

# README: hotels-vienna dataset

This is a README file for the `hotels-vienna` dataset that includes information on price and features of hotels in Vienna for one date.

Used in case studies Finding a good deal among hotels: data collection Finding a good deal among hotels: data preparation Finding a good deal among hotels: data exploration Finding a good deal among hotels with simple regression Measurement error in hotel ratings Finding a good deal among hotels with non-linear function Finding a good deal among hotels with multiple regression

## Data source

---

Scraped from a price comparison website. It was anonymized and slightly altered to ensure confidentiality. It contains all the necessary information about the location and rating that helps to distinguish them.

## Data access and copyright

---

The data was collected by the authors and may be used for education purposes only.

## About the data

---

### Raw data tables

`hotelbookingdata-vienna.csv`

The file contains data about hotel prices and features from a price comparison website.

- for Vienna, Austria,
- for a single weekday night in November 2017
- The dataset has N=430 observations.
- ID variable: `hotel_id`

### Tidy data table

`hotels-vienna` is just a slightly cleaned version of the raw data excluding duplicates.

- The dataset has N=428 observations.
- ID variable: `hotel_id`

variable name	info	type
hotel_id	Hotel ID	numeric
accommodation_type	Type of accomodation	factor
country	Country	string
city	City based on search	string
city_actual	City actual of hotel	string
neighbourhood	Neighbourhood	string
center1label	Centre 1 - name of location for distance	string
distance	Distance - from main city center	numeric
center2label	Centre 2 - name of location for distance_alter	string
distance_alter	Distance - alternative - from Centre 2	numeric
stars	Number of stars	numeric
rating	User rating average	numeric
rating_count	Number of user ratings	numeric
ratingta	User rating average (tripadvisor)	numeric
ratingta_count	Number of user ratings (tripadvisor)	numeric
hotel_id	Hotel ID	numeric
year	Year (YYYY)	numeric
month	Month (MM)	numeric
weekend	Flag, if day is a weekend	binary
holiday	Flag, if day is a public holiday	binary
nnights	Number of nights	factor
price	Pricee in EUR	numeric
scarce_room	Flag, if room was noted as scarce	binary
offer	Flag, if there was an offer available	binary
offer_cat	Type of offer	factor

📅 Updated: September 18, 2024