

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN



HCMUTE

**BÁO CÁO ĐỒ ÁN MÔN HỌC
KHO DỮ LIỆU**

ĐỀ TÀI:

PHÂN TÍCH DỮ LIỆU TẤN CÔNG MẠNG

GVHD: ThS. Nguyễn Văn Thành

Nhóm sinh viên thực hiện:

- | | |
|-----------------------|----------------|
| 1. Lê Hồ Quốc Huy | MSSV: 22133025 |
| 2. Nguyễn Duy Nam | MSSV: 18133031 |
| 3. Trần Bảo Việt | MSSV: 22133065 |
| 4. Lê Quỳnh Nhựt Vinh | MSSV: 22133066 |

TP. HỒ CHÍ MINH, 05/2025

NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN

....., ngày.....tháng.....năm 2025

Người nhận xét

(Ký tên và ghi rõ họ tên)

PHÂN CÔNG NHIỆM VỤ

Nhiệm vụ	Trần Bảo Việt 22133065	Lê Quỳnh Nhựt Vinh 22133066	Lê Hồ Quốc Huy 22133025	Nguyễn Duy Nam 18133031
Tìm kiếm tập dữ liệu	X	X	X	X
Hiệu tập dữ liệu	X	X	X	X
Xác định Business Process	X	X	X	X
Xác định bảng Dim	X			X
Xác định bảng Fact	X			X
Đẩy dữ liệu từ CSV – SQL Server	X			X
Tạo nguồn kết nối dữ liệu	X			X
Staging và load và các dim, fact	X			X
Nhập dữ liệu vào SSAS và tạo Data Cube			X	
Phân tích SSAS			X	
Đặt các câu hỏi	X	X	X	X
Trả lời các câu hỏi bằng SSAS		X	X	
Trả lời các câu hỏi bằng Pivot Table		X	X	
Report PowerBI		X		
Viết báo cáo và trình bày	X	X	X	X

MỤC LỤC

1.1. Lý do chọn đề tài.	2
1.2.1. Tổng quan dataset	3
1.3. Tiền xử lý dữ liệu	4
1.4. Thiết kế kho dữ liệu	6
1.4.1. Lược đồ kho dữ liệu	6
1.4.2. Mô tả các bảng trong kho dữ liệu	7
1.5. Các câu truy vấn	9
2.1. Chuẩn bị các công cụ	10
2.2. Chuẩn bị cơ sở dữ liệu	10
2.3. Tạo mới project SSIS	11
2.4. Tạo bảng Dim và bảng Fact	11
2.4.1. Bảng Dim_Timestamp	15
2.4.2. Bảng Dim_Address	21
2.4.3. Bảng Dim_Attack	25
2.4.4. Bảng Dim_Victim	31
2.4.5. Bảng Fact_CyberSecurity	37
2.4.6. Tạo khoá ngoại	43
2.4.7. Chạy SSIS	45
2.5. Kiểm tra dữ liệu các bảng	49
2.6. Lược đồ sau khi hoàn thành	53
3.1. Tạo project SSAS mới	54
3.2. Xác định dữ liệu nguồn (Data Source)	55
3.3. Xác định khung nhìn dữ liệu nguồn (Data Source View)	59
3.4. Xây dựng các khối (Cube) và deploy Cube	64
3.4.1.1. Tạo Cube và Dimension	64
3.4.2. Thêm thuộc tính và chỉnh sửa property cho Dimension	69
3.4.3. Deploy project SSAS	73
3.5. Xác định các độ đo (Measures)	78
3.6. Phân cấp bảng chiều	80
3.7. Thực hiện các truy vấn sử dụng SSAS, Pivot table	87
3.7.1. Câu truy vấn 1	87
3.7.1.1. Sử dụng SSAS	87
3.7.1.2. Sử dụng Pivot table	88
3.7.2. Câu truy vấn 2	88
3.7.2.1. Sử dụng SSAS	89
3.7.2.2. Sử dụng Pivot table	89
3.7.3. Câu truy vấn 3	90
3.7.3.1. Sử dụng SSAS	90
3.7.3.2. Sử dụng Pivot table	91
3.7.4. Câu truy vấn 4	92
3.7.4.1. Sử dụng SSAS	92
3.7.4.2. Sử dụng Pivot table	93

3.7.5. Câu truy vấn 5	93
3.7.5.1. Sử dụng SSAS	94
3.7.5.2. Sử dụng Pivot table	96
3.7.6. Câu truy vấn 6	96
3.7.6.1. Sử dụng SSAS	96
3.7.6.2. Sử dụng Pivot table	97
3.7.7. Câu truy vấn 7	97
3.7.7.1. Sử dụng SSAS	97
3.7.7.2. Sử dụng Pivot table	98
3.7.8. Câu truy vấn 8	99
3.7.8.1. Sử dụng SSAS	99
3.7.8.2. Sử dụng Pivot table	99
3.7.9. Câu truy vấn 9	99
3.7.9.1. Sử dụng SSAS	100
3.7.9.2. Sử dụng Pivot table	101
3.7.10. Câu truy vấn 10	101
3.7.10.1. Sử dụng SSAS	102
3.7.10.2. Sử dụng Pivot table	102
3.7.11. Câu truy vấn 11	103
3.7.11.1. Sử dụng SSAS	103
3.7.11.2. Sử dụng Pivot table	104
4.1. Quá trình lập báo biểu bằng PowerBI	106
4.1.1. Chuẩn bị công cụ	106
4.1.2. Tạo mới report và kết nối	106
4.1.1. Tạo report 1	108
4.1.2. Tạo report 2	112
4.1.3. Tạo report 3	118
4.2. Quá trình lập báo biểu bằng Google Data Studio	122
4.2.1. Tạo report 1	122
4.2.2. Tạo report 2	126
4.2.3. Tạo report 3	133
DANH MỤC TÀI LIỆU THAM KHẢO	138

CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI

1.1. Lý do chọn đề tài.

Lĩnh vực an ninh mạng đang ngày càng trở nên quan trọng trong thời đại số hóa hiện nay. Với sự phát triển nhanh chóng của công nghệ và sự phụ thuộc ngày càng tăng của các tổ chức, doanh nghiệp và cá nhân vào hệ thống thông tin, việc bảo vệ dữ liệu và tài sản số trở thành một ưu tiên hàng đầu. Phân tích dữ liệu về các cuộc tấn công an ninh mạng giúp chúng ta hiểu rõ hơn về các mối đe dọa, phương thức tấn công và xu hướng mới nhất trong lĩnh vực này, từ đó có thể phát triển các chiến lược phòng thủ hiệu quả hơn.

Hơn nữa, việc nghiên cứu và phân tích dữ liệu về các cuộc tấn công an ninh mạng còn mang lại giá trị to lớn cho cộng đồng và xã hội. Kết quả từ các phân tích này có thể giúp cải thiện các hệ thống bảo mật, phát triển các công cụ phòng chống mới, và nâng cao nhận thức của người dùng về an ninh mạng. Điều này không chỉ góp phần bảo vệ thông tin cá nhân và tài sản của người dùng mà còn giúp doanh nghiệp và tổ chức giảm thiểu rủi ro và tổn thất tài chính do các cuộc tấn công gây ra. Ngoài ra, trong bối cảnh các cuộc tấn công mạng ngày càng tinh vi và quy mô lớn, việc nghiên cứu sâu về chủ đề này còn có thể đóng góp vào việc xây dựng các chính sách và quy định pháp lý liên quan đến an ninh mạng, tạo ra một môi trường số an toàn và đáng tin cậy hơn cho tất cả mọi người.

1.2. Giới thiệu về dataset.

1.2.1. Tổng quan dataset

Tên bộ dữ liệu: Cyber Security Attacks

Tác giả: Aashray Agur và Uma Venugopal

Bộ dữ liệu cung cấp thông tin về các cuộc tấn công mạng trên toàn cầu từ năm 2020 đến năm 2023

Nguồn tải dataset: <https://www.kaggle.com/datasets/teamincribo/cyber-security-attacks>

Bộ dữ liệu gồm có 25 cột và 40000 dòng

1.2.2. Mô tả thuộc tính

STT	Tên thuộc tính	Kiểu dữ liệu	Ý nghĩa
1	Timestamp	timestamp	Thời điểm chính xác khi sự kiện xảy ra
2	Source IP Address	varchar	Địa chỉ IP của thiết bị gửi dữ liệu
3	Destination IP Address	varchar	Địa chỉ IP của thiết bị nhận dữ liệu
4	Source Port	varchar	Số cổng trên thiết bị gửi dữ liệu
5	Destination Port	varchar	Số cổng trên thiết bị nhận dữ liệu
6	Protocol	varchar	Quy tắc truyền thông được sử dụng (TCP, UDP)
7	Packet Length	int	Kích thước gói dữ liệu
8	Packet Type	varchar	Chỉ định loại gói tin trong mạng (Data, Control)
9	Traffic Type	varchar	Phân loại lưu lượng mạng (HTTP, FTP, Email)
10	Payload Data	varchar	Nội dung của gói dữ liệu
11	Malware Indicators	varchar	Dấu hiệu của phần mềm độc hại
12	Anomaly Scores	float	Đánh giá mức độ bất thường
13	Alerts/Warnings	varchar	Cảnh báo về sự kiện nguy hiểm
14	Attack Type	varchar	Phương thức tấn công (DDoS, phishing)
15	Attack Signature	varchar	Đặc điểm nhận dạng của cuộc tấn công
16	Action Taken	varchar	Biện pháp đối phó (chặn, cách ly)
17	Severity Level	varchar	Mức độ nguy hiểm của sự cố
18	User Information	varchar	Thông tin người dùng
19	Device Information	varchar	Thông tin thiết bị liên quan
20	Network Segment	varchar	Phân đoạn mạng nơi sự kiện xảy ra
21	Geo-location Data	varchar	Vị trí địa lý của các IP liên quan
22	Proxy Information	varchar	Thông tin về proxy được sử dụng

23	Firewall Logs	varchar	Nhật ký hoạt động của tường lửa
24	IDS/IPS Alerts	varchar	Cảnh báo từ hệ thống phát hiện/ngăn chặn xâm nhập
25	Log Source	varchar	Nguồn gốc của nhật ký (Server, Firewall)

1.3. Tiết xuât lý dữ liệu

Bước 1: Import các thư viện cần thiết cho việc xử lý dữ liệu và đọc file csv

```
import pandas as pd
import requests
from time import sleep
import xlrd
from xlutils.copy import copy

file_path = '/Users/tranbaoviet/Downloads/file_moi.csv'
df = pd.read_csv(file_path)
```

Bước 2: Tiến hành điền dữ liệu null vào các ô trống của các cột như: Malware Indicators, IDS/IPS Alerts, Firewall Logs

```
df['IDS/IPS Alerts'] = df['IDS/IPS Alerts'].fillna('No data')
df['Firewall Logs'] = df['Firewall Logs'].fillna('No records')
df['Malware Indicators'].fillna('No detected')
df['Proxy Information'].fillna('No detected')
```

Bước 3: Chuyển đổi dữ liệu Source IP Address, Destination IP Address sang địa chỉ vật lý bằng cách gọi API của ipinfo.io

```
TOKEN = "bd75ae0f9db684"
```

```

def get_location(ip):
    try:
        response = requests.get(f"https://ipinfo.io/{ip}?token={TOKEN}")
        data = response.json()
        city = data.get('city', '')
        region = data.get('region', '')
        country = data.get('country', '')
        return f"{city}, {region}, {country}".strip(',')
    except:
        return "Unknown"
    finally:
        sleep(0.1) # Để tránh vượt quá giới hạn tốc độ của API

# Kiểm tra xem cột 'Source IP Address' có tồn tại không
if 'Source IP Address' in df.columns:
    # Tạo một cột mới 'Source Address' và thêm thông tin vị trí dựa trên IP
    df['Source Address'] = df['Source IP Address'].apply(get_location)

# Làm tương tự như Source IP Address
if 'Destination IP Address' in df.columns:
    df['Destination Address'] = df['Destination IP Address'].apply(get_location)

```

Bước 4: Sau khi chuyển đổi IP sang địa chỉ vật lý, sẽ tồn tại những record không chứa dữ liệu vì IP bị lỗi, tiến hành làm sạch dữ liệu

```
df_cleaned = df.dropna(subset=['Source Address', 'Destination Address'])
```

Bước 5: Sau khi chuyển đổi IP dữ liệu sẽ có dạng city,region,country. Nhóm em sẽ thống kê theo country nên sẽ tách cột country ra

```
def extract_country(address):
    return address.split(",")[-1].strip()

df['Source Country'] = df['Source Address'].apply(extract_country)
df['Destination Country'] = df['Destination Address'].apply(extract_country)
```

Bước 6: Xoá các cột không có dữ liệu country.

```
df = df[df['Source Address'] != 'Unknown']
df = df[df['Destination Address'] != 'Unknown']
```

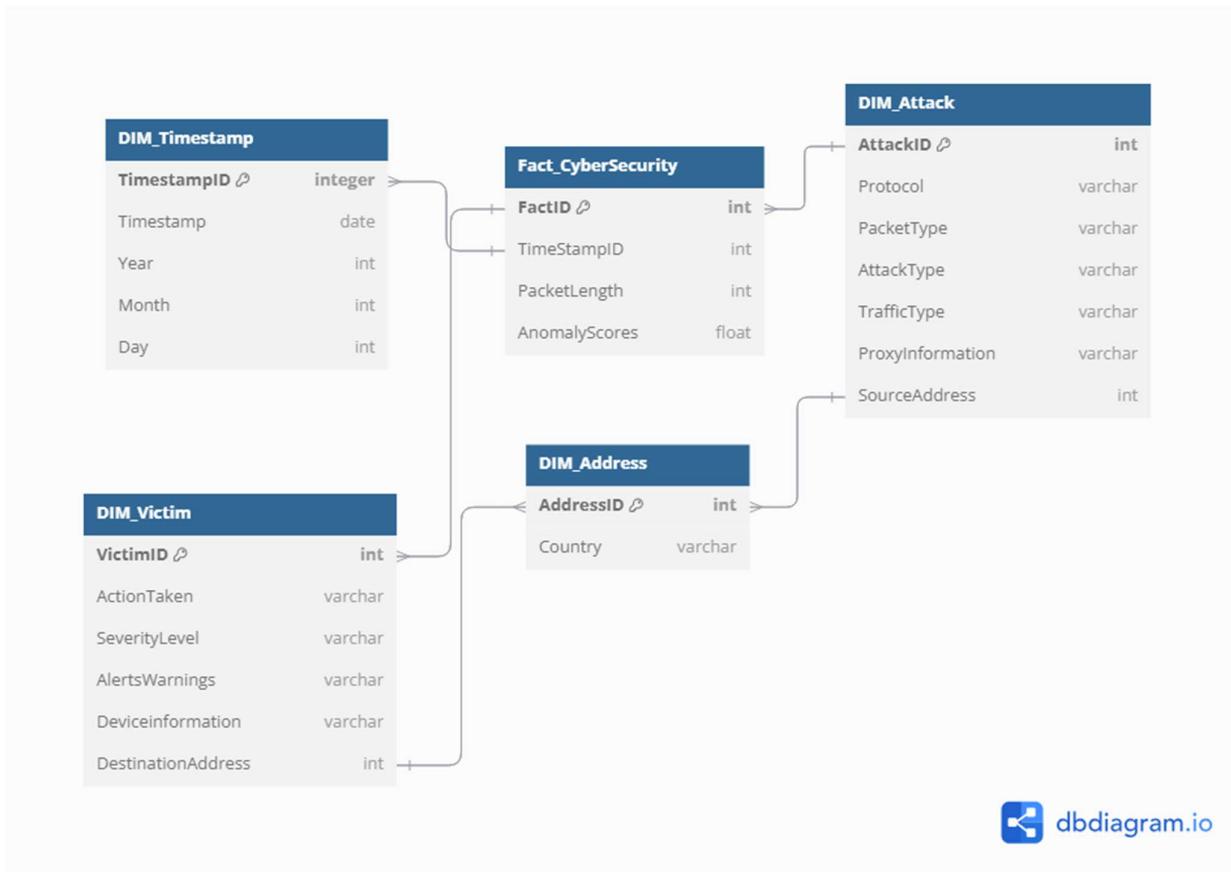
Bước 7: Lưu file csv.

```
df.to_csv('cyber_security_final1.csv', index=False)
```

Bộ dữ liệu sau khi xử lý sẽ gồm 30 cột và 39560 dòng.

1.4. Thiết kế kho dữ liệu

1.4.1. Lược đồ kho dữ liệu



1.4.2. Mô tả các bảng trong kho dữ liệu

1.4.2.1. Bảng Fact_CyberSecurity

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Mô tả thuộc tính
1	FactID	int	Khoá chính	Mã fact
2	TimeStampID	int	Khoá ngoại	Mã thời gian xảy ra cuộc tấn công
3	PacketLength	int		Độ dài gói tin được gửi đi
4	AnomalyScores	float		Chỉ số bất thường

1.4.2.2. Bảng Dim_Attack

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Mô tả thuộc tính
1	AttackID	int	Khoá chính	Mã tấn công

2	Protocol	varchar		Quy tắc truyền thông được sử dụng (TCP, UDP)
3	PacketType	varchar		Chỉ định loại gói tin trong mạng (Data, Control)
4	AttackType	varchar		Phương thức tấn công (DDoS, phishing)
5	TrafficType	varchar		Phân loại lưu lượng mạng (HTTP, FTP, Email)
6	ProxyInformation	varchar		Thông tin về proxy được sử dụng
7	SourceAddress	int	Khoá ngoại	Mã địa chỉ người gửi gói tin

1.4.2.3. Bảng Dim_Victim

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Mô tả thuộc tính
1	VictimID	int	Khoá chính	Mã nạn nhân
2	ActionTaken	varchar		Biện pháp đối phó (Blocked, Ignored)
3	SeverityLevel	varchar		Mức độ nguy hiểm của sự cố
4	AlertsWarnings	varchar		Cảnh báo về sự kiện nguy hiểm
5	DeviceInformation	varchar		Thông tin thiết bị liên quan
6	DestinationAddress	int	Khoá ngoại	Mã địa chỉ người nhận gói tin

1.4.2.4. Bảng Dim_Address

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Mô tả thuộc tính

1	AddressID	int	Khoá chính	Mã địa chỉ
2	Country	varchar		Quốc gia

1.4.2.5. Bảng Dim_Timestamp

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Mô tả thuộc tính
1	TimestampID	int	Khoá chính	Mã thời gian
2	Timestamp	date		Ngày Tháng Năm
3	Day	int		Ngày
4	Month	int		Tháng
5	Year	int		Năm

1.5. Các câu truy vấn

1. Truy vấn drill down tổng số cuộc tấn công bằng phương thức TCP theo từng tháng năm 2022.
2. Liệt kê top 5 cuộc tấn công có điểm cao nhất và phương thức của chúng.
3. Liệt kê tổng số cuộc tấn công từng tháng trong năm với loại là mã độc.
4. Truy vấn tổng điểm Anomaly Scores của các cuộc tấn công theo từng Phương thức (Protocol) và Năm với phương thức là Malware.
5. Truy vấn tính điểm trung bình của các cuộc tấn công Malware theo từng tháng trong năm 2022.
6. Liệt kê số cuộc tấn công bị chặn theo năm.
7. Truy vấn chiều dài các gói tin đã bị hệ thống bảo mật bỏ qua theo từng năm.
8. Liệt kê số lượng các cuộc tấn công thành công theo từng phương thức.
9. Liệt kê top 10 những cuộc tấn công có độ dài lớn nhất của hai nước Nhật Bản và Mỹ.
10. Truy vấn cuộc tấn công đã được cảnh báo nhưng bị bỏ qua trong năm 2020.
11. Tổng độ dài các gói tin của từng phương thức của 20 cuộc tấn công có điểm AnomalyScore cao nhất.

CHƯƠNG 2. QUÁ TRÌNH XÂY DỰNG KHO DỮ LIỆU (SSIS)

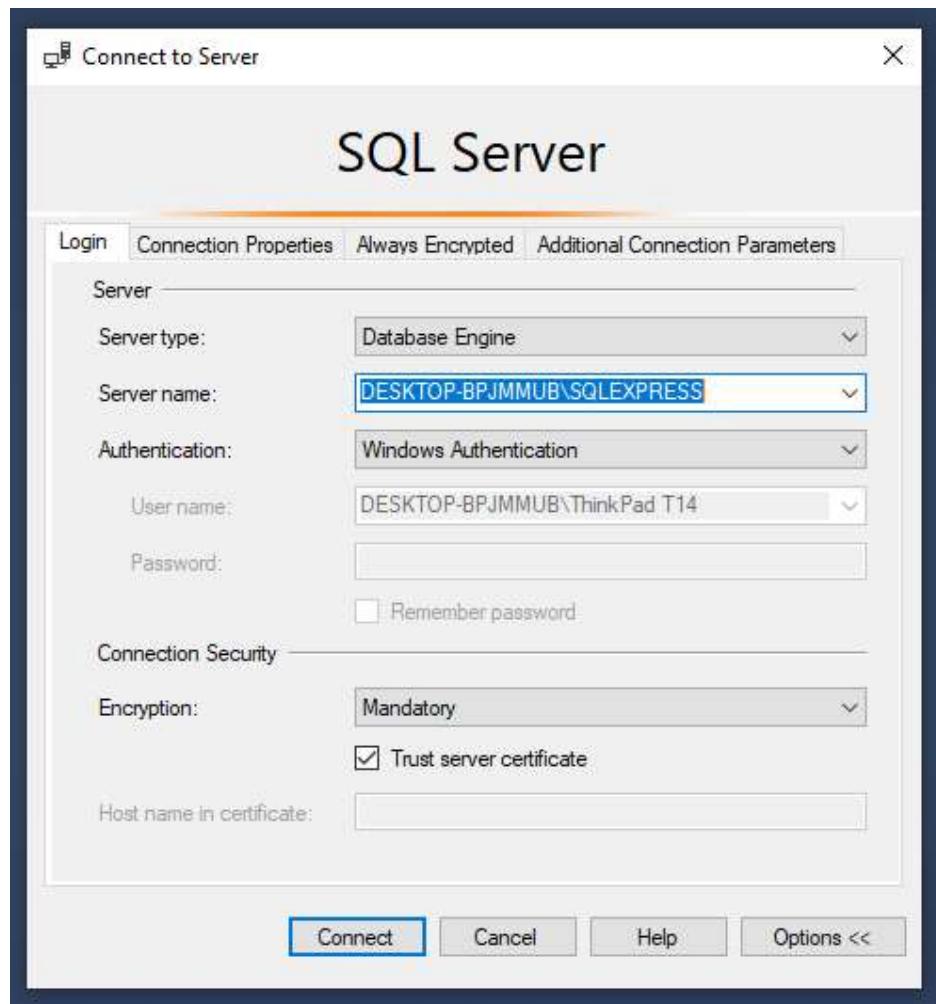
2.1. Chuẩn bị các công cụ

Để thực hiện được quá trình SSIS ta cần chuẩn bị và cài đặt các công cụ sau:

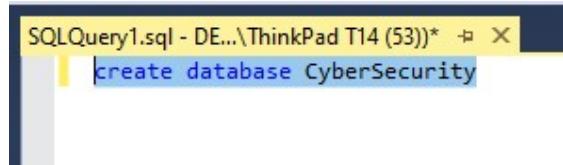
- Visual studio Community 2022
- SQL Server Integration Services Project

2.2. Chuẩn bị cơ sở dữ liệu

Bước 1: Mở SQL Server v20.2 và kết nối với server bằng tài khoản user của window (Windows Authentication).



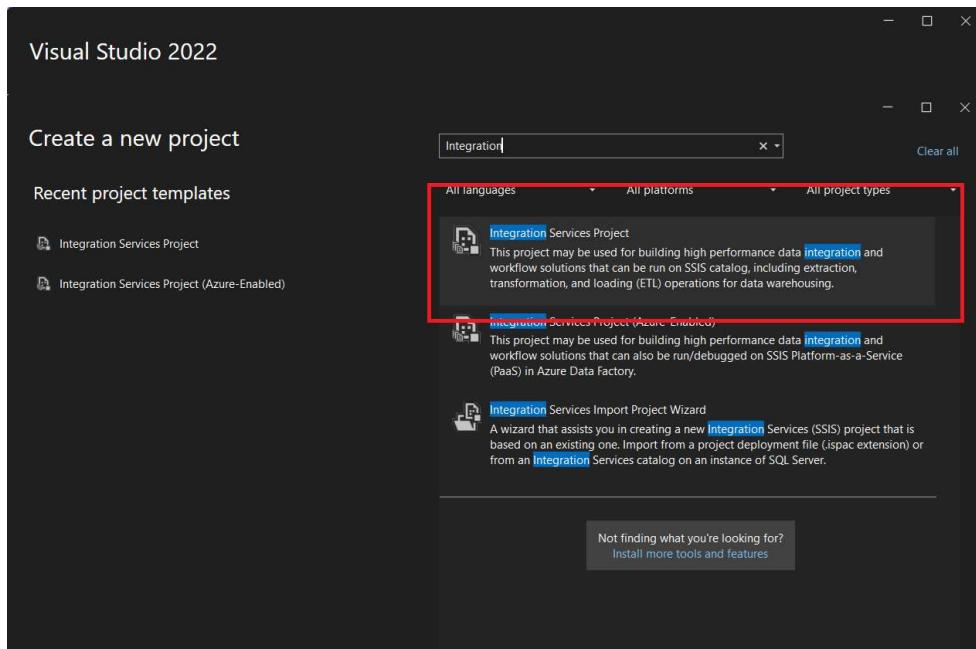
Bước 2: Khởi tạo cơ sở dữ liệu có tên CyberSecurity.



```
SQLQuery1.sql - DE...\ThinkPad T14 (53)* ↗ X
create database CyberSecurity
```

2.3. Tạo mới project SSIS

Bước 1: Mở Visual Studio 2022 và chọn “Create a new project”, chọn Integration Services Project và chọn Next.



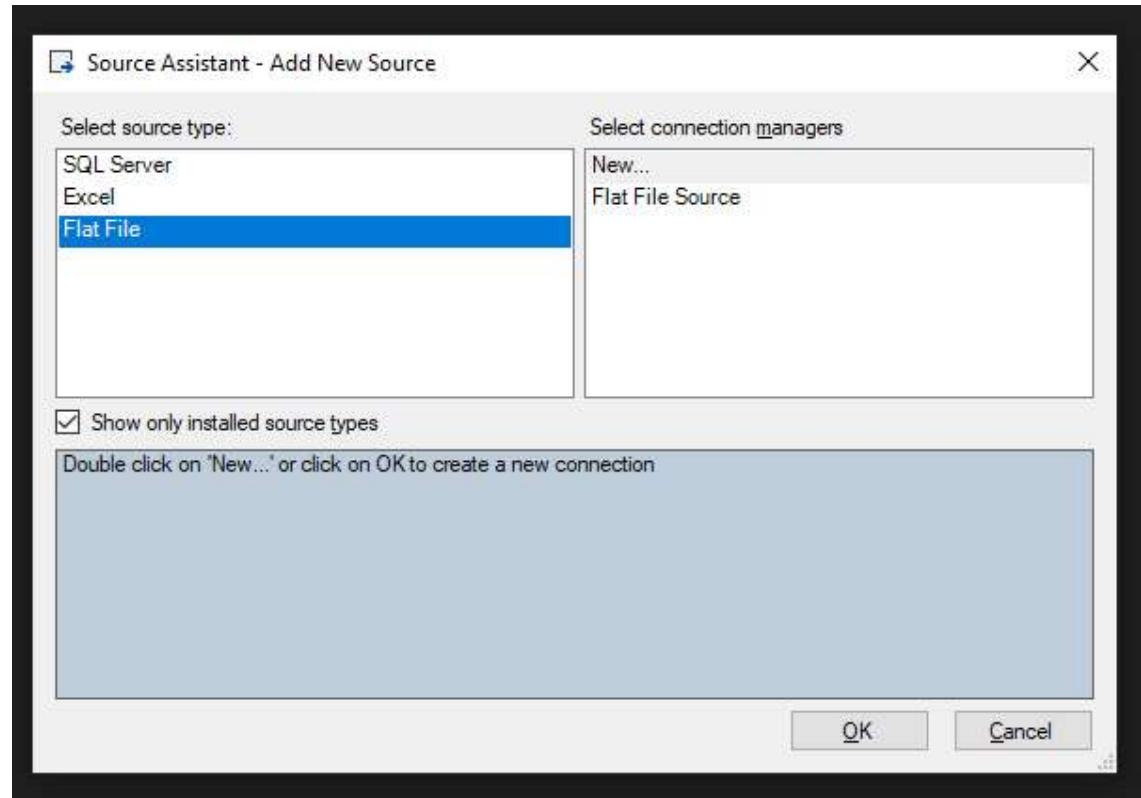
Bước 2: Đặt tên và thiết lập đường dẫn cho Project.

2.4. Tạo bảng Dim và bảng Fact

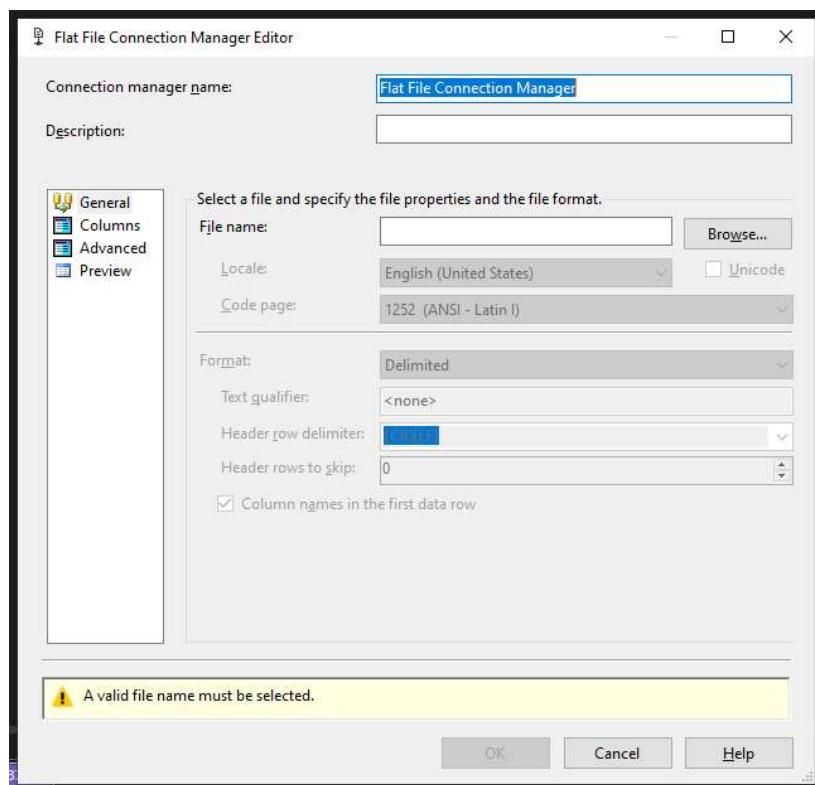
- Tạo Data Flow Task có tên là “Load Dim and Fact table” để chuẩn bị cho việc tạo các bảng Dim và Fact.



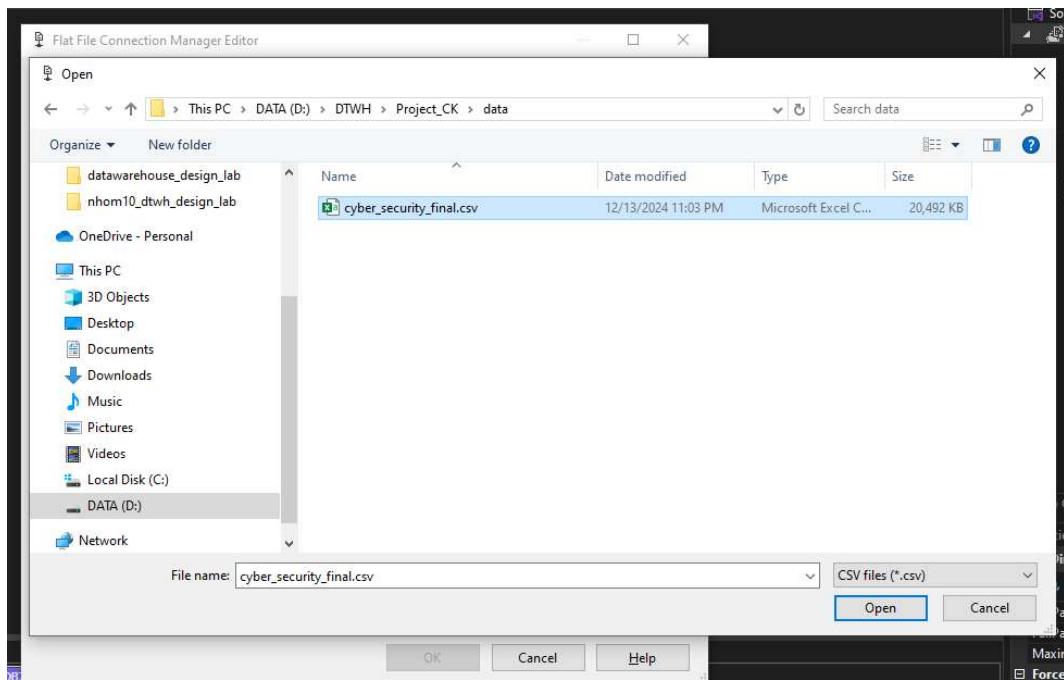
- Sang mục Data Flow, tạo mới một “Flat File Source” bằng cách add new Source Assistant từ toolbox.



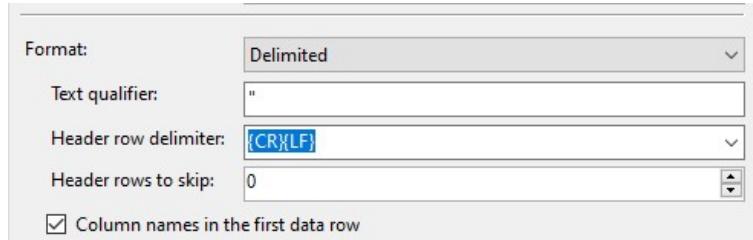
- Sau đó chọn vào “New” để tạo mới một “Flat file connection”, đọc dữ liệu từ file .csv chứa dataset.



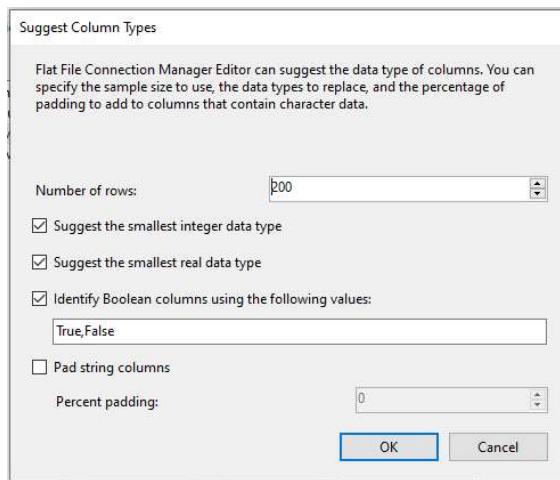
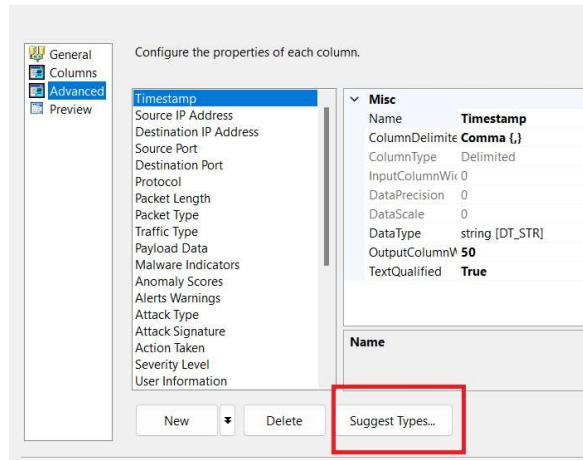
- Chọn “Browse”, chọn file .csv chứa dataset.



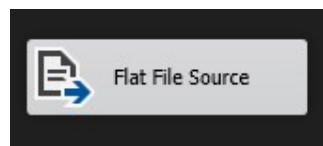
- Sau khi đã chọn file .csv, điền kí tự “ vào mục Text qualifier để tránh bị lỗi khi đọc dữ liệu

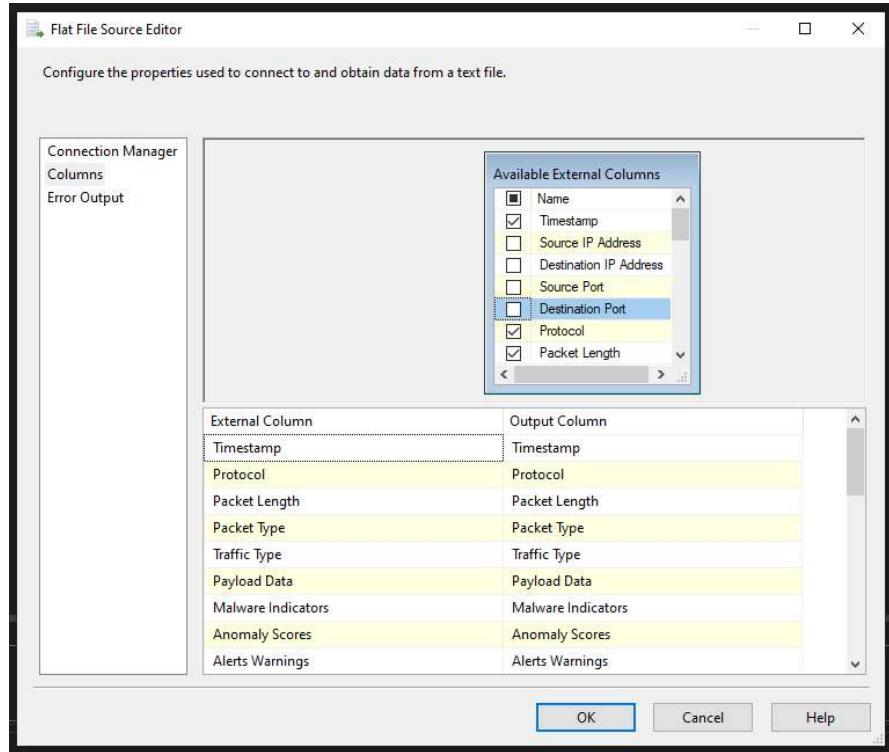


- Vào mục “Advanced” chọn “Suggest Types” để công cụ tự động chọn kiểu dữ liệu phù hợp cho từng cột và nhấn “OK”.

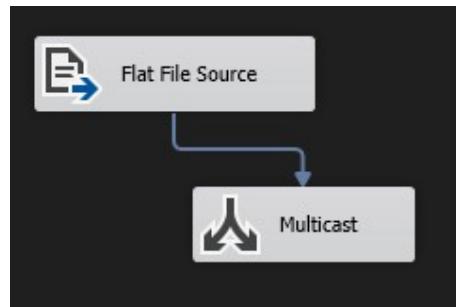


- Nhấn OK để tạo file Flat File Source. Sau đó edit vào mục “Column” Và chọn những cột trong dataset sẽ dùng.





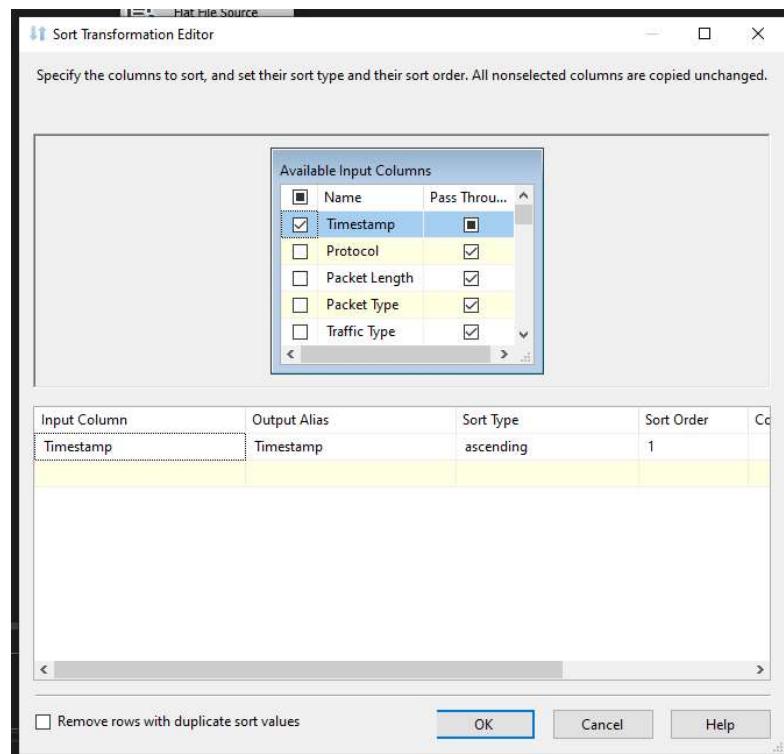
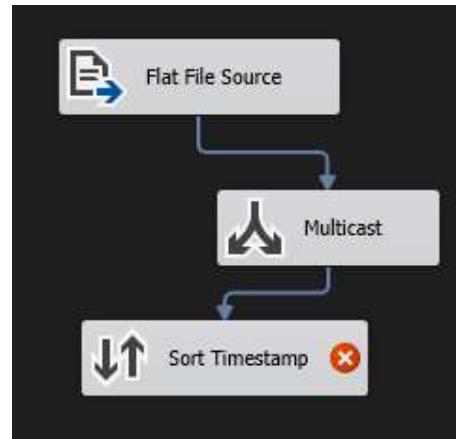
- Chọn “Multicast” để tiến hành đổ dữ liệu vào kho.



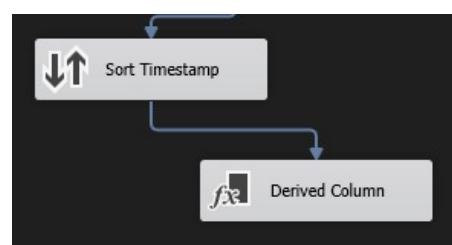
2.4.1. Bảng Dim_Timestamp

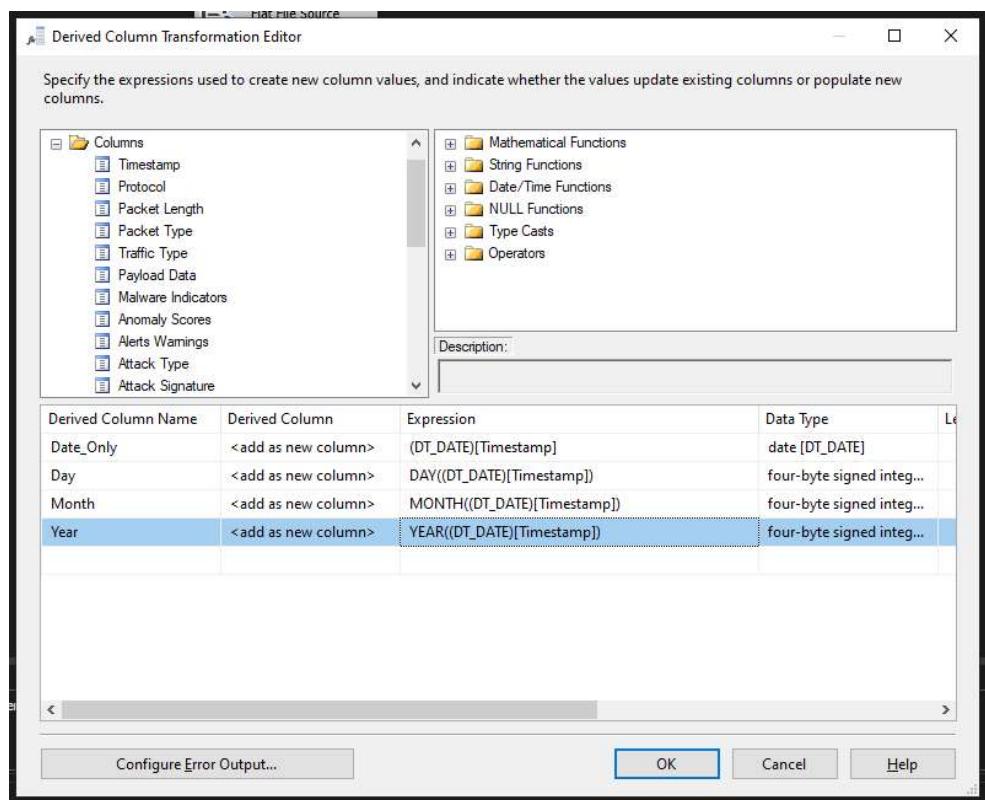
DIM_Timestamp	
TimestampID	integer
Timestamp	date
Year	int
Month	int
Day	int

Bước 1: Tạo mới một “Sort” có tên là Sort Timestamp để lấy ra các cột dữ liệu cần thiết cho DIM_Timestamp. Nhấn chuột phải và Edit để chọn Timestamp làm cột dữ liệu cho Sort Timestamp.

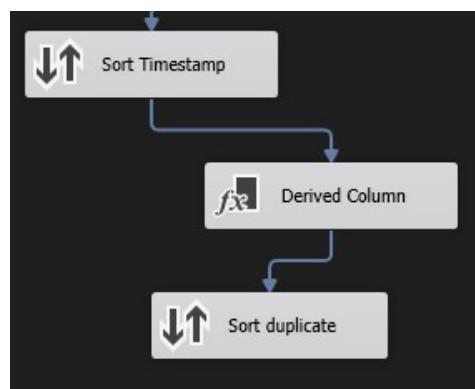


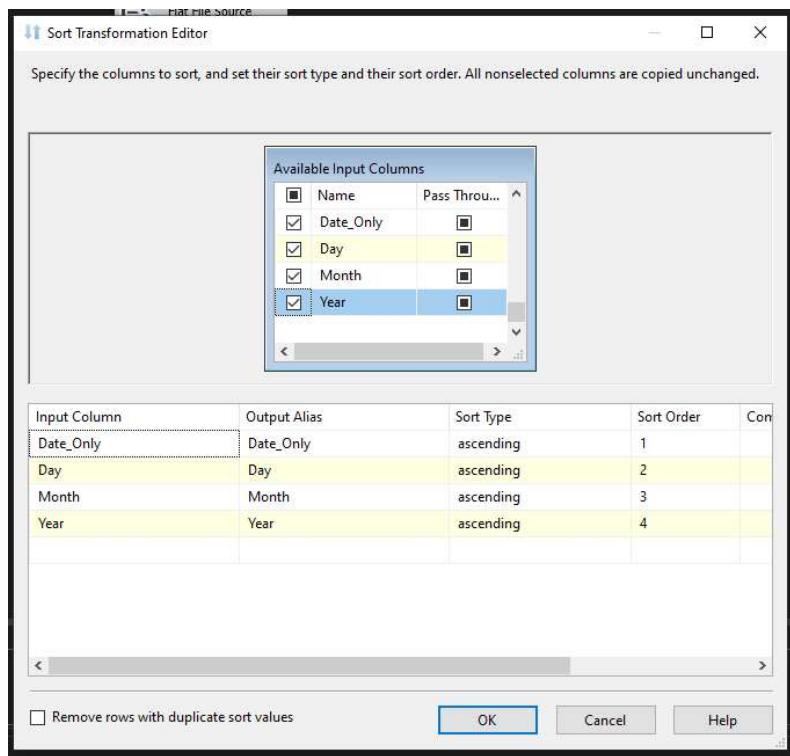
Bước 2: Thêm thành phần “Derived Column” và chọn Edit để chia cột dữ liệu. Ta chia Timestamp thành các cột dữ liệu Day, Month, Year và đồng thời ép kiểu dữ liệu của Timestamp thành Date



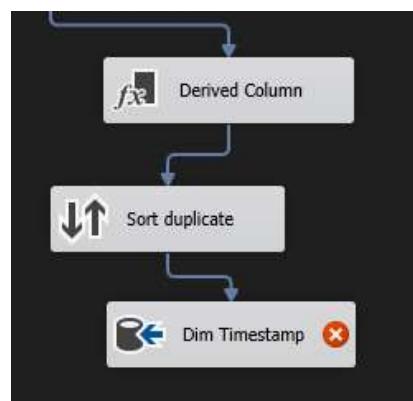


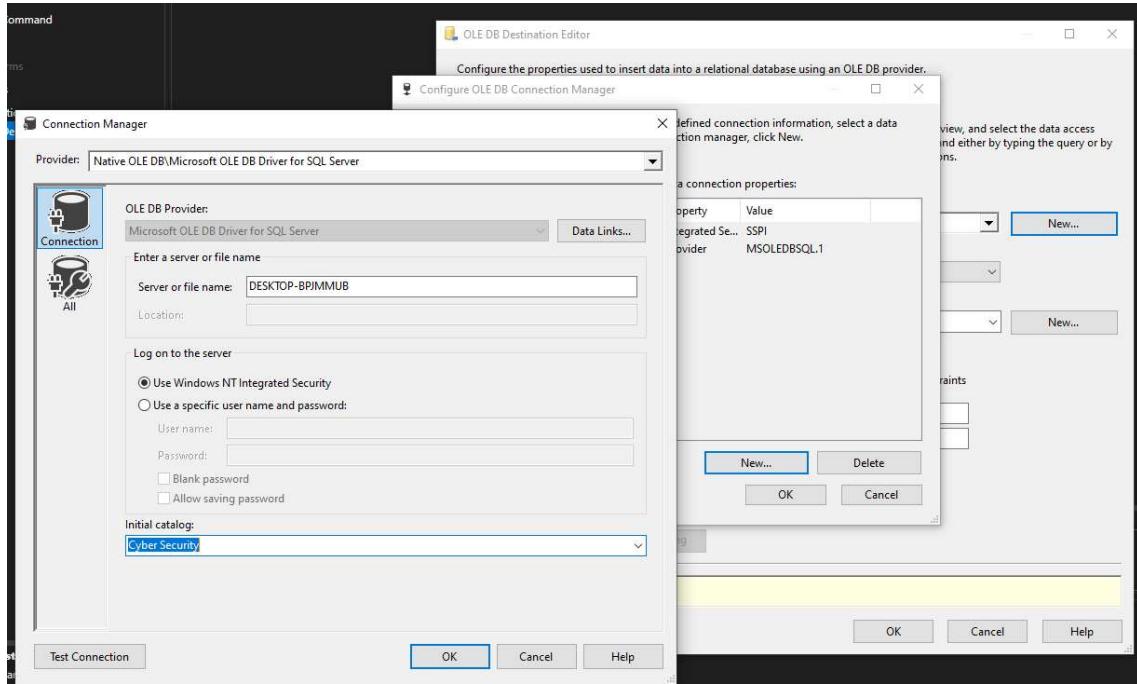
Bước 3: Thêm thành phần Sort để xoá các dòng dữ liệu bị trùng.



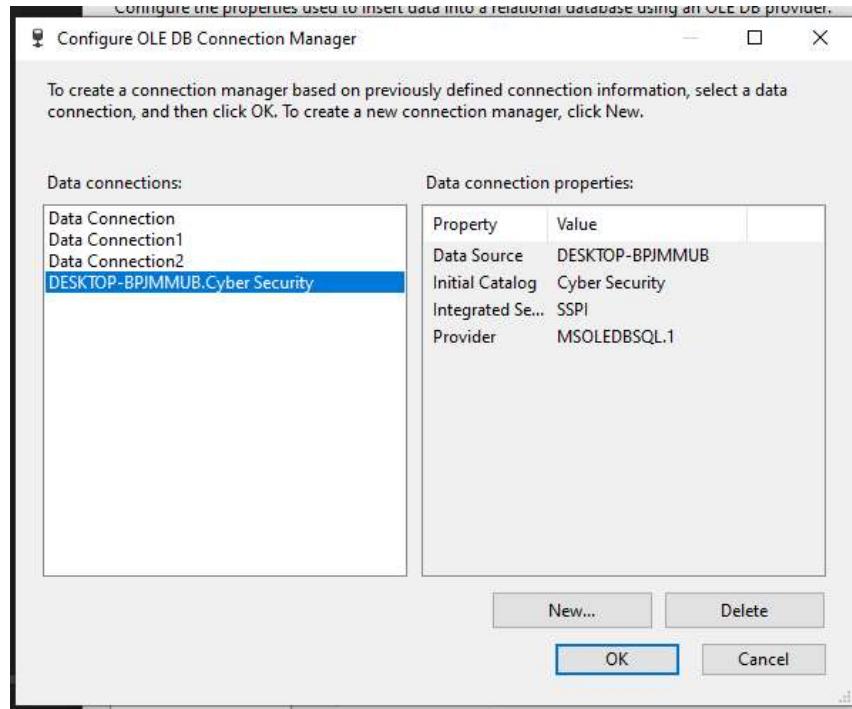


Bước 4: Tạo Dim_Timestamp từ một “OLE DB Destination”. Double click vào “OLE DB Destination” để tạo một connection mới đến MS SQL Server.

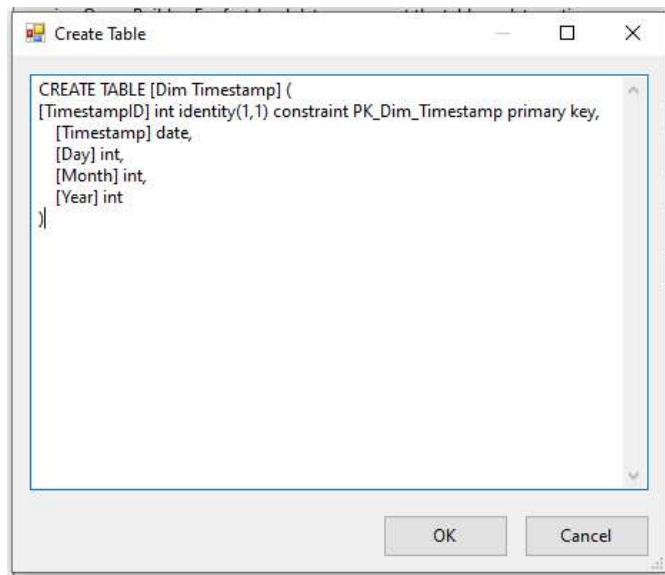
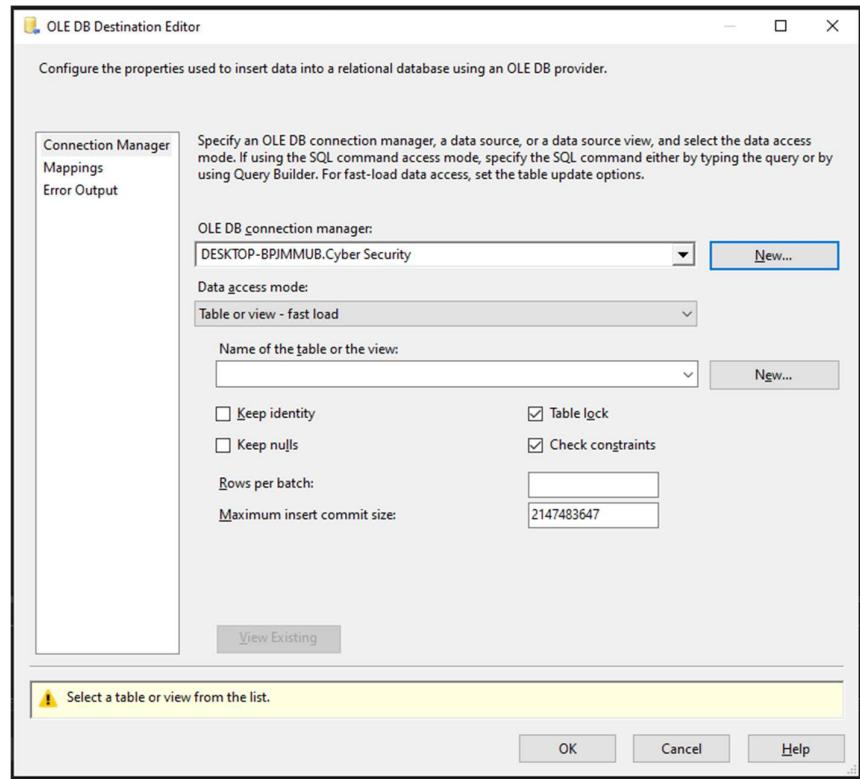




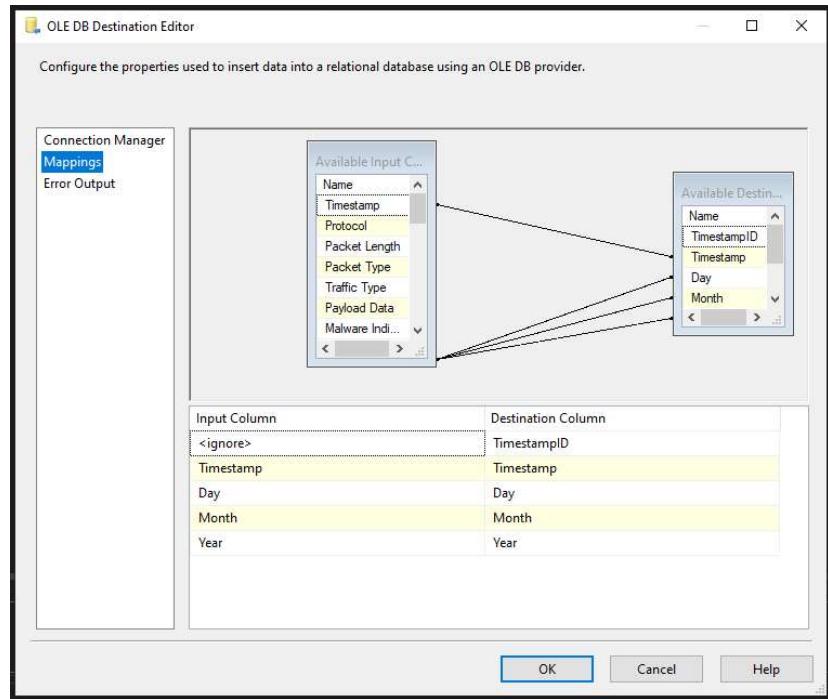
Bước 5: Chọn tên server name trùng với server name MS SQL Server để ta có thể kết nối đến datawarehouse CyberSecurity vừa tạo. Kết nối đến server bằng tài khoản window mặc định (Windows Authentication) nhấn Test Connection để kiểm tra kết nối.



Bước 6: Chọn New để tạo bảng Dim_Timestamp



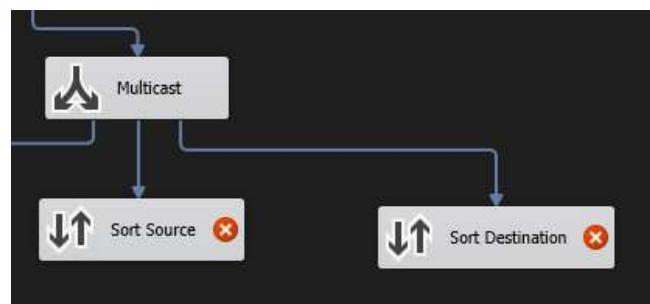
Bước 7: Vào mục “Mappings” để kiểm tra ánh xạ cột dữ liệu. Nhấn “OK” để hoàn tất

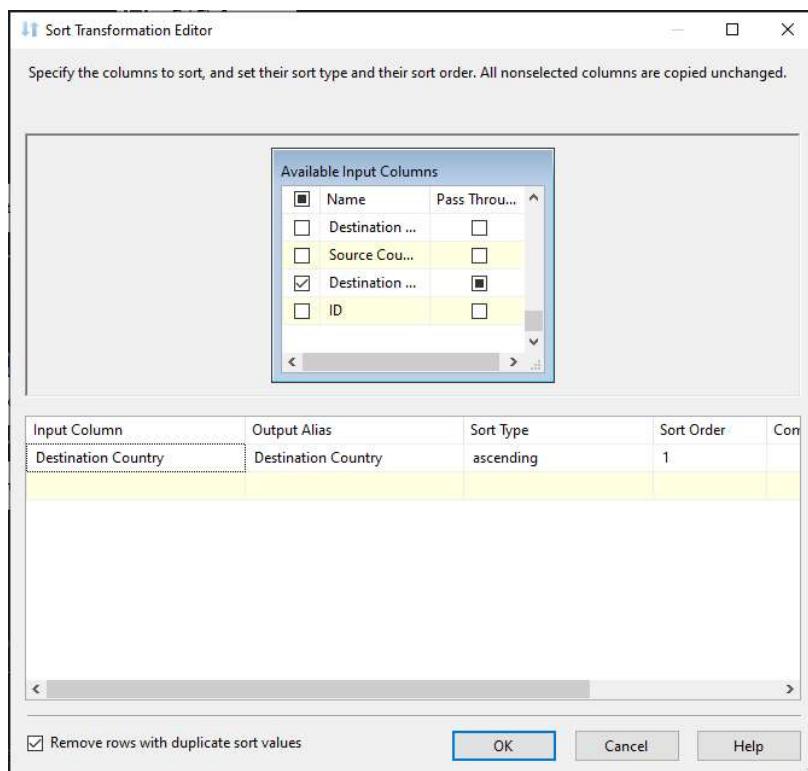
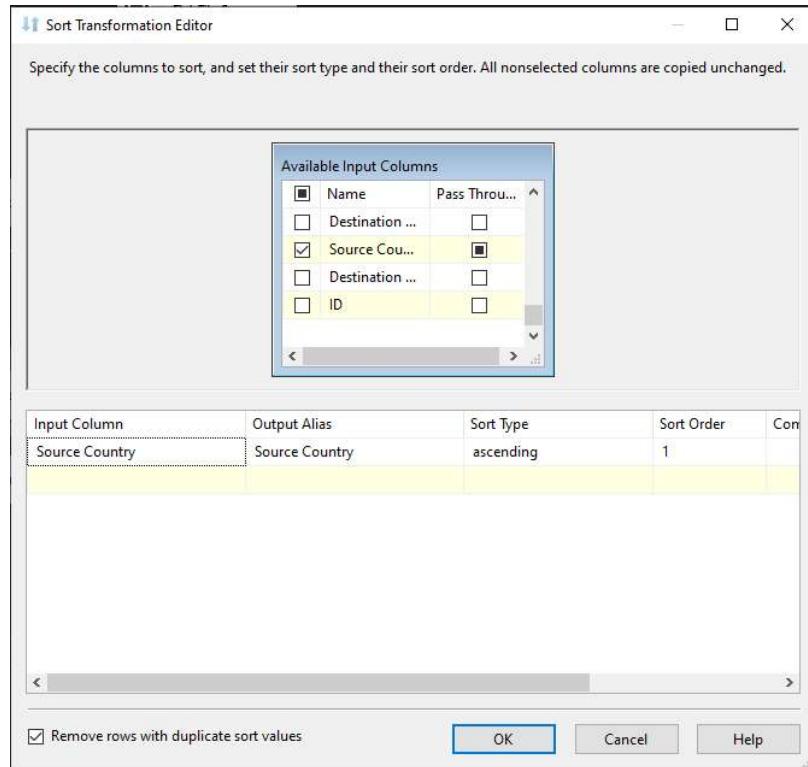


2.4.2. Bảng Dim_Address

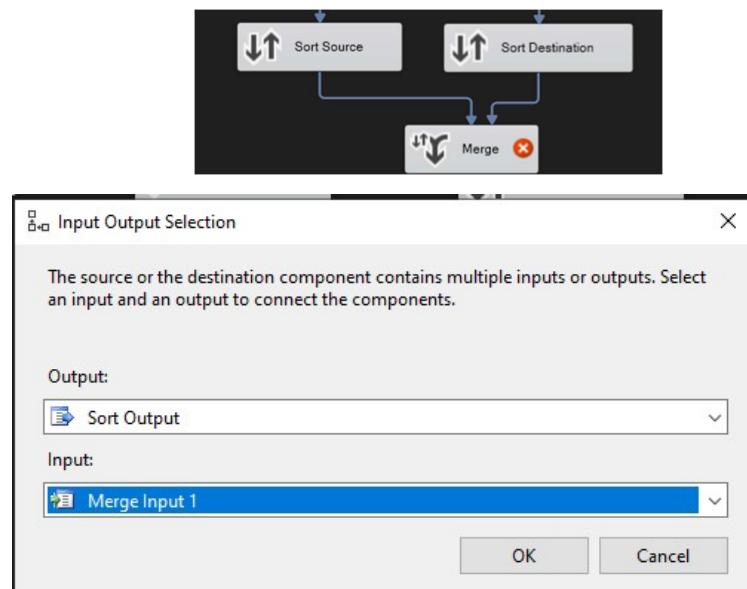
DIM_Address		
	AddressID	int
	Country	varchar

Bước 1: Tạo Sort Source và Sort Destination để lọc dữ liệu. Chọn xoá các dòng dữ liệu bị trùng.

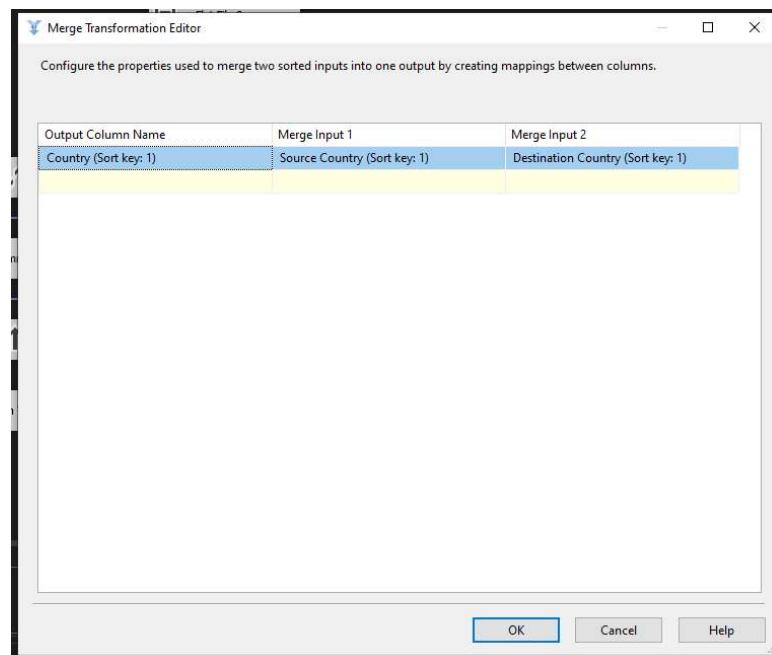




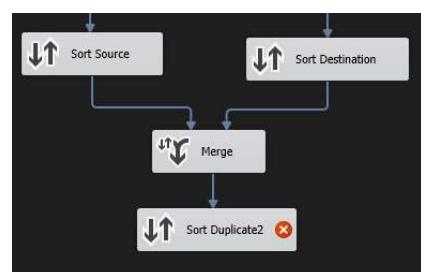
Bước 2: Chọn thành phần Merge để kết hợp 2 cột dữ liệu lại.

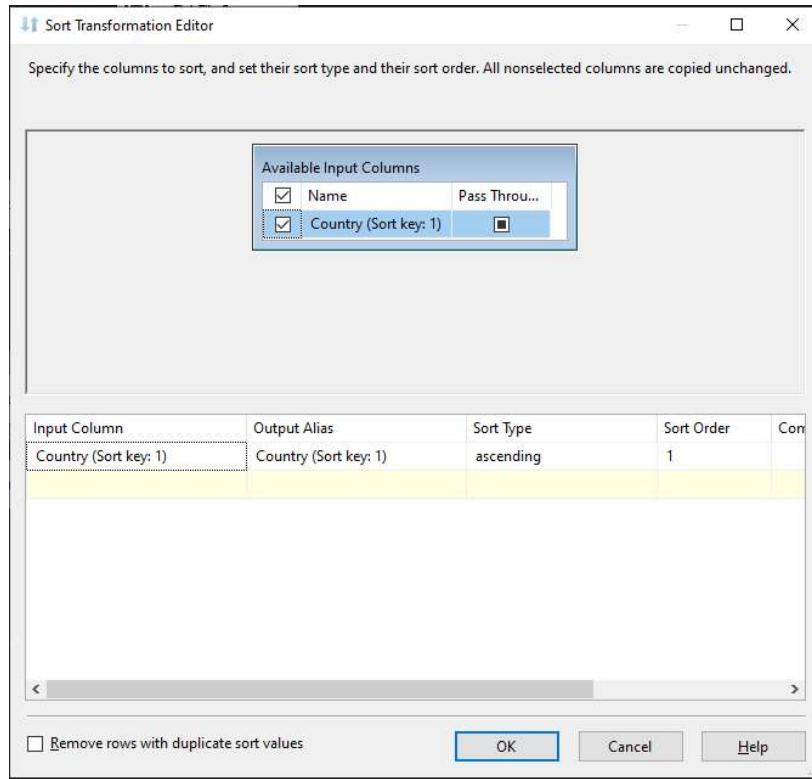


Chọn Merge Input 1 là Source Country và Merge Input 2 là Destination Country. Nhập tên Output là Country

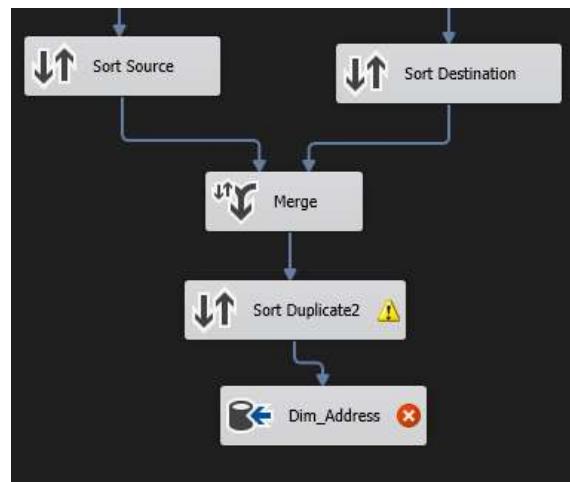


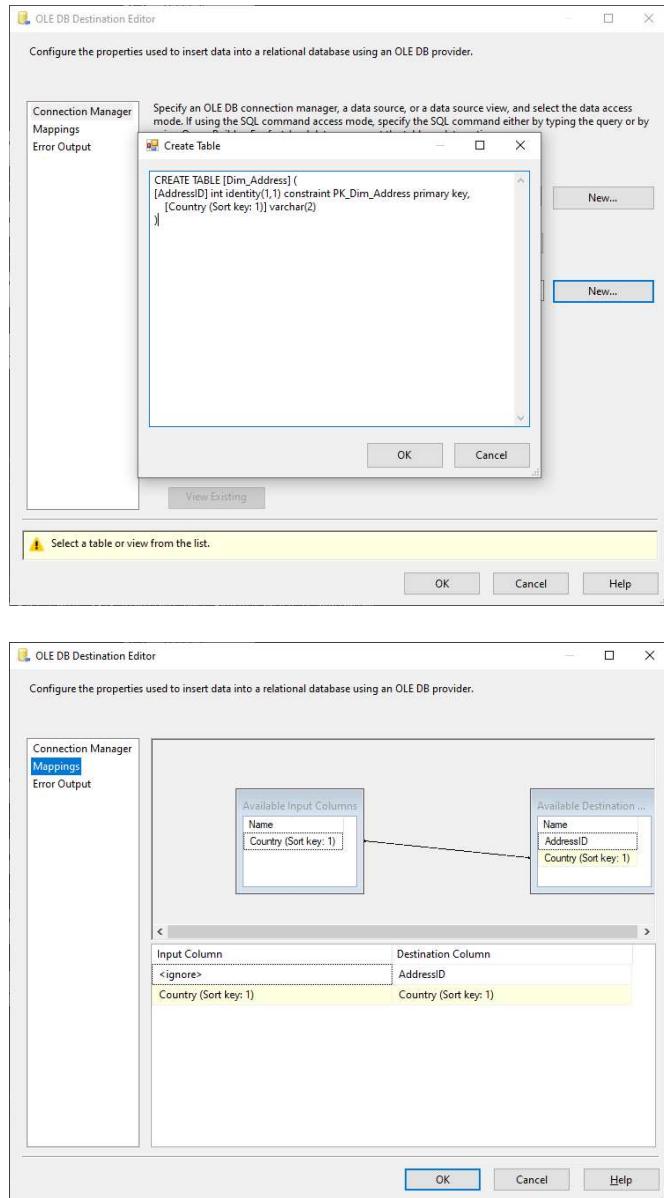
Bước 3: Tạo Sort để xoá các dòng dữ liệu trùng sau khi đã merge.





Bước 4: Tạo “OLE DB Destination” đặt tên là Dim_Address để lưu kết quả. Nhấn New và tạo mới table. Vào Mappings kiểm tra ánh xạ dữ liệu và nhấn OK.

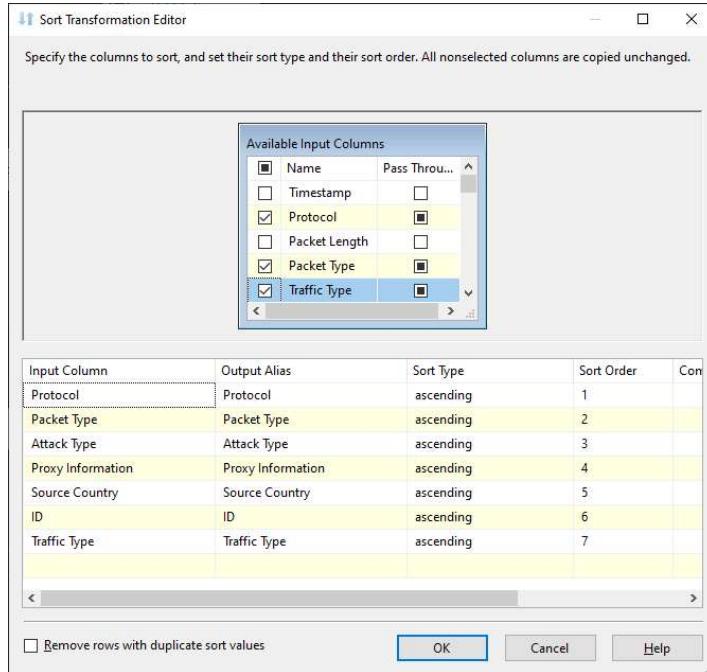
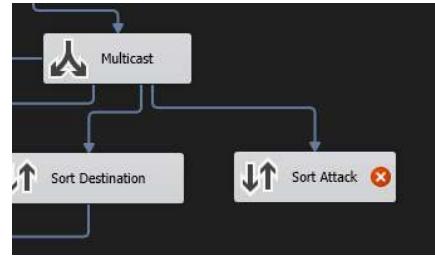




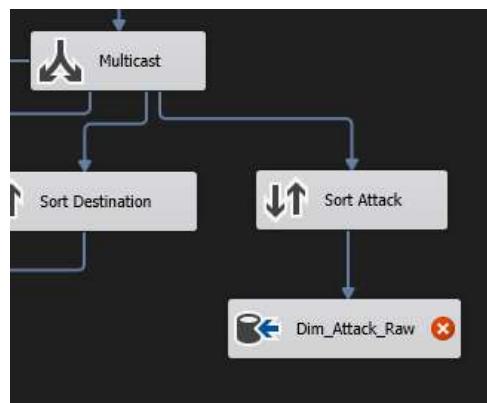
2.4.3. Bảng Dim_Attack

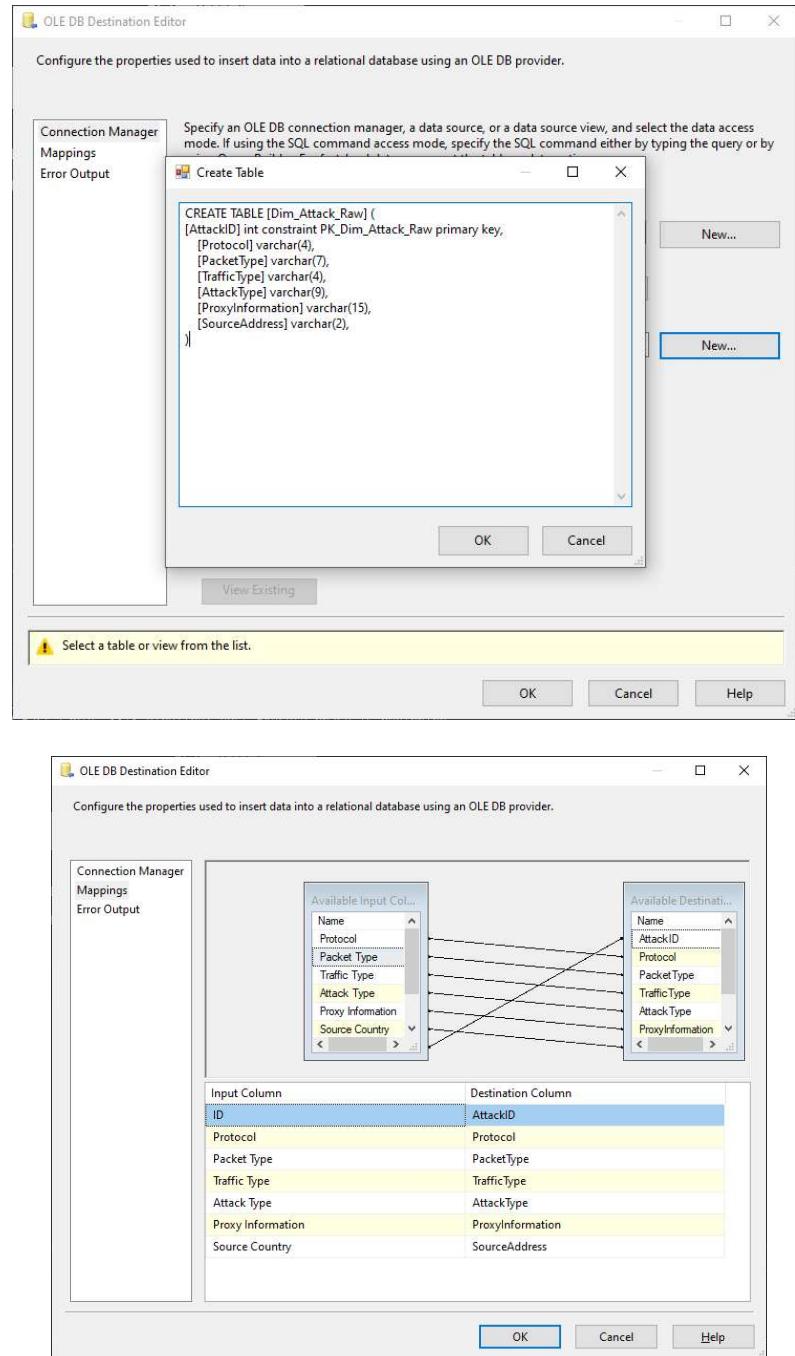
DIM_Attack	
AttackID	int
Protocol	varchar
PacketType	varchar
AttackType	varchar
TrafficType	varchar
ProxyInformation	varchar
SourceAddress	int

Bước 1: Tạo Sort để lọc các thuộc tính cần thiết cho Dim_Attack.

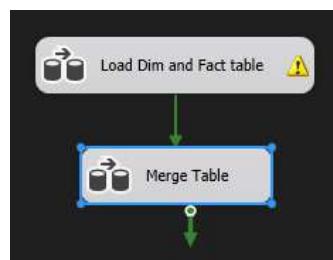


Bước 2: Tạo “OLE DB Destination” đặt tên là Dim_Attack_Raw để lưu dữ liệu. Nhấn New và tạo mới table. Vào Mappings kiểm tra ánh xạ dữ liệu và nhấn OK.

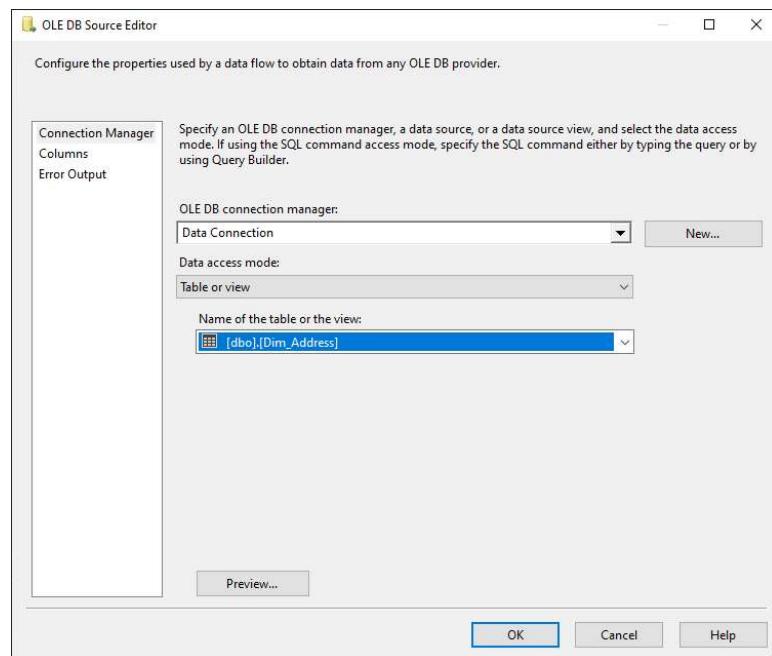
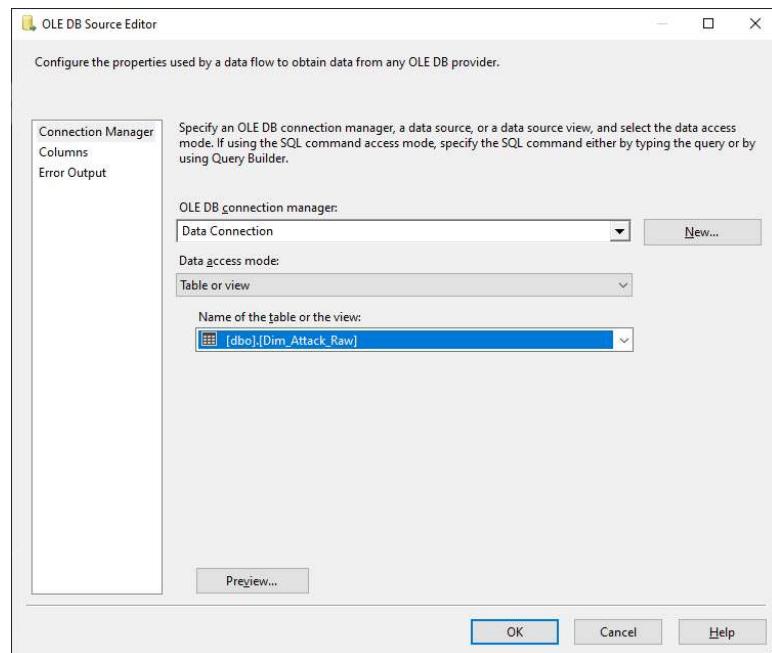




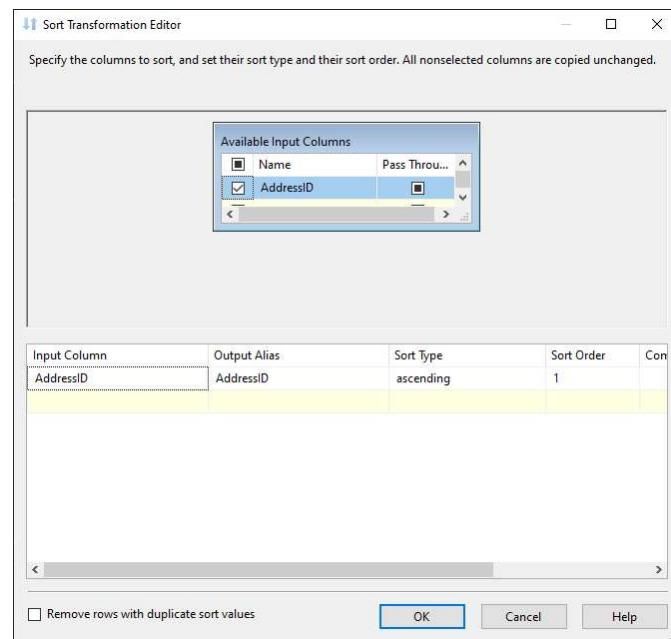
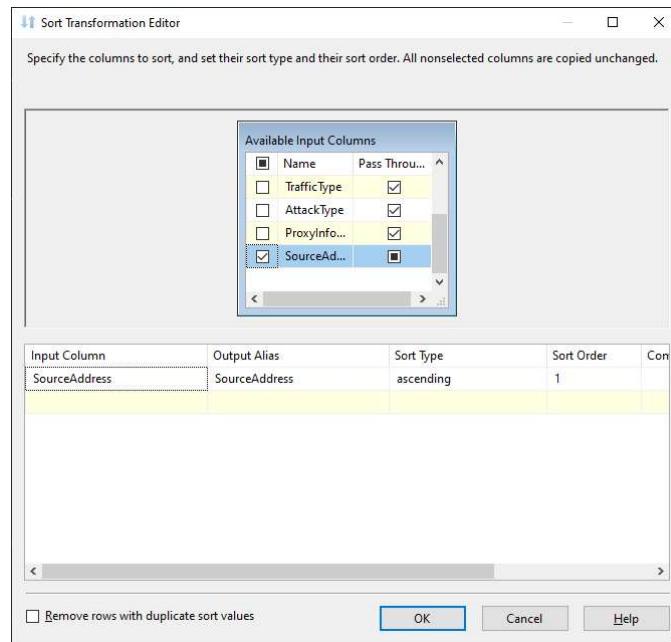
Bước 3: Thêm một thành phần “Data Flow Task” mới để tiến hành đổi SourceAddress lấy khoá ngoại ở bảng Dim_Address.



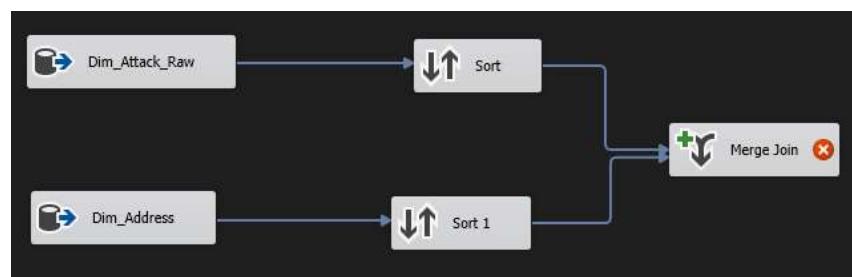
Bước 4: Trong Data Flow của “Merge table”, thêm 2 thành phần “OLE DB Source” của Dim_Attack_Raw và Dim_Address.

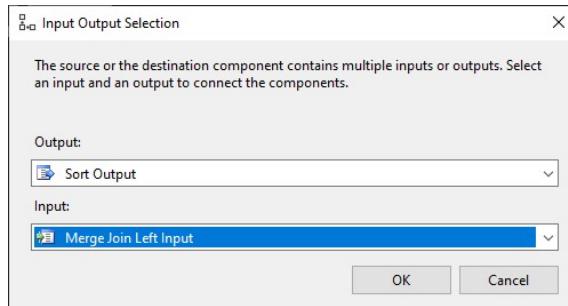


Bước 5: Tiến hành Sort từng table theo cột muốn đổi là SourceAddress và Country.

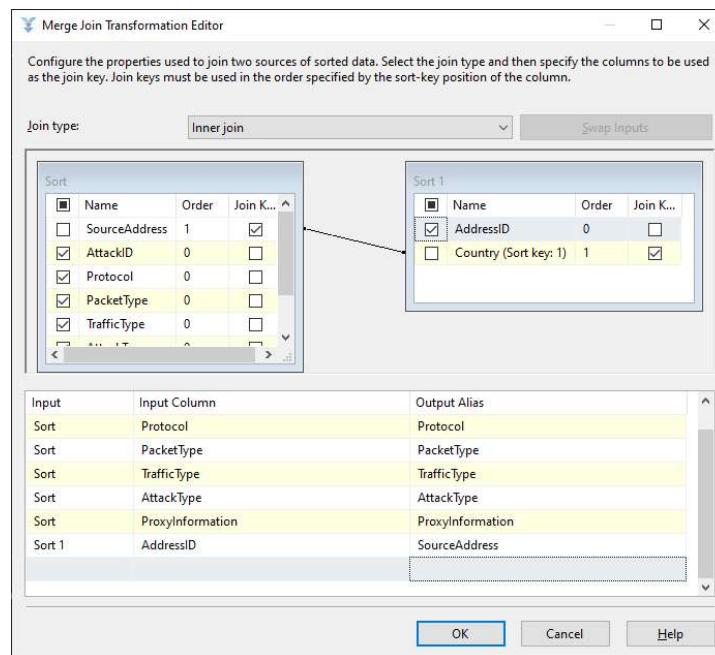


Bước 6: Thêm Look up để kết nối 2 nguồn dữ liệu bởi vì 2 bảng này không có chung cột thuộc tính nào.

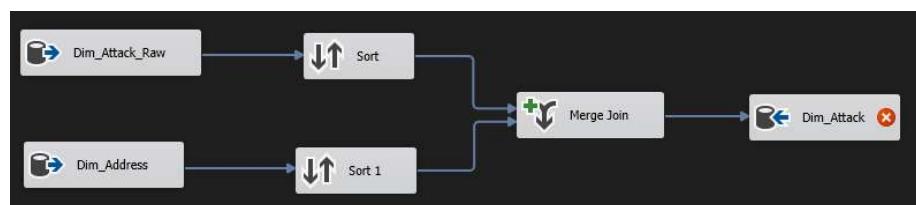


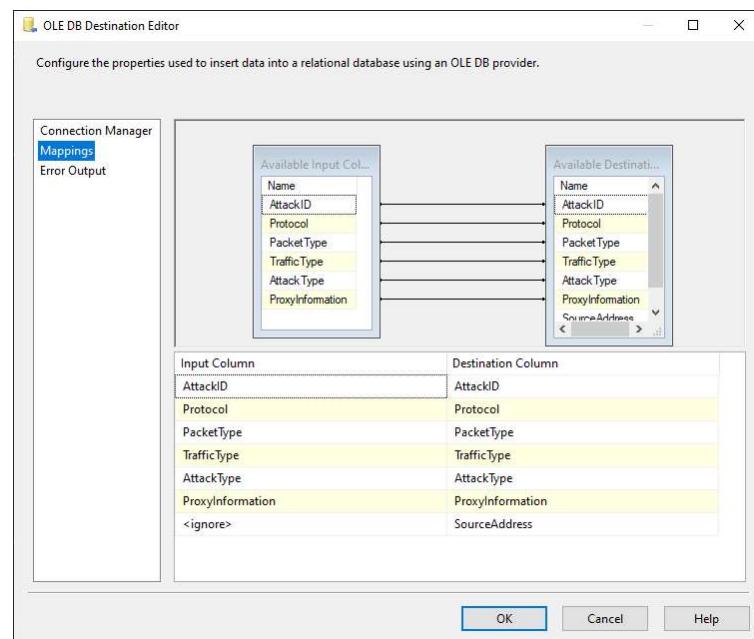
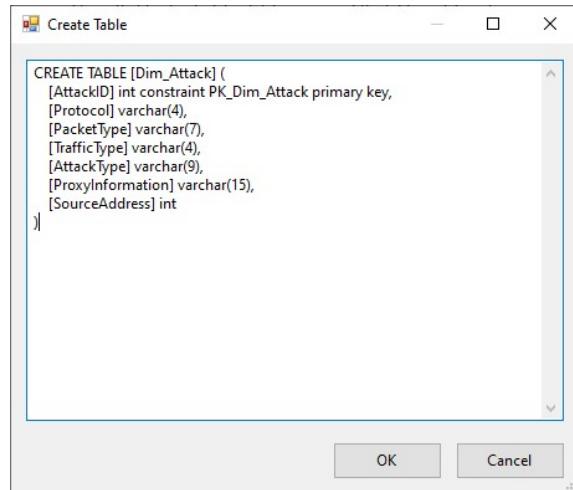


Bước 7: Trong Merge Join, ta chọn tất cả các cột của Sort trừ SourceAddress. Tiếp theo ta chọn AddressID của Sort1. Đặt tên cho Output của Sort1 là SourceAddress.



Bước 8: Thêm thành phần “OLE DB Destination” mới với tên là Dim_Attack để lưu lại kết quả. Tạo table Dim_Attack, kiểm tra Mappings và nhấn OK.

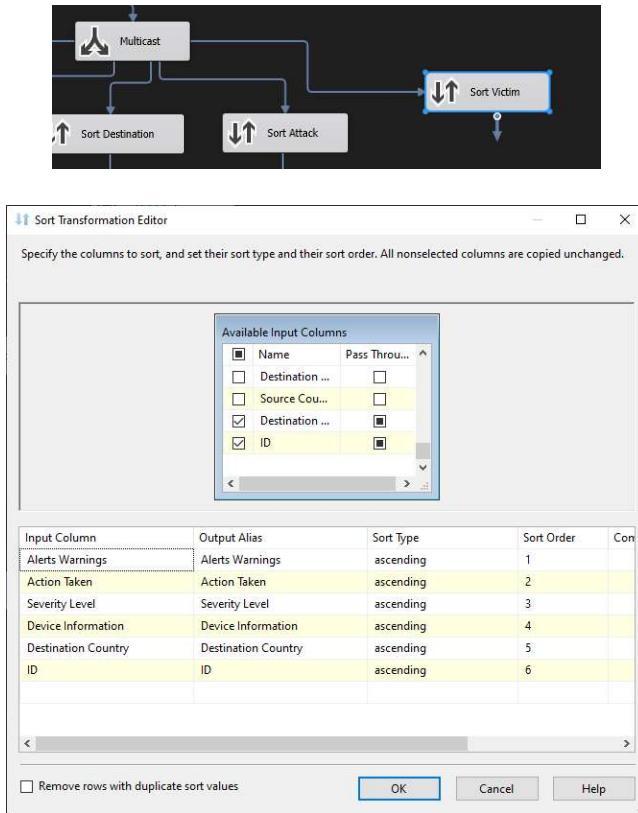




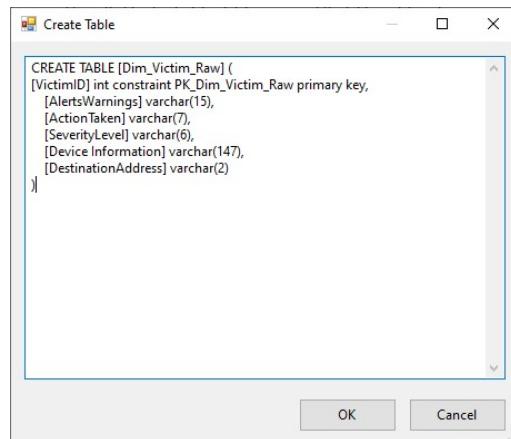
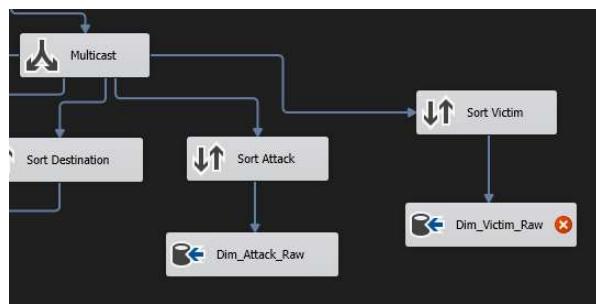
2.4.4. Bảng Dim_Victim

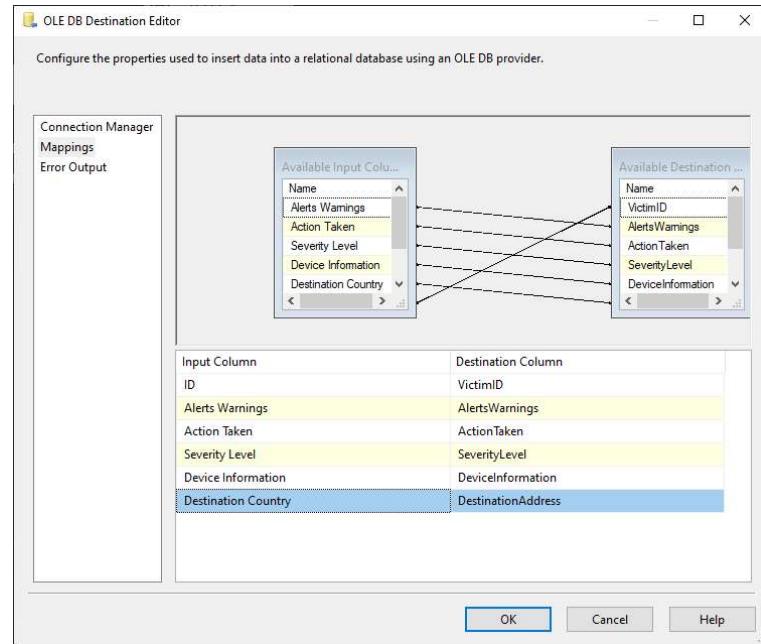
DIM_Victim	
VictimID	int
ActionTaken	varchar
SeverityLevel	varchar
AlertsWarnings	varchar
Deviceinformation	varchar
DestinationAddress	int

Bước 1: Tạo Sort để lọc các thuộc tính cần thiết cho Dim_Victim.

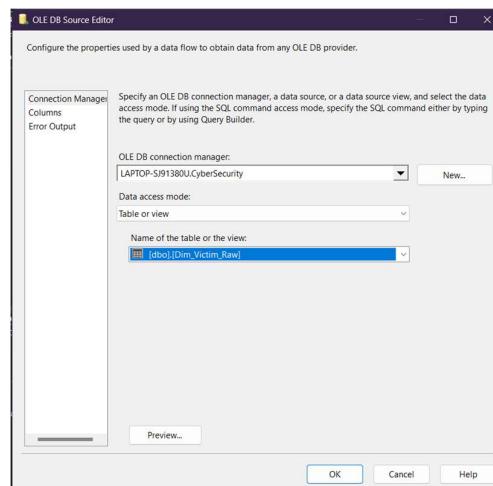


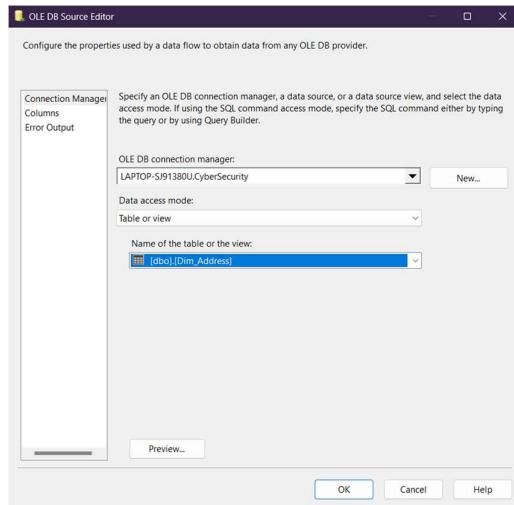
Bước 2: Tạo “OLE DB Destination” đặt tên là Dim_Victim_Raw để lưu dữ liệu. Nhấn New và tạo mới table. Vào Mappings kiểm tra ánh xạ dữ liệu và nhấn OK.



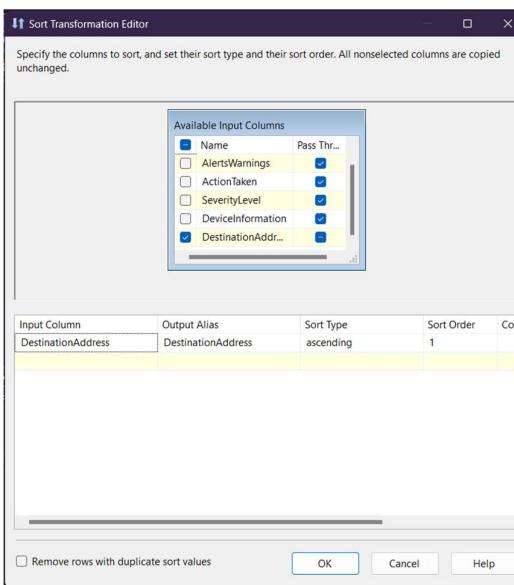


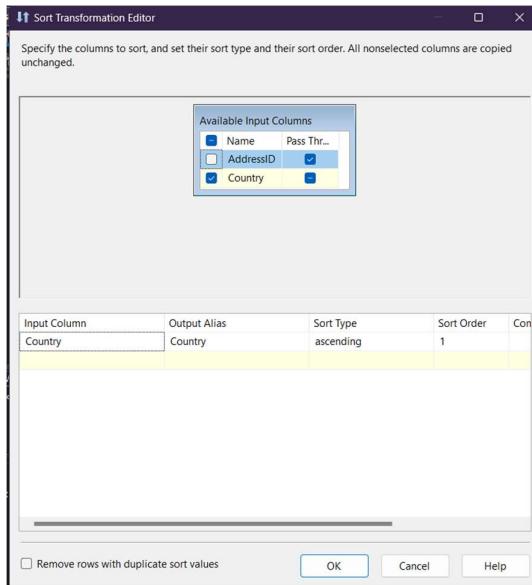
Bước 3: Trong Data Flow của “Merge table”, thêm 2 thành phần “OLE DB Source” của Dim_Victim_Raw và Dim_Address.



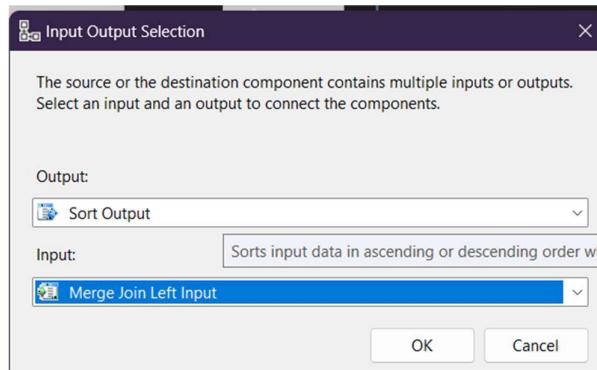
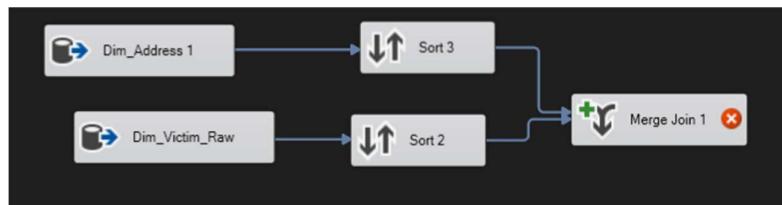


Bước 4: Tiến hành Sort từng table theo cột muốn đổi là DestinationAddress và Country.

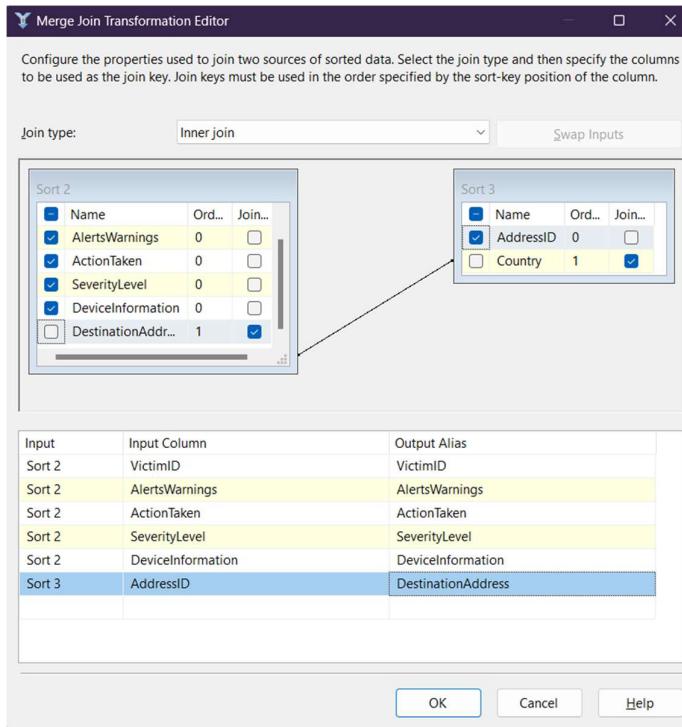




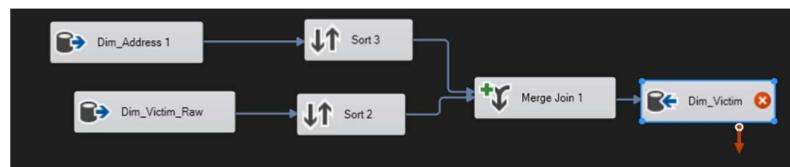
Bước 5: Thêm Merge Join, chọn Merge Join Left Input để giữ lại toàn bộ các dòng trong bảng Dim_Victim_Raw khi thực hiện phép kết trái với cột ID của bảng Dim_Address.



Bước 6: Trong Merge Join, ta chọn tất cả các cột của Sort2 trừ DestinationAddress. Tiếp theo ta chọn AddressID của Sort3. Đặt tên cho Output của Sort3 là DestinationAddress.



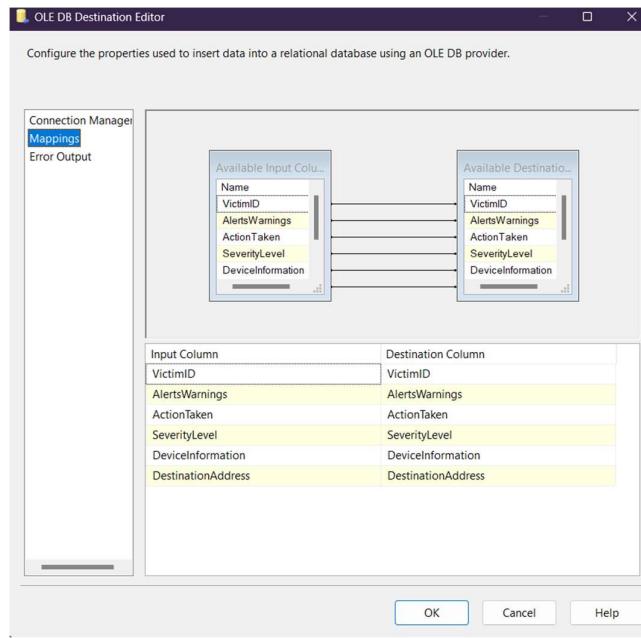
Bước 7: Thêm thành phần “OLE DB Destination” mới với tên là Dim_Victim để lưu lại kết quả. Tạo table Dim_Victim, kiểm tra Mappings và nhấn OK.



Create Table [Dim_Victim]

```

CREATE TABLE [Dim_Victim] (
    [VictimID] int constraint PK_Dim_Victim primary key,
    [AlertsWarnings] varchar(15),
    [ActionTaken] varchar(7),
    [SeverityLevel] varchar(6),
    [DeviceInformation] varchar(147),
    [DestinationAddress] int
)
  
```

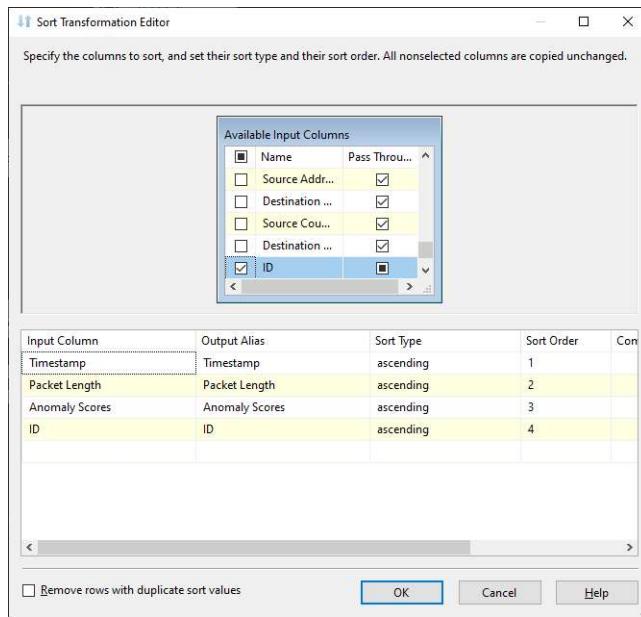


2.4.5. Bảng Fact_CyberSecurity

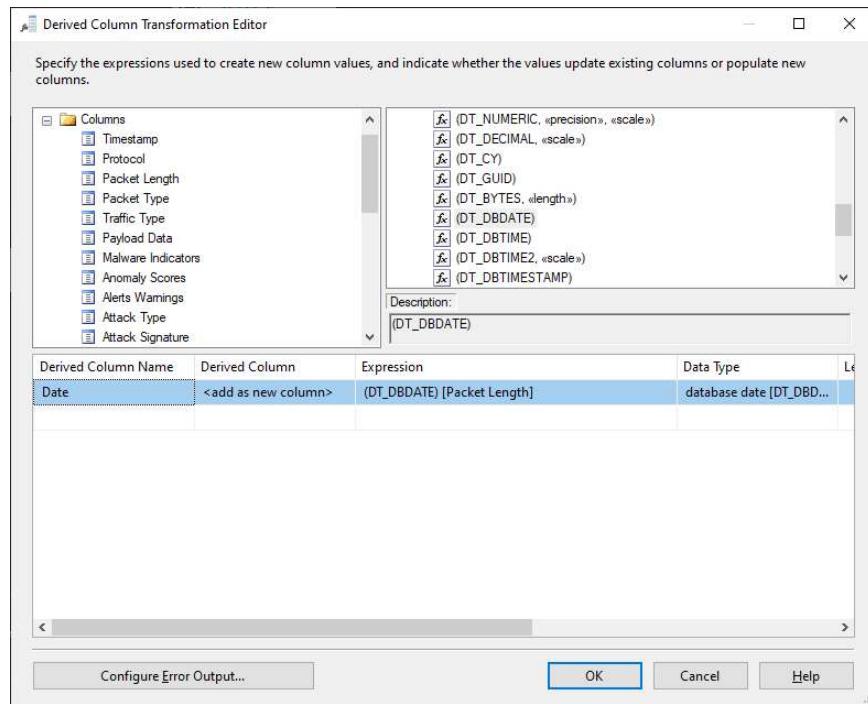
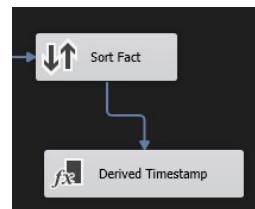
FactCyberSecurity	
- FactID	int
- TimeStampID	int
PacketLength	int
AnomalyScores	float

Bước 1: Tạo Sort để lọc các thuộc tính cần thiết cho Fact_CyberSecurity.

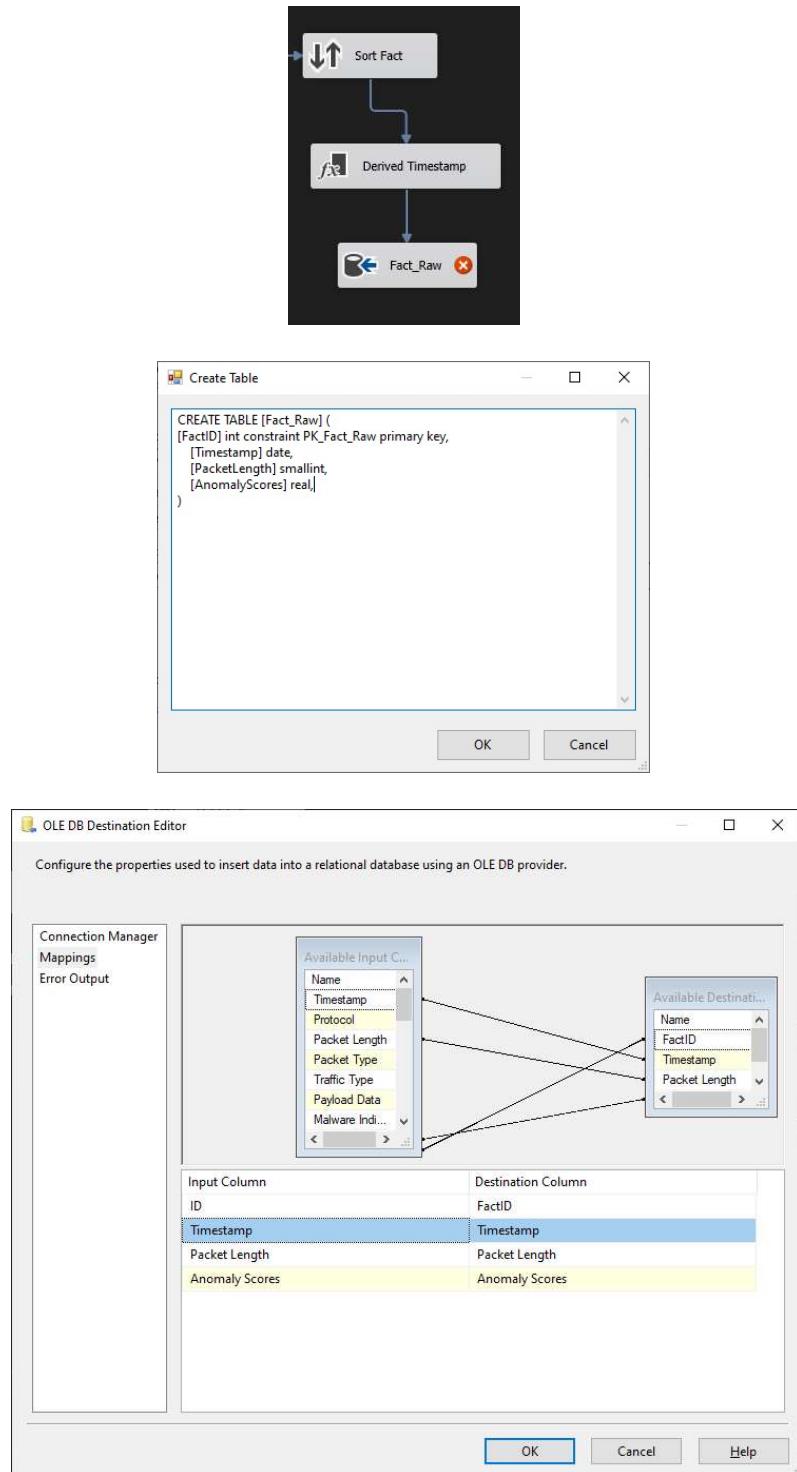




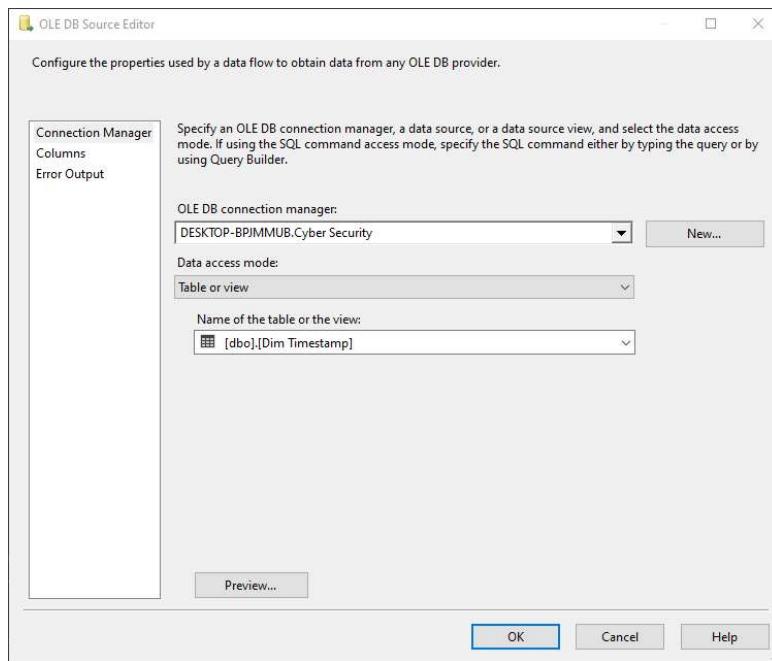
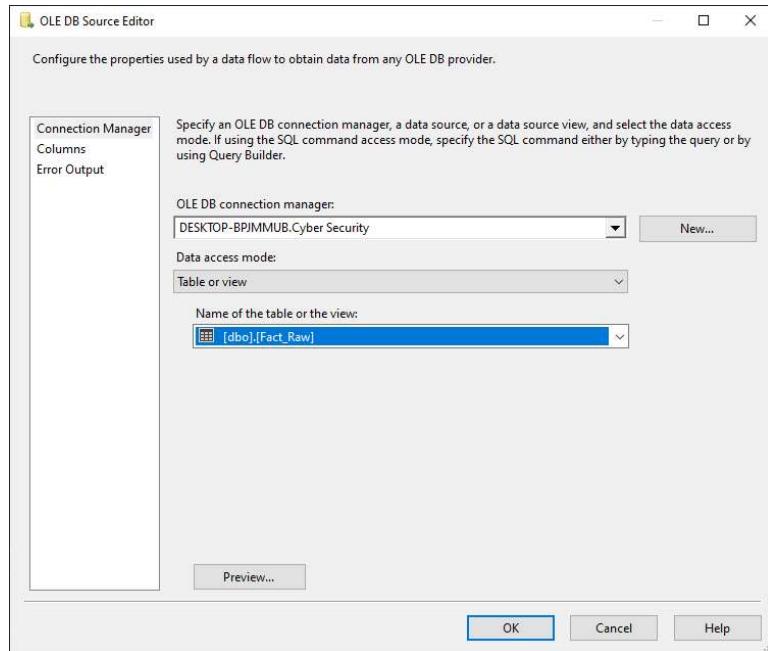
Bước 2: Thêm thành phần Derived Column để ép kiểu dữ liệu Timestamp thành Date hỗ trợ cho việc đổi dữ liệu lấy khoá ngoại với Dim_Timestamp.



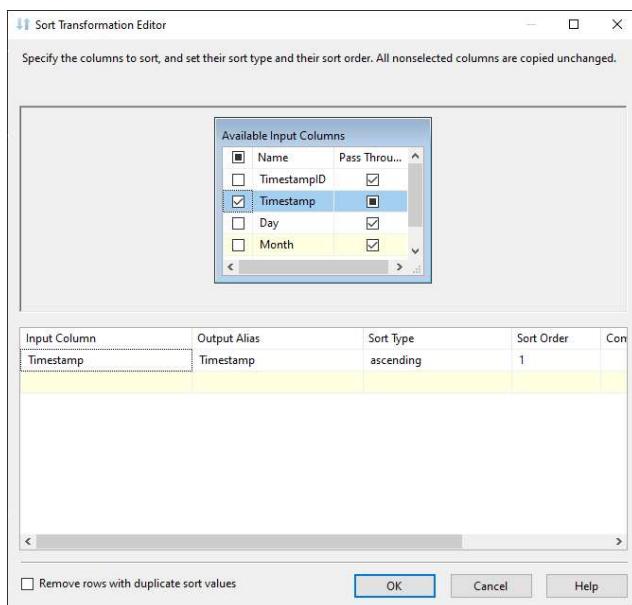
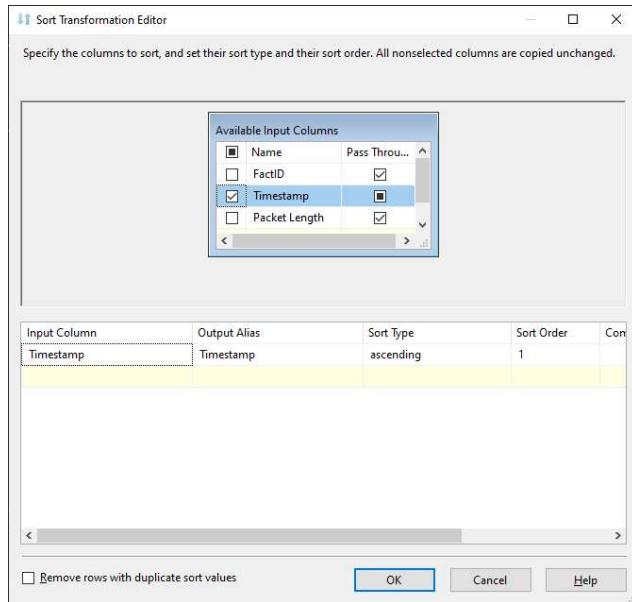
Bước 3: Tạo “OLE DB Destination” đặt tên là Fact_Raw để lưu dữ liệu. Nhấn New và tạo mới table. Vào Mappings kiểm tra ánh xạ dữ liệu và nhấn OK.



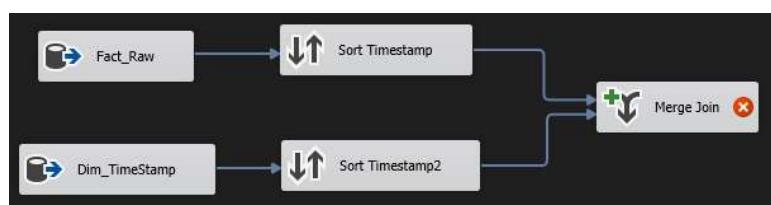
Bước 4: Trong Data Flow của “Merge table”, thêm 2 thành phần “OLE DB Source” của Fact_Raw và Dim_Timestamp.

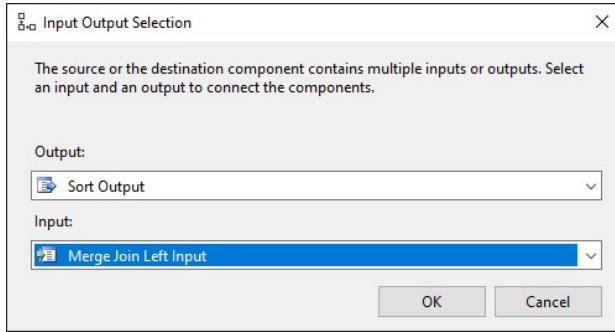


Bước 5: Tiến hành Sort từng table theo cột muốn đổi là Timestamp.

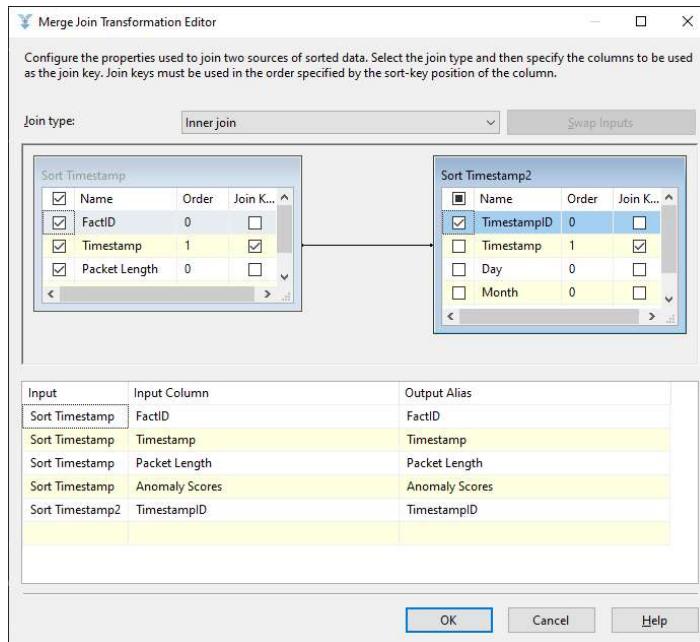


Bước 6: Thêm Merge Join, chọn Merge Join Left Input để giữ lại toàn bộ các dòng trong bảng Fact_Raw khi thực hiện phép kết trái với cột ID của bảng Dim_Timestamp

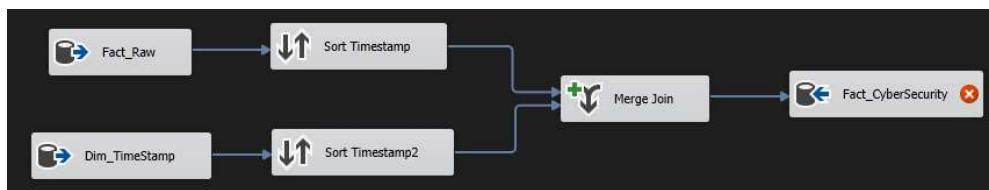


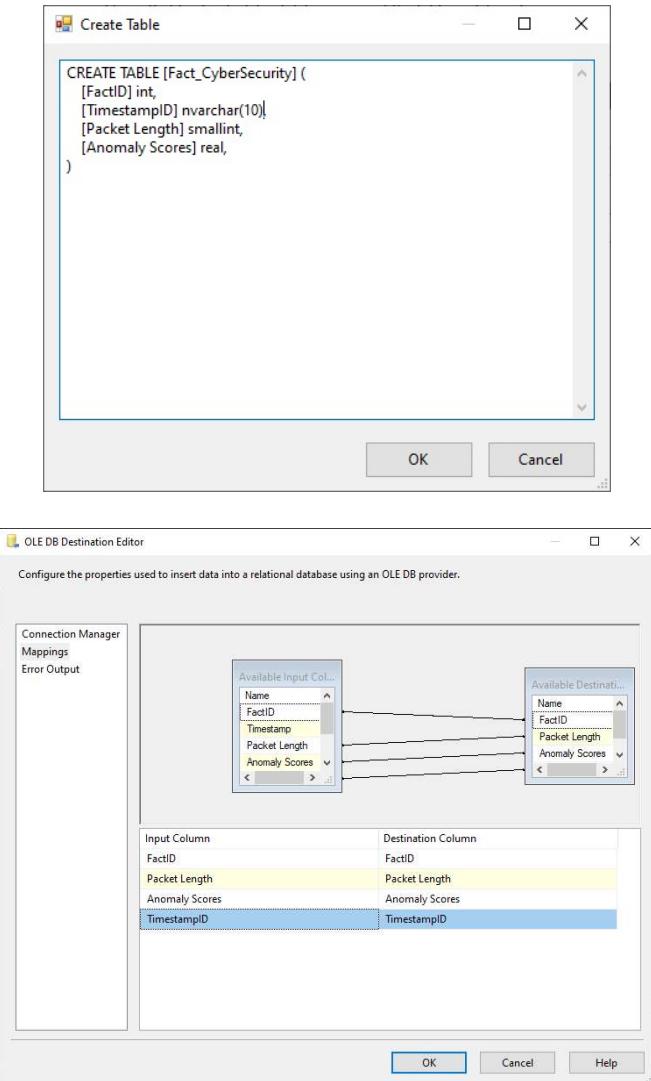


Bước 7: Trong Merge Join, ta chọn tất cả các cột của Sort4 trừ Timestamp. Tiếp theo ta chọn Timestamp ID của Sort5.



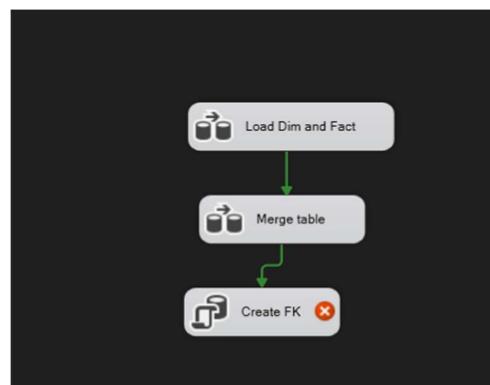
Bước 8: Thêm thành phần “OLE DB Destination” mới với tên là Fact_CyberSecurity để lưu lại kết quả. Tạo table Fact_CyberSecurity, kiểm tra Mappings và nhấn OK.



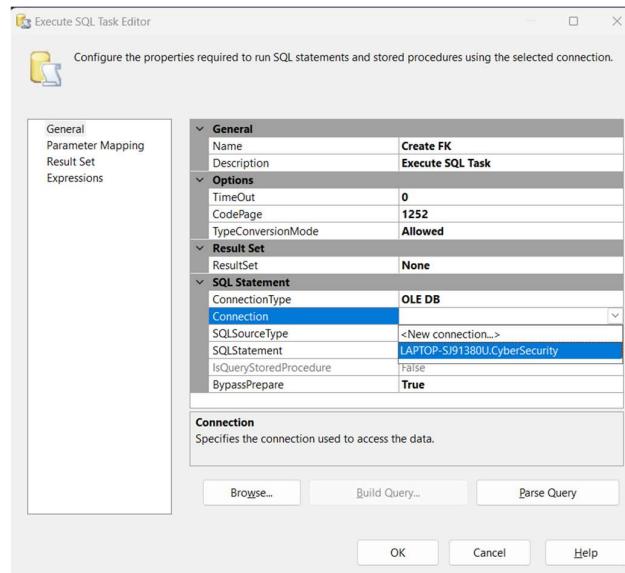


2.4.6. Tạo khoá ngoại

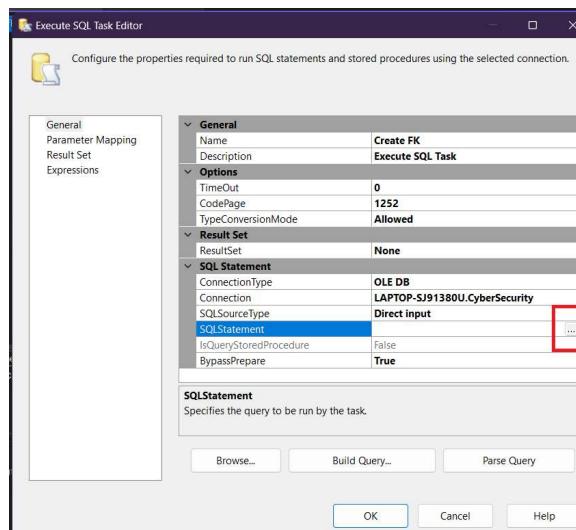
Bước 1: Thêm một “Execute SQL Task” dưới Merge table để tạo khoá ngoại. Đặt tên là Create FK.

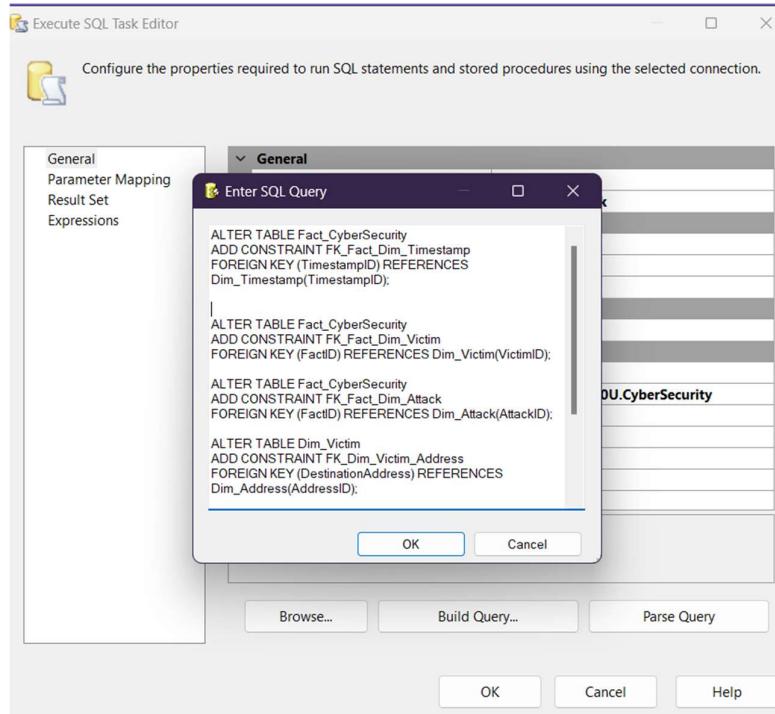


Bước 2: Nhấn chuột phải vào Create FK và chọn Edit. Chọn connection đã được thiết lập đến database.



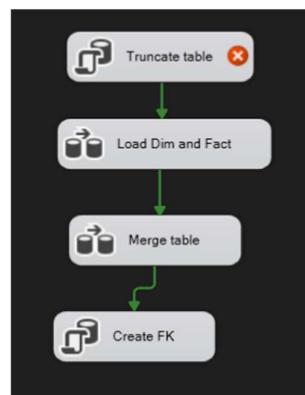
Bước 3: Nhấn vào nút ba chấm ở SQLStatement và viết câu lệnh SQL để tạo khoá ngoại cho các table. Nhấn OK



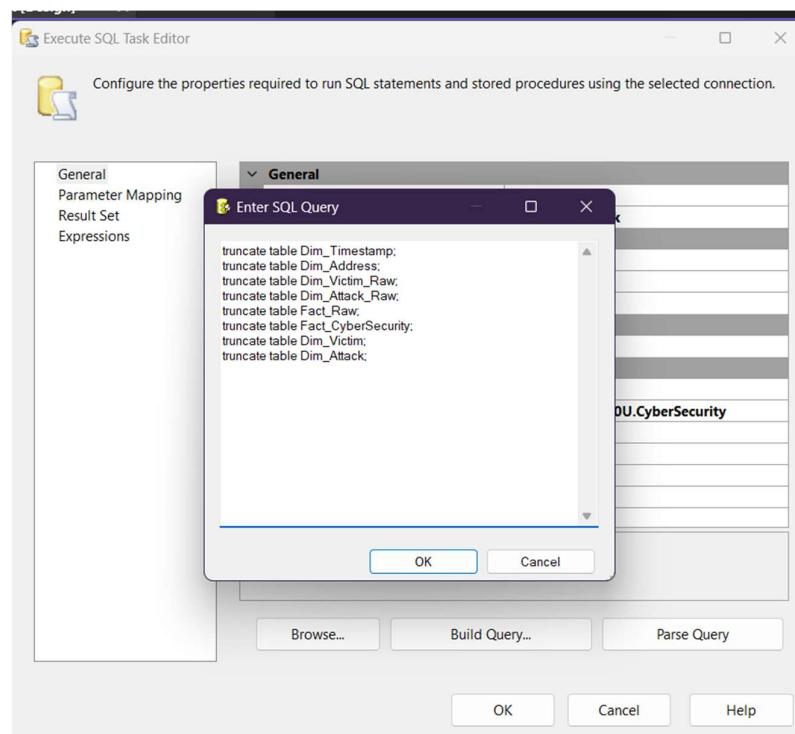
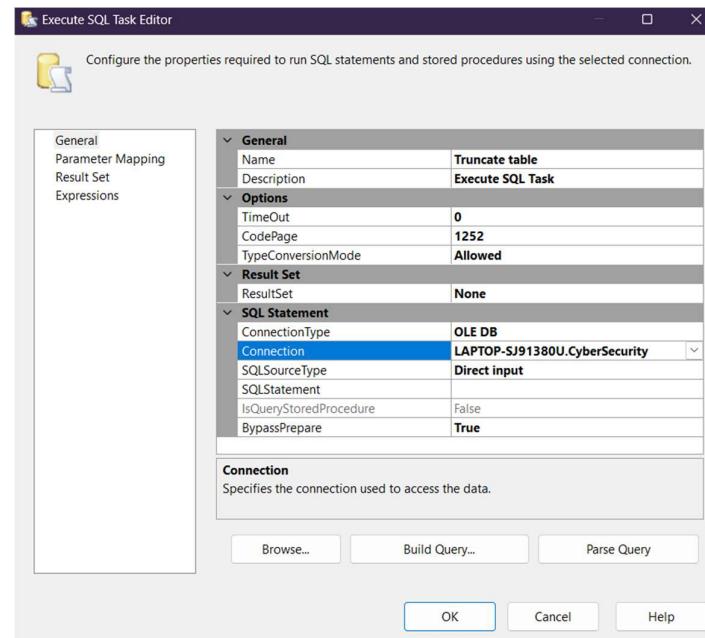


2.4.7. Chạy SSIS

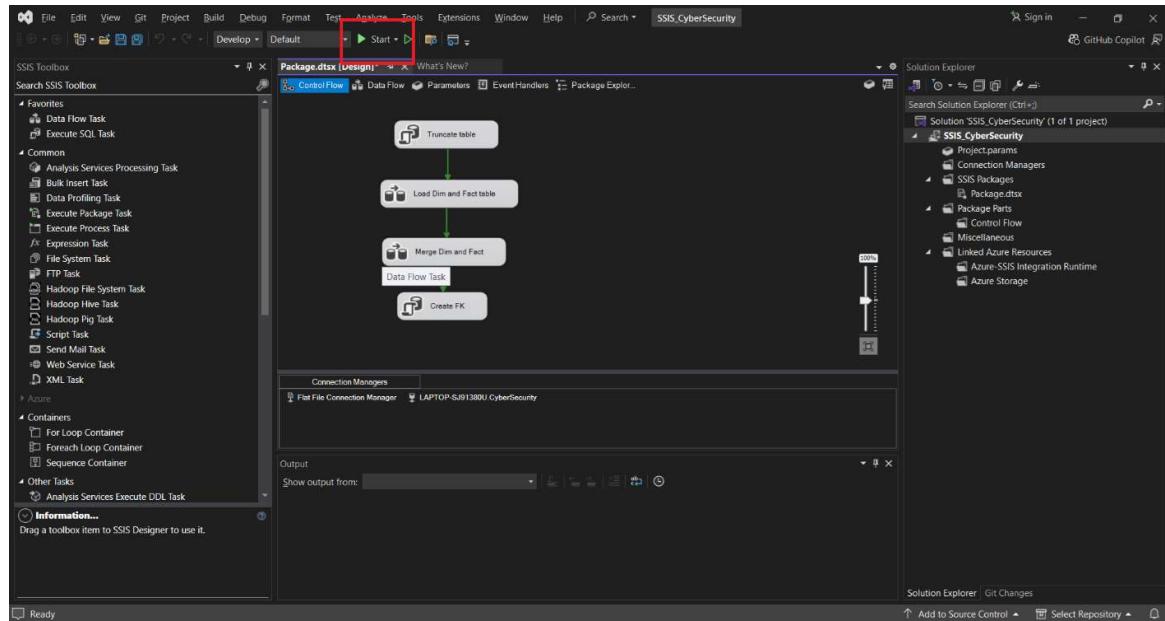
Bước 1: Thêm vào một “Execute SQL Task” ở trước Load Dim and Fact để làm sạch dữ liệu trong table tránh trường hợp dữ liệu bị chồng chéo mỗi lần chạy lại project.



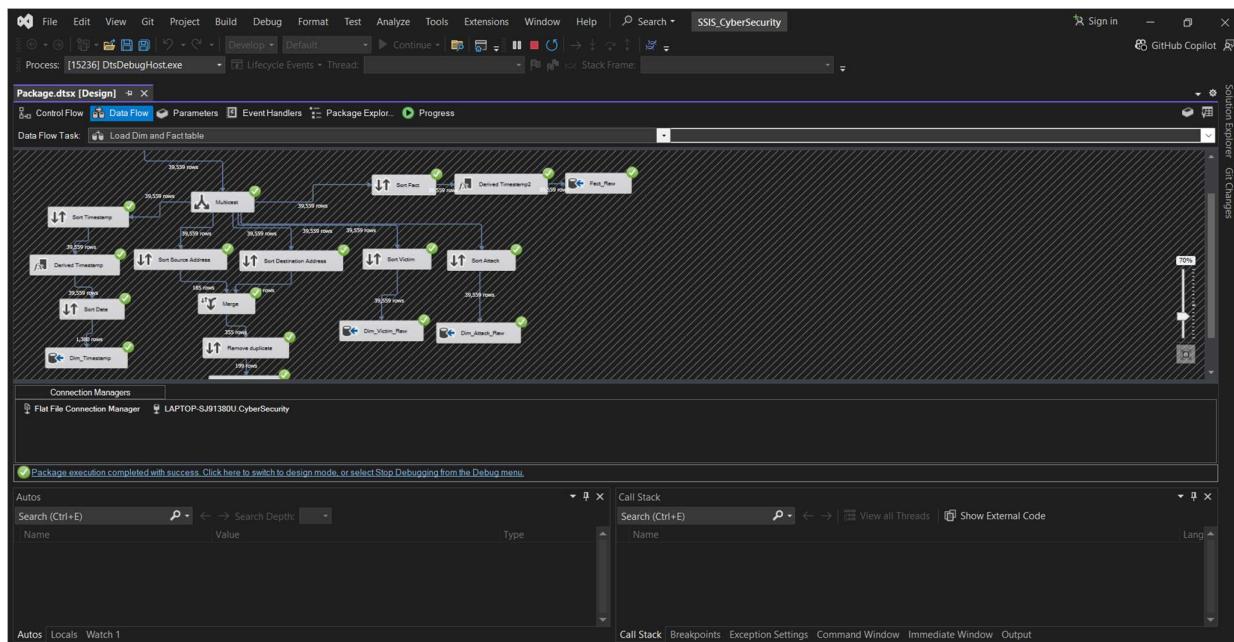
Bước 2: Chọn connection đã được kết nối và viết các câu lệnh SQL để làm sạch table.

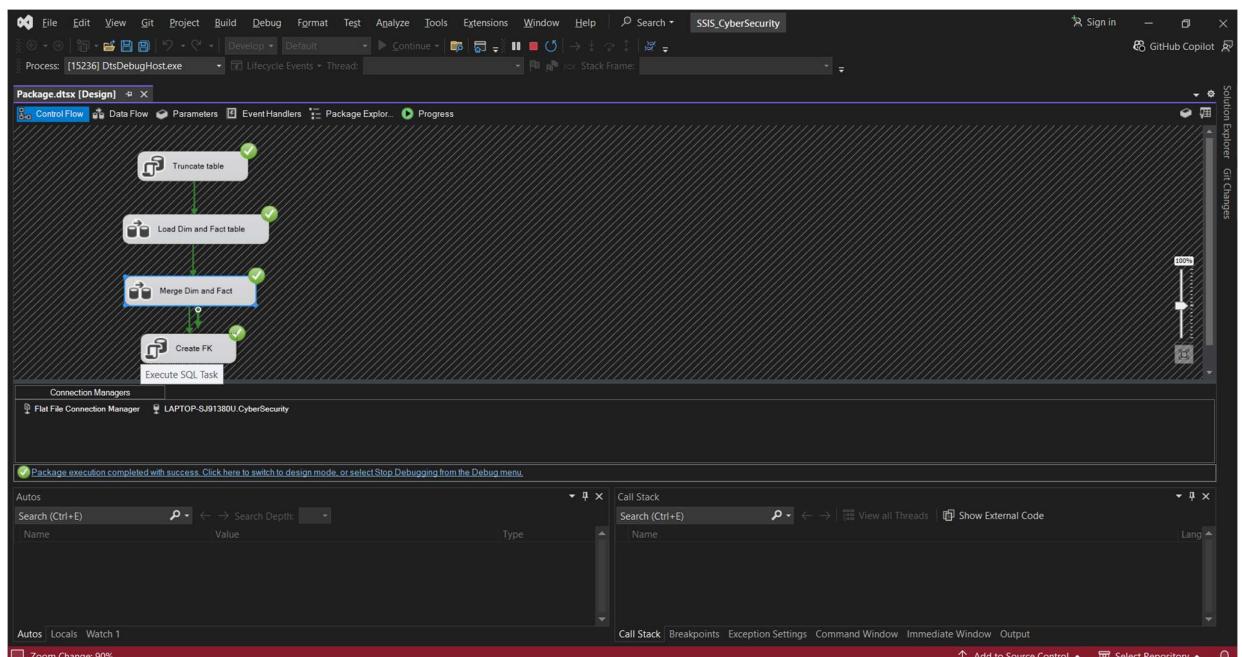
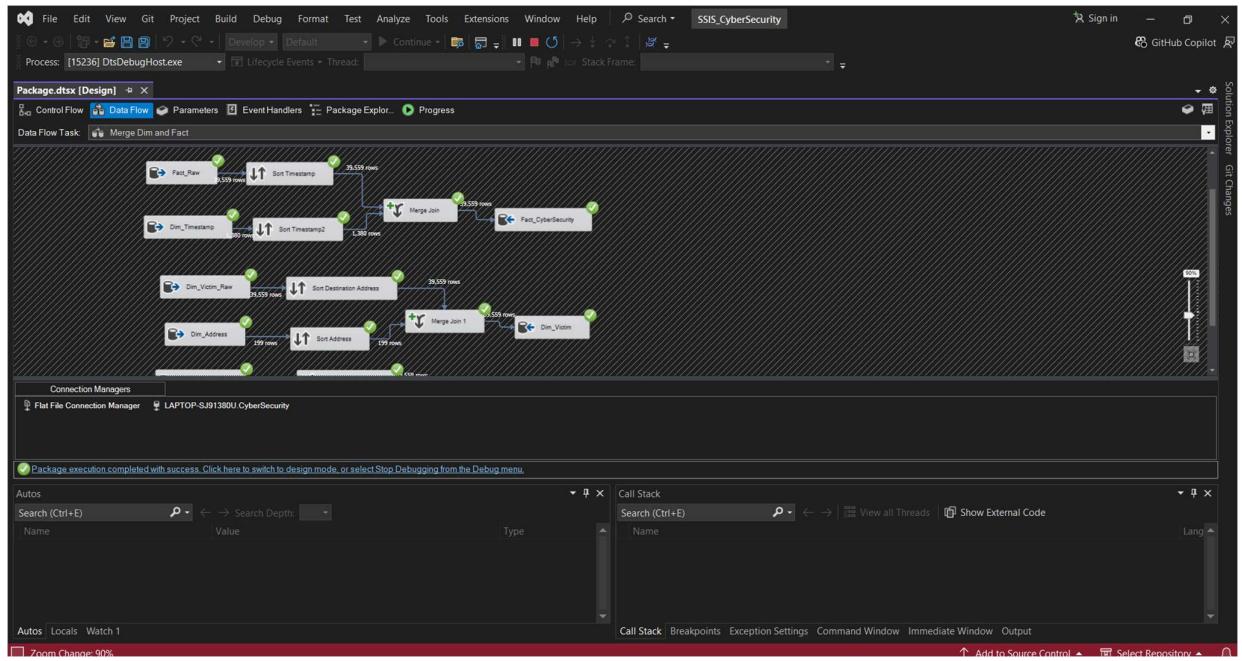


Bước 3: Nhấn nút Start trên thanh công cụ để chạy dự án.

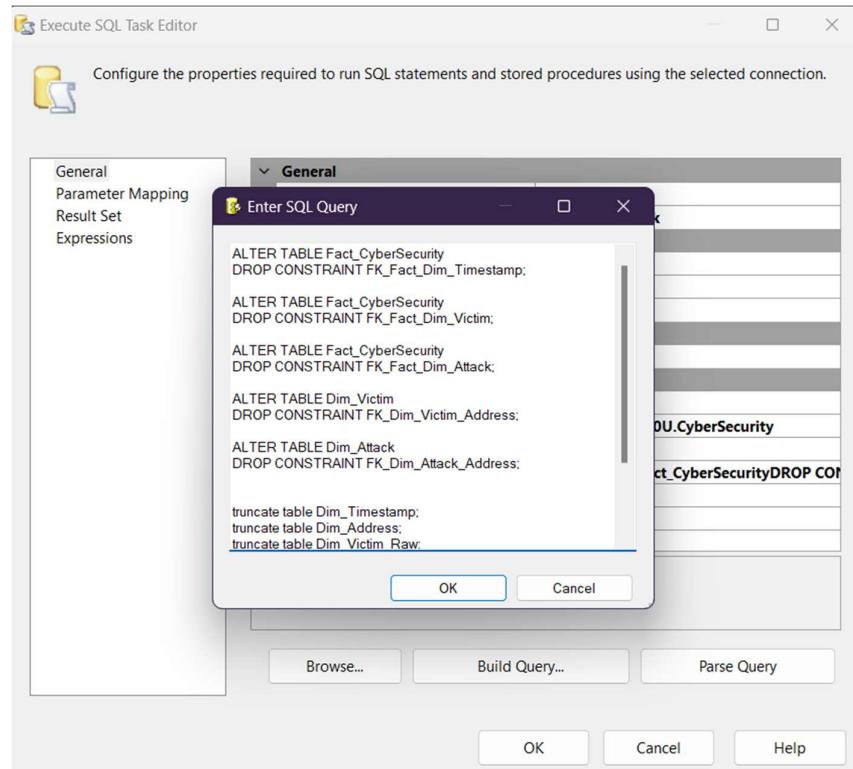


Kết quả chạy project:





Bước 4: Sau khi chạy project lần đầu tiên, các khóa ngoại sẽ được tạo ra. Để tránh việc vi phạm các ràng buộc khi xoá dữ liệu của table, ta sẽ thêm các câu lệnh SQL để xoá khóa ngoại trước khi xoá dữ liệu trong table trong phần Truncate table.



2.5. Kiểm tra dữ liệu các bảng

Dữ liệu bảng Dim_Timestamp:

	TimestampID	Timestamp	Day	Month	Year
1	1	2020-01-01	1	1	2020
2	2	2021-01-01	1	1	2021
3	3	2022-01-01	1	1	2022
4	4	2023-01-01	1	1	2023
5	5	2020-02-01	1	2	2020
6	6	2021-02-01	1	2	2021
7	7	2022-02-01	1	2	2022
8	8	2023-02-01	1	2	2023
9	9	2020-03-01	1	3	2020
10	10	2021-03-01	1	3	2021
11	11	2022-03-01	1	3	2022
12	12	2023-03-01	1	3	2023
13	13	2020-04-01	1	4	2020
14	14	2021-04-01	1	4	2021
15	15	2022-04-01	1	4	2022
16	16	2023-04-01	1	4	2023
17	17	2020-05-01	1	5	2020
18	18	2021-05-01	1	5	2021
19	19	2022-05-01	1	5	2022
20	20	2023-05-01	1	5	2023
21	21	2020-06-01	1	6	2020
22	22	2021-06-01	1	6	2021
23	23	2022-06-01	1	6	2022
24	24	2023-06-01	1	6	2023
25	25	2020-07-01	1	7	2020
26	26	2021-07-01	1	7	2021
27	27	2022-07-01	1	7	2022
28	28	2023-07-01	1	7	2023
29	29	2020-08-01	1	8	2020
30	30	2021-08-01	1	8	2021
31	31	2022-08-01	1	8	2022
32	32	2023-08-01	1	8	2023
33	33	2020-09-01	1	9	2020

Dữ liệu bảng Dim_Address:

	AddressID	Country
1	1	AD
2	2	AE
3	3	AF
4	4	AG
5	5	AL
6	6	AM
7	7	AO
8	8	AR
9	9	AT
10	10	AU
11	11	AW
12	12	AX
13	13	AZ
14	14	BA
15	15	BB
16	16	BD
17	17	BE
18	18	BF
19	19	BG
20	20	BH
21	21	BI
22	22	BJ
23	23	BN
24	24	BO
25	25	BQ
26	26	BR
27	27	BS
28	28	BT

Dữ liệu bảng Dim_Attack:

	AttackID	Protocol	PacketType	TrafficType	AttackType	ProxyInformation	SourceAddress
1	1	ICMP	Data	HTTP	Malware	150.97.135	39
2	2	ICMP	Data	HTTP	Malware	No detected	60
3	3	UDP	Control	HTTP	DDoS	114.133.48.179	188
4	4	UDP	Data	HTTP	Malware	No detected	91
5	5	TCP	Data	DNS	DDoS	149.6.110.119	188
6	6	UDP	Data	HTTP	Malware	No detected	188
7	7	TCP	Data	DNS	DDoS	No detected	188
8	8	ICMP	Data	DNS	Intrusion	192.31.159.5	188
9	9	TCP	Control	FTP	Intrusion	No detected	83
10	10	UDP	Data	HTTP	Malware	87.128.245.244	176
11	11	ICMP	Data	HTTP	Malware	29.161.99.247	26
12	12	TCP	Control	HTTP	Malware	No detected	161
13	13	ICMP	Control	DNS	Intrusion	No detected	183
14	14	ICMP	Data	HTTP	Malware	59.131.15.72	55
15	15	UDP	Data	DNS	Malware	155.121.88.187	188
16	16	TCP	Data	DNS	Intrusion	No detected	188
17	17	TCP	Control	HTTP	DDoS	No detected	188
18	18	TCP	Data	HTTP	DDoS	20.252.145.34	156
19	19	TCP	Control	DNS	DDoS	No detected	39
20	20	TCP	Data	HTTP	DDoS	106.203.136.99	188
21	21	ICMP	Control	DNS	Malware	71.135.96.151	39
22	22	UDP	Control	DNS	Malware	No detected	188
23	23	UDP	Data	DNS	DDoS	169.32.12.215	80
24	24	UDP	Data	HTTP	Malware	50.108.129.83	161
25	25	TCP	Control	DNS	DDoS	No detected	188
26	26	UDP	Data	DNS	No detected	No detected	188

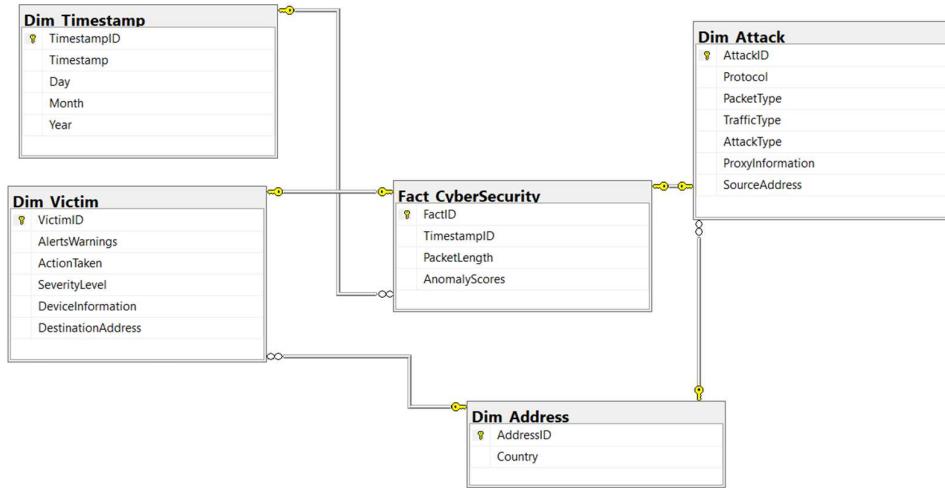
Dữ liệu bảng Dim_Victim:

	VictimID	AlertsWarnings	ActionTaken	SeverityLevel	DeviceInformation	DestinationAddress
1	1	No Alert	Logged	Low	Mozilla/5.0 (compatible; MSIE 8.0; Windows NT 6.2; Trident/5.0)	62
2	2	No Alert	Blocked	Low	Mozilla/5.0 (compatible; MSIE 8.0; Windows NT 6.2; Trident/4.0)	188
3	3	Alert Triggered	Ignored	Low	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.2; Trident/5.0)	188
4	4	Alert Triggered	Blocked	Medium	Mozilla/5.0 (Macintosh; PPC Mac OS X_10_11_5; rv:1.9.6.20) Gecko/2583-02-14 13:30:10 Firefox/11.0	39
5	5	Alert Triggered	Blocked	Low	Mozilla/5.0 (compatible; MSIE 5.0; Windows NT 6.2; Trident/3.0)	128
6	6	No Alert	Logged	Medium	Opera/8.58.(X11; Linux i686; fa-IR) Presto/2.9.175 Version/12.00	46
7	7	No Alert	Ignored	High	Opera/9.24.(X11; Linux i686; fa-IR) Presto/2.9.175 Version/10.00	135
8	8	Alert Triggered	Logged	High	Mozilla/5.0 (Macintosh; U; PPC Mac OS X_10_7_6) AppleWebKit/534.2 (KHTML, like Gecko) Chrome/45.0.865.0 Safari/534.2	182
9	9	Alert Triggered	Blocked	High	Mozilla/5.0 (Macintosh; U; PPC Mac OS X_10_5_8) AppleWebKit/536.1 (KHTML, like Gecko) Chrome/38.0.861.0 Safari/536.1	188
10	10	Alert Triggered	Blocked	Medium	Mozilla/5.0 (Windows; U; Windows NT 6.0) AppleWebKit/533.28.5 (KHTML, like Gecko) Version/4.0 Safari/533.28.5	188
11	11	Alert Triggered	Ignored	Medium	Mozilla/5.0 (Pod; U; CPU iPhone OS 4_3 like Mac OS X; zh-ZA) AppleWebKit/534.25.1 (KHTML, like Gecko) Version/4.0.5 Mobile/8B111 Safari/6534.25.1	54
12	12	No Alert	Ignored	Low	Opera/8.38.(X11; Linux x86_64; pt-BR) Presto/2.9.175 Version/12.00	39
13	13	Alert Triggered	Blocked	Medium	Opera/8.54.(Windows NT 6.0; tg-TJ) Presto/2.9.170 Version/12.00	188
14	14	No Alert	Ignored	Low	Mozilla/5.0 (compatible; MSIE 7.0; Windows NT 10.0; Trident/4.1)	39
15	15	Alert Triggered	Logged	High	Mozilla/5.0 (X11; Linux i686; rv:1.9.7.20) Gecko/5845-10-22 09:50:59 Firefox/3.8	134
16	16	Alert Triggered	Logged	Medium	Opera/8.85.(X11; Linux x86_64; sk-SK) Presto/2.9.163 Version/11.00	188
17	17	Alert Triggered	Ignored	High	Opera/8.72.(Windows NT 6.1; ia-FR) Presto/2.9.160 Version/12.00	32
18	18	No Alert	Logged	Low	Mozilla/5.0 (X11; Linux i686) AppleWebKit/536.0 (KHTML, like Gecko) Chrome/46.0.873.0 Safari/536.0	188
19	19	Alert Triggered	Ignored	High	Mozilla/5.0 (compatible; MSIE 6.0; Windows NT 10.0; Trident/4.0)	39
20	20	No Alert	Ignored	High	Mozilla/5.0 (iPhone; CPU iPhone OS 12_4_8 like Mac OS X) AppleWebKit/533.2 (KHTML, like Gecko) FxiOS/18.14735.0 Mobile/92M384 Safari/533.2	39
21	21	Alert Triggered	Blocked	Medium	Mozilla/5.0 (Macintosh; U; PPC Mac OS X_10_11_1 rv:4.0; en-US) AppleWebKit/533.41.7 (KHTML, like Gecko) Version/4.0.4 Safari/533.41.7	188
22	22	No Alert	Blocked	Low	Mozilla/5.0 (Windows; U; Windows NT 5.0.01) AppleWebKit/531.13.4 (KHTML, like Gecko) Version/4.0.4 Safari/531.13.4	197
23	23	No Alert	Logged	High	Mozilla/5.0 (compatible; MSIE 5.0; Windows 95; Trident/5.1)	188
24	24	No Alert	Logged	Medium	Mozilla/5.0 (compatible; MSIE 0.0; Windows NT 5.0; Trident/5.1)	188

Dữ liệu bảng Fact_CyberSecurity:

	FactID	TimestampID	PacketLength	AnomalyScores
1	1	1329	503	28.67
2	2	1165	1174	51.5
3	3	593	306	87.42
4	4	74	385	15.79
5	5	714	1462	0.52
6	6	1265	1423	5.76
7	7	705	379	31.55
8	8	514	1022	54.05
9	9	1205	1281	56.34
10	10	671	224	16.51
11	11	893	661	24.91
12	12	1159	281	86.07
13	13	1342	64	74.2
14	14	553	1341	62.14
15	15	8	913	72.65
16	16	677	425	91.56
17	17	792	838	43.83
18	18	919	969	40.46
19	19	1311	124	44.75
20	20	1206	461	13.19
21	21	712	845	89.99
22	22	1283	410	86.91
23	23	1054	1425	72.25
24	24	677	118	96.27
25	25	270	1386	67.43
26	26	270	433	89.86
27	27	1356	531	87.23
28	28	1255	282	44.4
29	29	1286	1487	45.42
30	30	1328	563	25.7

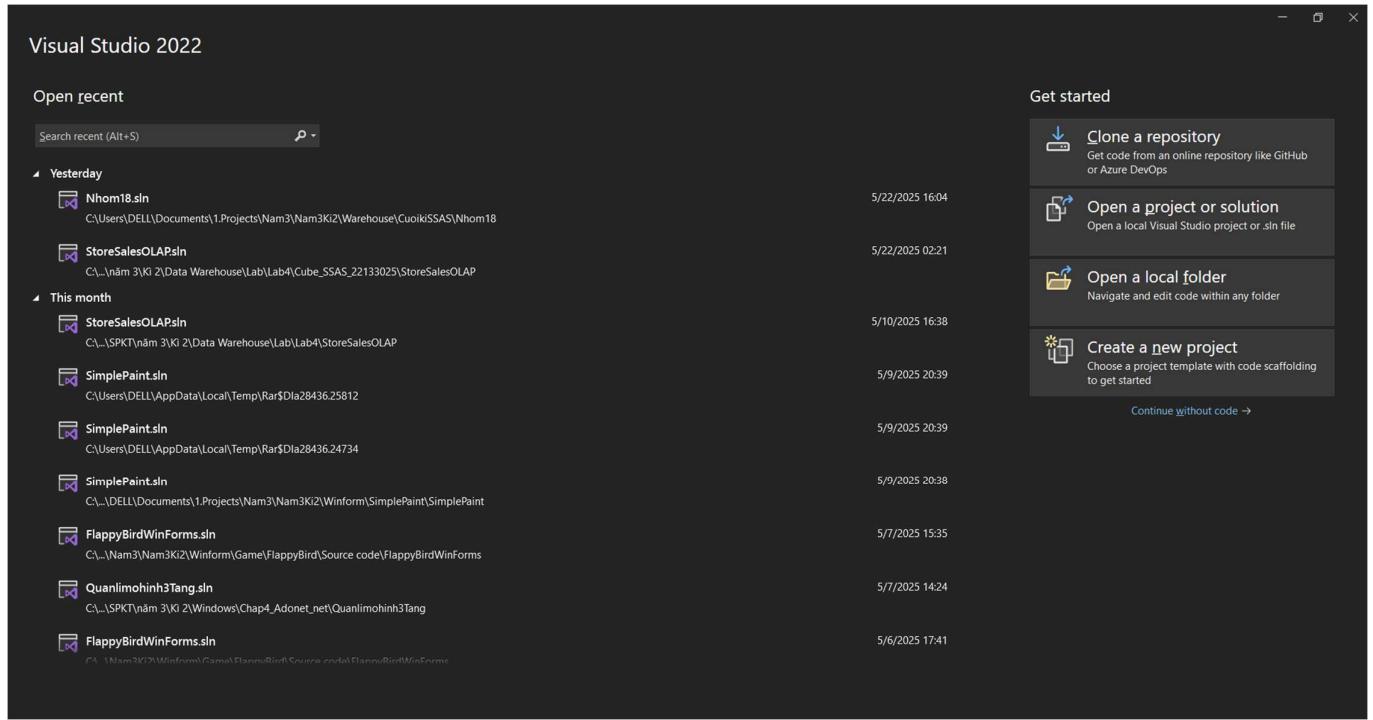
2.6. Lược đồ sau khi hoàn thành



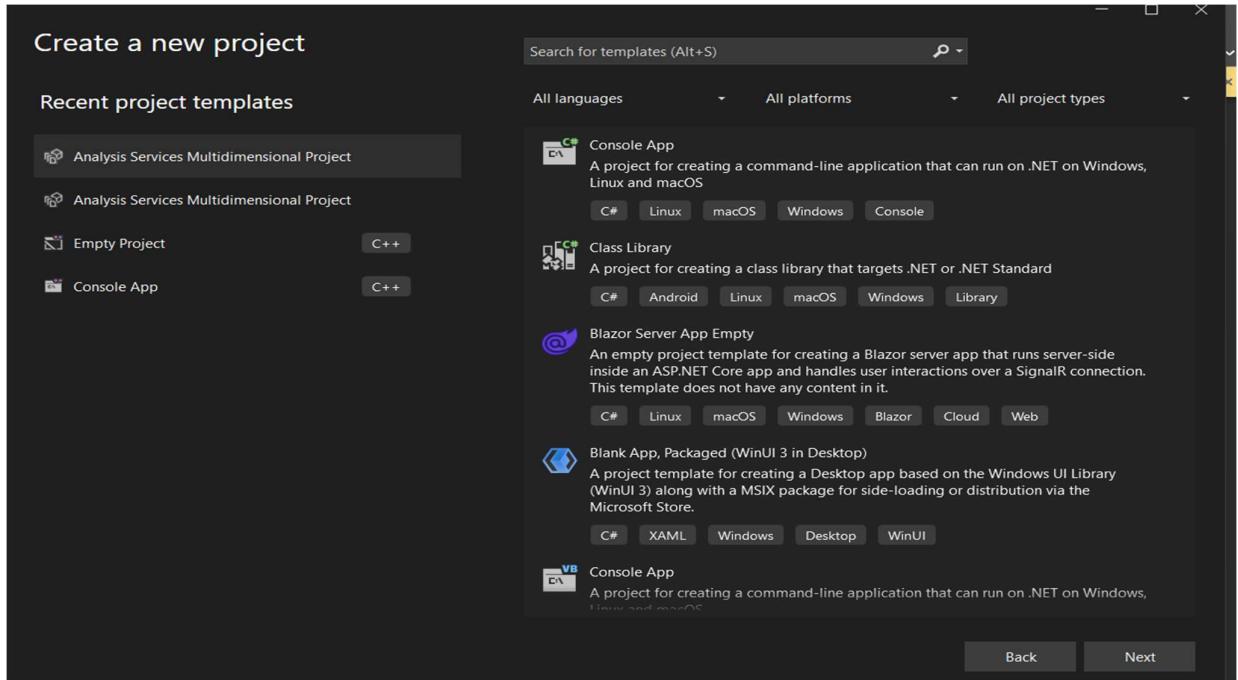
CHƯƠNG 3. THỰC HIỆN PHÂN TÍCH TRÊN KHO DỮ LIỆU (SSAS)

3.1. Tạo project SSAS mới

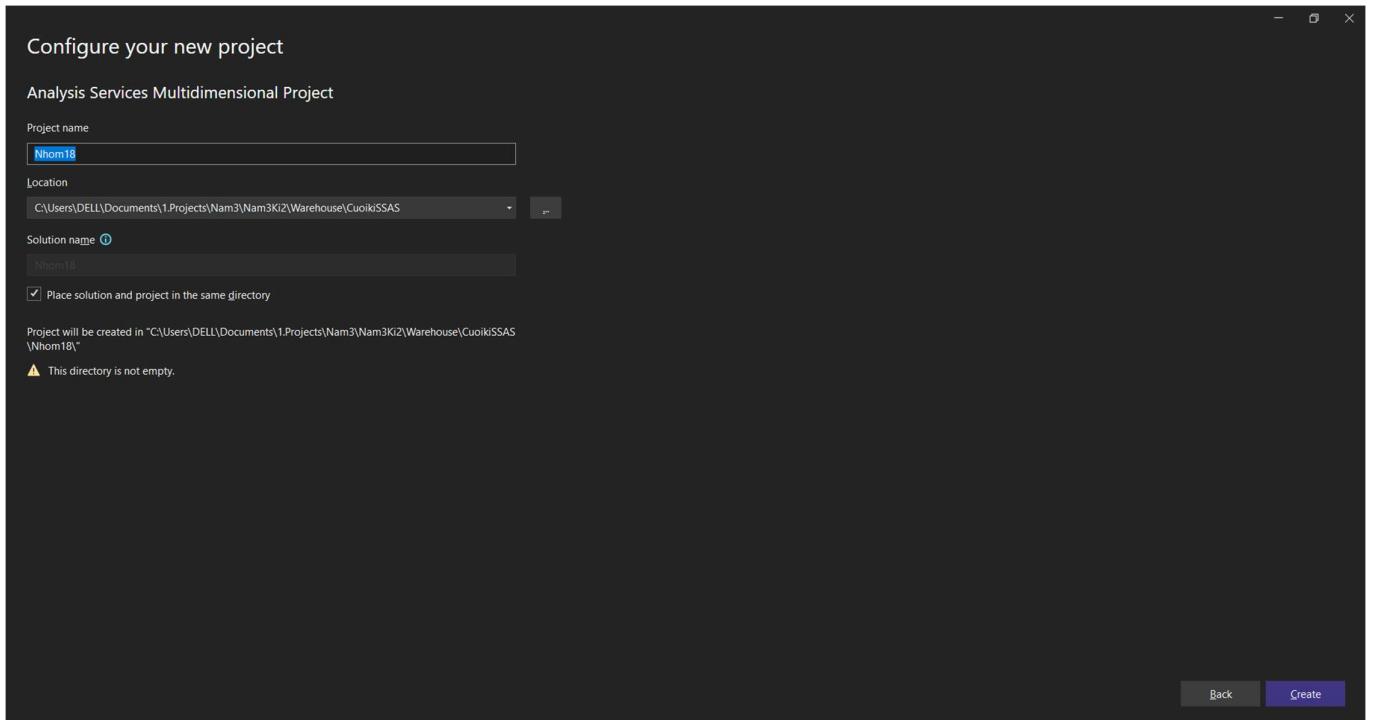
Bước 1: Mở Visual Studio và chọn “Create a new project”.



Bước 2: Chọn Analysis Services Multidimensional Project và chọn Next.

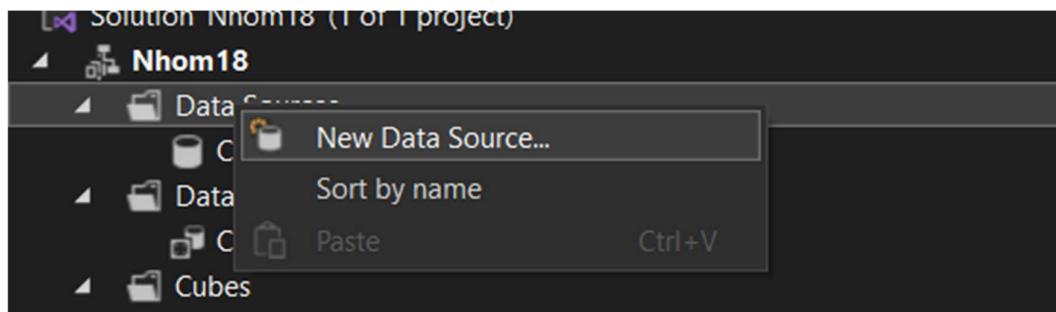


Bước 3: Đặt tên và thiết lập đường dẫn cho Project. Sau đó chọn Create.



3.2. Xác định dữ liệu nguồn (Data Source)

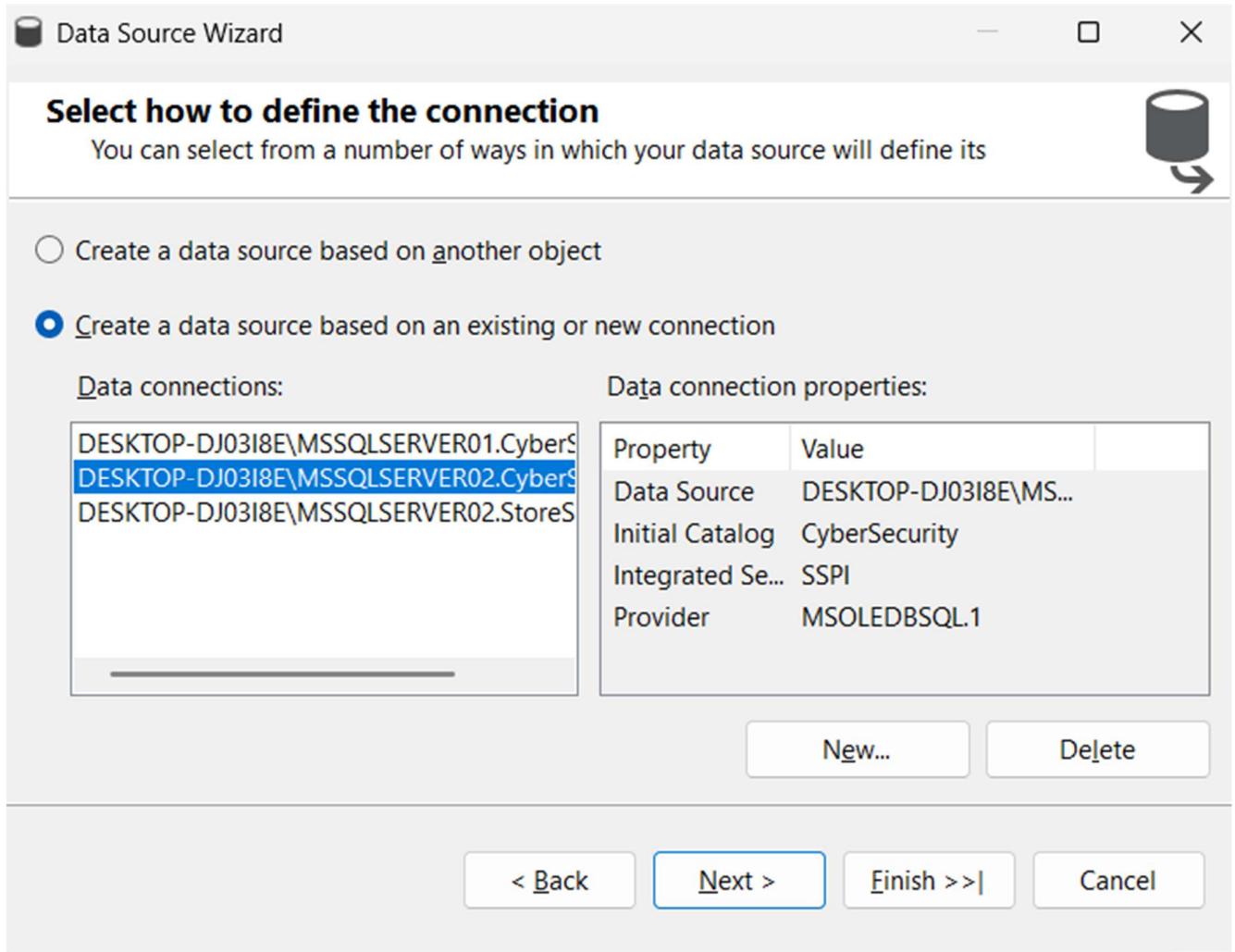
Bước 1: Tại Solution Explorer, right-click vào thư mục Data Sources, chọn New Data Source.



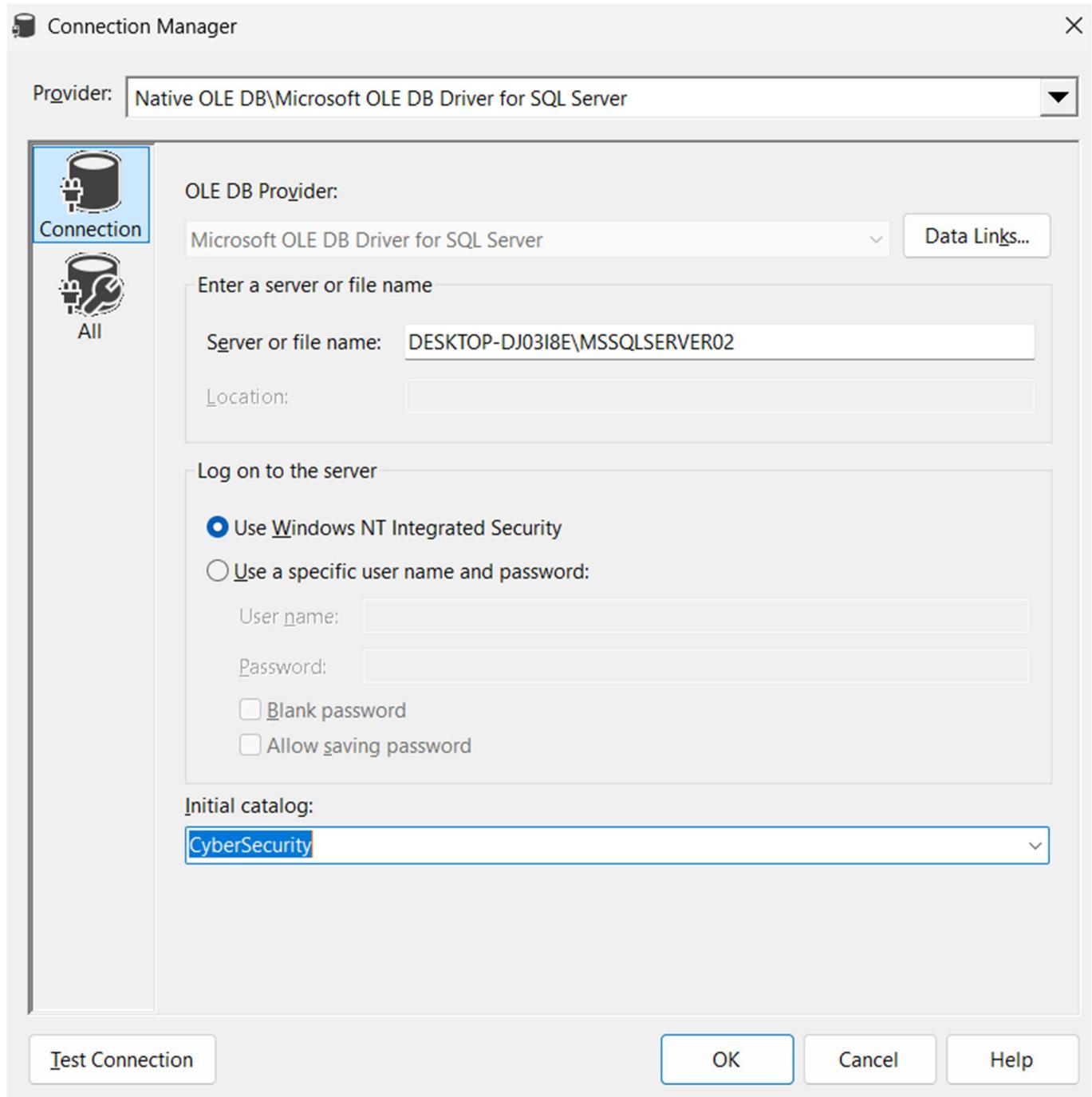
Bước 2: Hộp thoại Data Source Wizard xuất hiện, chọn Next để tiếp tục.



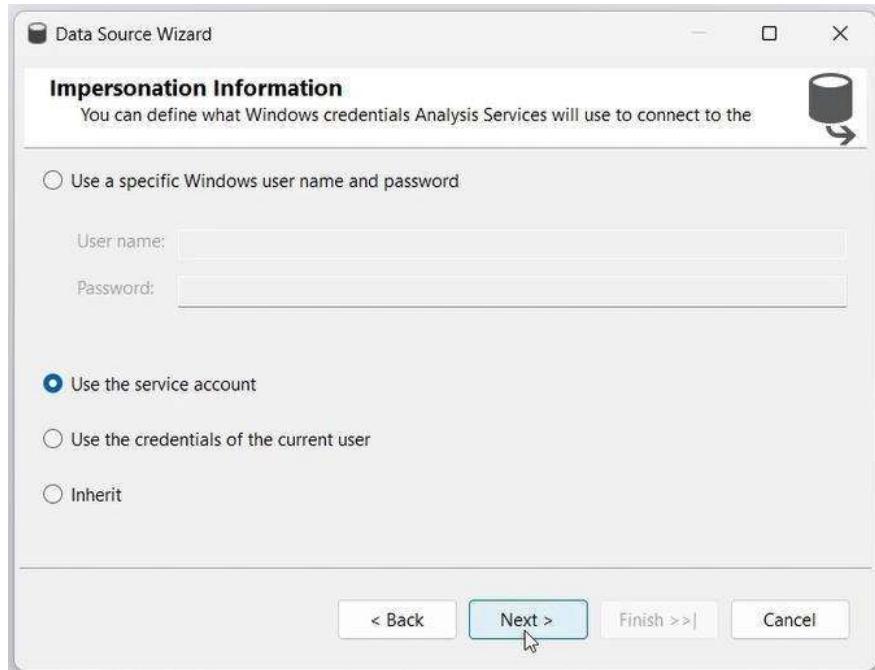
Bước 3: Ở trang Select how to define the connection, chọn “Create a data source based on an existing or new connection”, chọn New để tạo mới một Data connection.



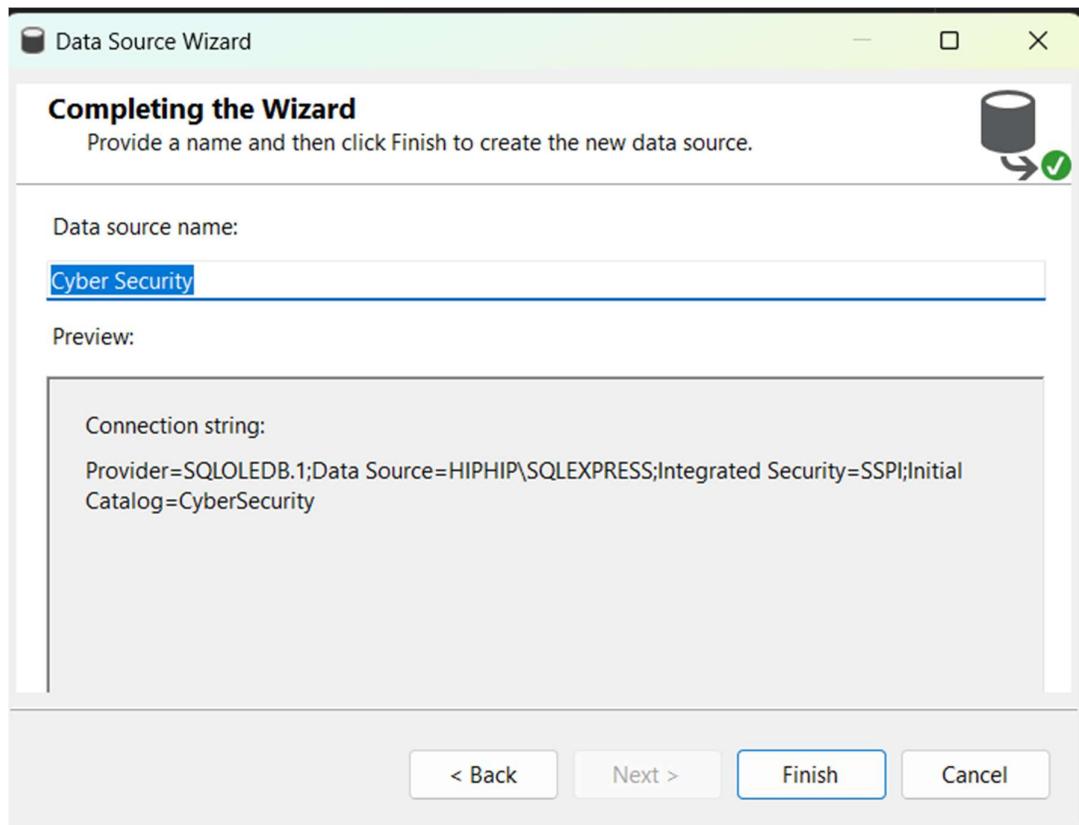
Bước 4: Chọn provider, server name và database để kết nối. Nhấn Next để tiếp tục.



Bước 5: Chọn “Use the service account”, sau đó chọn Next để tiếp tục.

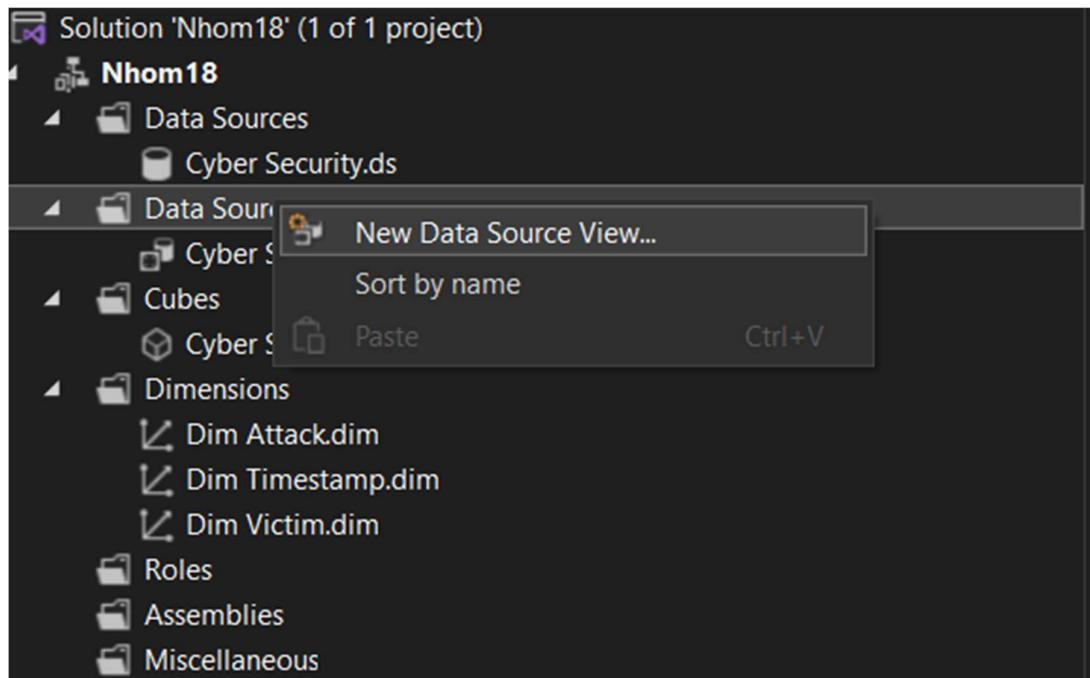


Bước 6: Cuối cùng ta chọn Finish để hoàn tất quy trình định nghĩa nguồn dữ liệu.

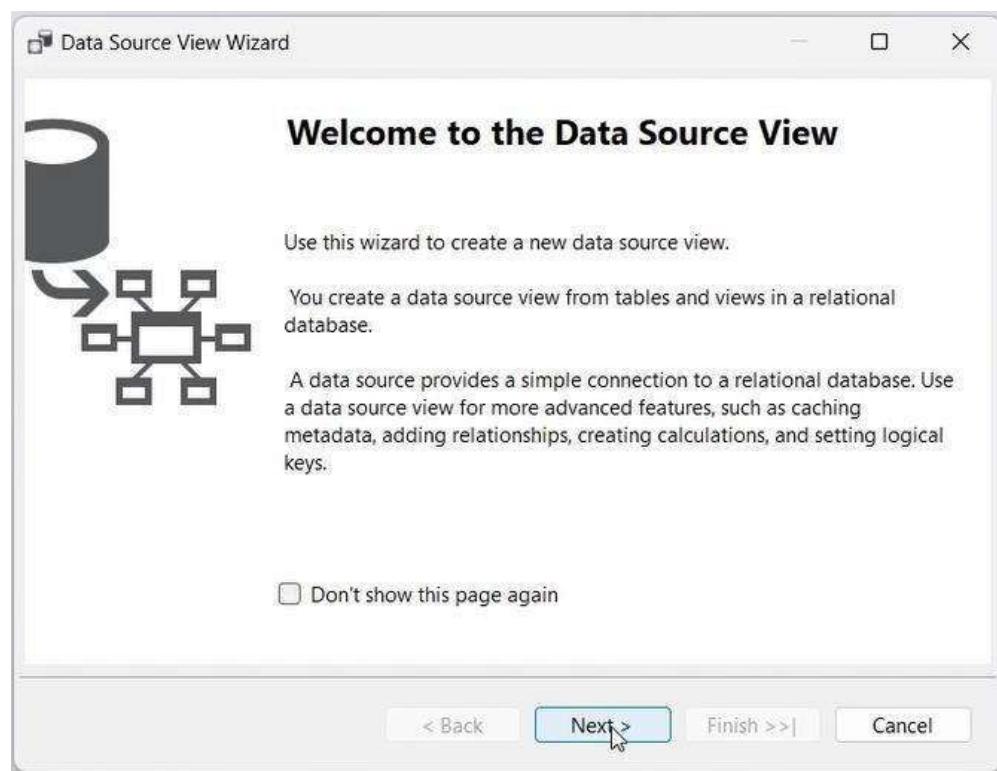


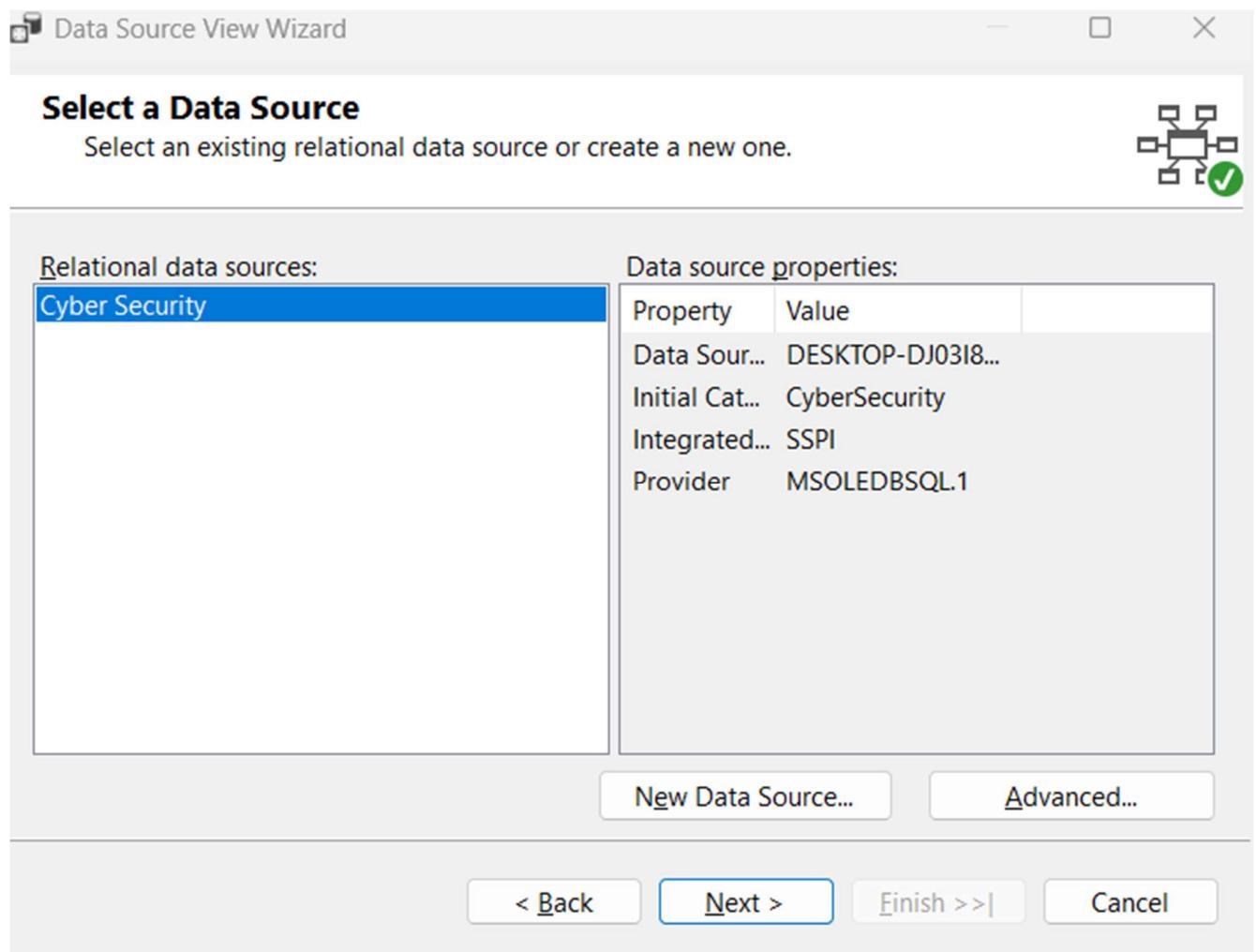
3.3. Xác định khung nhìn dữ liệu nguồn (Data Source View)

Bước 1: Tại Solution Explorer, right-click vào thư mục Data Source Views và chọn New Data Source View.

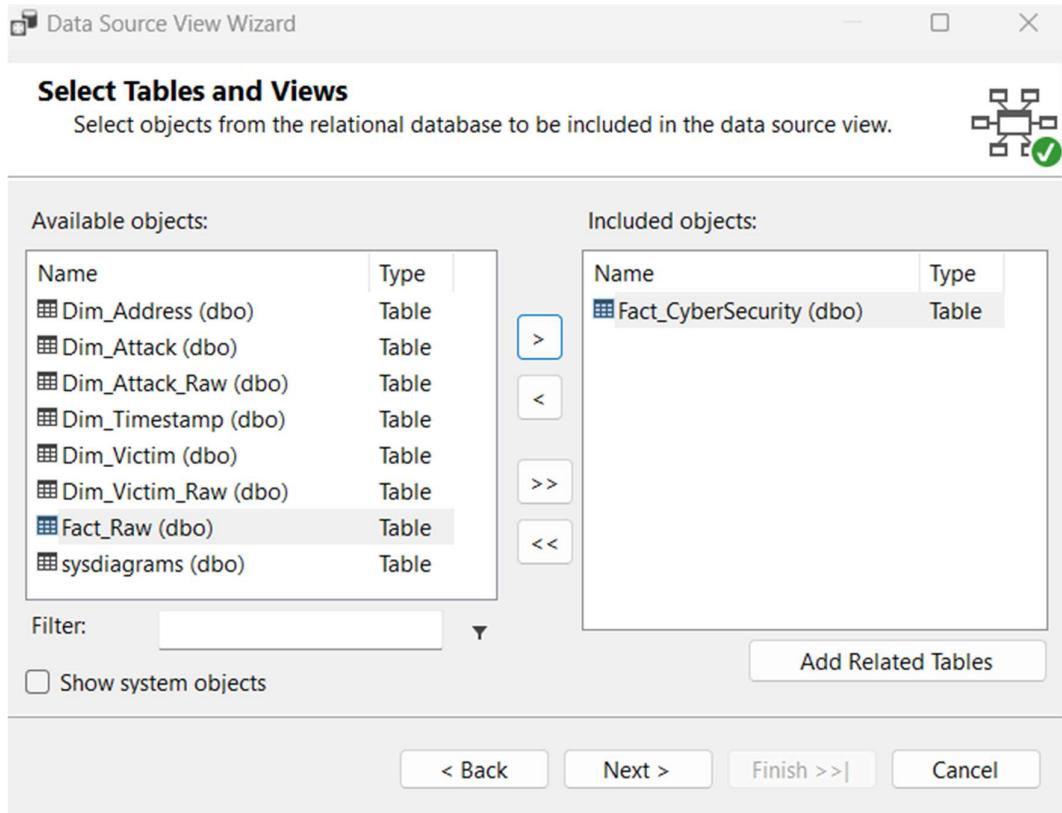


Bước 2: Hộp thoại Data Source View Wizard xuất hiện, chọn Next để tiếp tục. Chọn Data source vừa tạo và chọn Next để tiếp tục.

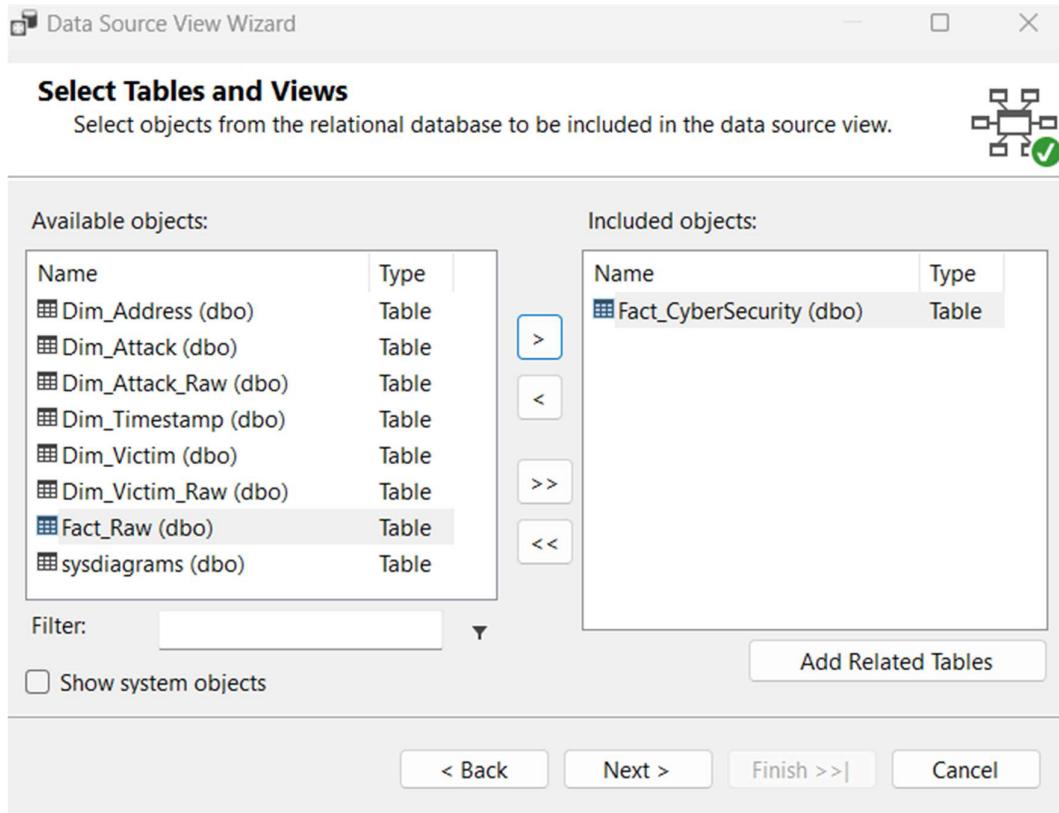




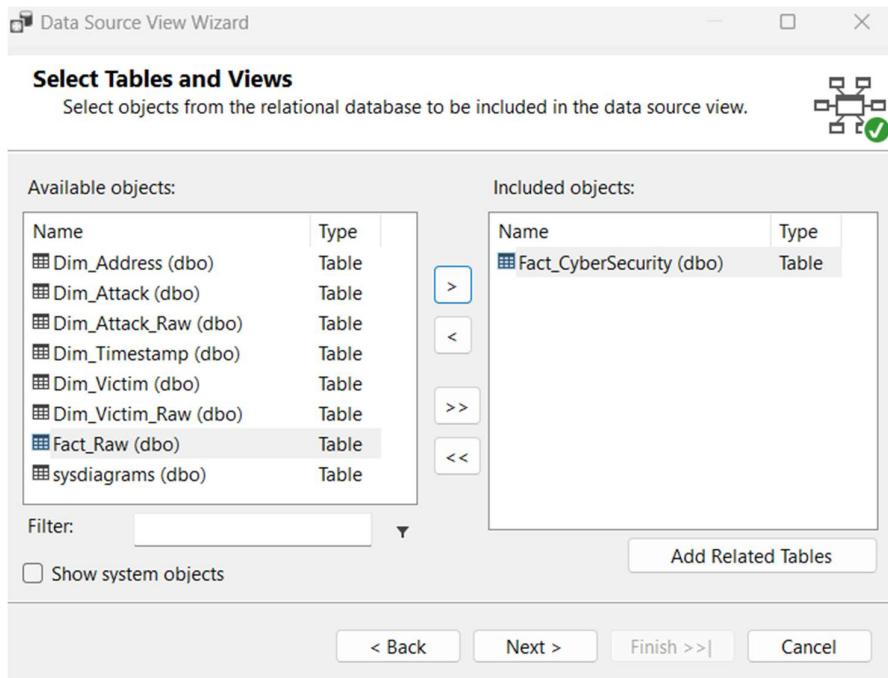
Bước 3: Chọn bảng Fact, sau đó chọn nút “>” để thêm bảng Fact vào Data source view.



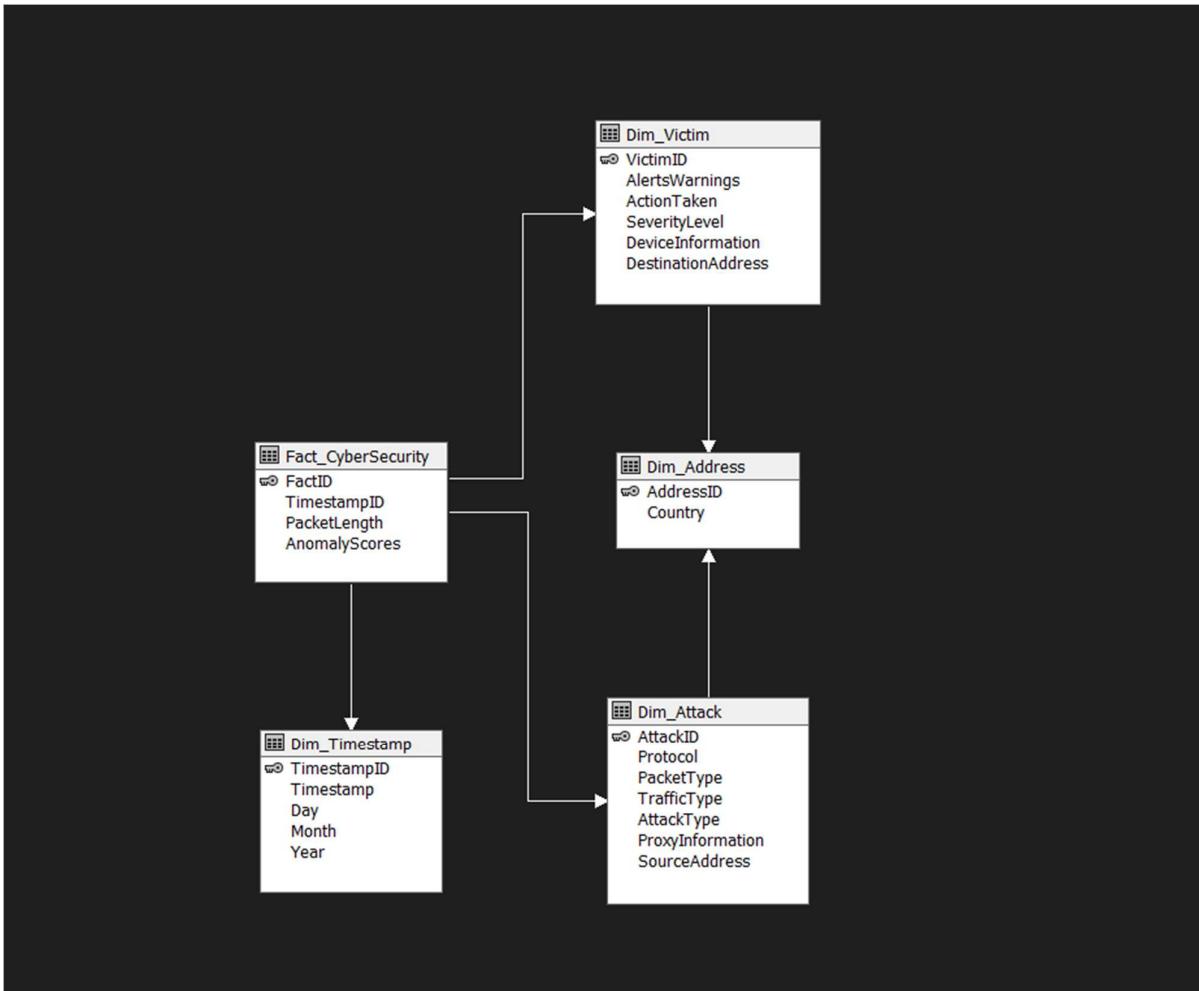
Bước 4: Chọn Add Related Tables để thêm tất cả bảng Dim vào Data source view và chọn Next để tiếp tục.



Bước 5: Chọn Finish để hoàn tất.



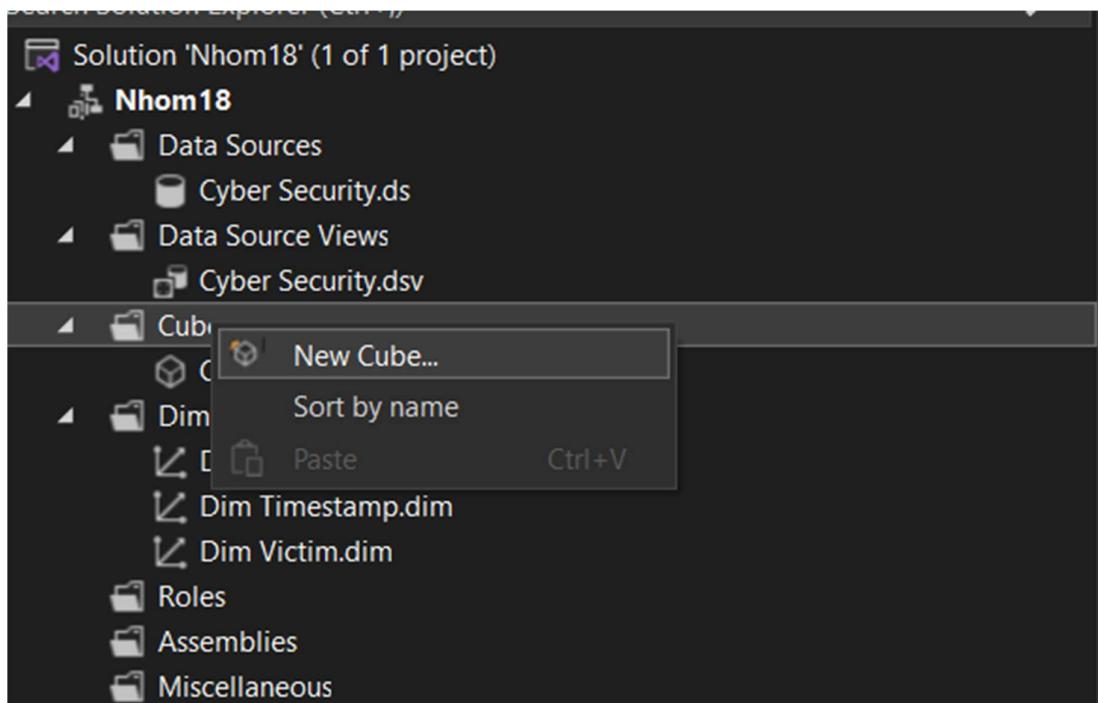
Sau khi kết thúc quá trình này, ta sẽ được data source view như hình sau:



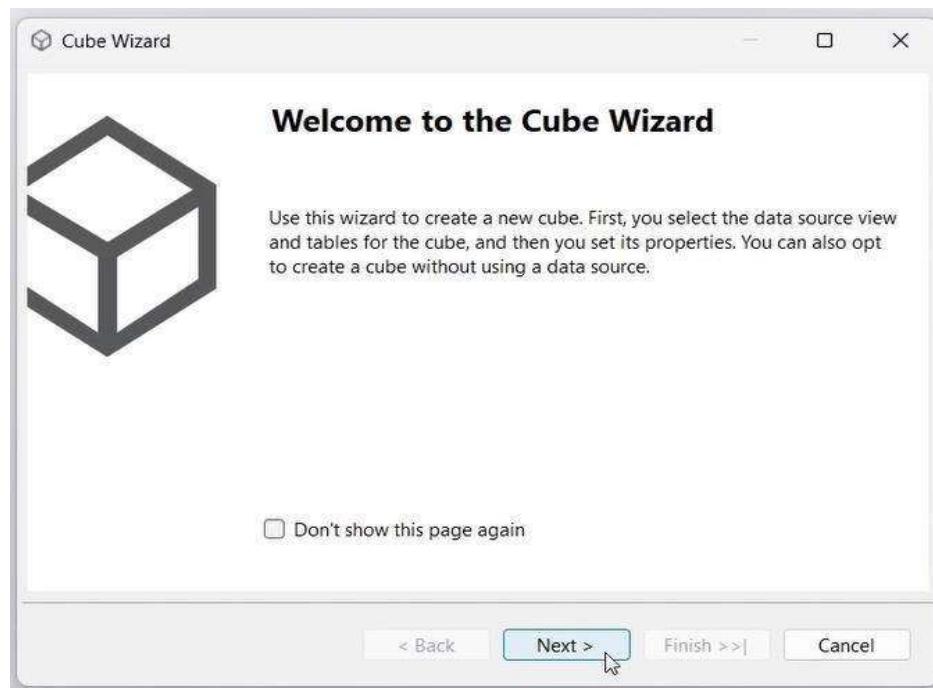
3.4. Xây dựng các khối (Cube) và deploy Cube

3.4.1.1. Tạo Cube và Dimension

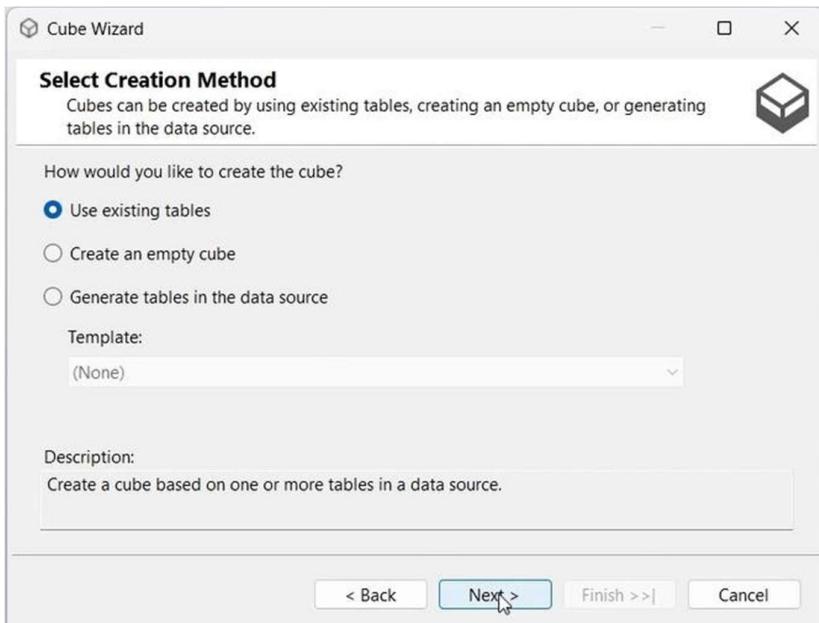
Bước 1: Tại Solution Explorer, right-click vào thư mục Cubes và chọn New Cube.



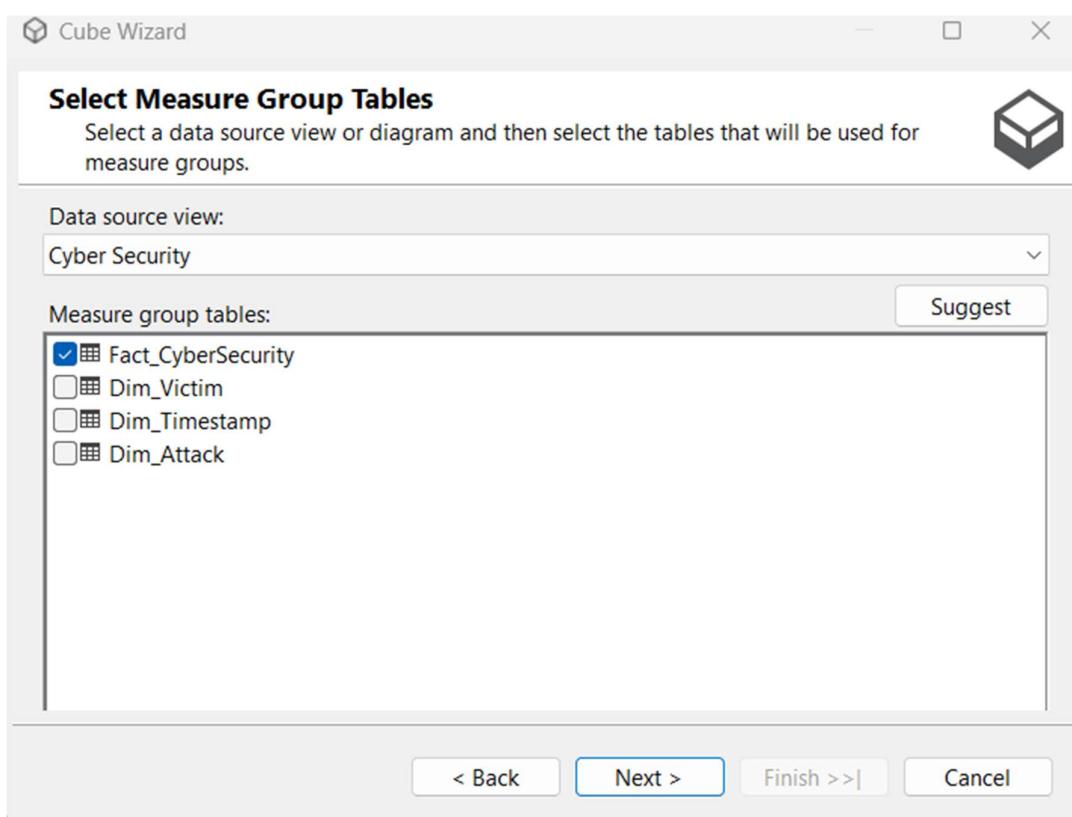
Bước 2: Hộp thoại Cube Wizard xuất hiện, chọn Next để tiếp tục



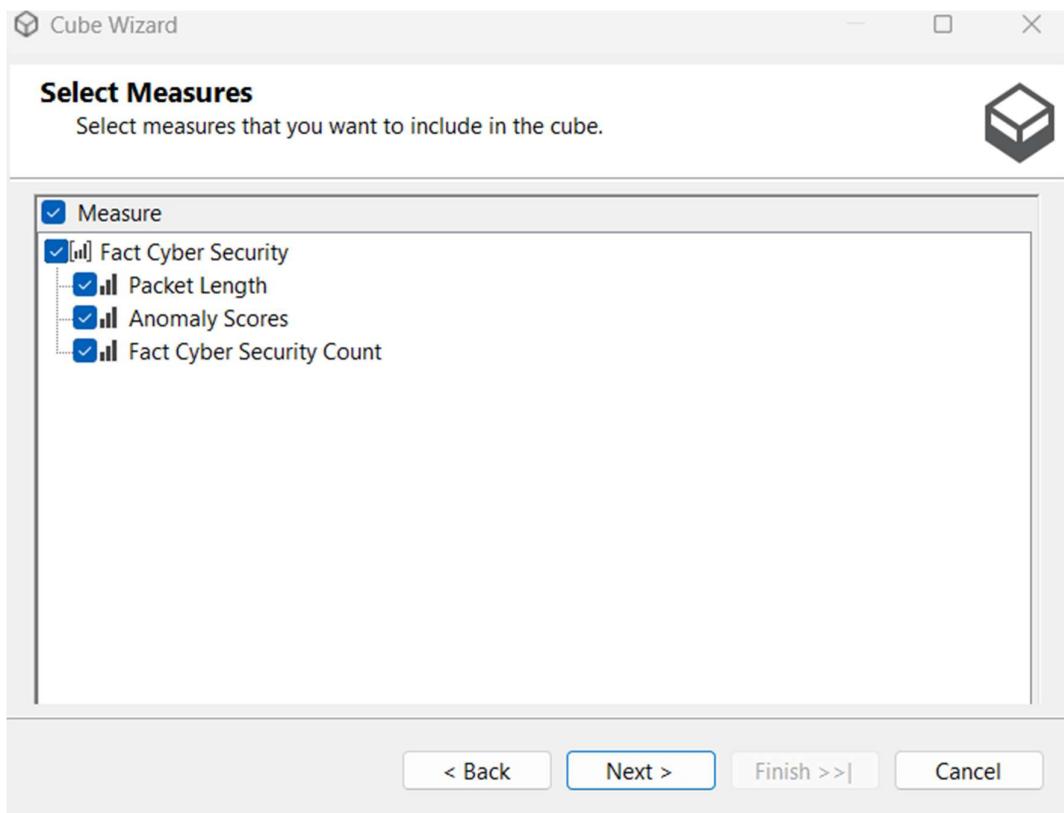
Bước 3: Chọn use existing tables, sau đó chọn Next để tiếp tục.



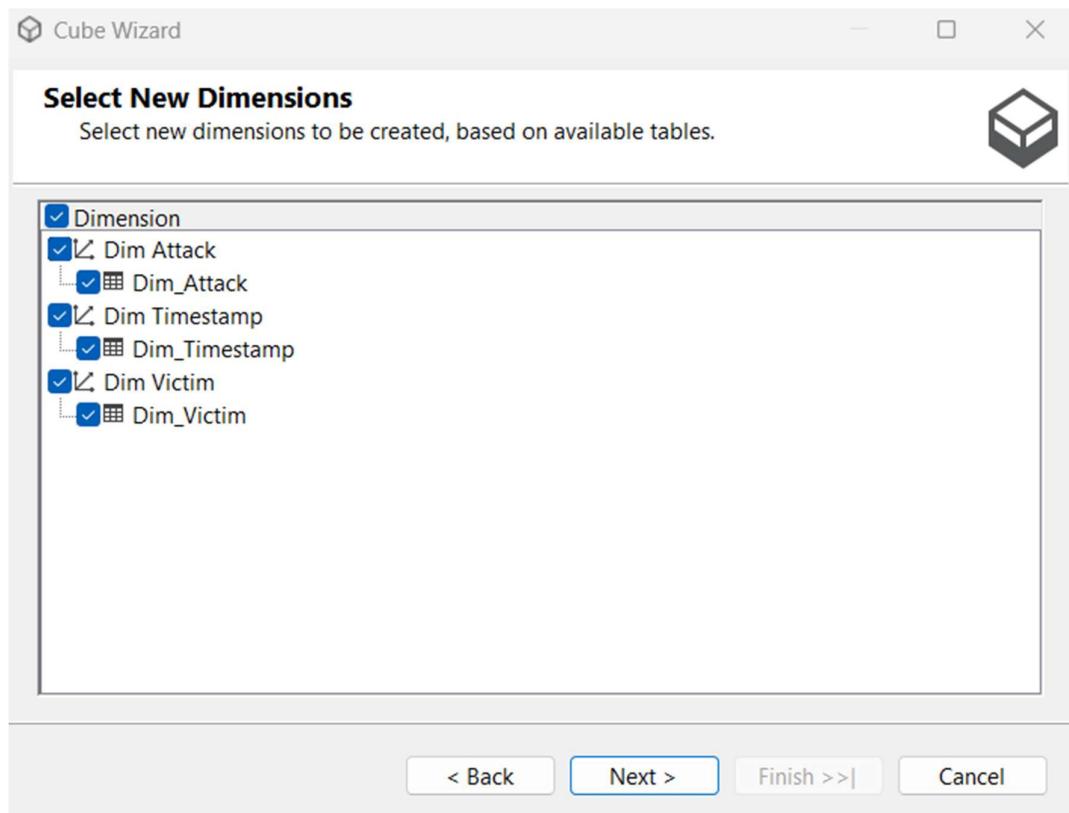
Bước 4: Chọn bảng Fact và các bảng Bridge để phân chia các measure group.



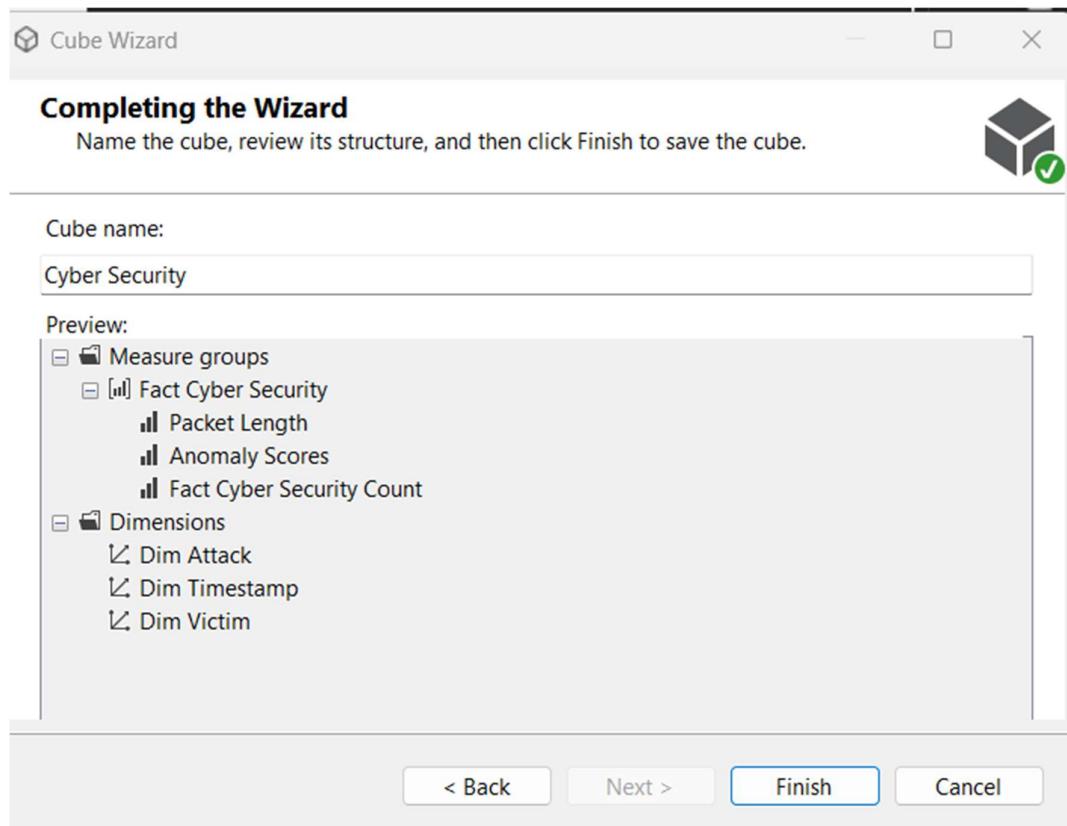
Bước 5: Chọn những độ đo đề xuất, sau đó chọn Next để tiếp tục.



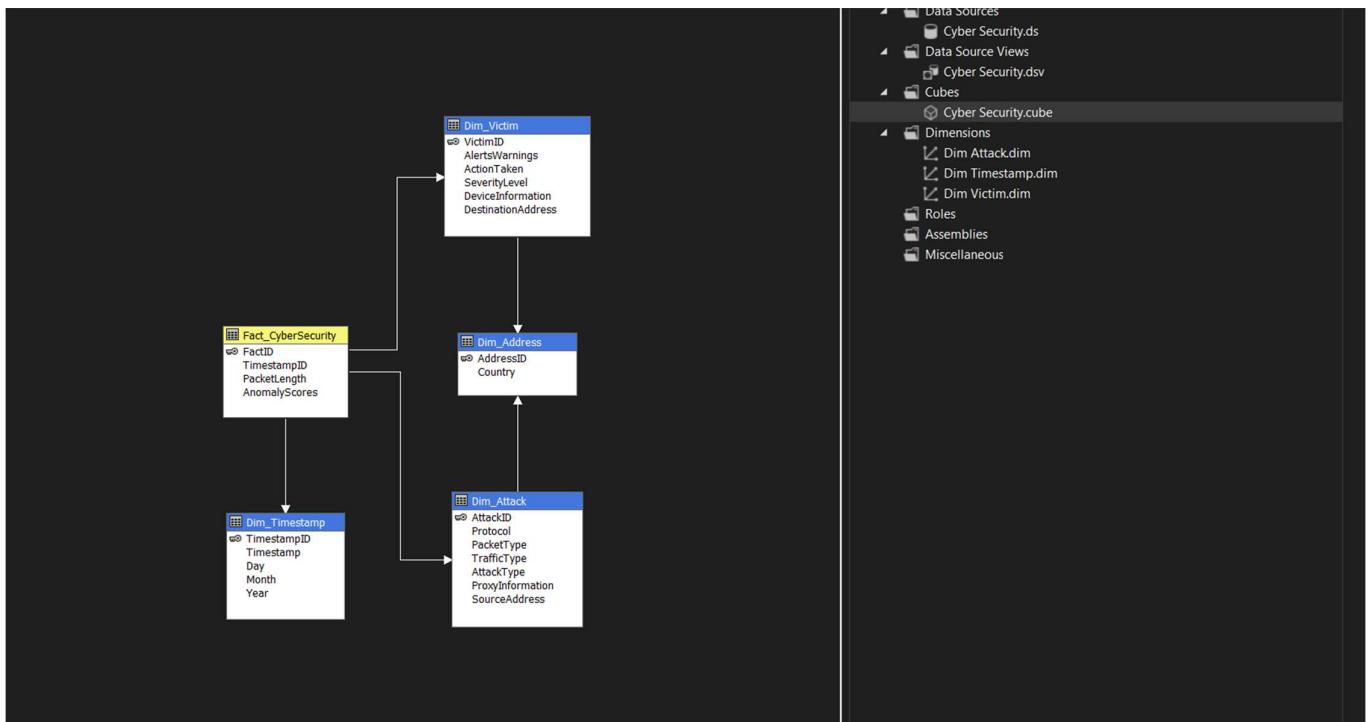
Bước 6: Chọn danh sách bảng Dimension, sau đó chọn Next để tiếp tục.



Bước 7: Xem lại các độ đo được tạo, các bảng Dimension và chọn Finish để hoàn tất.

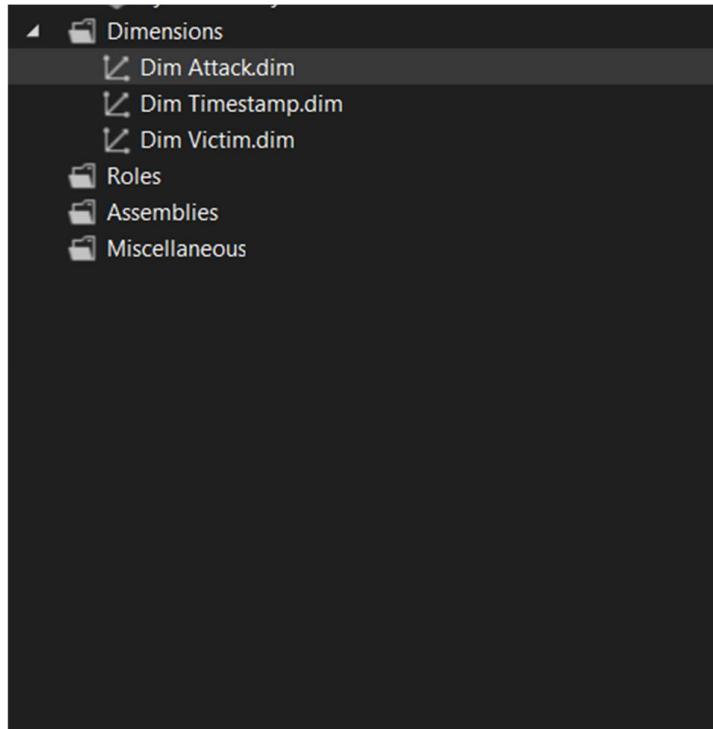


Sau khi kết thúc quá trình này, ta sẽ được kết quả như hình sau:



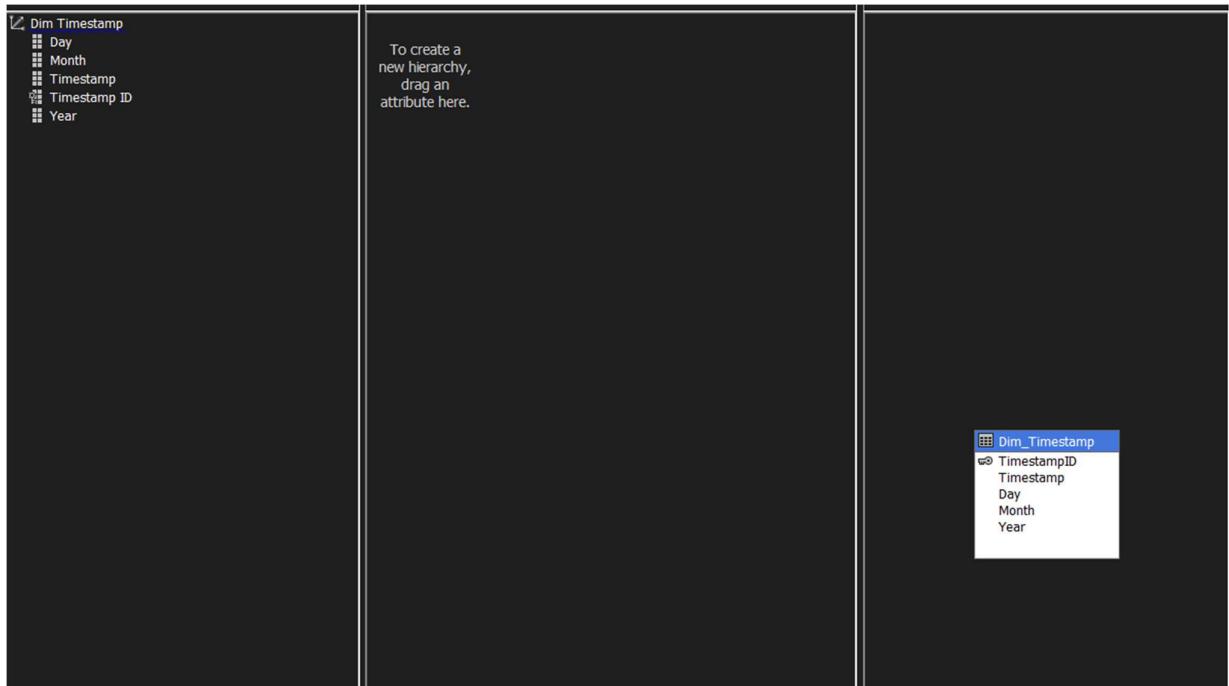
3.4.2. Thêm thuộc tính và chỉnh sửa property cho Dimension

Ở khung Dimensions, lần lượt chọn vào từng bảng Dimension Edit bảng Dimension để tiến hành chỉnh sửa bảng Dimension:

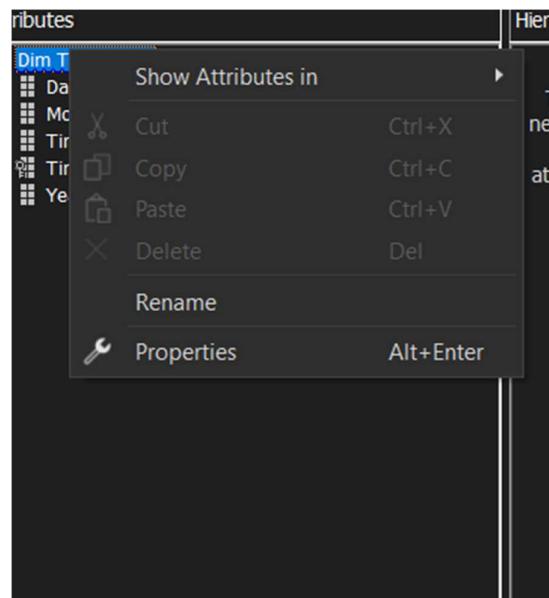


Bảng Dim TimeStamp:

Bước 1: Kéo thả thuộc tính TimeStamp, Day, Month, Year ở cột Data Source View sang cột Attributes.



Bước 2: Right-click tại bảng Dim TimeStamp và chọn Properties. Tại cửa sổ Properties, chỉnh sửa giá trị ErrorConfiguration KeyDuplicate là IgnoreError.

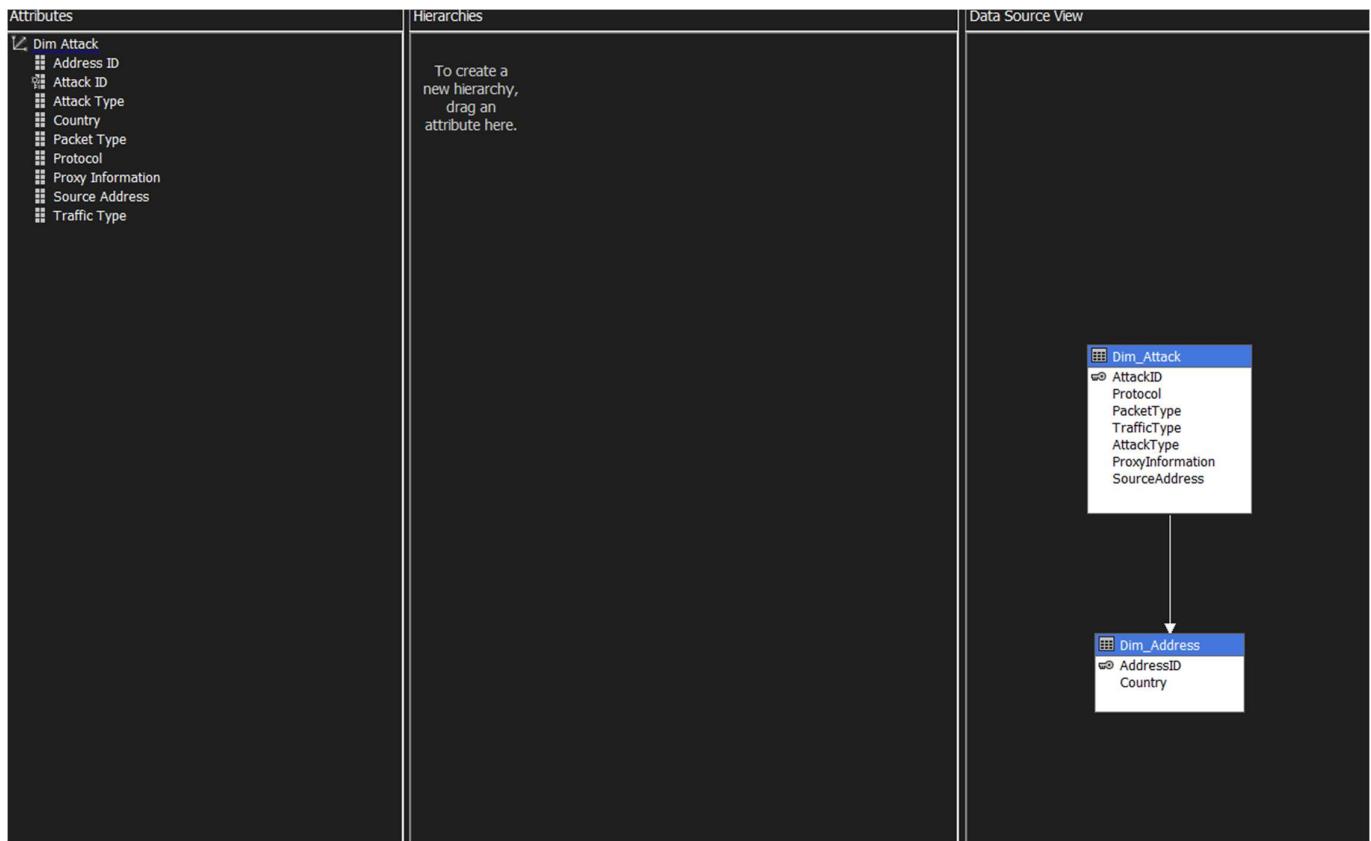


Bước 3: Tại cửa sổ Properties, chỉnh sửa giá trị UnknownMember là Hidden.

Error Configuration	(Custom)	ProcessingState	Unprocessed
CalculationError	IgnoreError	Source	Cyber Security (Data source \
KeyDuplicate	IgnoreError	StringStoresCompatibility	1050
KeyErrorAction	ConvertToUnknown	UnknownMember	<input checked="" type="checkbox"/> Hidden
KeyErrorLimit	0	UnknownMemberName	
KeyErrorLimitAction	StopProcessing	WriteEnabled	False

Bảng Dim Attack:

Bước 1: Kéo thả thuộc tính ở cột Data Source View sang cột Attributes.



Bước 2: Right-click tại bảng Dim attack và chọn Properties. Tại cửa sổ Properties, chỉnh sửa giá trị ErrorConfiguration KeyDuplicate là IgnoreError.

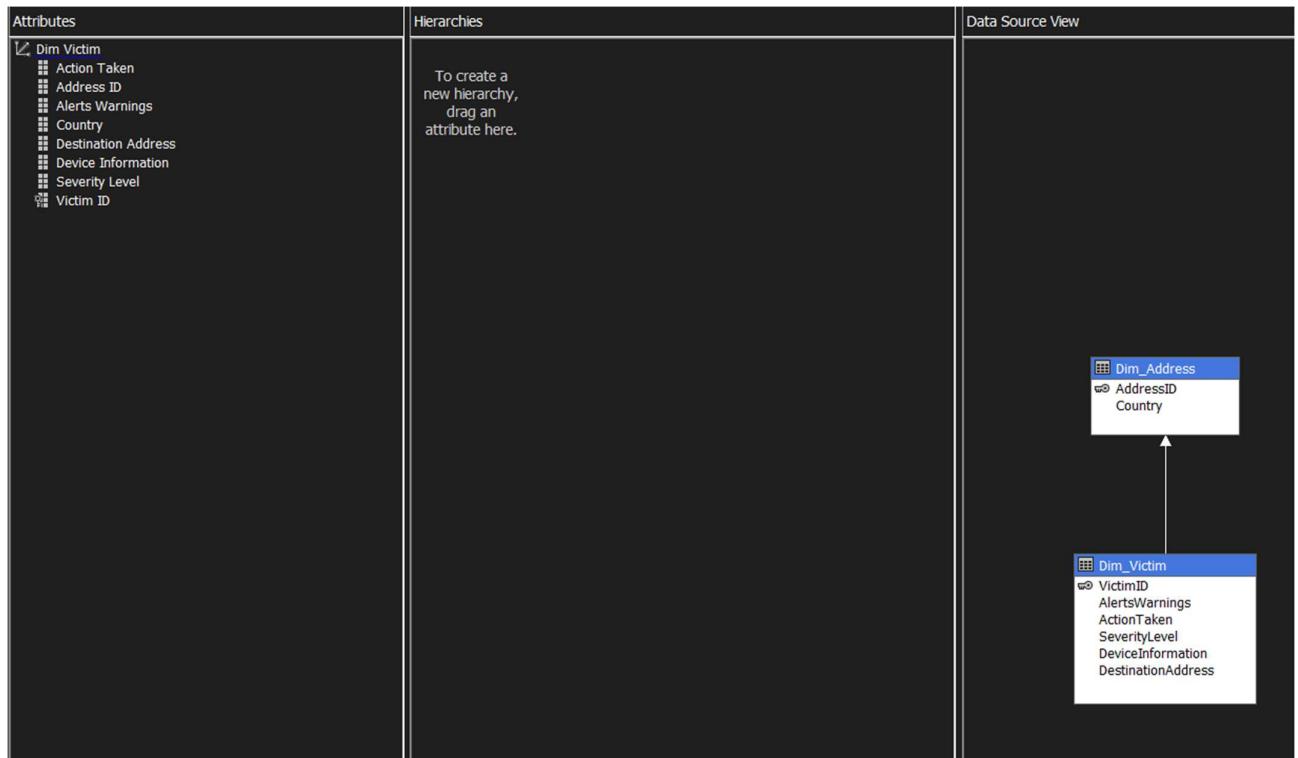
ErrorConfiguration	(custom)
CalculationError	IgnoreError
KeyDuplicate	IgnoreError
KeyErrorAction	ConvertToUnknown
KeyErrorLimit	0
KeyErrorLimitAction	StopProcessing

Bước 3: Tại cửa sổ Propterties, chỉnh sửa giá trị UnknownMember là Hidden.

ProcessingState	Unprocessed
Source	Cyber Security (Data source view)
StringStoresCompatibility	1050
UnknownMember	Hidden
UnknownMemberName	
WriteEnabled	False
Basic	

Bảng Dim Victim:

Bước 1: Kéo thả thuộc ở cột Data Source View sang cột Attributes.



Bước 2: Right-click tại bảng Dim_Movie và chọn Properties. Tại cửa sổ Properties, chỉnh sửa giá trị ErrorConfiguration [KeyDuplicate là IgnoreError.

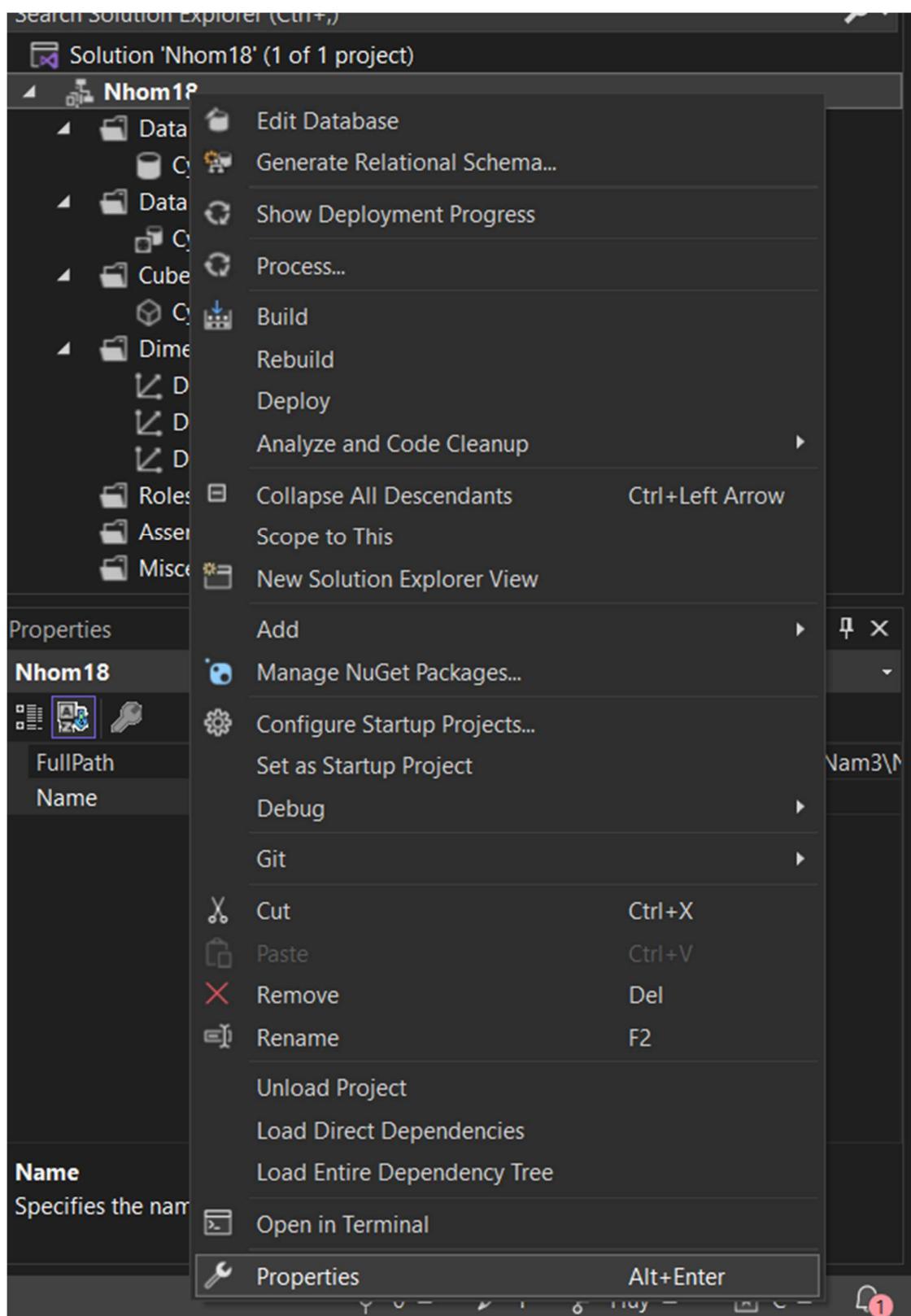
ErrorConfiguration (custom)	
CalculationError	IgnoreError
KeyDuplicate	IgnoreError
KeyErrorAction	ConvertToUnknown
KeyErrorLimit	0
KeyErrorLimitAction	StopProcessing

Bước 3: Tại cửa sổ Propterties, chỉnh sửa giá trị UnknownMember là Hidden.

ProcessingState	Unprocessed
Source	Cyber Security (Data source)
StringStoresCompatibility	1050
UnknownMember	Hidden
UnknownMemberName	
WriteEnabled	False

3.4.3. Deploy project SSAS

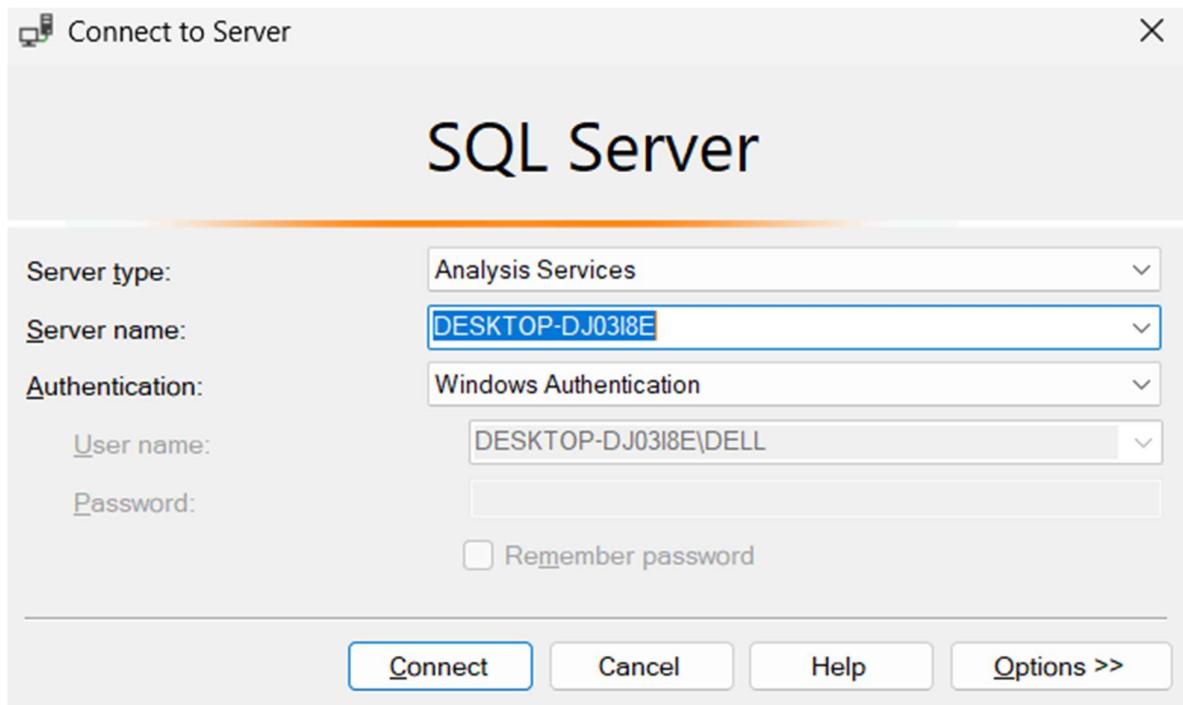
Bước 1: Tại Solution Explorer, right-click ở tên project và chọn Properties để chỉnh sửa kết nối đến Analysis Service của SQL Server.

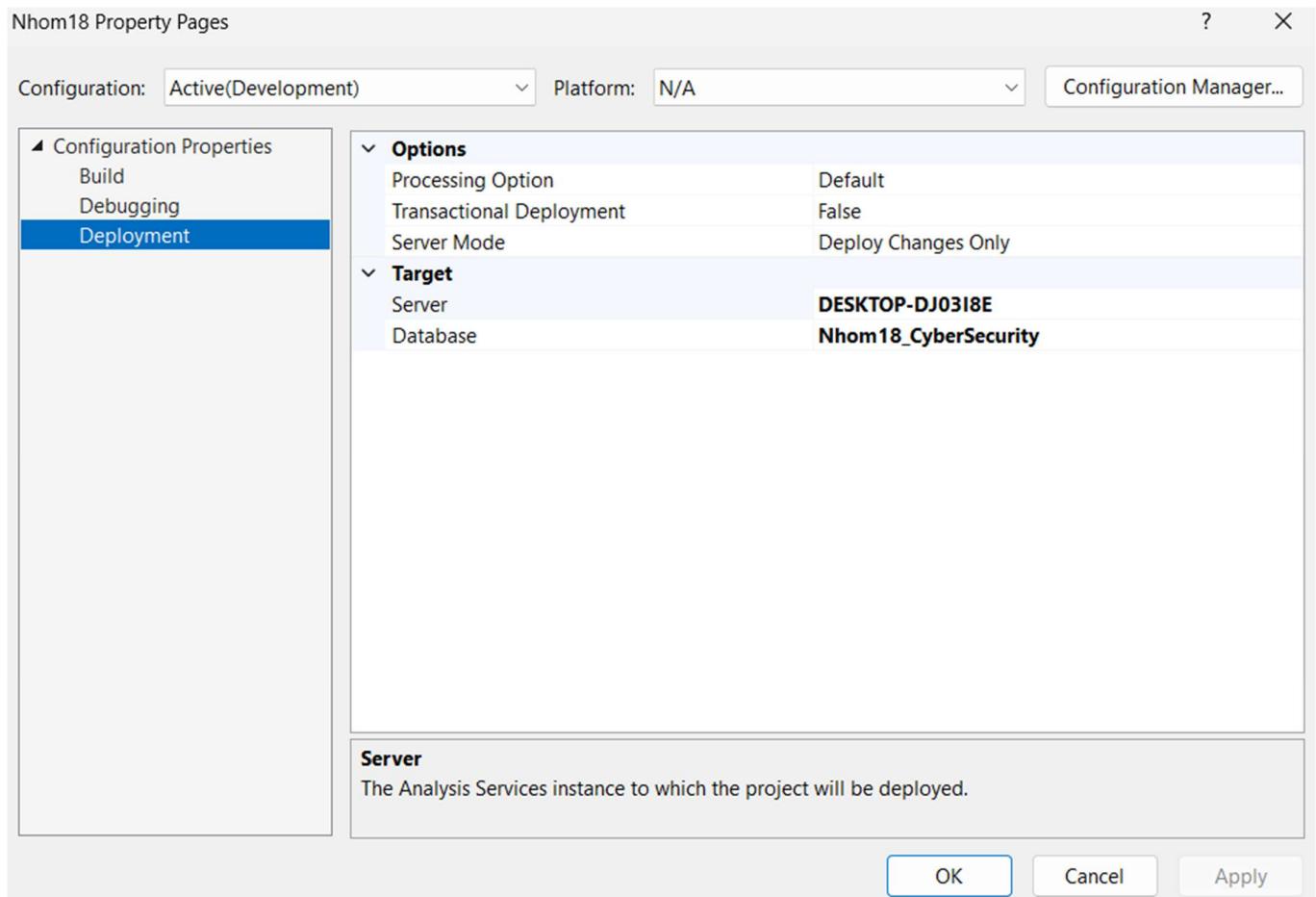


Bước 2: Ở cửa sổ vừa mở, đi đến Deployment và thực hiện đổi tên Server theo tên trong SQL Server.

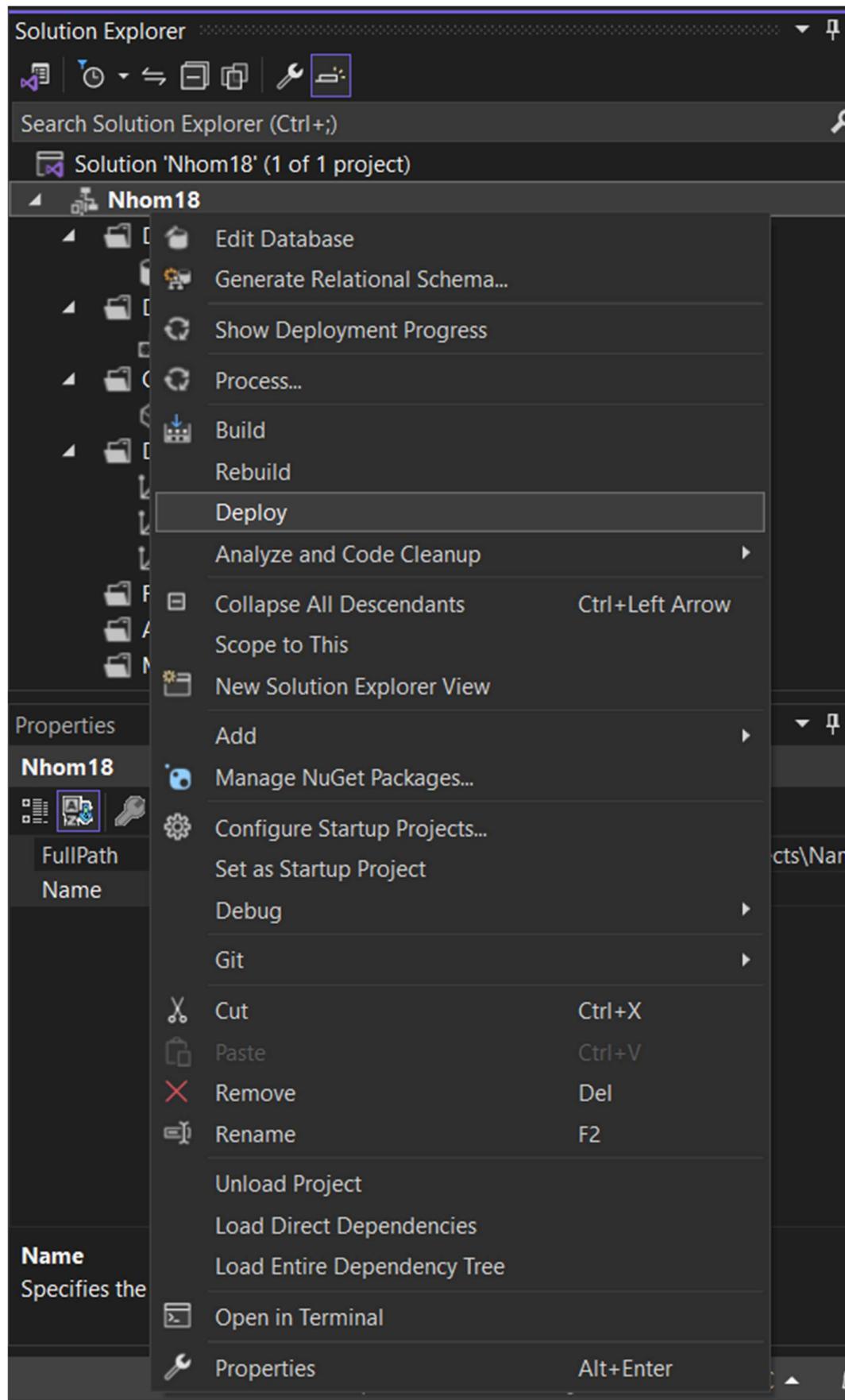
Server ban đầu là localhost:

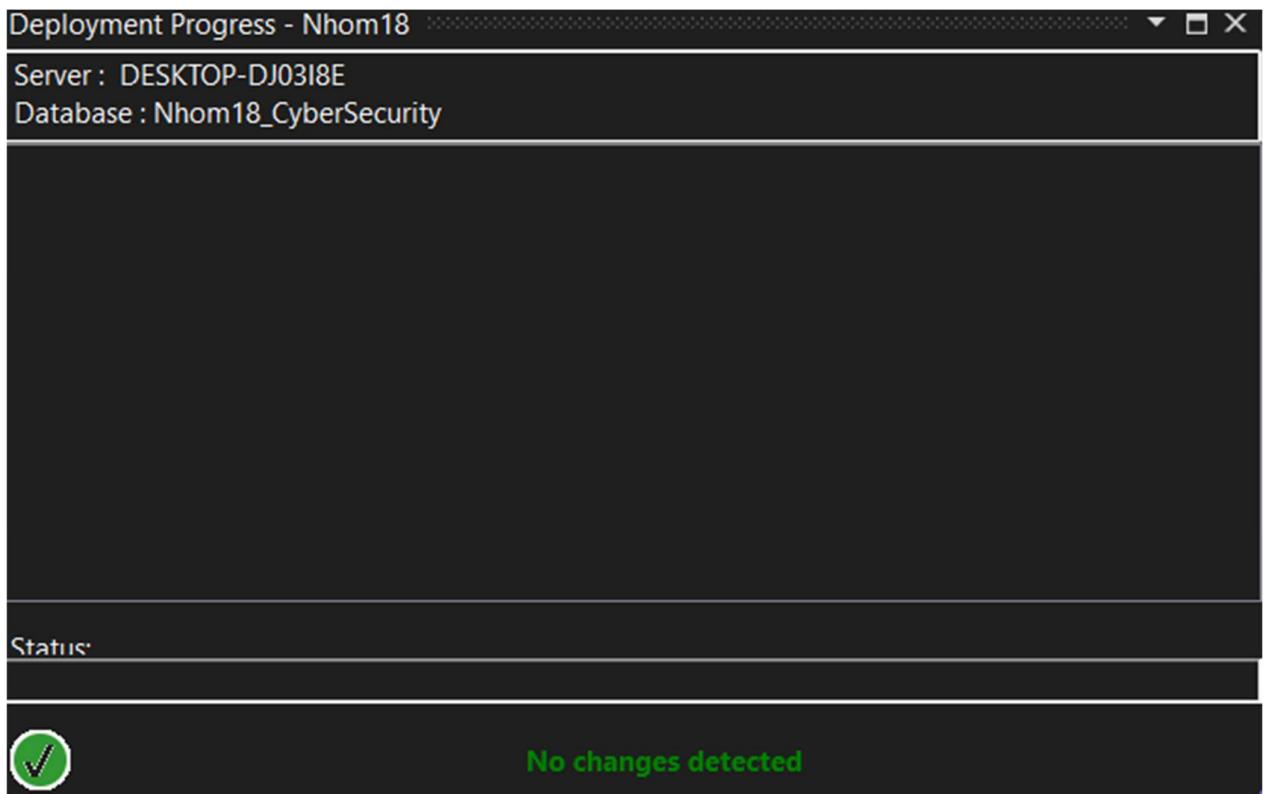
Mở SQL Server và lấy tên Server của Analysis Services như sau:





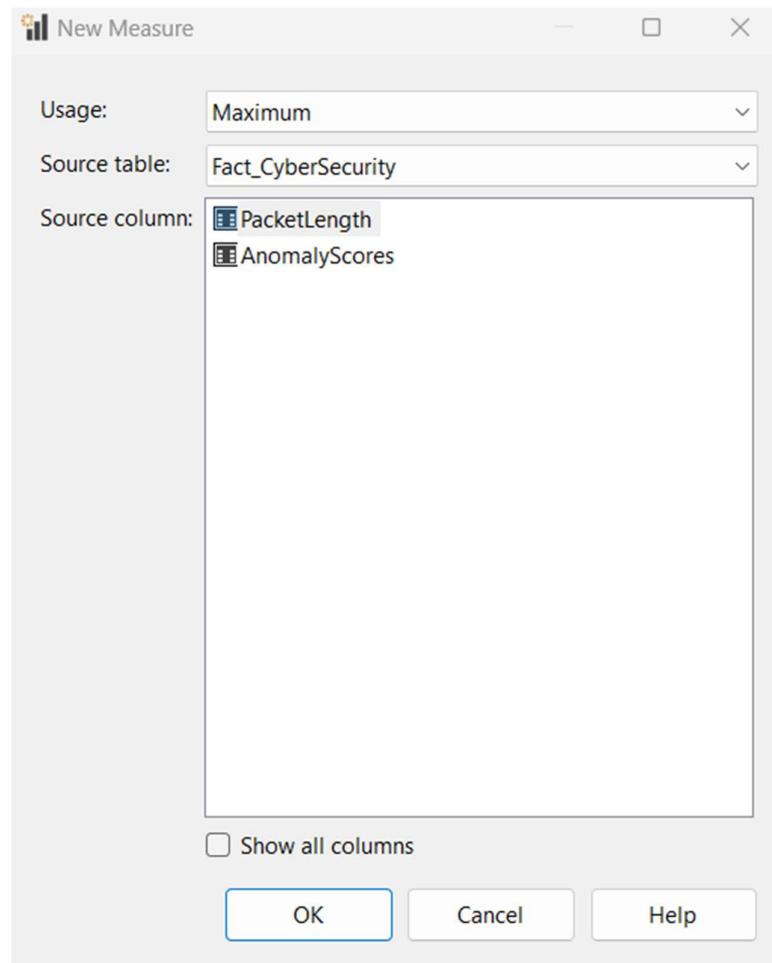
Bước 3: Right-click vào project đang hiện hành và chọn Deploy project.



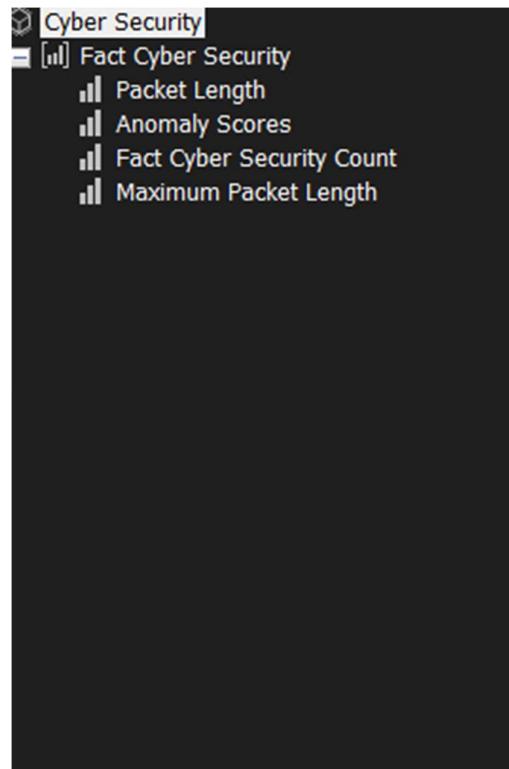


3.5. Xác định các độ đo (Measures)

Tại khung vừa tạo, chọn Show Measures Grid để hiển thị chi tiết các độ đo. Để tạo ra các độ đo mới, ta chọn Add new measure, sau đó tạo ra thêm một độ đo như sau:



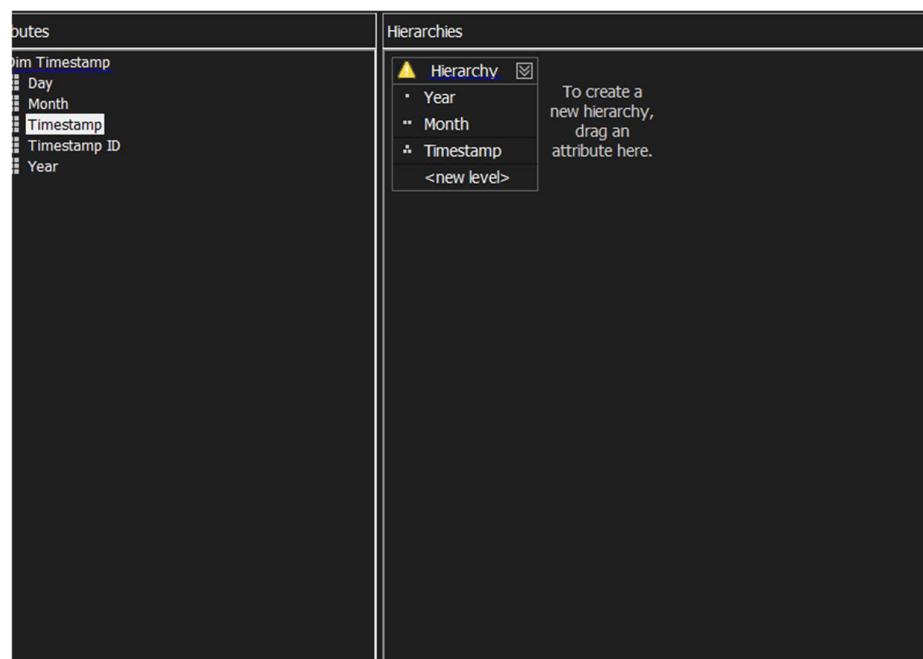
Kết quả sau khi thực hiện thêm độ đo mới:



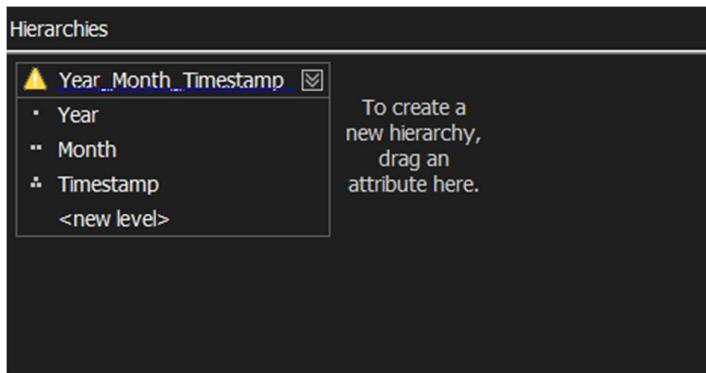
3.6. Phân cấp bảng chiều

Thực hiện tạo các thuộc tính phân cấp (Hierarchies) và định nghĩa Attribute Relationships cho bảng Dim Date. Tạo Hierarchy phân cấp theo Year ⇨ Month ⇨ TimeStamp:

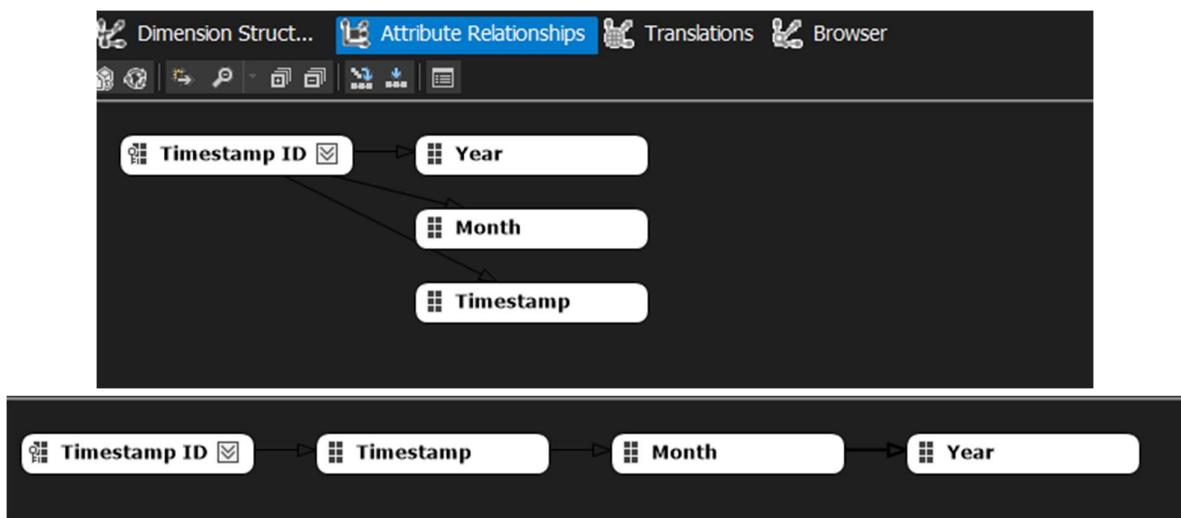
Bước 1: Kéo những thuộc tính Year, Month, TimeStamp qua cửa sổ Hierarchies với thứ tự từ trên xuống là phân cấp từ cao tới thấp.



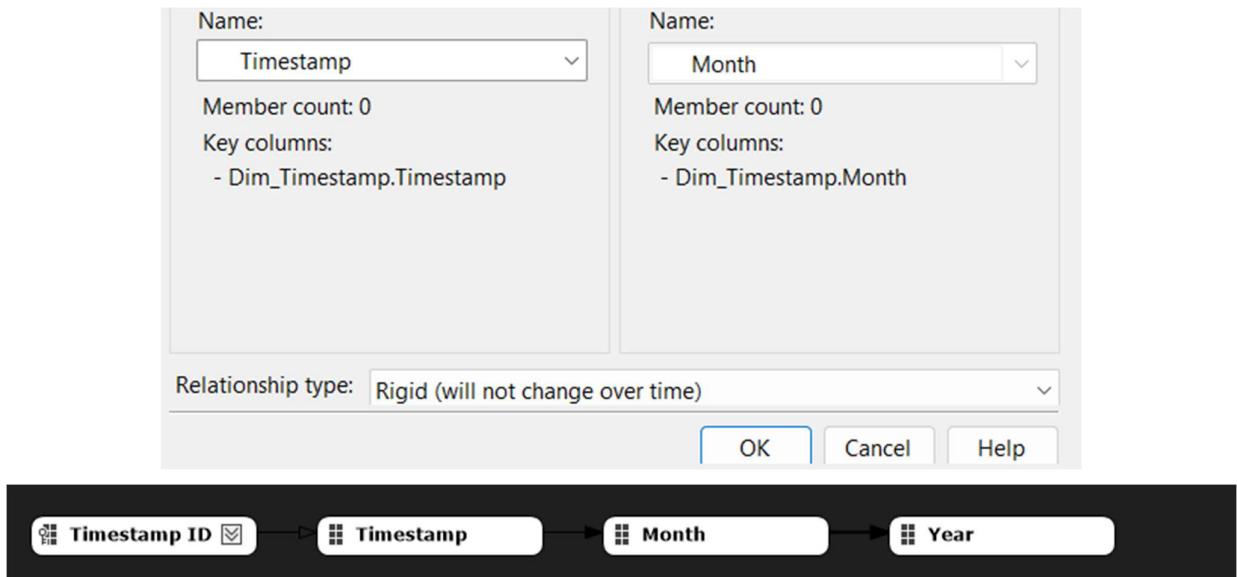
Bước 2: Đổi tên Hierarchy.



Bước 3: Tại tab Attribute Relationships, định nghĩa các mối quan hệ. Thực hiện kéo thả phân cấp từ nhỏ đến lớn theo thứ tự từ trái sang phải.

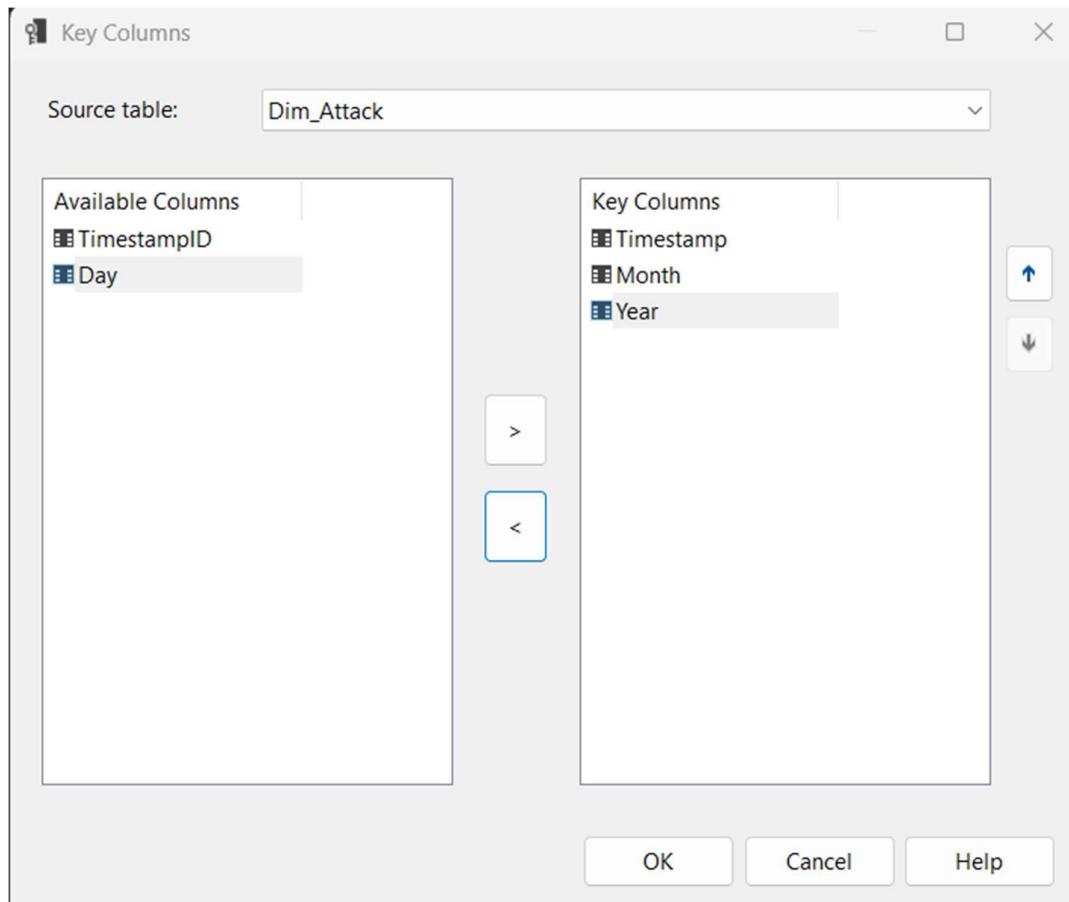


Chỉnh sửa Relationship Type thành Rigid:



Bước 4: Chỉnh khóa cột (KeyColumns) và tên cột (Name Column) của thuộc tính TimeStamp. Vì thuộc tính này sẽ lấy khóa cột gồm chính nó và những thuộc tính cấp cao hơn.

Chuyển sang tab Dimension Structure, cột Attributes, right-click vào thuộc tính Day và chọn Properties. Tại cửa sổ Properties, chọn KeyColumns. Thêm các thuộc tính cấp cao hơn vào KeyColumns, sau đó chọn OK để hoàn tất:



MembersWithDataCap	
NamingTemplate	
RootMemberIf	ParentIsBlankSelfOrMissing
UnaryOperatorColumn	(none)
Source	
CustomRollupColumn	(none)
CustomRollupProperties	(none)
KeyColumns	Dim_Timestamp.Timestamp
NameColumn	(none)
ValueColumn	(none)

Tại cửa sổ Properties, chọn NameColumn và chọn tên thuộc tính sẽ hiển thị trên Hierachy là TimeStamp:

Chỉnh thuộc tính OrderBy thành Key để Timestamp được sắp xếp theo thứ tự tăng dần:

The screenshot shows two windows from a data modeling tool.

Name Column Dialog:

- Binding type: Column binding
- Source table: Dim_Timestamp
- Source column: A list box containing:
 - TimestampID
 - Timestamp
 - Day
 - Month
 - Year
- Buttons: OK, Cancel, Help

KeyColumns Properties Grid:

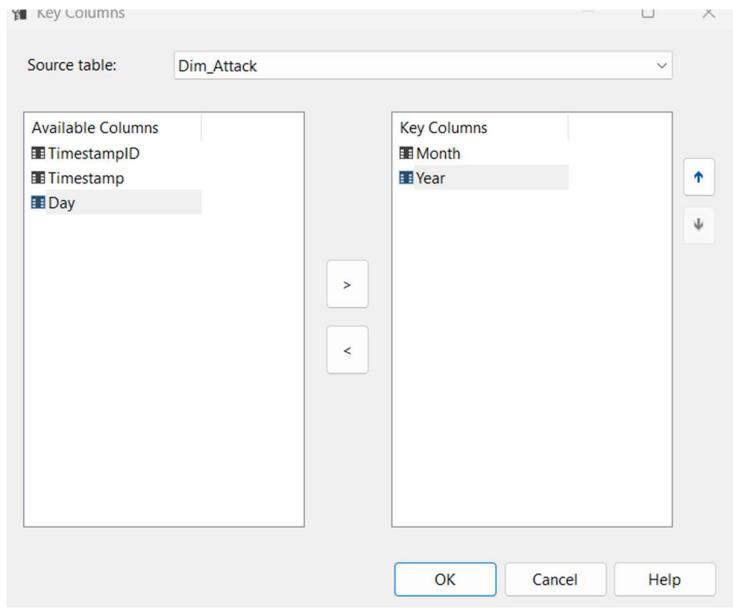
KeyColumns (none)	
Source	
CustomRollupColumn	(none)
CustomRollupProperties	(none)
KeyColumns	(Collection)
NameColumn	Dim_Timestamp.Timestamp
ValueColumn	(none)

Below the grid is a table of properties:

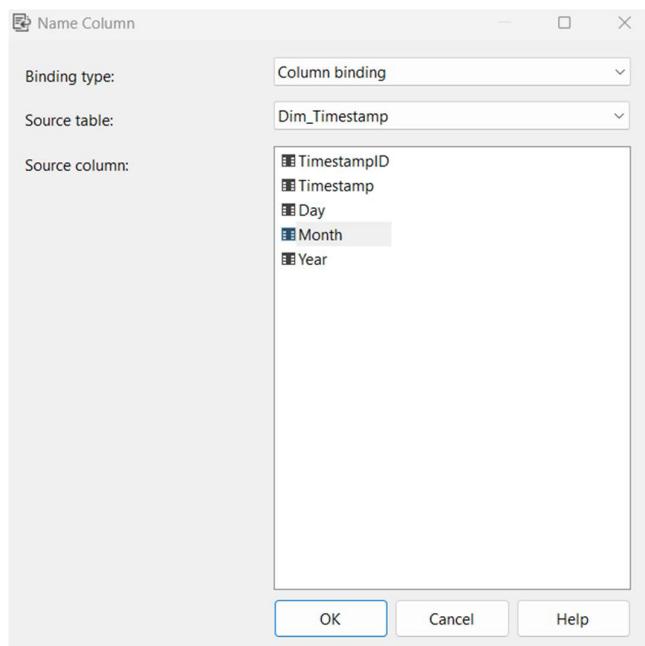
EstimatedCount	0
HasLineage	False
IsAggregatable	True
OrderBy	Key
OrderByAttribute	
ProcessingState	Unprocessed

Bước 5: Chính khóa cột (KeyColumns) và tên cột (Name Column) của thuộc tính Month. Vì thuộc tính này sẽ lấy khóa cột gồm chính nó và những thuộc tính cấp cao hơn.

Chuyển sang tab Dimension Structure, cột Attributes, right-click vào thuộc tính Day và chọn Properties. Tại cửa sổ Properties, chọn KeyColumns. Thêm các thuộc tính cấp cao hơn vào KeyColumns, sau đó chọn OK để hoàn tất.



Tại cửa sổ Properties, chọn NameColumn và chọn tên thuộc tính sẽ hiển thị trên Hierachy là Month:



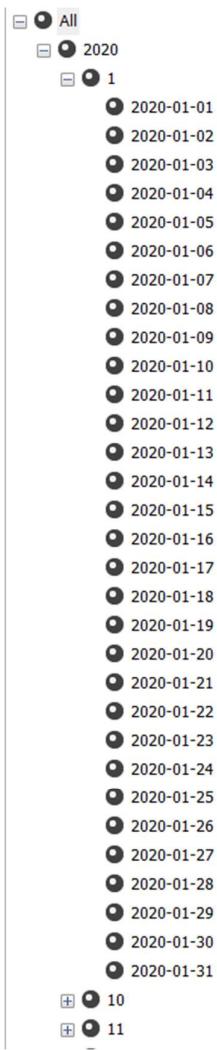
AttributeHierarchyOrder	True
ExtendedType	
GroupingBehavior	EncourageGrouping
InstanceSelection	None
MemberNamesUnique	False
VisualizationProperties	
Parent-Child	
MembersWithData	NonLeafDataVisible
MembersWithDataCap	
NamingTemplate	
RootMemberIf	ParentsBlankSelfOrMissing
UnaryOperatorColumn	(none)
Source	
CustomRollupColumn	(none)
CustomRollupProperties	(none)
KeyColumns	(Collection)
NameColumn	Dim_Timestamp.Month ...
ValueColumn	(none)

Chỉnh thuộc tính OrderBy thành Key để Month được sắp xếp theo thứ tự tăng dần:

EstimatedCount	0
HasLineage	False
IsAggregatable	True
OrderBy	Key
OrderByAttribute	
ProcessingState	Unprocessed

Bước 6: Thực hiện deploy project để đảm bảo không có lỗi xảy ra sau quá trình phân cấp.

Chuyển sang tab Browser để kiểm tra xem các thuộc tính đã phân cấp đúng hay chưa:



3.7. Thực hiện các truy vấn sử dụng SSAS, Pivot table

3.7.1. Câu truy vấn 1

Nội dung câu truy vấn: truy vấn drill down tổng số cuộc tấn công bằng phương thức TCP theo từng tháng năm 2022

3.7.1.1. Sử dụng SSAS

Bước 1: Vào mục Browser, kéo từ bảng Dim Attack thuộc tính Protocol. Tại cột Operator chọn “Equal” và Filter Expression chọn vào TCP, kéo từ bảng Dim Timestamp thuộc tính Year. Tại cột Operator chọn “Equal” và Filter Expression chọn vào 2022. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Param...
Dim Attack	Protocol	Equal	{ TCP }	<input type="checkbox"/> <input checked="" type="checkbox"/>
Dim Timestamp	Year	Equal	{ 2022 }	<input type="checkbox"/> <input checked="" type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input checked="" type="checkbox"/>
Year	Month	Fact Cyber Security Co...		
2022	1	295		
2022	10	321		
2022	11	272		
2022	12	295		
2022	2	273		
2022	3	305		
2022	4	272		
2022	5	300		
2022	6	293		
2022	7	345		
2022	8	292		
2022	9	291		

3.7.1.2. Sử dụng Pivot table

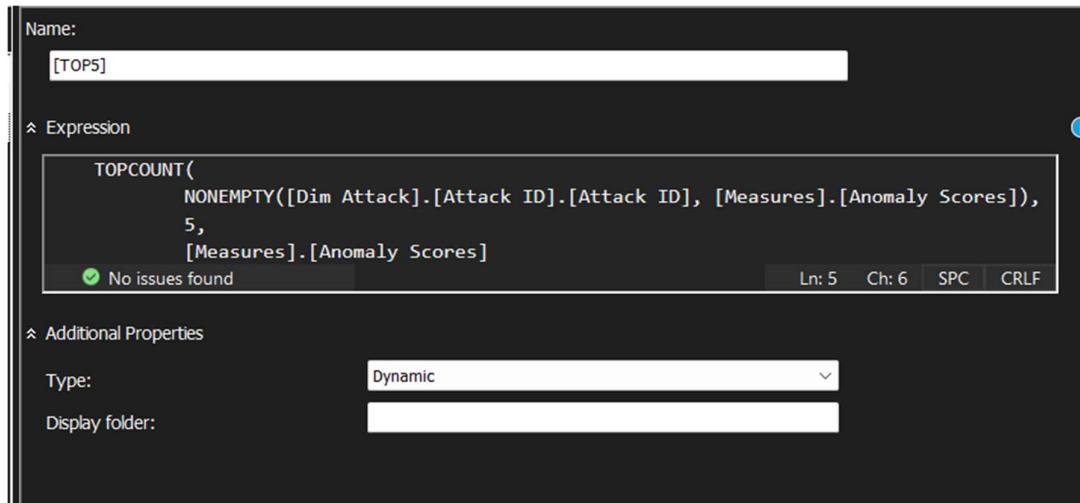
Protocol	TCP
Year	2022
Row La	Fact Cyber Security Count
	2022
⊕ 1	295
⊕ 2	273
⊕ 3	305
⊕ 4	272
⊕ 5	300
⊕ 6	293
⊕ 7	345
⊕ 8	292
⊕ 9	291
⊕ 10	321
⊕ 11	272
⊕ 12	295
Grand Tot	3554

3.7.2. Câu truy vấn 2

Nội dung câu truy vấn: Liệt kê top 5 cuộc tấn công có điểm cao nhất và phương thức của chúng.

3.7.2.1. Sử dụng SSAS

Bước 1: Trong Cube, vào mục Calculation, tạo mới nameset:



Bước 2: Vào mục Browser, kéo từ bảng Dim Attack thuộc tính AttackId. Tại cột Operator chọn “In” và Filter Expression chọn vào nameset “TOP5” vừa tạo. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Param..
Dim Attack	# Attack ID	In	TOP5	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Attack ID	Protocol	Anomaly Scores
1971	TCP	99.99
11771	ICMP	99.99
17786	ICMP	100
27605	ICMP	99.99
28740	TCP	99.99

3.7.2.2. Sử dụng Pivot table

The screenshot shows a Microsoft Power BI interface. On the left is a data grid with columns A through F. Row 1 is labeled "Row Labels" and "Anomaly Scores". Rows 2 through 11 show data points: 1971 (TCP, 99.98999786), 11771 (ICMP, 99.98999786), 17786 (ICMP, 100), 27605 (ICMP, 99.98999786), and 28740 (TCP, 99.98999786). The right side is the "PivotTable Fields" pane, which includes a search bar and a list of fields under "Fact Cyber Security": Anomaly Scores (selected), Fact Cyber Security Count, Maximum Packet Length, and Packet Length. Below this is a section for "Dim Attack" and a "Filters" section.

A	B	C	D	E	F
1 Row Labels	Anomaly Scores				
2 1971					
3 TCP	99.98999786				
4 11771					
5 ICMP	99.98999786				
6 17786					
7 ICMP	100				
8 27605					
9 ICMP	99.98999786				
10 28740					
11 TCP	99.98999786				
12					
13					
14					
15					
16					
17					

3.7.3. Câu truy vấn 3

Nội dung câu truy vấn: Liệt kê tổng số cuộc tấn công từng tháng trong năm với loại là mã độc.

3.7.3.1. Sử dụng SSAS

Bước 1: Vào mục Browser, kéo từ bảng Dim Attack thuộc tính Attack Type. Tại cột Operator chọn “Equal” và Filter Expression chọn vào Malware. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Param
Dim Attack	Attack Type	Equal	{ Malware }	<input type="checkbox"/>
<Select dimension>				<input type="checkbox"/>

Year	Month	Fact Cyber Security Co...
2020	1	284
2020	10	310
2020	11	306
2020	12	293
2020	2	263
2020	3	276
2020	4	252
2020	5	305
2020	6	258
2020	7	300
2020	8	313
2020	9	278
2021	1	285
2021	10	298
2021	11	296
2021	12	285
2021	2	249
2021	3	295
2021	4	299
2021	5	266
2021	6	294
2021	7	304
2021	8	303
2021	9	303
2022	1	296
2022	10	298
2022	11	322
2022	12	300
2022	2	266
2022	3	304

3.7.3.2. Sử dụng Pivot table

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Attack Type	Malware											
2													
3	Row Labels	Fact Cyber Security Count											
4	2020												
5	#1	284											
6	#10	310											
7	#11	306											
8	#12	293											
9	#2	263											
10	#3	276											
11	#4	252											
12	#5	305											
13	#6	258											
14	#7	300											
15	#8	313											
16	#9	278											
17	2021												
18	#1	285											
19	#10	298											
20	#11	296											
21	#12	285											
22	#2	249											
23	#3	295											
24	#4	299											
25	#5	266											

3.7.4. Câu truy vấn 4

Nội dung câu truy vấn: Truy vấn tổng điểm Anomaly Scores của các cuộc tấn công theo từng Phương thức (Protocol) và Năm với phương thức là Malware.

3.7.4.1. Sử dụng SSAS

Bước 1: Vào mục Browser, kéo từ bảng Dim Attack thuộc tính Attack Type. Tại cột Operator chọn “Equal” và Filter Expression chọn vào Malware. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension Hierarchy Operator Filter Expression Param...

Dim Attack Attack Type Equal { Malware }

<Select dimension>

Protocol	Year	Anomaly Scores
ICMP	2020	58840.98
ICMP	2021	58625.1
ICMP	2022	60400.5
ICMP	2023	43600.18
TCP	2020	57746.19
TCP	2021	58543.19
TCP	2022	59715.57
TCP	2023	42613.45
UDP	2020	55929.94
UDP	2021	58354.86
UDP	2022	59693.44
UDP	2023	45138.82

3.7.4.2. Sử dụng Pivot table

The screenshot shows a PivotTable setup in Excel. The PivotTable Fields pane on the right lists fields: Dim Attack (Attack ID, Attack Type), Protocol, and Anomaly Scores. The Rows area contains 'Attack Type' and 'Malware'. The Values area contains 'Anomaly Scores'. The Data area shows a hierarchical breakdown of Anomaly Scores by Protocol (ICMP, TCP, UDP) and Year (2020, 2021, 2022, 2023). A Grand Total of 659202.1875 is displayed.

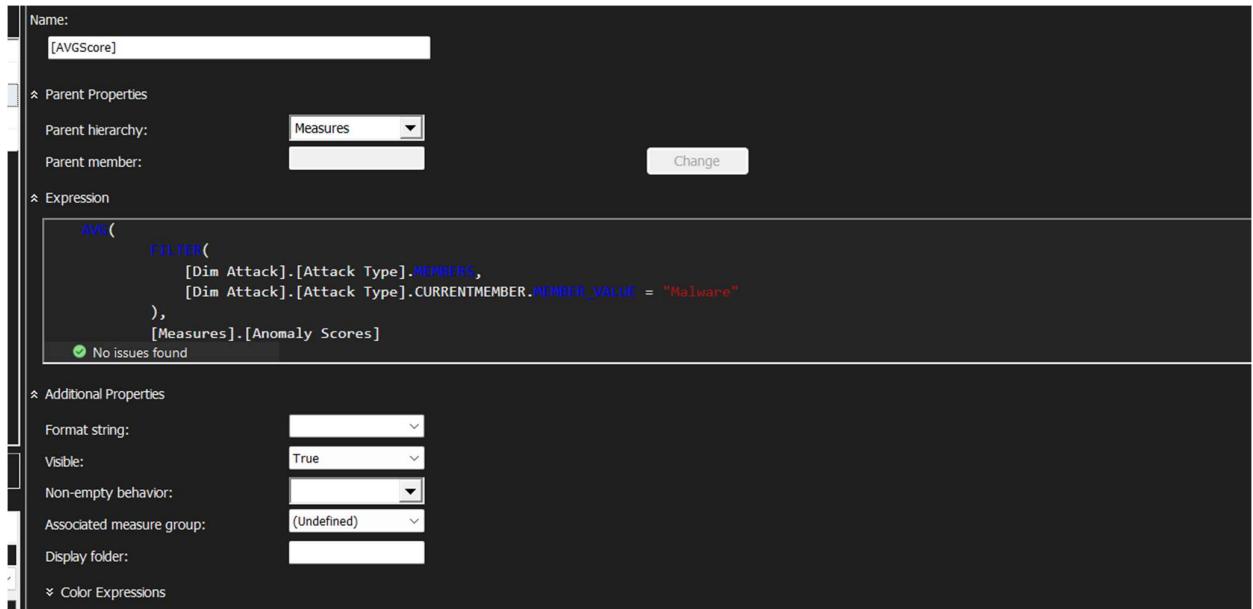
Attack Type	Malware	Row Labels	Anomaly Scores
ICMP		2020	58840.98438
ICMP		2021	58625.09766
ICMP		2022	60400.49609
ICMP		2023	43600.17969
TCP		2020	57746.19141
TCP		2021	58543.19141
TCP		2022	59715.57031
TCP		2023	42613.45313
UDP		2020	55929.94141
UDP		2021	58354.85938
UDP		2022	59693.4375
UDP		2023	45138.82031
Grand Total		659202.1875	

3.7.5. Câu truy vấn 5

Nội dung câu truy vấn: Truy vấn tính điểm trung bình của các cuộc tấn công Malware theo từng tháng trong năm 2022.

3.7.5.1. Sử dụng SSAS

Bước 1: Tạo Calculation tính trung bình thời lượng phim của top 5 phim có điểm bình chọn cao nhất:



The screenshot shows the 'Calculation Properties' dialog in SSAS. The 'Name' field is set to '[AVGScore]'. Under 'Parent Properties', 'Parent hierarchy' is set to 'Measures'. The 'Expression' section contains the following DAX code:

```
AVG(
    FILTER(
        [Dim Attack].[Attack Type].MEMBERS,
        [Dim Attack].[Attack Type].CURRENTMEMBER.MEMBER_VALUE = "Malware"
    ),
    [Measures].[Anomaly Scores]
)
```

Below the expression, a green checkmark indicates 'No issues found'. The 'Additional Properties' section includes fields for 'Format string', 'Visible' (set to 'True'), 'Non-empty behavior', 'Associated measure group' (set to '(Undefined)'), and 'Display folder'. The 'Color Expressions' section is collapsed.

Bước 2: Vào mục Browser, kéo từ bảng Dim Time Stamp thuộc tính Attack Type. Tại cột Operator chọn “Equal” và Filter Expression chọn vào 2022. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression
Dim Timestamp	# Year	Equal	{ 2022 }
Month	AVGScore		
1	14487.0517578125		
10	14801.4140625		
11	17114.970703125		
12	14495.2421875		
2	13636.30078125		
3	16021.0830078125		
4	13738.482421875		
5	14688.650390625		
6	16376.890625		
7	15763.9013671875		
8	14881.41015625		
9	13804.1796875		

3.7.5.2. Sử dụng Pivot table

	A	B	C	D	E	F	G	H
1	Year	2022						
2	Attack Type	Malware						
3								
4	Row Labels	AVGScore						
5	1	14487.05176						
6	10	14801.41406						
7	11	17114.9707						
8	12	14495.24219						
9	2	13636.30078						
10	3	16021.08301						
11	4	13738.48242						
12	5	14688.65039						
13	6	16376.89063						
14	7	15763.90137						
15	8	14881.41016						
16	9	13804.17969						
17	Grand Total	179809.5781						
18								
19								
20								

3.7.6. Câu truy vấn 6

Nội dung câu truy vấn: Liệt kê số cuộc tấn công bị chặn theo năm.

3.7.6.1. Sử dụng SSAS

Bước 1: Vào mục Browser, kéo từ bảng Dim Victim thuộc tính Action Taken. Tại cột Operator chọn “Equal” và Filter Expression chọn vào Blocked. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Param...
Dim Victim	Action Taken	Equal	{ Blocked }	<input type="checkbox"/> <input checked="" type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input checked="" type="checkbox"/>

Year	Fact Cyber Security Co...
2020	3535
2021	3480
2022	3634
2023	2742

3.7.6.2. Sử dụng Pivot table

Action Taken	Blocked
Row Labels	Fact Cyber Security Count
2020	3535
2021	3480
2022	3634
2023	2742
Grand Total	13391

3.7.7. Câu truy vấn 7

Nội dung câu truy vấn: Truy vấn chiều dài các gói tin đã bị hệ thống bảo mật bỏ qua theo từng năm.

3.7.7.1. Sử dụng SSAS

Bước 1: Tạo Calculation tính trung bình thời lượng phim của top 5 phim có điểm bình chọn cao nhất:

Name: [Average Packet Length Ignored]

Parent Properties (i)

Parent hierarchy: Measures ▼

Parent member: Change

Expression

```
AVG(
    FILTER(
        [Dim Victim].[Action Taken].MEMBERS,
        [Dim Victim].[Action Taken].CURRENTMEMBER IS [Dim Victim].[Action Taken].&
            [Ignored]
    )
)
```

✓ No issues found Ln: 8 Ch: 10 SPC MIXED

Additional Properties

Format string: ▼

Bước 2 Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Param...
<Select dimension>				<input type="checkbox"/> <input checked="" type="checkbox"/>
Year	Average Packet Length...			
2020	2735139			
2021	2726189			
2022	2704727			
2023	2089351			

3.7.7.2. Sử dụng Pivot table

3.7.8. Câu truy vấn 8

Nội dung câu truy vấn: Liệt kê số lượng các cuộc tấn công thành công theo từng phương thức.

3.7.8.1. Sử dụng SSAS

Bước 1: Vào mục Browser, kéo từ bảng Dim Victim thuộc tính Action Taken. Tại cột Operator chọn “Not Equal” và Filter Expression chọn vào Blocked

Bước 2 Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Protocol	Fact Cyber Security Co...
ICMP	8820
TCP	8689
UDP	8659

3.7.8.2. Sử dụng Pivot table

Action Taken	(Multiple Items)
Row Labels	Fact Cyber Security Count
ICMP	8820
TCP	8689
UDP	8659
Grand Total	26168

3.7.9. Câu truy vấn 9

Nội dung câu truy vấn: Liệt kê top 10 những cuộc tấn công có độ dài lớn nhất của hai nước Nhật Bản và Mỹ.

3.7.9.1. Sử dụng SSAS

Bước 1: Trong Cube, vào mục Calculation, tạo mới nameset:

Name: [Top 10 Attacks by Packet Length]

Expression:

```
TOPCOUNT(
    FILTER(
        NONEMPTY(
            [Dim Attack].[Attack ID].[Attack ID].MEMBERS,
            ([Measures].[Packet Length])
        ),
        [Dim Attack].[Country].CURRENTMEMBER IS [Dim Attack].[Country].[JP]
        OR [Dim Attack].[Country].CURRENTMEMBER IS [Dim Attack].[Country].[US]
    ),
    10,
    ([Measures].[Packet Length])
)
```

No issues found

Type: Dynamic

Bước 2: Vào mục Browser, kéo từ bảng Dim Attack thuộc tính AttackId. Tại cột Operator chọn “In” và Filter Expression chọn vào nameset “[Top 10 Attacks by Packet Length]” vừa tạo. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Para...
Dim Attack	Attack ID	In	Top 10 Attacks by Packet Length	<input type="checkbox"/> <input checked="" type="checkbox"/>
<Select dimension>				

Attack ID	Packet Length	Anomaly Scores
4825	1500	56.79
5884	1500	46.8
6947	1500	26.67
17313	1500	99.9
17652	1500	3.92
17804	1500	53.68
20235	1500	28.26
23365	1500	50.73
31008	1500	56.17
34350	1500	98.54

3.7.9.2. Sử dụng Pivot table

The screenshot shows a Microsoft Excel spreadsheet with a PivotTable. The PivotTable Fields pane on the right indicates that 'Country' is assigned to Row Labels, and 'Anomaly Scores' and 'Packet Length' are assigned to the Values area. A 'Sets' item named 'Top 10 Attacks by Packet Length' is selected. The main table displays the following data:

Row Labels	Anomaly Scores	Packet Length
4825	56.79000092	1500
5884	46.79999924	1500
6947	26.67000008	1500
17313	99.90000153	1500
17652	3.920000076	1500
17804	53.68000031	1500
20235	28.26000023	1500
23365	50.72999954	1500
31008	56.16999817	1500
34350	98.54000092	1500

3.7.10. Câu truy vấn 10

Nội dung câu truy vấn: Truy vấn cuộc tấn công đã được cảnh báo nhưng bị bỏ qua trong năm 2020

3.7.10.1. Sử dụng SSAS

Bước 1: Vào mục Browser, kéo từ bảng Dim Victim thuộc tính Alert Warnings, Action Taken, Timestamp. Tại cột Operator chọn Equal Alert Triggered, Ignored, 2020

Dimension	Hierarchy	Operator	Filter Expression	Param...
Dim Victim	Alerts Warnings	Equal	{ Alert Triggered }	<input type="checkbox"/> <input checked="" type="checkbox"/>
Dim Victim	Action Taken	Equal	{ Ignored }	<input type="checkbox"/> <input checked="" type="checkbox"/>
Dim Timestamp	Year_Month_Timestamp	Equal	{ 2020 }	<input type="checkbox"/> <input checked="" type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input checked="" type="checkbox"/>
				<input type="checkbox"/> <input checked="" type="checkbox"/>
Attack ID	Year	Packet Length	Anomaly Scores	

Bước 2: Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Attack ID	Year	Packet Length	Anomaly Scores
19	2020	124	44.75
33	2020	373	3.87
35	2020	521	60.57
84	2020	173	73.69
118	2020	1292	61.3
119	2020	1026	36.95
150	2020	1103	91.04
173	2020	1090	2.09
178	2020	350	28.97
199	2020	432	30.26
219	2020	798	53.56
221	2020	452	71.97
239	2020	1303	3.03
281	2020	98	99.08
295	2020	1188	36.73
331	2020	815	40.9
355	2020	369	67.22
359	2020	1373	86.44
410	2020	1433	48.06
423	2020	650	31.63
446	2020	1284	71.9
540	2020	245	18.64
556	2020	126	79.85
590	2020	1356	52.15
601	2020	1350	5.31
603	2020	1204	25.52
666	2020	135	3.17
678	2020	634	65
729	2020	273	31.51
736	2020	152	87.88

3.7.10.2. Sử dụng Pivot table

The screenshot shows a Microsoft Power BI interface. On the left is a data grid with columns A, B, C, D, E. Rows include 'Alerts Warnings', 'Action Taken', and 'Year_Month_Timestamp'. Below this is a table with 'Row Labels', 'Anomaly Scores', and 'Packet Length' columns. The table contains numerous rows of data. To the right is the 'PivotTable Fields' pane, which includes a search bar, a list of available fields (Fact Cyber Security Count, Packet Length, Dim Attack, Attack ID, Address ID, Attack Type, Counter), and sections for Filters, Columns, Rows, and Values. The 'Values' section under 'Columns' has 'Σ Values' selected, and under 'Rows' it has 'Attack ID' selected. Buttons for 'Defer Layout Update' and 'Update' are at the bottom.

A	B	C	D	E
Alerts Warnings	Alert Triggered			
Action Taken	Ignored			
Year_Month_Timestamp	2020			
Row Labels	Anomaly Scores	Packet Length		
19	44.75	124		
33	3.869999886	373		
35	60.56999969	521		
84	73.69000244	173		
118	61.29999924	1292		
119	36.95000076	1026		
150	91.04000092	1103		
173	2.08999914	1090		
178	28.96999931	350		
199	30.26000023	432		
219	53.56000137	798		
221	71.97000122	452		
239	3.02999971	1303		
281	99.08000183	98		
295	36.72999954	1188		
331	40.90000153	815		
355	67.22000122	369		
359	86.44000244	1373		
410	48.06000137	1433		
423	31.62999916	650		
446	71.90000153	1284		
540	18.63999939	215		

3.7.11. Câu truy vấn 11

Nội dung câu truy vấn: Tổng độ dài các gói tin của từng phương thức của 20 cuộc tấn công có điểm AnomalyScore cao nhất.

3.7.11.1. Sử dụng SSAS

Bước 1: Trong Cube, vào mục Calculation, tạo mới nameset:

Name: [Top 20 Attacks]

Parent Properties

Parent hierarchy: Measures

Parent member: Change

Expression

```
TOPCOUNT(
    EXISTING [Dim Attack].[Attack ID].Members,
    20,
    [Measures].[Anomaly Scores]
)
No issues found
```

Ln: 5 Ch: 10 SPC LF

Additional Properties

Bước 2: Vào mục Browser, kéo từ bảng Dim Attack thuộc tính AttackId. Tại cột Operator chọn “In” và Filter Expression chọn vào nameset “[Top 20 Attacks]” vừa tạo. Sau đó kéo các thuộc tính để hiển thị câu truy vấn, ta được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Param...
Dim Attack	# Attack ID	In	Top 20 Attacks	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>
Protocol	Packet Length			
ICMP	10415370			
TCP	10203601			
UDP	10281887			

3.7.11.2. Sử dụng Pivot table

Packet Length C

Row Labels 7 8 11 13 16 11 77 1 1 12 (22 27 0 25 17 3 Grand Total

	7	8	11	13	16	11	77	1	1	12	(22	27	0	25	17	3	Grand Total
ICMP			1010	#	1347												10415370
TCP	#		1248		#	#											10203601
UDP	#	#	#														10281887
Grand Total	###	##	1248	1010	#	#	1347	994	#	30900858							

PivotTable Fields

Choose fields to add to report:

Search

Traffic Type

Sets

Top 10 Attacks by Packet Length

Top 20 Attacks

Dim Timestamp

> Year_Month_Timestamp

> More Fields

CHƯƠNG 4. QUÁ TRÌNH LẬP BÁO BIỂU

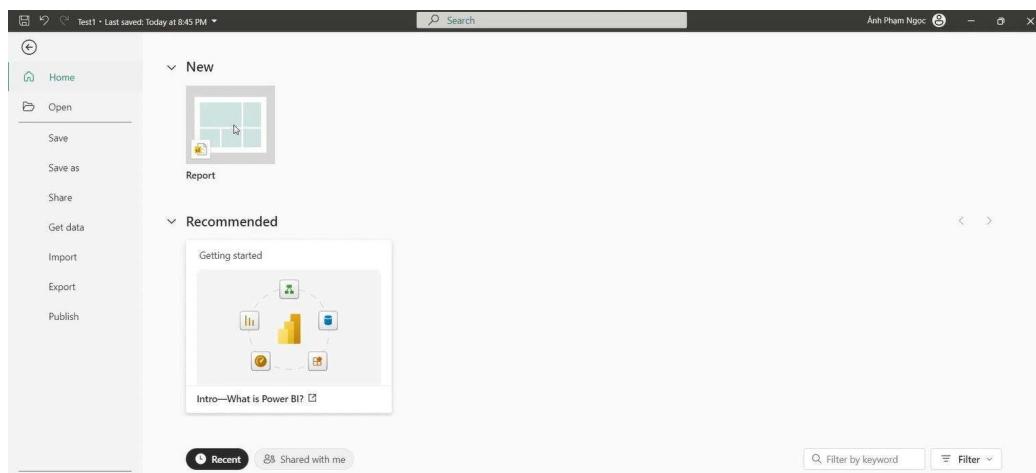
4.1. Quá trình lập báo biểu bằng PowerBI

4.1.1. Chuẩn bị công cụ

Đăng ký tài khoản và tải Power BI theo hướng dẫn tại đường link: [Sign in and download Power BI Desktop](#)

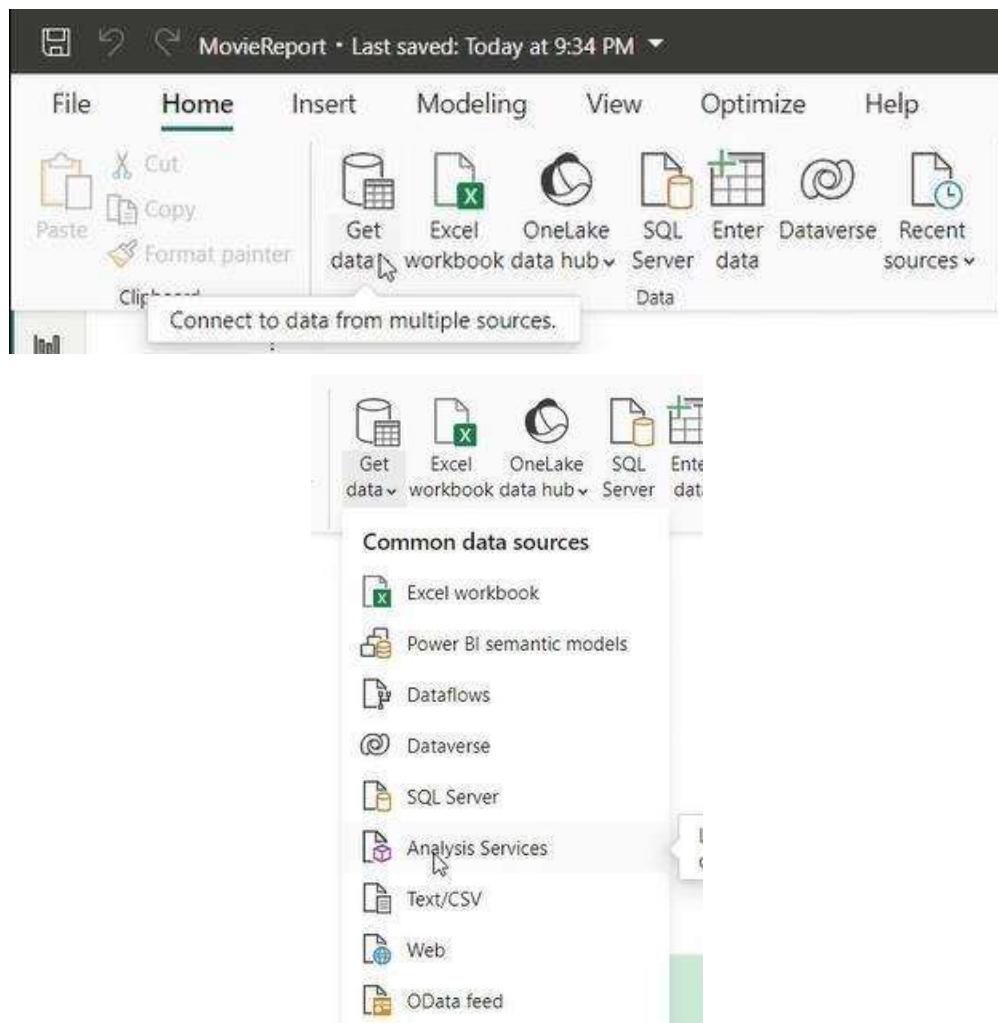
4.1.2. Tạo mới report và kết nối

- Tại mục New chọn Report để tạo mới.



- Kết nối với cube trong SSAS.

Bước 1: Chọn Get data tại mục Home trên thanh công cụ. Chọn vào Analysis Services.



Bước 2: Nhập Server và Database.

SQL Server Analysis Services database

Server ⓘ
HIPHIP

Database (optional)
Hello6

Import
 Connect live

▷ MDX or DAX query (optional)

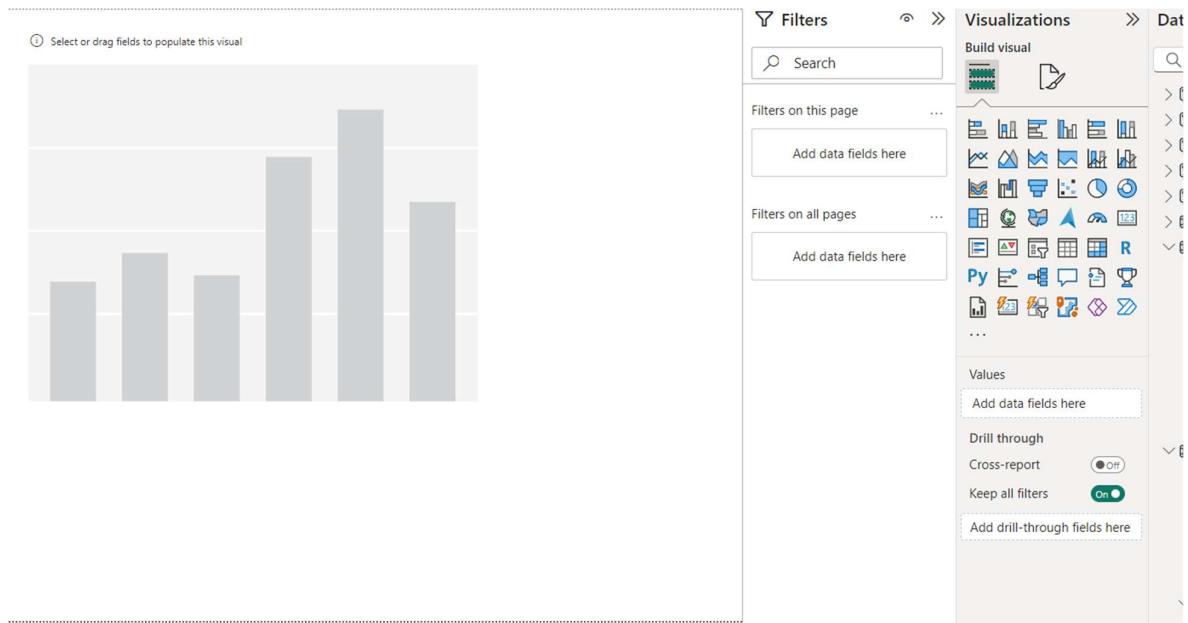
OK **Cancel**

- Server là server name kết nối Analysis Services trong SQL Server.

4.1.1. Tạo report 1

Câu truy vấn: Truy vấn drill down tổng số cuộc tấn công của từng phương thức theo từng tháng năm có kiểu tấn công là Mã độc.

Bước 1: Chọn biểu đồ Chart trong mục Visualizations.



Bước 2: Chọn từ Data các thuộc tính: Fact Cyber Security Count.

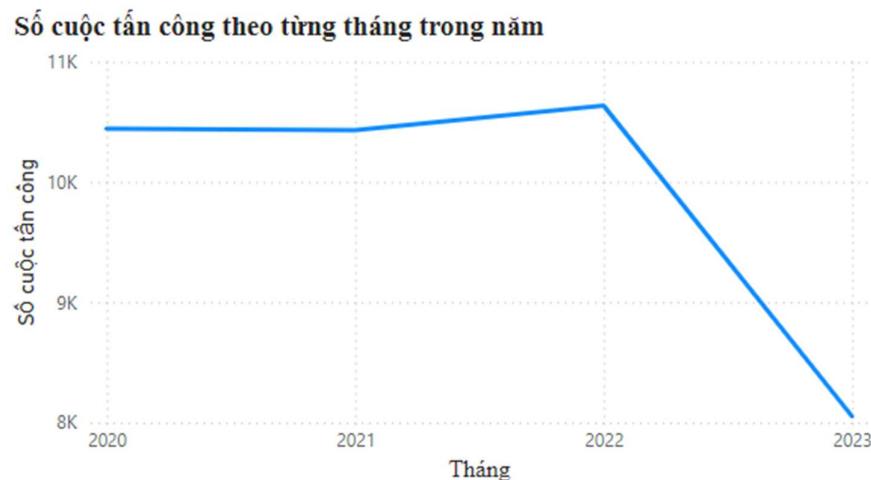
- Thêm điều kiện lọc cho AttackType: chọn vào “Malware”

The screenshot shows the Power BI Data View interface. On the left, there's a filter pane for 'Attack Type' set to 'is Malware'. Below it is a 'Basic filtering' section with a search bar and checkboxes for 'Select all', 'DDoS', 'Intrusion', and 'Malware', with 'Malware' checked. A 'Require single selection' checkbox is also present. In the center, there's a visualization area with a 'X-axis' set to 'Year_Month_Timestamp' (with 'Year' and 'Month' selected) and a 'Y-axis' set to 'Fact Cyber Security C...'. On the right, the 'Data' pane displays a hierarchy of dimensions and measures. Under 'Dim Attack', 'Fact Cyber Security' is expanded, showing 'Anomaly Scores' (unchecked), 'Fact Cyber Security Count' (checked), and 'Packet Length' (unchecked). Other dimensions like 'Dim Victim' and 'Dim Address' are also listed.

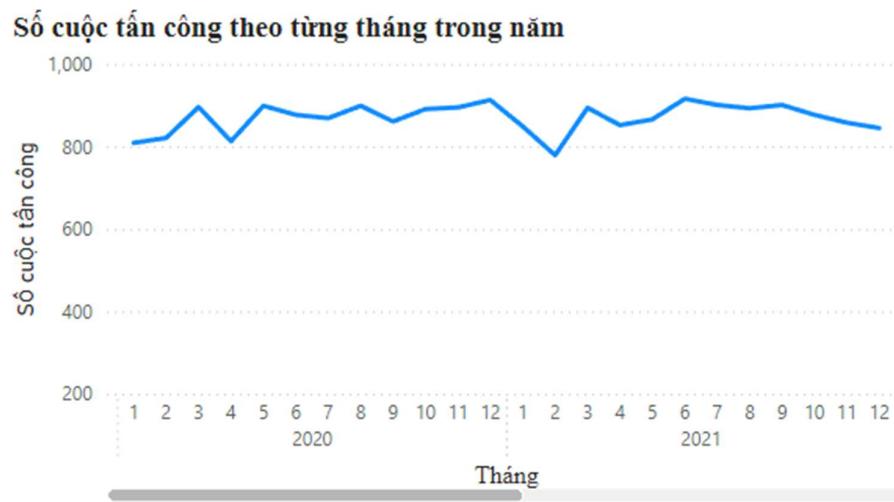
Bước 3: Thêm Legend các thuộc tính Year_Month_Timestamp-Year, Year_Month_Timestamp-Month

This screenshot shows the Power BI Data View with the same filter and visualization setup as the previous one. The 'X-axis' is still set to 'Year_Month_Timestamp' with 'Year' and 'Month' selected. In the 'Legend' section of the visualization pane, there is a dropdown menu with options: 'Add data fields here', 'Day', 'Month', 'Timestamp', 'Timestamp ID', 'Year', 'Year Month Timestamp', and 'Year Month'. The 'Year Month Timestamp' option is checked. The 'Y-axis' remains set to 'Fact Cyber Security C...'. The 'Data' pane on the right shows the expanded 'Fact Cyber Security' dimension with its three measures.

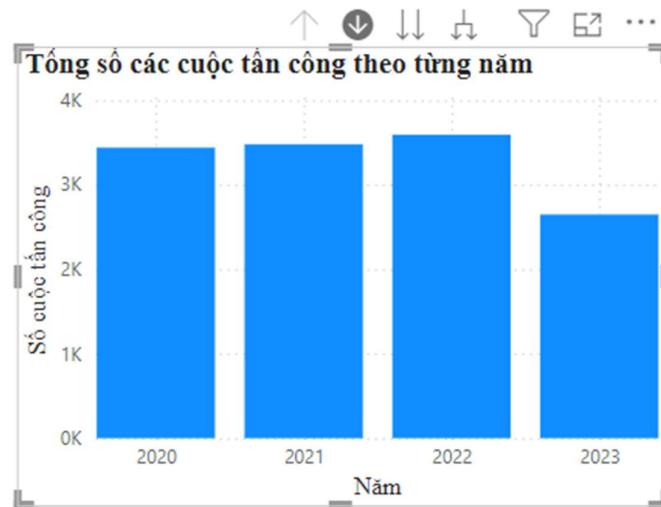
- Kết quả với biểu đồ tròn:



- Drill Down xuống tháng:

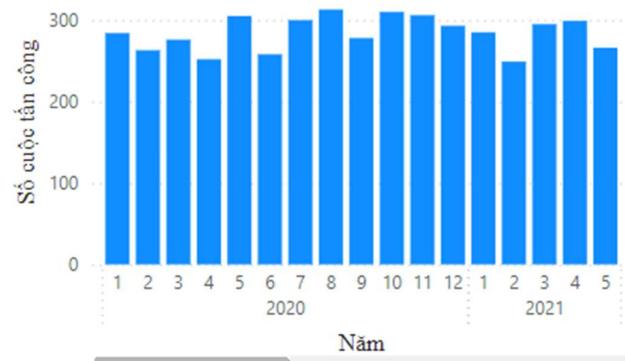


- Sơ đồ hình cột:



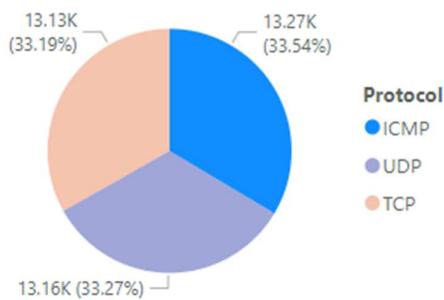
- Drill Down xuống tháng:

Tổng số các cuộc tấn công theo từng năm



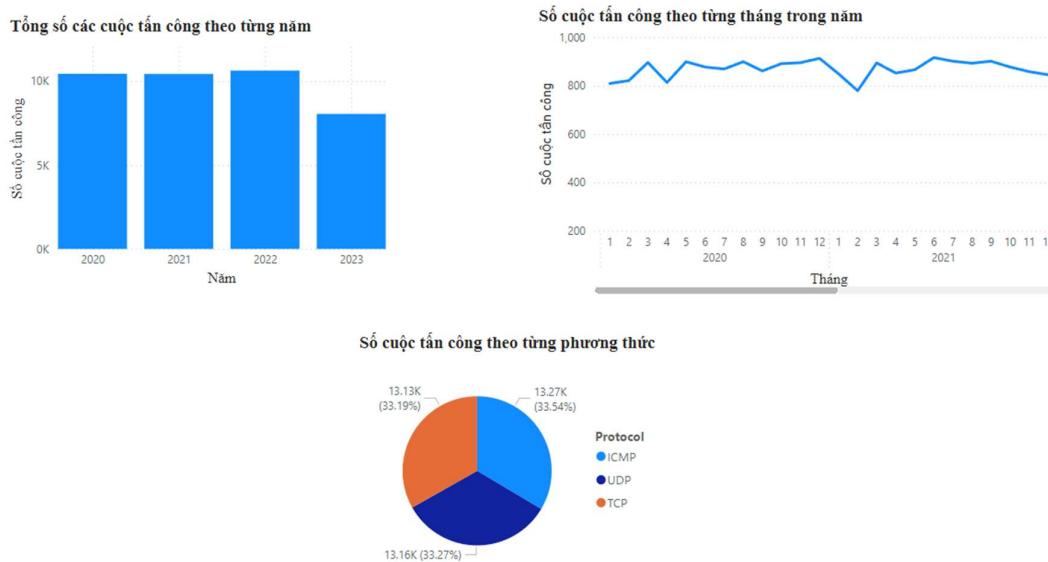
Sơ đồ hình tròn thể hiện, phương thức tấn công ta có thể chọn từng phương thức để thể hiện tổng số cuộc tấn công của phương thức đó theo từng tháng, năm.

Số cuộc tấn công theo từng phương thức



Bước 4: Tổng hợp lại report

Truy vấn drill down tổng số cuộc tấn công của từng phương thức theo từng tháng năm có kiểu tấn công là Mã độc



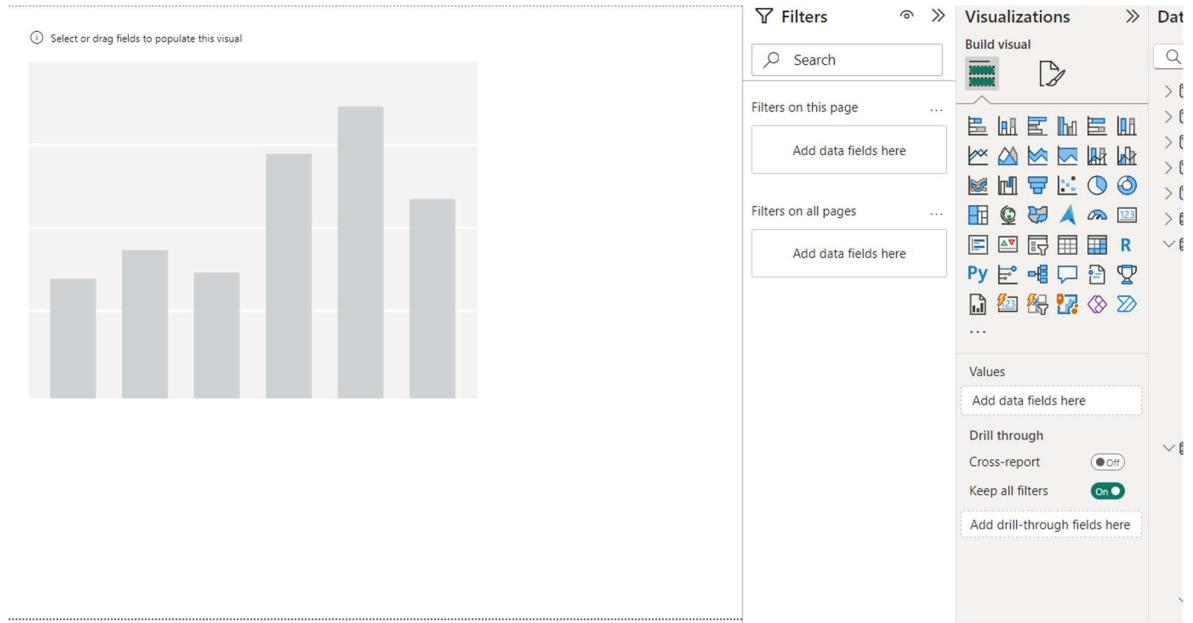
Nhận xét: từ biểu đồ ta có thể thấy:

- Phương thức được dung để tấn công nhiều nhất là ICMP
- Mỗi năm có thể có lên tới 10000 cuộc tấn công
- Tháng 3, 11, 6 là những tháng được sử dụng để tấn công nhiều nhất.

4.1.2. Tạo report 2

Câu truy vấn: Tổng độ dài các gói tin của từng phương thức hoặc từng năm của 20 cuộc tấn công có điểm AnomalyScore cao nhất.

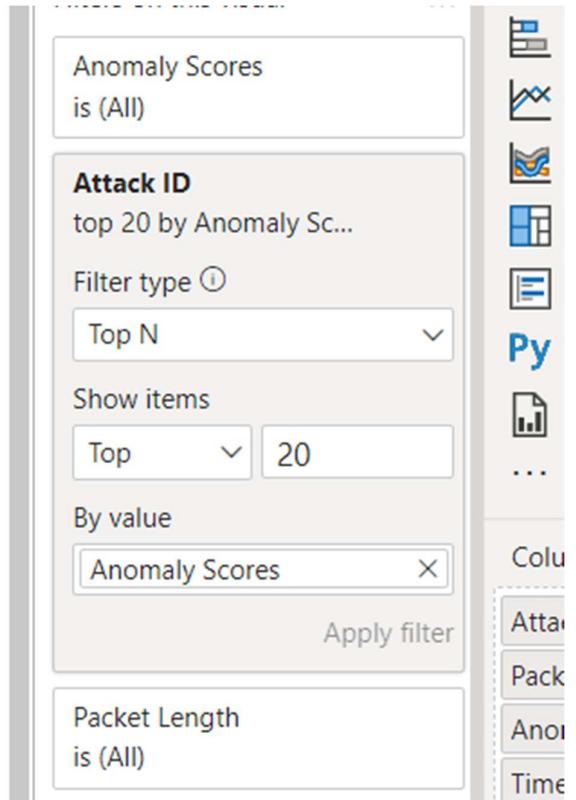
Bước 1: Chọn biểu đồ Chart trong mục Visualizations.



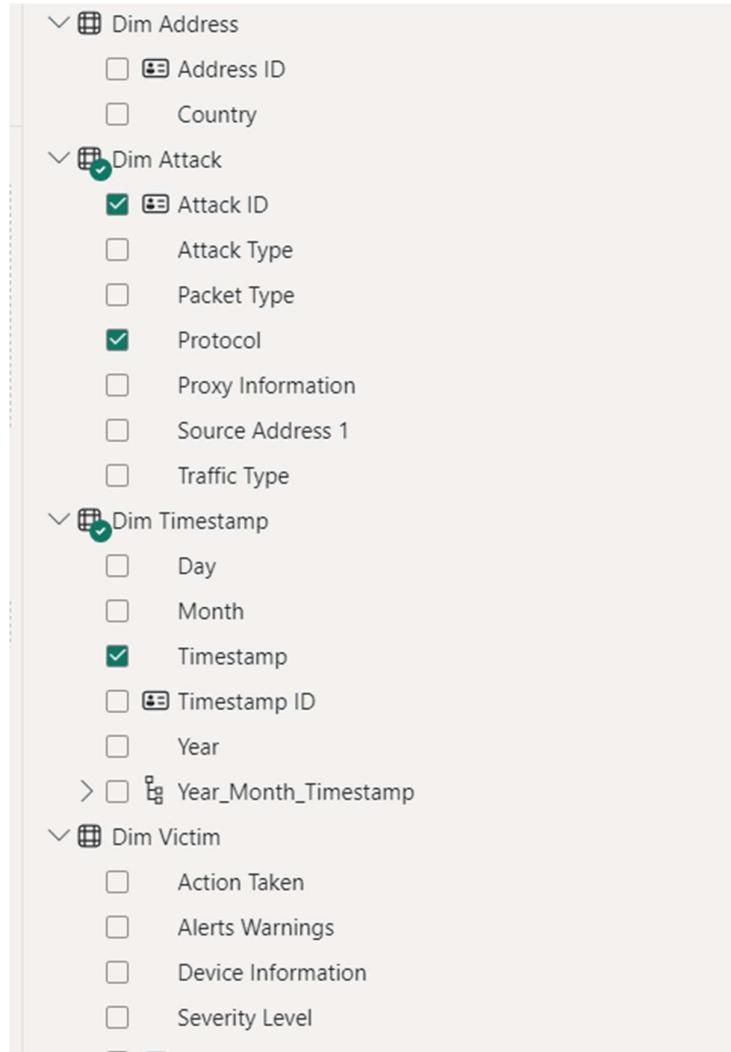
Bước 2: Chọn từ Data các thuộc tính: Packet Length , AnomalyScore,



- Thêm điều kiện lọc cho AttackId: chọn Show Items Top N , Nhập top 20, chọn thuộc tính tính toán là Anomaly Score .



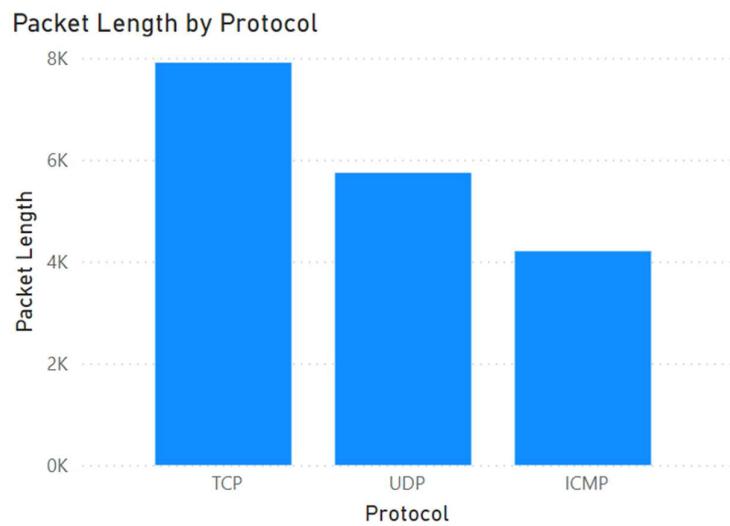
Bước 3: Thêm Legend các thuộc tính Timestamp, Protocol, AttackId



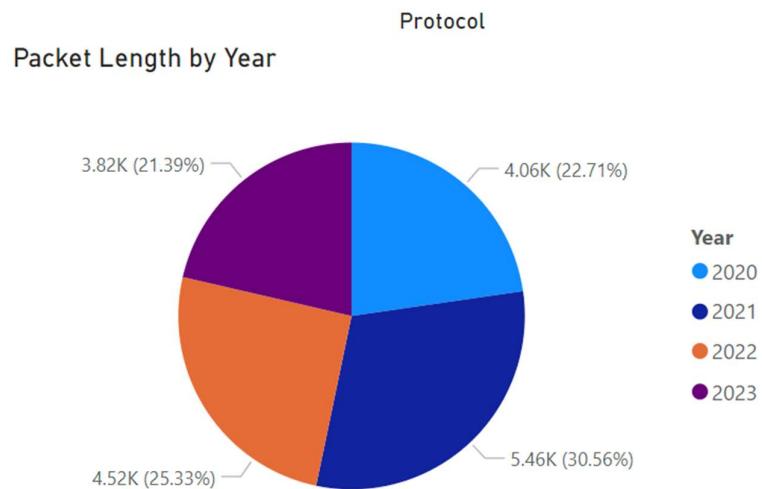
- Kết quả với biểu đồ bảng:

Attack ID	Packet Length	Anomaly Scores	Timestamp	Protocol
17786	384	100.00	2023-01-29	ICMP
1971	968	99.99	2020-06-09	TCP
11771	1010	99.99	2021-07-30	ICMP
27605	1173	99.99	2021-12-20	ICMP
28740	779	99.99	2021-01-06	TCP
34281	288	99.99	2021-01-06	ICMP
37869	571	99.99	2023-07-17	TCP
784	986	99.98	2022-10-27	UDP
11314	829	99.98	2021-12-20	UDP
12359	976	99.98	2022-09-07	TCP
33311	971	99.98	2020-08-30	TCP
35120	1003	99.98	2022-07-04	UDP
8933	269	99.97	2023-02-20	UDP
11316	1248	99.97	2023-01-22	TCP
25173	994	99.97	2021-07-10	TCP
27796	383	99.97	2021-08-07	UDP
35939	954	99.97	2020-07-09	UDP
20005	1400	99.95	2022-09-20	TCP
22270	1347	99.95	2023-01-03	ICMP
23856	157	99.95	2022-05-05	UDP
29172	264	99.95	2020-02-26	UDP
38780	898	99.95	2020-11-27	UDP
Total	17852	2,199.44		

- Kết quả với biểu đồ cột với tổng gói tin theo từng phương thức:



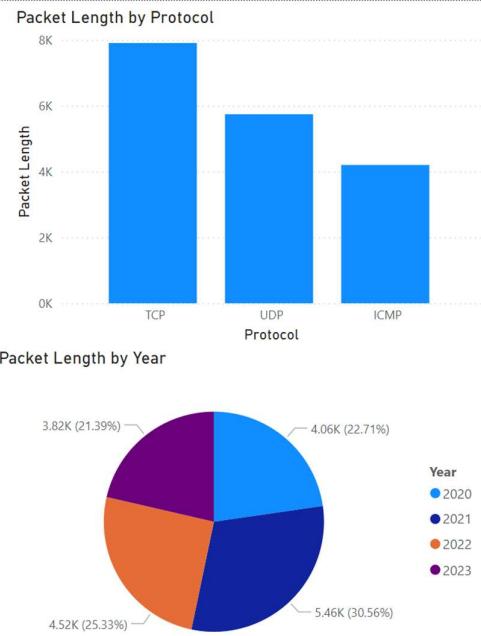
- Kết quả với biểu đồ cột với tổng gói tin theo từng năm:



Bước 4: Tổng hợp lại report

**TỔNG GÓI TIN CỦA TÙNG PHƯƠNG THỨC
HOẶC TÙNG NĂM CỦA TOP 20 CUỘC TẤN
CÔNG CÓ ĐIỂM CAO NHẤT**

Attack ID	Packet Length	Anomaly Scores	Timestamp	Protocol
17786	384	100.00	2023-01-29	ICMP
1971	968	99.99	2020-06-09	TCP
11771	1010	99.99	2021-07-30	ICMP
27605	1173	99.99	2021-12-20	ICMP
28740	779	99.99	2021-01-06	TCP
34281	288	99.99	2021-01-06	ICMP
37869	571	99.99	2023-07-17	TCP
784	986	99.98	2022-10-27	UDP
11314	829	99.98	2021-12-20	UDP
12359	976	99.98	2022-09-07	TCP
33311	971	99.98	2020-08-30	TCP
35120	1003	99.98	2022-07-04	UDP
8933	269	99.97	2023-02-20	UDP
11316	1248	99.97	2023-01-22	TCP
25173	994	99.97	2021-07-10	TCP
27796	383	99.97	2021-08-07	UDP
35939	954	99.97	2020-07-09	UDP
20005	1400	99.95	2022-09-20	TCP
22270	1347	99.95	2023-01-03	ICMP
23856	157	99.95	2022-05-05	UDP
29172	264	99.95	2020-02-26	UDP
38780	898	99.95	2020-11-27	UDP
Total	17852	2,199.44		



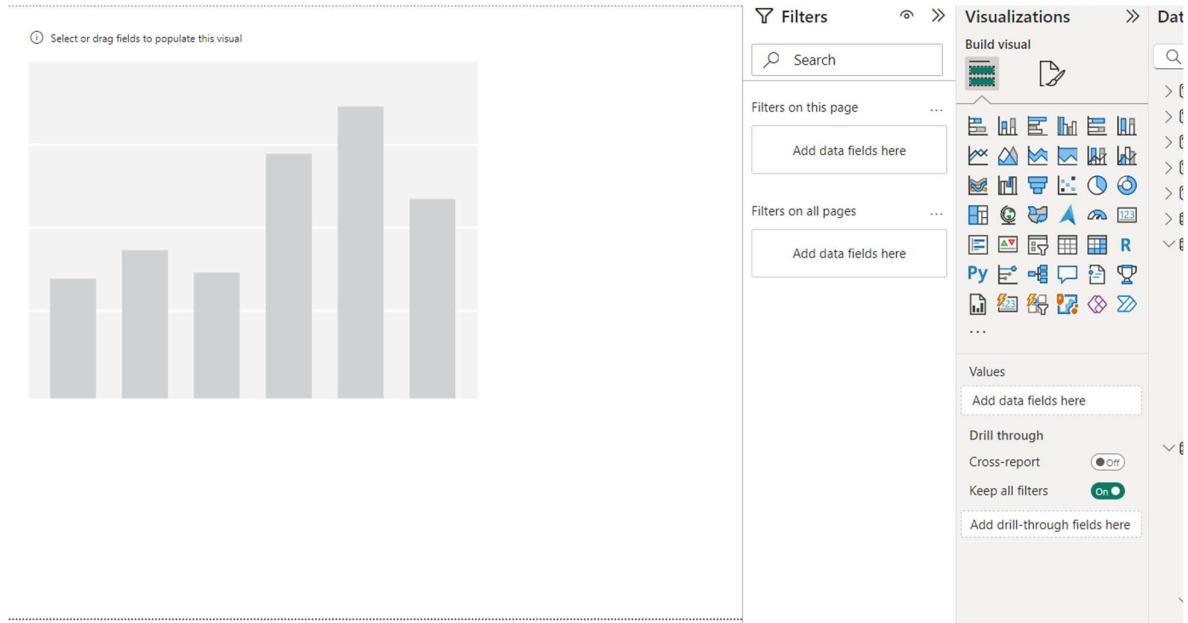
Nhận xét: từ biểu đồ ta có thể thấy:

- Trong 20 cuộc tấn công có điểm bất thường cao, thì phương thức TCP độ dài lớn nhất
- Trong 20 cuộc tấn công có điểm bất thường cao, thì tổng độ dài các gói tin trong năm 2021 là lớn nhất.
- Độ chênh lệch điểm 20 gói tiên lớn nhất không nhiều từ 0.05- 0.01

4.1.3. Tạo report 3

Câu truy vấn: Truy vấn tổng chiều dài gói tin theo từng phương thức trong từng tháng, năm và phân bổ của nó trên thế giới

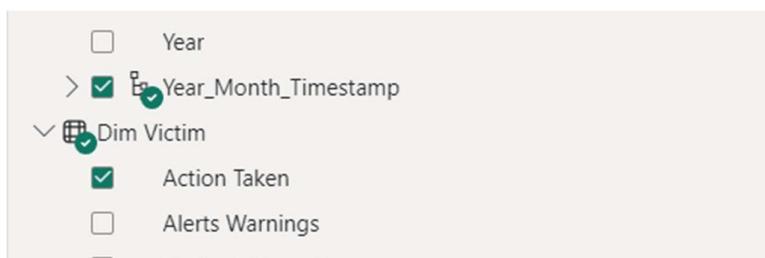
Bước 1: Chọn biểu đồ Chart trong mục Visualizations.



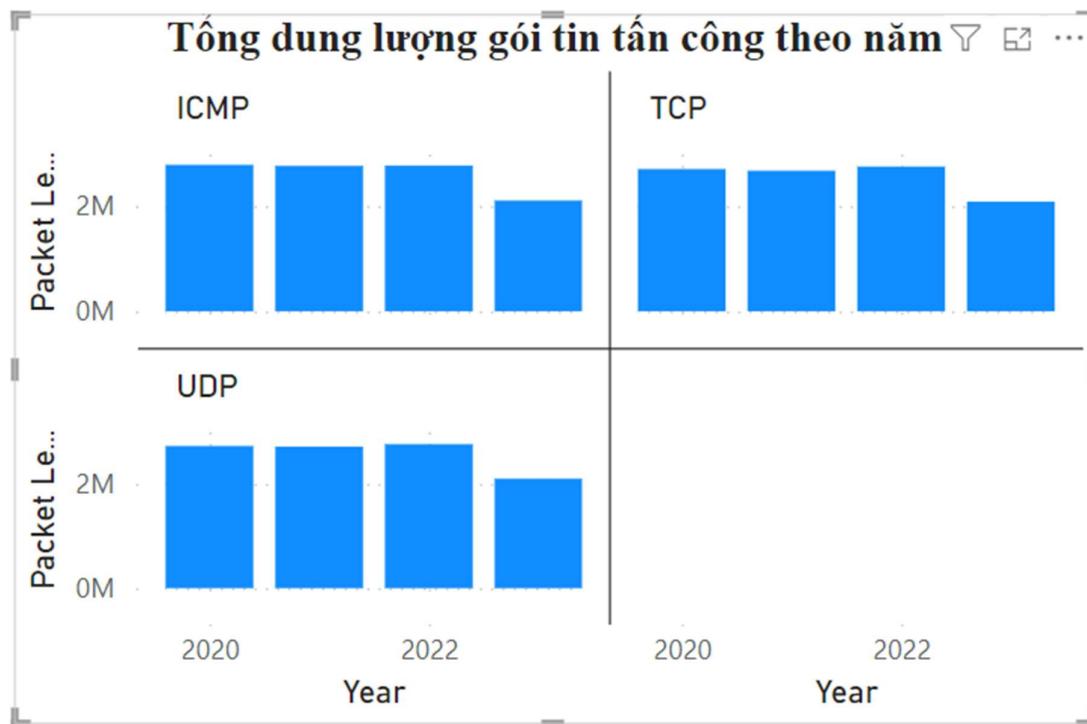
Bước 2: Chọn từ Data các thuộc tính: Packet Length



Bước 3: Thêm Legend các thuộc tính Year_Month_Timestamp, Action Taken.



- Kết quả với biểu đồ cột chồng



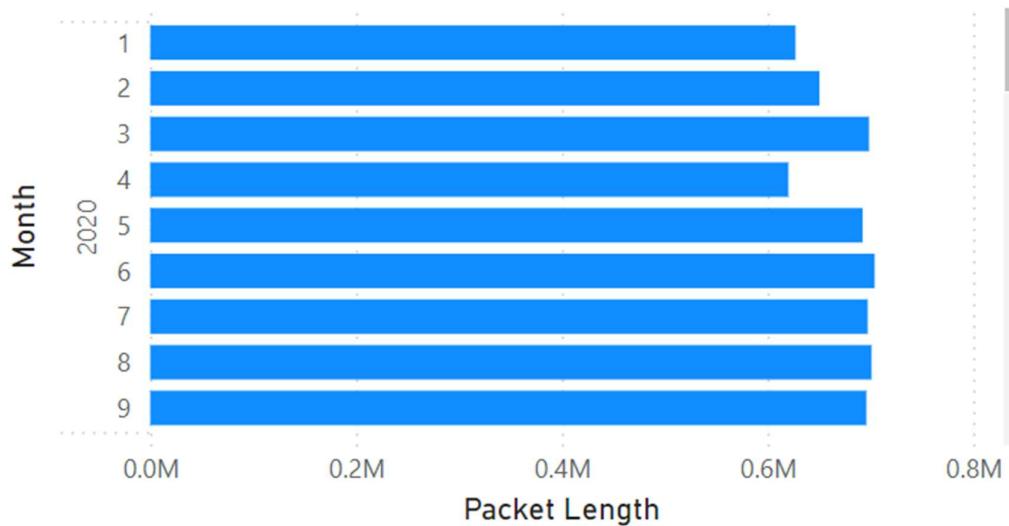
- Phân bố gói tin theo từng quốc gia:

Phân bố tổng dung lượng các gói tin được gửi trên toàn thế giới

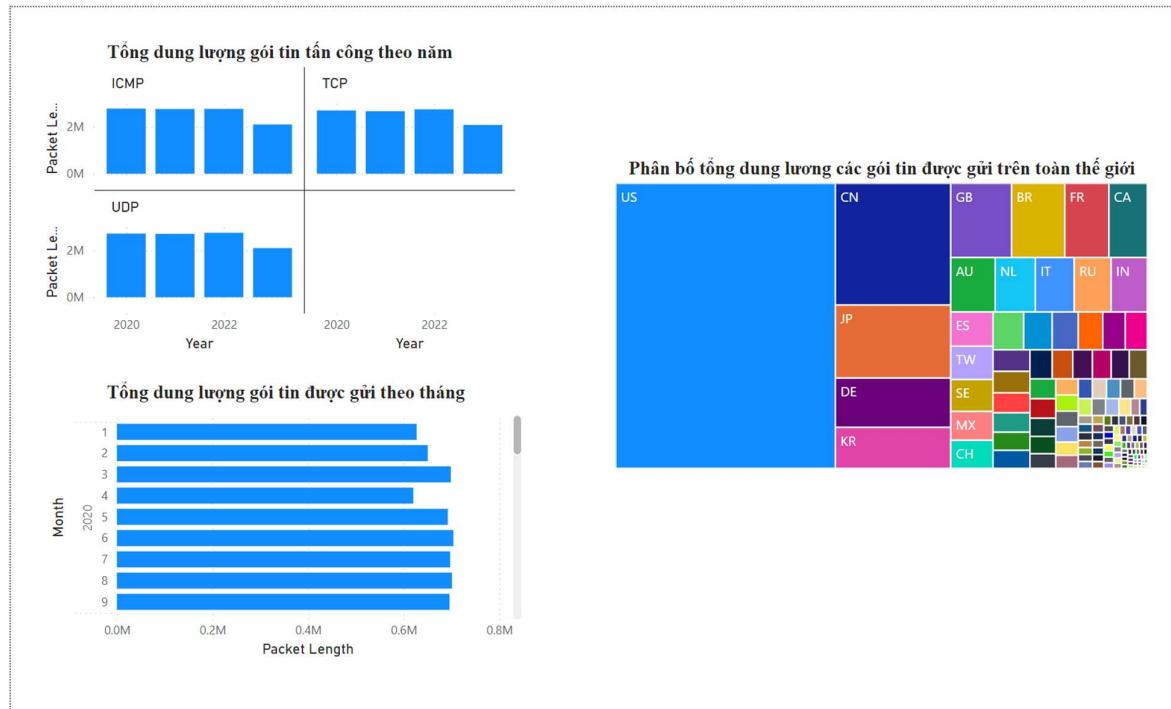


- Sơ đồ hình cột:

Tổng dung lượng gói tin được gửi theo tháng



Bước 4: Tổng hợp lại report



Nhận xét: từ biểu đồ ta có thể thấy:

- Mỹ là nước nhận được tổng độ dài gói tin nhiều nhất.

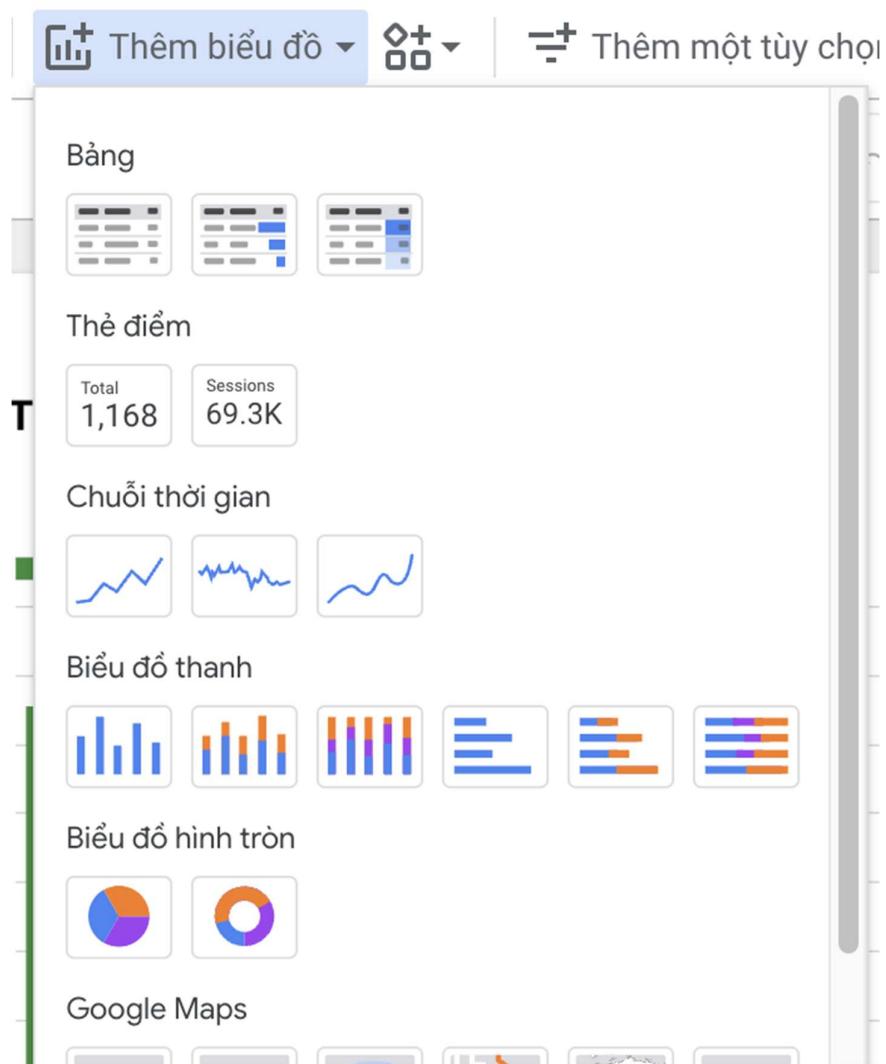
- Tháng 4 là tháng ít được sử dụng để tấn công nhất, tổng độ dài gói tin thấp nhất.
- Tổng độ dài gói tin có phương thức là TCP vào năm 2021 là cao nhất.

4.2. Quá trình lập báo biểu bằng Google Data Studio

4.2.1. Tạo report 1

Câu truy vấn: Truy vấn tổng điểm Anomaly Scores của các cuộc tấn công theo từng Phương thức (Protocol) và tháng với phương thức là Malware trong năm 2021

Bước 1: Chọn vào thêm biểu đồ, chọn vào biểu đồ Bảng và biểu đồ cột.



Bước 2: Chọn vào sơ đồ vừa tạo, tiến hành tạo bộ lọc cho sơ đồ

- Tạo bộ lọc mới gồm Tag là “Bao gồm”, Chọn trường là AttackType, chọn điều kiện là Bảng, Giá trị là “Malware”

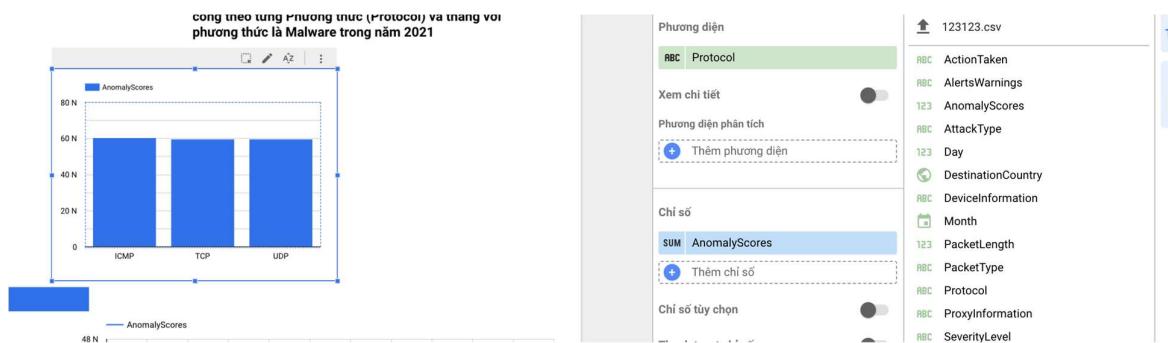
- Tạo bộ lọc mới gồm Tag là “Bao gồm”, Chọn trường là Year, chọn điều kiện là Bằng. Giá trị là “2022”

VÀ

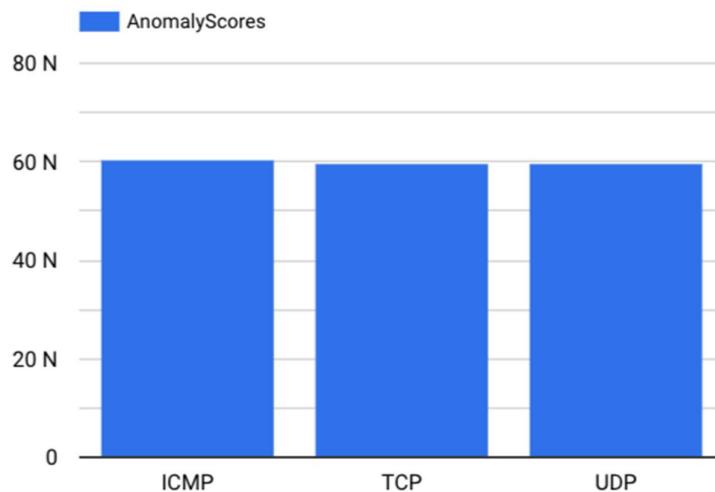
lọc này có 2 mèngh để

LƯU

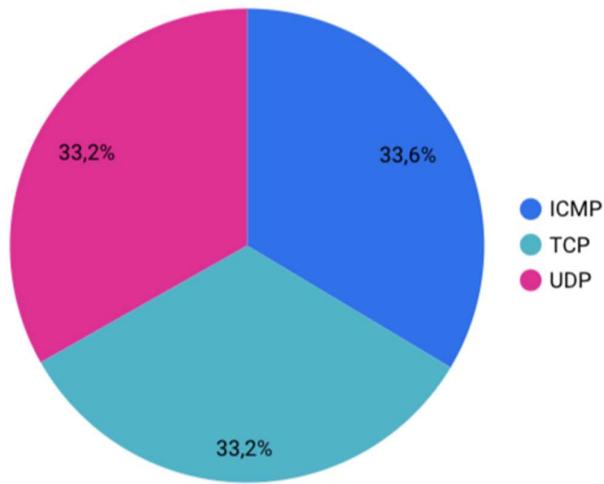
Bước 3: kéo Protocol từ bảng dữ liệu vào phần phương diện, kéo AnomalyScore vào mục chỉ số.



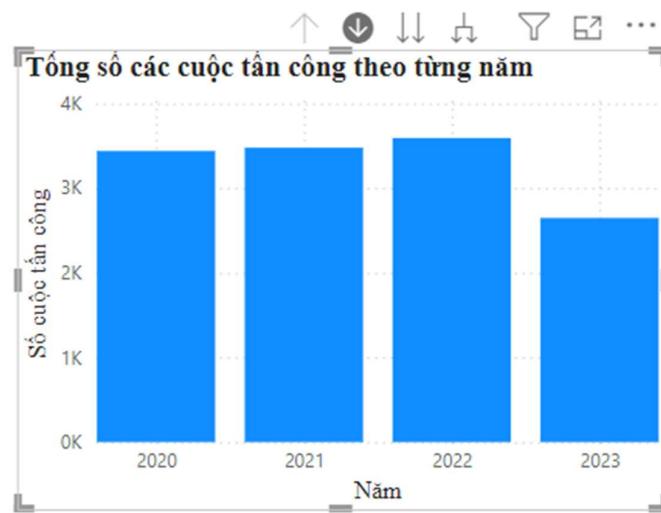
- Kết quả với biểu đồ hình cột:



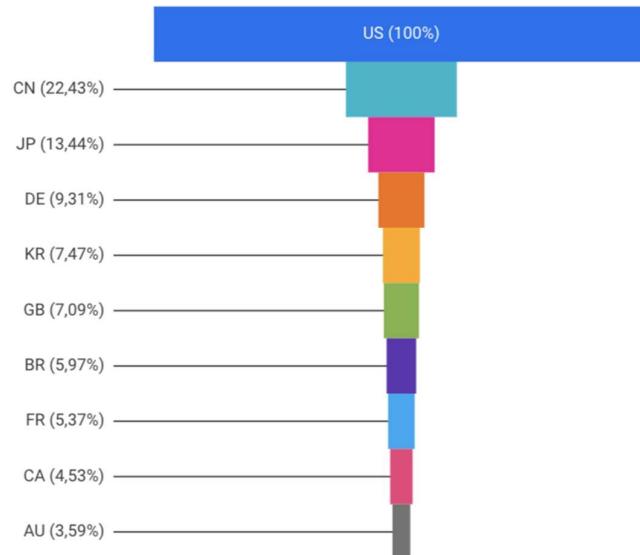
- Kết quả với biểu đồ hình tròn:



- Sơ đồ hình cột:

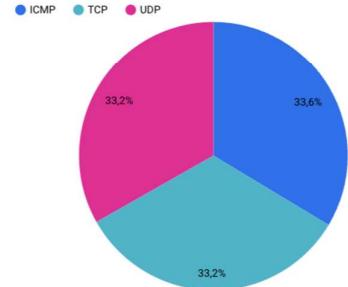
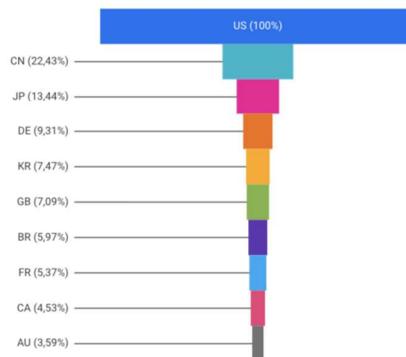
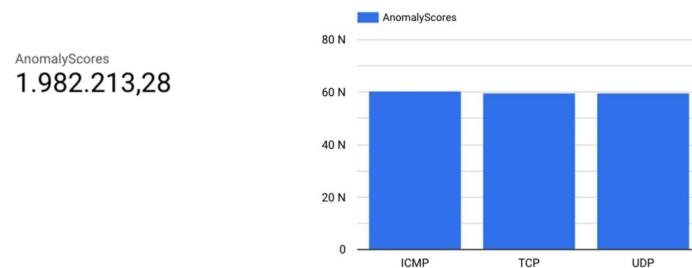


Sơ đồ hình tháp đồ thể hiện các quốc gia nguồn gốc của cuộc tấn công ta có thể chọn từng quốc gia để thể hiện tổng điểm Anomaly Scores của các cuộc tấn công theo từng Phương thức.



Bước 4: Tổng hợp lại report

Truy vấn tổng điểm Anomaly Scores của các cuộc tấn công theo từng Phương thức (Protocol) và tháng với phương thức là Malware trong năm 2021



Nhận xét: từ biểu đồ ta có thể thấy:

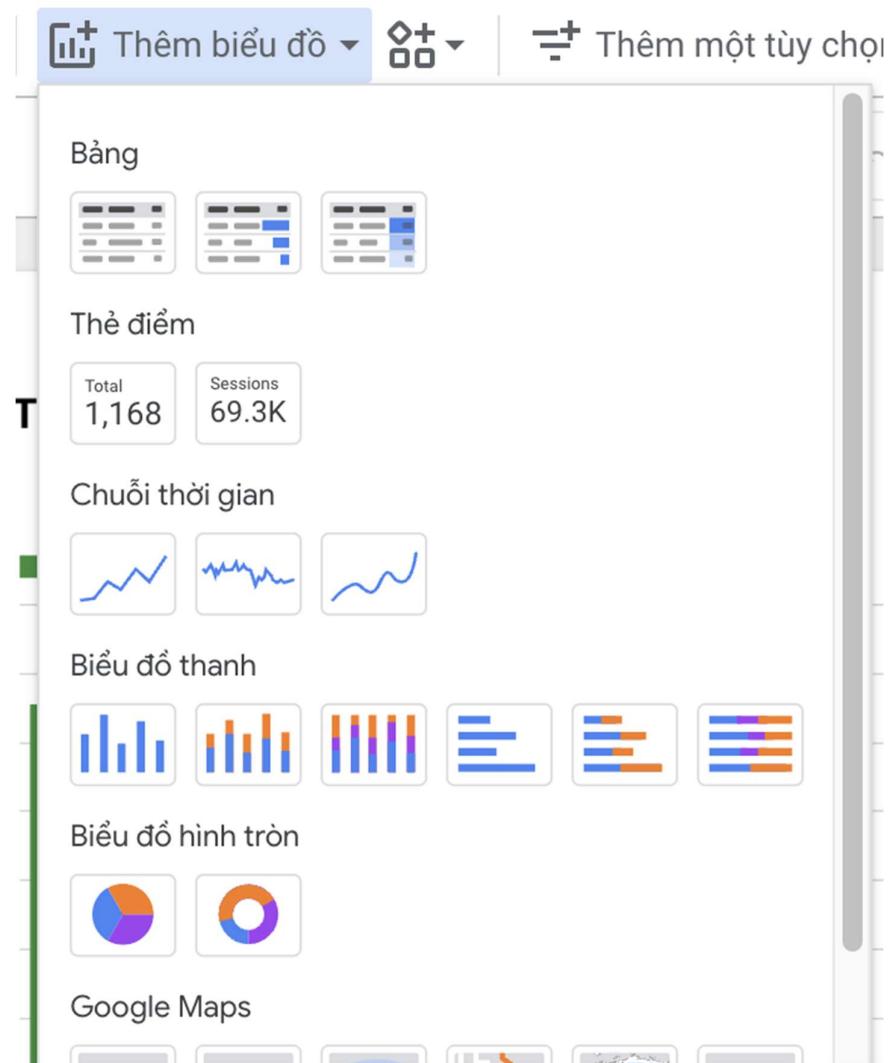
- Phương thức ICMP có khả nhận được nhận biết cao nhất

- Các phương thức có tổng điểm tương đương nhau
- Riêng nước mỹ có tổng điểm AnomalyScore cao nhất với hơn 600 ngàn điểm.

4.2.2. Tạo report 2

Câu truy vấn: Top 10 Quốc Gia Có Nhiều Cuộc Tấn Công Mạng Nhất trong năm 2022

Bước 1: Chọn vào thêm biểu đồ, chọn vào biểu đồ Bảng và biểu đồ cột.



Bước 2: Chọn vào sơ đồ vừa tạo, tiến hành tạo bộ lọc cho sơ đồ

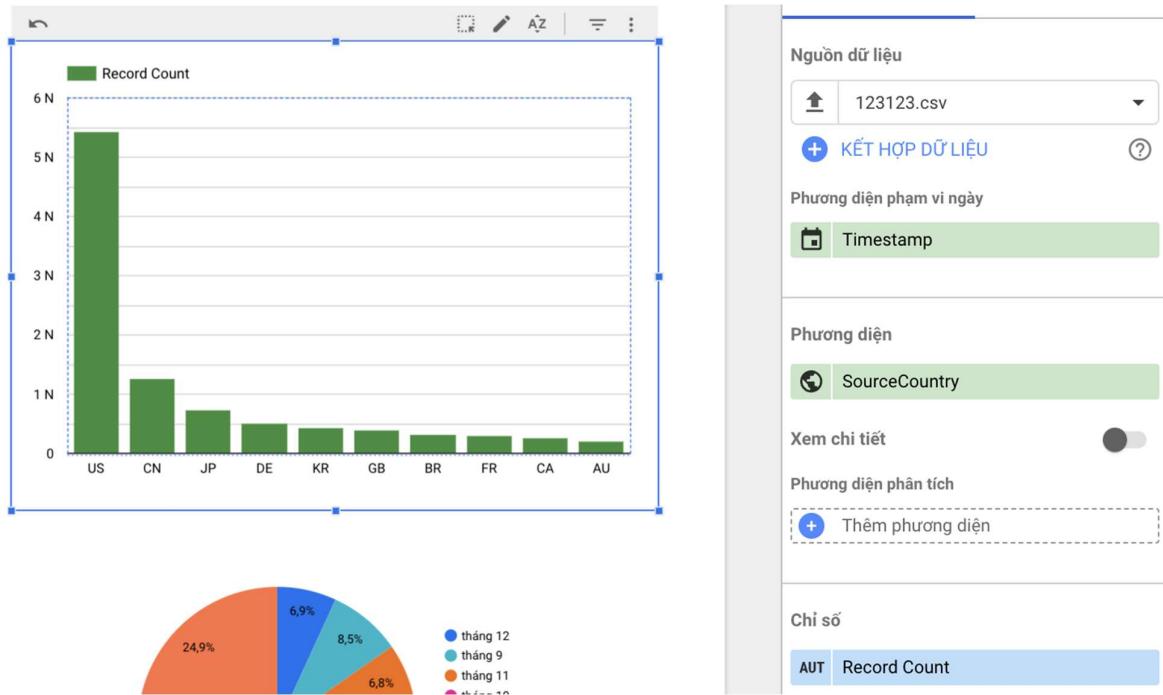
- Tạo bộ lọc mới gồm Tag là “Bao gồm”, Chọn trường là Year, chọn điều kiện là Bằng. Giá trị là “2022”

Tên Hiện giá trị để xuất trong khi nhập

Bao gồm Bảng (=)

VÀ

Bước 3: kéo Source Country từ bảng dữ liệu vào phần phương diện, kéo Record Count vào mục chỉ số.



Bước 4: Điều chỉnh mục sắp xếp giảm dần và điều chỉnh số hàng trên 1 trang là 10.

Số hàng

Phân trang

Hàng trên cùng

Trên mỗi trang

10



Hàng tóm tắt

Hiển thị hàng tóm tắt

Sắp xếp

AUT Record Count

Giảm dần

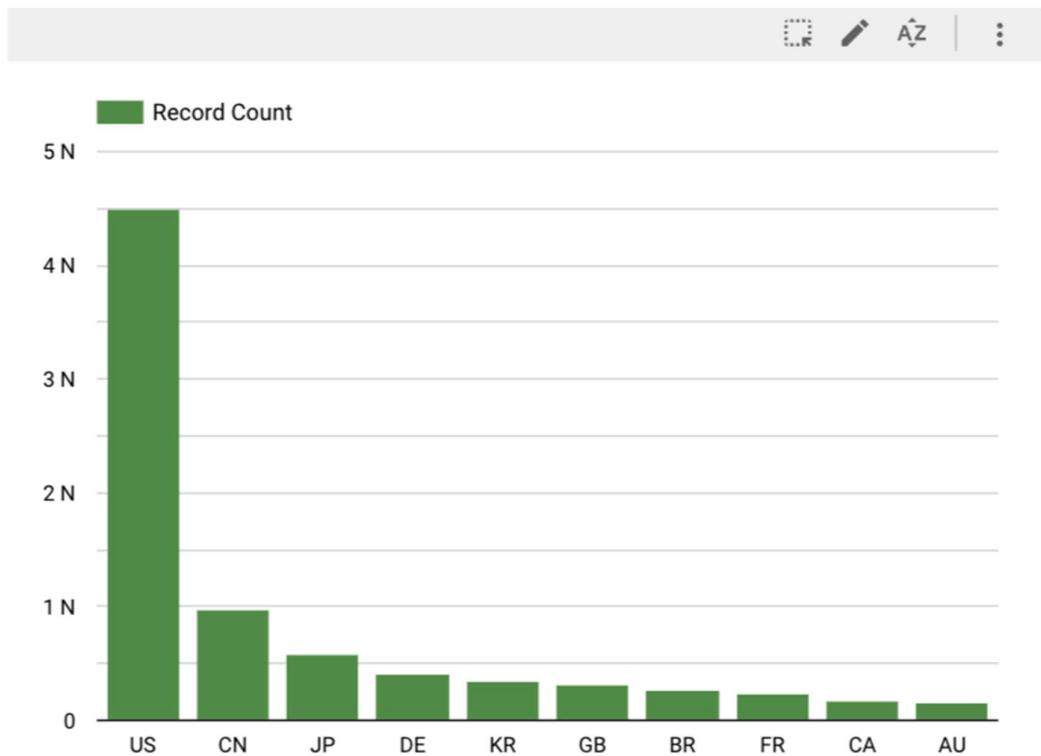
Tăng dần

- Kết quả với biểu đồ bảng:

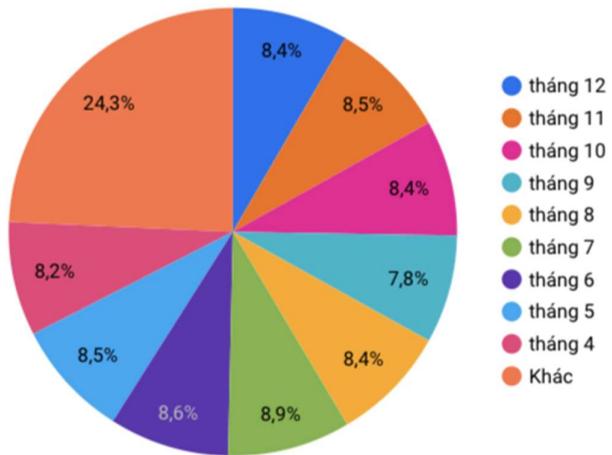
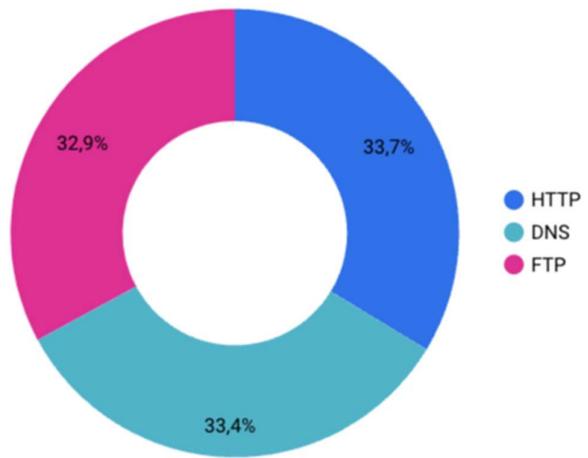
	SourceCountry	Record Count ▾
1.	US	4.496
2.	CN	970
3.	JP	574
4.	DE	403
5.	KR	340
6.	GB	317
7.	BR	264
8.	FR	242
9.	CA	178
10.	AU	151

1 - 10 / 145 < >

- Kết quả với biểu đồ hình cột



Bước 5: Sơ đồ hình tròn thể hiện các phương thức và các tháng trong năm ta có thể chọn filter theo từng phương thức và từng tháng trong năm.



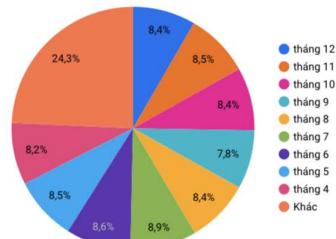
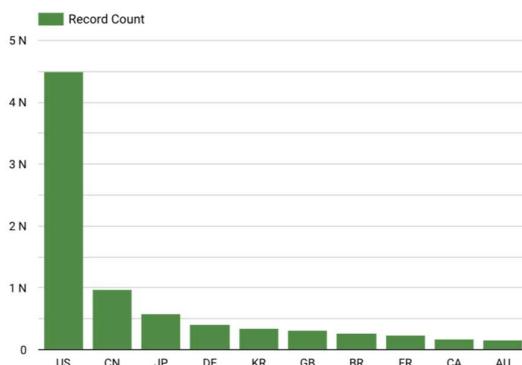
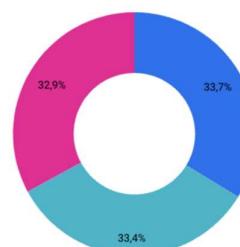
Bước 6: Tổng hợp lại report

Record Count
10.634

Top 10 Quốc Gia Có Nhiều Cuộc Tấn Công Mạng Nhất trong năm 2022

SourceCountry	Record Count
1. US	4.496
2. CN	970
3. JP	574
4. DE	403
5. KR	340
6. GB	317
7. BR	264
8. FR	242
9. CA	178
10. AU	151

1 - 10 / 145 < >

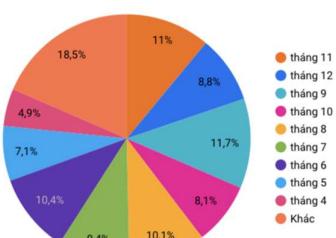
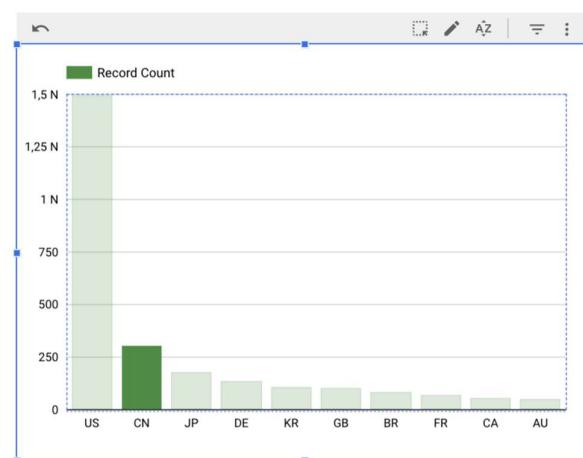
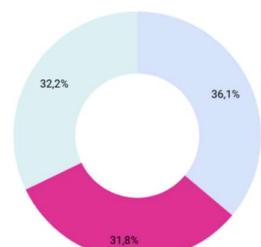


Record Count
308

Top 10 Quốc Gia Có Nhiều Cuộc Tấn Công Mạng Nhất trong năm 2022

SourceCountry	Record Count
1. CN	308

1 - 1 / 1 < >



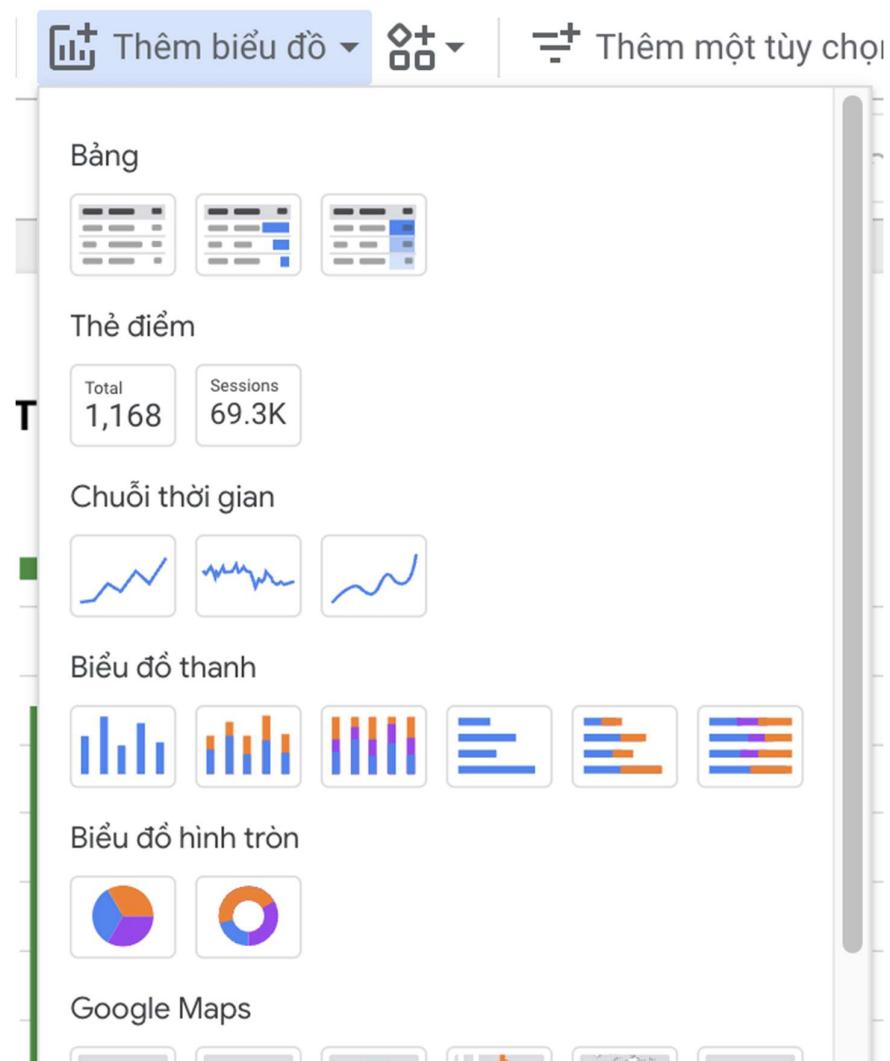
Nhận xét: từ biểu đồ ta có thể thấy:

- Mỹ là quốc gia có nhiều cuộc tấn công nhất lên đến 4496 cuộc tấn công trong năm 2022.
- Tháng 11 là tháng có cao điểm của các cuộc tấn công của các nước trên thế giới
- Tại Nhật Bản, thì phương thức HTTP được sử dụng phổ biến để tấn công

4.2.3. Tạo report 3

Câu truy vấn: Tổng dung lượng gói tin của từng hành động khi gói tin được gửi trong từng tháng năm 2021

Bước 1: Chọn biểu đồ Chart trong mục Visualizations.

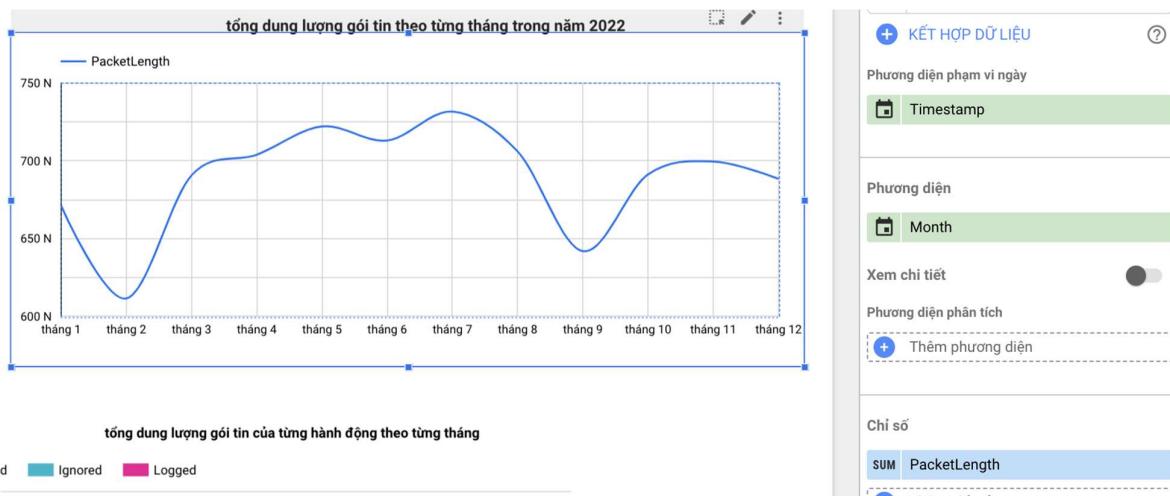


Bước 2: Chọn vào sơ đồ vừa tạo, tiến hành tạo bộ lọc cho sơ đồ

- Tạo bộ lọc mới gồm Tag là “Bao gồm”, Chọn trường là Year, chọn điều kiện là Bằng. Giá trị là “2021”

The screenshot shows the 'Chỉnh sửa bộ lọc' (Edit Filter) dialog box. At the top, there's a search bar with 'Bao gồm' selected. Below it, there are two dropdown menus: one for 'Year' with '2021' selected, and another for 'Condition' with 'Bằng (=)' selected. To the right of these, there's a 'HOẶC' (OR) button and a 'VÀ' (AND) button below it.

Bước 3: kéo Month từ bảng dữ liệu vào phần phương diện, kéo Packet Length vào mục chỉ số.

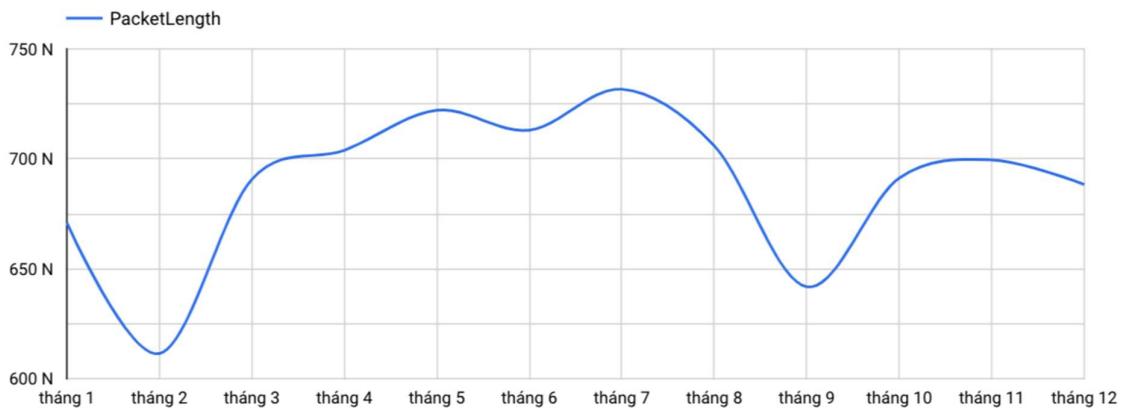


- Tổng dung lượng gói tính của toàn bộ tháng:

PacketLength
8.271.009

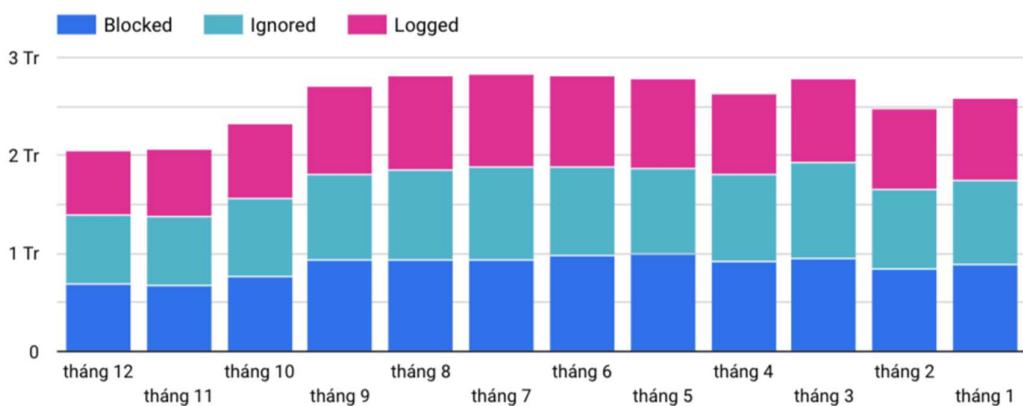
- Kết quả với biểu đồ đường:

tổng dung lượng gói tin theo từng tháng trong năm 2022



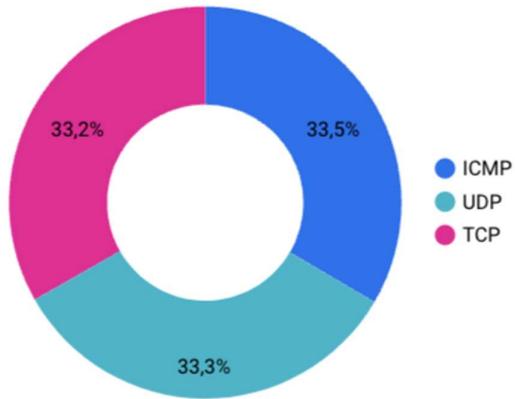
- Kết quả với biểu đồ cột chồng :

tổng dung lượng gói tin của từng hành động theo từng tháng



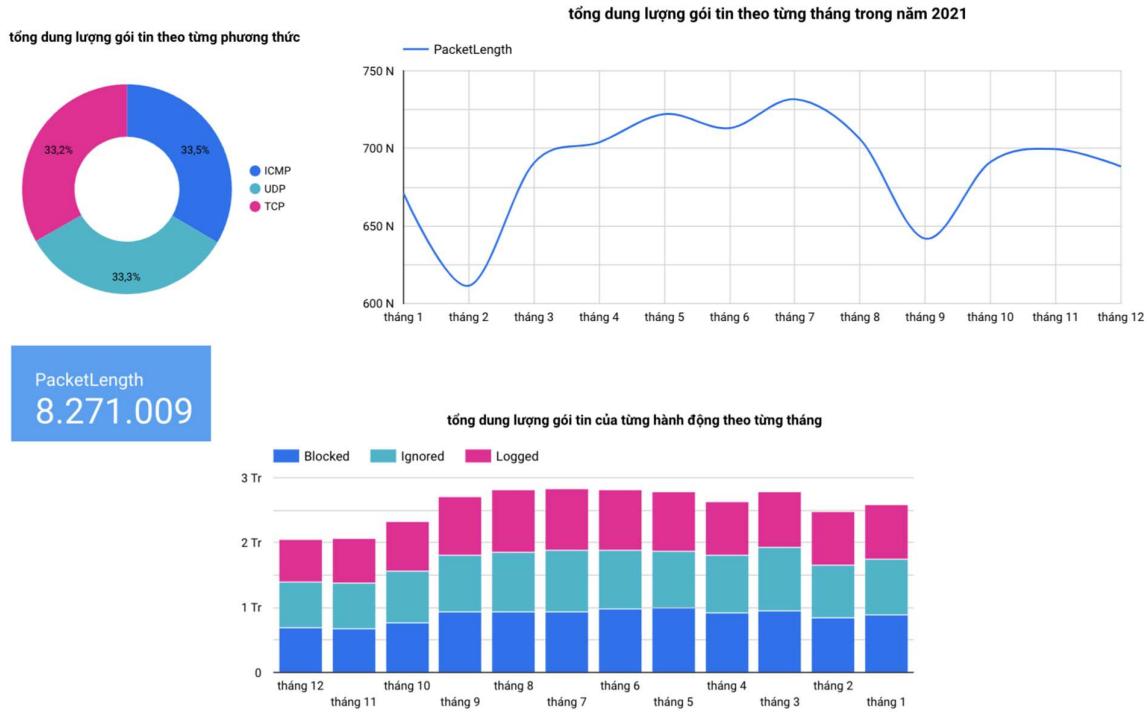
Sơ đồ hình tròn thể hiện các phương thức và các tháng trong năm ta có thể chọn filter theo từng phương thức và từng tháng trong năm.

tổng dung lượng gói tin theo từng phương thức



Bước 4: Tổng hợp lại report

TỔNG DUNG LƯỢNG GÓI TIN ĐƯỢC GỬI TỪNG THÁNG TRONG NĂM 2021



Nhận xét: từ biểu đồ ta có thể thấy:

- Tổng dung lượng gói tin cao điểm được gửi vào mùa hè từ tháng 4 đến tháng 8.

- Tổng dung lượng các gói tin của các gói bị Ignore của từng tháng cao hơn Blocked, và Logged
- Tổng dung lượng gói tin ICMP chiếm tỉ lệ cao nhất.

DANH MỤC TÀI LIỆU THAM KHẢO