Group Project requirements:

* type everything, including your code and the report, in R Markdown
* encourage you to create an R Shiny app to show during your presentation.
* The paper will likely be about 10 to 20 pages in length, including graphs, tables, and figures. (80% of the project grade.)
* Your group should plan to speak for 40-45 minutes. you should explain your algorithm, your data set, and walk us through your analysis, including showing graphs and tables. (20% of the project grade.)
* Date: April 17, April 22.

**Topic**: Hierarchical Clustering (top-down version) \*specifically\* Divisive Algorithms

**Report:**

The report can be broken up into several components:

**1. Introduction**

* Briefly introduce the concept of clustering and its importance in data analysis.
* Define hierarchical clustering and mention that you'll specifically focus on divisive algorithms.
* Outline the objectives of the report.

**2. Mathematics behind Divisive Hierarchical Clustering**

* Explain the basic principles of hierarchical clustering.
* Describe the divisive approach, where clusters are divided recursively.
* Discuss the mathematical criteria used for splitting clusters (e.g., maximizing inter-cluster dissimilarity).
* Provide equations or algorithms that illustrate the process.

**3. Analysis on a Data Set**

* Describe the data set you'll be using for analysis (e.g., its size, attributes, and source).
* Apply divisive hierarchical clustering to the data set using appropriate algorithms or software.
* Present the results of the clustering, including any visualizations such as dendrograms or cluster heatmaps.
* Interpret the clusters obtained and discuss their relevance or insights gained from the analysis.

**4. Diagnostic and Remedial Measures**

* Discuss how to assess the quality of hierarchical clustering results.
* Describe diagnostic measures such as silhouette scores or dendrogram visualizations.
* Explain how to identify and address issues like suboptimal cluster separation or excessive cluster sizes.

**5. Strengths, Weaknesses, and Uses**

* Highlight the strengths of divisive hierarchical clustering, such as its ability to handle large datasets and its interpretability.
* Discuss weaknesses, such as sensitivity to initial conditions and computational complexity.
* Explain common use cases, such as market segmentation, biological classification, or image segmentation.

**6. Conclusion**

* Summarize the key points discussed in the report.
* Reflect on the effectiveness of divisive hierarchical clustering for the given data set.
* Offer recommendations for further exploration or improvements.

**References**

* Provide citations for any sources or literature referenced in the report.