Ecommerce Analytics

Requirements

I am hired as junior analyst in E-commerce company, and I should analyze their raw transaction logs to find insights. The dataset is in Google spreadsheets, where each row represents an activity, or event, by a user on the company's website. Each time a user views a product page, opens their shopping cart, or completes a purchase, the event is captured in the activity logs.

1. Finding Conversions Rates and Retention Rates based on their first purchase through their month by month.
2. conversion rates and Retention Rates values are decreased from view to purchase and from cohort's month by month.

Data Source

- The primary dataset used for this analysis is raw user activity and purchase activity in the form of google sheets which contains user_id, event_type, category_code, brand, price, event_date.

Tools

- Excel – Data Cleaning
- Pivot tables – Conversion funnel, Retention Rates, cohort analysis, first_purchase_activity

Data Cleaning and Preparation

Part 1: Build a conversion funnel

In the initial stage, I performed the following tasks

- Data loading and inspection
- created the funnel in a pivot table as a new sheet called "conversion_funnel".
- counting the unique users for each stage of the funnel using pivot table.
- Used formulas to create two new columns in pivot table: total conversion rates and conversion rates.
    1. Total conversion rates: =B2/$B$2
    2. Conversion rates: =B3/B2

Part 2: Prepare data for cohort analysis

In the middle stage, **data preparation for cohort analysis**

Created **filter purchases** and **calculate first purchase dates**

- Created a new blank sheet tab called “purchase_activity” with pivot table.
- calculated the minimum event_date for each user.

Next, I transferred the first purchase dates to a new column in the purchase activity data.

- Created “first_purchase_date” as a column header in cell G1 of the “purchase_activity” sheet
- Used a formula VLOOKUP () function to find the date from the “first_purchase” sheet that corresponds to the user ID in cell A2
    - =VLOOKUP (A2, first_Purchase! $A$2: $B$1082, 2)
- To check the results first_purchase_date values are greater than the purchase_date values.

<u>Build and track cohorts on monthly basis</u>

- Group the users and transactions by month for the cohort analysis.
- Create three new columns in the “purchase_activity” sheet to help build the cohorts: event_month, first_purchase_month, and cohort_age.

Used formulas for

- Event _month: =TEXT(F2,"YYYY-MM")
- First_purchase_activity: =TEXT(G2,"YYYY-MM")
- Cohort_age: =DATEDIF (I2, H2,"M")

Part 3: Calculate retention rates

The last steps of the analysis are to aggregate the purchase data into cohorts and then calculate retention rates for each cohort month by month.

Group data into cohorts

- Using the data from the “purchase_activity” sheet, inserted another pivot table as a new sheet called “cohort_analysis”.
- In pivot table each represents one cohort, which are based on the month in which customers made their first purchase and also have the count of unique users for each cohort_age in the columns.

Calculate overall retention rates

Formula Used for retention rates: =IFERROR (Cohort_Analysis! C3/Cohort_Analysis! $G3,"-")

calculated the retention rate for each cohort at each cohort age in the table you created, based on the starting cohort sizes.

Analysis:

Over all analysis where users are decreasing from month to month with different cohorts for purchase and the retention rates came low based on the analysis.

Raw Data: The spreadsheet contains data with columns user ID, event type contains "view", "purchase" and "shopping_cart", category code includes with product, brands column with specific brands, prices with wide range prices and event dates ranges from 2029-09-24 to 2021-02.28.

Conversion Funnel: The conversion rate from viewing to adding to cart is 29% which is low at this stage and from shopping cart to purchase is 35.61% which is high conversion rate at this stage. The total conversion rate overall from view to purchase is 10% which is low conversion rate.

Retention Rates: The Retention rates are decreasing while the cohort ages increase in all monthly cohorts. The Highest Cohort is on the month 2020-09 with Cohort Age 1 as 13%. The Lowest Cohort age is 4 with the lowest value 0%. Cohorts with cohort age is decreasing month by month.

Results:

Conversion Funnel: This table tracks Unique users including views, shopping cart and purchases

with conversion rates. With various Event Types conversion Rates are listed.

Retention Rates: Based on the User Interactions through their first purchase month with column cohort age calculated Retention rates to examine user behavior over time.