DA3
Assignment 2
Airbnb prediction models

35 points
Individual or in Pairs

Tasks

- Business: Your task will be to help a company operating small and mid-size **apartments** hosting **2-6 guests**. The company is set to price their new apartments not on the market.

- Build a price prediction model similarly to how we did in our case study for London.

- Discuss your modeling decisions and compare your results to those of the case study.

Getting data

- Visit the website we got the airbnb dataset from: http://insideairbnb.com/get-the-data.html and download data from **another city**

- *Any city with at least 10k rows (not London) any date*

- *Everybody will have a different city - **Please sign up for a city HERE**.*

- May use any date

Regarding tasks

- You may use other variables we used in class,
- You may do different feature engineering
- You may make other sample design decisions
- But document your steps

- Have at least 3 different models and compare performance
  - Argue for your choice of models

Extra task for pairs

- Extra task **If** you do it in pairs.
- Please pick two dates, one before and one after lockdowns in that city.
- Carry out the exercise before the lockdown
- Take your selected model and use it on the post-lockdown date
- Compare the predictive power of the model. Discuss briefly in the report

Submit two documents to moodle

- A summary report (pdf), max 3 pages including tables and graphs discussing your work. It is targeted to data science team leaders
  - Can use technical language
  - But need to be the point
  - Focus on key decision points, results, interpretation, decision
  - For pair projects, write one extra page on the additional task
- Technical report -- a markdown / notebook in pdf/html with more technical discussion.
  - May include code snitppets
  - May include additional tables and graphs
  - Detail all decisions you made
- Reports should link to code in github

Scoring weights

- Data prep, label and feature engineering (30%)
- Model building, prediction and model selection (40%)
- Discussion of steps, decisions and results (15%)
- Quality of the write-up (15%)