

Checkpoint call - Monday May 22nd: 13:30 - 14:30 (GMT+1)

Attendees

- Anastasia Rizzo
- Astrid van Toor
- Leigh Feaviour

Notes

- We discussed our three independent EDAs and agreed there were useful elements from all of them that will be combined into a single EDA report for the project.
- We agreed that the purpose of the EDA, to provide some initial observations about the data so we can build an effective model, has been achieved.
- It could be argued that our initial question about identifying groups of listings by location, and whether that has an impact on price, can already be answered from the EDA (we know that Manhattan has the most expensive properties). However, we believe that by using the learning from the EDA to build one or more clustering algorithms we will be able to provide much better analytics without fundamentally changing the question.
- We will explore if there are clusters of neighbourhoods within Manhattan, possibly by room type, to provide more insights about price. We will then apply the same to the other neighbourhood groups, and then to New York as a whole, so see how the data clusters and what would be useful to report back to Airbnb.

Actions

Anastasia	Prepare consolidated EDA report, including proposed features to drop, and distribute on 23rd May
All	Review the consolidated EDA on 24th May and feedback over WhatsApp (or a call if needed) to finalise so we can move to modelling
Astrid	Set up Github repository
All	Research clustering models to use