DATA MASTERCLASS

GROUP PROJECT
AIRBNB CASE STUDY

TEAM 04

Pfizer

Code.Hub

Code.Hub

THE TEAM

MELINA ZIKOU

GIANNIS LAZARIDIS

KYRIAKI CHRISTODOULIDOU

CHRISTOS S. MOSCHOFIDIS

Pfizer

A CASE STUDY FOR CREATING A SERVICE FOR AIRBNB HOSTS IN ATHENS

PROJECT OVERVIEW

We are employed as Data Scientists at Airbnb. Airbnb wants to create a service for hosts with top-rated undervalued listings that will suggest they increase their prices. Our team is tasked with building a POC for this service.

Pfizer

Code.Hub

EXPLORATORY DATA ANALYSIS

A first glance in our data:

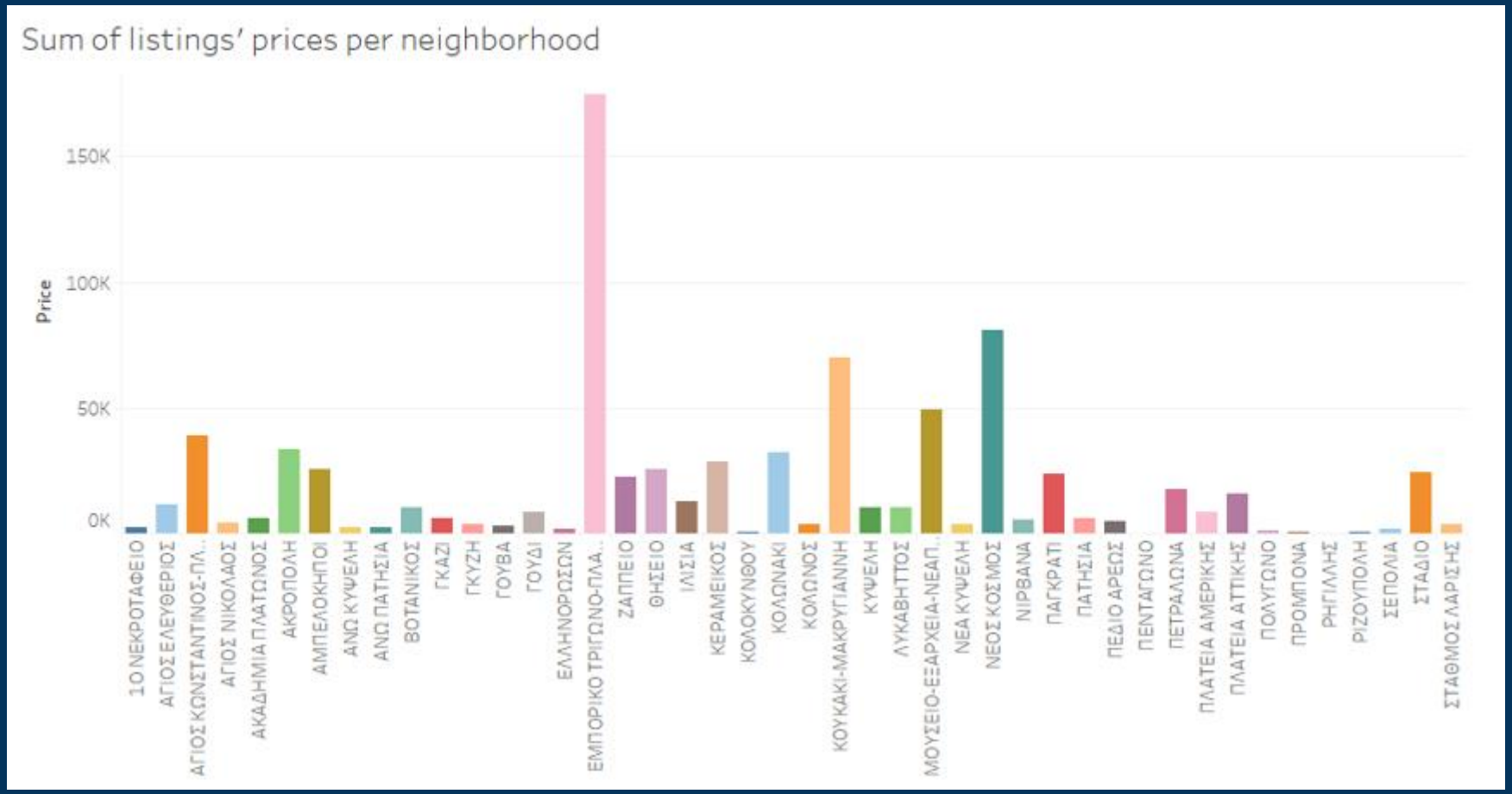
9582 samples and **67** features
56123 missing values
9582 rows affected
29 columns affected
6928 identified hosts

FURTHER EXPLORATION

"TOP 20" AMENITIES

Essentials
Wifi
Air conditioning
Long term stays allowed
Hangers
Hair dryer
Iron
Shampoo
Kitchen
Heating
Hot water
TV
Dishes and silverware
Cooking basics
Refrigerator
Coffee maker
Dedicated workspace
Bed linens
Washer
Elevator

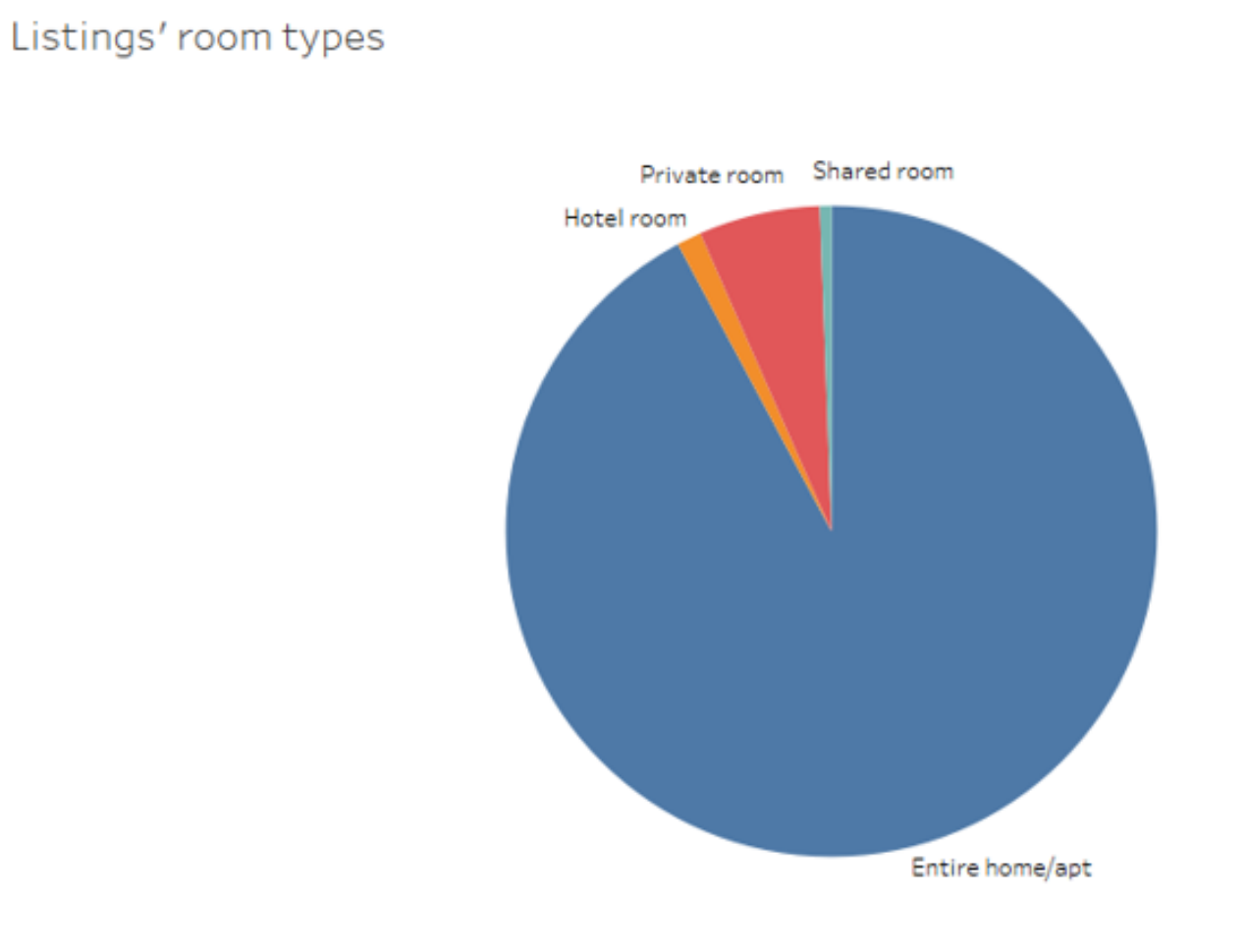
FURTHER EXPLORATION



Pfizer

Code.Hub

FURTHER EXPLORATION



PREPROCESSING

BRING THE DATASET IN A FORMAT UNDERSTANDABLE BY ML ALGORITHMS

WHICH DATA DO WE REALLY NEED?

- **COMMON LOGIC** FROM OUR OWN EXPERIENCE
- **MATHEMATICS** (PEARSON)

Pfizer

Code.Hub

PREPROCESSING

- Bathrooms
- Bedrooms
- Amenities
- Superhost
- Neighborhoods
- Number of reviews
- Accommodates
- Room Type
- Minimum Nights

PREPROCESSING

- **Missing values** were either dropped or replaced with the median.
- **Categorical features** were turned to numerical with "one-hot" encoding method.
- **New columns** were made by combining data such as neighborhoods' total value and "Top 20" amenities score.

PREPROCESSING

These methods were implemented during the modelling part in order to improve our model.

FEATURE SCALING

- STANDARDIZATION

HANDLING OUTLIERS

- IQR ALGORITHM WITH USE OF THE MEDIAN VALUE

Pfizer

Code.Hub

MODELLING

ML MODELS TESTED

Dummy Regressor
Linear Regression
KNN Regression
Random Forest
SVM
Gradient Boosting
Ada Boosting

Pfizer

Code.Hub

BEST MODELS

1. RANDOM FOREST		2. GRADIENT BOOSTING	
MAE	MAPE	MAE	MAPE
17.02	0.34	17.42	0.35

Pfizer

Code.Hub

THE WORK AMONG TEAM MEMBERS

EXPLORATION



PREPROCESSING


MODELLING


API


WEB APP

Pfizer

Code.Hub

LET US SHOW YOU OUR WEB APPLICATION

Pfizer

Code.Hub

Team 04

Melina Zikou

My name is Melina Zikou! I am a Computer Science student, specilizing in Data Sciences and Artificial Intelligence, eager to explore the world of data.

Kyriaki Christodoulidou

My name is Kyriaki Christodoulidou but feel free to call me Kiki! I am a Mathematician in the making and an aspiring Data Scientist getting ready to serve the magic of Data.

Giannis Lazaridis

My name is Ioannis Lazaridis and I am an enthusiastic, highly-motivated Electrical Engineering graduate passionate about data and computer engineering with MSc in Medical Informatics.

Christos - Spyridon Moschofidis

My name is Christos Moschofidis. I am an Economist who turned to IT via the MSc in Business Intelligence and Data Science and works as a BI Engineer. My interests lie in the world of Data Engineering and the entire Data Management ecosystem.

Complete the Form

No. of accommodates: 4

No. of Bedrooms: 2

No. of Beds: 2

No. of Reviews: 43

Minimum Nights: 2

Type of Airbnb: Home or Apartment

No. of Bathrooms: 2

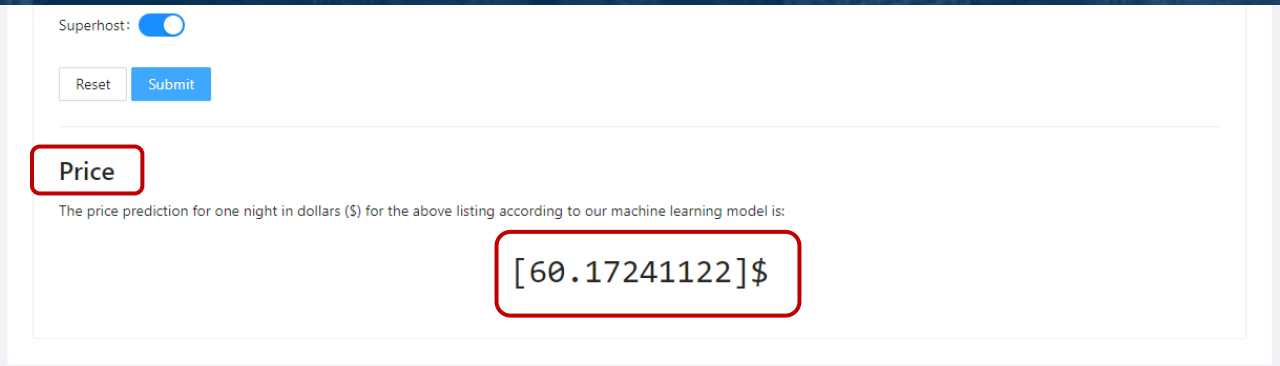
Bathroom Type: Private Bathroom

Amenities: Essentials × Wifi × Air conditioning × Kitchen × Heating × Hot water × TV ×

Neighbourhood: ΛΥΚΑΒΗΤΤΟΣ

Superhost:

Reset Submit



Code.Hub - Pfizer Data Masterclass
Airbnb Case Study
Team 04

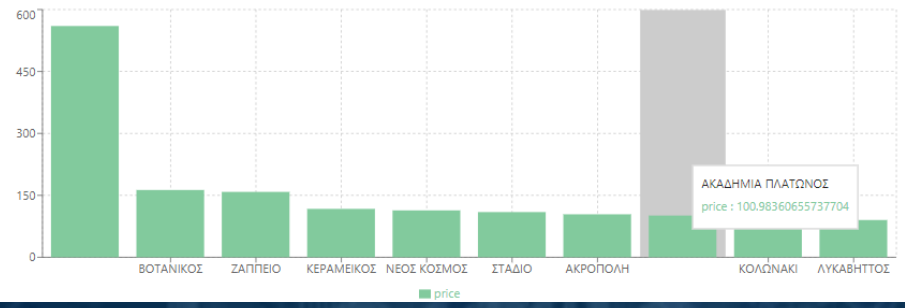
Pfizer

Code.Hub

Stats page

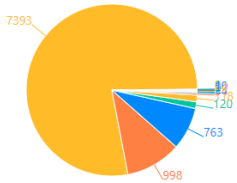
Average price per neighborhood

This diagram shows the average price of the listings in each of the top 20 neighborhoods.



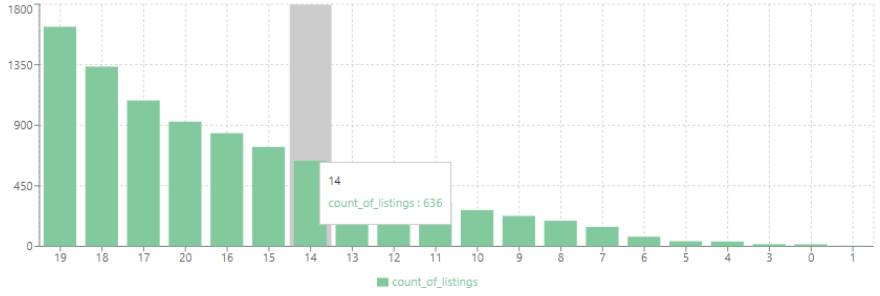
Bathroom Type

This pie diagram shows how many listings there are for each type of bathroom.



Number of amenities in each listing

This diagram shows how many listings there are, in correlation to the top 20 amenities extracted from the dataset.



PLEASE ENJOY OUR LIVE DEMO

Pfizer

Code.Hub

"My model for business is The Beatles. They were four guys who kept each other's kind of negative tendencies in check. They balanced each other and the total was greater than the sum of its parts. That's how I see business: great things in business are never done by one person. They're done by a team of people."

Steve Jobs

WHAT WE COULD DO FURTHER

IF WE HAD THE TIME...

- Descriptive statistics for better understanding.
- Testing different ways of preprocessing to see how the model responds to them.
- Deeper research in the models used.
- More appealing and more interactive Back and Front End Architecture.

LET US TAKE A MINUTE AND TAKE YOU A TRIP TO OUR FUTURE ...

THANKS FOR YOUR ATTENTION

Pfizer

Code.Hub


TEAM 04 IS READY
FOR A QNA ...
Pfizer
Code.Hub