VIETNAM-KOREA UNIVERSITY OF INFORMATION AND COMMUNICATION TECHNOLOGY

**COMPUTER SCIENCE DEPARTMENT**



**DATA WAREHOUSE**

**TOPIC: FLIGHT DATA ANALYSIS AND VISUALIZATION**

Student: Đàm Phương Nam-22IT.B139

Bùi Quốc Văn-22IT.B239

Nguyễn Thúy Hằng-22IT.B067

Cao Thị Thu Huyền-22IT.B097

Instructor: MSC.Trần Thanh Liêm

**Danang, November 2024**

VIETNAM-KOREA UNIVERSITY OF INFORMATION AND COMMUNICATION TECHNOLOGY

**COMPUTER SCIENCE DEPARTMENT**



**DATA WAREHOUSE**

**TOPIC: FLIGHT DATA ANALYSIS AND VISUALIZATION**

Student: Đàm Phương Nam-22IT.B139

Bùi Quốc Văn-22IT.B239

Nguyễn Thúy Hằng-22IT.B067

Cao Thị Thu Huyền-22IT.B097

Instructor: MSC.Trần Thanh Liêm

**Danang, November 2024**

# **ACKNOWLEDGMENTS**

During the implementation of this project "Flight data analysis and visualization", we would like to send our most sincere thanks to all those who have helped and supported the team both in terms of knowledge and knowledge. spirit during implementation.

First of all, we would like to sincerely thank our teacher - MSc. Tran Thanh Liem - Lecturer of the Faculty of Information Technology, Vietnam - Korea University of Information and Communications Technology, who directly guided, commented, and helped the group throughout the project implementation process. Next, my team would like to thank the teachers in the Faculty of Information Technology for their contributions and help during the project process.

Due to the limited time to carry out the project and limited knowledge, the implementation of the project certainly cannot avoid shortcomings. My group is looking forward to receiving feedback from Mr. Tran Thanh Liem, teachers in the school and friends to gain more experience and continue to perfect our project.

Our team sincerely thanks you!

Student

**Group 17**

[**LECTURER**](https://dictionary.cambridge.org/dictionary/english-vietnamese/lecturer)**'**[**S**](https://dictionary.cambridge.org/dictionary/english-vietnamese/s) **COMMENTS**

................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

**TABLE OF CONTENTS**

[**ACKNOWLEDGMENTS** 1](#_Toc181953511)

[**CHAPTER 1: TOPIC OVERVIEW** 1](#_Toc181953512)

[**1.1.** **Reason for choose the topic:** 1](#_Toc181953513)

[**1.1.1. Solution** 2](#_Toc181953514)

[**1.1.2. Objectives and meaning of the project** 2](#_Toc181953515)

[**1.2.1. Introducing data** 2](#_Toc181953516)

[**1.2.2 Introducing where the data is provided** 2](#_Toc181953517)

[**1.2.3. Dataset parameters** 3](#_Toc181953518)

[**1.2.4. Data after extraction (from excel file)** 3](#_Toc181953519)

[**1.2.5. Detailed description of the attributes in the dataset** 4](#_Toc181953520)

[**1.2.6. Data warehouse design** 5](#_Toc181953521)

[**CHAPTER 2: DESIGN ELT AND CONSTRUCTION OF DATABASE** 7](#_Toc181953522)

[**2.1 Data preprocessing** 7](#_Toc181953523)

[**2.2 Conceptual ETL design:** 7](#_Toc181953524)

[**2.3 ETL development by using SSIS** 7](#_Toc181953525)

[**CHAPTER 3: SQL SERVER ANALYSIS SERVICES** 17](#_Toc181953526)

[**3.1. SSAS Overview** 17](#_Toc181953527)

[**CHAPTER 4: SQL SERVER REPORTING SERVICES** 21](#_Toc181953528)

[**CHAPTER 5: CONCLUDE** 24](#_Toc181953529)

[**5.1 Results achieved:** 24](#_Toc181953530)

[**5.2 Limitations:** 24](#_Toc181953531)

[**REFERENCES** 25](#_Toc181953532)

**IMAGE CATALOG**

[Figure 1: Data from US Airline Flight Routes and Fares table 8](#_heading=h.gjdgxs)

[Figure 2: Data from the IATA Arlines table 8](#_heading=h.1ksv4uv)

[Figure 3: Table of detailed description of the attributes in the dataset 10](#_heading=h.30j0zll)

[Figure 4: Logical data model 10](#_heading=h.1fob9te)

[Figure 5: Star Schema 11](#_heading=h.3znysh7)

[Figure 6: SSIS 11](#_heading=h.2et92p0)

[Figure 7: Origin\_Dim 12](#_heading=h.tyjcwt)

[Figure 8: Data from the SQL of the table Origin\_Dim 12](#_heading=h.3dy6vkm)

[Figure 9: Destination\_Dim 12](#_heading=h.1t3h5sf)

[Figure 10: Data from the SQL of the table Destination\_Dim 13](#_heading=h.4d34og8)

[Figure 11: Farelow\_Dim 13](#_heading=h.2s8eyo1)

[Figure 12: Data from the SQL of the table Farelow\_Dim 14](#_heading=h.17dp8vu)

[Figure 13: Farelg\_Dim 14](#_heading=h.3rdcrjn)

[Figure 14: Data from the SQL of the table Farelg\_Dim 15](#_heading=h.26in1rg)

[Figure 15: Schedule\_Dim 16](#_heading=h.lnxbz9)

[Figure 16: Data from the SQL of the table Schedule\_Dim 16](#_heading=h.35nkun2)

[Figure 17: Flight\_Dim 17](#_heading=h.1ksv4uv)

[Figure 18: Data from the SQL of the table Flight\_Dim 17](#_heading=h.44sinio)

[Figure 19: Actual\_Dim 18](#_heading=h.2jxsxqh)

[Figure 20: Data from the SQL of the table Actual\_Dim 18](#_heading=h.z337ya)

[Figure 21: Time\_Dim 19](#_heading=h.3j2qqm3)

[Figure 22: Data from the SQL of the table Time\_Dim 20](#_heading=h.1y810tw)

[Figure 23: Date\_Dim 20](#_heading=h.4i7ojhp)

[Figure 24: Ariline\_Dim 21](#_heading=h.2xcytpi)

[Figure 25: Data from the SQL of the table Ariline\_Dim 21](#_heading=h.1ci93xb)

[Figure 26: Route\_Fact 22](#_heading=h.3whwml4)

[Figure 27: Data from the SQL of the table Route\_Fact 22](#_heading=h.2bn6wsx)

[Figure 28: Fare\_Fact 23](#_heading=h.qsh70q)

[Figure 29: Data from the SQL of the table Fare\_Fact 23](#_heading=h.3as4poj)

[**CATEGORIES**](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/category)[**OF**](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/of)**ACRONYMS**

|  |  |
| --- | --- |
| SSIS | SQL Server Integration Services |
| SSAS | SQL Server Analysis Services |
| SSRS | SQL Server Reporting Services |
| OLAP | Online Analytical Processing |
| ETL | Extract, Transform và Load |
| RDSMS | Relational Database Management System |

# 

# **CHAPTER 1: TOPIC OVERVIEW**

## **Reason for choose the topic:**

Providing detailed information about flight performance is an important reason. Data analysis allows evaluating the performance of airlines through indicators such as flight delay rate, flight cancellation rate and average flight time. This information helps improve service quality and operational processes.

Optimizing flight schedules is also a necessary factor. Analytical data can help airlines detect patterns and trends in flight routes, thereby adjusting flight frequency and improving aircraft utilization. This helps save costs and improve operational efficiency.

Analyzing the impact of external factors is an aspect that cannot be overlooked. Flight data allows analyzing the impact of weather, special events or traffic conditions at the airport. Understanding these factors helps airlines better prepare and minimize risks.

Demand forecasting is an important part of business strategy. Through predictive analytics methods, flight data can help companies identify travel trends, thereby predicting future demand. This aids in planning the number of flights and managing resources effectively.

Enhancing customer experience is also an important reason. Analyzing passenger feedback data helps identify specific problems they encounter. Thereby, airlines can improve services such as check-in processes, providing flight information, and better customer support.

Supporting strategic decisions is an indispensable benefit. Analytical data provides insight into market trends, helping managers make more strategic decisions on investments, network expansion and product development.

Finally, improving sustainability is an increasingly important factor. Flight data analysis can assist in the research and development of more sustainable solutions, such as optimizing fuel consumption and minimizing carbon emissions, which is becoming increasingly necessary in the context of environmental protection.

### **1.1.1. Solution**

The flight data visualization and analysis solution is to build an integrated data warehouse that stores historical information about flights from multiple sources, through an ETL process to ensure data consistency. Predict next year's trends, forecast passenger numbers and potential delays, helping airlines and managers make accurate strategic decisions.

### **1.1.2. Objectives and meaning of the project**

Objective: Analyze and visualize flight data to build a data warehouse and OLAP application to analyze detailed information about flights. The goal is to provide tools to help understand trends in the airline industry, such as changes in the frequency of delays, popular routes, and factors affecting operations. This helps airlines and regulators make strategic decisions and improve services.

What it means: This solution helps managers and airlines better understand operational performance, identifying key factors affecting flights such as weather or rush hour. At the same time, it also helps compare operational efficiency between routes or between different periods. This thereby supports better planning, optimized schedules, and improved customer satisfaction through predicting and handling potential problems.

**1.2. Overview of the dataset**

### **1.2.1. Introducing data**

Data sources are collected from stackoverflow.com, kaggle.

### **1.2.2 Introducing where the data is provided**

Kaggle is an online platform that allows users to share ideas, find inspiration, and participate in competitions with other data scientists. This is also a place to learn new information and programming tips, while providing real-world examples of data science applications. Kaggle has a wide variety of data sets from different sectors, from simple data like video game sales to complex data like air pollution. This data has a reliable source, creating opportunities for users to train and test models on highly applicable projects.

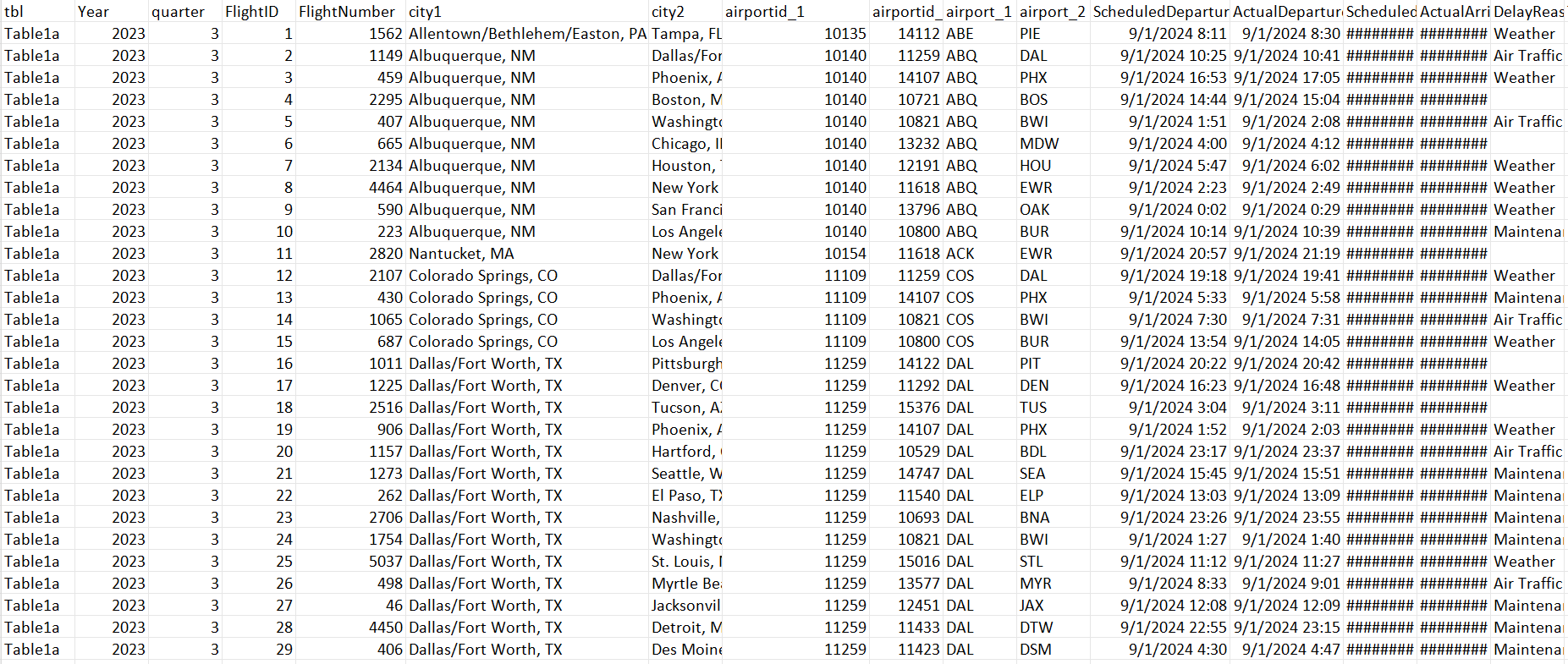
Stackoverflow.com is an online website that helps developers work more effectively through "crowdsourcing", where the developer community contributes ideas and solves problems publicly. The site has more than 50 million visits per month, with an average of more than 50 thousand online developers, each visiting about 7 times per month. Stack Overflow has more than 14 million programming questions and 19 million answers, with about 7.5 billion solved problems, helping the software development community optimize their time and effort at work.

### **1.2.3. Dataset parameters**

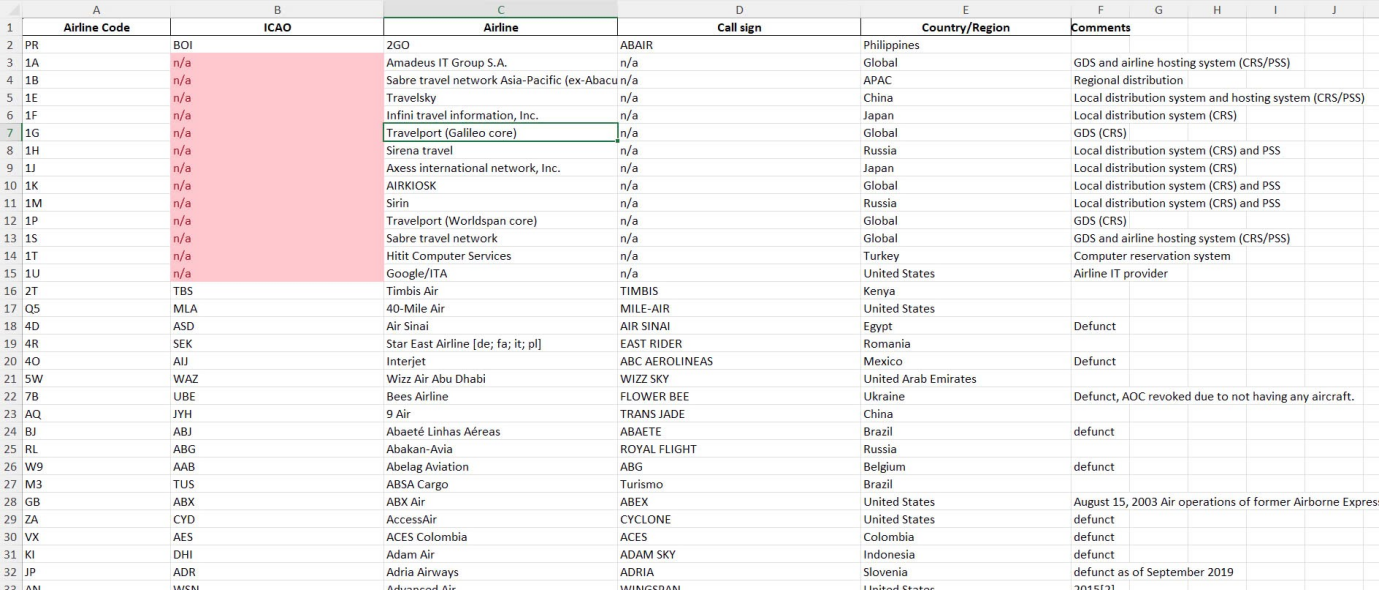
The data includes two tables: US Airline Flight Routes and Fares and IATA Arlines. Responses were reviewed by Stackoverflow, qualifying them for analysis based on completion rate and time spent on a survey. A total of 67,441 people completed the entire survey.

### **1.2.4. Data after extraction (from excel file)**

Perform data extraction from US Airline Flight Routes and Fares and IATA Airlines tables. Serves the purpose of building a data warehouse: Data after extracting from original data:



*Figure 1:* [*Data*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/from)[*US*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/us)[*Airline*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/airline)[*Flight*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/flight)*Routes*[*and*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/and)*Fares*[*table*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/table)

**

*Figure 2:* [*Data*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/the)*IATA Arlines*[*table*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/table)

### **1.2.5. Detailed description of the attributes in the dataset**

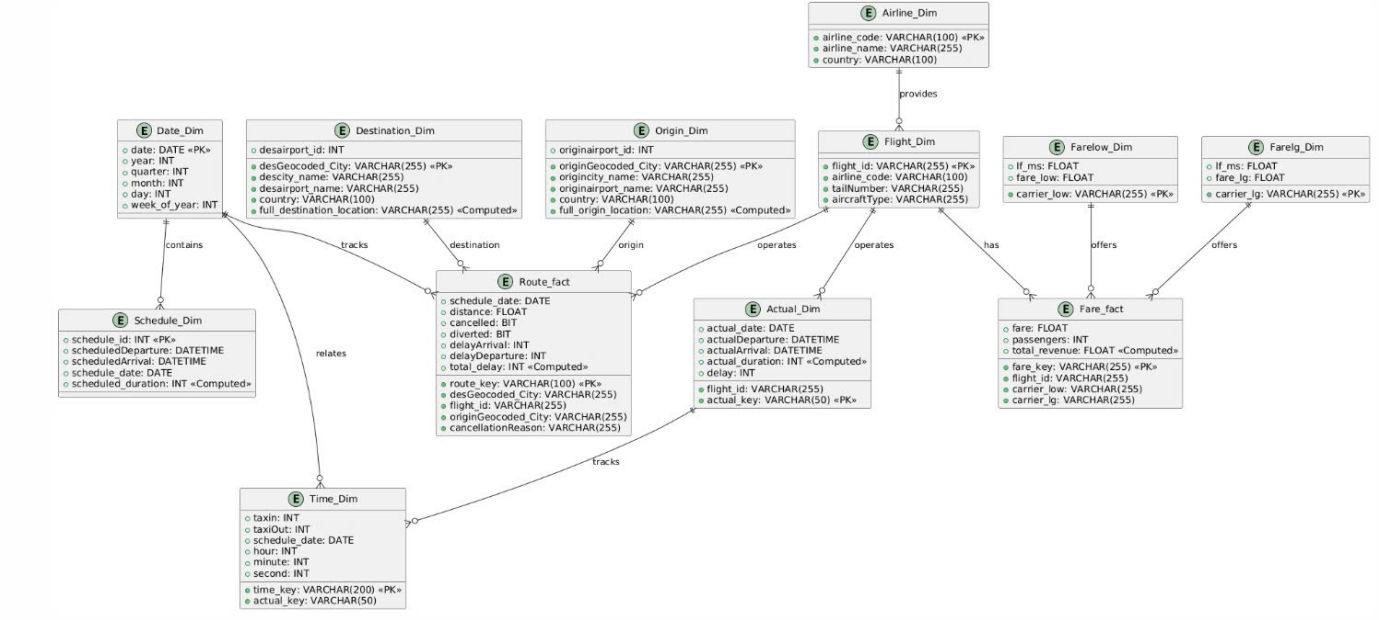
|  |  |  |
| --- | --- | --- |
| **No.** | **Property Name** | **Description** |
| 1 | Year | The year in which the flight took place. |
| 2 | Quarter | The quarter of the year (1-4) in which the flight took place. |
| 3 | FlightID | A unique identifier for each flight. |
| 4 | Airline | The name of the airline operating the flight. |
| 5 | airline\_code | The code representing the airline. |
| 6 | FlightNumber | The flight number assigned by the airline. |
| 7 | city1 | The departure city. |
| 8 | city2 | The arrival city. |
| 9 | airportid\_1 | The unique identifier of the departure airport. |
| 10 | airportid\_2 | The unique identifier of the arrival airport. |
| 11 | airport\_1 | The name of the departure airport. |
| 12 | airport\_2 | The name of the arrival airport. |
| 13 | ScheduledDeparture | The scheduled departure time. |
| 14 | ActualDeparture | The actual departure time. |
| 15 | ScheduledArrival | The scheduled arrival time. |
| 16 | ActualArrival | The actual arrival time. |
| 17 | DelayMinutes | The total number of minutes the flight was delayed. |
| 18 | DelayReason | The reason for the delay. |
| 19 | TaxiIn | The time it took to taxi to the gate after landing. |
| 20 | TaxiOut | The time it took to taxi to the runway before takeoff. |
| 21 | Cancelled | Indicates if the flight was canceled (yes/no). |
| 22 | Diverted | Indicates if the flight was diverted to a different airport (yes/no). |
| 23 | AircraftType | The type of aircraft used for the flight. |
| 24 | TailNumber | A unique identifier for the aircraft. |
| 25 | Distance | The distance between the departure and arrival airports. |
| 26 | passengers | The number of passengers on the flight. |
| 27 | fare | The fare paid by passengers. |
| 28 | carrier\_lg | The code of the large airline. |
| 29 | large\_ms | The market share of the large airline. |
| 30 | fare\_lg | The fare of the large airline. |
| 31 | carrier\_low | The code of the low-cost airline. |
| 32 | lf\_ms | The market share of the low-cost airline. |
| 33 | fare\_low | The fare of the low-cost airline. |
| 34 | Geocoded\_City1 | The geographic location of the departure city. |
| 35 | Geocoded\_City2 | The geographic location of the arrival city. |
| 36 | route\_code | The code representing the flight route. |
| 37 | Hour, minute, second | Time of Time |

*Figure 3: Table of d*[*etailed*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/detailed)[*description*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/description)[*of*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/the) *attributes* [*in*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/in)[*the*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/the) *dataset*

### **1.2.6. Data warehouse design**

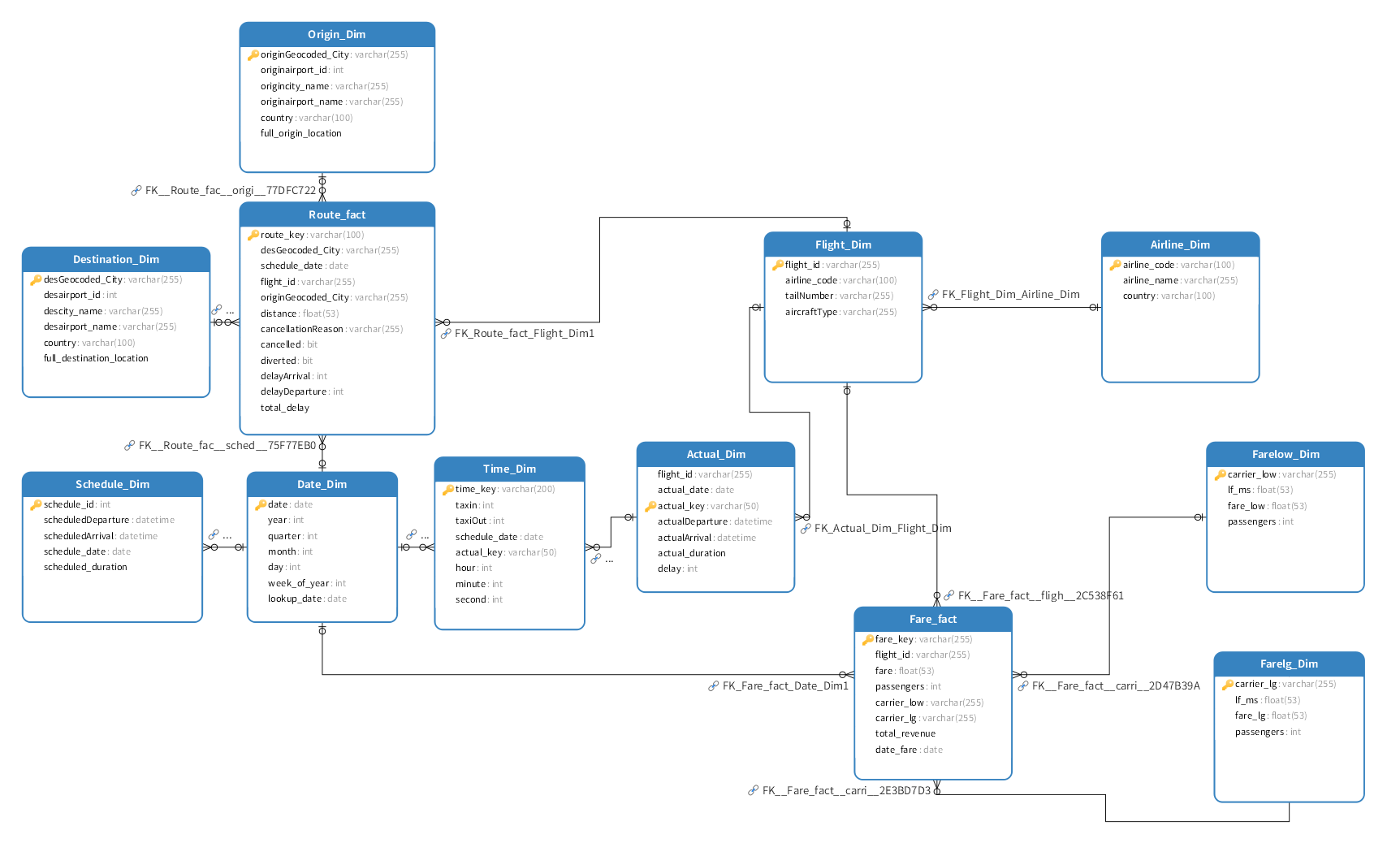
The Fact table includes Route\_fact and Fare\_fact which are fact tables containing foreign keys of Dimension tables and other attributes. Flight\_Dim, Origin\_Dim, Destination\_Dim, Time\_Dim, Farelg\_Dim, Farelow\_Dim, Actual\_Dim, Schedule\_Dim are dimensional tables containing primary key attributes and other attributes of the table.

#### **1.2.6.1. Logical design**



*Figure 4: Logical data model*

#### **1.2.6.2 Physical design**



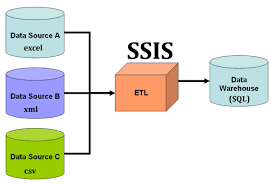
*Figure 5:* [*Star*](https://dictionary.cambridge.org/vi/dictionary/english-vietnamese/star)*Schema*

# **CHAPTER 2: DESIGN ELT AND CONSTRUCTION OF DATABASE**

## **2.1 Data preprocessing**

Extract and select necessary attributes for the main fact table:

## **2.2 Conceptual ETL design:**

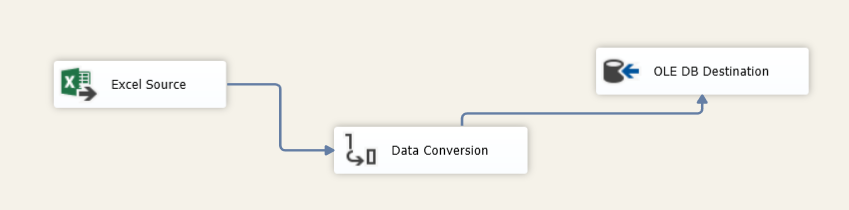


*Figure 6: SSIS*

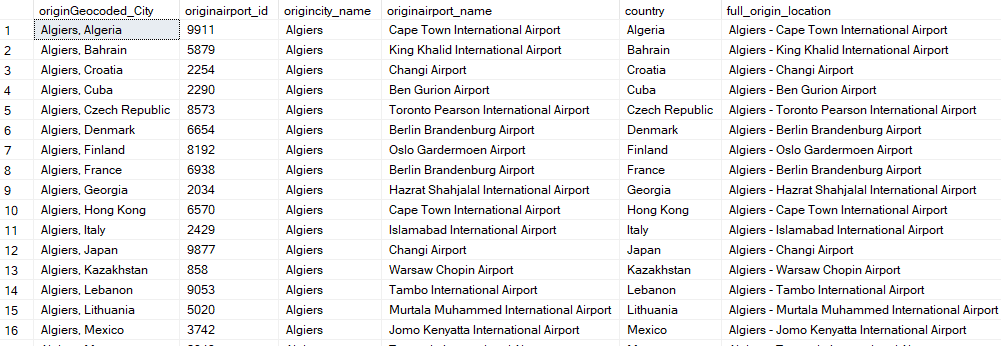
Control flow: Flight\_Dim,Actual\_Dim,Farelg\_Dim,Farelow\_Dim,Schedule\_Dim,Origin\_Dim,*Destination\_Dim*,Time\_Dim,Airline\_Dim,Fare\_fact,Route\_fact.

## **2.3 ETL development by using SSIS**

*Data Flow Task 1(Origin\_Dim)*

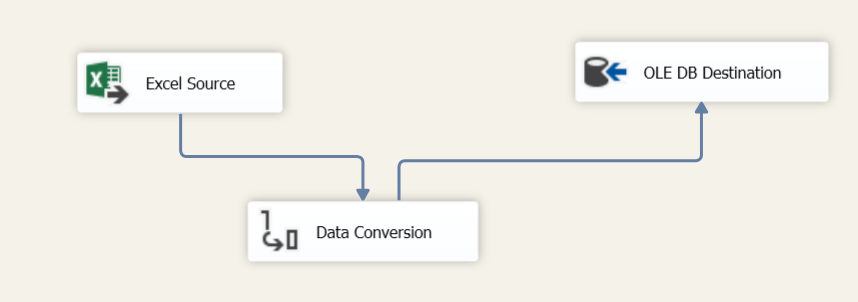
**

*Figure 7: Origin\_Dim*



*Figure 8:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Origin\_Dim*

*Data Flow Task 2(Destination\_Dim)*

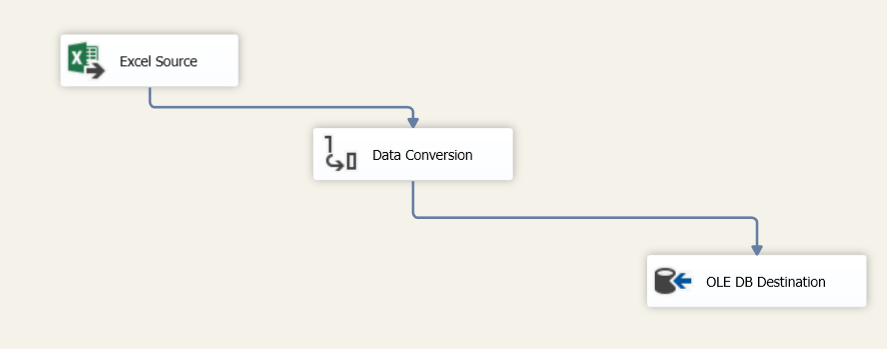
**

*Figure 9: Destination\_Dim*

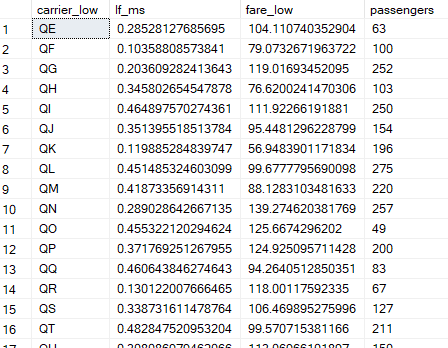


*Figure 10:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Destination\_Dim*

*Data Flow Task 3(Farelow\_Dim)*

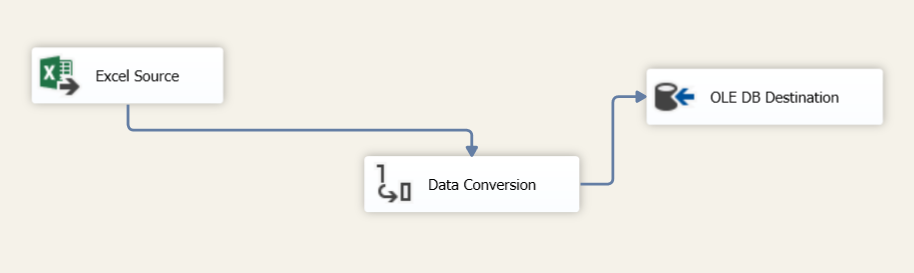
**

*Figure 11: Farelow\_Dim*

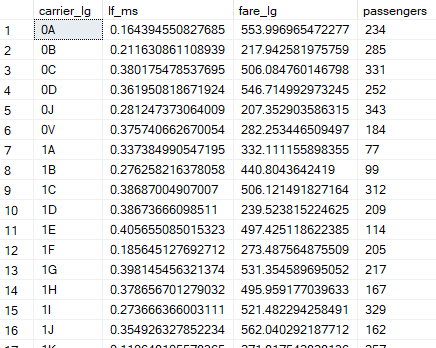


*Figure 12:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Farelow\_Dim*

*Data Flow Task 4(Farelg\_Dim)*

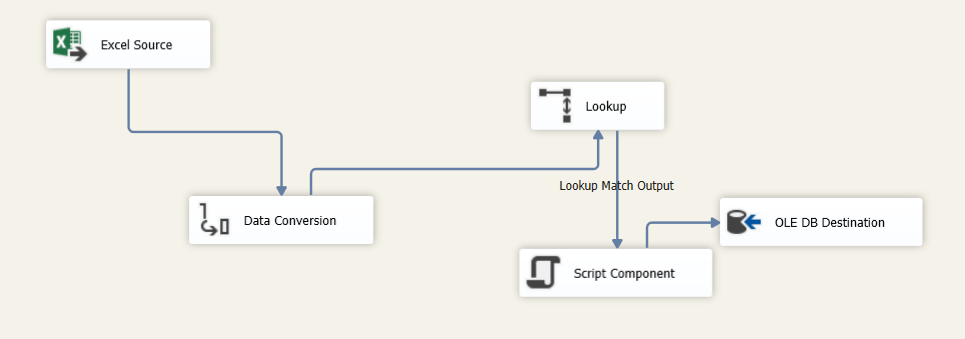
**

*Figure 13: Farelg\_Dim*

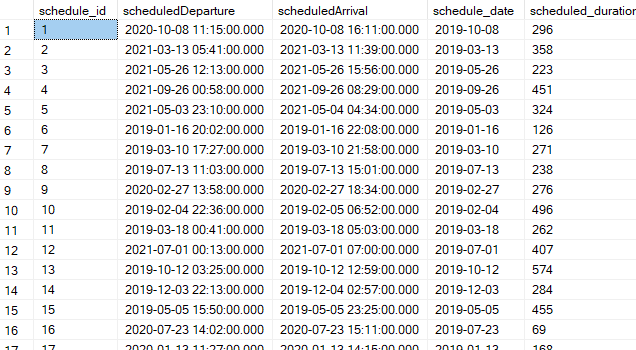


*Figure 14:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Farelg\_Dim*

*Data Flow Task 5(Schedule\_Dim)*

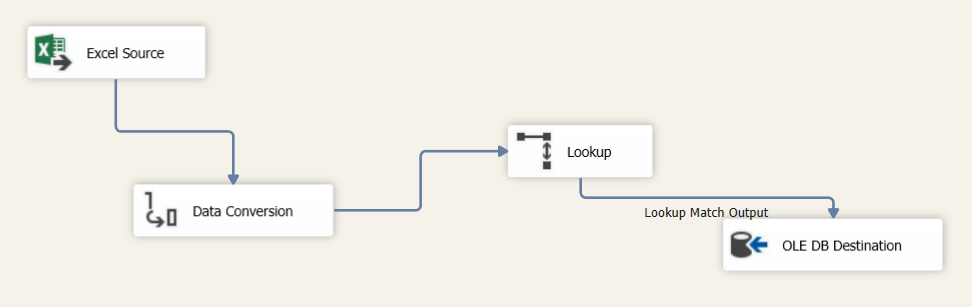
**

*Figure 15: Schedule\_Dim*

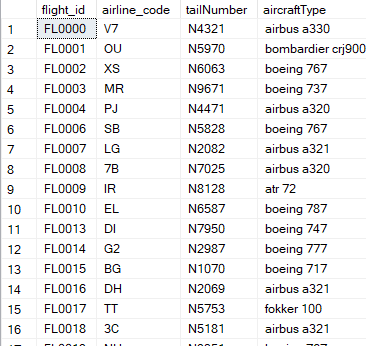


*Figure 16:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Schedule\_Dim*

*Data Flow Task 6(Flight\_Dim)*

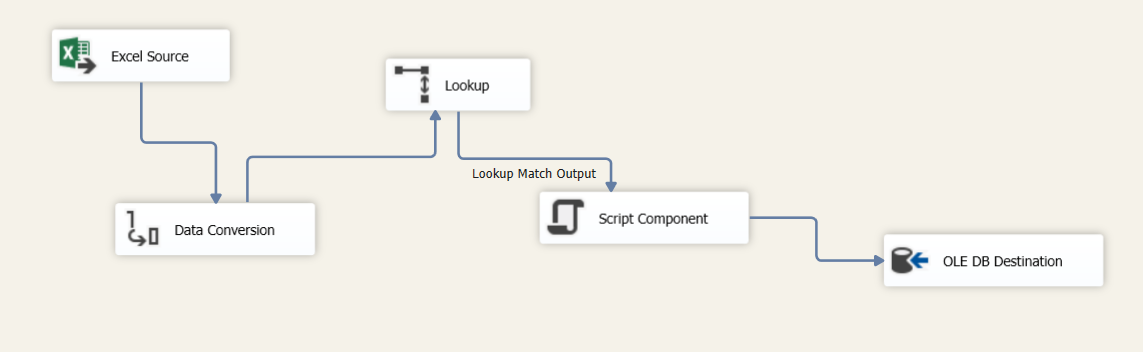
**

*Figure 17: Flight\_Dim*

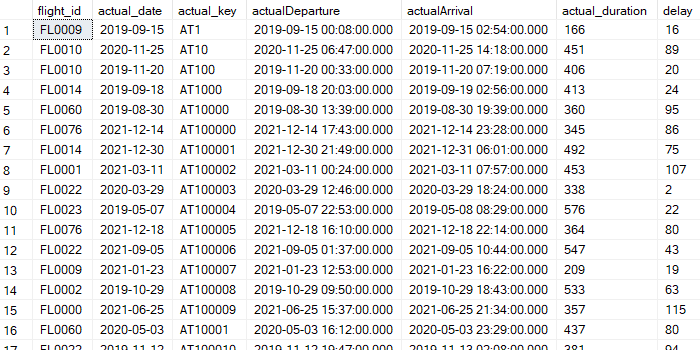


*Figure 18:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Flight\_Dim*

*Data Flow Task 7(Actual\_Dim)*

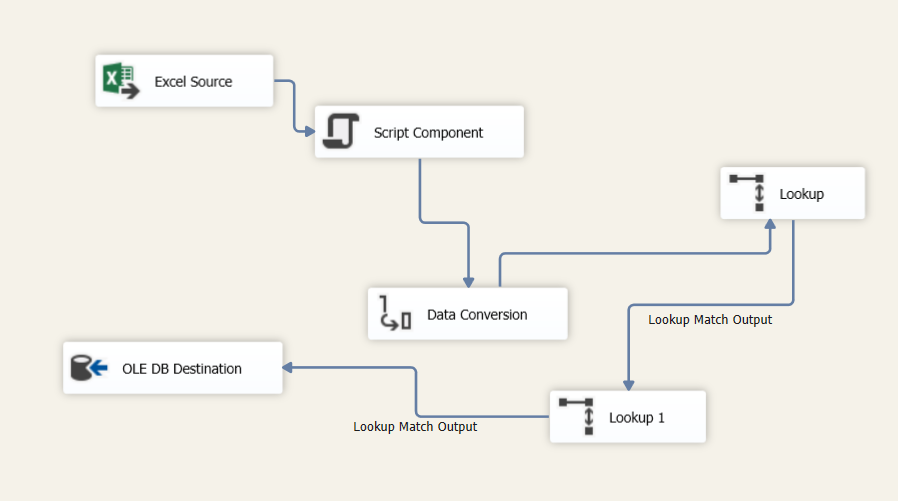
**

*Figure 19: Actual\_Dim*

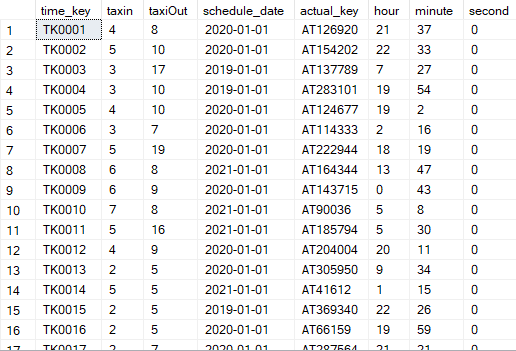


*Figure 20:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Actual\_Dim*

*Data Flow Task 8(Time\_Dim)*

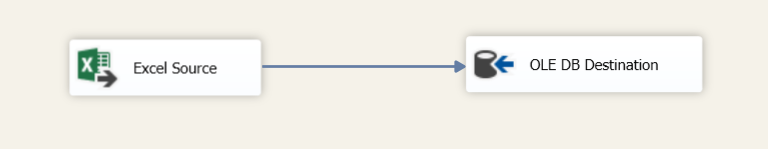
**

*Figure 21: Time\_Dim*



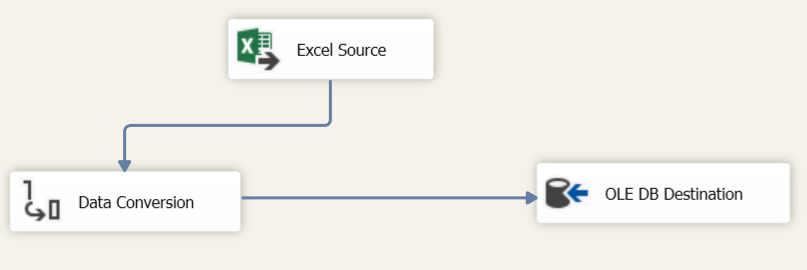
*Figure 22:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Time\_Dim*

*Data Flow Task 9(Date Dim)*

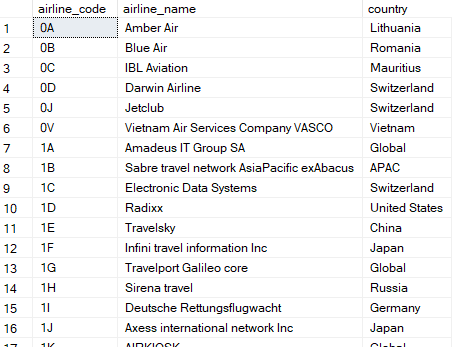


*Figure 23: Date\_Dim*

Data Flow Task 10(Airline Dim)

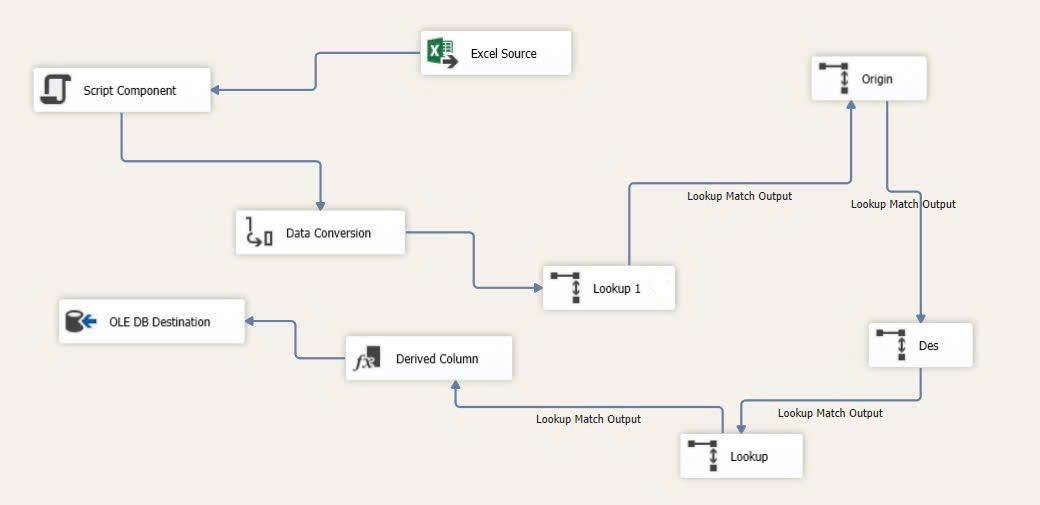


*Figure 24: Ariline\_Dim*

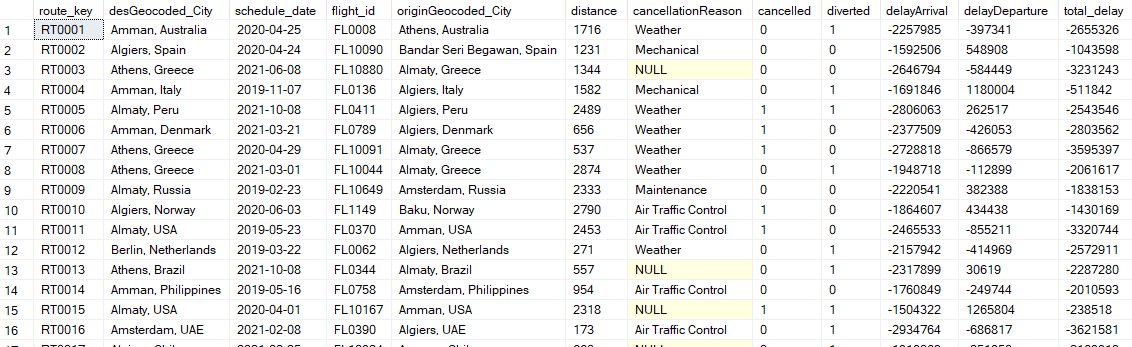


*Figure 25:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Ariline\_Dim*

*Data Flow Task 10(Route\_Fact)*

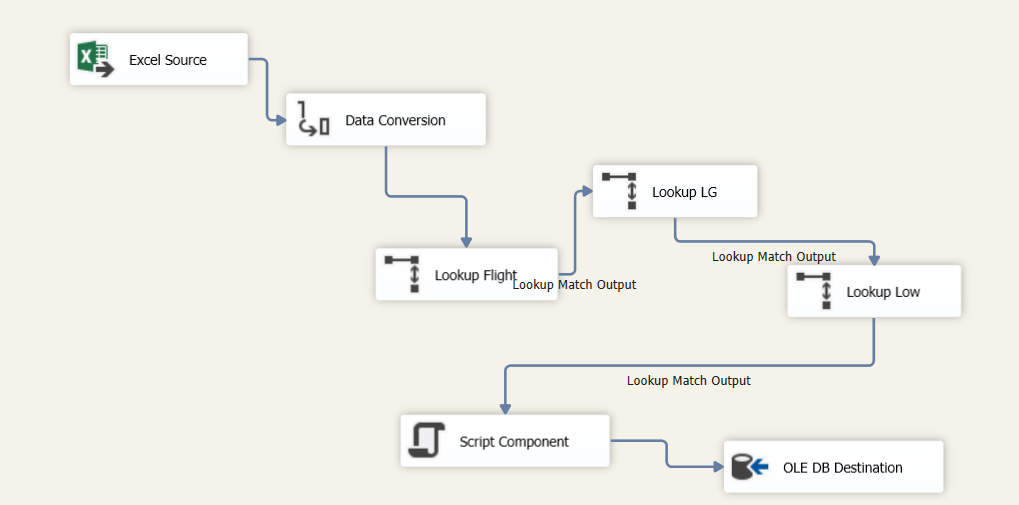
**

*Figure 26: Route\_Fact*

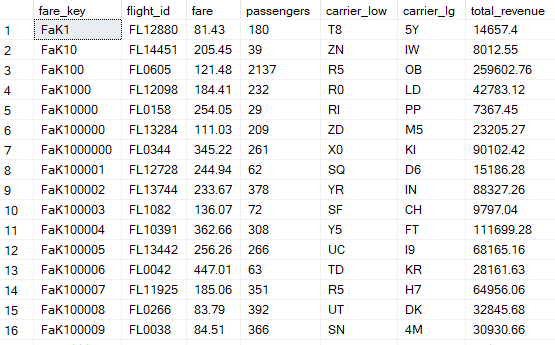
**

*Figure 27:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Route\_Fact*

*Data Flow Task 11(Fare\_Fact)*

**

*Figure 28: Fare\_Fact*

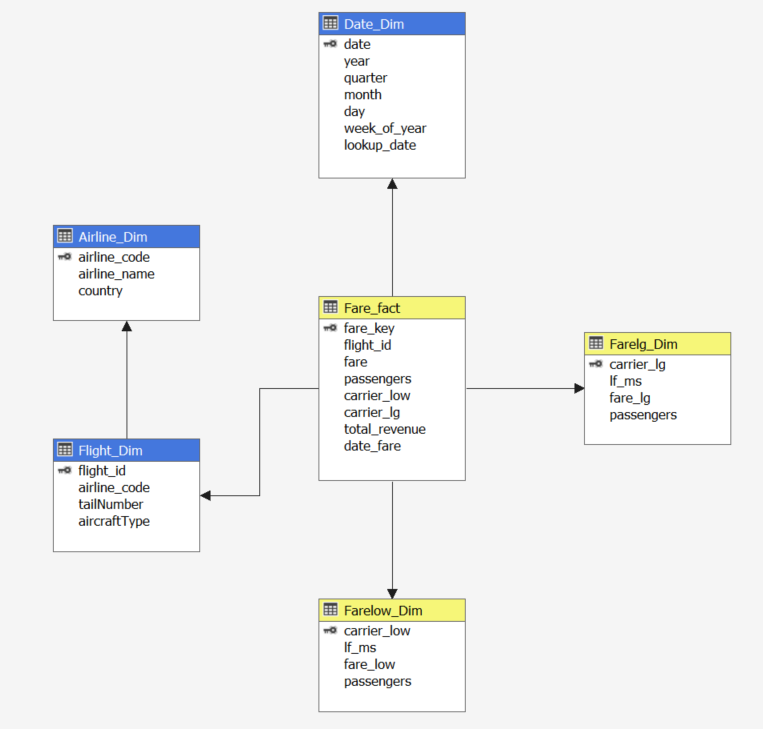


*Figure 29:* [*Data*](https://dictionary.cambridge.org/dictionary/english-vietnamese/data)[*from*](https://dictionary.cambridge.org/dictionary/english-vietnamese/from)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)*SQL*[*of*](https://dictionary.cambridge.org/dictionary/english-vietnamese/of)[*the*](https://dictionary.cambridge.org/dictionary/english-vietnamese/the)[*table*](https://dictionary.cambridge.org/dictionary/english-vietnamese/table) *Fare\_Fact*

# **CHAPTER 3: SQL SERVER ANALYSIS SERVICES**

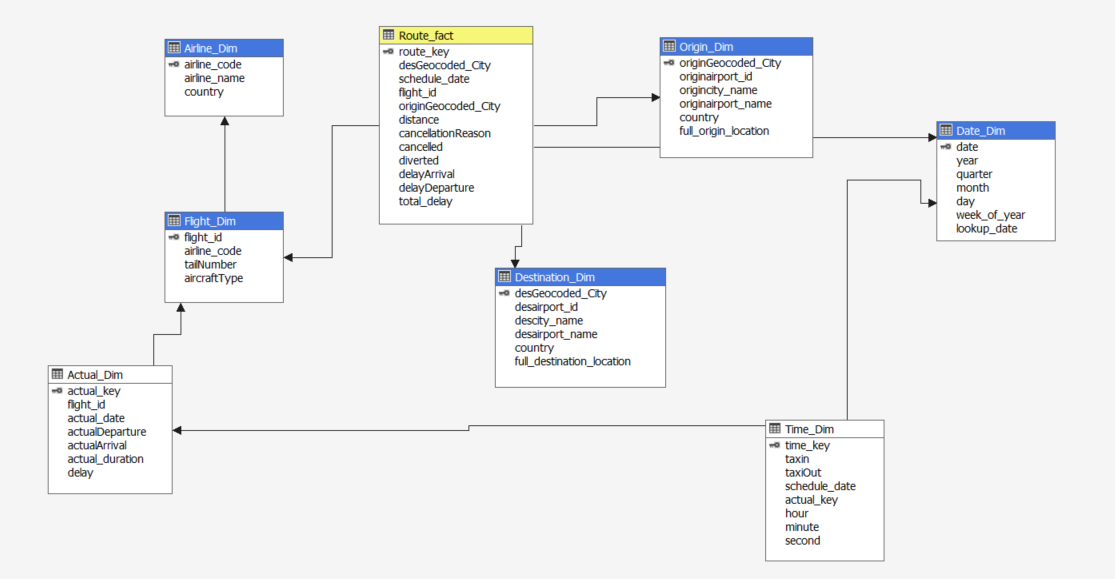
## **3.1. SSAS Overview**

**3.1.1. Cube Fare**



*Figure 30: Cubes Fare fact*

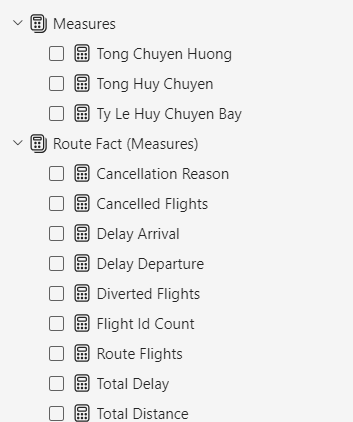
**3.1.2. Cubes route**



*Figure 31: Cubes Route fact*

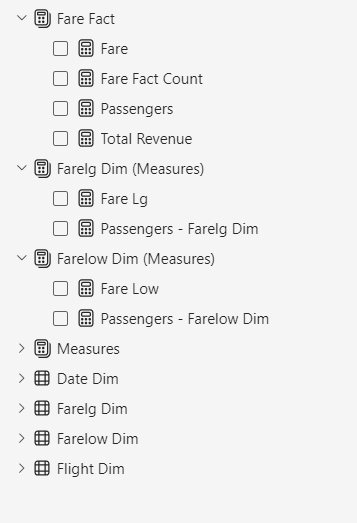
**3.2. Calculations measures**

**3.2.1. Fact measures**



*Figure 32: Measures Route fact*

**3.2.2. Cube measures**



*Figure 33: Measures Fare fact*

**3.3. MDX code for calculations**

**3.3.1. Departure and arrival delay analysis by airline**

SELECT NON EMPTY

{ [Measures].[Total Delay],

[Measures].[Delay Departure],

[Measures].[Delay Arrival] }

ON COLUMNS,

NON EMPTY

{ ([Flight Dim].[Airline Name].[Airline Name].ALLMEMBERS ) }

DIMENSION PROPERTIES MEMBER\_CAPTION, MEMBER\_UNIQUE\_NAME

ON ROWS

FROM [Route]

CELL PROPERTIES VALUE, BACK\_COLOR, FORE\_COLOR, FORMATTED\_VALUE, FORMAT\_STRING, FONT\_NAME, FONT\_SIZE, FONT\_FLAGS

**3.3.2. Analysis of flight cancellation reasons by reason**

SELECT NON EMPTY

{ [Measures].[Tong Huy Chuyen] }

ON COLUMNS,

NON EMPTY

{ ([Route Fact].[Cancellation Reason].[Cancellation Reason].ALLMEMBERS ) }

DIMENSION PROPERTIES MEMBER\_CAPTION, MEMBER\_UNIQUE\_NAME

ON ROWS

FROM [Route]

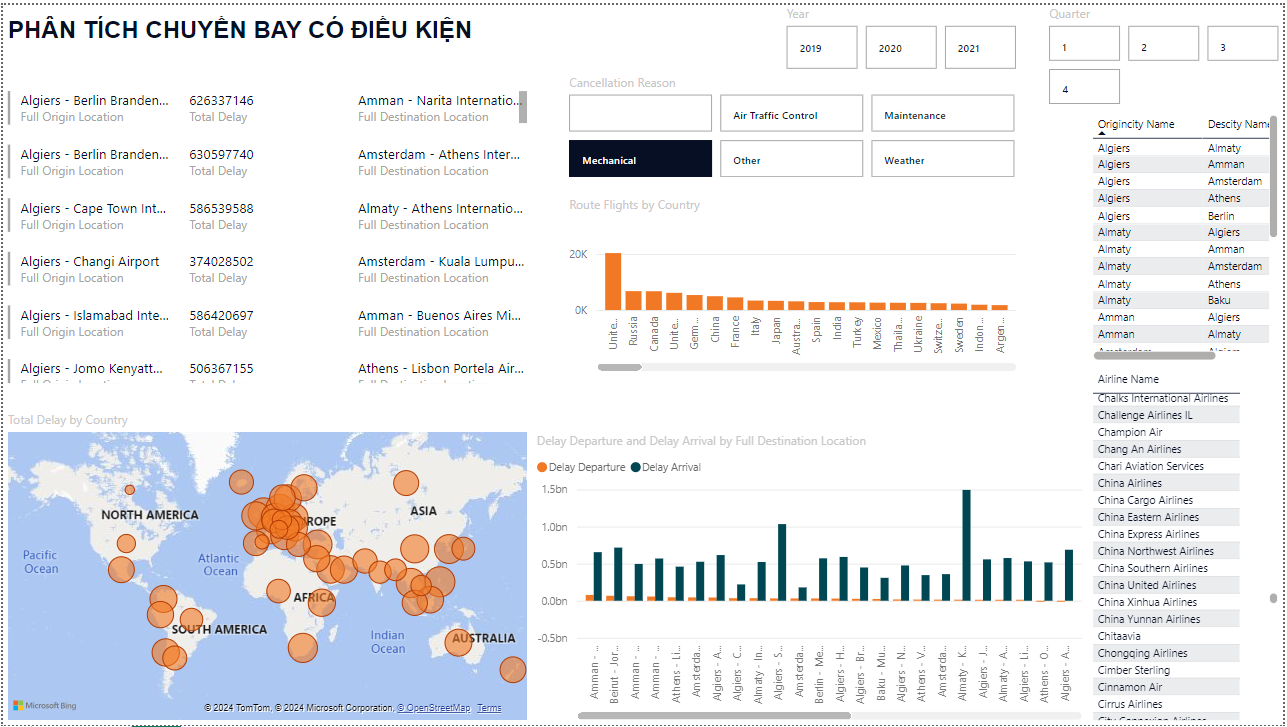
CELL PROPERTIES VALUE, BACK\_COLOR, FORE\_COLOR, FORMATTED\_VALUE, FORMAT\_STRING, FONT\_NAME, FONT\_SIZE, FONT\_FLAGS

)

# **CHAPTER 4: SQL SERVER REPORTING SERVICES**

**4.1. Dashboard by Route**

This dashboard provides a detailed overview of global flight cancellations, including trends over time, geographic distribution, and country-specific cause analysis. This information is useful for airlines and stakeholders in evaluating and improving flight performance.



*Figure 34: Dashboard Analysis 2019 -2021*

Explain:

**Cancellations:**

Air traffic control is the main cause of cancellations on most routes.

Some routes are significantly affected by Weather (e.g. Algiers - Almaty), Maintenance (e.g. Algiers - Amsterdam) or Mechanical (e.g. Algiers - Amman).

Other factors account for a small proportion of cancellations.

**Routes:**

The chart focuses on routes with Algiers as their destination or departure, suggesting that there may be issues with the airport or the airlines operating there.

Algiers - Almaty is the route with the highest cancellation rate, which requires further investigation to determine the specific cause.

**Delays**:

The bottom right chart compares departure and arrival delays on each route.

There are clear differences between routes. Some routes are frequently delayed in departure (e.g. Algiers - Amman), while others are delayed in arrival (e.g. Algiers - Bandar Seri Begawan).

**Geographic distribution:**

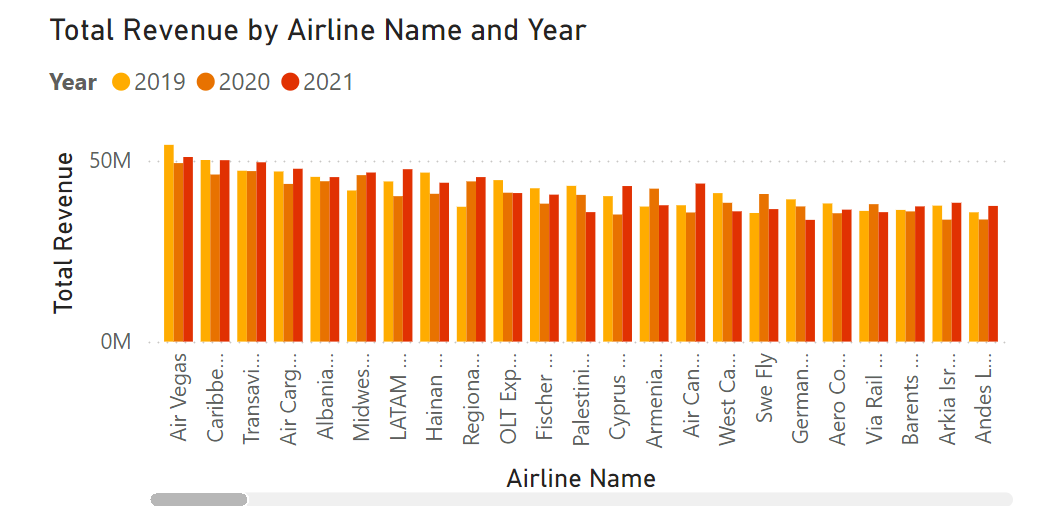
The world map shows the regions where flights are affected. Europe, North America and parts of Asia are the areas with the highest concentration of red spots, indicating a high number of delayed or canceled flights.

Flight traffic:

The "Route Flights by Country" chart shows the number of flights between Algiers and other countries.

**4.2. Fare Analysis**

The chart provides important information about the impact of the pandemic on the airline industry and shows the difference in the recovery of airlines between 2019 and 2021.



*Figure 35:*  Airline revenue chart 2019 - 2021

Most airlines saw their revenue decline sharply from 2019 to 2020, reflecting the severe impact of the COVID-19 pandemic on the airline industry.

Major airlines such as American Airlines, Delta, United, Southwest and JetBlue all reported significant revenue declines in 2020.

Some airlines such as Alaska, Allegiant, Spirit and Frontier Airlines saw their revenue declines be smaller, suggesting they were less affected than the major airlines.

Most airlines’ 2021 revenue forecasts have yet to fully recover from pre-pandemic levels, suggesting the industry is still in a slow recovery.

Some airlines such as Spirit, Frontier and Allegiant have 2021 revenue close to or exceeding 2019 levels, suggesting they are more resilient than the major airlines.

# **CHAPTER 5: CONCLUDE**

## **5.1 Results achieved:**

During the last semester, the group learned and applied knowledge about building data warehouses and OLAP and achieved the following results:

* Understand the basic concepts of data warehouse and OLAP, and the necessary properties of a data warehouse
* Strong knowledge and ability to apply and build a complete data warehouse used for data mining
* Equipped with knowledge of SSIS, SSAS, SSRS tools
* Build a complete data warehouse
* Presenting query optimization
* Built a complete data warehouse with the topic of Survey Results of developers on Stackoverflow, kaggle.

## **5.2 Limitations:**

- Difficulties in data cleaning and preparation: The process of cleaning and normalizing data can be complex and time consuming, especially with large and heterogeneous data.

- Query optimization: Query optimization for a data warehouse can be difficult when handling large or complex data, affecting system performance.

- Limited in-depth knowledge: A lack of in-depth knowledge of advanced techniques such as complex analytic, machine learning, or performance optimization can reduce the ability to build an effective data warehouse solution.

# **REFERENCES**

1. SQL Server 2012 Tutorials: Analysis Services - Multidimensional Modeling [<https://courses.uit.edu.vn/pluginfile.php/509409/mod_resource/content/1/2.SQL%20Server%202012%20Tutorials%20-%20Analysis%20Services%20Multidimensional%20Modeling%20%281%29.pdf> ]
2. Slide bài giảng, tài liệu giảng dạy của ThS. Trần Thanh Liêm
3. SQL Server Integration Service - SSIS (03/01/2023) [<https://learn.microsoft.com/en-us/sql/integration-services/lesson-1-1-creating-a-new-integration-services-project?view=sql-server-ver16> ]
4. Analysis Services - SSAS [<https://learn.microsoft.com/en-us/analysis-services/multidimensional-models/create-an-analysis-services-project-ssdt?view=asallproducts-allversions> ]
5. Jin Do (21/03/2023), Hướng dẫn tạo project SSIS SSAS đơn giản với SQL server và Visual Studio 2022 [Hướng dẫn tạo project SSIS SSAS đơn giản với SQL server và Visual Studio 2022](https://youtu.be/Vz8ttUSKpYo?si=4882aAVlURXXLg6B)
6. Power BI Documentation (21-23/03/2024) [<https://learn.microsoft.com/en-us/power-bi/> ]