Data Mining, Sheet 8

In this exercise sheet we will use clustering on real-world data. As always, you are free to use any technology that you deem fit.

Exercise 1

Read one of the tutorials on clustering, e.g.:
`https://www.kaggle.com/code/ryanholbrook/clustering-with-k-means`
or
`https://www.kaggle.com/code/satishgunjal/tutorial-k-means-clustering`

Exercise 2

Obtain data on historical buildings in Hamburg from `https://www.hamburg.de/bkm/denkmalliste/`. The OpenData format contains geographical locations.

Exercise 3

Cluster the data according to their location and visualize the result. Use the ellbow-method to find the optimal number of clusters. Try to improve your results based on the dates of the historical sites. Can you also use similarity of the descriptions of the historical sites?

Exercise 4

Write a short report on your findings.