**1. DATA:**

**Data Source:**

This data is publicly available open-source data. It was downloaded from the Website: Zenodo.org <https://zenodo.org/record/4446043>

**Data Collection:**

An automated experiment was conducted based on web scraping. Using a web-automation framework (*Selenium WebDriver*), search queries were executed on the Airbnb platform that referred to accommodations in 10 major European cities for two people and two nights.

The offers were collected four to six weeks before the travel dates, and the collected prices refer to the full amount due for the accommodation, including the reservation fee and cleaning fee.

**Data Contents:**

The dataset includes accommodation in 10 major European cities for two people and two nights.

**Data Profile:**

* The original data has 21 columns/51708 rows. After the Data Quality check, the dataset contains 12 columns and 51611 rows.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Index** | **Columns** | **Description** | **Time Variant/ Invariant** | **Data type** |
| 1 | City | List of 10 European cities | Invariant | Qualitative |
| 2 | Days | Weekdays\Weekends | Invariant | Qualitative |
| 3 | Accommodation | the full price of accommodation for two people and two nights in EUR | Invariant | Quantitative |
| 4 | Room\_type | the type of the accommodation | Invariant | Qualitative |
| 5 | Person\_capacity | the maximum number of guests | Invariant | Quantitative |
| 6 | Cleanliness\_rating | cleanliness rating | Invariant | Qualitative |
| 7 | Guest\_rating | overall rating of the listing | Invariant | Quantitative |
| 8 | Bedrooms | number of bedrooms (0 for studios) | Invariant | Quantitative |
| 9 | City\_distance | distance from the city center in km | Invariant | Quantitative |
| 10 | Metro\_distance | distance from the nearest metro station in km | Invariant | Quantitative |
| 11 | Longitude | longitude of the listing location | Invariant | Quantitative |
| 12 | Latitude | latitude of the listing location | Invariant | Quantitative |

**2. Limitations and Ethics:**

* + The data has no missing information or values.
  + In terms of ethics, the data does not contain any personal information of property owners or guests. It contains only the geographical location of the property.

**3.Questions to explore:**

1. Does the City/Metro distance from the apt/room affect the guest rate?
2. What are the cheapest cities?
3. Do the highest accommodations cities have a better guest rating?
4. Does the distance from the City Centre/Metro affect the accommodation rate?