Assignment: Programming II

EDH7916

Benjamin Skinner

Using `kmeans.RDS`, please answer the following questions. Use comments to respond as necessary.

Questions

1. After making the initial assignments with the "quick-and-dirty" code, merge the assignments back to the data (everything should be in order for a quick `cbind()` or `bind_cols()`) and then plot, assigning a unique color to each group. How did we do the first iteration?
2. The k-means algorithm can be sensitive to the starting points since it finds locally optimal solutions (no guarantee that the solution is the best of all possible solutions). Run the initial code a couple of times and see how your fit changes. Do some points move between groups?
3. Look through the `my_kmeans()` function and give a short answer on how the function will run with the following arguments:
   - `my_kmeans(data, 3)`
   - `my_kmeans(data, 3, standardize = TRUE)`
   - `my_kmeans(data, 3, nstarts = 10)`
4. What happens if you assume 4 or 5 groups? Run the `my_kmeans()` function with those options and make a plotly plot. How does it look — better fit? worse fit?

Submission details

- Save your script (`<lastname>_assignment_programming_ii.R`) in your `scripts` directory.
- Push changes to your repo (the new script and new folder) to GitHub prior to the next class session.