HDDA Tutorial: Distance

Department of Econometrics and Business Statistics, Monash University

Tutorial 3

Work in groups of two people:

1. Consider the *age* (in years) and *height* (in cm) of both you and the other person (you are allowed to lie about these). Compute the Euclidean distance between you and the other person for these two variables.
2. Repeat question 1 but use the Manhattan distance.
3. Repeat question 2 but measure height in metres.

Select from the following list the types of cuisines that you enjoy:

- Chinese food
- Indian food
- Italian food
- Japanese food
- Lebanese food
- Mexican food
- Thai food
- British food

4. Compute a Jaccard similarity between you and the other person with regards to your taste in food.
5. Compute a Jaccard distance between you and the other person with regards to your taste in food.
6. How would you define a distance between you and the other person that takes into account height, age and food preference.
7. Load in the Beer Dataset. Using numerical variables only, find the Euclidean distance between Pabst Extra Light and Augsberger. Do NOT use the `dist` function and for now, do NOT standardise the data.
8. Repeat the previous question but this time standardise the data.