

Power BI





NỘI DUNG

- Phần 1: Giới thiệu về Power BI
 - Phần 2: Kết nối và xử lý dữ liệu
 - Phần 3: Data model
 - Phần 4: Phân tích dữ liệu với DAX
 - Phần 5: Trực quan hóa báo cáo
 - Phần 6: Bài tập thực tế theo từng tuần
- 



Yêu cầu

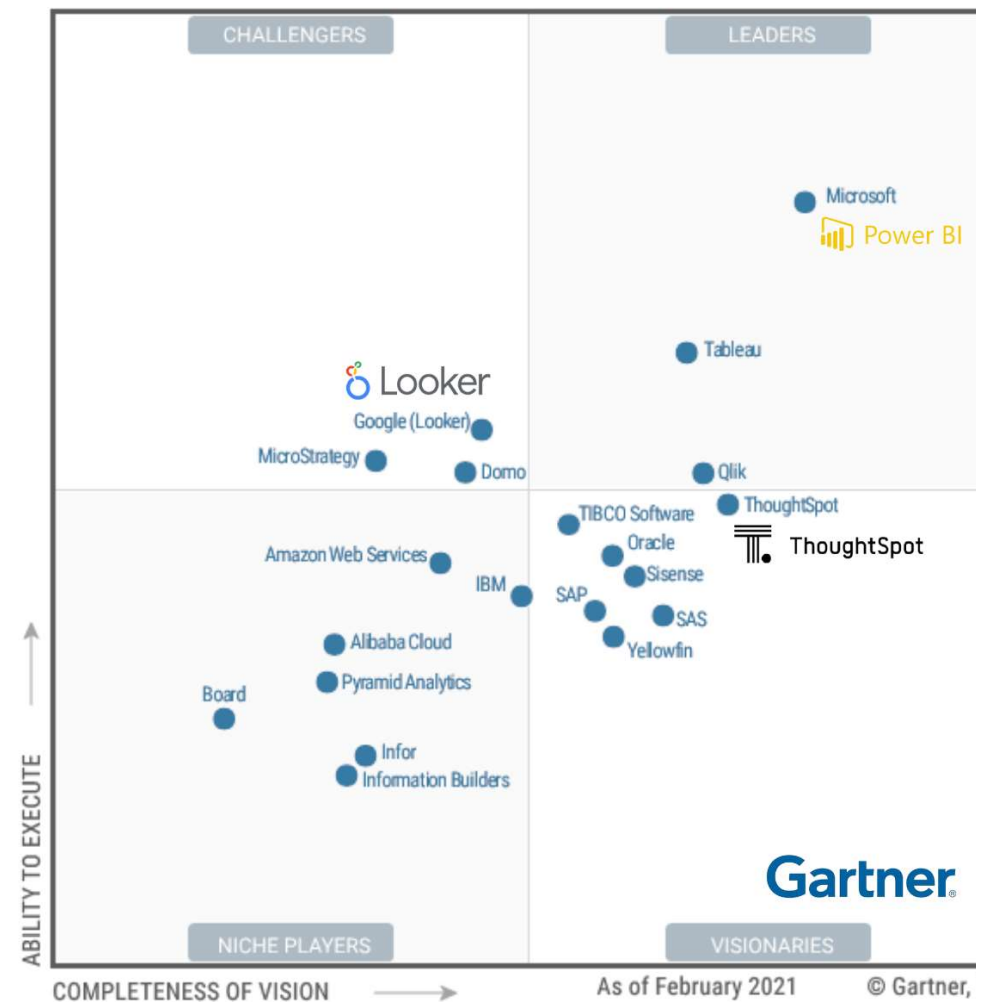
- Nắm được các kiến thức và thao tác cơ bản trên Power BI
 - Bài tập hàng tuần: áp dụng các nội dung cơ bản trong buổi học => 20% điểm toàn môn học
 - Bài tập giữa kỳ: bài tập lớn với kiến thức cơ bản:
 - Bài tập cá nhân
 - 30% điểm toàn môn học
 - Bài tập cuối kỳ: bài tập lớn với các kiến thức cơ bản và một phần nội dung nâng cao
 - Bài tập làm theo nhóm (2-3 bạn, đăng ký sau khi nộp bài giữa kỳ 1 tuần)
 - 50% điểm toàn khóa học
- 

Giới thiệu về Power BI



Power BI là gì?

Power BI là một sản phẩm văn phòng thuộc nhóm kinh doanh thông minh của Microsoft bao gồm phần mềm trên máy tính, phiên bản web, phiên bản trên điện thoại thông minh, nhằm trực quan hóa dữ liệu

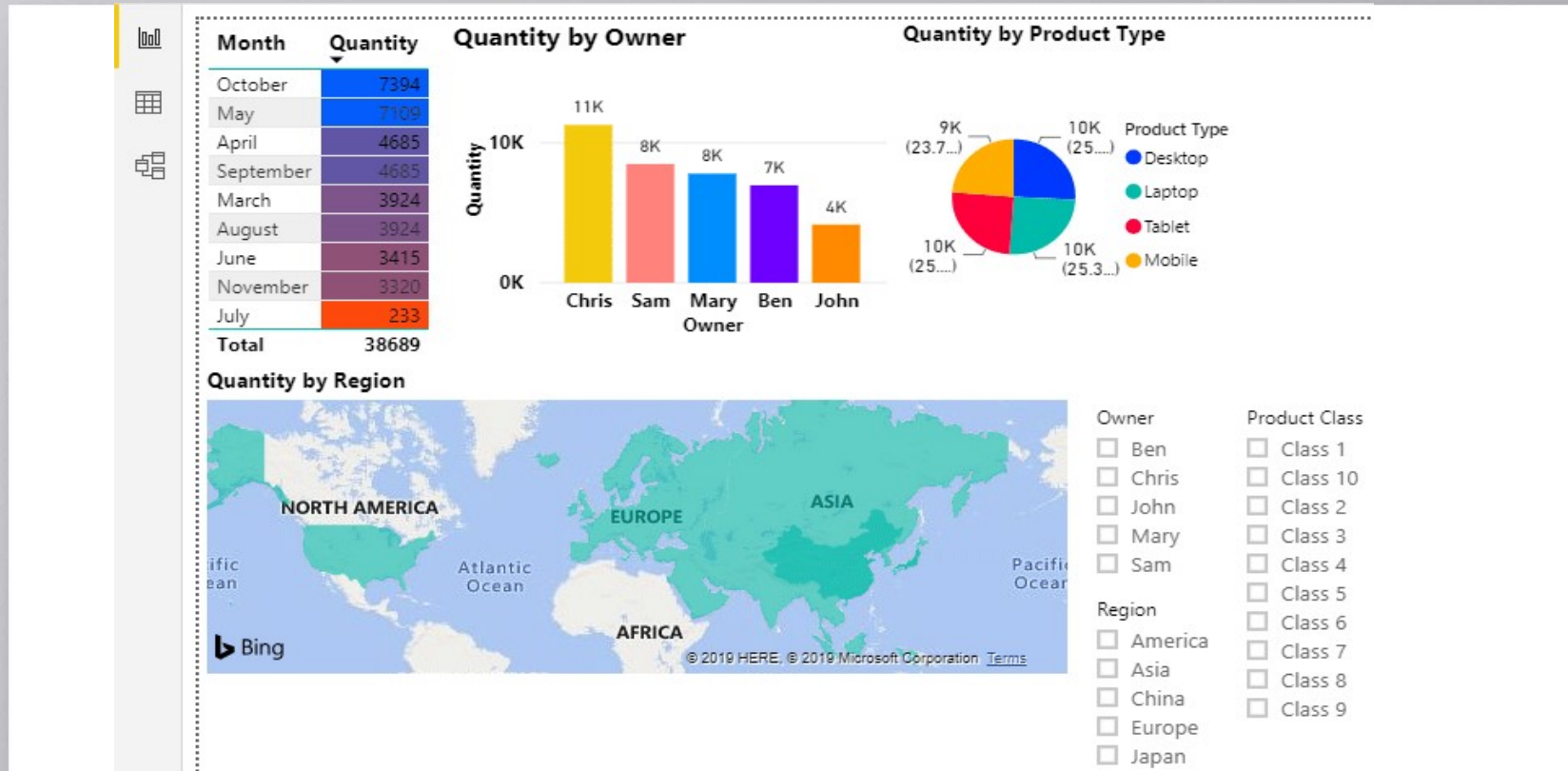


Gartner

As of February 2021

© Gartner,

Gartner



TEMPLATE 1

462.85K
kpi1

471.48K
kpi2

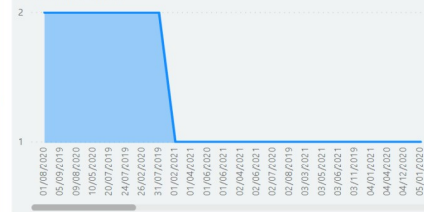
529.51K
kpi3

502.63K
kpi4

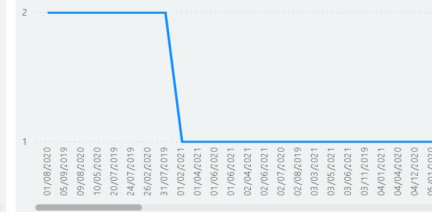
47.67K
kpi5

406.28K
kpi6

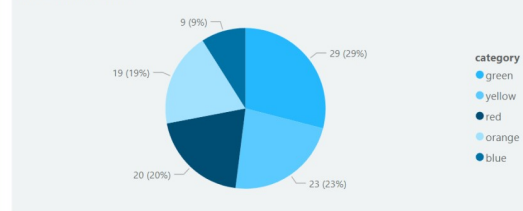
First Visualization



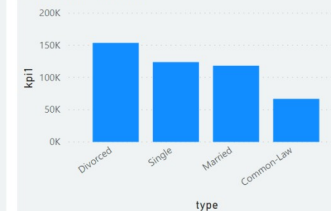
Second Visualization



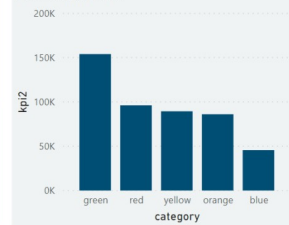
Third Visualization



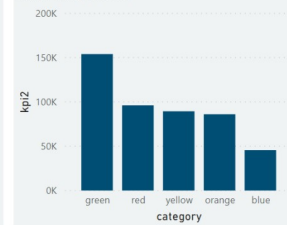
Fourth Visualization



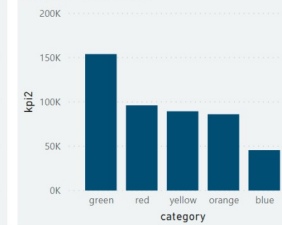
Fifth Visualization

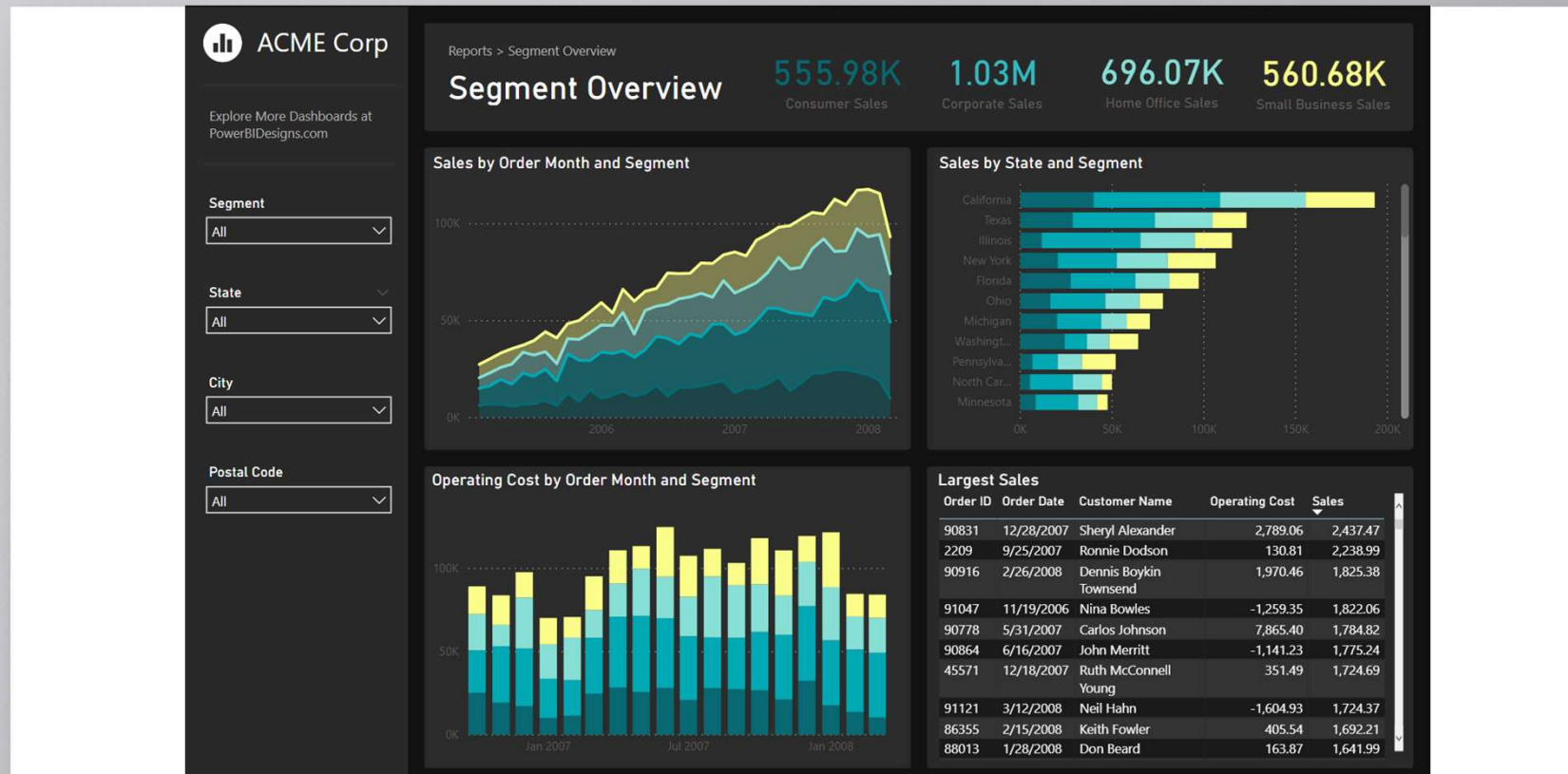


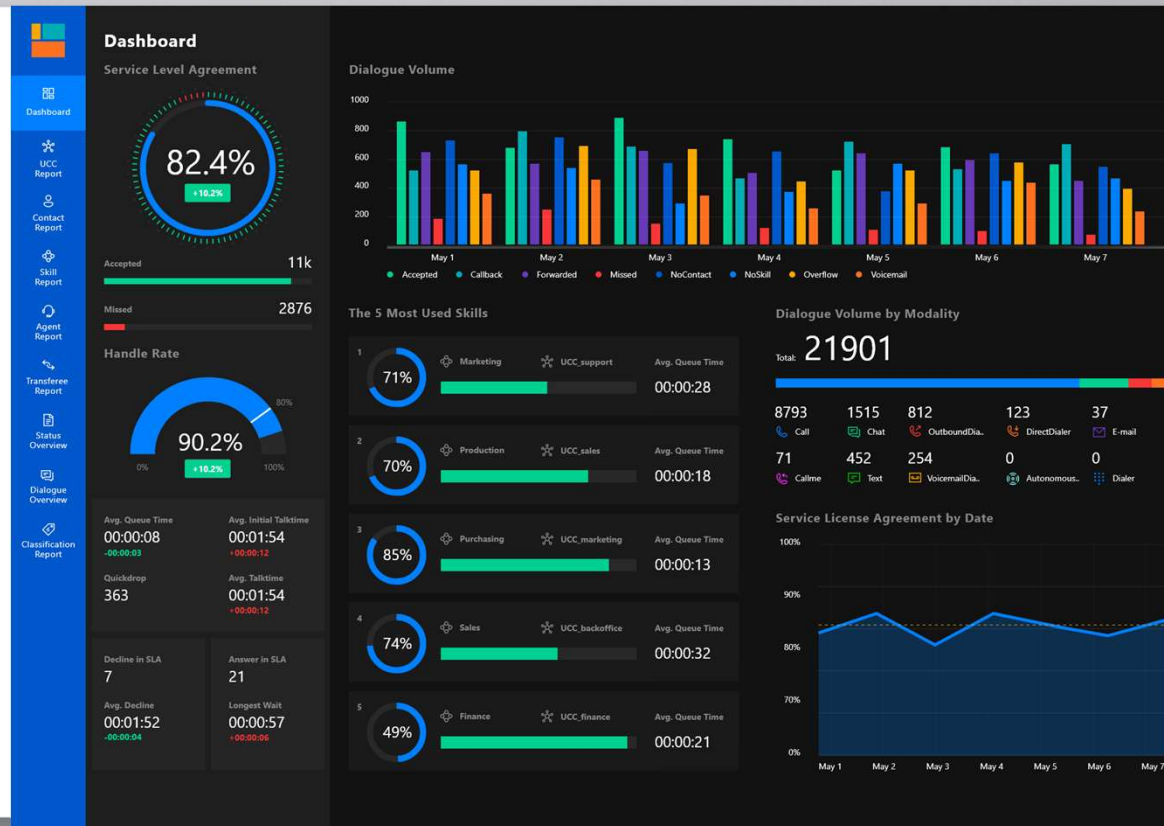
Sixth Visualization



Seventh Visualization







Tại sao nên sử dụng Power BI?

Kết nối, chuyển đổi và phân tích dữ liệu lớn

Truy cập dữ liệu từ nhiều nguồn (table, sheet, cloud, folder, v.v), tạo các quy trình định hình và tải dữ liệu hoàn toàn tự động (ETL)

Xây dựng các mô hình quan hệ để trộn dữ liệu từ nhiều nguồn

Tạo ra các relationships để phân tích tổng thể trên toàn bộ dữ liệu

Hỗ trợ phân tích dữ liệu bằng biểu thức phân tích (DAX)

Hỗ trợ phân tích bộ dữ liệu và cho phép nhiều phân tích nâng cao với DAX

Trực quan dữ liệu với các báo cáo và bảng điều khiển tương tác

Xây dựng các công cụ phân tích kinh doanh tùy chỉnh với các tính năng trực quan và bảng điều khiển tốt nhất

Power BI là công cụ hàng đầu và tương lai của tương tác

Microsoft Power BI trực quan, mạnh mẽ và hoàn toàn miễn phí để bắt đầu.

Giao diện của Power BI

3 phần chính:

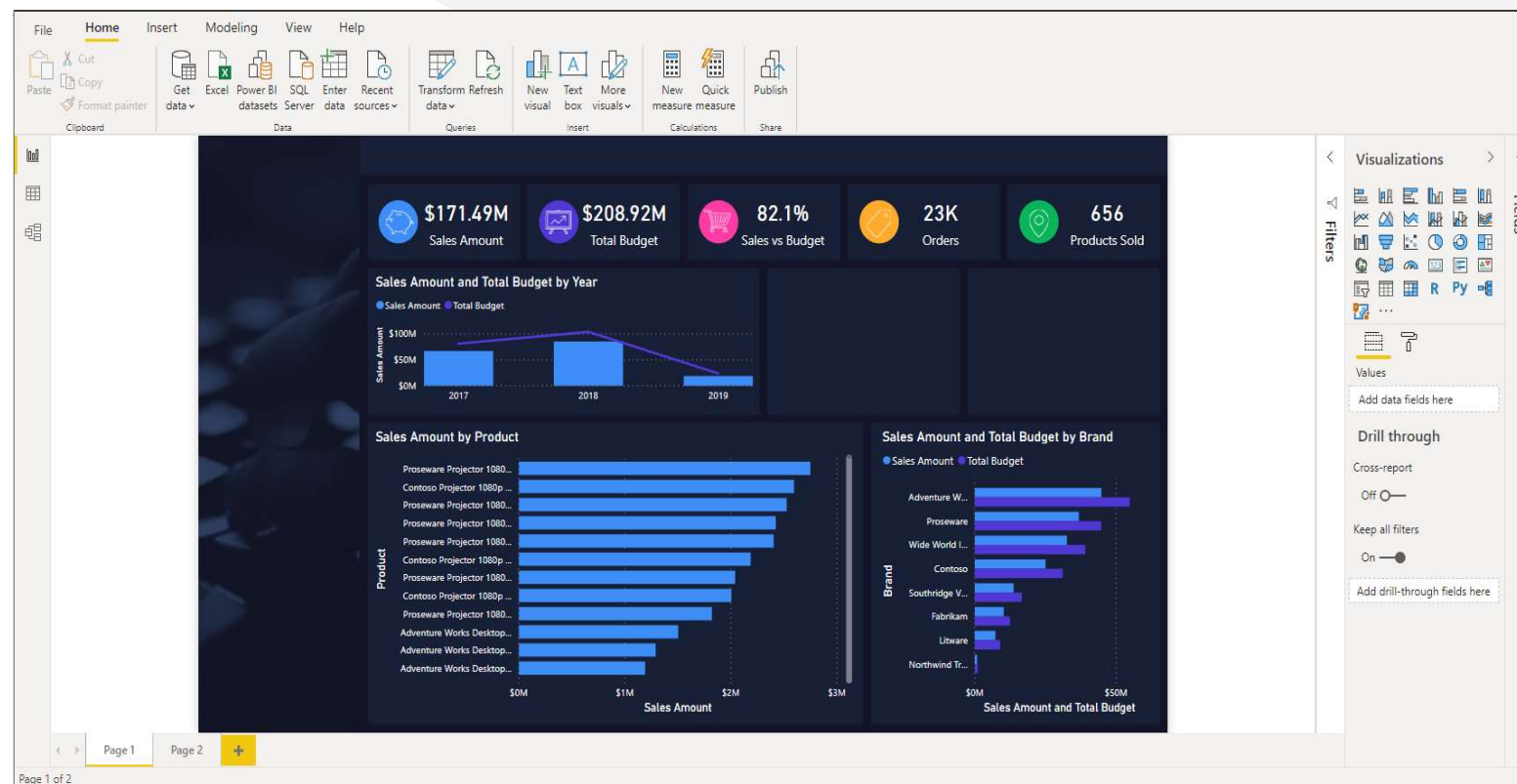
Report



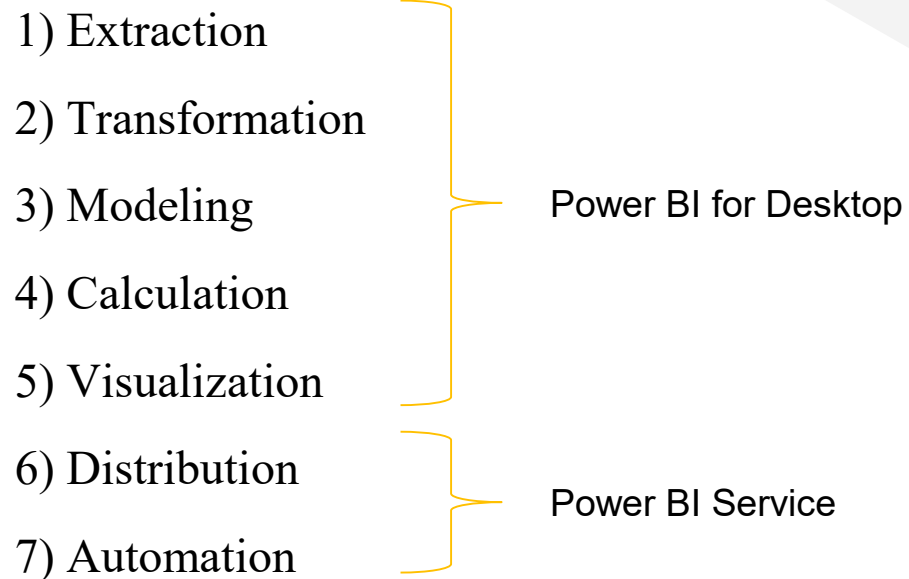
Data



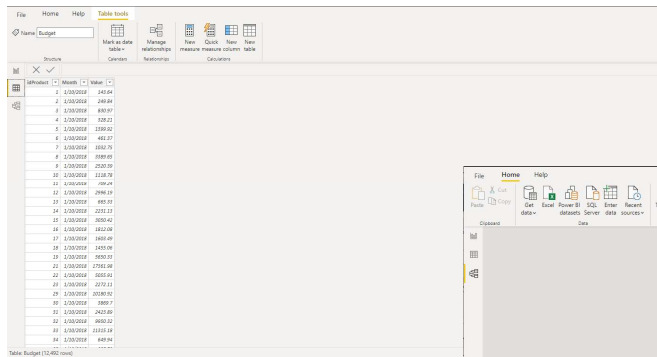
Model



7 bước để sử dụng được Power BI



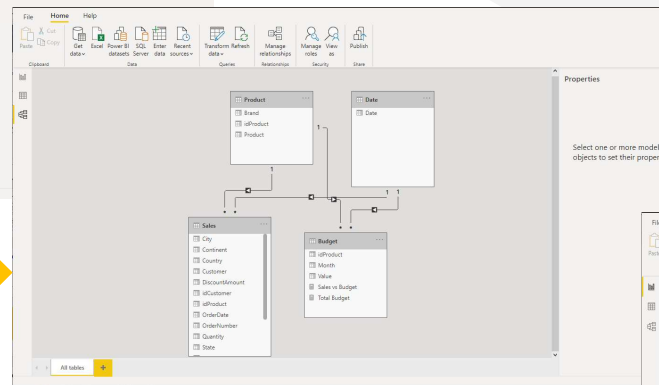
Quy trình làm việc



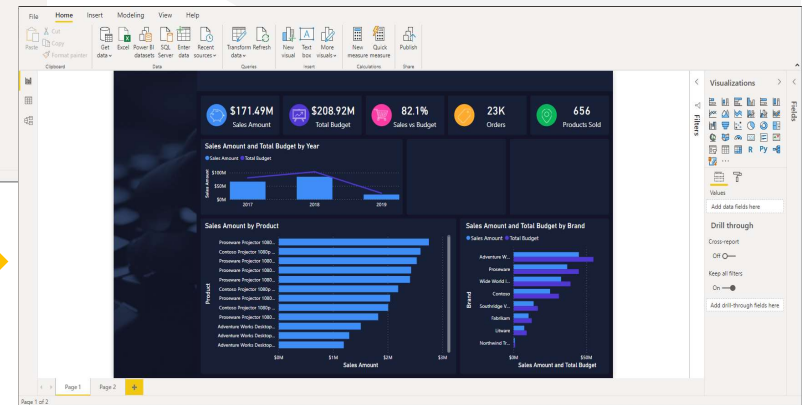
Year	Month	Sales
2	1/2010	145.64
2	2/2010	248.61
2	3/2010	401.91
2	4/2010	338.21
2	5/2010	339.61
2	6/2010	482.21
2	7/2010	342.71
2	8/2010	388.01
2	9/2010	250.19
2	10/2010	218.79
2	11/2010	194.41
2	12/2010	198.41
2	1/2011	221.21
2	2/2011	202.41
2	3/2011	242.09
2	4/2011	400.09
2	5/2011	243.09
2	6/2011	400.11
2	7/2011	2781.48
2	8/2011	1051.41
2	9/2011	2071.11
2	10/2011	2018.71
2	11/2011	1989.1
2	12/2011	2420.81
2	1/2012	4001.11
2	2/2012	2333.14
2	3/2012	449.11



Tổ chức dữ liệu từ raw data



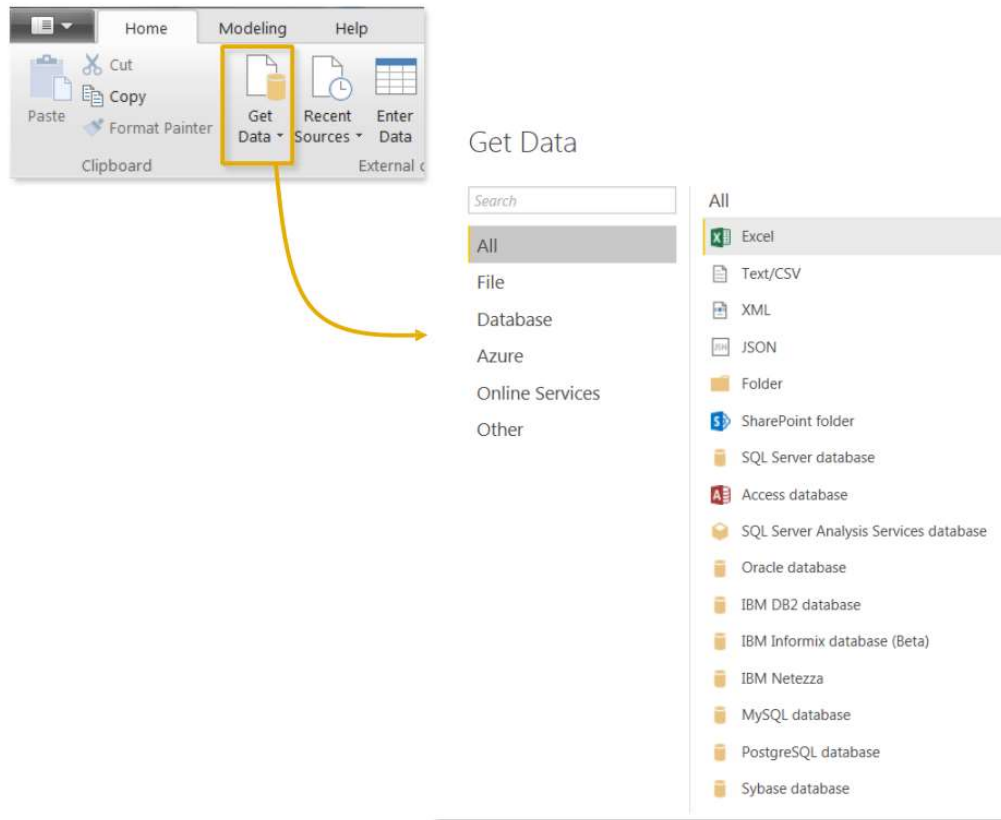
Tạo relationships để kết nối các nguồn dữ liệu



Trực quan hóa các báo cáo và tạo tương tác

Kết nối và xử lý dữ liệu

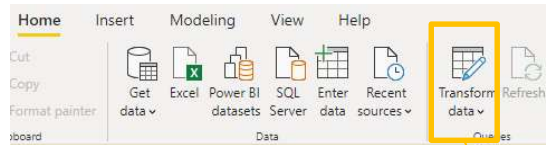




Power BI có thể kết nối tới rất nhiều nguồn data, bao gồm (nhưng không hạn chế):

- **Flat files & Folders** (*csv, text, xls, etc*)
- **Databases** (*SQL, Access, Oracle, IBM, Azure, etc*)
- **Online Services** (*Sharepoint, GitHub, Dynamics 365, Google Analytics, Salesforce, Power BI Service, etc*)
- **Others** (*Web feeds, R scripts, Spark, Hadoop, etc*)

Query Editor



Query editing tool

Formula Bar

Query list

The main interface of the Query Editor is shown. The top ribbon includes tabs: File, Home, Transform, Add Column, View, Tools, and Help. The 'Transform' tab is active, showing various data transformation options like 'Choose Columns', 'Remove Columns', 'Keep Rows', 'Remove Rows', 'Sort', 'Split Column', 'Group By', 'Data Type: Date', 'Merge Queries', 'Append Queries', 'Combine Files', 'Text Analytics', 'Vision', and 'Azure Machine Learning'. Below the ribbon is the 'Formula Bar' showing the query name and its M code: `Table.TransformColumnTypes(\*Promoted Headers\*,{{"Ngày", type date}, {"Mã Nhân viên", Int64.Type}, {"Mã sản phẩm", type text}, {"Số lượng bán", Int64.Type}})`. To the left of the main data area is the 'Query list' showing a list of queries: 'Bán hàng', 'Sản phẩm', 'TT Nhân viên', and 'File thực hành'. To the right of the main data area are two panels: 'PROPERTIES' showing the 'Name' as 'Bán hàng' and 'All Properties' link, and 'APPLIED STEPS' showing a list of steps: 'Source', 'Navigation', 'Promoted Headers', and 'Changed Type' (highlighted with a yellow box). The main data area displays a table with columns: 'Ngày', 'Mã Nhân viên', 'Mã sản phẩm', and 'Số lượng bán'. The table contains 30 rows of data.

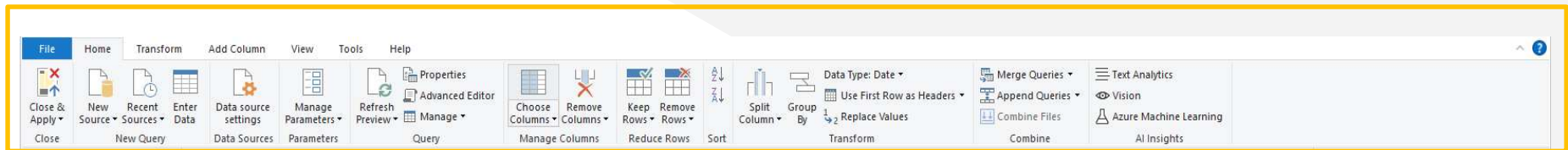
	Ngày	Mã Nhân viên	Mã sản phẩm	Số lượng bán
1	1/1/2019	10019	P10007	199
2	1/1/2019	10005	P10001	71
3	1/1/2019	10004	P10019	201
4	1/1/2019	10014	P10022	246
5	1/1/2019	10005	P10004	106
6	1/1/2019	10005	P10018	107
7	1/1/2019	10009	P10026	129
8	1/1/2019	10013	P10021	63
9	1/1/2019	10008	P10008	59
10	1/1/2019	10001	P10023	156
11	1/1/2019	10001	P10027	168
12	1/1/2019	10013	P10029	221
13	1/1/2019	10009	P10018	221
14	1/1/2019	10015	P10027	218
15	1/1/2019	10013	P10022	78
16	1/1/2019	10001	P10015	250
17	1/1/2019	10019	P10024	252
18	1/1/2019	10003	P10022	70
19	1/1/2019	10010	P10025	258
20	1/1/2019	10001	P10029	167
21	1/1/2019	10003	P10014	300
22	1/1/2019	10015	P10007	100
23	1/1/2019	10020	P10017	123
24	1/1/2019	10003	P10006	160
25	1/1/2019	10004	P10009	198
26	1/1/2019	10008	P10020	50
27	1/1/2019	10003	P10020	147
28	1/1/2019	10011	P10017	156
29	1/1/2019	10004	P10005	203
30	1/1/2019	10014	P10030	160

Table name & Properties

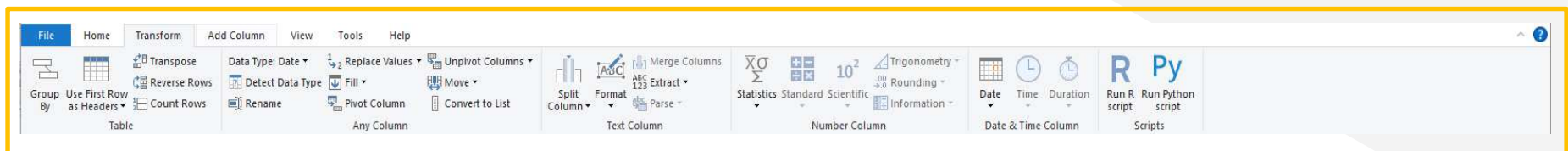
Applied steps

Query Editing Tools

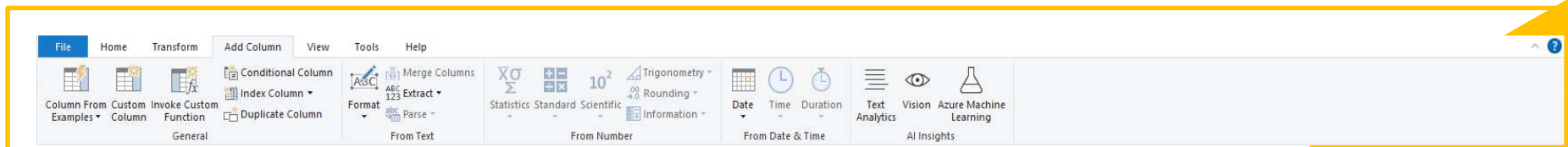
Tab HOME bao gồm các cài đặt chung và các công cụ chỉnh sửa bảng thông dụng



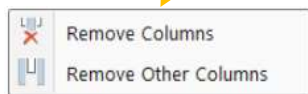
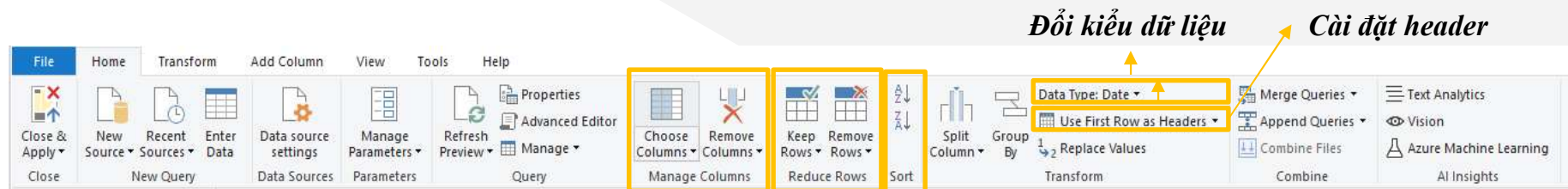
Tab TRANSFORM bao gồm các công cụ chỉnh sửa các bảng đã có sẵn



Tab ADD COLUMN bao gồm các công cụ tạo cột mới dựa trên các yêu cầu

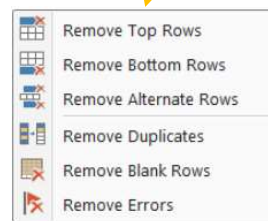


Chỉnh sửa bảng cơ bản



Chọn hoặc xóa cột

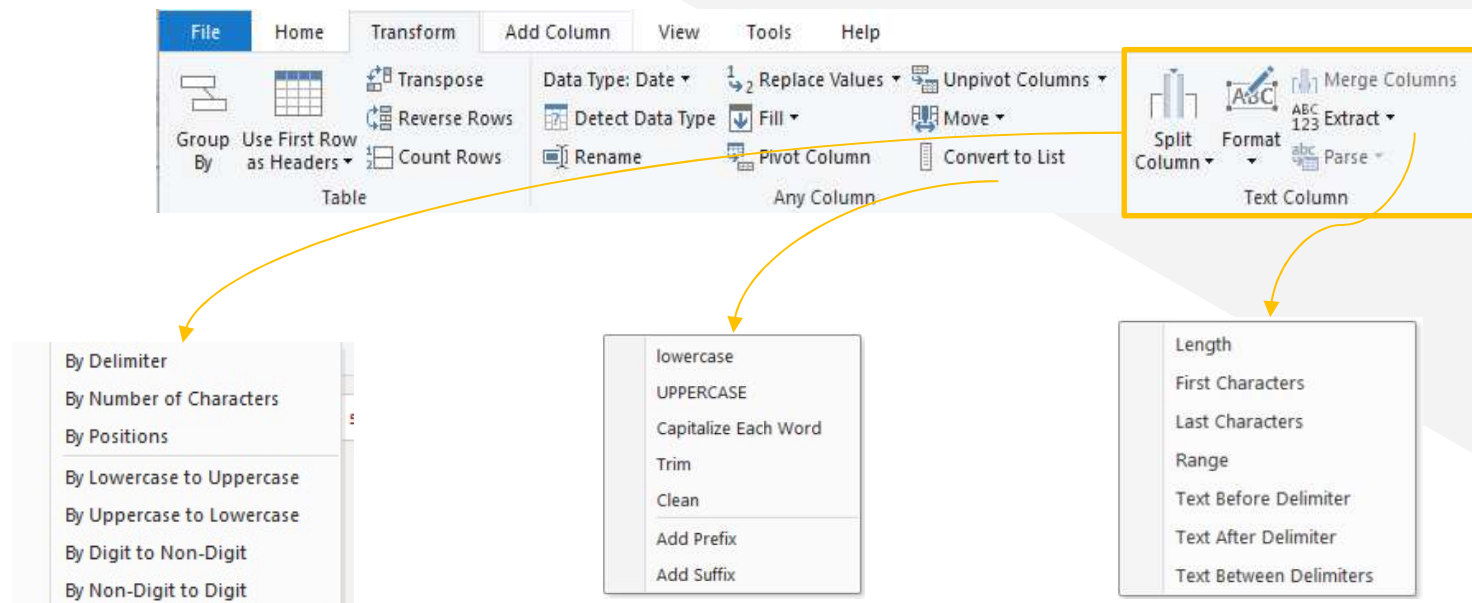
Tip: sử dụng
“Remove Other
Columns” khi chỉ
muốn giữ lại 1 cột



Giữ hoặc xóa các hàng

Tip: sử dụng
“Remove Duplicates”
để xóa các dữ liệu
trùng từ bảng cũ

Text tools



Chia cột text theo điều kiện

Dựa vào khoảng trống ở giữa, ký tự đặc biệt, con số, etc

Format text column

Tips: Sử dụng **Trim** để loại bỏ khoảng trắng hoặc **Clean** để xóa các ký tự ko thể in

Extract characters từ một text column

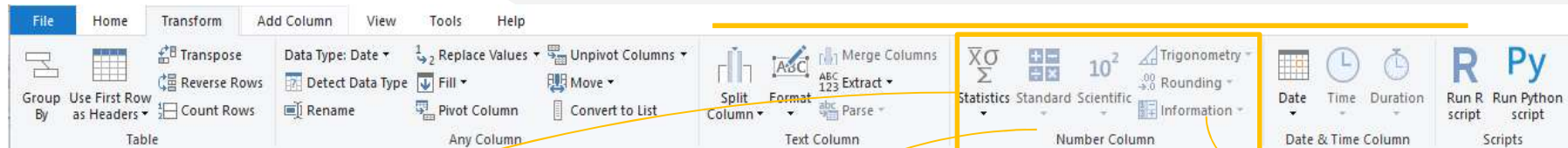
dựa trên chiều dài, ký tự đầu/cuối, dấu ngăn cách, etc

NOTE:

Transform: thao tác trên cột sẵn có

Add column: thao tác khi thêm cột mới

Number tools



Sum
Minimum
Maximum
Median
Average
Standard Deviation
Count Values
Count Distinct Values

Statistics functions cho phép bạn thực hiện một số tính toán cơ bản từ cột được chọn (sum, min/max, average, count, ...)

Note: Những tools này sẽ trả về một Kết quả đơn (SINGLE value), và thường sử dụng để tính toán trên bảng thay vì sử dụng ở bước xử lý data (load data)

Add
Multiply
Subtract
Divide
Integer-Divide
Modulo
Percentage
Percent Of

Standard

Absolute Value
Power
Square Root
Exponent
Logarithm
Factorial

Scientific

Sine
Cosine
Tangent
Arcsine
Arccosine
Arctangent

Trigonometry

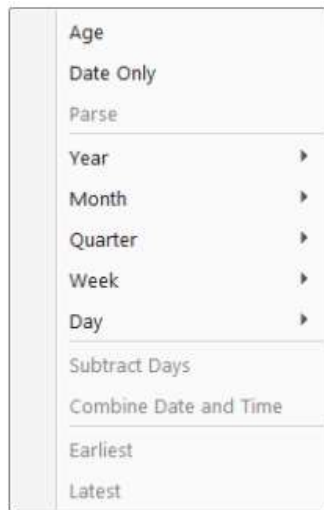
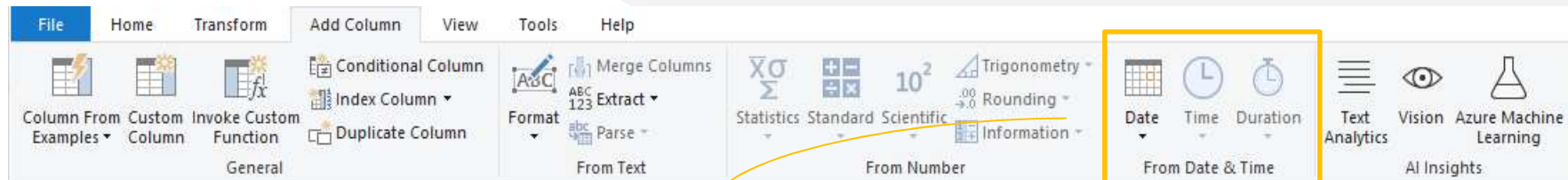
Is Even
Is Odd
Sign

Information tools cho phép bạn xác định kết quả Logic hoặc Nhị phân (TRUE/FALSE or 1/0) dựa trên giá trị Chẵn/Lẻ

Standard, Scientific và Trigonometry tools cho phép bạn apply các công thức ở mức độ cơ bản (cộng, trừ, nhân, chia, ...) hoặc nâng cao (power, logarithm, sine, tangent, ...) cho các giá trị ở trong cột

Note: Khác với Statistics options, các công cụ này áp dụng cho các giá trị trong cột chứ không trả về một Kết quả đơn.

Date Tool



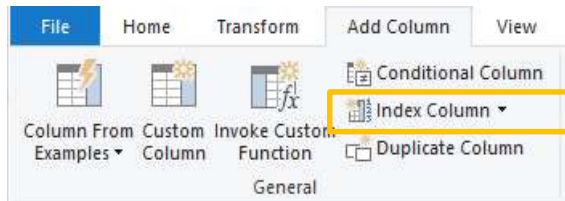
Date & Time tools tương đối đơn giản, bao gồm một số tùy chọn như:

- **Age:** sự khác nhau giữa thời gian hiện tại và thời gian trong cột
- **Date Only:** Bỏ đi các thành phần khác ngoài định dạng Date trong cột
- **Year/Month/Quarter/Week/Day:** Xuất ra các thành phần từ trường Date
- **Earliest/Latest:** Trả về kết quả về dữ liệu Date sớm nhất hoặc muộn nhất từ Date Column dưới dạng Kết quả đơn

Note: Với dữ liệu dạng Date, chúng ta nên thực hiện “Add Column” để thêm một trường/cột mới, hơn là chỉnh sửa trực tiếp dữ liệu trong “Transform”

Nên tạo một bảng chỉ chứa một cột **date column** và sử dụng Date tools để tạo ra một **calendar table** – sẽ cực kỳ hữu ích khi làm việc trong một file PowerBI có nhiều bảng

Thêm Index column



Index Columns chứa một list giá trị để xác định vị trí trong cột hoặc trong bảng (*điển hình thường bắt đầu từ 0 hoặc 1*)

Cột này cũng thường được dùng để tạo ra các **IDs** để có thể xác định quan hệ giữa các bảng

	i²₃ Index	OrderDate	StockDate	A⁸_C OrderNumber	i²₃ ProductKey	i²₃ CustomerKey
1	1	1/1/2015	9/21/2001	SO45080	332	14657
2	2	1/1/2015	12/5/2001	SO45079	312	29255
3	3	1/1/2015	10/29/2001	SO45082	350	11455
4	4	1/1/2015	11/16/2001	SO45081	338	26782
5	5	1/2/2015	12/15/2001	SO45083	312	14947
6	6	1/2/2015	10/12/2001	SO45084	310	29143
7	7	1/2/2015	12/18/2001	SO45086	314	18747
8	8	1/2/2015	10/9/2001	SO45085	312	18746
9	9	1/3/2015	10/3/2001	SO45093	312	18906
10	10	1/3/2015	9/29/2001	SO45090	310	29170
11	11	1/3/2015	12/11/2001	SO45088	345	11398
12	12	1/3/2015	10/24/2001	SO45092	313	18899
13	13	1/3/2015	12/16/2001	SO45089	351	25977
14	14	1/3/2015	10/26/2001	SO45091	314	18909
15	15	1/3/2015	9/11/2001	SO45087	350	11388
16	16	1/3/2015	9/11/2001	SO45094	310	22785
17	17	1/4/2015	10/30/2001	SO45096	312	12483
18	18	1/4/2015	10/30/2001	SO45097	313	29151

Thêm Conditional column

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name
QuantityType

	Column Name	Operator	Value		Output
If	OrderQuantity	equals	1	Then	Single Item
Else If	OrderQuantity	is greater than	1	Then	Multiple Items

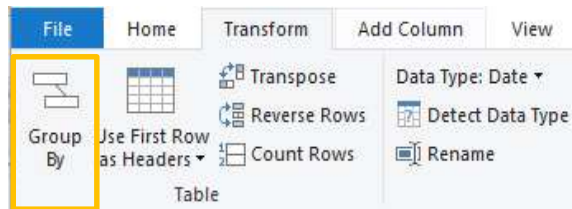
Add rule

Otherwise
Other

OK Cancel

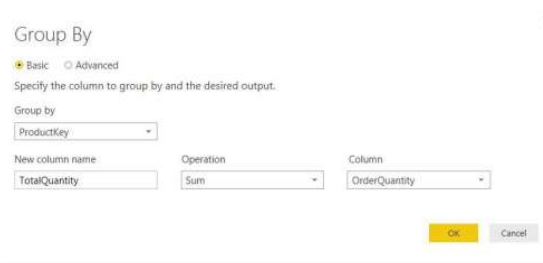
Conditional Columns cho phép bạn tạo ra một trường mới dựa trên một nguyên tắc hoặc điều kiện nào đó (giả định NẾU/THÌ)

Grouping data



Group By cho phép bạn sắp xếp data theo các cấp độ hay các nhóm (data từ chi tiết ngày về theo tháng, giao dịch theo cửa hàng, ...)

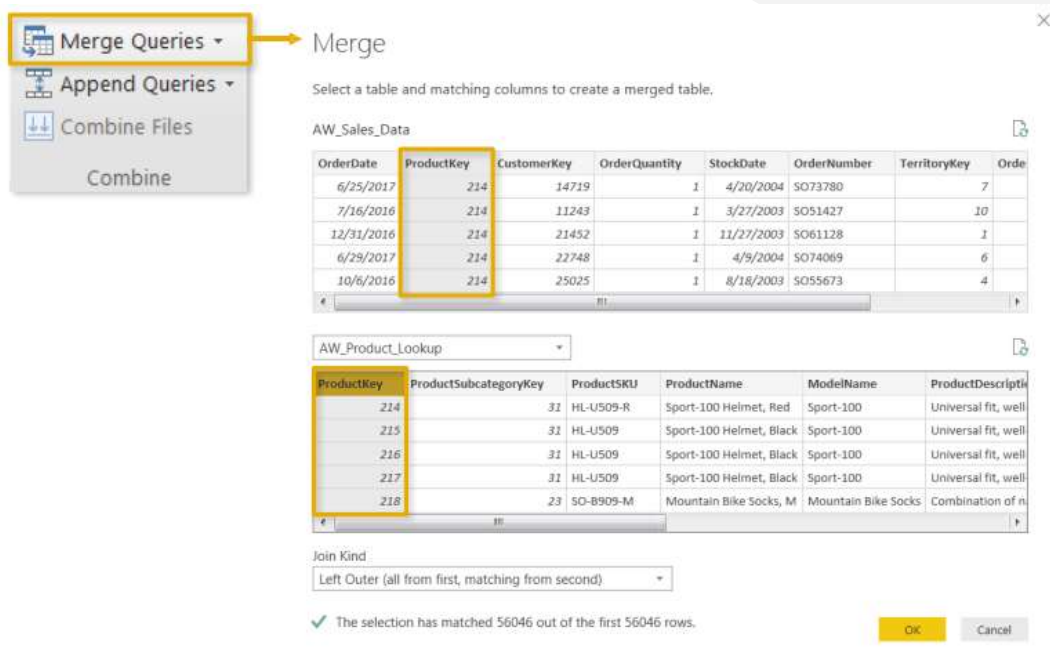
	OrderDate	ProductKey	CustomerKey	OrderQuantity
1	6/25/2017	214	14719	1
2	7/16/2016	214	11243	1
3	12/31/2016	214	21452	1
4	6/29/2017	214	22748	1
5	10/6/2016	214	25025	1
6	10/7/2016	214	16504	1
7	10/13/2016	214	13043	1
8	1/19/2017	214	23101	1
9	9/7/2016	214	24900	1
10	1/19/2017	214	24196	1
11	6/29/2017	214	12963	1
12	11/6/2016	214	14570	1
13	11/13/2016	214	16999	1
14	7/31/2016	214	12281	1
15	10/9/2016	214	15685	1
16	8/1/2016	214	16982	1
17	12/4/2016	214	12835	1



Ví dụ: Nhóm data và tạo cột mới tên **“TotalQuantity”** dựa trên **“ProductKey”**, sau đó tính tổng
NOTE: những trường không thuộc trường điều kiện nhóm hoặc trường kết quả sẽ bị xóa

	ProductKey	TotalQuantity
1	214	2099
2	217	1940
3	222	1995
4	225	4151
5	228	392
6	231	408
7	234	424
8	237	381
9	310	169
10	311	139
11	312	179
12	313	168
13	314	157
14	320	10
15	321	55
16	322	5
17	323	34

MERGING QUERIES

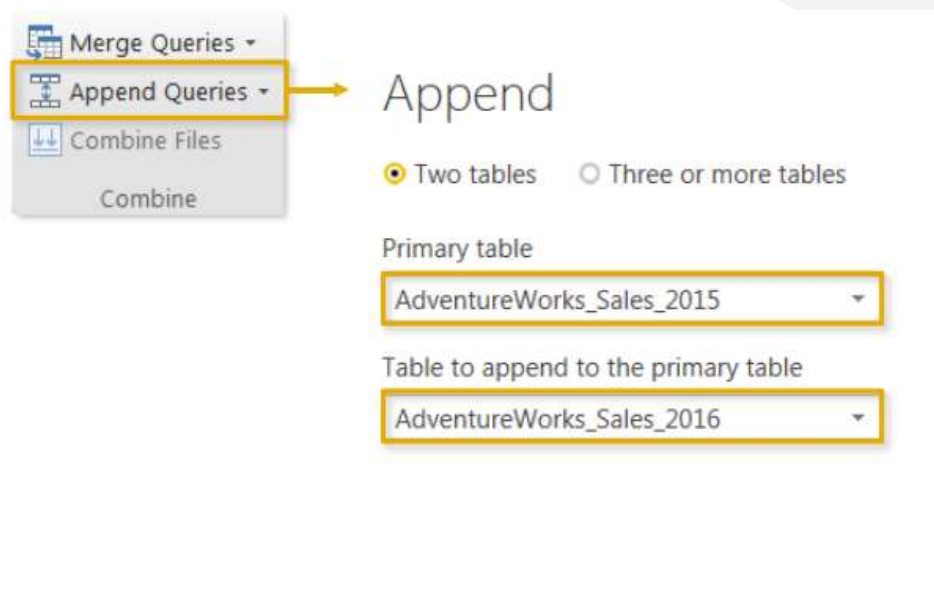


Merging queries cho phép bạn **gộp bảng** dựa trên những cột chung (giống hàm VLOOKUP trong Excel)

NOTE: Merging queries *thêm cột* tới một bảng đã có sẵn

Không nên gộp bảng quá nhiều để giảm thiểu số lượng bảng.
Thực tế trong PBI, sẽ tốt hơn nếu chia ra thành các bảng và tạo **relationships** cho chúng.

APPENDING QUERIES

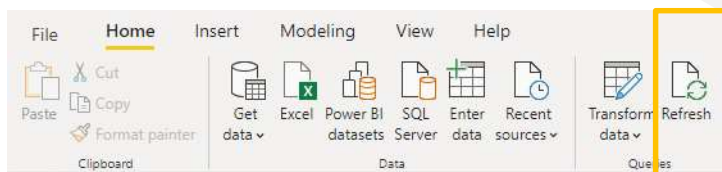


Appending queries cho phép bạn **combine** tables có cùng kiểu cấu trúc và kiểu định dạng

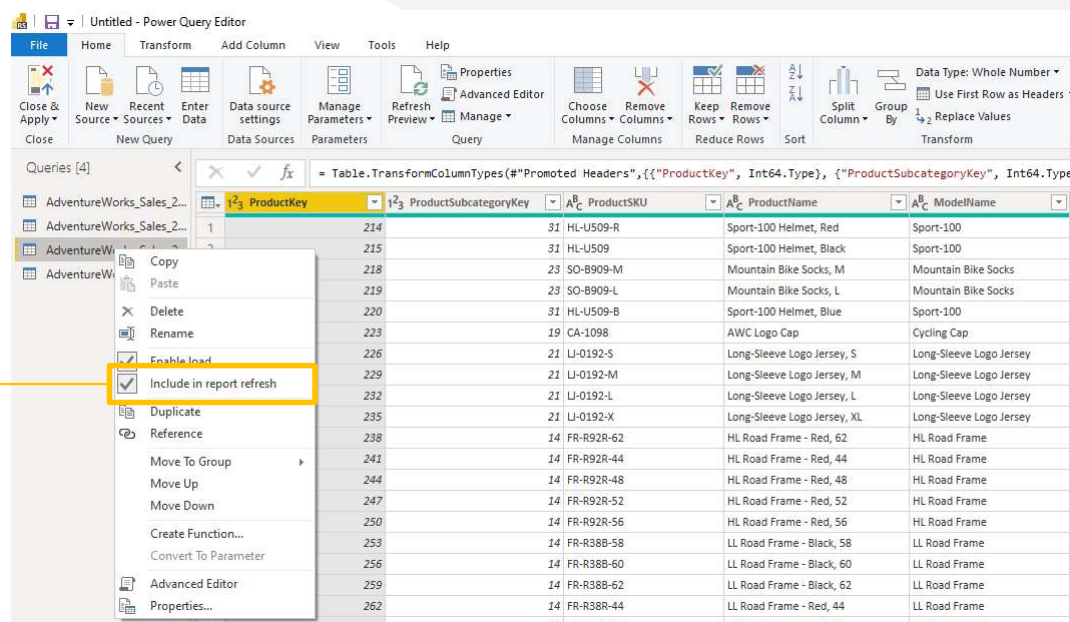
NOTE: Appending *thêm dòng* tới một bảng đã có sẵn

Sử dụng “**Folder**” option (Get Data > More > Folder) để gộp các file trong một folder (nhưng phải có chung cấu trúc); nếu bạn thêm 01 file mới, bạn chỉ cần refresh queries và mọi thứ sẽ tự động update!

REFRESHING QUERIES



Các queries trong model sẽ được refresh khi bạn sử dụng command “Refresh” từ tab **Home**



Có thể cài đặt việc Refresh thông qua phần “*Include in report refresh*” khi bấm vào các Queries