Exploring Airbnb Market Trends

Goal:

1. three csv files: airbnb_price.csv, airbnb_room_type.xlsx, and airbnb_last_review.tsv. airbnb_price.csv contains listing_id (unique identifier), price, neighborhood. airbnb_room_type.xlsx contains listing_id (unique identifier), listing description, room_type: shared, private, entir home/apt. airbnb_last_review.tsv contains listing_id (unique identifier), host_name, last_review (date when the listing was reviewd).
2. We are being asked to answer the following questions:
   a. When is the earliest date of review and the most recent date of review?
   b. How many prive room are in room_type
   c. Merge the price and room_type data frame together
   d. What is the average price of all rooms?

Technologies: Python, Pandas, Numpy

Description:

1. Read in the data using pandas.read_csv() for airbnb_price.csv and airbnb_last_review.tsv (with sep = '\t') and using pandas.ExcelFile() for airbnb_room_type.xlsx
2. Converted the last_review column by using pandas.to_datetime() and used .min() and .max() to retriet the earliest date and most recent date
3. Inspected sheetnames using .sheet_names and extracted sheet data using .parse()
4. Applied .str.lower() to clean categorical data for our value_counts() and retrieved the desired answer
5. Cleaned the 'price' column data through .str.strip(), str.replace(), and astype('float') to calculate average price .mean()
6. Merged price.csv and room_type.csv together through .merge() method
7. Delivered results