**Mở bài**

Trong thời đại số hóa, dữ liệu không chỉ là tài sản quan trọng mà còn là nguồn sức mạnh để đưa ra các quyết định chiến lược. Chính vì thế, nhóm chúng tôi đã tiếp cận bộ dữ liệu sản phẩm Amazon từ nguồn Kaggle với mục tiêu khám phá hành vi tiêu dùng và đánh giá chất lượng sản phẩm. Tuy nhiên, bộ dữ liệu này không đơn thuần là một kho thông tin khổng lồ mà còn đặt ra nhiều thách thức lớn, từ quy mô rộng lớn, đa dạng sản phẩm đến tỷ lệ giá trị thiếu lên tới 30%. Qua báo cáo này, chúng tôi muốn chia sẻ cách mà nhóm đã và sử dụng dữ liệu để mang lại những góc nhìn sâu sắc và giá trị.

**Thân bài**:

1. **Kiểm tra dữ liệu**

Bước đầu tiên trong quá trình xử lý dữ liệu là tải dữ liệu và đưa vào Python để phân tích. Điều này không chỉ đơn thuần là nhập liệu, mà là bước khởi đầu quan trọng để đảm bảo chất lượng của toàn bộ quy trình. Sau khi nhập dữ liệu, việc kiểm tra cấu trúc, loại dữ liệu, cũng như xác định các dòng bị lỗi hoặc thiếu không chỉ giúp có cái nhìn tổng quan mà còn là bước thiết yếu để đảm bảo tính chính xác và toàn vẹn của dữ liệu. Suy cho cùng, độ chính xác của dữ liệu đầu vào sẽ quyết định độ tin cậy của kết quả phân tích.

2. **Xử lý các lỗi định dạng**

Sau khi rà soát dữ liệu, nhóm đã tiến hành xử lý lỗi định dạng tại các cột no\_of\_ratings, actual\_price, và discount\_price. Các cột này chứa các ký tự không phù hợp như dấu phẩy và ký hiệu tiền tệ, dẫn đến việc Python không thể xác định đúng loại dữ liệu. Sử dụng các thuật toán Python từ các thư viện như Pandas và re, lỗi này đã được khắc phục hiệu quả. Điều này minh họa rõ ràng sự linh hoạt và mạnh mẽ của các công cụ xử lý dữ liệu hiện đại.

Trong quá trình xử lý, nhóm cũng phát hiện một số dòng văn bản tại các cột ratings và no\_of\_ratings. Những dòng này đã được phân tích cẩn thận để đảm bảo rằng chúng không làm sai lệch kết quả. Cuối cùng, chúng tôi quyết định chuyển chúng thành giá trị Null, một lựa chọn giúp giữ nguyên tính toàn vẹn của dữ liệu mà không bỏ qua các giá trị cần thiết.

3. **Xử lý các dữ liệu bị lặp**

Tiếp theo, nhóm loại bỏ các dòng dữ liệu bị lặp. Việc này không chỉ giúp làm giảm nhiễu trong phân tích mà còn tăng cường tính chính xác của kết quả. Trong ngành tài chính hoặc y tế, việc loại bỏ dữ liệu lặp là một yếu tố sống còn để đảm bảo rằng phân tích mang tính đại diện và không sai lệch.

4. **Xử lý các dữ liệu bị thiếu**

Một bước đầy thách thức khác là xử lý các giá trị thiếu (missing values). Với dữ liệu đa dạng và phức tạp, việc bổ sung bằng trung bình hay trung vị sẽ không hiệu quả. Do đó, nhóm áp dụng phương pháp KNN Imputation - một kỹ thuật hiện đại dựa trên các mẫu dữ liệu lân cận để ước lượng giá trị thiếu. Bằng cách thực hiện riêng cho từng sub\_category, phương pháp này giúp giảm thiểu sai số tiềm ẩn và nâng cao chất lượng dữ liệu.

5. **Xử lý các dữ liệu ngoại biên**

Ngoài ra, để xử lý các ngoại lệ (outliers), phương pháp IQR (Interquartile Range) đã được sử dụng. Đây là một giải pháp tối ưu, tập trung vào phân bố tự nhiên của dữ liệu và đảm bảo rằng các giá trị bất thường không ảnh hưởng đến phân tích tổng thể.

6. **Tạo đặc trưng**

Bên cạnh việc xử lý dữ liệu, nhóm còn thực hiện kỹ thuật tạo đặc trưng (feature engineering). Đây là giai đoạn quan trọng để làm giàu bộ dữ liệu và cải thiện khả năng phân tích. Ví dụ, nhóm đã trích xuất supplier từ cột link, sử dụng thuật toán K-means clustering để phân nhóm giá thành, và thêm cột discount\_percentage. Những đặc trưng này không chỉ giúp tăng cường khả năng trực quan hóa mà còn cung cấp các góc nhìn mới mẻ cho quá trình phân tích.

7. **So sánh kết quả trước và sau khi xử lý**

Cuối cùng, sau mỗi bước xử lý, nhóm đều kiểm tra kỹ lưỡng và lưu lại kết quả trong các tệp riêng biệt. Điều này không chỉ giúp đảm bảo tính minh bạch mà còn tạo điều kiện dễ dàng cho việc tái hiện và kiểm chứng. Kết quả cuối cùng cho thấy, sau các bước xử lý kỹ lưỡng, bộ dữ liệu không chỉ trở nên sạch hơn mà còn đảm bảo tính đáng tin cậy và sẵn sàng cho phân tích sâu hơn.

A screenshot of a computer screen

AI-generated content may be incorrect.

Nguồn: Python Source Code

**Trực quan hóa dữ liệu**

Trong bối cảnh thương mại điện tử ngày càng cạnh tranh, việc hiểu sâu các yếu tố ảnh hưởng đến chiến lược giảm giá không chỉ là lợi thế cạnh tranh mà còn là yếu tố then chốt để đạt được hiệu quả kinh doanh tối ưu và gia tăng sự hài lòng của khách hàng. Nhóm nghiên cứu của chúng tôi đã thực hiện phân tích chi tiết nhằm khám phá mối quan hệ giữa các biến số quan trọng như tỷ lệ giảm giá (%Discount\_price), số lượng đánh giá (No\_of\_ratings), đánh giá (Ratings), và giá thực tế (actual price).

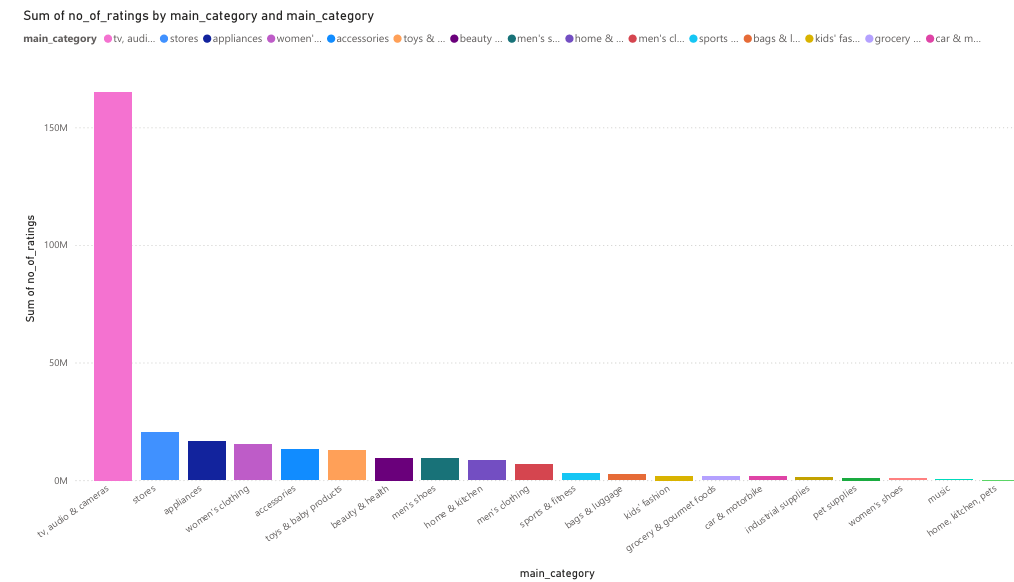
**1. Số lượng đánh giá (No\_of\_ratings) và các ngành hàng chính (main\_category)**

Nhóm nghiên cứu bắt đầu bằng việc phân tích số lượng đánh giá (No\_of\_ratings) theo ngành hàng chính (main\_category). Kết quả cho thấy rằng các ngành hàng có tỷ lệ giảm giá cao thường thu hút nhiều đánh giá hơn, đặc biệt là ở các ngành hàng liên quan đến nhu yếu phẩm và thời trang. Điều này chứng tỏ rằng các ưu đãi giá cả có thể trực tiếp kích thích sự quan tâm và tăng cường trải nghiệm mua sắm của khách hàng.

* 1. **Biểu đồ cột**

- Nhóm chọn sử dụng biểu đồ cột để biểu diễn mức độ phân bố số lượt đánh giá (trục x) giữa các ngành hàng chính (trục y)

- Lý do: để có thể quan sát rõ hơn mức độ phân bố các lượt đánh giá của khách hàng trong các danh mục cụ thể



- **Câu hỏi:**

+ các mặt hàng được nhiều lượt đánh giá

- **Kết quả từ biểu đồ:**

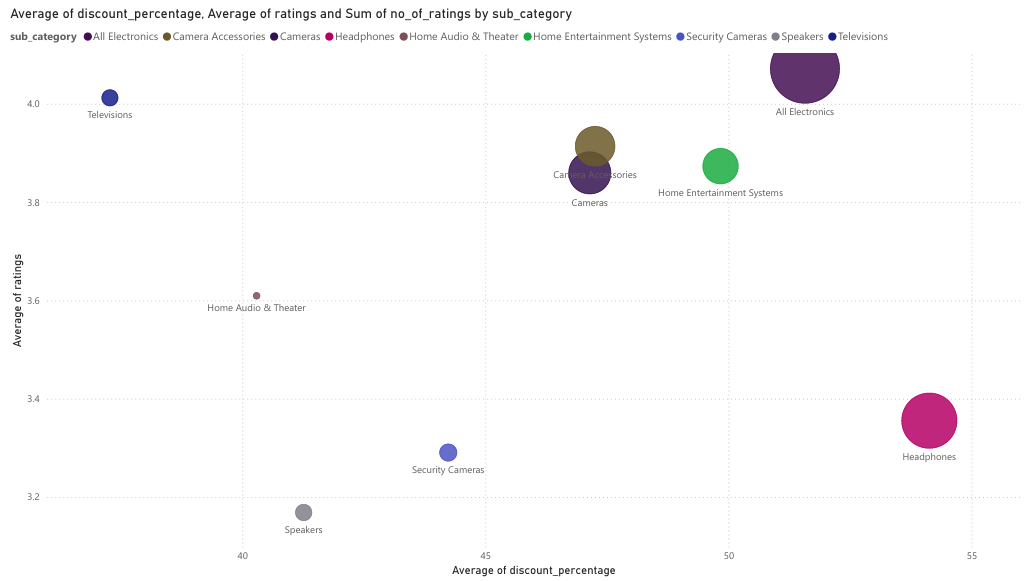
Sau khi trực quan hoá Main\_category tv,audio & cameras là danh mục có lượt đánh giá vượt trội khi so với các danh mục khác

• **Tổng kết**: Khách hàng có xu hướng đánh giá nhiều hơn đôi với các sản phẩm thuộc danh mục tv, audio & cameras

**1.2 Biểu đồ phân bố của danh mục phụ của danh mục tv, audio & cameras**

- Nhóm chọn sử dụng biểu đồ scatter plot để có thể biểu diễn được trung bình giảm giá (trục x), trung bình đánh giá(trục y), số lượt đánh giá(kích thước bong bóng) theo ngành hàng phụ của danh mục tv, audio & cameras

- **Lý do**: Từ biểu này có thể quan sát được trung bình giảm giá có ảnh hưởng đến đánh giá của khách hàng đối với sản phẩm không.



- **Câu hỏi:**

+ Các mặt hàng được giảm giá nhiều có nhận được đánh giá tốt không

+ Giảm giá có ảnh hưởng thế nào đến số lượt đánh giá của ngành hang

- **Kết quả từ biểu đồ:**

+ Tai nghe (Headphones): Có tỷ lệ chiết khấu trung bình cao nhất (~55%), nhưng điểm đánh giá trung bình chỉ khoảng 3.6. Kích thước bong bóng lớn cho thấy sản phẩm này có số lượng đánh giá cao, chứng tỏ được nhiều người mua quan tâm.

+ Tivi (Televisions): Đạt điểm đánh giá trung bình cao nhất (4.0), cho thấy sự hài lòng của khách hàng rất tốt. Tuy nhiên, tỷ lệ chiết khấu trung bình thấp hơn (~40%), và kích thước bong bóng nhỏ hơn, cho thấy số lượng người mua ít hơn so với các danh mục khác.

+ Camera và Phụ kiện camera (Cameras & Camera Accessories): Nằm ở khoảng giữa về cả tỷ lệ chiết khấu (~45%) và điểm đánh giá (~3.8). Kích thước bong bóng tương đối lớn, cho thấy sản phẩm này có mức độ phổ biến cao.

+ Hệ thống giải trí gia đình (Home Entertainment Systems): Có tỷ lệ chiết khấu trung bình khá (~45%), với điểm đánh giá khoảng 3.8. Kích thước bong bóng vừa phải, cho thấy mức độ quan tâm từ người mua không quá nổi bật.

+ Loa (Speakers) và Camera an ninh (Security Cameras): Điểm đánh giá thấp nhất trong các danh mục (~3.2–3.4). Tỷ lệ chiết khấu trung bình cũng thấp hơn so với các danh mục khác (~40–45%). Kích thước bong bóng nhỏ, cho thấy ít người mua hoặc ít nhận được sự chú ý.

+ Danh mục tổng hợp (All Electronics): Đại diện cho toàn bộ danh mục điện tử, với điểm đánh giá trung bình khoảng 3.8 và tỷ lệ chiết khấu ~45%. Kích thước bong bóng lớn nhất, vì nó bao gồm tổng số lượt đánh giá từ tất cả các danh mục con.

**• Tổng kết:**

+ Các sản phẩm có tỷ lệ chiết khấu cao như tai nghe thường thu hút nhiều lượt đánh giá nhưng không đảm bảo điểm số cao.

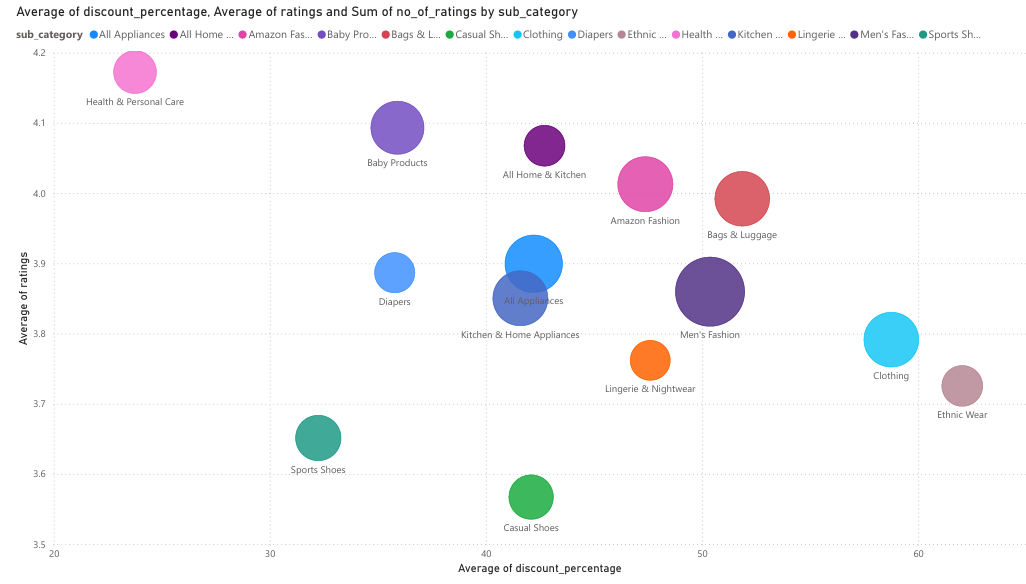
+ Tivi là danh mục có chất lượng vượt trội dựa trên điểm đánh giá, mặc dù không được giảm giá nhiều.

+ Các danh mục như loa và camera an ninh cần cải thiện chất lượng hoặc chiến lược tiếp thị để tăng sự hài lòng của khách hang

* 1. **Biểu đồ phân bố các danh mục phụ có nhiều lượt quan tâm nhất**

- Nhóm chọn sử dụng biểu đồ scatter plot để có thể biểu diễn được trung bình giảm giá (trục x), trung bình đánh giá (trục y), số lượt đánh giá(kích thước bong bóng) theo các danh mục phụ nhận được nhiều lượt đánh giá (14 danh mục)

- **Lý do**: Từ biểu này có thể quan sát được trung bình giảm giá có ảnh hưởng đến đánh giá của khách hàng đối với sản phẩm không.



- **Câu hỏi**:

+ Các mặt hàng được giảm giá nhiều thì có được quan tâm hơn hay không?

+ Giảm giá có ảnh hưởng đến đánh giá của khách hàng hay không?

+ Các chiến lược kinh doanh của từng ngành hang

- **Kết quả từ biểu đồ:**

+ Danh mục Health & Personal Care có xếp hạng trung bình cao nhất và phần trăm chiết khấu trung bình thấp nhất

+ Danh mục Sport shoes và casual shoes có phần trăm chiết khấu trung bình thấp và xếp hạng trung bình thấp.

+ Amazon fashion, Bags & Luggage có phần trăm chiết khấu trung bình và xếp hạng trung bình cao.

+ Các danh mục Clothing, Ethnic wear có phần trăm chiết khấu trung bình cao và xếp hạng trung bình thấp.

• **Tổng kết**

+ Phần trăm chiết khấu: Có một số mối tương quan giữa phần trăm chiết khấu trung bình và xếp hạng trung bình. Ví dụ: danh mục Sức khỏe & Chăm sóc cá nhân có xếp hạng trung bình cao nhất và phần trăm chiết khấu trung bình thấp nhất, trong khi các danh mục Quần áo và Trang phục dân tộc có phần trăm chiết khấu trung bình cao và xếp hạng trung bình thấp. Điều này có thể là do khách hàng có nhiều khả năng cho các sản phẩm được chiết khấu cao hơn xếp hạng thấp hơn, hoặc có thể là do các sản phẩm được chiết khấu cao hơn thường có chất lượng thấp hơn.

+ Chất lượng sản phẩm: Chất lượng sản phẩm là một yếu tố quan trọng có thể ảnh hưởng đến xếp hạng của sản phẩm. Các sản phẩm chất lượng cao có nhiều khả năng nhận được xếp hạng cao hơn, trong khi các sản phẩm chất lượng thấp có nhiều khả năng nhận được xếp hạng thấp hơn.

+ Sự hài lòng của khách hàng: Sự hài lòng của khách hàng là một yếu tố quan trọng khác có thể ảnh hưởng đến xếp hạng của sản phẩm. Khách hàng hài lòng với sản phẩm có nhiều khả năng cho sản phẩm đó xếp hạng cao hơn, trong khi khách hàng không hài lòng với sản phẩm có nhiều khả năng cho sản phẩm đó xếp hạng thấp hơn.

+ Sự quen thuộc của khách hàng: Khách hàng đã quen thuộc với sản phẩm có nhiều khả năng cho sản phẩm đó xếp hạng cao hơn, trong khi khách hàng không quen thuộc với sản phẩm có nhiều khả năng cho sản phẩm đó xếp hạng thấp hơn.

**2. Mối tương quan giữa đánh giá (Ratings) và tỷ lệ phần trăm giảm giá (%Discount\_price)**

Phân tích tiếp theo tập trung vào mối quan hệ giữa đánh giá (Ratings) và %Discount\_price theo ngành hàng chính (main\_category) và phân khúc giá. Dữ liệu cho thấy, trong các phân khúc giá thấp, tỷ lệ giảm giá cao thường thu hút các đánh giá tích cực. Ngược lại, ở phân khúc giá cao, yếu tố giảm giá không có tác động lớn đến sự thay đổi trong mức độ hài lòng.

Nhóm chúng tôi tiến hành nghiên cứu để tìm hiểu mối quan hệ giữa mức chiết khấu, mức độ tương tác, và đánh giá của người tiêu dùng theo từng phân khúc giá sản phẩm. Bước đầu tiên là áp dụng thuật toán K-means clustering để phân chia các sản phẩm thành ba mức giá cụ thể:

● Phân khúc giá rẻ: 0 - 13,449

● Phân khúc giá bình thường: 13,449 - 42,899

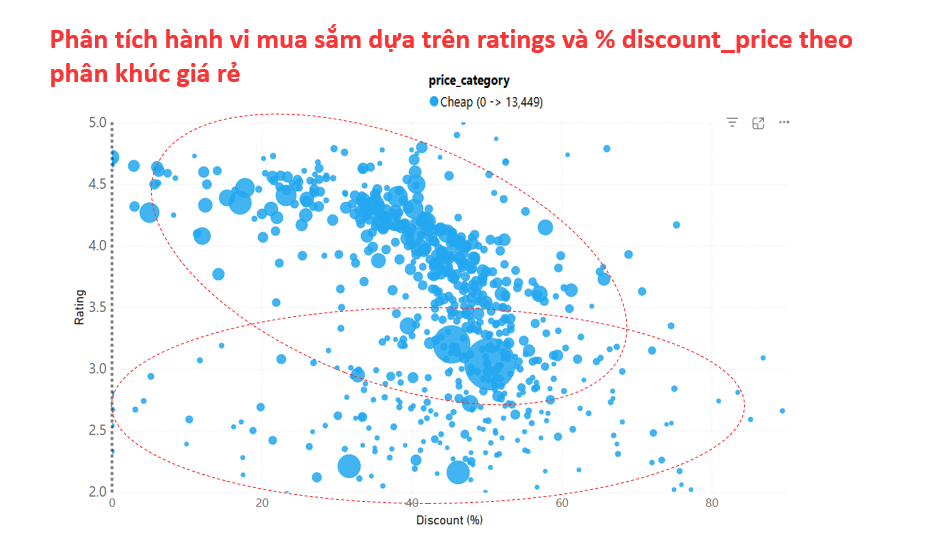
● Phân khúc giá đắt: Trên 42,899

**Kết quả phân tích chi tiết**

1. **Phân khúc giá rẻ:**

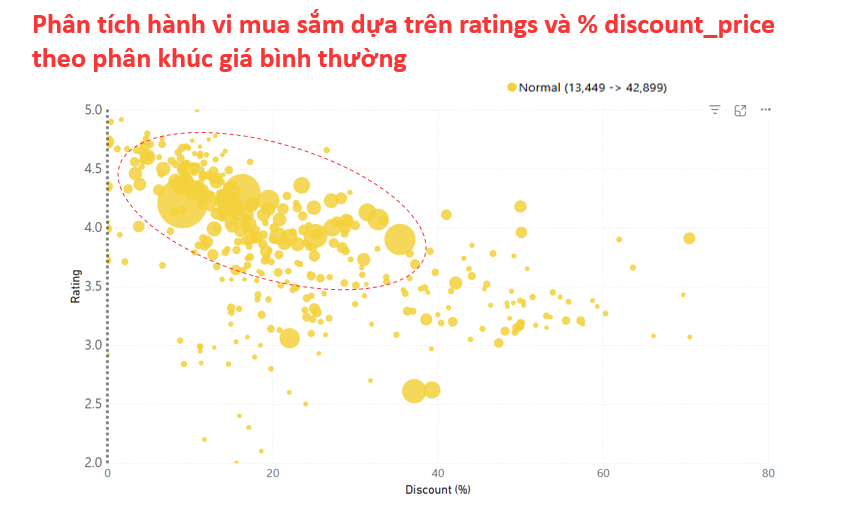
● Các sản phẩm giá rẻ thường được chiết khấu cao hơn, nhưng điều này lại đi kèm với xu hướng xếp hạng thấp hơn

● Dữ liệu này có mức độ phân tán lớn, cho thấy sự khác biệt rõ rệt trong đánh giá người tiêu dùng. Chiết khấu cao có thể làm tăng sự quan tâm và mua hàng, nhưng cũng có nguy cơ ảnh hưởng tiêu cực đến nhận thức về chất lượng sản phẩm.



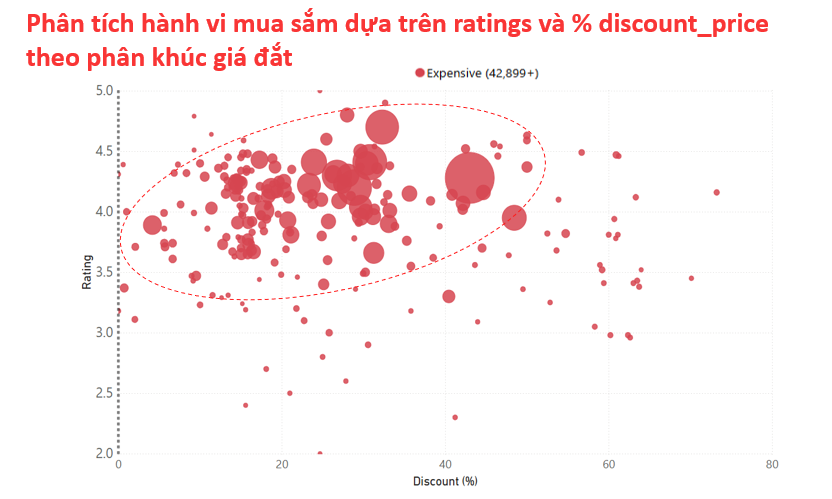
2. **Phân khúc giá bình thường:**

● Các sản phẩm trong phân khúc này có xu hướng tập trung ở mức chiết khấu thấp, với xếp hạng trung bình đến cao. Điều này cho thấy người tiêu dùng ở phân khúc này chú trọng vào chất lượng sản phẩm thay vì chỉ dựa vào mức ưu đãi.



3. **Phân khúc giá đắt:**

● Các sản phẩm giá đắt thường đi kèm với chiết khấu thấp và xếp hạng cao. Mặc dù số lượng sản phẩm trong phân khúc này ít, ưu đãi hợp lý vẫn đóng vai trò quan trọng trong việc thúc đẩy quyết định mua sắm.



**Các góc nhìn chiến lược**

**1. Ảnh hưởng của chiết khấu đến giá trị cảm nhận:**

● Phân khúc giá cao: Người tiêu dùng sẵn sàng chấp nhận chiết khấu thấp nếu họ tin rằng sản phẩm mang lại giá trị cao. Tuy nhiên, ưu đãi giá vẫn là yếu tố quan trọng để thúc đẩy quyết định mua hàng.

● Phân khúc giá rẻ: Chiết khấu cao có thể là yếu tố hấp dẫn, nhưng dễ làm người tiêu dùng cảm thấy chất lượng sản phẩm kém. Điều này đòi hỏi một sự cân nhắc kỹ lưỡng khi sử dụng chiết khấu lớn.

**2. Tâm lý khách hàng và mức giảm giá tối ưu:**

● Đối với các sản phẩm giá rẻ, người tiêu dùng thường ít đắn đo hơn do rủi ro tài chính thấp. Tuy nhiên, nếu mức chiết khấu vượt quá 60%, hiệu quả dường như giảm, vì khách hàng có thể bắt đầu nghi ngờ về chất lượng thực sự của sản phẩm.

● Với các sản phẩm giá bình thường, khách hàng đề cao sự cân bằng giữa giá cả và chất lượng. Một mức ưu đãi vừa phải, đi kèm với sự ổn định về chất lượng, sẽ xây dựng lòng tin lâu dài.

● Ở phân khúc giá cao, khi mức chiết khấu hợp lý kết hợp cùng đánh giá tích cực, sẽ tạo động lực mạnh mẽ để người tiêu dùng đưa ra quyết định nhanh chóng.

**3. Mức độ tương tác và chiến lược khuyến mãi:**

● Phân khúc giá rẻ và giá cao ghi nhận mức độ tương tác cao hơn khi chiết khấu tăng, nhưng với phân khúc giá bình thường, ưu đãi thấp dường như hiệu quả hơn. Điều này nhấn mạnh sự cần thiết của các chiến lược chiết khấu linh hoạt cho từng phân khúc.

**Kết luận**

Biểu đồ và phân tích của chúng tôi cho thấy rằng mức chiết khấu cần được điều chỉnh phù hợp với từng phân khúc giá. Đối với sản phẩm giá rẻ, chiết khấu cao có thể thu hút khách hàng, nhưng cần được cân nhắc để tránh làm mất giá trị sản phẩm. Với sản phẩm giá bình thường, chiến lược ưu tiên chất lượng và sự ổn định giá cả được đánh giá cao. Cuối cùng, sản phẩm giá đắt có thể tối ưu hóa doanh số bằng cách kết hợp giữa ưu đãi giá hợp lý và đánh giá tích cực từ khách hàng

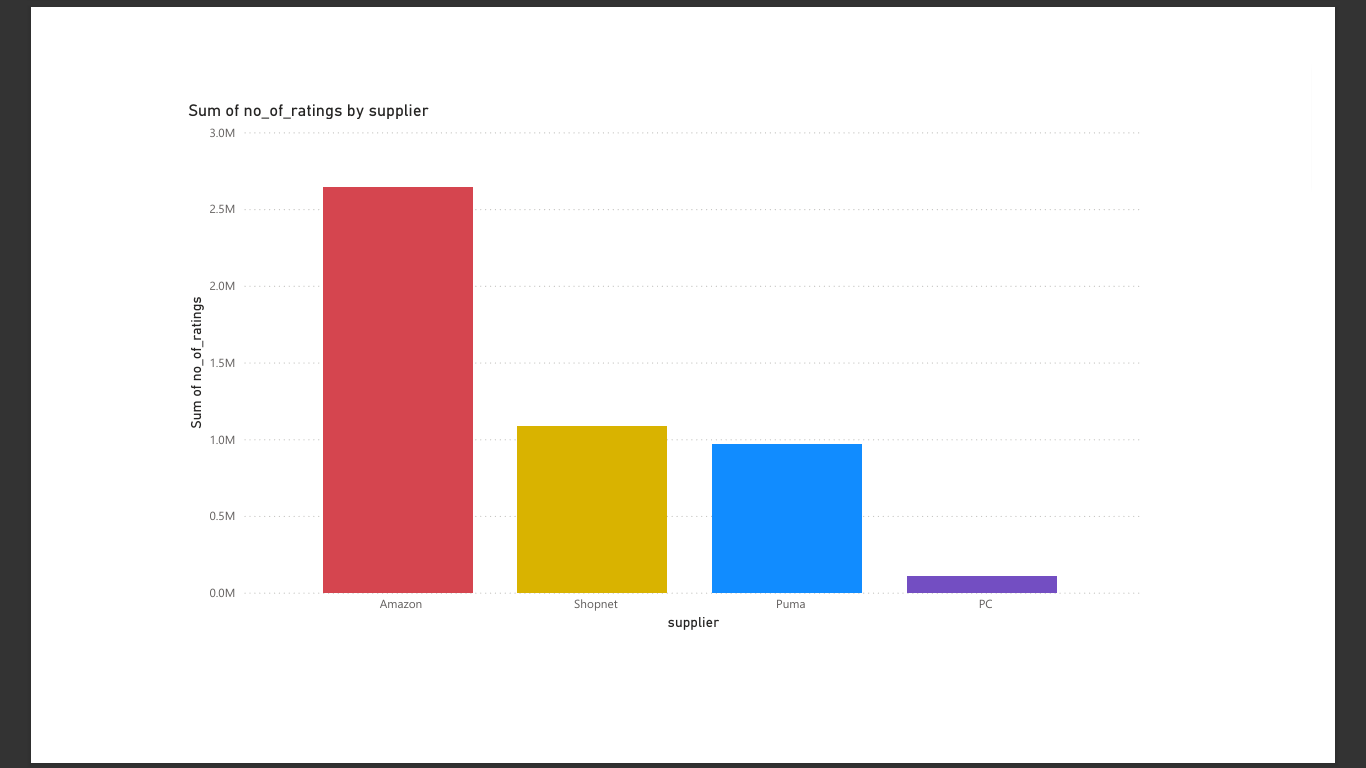
**3. Số lượng đánh giá (No\_of\_ratings) và tỷ lệ giảm giá (%Discount\_price) theo nhà cung cấp (supplier)**

Khi phân tích mối quan hệ giữa No\_of\_ratings và %Discount\_price theo từng nhà cung cấp, nhóm nghiên cứu nhận thấy rằng các nhà cung cấp có chiến lược giảm giá linh hoạt thường nhận được lượng đánh giá cao hơn. Điều này cho thấy sự sáng tạo trong chiến lược giảm giá có thể đóng vai trò quan trọng trong việc tăng cường tương tác với khách hàng.

**1. Biểu đồ xếp hạng số lượt đánh giá theo nhà cung cấp**

- Nhóm chọn biểu đồ cột để so sánh số lượng đánh giá (trục X) giữa các nhà cung cấp (trục Y)

- **Lý do**: Từ biểu này có thể quan sát được trung bình giảm giá có ảnh hưởng đến đánh giá của khách hàng đối với sản phẩm không.



- **Câu hỏi**: Nhà cung cấp nào có lượt đánh giá cao nhất

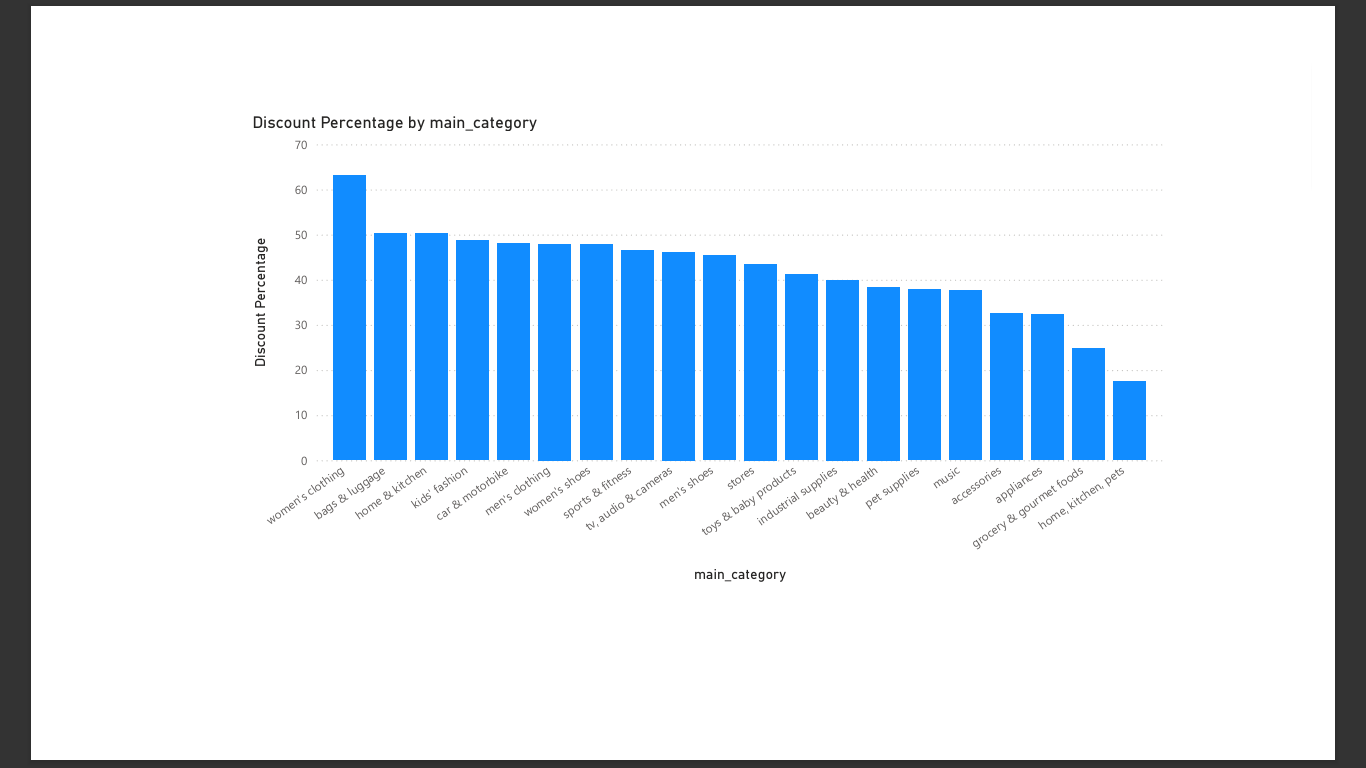
- **Kết quả** Amazon là nhà cung cấp được đánh giá nhiều nhất vượt trội khoảng hơn 2,5 triệu lượt gần như gấp đôi nhà cung cấp đứng thứ 2 khoảng hơn 1 triệu lượt và thứ 3 khoảng gần 1 triệu lượt.

**• Kết luận:** Chúng ta thấy được sự hiện diện và mức độ phổ biến của Amazon

**2. Biểu đồ xếp hạng phần trăm chiết khấu theo từng danh mục:**

- Nhóm chọn biểu đồ cột để xem phần trăm chiết khấu trung bình (trục X) trong từng danh mục sản phẩm (trục Y)

- **Lý do**: Dễ quan sát được danh mục nào chiết khấu cao nhất , danh mục chiết khấu thấp nhất , chiết khấu dao động từ bao nhiêu đến bao nhiêu



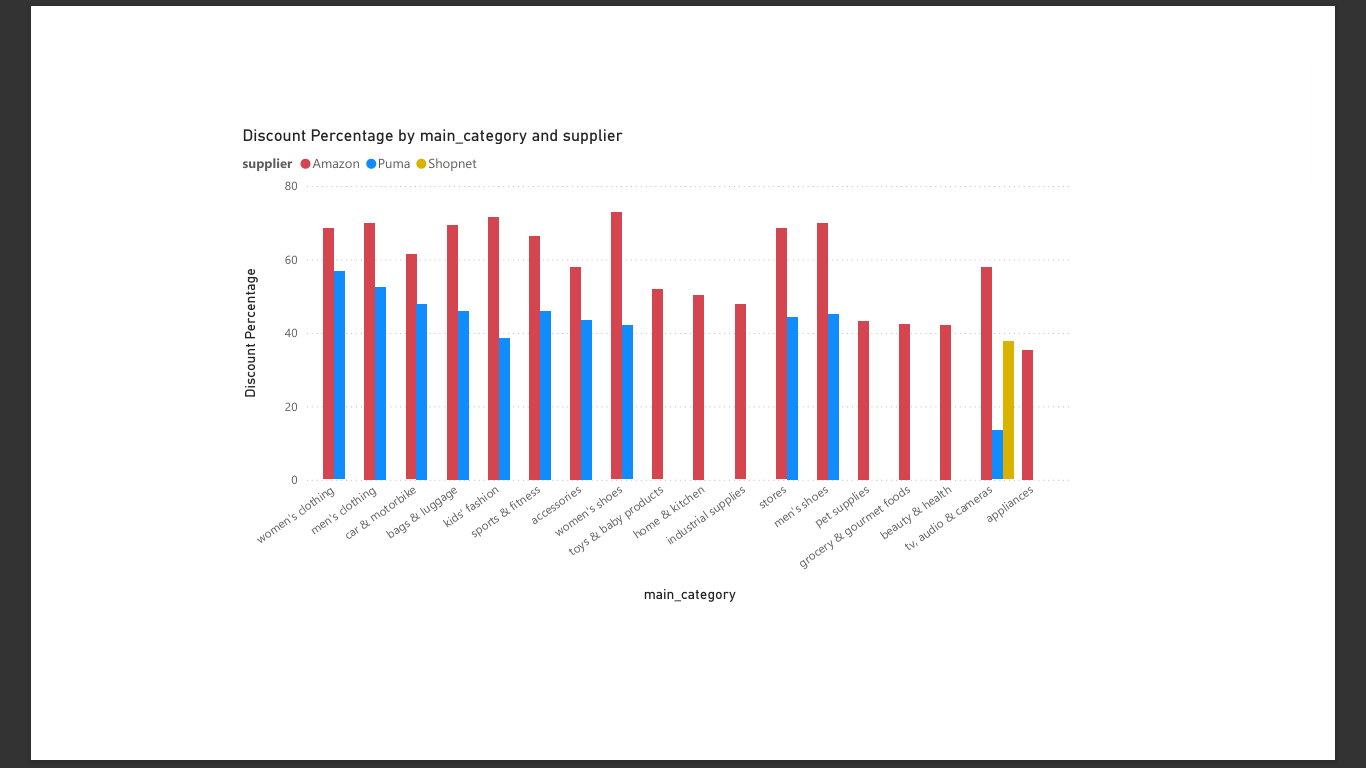
- **Câu hỏi**: Danh mục nào chiết khấu cao nhất , danh mục nào thấp nhất và dao động từ bao nhiêu đến bao nhiêu

- **Kết quả**: Quần áo phụ nữ có phần trăm chiết khấu cao nhất khoảng hơn 60% , danh mục chiết khấu thấp nhất home,kitchen,pets khoảng 18%

- **Kết luận**: Có một xu hướng giảm dần di chuyển dọc theo các danh mục

**3. Biểu đồ xếp hạng phần trăm chiết khấu từng danh mục theo 3 nhà cung cấp có lượt đánh giá cao nhất :**

- Nhóm chọn biểu đồ cột để xem phần trăm chiết khấu (trục X) theo từng danh mục (trục Y) và so sánh phần trăm chiết khấu trong từng danh mục theo nhà cung cấp

- **Lý do**: Để dễ quan sát phần trăm chiết khấu , các danh mục của các nhà cung cấp

- **Câu hỏi**: Các nhà cung cấp chiết khấu nhiều hay ít , các danh mục nào

- **Kết quả**: Amazon chiết khấu hầu hết các danh mục trừ danh mục âm nhạc và so sánh với các nhà cung cấp còn lại thì vượt trội hẳn . Trong khi đó nhà cung cấp Shopnet chỉ tập chung vào danh mục tv,audio và cameras và phầm trăm chiết khấu là khoảng 37,55% so với chiết khấu trung bình của danh mục này là khoảng 46,12% cho thấy được phần trăm chiết khẩu của họ khá ổn . Đối với nhà cung cấp Puma họ chiết khấu cũng rất nhiều danh mục nhưng chủ yếu tập chung vào ngành thời trang đồng thời tham gia vào danh mục tv,audio và cameras nhưng họ chiết khấu khá thấp chỉ 13.34%

- **Kết luận**: Amazon có tỷ lệ chiết khấu cao và nhiều danh mục hơn so với 2 nhà cung cấp còn lại

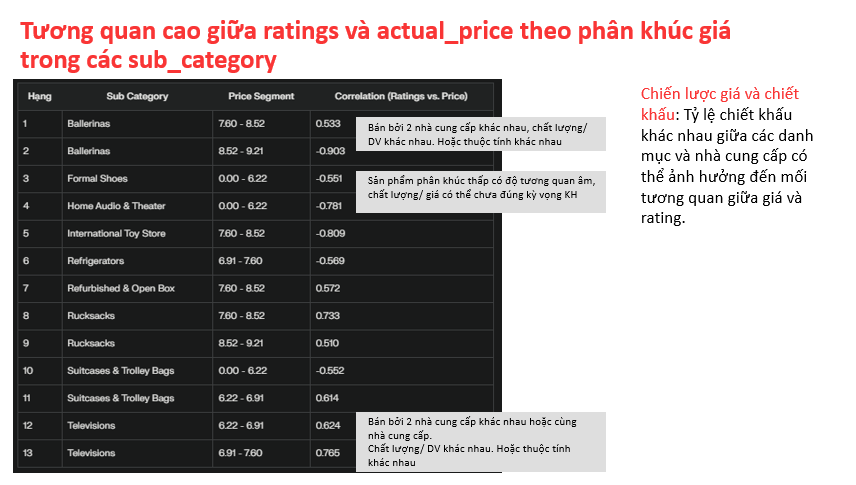
- **Kết luận:** Ta thấy được giữa số lượt đánh giá và phần trăm chiết khấu có tương quan với nhau như Amazon chiết khấu cao vào chiết khấu nhiều danh mục nhất thì có lượt rating nhiều nhất.

- **Tổng kết:** Shopnet – Tập Trung Đồ Điện Tử Lượng ratings cao thứ 2, cho thấy khách hàng có mức độ tương tác tốt. Mức giảm giá trung bình 46%, khá cao, cho thấy họ có chiến lược giảm giá mạnh để cạnh tranh. Tập trung vào đồ điện tử, điều này có thể giúp họ chuyên sâu hơn, nhưng cũng có rủi ro nếu thị trường này biến động mạnh. Chiến lược đề xuất: Tận dụng mức độ quan tâm cao để đẩy mạnh doanh số bằng các chương trình up-sell & cross-sell. Xem xét giảm bớt mức chiết khấu nếu biên lợi nhuận bị ảnh hưởng. Puma – Chủ Yếu Về Quần Áo Lượng ratings cao thứ 3, có sự quan tâm đáng kể từ khách hàng. Giảm giá không quá cao nhưng bao phủ nhiều danh mục, cho thấy chiến lược đa dạng sản phẩm. Chiến lược đề xuất: Tập trung hơn vào danh mục quần áo có lợi nhuận cao nhất. Tạo các chương trình bán combo hoặc ưu đãi theo mùa để tăng doanh thu. Amazon – Trải Rộng Tất Cả Danh Mục Ratings và giảm giá nhiều hơn trung bình, nhưng tổng giá trị các danh mục thấp. Chiến lược “đánh rộng” có thể giúp Amazon chiếm lĩnh thị trường, nhưng doanh thu trên từng danh mục không cao. Chiến lược đề xuất: Xác định danh mục có tiềm năng tăng trưởng cao nhất để tập trung đầu tư. Tận dụng lượng ratings cao để quảng bá sản phẩm nổi bật hơn. Kiểm tra lại chiến lược giá để xem có nên điều chỉnh mức giảm giá hay không.

**4. Mối tương quan giữa đánh giá (Ratings) và giá thực tế (actual price)**

Cuối cùng, nhóm nghiên cứu đánh giá mối tương quan giữa Ratings và actual\_price. Kết quả chỉ ra rằng sản phẩm có giá thực tế cao và được giảm giá sâu thường nhận được đánh giá tích cực, nhưng hiệu ứng này không phải lúc nào cũng đúng với các sản phẩm giá thấp.

Mục tiêu của việc phân tích là làm rõ mối quan hệ phức tạp giữa các yếu tố ảnh hưởng đến chiến lược giảm giá trong thương mại điện tử. Các doanh nghiệp cần hiểu rõ đặc điểm của ngành hàng, khách hàng mục tiêu và xu hướng đánh giá để đưa ra quyết định giảm giá hiệu quả. Một chiến lược giảm giá tốt không chỉ giúp gia tăng doanh thu mà còn củng cố mối quan hệ lâu dài với khách hàng.



Mối Tương Quan Giữa Giá và Rating Trong Phân Khúc Sản Phẩm Thấp

Mối quan hệ giữa giá cả và đánh giá (rating) trong phân khúc sản phẩm giá thấp là một chủ đề phức tạp, với nhiều yếu tố ảnh hưởng đáng kể. Các kết quả nghiên cứu cho thấy độ tương quan âm trong phân khúc này có thể phản ánh sự bất cân bằng giữa giá trị kỳ vọng của khách hàng và giá trị thực tế mà sản phẩm mang lại.

**1. Chất lượng và giá cả không đáp ứng kỳ vọng của khách hàng**

Trong nhiều trường hợp, sản phẩm giá thấp thường phải đánh đổi về mặt chất lượng để duy trì mức giá cạnh tranh. Khi kỳ vọng của khách hàng không được đáp ứng, họ dễ có xu hướng để lại các đánh giá tiêu cực, từ đó góp phần làm tăng độ tương quan âm giữa giá cả và rating.

**2. Tác động từ sự khác biệt trong chất lượng và dịch vụ**

Không phải tất cả sản phẩm giá thấp đều được tạo ra giống nhau. Những sự khác biệt trong chất lượng sản phẩm, dịch vụ hỗ trợ khách hàng, và chính sách bảo hành có thể ảnh hưởng lớn đến cảm nhận của khách hàng. Một số nhà cung cấp có thể cung cấp dịch vụ vượt trội, trong khi những nhà cung cấp khác không đảm bảo được mức độ phục vụ tối thiểu, dẫn đến các phản hồi trái chiều.

**3. Đặc điểm của sản phẩm và nhà cung cấp**

Các thuộc tính của sản phẩm, từ thiết kế, tính năng cho đến mức độ sử dụng thực tế, đóng vai trò lớn trong quyết định của khách hàng. Đồng thời, việc sản phẩm được cung cấp bởi các nhà cung cấp khác nhau cũng là yếu tố quan trọng. Các nhà cung cấp uy tín có thể đạt được đánh giá tích cực hơn, ngay cả với sản phẩm tương tự, nhờ cam kết về chất lượng và dịch vụ.

**4. Ảnh hưởng của chiến lược giá và chiết khấu**

Kết luận về ảnh hưởng của chiến lược giá và chiết khấu đến mối quan hệ giữa giá và rating được củng cố khi xem xét các yếu tố như chất lượng, dịch vụ, và đặc điểm sản phẩm. Chiến lược chiết khấu tốt, khi đi kèm với chất lượng đảm bảo, không chỉ giúp sản phẩm nổi bật hơn mà còn tăng mức độ hài lòng của khách hàng, từ đó cải thiện rating tổng thể.

⇒ Mức độ thành công của sản phẩm trên Amazon phụ thuộc vào sự kết hợp phức tạp giữa nhiều yếu tố, bao gồm giá cả, chiết khấu, chất lượng sản phẩm, nhận xét của khách hàng, và chiến lược của nhà cung cấp. Amazon chiếm ưu thế về số lượng ratings, nhưng các nhà cung cấp khác có thể cạnh tranh bằng cách tối ưu hóa chiến lược giá và tập trung vào các danh mục cụ thể.

**Insights**:

**Số lượng ratings quan trọng**: Số lượng sản phẩm lớn hơn có xu hướng thu hút nhiều đánh giá hơn, điều này mang lại lợi thế cho Amazon nhờ quy mô và độ phổ biến của nền tảng.

**Chiết khấu không đảm bảo thành công**: Tỷ lệ chiết khấu cao không luôn dẫn đến đánh giá tốt hơn; chất lượng sản phẩm và giá trị cảm nhận đóng vai trò quan trọng hơn trong mắt khách hàng.

**Phân khúc giá ảnh hưởng đến kỳ vọng**: Sản phẩm giá rẻ thu hút nhiều người mua bất kể đánh giá thấp. Sản phẩm đắt tiền phải đảm bảo chất lượng cao để đáp ứng kỳ vọng.

**Danh mục "Health & Personal Care" nổi bật**: Đây là danh mục có xếp hạng trung bình cao nhất và tỷ lệ chiết khấu thấp, cho thấy khách hàng sẵn lòng trả giá cao cho chất lượng.

**"Quần áo nữ" và "Túi xách & hành lý" cạnh tranh**: Các danh mục này có tỷ lệ chiết khấu cao do tính cạnh tranh, nhưng vẫn cần đảm bảo không làm suy giảm chất lượng sản phẩm.

**Chiến lược giá linh hoạt**: Một chiến lược giá hiệu quả phải cân nhắc đến phân khúc giá, chất lượng sản phẩm và kỳ vọng khách hàng để đạt được thành công bền vững.

**Chiến lược cho doanh nghiệp**

#### **1. Tối ưu hóa chiến lược chiết khấu**

Để tối ưu hóa chiến lược chiết khấu, việc thực hiện các phân tích chuyên sâu là cần thiết:

* **Phân tích tỷ lệ chiết khấu theo danh mục và nhà cung cấp** nhằm xác định cơ hội nâng cao hiệu quả.
* Đảm bảo **chất lượng sản phẩm được giảm giá cao** không bị suy giảm để tránh nhận các phản hồi tiêu cực từ khách hàng.
* **Thử nghiệm các mức chiết khấu khác nhau** để tối ưu hóa doanh số và cải thiện sự hài lòng của khách hàng, tạo ra sự cân bằng giữa giá trị và lợi nhuận.

#### **2. Tập trung vào chất lượng sản phẩm**

Chất lượng sản phẩm đóng vai trò quan trọng trong việc xây dựng lòng tin và tạo lòng trung thành từ khách hàng:

* **Kiểm soát chất lượng nghiêm ngặt** để đảm bảo tất cả sản phẩm đạt tiêu chuẩn cao nhất.
* **Thu thập phản hồi từ khách hàng** thông qua khảo sát và nhận xét để liên tục cải tiến.
* Tập trung **nâng cấp các sản phẩm có đánh giá thấp**, cải thiện trải nghiệm khách hàng và tăng khả năng giữ chân người mua.

#### **3. Phân tích hành vi khách hàng**

Hiểu rõ hành vi và kỳ vọng của khách hàng giúp xây dựng các chiến lược hiệu quả:

* **Theo dõi dữ liệu bán hàng** để phát hiện các xu hướng tiêu dùng.
* **Phân tích đánh giá và nhận xét** từ khách hàng để nắm bắt sở thích và điểm hạn chế của sản phẩm.
* Sử dụng **A/B testing** để kiểm tra các chiến lược giá, chiết khấu, và tiếp thị khác nhau nhằm xác định phương pháp tối ưu.

#### **4. Tận dụng lợi thế của Amazon**

Amazon mang đến một hệ sinh thái độc đáo giúp các doanh nghiệp tiếp cận hiệu quả hơn:

* **Tối ưu hóa danh sách sản phẩm** với từ khóa và hình ảnh chất lượng để tăng khả năng hiển thị.
* Cung cấp sản phẩm qua **Amazon Prime** để nâng cao trải nghiệm khách hàng và xây dựng độ tin cậy.
* Tận dụng **quảng cáo của Amazon** để mở rộng thị trường và thu hút thêm nhiều người mua tiềm năng.

#### **5. Tập trung vào phân khúc "Health & Personal Care"**

Danh mục này thể hiện tiềm năng lớn với khách hàng sẵn sàng chi trả cho chất lượng:

* Tiến hành **nghiên cứu thị trường** nhằm tìm kiếm các sản phẩm sáng tạo và hấp dẫn.
* **Xây dựng thương hiệu mạnh mẽ và đáng tin cậy** trong lĩnh vực chăm sóc sức khỏe và cá nhân.
* Cung cấp **dịch vụ khách hàng xuất sắc** để gia tăng lòng trung thành và sự tín nhiệm từ phía người tiêu dùng.

#### **6. Cải thiện các danh mục có xếp hạng thấp**

Xử lý các danh mục có xếp hạng thấp là cách hiệu quả để xây dựng hình ảnh tích cực hơn:

* **Khảo sát khách hàng** để tìm hiểu lý do khiến đánh giá sản phẩm không cao.
* Hợp tác chặt chẽ với nhà cung cấp để **nâng cao chất lượng sản phẩm**.
* Cân nhắc **tái định vị sản phẩm hoặc thay đổi chiến lược tiếp thị** nhằm làm mới hình ảnh và thu hút khách hàng mới.

#### **7. Thương hiệu và tiếp thị**

Xây dựng thương hiệu là yếu tố không thể thiếu trong việc cạnh tranh trên thị trường:

* **Truyền tải giá trị thương hiệu** thông qua các chiến dịch sáng tạo và hấp dẫn.
* Nhấn mạnh chất lượng, tính độc đáo, và các lợi ích khác biệt của sản phẩm để tạo dấu ấn với khách hàng.

**Limitations:**

**1. Dữ liệu không hoàn chỉnh**  
Chất lượng dữ liệu là một yếu tố quan trọng. Trong quá trình phân tích, chúng tôi nhận thấy rằng việc xử lý dữ liệu còn một số hạn chế, chẳng hạn như dữ liệu thiếu hoặc không đúng định dạng. Điều này đòi hỏi chúng tôi phải áp dụng các phương pháp xử lý dữ liệu chuyên sâu hơn để nâng cao độ tin cậy và tính toàn vẹn của kết quả.

**2. Kiến thức hạn chế về lĩnh vực bán hàng trực tuyến**  
Sự hiểu biết sâu sắc về lĩnh vực bán hàng trực tuyến là rất quan trọng. Chúng tôi nhận thấy rằng việc nâng cao kiến thức chuyên môn trong lĩnh vực này có thể giúp chúng tôi phân tích và diễn giải dữ liệu một cách hiệu quả hơn, đồng thời nắm bắt các yếu tố quan trọng ảnh hưởng đến hành vi khách hàng và xu hướng thị trường.

**3. Hạn chế về kỹ năng lập trình**  
Kỹ năng lập trình là một yếu tố không thể thiếu. Chúng tôi nhận thấy rằng việc nâng cao khả năng viết code phức tạp sẽ giúp chúng tôi kiểm tra và điều chỉnh mã nguồn một cách linh hoạt hơn, từ đó tối ưu hóa hiệu quả của các thuật toán và phương pháp phân tích được áp dụng.