

comparebnb

A NEW WAY TO SEARCH AND COMPARE
YOUR HOME AWAY FROM HOME

Miguel Gutierrez, Ashourrina Henshaw,
Yue Shen, Clara Zhang



01

Who we are. What we do.

comparebnb was designed with the traveler in mind.

From searching for a rental for your business trip to finding a room to share for one night to needing short term housing, our site allows you to easily search and compare all rentals matching your needs so you don't over pay.



comparebnb was built to take the stress out of traveling.

We believe that you should never have to over pay to experience the joy of traveling and that includes your travel accommodations.

Save money, live more.

Agenda

OUR PROCESS

WEBSITES

p. 02-06

p. 21-25

EDA DEEP DIVE

INFOGRAPHICS & POSTER

p. 07-12

p. 26-30

TABLEAU ANALYSIS

SUMMARY

p. 13-20

p. 31-40



EDA Process

02

In the process of EDA, we are trying to perform an initial investigations on data so as to discover patterns, to spot anomalies and to check assumptions with the help of summary statistics and graphical representations

PHASE 1

COMPILING/LOADING

We downloaded the data from the Inside Airbnb website. There are three files in csv format, which are listings, calendar and reviews.

PHASE 2

CLEANING

We cleaned up the data by cleaning null values and columns, detecting and dropping outliers, and transferring categorical variables into numeric format.

PHASE 3

ANALYZING

After data visualisations, we had a chance to understand the distribution and correlations of the variables within the dataset.

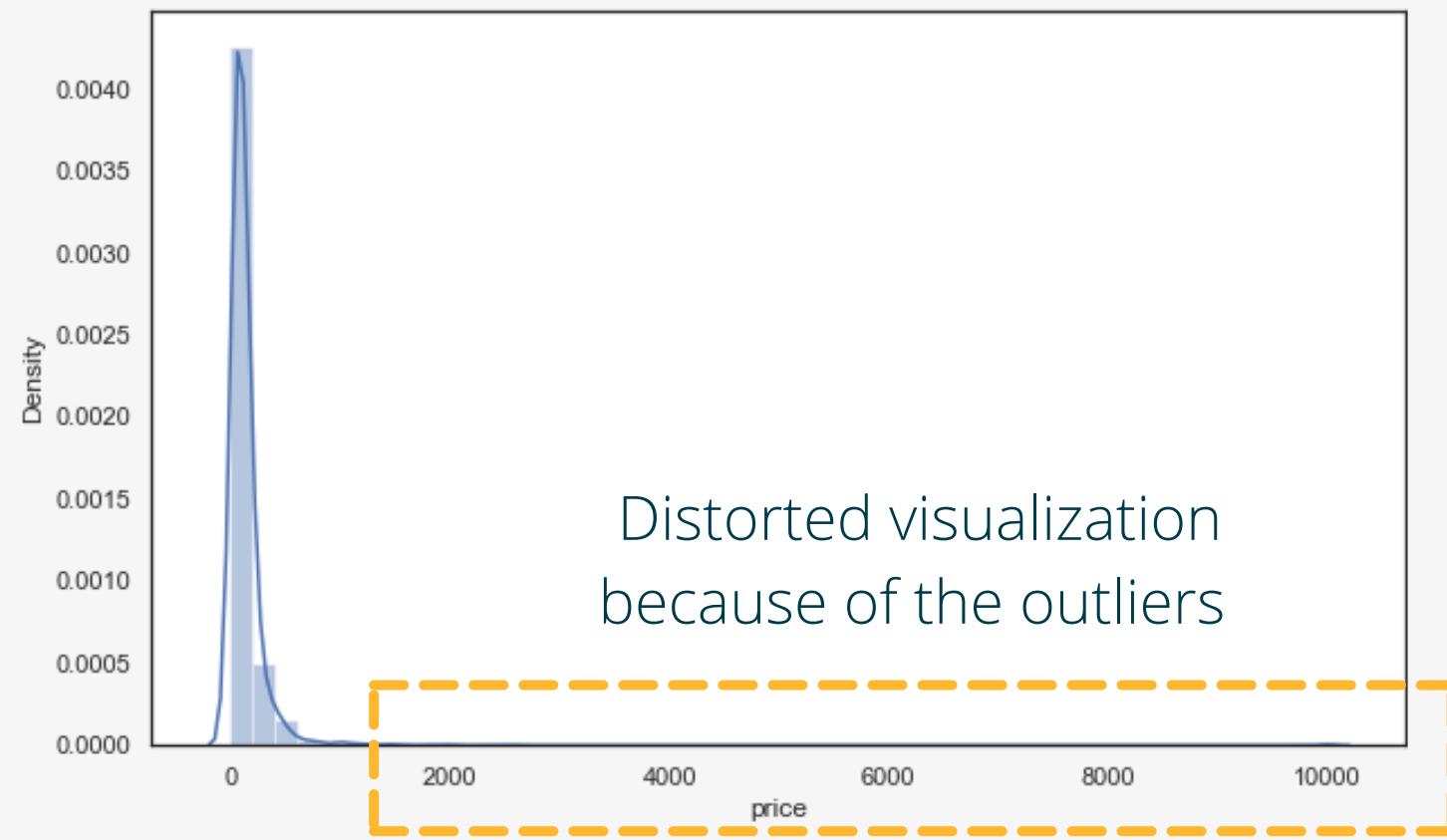
PHASE 4

RESULTS

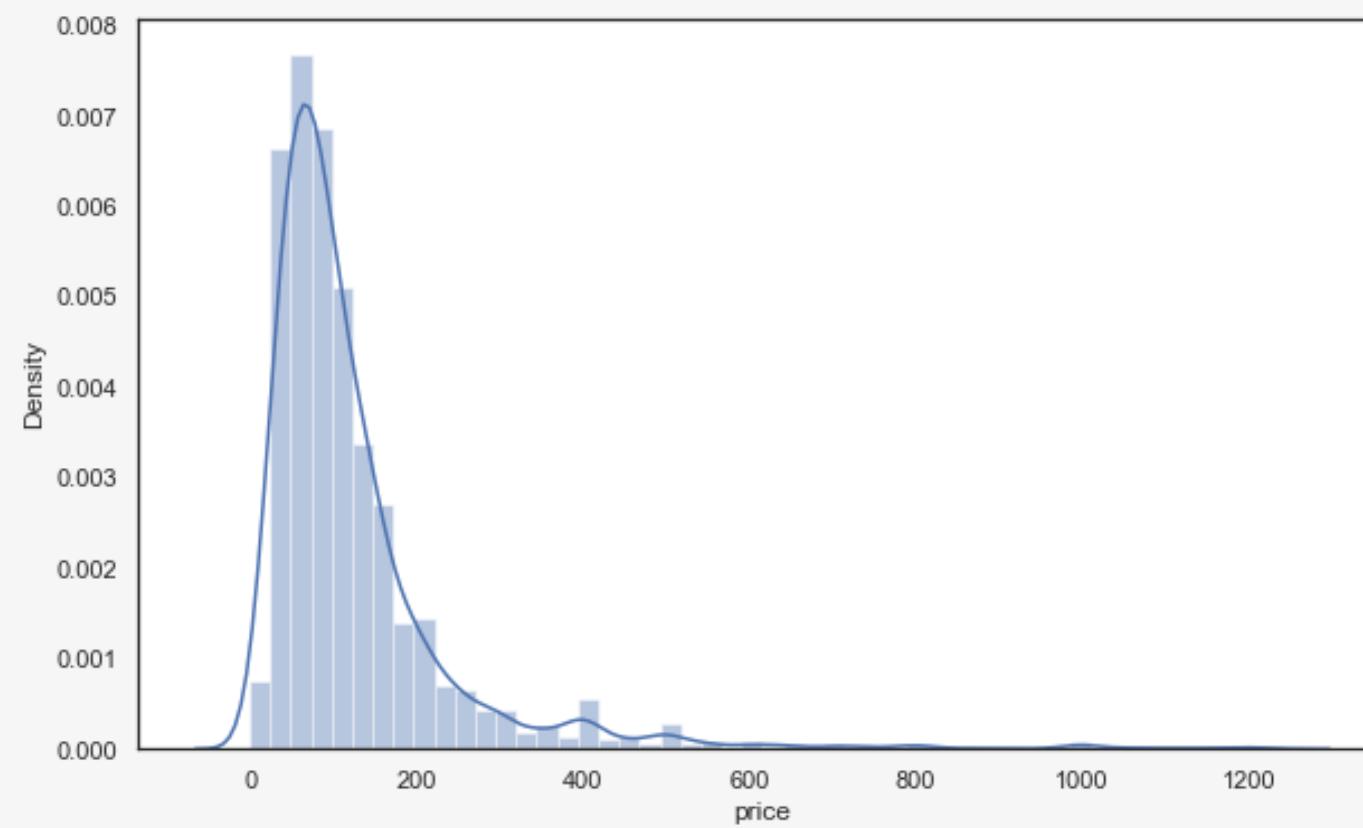
From the EDA process, we have a basic understanding of the dataset, and the dataset is ready for further building analytical models and visualisations.



Price Distribution before Dropping Outliers



Price Distribution after Dropping Outliers



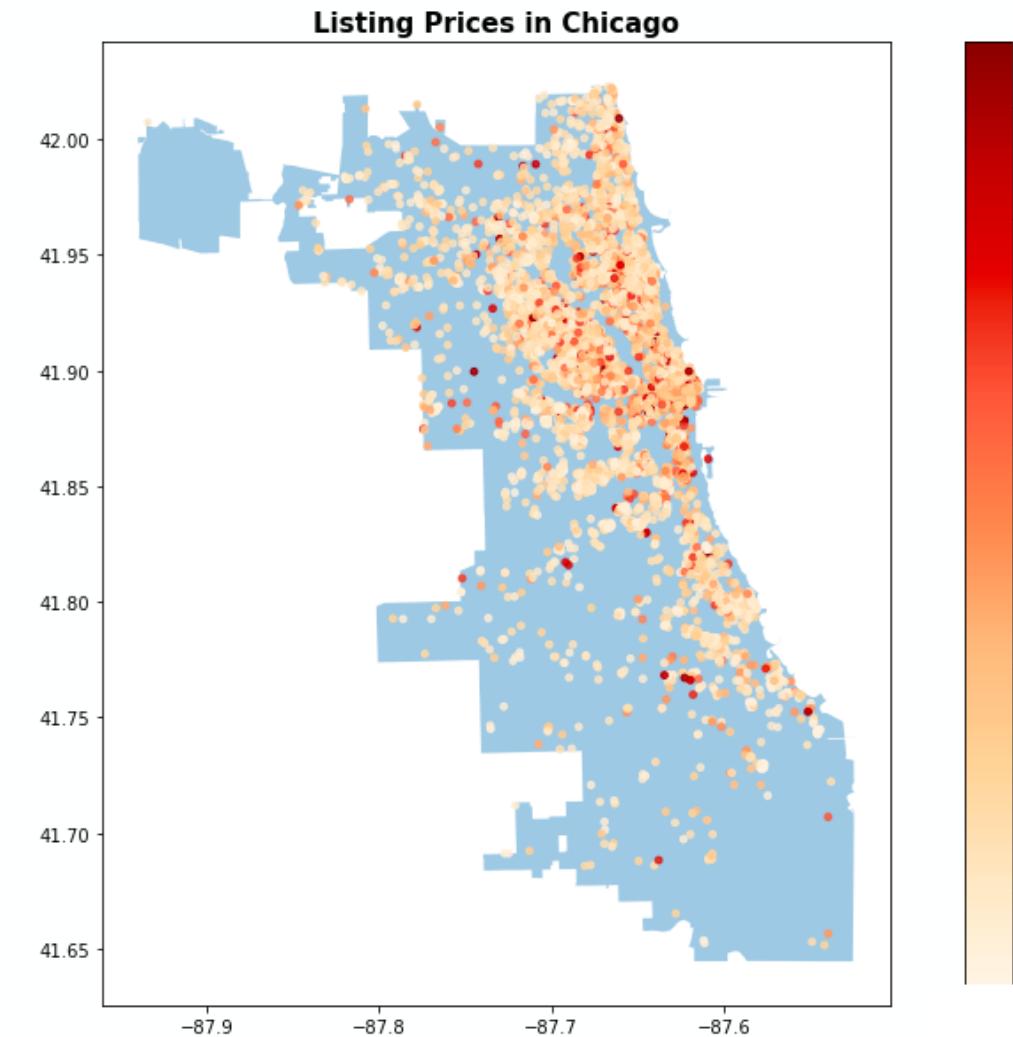
LOADING & CLEANING

- The datasets are loaded from the Inside Airbnb website. The three files contain detailed **listing**, **calendar** and **review** data for listing homes in Chicago.
- Dataset attributes:
 - Listings file: 16 columns, 6372 rows;
 - Calendar file: 7 columns, one million+ rows;
 - Review file: 6 columns, 267k+ rows;
 - **Key output feature: listing price.**
- An initial investigation shows that there is **no duplicated values** in the dataset, but there are both null columns and null values, so we **dropped the null columns** and **fill the null values** in the numeric columns with number 0.
- The investigation also indicates that there are outliers in the datasets, so we **dropped outliers with z-score greater than 3.**

ANALYZING

• Step 1: Exploring Distributions

- Plot price distribution
 - Show statistic summary with `data.describe()`
 - Visualize the price distribution on the Chicago city map
 - Create a word-cloud to show the keywords appeared in the reviews comments



- ▶ The listing price on map shows that most of listings are located in the urban area of Chicago, and more listings with high prices are located in the Northern and Central areas(downtown) of Chicago.

WordCloud of Comments in Reviews

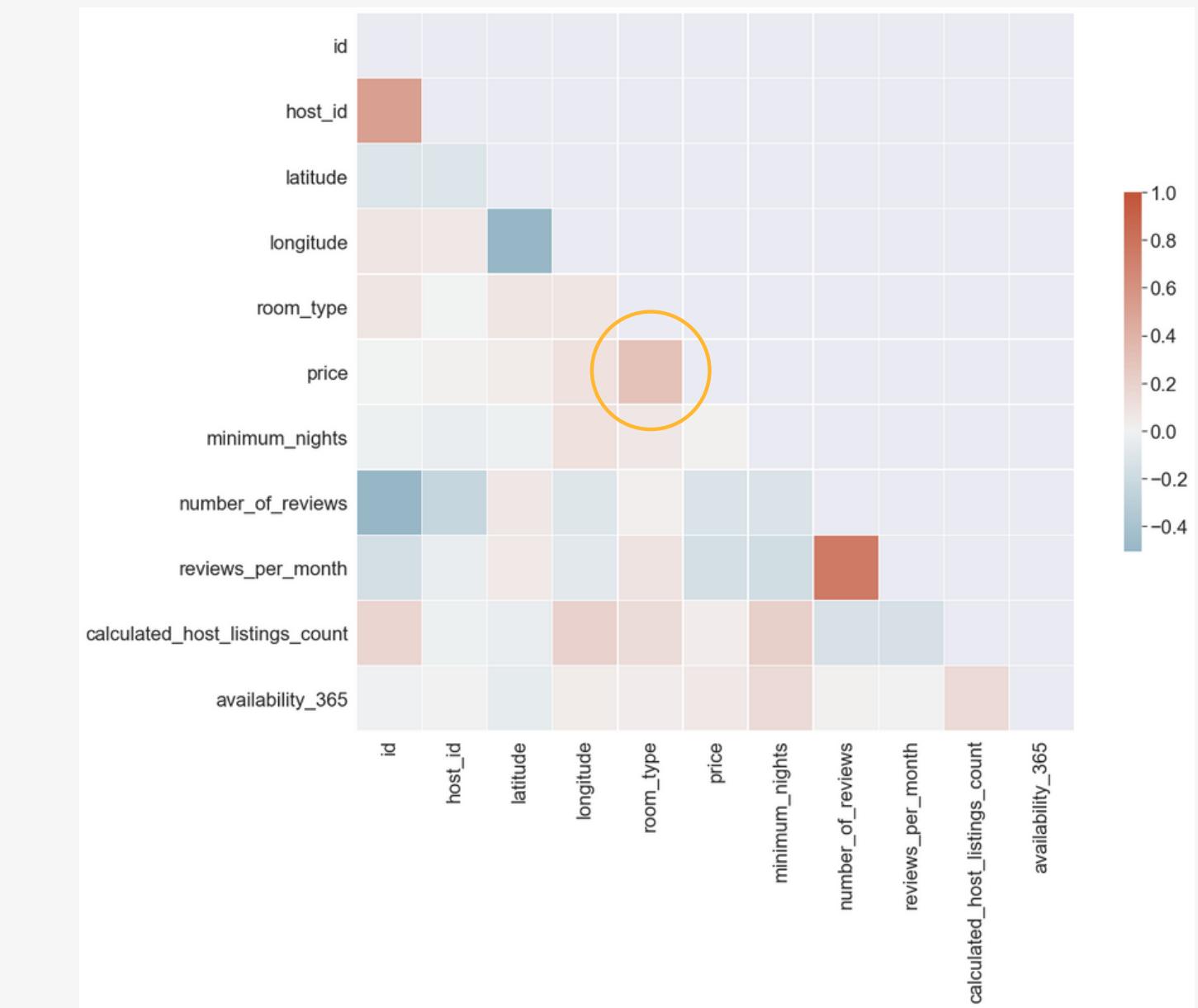
With the word cloud in comments of reviews, we can see that most reviews are positive, people care about the location, bed, and kitchen.



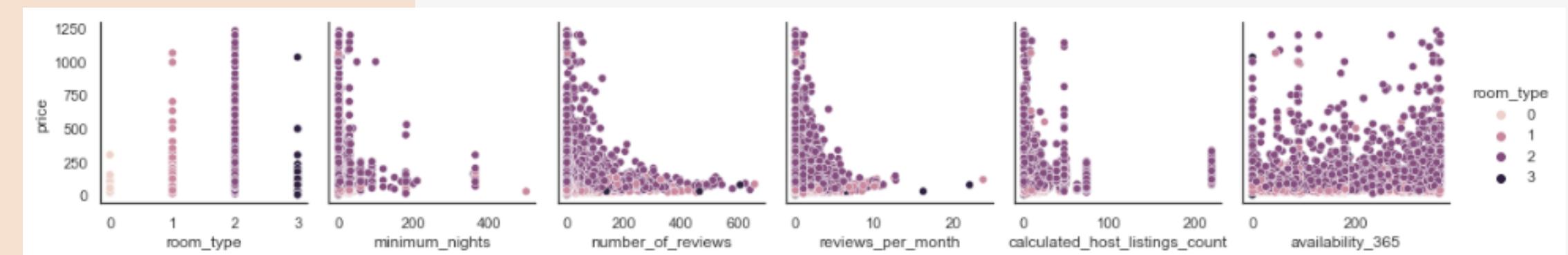
ANALYZING

- **Step 2: Detecting correlations**
 - Transfer key categorical variables into numeric format (i.e. room_type)
 - Draw correlation heat-map
 - Draw pair-plots
- **Key Findings:**
 - Price distribution is influenced by the room type. Generally, the price of shared rooms < private rooms < entire home/apt
 - Surprisingly, the number of reviews has a negative correlation with price

Listings Data Variables Correlation Heatmap



Listings Data Variables Pairplot



Next Steps

Put a visual on this slide and some bullet points about what we did after the Eda as a lead into the next section

TITLE

PROPERTY SALE

Presentations are tools
that can be used as
lectures, speeches,
reports, and more.

PROPERTY LEASE

Presentations are tools
that can be used as
lectures, speeches,
reports, and more.

PROPERTY AND PORTFOLIO MANAGEMENT

Presentations are tools
that can be used as
lectures and more.

05



06



Property features

Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more.



TITLE

07

LOCATION

Presentations are communication tools that can be used as lectures, speeches, reports, and more.

PRICE

Presentations are communication tools that can be used as lectures, speeches, reports, and more.

PROPERTY SIZE AND CONDITION

Presentations are communication tools that can be used as lectures, speeches, reports, and more.



EXPLORE ALPHA ANALYTICS!

Alpha Analytics

We provide data solutions and develop custom reporting and dashboards for your business needs

About Us

Alpha Analytics is dedicated to creating custom data science solutions to the client's specific needs. Interpreting data may be confusing – Alpha Analytics was created with this in mind. We strive to create understandable and actionable insights based on data you provide.

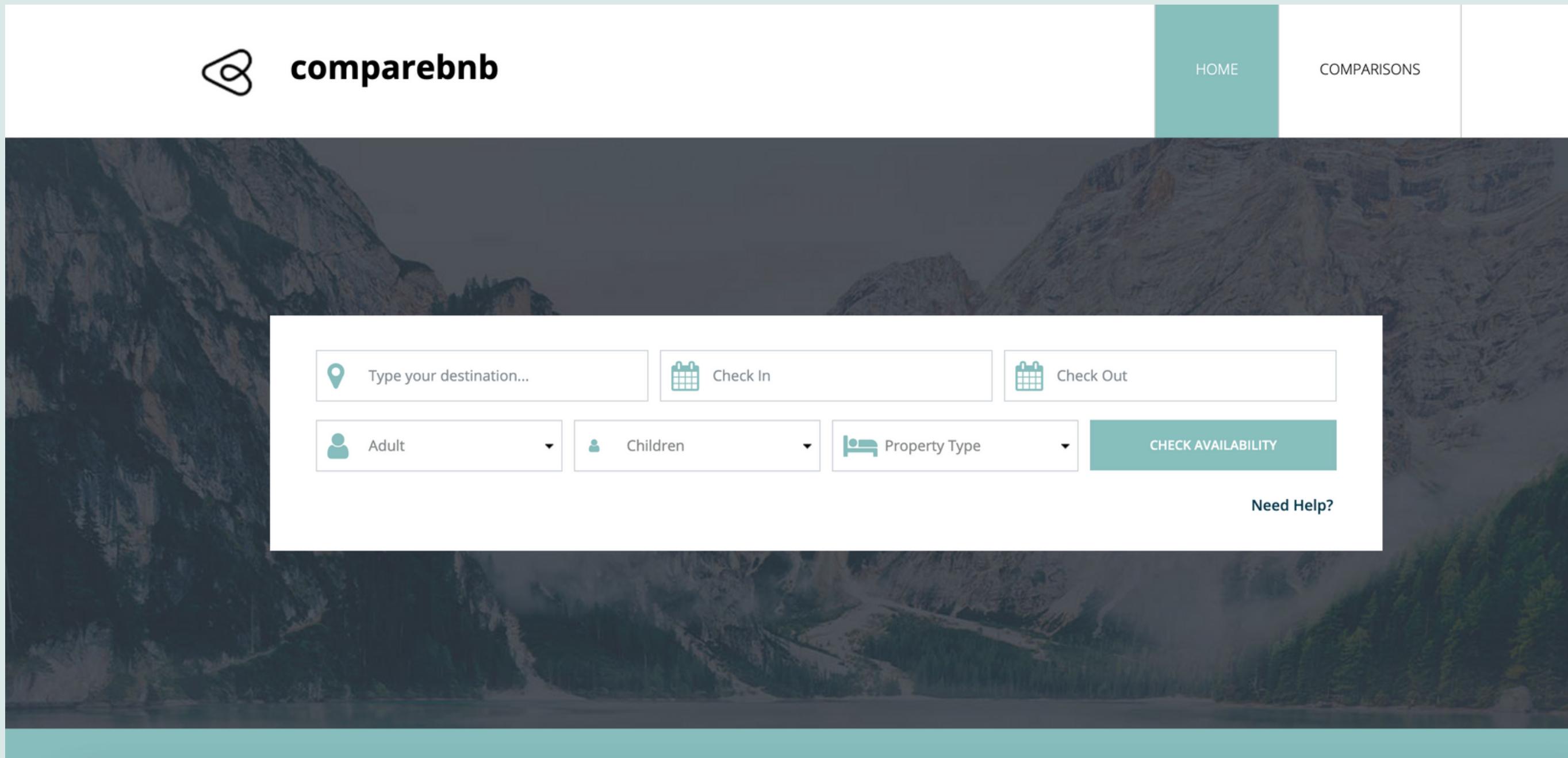
MEET OUR TEAM

With Alpha Analytics, you get

- ✓ Improved efficiency
- ✓ Increased productivity
- ✓ Ensured security
- ✓ Point-and-click interface

04

EXPLORE COMPAREBNB!



04

LOCATION

Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more.

PRICE

Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more.

PROPERTY SIZE AND CONDITION

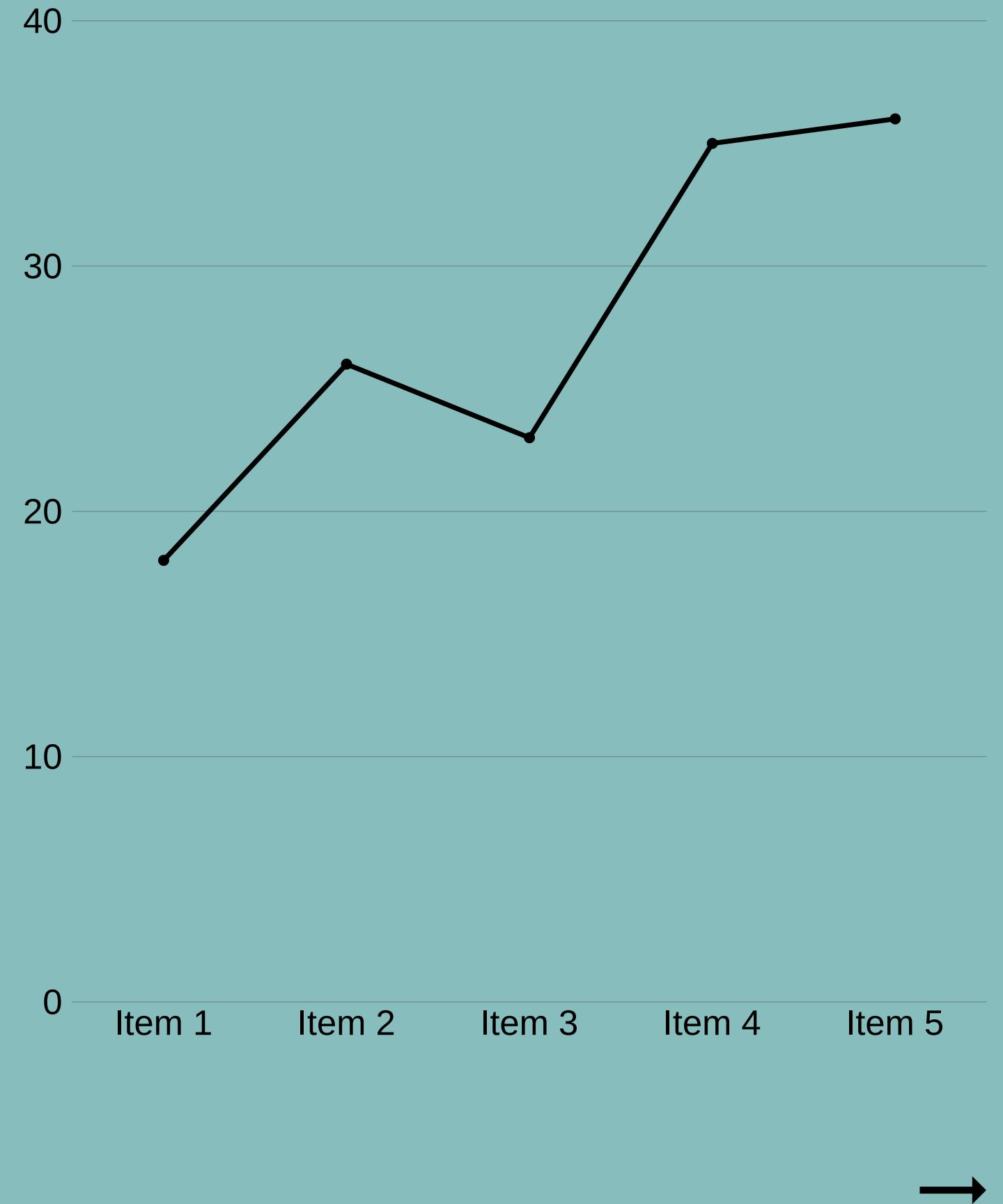
Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more.

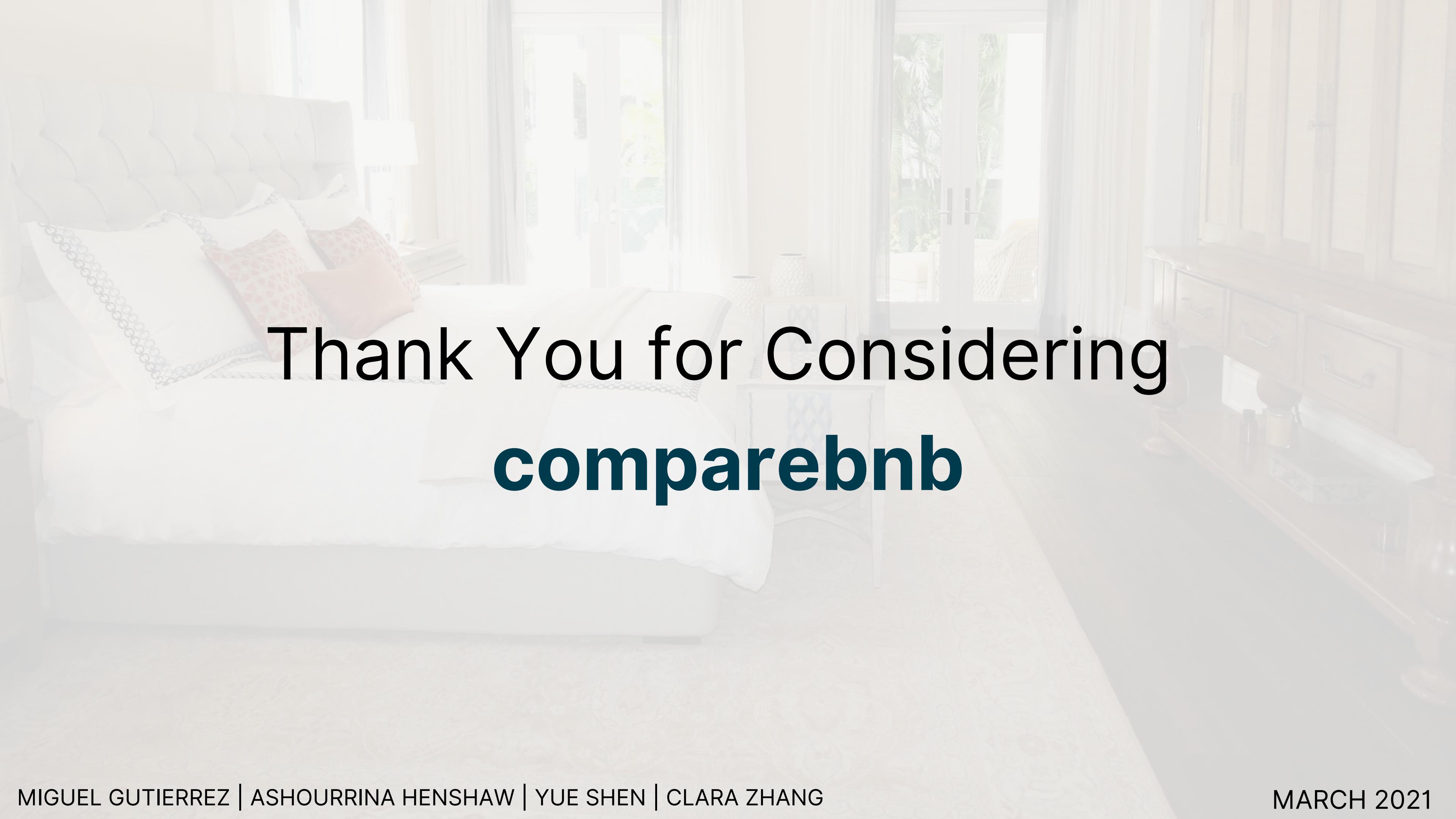
TITLE

08 →

Comparative analysis

Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more. It is mostly presented before an audience.





Thank You for Considering
comparebnb