

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HCM
KHOA CÔNG NGHỆ THÔNG TIN

**



HCMUTE

KHO DỮ LIỆU – DAWH_01
XÂY DỰNG KHO DỮ LIỆU VỀ NGÀNH HÀNG KHÔNG

Giáo viên hướng dẫn: thầy Nguyễn Văn Thành

Thành viên nhóm 2:

Đỗ Phi Long 20110670

Nguyễn Huỳnh Thắng 20110726

Nguyễn Khoa Quang Thắng 20133090

Hoàng Uyên 20133008

Tp. Hồ Chí Minh, tháng 5 năm 2023

Lời cảm ơn

Đầu tiên, nhóm chúng em xin gửi lời cảm ơn đến Thầy – Th.S Nguyễn Văn Thành – Giảng viên phụ trách môn Kho Dữ Liệu – Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh vì những đóng góp của thầy trong thời gian qua. Thầy đã cung cấp những kiến thức, hướng dẫn, đóng góp, xây dựng trong quá trình học để nhóm em có thể hoàn thành đồ án môn học một cách tốt nhất.

Kho dữ liệu là một trong những kiến thức đang ngày càng phổ biến cùng với OLAP, nhóm chúng em đã lên ý tưởng thực hiện xây dựng một kho dữ liệu cùng với các quy trình nghiệp vụ. Nhóm chúng em quyết định xây dựng đồ án “Xây dựng và khai thác kho dữ liệu về lĩnh vực hàng không”. Trong quá trình thực hiện đồ án, vận dụng những kiến thức học được từ lý thuyết đến thực hành trong quá trình học tập, cùng với sự tìm hiểu các kiến thức, công cụ mới, nhóm chúng em đã cố gắng hoàn thiện đồ án một cách tốt nhất.

Nếu có những chỗ chưa hoàn thiện, nhóm chúng em mong nhận được sự đóng góp từ thầy nhằm cải thiện và coi đó là một kinh nghiệm để có thể hoàn thành tốt hơn.

Nhóm chúng em một lần nữa xin chân thành cảm ơn thầy!

Lời Nhận Xét Của Giảng Viên

MỤC LỤC

CHƯƠNG 1: LÝ DO CHỌN ĐỀ TÀI VÀ GIỚI THIỆU TỔNG QUAN DATASET	1
1.1 Lý do hình thành đồ án	1
1.2 Giới thiệu tổng quan dataset	2
CHƯƠNG 2: THIẾT KẾ KHO DỮ LIỆU DÙNG KỸ THUẬT MÔ HÌNH HÓA DỮ LIỆU THEO CHIỀU (DIMENSIONAL MODELING).....	10
2.1 Thiết kế kho dữ liệu mức cao (High Level Design)	10
2.2 Thiết kế kho dữ liệu mức chi tiết (Detailed Level Design)	11
2.3 Các lược đồ mô tả kho dữ liệu.....	22
CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)	25
3.1 Quá trình tạo mới project SSIS	25
3.2 Quá trình đổ dữ liệu vào database Stage	27
Nhấn chuột phải chọn New Connection:	27
3.3 Quá trình Load dữ liệu vào database DW	61
CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)	72
4.1. Danh sách truy vấn	72
4.2. Xây dựng mô hình	72
4.3. Xây dựng khối – Cube	78
4.4. Cấu hình Dimensions	82
4.5. Thực hiện truy vấn	87
CHƯƠNG 5: REPORT	100
5.1 Báo cáo 1 – Tổng booking và giá bán theo quý năm 2015	100
5.2 Báo cáo 2 – Số lượng booking theo nơi đến Arrival	102
5.3 Báo cáo 3 – Số lượng booking theo nơi đi quốc gia.....	103
5.4 Báo cáo 4 – Số lượng booking theo các hãng hàng không.....	104

5.5 Báo cáo 5 – Tổng số lượng chuyến bay của các hãng hàng không	105
5.6 Báo cáo 6 – Tổng chuyến bay toàn quốc theo từng quý năm 2015.....	106
5.7 Báo cáo 7 – Tổng và trung bình lương nhân viên hàng không theo từng quốc gia.....	107
5.8 Báo cáo 8 – Lương nhân viên tại các phòng ban.....	108
CHƯƠNG 6: KẾT LUẬN	110
6.1 Kết quả đạt được	110
6.2 Hạn chế	110
6.3 Phân công nhiệm vụ.....	111
TÀI LIỆU THAM KHẢO	112

CHƯƠNG 1: LÝ DO CHỌN ĐỀ TÀI VÀ GIỚI THIỆU TỔNG QUAN DATASET

1.1 Lý do hình thành đồ án

1.1.1 Vấn đề đặt ra

Ngành hàng không là một trong những ngành trọng điểm đối với thế giới cũng như Việt Nam. Trải qua lịch sử hơn 100 năm đầy biến động, ngành hàng không đã có những biến động cả về công nghệ lẫn dịch vụ. [1]

Từ năm 1994 đến năm 2002 nổi lên với sự phát triển của mô hình hàng không giá rẻ, có ý nghĩa quan trọng với ngành hàng không ngày nay. Các công ty con và trung tâm phân phối được mở rộng toàn thế giới cùng với sự phát triển của thương mại điện tử cho các hệ thống bán vé máy bay trực tuyến.

Tháng 12/2019, tâm dịch Covid-19 nhanh chóng lan rộng ra toàn thế giới khiến tất cả các nước phải áp dụng các biện pháp hạn chế di chuyển, điều này đặc biệt nghiêm trọng cho ngành hàng không thế giới. American Airlines cắt giảm 75% số chuyến bay quốc tế, Korean Air giảm 80% lịch trình,... tóm lại, hầu như các chuyến bay trên thế giới đều bị cắt giảm đáng kể.

Trong 2 năm vừa qua, ngành hàng không quốc tế cũng như Việt Nam đang có dấu hiệu phục hồi trở lại. Các kiến thức về ngành hàng không rất đa dạng như thông tin về các hãng hàng không trên toàn thế giới, các sân bay, chuyến bay,... và phải kể đến quy trình đặt vé máy bay hay còn gọi là booking. Những data kể trên rất phù hợp để xây dựng một kho dữ liệu về lịch trình chuyến bay cũng như booking.

Vì vậy, nhóm chúng em quyết định xây dựng một kho dữ liệu về các chuyến bay và đặt vé máy bay của ngành hàng không thế giới.

1.1.2 Giải pháp

Dựa trên nhu cầu phân tích, thống kê và khai thác dữ liệu về ngành hàng không thế giới. Giải pháp được đưa ra là xây dựng một kho dữ liệu nhằm mục đích khai thác, phân tích và tạo báo cáo tổng về các chuyến bay và booking.

1.1.3 Mục tiêu và ý nghĩa dự án

Đầu tiên là xây dựng một kho dữ liệu và OLAP Data Warehouse hay kho dữ liệu lưu trữ data từ nhiều nguồn, giúp tăng cường hiệu suất báo cáo và phân tích. Dựa vào điều đó cũng như các kiến thức đã học được trong quá trình học tập, nhóm chúng em quyết định xây dựng một kho dữ liệu OLAP để thấy được hiệu quả của nó.

Thứ hai là khai thác và phân tích data về hàng không trong thực tế. Khai thác các thông tin của hàng không quốc tế nhằm có cái nhìn chung về kiến thức cũng như quy trình vận hành cơ bản của một hệ thống đặt vé máy bay.

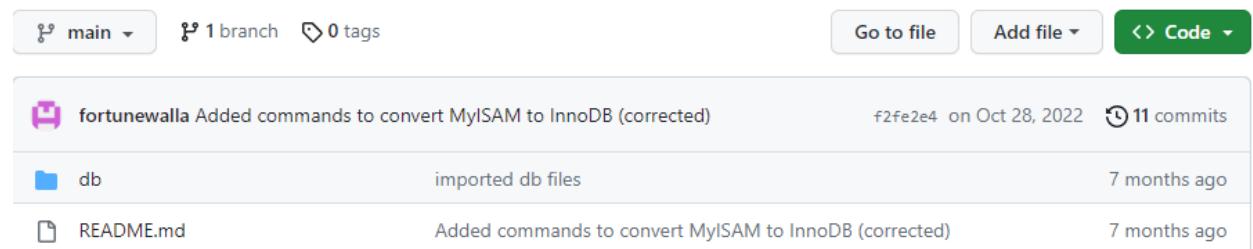
1.2 Giới thiệu tổng quan dataset

1.2.1 Nguồn dữ liệu sử dụng

Nguồn dataset được lấy từ github.com/fortunewalla/airportdb

1.2.2 Tiền xử lý dữ liệu

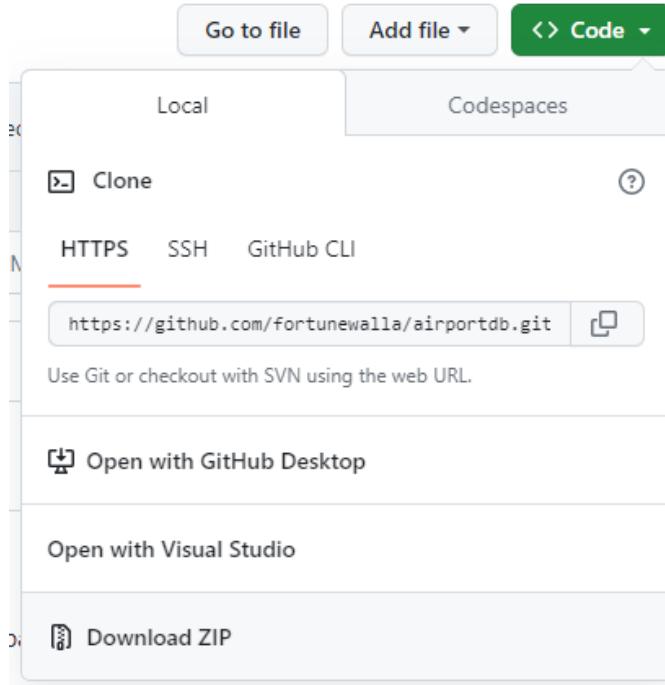
Vào đường dẫn github.com/fortunewalla/airportdb



The screenshot shows a GitHub repository page for 'fortunewalla/airportdb'. At the top, there are buttons for 'main' (branch), '1 branch', '0 tags', 'Go to file', 'Add file', and a green 'Code' button. Below this is a commit history table:

Author	Message	Date	Commits
fortunewalla	Added commands to convert MyISAM to InnoDB (corrected)	f2fe2e4 on Oct 28, 2022	11 commits
db	imported db files	7 months ago	
README.md	Added commands to convert MyISAM to InnoDB (corrected)	7 months ago	

Nhấn vào **Code** rồi **Download ZIP** để tải toàn bộ tài liệu về dataset



Vào SQL Server, tạo New Query, lần lượt tạo database và các table

```
create database FlightDB
use FlightDB
CREATE TABLE passenger (
    passenger_id int NOT NULL,
    passportno char(9) NOT NULL UNIQUE,
    firstname varchar(100) NOT NULL,
    lastname varchar(100) NOT NULL,
    PRIMARY KEY (passenger_id)
)
CREATE TABLE passengerdetails (
    passenger_id int NOT NULL REFERENCES passenger(passenger_id),
    birthdate date NOT NULL,
    sex char(1) DEFAULT NULL,
    street varchar(100) NOT NULL,
    city varchar(100) NOT NULL,
    zip smallint NOT NULL,
    country varchar(100) NOT NULL,
    emailaddress varchar(120) DEFAULT NULL,
    telephoneno varchar(30) DEFAULT NULL,
    PRIMARY KEY (passenger_id)
)
```

```
CREATE TABLE airport (
    airport_id smallint NOT NULL,
    iata char(3) DEFAULT NULL,
    icao char(4) NOT NULL UNIQUE,
    [name] varchar(50) NOT NULL,
    PRIMARY KEY (airport_id),
)
CREATE TABLE airport_geo (
    airport_id smallint NOT NULL REFERENCES airport(airport_id),
    [name] varchar(50) NOT NULL,
    city varchar(50) DEFAULT NULL,
    country varchar(50) DEFAULT NULL,
    latitude decimal(11,8) NOT NULL,
    longitude decimal(11,8) NOT NULL,
    geolocation text NULL,
)
CREATE TABLE airline (
    airline_id smallint NOT NULL,
    iata char(2) NOT NULL UNIQUE,
    airlinename varchar(30) DEFAULT NULL,
    base_airport smallint NOT NULL,
    PRIMARY KEY (airline_id),
    CONSTRAINT fluglinie_ibfk_1 FOREIGN KEY (base_airport)
        REFERENCES airport (airport_id)
)
```

```

CREATE TABLE airplane_type (
    [type_id] int NOT NULL,
    identifier varchar(50) DEFAULT NULL,
    [description] text,
    PRIMARY KEY ([type_id]),
)

CREATE TABLE airplane (
    airplane_id int NOT NULL,
    capacity int NOT NULL,
    [type_id] int NOT NULL REFERENCES airplane_type([type_id]),
    airline_id int NOT NULL,
    PRIMARY KEY (airplane_id)
)

CREATE TABLE flight (
    flight_id int NOT NULL,
    flightno char(8) NOT NULL,
    [from] smallint NOT NULL REFERENCES airport(airport_id),
    [to] smallint NOT NULL REFERENCES airport(airport_id),
    departure datetime NOT NULL,
    arrival datetime NOT NULL,
    airline_id smallint NOT NULL REFERENCES airline(airline_id),
    airplane_id int NOT NULL REFERENCES airplane(airplane_id),
    PRIMARY KEY (flight_id)
)

CREATE TABLE booking (
    booking_id int NOT NULL,
    flight_id int NOT NULL REFERENCES flight(flight_id),
    seat char(4) DEFAULT NULL,
    passenger_id int NOT NULL REFERENCES passenger(passenger_id),
    price decimal(10,2) NOT NULL,
    PRIMARY KEY (booking_id),
    CONSTRAINT sitzplan_unq UNIQUE (flight_id,seat)
)

CREATE TABLE flightschedule (
    flightno char(8) NOT NULL,
    [from] smallint NOT NULL,
    [to] smallint NOT NULL,
    departure time NOT NULL,
    arrival time NOT NULL,
    airline_id smallint NOT NULL,
    monday bit DEFAULT 0,
    tuesday bit DEFAULT 0,
    wednesday bit DEFAULT 0,
    thursday bit DEFAULT 0,
    friday bit DEFAULT 0,
    saturday bit DEFAULT 0,
    sunday bit DEFAULT 0,
    PRIMARY KEY (flightno)
)

```

```

CREATE TABLE employee (
    employee_id int NOT NULL,
    firstname varchar(100) NOT NULL,
    lastname varchar(100) NOT NULL,
    birthdate date NOT NULL,
    sex char(1) DEFAULT NULL,
    street varchar(100) NOT NULL,
    city varchar(100) NOT NULL,
    zip smallint NOT NULL,
    country varchar(100) NOT NULL,
    emailaddress varchar(120) DEFAULT NULL,
    telephoneno varchar(30) DEFAULT NULL,
    salary decimal(8,2) DEFAULT NULL,
    department varchar(20) DEFAULT NULL,
    username varchar(20) DEFAULT NULL UNIQUE,
    [password] char(32) DEFAULT NULL,
    PRIMARY KEY (employee_id)
)

```

Mở file ffdb_data.sql

ffdb	11/4/2023 9:12 PM	File folder	
ffdb_data.sql	7/10/2022 11:35 AM	SQL Source File	1,427,440 KB
ffdb_data.sql.gz	11/4/2023 9:12 PM	WinRAR archive	99,056 KB
ffdb_schema.sql	11/4/2023 9:12 PM	SQL Source File	17 KB

Copy những câu lệnh insert tương ứng với các table

```

LOCK TABLES `airline` WRITE;
/*!40000 ALTER TABLE `airline` DISABLE KEYS */;
INSERT INTO `airline` (`airline_id`, `iata`, `airlinename`, `base_airport`) VALUES (1,'AF','Afghanistan Airlines',850);
INSERT INTO `airline` (`airline_id`, `iata`, `airlinename`, `base_airport`) VALUES (2,'AL','Albania Airlines',10274);
INSERT INTO `airline` (`airline_id`, `iata`, `airlinename`, `base_airport`) VALUES (3,'AM','American Samoa Airlines',9210);
INSERT INTO `airline` (`airline_id`, `iata`, `airlinename`, `base_airport`) VALUES (4,'AN','Angola Airlines',190);
INSERT INTO `airline` (`airline_id`, `iata`, `airlinename`, `base_airport`) VALUES (5,'AR','Argentina Airlines',285);

```

1.2.3 Mô tả chi tiết dữ liệu

1.2.3.1 Tổng quát

Dataset có tổng cộng 12 tables với hơn 450.000 dòng mô tả các hoạt động của các chuyến bay trên toàn thế giới vào năm 2015, mô tả hoạt động đặt vé máy bay cùng với một tập khách hàng, mô tả các sân bay, hàng hàng không, máy bay trên toàn thế giới, ngoài ra dataset còn thống kê các nhân viên làm việc cho các phòng ban khác nhau cùng với mức lương của họ.

1.2.3.2 Mô tả các bảng dataset

Tên bảng	Mô tả
airport	Mô tả các sân bay theo Hiệp hội Vận tải Hàng không Quốc tế IATA và theo Tổ chức Hàng không Dân dụng Quốc tế ICAO
airport_geo	Mô tả chi tiết thành phố, đất nước, vị trí các sân bay
airline	Mô tả các hãng hàng không theo Hiệp hội Vận tải Hàng không Quốc tế IATA
airplane	Mô tả các máy bay thuộc các hãng hàng không
airplane_type	Mô tả các kiểu máy bay
flight	Mô tả hoạt động các chuyến bay bao gồm số hiệu chuyến bay, nơi đi, nơi đến, giờ cất cánh, hạ cánh,...
flight_schedule	Mô tả lịch trình bay của các chuyến bay
booking	Mô tả việc đặt vé máy bay của các khách hàng
passenger	Mô tả khách hàng
passenger_detail	Mô tả chi tiết thông tin khách hàng
employee	Mô tả thông tin nhân viên

1.2.4 Các công cụ được sử dụng

1.2.4.1 Visual Studio 2022

Visual studio là một phần mềm hỗ trợ đắc lực hỗ trợ công việc lập trình website. Công cụ này được tạo lên và thuộc quyền sở hữu của ông lớn công nghệ Microsoft. Năm 1997, phần mềm lập trình nay có tên mã Project Boston. Nhưng sau đó, Microsoft đã kết hợp các công cụ phát triển, đóng gói thành sản phẩm duy nhất.

Phần mềm lập trình Visual Studio của Microsoft hỗ trợ sử dụng trên nhiều nền tảng khác nhau. Không giống như các trình viết code khác, Visual Studio sử dụng được trên cả Windows, Linux và Mac Systems. Điều này cực kỳ tiện lợi cho lập trình viên trong quá trình ứng dụng.

Không chỉ hỗ trợ đa nền tảng, Visual Studio cũng cho phép sử dụng nhiều ngôn ngữ lập trình khác nhau từ C#, F#, C/C++, HTML, CSS, Visual Basic,

JavaScript,... Bởi vậy, Visual Studio có thể dễ dàng phát hiện và thông báo cho bạn khi các chương trình có lỗi.

Mặc dù Visual Studio có hệ thống các ngôn ngữ hỗ trợ lập trình khá đa dạng. Nhưng nếu lập trình viên muốn sử dụng một ngôn ngữ khác, bạn có thể dễ dàng tải xuống các tiện ích mở rộng. Tính năng hấp dẫn này được hoạt động như một phần chương trình độc lập nên không lo làm giảm hiệu năng của phần mềm.

Cho đến nay, nhà phát triển Microsoft vẫn đang không ngừng cải tiến và ứng dụng các công nghệ mới cho Visual Studio. Bởi vậy, đây là “cánh tay đắc lực” hỗ trợ cho các developers, data administrators, testers và bất kỳ ai có nhu cầu tiếp cận, tìm hiểu về lập trình đều có thể sử dụng Visual Studio.

1.2.4.2 Microsoft SQL Server 2019 và Management Studio 18

SQL Server là một hệ quản trị cơ sở dữ liệu quan hệ (Relational Database Management System (RDBMS)) sử dụng câu lệnh SQL (Transact-SQL) để trao đổi dữ liệu giữa máy Client và máy cài SQL Server. Một RDBMS bao gồm databases, database engine và các ứng dụng dùng để quản lý dữ liệu và các bộ phận khác nhau trong RDBMS.

Microsoft đã xây dựng SQL Server 2019 dựa trên các tính năng xử lý truy vấn thông minh đã được triển khai trong các phiên bản SQL Server trước đó. Đồng thời, Microsoft cũng đã mở rộng SQL Server sang các phần mới của công cụ cơ sở dữ liệu với những ưu điểm nổi bật hơn.

SQL Server 2019 mang đến các tính năng tuân thủ và bảo mật sáng tạo, hiệu suất hàng đầu trong ngành, tính khả dụng quan trọng của sứ mệnh và phân tích nâng cao cho tất cả khối lượng công việc dữ liệu của bạn, hiện đã có hỗ trợ cho dữ liệu lớn được tích hợp sẵn, giúp người dùng tiết kiệm rất nhiều thời gian khi bắt đầu một dự án.

SSMS là một ứng dụng phần mềm thiết kế bởi Microsoft, ra mắt lần đầu năm 2005. Ứng dụng này cho phép lập trình viên cấu hình, quản lý và quản trị bộ máy cơ sở dữ liệu (database engine) SQL Server.

1.2.4.3 SQL Server Integration Services (SSIS)

SSIS là viết tắt của SQL Server Integration Services – một công cụ tích hợp và quản lý dữ liệu đi kèm với cơ sở dữ liệu Microsoft SQL Server.

SSIS giúp lưu trữ dữ liệu linh hoạt và nhanh chóng hơn với các tính năng như 1 công cụ ETL (trích xuất lượng lớn dữ liệu từ nhiều nguồn dữ liệu trong các định dạng khác nhau, chỉnh sửa và làm sạch dữ liệu sau đó tải chúng vào kho dữ liệu đã được thiết kế).

Một trong những tính năng vượt trội nhất của SSIS là khả năng trích xuất và đồng bộ dữ liệu nhanh chóng và linh hoạt. SSIS cung cấp một nền tảng thuận tiện để di chuyển dữ liệu một cách dễ dàng từ nguồn này sang nguồn khác.

Có một số thành phần cốt lõi tạo nên SSIS GUI (Đồ họa giao diện người dùng) trong cấu trúc SSIS: Control Flow (Luồng điều khiển), Data Flow (Luồng dữ liệu), Event Handlers (Xử lý sự kiện), Package Explorer

SSIS mang một tính linh hoạt cao giống như một công cụ ETL hay một công cụ chuyển đổi dữ liệu. Việc kết nối giao diện người dùng với cấu hình SSIS vô cùng đơn giản. SSIS cung cấp một số tính năng cho phép bạn tạo tài liệu mới khi bạn chuyển đổi dữ liệu ban đầu.

1.2.4.4 SQL Server Analysis Services (SSAS)

SSAS trong là một trong những dịch vụ chính của SQL Server dùng để xây dựng các chiêu và cubes cho kho dữ liệu, trong phiên bản mới này còn hỗ trợ một số thuật toán khai phá dữ liệu điều này cung cấp cho người quản lý có cái nhìn sâu sắc về dữ liệu của họ hơn. SSAS là một phần của nền tảng quản lý doanh nghiệp thông minh (BI), nó không chỉ là một thành phần của SQL Server, nó còn được sử dụng trên .NET Framework và môi trường phát triển Visual Studio.

Dịch vụ SSAS của SQL Server cung cấp cho người dùng các tính năng mạnh mẽ trong phân tích dữ liệu. Đây là công cụ mạnh để xây dựng các hệ thống xử lý giao dịch trực tuyến.

CHƯƠNG 2: THIẾT KẾ KHO DỮ LIỆU DÙNG KỸ THUẬT MÔ HÌNH HÓA DỮ LIỆU THEO CHIỀU (DIMENSIONAL MODELING)

2.1 Thiết kế kho dữ liệu mức cao (High Level Design)

2.1.1 Xây dựng Business Process

Các yêu cầu business đặt ra:

- Báo cáo doanh số đặt vé máy bay (Booking Sales Reporting): theo dõi doanh số bán vé máy bay (booking) dựa trên khách hàng (passenger), chuyến bay (flight) để biết được rằng khách hàng nào đặt vé nhiều nhất, chuyến bay nào được đặt nhiều nhất.
- Lịch trình chuyến bay (Flight Schedule Monitoring): ghi lại hoạt động của các chuyến bay(flight) theo ngày để biết được rằng chuyến bay bay nhiều nhất, hãng hàng không (airline) nào có nhiều chuyến bay nhất, máy bay (airplane) nào bay nhiều nhất, điểm đến (airport) nhiều nhất.
- Phân tích lương nhân viên (Employee Salary Analysis): thống kê lương nhân viên (employee) theo tháng làm việc tại ngành hàng không để biết được rằng nhân viên thuộc phòng ban (department) nào được trả lương cao nhất.

2.1.2 Xác định mức độ chi tiết (Grain) cho các business process

Business Process Name	Fact Table	Fact Grain Type	Granularity
Fact Booking Sales	FactBookingSales	Transaction	One row per booking
Fact Flight Schedule	FactFlightSchedule	Periodic Snapshot	One row per daily flight snapshot
Fact Employee Salary	FactEmployeeSalary	Transaction	One row per employee

- Hoạt động Booking Sales Reporting gắn với Fact Booking Sales ghi lại mỗi dòng thể hiện việc bán một vé máy bay, thuộc loại Transaction.
- Hoạt động Flight Schedule Monitoring gắn với Fact Flight Schedule ghi mỗi dòng cho hoạt động bay của một chuyến bay, thuộc loại Periodic Snapshot.
- Hoạt động Employee Salary Analysis gắn với Fact Employee Salary ghi lại mỗi dòng với mỗi lương của một nhân viên, thuộc loại Transaction.

2.1.3 Xác định các Dimension

Business Process Name	Fact Table	Date	Airport	Airline	Airplane	Flight	Passenger	Employee	Address	Department
Fact Booking Sales	FactBookingSales	X				X	X		X	
Fact Flight Schedule	FactFlightSchedule	X	X	X	X	X				
Fact Employee Salary	FactEmployeeSalary							X	X	X

- Fact Booking Sales liên quan đến bảng Flight, Passenger, phân cấp khách hàng theo bảng Address, bảng Date.
- Fact Flight Schedule liên quan đến bảng Airport, Airline, Airplane, Flight và Date.
- Fact Employee Salary liên quan đến bảng Employee, phân cấp Employee theo Address, Department.

2.1.4 Xác định các Fact Measure

Business Process Name	Fact Table	Fact Grain Type	Granularity	Facts	Date	Airport	Airline	Airplane	Flight	Passenger	Employee	Address	Department
Fact Booking Sales	FactBookingSales	Transaction	One row per booking	Seat, Price	X								
Fact Flight Schedule	FactFlightSchedule	Periodic Snapshot	One row per daily flight snapshot	Departure Date, Arrival Date	X	X	X	X	X				
Fact Employee Salary	FactEmployeeSalary	Transaction	One row per employee	Salary							X	X	X

- Fact Booking Sales ghi lại Seat (ghế ngồi) và Price (giá vé theo USD) của một khách hàng khi đặt vé.
- Fact Flight Schedule ghi lại Departure Date (ngày bay) và Arrival Date (ngày hạ cánh) của một chuyến bay
- Fact Employee Salary ghi lại Salary (lương theo tháng) ứng với mỗi nhân viên

2.2 Thiết kế kho dữ liệu mức chi tiết (Detailed Level Design)

2.2.1 Mô tả các bảng Dimension

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimAirport	AirportId	Natural Key		1
	Iata	Mã sân bay theo Hiệp hội Vận tải Hàng không Quốc tế (International Air Transport Association)		BMV
	Icao	Mã sân bay theo Tổ chức Hàng không Dân dụng Quốc tế (International Civil Aviation Organization)		VVBM
	AirportName	Tên sân bay		BUONMATHUOT
	City	Thành phố sân bay tọa lạc		BUONMATHUOT
	Country	Đất nước sân bay tọa lạc		VIETNAM
	Latitude	Vĩ độ		12.666111
	Longitude	Kinh độ		108.114444
	AirportKey	Surrogate Key		1

Table Name DimAirport

Table Type	Dimension
Display Name	Airport
Database Schema	FlightDB
Table Description	Airports dimension
Comment	comes from airport table in FlightDB
Biz Filter Logic	
Size	one for each airport
Generate Script?	Y

Attributes and Measures

Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type
AirportId	AirportId	Natural key	-1	1	key
Iata	Iata	Mã sân bay theo Hiệp hội Vận tải Hàng không Quốc tế (International Air Transport Association)	None	BMV	1
Icao	Icao	Mã sân bay theo Tổ chức Hàng không Dân dụng Quốc tế (International Civil Aviation Organization)	None	VVBM	1
AirportName	AirportName	Tên sân bay	None	BUONMATHUOT	1
City	City	Thành phố sân bay tọa lạc	None	BUONMATHUOT	1
Country	Country	Đất nước sân bay tọa lạc	None	VIETNAM	1
Latitude	Latitude	Vĩ độ	None	12.666111	1
Longitude	Longitude	Kinh độ	None	108.114444	1
AirportKey	AirportKey	Surrogate primary key	-1	1	key

Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Target		Source			
							Source System	Source Schema	Source Table	Source Field Name	Source	Datatype
int					N		FlightDB	dbo	Airport	AirportId	int	
nvarchar	3				Y		FlightDB	dbo	Airport	Iata	nvarchar	
nvarchar	4				N		FlightDB	dbo	Airport	Icao	nvarchar	
nvarchar	50				N		FlightDB	dbo	Airport	AirportName	nvarchar	
nvarchar	50				Y		FlightDB	dbo	Airport	City	nvarchar	
nvarchar	50				Y		FlightDB	dbo	Airport	Country	nvarchar	
decimal	(11,8)				N		FlightDB	dbo	Airport	Latitude	decimal	
decimal	(11,8)				N		FlightDB	dbo	Airport	Longitude	decimal	
int				PK ID	N		Derived					

- Bảng DimAirport gồm các trường AirportId, Iata, Icao, AirportName, City, Country, Latitude, Longitude, AirportKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimAirline	AirlineId	Natural Key		1
	Iata	Mã sân bay theo Hiệp hội Vận tải Hàng không Quốc tế (International Air Transport Association)		BMV
	AirlineName	Tên hãng hàng không		Vietnam Airlines
	Base Airport	Sân bay cở sở		1232
	AirlineKey	Surrogate Key		1

Table Name	DimAirline
Table Type	Dimension
Display Name	Airline
Database Schema	FlightDB
Table Description	Airlines dimension
Comment	comes from airline table in FlightDB
Biz Filter Logic	
Size	one for each airline
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	
AirlineId	AirlineId	Natural key	-1	1	key	
Iata	Iata	Mã sân bay theo IATA	None	BMV	1	
AirlineName	AirlineName	Tên hàng hàng không	None	Vietnam Airlines	1	
Base Airport	Base Airport	Sân bay cảng sở	None	1232	1	
AirlineKey	AirlineKey	Surrogate primary key	-1	1	key	

Target					Source						
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int					N		FlightDB	dbo	Airline	AirlineId	int
nvarchar	2				N		FlightDB	dbo	Airline	Iata	nvarchar
nvarchar	30				Y		FlightDB	dbo	Airline	AirlineName	nvarchar
int					N		FlightDB	dbo	Airline	Base Airport	int
			PK ID		N		Derived				

- Bảng DimAirline gồm các trường AirlineId, Iata, AirlineName, BaseAirport, AirlineKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimAirplane	AirplaneId	Natural Key		1
	Capacity	Mã định danh máy bay		420
	TypeId	Kiểu máy bay		60
	Airline	Hãng hàng không sở hữu máy bay		35
	Identifier	Tên nhận dạng kiểu máy bay		Airbus-A320-Familie
	Description	Mô tả kiểu máy bay		Als Airbus-A320-Familie werden die
	AirplaneKey	Surrogate Key		1

Table Name	DimAirplane
Table Type	Dimension
Display Name	Airplane
Database Schema	FlightDB
Table Description	Airplane dimension
Comment	comes from airplane table in FlightDB
Biz Filter Logic	
Size	one for each airplane
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	
AirplaneId	AirplaneId	Natural Key	-1	1		key
Capacity	Capacity	Mã định danh máy bay	None	420		1
TypeId	TypeId	Kiểu máy bay	None	60		1
Airline	Airline	Hãng hàng không sở hữu máy bay	None	35		1
Identifier	Identifier	Tên nhân dạng kiểu máy bay	None	Airbus-A320-Familie		1
Description	Description	Mô tả kiểu máy bay	None	Als Airbus-A320-Familie		1
AirplaneKey	AirplaneKey	Surrogate primary key	-1	1		key

Target					Source						
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int					N		FlightDB	dbo	Airplane	AirplaneId	int
int					N		FlightDB	dbo	Airplane	Capacity	int
int					N		FlightDB	dbo	Airplane	TypeId	int
int					N		FlightDB	dbo	Airplane	Airline	int
nvarchar	50				Y		FlightDB	dbo	Airplane	Identifier	nvarchar
text					Y		FlightDB	dbo	Airplane	Description	text
int				PK ID	N		Derived	dbo	Airplane		

- Bảng DimAirplane gồm các trường AirplaneId, Capacity, TypeId, Airline, Identifier, Description, AirplaneKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimFlight	Fight No	Số hiệu chuyến bay		AF1377
	From	Nơi đi		A CORUNA
	To	Nơi đến		BUONMATHUOT
	Departure	Thời điểm bay		10:15:00 AM
	Arrival	Thời điểm đến		8:46:00 PM
	Airline	Hãng hàng không		Vietnam Airlines
	FlightKey	Surrogate Key		1

Table Name	DimFlight
Table Type	Dimension
Display Name	Flight
Database Schema	FlightDB
Table Description	Flights dimension
Comment	comes from flight table in FlightDB
Biz Filter Logic	
Size	one for each flight
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	
Fight No	Fight No	Số hiệu chuyến bay	None	AF1377		key
From	From	Nơi đi	None	A CORUNA		1
To	To	Nơi đến	None	BUONMATHUOT		1
Departure	Departure	Thời điểm bay	None	10:15:00 AM		1
Arrival	Arrival	Thời điểm đến	None	8:46:00 PM		1
Airline	Airline	Hãng hàng không	None	Vietnam Airlines		1
FlightKey	FlightKey	Surrogate primary key	-1	1		key

Target						Source					
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
nvarchar	8				N		FlightDB	dbo	Flight	Flight No	nvarchar
int					N		FlightDB	dbo	Flight	From	int
int					N		FlightDB	dbo	Flight	To	int
time	7				N		FlightDB	dbo	Flight	Departure	time
time	7				N		FlightDB	dbo	Flight	Arrival	time
int					N		FlightDB	dbo	Flight	Airline	int
int				PK ID	N		Derived				

- Bảng DimFlight gồm các trường FlightNo, From, To, Departure, Arrival, Airplane, FlightKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimDate	Date	Ngày, tháng, năm		5/1/2005
	DayOfWeek	Ngày thứ mấy trong tuần		5
	DayOfMonth	Ngày thứ mấy trong tháng		5
	DayName	Tên ngày		Friday
	WeekOfYear	Tuần thứ mấy trong năm		1
	MonthOfYear	Tháng		1
	MonthName	Tên tháng		January
	Quarter	Quý		1
	Year	Năm dương lịch		1996

Table Name	DimDate
Table Type	Dimension
Display Name	Date
Database Schema	northwind
Table Description	Date dimension contains one row for every day.
Comment	The Date dimension is derived; it is not populated from any transaction system.
Biz Filter Logic	
Size	365/year
Generate Script?	Y

Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules
DateKey	DateKey	Surrogate primary key	-1	20041123 38314			In the form: yyyyymmdd
Date	Date	Full date as a SQL date			Day		
DayOfWeek	DayOfWeek	Number of the day of week; Sunday = 1	0	1..7	1	Day	
DayName	DayName	Day name of week, eg Monday	Unk date	Sunday	1	Day	
DayOfMonth	DayOfMonth	Number of the day in the month	0	1..31	1	Day	
DayOfYear	DayOfYear	Number of the day in the year	0	1..365	1	Day	
WeekOfYear	WeekOfYear	Week of year, 1..53	0	1..52 or 53	1	Calendar	
MonthName	MonthName	Month name, eg January	Unk month	November	1	Calendar	
MonthOfYear	MonthOfYear	Month of year, 1..12	0	1, 2, ..., 12	1	Calendar	
Quarter	Quarter	Calendar quarter, 1..4	0	1, 2, 3, 4	1	Calendar	
QuarterName	QuarterName	Quarter name eg. First	Unk qtr	November	1	Calendar	
Year	Year	Calendar year, eg 2010	0	2004	1	Calendar	
IsWeekday	IsWeekday	Is today a weekday	0	1, 0	1	Day	

Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source				
							Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int				PK	N		Derived				
date					Y		Derived				
tinyint					N		Derived				
nchar	10				N		Derived				
tinyint					N		Derived				
smallint					N		Derived				
tinyint					N		Derived				
nchar	10				N		Derived				
tinyint					N		Derived				
tinyint					N		Derived				
nchar	10				N		Derived				
smallint					N		Derived				
bit					N	0	Derived				

- Bảng DimDate mô tả một dòng là một ngày

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimPassenger	PassengerId	Natural Key	1	
	PassportNo	Mã passport	P103023	
	First Name	First Name của khách hàng	Buzz	
	Last Name	Last Name của khách hàng	Aldrin	
	Birthdate	Ngày sinh	12/1/1930	
	Sex	Giới tính	Mara.Corday@cvcpaging.com	
	Street	Địa chỉ (đường)	Archenweg 82	
	City	Địa chỉ (thành phố)	Behamberg	
	Zip	Mã bưu chính	4441	
	Country	Địa chỉ (quốc tịch)	Korea	
	EmailAddress	Email	Mara.Corday@cvcpaging.com	
	Telephone No	Số điện thoại	03133 10894292	
	PassengerKey	Surrogate Key	1	

Table Name	DimPassenger
Table Type	Dimension
Display Name	Passenger
Database Schema	FlightDB
Table Description	Passengers dimension
Comment	comes from passenger table in FlightDB
Biz Filter Logic	
Size	one for each passenger
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	
PassengerId	PassengerId	Natural Key	-1	1		key
PassportNo	PassportNo	Mã passport	None	P103023		1
First Name	First Name	First Name của khách hàng	None	Buzz		1
Last Name	Last Name	Last Name của khách hàng	None	Aldrin		1
Birthdate	Birthdate	Ngày sinh	None	12/1/1930		1
Sex	Sex	Giới tính	None	Mara.Corday@cvcpaging.com		1
Street	Street	Địa chỉ (đường)	None	Archenweg 82		1
City	City	Địa chỉ (thành phố)	None	Behamberg		1
Zip	Zip	Mã bưu chính	None	4441		1
Country	Country	Địa chỉ (quốc tịch)	None	Korea		1
EmailAddress	EmailAddress	Email	None	Mara.Corday@cvcpaging.com		1
Telephone No	Telephone No	Số điện thoại	None	03133 10894292		1
PassengerKey	PassengerKey	Surrogate primary key	-1	1		key

Target						Source					
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int					N		FlightDB	dbo	Passenger	PassengerId	int
nvarchar	9				N		FlightDB	dbo	Passenger	PassportNo	nvarchar
nvarchar	100				N		FlightDB	dbo	Passenger	First Name	nvarchar
nvarchar	100				N		FlightDB	dbo	Passenger	Last Name	nvarchar
date					N		FlightDB	dbo	Passenger	Birthdate	date
nvarchar	1				Y		FlightDB	dbo	Passenger	Sex	nvarchar
nvarchar	100				N		FlightDB	dbo	Passenger	Street	nvarchar
nvarchar	100				N		FlightDB	dbo	Passenger	City	nvarchar
smallint					N		FlightDB	dbo	Passenger	Zip	smallint
nvarchar	100				N		FlightDB	dbo	Passenger	Country	nvarchar
nvarchar	120				Y		FlightDB	dbo	Passenger	EmailAddress	nvarchar
nvarchar	30				Y		FlightDB	dbo	Passenger	Telephone No	nvarchar
int				PK ID	N		Derived				

- Bảng DimPassenger gồm các trường PassengerId, PassportNo, FirstName, LastName, Birthdate, Sex, Street, City, Zip, Country, EmailAddress, TelephoneNo, PassengerKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimAddress	Street	Địa chỉ (đường)		Rosenweg 78
	City	Địa chỉ (thành phố)		Unterthiergrub
	Zip	Mã bưu chính		4162
	Country	Địa chỉ (quốc tịch)		Spain
	AddressKey	Surrogate Key		1

Table Name	DimAddress
Table Type	Dimension
Display Name	Address
Database Schema	FlightDB
Table Description	Addresses dimension
Comment	comes from employee table in FlightDB
Biz Filter Logic	
Size	one for each address
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	
Street	Street	Địa chỉ (đường)	-1	Rosenweg 78	key	
City	City	Địa chỉ (thành phố)	None	Unterthiergrub	1	
Zip	Zip	Mã bưu chính	None	4162	1	
Country	Country	Địa chỉ (quốc tịch)	None	Spain	1	
AddressKey	AddressKey	Surrogate primary key	-1	1	key	

Target						Source					
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
nvarchar	100				N		FlightDB	dbo	Employee	Street	nvarchar
nvarchar	100				N		FlightDB	dbo	Employee	City	nvarchar
smallint					N		FlightDB	dbo	Employee	Zip	smallint
nvarchar	100				N		FlightDB	dbo	Employee	Country	nvarchar
int				PK ID	N		Derived				

- Bảng DimAddress gồm các trường Street, City, Zip, Country, AddressKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimDepartment	Department	Vị trí việc làm		Marketing
	DepartmentKey	Surrogate Key		3

Table Name DimDepartment

Table Type	Dimension
Display Name	Department
Database Schema	FlightDB
Table Description	Departments dimension
Comment	comes from employee table in FlightDB
Biz Filter Logic	
Size	one for each department
Generate Script?	Y

Attributes and Measures

Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type
Department	Department	Vị trí việc làm	None	Marketing	1
DepartmentKey	DepartmentKey	Surrogate primary key	-1	3	key

Target					Source						
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
nvarchar	20			PK ID	Y	N	FlightDB	dbo	Airport	Employee	nvarchar
int							Derived				

- Bảng DimDepartment gồm các trường Department, DepartmentKey với mô tả và giá trị ví dụ

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
DimEmployee	EmployeeId	Natural Key		1
	FirstName	First Name của khách hàng		Buzz
	LastName	Last Name của khách hàng		Aldrin
	Birthdate	Ngày sinh		12/1/1930
	Sex	Giới tính		Mara.Corday@cvcpaging.com
	EmailAddress	Email		Mara.Corday@cvcpaging.com
	TelephoneNo	Số điện thoại		03133_10894292
	UserName	Username		Michael1
	Password	Mật khẩu		3e06fa3927cbdf4e9d93ba4541acce86
EmployeeKey	Email		1	

Table Name	DimEmployee
Table Type	Dimension
Display Name	Employee
Database Schema	FlightDB
Table Description	Employees dimension
Comment	comes from employee table in FlightDB
Biz Filter Logic	
Size	one for each employee
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Unknown Member	Example Values		SCD Type
EmployeeId	EmployeeId	Natural Key	-1	1		key
FirstName	FirstName	First Name của khách hàng	None	Buzz		1
LastName	LastName	Last Name của khách hàng	None	Aldrin		1
Birthdate	Birthdate	Ngày sinh	None	12/1/1930		1
Sex	Sex	Giới tính	None	Mara.Corday@cvcpaging.cc	1	1
EmailAddress	EmailAddress	Email	None	Mara.Corday@cvcpaging.cc	1	1
TelephoneNo	TelephoneNo	Số điện thoại	None	03133 10894292		1
UserName	UserName	Username	None	Michael1		1
Password	Password	Mật khẩu	None	3e06fa3927cbdf4e9d93ba	1	1
EmployeeKey	EmployeeKey	Surrogate primary key	-1	1		key

Target					Source						
Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int					N		FlightDB	dbo	Airport	EmployeeId	int
nvarchar	100				N		FlightDB	dbo	Airport	FirstName	nvarchar
nvarchar	100				N		FlightDB	dbo	Airport	LastName	nvarchar
date					N		FlightDB	dbo	Airport	Birthdate	date
nvarchar	1				Y		FlightDB	dbo	Airport	Sex	nvarchar
nvarchar	120				Y		FlightDB	dbo	Airport	EmailAddress	nvarchar
nvarchar	30				Y		FlightDB	dbo	Airport	TelephoneNo	nvarchar
nvarchar	20				Y		FlightDB	dbo	Airport	UserName	nvarchar
nvarchar	32				Y		FlightDB	dbo	Airport	Password	nvarchar
int				PK ID	N		Derived				

- Bảng DimEmployee gồm các trường EmployeeId, FirstName, LastName, BirthDate, Sex, EmailAddress, TelephoneNo, UserName, Password, EmployeeKey với mô tả và giá trị ví dụ

2.2.2 Mô tả các bảng Fact

Dimension / Fact Table	Attribute / Fact Name	Description	Alternate Names	Sample Values
FactBookingSales	Seat	Số ghế ngồi		21A
	Price	Giá tiền booking tính theo USA		1200.12
	Departure Date	Ngày bay		5/1/2005
	Arrival Date	Ngày đến		5/1/2005

Table Name FactBookingSales

Table Type	Fact
Display Name	BookingSales
Database Schema	FlightDB
Table Description	Booking Sales information
Comment	
Biz Filter Logic	
Size	
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Example Values	Display Folder	ETL Rules	
FlightKey	FlightKey	Key to Flight	103	key	Key lookup from DimFlight.FlightKey	
PassengerKey	PassengerKey	Key to Passenger	209	key	Key lookup from DimPassenger.PassengerKey	
BookingId	BookingId	Mã booking	1433230	Exclude from cube		
Seat	Seat	Số ghế ngồi	21A	Exclude from cube		
Price	Price	Giá tiền booking tính theo USA	1200.12	Exclude from cube		
DepartmentDate	DepartmentDate	Key to Date	20150101	key	Key lookup from DimDate.DateKey	
ArrivalDate	ArrivalDate	Key to Date	20150101	key	Key lookup from DimDate.DateKey	

Datatype	Size	Target			Source						
		Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int				FK,PK	DimFlight.FlightKey	N					
int				FK,PK	DimPassenger.PassengerKey	N					
int				FK,PK		N					
nvarchar	4			FK		N	FlightDB	dbo	Booking	Seat	nvarchar
decimal	(10,2)			FK	DimDate.DateKey	N	FlightDB	dbo	Booking	Price	decimal
int				FK	DimDate.DateKey	N					
int				FK	DimDate.DateKey	N					

- FactBookingSales ghi lại Seat, Price, DepartureDate, ArrivalDate cho mỗi booking

Dimension / Fact Table	Attribute /			Alternate Names	Sample Values
	Fact Name	Description			
FactFlightSchedule	Departure Date	Ngày bay			5/1/2005
	Arrival Date	Ngày đến			5/1/2005

Table Name FactFlightSchedule

Table Type	Fact
Display Name	FlightSchedule
Database Schema	FlightDB
Table Description	FlightSchedule information
Comment	
Biz Filter Logic	
Size	
Generate Script?	Y

Attributes and Measures						
Column Name	Display Name	Description	Example Values	Display Folder	ETL Rules	
FlightId	FlightId	Mã Id Flight	100	Exclude from cube		
FlightKey	FlightKey	Key to Flight	103	key	Key lookup from DimFlight.FlightKey	
From	From	Key to Airport	9583	key	Key lookup from DimAirport.AirportKey	
To	To	Key to Airport	1924	key	Key lookup from DimAirport.AirportKey	
AirlineKey	AirlineKey	Key to Airline	49	key	Key lookup from DimAirline.AirlineKey	
AirplaneKey	AirplaneKey	Key to Airplane	2480	key	Key lookup from DimAirplane.AirplaneKey	
DepartmentDate	DepartmentDate	Key to Date	20150101	key	Key lookup from DimDate.DateKey	
ArrivalDate	ArrivalDate	Key to Date	20150101	key	Key lookup from DimDate.DateKey	

Datatype	Size	Target				Source					
		Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int			PK	DimFlight.FlightKey	N		FlightDB	dbo	Flight	FlightId	int
int			FK,PK	DimAirport.AirportKey	N						
int			FK,PK	DimAirport.AirportKey	N						
int			FK,PK	DimAirline.AirlineKey	N						
int			FK,PK	DimAirplane.AirplaneKey	N						
int			FK	DimDate.DateKey	N						
int			FK	DimDate.DateKey	N						

- FactFlightSchedule ghi lại DepartureDate và ArrivalDate cho mỗi chuyến bay

Dimension / Fact Table	Attribute / Fact Name			Description	Alternate Names	Sample Values
	Fact Name	Description	Unit			
FactEmployeeSalary	Salary	Lương theo tháng tính theo USD				1203.19

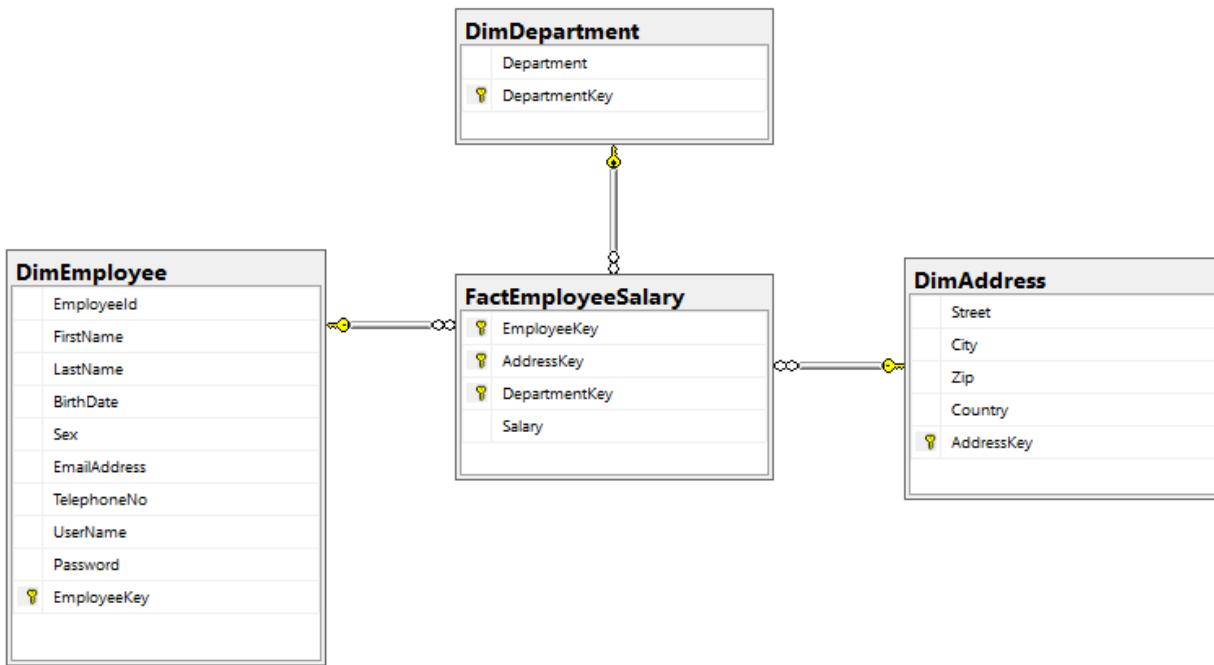
Table Name	FactEmployeeSalary
Table Type	Fact
Display Name	EmployeeSalary
Database Schema	FlightDB
Table Description	Employee Salary information
Comment	
Biz Filter Logic	
Size	
Generate Script?	Y

Column Name	Display Name	Attributes and Measures						ETL Rules
		Description	Example Values	Display Folder				
EmployeeKey	EmployeeKey	Key to Employee	19	key	Key lookup from DimEmployee.EmployeeKey			
AddressKey	AddressKey	Key to Address	584	key	Key lookup from DimAddress.AddressKey			
DepartmentKey	DepartmentKey	Key to Department	3	key	Key lookup from DimDepartment.DepartmentKey			
Salary	Salary	Lương theo tháng USD	1431.93	Exclude from cube				

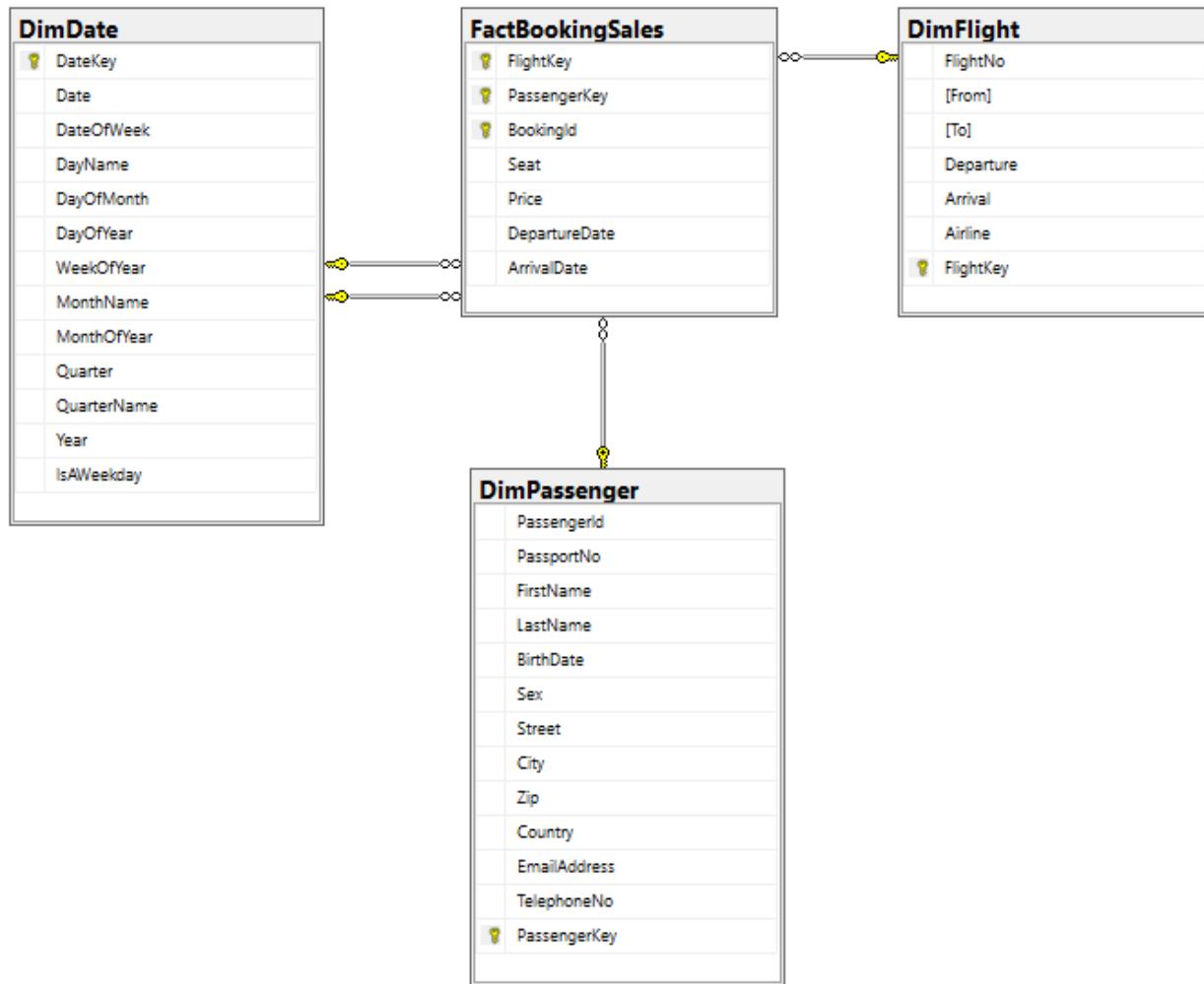
Datatype	Size	Target				Source					
		Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
int			FK,PK	DimEmployee.EmployeeKey	N		FlightDB	dbo	Employee	Salary	decimal
int			FK,PK	DimAddress.AddressKey	N						
int			FK,PK	DimDepartment.DepartmentKey	N						
decimal	(8,2)		FK		Y						

- FactEmployeeSalary ghi lại Salary cho mỗi nhân viên

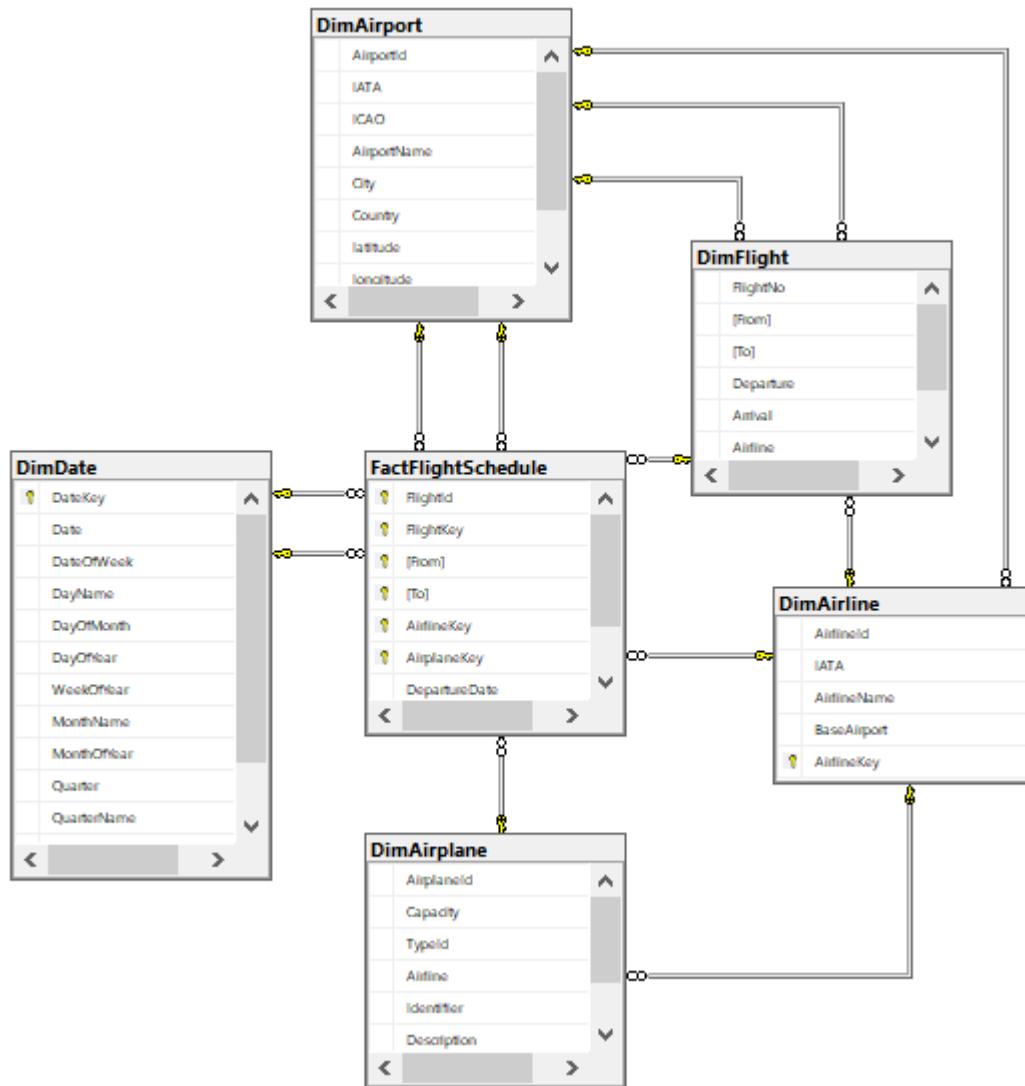
2.3 Các lược đồ mô tả kho dữ liệu



Lược đồ hình sao Star Schema mô tả FactEmployeeSalary và các bảng Dimension



Lược đồ hình sao Star Schema mô tả FactBookingSales và các bảng Dimension



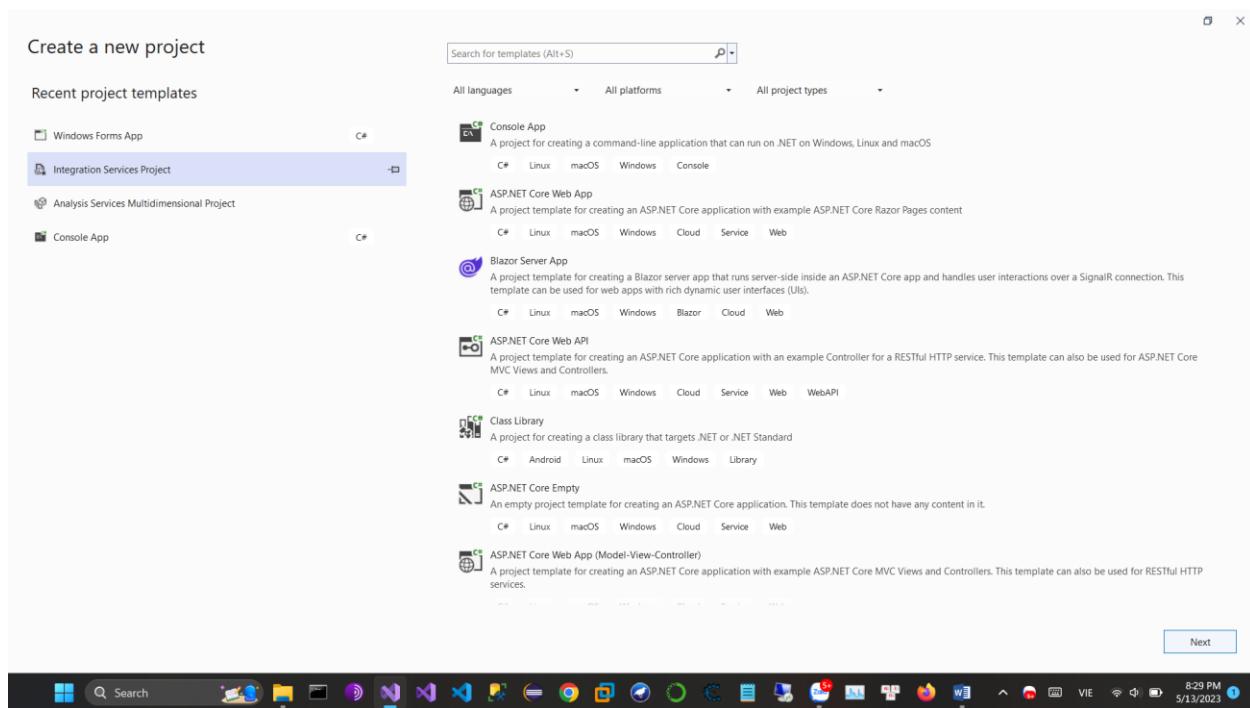
Lược đồ hình bông tuyết Snowflake mô tả FactFlightSchedule và các bảng Dimension

CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)

Ở chương này, Sử dụng công cụ SSIS xây dựng các gói tích hợp dữ liệu, cài đặt các gói chạy tự động. Viết truy vấn đưa dữ liệu từ OLTP đến kho dữ liệu.

3.1 Quá trình tạo mới project SSIS

- Chọn tag **File -> New -> Project..**
- Một dialog xuất hiện. Chọn **Integration Services -> Integration Services Project -> Nhập Name và Solution name. -> OK**



Configure your new project

Integration Services Project

Project name

Location

Solution

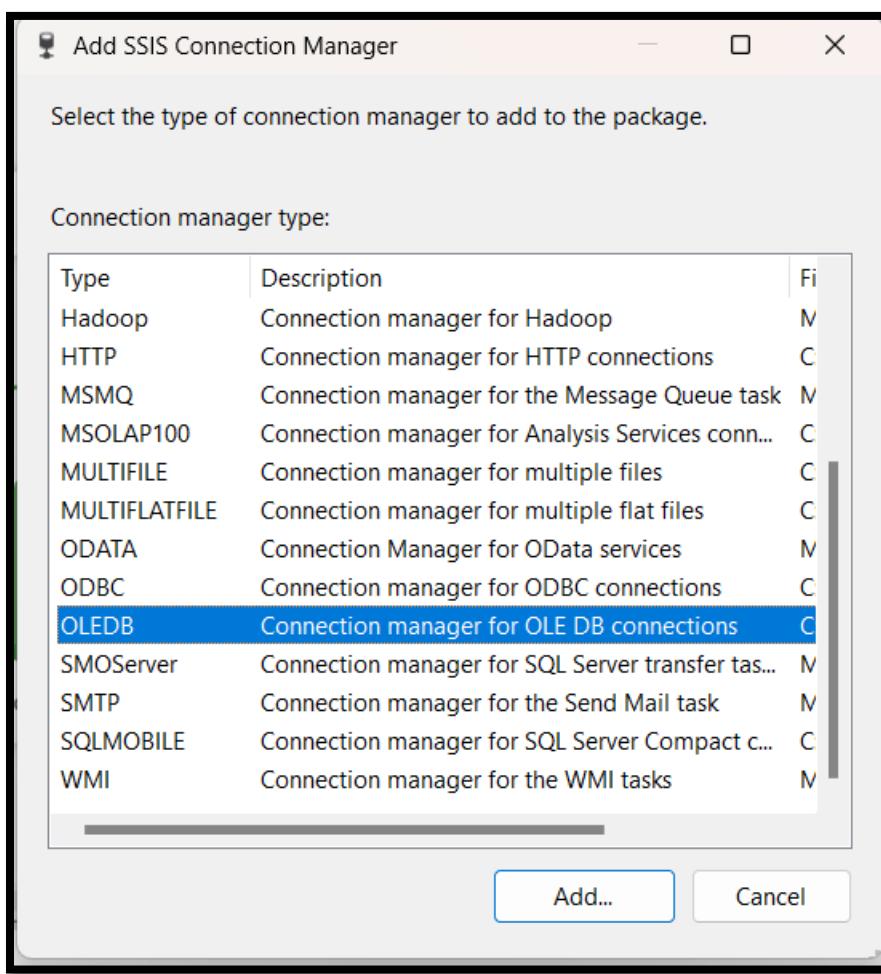
Solution name (i)

Place solution and project in the same directory

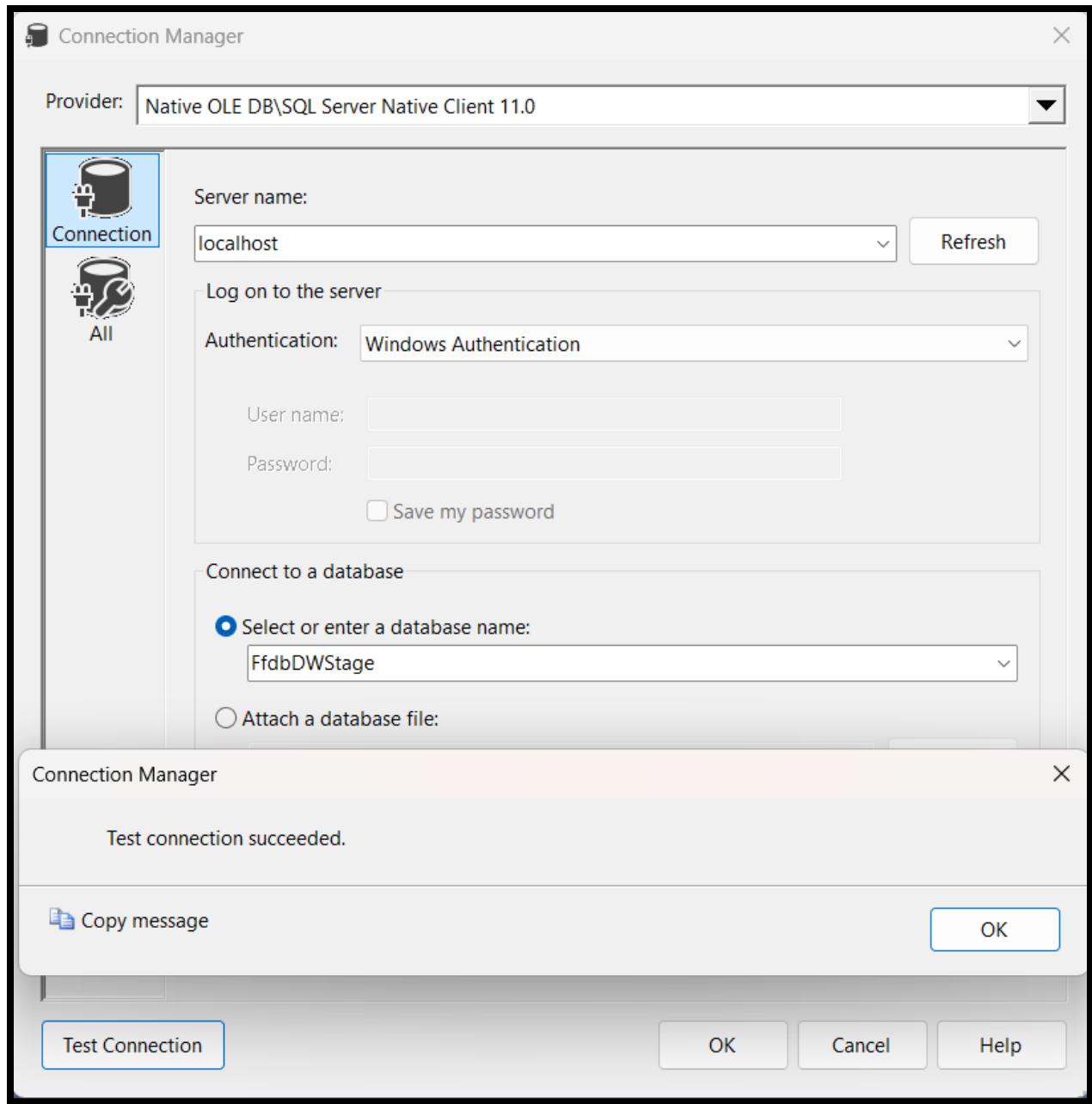
3.2 Quá trình đổ dữ liệu vào database Stage

3.2.1 Quá trình tạo Connection Manager

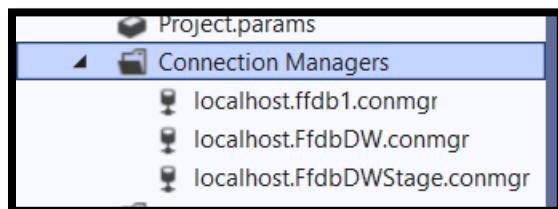
Nhấn chuột phải chọn New Connection:



Chọn New...

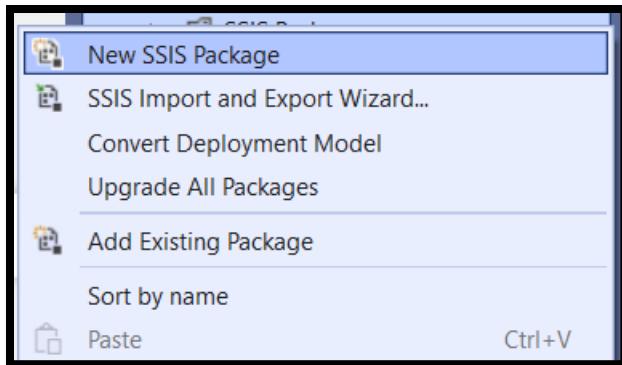


Tương tự với 2 DataBase còn lại

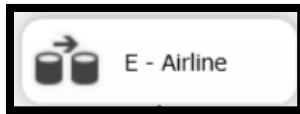


3.2.2 Quá trình đẩy dữ liệu vào Stage

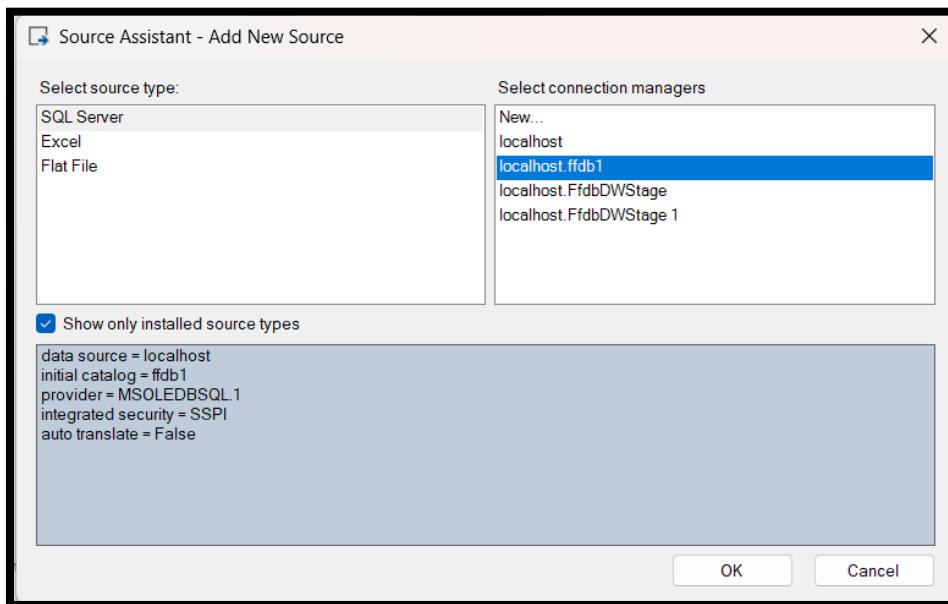
- Tạo SSIS Package mới tên stg_ffdb.dtsx



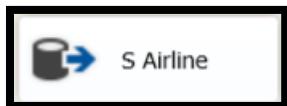
- Từ SSIS Toolbox, kéo Data Flow Task và đặt tên: E(extract)-Name(table)
➤ E - Airline



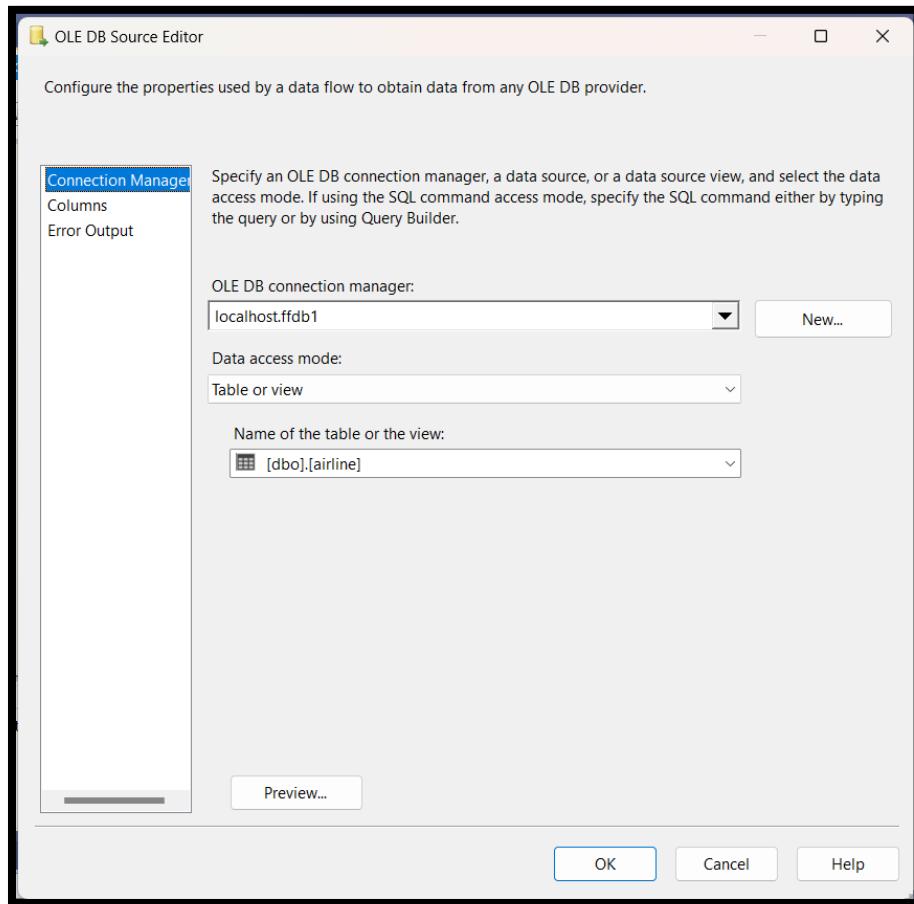
- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



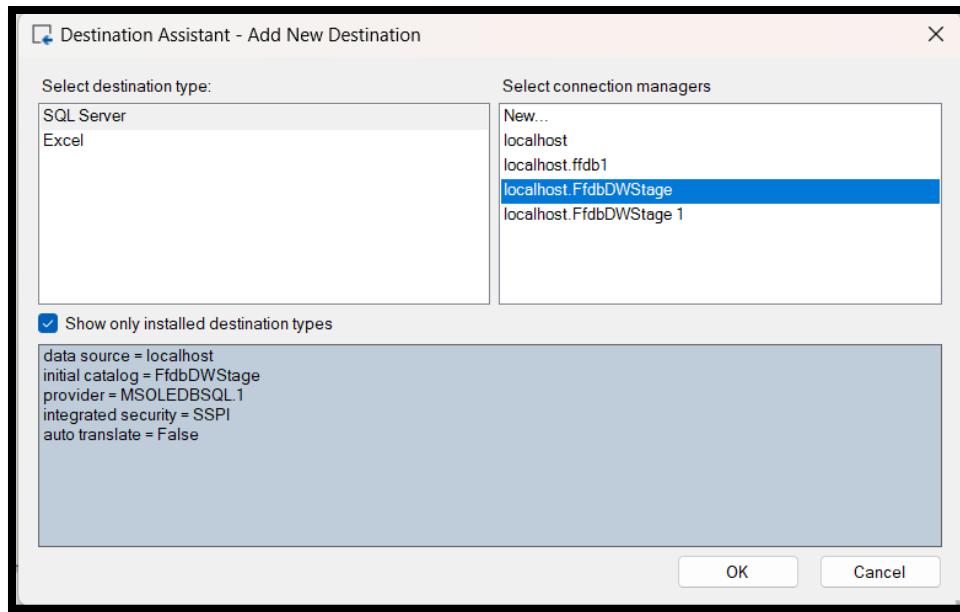
- Đặt tên S(Source) + Name(table)



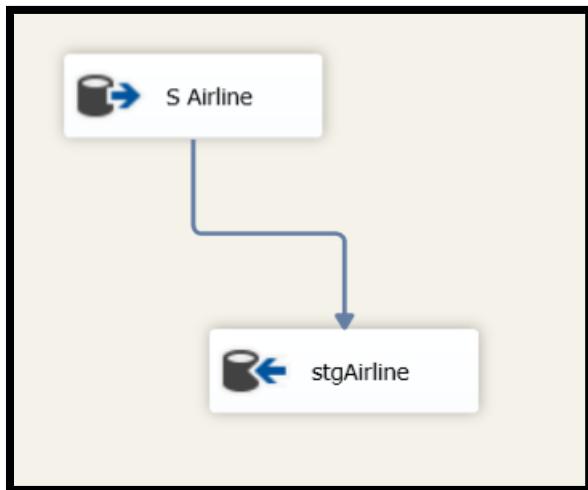
- Chọn table Airline, OK.



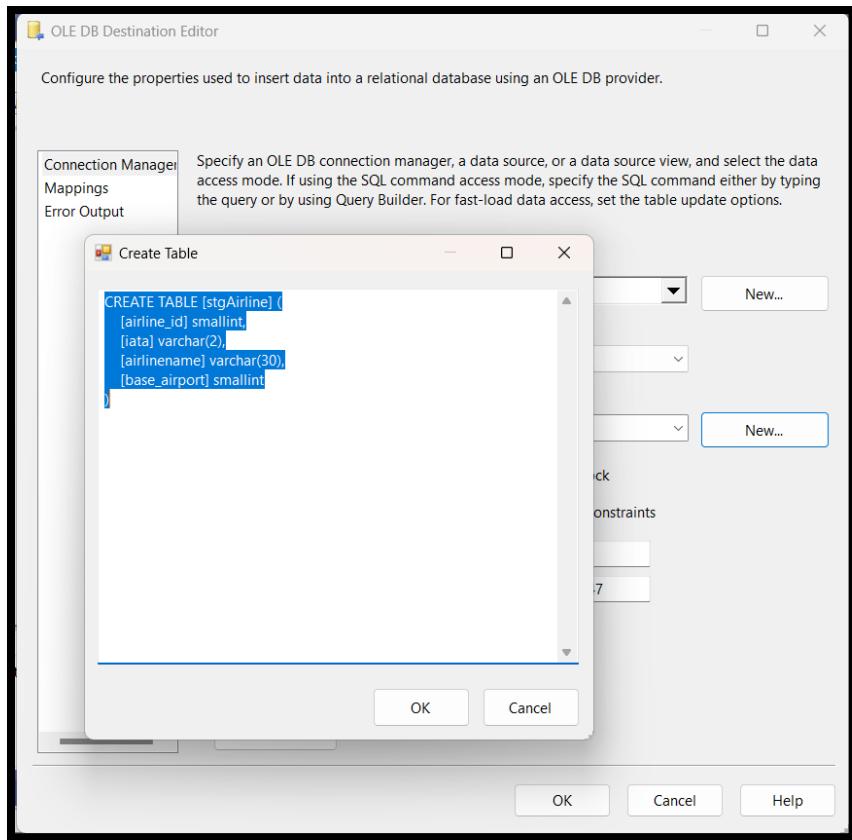
- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source

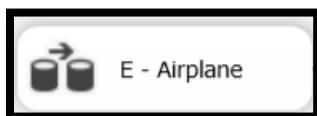


- Chọn New... để tạo table stg mới

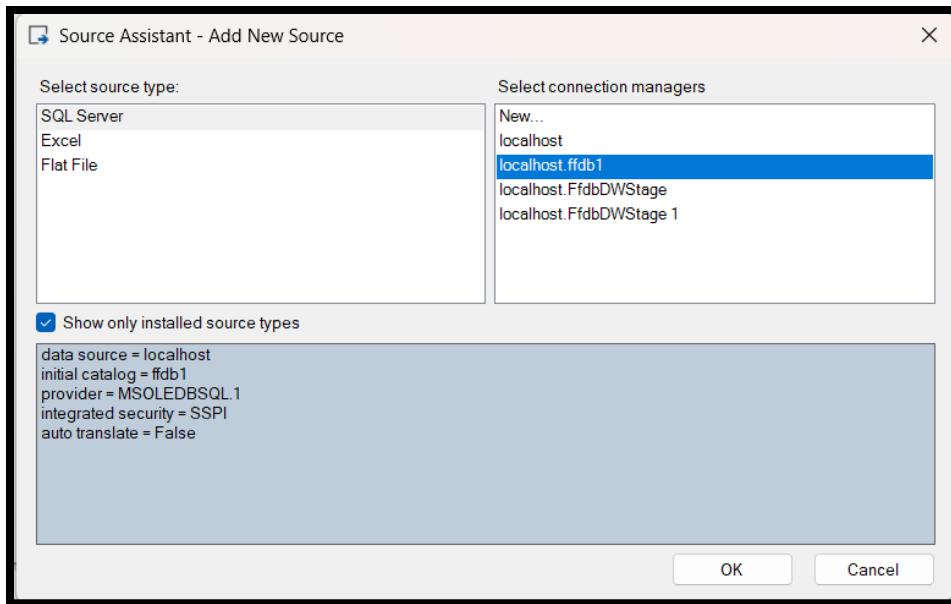


Nhấn F5 để run

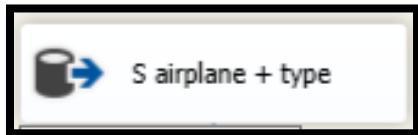
➤ E – Airplane



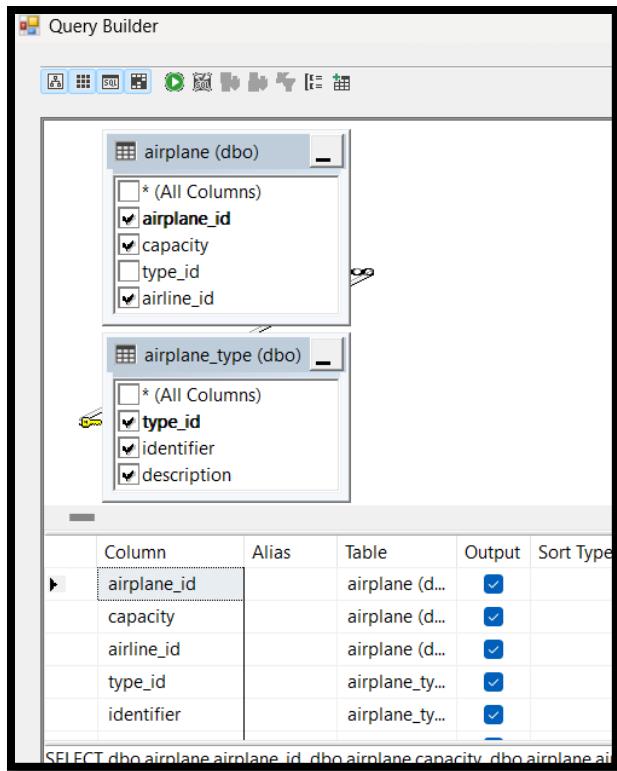
- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



- Đặt tên S(Source) + Name(table)



- Tại Data access mode, chọn SQL command, chọn Build Query, chọn những column cần thiết.



Data access mode:

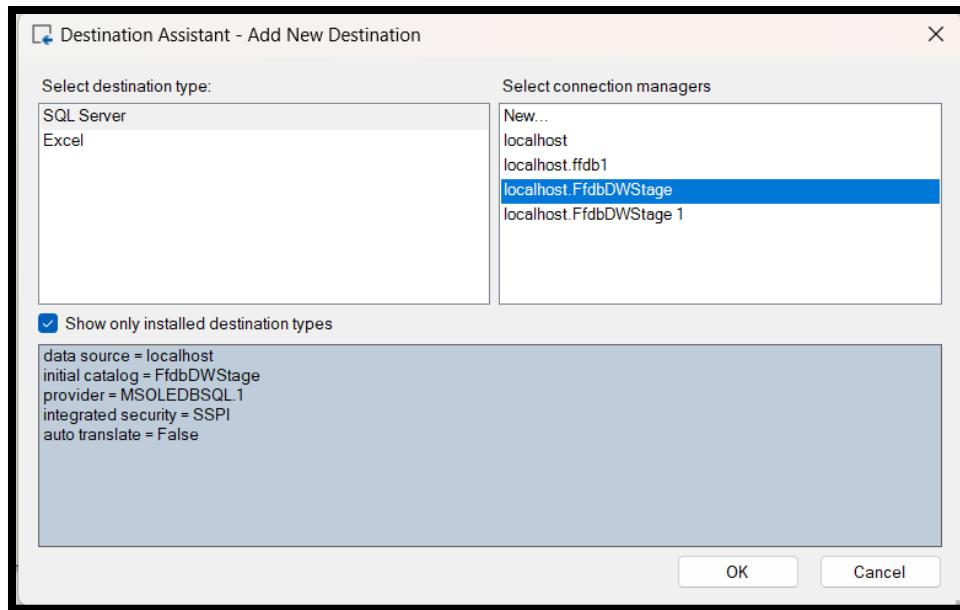
SQL command

SQL command text:

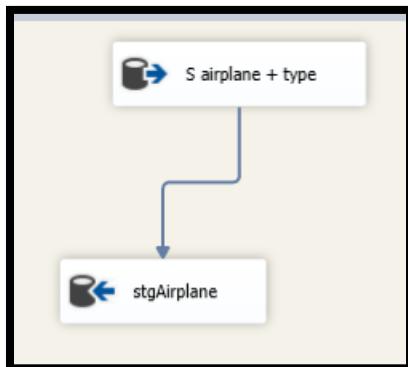
```
SELECT dbo.airplane.airplane_id, dbo.airplane.capacity,
dbo.airplane.airline_id, dbo.airplane_type.type_id,
dbo.airplane_type.identifier, dbo.airplane_type.description
FROM   dbo.airplane, dbo.airplane_type
WHERE  dbo.airplane.type_id = dbo.airplane_type.type_id
```

Parameters...
Build Query...
Browse...
Parse Query

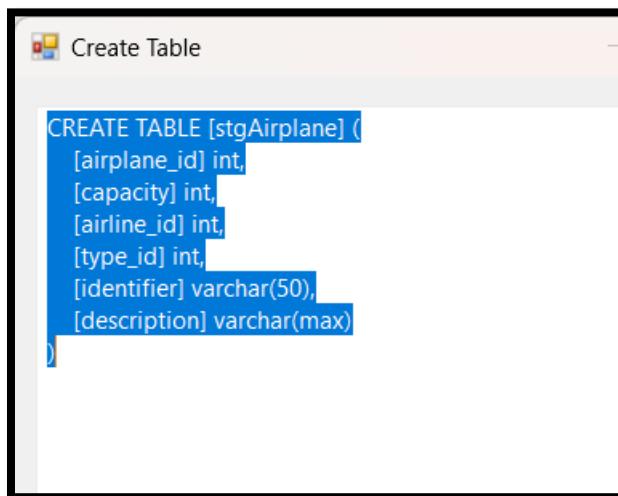
- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source

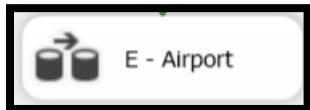


- Chọn New... để tạo table stg mới

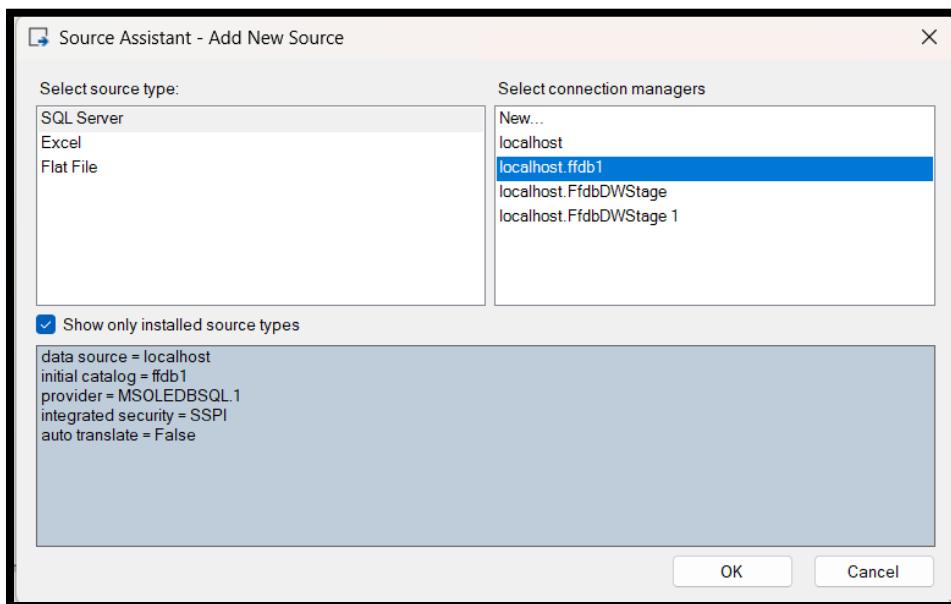


Nhấn F5 để run

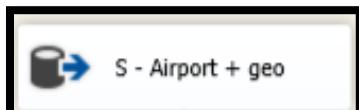
➤ E – Airport



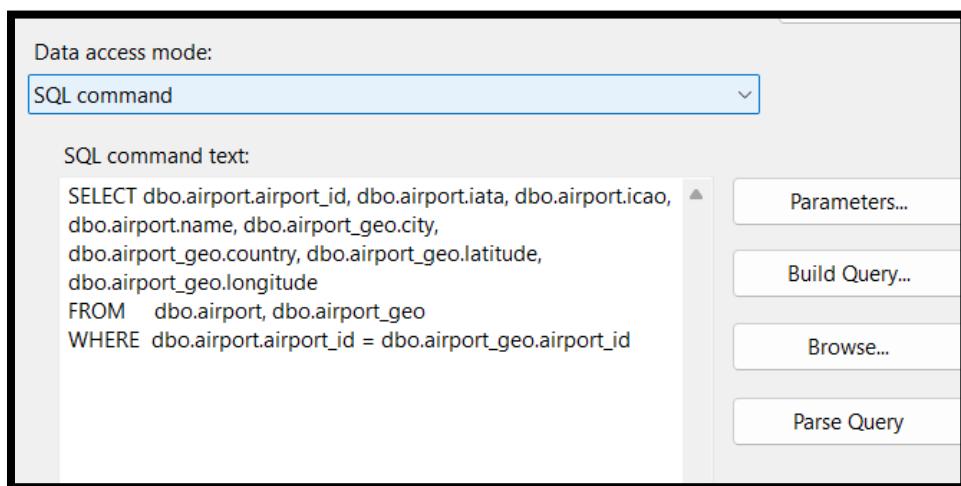
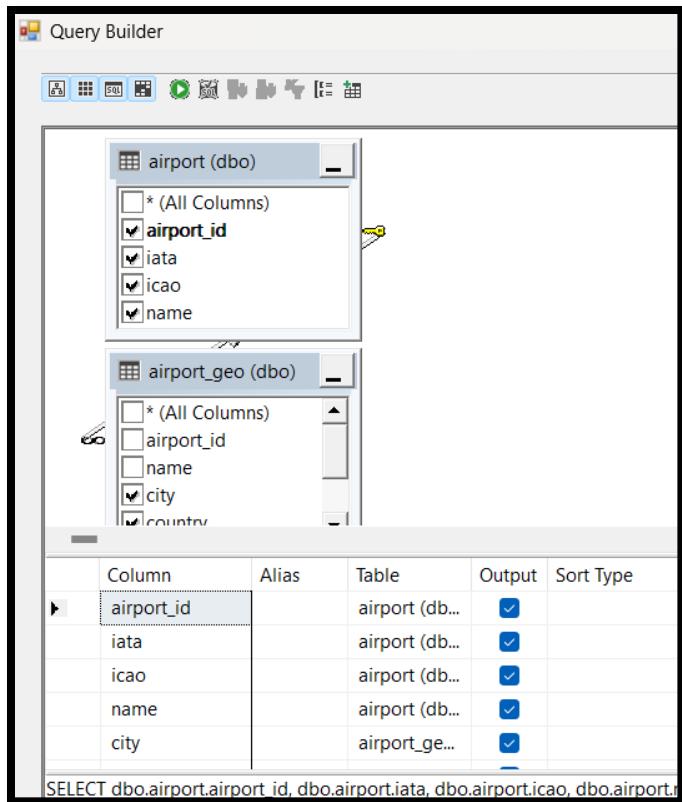
- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



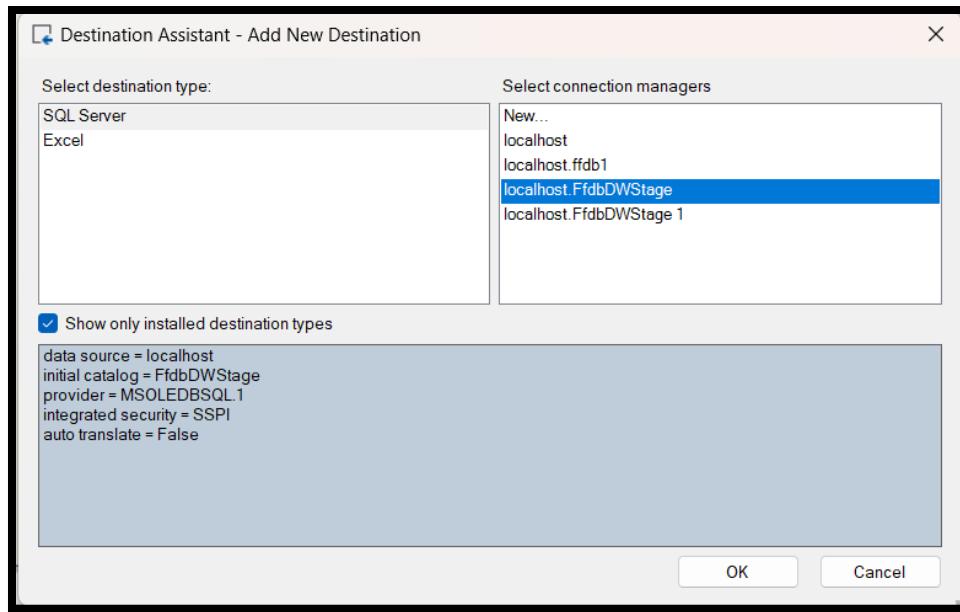
- Đặt tên S(Source) + Name(table)



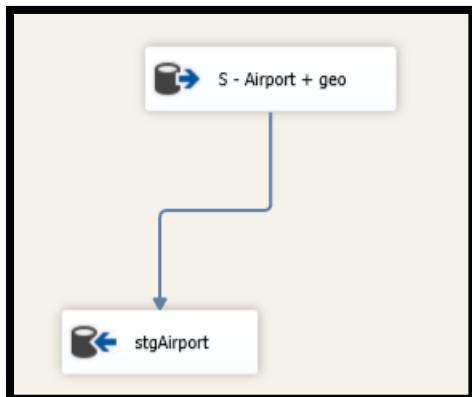
- Tại Data access mode, chọn SQL command, chọn Build Query, chọn những column cần thiết.



- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source

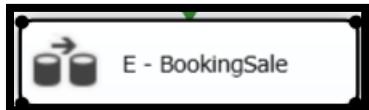


- Chọn New... để tạo table stg mới

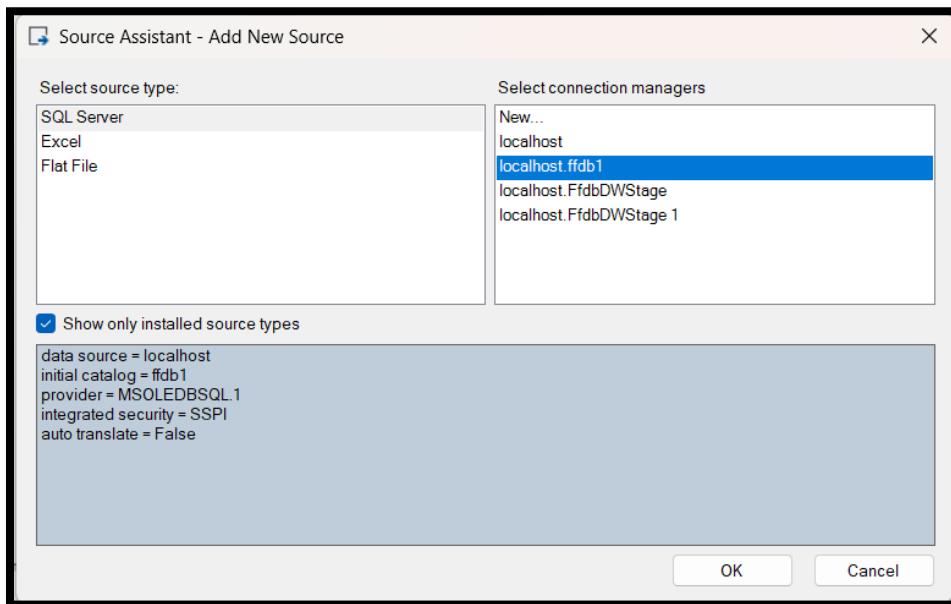


Nhấn F5 để run

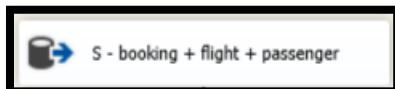
➤ E – BookingSale



- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



- Đặt tên S(Source) + Name(table)



- Tại Data access mode, chọn SQL command, chọn Build Query, chọn những column cần thiết.

The screenshot shows the Query Builder interface with two tables selected:

- booking (dbo)** table with columns: * (All Columns), booking_id, flight_id, seat, passenger_id. The booking_id and seat columns are checked.
- flight (dbo)** table with columns: * (All Columns), flight_id, flightno, from, to. The flightno column is checked.

The resulting query table below lists the selected columns and their aliases:

Column	Alias	Table	Output	Sort Type
booking_id		booking (d...)	✓	
flightno		flight (dbo)	✓	
passenger_id		passenger ...	✓	
seat		booking (d...)	✓	
price		booking (d...)	✓	

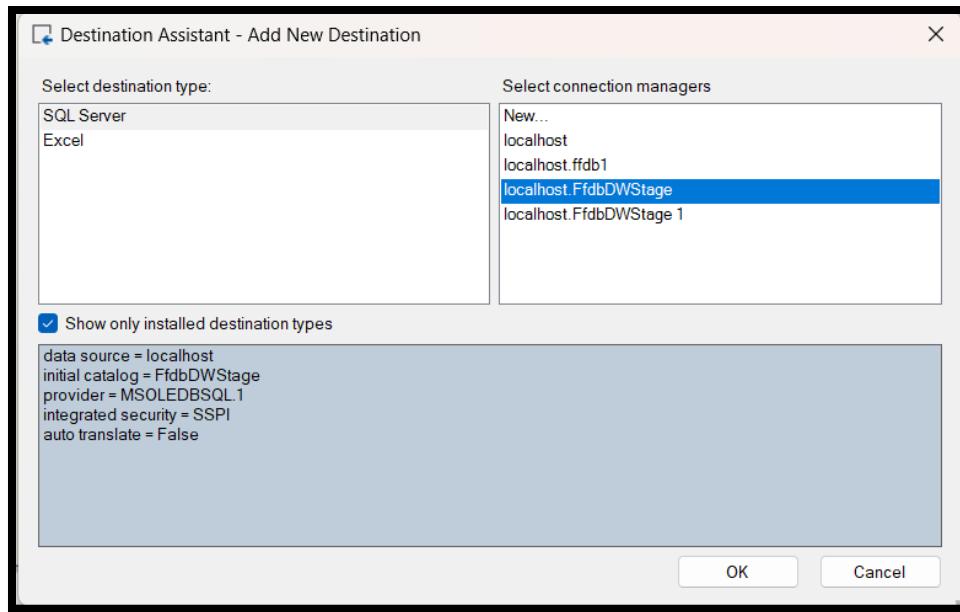
Data access mode: SQL command

SQL command text:

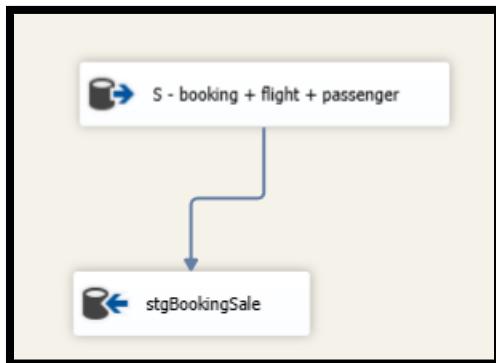
```

SELECT dbo.booking.booking_id, dbo.flight.flightno,
dbo.passenger.passenger_id, dbo.booking.seat,
dbo.booking.price, dbo.flight.departure, dbo.flight.arrival
FROM dbo.booking, dbo.flight, dbo.passenger
WHERE dbo.booking.flight_id = dbo.flight.flight_id AND
dbo.booking.passenger_id = dbo.passenger.passenger_id
    
```

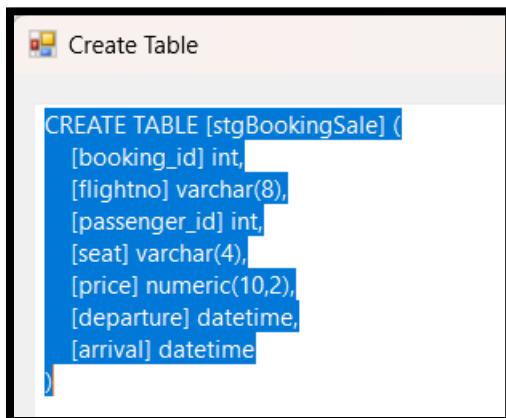
- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source

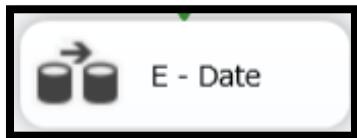


- Chọn New... để tạo table stg mới

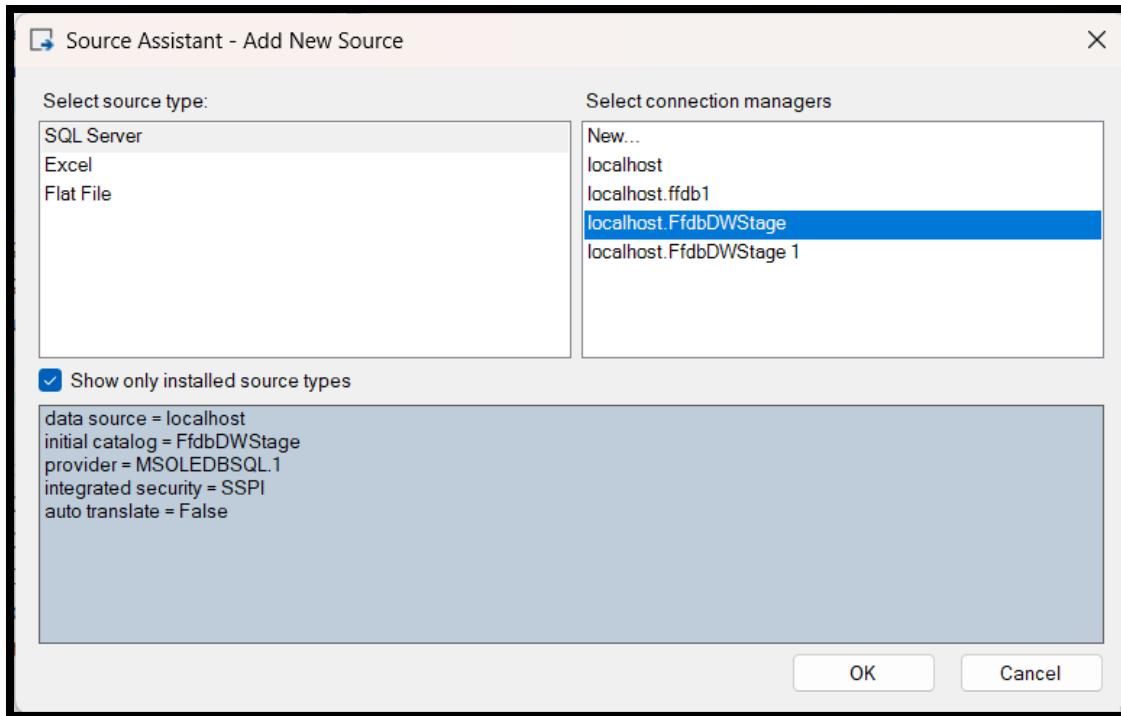


Nhấn F5 để run

➤ E – Date



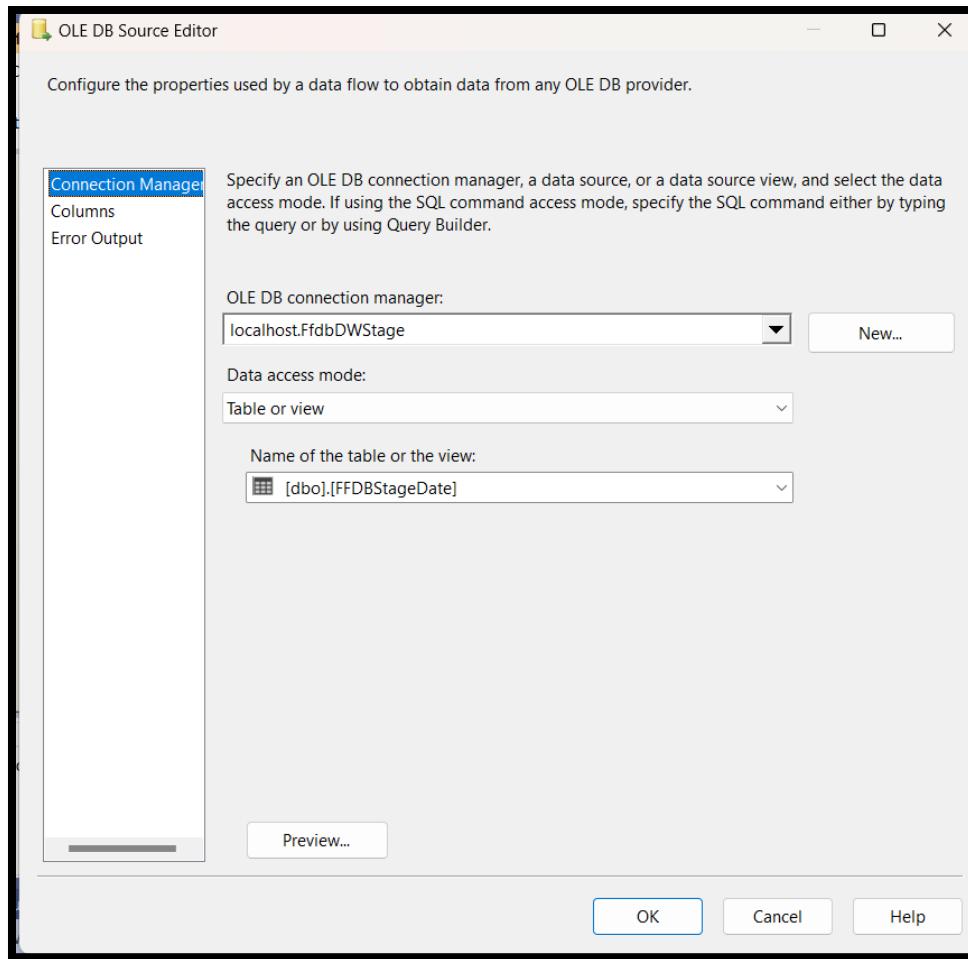
- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.FfdbDWStage



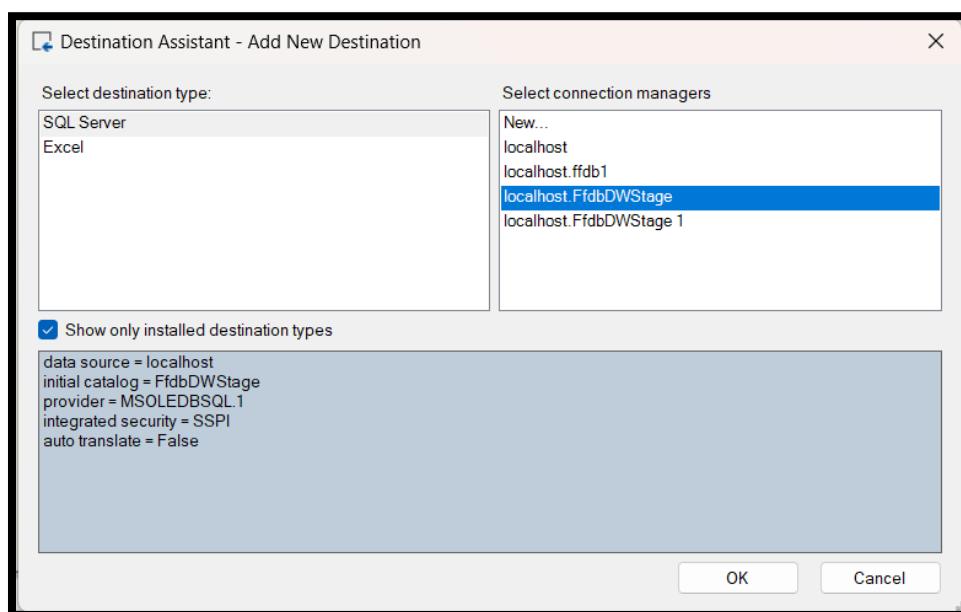
- Đặt tên S(Source) + Name(table)



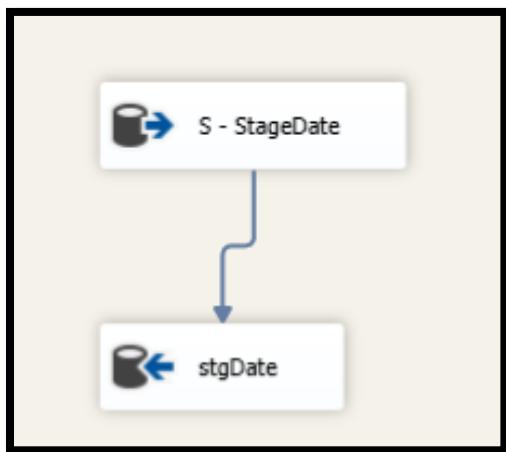
- Chọn table FFDBStageDate, OK.



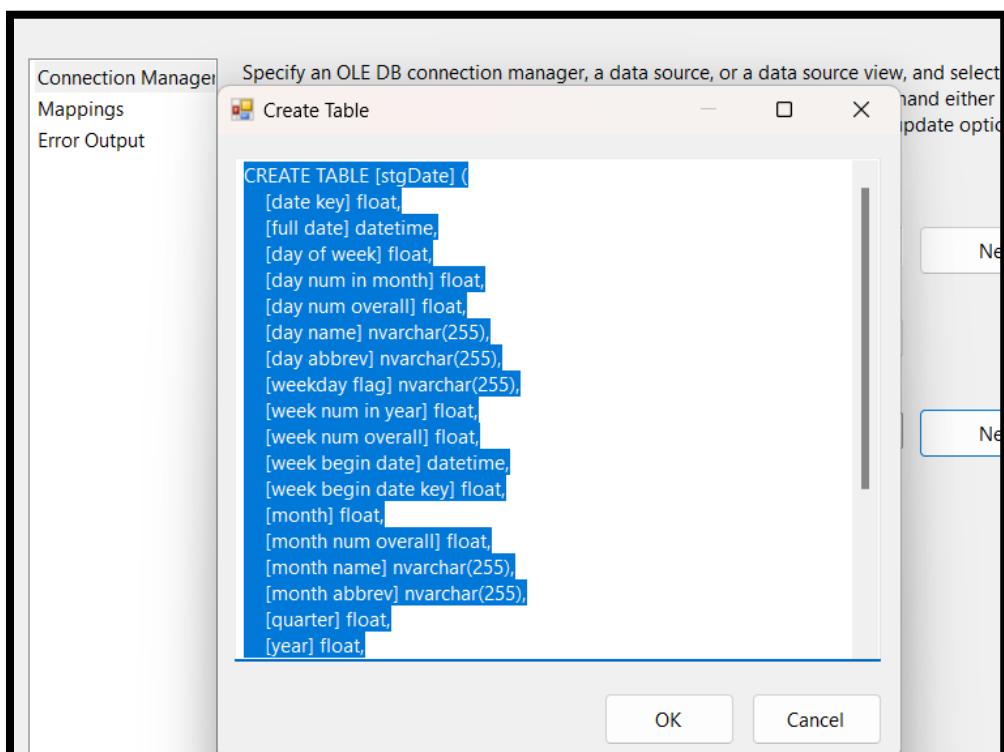
- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source



- Chọn New... để tạo table stg mới



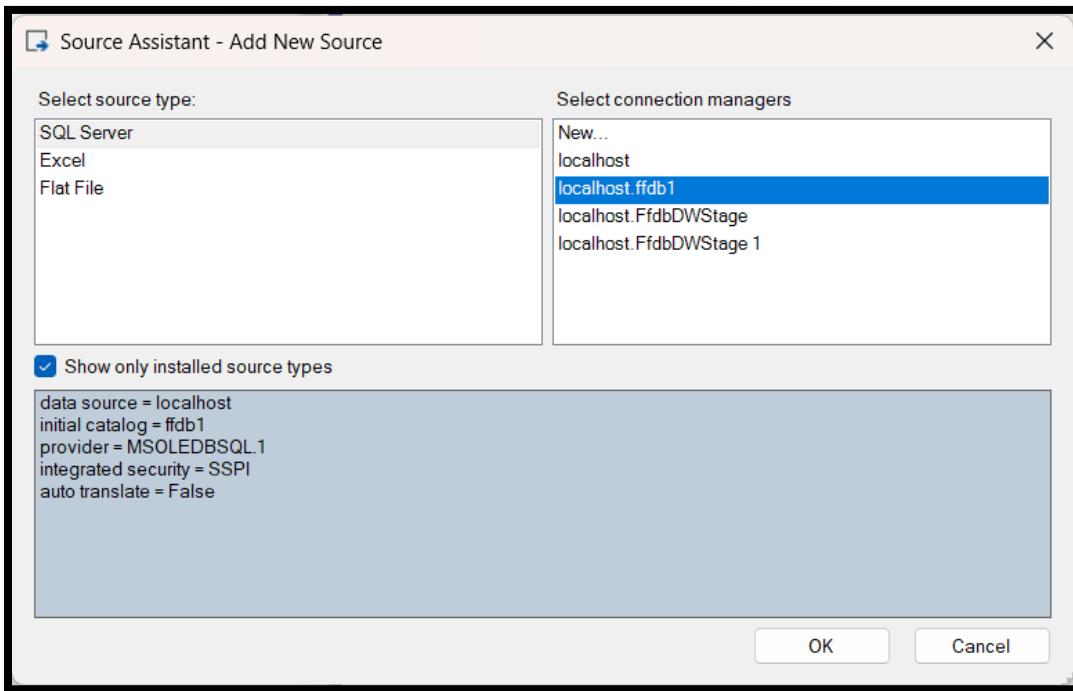
Nhấn F5 để run

➤ E – Employee

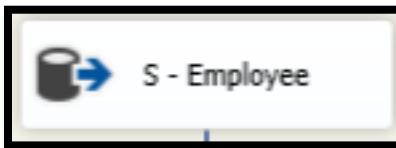


- Trong Data Flow:

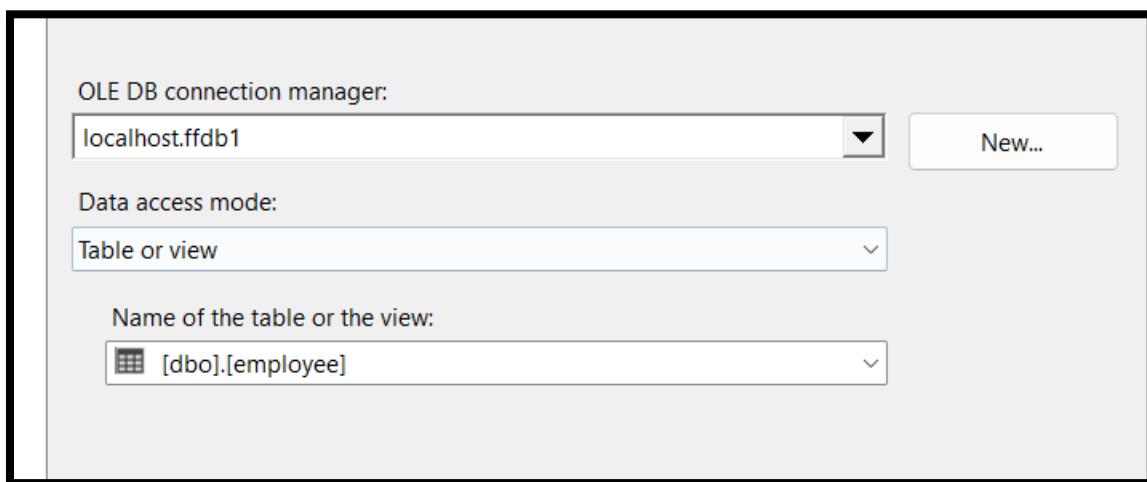
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



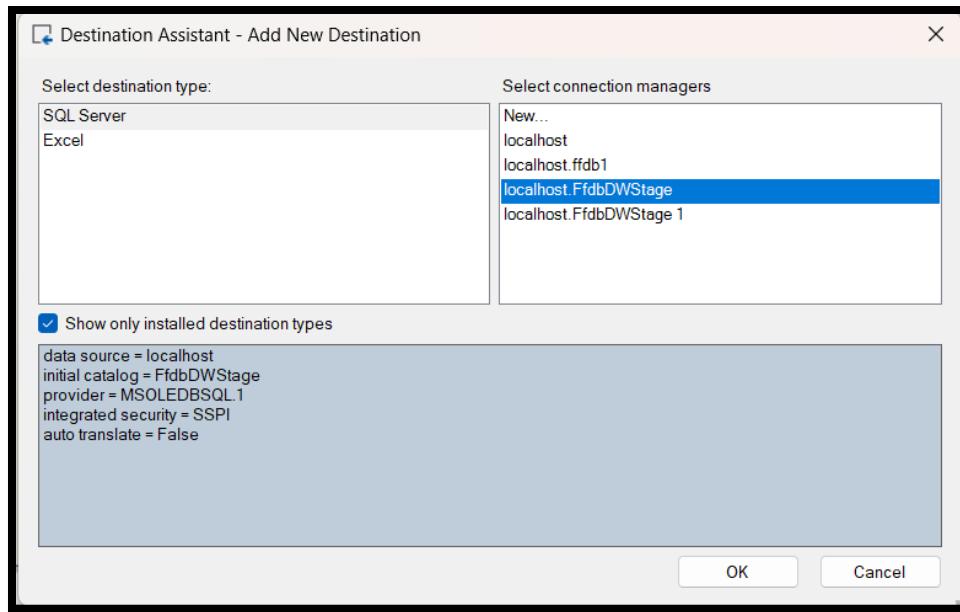
- Đặt tên S(Source) + Name(table)



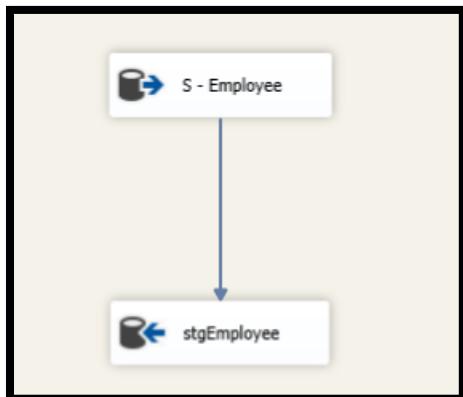
- Chọn table Employee, OK.



- Kéo thả Destination Assistant và chọn SQL localhost.FfdbDWStage



- Đặt tên và tạo kết nối từ Source

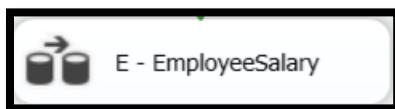


- Chọn New... để tạo table stg mới

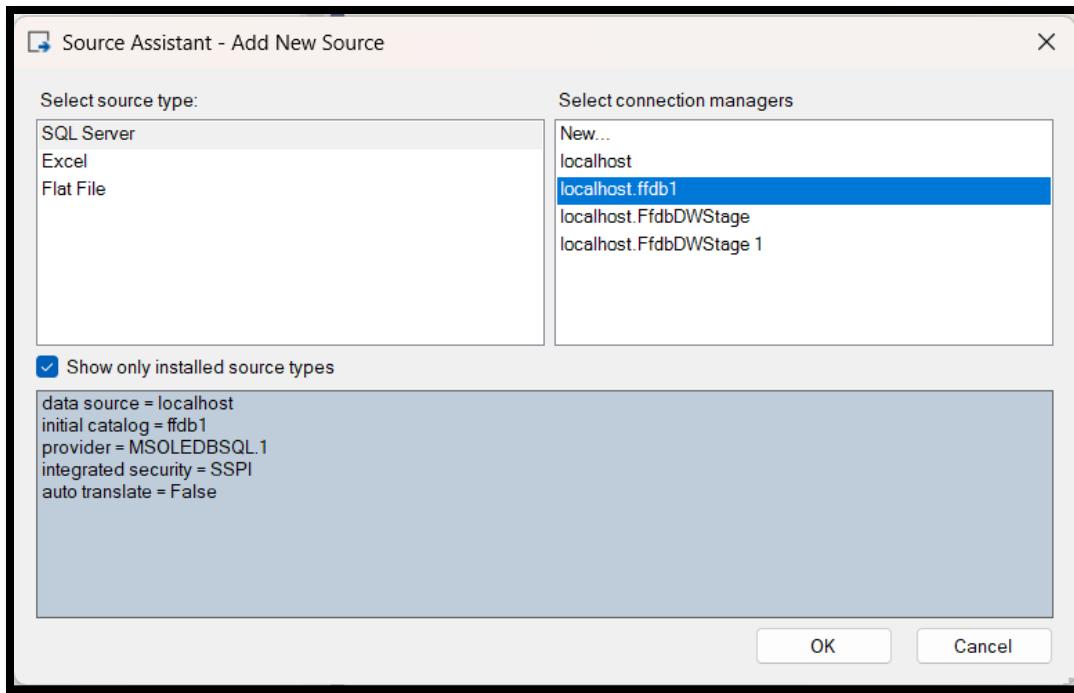
```
CREATE TABLE [stgEmployee] (
[employee_id] int,
[firstname] varchar(100),
[lastname] varchar(100),
[birthdate] date,
[sex] varchar(1),
[street] varchar(100),
[city] varchar(100),
[zip] smallint,
[country] varchar(100),
[emailaddress] varchar(120),
[telephoneno] varchar(30),
[salary] numeric(8,2),
[department] varchar(20),
[username] varchar(20),
[password] varchar(32)
)
```

Nhấn F5 để run

➤ E – EmployeeSalary



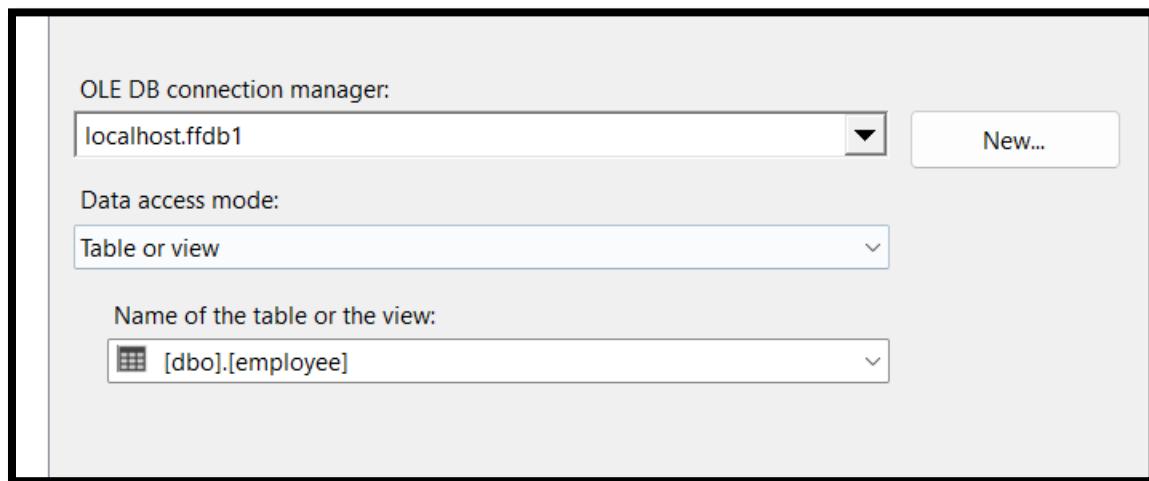
- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



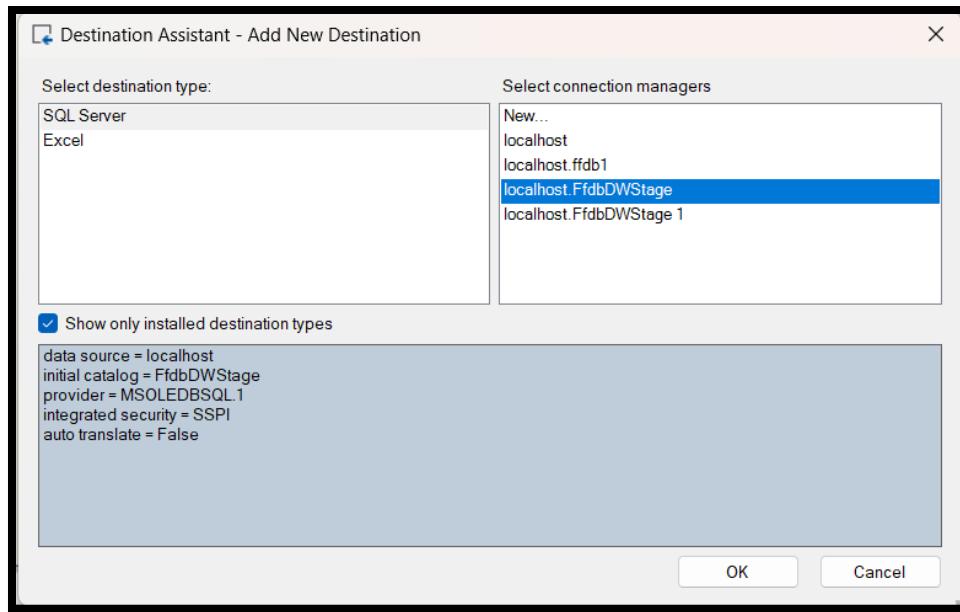
- Đặt tên S(Source) + Name(table)



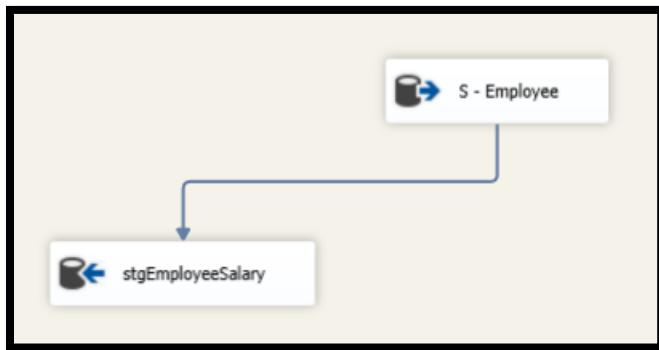
- Chọn table Employee, OK.



- Kéo thả Destination Assistant và chọn SQL localhost.FfdbDWStage



- Đặt tên và tạo kết nối từ Source

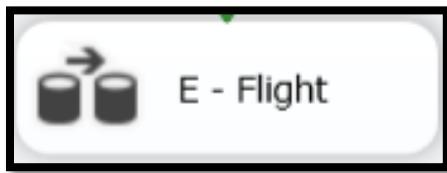


- Chọn New... để tạo table stg mới

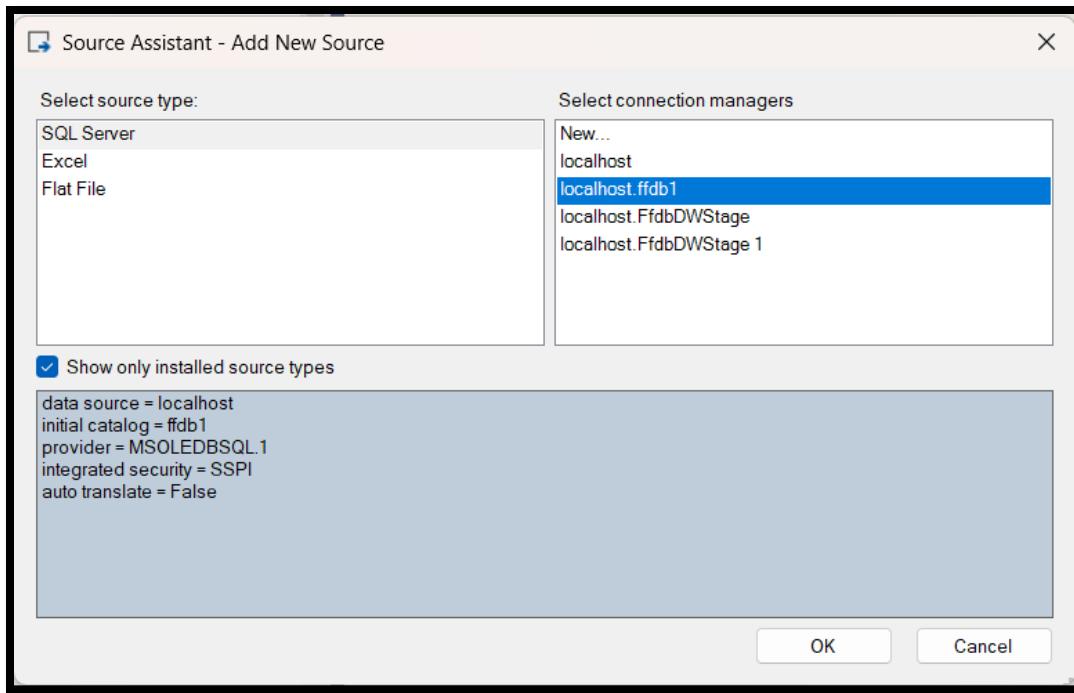
```
CREATE TABLE [stgEmployee] (
    [employee_id] int,
    [firstname] varchar(100),
    [lastname] varchar(100),
    [birthdate] date,
    [sex] varchar(1),
    [street] varchar(100),
    [city] varchar(100),
    [zip] smallint,
    [country] varchar(100),
    [emailaddress] varchar(120),
    [telephoneno] varchar(30),
    [salary] numeric(8,2),
    [department] varchar(20),
    [username] varchar(20),
    [password] varchar(32)
)
```

Nhấn F5 để run

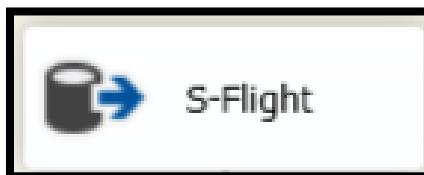
➤ E – Flight



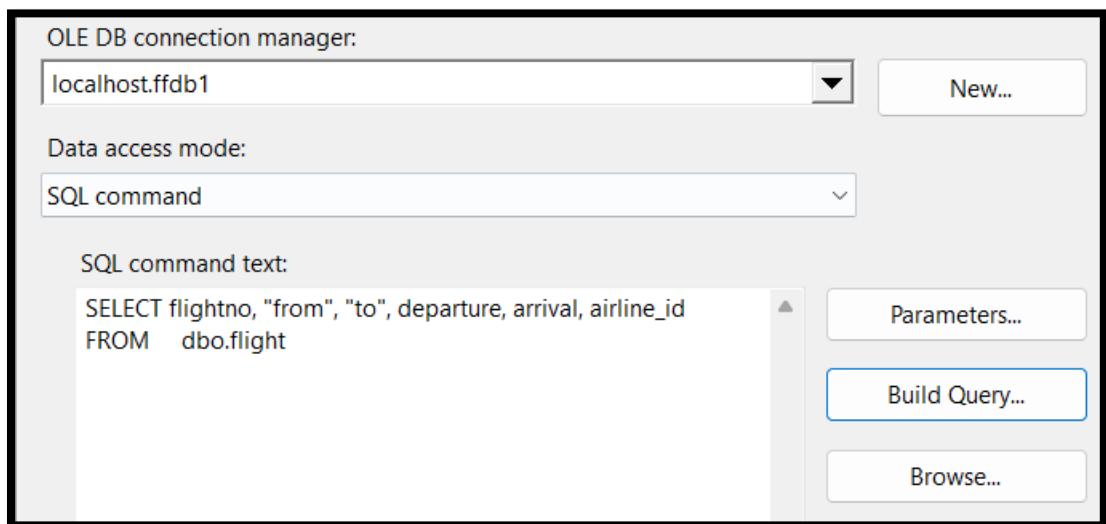
- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1

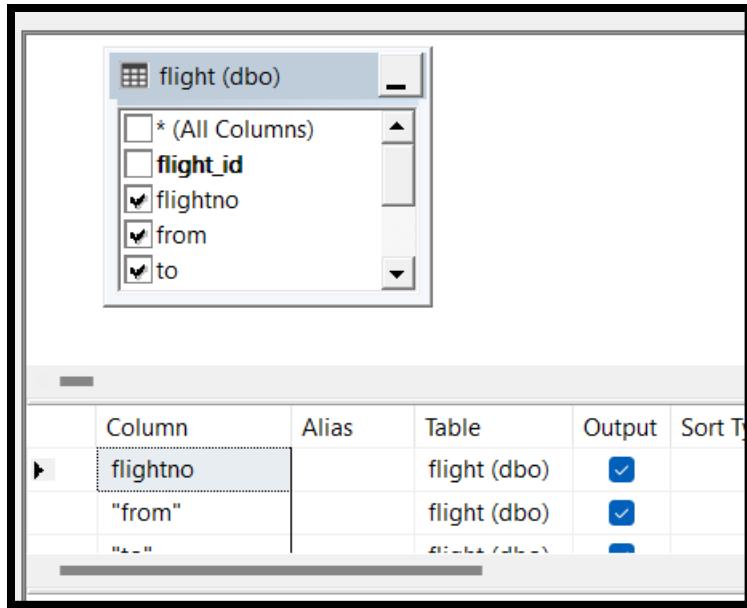


- Đặt tên S(Source) + Name(table)

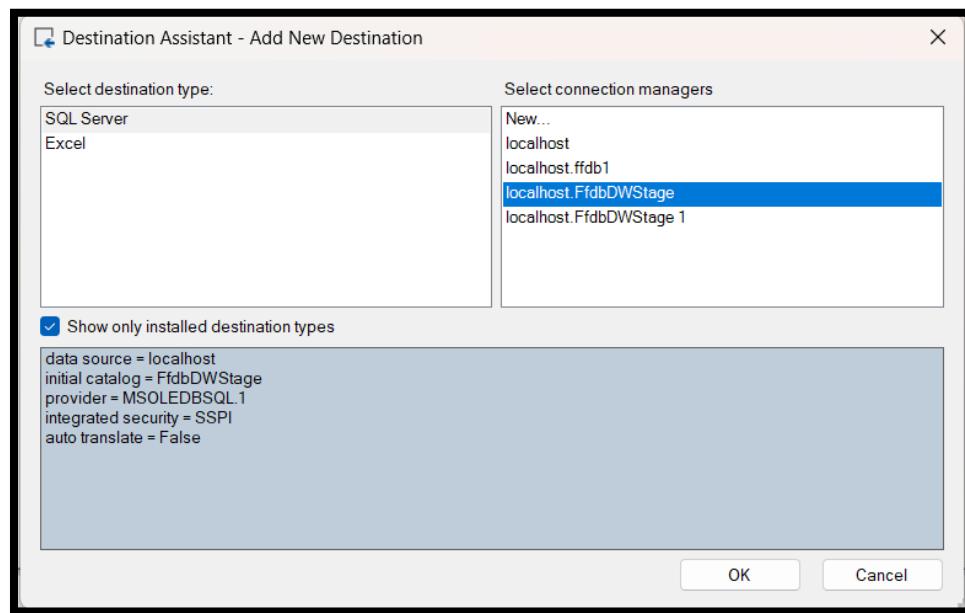


- Tại Data access task, chọn SQL command, Build Query... , chọn column mong muốn.

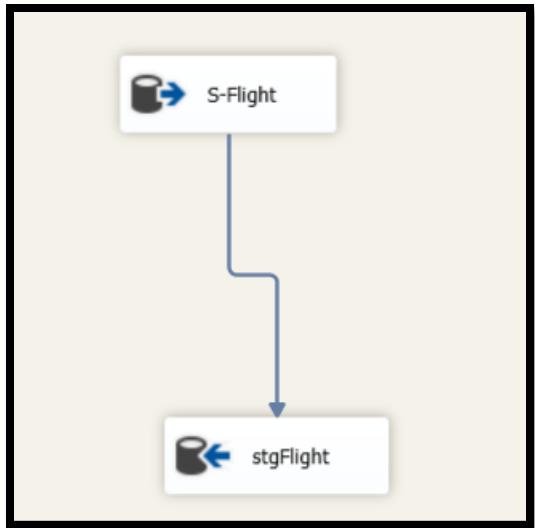




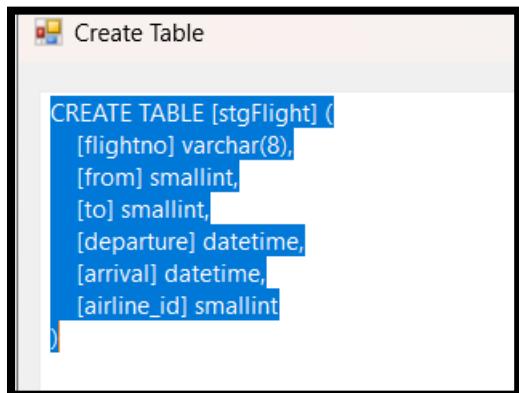
- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source

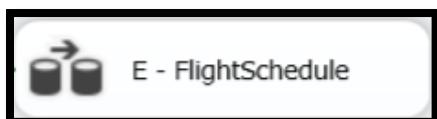


- Chọn New... để tạo table stg mới

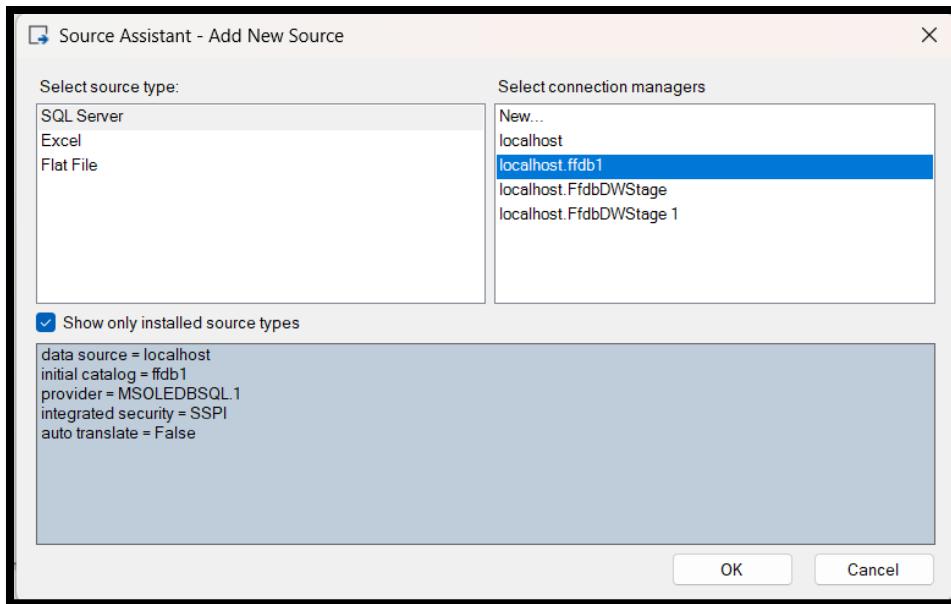


Nhấn F5 để run

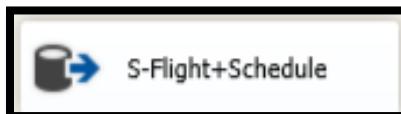
➤ E – FlightSchedule



- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1

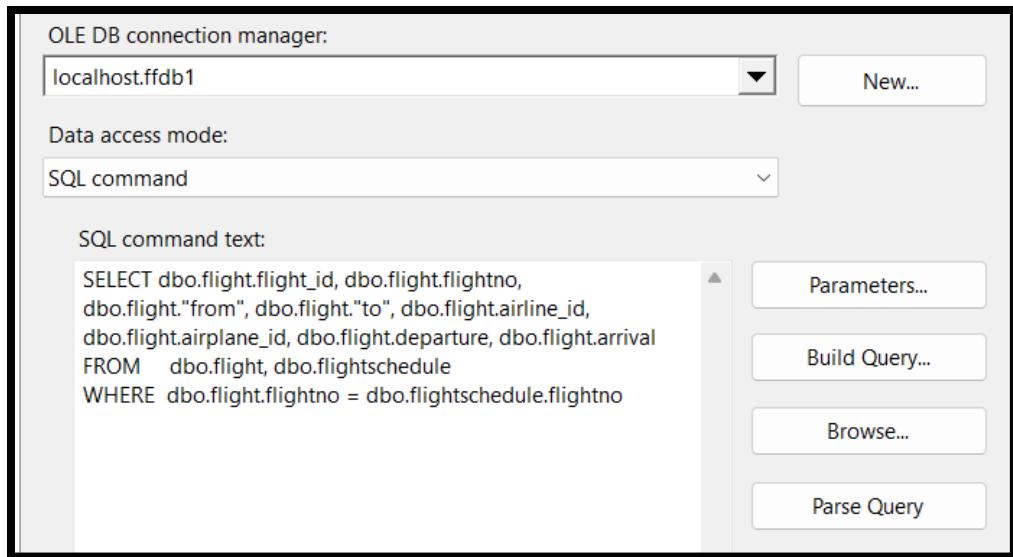


- Đặt tên S(Source) + Name(table)

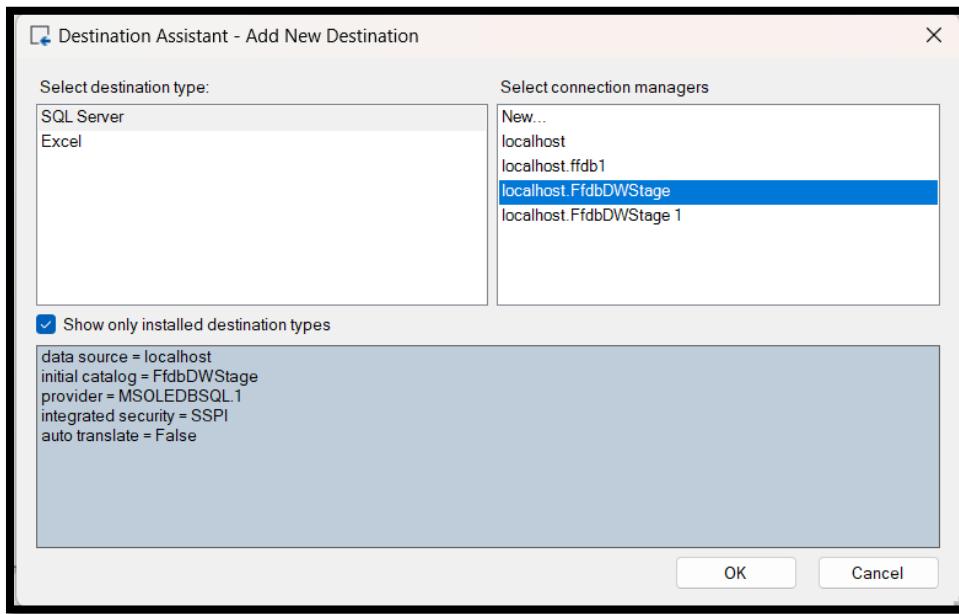


- Tại Data access mode, chọn SQL command, chọn Build Query, chọn những column cần thiết.

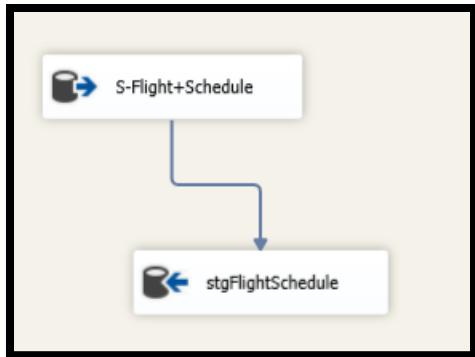
	Column	Alias	Table	Output	Sort Type	Sort Order
▶	flight_id		flight (dbo)	<input checked="" type="checkbox"/>		
	flightno		flight (dbo)	<input checked="" type="checkbox"/>		



- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source



- Chọn New... để tạo table stg mới

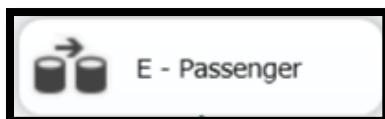
```

CREATE TABLE [stgFlightSchedule] (
    [flight_id] int,
    [flightno] varchar(8),
    [from] smallint,
    [to] smallint,
    [airline_id] smallint,
    [airplane_id] int,
    [departure] datetime,
    [arrival] datetime
)

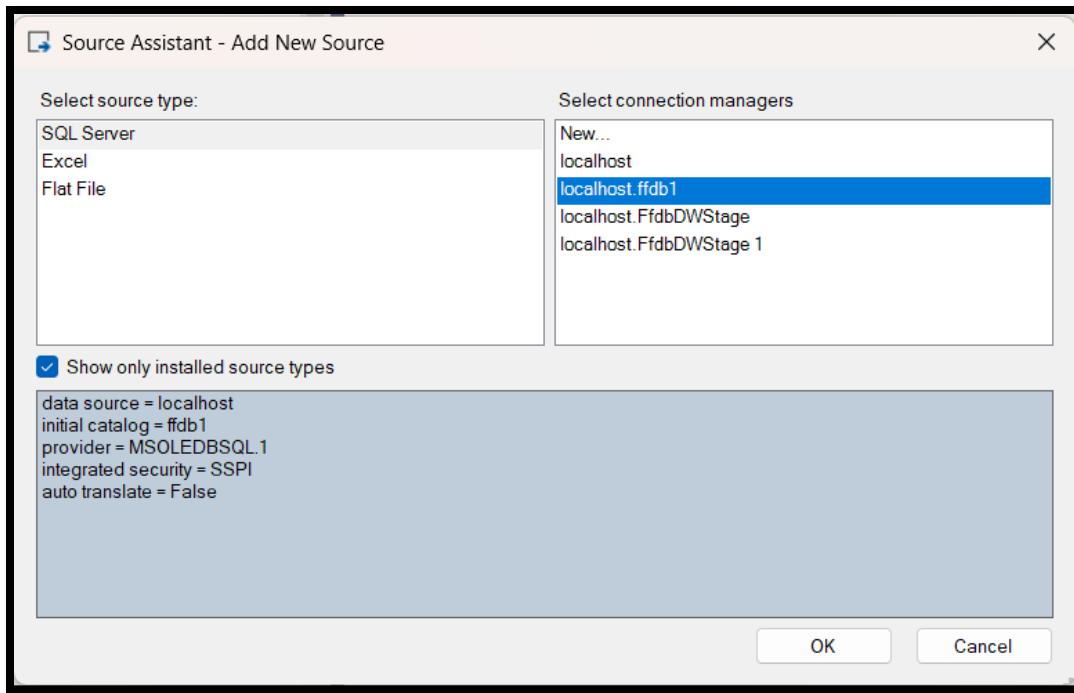
```

Nhấn F5 để run

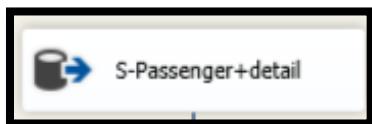
➤ E – Passenger



- Trong Data Flow:
- Kéo thả Source Assistant và chọn SQL localhost.ffdb1



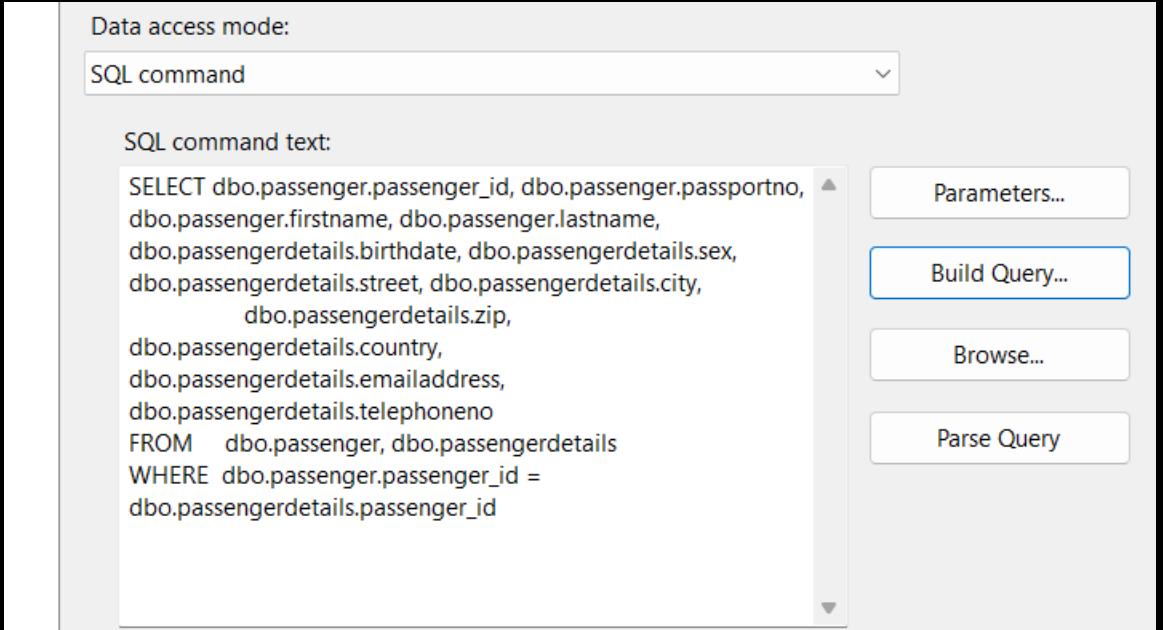
- Đặt tên S(Source) + Name(table)



- Chọn table Employee, OK.

The screenshot shows the 'passenger (dbo)' table in the object explorer. The 'passenger_id' column is selected. Below it, the 'passengerdetails (...)' table is shown. The main pane displays a mapping grid:

Column	Alias	Table	Output	Sort Type	Sort Order
passenger_id		passenger ...	<input checked="" type="checkbox"/>		
passportno		passenger ...	<input checked="" type="checkbox"/>		

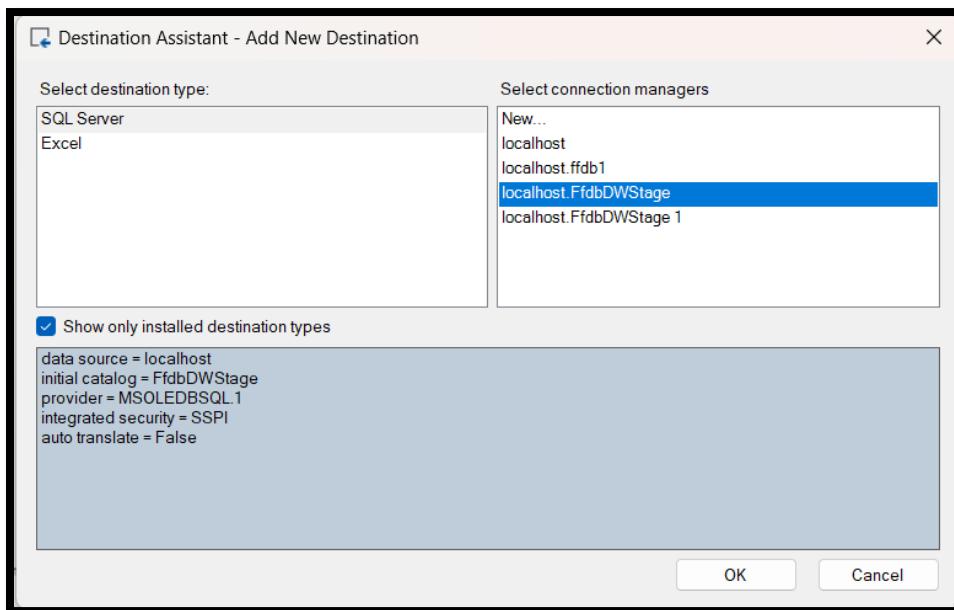


The screenshot shows a SQL command editor interface. At the top, it says "Data access mode: SQL command". Below that is a large text area containing a complex SQL query:

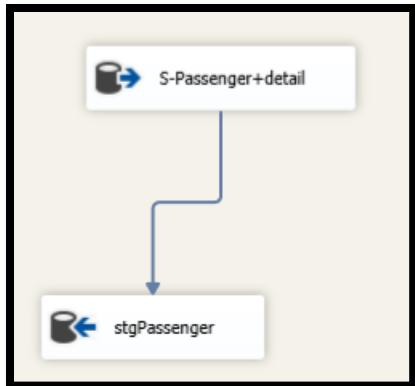
```
SELECT dbo.passenger.passenger_id, dbo.passenger.passportno,
dbo.passenger.firstname, dbo.passenger.lastname,
dbo.passengerdetails.birthdate, dbo.passengerdetails.sex,
dbo.passengerdetails.street, dbo.passengerdetails.city,
dbo.passengerdetails.zip,
dbo.passengerdetails.country,
dbo.passengerdetails.emailaddress,
dbo.passengerdetails.telephoneno
FROM dbo.passenger, dbo.passengerdetails
WHERE dbo.passenger.passenger_id =
dbo.passengerdetails.passenger_id
```

To the right of the query text are several buttons: "Parameters...", "Build Query...", "Browse...", and "Parse Query".

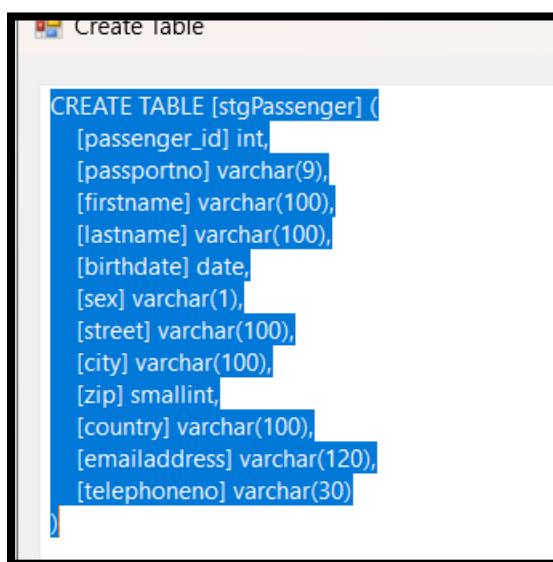
- Kéo thả Destination Assistant và chọn SQL localhost.ffdbDWStage



- Đặt tên và tạo kết nối từ Source

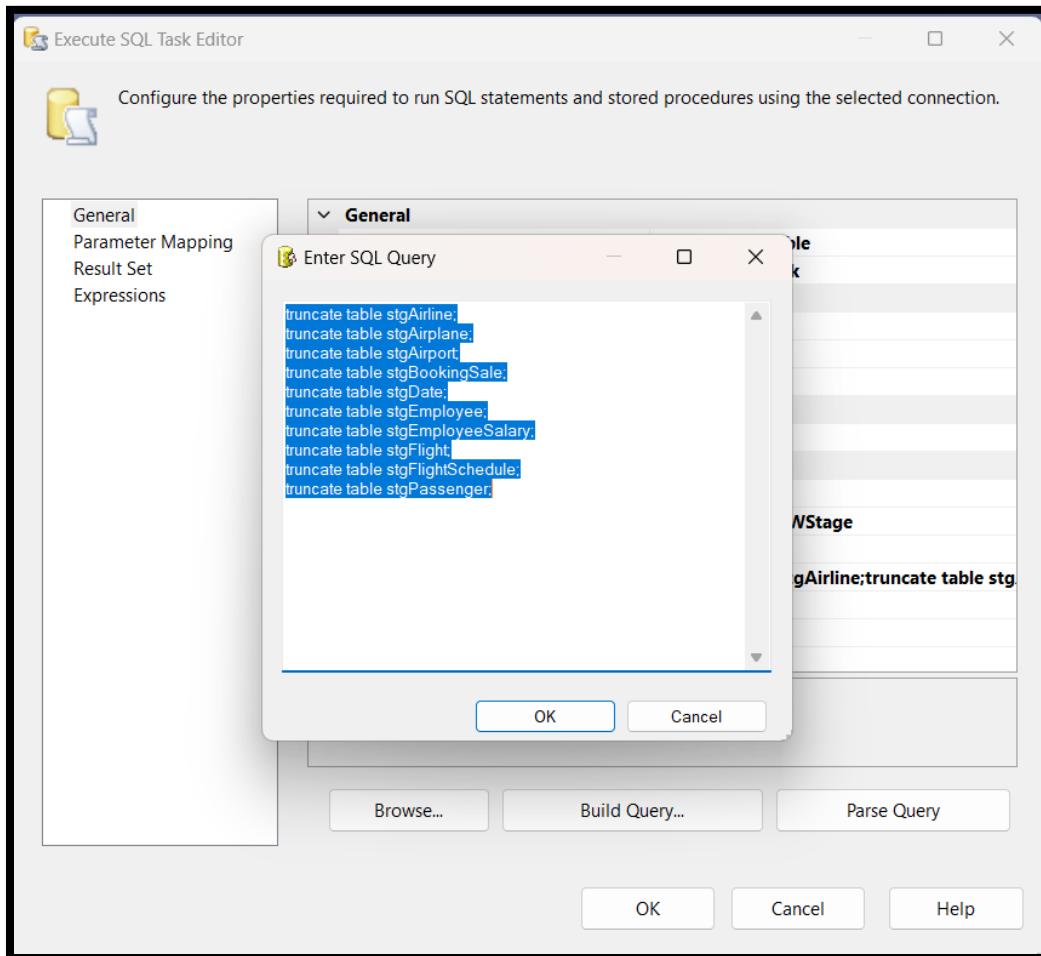
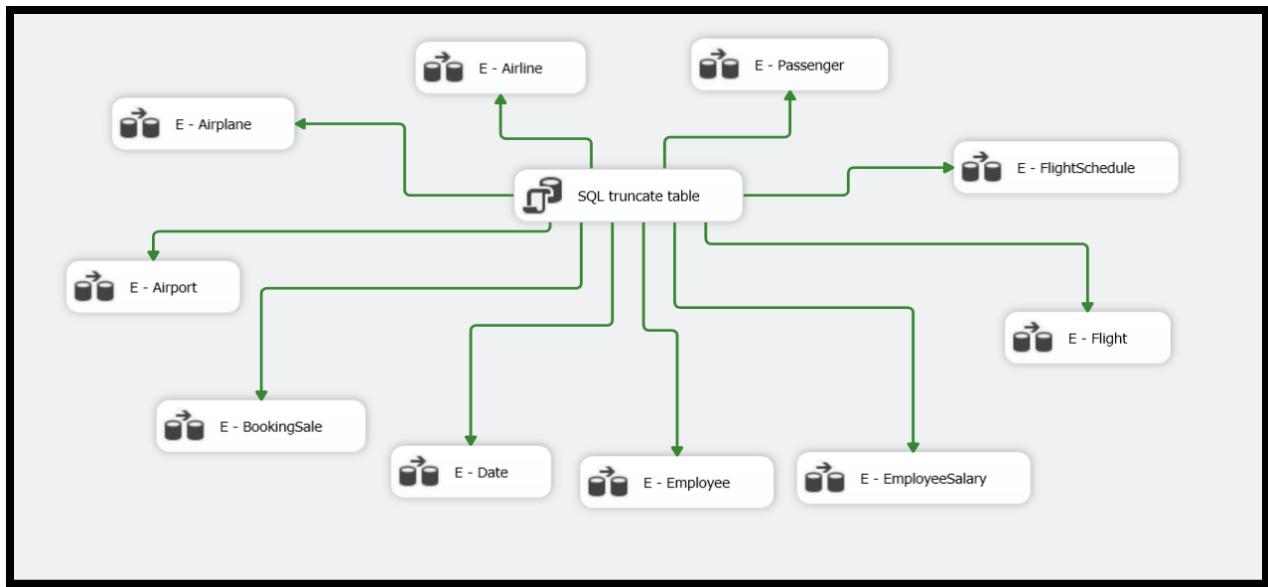


- Chọn New... để tạo table stg mới

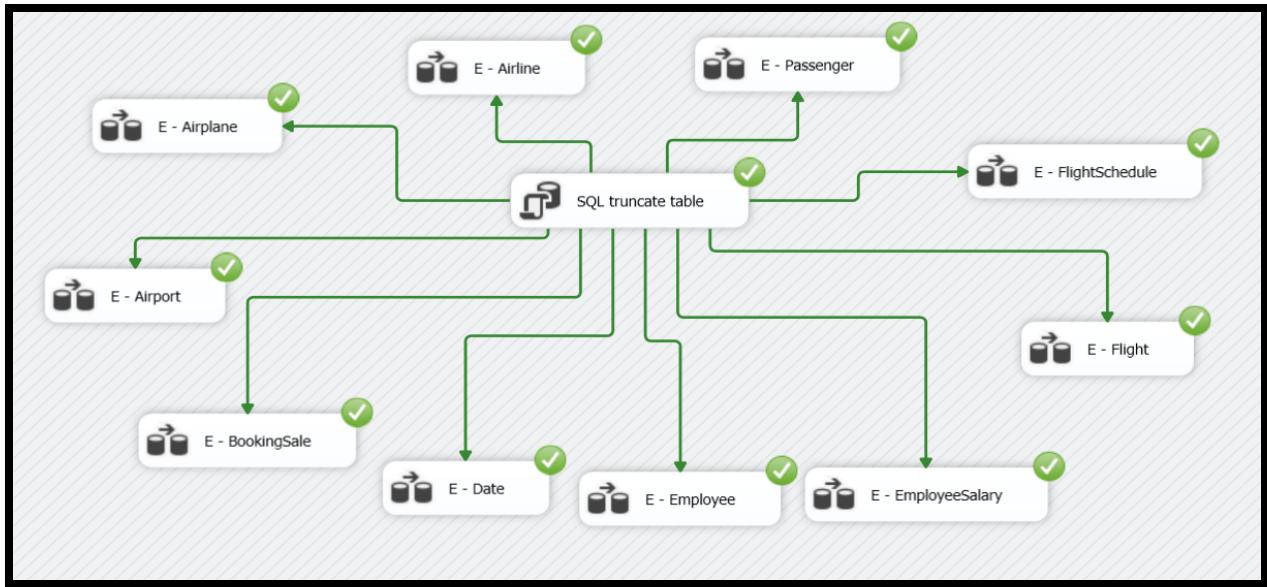


Nhấn F5 để run

Tại Control Flow, chọn SQL truncate table và kết nối với tất cả Data Flow



Hoàn tất quá trình E (Extract) và T (Transform):



3.3 Quá trình Load dữ liệu vào database DW

- Tạo 1 SSIS Package mới, DW_ffdb.dtsx

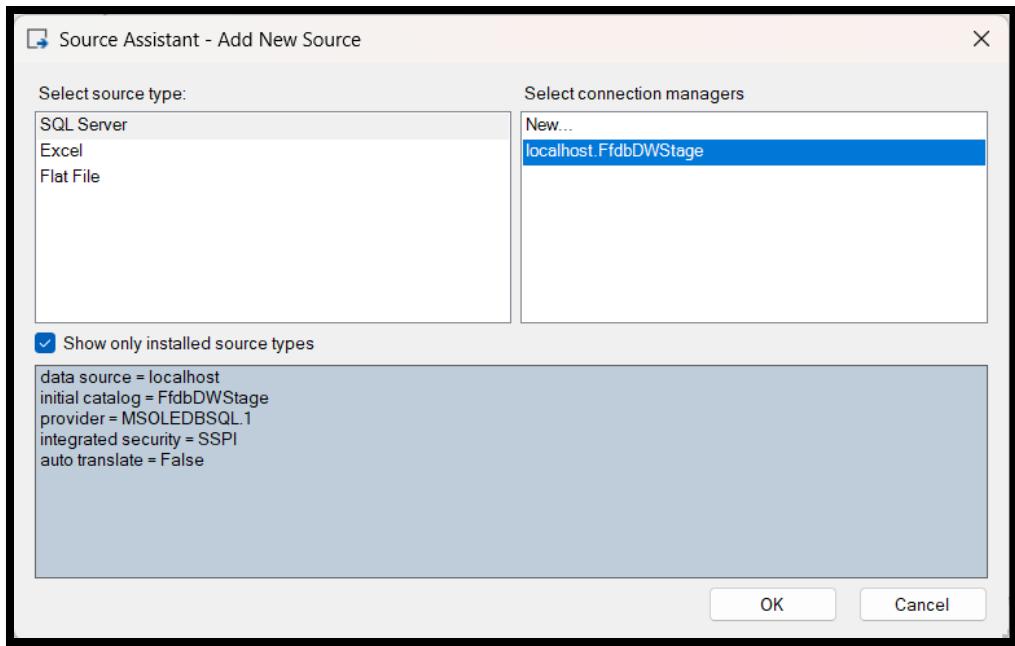


➤ L – Address

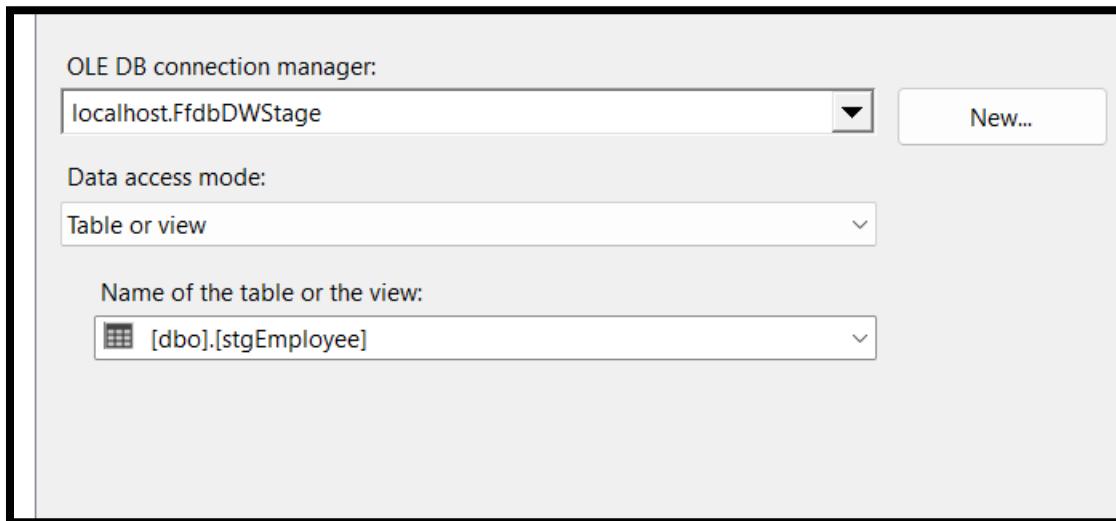
- Từ SSIS Toolbox, kéo thả Data Flow Task, đặt tên L(Load) + Name(table)



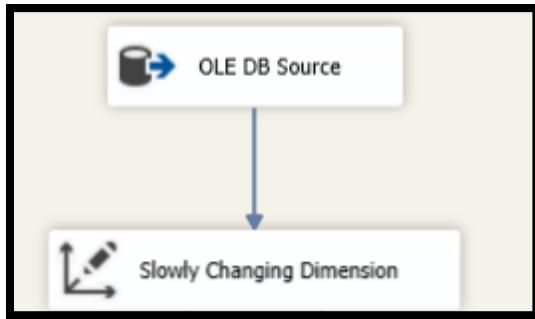
- Tại Data Flow, kéo thả Source Assistant, chọn SQL localhost.ffdbDWStage



- Tại OLE DB Source, chọn table Employee



- Kéo thả Slowly Changing Dimension, kết nối với Source



- Connection Manager, chọn ffdbDW, table or view, chọn DimAddress

Connection manager: localhost.FfdbDW

Table or view: [dbo].[DimAddress]

Input Columns	Dimension Columns	Key Type
city	City	Not a key column
country	Country	Not a key column
street	Street	Not a key column
zip	Zip	Business key

Next

Fixed Attribute
Select this type when the value in a column should not change. Changes are treated as errors.

Changing Attribute
Select this type when changed values should overwrite existing values. This is a Type 1 change.

Historical Attribute
Select this type when changes in column values are saved in new

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
City	Changing a...
Country	Changing a...
Street	Changing a...

Next

Fixed attributes

Fail the transformation if changes are detected in a fixed attribute

Changing attributes

Change all the matching records, including outdated records, when changes are detected in a changing attribute

Next

When data for the inferred member is loaded, you can update the existing record rather than create a new one.

Enable inferred member support

All columns with a change type are null

Use a Boolean column to indicate whether the current record is an inferred member

Inferred member indicator:

Next

The following outputs will be configured:

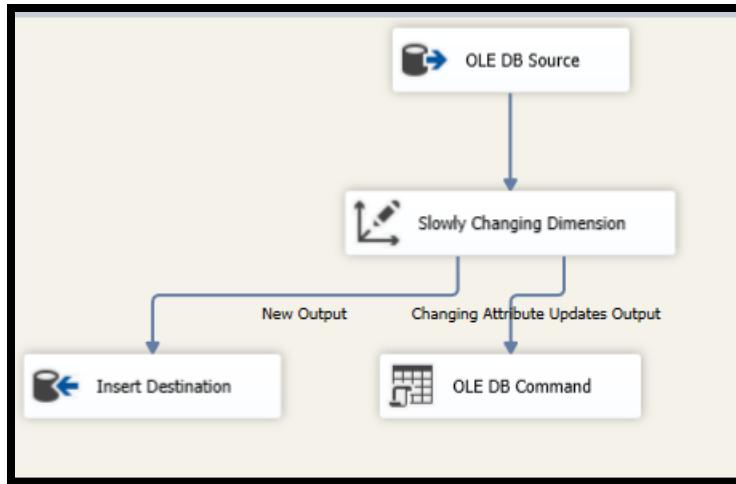
Slowly Changing Dimension Transformation

New Records

 New Record Output

 Other Outputs

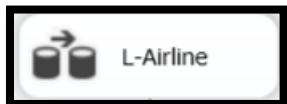
Finish.



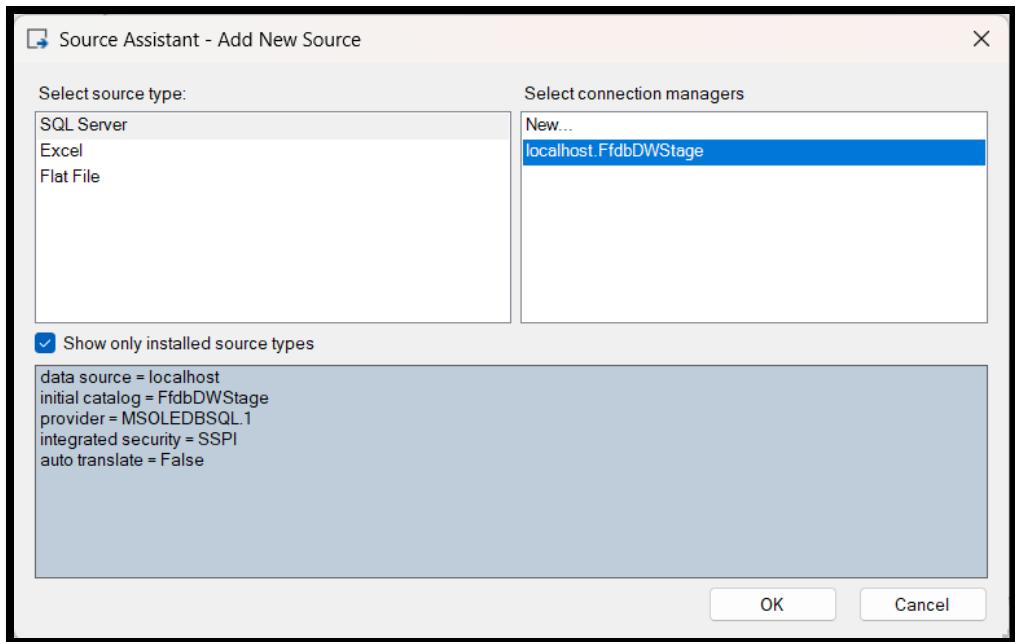
F5 để run.

➤ L – Address

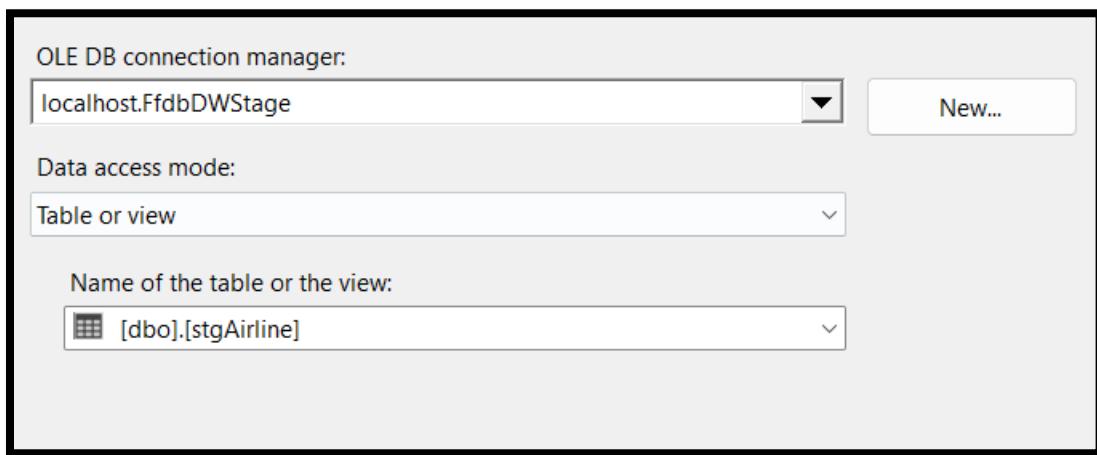
- Từ SSIS Toolbox, kéo thả Data Flow Task, đặt tên L(Load) + Name(table)



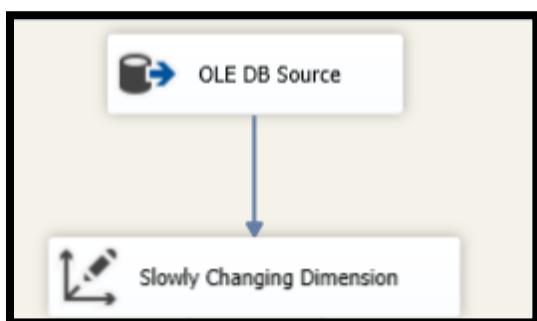
- Tại Data Flow, kéo thả Source Assistant, chọn SQL localhost.ffdbDWStage



- Tại OLE DB Source, chọn table Employee



- Kéo thả Slowly Changing Dimension, kết nối với Source



- Connection Manager, chọn ffdbDW, table or view, chọn DimAirline

Connection manager:

localhost.FfdbDW

Table or view:

[dbo].[DimAirline]

Input Columns	Dimension Columns	Key Type
airline_id	Airlineld	Business key
airlinename	AirlineName	Not a key column
base_airport	BaseAirline	Not a key column
iata	IATA	Not a key column

Next

Fixed Attribute

Select this type when the value in a column should not change. Changes are treated as errors.

Changing Attribute

Select this type when changed values should overwrite existing values. This is a Type 1 change.

Historical Attribute

Select this type when changes in column values are saved in new

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
AirlineName	Changing a...
BaseAirline	Changing a...
IATA	Changing a...

Next

Fixed attributes

Fail the transformation if changes are detected in a fixed attribute

Changing attributes

Change all the matching records, including outdated records, when changes are detected in a changing attribute

Next

When data for the inferred member is loaded, you can update the existing record rather than create a new one.

Enable inferred member support

All columns with a change type are null

Use a Boolean column to indicate whether the current record is an inferred member

Inferred member indicator:



Next

The following outputs will be configured:

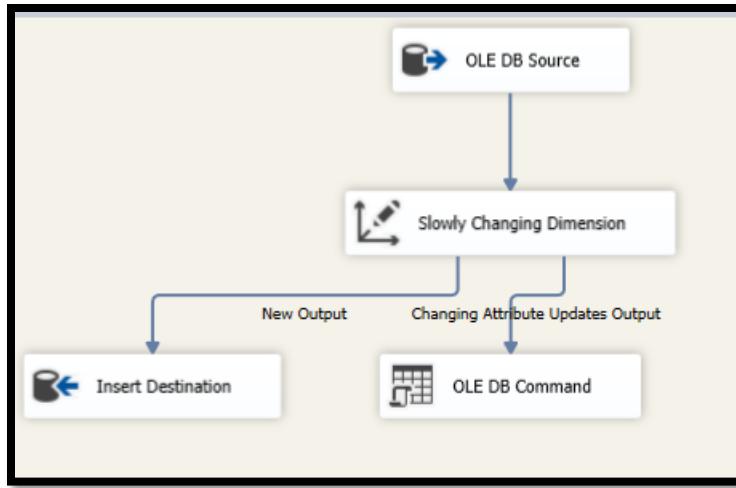
Slowly Changing Dimension Transformation

New Records

New Record Output

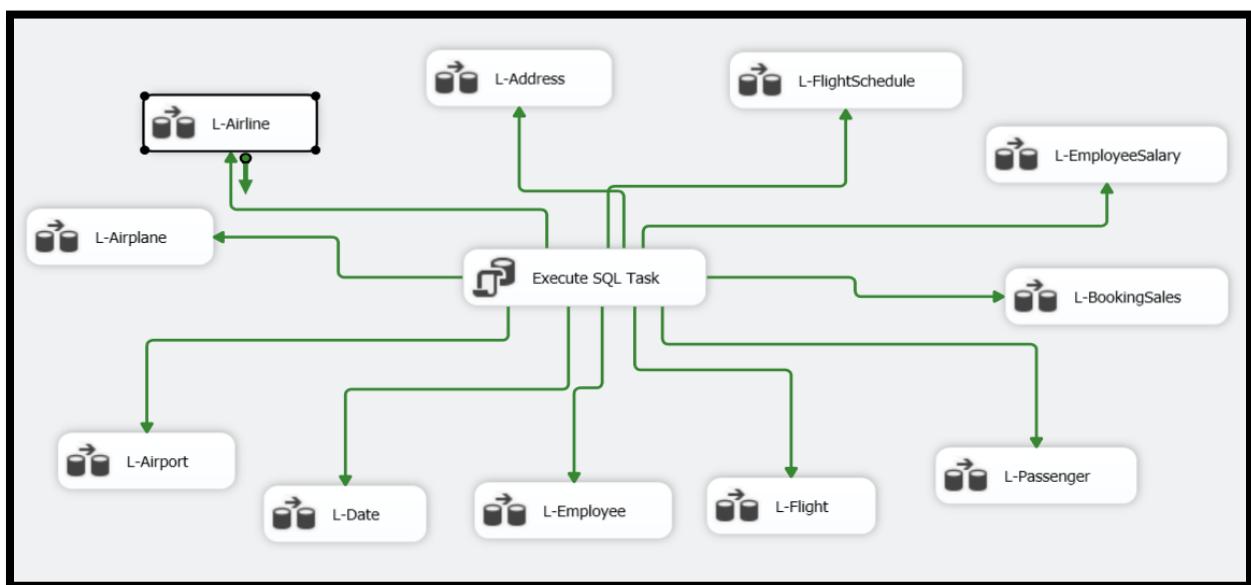
Other Outputs

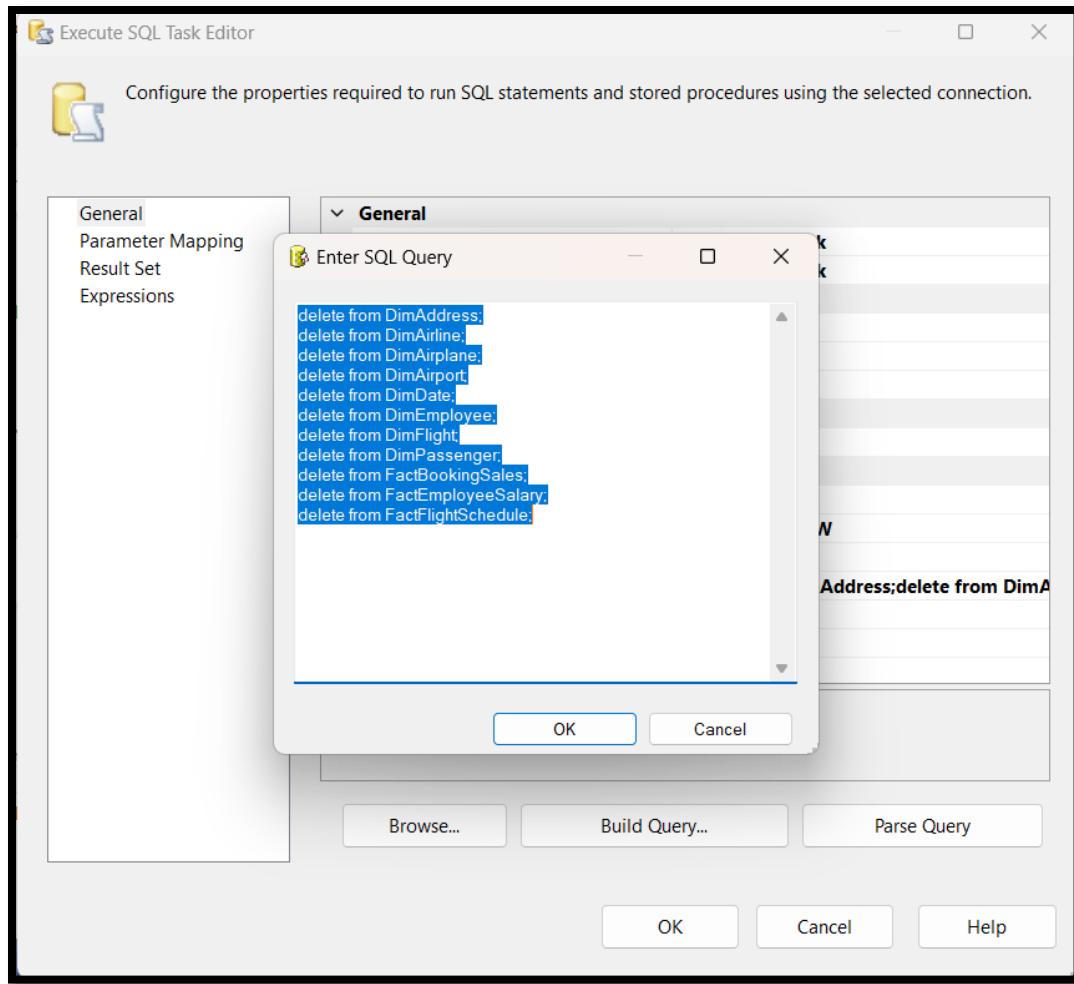
Finish.



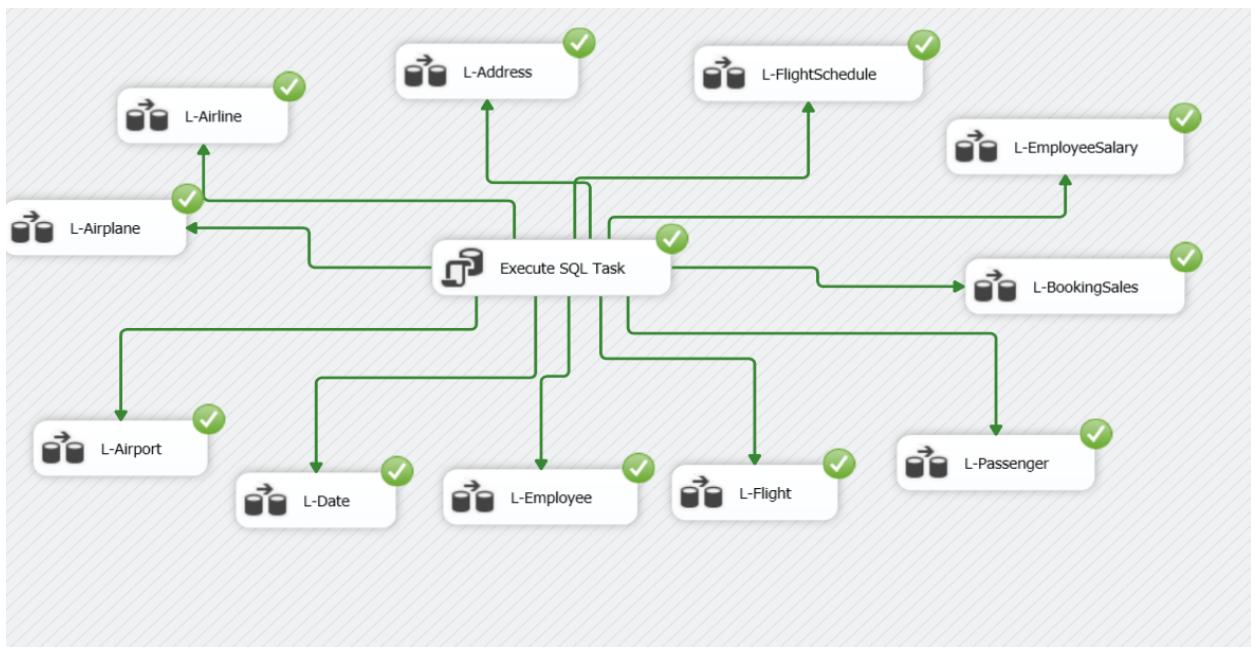
F5 để run.

- ⇒ Thao tác tương tự với L – Airplane, L – Airport, L – Date, L – Employee, L – Flight, L – Passenger, L – BookingSales, L – EmployeeSalary, L – FlightSchedule.
- Kéo thả SQL truncate table, kết nối với tất cả Data Flow Task





Hoàn thành quá trình L(Load) dữ liệu vào database ffdbDW:



CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)

Sử dụng công cụ SSAS xây dựng mô hình, xây dựng khối cube, phân tích dữ liệu trên cube. Sử dụng pivot table trong Excel, ngôn ngữ truy vấn MDX và tiền hành đổi chiều kết quả.

4.1. Danh sách truy vấn

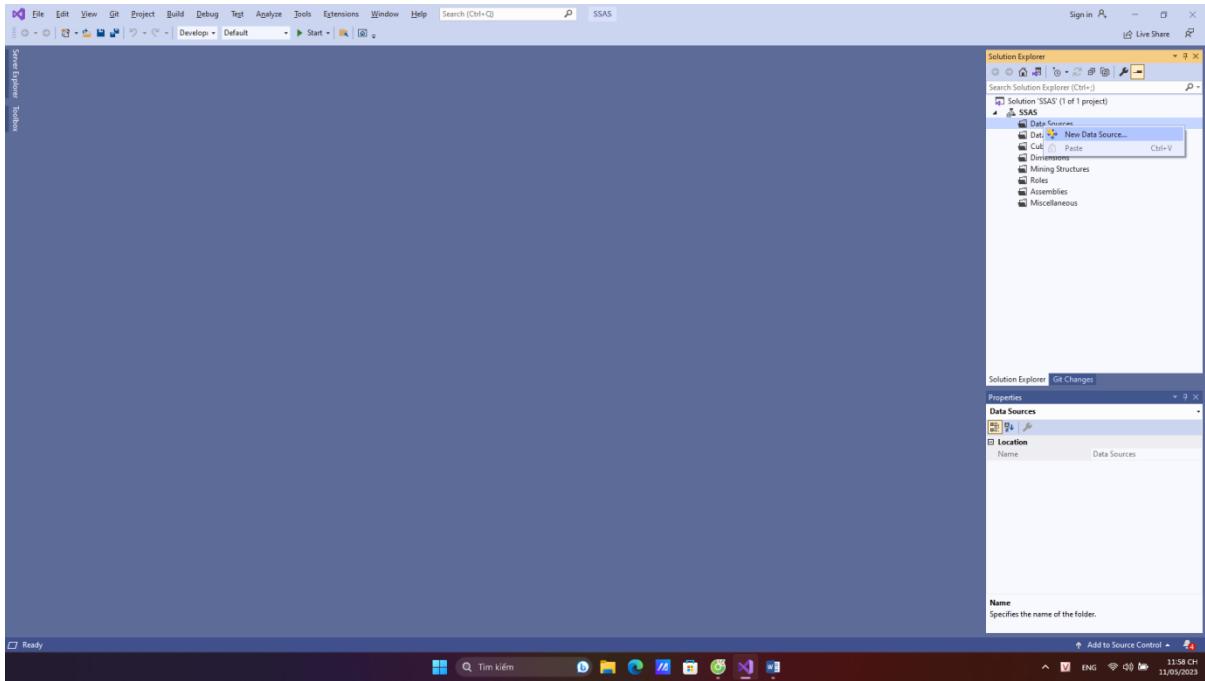
❖ Sử dụng SSAS, Pivot Table và DMX

- Thống kê số lượng vé và tổng giá vé trong năm 2015
- Thống kê số lượng vé theo nơi đến (quốc gia)
- Thống kê số lượng hành khách theo quốc gia
- Thống kê số lượng vé đã được đặt của các hãng hàng không
- Thống kê số lượng chuyến bay của các hãng hàng không
- Thống kê số lượng chuyến bay các quý trong năm 2015
- Thống kê tổng và trung bình lương của nhân viên theo quốc gia
- Thống kê trung bình lương của nhân viên theo phòng ban

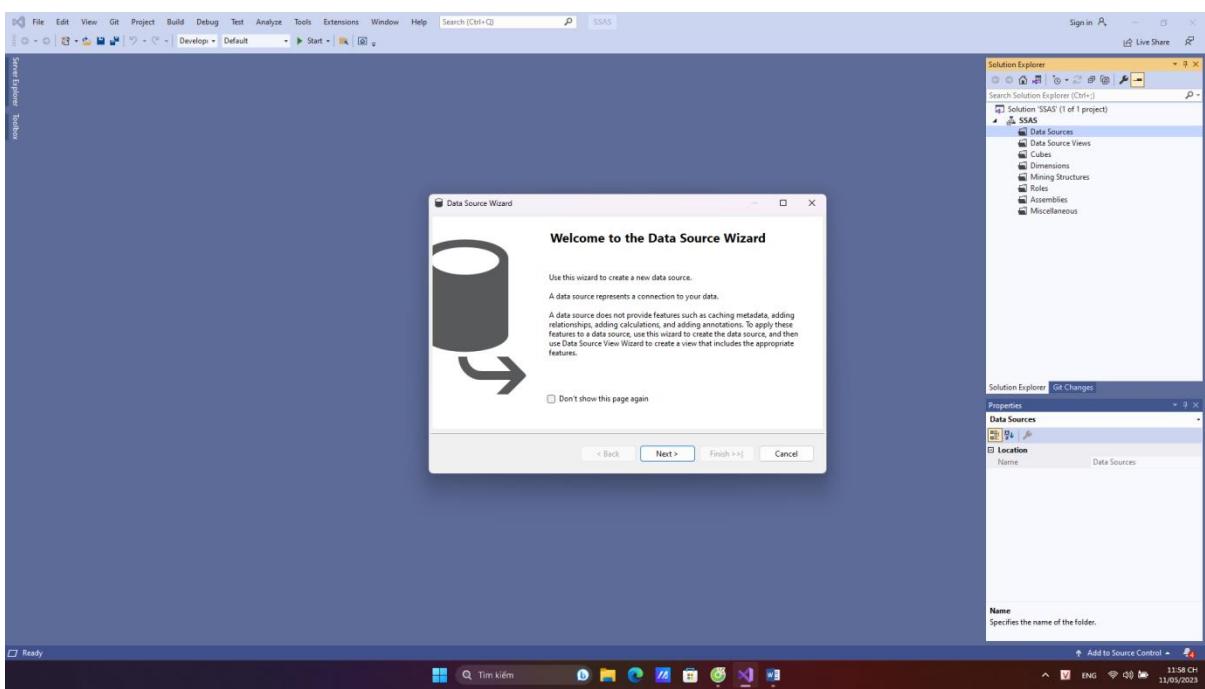
4.2. Xây dựng mô hình

4.2.1 Data Source

Chọn New Data Source

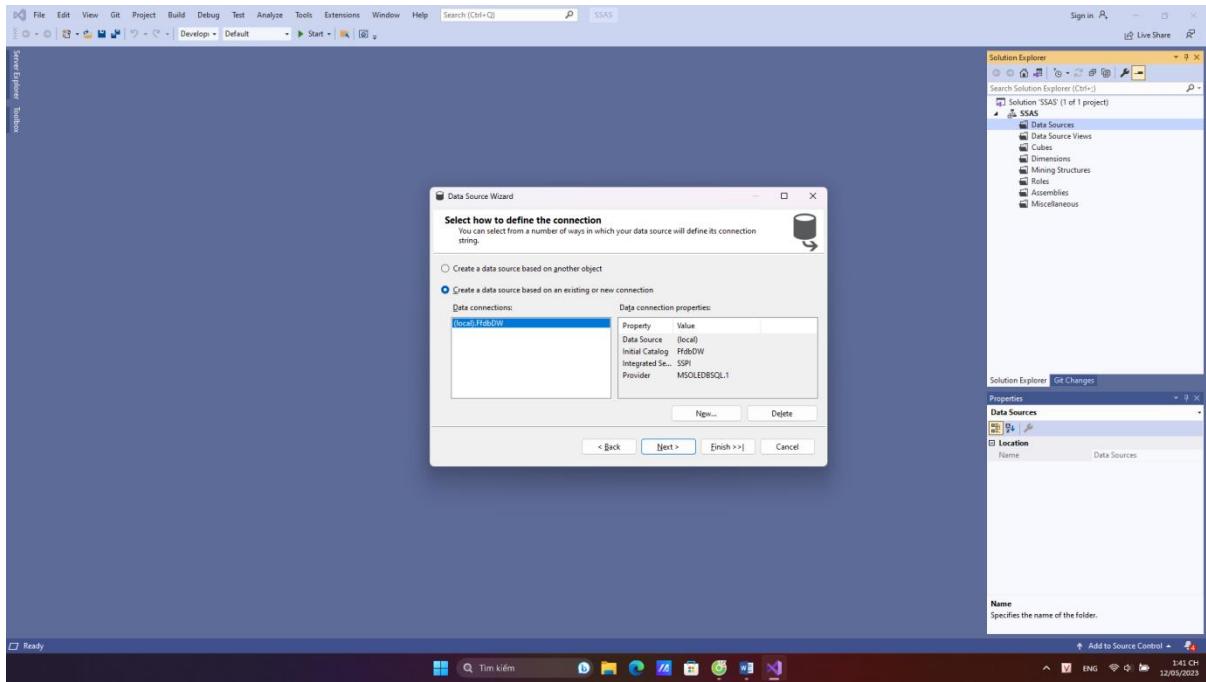


Nhấn Next

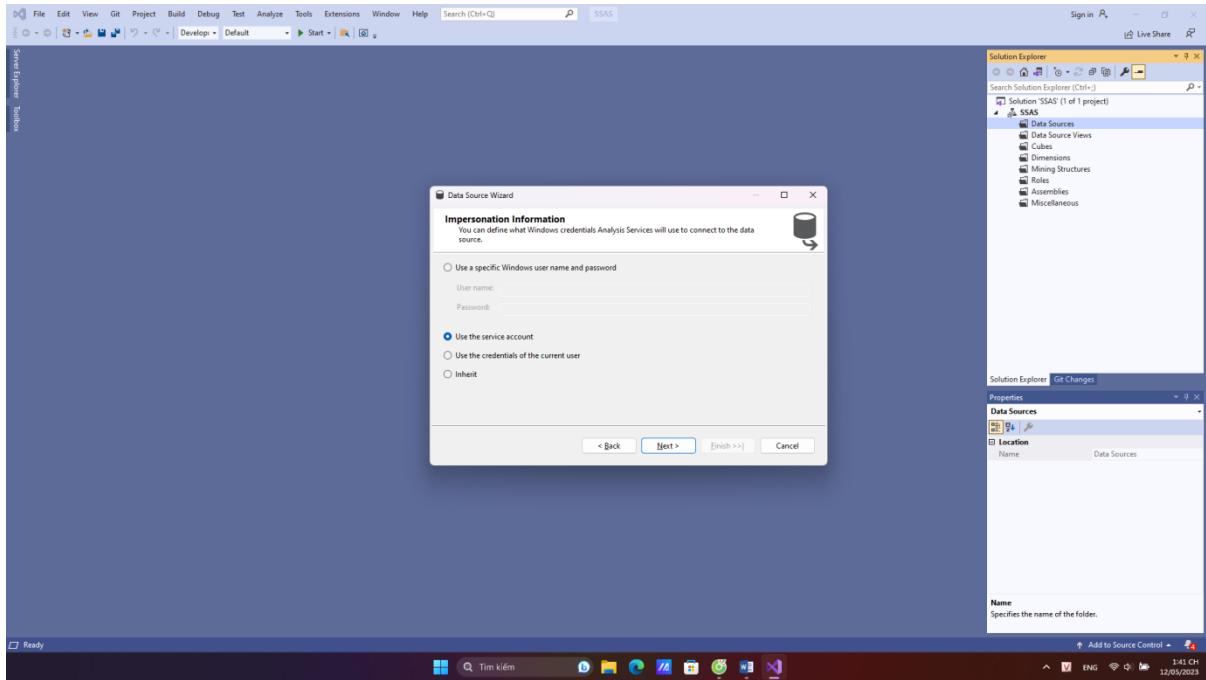


Chọn Create a data source based on an existing or new connection -> Chọn
New để tạo connection

Nhấn Next



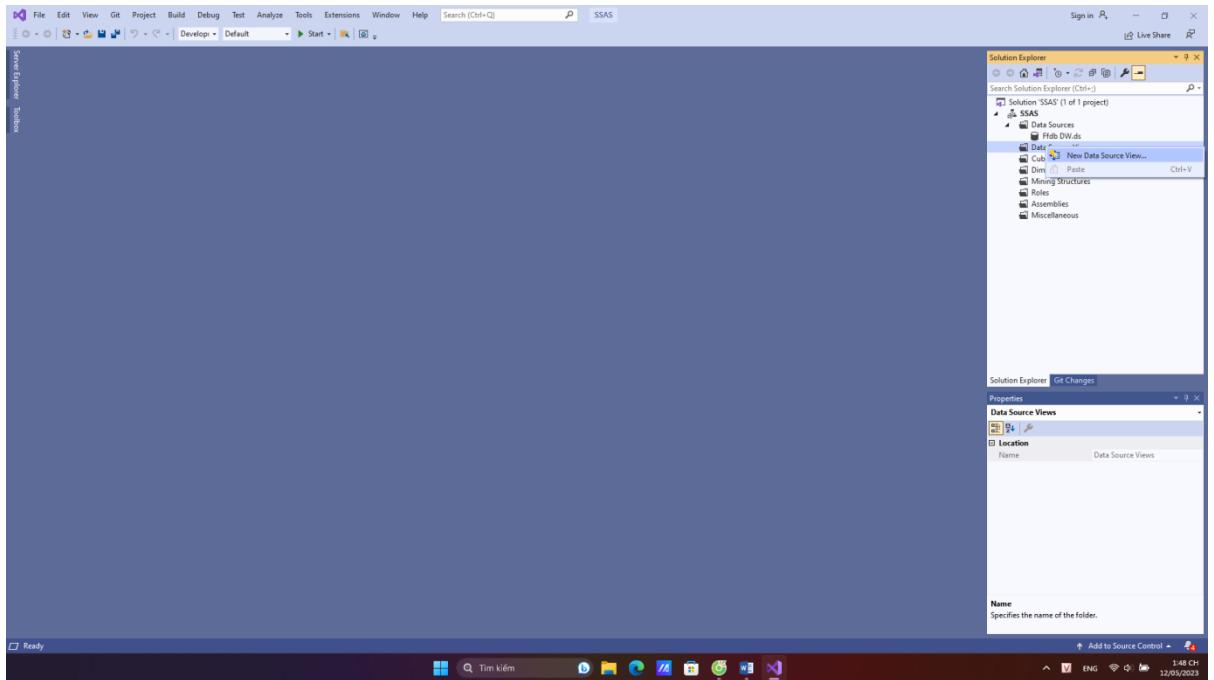
Chọn Use the service account và nhấn Next



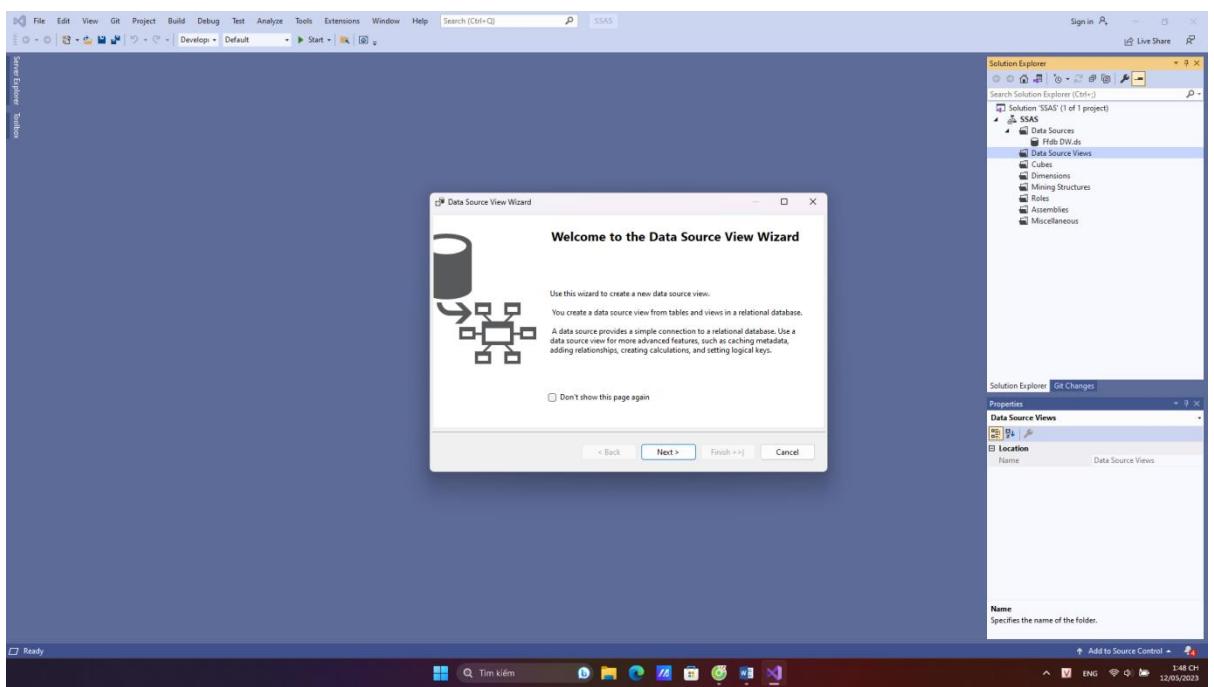
Nhấn Finish để kết thúc quá trình tạo Data Source

4.2.2 Data Source View

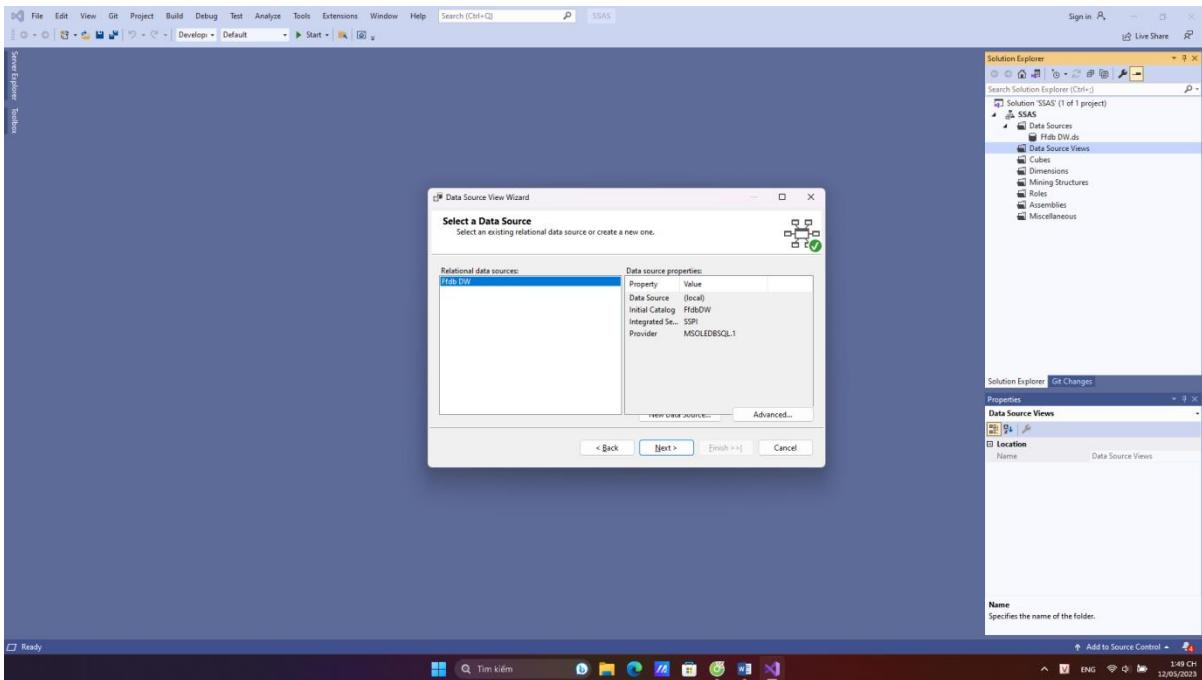
Chọn New Data Source View



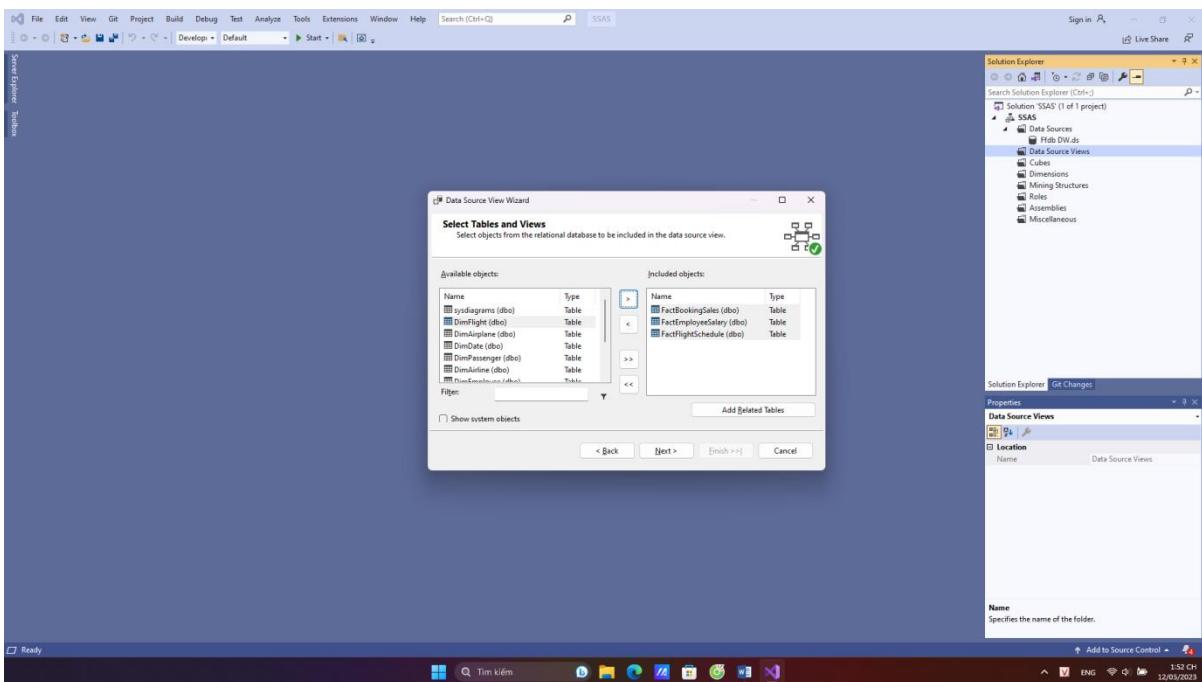
Nhấn Next



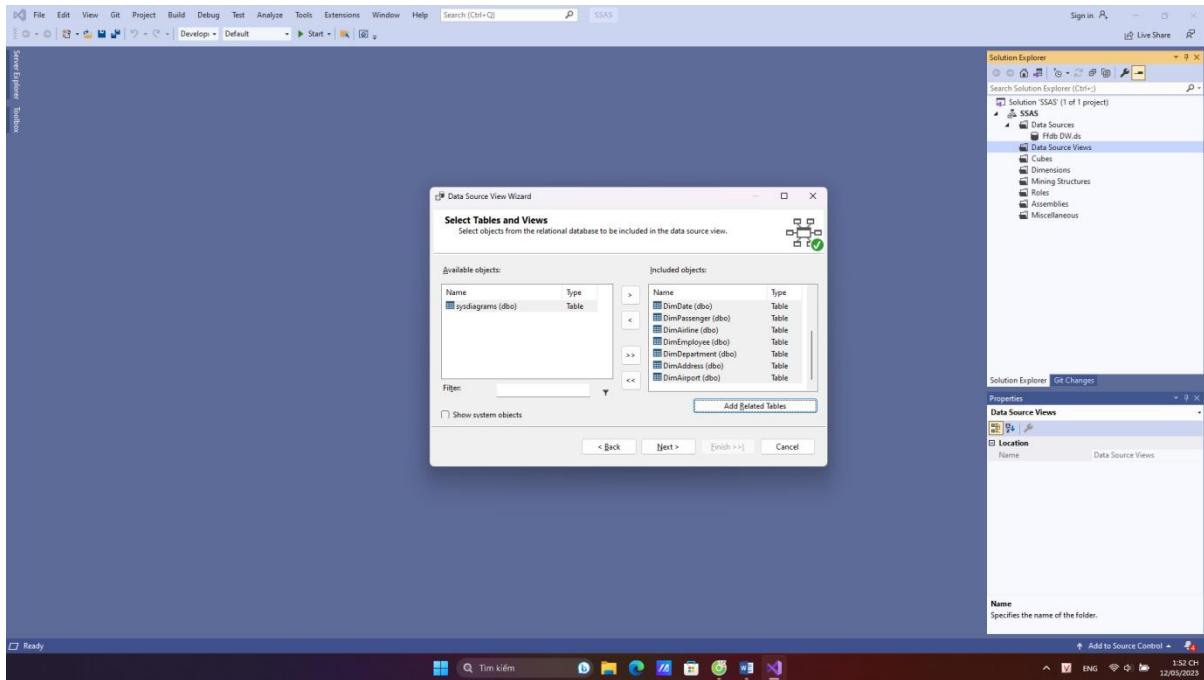
Chọn Data Source đã tạo ở trên -> Nhấn Next



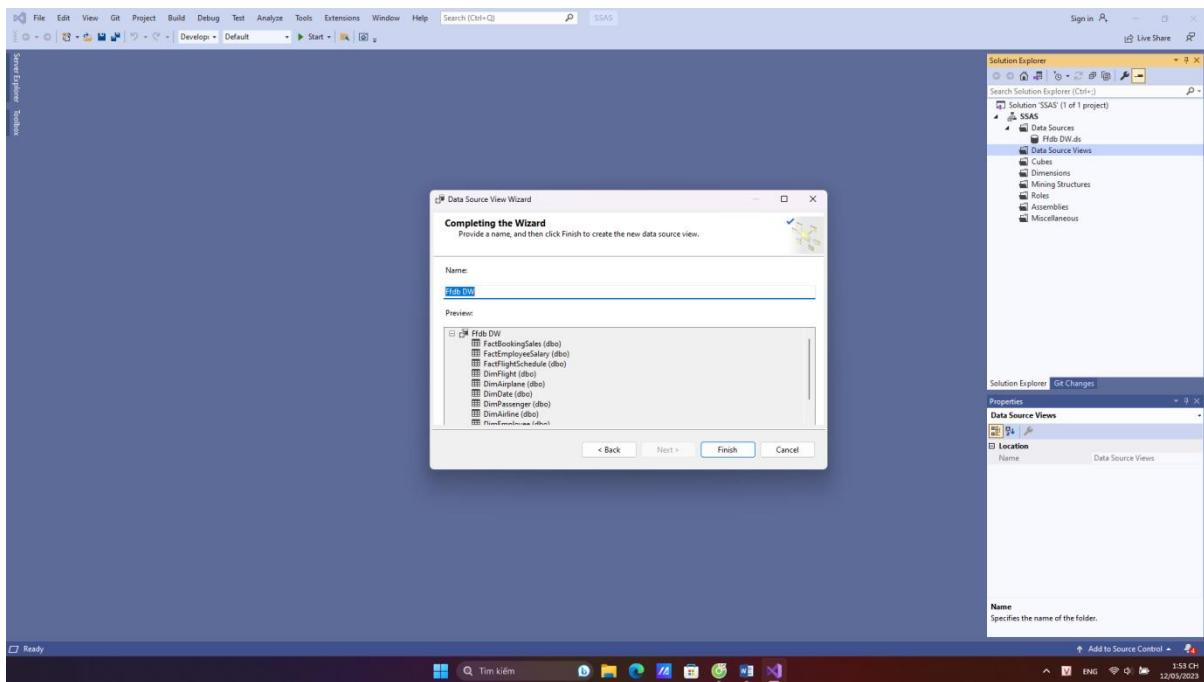
Chọn các bảng fact



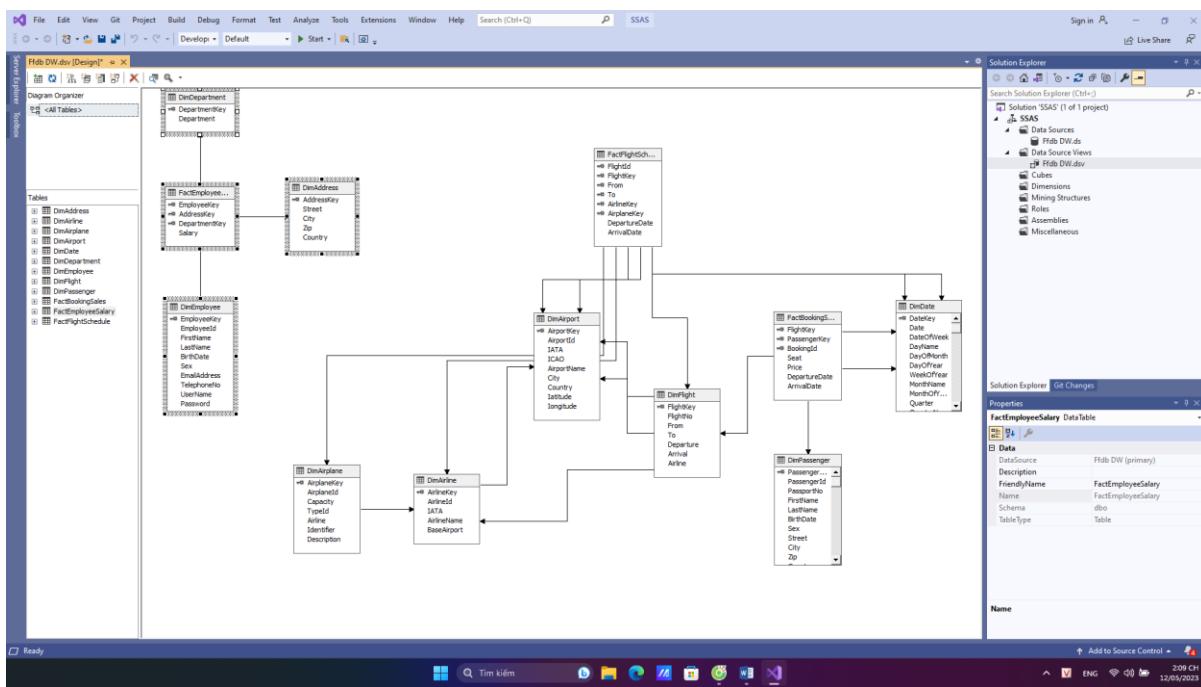
Nhấn Add Related Tables để thêm các bảng Dim có liên quan



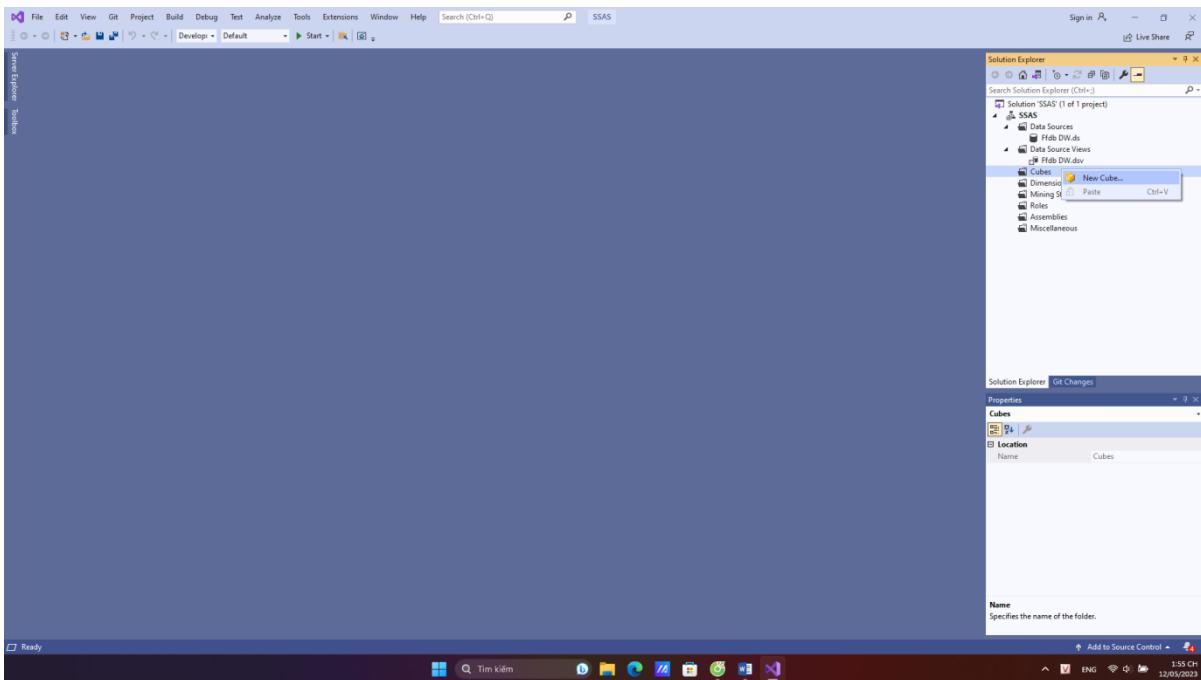
Nhấn Next



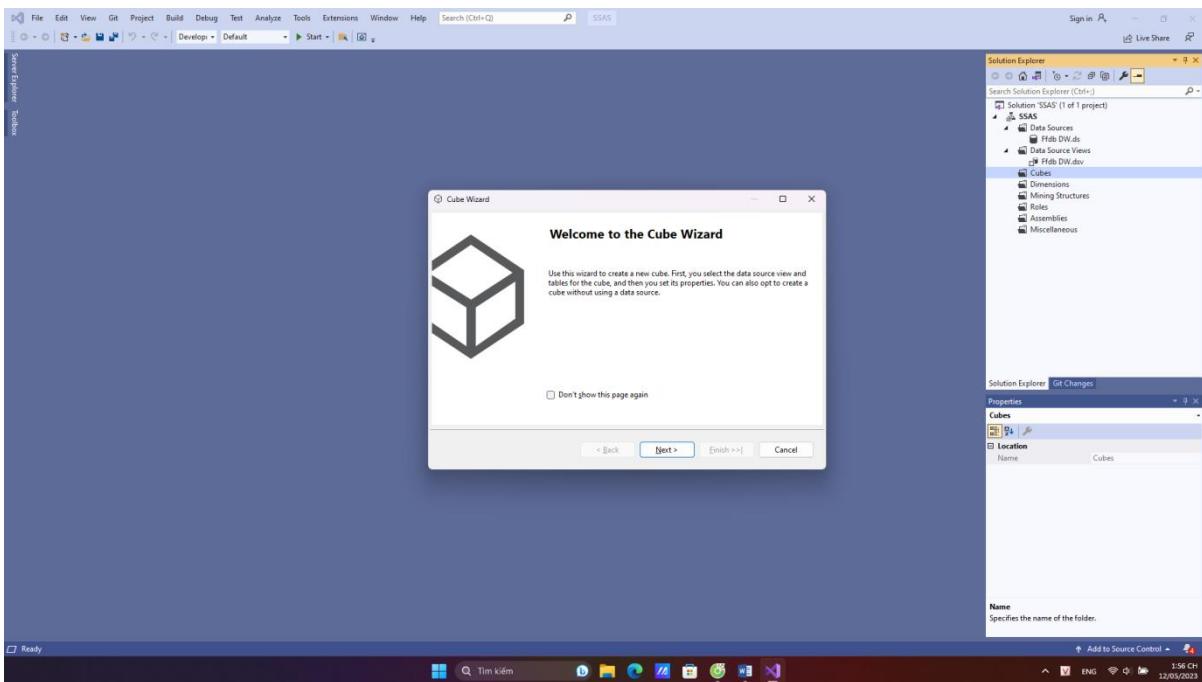
Nhấn Finish để kết thúc quá trình tạo Data Source View



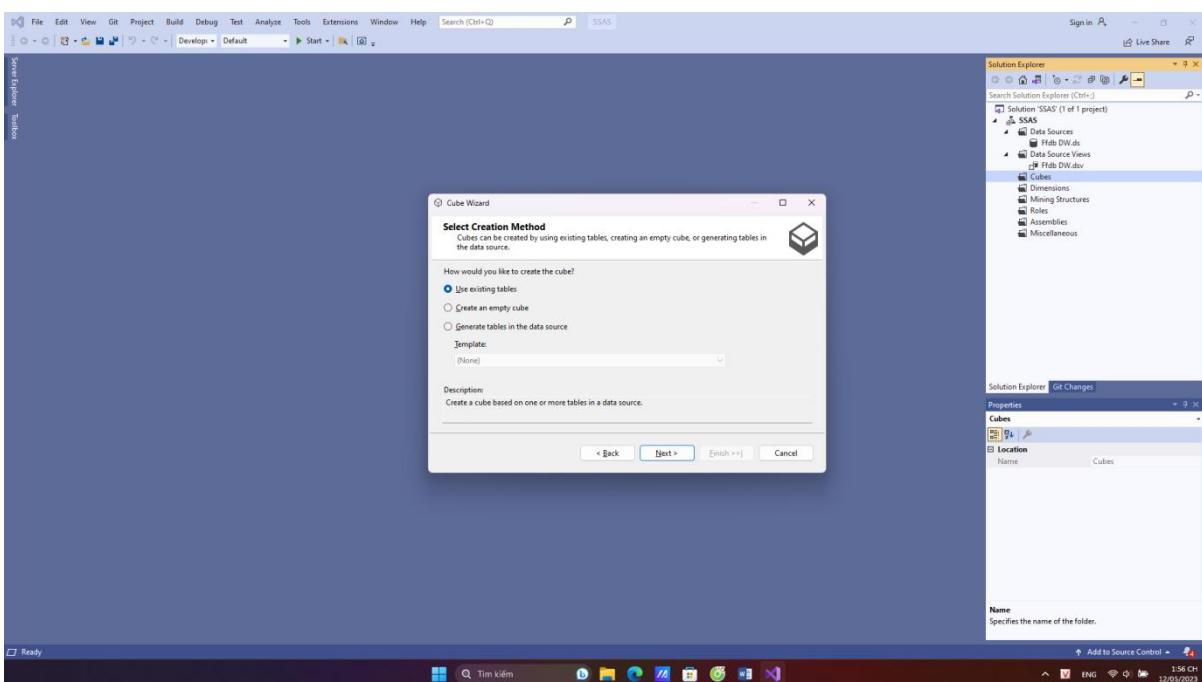
4.3. Xây dựng khối – Cube Chọn New Cube



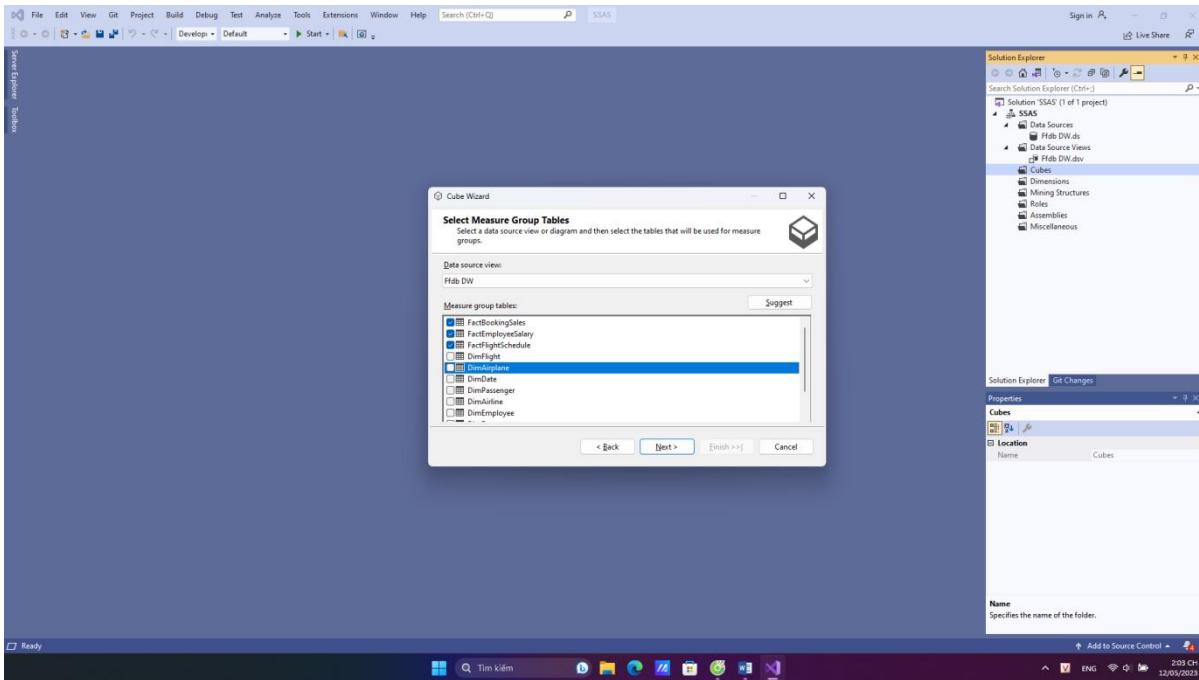
Nhân Next



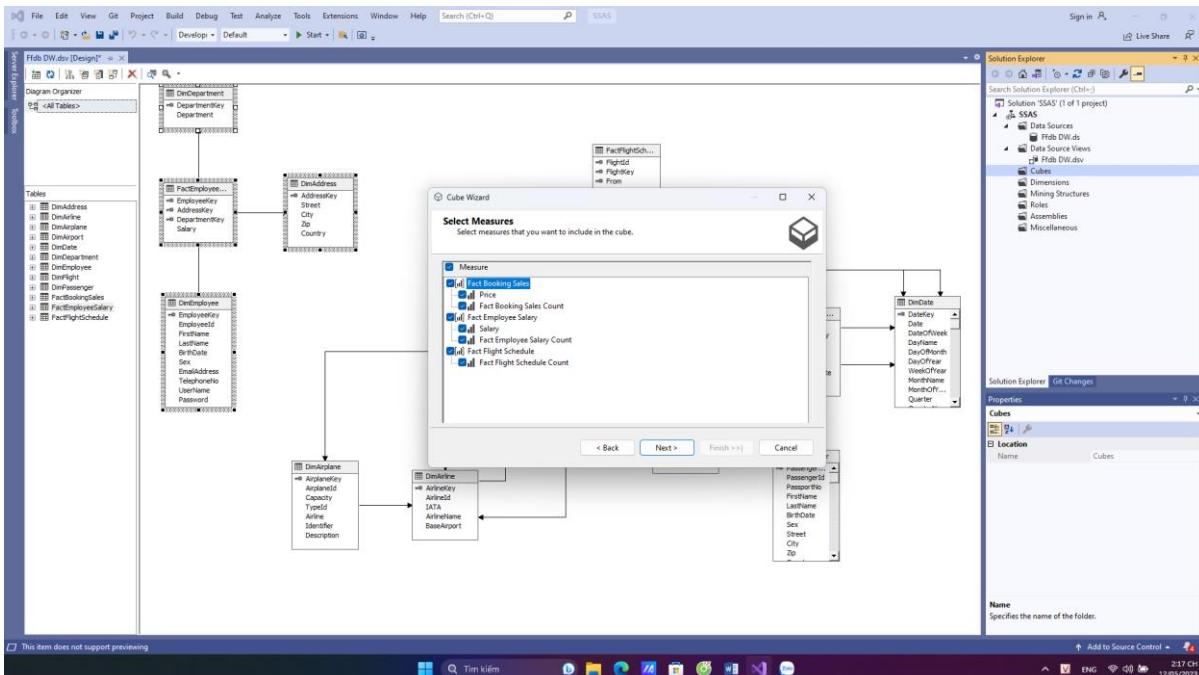
Chọn Use existing tables và nhấn Next



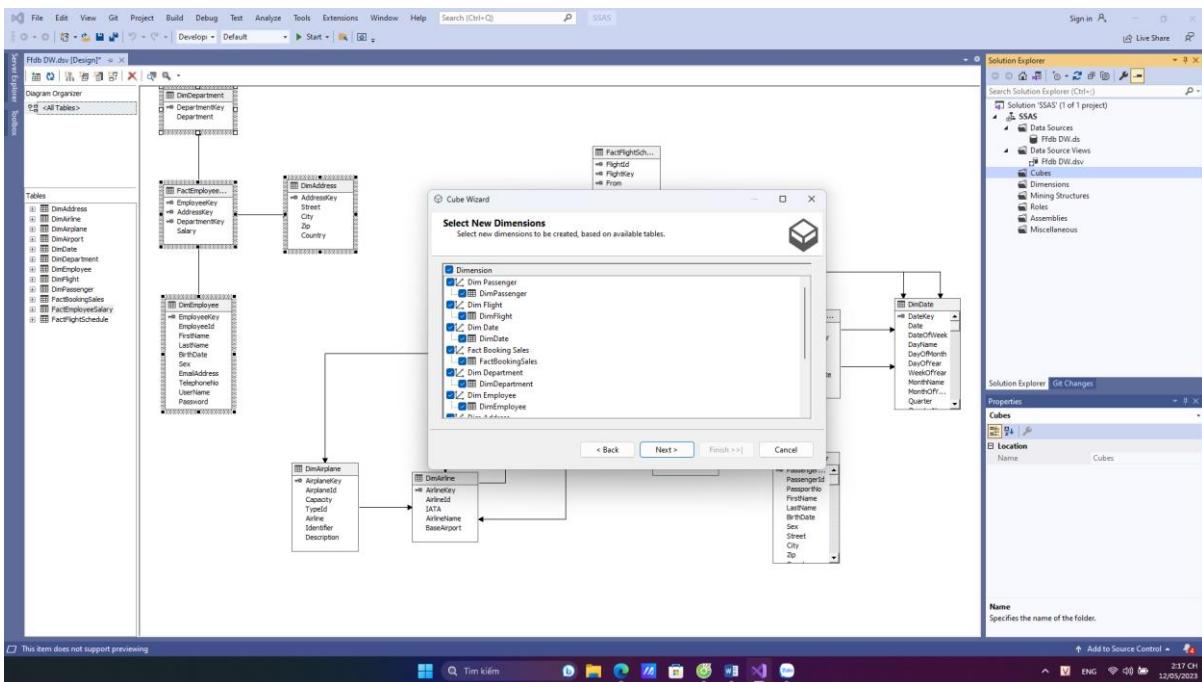
Chọn các bảng Fact -> Nhấn Next



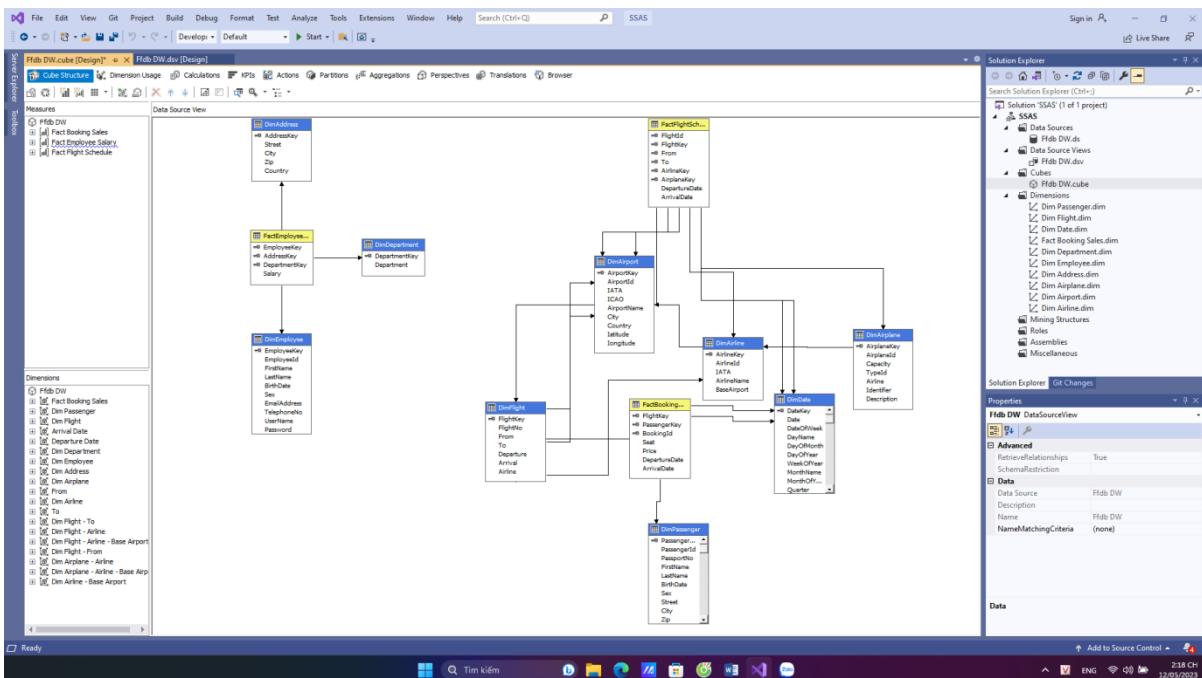
Nhấn Next



Nhấn Next

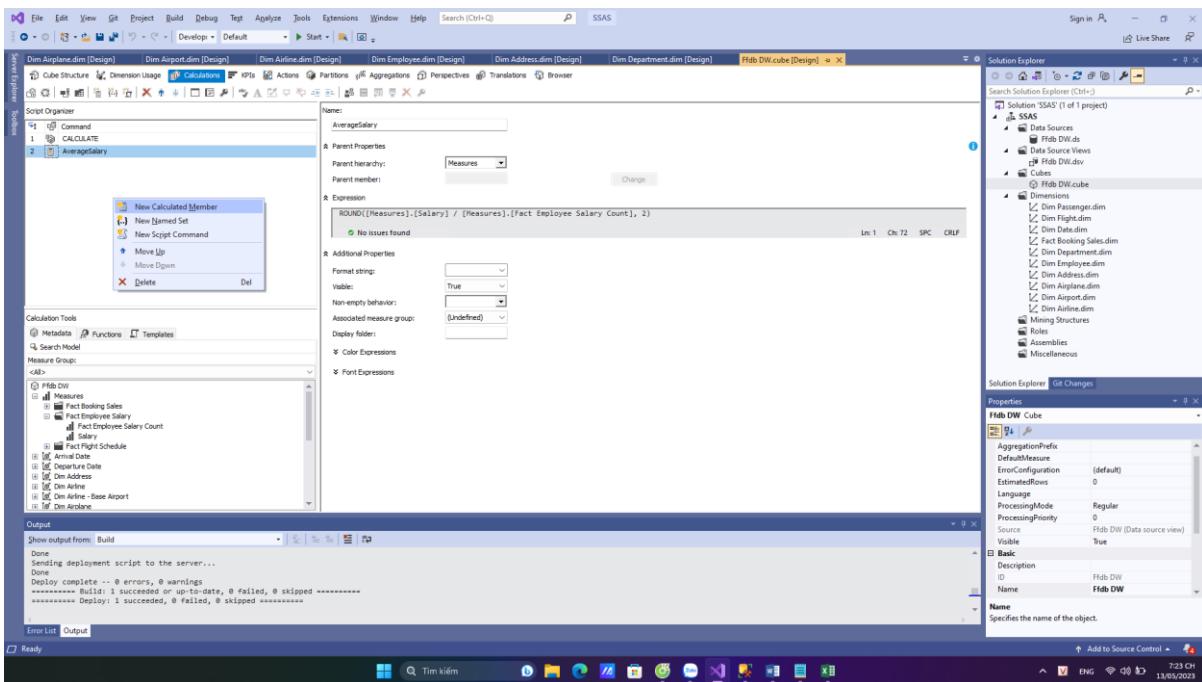


Nhấn Finish để kết thúc quá trình tạo cube

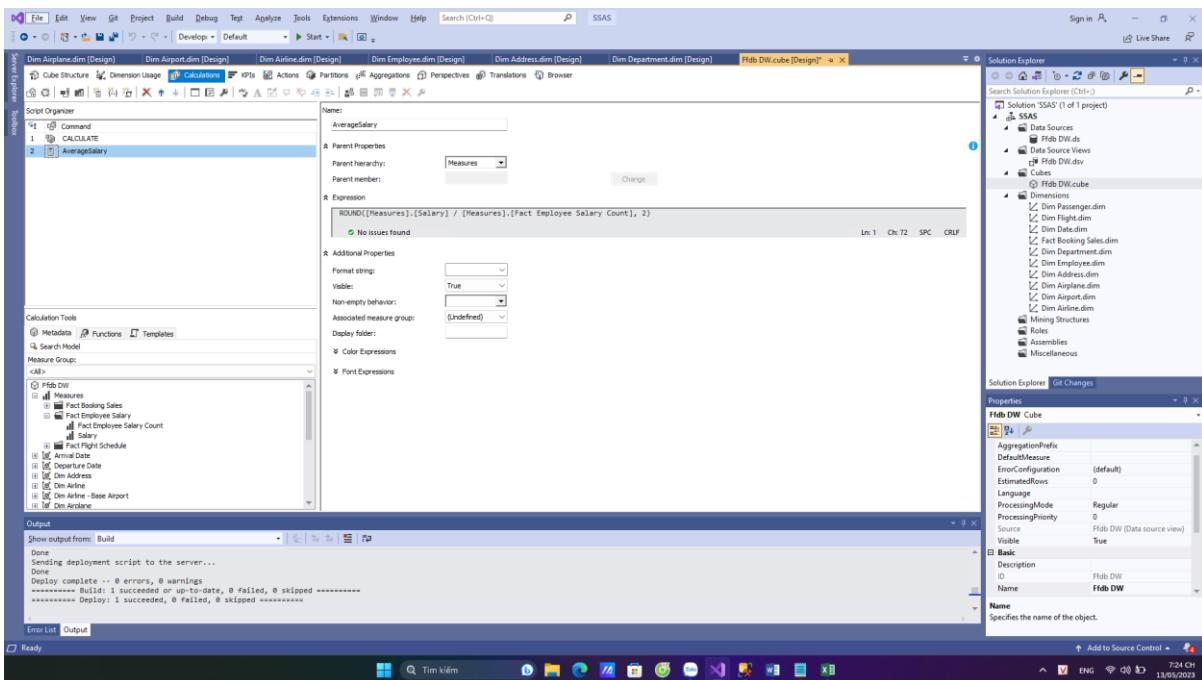


Thêm calculations:

Chọn New Calculated Member

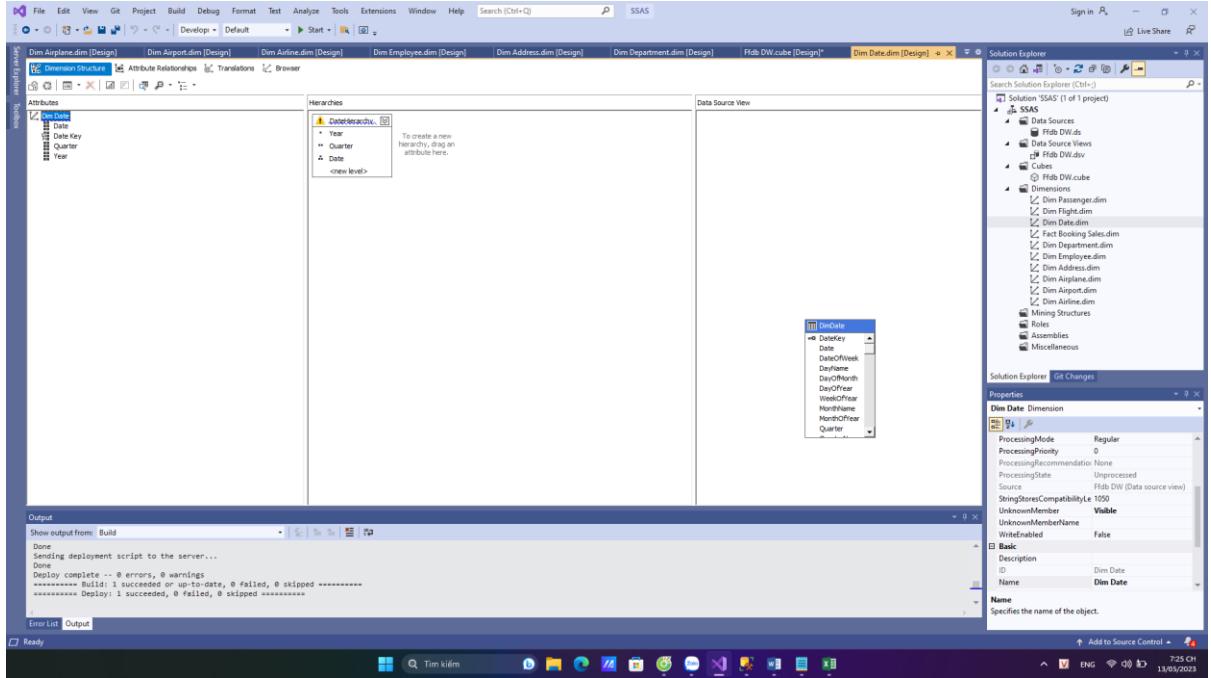


Thêm vào các thuộc tính như sau

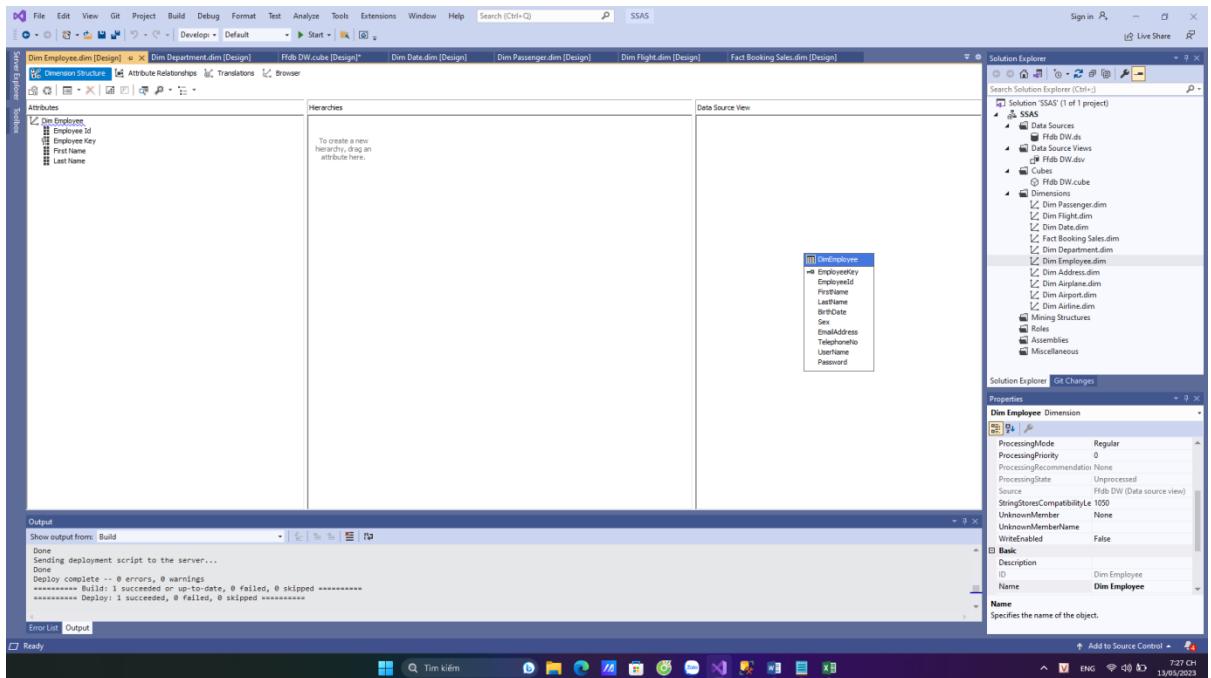


4.4. Cấu hình Dimensions

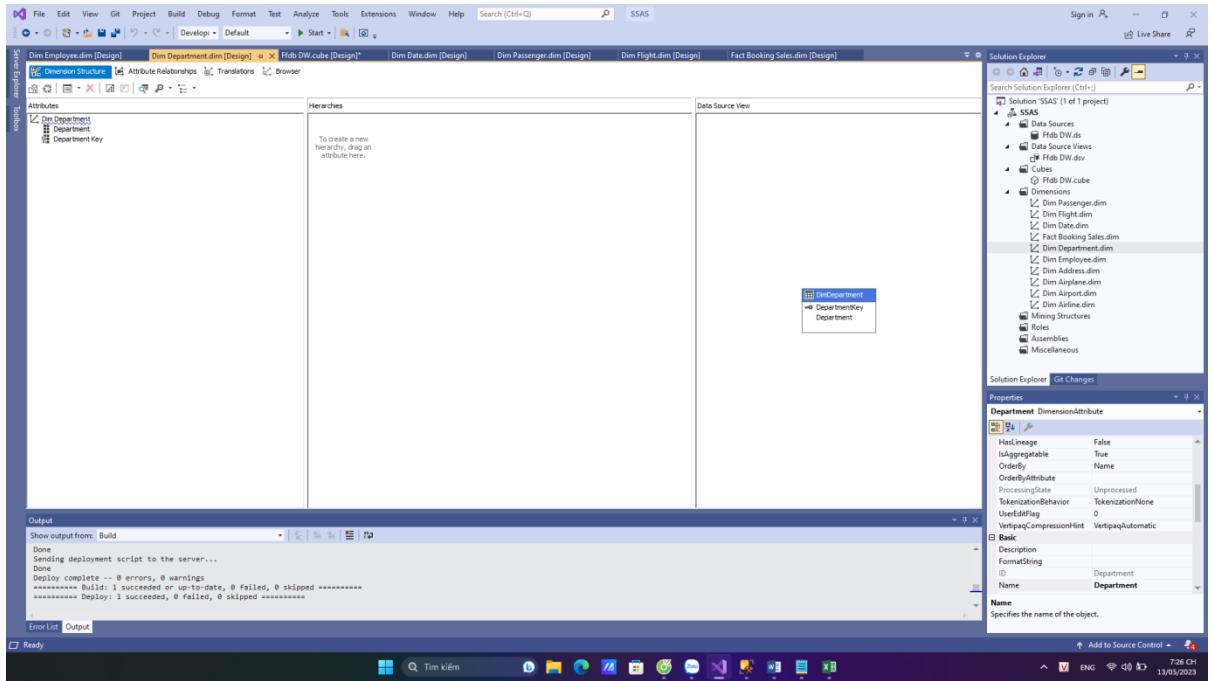
4.4.1 Dim Date



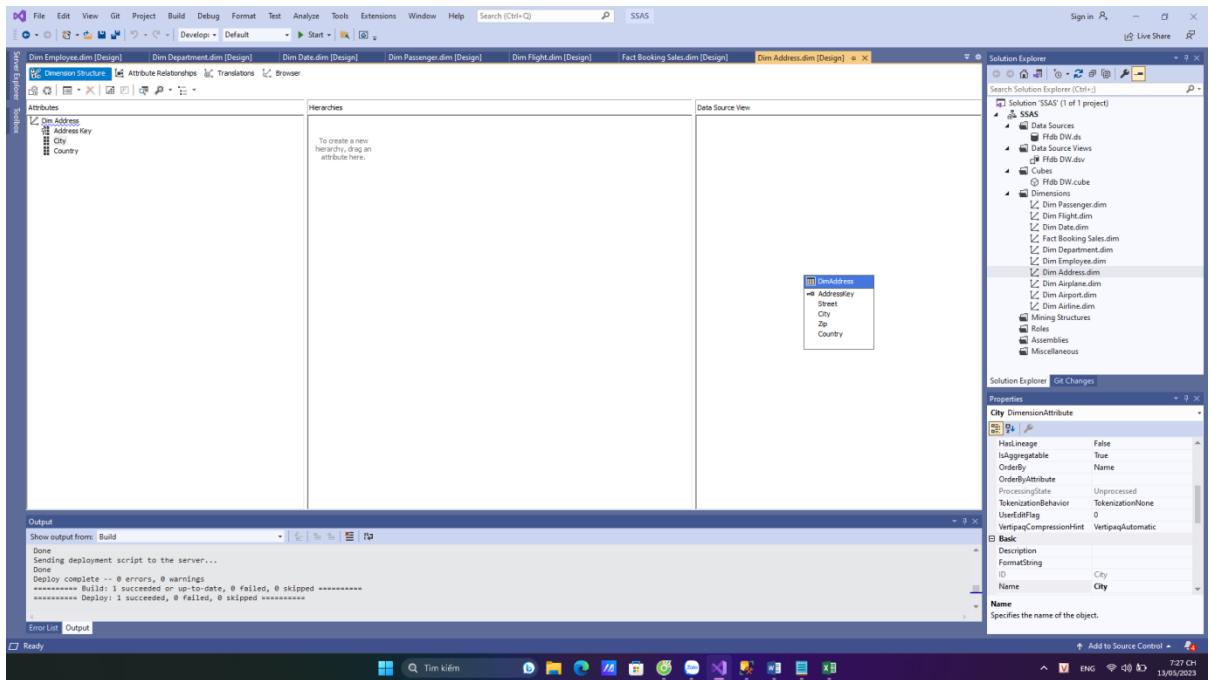
4.4.2 Dim Employee



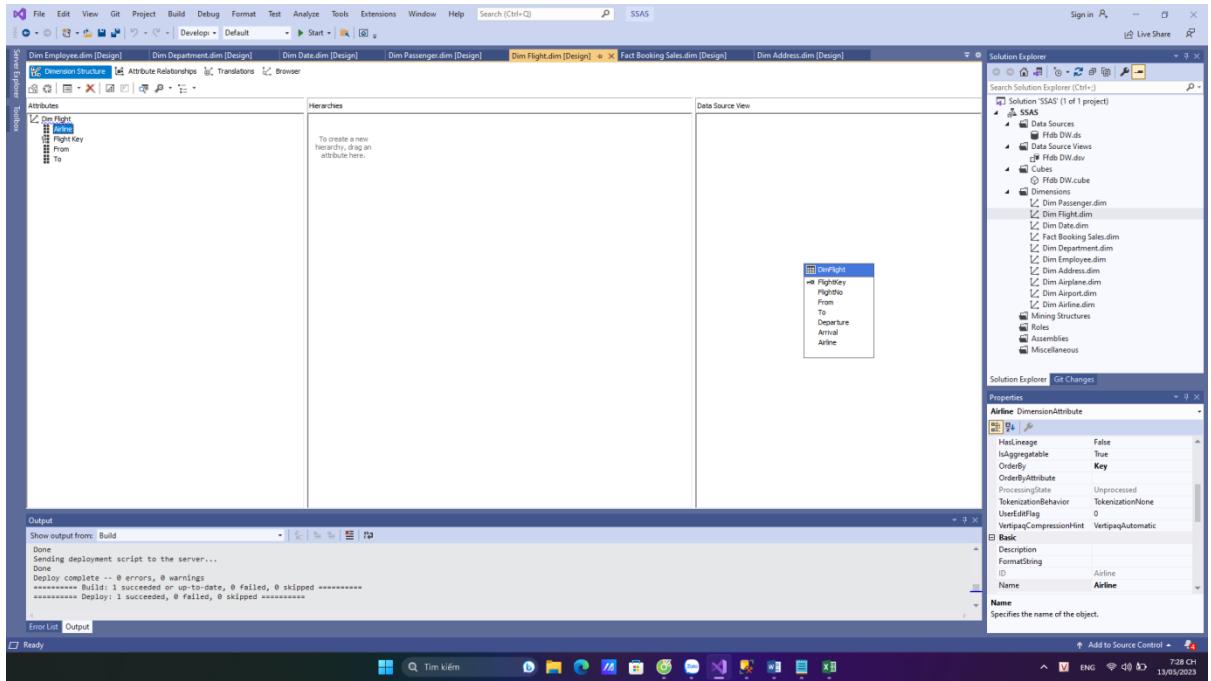
4.4.3 Dim Department



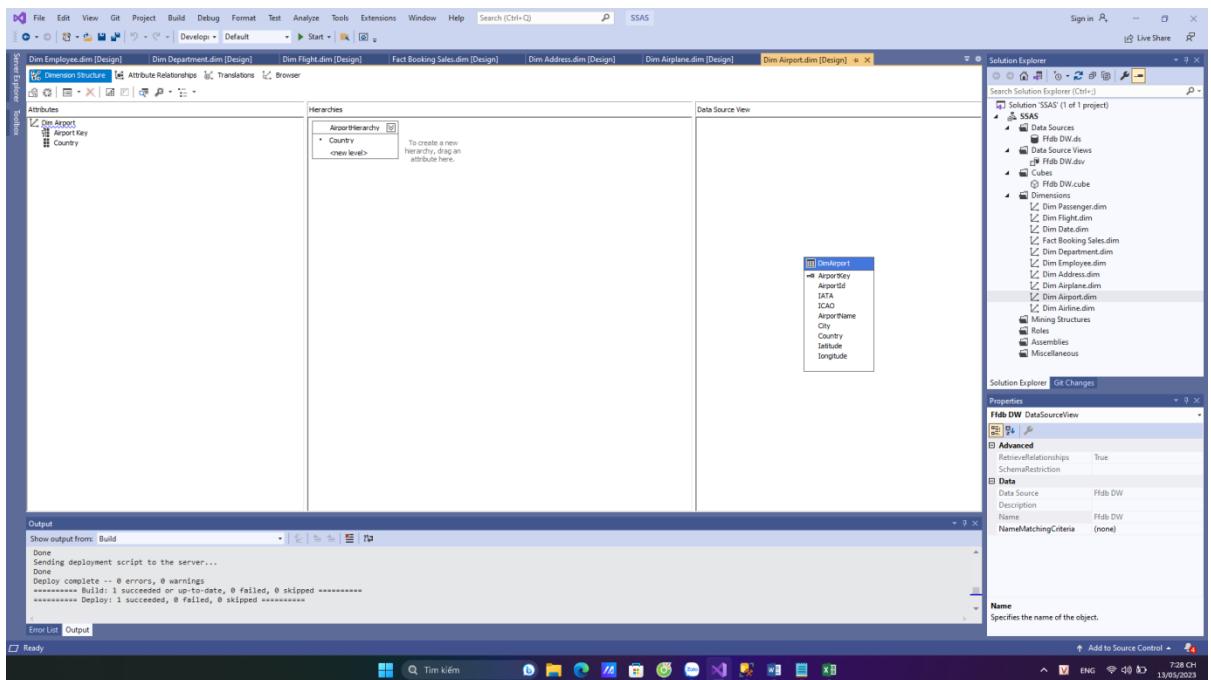
4.4.4 Dim Address



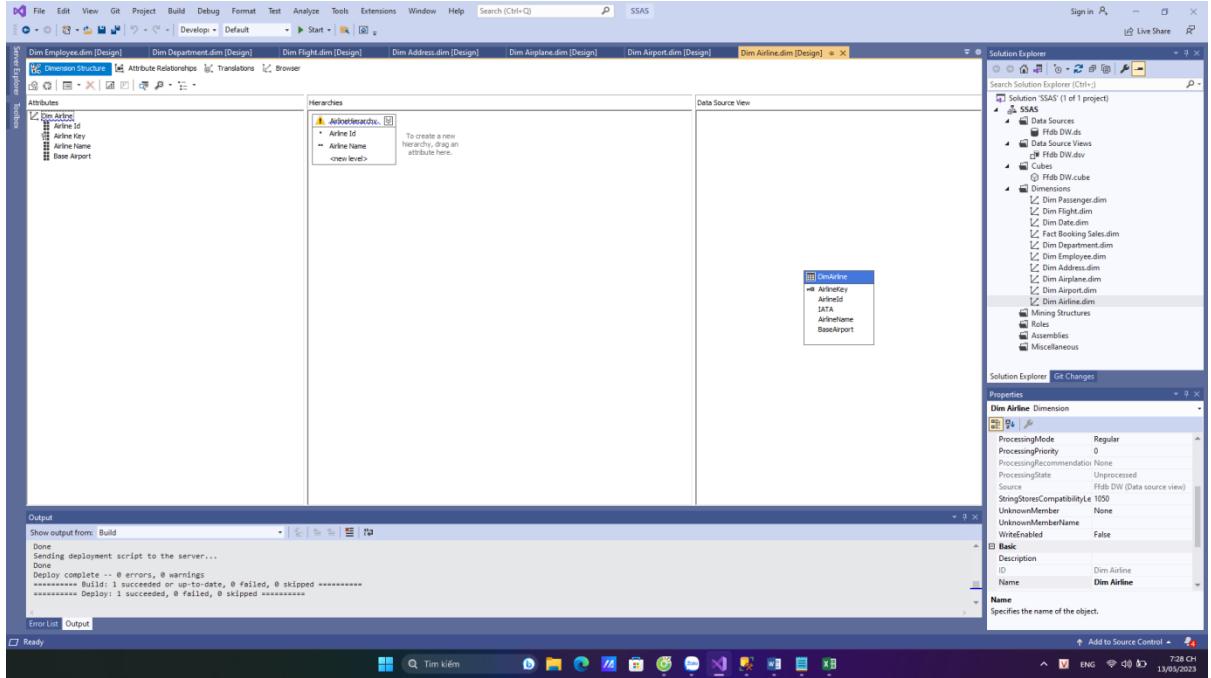
4.4.5 Dim Flight



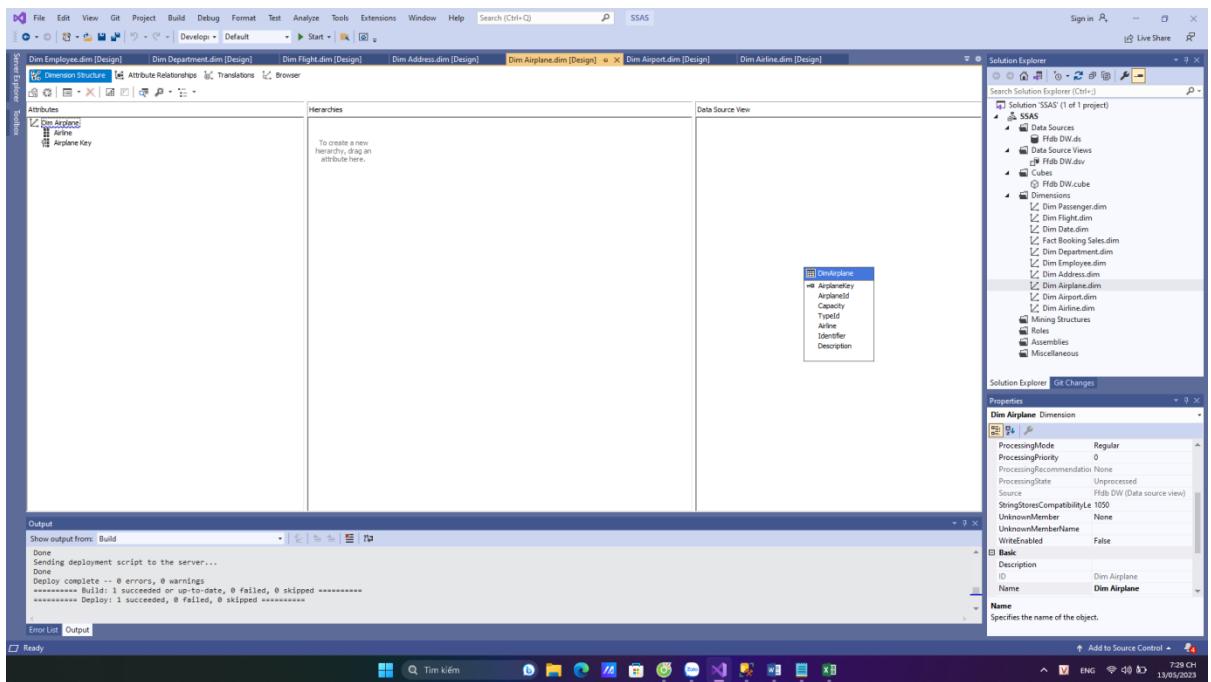
4.4.6 Dim Airport



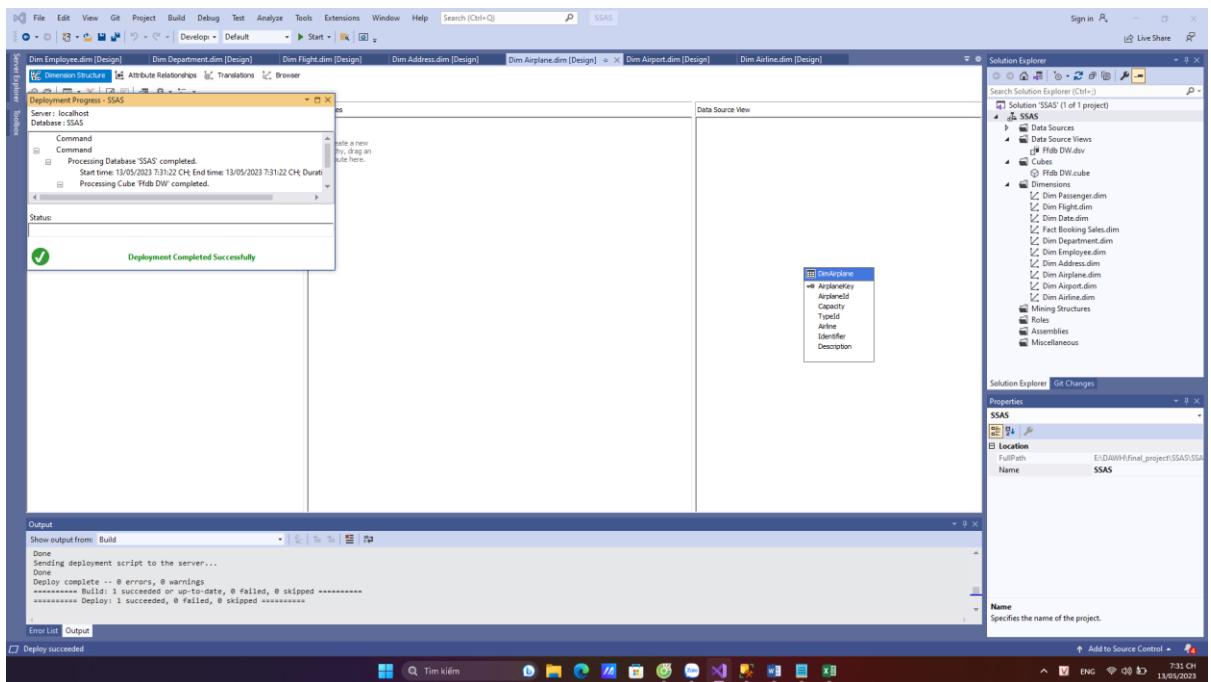
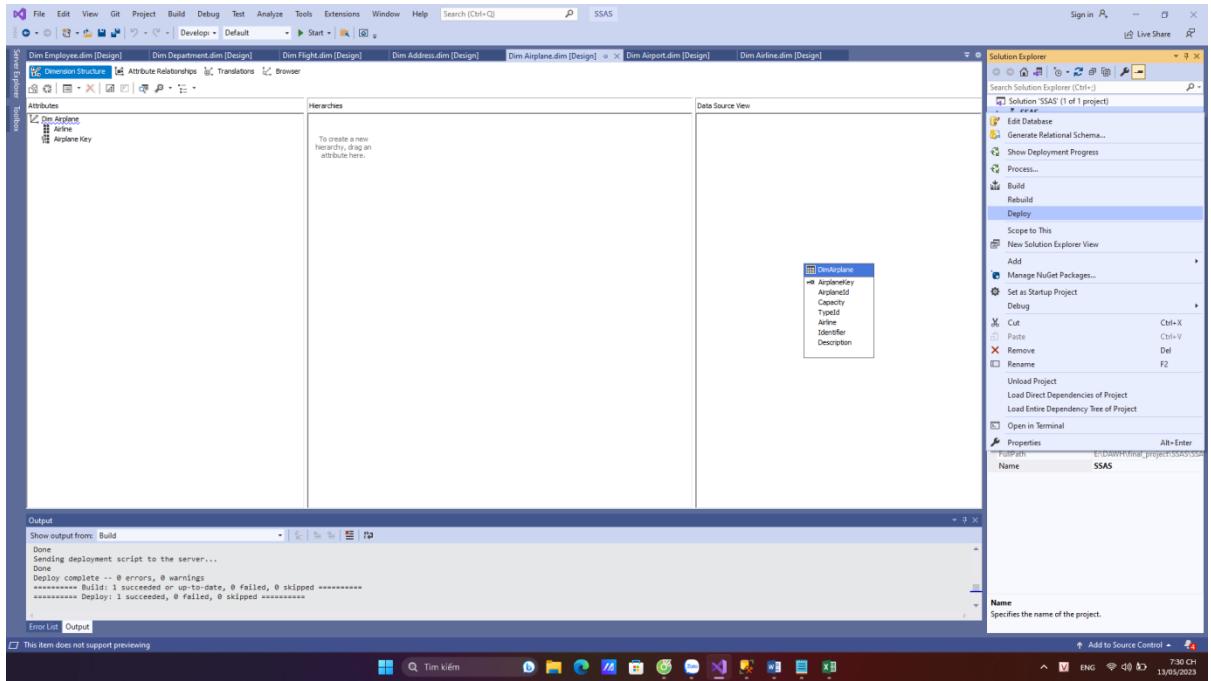
4.4.7 Dim Airline



4.4.8 Dim Airplane



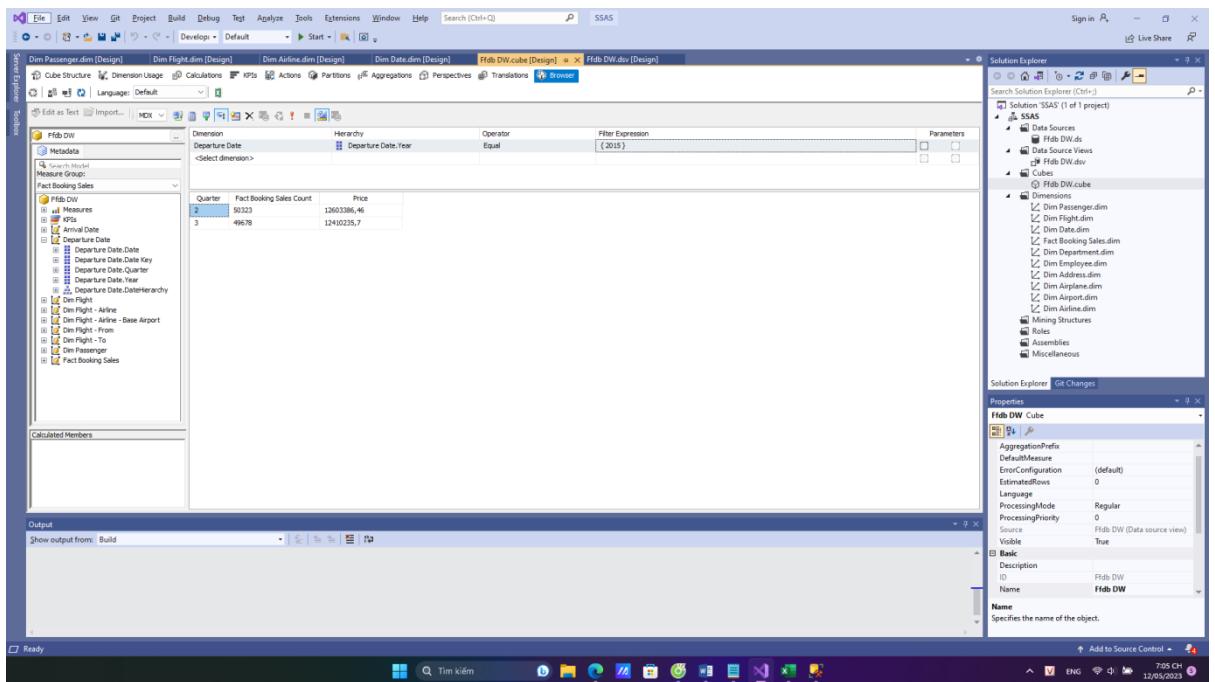
Delay Project:



4.5. Thực hiện truy vấn

4.5.1 Tổng số lượng vé và số tiền vé theo quý trong năm 2015

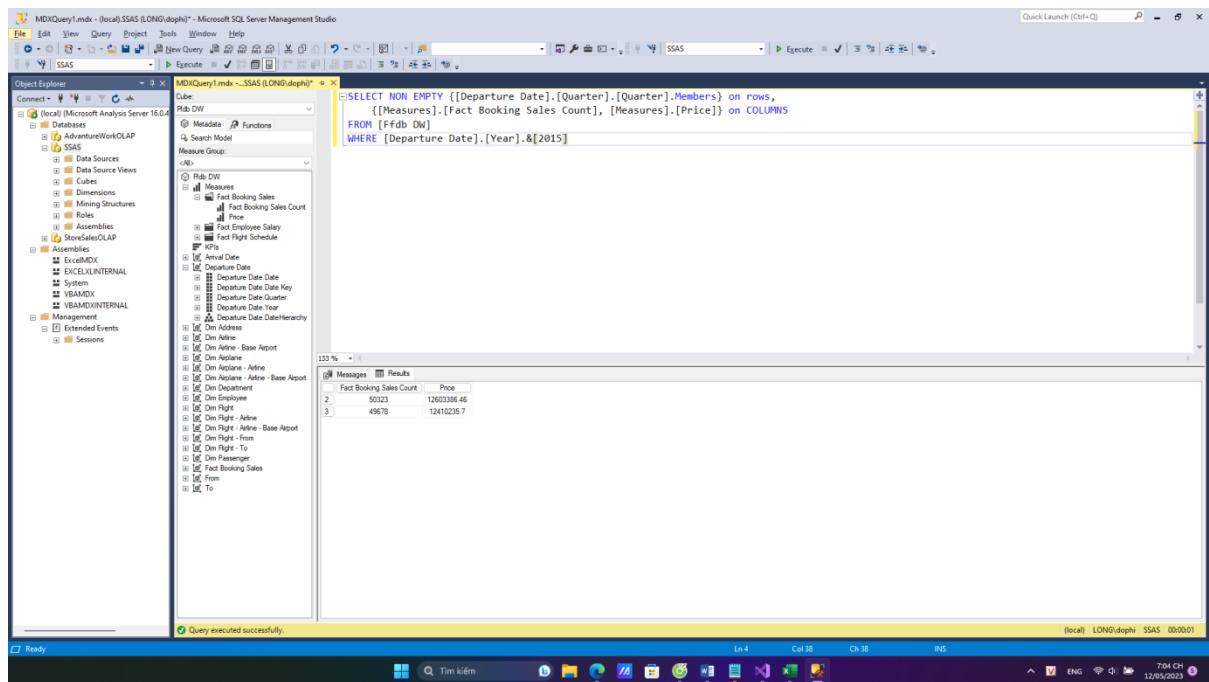
- Sử dụng công cụ SSAS:



- Sử dụng pivot table trong Excel

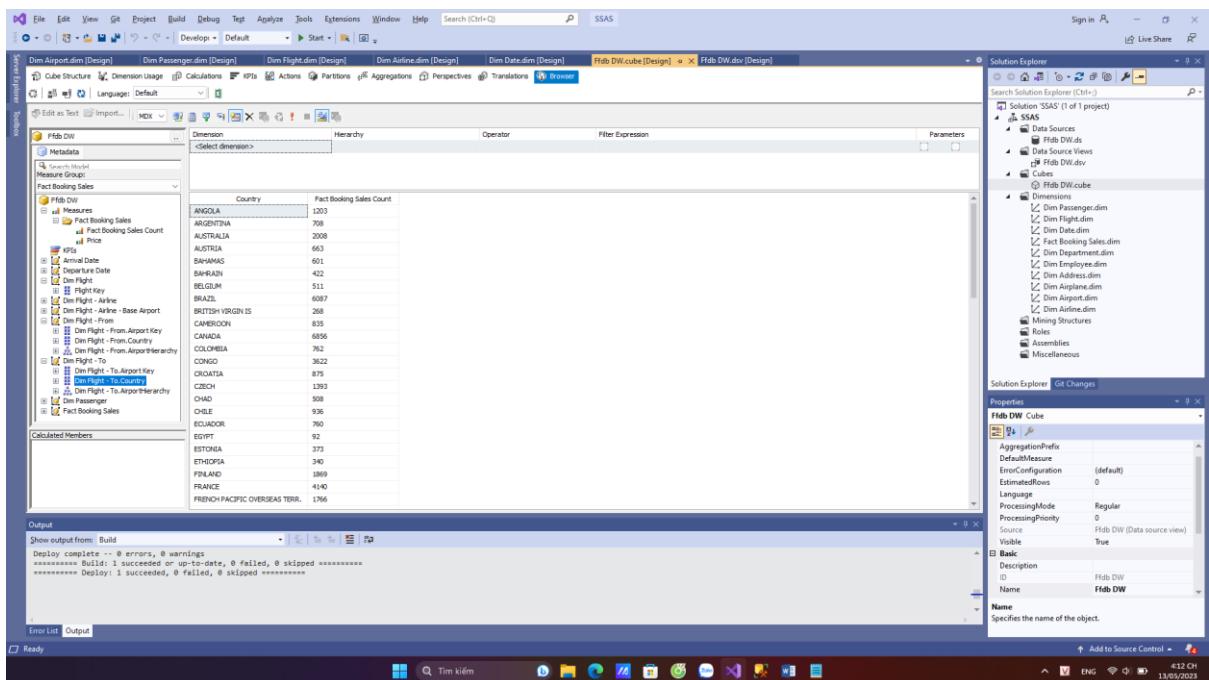
	Departure Date.Year	Fact Booking Sales	Count	Price
1	2015	50323	12803386,46	
2		49678	12410235,7	
3	Grand Total	100001	25013622,16	

- Sử dụng truy vấn MDX



4.5.2 Thống kê số lượng vé theo nơi đến (quốc gia)

- Sử dụng công cụ SSAS



- Sử dụng Pivot Table trong Excel

Row Labels	Fact Booking Sales Count
ANGOLA	1203
ARGENTINA	708
AUSTRALIA	2008
AUSTRIA	663
BAHAMAS	601
BAHRAIN	422
BELGIUM	511
BRAZIL	6007
BRITISH VIRGIN IS	268
CAMEROON	835
CANADA	6856
COLOMBIA	762
CONGO	3622
CROATIA	875
CZECH	1393
CHAD	508
CHILE	936
ECUADOR	760
Egypt	92
ESTONIA	3
ETHIOPIA	340
FINLAND	1889
FRANCE	4140
FRENCH PACIFIC OVERSEAS TERR.	1766
GERMANY	5522
GIBRALTAR	640
GUANA	172
HUNGARY	243
ICELAND	922
INDIA	124
INDONESIA	1176
IRAN	618
IRELAND	639
ITALY	1004
IVORY COAST	377
JAPAN	1368
KAZAKHSTAN	108
LESOTHO	565
LIBYA	1096
MADAGASCAR	59

- Sử dụng truy vấn MDX

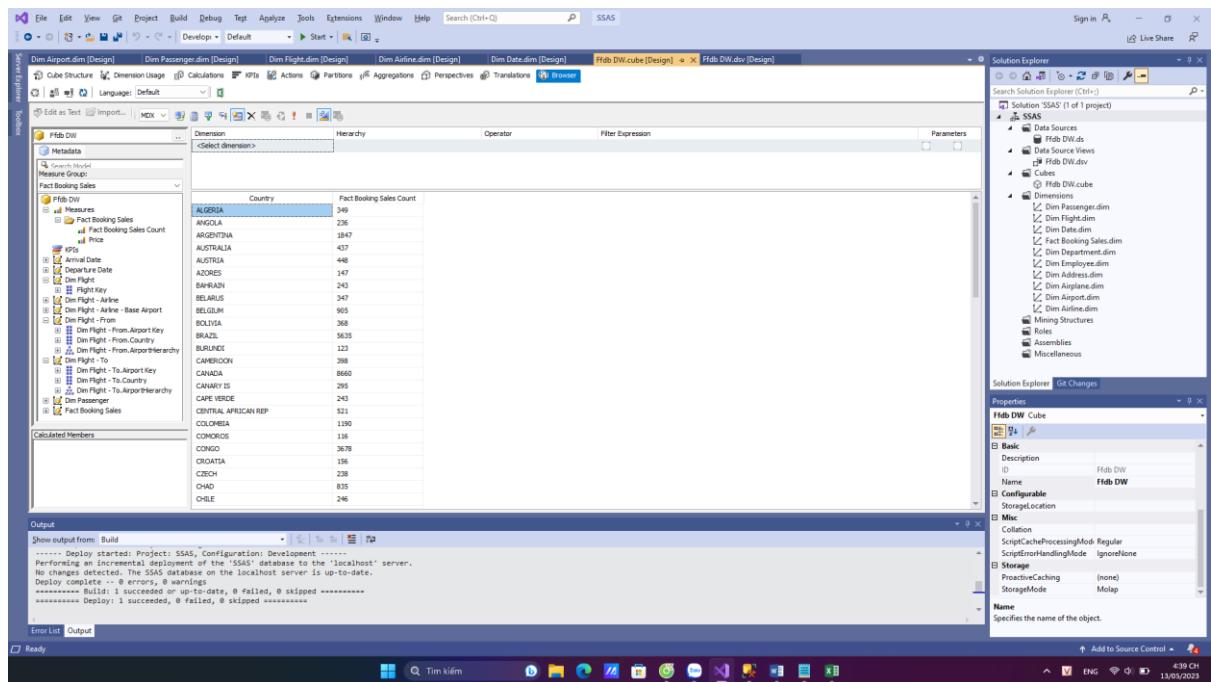
```
-- Câu 1: Thống kê số lượng vé và tổng giá vé trong năm 2015 --
SELECT
    NON EMPTY [Departure Date].[Quarter].[Quarter] ON ROWS,
    {[Measures].[Fact Booking Sales Count], [Measures].Price} ON COLUMNS
FROM [Ffdb Dw]
WHERE [Departure Date].[Year].&[2015]

-- Câu 2: --
SELECT
    NON EMPTY [Dim Flight - To].[Country].[Country] ON ROWS,
    [Measures].[Fact Booking Sales Count] ON COLUMNS
FROM [Ffdb Dw]
```

Country	Fact Booking Sales Count
ANGOLA	1203
ARGENTINA	708
AUSTRALIA	2008
AUSTRIA	663
BAHAMAS	601
BAHRAIN	422
BELGIUM	511
BRAZIL	6007
BRITISH VIRGIN IS	268
CAMEROON	835
CANADA	6856
COLOMBIA	762
CONGO	3622
CROATIA	875
CZECH	1393
CHAD	508
CHILE	936
ECUADOR	760
Egypt	92

4.5.3 Thống kê số lượng vé đã đặt theo quốc gia

- Sử dụng công cụ SSAS



- Sử dụng Pivot Table trong Excel

Country	Fact Booking Sales Count
ALGERIA	349
ANGOLA	236
ARGENTINA	1847
AUSTRALIA	437
AUSTRIA	448
AZORES	147
BAHAMAS	247
BELARUS	905
BELGIUM	368
BOLIVIA	5635
BURUNDI	123
CAMEROON	398
CANADA	8660
CANARY IS	295
CAPE VERDE	243
CENTRAL AFRICAN REP	521
COLOMBIA	1190
COMOROS	116
CONGO	3678
CROATIA	156
CZECH	238
CHAD	835
CHILE	246
DENMARK	657
DJIBOUTI	397
DOMINICAN REPUBLIC	241
Egypt	293
EL SALVADOR	573
FINLAND	559
FRANCE	4743
GERMANY	6204
GREECE	346
HUNGARY	422
ICELAND	262
INDIA	1283
INDONESIA	335
IRAN	2540
IRELAND	333
JORDAN	303

- Sử dụng truy vấn MDX

-- Câu 3: Thống kê số lượng hành khách theo quốc gia --

```

SELECT
    NON EMPTY [Dim Flight - From].[Country].[Country] ON ROWS,
    [Measures].[Fact Booking Sales Count] ON COLUMNS
FROM [Ffdb DW]

```

	Fact Booking Sales Count
ALGERIA	349
ANGOLA	256
ARGENTINA	1047
AUSTRALIA	437
AUSTRIA	446
AZORES	147
BAHRAIN	243
BELARUS	347
BEDROZIM	965
BOLIVIA	368
BRAZIL	5635
BURUNDI	123
CAMEROON	398
CANADA	8650
CAPE VERDE	295
CENTRAL AFRICAN REP	521
COLOMBIA	1190
COMOROS	116

4.5.4 Thống kê số lượng vé đã được đặt của các hãng hàng không

- Sử dụng công cụ SSAS

The screenshot shows the Microsoft Visual Studio IDE with the SSAS project 'Ffdb DW' open. The Solution Explorer on the right lists the project structure, including Data Sources, Dimensions, and Cubes. The Ffdb DW cube is selected. The main workspace displays the cube's structure, showing dimensions like Dim Passenger.dim, Dim Flight.dim, Dim Airline.dim, and Dim Date.dim, along with the Fact Booking Sales fact table and its measures. The Properties window on the right shows the cube's properties, such as Aggregation-Prefix, Default-Measure, and ProcessingMode.

- Sử dụng Pivot Table trong Excel

PivotTable Fields

- Show fields: (All)
- Dim Department
- Dim Employee
- Dim Flight
- Dim Flight - Airline
 - Airline Id
 - Airline Name
- Measures

ROWS: Airline Name

VALUES: Fact Booking Sales Count

	Fact Booking Sales Count
1 Row Labels	Fact Booking Sales Count
2	12800
3	507
4	43239
5	20462
6	19238
7	3725
8 Grand Total	100001
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	

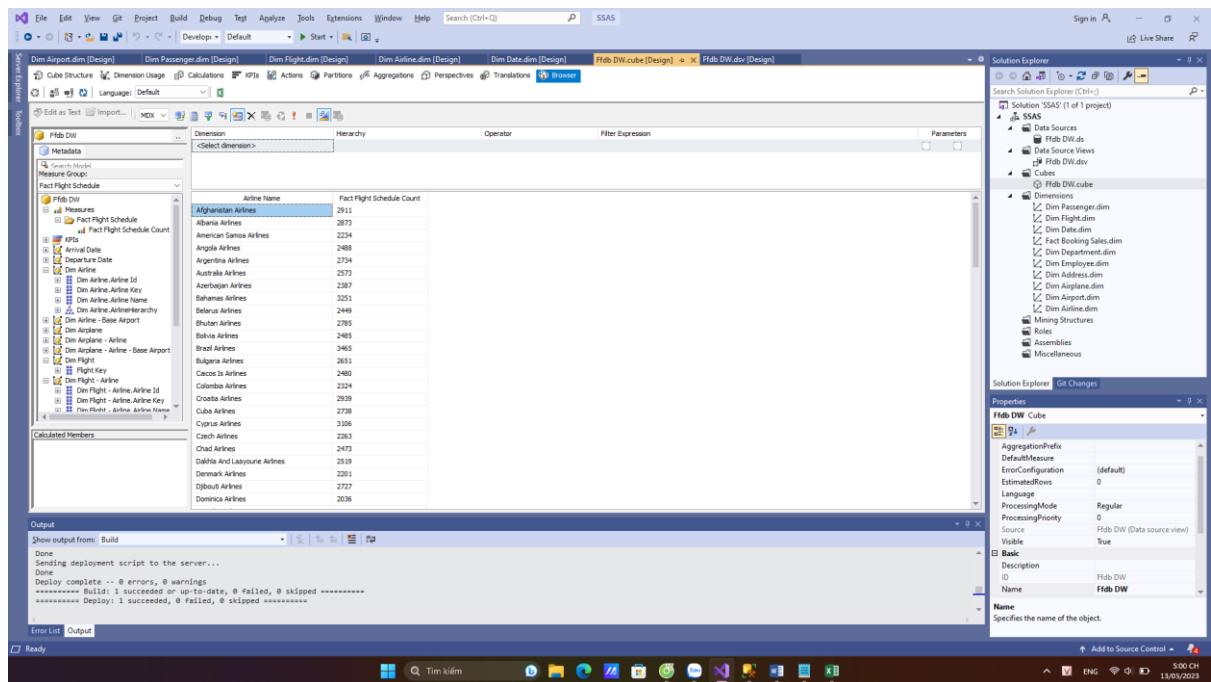
- Sử dụng truy vấn MDX

```
-- Câu 4: Thống kê số lượng vé đã được đặt của các hãng hàng không --
SELECT
    NON EMPTY [Dim Flight - Airline].[Airline Name].[Airline Name] ON ROWS,
    [Measures].[Fact Booking Sales Count] ON COLUMNS
FROM [Ffdb DW]
```

Airline Name	Fact Booking Sales Count
Djibouti Airlines	12800
Dominica Airlines	507
Falkland Is Airlines	43239
Kyrgyzstan Airlines	20462
Sudan Airlines	19238
Swaziland Airlines	3725

4.5.5 Thông kê số lượng chuyến bay của các hãng hàng không

- Sử dụng công cụ SSAS



- Sử dụng Pivot Table trong Excel

Airline Name	Fact Flight Schedule Count
Afghanistan Airlines	2911
Albania Airlines	2873
American Samoa Airlines	2234
Angola Airlines	2488
Argentina Airlines	2734
Australia Airlines	2573
Azerbaijan Airlines	2387
Bahrain Airlines	3251
Bolivia Airlines	2449
Bhutan Airlines	2765
Bolivia Airlines	2485
Brazil Airlines	2465
Bulgaria Airlines	2651
Cacica Is Airlines	2480
Colombia Airlines	2324
Croatia Airlines	2939
Cuba Airlines	2739
Cyprus Airlines	3106
Czech Airlines	2363
Chad Airlines	2473
Dakha And Layounie Airlines	2519
Denmark Airlines	2201
Djibouti Airlines	2727
Dominica Airlines	2036

- Sử dụng truy vấn MDX

-- Câu 5: Thống kê số lượng chuyến bay của các hãng hàng không --

```

SELECT
    NON EMPTY [Dim Flight - Airline].[Airline Name].[Airline Name] ON ROWS,
    [Measures].[Fact Flight Schedule Count] ON COLUMNS
FROM [Ffdb DW]

```

Airline Name	Fact Flight Schedule Count
Afghanistan Airlines	2911
Airbus	2973
Airbus Services	2234
Airline Airlines	2403
Angola Airlines	2734
Australia Airlines	2573
Azerbaijan Airlines	2387
Bahamas Airlines	3650
Bahrain Airlines	2449
Bhutan Airlines	2785
Bolivia Airlines	2405
Brazil Airlines	3465
Bulgaria Airlines	2651
Cacicos Airlines	2480
Croatia Airlines	2623
Cuba Airlines	2738
Cyprus Airlines	3106
Czech Airlines	2633
Chad Airlines	2473
Dakha Lasyoune Airlines	2919
Democratic Airlines	2201
Djibouti Airlines	2777
Dominica Airlines	2006
Ecuador Airlines	2550

4.5.6 Thống kê số lượng chuyến bay các quý trong năm 2015

- Sử dụng công cụ SSAS

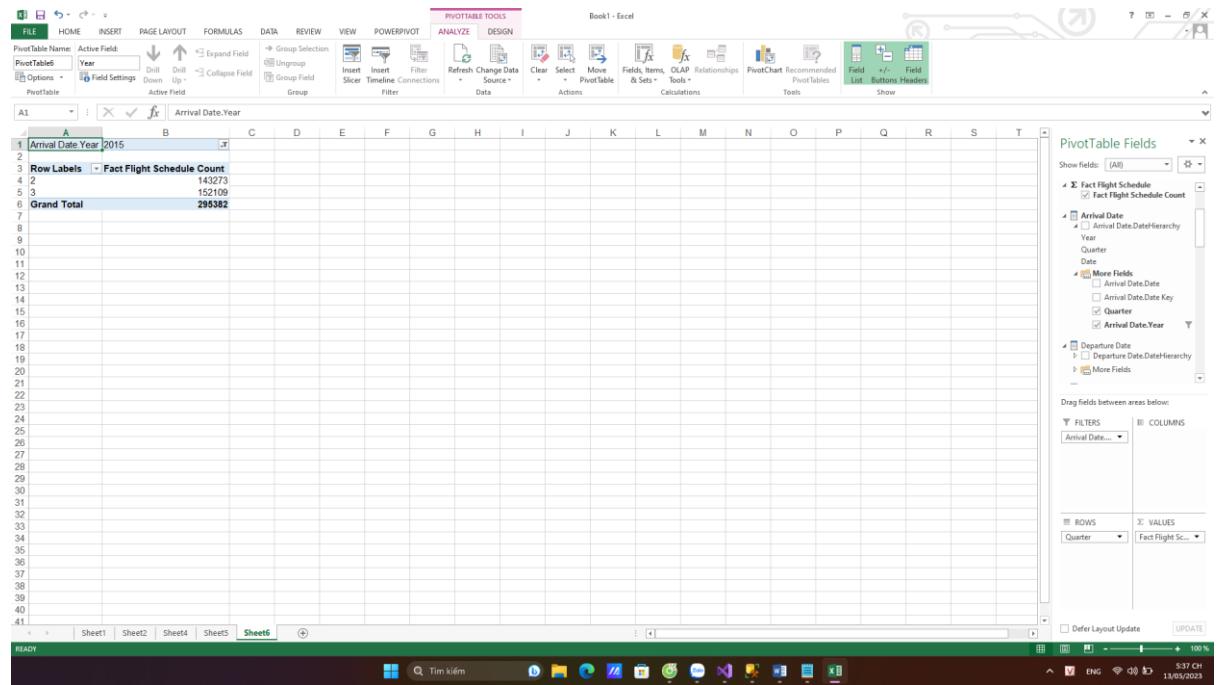
Output

```

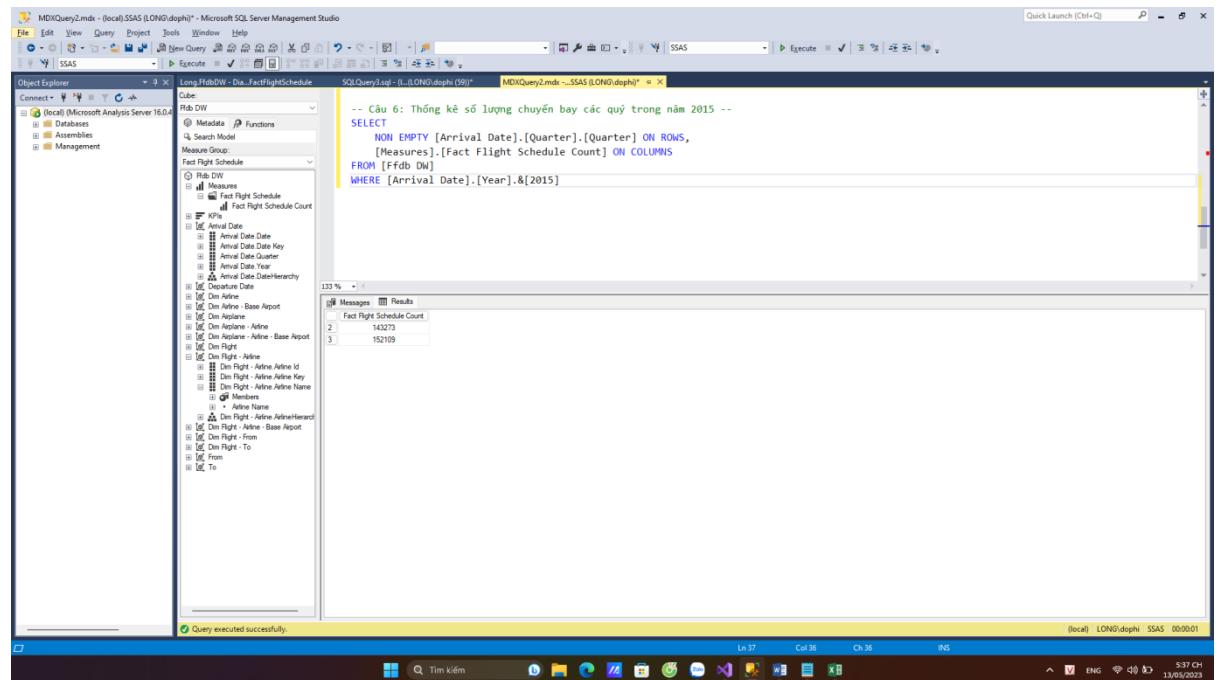
Show output from: Build
Done
Starting deployment script to the server...
Done
Deploy complete - 0 errors, 0 warnings
*****Build: 1 succeeded or up-to-date, 0 failed, 0 skipped *****
*****Deploy: 1 succeeded, 0 failed, 0 skipped *****

```

- Sử dụng Pivot Table trong Excel

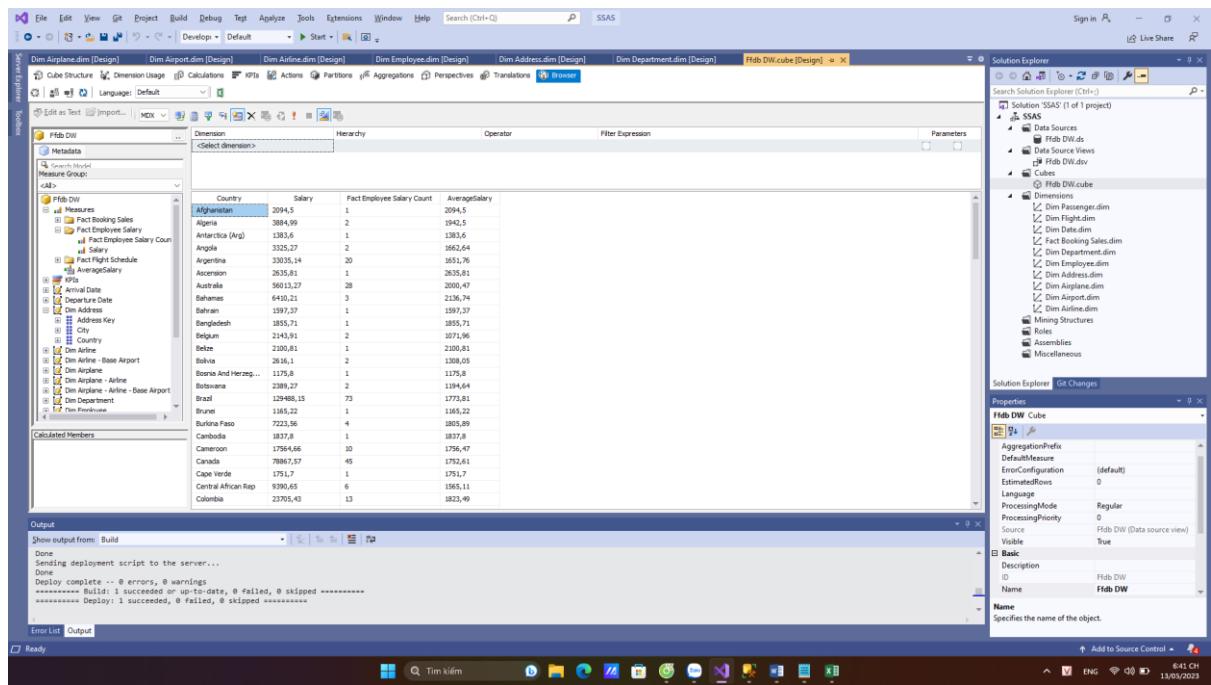


- Sử dụng truy vấn MDX



4.5.7 Thống kê tổng và trung bình lương nhân viên của các quốc gia

- Sử dụng công cụ SSAS



- Sử dụng Pivot Table trong Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Row Labels		Salary	Fact Employee Salary Count	AverageSalary													
1																	
2	Afghanistan	2094.5	1	2094.5													
3	Algeria	3884.99	2	1942.5													
4	Antarctica (Arg.)	1383.6	1	1383.6													
5	Angola	3325.27	2	1662.64													
6	Argentina	33035.14	20	1651.76													
7	Ascension	2635.81	1	2635.81													
8	Australia	56013.27	28	2000.47													
9	Bahamas	6401.21	3	2136.74													
10	Bahrain	1597.37	1	1597.37													
11	Bangladesh	1855.71	1	1855.71													
12	Belgium	2143.91	2	1071.96													
13	Belize	2100.0	1	2100.0													
14	Bolivia	2616.1	2	1308.05													
15	Bosnia And Herzegovina	1175.8	1	1175.8													
16	Botswana	2389.27	2	1194.64													
17	Brazil	129480.15	73	1773.81													
18	Burkina Faso	1165.22	1	1165.22													
19	Burkina Faso	7223.56	4	1805.89													
20	Cambodia	1837.8	1	1837.8													
21	Cameron	1754.66	10	1754.67													
22	Canada	78867.57	45	1752.61													
23	Cape Verde	1751.7	1	1751.7													
24	Central African Rep.	8936.05	6	1556.11													
25	Colombia	23705.43	13	1823.49													
26	Congo	26982.5	14	1919.75													
27	Cook Is.	2997.78	2	1148.89													
28	Costa Rica	1380.69	1	1380.69													
29	Cuba	6076.61	3	2025.54													
30	Cyprus	1028.52	1	1028.52													
31	Chad	12902.22	7	1414.56													
32	Chile	7088.88	4	1772.22													
33	China	5273.67	3	1757.89													
34	Denmark	13560.83	7	1937.26													
35	Djibouti	4242.3	2	2121.15													
36	Ecuador	2262.81	1	2262.81													
37	Egypt	10001.96	6	1666.99													
38	Equatorial Guinea	2208.42	1	2208.42													
39	Eritrea	3388.66	2	1694.33													
40	Ethiopia	17162.71	11	1560.25													
41	Finland	5148.38	3	1716.13													

- Sử dụng truy vấn MDX

```

MDXQuery2.mdx - [local].SSAS ([LONG\dophi]) - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
Object Explorer
Connections > [local].SSAS (LONG\dophi) -> Long FfdB DW - Dim_FactEmployeeSalary
    > Rdb DW
        > Metadata > Functions
        > Search Model
        > Measure Group:
        > Rdb DW
            > <All>
                > Measures
                    > Fact Booking Sales
                    > Fact Employee Salary
                    > Fact Employee Salary Count
                    > Fact Flight Schedule
                    > AverageSalary
                > KPIs
                > Date
                > Department
                > Address
                > Airline
                > Airline - Base Airport
                > Dim Airplane
                > Dim Airplane - Airline
                > Dim Airplane - Airline - Base Airport
                > Dim Department
                > Department
                > Department Key
                > Dim Employee
                > Dim Employee
                > Dim Right - Airline
                > Dim Right - Airline - Base Airport
                > Dim Passenger
                > Dim Passenger
                > Fact Booking Sales
                > To
        > Messages Results
        > Results
SELECT
    [Dim Address].[Country].[Country] ON ROWS,
    {[Measures].[Salary], [Measures].[Fact Employee Salary Count], [Measures].[AverageSalary]} ON COLUMNS
FROM [FfdB DW]
WHERE [Arrival Date].[Year].&[2015]

-- Câu 7: Thống kê tổng lương của các nhân viên theo quốc gia --
SELECT
    [Dim Address].[Country].[Country] ON ROWS,
    {[Measures].[Salary], [Measures].[Fact Employee Salary Count], [Measures].[AverageSalary]} ON COLUMNS
FROM [FfdB DW]

-- Câu 8: Thống kê lương trung bình của nhân viên trong các phòng ban --
SELECT
    [Dim Department].[Department] ON ROWS

```

Query executed successfully.

4.5.8 Thống kê lương trung bình của nhân viên trong các phòng ban

- Sử dụng công cụ SSAS

The screenshot shows the Microsoft Visual Studio IDE with the FfdB DW cube design open. The cube structure is being edited, showing dimensions such as Department, Salary, and FactEmployeeSalaryCount, and measures like AverageSalary. The Solution Explorer and Properties panes are visible on the right side of the interface.

- Sử dụng Pivot Table trong Excel

Fact Employee Salary Count, AverageSalary

	Row Labels	Column Labels	Values
1	Row Labels	Salary	Fact Employee Salary Count, AverageSalary
2	Accounting	189	1089.66
3	Airfield	214	1813.28
4	Logistics	202	1803.5
5	Management	187	1827.63
6	Marketing	208	1794.13
7	Grand Total	1824280.39	1000
8			1824.25
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			

- Sử dụng truy vấn MDX

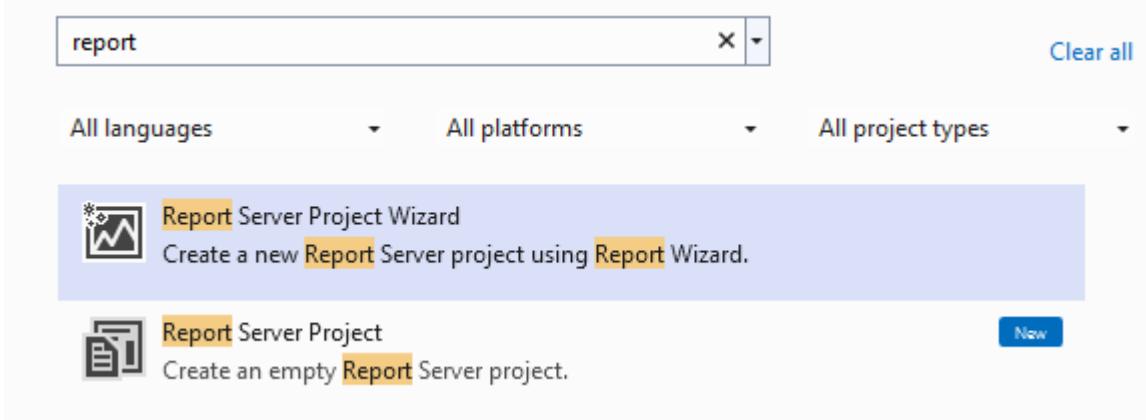
```
-- Câu 8: Thống kê lương trung bình của nhân viên trong các phòng ban --
SELECT
    [Dim Department].[Department].[Department] ON ROWS,
    {[Measures].[Salary], [Measures].[Fact Employee Salary Count], [Measures].[AverageSalary]} ON COLUMNS
FROM [Fdb DW]
```

Department	Salary	Fact Employee Salary Count	AverageSalary
Accounting	356956.59	189	1089.66
Airfield	388040.89	214	1813.28
Logistics	364306.57	202	1803.5
Management	341766.91	187	1827.63
Marketing	373179.43	208	1794.13

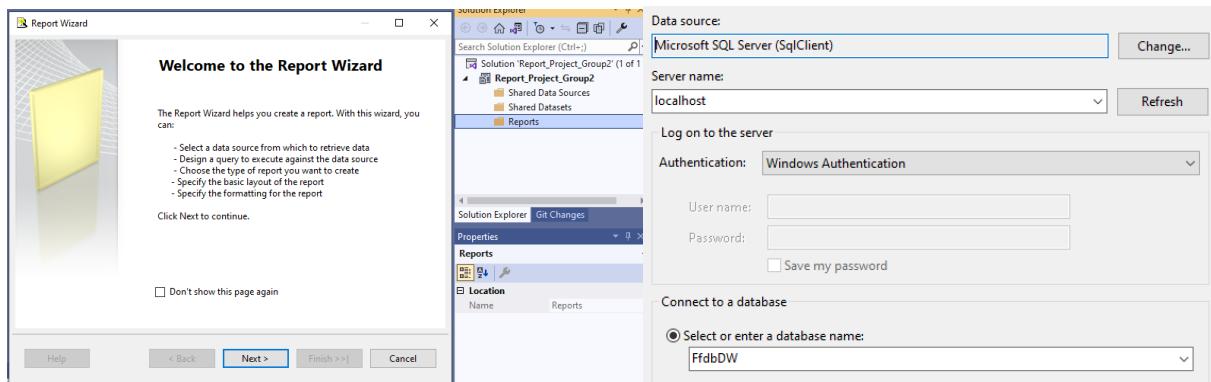
Nhận xét: tất cả truy vấn thực hiện bằng công cụ SSAS, Pivot Table, MDX đều cho kết quả giống nhau.

CHƯƠNG 5: REPORT

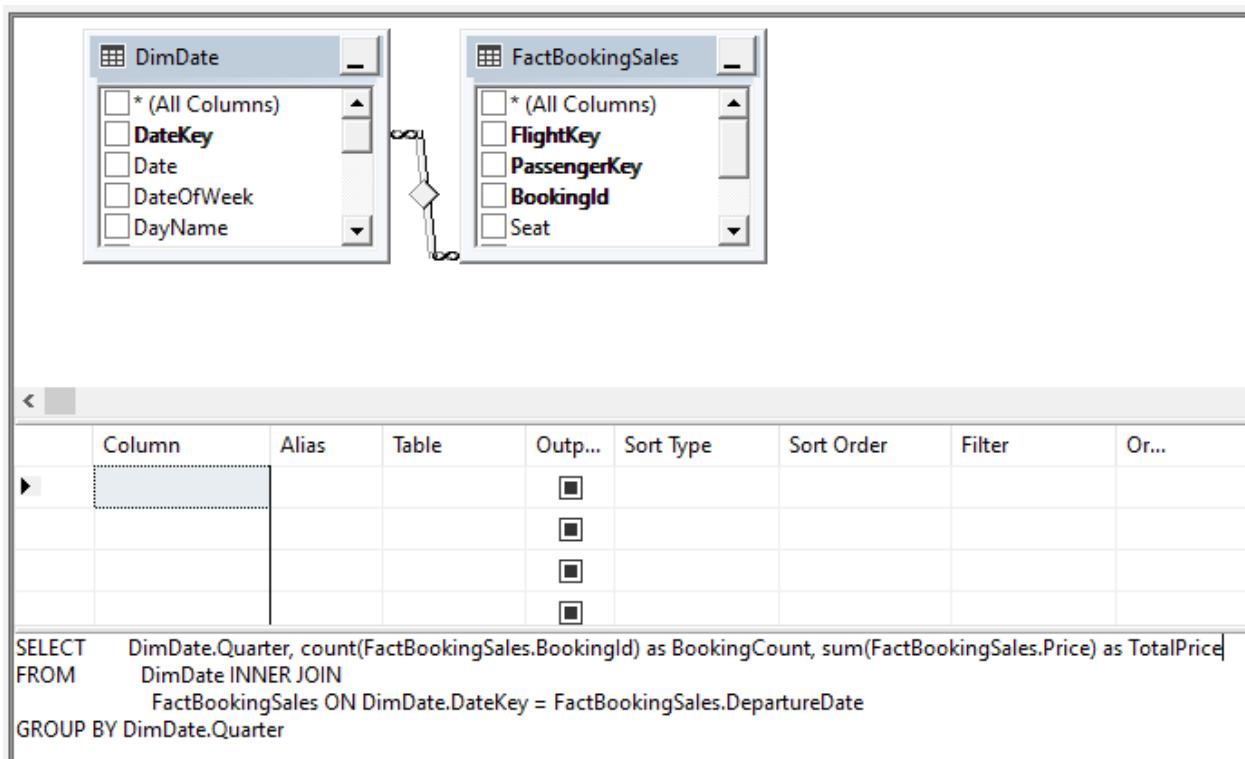
5.1 Báo cáo 1 – Tổng booking và giá bán theo quý năm 2015



Tạo Report Server Project Wizard trong Visual Studio



Tạo Report 1 với data source FfdbDW



Sử dụng bảng FactBookingSales và DimDate cùng với query

Design the Table
Choose how to group the data in the table.

Available fields:

- Page>
- Group>
- Details>

Displayed fields:

- Quarter
- BookingCount
- TotalPrice

XXXX
XXXX
XXXX
XXXX XXX XXX
XXX XXX XXX
XXXX XXX XXX
XXXX XXX XXX

Gom nhóm cột Quarter, chi tiết cột BookingCount và TotalPrice

Design Preview

1 of 1 | Find | Next

Report1

Quarter	Booking Count	Total Price
2	50323	12603386.46
3	49678	12410235.70

Xem kết quả

5.2 Báo cáo 2 – Số lượng booking theo nơi đến Arrival

The diagram shows three tables: DimAirport, DimFlight, and FactBookingSales. DimAirport has columns: AirportName, City, Country, latitude, longitude. DimFlight has columns: * (All Columns), FlightNo, From, To, Departure. FactBookingSales has columns: * (All Columns), FlightKey, PassengerKey, BookingId, Seat. Relationships are shown between DimFlight and DimAirport (From/To), and between DimFlight and FactBookingSales (FlightKey).

	Column	Alias	Table	Output	Sort Type	Sort Order	Filter	Or..
▶	DimAirport.Country			<input type="checkbox"/>				
				<input type="checkbox"/>				
				<input type="checkbox"/>				
				<input type="checkbox"/>				

```

SELECT DimAirport.Country, count(FactBookingSales.BookingId) as BookingCount
FROM DimAirport INNER JOIN
     DimFlight ON DimAirport.AirportKey = DimFlight.[To] INNER JOIN
     FactBookingSales ON DimFlight.FlightKey = FactBookingSales.FlightKey
GROUP BY DimAirport.Country
  
```

Báo cáo số lượng booking theo country dựa vào FactBookingSales, DimFlight và DimAirport

Report2

Country	Booking Count
ANGOLA	1203
ARGENTINA	708
AUSTRALIA	2008
AUSTRIA	663
BAHAMAS	601
BAHRAIN	422
BELGIUM	511
BRAZIL	6087

Xem preview kết quả số lượng booking theo nơi đến quốc gia

5.3 Báo cáo 3 – Số lượng booking theo nơi đi quốc gia

The diagram shows three tables: DimAirport, DimFlight, and FactBookingSales. DimAirport has columns: * (All Columns), AirportId, IATA, ICAO, and AirportName. DimFlight has columns: * (All Columns), FlightNo, From, To, and Departure. FactBookingSales has columns: * (All Columns), FlightKey, PassengerKey, BookingId, and Seat. There are relationships between DimFlight and DimAirport (From to AirportKey), and between DimFlight and FactBookingSales (FlightKey to FlightKey).

Column	Alias	Table	Output	Sort Type	Sort Order	Filter	Or...
			<input type="checkbox"/>				
			<input type="checkbox"/>				
			<input type="checkbox"/>				
			<input type="checkbox"/>				

```
SELECT DimAirport.Country, count(FactBookingSales.BookingId) as BookingCount
FROM DimAirport INNER JOIN
     DimFlight ON DimAirport.AirportKey = DimFlight.[From] INNER JOIN
     FactBookingSales ON DimFlight.FlightKey = FactBookingSales.FlightKey
GROUP BY DimAirport.Country|
```

Ta cần FactBookingSales, DimFlight và DimAirport

Report3

Country	Booking Count
ALGERIA	349
ANGOLA	236
ARGENTINA	1847
AUSTRALIA	437
AUSTRIA	448
AZORES	147

Xem kết quả

5.4 Báo cáo 4 – Số lượng booking theo các hãng hàng không

The diagram shows three tables: DimAirline, DimFlight, and FactBookingSales. DimAirline has columns: AirlineId, IATA, AirlineName, BaseAirport, and AirlineKey. DimFlight has columns: * (All Columns), FlightNo, From, To, and Departure. FactBookingSales has columns: * (All Columns), FlightKey, PassengerKey, BookingId, and Seat. Relationships are shown between DimAirline.AirlineKey and DimFlight.AirlineKey, and between DimFlight.FlightKey and FactBookingSales.FlightKey.

	Column	Alias	Table	Outp...	Sort Type	Sort Order	Filter	Or...
▶	AirlineName		DimAirline	<input checked="" type="checkbox"/>				
				<input type="checkbox"/>				
				<input type="checkbox"/>				
				<input type="checkbox"/>				

```

SELECT DimAirline.AirlineName, count(FactBookingSales.BookingId) as BookingCount
FROM DimAirline INNER JOIN
     DimFlight ON DimAirline.AirlineKey = DimFlight.Airline INNER JOIN
     FactBookingSales ON DimFlight.FlightKey = FactBookingSales.FlightKey
GROUP BY DimAirline.AirlineName
  
```

Nhóm theo AirlineName của DimAirline, tính số lượng BookingId của FactBookingSales

Report4

Airline Name	Booking Count
Djibouti Airlines	12830
Dominica Airlines	507
Falkland Is Airlines	43239
Kyrgyzstan Airlines	20462
Sudan Airlines	19238
Swaziland Airlines	3725

Kết quả số lượng booking theo từng hãng hàng không

5.5 Báo cáo 5 – Tổng số lượng chuyến bay của các hãng hàng không

Column	Alias	Table	Outp...	Sort Type	Sort Order	Group By
AirlineName		DimAirline	<input checked="" type="checkbox"/>			Group By
FlightId	Expr1	FactFlight...	<input checked="" type="checkbox"/>			Count
			<input type="checkbox"/>			

```
SELECT DimAirline.AirlineName, COUNT(FactFlightSchedule.FlightId) AS FlightCount
FROM DimAirline INNER JOIN
     FactFlightSchedule ON DimAirline.AirlineKey = FactFlightSchedule.AirlineKey
GROUP BY DimAirline.AirlineName
```

Tổng FlightId của FactFlightSchedule theo từng AirlineName của DimAirline

Report5

Airline Name	Flight Count
Afghanistan Airlines	2911
Albania Airlines	2873
American Samoa Airlines	2234
Angola Airlines	2488
Argentina Airlines	2734

Tổng số chuyến bay theo từng quốc gia

5.6 Báo cáo 6 – Tổng chuyến bay toàn quốc theo từng quý năm 2015

Column	Alias	Table	Outp...	Sort Type	Sort Order	Fil
Quarter		DimDate	<input checked="" type="checkbox"/>			

```
SELECT DimDate.Quarter, count(FactFlightSchedule.FlightId) as FlightCount
FROM DimDate INNER JOIN
     FactFlightSchedule ON DimDate.DateKey = FactFlightSchedule.ArrivalDate
GROUP BY DimDate.Quarter
```

Gom tổng số chuyến bay theo Quarter trong DimDate

Report6

Quarter	Flight Count
2	143273
3	152109

Xem kết quả

5.7 Báo cáo 7 – Tổng và trung bình lương nhân viên hàng không theo từng quốc gia

The screenshot shows the SSAS Data Source View (DSV) designer interface. At the top, there are two tables: 'DimAddress' and 'FactEmployeeSalary'. A relationship is established between them, specifically between the 'Country' column in 'DimAddress' and the 'AddressKey' column in 'FactEmployeeSalary'. Below the tables is a results grid where columns are mapped from the tables to aliases: 'Country' is mapped to 'DimAddress' and 'Group By'; 'Salary' is mapped to 'FactEmployeeSalary' and 'Sum'; and 'EmployeeKey' is mapped to 'FactEmployeeSalary' and 'Count'. At the bottom, the generated SQL query is displayed:

```
SELECT DimAddress.Country, SUM(FactEmployeeSalary.Salary) AS TotalSalary, COUNT(FactEmployeeSalary.EmployeeKey) AS EmployeeCount,
       AVG(FactEmployeeSalary.Salary) AS AverageSalary
FROM DimAddress INNER JOIN
     FactEmployeeSalary ON DimAddress.AddressKey = FactEmployeeSalary.AddressKey
GROUP BY DimAddress.Country
```

Tổng salary, số lượng nhân viên, trung bình salary theo Country của DimAddress

Report7

Country	Total Salary	Employee Count	Average Salary
Afghanistan	2094.50	1	2094.50000
Algeria	3884.99	2	1942.49500
Angola	3325.27	2	1662.63500
Antarctica (Arg)	1383.60	1	1383.60000
Argentina	33035.14	20	1651.75700
Ascension	2635.81	1	2635.81000
Australia	56013.27	28	2000.473928
Bahamas	6410.21	3	2136.736666

5.8 Báo cáo 8 – Lương nhân viên tại các phòng ban

The screenshot shows the SSAS Data Source View (DSV) designer interface. It displays two tables: DimDepartment and FactEmployeeSalary. A relationship is being established between them, indicated by a line connecting the DepartmentKey column in DimDepartment to the DepartmentKey column in FactEmployeeSalary. Below the tables is a query grid where the 'Department' column from DimDepartment is selected. At the bottom, the generated SQL SELECT statement is shown:

```
SELECT DimDepartment.Department, sum(FactEmployeeSalary.Salary) as TotalSalary, count(FactEmployeeSalary.EmployeeKey) as EmployeeCount,
       avg(FactEmployeeSalary.Salary) as AverageSalary
  FROM DimDepartment INNER JOIN
       FactEmployeeSalary ON DimDepartment.DepartmentKey = FactEmployeeSalary.DepartmentKey
 GROUP BY DimDepartment.Department
```

Tổng salary, số lượng nhân viên, trung bình salary theo phòng ban
Department của DimDepartment

Report8

Department	Total Salary	Employee Count	Average Salary
Accounting	356956.59	189	1888.659206
Airfield	388040.89	214	1813.275186
Logistics	364306.57	202	1803.497871
Management	341766.91	187	1827.630534
Marketing	373179.43	208	1794.131875

Xem kết quả

CHƯƠNG 6: KẾT LUẬN

6.1 Kết quả đạt được

Trong quá trình học tập cho tới thời điểm hoàn thành đồ án, nhóm đã tìm hiểu và ứng dụng những kiến thức về kho dữ liệu và OLAP để đạt được những kết quả:

- Ôn lại kiến thức về câu lệnh truy vấn SQL và hệ quản trị cơ sở dữ liệu SQL Server
- Hiểu rõ các khái niệm cơ bản về một kho dữ liệu Datawarehouse
- Hiểu được quy trình xây dựng một kho dữ liệu theo kỹ thuật mô hình hóa theo chiều Dimensional Modeling từ level cao đến level chi tiết
- Xây dựng được quy trình tích hợp dữ liệu vào kho thông qua câu lệnh truy vấn SQL và công cụ hỗ trợ SSIS
- Phân tích được những câu hỏi, vấn đề được đặt ra trong kho dữ liệu đã xây dựng
- Từ kho dữ liệu giúp nắm rõ và hình dung được một hệ thống quản lý bán vé máy bay trên thực tế diễn ra như thế nào

6.2 Hạn chế

Trong quá trình làm đồ án cho tới khi hoàn thành, nhóm nhận thấy có những hạn chế như sau:

- Thiết kế diagram về kho dữ liệu, các bảng Dim có nhiều khóa ngoại liên kết với nhau
- Quá trình SSIS còn vài chỗ kiểu dữ liệu còn chưa hoàn chỉnh về kiểu dữ liệu

6.3 Phân công nhiệm vụ

Công việc	Chi tiết	Nguyễn Khoa Quang Thắng	Nguyễn Huỳnh Thắng	Đỗ Phi Long	Hoàng Uyên	Mức độ hoàn thành
Tiền xử lí Dataset	Tạo database và các table	X				100%
	Tạo scripts và thêm data từ dataset	X				50%
				X		50%
	Xóa table không cần thiết, thay đổi kiểu dữ liệu phù hợp	X			X	50%
Thiết kế kho dữ liệu	Đặt ra yêu cầu business và xác định mức độ chi tiết grain, đưa ra các fact measure	X				100%
	Xác định các dimension	X				100%
	Thiết kế chi tiết definition, target, source các dimension	X				40%
			X			30%
				X		30%
	Thiết kế chi tiết definition, target, source các fact	X				70%
				X		30%
Tích hợp dữ liệu vào kho SSIS	Tạo các bảng Staging từ dataset với sql query	X				70%
				X		30%
	Extract và Transform, tạo các bảng staging với Visual Studio		X			100%
Phân tích dữ liệu SSAS	Loading		X			100%
	Đặt câu hỏi cho business	X				50%
			X			50%
	Xây dựng datasource view		X			30%
				X		70%
	Xây dựng Cube			X		100%
Tạo report	Truy vấn với pivot table và MDX		X			30%
				X		70%
	Tạo các báo cáo cho kết quả của quá trình SSAS	X				50%
Xây dựng đồ án	Format bài báo cáo				X	100%

TÀI LIỆU THAM KHẢO

- [1]. Vấn đề đặt ra: <https://thinksmart.com.vn/hon-100-nam-ra-doi-va-phat-trien-nganh-hang-khong/>
- [2]. Visual Studio 2022: <https://visualcpp.net/visual-studio-la-gi-gioi-thieu-phan-mem-visual-studio/>
- [3]. SQL Server Integration Service: <https://blog.cole.vn/sql-server-integration-services-ssis/>
- [4]. SSMS: <https://microsoft.fptcloud.com/kien-thuc/uu-diem-sql-server-2019/>