|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC NHA TRANG**  **KHOA CÔNG NGHỆ THÔNG TIN**  A blue and white circle with text and a star  Description automatically generated  **CHUYÊN ĐỀ TỐT NGHIỆP**  *Ngành Hệ Thống Thông Tin Quản Lý*  **NGHIÊN CỨU TABLEAU VÀ ỨNG DỤNG CHO BÀI TOÁN PHÂN KHÚC KHÁCH HÀNG**     |  |  | | --- | --- | | **Giảng viên hướng dẫn** | **: Th.s Mai Cường Thọ** | | **Sinh viên thực hiện** | **: Nguyễn Sơn Tùng** | | **Mã số sinh viên** | **: 62132552** |   **KHÁNH HÒA - 2024** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC NHA TRANG**  **KHOA CÔNG NGHỆ THÔNG TIN**  A blue and white circle with text and a star  Description automatically generated  **CHUYÊN ĐỀ TỐT NGHIỆP**  *Ngành Hệ Thống Thông Tin Quản Lý*  **NGHIÊN CỨU TABLEAU VÀ ỨNG DỤNG CHO BÀI TOÁN PHÂN KHÚC KHÁCH HÀNG**     |  |  | | --- | --- | | **Giảng viên hướng dẫn** | **: Th.s Mai Cường Thọ** | | **Sinh viên thực hiện** | **: Nguyễn Sơn Tùng** | | **Mã số sinh viên** | **: 62132552** |   **KHÁNH HÒA - 2024** |

# **LỜI CAM ĐOAN**

“Em xin cam đoan chuyên đề tốt nghiệp **“*Nghiên cứu tableau và ứng dụng cho bài toán phân khúc khách hàng*”** là chuyên đề nghiên cứu độc lập dưới sự hướng dẫn của Thầy Mai Cường Thọ. Các nội dung trong để tài này được em thu thập từ các nguồn cung cấp dữ liệu cho nghiên cứu và học tập. Tất cả mọi thứ liên quan đến nội dung, cũng như kết quả thu được trong để tài này là trung thực và chưa được công bố ở bất kỳ đâu.

Nếu như có bất kỳ sự gian lận nào em xin hoàn toàn chịu trách nhiệm trước bộ môn, hội đồng, khoa Công nghệ thông tin về nội dung báo cáo của mình, cũng như các hình thức kỷ luật theo quy định mà Nhà trường đã đề ra.

*Khánh Hòa, ngày 03 tháng 06 năm 2024*

**Sinh viên cam kết**

**Nguyễn Sơn Tùng**

# **LỜI CẢM ƠN**

Lời đầu tiên em xin chân thành cảm ơn tất cả giảng viên đến từ Khoa Công nghệ thông tin, những người thầy, người cô là người thầm lặng suốt những năm tháng, đã chỉ bảo và cho em những bài học vô giá và kiến thức rất bổ ích và cũng cảm ơn đến các bạn lớp 62.TTQL-2.

Và quan trọng hơn em xin gửi lời cảm ơn chân thành đến giảng viên ThS. Mai Cường Thọ, là người Thầy đã hướng dẫn và tạo mọi điều kiện giúp đỡ, chỉ ra những cái sai của em trong suốt ba tháng làm chuyên đề tốt nghiệp vừa qua.

Em xin cảm ơn đến gia đình, bạn bè của em đã động viên, cổ vũ và đóng góp nhiều ý kiến trong quá trình học tập, đời sống cá nhân cũng như quá trình thực hiện chuyên đề. Do thời gian còn hạn chế và kiến thức bản thân của em còn nhiều hạn chế nên chuyên đề tốt nghiệp của em thực hiện sẽ không tránh khỏi những sai lầm thiếu sót. Rất mong nhận được những góp ý từ thầy cô giảng viên và các bạn để em có thêm nhiều kinh nghiệm trong phân tích và thiết kế để chuẩn bị thật tốt cho công việc sau này.

Em xin chân thành cảm ơn!

**MỤC LỤC**

[LỜI CAM ĐOAN 1](#_Toc167693825)

[LỜI CẢM ƠN 2](#_Toc167693826)

[DANH MỤC HÌNH ẢNH 4](#_Toc167693827)

[LỜI MỞ ĐẦU 6](#_Toc167693828)

[CHƯƠNG I: CƠ SỞ LÝ THUYẾT 7](#_Toc167693829)

[1.1. Tổng quan về BI 7](#_Toc167693830)

[1.2. Dashboard 8](#_Toc167693831)

[1.3. Tiền xử lý dữ liệu 9](#_Toc167693832)

[1.4. Trực quan hoá dữ liệu 10](#_Toc167693833)

[1.5. Phân khúc khách hàng sử dụng RFM 10](#_Toc167693834)

[1.5.1 Tìm hiểu RFM 11](#_Toc167693835)

[1.5.2. Lợi ích của RFM 11](#_Toc167693836)

[CHƯƠNG II: TỔNG QUAN VỀ CÔNG CỤ BI TABLEAU 14](#_Toc167693837)

[2.1. Tableau BI 14](#_Toc167693838)

[2.1.1. Giới thiệu về Tableau 14](#_Toc167693839)

[2.1.2. Các tính năng của Tableau 14](#_Toc167693840)

[2.1.3. Lợi ích khi sử dụng Tableau 15](#_Toc167693841)

[2.1.4. Các công cụ của Tableau 15](#_Toc167693842)

[2.2. Xử lý dữ liệu trên Tableau 16](#_Toc167693843)

[2.2.1. Nguồn dữ liệu kết nối với Tableau 16](#_Toc167693844)

[2.2.2. Xử lí dữ liệu 17](#_Toc167693845)

[2.3. Tiền xử lý dữ liệu với Tableau Prep Builder 21](#_Toc167693846)

[2.3.1. Tableau Prep Builder là gì ? 21](#_Toc167693847)

[2.3.2. Xử lý dữ liệu Tableau Prep Builder 21](#_Toc167693848)

[CHƯƠNG 3. PHÂN KHÚC THỊ TRƯỜNG KHÁCH HÀNG 25](#_Toc167693849)

[3.1. Phân khúc khách hàng sử dụng phân tích RFM 25](#_Toc167693850)

[3.1.1. Mô tả tập dữ liệu Online Retail 25](#_Toc167693851)

[3.1.2. Quy trình triển khai phân tích RFM 25](#_Toc167693852)

[3.2. Kết quả 26](#_Toc167693853)

[3.3. Tiện ích Tableau so với bảng tính thông thường 31](#_Toc167693854)

[KẾT LUẬN 33](#_Toc167693855)

# **DANH MỤC HÌNH ẢNH**

[Hình 1. 1 Các thành phần của 1 hệ thống BI 9](#_Toc167909787)

[Hình 1. 2 Các biểu đồ trực quan hoá dữ liệu 11](#_Toc167909788)

[Hình 1. 3 Biểu đồ minh hoạ nguyên tắc Pareto (nguyên tắc 80/20). 14](#_Toc167909789)

[Hình 2. 1 Các nguồn dữ liệu trong Tableau 18](#_Toc167909805)

[Hình 2. 2 Giao diện chính của Tableau 19](#_Toc167909806)

[Hình 2. 3 Kết nối dữ liệu với file Microsoft Excel 19](#_Toc167909807)

[Hình 2. 4 Không gian làm việc của Tableau 20](#_Toc167909808)

[Hình 2. 5 Bảng dữ liệu 21](#_Toc167909809)

[Hình 2. 6 Giao diện chính của Tableau Prep Builder 23](#_Toc167909810)

[Hình 2. 7 Kết nối dữ liệu với Microsoft Excel 23](#_Toc167909811)

[Hình 2. 8 Kết nối dữ liệu hiện thị trên Tableau Prep Builder 24](#_Toc167909812)

[Hình 2. 9 Các thao tác làm sạch dữ liệu 24](#_Toc167909813)

[Hình 2. 10 Các nguồn dữ liệu trong Tableau Prep Builder 26](#_Toc167909814)

[Hình 3. 1 Biểu đồ thể hiện doanh thu theo đất nước ở dạng treemaps 28](#_Toc167910143)

[Hình 3. 2 Biểu đồ thể hiện doanh thu theo sản phẩm 29](#_Toc167910144)

[Hình 3. 3 Biểu đồ thể hiện doanh thu theo trạng thái 30](#_Toc167910145)

[Hình 3. 4 Biểu đồ thể hiện phân khúc khách hàng theo đất nước 31](#_Toc167910146)

[Hình 3. 5 Biểu đồ thể hiện doanh số hàng tháng theo quốc gia 32](#_Toc167910147)

[Hình 3. 6 Biểu đồ thể hiện doanh số theo dòng sản phẩm và năm đặt hàng 33](#_Toc167910148)

[Hình 3. 7 Dashboard tổng quan báo cáo doanh thu theo vùng 34](#_Toc167910149)

[Hình 3. 8 Dashboard tổng quan về hoạt động kinh doanh toàn cầu của công ty. 34](#_Toc167910150)

[Hình 3. 9 Dashboard thể hiện hiệu suất bán hàng và phân tích kinh doanh 35](#_Toc167910151)

[Hình 3. 10 Dashboard trực quan hóa sản phẩm bán lẻ trực tuyến 35](#_Toc167910152)

[Hình 3. 11 Dashboard phân tích nhóm khách hàng theo tiêu chí RFM 36](#_Toc167910153)

# 

# **LỜI MỞ ĐẦU**

**1. Lý do chọn đề tài**

Hiện nay, do nhu cầu trực quan hoá dữ liệu ngày càng trở nên vô cùng phát triển mạnh mẽ cho các doanh nghiệp, các trực quan hoá dữ liệu không chỉ là tập hợp những con số hay chữ, thay vào đó việc sử dụng công cụ Tableau tổng hợp các dữ liệu chuyển những dữ liệu từ các dãy số thành hình ảnh, giúp người xem so sánh, tổng kết, đánh giá và đưa ra những quyết định chính xác cao, giảm thiểu tối đa sai sót khi viết bằng tay. Khi các doanh nghiệp có những báo cáo ngày càng nhiều và mất nhiều thời gian để so sánh và đánh giá. Đó cũng là lúc cần một công cụ hỗ trợ làm việc chính là Tableau được ứng dụng vào các doanh nghiệp để phân tích và trực quan dữ liệu.

Ngoài ra, cùng với sự định hướng của Thầy Mai Cường Thọ em đã chọn đề tài **“Nghiên cứu tableau và ứng dụng cho bài toán phân khúc khách hàng”** làm chuyên đề tốt nghiệp của mình.

**2. Mục tiêu chuyên đề**

Bằng công cụ Tableau phân tích số liệu kinh doanh và cập nhật tự động, có nhiều góc nhìn về tình hình kinh doanh của doanh nghiệp đến nhà quản trị.

Có một dashboard quản trị giúp cho các sếp có thể truy cập vào dashboard này mọi lúc mọi nơi, có thể so sánh, tổng kết, đánh giá và đưa ra quyết định nhanh gọn, chính xác.

**3. Cấu trúc của chuyên đề**

Bố cục của đồ án gồm có 2 chương:

- Chương 1. Cơ sở lý thuyết

- Chương 2. Tổng quan về Tableau

- Chương 3. Ứng dụng phân khúc thị trường Khách hàng

# **CHƯƠNG I: CƠ SỞ LÝ THUYẾT**

**1.1. Tổng quan về BI**

Business Intelligence (BI): là giải pháp quản trị doanh nghiệp thông minh, là một hệ thống báo cáo cho phép các tổ chức, doanh nghiệp khai thác dữ liệu từ nhiều nguồn khác nhau như về khách hàng, thị trường, nhà cung cấp, đối tác, nhân sự, … phân tích và sử dụng các dữ liệu đó thành các nguồn thông tin có ý nghĩa nhằm hỗ trợ việc ra quyết định.

Hệ thống BI gồm 6 thành phần chính (hình 1.1):

- Data Sources (nguồn dữ liệu): là cơ sở dữ liệu thô (thường là cơ sở dữ liệu quan hệ) đến từ nhiều nguồn khác nhau như các ứng dụng business, phần mềm bán hàng, website thương mại điện tử, …

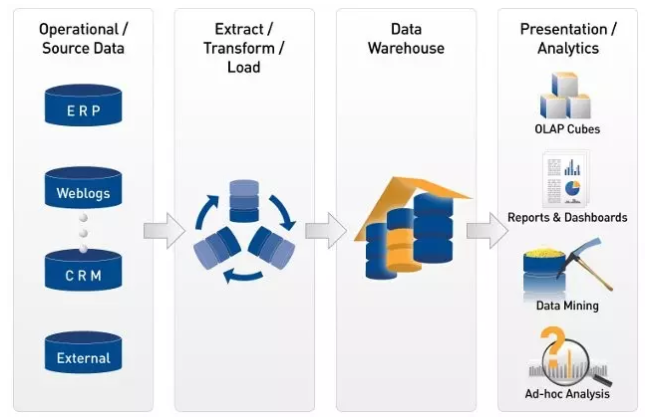
- Data Warehouse (kho dữ liệu): là cơ sở dữ liệu tích hợp, chứa dữ liệu tổng hợp của doanh nghiệp.

- Data Exploration (khám phá dữ liệu): là công việc đầu tiên khi chúng ta bắt đầu vào phân tích, tìm hiểu dữ liệu.

- Data Mining (khai phá dữ liệu): các kỹ thuật dùng để khai phá dữ liệu và phát hiện tri thức như phân loại, phân cụm, phát hiện luật kết hợp, dự báo, …

- Data Presentations & Visualization (trực quan hoá dữ liệu): tạo ra các báo cáo, biểu đồ từ quá trình data mining để phục vụ cho nhu cầu của người dùng cuối,…

- Decisions Making (ra quyết định): là quá trình bao gồm nhiều bước liên quan đến việc xác định vấn đề, hình thành, lựa chọn phương án thông qua văn bản quyết định.



Hình 1. 1 Các thành phần của 1 hệ thống BI

1.2. Dashboard

Dashboard thường được lập ra bởi các nhà phân tích kinh tế chuyên nghiệp. Trong một dashboard có rất nhiều báo cáo (Report), trong đó (Dashboard là bảng thông tin tổng hợp kết hợp nhiều báo cáo (Report) trên một màn hình hiển thị, giúp cho người xem có cái nhìn tổng quát và nhanh chóng có nhiều thông tin quản trị khác nhau.

**Báo cáo (Report)** đơn giản là những báo cáo quản trị thể hiện các số liệu ở nhiều nội dung và nhiều góc độ khác nhau tuỳ theo nhu cầu của người quản lý.

**Vậy Dashboard là gì?** Dashboard có thể hiểu được là một bảng điều khiển kỹ thuật số (digital control), hay còn được coi là một giao diện số được dùng để thu thập và tổng hợp dữ liệu của toà bộ tổ chức.

“Trong dashboard không chỉ cung cấp các dữ liệu chuyên sâu trong quá trình sản xuất kinh doanh, đồng thời dashboard còn đưa ra một cái nhìn tổng quát về năng suất cuẩ từng bộ phận, các xu hướng, các hoạt động, các chỉ số **KPI (Key Performance Indicator** – hay còn gọi là chỉ số đánh giá thực hiện công việc).”

Như vậy, Dashboard đưa ra cách nhìn tổng quan về mọi mặt trong hoạt động kinh doanh của tổ chức sao cho việc trình bày phải dễ hiểu và toàn diện chỉ trong 1 trang báo cáo.

Nhà quản lý có thể thấy được các thước đo chỉ số hoạt động của doanh nghiệp mình bằng hình ảnh, đồ thị đa chiều.

Một số ưu điểm mà Dashboard mang lại:

- Trực quan và sinh động giúp cho các nhà phân tích có thể tìm ra vấn đề một cách nhanh chóng.

- Giảm áp lực cho người trình bày khi đọc vì báo cáo chỉ tóm gọn trong một màn hình trình chiếu hoặc một trang giấy.

- Linh hoạt, dễ dàng cho phép người dùng tương tác để lựa chọn,

- Trình bày thông tin mang tính hỗ trợ đưa ra quyết định.

- Tiết kiệm thời gian lập và trình bày báo cáo vì tính tự động hoá của Dashboard qua việc tạo và lập trên những form mẫu có sẵn trên máy tính.

**1.3. Tiền xử lý dữ liệu**

Chúng ta thường bắt đầu việc xử lý dữ liệu thô, chứ không có dữ liệu chuẩn bị sẵn. Phần lớn công việc xây dựng một kho dữ liệu là trích chọn, làm sạch và chuyển đổi dữ liệu.

Dữ liệu có chất lượng cao nếu như phù hợp với mục đích sử dụng trong điều hành, ra quyết định và lập kế hoạch.

Các công việc của tiền xử lý dữ liệu:

+ Làm sạch dữ liệu (data cleaning): loại bỏ nhiễu, hiệu chỉnh những phần dữ liệu không nhất quán.

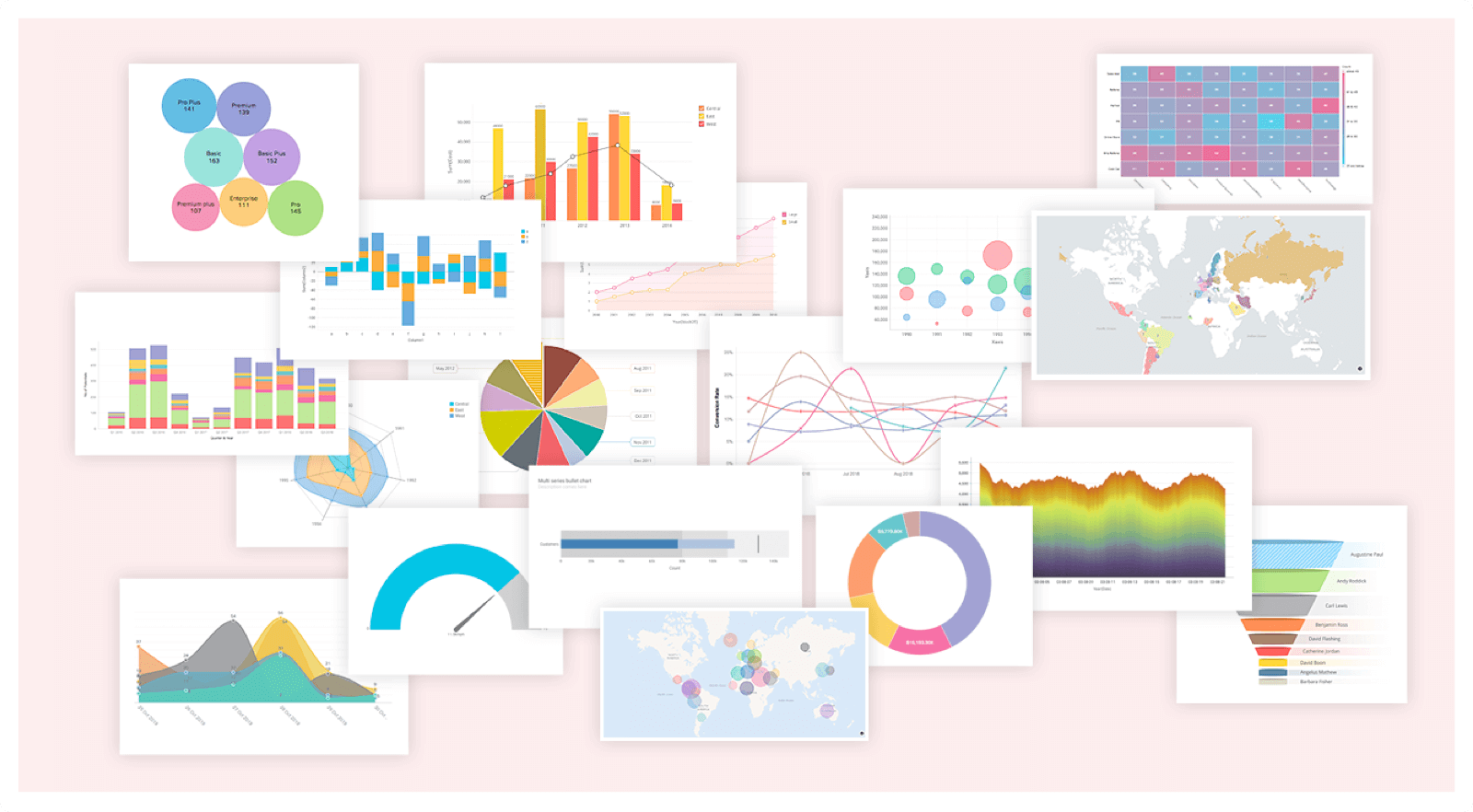
+ Tích hợp dữ liệu (data integration): trộn dữ liệu nhiều nguồn khác nhau vào một kho dữ liệu.

+ Chuyển đổi dữ liệu (data transformation): là chuẩn hoá dữ liệu (data normalization).

+ Rút gọn dữ liệu (data reduction): thu giảm kích thước dữ liệu (giảm số phần tử) bằng cách kết hợp dữ liệu (data aggregtion), loại bỏ các đặc điểm dư thừa (giảm số chiều/ thuộc tính dữ liệu), gom cụm dữ liệu.

**1.4. Trực quan hoá dữ liệu**

Là một phương thức biểu diễn trực quan của dữ liệu dưới các hình ảnh, biểu đồ, bảng đồ và đồ thị trực quan để truyền tải thông tin đến người dùng một cách dễ hiểu hơn, sinh động hơn.Sfdfs sdf sdfsdf sđfsd sé s Sfdfs sdf sdfsdf sđfsd sé s sdfé s sdfsdf sđfsd



Hình 1. 2 Các biểu đồ trực quan hoá dữ liệu

**1.5. Phân khúc khách hàng sử dụng RFM**

Từ những khách hàng tiêu dùng trung thành cho đến những khách hàng rời bỏ công ty, tất cả những khách hàng đều có nhu cầu và mong muốn đa dạng. Công ty muốn khách hàng chi tiêu nhiều hơn từ những chiến dịch tiếp thị chương trình, sản phẩm mới tới khách hàng theo những cách khác nhau. Tuy nhiên, câu hỏi đặt ra là làm thế nào để đưa ra được các chiến dịch tiếp thị phù hợp với những nhóm khách hàng đang có nhu cầu để từ đó tăng tỷ lệ phản hồi từ khách hàng và từ đó tăng doanh số bán hàng. Bài toán đặt ra là làm thế nào để có thể phân khúc khách hàng một cách tương đối chính xác dựa trên hành vi giao dịch lịch sử của khách hàng, thuật toán RFM sẽ giúp chúng ta giải quyết vấn đề này một cách nhanh chóng và hiệu quả.

### **1.5.1 Tìm hiểu RFM**

Mô hình này được gọi là RFM, là một kỹ thuật tiếp thị được sử dụng để đánh giá hành vi của khách hàng của công ty. Ví dụ: nó có thể được sử dụng để đánh giá thời gian gần đây khách hàng đã mua sản phẩm hoặc dịch vụ (lần truy cập gần đây), số lần khách hàng đã mua hàng kể từ một ngày (tần suất) nhất định và khách hàng đã chi bao nhiêu trong một ngày nhất định một khoảng thời gian (tiền tệ).

- Lần mua hàng gần đây – R (Recency).

- Tần suất mua hàng của khách hàng – F (Frequency).

- Tiền, giá trị, khách hàng chi bao nhiêu tiền – M (Monetary value).

RFM là một mô hình rất phổ biến vì các lý do:

- Nó sử dụng thang đo một cách khách quan.

- Doanh nghiệp sử dụng mô hình này một cách hiệu quả và dễ dàng vì đây là mô hình đơn giản.

- Trực quan - đầu ra của phương pháp phân khúc này là dễ hiểu và giải thích.

### **1.5.2. Lợi ích của RFM**

*a. Hiểu về khách hàng của bạn*

Nhờ vào mô hình RFM, các công ty hiểu hơn về khách hàng của họ: phân loại được nhóm khách hàng trung thành của họ, những khách hàng không hài lòng với dịch vụ của họ, khách hàng cũ đã rời bỏ họ và từ đó hiểu hơn về sản phẩm của họ.

*b. Tăng doanh số của bạn*

Bằng cách xem xét khách hàng thông qua ba khía cạnh chính của mô hình phân tích RFM, doanh nghiệp của chúng ta có thể cải thiện đáng kể các kỹ thuật tiếp thị của mình và nhắm mục tiêu đưa ra các chiến dịch của doanh nghiệp mình một cách chính xác hơn theo mong muốn và nhu cầu của khách hàng, điều này cuối cùng sẽ tăng tỷ lệ phản hồi và doanh số bán hàng.

*c. Giảm đáng kể lượng khách hàng rời bỏ công ty/ doanh nghiệp*

Với phân tích RFM, doanh nghiệp có thể phân loại khách hàng của mình thành một số danh mục. Số lượng danh mục phụ thuộc vào bản chất của doanh nghiệp và bản năng của nhà phân tích. Từ kết quả phân tích doanh nghiệp có thể tìm ra nhóm khách hàng trung thành cũng như nhóm khách hàng có khả năng rời bỏ công ty hoặc đã rời bỏ công ty và để lấy lại được lượng khách hàng rời bỏ công ty thì công ty hay doanh nghiệp cần có các hành động khác nhau để lấy lại khách hàng. Ví dụ: Công ty hoặc doanh nghiệp có thể gửi e-mail cá nhân cho khách hàng đến những người có nguy cơ rời đi để kết nối lại với doanh nghiệp.

*d. Quy tắc 80/20*

Đây là quy tắc ngón tay cái tuyên bố rằng 80% kết quả có thể được quy cho 20% nguyên nhân cho một sự kiện nhất định. Thông thường, một doanh nghiệp sử dụng phương pháp này để chỉ ra rằng 80% doanh thu được tạo ra bởi 20% tổng số khách hàng của mình. Cách tiếp cận này giúp những người ra quyết định tập trung vào các yếu tố quan trọng nhất ảnh hưởng đến tài nguyên của họ. Ví dụ (hình 1.3):

Chart

Description automatically generated

Hình 1. 3 Biểu đồ minh hoạ nguyên tắc Pareto (nguyên tắc 80/20).

Nếu doanh nghiệp có thể xác định những người chi tiêu hàng đầu của mình, doanh nghiệp có thể đưa ra các chiến lược tiếp thị phù hợp để giữ chân khách hàng hiện tại hoặc có được những khách hàng có hồ sơ tương tự.

Tuy nhiên so với RFM thì quy tắc 80/20 vẫn còn hạn chế vì nó chỉ nhìn vào tần suất chi tiêu, và không giải quyết được sự suy thoái.

**CHƯƠNG II: TỔNG QUAN VỀ CÔNG CỤ BI TABLEAU**

**2.1. Tableau BI**

### **2.1.1. Giới thiệu về Tableau**

Tableau là một phần mền được tạo bởi một công ty cùng tên ở Washington và được thành lập vào năm 2003. Tableau là công cụ phần mềm hỗ trợ phân tích và trực quan hoá dữ liệu (Data Visualization), được dùng nhiều trong lĩnh vực BI (Business Intelligence). Cũng giống như Excel, Tableau giúp tổng hợp các dữ liệu nhưng ở một cấp độ cao hơn khi chuyển những dữ liệu này từ các dãy số văn bản thành những hình ảnh, biểu đồ trực quan xây dựng các dashboard và phân tích.

Tableau thiết kế cho tất cả mọi người sử dụng một cách dễ dàng, không yêu cầu mọi người phải giỏi về lập trình. Vì vậy, bất cứ một ai cũng có thể sử dụng Tableau cho những vấn đề họ cần tìm hiểu về các doanh nghiệp mà họ muốn với các thao tác kéo và thả chuột linh hoạt. Giao diện Tableau dễ dàng sử dụng mà bất cứ người dùng nào thành thạo Excel thì có thể tìm hiểu và sử dụng các thao tác này bằng chuột một cách linh hoạt.

### **2.1.2. Các tính năng của Tableau**

Khi sử dụng Tableau người dùng có thể thực hiện một số các chức năng :

- Mô phỏng và phân tích dữ liệu.

- Phân tích dữ liệu với Big Data.

- Phân tích theo thời gian.

- Hỗ trợ tạo các truy vấn bằng thao tác đơn giản.

- Nhập dữ liệu với kích thước lớn, quản lý siêu dữ liệu.

- Kéo thả để xây dựng các Business Dashboard từ đơn giản đến phức tạp.

- Kết nối và chia sẻ thông qua các ứng dụng trực tuyến thời gian thực.

- Biên dịch các truy vấn thành những hình ảnh, biểu đồ.

- Quản lý các dữ liệu hoạt động (Data Stories).

### **2.1.3. Lợi ích khi sử dụng Tableau**

Sử dụng BI Tableau sẽ có các lợi ích :

- Tổng hợp và phân tích dữ liệu từ nhiều nguồn, khả năng tích hợp dữ liệu đa dạng và đưa ra các góc nhìn về hoạt động doanh nghiệp.

- Các nhà doanh nghiệp, quản trị luôn có những chuỗi dữ liệu thời gian giúp họ nhận ra các sự khác biệt và các phương án ứng xử đó. Đồng thời giúp họ đưa ra nền tảng và quyết định chính xác và đúng thời điểm.

- Tableau xây dựng một môi trường làm việc dựa trên nền tảng số liệu, trên nền tảng phân tích, tạo nên một sực mạnh tập thể cho doanh nghiệp.

- Ai cũng có thể xem và phân tích và các số liệu của mình quản lý, làm cho doanh nghiệp của người dùng vận hành trong môi trường ưu tú nhất và hiệu quả nhất.

- Tableau được triển khai nhanh, khả năng tiếp cận nhanh, giảm tối đa chi phí và dễ dàng mở rộng theo một mô hình.

### **2.1.4. Các công cụ của Tableau**

Tableau hỗ trợ nhiều công cụ cho người dùng lựa chọn:

**- Tableau Prep :** là công cụ được dùng để chuẩn bị dữ liệu, trích xuất và xử lý dữ liệu (ETL). Mang đến sự thay đổi quan trọng trong việc tổ chức dữ liệu, so với phương pháp truyền thống có nhiều cải tiến. Tableau Prep giúp mọi người dùng doanh nghiệp và nhà phân tích định hình dữ liệu nhanh chóng, cho phép thực hiện các truy vấn, kết hợp và làm sạch dữ liệu cực kì đơn giản và tiện lợi. Sử dụng Tableau Prep giúp dữ liệu có tổ chức, rõ ràng, dễ quản lý hơn. Hiện nay có hai công cụ là Tableau Prep Builder để xây dựng luồng dữ liệu và Tableau Prep Conductor để quản lý các luồng.

**- Tableau Desktop** cung cấp giao diện trực quan cùng các tính năng đa dạng mã hoá và phân tích dữ liệu. Phần lớn là các thao tác kéo thả chuột và cũng không yêu cầu quá nhiều việc về mặt kỹ thuật hay lập trình. Tableau Desktop có khả năng kết nối rộng rãi đến nhiều định dạng file khác nhau, để đáp ứng tốt nhất về nhu cầu phân tích trong nhiều ngành nghề, lĩnh vực.

**- Tableau Online :** là nền tảng phân tích của bạn được lưu trữ hoàn toàn trên đám mây. Xuất bản trang tổng quan và chia sẻ khám phá của bạn với bất kỳ ai. Mời đồng nghiệp hoặc khách hàng tìm hiểu và khám phá các cơ hội tiểm ẩn bằng hình ảnh tương tác và dữ liệu chính xác. Tất cả đều có thể truy cập dễ dàng từ trình duyệt hoặc khi đang di chuyển bằng các ứng dụng dành cho điện thoại di động. Phần lớn sử dụng Tableau online chủ yếu sử dụng trong việc học tập.

**- Tableau Server :** mở rộng giá trị dữ liệu trên toàn bộ tổ chức với máy chủ tableau. Là nơi chia sẻ các báo cáo phân tích của doanh nghiệp, những dữ liệu trên đây được bảo mật cẩn thận và cấp quyền truy cập.

**2.2. Xử lý dữ liệu trên Tableau**

### **2.2.1. Nguồn dữ liệu kết nối với Tableau**

Tableau hỗ trợ kết nối với nhiều loại dữ liệu, được lưu trữ ở nhiều nơi khác nhau. Cung cấp các trình tự kết nối gốc được xây dựng và tối ưu hoá cho các loại dữ liệu đó. Dưới đây là một số loại nguồn dữ liệu kết nối đến Tableau :

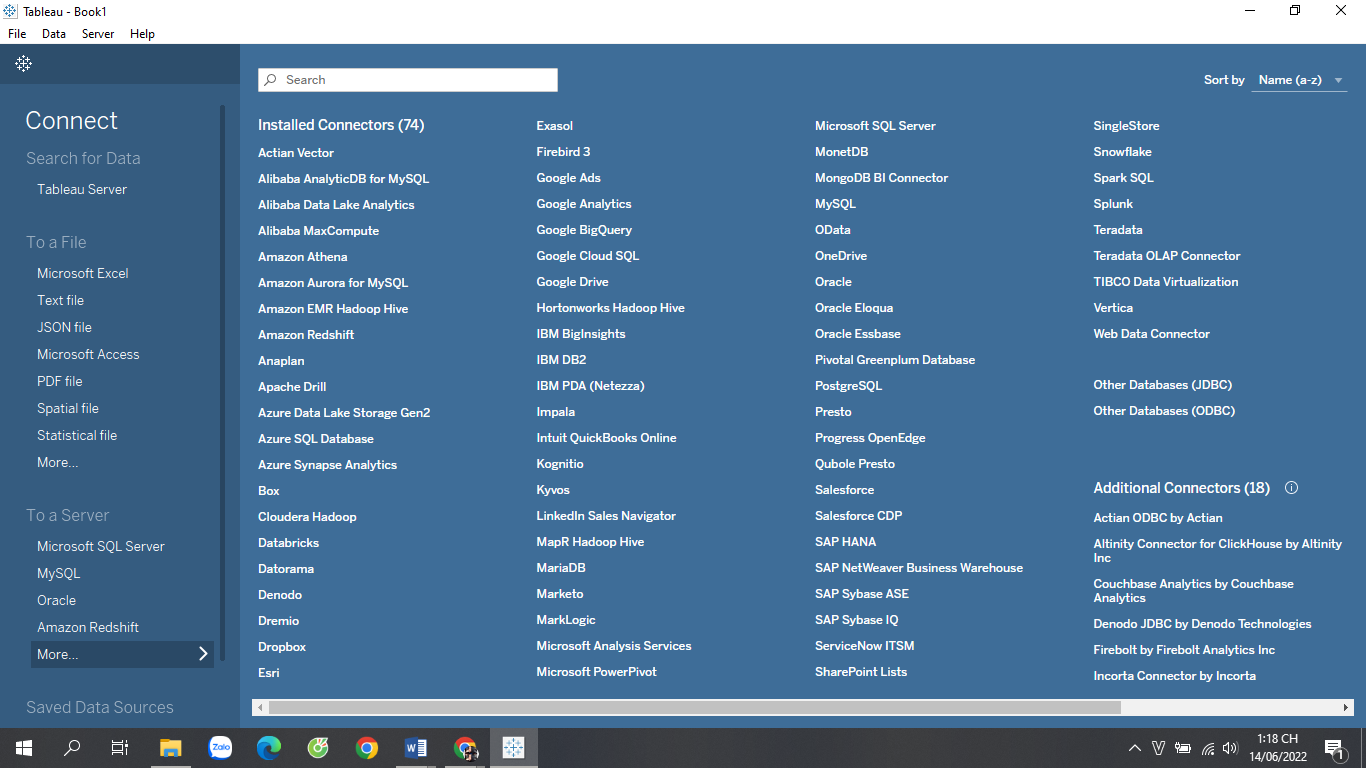
- Có thể kết nối với cơ sở dữ liệu và bảng từ Tableau Server trên ngăn Connect trong Tableau Desktop.

- Tìm dữ liệu bằng máy chủ tableau hoặc tableau online.

- Các tệp với cơ sở dữ liệu được hỗ trợ thì Tableau cung cấp các trình tự kết nối gốc được xây dựng và tối ưu hoá cho các loại dữ liệu.

- Các kết nối riêng : Other Database (JDBC), Other Databases (ODBC), Web Data Connector hoặc Connector Plugin được xây dựng bằng Tableau Connector SDK.

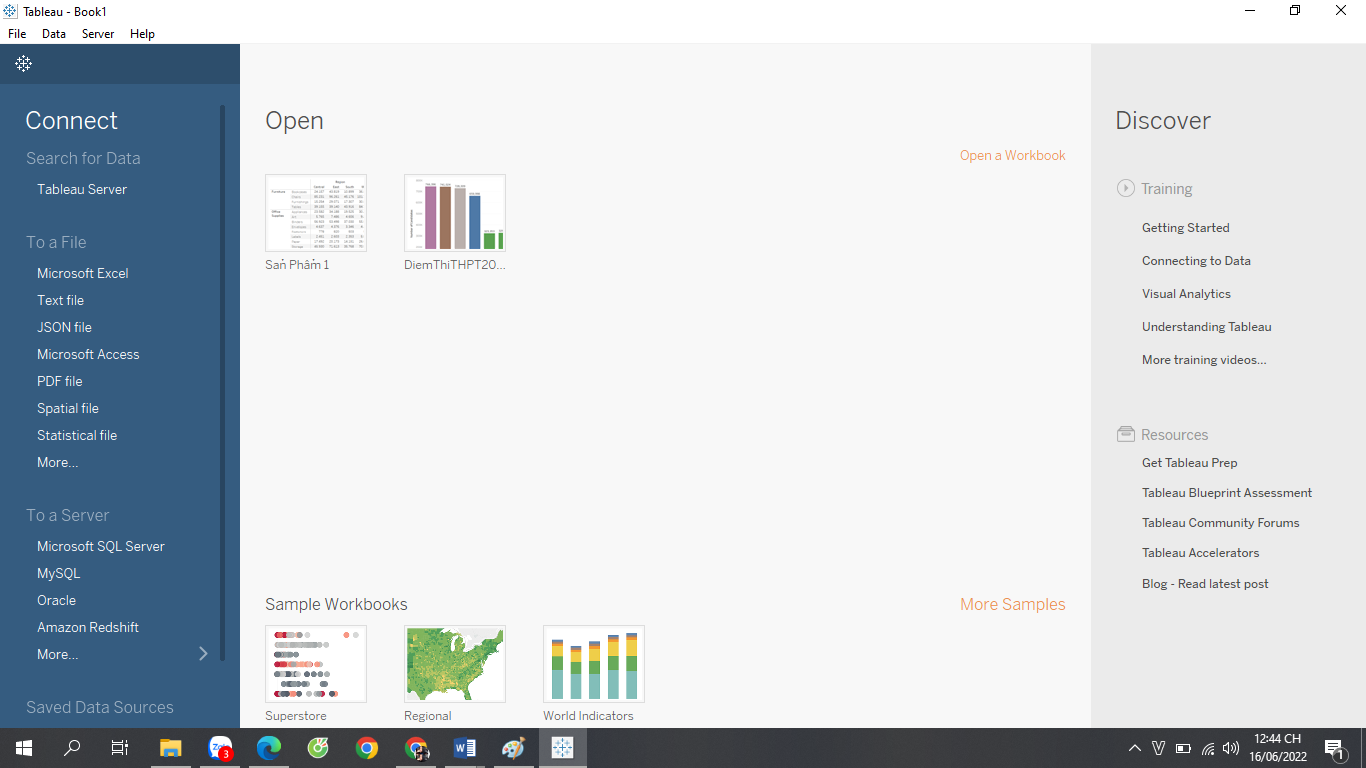
- Các tập tin, tải file Excel hoặc các nguồn dữ liệu dựa trên văn bản (.xlsx, .csv, .tsv). Truy cập và kết hợp dữ liệu khác nhau mà không cần viết mã.

- Kết nối với dữ liệu trước hoặc trên đám mây cho dù dữ liệu đó có lớn, cơ sở dữ liệu SQL, bảng tính hay các ứng dụng đám mây như Google Analytics và Salesforcesd fsdsdfsdsdf sdf sdf sd fsdsdfdsdf sdf sdf sd fsdsdf sdf sdf sd fsdsdf Hình 2. 1 Các nguồn dữ liệu trong Tableau

### **2.2.2. Xử lí dữ liệu**

**2.2.2.1. Kết nối dữ liệu**

Dưới đây chính là giao diện đầu tiên của Tableau Desktop.



Hình 2. 2 Giao diện chính của Tableau

- Sau khi ta mở Tableau Desktop lên thì giao diện đầu tiên người dùng sẽ thấy được là như hình trên. Ở đây, người dùng sẽ chọn trình kết nối dữ liệu mà người dùng muốn sử dụng.

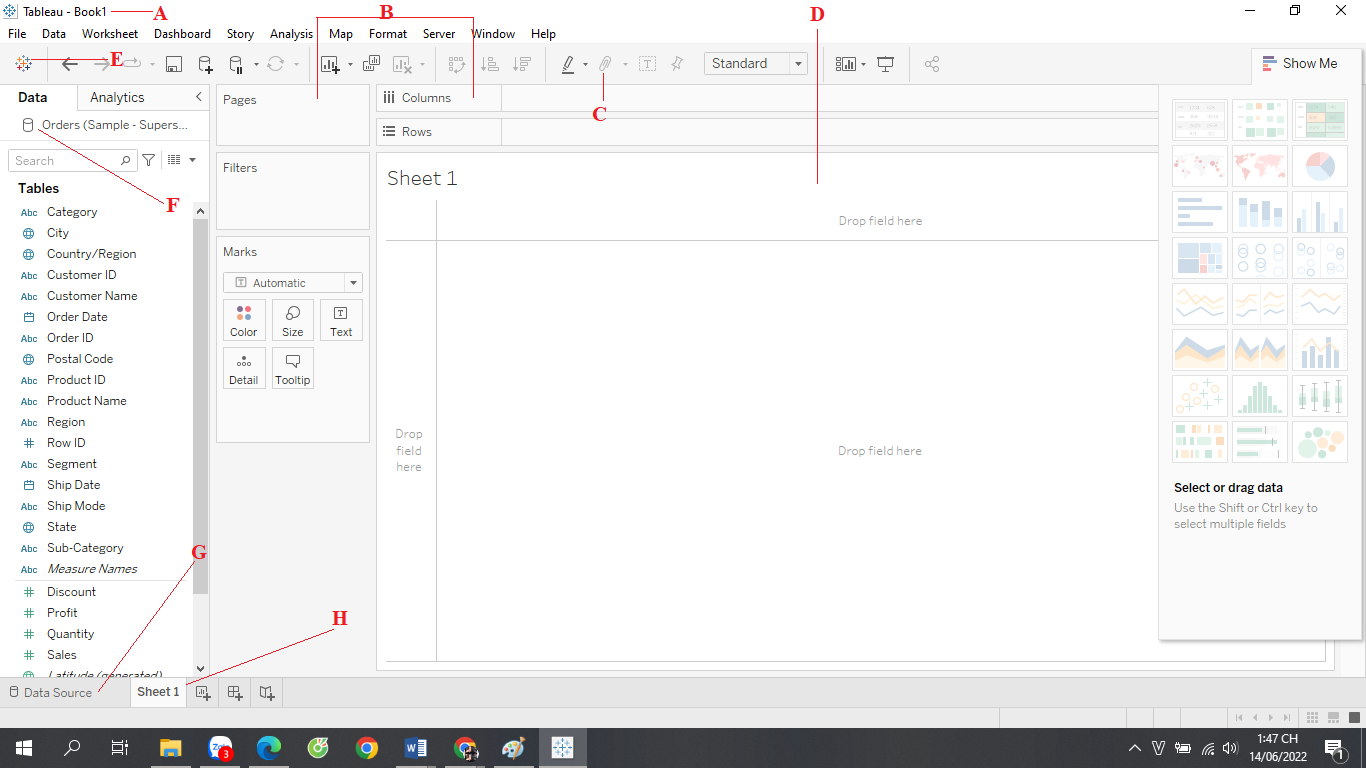
- Bên tay trái màn hình giao diện có các loại dữ liệu để connect. Chọn file excel và nhấn Open.



Hình 2. 3 Kết nối dữ liệu với file Microsoft Excel

- Hoặc kết nối với nhiều loại dữ liệu khác có sẵn.

- Nhấp vào tab Sheet 1 để chuyển đến Worksheet mới và bắt đầu phân tích.



Hình 2. 4 Không gian làm việc của Tableau

A. Tên cửa sổ làm việc. Một cửa sổ làm việc chứa các trang tính. Trang tính có thể là worksheet, dashboard, story.

B. Cards and shelves – Kéo các trường vào thẻ và giá trong không gian làm việc để thêm dữ liệu vào chế độ xem.

C. Toolbar – Sử dụng thanh công cụ để truy cập các lệnh và các công cụ phân tích và điều hướng,

D. View – Là không gian làm việc, nơi bạn tạo ra hình ảnh trực quan.

E. Nhấp vào biểu tượng này là chuyển đến trang Start Page, nơi bạn có thể kết nối với dữ liệu.

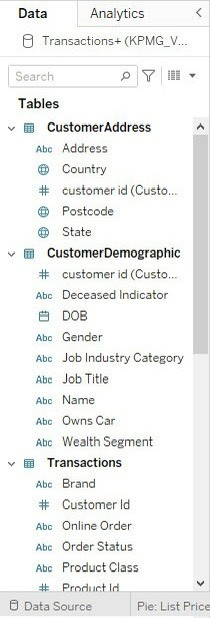
F. Side Bar – Trong một trang tính, khu vực thanh bên chứa ngăn Data pane và Analytics pane, Dashboard và Story.

G. Nhấp vào tab này sẽ chuyển đến trang Data Source Page và xem dữ liệu.

H. Sheet tabs – Các tab đại diện cho từng trang tính trong sổ làm việc. Điều này có thể bao gồm Worksheet, Dashboard, Story.

**2.2.2.2. Giới thiệu Data**

Trong 1 trang tính, các cột từ nguồn dữ liệu của bạn được hiển thị dưới dạng các trường bên trái trong ngăn Data. Ngăn Data chứa nhiều trường được sắp xếp theo bảng.

**

Hình 2. 5 Bảng dữ liệu

**2.2.2.3. Các cách để bắt đầu xây dựng View**

View là một hình ảnh trực quan mà bạn đã tạo trong Tableau. Nó có thể là một biểu đồ, một đồ thị, một bản đồ hoặc một bảng văn bản.

Mọi chế độ xem xây dựng trong Tableau nên bắt đầu bằng một câu hỏi. Bạn muốn biết gì ?

- Mỗi khi kéo một trường vào chế độ xem là đang đặt câu hỏi về dữ liệu. Câu hỏi sẽ khác nhau tuỳ thuộc vào vị trí kéo các trường khác nhau, loại trường và thứ tự kéo các trường vào dạng xem.

- Đối với mọi câu hỏi, chế độ xem thay đổi để thể hiện câu trả lời một cách trực quan với các dấu (hình dạng, văn bản, phân cấp, cấu trúc bảng, trục, màu).

- Khi tạo một dạng xem thì bạn thêm các trường từ ngăn Data. Có thể làm điểu này theo nhiều cách khác nhau.

- Kéo các trường từ ngăn Data và thả chúng vào Columns, Rows là một cách.

- Hoặc Click vào một hoặc nhiều trường trong ngăn Data.

- Để hoàn tác hoặc làm lại thì bạn nhấp vào biểu tượng  hoặc  trên thanh công cụ.

**2.2.2.4. Xây dựng View từ đầu**

Các bước này chỉ ra cách tạo chế độ xem cơ bản hiển thị lợi nhuận qua từng năm.

- Từ ngăn Data, Kéo Order Date vào Columns.

- Bảng kết quả có bốn cột và một hàng. Mỗi tiêu đề cột đại diện cho một thành viên của trường Order Date (cấp ngày mặc định là YEAR). Mỗi ô chứa một nhãn “Abc”, cho biết rằng loại nhãn hiện tại cho chế độ xem này là văn bản.

Trường có màu xanh lam, điều này cho thấy rằng nó là rời rạc.

- Để ẩn tiêu đề trên trang tính, chọn Menu thả xuống ở bên phải của tiêu đề (Sheet 1) và HideTitle.

- Từ ngăn Data, kéo trường Profit vào Rows.

- Trường Profit có màu xanh lục trên Rows, cho biết rằng trường này liên tục. Ngoài ra, tên trường thay đổi thành SUM (Profit). Đó là vì các số đo được tổng hợp tự động khi thêm chúng vào chế độ xem và tổng hợp mặc định cho số đo này là SUM.

Tableau biến bảng thành biểu đồ đường và tạo trục tung (dọc theo bên trái) cho số đo.

* 1. **Tiền xử lý dữ liệu với Tableau Prep Builder**

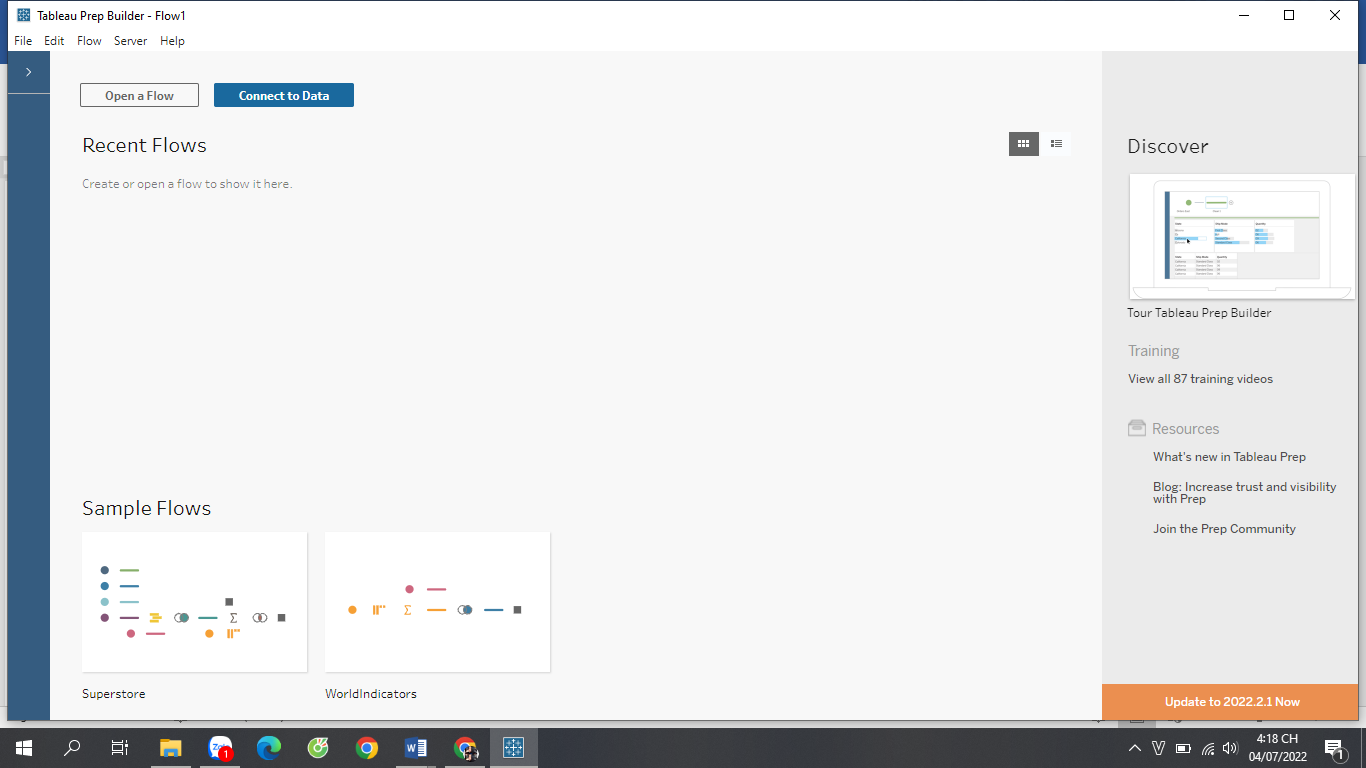
### **2.3.1. Tableau Prep Builder là gì ?**

Tableau Prep Builder là một tool trong Tableau được thiết kế để giúp chuẩn bị dữ liệu và trực quan một cách dễ dàng. Sử dụng Tableau Prep Builder để kết hợp, định hình và làm sạch dữ liệu để phân tích trong Tableau.

### **2.3.2. Xử lý dữ liệu Tableau Prep Builder**

**2.3.2.1. Kết nối dữ liệu**

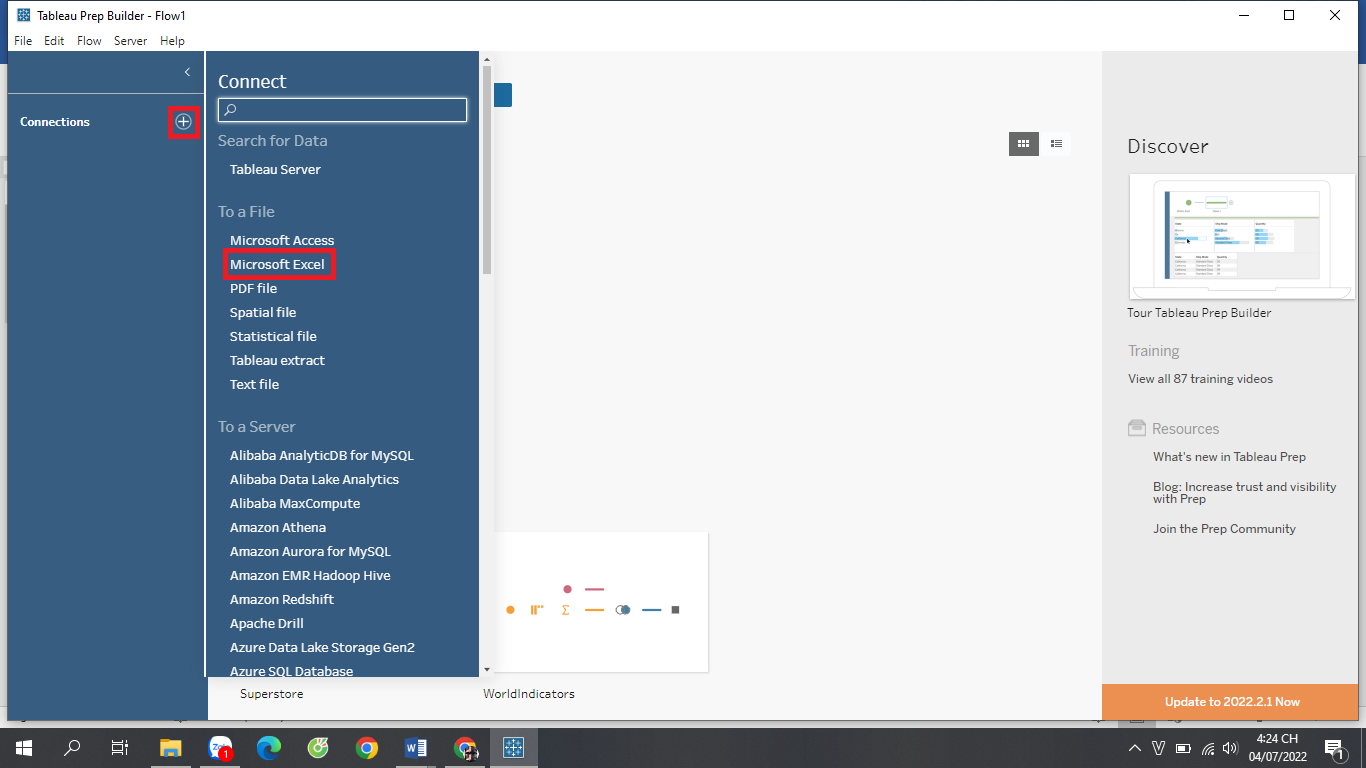
Dưới đây chính là giao diện đầu tiên của Tableau Prep Builder.



Hình 2. 6 Giao diện chính của Tableau Prep Builder

- Sau khi ta mở Tableau Prep Builder lên thì giao diện đầu tiên người dùng sẽ thấy được là như hình trên. Ở đây, người dùng sẽ chọn trình kết nối dữ liệu mà người dùng muốn sử dụng.

- Bên tay trái màn hình giao diện connect có dấu  ấn vào để hiển thị dữ liệu chúng ta muốn chọn. Chọn file excel và nhấn Open.



Hình 2. 7 Kết nối dữ liệu với Microsoft Excel

- Hoặc kết nối với nhiều loại dữ liệu khác có sẵn.

**2.3.2.2. Hình dạng và kết nối dữ liệu**

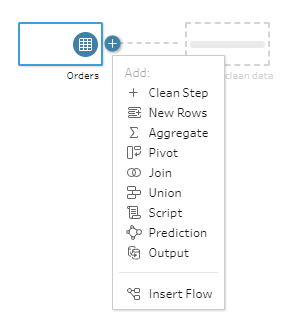
- Sau khi kết nối với dữ liệu Excel, kéo và thả hoặc double – click chuột đưa các bảng dữ liệu vào khung Flow.

A screenshot of a computer

Description automatically generated

Hình 2. 8 Kết nối dữ liệu hiện thị trên Tableau Prep Builder

- Sử dụng các thao tác quen thuộc như split, rename, filter, pivot, union và join để làm sạch dữ liệu và định hình dữ liệu mà chúng ta mong muốn.



Hình 2. 9 Các thao tác làm sạch dữ liệu

- Mỗi bước trong quy trình được thể hiện trực quan trong biểu đồ flow mà chúng ta đã tạo và kiểm soát. Tableau Prep Builder theo dõi từng thao tác để kiểm tra công việc và thực hiện thay đổi tại bất kỳ điểm nào trong flow.

- Khi hoàn thành flow, hãy chọn Run Flow để áp dụng các thao tác mà chúng ta đã thực hiện cho toàn bộ dữ liệu.

- Không gian làm việc của Tableau Prep Builder bao gồm:

+ Connections (A): Hiện thị cơ sở dữ liệu và files đang kết nối.

+ Flow (B): Thể hiện trực quan các bước thao tác, chuẩn bị dữ liệu.

+ Profile (C): Tóm tắt từng field trong data sample. Xem hình dạng dữ liệu và nhanh chóng tìm ra các null và outlier (ngoại lệ).

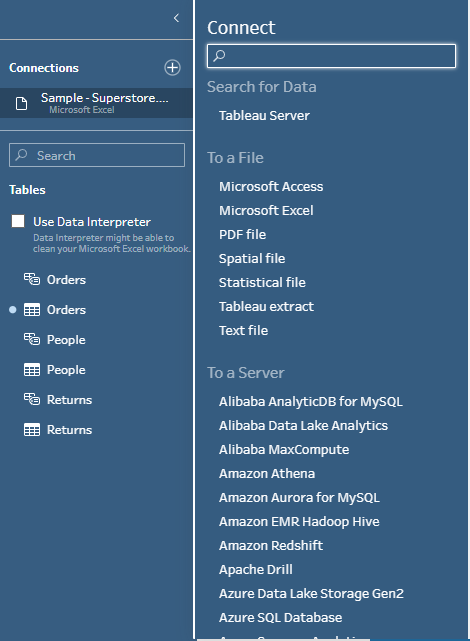
+ Data grid (D): Chi tiết row level cho dữ liệu.

**2.3.2.3. Xây dựng Flow từ đầu**

Các bước này chỉ ra cách tạo chế độ xem cơ bản:

+ Bước 1: Mở Tableau Prep Builder và click vào Add connections.

+ Bước 2: Từ list của trình kết nối, chọn file hoặc sever lưu trữ data.

Dsgsd sd fsdsdf 

Hình 2. 10 Các nguồn dữ liệu trong Tableau Prep Builder

+ Bước 3: Từ Connections, thực hiện một trong các thao tác sau:

\* Nếu đã kết nối với file, double-click chuột hoặc kéo một bảng vào khung Flow để bắt đầu thực hiện làm sạch dữ liệu. Đối với các bảng đơn, Tableau Prep Builder sẽ tự động tạo một bước Input trong Flow khi thêm data vào Flow.

\* Nếu kết nối với database, hãy chọn một database hoặc schema, sau đó double-click hoặc kéo bảng vào Flow để bắt đầu thực hiện flow.

CHƯƠNG 3. PHÂN KHÚC THỊ TRƯỜNG KHÁCH HÀNG

## **3.1. Phân khúc khách hàng sử dụng phân tích RFM**

### **3.1.1. Mô tả tập dữ liệu Online Retail**

Đây là tập dữ liệu xuyên quốc gia chứa tất cả các giao dịch xảy ra từ ngày 01/06/2003 đến ngày 31/05/2005 cho một cửa hàng bán lẻ trực tuyến không có cửa hàng và có trụ sở tại Vương quốc Anh. Công ty chủ yếu bán quà tặng nhân dịp độc đáo. Nhiều khách hàng của công ty là những người buôn bán. Tập dữ liệu này gồm có hơn 500 dòng và 8 cột:

- InvoiceNo: Số hóa đơn. Danh nghĩa, một số tích phân gồm 6 chữ số được chỉ định duy nhất cho mỗi giao dịch. Nếu mã này bắt đầu bằng chữ cái 'c', nó chỉ ra một sự hủy bỏ.

- StockCode: Mã sản phẩm (mặt hàng). Danh nghĩa, một số tích phân gồm 5 chữ số được gán duy nhất cho mỗi tích riêng biệt.

- Description: Tên sản phẩm (mặt hàng). Trên danh nghĩa.

- Quantity: Số lượng của mỗi sản phẩm (mặt hàng) trên mỗi giao dịch. Số.

- InvoiceDate: Ngày và giờ làm việc. Số, ngày và giờ khi mỗi giao dịch được tạo.

- UnitPrice: Đơn giá. Số, Giá sản phẩm trên mỗi đơn vị bằng đồng bảng Anh.

- CustomerID: Số khách hàng. Danh nghĩa, một số nguyên gồm 5 chữ số được chỉ định duy nhất cho mỗi khách hàng.

- Country: Tên quốc gia. Trên danh nghĩa, tên quốc gia nơi mỗi khách hàng cư trú.

**3.1.2. Quy trình triển khai phân tích RFM**

Quy trình triển khai phân tích RFM được tiến hành thông qua 4 bước:

- Bước 1: Xác định nguồn dữ liệu phân tích

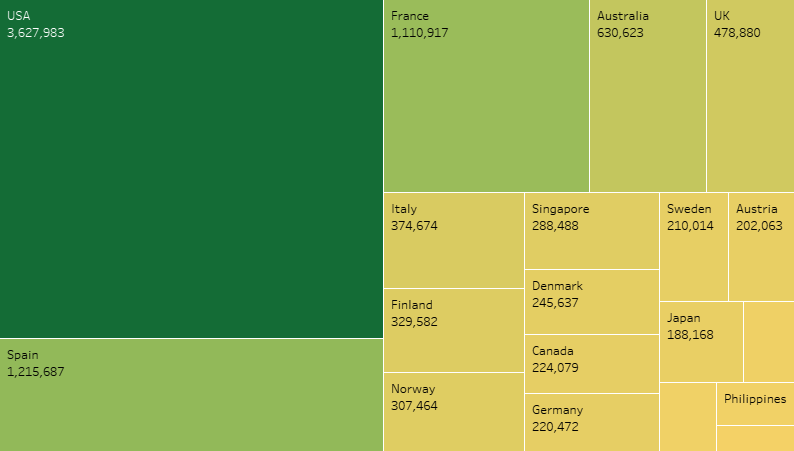
- Bước 2: Tính toán các giá trị R, F và M

- Bước 3: Nhóm các khách hàng vào các nhóm theo R, F và M

- Bước 4: Phân tích kết quả và đề xuất chiến lược để chăm sóc khách hàng phù hợp với từng nhóm khách hàng.

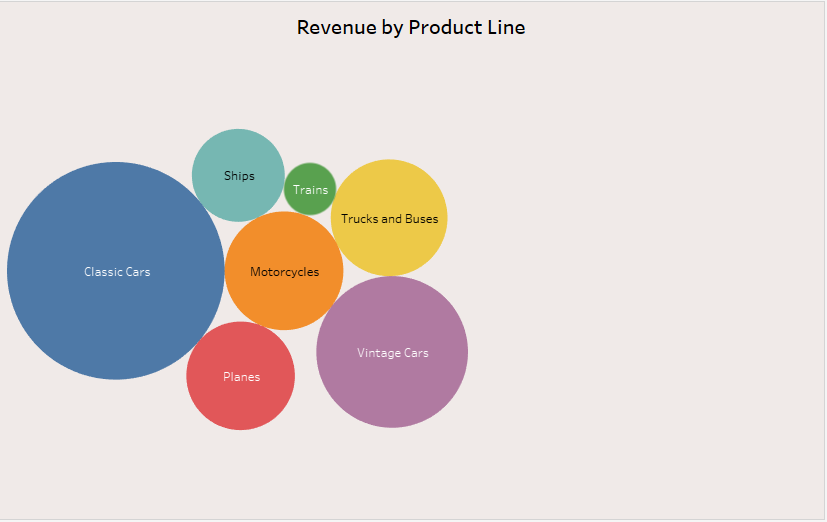
**3.2. Kết quả**

- Biểu đồ doanh thu theo đất nước ở dạng treemaps, sẽ thấy được doanh thu cụ thể của từng đất nước cùng với doanh thu đóng góp.



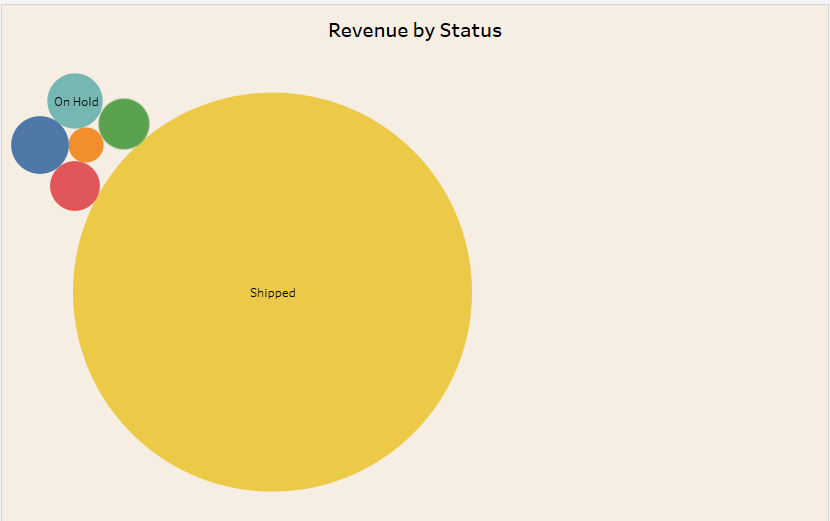
Hình 3. 1 Biểu đồ thể hiện doanh thu theo đất nước ở dạng treemaps

- Biểu đồ doanh thu theo dòng sản phẩm ở dạng packed bubble



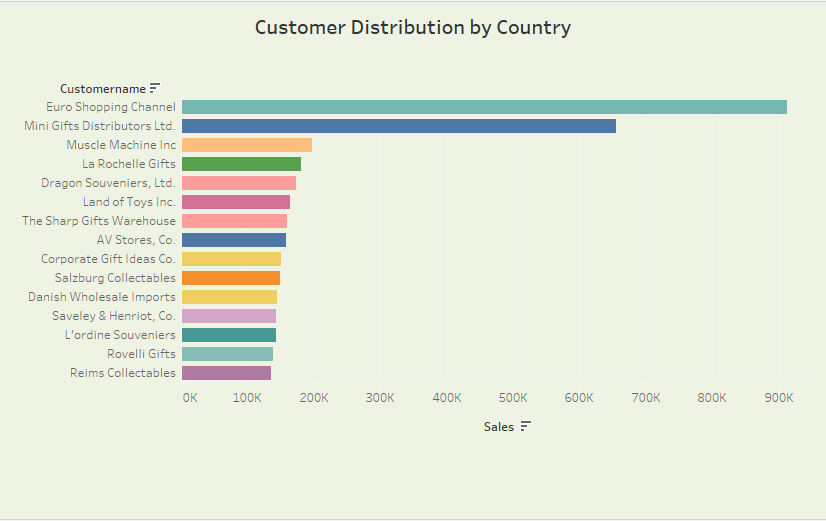
Hình 3. 2 Biểu đồ thể hiện doanh thu theo sản phẩm

- Biểu đồ doanh thu theo trạng thái ở dạng packed bubble



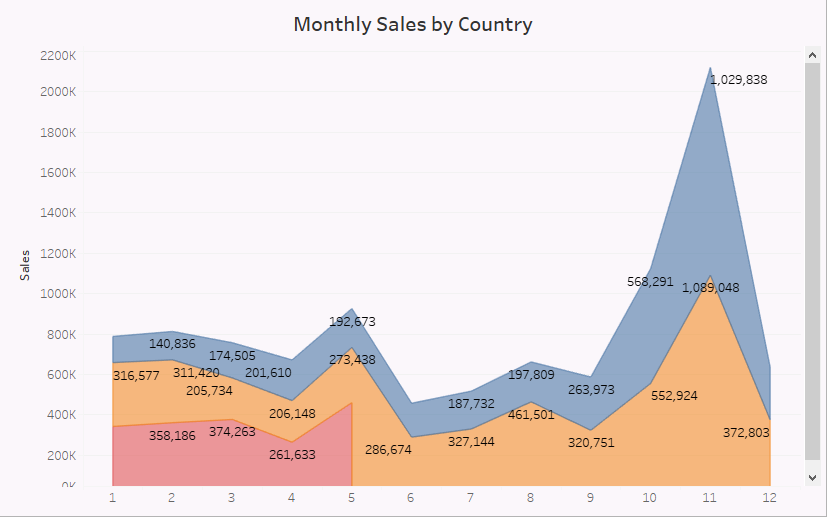
Hình 3. 3 Biểu đồ thể hiện doanh thu theo trạng thái

- Biểu đồ phân khúc khách hàng theo đất nước ở dạng horizontal bars

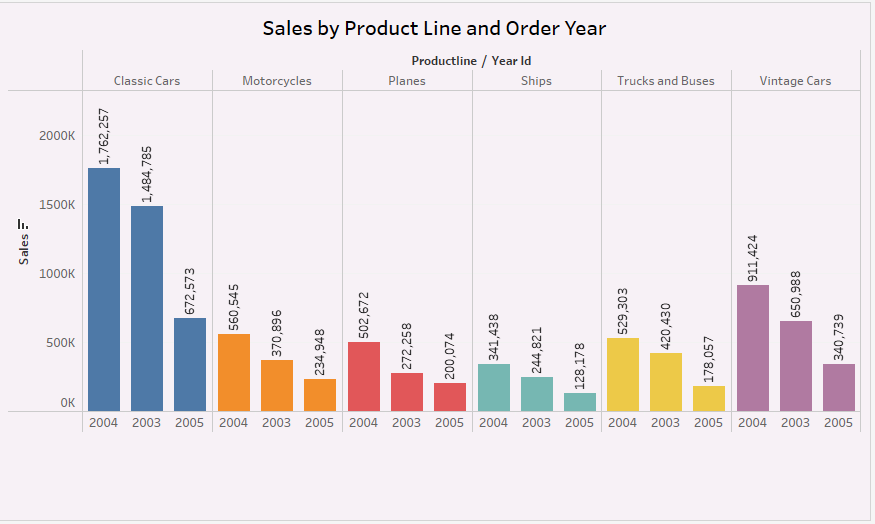


Hình 3. 4 Biểu đồ thể hiện phân khúc khách hàng theo đất nước

- Biểu đồ doanh số hàng tháng theo quốc gia ở dạng area charts



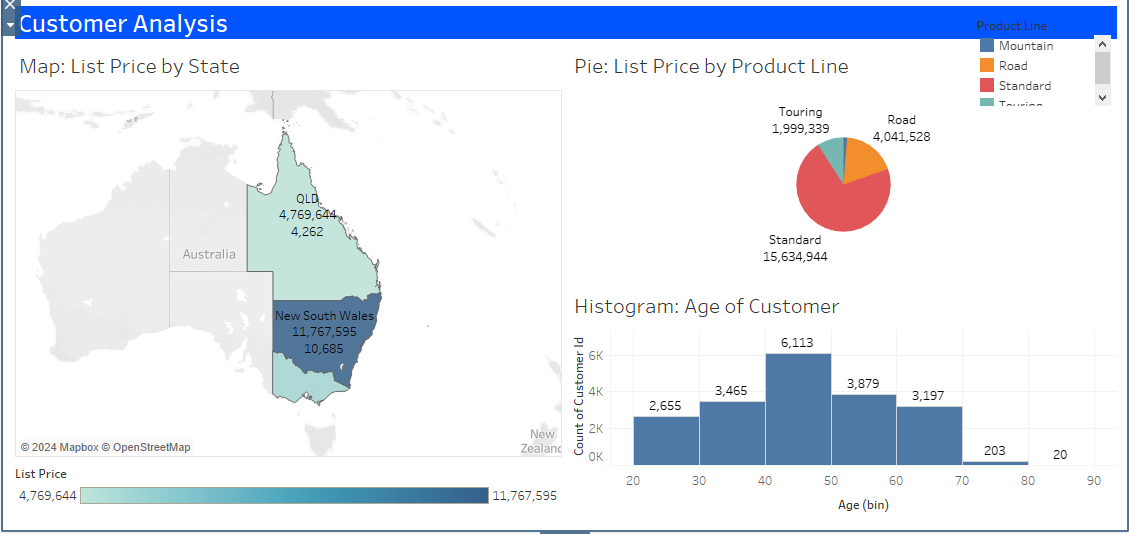
Hình 3. 5 Biểu đồ thể hiện doanh số hàng tháng theo quốc gia

- Biểu đồ doanh số theo dòng sản phẩm và năm đặt hàng ở dạng horizontal bars

Hình 3. 6 Biểu đồ thể hiện doanh số theo dòng sản phẩm và năm đặt hàng

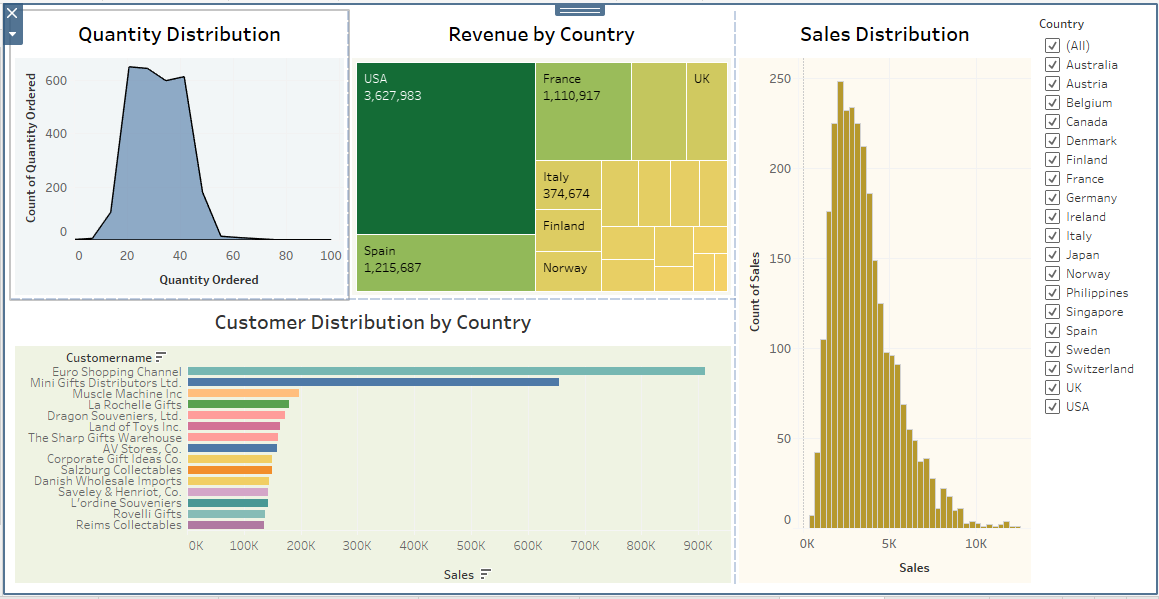
Kết quả các Dashborad

Customer Analysis



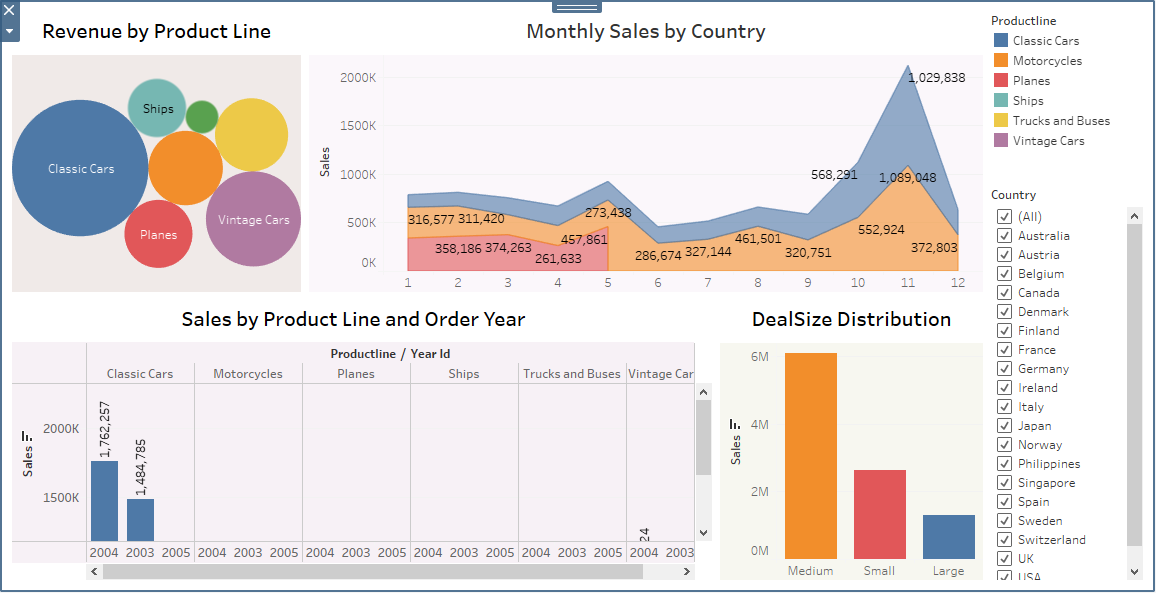
Hình 3. 7 Dashboard tổng quan báo cáo doanh thu theo vùng

Global Sales Analytics Dashboard



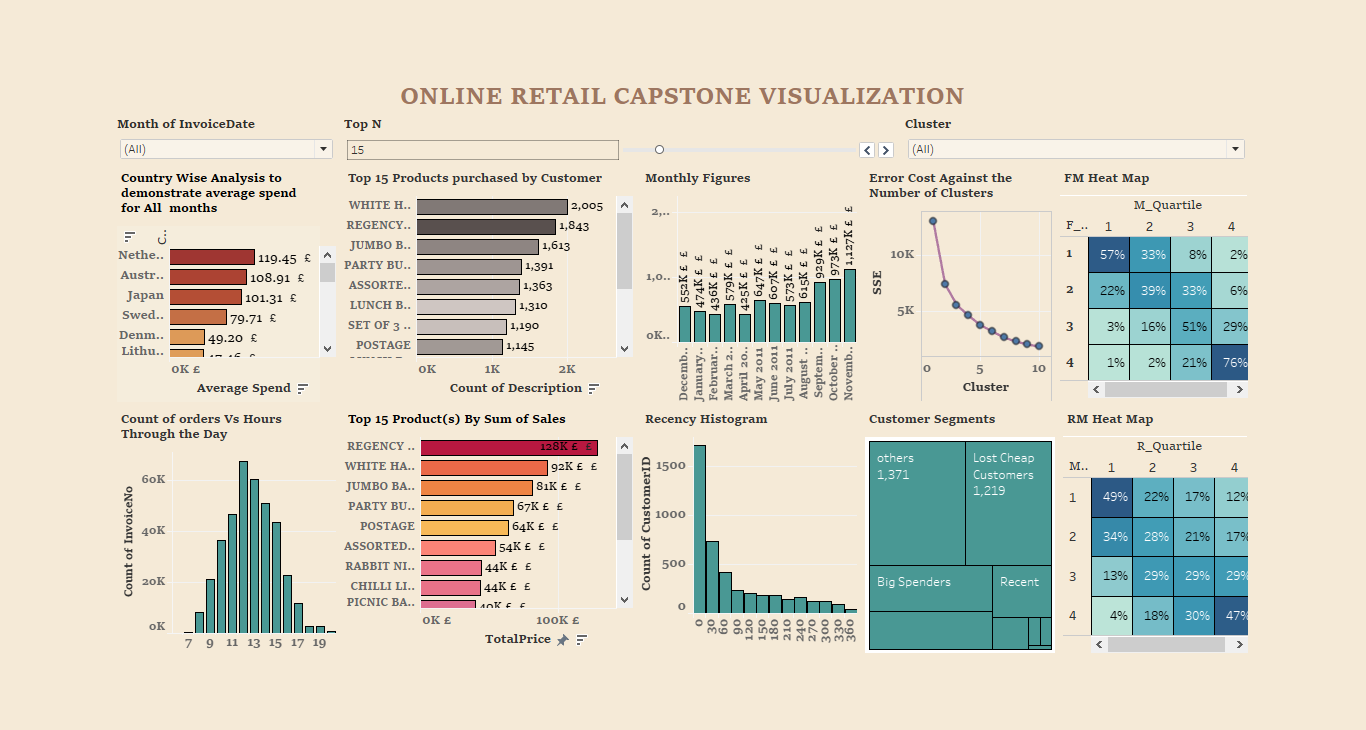
Hình 3. 8 Dashboard tổng quan về hoạt động kinh doanh toàn cầu của công ty.

Sales Performance & Analytics Dashboard



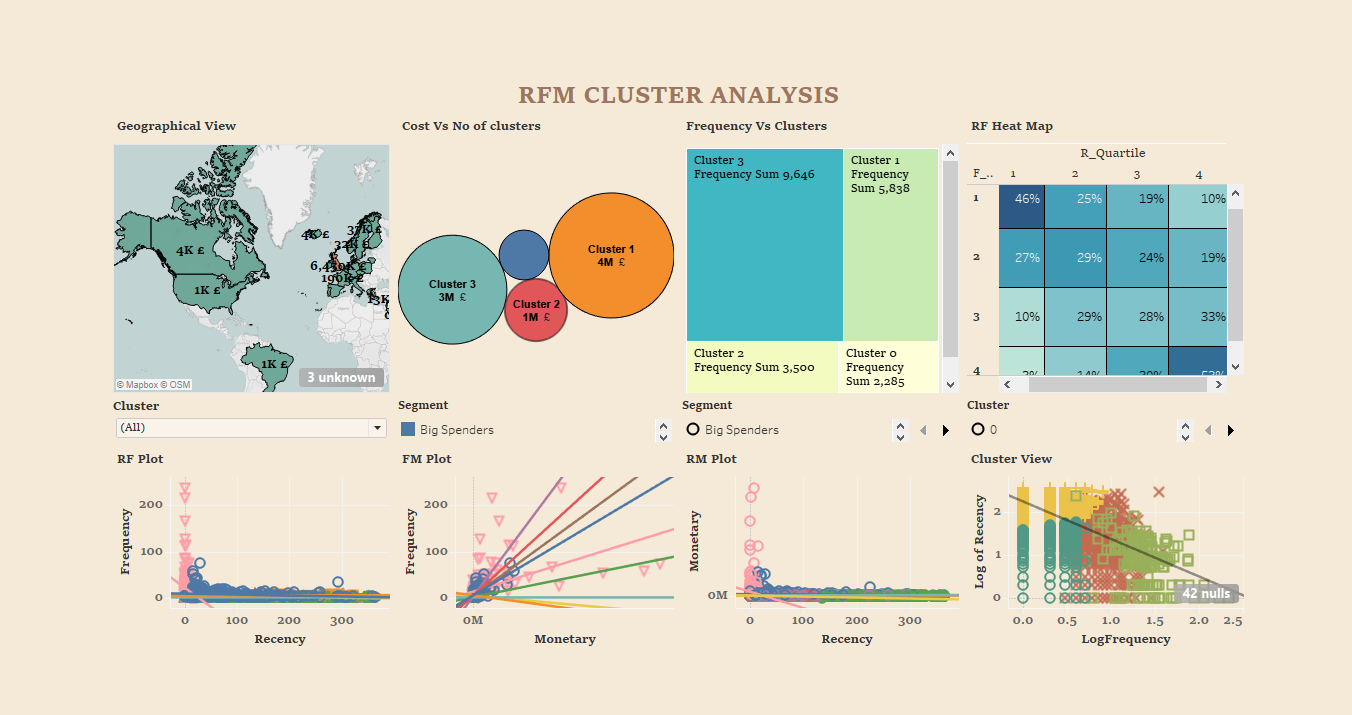
Hình 3. 9 Dashboard thể hiện hiệu suất bán hàng và phân tích kinh doanh

Online retail capstone visualization



Hình 3. 10 Dashboard trực quan hóa sản phẩm bán lẻ trực tuyến

RFM cluster analysis



Hình 3. 11 Dashboard phân tích nhóm khách hàng theo tiêu chí RFM

**3.3. Tiện ích Tableau so với bảng tính thông thường**

- Bảng tính sẽ có những tiện ích như sau:

+ Với bảng tính thông thường, một công cụ tốt nhưng chưa hoàn chỉnh, không thể tích hợp tất cả dữ liệu, không thể phân tích được dữ liệu lớn và từ nhiều nguồn, dễ xảy ra lỗi và thường có thể dẫn đến sự cố sập chương trình, khiến bạn phải chờ đợi và thất vọng.

+ Về việc kết hợp và dọn dẹp dữ liệu: hầu hết các bảng tính thông thường đều lưu trữ ngày theo các dòng trong cột. Điều này khiến cho việc xử lý dữ liệu ngày trở nên phức tạp và tốn nhiều công sức.

+ Trực quan hóa dữ liệu nghĩa là giúp dữ liệu dễ đọc hơn và đọc nhanh hơn. Bảng tính dựa trên số thường làm điều ngược lại. Các công cụ của bảng tính cho phép tạo một số bảng biểu cơ bản, tích hợp sẵn. Tuy nhiên, trong thị trường ngày nay, những đồ thị tiêu chuẩn này không còn phát huy nhiều tác dụng nữa, thay vào đó là nhiều loại hình ảnh nâng cao giúp đạt được giá trị thực thụ từ dữ liệu.

+ Không phải dữ liệu nào cũng cần được biểu diễn ở dạng biểu đồ cột hoặc biểu đồ thanh. Não người xử lý số, hình ảnh trực quan hoặc ảnh dưới dạng một “mẩu” thông tin.

- Còn với Tableau:

+ Tableau với nội dung trực quan đẹp mắt, công phu bằng nhiều loại đồ thị và biểu đồ khác nhau giúp làm sáng tỏ mọi vấn đề. Có thể nhanh chóng hiểu rõ và chia sẻ cả một bức tranh hoàn chỉnh cũng như giải đáp được tất cả những câu hỏi mà báo cáo thông thường không lường trước được.

+ Bảng giao diện tương tác với nguồn cấp dữ liệu mới nhất có thể hỗ trợ các nhiệm vụ phân tích quan trọng, chẳng hạn như lọc dạng xem, điều chỉnh tham số, thực hiện tính toán nhanh và xem chi tiết để kiểm tra dữ liệu cơ bản. Trả lời nhanh các câu hỏi ngoài dự kiến chỉ với vài cú nhấp chuột. Dữ liệu có thể được lấy từ quý này, tuần này, giờ này hay thậm chí 30 giây trước. Chắc chắn, nguồn cấp dữ liệu mới nhất hoặc nguồn cấp trong dòng thời gian phù hợp với tổ chức sẽ đẩy nhanh tốc độ giải quyết vấn đề và hoạt động hằng ngày của tổ chức.

+ Tableau tổng hợp các thông tin liên quan ở một nơi, tự động cập nhật dữ liệu mọi lúc. Tableau giúp mọi người xem và hiểu dữ liệu, bất kể dữ liệu lớn cỡ nào hay được lưu trữ ở đâu. Nhanh chóng kết nối, kết hợp, dọn dẹp, trực quan hóa và chia sẻ bảng giao diện số từ máy tính đến IPad với trải nghiệm liền mạch. Tạo và phát hành bảng giao diện số có khả năng tự động cập nhật dữ liệu rồi chia sẻ với đồng nghiệp, đối tác hoặc khách hàng mà không cần kỹ năng lập trình.

+ Chắc chắn một chồng bảng tính thông thường in ra sẽ không được xem là đặt tất cả các yếu tố có liên quan ở một nơi. Bảng tính của bạn chỉ hữu ích cho đến lần trích xuất hoặc cập nhật dữ liệu tiếp theo.”

KẾT LUẬN

1. **Kết quả**

- Hiểu được tổng quan về Tableau.

- Tìm hiểu được các cách đưa dữ liệu và xử lý dữ liệu với Tableau.

- Biết được các chart cơ bản, phục vụ cho trực quan hoá dữ liệu.

- Xây dựng được những báo cáo đầy đủ thông tin giúp nhà lãnh đạo nắm được những thông tin quan trọng một cách nhanh nhất và chính xác nhất.

1. **Kết quả thực nghiệm**

- Cài đặt được Tableau Desktop và Tableau Prep Builder.

- Xây dựng một dashboard tổng quan về kinh doanh của doanh nghiệp.

1. **Hạn chế**

- Tableau là một công cụ trực quan hóa dữ liệu mạnh mẽ và tinh tế, nhưng để sử dụng nó một cách đầy đủ và hiệu quả, doanh nghiệp cần đầu tư một khoản chi phí không nhỏ để hoàn thiện và tối ưu hóa quy trình sử dụng.

1. **Định hướng phát triển**

- Để đảm bảo dữ liệu được cập nhật hàng ngày cho các nhà lãnh đạo, cần tiến hành nghiên cứu để tối ưu quy trình xử lý và phân tích dữ liệu trong Tableau.

- Cần nắm vững các quy trình xử lý dữ liệu, lựa chọn biểu đồ phù hợp và tìm hiểu thêm các tài liệu liên quan để nâng cao khả năng sử dụng Tableau.”

TÀI LIỆU THAM KHẢO

[1]. Nguyễn Thị Hương Lý (2022), “Bài giảng khai phá dữ liệu kinh doanh”, Khoa CNTT, Trường Đại học Nha Trang.

[2]. Micheal Alenxander & John Walkenbach. Excel Dashboards and Reports. Wiley.

[3]. Jordan Goldmeier. Dashboards for Excel. Apress.