Welcome

Workshop 2: Exploring Madrid's Housing Market using Spatial Data Science

About me

- From Sheffield
- PhD student at the University of Liverpool
- Data Analytics and Society CDT, funded by the ESRC (2020)
- Undergraduate degree in Human Geography (University of Liverpool, 2017)
- Play football as part of a community project in Liverpool providing integration and support for female refugees and asylum seekers

Data Analytics & Society

Refugee women empowered and united by Liverpool football club - BBC News

Comfort Angels - Shado Magazine (shado-mag.com)

Geographic Data Science Lab (GDSL)

Twitter: geodatascience
Our People - Geographic Data Science Lab - University of Liverpool

- The GDSL is a research centre at the University of Liverpool, working at the intersection between Data Science and Geography.
- We are interested in making sense of how our world works; turning new forms of spatial data into information that we can use to make intelligent decisions.
- Collectively, we span the disciplines of econometrics, spatial analysis, spatial statistics, demography, population geography and computer science.

Hedonic Modelling with Machine Learning

- A hedonic model is used to estimate the market price for a commodity, by regressing the implicit value of its attributes (Rosen, 1974).
- Standard government practice when assessing properties for tax purposes
- A house is seen as a bundle of characteristics
- House price = Sum of locational (access to schools, parks, hospitals), structural (housing features) and neighbourhood attributes (air quality, crime rate, life expectancy)
- Using data from Idealista (Spain, Italy, Portugal)

idealista

Why study the housing market?

- Link with house prices and the economy
- Houses are the largest source of debt in the UK
- 2007 US subprime mortgage crisis caused the financial crash
- New pressure on the rental market, housing affordability
- Where you live is a social determinant (residential sorting/ segregation)
- Key part of wealth that is passed down through generations
- Understand the growth of cities

Machine Learning

Prediction tasks: predicting house price change, mass appraisal

- Random Forest (Rico-Juan and La Paz, 2021)

Classification tasks: creating housing submarkets

- Principle Component Analysis and Clustering (Wu and Sharma, 2012)

Deep learning: satellite and street view image classification

- Deep Neural Network (Law *et al.*, 2019)

Workshop: Exploring the housing market in Madrid

- The workshop is split into 2 sections – exploratory (spatial) analysis and unsupervised machine learning
- There are links in the text to explain some of the techniques
- Optional Tasks
- Feel free to ask questions
- Let me know if you spot any mistakes
- More data on github if you are interested

References

Law, S., Paige, B. and Russell, C., 2019. Take a look around: using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*(5), pp.1-19.

Rico-Juan, J.R. and de La Paz, P.T., 2021. Machine learning with explainability or spatial hedonics tools? An analysis of the asking prices in the housing market in Alicante, Spain. *Expert Systems with Applications*, *171*, p.114590.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, *82*(1), pp.34-55.

Wu, C. and Sharma, R., 2012. Housing submarket classification: The role of spatial contiguity. *Applied Geography*, *32*(2), pp.746-756.