Executive Summary

The following data analysis of the Manhattan Vacation Rental market aims to assist investors in identifying the ideal properties for investment purchase based on neighborhood desirability, property size, and revenue generation.

In order to eventually isolate listings with optimal revenue generation, the following data analysis identifies the top 10 most desirable neighborhoods. Desirability was derived as an intersection of number of reviews and most popular property size for the neighborhood.

Once the most desirable neighborhoods were identified, these neighborhoods were cross examined against the most popular property size in said neighborhoods to identify the top listings.

Once the top listings were identified, further analysis of their calendar bookings helped create data about revenue generation.

Finally, the aforementioned investigations led to the idenfication of the listing with the highest revenue generation that was the most popular property size in a top 10 desirable neighborhood.

Table of Contents

Assumptions

In this analysis, the following assumptions must be made about all shown data and derived insights:

1. A neighborhood's desirability is directly correlated to the number of reviews for that borough.

2. The popularity of a property's size is cross-derivative of the number of bedrooms in the unit and the number of reviews for the listing.

3. The top listings are cross-derivative of the most popular property sizes in the pre-determined most desirable neighborhoods.

4. Annual revenue of listings is a projection derived from monthly earnings in the calendar data.

5. Top earners were identified from listings with the highest revenue earned on a nightly basis.

Change Log

Version 1.1 - Data Cleaning: *Comprehensive* - Assesed and corrected for duplicates, inconsistencies, formatting errors in Prepared Data - Listings

Version 1.2 - Data Cleaning: *Specific Column* - Neighborhood - Edited for capitalization and formatting errors. New column renamed as neighborhood_clean.

Version 1.2.1 - Data Cleaning: *Specific Column* - Bedrooms - Identified empty cells and replaced with '0'. New column renamed as bedrooms_clean.

Version 1.3 - Pivot Table: *Created - Most Desirable Neighborhoods* - Created Pivot Table using cleaned data to identify which neighborhoods were most attractive, where attractiveness was established as correlated to 'number_of_reviews_ltm'

Version 1.4 - Pivot Table: *Created - Most Popular Property Sizes* - Created Pivot Table using cleaned data to identify which properties were most attractive, where attractiveness was established as correlated to the number of reviews per each incidence of number of bedrooms

Version 1.5 - Data Entered - Prepared Data Listings: *Column Created - Top Neighborhood* - Column created with congenate function to account for top 10 neighborhoods from Pivot: Most Desireable Neighborhoods based on reviews and property size

Version 1.6 - Data Entered - Prepared Data Listings: *Column Created - Top Property Size* - Column created to identify which property size was most popular, with applied function based on number of reviews and top neighborhoods.

Version 1.7 - Data Entered - Prepared Data Listings: *Column Created - Top Listings* - Column to identify which listings were most popular, based on a function with arguments relying on the aforementioned outcome of functions in created columns "Top Neighborhood" and "Top Property Size"

Version 1.8 - Data Entered - Raw Data Calendar: *Column Created - Revenue_Earned* -

*Please note that my entire chrome application would freeze and force quit everytime I tried to duplicate this sheet to create a format of "PreparedData_Calendar". For this reason, I added a column to the 'Raw Data Calendar' sheet and worked from there. Ordinarily, I would have duplicated and renamed the sheet to work with the data, as well as create a link to the prepared data sheet for calendars in the table of contents.

Version 1.9 - Pivot Table: *Created - Top Listings* - Pivot table created to identify the top earning listing, as a derivative of listing ID and sum of revenue (based on column created - revenue_earned)

Version 2.0 - Data Cleaning: *Comprehensive* - Hid columns in Prepared Data Listings that were not used, verified that formatting was consistent throughout sheets

Version 2.1 - Data Entered: *Start Here*, *Assumptions + Change Log* - Assumptions entered into corresponding category, links for pivot tables created in Table of Contents, Executive Summary entered into corresponding category

+ ≡   Start Here ▾   **Assumptions + Change Log** ▾   Data Dictionary ▾   RawData_Listings ▾

Data Dictionary for Listings

Field	Type	Calculated	Description
id	integer		Airbnb's unique identifier for the listing
listing_url	text	y	
scrape_id	bigint	y	Inside Airbnb "Scrape" this was part of
last_scraped	datetime	y	UTC. The date and time this listing was "scraped".
source	text		One of "neighbourhood search" or "previous scrape". "neighbourhood search" means that the listing was found by searching the city, while "previous scrape" means that the listing was seen in another scrape performed in the last 65 days, and the listing was confirmed to be still available on the Airbnb site.
name	text		Name of the listing
description	text		Detailed description of the listing
neighborhood_overview	text		Host's description of the neighbourhood
picture_url	text		URL to the Airbnb hosted regular sized image for the listing
host_id	integer		Airbnb's unique identifier for the host/user
host_url	text	y	The Airbnb page for the host
host_name	text		Name of the host. Usually just the first name(s).
host_since	date		The date the host/user was created. For hosts that are Airbnb guests this could be the date they registered as a guest.
host_location	text		The host's self reported location
host_about	text		Description about the host
host_response_time			
host_response_rate			
host_acceptance_rate			That rate at which a host accepts booking requests.
host_is_superhost	boolean [t=true; f=false]		
host_thumbnail_url	text		
host_picture_url	text		
host_neighborhood	text		
host_listings_count	text		The number of listings the host has (per Airbnb calculations)
host_total_listings_count	text		The number of listings the host has (per Airbnb calculations)
host_verifications			
host_has_profile_pic	boolean [t=true; f=false]		
host_identity_verified	boolean [t=true; f=false]		
neighborhood	text	y	The neighbourhood as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.
neighbourhood_group	text	y	The neighbourhood group as geocoded using the latitude and longitude against

id	neighborhood	bedrooms	beds	number_of_reviews_ltm
82638	East Village	1	1	3
9357	Hells Kitchen		1	0
57618	hells kitchen	1	2	7
57754	Flatiron District	1	1	28
57874	East Harlem	1	1	55
15341	LOWER EAST SIDE	1	2	6
59014	Lower East Side	1	1	1
59709	chinatown	2	2	23
60164	Tribeca	3	3	3
16974	East Harlem	4	6	26
60611	East Harlem	1	1	1
107895	Upper West Side		1	2
17037	East Village	1	1	0
60680	East Village	1	2	2
61509	Midtown	1	1	0
62427	East Village	1	1	1
19159	Harlem	2	2	0
62891	East Village		2	9
116551	Harlem		2	2
63693	east village	2	2	0
23686	West Village	3	3	3
123784	Harlem		2	33
126443	east village	2	4	60
27644	Harlem	1	4	18
65425	Nolita	2	2	10
29683	Lower East Side	1	1	2
66251	East Village		1	32
140973	East Village	2	3	4
67288	East Harlem	1	1	7
144087	Harlem	2	3	7
36703	West Village	1	2	5
148825	UPPER WEST SIDE	2	2	36
68974	Nolita	3	3	11

Start Here   Assumptions + Change Log   Data Dictionary   RawData_Listings

	A	B	C	D	E	F	G	H
	listing_id	date	available	price	adjusted_price	minimum_nights	maximum_nights	revenue_earned
	9357	2022-09-07	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-08	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-09	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-10	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-11	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-12	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-13	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-14	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-15	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-16	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-17	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-18	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-19	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-20	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-21	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-22	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-23	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-24	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-25	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-26	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-27	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-28	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-29	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-09-30	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-10-01	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-10-02	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-10-03	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-10-04	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-10-05	f	$175.00	$175.00	30	1125	$175.00
	9357	2022-10-06	f	$175.00	$175.00	30	1125	$175.00
	15341	2022-09-08	f	$205.00	$205.00	30	1125	$205.00
	15341	2022-09-09	f	$221.00	$221.00	30	1125	$221.00
	15341	2022-09-10	f	$216.00	$216.00	30	1125	$216.00

id	neighborhood	neighborhood_cl	bedrooms	bedrooms_clean	number_of_reviews_ltm	number_of_revie	top_listing	top_neighborho	top_propertysize	revenue_earned
82638	East Village	East Village	1	1	3	0	1	TRUE	TRUE	3549
9357	Hells Kitchen	Hells Kitchen		0	0	0	0	TRUE	FALSE	5250
57618	hells kitchen	Hells Kitchen	1	1	7	0	1	TRUE	TRUE	7500
57754	Flatiron District	Flatiron District	1	1	28	2	0	FALSE	TRUE	9570
57874	East Harlem	East Harlem	1	1	55	3	1	TRUE	TRUE	4320
15341	LOWER EAST S	Lower East Side	1	1	6	0	1	TRUE	TRUE	2729
59014	Lower East Side	Lower East Side	1	1	1	0	1	TRUE	TRUE	5750
59709	chinatown	Chinatown	2	2	23	2	0	FALSE	FALSE	8100
60164	Tribeca	Tribeca	3	3	3	0	0	FALSE	FALSE	3500
16974	East Harlem	East Harlem	4	4	26	0	0	TRUE	FALSE	7350
60611	East Harlem	East Harlem	1	1	1	1	1	TRUE	TRUE	3300
107895	Upper West Side	Upper West Side		0	2	0	0	TRUE	FALSE	4980
17037	East Village	East Village	1	1	0	0	1	TRUE	TRUE	6600
60680	East Village	East Village	1	1	2	0	1	TRUE	TRUE	4800
61509	Midtown	Midtown	1	1	0	0	0	TRUE	FALSE	3300
62427	East Village	East Village	1	1	1	1	1	TRUE	TRUE	3120
19159	Harlem	Harlem	2	2	0	0	0	TRUE	FALSE	3190
62891	East Village	East Village		0	9	2	0	TRUE	FALSE	4725
116551	Harlem	Harlem		0	2	0	0	TRUE	FALSE	2100
63693	east village	East Village	2	2	0	0	0	TRUE	FALSE	11700
23686	West Village	West Village	3	3	3	0	0	TRUE	FALSE	14250
123784	Harlem	Harlem		0	33	1	0	TRUE	FALSE	2739
126443	east village	East Village	2	2	60	4	0	TRUE	FALSE	12852
27644	Harlem	Harlem	1	1	18	4	1	TRUE	TRUE	4750
65425	Nolita	Nolita	2	2	10	3	0	FALSE	FALSE	3075
29683	Lower East Side	Lower East Side	1	1	2	0	1	TRUE	TRUE	5550
66251	East Village	East Village		0	32	1	0	TRUE	FALSE	9776
140973	East Village	East Village	2	2	4	1	0	TRUE	FALSE	6450
67288	East Harlem	East Harlem	1	1	7	0	1	TRUE	TRUE	3665
144087	Harlem	Harlem	2	2	7	0	0	TRUE	FALSE	9584
36703	West Village	West Village	1	1	5	0	1	TRUE	TRUE	6000
148825	UPPER WEST S	Upper West Side	2	2	36	2	0	TRUE	FALSE	5021
68974	Nolita	Nolita	3	3	11	0	0	FALSE	FALSE	17275

Start Here ▾   Assumptions + Change Log ▾   Data Dictionary ▾   RawData_Listings ▾   RawData_Calendar ▾   **PreparedData_Listings** ▾

bedrooms_clean	neighborhood_c	SUM of number_
1	Lower East Side	5399
	Hells Kitchen	2204
	Harlem	2126
	Upper West Side	1639
	East Village	1332
	Nolita	1263
	Chelsea	1095
	Midtown	908
	West Village	907
	East Harlem	899
	Upper East Side	885
	Washington Heig	432
	Chinatown	378
	Kips Bay	371
	Gramercy	299
	Inwood	266
	Financial District	218
	Murray Hill	208
	Ellis Island	206
	Soho	202
	Battery Park City	202
	Little Italy	148
	Greenwich Villag	117
	Tribeca	109
	Noho	109
	Flatiron District	52
	Two Bridges	38
	Theater District	37
	Central Park	21
	Stuyvesant Town	5
	Morningside Heig	4
	Roosevelt Island	1
	Civic Center	0
1 Total		22080
2	Hells Kitchen	2470

+ ≡   Data Dictionary ▾   RawData_Listings ▾   RawData_Calendar ▾   PreparedData_Listings ▾   **Pivot: Most Popular Property Sizes** ▾

neighborhood_clean	SUM of number_	
Lower East Side	6242	AC2="Lower East Side",
Hells Kitchen	5506	AC2="Hells Kitchen",
Harlem	5157	AC2="Harlem",
Midtown	4128	AC2="Midtown",
Upper West Side	3497	AC2="Upper West Side",
Chelsea	2913	AC2="Chelsea",
East Village	2572	AC2="East Village",
East Harlem	2175	AC2="East Harlem",
West Village	1735	AC2="West Village",
Upper East Side	1696	AC2="Upper East Side",
Nolita	1505	
Financial District	1112	
Chinatown	894	
Washington Heights	773	
Kips Bay	622	
Gramercy	573	
Soho	481	
Inwood	361	
Ellis Island	358	
Battery Park City	348	
Murray Hill	341	
Greenwich Village	311	
Little Italy	246	
Theater District	209	
Tribeca	173	
Noho	155	
Flatiron District	112	
Marble Hill	74	
Two Bridges	43	
Central Park	37	
Stuyvesant Town	5	
Morningside Heights	4	
Civic Center	2	
Roosevelt Island	1	
Grand Total	44361	

id	SUM of revenue
49946551	29940
44267714	26100
6833395	22800
46432776	20300
41295029	15600
53057804	15300
53058987	14425
4859903	14233
43118961	13601
4967114	13500
289665	13500
50618605	13050
1060019	13021
5800130148099…	12930
30087377	12810
25313748	12785
34384453	12609
49928900	12236
53167175	12225
53602040	12135
41890621	11970
37251520	11850
48374098	11610
52755593	11600
52336155	11600
39225271	11575
29687597	11571
13058232	11500
29495933	11455
21366948	11268
6640144	11250
4644669	11250
258688	11250
46351532	11111
52257683	11094

+ ☰ | RawData_Calendar ▾ | PreparedData_Listings ▾ | Pivot: Most Popular Property Sizes ▾ | Pivot: Most Desirable Neighborhoods ▾ | **Pivot: Top Listings** ▾