**Project Report – E commerce Analysis**
*Tools used: R, Excel*

## Aim of the Assignment

The aim of this assignment is to analyze the users and page views dataset from an eCommerce company and provide recommendations to the product team to boost the sales.

This project also aims to find out if there are any issues with the dataset used for the assessment.

## Key Observations from the Data

- The overall conversion rate (homepage to payment confirmation) is 0.48%. Only 7.35% users are actually buying the product(s) after visiting the payment page.
- 13.31% users are moving to payment page after landing on the search page.
- Major chunk (more than 50%) of the traffic comes from search (organic and paid) and social media.
- Some users visited the homepage 8 times, search page 7 times, payment page 2 times but still didn't buy any product.
- Share of devices (desktop & mobile) is more or less equal.
- Gender ratio of the visitors seems 50:50

## Insights Drawn from the Observations

While first-time visit doesn't necessarily lead to sale, a conversion rate of 0.48% seems unusually low. [Note: According to invespcro.com, the average eCommerce conversion rate in the US is ~ 2-3%. According to wordstream.com, the top 25% landing pages have a conversion rate of 5.31% or higher. Top 10% are are even converting at 11.45% or higher.]

A more meaningful way to summarise the funnel conversion data would be to find out how many users are purchasing the item after going to the payment page. This comes to 7.35%, which is still lower than the industry benchmark.

A low conversion rate can be a sign of a many things. I have noted down a few that, if valid, need immediate action.

Business focused

- The site has a high page load time.
- The site has navigation issues or the content on the website is confusing.
- Users are getting poor search results.
- There's a glitch in the on-click actions on the site.  payment gateway is broken.
- This company is not reaching out to the correct audience.

- The site may be infected with malware that is keeping the users out.

Consumer focused

- The customers perceive the product prices/shipping cost to be too high
- Users' coupons/referral code/promo code not applicable
- Users' preferred mode of payment is not supported
- Items are unavailable to be shipped to users' location
- Drop-offs on the payment page can also happen if users don't have their card details handy

**Recommendations**

- Check if the website is optimised for all platforms and is loading quickly. Page load time is crucial for overall user experience. The revenues are impacted adversely as the page views go down due to a slow site. I can comment more on this after analysing the site speed report.

- Tackle navigation issues that are hampering visitors' experience on the site. Revisit the design if the visitors are not able to understand what to do next after landing. Use prominent CTAs(call to action, for example: click here for site tour), and tool-tips for the new users.

- Analyse if the users are able to find the items they want to buy. Ask "did you find what you were looking for" on the search results page. You might need to tweak the search algorithms based on the responses.

- Paid and organic search are doing well in redirecting web traffic. This means there is good brand familiarity but at the same time it's not adding to sales. The potential reasons for this could be that a) the search keywords are irrelevant, excessive or misleading b) the ads are not reaching the right audience. I'd recommend the team to redefine the target audience and reach out to them with strong and focused product keywords.

- The Marketing mix can also be revisited since paid search traffic alone is not transcending into conversions. Maybe increase the spends on other channels such as referrals and display ads.

- Re-write the content on the site to make it informative and persuasive. Provide SEO-friendly, detailed product descriptions. Show shipping, cancellations, returns, and warranty information upfront.

- If users abandon their carts, send them a reminder within 30 minutes. If possible, sweeten the deal with x% off coupon if ordered within the next 24 hours. Send a couple more reminders within the next 3 days and follow up with re-targeting ads. Analyse feedback. If the users find shipping cost too high, run limited period discounts.

- Ensure that the coupons/referral code/promo code/discount vouchers are working on the checkout page.

- Save users' browsing information using cookies (after asking their explicit permission). It would be good to check if the users are browsing similar sites for price comparison. It will help the team arrive at a competitive pricing strategy.

- Check the site's health on a regular basis. Run malware diagnostics.

- Fix glitches on buttons, if any. Wait, did you check if the payment gateway is working properly? It doesn't matter if everything else is working but the customers are not able to checkout and pay.

**Existing and Possible Issues with the Dataset**

- There are duplicate rows in the users table.

  Workaround: I removed all duplicates before loading the dataset in RStudio.

- Some users have been registered as male as well as female in users table. The possibilities for this could be:

  a) The site is asking users' gender as soon as they enter. Some people don't want to give out their personal details. Hence, they may be choosing random values every time.
  b) The site might be tracking users' behavior and assigning them genders. For example: if a female is browsing items in the male section and moves back to the female section, the system might be assigning female-male-female gender to the same person in different sessions.

  Workaround: Cleaning the data is important in such cases. I assigned "NaN" to the gender of such rows and removed one row. By removing both the rows I would be losing information and by keeping both the rows, I would be compromising with data cleanliness.

- Page views assumed to be unique users

  A page view does not necessarily represent a unique user. The same user may visit the same page twice in one session or even several times in the given time duration.

  Workaround: I created a dataframe of unique users for further analysis by creating pivot table in MS Excel and loading in RStudio.

- The data values for share of devices, and gender ratio is ~50:50

  We need to recheck how are we collecting these values, as a 50-50 split seems like an ideal-world scenario, which, especially in the case of eCommerce, is too farfetched.