Bài tập: Kết nối và Xử lý Data

Sử dụng bộ file thực hành đã được đính kèm **Adventure Works**, hoàn thành các nhiệm vụ sau:

- 1) Kết nối dữ liệu từ file AdventureWorks_Product_Categories và AdventureWorks_Product_Subcategories từ bộ file thực hành:
 - Đặt tên queries là AW_Product_Category_Lookup và
 AW_Product_Subcategory_Lookup
 - Đảo bảo rằng tất cả các trường đều có header và loại dữ liệu trong các trường dữ liệu là chuẩn xác
 - Disable chức năng refresh cho cả 2 bảng dữ liệu
- 2) Tùy biến dữ liệu trong Query Editor cho AW_Product_Lookup:
 - Thêm một calculated column để chiết xuất ra toàn bộ ký tự trước dấu gạch ngang ("") trong trường (cột) **ProductSKU**, đặt tên là "**SKUType**"
 - Cập nhật tính toán cho trường **SKUType** để chiết xuất các ký tự trước dấu gạch ngan thứ 2, thay vì dầu gạch ngang đầu tiên
 - Thay thế các giá trị 0 trong trường **ProductStyle** bằng "*NA*"
 - Cập nhật dữ liệu trường **DiscountPrice** để tính giá khuyến mại 15%, bằng cách nhân **ProductPrice** với 0.85 *(thay vì 0.9)*
- 3) Kết nối dữ liệu Sale từ 3 file dữ liệu gốc: *AdventureWorks_Sales_2015*, *AdventureWorks_Sales_2016* và *AdventureWorks_Sales_2017* (có thể sử dụng Get data từ Folder hoặc sử dụng tính năng Append Queries)
- 4) Sử dụng Statistics tools trong Query Editor, thực hiện các nhiệm vụ sau:
 - Tính chi phí sản phẩm trung bình product cost (Kết quả: \$413.66)
 - Đếm số màu sản phẩm (distinct product colors) (Kết quả: **10**)
 - Đếm số tên khách hàng (distinct customer names) (Kết quả: 18,110)

- Thu nhập lớn nhất của một khách hàng (Kết quả: \$170,000)
- Đếm số đặt hàng (Kết quả: **56,046**)
- Đếm số đặt hàng khác nhau (Kết quả: **25,164**)

5) Tùy biến dữ liệu trong Query Editor cho AW_Customer_Lookup

- Thêm một calculated column để tính năm sinh (đặt tê là "BirthYear"), dựa trên dữ liệu từ trường BirthDate
- Thêm một conditional column để phân loại khách hàng theo thu nhập (đặt tên là "**IncomeLevel**"), dựa trên các điều kiện sau:
- Nếu **AnnualIncome** >= \$150,000, thì **IncomeLevel** = "Very High"
- Nếu **AnnualIncome** >= \$100,000, thì **IncomeLevel** = "*High*"
- Nếu **AnnualIncome** >= \$50,000, thì **IncomeLevel** = "Average"
- Các trường hợp còn lại **IncomeLevel** = "Low"
- **6)** Apply các thay đổi, và đảm bảo các trường và các bảng đều có trong cả **Data** view và **Relationships** view. Lưu lại file PBI.