

Question 1 :

$$\text{mean}(C_J) = \frac{1}{|C_J|} \times \sum_{x \in C_J} x$$

$$\begin{aligned} \text{cost}(C_1, C_2, \dots, C_k; z_1, z_2, \dots, z_k) &= \sum_{j=1}^k \sum_{x \in C_j} \|x - z_j\|^2 \\ &= \sum_{j=1}^k \sum_{x \in C_j} (\|x - \text{mean}(C_j) + \text{mean}(C_j) - z_j\|^2) = \\ &= \sum_{j=1}^k \sum_{x \in C_j} \left(\underbrace{(x - \text{mean}(C_j))}_a + \underbrace{(\text{mean}(C_j) - z_j)}_b \right)^2 \\ &= \sum_{j=1}^k \sum_{x \in C_j} (x - \text{mean}(C_j))^2 + \underbrace{2 \sum_{j=1}^k \sum_{x \in C_j} (x - \text{mean}(C_j)) (\text{mean}(C_j) - z_j)}_0 \\ &\quad + \sum_{j=1}^k \sum_{x \in C_j} (\text{mean}(C_j) - z_j)^2 \\ &= \underbrace{\sum_{j=1}^k \sum_{x \in C_j} (x - \text{mean}(C_j))^2}_{\text{Cost}(C_j; \text{mean}(C_j))} + \sum_{j=1}^k \sum_{x \in C_j} (\text{mean}(C_j) - z_j)^2 \\ &= \text{cost}(C_j; \text{mean}(C_j)) + \sum_{j=1}^k \sum_{x \in C_j} (\text{mean}(C_j) - z_j)^2 \end{aligned}$$