# Introduction to the dataset

The data was collected from [www.insideairbnb.com](http://www.insideairbnb.com), which is a non-commercial website not associated with or endorsed by Airbnb or any of Airbnb’s competitors[[1]](#footnote-1).

For creating a map of the location of the listings in Copenhagen, we use the latitude and longitude of each listing. Even though each listing also has a specified zip code and neighbourhood name, there are some discrepancies between these two variables in the dataset. The accuracy of the location on Airbnb is generally a problem because location information for listing are anonymized by Airbnb[[2]](#footnote-2). For example, listings in the same building are anonymized individually such that the locations of the listings appear scattered in the area surrounding the actual address. According to Inside Airbnb, this means that the location of a listing in the available data on insideairbnb.com in practice will have and error margin of 0-150 metres of the actual address[[3]](#footnote-3). In the dataset, Inside Airbnb have compiled neighbourhood names for each listing by comparing listing’s geographic coordinates with a city’s definition of neighbourhoods[[4]](#footnote-4), i.e. neighbourhood names specified by the user on the listing is not used in this analysis.

1. <http://insideairbnb.com/about.html#disclaimers> [↑](#footnote-ref-1)
2. <http://insideairbnb.com/about.html#disclaimers> [↑](#footnote-ref-2)
3. <http://insideairbnb.com/about.html#disclaimers> [↑](#footnote-ref-3)
4. <http://insideairbnb.com/about.html#disclaimers> [↑](#footnote-ref-4)