

ANA502 Draft Proposal: Understanding What Drives Occupancy in the Short-Term Rental Economy

KOHEI NISHITANI

1. Introduction and Motivation

The number of international tourists visiting Japan has dramatically increased thus making the country a worldwide destination for travelers. The substantial economic expansion has created a serious deficit of available lodging particularly in locations with high tourist interests and urban centers. Airbnb and similar platforms have provided essential substitute accommodations by enabling users to book rooms in various sized facilities spread throughout Japan.

The increasing demand promotes new investment possibilities and innovative solutions in the short-term rental business. The proposed analysis focuses on Airbnb booking trends within Japan's specific area to identify marketplace requirements for customer demand and generate mission-critical business intelligence. The research aims to support strategic choices through data analytics and machine learning methods for shortened rental market investors and owners alongside hospitality entrepreneurs looking to excel in this business environment.

As part of my CPT, I'm exploring new business opportunities in the hospitality and property management space. This project aims to analyze Airbnb booking occupancy in a specific area in Japan, providing actionable insights that could support investment or operational decisions.

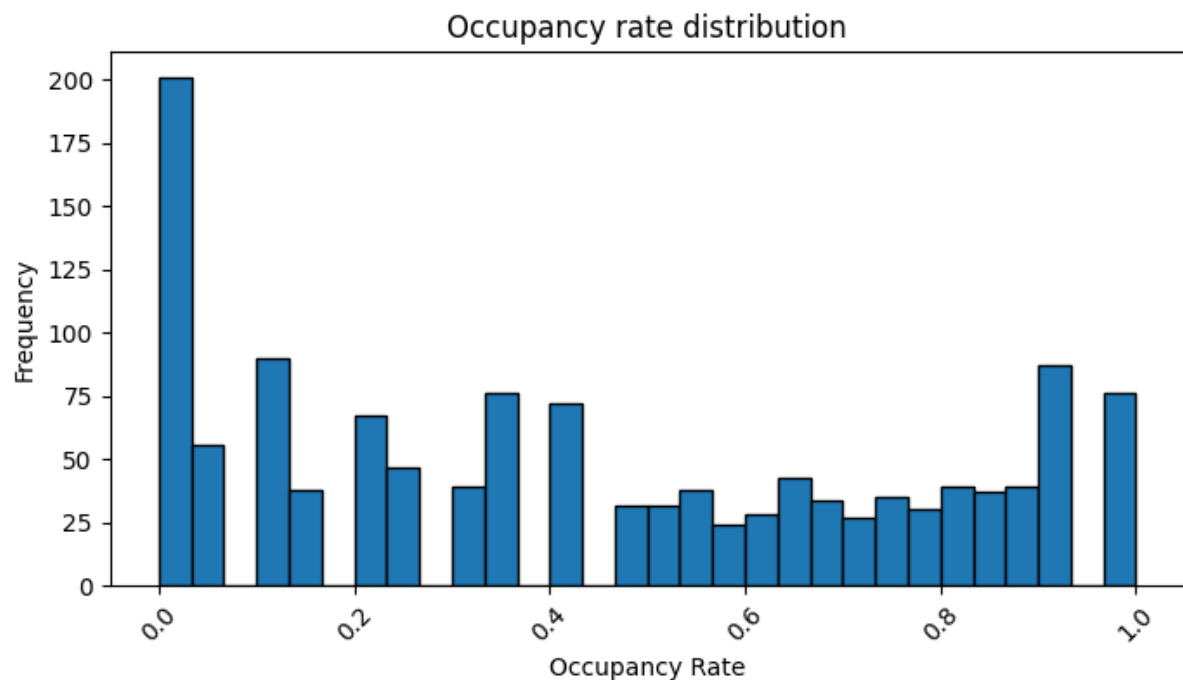
2. Research Question

What are the key factors that influence booking occupancy rates on Airbnb in Japan, and how can they inform strategic business opportunities?

3. Dataset Selection

The dataset used for this project includes Airbnb listings in a specific region in Japan, using [web scraping](#). The dataset has 1287 records with 41 columns, including occupancy rate, price per night and other property information. Some columns includes missing values, we need to fill these missing values 0 or delete it based on employed models.

Regarding target variable occupancy_rate's basic statistics, mean is 0.44 and median is 0.40.



4. Methodology

4.1 Data Cleaning & Preparation

- Handle missing values, especially in review or availability columns.
- Feature engineering: average occupancy, distance from city center, extract qualitative info from meta data etc.

4.2 Exploratory Data Analysis (EDA)

- Visualize relationships between occupancy and features like price, location, season.
- Identify high-performing listings and commonalities.

4.3 Predictive Modeling

- Use regression models (e.g., Random Forest, XGBoost, Linear Regression) to predict occupancy rates.
- Cluster listings based on performance or pricing using k-means or DBSCAN.

4.4 Business Insights

- Recommend pricing strategies based on season or listing type.
- Identify underserved areas or segments for new business opportunities.

5. Expected and Anticipated Outcomes

- Strategic insights for business owners to optimize operations and marketing
- Limitations may include data completeness or lack of booking confirmation records (vs. availability).
- Challenges to predict occupancy rates based on image impression, not listed written information.

6. Conclusion & Next Steps

- This project aims to combine data analytics and business intelligence to uncover actionable trends in Airbnb occupancy in Japan. The outcomes can serve both hosts and real estate stakeholders by guiding decisions around pricing, listing optimization, and expansion.
- From initial analysis, simple linear regression indicates poor fitting based on Mean Squared Error: 0.103, R² Score: 0.003. This exercise will require normalization and apply different models.

Annotated bibliography

Granicus. Short-term vacation rentals after COVID-19: What will change?

Annotation:

This blog post from Granicus discusses how the short-term vacation rental market may evolve in the wake of the COVID-19 pandemic. The piece addresses changing guest preferences such as heightened safety expectations, flexible cancellation policies, and the demand for longer stays as remote work continues. It also touches on the potential for increased regulatory scrutiny, suggesting local governments may enforce stricter guidelines around short-term rentals for public health and community impact reasons. For an Airbnb occupancy analysis project, this blog provides contemporary, practical insights into market shifts that might affect listing performance and booking patterns. Although it is not a peer-reviewed study, its timely industry viewpoint can help contextualize any observed changes in occupancy rates or host strategies post-pandemic.

Airbnb. (2021, May). Airbnb report on travel & living

Annotation:

This official PDF report from Airbnb examines emerging travel and living patterns in 2021, especially considering the pandemic's influence on remote work, "workations," and extended stays. It provides first-party data on shifting guest behaviours—such as average length of stay, trending destinations, and changes in off-peak booking patterns. By highlighting how guests increasingly blend work, travel, and leisure, the report offers direct insight into Airbnb's own occupancy metrics and future expectations for host demand. While the document serves as a marketing-oriented publication, it still contains valuable quantitative data (e.g., charts, booking trends) and qualitative commentary relevant to an Airbnb occupancy analysis. Researchers can compare these figures with independent datasets or third-party platforms (like Inside Airbnb) to examine whether Airbnb's reported shifts align with broader industry findings.

[AirDNA: Short-term rental trends](#)

Annotation:

This AirDNA blog post provides up-to-date analyses of global short-term rental (STR) trends, focusing on occupancy rates, average daily rates (ADR), and demand patterns across different regions. AirDNA is a well-regarded analytics platform in the STR industry, gathering large-scale listing data from Airbnb and similar sites. For an Airbnb occupancy project, the real-world metrics and market segmentation breakdowns in this blog offer practical benchmarks that can help you compare your specific dataset against broader industry performance. While it may not undergo the scrutiny of peer-reviewed research, AirDNA's monthly or quarterly snapshots of STR performance are often used by both researchers and property managers to understand current occupancy dynamics and forecast seasonal trends.

[Grand View Research Short-term vacation rental market size, share & trends analysis report](#)

Annotation:

Grand View Research offers comprehensive market research and industry analyses, and this report focuses specifically on short-term vacation rentals, covering market size, growth projections, and evolving consumer patterns. Though the full report may require a purchase, the publicly accessible summary outlines key drivers (e.g., technology platforms, changing traveler preferences) and regional variations in demand. For an Airbnb occupancy study, this resource helps situate your project within the global STR market, highlighting macro-level factors—such as economic conditions or traveler behaviors—that can influence Airbnb booking rates. Its industry-oriented

approach makes it valuable if your analysis aims to connect occupancy findings to broader commercial implications, including revenue forecasts and market competitiveness.

[Minut. \(n.d.\). Short-term rental statistics](#)

Annotation :

Minut is a property monitoring service that publishes periodic statistics and insights on the short-term rental sector. In this blog post, the company discusses recent data on demand growth, average booking windows, and guest demographics across various STR platforms. The blog highlights shifting traveler preferences, with remote and flexible work styles driving longer stays in certain regions. Although not an academic publication, it offers practical, real-world snapshots of key performance indicators that correlate with occupancy (e.g., how quickly listings fill up). Incorporating these statistics into your project can provide contemporary context for analyzing any anomalies or confirming industry-wide patterns in Airbnb occupancy rates.