## Goal

The objective of this exercise is to gain further practise in understanding and configuring relationships in PowerBI to gain data insights. We will import some real-world AirBnB data into PowerBI and perform some analysis to gain insight from the data collected.

## Steps

1. Import the following data files into PowerBI or SQL. In SQL you could take the opportunity to resolve some of the errors that you might encounter importing the data into PowerBI (hint : the property name contains characters that can be problematic when working with comma separated files) . Alternatively, choosing to Transform the data prior to Loading will allow you to resolve any load or type conversion errors you encounter. You can drop any rows which import with data errors if they represent only a small portion of the loaded data.
   * Property.csv
   * Property\_Summary.csv
   * Property\_Type.csv
   * Host.csv
   * Neighbourhood.csv
2. Review the data and draw out a sketch of the correct inter table relationships
   * Cardinality between tables
   * Primary and foreign keys (what are the joining fields between the tables)
3. In PowerBI add transformation steps to each source query which rename the primary keys of each table and use the data model view to amend and test the correct relationship configuration between the tables.

Once the tables are correctly joined you can begin your analysis and visualisation:

1. As a marketing analyst for AirBnB you’re interested in finding out which hosts have the most properties in the New York area. To start with, create a visualisation for the total number of properties you have in this area.
2. Drill down to find out which hosts have the most properties. Create a visualisation to show the top 10 hosts in terms of the number of properties they host. Hint: host name is not unique. You should utilise host ID to identify the top host.

Chart

Description automatically generated

1. For the top host, create a visualisation to show a breakdown of their properties by type
2. Create a map visualisation showing where these properties are using the latitude and longitude values. Hint : you might encounter some misleading values in these columns which alter your map. If that’s the case, consider filtering to an appropriate range of longitude and latitude values.

A picture containing graphical user interface

Description automatically generated

1. You want to see the Average Monthly Reviews for this top host. The attribute provided is too granular for meaningful analysis. Create a new column in the underlying database table and put the Average Monthly Review into suitable bands/buckets. Create a visualisation to show how many of their properties fall into each Average Monthly Review band (hint: you will need to do some data preparation to handle null values in the data).

Chart, bar chart

Description automatically generated

1. Now turn your attention to the Last Review Date for this host’s properties. Create a visualisation to show the count of properties by Last Review Date.

Chart, line chart

Description automatically generated

1. Now see if there are any properties for this host which are getting high numbers of reviews (we don’t know which are positive or negative at this stage!).Create a visualisation to identify the top 2 properties for this host based on the number of reviews received.