

Homework 5

Kelsey Neis - neis@umn.edu - 3942198

11/15/2021

Question 1

For each of the described data sets, decide what type of clustering should be used (hierarchical or partitional, exclusive or non-exclusive (overlapping/fuzzy), complete or partial (incomplete)). Briefly explain your reasoning if you feel there may be several possible answers. Note: we are using partitional and hierarchical in the more relaxed use of the terms to mean un-nested or nested, respectively.

a) Grouping of cities based on the various locations, namely county (local region), state or province, and country.

hierarchical, exclusive, complete

b) A supermarket manager wants to group all goods into several categories, each of which has multiple subcategories.

hierarchical, non-exclusive, complete

c) You want to group all locations on Earth as to whether they belong to a tropical rainforest, a deciduous forest, or an evergreen forest. Here, each location corresponds to a grid cell of surface area 1km by 1km, and a location can have more than one variety of forest or may be unforested.

partitional, non-exclusive (fuzzy), partial. Partial because there will be many grid cells that do not have any forest and will not be interesting to my clustering experiment.