41509.94006347656
Reggio Emilia   3       56421.650390625
Reggio Emilia   4       44669.740478515625
Reims   1       52029.07043457031
Reims   2       18971.959716796875
Reims   3       15146.31982421875
Reims   4       48895.59014892578
Salzburg        2       98104.24005126953
Salzburg        3       6693.2802734375
Salzburg        4       45001.10986328125
San Diego       1       87489.23010253906
San Francisco   1       72899.19995117188
San Francisco   4       151459.4805908203
San Jose        2       160010.27026367188
San Rafael      1       267315.2586669922
San Rafael      2       7261.75
San Rafael      3       216297.40063476562
San Rafael      4       163983.64880371094
Sevilla 4       54723.621154785156
Singapore       1       28395.18994140625
Singapore       2       92033.77014160156
Singapore       3       90250.07995605469
Singapore       4       77809.37023925781
South Brisbane  1       21730.029907226562
South Brisbane  3       10640.290161132812
South Brisbane  4       27098.800048828125
Stavern 1       54701.999755859375
Stavern 4       61897.19006347656
Strasbourg      2       80438.47985839844
Torino  3       94117.25988769531
Toulouse        1       15139.1201171875
Toulouse        3       17251.08056640625
Toulouse        4       38098.240234375
Tsawassen       2       31302.500244140625
Tsawassen       3       43332.349609375
Vancouver       4       75238.91955566406
Versailles      1       5759.419921875
Versailles      4       59074.90026855469
White Plains    4       85555.98962402344
Time taken: 19.794 seconds, Fetched: 182 row(s)
hive>
```

7.Find a month for each year in which the maximum number of quantities were sold.

```
>select year_id,month_id,sum(sales) as max_sales from
vehical_sales_order_data_orc_v1 group by year_id,month_id having
year_id=2003 order by max_sales desc limit 1
Union all
select year_id,month_id,sum(sales) as max_sales from
vehical_sales_order_data_orc_v1 group by year_id,month_id having
year_id=2004 order by max_sales desc limit 1
Union all
select year_id,month_id,sum(sales) as max_sales from
vehical_sales_order_data_orc_v1 group by year_id,month_id having
year_id=2005 order by max_sales desc limit 1
```