Grouping Data

⚠ Changes before 22/23
 □ Add Learning Outcomes for each week □ Point to clustering code and show DBSCAN, k-Means, etc. as applied to the Airbnb data. □ Stress different view of clustering as part of a pipeline, not an absolute "just k-means it"; link to thinking about paradigms in CASA0001; □ thinking about proxies (why would you find ppsqm more/less useful than price? what are you really measuring?
Connects to both CASA0001 and CASA0007 and CASA0005.

Overview

Important
This week's Learning Outcomes are:
1.
2.3.

Preparation

Videos

Come to class prepared to briefly present:

Video	Markdown for Note-taking
Classification	Notes
Clustering	Notes

Readings

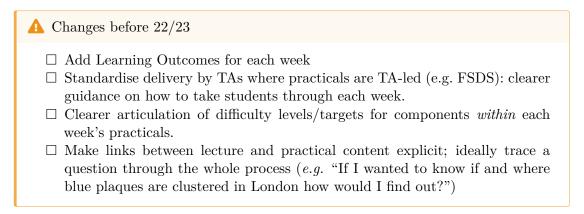
Come to class prepared to briefly present:

- (Shapiro and Yavuz 2017) URL
- (Wolf et al. 2021) DOI
- $\bullet\,$ (Singleton and Arribas-Bel 2021) DOI

Activities

- Padlet: [Collaborative Agenda]
- Complete the short Moodle quiz associated with this week's activities.

Practical



The practical can be downloaded from GitHub.

References

Shapiro, W., and M. Yavuz. 2017. "Rethinking 'distance' in New York City." Medium. https://medium.com/topos-ai/rethinking-distance-in-new-york-city-d17212d249 19.

Singleton, Alex, and Daniel Arribas-Bel. 2021. "Geographic Data Science." Geographical Analysis 53 (1):61–75. https://doi.org/10.1111/gean.12194.

Wolf, Levi John, Sean Fox, Rich Harris, Ron Johnston, Kelvyn Jones, David Manley, Emmanouil Tranos, and Wenfei Winnie Wang. 2021. "Quantitative Geography III: Future Challenges and Challenging Futures." *Progress in Human Geography* 45 (3). SAGE Publications Sage UK: London, England:596–608.