

# **Meta Analytics for the Advertising Panel for Mexico**

**MKTG, Marketing Analytics Project**

Anson Hsu, Aidan Nguyen, Ryan Wong, James  
Baguley

## **Table of Contents:**

- 1. Introduction to Company**
- 2. The Most Successful Ad**
- 3. Trends in Terms of Audience Segmentation**
- 4. Pages that have more Traffic by Region**
- 5. Conclusion/Recommendations**
- 6. References**
- 7. Appendix**

## **1. Introduction to Company**

Sistema Biobolsa is a privately held company focusing on providing renewable energy solutions through biodigester technology. Biodigesters are systems that biologically digests organic material, converting waste into renewable energy and organic fertilizer. Their main customers are small-holder farmers, who after being sold the product are assisted by the company in installation of the system directly onto their properties. Sistema serves 63,363 farms in 31 countries, with their largest customer bases being located in Latin America and Africa. Their goals as a company include positively impacting over 290,000 farms by 2025, and reducing annual global greenhouse gas emissions by 1% by 2030 ([Linkedin](#)).

Sistema is in the process of transitioning from traditional marketing methods, which has dominated their marketing strategy thus far, to online methods such as social media advertisement that will hopefully work to expand their customer base. As they are new to this area of marketing, they do not have a clear idea of which ad campaigns are most successful or which demographics are predominantly engaging with them. This project aims to help answer these questions through analysis of data provided by Sistema, which contains the information they have collected from Meta Analytics regarding engagement with their facebook ads. This data contains information on the engagement, cost and demographic reach of various advertising campaigns conducted by Sistema from the start of 2022 to the end of 2023.

## **2. The Most Successful Ad:**

We're assessing various advertisements' effectiveness by analyzing reach, indicating participant engagement. Our aim is to identify the most accurate predictive model for estimating reach, unveiling key attributes impacting campaign success. By leveraging this model, we'll provide forecasting to Sistema, enabling data-driven decisions for future campaigns and optimizing reach through strategic attribute promotion. This analysis is vital for understanding campaign success drivers and shaping targeted marketing strategies based on data insights.

### **2.1 EDA:**

We are leveraging the datasets "2022\_Mx\_campaigns-Engagement", "2022\_Mx\_Campaigns-performance", "2023\_Mx\_campaigns-Engagement", and "2023\_Mx\_Campaigns-performance" to amalgamate these four datasets into the unified dataset named "mxn". As illustrated in [Figure1](#), it is evident that several variables such as post\_reactions, post\_comments, post\_shares, link\_clicks, follows\_or\_likes, and cpc\_cost\_per\_link\_click contain missing values. The dataset encompasses both categorical and numerical data, as shown in [Figure3](#), enabling us to assess the distribution of each numerical variable. Notably, all numerical data exhibit a right-skewed distribution.

[Figure4](#) provides insights into the frequency distribution of each campaign, revealing that the majority have a frequency of 4, with an average frequency of 3.25 across campaigns. Moreover, [Figure5](#) highlights that most advertisements (ADs) have a frequency of less than 5, with only 6 ADs exceeding a frequency of 10. This suggests that the majority of ADs are associated with one

or two campaigns, and each AD's campaigns are unique. [Figure6](#) to [Figure14](#) present plots depicting the distribution of numerical features grouped by AD. Notably, features such as "reach", "impressions", "post\_reactions", "post\_comments", "post\_shares", and "link\_clicks" exhibit similar distributions, with ADs like "45L" and "Ahorros 20 años" showing higher values likely due to their higher frequency. Consequently, when aggregating by AD, mean aggregation may be more appropriate than summation. The unique aspect of "cost\_per\_link" is its indication that higher costs per link are associated with ADs featuring lower frequencies, as evidenced in the plots.

## 2.2 Preprocessing

In the exploratory data analysis (EDA) phase, it was observed that certain features contain missing values, prompting the need for imputation. For numerical data, missing values will be filled accordingly with mean value. Specifically, for the "follows\_or\_likes" feature, NA values will be replaced with 0, as indicated. Following this, the data will be grouped by "Ad\_name" since the objective of this study is to evaluate the effectiveness of different advertisements (ADs). The aggregation process involves computing the average of metrics such as "post\_reactions", "post\_comments", "post\_shares", "follows\_or\_likes", "cpc\_cost\_per\_link\_click", "impressions", "clicks\_all", "cpc\_all", "amount\_spent", and the target variable "reach". Additionally, "link\_clicks" will be aggregated as a total, and the "ad\_delivery" feature will be categorized as "inactive" if there is no ad delivery within the group. A frequency count will be conducted for each AD, and two new columns, "year\_2022" and "year\_2023", will be created to indicate the number of campaigns in each AD for 2022 and 2023, respectively. The resulting aggregated dataset, named "ad\_mxn", is depicted in [Figure15](#). Furthermore, [Figure17](#) and [Figure18](#) illustrate that most features exhibit a right-skewed distribution, necessitating transformation using the logarithm function (log()). Subsequently, all numerical features will be scaled, and the "inactive" column will be factorized to facilitate further analysis and modeling.

## 2.3 Linear Regression Modeling

We employed stepwise regression to select the most relevant features impacting reach. The final model ([Figure19](#)) includes features such as "avg\_impressions", "tot\_link\_clicks", and "year\_2023" as significant predictors of reach. These features were selected based on their contribution to minimizing the Akaike Information Criterion (AIC) in the stepwise process.

A linear regression model was trained using the selected features on a training dataset, and the model performance was evaluated using n-fold cross-validation. [Figure20](#) displays the coefficients and their significance levels and the R-square equals to 0.908. Notably, the coefficient estimates provide valuable insights into how we determine the successful AD by different value of:

The impact of various factors on reach can be quantified through our analysis. For every one percent increase in followers or likes, there is an estimated 1.04% increase in reach. Similarly, a one percent increase in comments corresponds to a 1.01% increase in reach. Inactive campaigns exhibit a significant 38.67% increase in reach, albeit on a logarithmic scale. For clicks, each one percent increase results in a 2.35% increase in reach. However, a one percent increase in ad frequency leads to a 1.22% decrease in reach. Shares and post reactions contribute positively,

with each one percent increase translating to a 1.23% and 1.03% increase in reach, respectively. On the expenditure side, a one percent increase in average cost per click and average amount spent corresponds to a 1.66% and 1.04% increase in reach, respectively. Conversely, an increase in the number of campaigns in 2022 and average cost per link click results in a 1.01% and 1.44% decrease in reach, respectively.

While examining statistical significance, we find that only avg\_cpc\_cost\_per\_link\_click is significant at the 10% level. Additionally, based on p-values, avg\_cpc\_cost\_per\_link\_click, avg\_post\_shares, and freq\_ad are the top three features with the lowest p-values. These findings suggest that these features may have a more direct impact on our target variable "reach".

Finally, the mean absolute percentage error (MAPE) and mean absolute error (MAE) on training are 3.71% and 0.286 respectively, and the MAPE and MAE on testing are 4.28% and 0.350 respectively. And the QQ-plot shown on [Figure21](#) indicates that our model's residual is normal.

### **3. Trend in Terms of Audience Segmentation:**

In this section of our marketing analytics report, our goal is to assess the effectiveness of Sistema's advertising efforts across diverse demographic groups. Through a thorough analysis of crucial metrics such as "reach", "impressions," and "clicks," we aim to identify trends that shed light on the comparative success of various ad campaigns. Additionally, by examining demographic factors like "age" and "gender," we seek to gain deeper insights into viewer preferences. This comprehensive approach enables us to provide Sistema with valuable insights to tailor future ad campaigns more precisely to different audience segments. For further segmentation, we plan to employ Cluster Analysis, allowing us to group ads with similar performance metrics.

#### **3.1 EDA:**

We used the datasets “2022\_Mx\_Campaigns-by-demographics” and “2023\_MX\_Campaigns-by-Demographics.” We merged these two datasets in order to get a complete picture for these two years. For this analysis we focused on the columns describing the reach, impressions, clicks for each ad campaign, separated by gender and age.

Preliminary exploration of the data reveals there are 6 values for age groups and 3 genders (“male,” “female” and “unknown”), dividing each campaign name into 17 different rows, with each row containing reach, impression and click information. As the data rendered entries that were equal to zero as NA, we replaced those with 0 in order to enable calculation of averages. Completely empty rows and the empty columns “ad\_delivery,” “starts” and “ends” were removed. For each campaign, impressions is the highest metric, with reach usually 50 to 75% of it, and clicks a much smaller proportion of both ([figure22](#)). For age, the majority of engagement comes from people 25-54, and a minority from people aged 55 to 65+ and 18 to 24, with a similar proportion of overall impressions/reach/clicks. People aged from 35 to 54 maintain a higher proportion of clicks than other age groups, indicating they might be more susceptible to this kind of advertisement ([figure23](#)). Exploration of the gender column reveals that there are far more males engaging with the ads compared to females, as well as a very small proportion of people whose gender is unknown ([figure24](#)).

Additionally, we conducted an analysis focusing on the "clicks" metric, as it provides insights into the audience's level of interest in our ads, indicated by their willingness to click. Our examination reveals that the majority of clicks originate from individuals aged 25-44 ([figure25](#)), with some outliers and most of these individuals are male ([figure26](#)). There is also a noticeable number of clicks coming from individuals aged 45-54.

From these findings, it becomes apparent that individuals in older age groups and females show less interest in the ads compared to their counterparts. This insight allows us to tailor our marketing strategies to better target the audience segments that exhibit higher engagement levels.

### 3.2 Preprocessing

We took several steps to prepare the data for a cluster analysis that would help determine the overall engagement between different demographics. As mentioned before entries reading NA were set equal to 0 and completely empty rows and the empty columns were removed. As the "unknown" gender made up a very small proportion of the overall engagement and would complicate interpretation of any clustering models, we decided to remove it before performing any modeling. We then had to convert categorical variables such as age and gender into formats that the models could work with. To this end we converted age into a numeric variable with one equaling male and zero equaling female, and took the median of each age range in order to convert them into numeric variables as well. As engagement among different demographics was the target, we selected only reach, engagement, and clicks to be inserted into the clustering model.

### 3.3 Hierarchical Clustering

For Hierarchical Clustering, we first try to determine the ideal number of clusters by using the Within-Cluster Sum of Squares (WSS) ([figure27](#)) and Silhouette ([figure28](#)) methods. After examining the graphs obtained from both techniques, we decided to try to split the data into two different numbers of clusters: one with 5 clusters and the other 6 clusters, based on the metrics "reach," "impression," and "click". Afterwards, we compared the results from both methods to ascertain the optimal number of clusters for this clustering method.

In analyzing the dendrogram ([figure29](#)) for 5 clusters, we observe distinct patterns among the clusters. Notably, clusters 2 (blue) and 4 (purple) exhibit similarities, showing closer resemblance to cluster 1 (green) than to the remaining clusters. On the contrary, clusters 3 and 5 appear different, as evidenced by their dendrogram lengths. To delve deeper into these clusters, we pruned the tree and tabulated the number of observations in each cluster. The results reveal varying sizes: 14 observations in cluster 1, 1 observation in cluster 2, 52 observations in cluster 3, 4 observations in cluster 4, and 495 observations in cluster 5.

Finally, we computed the mean age and gender for each cluster to understand their demographics and engagement levels ([figure30](#)). Notably, cluster 5 emerges as the largest but least engaged demographic, with an average age exceeding 41 and predominantly female participants. In contrast, cluster 2 exhibits the highest engagement, with an average age around 29 and predominantly male participants. The remaining clusters fall in between, with average ages in the 30s and a predominantly male demographic. Despite its large size, cluster 5 exhibits the lowest

engagement rate compared to the others.

Examining the dendrogram ([figure31](#)) for 6 clusters reveals a striking similarity to the 5-cluster dendrogram, with the key difference being the split of cluster 5 into two distinct clusters. These two clusters exhibit similarities with cluster 3, while the remaining clusters (1, 2, and 4) maintain their relationships as observed in the 5-cluster analysis. The composition of observations within clusters 1, 2, 3, and 4 remains consistent with our previous findings. However, cluster 5 has now been split into two clusters, resulting in 43 observations in cluster 5 and 452 observations in cluster 6. We then proceeded to compute the mean age and gender for each cluster to gain insights into their demographics and engagement levels ([figure32](#)). Notably, dividing the dataset into 6 clusters provided additional granularity, particularly with the split in cluster 5. This division resulted in two distinct groups with similar mean ages around 41 but different gender distributions. Specifically, cluster 5 predominantly comprises males, while cluster 6 is predominantly female. Despite cluster 6 having the highest number of observations, it exhibits the lowest engagement compared to the other clusters. Based on these findings, we concluded that dividing the dataset into 6 clusters offers a more comprehensive understanding of the data, enabling better segmentation for targeted marketing strategies.

### 3.4 K-means Clustering

Attempting to find the optimal number of clusters with WSS and Silhouette puts the optimal number of clusters for k means between four and five, as this is when the slope begins to level out ([figure33](#), [figure34](#)). We thus decided to compare the clusters formed between k=4 and k=5 for this model. Computing centers at k=4 reveals 4 different groups of varying engagement (impressions, reach and clicks). The clusters demonstrating the lowest levels of engagement were the ones where the data was most likely to be over 40 and predominantly female, while the ones demonstrating the highest engagement contained data that was overwhelmingly male and early 30s, with those in the middle tending to be mid 30s to 40 and male ([figure35](#)). As we took the median of each age group, the 25-34 age group is counted as 29.5, and 35 to 44 as 39.5. Therefore we can probably extend this early 30s group to include ages such as mid twenties and the over forties group to include people closer to 50. Regarding click-through rates, the lowest group, those mostly female and over 40, are the most likely to click on the ads by a small amount, from .7 to .4% more likely.

Adjusting k to equal five results in little overall change in the clusters, with the early 30s/late 20s male demographic still making up the majority of engagement and females around 40 making up the least. However, it does confirm that males around and over 40 make up much less of engagement compared to males in their 30s, with the second lower cluster being mostly male and mid 40s and the third lowest all male and early 40s. Click-through rates are similarly unchanged, with the lowest engagement group still having the highest proportion of clicks ([figure36](#)). Performing ANOVA on this cluster analysis reveals that the feature most important for separation of clusters is reach at 7% significance, then impressions at 9% and then clicks at 12%. The explained variance ratio indicates that together these features account for 94% of the variance in the model. Finally, a 2-dimensional graph displays some other qualities of the clusters, such as that clusters 1, 2 and 3 have wider variance and that the clusters with the lowest engagement overlap, most likely where groups are similarly female or over 40 ([figure37](#)).

## 4. Pages that have more Traffic by Region

Looking at Sistema's Google Analytics page, They display the most popular pages on their website and the countries that visit them the most [Figure 38]. They also display the most popular in the last 90 days, which are similar to the last 12 months [Figure 43]. So we can conclude that their popularity and reach generally remain the same throughout the year when sorted by number of views. There were a total of 47,605 users on the whole Sistema website within the last 12 months [Figure 39] and 20,927 users within the last 90 days [Figure 44]. Notably, the inclusion of Masovian Voivodeship (Poland) as a new region on the Biogester page within the last 90 days suggests an expanding reach and engagement in previously untapped areas.

When sorted by the pages by region based on the average views per user, the pages and regions listed are different than described above. [Figure 40](#) displays the most frequent regions and page visits.. This underscores the platform's potential for growth in diverse geographical markets. Moreover, the analysis based on average views per user highlights varying levels of engagement across regions, underscoring the need for tailored strategies to optimize user experience and retention. For the last 90 days, about half are the same as within the last 12 months but the other half are different. Interestingly, none of the regions are duplicates in the top 10 within the last 90 days; each region is unique[[Figure 45](#)]. These insights underscore the significance of considering the unique considerations of customers in different geographical areas.

## 5. Conclusion/Recommendations

Key factors like cost efficiency of attracting clicks, the level of engagement through shares, and frequency of ad placements significantly impact reach. Leveraging these insights, Sistema can optimize campaigns, allocate resources effectively, and enhance audience engagement through strategic attribute promotion. The model's performance metrics, including MAPE and MAE, demonstrate its accuracy in predicting reach, while the QQ-plot confirms the model's reliability.

As demographic information is dominated by Males aged late twenties to mid thirties, Sistema may want to broaden its online marketing approach to reach more females and older people. This could entail placing their advertisements on Facebook pages that receive more traffic from these demographics, or expanding their digital presence to other social media platforms such as Instagram and X.

Based on the tables in the Google Analytics platform, certain pages such as the Biogester solution and Opportunities consistently attract significant traffic across all regions. Regional variations in user engagement should also be taken into account. For example, among the regions analyzed, Maharashtra and Karnataka in India, along with Mexico City in Mexico, emerge as particularly significant due to their consistent presence in the top-ranking pages by views and users. These regions represent key markets so having advertisements from those regions on those pages would make sense. Sistema should invest in region-specific content and marketing strategies to cater effectively to diverse audiences, such as having products that are more popular from those regions and sponsoring companies from those countries.

## **6. References**

Sistema.bio / Sistema BIOBOLSA. LinkedIn.  
(n.d.)<https://www.linkedin.com/company/sistema-biobolsa/>)

## 7. Appendix

### 2. The Most Successful AD:

Figure 1

	campaign_name	account_name	ad_set_name	ad_name	ad_delivery
post_reactions	0	0	0	0	0
post_comments	5	43	59	2	171
currency	cpc_cost_per_link_click		reach	impressions	clicks_all
cpc_all	0	2	0	0	0
amount_spent	0	0	year		
year	0	0			

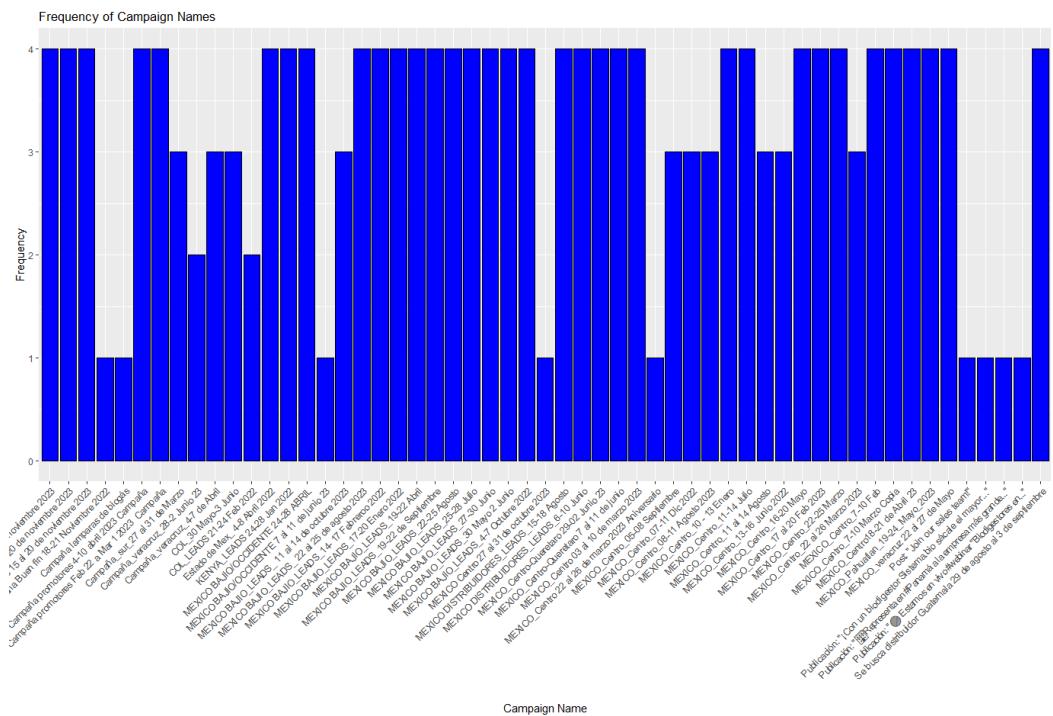
Figure 2

```
> str(mx)
'data.frame': 179 obs. of 18 variables:
 $ campaign_name      : chr "Campana Buen fin 18-21 Noviembre 2022" "MEXICO BAJIO LEADS_19-22 Abril" "MEXICO BAJIO LEADS_14-17 Febrero 2022" ...
 $ account_name        : chr "Buen Manejo del Campo SA de CV" "Buen Manejo del Campo SA de CV" "Buen Manejo del Campo SA de CV" ...
 $ ad_set_name          : chr "conjunto Buen fin" "19-22 Abril 2022" "30 Mayo al 2 Junio 2022" "14-17 Febrero 2022" ...
 $ ad_name              : chr "70% descuento" "Ahorros 20 años" "Ahorros 20 años" "Ahorros 20 años"
 $ ad_delivery          : chr "not_delivering" "not_delivering" "not_delivering" "not_delivering"
 $ post_reactions       : num 77 182 188 211 330 206 341 339 224 258 ...
 $ post_comments         : num 18 32 38 67 67 39 133 104 56 84 ...
 $ post_shares          : num 5 23 32 17 64 37 59 44 34 27 ...
 $ link_clicks          : num 240 526 457 656 903 ...
 $ follows_or_likes     : num 1 NA NA NA NA NA NA NA NA ...
 $ currency              : chr "MXN" "MXN" "MXN" "MXN" ...
 $ cpc_cost_per_link_click: num 3.333 1.378 1.436 0.942 0.677 ...
 $ reach                 : num 23785 28656 37352 52367 51632 ...
 $ impressions           : num 37103 36667 45272 72157 64351 ...
 $ clicks_all            : num 1458 2989 2973 3853 4606 ...
 $ cpc_all               : num 0.549 0.242 0.221 0.16 0.133 ...
 $ amount_spent          : num 800 725 656 618 611 ...
 $ year                  : num 2022 2022 2022 2022 2022 ...
```

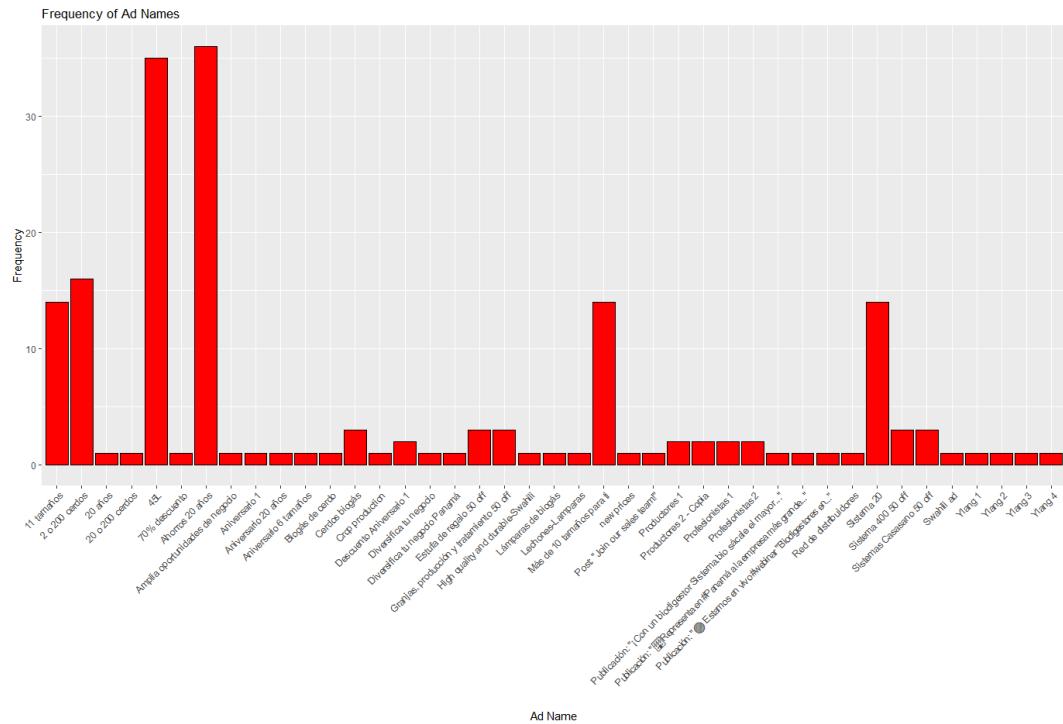
Figure 3

```
> summary(mx)
campaign_name      account_name      ad_set_name      ad_name      ad_delivery      post_reactions      post_comments
Length:179          Length:179          Length:179          Length:179          Length:179          Min.   : 1.00      Min.   : 1.00
Class :character    Class :character    Class :character    Class :character    Class :character    1st Qu.: 9.00      1st Qu.: 2.00
Mode  :character    Mode  :character    Mode  :character    Mode  :character    Mode  :character    Median : 35.50      Median : 10.00
                                         Mode  :character    Mode  :character    Mode  :character    Mean   : 65.75      Mean   : 20.67
                                         Mode  :character    Mode  :character    Mode  :character    3rd Qu.: 95.75      3rd Qu.: 29.00
                                         NA's   :5          NA's   :5          NA's   :5          Max.  :341.00      Max.  :133.00
                                         NA's   :5          NA's   :5          NA's   :5          NA's   :5          NA's   :43
post_shares      link_clicks      follows_or_likes      currency      cpc_cost_per_link_click      reach      impressions
Min.   : 1.00      Min.   : 1.00      Min.   :1.000      Length:179      Min.   : 0.04571      Min.   : 26      Min.   : 34
1st Qu.: 2.00      1st Qu.: 18.00     1st Qu.:1.000      Class :character  1st Qu.: 0.79088      1st Qu.: 1390     1st Qu.: 1877
Median : 6.00      Median : 94.00     Median :1.000      Mode  :character  Median : 1.30984      Median : 5428      Median : 8500
Mean   :11.8       Mean   :201.7      Mean   :1.125      NA's   :171        Mean   : 1.97502      Mean   :12457      Mean   :17392
3rd Qu.:18.0       3rd Qu.:283.0      3rd Qu.:1.000      NA's   :2          Mean   : 2.06444      3rd Qu.:17912     3rd Qu.:24469
Max.  :64.0        Max.  :1188.0      Max.  :2.000      NA's   :2          Max.  :14.28286      Max.  :81729      Max.  :98739
NA's   :59         NA's   :2          NA's   :171      NA's   :2          NA's   :2          NA's   :2          NA's   :2
```

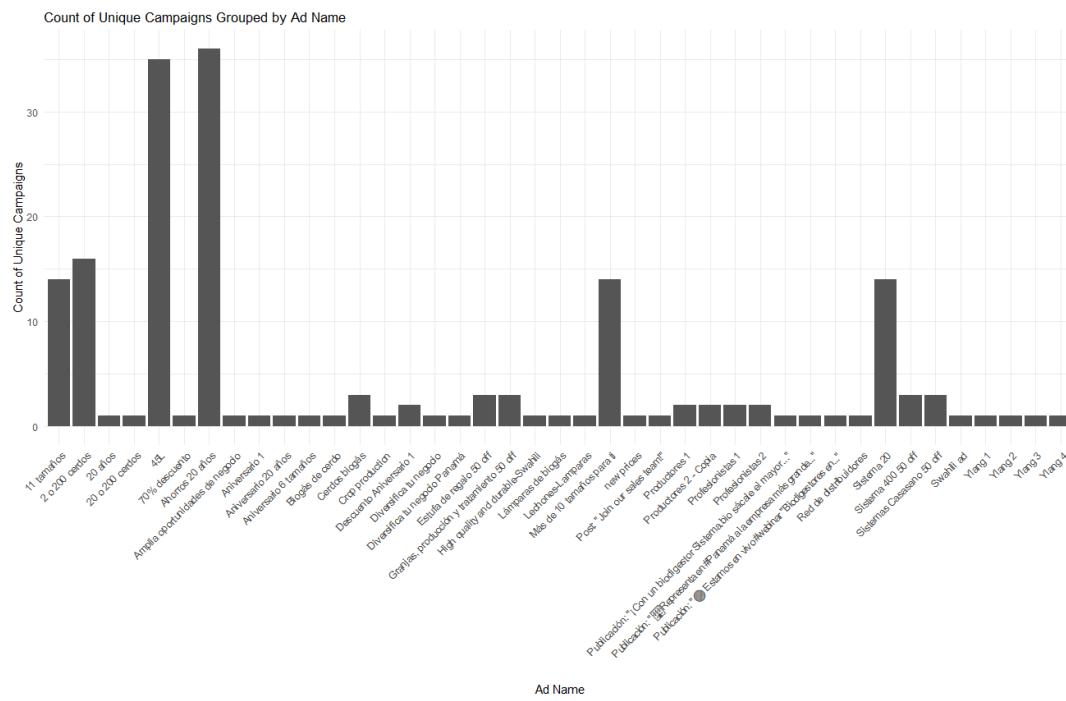
**Figure 4**



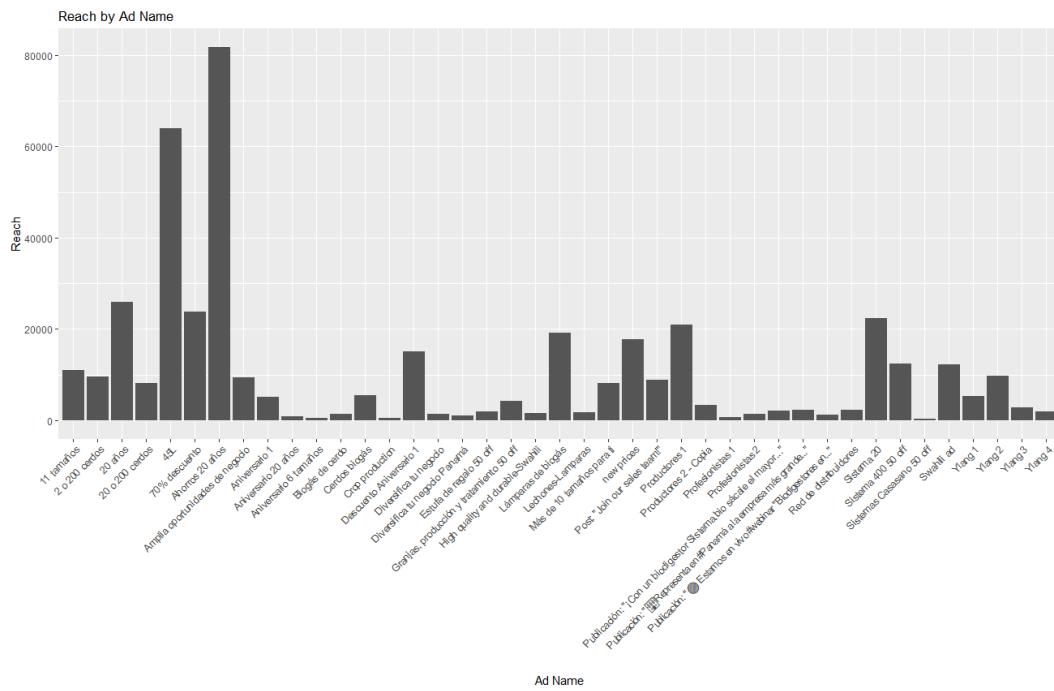
**Figure 5**



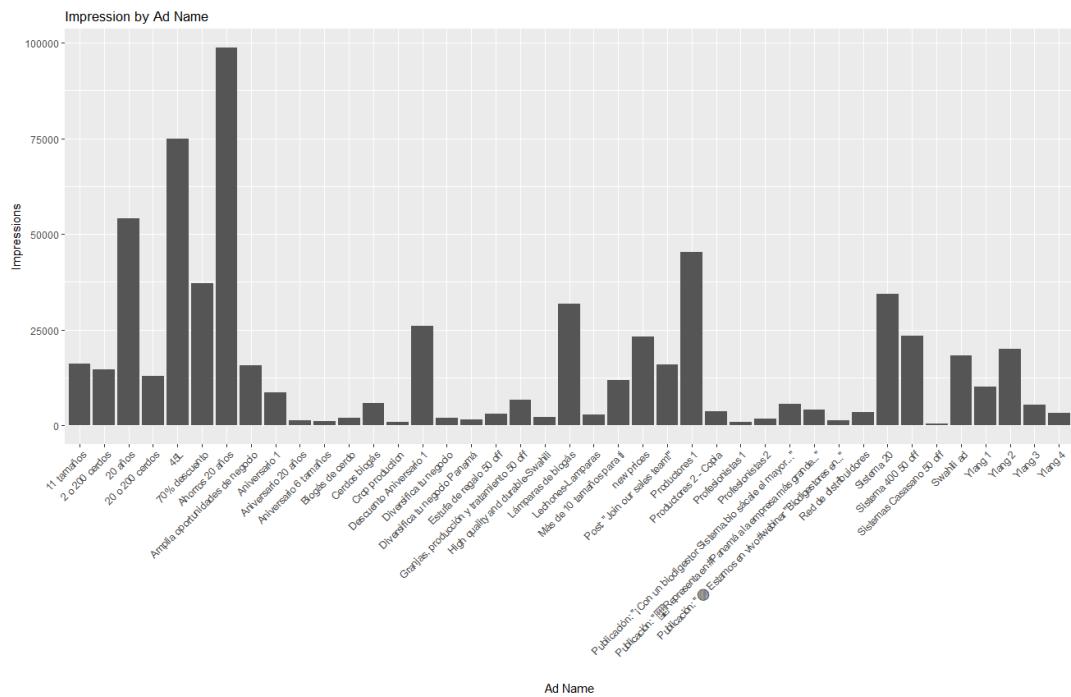
**Figure 6**



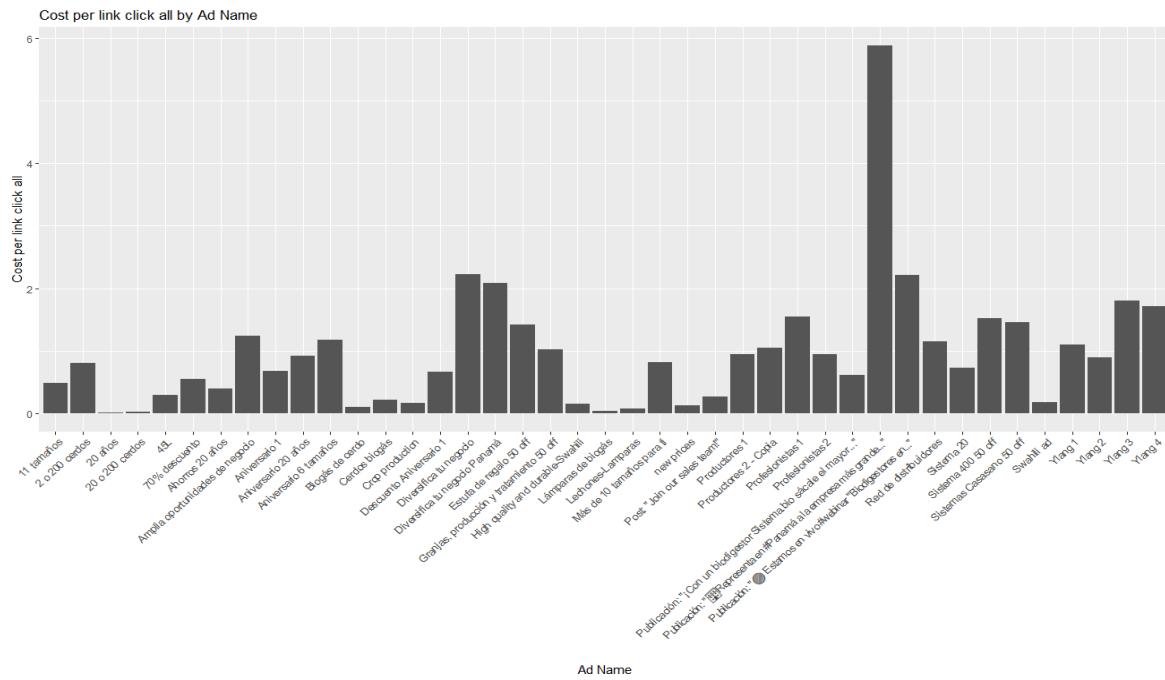
**Figure 7**



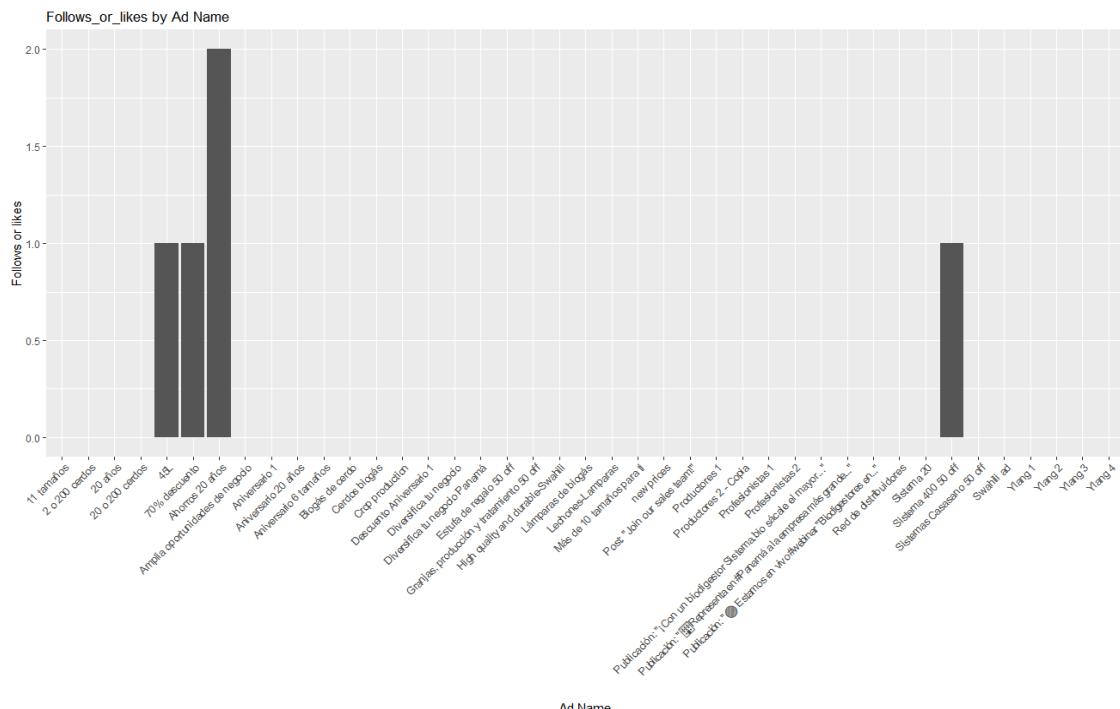
**Figure 8**



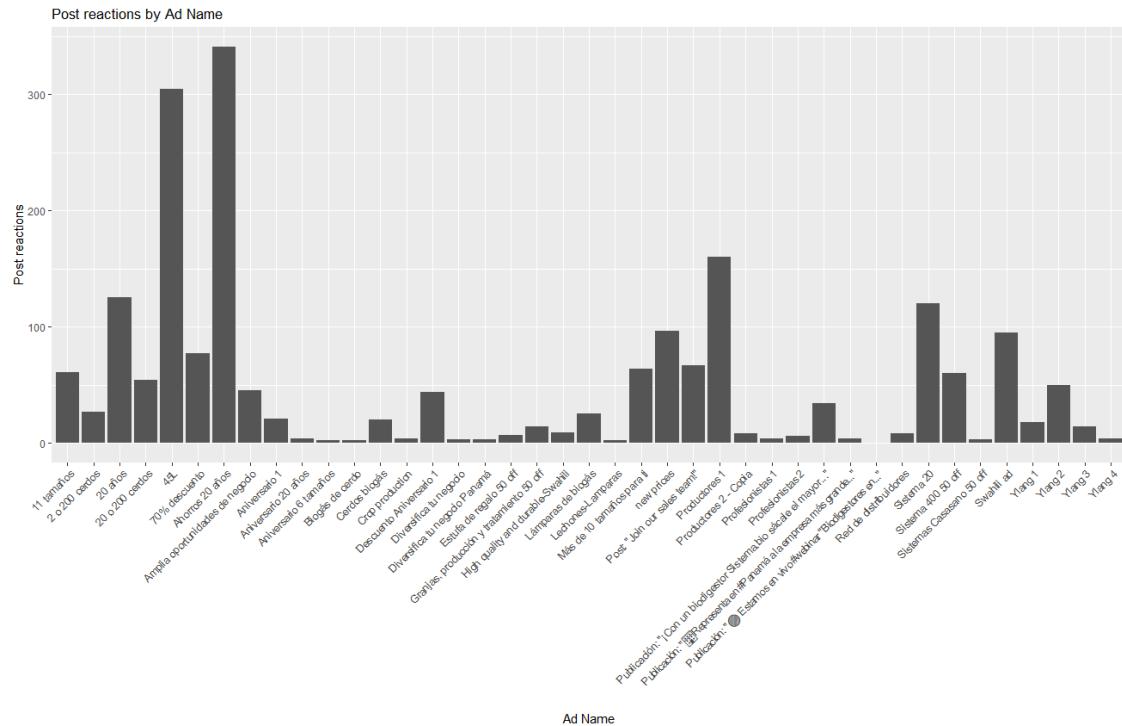
**Figure 9**



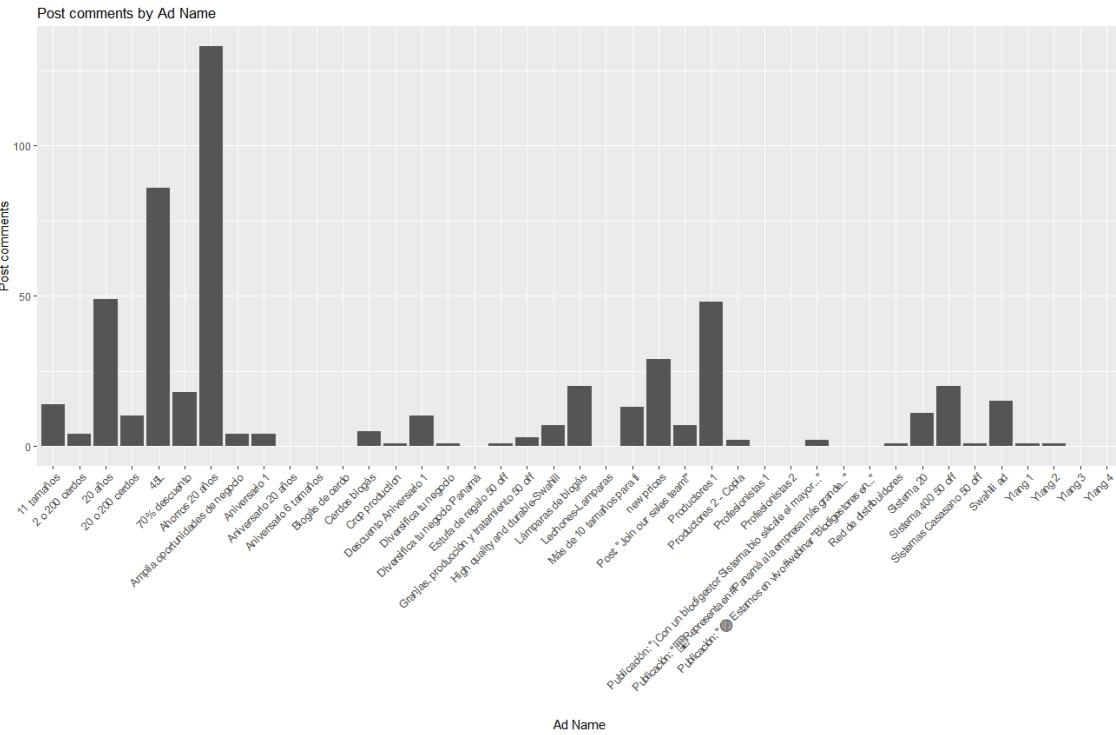
**Figure 10**



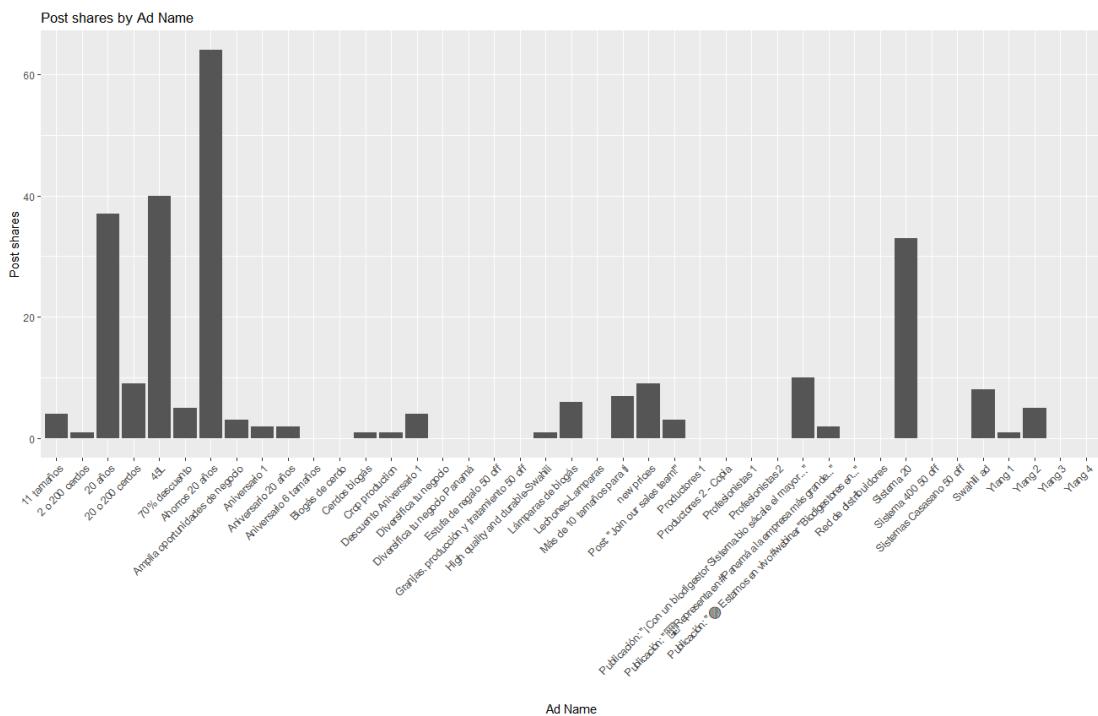
**Figure 11**



**Figure 12**



**Figure 13**



**Figure 14**

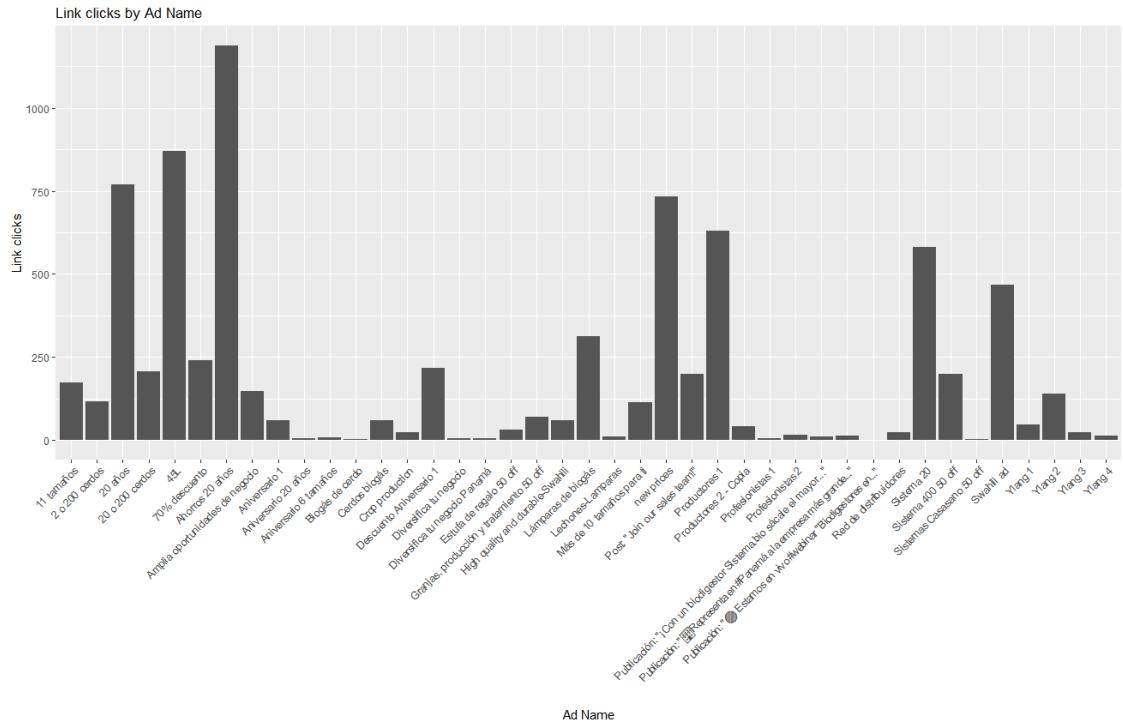
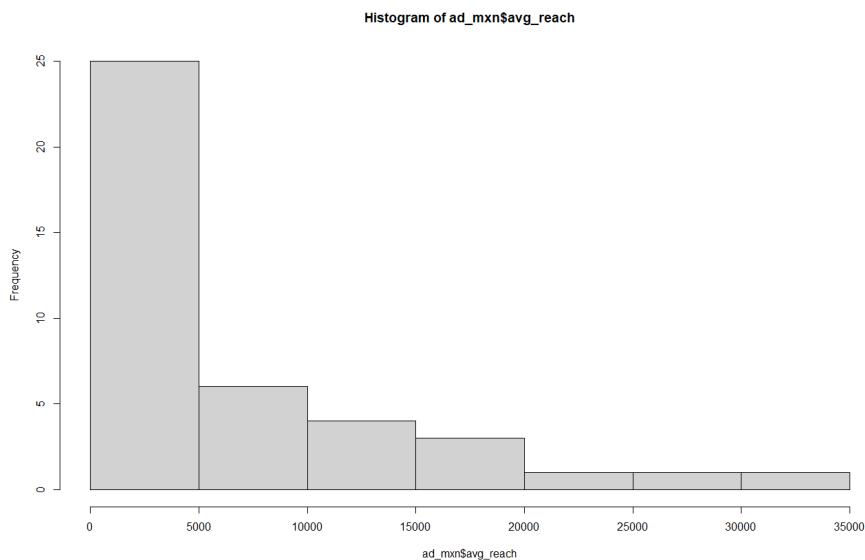


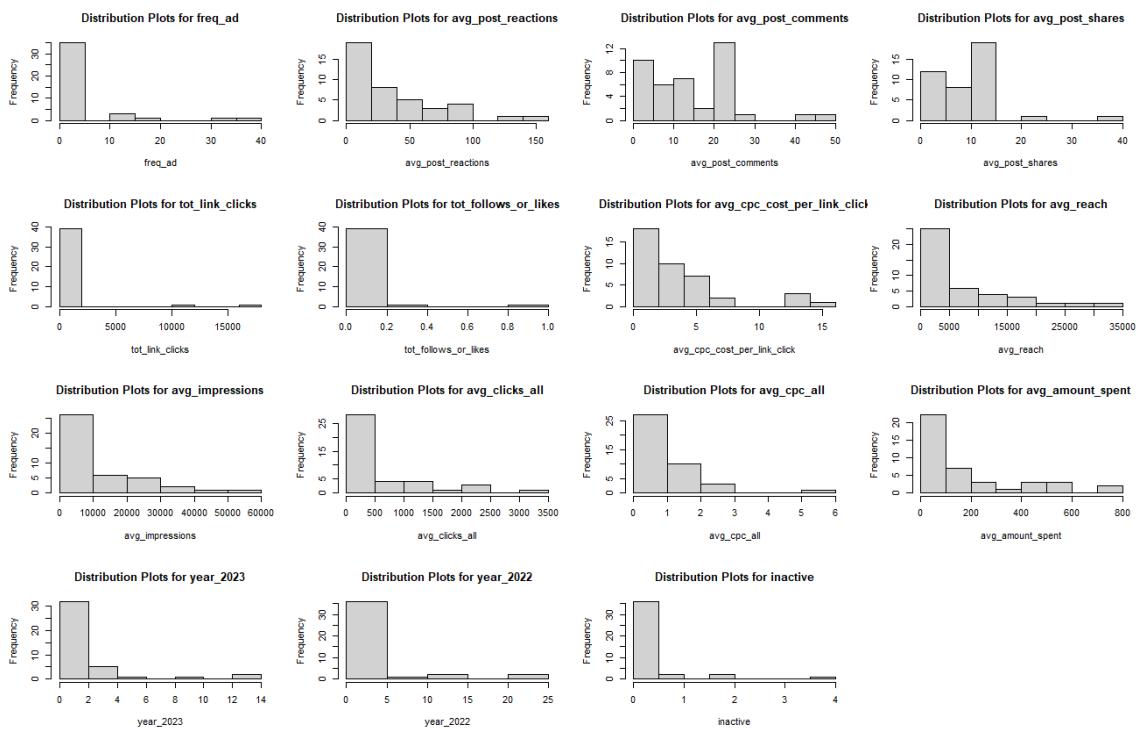
Figure 15

```
tibble [41 x 16] (S3: tbl_df/tbl/data.frame)
$ ad_name          : chr [1:41] "11 tamaños" "2 o 200 cerdos" "20 años" "20 o 200 cerdos"
$ freq_ad          : int [1:41] 14 16 1 1 35 1 36 1 1 1 ...
$ avg_post_reactions : num [1:41] 26.4 NA 125 54 NA ...
$ avg_post_comments : num [1:41] NA NA 49 10 NA ...
$ avg_post_shares  : num [1:41] NA NA 37 9 NA ...
$ tot_link_clicks  : num [1:41] 953 370 771 207 10122 ...
$ tot_follows_or_likes : num [1:41] NA NA NA NA 1 NA NA NA NA ...
$ avg_cpc_cost_per_link_click: num [1:41] 1.5528 1.0873 0.0457 0.0529 1.0323 ...
$ avg_reach         : num [1:41] 4434 1565 25936 8154 18601 ...
$ avg_impressions  : num [1:41] 6263 2055 54137 12889 24489 ...
$ avg_clicks_all   : num [1:41] 386.8 90.2 2368 571 2049.8 ...
$ avg_cpc_all       : num [1:41] 0.288 0.3245 0.0149 0.0192 0.1423 ...
$ avg_amount_spent : num [1:41] 108.6 28 35.2 11 242.1 ...
$ year_2023         : int [1:41] 3 2 0 0 13 0 14 1 1 1 ...
$ year_2022         : int [1:41] 11 14 1 1 22 1 22 0 0 0 ...
$ inactive          : int [1:41] 1 0 0 2 0 0 0 0 1 ...
```

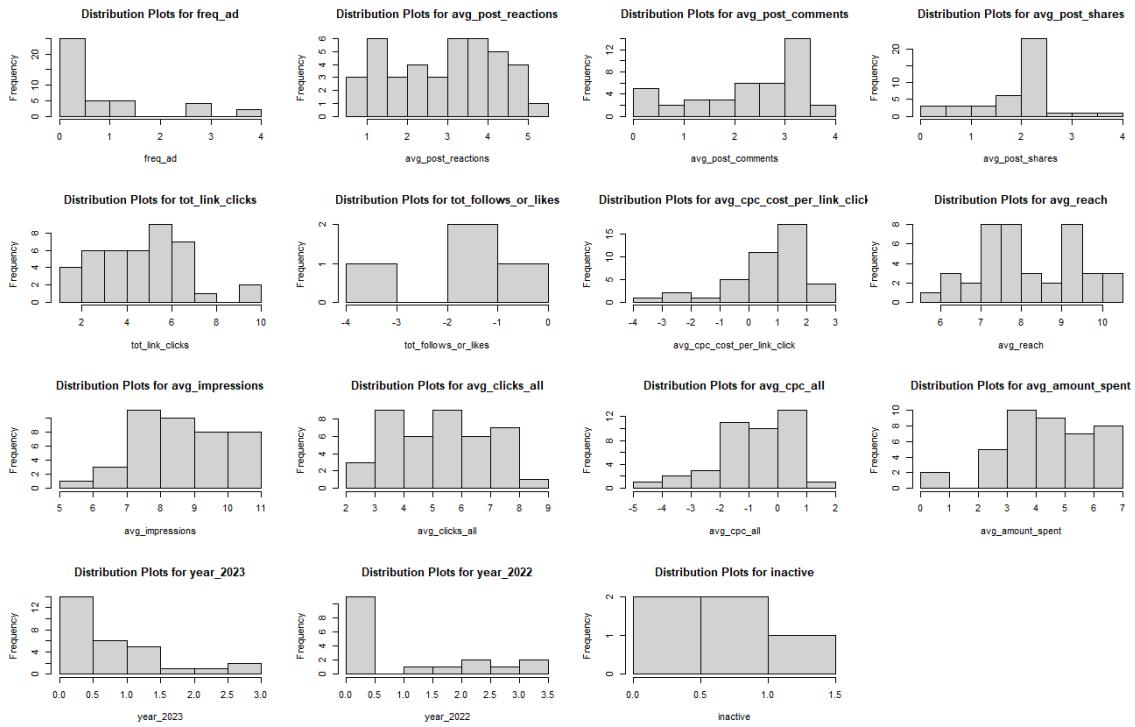
Figure 16



**Figure 17**



**Figure 18**



**Figure 19**

Step: AIC=-154.49  
 $\text{avg\_reach} \sim \text{avg\_impressions} + \text{tot\_link\_clicks} + \text{year\_2023}$

	df	sum of sq	RSS	AIC
<none>			0.7791	-154.493
+ tot_follows_or_likes	1	0.0346	0.7444	-154.356
+ avg_post_comments	1	0.0262	0.7529	-153.894
+ inactive	1	0.0233	0.7558	-153.737
+ avg_clicks_all	1	0.0202	0.7589	-153.570
+ freq_ad	1	0.0172	0.7619	-153.407
+ avg_post_shares	1	0.0130	0.7660	-153.183
- year_2023	1	0.0656	0.8446	-153.179
+ avg_cpc_all	1	0.0104	0.7686	-153.045
+ avg_post_reactions	1	0.0057	0.7733	-152.795
+ year_2022	1	0.0055	0.7736	-152.782
+ avg_amount_spent	1	0.0008	0.7783	-152.534
+ avg_cpc_cost_per_link_click	1	0.0000	0.7790	-152.495
- tot_link_clicks	1	0.3239	1.1030	-142.238
- avg_impressions	1	25.2501	26.0292	-12.628

**Figure 20**

```

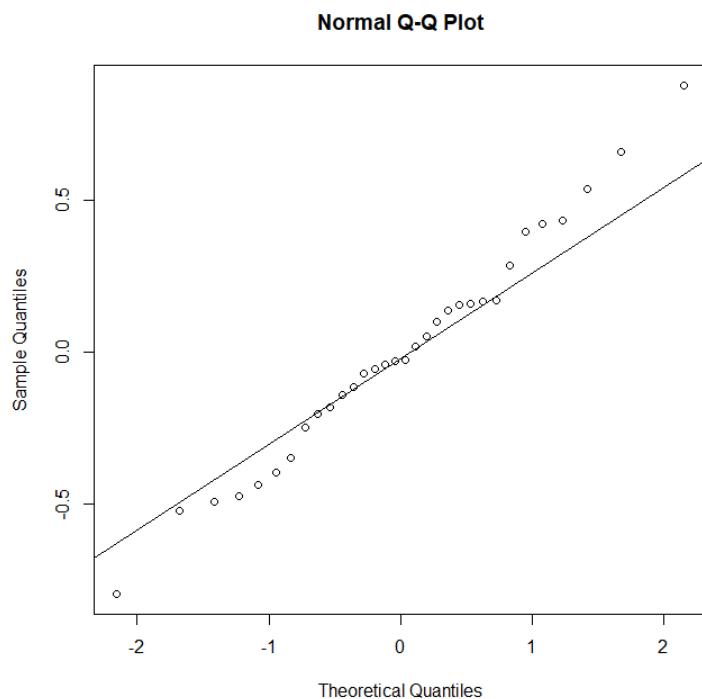
Residuals:
    Min      1Q  Median      3Q     Max 
-0.79525 -0.21335 -0.02811  0.16679  0.87755 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept) 4.814444   0.756937  6.360 3.31e-06 ***
tot_follows_or_likes 0.036070   0.028876  1.249  0.2260  
avg_post_comments 0.007792   0.097852  0.080  0.9373  
inactive1 0.386652   0.347774  1.112  0.2794  
avg_clicks_all 0.852310   0.767966  1.110  0.2802  
freq_ad -0.198324   0.149666 -1.325  0.2001  
avg_post_shares 0.203905   0.145575  1.401  0.1766  
avg_cpc_all 0.509022   0.815986  0.624  0.5398  
avg_post_reactions 0.027724   0.106019  0.262  0.7964  
year_2022 -0.011822   0.027346 -0.432  0.6701  
avg_amount_spent -0.126964  0.766613 -0.166  0.8701  
avg_cpc_cost_per_link_click -0.361411  0.199031 -1.816  0.0844 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4614 on 20 degrees of freedom
Multiple R-squared:  0.908,    Adjusted R-squared:  0.8574 
F-statistic: 17.95 on 11 and 20 DF,  p-value: 5.163e-08

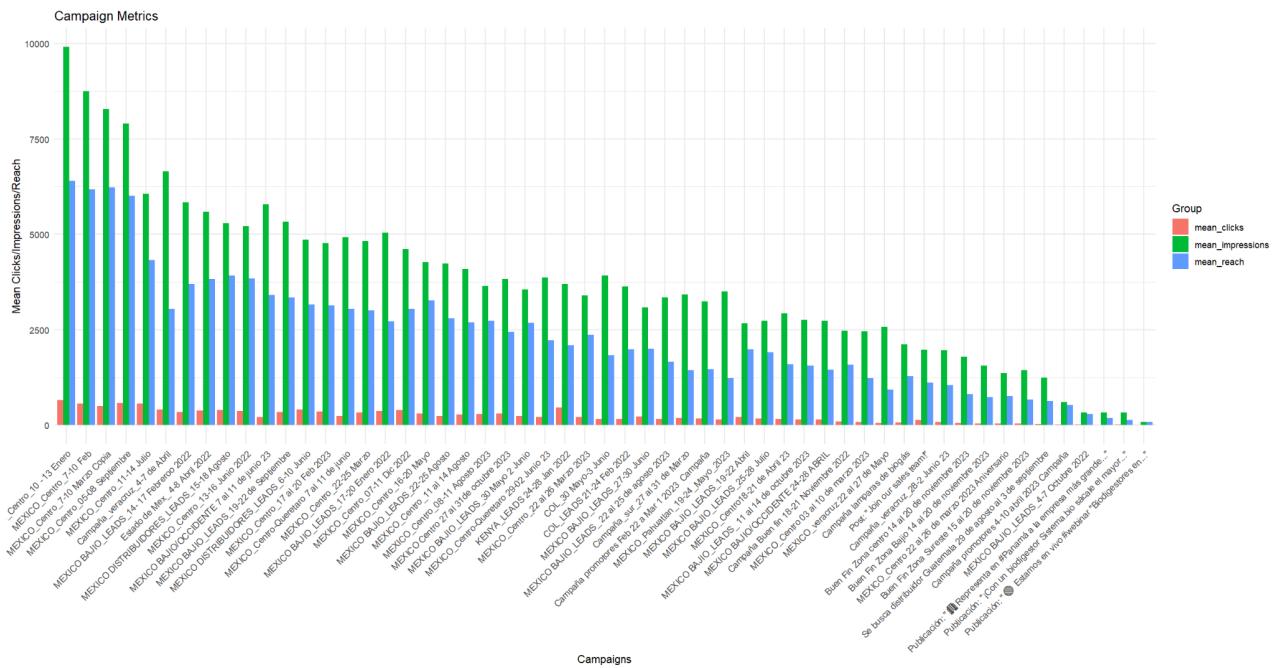
```

**Figure 21**

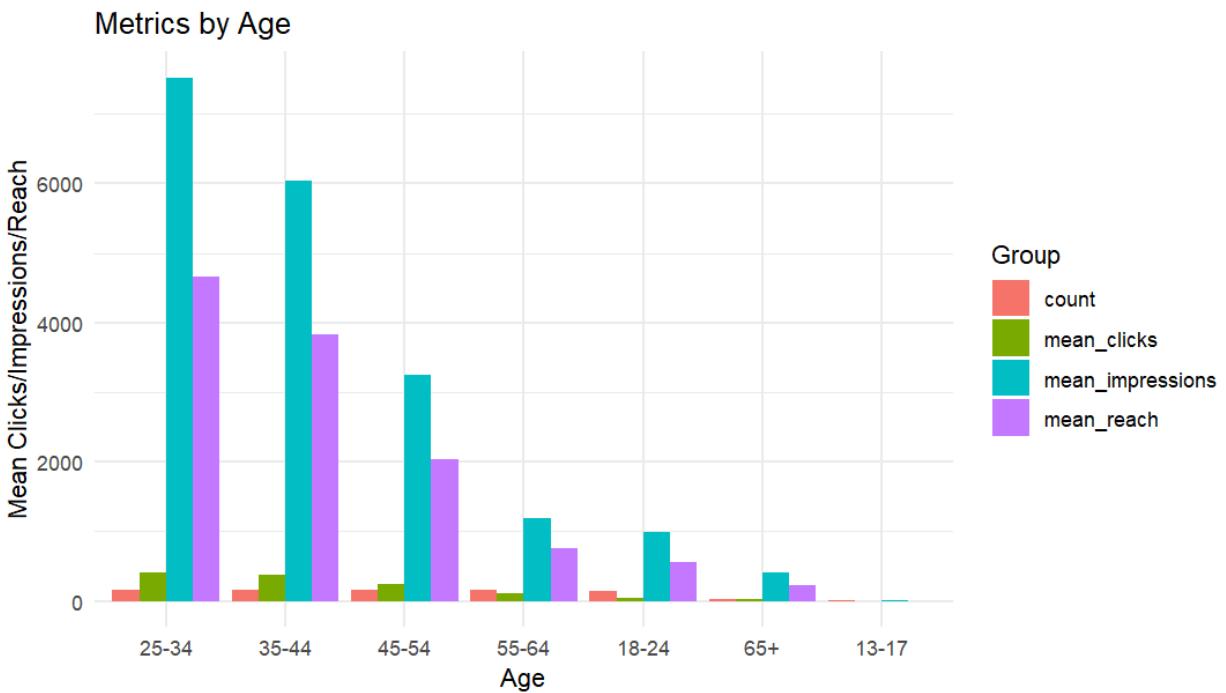


### 3. Trend in Terms of Audience Segmentation:

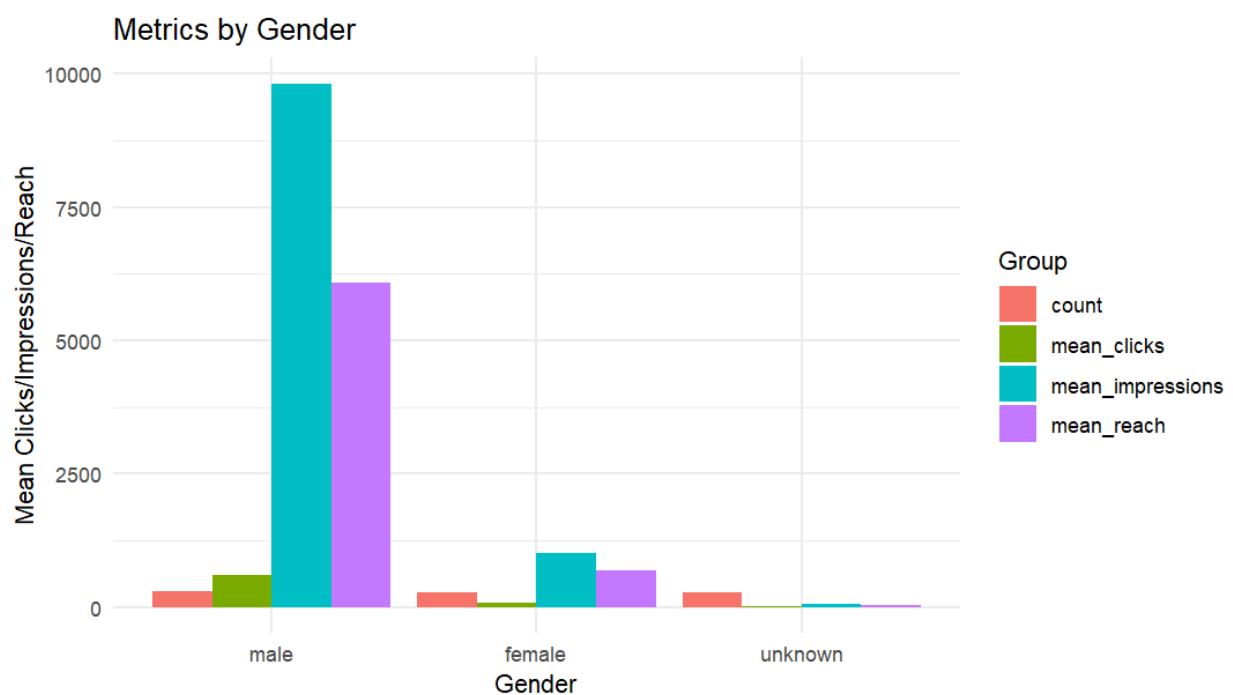
**Figure 22**



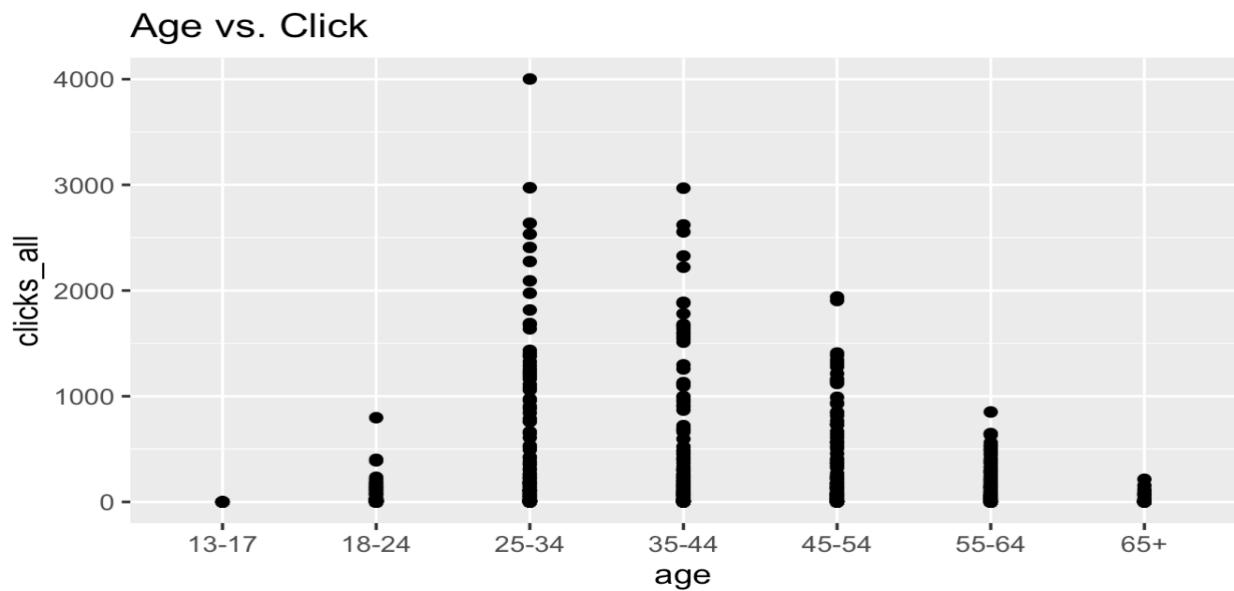
**Figure 23**



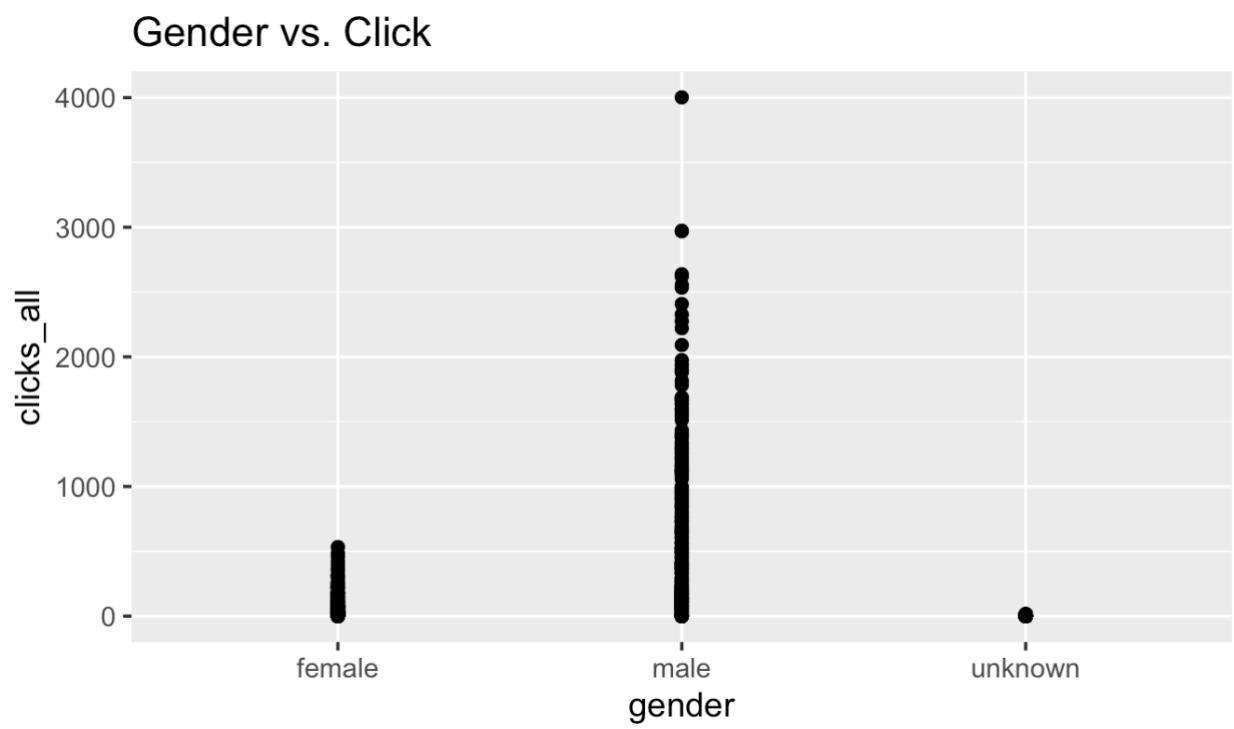
**Figure 24**



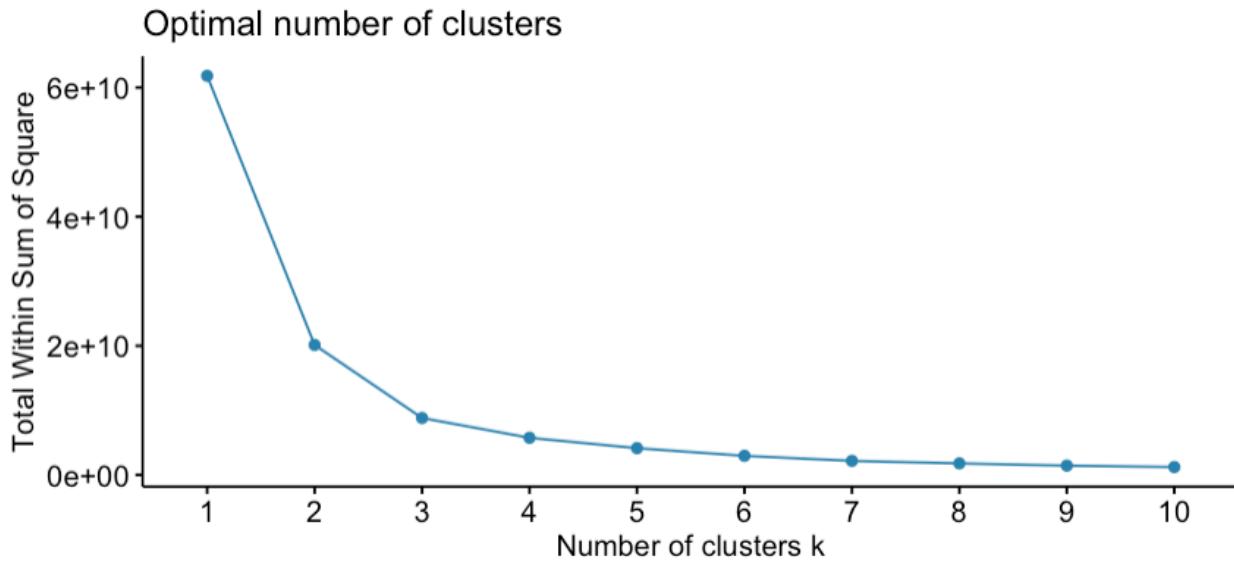
**Figure 25**



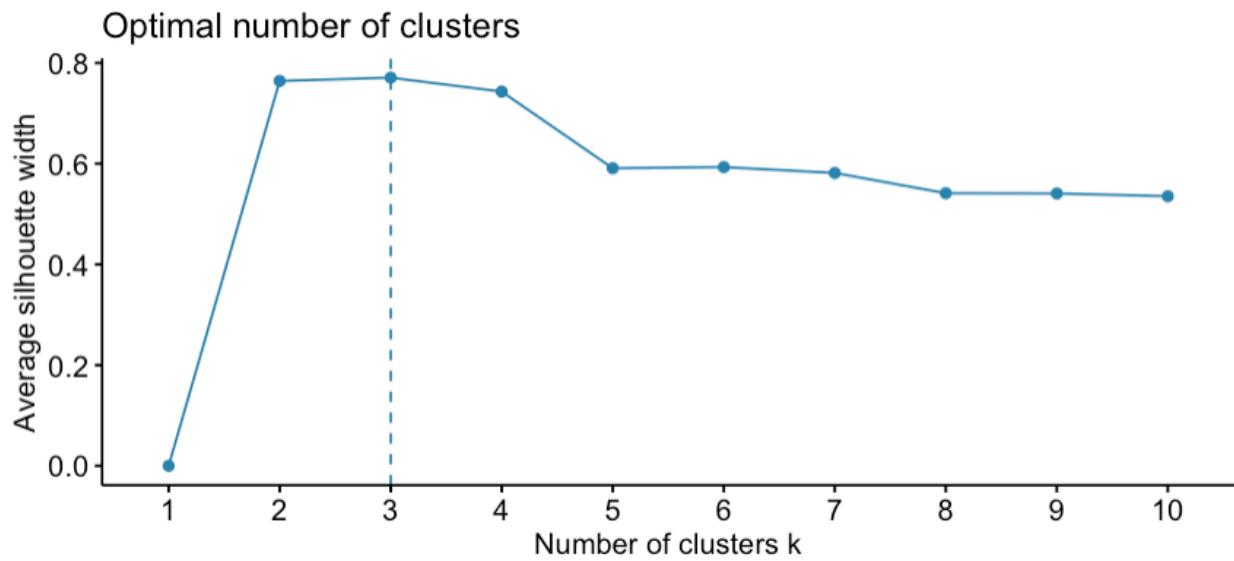
**Figure 26**



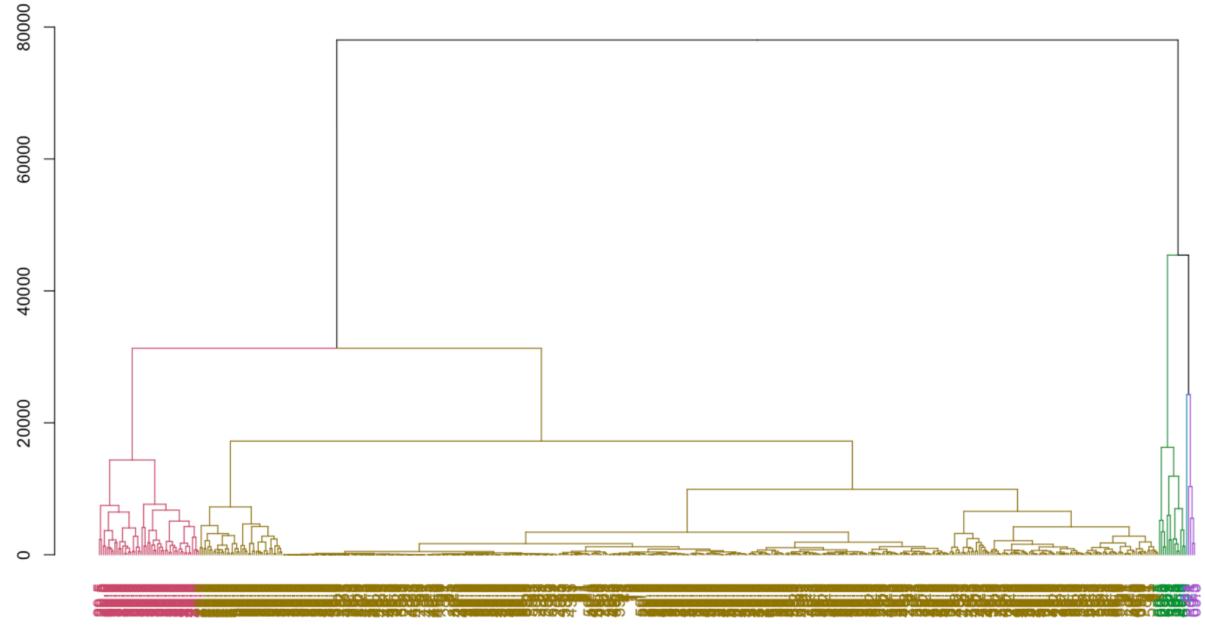
**Figure 27**



**Figure 28**



**Figure 29**



**Figure 30**

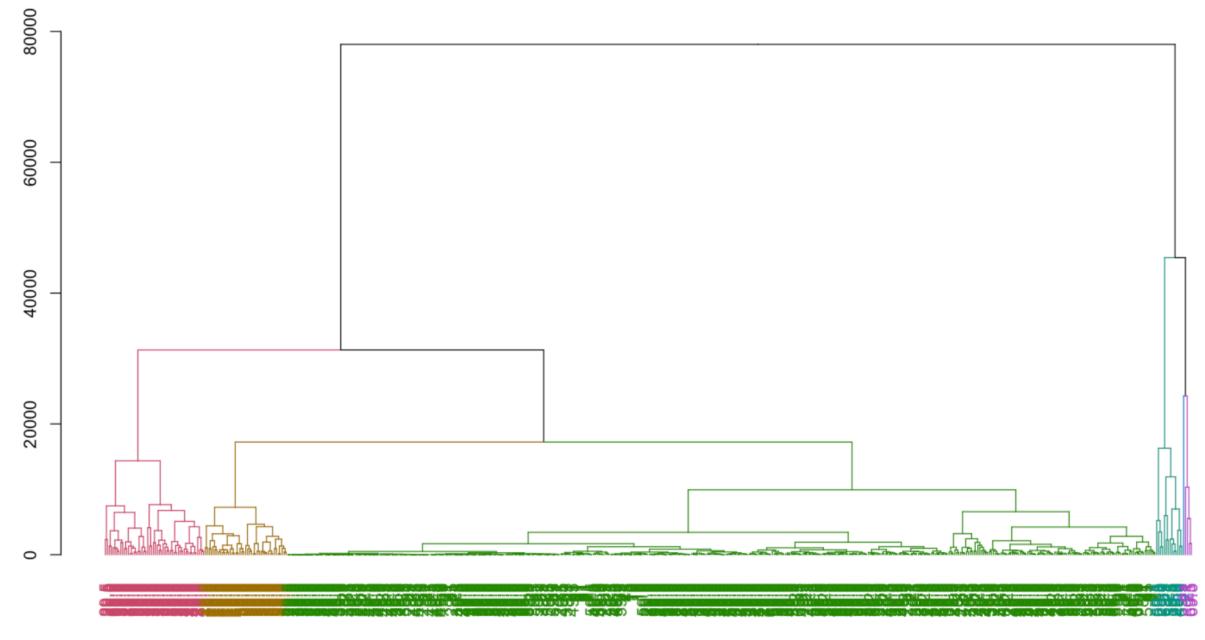


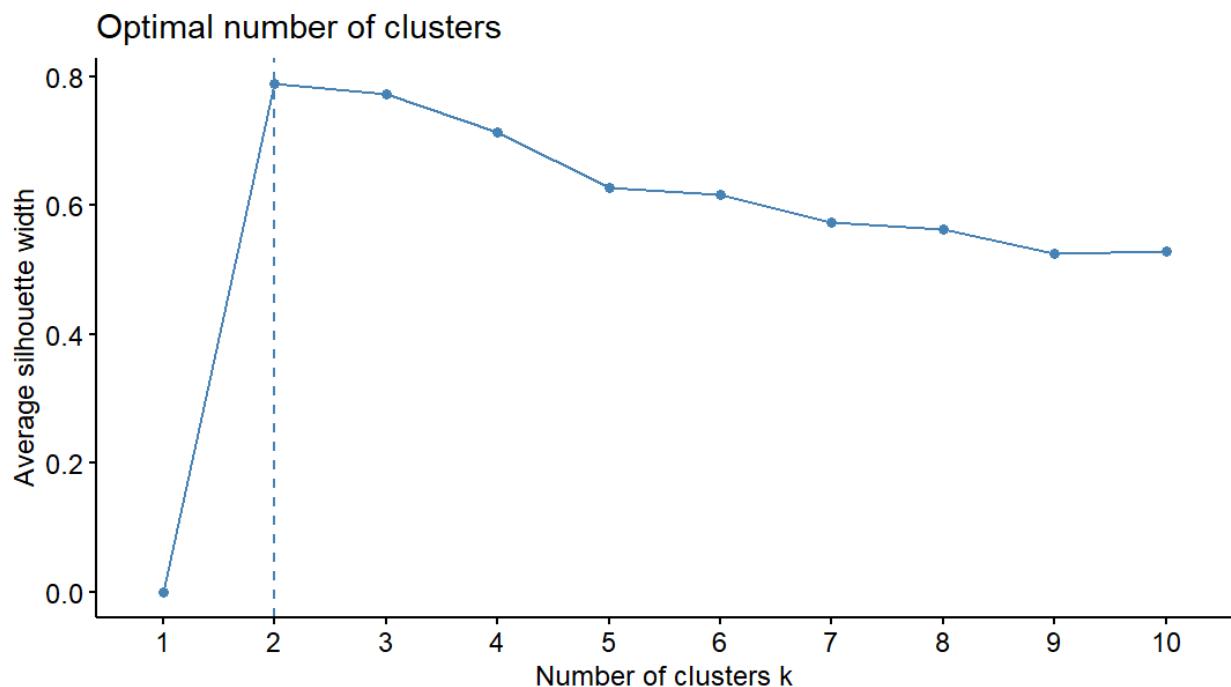
Figure 31

```
> print(hcluster_means_5)
  hier_cluster_5    reach impressions clicks_all      age   gender
1              1 21698.29    33664.214  1989.2143 33.07143 1.0000000
2              2 41824.00    65756.000  4002.0000 29.50000 1.0000000
3              3 12725.44    20065.692  1246.1538 36.80769 1.0000000
4              4 33567.25    46676.500  2746.7500 32.00000 1.0000000
5              5 1602.59     2690.636   173.9152 41.37778 0.4363636
```

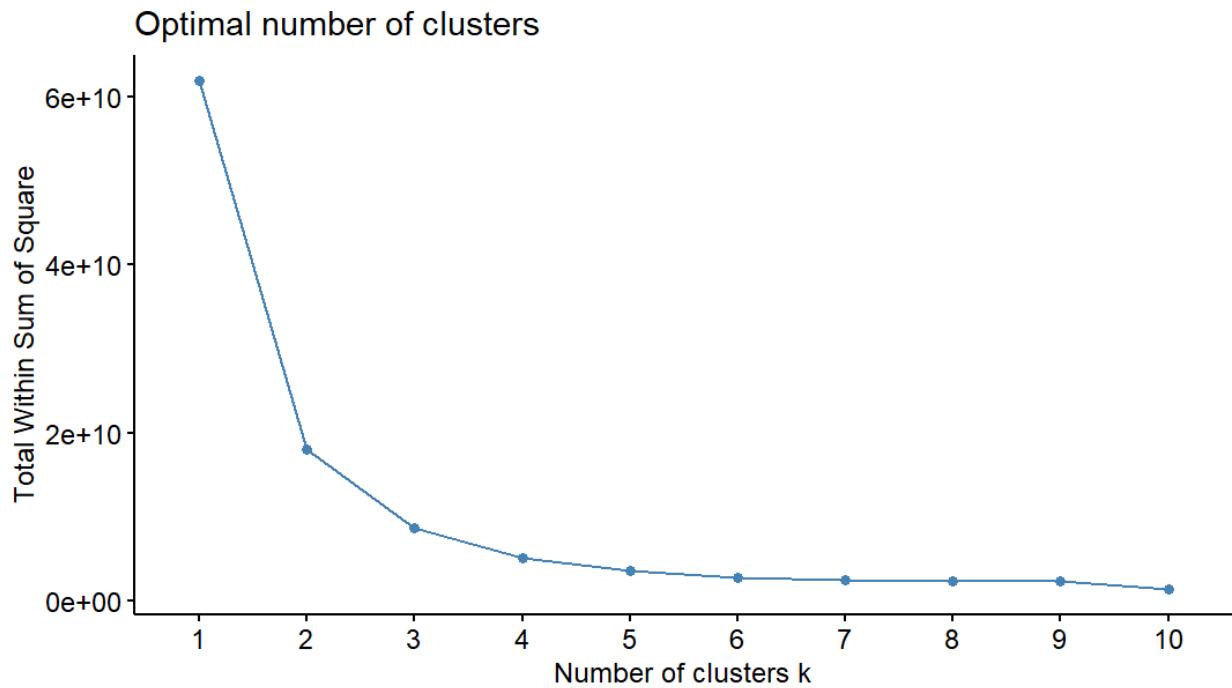
Figure 32

```
> print(hcluster_means)
  hier_cluster_6    reach impressions clicks_all      age   gender
1              1 21698.286   33664.214  1989.2143 33.07143 1.0000000
2              2 41824.000   65756.000  4002.0000 29.50000 1.0000000
3              3 12725.442   20065.692  1246.1538 36.80769 1.0000000
4              4 33567.250   46676.500  2746.7500 32.00000 1.0000000
5              5 6732.791    11640.023   705.5349 40.69767 1.0000000
6              6 1114.540    1839.257   123.3407 41.44248 0.3827434
```

Figure 33



**Figure 34**



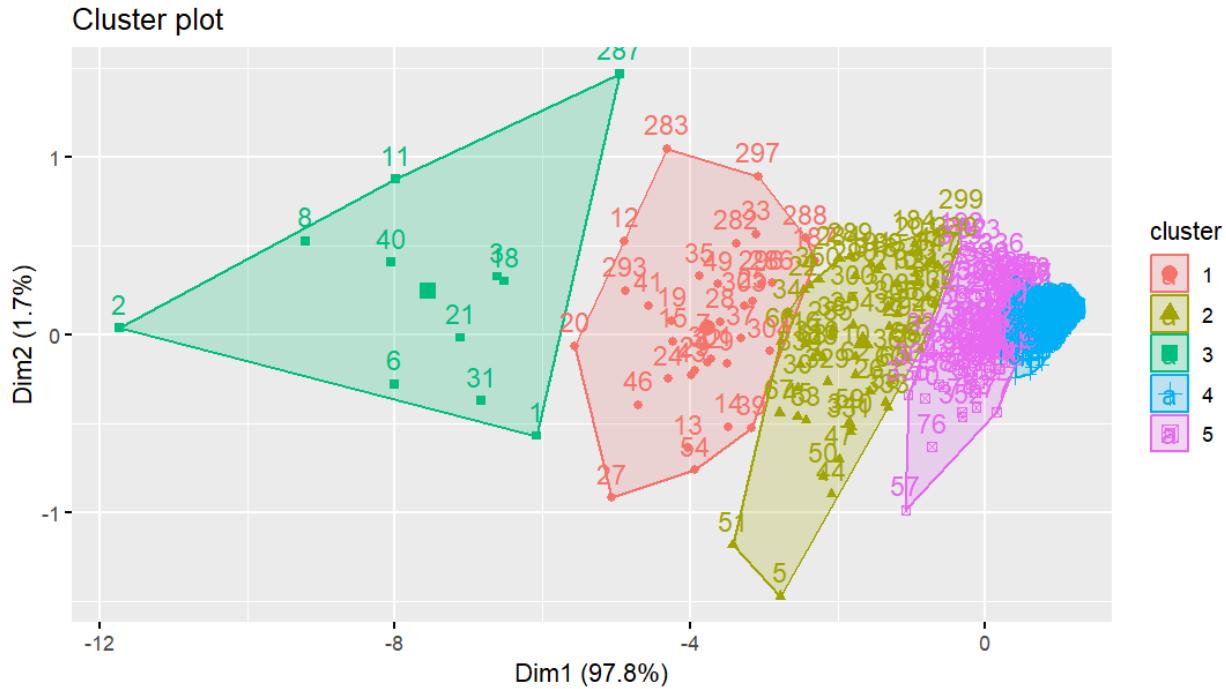
**Figure 35**

```
> print(kcluster_means)
  kmean_cluster      reach impressions clicks_all      age      gender
1              1 13739.077    21254.269   1320.8077 36.23077 1.0000000
2              2  5911.905    10343.365    636.0405 41.85811 0.9729730
3              3 27482.071    41686.929   2428.2143 32.35714 1.0000000
4              4   934.770     1513.601    103.2441 41.27230 0.3497653
```

**Figure 36**

```
> print(kcluster_means)
  kmean_cluster      reach impressions clicks_all      age      gender
1              1 15917.1176   24789.4412  1501.32353 34.79412 1.0000000
2              2  8499.3871   14141.3226   884.59677 39.20161 1.0000000
3              3 29578.0000   43874.8182  2586.27273 33.13636 1.0000000
4              4   578.3971     911.4058    64.46087 40.70435 0.2550725
5              5  2998.4649    5108.6053  325.85965 43.82018 0.8070175
```

**Figure 37**



#### 4. Pages that have more Traffic by Region using Google Analytics

Figure 38

Table showing the top 10 pages by region in order from most amount of views to the least amount of views within the last 12 months:

Pages and screens: Page title and screen class

Last 12 months Mar 18, 2023 - Mar 18, 2024

Search... Rows per page: 50 Go to: 1 1-50 of 9976

	Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event
			121,057 100% of total	47,605 100% of total	2.54 Avg 0%	1m 45s Avg 0%	100%
1	Sistema.bio : The Biogester solution.	Nairobi County	9,889	5,296	1.87	1m 27s	
2	Sistema.bio : The Biogester solution.	Maharashtra	9,106	5,418	1.68	1m 06s	
3	Opportunities : Sistema.bio	Nairobi County	5,296	2,588	2.05	1m 14s	
4	Opportunities : Sistema.bio	Maharashtra	2,086	1,077	1.94	59s	
5	Sistema.bio : The Biogester solution.	Karnataka	1,739	1,065	1.63	1m 09s	
6	Contact Us : Sistema.bio	Maharashtra	1,638	1,066	1.54	42s	
7	Sistema.bio : La inversión que funciona para agricultores y ganaderos	Mexico City	1,627	982	1.66	1m 14s	
8	Sistema.bio : The Biogester solution.	Uttar Pradesh	1,518	982	1.55	1m 18s	
9	Sistema.bio : The Biogester solution.	(not set)	1,482	954	1.55	1m 04s	
10	Who we are : Sistema.bio	Nairobi County	1,398	846	1.65	1m 45s	

**Figure 39**

**Table showing the top 10 pages by region in order from the most amount of users to the least amount of users within the last 12 months:**

Pages and screens: Page title and screen class

Last 12 months Mar 18, 2023 - Mar 18, 2024

	Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
			121,057 100% of total	47,605 100% of total	2.54 Avg 0%	1m 45s Avg 0%	495 100% of
1	Sistema.bio : The Biogester solution.	Maharashtra	9,106	5,418	1.68	1m 06s	36
2	Sistema.bio : The Biogester solution.	Nairobi County	9,889	5,296	1.87	1m 27s	40
3	Opportunities : Sistema.bio	Nairobi County	5,296	2,588	2.05	1m 14s	26
4	Opportunities : Sistema.bio	Maharashtra	2,086	1,077	1.94	59s	8
5	Contact Us : Sistema.bio	Maharashtra	1,638	1,066	1.54	42s	5
6	Sistema.bio : The Biogester solution.	Karnataka	1,739	1,065	1.63	1m 09s	7
7	Sistema.bio : La inversión que funciona para agricultores y ganaderos	Mexico City	1,627	982	1.66	1m 14s	6
8	Sistema.bio : The Biogester solution.	Uttar Pradesh	1,518	982	1.55	1m 18s	6
9	Sistema.bio : The Biogester solution.	(not set)	1,482	954	1.55	1m 04s	6
10	Who we are : Sistema.bio	Nairobi County	1,398	846	1.65	1m 45s	5

Figure 40

Table showing the top 10 pages by region in order from most amount of average views per user to the least amount of average views per user within the last 12 months:

Pages and screens: Page title and screen class

Last 12 months Mar 18, 2023 - Mar 18, 2024

	Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
			121,057 100% of total	47,605 100% of total	2.54 Avg 0%	1m 45s Avg 0%	10
1	Test Form 2024 : Sistema.bio	Guerrero	58	1	58.00	46m 31s	
2	You searched for Team : Sistema.bio	Maharashtra	44	1	44.00	7m 24s	
3	Sistema.bio team news from Q3 : Sistema.bio	Maharashtra	78	2	39.00	1m 07s	
4	You searched for Aniket shinde : Sistema.bio	Maharashtra	37	1	37.00	1m 27s	
5	Thank you : Sistema.bio	Guerrero	58	2	29.00	1m 41s	
6	2024 - Noticias Sistema.bio : Sistema.bio	Valle del Cauca	43	2	21.50	3m 39s	
7	Blog : Sistema.bio	Flanders	56	3	18.67	2m 50s	
8	You searched for Gold : Sistema.bio	Karnataka	17	1	17.00	5m 30s	
9	Sistema.bio : La inversión que funciona para agricultores y ganaderos	Lampung	15	1	15.00	17m 32s	
10	Sistema.bio : The Biogester solution.	Quindio	29	2	14.50	41s	

Figure 41

Table showing the top 10 pages by region in order from highest average engagement time to the lowest average engagement time within the last 12 months:

Pages and screens: Page title and screen class Last 12 months Mar 18, 2023 – Mar 18, 2024

Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
		121,057 100% of total	47,605 100% of total	2.54 Avg 0%	1m 45s Avg 0%	All
1 Women and biodigestion: the sustainable farming solution : Sistema.bio	Uttar Pradesh	3	1	3.00	56m 09s	
2 Test Form 2024 : Sistema.bio	Guerrero	58	1	58.00	46m 31s	
3 ¿Cuál es el impacto real de los biodigestores en el medio ambiente? : Sistema.bio	Cortes Department	2	1	2.00	26m 54s	
4 Kontaktieren Sie uns : Sistema.bio	Hamburg	0	1	0.00	24m 55s	
5 ¿Cuál es el impacto real de los biodigestores en el medio ambiente? : Sistema.bio	Intibucá Department	1	1	1.00	24m 17s	
6 Sistema.bio : Die Biofermenter-Lösung.	Hamburg	0	1	0.00	24m 00s	
7 Contact Us : Sistema.bio	San Jose Province	6	1	6.00	20m 53s	
8 ¿Cuál es el impacto real de los biodigestores en el medio ambiente? : Sistema.bio	Puno	1	1	1.00	19m 28s	
9 Biogás y los ODS: Informe de Shell Foundation : Sistema.bio	Andalusia	3	1	3.00	19m 21s	
10 2024 - Noticias Sistema.bio : Sistema.bio	Quindío	1	1	1.00	18m 57s	

**Figure 42**

**Table showing the top 10 pages by region in order from highest event count to the lowest event count within the last 12 months:**

Pages and screens: Page title and screen class

Last 12 months Mar 18, 2023 - Mar 18, 2024

Rows per page: 50 Go to: 1 < 1-50 of 9976 >

Page title and screen class	Region	Average engagement time	Event count All events	Conversions All events	Total revenue
		1m 45s Avg 0%	495,213 100% of total	0.00	\$0.00
1 Sistema.bio : The Biogester solution.	Nairobi County	1m 27s	40,788	0.00	\$0.00
2 Sistema.bio : The Biogester solution.	Maharashtra	1m 06s	36,975	0.00	\$0.00
3 Opportunities : Sistema.bio	Nairobi County	1m 14s	26,165	0.00	\$0.00
4 Opportunities : Sistema.bio	Maharashtra	59s	8,955	0.00	\$0.00
5 Sistema.bio : The Biogester solution.	Karnataka	1m 09s	7,273	0.00	\$0.00
6 Sistema.bio : The Biogester solution.	Uttar Pradesh	1m 18s	6,956	0.00	\$0.00
7 Sistema.bio : La inversión que funciona para agricultores y ganaderos	Mexico City	1m 14s	6,567	0.00	\$0.00
8 Sistema.bio : The Biogester solution.	(not set)	1m 04s	6,014	0.00	\$0.00
9 Contact Us : Sistema.bio	Maharashtra	42s	5,968	0.00	\$0.00
10 Who we are : Sistema.bio	Nairobi County	1m 45s	5,827	0.00	\$0.00

**Figure 43**

**Table showing the top 10 pages by region in order from most amount of views to the least amount of views within the last 90 days:**

Pages and screens: Page title and screen class

Last 90 days Dec 20, 2023 - Mar 18, 2024 |

Search... Rows per page: 50 Go to: 1 < 1-50 of 5637 >

	Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
			53,171 100% of total	20,927 100% of total	2.54 Avg 0%	1m 49s Avg 0%	242 100% of
1	Sistema.bio : The Biogester solution.	Nairobi County	4,300	2,394	1.80	1m 27s	19
2	Sistema.bio : The Biogester solution.	Maharashtra	3,726	2,139	1.74	1m 13s	16
3	Opportunities : Sistema.bio	Nairobi County	2,217	1,100	2.02	1m 18s	11
4	Sistema.bio : The Biogester solution.	Karnataka	785	477	1.65	1m 08s	3
5	Sistema.bio : The Biogester solution.	Uttar Pradesh	774	514	1.51	1m 28s	3
6	Sistema.bio : La inversión que funciona para agricultores y ganaderos	Mexico City	762	440	1.73	1m 15s	3
7	Sistema.bio : The Biogester solution.	(not set)	705	425	1.66	1m 12s	3
8	Opportunities : Sistema.bio	Maharashtra	674	359	1.88	1m 05s	3
9	Contact Us : Sistema.bio	Maharashtra	660	447	1.48	50s	2
10	Sistema.bio : The Biogester solution.	Gujarat	637	383	1.66	1m 31s	3

**Figure 44**

**Table showing the top 10 pages by region in order from most amount of users to the least amount of users within the last 90 days:**

Pages and screens: Page title and screen class

Last 90 days Dec 20, 2023 – Mar 18, 2024

	Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
			53,171 100% of total	20,927 100% of total	2.54 Avg 0%	1m 49s Avg 0%	242 100% of
1	Sistema.bio : The Biogester solution.	Nairobi County	4,300	2,394	1.80	1m 27s	19
2	Sistema.bio : The Biogester solution.	Maharashtra	3,726	2,139	1.74	1m 13s	16
3	Opportunities : Sistema.bio	Nairobi County	2,217	1,100	2.02	1m 18s	11
4	Sistema.bio : The Biogester solution.	Uttar Pradesh	774	514	1.51	1m 28s	3
5	Sistema.bio : The Biogester solution.	Karnataka	785	477	1.65	1m 08s	3
6	Contact Us : Sistema.bio	Maharashtra	660	447	1.48	50s	2
7	Sistema.bio : La inversión que funciona para agricultores y ganaderos	Mexico City	762	440	1.73	1m 15s	3
8	Sistema.bio : The Biogester solution.	(not set)	705	425	1.66	1m 12s	3
9	Sistema.bio : The Biogester solution.	Masovian Voivodeship	407	407	1.00	0s	1
10	Who we are : Sistema.bio	Nairobi County	636	404	1.57	1m 35s	3

**Figure 45**

**Table showing the top 10 pages by region in order from most amount of average views per user to the least amount of average views per user within the last 90 days:**

Pages and screens: Page title and screen class Last 90 days Dec 20, 2023 - Mar 18, 2024

Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
		53,171 100% of total	20,927 100% of total	2.54 Avg 0%	1m 49s Avg 0%	242 100% of
1 Test Form 2024 : Sistema.bio	Guerrero	58	1	58.00	46m 31s	
2 Blog : Sistema.bio	Flanders	53	1	53.00	7m 02s	
3 Thank you : Sistema.bio	Guerrero	58	2	29.00	1m 41s	
4 Sistema.bio : The Biogester solution.	Souss-Massa	22	1	22.00	7m 06s	
5 2024 - Noticias Sistema.bio : Sistema.bio	Valle del Cauca	43	2	21.50	3m 39s	
Sistema.bio : La inversión que funciona para agricultores y ganaderos	Connecticut	20	1	20.00	11m 51s	
7 Blog : Sistema.bio	Utah	18	1	18.00	5m 56s	
8 Contact Us : Sistema.bio	Guangdong Province	17	1	17.00	1m 21s	
9 Sistema.bio : La inversión que funciona para agricultores y ganaderos	Lampung	15	1	15.00	17m 32s	
10 Downloads & Resources : Sistema.bio	Aichi	11	1	11.00	0s	

**Figure 46**

**Table showing the top 10 pages by region in order from highest average engagement time to the lowest average engagement time within the last 90 days:**

Pages and screens: Page title and screen class

Last 90 days Dec 20, 2023 - Mar 18, 2024

Search... Rows per page: 50 Go to: 1 < 1-50 of 5637 >

	Page title and screen class	Region	Views	Users	Views per user	Average engagement time	Event count
			53,171 100% of total	20,927 100% of total	2.54 Avg 0%	1m 49s Avg 0%	242 100% of
1	Test Form 2024 : Sistema.bio	Guerrero	58	1	58.00	46m 31s	
2	Casos de éxito : Sistema.bio	Guerrero	1	1	1.00	25m 10s	
3	Kontaktieren Sie uns : Sistema.bio	Hamburg	0	1	0.00	24m 55s	
4	Sistema.bio : Die Biofermenter-Lösung.	Hamburg	0	1	0.00	24m 00s	
5	Women and biodigestion: the sustainable farming solution : Sistema.bio	Uttar Pradesh	1	1	1.00	23m 09s	
6	Sistema.bio : The Biodigester solution.	Demerara-Mahaica	3	1	3.00	21m 51s	
7	Contact Us : Sistema.bio	San Jose Province	6	1	6.00	20m 53s	
8	Biogás y los ODS: Informe de Shell Foundation : Sistema.bio	Andalusia	3	1	3.00	19m 21s	
9	2024 - Noticias Sistema.bio : Sistema.bio	Quindío	1	1	1.00	18m 57s	
10	10 Años 10 Voces : Sistema.bio	Gujarat	8	1	8.00	18m 23s	

**Figure 47**

**Table showing the top 10 pages by region in order from highest event count to the lowest event count within the last 90 days:**

A Pages and screens: Page title and screen class Last 90 days Dec 20, 2023 - Mar 18, 2024

	Page title and screen class	Region	Average engagement time	Event count	Conversions	Total revenue
			1m 49s Avg 0%	242,529 100% of total	0.00	\$0.00
1	Sistema.bio : The Biogester solution.	Nairobi County	1m 27s	19,877	0.00	\$0.00
2	Sistema.bio : The Biogester solution.	Maharashtra	1m 13s	16,858	0.00	\$0.00
3	Opportunities : Sistema.bio	Nairobi County	1m 18s	11,895	0.00	\$0.00
4	Sistema.bio : The Biogester solution.	Uttar Pradesh	1m 28s	3,878	0.00	\$0.00
5	Opportunities : Sistema.bio	Maharashtra	1m 05s	3,721	0.00	\$0.00
6	Sistema.bio : The Biogester solution.	Karnataka	1m 08s	3,540	0.00	\$0.00
7	Sistema.bio : La inversión que funciona para agricultores y ganaderos	Mexico City	1m 15s	3,355	0.00	\$0.00
8	Sistema.bio : The Biogester solution.	(not set)	1m 12s	3,071	0.00	\$0.00
9	Sistema.bio : The Biogester solution.	Gujarat	1m 31s	3,051	0.00	\$0.00
10	Who we are : Sistema.bio	Nairobi County	1m 35s	3,041	0.00	\$0.00