Hierarchical Clustering.

Agglomerative:

1. Start with every point in its own cluster
2. At each step, merge the 2 closest clusters.
3. Stop when

$$\sum_{p_i \in C_1} d(p_i, \mu_1) \qquad \sum_{x_i \in C_2} d(x_i, \mu_2)$$

$$\sum_{y_i \in C_{12}} d(y_i, \mu_{12})$$

$$\mu_1 = \frac{\sum_{p_i \in C_1} p_i}{|C_1|} \qquad \mu_2 = \frac{\sum_{x_i \in C_2} x_i}{|C_2|} \qquad \mu_{12} = \frac{\sum_{y_i \in C_{12}} y_i}{|C_1|+|C_2|}$$

$$\sum_{p \in C_1} d(p, \mu_1) + \sum_{p \in C_2} d(p, \mu_2)$$

$$= \sum_{p_i \in C_1} \left\| p_i - \frac{\sum_{p_i \in C_1} p_i}{|C_1|} \right\| + \sum_{p \in C_2} \left\| p - \frac{\sum_{p \in C_2} x_i}{|C_2|} \right\|$$

$$\sum_{p \in C_{12}} d(p, \mu_{12}) = \sum_{p \in C_{12}} \left\| p - \frac{\sum_{p \in C_1} \sum_{p \in C_2} p}{|C_1|+|C_2|} \right\|$$