DATA ANALYTICS: POWER BI

SUBMITTED TO: PROF. KRISHAN KUMAR

Abhishek Goyal
UE208007
IT
5th Semester
Abhishekgoyal274@gmail.com

Table Of ContentS

DATA ANALYSIS

In today's world data is everything. It is everywhere and in big quantities, but how to get usefull information from this data? This is where data analysis comes. It **extracts, transforms and load** data in most usefull and easy to understand way.

The rich variety of data that enterprises generate contains valuable insights, and **data analytics** is the way to unlock them. Data analytics can help an organization with everything from personalizing a marketing pitch for an individual customer to identifying and mitigating risks to its business.

Data Analysis:

Data analysis is a process of **inspecting, cleansing, transforming, and modeling** data with the goal of discovering useful information, informing conclusions, and supporting decision-making.

Example:

Assume we have large quantity of data stored in an excel file (or any other format like XML, JSON, sql dump, database etc as data must be stored in some format) that stores sales, customer, income etc of my company of past 10 years. As owner I want to analyse how my buisness is doing using that data, but I can't as I can not conclude anything from that. Data analysis if process which will extract (use that excel file into such that I can perform operations on that), it will clean all wrong, useless data and present it in a form that I can understand.

Importance of data analysis:

Here is a list of reasons why data analysis is such a crucial part of doing business today.

- Better Customer Targeting: You don't want to waste your business's precious time, resources, and money putting together advertising campaigns targeted at demographic groups that have little to no interest in the goods and services you offer. Data analysis helps you see where you should be focusing your advertising efforts.
- You Will Know Your Target Customers Better: Data analysis tracks how well your products and campaigns are performing within your target demographic. Through data analysis, your business can get a better idea of your target audience's spending habits, disposable income, and most likely areas of interest. This data helps businesses set prices, determine the length of ad campaigns, and even help project the quantity of goods needed.
- Reduce Operational Costs: Data analysis shows you which areas in your business need more resources and money, and which areas are not producing and thus should be scaled back or eliminated outright.

- Better Problem-Solving Methods: Informed decisions are more likely to be successful decisions. Data provides businesses with information. You can see where this progression is leading. Data analysis helps businesses make the right choices and avoid costly pitfalls.
- You Get More Accurate Data: If you want to make informed decisions, you need data, but there's more to it. The data in question must be accurate. Data analysis helps businesses acquire relevant, accurate information, suitable for developing future marketing strategies and business plans.

Problem:

Today there are many methods and tools used for data analysis, data cleaning, data visulization, such as Excel, Python, R, Looker, Rapid Miner, Chartio, Metabase, Redash. We have to use many tools for data cleaning, analysis and visulization, but power BI by microsoft is all in one tool.

- Hard to use
- To many Commands
- Bad interface
- Hard to learn

Solution:

Power BI is an interactive data cleaning, analysis and visualization software product developed by Microsoft with a primary focus on business intelligence. It is part of the Microsoft Power Platform. Power BI is a collection of software services, apps, and connectors that work together to turn unrelated sources of data into coherent, visually immersive, and interactive insights. Data may be input by reading directly from a database, webpage, or structured files such as spreadsheets, CSV, XML, and JSON.

- Power BI helps you to access your data instantly with less manual work. It can handle a huge amount of data making it easy to decipher using advanced visualizations. It allows you to get data from different data sources by automatically connecting with them, saving you time and effort.
- Power BI helps you quickly identify data quality issues and provides numerous ways to address them. Power Query provides you with exciting features to clean and prepare data for analysis.

- Power BI overcomes these issues by leveraging Azure Active Directory for authentication and Power BI login credentials to access the resources.
- On the other hand, Microsoft Power BI is a user-friendly and simple tool and can be used by anyone, even by non-experienced BI people.

Technology stack:

Sql: This is used to match the results of our querry and mysql server is used to extract data to Power BI for analysis.

Basic Sql queries:

- **SELECT**

SELECT is probably the most commonly-used SQL statement. You'll use it pretty much every time you query data with SQL. It allows you to define what data you want your query to return.

For example, in the code below, we're selecting a column called `name` from a table called `customers`.

```
SELECT name
FROM customers;
```

- **AGGREGATE FUNCTIONS :**

COUNT/SUM/AVG/MIN/MAX

Example:

SUM

SUM returns the total sum of a numeric column.

```
SELECT SUM(age)
FROM customers;
```

JOINS (INNER, LEFT, RIGHT, FULL)

A JOIN clause is used to combine rows from two or more tables. The four types of JOIN are INNER, LEFT, RIGHT and FULL.

INNER JOIN

INNER JOIN selects records that have matching values in both tables.

```
SELECT name
```

```
FROM customers
INNER JOIN orders
ON customers.customer_id = orders.customer_id;
```

- **UNION**

UNION combines multiple result-sets using two or more SELECT statements and eliminates duplicate rows.

SELECT name FROM customersUNIONSELECT name FROM orders;

- **UNION ALL**

UNION ALL combines multiple result-sets using two or more SELECT statements and keeps duplicate rows.

```
SELECT name FROM customers
UNION
SELECT name FROM orders;
```

- **WildCards:**

LIKE

LIKE searches for a specified pattern in a column. In the example code below, any row with a name that included the characters Bob would be returned.

```
SELECT name
FROM customers
WHERE name LIKE '%Bob%';
```

Other operators for LIKE:

- %x — will select all values that begin with x
- %x% — will select all values that include x
- x% — will select all values that end with x
- x%y — will select all values that begin with x and end with y
- _x% — will select all values have x as the second character
- x_% — will select all values that begin with x and are at least two characters long. You can add additional _ characters to extend the length requirement, i.e. x__%

And so on.

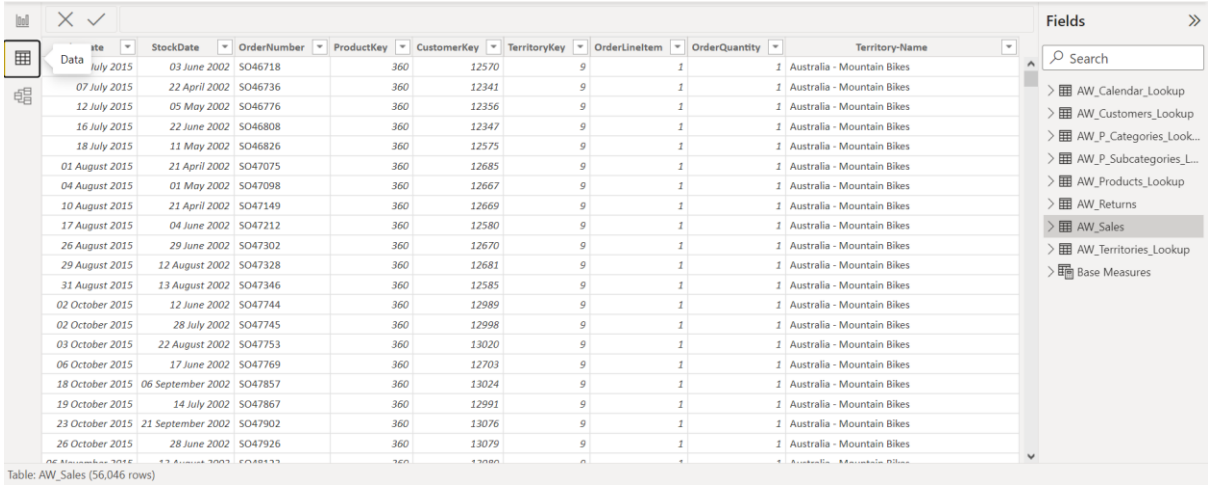
Power BI: this is the main app we use to analyse our data.

It can be devided into three sections:

- **Report**
- **Data**
- **Model**

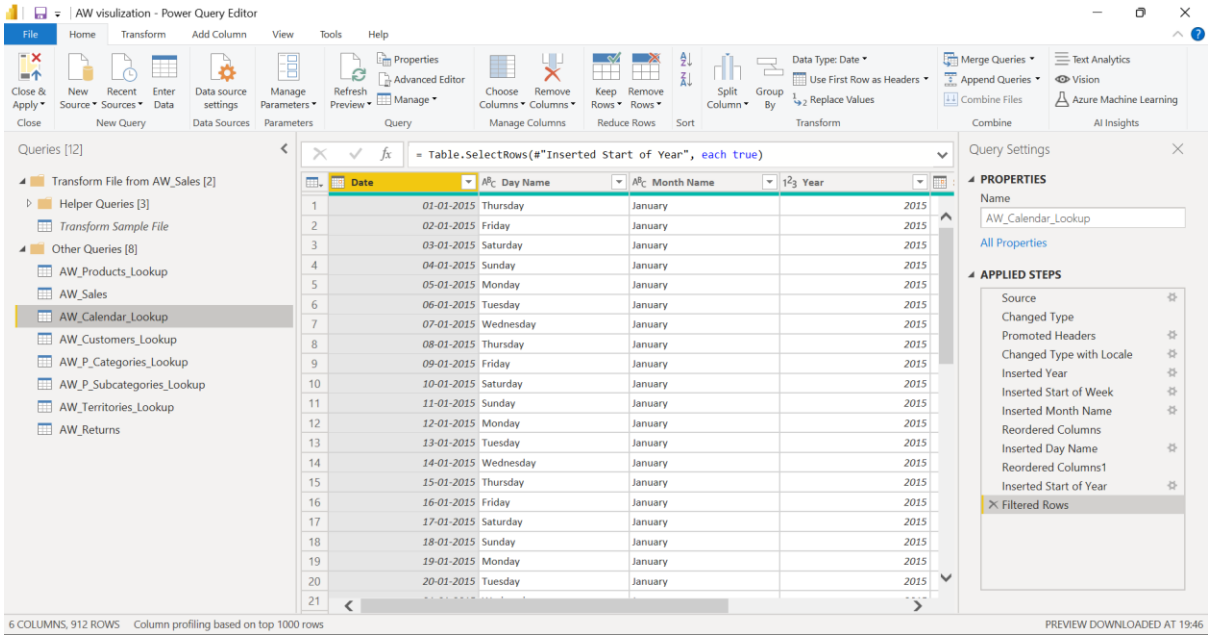
- ## Data View:

In this view we change all data acording to our need, clear errored data, add new columns, change data types and anything we want with to change with tables.



StockDate	OrderNumber	ProductKey	CustomerKey	TerritoryKey	OrderLineItem	OrderQuantity	Territory-Name
03 June 2002	SO46718	360	12570	9	1	1	Australia - Mountain Bikes
22 April 2002	SO46736	360	12341	9	1	1	Australia - Mountain Bikes
05 May 2002	SO46776	360	12356	9	1	1	Australia - Mountain Bikes
22 June 2002	SO46808	360	12347	9	1	1	Australia - Mountain Bikes
11 May 2002	SO46826	360	12575	9	1	1	Australia - Mountain Bikes
21 April 2002	SO47075	360	12685	9	1	1	Australia - Mountain Bikes
01 May 2002	SO47098	360	12667	9	1	1	Australia - Mountain Bikes
21 April 2002	SO47149	360	12669	9	1	1	Australia - Mountain Bikes
04 June 2002	SO47212	360	12580	9	1	1	Australia - Mountain Bikes
29 June 2002	SO47302	360	12670	9	1	1	Australia - Mountain Bikes
12 August 2002	SO47328	360	12681	9	1	1	Australia - Mountain Bikes
13 August 2002	SO47346	360	12585	9	1	1	Australia - Mountain Bikes
12 June 2002	SO47744	360	12989	9	1	1	Australia - Mountain Bikes
28 July 2002	SO47745	360	12998	9	1	1	Australia - Mountain Bikes
22 August 2002	SO47753	360	13020	9	1	1	Australia - Mountain Bikes
17 June 2002	SO47769	360	12703	9	1	1	Australia - Mountain Bikes
06 September 2002	SO47857	360	13024	9	1	1	Australia - Mountain Bikes
14 July 2002	SO47867	360	12991	9	1	1	Australia - Mountain Bikes
21 September 2002	SO47902	360	13076	9	1	1	Australia - Mountain Bikes
28 June 2002	SO47926	360	13079	9	1	1	Australia - Mountain Bikes

Power Query: We use this to perform operations on data more efficiently.



- **Model view:**

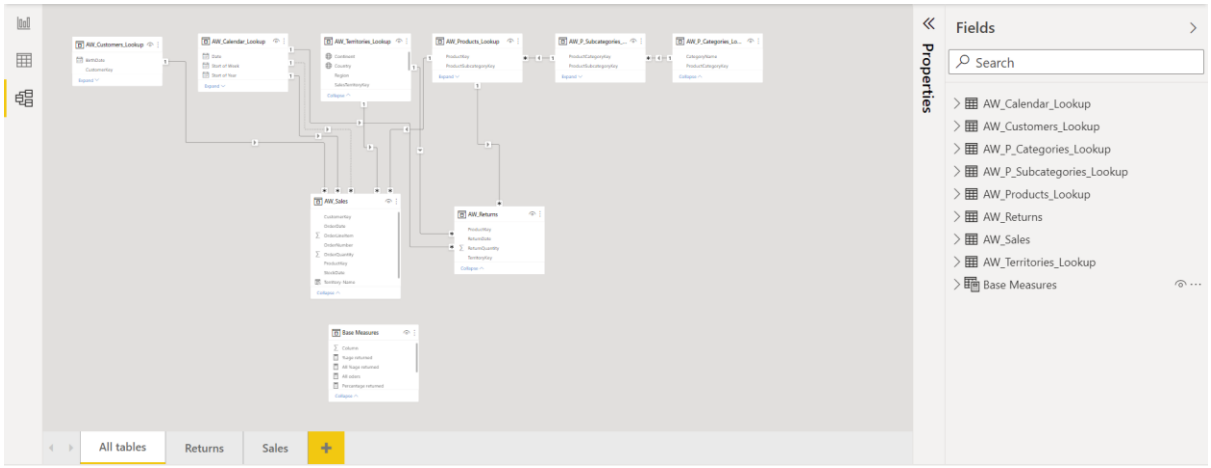
Once we have data structured according to our needs we can create relationships between data.

- **Ont to one**
- **Onr to many and so on.**

We usually devide tables into two category lookup table and data table.

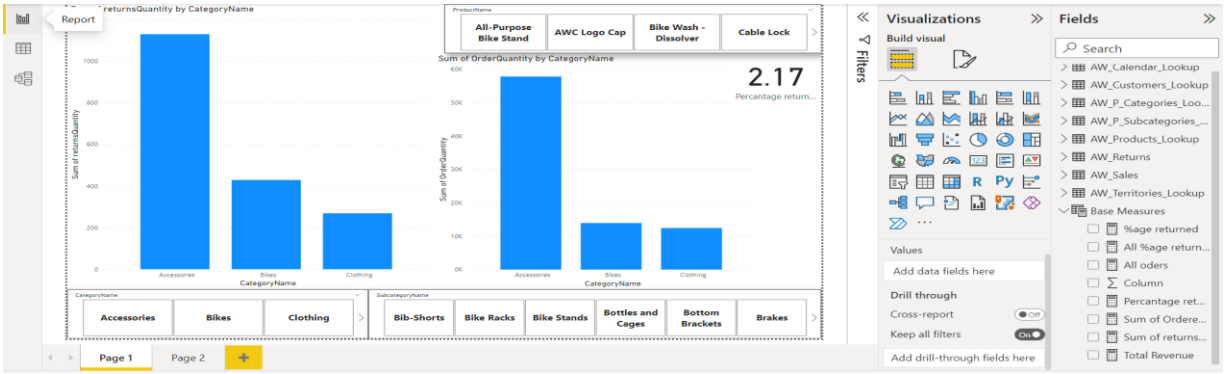
Lookup Tables: it contains additional info about a particular thing. Like territory table contains data about its location, region, country etc.

Data table: It contains our main data usually many foreign keys and data to visulize. We use look up tables with this to visulize data.



- **Report View:**

Once all data and relationships have been created we can now visulize data using report view. Iy has many tools and smart features t create a great desktop and mobile rport view.

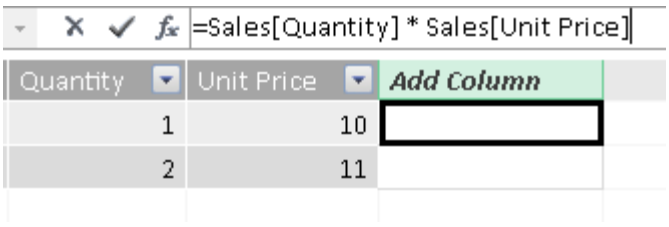


- **DAX**: Data Analysis Expression

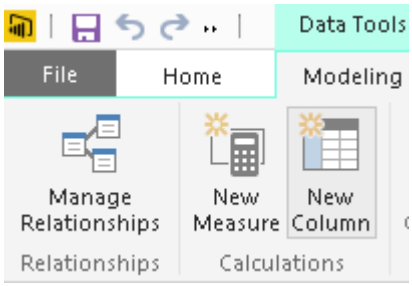
It uses its own language for creating calculated columns and measures used for cleaning and visulisaton.

Calculated columns

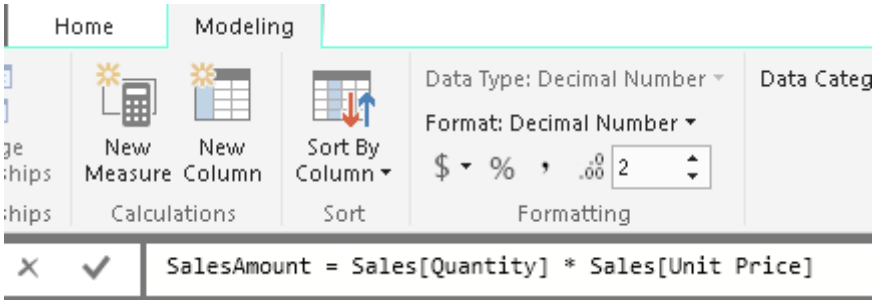
When you create a data model in Power Pivot for Excel, Analysis Services Tabular, or Power BI Desktop, you can extend a table by creating new columns. The content of the columns is defined by a DAX expression evaluated row by row. The user interface is different depending on the tools you use. Excel and Analysis Services require you to write the expression in the formula textbox when you select the last column on the right – "Add Column". You can rename the new column before or after defining the expression by right-clicking the new column and selecting the Rename Column menu item. As you see in the following picture, the DAX formula you write does not contain the column name and starts with the assignment symbol (=).



In Power BI Desktop, you have a different user interface. You have to click the New Column button in order to create a new column.



The new column name is part of the formula you write in the formula textbox.



The user interface allows you to simply define a new column, but we talk about calculated column to make a distinction between native columns (those read from the data source or evaluated by a query written in Power Query or Power BI) and calculated columns (those created extending a table in the data model).

A calculated column is just like any other column in a table and you can use it in any part of a report. You can also use a calculated column to define a relationship if needed. The DAX expression defined for a calculated column operates in the context of the current row across that table. Any reference to a column returns the value of that column for the current row. You cannot directly access the values of other rows.

Measures

There is another way of defining calculations in a DAX model, useful whenever you do not want to compute values for each row but, rather, you want to aggregate values from many rows in a table. These calculations are measures. This is the same name used in the user interface, with the exception of Excel 2013, which uses the term "calculated field" instead of "measures". Excel 2016 reverted back to "measures", which is the term used in DAX and originally used in Power Pivot for Excel 2010, too.

In the previous example you learned how to define the GrossMargin column in the Sales table to compute the gross margin amount. However, what happens if you want to show the gross margin as a percentage of the sales amount? You could create a calculated column with the following formula:

```
Sales[GrossMarginPct] = DIVIDE ( Sales[GrossMargin], Sales[SalesAmount] )
```

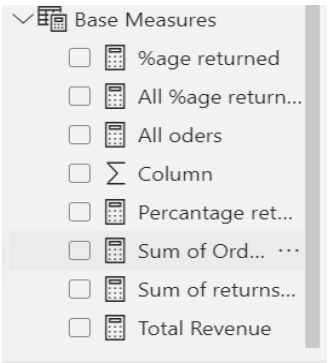
This formula computes the right value at the row level, as you can see in the following picture.

Unit Cost	Quantity	Unit Price	SalesAmount	TotalProductCost	GrossMargin	GrossMarginPct
3.00	1	8.000	8.00	3.00	5.00	63 %
3.00	2	7.500	15.00	6.00	9.00	60 %
3.00	3	7.200	21.60	9.00	12.60	58 %
3.00	4	6.800	27.20	12.00	15.20	56 %
2.50	5	7.000	35.00	12.50	22.50	64 %
2.50	6	5.300	31.80	15.00	16.80	53 %
2.50	7	5.400	37.80	17.50	20.30	54 %
2.50	8	5.100	40.80	20.00	20.80	51 %
2.50	9	5.100	45.90	22.50	23.40	51 %
2.90	8	5.000	40.00	23.20	16.80	42 %

Nevertheless, when computing the aggregate value of a percentage, you cannot rely on calculated columns. instead you need to compute the aggregate value as the sum of gross margin divided by the sum of sales amount. Therefore, in this case, you need to compute the ratio on the aggregates – you cannot use an aggregation of calculated columns. In other words, you compute the ratio of the sums, not the sum of the ratio.

```
1 All %age returned = ([Sum of returnsQuantity]/ [All oders]) * 100
```

You cannot use a calculated column for this operation. If you need to operate on aggregate values instead of on a row-by-row basis, you must create measures.




```
1 All oders = CALCULATE(SUM(AW_Sales[OrderQuantity]), ALL(AW_Products_Lookup))
```

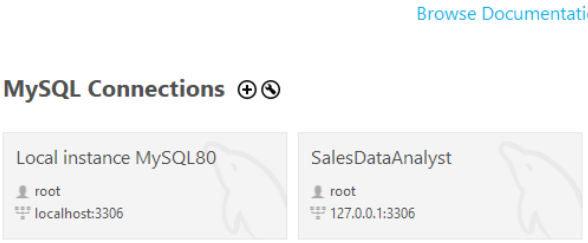
Example:

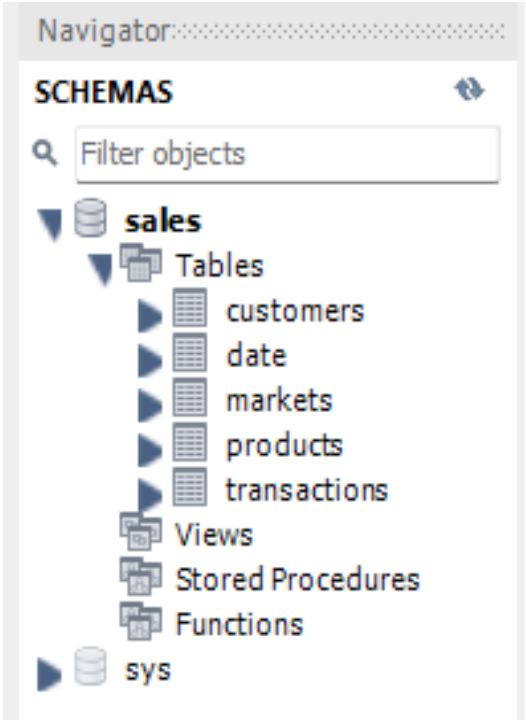
SUM, SUMX, ALL, FILTER

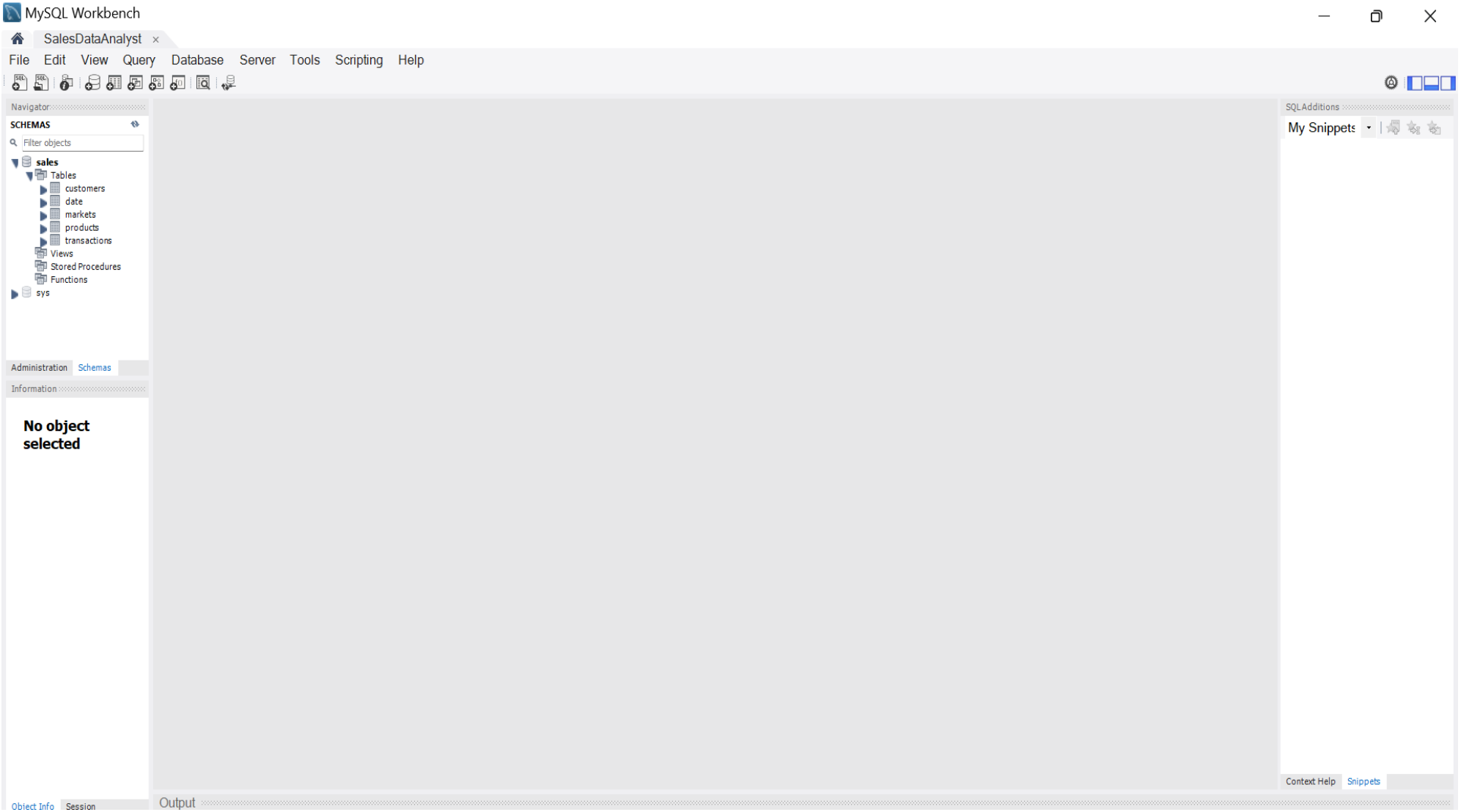
Project 1 (Buisness Sales Sample Data)

Data Used:

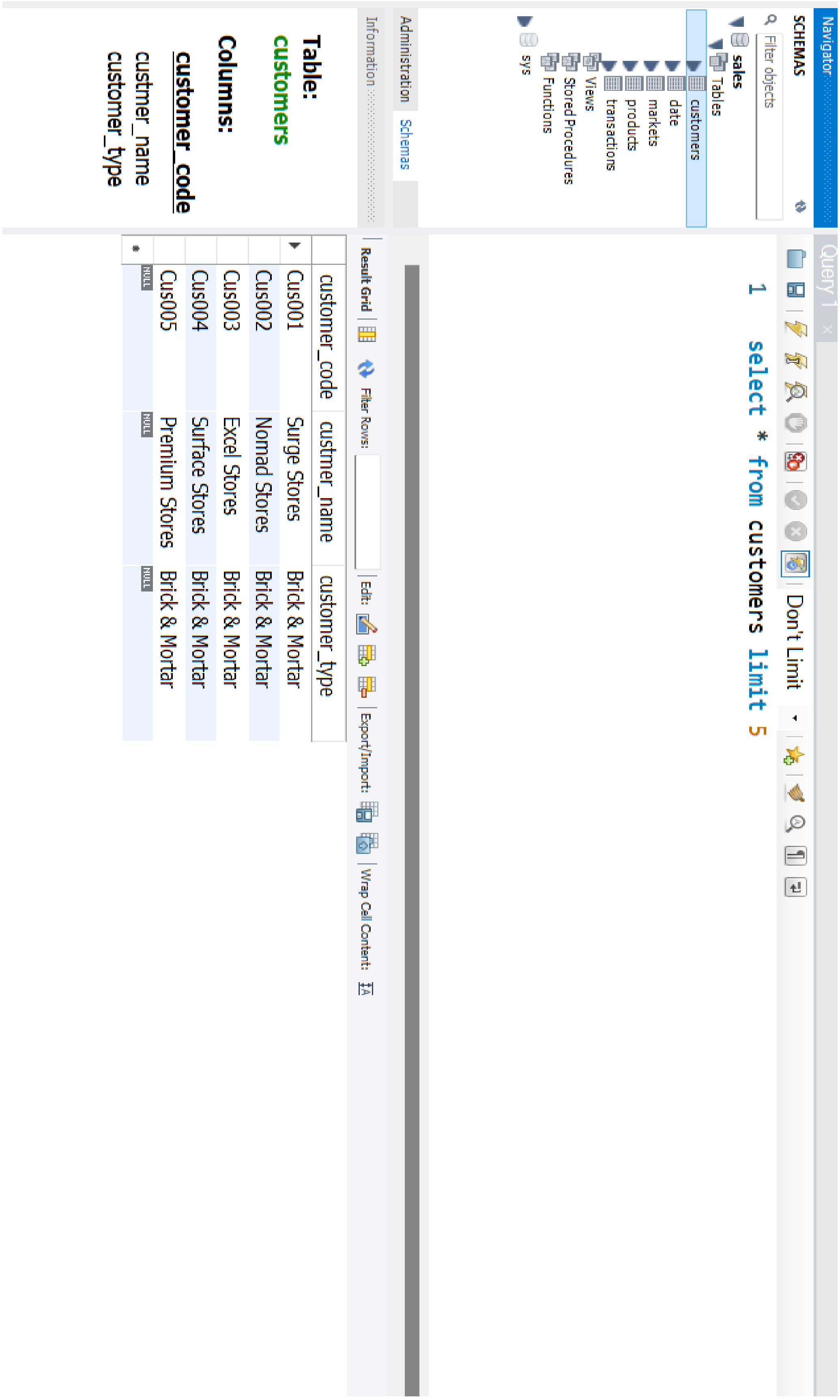
db_dump.sql






MySQL Workbench
SalesDataAnalyst
File Edit View Query Database Server Tools Scripting Help
Navigator
SCHEMAS
Filter objects
sales
Tables
customers
date
markets
products
transactions
Views
Stored Procedures
Functions
sys
Administration Schemas
Information
No object selected
Object Info Session
SQLAdditions
My Snippets
Context Help Snippets
Output

Customers:


Navigator
SCHEMAS
Filter objects
sales
Tables
customers
date
markets
products
transactions
Views
Stored Procedures
Functions
sys
Administration
Schemas
Information
Table: customers
Columns:
customer_code
custmer_name
customer_type
Query 1
Don't Limit
1 select * from customers limit 5
Result Grid
Filter Rows:
Edit:
Export/Import:
Wrap Cell Content:
customer_code
custmer_name
customer_type
Cus001
Surge Stores
Brick & Mortar
Cus002
Nomad Stores
Brick & Mortar
Cus003
Excel Stores
Brick & Mortar
Cus004
Surface Stores
Brick & Mortar
Cus005
Premium Stores
Brick & Mortar
NULL
NULL
NULL

Products:

product_code	product_type
Prod001	Own Brand
Prod002	Own Brand
Prod003	Own Brand
Prod004	Own Brand
Prod005	Own Brand
NULL	NULL

Date:

date	cy_date	year	month_name	date_yy_mmm
2017-06-01	2017-06-01	2017	June	17-Jun
2017-06-02	2017-06-01	2017	June	17-Jun
2017-06-03	2017-06-01	2017	June	17-Jun
2017-06-04	2017-06-01	2017	June	17-Jun
2017-06-05	2017-06-01	2017	June	17-Jun
NULL	NULL	NULL	NULL	NULL

Markets:

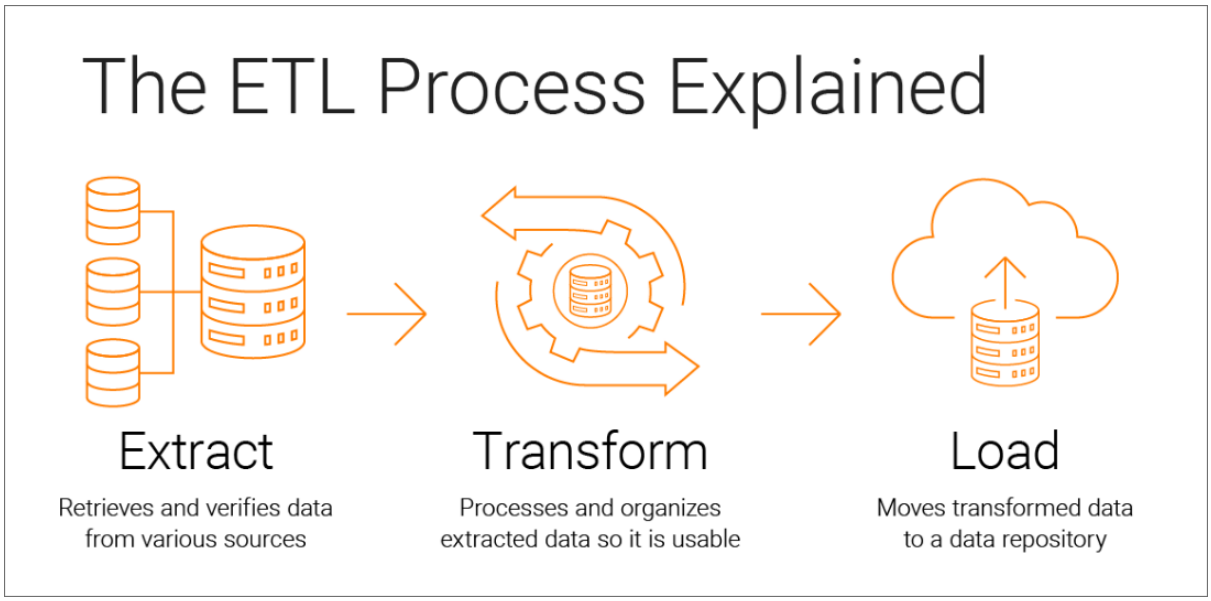
markets_code	markets_name	zone
Mark001	Chennai	South
Mark002	Mumbai	Central
Mark003	Ahmedabad	North
Mark004	Delhi NCR	North
Mark005	Kanpur	North
NULL	NULL	NULL

Transactions:

product_code	customer_code	market_code	order_date	sales_qty	sales_amount	currency
Prod001	Cus001	Mark001	2017-10-10	100	41241	INR
Prod001	Cus002	Mark002	2018-05-08	3	-1	INR
Prod002	Cus003	Mark003	2018-04-06	1	875	INR
Prod002	Cus003	Mark003	2018-04-11	1	583	INR
Prod002	Cus004	Mark003	2018-06-18	6	7176	INR

ETL (Extract, Transform and Load):

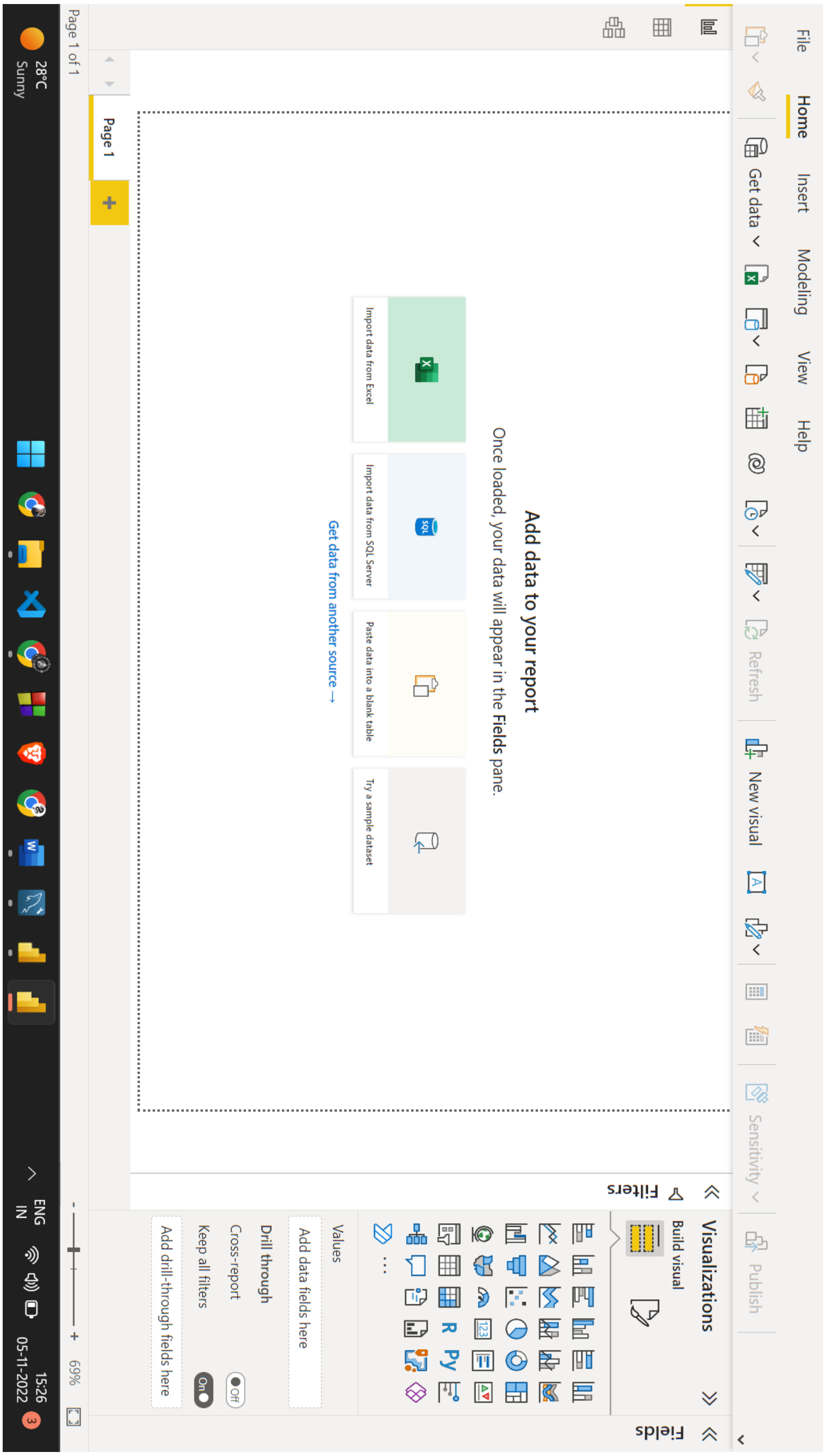
The easiest way to understand how ETL works is to understand what happens in each step of the process.



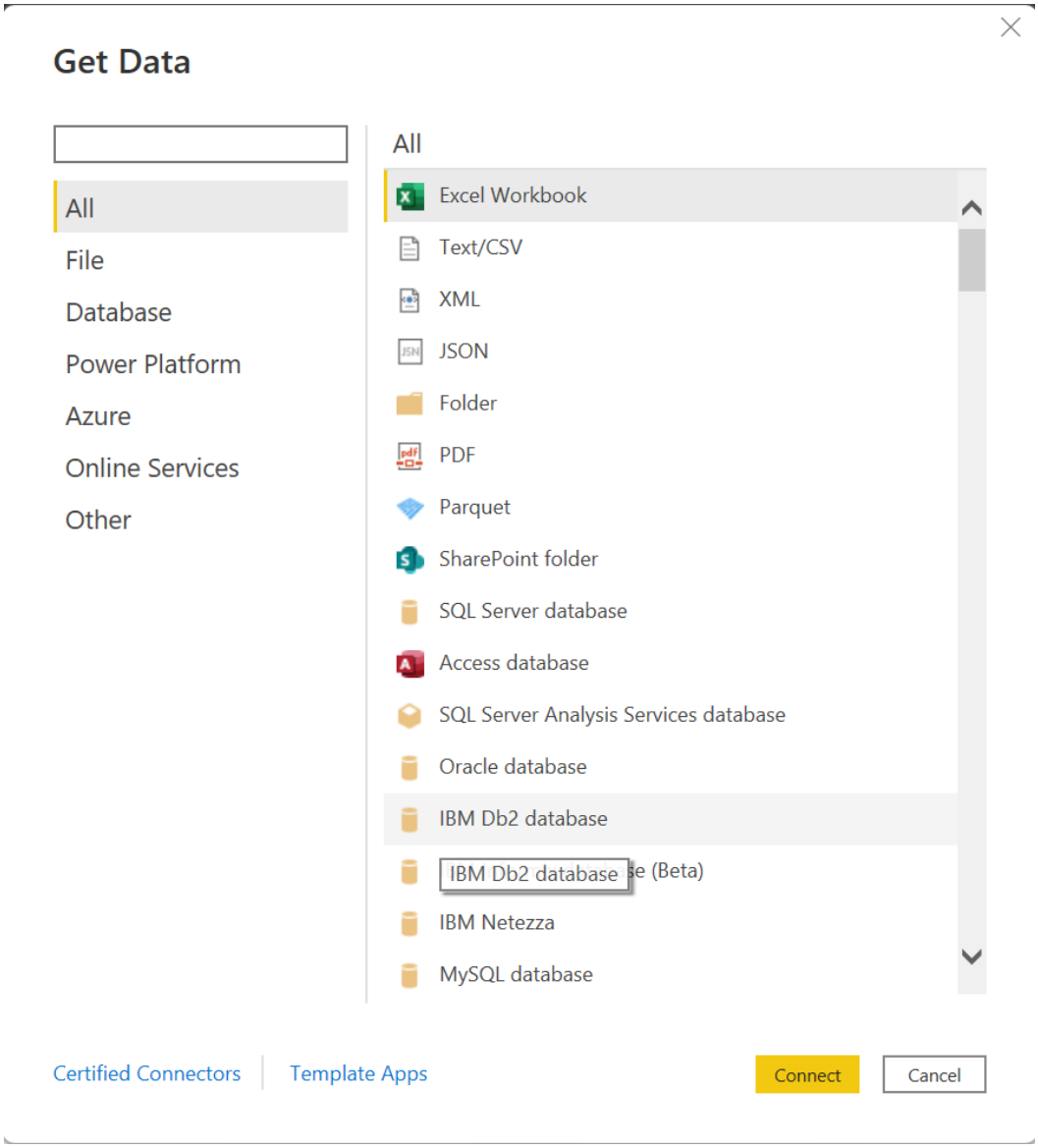
Extract

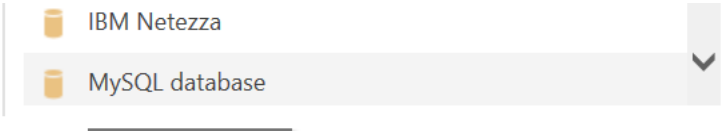
During data extraction, raw data is copied or exported from source locations to a staging area. Data management teams can extract data from a variety of data sources, which can be structured or unstructured. Those sources include but are not limited to:

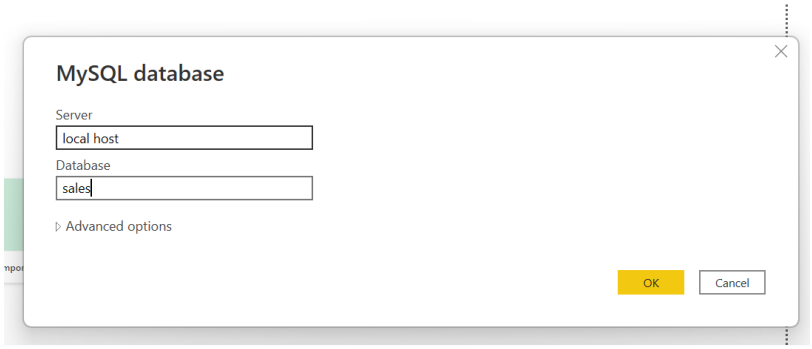
- SQL or NoSQL servers
- CRM and ERP systems
- Flat files
- Email
- Web pages

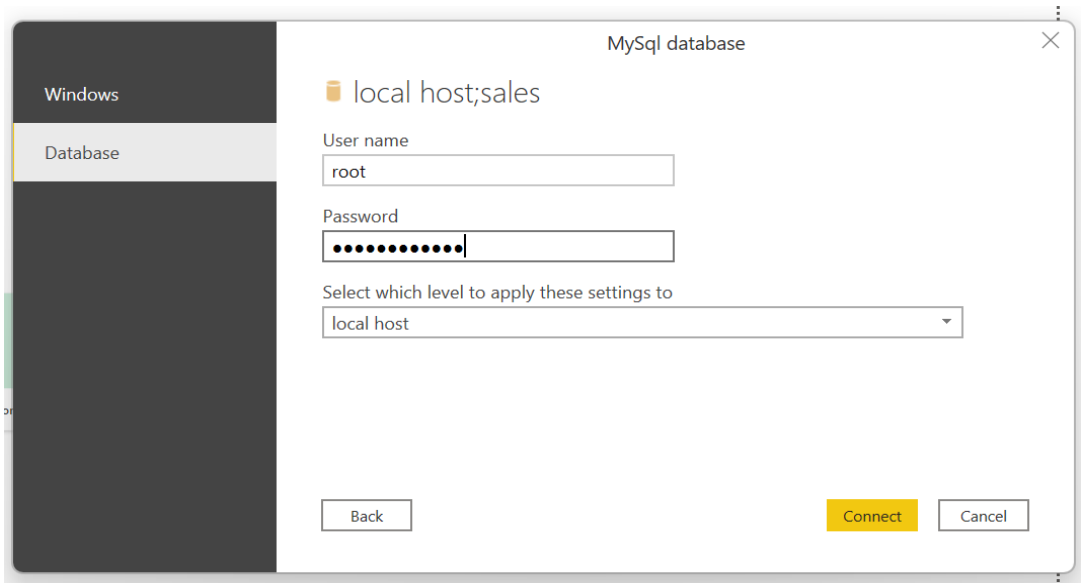


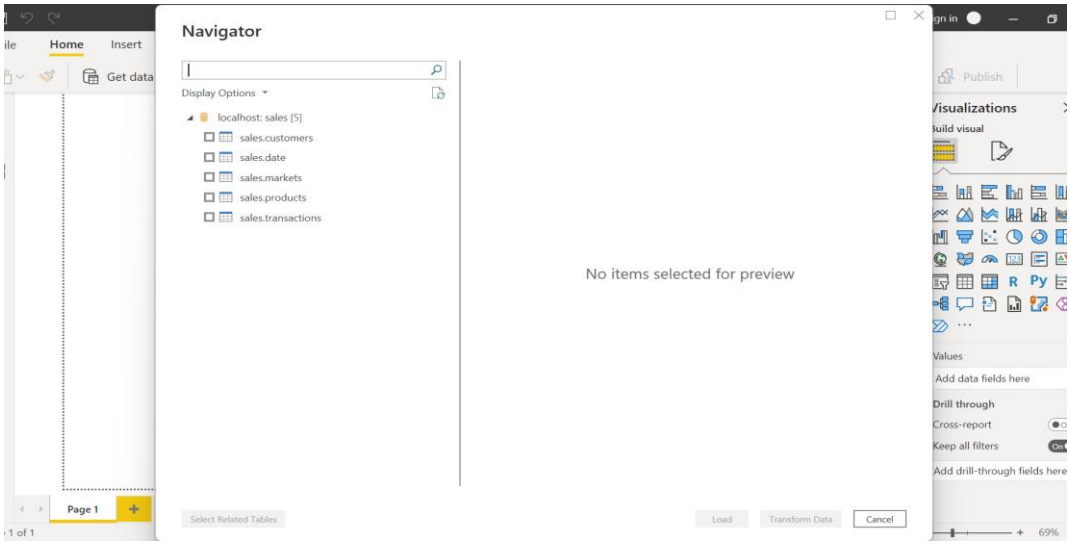
Get data from another source →

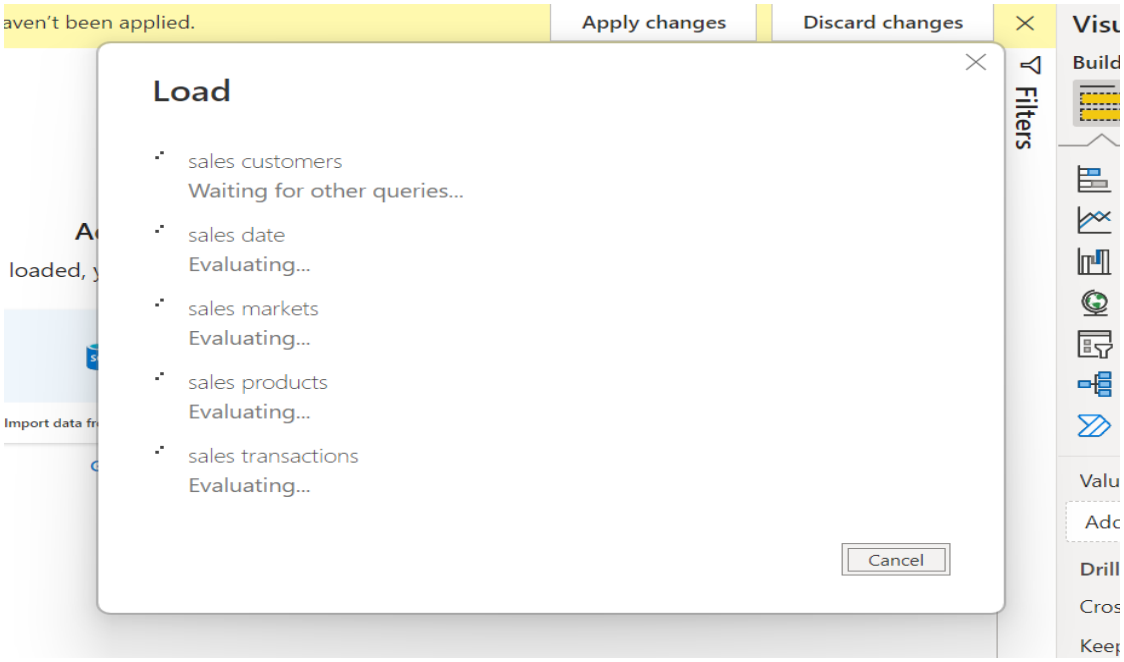

Get Data
All
File
Database
Power Platform
Azure
Online Services
Other
All
Excel Workbook
Text/CSV
XML
JSON
Folder
PDF
Parquet
SharePoint folder
SQL Server database
Access database
SQL Server Analysis Services database
Oracle database
IBM Db2 database
IBM Db2 database
(Beta)
IBM Netezza
MySQL database
Certified Connectors
Template Apps
Connect
Cancel


IBM Netezza
MySQL database


MySQL database
Server
local host
Database
sales
Advanced options
OK
Cancel


MySql database
Windows
Database
local host;sales
User name
root
Password
Select which level to apply these settings to
local host
Back
Connect
Cancel

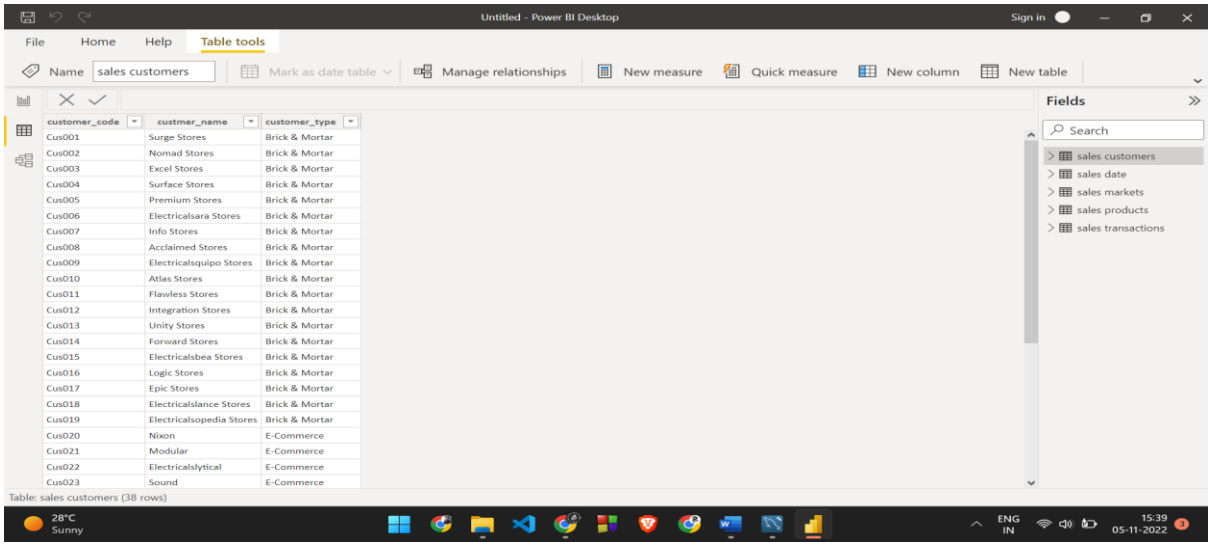

Navigator
Home
Insert
Get data
Display Options
localhost: sales [5]
sales.customers
sales.date
sales.markets
sales.products
sales.transactions
No items selected for preview
Page 1
Select Related Tables
Load
Transform Data
Cancel
Publish
Visualizations
Build visual
Values
Add data fields here
Drill through
Cross-report
Keep all filters
Add drill-through fields here

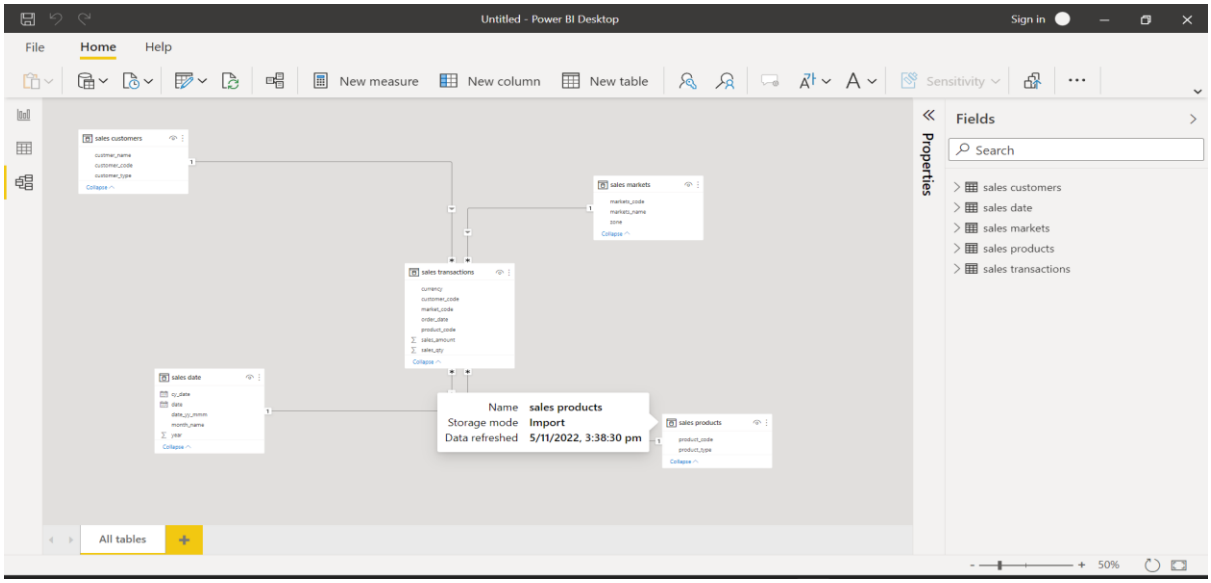

aven't been applied.
Apply changes
Discard changes
Load
sales customers
Waiting for other queries...
sales date
Evaluating...
sales markets
Evaluating...
sales products
Evaluating...
sales transactions
Evaluating...
Cancel
Filters

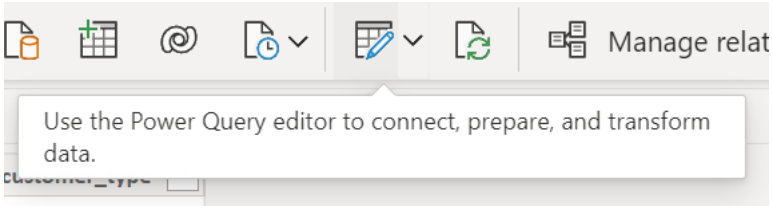
Transform

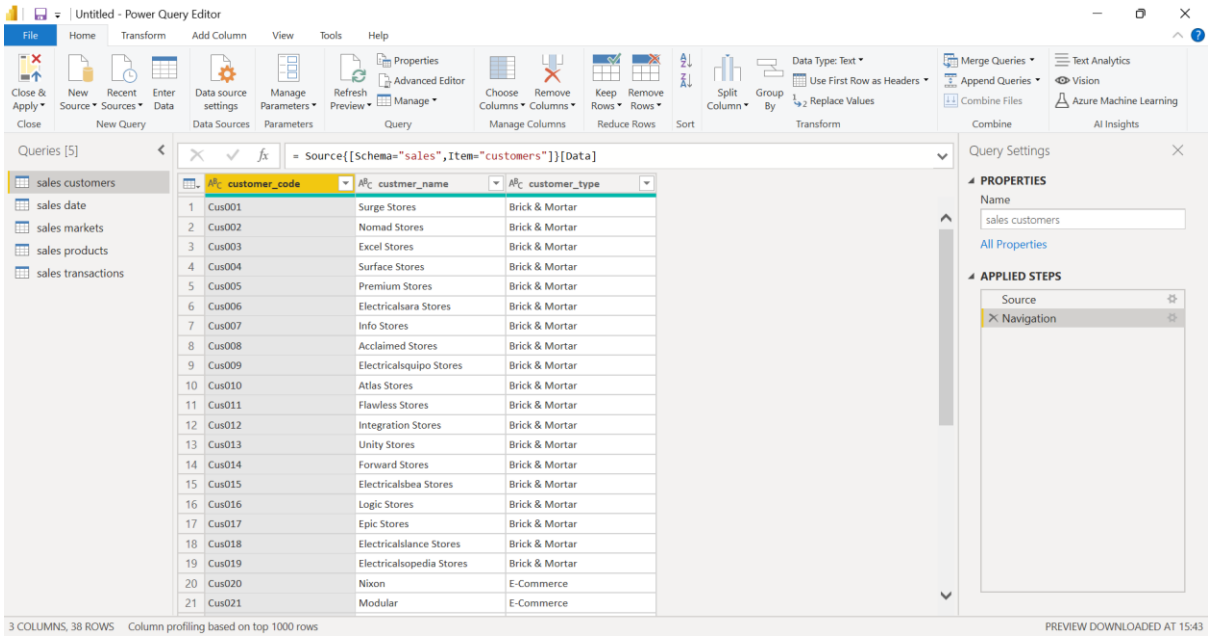
In the staging area, the raw data undergoes data processing. Here, the data is transformed and consolidated for its intended analytical use case. This phase can involve the following tasks:

- Filtering, cleansing, de-duplicating, validating, and authenticating the data.
- Performing calculations, translations, or summarizations based on the raw data. This can include changing row and column headers for consistency, converting currencies or other units of measurement, editing text strings, and more.
- Conducting audits to ensure data quality and compliance
- Removing, encrypting, or protecting data governed by industry or governmental regulators
- Formatting the data into tables or joined tables to match the schema of the target data warehouse.

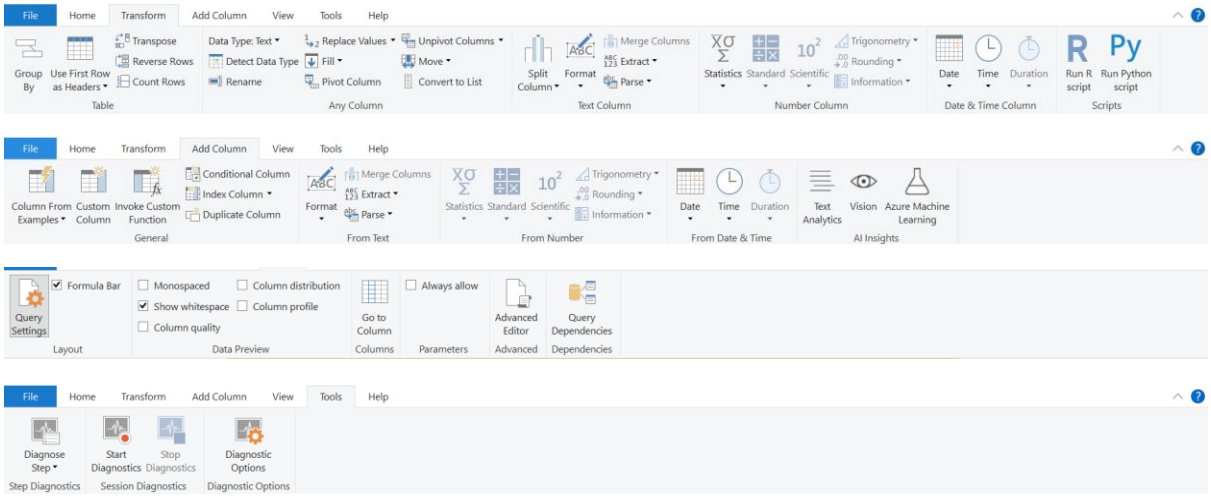






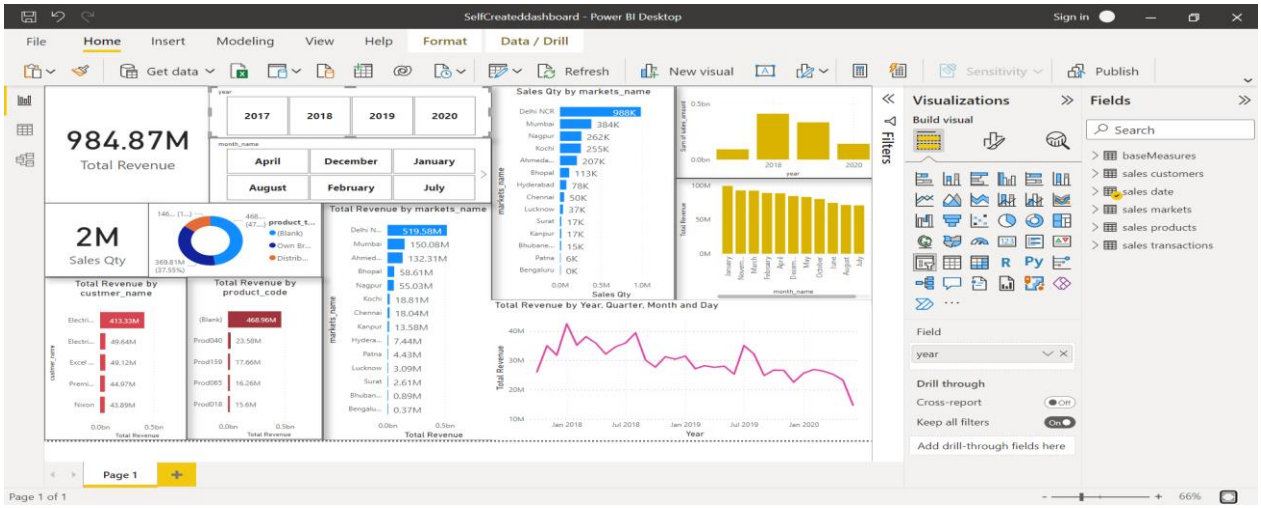


As we can see Power BI has many many great tools to transform data.



LOAD

In this last step, the transformed data is moved from the staging area into a target data warehouse. Typically, this involves an initial loading of all data, followed by periodic loading of incremental data changes and, less often, full refreshes to erase and replace data in the warehouse. For most organizations that use ETL, the process is automated, well-defined, continuous and batch-driven. Typically, ETL takes place during off-hours when traffic on the source systems and the data warehouse is at its lowest.

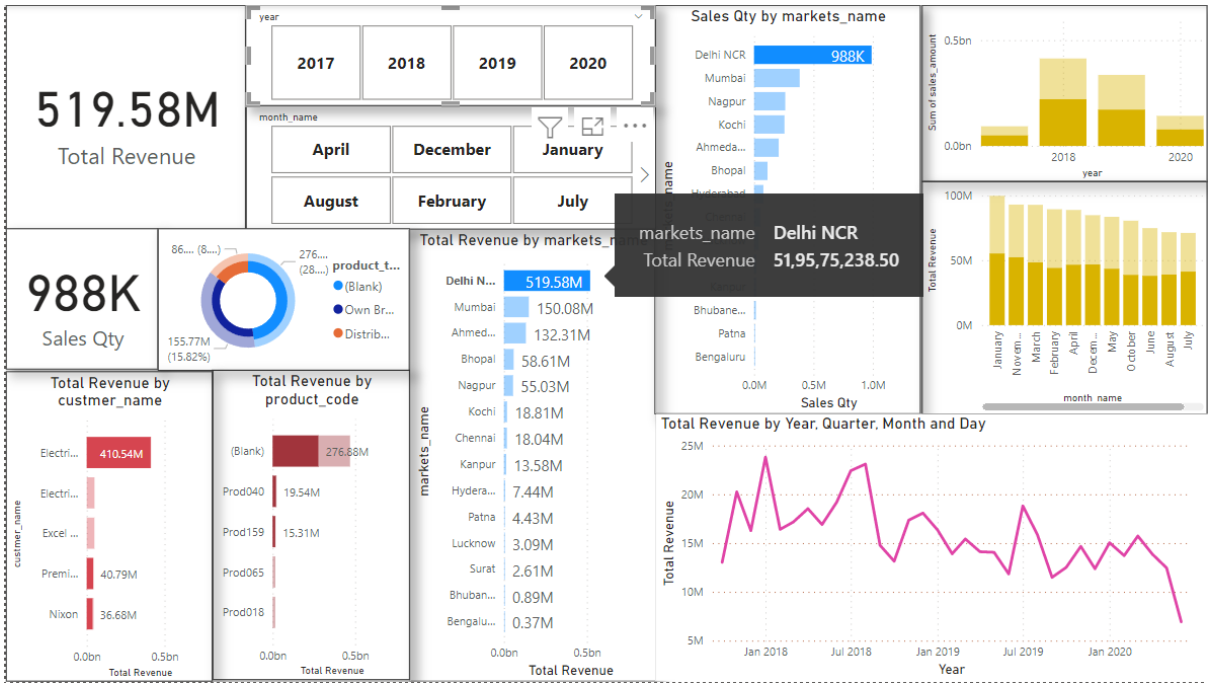


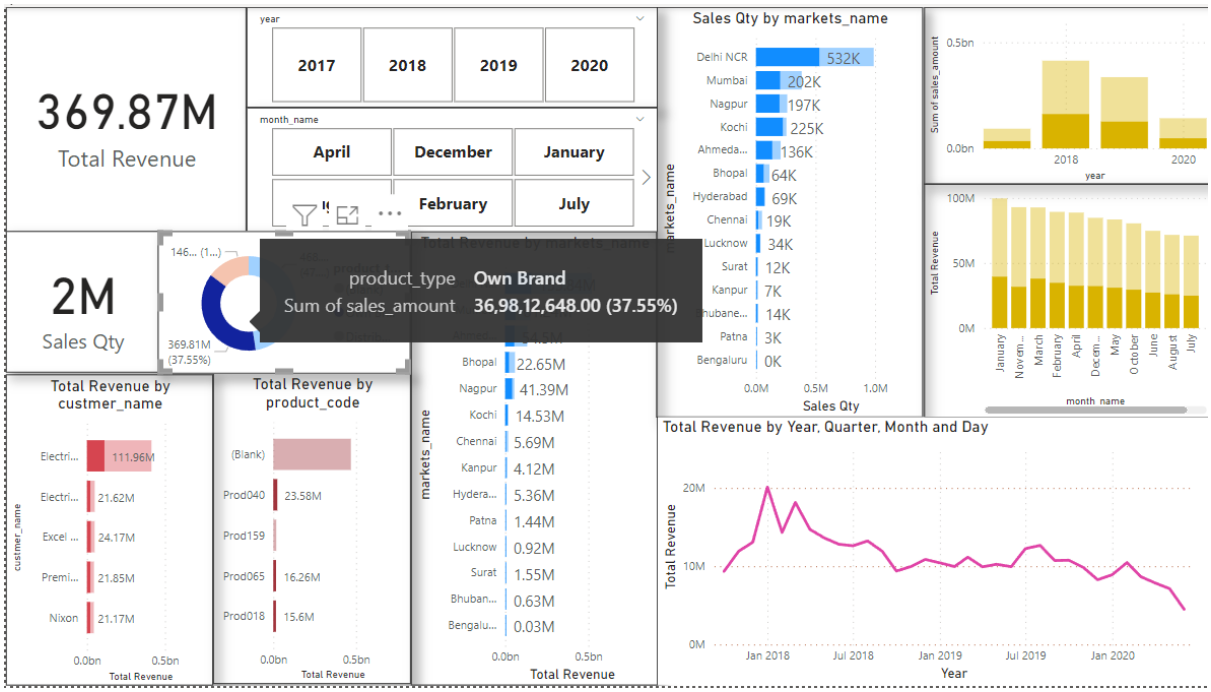
Final Result:

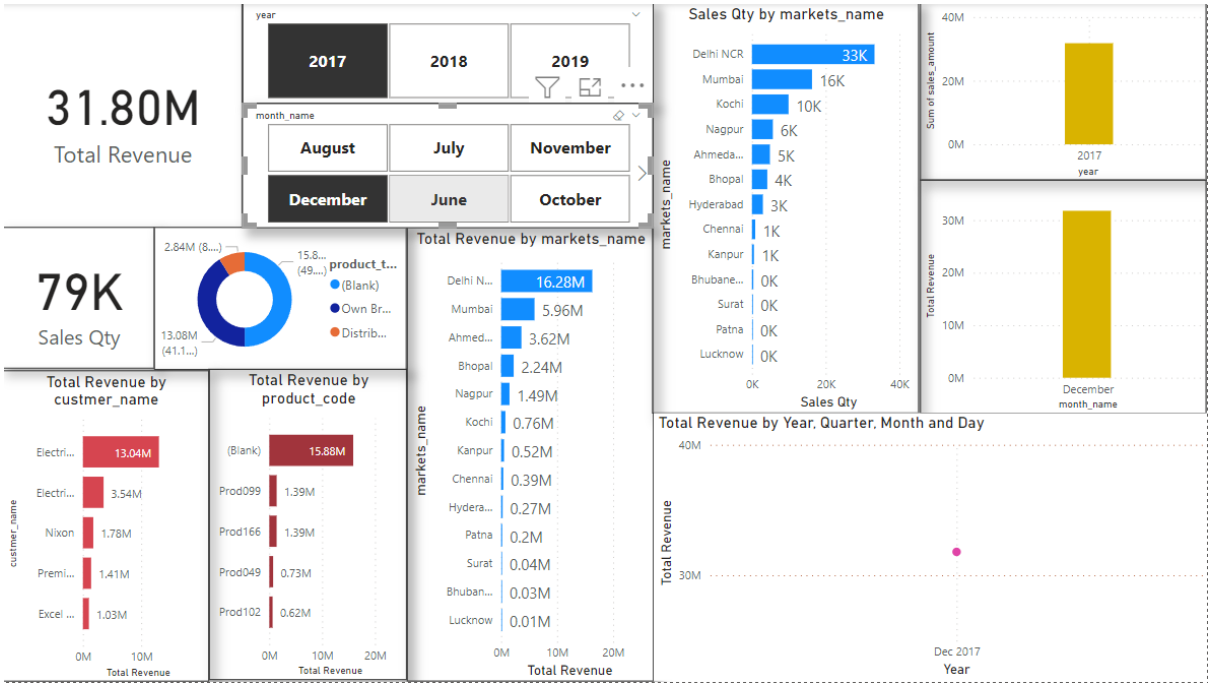
Final Result of this project easy, good and clean dashboard representing all data required for buisness analysis.

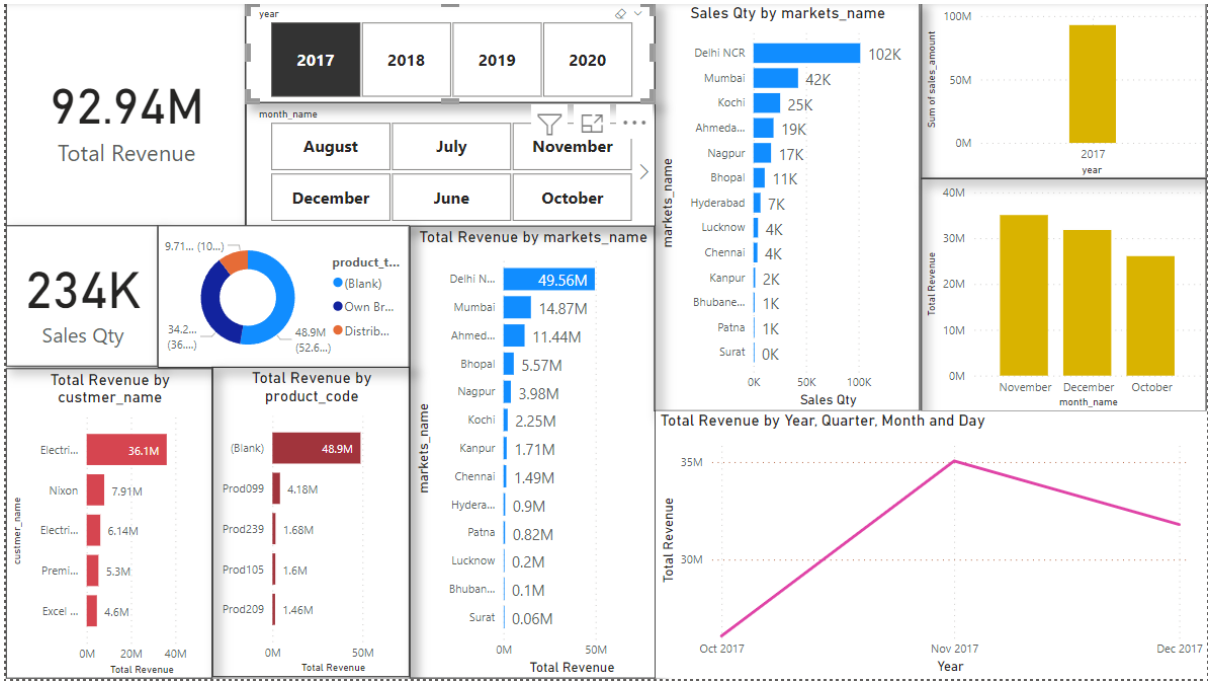
Desktop Dashboard:

Filter is apllied to all visulizations according to the option we choose. We can choose either to keep the filter or remove it.

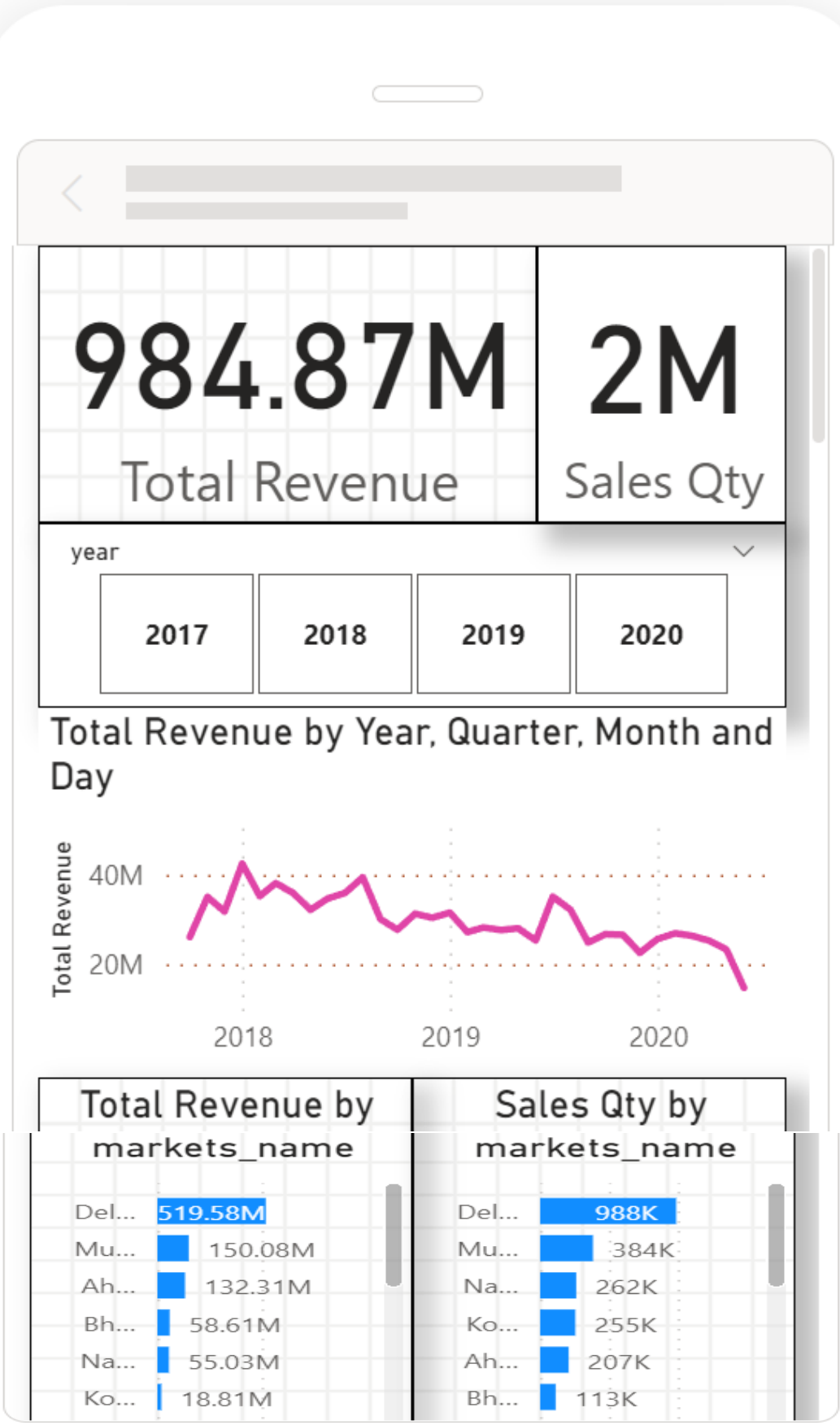





31.80M
Total Revenue
year
2017
2018
2019
month_name
August
July
November
December
June
October
79K
Sales Qty
product_t...
(Blank)
Own Br...
Distrib...
Total Revenue by markets_name
Delhi N... 16.28M
Mumbai 5.96M
Ahmed... 3.62M
Bhopal 2.24M
Nagpur 1.49M
Kochi 0.76M
Kanpur 0.52M
Chennai 0.39M
Hydera... 0.27M
Patna 0.2M
Surat 0.04M
Bhuban... 0.03M
Lucknow 0.01M
Total Revenue by custmer_name
Electri... 13.04M
Electri... 3.54M
Nixon 1.78M
Premi... 1.41M
Excel ... 1.03M
Total Revenue by product_code
(Blank) 15.88M
Prod099 1.39M
Prod166 1.39M
Prod049 0.73M
Prod102 0.62M
Sales Qty by markets_name
Delhi NCR 33K
Mumbai 16K
Kochi 10K
Nagpur 6K
Ahmeda... 5K
Bhopal 4K
Hyderabad 3K
Chennai 1K
Kanpur 1K
Bhubane... 0K
Surat 0K
Patna 0K
Lucknow 0K
Total Revenue by Year, Quarter, Month and Day
Dec 2017
Year


92.94M
Total Revenue
year
2017
2018
2019
2020
month_name
August
July
November
December
June
October
234K
Sales Qty
product_t...
(Blank)
Own Br...
Distrib...
Total Revenue by markets_name
Delhi N... 49.56M
Mumbai 14.87M
Ahmed... 11.44M
Bhopal 5.57M
Nagpur 3.98M
Kochi 2.25M
Kanpur 1.71M
Chennai 1.49M
Hydera... 0.9M
Patna 0.82M
Lucknow 0.2M
Bhuban... 0.1M
Surat 0.06M
Total Revenue by custmer_name
Electri... 36.1M
Nixon 7.91M
Electri... 6.14M
Premi... 5.3M
Excel ... 4.6M
Total Revenue by product_code
(Blank) 48.9M
Prod099 4.18M
Prod239 1.68M
Prod105 1.6M
Prod209 1.46M
Sales Qty by markets_name
Delhi NCR 102K
Mumbai 42K
Kochi 25K
Ahmeda... 19K
Nagpur 17K
Bhopal 11K
Hyderabad 7K
Lucknow 4K
Chennai 4K
Kanpur 2K
Bhubane... 1K
Patna 1K
Surat 0K
Total Revenue by Year, Quarter, Month and Day
Oct 2017
Nov 2017
Dec 2017
Year

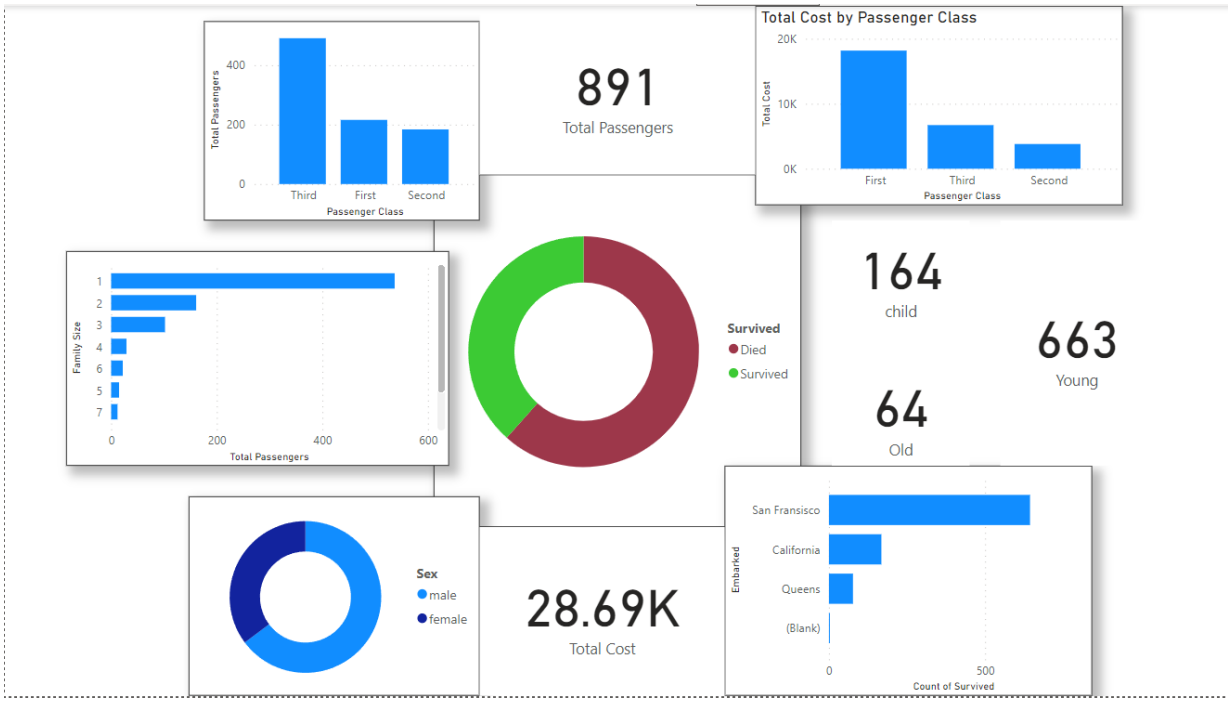
Mobile Dashboard



Project 2: (Titanic Survival Data)
Ms Exel data set:
Cleaning and Visulization

PassengerI	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Surv
1	3	Braund, Mr. Owen Harris	male	22	1	0	A/5 21171	7.25		S	
2	1	Cumings, Mrs. John Bradley (Florence Briggs Thayer)	female	38	1	0	PC 17599	71.2833	C85	C	
3	3	Heikkinen, Miss. Laina	female	26	0	0	STON/O2.	7.925		S	
4	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35	1	0	113803	53.1	C123	S	
5	3	Allen, Mr. William Henry	male	35	0	0	373450	8.05		S	
6	3	Moran, Mr. James	male		0	0	330877	8.4583		Q	
7	1	McCarthy, Mr. Timothy J	male	54	0	0	17463	51.8625	E46	S	
8	3	Palsson, Master. Gosta Leonard	male	2	3	1	349909	21.075		S	
9	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27	0	2	347742	11.1333		S	
10	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14	1	0	237736	30.0708		C	
11	3	Sandstrom, Miss. Marguerite Rut	female	4	1	1	PP 9549	16.7	G6	S	
12	1	Bonnell, Miss. Elizabeth	female	58	0	0	113783	26.55	C103	S	
13	3	Saundercock, Mr. William Henry	male	20	0	0	A/5. 2151	8.05		S	
14	3	Andersson, Mr. Anders Johan	male	39	1	5	347082	31.275		S	
15	3	Vestrom, Miss. Hulda Amanda Adolfina	female	14	0	0	350406	7.8542		S	
16	2	Hewlett, Mrs. (Mary D Kingcome)	female	55	0	0	248706	16		S	
17	3	Rice, Master. Eugene	male	2	4	1	382652	29.125		Q	
18	2	Williams, Mr. Charles Eugene	male		0	0	244373	13		S	
19	3	Vander Planke, Mrs. Julius (Emelia Maria Vandemoortele)	female	31	1	0	345763	18		S	
20	3	Masselmani, Mrs. Fatima	female		0	0	2649	7.225		C	
21	2	Fynney, Mr. Joseph J	male	35	0	0	239865	26		S	
22	2	Beesley, Mr. Lawrence	male	34	0	0	248698	13	D56	S	
23	3	McGowan, Miss. Anna "Annie"	female	15	0	0	330923	8.0292		Q	
24	1	Sloper, Mr. William Thompson	male	28	0	0	113788	35.5	A6	S	
25	3	Palsson, Miss. Torborg Danira	female	8	3	1	349909	21.075		S	
26	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia Johansson)	female	38	1	5	347077	31.3875		S	
27	3	Emir, Mr. Farred Chehab	male		0	0	2631	7.225		C	
28	1	Fortune, Mr. Charles Alexander	male	19	3	2	19950	263	C23 C25 C2	S	
29	3	O'Dwyer, Miss. Ellen "Nellie"	female		0	0	330959	7.8792		Q	
30	3	Todoroff, Mr. Lalio	male		0	0	349216	7.8958		S	

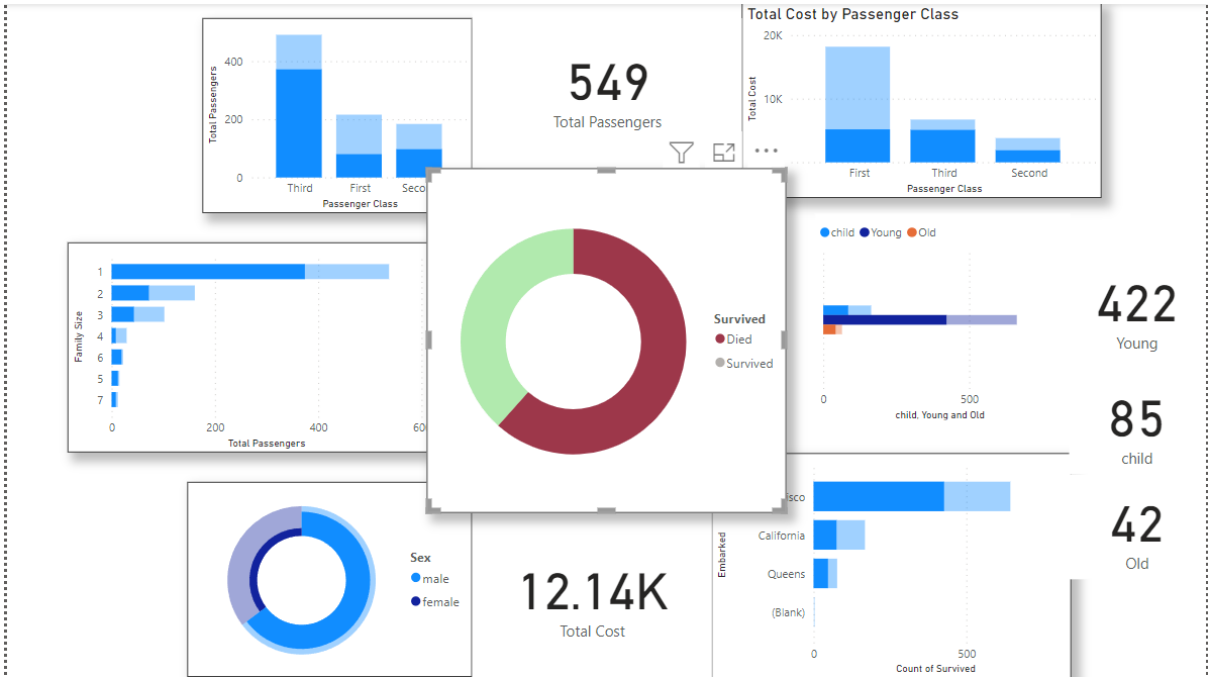
Desktop View: (Titanic Survival Data Visulization)



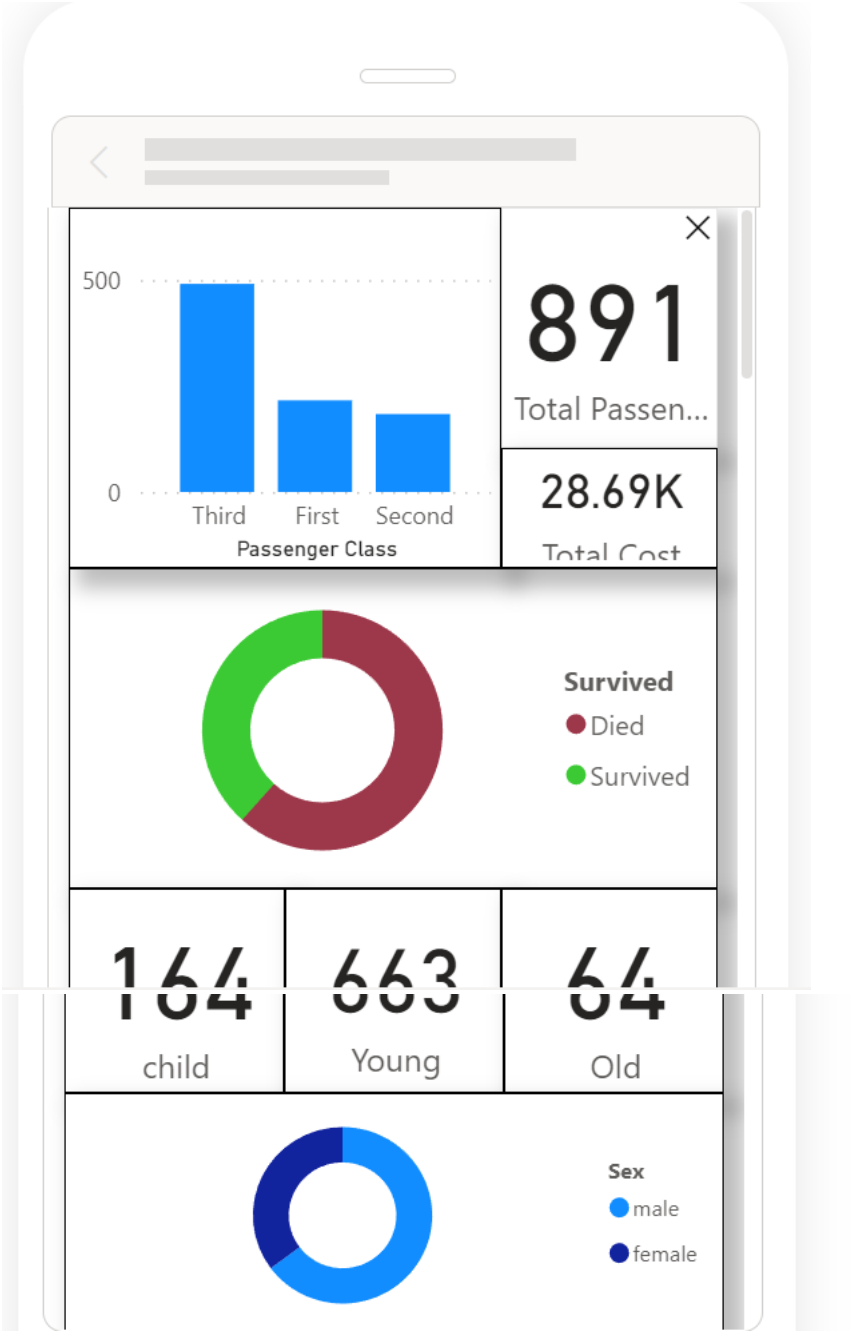
As we can see visuliazation by Power BI is

Dynamic in nature and changes accrding to relationa

Showing most appropriate results.

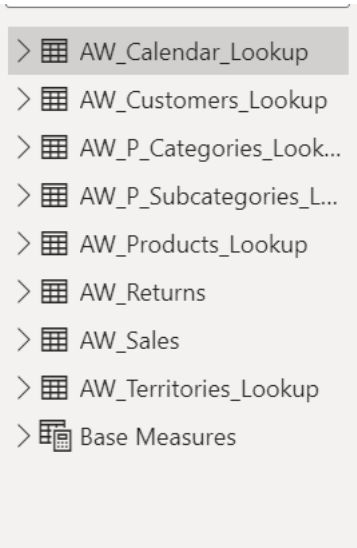


Mobile View:

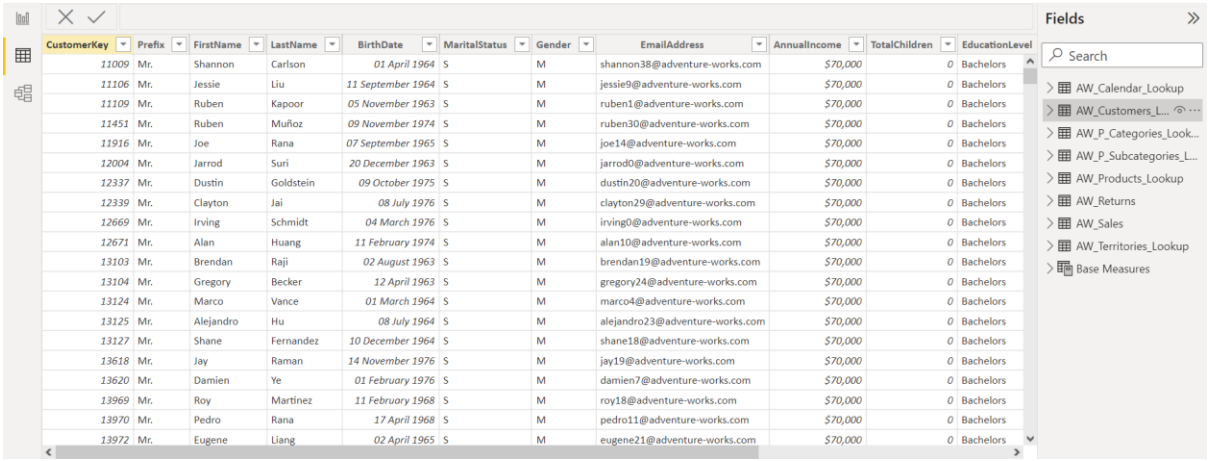

500
0
Third
First
Second
Passenger Class
891
Total Passen...
28.69K
Total Cost
Survived
Died
Survived
164
child
663
Young
64
Old
Sex
male
female

Project 3: Adventure Works Sale

Data:

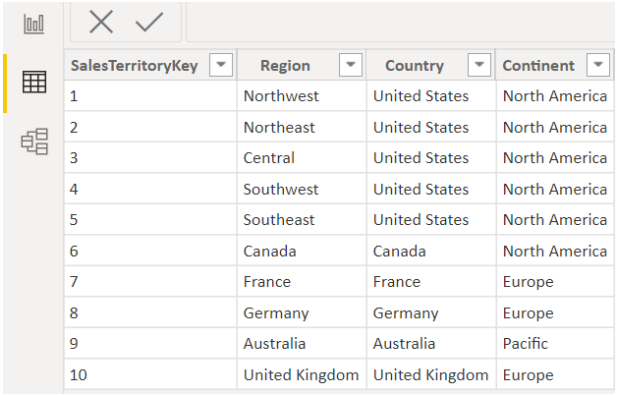


Customer Table:



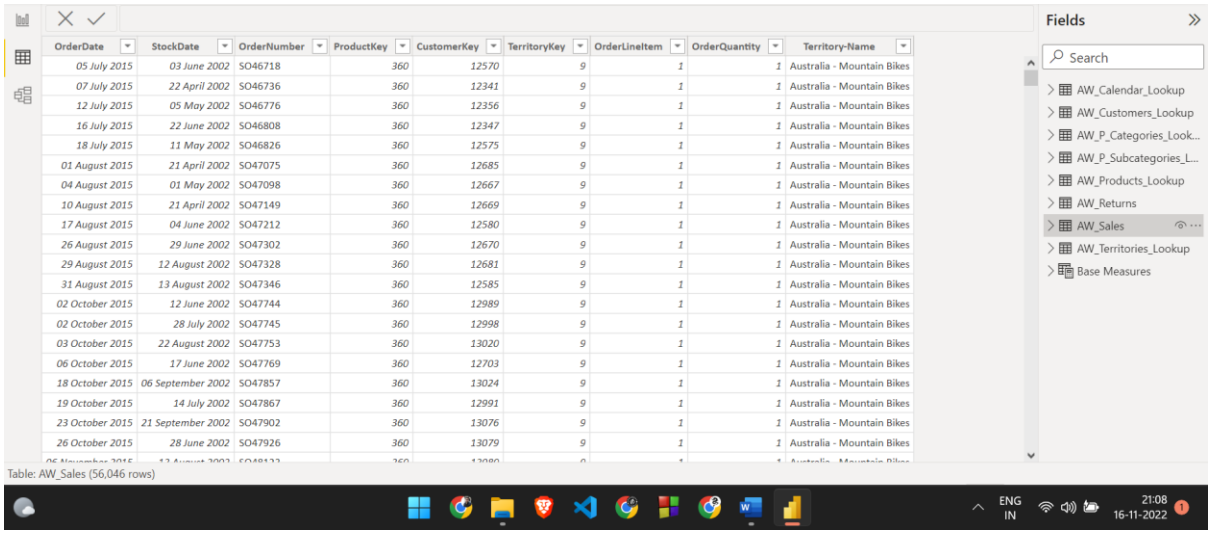
CustomerKey	Prefix	FirstName	LastName	BirthDate	MaritalStatus	Gender	EmailAddress	AnnualIncome	TotalChildren	EducationLevel
11009	Mr.	Shannon	Carlson	01 April 1964	S	M	shannon38@adventure-works.com	$70,000	0	Bachelors
11106	Mr.	Jessie	Liu	11 September 1964	S	M	jessie9@adventure-works.com	$70,000	0	Bachelors
11109	Mr.	Ruben	Kapoor	05 November 1963	S	M	ruben1@adventure-works.com	$70,000	0	Bachelors
11451	Mr.	Ruben	Muñoz	09 November 1974	S	M	ruben30@adventure-works.com	$70,000	0	Bachelors
11916	Mr.	Joe	Rana	07 September 1965	S	M	joe14@adventure-works.com	$70,000	0	Bachelors
12004	Mr.	Jarrod	Suri	20 December 1963	S	M	jarrod0@adventure-works.com	$70,000	0	Bachelors
12337	Mr.	Dustin	Goldstein	09 October 1975	S	M	dustin20@adventure-works.com	$70,000	0	Bachelors
12339	Mr.	Clayton	Jai	08 July 1976	S	M	clayton29@adventure-works.com	$70,000	0	Bachelors
12669	Mr.	Irving	Schmidt	04 March 1976	S	M	irving0@adventure-works.com	$70,000	0	Bachelors
12671	Mr.	Alan	Huang	11 February 1974	S	M	alan10@adventure-works.com	$70,000	0	Bachelors
13103	Mr.	Brendan	Raji	02 August 1963	S	M	brendan19@adventure-works.com	$70,000	0	Bachelors
13104	Mr.	Gregory	Becker	12 April 1963	S	M	gregory24@adventure-works.com	$70,000	0	Bachelors
13124	Mr.	Marco	Vance	01 March 1964	S	M	marco4@adventure-works.com	$70,000	0	Bachelors
13125	Mr.	Alejandro	Hu	08 July 1964	S	M	alejandro23@adventure-works.com	$70,000	0	Bachelors
13127	Mr.	Shane	Fernandez	10 December 1964	S	M	shane18@adventure-works.com	$70,000	0	Bachelors
13618	Mr.	Jay	Raman	14 November 1976	S	M	jay19@adventure-works.com	$70,000	0	Bachelors
13620	Mr.	Damien	Ye	01 February 1976	S	M	damien7@adventure-works.com	$70,000	0	Bachelors
13969	Mr.	Roy	Martinez	11 February 1968	S	M	roy18@adventure-works.com	$70,000	0	Bachelors
13970	Mr.	Pedro	Rana	17 April 1968	S	M	pedro11@adventure-works.com	$70,000	0	Bachelors
13972	Mr.	Eugene	Liang	02 April 1965	S	M	eugene21@adventure-works.com	$70,000	0	Bachelors

Territory:



SalesTerritoryKey	Region	Country	Continent
1	Northwest	United States	North America
2	Northeast	United States	North America
3	Central	United States	North America
4	Southwest	United States	North America
5	Southeast	United States	North America
6	Canada	Canada	North America
7	France	France	Europe
8	Germany	Germany	Europe
9	Australia	Australia	Pacific
10	United Kingdom	United Kingdom	Europe

Sales Tabels:



OrderDate	StockDate	OrderNumber	ProductKey	CustomerKey	TerritoryKey	OrderLineItem	OrderQuantity	Territory-Name
05 July 2015	03 June 2002	SO46718	360	12570	9	1	1	Australia - Mountain Bikes
07 July 2015	22 April 2002	SO46736	360	12341	9	1	1	Australia - Mountain Bikes
12 July 2015	05 May 2002	SO46776	360	12356	9	1	1	Australia - Mountain Bikes
16 July 2015	22 June 2002	SO46808	360	12347	9	1	1	Australia - Mountain Bikes
18 July 2015	11 May 2002	SO46826	360	12575	9	1	1	Australia - Mountain Bikes
01 August 2015	21 April 2002	SO47075	360	12685	9	1	1	Australia - Mountain Bikes
04 August 2015	01 May 2002	SO47098	360	12667	9	1	1	Australia - Mountain Bikes
10 August 2015	21 April 2002	SO47149	360	12669	9	1	1	Australia - Mountain Bikes
17 August 2015	04 June 2002	SO47212	360	12580	9	1	1	Australia - Mountain Bikes
26 August 2015	29 June 2002	SO47302	360	12670	9	1	1	Australia - Mountain Bikes
29 August 2015	12 August 2002	SO47328	360	12681	9	1	1	Australia - Mountain Bikes
31 August 2015	13 August 2002	SO47346	360	12585	9	1	1	Australia - Mountain Bikes
02 October 2015	12 June 2002	SO47744	360	12989	9	1	1	Australia - Mountain Bikes
02 October 2015	28 July 2002	SO47745	360	12998	9	1	1	Australia - Mountain Bikes
03 October 2015	22 August 2002	SO47753	360	13020	9	1	1	Australia - Mountain Bikes
06 October 2015	17 June 2002	SO47769	360	12703	9	1	1	Australia - Mountain Bikes
18 October 2015	06 September 2002	SO47857	360	13024	9	1	1	Australia - Mountain Bikes
19 October 2015	14 July 2002	SO47867	360	12991	9	1	1	Australia - Mountain Bikes
23 October 2015	21 September 2002	SO47902	360	13076	9	1	1	Australia - Mountain Bikes
26 October 2015	28 June 2002	SO47926	360	13079	9	1	1	Australia - Mountain Bikes

Table: AW_Sales (56,046 rows)

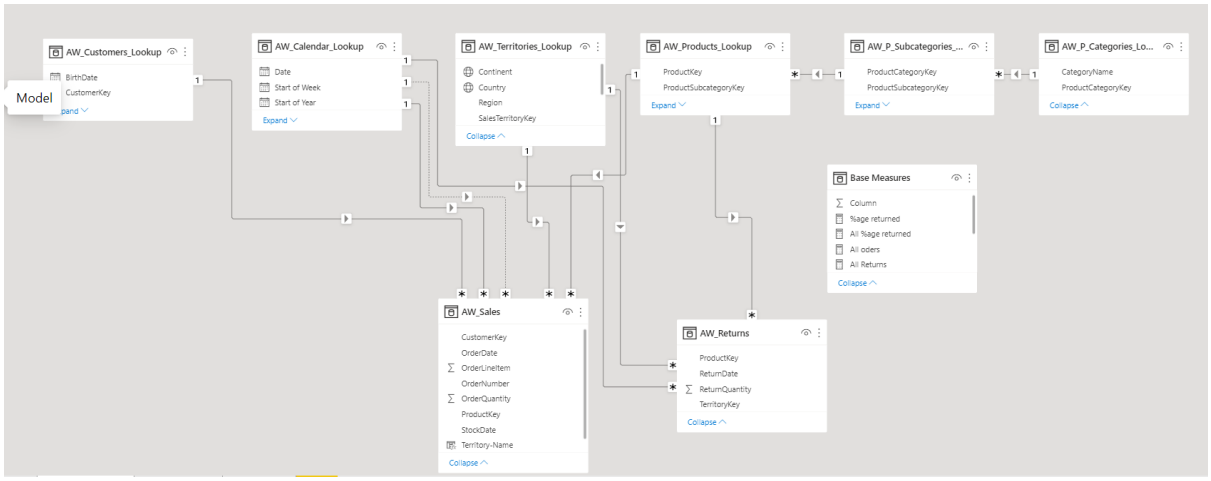
Fields

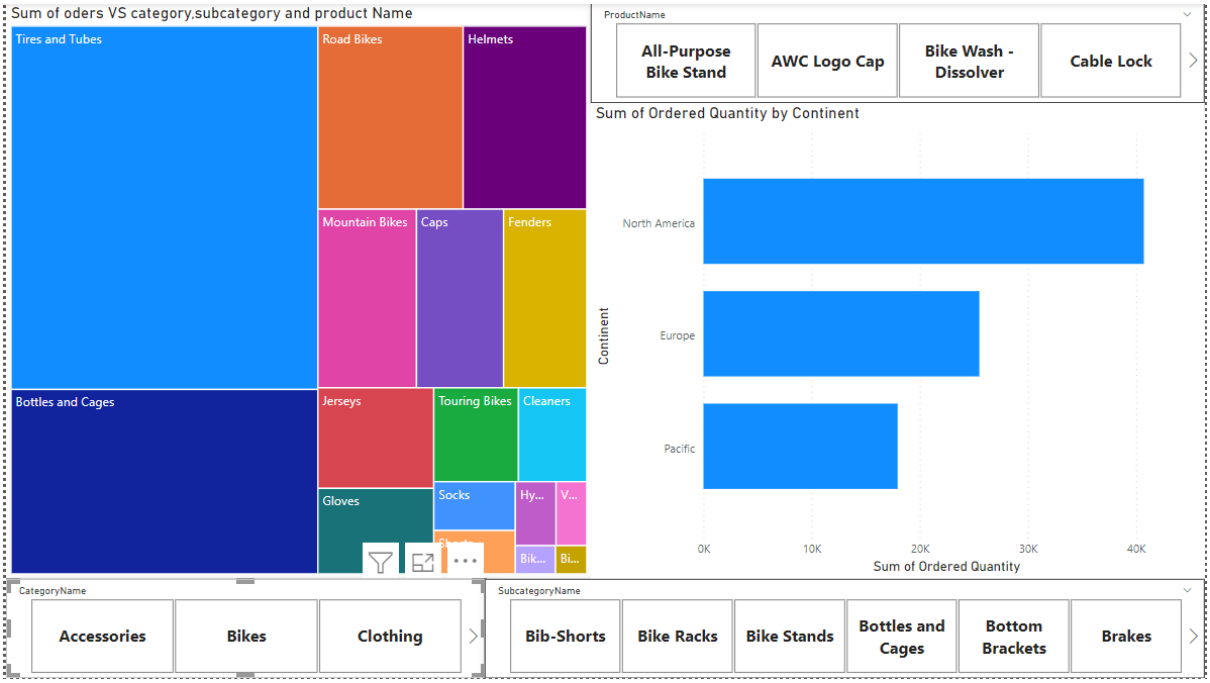
Search

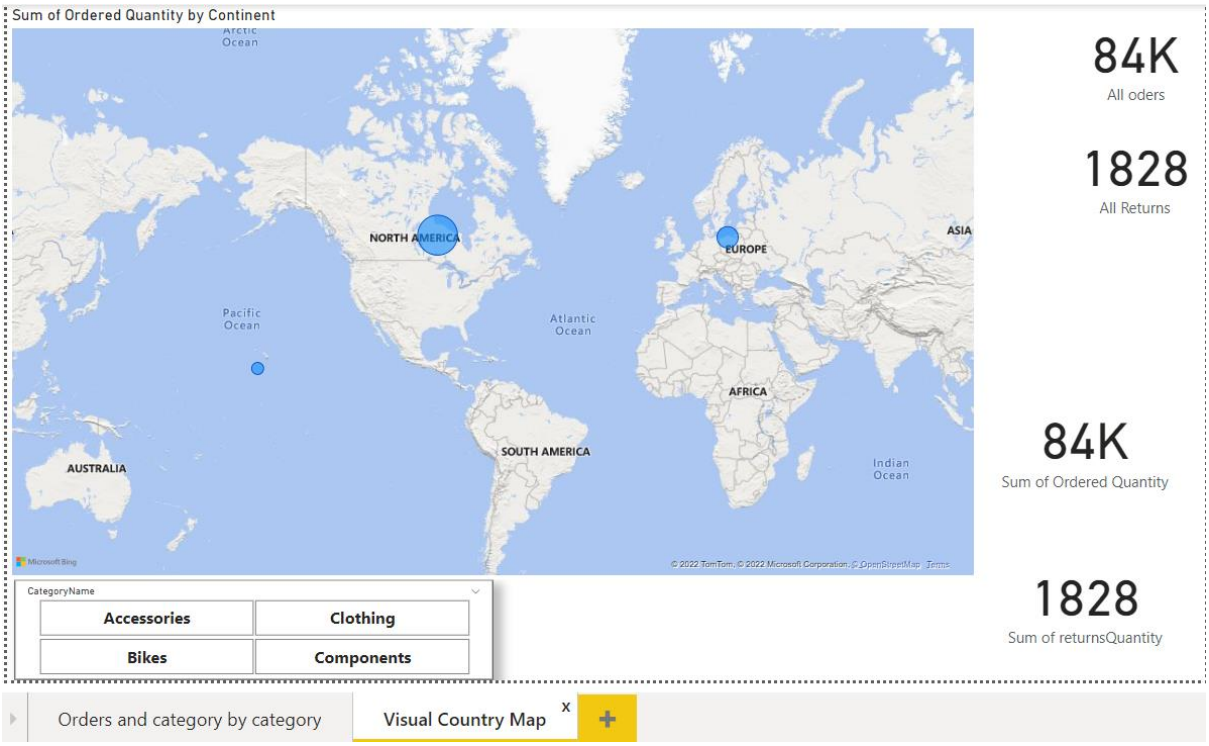
AW_Calendar_Lookup
AW_Customers_Lookup
AW_P_Categories_Look...
AW_P_Subcategories_L...
AW_Products_Lookup
AW_Returns
AW_Sales
AW_Territories_Lookup
Base Measures

ENG IN 21:08 16-11-2022

Table Relations:

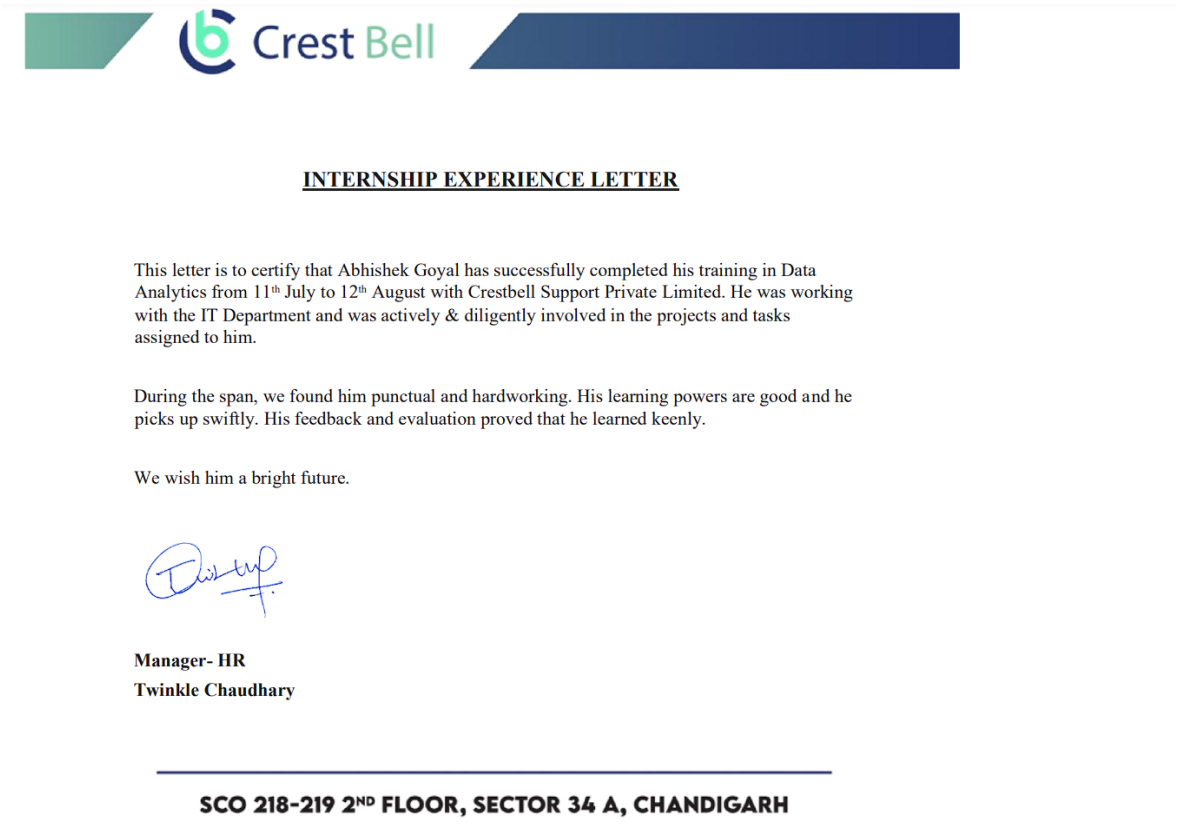



Sum of oders VS category,subcategory and product Name
Tires and Tubes
Road Bikes
Helmets
Mountain Bikes
Caps
Fenders
Bottles and Cages
Jerseys
Touring Bikes
Cleaners
Gloves
Socks
ProductName
All-Purpose Bike Stand
AWC Logo Cap
Bike Wash - Dissolver
Cable Lock
Sum of Ordered Quantity by Continent
North America
Europe
Pacific
Continent
0K
10K
20K
30K
40K
Sum of Ordered Quantity
CategoryName
Accessories
Bikes
Clothing
SubcategoryName
Bib-Shorts
Bike Racks
Bike Stands
Bottles and Cages
Bottom Brackets
Brakes


Sum of Ordered Quantity by Continent
NORTH AMERICA
EUROPE
ASIA
AFRICA
SOUTH AMERICA
AUSTRALIA
Pacific Ocean
Atlantic Ocean
Indian Ocean
84K
All oders
1828
All Returns
84K
Sum of Ordered Quantity
1828
Sum of returnsQuantity
CategoryName
Accessories
Clothing
Bikes
Components
Orders and category by category
Visual Country Map

Conclusion:

Tools like Power BI can be used to perform etl on large data set easily and fast and can be used to represent in easy to understand and eye catching manner. This tool gives many features and data analytic work can be done fastly so that buisness can focus its attention on increasing its customers and providing better customer service.



Crest Bell

INTERNSHIP EXPERIENCE LETTER

This letter is to certify that Abhishek Goyal has successfully completed his training in Data Analytics from 11th July to 12th August with Crestbell Support Private Limited. He was working with the IT Department and was actively & diligently involved in the projects and tasks assigned to him.

During the span, we found him punctual and hardworking. His learning powers are good and he picks up swiftly. His feedback and evaluation proved that he learned keenly.

We wish him a bright future.

Manager- HR
Twinkle Chaudhary

SCO 218-219 2ND FLOOR, SECTOR 34 A, CHANDIGARH

Biblography:

IBM ETL

Udemy

Importance of data analysis

Other tools for data analysis

Power BI Microsoft