

Mukhamejan Assan

## Assignment 2

### AirBnB listings in Athens

In this assignment, I developed a pricing model for a company managing small and medium-sized apartments that can accommodate 2-6 guests. The goal was to enable the company to price new apartments independently of market trends.

The dataset analyzed had 73 variables, but only a subset of them were relevant to the new listings. To prepare the data, I encoded the property type and removed any rooms that were not entire apartments. I also removed missing values for the beds variable and deemed other variables with missing values unnecessary for the analysis.

Next, I grouped similar apartments by mapping their location coordinates to the electoral districts of Athens and plotted the average log prices on a map. The visualization showed that prices tended to be higher in areas closer to landmarks.

I then classified the selected features into five categories, including basic property descriptors, location dummies, amenities dummies, description scores, and booking specifications. According to the correlation matrix, the top predictors were the number of bathrooms, accommodated people, number of bedrooms, beds, and various neighborhood and amenity features.

To determine the best performing model, I evaluated four models, including OLS, LASSO, CART, and Random Forest, using RMSE Train and RMSE Test metrics. The LASSO model performed the best, with an RMSE Train of 0.5754 and an RMSE Test of 0.5784.

To select the most important features, I used LASSO regularization, which shrinks the coefficients of the less important features to zero, effectively removing them from the model. The resulting model included the most important features, which were service amenities such as Wi-Fi, TV, iron, and air conditioning, as well as the number of bathrooms, accommodated people, number of bedrooms, and various neighborhood and amenity features.

Based on the analysis, these service amenities were the most critical features to potential clients, as they increase the comfort of the stay and justify higher prices. The listing description also plays a vital role in attracting clients, while lowering barriers to booking the stay increases the listing's attractiveness. Regarding location, since Athens has several historical landmarks, there is no single "hot area," and the listing could add a markup.