SALES INSIGHT

video - https://www.youtube.com/watch?v=hhZ62IlTxYs&list=PLeo1K3hjS3uva8pk1FI3iK9kCOKQdz1I9

Theory - https://medium.com/@thpereiraduarte/sales-insights-atliq-hardware-ee2ab6f85f2a

Source files - https://codebasics.io/resources/sales-insights-data-analysis-project

In this project I’ve simulated a business problem resolution to a hardware company in India.

## **AtliQ Hardware**

- Is a company which supplies computer hardware and peripherals to many clients across India;  
- The company has a head office in Dehli and regional offices throughout India.

## **Business Issue**

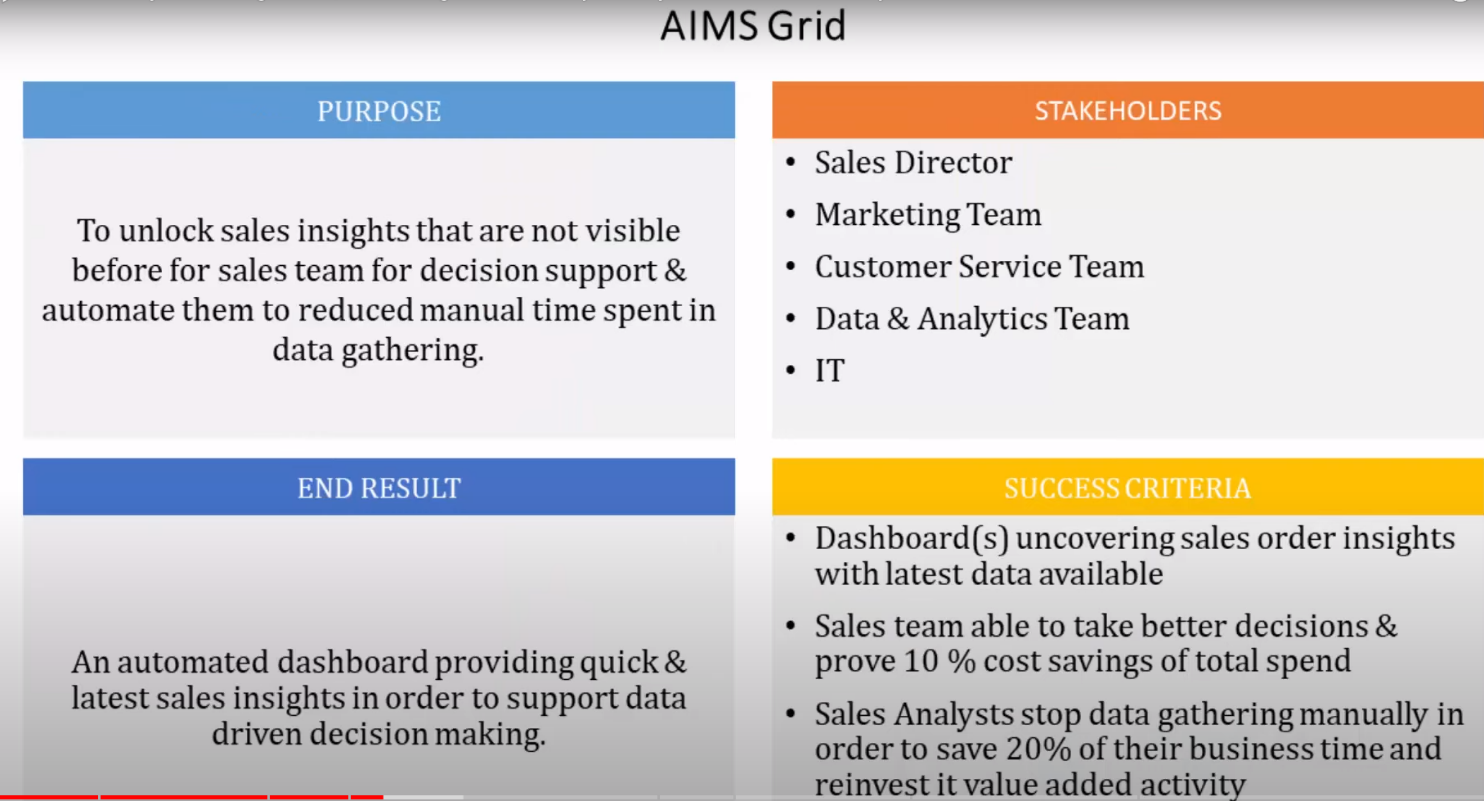
The sales director is facing a lot of challenges such as:  
The marketing is growing dynamically and then he’s struggling in terms of tracking the sales, needing more accurate insights about the company sales, and then take the necessary decisions.

*Imagine the scenario:*  
*The sales director need to know how the sales are going in all operations and the information provided by the regional sales managers are not being enough.*  
*Just hearing the numbers or receive tons of excel files is far from being effective in terms of having a reliable overview of the business.*  
*Instead, he want to be able to look at the data and understand what’s going on right away.*

show me top 5 customers, Top 5 product by revenue number, revenue breakdown by cities, which states have less sales.

Purpose = to unlock sales insights that are not visible before for sales team for decision support and automate them to reduce manual time spent in data gathering

Result = An automated dashboard providing quick and latest sales insights in order to support data driven decision making



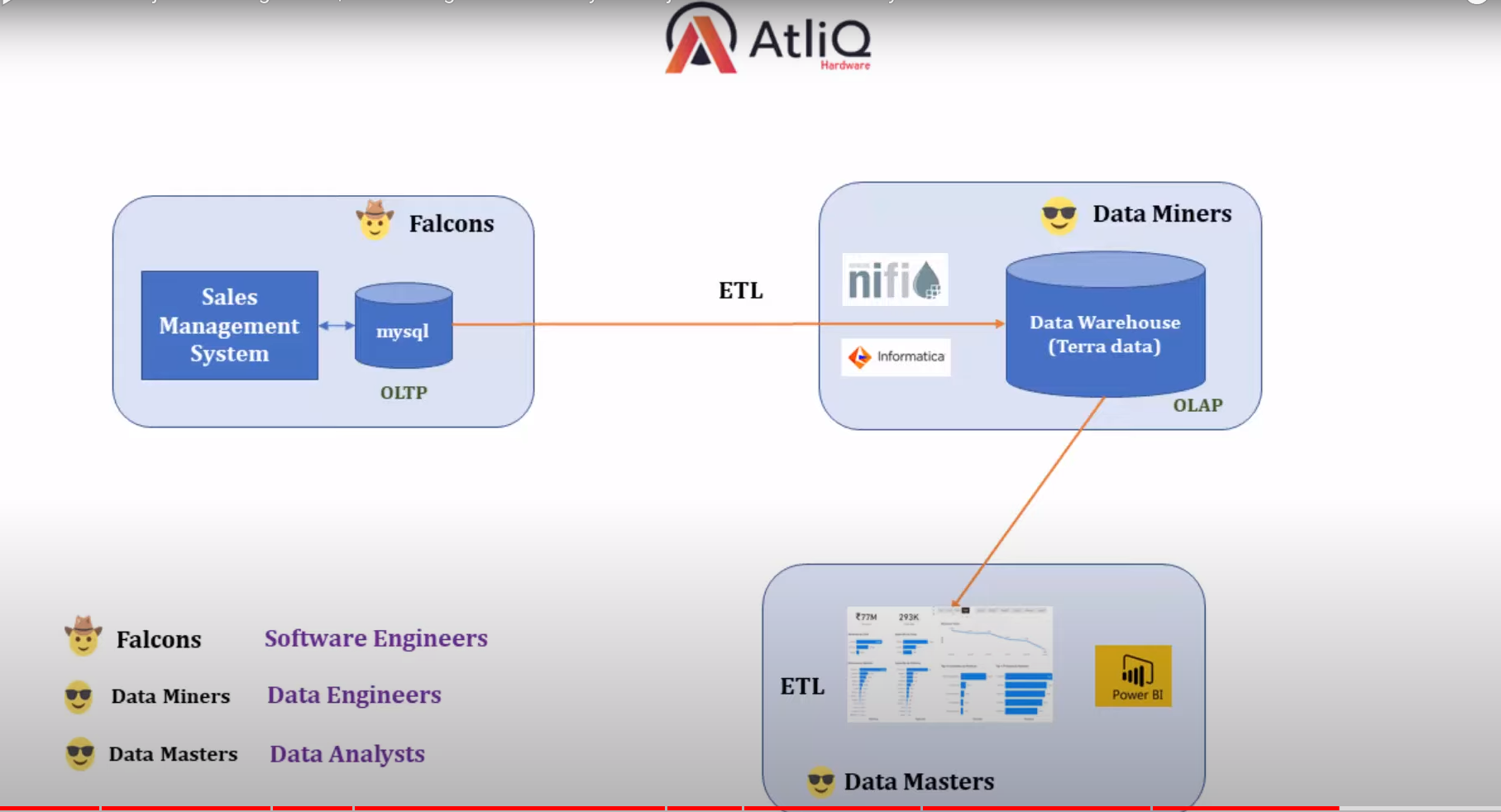
Star schema (https://www.databricks.com/glossary/star-schema),

DAX (Data Analysis Expressions)(programming language of power bi)

sales management system ----> MYsql (software engineers) [ETL]

Data warehouse (Terra data [relational data]) ----> Data engineers [ETL]

building dashboards ----> Data analyst



process

Data analysis using SQL

Data cleaning (like some currency are in INR and USD) in Power BI

Building dashboard

**Solution:**  
**Create a simple and informative dashboard about the company sales.**

I used **SQL** queries in **MySQL Workbench** to take a look into the data and **Power BI** for *ETL* and visualizations to create the insights.

## **Overview**

**Data Analysis Using SQL**

Show all customer records

SELECT \* FROM customers;

Show total number of customers

SELECT count(\*) FROM customers;

Show transactions for Chennai market (market code for chennai is Mark001

SELECT \* FROM transactions where market\_code='Mark001';

Show distrinct product codes that were sold in chennai

SELECT distinct product\_code FROM transactions where market\_code='Mark001';

Show transactions where currency is US dollars

SELECT \* from transactions where currency="USD"

Show transactions in 2020 join by date table

SELECT transactions.\*, date.\* FROM transactions INNER JOIN date ON transactions.order\_date=date.date where date.year=2020;

Show total revenue in year 2020,

SELECT SUM(transactions.sales\_amount) FROM transactions INNER JOIN date ON transactions.order\_date=date.date where date.year=2020 and transactions.currency="INR\r" or transactions.currency="USD\r";

Show total revenue in year 2020, January Month,

SELECT SUM(transactions.sales\_amount) FROM transactions INNER JOIN date ON transactions.order\_date=date.date where date.year=2020 and and date.month\_name="January" and (transactions.currency="INR\r" or transactions.currency="USD\r");

Show total revenue in year 2020 in Chennai

SELECT SUM(transactions.sales\_amount) FROM transactions INNER JOIN date ON transactions.order\_date=date.date where date.year=2020 and transactions.market\_code="Mark001";

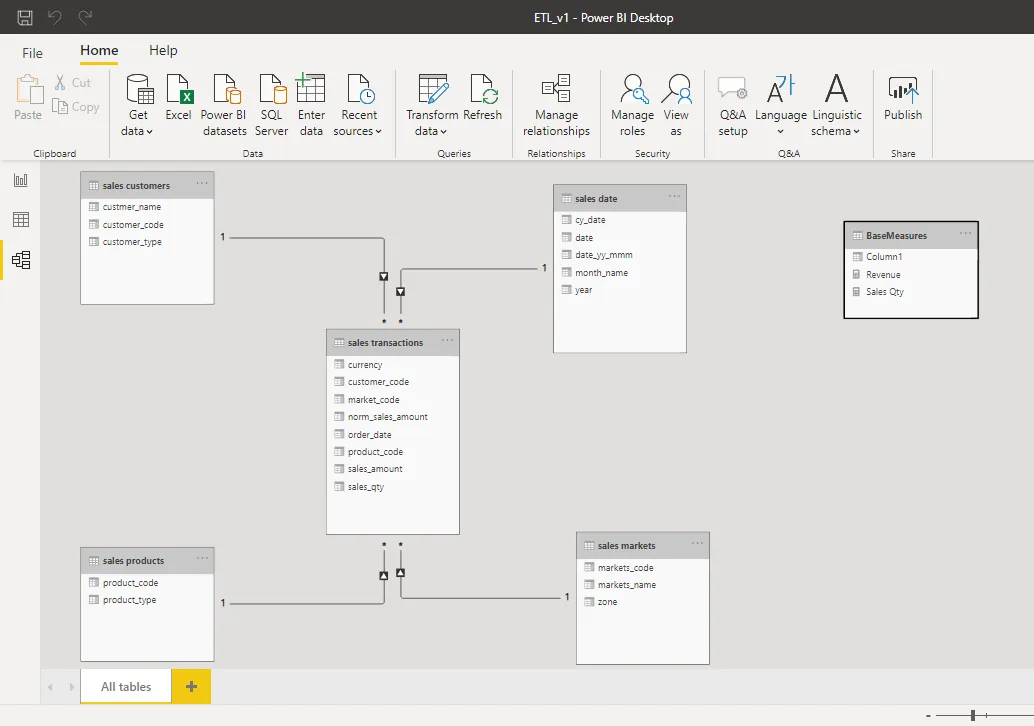
**After a quick data exploration in MySQL, here are some initial findings:**

* The database contains 5 tables: **customers, date, markets, products, and transactions;**
* There are **17 markets, 279 products, and 38 customers**;
* The observation period is from **january 2018 to june 2020;**
* The total revenue in 2020 was **$ 142,23 Mi**, 42% less than 2019, which was **$ 336,45 Mi**;
* Most of the transactions data are in **INR currency**, but we have 4 records in **U$ currency**.  
  And we got Paris and New York on the **“sales markets”** table. We’re going to deal with it in the ETL process.

## **ETL (Extract, Transform, Load)**

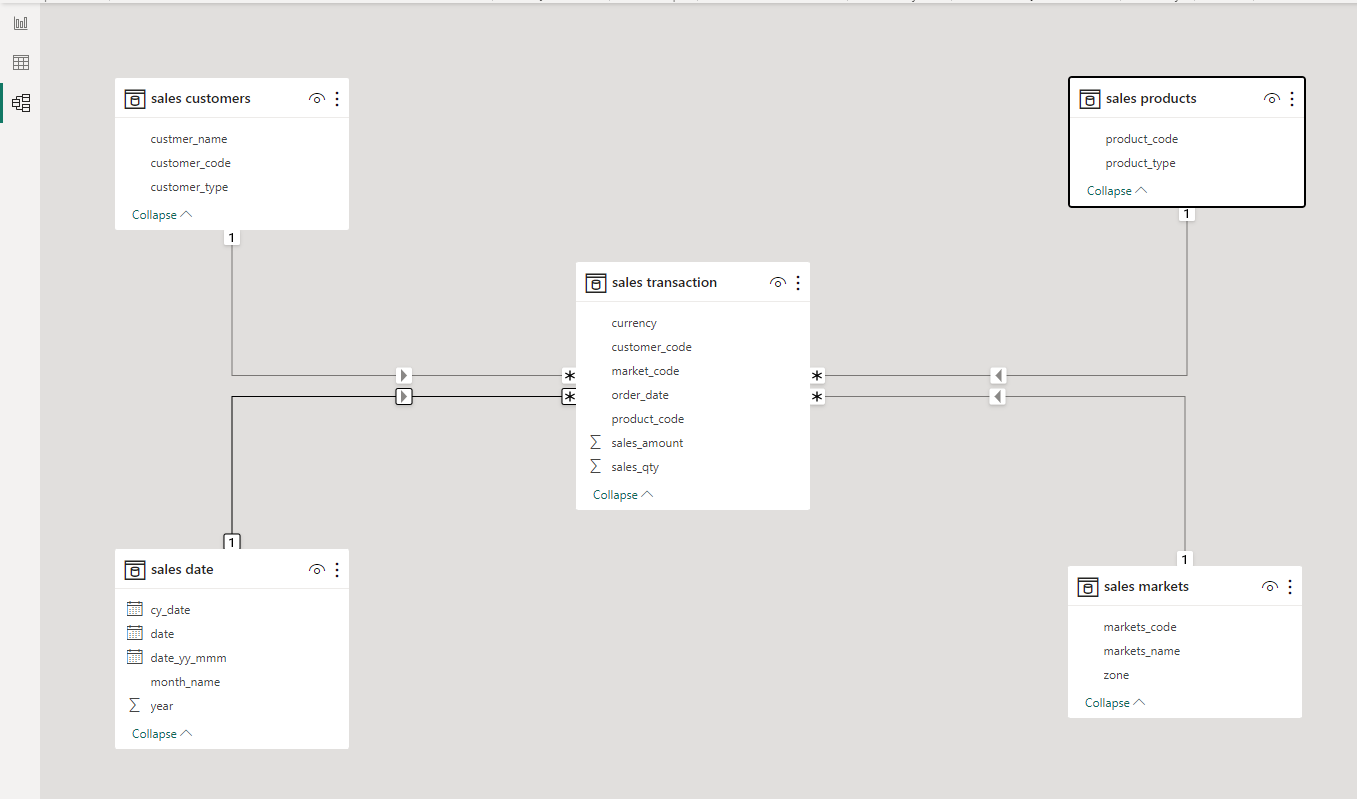
Once I know the basic features of the data I have to work with, I imported the MySQL database into Power BI to do the necessary transformations and end up with a simple, reliable, and useful dashboard.

*Data Modeling Step*  
*We got five tables and we need to ensure that the tables are correctly connected.*



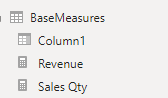
Star Schema

A star schema is a multi-dimensional data model used to organize data in a database so that it is easy to understand and analyze.



**[sales transactions]** — main table  
TABLES CONNECTED  
- sales customers > connected by **“customer\_code”** column;  
- sales date> connected by **“date”** column;  
- sales products > connected by **“product\_code”** column;  
- sales markets > connected by **“market\_code”** column.

*Measures Created*  
*“Revenue”* and **“Sales Qty”** measures to get the sum of each column instantly in the dashboard.

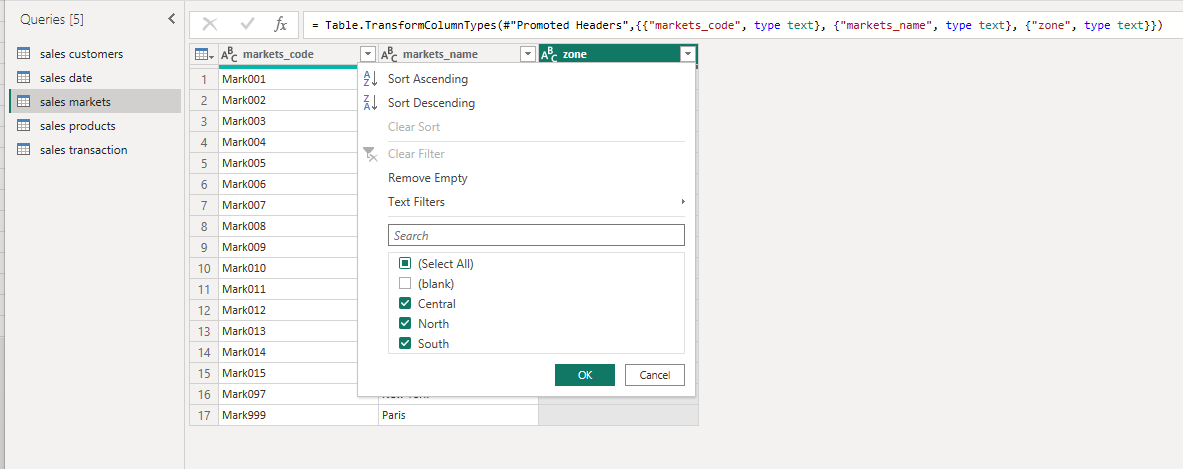


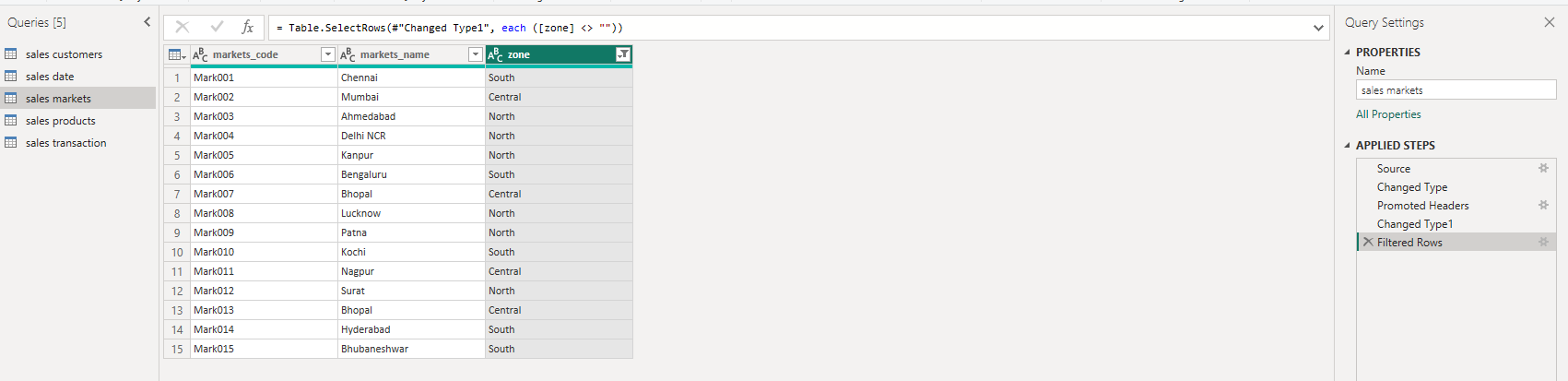
*Filtering*  
*Once the company is operating only in India, Ifiltered out “Paris” and “New York” in the* **sales markets** table by unchecking the “(blank)” field in the “zone” column.

A screenshot of a computer

Description automatically generated

We will go on transform data then power query editor will get open





*Cleaning and adding new column*  
*-* The **“sales\_amount”** column has **-1** and **0** values. I removed them in this case so we can see only the actual sales numbers on the dashboard (sometimes “0” means promotional sales and giveaways, but to know that, we should have further information, which is not the case of this project).

A screenshot of a computer

Description automatically generated

Filtered table power query

= Table.SelectRows(#"Changed Type", each ([sales\_amount] <> -1 and [sales\_amount] <> 0))

- I find out that the **“currency”** column (sales transactions table) have **4 USD currency values**, 2 of them with hidden characters, so i had to include an argument into the conditional formula, in order to create a new column called “norm\_sales\_amount”, where it converts the **USD currency value** into **INR currency value**.

Formula to create norm\_amount column

= Table.AddColumn(#"Filtered Rows", "norm\_amount", each if [currency] = "USD" or [currency] ="USD#(cr)" then [sales\_amount]\*75 else [sales\_amount], type any)

A screenshot of a computer

Description automatically generated

Change query null to 0

= Table.AddColumn(#"Filtered Rows", "Custom", each if [currency] = "USD" then 1 else null)

To

= Table.AddColumn(#"Filtered Rows", "norm\_sales\_amount", each if [currency] = "USD" then 1 else 0)

To

= Table.AddColumn(#"Filtered Rows", "norm\_sales\_amount", each if [currency] = "USD" then [sales\_amount]\*75 else [sales\_amount])

(in above query Converting USD to INR)

A screenshot of a computer

Description automatically generated

Since there are two USD some USD are “USD#(cr)”

= Table.AddColumn(#"Filtered Rows", "norm\_sales\_amount", each if [currency] = "USD" or [currency]="USD#(cr)" then [sales\_amount]\*75 else [sales\_amount])

In sql we found that there are two types of INR

SELECT \* from transactions where currency="INR/r"

Count -150000

SELECT \* from transactions where currency="INR"

Count - 279

SELECT \* from transactions where currency="USD" AND currency="USD/r"

Since we observed from above sql query the same transaction is repeating with INR and INR/r also with USD and USD/r so we will filter out USD and INR from following power bi query

= Table.SelectRows(#"Remove -1/0 in sales amount", each ([currency] = "INR#(cr)" or [currency] = "USD#(cr)"))

Then

Home->Enter data->Create table->BaseMeasures->load

A screenshot of a computer

Description automatically generated

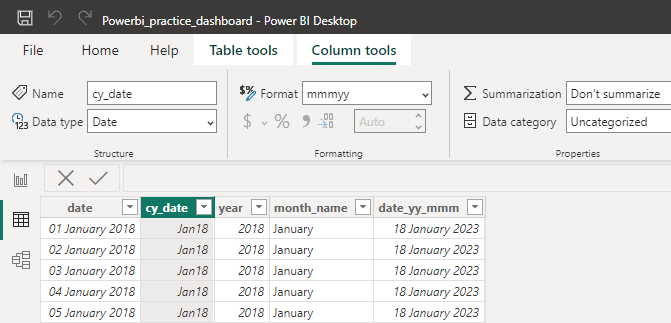
Then we created two measures

Revenue

Revenue = SUM('sales transaction'[sales\_amount])

Sales Qty

Sales QTY = SUM('sales transaction'[sales\_qty])



After all the analysis and transformation, my data is reliable enough to build the dashboard:

A screenshot of a graph

Description automatically generated

The dashboard shows all the main information about the company sales, such as **Revenue, Sales Quantity, Revenue and Sales by Market, Revenue Trend, Top 5 Customers and Top 5 Products.**  
**It can be filtered by YEAR and MONTH inside the observation period, so the sales director can have a deeper and quick view of the sales to support his decision making process.**