

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN - ĐHQG TP HCM
KHOA CÔNG NGHỆ THÔNG TIN

—o0o—



BÁO CÁO ĐỒ ÁN GIỮA KỲ
PHÂN TÍCH SỰ HÀI LÒNG CỦA KHÁCH HÀNG

GVHD: TS Bùi Tiến Lên
Nhóm: 3

Mục lục

1	Thông tin nhóm	2
2	Nội dung	2
2.1	Bối cảnh	2
2.2	Giới thiệu về dữ liệu	2
2.3	Sử dụng Dashboard	3
2.4	Phân tích Dashboard	5
2.5	Một số cách tiếp cận khác	7
2.6	Đánh giá Dashboard	8
3	Tham khảo	10

1 Thông tin nhóm

MSSV	Họ và tên	Nhiệm vụ
21120045	Bùi Hồng Đăng	Phần nội dung: Đánh giá Dashboard Làm Slide, Dashboard, Thuyết trình
21120108	Nguyễn Tiến Nhật	Phần nội dung: Sử dụng Dashboard, Hướng tiếp cận khác Làm Slide, Dashboard
21120123	Lê Thanh Thái Quảng	Phần nội dung: Phân tích Dashboard Làm Slide, Dashboard, Thuyết trình
21120169	Thái Chí Vỹ	Phần nội dung: Bối cảnh, dữ liệu, Sử dụng Dashboard Làm Slide, Dashboard
21120201	Bùi Đình Bảo	Phần nội dung: Phân tích Dashboard, Đánh giá Dashboard Làm Slide, Dashboard

2 Nội dung

2.1 Bối cảnh

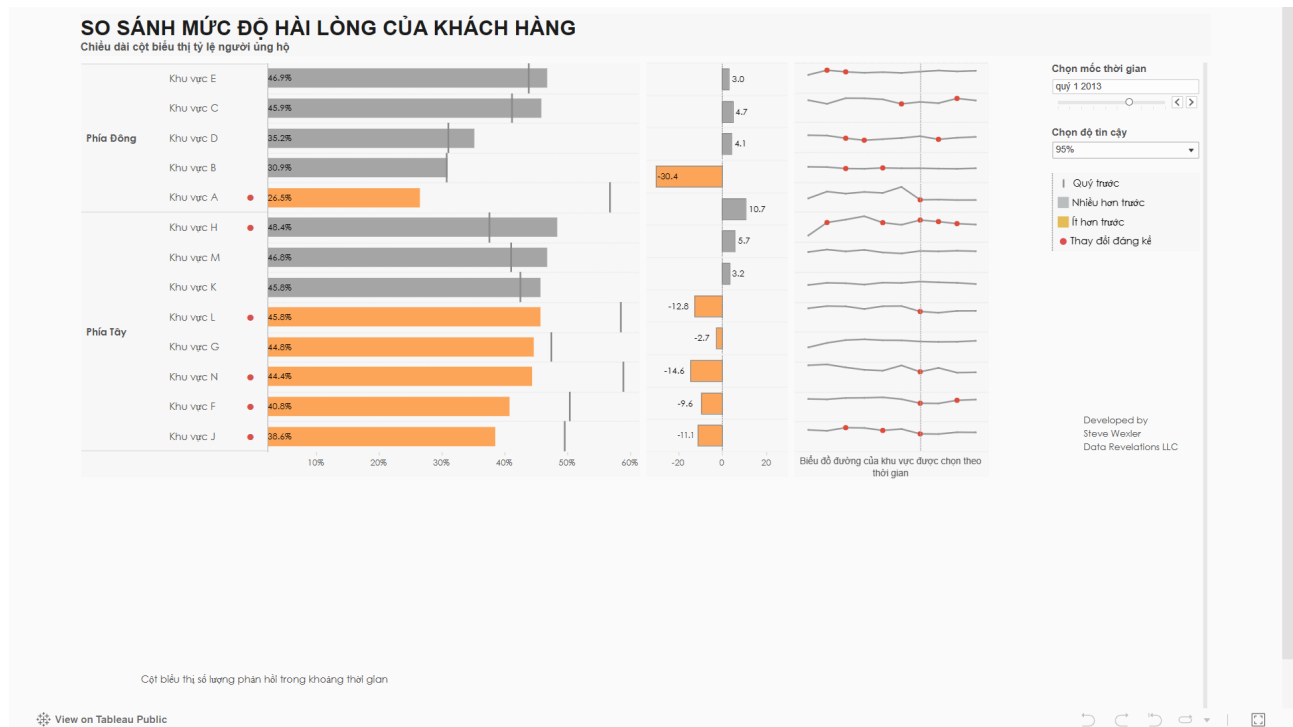
- Vấn đề: Công ty quan tâm đến sự hài lòng của khách hàng và muốn theo dõi điều này hàng quý theo các khu vực. Ngoài ra, công ty cũng muốn xem xét sự thay đổi mức độ hài lòng, đặc biệt là các biến động lớn, để kịp thời phản ứng với các vấn đề phát sinh.
- Nhiệm vụ:
 - Tổng hợp tỷ lệ phần trăm khách hàng rất hài lòng ("Người ủng hộ") theo phân khu và khu vực.
 - Sắp xếp dữ liệu theo khu vực để so sánh sự hài lòng giữa các khu vực.
 - Phân tích hiệu suất theo thời kỳ, ví dụ, so sánh quý này với quý trước.
 - Đánh giá sự thay đổi có ý nghĩa không bằng cách sử dụng các thước đo thống kê mà công ty đang áp dụng.
 - Xếp hạng doanh số bán hàng cho các sản phẩm và dịch vụ, phân tích theo từng bang, và so sánh chúng với một giai đoạn hoặc các giai đoạn trước đó.

2.2 Giới thiệu về dữ liệu

- Dataset được tham khảo từ quyển The Big Book of Dashboards
- Dataset có 119893 dòng, là đánh giá của một khách hàng, với 5 thuộc tính:
 - Công ty: chỉ công ty có dịch vụ được đánh giá
 - ID: thể hiện ID của khách hàng
 - Khu vực: thể hiện khu vực công ty được đánh giá bởi khách hàng sử dụng dịch vụ của công ty đó
 - Ngày: thể hiện ngày mà khách hàng đánh giá
 - Đánh giá: Mức độ hài lòng của khách hàng đối với dịch vụ của công ty (0-10) (Một khách hàng được tính là người ủng hộ nếu có "Đánh giá" ≥ 9)

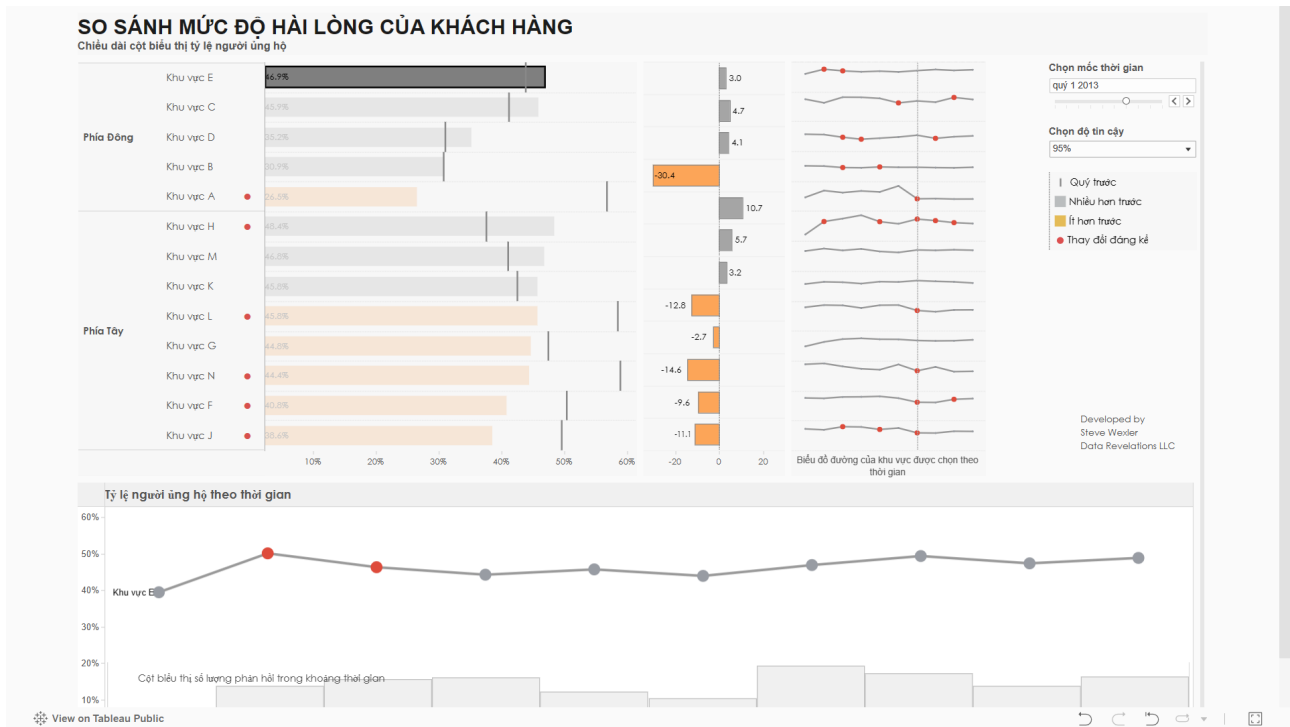
2.3 Sử dụng Dashboard

Trong phần này, nhóm sẽ trình bày tổng quan về giao diện của Dashboard cũng như cách sử dụng Dashboard này.



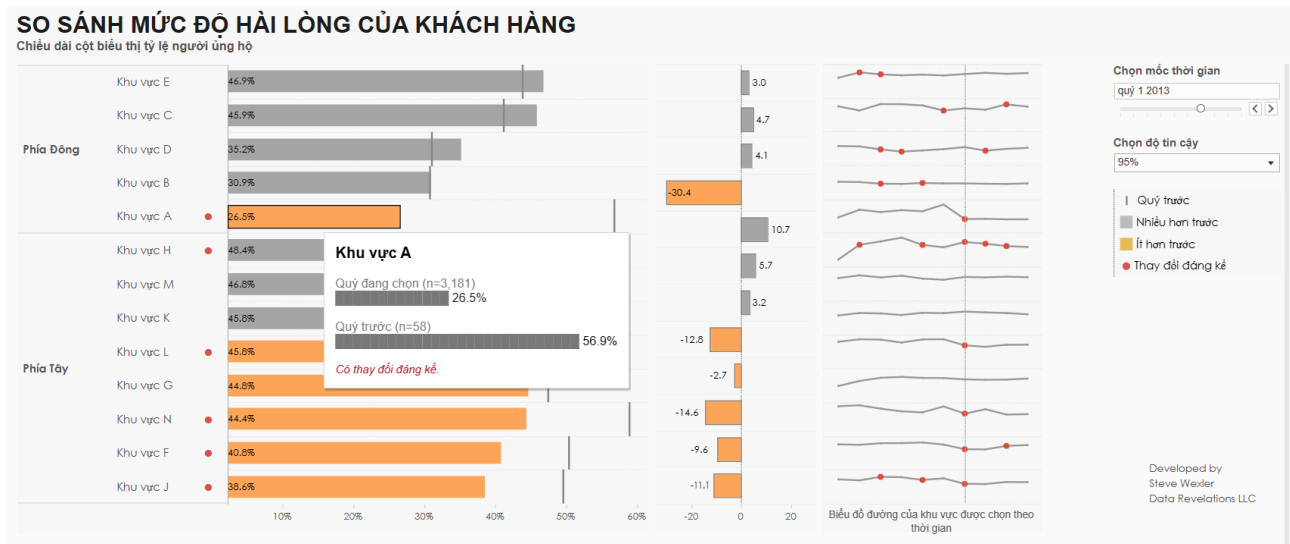
Hình 1: *Giao diện chính của Dashboard*

Hình 1 cho thấy giao diện chính của Dashboard. Phần chính của Dashboard bao gồm, biểu đồ cột ngang so sánh mức độ hài lòng của các khu vực được sắp xếp giảm dần, bên cạnh đó có biểu đồ cột cho thấy phần trăm tăng giảm mức độ hài lòng và biểu đồ đường cho thấy tổng quan về mức độ hài lòng của một khu vực. Ngoài ra, phía ngoài cùng bên phải có một bảng điều khiển để chọn thông tin về thời gian, độ tin cậy và chú thích màu sắc của biểu đồ. Phần dưới là phần hiển thị thông tin chi tiết của khu vực.



Hình 2: Chọn một khu vực

Khi chọn một khu vực thì thông tin của khu vực sẽ được hiển thị rõ hơn ở các biểu đồ bên dưới như *Hình 2*. Thông tin hiển thị bao gồm các biểu đồ: biểu đồ đường thể hiện tổng quan tất cả các mốc thời gian của một khu vực, biểu đồ cột thể hiện số lượng phản hồi của từng khu vực.

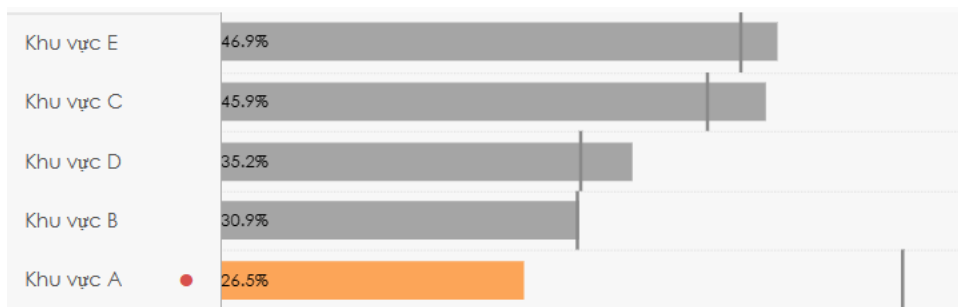


Hình 3: Đưa chuột lên một khu vực

Khi đưa chuột lên một khu vực, một cửa sổ bật lên cung cấp thêm thông tin về khu vực đó như *Hình 3*. Thông tin bao gồm số lượng phản hồi của khu vực tại quý đang chọn và quý trước và cho biết khu vực có sự thay đổi đáng kể hay không.

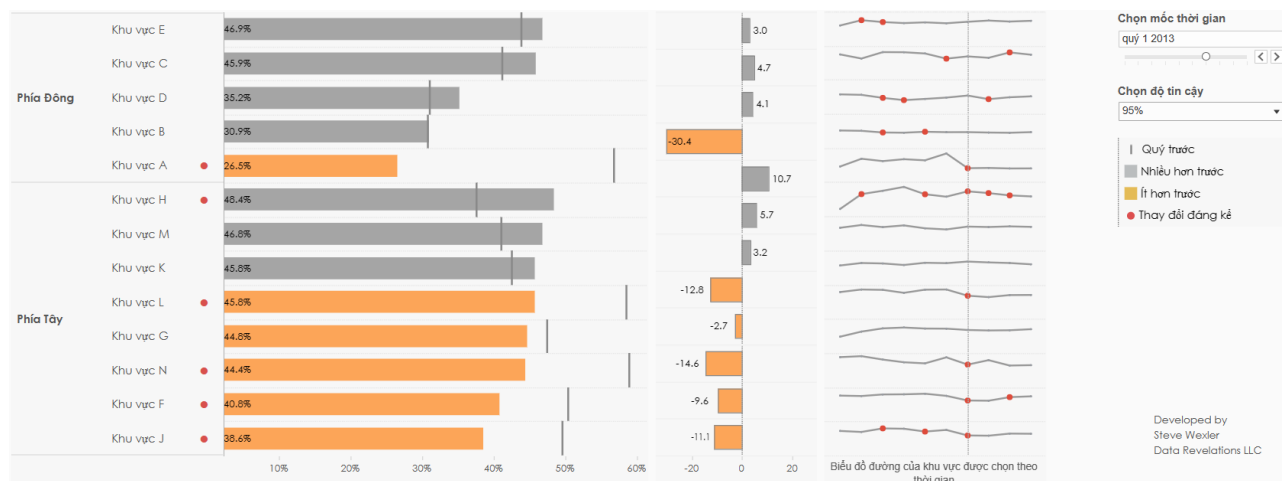
2.4 Phân tích Dashboard

Tiếp theo, nhóm sẽ đưa ra phân tích về Dashboard để trả lời cho câu hỏi tại sao Dashboard này lại thành công.



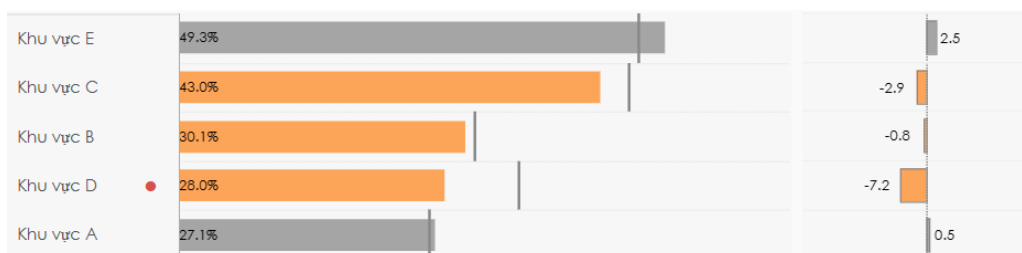
Hình 4: So sánh giữa các khu vực và so sánh qua thời kỳ

Việc sử dụng biểu đồ cột ngang giúp việc so sánh các khu vực với nhau trở nên dễ dàng hơn. Khi nhìn *Hình 4*, ta có thể dễ dàng nhận thấy một khu vực có tỷ lệ người ủng hộ cao hơn hoặc thấp hơn khu vực nào. Ngoài ra, các đường thẳng dọc ở từng cột của biểu đồ biểu thị mức độ hài lòng của khách hàng ở thời gian khảo sát quý trước đó. Từ đó ta có thể so sánh với mức độ hài lòng của khách hàng ở hiện tại với quý trước đó.



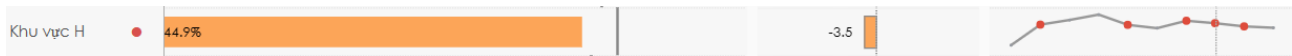
Hình 5: Màu sắc

Việc biểu diễn màu sắc của cột (màu xám tương ứng với sự tăng lên và màu vàng ứng với sự sụt giảm) về mức độ hài lòng của khách hàng so với giai đoạn trước giúp người xem có thể dễ dàng phân biệt từng loại khu vực theo sự tăng giảm về mức độ hài lòng của khách hàng một cách dễ dàng hơn.



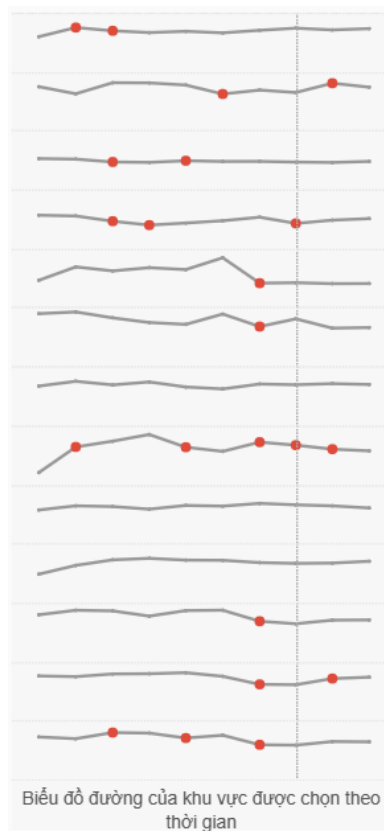
Hình 6: Số liệu tăng giảm cụ thể

Biểu độ phụ bên trái biểu đồ chính thể hiện tỷ lệ phần trăm tăng giảm của khu vực tại quý đang chọn so với quý trước *Hình 6*. Việc biểu diễn tỷ lệ phần trăm tăng giảm mức độ hài lòng của khách hàng ở mỗi khu vực so với giai đoạn trước giúp người xem có thể so sánh trong trường hợp chiều dài của các cột không có sự khác biệt quá lớn thì lúc đó các chỉ số về lượng sẽ trở nên hữu ích.



Hình 7: Chấm đỏ thể hiện thay đổi đáng kể

Các chấm đỏ trên biểu đồ dùng để biểu thị rằng vào quý nào thì sự thay đổi về mức độ hài lòng của khách hàng là đáng kể so với giai đoạn trước đó, sử dụng kiểm định Z-test với độ tin cậy tùy chọn để có thể tìm ra những cột có thay đổi đáng kể.



Hình 8: Biểu đồ đường

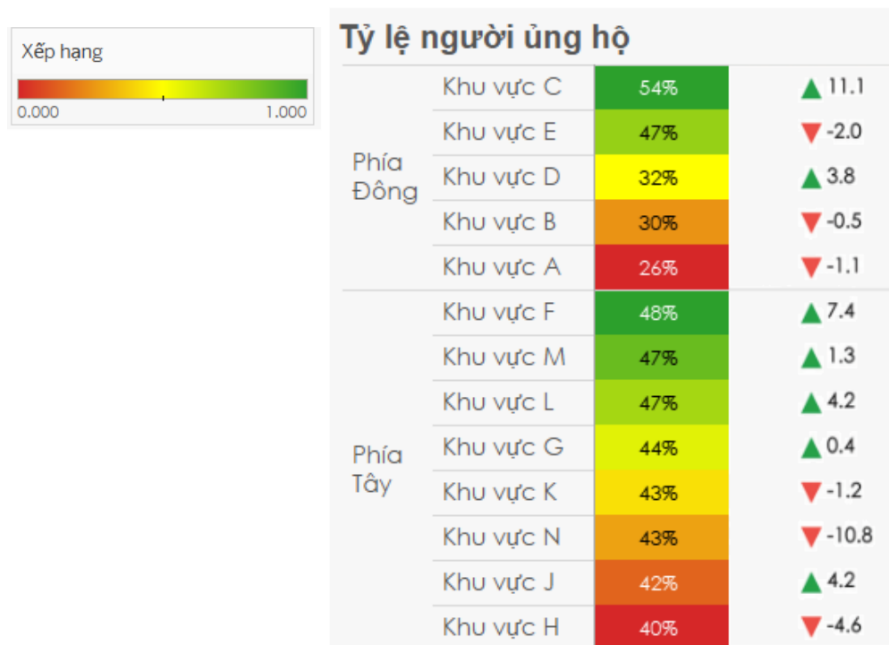
Biểu đồ đường cho ta thấy được tổng quan về sự thay đổi về mức độ hài lòng của khách hàng qua từng quý. Ngoài ra còn cho ta biết tại thời gian nào có sự thay đổi đáng kể (chấm đỏ tại thời gian đó).

Hình 9: Bảng điều khiển

Các tính năng khác của dashboard bao gồm: Ta có thể tùy chọn khoảng thời gian bằng thanh trượt phía góc trên bên phải và điều chỉnh độ tin cậy trong việc chọn ra các sự thay đổi đáng kể về mức độ hài lòng của khách hàng. Chú thích của biểu đồ cũng được thể hiện ở bên phải.

2.5 Một số cách tiếp cận khác

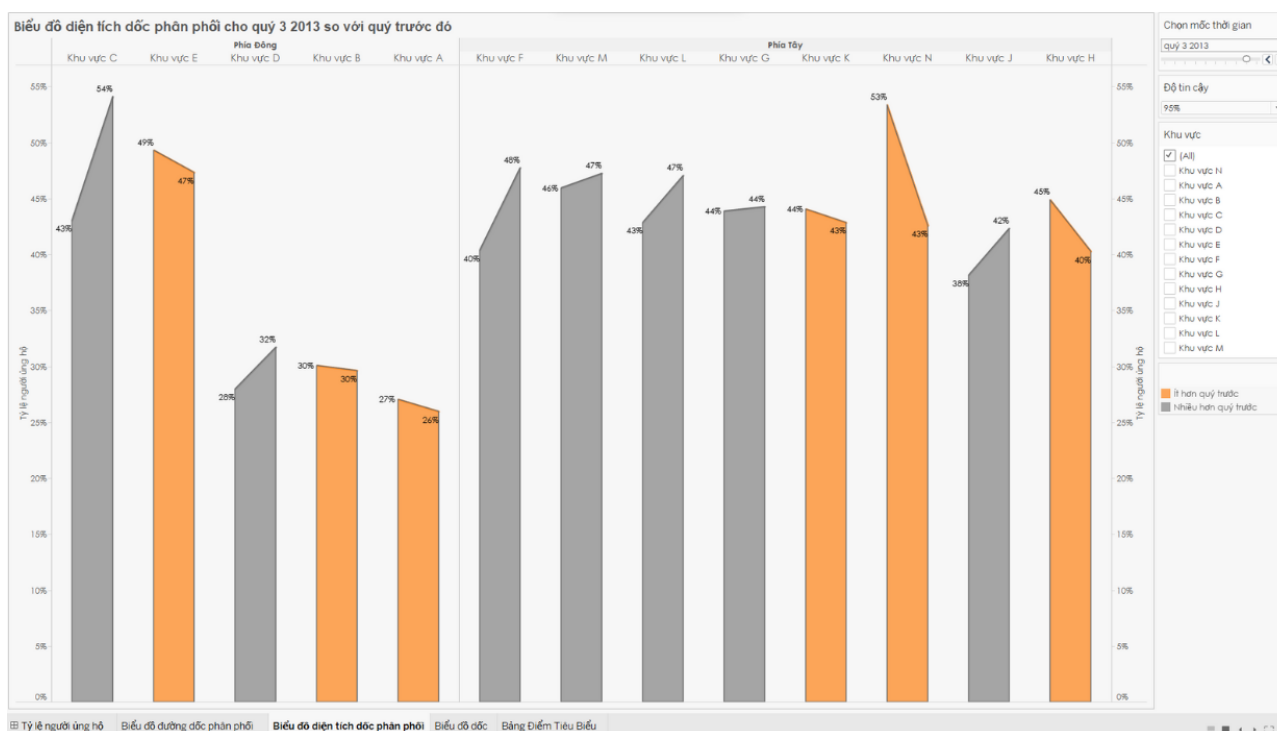
Để đưa ra lời giải cho bài toán trực quan "Phân tích mức độ hài lòng của khách hàng", có rất nhiều cách tiếp cận thường được sử dụng. Do đó trong phần tiếp theo này nhóm đề cập về một số cách tiếp cận khác đối với bài toán và tại sao những cách tiếp cận này lại không nên được sử dụng.



Hình 10: Score card

Biểu đồ Score card ở Hình 10 là một biểu đồ phổ biến, nhưng trong bài toán này nên tránh sử dụng vì các lý do sau:

- Việc so sánh độ hài lòng giữa các vùng miền bằng màu sắc không hiệu quả, cụ thể là ta khó có thể nhìn thấy khu vực C sẽ có độ hài lòng gấp đôi khu A
- Việc sử dụng màu sắc để phân biệt các vùng trong cùng một vai trò (phía đông và phía tây) sẽ dễ gây hiểu lầm cho người xem, cụ thể là giữa khu vực C của phía đông và khu vực F của phía tây có màu sắc giống nhau nhưng giá trị phần trăm lại khác nhau.
- Việc sắp xếp độ hài lòng theo từng vai trò (phía đông và phía tây) bằng mức độ màu sắc (với màu tốt nhất là màu xanh lá và màu xấu nhất là màu đỏ) sẽ gây ra vấn đề: Xem xét region H với 40.3%, nếu nó ở miền Đông thì nó sẽ có màu xanh nhạt hoặc vàng.
- Việc thể hiện sự tăng giảm có vấn đề đó là ví dụ như khu vực C tăng 11.1 thì người xem sẽ không biết được là 11.1 đó thể hiện phần trăm tăng hay là số đánh giá tốt tăng về mức độ hài lòng của khách hàng.
- Có sự xung đột giữa việc biểu diễn màu xanh cho rank (chỉ thứ hạng) và màu xanh cho sự thay đổi (chỉ sự tăng).



Hình 11: Biểu đồ diện tích dọc phân phối

Ngoài ra, còn một loại biểu đồ thường được sử dụng nữa đó chính là biểu đồ diện tích dọc phân phối như *Hình 11*. Tuy nhiên, dùng biểu đồ cột ngang như đã trình bày ở phần trước vẫn là sự lựa chọn tốt hơn vì:

- Hoạt động tốt hơn với thành phần khác của trang
- Cung cấp cái nhìn dọc theo thời gian, sẽ không phù hợp với biểu đồ độ dốc

2.6 Đánh giá Dashboard

Trong phần này, nhóm trình bày phần đánh giá Dashboard theo 10 tiêu chí

1. Kết hợp nguồn dữ liệu đáng tin cậy

- Nguồn dữ liệu của dashboard được tham khảo từ cuốn The Big Book of Dashboards
- Quy trình xử lý dữ liệu của dự án được tham khảo từ cuốn The Big Book of Dashboards

2. Phù hợp với mục đích

- Trực quan hóa phải phản ánh mục đích cụ thể của nó:
 - Biểu đồ cột ngang: So sánh mức độ hài lòng giữa các khu vực, có đường kẻ dọc để so sánh hiện tại và trước đây, cũng như thể hiện sự tăng hoặc giảm bằng màu sắc.
 - Biểu đồ đường: Thể hiện mức độ hài lòng qua các năm, có thể sử dụng dấu chấm đỏ để đánh dấu vùng có sự tăng đáng kể.
- Phù hợp với mục tiêu vì thông qua trực quan có thể:
 - So sánh mức độ hài lòng trong các khu vực
 - So sánh mức độ hài lòng trong 1 khu vực qua các thời gian khác nhau
 - Thể hiện xu hướng tăng/giảm của mức độ hài lòng của 1 khu vực
 - Làm nổi bật khu vực tăng đáng kể

3. Rõ ràng và dễ hiểu

- Trực quan hóa đã truyền đạt thông tin rất rõ ràng và dễ hiểu, mọi biểu đồ trên dashboard đều thể hiện sự so sánh: biểu đồ cột ngang so mức độ hài lòng giữa các khu vực và tiểu khu vực; biểu đồ đường so giữa các quý, đặc biệt là so tăng trưởng, suy thoái giữa 2 kỳ liên tiếp theo biểu đồ cột.
- Biểu đồ đã được trực quan hóa một cách hợp lý:
 - Biểu đồ cột ngang hiển thị mức độ hài lòng của khách hàng theo khu vực và tiểu khu vực, kèm theo biểu đồ cột phụ để thể hiện sự biến đổi so với quý trước.
 - Biểu đồ đường và cột được đơn giản hóa, phần cột so sánh mức độ hài lòng giữa các quý.
 - Dấu chấm đỏ thể hiện sự thay đổi có ý nghĩa của mức độ hài lòng.

4. Sự tích hợp và liên kết

- Các biểu đồ cột, đường được liên kết với nhau thông qua việc sử dụng bộ dữ liệu và cùng mục tiêu phân tích - đánh giá sự hài lòng của khách hàng
- Bên cạnh biểu đồ cột ngang thể hiện mức độ hài lòng của khách hàng, có 1 biểu đồ cột ngang khác thể hiện sự biến đổi so với quý trước của mỗi khu vực. Ngoài ra, biểu đồ đường tương ứng với mỗi khu vực cho thấy quá trình biến động mức độ hài lòng qua các quý.
- Khi lựa chọn một khu vực trên dashboard tương tác, biểu đồ cột sẽ chi tiết phần trăm hài lòng qua các quý và biểu đồ đường cho thấy xu hướng biến đổi, với điểm đỏ đánh dấu thay đổi lớn.

5. Phân tích được sự thay đổi và xu hướng

- Điểm mạnh
 - Trực quan hóa thời gian: sử dụng biểu đồ đường hoặc cột để thể hiện sự thay đổi theo thời gian, giúp dễ dàng nhận biết xu hướng và mẫu thay đổi

- Mỗi quan hệ rõ ràng: Sự sắp xếp và thiết kế màu sắc, kích thước và vị trí của dashboard cho phép hiểu được mối quan hệ giữa các biến số
- Điểm yếu
 - Giới hạn dữ liệu: Nếu dashboard không được cập nhật thường xuyên việc phân tích xu hướng có thể bị hạn chế
 - Phức tạp trong hiểu biết: Đối với người dùng không chuyên, việc hiểu và phân tích các mối quan hệ phức tạp giữa các biến có thể khó khăn nếu không có hướng dẫn cụ thể

6. Tương tác và điều hướng

- Dashboard tương tác cho phép lựa chọn khu vực để hiểu rõ thông tin.
- Mức độ hài lòng của quý trước đó được thể hiện qua biểu đồ cột.
- Các số liệu cụ thể được cung cấp qua các điểm thể hiện dữ liệu trên các biểu đồ đường

7. Thiết kế hấp dẫn

- Việc sử dụng thiết kế đồ họa và màu sắc trong trực quan hóa trên dashboard tạo nên sự thu hút và giúp dễ dàng tập trung. Khi áp dụng màu sắc xám và cam trên các biểu đồ cột và đường, những màu này không những tạo nên vẻ đẹp hài hòa, mà còn giúp người dùng dễ dàng phân biệt và theo dõi dữ liệu. Ngoài ra, hai màu sắc này giúp người bị mù màu dễ dàng phân biệt được.
- Sử dụng màu có chủ ý trong thiết kế có thể nâng cao khả năng nhận biết về thông tin được trình bày. Chẳng hạn, áp dụng xám cho một loại dữ liệu và cam cho loại dữ liệu khác giúp người xem lập tức nhận diện và hiểu được sự khác biệt và mối quan hệ. Đồng thời, cách tiếp cận này cũng tránh gây quá tải với quá nhiều màu sắc, giúp thông tin không bị lan man và giữ được sự tập trung cho người xem

8. Tính phản hồi

- Tùy thuộc vào người xem muốn làm rõ vấn đề nào mà có thể tùy chỉnh dashboard để chỉ ra rõ vấn đề, trả lời câu hỏi của người xem
- Có thể dựa vào ý kiến của người xem để điều chỉnh dashboard. Ví dụ điều chỉnh về các biến trực quan

9. Khả năng tích hợp, chia sẻ và hiệu suất

Trong phần trình bày dự án này, tác giả chưa có đề cập về hai tiêu chí khả năng tích hợp, chia sẻ và hiệu suất nên nhóm không đánh giá hai tiêu chí này.

3 Tham khảo

Tài liệu

- [1] Steve Wexler, Jeffrey Shaffer, andy Cotgreave, *The Big Book of Dashboard*.
- [2] Steve Wexler, *Dashboard Customer Satisfaction Comparision*.