

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN



BÁO CÁO ĐỒ ÁN
QUAN HÓA DỮ LIỆU

Môn học: Trực quan hóa dữ liệu

Thành viên nhóm:

- | | |
|-----------------------|----------------|
| – Lê Nguyễn Kiều Oanh | MSSV: 21127129 |
| – Đoàn Ngọc Mai | MSSV: 21127104 |

Giảng viên hướng dẫn:

- Bùi Tiến Lên
- Lê Ngọc Thành
- Lê Nguyễn Nhật Trường

Thành phố Hồ Chí Minh, ngày 7 tháng 4 năm 2024

Mục lục

I. Thông tin nhóm.....	3
II. Phân công công việc	3
III. Những phát hiện ý nghĩa.....	3

I. Thông tin nhóm

Họ và tên	Mã số sinh viên
Đoàn Ngọc Mai	21127104
Lê Nguyễn Kiều Oanh	21127129

II. Phân công công việc

Họ và tên	Mã số	Công việc	Mức độ hoàn thành
Đoàn Ngọc Mai	21127104	Hoàn thành mục EDA 3D	100%
		Hoàn thành phần Insight	100%
		Viết Báo cáo và trang trí lại notebook	100%
Lê Nguyễn Kiều Oanh	21127129	Hoàn thành Data Understanding	100%
		Hoàn thành EDA 1D	100%
		Hoàn thành EDA 2D	100%

III. Những phát hiện ý nghĩa

Bữa ăn **Bed & Breakfast** và loại phòng **A** dường như là lựa chọn phổ biến, và hầu hết khách hàng đặt phòng thông qua các **kênh trực tuyến** và **không yêu cầu tiền đặt cọc**.

Số lượng **adults** trong các đặt phòng thường cao hơn so với **children** và **babies**.

Số lượng đặt phòng thông qua **agent** cao hơn nhiều so với **company**, cho thấy sự ưu tiên sử dụng đại lý du lịch.

Total_of_special_requests có tương quan tiêu cực với **is_canceled**, điều này có thể chỉ ra rằng càng nhiều yêu cầu đặc biệt được đưa ra, khả năng khách hàng sẽ không hủy đặt phòng.

Lead_time và **is_canceled** có tương quan tích cực mạnh, cho thấy thời gian từ khi đặt đến ngày đến càng lâu, khả năng hủy đặt phòng càng cao.

Previous_cancellations có tương quan tích cực với **is_canceled**, nghĩa là khách hàng đã từng hủy bỏ đặt phòng trước đây có nhiều khả năng sẽ hủy lại.

Đa số các đặt phòng không phải chờ đợi quá lâu trước khi được xác nhận hoặc chỉ thực hiện một số ít thay đổi.

Những khách hàng loại **Transient-Party** có thời gian chờ đợi trong danh sách chờ lâu nhất.

Sự đặt phòng không tập trung vào một thời điểm cụ thể trong năm.

Hầu hết các đặt phòng, bất kể loại khách sạn, có ít hơn 5 thay đổi. Điều này cho thấy khách hàng thông thường không thường xuyên thực hiện nhiều thay đổi cho đặt phòng của họ.

Khách hàng loại **Contract** và **Transient-Party** thường có thời gian dẫn **lead time** dài hơn so với **Group** và **Transient**, điều này có thể liên quan đến tính chất đặt phòng trước và kế hoạch cụ thể hơn của những nhóm khách hàng này.

Khách hàng **Transient**, với thời gian dẫn trung bình ngắn hơn và phạm vi lớn, có thể phản ánh sự linh hoạt hơn trong việc lên kế hoạch du lịch.

Khách hàng **Contract** và **Transient** có xu hướng yêu cầu nhiều dịch vụ đặc biệt hơn so với các nhóm khách khác.

Mặc dù khách hàng **Transient** có số lần hủy trước đó ít, nhưng số lần hủy tối đa cao bất thường (26 lần), có thể phản ánh sự không chắc chắn hoặc thay đổi kế hoạch thường xuyên của nhóm này.

Khách hàng **Transient-Party** ít yêu cầu dịch vụ đặc biệt và có tỷ lệ hủy trước đó thấp, phản ánh tính linh hoạt và quyết định đặt phòng gần ngày lưu trú hơn.