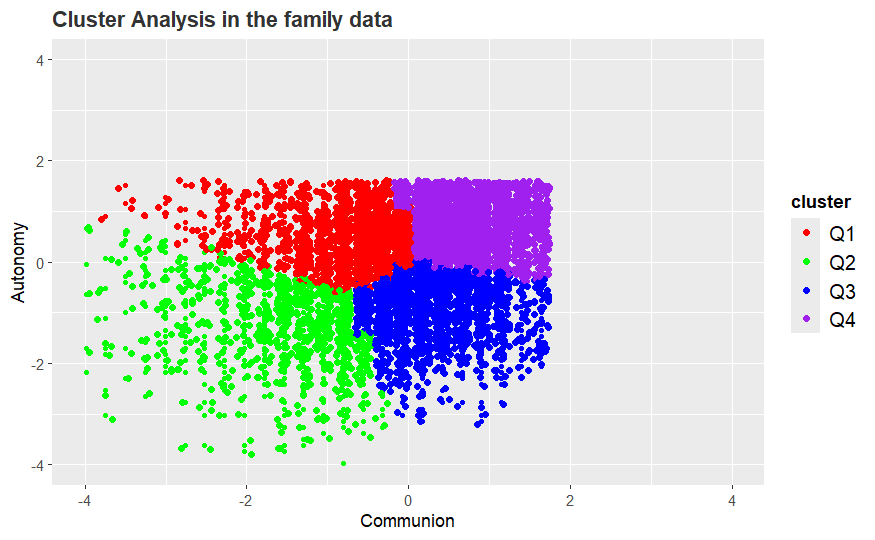
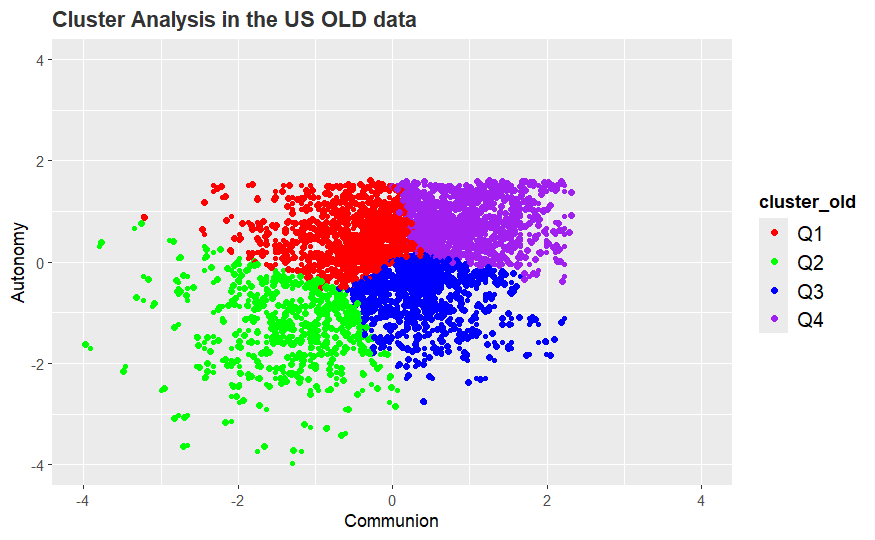
For goal 3 analysis

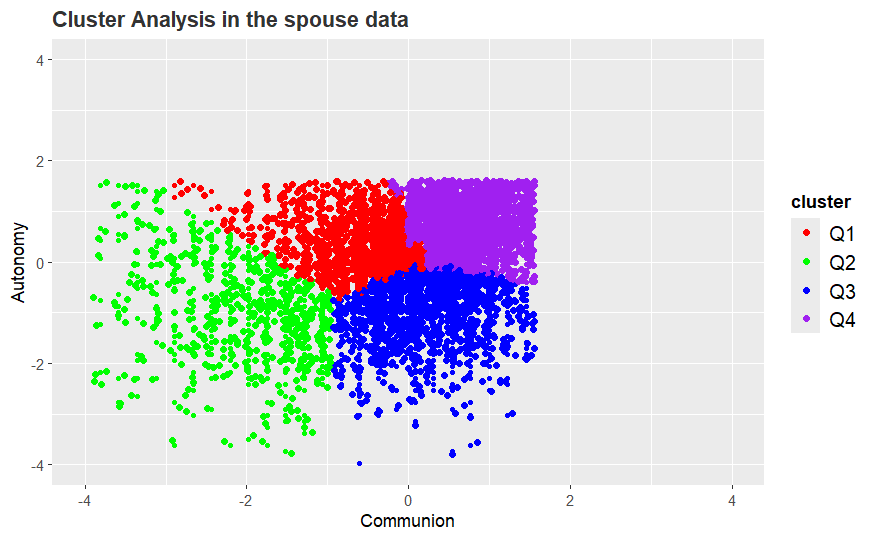
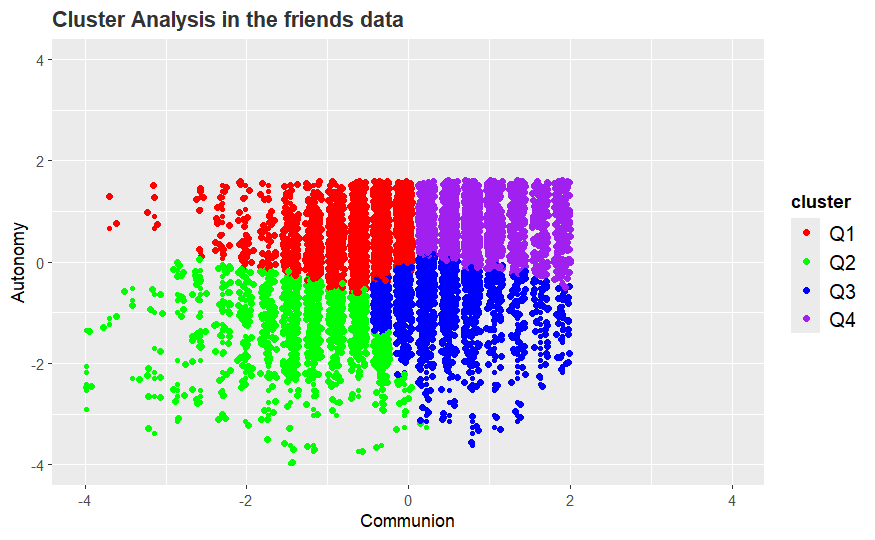
We draw the plots and use the API to see the similarity between all communion clustering and factorial communion (family, friends and spouse)

*The Adjusted Rand Index (ARI) ranges between -1 and 1 and is used to evaluate the performance of clustering. Here are some general guidelines for interpreting ARI values in terms of similarity:*

* ***ARI ≈ 1****: Perfect agreement, indicating that the clustering result is identical to the true classification.*
* ***0.5 ≤ ARI < 1****: High similarity, indicating a strong agreement between the clustering result and the true classification.*
* ***0 < ARI < 0.5****: Moderate similarity, indicating some agreement between the clustering result and the true classification, but with noticeable differences.*
* ***ARI = 0****: No similarity, indicating that the clustering result is equivalent to random classification.*
* ***ARI < 0****: Worse than random, indicating that the clustering result is worse than a random classification.*

In practical applications, you can set different thresholds based on your needs to determine what can be considered similar. Generally, an ARI greater than 0.5 can be considered to indicate a high level of similarity, which meets the requirements of most practical applications.





These are the 4 plots of Kmeans clusters about all communion, family\_communion, friends\_comminion and spouse\_comminon. Then I did ARI index to see the similarity between the first clustering method with the other 3 methods,

Although they look alive, when I combined the clustering results with M2ID, the results differ a lot.

For family communion,

[1]"Adjusted Rand Index: 0.455004545246945"

For friends communion,

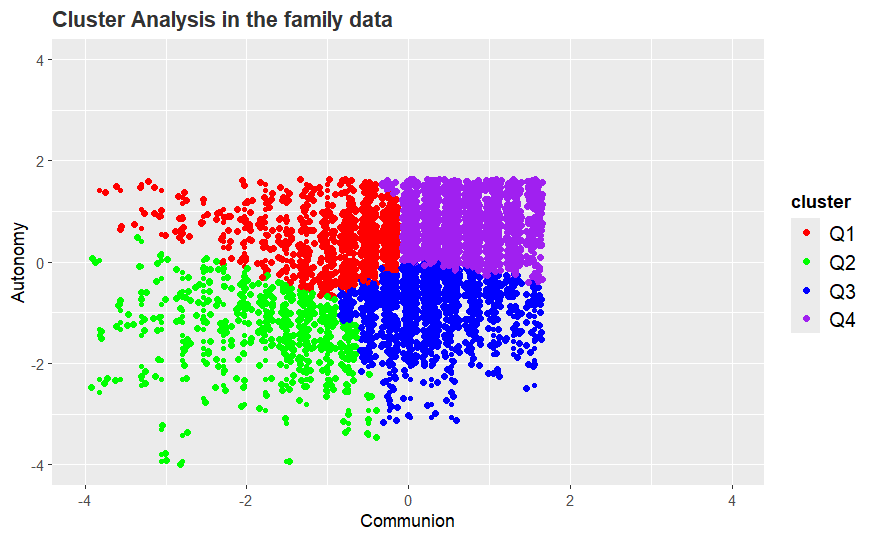
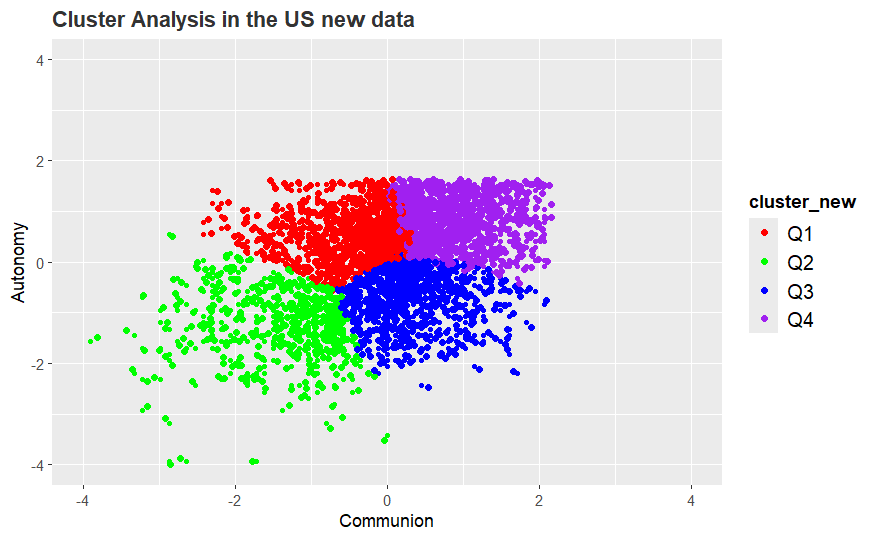
[1] "Adjusted Rand Index: 0.413188554201495"

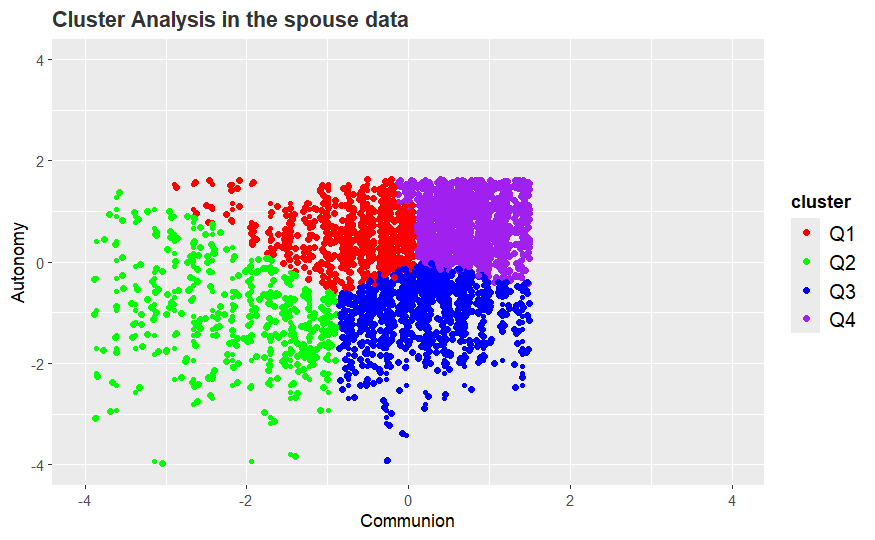
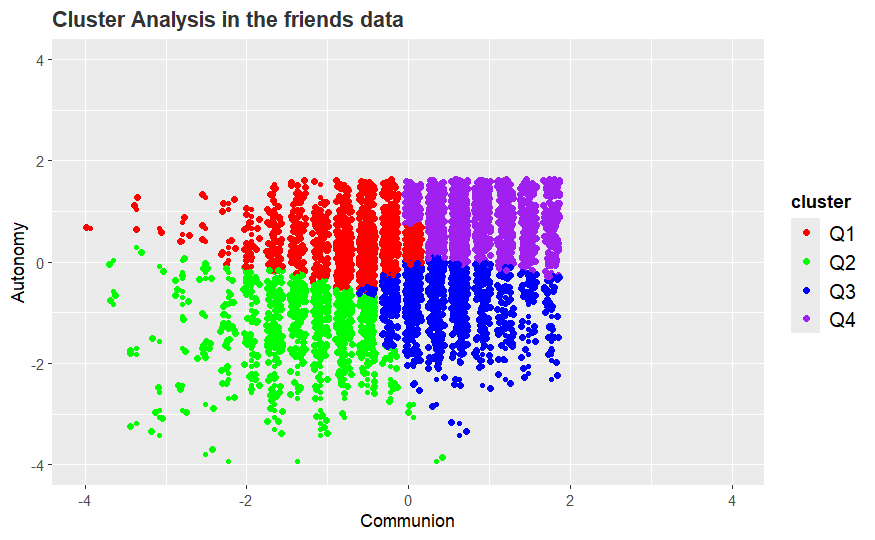
For spouse communion,

[1] "Adjusted Rand Index: 0.395971459245447"

**Which means they differ with each other.** **All the 3 factors( family, friends and spouse) combined contribute to the distribution of all communion.**

In order to confirm this result, I used the midus2 data as a test to check if it is the same way with the midus2 data.





FOR family communion,

[1] "Adjusted Rand Index: 0.472170409095875"

For friends communion,

[1] "Adjusted Rand Index: 0.485035796652963"

For spouse communion,

[1] "Adjusted Rand Index: 0.42697784547028"

Which confirms our result that **All the 3 factors( family, friends and spouse) combined contribute to the distribution of all communion. And no single factor has significant influence on the clustering results.**