

**Project 2**

For this project, we decided to look at airbnb data in Seattle. The two tables we intend to use are the calendar.csv and listings.csv joined with “listing\_id” in the calendar table and “id” in the listings table. Our reason for linking these tables is to be able to compare the price in the calendars table with the various information in the listings table such as room type, square feet, zip code and reviews rating. After cleaning the relevant columns we plan to upload the data to PostgreSQL.

**Listings.csv - Columns to Clean**

A - id

X - host\_response\_rate

Y - host\_acceptance\_rate

Z - host\_is\_superhost

AO - zipcode

AW - property\_type

AX - room\_type

AY - accommodates

BE - square\_feet

BF - price

BG - weekly\_price

BH - monthly\_price

BV - number\_of\_reviews

BY - review\_scores\_rating

**Calendar.csv - Columns to Clean**

A - listing\_id

C - available

D - Price

Datasource: <https://www.kaggle.com/datasets/airbnb/seattle>

Calendar.csv

Listings.csv

**Members**

Nick Taheri

Jose Sandoval

Vanessa Kemp

Aidan Thompson