

BÁO CÁO TỔNG HỢP DỰ ÁN

(H2N -Stock Dashboard Project)

Duy Duong – Data Analyst Fresher

Liên hệ: duyduong1206@gmail.com / 058 345 4703

Mục lục

BÁO CÁO TỔNG HỢP DỰ ÁN.....	1
Giới Thiệu và Cài Đặt Ban Đầu.....	2
Quản Lý Cơ Sở Dữ Liệu và Staging Database	3
Chuyển Đổi và Xử Lý Dữ Liệu	3
Mô Hình Hóa Dữ Liệu Dự Án.....	4
Xử Lý Dữ Liệu với Ngôn Ngữ DAX	5
Trực Quan Hóa Báo Cáo Chi Tiết và Tổng Thể	5
Tạo Dashboard và Tương Tác Người Dùng	6
Xuất Bản Báo Cáo và Trình Bày Dự Án.....	7
KẾT LUẬN.....	7

Báo cáo này trình bày toàn bộ quá trình thực hiện dự án mẫu, tập trung vào xây dựng, xử lý, mô hình hóa dữ liệu và trực quan hóa kết quả qua Power BI. Mục tiêu là cung cấp cái nhìn tổng thể về quy trình làm việc và kỹ năng xử lý dữ liệu.

Giới Thiệu và Cài Đặt Ban Đầu

Nội dung chính:

- **Giải thích cấu trúc dữ liệu dự án:**

Giới thiệu các cột dữ liệu như mã cổ phiếu, ngày giao dịch, giá mở cửa, giá cao nhất, giá thấp nhất, khối lượng giao dịch, thông tin nhóm ngành, thông tin công ty và các thông số liên quan trên sàn chứng khoán.

- **Cài đặt phần mềm:**

Chọn và cài đặt các phần mềm cần thiết gồm SQL Server Express Edition 2019, SSMS 19.1 và Power BI Desktop.

- **Import dữ liệu:**

Thực hiện các bước import dữ liệu mẫu **sample.bak** vào hệ thống thông qua SQL Server.

- **Các trường dữ liệu:**

- **MaCK:** Mã để xác định tính duy nhất của các cổ phiếu
- **NGAYGIAODICH:** Ngày giao dịch của sàn chứng khoán
- **GIAMOCUA:** Giá khởi điểm của ngày giao dịch
- **GIACAONHAT:** Giá cao nhất trong ngày giao dịch
- **GIATHAPNHAT:** Giá thấp nhất trong ngày giao dịch
- **KHOILUONGGIAODICH:** Tổng cổ phiếu bán ra trong ngày
- **TENNHOMNGANH:** Các nhóm ngành trên thị trường
- **MANHOMNGANH:** Mã của TENNHOMNGANH
- **THONGTINCONGTY:** Thông tin chi tiết về các công ty trên sàn chứng khoán
- **SÀN:** Sàn giao dịch chứng khoán
- **CONGTY:** Tên của công ty trên sàn CK
- **TENSAN_VIET:** Tên sàn giao dịch bằng tiếng Việt

- **TENSAN_ANH:** Tên sàn giao dịch bằng tiếng Anh
 - **BIENDODAODONG:** Tỷ lệ tăng hoặc giảm tối đa của một mã chứng khoán
 - **GIADIEUCHINHDONGCUA:** Giá đóng cửa trên sàn giao dịch
-

Quản Lý Cơ Sở Dữ Liệu và Staging Database

Nội dung chính:

- **Nhân bản cơ sở dữ liệu:**

Trình bày ý nghĩa và lợi ích của việc nhân bản (mirror database) nhằm tiết kiệm chi phí, dễ thao tác và phục hồi dữ liệu.

- **Tạo Staging Database:**

Sử dụng trình wizard và Task Copy Database để tạo staging database từ SAMPLE database, đảm bảo dữ liệu sẵn sàng cho các bước xử lý tiếp theo.

- **Làm sạch dữ liệu:**

Thực hiện các bước làm sạch dữ liệu trong Staging database để loại bỏ sai sót và dữ liệu dư thừa.

Chuyển Đổi và Xử Lý Dữ Liệu

Nội dung chính:

- **Kỹ thuật chuyển đổi dữ liệu:**

Trình bày các phương pháp như:

- Revising, Data Cleaning, Format Revision/Conversion
- Key Restructing, Deduplication, Removing unused columns
- Data Validation, Derivation, Summarization
- Splitting, Pivoting, Filtering, Merging
- Scaling/Normalization, Sorting/Indexing, Joining

- **Sửa lỗi và thêm cột mới:**

Sử dụng hàm Format để định dạng dữ liệu, tính giá sản và thêm các cột phục vụ cho phân tích nghiệp vụ.

- **Hướng dẫn tổ chức dự án:**

Sắp xếp các file dự án bao gồm file backup và tài liệu trình bày.

Ý nghĩa:

Phần này nhấn mạnh tầm quan trọng của việc chuyển đổi và xử lý dữ liệu, đảm bảo dữ liệu đầu vào sạch và nhất quán, tạo nền tảng vững chắc cho việc phân tích và báo cáo.

Mô Hình Hóa Dữ Liệu Dự Án

Nội dung chính:

- **Tạo cơ sở dữ liệu Spend:**

Sử dụng staging database để tạo Spend database qua trình wizard.

- **Kết nối Power BI với SQL Server:**

Thiết lập kết nối để truy xuất dữ liệu.

- **Chuẩn hóa dữ liệu bằng Power Query:**

Đổi tên bảng, xóa cột dư thừa, định dạng lại các trường dữ liệu và tách cột ngày, tháng, năm.

- **Mô hình hóa dữ liệu:**

Tạo các bảng phụ (TB_SAN, TB_NGANH, TB_CHUNGKHOAN) thông qua việc

duplicate bảng và giữ lại các cột cần thiết, đồng thời thực hiện grouping, đánh số thứ tự, tạo index và thiết lập các tham số kết nối.

Ý nghĩa:

Phần này chuyển đổi dữ liệu thô thành mô hình dữ liệu rõ ràng và hiệu quả, giúp tối ưu hóa mối quan hệ giữa các bảng và tạo nền tảng cho việc trực quan hóa dữ liệu.

Xử Lý Dữ Liệu với Ngôn Ngữ DAX

Nội dung chính:

- **Phân biệt Calculated Column và Measures:**
 - Calculated Column: Tính toán cho từng dòng dữ liệu.
 - Measures: Tính toán tổng hợp như SUM, AVG, COUNT, v.v.
- **Thêm các Measures:**

Tạo các cột [DATE], [NAM], [THANG], [TENTHANG], [NGAYTRONGTUAN], [QUY], [NAMQUY] từ dữ liệu ngày giao dịch và các Measure tính số lượng giao dịch, tổng khối lượng giao dịch, tổng giá trị giao dịch, v.v.
- **Thêm các Calculated Columns:**

Tạo các cột như [TENNHOMNGANH], [BIENDOGIATRONGNGAY], [GIADIEUCHINHHDONGCUA], [THANHTIEN], [TRANGTHAI], [SoLuongMaCK], [TongKhoiLuongGiaoDich].

Ý nghĩa:

Phần này khẳng định khả năng sử dụng ngôn ngữ DAX trong Power BI, giúp phân chia rõ ràng giữa các loại tính toán và cải thiện khả năng trực quan hóa dữ liệu.

Trực Quan Hóa Báo Cáo Chi Tiết và Tổng Thể

Nội dung chính:

- **Thêm visual mới (Candlestick chart):**
Hướng dẫn thêm visual hiển thị biểu đồ nến.
- **Tạo báo cáo chi tiết:**
Sử dụng bộ lọc theo ngày giao dịch, ngành, mã chứng khoán và thống kê số liệu giao dịch theo sàn, giá trị giao dịch, số mã tăng giảm.
- **Tạo báo cáo tổng thể:**
Xây dựng các báo cáo với bộ lọc nâng cao, tạo Measure đếm số lượng giao dịch, biểu đồ giá trong ngày so với giá tham chiếu và thống kê khối lượng giao dịch trên nhiều sàn.

Ý nghĩa:

Phần này mở rộng khả năng trình bày báo cáo trong Power BI, từ phân tích chi tiết đến cái nhìn tổng thể, giúp hiển thị dữ liệu một cách trực quan và dễ hiểu.

Tạo Dashboard và Tương Tác Người Dùng

Nội dung chính:

- **Xây dựng Dashboard:**
Dựa trên các biểu đồ đã tạo, xây dựng trang chi tiết và trang tổng quan với các KPI.
- **Tạo bookmark và button điều khiển:**
Thiết lập các chức năng điều hướng như button Back, ẩn/hiện slicer, hiển thị tất cả ngày và hướng dẫn sử dụng, nhằm tăng tính tương tác cho báo cáo.

Ý nghĩa:

Phần này nâng cao trải nghiệm người dùng thông qua dashboard tương tác, giúp báo cáo trở nên linh hoạt và chuyên nghiệp.

Xuất Bản Báo Cáo và Trình Bày Dự Án

Nội dung chính:

- **Xuất bản báo cáo:**
Đăng nhập Power BI Service và xuất bản báo cáo lên My Workspace.
- **Hoàn thiện tài liệu dự án:**
Bao gồm tài liệu dự án và slide báo cáo, tổng hợp toàn bộ quá trình thực hiện.

Ý nghĩa:

Phần cuối cùng khép lại toàn bộ dự án bằng cách đưa báo cáo lên môi trường trực tuyến, thể hiện sự hoàn thiện trong quá trình xây dựng, xử lý và trình bày dữ liệu.

KẾT LUẬN

Báo cáo tổng hợp này ghi nhận quá trình từ khâu cài đặt, chuẩn bị dữ liệu, xử lý và mô hình hóa, cho đến trực quan hóa và xuất bản báo cáo trên Power BI. Quá trình thực hiện thể hiện khả năng sử dụng các công cụ SQL Server, Power Query và Power BI, cũng như kỹ năng phân tích dữ liệu chuyên sâu.