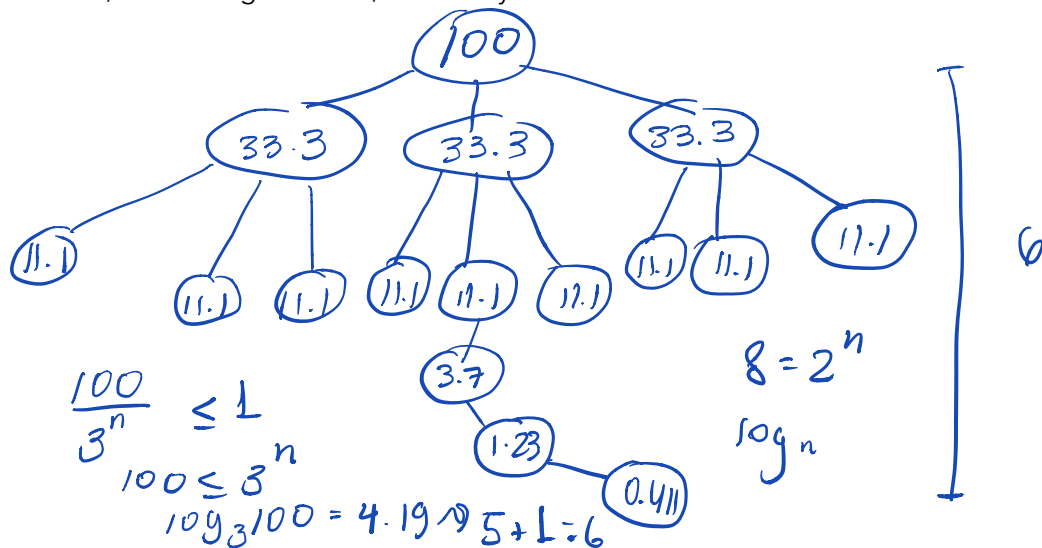


Hierarchical Clustering Quiz

- Hierarchical clustering methods produce
 - Clusters at multiple levels, where there are:
 - Singleton clusters at the bottom
 - Fewer larger clusters as you go up the hierarchy
 - A single cluster at the top
- I create a hierarchical clustered model, using top-down divisive clustering, from a dataset of 100 instances, and setting k to be 3, how many levels will there be in the model?



- In agglomerative hierarchical clustering, we start with singleton clusters and repeatedly merge clusters. The critical aspect in determining the structure of the resulting dendrogram is the metric used to measure how far apart two clusters are.

4. Distance measures for clustering

single link $\rightarrow \min_{x_1 \in C_1, x_2 \in C_2} D(x_1, x_2)$

complete link $\rightarrow \max_{x_1 \in C_1, x_2 \in C_2} D(x_1, x_2)$

average link $\rightarrow \frac{1}{|C_1|} \frac{1}{|C_2|} \sum_{x_1 \in C_1} \sum_{x_2 \in C_2} D(x_1, x_2)$

centroids $\rightarrow D\left[\left[\frac{1}{|C_1|} \sum_{x_1 \in C_1} x_1\right], \left[\frac{1}{|C_2|} \sum_{x_2 \in C_2} x_2\right]\right]$

Ward's method $\rightarrow \sum_{x_1 \in C_1, x_2 \in C_2} D(x_1, \mu_{C_1 \cup C_2})^2$