* **MYSQL:**

1. Imported the file inside MySQL from files, then executed whole file to load the data.
2. Performed basic exploration by using select statement on all data tables.
3. Viewing and understand the data structure, columns, values.
4. Finding cardinality between different tables using join statement on common columns between them.
5. There where null values in salesRepEmployeeNumber column in customer table, so to check weather this customers have ordered or made any payments in orders and payments table, used join statement between this tables. If there is no data of customer in other than customer column we can remove them to maintain 100% validity.

Query:

select \* from customers c left join payments p on c.customerNumber= p.customerNumber

left join orders o on c.customerNumber=o.customerNumber

where c.salesRepEmployeeNumber is null;.

1. To find Total number of orders placed by years which are shipped

(can also use different orders status).

Query:

Select Year(orderDate), sum(orderNumber) from orders

where status="Shipped"

group by 1 order by 1;

1. To find Total amount of payment done by customers on the basis of years.

Query:

Select Year(paymentDate), sum(amount) from payments group by 1 order by Year(paymentDate);

1. To find Total order quantity and amount for different product line.

Query:

Select a.productLine, sum(c.quantityOrdered),sum(c.quantityOrdered\*c.priceEach) from productlines a inner join products b on a.productLine=b.productLine

inner join orderdetails c on b.productCode=c.productCode

group by 1;

1. To see Total quantity and amount of different product line product which have shipping date null, this orders have status cancelled, disputed, On hold or may not have shipped yet.

Query:

Select a.productLine,d.shippedDate, sum(c.quantityOrdered),sum(c.quantityOrdered\*c.priceEach) from productlines a inner join products b on a.productLine=b.productLine

inner join orderdetails c on b.productCode=c.productCode

right join orders d on c.orderNumber=d.orderNumber

where d.shippedDate is null

group by a.productLine,d.shippedDate;

* **POWER QUERY EDITOR:**

1. Load data form GetData option in ribbon, select source as MYSQL, select server and database and then select transform database option to use Power Query editor.
2. After loading, explore data, use column quality and distribution option in view, this will provide more information about each columns like valid, null, error and distribution of values.
3. Removing address 2 and state columns from customers and offices tables dues to excess null values.
4. Changing datatype of credit limit column of customer table to fixed decimal for currency.
5. Changing datatype of price each column of orderdetails table to fixed decimal for currency.
6. Removing comments columns from orders tables dues to excess null values.
7. Changing datatype of amount column of payments table to fixed decimal for currency.
8. Removing html Description and image columns from productlines tables dues to excess null values.
9. Changing datatype of buyprice and MSRP column of payments table to fixed decimal for currency.
10. Removing all values from customers table for having salesRepEmployeeNumber null, because this values do not have relatable data in other tables.

* **Model View and Table view:**

1. In model view, we can see cardinality between the tables and can also manage them using manage relationship option.
2. This data is a snowflake schema, having orders table as fact table and other as dimensional and sub-dimensional tables.
3. Each table has one to many relationship with other.
4. Check each table is linked with the other with related column only
5. In table view we can view tables and there values in tabular form.
6. Here we can add custom columns, measure and quick measures using DAX formulas.
7. In orderdetails I have added following columns using NewColumn:

* Amount column which will give total order amount of each order with the product between priceEach and orderQuantity:

Amount = 'classicmodels orderdetails'[quantityOrdered]\*'classicmodels orderdetails'[priceEach]

1. In orderdetails I have added following NewMeasures:

* To find the total number of orders : Total Orders = SUM('classicmodels orderdetails'[orderNumber])
* To find total sales amount :Total sales = SUM('classicmodels orderdetails'[Amount])
* To find Average sales amount : Avergage Sales Amount = AVERAGE('classicmodels orderdetails'[Amount])
* To find Average Order quantity : Average Order Quantity = AVERAGE('classicmodels orderdetails'[quantityOrdered])

1. Created a new date dim table with reference to orderdate and payment date columns using CALENDARAUTO function, also created a new column for year and month.

Date Dim = CALENDARAUTO()

* **DASHBOARD:**

1. On the top left corner, there is a slicer for different years. This will segregate data for different years selected.
2. On the top right corner there are 3 slicers
   * 1. Top most slicer will segregate data for different ProductLines.
     2. 2nd one will segregate for orders from different country.
     3. 3rd will segregate on the basis of order status like shipped, on hold, etc.

All this slicers will effect all the charts and values on the Dash Board.

1. Below the Heading, there are cards showing information starting from left corner as follows:
2. Total Orders : shows total numbers of orders placed.
3. Total sales : shows sum of amount of orders.
4. Total Quantity in Stock : shows total quantity of products in Stock.
5. Avg.Order Amount : shows Average of amount of orders.
6. Total Payments : shows Total amount paid by customers.
7. In the centre portion there is the line chart showing graph for Total sales by Year, Quarter and Product Line as Legend. This shows distribution of sales each year for different Product Lines with different colours mentioned on the chart.
8. Besides the graph, there is funnel chart of Top 5 country with there order quantities. This shows which country has how many numbers of orders placed.
9. At Bottom there are two Donut charts with different colours for each product line as mentioned in line chart:
10. Total sales by different product lines : showing sales of each product lines.
11. Total orders by product lines : showing number of orders placed for each product line.
12. Also there are bar charts for:
13. Total Sales by year and product line : showing sales by each year and contribution of each product lines.
14. Payment Amount by year : showing payments received from customer.

* **Analysis:**

From the line chart, we can see that the sales are at peak during the 3rd and 4th quarters of 2003 and 2004, but also decreasing in 1st and 2nd quarters respectively. There is need to increase sales during this 1st and 2nd quarters.

USA stand on the top in the number of orders placed with large different in numbers as compare to other countries. Spain and UK being 2nd and 3rd on the list. We also need to expand and focus more on this countries to improve sales.

From the Product Lines, Classic car sales are the highest, with around 40% of total sales, following by Vintage cars and motorcycles.