Health.inc Case Study

Philip van Harreveld
pevanhar@ncsu.edu
802-318-3382

Variables: The variables collected by the business team in the pageview dataset are all important for different reasons. I used pageview_id as my primary key because it is the best option to uniquely identify each row in our database. For the purposes of this case study a primary key wasn't necessary because all of our queries in SQL were within a single table. But I still created one because it is good practice to have a primary key in case other tables are added to the database in the future.

Critical Variables: The critical variables for determining whether or not a user will buy a product on one of the assets are asset_shown, page_category, page_topic_description, and return_visitor. These variables are critical because they have a direct influence on conversion rates and revenue potential by capturing key factors that determine user engagement and purchasing behavior. By focusing on critical variables we can identify the best performing asset to user/page combinations. Which can reveal insights into how we can optimize asset assignment to maximize conversions and revenue. This will also naturally help the user by virtue of seeing assets more relevant to their needs and interests.

Non-Critical Variables: The non-critical variables for determining whether or not a user will buy a product on one of the assets are asset_loaded_time, session_start_time, user_id, and pageview_id. These variables are non-critical because they do not directly influence the users decision to purchase an asset but instead provide contextual or tracking information. Variables such as when the asset was loaded, when the session started, and unique identifiers for users or page views are not inherently impactful on conversion behavior from the user. But while the work we will be doing on this dataset does not require the use of these variables, they can be absolutely necessary in the context of another problem. So in the grand scheme of things it is important that they are collected along with the current critical variables.

Average Conversion Rate by Asset: There was a large discrepancy between variables in terms of average conversion rates. Asset A had 55,000 total page views and converted 10.97% of users. Asset B had 27,500 total page views and converted 7.13% of users. Asset C had 17,500 total page views and converted 17.06% of users.

- **Asset A:** 10.97%
- **Asset B:** 7.13%
- **Asset C:** 17.06%

Asset Conversion Rate by Return Visitor: Return visitor status had a major impact on conversion rates across all three asset types. The average return visitor conversion rate was 24.66%. While the average non-returner conversion rate was just 9.12%. When we get more specific we find the following information:

-Asset A: 15% increase in conversions when the user was a returner and accounted for 35% of Asset A's total revenue while adding up to just 16% of Asset A's total views.

-Asset B: 21.5% increase in conversions when the user was a returner and accounted for 35% of Asset B's total revenue while adding up to just 16% of Asset B's total views.

-Asset C: 21% increase in conversions when the user was a returner and accounted for 35% of Asset C's total revenue while adding up to just 17% of Asset C's total views.

Asset Conversion and Revenue Generation by Page Category: Certain asset types will generate more or less revenue when shown to specific page categories. This takes into account conversion rates as well as the revenue per conversion for each asset type to determine total revenue. We also calculated total potential revenue by theorizing about how many conversions and subsequently how much revenue a specific asset would generate if it was the sole asset shown to all users of a specific page category. These are the assets that align best with each page category:

-Otherinflammatoryboweldisease: Asset B
-Nauseaandvomiting: Asset A
-Ulcerativecolitis: Asset B
-Crohnsdisease: Asset B
-Otherdigestiveproblems: Asset C

Asset Conversion and Revenue Generation by Page Topic Description: Certain asset types will generate more or less revenue when paired with specific page topic descriptions. The same process applied to solve the above question(page category) was applied here. These are just a few examples of assets that become especially high revenue generators when paired with a specific page topic description:

-Thyroid: Asset B
-Food Poisoning: Asset C
-Bleeding and Blood Loss: Asset A

Asset Conversion Rates By Page Category: Looking at conversion rates for page categories helps us to see which assets perform better or worse on specific categories. Which gives insight into how we can optimize the assets later on. Notable observations:

1) Asset C has the largest conversion rates across all page categories, followed by asset A, and then asset B. There is a noticeable difference in the rates between assets A and B while asset C is in a league of its own and quite clearly much higher than the other two.
2) All categories follow similar conversion rate trends. For example Crohnsdisease has the highest conversion rate for all three assets, while Otherdigestiveproblems has the lowest conversion rate for all three assets.
3) Asset B has a 0% conversion rate on the Otherdigestiveproblems category.

Asset Conversion Rates By Page Topic Description: Looking at conversion rates for page topic descriptions helps us to see which assets perform better or worse on specific topics. Which gives insight into how we can optimize the assets later on. These are some observations:

1) Page topics followed trends with regards to conversion rates across all assets for the most part with a few outliers.
2) Asset C has significantly higher conversion rates than both A and B for all 20 page topics.
3) Asset B has similar conversion rates to asset A although slightly lower on average. But there are four page topics that asset B has a 0% conversion rate on. This is the only asset to have this happen.

Optimization Strategy: To optimize asset assignment across user type, page categories,and page topics I needed a consistent strategy. Which turned out to be comparing current revenue numbers to potential revenue numbers in an optimized system. I already knew how to find total revenue from previous queries, but now I have to find potential revenue as well. To calculate potential revenue I took the total views and multiplied it by the respective conversion rate, and then multiplied that number again with the revenue per conversion for the specific asset type. This determined the potential revenue by using the rate and revenue per conversion against the total amount of views instead of the current amount of views to see how much revenue would be generated by only using a single asset instead of a mix of all three. Then you can simply compare the total potential revenue numbers in the output to see how you should optimize each asset.

Optimizing Asset Assignment for user type(New or Returning user):

-Asset A: While being middle of the road in terms of conversion rate, asset A falls short of asset C and above asset B. Although asset A doesn't have the highest conversion rate it still has the highest potential revenue because of its $5 of revenue per conversion. The combination of second highest rate and second highest revenue per allows asset A to become the best in terms of total revenue generation for both new and returning users.

-Asset B: Has the lowest conversion rate of all three assets for both new and returning users. But because of its high $7 of revenue per conversion, asset B makes up for its low conversion rate. When optimized, it becomes the second highest in terms of total potential revenue for both new and returning customers. And it's a close second to asset A in first place.

-Asset C: While having the highest conversion rate for both new and returning customers, asset C falls dead last in terms of potential total revenue because of its lower $2.5 revenue per conversion. This means that although it has a high rate it is not very profitable and therefore comes last for both new and returning users.

Optimizing Asset Assignment For Page Category:
(Current=Revenue Generated In The Current System, Optimized=New Optimized Revenue)
- **-Otherinflammatoryboweldisease:** Current=$24,250, Optimized=$26,400 (Asset B)
- **-Nauseaandvomiting:** Current=$9,750, Optimized=$12,000 (Asset A)
- **-Ulcerativecolitis:** Current=$9,250, Optimized=$10,650 (Asset B)
- **-Crohnsdisease:** Current=$5,050, Optimized=$6,650 (Asset B)
- **-Otherdigestiveproblems:** Current=$3,000, Optimized=$5,400 (Asset C)

Optimized Asset Assignment For Page Topic: (Top 5)
- **-Thyroid:** Current=$5,800, Optimized=$6,000 (Asset B)
- **-Diabetes Nutrition:** Current=$6,250, Optimized=$6,900 (Asset B)
- **-Food Poisoning:** Current=$2,350, Optimized=$2,850 (Asset C)
- **-Abdominal Pain:** Current=$2,300, Optimized=$2,850 (Asset A)
- **-Joint Pain:** Current=$2,450, Optimized=$2,800 (Asset B)