

AIRBNB Price prediction

Introduction

- Exploratory Data Analysis of Airbnb rentals in USA using pyspark.
- The dataset consists details of Airbnb id, log_price, property_type, room_type, amenities, accommodates, bathrooms, bed_type, cancellation_policy, cleaning_fee, city, description, first_review, host_has_profile_pic, host_identity_verified, host_response_rate, host_since, instant_bookable, last_review, latitude, longitude, name, neighbourhood, number_of_reviews, review_scores_rating, thumbnail_url, zipcode, bedrooms, beds

DATA SOURCE

- **Dataset location:** <https://www.kaggle.com/stevezhenghp/airbnb-price-prediction>
- Number of columns : 29
- Number of Rows : 74112
- Dataset Size: 99MB

Goal of the Project

- Analysis Areas:

- Does the price of the Airbnb depend on the rating score?
- Which Airbnb has the highest number of ratings in a city?
- Dependency of price on property type and amenities
- Depending on the number of hosts reviewed, what is the average rating by the hosts.



THANK YOU