

The O'Reilly logo is centered on a vibrant orange background. The word "O'REILLY" is written in a bold, white, sans-serif font. A registered trademark symbol (®) is positioned at the top right of the letter "Y".

O'REILLY®



# Power BI Bootcamp

## Instructors:

Nicolás Lagreste Zucchini

Maria Florencia Hourcouripé



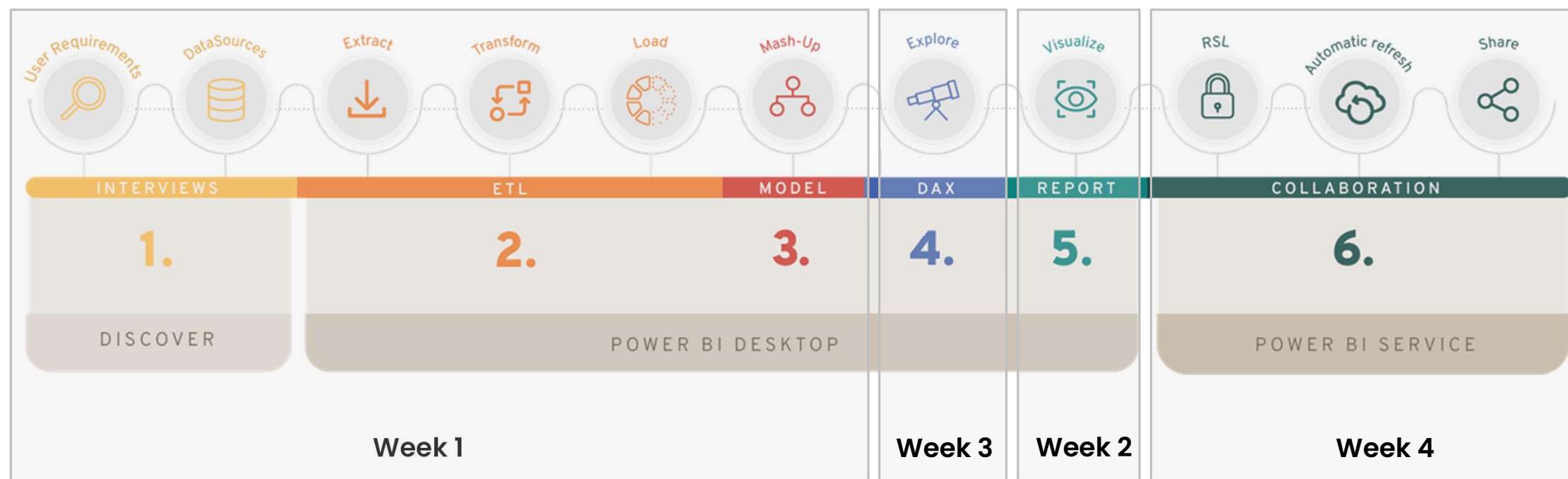


# Course Agenda

- **Week 1:** Introduction to Business Intelligence, Power BI and Power Query.
- **Week 2:** Advanced and Interactive Visuals.
- **Week 3:** Calculated measures, columns & custom tables with Power BI Programming Language (DAX).
- **Week 4:** Ways to collaborate and share in Power BI.



# Our Method: Flow of Report Development



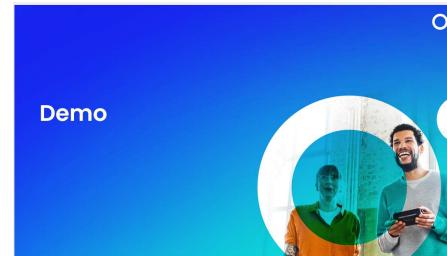


**Our Method on each topic:**

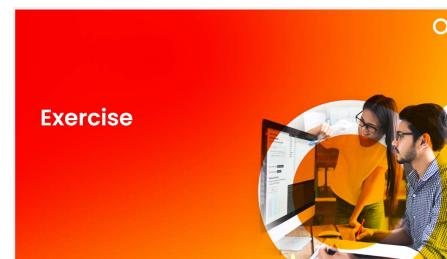
**Demo example, followed by exercises to practice!**

**We will develop two projects in PowerBI:**

- **DEMO – Flor & Nico**



- **EXERCISE – Your turn!**



\*Student should have the latest version of **Power BI Desktop installed.**  
**March 2023 or later.**



## About us:



Nicolás Lagreste Zucchini [in](#)

nico@analyticmood.com  
 www.analyticmood.com

analytic mood

- Co-founder Analytic Mood
- Data Expert and Business Analyst in BI Applications, NTT Data, Kabel Systems, Grupo Solutio and Analytic Mood.
- He is responsible for Microsoft Power BI training and full lifecycle development of next-generation software, including initial requirement gathering to design, coding, testing, documentation and implementation.
- Active participant in the Power BI community. He participated as a speaker in the last Power BI Summit.



María Florencia Hourcouripe [in](#)

flor@analyticmood.com  
 www.analyticmood.com

analytic mood

- Co-founder Analytic Mood
- Master's degree from the EAE Business School in "Business Intelligence & Innovación tecnológica".
- Consultant & trainer in Microsoft Power BI.
- Active participant in the Power BI community. She participated as a speaker in the last Power BI Summit.

O'REILLY®

# WEEK 01

Introduction to  
Business Intelligence,  
Power BI and Power  
Query





# Week 1: Introduction to Business Intelligence, Power BI and Power Query

- What is Business Intelligence and Power BI; What do we mean by Business Intelligence? A brief explanation of what is the maturity cycle of Business Intelligence & Data Science. Understanding that Power BI is a collection of software, services, apps, and connectors.
- What is an ETL and how to build a powerful ETL solution with no code. Defining and creating a Star Schema Model, and translation of all this into visual graphics.
- Brief explanation of the difference between Power BI Desktop and Service. Understand what the main features of each tool or service are.





# Introduction to Business Intelligence and Power BI

- Introduction to Business Intelligence and Power BI Platform.
- Maturity cycle of Business Intelligence & Data Science.
- Working with Power BI projects step by step – Full architecture.
- Exercise: Setting Up Power BI Desktop
- Q&A

## Exercise 0.0 – Initial set up

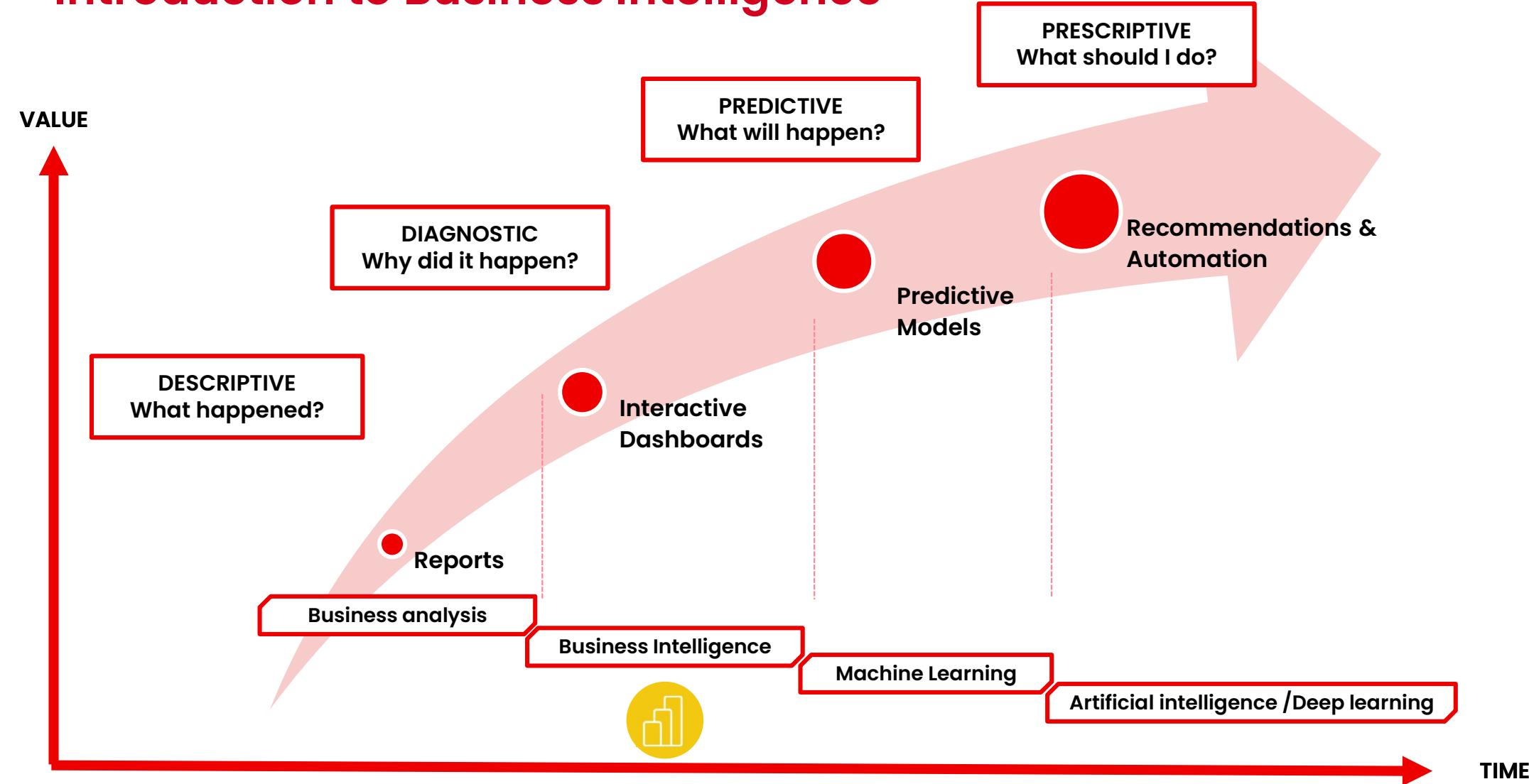
Download the shared .ZIP

- Link 01:  
[bit.ly/PowerBIBootcamp202309](https://bit.ly/PowerBIBootcamp202309)  
Pass: **OREILLY2023**
- Link 02:  
[bit.ly/PowerBIBootcamp202309W01](https://bit.ly/PowerBIBootcamp202309W01)



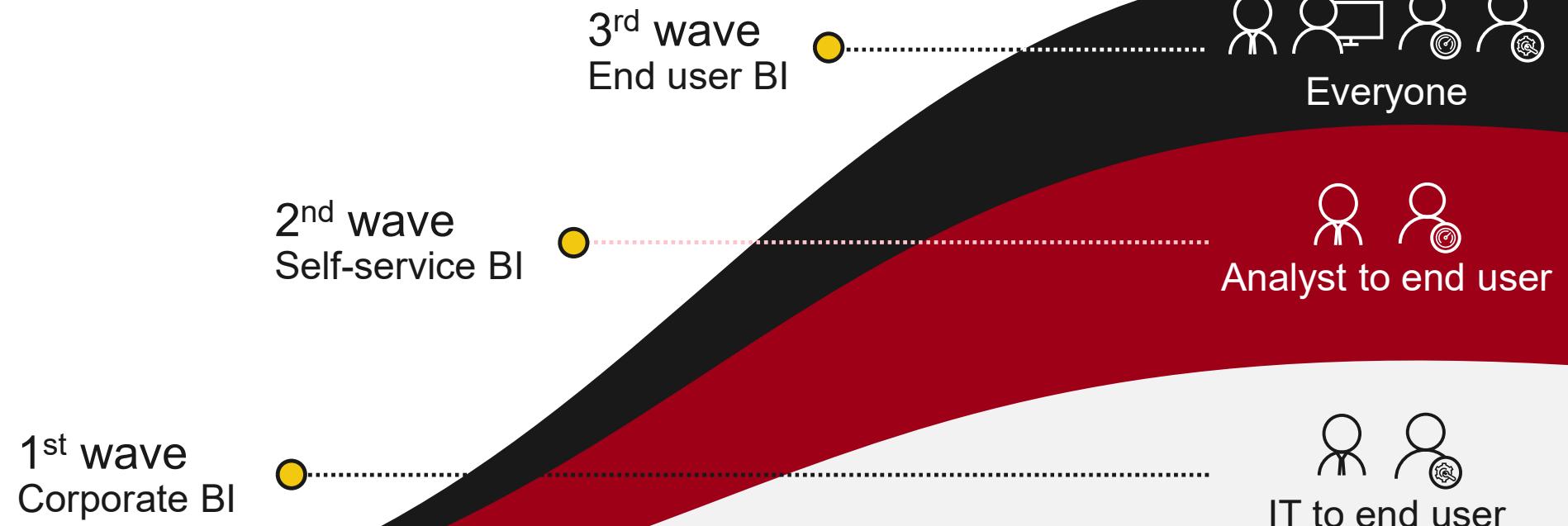
# Introduction to Business Intelligence

VALUE





## Evolution of BI





# Microsoft Power Platform

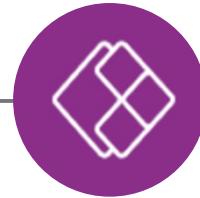
## Give everyone the power to innovate:

Empower everyone at your organization with an intuitive, collaborative, and extensible platform of low-code tools that makes it easy to create efficient and flexible solutions.

**Power BI**



**Power Apps**



**Power Automate**



**Power Virtual Agents**



**Analyze:**

Discover intelligent insights in diverse data

**Act:**

Build low-code solutions to business challenges.

**Automate**

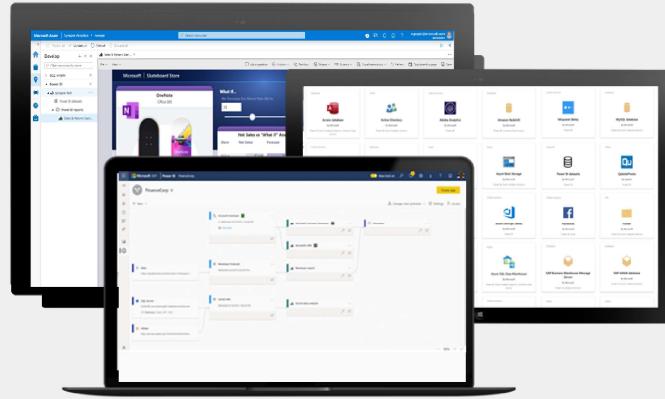
Streamline processes with no-code automation.

**Assist**

Handle routine inquiries at scale with conversational AI.

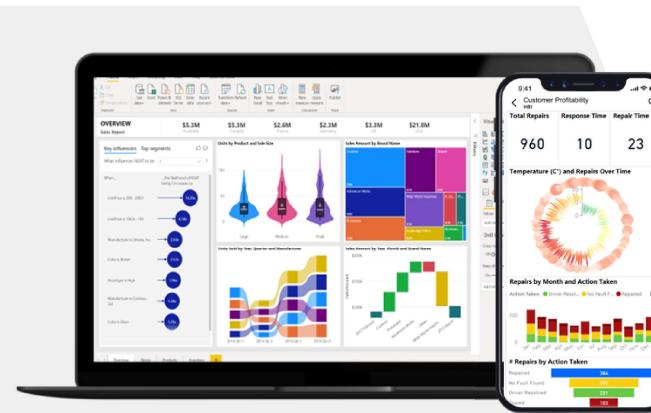


# Microsoft Power BI



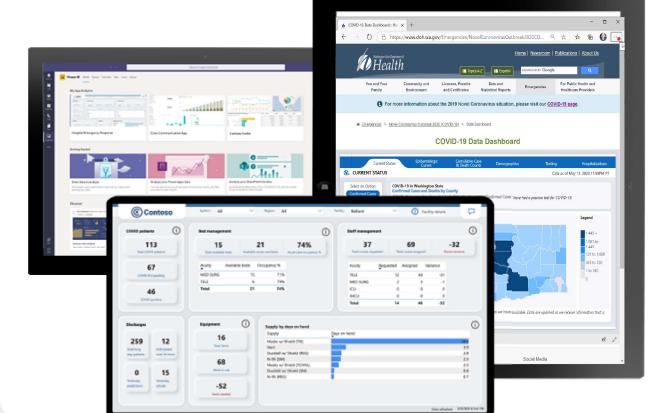
## All Data

Break down data silos and enrich your data with intelligence



## For Everyone

Create rich-interactive data experiences with AI infused insights



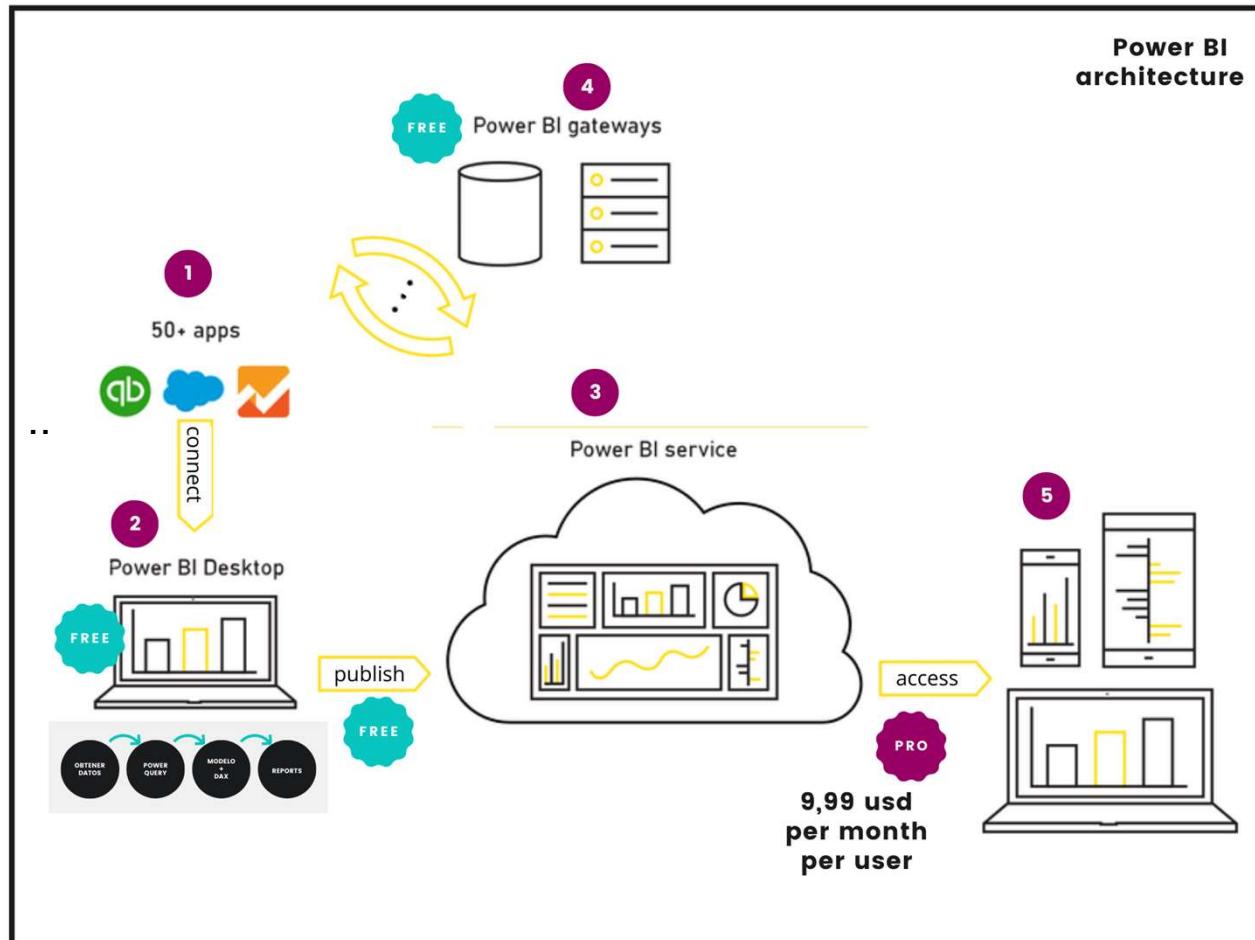
## For Every Decision

Integrate trusted and secure BI into the fabric of your organization and apps

Power BI drives a data culture for everyone and every decision



# Full Architecture





# Why using Power BI?

The screenshot displays a Microsoft Power BI Desktop interface. The main area shows a dashboard titled "Microsoft | Skateboard Store". The dashboard features a "What If..." section with a slider for "We Decrease Our Return Rate (%)" set to 25. It also includes a table titled "Net Sales vs 'What If' Analysis" and a line chart titled "What If" Analysis Forecast. The right side of the screen shows the "Visualizations" and "Fields" panes, which contain various data fields and filters. The ribbon at the top includes tabs for File, Home, Insert, Modeling, View, Help, External Tools, and Table tools.

Microsoft | Skateboard Store

Last Refresh: Jun 30th, 2019 / Chicago, IL, USA

What If...

We Decrease Our Return Rate (%) to: 25

Net Sales (Forecast) \$30,772 Extra Profit \$0 0.0% Profit Increase

Net Sales vs "What If" Analysis

Store	Net Sales	Forecast	Extra Profit
Abbas	\$2,646	\$2,646	\$0
Aliqui	\$3,724	\$3,724	\$0
Barba	\$1,568	\$1,568	\$0
Contoso	\$2,842	\$2,842	\$0
Fama	\$3,626	\$3,626	\$0
Leo	\$1,764	\$1,764	\$0
Natura	\$1,274	\$1,274	\$0
Palma	\$0	\$0	\$0
Pirum	\$2,744	\$2,744	\$0
Pomum	\$1,372	\$1,372	\$0
Quibus	\$2,940	\$2,940	\$0
Salvus	\$2,646	\$2,646	\$0
VanArdel	\$2,548	\$2,548	\$0
Victoria	\$1,078	\$1,078	\$0
Total	\$30,772	\$30,772	\$0

"What If" Analysis Forecast

Predicted Net Sales Forecast

Return Rate 24%

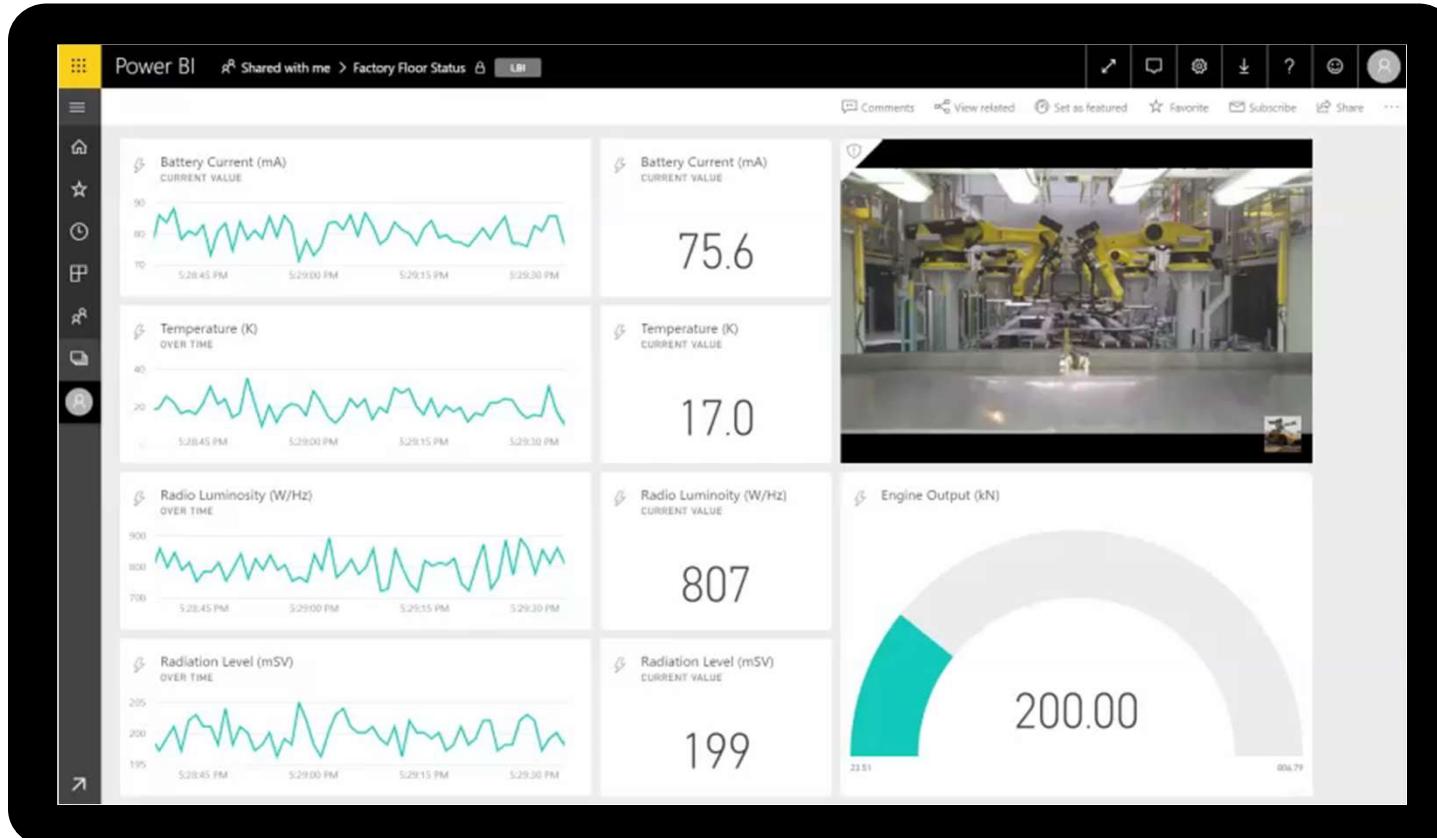
©Microsoft Corporation. All rights reserved.

Page 3 of 16

Power BI  
Desktop is free



## Why using Power BI?



Real time dashboards and interactive reports



# Why using Power BI?

The screenshot shows the Power BI desktop application interface. The top navigation bar includes File, Home, Insert, Modeling, View, Help, Format, Data / Drill, and various ribbon tabs like Home, Data, Queries, Insert, Calculations, and Share.

The main area displays three visualizations:

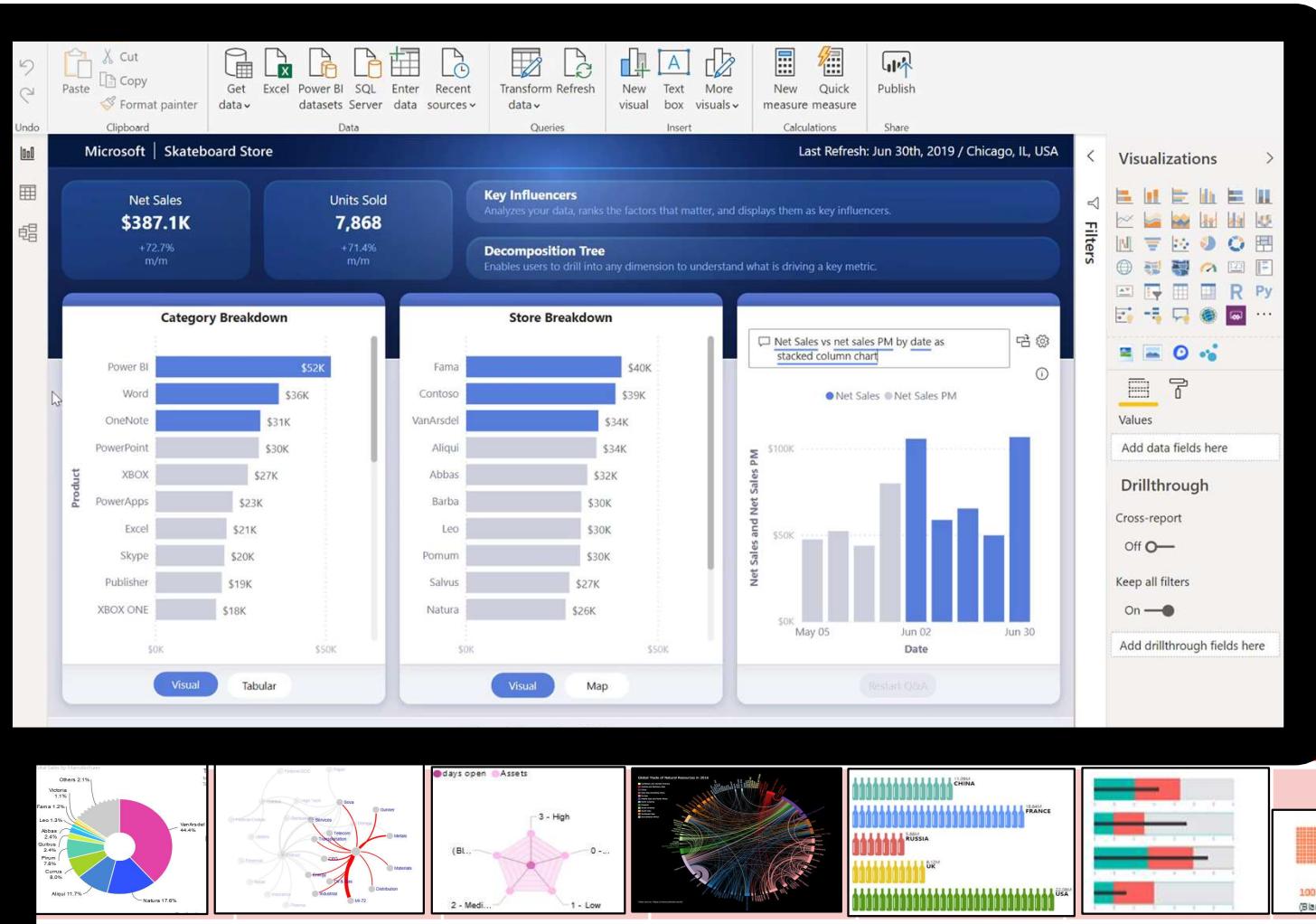
- Ask a question with Q&A**: A bar chart titled "show total global sales by genre". The Y-axis is "Genre" and the X-axis is "Global\_Sales". The data shows sales for various game genres.
- Number of games by publisher**: A donut chart showing the distribution of games by publisher. The segments represent different publishers with their counts and percentages.
- Sentiment by platform**: A bar chart showing sentiment scores for different platforms. The Y-axis is "Sentiment" and the X-axis is "Platform". The scores are: PS2 (0.63), DS (0.63), Wii (0.61), PS3 (0.56), and X360 (0.56).

A sidebar on the right lists "Visualizations" and "Filters" with various options.

## Natural Language query & AI insights



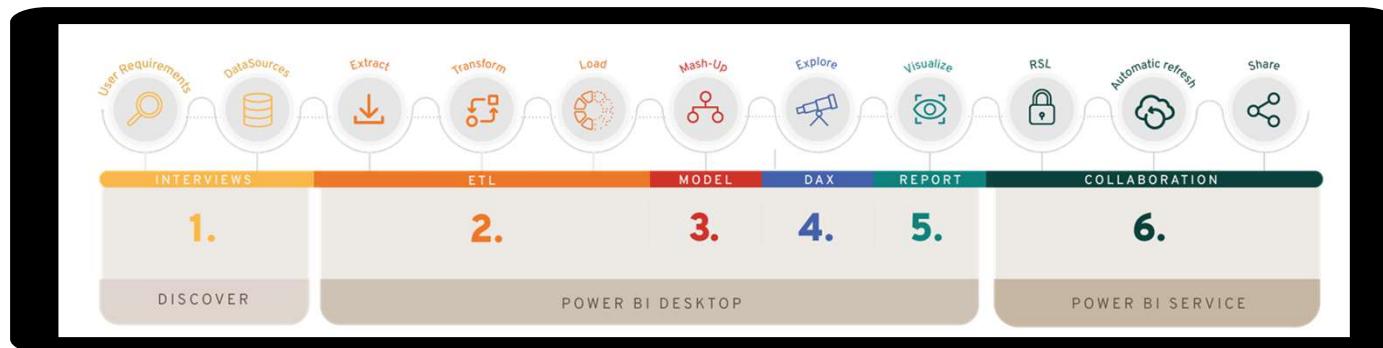
# Why using Power BI?



250+ Custom  
visuals Natural



# Why using Power BI?



All Business Intelligence process in one

The screenshot shows the Power BI desktop interface with the following components:

- Left pane:** A list of tables available for selection, including Suppliers, Customers, and Sales.
- Relationships pane:** Shows the current relationships between tables like Suppliers and Customers.
- Modeling ribbon:** Includes tabs for Archivo, Home, and Modeling. The Modeling tab is active, showing options for Relationships, Calculations, and Sort.
- Calculations pane:** Displays DAX code for calculating total sales:

```
% Ventas Producto =  
var SalesAmountTotal = sum(FactSales[SalesAmount])  
return  
FactSales[SalesAmount] / SalesAmountTotal
```
- Right pane:** A detailed view of the data model showing entities like Customer, Sales, and Product, along with their properties and relationships.



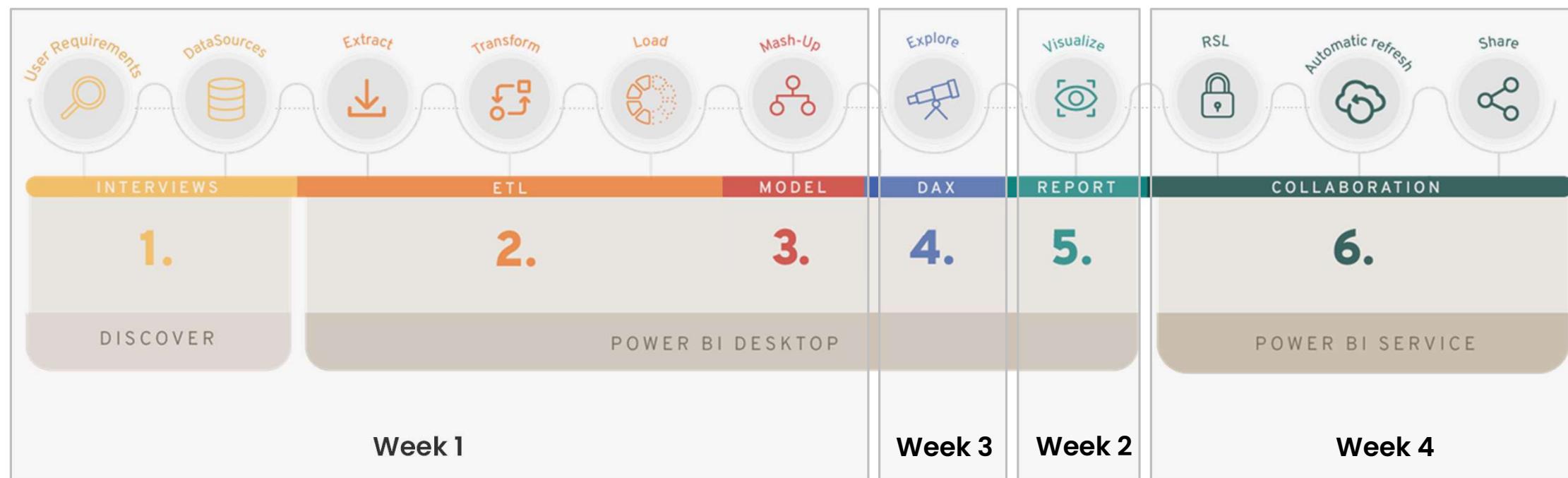
## Why using Power BI?



## Gartner Magic Quadrant for “Analytics and Business Intelligence Platforms”

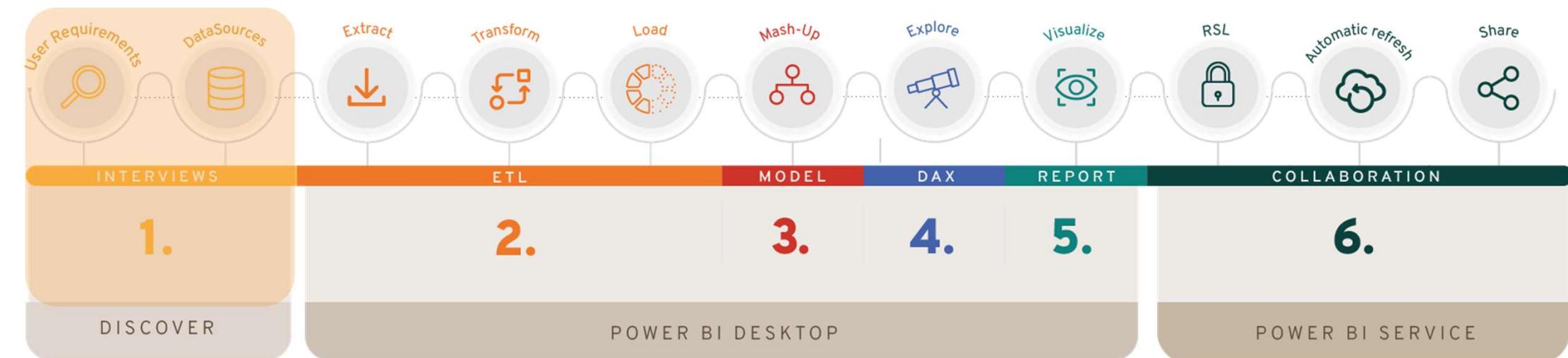


# Our Method: Flow of Report Development





# Our Method: Flow of Report Development





# Interviews

Understand processes, needs and data sources

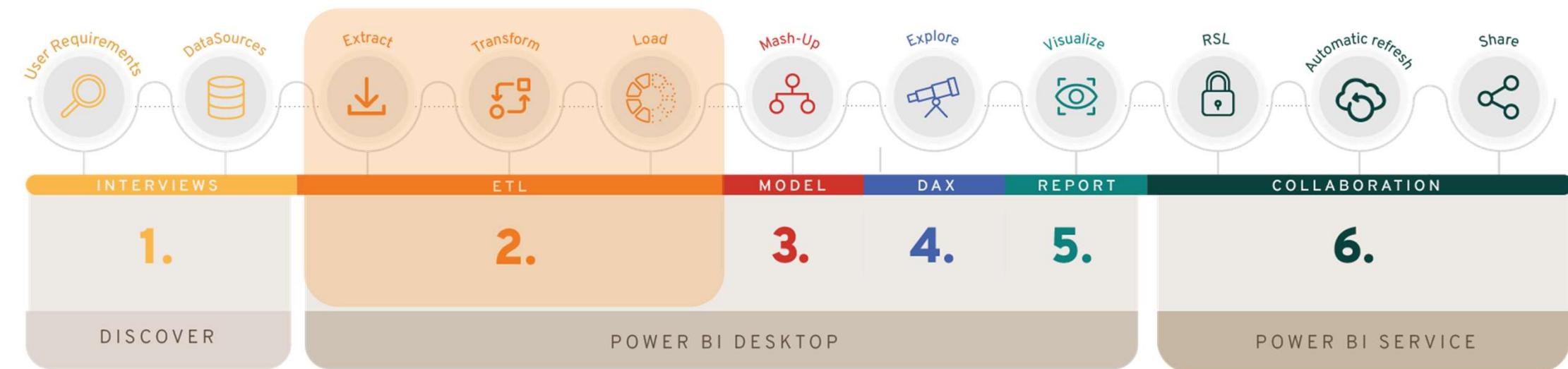


## DataSources

- 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 Informix database (Beta)
- IBM Netezza
- MySQL database



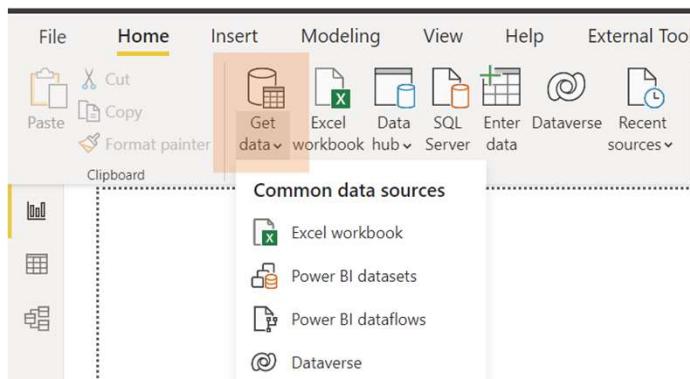
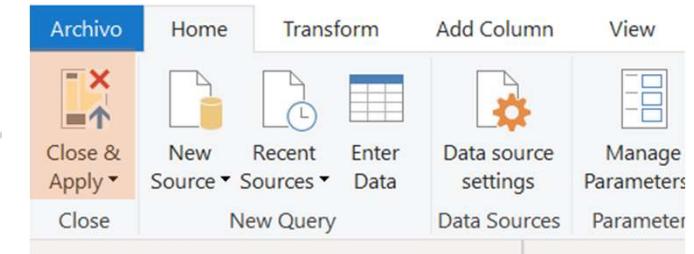
# Our Method: Flow of Report Development





# Power BI Desktop: Power Query

Clean and transform data sources



	SalesKey	DateKey	channelKey	StoreKey	ProductKey
1	3398163	04/07/2019 0:00:00	1	247	
2	3395082	04/10/2019 0:00:00	1	52	
3	3392414	15/05/2019 0:00:00	1	40	
4	3388423	27/06/2019 0:00:00	1	67	
5	3383822	20/10/2019 0:00:00	1	70	
6	3382789	12/07/2019 0:00:00	1	238	
7	3378463	26/10/2019 0:00:00	1	282	
8	3371111	23/05/2019 0:00:00	1	21	
9	3369889	29/04/2019 0:00:00	1	167	
10	3365947	25/05/2019 0:00:00	1	48	
11	3353916	29/04/2019 0:00:00	1	194	
12	3349673	30/06/2019 0:00:00	1	8	
13	3346418	02/05/2019 0:00:00	1	102	
14	3329410	29/10/2019 0:00:00	1	146	
15	3327975	04/06/2019 0:00:00	1	93	

INTERVIEWS

ETL

MODEL

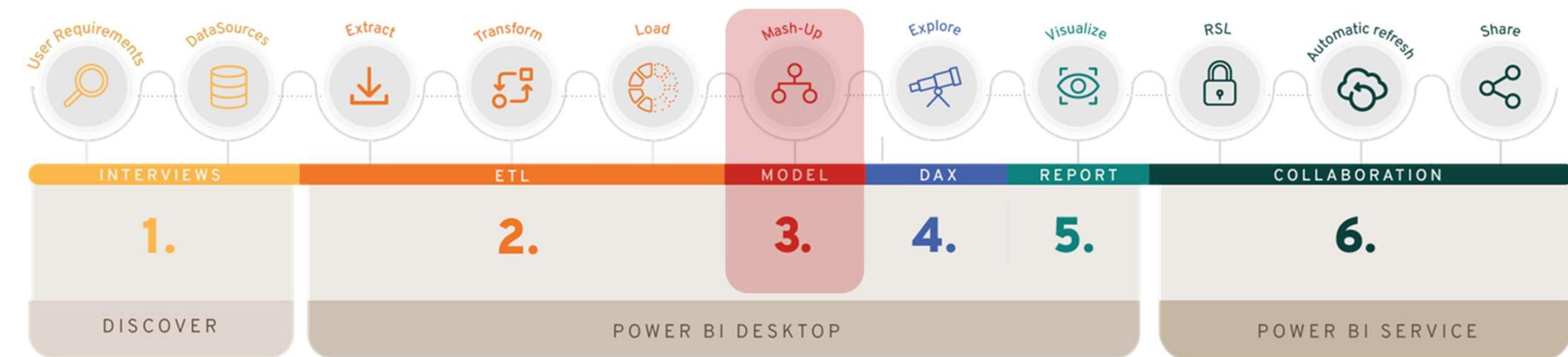
DAX

REPORT

COLLABORATION



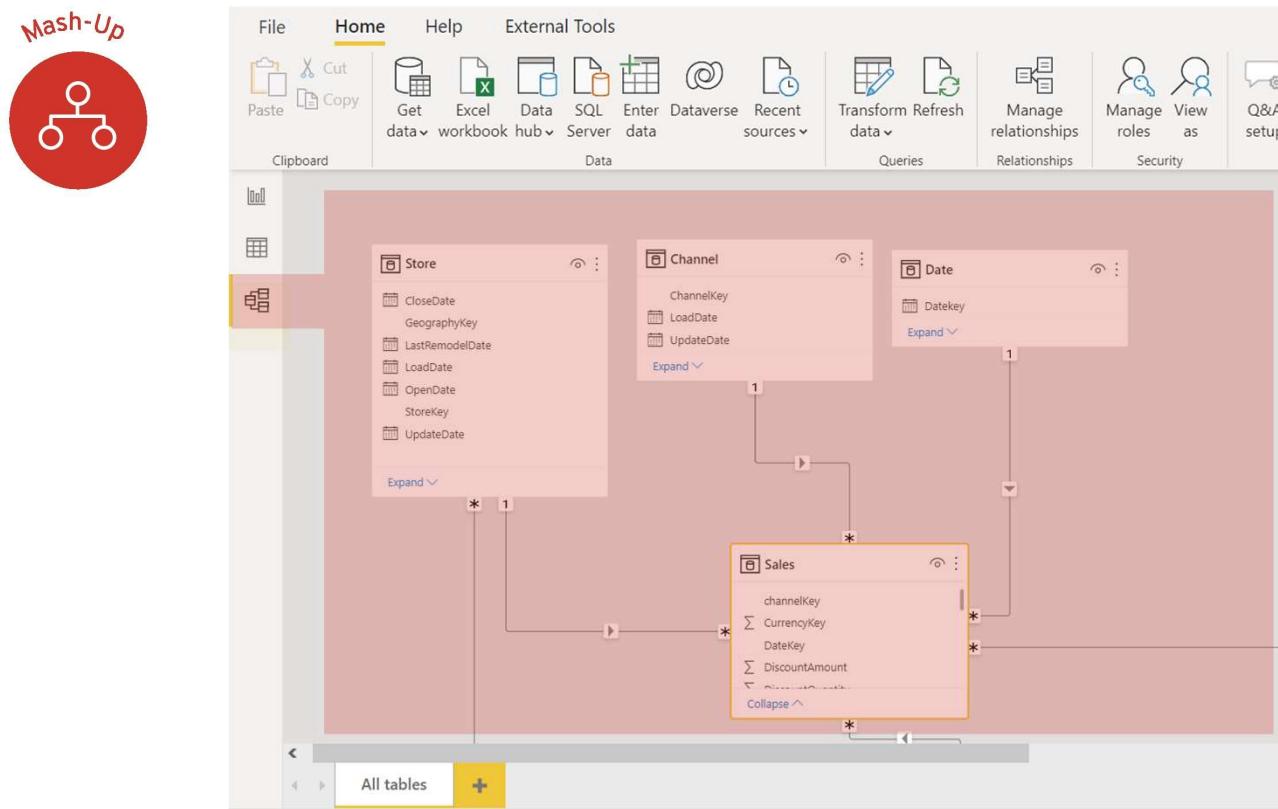
# Our Method: Flow of Report Development





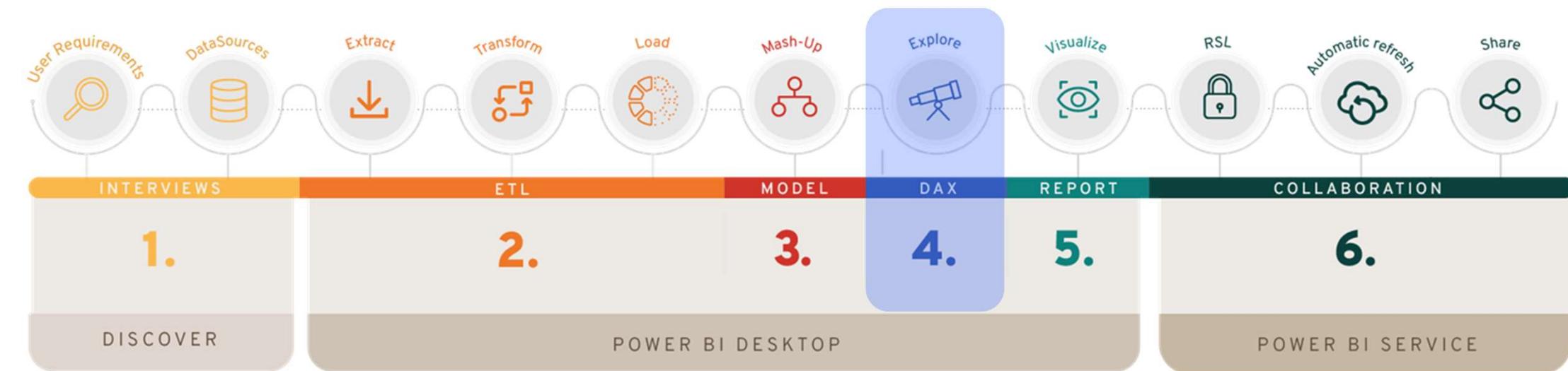
# Power BI Desktop: Model

Mash-Up tables to create a model that allows different analysis





# Our Method: Flow of Report Development





# Power BI Desktop: DAX

KPIs to answer business questions

Explore



The screenshot shows the Power BI Desktop interface. At the top, there's a ribbon with tabs: Structure, Formatting, Properties, Sort, Groups, Relationships, and Calculations. Below the ribbon, the 'Explore' tab is selected. On the left, there's a blue circular icon with a telescope symbol. The main area has a grid of data with columns: Quantity, ReturnAmount, DiscountQuantity, DiscountAmount, TotalCost, SalesAmount, ETLLoadID, LoadDate, UpdateDate, Year, and Kind of purchase. The 'Kind of purchase' column contains values like 'Retail' and 'Wholesale'. Above the grid, the DAX code is visible:

```
1 Kind of purchase =  
2 SWITCH(TRUE(),  
3 Sales[SalesQuantity] < 20, "Retail",  
4 "Wholesale")
```

INTERVIEWS

ETL

MODEL

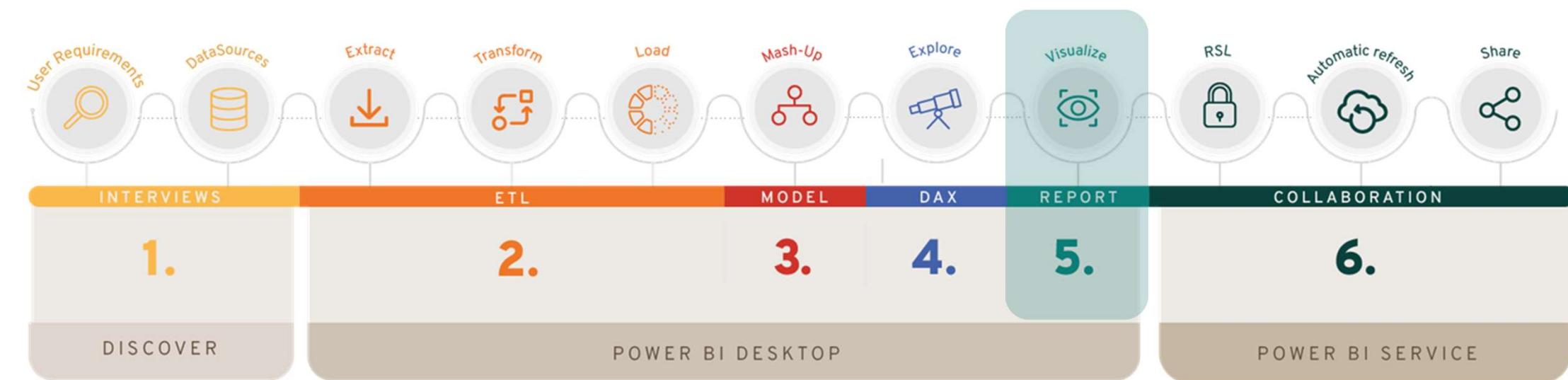
DAX

REPORT

COLLABORATION



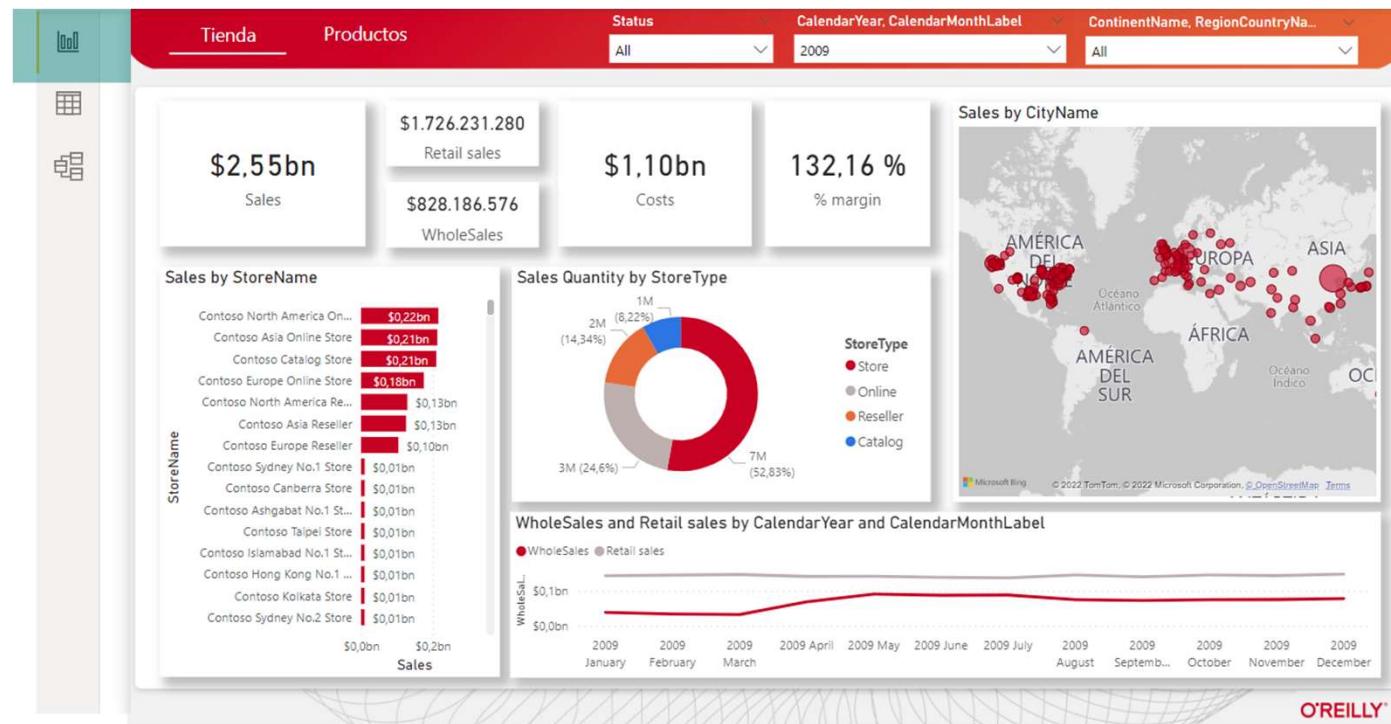
# Our Method: Flow of Report Development





# Power BI Desktop: Report

Create visuals for business users to understand their business



INTERVIEWS

ETL

MODEL

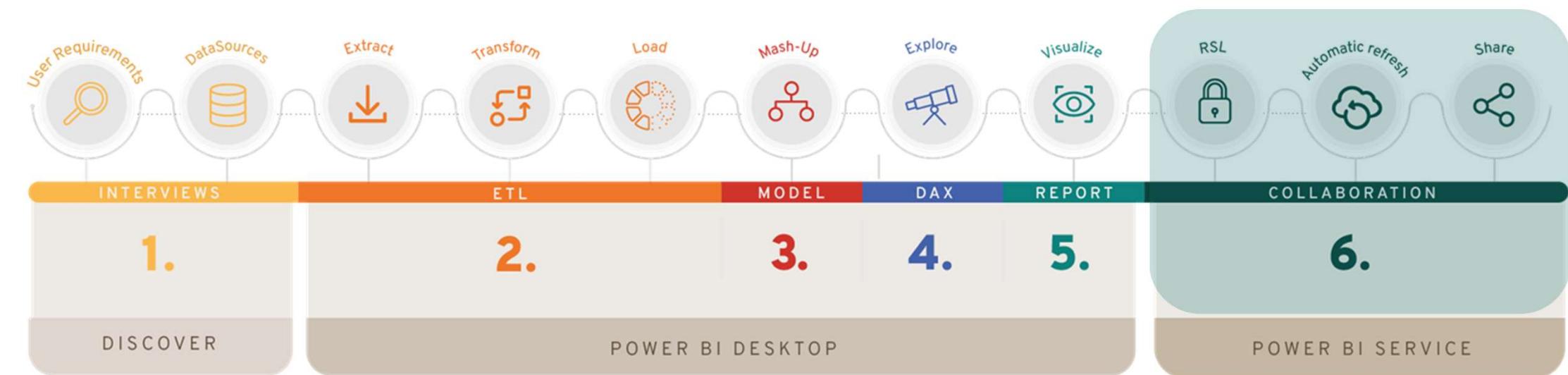
DAX

REPORT

COLLABORATION



# Our Method: Flow of Report Development

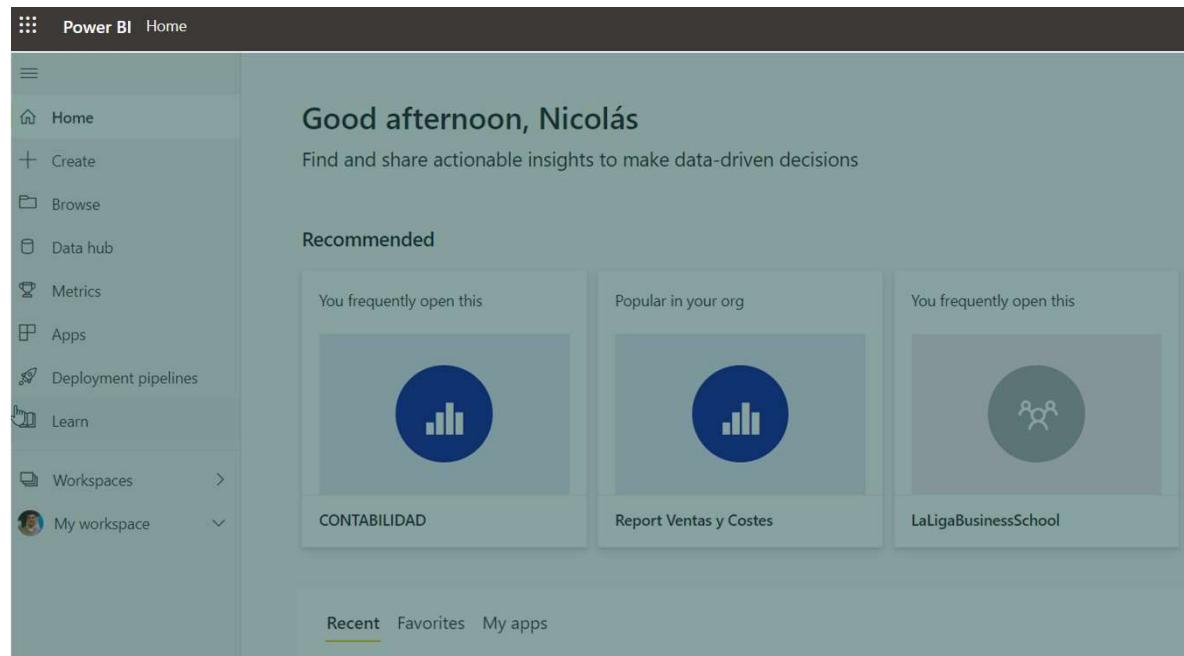




# Web Service Power BI

Share reports with colleagues

- RSL 
- Automatic refresh 
- Share 



The screenshot shows the Power BI Home page. At the top, it greets the user with "Good afternoon, Nicolás" and encourages finding and sharing actionable insights. Below this, a "Recommended" section displays three cards: "You frequently open this" (CONTABILIDAD), "Popular in your org" (Report Ventas y Costes), and "You frequently open this" (LaLigaBusinessSchool). The sidebar on the left includes links for Home, Create, Browse, Data hub, Metrics, Apps, Deployment pipelines, Learn, Workspaces (with "My workspace" selected), and Help.

INTERVIEWS

ETL

MODEL

DAX

REPORT

COLLABORATION



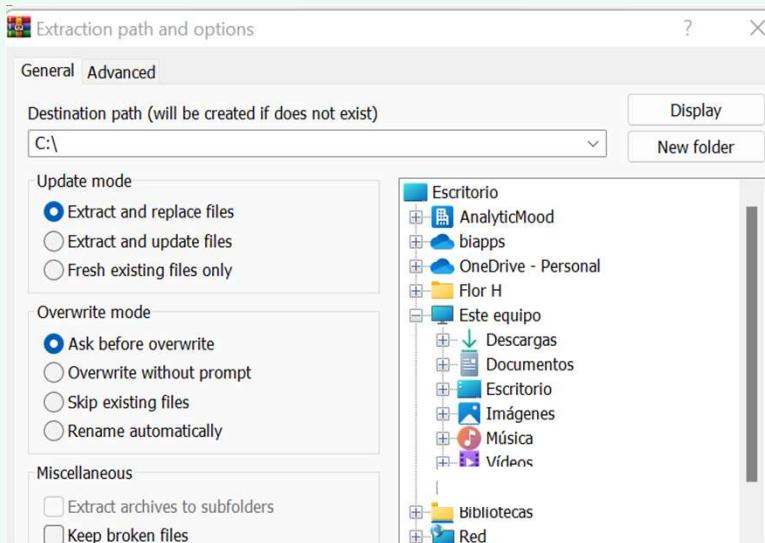
# Exercise 0.0



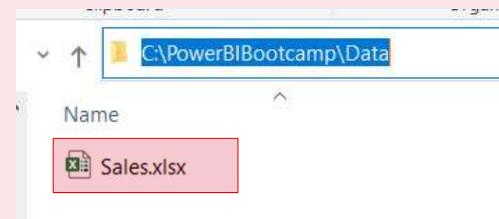
## Exercise 0.0 – Initial set up

1 First you must **download** the shared .ZIP and extract it to the Local C:\ path.

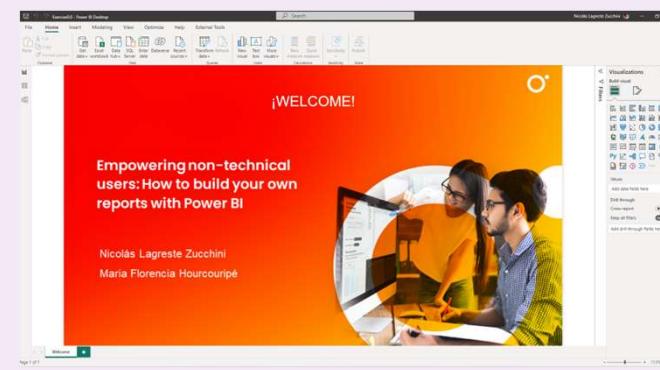
- Link 01:  
[bit.ly/PowerBIBootcamp202309](https://bit.ly/PowerBIBootcamp202309)
  
- Pass: **OREILLY2023**
  
- Link 02:  
[bit.ly/PowerBIBootcamp202309W01](https://bit.ly/PowerBIBootcamp202309W01)



2 Make sure the path to the **data sources (Sales.xlsx file)** is as follows: **C:\PowerBIBootcamp\Data**



3 Lets test Power BI Desktop by opening **Exercise0.0** that is located in  
**C:\PowerBIBootcamp\Week01\Exercises\Exercise0.0.pbix**



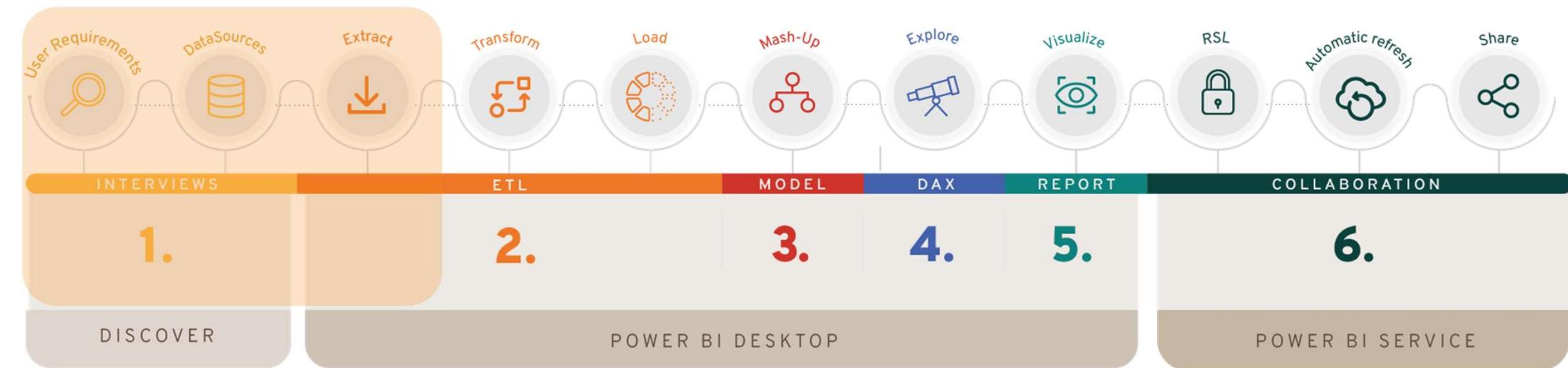


# Basic Transformation in Power Query

- Demo 1.1: Connect to 3 Excel Sheets and do basic transformation in Power Query (delete columns, filter rows, change names, add calculated columns, change data types).
- Load transform data to Power BI.
- Exercise 1.1: Basic Transformation in Power Query.
- Q&A



# Our Method: Flow of Report Development





# Demo 1.0





# Interviews

Understand processes, needs and data sources



## You are not the user

- Who is the audience?
- What do they need?

## How do you know your audience?

- |                |  |
|----------------|--|
| • Observing    | • Understand the business.                 |
| • Listening    | • What are our user's goals?               |
| • Interviewing | • What is the objective of this dashboard? |

Try to understand what they need, not what they want.



# Interviews

Understand processes, needs and data sources

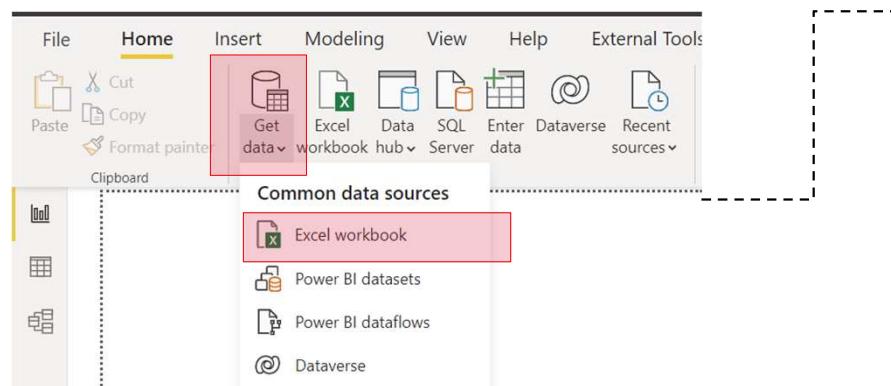


- Excel Workbook
  - Text/CSV
  - XML
  - JSON
  - Folder
  - PDF
  - Parquet
- SharePoint folder
- SQL Server database
- Access database
- SQL Server Analysis Services database
- Oracle database
- IBM Db2 database
- IBM Informix database (Beta)
- IBM Netezza
- MySQL database



# Power BI Desktop: Power Query: Extract

Extract



Navigator

DIMs.xlsx [8]

- CalendarTable
- ChannelTable
- ProductTable
- StoreTable

Calendar

Channel

Product

Store

Load Transform Data Cancel

INTERVIEWS

ETL

MODEL

DAX

REPORT

COLLABORATION



# Exercise 1.0



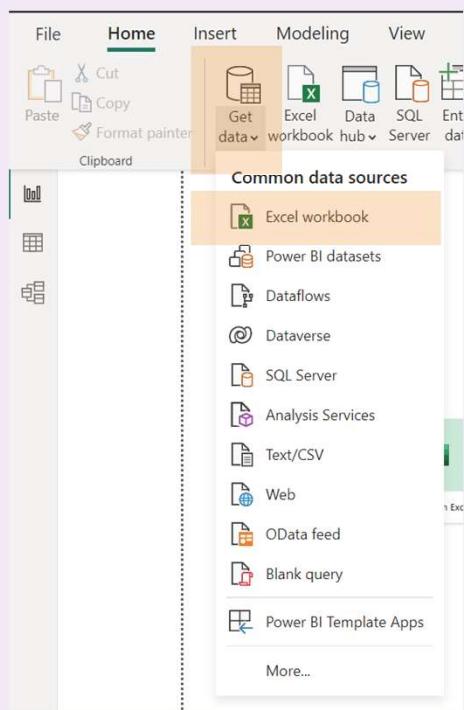
## Exercise 1.0.

Goal: Become familiar with the different data sources.

In order to do it, open a new **Power BI Desktop** and connect it to the following data sources. Just select *transform* option before *loading* them.

### 1 Source: Excel

File: **Sales.xlsx** (Location: "C:\PowerBIBootcamp\Data\Sales.xlsx")



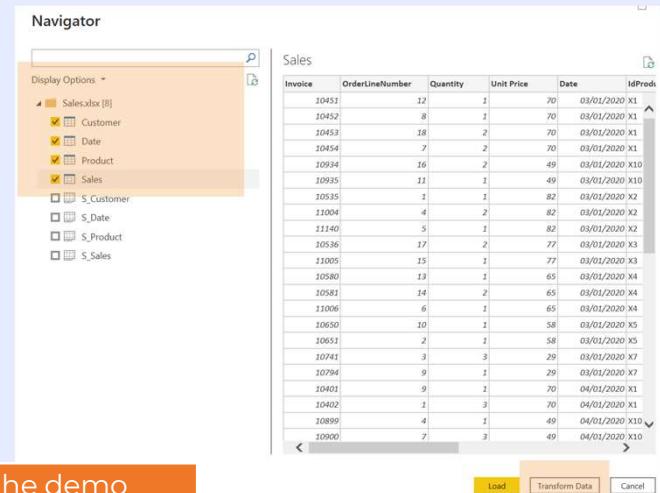
### 2 Tables:

a **Sales**

b **Customer**

c **Date**

d **Product**

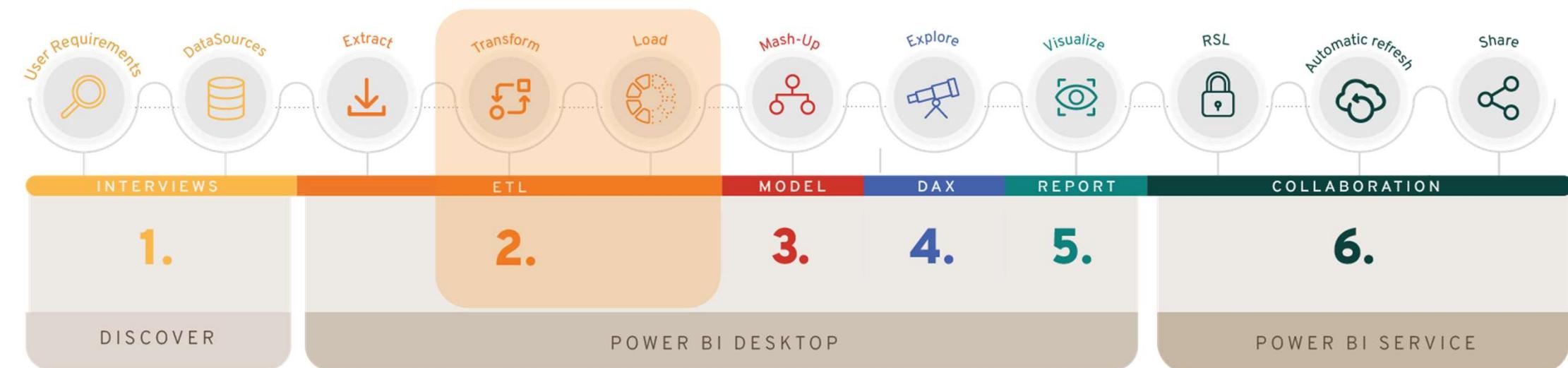


Invoice	OrderLineNumber	Quantity	Unit Price	Date	IdProd
10451	12	1	70	03/01/2020 X1	
10452	8	1	70	03/01/2020 X1	
10453	18	2	70	03/01/2020 X1	
10454	7	2	70	03/01/2020 X1	
10534	16	2	49	03/01/2020 X10	
10935	11	1	49	03/01/2020 X10	
10535	1	1	82	03/01/2020 X2	
11004	4	2	82	03/01/2020 X2	
11140	5	1	82	03/01/2020 X2	
10536	17	2	77	03/01/2020 X3	
11005	15	1	77	03/01/2020 X3	
10580	13	1	65	03/01/2020 X4	
10581	14	2	65	03/01/2020 X4	
11006	6	1	65	03/01/2020 X4	
10650	10	1	58	03/01/2020 X5	
10651	2	1	58	03/01/2020 X5	
10741	3	3	29	03/01/2020 X7	
10794	9	1	29	03/01/2020 X7	
10401	9	1	70	04/01/2020 X1	
10402	1	3	70	04/01/2020 X1	
10899	4	1	49	04/01/2020 X10	
10900	7	3	49	04/01/2020 X10	

\*Note: these data are different from those of the demo



# Our Method: Flow of Report Development





# Demo 1.1





# Presentation: Parts of Power query (report, data, relations)

The screenshot shows the Microsoft Power Query interface with several key components highlighted:

- Transform** button in the ribbon.
- Transform columns** button in the ribbon.
- Queries [10]** list on the left, with **Sales** selected.
- Add new column to table** button in the ribbon.
- Data** grid showing sales data with columns: SalesKey, DateKey, channelKey, StoreKey, ProductKey.
- Query Settings** pane on the right.
- Properties** section in the Query Settings pane.
- Applied Steps** section in the Query Settings pane, listing steps like Source, Filtered Hidden Files1, Invoke Custom Function1, Renamed Column, Removed Other, Expanded Table, Removed Columns, Changed Type, Extracted Text, Changed Type, Changed Type, and Filtered Rows.



# Use first Row as Headers

The screenshot shows the Power BI desktop interface with the 'Transform' tab selected. A red box highlights the 'Use First Row as Headers' button in the ribbon, which is being pointed to by a large black arrow. Below the ribbon, the 'Queries [10]' pane shows a list of queries, and the main area displays a table with four columns: Column1, Column2, Column3, and Column4. The first row contains the column names: ProductKey, ProductLabel, ProductName, and ProductDescription. The following rows contain data for various products, such as Mouse E50 Grey and Advanced 2.4 GH. A callout box with a red border points to the first row of data with the text: 'If column's names appear in the first row, you should use this transformation'.

ProductKey	ProductLabel	ProductName	ProductDescription
873		Mouse E50 Grey	Advanced 2.4 GH
874		Mouse Optical Mouse X205 Black	Wireless notebook
875		Mouse Optical Mouse X205 White	Wireless notebook
880	0308138	Contoso Optical Wheel OEM PS/2 Mouse E60 Black	PS/2 mouse, 6 fe
884	0308142	Contoso Bluetooth Notebook Mouse X305 White	Transceiver-free
886	0308144	Contoso Bluetooth Notebook Mouse X305 Grey	Transceiver-free
889	0308147	Contoso Bluetooth Notebook Mouse E70 Silver	Transceiver-free
891	0308149	SV Rechargeable Bluetooth Notebook Mouse E80 Silver	33 ft range, recha
893	0308151	SV Rechargeable Bluetooth Notebook Mouse E80 White	33 ft range, recha



# Change names

= Table.SelectRows(#"Changed Type", each ([StoreKey] <> null))

		ContinentName	CityName	StateProvinceName	RegionCountryName
1	17/06/2009	North America	Renton	Washington	United States
2	16/06/2009	North America	Albany	New York	United States
3	17/06/2009	North America			United States
4	17/06/2009	North America			United States
5	18/06/2009	North America			United States
6	16/06/2009	North America			United States
7	17/06/2009	North America	Greeley	Colorado	United States
8	16/06/2009	North America	Greeley	Colorado	United States
9	18/06/2009	North America	Appleton	Wisconsin	United States
10	17/06/2009	North America	Lafayette	Colorado	United States
11	18/06/2009	North America	Arlington	Texas	United States

Doble click in the column name if you want to change the name.

Queries [10]

- ▷ Transform File from Facts [2]
- ▷ Other Queries [6]
  - Channel
  - Calendar
  - Product
  - Sales
  - Store
  - Measure

Doble click in the query name if you want to change the name.



# Multiply Columns

The screenshot shows the Power BI Data Editor interface. On the left, the 'Queries [10]' pane is open, displaying 'Transform File from Facts [2]', 'Other Queries [6]', and several tables: Channel, Calendar, Product, Sales, Store, and Measure. The 'Sales' query is currently selected. In the center, a table view shows columns for Cost, UnitPrice, and SalesQuantity. A context menu is open over the first two columns, listing options: Multiply, Subtract, Divide, Divide (Integer), Modulo, Percentage, and Percent Of. An arrow points from the text in the bottom right towards this menu. A red box highlights the text: 'If you need to multiply two columns, select them using "Ctrl" key, and then "Multiply" button'. The top ribbon bar includes tabs for Archivo, Home, Transform, Add Column, View, Tools, Help, and various icons for column operations like Conditional Column, Merge Columns, Statistics, and Standard.

If you need to multiply two columns, select them using "Ctrl" key, and then "Multiply" button



## Filter Rows

If you filter one value in one specific column, that full row will be deleted from the query.

Note: this does not affect the data source



# Split Column

You can separate columns delimited by comma or other separator using this option.

The screenshot shows the Power BI ribbon with the 'Transform' tab selected. In the 'Text' section of the ribbon, the 'Split Column' button is highlighted with a red box and a black arrow pointing from the explanatory text above to it. Below the ribbon is a table preview showing columns 'IdProduct' and 'Category'. The table contains 13 rows of product information, such as 'Real Madrid youth third jersey by Adidas' and 'Barcelona I96 Jacket by Nike'.

## Split Column by Delimiter

Specify the delimiter used to split the text column.

Select or enter delimiter

--Custom--  
,

Split at

- Left-most delimiter  
 Right-most delimiter  
 Each occurrence of the delimiter

Advanced options

Quote Character

"

Split using special characters

Insert special character ▾

OK

Cancel



# Remove Other Columns

The screenshot shows the Power BI Data Editor interface. A tooltip for the 'Remove Other Columns' option is displayed, stating: 'Remove all columns from this table, except the currently selected ones.' A red arrow points from this tooltip to a callout box below it.

	Invoice	Quantity	Unit Price	Order Date	Customer	Product
1	10451	1	70	03/01/2020	X1	
2	10452	1	70	03/01/2020	X1	
3	10453	2	70	03/01/2020	X1	
4	10454	2	70	03/01/2020	X1	
5	10934	2	49	03/01/2020	X10	
6	10935	2	49	03/01/2020	X10	
7	10535	7	82	03/01/2020	X2	

It is a good practice to remove all unnecessary columns.

This option leaves only the necessary columns of the model.



# Extract

The screenshot shows the Power BI Data Editor interface. A red box highlights the 'ABC Extract' button in the ribbon's 'Text' tab. A tooltip for 'Text Before Delimiter' is displayed, stating: 'Return the text that occurs before a delimiter.' Below the ribbon, a table has a row selected with a red box. The formula bar shows: `= Source{[Item="DateTable", Kind="Table"]}`. The table has columns 'Date' and 'Year'. The 'Date' column contains dates like '01.01.2020-' and '02.01.2020-'. The 'Year' column contains '2020'.

Characters can be extracted from the different values in order to clean the data.

This screenshot is similar to the one above, showing the Data Editor with the 'ABC Extract' button highlighted. A tooltip for 'First Characters' is displayed, stating: 'Return a specified number of characters from the start of each value in this column.' The table structure and data are identical to the first screenshot.

Extract First Characters

Enter how many starting characters to keep.

Count

It is also possible to extract the first characters only.



# Change Data Types

Customer	City
1.2 Decimal Number	Frar
\$ Fixed decimal number	Aus
123 Whole Number	Frar
% Percentage	Aus
Date/Time	Waverly
Date	Melbourne
Time	Brisbane
Date/Time/Timezone	USA
Duration	UK
Text	Nor
True/False	Gern
Binary	Frar
Using Locale...	Spa
	Can
	louver

It is important that each column has the correct data type for better compression.

Total Sales	Total Cost
1.2 Decimal Number	63
\$ Fixed decimal number	63
123 Whole Number	126
% Percentage	126
Date/Time	70
Date	65
Time	52
Date/Time/Timezone	104
Duration	52
Text	130
True/False	65
Binary	52
Using Locale...	52

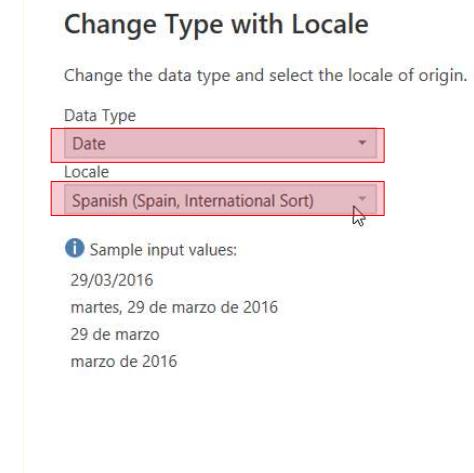
It is also important for correct calculations that the numerical columns are well categorized.



# Change Data Types (cont.)

	A B C	Date	A B C	Year
1		1.2 Decimal Number		
2		\$ Fixed decimal number		
3		123 Whole Number		
4		% Percentage		
5		Date/Time		
6		Date		
7		Time		
8		Date/Time/Timezone		
9		Duration		
10	A B C	Text		
11		True/False		
12		Binary		
13		Using Locale...		
14		15.01.2020		
15		16.01.2020		
16		17.01.2020		
17		18.01.2020		

Sometimes it is necessary to change using the locale of the source file.



You must select the data type and its locale



# Detect Data Type

The screenshot shows the Power BI desktop interface with the 'Transform' tab selected. In the ribbon, the 'Data Type' dropdown is set to 'Any'. A red box highlights the 'Detect Data Type' button, which is being clicked by a mouse cursor. A tooltip window titled 'Detect Data Type' appears, stating: 'Automatically detect the data type of the currently selected columns.' Below the ribbon, a table is displayed with four rows of data. The columns are labeled 'Year', 'Month Num', and 'Month'. The data shows dates from January 1, 2020, to January 4, 2020. To the right of the table, a red box contains a tip: 'Tip: Ctrl + A to select all columns at the same time'.

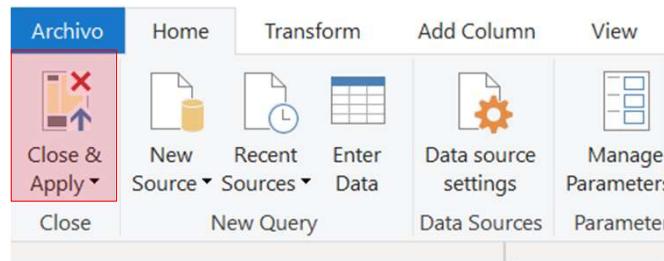
To detect all types of data automatically it is very useful to use this functionality.

Select the required columns and click Detect Data Type in the Transform tab.



## Load data to model

- Once all changes have been made, close and apply the changes to view them and generate a data model.





# Exercise 1.1



## Exercise 1.1. Open the file **Exercise1.1.pbix** (Location: C:\PowerBIBootcamp\Week01\Exercises\Exercise1.1.pbix) and then open Power Query

1

In Product table separate the product from the category. Rename columns.

2

In Sales table:

a Filter out null records.

b Remove the OrderLineNumber column.

3

In Date table:

a Extract first 10 characters from date column.

b Change Type with Date Locale.

4

Check the data types for each column.

Tip: Press Ctrl + A and use Detect Data Type in Transform tab.

5

Load the tables into the data model.

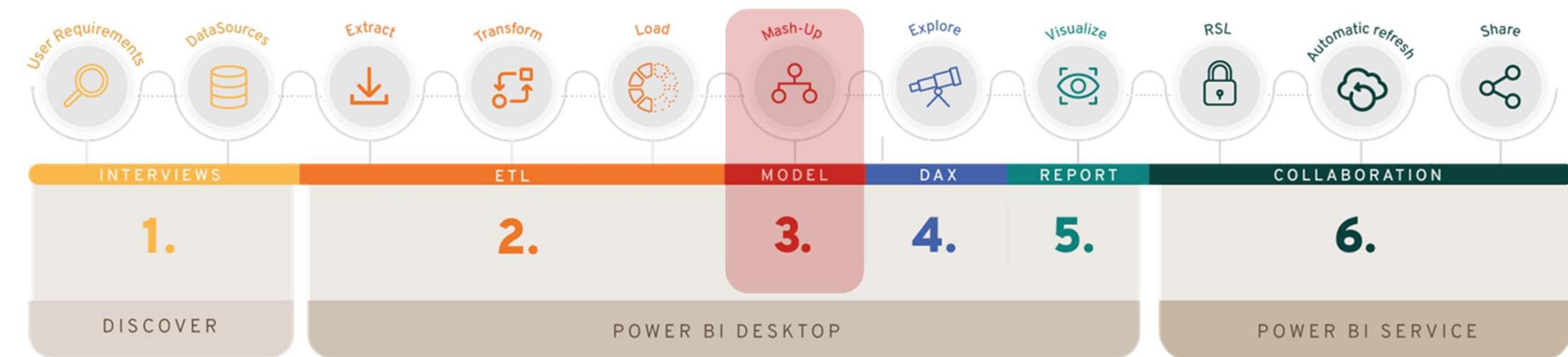


# Why building star schema models in Power BI

- What is a Star Schema Model?
- Facts and dimensions of star schema models.
- What do we mean with relationships between tables.



# Our Method: Flow of Report Development





What type of models are better suited for data analysis?



INTERVIEWS

ETL

MODEL

DAX

REPORT

COLLABORATION



# Denormalized models in Excel

	ProductID	Product	Date	CustomerID	Email	Last Name	First Name	Full Name	CampaignID	Units	1.2	CatSegID
1	676	Maximus UC-41	9/25/2011	70283	Farrah.Kent@xyz.com	Kent	Farrah	Farrah Kent	22	1	10	
2	585	Maximus UC-50	3/24/2014	70283	Farrah.Kent@xyz.com	Kent	Farrah	Farrah Kent	15	1	10	
3	585	Maximus UC-50	11/30/2014	138334	Martha.McClain@xyz..	McClain	Martha	Martha McClain	8	1	10	
4	585	Maximus UC-50	6/21/2015	27193	Hedda.McIntosh@xyz..	McIntosh	Hedda	Hedda McIntosh	22	1	10	
5	585	Maximus UC-50	1/6/2013	238970	Lunesa.Walker@xyz.com	Walker	Lunesa	Lunesa Walker	21	1	10	
6	585	Maximus UC-50	3/22/2013	182241	Upton.Page@xyz.com	Page	Upton	Upton Page	17	1	10	
7	449	Maximus OM-54	9/25/2011	195385	Drake.Wells@xyz..	Wells	Drake	Drake Wells	22	1	4	
8	449	Maximus OM-54	9/30/2014	168009	Wallace.Bender@xyz..	Bender	Wallace	Wallace Bender	17	1	4	
9	449	Maximus OM-54	8/12/2014	110391	Astra.Erickson@xyz..	Erickson	Astra	Astra Erickson	20	1	4	
10	449	Maximus OM-54	4/16/2014	49327	Echo.Bradley@xyz.com	Bradley	Echo	Echo Bradley	7	1	4	
11	449	Maximus OM-54	2/28/2013	65982	Yoko.Gross@xyz.com	Gross	Yoko	Yoko Gross	17	1	4	
12	449	Maximus OM-54	6/6/2013	97	Yoshi.Grant@xyz.com	Grant	Yoshi	Yoshi Grant	10	1	4	
13	449	Maximus OM-54	5/14/2013	56757	Brian.Carrillo@xyz..	Carrillo	Brian	Brian Carrillo	10	1	4	
14	449	Maximus OM-54	4/9/2015	248715	Mark.Hewitt@xyz.com	Hewitt	Mark	Mark Hewitt	19	1	4	
15	449	Maximus OM-54	4/28/2013	248715	Mark.Hewitt@xyz.com	Hewitt	Mark	Mark Hewitt	8	1	4	
16	449	Maximus OM-54	3/28/2014	240931	Oscar.Avila@xyz.com	Avila	Oscar	Oscar Avila	18	1	4	
17	449	Maximus OM-54	2/26/2014	201094	Duncan.Mcintosh@xyz..	Mcintosh	Duncan	Duncan McIntosh	19	1	4	
18	615	Maximus UC-80	5/14/2012	212445	Jacob.Santiago@xyz..	Santiago	Jacob	Jacob Santiago	22	1	10	
19	615	Maximus UC-80	5/14/2012	70666	Hilary.Coller@xyz..	Collier	Hilary	Hilary Collier	22	1	10	
20	615	Maximus UC-80	5/14/2012	114459	Chester.Mitchell@xyz..	Mitchell	Chester	Chester Mitchell	22	1	10	
21	615	Maximus UC-80	5/14/2012	221670	Sage.Yang@xyz.com	Yang	Sage	Sage Yang	22	1	10	
22	615	Maximus UC-80	6/3/2012	168009	Wallace.Bender@xyz..	Bender	Wallace	Wallace Bender	22	1	10	
23	615	Maximus UC-80	6/3/2012	154439	Illiana.Dunlap@xyz.c..	Dunlap	Illiana	Illiana Dunlap	22	1	10	
24	615	Maximus UC-80	6/4/2012	911391	Joelle.Lee@xyz.com	Lee	Joelle	Joelle Lee	22	1	10	

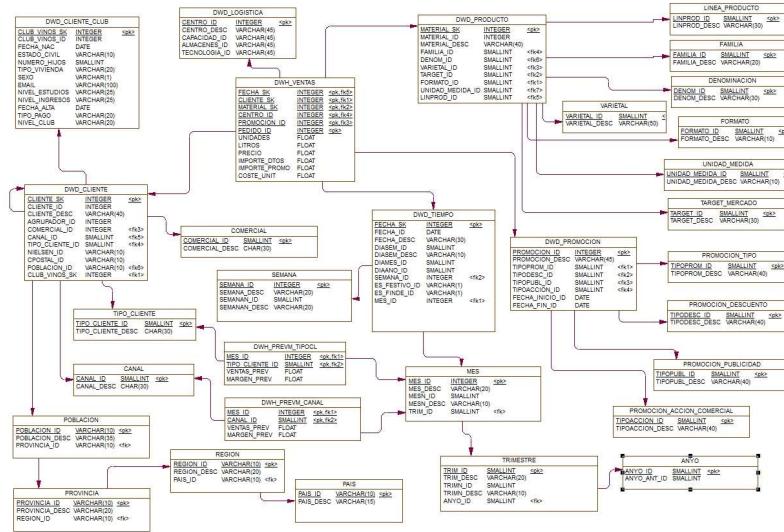
- **Data Duplication:** Excel often leads to data duplication as different sheets or workbooks may contain similar data, increasing the risk of inconsistencies and errors.

- **Performance Issues:** As the data grows, Excel's performance can degrade, leading to slow calculations and analysis, especially when using complex formulas or functions.

- **Lack of Data Integrity:** Excel lacks strict data integrity controls, making it easier for users to input incorrect or inconsistent data, impacting the accuracy of analysis and reporting.



# Normalized models in Data Bases



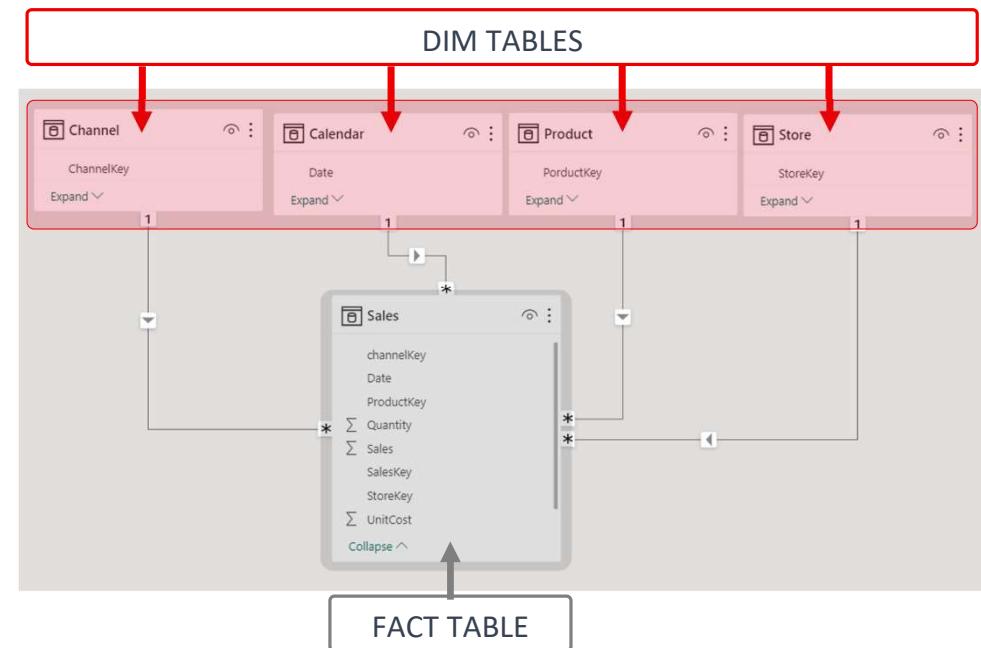
- Readability:** The normalized structure may result in cryptic column names and lack of context, making it harder for non-technical users to understand the data and generate reports.
- Complex Queries:** Normalized models require complex queries with multiple joins to retrieve meaningful data for reporting, leading to slower performance and increased query complexity.
- Performance Overhead:** The need for multiple joins and complex data retrieval can cause performance overhead, leading to slower reporting and longer processing times.



# Star Schema Overview

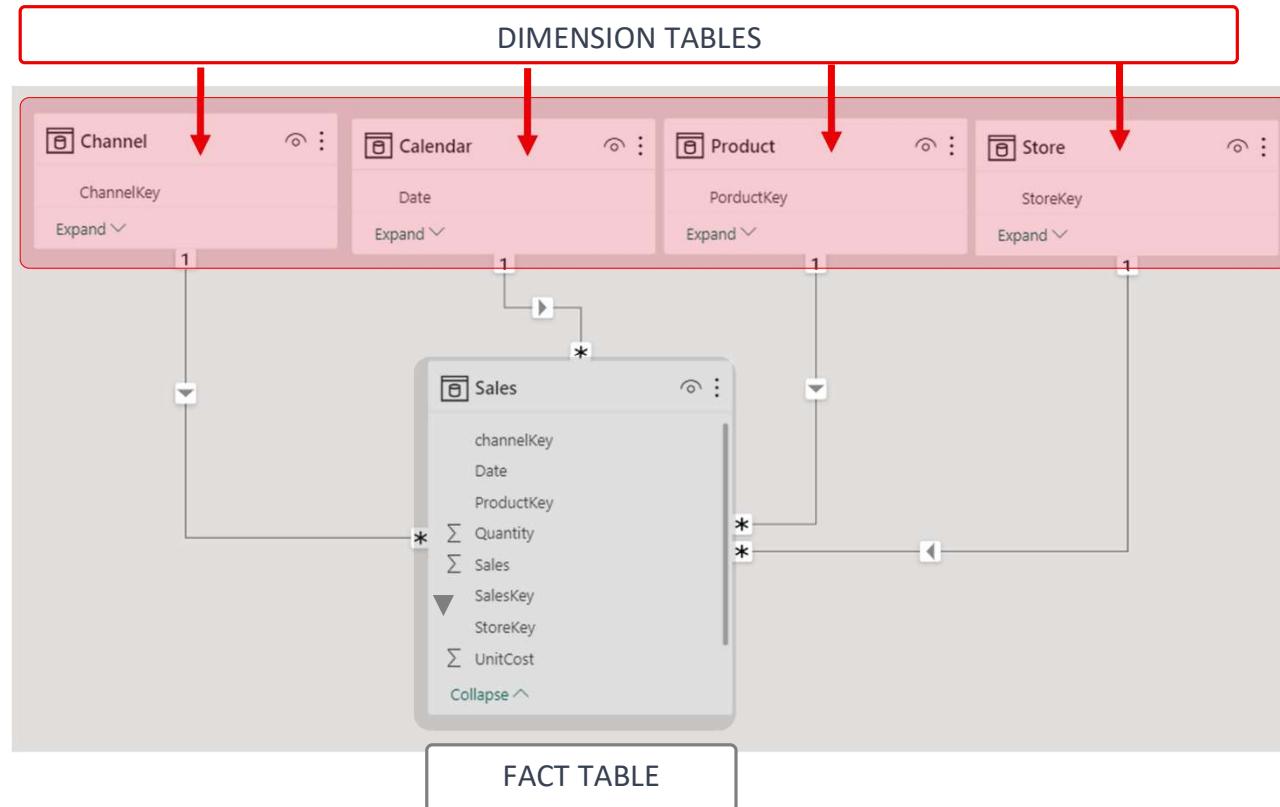
A star schema is a type of data modeling approach used in data warehouses and business intelligence systems.

It is designed for efficient data analysis and reporting.





# Dim & Facts



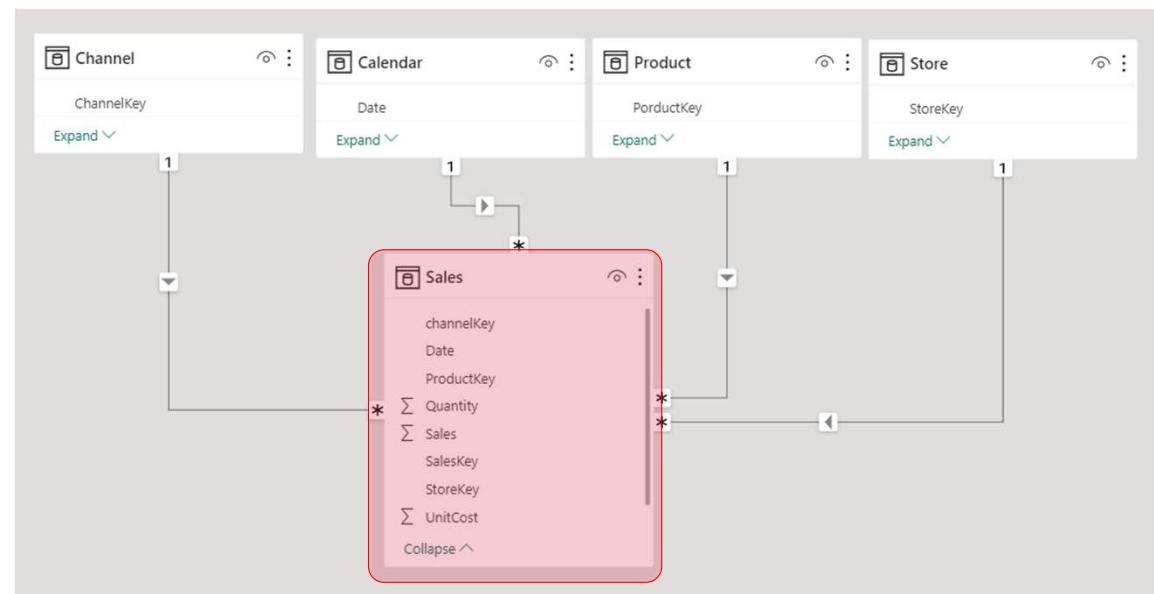


## Facts

A **fact table** represents individual transactions or events.

Each row contains numerical data to be aggregated.

Examples include sales, units sold, and costs.

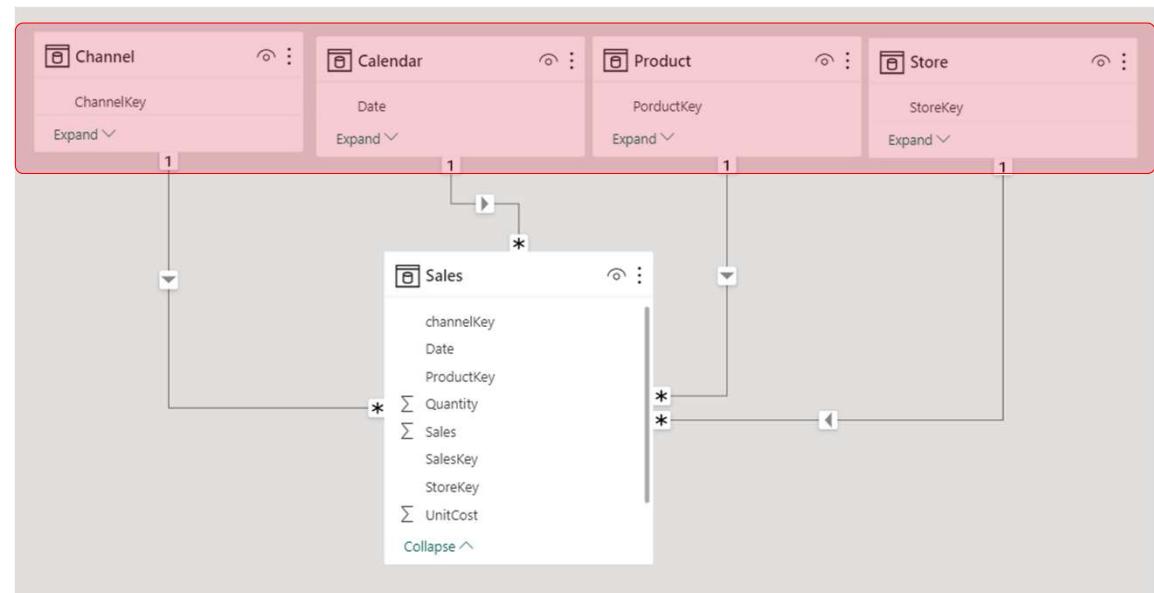




# Dimensions

A **Dim** (or Dimension) table contains descriptive attributes that define how a fact should roll up.

- Sales by customer
- Quantity by Product





## Facts & Dim

The **fact table** stores two types of information:

- Numeric values
- Dimensions ID

VALUE DATA			DIMENSION ID		
ORDER_ID	SALES_AMOUNT	QUANTITY	PRODUCT_ID	EMPLOYEE_ID	DATE_ID
101	150	2	22	12225	20230501
102	210	3	17	12588	20230505
103	250	1	12	14558	20230509
104	350	5	16	13225	20230512
105	70	1	16	16442	20230515

The **dim table** stores two types of information:

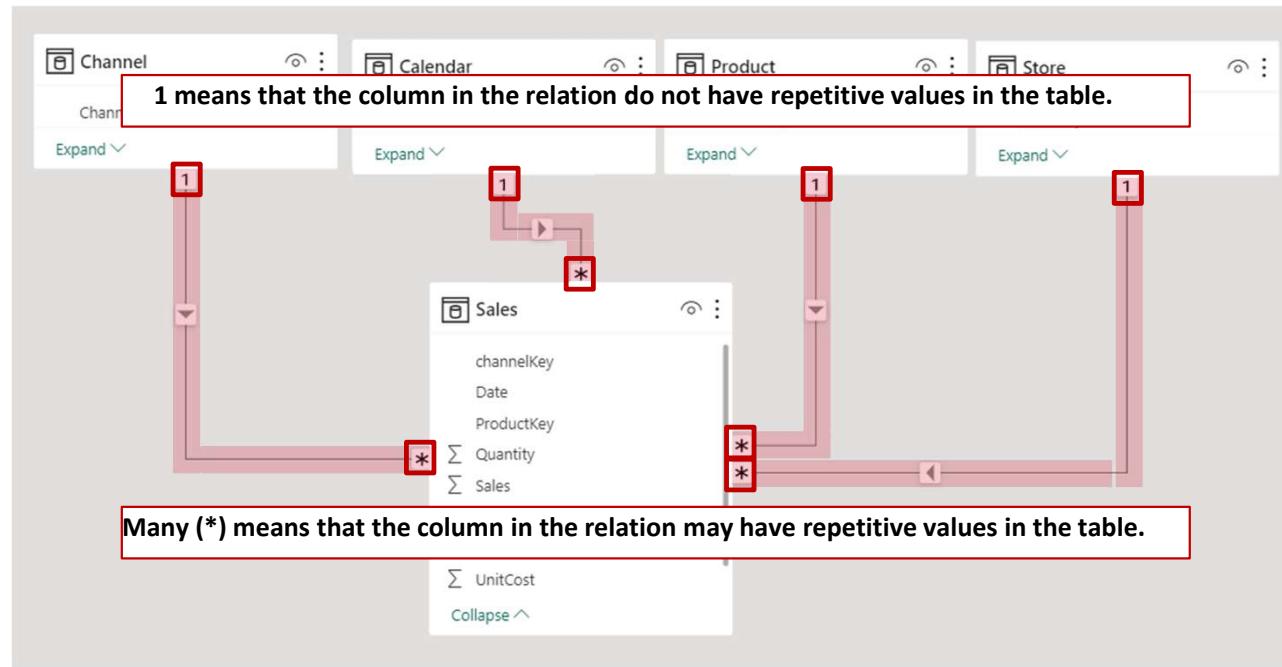
- Dimensions ID
- Descriptive columns

DIMENSION ID		DESCRIPTIVE COLUMNS			
PRODUCT_ID	PRODUCT_NAME	MANUFACTURER	RANGE	TYPE	
1	Apple Watch Series 7	Apple Inc.	Very Light	Smartwatch	
2	Lenovo IdeaCentre AIO 520	Lenovo	Light	All-in-One PC	
3	LG UltraGear 27GL850	LG	Very Light	Gaming Monitor	
4	ASUS RT-AX82U	ASUS	Medium	Wi-Fi Router	
5	Dell XPS 13	Dell	Light	Laptop	
6	TP-Link Kasa Smart Plug	TP-Link	Light	Smart Plug	
7	Sony PlayStation 5	Sony	Light	Gaming Console	
8	Epson Perfection V600	Epson	Very Light	Flatbed Scanner	
9	iPad Pro 12.9-inch	Apple Inc.	Medium	Tablet	
10	Philips Hue White & Color	Signify	Light	Smart Bulb	
11	Samsung Galaxy S22	Samsung	Very Light	Smartphone	



# Cardinality

A relationship in Power BI establishes a connection between two tables (typically a fact table and a dimension table) using a column from each table.

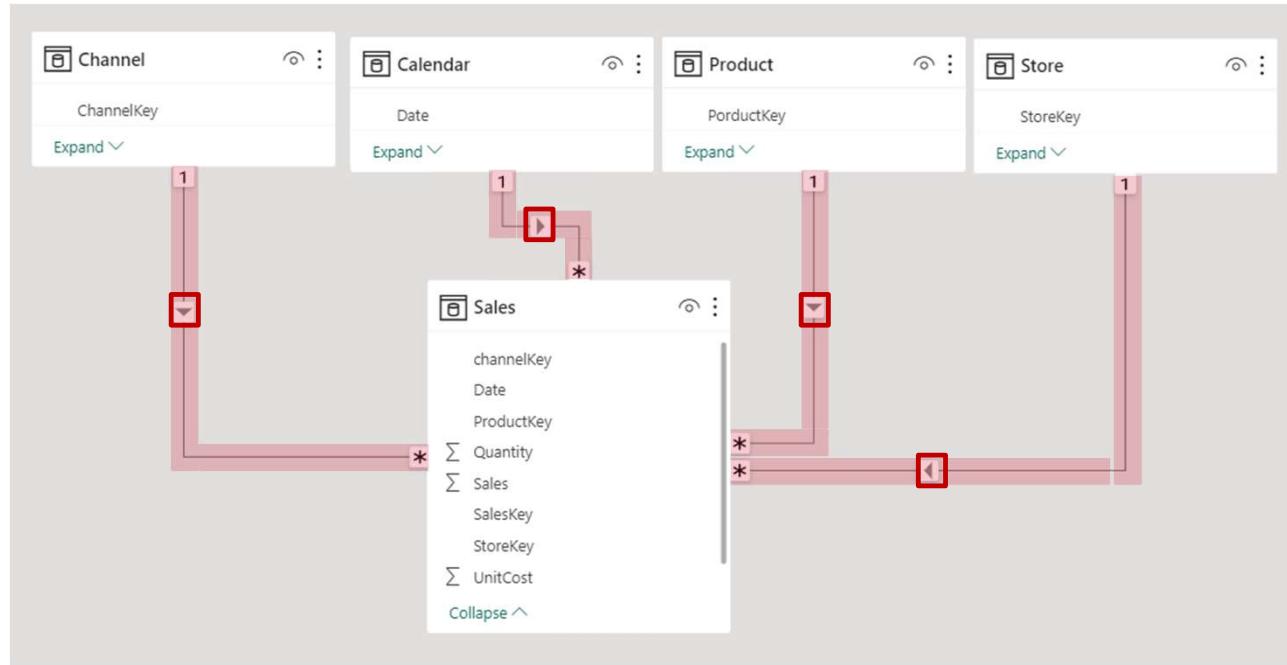




# Direction

The direction in a relationship determines how data flows between tables.

In a traditional Star model (image), dimensions filter the facts through the field that is part of the relationship.





# Building the model

- Presentation: Parts of Power BI (report, data, relations)
- Demo1.2: Create relations between 3 tables (products, sales & calendar) and draw a table in Power BI. The difference between number, text, and date columns is explained in order to validate and assign the correct data type to create the correct visualization.
- Exercise 1.2: Building a model
- Q&A



# Demo 1.2





# Presentation: Parts of Power BI (report, data, relations)

The screenshot shows the Microsoft Power BI desktop application interface. The ribbon at the top has tabs for File, Home (selected), Insert, Modeling, View, Optimize, Help, External tools, Data, Queries, Insert, Calculations, Sensitivity, and Share. The main area displays a 'Build a visual' pane for a 'Stacked column chart' with a 'Data' section containing a '+Add data' button. To the left is a 'Data' pane with sections for Filters and Data, both with search fields and 'Add data fields here' buttons. Below the ribbon, there are three red-bordered callout boxes with arrows pointing to specific areas:

- Click here to design your report** (points to the 'Insert' tab ribbon)
- Click here to see data in your tables** (points to the 'Data' pane)
- Click here to design the data model** (points to the 'Data' ribbon)

At the bottom, a horizontal navigation bar includes buttons for Page 1, ETL, MODEL, DAX, REPORT (which is highlighted in green), and COLLABORATION.

Visuals that you can use in your report

Tables and fields in your model, to use them in your report



# Your first visual in Power BI

The screenshot shows the Power BI desktop application interface. At the top, there's a ribbon with tabs: Home (selected), Insert, Modeling, View, Optimize, Help, External tools, Format, and Data / Drill. Below the ribbon, there are several icons for data sources like Excel, SQL Server, and Data. The main area shows a visualization of horizontal bars. A tooltip says "Select or drag data to populate this visual". To the right, a "Build a visual" pane is open, showing sections for Visual types (with a radio button for "Off" and "Suggest a type"), Y-axis, X-axis, Legend, Small multiples, and Tooltips, each with a "+Add data" button. A "Select data" dialog box is also open, listing categories like financials, Sales, COGS, Country, and Date, with "Date" selected. Arrows point from numbered callouts to specific elements: callout 1 points to the "Insert" tab in the ribbon, and callout 2 points to the "Date" field in the "Select data" dialog.

1 First select the visual you want to create

2 Then select the fields you want to visualize.

INTERVIEWS

ETL

MODEL

DAX

REPORT

COLLABORATION



# Exercise 1.2

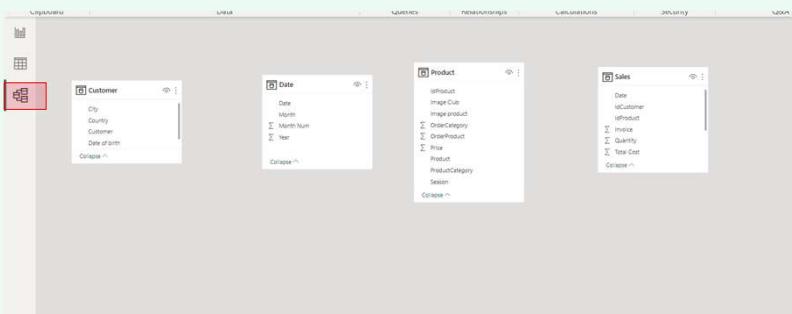


## Exercise 1.2.

Open the file **Exercise1.2.pbix** (Location: C:\PowerBIBootcamp\Week01\Exercises\Exercise1.2.pbix)

1

Check that the tables are properly loaded.



3

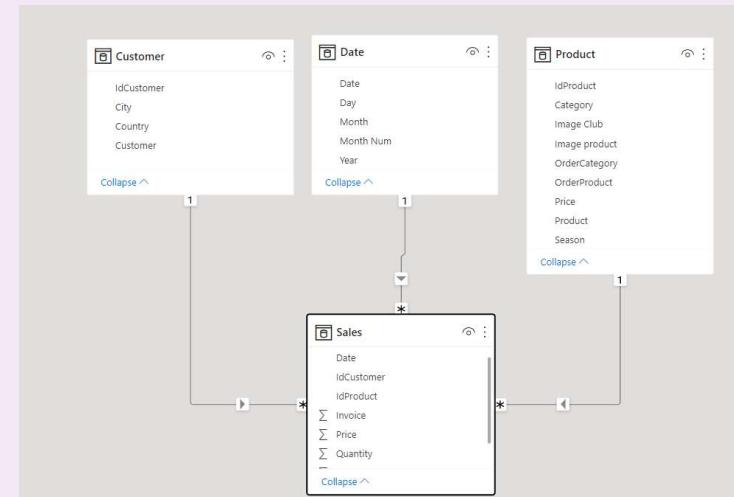
Insert a matrix to visualize the quantity sold per product.

The screenshot shows the "Build a visual" dialog for a Matrix visual. The visual type is set to "Matrix". The rows are defined by the "Product" table, and the values are defined by the "Quantity" column of the "Sales" table.

Product	Quantity
Athletic Bilbao home jersey	99
Atletico Madrid H86 cap	135
Atletico Madrid Home jersey	206
Barcelona home jersey	177
Barcelona 196 jacket	100
Barcelona pre match training jersey	117
Joao Felix signed boots	77
Real Madrid Icon jacket	180
Real Madrid youth third jersey	54
Valencia Casual hoodie	82
Valencia home jersey	112
<b>Total</b>	<b>1339</b>

2

Create the relationships between the **Sales** table and the rest of the tables through their main keys. For the relationship with the **Date** table, create an active relationship with the column date.



\*HINT: You may need to preview some of the tables to see what is in them.

*Think about:* What sort of data model are you creating?



## Poll

What type of model is recommended to use in Power BI?

- Star schema Model
- Denormalized
- Single table repeating data
- Normalized
- None of them





## Q&A

- Write your questions in the chat so that we can answer them and discuss them together.



The O'Reilly logo is centered on a vibrant orange background. The word "O'REILLY" is written in a bold, white, sans-serif font. A registered trademark symbol (®) is positioned at the top right of the letter "Y".

O'REILLY®