

**BỘ CÔNG THƯƠNG**  
**TRƯỜNG ĐẠI HỌC KINH TẾ**  
**- KỸ THUẬT CÔNG NGHIỆP**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO TỔNG KẾT**  
**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC SINH VIÊN**  
**NĂM HỌC 2023**

Tên đề tài:

**TÌM HIỂU VỀ NGÔN NGỮ TRUY VẤN ĐA CHIỀU**  
**MDX TRÊN HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU**  
**SQL SERVER**

**Giảng viên hướng dẫn:** Th.s Vũ Mỹ Hạnh

**Chủ nhiệm đề tài:** Đàm Văn Anh      Lớp ĐHTin14A12HN

**Thành viên:** Nguyễn Thiện Thắng      Lớp ĐHTin14A12HN  
Bùi Thị Huyền Trang      Lớp ĐHTin14A12HN

**HÀ NỘI 04/2023**

**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC KINH TẾ  
- KỸ THUẬT CÔNG NGHIỆP**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO TỔNG KẾT  
ĐỀ TÀI NGHIÊN CỨU KHOA HỌC SINH VIÊN  
NĂM HỌC 2023**

Tên đề tài:

**TÌM HIỂU VỀ NGÔN NGỮ TRUY VẤN ĐA CHIỀU  
MDX TRÊN HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU  
SQL SERVER**

**Giảng viên hướng dẫn: Th.s Vũ Mỹ Hạnh**

**Chủ nhiệm đề tài: Đàm Văn Anh                      Lớp ĐHTin14A12HN**

**Thành viên: Nguyễn Thiện Thắng                      Lớp ĐHTin14A12HN**

**Bùi Thị Huyền Trang                      Lớp ĐHTin14A12HN**

**HÀ NỘI 04/2023**

## DANH SÁCH NHỮNG NGƯỜI THỰC HIỆN

**Chủ nhiệm đề tài:** Sinh viên Đàm Văn Anh

**Lớp:** Tin 14A12 HN

**Giảng viên hướng dẫn:** Ths. Vũ Mỹ Hạnh

**Khoa:** Công nghệ thông tin

**Thành viên thực hiện:**

Họ và tên	Lớp	Khoa
Nguyễn Thiên Thắng	Tin 14A12 HN	Công nghệ thông tin
Bùi Thị Huyền Trang	Tin 14A12 HN	Công nghệ thông tin

## MỤC LỤC

DANH SÁCH NHỮNG NGƯỜI THỰC HIỆN .....	3
MỤC LỤC .....	4
DANH MỤC BIỂU, BẢNG .....	6
DANH MỤC HÌNH .....	7
TRANG KÝ HIỆU VÀ CHỮ VIẾT TẮT .....	9
LỜI MỞ ĐẦU .....	10
Chương 1: Đặt vấn đề.....	11
1.1. Giới thiệu các quy định tuyển sinh đại học của trường Đại học Kinh tế Kỹ thuật Công nghiệp.....	11
1.1.1. Thông tin ngành và chỉ tiêu tuyển sinh dự kiến .....	11
1.1.2. Phương thức tuyển sinh .....	13
1.2. Quy trình xét tuyển và cách thức hoạt động của hệ thống xử lý nguyện vọng xét tuyển đại học, cao đẳng .....	15
1.2.1. Nguyên tắc xét tuyển .....	15
1.2.2. Điều chỉnh nguyện vọng.....	16
1.3. Đề xuất bài toán và hướng giải quyết.....	17
Chương 2: Lý thuyết cơ sở .....	18
2.1. Kho dữ liệu .....	18
2.1.1. Kho dữ liệu là gì? .....	18
2.1.2. Ứng dụng của kho dữ liệu .....	20
2.1.3. Phân biệt Kho dữ liệu và Cơ sở dữ liệu truyền thống .....	21
2.1.4. Các mô hình biểu diễn kho dữ liệu.....	21
2.1.5. Các giai đoạn thiết kế kho dữ liệu .....	24
2.2. Ngôn ngữ truy vấn dữ liệu đa chiều MDX.....	25
2.2.1. Giới thiệu khái quát về ngôn ngữ MDX.....	25
2.2.2. Một số cú pháp cơ bản của ngôn ngữ MDX .....	25
2.3 Tổng quan về phân tích và xử lý trực tuyến OLAP .....	26
2.3.1. OLAP là gì? .....	26
2.3.2. Quy trình phân tích và xử lý trực tuyến bằng OLAP trên kho dữ liệu...	26

Chương 3: Mô phỏng quá trình xây dựng và truy vấn MDX trên kho dữ liệu đa chiều .....	28
3.1. Xây dựng và cài đặt kho dữ liệu tuyển sinh .....	28
3.2. Mô phỏng quá trình truy vấn MDX trên kho dữ liệu tuyển sinh.....	29
Chương 4: Triển khai ứng dụng .....	31
4.1. Triển khai và cài đặt các công cụ hỗ trợ.....	31
4.1.1. Cài đặt Microsoft SQL Server 2019 Developer .....	31
4.1.2. Cài đặt SQL Server Management Studio .....	32
4.1.3. Cài đặt Visual Studio 2017 Community.....	33
4.1.4. Cài đặt Microsoft SQL Server Data Tools .....	33
4.2. Xây dựng và cài đặt kho dữ liệu Tuyển sinh.....	34
4.3. Thiết kế giao diện chương trình .....	40
4.4. Lập trình chương trình Quản lí tuyển sinh .....	45
4.4.1 Các thư viện sử dụng trong lập trình WinForms .....	45
4.4.2 Triển khai lập trình .....	46
KẾT LUẬN VÀ KIẾN NGHỊ.....	48
TÀI LIỆU THAM KHẢO .....	49
LỜI CẢM ƠN.....	50

## **DANH MỤC BIỂU, BẢNG**

Bảng 1.1. Bảng thông tin ngành và chỉ tiêu tuyển sinh **Error! Bookmark not defined.**

Bảng 1.2. Mã tổ hợp các môn xét tuyển..... **Error! Bookmark not defined.**

Bảng 2.1. Bảng so sánh Kho dữ liệu và Cơ sở dữ liệu truyền thống ..... 21

## DANH MỤC HÌNH

Hình 2.1. Ví dụ về lược đồ hình sao.....	23
Hình 2.2. Ví dụ về lược đồ bông tuyết .....	24
Hình 3.1. Lược đồ hình sao của hệ thống Tuyển sinh.....	28
Hình 4. 1. Cài đặt SQL Server.....	31
Hình 4. 2. Chọn tính năng trong SQL Server .....	31
Hình 4. 3. Chọn chế độ cho server của Analysis Services .....	32
Hình 4. 4. Cài SQL Server Management Studio .....	32
Hình 4. 5. Kết nối SQL Server trong SQL Server Management Studio .....	33
Hình 4. 6. Bảng chọn chức năng của phần mềm Visual Studio .....	33
Hình 4. 7. Cài đặt Microsoft SQL Server Data Tools .....	33
Hình 4. 8. Tạo Project Analysis Services mới trong Visual Studio .....	34
Hình 4. 9. Nhập tên Project để hoàn thành tạo Project Analysis Services .....	34
Hình 4. 10. Thêm Data Source trong Analysis Services .....	35
Hình 4. 11. Kết nối Database trong Analysis Services .....	35
Hình 4. 12. Thêm Data Source View trong Analysis Services .....	36
Hình 4. 13. Thêm Include objects trong Analysis Services .....	36
Hình 4. 14. Thêm Cubes trong Analysis Services.....	36
Hình 4. 15. Chọn measure group tables cho cube trong Analysis Service .....	37
Hình 4. 16. Dimension THISINH.....	37
Hình 4. 17. Dimension NGANH .....	37
Hình 4. 18. Dimension DOITUONGUUTIENT .....	38
Hình 4. 19. Dimension TOHOPMON .....	38
Hình 4. 20. Dimension KHUVUC .....	38
Hình 4. 21. Cấp quyền quản lý database cho user.....	39
Hình 4. 22.. Tiến hành Process Project .....	40
Hình 4. 23. Giao diện đăng nhập.....	40
Hình 4. 24. Giao diện chính của chương trình .....	41
Hình 4. 25. Giao diện chức năng thống kê .....	41
Hình 4. 26. Chức năng truy vấn đa chiều của hệ thống .....	42
Hình 4. 27. Chức năng truy vấn cho nhà quản lí.....	43
Hình 4. 28. Chức năng dự đoán điểm chuẩn .....	43
Hình 4. 29. Chức năng thông tin điểm chuẩn.....	44
Hình 4. 30. Chức năng tra cứu điểm thí sinh.....	44
Hình 4. 31. Quy trình kết nối và đưa dữ liệu lên Form .....	45
Hình 4. 32. Cửa sổ thêm thư viện ADOMD.NET.....	46

Hình 4. 33. Cửa sổ thêm thư viện DotNetBar2 .....	46
---	----



## TRANG KÝ HIỆU VÀ CHỮ VIẾT TẮT

Ký hiệu viết tắt	Ý nghĩa
CSDL	Cơ sở dữ liệu
WH	Data WareHouse – Kho dữ liệu
DM	Data Mining – Khai phá dữ liệu
MDX	MultiDemensional eXpression - Ngôn ngữ truy vấn MDX
DDL	Ngôn ngữ định nghĩa dữ liệu.
OLAP	Online Analysic Process - Xử lý phân tích trực tuyến

## LỜI MỞ ĐẦU

Hệ thống tuyển sinh đại học là quy trình quản lý và chọn lọc ứng viên để đăng ký và tham gia học tập tại các trường đại học. Hệ thống này thường được quản lý bởi các cơ quan chức năng của chính phủ hoặc bởi các trường đại học, và có thể có những sự khác biệt về quy trình tuyển sinh giữa các quốc gia và các trường học.

Các quy trình tuyển sinh đại học bao gồm đăng ký, thi và xếp hạng thí sinh dựa trên các tiêu chí như điểm thi, thành tích học tập, kinh nghiệm, đặc điểm cá nhân và tiềm năng. Ngoài ra, hệ thống tuyển sinh cũng có thể yêu cầu các thủ tục như viết đơn xin học, nộp hồ sơ, phỏng vấn và kiểm tra năng lực tiếng Anh.

Quá trình tuyển sinh đại học có ý nghĩa quan trọng trong việc xác định những người sẽ được đào tạo và trở thành những chuyên gia trong các lĩnh vực khác nhau. Do đó, hệ thống tuyển sinh đại học phải được thực hiện một cách công bằng và đúng quy trình để đảm bảo sự chất lượng và công bằng cho tất cả các thí sinh.

Với thực trạng của hệ thống tuyển sinh thực tế, nhằm hỗ trợ quá trình truy xuất và phân tích dữ liệu đa dạng và phức tạp của Thí sinh và Nhà trường, nhằm trợ giúp các quyết định đúng đắn. Nhóm chúng em đã cùng đề xuất đề tài “Tìm hiểu về ngôn ngữ truy vấn đa chiều MDX trên hệ quản trị cơ sở dữ liệu SQL Server”, từ đó ứng dụng trong việc xây dựng Kho dữ liệu tuyển sinh hỗ trợ truy vấn MDX và phân tích trực tuyến OLAP nhằm đưa ra các thống kê cũng như các phân tích về dữ liệu về tuyển sinh trợ giúp người dùng.

Đề tài gồm 4 chương, bao gồm:

Chương 1: Đặt vấn đề

Chương 2: Lý thuyết cơ sở

Chương 3: Mô phỏng quá trình xây dựng và truy vấn MDX trên kho dữ liệu đa chiều.

Chương 4: Triển khai ứng dụng

Trong đó, *Chương 1*: mô tả hệ thống tuyển sinh đại học và quy trình xét tuyển đại học theo nguyện vọng từ đó nêu ra bài toán và hướng giải quyết. *Chương 2*: Trình bày lý thuyết cơ sở về Kho dữ liệu, ngôn ngữ MDX, phân tích trực tuyến OLAP. *Chương 3*: Trình bày về quá trình xây dựng kho dữ liệu Tuyển sinh và mô phỏng quá trình truy vấn MDX trên kho dữ liệu tuyển sinh. *Chương 4*: Trình bày về việc xây dựng chương trình mô phỏng hỗ trợ tìm hiểu về ngôn ngữ MDX và hỗ trợ Thí sinh và Nhà trường trong công tác tuyển sinh và hỗ trợ tuyển sinh.

## Chương 1: Đặt vấn đề

### 1.1. Giới thiệu các quy định tuyển sinh đại học của trường Đại học Kinh tế Kỹ thuật Công nghiệp

Tìm hiểu về hệ thống tuyển sinh đại học và các quy định tuyển sinh của trường Đại học Kinh tế Kỹ thuật Công nghiệp, chúng em thu thập được những thông tin sau:

#### 1.1.1. Thông tin ngành và chỉ tiêu tuyển sinh dự kiến

Bảng 1.1. Bảng thông tin ngành và chỉ tiêu tuyển sinh

TT	Mã ngành	Tên ngành	Tổ hợp xét tuyển
1	7220201	Ngành Ngôn ngữ anh - Chuyên ngành Biên phiên dịch - Chuyên ngành Phương pháp giảng dạy	A01, D01, D09 và D14
2	7340101	Ngành Quản trị kinh doanh - Chuyên ngành Quản trị kinh doanh tổng hợp - Chuyên ngành Quản trị nhân lực - Chuyên ngành Quản trị bán hàng	A00; A01; C01 và D01
3	7340115	Ngành Marketing - Chuyên ngành Quản trị Marketing - Chuyên ngành Truyền thông Marketing	A00; A01; C01 và D01
4	7340121	Ngành Kinh doanh thương mại - Chuyên ngành Kinh doanh thương mại - Chuyên ngành Thương mại điện tử - Chuyên ngành Kinh doanh quốc tế	A00; A01; C01 và D01
5	7510605	Ngành Logistics & Quản lý chuỗi cung ứng	A00; A01; C01 và D01
6	7340201	Ngành Tài chính - Ngân hàng - Chuyên ngành Tài chính doanh nghiệp - Chuyên ngành Ngân hàng - Chuyên ngành Công nghệ tài chính - Chuyên ngành Đầu tư tài chính	A00; A01; C01 và D01
7	7340204	Ngành Bảo hiểm - Chuyên ngành Định phí bảo hiểm - Chuyên ngành Kinh tế bảo hiểm - Chuyên ngành Bảo hiểm xã hội	A00; A01; C01 và D01
8	7340301	Ngành Kế toán - Chuyên ngành Kế toán doanh nghiệp - Chuyên ngành Kế toán nhà hàng khách sạn	A00; A01; C01 và D01

		- Chuyên ngành Kế toán công	
9	7340302	Kiểm toán	A00; A01; C01 và D01
10	7480102	Ngành Mạng máy tính và TTDL - Chuyên ngành Mạng máy tính & truyền thông dữ liệu - Chuyên ngành An toàn thông tin	A00; A01; C01 và D01
11	7480108	Công nghệ kỹ thuật máy tính - Chuyên ngành Công nghệ nhúng và lập trình di động - Chuyên ngành Trí tuệ nhân tạo và điều khiển thông minh	A00; A01; C01 và D01
12	7480201	Ngành Công nghệ thông tin - Chuyên ngành Công nghệ thông tin - Chuyên ngành Hệ thống thông tin	A00; A01; C01 và D01
13	7510201	Ngành CNKT cơ khí - Chuyên ngành Công nghệ chế tạo máy - Chuyên ngành Máy và cơ sở thiết kế máy - Chuyên ngành Kỹ thuật đo và gia công thông minh	A00; A01; C01 và D01
14	7510203	Ngành CNKT cơ điện tử - Chuyên ngành Kỹ thuật Robot - Chuyên ngành Hệ thống Cơ điện tử - Chuyên ngành Hệ thống Cơ điện tử trên Ô tô	A00; A01; C01 và D01
15	7510205	Ngành Công nghệ kỹ thuật Ô tô - Chuyên ngành Công nghệ kỹ thuật Ô tô - Chuyên ngành Máy – Gầm ô tô - Chuyên ngành Điện – Cơ điện tử ô tô	A00; A01; C01 và D01
16	7510301	Ngành CNKT điện, điện tử - Chuyên ngành Điện dân dụng và công nghiệp - Chuyên ngành Hệ thống cung cấp điện - Chuyên ngành Thiết bị điện, điện tử	A00; A01; C01 và D01
17	7510302	Ngành CNKT điện tử – viễn thông - Chuyên ngành Hệ thống nhúng và điện tử ứng dụng - Chuyên ngành Truyền thông số và IoT	A00; A01; C01 và D01
18	7510303	Ngành CNKT điều khiển và tự động hoá - Chuyên ngành Tự động hóa công nghiệp - Chuyên ngành Kỹ thuật điều khiển và hệ thống thông minh - Chuyên ngành Thiết bị đo thông minh	A00; A01; C01 và D01

19	7540202	Ngành Công nghệ sợi, dệt	A00; A01; C01 và D01
20	7540204	Ngành Công nghệ dệt, may - Chuyên ngành Công nghệ may - Chuyên ngành Thiết kế Thời trang	A00; A01; C01 và D01
21	7540101	Ngành Công nghệ thực phẩm - Chuyên ngành Công nghệ lên men - Chuyên ngành Công nghệ bảo quản và Chế biến thực phẩm - Chuyên ngành Quản lý chất lượng thực phẩm	A00; A01; B00 và D07
22	7460108	Ngành Khoa học dữ liệu - Chuyên ngành Khoa học dữ liệu - Chuyên ngành Phân tích dữ liệu	A00; A01; C01 và D01
23	7810103	Ngành Quản trị dịch vụ du lịch và lữ hành - Chuyên ngành Quản trị lữ hành - Chuyên ngành Quản trị khách sạn	A00; A01; C00 và D01
24	7810201	Ngành Quản trị khách sạn	A00; A01; C00 và D01

**Ghi chú:**

- Tổng chỉ tiêu tuyển sinh năm đại học 2023 dự kiến là: 4.868 chỉ tiêu;
- Dành tối đa 20% chỉ tiêu để xét tuyển tại cơ sở Nam Định;
- Nếu chỉ tiêu xét tuyển theo từng phương thức trước chưa đủ, chỉ tiêu còn lại được chuyển sang xét tuyển theo các phương thức xét tuyển sau.

\* Bảng mã các tổ hợp xét tuyển:

Bảng 1.2. Mã tổ hợp các môn xét tuyển

<b>Mã tổ hợp các môn xét tuyển</b>			
<b>kết quả thi tốt nghiệp THPT và kết quả học tập THPT</b>			
A00	Toán - Vật lý - Hóa học	D01	Toán - Anh văn - Ngữ văn
A01	Toán - Vật lý - Anh văn	D07	Toán - Hóa - Anh văn
B00	Toán - Sinh học - Hóa học	D09	Toán - Lịch sử - Anh văn
C01	Toán - Vật lý - Ngữ văn	D10	Toán - Địa lý - Anh văn
C00	Ngữ văn - Lịch sử - Địa lý	D14	Ngữ văn - Lịch sử - Anh văn

\* Quy định chênh lệch điểm xét tuyển giữa các tổ hợp: Các tổ hợp môn xét tuyển có giá trị ngang nhau trong cùng đợt xét tuyển.

**1.1.2. Phương thức tuyển sinh**

Năm 2023, Trường Đại học Kinh tế - Kỹ thuật Công nghiệp tổ chức tuyển sinh đại học hệ chính quy theo các phương thức:

**Phương thức 1:** Xét tuyển theo kết quả Kỳ thi đánh giá tư duy do Đại học Bách khoa Hà Nội chủ trì tổ chức

\* *Đối tượng và điều kiện xét tuyển:*

- Thí sinh tham dự thi kỳ thi đánh giá tư duy năm 2023 do Đại học Bách khoa Hà Nội chủ trì tổ chức;

- Đủ điều kiện tốt nghiệp THPT theo quy định;

\* *Chỉ tiêu tuyển sinh, ngành và tổ hợp môn xét tuyển:* Dành tối đa 5 % trong tổng chỉ tiêu tuyển sinh được xác định để xét tuyển theo phương thức này.

\* *Xác định điểm xét tuyển:* Điểm xét từng ngành được xác định theo điểm tổng của bài thi đánh giá tư duy được quy đổi về thang điểm 30.

+ *Lịch tuyển sinh và cách thức đăng ký dự tuyển:* Theo quy định của Bộ Giáo dục và Đào tạo và Thông báo của Đại học Bách khoa Hà Nội

**Phương thức 2:** Xét kết quả kỳ thi đánh giá năng lực do ĐH Quốc gia Hà Nội tổ chức

\* *Đối tượng và điều kiện xét tuyển:*

- Thí sinh tham dự thi kỳ thi đánh giá năng lực năm 2023 do Đại học Quốc gia Hà Nội chủ trì tổ chức;

- Đủ điều kiện tốt nghiệp THPT theo quy định;

\* *Chỉ tiêu tuyển sinh:* Dành tối đa 5% trong tổng chỉ tiêu tuyển sinh được xác định để xét tuyển theo phương thức này.

\* *Xác định điểm xét tuyển:* Điểm xét từng ngành được xác định theo điểm tổng của bài thi đánh giá năng lực được quy đổi về thang điểm 30.

\* *Lịch tuyển sinh và cách thức đăng ký dự tuyển:* Kỳ thi đánh giá năng lực dự kiến tổ chức thành nhiều đợt và tại các điểm thi khác nhau. Thời gian đăng ký xét tuyển và xét tuyển theo quy định của Bộ Giáo dục và Đào tạo.

**Phương thức 3:** Xét tuyển kết quả kỳ thi tốt nghiệp THPT năm 2023

\* *Đối tượng và điều kiện xét tuyển:*

- Thí sinh tham dự thi kỳ thi tốt nghiệp THPT năm 2023 và có đăng ký sử dụng kết quả để xét tuyển đại học;

- Đủ điều kiện tốt nghiệp THPT theo quy định;

- Đạt ngưỡng điểm xét tuyển tối thiểu do Bộ Giáo dục và Đào tạo quy định.

- Đối với ngành Ngôn ngữ anh, điểm môn Tiếng anh trong tổ hợp xét tuyển tối thiểu đạt từ 6,00 điểm trở lên.

\* *Chỉ tiêu tuyển sinh, ngành và tổ hợp môn xét tuyển:* Dành tối đa 70% trong tổng chỉ tiêu tuyển sinh được xác định để xét tuyển theo phương thức này.

\* *Xác định điểm trúng tuyển (ĐTT)*

- Xét kết quả kỳ thi tốt nghiệp THPT năm 2023.

-  $ĐTT = ĐXT + ĐUT$ , trong đó:

+ ĐTT: Điểm trúng tuyển;

- + ĐXT (đối với những ngành không có môn chính): Điểm xét tuyển được xác định bằng tổng điểm 3 môn thuộc tổ hợp môn đăng ký xét tuyển;
- + ĐXT (đối với Ngành ngôn ngữ anh): Điểm xét tuyển được xác định bằng (điểm môn 1 + điểm môn 2 + môn tiếng anh x 2) x 3/4 thuộc tổ hợp môn đăng ký xét tuyển;
- + ĐUT: Điểm ưu tiên theo quy định hiện hành trong Quy chế tuyển sinh đại học hệ chính quy của Bộ GD&ĐT.

\* *Hồ sơ tuyển sinh*: Phiếu đăng ký xét tuyển theo mẫu của Bộ GD&ĐT.

\* *Lịch tuyển sinh*: Theo quy định của Bộ Giáo dục và Đào tạo.

\* *Cách thức đăng ký xét tuyển*: Theo quy định của Bộ GD&ĐT.

- Thí sinh có nguyện vọng học tập tại cơ sở Hà Nội của Nhà trường ghi mã đăng ký xét tuyển vào sau mã ngành chuẩn.
- Thí sinh có nguyện vọng học tập tại cơ sở Nam Định của Nhà trường ghi mã đăng ký xét tuyển DKD vào sau mã ngành chuẩn.

**Phương thức 4:** Xét tuyển kết quả học bạ THPT

\* Đối tượng và điều kiện xét tuyển:

- Thí sinh đã tốt nghiệp THPT hoặc tương đương;

\* Thời gian và cách thức đăng ký xét tuyển: Theo quy định của Bộ GD&ĐT và theo thông báo trong từng đợt xét tuyển được Nhà trường đăng tải trên trang thông tin điện tử theo địa chỉ [www.uneti.edu.vn](http://www.uneti.edu.vn)

**Phương thức khác:** Xét tuyển thẳng đối với những thí sinh thuộc các đối tượng được quy định tại khoản 2 điều 7 Quy chế tuyển sinh trình độ đại học; tuyển sinh trình độ cao đẳng ngành đào tạo giáo viên mầm non hiện hành.

- Tuyển thẳng theo đề án tuyển sinh riêng của Nhà trường : Tiêu chí xét tuyển thẳng áp dụng theo Đề án tuyển sinh của Trường Đại học Kinh tế - Kỹ thuật Công nghiệp đối với thí sinh tốt nghiệp THPT năm 2023 và các năm trước. Chi tiết các tiêu chí trong thông báo tuyển sinh sẽ được đăng tải trên hệ thống website: [www.tuyensinh.uneti.edu.vn](http://www.tuyensinh.uneti.edu.vn)

## **1.2. Quy trình xét tuyển và cách thức hoạt động của hệ thống xử lý nguyện vọng xét tuyển đại học, cao đẳng**

Tham khảo quy trình tuyển sinh của Bộ Giáo dục và Đào tạo năm 2022, chúng em thu được những thông tin cụ thể như sau:

### **1.2.1. Nguyên tắc xét tuyển**

Năm 2022, các cơ sở đào tạo tự tổ chức xét tuyển theo các phương thức và tiêu chí xét tuyển riêng hoặc tự nguyện phối hợp theo nhóm để tổ chức xét tuyển theo các phương thức và tiêu chí xét tuyển chung. Tuy vậy, các cơ sở đào tạo vẫn cần tuân thủ các nguyên tắc xét tuyển cơ bản theo như Quy chế tuyển sinh đại học, cao đẳng mà Bộ đã đưa ra trước đó, bao gồm:

Điểm trúng tuyển vào trường không thấp hơn ngưỡng đầu vào do Bộ công bố.

Không phân biệt thứ tự ưu tiên của nguyện vọng đăng ký đối với một ngành đào tạo theo một phương thức và tổ hợp môn. Trong trường hợp thí sinh đỗ nhiều nguyện vọng, nguyện vọng cao nhất sẽ được ghi nhận và tự động loại bỏ nguyện vọng thấp.

Trường hợp có nhiều thí sinh bằng điểm nhau khi xét tuyển ở cuối danh sách, các cơ sở đào tạo có thể sử dụng các tiêu chí phụ khác nhau, trong đó có thể sử dụng thứ tự nguyện vọng ưu tiên của thí sinh.

Việc sử dụng điểm ưu tiên khu vực tại các cơ sở đào tạo phải đảm bảo đồng bộ, thống nhất và bám sát quy định về điểm ưu tiên do Bộ quy định.

### **1.2.2. Điều chỉnh nguyện vọng.**

Sau khi thí sinh đăng ký nguyện vọng trên hệ thống, các cơ sở đào tạo tải thông tin, dữ liệu từ hệ thống để phục vụ xét tuyển. Dữ liệu của thí sinh bao gồm: Dữ liệu đăng ký nguyện vọng và kết quả thi tốt nghiệp trung học phổ thông, kết quả học tập cấp trung học phổ thông của những thí sinh dự tuyển vào cơ sở đào tạo theo các phương thức tuyển sinh.

Hệ thống xử lý nguyện vọng sẽ tự động loại bỏ khỏi danh sách những nguyện vọng thấp của thí sinh đủ điều kiện trúng tuyển nhiều nguyện vọng, trả lại danh sách thí sinh dự kiến trúng tuyển theo nguyện vọng cao nhất. Vì vậy, thí sinh cần cân nhắc thứ tự đặt nguyện vọng, trong trường hợp thí sinh trúng tuyển nhiều nguyện vọng, nguyện vọng cao nhất sẽ được ghi nhận và những nguyện vọng sau sẽ bị loại bỏ.

Căn cứ vào kết quả xử lý nguyện vọng, cơ sở đào tạo lặp lại quy trình xét tuyển ở chu kỳ sau, điều chỉnh điểm trúng tuyển sao cho phù hợp với chỉ tiêu trong thời hạn quy định. Ở chu kỳ cuối, sẽ quyết định điểm trúng tuyển vào các ngành, chương trình đào tạo (theo các phương thức tuyển sinh) và tải lên hệ thống danh sách (chính thức) thí sinh đủ điều kiện trúng tuyển.

Trên cơ sở kết quả xử lý nguyện vọng cuối cùng, danh sách thí sinh trúng tuyển vào các ngành và chương trình đào tạo sẽ được công bố chính thức kèm theo công bố điểm trúng tuyển (và các điều kiện, tiêu chí phụ nếu có) vào các ngành, chương trình đào tạo theo các phương thức tuyển sinh; tổ chức cho thí sinh tra cứu kết quả xét tuyển của cá nhân trên trang thông tin điện tử của cơ sở đào tạo (hoặc của nhóm cơ sở đào tạo phối hợp xét tuyển).

Năm 2022, thí sinh có chưa đầy 1 tháng để đăng ký và thay đổi nguyện vọng xét tuyển. Trong khoảng thời gian xét tuyển, thí sinh cần theo dõi sát sao hệ thống danh sách thí sinh dự kiến trúng tuyển mà các cơ sở đào tạo tải lên Cổng thông tin tuyển sinh để cân nhắc tỷ lệ chọi. Thí sinh cần xem xét kỹ lưỡng với số điểm mình đạt được, kết hợp đối chiếu với chỉ tiêu của nhà trường để dự đoán xác suất đỗ của mình là bao nhiêu. Thí sinh cũng cần phân tích số lượng thí sinh nộp hồ sơ vào, tham chiếu điểm chuẩn của các năm trước để có những chọn lựa đúng đắn.



### **1.3. Đề xuất bài toán và hướng giải quyết**

Trong quá trình tìm hiểu về các quy định tuyển sinh của trường Trường Đại học Kinh tế - Kỹ thuật Công nghiệp và quy trình xét tuyển đại học (*tham khảo năm 2022*). Chúng em nhận thấy, để một trường có thể chủ động trong công tác tuyển sinh và hỗ trợ tư vấn tuyển sinh, rất cần một hệ thống hỗ trợ phân tích dữ liệu tuyển sinh nhằm trợ giúp thí sinh đưa ra quyết định lựa chọn và thay đổi nguyện vọng đúng đắn. Đồng thời, trợ giúp cho Ban tuyển sinh các trường có thể thuận tiện hơn trong việc quyết định đưa ra điểm chuẩn phù hợp cho các ngành khác nhau trong trường.

Trên nhu cầu thực tế đó, chúng em đề xuất hướng giải quyết cho bài toán này đồng thời là hướng phát triển cho đề tài như sau:

- Xây dựng kho dữ liệu về dữ liệu tuyển sinh
- Thực hiện các truy vấn đa chiều bằng MDX trên kho dữ liệu và kết hợp với kỹ thuật phân tích trực tuyến OLAP giúp đưa ra những thông kê mang tính nhất thời nhằm hỗ trợ người dùng ra quyết định phù hợp
- Xây dựng chương trình ứng dụng cho người dùng thuận tiện trong tương tác.

## Chương 2: Lý thuyết cơ sở

### 2.1. Kho dữ liệu

#### 2.1.1. Kho dữ liệu là gì?

##### 2.1.1.1. Định nghĩa kho dữ liệu

**Kho dữ liệu** ban đầu được định nghĩa là một tập dữ liệu được dùng để hỗ trợ quy trình ra quyết định. Theo quá trình phát triển, Kho dữ liệu được định nghĩa là một môi trường thông tin (information environment) (từ môi trường thông tin thể hiện nó không phải là một sản phẩm (product))

##### 2.1.1.2. Đặc điểm của kho dữ liệu

Dữ liệu được lưu trong Kho dữ liệu không được tạo ra trực tiếp từ người dùng mà được lấy từ các nguồn dữ liệu sẵn có và mục đích là phục vụ tạo ra các báo cáo quản trị do đó nó có các tính chất sau:

##### **Hướng chủ đề (subject-oriented)**

Mục đích của Kho dữ liệu là phục vụ các yêu cầu phân tích, hoặc khai phá cụ thể được gọi là chủ đề. Ví dụ với chủ đề phân tích nhân sự thì có thể bao gồm các độ đo về doanh thu của từng người, số ngày nghỉ trong tháng, số dự án tham gia trong tháng, theo các chiều phân tích: thời gian, chi nhánh, sản phẩm, ...

Một sự so sánh dễ hiểu, giống như chẩn đoán một bệnh ví dụ bệnh liên quan đến tim, thì bác sỹ cần quan tâm không chỉ một mà một vài chỉ số như các chỉ số liên quan đến máu, chỉ số về huyết áp, nhịp tim, điện tâm đồ. Ngoài ra còn cần theo dõi theo thời gian (có thể là hàng ngày) để xem xét sự thay đổi mà có phương pháp điều trị kịp thời. Trong trường hợp này thời gian được gọi là chiều phân tích. Để chẩn đoán được chính xác thì cần đầy đủ các thông tin về các chỉ số trên, và cũng không cần các chỉ số khác lẫn vào làm nhiễu quá trình chẩn đoán và cũng không cần thiết. Việc tổ chức dữ liệu theo chủ đề này sẽ dẫn đến nhu cầu tổ chức lưu trữ dữ liệu khác với các cơ sở dữ liệu tác nghiệp.

##### **Được tích hợp (integrated)**

Tại một bệnh viện, các phòng khác nhau sẽ thực hiện các xét nghiệm khác nhau, do đó để có được đầy đủ thông tin phục vụ chẩn đoán thì cần thu thập được kết quả từ nhiều nguồn. Điều này hoàn toàn tương tự như tại doanh nghiệp, dữ liệu cần để phân tích có thể nằm rải rác ở nhiều hệ thống tác nghiệp khác nhau, và vì vậy cần tích hợp lại. Quá trình tích hợp này sẽ được thực hiện trong quá trình ETL như đã trình bày ở bài trước. Việc tổng hợp dữ liệu từ nhiều nguồn vào một kho dữ liệu cho phép chúng ta có thể xem đồng thời nhiều nhóm chỉ tiêu khác nhau (từ nhiều hệ thống nghiệp vụ khác nhau), ví dụ ta có thể xem chỉ tiêu doanh thu ở nhiều mảng nghiệp vụ khác nhau để có thể so sánh được sự tương quan giữa các mảng nghiệp vụ này. Điều này cũng giống trong chẩn đoán bệnh ta có thể cần nhiều xét nghiệm (thử máu, thử nước tiểu, siêu âm, ...) và kiểm tra khác nhau để có thể đưa ra kết luận chính xác.

### ***Có gán nhãn thời gian (time variant):***

Như đã đề cập, với các chỉ số thay đổi liên tục (như huyết áp, nhịp tim) việc chẩn đoán bệnh sẽ cần dữ liệu của các hôm trước để so sánh phục vụ quá trình điều trị. Do đó hàng ngày cần phải lưu lại giá trị của các chỉ số này. Hay nói cách khác các chỉ số này khi lưu sẽ được gán 1 nhãn thời gian tương ứng. Tương tự như vậy, dữ liệu lịch sử có tầm quan trọng đặc biệt trong phân tích dữ liệu, cùng một độ đo sẽ có nhiều giá trị khác nhau trong lịch sử có thể dùng để so sánh với nhau để biết được sự thay đổi là tốt hay xấu.

Ví dụ, độ đo doanh thu của một mặt hàng của tháng hiện tại, nếu đem so sánh với doanh thu của mặt hàng đó trong tháng trước, tháng này năm trước thì sẽ có nhiều thông tin hơn để đánh giá doanh thu của mặt hàng đó là tốt hay không, trên cơ sở đó sẽ có các quyết định phù hợp. Ngoài ra, dữ liệu lịch sử còn cho phép dự báo được tương lai khi ứng dụng khai phá dữ liệu.

### ***Bất biến (non-volatile):***

Khác với các cơ sở dữ liệu (CSDL) giao dịch, nơi thông tin của một đối tượng có thể được cập nhật thay đổi hàng ngày, dữ liệu trong Kho dữ liệu có chức năng báo cáo lại các chỉ số về hoạt động thực tế đã xảy ra. Do đó dữ liệu trong Kho dữ liệu không thể cập nhật, thay đổi vì nó sẽ không phản ánh đúng thực tế. Do đó với kho dữ liệu chỉ có 2 thao tác chính là tải dữ liệu vào kho và truy cập (đọc) dữ liệu từ kho.

Trường hợp sau khi tổng hợp dữ liệu mà dữ liệu ở trong nguồn bị thay đổi, khi đó một giải pháp xử lý là thực hiện lại quá trình ETL để tải lại dữ liệu từ nguồn vào Kho dữ liệu chứ không cho phép sửa đổi dữ liệu ở trong Kho. Một ý nghĩa khác của tính chất này là dữ liệu lịch sử vẫn được bảo tồn, vẫn có ý nghĩa chứ không như dữ liệu tác nghiệp ở các cơ sở dữ liệu, dữ liệu cũ (năm trước, hoặc tháng trước) không có ý nghĩa phục vụ hoạt động hàng ngày.

Như vậy, nếu như CSDL tác nghiệp được ví như cái tủ sách cá nhân, nơi người ta thường xuyên tra cứu, cập nhật, hiệu đính, ghi chú vào lề, thêm mới hoặc chuyển sách đi. Thì Kho dữ liệu lại được so sánh với thư viện quốc gia, nơi các tài liệu kinh điển được đưa đến liên tục để lưu trữ và tham khảo, không ai sửa chữa hoặc chuyển chúng qua chỗ nào khác cả.

#### ***2.1.1.3. Các chức năng của kho dữ liệu***

***Khả năng cân bằng (scalable) :*** Kho dữ liệu có khả năng lưu trữ và quản lý một số lượng khổng lồ các giao dịch và dữ liệu tổng hợp. Kho dữ liệu có thể lớn lên theo thời gian, có thêm dữ liệu mới hoặc giữ lại dữ liệu lịch sử cho một thời gian dài mà không bị quá tải, vẫn đảm bảo tính ổn định hoạt động

***Khả năng quản trị (manageable):*** Kho dữ liệu có khả năng quản trị việc tạo quyết định

***Khả năng sẵn sàng (available) :*** Kho dữ liệu luôn sẵn sàng bất kỳ khi nào cần

*Khả năng mở rộng (extensible)*: Thực hiện 1 cách dễ dàng khi thêm các loại dữ liệu mới, dữ liệu kết hợp và dữ liệu tổng hợp và kho dữ liệu

*Khả năng mềm dẻo (Flexible)*: hỗ trợ nhiều cách truy nhập dữ liệu chi tiết, dữ liệu tổng hợp theo phân tích nhiều chiều

*Khả năng tích hợp (Integrated)*: kho dữ liệu phải có khả năng tích hợp đầy đủ với các hệ thống sẵn có, môi trường tác nghiệp. Do vậy, có thể tải dữ liệu từ nhiều nguồn

*Khả năng truy cập sử dụng được (Accessible)*: Kho dữ liệu phải có khả năng truy cập sử dụng được, từ các công cụ mềm dẻo tới phạm vi rộng rãi người sử dụng

*Khả năng tin cậy (Reliable)*: dữ liệu tải từ nhiều nguồn khác nhau phải được hợp nhất, chuẩn hóa bảo đảm tính toàn vẹn dữ liệu và hợp lệ tại một thời điểm bất kỳ

#### 2.1.1.4. Các loại dữ liệu trong Datawarehouse

**Dữ liệu nghiệp vụ**: Dữ liệu nghiệp vụ (Business data – BD) là dữ liệu dùng để vận hành và quản lý một doanh nghiệp hoặc một tổ chức. Nó phản ánh những hoạt động của doanh nghiệp và những đối tượng trong thế giới thực như là khách hàng, địa điểm, sản phẩm v.v.. Nó được tạo ra và sử dụng bởi các hệ thống xử lý giao tác cũng như các hệ thống hỗ trợ quyết định.

**Siêu dữ liệu (Metadata)**: Metadata là dữ liệu về dữ liệu được sử dụng trong DW (hay gọi là siêu dữ liệu) trả lời các câu hỏi ai, cái gì, khi nào, tại sao, như thế nào về dữ liệu. Nó được sử dụng cho việc xây dựng, duy trì, quản lý và sử dụng DW. Metadata là một trong những phương diện quan trọng nhất của DW.

#### 2.1.2. Ứng dụng của kho dữ liệu

**Hàng không**: Xây dựng hệ thống quản lý các công việc hàng không như quản lý tuyến bay, lịch bay, phân công phi hành đoàn, quản lý các chương trình ưu đãi dành cho khách hàng...

**Ngân hàng**: quản lý dòng tiền, quản lý các quỹ đầu tư, cho vay, thời hạn thanh toán...

**Y – Dược**: quản lý thông tin bệnh nhân, tình trạng bệnh án, phác đồ điều trị, đơn thuốc được kê cho họ, thời gian điều trị bệnh...

**Bảo hiểm**: phân tích xu hướng của khách hàng và để theo dõi sự thay đổi của thị trường một cách nhanh chóng

**Bán lẻ**: quản lý mặt hàng nhập vào – bán ra, quản lý chuỗi phân phối, xác định mô hình mua hàng của khách hàng, các chương trình khuyến mãi cũng như các chính sách giá

**Giáo dục**: quản lý thông tin học sinh – giáo viên – công nhân viên của trường, quản lý quá trình học tập, giáo án, bài giảng, kết quả học tập của học sinh... Là nền tảng để xây dựng các phần mềm dạy học hiện nay.

**Quản lý quan hệ khách hàng:** tổng hợp, xử lý, phân tích và giúp dự đoán insight của khách hàng và xu hướng thị trường trong tương lai

**Thương mại điện tử:** là lĩnh vực có sự ứng dụng rõ rệt nhất của kho dữ liệu thông qua việc quản lý thông tin hàng hóa, người bán, người mua, tình trạng đơn hàng, các chương trình khuyến mãi...

### 2.1.3. Phân biệt Kho dữ liệu và Cơ sở dữ liệu truyền thống

Bảng 2.1. Bảng so sánh Kho dữ liệu và Cơ sở dữ liệu truyền thống

	Kho dữ liệu	Cơ sở dữ liệu truyền thống
Giống nhau	1. Cả hai cùng quy trình xây dựng và chúng có thể được triển khai theo giản đồ hình sao hay bông tuyết để thiết kế quy trình mô hình 2. Cùng chung tài liệu nghiệp vụ hay dữ liệu từ nguồn hỗn tạp 3. Người sử dụng có thể xử lý phân tích trực tuyến thông qua kết quả được thiết lập	
Khác nhau	Có ý định đạt được tri thức nghiệp vụ định hướng văn bản (text-oriented)	Có ý định đạt được tri thức nghiệp vụ định hướng số (numeric-oriented)
	Tài nguyên được thu thập từ báo cáo điều tra thị trường, báo cáo trạng thái dự án, lời phàn nàn của khách hàng, thư điện tử, quảng cáo cạnh tranh	Tài nguyên được thu thập từ CSDL bên trong các điểm bán hàng, hệ thống ERP, hệ thống thanh toán, hay hệ thống quản lý tài chính
	Nó lọc những tài liệu không cần thiết và định hướng giúp người sử dụng trả lời câu hỏi vì sao	Nó tích hợp dữ liệu theo một vài chiều và có khuynh hướng giúp người sử dụng trả lời một số câu hỏi ai, cái gì, khi nào, ở đâu và người nào / cái nào
	Nâng cao chất lượng với công nghệ khai phá văn bản cho việc tóm tắt tài liệu hay phân loại tài liệu	Nâng cao chất lượng với công nghệ khai phá dữ liệu cho việc tóm tắt, phân lớp, phân cụm dữ liệu đã được định dạng hay tìm sự kết hợp
	Nguồn dữ liệu nên được tích hợp trong file hệ thống hay CSDL XML	Nguồn dữ liệu có thể được tích hợp trong CSDL quan hệ

### 2.1.4. Các mô hình biểu diễn kho dữ liệu

#### 2.1.4.1. Dữ liệu khối đa chiều

Khối dữ liệu đa chiều hay còn gọi là *Data Cube*, là một cấu trúc dữ liệu được sử dụng trong việc lưu trữ và phân tích dữ liệu. Khối dữ liệu đa chiều bao gồm nhiều chiều dữ liệu, mỗi chiều biểu diễn một thuộc tính của dữ liệu, và cho phép người dùng truy cập và phân tích dữ liệu theo nhiều khía cạnh khác nhau.

Các chiều dữ liệu trong khối dữ liệu đa chiều thường bao gồm thời gian, địa điểm, sản phẩm, khách hàng và các chỉ số khác. Ví dụ, một khối dữ liệu đa chiều về doanh số bán hàng có thể bao gồm các chiều là thời gian, sản phẩm và khu vực.

Cấu trúc khối dữ liệu đa chiều được biểu diễn bằng một bảng có nhiều cột và hàng, trong đó mỗi hàng biểu diễn một bộ dữ liệu với các giá trị tương ứng trong các chiều khác nhau. Việc phân tích dữ liệu trên khối dữ liệu đa chiều thường được thực hiện bằng các phép lọc, nhóm, tổng hợp và tính toán trên các giá trị trong các chiều khác nhau.

Các ứng dụng của khối dữ liệu đa chiều bao gồm phân tích dữ liệu thời gian, phân tích kinh doanh, dự báo và quản lý chuỗi cung ứng. Khối dữ liệu đa chiều cung cấp cho người dùng khả năng phân tích dữ liệu một cách nhanh chóng và dễ dàng, giúp họ đưa ra các quyết định kinh doanh thông minh và đạt được hiệu quả cao hơn trong hoạt động của mình.

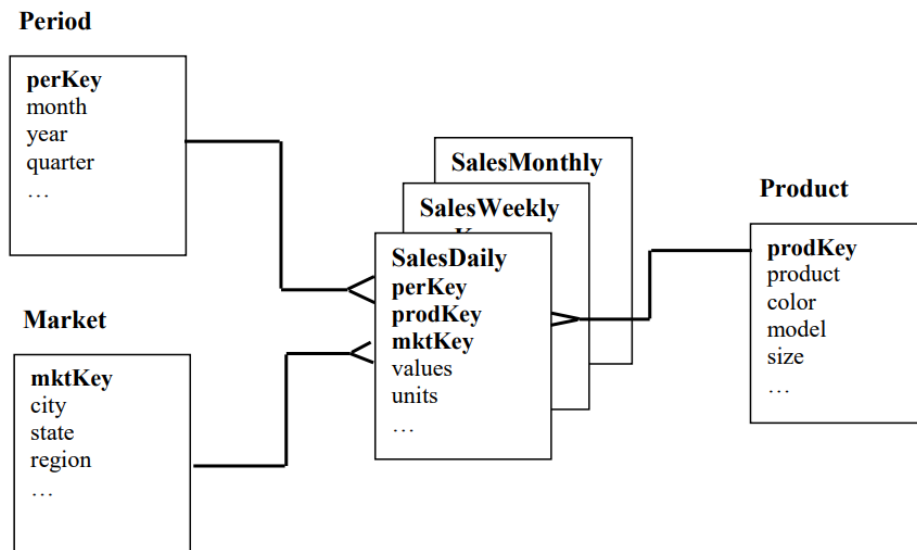
#### *2.1.4.2. Mô hình kho dữ liệu*

##### **Lược đồ hình sao (Star Schema):**

Sơ đồ hình sao được đưa ra lần đầu tiên bởi Dr. Ralph Kimball nhờ là một lựa chọn thiết kế cơ sở dữ liệu cho DW. Nó được gọi là sơ đồ hình sao bởi vì các sự kiện nằm ở trung tâm của mô hình và được bao quanh bởi các phạm vi liên quan, rất giống với các điểm của một ngôi sao. Sơ đồ hình sao cho phép một hệ thống đối tượng có thể kết nối với nhiều đối tượng khác. Mô hình này thể hiện cách nhìn của người sử dụng về nhiều vấn đề trong tác nghiệp. Trong sơ đồ hình sao, dữ liệu được xác định và phân loại theo 2 kiểu:

Các sự kiện được tổ chức thành bảng Fact. Bảng này chứa các thông tin cơ sở ở mức giao tác ở trong nghiệp vụ mà các ứng dụng cần thiết. Ví dụ, khi phân tích dữ liệu kinh doanh thì cần những dữ liệu về những mặt hàng đã bán được trong các giao dịch bán hàng về số lượng, chủng loại, giá thành, v.v.. Những dữ liệu này đều được lưu ở bảng Fact của kho dữ liệu. Tuy nhiên, trước khi các dữ liệu này được đưa vào kho dữ liệu thì cần phải chọn một trường dữ liệu nào đó thường sử dụng trong các chiều phân tích để tham chiếu (xem như khoá ngoại trong các quan hệ liên kết) và sau đó đưa vào bảng các chiều. Các sự kiện là các đại lượng số của công việc. Các bảng Fact thường rất lớn, chứa hàng triệu dòng mà phần lớn là số.

Phạm vi hay các chiều của dữ liệu, được tổ chức thành các bảng Dimension. Bảng Dimension, ngược lại, thường tương đối nhỏ so với các bảng Fact, chứa các thông tin mô tả. Đó là các bộ lọc hoặc các ràng buộc của những sự kiện ở bảng Fact. Bảng Dimension chứa các dữ liệu cần thiết cho việc thực hiện các giao tác nghiệp vụ theo một chiều, hay phạm vi nào đó. Ví dụ, trong ứng dụng phân tích kinh doanh, bảng Dimension bao gồm: thời gian, vùng bán hàng, loại sản phẩm, v.v.



Hình 2.1. Ví dụ về lược đồ hình sao

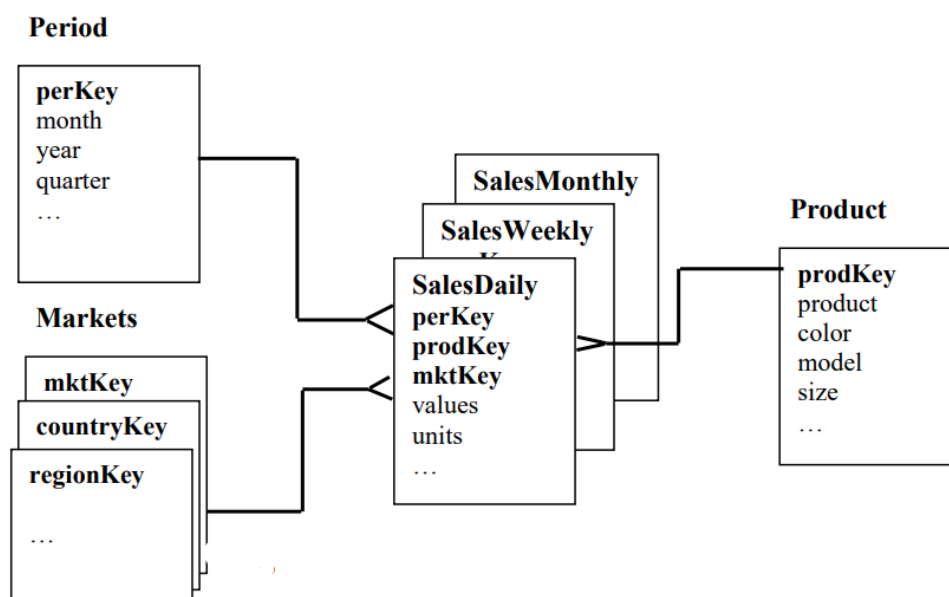
#### *Ưu điểm của lược đồ hình sao*

- Hỗ trợ rất đa dạng các câu truy vấn và xử lý khá hiệu quả những câu truy vấn đó.
- Phù hợp với thói quen của người sử dụng nhận và sử dụng dữ liệu, nên dữ liệu được hiểu trực quan hơn.
- Sơ đồ này rất trực quan, dễ sử dụng, thể hiện khung nhìn đa chiều của dữ liệu dùng ngữ nghĩa của cơ sở dữ liệu quan hệ. Khóa của bảng sự kiện được tạo bởi những khóa của các bảng chứa thông tin theo từng phạm vi (Dimension Table). Tất cả các khóa đều được xác định với cùng một chuẩn đặt tên.

#### **Lược đồ bông tuyết (Snowflake)**

Lược đồ bông tuyết là một sự mở rộng của lược đồ hình sao tại đó mỗi cánh sao không phải là một bảng Dimension mà là nhiều bảng. Trong dạng sơ đồ này, mỗi bảng theo chiều của sơ đồ hình sao được chuẩn hóa hơn. Lược đồ bông tuyết cải thiện năng suất truy vấn, tối thiểu không gian đĩa cần thiết để lưu trữ dữ liệu và cải thiện năng suất nhờ việc chỉ phải kết hợp những bảng có kích thước nhỏ hơn thay vì phải kết hợp những bảng có kích thước lớn lại không chuẩn hóa. Nó cũng làm tăng tính linh hoạt của các ứng dụng bởi sự chuẩn hóa và ít mang bản chất theo chiều hơn. Nó

làm tăng số lượng các bảng và làm tăng tính phức tạp của một vài truy vấn cần có sự tham chiếu tới nhiều bảng.



Hình 2.2. Ví dụ về lược đồ bông tuyết

*Nhược điểm của lược đồ bông tuyết:*

- Lược đồ ít trực quan hơn và phức tạp
- Khả năng duyệt qua/truy vấn các nội dung khó
- Hiệu suất truy vấn thấp.

### 2.1.5. Các giai đoạn thiết kế kho dữ liệu

Có nhiều phương pháp thiết kế kho dữ liệu, mặc dù khác nhau nhưng nhìn chung các phương pháp đều chứa một số công việc cơ bản cho việc thiết kế kho dữ liệu gồm các giai đoạn sau:

*Xác định mô hình nghiệp vụ (Defining the business model):* Phân tích các định hướng chiến lược để rút ra các quá trình nghiệp vụ được thực hiện trong kho dữ liệu. Các yêu cầu nghiệp vụ phải được phân tích, xác định và tài liệu hóa đơn vị đo và chiều nghiệp vụ (business measures và business dimensions) cho mỗi quá trình nghiệp vụ

*Xác định mô hình logic (Defining the logical model):* Dùng kỹ thuật đồ họa để thể hiện các định nghĩa, đặc điểm và mối quan hệ của dữ liệu trong nghiệp vụ, hoặc khái niệm chuyên môn, nhằm mục đích mô tả hệ thống cho người sử dụng cuối. Thông thường phương pháp này dùng chuẩn 3NF

*Xác định mô hình Dimensional (Defining the dimensional model)* Mô hình nghiệp vụ được chuyển thành mô hình Dimensional, lọc bỏ các bảng và thuộc tính của bảng được định nghĩa, liên kết giữa các bảng được tạo dựng và nguồn của kho dữ liệu được xác định.



*Xác định mô hình vật lý (Defining the physical model)* Mô hình Dimensional được chuyển thành mô hình vật lý, bao gồm các khai báo về định nghĩa kiểu dữ liệu của các thuộc tính, kích thước dự tính, chiến lược tạo chỉ mục (index) và lưu trữ dữ liệu

## **2.2. Ngôn ngữ truy vấn dữ liệu đa chiều MDX**

### **2.2.1. Giới thiệu khái quát về ngôn ngữ MDX**

Ngôn ngữ MDX (*MultiDimensional eXpression*) là ngôn ngữ truy vấn cho cơ sở dữ liệu đa chiều, nó tương tự ngôn ngữ SQL cho cơ sở dữ liệu quan hệ. Tuy nhiên, đây là ngôn ngữ tính toán vì thế nó tương tự có cú pháp giống công thức của bảng tính. MDX do Microsoft giới thiệu vào năm 1997 cùng với SQL Server 7.0 nhanh chóng được các nhà phát triển chấp nhận đã trở thành chuẩn cho hệ thống OLAP. Từ năm 2001 khi XMLA ra đời MDX được mở rộng và hỗ trợ truy vấn với MDX (MDXML).

Ngôn ngữ MDX hỗ trợ mạnh cho tính toán trên khối dữ liệu đa chiều, nó hỗ trợ việc truy cập dữ liệu một cách trực quan và dễ dàng. Người dùng có thể trích xuất dữ liệu từ cube với số chiều bất kỳ. MDX có thể trả lời được các câu hỏi phức tạp trong kinh doanh ở thế giới thực như: “trong các khu vực bán được nhiều sản phẩm nhất đâu là dòng sản phẩm bán chạy nhất?”.

MDX được xây dựng cho các hệ thống phân tích dữ liệu chính vì thế nó có thể xử lý được những truy vấn một cách dễ dàng. MDX có thể xử lý dữ liệu một cách mềm dẻo, kết quả trả về của MDX cũng rất linh hoạt. Các phép truy vấn phức tạp trong SQL như pivot được sử dụng một cách dễ dàng trong MDX. MDX không hỗ trợ các ngôn ngữ định nghĩa dữ liệu (DDL) tuy nhiên ngôn ngữ thao tác dữ liệu (DML) của MDX có hỗ trợ thực thi một số thao tác định nghĩa dữ liệu.

### **2.2.2. Một số cú pháp cơ bản của ngôn ngữ MDX**

*Cấu trúc của MDX:* giống như SQL nhưng mở rộng hơn để thao tác với cơ sở dữ liệu đa chiều. Câu truy vấn MDX có cấu trúc như sau :

**[WITH**

**MEMBER ...**

**SET ...]**

**SELECT**

**[<mô tả về chiều thứ nhất>] ,**

**[<mô tả về chiều thứ hai>] ...**

**FROM**

**[<mô tả về khối dữ liệu>]**

**WHERE**

**[<điều kiện cắt lớp>]**

Trong đó:

**WITH:** dùng để định nghĩa thêm chiều mới từ các chiều đã tồn tại.

**Mệnh đề SELECT:** dùng để xác định các chiều của tập hợp kết quả. Mệnh đề này còn hỗ trợ các cú pháp tùy chọn khác, như là từ khóa WITH và việc sử dụng các hàm MDX để xây dựng các thành viên bằng việc tính toán để thêm vào 1 trục hoặc 1 chiều cắt.

**Mệnh đề FROM:** xác định nguồn dữ liệu (cube) dùng để lấy dữ liệu.

**Mệnh đề WHERE:** dùng để xác định chiều cắt dữ liệu, nhằm lọc dữ liệu đầu ra.

*Quy ước về cú pháp MDX*

- Hàm
  - + Một số sử dụng ký hiệu hàm theo phong cách truyền thống
    - Ví dụ: hàm Order()
  - + Một số sử dụng định dạng đối tượng
    - Ví dụ: BoMon.Members - Members là hàm
- Thứ tự trục
  - + COLUMNS, ROWS, PAGES, CHAPTERS, SECTIONS
  - + Mỗi trục được đánh số thứ tự bắt đầu từ 0
    - Ví dụ: { [Dim\_ThoiGian].Members } on Axis(0)
- Dấu phẩy – xây dựng một tập hợp bằng cách liệt kê các các bộ
  - + Ví dụ: {[Nam].[2001], [Nam].[2002], ...}
- Dấu hai chấm – xây dựng một tập hợp bằng cách chỉ định một phạm vi
  - + Ví dụ: { [Nam].[2001] : [Nam].[2005] }
- Members – trả về tập tất cả các thành viên
  - + Ví dụ: [BoMon].Members

## 2.3 Tổng quan về phân tích và xử lý trực tuyến OLAP

### 2.3.1. OLAP là gì?

OLAP – Online Analysis Process (Xử lý phân tích trực tuyến): Là các công cụ phần mềm cho phép các nhà quản lý cấp cao hiểu biết sâu sắc về dữ liệu thông qua truy cập tương tác nhanh, nhất quán, trong nhiều chế độ xem thông tin có thể được chuyển đổi từ dữ liệu thô sang phản ánh chiều hướng thực sự của doanh nghiệp theo cách hiểu của người dùng.

### 2.3.2. Quy trình phân tích và xử lý trực tuyến bằng OLAP trên kho dữ liệu

*Chuẩn bị dữ liệu:* Dữ liệu từ các nguồn khác nhau được thu thập và tích hợp vào kho dữ liệu OLAP. Các công cụ ETL (Extract, Transform, Load) được sử dụng để xử lý dữ liệu trước khi đưa vào kho dữ liệu.

*Thiết kế và tạo Cube:* Cube là một cấu trúc dữ liệu đa chiều được sử dụng để lưu trữ dữ liệu trong kho dữ liệu OLAP. Cube được thiết kế dựa trên các yêu cầu phân tích dữ liệu cụ thể.

*Phân tích và thiết kế các Dimension:* Dimension là các thuộc tính của dữ liệu, được sử dụng để phân loại dữ liệu trong Cube. Các Dimension được thiết kế dựa trên các yêu cầu phân tích dữ liệu cụ thể.

*Xây dựng các Measures:* Measures là các thông số thống kê hoặc các chỉ số được sử dụng để phân tích dữ liệu trong Cube. Các Measures được thiết kế dựa trên các yêu cầu phân tích dữ liệu cụ thể.

*Triển khai và quản lý Cube:* Cube được triển khai và quản lý trên kho dữ liệu OLAP. Quản lý Cube bao gồm cập nhật dữ liệu, tối ưu hóa hiệu suất, và quản lý bảo mật.

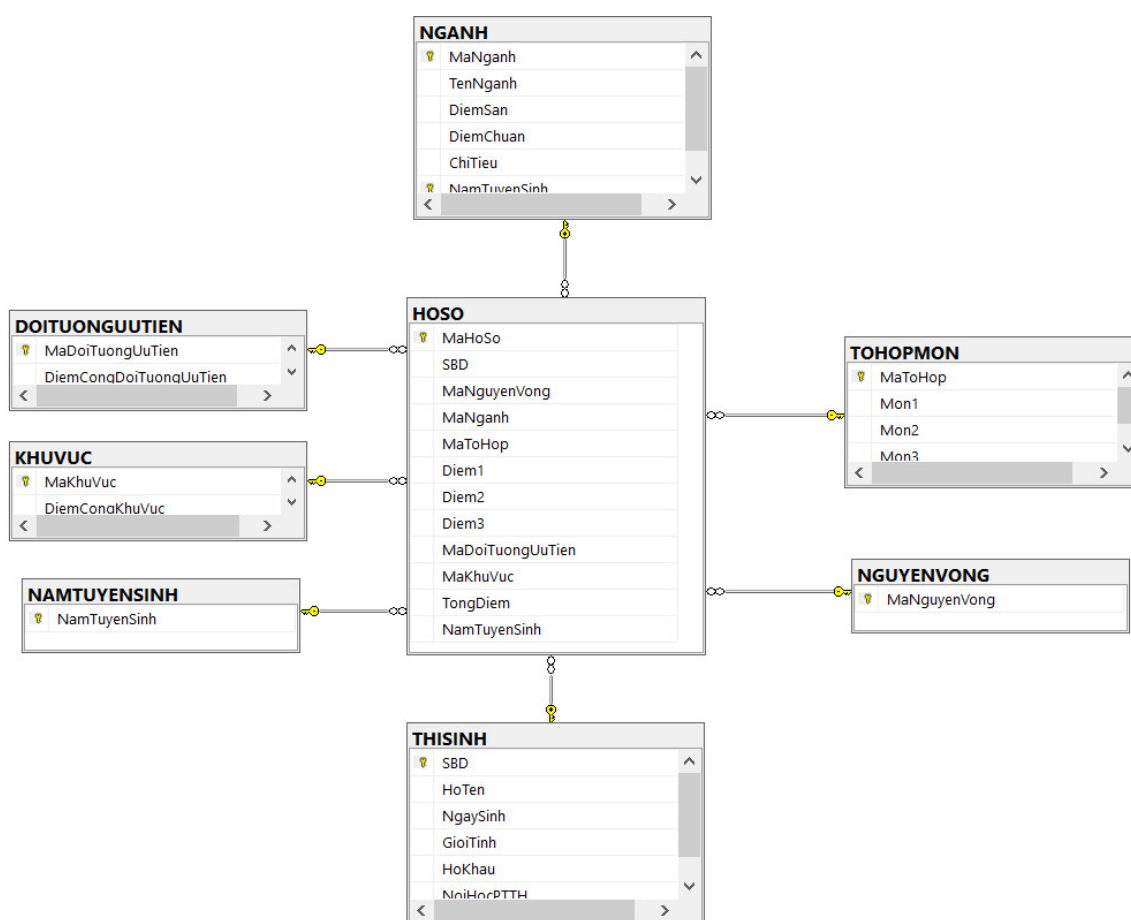
*Phân tích và truy vấn dữ liệu:* Các người dùng có thể truy vấn dữ liệu trên kho dữ liệu OLAP bằng các công cụ phân tích và truy vấn dữ liệu. Các công cụ này cho phép người dùng truy vấn dữ liệu theo nhiều góc độ khác nhau, sử dụng các kỹ thuật như drill down, roll up, slice and dice, và pivot.

*Phân tích dữ liệu và đưa ra quyết định:* Dữ liệu được phân tích và đưa ra quyết định kinh doanh dựa trên các thông tin được trích xuất từ kho dữ liệu OLAP. Quyết định kinh doanh được đưa ra nhanh chóng và chính xác hơn nhờ vào khả năng phân tích và truy vấn dữ liệu của OLAP.

## Chương 3: Mô phỏng quá trình xây dựng và truy vấn MDX trên kho dữ liệu đa chiều

### 3.1. Xây dựng và cài đặt kho dữ liệu tuyển sinh

Trên cơ sở bài toán phân tích dữ liệu tuyển sinh nhằm trợ giúp thí sinh đưa ra quyết định lựa chọn và thay đổi nguyện vọng đúng đắn, đồng thời, trợ giúp cho Ban tuyển sinh các trường có thể thuận tiện hơn trong việc quyết định đưa ra điểm chuẩn phù hợp cho các ngành khác nhau trong trường đã được đề cập ở trên. Qua tìm hiểu về hệ thống tuyển sinh của trường Đại học Kinh tế Kỹ thuật Công nghiệp và quy trình tuyển sinh theo nguyện vọng của Bộ Giáo dục có thể nhận thấy một số nhu cầu xuất hiện trên một số đối tượng người dùng như sau:



Hình 3.1. Lược đồ hình sao của hệ thống Tuyển sinh

Đối tượng là thí sinh đăng ký nguyện vọng tại trường: Sau khi các trường đại học nói chung và trường Đại học Kinh tế Kỹ thuật Công nghiệp nói riêng công bố điểm sàn đại học. Thí sinh dựa trên điểm số ba môn tổ hợp của mình để thay đổi nguyện vọng sao cho khả năng trúng tuyển cao nhất. Để hỗ trợ đối tượng này, hệ thống rất cần cung cấp các thông tin về dữ liệu điểm sàn và điểm chuẩn của các năm trước để thí sinh tham khảo làm cơ sở cho việc thay đổi nguyện vọng hợp lý. Mặt khác, dữ liệu hỗ trợ thí sinh nói trên liên quan đến các thông tin chính như: Các ngành

đào tạo của trường, tổ hợp môn xét tuyển theo các ngành khác nhau, điểm sàn và điểm chuẩn thay đổi từng năm và thay đổi theo từng khu vực và đối tượng ưu tiên.

Đối tượng là Ban tuyển sinh của Nhà trường cần dựa vào dữ liệu điểm thi của các thí sinh đăng ký các nguyện vọng 1, 2, 3 tại trường, các thí sinh đã đăng ký nguyện vọng 1 và 2 ở trường khác không đạt nhưng có đăng ký nguyện vọng 3 tại Nhà trường. Đồng thời, căn cứ vào chỉ tiêu tuyển sinh của mỗi ngành, từ đó phân tích để đưa ra điểm chuẩn phù hợp, đảm bảo không vượt quá nhiều hay quá ít so với chỉ tiêu đề ra.

Trên cơ sở phân tích trên, chúng em nhận thấy, cần xây dựng Kho dữ liệu Tuyển sinh có dạng lược đồ hình sao với các thông tin cụ thể ở các bảng Fact HỒ SƠ gồm các thông tin:

*HOSO (MaHoSo, SBD, MaNguyenVong, MaNganh, MaToHop, Diem1, Diem2, Diem3, MaDoiTuongUuTien, MaKhuVuc, TongDiem, NamTuyenSinh)*

và bảng Dimension: *Đối tượng ưu tiên, Khu vực, Ngành, Tổ hợp môn, Nguyện vọng, Năm tuyển sinh, Thí sinh*. Đồng thời đưa các khóa của bảng Dimension vào làm khóa ngoại của bảng Fact. [Hình 3.1]

### 3.2. Mô phỏng quá trình truy vấn MDX trên kho dữ liệu tuyển sinh.

Một số truy vấn MDX trên kho dữ liệu tuyển sinh dựa vào nhu cầu truy vấn thực tế:

*Câu truy vấn 1: Thống kê số lượng hồ sơ đăng kí ngành CNTT theo nguyện vọng ở tất cả các khu vực*

```
SELECT nonempty([NGUYENVONG].[Ma Nguyen Vong].[Ma Nguyen Vong]) on 0,
[NGANH].[Ten Nganh].&[Công nghệ thông tin]*[KHUVUC].[Ma Khu Vuc].[Ma Khu Vuc] on 1
from [QUAN LI TUYEN SINH]
where [Measures].[HOSO Count];
```

		1	2
Công nghệ thông tin	KV1	19	20
Công nghệ thông tin	KV2	28	28
Công nghệ thông tin	KV2NT	14	9
Công nghệ thông tin	KV3	7	7

Hình 3. 2 Kết quả câu truy vấn 1

### Câu truy vấn 2: Thống kê số lượng hồ sơ theo từng ngành, theo năm

```
SELECT {nonempty([NAMTUYENSINH].[Nam Tuyen Sinh].[Nam Tuyen Sinh]),[NAMTUYENSINH].[Nam Tuyen Sinh].[All]} on 0,
{nonempty([NGANH].[Ten Ngành].[Ten Ngành]),[NGANH].[Ten Ngành].[All]} on 1
from [QUAN LI TUYEN SINH]
where [Measures].[HOSO Count];
```

	2020	2021	2022	All
CN kỹ thuật cơ khí	37	25	38	100
Công nghệ thông tin	36	47	49	132
Công nghệ thực phẩm	23	24	37	84
Tài chính ngân hàng	24	24	36	84
All	120	120	160	400

Hình 3. 3 Kết quả câu truy vấn 2

### Câu truy vấn 3: Thống kê số hồ sơ đăng kí nguyện vọng 1 và 2 theo các tổ hợp môn của tất cả các ngành

```
SELECT nonempty([TOHOPMON].[Ma To Hop].[Ma To Hop]) on 0,
{[NGUYENVONG].[Ma Nguyen Vong].&[1],[NGUYENVONG].[Ma Nguyen Vong].&[2]} on 1
from [QUAN LI TUYEN SINH]
where [Measures].[HOSO Count];
```

	A00	A01	B00	C01	D01
1	65	68	3	1	66
2	53	57	2	5	80

Hình 3. 4 Kết quả câu truy vấn 3

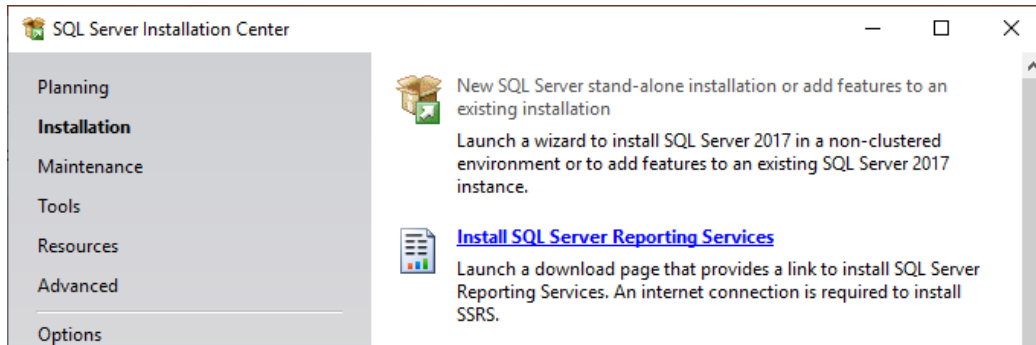
## Chương 4: Triển khai ứng dụng

### 4.1. Triển khai và cài đặt các công cụ hỗ trợ

#### 4.1.1. Cài đặt Microsoft SQL Server 2019 Developer

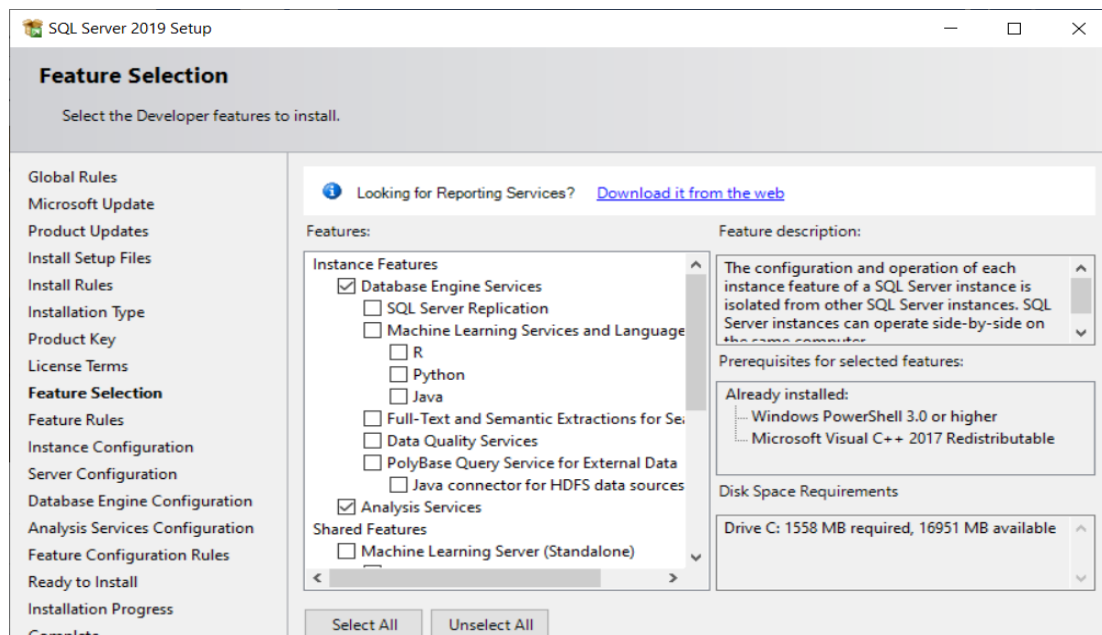
B1: Tiến hành tải SQL Server 2019 trên trang [Microsoft](#) và chọn bản Developer.

B2: Mở File cài đặt, chọn Installation / New SQL Server stand-alone installation or add features to an existing installation.



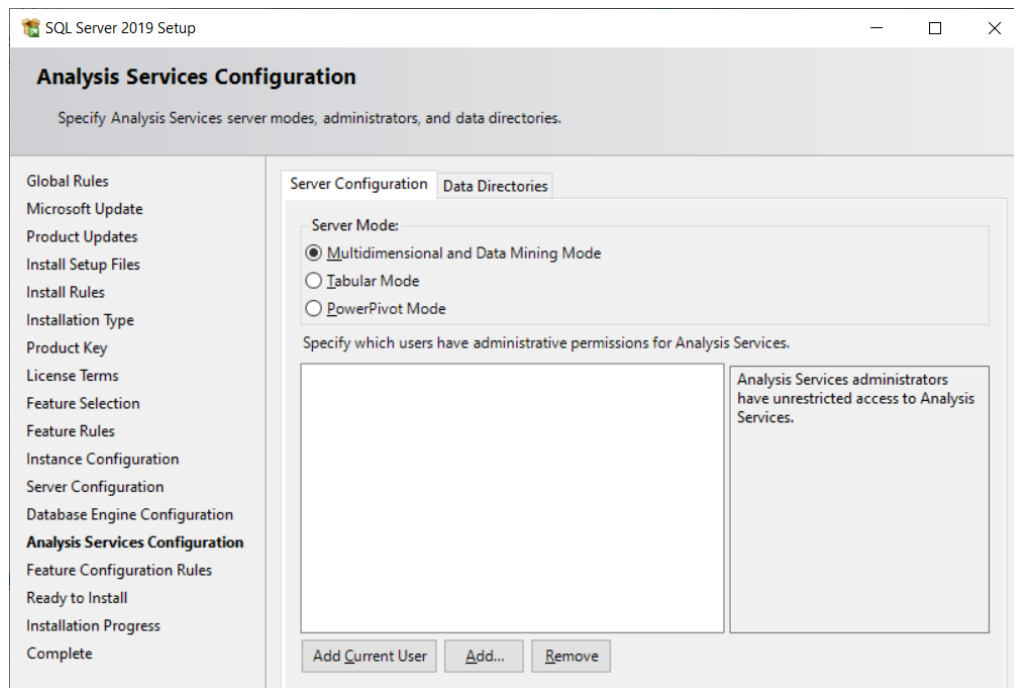
Hình 4. 1. Cài đặt SQL Server

B3. Trong SQL Server 2019 Setup, phần Feature trong **Feature Selection** chọn **Instance Features** như sau:



Hình 4. 2. Chọn tính năng trong SQL Server

B4. Ở phần Analysis Services Configuration/ Sever Configuration tại Sever Mode chọn Multidimensional and Data Mining Mode

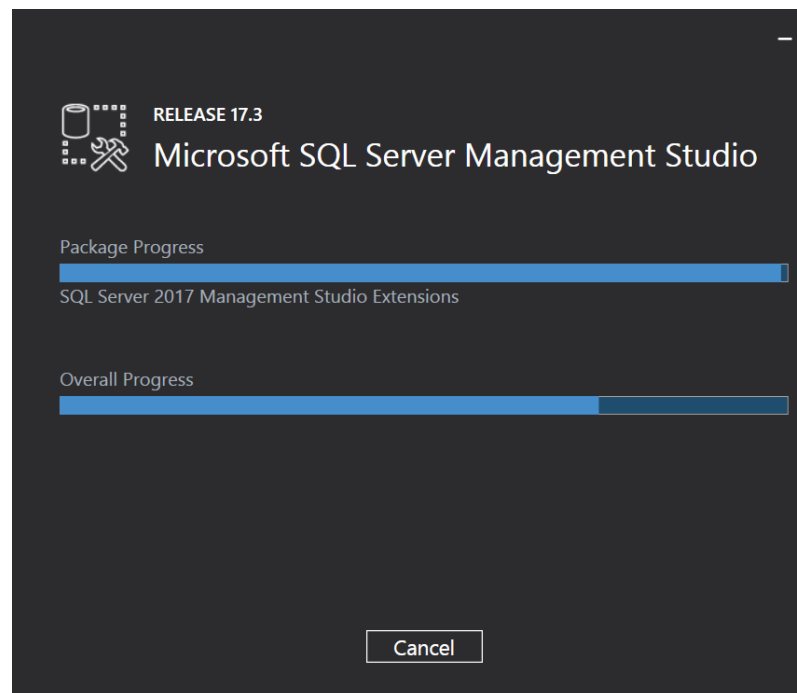


Hình 4. 3. Chọn chế độ cho server của Analysis Services

#### 4.1.2. Cài đặt SQL Server Management Studio

B1. Tiến hành tải SQL Server Management Studio trên trang Microsoft. Chọn Install để cài đặt.

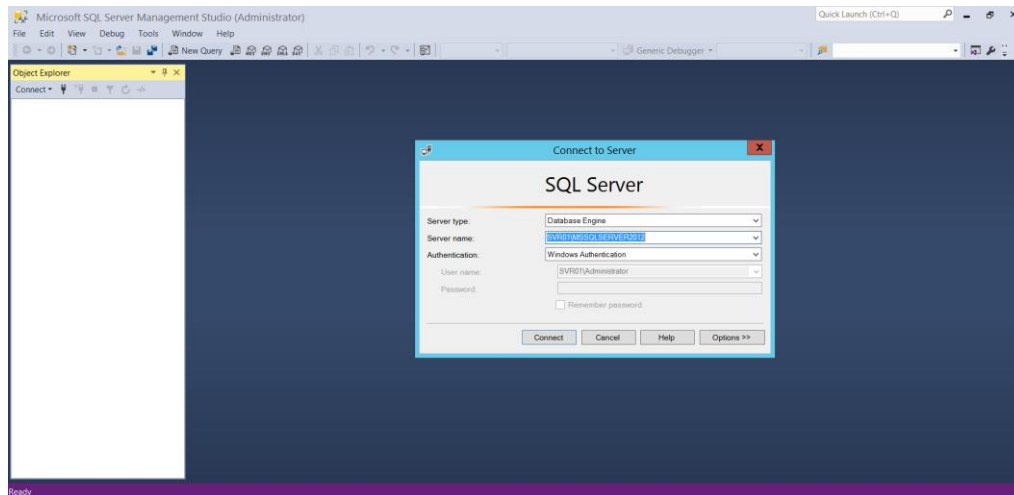
B2. Phần mềm tiến hành cài tự động, đợi cho đến khi phần mềm yêu cầu khởi động lại máy.



Hình 4. 4. Cài SQL Server Management Studio

B3. Kết nối đến SQL Server.

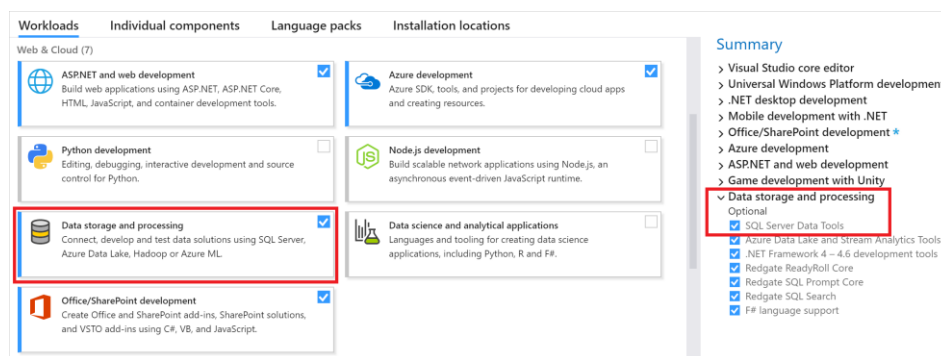




Hình 4. 5. Kết nối SQL Server trong SQL Server Management Studio

#### 4.1.3. Cài đặt Visual Studio 2017 Community

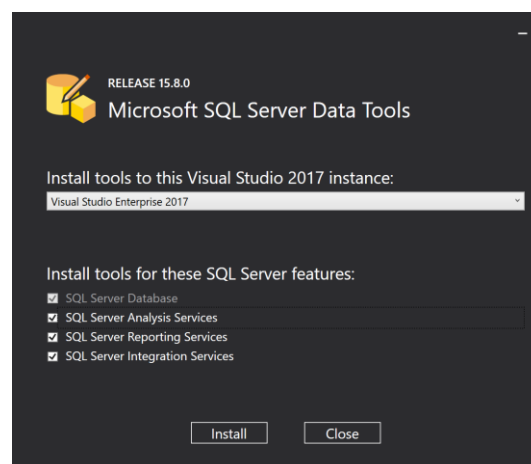
Tiến hành tải bộ cài đặt Visual Studio 2017 trên trang Microsoft và chọn các chức năng như hình.



Hình 4. 6. Bảng chọn chức năng của phần mềm Visual Studio

#### 4.1.4. Cài đặt Microsoft SQL Server Data Tools

Tiến hành tải bộ cài đặt Microsoft SQL Server Data Tools trên trang Microsoft. Cài đặt và chọn các tính năng như hình:

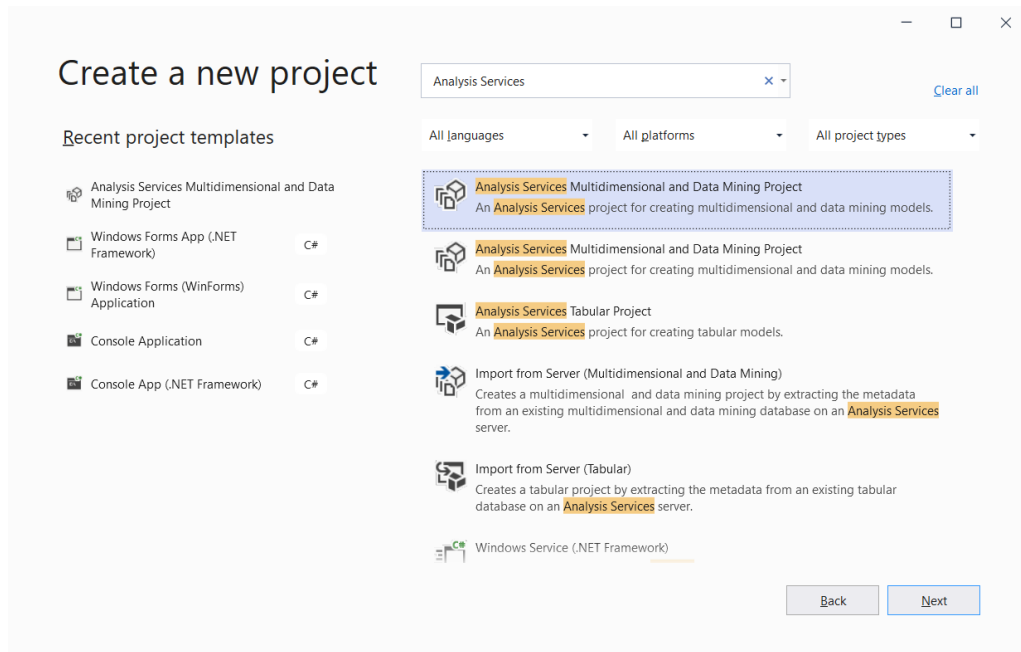


Hình 4. 7. Cài đặt Microsoft SQL Server Data Tools

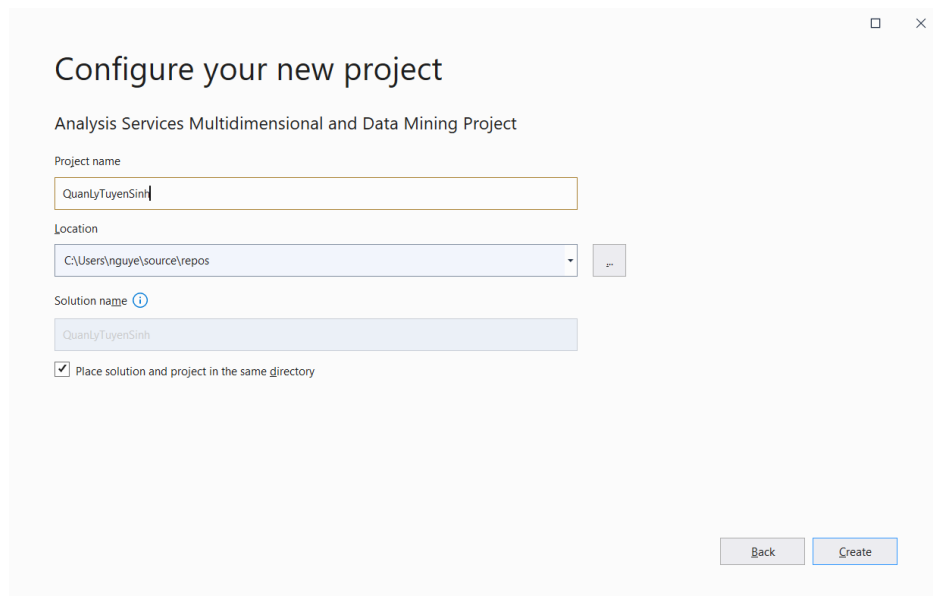
## 4.2. Xây dựng và cài đặt kho dữ liệu Tuyển sinh

B1. Sử dụng SQL Server Tools 18 để tạo Database, bảng và thêm dữ liệu mẫu để khai phá.

B2. Tạo một Project mới trong Visual Studio 2019, chọn **Analysis Services Multidimensional and Data Mining Project**.

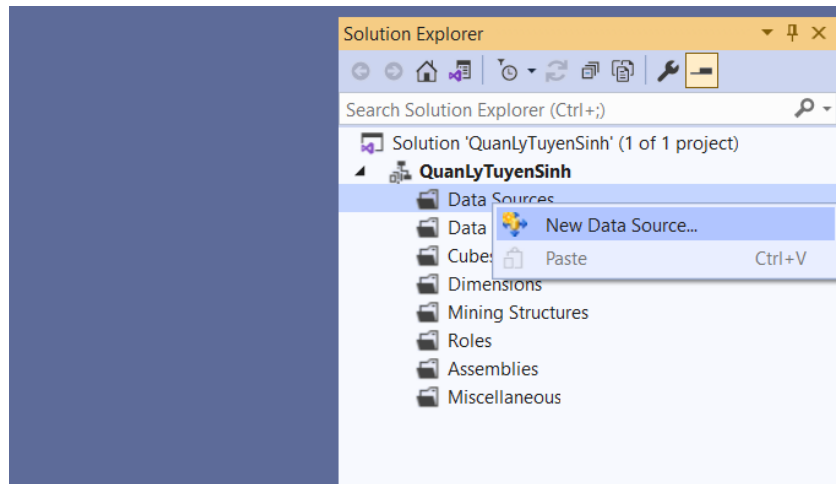


Hình 4. 8. Tạo Project Analysis Services mới trong Visual Studio



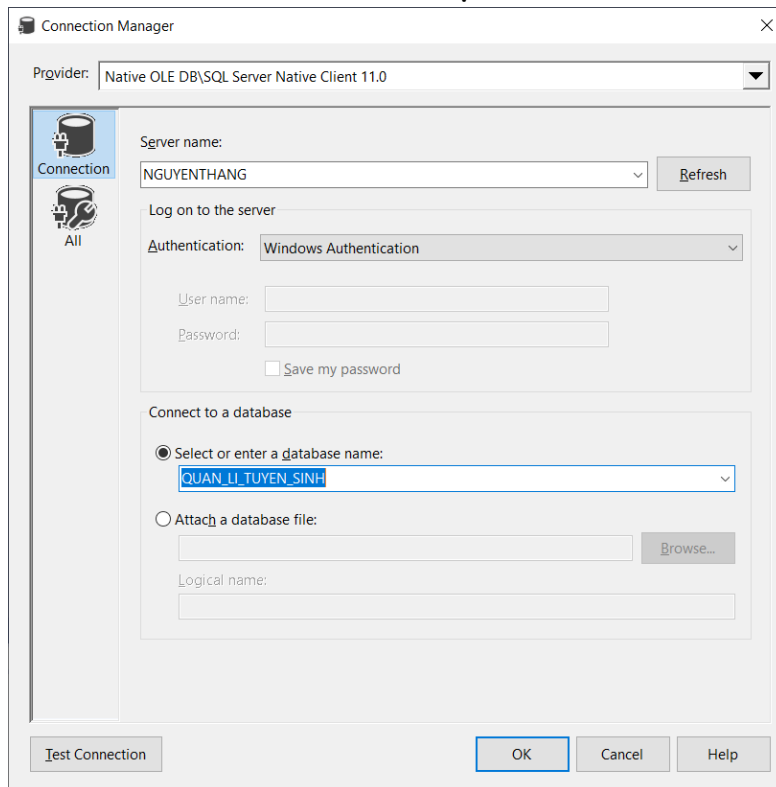
Hình 4. 9. Nhập tên Project để hoàn thành tạo Project Analysis Services

B3. Thêm Data Sources



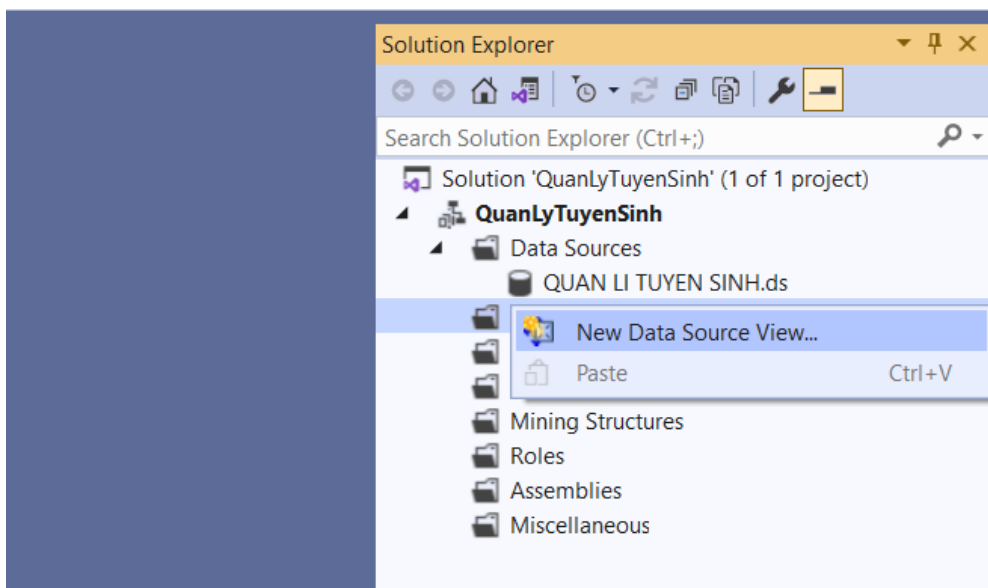
Hình 4. 10. Thêm Data Source trong Analysis Services

B4. Tiến hành kết nối đến Database vừa tạo ở B1.



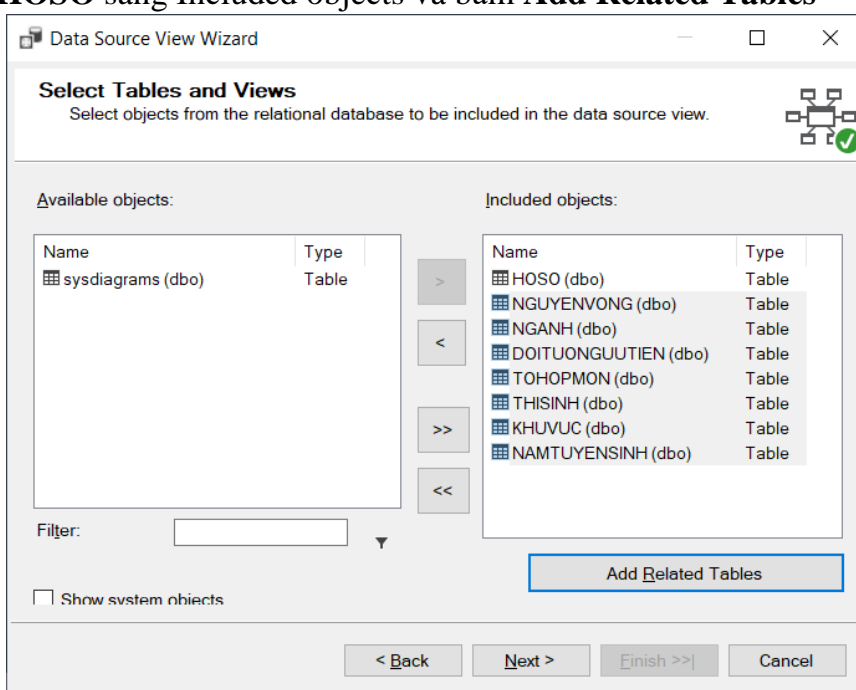
Hình 4. 11. Kết nối Database trong Analysis Services

B5. Thêm **Data Source Views**, chọn bảng đã tạo ở B1 để tiến hành tạo View mới cho khai phá.



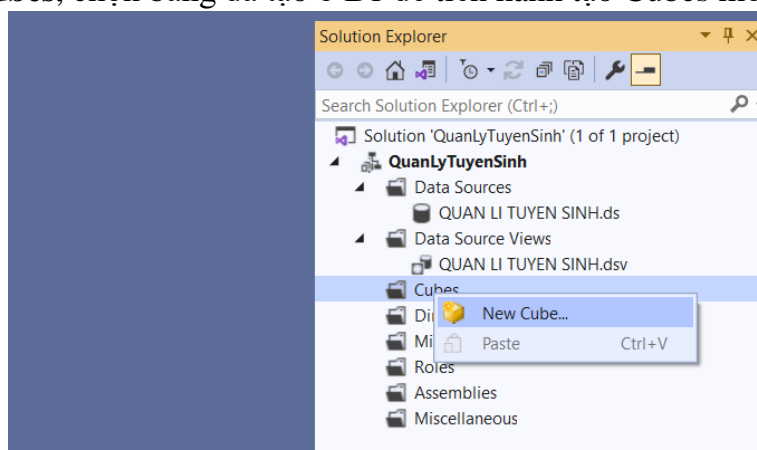
Hình 4. 12. Thêm Data Source View trong Analysis Services

Thêm bảng **HOSO** sang Included objects và bấm **Add Related Tables**



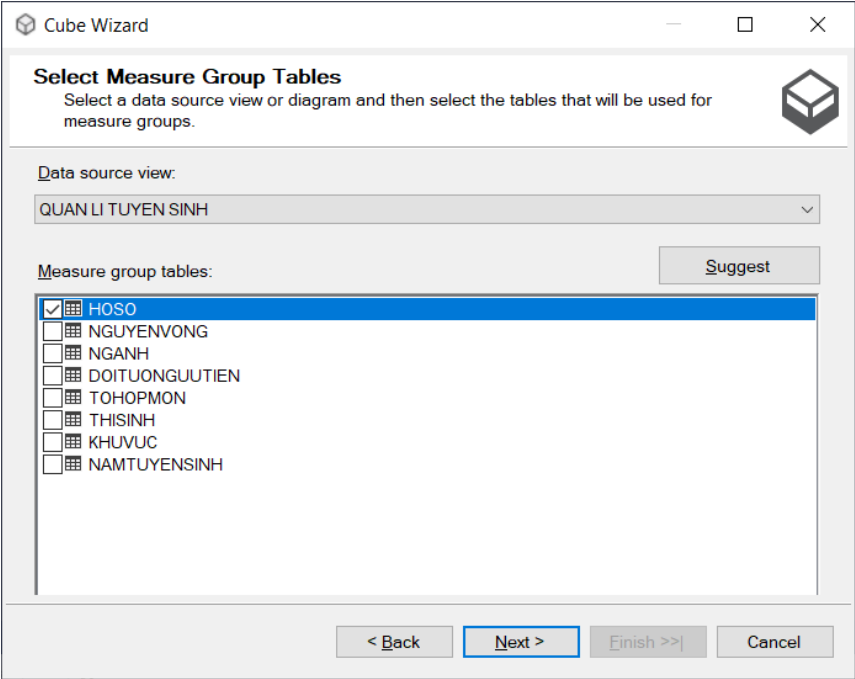
Hình 4. 13. Thêm Include objects trong Analysis Services

B6. Thêm **Cubes**, chọn bảng đã tạo ở B1 để tiến hành tạo Cubes mới cho khai phá.



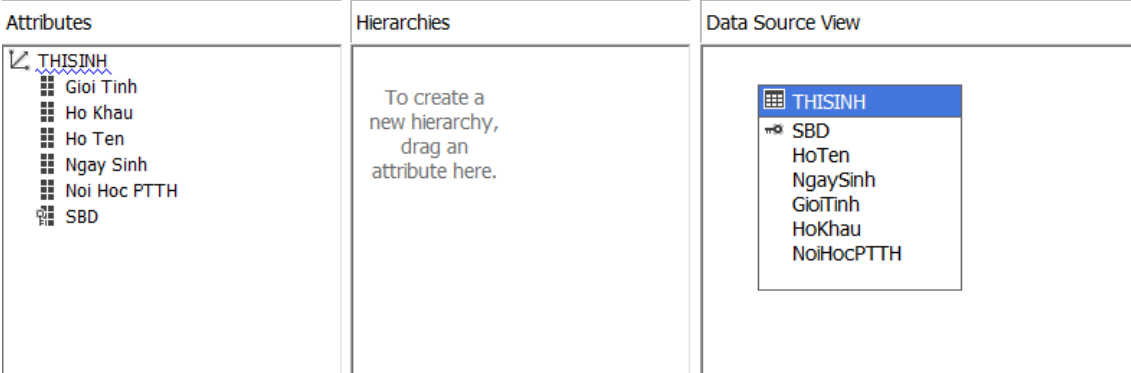
Hình 4. 14. Thêm Cubes trong Analysis Services

Chọn bảng **HOSO** làm **Measure group tables**

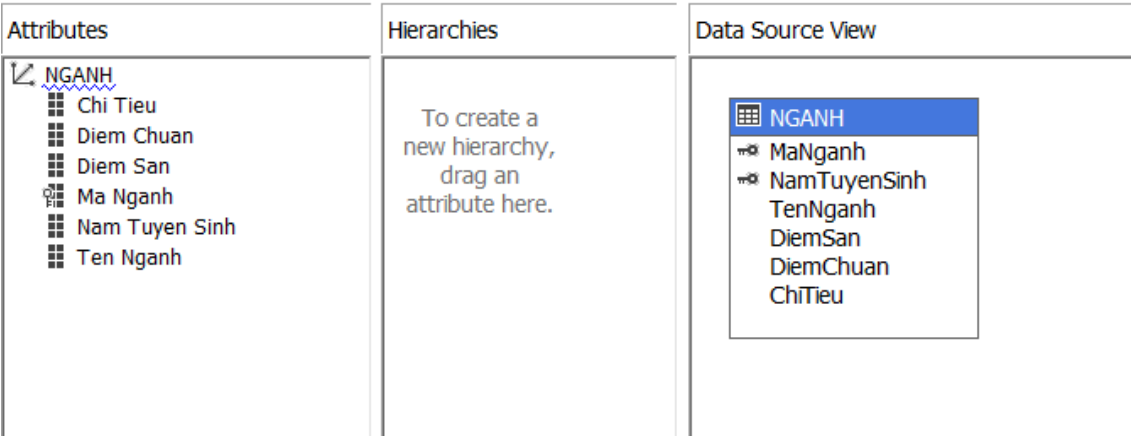


Hình 4. 15. Chọn measure group tables cho cube trong Analysis Service

B7. Thêm các thuộc tính cho các dimension từ Data Source View trong Cubes



Hình 4. 16. Dimension THISINH



Hình 4. 17. Dimension NGANH

Attributes	Hierarchies	Data Source View
<b>DOITUONGUUTIENT</b> Diem Cong Doi Tuong Uu Tien Ma Doi Tuong Uu Tien	To create a new hierarchy, drag an attribute here.	<b>DOITUONGUUTIENT</b> MaDoiTuongUuTien DiemCongDoiTuongUuTien

Hình 4. 18. Dimension DOITUONGUUTIENT

Attributes	Hierarchies	Data Source View
<b>TOHOPMON</b> Ma To Hop Mon1 Mon2 Mon3	To create a new hierarchy, drag an attribute here.	<b>TOHOPMON</b> MaToHop Mon1 Mon2 Mon3

Hình 4. 19. Dimension TOHOPMON

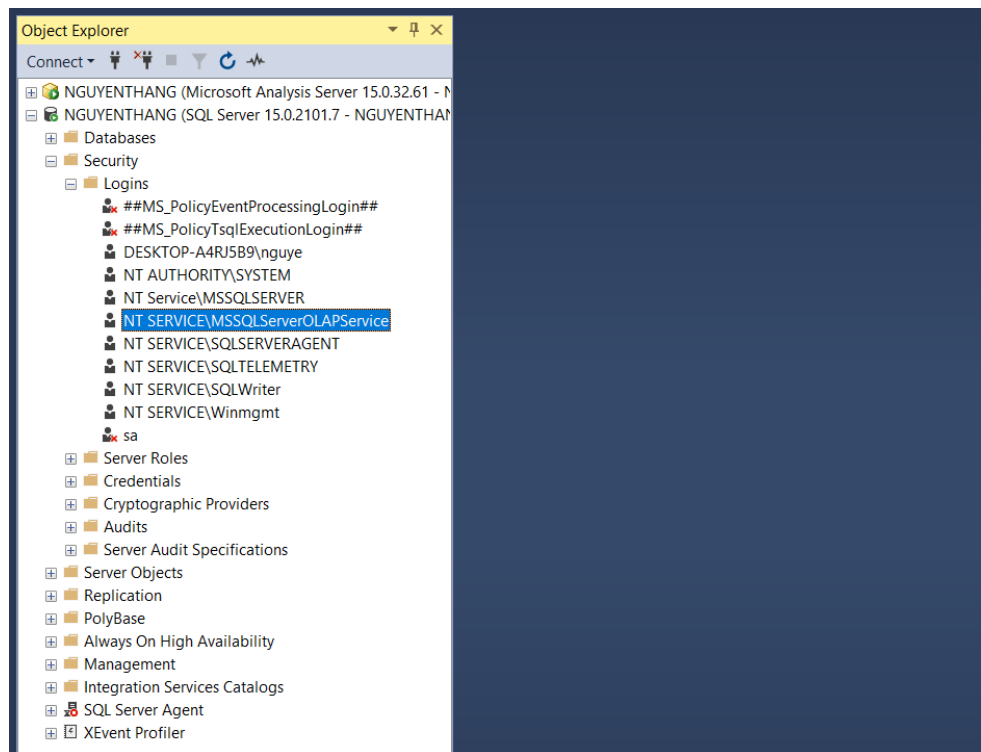
Attributes	Hierarchies	Data Source View
<b>KHUVUC</b> Diem Cong Khu Vuc Ma Khu Vuc	To create a new hierarchy, drag an attribute here.	<b>KHUVUC</b> MaKhuVuc DiemCongKhuVuc

Hình 4. 20. Dimension KHUVUC

#### B8. Tiến hành Process Project

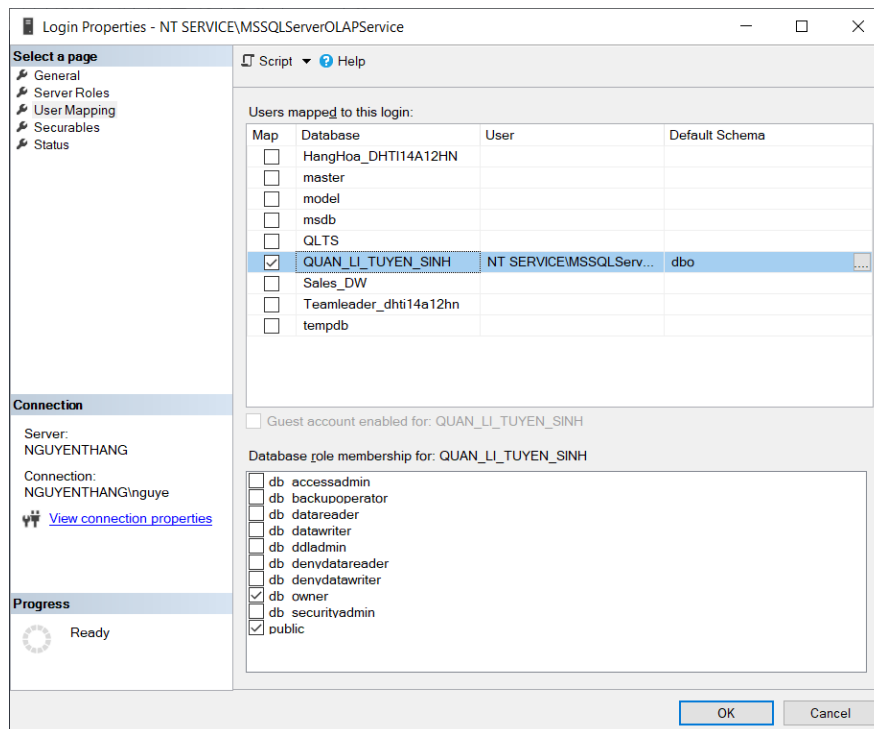
Trước tiên cần cấp quyền quản lý database QUAN\_LI\_TUYEN\_SINH cho login NT SERVICE\MSSQLServerOLAPService (“NT SERVICE “ sẽ khác nhau trên mỗi máy)

Kết nối database engine -> Security -> Logins



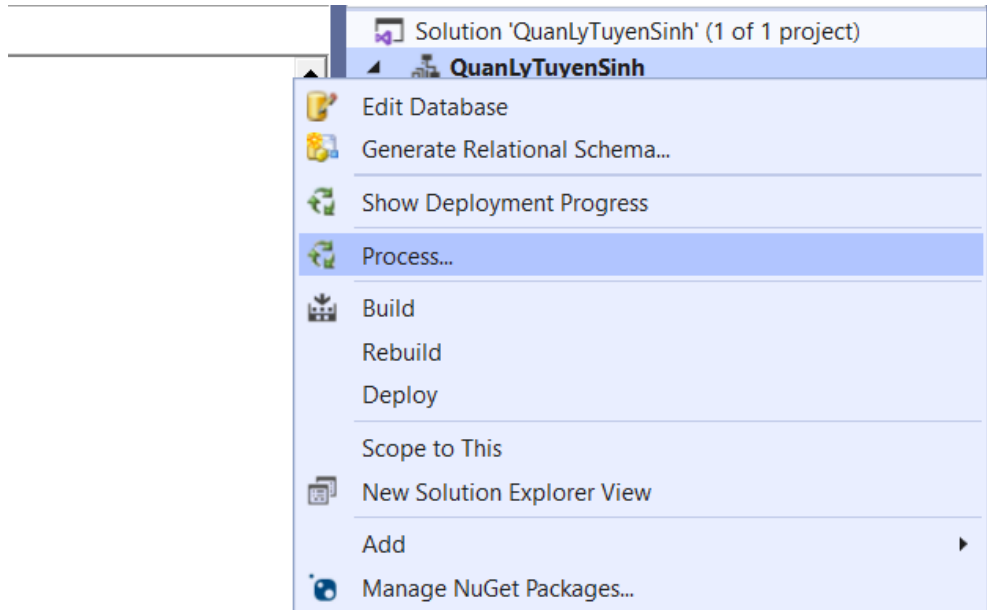
Hình 4.20. Mở login trong SQL server

Mở Properties -> User Mapping -> Tại Users mapped to this login tích chọn QUAN\_LI\_TUYEN\_SINH -> Tại Database role membership for: QUAN\_LI\_TUYEN\_SINH chọn db owner.



Hình 4. 21. Cấp quyền quản lý database cho user

Quay lại Visual Studio 2019, nhấn chuột phải vào Project và chọn Process sau đó nhấn Run để hoàn thành Process



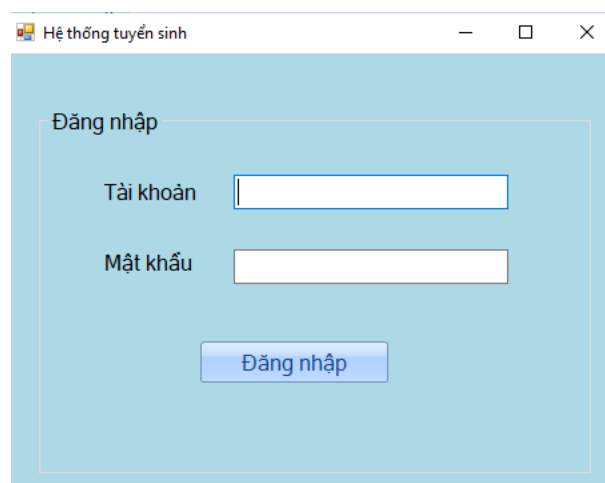
Hình 4. 22.. Tiến hành Process Project

#### 4.3. Thiết kế giao diện chương trình

Mục đích xây dựng chương trình:

Chương trình “Quản lí tuyển sinh” được xây dựng với mục đích ứng dụng ngôn ngữ lập trình C# và ngôn ngữ truy vấn đa chiều MDX để mô phỏng các truy vấn trên kho dữ liệu tuyển sinh. Cung cấp các chức năng hiển thị cơ sở dữ liệu tuyển sinh, giao diện OLAP đơn giản, các giao diện hiển thị kết quả truy vấn đa chiều, không cung cấp các chức năng thêm, sửa, xóa can thiệp vào cơ sở dữ liệu, mà chỉ mang tính thống kê.

**Giao diện đăng nhập:**

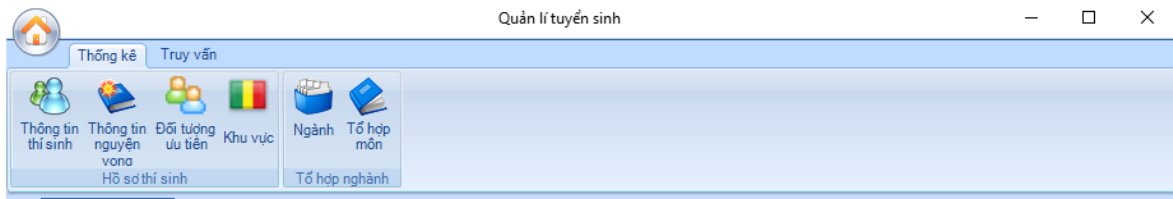


Hình 4. 23. Giao diện đăng nhập



### Giao diện sử dụng chính :

Giao diện chính bao gồm 2 tab lớn chứa 2 chức năng lớn của chương trình: thống kê, truy vấn.



Hình 4. 24. Giao diện chính của chương trình

**Thống kê:** là chức năng của hệ thống, nhằm hiển thị thông tin về cơ sở dữ liệu của hệ thống tuyển sinh. Gồm các nút chức năng nhỏ để người dùng chọn thông tin họ muốn được thống kê như: thông tin thí sinh, thông tin hồ sơ nguyện vọng, các đối tượng ưu tiên và điểm cộng, các khu vực ưu tiên và điểm cộng, các ngành, các tổ hợp. Ngoài cung cấp chức năng thống kê, các bảng thống kê còn cung cấp chức năng sắp xếp, bằng cách nhấn vào tên của cột.

MaHoSo	SBD	MaNguyenVc	MaNganh	MaToHop	Diem1	Diem2	Diem3	MaDoiTuong	MaKhuVuc	TongDiem	N
HS001	73419810...	2	7340201	D01	7,25	9,25	8,75	UT0	KV2NT	25,75	20
HS002	73419810...	2	7540101	D01	8,75	8,00	7,50	UT0	KV2NT	24,75	20
HS003	73419810...	2	7340201	D01	6,00	9,75	7,75	UT0	KV2NT	24,00	20
HS004	73419810...	1	7480201	A01	9,50	5,50	9,25	UT1	KV2NT	26,75	20
HS005	73419810...	1	7480201	D01	8,00	9,00	8,00	UT0	KV2NT	25,50	20
HS006	73419810...	1	7480201	D01	9,00	10,00	7,75	UT0	KV2NT	27,25	20
HS007	73419810...	2	7540101	A01	9,75	6,75	6,00	UT0	KV2NT	23,00	20
HS008	73419810...	1	7540101	B00	9,25	5,50	6,00	UT0	KV2NT	21,25	20
HS009	73419810...	1	7540101	A00	8,75	7,75	7,00	UT0	KV2NT	24,00	20
HS010	73419810...	1	7480201	A01	9,75	8,00	9,25	UT0	KV2NT	27,50	20
HS011	73419810...	2	7540101	A00	6,50	8,00	5,75	UT0	KV2NT	20,75	20
HS012	73419810...	2	7540101	D01	5,50	9,25	6,75	UT0	KV2NT	22,00	20
HS013	73419810...	1	7510201	D01	8,75	5,75	6,50	UT0	KV2NT	21,50	20
HS014	73419810...	2	7340201	A01	6,25	5,50	8,75	UT0	KV2NT	21,00	20
HS015	73419810...	2	7340201	D01	8,50	7,25	8,00	UT0	KV2NT	24,25	20
HS016	73419810...	1	7480201	A00	8,00	9,75	9,25	UT0	KV2NT	27,50	20

Hình 4. 25. Giao diện chức năng thống kê

**Truy vấn:** là chức năng của hệ thống, cung cấp cho người dùng khả năng thao tác với cơ sở dữ liệu đa chiều thông qua giao diện olap đơn giản, cung cấp khả năng hiển thị kết quả truy vấn đa chiều thành dạng bảng. Đồng thời cung cấp các chức năng hỗ trợ dự báo điểm chuẩn, truy vấn cơ sở dữ liệu tuyển sinh bằng các câu truy vấn có sẵn, chức năng tra cứu điểm cho sinh viên.

Quản lý tuyển sinh

Thông kê
Truy vấn

Truy vấn đa chiều

Truy vấn trực tiếp

Các câu truy vấn dữ liệu tuyển sinh

Truy vấn dành cho nhà quản lý

Dự đoán điểm chuẩn

Truy vấn dành cho nhà quản lý

Thông tin điểm chuẩn năm gần đây

Truy vấn cho sinh viên

Tra cứu điểm thi sinh

Truy vấn cho sinh viên

Truy vấn
Truy vấn cho nhà quản lý
Dự đoán điểm chuẩn
Thông tin điểm chuẩn
Tra cứu điểm thi sinh

Chọn thước đo : Thước đo Tổng điểm Thêm thước đo

Chọn chiều : Số báo danh Thêm chiều

Chọn điều kiện : Tên ngành Công nghệ thông tin Thêm điều kiện

SELECT [Measures].[Tong Diem] on 0,  
 {[THISINH].[SBD].[SBD]} on 1  
 FROM [QUAN LY TUYEN SINH]  
 Where [NGANH].[Ten Nganh].&[Công nghệ thông tin]

Chạy lệnh
Xóa

Column1	KeyColumn1	Tong Diem
7341981001	[THISINH].[SBD].&[7341981001]	
7341981002	[THISINH].[SBD].&[7341981002]	
7341981003	[THISINH].[SBD].&[7341981003]	
7341981004	[THISINH].[SBD].&[7341981004]	26.75
7341981005	[THISINH].[SBD].&[7341981005]	25.5
7341981006	[THISINH].[SBD].&[7341981006]	27.25
7341981007	[THISINH].[SBD].&[7341981007]	
7341981008	[THISINH].[SBD].&[7341981008]	
7341981009	[THISINH].[SBD].&[7341981009]	
7341981010	[THISINH].[SBD].&[7341981010]	27.5
7341981011	[THISINH].[SBD].&[7341981011]	
7341981012	[THISINH].[SBD].&[7341981012]	

*Hình 4. 26. Chức năng truy vấn đa chiều của hệ thống*

**Chức năng truy vấn đa chiều của hệ thống:**

- Nút “Thêm thước đo”: thêm thước đo được chọn từ combobox chọn thước đo vào textbox câu truy vấn, không cho chọn trùng thước đo.
- Nút “thêm chiều”: thêm chiều được chọn từ combobox chọn chiều vào textbox câu truy vấn, không cho chọn trùng chiều.
- Nút “Thêm điều kiện”: thêm lát cắt được chọn từ 2 combobox chọn điều kiện, không cho chọn trùng điều kiện, không cho chọn điều kiện khác cấp với điều kiện trước.
- Nút “Chạy lệnh”: Thiết lập kết nối với kho dữ liệu, lấy kết quả câu truy vấn đa chiều trả về dạng bảng trong datagridview.
- Nút “Xóa”: Xóa các trạng thái hiện tại.
- Textbox “câu truy vấn”: có thể sử dụng để viết trực tiếp câu lệnh truy vấn, câu lệnh sẽ được thực hiện sau khi nhấn nút “chạy lệnh”.

Column1	KeyColumn1	2020	2021	2022	All
CN kỹ thuật cơ khí	[NGANH].[Ten Nganh].&[...]	37	25	38	100
Công nghệ thông tin	[NGANH].[Ten Nganh].&[...]	36	47	49	132
Công nghệ thực phẩm	[NGANH].[Ten Nganh].&[...]	23	24	37	84
Tài chính ngân hàng	[NGANH].[Ten Nganh].&[...]	24	24	36	84
All	[NGANH].[Ten Nganh].[All]	120	120	160	400

Hình 4. 27. Chức năng truy vấn cho nhà quản lí

### Chức năng truy vấn cho nhà quản lí:

Combobox “Chọn câu truy vấn”: chọn câu truy vấn và kết quả được hiển thị lên datagrid view.

Dự đoán điểm chuẩn năm 2023

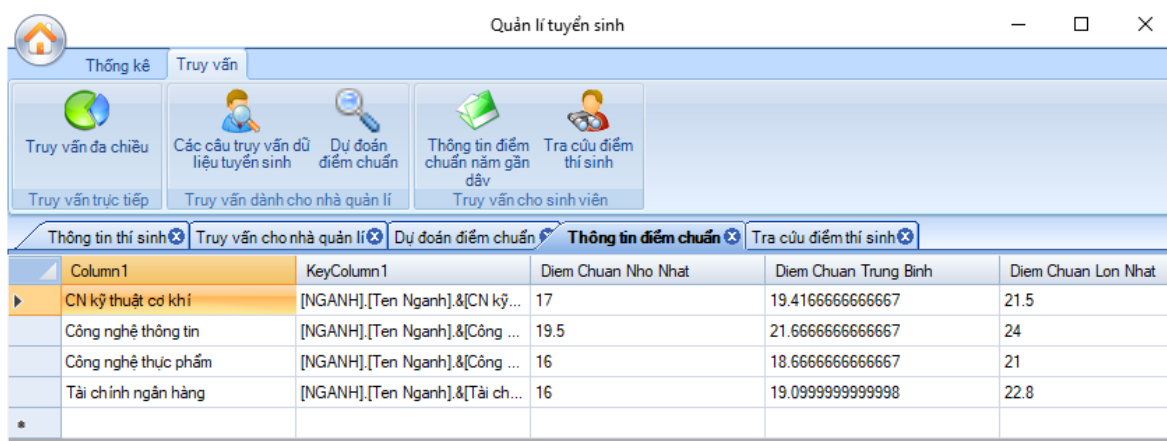
Chọn ngành dự đoán:  Điểm chuẩn dự đoán:  - 0.05 +

Số thí sinh lấy được:  - 1 +

Hình 4. 28. Chức năng dự đoán điểm chuẩn

### Chức năng dự đoán điểm chuẩn:

- Combobox “Chọn ngành dự đoán”: sau khi chọn ngành được hiển thị trên combobox, textbox “Điểm chuẩn dự đoán” sẽ nhận điểm chuẩn trung bình những năm gần đây của ngành được chọn và textbox “Số thí sinh lấy được” sẽ hiển thị số thí sinh đạt điểm chuẩn của ngành đó.
- Nút “+”: tăng giá trị của các text box theo bước nhảy được nhập trong textbox ở giữa.
- Nút “-”: giảm giá trị của các text box theo bước nhảy được nhập trong textbox ở giữa.
- Textbox “Điểm chuẩn dự đoán”: giá trị của điểm thay đổi sẽ dẫn tới giá trị của textbox “Số thí sinh lấy được” thay đổi.
- Textbox “Số thí sinh lấy được”: giá trị của số lượng thí sinh thay đổi sẽ dẫn tới giá trị của Textbox “Điểm chuẩn dự đoán” thay đổi

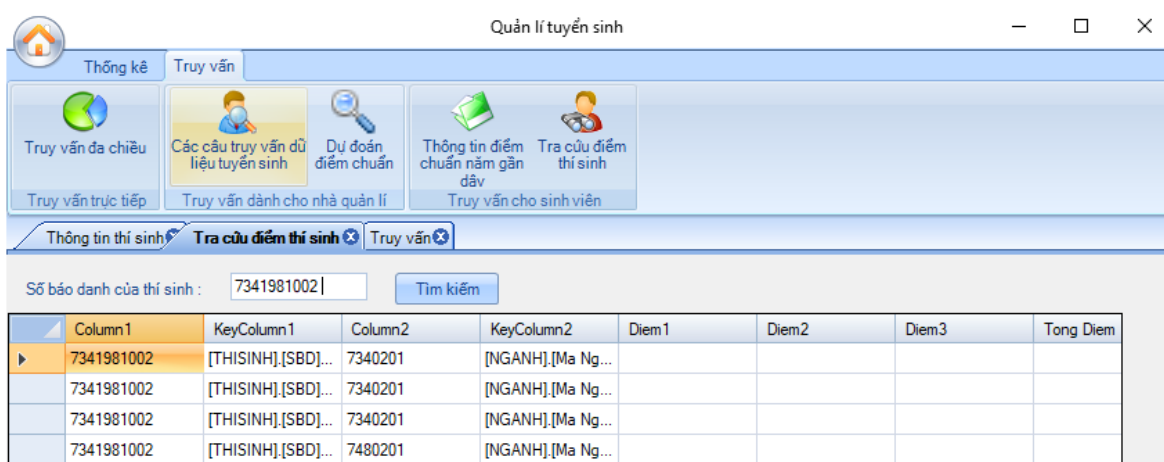


Column1	KeyColumn1	Diem Chuan Nho Nhất	Diem Chuan Trung Binh	Diem Chuan Lon Nhất
CN kỹ thuật cơ khí	[NGANH].[Ten Nganh].&[CN kỹ...	17	19.4166666666667	21.5
Công nghệ thông tin	[NGANH].[Ten Nganh].&[Công ...	19.5	21.6666666666667	24
Công nghệ thực phẩm	[NGANH].[Ten Nganh].&[Công ...	16	18.6666666666667	21
Tài chính ngân hàng	[NGANH].[Ten Nganh].&[Tài ch...	16	19.0999999999999	22.8

Hình 4. 29. Chức năng thông tin điểm chuẩn

### Chức năng thông tin điểm chuẩn:

- Xuất thông tin điểm chuẩn nhỏ nhất, trung bình, lớn nhất, theo ngành những năm gần đây.



Column1	KeyColumn1	Column2	KeyColumn2	Diem1	Diem2	Diem3	Tong Diem
7341981002	[THISINH].[SBD]...	7340201	[NGANH].[Ma Ng...				
7341981002	[THISINH].[SBD]...	7340201	[NGANH].[Ma Ng...				
7341981002	[THISINH].[SBD]...	7340201	[NGANH].[Ma Ng...				
7341981002	[THISINH].[SBD]...	7480201	[NGANH].[Ma Ng...				

Hình 4. 30. Chức năng tra cứu điểm thi sinh

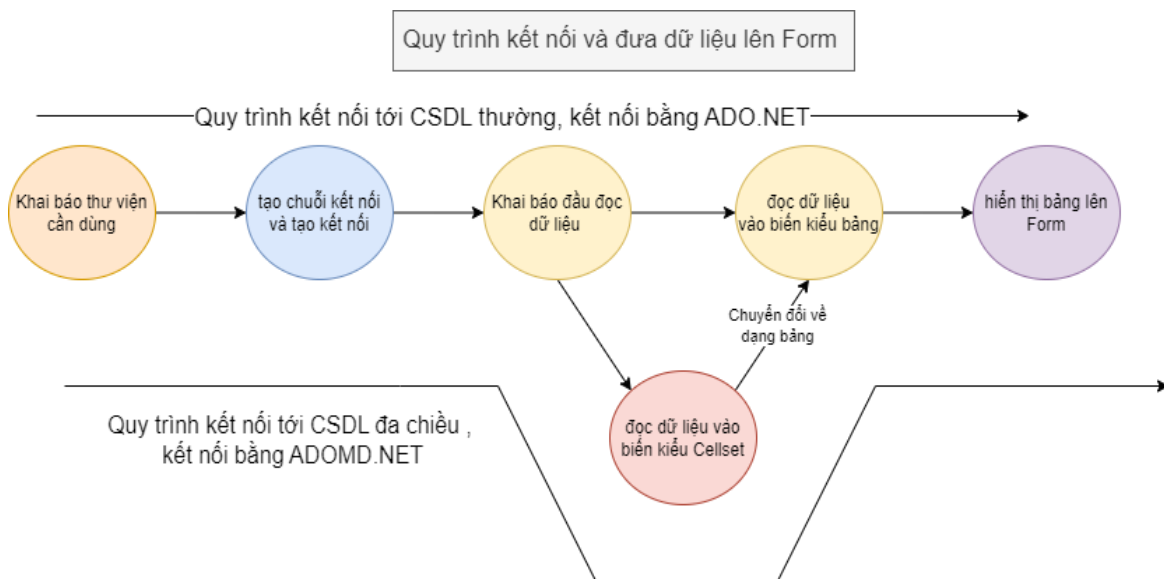
### Chức năng tra cứu điểm thi sinh:

Textbox “Số báo danh của thí sinh”: Nhập số báo danh và ấn Enter hoặc bấm nút “tìm kiếm” để tìm kiếm thông tin điểm thi của thí sinh.

### 4.4.Lập trình chương trình Quản lý tuyển sinh

Chương trình được viết bằng ngôn ngữ C# trên nền tảng lập trình .NET, thiết kế bằng WinForms.

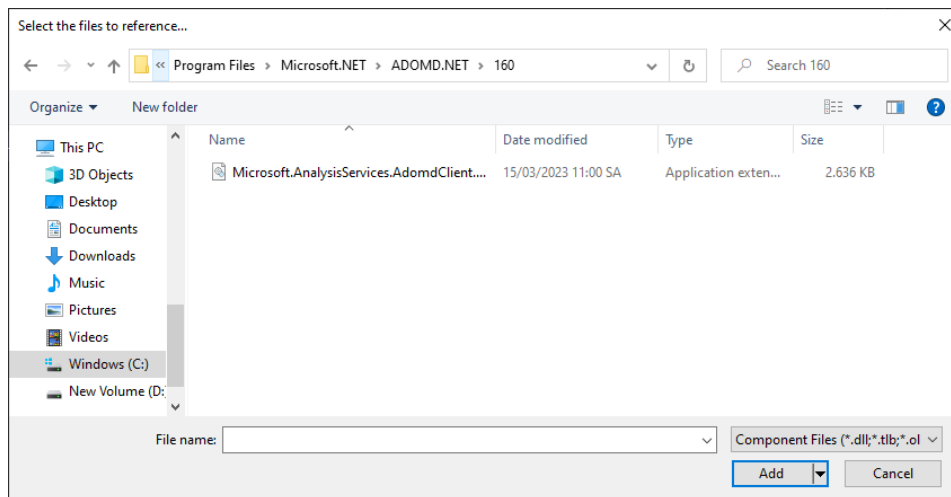
Trong chương trình cần xử lý quá trình trao đổi và kết nối dữ liệu giữa from giao diện và kết quả truy vấn MDX trên Kho dữ liệu theo một quy trình sau:



Hình 4. 31. Quy trình kết nối và đưa dữ liệu lên Form

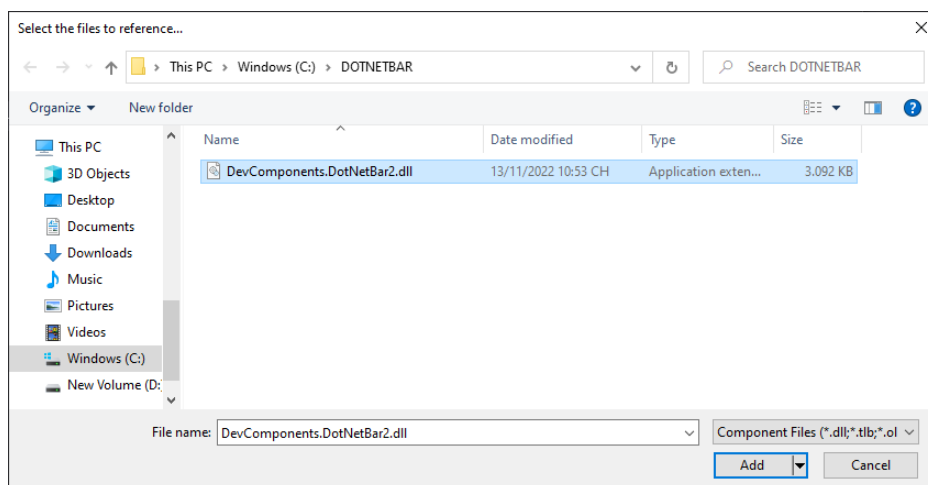
#### 4.4.1 Các thư viện sử dụng trong lập trình WinForms

- Thư viện **ADO.net** trong .NET Framework, được phát triển bởi Microsoft, để hỗ trợ việc truy cập và quản lý dữ liệu từ các nguồn dữ liệu khác nhau. Nó cung cấp các lớp và đối tượng để thực hiện các thao tác liên quan đến cơ sở dữ liệu, chẳng hạn như kết nối, truy vấn, thêm, sửa, xóa dữ liệu.
- Thư viện **ADOM.net** trong .NET Framework, được phát triển bởi Microsoft, để hỗ trợ việc truy xuất và xử lý dữ liệu đa chiều (*multidimensional data*) từ các cube OLAP thông qua kết nối ADO.NET. ADOMD.NET cung cấp các đối tượng và lớp để lập trình viên có thể phát triển các ứng dụng truy xuất dữ liệu OLAP.



Hình 4. 32. Cửa sổ thêm thư viện ADOMD.NET

- Thư viện **DotNetBar2** được phát triển bởi DevComponents LLC, cung cấp một loạt các thành phần giao diện người dùng (UI controls) tùy chỉnh cho các ứng dụng Windows Forms trong .NET Framework.



Hình 4. 33. Cửa sổ thêm thư viện DotNetBar2

#### 4.4.2 Triển khai lập trình

Kết nối với CSDL bằng thư viện ADO.net truy xuất CSDL thường phục vụ các chức năng trong chức năng “**Thống kê**”:

- + *Sử dụng thêm các thư viện:*

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

- + *Khai báo chuỗi kết nối:*

```
private static string StringCon = @"Data Source=LAPTOP-
V6MP41NS;Initial Catalog=QUAN_LI_TUYEN_SINH;Integrated Security=True";
```

- + *Khai báo hàm tạo kết nối để kết nối tới csdl qua chuỗi kết nối :*

```
private static SqlConnection NewConnection()
{
    return new SqlConnection(StringCon);
}
```

- + *Hàm trả về bảng thông qua tham số là câu lệnh truy vấn SQL:*

```
public static DataTable GetDataTable(string sqlCommand){...}
```

Kết nối với kho dữ liệu bằng thư viện ADOMD.net phục vụ cho các chức năng trong chức năng “**Truy vấn**”:

- + *Sử dụng thêm các thư viện:*

```
using System.Data;
using System.Data.SqlClient;
using Microsoft.AnalysisServices.AdomdClient;
using Tuple = Microsoft.AnalysisServices.AdomdClient.Tuple;
```

- + *Khai báo chuỗi kết nối:*

```
static string con = @"Data Source=LAPTOP-V6MP41NS;Catalog=QUANLITUYENSINH";
```

- + *Tạo kết nối mới từ chuỗi kết nối*

```
static AdomdConnection adomdcon;
```

- + *Hàm trả về kiểu dữ liệu Cellset là kết quả của truy vấn đa chiều:*

```
public static CellSet GetCellset(string command) {...}
```

- + *Hàm chuyển đổi từ kiểu dữ liệu Cellset sang DataTable để có thể hiển thị trên dataGridView:*

```
public static DataTable ToDataTable(CellSet cs){...}
```

Toàn bộ nội dung demo về Thiết kế giao diện bằng thư viện **DotNetBar2**, và lập trình các module trong chương trình được nhóm em chia sẻ trong đường link dưới đây:

Source code : [https://drive.google.com/drive/folders/139TDvgWY09-70Mg-8qiJ10\\_k3PLq6ppt?usp=sharing](https://drive.google.com/drive/folders/139TDvgWY09-70Mg-8qiJ10_k3PLq6ppt?usp=sharing)

# KẾT LUẬN VÀ KIẾN NGHỊ

## Kết luận

Trải qua quá trình nghiên cứu và triển khai đề tài, nhóm chúng em đã đạt được một số kết quả sau:

- Nghiên cứu lý thuyết: Ngôn ngữ MDX, Kho dữ liệu, và Phân tích trực tuyến OLAP
- Xây dựng kho dữ liệu mô phỏng hệ thống tuyển sinh
- Mô phỏng cách thực thi các truy vấn MDX trên SQL Server
- Sử dụng kỹ thuật phân tích trực tuyến OLAP trong phân tích dữ liệu tuyển sinh
- Xây dựng chương trình với tính năng mô phỏng ngôn ngữ MDX và hỗ trợ phân tích điểm chuẩn, lựa chọn nguyện vọng đăng ký của thí sinh.

## Kiến nghị

Đề tài đang triển khai trên dữ liệu tự xây dựng mô phỏng dữ liệu tuyển sinh đạt học. Tuy nhiên, nếu được hỗ trợ cung cấp dữ liệu tuyển sinh thực tế của các năm học đã qua và thông tin thí sinh đăng ký hiện tại từ Nhà trường, chương trình sẽ mang tính thực tế hơn.



## TÀI LIỆU THAM KHẢO

- [1] Tóm tắt luận văn “*DataWarehouse và ứng dụng trong bài toán quản lý cửa hàng bán lẻ xăng dầu*”, Nguyễn Quang Huy, ĐH Công nghệ - ĐH Quốc Gia HN, 2012.
- [2] Wiley, *Buiding the DataWarehouse*, Fourth Editic, William H.Inmon.
- [3] Ngôn ngữ MDX, <https://learn.microsoft.com/vi-vn/sql/mdx/mdx-function-reference-mdx?view=sql-server-2017> , ngày 03/01/2023.
- [4] Quy định tuyển sinh đại học năm 2022, <https://tuyensinh.uneti.edu.vn/tin-tuc/nhung-luu-y-cho-thi-sinh-dang-ky-xet-tuyen-nguyen-vong-nam-2022-296.html>,
- [5] <http://bis.net.vn/forums/p/560/1083.aspx>
- [6] Trang tuyển sinh của trường ĐHKTKTCN, <https://tuyensinh.uneti.edu.vn/thong-bao/thong-bao-tuyen-sinh-dai-hoc-he-chinh-quy-nam-2022-289.html>, ngày 21/6/2022

## LỜI CẢM ƠN

Chúng em xin gửi lời cảm ơn chân thành đến Nhà trường và Khoa Công nghệ thông tin đã tạo ra sân chơi để chúng em có cơ hội học tập và nghiên cứu các kiến thức khoa học mới và bổ ích.

Chúng em cũng xin gửi lời cảm ơn đến Cô Vũ Mỹ Hạnh đã trực tiếp hướng dẫn, tận tình chỉ bảo, truyền đạt nhiều kiến thức, kinh nghiệm cho chúng em trong quá trình làm đề tài này.

Cuối cùng, chúng em xin kính chúc quý Thầy Cô thật nhiều sức khỏe, thành công, tiếp tục thực hiện sứ mệnh cao đẹp của mình là truyền đạt kiến thức cho thế hệ mai sau.

*Chúng em xin chân thành cảm ơn!*