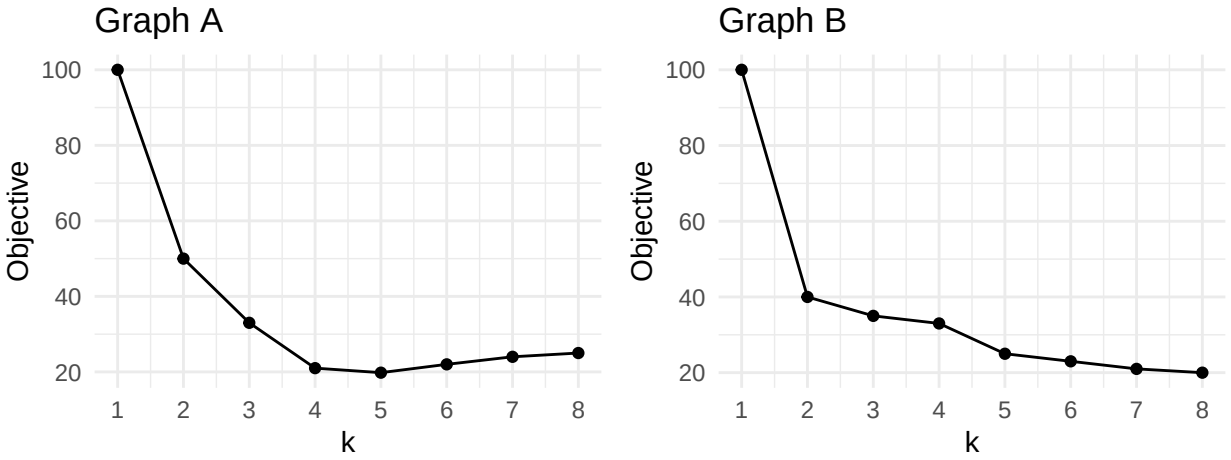
Sta-395 (Sp24) Quiz #1 - 1/25/24

Recall that the k-means objective function is defined as follows:

$$Q = \sum_{i=1}^{n} ||\mathbf{x}_i - \mathbf{z}_{A(\mathbf{x}_i)}||^2$$

Where the goal of the k-means algorithm is to find a set of k prototypes, $\{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$, that minimize Q for a pre-specified value of k and fixed assignment function, $A()$.

Question #1: Graphs A and B (below) show hypothetical values of this objective function at the termination of the k-means algorithm for different values of k. One of these graphs is realistic while the other isn't. Write the corresponding letter of the *realistic graph*. You may assume $n \gg k$ for all displayed values of k.



Question #2: In your own words, briefly describe each step of the k-means algorithm. You do not need to describe the initialization or termination of the algorithm, only the algorithm's core steps. You may assume that k prototypes have been already been intialized. Limit your answer to a single short sentence describing each step.