Nick & Molly

**Project 2 Proposal: Airbnb Comparison Between Seattle and Boston**

Through our data, we aim to look at a comparison of Airbnb data between Seattle and Boston. We will be exploring the two cities' data by looking at their listings and reviews.

**Extract:**

All files are CSV files extracted by using Pandas.

**Transform:**

Rename the original datasets to reflect the city they represent.

Realize that the listing id in reviews corresponds with the id value in listings.

Change the listing id column to id and delete the id column in reviews. Do this for both Boston and Seattle CSVs.

Merge the listings and reviews data sets on id as the primary key for Boston.

Merge the listings and reviews data sets on id as the primary key for Seattle.

Cleaned the data by including columns that we deemed as important information.

Remove all commas from our DataFrames and replace them with an empty space.

Convert Pandas DataFrame into CSV.

**Load:**

Created tables in PostgreSQL

Identified ID as primary keys

Import both CSVs into tables in PostgreSQL

**Sources:**

<https://www.kaggle.com/datasets/airbnb/boston?select=listings.csv>

<https://www.kaggle.com/datasets/airbnb/boston?select=reviews.csv>

<https://www.kaggle.com/datasets/airbnb/seattle?select=listings.csv>

<https://www.kaggle.com/datasets/airbnb/seattle?select=reviews.csv>