

Datengetriebene Analyse zur Optimierung von Airbnb- Investitionen für "InvestZurich AG"

BINA – Gruppe Case Study 03– FS2025

Table of Content

Step 1

Defining Objectives and Information Needs

Step 2

Collecting Data

Step 3

Analyzing Data

Step 4

Presenting Results

Step 5

Deployment & Next Steps

Step 1

Defining Objectives and Information Needs

Use Case 1

Marktpotenzial & Standortanalyse -> Bäse / Marc

Ziel: Wo in Zürich bieten sich die besten Investmentchancen? Welche Quartiere haben hohe Preise/Nachfrage? Welche Wohnungstypen (Grösse, Zimmerzahl) sind pro Quartier gefragt/unterversorgt?

Methoden-Ansatz: Descriptive Statistics, Clustering, Datenvisualisierung.

Use Case 2

Preisstrategie & Ertragsprognose -> Tobi

Ziel: Was sind die Haupttreiber für Airbnb-Preise in Zürich? Welchen Preis kann InvestZurich AG für ein spezifisches Objekt realistisch erwarten? Gibt es saisonale Unterschiede bei Buchungen?

Methoden-Ansatz: Regression, Time Series Analyse, Descriptive Statistics.

Use Case 3

Performance Optimierung & Benchmarking -> JK

Ziel: Was unterscheidet Top-Performer (Superhosts) von anderen? Wie kann InvestZurich AG diesen Status für ihre Objekte erreichen (klare Handlungspfade)?

Methoden-Ansatz: Classification (z.B. Decision Trees), Descriptive Statistics.

Use Case 4

Listing-Optimierung durch Textanalyse (Optional) -> Robin

Ziel: Hat der Tonfall/Inhalt der Beschreibungstexte Einfluss auf Buchungen/Bewertungen?

Methoden-Ansatz: NLP (Sentiment Analyse, Topic Modeling).

Step 1

Defining Objectives and Information Needs

- Geospatial Verteilung der Listings
- Preisdaten (zusätzlicher Datensatz für Preisdaten über Zeit)
- Verfügbarkeits-/Buchungsindikatoren
- Objektmerkmale
- Standortmerkmale
- Host-Informationen
- Review-Daten
- Textdaten der Listings

Step 2

Collecting Data

Data source

*Die Analyse basiert auf öffentlich verfügbaren Daten von **Inside Airbnb** für Zürich (Stand: 30. Dezember 2024).*

Datasets

`listings.csv.gz`: Detailed Listings data (inkl. Textfelder für NLP)

`calendar.csv.gz`: Detailed Calendar Data (Preis/Verfügbarkeit über Zeit → Time Series)

`reviews.csv.gz`: Detailed Review Data (inkl. Text und Datum → Time Series, NLP)

`listings.csv`: Summary information and metrics for listings in Zurich (good for visualisations)

`reviews.csv`: Summary Review data and Listing ID (to facilitate time-based analytics)

`neighbourhoods.csv`: Neighbourhood list for geo filter.

`neighbourhoods.geojson`: GeoJSON file of neighbourhoods of the city (für Kartenvisualisierung)

Step 3

Analyzing Data

In diesem Schritt werden die vorbereiteten Daten analysiert, um die in Step 1 definierten Use Cases für InvestZurich AG zu beantworten.

Descriptive Statistics

Anwendung: Grundlage für alle Analysen. Berechnung von Kennzahlen (Mittelwerte, Median, Häufigkeiten) für Preise, Bewertungen, Auslastungsindikatoren etc. Vergleich von Quartieren und Host-Typen.
Ziel: Basisverständnis der Daten schaffen.

Regression

Anwendung: Identifikation der Haupttreiber für Preise und Bewertungen.
Ziel: Prognose von Preisen und Bewertungen.

Clustering

Anwendung: Gruppierung von Quartieren oder Listings basierend auf Ähnlichkeiten.
Ziel: Identifikation von Mustern und Segmenten.

Classification

Anwendung: Vorhersage von Superhost-Status oder anderen binären Zielvariablen.
Ziel: Handlungsempfehlungen für Performance-Optimierung.

Time Series Analysis

Anwendung: Analyse von saisonalen Mustern und Trends in Preisen und Verfügbarkeiten.
Ziel: Zeitbasierte Prognosen erstellen.

NLP

Anwendung: Analyse von Beschreibungstexten und Reviews.
Ziel: Insights aus Textdaten gewinnen (z.B. Sentiment, Themen).

Step 4

Presenting Results



- Diagramme auf einer Webseite darstellen
- Interaktion durch Hilfe von Javascript
- Google Maps Plattform bietet eine Heatmap-Funktion an

Heatmap

