Theoretical task 10.

Recommendations: all solutions should be short, mathematically strict (unless qualitative explanation is needed), precise with respect to the stated question and clearly written. Solutions may be submitted in any readable format, including images.
Submission link: here

1. How would the steps of K-Means algorithm change, if in minimization criterion the euclidean distance is replaced by Manhattan distance (L1 distance)? What about its computational complexity?

2. (a) Show that K-Means criterion and the following criterion are equivalent:

$$Q(C) = \sum_k \sum_{x_i, x_j \in C_k} \frac{x_i^\top x_j}{|C_k|},$$

   where C_k is cluster with label k

   (b) Given that result, what technique can be used in K-Means algorithm?