Data Analyst Project- Internet Sales

By Sherry Wang

Date: 05/09/2023

Contents

[Overview 3](#_Toc144825638)

[Business Request & User Stories 4](#_Toc144825639)

[Data cleansing and transformation (SQL) 4](#_Toc144825640)

[Extracting data: 4](#_Toc144825641)

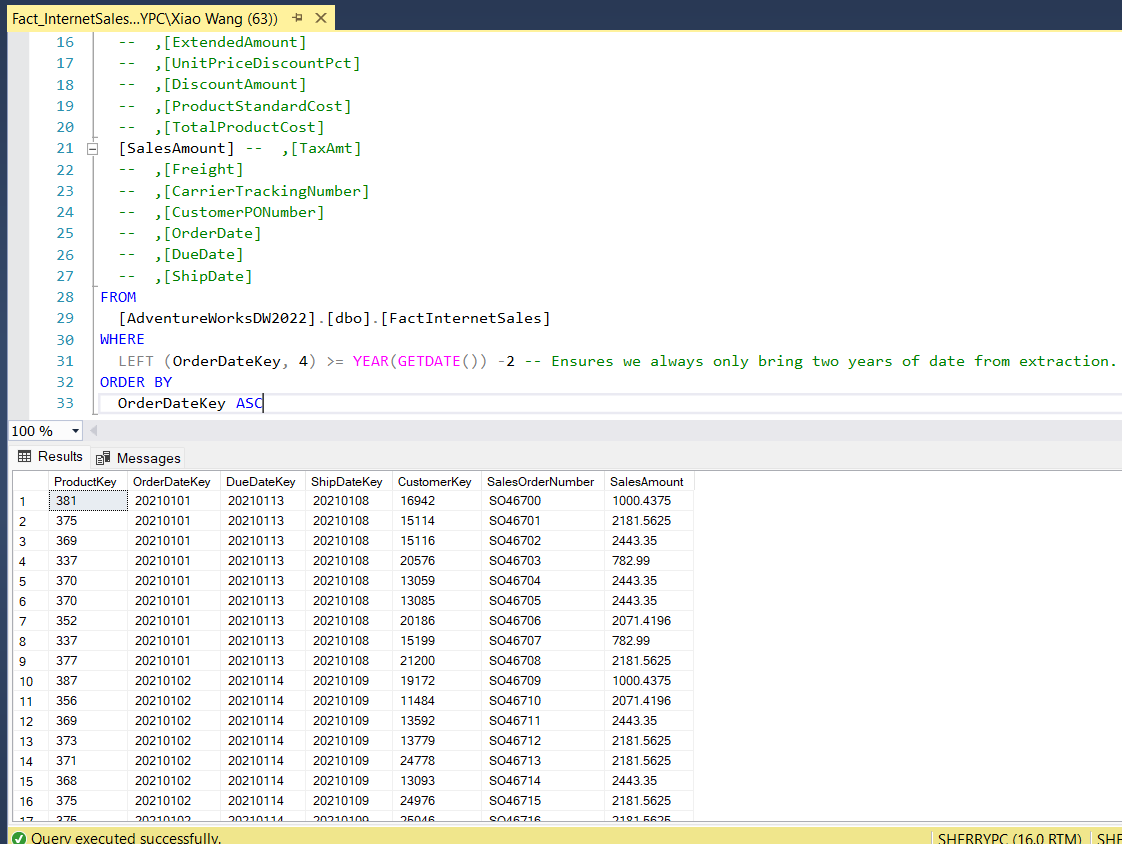
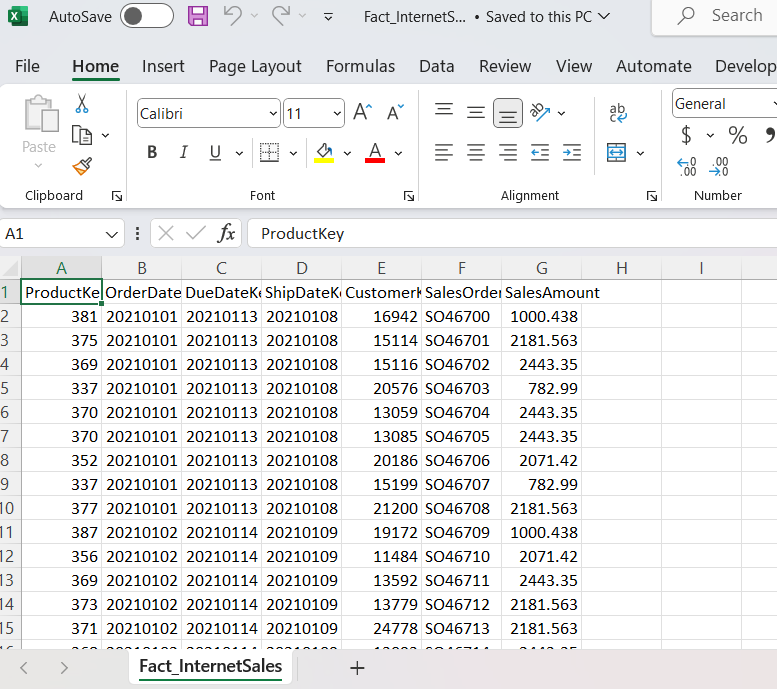
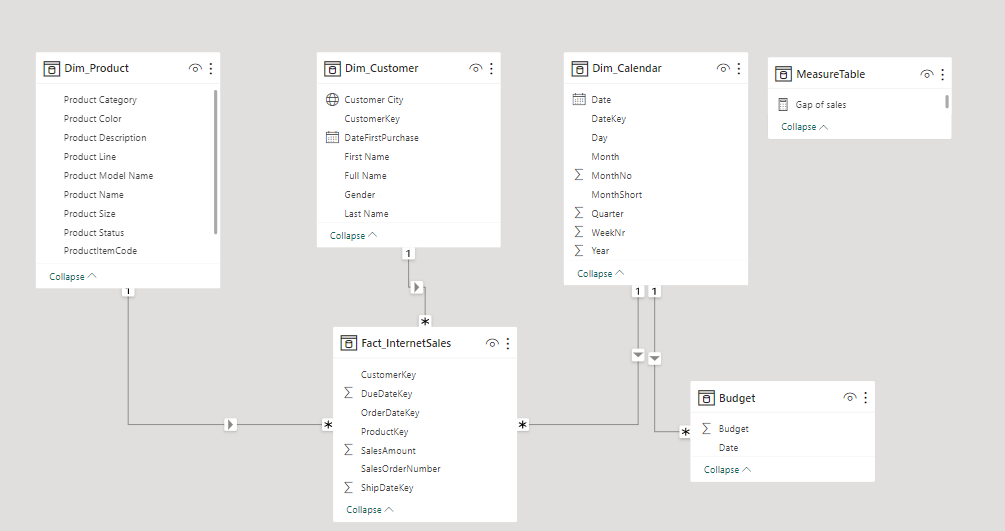
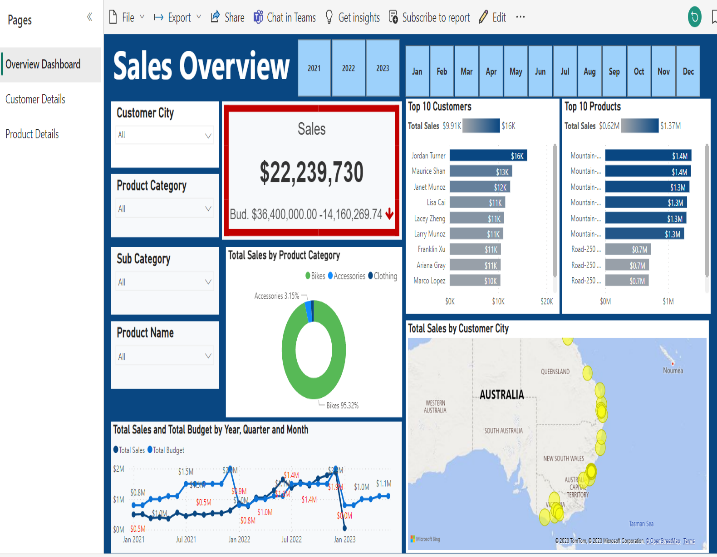
[Deal with missing value and null value 5](#_Toc144825642)

[Data Modelling 5](#_Toc144825643)

[Visualization 6](#_Toc144825644)

[Reference: 6](#_Toc144825645)

# Overview



# Business Request & User Stories

The business request generated from sales team. Based on the requests that was made from the business we following user stories were defined to fulfill delivery and ensure that acceptance criteria were maintained throughout the project.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No# | As a(role) | I want(request/demand) | So that I (user value) | Acceptance criteria |
| 1 | sales manager | To get a dashboard overview of internet sales | Can follow better which customers and products sells the best | A Power BI dashboard which updates data once a day |
| 2 | Sales representative | A detailed overview of internet sales per customer | Can follow up my customers that buys the most and who we can sell more | A Power BI dashboard which allows me to filter data for each customer |
| 3 | Sales representative | A detailed overview of internet sales per product | Can follow up my products that sells the most | A Power BI dashboard which allows me to filter data for each product |
| 4 | Sales manager | A dashboard overview of internet sales | Can follow sales over time against budget | A Power BI dashboard with graphs and KPIs comparing against budget |

# Data cleansing and transformation (SQL)

Data source was scratching from Microsoft website. (AdventureWorksDW 2022)

To build the necessary data model for doing analysis and fulfilling the business needs defined in the user stories the following tables were extracted using SQL.

One data source (sales budget) was provided in Excel format and was connected in the data model in a later step of the process.

## Extracting data:

There are one fact table and three dimension tables extracting from the database as following:

Dim\_Calendar table: transform data type (month, weekly, ..)

Dim\_Customer table: concat name string, join the table from the dim geography table for analysing different customer location.

Dim\_Product table: connect with dim main category table and subcategory table to get product details, Product status column we apply “outdated” status to replace Null value.

Fact\_InternetSales table: extract data 2 year before current year, the column field from [ProductKey], [OrderDateKey], [DueDateKey], [ShipDateKey], [CustomerKey], [SalesOrderNumber], [SalesAmount]

## Deal with missing value and null value

* Missing value (challenge of blank value in Dim\_Calendar table) solved measure.
* Replacing 209 null value (firstly, product category and subcategory replace to components regarding to product name. Secondly, subcategory null value with “other”. Finally, replacing null value in product colour, size, model name and description column field as N/A)
* Transformed customer city data into city categories for analysis and categorization.

# Data Modelling

Below is a screenshot of the data model after and prepared tables were read into Power BI.

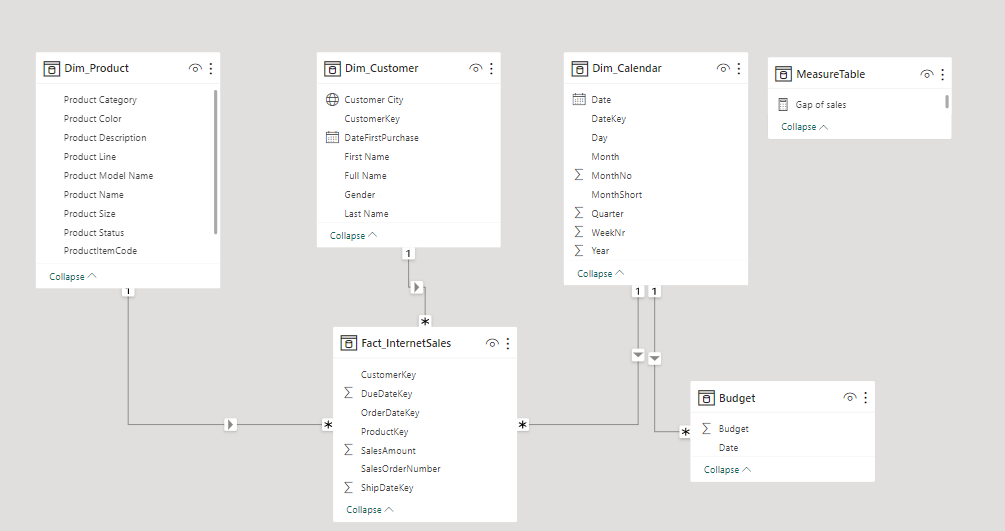
This data model also shows Fact\_Budget has been connected to Fact\_InternetSales and other necessary Dimension tables.

In addition, three measures have been created which used to analyse relationship between actual sales and sales budget as following:

Total Sales = sum(Fact\_InternetSales[SalesAmount])

Total Budget = sum(Budget[Budget])

Gap of sales = [Total Sales]-[Total Budget]



# Visualization

The final dashboard with one page sales overview, two pages focusing on combining tables for necessary details and visualization to show sales over time, per customer and per product.

Click the link to open the dashboard, thank you.

<<https://app.powerbi.com/view?r=eyJrIjoiMzEzZjk4MzEtYWJhMC00NTlhLWJjMjEtNmYyYTYzZTRmZDhlIiwidCI6ImFiYTgxYzhkLTNlYWUtNDg1OS1hYTBhLTlmNGE5ZTAxMmU0NiJ9>>

# Reference:

AdventureWorksDW 2022

<<https://learn.microsoft.com/en-us/sql/samples/adventureworks-install-configure?view=sql-server-ver16&tabs=ssms>>