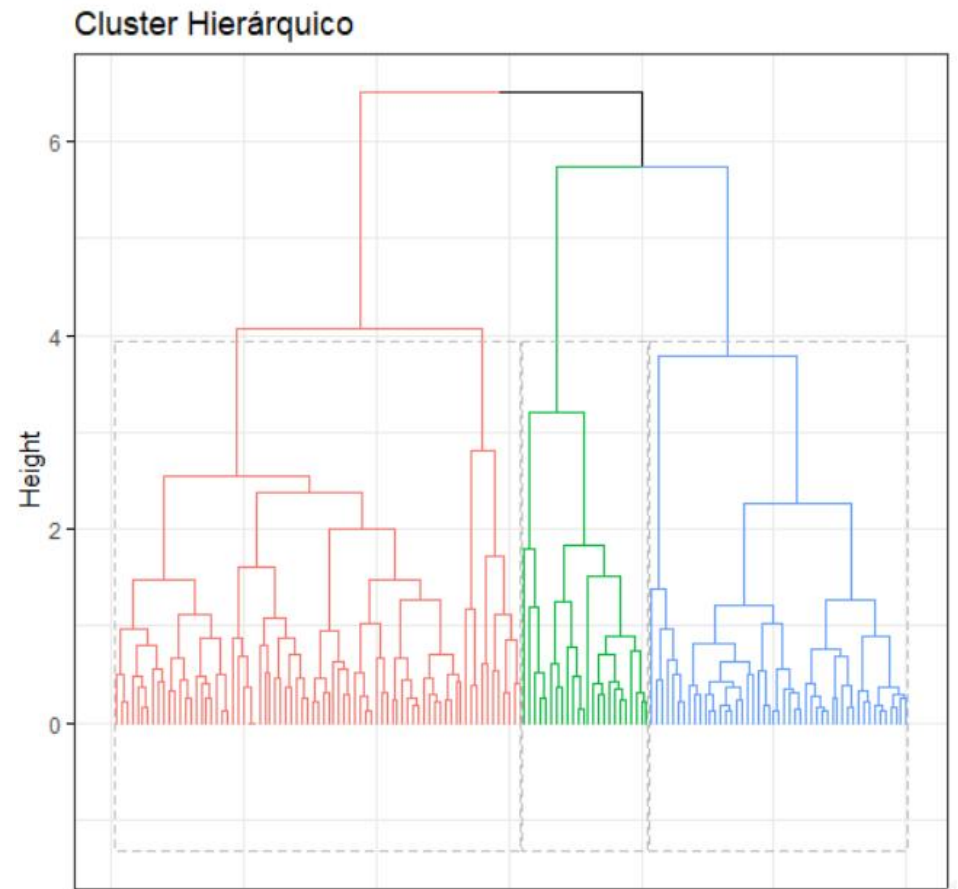


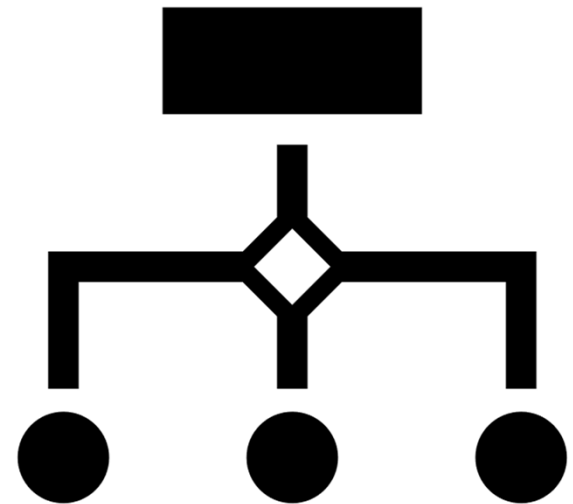
Cluster Hierárquico

- Aglomerativa: começa com pontos em grupos individuais e a cada etapa funde os pares mais próximos. Requer uma noção de proximidade. Mais comuns
- Divisiva: Começa incluindo todos, e a cada etapa divide até que reste apenas grupos únicos
- Dendograma



Bisecting k-means

- Cluster hierárquico do tipo divisivo



Hiper Parâmetros

- distanceMeasure: Medida de distância: Opções: “euclidean” e “cosine”. (padrão: euclidean)
- k: Número de clusters. (padrão: 4)
- maxIter: número máximo de iteração. (padrão: 20)
- minDivisibleClusterSize: o número mínimo de divisões se ≥ 1 ou a proporção mínima se < 1 (padrão: 1)

Avaliação da Performance

- Coeficiente Silhouette
- Medida do quanto cada ponto em um cluster está próximo do cluster vizinho
- Valor entre -1 e 1
- Próximo a 0 indica que os pontos estão próximos do limite
- Próximo a 1 indica que os pontos estão distantes do cluster vizinho
- Próximo a -1 indica que os pontos estão no cluster errado

