**Overview of the Airbnb Dataset for Cape Town**

The dataset you provided contains detailed information about Airbnb listings in Cape Town. Below is a description of the dataset, focusing on the fields most relevant to your analysis.

**Data Structure**

* **Total Rows (Listings):** 23,564
* **Total Columns:** 18

**Key Attributes**

1. **ID Fields:**
   * **id**: Unique identifier for each listing.
   * **host\_id**: Unique identifier for each host.
2. **Property Information:**
   * **name**: The name of the listing, which typically describes the property.
   * **room\_type**: Indicates the type of room being offered (e.g., Entire home/apt, Private room, Shared room).
   * **neighbourhood**: The specific area or neighbourhood where the listing is located.
   * **latitude & longitude**: Geographical coordinates of the property, useful for mapping the locations.
3. **Pricing and Availability:**
   * **price**: The price per night for the listing, in the local currency (presumably South African Rand).
   * **minimum\_nights**: The minimum number of nights required for booking.
   * **availability\_365**: Number of days in the year that the listing is available for booking.
4. **Review Data:**
   * **number\_of\_reviews**: Total number of reviews the listing has received.
   * **last\_review**: Date of the most recent review.
   * **reviews\_per\_month**: Average number of reviews the listing receives per month.
   * **number\_of\_reviews\_ltm**: Number of reviews in the last 12 months.
5. **Host Information:**
   * **host\_name**: The name of the host managing the listing.
   * **calculated\_host\_listings\_count**: Number of listings the host manages.
6. **Regulatory Information:**
   * **license**: License number associated with the listing, if available. This field is sparsely populated.

**Summary of Key Attributes**

* **Price**:
  + Range: 160 to 999,538 (currency unit is not specified, but typically ZAR).
  + Median: 1,250.5
* **Minimum Nights**:
  + Range: 1 to 999 nights
  + Median: 2 nights
* **Number of Reviews**:
  + Range: 0 to 756 reviews
  + Median: 6 reviews
* **Availability**:
  + Range: 0 to 365 days
  + Median: 220 days

**Data Format and Domain**

* **Format**: The dataset is in CSV format with a mix of numeric, categorical, and date fields.
* **Data Types**:
  + id, host\_id: Integer
  + name, host\_name, neighbourhood, room\_type, license: String
  + latitude, longitude, price, minimum\_nights, number\_of\_reviews, reviews\_per\_month, calculated\_host\_listings\_count, availability\_365, number\_of\_reviews\_ltm: Numeric
  + last\_review: Date

**Relevance to Your Application**

For your application that focuses on finding the best places according to the number of people, price, and reviews:

* **price**, **number\_of\_reviews**, **room\_type**, and **availability\_365** are directly relevant.
* **latitude** and **longitude** can be used to filter or map listings by location.
* **minimum\_nights** might influence the suitability of a listing for different types of stays.
* **reviews\_per\_month** and **last\_review** can help gauge the recency and frequency of reviews, indicating popularity and reliability.

**Overview of the Dataset**

The dataset contains various attributes related to Airbnb listings in Cape Town. Below is a detailed description of all the available data fields, including their format, range, and domain.

**Data Fields**

1. **id**
   * **Format**: Integer
   * **Range**: Varies per listing (Unique ID for each listing)
   * **Domain**: Identifies each Airbnb listing uniquely.
2. **name**
   * **Format**: String
   * **Domain**: Descriptive name of the Airbnb listing, which often includes key features or location.
3. **host\_id**
   * **Format**: Integer
   * **Range**: Varies per host (Unique ID for each host)
   * **Domain**: Identifies each host uniquely.
4. **host\_name**
   * **Format**: String
   * **Domain**: The name of the person or entity hosting the listing.
5. **neighbourhood\_group**
   * **Format**: String
   * **Domain**: This field is often used to group neighbourhoods, but it appears to be mostly NaN in this dataset.
6. **neighbourhood**
   * **Format**: String
   * **Domain**: Specific area or neighbourhood where the listing is located (e.g., "Ward 57").
7. **latitude**
   * **Format**: Float
   * **Range**: Approx. -34 to -33 (Cape Town coordinates)
   * **Domain**: Geographic coordinate for the latitude of the listing location.
8. **longitude**
   * **Format**: Float
   * **Range**: Approx. 18 to 19 (Cape Town coordinates)
   * **Domain**: Geographic coordinate for the longitude of the listing location.
9. **room\_type**
   * **Format**: String
   * **Domain**: Type of room offered (e.g., "Entire home/apt", "Private room", "Shared room").
10. **price**
    * **Format**: Float
    * **Range**: 160 to 999,538
    * **Domain**: Price per night for the listing, likely in South African Rand (ZAR).
11. **minimum\_nights**
    * **Format**: Integer
    * **Range**: 1 to 999
    * **Domain**: Minimum number of nights required to book the listing.
12. **number\_of\_reviews**
    * **Format**: Integer
    * **Range**: 0 to 756
    * **Domain**: Total number of reviews that the listing has received.
13. **last\_review**
    * **Format**: Date (YYYY-MM-DD)
    * **Range**: Various dates, with the most recent around June 2024
    * **Domain**: The date of the most recent review.
14. **reviews\_per\_month**
    * **Format**: Float
    * **Range**: 0.01 to 15.18
    * **Domain**: Average number of reviews the listing receives per month.
15. **calculated\_host\_listings\_count**
    * **Format**: Integer
    * **Range**: 1 to 134
    * **Domain**: Number of listings managed by the host.
16. **availability\_365**
    * **Format**: Integer
    * **Range**: 0 to 365
    * **Domain**: Number of days in the year that the listing is available for booking.
17. **number\_of\_reviews\_ltm**
    * **Format**: Integer
    * **Range**: 0 to 178
    * **Domain**: Number of reviews received in the last 12 months.
18. **license**
    * **Format**: String
    * **Domain**: License number associated with the listing, if available. This field is mostly sparse.

**Summary of Key Attributes**

* **Numerical Data**: Includes fields like price, minimum\_nights, number\_of\_reviews, reviews\_per\_month, calculated\_host\_listings\_count, availability\_365, number\_of\_reviews\_ltm. These fields are mostly integers or floats and are critical for quantitative analysis.
* **Categorical Data**: Includes fields like name, host\_name, neighbourhood, room\_type, license. These fields are strings and provide descriptive or categorical information about the listings.
* **Geospatial Data**: Latitude and longitude fields allow for spatial analysis and mapping of listings.
* **Date Data**: last\_review provides temporal information regarding the recency of reviews.