**CAPSTONE PROJECT**

**PROJECT ONE: SALES PERFORMANCE ANALYSIS FOR A RETAIL STORE**

* Open the project in excel
* Rename each sheet to the title of the project. Sheet 1 to Sales performance project and sheet 2 to Customer Segmentation project.
* On Sales performance project (SPP), I removed all the repeated data by selecting an active cell and then on the excel pane, selecting the Data icon, and clicking on remove duplicate to remove all the duplicated data. The same process is repeated for Customer Segmentation Project (CSP).
* I calculated the total revenue by multiplying the quantity by price. I, then created a Pivot table by selecting Insert on the Excel browser, then clicking on Pivot table. I had selected an active cell or chose a cell on the data being analyze, this will ensure that the pivot table is created for the data provided.
* Pivot tables created are: Total sales by product, total sales by region, total sales by monthly summary, others include total sales by region by revenue generated (this will help to know the amount each region generated, the region with the highest and lowest revenue for example are South and West respectively), Total sales by region, by product, by revenue (here we can view the products with the highest sales and which region this products is highly demanded for most for example: Shoes are the most demanded product in the south, while in the North, shirt is the most move-able product. And finally the product, the total quantity, and total revenue. (Similar steps was done to Customer Distribution Data).

SQL:

PROJECT 1

select \* from [dbo].[sales data2]

--------

-----TOTAL SALES FOR EACH PRODUCT CATEGORY----

select product,sum(quantity) as 'total sales'

From [dbo].[sales data2]

Group by product;

-----

-------NUMBER OF SALES TRANSACTIONS IN EACH YEAR------

select region, count(\*) as 'transaction sales'

From[dbo].[sales data2]

Group by Region;

------------

-----HIGHEST SELLING PRODUCT BY TOTAL SALES VALUE----

select top 1 product, sum(quantity\*unitprice)

from [dbo].[sales data2]

Group by product;

alter table [dbo].[sales data2]

alter column unitprice bigint;

----------------

-----TOTAL REVENUE PER PRODUCT------

select product, sum(TOTAL\_REVENUE) As 'Total revenue'

from[dbo].[sales data2]

Group by product;

-------------

------MONTHLY SALES TOTAL FOR THE CURRENT YEAR----

SELECT DATENAME(MONTH,ORDERDATE) AS 'MONTHLYSALES',

Sum(quantity) as 'totalsales'

from[dbo].[sales data2]

where year(orderdate)= year(getdate())

group by datename(month,orderdate)

order by totalsales asc;

-------

------TOP FIVE CUSTOMERS BY TOTAL PURCHASE AMOUNT-----

select top 5 Customer\_id, sum(total\_revenue) as 'Total Purchase'

from [dbo].[sales data2]

group by customer\_id;

---------

------PERCENTAGE OF TOTAL SALES CONTRIBUTED BY EACH REGION------

Select Region,

CAST((sum(quantity)\*100/(select sum(quantity) from[dbo].[sales data2])) as varchar(10)) +'%' as salespercentage

from[dbo].[sales data2]

group by region order by salespercentage desc;

------------

-----PRODUCTS WITH NO SALES IN THE LAST QUARTER------

Select Distinct PRODUCT

from[dbo].[sales data2]

where PRODUCT not in (

select product from[dbo].[sales data2]

where orderdate >= DATEADD(quarter,-1,getdate())

);

PROJECT 2

SELECT \* FROM [dbo].[customer data]

---------------------------------

------TOTAL NUMBER OF CUSTOMER PER REGION-----------

Select region, count(\*) as 'no of customer'

from [dbo].[customer data]

group by Region

----------

----THE MOST POPULAR SUBSCRIPTION TYPE BY THE NUMBER OF CUSTOMER---

Select subscriptiontype,

count (customerid) as No\_Of\_Customer

From [dbo].[customer data]

Group by SubscriptionType

Order by No\_Of\_Customer Asc;

-----------

----CUSTOMERS WHO CANCELLED THEIR SUBSCRPTIONS WITHIN 6 MONTHS----

Select CustomerID,

Region,CustomerName

SubscriptionType,

SubscriptionStart,

SubscriptionEnd,

Canceled,

Revenue

From [dbo].[customer data]

WHERE

Canceled='0'

And DATEDIFF(month, SubscriptionStart, SubscriptionEnd) <= 6

Order by SubscriptionEnd Asc;

Select CustomerId

From[dbo].[customer data]

Where Datediff(month, SubscriptionStart, SubscriptionEnd) <= 6;

-------------

----FIND THE AVERAGE SUBSCRIPTION DURATION FOR ALL CUSTOMERS---

Select CustomerName, AVG(SUBSCRIPTION\_DURATION) as 'AVG SUBSCRIPTION DURATION'

From[dbo].[customer data]

Group by CustomerName;

----------

------------

-----FIND CUSTOMERS WITH SUBSCRIPTIONS MORE THAN 12 MONTHS-----

SELECT

CustomerID,

CustomerName,

Region,

SubscriptionType,

SubscriptionStart,

SubscriptionEnd,

Canceled,

Revenue

From [dbo].[customer data]

Where

Canceled = '1'

And DATEDIFF(Month,SubscriptionEnd,SubscriptionStart)>12

Order by

SubscriptionStart;

---------------

-------TOTAL REVENUE BY SUBSCRIPTION TYPE------

SELECT SubscriptionType,

Sum(revenue) as Total\_Revenue

From[dbo].[customer data]

group by SubscriptionType;

---------------

------TOP THREE REGIONS BY SUBSCRIPTION CANCELATION

SELECT Top 3 Region,

Count(\*) as subscriptionEnd\_Count

From[dbo].[customer data]

Where SubscriptionEnd is null

Group by Region

Order by SubscriptionEnd\_count;

-------------------

-------TOTAL NUMBER OF ACTIVE AND CANCELED SUBSCRIPTIONS-----

SELECT

Canceled,

COUNT(\*) TotalSubscriptions

From[dbo].[customer data]

GROUP BY

CANCELED;

TOOLS:

HP LAPTOP

Microsoft excel 2013

Miceosoft word 2013

SQL

Power Bi