

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN  
KHOA CÔNG NGHỆ THÔNG TIN

-----□□-----



**BÁO CÁO ĐỒ ÁN THỰC HÀNH**

**GD1: ETL source-> stage-> NDS**

**Môn: Hệ thống thông tin phục vụ trí tuệ kinh doanh**

*29/9/2024 – 23/11/2024*

**MÃ HỌC PHẦN: CSC12107**

**Nhóm: 3**

**Thành viên:**

<b>ID</b>	<b>Họ tên</b>
19127536	Võ Lâm Hải Quốc
21127211	Nguyễn Vũ Tường An
21127450	Võ Trung Tín
21127699	Lô Thủy Tiên

**Giảng viên:**

ThS. Hồ Thị Hoàng Vy

ThS. Tiết Gia Hồng

ThS. Nguyễn Ngọc Minh Châu

**Báo cáo:**

*Lô Thủy Tiên*

Thành phố Hồ Chí Minh – 2024

---

## MỤC LỤC

<b>THÔNG TIN VỀ ĐỒ ÁN.....</b>	<b>3</b>
<b>THÔNG TIN NHÓM.....</b>	<b>5</b>
<b>I. MÔ TẢ DỮ LIỆU NGUỒN:.....</b>	<b>6</b>
<b>II. THIẾT KẾ CẤU TRÚC CÀI ĐẶT:.....</b>	<b>8</b>
<b>III. GIAI ĐOẠN SOURCE → STAGE:.....</b>	<b>9</b>
1. Profiling dữ liệu:.....	9
2. Cấu trúc Metadata:.....	11
3. Data Flow:.....	12
<b>IV. GIAI ĐOẠN : STAGE → NDS.....</b>	<b>13</b>
1. Cấu trúc NDS:.....	13
2. Chi tiết từng bảng trong NDS:.....	13
a. Bảng NDS_States:.....	13
b. Bảng NDS_Counties:.....	14
c. Bảng NDS_AirData:.....	15
<b>PHÂN CÔNG CÔNG VIỆC.....</b>	<b>17</b>
<b>ĐÁNH GIÁ THÀNH VIÊN.....</b>	<b>20</b>
<b>TÀI LIỆU THAM KHẢO.....</b>	<b>21</b>
<i>Công cụ và phần mềm hỗ trợ:.....</i>	<i>21</i>
<i>Tài liệu tham khảo:.....</i>	<i>21</i>

---

## THÔNG TIN VỀ ĐỒ ÁN

**Mã học phần:** CSC12107

**Tên học phần:** Hệ thống thông tin phục vụ trí tuệ kinh doanh

**Tên :** Đồ án thực hành - 2425.BI.DATH

**Hình thức:**

- Báo cáo (.doc, ppt)
- Source demo
- Video demo

**Mô tả:**

Dữ liệu sẽ được phân tích để phát hiện các xu hướng và mẫu trong chất lượng không khí của Hoa Kỳ từ năm 2021 đến 2023. Đồ án yêu cầu xây dựng một kho dữ liệu (DW) từ các nguồn dữ liệu thô, sau đó triển khai các quy trình ETL, thiết kế mô hình OLAP, thực hiện phân tích dữ liệu và tạo các báo cáo.

- **Dữ liệu:** Dữ liệu chất lượng không khí hàng ngày của EPA, phân chia theo quận từ năm 2021 đến 2023, kết hợp với dữ liệu địa lý và định nghĩa phân loại AQI.
- **Thiết kế kho dữ liệu:** Xây dựng các bảng chiều như Địa lý (State > County), Thời gian (Year > Quarter > Month > Day) và các chiều khác để phục vụ yêu cầu báo cáo.
- **Yêu cầu báo cáo và phân tích:**
  - Tạo các báo cáo biểu đồ để trình bày sự biến động AQI theo thời gian.
  - Phân tích các câu hỏi mở để đưa ra đánh giá về biến động AQI, ngày ô nhiễm nặng, và các yếu tố có ảnh hưởng đến chất lượng không khí.
  - Tích hợp bản đồ khu vực để biểu diễn trung bình AQI qua các khu vực.

- **Khai phá dữ liệu:** Ứng dụng các mô hình khai phá dữ liệu nhằm dự đoán chất lượng không khí trong các kỳ tới (Q1-2024, tháng 01-2024, ...), giải thích thuật toán, lý do chọn lựa, và trình bày kết quả.
- **Tổng quan và kết quả:** Đưa ra đánh giá tổng quan về chất lượng không khí ở các quận Hoa Kỳ vào năm 2023. Kết luận các thành tựu đạt được trong đồ án, cũng như đề xuất hướng cải thiện cho chất lượng không khí trong tương lai.

**Mục tiêu:**

- Thiết kế mô hình dữ liệu: Sử dụng mô hình ngôi sao hoặc bông tuyết để đáp ứng yêu cầu phân tích.
- Quy trình ETL: Sử dụng công cụ SSIS để thu thập, làm sạch và tích hợp dữ liệu từ các nguồn khác nhau vào kho dữ liệu.
- OLAP và Trục quan hóa: Sử dụng công cụ SSAS để khai thác các công nghệ OLAP cơ bản và tạo báo cáo bằng SSRS hoặc Excel.
- Khai phá dữ liệu: Áp dụng các thuật toán khai phá dữ liệu với SSAS để phân tích và dự đoán chất lượng không khí.

**Giảng viên phụ trách: Cô Hồ Thị Hoàng Vy, Cô Tiết Gia Hồng, Cô Nguyễn Ngọc Minh Châu**

---

**THÔNG TIN NHÓM**

Nhóm: 3

<b>MSSV</b>	<b>Họ tên</b>	<b>Email</b>	<b>Ghi chú</b>
19127536	Võ Lâm Hải Quốc	vlhquoc19@clc.fitus.edu.vn	
21127211	Nguyễn Vũ Tường An	nvtan21@clc.fitus.edu.vn	
21127450	Võ Trung Tín	vttin21@clc.fitus.edu.vn	
<b>21127699</b>	<b>Lô Thủy Tiên</b>	<b>littien21@clc.fitus.edu.vn</b>	

## I. MÔ TẢ DỮ LIỆU NGUỒN:

Mô tả dữ liệu: Dữ liệu trong đồ án gồm các tệp chứa thông tin về chất lượng không khí tại các quận của Hoa Kỳ, do Cơ quan Bảo vệ Môi trường Hoa Kỳ (EPA) cung cấp. Mục tiêu chính là tạo kho dữ liệu (DW) từ những dữ liệu này, phân tích và xác định xu hướng chất lượng không khí qua ba năm (2021-2023). Các nguồn dữ liệu chính bao gồm:

- **US Daily AQI Report by County (2021-2023):**

- Dữ liệu: 3 tệp CSV chứa thông tin chỉ số chất lượng không khí (AQI) được ghi nhận hàng ngày tại các quận của 10 tiểu bang từ 2021 đến 2023.
- Source:

Tên thuộc tính	Ý nghĩa
State Name	Tên của tiểu bang tại Hoa Kỳ nơi dữ liệu AQI được ghi nhận.
County Name	Tên của quận trong tiểu bang nơi chỉ số AQI được đo.
State Code	Mã FIPS đại diện duy nhất cho mỗi tiểu bang ở Hoa Kỳ.
County Code	Mã FIPS đại diện duy nhất cho mỗi quận trong tiểu bang.
Date	Ngày khi chỉ số AQI được ghi nhận.
AQI	Chỉ số AQI biểu thị mức độ ô nhiễm không khí, với các giá trị cao biểu thị chất lượng không khí kém hơn.
Category	Phân loại AQI theo mức độ ảnh hưởng đến sức khỏe (Tốt, Trung bình, Không lành mạnh, v.v.).
Defining Parameter	Chất gây ô nhiễm chính (ví dụ: ozone, PM2.5) ảnh hưởng lớn nhất đến chỉ số AQI tại địa điểm đo.
Defining Site	Tên hoặc mã của trạm đo AQI nơi dữ liệu trong ngày được lấy.
Number of Sites Reporting	Tổng số trạm đo trong quận báo cáo dữ liệu trong ngày đó.

Created	Ngày và giờ khi bản ghi AQI này được tạo ra trong hệ thống.
Last Updated	Ngày và giờ khi bản ghi AQI này được cập nhật lần cuối trong hệ thống.

- **Dữ liệu Địa lý**

- Dữ liệu: Tập CSV chứa thông tin về các quận tại Hoa Kỳ.
- Source: **(2B)uscounties.csv**

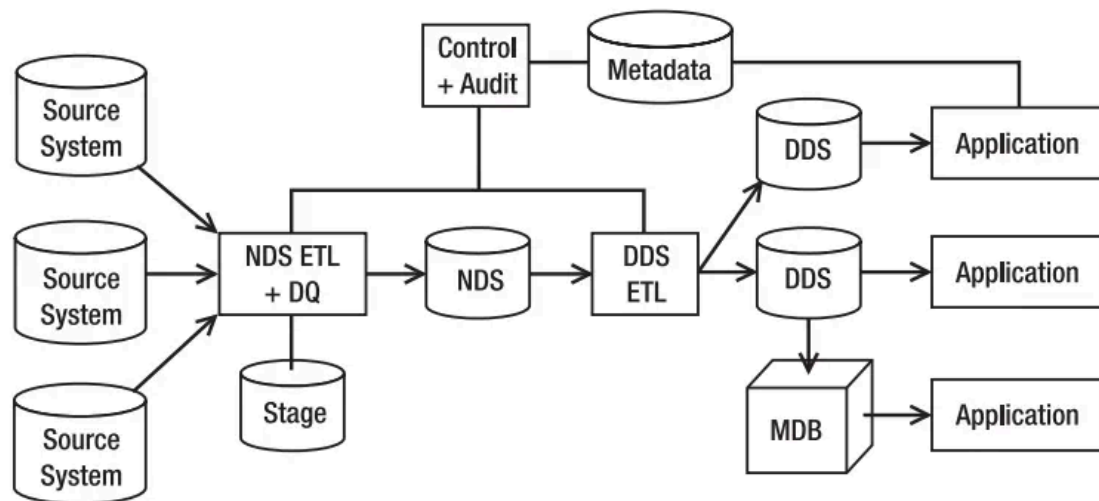
Tên thuộc tính	Ý nghĩa
County	Tên của quận.
County ASCII	Tên của quận không có ký tự đặc biệt, giúp dễ dàng xử lý trong các hệ thống lưu trữ dữ liệu.
County Full	Tên đầy đủ của quận, có thể bao gồm cả phần miêu tả hoặc phần mở rộng của tên.
County FIPS	Mã FIPS duy nhất xác định quận trong hệ thống phân cấp liên bang tại Hoa Kỳ.
State ID	Mã định danh của tiểu bang mà quận thuộc về.
State Name	Tên của tiểu bang mà quận thuộc về.
Lat	Vĩ độ của vị trí trung tâm quận, dùng để định vị trên bản đồ.
Lng	Kinh độ của vị trí trung tâm quận, dùng để định vị trên bản đồ.
Population	Dân số của quận, biểu thị tổng số cư dân sinh sống trong khu vực đó.

- **AQI Category Definition**

AQI Basics for Ozone and Particle Pollution			
Daily AQI Color	Levels of Concern	Values of Index	Description of Air Quality
Green	Good	0 to 50	Air quality is satisfactory, and air pollution poses little or no risk.
Yellow	Moderate	51 to 100	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution.
Orange	Unhealthy for Sensitive Groups	101 to 150	Members of sensitive groups may experience health effects. The general public is less likely to be affected.
Red	Unhealthy	151 to 200	Some members of the general public may experience health effects; members of sensitive groups may experience more serious health effects.
Purple	Very Unhealthy	201 to 300	Health alert: The risk of health effects is increased for everyone.
Maroon	Hazardous	301 and higher	Health warning of emergency conditions: everyone is more likely to be affected.

## II. THIẾT KẾ CẤU TRÚC CÀI ĐẶT:

Lựa chọn kiến trúc cài đặt NDS + DDS, cụ thể như sau:



**Figure 2-5.** *NDS + DDS data flow architecture*

Dữ liệu từ các nguồn sẽ được rút trích và tải vào **Stage** theo phương pháp Incremental trước khi đưa vào **NDS**. Sau khi dữ liệu ở **Stage**, tiến hành **Profiling** để kiểm tra và làm sạch, biến đổi dữ liệu nếu cần. **NDS** là cơ sở dữ liệu chuẩn hóa



lưu trữ toàn bộ dữ liệu. Dữ liệu từ **NDS** sẽ được rút trích ra các **DDS** hoặc **Data Mart** phục vụ cho phân tích.

### III. GIAI ĐOẠN SOURCE → STAGE:

#### 1. Profiling dữ liệu:

[chi tiết Profiling dữ liệu](#)

*File 10\_state\_api\_2021*

Tên thuộc tính	SL giá trị phân biệt	Tỉ lệ NULL	Ghi chú
State Name	277	0%	
County Name	6	0%	
State Code	116	0%	
County Code	182	0%	
Date	365	0%	<b>Ngày (định dạng: mm/dd/yyyy)</b>
AQI	365	0%	Chỉ số chất lượng không khí: Giá trị AQI (0-500). Xem bảng phân loại
Category	5	0%	Phân loại chất lượng không khí (ví dụ: Good, Moderate, Unhealthy).
Defining Parameter	616	0%	
Defining Site	1	0%	
Number of Sites Reporting	24	0%	
Created	10	0%	
Last Updated	10	0%	

*File 10\_state\_api\_2022*

Tên thuộc tính	SL giá trị phân biệt	Tỉ lệ NULL	Ghi chú
State Name	277	0%	

County Name	6	0%	
State Code	116	0%	
County Code	190	0%	
Date	365	0%	<b>Ngày (định dạng: mm/dd/yyyy)</b>
AQI	365	0%	Chỉ số chất lượng không khí: Giá trị AQI (0-500). Xem bảng phân loại
Category	5	0%	Phân loại chất lượng không khí (ví dụ: Good, Moderate, Unhealthy).
Defining Parameter	620	0%	
Defining Site	1	0%	
Number of Sites Reporting	23	0%	
Created	18	0%	
Last Updated	18	0%	

***File 10\_state\_api\_2023***

<b>Tên thuộc tính</b>	<b>SL giá trị phân biệt</b>	<b>Tỉ lệ NULL</b>	<b>Ghi chú</b>
State Name	277	0%	
County Name	6	0%	
State Code	115	0%	
County Code	187	0%	
Date	365	0%	<b>Ngày (định dạng: mm/dd/yyyy)</b>
AQI	365	0%	Chỉ số chất lượng không khí: Giá trị AQI (0-500). Xem bảng phân loại
Category	5	0%	Phân loại chất lượng không khí (ví dụ: Good, Moderate, Unhealthy).
Defining	595	0%	

Parameter			
Defining Site	1	0%	
Number of Sites Reporting	23	0%	
Created	16	0%	
Last Updated	16	0%	

**(2B)uscounties.csv**

Tên thuộc tính	SL giá trị phân biệt	Tỉ lệ NULL	Ghi chú
County	1838	0%	
County ASCII	1838	0%	
County Full	3144	0%	
County FIPS	1882	0%	
State ID	51	0%	
State Name	51	0%	
Lat		0%	
Lng		0%	
Population	3080	0%	

## 2. Cấu trúc Metadata:

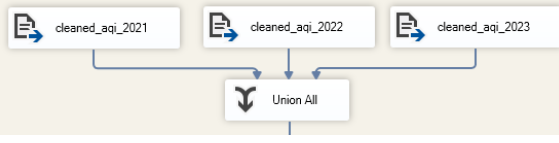
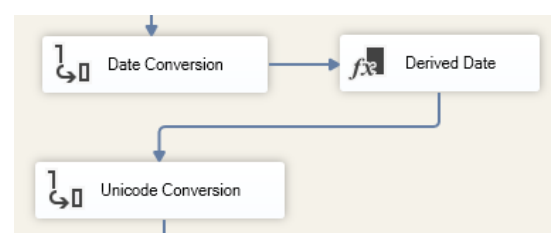
Tên cột	Kiểu dữ liệu	Ý nghĩa	Ghi chú
ID	Số	Khóa chính tự tăng	
Name	Chuỗi	Tên data flow	
LSET	Ngày tháng	Thời gian rút trích thành công gần nhất	
CET	Ngày tháng	Thời gian bắt đầu rút trích	

### 3. Data Flow:

Để tăng tốc độ rút trích từ nguồn, các bảng ở Stage có cấu trúc tương tự với nguồn. Quá trình ETL sẽ thực hiện như sau:

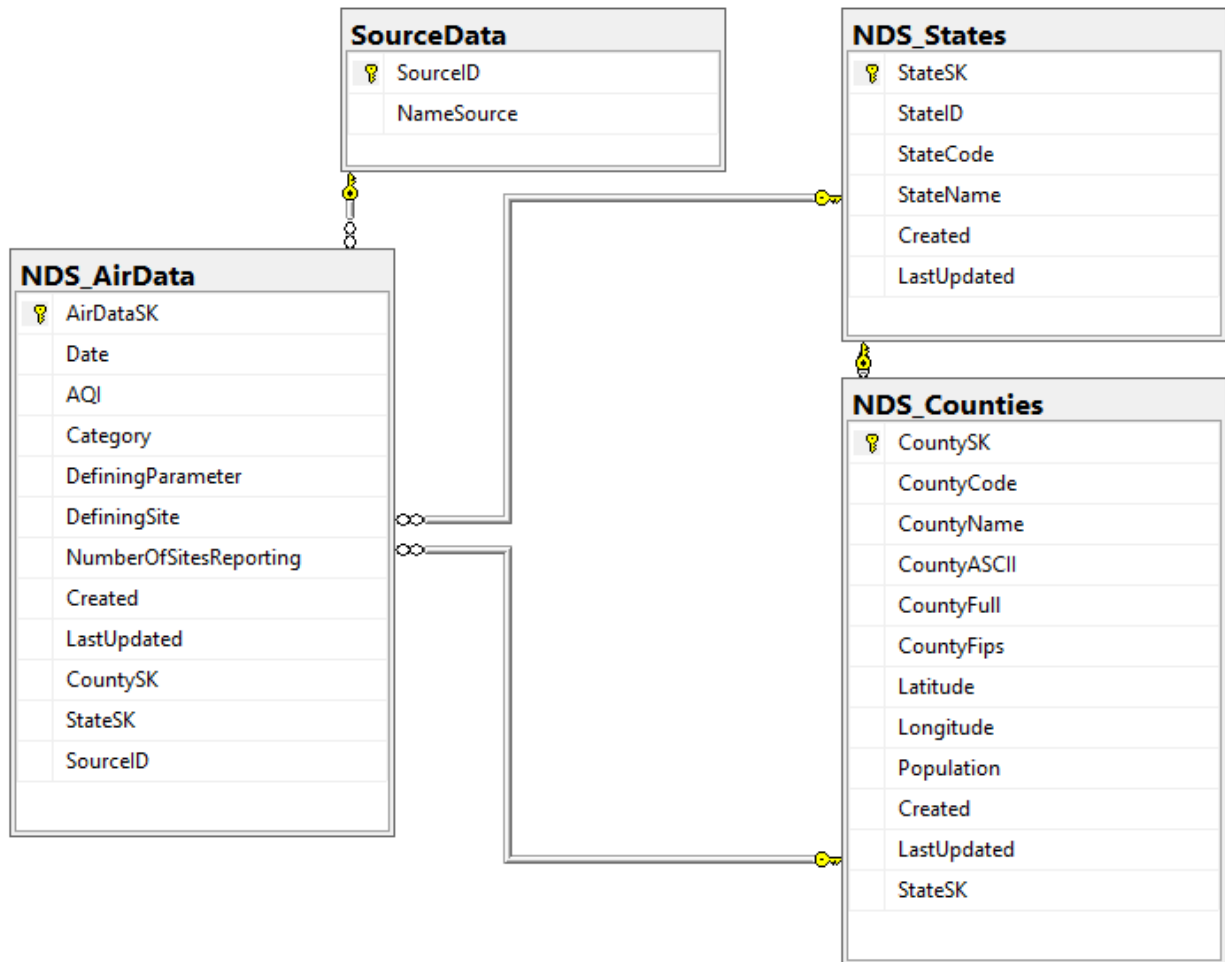
1. Ghi lại thời điểm **CET** tại Metadata khi bắt đầu ETL từ nguồn vào Stage.
2. Lấy giá trị **LSET** của bảng chuẩn bị nạp vào Stage từ Metadata.
3. Chọn các dòng dữ liệu được tạo ra hoặc cập nhật từ **LSET** đến **CET** từ nguồn và đổ vào Stage.
4. Dọn dẹp dữ liệu cũ trong Stage trước khi tải dữ liệu mới.
5. Tiến hành **load** dữ liệu mới vào Stage.
6. **Profiling** dữ liệu vừa được load vào Stage.
7. Cập nhật **LSET** bằng giá trị **CET** hiện tại vào Metadata.

Danh sách các transformation đã sử dụng

Transformation	Mục đích sử dụng	Ghi chú
Union All	Gộp nhiều luồng dữ liệu thành một. - Thực hiện để gộp 3 file ( <i>File 10_state_api_2021 + File 10_state_api_2022 + File 10_state_api_2023</i> )	
Data Conversion	Chuyển đổi kiểu dữ liệu của cột. Thống nhất kiểu định dạng của DATE và chuyển các chuỗi sang UNICODE STRING	

## IV. GIAI ĐOẠN : STAGE → NDS

### 1. Cấu trúc NDS:



### 2. Chi tiết từng bảng trong NDS:

#### a. Bảng SourceData:

STT	Tên thuộc tính	Mô tả	Transformation rules	Nguồn	SCD
1	SourceID	Mã của source (Natural Key)			
2	SourceName	Tên của Source			

**b. Bảng NDS\_States:**

STT	Tên thuộc tính	Mô tả	Transformation rules	Nguồn	SCD
1	StateSK	Khóa thay thế duy nhất cho mỗi stage	Tạo từ StageID_NK bằng cách sinh ID duy nhất trong NDS (Khóa tự tăng)		
2	StateID	Mã của bang (Natural Key)		state_api	
3	StateCode	Mã của bang		uscounties.csv	2
4	StateName	Tên của bang		state_api	2
5	Created	Ngày tạo bản ghi		ETL Process	
6	LastUpdated	Ngày cập nhật bản ghi		ETL Process	
7	SourceID	ID của nguồn dữ liệu		ETL Process	

**c. Bảng NDS\_Counties:**

STT	Tên thuộc tính	Mô tả	Transformation rules	Nguồn	SCD
1	CountySK	Khóa thay thế duy nhất cho mỗi County	Tạo từ CountyName bằng cách sinh ID duy nhất trong NDS (Khóa tự tăng)		
2	CountyCode	Mã của quận		state_api	1
3	CountyName	Tên của quận		uscounties.csv	2
4	CountyASCII	Tên ASCII của quận		state_api	2
5	CountyFull	Tên đầy đủ của quận		state_api	2

6	CountyFips	Mã FIPS của quận (Natural Key)		state_api	
7	Latitude	Vĩ độ của quận		state_api	
8	Longitude	Kinh độ của quận		state_api	
9	Population	Dân số của quận		state_api	
10	Created	Ngày tạo bản ghi		ETL Process	
11	LastUpdated	Ngày cập nhật bản ghi		ETL Process	
12	StateSK	Khóa thay thế của Stage mà Counties thuộc	Ánh xạ từ [state_id] trong "Counties_STAGE" với StateSK trong <i>NDS_Counties</i>		
13	SourceID	ID của nguồn dữ liệu		ETL Process	

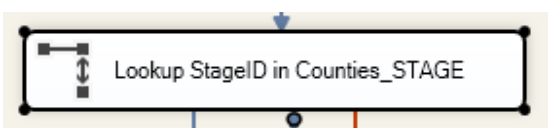
**d. Bảng NDS\_AirData:**

STT	Tên thuộc tính	Mô tả	Transformation rules	Nguồn	SCD
1	AirDataSK	Khóa thay thế duy nhất cho mỗi AirData	Tạo từ Date và CountyName bằng cách sinh ID duy nhất trong NDS (Khóa tự tăng)		
2	Date	Ngày (Natural Key)	Ngày, chuyển về chung một định dạng (định dạng: yyyy/mm/dd)	state_api	
3	AQI	Chỉ số chất lượng không khí (Air Quality Index)	Chỉ số chất lượng không khí: Giá trị AQI (0-500). Xem bảng phân loại	state_api	


4	Category	Loại chỉ số AQI	Phân loại chất lượng không khí (ví dụ: Good, Moderate, Unhealthy).	state_api	
5	DefiningParameter	Thông số xác định		state_api	
6	DefiningSite	Mã của địa điểm xác định		state_api	
7	NumberOfSites Reporting	Số lượng địa điểm báo cáo		state_api	
8	Created	Ngày tạo bản ghi		ETL Process	
9	LastUpdated	Ngày cập nhật bản ghi		ETL Process	
10	CountySK	Khóa thay thế của County mà AirData thuộc	Ánh xạ từ [county_Name] trong "Air_Quality_STAGE" với CountySK trong <b>NDS_AirData</b>	ETL Process	
11	StateSK	Khóa thay thế của Stage mà County thuộc	Ánh xạ từ [State Name] trong "Air_Quality_STAGE" với StateSK trong <b>NDS_AirData</b>		
12	SourceID	ID của nguồn dữ liệu		ETL Process	

### 3. Data Flow:

Danh sách các transformation đã sử dụng

Transformation	Mục đích sử dụng	Ghi chú
Lookup	Tra cứu dữ liệu từ bảng khác và thêm giá trị cột mới. (Sử dụng để liên kết với bảng Stage lấy khóa NK hoặc tạo khóa ngoại trong NDS)	



Derived Column	Tính toán giá trị mới hoặc sửa đổi giá trị cột.	 ADD SourceID column
----------------	---	---

## PHÂN CÔNG CÔNG VIỆC

STT	Công việc	Người thực hiện	Ngày hoàn thành
<b>GIAI ĐOẠN 1: SOURCE</b>			
1	Tạo template Báo cáo, phân công	Lô Thủy Tiên	00:00 AM 27/10/2024
2	<b>MÔ TẢ DỮ LIỆU NGUỒN: <i>Phân tích chi tiết data sources:</i></b> <ul style="list-style-type: none"> <li>- US Daily AQI report (2021-2023)</li> <li>- Geography data (uscounties.csv)</li> <li>- AQI Category Definition</li> </ul>	Lô Thủy Tiên	00:00 AM 27/10/2024
<b>GIAI ĐOẠN 2: SOURCE → STAGE</b>			
3	<b><i>Profiling dữ liệu:</i></b> <ul style="list-style-type: none"> <li>- Thực hiện data profiling ban đầu</li> <li>- Hoàn thành báo cáo về profiling</li> </ul>	Võ Trung Tín	00:00 AM 27/10/2024
4	<b><i>Data Cleaning &amp; Transformation</i></b> <ul style="list-style-type: none"> <li>- Chuẩn hóa định dạng ngày tháng</li> <li>- Xử lý missing values</li> <li>- Validate data ranges với AQI Category Definition</li> </ul>	Nguyễn Vũ Tường An	00:00 AM 27/10/2024
5	<b><i>Thiết kế schema:</i></b> <ul style="list-style-type: none"> <li>- Thiết kế schema cho staging area</li> </ul>	Nguyễn Vũ Tường An	00:00 AM 27/10/2024
6	<b><i>Xây dựng SSIS package để load dữ liệu từ:</i></b>	Lô Thủy Tiên (Thực hiện chính)	00:00 AM 30/10/2024

	<ul style="list-style-type: none"> <li>- 3 file CSV Air Quality Data (2021-2023) vào staging</li> <li>- File uscounties.csv vào staging</li> </ul> <p><i>Xác định và xử lý các vấn đề về data quality trong staging</i></p>		
7	<b><i>Xây dựng SSIS package</i></b>	Võ Lâm Hải Quốc Nguyễn Vũ Tường An Võ Trung Tín (Tham gia google meet)	00:00 AM 30/10/2024
<b>GIAI ĐOẠN 3: STAGE → NDS</b>			
8	<b><i>Thiết kế schema:</i></b> <ul style="list-style-type: none"> <li>- Thiết kế schema cho NDS</li> <li>- Vẽ sơ đồ <b>Cấu trúc NDS</b></li> </ul>	Võ Lâm Hải Quốc	00:00 AM 30/10/2024
9	<b><i>Báo cáo:</i></b> <ul style="list-style-type: none"> <li>- Hoàn thành báo cáo <b>Chi tiết từng bảng trong NDS</b></li> </ul>	Võ Trung Tín	00:00 AM 1/11/2024
10	<b><i>Thiết kế schema:</i></b> <ul style="list-style-type: none"> <li>- Tạo scripts tạo database và tables (Stage + NDS)</li> </ul>	Lô Thủy Tiên	00:00 AM 9/11/2024
11	<b><i>Xây dựng SSIS package để load data từ staging vào NDS</i></b>	Lô Thủy Tiên (Thực hiện chính)	00:00 AM 16/11/2024
12	<b><i>Xây dựng SSIS package để load data từ staging vào NDS</i></b>	Võ Lâm Hải Quốc Nguyễn Vũ	00:00 AM 16/11/2024

		Tường An Võ Trung Tín (Tham gia google meet)	
13	Hoàn thành báo cáo tổng hợp giai đoạn 1 ETL source-> stage-> NDS	Võ Lâm Hải Quốc Nguyễn Vũ Tường An Võ Trung Tín Lô Thủy Tiên	00:00 AM 23/11/2024
14	<b><i>Video giải thích từng component trong ETL process</i></b>	Lô Thủy Tiên	00:00 AM 23/11/2024

Báo cáo từ Github:

Link: <https://github.com/Locung201/BI2425>

Demo:

Link: <https://youtu.be/hoX4RLBOON0>



## ĐÁNH GIÁ THÀNH VIÊN

Nhóm : 03

MSSV	Họ tên	Hoàn thành	Nhận xét
19127536	Võ Lâm Hải Quốc	100%	
21127211	Nguyễn Vũ Tường An	100%	
21127450	Võ Trung Tín	100%	
21127699	Lô Thủy Tiên	100%	

**TÀI LIỆU THAM KHẢO****Công cụ và phần mềm hỗ trợ:**

STT	Chức năng	Công cụ
[1]	Thiết kế SSIS	Microsoft Visual Studio 2022
[2]	Báo cáo	Google Docs
[3]	Quản lý, trao đổi	Facebook, Messenger
[4]	Họp định kỳ	Google Meet
[5]	Quay Video demo	OBS Open Broadcaster Software
[6]	Thuyết trình	Canva
[7]	Sản phẩm Demo	Youtube
[8]	AI support	ChatGPT, Copilot, Perplexity, Gemini
[9]	Quản lý dự án	SSMS 20
[10]	Thiết kế CSDL	SQL server

**Tài liệu tham khảo:**

- [0] Tài liệu môn học Hệ thống thông tin phục vụ trí tuệ kinh doanh - 21HTTT2
- [1] Cô Sol ( 3 thg 11, 2024). [BI] - *DATH - Giải thích các component trong ETL (Source - Stage - NDS)* [Video]. YouTube. <https://www.youtube.com/watch?v=lqe5-Fn2rbc>.
- [2] Microsoft. (2024). *SQL Server Integration Services*.  
<https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver16>
- [3] Đông Chí . (10 thg 12, 2019). *PHẦN 1: Đổ dữ liệu từ stage sang NDS (Chicago)*. YouTube. <https://www.youtube.com/watch?v=TXqDN1aIdjo&t=376s>
- [4] Đông Chí . (24 thg 12, 2019). *Phần 2: Đổ dữ liệu từ NDS sang DDS* YouTube. <https://www.youtube.com/watch?v=128Oldh5uMo>.
- [5] Chuc Nguyen Van. (3 thg 7, 2021). *ETL Project From Excel Data Source to Star Schema with SSIS*. YouTube. <https://www.youtube.com/watch?v=Yp8fXLnVCp8&t=597s>
- [6] Huy Bui. (2022, 30 tháng 9). *SQL Server Integration Services – SSIS*. Cole.edu.vn. <https://cole.edu.vn/sql-server-integration-services-ssis/>
- [6] Microsoft. (n.d.). *Microsoft Learn*. Link: <https://learn.microsoft.com/en-us/> .