

```
import pandas as pd
df = pd.read_csv(r'C:\Users\derek\Documents\CSE351\AB_NYC_2019.csv')
# First step is to clean the data of any duplicate rows.
df.drop_duplicates()
#
```

host_id \	id	name
0	2539	Clean & quiet apt home by the park
2787		
1	2595	Skylit Midtown Castle
2845		
2	3647	THE VILLAGE OF HARLEM....NEW YORK !
4632		
3	3831	Cozy Entire Floor of Brownstone
4869		
4	5022	Entire Apt: Spacious Studio/Loft by central park
7192		
...
...		
48890	36484665	Charming one bedroom - newly renovated rowhouse
8232441		
48891	36485057	Affordable room in Bushwick/East Williamsburg
6570630		
48892	36485431	Sunny Studio at Historical Neighborhood
23492952		
48893	36485609	43rd St. Time Square-cozy single bed
30985759		
48894	36487245	Trendy duplex in the very heart of Hell's Kitchen
68119814		

\	host_name	neighbourhood_group	neighbourhood	latitude
0	John	Brooklyn	Kensington	40.64749
1	Jennifer	Manhattan	Midtown	40.75362
2	Elisabeth	Manhattan	Harlem	40.80902
3	LisaRoxanne	Brooklyn	Clinton Hill	40.68514
4	Laura	Manhattan	East Harlem	40.79851
...
48890	Sabrina	Brooklyn	Bedford-Stuyvesant	40.67853
48891	Marisol	Brooklyn	Bushwick	40.70184

48892	Ilgar & Aysel	Manhattan	Harlem	40.81475
48893	Taz	Manhattan	Hell's Kitchen	40.75751
48894	Christophe	Manhattan	Hell's Kitchen	40.76404

	longitude	room_type	price	minimum_nights
number_of_reviews \				
0	-73.97237	Private room	149	1
9				
1	-73.98377	Entire home/apt	225	1
45				
2	-73.94190	Private room	150	3
0				
3	-73.95976	Entire home/apt	89	1
270				
4	-73.94399	Entire home/apt	80	10
9				
...
...				
48890	-73.94995	Private room	70	2
0				
48891	-73.93317	Private room	40	4
0				
48892	-73.94867	Entire home/apt	115	10
0				
48893	-73.99112	Shared room	55	1
0				
48894	-73.98933	Private room	90	7
0				

	last_review	reviews_per_month	
calculated_host_listings_count \			
0	2018-10-19	0.21	6
1	2019-05-21	0.38	2
2	NaN	NaN	1
3	2019-07-05	4.64	1
4	2018-11-19	0.10	1
...
48890	NaN	NaN	2
48891	NaN	NaN	2

48892	NaN	NaN	1
48893	NaN	NaN	6
48894	NaN	NaN	1

	availability_365
0	365
1	355
2	365
3	194
4	0
...	...
48890	9
48891	36
48892	27
48893	2
48894	23

[48895 rows x 16 columns]

created missing indicator for: name
created missing indicator for: host_name
created missing indicator for: last_review
created missing indicator for: reviews_per_month

<AxesSubplot:xlabel='index'>

