CASA0013 – Foundations of Spatial Data Science

Assessment #3
Data-Led Executive Briefing

Deadline: 11 January 17:00
Weighting: 65% of the total mark for the module
Submission: Via Moodle

Assessment outline:

The focus of this assessment is to test the student's ability to make use of concepts and programming methods covered in class as part of an analytical process to support decision-making in a non-academic context.
There are two parts to this assessment, and you must submit both parts:

1. An **Executive Briefing** that may be written using any writing tool, but we would encourage you to consider using Markdown, RMarkdown, or LaTeX (this part counts for 60% of total mark for this submission)
2. A **Reproducible Analysis** that must be a Jupyter Notebook (this part counts for 40% of total mark for this submission)

The word limit for the Executive Briefing is 1,500 words. There is no limit to the amount of code in the Reproducible Analysis notebook. The score bands for marking your work are presented on the Marking Scheme Matrix posted on Moodle.

Assessment task:

Students are required to present an analysis of data from the Inside Airbnb web site for London. The **Executive Briefing** consists of presenting the challenges and opportunities relating to Airbnb's operations in London using programming and data science methods to support your arguments. You're expected to choose and address your briefing to a specific audience: The Mayor of the Greater London Authority or the Chief Executive of an Investment Company, both of which should be considered senior-level occupations with limited time. As an example, in their analysis students may wish to develop the evidence either for or against the regulation of listings on Airbnb in London, or for or against investing in the Airbnb platform in London, in which case the briefing document should reference existing policies, where relevant, and should make recommendations based on the analysis undertaken. Some possible briefing topics and examples of executive briefings are provided below.

The briefing may be written without substantially new modelling or coding other than the code written in practicals. For example, your analysis can be developed based on the sensible use of descriptive statistics. The use of charts, maps and tables is recommended to advance your arguments and convey your message, with attention to the thoughtful use of captions. All charts, maps and tables must be generated from Python and we should be able to confirm this by finding the appropriate table, chart, or map in the Reproducible Analysis.

The structure and sections of the Executive Briefing are the following:

- Executive summary (summary of your briefing creating a summary of each of the sections below, 200-300 words)
- Background (problem framing, 300-400 words)

- Data Analysis (description of the analysis, approach and results, 400-500 words)
- Conclusion (what was learned, limitations and what are your recommendations, 300-400 words)
- References (use the standard APA/Harvard-style referencing, does not count towards word count)

The **Reproducible Analysis** must be written in Python and may draw on concepts and methods covered in both 'Quantitative Methods' and 'GIS and Science'. It should be possible to 'Restart Kernel and Run All' such that all charts and tables used in the briefing document are produced. The reproducibility will be tested by selecting "Restart Kernel and Run All" using the sds:2021 Docker environment to reproduce your entire analysis (e.g. data extraction, cleaning, transformation, clustering, charts, tables, etc.). We will not install libraries 'by-hand'. You may wish to make any data that you use available via GitHub or Dropbox to ensure its accessibility to your notebook code. A Python notebook template will be provided for guidance. You should zip up your notebook prior to submission and then submit this as a zip file so that it is not corrupted by Moodle.

Possible Briefing Topics:

These are indicative topics, and you should feel free to strike out if some other aspect of the topic and data are of interest:

- **Impact of Airbnb on local area rental markets** — this would require some assumptions about listings and lettings based on available data but as long as these are clearly stated this would be a strong approach; there are good examples of models used in other cities that it may be possible to draw on, or adapt to, London. You may want to consider things like the type of listing and the issues around the Short- and Long-Term Rental markets.
- **Impact of Airbnb on London's Tourism Economy** — this would look at the distribution of London's tourism venues and, possibly, hotels alongside Airbnb listings in order to evaluate the extent to which tourism 'dollars' might be spent in ways that positively impact less tourist-oriented areas if we assume (again, detail the assumptions) that some percentage of a tourist's dollars are spent locally in an area. Again, there may be models developed elsewhere that could be adapted for the London context.
- **Opportunities and Risks arising from Covid-19** — it should/may be possible to assess the impact of Covid-19 on London's short- and long-term rental markets by looking at entry to/exit from the Airbnb marketplace by comparing more than one snapshot of London data. Again, this will require some reasonable assumptions to be drawn (are all flats withdrawn from Airbnb going back on to the Long-Term Rental Sector?) but these can be documented and justified.
- **Opportunities for Place- or Listing-Branding** — identifying key terms and features/amenities used to market listings by area and using these to identify opportunities for investment or branding. This would benefit from the use of NLP approaches and, potentially, word embeddings to identify distinctive patterns of word use as well as, potentially, One-Hot encoding to identify specific amenities that appear associated in some way with particular areas.
- **The Challenge of Ghost Hotels** — evaluating ways to automatically identify ghost hotels from the InsideAirbnb data and then, potentially, assessing their extent and impact on local areas where they dominate either 'proper' hotel provision or other types of listings. You will need to consider the way that Airbnb randomly shuffles listings to prevent exactly this type of application and textual similarity via NLP is an obvious application.
- **The Professionalisation of Airbnb** — this could be treated either as a regulatory challenge (is Airbnb not benefiting locals) or an investment opportunity (is this a way to 'scale' or develop new service offers for small hosts) depending on your interests. You will need to consider the different types of hosts and evaluate ways of distinguishing between them (e.g., number of listings, spatial extent, etc.).
- **Impact Profiles** — a geodemographic classification of London neighbourhoods based on how they have, or have not, been impacted by Airbnb. This would require you to think about how to develop a

classification/clustering of London neighbourhoods and use data to develop 'pen portraits' of each so that policy-makers could better-understand the range of environments in which Airbnb operates and why a 1-size-fits-all regulatory approach may be insufficient. Again, this could be argued from either standpoint or even both simultaneously: these areas are already so heavily impacted that regulation is too little, too late, while these other areas are 'at risk'.

Examples of Executive Briefings:

Although the following examples are all much longer than permitted under the assessment format, they are exemplary in their communication of the data and key findings in a manner that is clear, straightforward, and well-illustrated:

- Smith, D.A. (2010), Valuing housing and green spaces: Understanding local amenities, the built environment and house prices in London, GLA Economics; URL: https://www.london.gov.uk/sites/default/files/gla_migrate_files_destination/GLAE-wp-42.pdf
- Travers, T. Sims, S. and Bosetti, N. (2016), Housing and Inequality in London, Centre for London; URL: https://www.centreforlondon.org/publication/housing-and-inequality-in-london/
- Bivens, J. (2019), The economic costs and benefits of Airbnb, Economic Policy Institute; URL: https://www.epi.org/publication/the-economic-costs-and-benefits-of-airbnb-no-reason-for-local-policymakers-to-let-airbnb-bypass-tax-or-regulatory-obligations/
- Wachsmuth, D., Chaney, D., Kerrigan, D. Shillolo, A. and Basalaev-Binder, R. (2018), The High Cost of Short-Term Rentals in New York City, Urban Politics and Governance research group, McGill University; URL: https://www.mcgill.ca/newsroom/files/newsroom/channels/attach/airbnb-report.pdf
- Chapple, K. (2009), Mapping Susceptibility to Gentrification: The Early Warning Toolkit, Centre for Community Innovation; URL: https://communityinnovation.berkeley.edu/publications

Useful references:

Students might want to review the partial bibliography available here; this is by no means complete and you will likely find other relevant work but you may find it useful for spurring your thinking on what to study and how to study it. You might also want to have a look at the following manuscripts related to London and the UK:

- KeyNest (2019), Understanding Airbnb regulations in London, KeyNest; URL: https://keynest.com/blog/airbnb-regulations-london
- Airbnb (n.d.), I rent out my home in London. What short-term rental laws apply?, Airbnb; URL: https://www.airbnb.co.uk/help/article/1340/i-rent-out-my-home-in-london-what-shortterm-rental-laws-apply
- Hostmaker (2018), Important Airbnb regulations and laws you should know about in London, Hostmaker; URL: https://hostmaker.com/blog/important-airbnb-regulations-country-laws-know-london/
- Bailey & Milton (2017), The suburbanisation of poverty in British cities, 2004-16: extent, processes and nature. URL: https://doi.org/10.1080/02723638.2017.1405689

Frequently Asked Questions:

Do references, declaration of authorship and captions count towards the word count?

No, only the text in the body of your document counts towards the word count. The 1,500 words limit is indicative and students are allowed up to 150 words more or less that limit.