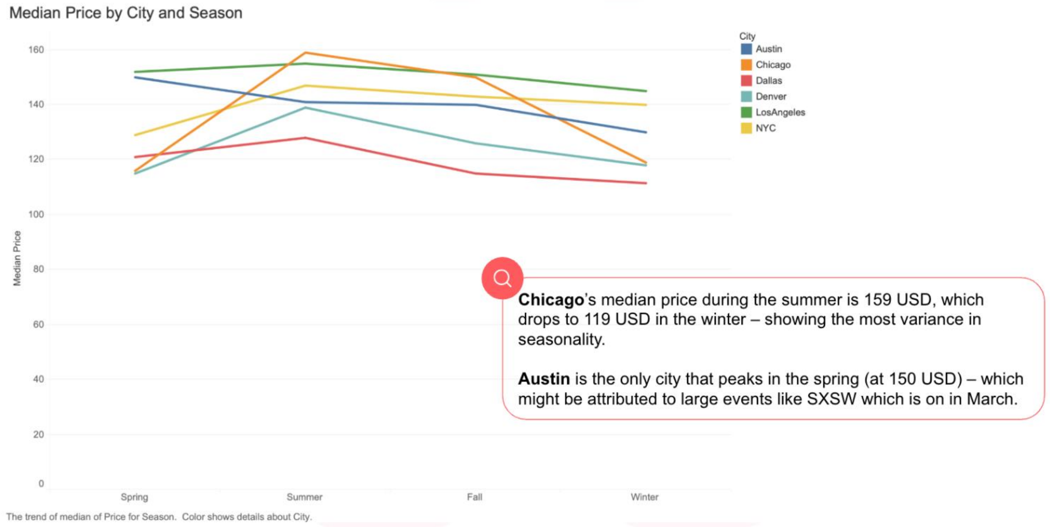
Right Message, Right Audience: Optimizing Airbnb’s Campaign with Segmentation

Kirana Ananda

Q1'25 Recap

Seasonal demand influences shifts in **pricing trends**, this stood out particularly in



'Room type' and **'accommodates'** have the highest influence on price across the key cities

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
season	3	359	120	266.25	< 2e-16	***
city	5	1507	301	670.54	< 2e-16	***
room_type	3	37435	12478	27760.85	< 2e-16	***
season:city	15	453	30	67.22	< 2e-16	***
season:room_type	9	87	10	21.49	< 2e-16	***
city:room_type	15	999	67	148.13	< 2e-16	***
season:city:room_type	44	57	1	2.88	6.07e-10	***
Residuals	273671	123013	0			

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6704 on 273671 degrees of freedom
Multiple R-squared:  0.2495,	Adjusted R-squared:  0.2493
F-statistic: 967.9 on 94 and 273671 DF,  p-value: < 2.2e-16
```

The new model (+accommodates) improves prediction accuracy by 22%.

Location-based marketing strategy, proof of concept starting out with Chicago

Persona-based marketing strategy, creation of distinct personas based on additional variables that reflect stay characteristics.

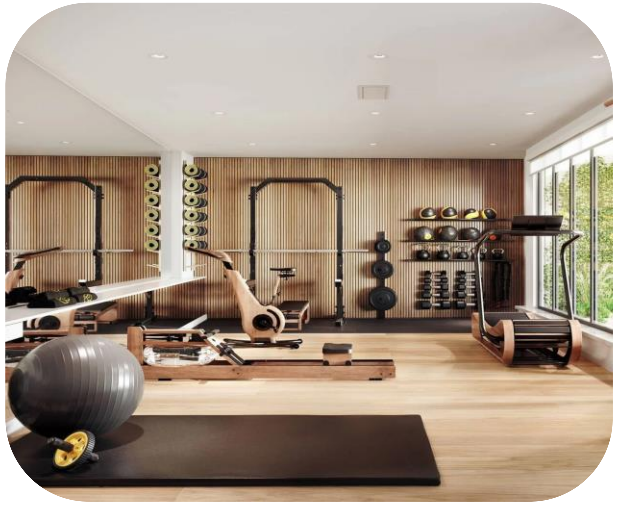
Q2’25 strategic question:
Which of these properties is most appealing to you?

1

2

3

4

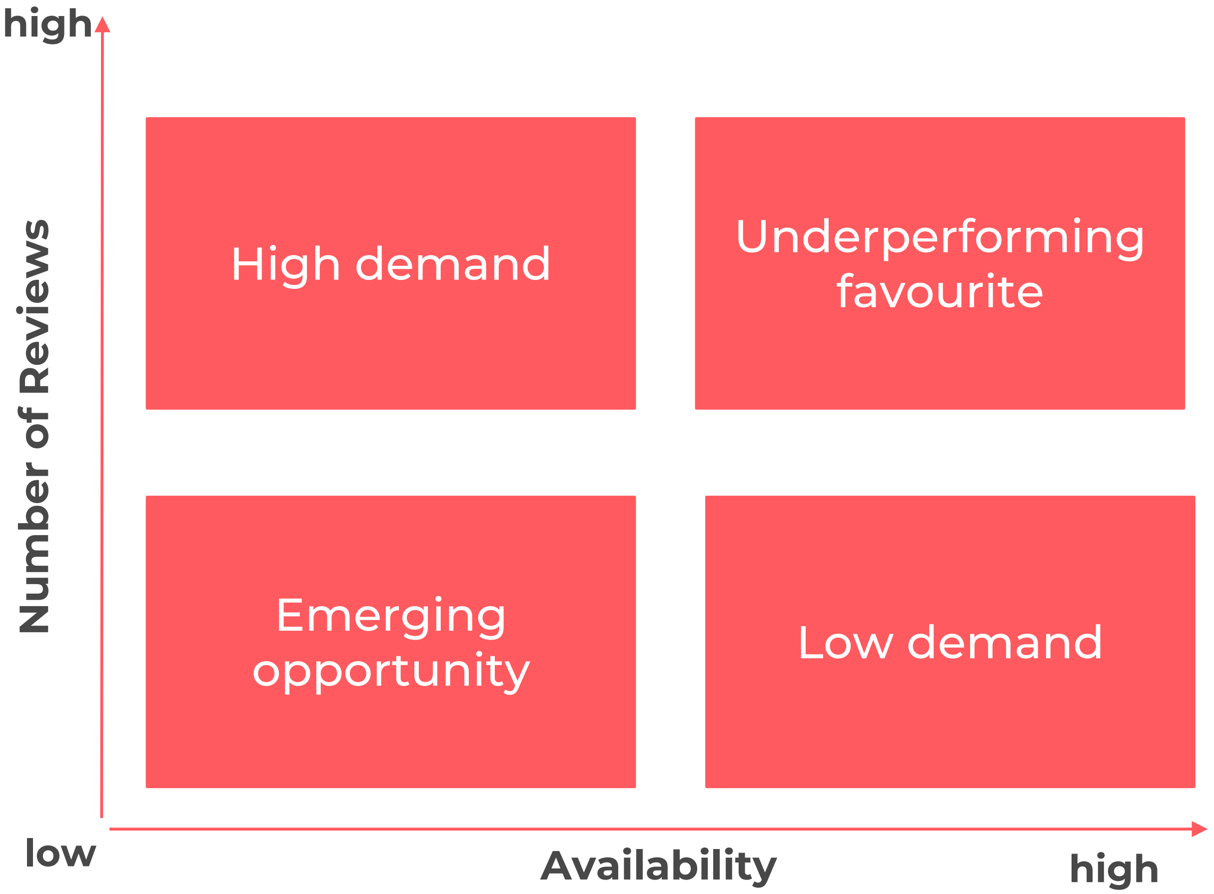


5

Data set in the analysis

1. City and neighborhood group
2. Price and price cluster (k-means)
3. Season
4. Room type and accommodates
5. Amenities
6. Text analysis
    - Review dataset
    - Neighborhood overview
7. Demand, a proxy based on availability_365 and number_of_reviews_ltm

Market Map



Data preparation

1 Restructure the ‘amenities’ column and exclude amenities type <100 occurrences.

```
[45]: # 3. More efficient approach to check amenities against keywords
for category, keywords in category_keywords.items():
    # Convert keywords to lowercase for case-insensitive matching
    keywords_lower = [keyword.lower() for keyword in keywords]

    for col in amenity_columns:
        # Get values from the amenity column, convert to lowercase
        amenity_values = amenities_expanded[col].fillna("").astype(str).str.lower()

        # Find rows where the amenity value matches any keyword in this category
        for keyword in keywords_lower:
            # Find matches where this keyword appears in the amenity value
            matches = amenity_values.str.contains(keyword, regex=False)
            # Set category flag to 1 for matches
            category_flags.loc[matches, category] = 1
```

```
[47]: category_flags
```

[47]:

	Safety	WiFi	Kitchen	General Facilities	Bathroom	Clothing storage	Laundry at home	Heating & cooling system	Kitchen utensils	Bedroom	...	Recreation spot direct access	Waterfront	Kitchenette	Playground	Outdo kitch
0	1	1	1	1	1	1	1	1	1	1	...	0	0	0	0	
1	1	0	1	1	1	1	1	1	1	1	...	0	0	0	0	
2	1	0	1	1	1	1	1	1	1	1	...	0	0	0	0	
3	1	1	1	1	1	1	1	1	1	1	...	0	0	0	0	
4	1	1	1	1	1	1	1	1	1	1	...	0	0	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
469524	1	1	1	0	0	0	1	1	0	0	...	0	0	0	0	
469530	1	1	1	0	0	0	1	0	0	0	...	0	0	0	0	
469552	1	1	0	1	1	0	1	1	1	1	...	0	0	0	0	
469553	1	1	0	1	1	0	1	1	1	1	...	0	0	0	0	
469554	1	1	1	1	1	1	1	1	1	1	...	0	0	0	0	

153833 rows × 40 columns

2 Qualitatively categorize the ‘amenities’ from 514 to 40.

469555 rows × 3 columns

```
[25]: cleaned_df = cleaned_df.drop_duplicates(subset=['id'])
```

```
[27]: cleaned_df
```

[27]:

	id	city	amenities
0	69810	Austin	["Cooking basics", "Bed linens", "Backyard", "...
1	70812	Austin	["Cooking basics", "Bed linens", "EV charger",...
2	72833	Austin	["Courtyard view", "Cooking basics", "Bed line...
3	73005	Austin	["Clothing storage", "Central air conditioning...
4	73289	Austin	["Central air conditioning", "Cooking basics",...
...	...	...	...
469524	1112482646283343030	Chicago	["Carbon monoxide alarm", "Fire extinguisher",...
469530	1114571672303000146	Chicago	["First aid kit", "Carbon monoxide alarm", "Fi...
469552	1117438873812832956	Chicago	["Private entrance", "TV with standard cable",...
469553	1117439997700638787	Chicago	["Private entrance", "TV with standard cable",...
469554	1117896281364781567	Chicago	["Ceiling fan", "Fire pit", "Heating", "Lockbo...

153833 rows × 3 columns

3 Regrouping 40 to 27 amenities to ease clustering process.

```
[53]: count_ones

[53]: Safety                            147997
      WiFi                              143285
      Kitchen                           138209
      General Facilities                135235
      Bathroom                          132261
      Clothing storage                  117262
      Laundry at home                   140071
      Heating & cooling system          149646
      Kitchen utensils                  119062
      Bedroom                           100320
      Workspace                          91182
      Entertainment                     132356
      Check-in                          113995
      Parking                           116507
      Pets                               49232
      Outdoor activities                 54262
      BBQ                                37427
      Elevator                           25233
      Patio or balcony                   52023
      Backyard                           42511
      Pool                               29257
      Fireplace                          23026
      Housekeeping                       17841
      Sports                             21073
      Gym                                20990
      Family needs                       32698
      Hot tub                            16663
      View                               20861
      Lounge                              8139
      EV charger                          6643
      Recreation spot direct access       9814
      Waterfront                          2333
      Kitchenette                         1879
      Playground                          1836
      Outdoor kitchen                     3187
      Movie theater                       1334
      Spa and wellness                    1706
      Water sports                         844
      Mini golf                            444
      Rooftop                              439
      dtype: int64
```

Criteria: grouping together if > 100k

4 Resampling n=5000 per cities (LA, NYC, Austin, Chicago, Denver, Dallas) to enter cluster analysis.

Cluster analysis

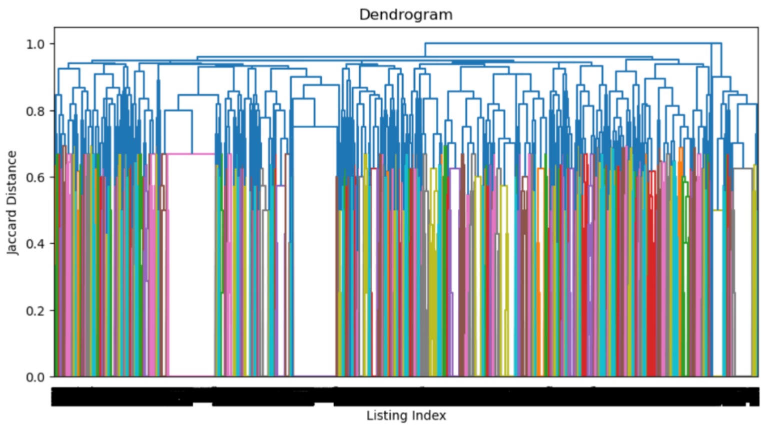
1 Cluster analysis with complete method using pairwise Jaccard distances*

```
[161]: X = check.drop(columns=['id', 'city'])  # only binary features used for clustering

[163]: # Condensed distance array
       dist_array = pdist(X, metric='jaccard')

[165]: # Hierarchical clustering
       linkage_matrix = linkage(dist_array, method='complete')

       # Plot dendrogram
       plt.figure(figsize=(10, 5))
       dendrogram(linkage_matrix)
       plt.title("Dendrogram")
       plt.xlabel("Listing Index")
       plt.ylabel("Jaccard Distance")
       plt.show()
```



2 Cross-checking the generated cluster**

7 (n=32) is outlier because this cluster shows no meaningful amenities.

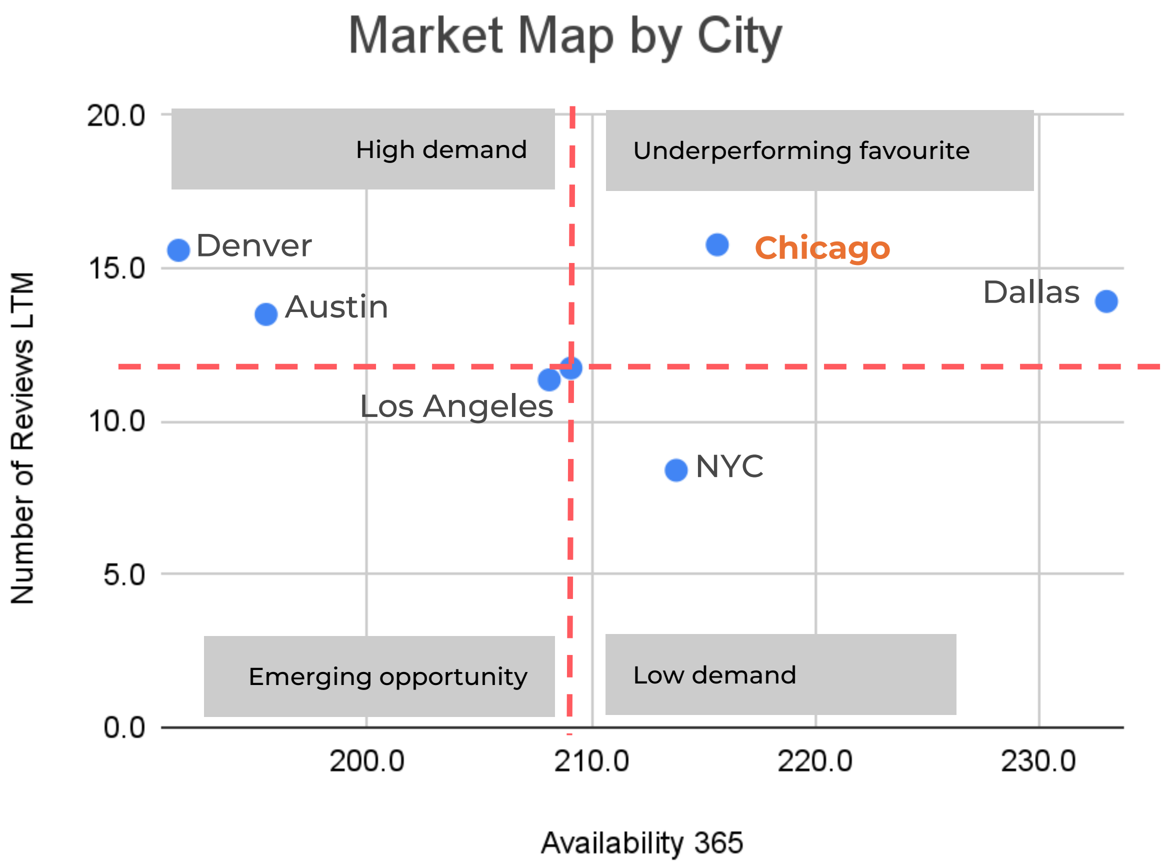
Row Labels	Grand Total	Family retreat	Workplace serenity	Patio paradise	Backyard bliss	Active leisure	BBQ escape	7	Pet haven
Basic airbnb	29966	2.72	4.09	3.45	2.79	1.76	1.95	#DIV/0!	3.08
Rooftop	172	-0.82	-0.54	-0.74	-0.75	-0.97	-1.00	#DIV/0!	-0.46
Mini golf	124	-0.81	-0.55	-0.74	-0.74	-1.22	-0.95	#DIV/0!	-0.46
Water sports	224	-0.81	-0.54	-0.74	-0.74	-1.11	-0.92	#DIV/0!	-0.46
Spa and wellness	314	-0.82	-0.54	-0.72	-0.72	-0.99	-0.94	#DIV/0!	-0.46
Movie theater	323	-0.81	-0.53	-0.72	-0.73	-1.05	-0.91	#DIV/0!	-0.45
Outdoor kitchen	705	-0.80	-0.55	-0.74	-0.67	-0.67	-0.81	#DIV/0!	-0.46
Playground	425	-0.72	-0.52	-0.70	-0.73	-1.16	-0.87	#DIV/0!	-0.46
Waterfront	487	-0.72	-0.53	-0.71	-0.71	-1.02	-0.88	#DIV/0!	-0.46
Recreation spot direct access	1819	-0.26	-0.49	-0.52	-0.60	-0.64	-0.65	#DIV/0!	-0.46
EV charger	1580	-0.65	-0.39	-0.52	-0.67	-0.30	-0.70	#DIV/0!	-0.43
Lounge	1851	-0.63	-0.51	-0.67	-0.59	-0.23	-0.34	#DIV/0!	-0.45
View	5389	0.22	-0.43	-0.43	0.06	0.20	0.27	#DIV/0!	-0.43
Hot tub	2430	-0.72	-0.31	-0.54	-0.47	-0.38	-0.41	#DIV/0!	-0.40
Family needs	8415	1.14	0.42	0.14	0.10	-0.03	0.40	#DIV/0!	-0.36
Gym	4315	-0.76	0.22	0.26	-0.58	1.54	-0.66	#DIV/0!	-0.35
Sports	4625	-0.56	-0.26	0.06	-0.29	1.09	-0.17	#DIV/0!	-0.05
Housekeeping	4073	1.20	-0.42	-0.61	-0.49	-0.38	-0.25	#DIV/0!	-0.44
Fireplace	5241	0.14	-0.01	-0.30	-0.21	-0.84	0.38	#DIV/0!	-0.37
Pool	6052	-0.52	0.05	0.95	-0.32	1.24	0.03	#DIV/0!	-0.07
Backyard	11386	0.50	-0.31	0.60	1.69	-0.51	1.32	#DIV/0!	-0.20
Patio or balcony	13818	1.30	-0.08	2.04	1.25	1.05	1.83	#DIV/0!	-0.12
Elevator	4538	-0.65	0.63	-0.29	-0.61	1.22	-0.77	#DIV/0!	-0.26
BBQ	9194	-0.03	-0.22	0.51	0.63	0.92	1.27	#DIV/0!	0.04
Outdoor activities	13413	1.07	-0.19	0.09	1.81	1.37	1.71	#DIV/0!	-0.05
Pets	10651	0.76	0.31	0.15	0.57	-0.30	0.29	#DIV/0!	3.09
Workspace	20802	2.04	2.21	1.43	1.74	1.41	1.76	#DIV/0!	1.90

* I tried ward and average method too, but ward is not suitable for dummy data and average produce monoton cluster. Complete method is chosen because I want to avoid merging very distant outliers.

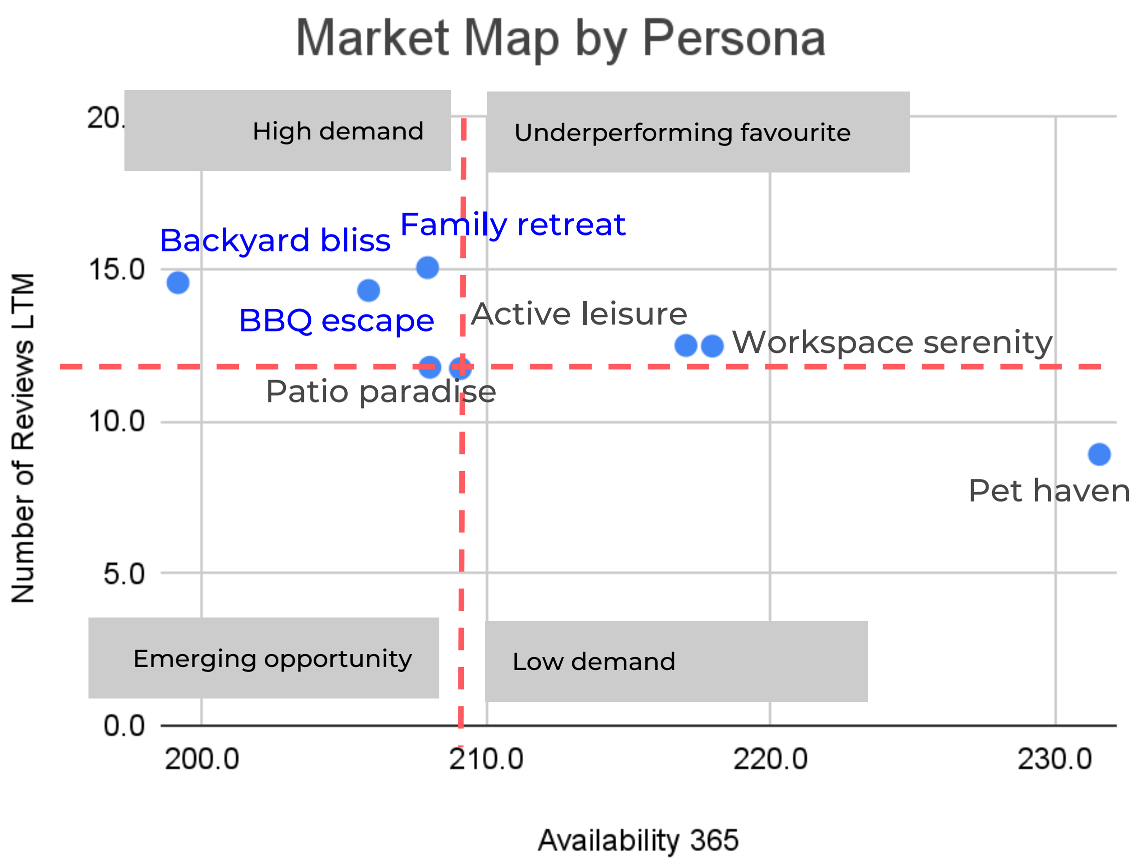
** I tested 8 to 10 clusters, and 8 proved best, as it showed clearer distinctiveness. With more than 8, clusters have some similar strengths judged by their z-score.

Market Overview

1 Baseline for location-based marketing



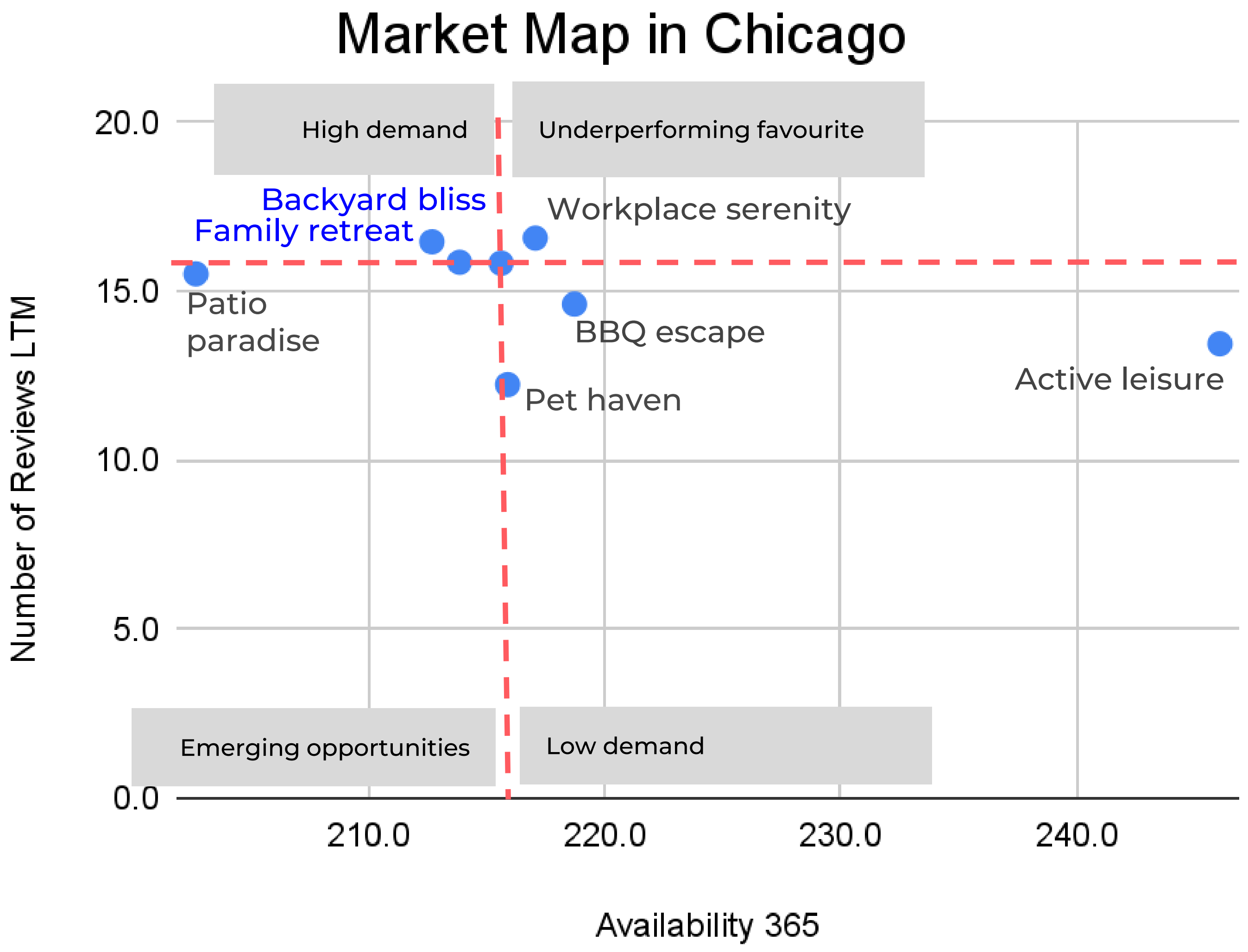
2 Baseline for persona-based marketing



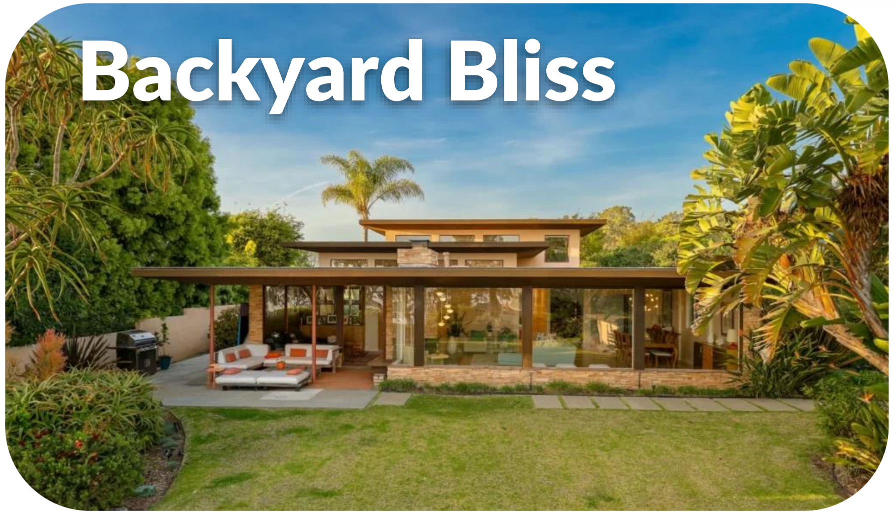
Chicago Overview

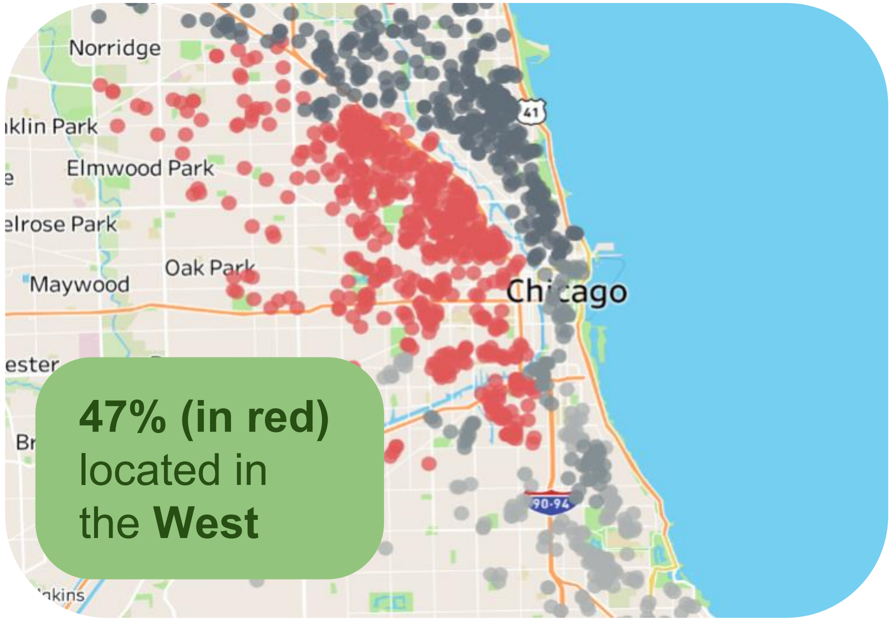
Mission

Create a segmented marketing strategy, focusing on the high demand personas to lift Chicago from an ‘underperforming favorite’ to a ‘high demand’ city

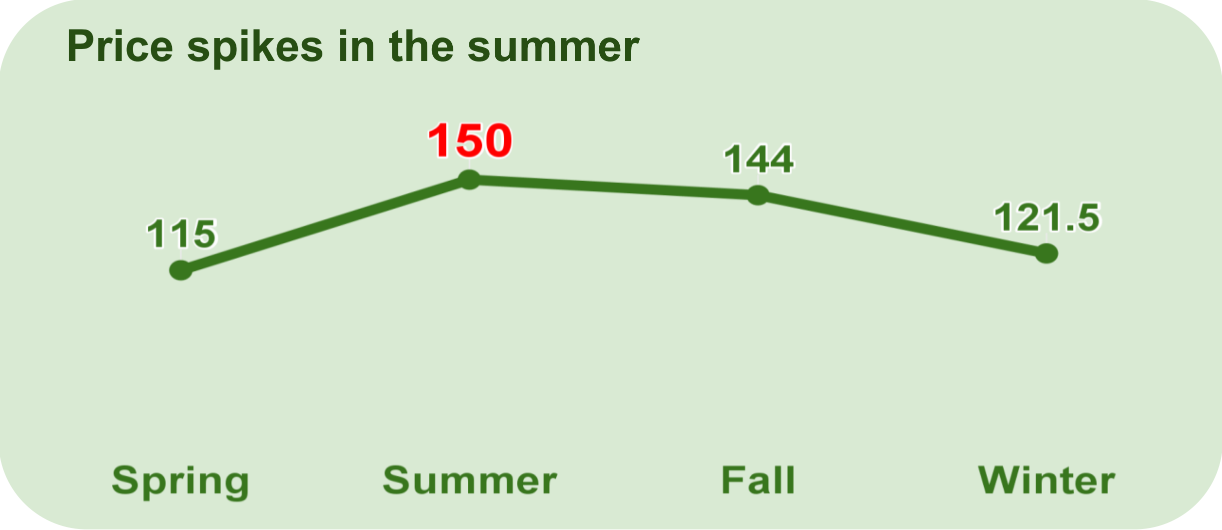


Persona





Median Price: 132 USD; 74% under 200 USD per night



76% entire home/apt

Accommodates **4**

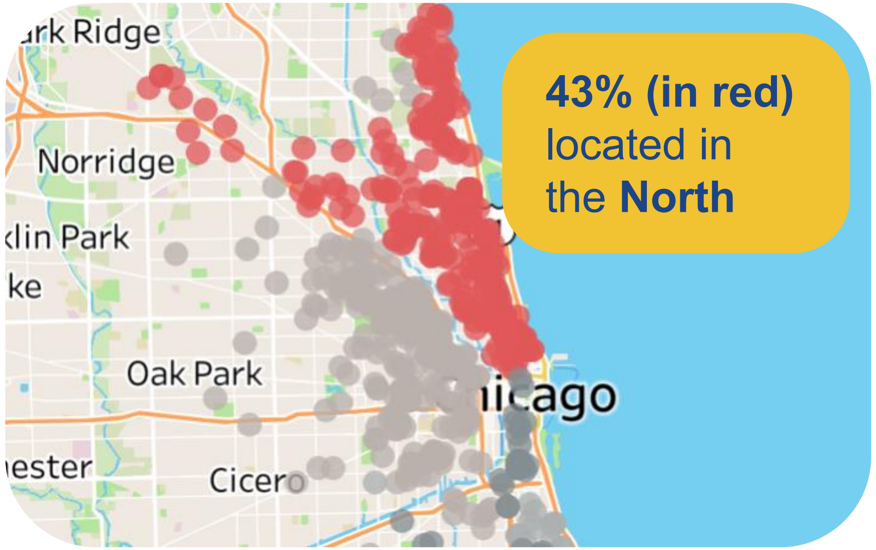
"Serenity"

Key amenities:

- Backyard
- Outdoor activities
- Patio/balcony view

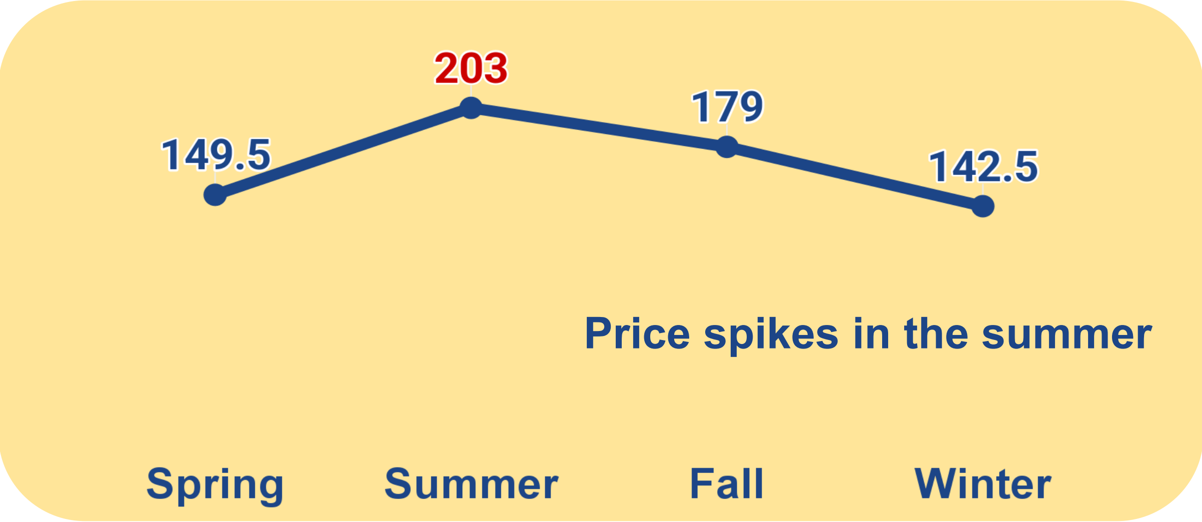
Persona

Family Retreat



43% (in red) located in the North

Median Price: 165 USD; 60% under 200 USD per night



Price spikes in the summer

82%
entire home/apt

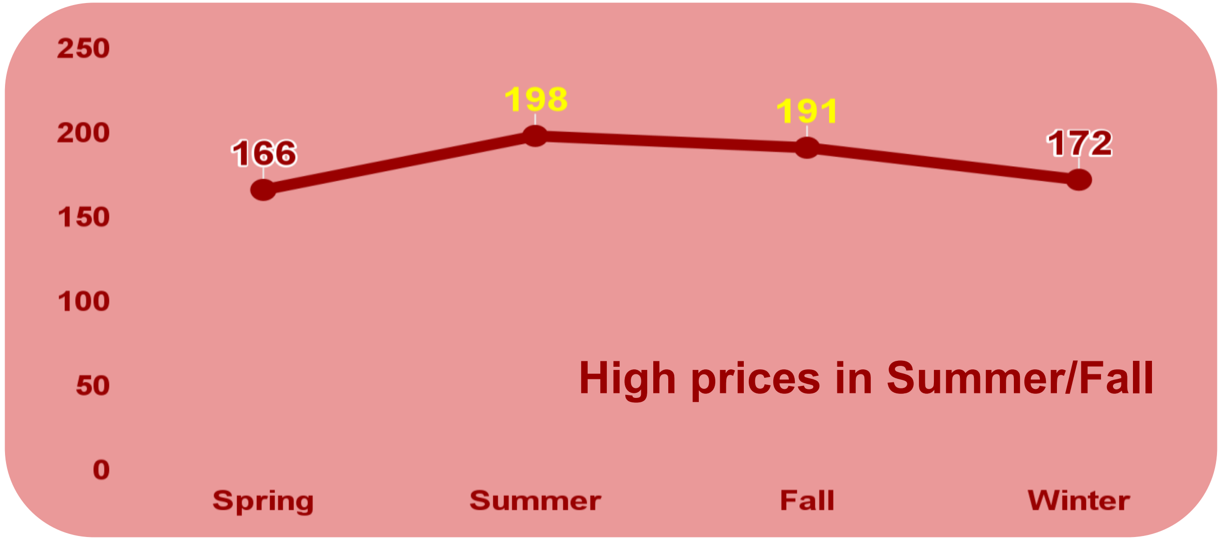
Accommodates
5

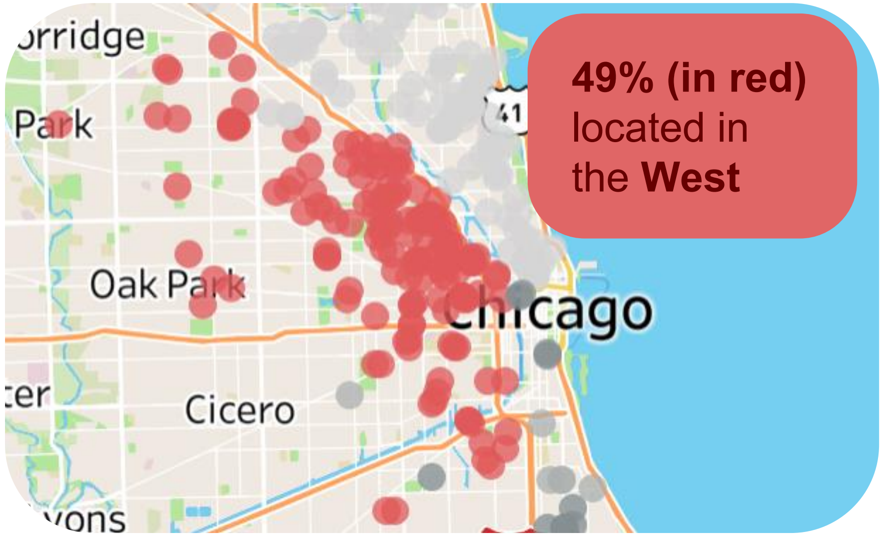
"Great"

Key amenities:
- Housekeeping
- Playground
- Direct access to recreation spot

Persona

Median Price: 181 USD; 56% under 200 USD per night





49% (in red) located in the **West**

81% **entire home/apt**

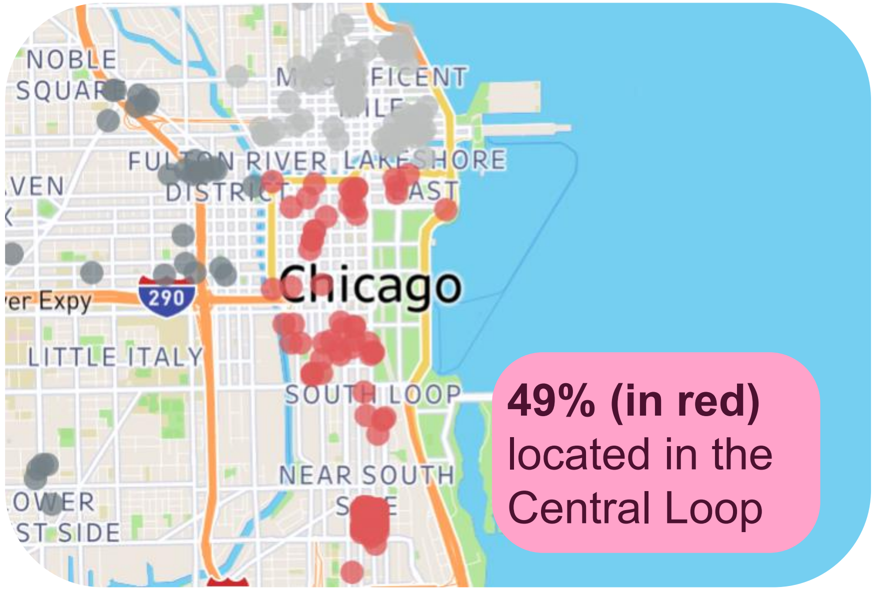
Accommodates **5**

"Comfortable"

Key amenities:
- **BBQ**
- **Lounge**

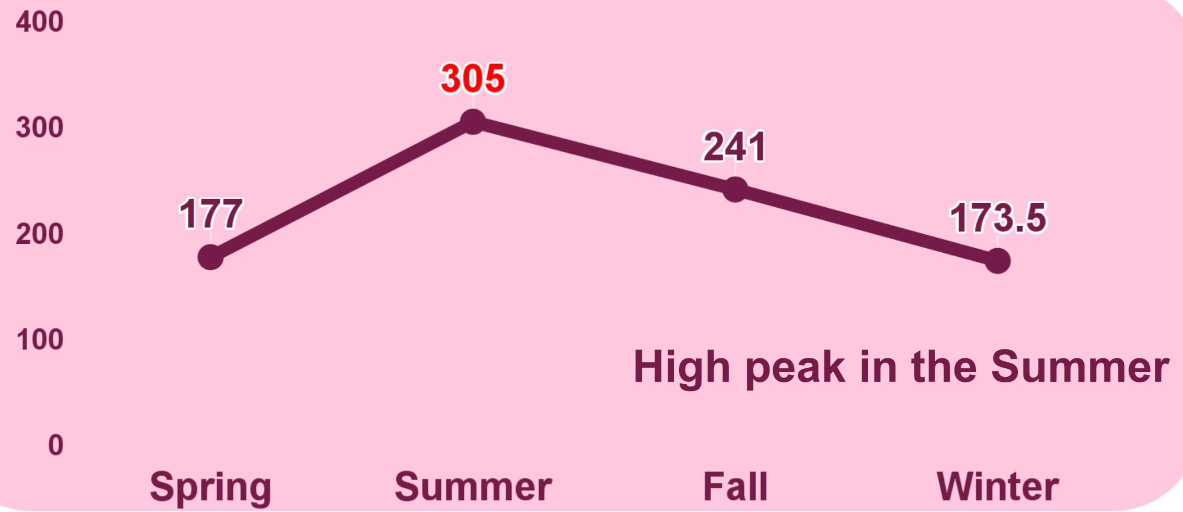
Persona

Active Leisure



49% (in red) located in the Central Loop

Median Price: 213.5 USD; 45% under 200 USD per night



99% entire home/apt

Accommodates **5**

Key amenities:

- Sports
- Hot tub
- Pool
- Spa and Wellness

Content Plan

- Seasonal blog series highlighting key trends, attractions, activities
- Omni-channel social campaigns with UGC highlighting seasonal experiences

Spring

Workspace serenity

The cheapest type of property, ideal for the ‘work from anywhere’ productive escape

Family retreat

Family outdoor fun, sunny escapes, taking advantage of the summer attractions

Summer

Backyard bliss

Experience ‘summer in the city’ with your friends and family

BBQ Escape

Perfect for a summer staycation

Fall

Patio paradise

Book early for an unforgettable fall breeze in Chicago!

Winter

Active leisure

Stay active in Winter with the best price!

Thank You

Kirana Ananda