**Axon Cars Retailer Power BI Documentation**

The Main Objective of this project is to connecting the database using MYSQL and establish a connection between MYSQL and Power BI.

Simply, establishing a connection between MYSQL and Power BI to import the dataset into Power BI.

And the project says to do a **Sales Analysis Dashboard** for the client to manage and analyze the data. The management needs to get the up-to-date sales reports for **decision making** process.

**Process of Capstone Project**

**Establishing Connection**

Downloaded the SQL file into MYSQL Workbench which is provided by the Organization for the purpose of data visualization. Observed all the data from the tables in this database named as “classicmodels”. Imported this database into Power BI Desktop using the server: 127.0.0.1:3306 and Database: classicmodels. Now selected all the tables available in the database which are as follows:

Classicmodels.customers

Classicmodels.employees

Classicmodels.offices

Classicmodels.orderdetails

Classicmodels.orders

Classicmodels.payments

Classicmodels.productlines

Classicmodels.products

**Data Transformation**

**Classicmodels.customers**

Removed unnecessary columns which are disturbing the data model. Existing columns are: Customer Number, Store Name, Last Name, First Name, Phone, AddressLine1, City, Postal Code, Country, SalesRep Employee Number, Credit Limit

Renamed with capitalizing all the column names.

Filtered Rows of Postal Code and SalesRep Employee Number columns by removing Null values.

**Classicmodels.employees**

Removed unnecessary columns which are disturbing the data model. Existing columns are: Employee Number, Last Name, First Name, Extension, Email, Office Code, Reports To, Job Title.

Renamed with capitalizing all the column names.

**Classicmodels.offices**

Removed unnecessary columns which are disturbing the data model. Existing columns are: Office Code, City, Phone, AddressLine1, State, Country, Postal Code, Territory

Renamed with capitalizing all the column names.

**Classicmodels.orderdetails**

Removed unnecessary columns which are disturbing the data model. Existing columns are: OrderLineNumber, PriceEach, Quantity Ordered, Product Code, Order Number

Renamed with capitalizing all the column names.

**Classicmodels.orders**

Removed unnecessary columns which are disturbing the data model. Existing columns are: Order Number, Order Date, Delivered Date, Shipped Date, Status, Comments, Customer Number.

Removed null values by filtering in Shipped Date column.

Renamed with capitalizing all the column names.

**Classicmodels.payments**

Removed unnecessary columns which are disturbing the data model. Existing columns are: Customer Number, Check Number, Payment Date, Amount

Renamed with capitalizing all the column names.

**Classicmodels.productlines**

Removed unnecessary columns. Existing columns are: Product Line, Text Description.

Renamed with capitalizing all the column names.

**Classicmodels.products**

Removed unnecessary columns. Existing columns are: Product Code, Product Name, Product Line, Product Scale, Product Vendor, Product Description, Quantity InStock, Buy Price, MSRP.

Renamed with capitalizing all the column names.

After making all the changes **Transformed the model by “Close and Apply” option**

**Power BI Desktop**

After observing the data model in the “Model View”, I found that this is neither a star schema nor a snowflake schema because, there is no Fact table in it and too many tables for one kind of data. So, now I need to merge some tables to make it one table for one kind of data and need to delete some tables if required and bring a Fact table to make it a star schema. For this, the following changes has been done:

1. Renamed all tables for better understanding
   1. Customers
   2. Employees
   3. Offices
   4. Orders
   5. Payments
   6. Products
2. Deleted “Product Line” table because the same columns are already existed in “Products” table
3. Merged “Orders” table and “Order Details” table
4. Created some measures
   1. Total Price by quantity
   2. Profit
   3. Total Discount
   4. Average Sales Amount
   5. MAX of Sales Amount
   6. Min of Sales Amount
5. Created Region Hierarchy
6. Created Date Hierarchy
7. Created a column for “Cancelled” status using DAX function
8. Created a calculated column “Discount” in Products table
9. Created a Quick Measure “Sales Amount YOY%” for Year-on-Year Percentage change
10. Created Target measure for KPI’s analysis as it has not been given by the organization
11. Created a calculated column using DAX Function for Distinct count of Products

**Merging queries**

When I tried to merge queries in power query editor, I found some data discrepancy in “Order Number” in both tables “Orders” and “Order details” table. This could happen for below reasons:

1. **Duplicate Order Numbers:** There might be duplicate order numbers within one or both of the tables. If the same order number appears multiple times in either table, it would contribute to the inflated row count after merging.
2. **Data Integrity Issues:** There could be discrepancies or inconsistencies in the data. For example, some order numbers in one table might not have corresponding entries in the other table due to data entry errors, missing data, or incomplete records.

As the dataset is not appropriate, this merge condition may not satisfy perfect analytical visualization. Also, there is no fact table itself. So, performed the correct approach as I know, irrespective of data discrepancy in the dataset.

**Visualizations Used in the Dashboards**

1. Dashboard Name: **Axon classic cars Dashboard**

**KPI:**

Used Value as Sales Amount

Used Trend Axis as Year

Used Target as created Target Measure

**Cards:**

Used measures are “Sum of Quantity In Stock”, “Sum of Profit” and “Count of Order Number”. It displays a single value in each card respectively.

**Guage Chart:**

Guage chart based on “Sum of Sales Revenue”. It shows the performance of a measure like sales revenue etc. Used a target line and data labels for proper identification of the data. It quickly displays Total Sales Revenue.

**Area Chart:**

It is used to display the sum of sales based on Quarter, month and Year. Year slicer is used outside this chart for page interaction.

**Donut Chart:**

It shows the Sales based on Product Category. Displays the sum of sales and percentage of each category outside of the Donut. Used legend as Product Category.

**Line and Stacked Column Chart:**

It Displays the Sum of Sales and Sum of Profit based on Year and Status of the Product. Here Sum of profit is shown as Stacked column chart and Sum of Sales Amount is shown as Line chart.

**Clustered Bar Chart:**

It Shows the Sum of Sales by Country. Zoom Slider is used in this visualization for better visual purpose.

**Clustered Column Chart:**

It displays Revenue by Product category and Product Scale. Used Legend as Product Scale and used data labels for all categories in X – Axis, Revenue in Y – Axis.

**Slicer:**

It changes the values of each chart based on Year used in this Slicer. It interacts whole report with all charts available here. When Slicer changes, remaining all visuals will change accordingly because it interacts with all the visualizations.

1. **Report Name: Data in Matrix and Area Chart Visuals**

**Area Chart:**

It Display the Sum of Sales by year and Product line category. Across Product Line Category, Vintage Cars had the most interesting recent trend and started trending down on 2003, falling by 63.19% (1,885,609.90) in 2 years.﻿

Matrix chart:

It shows the same data as Area chart in a tabular format with each column totals and each row totals at the end. It has Drill through option available to check.

Narrative:

Used Narrative for smart narration about the visual to understand better. It explains the charts or tables visualization clearly using AI.

1. **Report Name: Scatter Plot Using Python and MAP Visual**

**Python Visual:**

The Scatter Plot Displays Sales amount Country Wise using Python code which is interpreted in this page. This scatter plot will shows up when we have Python installed in our Local Machine where we're trying to see this chart. It shows Country wise Total Sales.

**Map Visual:**

It shows sum of sales Country wise and State wise.

**Line Chart:**

It shows the sum of total price by Quantity and by Year.

**Table:**

It shows the data of Sales by Time which means, Year and Month with Grand Total.

**Narrative:**

It shows the detailed explanation of the report page.

1. **Report Name: Key Performance Indicator**

**Ker Influencer:**

This Key Influencer is displaying visuals based on the dataset. But the dataset is not proper to show proper key Influencer here. So, only available columns are used to create it, but this is not appropriate. It shows the key performance of particular measures in our dataset.

1. **Report Name: Area Chart with Narrative**

**Area Chart:**

It shows the count of cancellation and count of status and count of comments by Year, Quarter and Month.

**Narrative:**

It shows the detailed explanation of the report page.