**MSCA 37014 Final Project**

Airbnb is interested in better understanding data relating to price of listings on their website. They have hired you to explore the dataset provided to gain insight into its usefulness in the listing assessment process. Please download **listings.csv.gz** under Amsterdam from **http://insideairbnb.com/get-the-data.html**. The dataset consists of a random sample of homes that have been booked in Amsterdam during December, 2020.

Your tasks are to:

1. Carefully review the data and get an understanding of what you are working with.
2. If you choose (this is not required), incorporate other data that might be useful
3. Explore your data by analyzing the location and dispersion of relevant variables and process it as appropriate. This may mean, for example, removing outliers, imputing null values, creating composite variables, or considering the logs of values.
4. Create visualizations that will help the assessor and other stakeholders understand the data
5. Conduct tests to determine which variables are worth exploring
6. Build models using OLS regression to predict listing price
7. Perform diagnostics and goodness of fit tests on the model

At every step, explain the assumptions, limitations, and ramifications of what you are doing.

At the end, include a summary of what you did and what you found. Analyze the dataset and provide a report explaining your work and its results.

Please upload your final project report as a Jupyter Notebook with all the code cells and their outputs, along with commentary texts, tables or images in the markdown cells. Upload external data files if you have used any. If you used multiple files, upload them in a compressed folder.