Max
3/10/2024

Advertising Conversions in SQL

The goal of this project was to analyze a kaggle dataset detailing a social media ad campaign to generate marketing insights using SQL for data manipulation and querying. I wanted to know which advertisement was most effective and cost efficient with which demographics, and from this information try to predict what the product was being advertised. This anonymized data contained 1143 observations and 11 variables:

1.) ad_id: an unique ID for each ad.
2.) xyz_campaign_id: an ID associated with each ad campaign of XYZ company.
3.) fb_campaign_id: an ID associated with how Facebook tracks each campaign.
4.) age: age of the person to whom the ad is shown.
5.) gender: gender of the person to whom the add is shown
6.) interest: a code specifying the category to which the person's interest belongs (interests are as mentioned in the person's Facebook public profile).
7.) Impressions: the number of times the ad was shown.
8.) Clicks: number of clicks on for that ad.
9.) Spent: Amount paid by company xyz to Facebook, to show that ad.
10.) Total conversion: Total number of people who enquired about the product after seeing the ad.
11.) Approved conversion: Total number of people who bought the product after seeing the ad.

I began the analysis by dividing the data into third normal form so that each field was dependent on one and only one primary key in each dataset. This resulted in four separate tables with ad_id as a primary key in each:

1) ad_table: ad_id, xyz_campaign_id, fb_campaign_id, Spent
2) audience_table: ad_id, age, gender, interest
3) conversion_table: ad_id, Total_Conversion, Approved_Conversion
4) impression_table: ad_id, Impressions, Clicks

Next I explored and cleaned the data. There were no duplicate ad_ids and no nulls in the data. There were 691 unique facebook_campaign_id values and 3 unique xyz_campaign values: 916 (47 unique campaigns), 936 (367 unique campaigns), and 1178 (277 unique campaigns). After cleaning and exploring the data I used aggregate functions to calculate four KPI's for each ad:

1) CTR (Click-Through Rate = Clicks / Impressions * 100)
2) CPC (Cost Per Click = Spent / Clicks * 100)
3) Total Conversion Rate (Total Conversions/Interactions * 100)
4) Approved Conversion Rate (Approved Conversions/Interactions * 100)

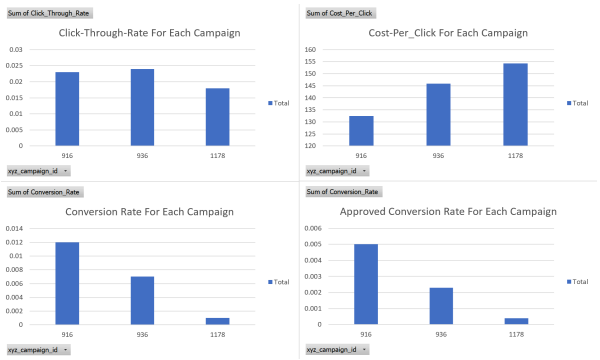
By comparing these values across all three ads I generated several insights. 916 is the cheapest ad in CPC and with the best approved conversion rate, if I were to recommend only one advertisement to use it would be this one. Ad 936 has the highest CTR and an especially bad ACR indicates that it's doing a great job of drawing attention but people are losing interest after they click on it. Ad 1178 has both a high CPC, low CTR, and low conversion rates. This likely indicates ineffective messaging and the least successful ad despite being the second most used. However it may be possible that one of the three ads was better with a particular demographic than another. To find this information I compared age, interest, and gender to my KPIs for each ad to gain a deeper understanding of these trends. Using CTEs, Subqueries, and Temporary Tables I analyzed each of these factors.

All ads had higher CTRs with the 45-49 year old age bucket, and generally tended to favor an older audience. Ad 916 had medium CTR for each age bucket except 45-49 which was high. Ad 936 had a high CTR with ages 40 and above but low CTR with 30-34 year olds. Ad 1178 had medium CTR with ages 45-49 but low CTR across all other ages. Women notably had higher CTR than men but lower conversion rates which indicates that women are losing interest after they click on the ad while men don't. CPC was similar across genders except for ad 1178 which tended to be more expensive per click when targeting men. In terms of interests, Ad 1178 was the only one targeted at interests over 100, but had a lower click-through-rate across all other interests which indicates the other two ads should be targeted at those other interests as well. The other two ads performed better when targeted at users with different interests. Ad 916 had especially high CTR at interests 10, 20, 21, 24, 25, 27, 30, 32, 63, and 64, while ad 936 had better CTR with all other interests under 100. Complementary to these insights I built a stored procedure in the database which functions as a facebook id-lookup tool allowing a manager to zoom in on one particular ad at a time and find its individual KPI and demographic information.

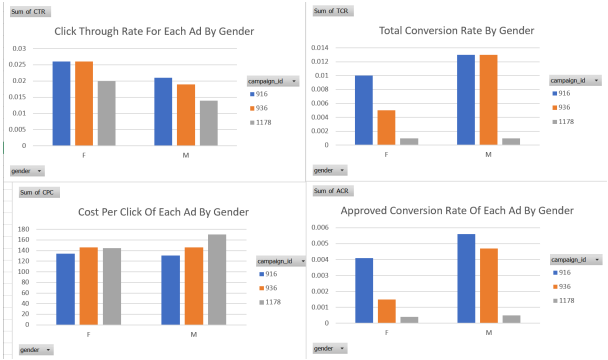
Bringing it all together, Ad 916 is the most efficient, offering the lowest cost with the highest approved conversion rate. It performs moderately across all age groups, and should be especially successful among men aged 45-49. Although ad 936 leads in CTR, its higher CPC and lower conversion rate compared to ad 916 make it less desirable but still a strong option for certain interests. It does notably better with individuals over 40 and struggles with the 30-34 age group. Ad 1178 is the least effective, characterized by high CPC, poor conversion rates, and the worst CTR, suggesting potential issues like misleading content, unexpected costs, or ineffective messaging. It's advised to be the first ad to drop, performing especially poorly across demographics but slightly better with the oldest age groups.

For optimal performance ad 1178 should be dropped and ads 916 and 936 should be targeted at an older male demographic which tends to have higher conversion rates. Ad 916 is best suited for specific interests (10, 20, 21, 24, 25, 27, 30, 32, 63, 64), and ad 936 should target the remaining interests under 100 not covered by ad 916. More resources should be devoted to solving the overarching issue of women losing interest after clicking on the ad. They are likely encountering some pain-point after the click that men are not. Their high CTR could prove a very strong asset if this issue is resolved. Ads 916 and 936 should be tested for interests above 100 since they well outperform 1178 on all other interests. Finally, the fb_id_lookup tool should be used to measure the effectiveness of individual campaigns across each KPI and demographic. Visual representations of the data discussed here and the data set from kaggle are provided below.

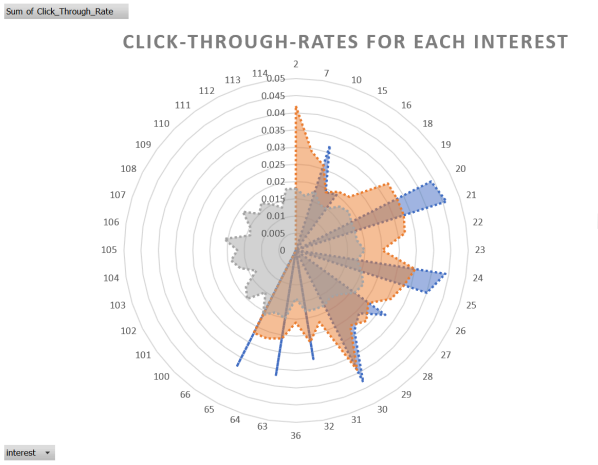
KPI Dashboard:



KPIs by Gender



CTR by Interest



CTR by Age

xyz_campaign_	age	CTR
916	30-34	Medium
916	35-39	Medium
916	40-44	Medium
916	45-49	Higher
936	30-34	Lower
936	35-39	Medium
936	40-44	Higher
936	45-49	Higher
1178	30-34	Lower
1178	35-39	Lower
1178	40-44	Lower
1178	45-49	Medium

- Data Source: https://www.kaggle.com/datasets/loveall/clicks-conversion-tracking