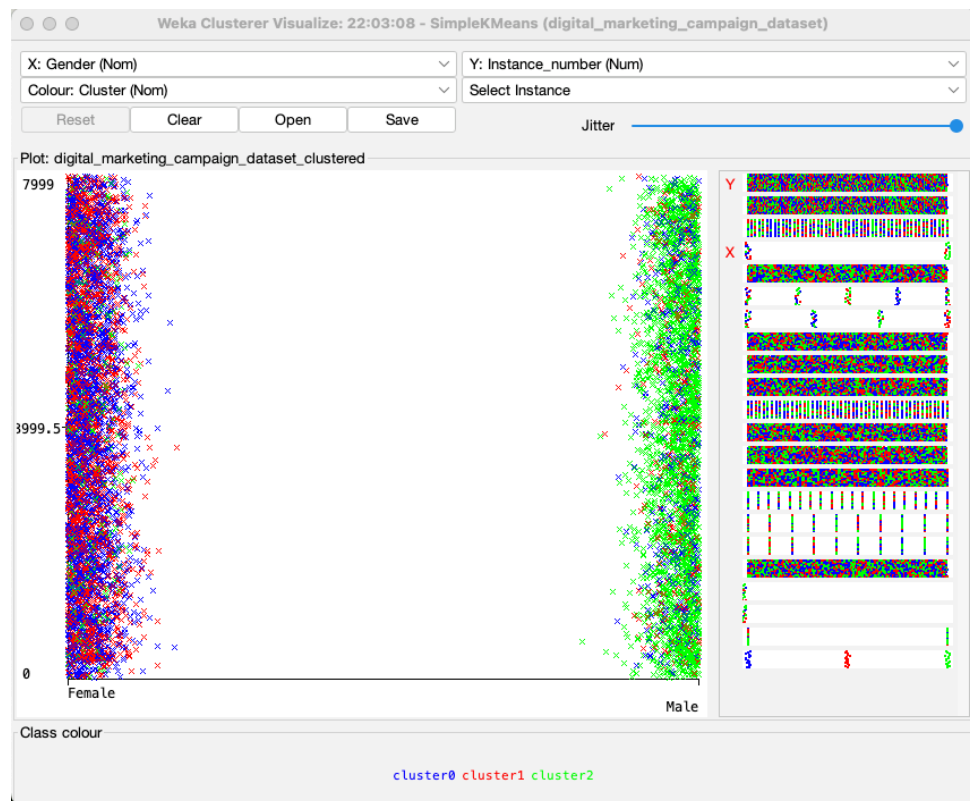


Lab5 by Isaac Braun, Joshy Kasahara

K-means - Weka

Cluster count: 3

Genders

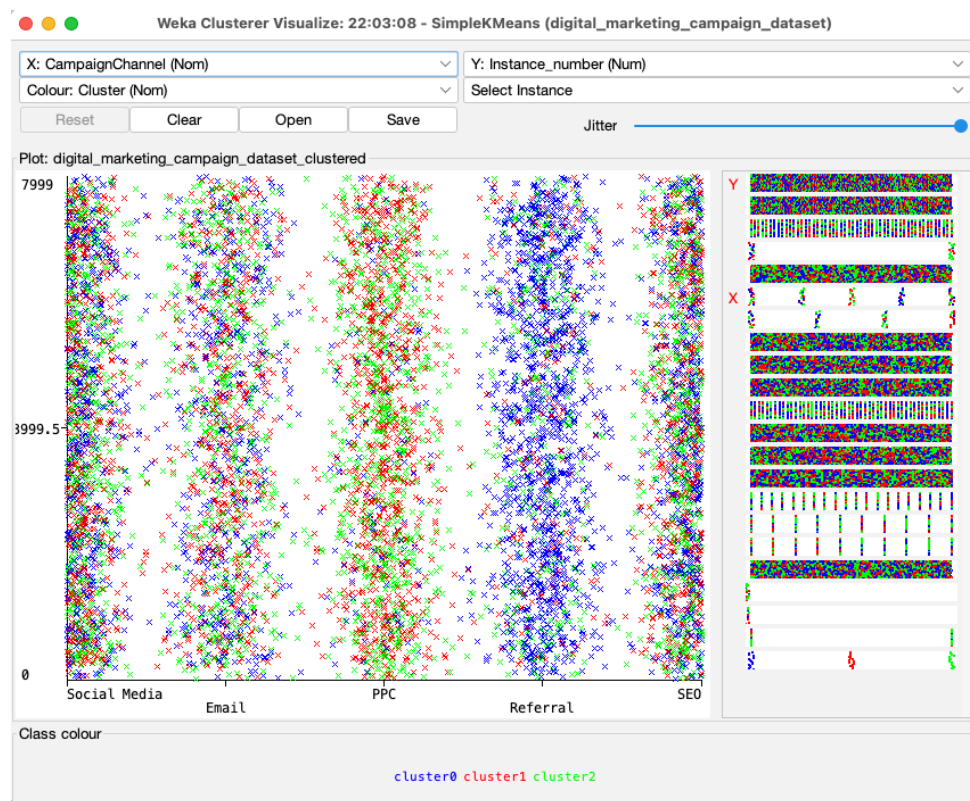


First of all, it is notable that the clusters are divided by gender. While the majority of users in clusters 0 and 1 are female, cluster 2 mostly consists of male users.

Cluster Characteristics Summary

- Cluster 0 - Majority female
- Cluster 1 - Majority female
- Cluster 2 - Majority male

Campaign Channel



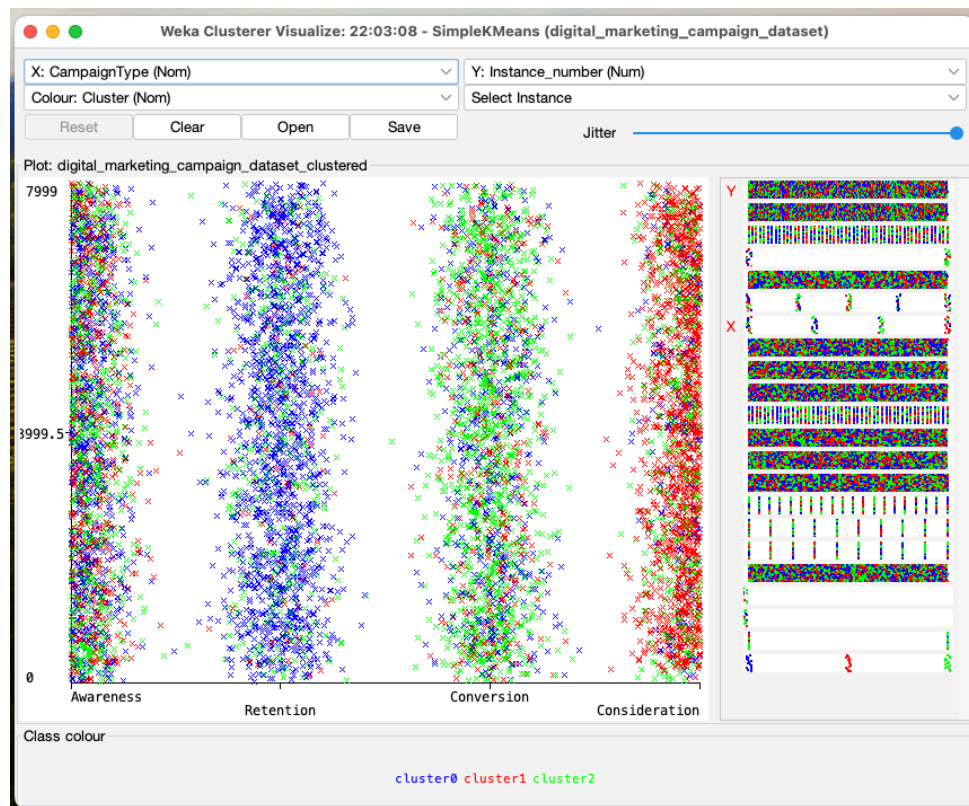
Campaign channels visualize how users discovered the campaign—Social Media, Email, PPC (Pay-Per-Click online ads), Referral (link posted in another online medium), and SEO.

Cluster 0, which consists mostly of female users, heavily relies on Referrals, unlike the other clusters. Clusters 1 (females) and 2 (males) are more evenly distributed and tend to rely on PPC, whereas Cluster 0 does not.

Cluster Characteristics Summary

- **Cluster 0** - Majority female - Referral lovers / ads haters
- **Cluster 1** - Majority female - Ad clickers (PPC) / No referral
- **Cluster 2** - Majority male - Ad clickers (PPC) / No referral

Campaign Types:



Campaign types are the common web campaign categorization which can be described as following:

- **Awareness:** Campaigns aimed at increasing brand recognition and visibility among potential customers. These campaigns focus on reaching a broad audience to create awareness about the brand or product.
- **Retention:** Campaigns designed to keep existing customers engaged and loyal to the brand. These often include personalized offers, loyalty programs, and regular communication to maintain customer interest.
- **Conversion:** Campaigns focused on encouraging potential customers to take a specific action, such as making a purchase or signing up for a service. These campaigns are typically targeted at users who are already aware of the brand and are in the decision-making stage.
- **Consideration:** Campaigns that aim to move potential customers from awareness to the consideration stage. These campaigns

provide more detailed information about the product or service, often highlighting benefits and features to persuade customers to consider making a purchase.

Analysis

Cluster 0 (majority females) consists of retention lovers who are keen on the brand.

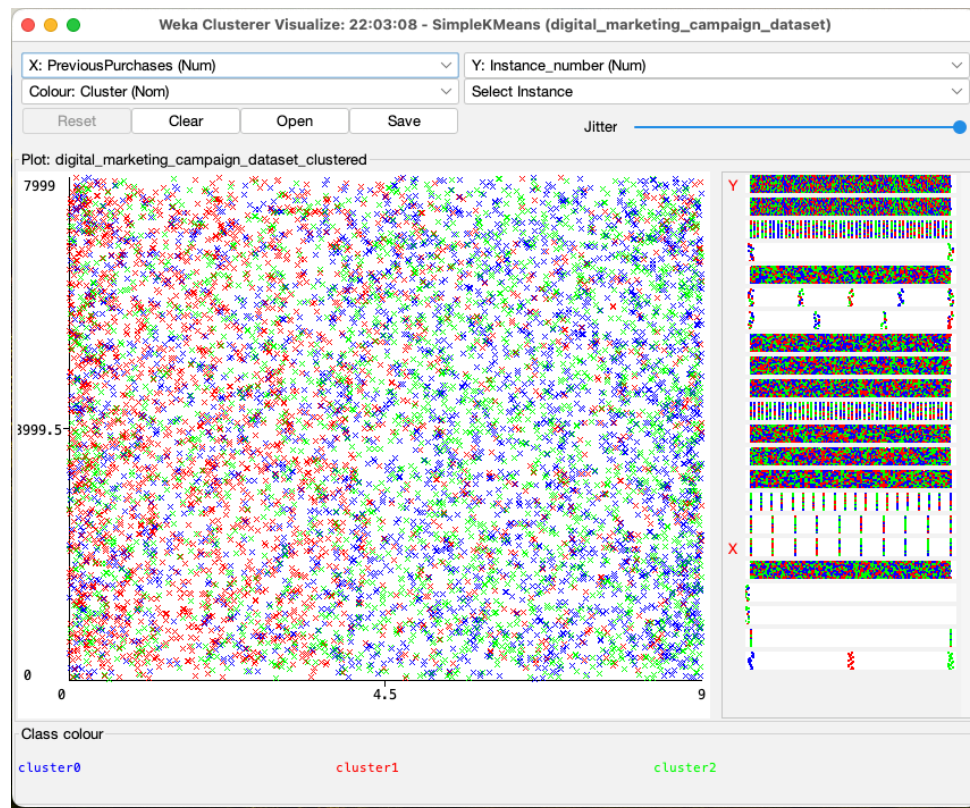
Cluster 1 (majority females) consists of skeptical users who respond to consideration campaign types.

Cluster 2 (majority males) may not be deep thinkers but are likely to respond to call-to-actions. They do not spend much time thinking or have specific loyalty to the brand. If the opportunity and the campaign call match, they are likely to convert spontaneously.

Cluster Characteristics Summary

- **Cluster 0** - Majority female - Referral lovers - **Retention lovers (Loyal customers)**
- **Cluster 1** - Majority female - Ad clicker (PPC) / No referral - **Consideration (careful buyers)**
- **Cluster 2** - Majority male - Ad clicker (PPC) / No referral - **Conversion Lover (Answering call-for-actions)**

Previous Purchasers

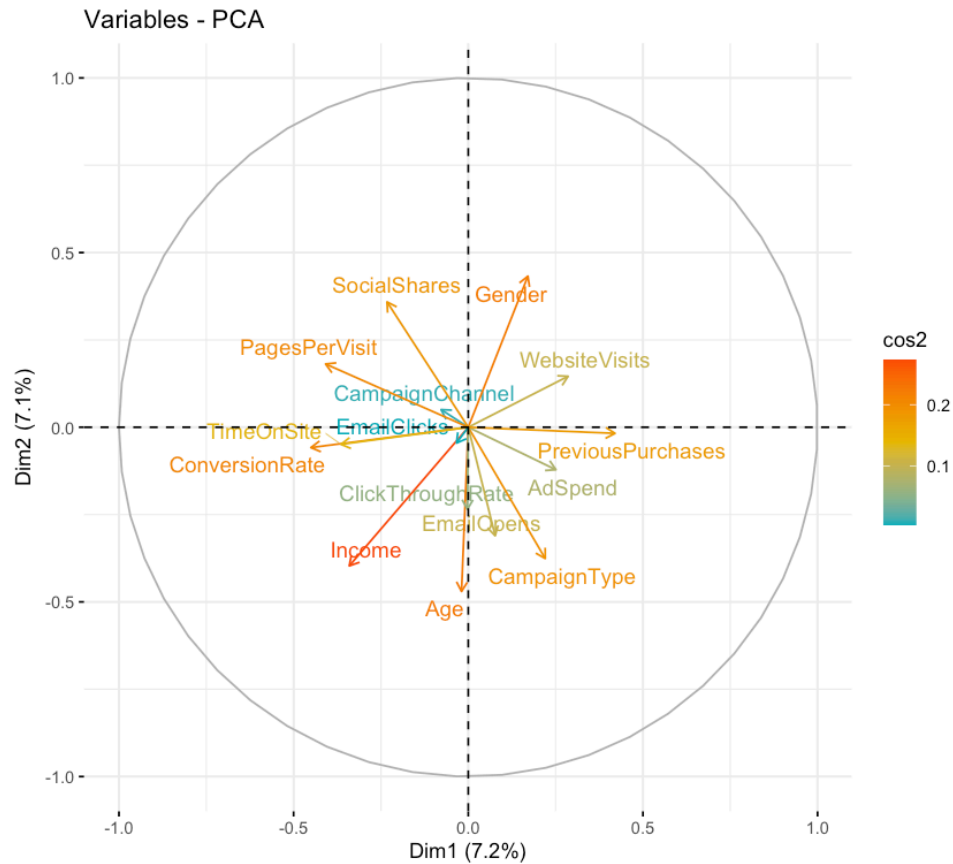


Cluster 0 (majority females) consists of skeptical user groups, as described previously in Campaign Types. The results from this section show they have less previous purchase history compared to clusters 1 and 2. In the long term, they may eventually become part of cluster 1, which consists of users with more previous purchases and greater brand loyalty. Clusters 1 and 2 have a more evenly distributed and greater purchase history than cluster 0.

Cluster Characteristics Summary

- **Cluster 0** - Majority female - Referral lovers - Retention lovers (Loyal customers) - **Less previous purchases**
- **Cluster 1** - Majority female - Ad clicker (PPC) / No referral - Consideration (careful buyers) - **More previous purchases**
- **Cluster 2** - Majority male - Ad clicker (PPC) / No referral - Conversion Lover (Answering call-for-actions) - **More previous purchases (more than cluster 1)**

PCA



1. The critical metric of a customer's conversion rate (ConversionRate) shows a direct positive correlation with the amount of time customers spend on the website (TimeOnSite). Knowing the website is working well is fantastic. Still, it highlights an area that can be capitalized. Suppose more effort is put into making the website visually appealing, reducing the number of steps it takes to purchase a product, and other methods to increase the time spent on the website. In that case, there is a high probability of increased sales.
2. It can be observed that a customer's age (Age) and the number of emails they open (EmailOpens) have a positive correlation. The company could use this knowledge to tune its emails to appeal to older audiences better and provide a clear call to action. While customers are opening emails, it can be seen from the low contribution of EmailClicks that they are not taking the call to action.

The goal would be to get older customers to the website, where they have a higher statistical chance of making a conversion.

3. With an almost direct negative correlation between customers who have purchased products before (PreviousPurchases) and the conversion rate (ConversionRate), we see a market slice that may not be served well. Since more time on the website typically creates a conversion, as seen previously, the company could tune specific emails and ads for previous customers calling them to the website.
4. While the CampaignChannel and EmailClicks variables have extremely low representation, and one might think to remove them, it can be argued that this displays areas where more effort should be placed. The company could research how to utilize the unique capabilities of the different channels and improve its email appeals to help achieve more clicks.