# Air BnB
# Price Comparison

# Background

AirBnB is online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities.

With the supplied data, we want to know :

1. What is the difference between lodging price in several cities in the U.S.?
2. Where is the highest review per month lodging?
3. Is there any correlation between price and the number of reviews?

# The Data : Before Cleaning

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 48895 entries, 0 to 48894
Data columns (total 15 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   id                              48895 non-null  object
 1   host_id                         48895 non-null  object
 2   host_name                       48874 non-null  object
 3   neighbourhood_group             48895 non-null  object
 4   neighbourhood                   48895 non-null  object
 5   latitude                        48895 non-null  float64
 6   longitude                       48895 non-null  float64
 7   room_type                       48895 non-null  object
 8   price                           48895 non-null  float64
 9   minimum_nights                  48895 non-null  int64
 10  number_of_reviews               48895 non-null  int64
 11  last_review                     38843 non-null  datetime64[ns]
 12  reviews_per_month               38843 non-null  float64
 13  calculated_host_listings_count  48895 non-null  int64
 14  availability_365                48895 non-null  int64
dtypes: datetime64[ns](1), float64(4), int64(4), object(6)
memory usage: 5.6+ MB
```

# Relevant Data (Samples After Cleaning)

| | id | neighbourhood_group | room_type | price | reviews_per_month |
|---|---|---|---|---|---|
| 1017 | 413775 | Manhattan | Entire home/apt | 275.0 | 0.24 |
| 46504 | 35282511 | Manhattan | Entire home/apt | 153.0 | NaN |
| 10342 | 7916103 | Brooklyn | Private room | 60.0 | 0.07 |
| 38268 | 30161460 | Manhattan | Private room | 199.0 | 3.27 |
| 27263 | 21539835 | Manhattan | Entire home/apt | 300.0 | 0.05 |
| 45103 | 34535938 | Queens | Entire home/apt | 150.0 | 3.00 |
| 34929 | 27695485 | Queens | Private room | 58.0 | 2.96 |
| 19575 | 15660209 | Queens | Entire home/apt | 95.0 | 0.23 |
| 2464 | 1240586 | Manhattan | Entire home/apt | 99.0 | 0.09 |
| 35345 | 28042014 | Queens | Private room | 80.0 | 7.56 |

# Price Distribution

Manhattan is expensive!

Where do you go if you want
the cheapest lodging?

- Private room : go to Staten Island
- Entire home : go to Bronx
- Shared room : go to Brooklyn
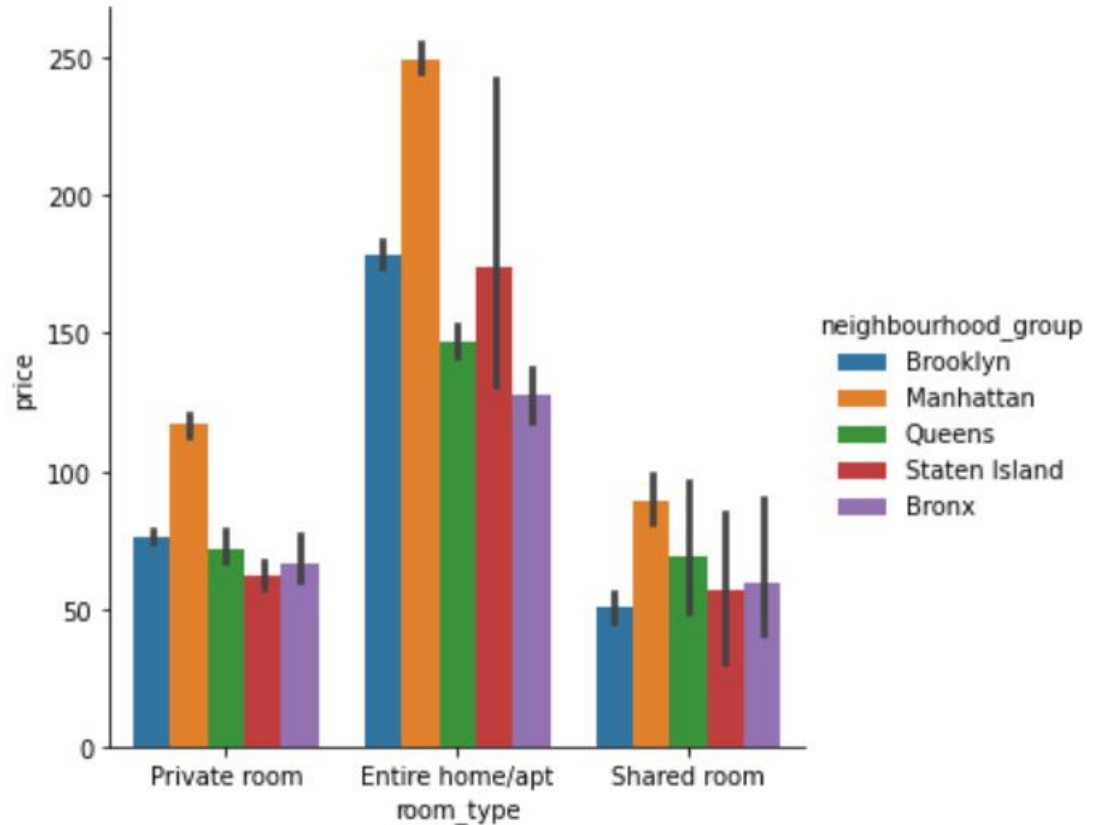
# Reviews

The number of reviews can be an indicator whether the lodging is good (lots of returning customers and reviews) or not (less reviews).

| review |  |
|---|---|
| 0.00 | 10052 |
| 0.01 | 42 |
| 0.02 | 919 |
| 0.03 | 804 |
| 0.04 | 655 |
| ... | |
| 17.82 | 1 |
| 19.75 | 1 |
| 20.94 | 1 |
| 27.95 | 1 |
| 58.50 | 1 |

Name: reviews_per_month, Length: 938, dtype: int64

Less reviews does not mean it is bad. Maybe it just started, or maybe the price is too expensive that no customers rent it.
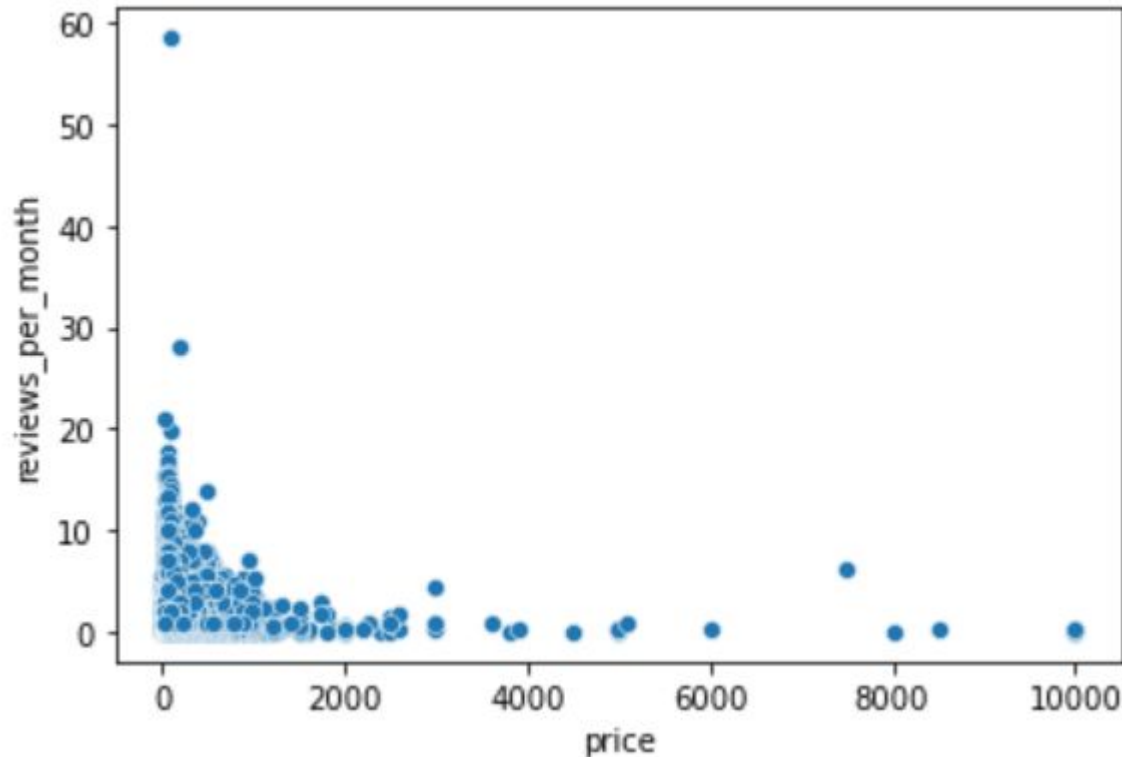
There are **10052 lodgings** that has no review at all.

There are **72 lodgings** that has 10 or more reviews per month
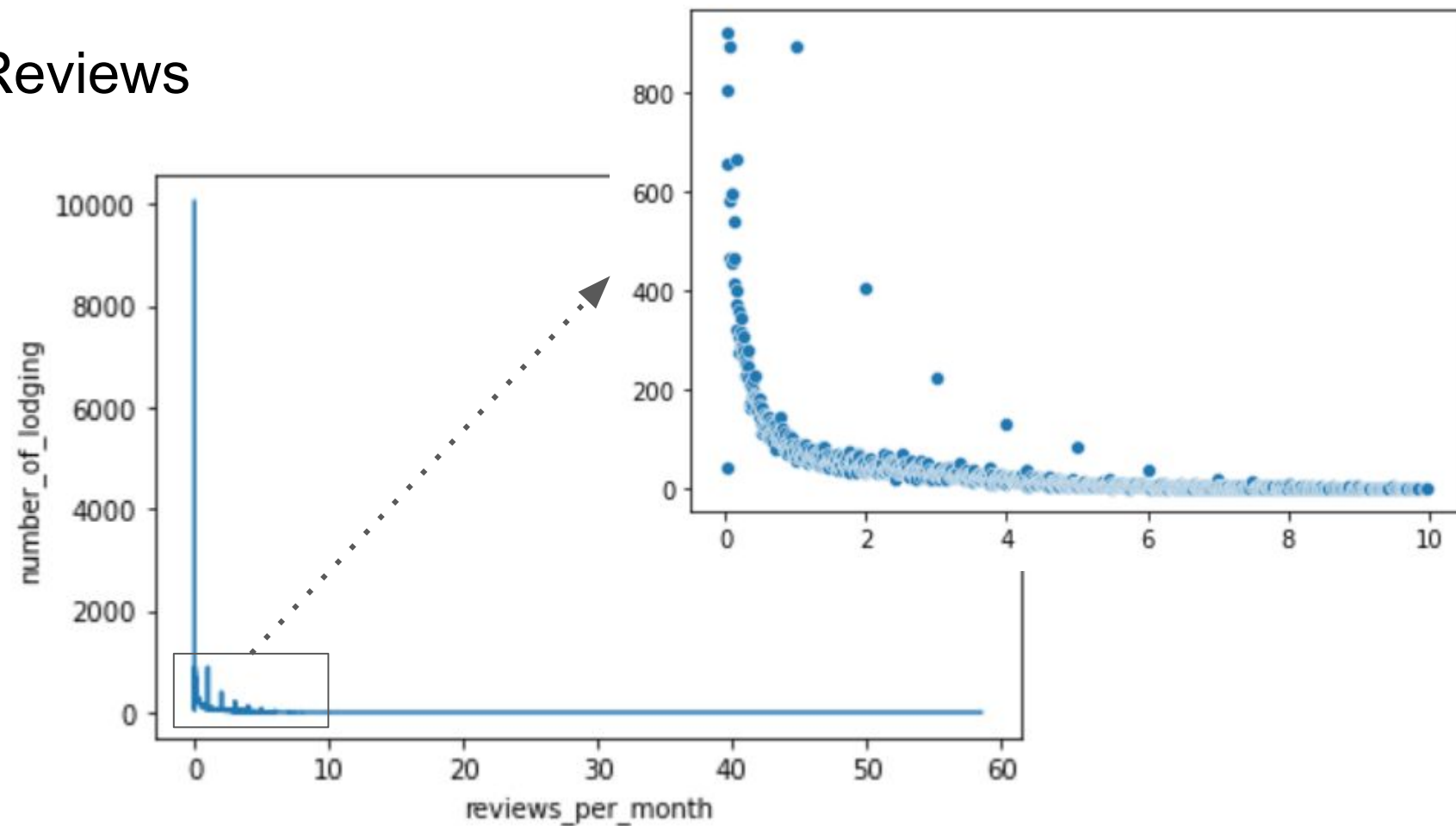
# Reviews

# Correlation Between Price and Reviews



Weak Correlation

ρ = -0.031