**Social Media Usage of the Class**

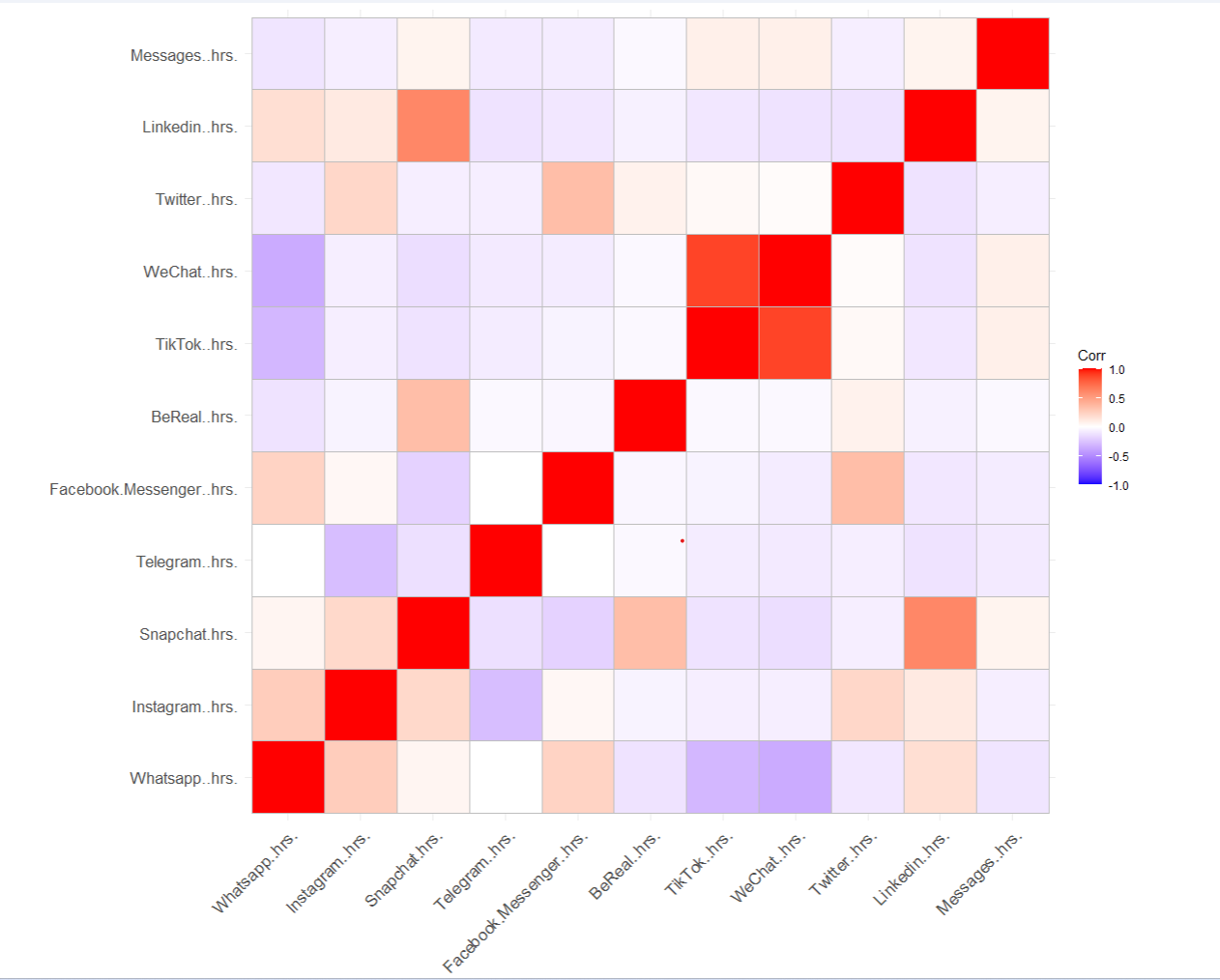
The dataset includes information on the amount of time (in hours) that individuals spend on different social media platforms weekly.

1. the relationship between the different social media apps in terms of usage time?

2. identifying clusters of social media that are similar?

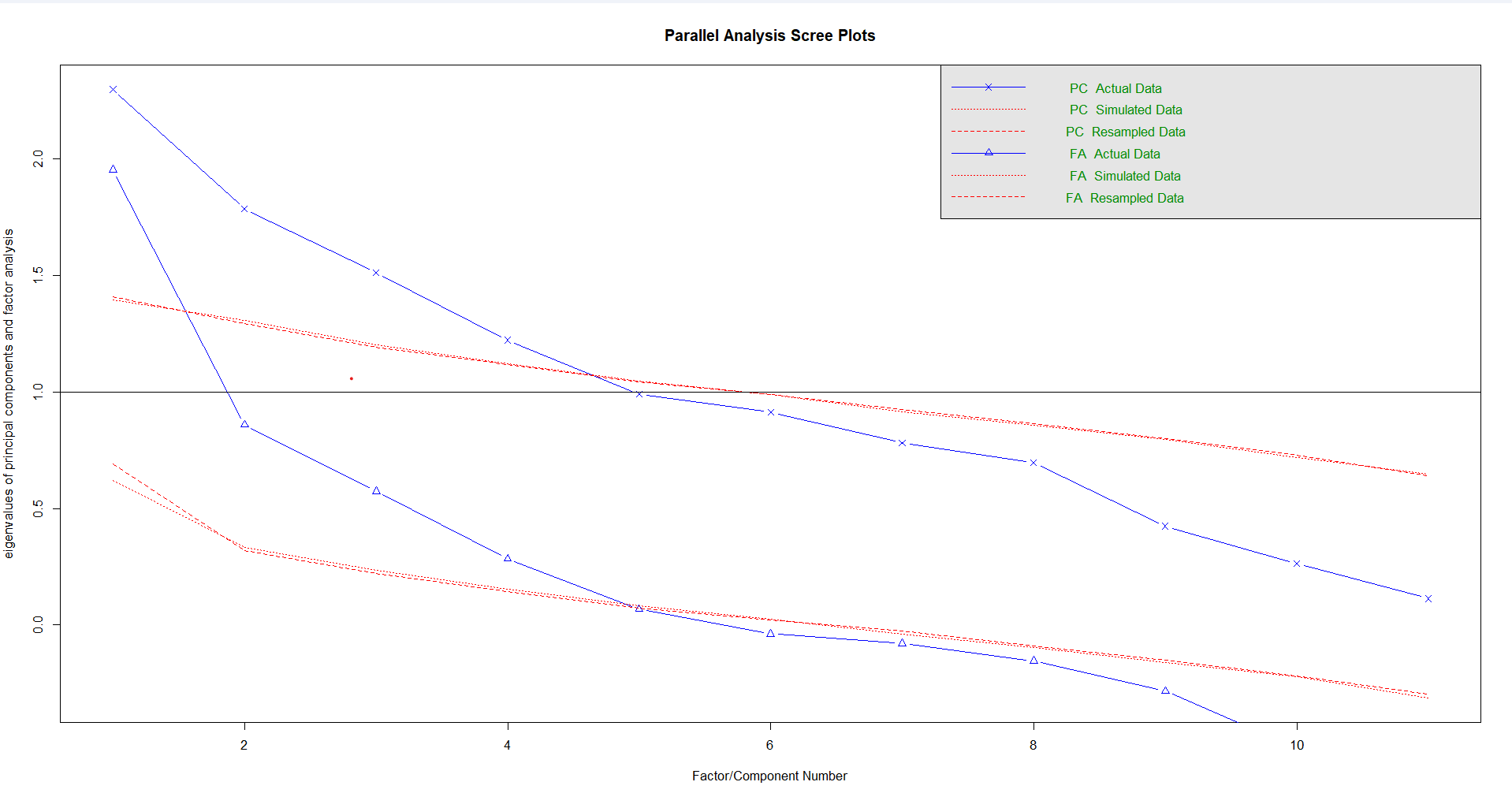
3. Are there any underlying factors that drive social media usage?

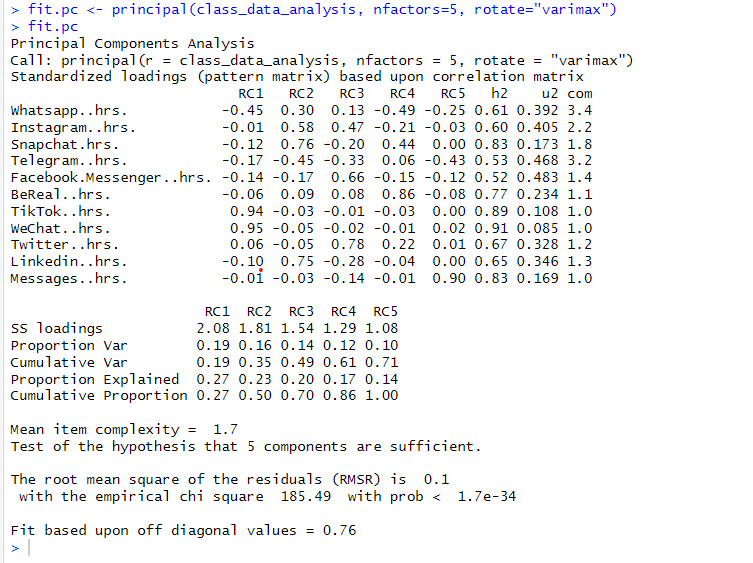
**Correlation Matrix**

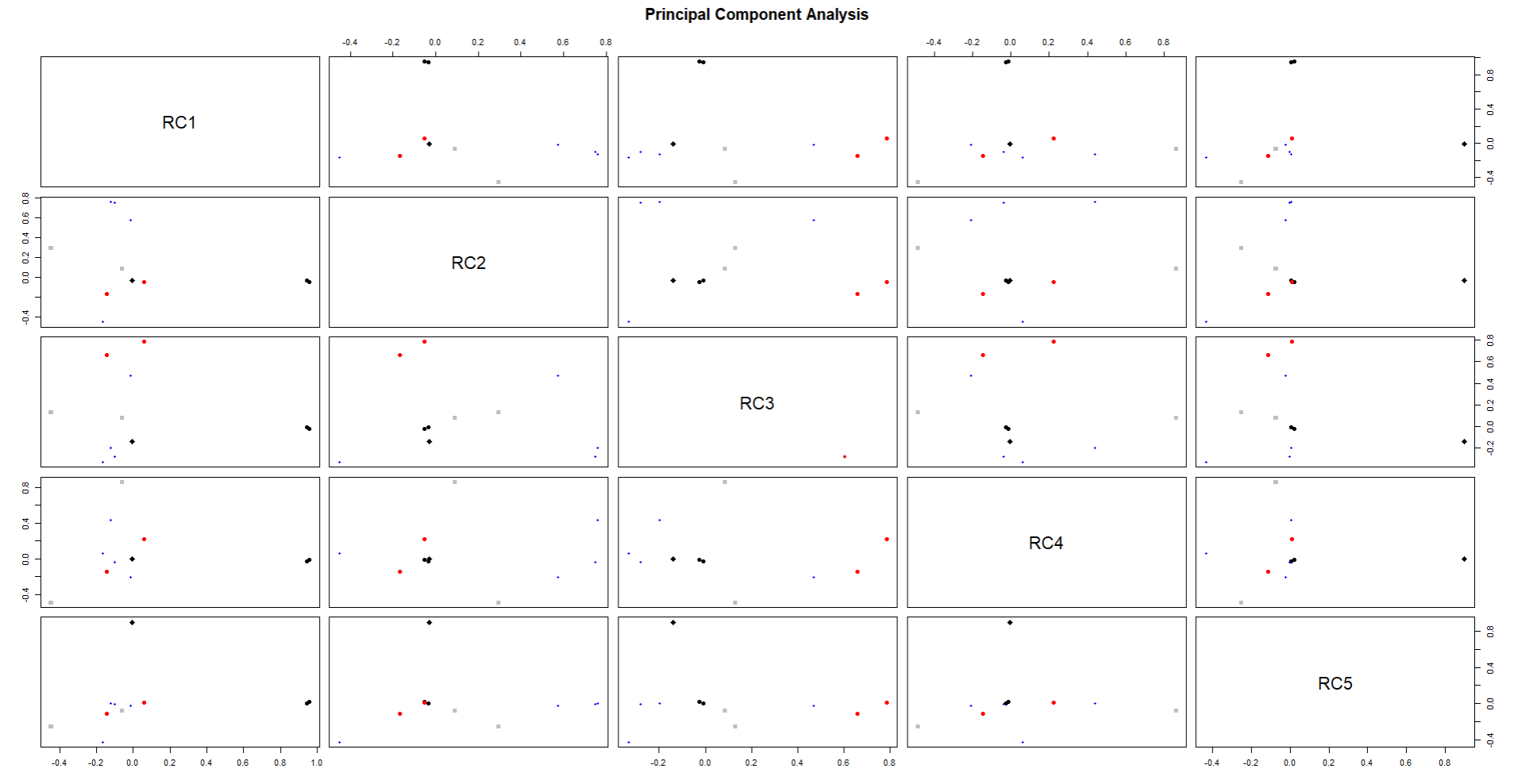


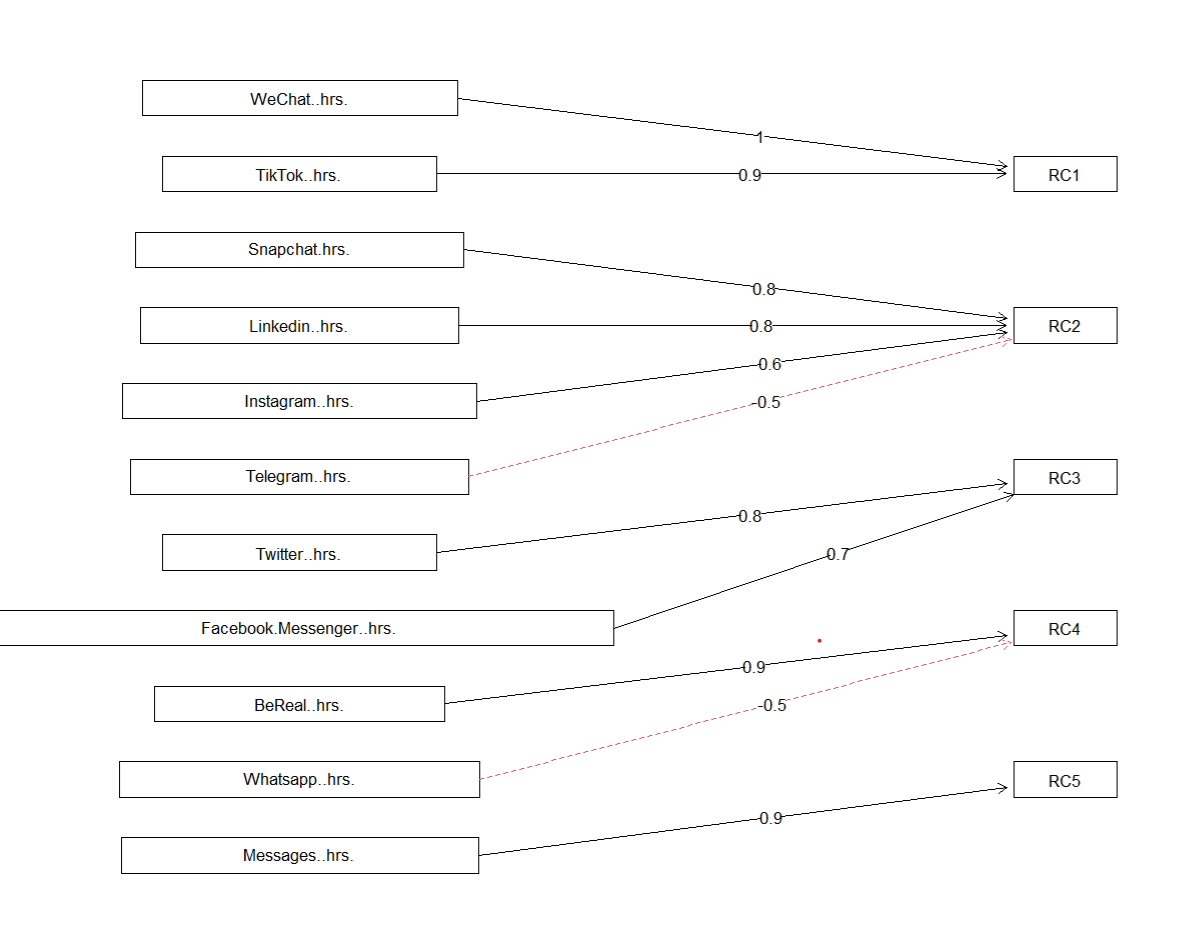
WeChat and Tiktok usage time(in hrs) are highly correlated which suggests that these variables may be measuring similar constructs or concepts.

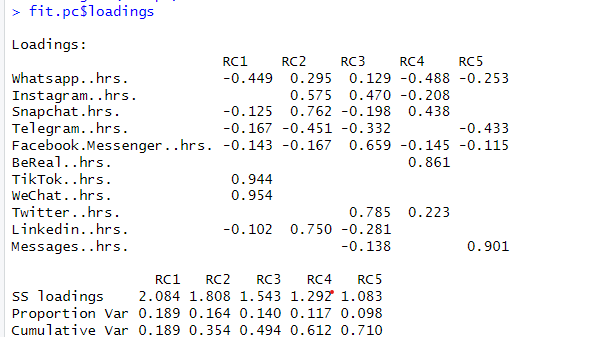
**Exploratory Factor Analysis**











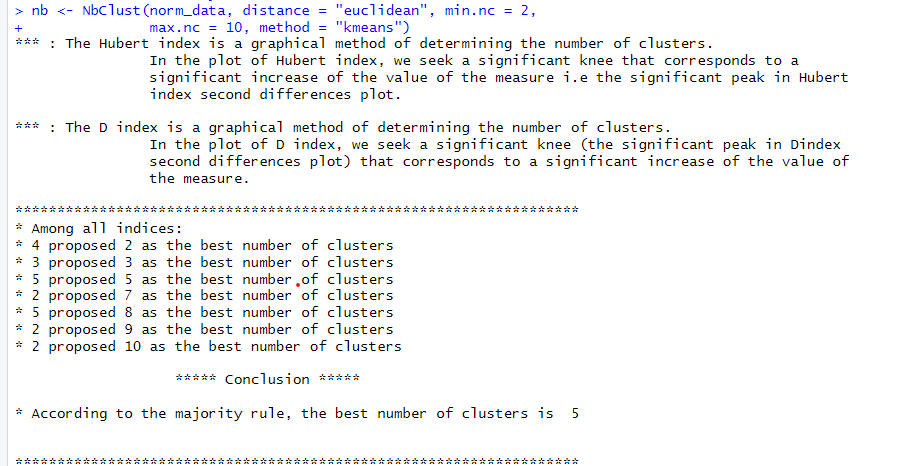
Factor 1 includes WeChat and Tiktok these two platforms share similar characteristics or are used for similar purposes. This could suggest that both WeChat and TikTok are social media platforms that are primarily used for communication and socialization purposes.

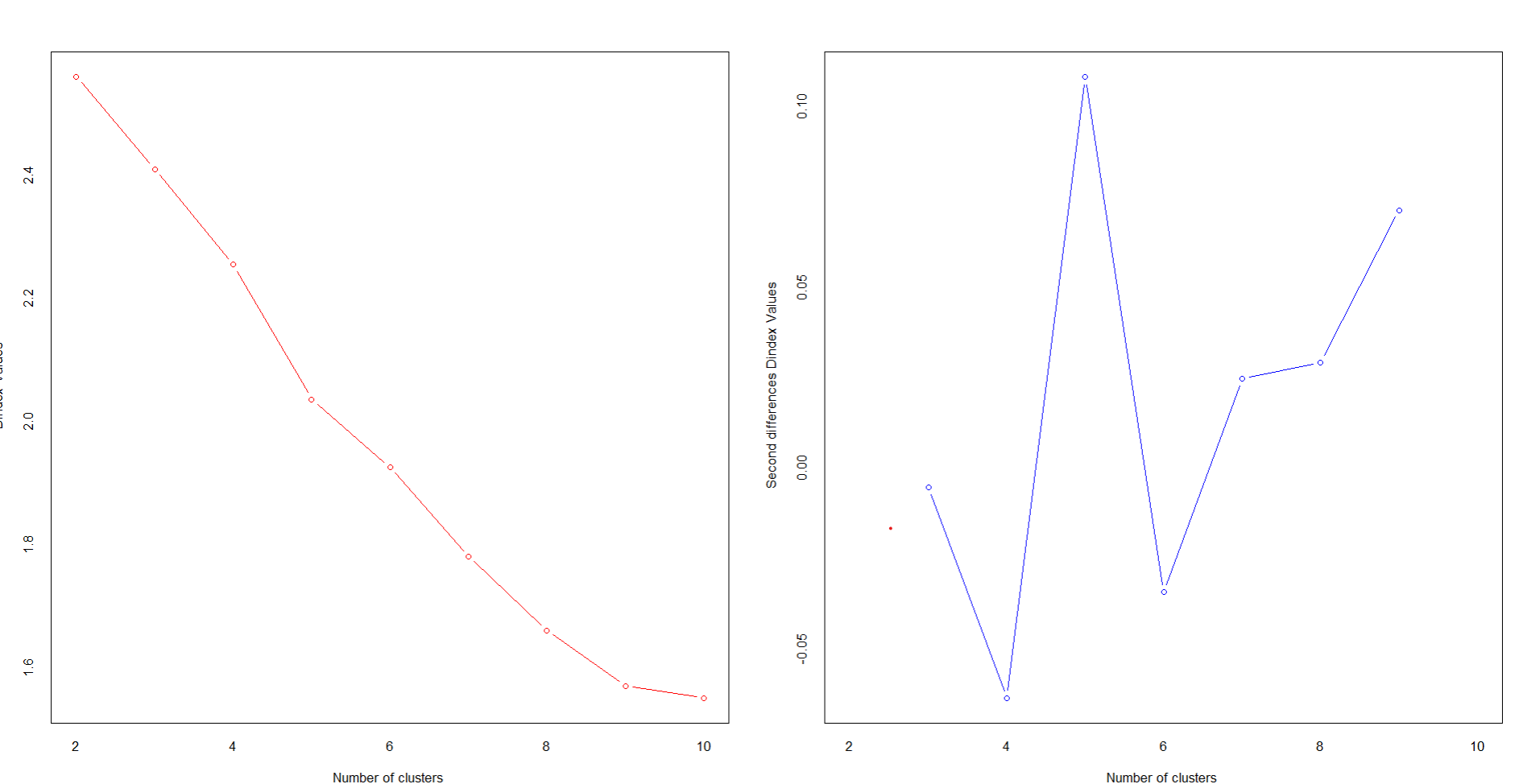
Factor 2 includes Snapchat, Linkedin, Instagram reflects a social networking factor, where these platforms are used primarily for building and maintaining social connections, sharing personal updates and experiences, and staying informed about others' lives.

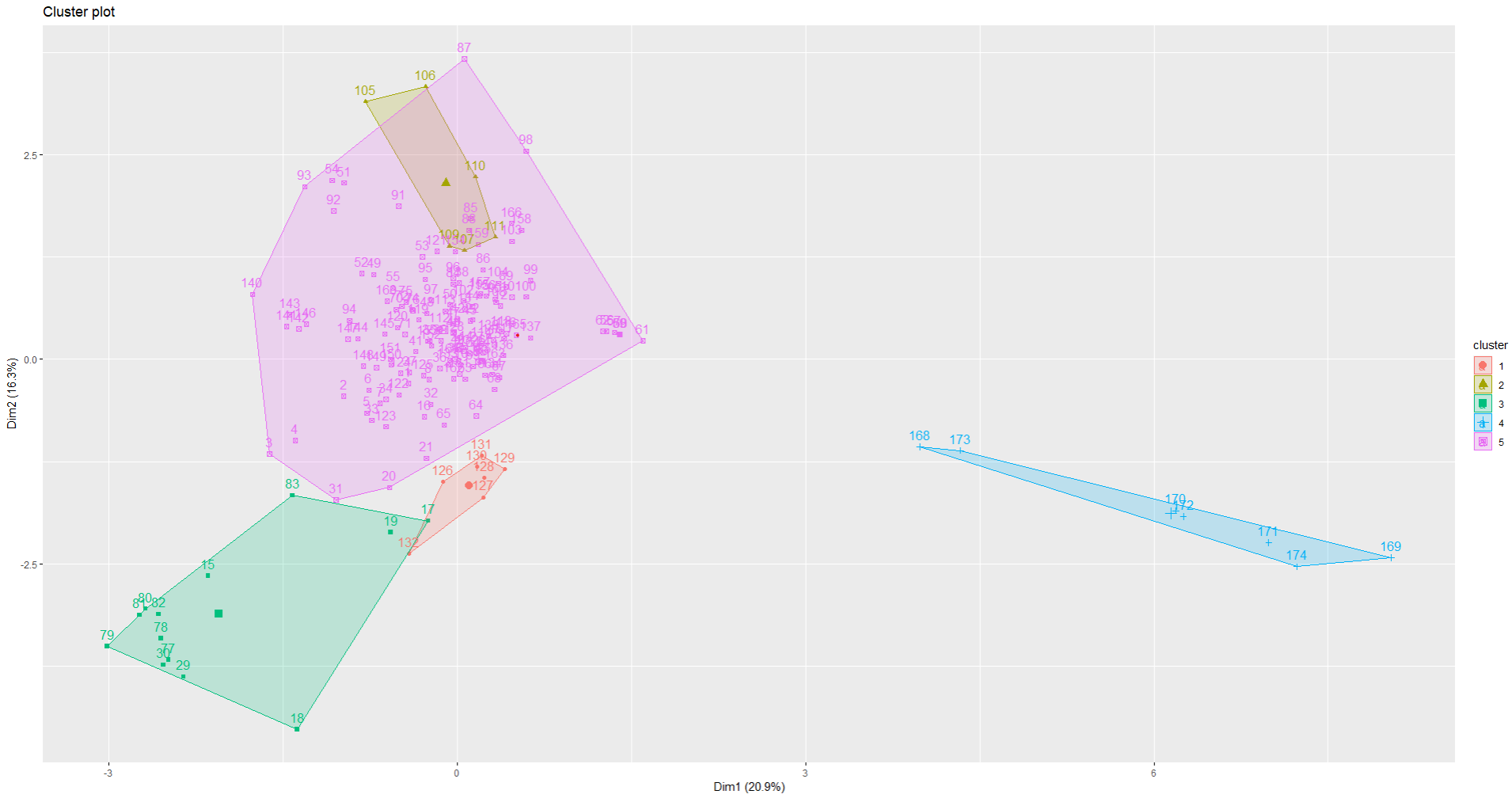
Factor 4 includes BeReal usage time, which is a social media platform

Factor 3 includes Twitter and Facebook, represent the public and real-time nature of information sharing on these platforms, where users can share their thoughts and opinions with a wider audience.

**Cluster Analysis**



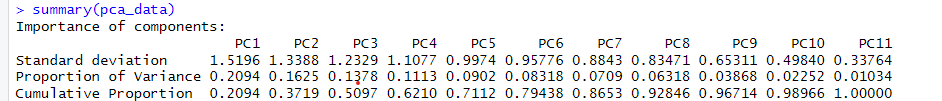


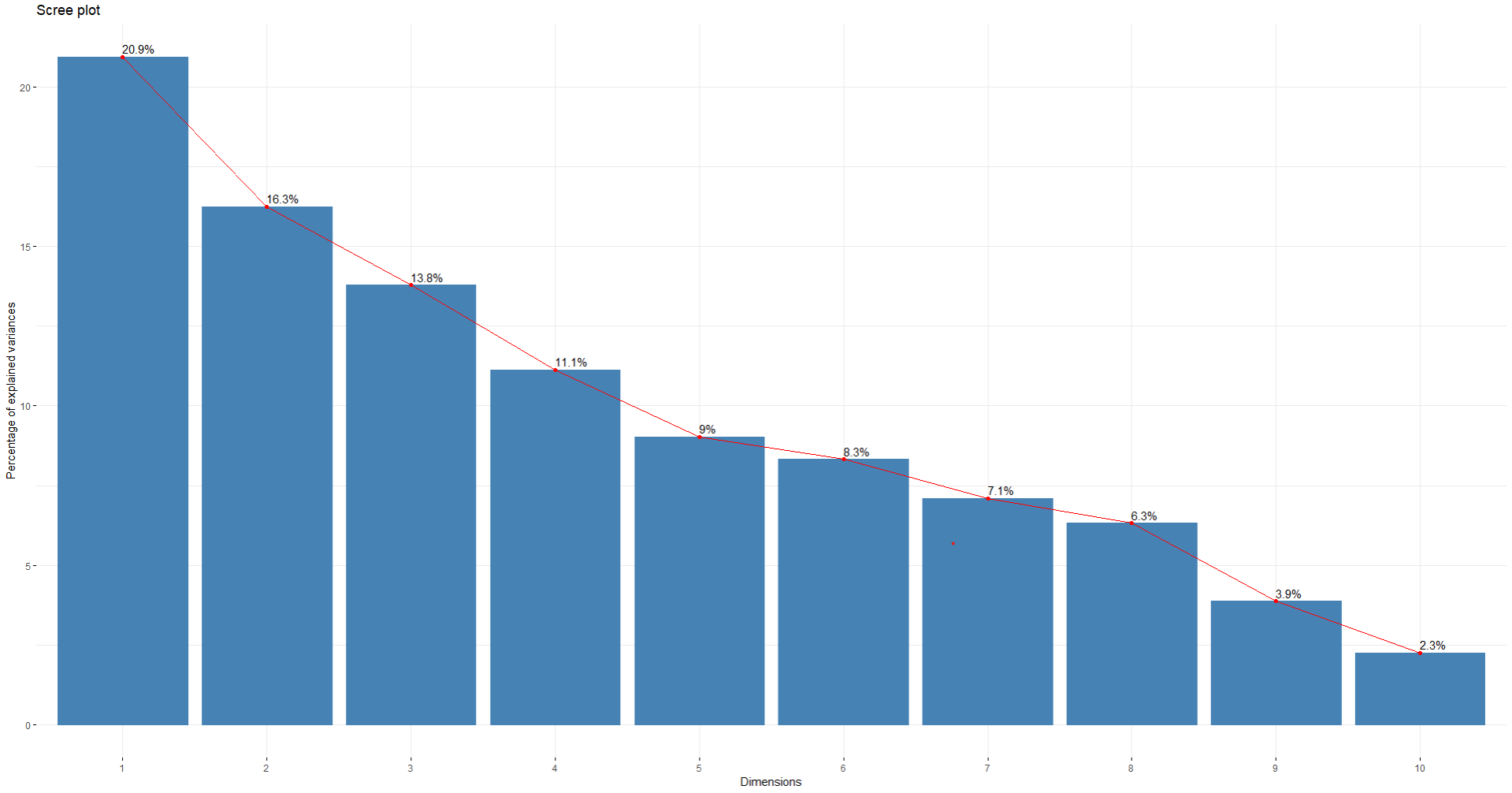


Number of clusters suggested by nbclust is 5

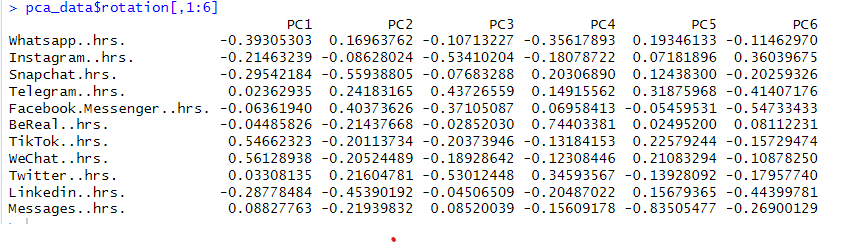
cluster analysis has identified groups of users with similar usage patterns or preferences for social media apps.

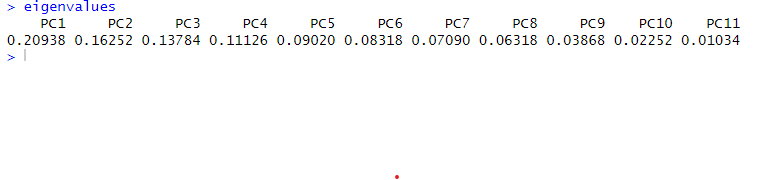
**PCA**

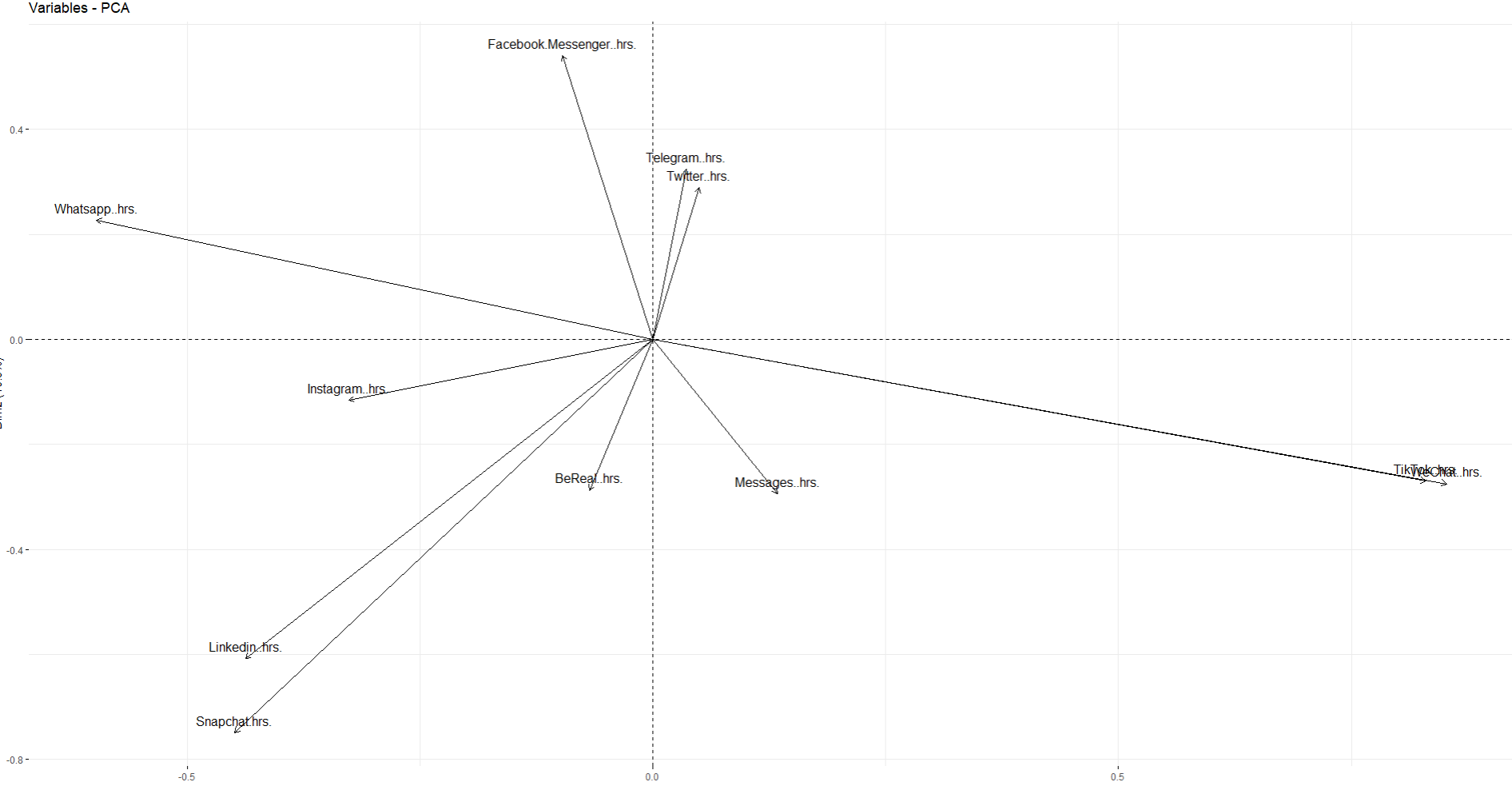


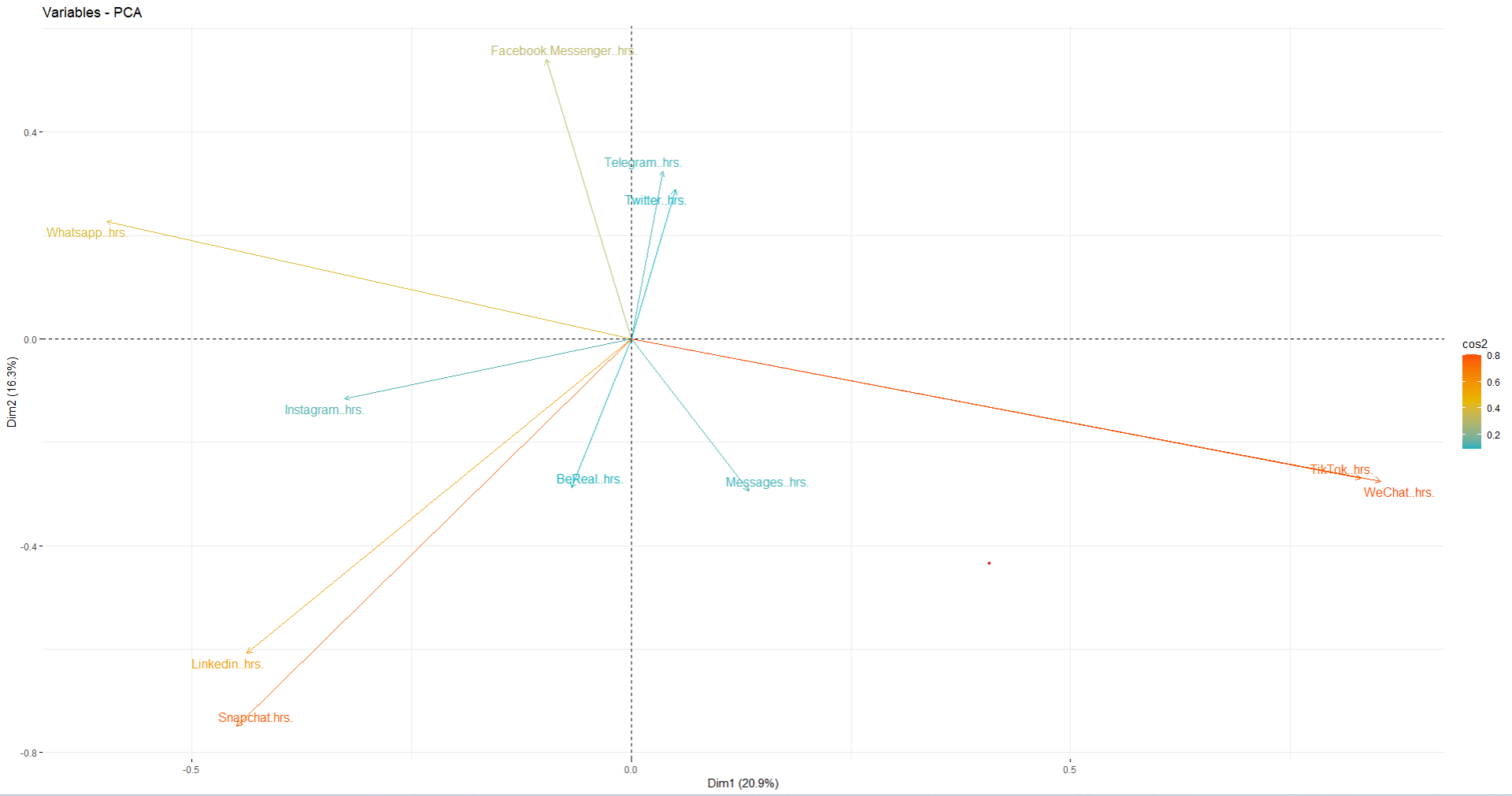


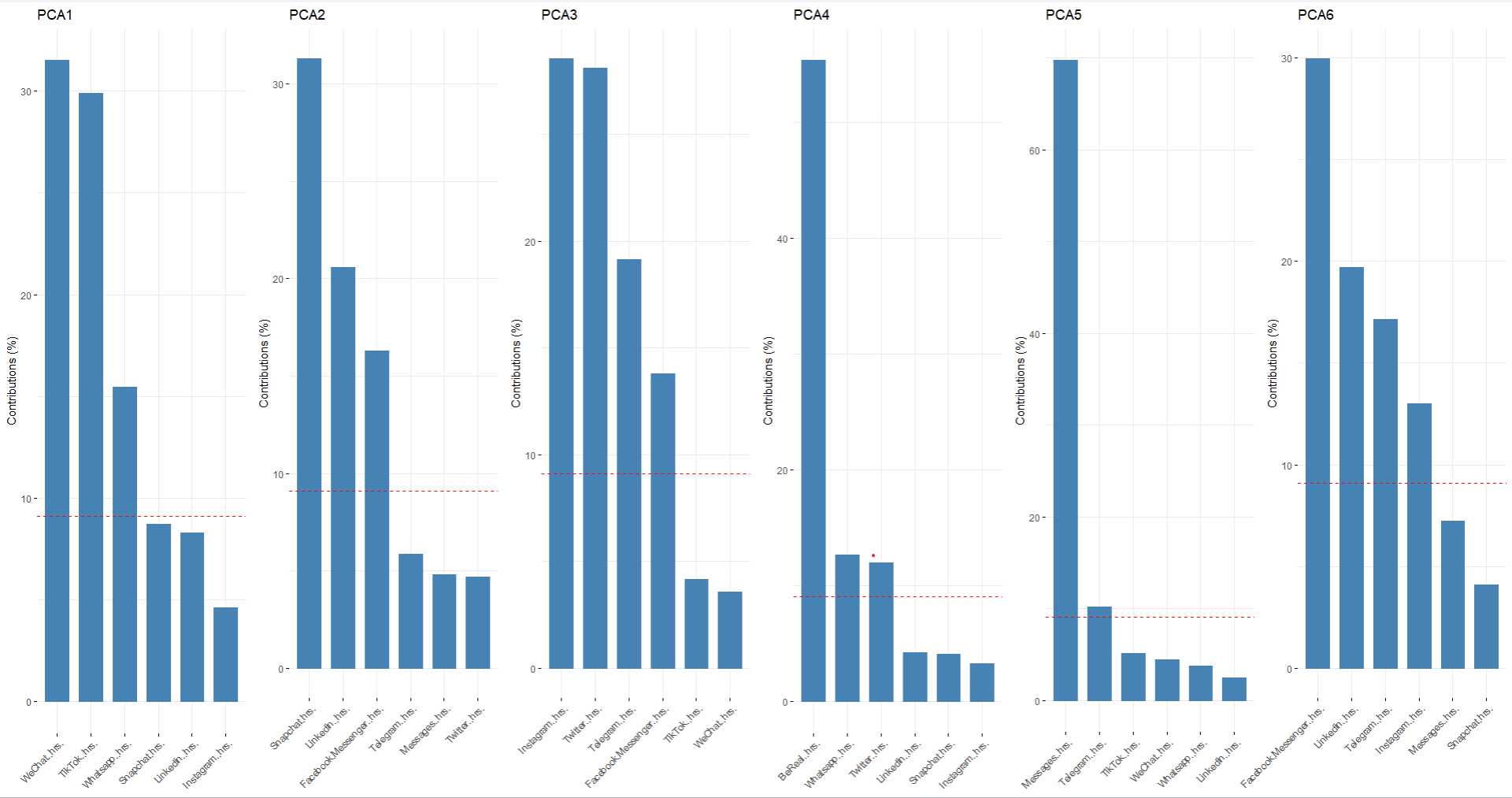
From the scree plot, it can be observed that the first six PCs explain over 80% of the total variance in the data.











Based on the PCA analysis performed on the dataset, several conclusions can be made:

The first six principal components explain over 80% of the total variance in the data, with PC1 contributing the most at 31%.

PC1: The variables that have the highest positive correlation with PC1 are WeChat, Whatsapp, and TikTok, indicating that these platforms have a significant impact on the overall social media usage patterns.

PC2: The variables that contribute most to this component are Snapchat, Linkedin, and Facebook. This component can be interpreted as a measure of the usage of visually-oriented social media platforms.

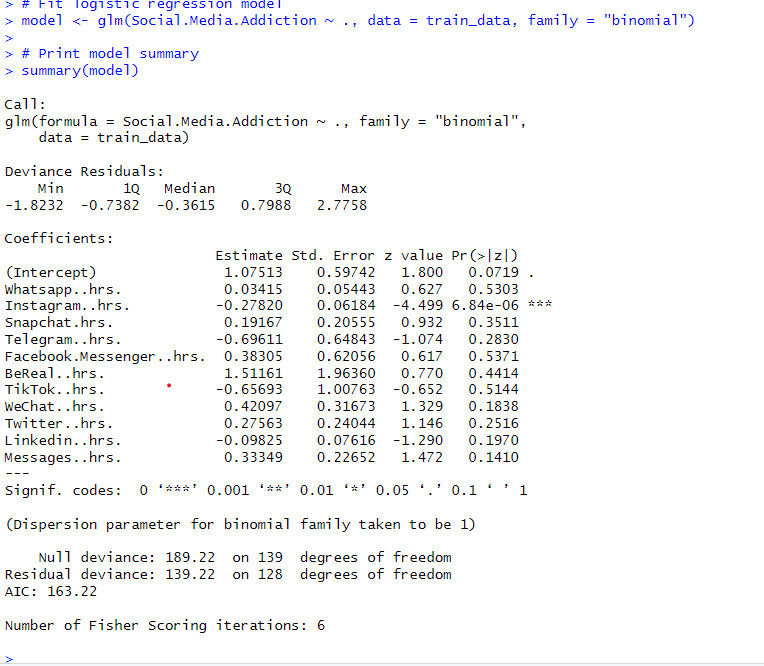
PC3: The variables that contribute most to this component are Instagram, Twitter, Telegram, and Facebook. This component can be interpreted as a measure of the usage of social media applications to get to know others' social lives.

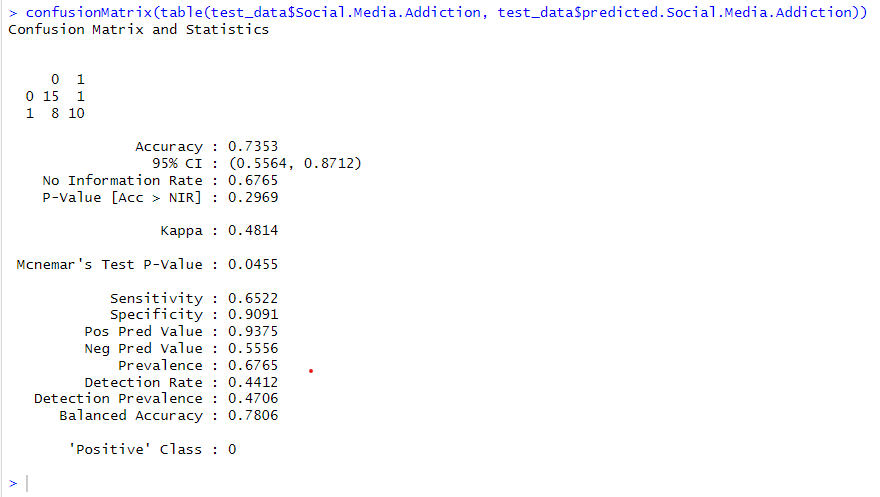
PC4: The variables that contribute most to this component are BeReal, Whatsapp, and Twitter. This component can be interpreted as a measure of the usage of social media platforms for interaction with others.

PC5: The variable that contributes most to this component is Messages. This component can be interpreted as a measure of the usage of texting others not using the internet.

PC6: The variables that contribute most to this component are Facebook, Linkedin, and Telegram. This component is much very similar to PC3.

**Logistic Regression**





Accuracy is 73% and the AIC score is 163.22