***I. Hoạt động của trang thương mại***

- Từ năm 2017 – 2020, có 793 khách hàng khác nhau đặt hàng từ sàn thương mại điện tử X. Trong từng năm từ giai đoạn 2017 – 2020, có khoảng (570 – 693 khách hàng khác nhau) => lượng repeated customers cao, khả năng thu hút những khách hàng mới cũng cao

- Trung bình chỉ tốn 3.91 ngày cho 1 đơn giao hàng, nhiều nhất là 8 ngày => sàn thương mại điện tử X đã tối ưu trong việc giao hàng => thu hút được nhiều khách hàng cũng như tối ưu được quá trình hoạt động của cả bộ máy

- Tổng doanh thu chỉ tập trung vào những tháng cuối năm, hơn nữa, lượng lợi nhuận không đồng đều ở các vùng khác nhau => các chính sách, ưu đãi thu hút người mua chưa hợp lí.

- Phần lớn lợi nhuận tập trung vào mảng công nghệ, các nhánh còn lại như (Nội thất, văn phòng phẩm, …) có dao động không ổn định, đặc biệt là nội thất chịu lỗ quanh năm => cách thu hút khách hàng chưa hợp lí, cần cân nhắc lại giá cả và các phương pháp bán hàng

***II. Dữ liệu khách hàng***

- Từ treemap, phần lớn khách hàng đặt mua đều từ thành phố lớn, các vùng trung tâm như (New York: 915 lượt mua, Los Angeles: 747 lượt mua, Philadelphia, … )

- Chỉ riêng New York đã chiếm gần nửa doanh số so với tổng thu được và cũng là nguồn lợi nhuận chính cho sàn thương mại điện tử X

***III. Tự động hóa***

- Do em chưa thuần thục sử dụng SQL nên vẫn còn gặp nhiều vướng mắc khi import và tự động hóa file excel nên em sử dụng Python để tự động hóa dữ liệu. Mỗi ngày, set task scheduler vào 9h sáng để lấy báo cáo cho ngày mới, báo cáo update sẽ từ file gốc (report\_data.xlsx) và làm mới sang (reportDataClean.xlsx). Task mangager sẽ chạy file runPython.bat

- Nếu muốn tự động hóa powerBI thì cần sử dụng dữ liệu trên web hoặc từ SQL -> sử dụng (from google.collab import drives) để upload dữ liệu lên drive và tiến hành các bước tự động hóa. Publish báo cáo lên web và vào phần dataset, chọn scheduled refresh để tự động hóa

***IV. Cách bố trí powerBI***

- Tạo các cột chứa phép tính để tính tổng Sale và tổng Profit (nhân với quantity)

- Tính ship time = datediff(order date, ship date, day)

- Tạo các measure để tính các giá trị đặc biệt như giá trị trung bình, …

- Sử dụng các slicers để biểu thị các thông tin cụ thể theo riêng từng mục

- Dùng line & column chart để thể hiện lợi nhuận và doanh thu trên cùng biểu đồ, có data label để quan sát cụ thể

- Dùng treemap để xem phân bố các thành phố, người mua thường tập trung ở đâu

- Có các card để biểu thị các thông số thống kê