From Renting to Revenue: Tableau to Analysis Airbnb in New York

Created by:

Rishabh Sureshkumar Shah (CIN: 402026791) – rshah39@calstatela.edu

Hiraj Thakkar (CIN: 402652338) – hthakka4@calstatela.edu

1) *Introduction*

Airbnb, Inc is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. Based in San Francisco, California, the platform is accessible via website and mobile app. Airbnb does not own any of the listed properties; instead, it profits by receiving commission from each booking. The company was founded in 2008. Airbnb is a shortened version of its original name, AirBedandBreakfast.com.

The Airbnb Open Data dataset contains a wealth of information about Airbnb listings in major cities worldwide. This dataset contains data on various listing attributes such as host details, property type, price, availability, and review scores. The dataset can be used to investigate various aspects of Airbnb's business, such as pricing trends, customer preferences, and market saturation.

Motivation:

Analyzing the Airbnb Open Data dataset using Tableau can provide valuable insights for individuals and businesses alike. For instance, it can help property owners understand pricing trends in their area and optimize their prices accordingly. Similarly, it can help potential Airbnb hosts identify areas with high demand and optimize their listings to attract more customers. For businesses, this dataset can be used to understand customer preferences and market saturation, which can inform marketing strategies and help identify new market opportunities.

Overall, by using Tableau to analyze this dataset, individuals and businesses can gain a deeper understanding of the Airbnb market and make data-driven decisions to improve their business outcomes.

The dataset can also investigate the relationship between various factors and Airbnb booking rates. Researchers could, for example, examine the effect of location, listing type, and reviews on booking rates. The dataset can also be used to investigate the impact of external factors on Airbnb bookings, such as seasonality, events, and economic trends.

2) *Data Description*

Dataset Link:

[Airbnb Open Data | Kaggle]

Name of the field	Description of the field
Name	Name of the Airbnb listing.
Host_Id	Airbnb's unique identifier for the listing.
Host Identity Verified	Verified or Unconfirmed.
Host Name	Name of the host. Usually just the first name(s).
Neighbourhood group	The neighborhood group as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.
Neighborhood	The neighborhood as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.
Latitude	Uses the World Geodetic System (WGS84) projection for latitude and longitude.
Longitude	Uses the World Geodetic System (WGS84) projection for latitude and longitude.
Country	Name of the country (United States).
Country Code	US.
Instant Bookable	[t=true; f=false] Whether the guest can automatically book the listing without the host requiring accepting their booking request. An indicator of a commercial listing.
Cancellation Policy	Types of cancellation policy: strict, moderate, flexible.
Room Type	All homes are grouped into the following three-room types: Private room, Shared room, Entire place.
Construction Year	The year that listing was built.
Price	Daily price in local currency.
Service fee	Service fee of that listing.
Minimum Night	Minimum number of night stay for the listing (calendar rules may be different).
Number of Reviews	The number of reviews the listing has.
Last Review	The date of the last/newest review.

Reviews per Month	The number of reviews the listing has over the lifetime of the listing.
Review Rate Number	Reviews starting from 1 to 5.
Calculated host listing count	The number of listings the host has in the current scrape, in the city/region geography.
Availability 365	The availability of the listing x days in the future as determined by the calendar. Note a listing may not be available because it has been booked by a guest or blocked by the host.
House rules	House rules that listing has.

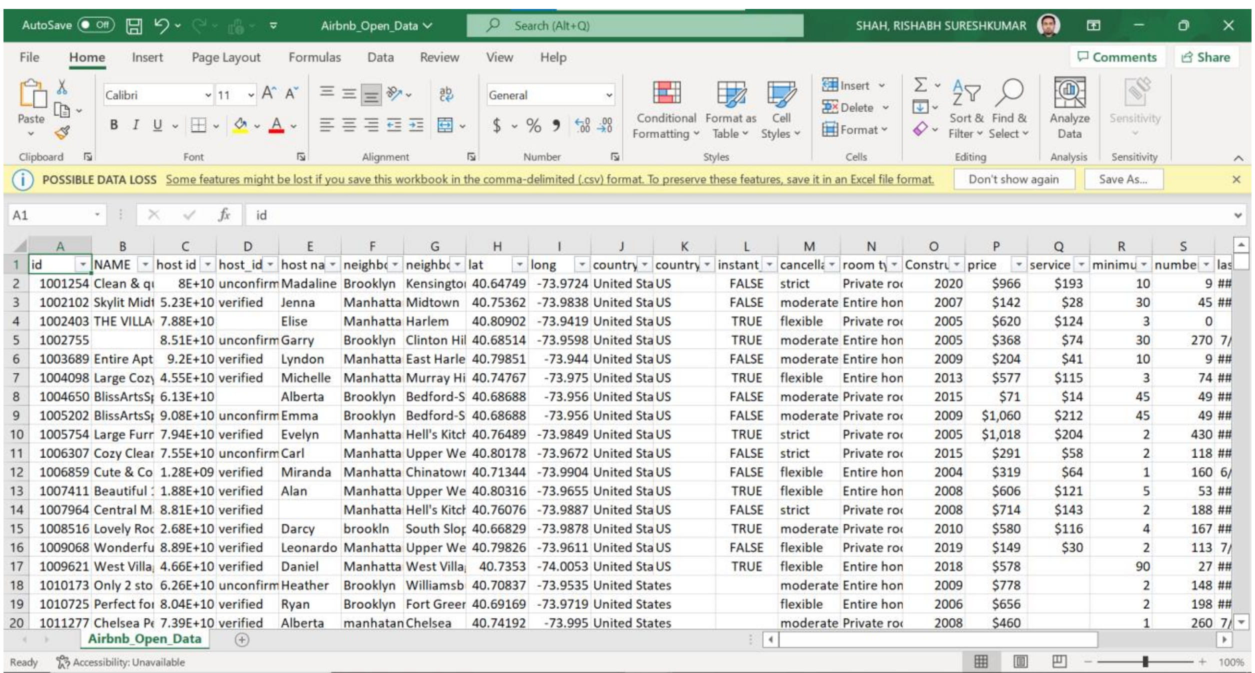
3) *Data Cleaning*

1. ***Deleting Unwanted Columns***:

   The "id" column most likely refers to a unique identifier for each Airbnb listing. The "license" column may refer to a license number or permit required for Airbnb hosts to legally rent out their properties. While this information may be important for regulatory or legal purposes, it may not be relevant to many other analyses.

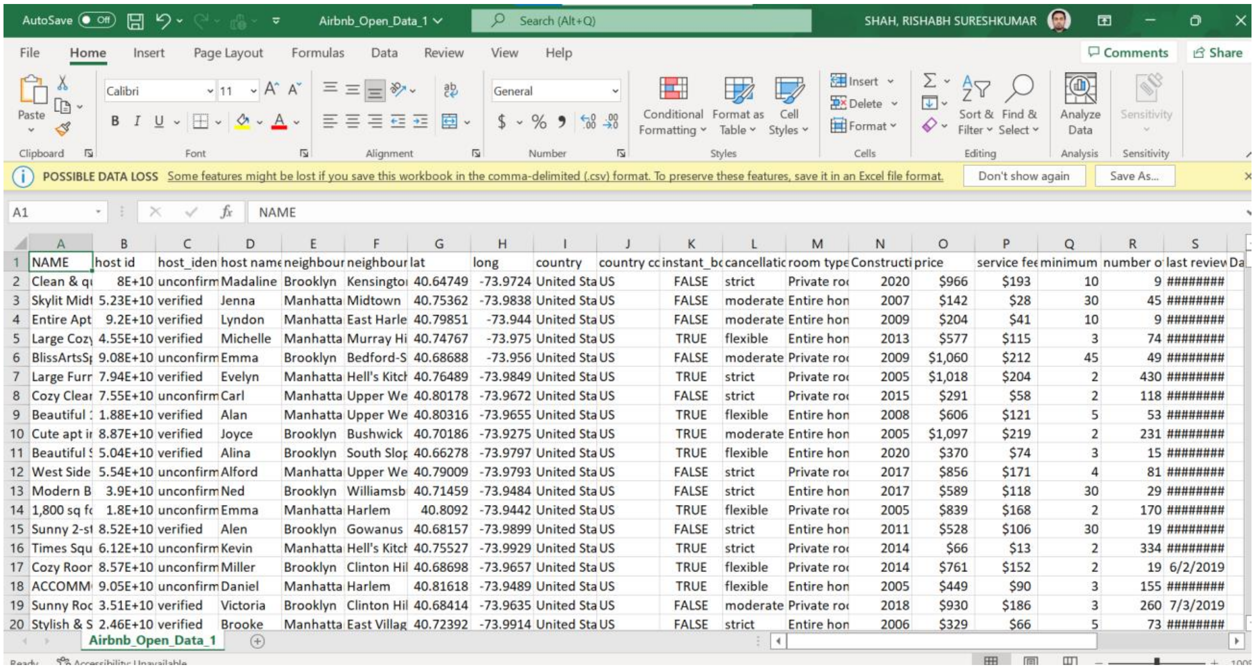
   Overall, deleting unwanted columns can help to simplify the dataset and reduce the amount of noise in the data, making it easier to analyze and draw meaningful insights.

Before:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	id	NAME	host id	host_id	host na	neighb	neighb	lat	long	country	country	instant_	cancell	room ty	Constru	price	service	minimu	numbe	las
2	1001254	Clean & q	8E+10	unconfirm	Madaline	Brooklyn	Kensingto	40.64749	-73.9724	United Sta	US	FALSE	strict	Private ro	2020	$966	$193	10	9	##
3	1002102	Skylit Mid	5.23E+10	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	##
4	1002403	THE VILLA	7.88E+10		Elise	Manhatta	Harlem	40.80902	-73.9419	United Sta	US	TRUE	flexible	Private ro	2005	$620	$124	3	0	
5	1002755		8.51E+10	unconfirm	Garry	Brooklyn	Clinton Hil	40.68514	-73.9598	United Sta	US	TRUE	moderate	Entire hon	2005	$368	$74	30	270	7/
6	1003689	Entire Apt	9.2E+10	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	##
7	1004098	Large Coz	4.55E+10	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	##
8	1004650	BlissArtsS	6.13E+10		Alberta	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private ro	2015	$71	$14	45	49	##
9	1005202	BlissArtsS	9.08E+10	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private ro	2009	$1,060	$212	45	49	##
10	1005754	Large Furr	7.94E+10	verified	Evelyn	Manhatta	Hell's Kitc	40.76489	-73.9849	United Sta	US	TRUE	strict	Private ro	2005	$1,018	$204	2	430	##
11	1006307	Cozy Clear	7.55E+10	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private ro	2015	$291	$58	2	118	##
12	1006859	Cute & Co	1.28E+09	verified	Miranda	Manhatta	Chinatow	40.71344	-73.9904	United Sta	US	FALSE	flexible	Entire hon	2004	$319	$64	1	160	6/
13	1007411	Beautiful	1.88E+10	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	##
14	1007964	Central M	8.81E+10	verified		Manhatta	Hell's Kitc	40.76076	-73.9887	United Sta	US	FALSE	strict	Private ro	2008	$714	$143	2	188	##
15	1008516	Lovely Ro	2.68E+10	verified	Darcy	brookln	South Slo	40.66829	-73.9878	United Sta	US	TRUE	moderate	Private ro	2010	$580	$116	4	167	##
16	1009068	Wonderfu	8.89E+10	verified	Leonardo	Manhatta	Upper We	40.79826	-73.9611	United Sta	US	FALSE	flexible	Private ro	2019	$149	$30	2	113	7/
17	1009621	West Villa	4.66E+10	verified	Daniel	Manhatta	West Villa	40.7353	-74.0053	United Sta	US	TRUE	flexible	Entire hon	2018	$578		90	27	##
18	1010173	Only 2 sto	6.26E+10	unconfirm	Heather	Brooklyn	Williamsb	40.70837	-73.9535	United States			moderate	Entire hon	2009	$778		2	148	##
19	1010725	Perfect fo	8.04E+10	verified	Ryan	Brooklyn	Fort Gree	40.69169	-73.9719	United States			flexible	Entire hon	2006	$656		2	198	##
20	1011277	Chelsea P	7.39E+10	verified	Alberta	manhatan	Chelsea	40.74192	-73.995	United States			moderate	Private ro	2008	$460		1	260	7/

After:



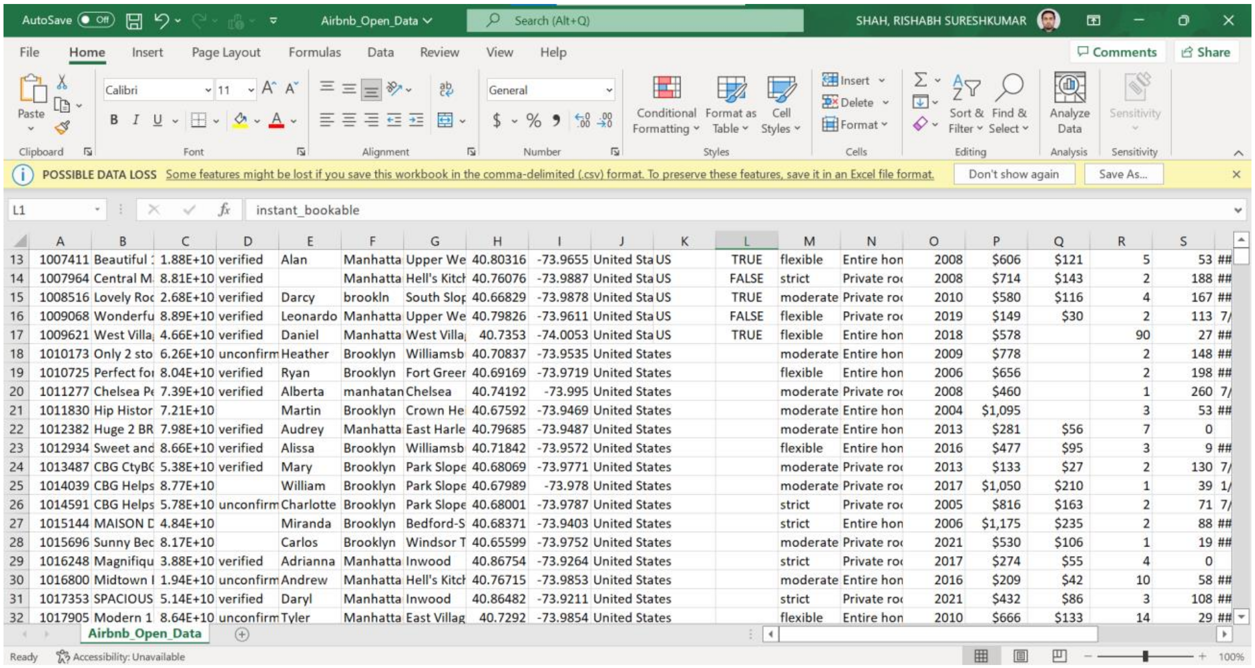
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	NAME	host id	host_iden	host name	neighbour	neighbour	lat	long	country	country cc	instant_bc	cancellatic	room type	Constructi	price	service fee	minimum	number o	last review	Da
2	Clean & qu	8E+10	unconfirm	Madaline	Brooklyn	Kensingtoi	40.64749	-73.9724	United Sta	US	FALSE	strict	Private roc	2020	$966	$193	10	9	########	
3	Skylit Midt	5.23E+10	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	########	
4	Entire Apt	9.2E+10	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	########	
5	Large Cozy	4.55E+10	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	########	
6	BlissArtsSp	9.08E+10	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private roc	2009	$1,060	$212	45	49	########	
7	Large Furr	7.94E+10	verified	Evelyn	Manhatta	Hell's Kitcł	40.76489	-73.9849	United Sta	US	TRUE	strict	Private roc	2005	$1,018	$204	2	430	########	
8	Cozy Clear	7.55E+10	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private roc	2015	$291	$58	2	118	########	
9	Beautiful 1	1.88E+10	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	########	
10	Cute apt ir	8.87E+10	verified	Joyce	Brooklyn	Bushwick	40.70186	-73.9275	United Sta	US	TRUE	moderate	Entire hon	2005	$1,097	$219	2	231	########	
11	Beautiful S	5.04E+10	verified	Alina	Brooklyn	South Slop	40.66278	-73.9797	United Sta	US	TRUE	flexible	Entire hon	2020	$370	$74	3	15	########	
12	West Side	5.54E+10	unconfirm	Alford	Manhatta	Upper We	40.79009	-73.9793	United Sta	US	FALSE	strict	Private roc	2017	$856	$171	4	81	########	
13	Modern B	3.9E+10	unconfirm	Ned	Brooklyn	Williamsb	40.71459	-73.9484	United Sta	US	FALSE	strict	Entire hon	2017	$589	$118	30	29	########	
14	1,800 sq fc	1.8E+10	unconfirm	Emma	Manhatta	Harlem	40.8092	-73.9442	United Sta	US	TRUE	flexible	Private roc	2005	$839	$168	2	170	########	
15	Sunny 2-st	8.52E+10	verified	Alen	Brooklyn	Gowanus	40.68157	-73.9899	United Sta	US	FALSE	strict	Entire hon	2011	$528	$106	30	19	########	
16	Times Squ	6.12E+10	unconfirm	Kevin	Manhatta	Hell's Kitcł	40.75527	-73.9929	United Sta	US	TRUE	strict	Private roc	2014	$66	$13	2	334	########	
17	Cozy Roor	8.57E+10	unconfirm	Miller	Brooklyn	Clinton Hil	40.68698	-73.9657	United Sta	US	TRUE	flexible	Private roc	2014	$761	$152	2	19	6/2/2019	
18	ACCOMM	9.05E+10	unconfirm	Daniel	Manhatta	Harlem	40.81618	-73.9489	United Sta	US	TRUE	flexible	Entire hon	2005	$449	$90	3	155	########	
19	Sunny Roc	3.51E+10	verified	Victoria	Brooklyn	Clinton Hil	40.68414	-73.9635	United Sta	US	FALSE	moderate	Private roc	2018	$930	$186	3	260	7/3/2019	
20	Stylish & S	2.46E+10	verified	Brooke	Manhatta	East Villag	40.72392	-73.9914	United Sta	US	FALSE	strict	Entire hon	2006	$329	$66	5	73	########	

2. ***Removed missing/null values:***

Only full or null values in a dataset can lead to accurate analyses and results. Therefore, it is vital to identify and remove these values before performing any analysis. In Excel, you can check for missing values in a column using the filter function and delete any rows with missing values.

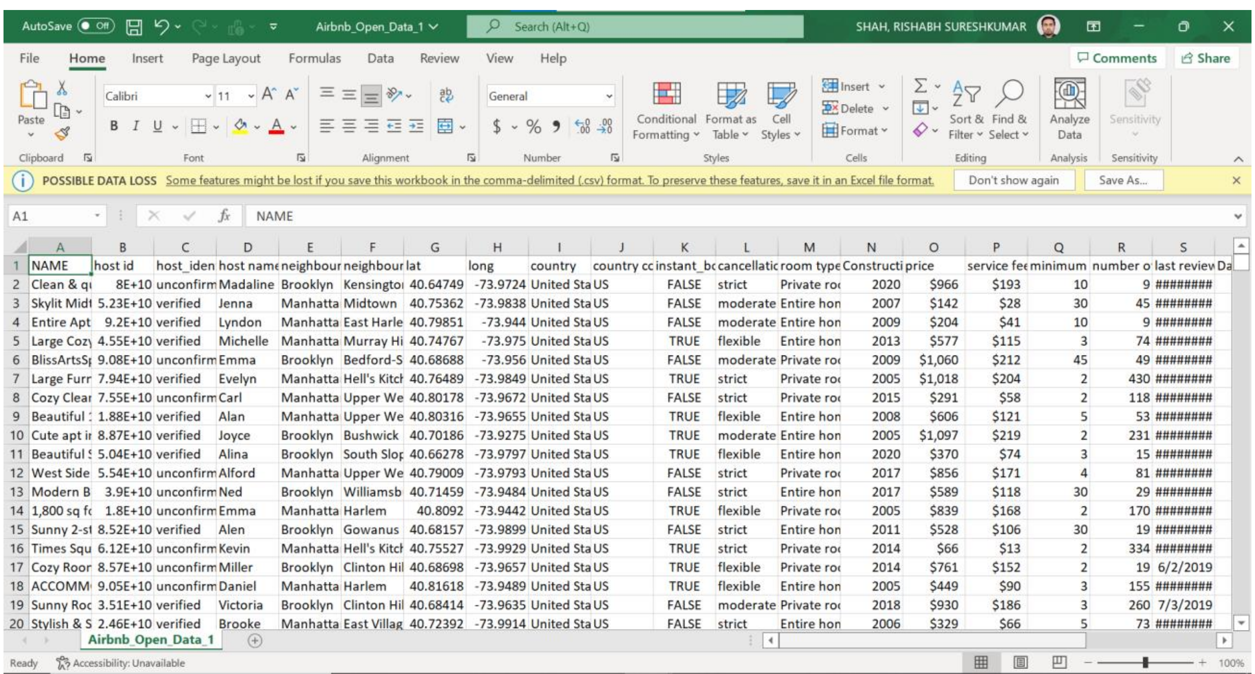
Furthermore, we checked if there were any missing values in the dataset by using the filter function to check for blank cells in each column. I clicked on the column header, then clicked on the filter button. I unchecked the "Select All" option in the dropdown list, then checked the "Blanks" option. Excel showed us all the blank cells in that column. Then delete all rows for those cells.

Before:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
13	1007411	Beautiful :	1.88E+10	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	##
14	1007964	Central M.	8.81E+10	verified		Manhatta	Hell's Kitcl	40.76076	-73.9887	United Sta	US	FALSE	strict	Private ro	2008	$714	$143	2	188	##
15	1008516	Lovely Roc	2.68E+10	verified	Darcy	brookln	South Slop	40.66829	-73.9878	United Sta	US	TRUE	moderate	Private ro	2010	$580	$116	4	167	##
16	1009068	Wonderfu	8.89E+10	verified	Leonardo	Manhatta	Upper We	40.79826	-73.9611	United Sta	US	FALSE	flexible	Private ro	2019	$149	$30	2	113	7/
17	1009621	West Villa	4.66E+10	verified	Daniel	Manhatta	West Villa	40.7353	-74.0053	United Sta	US	TRUE	flexible	Entire hon	2018	$578		90	27	##
18	1010173	Only 2 sto	6.26E+10	unconfirm	Heather	Brooklyn	Williamsb	40.70837	-73.9535	United States			moderate	Entire hon	2009	$778		2	148	##
19	1010725	Perfect for	8.04E+10	verified	Ryan	Brooklyn	Fort Greer	40.69169	-73.9719	United States			flexible	Entire hon	2006	$656		2	198	##
20	1011277	Chelsea Pe	7.39E+10	verified	Alberta	manhatan	Chelsea	40.74192	-73.995	United States			moderate	Private ro	2008	$460		1	260	7/
21	1011830	Hip Histor	7.21E+10		Martin	Brooklyn	Crown He	40.67592	-73.9469	United States			moderate	Entire hon	2004	$1,095		3	53	##
22	1012382	Huge 2 BR	7.98E+10	verified	Audrey	Manhatta	East Harle	40.79685	-73.9487	United States			moderate	Entire hon	2013	$281	$56	7	0	
23	1012934	Sweet and	8.66E+10	verified	Alissa	Brooklyn	Williamsb	40.71842	-73.9572	United States			flexible	Entire hon	2016	$477	$95	3	9	##
24	1013487	CBG CtyBC	5.38E+10	verified	Mary	Brooklyn	Park Slope	40.68069	-73.9771	United States			moderate	Private ro	2013	$133	$27	2	130	7/
25	1014039	CBG Helps	8.77E+10		William	Brooklyn	Park Slope	40.67989	-73.978	United States			moderate	Private ro	2017	$1,050	$210	1	39	1/
26	1014591	CBG Helps	5.78E+10	unconfirm	Charlotte	Brooklyn	Park Slope	40.68001	-73.9787	United States			strict	Private ro	2005	$816	$163	2	71	7/
27	1015144	MAISON D	4.84E+10		Miranda	Brooklyn	Bedford-S	40.68371	-73.9403	United States			strict	Entire hon	2006	$1,175	$235	2	88	##
28	1015696	Sunny Bec	8.17E+10		Carlos	Brooklyn	Windsor T	40.65599	-73.9752	United States			moderate	Private ro	2021	$530	$106	1	19	##
29	1016248	Magnifiqu	3.88E+10	verified	Adrianna	Manhatta	Inwood	40.86754	-73.9264	United States			strict	Private ro	2017	$274	$55	4	0	
30	1016800	Midtown I	1.94E+10	unconfirm	Andrew	Manhatta	Hell's Kitcl	40.76715	-73.9853	United States			moderate	Entire hon	2016	$209	$42	10	58	##
31	1017353	SPACIOUS	5.14E+10	verified	Daryl	Manhatta	Inwood	40.86482	-73.9211	United States			strict	Private ro	2021	$432	$86	3	108	##
32	1017905	Modern 1	8.64E+10	unconfirm	Tyler	Manhatta	East Villag	40.7292	-73.9854	United States			flexible	Entire hon	2010	$666	$133	14	29	##

After:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	NAME	host id	host_iden	host name	neighbour	neighbour	lat	long	country	country cc	instant_bc	cancellatic	room type	Constructi	price	service fee	minimum	number o	last review	Da
2	Clean & qu	8E+10	unconfirm	Madaline	Brooklyn	Kensingto	40.64749	-73.9724	United Sta	US	FALSE	strict	Private ro	2020	$966	$193	10	9	########	
3	Skylit Midt	5.23E+10	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	########	
4	Entire Apt	9.2E+10	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	########	
5	Large Cozy	4.55E+10	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	########	
6	BlissArtsSp	9.08E+10	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private ro	2009	$1,060	$212	45	49	########	
7	Large Furr	7.94E+10	verified	Evelyn	Manhatta	Hell's Kitcl	40.76489	-73.9849	United Sta	US	TRUE	strict	Private ro	2005	$1,018	$204	2	430	########	
8	Cozy Clear	7.55E+10	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private ro	2015	$291	$58	2	118	########	
9	Beautiful :	1.88E+10	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	########	
10	Cute apt ir	8.87E+10	verified	Joyce	Brooklyn	Bushwick	40.70186	-73.9275	United Sta	US	TRUE	moderate	Entire hon	2005	$1,097	$219	2	231	########	
11	Beautiful S	5.04E+10	verified	Alina	Brooklyn	South Slop	40.66278	-73.9797	United Sta	US	TRUE	flexible	Entire hon	2020	$370	$74	3	15	########	
12	West Side	5.54E+10	unconfirm	Alford	Manhatta	Upper We	40.79009	-73.9793	United Sta	US	FALSE	strict	Private ro	2017	$856	$171	4	81	########	
13	Modern B	3.9E+10	unconfirm	Ned	Brooklyn	Williamsb	40.71459	-73.9484	United Sta	US	FALSE	strict	Entire hon	2017	$589	$118	30	29	########	
14	1,800 sq fo	1.8E+10	unconfirm	Emma	Manhatta	Harlem	40.8092	-73.9442	United Sta	US	TRUE	flexible	Private ro	2005	$839	$168	2	170	########	
15	Sunny 2-st	8.52E+10	verified	Alen	Brooklyn	Gowanus	40.68157	-73.9899	United Sta	US	FALSE	strict	Entire hon	2011	$528	$106	30	19	########	
16	Times Squ	6.12E+10	unconfirm	Kevin	Manhatta	Hell's Kitcl	40.75527	-73.9929	United Sta	US	TRUE	strict	Private ro	2014	$66	$13	2	334	########	
17	Cozy Roor	8.57E+10	unconfirm	Miller	Brooklyn	Clinton Hil	40.68698	-73.9657	United Sta	US	TRUE	flexible	Private ro	2014	$761	$152	2	19	6/2/2019	
18	ACCOMM	9.05E+10	unconfirm	Daniel	Manhatta	Harlem	40.81618	-73.9489	United Sta	US	TRUE	flexible	Entire hon	2005	$449	$90	3	155	########	
19	Sunny Roc	3.51E+10	verified	Victoria	Brooklyn	Clinton Hil	40.68414	-73.9635	United Sta	US	FALSE	moderate	Private ro	2018	$930	$186	3	260	7/3/2019	
20	Stylish & S	2.46E+10	verified	Brooke	Manhatta	East Villag	40.72392	-73.9914	United Sta	US	FALSE	strict	Entire hon	2006	$329	$66	5	73	########	

3. ***Splitting columns:***

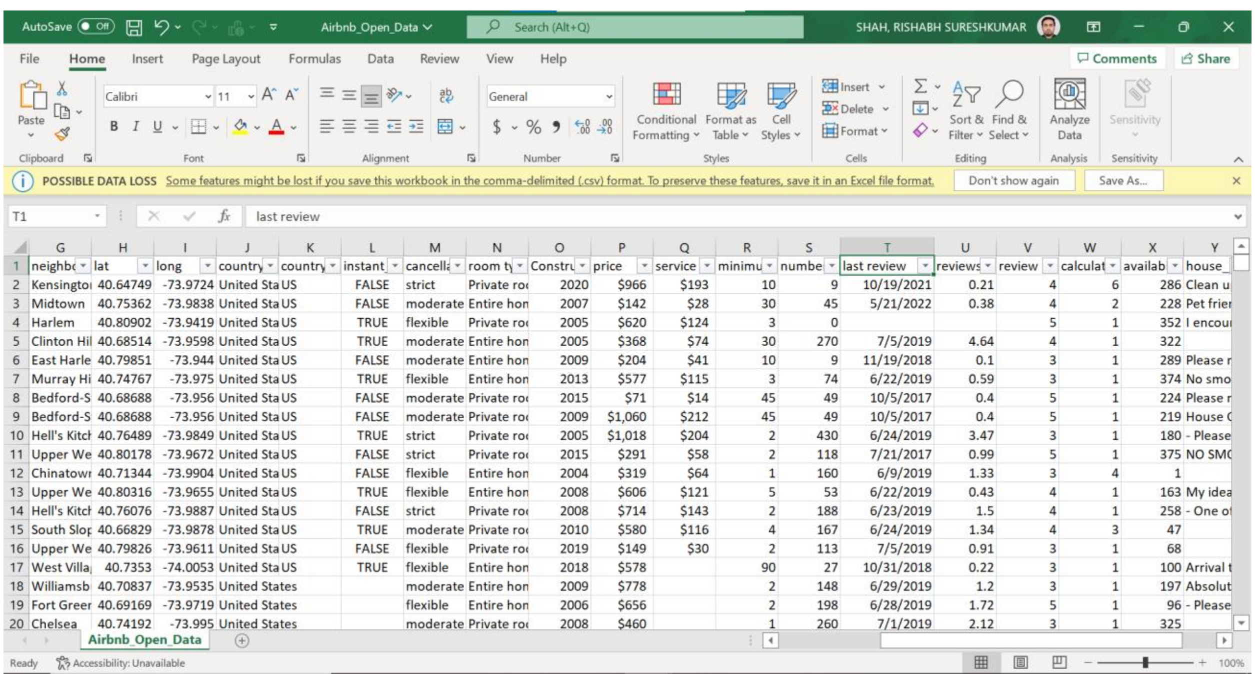
Separate Day, Month, and Year

Inserted 3 new column T, U, and V and named as Day, Month, and Year respectively.

- To extract the day, at T2 cell, applied "= DAY(S2)" and applied for all cells in this column.
- To extract the month, at U2 cell, applied "= MONTH(S2)" and applied for all cells in this column.
- To extract the year, at V2 cell, applied "= YEAR(S2)" and applied for all cells in this column.

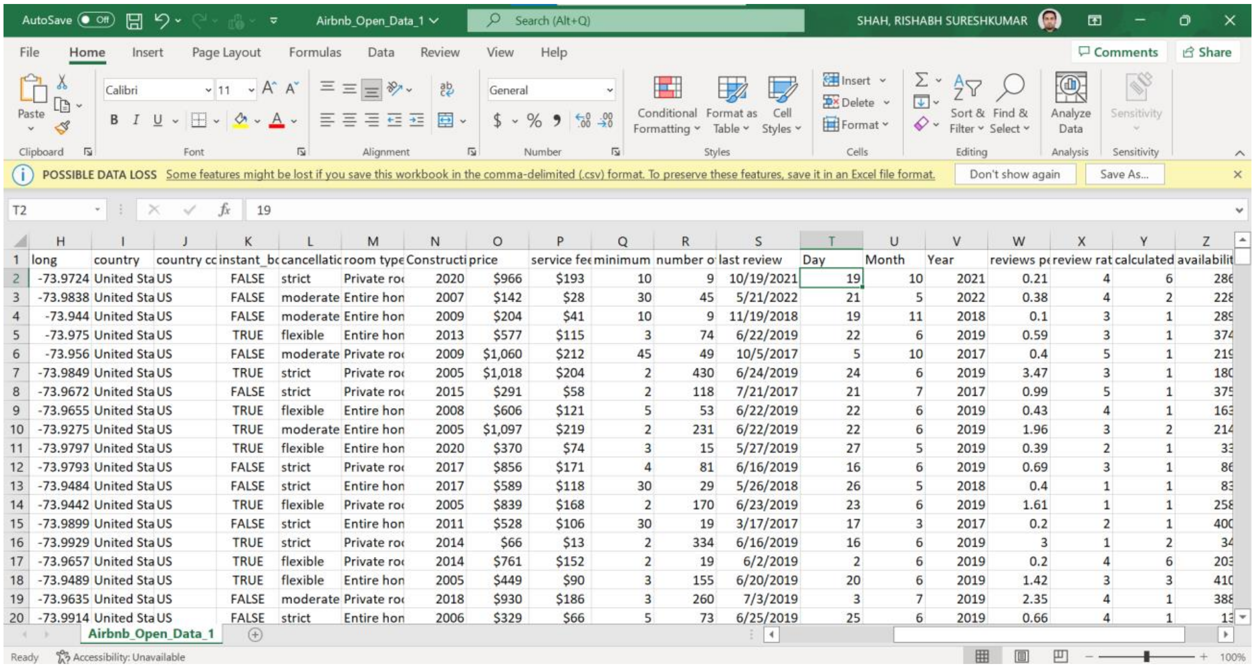
By splitting the original date column into three separate columns, it becomes easier to work with and analyze the data in a more efficient manner.

Before:



	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	neighbc	lat	long	country	country	instant_	cancell	room ty	Constru	price	service	minimu	numbe	last review	reviews	review	calculat	availab	house_
2	Kensingto	40.64749	-73.9724	United Sta	US	FALSE	strict	Private ro	2020	$966	$193	10	9	10/19/2021	0.21	4	6	286	Clean u
3	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	5/21/2022	0.38	4	2	228	Pet frier
4	Harlem	40.80902	-73.9419	United Sta	US	TRUE	flexible	Private ro	2005	$620	$124	3	0			5	1	352	I encou
5	Clinton Hil	40.68514	-73.9598	United Sta	US	TRUE	moderate	Entire hon	2005	$368	$74	30	270	7/5/2019	4.64	4	1	322	
6	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	11/19/2018	0.1	3	1	289	Please r
7	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	6/22/2019	0.59	3	1	374	No smo
8	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private ro	2015	$71	$14	45	49	10/5/2017	0.4	5	1	224	Please r
9	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private ro	2009	$1,060	$212	45	49	10/5/2017	0.4	5	1	219	House C
10	Hell's Kitc	40.76489	-73.9849	United Sta	US	TRUE	strict	Private ro	2005	$1,018	$204	2	430	6/24/2019	3.47	3	1	180	- Please
11	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private ro	2015	$291	$58	2	118	7/21/2017	0.99	5	1	375	NO SMC
12	Chinatowr	40.71344	-73.9904	United Sta	US	FALSE	flexible	Entire hon	2004	$319	$64	1	160	6/9/2019	1.33	3	4	1	
13	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	6/22/2019	0.43	4	1	163	My idea
14	Hell's Kitc	40.76076	-73.9887	United Sta	US	FALSE	strict	Private ro	2008	$714	$143	2	188	6/23/2019	1.5	4	1	258	- One of
15	South Slo	40.66829	-73.9878	United Sta	US	TRUE	moderate	Private ro	2010	$580	$116	4	167	6/24/2019	1.34	4	3	47	
16	Upper We	40.79826	-73.9611	United Sta	US	FALSE	flexible	Private ro	2019	$149	$30	2	113	7/5/2019	0.91	3	1	68	
17	West Villa	40.7353	-74.0053	United Sta	US	TRUE	flexible	Entire hon	2018	$578		90	27	10/31/2018	0.22	3	1	100	Arrival t
18	Williamsb	40.70837	-73.9535	United States			moderate	Entire hon	2009	$778		2	148	6/29/2019	1.2	3	1	197	Absolut
19	Fort Greer	40.69169	-73.9719	United States			flexible	Entire hon	2006	$656		2	198	6/28/2019	1.72	5	1	96	- Please
20	Chelsea	40.74192	-73.995	United States			moderate	Private ro	2008	$460		1	260	7/1/2019	2.12	3	1	325	

After:



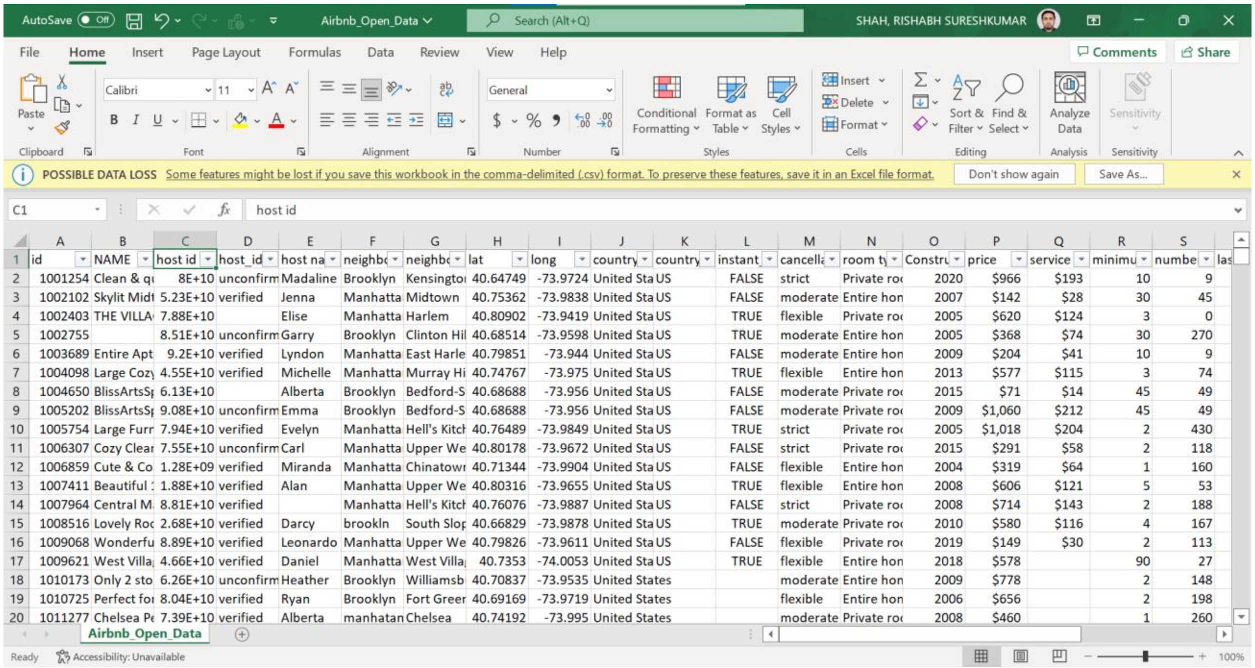
	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	long	country	country cc	instant_bc	cancellatic	room type	Constructi	price	service fee	minimum	number o	last review	Day	Month	Year	reviews pe	review rat	calculated	availabilit
2	-73.9724	United Sta	US	FALSE	strict	Private ro	2020	$966	$193	10	9	10/19/2021	19	10	2021	0.21	4	6	286
3	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	5/21/2022	21	5	2022	0.38	4	2	228
4	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	11/19/2018	19	11	2018	0.1	3	1	289
5	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	6/22/2019	22	6	2019	0.59	3	1	374
6	-73.956	United Sta	US	FALSE	moderate	Private ro	2009	$1,060	$212	45	49	10/5/2017	5	10	2017	0.4	5	1	219
7	-73.9849	United Sta	US	TRUE	strict	Private ro	2005	$1,018	$204	2	430	6/24/2019	24	6	2019	3.47	3	1	180
8	-73.9672	United Sta	US	FALSE	strict	Private ro	2015	$291	$58	2	118	7/21/2017	21	7	2017	0.99	5	1	375
9	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	6/22/2019	22	6	2019	0.43	4	1	163
10	-73.9275	United Sta	US	TRUE	moderate	Entire hon	2005	$1,097	$219	2	231	6/22/2019	22	6	2019	1.96	3	2	214
11	-73.9797	United Sta	US	TRUE	flexible	Entire hon	2020	$370	$74	3	15	5/27/2019	27	5	2019	0.39	2	1	33
12	-73.9793	United Sta	US	FALSE	strict	Private ro	2017	$856	$171	4	81	6/16/2019	16	6	2019	0.69	3	1	86
13	-73.9484	United Sta	US	FALSE	strict	Entire hon	2017	$589	$118	30	29	5/26/2018	26	5	2018	0.4	1	1	83
14	-73.9442	United Sta	US	TRUE	flexible	Private ro	2005	$839	$168	2	170	6/23/2019	23	6	2019	1.61	1	1	258
15	-73.9899	United Sta	US	FALSE	strict	Entire hon	2011	$528	$106	30	19	3/17/2017	17	3	2017	0.2	2	1	400
16	-73.9929	United Sta	US	TRUE	strict	Private ro	2014	$66	$13	2	334	6/16/2019	16	6	2019	3	1	2	34
17	-73.9657	United Sta	US	TRUE	flexible	Private ro	2014	$761	$152	2	19	6/2/2019	2	6	2019	0.2	4	6	203
18	-73.9489	United Sta	US	TRUE	flexible	Entire hon	2005	$449	$90	3	155	6/20/2019	20	6	2019	1.42	3	3	410
19	-73.9635	United Sta	US	FALSE	moderate	Private ro	2018	$930	$186	3	260	7/3/2019	3	7	2019	2.35	4	1	388
20	-73.9914	United Sta	US	FALSE	strict	Entire hon	2006	$329	$66	5	73	6/25/2019	25	6	2019	0.66	4	1	13

4. ***Resize the column width:***

Resizing column width is a simple and quick technique that can significantly improve a dataset's readability and usability. By ensuring that all data is visible in the spreadsheet, analysts can work with the data more easily, and errors caused by incomplete, or cut-off data are reduced.

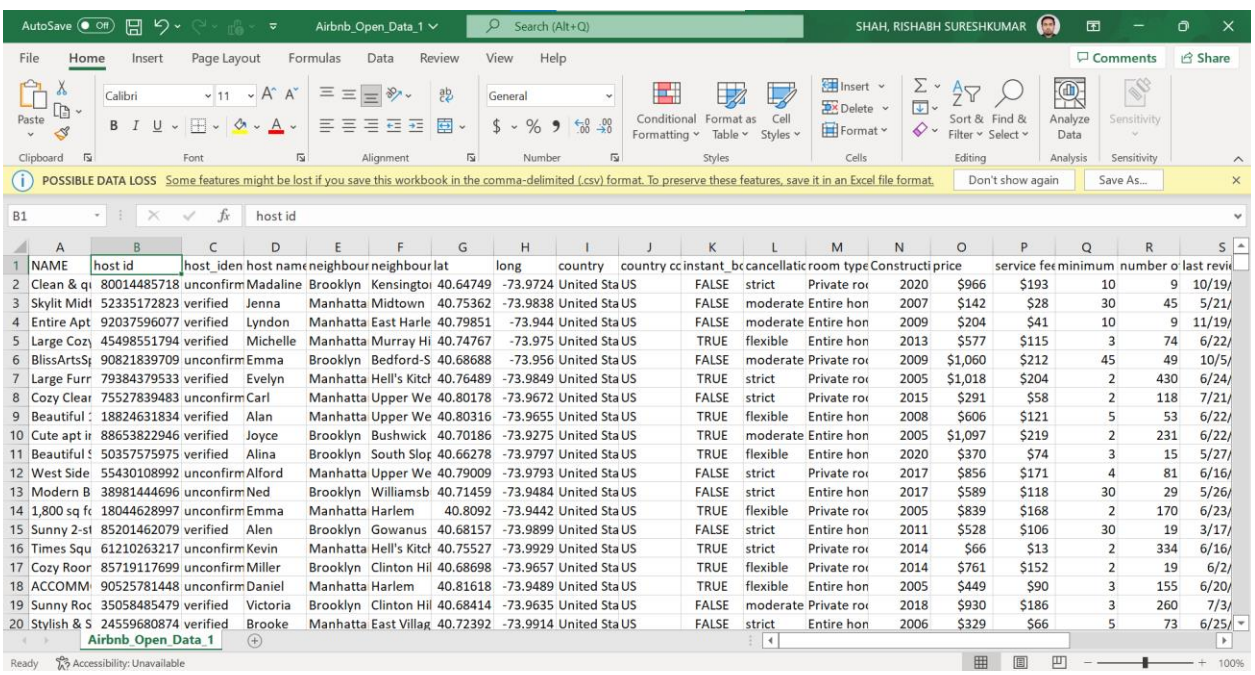
The column width of the "host id" column was increased in the given example to ensure that the entire value of the host ID is visible in the cell. This is significant because omitting the host ID may result in errors when analyzing or calculating the data.

Before:



1	id	NAME	host id	host_id	host na	neighb	neighb	lat	long	country	country	instant_	cancella	room ty	Constru	price	service	minimu	numbe	las
2	1001254	Clean & qu	8E+10	unconfirm	Madaline	Brooklyn	Kensingto	40.64749	-73.9724	United Sta	US	FALSE	strict	Private roo	2020	$966	$193	10	9	
3	1002102	Skylit Midt	5.23E+10	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	
4	1002403	THE VILLA	7.88E+10		Elise	Manhatta	Harlem	40.80902	-73.9419	United Sta	US	TRUE	flexible	Private roo	2005	$620	$124	3	0	
5	1002755		8.51E+10	unconfirm	Garry	Brooklyn	Clinton Hil	40.68514	-73.9598	United Sta	US	TRUE	moderate	Entire hon	2005	$368	$74	30	270	
6	1003689	Entire Apt	9.2E+10	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	
7	1004098	Large Cozy	4.55E+10	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	
8	1004650	BlissArtsSp	6.13E+10		Alberta	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private roo	2015	$71	$14	45	49	
9	1005202	BlissArtsSp	9.08E+10	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private roo	2009	$1,060	$212	45	49	
10	1005754	Large Furr	7.94E+10	verified	Evelyn	Manhatta	Hell's Kitch	40.76489	-73.9849	United Sta	US	TRUE	strict	Private roo	2005	$1,018	$204	2	430	
11	1006307	Cozy Clear	7.55E+10	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private roo	2015	$291	$58	2	118	
12	1006859	Cute & Co	1.28E+09	verified	Miranda	Manhatta	Chinatown	40.71344	-73.9904	United Sta	US	FALSE	flexible	Entire hon	2004	$319	$64	1	160	
13	1007411	Beautiful 1	1.88E+10	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	
14	1007964	Central M.	8.81E+10	verified		Manhatta	Hell's Kitch	40.76076	-73.9887	United Sta	US	FALSE	strict	Private roo	2008	$714	$143	2	188	
15	1008516	Lovely Roo	2.68E+10	verified	Darcy	brookln	South Slop	40.66829	-73.9878	United Sta	US	TRUE	moderate	Private roo	2010	$580	$116	4	167	
16	1009068	Wonderfu	8.89E+10	verified	Leonardo	Manhatta	Upper We	40.79826	-73.9611	United Sta	US	FALSE	flexible	Private roo	2019	$149	$30	2	113	
17	1009621	West Villag	4.66E+10	verified	Daniel	Manhatta	West Villag	40.7353	-74.0053	United Sta	US	TRUE	flexible	Entire hon	2018	$578		90	27	
18	1010173	Only 2 sto	6.26E+10	unconfirm	Heather	Brooklyn	Williamsb	40.70837	-73.9535	United States			moderate	Entire hon	2009	$778		2	148	
19	1010725	Perfect for	8.04E+10	verified	Ryan	Brooklyn	Fort Greer	40.69169	-73.9719	United States			flexible	Entire hon	2006	$656		2	198	
20	1011277	Chelsea Pe	7.39E+10	verified	Alberta	manhatan	Chelsea	40.74192	-73.995	United States			moderate	Private roo	2008	$460		1	260	

After:



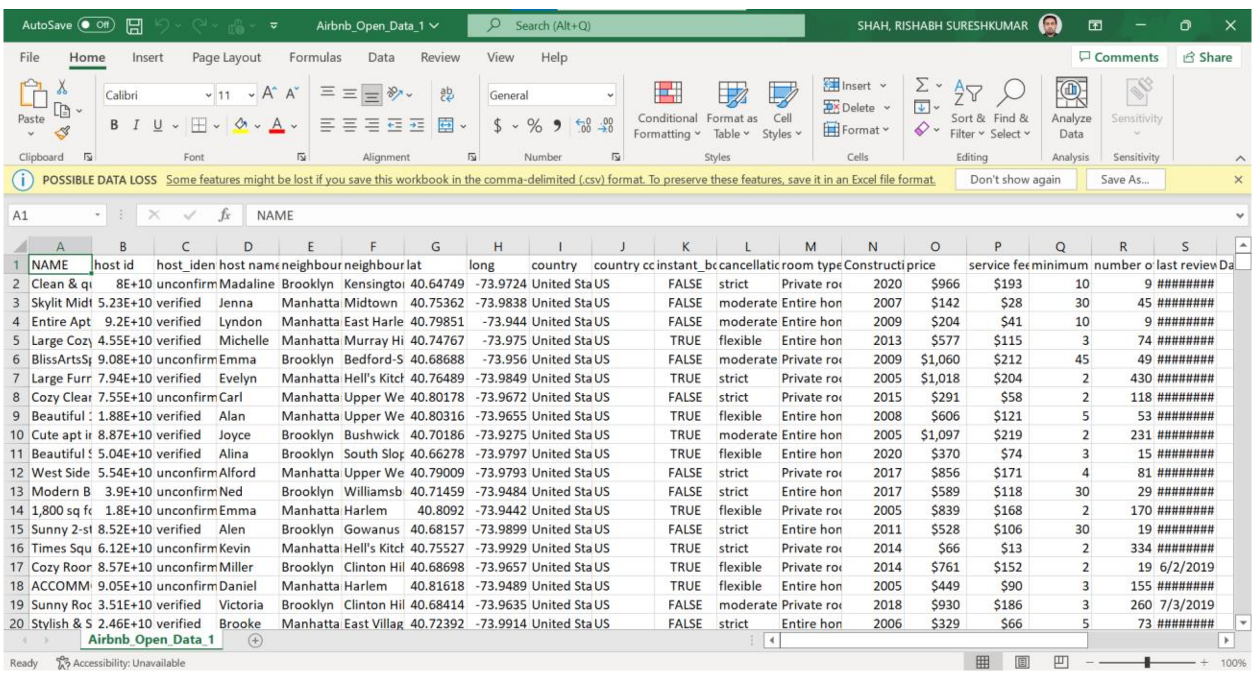
1	NAME	host id	host_iden	host name	neighbour	neighbour	lat	long	country	country cc	instant_bc	cancellatic	room type	Constructi	price	service fee	minimum	number o	last revie
2	Clean & qu	80014485718	unconfirm	Madaline	Brooklyn	Kensingto	40.64749	-73.9724	United Sta	US	FALSE	strict	Private roo	2020	$966	$193	10	9	10/19/
3	Skylit Midt	52335172823	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	5/21/
4	Entire Apt	92037596077	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	11/19/
5	Large Cozy	45498551794	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	6/22/
6	BlissArtsSp	90821839709	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private roo	2009	$1,060	$212	45	49	10/5/
7	Large Furr	79384379533	verified	Evelyn	Manhatta	Hell's Kitch	40.76489	-73.9849	United Sta	US	TRUE	strict	Private roo	2005	$1,018	$204	2	430	6/24/
8	Cozy Clear	75527839483	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private roo	2015	$291	$58	2	118	7/21/
9	Beautiful 1	18824631834	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	6/22/
10	Cute apt ir	88653822946	verified	Joyce	Brooklyn	Bushwick	40.70186	-73.9275	United Sta	US	TRUE	moderate	Entire hon	2005	$1,097	$219	2	231	6/22/
11	Beautiful S	50357575975	verified	Alina	Brooklyn	South Slop	40.66278	-73.9797	United Sta	US	TRUE	flexible	Entire hon	2020	$370	$74	3	15	5/27/
12	West Side	55430108992	unconfirm	Alford	Manhatta	Upper We	40.79009	-73.9793	United Sta	US	FALSE	strict	Private roo	2017	$856	$171	4	81	6/16/
13	Modern B	38981444696	unconfirm	Ned	Brooklyn	Williamsb	40.71459	-73.9484	United Sta	US	FALSE	strict	Entire hon	2017	$589	$118	30	29	5/26/
14	1,800 sq fo	18044628997	unconfirm	Emma	Manhatta	Harlem	40.8092	-73.9442	United Sta	US	TRUE	flexible	Private roo	2005	$839	$168	2	170	6/23/
15	Sunny 2-st	85201462079	verified	Alen	Brooklyn	Gowanus	40.68157	-73.9899	United Sta	US	FALSE	strict	Entire hon	2011	$528	$106	30	19	3/17/
16	Times Squ	61210263217	unconfirm	Kevin	Manhatta	Hell's Kitch	40.75527	-73.9929	United Sta	US	TRUE	strict	Private roo	2014	$66	$13	2	334	6/16/
17	Cozy Roor	85719117699	unconfirm	Miller	Brooklyn	Clinton Hil	40.68698	-73.9657	United Sta	US	TRUE	flexible	Private roo	2014	$761	$152	2	19	6/2/
18	ACCOMM	90525781448	unconfirm	Daniel	Manhatta	Harlem	40.81618	-73.9489	United Sta	US	TRUE	flexible	Entire hon	2005	$449	$90	3	155	6/20/
19	Sunny Roo	35058485479	verified	Victoria	Brooklyn	Clinton Hil	40.68414	-73.9635	United Sta	US	FALSE	moderate	Private roo	2018	$930	$186	3	260	7/3/
20	Stylish & S	24559680874	verified	Brooke	Manhatta	East Villag	40.72392	-73.9914	United Sta	US	FALSE	strict	Entire hon	2006	$329	$66	5	73	6/25/

5. *Spellings and Formatting:*

Spellings and Formatting is a category of data cleaning that refers to checking and correcting any inconsistent or erroneous spellings, capitalization, punctuation, or formatting in the dataset. This is especially important for categorical data, such as property types, amenities, or neighborhood names, to ensure consistency and standardization for analysis.

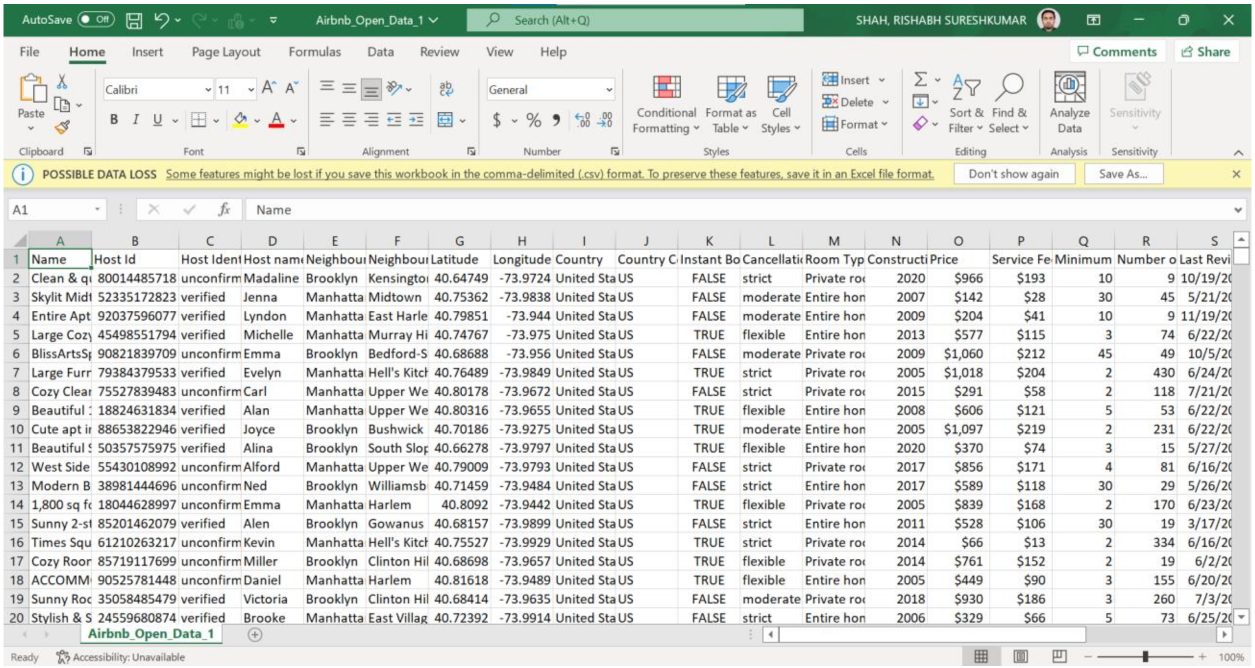
In the Airbnb dataset, for example, the "neighbourhood_group" column may contain inconsistent spellings or abbreviations for the same neighborhood, such as "Brooklyn," "Brookln," or "Bklyn." To clean the data in Excel, use tools like "Find and Replace" to correct any misspellings or inconsistencies, and standardize the formatting of the neighborhood names across the entire dataset.

Before:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	NAME	host id	host_iden	host name	neighbour	neighbour	lat	long	country	country cc	instant_bc	cancellatic	room type	Constructi	price	service fee	minimum	number o	last review	Da
2	Clean & qu	8E+10	unconfirm	Madaline	Brooklyn	Kensingtoı	40.64749	-73.9724	United Sta	US	FALSE	strict	Private roı	2020	$966	$193	10	9	########	
3	Skylit Midt	5.23E+10	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	########	
4	Entire Apt	9.2E+10	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	########	
5	Large Cozy	4.55E+10	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	########	
6	BlissArtsSp	9.08E+10	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private roı	2009	$1,060	$212	45	49	########	
7	Large Furr	7.94E+10	verified	Evelyn	Manhatta	Hell's Kitcł	40.76489	-73.9849	United Sta	US	TRUE	strict	Private roı	2005	$1,018	$204	2	430	########	
8	Cozy Clear	7.55E+10	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private roı	2015	$291	$58	2	118	########	
9	Beautiful 1	1.88E+10	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	########	
10	Cute apt ir	8.87E+10	verified	Joyce	Brooklyn	Bushwick	40.70186	-73.9275	United Sta	US	TRUE	moderate	Entire hon	2005	$1,097	$219	2	231	########	
11	Beautiful S	5.04E+10	verified	Alina	Brooklyn	South Slop	40.66278	-73.9797	United Sta	US	TRUE	flexible	Entire hon	2020	$370	$74	3	15	########	
12	West Side	5.54E+10	unconfirm	Alford	Manhatta	Upper We	40.79009	-73.9793	United Sta	US	FALSE	strict	Private roı	2017	$856	$171	4	81	########	
13	Modern B	3.9E+10	unconfirm	Ned	Brooklyn	Williamsb	40.71459	-73.9484	United Sta	US	FALSE	strict	Entire hon	2017	$589	$118	30	29	########	
14	1,800 sq fc	1.8E+10	unconfirm	Emma	Manhatta	Harlem	40.8092	-73.9442	United Sta	US	TRUE	flexible	Private roı	2005	$839	$168	2	170	########	
15	Sunny 2-st	8.52E+10	verified	Alen	Brooklyn	Gowanus	40.68157	-73.9899	United Sta	US	FALSE	strict	Entire hon	2011	$528	$106	30	19	########	
16	Times Squ	6.12E+10	unconfirm	Kevin	Manhatta	Hell's Kitcł	40.75527	-73.9929	United Sta	US	TRUE	strict	Private roı	2014	$66	$13	2	334	########	
17	Cozy Roor	8.57E+10	unconfirm	Miller	Brooklyn	Clinton Hil	40.68698	-73.9657	United Sta	US	TRUE	flexible	Private roı	2014	$761	$152	2	19	6/2/2019	
18	ACCOMM	9.05E+10	unconfirm	Daniel	Manhatta	Harlem	40.81618	-73.9489	United Sta	US	TRUE	flexible	Entire hon	2005	$449	$90	3	155	########	
19	Sunny Roc	3.51E+10	verified	Victoria	Brooklyn	Clinton Hil	40.68414	-73.9635	United Sta	US	FALSE	moderate	Private roı	2018	$930	$186	3	260	7/3/2019	
20	Stylish & S	2.46E+10	verified	Brooke	Manhatta	East Villag	40.72392	-73.9914	United Sta	US	FALSE	strict	Entire hon	2006	$329	$66	5	73	########	

After:

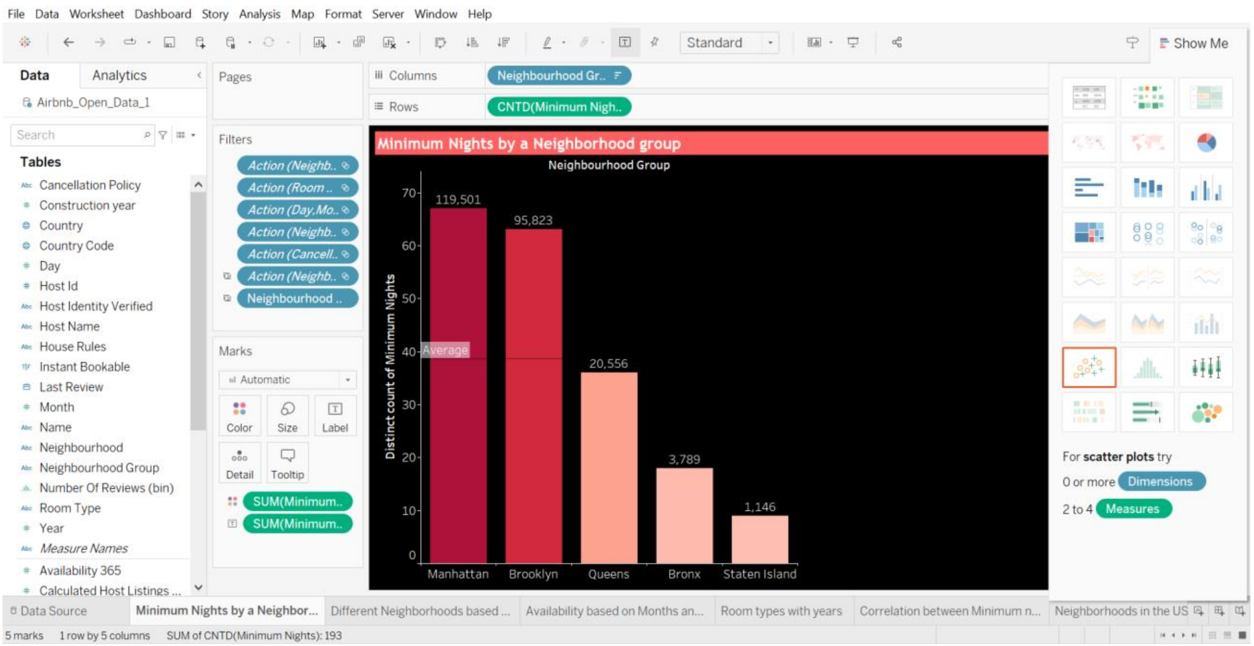


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Name	Host Id	Host Ident	Host name	Neighbour	Neighbour	Latitude	Longitude	Country	Country C	Instant Bo	Cancellati	Room Typ	Constructi	Price	Service Fe	Minimum	Number o	Last Revi
2	Clean & qu	80014485718	unconfirm	Madaline	Brooklyn	Kensingtor	40.64749	-73.9724	United Sta	US	FALSE	strict	Private roo	2020	$966	$193	10	9	10/19/20
3	Skylit Midt	52335172823	verified	Jenna	Manhatta	Midtown	40.75362	-73.9838	United Sta	US	FALSE	moderate	Entire hon	2007	$142	$28	30	45	5/21/20
4	Entire Apt	92037596077	verified	Lyndon	Manhatta	East Harle	40.79851	-73.944	United Sta	US	FALSE	moderate	Entire hon	2009	$204	$41	10	9	11/19/20
5	Large Cozy	45498551794	verified	Michelle	Manhatta	Murray Hi	40.74767	-73.975	United Sta	US	TRUE	flexible	Entire hon	2013	$577	$115	3	74	6/22/20
6	BlissArtsSp	90821839709	unconfirm	Emma	Brooklyn	Bedford-S	40.68688	-73.956	United Sta	US	FALSE	moderate	Private roo	2009	$1,060	$212	45	49	10/5/20
7	Large Furr	79384379533	verified	Evelyn	Manhatta	Hell's Kitch	40.76489	-73.9849	United Sta	US	TRUE	strict	Private roo	2005	$1,018	$204	2	430	6/24/20
8	Cozy Clear	75527839483	unconfirm	Carl	Manhatta	Upper We	40.80178	-73.9672	United Sta	US	FALSE	strict	Private roo	2015	$291	$58	2	118	7/21/20
9	Beautiful 1	18824631834	verified	Alan	Manhatta	Upper We	40.80316	-73.9655	United Sta	US	TRUE	flexible	Entire hon	2008	$606	$121	5	53	6/22/20
10	Cute apt in	88653822946	verified	Joyce	Brooklyn	Bushwick	40.70186	-73.9275	United Sta	US	TRUE	moderate	Entire hon	2005	$1,097	$219	2	231	6/22/20
11	Beautiful S	50357575975	verified	Alina	Brooklyn	South Slop	40.66278	-73.9797	United Sta	US	TRUE	flexible	Entire hon	2020	$370	$74	3	15	5/27/20
12	West Side	55430108992	unconfirm	Alford	Manhatta	Upper We	40.79009	-73.9793	United Sta	US	FALSE	strict	Private roo	2017	$856	$171	4	81	6/16/20
13	Modern B	38981444696	unconfirm	Ned	Brooklyn	Williamsb	40.71459	-73.9484	United Sta	US	FALSE	strict	Entire hon	2017	$589	$118	30	29	5/26/20
14	1,800 sq fo	18044628997	unconfirm	Emma	Manhatta	Harlem	40.8092	-73.9442	United Sta	US	TRUE	flexible	Private roo	2005	$839	$168	2	170	6/23/20
15	Sunny 2-st	85201462079	verified	Alen	Brooklyn	Gowanus	40.68157	-73.9899	United Sta	US	FALSE	strict	Entire hon	2011	$528	$106	30	19	3/17/20
16	Times Squ	61210263217	unconfirm	Kevin	Manhatta	Hell's Kitch	40.75527	-73.9929	United Sta	US	TRUE	strict	Private roo	2014	$66	$13	2	334	6/16/20
17	Cozy Roor	85719117699	unconfirm	Miller	Brooklyn	Clinton Hil	40.68698	-73.9657	United Sta	US	TRUE	flexible	Private roo	2014	$761	$152	2	19	6/2/20
18	ACCOMM	90525781448	unconfirm	Daniel	Manhatta	Harlem	40.81618	-73.9489	United Sta	US	TRUE	flexible	Entire hon	2005	$449	$90	3	155	6/20/20
19	Sunny Roo	35058485479	verified	Victoria	Brooklyn	Clinton Hil	40.68414	-73.9635	United Sta	US	FALSE	moderate	Private roo	2018	$930	$186	3	260	7/3/20
20	Stylish & S	24559680874	verified	Brooke	Manhatta	East Villag	40.72392	-73.9914	United Sta	US	FALSE	strict	Entire hon	2006	$329	$66	5	73	6/25/20

D) Data Visualization:

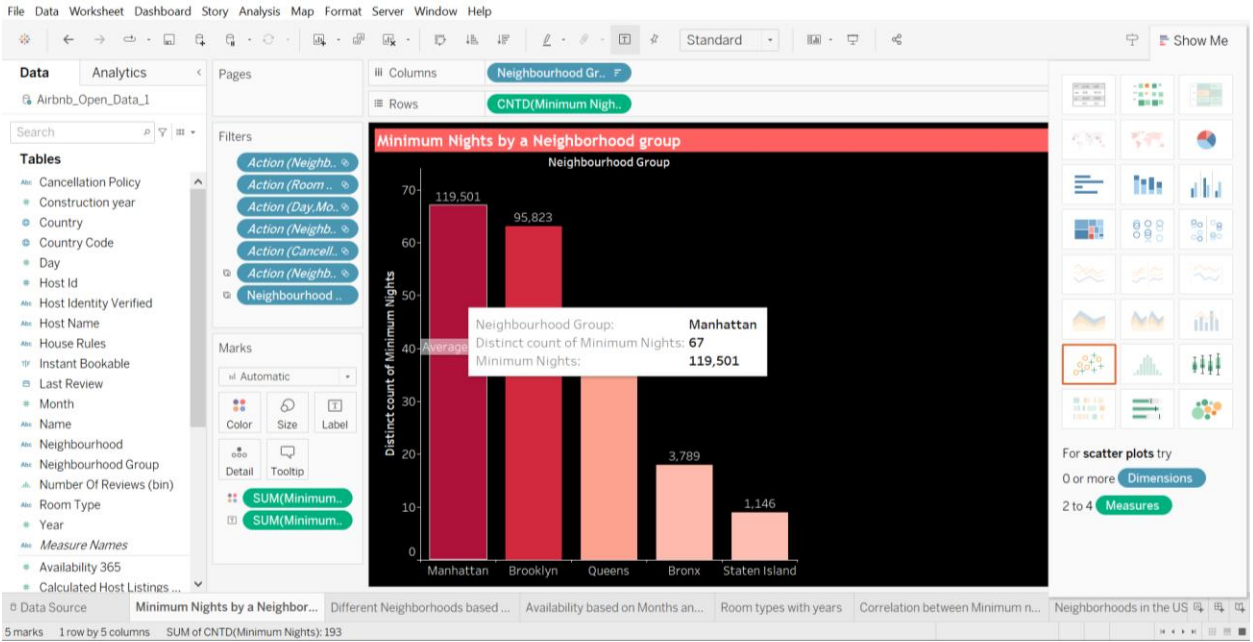
1. **How does the minimum number of nights required for booking an Airbnb property vary across different neighborhood groups?**

   Airbnb has divided its properties in 5 Neighborhood group named Bronx, Brooklyn, Manhattan, Queens, Staten Island.

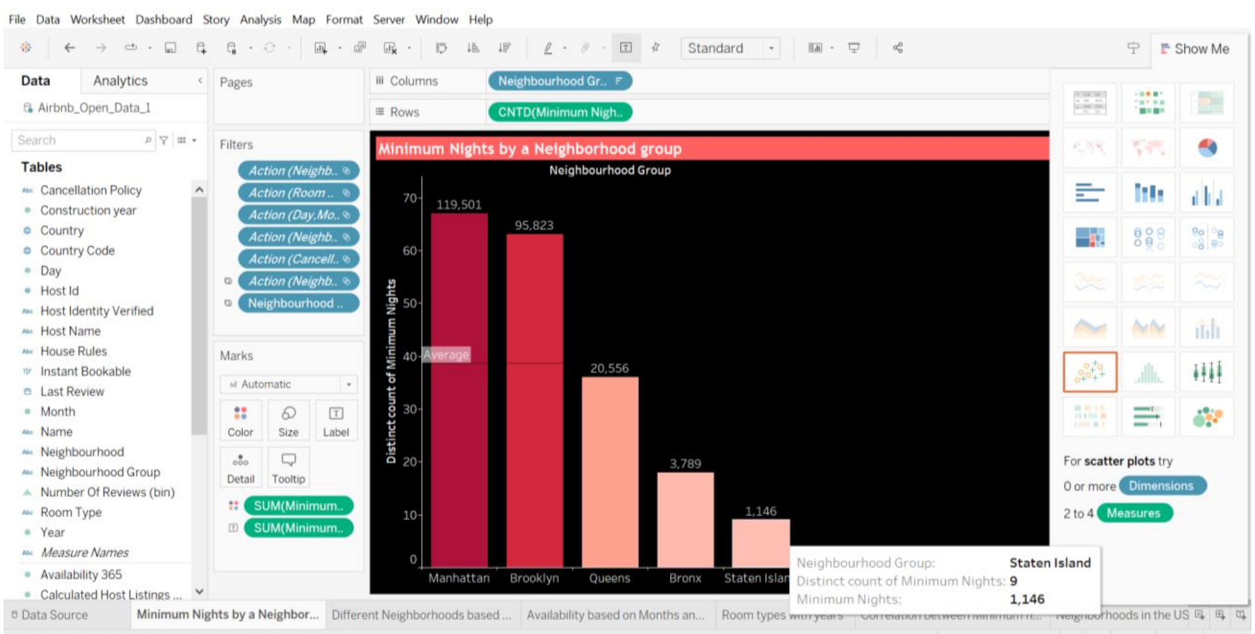


The data is presented in a Histogram format, with the horizontal axis representing neighborhood groups and the vertical axis showing the distinct count of minimum nights.

Each bar's height represents the number of minimum nights for each neighborhood group. The taller the bar, the more minimum nights that neighborhood group has.



Manhattan has the highest distinct count of minimum nights i.e 67, because its corresponding bar is the tallest of all the neighborhood groups. This indicates that properties in Manhattan tend to have longer minimum stay requirements compared to other neighborhood groups.

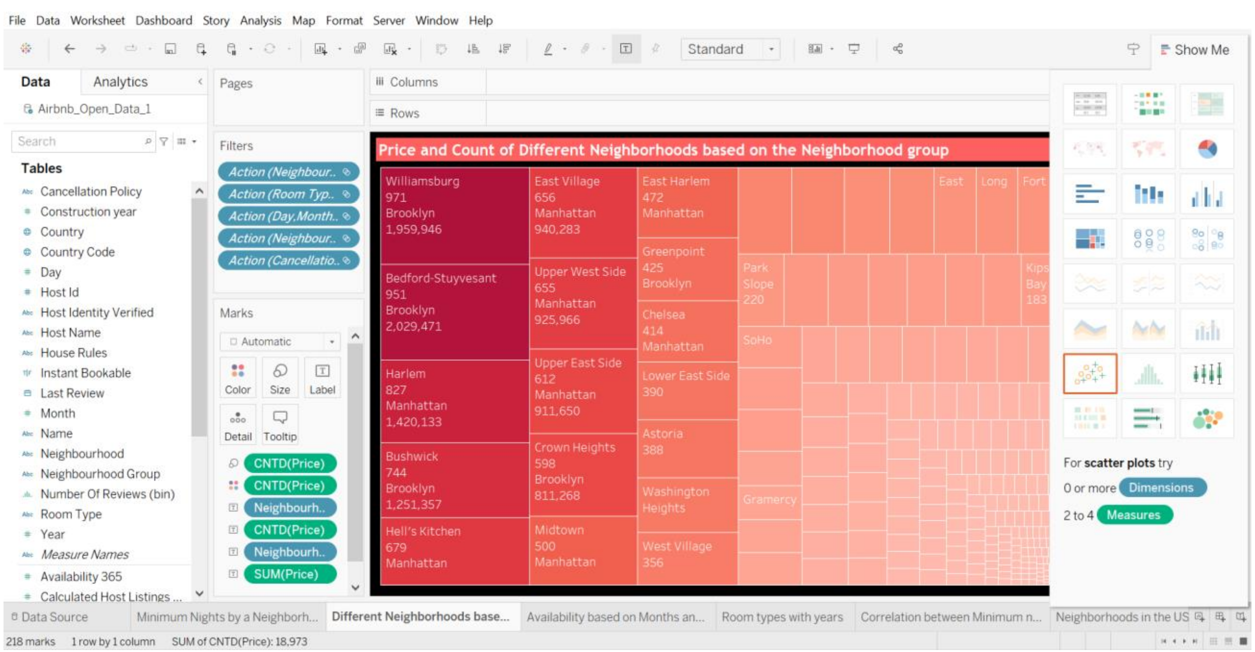


While, Staten Highland has the lowest distinct count of minimum nights i.e., 9, as its corresponding bar is the shortest among all the neighborhood groups. This implies that properties in Staten Island generally have shorter minimum stay requirements compared to other neighborhood groups.

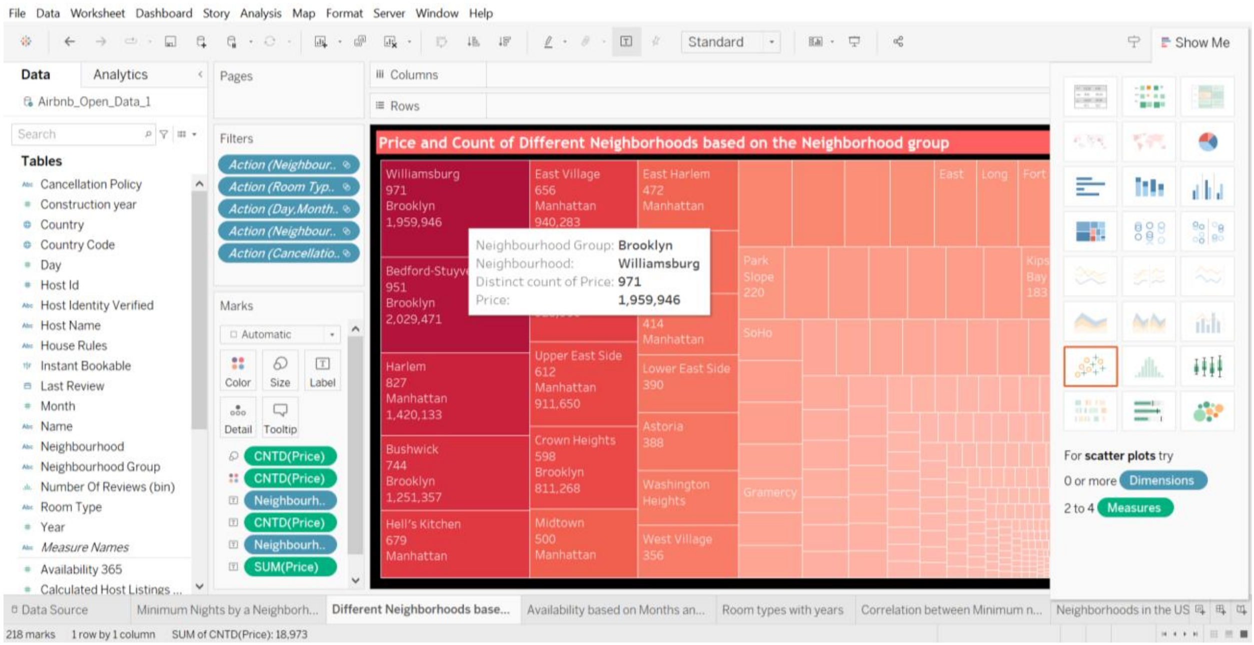
 # For the above visualization we have used Histogram.

2. **Which neighborhood occurs the most and least in the dataset based on neighborhood group and what is its total price?**

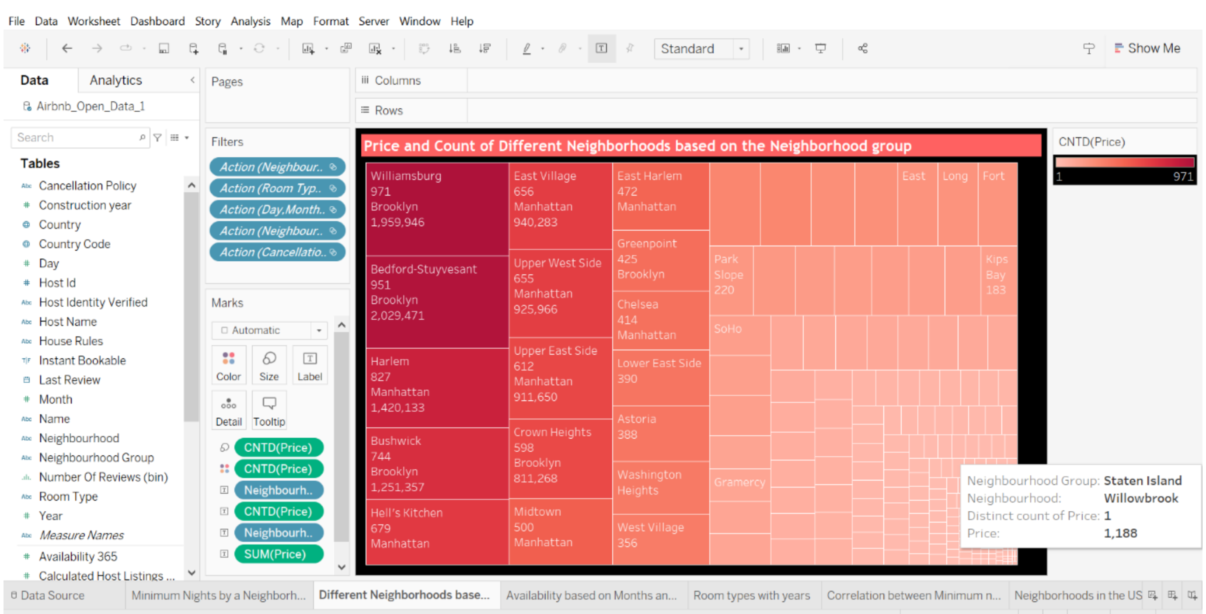
   The data visualization likely presents information about the relationship between neighborhoods (e.g., different areas or regions) and the distinct count of prices associated with them in the context of neighborhood versus a distinct price count in the tree maps. Tree maps are a type of hierarchical chart in which rectangles are used to represent data values, with the size of the rectangles indicating the magnitude of the data.



The main rectangular blocks or nodes in the chart would most likely be different neighborhoods in the tree maps. Each neighborhood would be represented by a rectangle, the size of which would indicate the distinct count of prices associated with that neighborhood.



The size of the rectangles would represent the unique price count for each neighborhood. Williamsburg represent higher counts for price of 971.



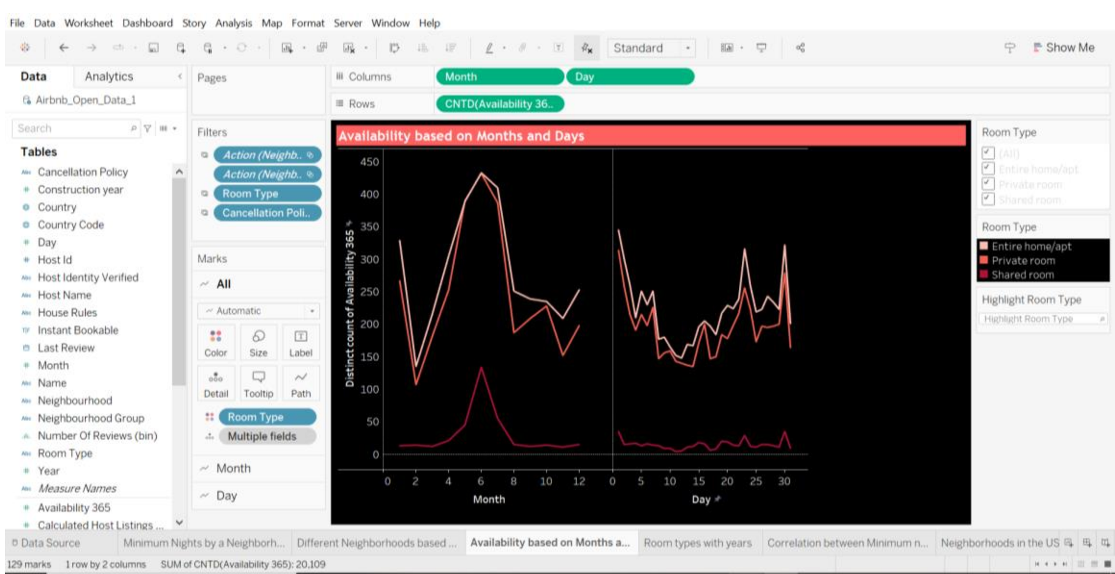
While smaller rectangles would indicate lower counts. This would allow for a visual comparison of the distinct price counts across different neighborhoods.

Color coding may also be used in tree maps to provide additional data information, such as different colors representing different price ranges or levels. This could aid in visually distinguishing neighborhoods with different price counts or pricing patterns.

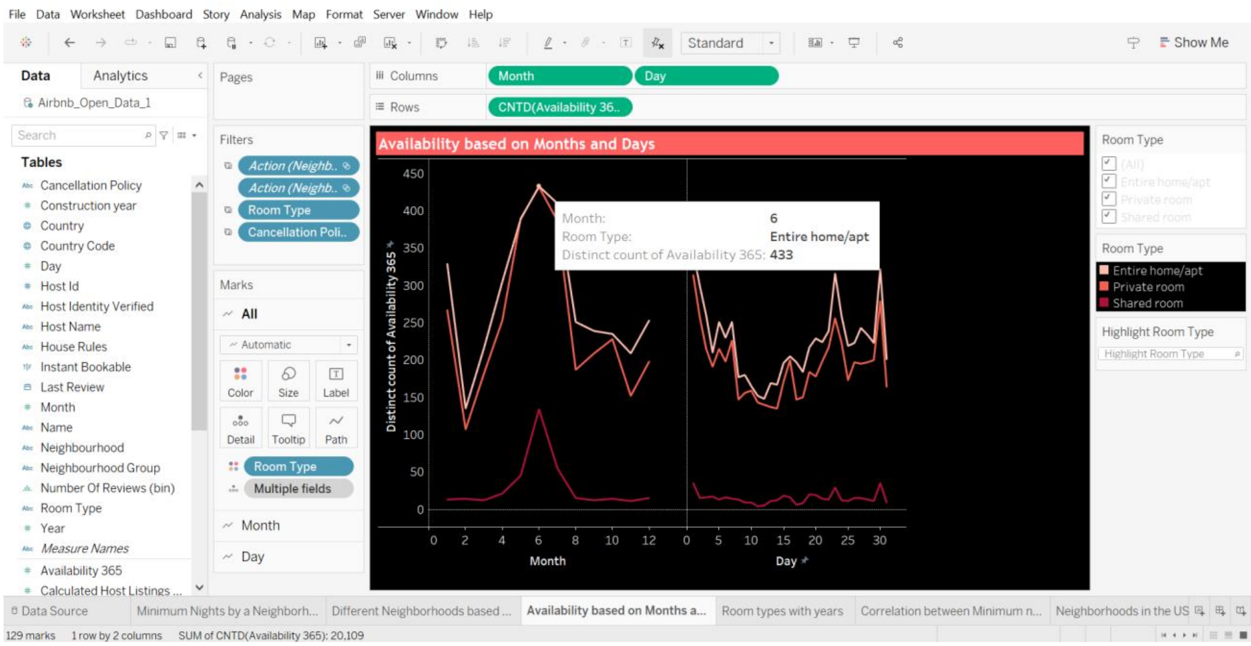
#For the above visualization we have used Tree maps and Rank.

3. **What is the trend of availability for 365 days across different months and days, segmented by room types in the Airbnb?**

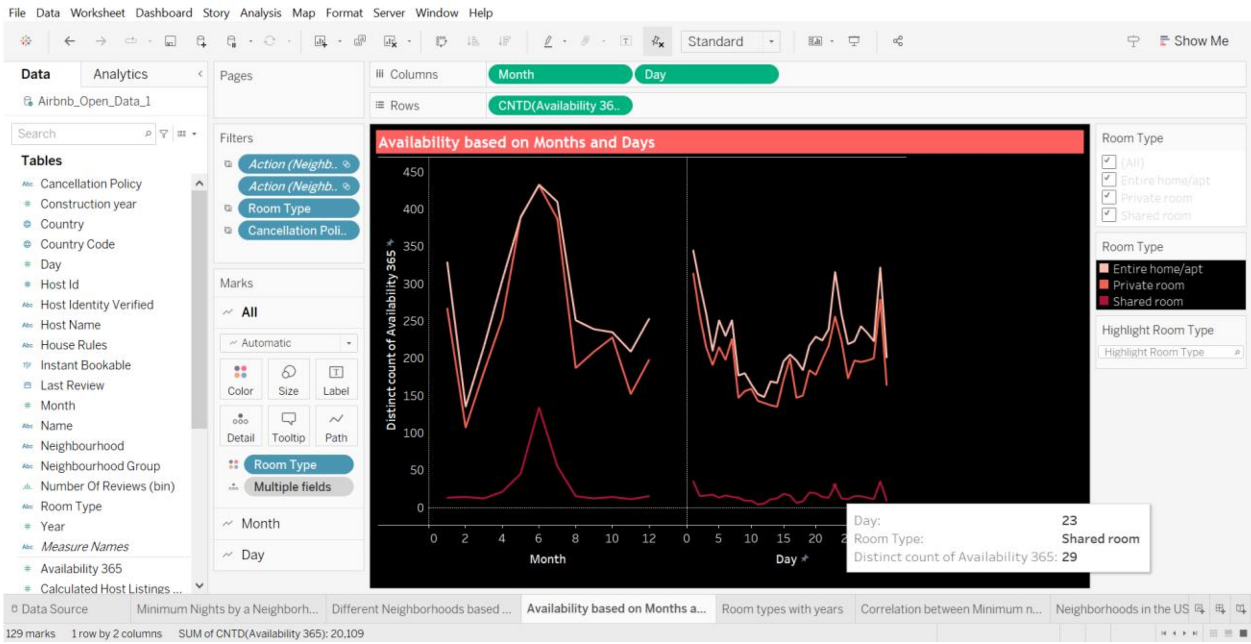
The line graph would have two axes, with the months and days on the x-axis and the distinct count of availability for the entire year on the y-axis. Each line on the graph represents one of the three room types (e.g., entire home/apt, private room, shared room), and its position on the y-axis represents the number of available listings for that room type on a given month and day.



The line graph would show the trend of availability for each room type over the course of the year, with changes in the number of available listings plotted along the y-axis. This could aid in the identification of any patterns or trends in the availability of various room types, such as seasonal fluctuations or changes in demand over time.



Room type comparison: By comparing the lines representing the various room types, any differences in availability patterns between the room types can be identified. One room type, for example, may have consistent availability throughout the year, whereas another may have more fluctuations or peaks and valleys.

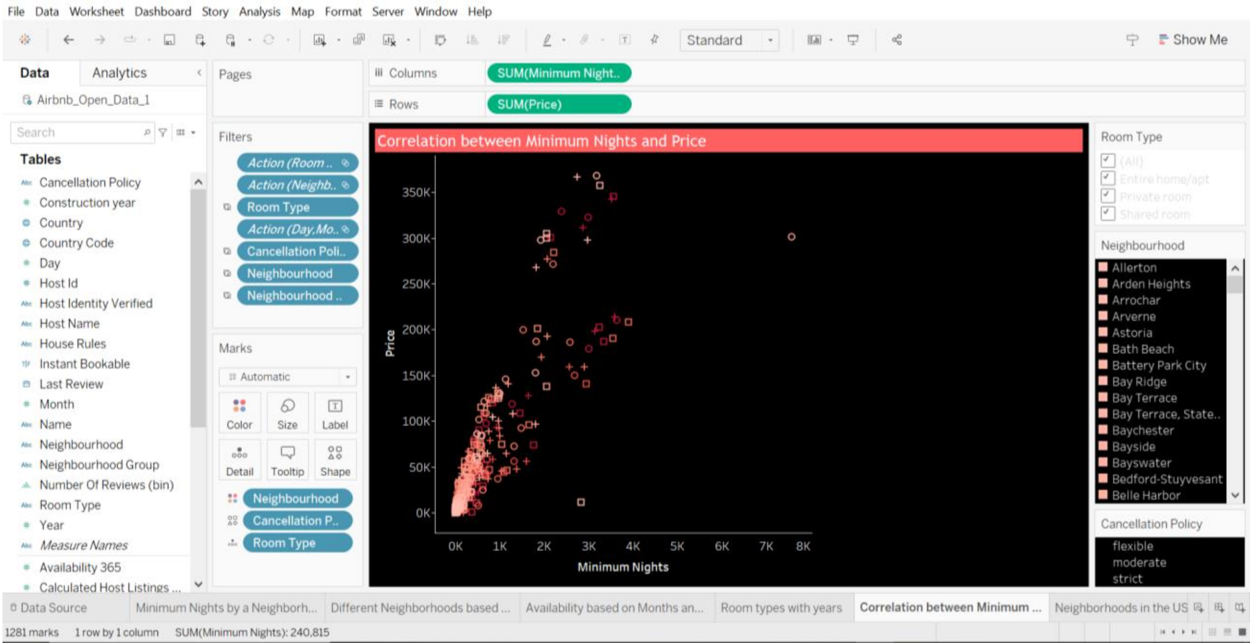


Seasonal patterns: The line graph can aid in the identification of any seasonal patterns in availability for each room type. Certain room types, for example, may have higher availability during certain months or days of the year, which could indicate seasonality in demand for those types of accommodations.

#For the above visualization we have used Line Chart , Forecast Trend Lines & Dates

4. **What is the relationship between the total sum of minimum nights and the total sum of price for Airbnb listings, based on neighborhood, cancellation policy, and room type?**

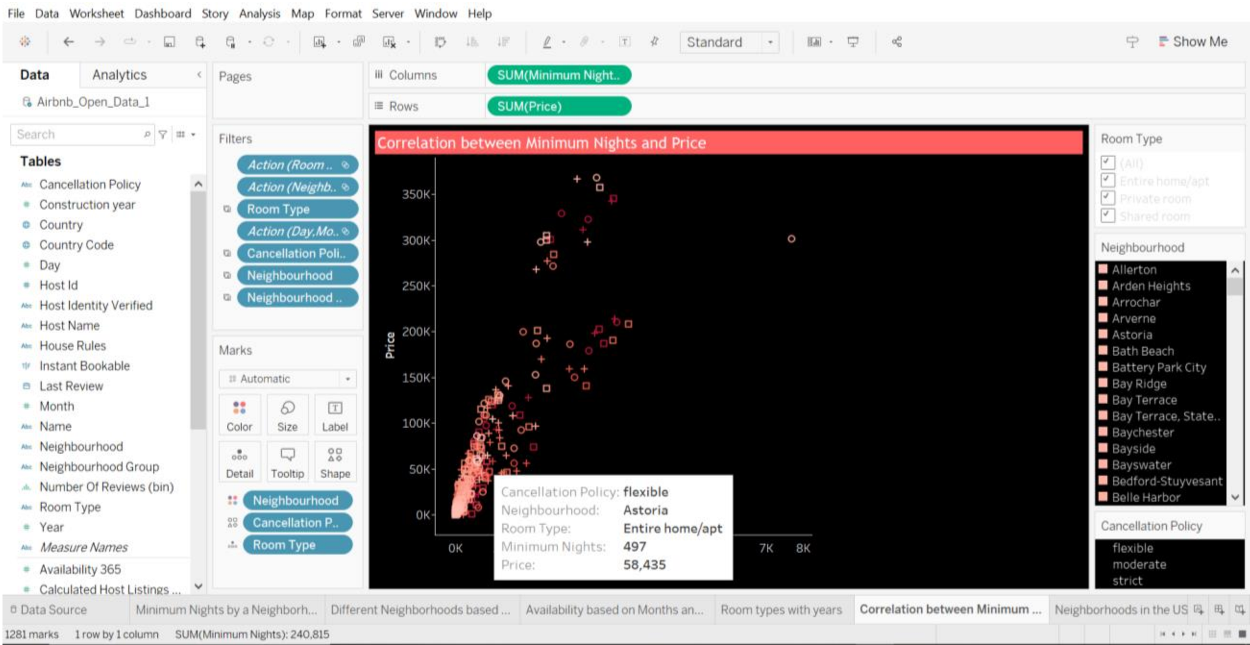
Analyzing the patterns, trends, and relationships between the sum of minimum nights, sum of price, and the various categories would be necessary for interpreting the information from the scatter plot graph. (Neighborhood, cancellation policy, and room type). Among the possible interpretations are:



The scatter plot may reveal whether there is a positive, negative, or no relationship between the sum of minimum nights and the sum of price. If the data points cluster along a diagonal line sloping upwards from left to right, it may indicate a positive correlation, implying that higher minimum nights are associated with higher prices, and vice versa.

If, on the other hand, the data points tend to cluster along a diagonal line sloping downwards from left to right, this could indicate a negative correlation, implying that higher minimum nights are associated with lower prices, and vice versa.

If no clear pattern emerges, it may imply that there is no significant relationship between the two variables.

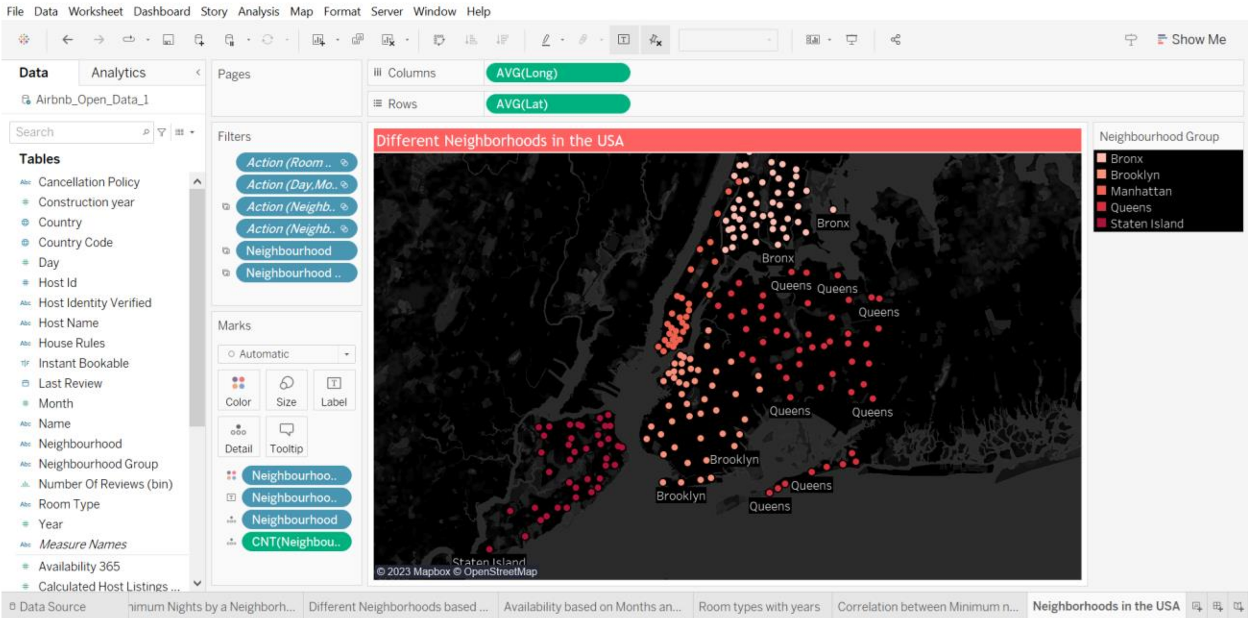


The scatter plot could also demonstrate how the sum of minimum nights and total price vary depending on cancellation policies and room types. The data points, like the neighborhood, could be color-coded or labeled to represent different cancellation policies and room types. This could reveal how different cancellation policies or room types affect the relationship between minimum nights and price.

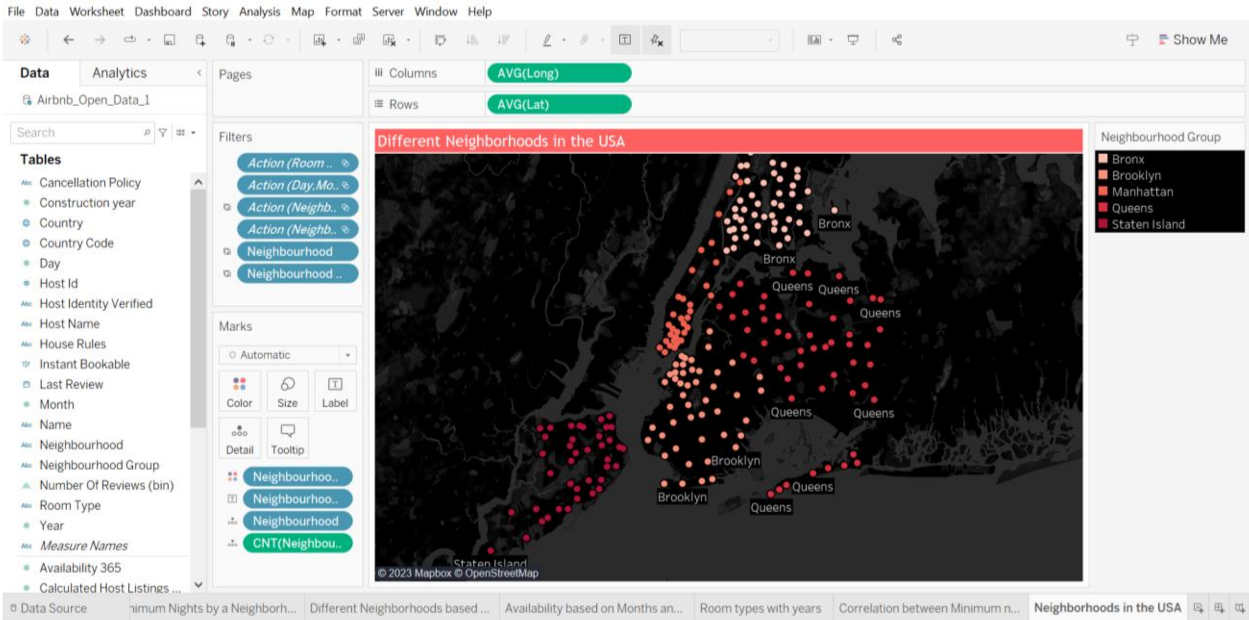
#For the above visualization we have used Scatter Plot

5. **What is the geographical distribution of Airbnb listings regarding longitude and latitude, segmented by neighborhood and neighborhood group?**

The symbols or markers would be color-coded or grouped based on information about the neighborhood and neighborhood groups. This would allow for easy visual identification of the map's various neighborhoods and neighborhood groups.



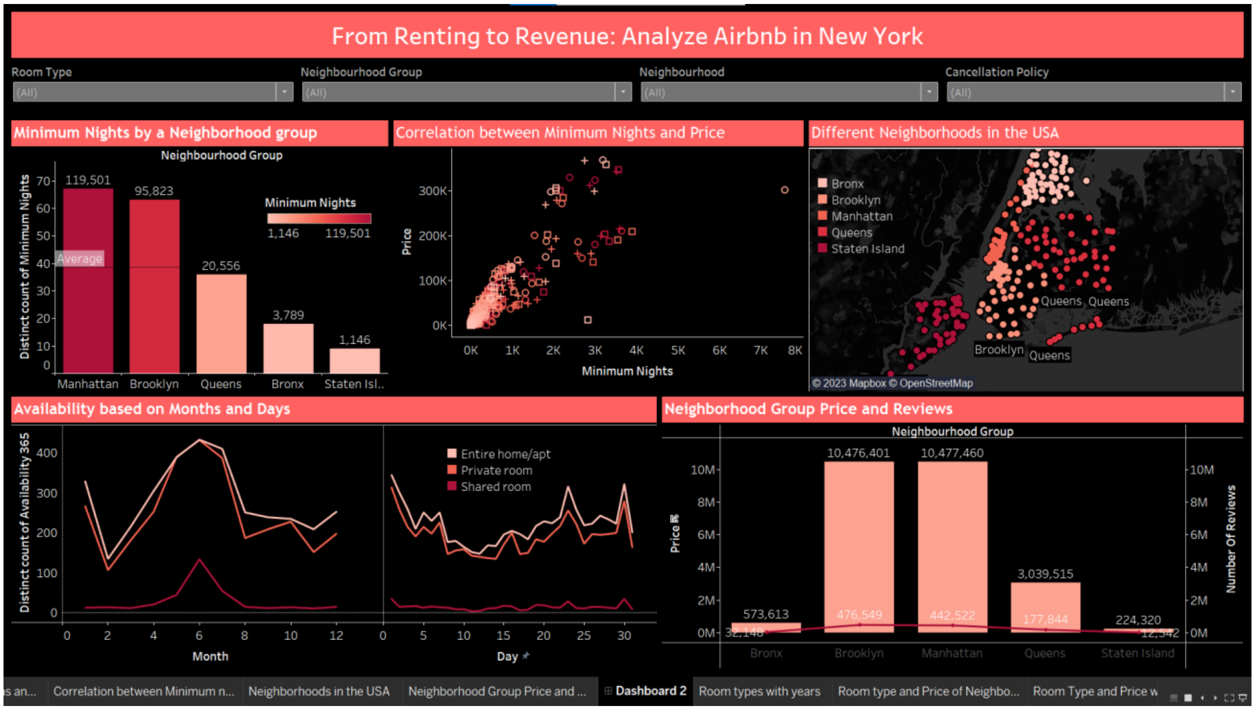
Analyzing the distribution of the symbols or markers on the map would be necessary for interpreting the information from the data visualization. Based on the average longitude and latitude values, the spatial arrangement of the symbols would provide insights into the geographic distribution of Airbnb listings. Clustering or patterns of symbols may indicate listing concentrations in specific areas, whereas sparse or scattered symbols may indicate lower listing densities in other areas.



This data visualization would aid in comprehending the geographic distribution of Airbnb listings in relation to the average longitude and latitude, as well as how this distribution varies across neighborhoods and neighborhood groups. It may also reveal any spatial patterns or trends in the data that can be used to make informed decisions or identify potential opportunities in the Airbnb market. So, analyzing the spatial distribution of the symbols, identifying patterns or trends, and gaining insights into the relationship between location, neighborhood/neighborhood group, and Airbnb listings would be part of the interpretation.

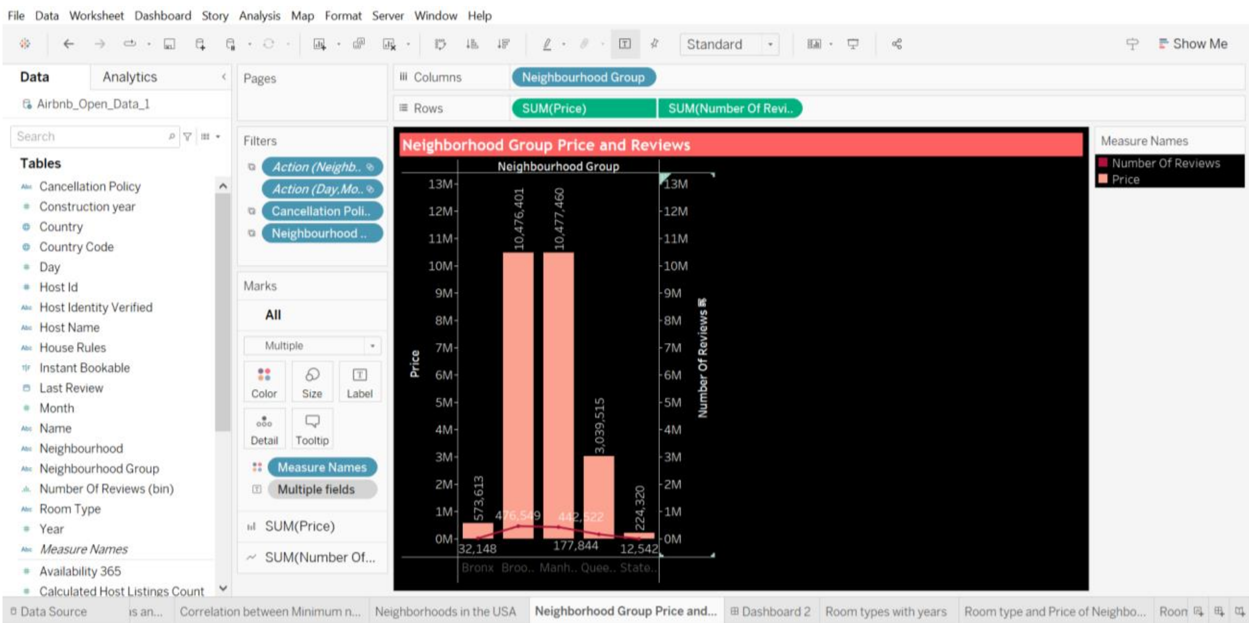
#For the above visualization we have used Geo Spatial Map

E) ***Dashboard:***



As you can see here it is an interactive dashboard and it has all the required charts that will allow us to analyze our data set so now let's talk about some of the insights that we can get from this dashboard first of all it is Histogram, we have the total bookings by minimum nights and neighborhood group and as you can see here this means simply that the names selected in here they do not have last review date and last review date is the only date available in the data set so we can just exclude this value .so we don't have our values and as you can see here we have most of the reservations are made in the summer and on June for all neighborhood groups and in the second position we have a then we have January for total reviews by year with more than million reviews if we use the map and the tree map we can see that Manhattan has the highest average price for the rent with a value which is nearly and if we use the same map and the horizontal bar chart we can see that the neighborhood that has the highest average price which is 800$ is Fort Wadsworth and it is located in here so we have the group Staten Island.

Also, we have added one additional visualization to the dashboard:

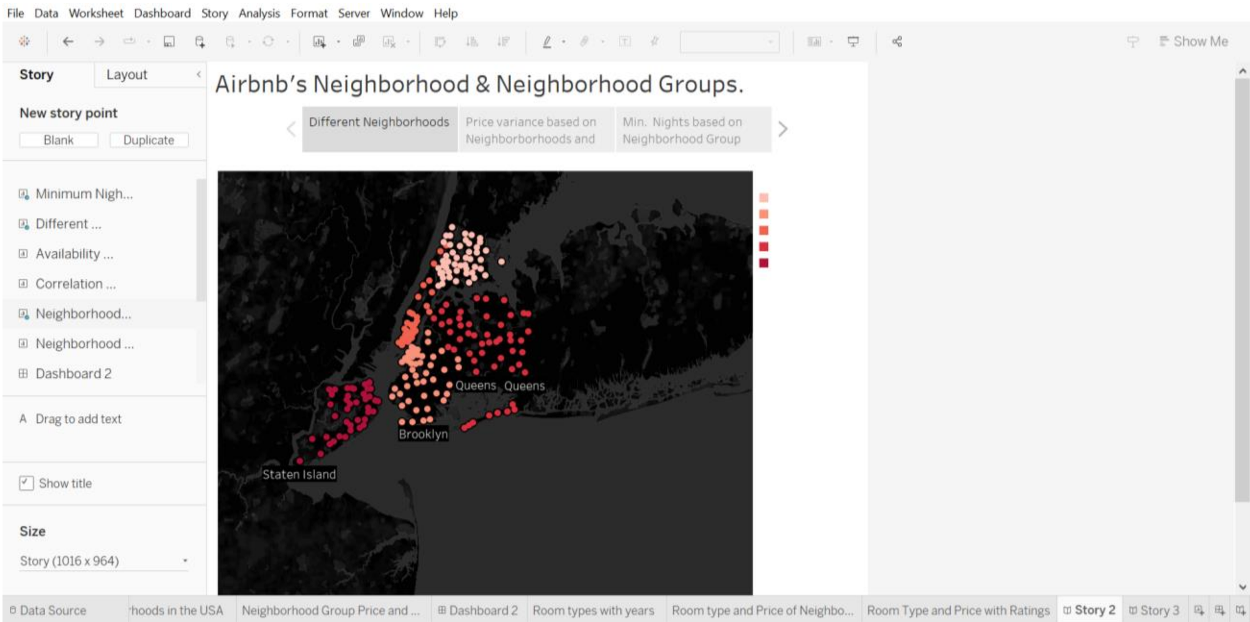


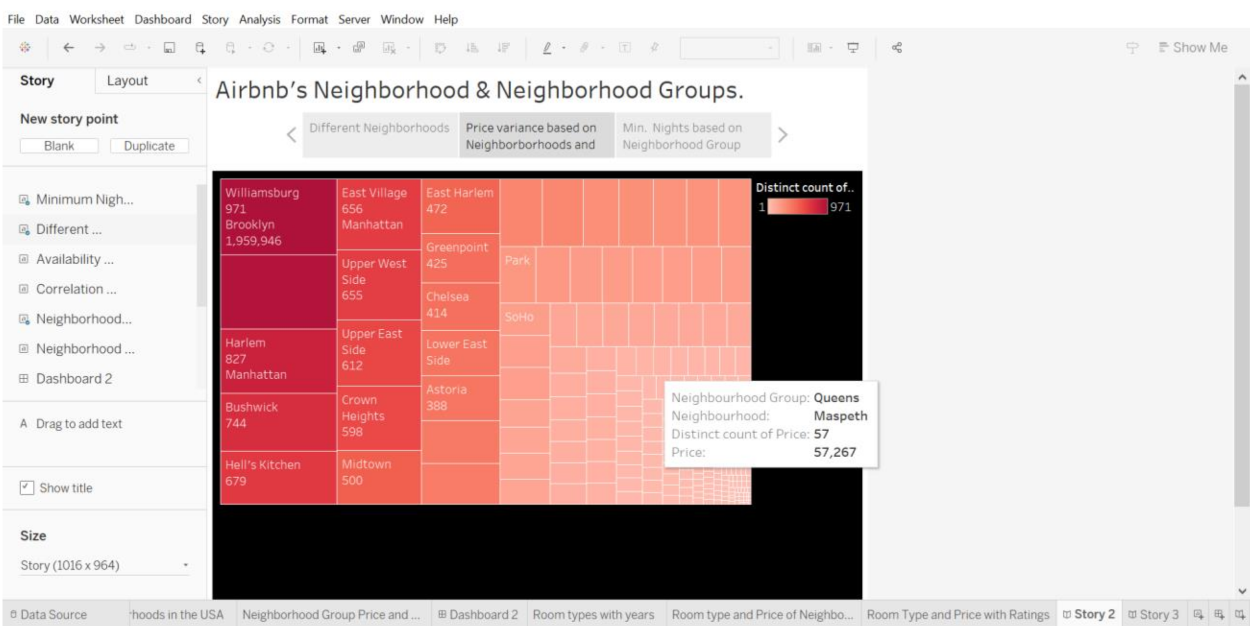
The dual axis chart that shows neighborhood group vs. price and number of reviews illustrates how these two variables are related to each other across different neighborhood groups. The chart has two axes: the left axis displays the average price of Airbnb listings in each neighborhood group, and the right axis displays the average number of reviews received by listings in each neighborhood group.

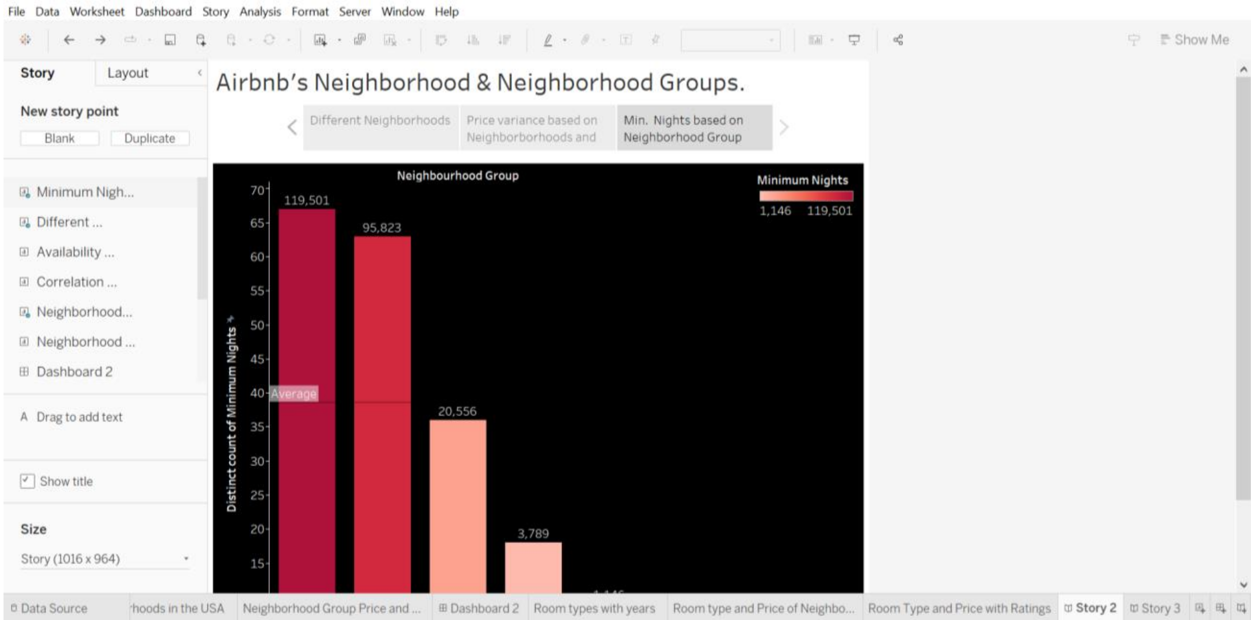
Each neighborhood group is represented by a different color on the chart, which is color-coded by neighborhood group. Each group is also represented by a horizontal bar, the length of which indicates the average price and the width of which indicates the average number of reviews.

F) Story Telling:

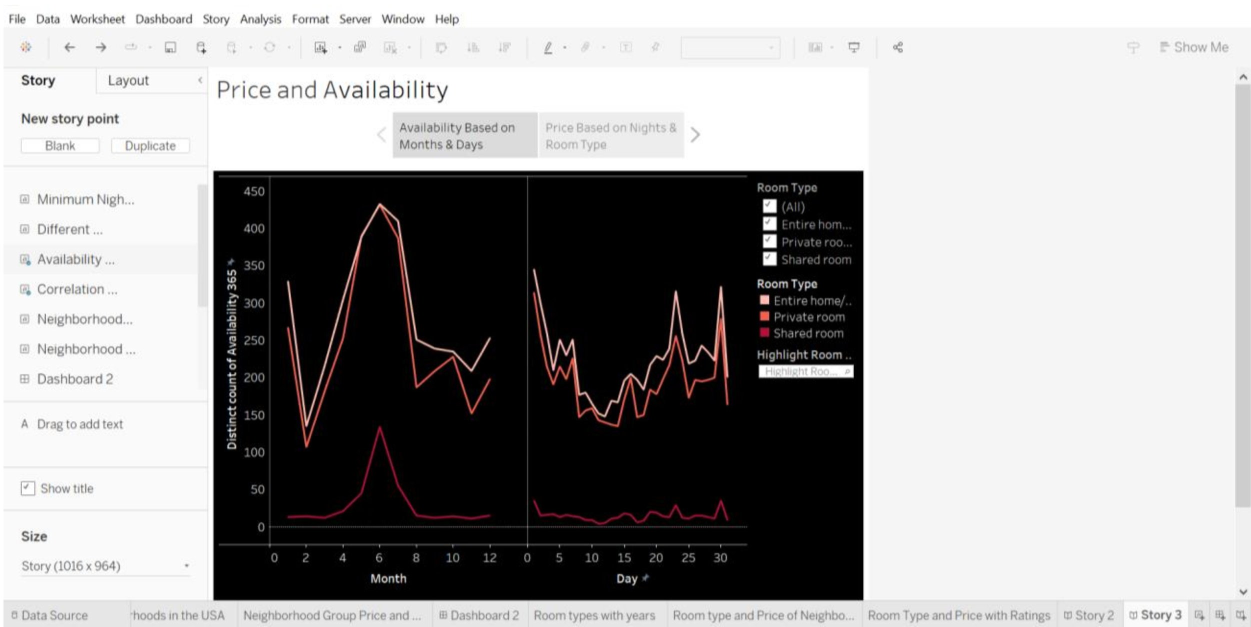
Story 1: Airbnb’s Neighborhood & Neighborhood Groups

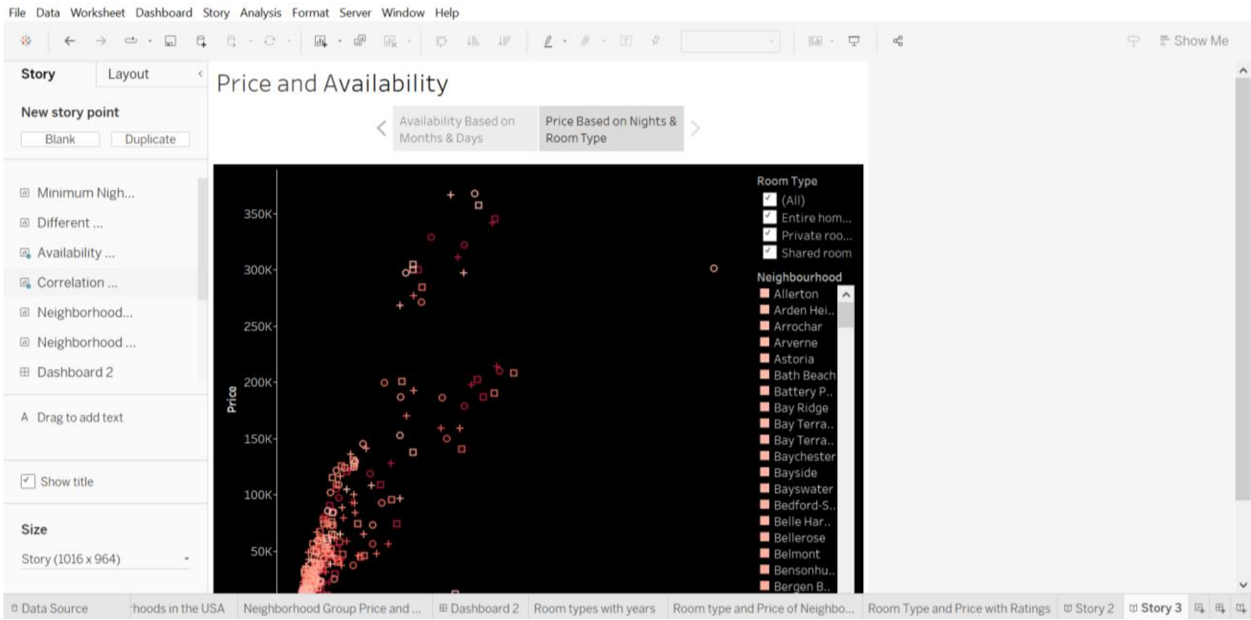






Story 2: Price and Availability





Only a few years ago, most travelers stayed in hotels. Airbnb changed that. As of 2018, the company offers over 5 million properties, in over 85,000 cities across the world, and its market valuation exceeds $30 billion. In 2017 alone, Airbnb users booked over 100 million nights.

We discovered several interesting findings after analyzing the Airbnb Data dataset in Tableau:

Location:

- Why travelers choose Airbnb? When reflecting upon the rapid emergence of Airbnb, perhaps the first question that arises is why millions of travelers are opting to stay in the (oftentimes unlicensed) homes of strangers found online, rather than just simply booking a hotel. In Guttentag's (2015) early look at Airbnb through the lens of disruptive innovation theory, he proposed that even though Airbnb may underperform in comparison with hotels when considering hotels' primary performance attributes (e.g., service quality and security), Airbnb offers an alternative value proposition centered around cost-savings, household amenities and the potential for a more authentic local experience.

- Airbnb guests and cost-savings to be their primary motivator. Nevertheless, the study focused explicitly on authenticity, and the researchers also found authenticity contributed to Airbnb's appeal, as related to three key areas – the accommodations, the social interactions with hosts and the interactions with local culture

- The impact of location on pricing was one of the most significant findings. We discovered that the average price of listings varies significantly by neighborhood, with SoHo and Tribeca having much higher prices than Harlem and the Bronx.

- This is most likely due to factors such as the neighborhood's desirability, the quality of the properties, and the level of amenities and services available. This finding is consistent with the findings of a Smart Asset study, which looked at data on Airbnb listings in 15 major U.S. cities and discovered that listings in neighborhoods with higher home values and income tended to be more expensive.

- The study also discovered that proximity to tourist attractions and public transportation were important pricing factors. (Using AirBnB for Real Estate Investing - SmartAsset)

Property & Room Type:

- we show that a 1% increase in Airbnb listings is causally associated with a 0.018% increase in rental rates and a 0.026% increase in house prices. While these effects may seem very small, consider that Airbnb's year-over-year average growth is about 44%.

- These results show that Airbnb does have an impact on the housing market. However, they don't tell the full story of how it is happening. In our study, we present two additional results that help explain:

- First, we show that zip codes with higher owner-occupancy rates (the fraction of properties occupied by the owners themselves) are less affected by Airbnb. Those rates are important because the landlords who switch their properties from long-term rentals to short-term rentals are those who don't live in the houses they rent. Owner-occupiers do use Airbnb, but they use it to rent out their spare rooms or perhaps the whole home while they are away. However, these homes are still primarily occupied by a long-term resident (the owner), so they are not the ones being reallocated as short-term rentals through Airbnb.

- Second, we present evidence that Airbnb affects the housing market through the reallocation of housing stock. By looking at housing vacancies, we show two things about the Airbnb supply: it is positively correlated with the share of homes that are vacant for seasonal or recreational use — which is how the Census Bureau classifies houses that are part of the short-term rental market — and negatively correlated with the share of homes in the market for long-term rentals.

- Taken together, our results are consistent with the story that, because of Airbnb, absentee landlords are moving their properties out of the long-term rental and for-sale markets and into the short-term rental market.
- We also discovered that the type of property has a significant impact on pricing. The most expensive options were entire homes and apartments, while shared rooms were the least expensive.
- This is most likely due to the level of privacy and amenities provided by each property type.
- This finding is consistent with a Priceonomics study, which found that entire homes and apartments were typically more expensive than private or shared rooms in major U.S. cities.
- The study also discovered that treehouses, yurts, and other unique properties were generally more expensive than traditional lodging. (The Case for Airbnb - Priceonomics)

Reviews & Ratings:

- How guests choose their Airbnb accommodation? In many destinations, Airbnb guests have a multitude of Airbnb listings to choose from, and numerous studies have examined how such decisions are made. Examined which listing attributes influenced Airbnb listing demand in Vienna and found that listing size, photo quantity and host response rates increased demand, whereas price, distance from city center and host response time decreased demand.

- Reviews are indeed a central feature of Airbnb, because they help establish the required trust between guest and host, and several researchers have specifically investigated how reviews impact Airbnb choices. conducted an experiment with several thousand Airbnb users in which the authors manipulated the demographics (age, gender, marital status and home state) and reputation (star ratings and review quantity) of hypothetical hosts and found that positive reputations successfully counteracted biased distrust that arose due to social distance.

- Finally, we discovered that there was no clear relationship between reviews and ratings and listing price. While we expected higher prices for properties with higher ratings and more positive reviews, we discovered that this was not always the case.

- This finding is consistent with a Harvard Business Review study that examined data from Airbnb listings in Boston and discovered that higher-rated properties did not necessarily command higher prices.

- Other factors, such as location and property type, were found to be more significant predictors of price in the study. (Research: When Airbnb Listings in a City Increase, So Do Rent Prices (hbr.org))

Conclusions:

Finally, our analysis of the "Airbnb " dataset yielded several insights into the New York City Airbnb market. Location, property type, number of bedrooms, and reviews and ratings all have a significant impact on pricing, according to our findings. These findings are consistent with other Airbnb market studies, emphasizing the importance of these factors for both hosts and guests. Understanding these factors allows hosts to better price their properties and guests to make more informed booking decisions.

References:

1) Research: When Airbnb Listings in a City Increase, So Do Rent Prices (hbr.org)

2) The Case for Airbnb - Priceonomics

3) Algorithmic management: The case of Airbnb - ScienceDirect

4) (PDF) Progress on Airbnb: a literature review (researchgate.net)

5) The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry by Georgios Zervas, Davide Proserpio, John Byers :: SSRN