





Figure 2. Most frequent words in neighborhood descriptions from the exorbitant dataset, a total of 3591 (around 7%)

Description: df [10 × 2]

	word <chr>	n <int>
1	miles	1955
2	beach	955
3	hollywood	772
4	park	583
5	neighborhood	579
6	restaurants	563
7	los	549
8	hills	539
9	museum	538
10	minutes	502

1-10 of 10 rows

negative



positive

negative



positive

Figure 3. Top mentioned positive/negative terms for regular data

Figure 4. Top mentioned positive/negative terms for exorbitant data

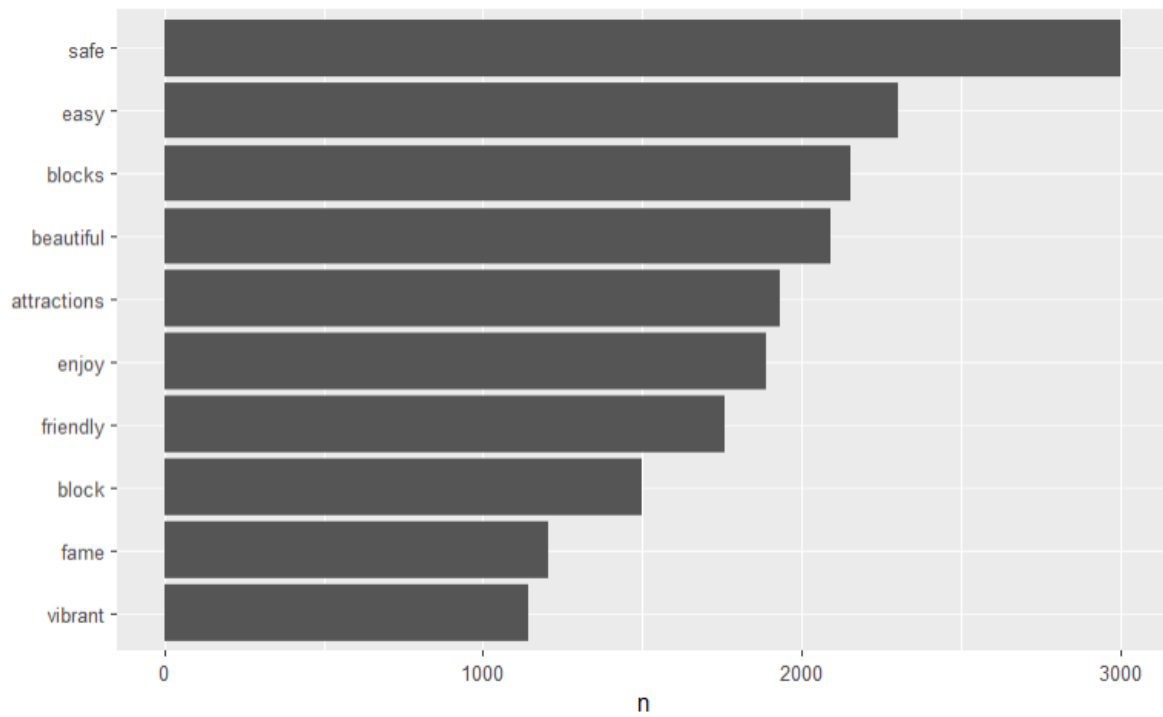


Figure 5. Top mentioned terms with respect to afind package for regular data

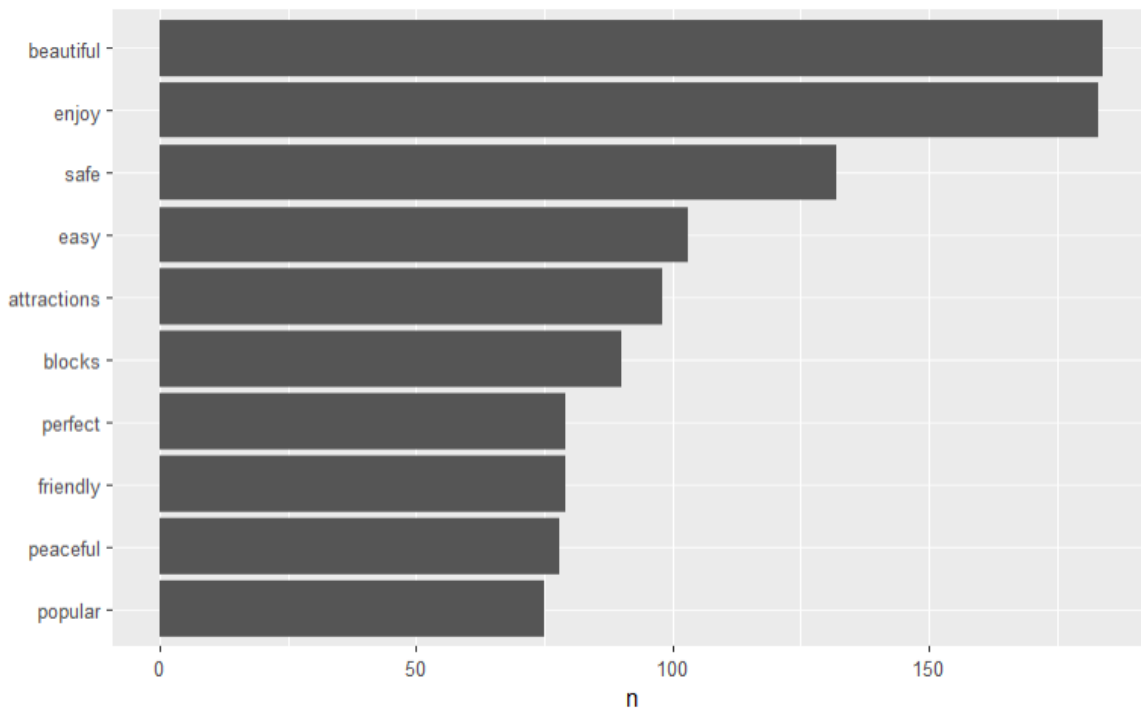


Figure 6. Top mentioned terms with respect to afind package for exorbitant data

Description: df [10 × 6]						
	price_level	word	n	tf	idf	tf_idf
	<chr>	<chr>	<int>	<float>	<float>	<float>
1	regular	furnished	637	0.0007654198	0.6931472	0.0005305486
2	regular	的	494	0.0005935909	0.6931472	0.0004114459
3	regular	超市	321	0.0003857139	0.6931472	0.0002673565
4	regular	outlets	232	0.0002787714	0.6931472	0.0001932296
5	regular	和	215	0.0002583442	0.6931472	0.0001730706
6	regular	至	206	0.0002475298	0.6931472	0.0001715746
7	regular	中心	205	0.0002463282	0.6931472	0.0001707417
8	regular	whittier	183	0.0002198930	0.6931472	0.0001524182
9	regular	rely	143	0.0001718289	0.6931472	0.0001191027
10	regular	elysian	142	0.0001706273	0.6931472	0.0001182699

Figure 7. Top unique words to regular dataset

Description: df [10 × 6]						
	price_level	word	n	tf	idf	tf_idf
	<chr>	<chr>	<int>	<float>	<float>	<float>
1	exorbitant	27.8	24	0.0003864112	0.6931472	0.0002678398
2	exorbitant	arm's	24	0.0003864112	0.6931472	0.0002678398
3	exorbitant	workspaces	24	0.0003864112	0.6931472	0.0002678398
4	exorbitant	nms	20	0.0003220093	0.6931472	0.0002231999
5	exorbitant	1.9miles	15	0.0002415070	0.6931472	0.0001673999
6	exorbitant	2.9miles	15	0.0002415070	0.6931472	0.0001673999
7	exorbitant	4.2miles	15	0.0002415070	0.6931472	0.0001673999
8	exorbitant	5.2miles	15	0.0002415070	0.6931472	0.0001673999
9	exorbitant	6.6miles	15	0.0002415070	0.6931472	0.0001673999
10	exorbitant	7.9miles	15	0.0002415070	0.6931472	0.0001673999

Figure 8. Top unique words to exorbitant dataset

Description: df [217,526 × 2]	
bigram	n
<chr>	<int>
los angeles	7284
in the	7069
of the	6792
to the	6380
is a	5424
walking distance	4409
santa monica	4156
miles br	4112
close to	3809
from the	3450

1-10 of 217,526 rows

Figure 9. Bigram with regular dataset

Description: df [13,572 × 2]	
bigram	n
<chr>	<int>
miles br	1421
los angeles	490
santa monica	394
beverly hills	328
walking distance	168
universal studios	158
br hollywood	152
venice beach	144
br santa	130
br los	115

1-10 of 13,572 rows

Figure 10. Bigram with exorbitant dataset

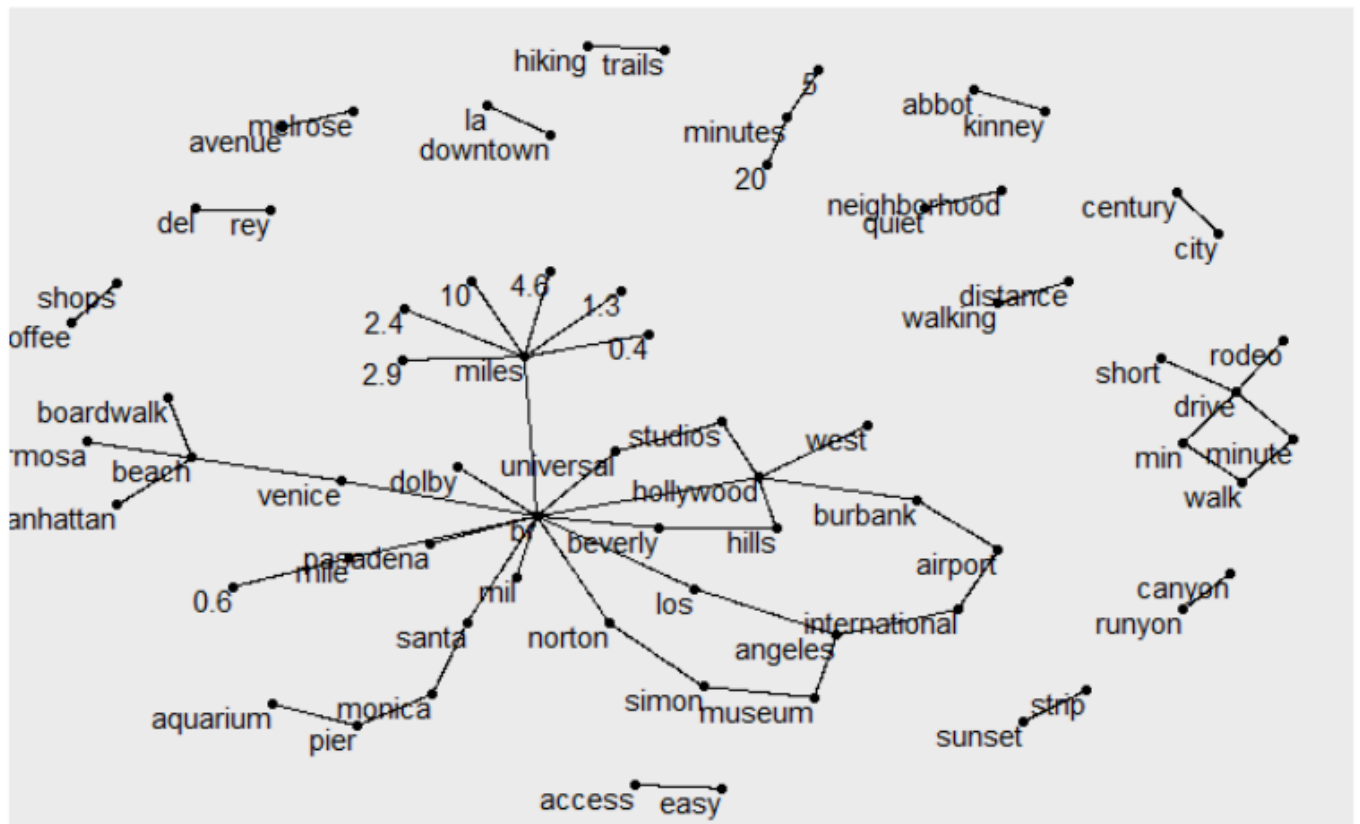
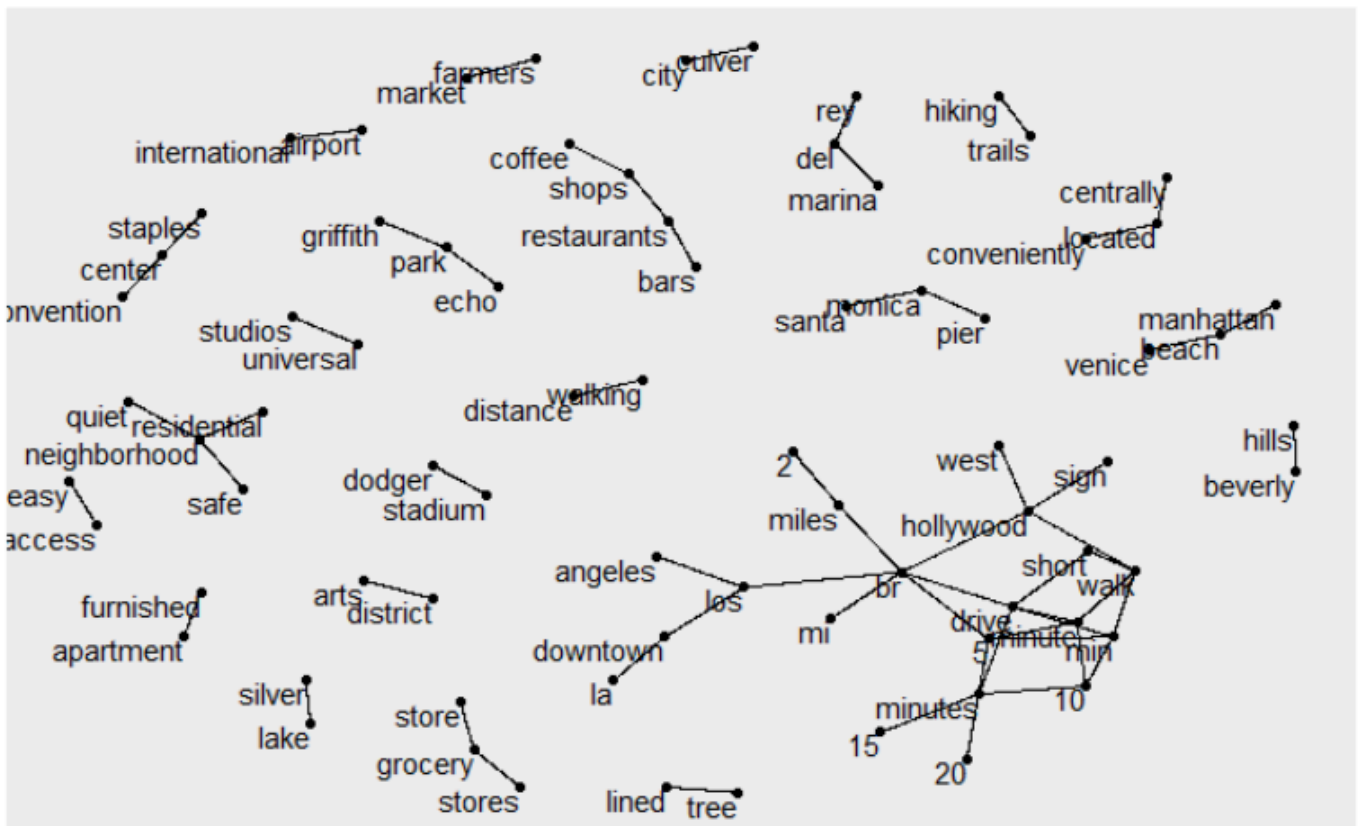


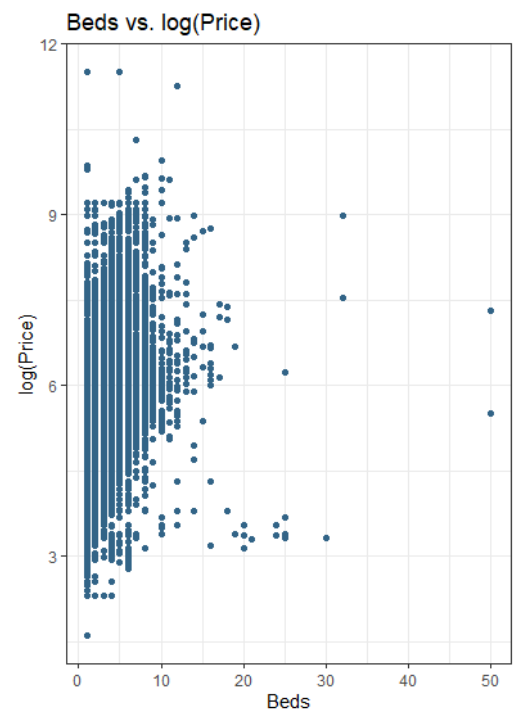
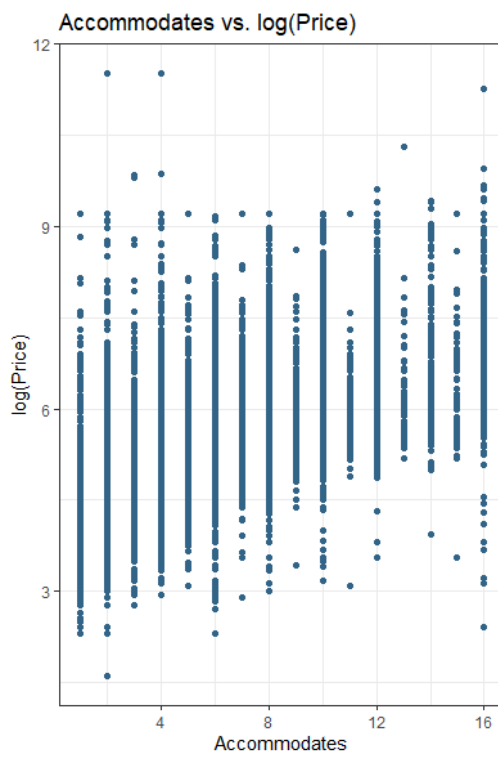
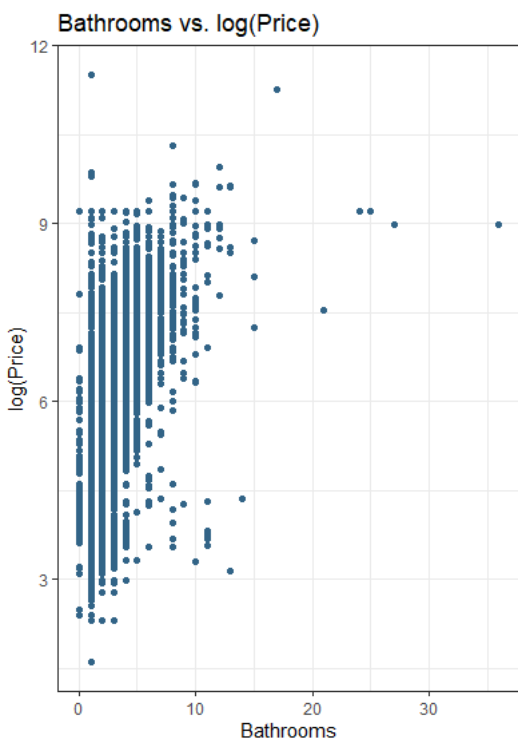
Figure 11. Strongly related words



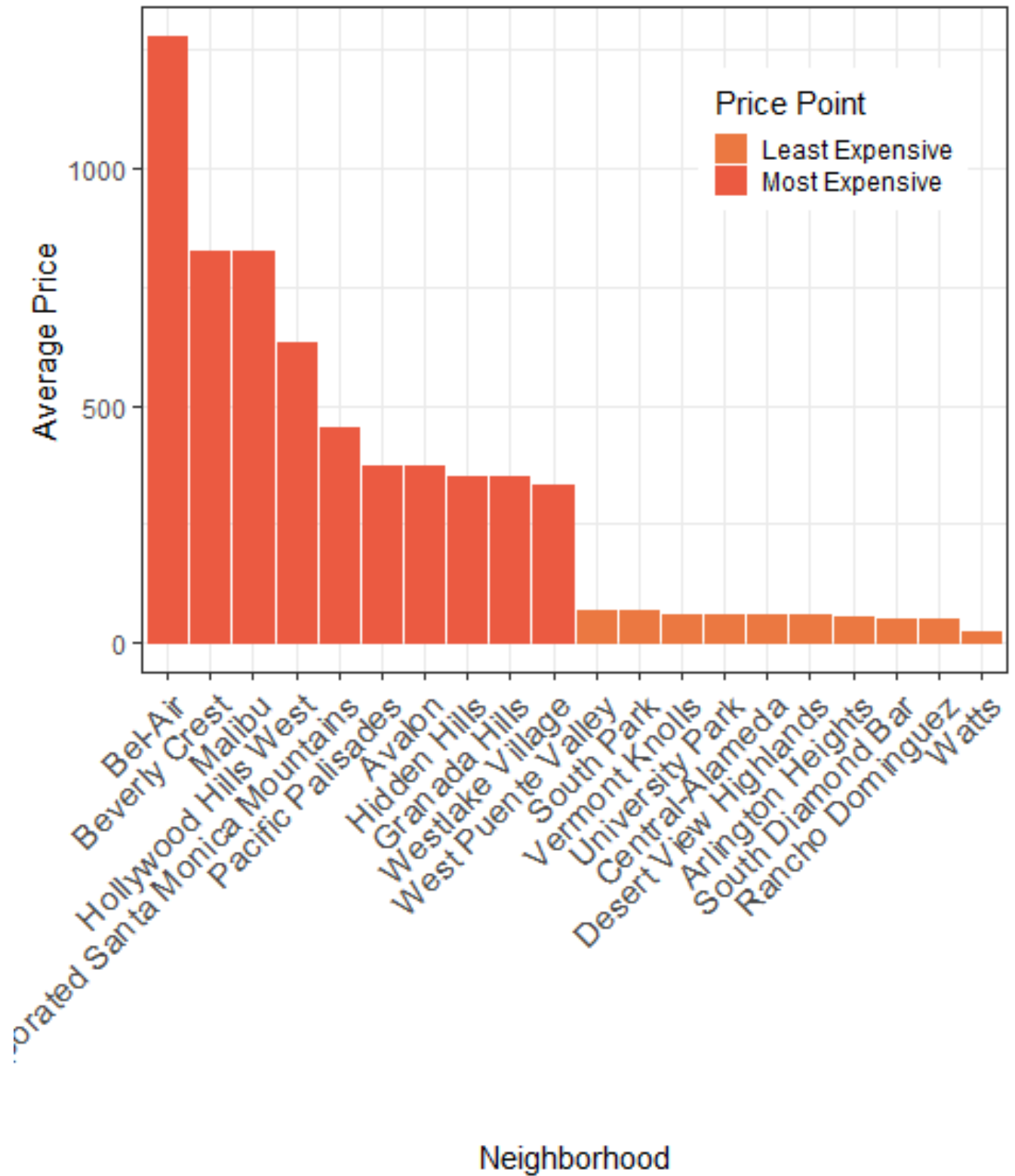
Figure 12. Untransformed price



Figure 13. Transformed price

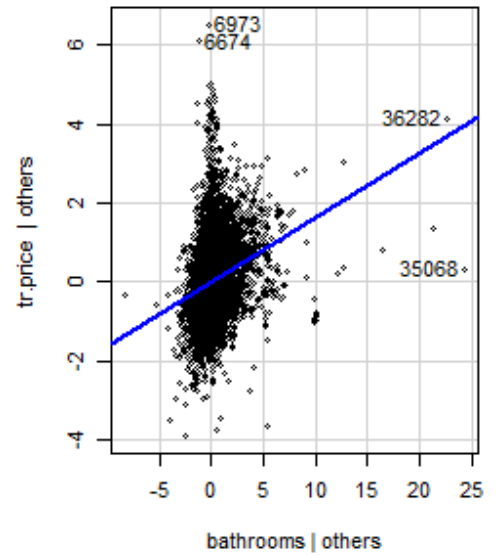
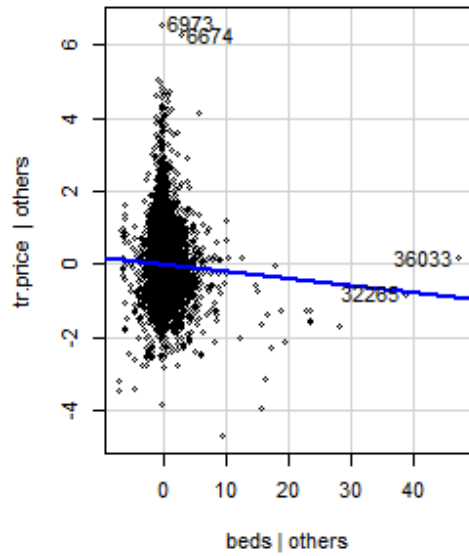
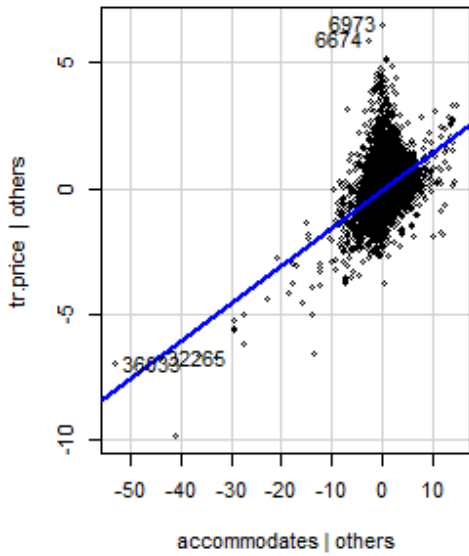


# Ten most expensive and least expensive Neighborhoods





Neighborhood	Booked Listings	Total Listings	Occupancy Rate
Sawtelle	207	416	0.4976
Burbank	244	569	0.4288
Rowland Heights	227	589	0.3854
Mid-Wilshire	208	553	0.3761
Glendale	302	824	0.3665
Pasadena	248	677	0.3663
Culver City	213	590	0.361
Hollywood	607	1696	0.3579
Venice	567	1611	0.352
Long Beach	586	1726	0.3395
Santa Monica	458	1384	0.3309
Hollywood Hills	244	759	0.3215
West Hollywood	414	1352	0.3062
Downtown	360	1313	0.2742



# Geospatial graphs

## Map of Airbnb's by Price and Location

