

**XN Project Individual Report**

ALY6080 Integrated Experiential Learning

Fall 2019

Submitted on: Dec 8th, 2019

By:

ANUJ GOLESAR

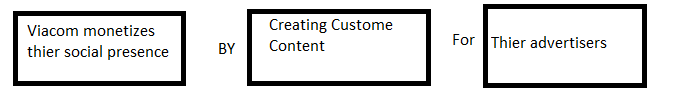
**Instructor: Prof. Nathaniel C Lin**

**INTRODUCTION**

In this assignment we are assisting the Viacom in Facebook ads targeting analysis. This report is about the analysis of the page level impression data and the CPM (cost per 1000 impression) data provided by Viacom. We will focus on the business model proposed by Viacom and will suggest whether the product offered by Viacom to advertisers is worth or not. Wherein Viacom has great organic reach and could draw enough impressions to the advertisers, however Viacom might need to buy paid impressions from social platforms like Facebook to reach particular demographics.

**RATIONAL**

Viacom monetizes their presence and draws the page impressions for their advertisers using their organic reach as follows:

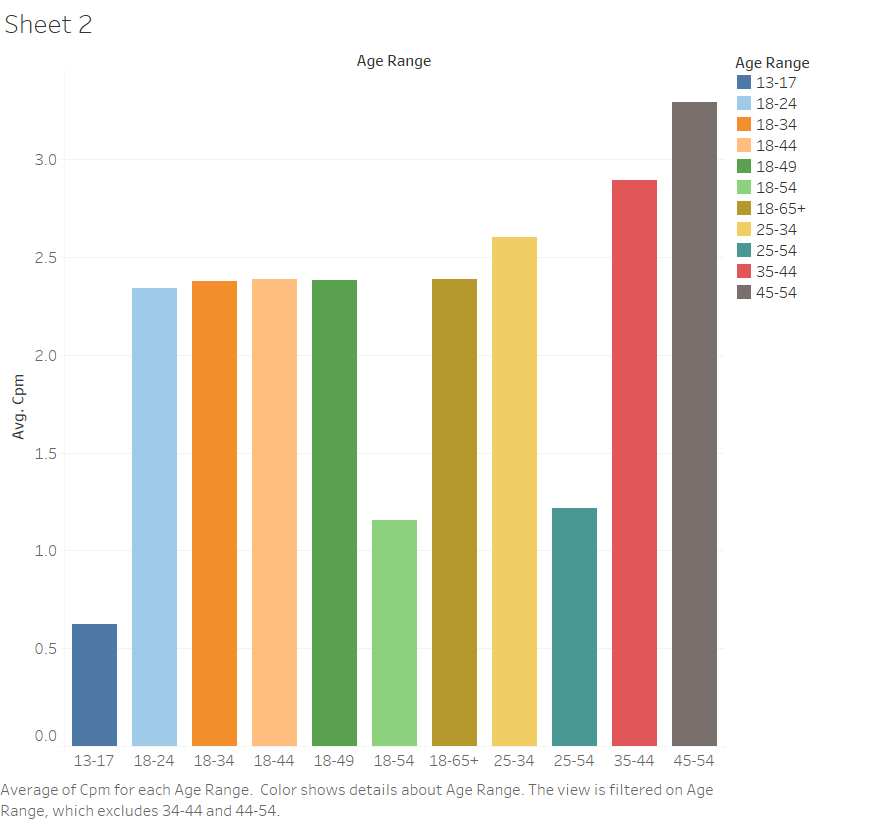


However, it is not possible sometimes for Viacom to target the particular demographics which are important for the advertisers. for the purpose Viacom might need to buy the paid impressions from the social sites like Facebook and earn revenue. However, for that purpose purchase price of the paid impressions must be less than that of the price. We will dive into our analysis using python and tableau to find out whether this product is worth offering or not for Viacom.

**ANALYSIS**

We have page level datasets which holds value for 19 different demographic data from Facebook. And CPM\_estimate dataset which has estimates of CPM on a particular page according to Facebook. hIDs are one of the important attributes of the dataset, it represents the id of website from which the impression was created.

To start with analysis, we performed basic EDA with our data and found out following facts,



From following figure , age group 45-54 are highest contributors of the CPM value where as other age groups need to be kept in focus and Viacom should create more specific content to target them.

We will build a predictive model and impute the CPM values with age\_gender\_unique dataset to determine impact of various factors on the CPM values. We will use the CPM estimates dataset and will divide it into 2018 and 2019 data. To build model and test its accuracy over different parameters. We included age, gender , organic traffic per hID , year as out Important variables for building a Random Forest Model.

We found out that:

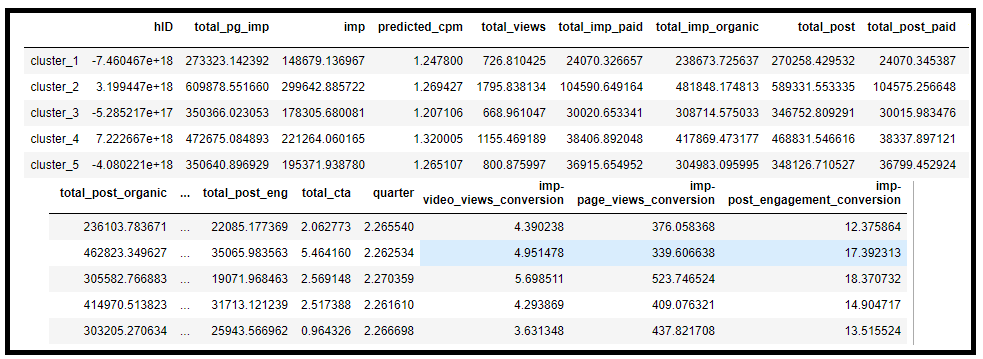
1. Viacom must focus on several hIDs as those hIDs brings biggest share of revenue to Viacom. VH1, MTV, Comedy Central, Nickelodeon, CMT, Logo are some of them to be focused on, however we learned that age group which is younger ( 24-34) tends to avoid the post engagement after clicking or creating and page impression mostly because of content displayed is not of their interest.
2. The page impressions from organic reach does plays a major role in CPM predictions as the accuracy of the model increases after considering the organic values in model building and thus it predicts more accurate CPM values.
3. In prediction of CPM values age range, gender, year are important features.
4. After we imputed the CPM values on the page level dataset, we were able to figure out that Female had the higher cpm values across the year.
5. We need to further dig into our data as demographics data has significant to do with this business model.

|  |  |  |
| --- | --- | --- |
|  | Accuracy : |  |
| Data | With Organic Value | Without Organic value |
| Model 1  [ CPM data 2018 and 2019 , Organic Data for 2018 ] | Accuracy : 88 % | Accuracy : 80 % |
| Model 2  [ CPM data for 2018 , Organic data for 2018] | Accuracy : 84 .34% | Accuracy : 81.4 % |
| Model 3  [ CPM data for 2019 . Organic data 2019 ] | Accuracy : 61.5 % | Accuracy : 33.34 % |

The accuracy changes suggest that we can assume that there could be change in pricing model of Facebook

In next step of the analysis we performed the Demographic clustering and Conversion funneling, to determine which demographics cohorts are most resourceful and are responsible for conversion of impressions into profitable CTA. In demographic clustering we found that,

The age range 25-34 clearly hitting the highest value of CPM=6.7, and more interesting fact is age group 65-100 is holding second highest CPM value as compared to other younger age groups. This insight might be useful for Viacom as most other social sites or shopping platform tend to not produce any content for age group 65 +. Viacom should target more, by producing more specific content for this age group.



We can see the clusters in following table to the left. Cluster 3 has highest conversion for post engagement conversion and has about 309080 total organic impression value.

**CONCLUSION**

From the analysis we can say that Viacom has very good organic reach and it is found out to be drawing huge number of page impressions and also important feature for CPM predictions. There are several hIDs which are most valuable to Viacom such as MTV or VH1 , however Viacom needs to make changes in their customized content in order to take advantage of the younger age groups to indulge into and involve into post engagement which can gain profitable CTA . Moreover, in the race Viacom seems to be falling short of making impact of advertisement of their contents whereas rival companies like Netflix or Apple TV are very much accessible to customers in few clicks. Viacom should think of mobile platform which will have access to all of its content and can be made available to audience according to their interests.