**BUSINESS PROBLEM**

Kiosk an e-commerce platform tasked us with analyzing their day-to-day sales activities and deducing patterns, trends, and business insights meant to:

1. Understand user preferences and user behavior
2. Increase success probability from targeted marketing
3. Optimize shopping experience
4. Enhance customer satisfaction, and
5. Drive business growth.

**DATA OVERVIEW**

The dataset was sourced from Kaggle. It has two tables, df1 and df2 which we analyzed differently.

Df1 includes extensive data, considering it is an e-commerce platform which requires numerous data such as: Age, Gender, Payment Method, Previous Purchases, Promo Code Used, and many other data points we will see.

These data points are crucial as they help us map the customer journey thus we can plot to make it more efficient and engaging.

**DATA CLEANING**

The dataset had no null values and no data in need of transformation.

**STATISTICAL DATA ANALYSIS**

From the descriptive analysis, we deduced:

1. The Mean and middle age of the users is 44 years, the youngest user is 18 years while the eldest at 70 years
2. The mean price for the products is $59, the cheapest at 20 and the most expensive at 100
3. The highest possible rating is 5, the mean rate point is 3.7 and the lowest rate point is 2.5.
4. The average number of previous purchases amongst the users is 25, the maximum number of previous purchases from a user is 50 and the least is 1.

We analyzed each field separately:

1. **Age**

* The most popular age among customers is 69
* The least popular age among customers is 44
* Age group 49 generated the highest revenues at $5,552
* Age group 67 generated the least revenues at $3,258

2. **Items purchased**

* Blouses generated the highest revenues at $10,410
* Jeans generated the least revenue at $ 7,548
* The most selling items were for women
* Blouses, Jewelry, and Pants were the fastest-moving goods with 171 pieces sold
* Jeans were the slowest-moving goods with only 124 pieces sold
* Jeans**,** gloves**,** and sneakers generated the least revenue and were also slow movers. Perhaps Kiosk should launch attractive promotions/discounts for these items to drive sales.

3. **Categories**

* The clothing category was the bestseller, raking up $104,264
* Outwear generated the least revenue by far at $ 18,524
* The clothing category was the fastest moving as well at 1737 pieces sold while only 324 outwear pieces were sold.
* We can say the faster and more the sales, the more revenues.

4. **Location** - we will use state abbreviations

* The highest revenues were raked up from MT, IL, CA, ID, and NV respectively
* The smallest revenues were generated from KS, HI, FL, NJ, and RI respectively
* Locations with the highest number of customers are MT, ID, IL, CA, and AL respectively
* Locations with the least number of customers are HI, RI, FL, and MA respectively.
* The states with the highest revenues have the highest customer count as well and the opposite is true as well from the analysis.
* Using these insights, Kiosk can effectively plan how to serve these areas effectively with the intent of cost-cutting, efficiency, and time management taking into consideration the distances apart from each other.
* Central fulfillment centers could be designed to serve certain markets that complement each other.
* As we will later see, customers value short delivery times thus it would be efficient to consider this factor while setting up fulfillment centers around areas with high customer count.

5. **SIZE**

* The most popular size bought was Medium with 1755 purchases followed by L with 1053 pieces sold, while XL was the least popular with 429 pieces sold.
* Kiosk should always have Medium sizes stocked in plenty and small numbers of XL in anticipation of good sales
* There were only two categories that XL performed well which are Sneakers and Shirts and thus should be more in inventory than size S.
* The size insights help a great deal in good inventory management and procurement prioritization which trickle down to increased profits and efficient use of storage space.

**SEASON**

* The Fall season has the highest revenues earned at $ 60,018 with Summer generating the least at $55,777.
* People normally prepare for winter during the fall season thus the high sales.
* It’s evident that people normally prepare ahead for a season considering the good sales in Spring at $58,679 and Fall at $ 60,018
* T-shirts were best-sellers across all Seasons thus t-shirts should always be in plenty.
* Backpacks were the worst selling across all seasons thus they shouldn’t be in high quantities in the fulfillment centers to save up on storage for the more fast moving goods like t-shirts.

**RATING**

* We assumed a good rating is one > 4 and a bad rating one < 2.6(the lowest rating in our dataset is 2.5)
* Sandals had the most high ratings count, this may reflect on their good quality and comfort customers experienced.
* Belts, Jackets and Gloves had the highest number of bad ratings. Kiosk should investigate what was the source of dissatisfaction from the customers.

**Colors**

We deduced the colors associated with high good ratings to define which colors did the customers like. Such an insight will guide Kiosk in placing such colors on the front row on the app to increase the probability of more sales.

* Colors gray and yellow had the highest ratings while blue and violet had the most bad ratings. This insight reflects customers preferred gray and yellow colored attires and tended to avoid blue/violet attires.

Kiosk could reduce the number of these colors to reduce stock hoarding

We then deduced the color associated with the highest frequency of purchased items i.e., which color was on the attires most bought.

* Colors Olive and yellow were the most frequent on purchased goods
* Gold, Brown and White were on the slowest moving goods.
* Kiosk may increase the count of these preferred colors whilst reducing the avoided colors gradually. This will in turn increase profits as money will be invested in fast moving attires.
* Kiosk should also dominantly use these colors on the app since there is a significant probability that these colors attract customers.

**Subscriptions**

Insights from this analysis will help Kiosk decide on migrating numbers to the subscription plan.

* Most revenues, 73.1% of all revenues were earned from unsubscribed customers
* Subscribed customers generated 26.9% of total revenues
* 73% of Kiosk’s users are not subscribed
* Ages with most subscribers are 64 and 54 with 28 users, 55 and 50 with 27 and 43 with 26.
* Ages 39 and 18 were lowest with 14, 60 with 13 and 44 with 11.
* Perhaps Kiosk should drive the unsubscribed age group 60 and 44 into subscription as most of their peers are on a plan.
* Kiosk should have incentives that attract customers into the plan, incentives that guarantee convenience.
* Subscription plans generate predictable income thus a good business resource, it would be important for Kiosk to migrate these numbers into a plan.

**Discounts**

Discount analysis will help us establish a correlation between discounts and sales.

* Non-discounted goods generated $133,670 where as discounted goods generated $99,411
* Most of the items on sale are non-discounted at a total of 2223 and discounted items are 1677
* Items with the most discounts are Pants, Coats, Sweaters and Hats.
* Items with the least discounts are Gloves, Sandals, Blouses, and Jeans.
* It’s worth to note Jeans and Gloves were the worst sellers. Thus Kiosk would decide whether to discount these items more to test if their sales will improve.

**Payment Methods**

Insights from this analysis will help Kiosk’s team enhance app compatibility of different preferred payment methods.

Increased app compatibility enhances customer convenience and a good customer journey.

* Most of the revenues were paid via Credit card at $40,310 followed by Paypal at $40,109.
* A good sum was collected in cash as well ,$40,002, a good statistic to have at hand as Kiosk could strategize on making cash payments much easier and secure for customer convenience.
* Using these insights, Kiosk may plan on moving numbers to methods they deem convenient.

**Frequency of purchase**

This value represents the number of times a user purchases items in a period of time.

It’s a useful metric that may help anticipate sales, plan for procurement, discount planning and carry targeted marketing.

* Every 3 month frequency has the most user count whilst weekly has the least user count
* Users who returned annually generated the most revenues, a sum of $34,419 whilst weekly users generate the least at $31,786
* These insights may help Kiosk do targeted marketing when a customer’s frequency nears, to increase sales probability

**Shipping Methods**

* The most used shipping method is Free shipping with 675 users whilst 2-day shipping is the least preferred with 627 users.
* This metric may reflect users prefer fast and free shipping thus such an insight can be used to woo customers on a certain side
* From the other preferred methods that are store pickups and standard, we can safely deduce customers put a premium on convenience and speed, an insight Kiosk can maximize on.

**CONCLUSION**

When put into consideration, the insights above provide a new and fresh perspective on the daily operations.

Customer preferences have been deduced from items purchased, ratings and colors, profits can be maximized by reducing slow moving inventory like jeans, app compatibility can be enhanced to smoothen credit card payments, targeted marketing can be used on certain age brackets to woo them into buying a subscription plan,

Certain items can be displayed on the front row at a certain period of time the customer will most likely come back for a purchase and many other instances.

We believe with this insights, Kiosk will experience a healthy growth process onwards.

,

,