**BỘ CÔNG THƯƠNG**



**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP.HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO BÀI TẬP NHÓM**

**THỰC HÀNH KHO DỮ LIỆU VÀ OLAP**

**Đề tài:XÂY DỰNG KHO DỮ LIỆU PHÂN TÍCH KINH DOANH CỦA CỬA HÀNG BÁN QUẦN ÁO**

Giảng viên hướng dẫn : Đinh Thị Mận

TP. HỒ CHÍ MINH – 2024

**BẢNG PHÂN CÔNG NHIỆM VỤ**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | MSSV | Họ và Tên | Phân công | Mức độ hoàn thành |
| 1 | 2001215906 | Nguyễn Ngọc Lâm | Phân tích dữ liệu SSAS, SSIS | 100% |
| 2 | 2001210503 | Nguyễn Lê Phong | Kết nối SSAS với WindowForms và xử lý | 100% |
| 3 | 2001215638 | Phạm Hữu Chiến | Chỉnh sửa word và truy vấn mdx | 100% |
| 4 | 2001210328 | Trần Gia Bảo | Kết nối SSAS với WindowForms và xử lý | 100% |
| 5 | 2001216268 | Võ Minh Trường | Chỉnh sửa word và truy vấn mdx | 100% |

**MỤC LỤC**

[CHƯƠNG 1 TỔNG QUAN ĐỀ TÀI 7](#_Toc185342688)

[1.1. GIỚI THIỆU 7](#_Toc185342689)

[**1.2. MỤC TIÊU VÀ PHẠM VI ĐỀ TÀI 8**](#_Toc185342690)

[1.2.1. Mục tiêu tổng quát 8](#_Toc185342691)

[1.2.2. Mục tiêu cụ thể 8](#_Toc185342692)

[1.2.3. Phạm vi về nội dung 9](#_Toc185342693)

[1.2.4. Phạm vi về không gian 9](#_Toc185342694)

[**1.3. KHẢO SÁT HỆ THỐNG 10**](#_Toc185342695)

[**1.3.1.** Khảo sát thực trạng 10](#_Toc185342696)

[1.4. ĐỐI TƯỢNG NGHIÊN CỨU 11](#_Toc185342697)

[1.4.1. Đối tượng nghiên cứu chính 11](#_Toc185342698)

[1.4.2. Đối tượng người dùng 12](#_Toc185342699)

[CHƯƠNG 2 CƠ SỞ LÝ THUYẾT 13](#_Toc185342700)

[2.1. ĐỐI TƯỢNG NGHIÊN CỨU 13](#_Toc185342701)

[2.1.1. Định nghĩa kho dữ liệu (Data Warehouse) 13](#_Toc185342702)

[2.1.2. Vai trò của kho dữ liệu trong kinh doanh 14](#_Toc185342703)

[2.2. OLAP TRONG KHO DỮ LIỆU 15](#_Toc185342704)

[2.2.1. Định nghĩa OLAP (Online Analytical Processing) 15](#_Toc185342705)

[2.2.2. Các mô hình OLAP 16](#_Toc185342706)

[**2.3. QUY TRÌNH ETL TRONG XÂY DỰNG KHO DỮ LIỆU OLAP** 17](#_Toc185342707)

[2.3.1. Extract (Trích xuất dữ liệu) 17](#_Toc185342708)

[2.3.2. Transform (Chuyển đổi dữ liệu) 17](#_Toc185342709)

[2.3.3. Load (Nạp dữ liệu vào kho dữ liệu) 18](#_Toc185342710)

[2.4. **CÔNG CỤ XÂY DỰNG KHO DỮ LIỆU OLAP** 18](#_Toc185342711)

[2.4.1. Công cụ ETL (Extract, Transform, Load) 18](#_Toc185342712)

[2.4.2. Công cụ lưu trữ dữ liệu OLAP 19](#_Toc185342713)

[2.4.3. Công cụ trực quan hóa và phân tích dữ liệu 19](#_Toc185342714)

[CHƯƠNG 3 THIẾT KẾ KHO DỮ LIỆU 21](#_Toc185342715)

[3.1. PHÂN TÍCH ĐỀ TÀI 21](#_Toc185342716)

[3.1.1. Phân tích yêu cầu 21](#_Toc185342717)

[3.1.2. Yêu cầu hệ thống 22](#_Toc185342718)

[3.2. THIẾT KẾ HỆ THỐNG 24](#_Toc185342719)

[3.2.1. Cơ sở dữ liệu 24](#_Toc185342720)

[3.2.2. Kho dữ liệu 29](#_Toc185342721)

[3.2.3. Lược đồ quan hệ 32](#_Toc185342722)

[3.2.4. SSIS 34](#_Toc185342723)

[3.2.4. SSAS 40](#_Toc185342724)

[CHƯƠNG 4 CÀI ĐẶT ỨNG DỤNG 42](#_Toc185342725)

[CHƯƠNG 5 KẾT LUẬN 53](#_Toc185342731)

[5.1. KẾT QUẢ 55](#_Toc185342732)

[5.1.1. Về mặt chức năng 55](#_Toc185342733)

[5.1.2. Về mặt kỹ thuật 55](#_Toc185342734)

[5.2. HẠN CHẾ 56](#_Toc185342735)

[5.2.1. Hạn chế về chức năng 56](#_Toc185342736)

[5.2.2. Hạn chế về kỹ thuật 56](#_Toc185342737)

[5.4. KẾT LUẬN 56](#_Toc185342738)

[TÀI LIỆU THAM KHẢO 57](#_Toc185342739)

**DANH MỤC HÌNH ẢNH**

[Hình 3.1: Database Diagram 32](#_Toc185345504)

[Hình 3.2: Kho dữ liệu mô hình chòm sao 33](#_Toc185345505)

[Hình 3.3: SSIS 34](#_Toc185345506)

[Hình 3.4: File Excel 34](#_Toc185345507)

[Hình 3.5: SSIS Dim\_Time 35](#_Toc185345508)

[Hình 3.6: SSIS Dim\_Coupon 36](#_Toc185345509)

[Hình 3.7: SSIS Dim\_Customer 36](#_Toc185345510)

[Hình 3.8: SSIS Dim\_Product 37](#_Toc185345511)

[Hình 3.9: SSIS Dim\_Geography 37](#_Toc185345512)

[Hình 3.10: SSIS Dim\_Warehouse 38](#_Toc185345513)

[Hình 3.11: SSIS Fact 39](#_Toc185345514)

[Hình 3.12: SSAS 40](#_Toc185345515)

[Hình 3.13: SSAS Cube 41](#_Toc185345516)

[Hình 4.1: Giao diện trang chủ 42](#_Toc185345517)

[Hình 4.2: Giao diện truy vấn MDX 43](#_Toc185345518)

[Hình 4.3: Giao diện thực hiện xuất PDF 44](#_Toc185345519)

[Hình 4.4: File PDF xuất ra 44](#_Toc185345520)

[Hình 4.5: Thống kê 1 45](#_Toc185345521)

[Hình 4.6: Thống kê 2 46](#_Toc185345522)

[Hình 4.7: Thống kê 3 47](#_Toc185345523)

[Hình 4.8: Thống kê 4 48](#_Toc185345524)

[Hình 4.9: Thống kê 5 49](#_Toc185345525)

[Hình 4.10: Giao diện Tableau 50](#_Toc185345526)

[Hình 4.11: Truy vấn tổng doanh thu theo năm và lọc theo doanh thu > 0 51](#_Toc185345527)

[Hình 4.12: Truy vấn tổng doanh thu theo thành phố và lọc theo doanh thu > 1000 51](#_Toc185345528)

[Hình 4.13: Truy vấn tổng doanh thu theo sản phẩm và khách hàng, để tính doanh thu trung bình 51](#_Toc185345529)

[Hình 4.14: Truy vấn lấy top 5 sản phẩm theo doanh thu trong năm hiện tại và sắp xếp theo thứ tự giảm dần 52](#_Toc185345530)

[Hình 4.15: Truy vấn tổng doanh thu theo quý và sắp xếp theo thứ tự giảm dần 52](#_Toc185345531)

[Hình 4.16: Truy vấn doanh thu trung bình theo dòng sản phẩm và năm 53](#_Toc185345532)

[Hình 4.17: Truy vấn tổng doanh thu và số lượng đơn hàng theo lãnh thổ 53](#_Toc185345533)

[Hình 4.18: Truy vấn tổng số lượng đơn hàng theo khách hàng 53](#_Toc185345534)

[Hình 4.20: Truy vấn tổng doanh thu theo quý 54](#_Toc185345535)

[Hình 4.19: Truy vấn tổng số lượng đơn hàng theo sản phẩm 54](#_Toc185345536)

# CHƯƠNG 1 TỔNG QUAN ĐỀ TÀI

## GIỚI THIỆU

Đề tài xây dựng kho dữ liệu phân tích kinh doanh cửa hàng bán quần áo với việc sử dụng dữ liệu để đưa ra các chiến lược kinh doanh, thống kê một cách chính xác và kịp thời đang là một yếu tố quan trọng trong lĩnh vực kinh doanh thời trang. Kho dữ liệu này không chỉ hỗ trợ các cửa hàng quần áo và lưu trữ dữ liệu từ nhiều nguồn khác nhau mà còn lập báo cáo và giám sát chi tiết để cải thiện phương pháp tối ưu hóa bán hàng và kiểm soát hàng tồn kho, đồng thời nâng cao dịch vụ khách hàng.

## 1.2. MỤC TIÊU VÀ PHẠM VI ĐỀ TÀI

### 1.2.1. Mục tiêu tổng quát

Xây dựng một hệ thống kho dữ liệu phân tích kinh doanh hiện đại và hiệu quả cho cửa hàng bán quần áo, giúp cải thiện khả năng quản lý, tối ưu hóa hoạt động kinh doanh và hỗ trợ ra quyết định dựa trên dữ liệu.

### 1.2.2. Mục tiêu cụ thể

Kho lưu trữ dữ liệu được kiến ​​trúc và tạo ra để đáp ứng đầy đủ các yêu cầu, thao tác và trích xuất chi tiết từ nhiều nguồn khác nhau như hệ thống quản lý bán lẻ, hồ sơ khách hàng và dữ liệu sản phẩm. Hệ thống còn đưa ra các báo cáo thể hiện cách thức hoạt động của doanh nghiệp. Các chỉ số như thu nhập, lượng hàng tồn kho và tác động tiếp thị được hiển thị một cách rõ ràng và dễ hiểu, mang lại sự hỗ trợ quyết định mạnh mẽ. Hiệu suất của hệ thống được đánh giá bằng các số liệu kinh doanh thực tế trong hoạt động hàng ngày của cửa hàng.

### 1.2.3. Phạm vi về nội dung

Đề tài tập trung vào việc xây dựng hệ thống kho dữ liệu phục vụ phân tích kinh doanh cho cửa hàng bán quần áo, bao gồm các nội dung:

* Thu thập dữ liệu**:** Thu thập từ excel của trang web bán quần áo và database của cửa hàng bán quần áo
* Xử lý dữ liệu**:** Triển khai quy trình ETL (Extract, Transform, Load) để làm sạch và chuẩn hóa dữ liệu từ nhiều nguồn khác nhau.
* Lưu trữ dữ liệu**:** Chuyển từ các quy trình xử lý dữ liệu ra mô hình DDS
* Phân tích dữ liệu**:** Phân tích từ SSAS

### 1.2.4. Phạm vi về không gian

Đề tài được triển khai trong phạm vi hoạt động của một cửa hàng bán quần áo cụ thể. Phạm vi không gian tập trung vào dữ liệu nội bộ của cửa hàng, bao gồm thông tin bán hàng, quản lý hàng hóa, và thông tin khách hàng trong hệ thống hiện tại.

## 1.3. KHẢO SÁT HỆ THỐNG

* + 1. **Khảo sát thực trạng**

**Thực trạng quản lý dữ liệu:**

Hiện tại, công việc quản lý dữ liệu tại cửa hàng bán quần áo vẫn còn khá lộn xộn và thiếu tính kết nối. Mọi thứ chủ yếu được thực hiện qua các hệ thống rời rạc và công cụ thủ công như file Excel. Dữ liệu từ các bộ phận như bán hàng, tồn kho và thông tin khách hàng không đồng bộ và không có sự liên kết chặt chẽ. Điều này khiến dữ liệu bị phân tán, khó kiểm soát, thậm chí còn bị trùng lặp hoặc sai lệch.

**Thực trạng phân tích và ra quyết định kinh doanh:**

Cửa hàng vẫn chưa áp dụng các công cụ phân tích hiện đại để hỗ trợ quá trình ra quyết định kinh doanh. Hiện tại, các báo cáo chỉ dừng lại ở mức tổng hợp đơn giản, chưa thể phản ánh toàn diện mọi khía cạnh hoạt động của cửa hàng.

Cửa hàng vẫn còn thiếu khả năng phân tích chuyên sâu, như xác định nhóm khách hàng tiềm năng, nắm bắt xu hướng mua sắm hay đo lường hiệu quả của các chương trình khuyến mãi. Thay vì dựa vào dữ liệu cụ thể, các quyết định phần lớn vẫn dựa vào kinh nghiệm cá nhân, khiến hiệu quả kinh doanh chưa thực sự ổn định.

**Thách thức lớn:**

Hiện tại, hoạt động quản lý dữ liệu tại cửa hàng bán quần áo vẫn còn phụ thuộc vào các hệ thống rời rạc và công cụ thủ công như file Excel. Dữ liệu từ các bộ phận như bán hàng, tồn kho và thông tin khách hàng vẫn đang "đứng riêng lẻ", chưa được kết nối một cách chặt chẽ. Điều này khiến dữ liệu trở nên phân tán, khó kiểm soát và không ít lần rơi vào tình trạng trùng lặp hoặc sai lệch.

## 1.4. ĐỐI TƯỢNG NGHIÊN CỨU

### 1.4.1. Đối tượng nghiên cứu chính

Đối tượng nghiên cứu chính của đề tài là dữ liệu kinh doanh liên quan đến hoạt động của cửa hàng bán quần áo. Các loại dữ liệu này bao gồm:

* Dữ liệu bán hàng: Ghi nhận doanh thu, số lượng sản phẩm bán ra, thời điểm giao dịch, và các chương trình khuyến mãi đi kèm.
* Dữ liệu hàng tồn kho: Thông tin về số lượng hàng hóa còn lại, tình trạng nhập - xuất hàng, và các vấn đề liên quan đến quản lý tồn kho.
* Dữ liệu khách hàng: Hồ sơ khách hàng, lịch sử mua sắm, các phân nhóm khách hàng (khách hàng tiềm năng, khách hàng trung thành).
* Dữ liệu từ thị trường: Xu hướng thời trang, hành vi tiêu dùng, và các yếu tố bên ngoài ảnh hưởng đến nhu cầu mua sắm.
* Dữ liệu từ vị trí địa lý: Thông tin về địa điểm mua hàng, từ mức quốc gia, tiểu bang, thành phố, quận.
* Dữ liệu từ coupon: Lấy dữ liệu thống kê về các mã giảm giá được sử dụng

### 1.4.2. Đối tượng người dùng

**Chủ cửa hàng:**

Chủ cửa hàng đóng vai trò là người đưa ra các quyết định chiến lược, sử dụng hệ thống để nắm bắt bức tranh tổng thể về hoạt động kinh doanh. Dựa trên các báo cáo chi tiết và phân tích xu hướng, chủ cửa hàng có thể xây dựng các kế hoạch dài hạn nhằm tối ưu hóa doanh thu và lợi nhuận.

**Quản lý cửa hàng:**

Quản lý cửa hàng chịu trách nhiệm điều hành các hoạt động vận hành hàng ngày. Họ sử dụng hệ thống để kiểm soát tình trạng tồn kho, điều phối hoạt động nhập hàng và đánh giá hiệu suất làm việc của nhân viên, đảm bảo mọi hoạt động diễn ra trơn truvà hiệu quả.

# CHƯƠNG 2 CƠ SỞ LÝ THUYẾT

## 2.1. ĐỐI TƯỢNG NGHIÊN CỨU

### 2.1.1. Định nghĩa kho dữ liệu (Data Warehouse)

Kho dữ liệu (Data Warehouse) là một hệ thống tập trung lưu trữ và tổ chức dữ liệu từ nhiều nguồn khác nhau nhằm phục vụ các mục đích phân tích và hỗ trợ ra quyết định. Đây là một nền tảng quan trọng trong việc triển khai các hệ thống thông tin hỗ trợ quyết định (DSS - Decision Support Systems), đặc biệt là các ứng dụng xử lý phân tích trực tuyến (OLAP - Online Analytical Processing).

Khác với cơ sở dữ liệu giao dịch (OLTP - Online Transaction Processing), nơi dữ liệu được xử lý để phục vụ các giao dịch hàng ngày, kho dữ liệu tập trung vào việc lưu trữ và phân tích dữ liệu lịch sử. Dữ liệu trong kho thường được tổ chức theo các chủ đề cụ thể (subject-oriented), không thay đổi (non-volatile), và mang tính thời gian (time-variant).

##### **Các đặc điểm chính của kho dữ liệu**

* **Tích hợp (Integration):**  
  Kho dữ liệu tích hợp dữ liệu từ nhiều nguồn khác nhau, bao gồm các cơ sở dữ liệu giao dịch, file Excel, và thậm chí cả dữ liệu từ mạng xã hội. Quy trình này đòi hỏi xử lý và chuẩn hóa dữ liệu để đảm bảo tính nhất quán.
* **Định hướng chủ đề (Subject-oriented):**  
  Dữ liệu trong kho được tổ chức xung quanh các lĩnh vực cụ thể trong doanh nghiệp, như bán hàng, khách hàng, sản phẩm, hoặc tồn kho. Điều này giúp doanh nghiệp dễ dàng truy cập và phân tích dữ liệu liên quan đến một chủ đề cụ thể mà không bị phân tán bởi các thông tin không cần thiết.
* **Không thay đổi (Non-volatile):**  
  Một khi dữ liệu đã được nạp vào kho, nó không bị chỉnh sửa hoặc xóa đi. Điều này giúp duy trì lịch sử dữ liệu và cho phép thực hiện các phân tích dài hạn dựa trên dữ liệu chính xác.
* **Dữ liệu thời gian (Time-variant):**  
  Dữ liệu trong kho thường chứa thông tin liên quan đến các mốc thời gian cụ thể, như ngày, tuần, tháng, hoặc năm. Điều này cho phép phân tích xu hướng, so sánh các giai đoạn khác nhau, và hỗ trợ dự đoán tương lai.

### 2.1.2. Vai trò của kho dữ liệu trong kinh doanh

Kho dữ liệu đóng vai trò trung tâm trong việc hỗ trợ các hoạt động kinh doanh hiện đại, đặc biệt là trong các doanh nghiệp bán lẻ và thời trang. Dưới đây là những lợi ích chính mà kho dữ liệu mang lại:

* **Hỗ trợ ra quyết định chiến lược:**  
  Dữ liệu được tổ chức và trình bày một cách rõ ràng, giúp nhà quản lý có được cái nhìn toàn diện về hoạt động kinh doanh và đưa ra các quyết định chiến lược dựa trên thông tin thực tế.
* **Phân tích xu hướng:**  
  Kho dữ liệu cho phép theo dõi và phân tích các xu hướng kinh doanh, chẳng hạn như xu hướng tiêu dùng, hành vi mua sắm của khách hàng, hoặc hiệu quả của các chương trình khuyến mãi.
* **Tăng hiệu quả hoạt động:**  
  Việc tích hợp và tự động hóa dữ liệu từ nhiều nguồn khác nhau giúp giảm thiểu thời gian và chi phí liên quan đến việc xử lý dữ liệu thủ công.
* **Nâng cao khả năng cạnh tranh:**  
  Với khả năng phân tích nhanh chóng và chính xác, doanh nghiệp có thể nắm bắt các cơ hội thị trường, cải thiện chiến lược bán hàng, và tối ưu hóa các quy trình kinh doanh.
* **Cải thiện dịch vụ khách hàng:**  
  Phân tích dữ liệu khách hàng giúp doanh nghiệp hiểu rõ hơn về nhu cầu và mong muốn của họ, từ đó cung cấp dịch vụ và sản phẩm phù hợp hơn.

## 2.2. OLAP TRONG KHO DỮ LIỆU

### 2.2.1. Định nghĩa OLAP (Online Analytical Processing)

OLAP (Online Analytical Processing) là một công nghệ quan trọng trong hệ thống kho dữ liệu, cho phép người dùng phân tích dữ liệu đa chiều để đưa ra các quyết định kinh doanh. Khác với OLTP (Online Transaction Processing) tập trung vào việc xử lý giao dịch, OLAP tập trung vào việc phân tích dữ liệu để hỗ trợ các truy vấn phức tạp.

##### **Các đặc điểm của OLAP**

* **Hỗ trợ phân tích đa chiều (Multidimensional Analysis):**  
  OLAP tổ chức dữ liệu dưới dạng các khối (cubes), với mỗi chiều (dimension) đại diện cho một khía cạnh phân tích, chẳng hạn như thời gian, sản phẩm, địa điểm, và doanh số.
* **Truy vấn phức tạp (Complex Querying):**  
  OLAP cho phép thực hiện các truy vấn phức tạp như tổng hợp, phân tích xu hướng, hoặc so sánh dữ liệu giữa các giai đoạn khác nhau.
* **Hiệu suất cao:**  
  Dữ liệu trong OLAP thường được tổ chức và lập chỉ mục trước, giúp tăng tốc độ truy vấn ngay cả khi làm việc với khối lượng lớn dữ liệu.

### ****2.2.2. Các mô hình OLAP****

* **MOLAP (Multidimensional OLAP):**  
  MOLAP lưu trữ dữ liệu trong các khối đa chiều và tối ưu hóa cho việc truy vấn nhanh. Đây là mô hình OLAP truyền thống với hiệu suất cao trong các phân tích phức tạp.
* **ROLAP (Relational OLAP):**  
  ROLAP lưu trữ dữ liệu trong cơ sở dữ liệu quan hệ và sử dụng SQL để thực hiện truy vấn. Phù hợp với các hệ thống có khối lượng dữ liệu lớn và thường xuyên thay đổi.
* **HOLAP (Hybrid OLAP):**  
  HOLAP kết hợp ưu điểm của MOLAP và ROLAP, cho phép lưu trữ dữ liệu tổng hợp trong khối đa chiều (MOLAP) và dữ liệu chi tiết trong cơ sở dữ liệu quan hệ (ROLAP).

**2.3. QUY TRÌNH ETL TRONG XÂY DỰNG KHO DỮ LIỆU OLAP**

### ****2.3.1. Extract (Trích xuất dữ liệu)****

Quy trình trích xuất dữ liệu là bước đầu tiên trong ETL, nhằm thu thập dữ liệu từ nhiều nguồn khác nhau. Các nguồn dữ liệu phổ biến bao gồm:

* Hệ thống POS (Point of Sale): Dữ liệu giao dịch bán hàng.
* Hệ thống ERP (Enterprise Resource Planning): Dữ liệu quản lý hàng tồn kho, nhà cung cấp.
* CRM (Customer Relationship Management): Dữ liệu khách hàng và hành vi tiêu dùng.

### ****2.3.2. Transform (Chuyển đổi dữ liệu)****

Dữ liệu thô sau khi được trích xuất sẽ trải qua quá trình làm sạch và chuẩn hóa, bao gồm:

* Loại bỏ dữ liệu bị lỗi, không hợp lệ hoặc trùng lặp.
* Chuẩn hóa định dạng dữ liệu, ví dụ: chuyển đổi tất cả ngày tháng về cùng một định dạng chuẩn.
* Tích hợp dữ liệu từ nhiều nguồn để tạo thành một tập dữ liệu duy nhất, nhất quán.

### ****2.3.3. Load (Nạp dữ liệu vào kho dữ liệu)****

Sau khi chuyển đổi, dữ liệu sẽ được nạp vào kho dữ liệu OLAP. Quá trình này có thể thực hiện theo hai phương pháp chính:

* **Nạp toàn bộ (Full Load):** Áp dụng khi xây dựng kho dữ liệu lần đầu tiên.
* **Nạp gia tăng (Incremental Load):** Chỉ nạp các thay đổi mới để cập nhật dữ liệu hiện có, giúp tiết kiệm thời gian và tài nguyên.

## 2.4. CÔNG CỤ XÂY DỰNG KHO DỮ LIỆU OLAP

Để xây dựng và vận hành kho dữ liệu OLAP hiệu quả, doanh nghiệp cần sử dụng các công cụ phù hợp trong từng giai đoạn. Dưới đây là các công cụ phổ biến trong việc triển khai kho dữ liệu OLAP.

### ****2.4.1. Công cụ ETL (Extract, Transform, Load)****

* **Informatica PowerCenter:**
  + Là một công cụ ETL hàng đầu, hỗ trợ tích hợp dữ liệu từ nhiều nguồn khác nhau.
  + Có khả năng xử lý các quy trình phức tạp và đảm bảo tính nhất quán của dữ liệu.
* **Talend Open Studio:**
  + Một công cụ mã nguồn mở phổ biến, dễ dàng sử dụng và hỗ trợ các định dạng dữ liệu đa dạng.
  + Phù hợp với các doanh nghiệp vừa và nhỏ nhờ chi phí thấp và tính linh hoạt cao.
* **Microsoft SQL Server Integration Services (SSIS):**
  + Công cụ tích hợp mạnh mẽ trong hệ sinh thái Microsoft, hỗ trợ xử lý ETL tự động.

### ****2.4.2. Công cụ lưu trữ dữ liệu OLAP****

* **Microsoft SQL Server Analysis Services (SSAS):**
  + Hỗ trợ xây dựng các khối OLAP đa chiều, tối ưu hóa cho việc truy vấn.
  + Tích hợp chặt chẽ với các công cụ phân tích như Power BI.
* **Oracle Essbase:**
  + Một trong những giải pháp OLAP hàng đầu, cung cấp hiệu suất cao cho các phân tích phức tạp.
* **SAP BW/4HANA:**
  + Nền tảng kho dữ liệu tiên tiến, phù hợp với các doanh nghiệp lớn, hỗ trợ phân tích thời gian thực.

### ****2.4.3. Công cụ trực quan hóa và phân tích dữ liệu****

* **Tableau:**
  + Công cụ trực quan hóa mạnh mẽ, hỗ trợ các phân tích tương tác.
  + Phù hợp với người dùng không chuyên nhờ giao diện thân thiện.
* **Power BI:**
  + Cung cấp các báo cáo động và trực quan, tích hợp tốt với các hệ thống của Microsoft.
* **QlikView:**
  + Công cụ trực quan hóa và phân tích mạnh mẽ, giúp doanh nghiệp tìm kiếm các thông tin ẩn giấu trong dữ liệu.

# CHƯƠNG 3 THIẾT KẾ KHO DỮ LIỆU

## 3.1. PHÂN TÍCH ĐỀ TÀI

### 3.1.1. Phân tích yêu cầu

**Mục tiêu chính:**

Đề tài hướng đến việc thiết kế và triển khai một hệ thống kho dữ liệu (Data Warehouse) nhằm lưu trữ và phân tích toàn diện dữ liệu kinh doanh của cửa hàng bán quần áo. Hệ thống sẽ hỗ trợ ban quản lý ra quyết định chiến lược, tối ưu hóa vận hành và nâng cao hiệu quả kinh doanh.

**Yêu cầu cụ thể:**

**Thu thập dữ liệu:**

Tích hợp dữ liệu từ nhiều nguồn khác nhau, bao gồm:

* Bán hàng**:** Dữ liệu về doanh thu, sản phẩm, thời gian, khu vực bán hàng.
* Khách hàng**:** Thông tin cá nhân, lịch sử mua hàng, hành vi tiêu dùng.
* Kho hàng**:** Số lượng hàng tồn, sản phẩm bán chạy, sản phẩm sắp hết hàng.
* Khuyến mãi**:** Thông tin các chương trình ưu đãi, giảm giá, hiệu quả chiến dịch quảng cáo.
* Địa lý: Thông tin về các địa điểm mua hàng, địa điểm tiêu dùng

**Xử lý dữ liệu:**

Làm sạch dữ liệu để đảm bảo tính chính xác và đồng nhất.

Chuyển đổi dữ liệu từ các định dạng khác nhau sang một cấu trúc thống nhất.

Áp dụng các kỹ thuật chuẩn hóa nhằm giảm thiểu trùng lặp và tối ưu hóa lưu trữ.

**Lưu trữ dữ liệu:**

Xây dựng kho dữ liệu tập trung, được thiết kế với cấu trúc hỗ trợ phân tích và báo cáo.

Đảm bảo kho dữ liệu có khả năng mở rộng khi khối lượng dữ liệu tăng lên.

**Phân tích và báo cáo:**

Xây dựng hệ thống báo cáo linh hoạt, tập trung vào các chỉ số kinh doanh cốt lõi như:

* Doanh thu theo ngày, tháng, năm.
* Sản phẩm bán chạy nhất và kém nhất.
* Phân tích hành vi mua sắm của khách hàng.

Triển khai các công cụ trực quan hóa dữ liệu (dashboards, biểu đồ) để hỗ trợ ra quyết định.

### 3.1.2. Yêu cầu hệ thống

**Chức năng chính:**

Hệ thống kho dữ liệu phải đảm bảo cung cấp đầy đủ các chức năng sau:

**Tự động hóa quy trình ETL:**

* Trích xuất (Extract) dữ liệu từ các nguồn đầu vào.
* Chuyển đổi (Transform) dữ liệu theo yêu cầu phân tích.
* Tải (Load) dữ liệu vào kho lưu trữ.

Lưu trữ và quản lý dữ liệu:

Hệ thống cần đảm bảo tính toàn vẹn dữ liệu và hỗ trợ truy vấn nhanh.

Thiết kế kho dữ liệu theo mô hình Star Schema hoặc Snowflake Schema để tối ưu hiệu suất truy xuất và phân tích.

**Phân tích và báo cáo:**

* Cung cấp giao diện trực quan cho người dùng cuối để xem các báo cáo và biểu đồ.
* Hỗ trợ các truy vấn động (dynamic queries) để phân tích dữ liệu theo nhiều tiêu chí khác nhau.

**Quản lý bảo mật và phân quyền:**

* Xây dựng hệ thống phân quyền người dùng để đảm bảo chỉ những người có thẩm quyền mới được truy cập dữ liệu nhạy cảm.
* Áp dụng mã hóa dữ liệu nhằm ngăn chặn rủi ro bảo mật.

**Yêu cầu kỹ thuật:**

**Phần mềm:**

* Sử dụng hệ quản trị cơ sở dữ liệu mạnh mẽ như MySQL, SQL Server, hoặc MongoDB.
* Triển khai công cụ ETL chuyên dụng như Talend, Apache Nifi hoặc SSIS.
* Công cụ trực quan hóa dữ liệu như Power BI, Tableau hoặc Google Data Studio.

**Phần cứng:**

* Máy chủ lưu trữ cần đảm bảo hiệu năng cao, với khả năng xử lý hàng triệu bản ghi dữ liệu mỗi ngày.
* Sử dụng hệ thống lưu trữ có khả năng sao lưu và phục hồi dữ liệu (backup & recovery).

**Tích hợp giao diện người dùng:**

* Giao diện phải thân thiện, dễ sử dụng, với khả năng tương thích trên các thiết bị như máy tính và điện thoại.
* Hiển thị các báo cáo theo thời gian thực (real-time) để hỗ trợ ra quyết định tức thời.

## 3.2. THIẾT KẾ HỆ THỐNG

### 3.2.1. Cơ sở dữ liệu

* **Users**(**User\_id**, First\_name, Last\_name, adr, phone, email);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **User\_id** | **int** | **Mã user** |  |
| First\_name | **Nvarchar(50)** | **Họ** |  |
| Last\_name | **Nvarchar(50)** | **Tên** |  |
| adr | **Nvarchar(200)** | **Địa chỉ** |  |
| phone | **varchar(50)** | **Số điện thoại** |  |
| email | **varchar(50)** | **Email** |  |

* **Account**(**Account\_name**, Account\_password, role, User\_id);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Account\_name** | **varchar(50)** | **Tên tài khoản** |  |
| Account\_password | **Varchar(255)** | **Mật khẩu** |  |
| role | **varchar(20)** | **Vai trò** |  |
| User\_id | **int** | **Mã user** |  |

* **Coupon** (**Coupon\_code**, discount\_percent, limit);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Coupon\_code** | **varchar(20)** | **Mã coupon** |  |
| discount\_percent | **int** | **Phần trăm giảm** |  |
| limit | **int** | **Giới hạn dùng** |  |

* **Invoice**(**Invoice\_id**, Account\_name, Invoice\_date, total\_amount, status, Coupon\_code, Discount\_amount, FinalTotal);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Invoice\_id** | **varchar(20)** | **Mã hóa đơn** |  |
| Account\_name | **varchar(50)** | **Tên tài khoản** |  |
| Invoice\_date | **date** | **Ngày tạo** |  |
| total\_amount | **money** | **Tổng tiền** |  |
| status | **Nvarchar(30)** | **Trạng thái** |  |
| Coupon\_code | **varchar(20)** | **Mã coupon** |  |
| Discount\_amount | **money** | **Tiền giảm** |  |
| FinalTotal | **money** | **Tổng tiền sau cùng** |  |

* **Categories** (**Category\_id**, category\_name);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Category\_id** | **int** | **Mã loại** |  |
| category\_name | **Nvarchar(50)** | **Tên loại** |  |

* **Product**(**Product\_id**, product\_name, description, price, image\_url);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Product\_id** | **int** | **Mã sản phẩm** |  |
| product\_name | **nvarchar(50)** | **Tên sản phẩm** |  |
| description | **Nvarchar(255)** | **Mô tả** |  |
| price | **money** | **Đơn giá** |  |
| image\_url | **varchar (100)** | **Hình ảnh** |  |

* **Categories\_Product**(**Product\_id**, category\_id);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Product\_id** | **int** | **Mã sản phẩm** |  |
| category\_id | **Nvarchar(50)** | **Mã Loại** |  |

* **Cart** (**Cart\_id**, product\_id, account\_name, quantity, color, size, total\_price);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Cart\_id** | **int** | **Mã giỏ hàng** |  |
| product\_id | **int** | **Mã sản phẩm** |  |
| account\_name | **varchar(50)** | **Tên tài khoản** |  |
| quantity | **int** | **Số lượng** |  |
| color | **nvarchar(50)** | **Màu** |  |
| size | **Char(10)** | **Kích thước** |  |
| total\_price | **money** | **Tổng giá** |  |

* **Warehouse**(**Product\_id**, color, size, quantity);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **Product\_id** | **int** | **Mã sản phẩm** |  |
| color | **nvarchar(20)** | **Màu** |  |
| size | **Char(10)** | **Kích thước** |  |
| quantity | **int** | **Số lượng** |  |

* **InvoiceDetails** (**invoicesdetails\_id**, invoice\_id, product\_id, quantity, color, size, total\_price);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **invoicesdetails\_id** | **int** | **Mã chi tiết hóa đơn** |  |
| invoice\_id | **Varchar(20)** | **Mã hóa đơn** |  |
| product\_id | **int** | **Mã sản phẩm** |  |
| quantity | **int** | **Số lượng** |  |
| color | **nvarchar(50)** | **Màu** |  |
| size | **Char(10)** | **Kích thước** |  |
| total\_price | **money** | **Tổng giá** |  |

* **Review** (**review\_id**, product\_id, account\_name, rating, comment, review\_date);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **review\_id** | **int** | **Mã đánh giá** |  |
| product\_id | **int** | **Mã sản phẩm** |  |
| account\_name | **Varchar(50)** | **Tên tài khoản** |  |
| rating | **int** | **Đánh giá sao** |  |
| comment | **nvarchar(max)** | **Bình luận** |  |
| review\_date | **datetime** | **Ngày đánh giá** |  |

### 3.2.2. Kho dữ liệu

* **Geography\_Dimension** (**geography\_key**, country, state, city, postal\_code, territory);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **geography\_key** | **int** | **Mã địa điểm** |  |
| country | **Nvarchar(50)** | **Quốc gia** |  |
| state | **Nvarchar(50)** | **Tiểu bang** |  |
| city | **Nvarchar(50)** | **Thành phố** |  |
| postal\_code | **Nvarchar(20)** | **Mã bưu chính** |  |
| territory | **Nvarchar(50)** | **Lãnh thổ** |  |

* **Customer\_Dimension** (**customer\_key**, customer\_name, phone, address);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **customer\_key** | **int** | **Mã khách hàng** |  |
| customer\_name | **Nvarchar(100)** | **Tên khách hàng** |  |
| phone | **Nvarchar(50)** | **Số điện thoại** |  |
| address | **Nvarchar(255)** | **Địa chỉ** |  |

* **Product\_Dimension** (**product\_code**, product\_name, product\_line, msrp);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **product\_code** | **Nvarchar(50)** | **Mã sản phẩm** |  |
| product\_name | **Nvarchar(50)** | **Tên sản phẩm** |  |
| product\_line | **Nvarchar(255)** | **Dòng sản phẩm** |  |
| msrp | **money** | **Giá bán lẻ đề xuất** |  |

* **Time\_Dimension** (**full\_date**, day\_of\_month, month, quarter, year);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **full\_date** | **Date** | **Thời gian** |  |
| day\_of\_month | **Int** | **Ngày trong tháng** |  |
| month | **Int** | **Tháng** |  |
| quarter | **Int** | **Quý** |  |
| year | **Int** | **Năm** |  |

* **Coupon\_Dimension** (**coupon\_key**, coupon\_code);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **coupon\_key** | **int** | **Khóa coupon** |  |
| coupon\_code | **Nvarchar(20)** | **Coupon** |  |

* **Warehouse\_Dimension** (**warehouse\_key**, product\_id, stock\_quantity);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **warehouse\_key** | **int** | **Mã kho** |  |
| product\_id | **Nvarchar(50)** | **Mã sản phẩm** |  |
| stock\_quantity | **int** | **Số lượng tồn** |  |

* **Sales\_Fact** (**product\_code**, product\_name, product\_line, msrp);

|  |  |  |  |
| --- | --- | --- | --- |
| **TÊN TRƯỜNG** | **KIỂU DỮ LIỆU** | **MÔ TẢ** | **KHÓA CHÍNH** |
| **order\_number** | **Int** | **Mã đơn hàng** |  |
| product\_code | **Nvarchar(50)** | **Mã sản phẩm** |  |
| customer\_key | **Int** | **Mã khách hàng** |  |
| time\_key | **Date** | **Thời gian** |  |
| coupon\_key | **Int** | **Mã coupon** |  |
| warehouse\_key | **Int** | **Mã kho** |  |
| geography\_key | **Int** | **Mã địa điểm** |  |
| quantity\_ordered | **Int** | **Số lượng đặt hàng** |  |
| total\_sales | **Money** | **Tổng doanh thu** |  |
| order\_status | **Nvarchar(50)** | **Trạng thái đơn hàng** |  |
| deal\_size | **Nvarchar(20)** | **Kích thước giao dịch** |  |

### 3.2.3. Lược đồ quan hệ

**Cơ sở dữ liệu:** Từ dữ liệu từ SQL.

A computer screen shot of a computer

Description automatically generated

Hình 3.1: Database Diagram

**Kho dữ liệu:** Phân tích được kho dữ liệu DDS mô hình chòm sao với các bảng Sales\_Fact, Warehouse\_Dimension, Coupon\_Dimension, Geography\_Dimension, Product\_Dimension, Customer\_Dimension, Time\_Dimension.

**A screenshot of a computer

Description automatically generated**

Hình 3.2: Kho dữ liệu mô hình chòm sao

### 3.2.4. SSIS

A screenshot of a computer

Description automatically generated

Hình 3.3: SSIS

Từ dữ liệu từ SQL trên và file Excel:

A screen shot of a computer error

Description automatically generated

Hình 3.4: File Excel

Phân tích được các Dimension như sau:

**Dim\_Time:** Dùng để phân tích dữ liệu theo thời gian.

**A diagram of a flowchart

Description automatically generated**

Hình 3.5: SSIS Dim\_Time

**Dim\_Coupon:** Dùng để phân tích dữ liệu theo coupon

**A diagram of a flowchart

Description automatically generated**

Hình 3.6: SSIS Dim\_Coupon

**Dim\_Customer:** Dùng để phân tích dữ liệu theo khách hàng.

**A diagram of a software process

Description automatically generated with medium confidence**

Hình 3.7: SSIS Dim\_Customer

**Dim\_Product:** Dùng để phân tích dữ liệu theo sản phẩm

**A screenshot of a computer

Description automatically generated**

Hình 3.8: SSIS Dim\_Product

**Dim\_Geography:** Dùng để phân tích dữ liệu theo địa lý

**A diagram of a flowchart

Description automatically generated**

Hình 3.9: SSIS Dim\_Geography

**Dim\_Warehouse:** Dùng để phân tích dữ liệu theo kho.

**A diagram of a process

Description automatically generated**

Hình 3.10: SSIS Dim\_Warehouse

**Fact:** Để phân tích dữ liệu bán hàng, tổng hợp và đo lường hiệu suất

**A screenshot of a computer

Description automatically generated**

Hình 3.11: SSIS Fact

### 3.2.4. SSAS

A screenshot of a computer

Description automatically generated

Hình 3.12: SSAS

**Cube:**

**A screenshot of a computer screen

Description automatically generated**

Hình 3.13: SSAS Cube

# CHƯƠNG 4 CÀI ĐẶT ỨNG DỤNG

## 4.1. GIAO DIỆN VÀ CHỨC NĂNG CỦA ỨNG DỤNG Giao diện trang chủ: Với 3 chức năng Truy vấn MDX, Thống kê và Tableau

A screenshot of a computer

Description automatically generated

Hình 4.1: Giao diện trang chủ

* Nút Truy vấn MDX sẽ hiện ra Giao diện truy vấn MDX để truy vấn.
* Nút Thống kê sẽ hiện Giao diện thống kê.
* Nút Tableau hiển thị trang phân tích Tableau.

**Giao diện truy vấn MDX:**

A screenshot of a computer

Description automatically generated

Hình 4.2: Giao diện truy vấn MDX

* Nút thực hiện sẽ thực hiện câu truy vấn MDX ở textbox trên.
* Nút Xuất PDF sẽ xuất ra file PDF dựa trên thống kê từ truy vấn MDX.

**File PDF được thống kê và xuất ra:** A screenshot of a computer

Description automatically generated

Hình 4.3: Giao diện thực hiện xuất PDF

A screenshot of a computer

Description automatically generated

Hình 4.4: File PDF xuất ra

**Chức năng thống kê với các lựa chọn thống kê:** Dựa trên từng lựa chọn, hệ thống sẽ phân tích và hiển thị thống kê báo cáo.

A screenshot of a computer

Description automatically generated

Hình 4.5: Thống kê 1

A screenshot of a graph

Description automatically generated

Hình 4.6: Thống kê 2

A screenshot of a graph

Description automatically generated

Hình 4.7: Thống kê 3

## A screenshot of a graph Description automatically generated

Hình 4.8: Thống kê 4

## A screenshot of a graph Description automatically generated

Hình 4.9: Thống kê 5

## Giao diện tableau: Dùng để phân tích dữ liệu bằng Tableau

## A screenshot of a computer Description automatically generated

Hình 4.10: Giao diện Tableau

## 4.2. TRUY VẤN MDX

* Truy vấn tổng doanh thu theo năm và lọc theo doanh thu > 0.

A screenshot of a computer

Description automatically generated

Hình 4.11: Truy vấn tổng doanh thu theo năm và lọc theo doanh thu > 0

* A screenshot of a computer

  Description automatically generatedTruy vấn tổng doanh thu theo thành phố và lọc theo doanh thu > 1000.

Hình 4.12: Truy vấn tổng doanh thu theo thành phố và lọc theo doanh thu > 1000

* A computer code with text

  Description automatically generated with medium confidenceTruy vấn tổng doanh thu theo sản phẩm và khách hàng, để tính doanh thu trung bình.

Hình 4.13: Truy vấn tổng doanh thu theo sản phẩm và khách hàng, để tính doanh thu trung bình

* A computer screen shot of a computer code

  Description automatically generatedTruy vấn lấy top 5 sản phẩm theo doanh thu trong năm hiện tại và sắp xếp theo thứ tự giảm dần .

Hình 4.14: Truy vấn lấy top 5 sản phẩm theo doanh thu trong năm hiện tại và sắp xếp theo thứ tự giảm dần

* A computer code with text

  Description automatically generated with medium confidenceTruy vấn tổng doanh thu theo quý và sắp xếp theo thứ tự giảm dần.

Hình 4.15: Truy vấn tổng doanh thu theo quý và sắp xếp theo thứ tự giảm dần

* A screenshot of a computer

  Description automatically generatedTruy vấn doanh thu trung bình theo dòng sản phẩm và năm.

Hình 4.16: Truy vấn doanh thu trung bình theo dòng sản phẩm và năm

* A screenshot of a computer code

  Description automatically generatedTruy vấn tổng doanh thu và số lượng đơn hàng theo lãnh thổ.

Hình 4.17: Truy vấn tổng doanh thu và số lượng đơn hàng theo lãnh thổ

* A close-up of a computer screen

  Description automatically generatedTổng số lượng đơn hàng theo khách hàng.

Hình 4.18: Truy vấn tổng số lượng đơn hàng theo khách hàng

* A computer screen shot of a computer code

  Description automatically generatedA close-up of a computer code

  Description automatically generatedTổng số lượng đơn hàng theo sản phẩm.

Hình 4.20: Truy vấn tổng doanh thu theo quý

Hình 4.19: Truy vấn tổng số lượng đơn hàng theo sản phẩm

* Tổng doanh thu theo quý.

# CHƯƠNG 5 KẾT LUẬN

## 5.1. KẾT QUẢ

### 5.1.1. Về mặt chức năng

Hệ thống kho dữ liệu đã được xây dựng để đáp ứng nhu cầu phân tích kinh doanh của cửa hàng bán quần áo. Hệ thống tích hợp các chức năng quan trọng như quản lý sản phẩm, thông tin khách hàng, doanh thu và chi phí, giúp việc quản lý trở nên dễ dàng và hiệu quả hơn.

Ngoài ra, hệ thống còn cung cấp các báo cáo phân tích trực quan, hỗ trợ chủ cửa hàng đưa ra các quyết định chiến lược chính xác và kịp thời. Với khả năng truy xuất dữ liệu nhanh chóng và chính xác, hệ thống giúp tối ưu hóa quy trình kinh doanh, tiết kiệm thời gian và nâng cao hiệu quả hoạt động.

### 5.1.2. Về mặt kỹ thuật

Hệ thống kho dữ liệu được xây dựng theo kiến trúc Data Warehouse, đảm bảo sự ổn định còn có khả năng mở rộng để đáp ứng nhu cầu kinh doanh trong tương lai. Công nghệ ETL được áp dụng để trích xuất, chuyển đổi và tải dữ liệu từ nhiều nguồn khác nhau, giúp dữ liệu được tập trung và xử lý một cách chính xác, nhất quán.

## 5.2. HẠN CHẾ

### 5.2.1. Hạn chế về chức năng

Hệ thống hiện tại vẫn còn một số hạn chế trong việc triển khai các chức năng phân tích nâng cao. Chẳng hạn, các tính năng như dự báo doanh thu hay phân tích chuyên sâu về hành vi khách hàng vẫn chưa được tích hợp. Bên cạnh đó, hệ thống chưa hỗ trợ tùy chỉnh báo cáo theo nhu cầu của từng người dùng, gây hạn chế trong việc cá nhân hóa trải nghiệm và đáp ứng các yêu cầu đặc thù.

### 5.2.2. Hạn chế về kỹ thuật

Về mặt kỹ thuật, hệ thống vẫn phụ thuộc vào một số công cụ của bên thứ ba, điều này có thể tiềm ẩn rủi ro về chi phí phát sinh hoặc ảnh hưởng đến tính ổn định lâu dài. Hiệu suất xử lý dữ liệu cũng có thể bị giảm nếu lượng dữ liệu tăng trưởng đột biến mà chưa có các giải pháp tối ưu hóa kịp thời.

Ngoài ra, vấn đề bảo mật dữ liệu chưa được hoàn thiện toàn diện. Hệ thống cần được bổ sung thêm các lớp bảo mật nâng cao để bảo vệ tốt hơn thông tin nhạy cảm và đảm bảo an toàn dữ liệu trong mọi tình huống.

### 5.4. KẾT LUẬN

Dự án xây dựng kho dữ liệu phân tích kinh doanh của cửa hàng bán quần áo đã đạt được những thành tựu đáng kể, đáp ứng phần lớn yêu cầu ban đầu. Hệ thống không chỉ hỗ trợ quản lý và phân tích dữ liệu hiệu quả mà còn mở ra tiềm năng phát triển trong tương lai. Tuy nhiên, để đạt được sự hoàn thiện tối ưu, cần tiếp tục cải thiện cả về chức năng và kỹ thuật, đồng thời hướng đến việc mở rộng quy mô và nâng cao trải nghiệm người dùng.

# TÀI LIỆU THAM KHẢO

* + Các tài liệu từ internet/website:
* Tài liệu SSAS connect windowform c#
* Tài liệu Tableau connect windowform c#