

Hệ thống thông tin phục vụ trí tuệ kinh doanh

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH



POWERBI

A Step-by-Step

Hệ thống thông tin phục vụ trí tuệ kinh doanh
Học kì 1 2021 - 2022

1. GIỚI THIỆU TỔNG QUAN VỀ POWERBI

1.1. Khái niệm PowerBI

Power BI là một tập hợp các dịch vụ phần mềm, ứng dụng di động, các công cụ kết nối và xử lý dữ liệu hoạt động cùng với nhau để có thể tập hợp dữ liệu từ nhiều nguồn, xử lý dữ liệu từ các nguồn này để đưa dữ liệu thành thông tin hữu ích dưới dạng hình ảnh, đồ thị, biểu đồ có khả năng tương tác trên nhiều thiết bị và nền tảng. Power BI giúp bạn kết nối và tập hợp dữ liệu từ nhiều hệ thống phần mềm cũng như dữ liệu riêng lẻ trên các files Excel, files text và các dạng dữ liệu khác, sau đó biến dữ liệu này thành thông tin hữu ích cho doanh nghiệp của bạn thông qua việc trực quan hoá dữ liệu giúp các nhà quản lý đưa ra được những quyết định dựa trên số liệu, giảm bớt sự cảm tính, tập trung vào những yếu tố dẫn tới thành công của tổ chức.

Power BI còn là sản phẩm dẫn đầu trong ngành analytics và business intelligence theo hai công ty nghiên cứu thị trường nổi tiếng Forrester và Gartner.

1.2. Các thành phần trong PowerBI

Power BI bao gồm:

- Power BI Desktop: là một phần mềm trên hệ điều hành Windows có tác dụng tập hợp dữ liệu, xử lý dữ liệu, xây dựng mô hình dữ liệu (**Data Model**) và trực quan hoá dữ liệu.
- Dịch vụ đám mây Power BI Service hay còn gọi là Power BI Online, là một dịch vụ phần mềm (**SaaS – Software as a Service**), là một phần của Power BI. Với Power BI Service, các báo cáo và dashboard

Hệ thống thông tin phục vụ trí tuệ kinh doanh

giúp bạn lắng nghe nhịp thở cũng như sức khỏe của doanh nghiệp, business của bạn.

- Ứng dụng di động Power BI apps cho các nền tảng Windows, iOS và Android: cho phép bạn truy cập vào các báo cáo, dashboard ngay trong lòng bàn tay.

Ba yếu tố – Power BI Desktop, Power BI Service, Power BI mobile apps – được thiết kế để giúp bạn tạo báo cáo dashboard, chia sẻ báo cáo dashboard và sử dụng những thông tin được rút ra từ dữ liệu kinh doanh trong doanh nghiệp của bạn một cách hiệu quả nhất.

Một thành phần thứ tư, Power BI Report Server, cho phép bạn xuất bản những báo cáo lên hệ thống Power BI Server đặt tại công ty của bạn (**on-premises**) sau khi đã hoàn thành các báo cáo này trong Power BI Desktop.

Ví dụ: Bạn có thể xem báo cáo và trang tổng quan trong **Power BI service** và đó có thể là tất cả những gì bạn làm với Power BI. Nhưng đồng nghiệp của bạn chuyên tạo báo cáo kinh doanh, lập báo cáo số, kinh doanh có thể sử dụng rộng rãi **Power BI Desktop** (và xuất bản các báo cáo Power BI Desktop lên Power BI service, sau đó bạn sử dụng để xem chúng). Và một đồng nghiệp khác, trong lĩnh vực bán hàng, có thể sử dụng ứng dụng điện thoại Power BI để theo dõi tiến độ về hạn ngạch bán hàng và đi sâu vào chi tiết khách hàng tiềm năng mới.

1.3 Lợi ích của PowerBI

- **Khả năng xử lý khối lượng dữ liệu lớn hơn Excel**

Có thể các bạn đã biết, đối với Excel mà nhiều doanh nghiệp còn đang sử dụng, chúng ta bị giới hạn số lượng dòng dữ liệu trên 1 Sheet . Đối với

Hệ thống thông tin phục vụ trí tuệ kinh doanh

những hoạt động của doanh nghiệp trước kia, có thể số lượng dòng dữ liệu này là đủ dùng, ... cho tới khi nhu cầu lưu trữ và xử lý dữ liệu lớn hơn mức này. Với Power BI thì giới hạn này đã bị bỏ qua do khả năng lưu trữ và xử lý dữ liệu đã được Microsoft thiết kế cho một thời kỳ kỷ nguyên dữ liệu, bạn ít khi phải lo lắng về việc quá nhiều dữ liệu dẫn tới công thức Excel bị chậm, máy bị treo hoặc không xử lý được nữa.

- Tích hợp cả các đoạn code Python và R dùng để trực quan hóa
- Power Query cung cấp nhiều lựa chọn liên quan đến việc sử dụng và làm sạch dữ liệu
- Workflow tự động từ dữ liệu tới báo cáo

Đối với những công việc báo cáo làm bằng Excel thì người làm báo cáo sẽ có quy trình (Workflow for Reporting) như sau:

- Tập hợp dữ liệu từ nhiều nguồn (Thông qua email, ổ cứng mạng (Shared drives), ...)
- Xử lý những dữ liệu này cho phù hợp với cấu trúc dữ liệu của báo cáo (Làm bằng tay, copy-paste, sử dụng công thức hoặc nâng cao hơn thì tự động hoá công việc này bằng ngôn ngữ lập trình VBA, ...)
- Tạo ra báo cáo bằng các công cụ Chart có sẵn trong Excel hoặc Addins bên ngoài (ThinkCell)

Trong nhiều năm qua, một Workflow như vậy vẫn đáp ứng được một phần nhu cầu báo cáo trong doanh nghiệp, tuy nhiên không tránh khỏi những hạn chế khi có sự thay đổi. Với giải pháp Power BI, bạn có thể hình dung được rằng, toàn bộ Workflow trên có thể được thay thế hoàn toàn,

Hệ thống thông tin phục vụ trí tuệ kinh doanh

từ khâu nạp dữ liệu, xử lý dữ liệu, tính toán và đưa ra các chỉ số, trực quan hoá các loại dữ liệu, tạo ra biểu đồ, báo cáo và Dashboard. Đối với giải pháp này, khi có dữ liệu mới, để cập nhật báo cáo, bạn chỉ việc bấm nút Refresh, sau ít giây, bạn sẽ có được một bản báo cáo với dữ liệu cập nhật mới nhất.

- Khai thác dữ liệu thông qua các ngôn ngữ truy vấn tự nhiên
- Các câu hỏi về khả năng kết nối với cơ sở dữ liệu mở rộng được công bố trên website
- Khả năng chia sẻ, tích hợp thông qua nền tảng Web, ứng dụng di động

Bộ giải pháp Power BI của Microsoft mang tới khả năng chia sẻ báo cáo tương tác qua hệ thống Web qua chức năng Publish to Web. Với chức năng này, người đọc báo cáo không nhất thiết phải cài đặt Power BI trên thiết bị của mình, mà chỉ cần một trình duyệt Web là đã có thể dễ dàng truy cập được các báo cáo nhanh chóng và dễ dàng. Kết hợp với khả năng phân quyền xem báo cáo trong Power BI gói dành cho tổ chức doanh nghiệp, bộ giải pháp Power BI sẽ thay đổi hoàn toàn hệ thống báo cáo của bạn theo chiều hướng tốt hơn, nhanh hơn, đẹp hơn, tiện dụng hơn.

- Khi bạn tải các dữ liệu vào dịch vụ web của Power BI, bạn có thể thiết lập lịch trình để cập nhật dữ liệu, mà không cần phải can thiệp thủ công.
- Giao diện Ribbon, Task Pane thân thiện dễ sử dụng

Cùng được Microsoft phát triển theo ngôn ngữ thiết kế của bộ Office đã rất quen thuộc với người dùng, chúng ta không mất nhiều thời gian để có

Hệ thống thông tin phục vụ trí tuệ kinh doanh

thể làm quen với giao diện của Power BI, làm giảm đáng kể thời gian học và làm việc với công cụ này, giúp năng suất làm việc tăng cao hơn, tạo ra nhiều giá trị hơn, trong thời gian ngắn hơn.

- **Đưa ra các bảng dashboard được cập nhật real-time.**
- **Tích hợp sâu rộng với nhiều công cụ doanh nghiệp**

Với định vị là một giải pháp linh hoạt, mạnh mẽ dành cho doanh nghiệp, Microsoft Power BI tích hợp chặt chẽ với các sản phẩm không chỉ từ Microsoft như SharePoint, Office 365, Dynamic CRM và cả những sản phẩm từ các hãng khác như Spark, Hadoop, Google Analytics, SAP, Salesforce, MailChimp, ...

- **Khả năng làm việc với hầu hết các nguồn dữ liệu thông dụng với Power Query**

Công cụ Power Query tích hợp trực tiếp trong Power BI Desktop đã tích hợp sẵn các connector (cổng kết nối) tới các nguồn dữ liệu phổ biến như:

- Kết nối tới tài liệu Excel (xls,xlsx,xlsb,xlsm, ...)
- Kết nối với các loại flat files (CSV, TSV, TXT, ...)
- Kết nối với các files PDF “generated”
- Kết nối với các hệ cơ sở dữ liệu: SQL Server, MySQL, IBM DB2, Oracle, ...
- Kết nối với các API dữ liệu chứng khoán, kinh tế, tài chính, ...
- Kết nối tới các trang Web và lấy dữ liệu
- Và nhiều kết nối khác ...

Hệ thống thông tin phục vụ trí tuệ kinh doanh

- Triển khai nhanh, đa cấu hình, và có môi trường bảo mật
- Khả năng trực quan hoá dữ liệu vượt trội

80% thời gian của bạn sẽ dành cho việc chuẩn bị kết quả cuối cùng là những biểu đồ (trong Power BI gọi là Visuals). Biểu đồ trong Power BI dựa trên nền tảng D3js nổi tiếng giúp cho việc biến dữ liệu thành các loại biểu đồ phong phú, đa dạng và dễ dàng. Chưa thấy thuyết phục bởi các biểu đồ sẵn có của Power BI? Không vấn đề gì bởi vì khả năng mở rộng dưới dạng Custom Visuals của Power BI sẽ cho phép bạn tự tay thiết kế và tạo ra những loại biểu đồ chỉ phục vụ cho riêng bạn và tổ chức của bạn. Có sẵn biểu đồ từ Python hoặc R? Power BI sẽ hỗ trợ bạn trong việc sử dụng các biểu đồ, Charts từ Python và R một cách đơn giản và dễ dàng.

- Việc kết nối đến nguồn dữ liệu qua đám mây hoặc ngay trên hệ thống của doanh nghiệp, được bảo mật và đáng tin cậy.
- Power BI được tạo nên từ nền tảng của AI (trí tuệ thông minh) và ML (máy học)

1.4. Nhược điểm của PowerBI

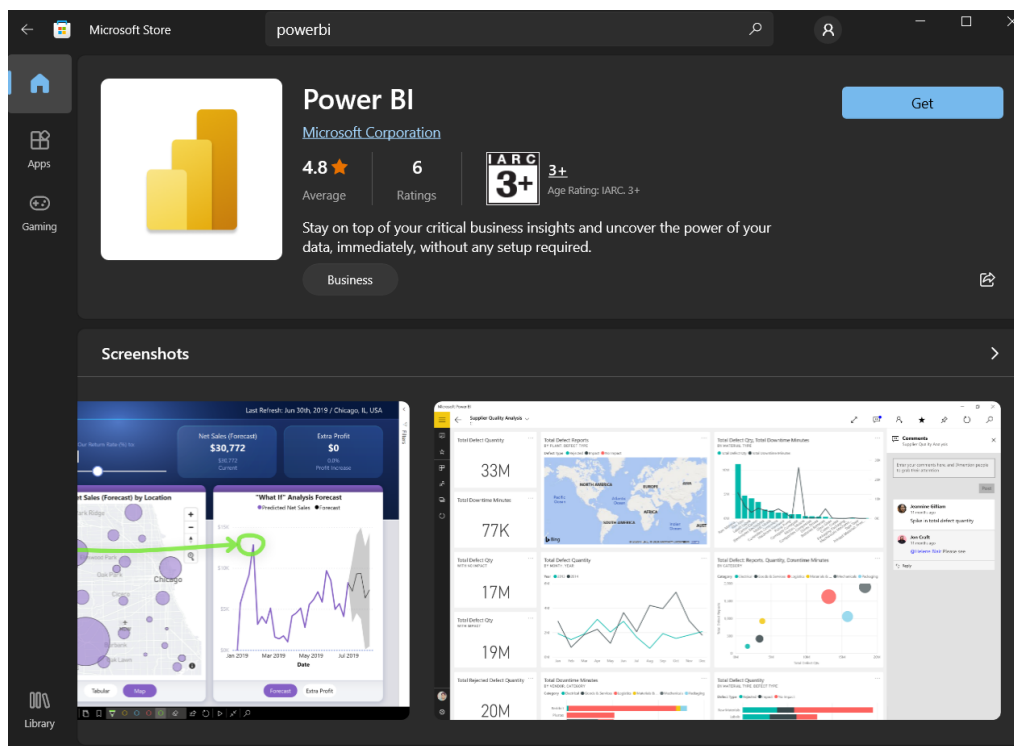
- Dashboard và báo cáo chỉ có thể được chia sẻ với những người dùng có cùng tên miền email.
- Power BI sẽ không kết hợp các dữ liệu được nhập vào từ những kết nối theo thời gian thực.
- Power BI không chấp nhận những file lớn hơn 250MB và các file được nén bởi các dữ liệu X trong các cơ sở dữ liệu trong bộ nhớ.

Hệ thống thông tin phục vụ trí tuệ kinh doanh

- Power BI không chấp nhận những file lớn hơn 1GB.
- Các dashboard này sẽ không thể chia sẻ cho người dùng, tài khoản khác hoặc các thông số thực tế khác.
- Có rất ít nguồn dữ liệu cho phép kết nối real-time với các báo cáo và dashboard trên Power BI.

1.5. Cài đặt PowerBI Desktop

- Cài đặt trực tiếp từ Microsoft Store



- Cài đặt qua trang chủ

<https://powerbi.microsoft.com/en-us/desktop/>

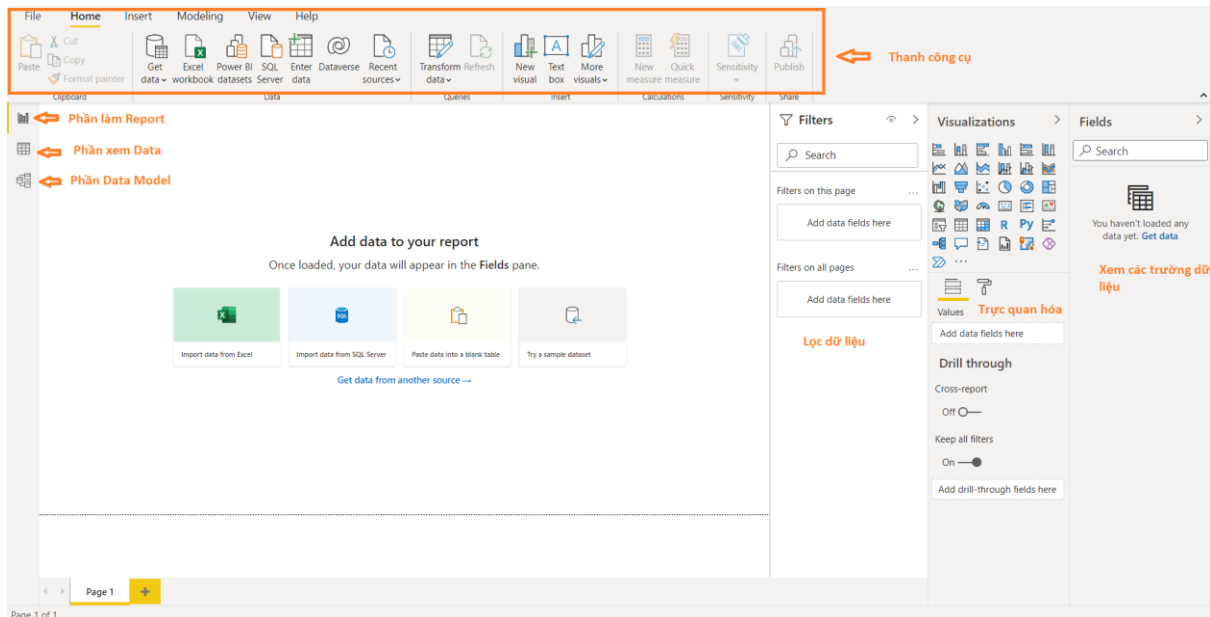
Hệ thống thông tin phục vụ trí tuệ kinh doanh

1.6. Data mẫu

Dữ liệu mẫu AdventureWorksDW2019 được lấy từ trang chủ của Microsoft.

[AdventureWorks sample databases - SQL Server | Microsoft Docs](#)

2.KẾT NỐI DỮ LIỆU VỚI POWER BI



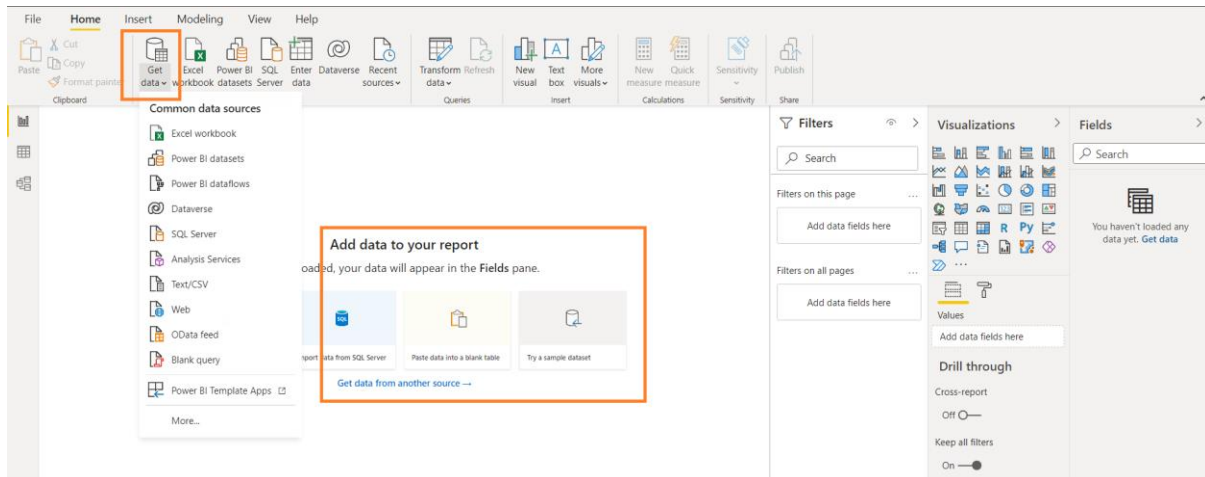
Hình 2.1 Giao diện Power BI

Power BI đã có thể kết nối với hơn 110 nguồn dữ liệu và kiểu kết nối khác nhau:

- File (Excel, Text/CSV, XML, JSON, PDF).
- Database (SQL Server, Oracle, IBM DB2, MySQL, PostgreSQL, Snowflake, etc).
- Power Platform (Power BI datasets, Power BI Dataflows, Common Data Services)
- Azure (SQL Database, Synapse Server, Analysis Services, Blob Storage, Data Lake, Cosmos DB, etc).
- Online Services (SharePoint, Dynamics 365, Dynamics Business Central, Azure DevOps, Salesforce, Google Analytics, Facebook, GitHub, etc).
- Other (Web, OData, ODBC, Hadoop, Spark, R script, Python script, etc).

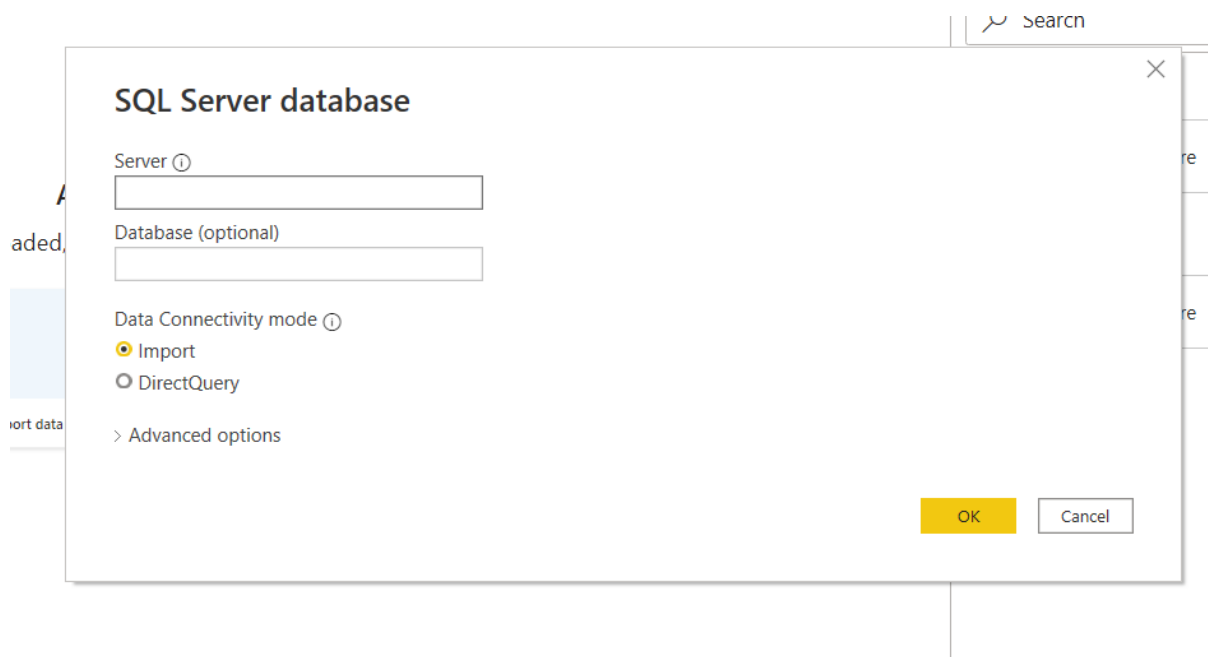
Hệ thống thông tin phục vụ trí tuệ kinh doanh

Để kết nối dữ liệu chúng ta chọn **Get data** từ thanh công cụ hoặc chọn từ **Add data to your report** từ phần màn hình Report chính của Power BI.



Hình 2.2 Các cách kết nối nguồn dữ liệu với Power BI

Trong bài viết này sẽ sử dụng SQL Server với Data AdventureWorksDW2019. Sau khi nhập tên Server name và Data chọn OK để tiến hành kết nối.

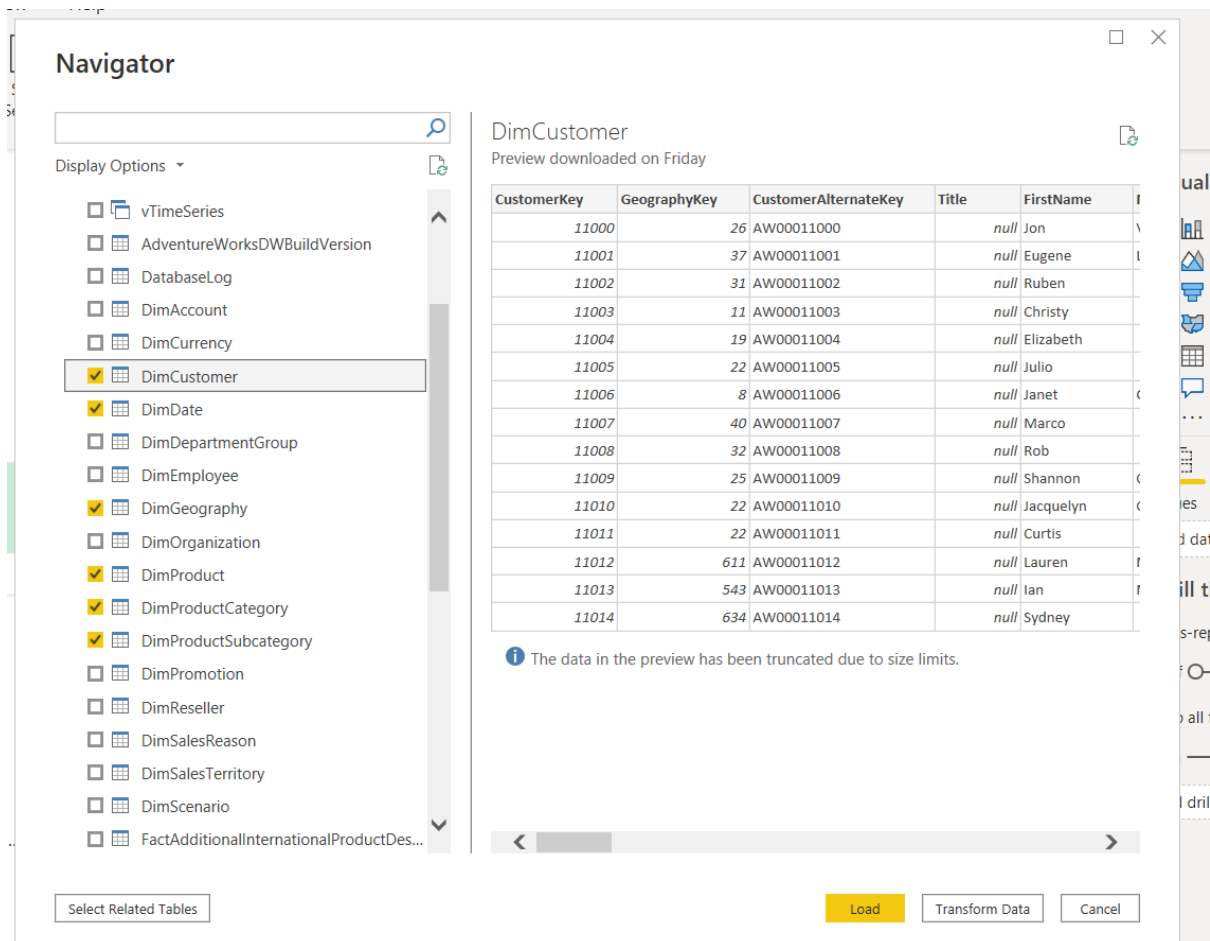


Hình 2.3 Nhập tên Server name và Data muốn kết nối từ SQL Server với Power BI

Hệ thống thông tin phục vụ trí tuệ kinh doanh

Sau khi chọn OK sẽ xuất hiện màn hình **Navigator** trong màn hình này sẽ xuất hiện tất cả các bảng dữ liệu có trong data

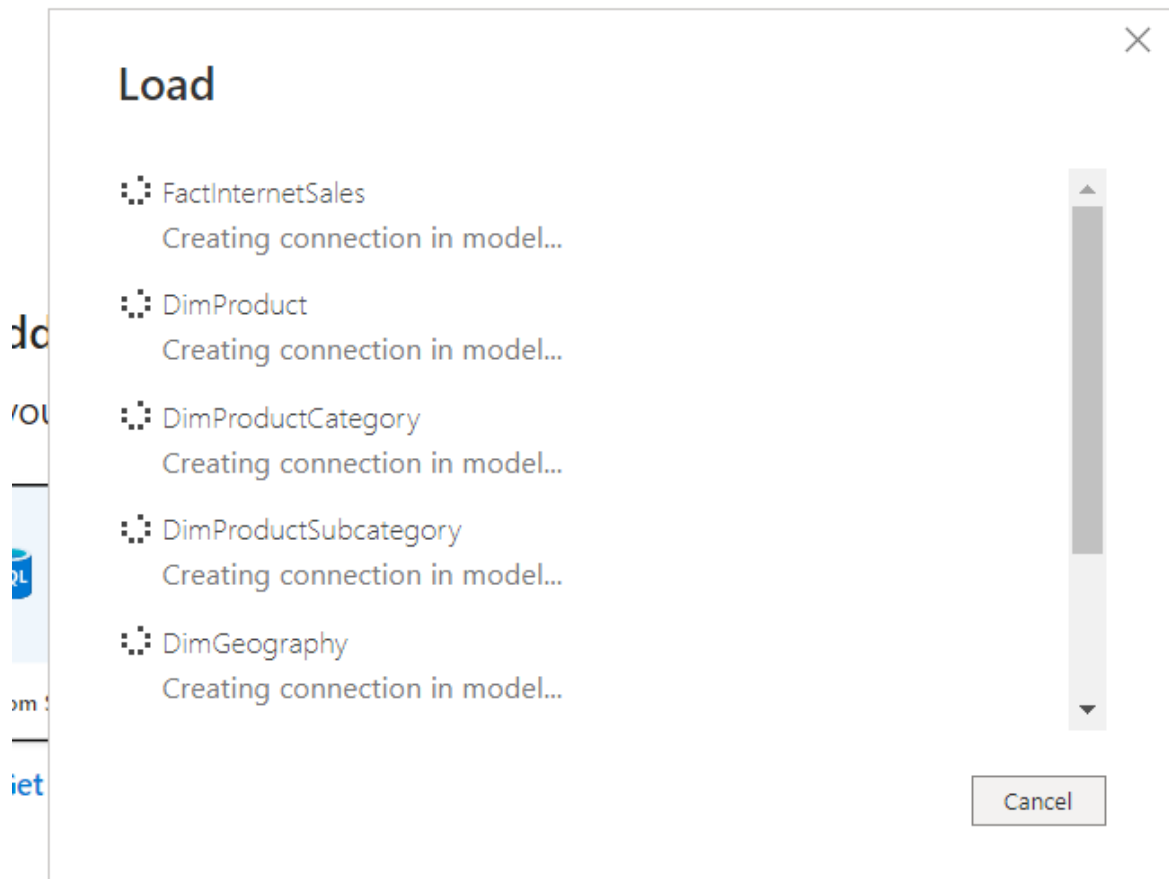
AdventureWorksDW2019. Trong bài viết này sẽ sử dụng các bảng DimDate, DimCustomer, DimGeography, DimProduct, DimProductCategory, DimProductSubCategory, FactInternetSales. Sau khi chọn các bảng cần sử dụng chọn **Load** để tiến hành Load dữ liệu vào Power BI.



Hình 2.4 Màn hình Load data vào Power BI

Chờ quá trình Load dữ liệu

Hệ thống thông tin phục vụ trí tuệ kinh doanh



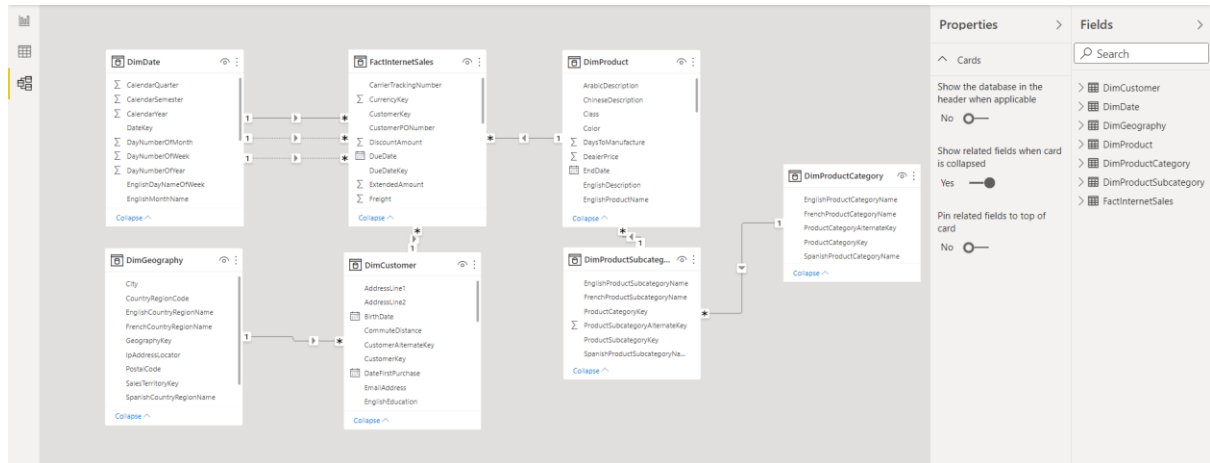
Hình 2.5 Quá trình Load dữ liệu

Page Data là nơi để quan sát dữ liệu và các bảng bạn hiện có.

CustomerKey	GeographyKey	CustomerAlternateKey	Title	FirstName	MiddleName	LastName	NameStyle	BirthDate	MaritalStatus	Suffix	Gender	EmailAddress
11602	135	AW00011602		Larry		Gill	False	Monday, October 11, 1982	S		M	larry16@adventure-works.com
11603	244	AW00011603		Geoffrey		Gonzalez	False	Friday, August 6, 1982	S		M	geoffrey16@adventure-works.com
11610	269	AW00011610		Blake		Collins	False	Monday, October 20, 1980	S		M	blake47@adventure-works.com
12517	133	AW00012517		Alexa		Watson	False	Tuesday, February 22, 1983	S		F	alexa0@adventure-works.com
12518	161	AW00012518		Jacquelyn		Dominguez	False	Sunday, March 27, 1983	S		F	jacquelyn13@adventure-works.com
12519	265	AW00012519		Casey		Gutierrez	False	Thursday, June 16, 1983	S		M	casey34@adventure-works.com
12714	157	AW00012714		Colleen		Lu	False	Sunday, January 14, 1979	S		F	colleen11@adventure-works.com
12728	131	AW00012728		Jeremiah		Stewart	False	Sunday, December 23, 1984	S		M	jeremiah44@adventure-works.com
12871	233	AW00012871		Leah		Li	False	Monday, April 5, 1982	S		F	leah2@adventure-works.com
13671	173	AW00013671		Frank		Ramos	False	Tuesday, August 7, 1979	S		M	frank25@adventure-works.com
13826	127	AW00013826		Candice		He	False	Wednesday, May 25, 1983	S		F	candice2@adventure-works.com
13830	237	AW00013830		Andrea		Cox	False	Monday, January 31, 1983	S		F	andrea11@adventure-works.com
13838	263	AW00013838		Jill		Rubio	False	Friday, December 25, 1981	S		F	jill29@adventure-works.com
14838	121	AW00014838		Darren		Alvarez	False	Sunday, January 23, 1983	S		M	darren26@adventure-works.com
14839	238	AW00014839		Natasha		Sanz	False	Monday, November 15, 1982	S		F	natasha20@adventure-works.com
14840	277	AW00014840		Autumn		Zhu	False	Friday, April 22, 1983	S		F	autumn13@adventure-works.com
14848	186	AW00014848		George		Louverdis	False	Monday, March 9, 1981	S		M	george5@adventure-works.com
14849	142	AW00014849		Dwayne		Martin	False	Friday, April 10, 1981	S		M	dwayne0@adventure-works.com
14991	272	AW00014991		Edwin		Zhao	False	Tuesday, August 25, 1981	S		M	edwin11@adventure-works.com
16317	263	AW00016317		Joel		Raman	False	Wednesday, March 23, 1983	S		M	joel11@adventure-works.com
16452	232	AW00016452		Cassie		Luo	False	Saturday, May 21, 1983	S		F	cassie4@adventure-works.com
16453	271	AW00016453		Mayra		Chandira	False	Thursday, October 7, 1982	S		F	mayra1@adventure-works.com
16458	192	AW00016458		Michelle		Ruiz	False	Friday, June 18, 1982	S		F	michelle37@adventure-works.com
16460	277	AW00016460		Jodi		Xu	False	Monday, May 17, 1982	S		F	jod5@adventure-works.com
16461	279	AW00016461		Rosa		Wu	False	Wednesday, August 5, 1981	S		F	rosa7@adventure-works.com
18038	207	AW00018038		Kelli		Rumar	False	Friday, November 14, 1980	S		F	kelli30@adventure-works.com
18048	229	AW00018048		Alisha		Sun	False	Friday, April 4, 1980	S		F	alisha23@adventure-works.com

Hình 2.6 Page Data

Hệ thống thông tin phục vụ trí tuệ kinh doanh

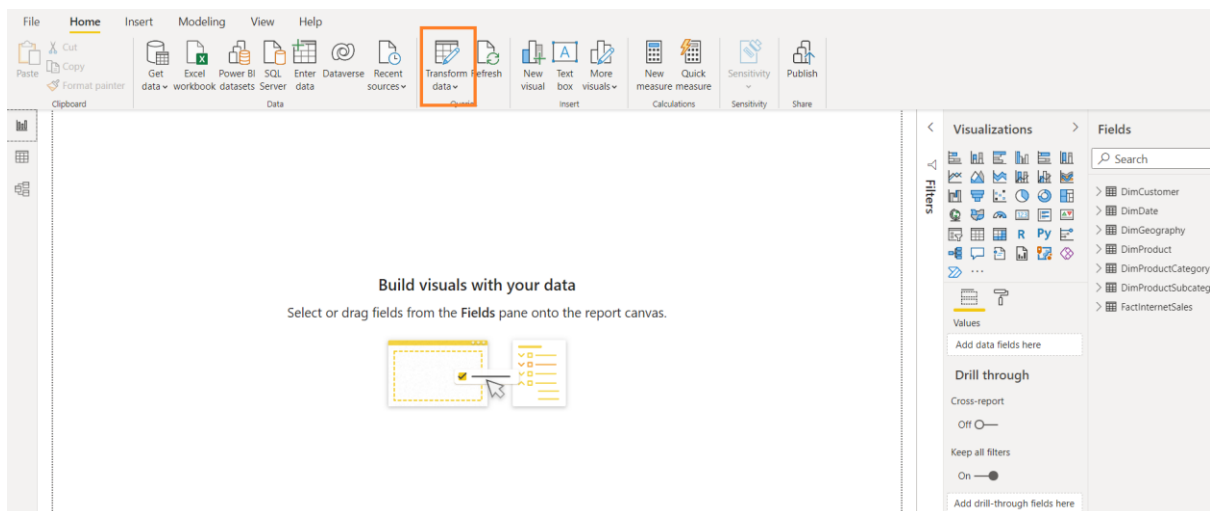


Hình 2.7 Page Data Model

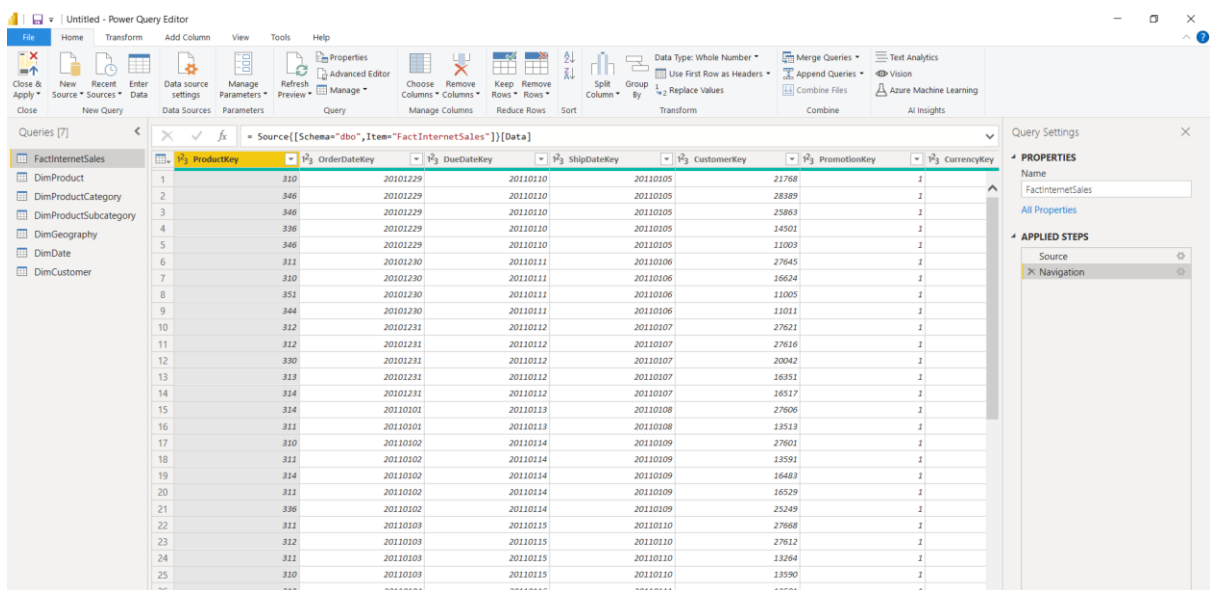
3. CHUẨN BỊ DỮ LIỆU - DATA PREPARATION

Sau khi đã có data trong Power BI chúng ta không thể vào phân tích vẽ biểu đồ liền được, vì không phải data lúc nào cũng đẹp, cũng sạch. Nên trước khi phân tích và vẽ biểu đồ ta cần làm sạch và biến đổi dữ liệu.

Để biến đổi dữ liệu chọn Transform data từ thanh công cụ, sau đó màn hình **Power Query Editor** sẽ xuất hiện.



Hình 3.1 Công cụ biến đổi dữ liệu



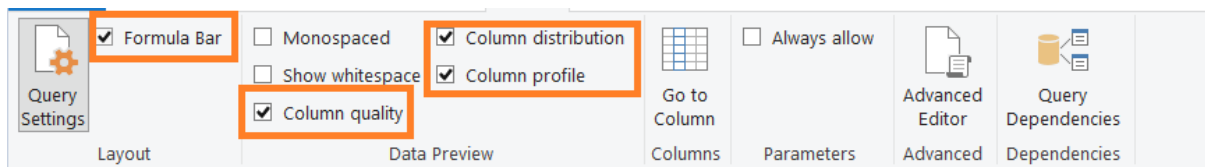
Hình 3.2 Power Query Editor

Hệ thống thông tin phục vụ trí tuệ kinh doanh

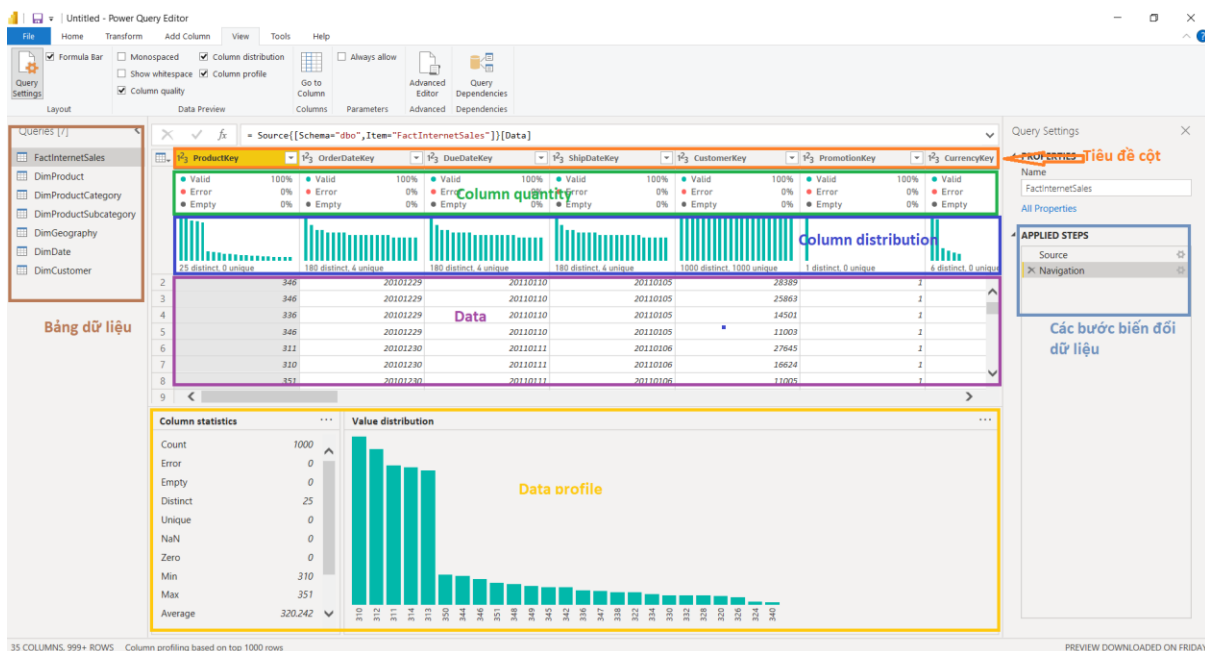
Bước đầu cho việc chuẩn bị dữ liệu là chúng ta cần hiểu hiện tại chúng ta đang có gì, bộ dữ liệu ra làm sao.

Trong công cụ **Power Query Editor** tại tab **View** click vào các tùy chọn:

- Formula Bar
- Column quality
- Column distribution



Hình 3.3 Power BI Desktop application interface



Hình 3.4 Bố cục dữ liệu trong Power Query Editor

- **Tiêu đề cột** : Hiển thị tên cột và kiểu dữ liệu.

Column quality : Hiển thị các giá trị valid (hợp lệ), empty (mất dữ liệu), error (lỗi).

- **Column distribution** : Hiển thị các giá trị khác biệt (distinct) và duy nhất (unique).
- **Các bước biến đổi dữ liệu**: hiển thị danh sách các phép chuyển đổi dữ liệu đã được áp dụng cho dữ liệu.

Hệ thống thông tin phục vụ trí tuệ kinh doanh

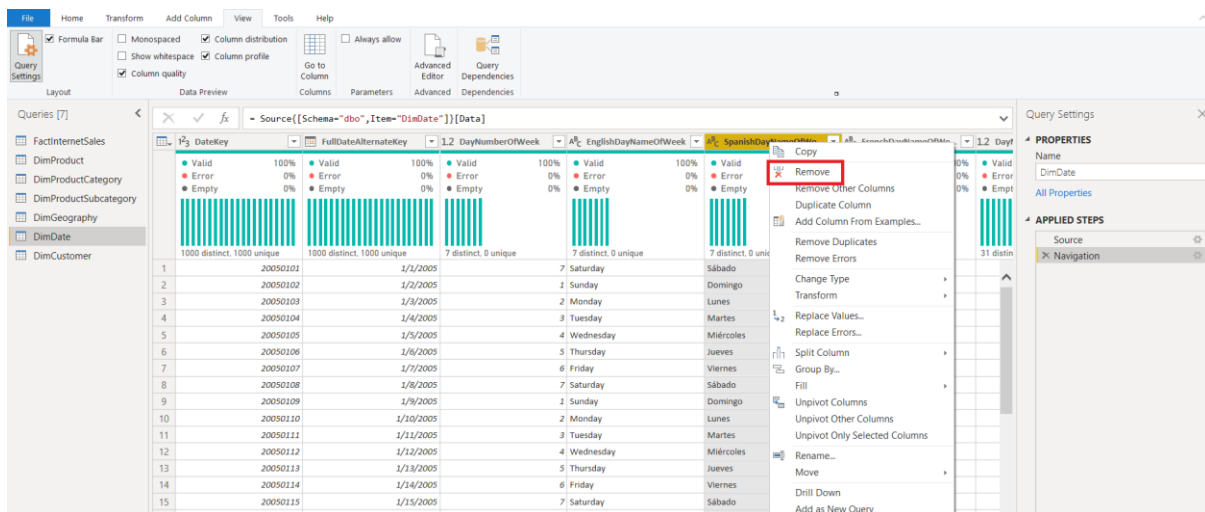
- Data: Hiển thị dữ liệu

3.1. Bảng DimDate

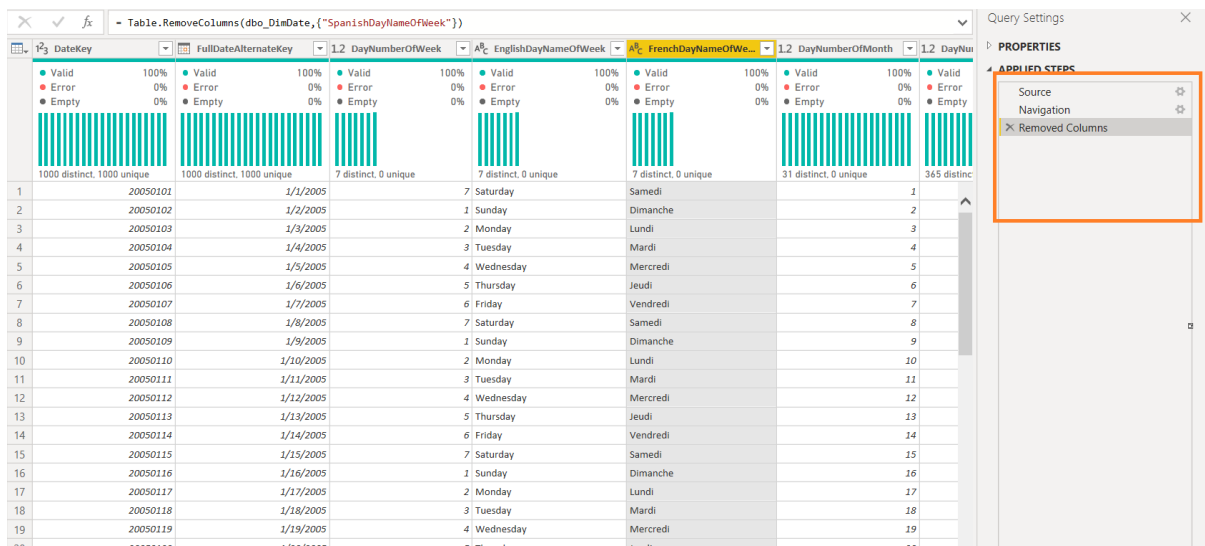
Xóa các cột dữ liệu - Removing Columns

Có những cột dữ liệu không sử dụng chúng ta có thể xóa nó đi, hoặc các cột dữ liệu bị missing value quá nhiều cũng có thể xóa cột đó đi vì dữ liệu bị mất quá nhiều không có giá trị cho phân tích.

Để xóa một cột dữ liệu ta nhấp chuột phải vào cột dữ liệu cần xóa và chọn **Remove**.



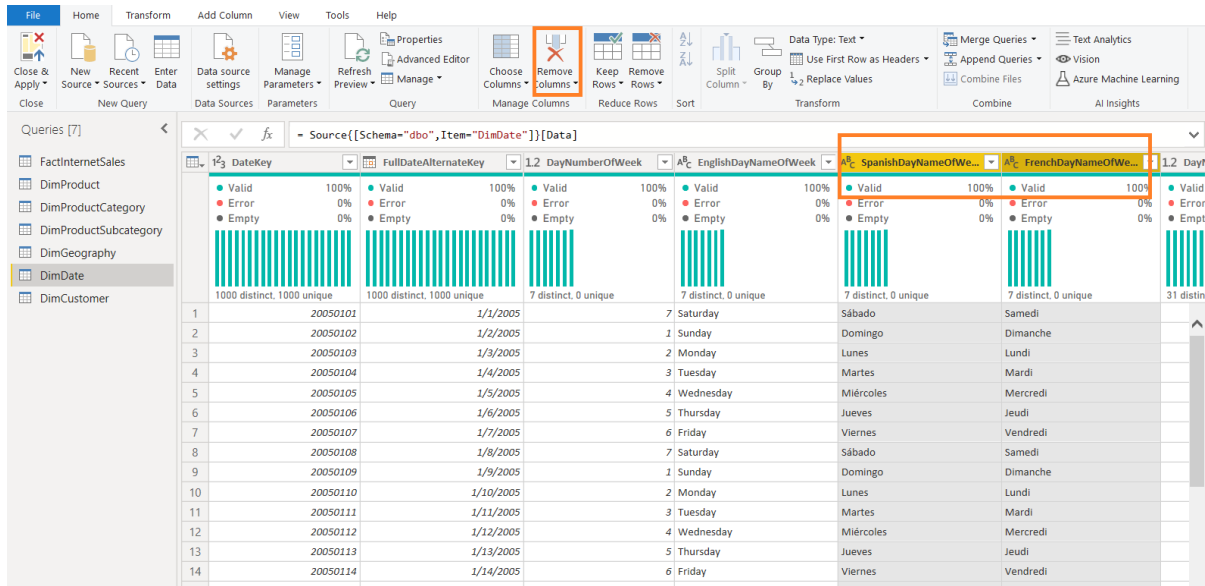
Hình 3.1.1 Xóa một cột dữ liệu - SpanishDayNameOfWeek



Hình 3.1.2 Lịch sử cột đã xóa được ghi lại

Hệ thống thông tin phục vụ trí tuệ kinh doanh

Để xóa nhiều cột dữ liệu ta giữ phím **Ctrl** + **cột muốn xóa** sau đó chọn **Remove Columns** trên thanh công cụ.



Hình 3.1.3 Xóa nhiều cột dữ liệu

Tương tự ta xóa các cột còn lại không sử dụng như:

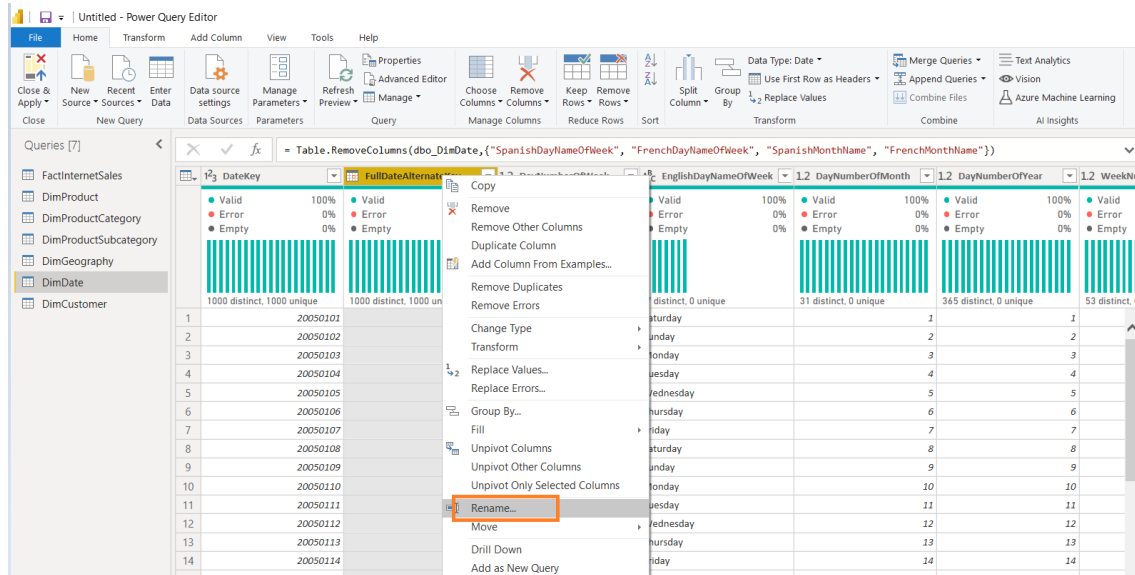
SpanishMonthName, FrenchMonthName, CalendarSemester, FiscalQuarter, FiscalYear, FiscalSemester, FactCallCenter, FactFinance, FactCurrencyRate, FactSurveyResponse, FactSalesQuota, FactResellerSales(DateKey) 3, FactResellerSales(DateKey) 2, FactResellerSales(DateKey), FactProductInventory.

Đổi tên cột dữ liệu - Renaming Columns

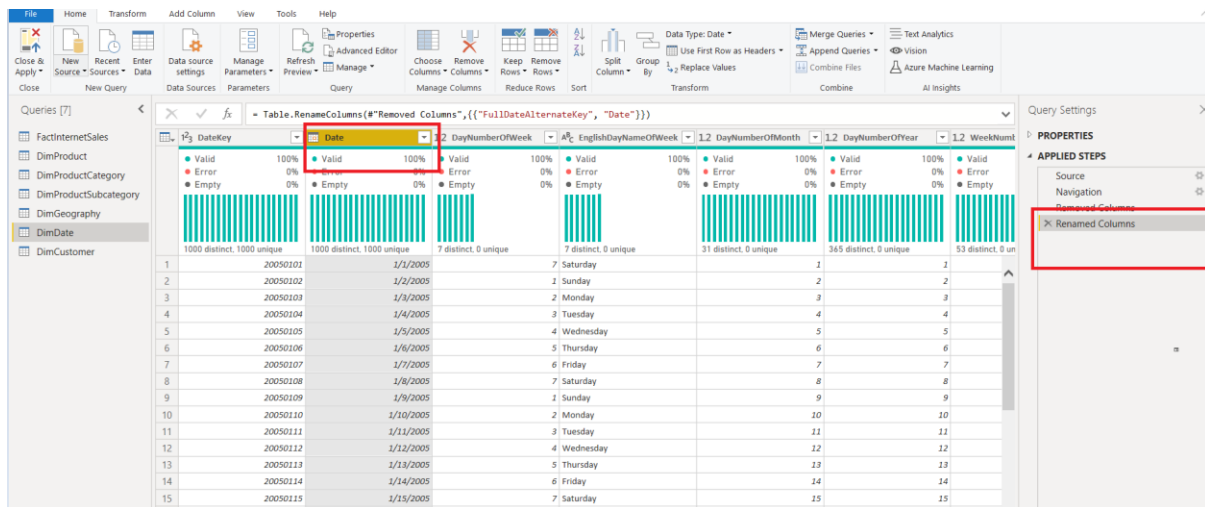
Tên cột dữ liệu có khi quá dài và ý nghĩa chưa hợp lệ hoặc thuật ngữ chúng ta không quen có thể tiến hành đổi tên cho dễ sử dụng hơn.

Để đổi tên cột ta nhấp đôi chuột vào cột muốn đổi sau đó đổi tên cột và nhấn Enter hoặc chuột phải vào cột muốn đổi và chọn **Rename**, đổi tên cột và Enter.

Hệ thống thông tin phục vụ trí tuệ kinh doanh



Hình 3.1.4 Đổi tên cột FullDateAlternateKey



Hình 3.1.5 Cột FullDateAlternateKey










sau khi đã đổi tên thành Date
và bước đổi tên đã được ghi lại.

Tương tự ta đổi tên cho các cột còn lại: **EnglishDayNameOfWeek** thành **Day**, **EnglishMonthName** thành **Month**, **CalendarQuarter** thành **Quarter**, **CalendarYear** thành **Year**.

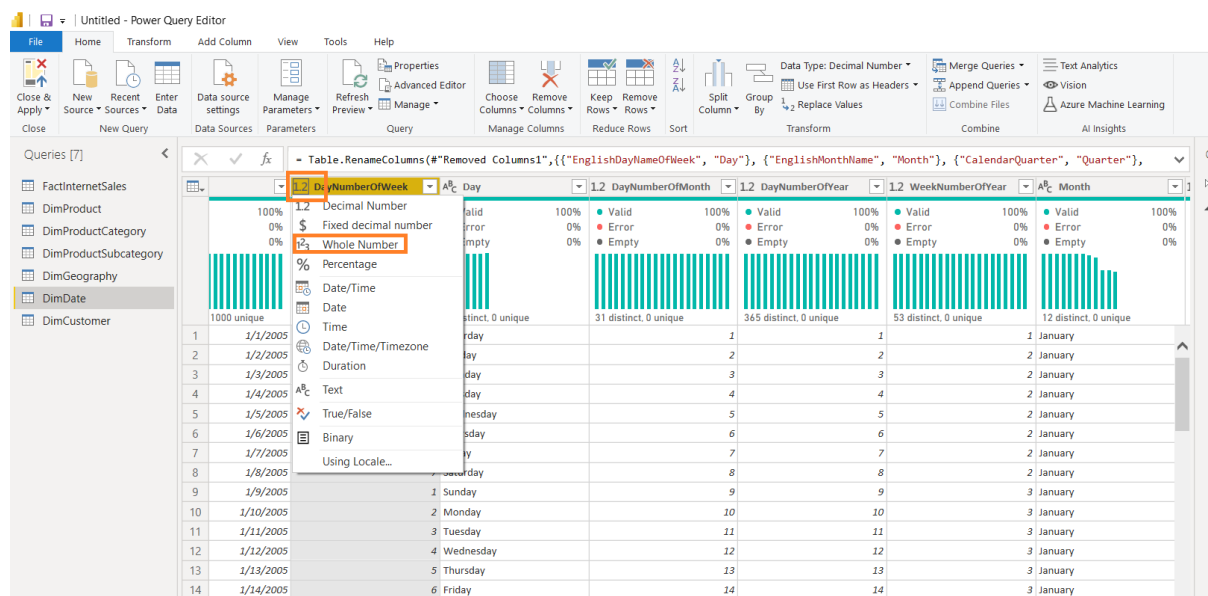
Thay đổi kiểu dữ liệu - Changing Data Types

Trong nhiều trường hợp các dữ liệu được đưa vào là chính xác, có loại dữ liệu không, hoặc các dạng định dạng dữ liệu ở các bảng là khác nhau. Nên ta cần thay đổi kiểu dữ liệu cho phù hợp.

Hệ thống thông tin phục vụ trí tuệ kinh doanh

-  Text - Kiểu Text
-  Kiểu True/ False
-  Whole Numbers - Kiểu số nguyên
-  Decimal Numbers - Kiểu số thập phân
-  Fixed Decimal Number - Kiểu số thập phân cố định
-  Percentage - Phần trăm
-  Date - ex: 15/02/2020
-  Time - ex: 12:35:00 PM
-  Date/Time - ex: 15/02/2020 10:25:00 AM

Để thay đổi kiểu dữ liệu ta chọn vài icon trước tên cột sau đó chọn kiểu dữ liệu muốn chuyển.



Hình 3.1.6 Chuyển đổi kiểu dữ liệu cột DayNumberOfWeek từ số thập phân sang số nguyên

Hệ thống thông tin phục vụ trí tuệ kinh doanh


Tương tự cho các cột còn lại: **DayNumberOfMonth** từ số thập phân sang số nguyên, **DayNumberOfYear** từ số thập phân sang số nguyên, **WeekNumberOfYear** từ số thập phân sang số nguyên.

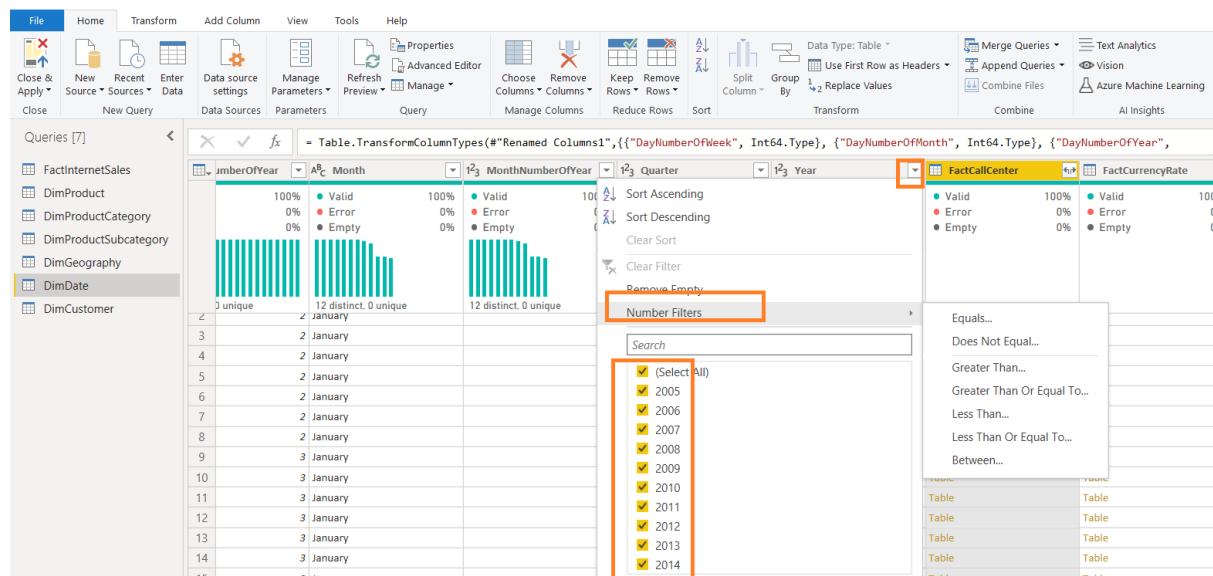
MonthNumberOfYear từ số thập phân sang số nguyên, **Quarter** từ số thập phân sang số nguyên, **Year** từ số thập phân sang số nguyên.

Lọc dữ liệu - Filters

Đôi khi dữ liệu không phải lúc nào ta cũng lấy hết để phân tích. Ví dụ như năm như năm chỉ cần thông tin những năm gần nhất ta có thể lấy 3 hoặc 4 năm gần nhất phân tích.

Lọc ra các năm phân tích 2012, 2013, 2014.

Chọn cột dữ liệu cần để các cột khác cũng lọc theo, bằng cách ấn vào icon  sau tên cột, có hai lựa chọn có thể chọn bỏ tích các ô vàng để lọc ra các dữ liệu không cần hoặc sử dụng **Number Filters** để lọc.



Hình 3.1.7 Lọc dữ liệu

Nếu chọn cách **Number Filters** để lọc ra dữ liệu 2012, 2013, 2014 thì ta chọn **Greater Than or Equal To..**

Tại ô **Enter or select a value** nhập 2012 sau đó nhấn **OK**.

Hệ thống thông tin phục vụ trí tuệ kinh doanh

×

Filter Rows

Apply one or more filter conditions to the rows in this table.

☒ Basic ☐ Advanced

Keep rows where 'Year'

is greater than or equal to ▾

Enter or select a value ▾

☒ And ☐ Or

▾

Enter or select a value ▾

OK

Cancel

Hình 3.1.8 Điều kiện lọc

The screenshot displays the Microsoft Power BI Desktop interface. On the left, the 'Queries' pane shows a list of tables: FactInternetSales, DimProduct, DimProductCategory, DimProductSubcategory, DimGeography, DimDate, and DimCustomer. The main area shows a data table with columns: FactInternetSales, DimProduct, DimProductCategory, DimProductSubcategory, DimGeography, DimDate, and DimCustomer. A context menu is open over the DimDate column, displaying options like 'Sort Ascending', 'Sort Descending', 'Clear Sort', 'Clear Filter', 'Remove Empty', and 'Number Filters'. The 'Number Filters' option is selected, and a sub-menu is visible showing a search bar and a list of years (2012, 2013, 2014) with checkboxes. The table data shows a list of dates from 1/1/2012 to 1/31/2014, with corresponding FactInternetSales values.

Hình 3.1.9 Dữ liệu sau khi lọc

3.2. Bảng Customer

Xóa các cột dữ liệu - Removing Columns

Các cột cần xóa là: FactSurveyResponse, Title, NameStyle, MaritalStatus, Suffix, CommuteDistance, DateFirstPurchase, AddressLine2, NumberCarsOwned, HouseOwnerFlag,

Hệ thống thông tin phục vụ trí tuệ kinh doanh

SpanishOccupation, FrenchOccupation, FrenchEducation, SpanishEducation, TotalChildren, NumberChildrenAtHome, Phone.

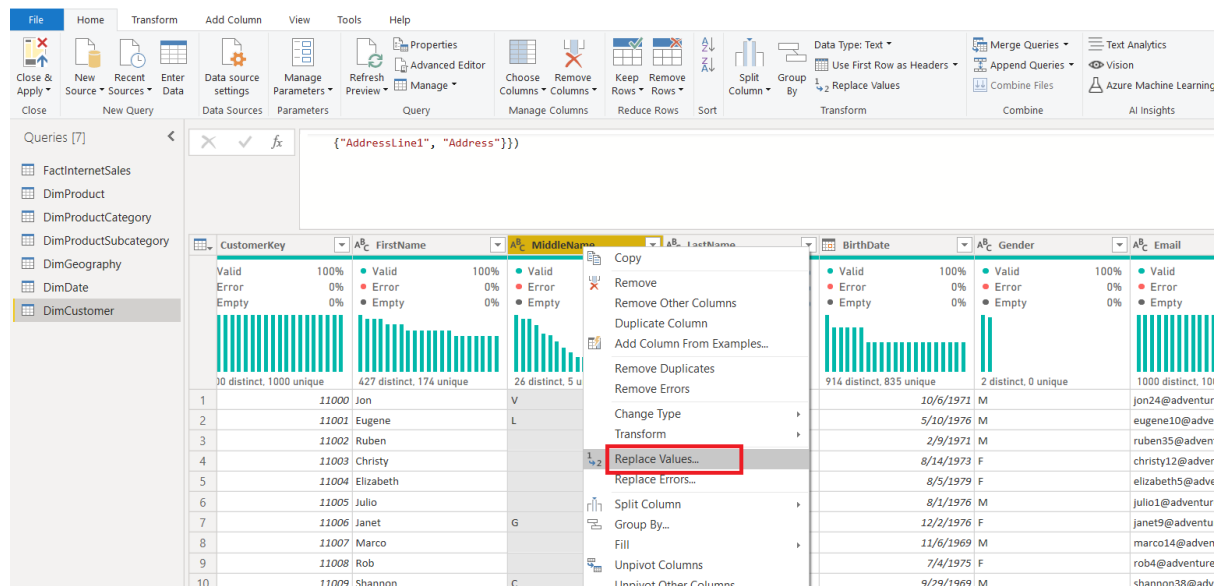
Đổi tên cột dữ liệu - Renaming Columns

Các cột cần đổi tên là: **EmailAddress** thành **Email**, **EnglishEducation** thành **Education**, **EnglishOccupation** thành **Occupation**, **AddressLine1** thành **Address**.

Thay đổi giá trị

Các dữ liệu đôi khi sẽ bị missing value ta có thể dùng giá trị khác để thay thế cho các giá trị bị mất, đối với các cột bị missing value quá nhiều thì ta có thể chọn xóa các cột đó hoàn toàn. Hoặc dữ liệu có thể là viết tắt, chúng ta có thể chỉnh chúng thành một giá trị ghi đầy đủ hơn.

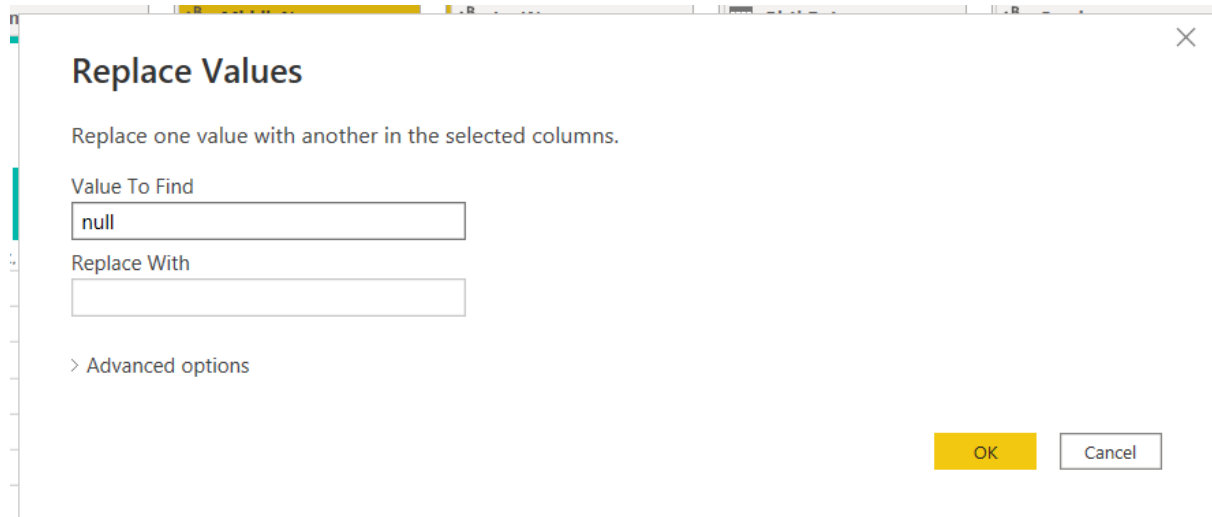
Chuột phải vào cột cần thay đổi giá trị chọn Replace value.



Hình 3.2.1 Replace value

Thay thế giá trị bị missing value cụ thể ở đây là null thành khoảng trắng sau đó chọn **OK**.

Hệ thống thông tin phục vụ trí tuệ kinh doanh



Hình 3.2.2 Thay thế giá trị null bằng khoảng trống

Tương tự ta thay thế cho cột **Gender** đổi giá trị **F** thành **Female**, **M** thành **Male**.

3.3. Bảng DimGeography

Xóa các cột dữ liệu - Removing Columns

Các cột cần xóa là: DimSalesTerritory, DimReseller, SpanishCountryRegionName, FrenchCountryRegionName.

Đổi tên cột dữ liệu - Renaming Columns

Các cột cần đổi tên là: **StateProvinceName** thành **StateProvince**, **EnglishCountryRegionName** thành **CountryRegion**.

3.4. Bảng DimProductCategory

Xóa các cột dữ liệu - Removing Columns

Các cột cần xóa là: SpanishProductCategoryName, FrenchProductCategoryName.

Hệ thống thông tin phục vụ trí tuệ kinh doanh

Đổi tên cột dữ liệu - Renaming Columns

Các cột cần đổi tên là: **EnglishProductCategoryName** thành **Product Category**

3.5. Bảng DimSubProductCategory

Xóa các cột dữ liệu - Removing Columns

Các cột cần xóa là: **FrenchProductSubcategoryName**,
SpanishProductSubcategoryName.

Đổi tên cột dữ liệu - Renaming Columns

Các cột cần đổi tên là: **EnglishProductSubcategoryName** thành **Product Subcategory**

3.6. Bảng Product

Xóa các cột dữ liệu - Removing Columns

Các cột cần xóa là: **FactResellerSales**, **FactProductInventory**, **Status**,
EndDate, **StartDate**, **TurkishDescription**, **JapaneseDescription**,
GermanDescription, **ThaiDescription**, **HebrewDescription**,
ArabicDescription, **ChineseDescription**, **FrenchDescription**,
WeightUnitMeasureCode, **SizeUnitMeasureCode**,
SpanishProductName, **FrenchProductName**, **StandardCost**,
FinishedGoodsFlag, **SafetyStockLevel**, **ReorderPoint**, **ListPrice**, **Size**,
Weight, **DaysToManufacture**, **SizeRange**, **DealerPrice**, **Style**, **Class**,
LargePhoto

Hệ thống thông tin phục vụ trí tuệ kinh doanh


Đổi tên cột dữ liệu - Renaming Columns

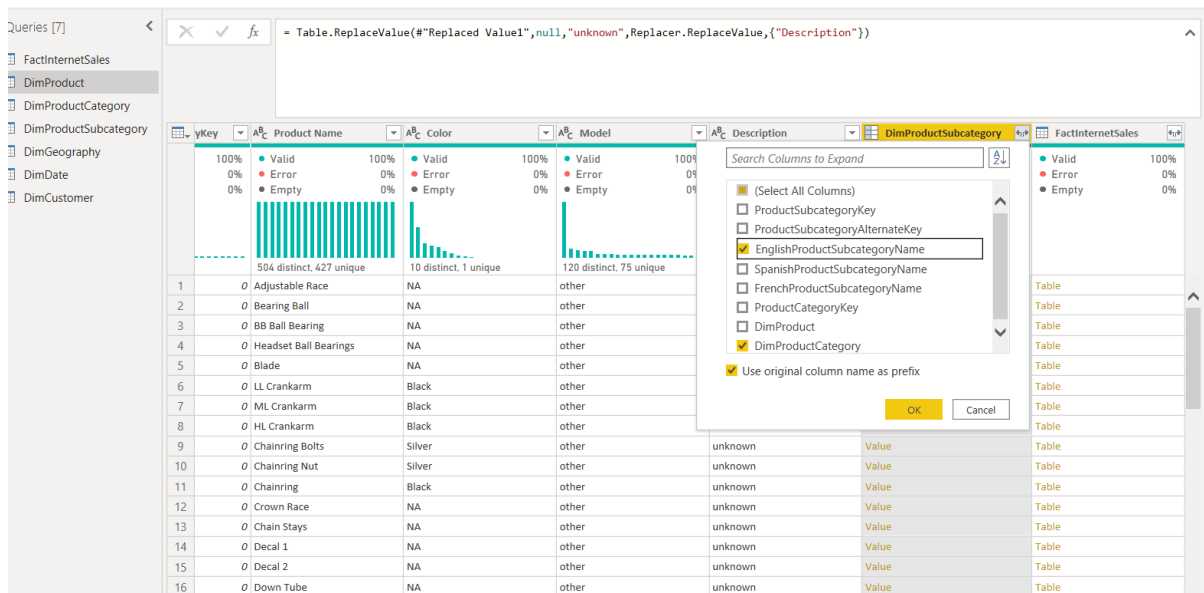
Các cột cần đổi tên là: **EnglishDescription** thành **Description**, **ModelName** thành **Model**, **EnglishProductName** thành **Product Name**.

Thay đổi giá trị

Các cột cần thay đổi giá trị: **ProductSubcategoryKey** đổi các giá null thành 0, **Model** đổi null thành other, **Description** đổi null thành unknown.

Lấy một cột từ bảng khác qua

Để lấy một cột từ bảng khác qua ta sử dụng icon  sau tên cột. Sau đó chọn tên cột muốn lấy và bảng liên kết.



Queries [7]

FactInternetSales

DimProduct

DimProductCategory

DimProductSubcategory

DimGeography

DimDate

DimCustomer

fx = Table.ReplaceValue("#Replaced Value1",null,"unknown",Replacer.ReplaceValue,{"Description"})

Search Columns to Expand

(Select All Columns)

☐ ProductSubcategoryKey

☐ ProductSubcategoryAlternateKey

☒ EnglishProductSubcategoryName

☐ SpanishProductSubcategoryName

☐ FrenchProductSubcategoryName

☐ ProductCategoryKey

☐ DimProduct

☒ DimProductCategory

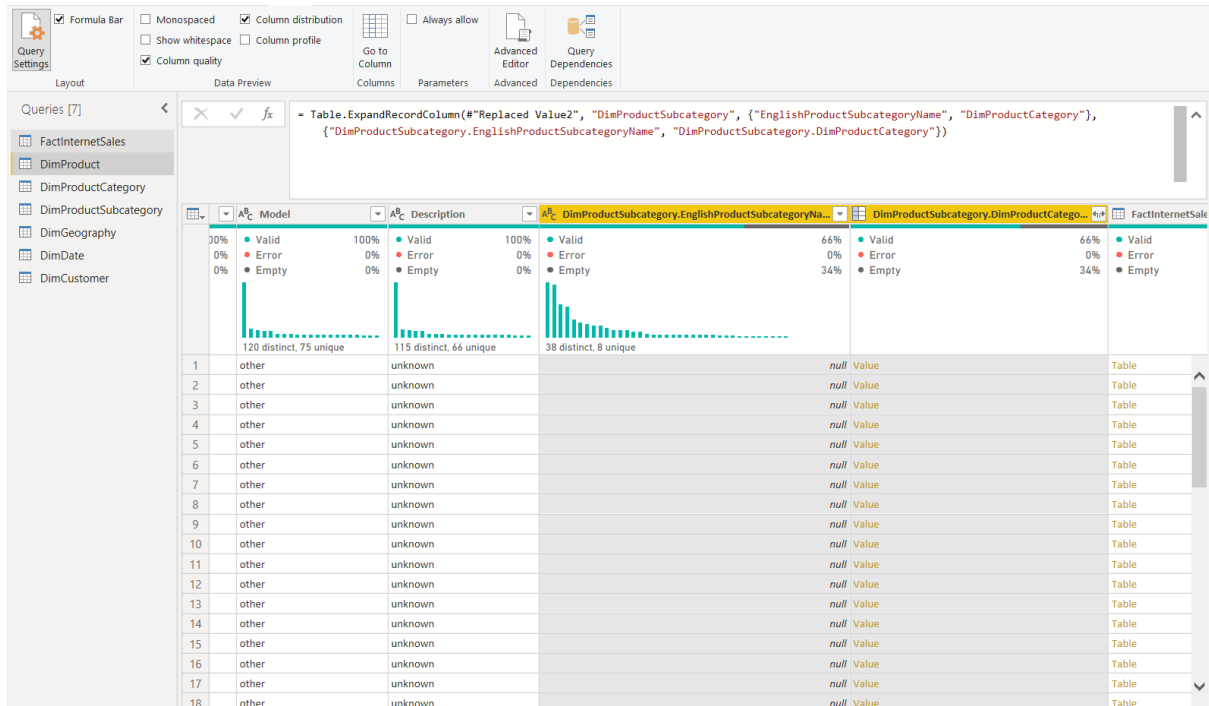
☒ Use original column name as prefix

OK Cancel

yKey	Product Name	Color	Model	Description	DimProductSubcategory	FactInternetSales
1	Adjustable Race	NA	other			
2	Bearing Ball	NA	other			
3	BB Ball Bearing	NA	other			
4	Headset Ball Bearings	NA	other			
5	Blade	NA	other			
6	LL Crankarm	Black	other			
7	ML Crankarm	Black	other			
8	HL Crankarm	Black	other			
9	Chainring Bolts	Silver	other	unknown	Value	
10	Chainring Nut	Silver	other	unknown	Value	
11	Chainring	Black	other	unknown	Value	
12	Crown Race	NA	other	unknown	Value	
13	Chain Stays	NA	other	unknown	Value	
14	Decal 1	NA	other	unknown	Value	
15	Decal 2	NA	other	unknown	Value	
16	Down Tube	NA	other	unknown	Value	

Hình 3.6.1 Lấy thêm cột EnglishProductSubCategoryName

Hệ thống thông tin phục vụ trí tuệ kinh doanh



Hình 3.6.2 Cột dữ liệu được thêm vào

Sau khi thêm cột tiến hành đổi tên cột thành Subcategory và thay thế các giá trị null thành unknown.

Thay đổi kiểu dữ liệu - Changing Data Types

Thay đổi dữ liệu cột **ProductSubcategoryKey** từ số thập phân sang số nguyên.

3.7. Bảng FactInternetSales

Xóa các cột dữ liệu - Removing Columns

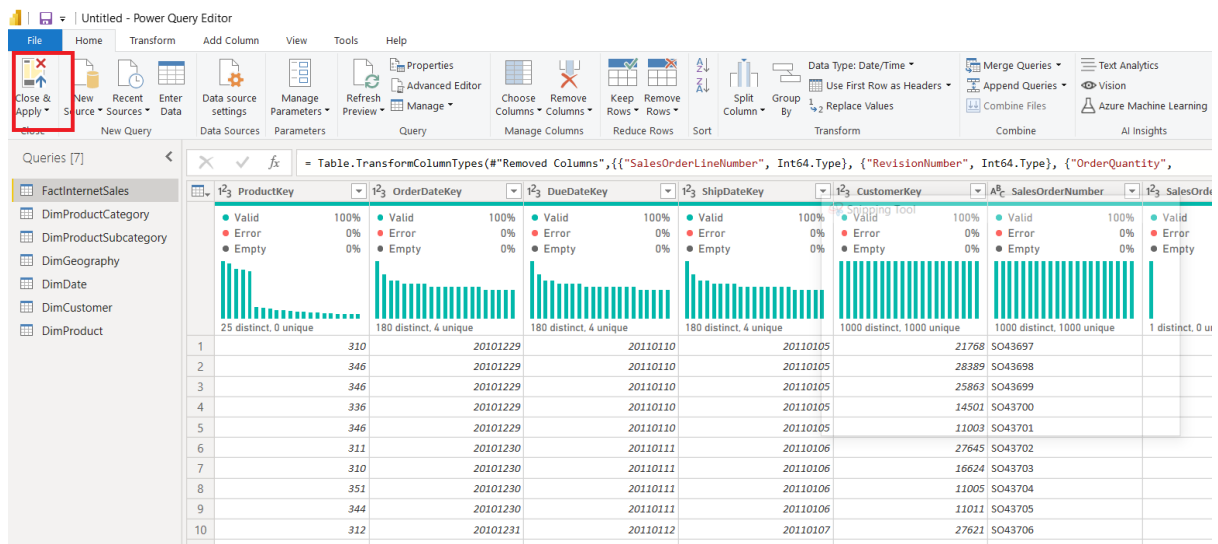
Các cột cần xóa là: CurrencyKey, SalesTerritoryKey, PromotionKey, FactInternetSalesReason, DimSalesTerritory, DimPromotion, CustomerPONumber, CarrierTrackingNumber, TaxAmt, Freight, UnitPrice, ExtendedAmount, UnitPriceDiscountPct, DiscountAmount

Hệ thống thông tin phục vụ trí tuệ kinh doanh

Thay đổi kiểu dữ liệu - Changing Data Types

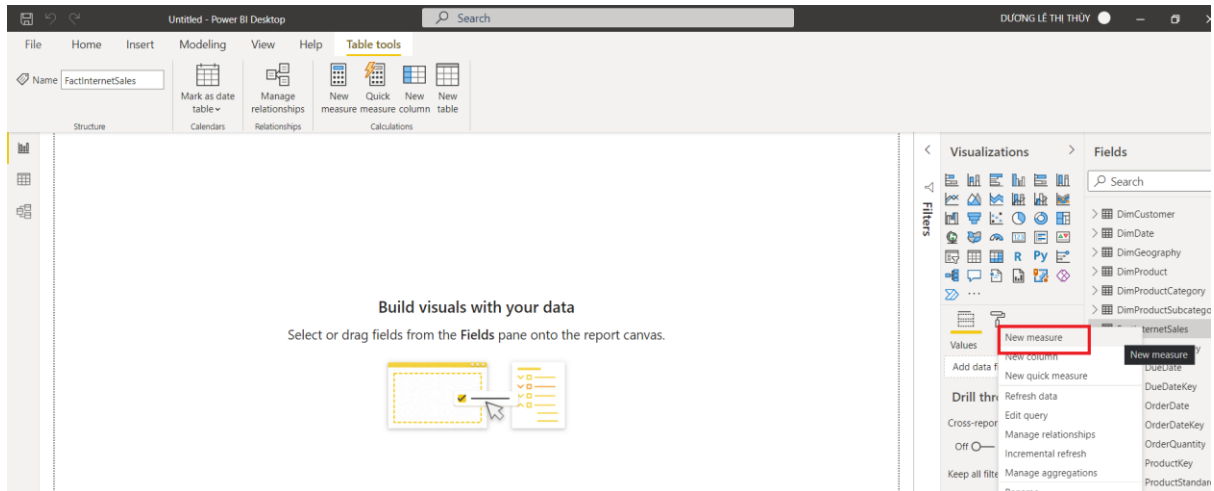
Thay đổi dữ liệu các cột **SalesOrderLineNumber** từ thập phân sang số nguyên, **RevisionNumber** từ thập phân sang số nguyên, **OrderQuantity** từ thập phân sang số nguyên.

Sau khi xử lý xong hết chọn **Apply and Close** để lưu lại thay đổi.

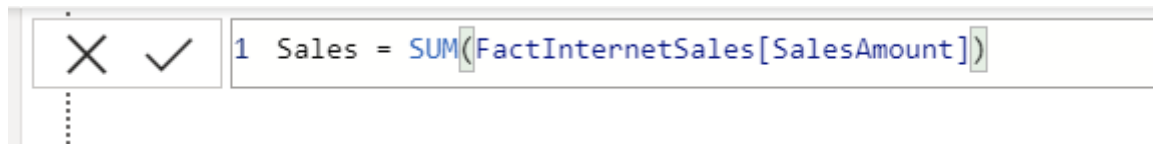


4. TRỰC QUAN HÓA DỮ LIỆU

4.1. Tạo một measure mới

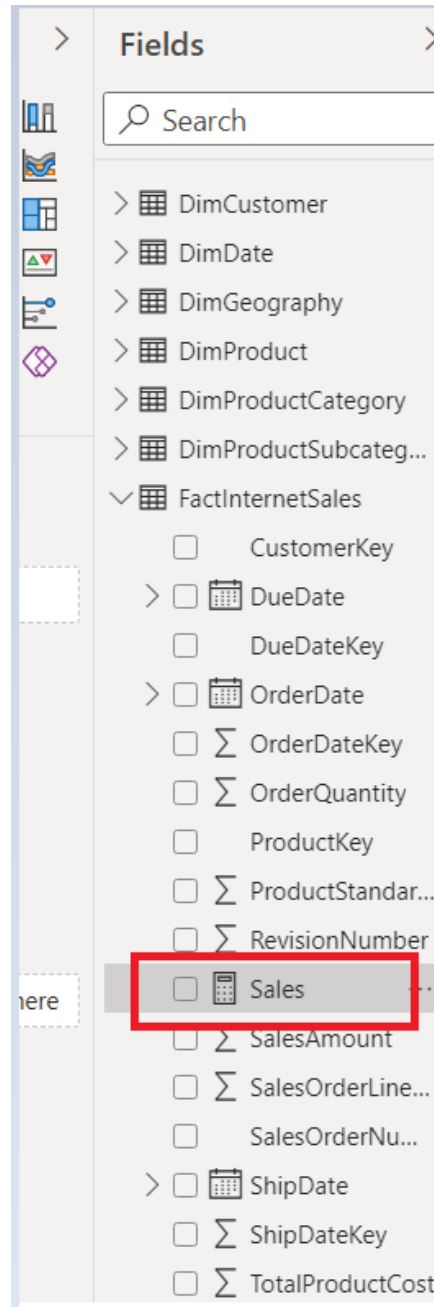


Hình 4.1 Tạo measure mới



Hình 4.2 Code DAX tạo một measure sales

Hệ thống thông tin phục vụ trí tuệ kinh doanh

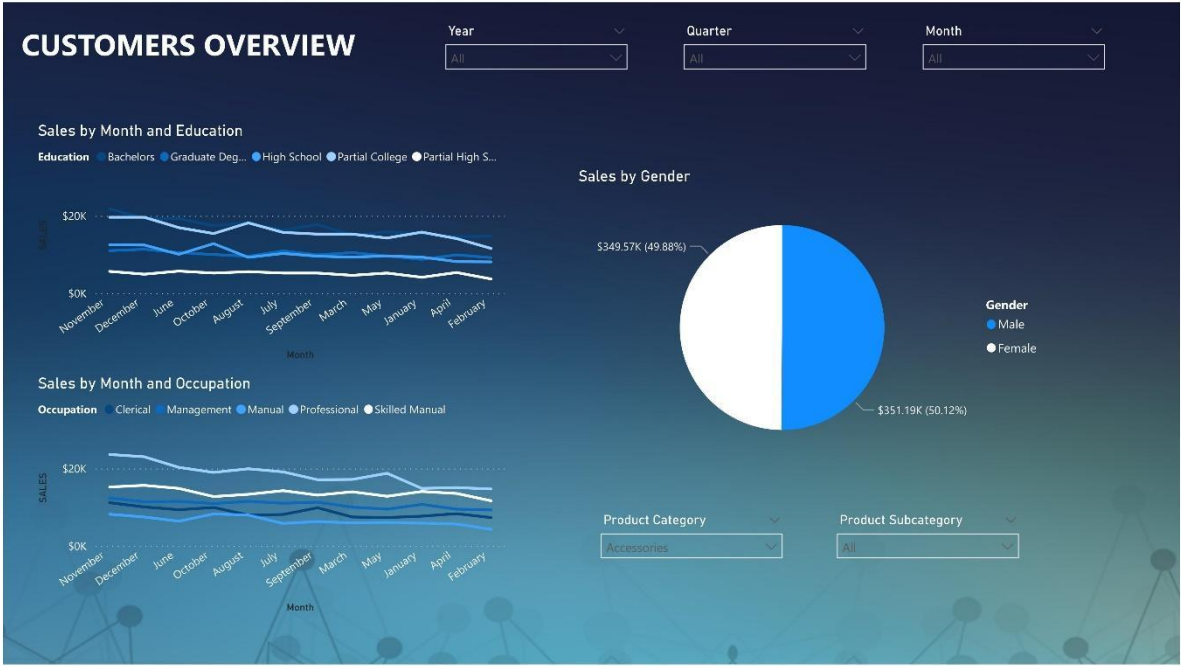


Hình 4.3 Measure được tạo

4.2. Trực quan hóa dữ liệu

- Report về Customers Overview

Hệ thống thông tin phục vụ trí tuệ kinh doanh

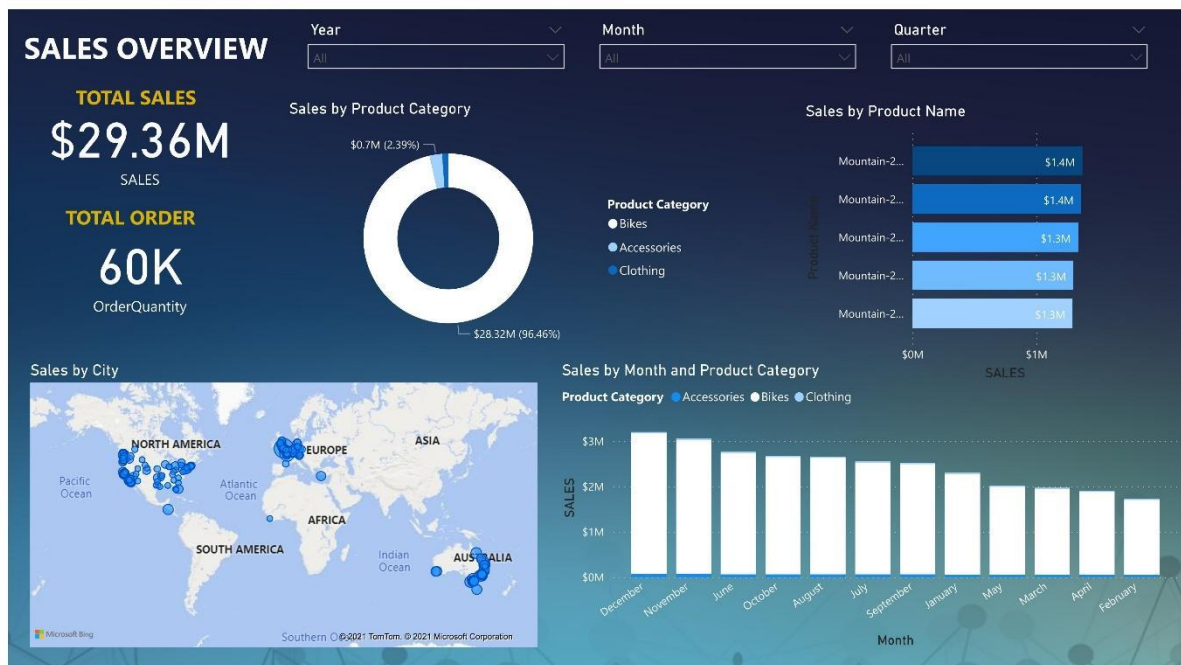


- Report về Products Overview



- Report về Sales Overview

Hệ thống thông tin phục vụ trí tuệ kinh doanh



5. NGUỒN THAM KHẢO

[Getting Started with Power BI and Power Query for Simple ETL - Part 2 \(mssqltips.com\)](#)

<https://alluringbi.com/blog-feed/page/2/>

[Background Gallery | Alluring Analytics \(alluringbi.com\)](#)

<https://www.theta.co.nz/search-results?keywords=Power BI For>

[Sale Analysis for AdventureWorks - Microsoft Power BI Community](#)

[The Data Visualisation Catalogue \(datavizcatalogue.com\)](#)

[Datawrapper: Create charts, maps, and tables](#)

[Visual Vocabulary ft. Vega \(gramener.github.io\)](#)