

BỘ Y TẾ
TRƯỜNG ĐẠI HỌC Y TẾ CÔNG CỘNG
KHOA: KHOA HỌC DỮ LIỆU



TÊN ĐỀ TÀI: DỮ LIỆU THỰC HÀNH

Môn học : Thực tập 1 (TH CSDL)
Mã môn học : 700902002
Giảng viên : Mai Văn Thủy
Nhóm : 4
Các thành viên : Nguyễn Hải An – 2211090001
Đinh Diệu Linh – 2211090022
Dương Mai Thu Hiền – 2211090012
Nguyễn Thị Thanh Nga - 2211090027
Đinh Lê Quỳnh Phương – 2211090031

MỤC LỤC

MỤC LỤC.....	2
DANH SÁCH HÌNH ẢNH.....	4
DANH SÁCH BẢNG BIỂU.....	4
LỜI MỞ ĐẦU.....	4
PHÂN CHIA NHIỆM VỤ.....	5
CHƯƠNG I. ĐẶT VẤN ĐỀ.....	7
CHƯƠNG II. THÔNG TIN BÁO CÁO.....	7
I. Quy trình chuyển đổi dữ liệu:.....	7
II. Thách thức và lợi ích:.....	7
CHƯƠNG III. NỘI DUNG BÁO CÁO.....	8
I. Làm sạch dữ liệu:.....	8
1. Định nghĩa, giải thích:.....	8
a) Làm Sạch Dữ Liệu: Quá Trình Cần Thiết Cho Phân Tích Dữ Liệu Chính Xác.....	8
b) Tại Sao Làm Sạch Dữ Liệu Lại Quan Trọng?.....	8
c) Các Bước Cơ Bản Trong Quá Trình Làm Sạch Dữ Liệu.....	8
2. Thực hiện:.....	9
II. Thống kê cơ bản:.....	11
1. Định nghĩa:.....	11
2. Thực hiện:.....	11
III. Nhập dữ liệu từ File Excel:.....	14
1. Sao dữ liệu từ SQL Server Import sang Import Wizard.....	14
2. Sao dữ liệu từ Visual Studio:.....	16
IV. Thiết kế cơ sở dữ liệu.....	18
A. Query tạo liên kết giữa các bảng:.....	18
1. Tạo cơ sở dữ liệu Thuctap1.....	18
2. Chỉnh sửa bảng VUNG.....	18
3. Thêm khoá chính vào bảng VUNG.....	18
4. Chỉnh sửa bảng TINH.....	18
5. Thêm khoá chính vào bảng TINH.....	18
6. Thêm khoá phụ vào bảng TINH.....	19
7. Chỉnh sửa bảng HUYEN.....	19
8. Thêm khoá chính vào bảng HUYEN.....	19
9. Thêm khoá phụ vào bảng HUYEN.....	19

10.	Chỉnh sửa bảng XA	19
11.	Thêm khoá chính vào bảng XA.....	19
12.	Thêm khoá phụ vào bảng XA.....	20
13.	Chỉnh sửa bảng DIABAN	20
14.	Thêm khoá chính vào bảng DIABAN	20
15.	Thêm khoá phụ vào bảng DIABAN.....	20
16.	Chỉnh sửa bảng THONGTINHO	20
17.	Thêm khoá chính vào bảng THONGTINHO	21
18.	Thêm khoá phụ vào bảng THONGTINHO	22
19.	Chỉnh sửa bảng THANHVIENTRONGHO.....	22
20.	Thêm khoá phụ vào bảng THANHVIENTRONGHO	22
21.	Chỉnh sửa bảng THONGTINTUVONG	22
23.	Thêm khoá phụ vào bảng THONGTINTUVONG.....	23
B.	Phân tích các bảng:.....	24
1.	Các khóa phụ, khóa chính trong bảng.....	24
2.	Sơ đồ liên kết giữa các bảng	24
3.	Mối quan hệ giữa các bảng	24
4.	Các kiểu dữ liệu được sử dụng:	25
5.	Sử dụng các cấu trúc điều khiển	26
5.1.	Sử dụng cấu trúc điều khiển if ... else	26
5.2.	Thủ tục và con trỏ	27
5.3.	Giao tác	28
5.4.	Hàm Function.....	28
5.5.	Trigger.....	28
5.6.	Tạo view.....	29
	CHƯƠNG IV: KẾT LUẬN	30
	LỜI CẢM ƠN	31

DANH SÁCH HÌNH ẢNH

Hình 1: Trước khi lọc dữ liệu bị lỗi phong chữ.....	9
Hình 2: Sau khi lọc dữ liệu bị lỗi phong chữ	9
Hình 3: Dữ liệu bị thiếu.....	9
Hình 4: Dữ liệu đã được cập nhật thêm.....	9
Hình 5: Bảng VUNG.....	10
Hình 6: Sử dụng hàm Vlookup để xác định rõ được dữ liệu.....	10
Hình 7: Trước khi thêm các mã liên kết giữa các bảng.....	10
Hình 8: Sau khi thêm các mã liên kết giữa các bảng	11
Hình 9: Các bước để nhập dữ liệu thống kê trên Power BI.....	11
Hình 10: Các dữ liệu sau khi được nhập lên 1	12
Hình 11: Các dữ liệu sau khi được nhập lên 2	12
Hình 12: Thống kê dữ liệu giữa năm tử vong và mã tử vong	12
Hình 13: Thống kê dữ liệu theo giới tính	13
Hình 14: Thống kê dữ liệu tử vong theo độ tuổi	13
Hình 15: Thống kê dữ liệu tử vong theo nguyên nhân.....	13
Hình 16: Các bước dùng SQL Server Import and Export 1	14
Hình 17: Các bước dùng SQL Server Import and Export 2	14
Hình 18: Các bước dùng SQL Server Import and Export 3	15
Hình 19: Các bước dùng SQL Server Import and Export 4	15
Hình 20: Các bước dùng SQL Server Import and Export 5	15
Hình 21: Sơ đồ liên kết giữa các bảng	24

DANH SÁCH BẢNG BIỂU

Bảng 1: Khóa chính và khóa phụ trong các bảng.....	24
Bảng 2: Các kiểu dữ liệu được dùng trong các bảng.....	26

LỜI MỞ ĐẦU

Trong bối cảnh phát triển mạnh mẽ của công nghệ 4.0, việc quản lý và xử lý dữ liệu đã trở thành một thách thức đối với các cơ quan nhà nước và doanh nghiệp. Để giải quyết vấn đề này, cần có những người am hiểu sâu rộng về công nghệ thông tin. Excel thường được sử dụng như một công cụ “quốc dân” để tổ chức dữ liệu, nhưng khi dữ liệu trở nên quy mô lớn và phức tạp, việc áp dụng các hệ thống quản lý cơ sở dữ liệu như SQL, C#, Python, ... trở nên cần thiết và hiệu quả hơn.

Báo cáo này sẽ trình bày quy trình chuyển đổi dữ liệu từ Excel sang SQL và C#, từ việc sử dụng dữ liệu có sẵn trong Excel, thiết kế cơ sở dữ liệu phù hợp, đến việc sử dụng các cấu trúc điều khiển và nhập dữ liệu vào SQL và C#. Sử dụng SQL không chỉ giúp tạo ra các truy vấn phức tạp, báo cáo tự động mà còn nâng cao hiệu quả phân tích dữ liệu.

Qua báo cáo này, chúng ta sẽ có cái nhìn tổng quan về quy trình chuyển đổi dữ liệu từ Excel sang SQL và C#, đồng thời nhận thức được giá trị mà việc sử dụng SQL mang lại trong việc quản lý và phân tích dữ liệu.

PHÂN CHIA NHIỆM VỤ

Thành viên	Mã sinh viên	Nhiệm vụ	Đóng góp
Nguyễn Hải An	2211090001	<ul style="list-style-type: none"> - Nhập dữ liệu từ SQL sang C# - Giải thích cách sao dữ liệu từ Visual Studio - Viết phần chương I: đặt vấn đề, chương II: thông tin báo cáo, chương IV: kết luận 	100%
Dương Mai Thu Hiền	2211090012	<ul style="list-style-type: none"> - Thiết kế khoá chính, khoá phụ, liên kết bảng bằng SQL - Viết phần lời mở đầu, lời cảm ơn, - Viết giải thích cách sao dữ liệu từ SQL Server Import sang Import Wizard 	100%
Đinh Diệu Linh	2211090022	<ul style="list-style-type: none"> - Viết các cấu trúc điều khiển if else , thủ tục, con trỏ bằng SQL - Viết mục thiết kế cơ sở dữ liệu phần Query tạo liên kết giữa các bảng(từ 1-11) - Viết cách sử dụng các cấu trúc điều khiển(từ 5.1-5.3) 	100%
Nguyễn Thị Thanh Nga	2211090027	<ul style="list-style-type: none"> - Viết các cấu trúc hàm, giao tác, trigger, sử dụng view thống kê dữ liệu từ vòng bằng SQL - Viết mục thiết kế cơ sở dữ liệu phần Query tạo liên kết giữa các bảng(từ 11-23) - Viết cách sử dụng các cấu trúc điều khiển(từ 5.4-5.6) 	100%
Đinh Lê Quỳnh Phương	2211090031	<ul style="list-style-type: none"> - Làm sạch dữ liệu trên Excel - Thiết kế cơ sở dữ liệu trên SQL - Thống kê cơ bản bằng Power BI - Viết chương III: Nội dung báo cáo phần làm sạch dữ liệu, thống kê cơ bản - Viết phần phân tích các bảng (các khóa chính, khóa phụ trong bảng, sơ đồ liên kết giữa các bảng, mối quan hệ giữa các bảng, các kiểu dữ liệu được sử dụng) - Viết phần mục lục, danh sách hình ảnh, danh sách bảng biểu 	100%

CHƯƠNG I. ĐẶT VẤN ĐỀ

Đi cùng với sự phát triển của công nghệ số hiện nay thì dữ liệu đóng vai trò quan trọng trong việc ra quyết định chiến lược và quản lý hoạt động. Một trong những thách thức lớn mà các tổ chức phải đối mặt là làm thế nào để việc quản lý và sử dụng dữ liệu trở nên dễ dàng, hiệu quả mà lại chính xác và nhanh chóng.

Mọi người thường sử dụng Microsoft Excel để tổ chức và xử lý các dữ liệu, nhưng nếu với quy mô lớn và phức tạp như bảng THONGTINH, THANHVIENTRONGHO, THONGTINTUVONG trong File excel có sẵn thì việc sử dụng File excel sẽ bị hạn chế và trở nên khó khăn cho người quản lý dữ liệu. Vì vậy, việc sử dụng hệ quản trị cơ sở dữ liệu như SQL được coi là “trợ lý đắc lực” cho việc quản lý và truy xuất dữ liệu

Qua bài báo cáo lần này, chúng ta sẽ biết về quá trình chuyển đổi dữ liệu từ File excel có sẵn sang cơ sở dữ liệu SQL. Chúng ta sẽ nhìn nhận rõ về các thách thức và lợi ích khi sử dụng hệ quản trị cơ sở dữ liệu SQL thông qua việc chuyển đổi dữ liệu ban đầu tới việc nhập dữ liệu vào SQL

CHƯƠNG II. THÔNG TIN BÁO CÁO

I. Quy trình chuyển đổi dữ liệu:

- **Chuẩn bị dữ liệu ban đầu trong File Excel:** Đầu tiên, dữ liệu cần được thu thập và tổ chức trong một File Excel. Điều này bao gồm việc nhập dữ liệu thô từ các nguồn khác nhau và sắp xếp chúng theo cách mà dễ dàng cho việc phân tích và chuyển đổi sau này.
- **Làm sạch dữ liệu trong File Excel:** Làm sạch dữ liệu là một bước quan trọng để đảm bảo rằng dữ liệu được chuyển đổi là chính xác và có thể sử dụng được. Quá trình này bao gồm việc loại bỏ các giá trị trùng lặp, sửa chữa các lỗi nhập liệu, và điền vào các giá trị thiếu.
- **Tiến hành quá trình nhập dữ liệu từ File Excel vào SQL:** Sau khi dữ liệu đã được làm sạch, bước tiếp theo là nhập chúng vào một cơ sở dữ liệu SQL. Điều này thường được thực hiện thông qua các công cụ nhập dữ liệu hoặc bằng cách sử dụng các lệnh SQL để nhập dữ liệu từ Excel vào SQL Server.
- **Thiết kế cơ sở dữ liệu SQL phù hợp với cấu trúc dữ liệu File Excel:** Cơ sở dữ liệu SQL cần được thiết kế để phản ánh cấu trúc dữ liệu của File Excel. Điều này bao gồm việc tạo ra các bảng và thiết lập các mối quan hệ giữa chúng, cũng như định nghĩa các khóa ngoại và chỉ mục để tối ưu hóa hiệu suất truy vấn.
- **Sử dụng các cấu trúc điều khiển:** Các cấu trúc điều khiển như vòng lặp và điều kiện có thể được sử dụng để tự động hóa quá trình nhập dữ liệu và xử lý các tình huống phức tạp trong quá trình chuyển đổi dữ liệu

II. Thách thức và lợi ích:

- **Thách thức:** Một số thách thức phổ biến trong quá trình chuyển đổi dữ liệu bao gồm việc đối mặt với dữ liệu không nhất quán, sự khác biệt về định dạng dữ liệu giữa Excel và SQL, và khó khăn trong việc tự động hóa quá trình chuyển đổi dữ liệu. Ngoài ra, việc sửa đổi tên cột và kiểu dữ liệu để phù hợp với cơ sở dữ liệu SQL cũng có thể là một thách thức.

- **Lợi ích:** Việc sử dụng SQL mang lại nhiều lợi ích như khả năng quản lý và phân tích dữ liệu một cách hiệu quả, truy vấn dữ liệu nhanh chóng và chính xác, và khả năng mở rộng cơ sở dữ liệu để xử lý lượng dữ liệu lớn. SQL cũng cung cấp các công cụ mạnh mẽ để tối ưu hóa và bảo mật dữ liệu³⁴.

CHƯƠNG III. NỘI DUNG BÁO CÁO

I. Làm sạch dữ liệu:

1. Định nghĩa, giải thích:

a) Làm Sạch Dữ Liệu: Quá Trình Cần Thiết Cho Phân Tích Dữ Liệu Chính Xác

Trong thế giới ngày nay, dữ liệu được coi là tài sản quý giá của các tổ chức và doanh nghiệp. Tuy nhiên, để có thể khai thác hiệu quả giá trị từ dữ liệu, việc đảm bảo chất lượng dữ liệu là điều cực kỳ quan trọng. Làm sạch dữ liệu, hay còn gọi là data cleaning, là một phần không thể thiếu trong quá trình chuẩn bị dữ liệu cho phân tích. Đây là quá trình loại bỏ hoặc sửa chữa các dữ liệu không chính xác, bị hỏng, định dạng sai, trùng lặp, hoặc không đầy đủ, nhằm tạo ra một tập dữ liệu chính xác và đáng tin cậy.

b) Tại Sao Làm Sạch Dữ Liệu Lại Quan Trọng?

Làm sạch dữ liệu quan trọng vì nó ảnh hưởng trực tiếp đến khả năng phân tích và đưa ra quyết định dựa trên dữ liệu. Dữ liệu không chính xác có thể dẫn đến các kết luận sai lệch, làm ảnh hưởng đến chiến lược và hoạt động kinh doanh. Một quyết định dựa trên dữ liệu không chính xác có thể gây ra hậu quả nghiêm trọng, từ việc lãng phí nguồn lực đến việc mất cơ hội kinh doanh.

c) Các Bước Cơ Bản Trong Quá Trình Làm Sạch Dữ Liệu

- **Xác định Dữ Liệu Cần Làm Sạch:** Trước tiên, cần phải xác định những phần dữ liệu nào cần được làm sạch. Điều này có thể được thực hiện thông qua việc phân tích dữ liệu và sử dụng các công cụ thống kê để tìm ra dữ liệu bất thường hoặc không đúng định dạng.
- **Loại Bỏ Dữ Liệu Trùng Lặp:** Dữ liệu trùng lặp có thể xuất hiện do nhiều nguyên nhân, từ lỗi nhập liệu đến việc hợp nhất các nguồn dữ liệu khác nhau. Việc loại bỏ dữ liệu trùng lặp giúp giảm thiểu sự méo mó trong phân tích.
- **Chỉnh Sửa Lỗi Định Dạng và Cấu Trúc:** Lỗi định dạng có thể bao gồm việc sử dụng sai định dạng ngày tháng, số tiền, hoặc đơn vị đo lường. Lỗi cấu trúc có thể liên quan đến việc sử dụng không nhất quán các tiêu chuẩn đặt tên hoặc phân loại.
- **Xử Lý Dữ Liệu Thiếu:** Dữ liệu thiếu có thể được xử lý bằng cách sử dụng các phương pháp thay thế như điền giá trị trung bình hoặc trung vị, hoặc loại bỏ các hàng hoặc cột không chứa thông tin đầy đủ.
- **Chuẩn Hóa Dữ Liệu:** Chuẩn hóa dữ liệu bao gồm việc đảm bảo rằng tất cả dữ liệu tuân theo một định dạng và tiêu chuẩn nhất định, giúp cho việc phân tích trở nên dễ dàng và chính xác hơn.

- **Xác Định và Xử Lý Các Giá Trị Ngoại Lệ:** Các giá trị ngoại lệ có thể làm méo mó kết quả phân tích và cần được xác định một cách cẩn thận để quyết định xem có nên giữ lại hay loại bỏ.

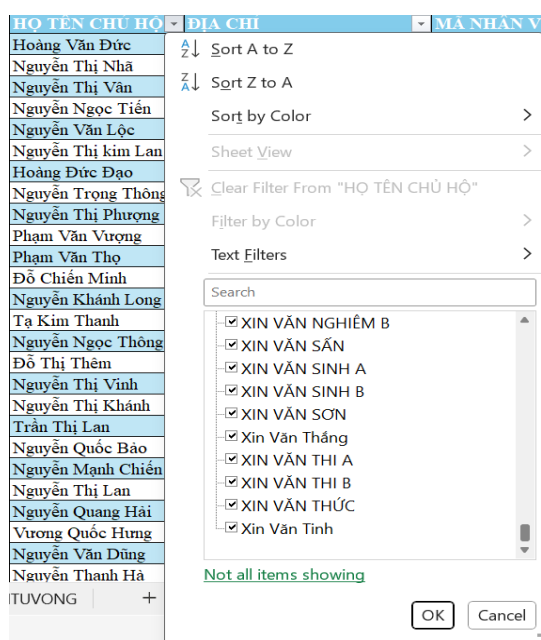
2. Thực hiện:

- Bước 1: Kiểm tra dữ liệu bị lặp

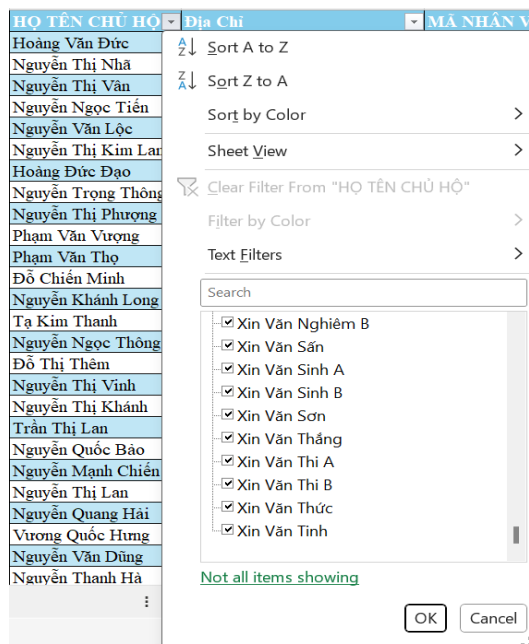
Chọn bất kì dữ liệu vào trong bảng, sau đó bấm Data và chọn Remove Duplicates, bấm Select All và chọn My data has headers rồi nhấn OK là sẽ hiển thị các bảng đã được lọc dữ liệu bị lặp

- Bước 2: Kiểm tra dữ liệu bị lỗi:

+ Kiểm tra dữ liệu bị lỗi phong chữ viết hoa và viết thường bằng cách sử dụng Ablebits Data, chọn Text và bấm Proper Case trong mục Change Case



Hình 1: Trước khi lọc dữ liệu bị lỗi phong chữ



Hình 2: Sau khi lọc dữ liệu bị lỗi phong chữ

+ Kiểm tra dữ liệu bị lỗi hiển thị thiếu ngày, tháng, năm, thời gian cụ thể: Chọn cột cụ thể rồi bấm Format Cells, chọn Custom và gõ dd/mm/yyyy hh:mm

NGÀY KẾT THÚC	NGÀY PHÒNG VÁN
2021-04-10	2021-04-06
2021-04-10	2021-04-06
2021-04-10	2021-04-06

Hình 3: Dữ liệu bị thiếu

NGÀY KẾT THÚC	NGÀY PHÒNG VÁN
10/04/2021 08:16:02	06/04/2021 16:37:41
10/04/2021 08:28:16	06/04/2021 16:47:38
10/04/2021 08:33:41	06/04/2021 16:59:08

Hình 4: Dữ liệu đã được cập nhật thêm

+ Kiểm tra dữ liệu tại sao bị lỗi (VD như có một số huyện bị nhập nhầm vào bảng XA thay vì vào bảng HUYEN nên xảy ra lỗi)

- Bước 3: Tạo ID (Tạo khóa chính của bảng): Tạo 1 cột Mavung đằng sau TÊN VÙNG và đánh số thứ tự từ V01 đến V05
- Bước 4: Tạo khóa phụ cho các bảng
 - * Có 2 lí do cần phải tạo khóa phụ trong bảng
 - + Kết nối các tỉnh: Mỗi tỉnh có một mã số. Nếu tên vùng thay đổi, mã số giữ nguyên nên các tỉnh vẫn liên kết được với nhau.
 - + Dữ liệu gọn gàng: Dùng mã số giúp bảng dữ liệu ngắn gọn, dễ quản lý.

TÊN VÙNG	MaVung
Đồng bằng sông Hồng	V01
Trung du và miền núi phía Bắc	V02
Bắc Trung Bộ và Duyên hải miền Trung	V03
Tây Nguyên	V04
Đồng Nam Bộ	V05
Đồng bằng sông Cửu Long	V06

Hình 5: Bảng VUNG

TÊN TỈNH	VÙNG	MaVung
Tỉnh Bình Phước	Đồng Nam Bộ	V05

Hình 6: Sử dụng hàm Vlookup để xác định rõ được dữ liệu

- * Cách điền khóa phụ vào các bảng:
 - + Tìm mã vùng: Sử dụng công thức VLOOKUP để tìm mã vùng cho từng tỉnh.
 - + Chuyển đổi dữ liệu: Khi đã tìm được mã, chuyển đổi chúng từ công thức sang dữ liệu cố định.
 - + Làm gọn bảng: Xóa cột vùng không cần thiết để bảng dễ quản lý hơn.

MA HỌ	TÊN TỈNH	TÊN HUYỆN	TÊN XÃ	TÊN VÙNG	TÊN ĐỊA BÀN	HỌ SO	NAM	HỌ TÊN CHỦ HỘ	ĐỊA CHỈ	MÃ NHÂN VIÊN
100100001012003	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	3	2021	Hoàng Văn Đức	Số 3/49/71 Tân Ấp	DTV01003
100100001012006	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	6	2021	Nguyễn Thị Nhã	Số 9/49/71 Tân Ấp	DTV01003
100100001012011	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	11	2021	Nguyễn Thị Vân	Số 14/49/71 Tân Ấp	DTV01003
100100001012014	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	14	2021	Nguyễn Ngọc Tiến	Số 8/49/71 Tân Ấp	DTV01003
100100001012018	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	18	2021	Nguyễn Văn Lộc	Số 2/49/71 Tân Ấp	DTV01003
100100001012021	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	21	2021	Nguyễn Thị kim Lan	Số 4/51/71 Tân Ấp	DTV01003
100100001012024	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	24	2021	Hoàng Đức Đạo	Số 53/71 Tân Ấp	DTV01003
100100001012028	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	28	2021	Nguyễn Trọng Thông	Số 6/55/71 Tân Ấp	DTV01003
100100001012031	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	31	2021	Nguyễn Thị Phương	Số 6/59/71 Tân Ấp	DTV01003
100100001012034	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	34	2021	Phạm Văn Vương	Số 59/71 Tân Ấp	DTV01003
100100001012037	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	37	2021	Phạm Văn Thọ	7/61/71 Tân Ấp	DTV01003
100100001012040	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	40	2021	Đỗ Chiến Minh	6/63/71 Tân Ấp	DTV01003
100100001012043	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	43	2021	Nguyễn Khánh Long	65/71 Tân Ấp	DTV01003
100100001012050	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	50	2021	Tạ Kim Thanh	75/71 Tân Ấp	DTV01003
100100001012053	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	53	2021	Nguyễn Ngọc Thông	77/71 Tân Ấp	DTV01003
100100001012056	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	56	2021	Đỗ Thị Thêm	79/71 Tân Ấp	DTV01003
100100001012059	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	59	2021	Nguyễn Thị Vinh	81/71 Tân Ấp	DTV01003
100100001012063	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	63	2021	Nguyễn Thị Khánh	21/12 Tân Ấp	DTV01003
100100001012066	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	66	2021	Trần Thị Lan	25/12 Tân Ấp	DTV01003
100100001012069	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	69	2021	Nguyễn Quốc Bảo	27/12 Tân Ấp	DTV01003
100100001012073	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	73	2021	Nguyễn Mạnh Chiến	5/12/27 Nghĩa Dũng	DTV01003
100100001012076	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	76	2021	Nguyễn Thị Lan	9/12/27 Nghĩa Dũng	DTV01003
100100001012079	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	79	2021	Nguyễn Quang Hải	11/12/27 Nghĩa Dũng	DTV01003
100100001012082	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	82	2021	Vương Quốc Hưng	21/12/27 Nghĩa Dũng	DTV01003
100100001012085	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	85	2021	Nguyễn Văn Dũng	25/12/27 Nghĩa Dũng	DTV01003
100100001012088	Thị xã Hà Nội	Quận Ba Đình	Phường Phúc Xá	Đồng bằng sông Hồng	Xóm đầu làng 1	88	2021	Nguyễn Thanh Hà	31/12/27 Nghĩa Dũng	DTV01003

Hình 7: Trước khi thêm các mã liên kết giữa các bảng

MA HO	Ma Tinh	Ma Huyen	Ma Xa	TÊN VÙNG	Ma Yang	TÊN ĐIỂM	Ma Di Ban	HỌ SO	NAM	HỌ TÊN CHỦ	HỌ TÊN
100100001012003	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	3	2021	Hoàng Văn Đức	S6 3/49/71 T
100100001012006	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	6	2021	Nguyễn Thị Nhà	S6 9/49/71 T
100100001012011	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	11	2021	Nguyễn Thị Văn	S6 14/49/71 T
100100001012014	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	14	2021	Nguyễn Ngọc Tiến	S6 8/49/71 T
100100001012018	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	18	2021	Nguyễn Văn Lộc	S6 2/49/71 T
100100001012021	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	21	2021	Nguyễn Thị Kim Lan	S6 4/51/71 T
100100001012024	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	24	2021	Hoàng Đức Đạo	S6 53/71 Tân
100100001012028	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	28	2021	Nguyễn Trọng Thông	S6 55/71 Tân
100100001012031	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	31	2021	Nguyễn Thị Phương	S6 6/55/71 T
100100001012034	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	34	2021	Phạm Văn Vương	S6 59/71 Tân
100100001012037	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	37	2021	Phạm Văn Thọ	7/61/71 Tân
100100001012040	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	40	2021	Đỗ Chiến Minh	6/63/71 Tân
100100001012043	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	43	2021	Nguyễn Khánh Long	65/71 Tân
100100001012050	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	50	2021	Ta Kim Thanh	75/71 Tân
100100001012053	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	53	2021	Nguyễn Ngọc Thông	77/71 Tân
100100001012056	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	56	2021	Đỗ Thị Thêm	79/71 Tân
100100001012059	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	59	2021	Nguyễn Thị Vinh	81/71 Tân
100100001012063	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	63	2021	Nguyễn Thị Khánh	21/12 Tân
100100001012066	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	66	2021	Trần Thị Lan	25/12 Tân
100100001012069	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	69	2021	Nguyễn Quốc Bảo	27/12 Tân
100100001012073	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	73	2021	Nguyễn Mạnh Chiến	5/12/27 Ngh
100100001012076	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	76	2021	Nguyễn Thị Lan	9/12/27 Ngh
100100001012079	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	79	2021	Nguyễn Quang Hải	11/12/27 Ngh
100100001012082	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	82	2021	Vương Quốc Hùng	21/12/27 Ngh
100100001012085	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	85	2021	Nguyễn Văn Dũng	25/12/27 Ngh
100100001012088	T21	H001	X001	Dòng bồng sông Hồng	V01	Xóm Dân Láng .1	D066	88	2021	Nguyễn Văn Hà	31/12/27 Ngh

Hình 8: Sau khi thêm các mã liên kết giữa các bảng

II. Thống kê cơ bản:

1. Định nghĩa:

Thống kê cơ bản trong dữ liệu là một lĩnh vực nghiên cứu về cách thu thập, phân tích, giải thích và trình bày dữ liệu. Nó bao gồm hai phần chính:

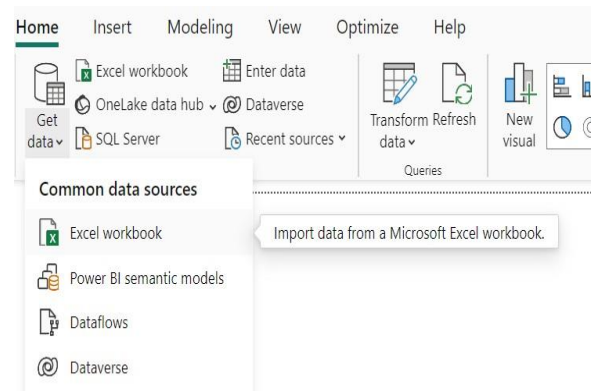
- Thống kê mô tả (Descriptive Statistics): Mô tả các đặc tính của dữ liệu, như trung bình, trung vị, mốt, phương sai, và độ lệch chuẩn. Phần này giúp ta hiểu được dữ liệu hiện tại mà không cần đến suy luận hay dự đoán
- Thống kê suy diễn (Inferential Statistics): Sử dụng dữ liệu từ một mẫu nhỏ để rút ra kết luận về một quần thể lớn hơn. Phần này thường liên quan đến việc ước lượng, kiểm định giả thuyết và dự đoán

2. Thực hiện:

Thống kê tử vong bằng Power BI

Bước 1: Chọn Get Data:

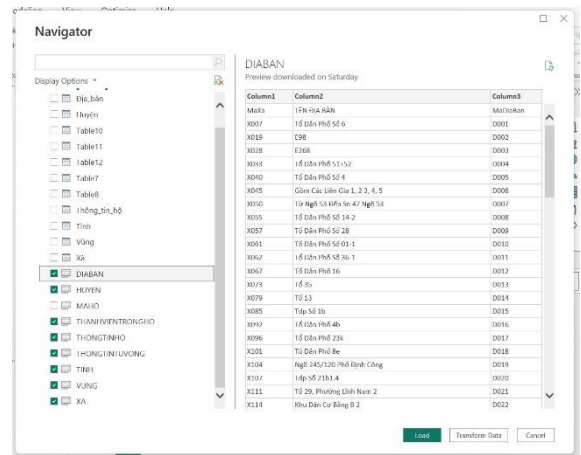
Bấm Excel Workbook để Import data from a Microsoft Excel workbook



Hình 9: Các bước để nhập dữ liệu thống kê trên Power BI

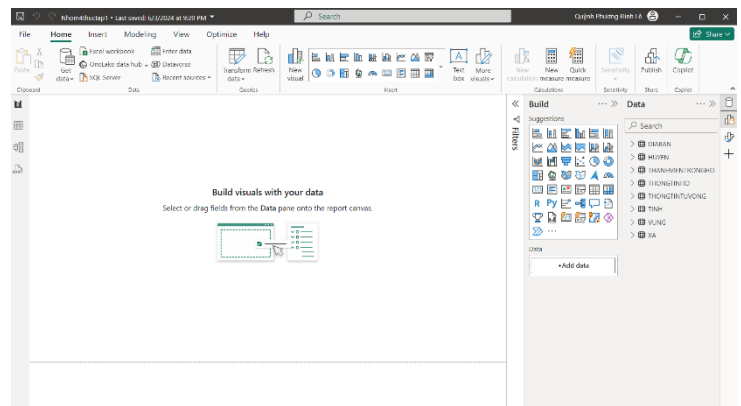
Bước 2:

- Chọn File Excel mình cần nhập
- Chọn các bảng để chuyển sang dữ liệu:
DIABAN, THANHVIENTRONGHO, THONGTINHO,
THONGTINTUVONG, TINH, VUNG, XA



Hình 10: Các dữ liệu sau khi được nhập lên 1

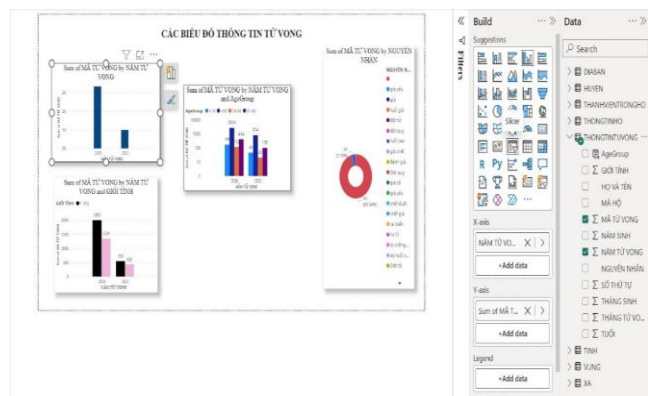
Bước 3: Sau khi chọn các bảng để chuyển sang dữ liệu trên thì mình sẽ chọn Close & Apply để xuất hiện các bảng dữ liệu trong mục Data



Hình 11: Các dữ liệu sau khi được nhập lên 2

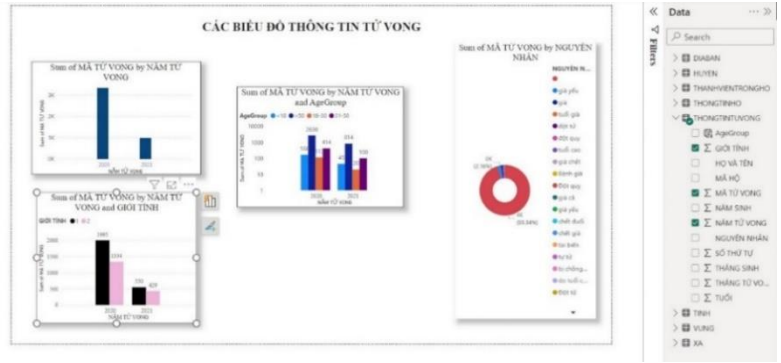
Bước 4: Thực hiện thống kê trong bảng THONGTINTUVONG

Sử dụng NĂM TỬ VONG làm trục x và SUM of MÃ TỬ VONG làm trục y để thống kê dữ liệu



Hình 12: Thống kê dữ liệu giữa năm tử vong và mã tử vong

Sử dụng NĂM TỬ VONG làm trục x, SUM of MÃ TỬ VONG làm trục y và GIỚI TÍNH làm Legend để thống kê dữ liệu tử vong theo giới tính



Hình 13: Thống kê dữ liệu tử vong theo giới tính

Sử dụng AgeGroup để phân tích dữ liệu theo độ tuổi bằng cách dùng Modelling- chọn New Column và gõ code như sau

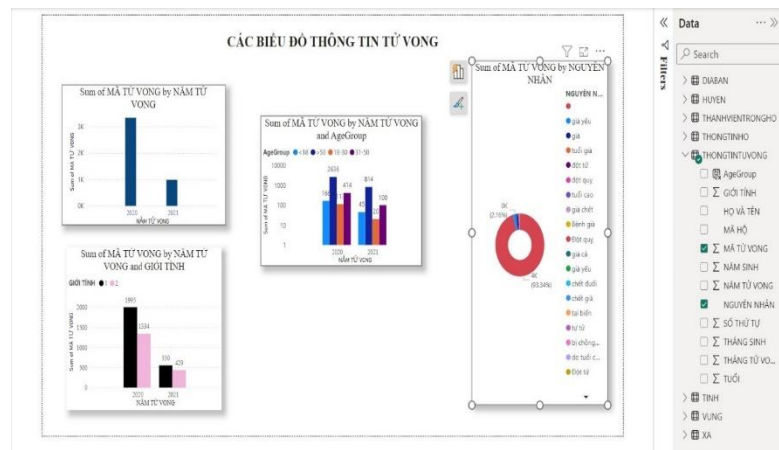
```
AgeGroup = IF (THONGTINTUVONG[TUỔI] < 18, "<18",
    IF (THONGTINTUVONG[TUỔI] <= 30, "18-30",
    IF (THONGTINTUVONG[TUỔI] <= 50, "31-50", ">50"))
```

Sau đó, chọn NĂM TỬ VONG làm trục x, SUM of MÃ TỬ VONG làm trục y và AgeGroup làm Legend



Hình 14: Thống kê dữ liệu tử vong theo độ tuổi

Sử dụng NGUYÊN NHÂN làm Legend và SUM of MÃ TỬ VONG làm Values để thống kê các nguyên nhân gây tử vong

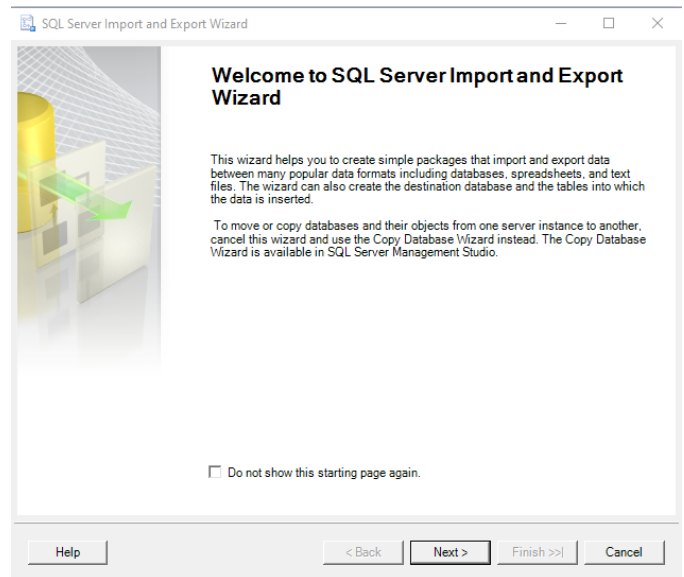


Hình 15: Thống kê dữ liệu tử vong theo nguyên nhân

III. Nhập dữ liệu từ File Excel:

1. Sao dữ liệu từ SQL Server Import sang Import Wizard

Bước 1: Mở chương trình Import, chọn database mình cần nhập bằng cách di chuyển chuột vào nút Tast, rồi vào Import Data, chương trình SQL Server Import and Export Wizard sẽ xuất hiện như hình ảnh trên. Sau đó bấm Next >



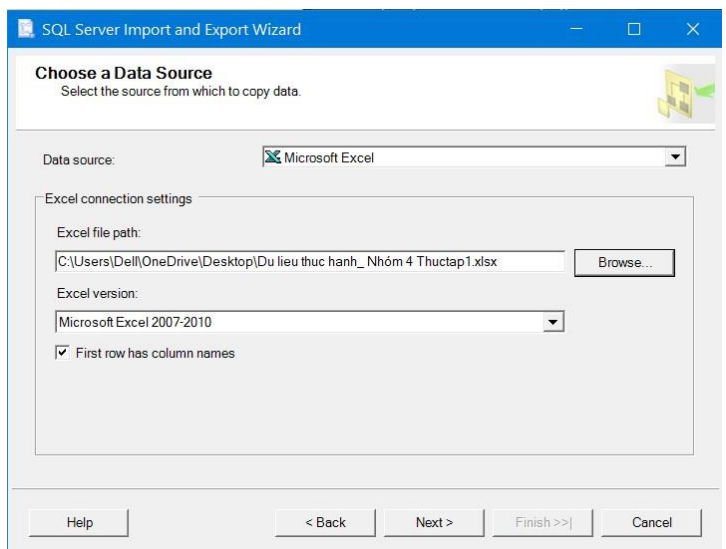
Hình 16: Các bước dùng SQL Server Import and Export 1

Bước 2: Chọn nguồn

- Trong mục Data Source, chọn Microsoft Excel
- Trong mục Excel file path, chọn Browse để chọn file mình cần nhập lên dữ liệu

**Nếu gặp lỗi: The 'Microsoft.ACE.OLEDB.12.0' provider is not registered on the local machine. (System.Data)*

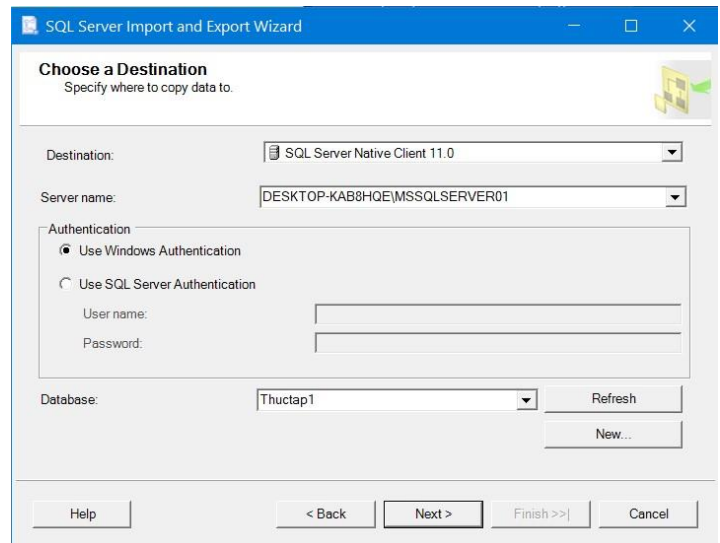
Tải và chạy: AccessRuntime_x86_en-us



Hình 17: Các bước dùng SQL Server Import and Export 2

Bước 3: Chọn địa điểm

- Trong mục Destination, chọn SQL Server Native Client 11.0
- Bấm Next > để chuyển sang trang tiếp theo
- Bấm Next > một lần nữa để chuyển sang trang tiếp theo



Hình 18: Các bước dùng SQL Server Import and Export 3

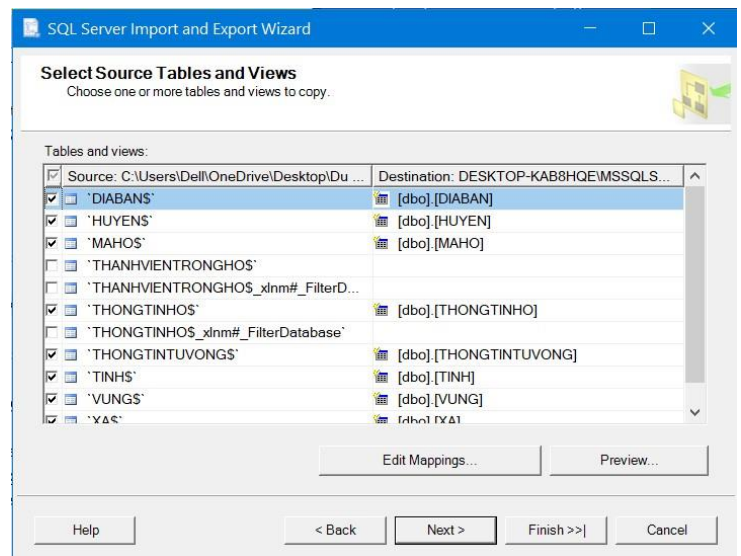
Bước 4: Chọn các bảng cần thiết

Tick các bảng cần chuyển qua, có thể đổi tên bảng

Bấm Next > để chuyển sang trang tiếp theo.

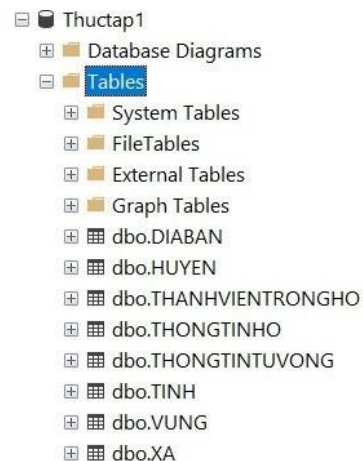
Bấm Next > một lần nữa để chuyển sang trang tiếp theo.

Bấm Finish > để chuyển chương trình chạy.



Hình 19: Các bước dùng SQL Server Import and Export 4

Bước 5: Kiểm tra xem database ThucTap1 đã có đầy đủ các bảng trong File excel hay chưa



Hình 20: Các bước dùng SQL Server Import and Export 5

2. Sao dữ liệu từ Visual Studio:

Khởi tạo thông tin đăng nhập:

- + Khởi động SQL Server Management Studio (SSMS).
- + Tạo kết nối đến instance của SQL Server, sử dụng server: NGUYEN-HAI-ANN\MSSQLSERVER01
- + Trong cửa sổ Object Explorer, điều hướng đến và mở rộng node của server.
- + Click chuột phải vào mục “Security” và chọn “New Login”.
- + Điền tên đăng nhập mong muốn, user: NGHAIAN4, và chọn phương thức xác thực, như SQL Server Authentication.
- + Thiết lập mật khẩu cho tài khoản: Password=HUPH2023
- + Cấu hình thêm các tùy chọn khác theo nhu cầu, như database mặc định hay server roles.
- + Nhấn “OK” để hoàn tất việc tạo login.

Tạo user trong database:

- + Truy cập vào database mà bạn muốn thêm user
- + Click chuột phải vào “Users” và chọn “New User”.
- + Nhập tên user: NGHAIAN4
- + Liên kết với login đã tạo ở bước trước, user: NGHAIAN4
- + Cài đặt các thuộc tính khác như schema mặc định hay permissions.
- + Nhấn “OK” để tạo user mới.

Mở Visual Studio, chọn Console App Template, vào trong Program.cs, gõ:

```
using System;
using Microsoft.Data.SqlClient;

namespace ImportData
{
    class Program
    {
        static void Main(string[] args)
        {
            string connectionString = " Server=NGUYEN-HAI-ANN\MSSQLSERVER01;Database=Thuctap1;UserId=NGHAIAN4;Password=HUPH2023;

            using SqlConnection connection = new(connectionString);
            try
            {
                connection.Open();
                Console.WriteLine("Kết nối thành công!");
                string query = "SELECT TOP 1 * FROM [dbo].[VUNG]";
```



```

using SqlCommand command = new(query, connection);
using SqlDataReader reader = command.ExecuteReader();
while (reader.Read())
{
    Console.WriteLine(reader.GetString(0));
}
}
catch (Exception ex)
{
    Console.WriteLine("Lỗi: " + ex.Message);
}
finally
{
    connection.Close();
    Console.WriteLine("Kết nối đã đóng.");
}
}
}

```

Sau đó bấm chạy để ứng dụng thực hiện việc sao chép

IV. Thiết kế cơ sở dữ liệu

A. Query tạo liên kết giữa các bảng:

1. Tạo cơ sở dữ liệu Thuctap1

```
CREATE DATABASE Thuctap1  
GO
```

```
USE Thuctap1  
GO
```

2. Chỉnh sửa bảng VUNG

```
---///TABLE [dbo].[VUNG]  
-- Alter columns to NOT NULL  
ALTER TABLE [dbo].[VUNG]  
ALTER COLUMN MaVung NVARCHAR(255) NOT NULL;
```

3. Thêm khoá chính vào bảng VUNG

```
-- Add PK Mavung  
ALTER TABLE [dbo].[VUNG]  
ADD PRIMARY KEY (MaVung);
```

4. Chỉnh sửa bảng TINH

```
---///TABLE [dbo].[TINH]  
-- Alter columns to NOT NULL  
-- Delete rows where MaTinh or MaVung is NULL or has placeholder values  
DELETE FROM [dbo].[TINH]  
WHERE MaTinh IS NULL OR MaTinh = 'Default_Value' OR MaVung IS NULL OR MaVung =  
'Default_Value';
```

```
-- Verify deletion  
SELECT * FROM [dbo].[TINH];  
ALTER TABLE [dbo].[TINH]  
ALTER COLUMN MaTinh NVARCHAR(255) NOT NULL;
```

```
ALTER TABLE [dbo].[TINH]  
ALTER COLUMN MaVung NVARCHAR(255) NOT NULL;
```

5. Thêm khoá chính vào bảng TINH

```
-- Add PK MaTinh  
ALTER TABLE [dbo].[TINH]  
ADD PRIMARY KEY ([MaTinh]);
```

6. Thêm khoá phụ vào bảng TINH

```
-- Add FK MaVung
ALTER TABLE [dbo].[TINH]
ADD CONSTRAINT FK_TINH_VUNG
FOREIGN KEY (MaVung) REFERENCES [dbo].[VUNG] (MaVung);
```

7. Chỉnh sửa bảng HUYEN

```
---///TABLE [dbo].[HUYEN]
-- Change Matinh column
ALTER TABLE [dbo].[HUYEN]
ALTER COLUMN MaTinh NVARCHAR(255) NOT NULL;

-- Change Mahuyen column
ALTER TABLE [dbo].[HUYEN]
ALTER COLUMN MaHuyen NVARCHAR(255) NOT NULL;
```

8. Thêm khoá chính vào bảng HUYEN

```
-- Add PK Mahuyen
ALTER TABLE [dbo].[HUYEN]
ADD PRIMARY KEY (MaHuyen);
```

9. Thêm khoá phụ vào bảng HUYEN

```
-- Add FK Matinh
ALTER TABLE [dbo].[HUYEN]
ADD CONSTRAINT FK_HUYEN_TINH
FOREIGN KEY (MaTinh) REFERENCES [dbo].[TINH](MaTinh);
```

10. Chỉnh sửa bảng XA

```
---///TABLE [dbo].[XA]
-- Alter columns to NOT NULL
ALTER TABLE [dbo].[XA]
ALTER COLUMN MaXa NVARCHAR(255) NOT NULL;

ALTER TABLE [dbo].[XA]
ALTER COLUMN MaHuyen NVARCHAR(255) NOT NULL;
```

11. Thêm khoá chính vào bảng XA

```
-- Add PK Maxa
ALTER TABLE [dbo].[XA]
```

ADD PRIMARY KEY (MaXa);

12. Thêm khoá phụ vào bảng XA

-- Add FK Mahuyen

ALTER TABLE [dbo].[XA]

WITH NOCHECK

ADD CONSTRAINT FK_XA_HUYEN

FOREIGN KEY (MaHuyen) REFERENCES [dbo].[HUYEN](MaHuyen);

13. Chỉnh sửa bảng DIABAN

---/// TABLE [dbo].[DIABAN]

-- Alter columns to NOT NULL

ALTER TABLE [dbo].[DIABAN]

ALTER COLUMN MaXa NVARCHAR(255) NOT NULL;

ALTER TABLE [dbo].[DIABAN]

ALTER COLUMN MaDiaBan NVARCHAR(255) NOT NULL;

14. Thêm khoá chính vào bảng DIABAN

-- Add PK Maxa

ALTER TABLE [dbo].[DIABAN]

ADD PRIMARY KEY (MaDiaBan);

15. Thêm khoá phụ vào bảng DIABAN

-- Add FK Mahuyen

ALTER TABLE [dbo].[DIABAN]

ADD CONSTRAINT FK_XA_DIABAN

FOREIGN KEY (MaXa) REFERENCES [dbo].[XA](MaXa);

16. Chỉnh sửa bảng THONGTINHO

---/// TABLE [dbo].[THONGTINHO]

-- Change MaHo column data type to BIGINT

ALTER TABLE [dbo].[THONGTINHO]

ALTER COLUMN MaHo BIGINT NOT NULL;

-- Change Maxa column data type to INT

ALTER TABLE [dbo].[THONGTINHO]

ALTER COLUMN MaXa NVARCHAR(255) NOT NULL;

-- Change [Hoso] column data type to INT

ALTER TABLE [dbo].[THONGTINHO]

ALTER COLUMN Hoso INT;

```

-- Change Nam column data type to SMALLINT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN Nam SMALLINT;

-- Change MaNhanVien column data type to VARCHAR(8)
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN MaNhanVien VARCHAR(8);

-- Change DienThoai column data type to VARCHAR
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN DienThoai VARCHAR(20);

-- Change [TSNK] column data type to INT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN TSNK INT;

-- Change [TSNAM] column data type to INT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN TSNAM INT;

-- Change [TSNU] column data type to INT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN [TSNU] INT;

-- Change [KT9] column data type to INT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN KT9 INT;

-- Change [C45] column data type to INT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN C45 INT;

-- Change [KT14] column data type to INT
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN KT14 INT;

-- Change NGƯỜI TẠO column data type to VARCHAR(8)
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN NguoiTao VARCHAR(8);

-- Change NgayTao column data type to DATETIME
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN NgayTao DATETIME;

-- Change Phienban column data type to VARCHAR
ALTER TABLE [dbo].[THONGTINHO]
ALTER COLUMN Phienban VARCHAR(10);

```

17. Thêm khoá chính vào bảng THONGTINHO

```
-- Set MaHo as the primary key
ALTER TABLE [dbo].[THONGTINHO]
ADD PRIMARY KEY (MaHo);
```

18. Thêm khoá phụ vào bảng THONGTINHO

```
-- Add foreign key constraint to XA table
ALTER TABLE [dbo].[THONGTINHO]
ADD CONSTRAINT FK_THONGTINHO_XA
FOREIGN KEY (MaXa) REFERENCES XA(MaXa);
```

19. Chỉnh sửa bảng THANHVIENTRONGHO

```
---/// TABLE [dbo].[THANHVIENTRONGHO]
-- Change MaHo column data type to BIGINT
ALTER TABLE [dbo].[THANHVIENTRONGHO]
ALTER COLUMN MaHo BIGINT NOT NULL;

-- Change MaThanhVien column data type to INT
ALTER TABLE [dbo].[THANHVIENTRONGHO]
ALTER COLUMN MaThanhVien INT;

-- Change SoThuTu column data type to INT
ALTER TABLE [dbo].[THANHVIENTRONGHO]
ALTER COLUMN SoThuTu INT;
```

20. Thêm khoá phụ vào bảng THANHVIENTRONGHO

```
-- Add foreign key constraint to THONGTINHO table
ALTER TABLE [dbo].[THANHVIENTRONGHO]
WITH NOCHECK
ADD CONSTRAINT FK_THANHVIENTRONGHO_THONGTINHO
FOREIGN KEY ([MaHo]) REFERENCES [dbo].[THONGTINHO]([MaHo]);
```

21. Chỉnh sửa bảng THONGTINTUVONG

```
---/// TABLE [dbo].[THONGTINTUVONG]
-- Change MaHo column data type to BIGINT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN MaHo BIGINT;

-- Change MaTuVong column data type to INT (not null)
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN MaTuVong INT;

-- Change SoThuTu column data type to INT
ALTER TABLE [dbo].[THONGTINTUVONG]
```

```

ALTER COLUMN SoThuTu INT;

-- Change Gioitinh column data type to BIT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN Gioitinh BIT;

-- Update existing data (1 for male, 2 for female)
UPDATE [dbo].[THONGTINTUVONG]
SET [Gioitinh] = CASE WHEN [Gioitinh] = 1 THEN 1 ELSE 0 END;

-- Change Thangtuvong column data type to INT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN Thangtuvong INT;

-- Change Namtuvong column data type to SMALLINT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN Namtuvong SMALLINT;

-- Change Thangsinh column data type to INT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN Thangsinh INT;

-- Change Namsinh column data type to SMALLINT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN Namsinh SMALLINT;

-- Change Tuoi column data type to INT
ALTER TABLE [dbo].[THONGTINTUVONG]
ALTER COLUMN Tuoi INT;

```

23. Thêm khoá phụ vào bảng THONGTINTUVONG

```

-- Add foreign key constraint to THONGTINHO table
ALTER TABLE [dbo].[THONGTINTUVONG]
WITH NOCHECK
ADD CONSTRAINT FK_MAHO_THONGTINTUVONG_THONGTINHO
FOREIGN KEY (MaHo) REFERENCES THONGTINHO(MaHo);

```


- + Một Tỉnh chứa nhiều Huyện.
- + Một Huyện gồm nhiều Xã.
- + Mỗi Xã có nhiều Mã hộ được ghi nhận trong bảng THONGTINHO.
- Còn giữa các bảng THONGTINHO, THANHVIENHO, và THONGTINTUVONG, chúng ta thấy mô hình quan hệ một-nhiều:
 - + Một Mã hộ từ bảng THONGTINHO có thể xuất hiện nhiều lần trong bảng THANHVIENHO do mỗi mã hộ có thể đại diện cho nhiều người.
 - + Một Mã hộ cũng có thể lặp lại trong bảng THONGTINTUVONG nếu có nhiều trường hợp tử vong liên quan đến cùng một mã hộ.
 - + Quan hệ giữa THANHVIENHO và THONGTINTUVONG là một mối quan hệ nhiều-một-nhiều.
- Về bảng DIABAN, do không có sự trùng khớp giữa thông tin địa bàn trong DIABAN và THONGTINHO, không thể kết nối trực tiếp bảng XA với THONGTINHO qua DIABAN. Ngoài ra, do thiếu dữ liệu trong cột Địa Bàn, không thể sử dụng DIABAN làm cầu nối giữa các bảng khác.

4. Các kiểu dữ liệu được sử dụng:

Kiểu Dữ liệu	Dùng trong các cột	Cách dùng
NVARCHAR	MaVung, MaTinh, MaHuyen, MaXa, MaDiaBan, TenVung, TenTinh,...	<p>- Là một kiểu dữ liệu chuỗi trong SQL Server, hỗ trợ Unicode và cho phép lưu trữ dữ liệu văn bản đa ngôn ngữ. Kiểu dữ liệu này có thể chứa ký tự từ bất kỳ bảng mã nào và thích hợp cho việc lưu trữ thông tin có sử dụng các ký tự đặc biệt</p> <p>- Được định nghĩa với một độ dài cụ thể, được biểu thị bằng số lượng ký tự tối đa mà nó có thể chứa, hoặc sử dụng MAX để chỉ định độ dài tối đa là 2GB.</p>
BIGINT	MaHo	<p>Trong lĩnh vực quản trị cơ sở dữ liệu, BIGINT đại diện cho một kiểu dữ liệu số nguyên, cho phép lưu trữ các giá trị với phạm vi rất rộng. Đối với các hệ thống cơ sở dữ liệu như SQL Server, BIGINT có thể lưu trữ các số từ -9,223,372,036,854,775,808 đến 9,223,372,036,854,775,807 khi được khai báo là kiểu có dấu. Trong trường hợp kiểu không dấu, BIGINT có thể chứa các giá trị từ 0 đến 18,446,744,073,709,551,615. Kiểu dữ liệu này thường được sử dụng khi các kiểu số nguyên nhỏ hơn như INT không đủ để biểu diễn các giá trị lớn</p>
SMALLINT	Nam, Namtuvong, Namsinh	<p>Là một kiểu dữ liệu số nguyên trong SQL, được sử dụng để lưu trữ các giá trị số nguyên có kích thước nhỏ. Kiểu dữ liệu SMALLINT chiếm 2 byte và có thể lưu trữ các giá trị từ -32,768 đến 32,767. Đây là một lựa chọn phù hợp khi bạn biết rằng dữ liệu số nguyên của mình không cần đến phạm vi lớn như INT hoặc BIGINT, giúp tiết kiệm không gian lưu trữ và tối ưu hóa hiệu suất của cơ sở dữ liệu.</p>

INT	MaThanhVien, Tuoi, Thangsinh, Thangtuvong,...	Là kiểu dữ liệu số nguyên, được sử dụng để lưu trữ các giá trị số nguyên, tức là các số không có phần thập phân. Kiểu dữ liệu int thường chiếm 4 byte (hoặc 32 bit) và có thể lưu trữ các giá trị từ -2,147,483,648 đến 2,147,483,647.
BIT	Gioitinh	Là dữ liệu số nguyên 0,1 hay NULL. Do giới tính trong dữ liệu chỉ có 2 trường hợp nên để Bit là hợp lý nhất.
FLOAT	KinhDo, ViDo	kiểu dữ liệu FLOAT được sử dụng để lưu trữ các số thực với độ chính xác cao. Kiểu dữ liệu này cho phép biểu diễn các giá trị nằm trong phạm vi từ -1,79E38 đến 1,79E38. FLOAT(n) quyết định kích thước của bộ nhớ được phân bổ cho trường dữ liệu, với hai lựa chọn là 4 byte hoặc 8 byte. Ví dụ, FLOAT(24) chỉ ra rằng trường dữ liệu sẽ chiếm 4 byte bộ nhớ, trong khi FLOAT(53) cho biết trường dữ liệu sẽ chiếm 8 byte. Nếu không có sự chỉ định cụ thể, giá trị mặc định của n sẽ là 53, tức là trường dữ liệu sẽ mặc định sử dụng 8 byte bộ nhớ để lưu trữ.
DATETIME	NgayPhongVan, NgayKetThuc	Là một kiểu dữ liệu được sử dụng để lưu trữ cả ngày và giờ cùng với phần giây có độ chính xác phân số dựa trên đồng hồ 24 giờ. Kiểu dữ liệu này cho phép bạn lưu trữ thông tin thời gian với độ chính xác đến 3 chữ số phân số giây và có phạm vi ngày từ 1 tháng 1 năm 1753 đến 31 tháng 12 năm 9999.
VARCHAR	MaNhanVien, Dienthoai, NguoiTao, Phienban	<ul style="list-style-type: none"> - VARCHAR là một kiểu dữ liệu được sử dụng để biểu diễn các giá trị chuỗi ký tự có độ dài biến đổi. - VARCHAR cho phép lưu trữ chuỗi ký tự với độ dài không cố định, nghĩa là bạn có thể lưu trữ chuỗi ký tự ngắn hơn so với độ dài tối đa đã khai báo mà không lãng phí không gian lưu trữ. Điều này khác biệt với kiểu dữ liệu CHAR, mà luôn chiếm đủ không gian lưu trữ dựa trên độ dài đã định nghĩa.

Bảng 2: Các kiểu dữ liệu được dùng trong các bảng

5. Sử dụng các cấu trúc điều khiển

5.1. Sử dụng cấu trúc điều khiển if ... else

```
-- Control structure if-else
-- Assume we want to check MaVung and update data if necessary
DECLARE @MaTinh NVARCHAR(255);
DECLARE @MaVung NVARCHAR(255);
SELECT @MaTinh = MaTinh, @MaVung = MaVung
FROM [dbo].[TINH]
WHERE MaTinh = 'T01';

IF @MaVung = 'V05'
BEGIN
    PRINT 'MaVung is V05';
```

```

UPDATE [dbo].[TINH]
SET TenTinh = 'New Province'
WHERE MaTinh = @MaTinh;
END
ELSE
BEGIN
    PRINT 'MaVung is not V05';
END

```

5.2.Thủ tục và con trỏ

```

-- Stored procedure
CREATE PROCEDURE THONGTIN(@MaHo BIGINT )
AS
BEGIN
    SELECT HoTenChuHo FROM THONGTINHO WHERE Hosoi = @MaHo
END

```

```

-- Procedure call
EXEC THONGTIN '198'

```

```

-- Cursor
CREATE PROCEDURE UpdateHuyen
AS
BEGIN
    DECLARE @MaHuyen NVARCHAR(255);
    DECLARE @TenHuyen NVARCHAR(255);
    DECLARE curHuyen CURSOR FOR
    SELECT MaHuyen, [TenHuyen]
    FROM Huyen
    WHERE MaHuyen = 'H012'

    OPEN curHuyen;
    FETCH NEXT FROM curHuyen INTO @MaHuyen, @TenHuyen;

    WHILE @@FETCH_STATUS = 0
    BEGIN
        UPDATE Huyen
        SET [TenHuyen] = 'New Name'
        WHERE MaHuyen = @MaHuyen;

        FETCH NEXT FROM curHuyen INTO @MaHuyen, @TenHuyen;
    END;

    CLOSE curHuyen;
    DEALLOCATE curHuyen;
END;

```

5.3. Giao tác

```
-- Transaction
BEGIN TRANSACTION;
BEGIN TRY
    UPDATE [dbo].[XA]
    SET TenXa = 'New Name'
    WHERE MaXa = 'X013';

    UPDATE [dbo].[DIABAN]
    SET TenDiaBan = 'New District'
    WHERE MaXa = 'X013';

    UPDATE [dbo].[THONGTINHO]
    SET HoTenChuHo = 'New Head of Household'
    WHERE MaXa = 'X013';

    COMMIT;
END TRY
BEGIN CATCH
    ROLLBACK;
    PRINT 'Error occurred. Transaction rolled back.';
END CATCH;
```

5.4. Hàm Function

```
-----HÀM FUNCTION ĐƯA RA MÃ HỘ CÓ MÃ Xã là X013---
-- Function
CREATE FUNCTION GetMaHoByMaxa (@Maxa NVARCHAR(255))
RETURNS TABLE
AS
RETURN (
    SELECT MaHo
    FROM [dbo].[THONGTINHO] WHERE MaXa = @Maxa
);

-- Function call
SELECT * FROM GetMaHoByMaxa('X013');
```

5.5. Trigger

```
-- Trigger
CREATE TRIGGER UpdateXaTrigger
ON [dbo].[XA]
AFTER UPDATE
AS
BEGIN
    IF UPDATE([TenXa])
```

```

BEGIN
  -- Check if Maxa = 'X013'
  IF EXISTS (SELECT 1 FROM inserted WHERE MaXa = 'X013')
  BEGIN
    UPDATE XA
    SET [TenXa] = 'No Data'
    WHERE MaXa = 'X013';
  END
END
END;

```

5.6. Tạo view

```

-- Create view for death statistics
CREATE VIEW VIEW_THONG_KE_TU_VONG AS
SELECT
  TTV.MaTuVong,
  TV.Hovaten,
  TTV.GioiTinh,
  TTV.NamSinh,
  TTV.NamTuVong,
  TTV.Tuoi,
  TTV.ThangTuVong,
  TTV.ThangSinh,
  TTV.SoThuTu
FROM
  [dbo].[THONGTINTUVONG] TTV
INNER JOIN
  [dbo].[THANHVIENTRONGHO] TV ON TTV.MaTuVong = TV.MaThanhVien;

-- Query the view to see results
SELECT * FROM VIEW_THONG_KE_TU_VONG;

```

CHƯƠNG IV: KẾT LUẬN

Kết thúc quá trình nghiên cứu và thực hành sâu rộng, chúng ta có thể thấy rằng cơ sở dữ liệu không chỉ là một công cụ lưu trữ thông tin mà còn là nền tảng vững chắc cho việc phân tích, quản lý và đưa ra quyết định trong mọi lĩnh vực của đời sống hiện đại. Qua thực hành, chúng ta đã có cơ hội áp dụng lý thuyết vào các tình huống thực tế, từ đó nâng cao kỹ năng và hiểu biết về việc thiết kế, triển khai và bảo trì cơ sở dữ liệu.

Những bài học được rút ra từ quá trình này không chỉ giới hạn ở kiến thức kỹ thuật mà còn mở rộng đến việc phát triển tư duy phản biện và giải quyết vấn đề. Cơ sở dữ liệu, với sự đa dạng về mô hình và công nghệ, đã chứng minh được sức mạnh trong việc hỗ trợ quyết định dựa trên dữ liệu và tối ưu hóa quy trình làm việc.

Nhìn về tương lai, việc tiếp tục học hỏi và cập nhật kiến thức về cơ sở dữ liệu sẽ là chìa khóa để chúng ta không chỉ thích nghi mà còn dẫn đầu trong kỷ nguyên số. Với sự phát triển không ngừng của công nghệ, việc sẵn sàng học hỏi và thích ứng sẽ đảm bảo rằng chúng ta có thể tận dụng tối đa tiềm năng của dữ liệu để phục vụ cho sự phát triển cá nhân và cộng đồng.

Tóm lại, tiểu luận này không chỉ là một hành trình khám phá kiến thức mà còn là một quá trình rèn luyện kỹ năng và thái độ làm việc chuyên nghiệp. Cơ sở dữ liệu không chỉ là một phần của công nghệ thông tin mà còn là một yếu tố quan trọng trong việc hình thành nền tảng cho thế hệ tương lai.

LỜI CẢM ƠN

Chúng em cảm ơn nhà trường và các thầy, cô đã tạo điều kiện cho chúng em được thực tập tại trường để chúng em được thực hành kiến thức SQL đã học

Nhóm 4 chúng em xin gửi lời cảm ơn chân thành tới thầy Mai Văn Thủy đã tận tình hướng dẫn, góp ý và giúp đỡ chúng em trong quá trình thực tập

Cuối cùng, nhóm 4 xin cảm ơn các thành viên trong nhóm vì đã hỗ trợ, giúp đỡ nhau hết mình để cùng nhau hoàn thành bài báo cáo thực tập, và đặc biệt cảm ơn các bạn cùng lớp vì đã giúp nhóm em những lúc nhóm em bị mắc lỗi sai

Chúng em xin chân thành cảm ơn,

Nhóm 4