AN ANALYSIS OF CURRENT INVENTORY IN:
NEW YORK CITY

RACHELLE PEREZ

SCOPE OF PROJECT

- Spreadsheet with 48,895 Listings
- 16 Data Points Per Listing
- TASK: **Exploratory Analysis**
  1. Formulate questions & hypotheses
  2. Answer these with multiple tools
     - Visualizations
     - Statistical Analysis (including p-value tests)
     - Query-based insights (SQL)
  3. Identify trends & areas of concern
  4. Establish relationships between variables

DATA POINTS

Listing ID
Name
Host ID
Host Name
Borough (Neighborhood Group)
Neighborhood
Latitude
Longitude
Room Type
Price
Minimum Nights
Number of Reviews
Last Review Date
Reviews Per Month
Number of Listings per Host
Availability (Last 365 Days)

Let’s look at:

AVAILABILITY

Why?

Limited availability can lead to **customer friction** and reduce **conversion.**

A review of New York City’s inventory can provide insight into possible **host behavior patterns** of similar and/or emerging cities.

ACTIVE VS. INACTIVE

CLAIM 1: Majority of Listings (75%) are available at least 1 day / year

Result of Hypothesis Test: Claim **rejected** with 95% confidence.
Unexpectedly, 36% of listings are not available ("inactive").

What IS available?

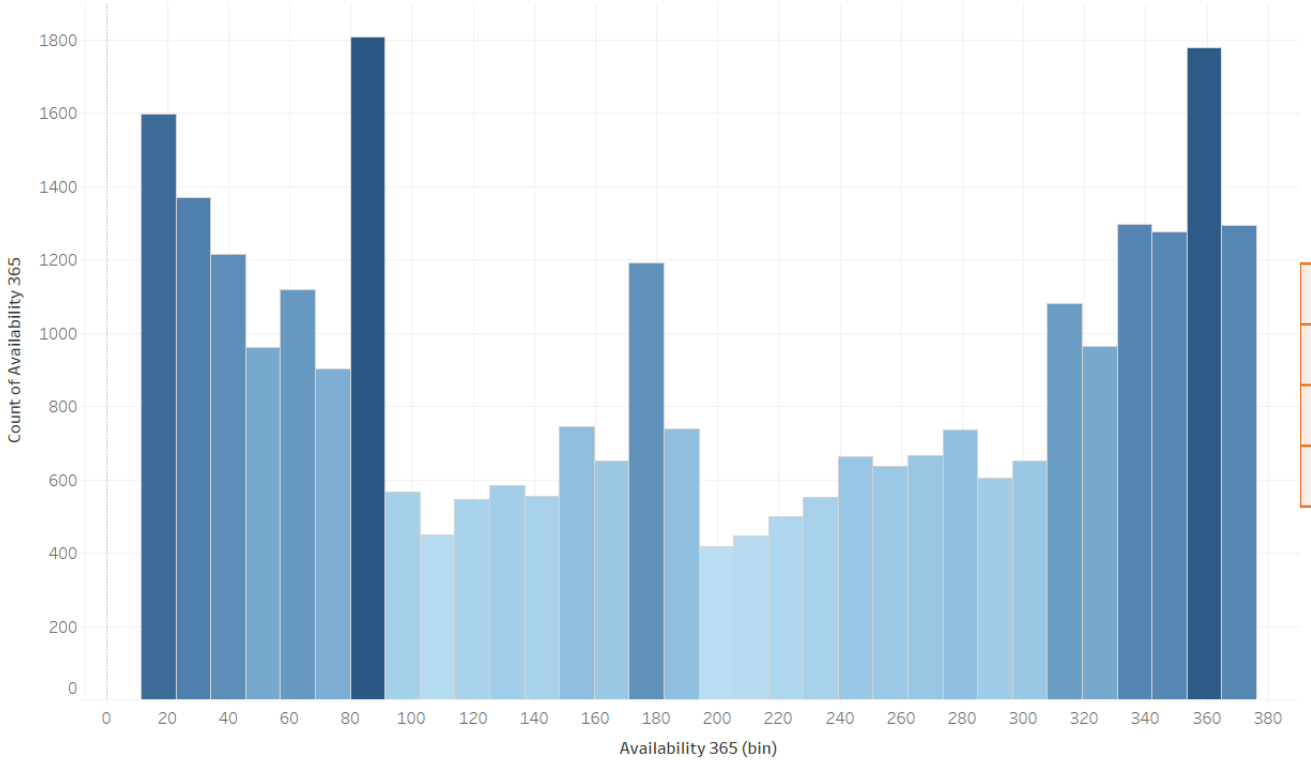
Status	Number of Listings	% of Total
Active (Available 1+ Days)	31,362	64%
Inactive (Available 0 Days)	17,533	36%
TOTAL	48,895	100%

HOW AVAILABLE IS A LISTING?

CLAIM 2: 50% of active are available 50+% of Year (183 days)

Result of Hypothesis Test: Claim **rejected** with 95% confidence

46% of Listings are available at least half the year



Availability	Listing Count	% of Total
Less Than Half the Year (1-182)	17,052	54%
More Than Half The year (183-365)	14,310	46%
TOTAL ACTIVE	**31,362**	100%

COMPARING LEVELS OF AVAILABILITY

Define Attribute: Availability Level

Availability (Days)	Attribute: Availability Level
0	Unavailable
1 -182 (Less than 50% Year)	Low Availability
183 - 365 (More than 50% Year)	High Availability

Proportion of Total Listings

Attribute: Availability Level	Listing Count (Source: SQL)	% of Total
Unavailable	17,533	36%
Low Availability	17,052	35%
High Availability	14,310	29%
TOTAL	**48,895**	**100%**

HOST PATTERNS

Average number of listings per host: **10.34** *
For **multiple-listing** hosts (at least 2), average is: **21.96***

Total Number of Hosts: 21,809*
Total Multiple-Listing hosts: 4,416*
20% of Hosts have multiple listings

Total Listings from Multiple-Listing hosts: 13,969
Total Active Listings: 31,362
45% of listings from multiple-listing hosts

** Source: SQL, Inactive Listings (Availability_365 = 0) Removed*

CONCLUSIONS

Conclusion	Potential Application / Next Steps
29% of Total Inventory is Inactive	**Why?** Does it take **less resources to re-activate a listing** or source a new one?
Less than half of the Active(46%), is available over 6 months of the year	What is the **effect on customer experience?** Should we review **resource allocation** to encouraging more availability from existing listings?
The average listing count per host is 10 list, but it goes up to 21 for multiple-listing hosts.	Indication of **commercial usage** in market. How are we dedicating resources in **host acquisition** to attract different types of hosts? A bigger question, **is this in line with vision?**
20% of hosts have multiple listings. These hosts control 45% of the inventory.	

Are these observations unique to New York City or can they do they apply to similar and/or emerging cities?

TOOLS USED:

- MICROSOFT EXCEL
- STRUCTURED QUERY LANGUAGE (SQL)
- TABLEAU

MORE INFORMATION ON THIS PROJECT:

- GITHUB
- TABLEAU PUBLIC
- LINKEDIN