Business Intelligence Data Challenge

Table of Contents

Table of Figures

1.Introduction

Usually when you sell something in the internet, the user/customer comes multiple times (e.g. visits) to the site, often also via different channels (e.g. Email, Facebook, Google, etc.), before he or she buys something (a conversion or transaction). In this analysis we try to explore the insights behind the conversion and explore the conversion on the basis of the attributed conversion fraction (IHC_Conv). Attributed conversion fraction is the fraction of the given conversion, which is attributed to a certain channel, sums up to 1.0 for every conversion.

1.1. Datasets

table_A_conversions.csv:	**table_B_attribution.csv:**
• example list of conversions/ sales • Conv_ID - transaction ID • Conv_Date - transaction date • Revenue - value of transaction • User_ID - an ID of a customer	• list of attribution results for conversions • Conv_ID - transaction ID (link to table A) • Channel - marketing channel • IHC_Conv - attributed conversion fraction by the IHC model

2.Exploratory Data Analysis

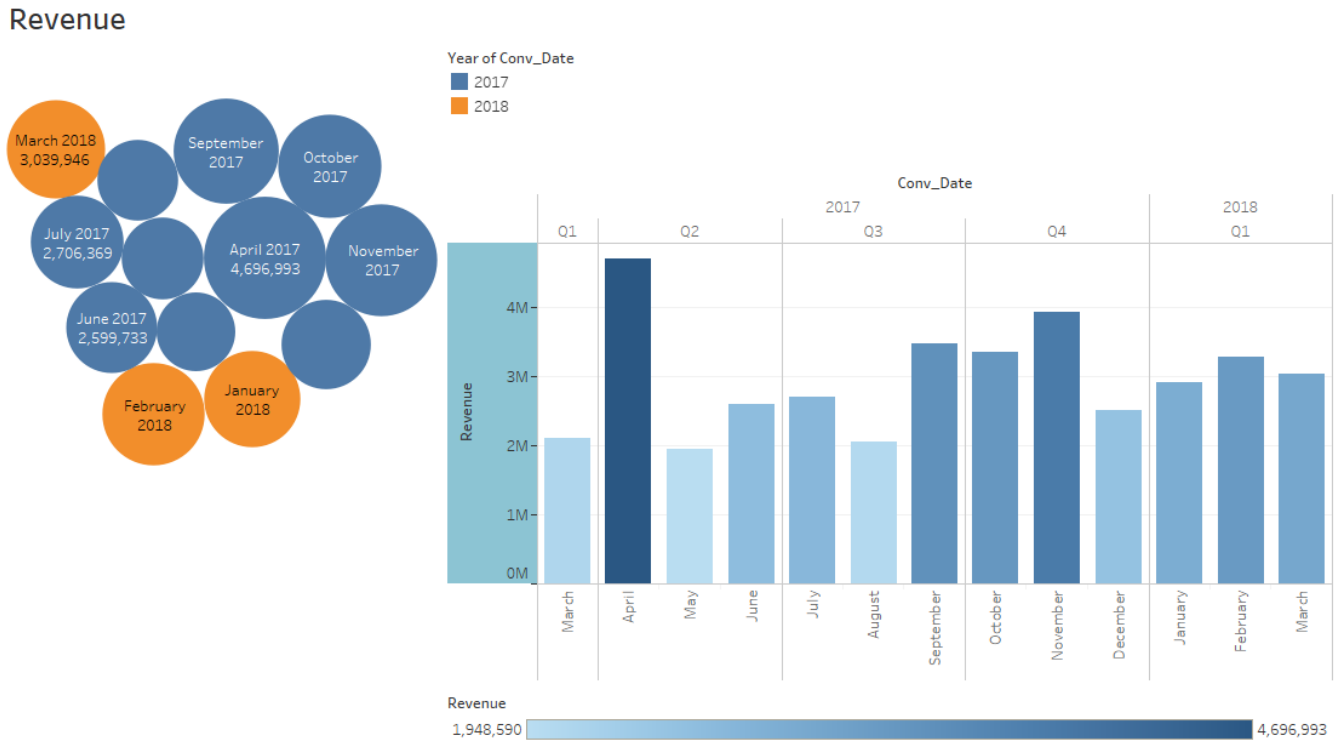

Figure 1- Revenue Comparison across different Quarters

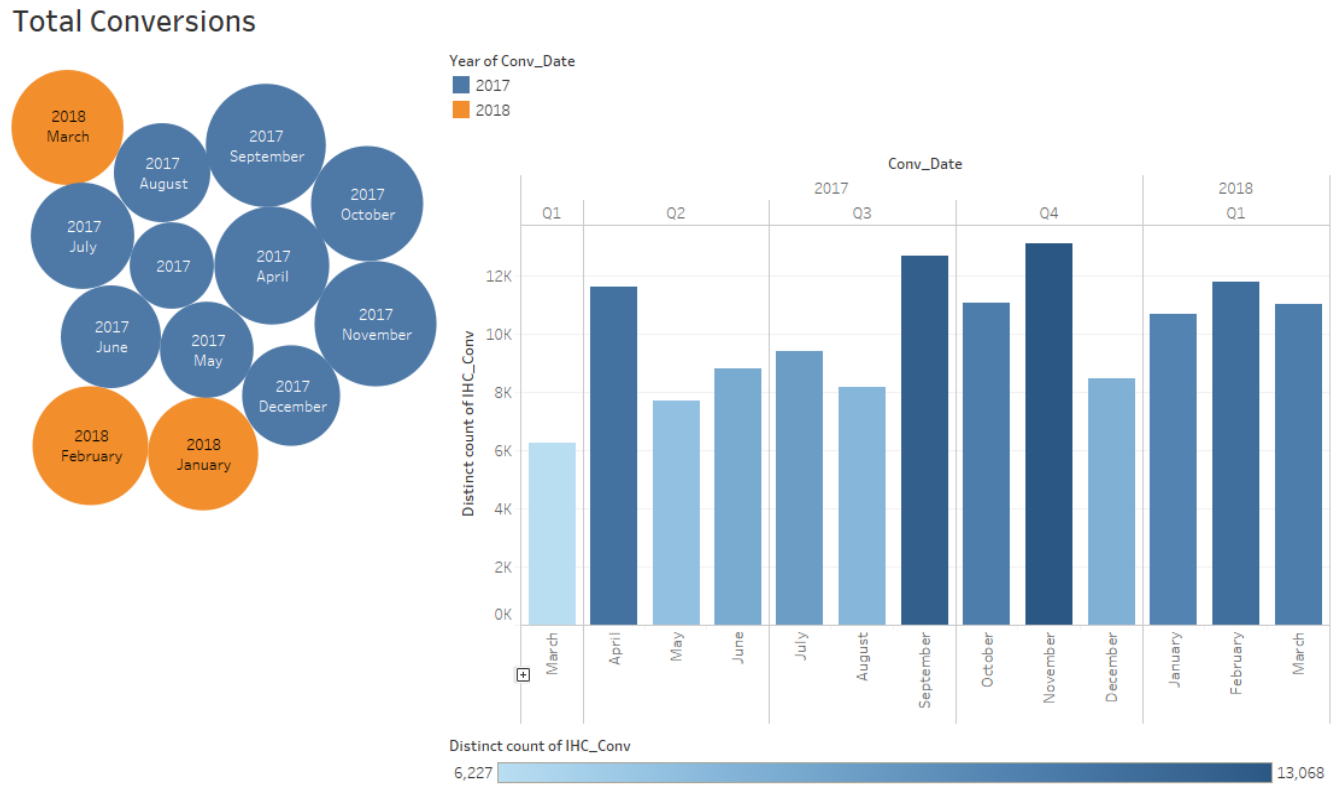

Figure 2- Total Conversions across each Quarters

Figure 1 and Figure 2 shows the total revenue and total number of conversions by each month, where the month of April stands out on total revenue generated and Quarter 4 records higher number of conversions happening. However, no significant seasonal pattern can be seen.

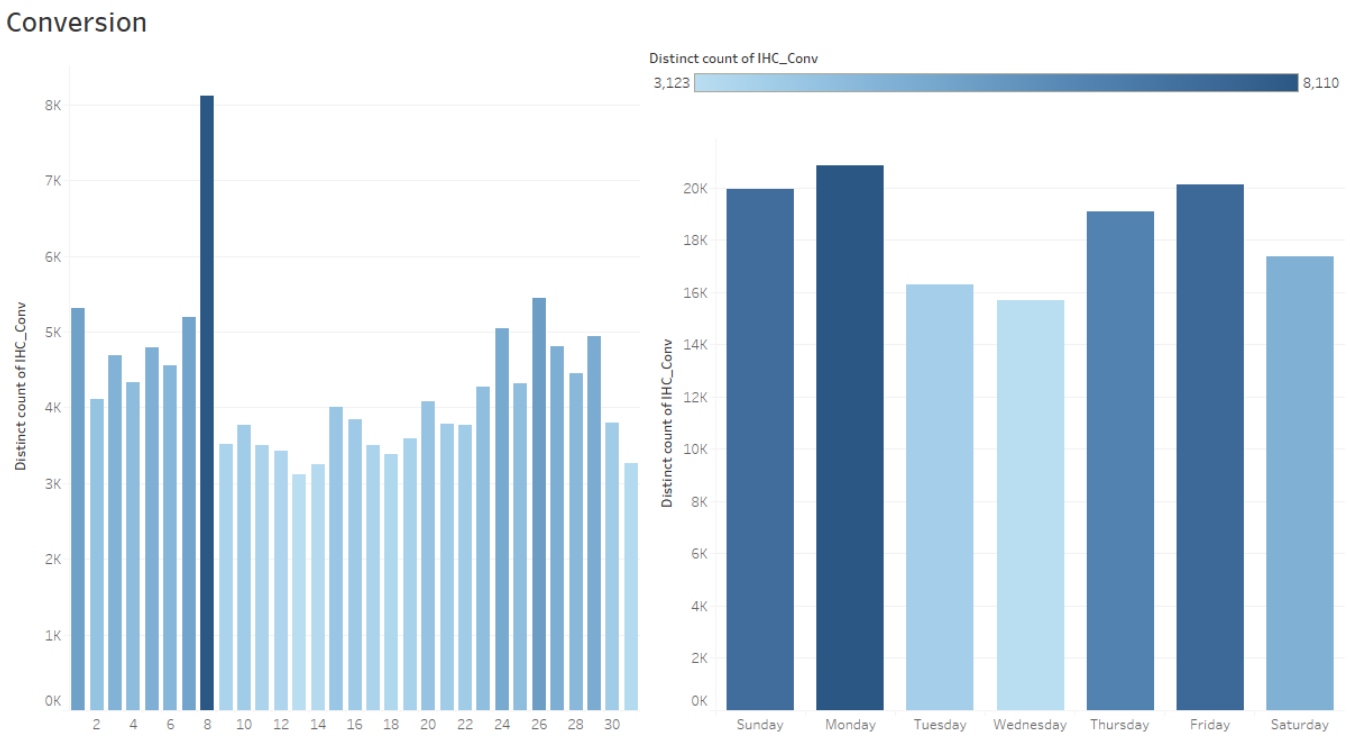

Figure 3- Total Conversion comparison across days of month and week

Figure 3 shows the number of conversions spread across Days of a month and weekdays. The date 8th of month records highest activity, however this could be a result of an offer or discount sale on that particular day of any month. The weekdays seem normal where it suggests people like to buy mostly on weekends.

3.Cohort Analysis

A cohort is a group of users who share something in common, for example sign-up date, first purchase month etc. Cohort analysis is the performed here to monitor the customer and revenue retention. The cohorts here are created based on a customer's first usage or transaction. Based on these cohorts, the retention rate and the revenue retention are analysed.

3.1. Cohort Analysis - Retention Rate

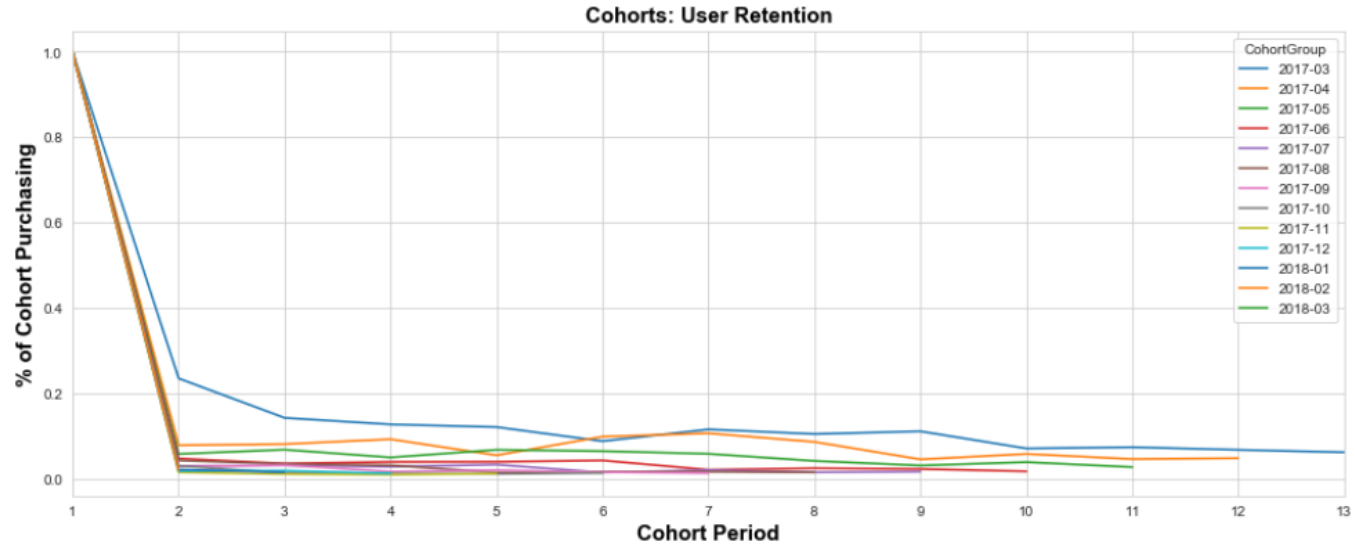

Figure 4- Cohort Analysis Customer Retention: Line Plot

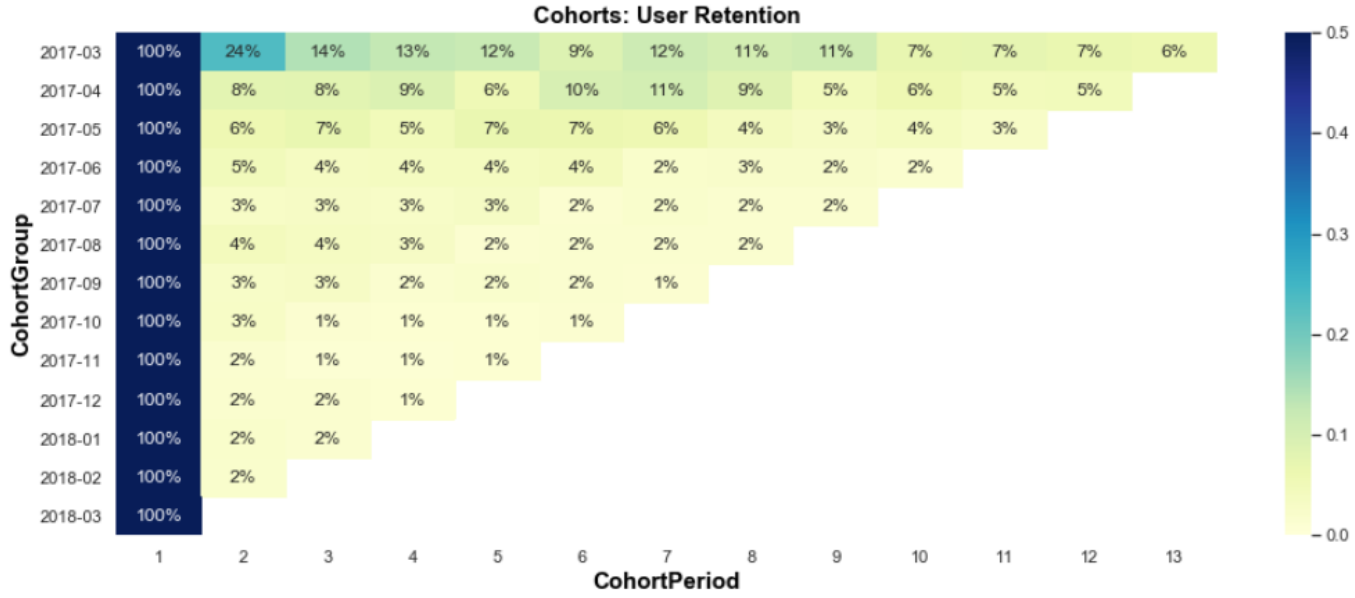

Figure 5- Cohort Analysis Customer Retention: Heat Map

Figure 4 and Figure 5 shows the retention rate of customers. we can see from the above chart that fewer users tend to purchase as time goes on. However, we can also see that the 2017-03 cohort is the strongest, which enables us to ask targeted questions about this cohort compared to others -- what other attributes (besides first purchase month) do these users share which might be causing them to stick around? How were most of these users acquired? Was there a specific marketing campaign that brought them in? Did they take advantage of a promotion at sign-up? The answers to these questions would inform future marketing and product efforts.

3.2. Cohort Analysis – Revenue Retention

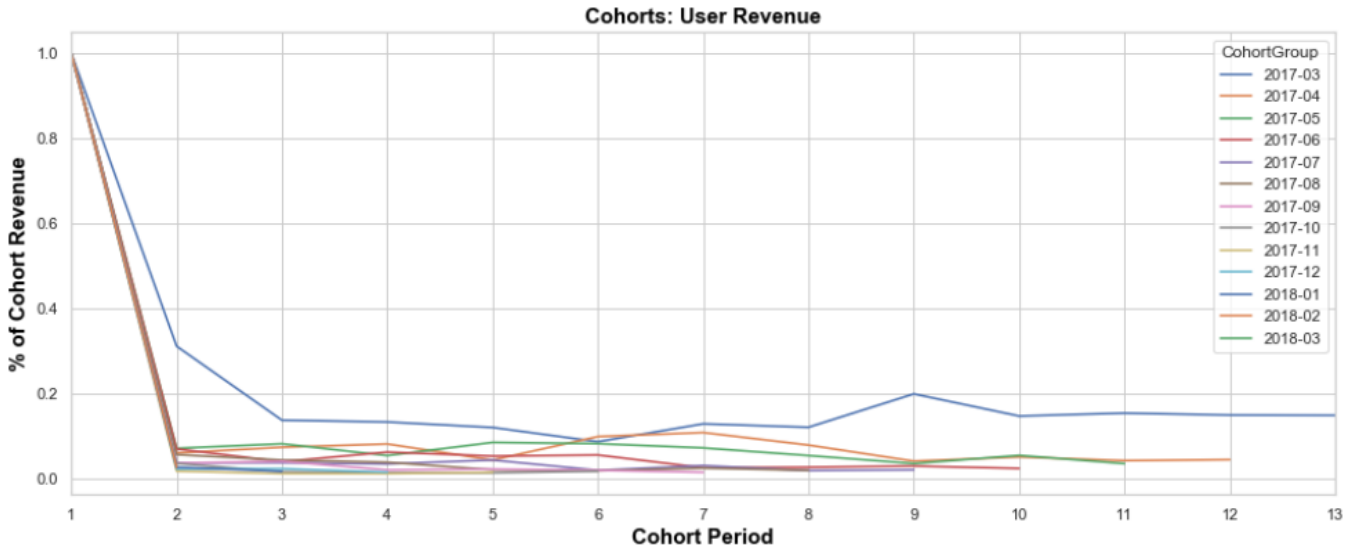

Figure 6- Cohort Analysis Revenue: Line Plot

Figure 6 and Figure 7 shows the retention rate of revenues. The situation for revenues is parallel with the consumer retention, which shows significant decline in revenues in recent months. From the graph, it can be concluded that from 2018/02 some strategies might have to be implemented to increase the return users or revenues. The decline shows about 50% of user from 2017/04 not returning/ revenue decreasing return on 2018/02. Marketing analysts or product analysts might further investigate if some campaigns or feature changes that can lead to a higher retention.

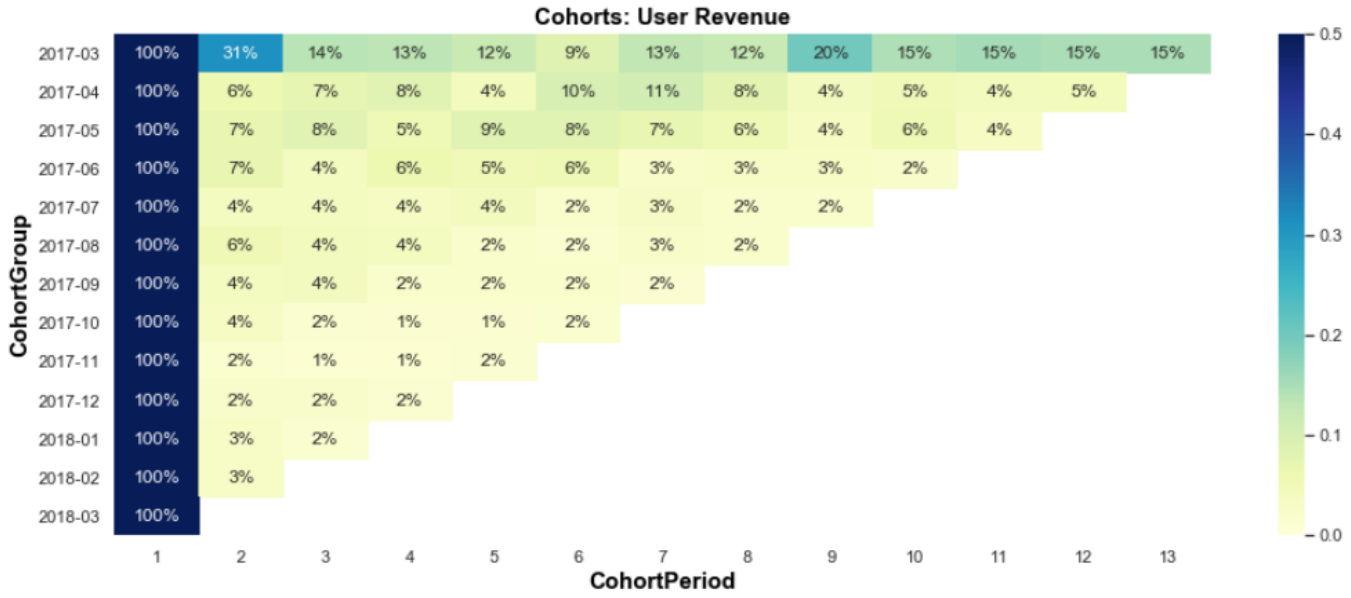

Figure 7- Cohort Analysis Revenue: Heat Map

4.Consumer Conversion Statistics

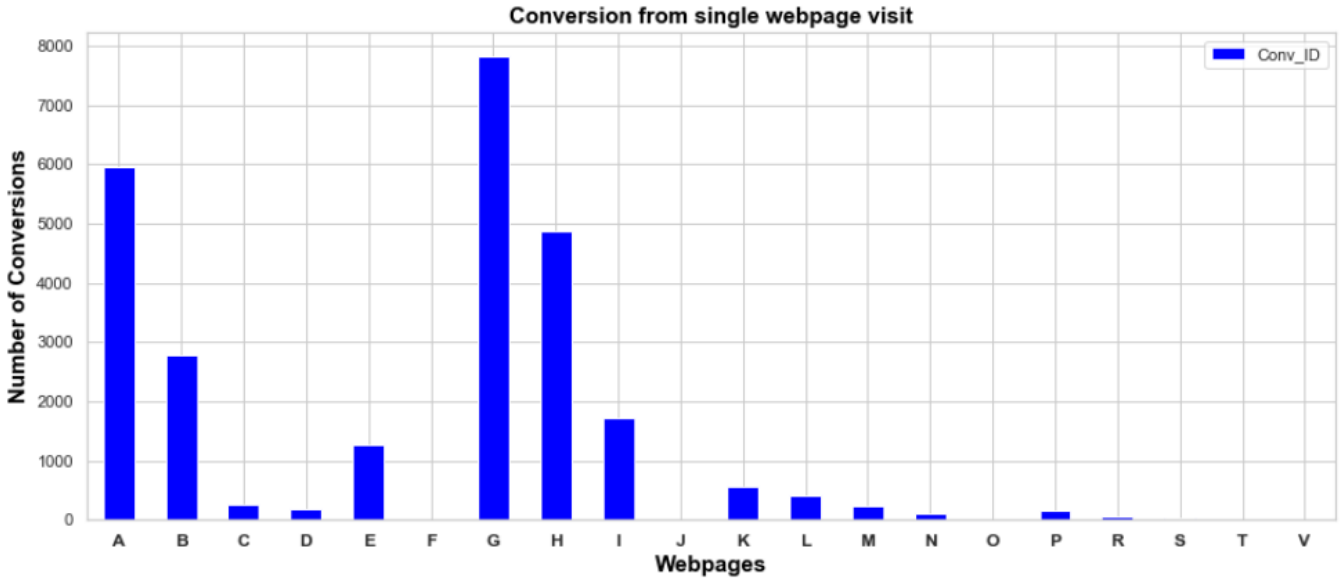

Figure 8- Total number of conversions from single webpage visit

Figure 8 shows the number of conversions that happened with a single web page visit. Website G is more influential in this regard followed by website A, H and B. The rest of the websites do not contribute significantly in single handed conversions.

Figure 9 shows the number of conversions that happened with the last webpage visit after intermediate visits. The last web page visit alone does not single headedly contribute to the conversion, but the entire intermediate visits contribute as well. But this visualization is just to see the impact of last web page visit on conversion. Website G, E and H are the most influential in this regard.

Similarly Figure 10 shows the number of websites one previous to the websites where conversions happened. This stat is considered because if we consider that the last page visited holds information, then the page that led to last webpage also can be considered significant to the conversion.

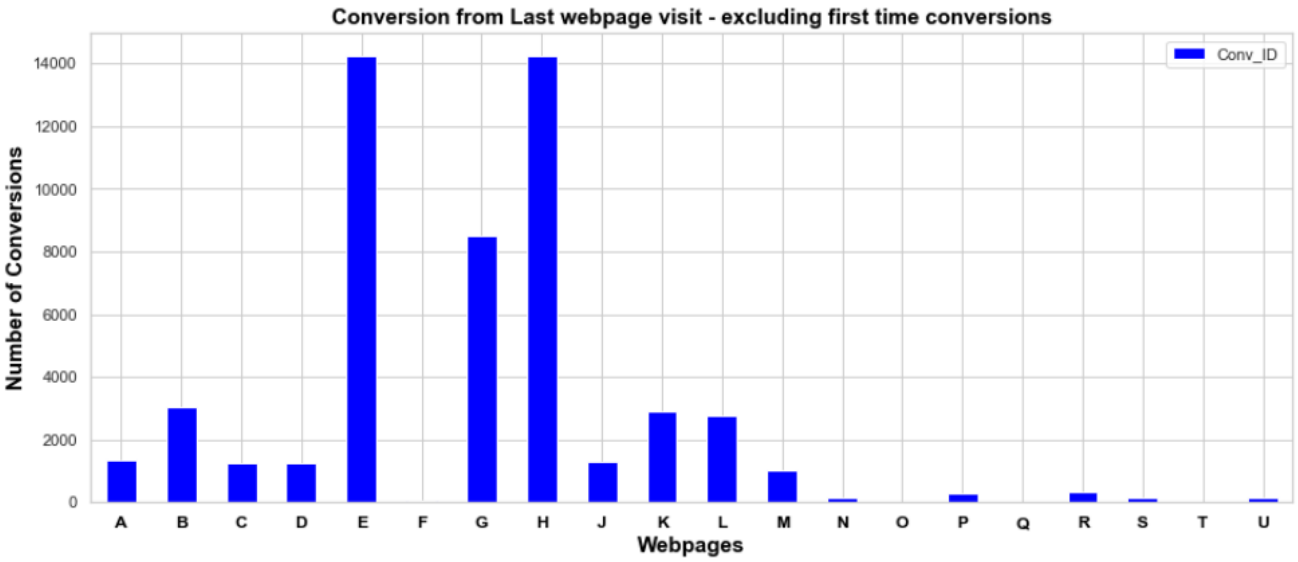

Figure 9- Webpages visited for conversion

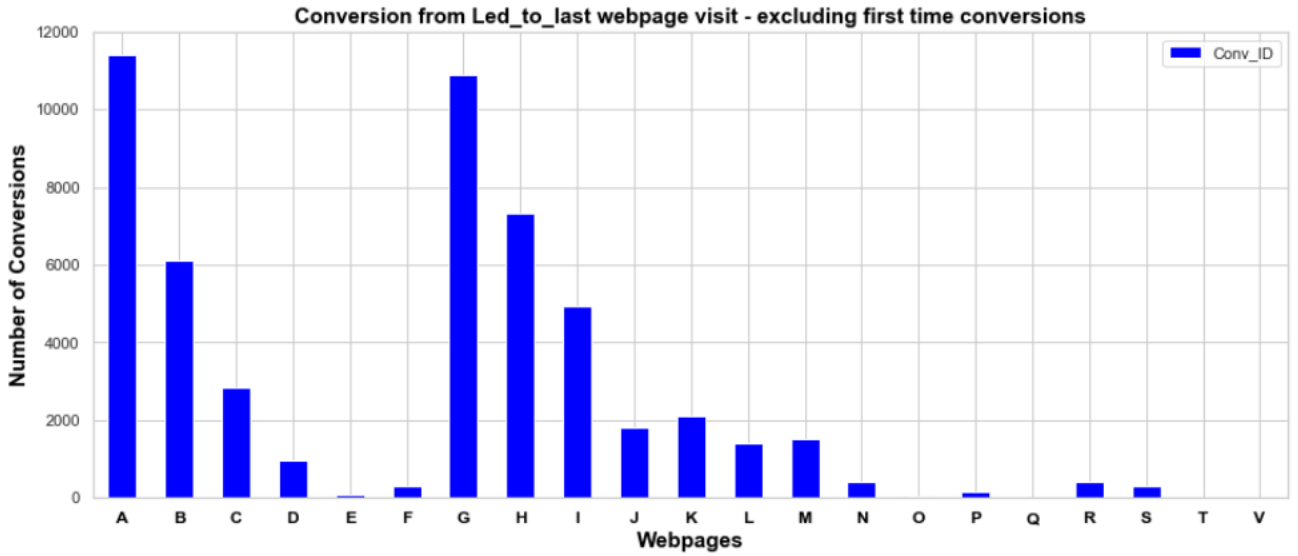

Figure 10- Webpages that led to the conversion (webpage visited before conversion)

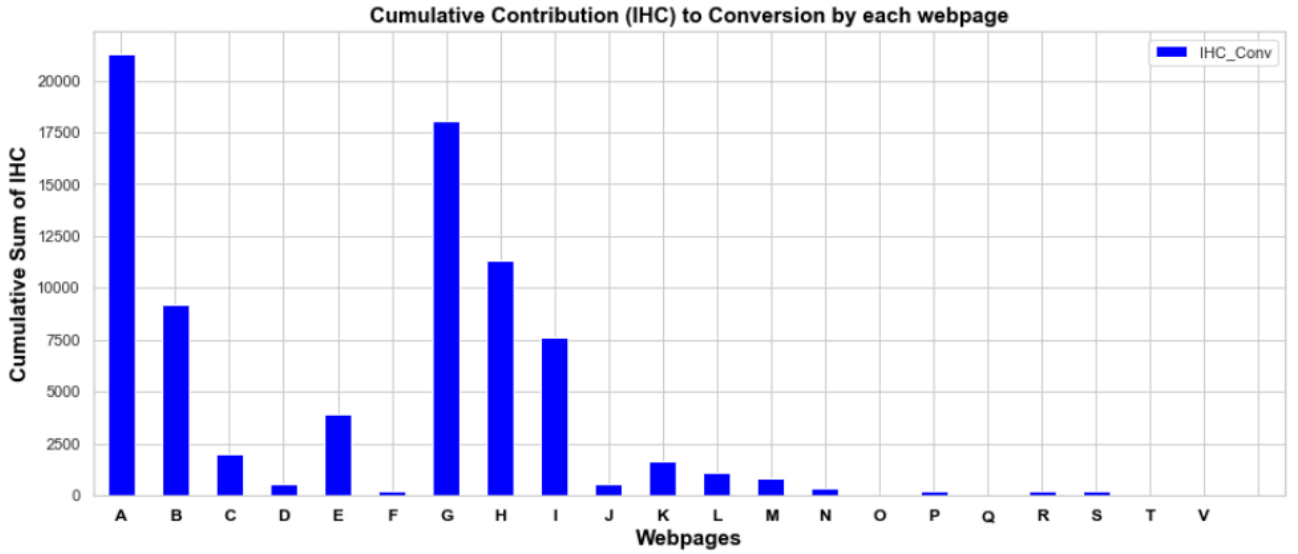

Figure 11- Cumulative sum of IHC of each webpage

Figure 11 shows the cumulative contribution of each individual webpage on the conversion. This visualization shows some significant strong players that dominate the conversion, A, B, G, H, and I. If the marketing team needs to eliminate or cut-off on advertising cost on multiple channels, they can consider this stat.

5.Customer Segmentation

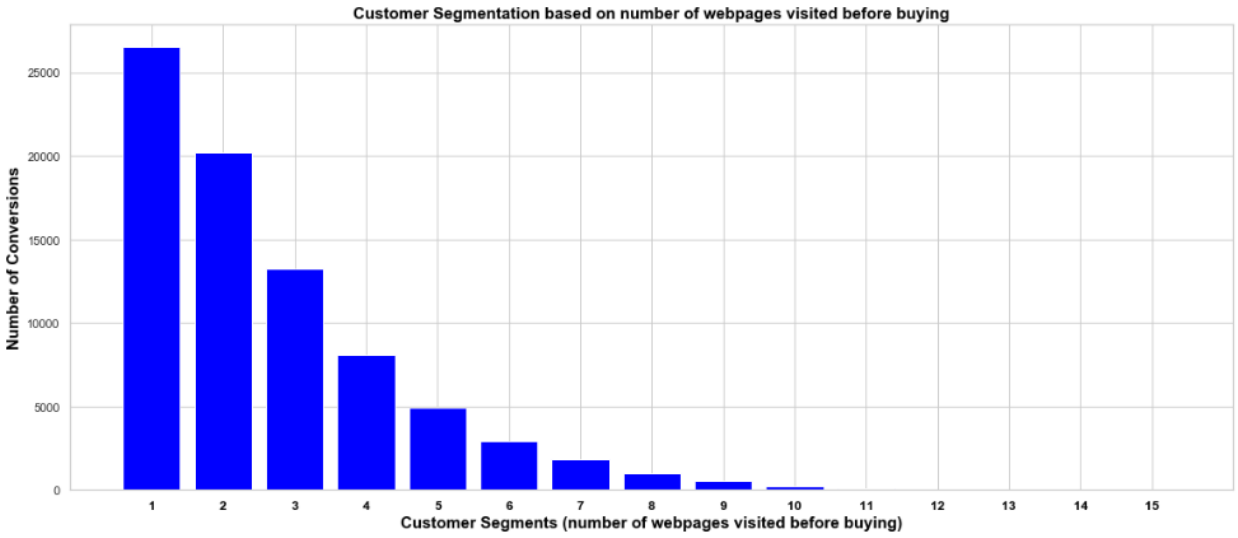

Figure 12- Customer Segmentation based on number of webpages visited before conversion

Figure 12 shows the customer segments based on number of webpages visited before confirming a conversion. It shows that the consumers like to visit as less webpages as possible and most of the consumers are satisfied with only one webpage visit. However, the consumers visiting 2, 3, 4, or 5 webpages (extra cautious consumers) are also significantly higher when put together and cannot be undermined.

Figure 13- Customer Segmentation based on number of days taken to make a conversion

Figure 13 shows the number of days a consumer researches before buying a product. People buy a product within a day are as twice as people who wait for a day before buying. It can be inferred from both these segments that people are predetermined on what they would like to buy. If they had spent more days, for example 4 or 5 days, that would show that a consumer bought a product randomly influenced by an advertisement which one had found amusing.