

```
In [1]: import pandas as pd
import numpy as np
df = pd.read_csv("AB_NYC_2019.csv")
df.head()
```

```
Out[1]:
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	nurr
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	40.64749	-73.97237	Private room	149		1
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home/apt	225		1
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	40.80902	-73.94190	Private room	150		3
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home/apt	89		1
4	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	40.79851	-73.94399	Entire home/apt	80		10

```
In [2]: df.describe()
```

```
Out[2]:
```

	id	host_id	latitude	longitude	price	minimum_nights	number_of_reviews	reviews_per_month	calculated_
count	4.889500e+04	4.889500e+04	48895.000000	48895.000000	48895.000000	48895.000000	48895.000000	38843.000000	
mean	1.901714e+07	6.762001e+07	40.728949	-73.952170	152.720687	7.029962	23.274466	1.373221	
std	1.098311e+07	7.861097e+07	0.054530	0.046157	240.154170	20.510550	44.550582	1.680442	
min	2.539000e+03	2.438000e+03	40.499790	-74.244420	0.000000	1.000000	0.000000	0.010000	
25%	9.471945e+06	7.822033e+06	40.690100	-73.983070	69.000000	1.000000	1.000000	0.190000	
50%	1.967728e+07	3.079382e+07	40.723070	-73.955680	106.000000	3.000000	5.000000	0.720000	

	id	host_id	latitude	longitude	price	minimum_nights	number_of_reviews	reviews_per_month	calculated_
75%	2.915218e+07	1.074344e+08	40.763115	-73.936275	175.000000	5.000000	24.000000	2.020000	
max	3.648724e+07	2.743213e+08	40.913060	-73.712990	10000.000000	1250.000000	629.000000	58.500000	



In [54]: `df.shape`

Out[54]: (48895, 16)

First Approach

In [4]: `min_quar , max_quar = df.price.quantile([0.01,0.99])`
`min_quar,max_quar`

Out[4]: (30.0, 799.0)

In [8]: `rissh = df[(df["price"]>min_quar)&(df["price"]<max_quar)]`

In [9]: `rissh.shape`

Out[9]: (47744, 16)

In [57]: `rissh.head(10)`

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	nurr
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	40.64749	-73.97237	Private room	149	1	
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home/apt	225	1	
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	40.80902	-73.94190	Private room	150	3	

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	nurr
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home/apt	89		1
4	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	40.79851	-73.94399	Entire home/apt	80		10
5	5099	Large Cozy 1 BR Apartment In Midtown East	7322	Chris	Manhattan	Murray Hill	40.74767	-73.97500	Entire home/apt	200		3
6	5121	BlissArtsSpace!	7356	Garon	Brooklyn	Bedford-Stuyvesant	40.68688	-73.95596	Private room	60		45
7	5178	Large Furnished Room Near B'way	8967	Shunichi	Manhattan	Hell's Kitchen	40.76489	-73.98493	Private room	79		2
8	5203	Cozy Clean Guest Room - Family Apt	7490	MaryEllen	Manhattan	Upper West Side	40.80178	-73.96723	Private room	79		2
9	5238	Cute & Cozy Lower East Side 1 bdrm	7549	Ben	Manhattan	Chinatown	40.71344	-73.99037	Entire home/apt	150		1



```
In [58]: rr = df["price"]
rr.head()
```

```
Out[58]: 0    149
1    225
2    150
3     89
4     80
Name: price, dtype: int64
```

Second Approach

```
In [59]: from scipy import stats
```

```
iit = np.abs(stats.zscore(df["price"]))
iit.shape
```

Out[59]: (48895,)

```
In [60]: iit
```

```
Out[60]: array([0.01549307, 0.30097355, 0.01132904, ..., 0.15707024, 0.4069123 ,
               0.2611711 ])
```

```
In [48]: z_score = df[iit<1.96]
```

```
In [49]: z_score.shape
```

Out[49]: (48124, 16)

```
In [50]: z_score.head()
```

```
Out[50]:
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	nurr
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	40.64749	-73.97237	Private room	149	1	
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home/apt	225	1	
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	40.80902	-73.94190	Private room	150	3	
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home/apt	89	1	
4	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	40.79851	-73.94399	Entire home/apt	80	10	

```
In [52]: outliers = df[iit>1.95]
```

```
outliers.shape
```

```
Out[52]: (771, 16)
```

```
In [53]: outliers
```

Out[53]:	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nig
	85	perfect for a family or small group	19601	74303	Maggie	Brooklyn	Brooklyn Heights	40.69723	-73.99268	Entire home/apt	800
	345	The Brooklyn Waverly	89427	116599	Sahr	Brooklyn	Clinton Hill	40.68613	-73.96536	Entire home/apt	650
	496	Luxury 2Bed/2.5Bath Central Park View	174966	836168	Henry	Manhattan	Upper West Side	40.77350	-73.98697	Entire home/apt	2000
	653	4 BEDROOM -2BATHRM WEST VILLAGE DUPLEX TOWNHOUSE	248871	605463	West Village	Manhattan	West Village	40.73066	-74.00287	Entire home/apt	700
	662	Heart & Soul of Greenwich Village	250801	1314834	Rhona	Manhattan	Greenwich Village	40.73129	-73.99944	Entire home/apt	850

	48523	Tasteful & Trendy Brooklyn Brownstone, near Train	36308562	217732163	Sandy	Brooklyn	Bedford-Stuyvesant	40.68767	-73.95805	Entire home/apt	1369
	48535	Stunning & Stylish Brooklyn Luxury, near Train	36311055	245712163	Urvashi	Brooklyn	Bedford-Stuyvesant	40.68245	-73.93417	Entire home/apt	1749

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nig
48697	36388720	HUGE LUXURY CONDO – INCREDIBLE WATER VIEWS	273619215	Layla	Manhattan	Upper West Side	40.77665	-73.98867	Entire home/apt	750	
48757	36419574	Luxury & Spacious 1500 ft² MANHATTAN Townhouse	11454384	Ellen	Manhattan	Tribeca	40.71815	-74.01145	Entire home/apt	700	
48839	36452721	Massage Spa. Stay overnight. Authors Artist dr...	274079964	Richard	Brooklyn	Sheepshead Bay	40.59866	-73.95661	Private room	800	

771 rows × 16 columns



In []: