Data-Driven Insights for E-commerce Revenue Growth

# Executive Summary

Company A is a fashion retailer with stores in London, Paris, Milan, and Berlin. Customers can purchase products in-store or online. The company also maintains an e-commerce website with product information and images. In many cases, website traffic increases when new products are launched.

Management has requested a data analysis report to answer the following questions:

1. Does website traffic affect revenue?
2. Which products generate the highest pageviews and revenue?
3. What are the different customer groups?

# Methodology

**Data Sources**

|  |  |  |
| --- | --- | --- |
| Source | Description | Link |
| Customers & Transactions | Customer information and transactions, both online and in-store | Customers |
| Website Traffic | Website traffic data for each product page (Duration: Jan to Dec 2020) | Traffic |

# Data Analysis

**1. Website Traffic and Revenue**

* **Hypothesis:** Website traffic has a positive impact on revenue.
* **Analysis:**
  + **Define variables:**
    - **Dependent variable:** Revenue
    - **Independent variable:** Website traffic
  + **Model:** Linear regression with Statsmodels
  + **Results:**
    - R-squared: 0.000
    - P-value: 0.375
  + **Interpretation:**
    - The R-squared value indicates the correlation between website traffic and sales. In this case, an R-squared of zero indicates no linear relationship between the two variables.
    - The p-value indicates the significance of the regression coefficient. A p-value greater than 0.05 (the usual significance threshold) indicates insufficient evidence to conclude that the regression coefficient is different from zero. This means that website traffic may not significantly affect sales in this model.

Based on the linear regression analysis results, it cannot be definitively concluded that website traffic has an effect on sales (SellPrice).

**2. Products with Highest Pageviews and Revenue**

* **Analysis:**
  + **Merge customer transactions and website traffic data.**
  + **Calculate pageviews and revenue for each product.**
  + **Identify the products with the highest pageviews and revenue.**
* **Results:**
  + Top 5 products by pageviews:

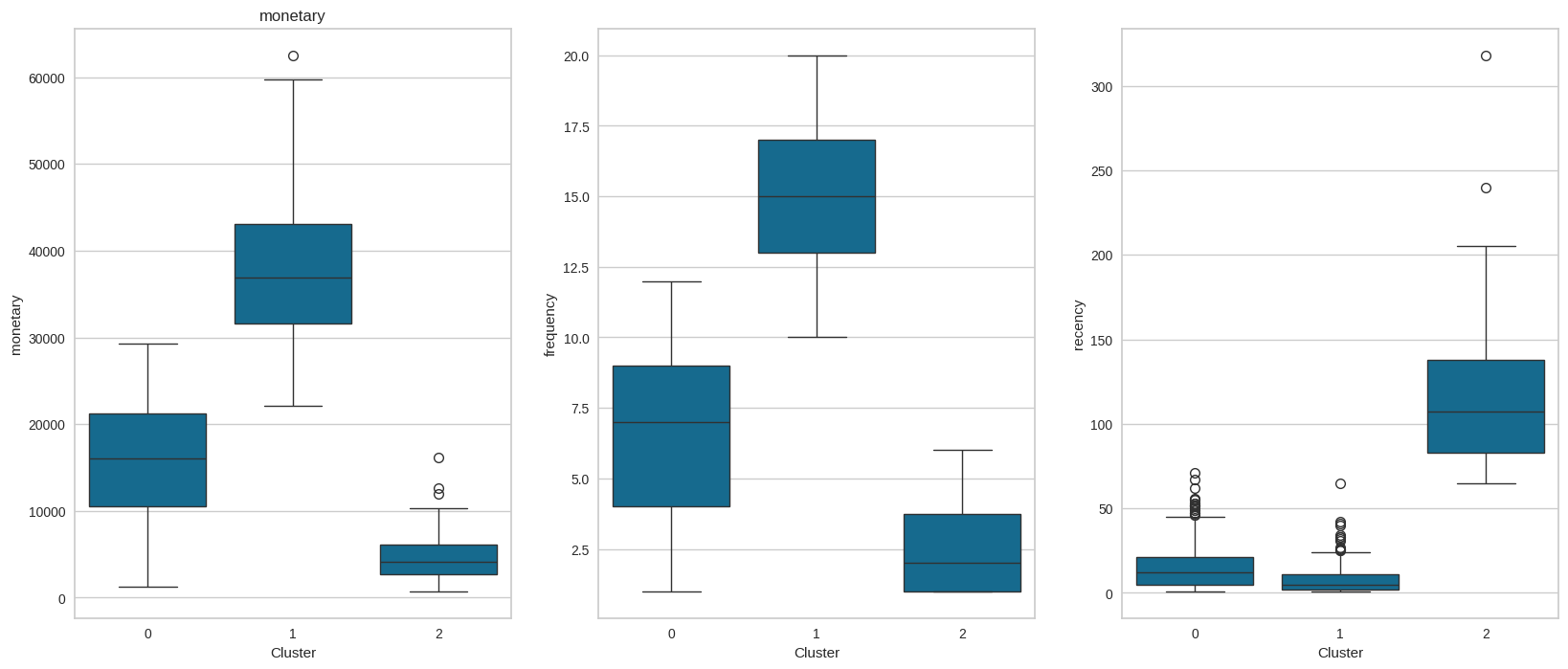
|  |  |  |
| --- | --- | --- |
|  | **Product** | **total\_pageviews** |
| **0** | sonEies-x0.139039192479287 | 215222.6 |
| **1** | porQlease0.507911745358356 | 155640.0 |
| **2** | tobH-info0.428382841646979 | 65819.4 |
| **3** | ninI-info0.46275630306324 | 60351.8 |
| **4** | crow-home0.0127810060365789 | 54677.8 |
| **5** | besfeek-50.280274313298435 | 53508.2 |
| **6** | nobHlease0.20690948418157 | 53041.2 |
| **7** | rogLrecap0.644396392507623 | 52381.2 |
| **8** | pumMvideo0.837931992780922 | 47779.6 |
| **9** | jefXaunch0.410067394635876 | 43667.4 |

* + Top 5 products by revenue:

|  |  |  |
| --- | --- | --- |
|  | **Product** | **total\_revenue** |
| **0** | hypsrview0.321288570724117 | 11556 |
| **1** | reeelease0.452821711209563 | 9205 |
| **2** | audE-info0.884915261087885 | 8428 |
| **3** | kinv-news0.29905739542661 | 7980 |
| **4** | ymc\_orson0.319251813809483 | 7794 |
| **5** | bmwk-20210.31690704020655 | 7686 |
| **6** | ubeYanada0.100113795307302 | 7552 |
| **7** | titM-info0.550757203954225 | 7520 |
| **8** | samLement0.78565851365938 | 7518 |
| **9** | dioa-info0.750756927974235 | 7422 |

**3. Customer Groups**

* **Analysis:**
  + **Perform customer segmentation using clustering algorithms.**
  + **Identify the key characteristics of each customer group.**
* **Results:**



* + Three customer groups were identified:

**Group 1:**

* Average spending: ~15,000
* Average number of purchases: ~7 times
* Recent purchase time: Within a reasonable timeframe

**Group 2:**

* Highest spending: ~35,000
* Most frequent purchases: ~15 times
* Most recent purchase time: Most recent

**Group 3:**

* Lowest spending: ~5,000
* Least frequent purchases: ~25 times
* Longest time since last purchase: ~100 days, with some customers not returning for nearly a year (considered lost)

**Additional Analysis:**

Group 1: These customers have a moderate spending level and purchase frequency. They are likely to be regular customers who are familiar with products or services.

Group 2: These customers are high-value customers. They spend the most money and make the most purchases. It is important to retain these customers and encourage them to continue spending.

Group 3: These customers have the lowest spending level and purchase frequency. They may be new customers or customers who have not been active in a while. It is important to reactivate these customers and encourage them to make more purchases.

**Recommendations**

* **Increase website traffic by:**
  + Investing in search engine optimization (SEO).
  + Running social media campaigns.
  + Implementing email marketing campaigns.
* **Promote high-margin products:**
  + Feature high-margin products in store.
  + Offer discounts and promotions on high-margin products.
* **Target specific customer groups with personalized marketing campaigns:**
  + Send targeted email campaigns to each customer group.
  + Create personalized product recommendations for each customer group.

**Conclusion**

This data analysis report has shown that website traffic does not impact on revenue. The company can increase revenue by increasing website promotion, promoting high-margin products, and targeting specific customer groups with personalized marketing campaigns.