

ANDRÉ PLANCHA, 105289

<ANDRE_PLANCHA@ISCTE-IUL.PT>

TOMÁS RIBEIRO, 105220

<TFROO1@ISCTE-IUL.PT>

AFONSO SILVA, 105208

<AGSOS@ISCTE-IUL.PT>

TODO AUTOR, NUMERO

<TODO EMAIL>

TODO TITULO

Hello Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris. Nulla malesuada porttitor

diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

O nosso trabalho tem como objetivo desenvolver um modelo que permita prever o preço a que os quartos são colocados á venda no Airbnb, com vista nesse objetivo foi nos disponibilizado uma base de dados que têm várias informações úteis para realizar esta tarefa. Usámos como suporte técnico a linguagem R através do uso de um computador pessoal quer para a limpeza de dados, o tratamento de dados, e as respetivas previsões efectuadas.

```
listings <- read.csv(here("data", "listings.csv"))
nrow(listings)

## [1] 6629
```

A base de dados que nos foi disponibilizada vem do projeto Inside Airbnb, fundado por Murray Cox com a missão de "[...] fornecer dados e defesa sobre o impacto do Airbnb em comunidades residenciais"¹.

A base de dados contém 6629 entradas, e cada uma delas representa um registo de um anúncio para o aluguer de um alojamento disponível no Airbnb, em São Francisco, Califórnia. Cada alojamento contém informação sobre o seu preço, localização, hospedeiro, o tipo de alojamento, as *reviews* do alojamento, e licença do alojamento.

```
data.frame(row.names = colnames(listings),
           type = sapply(listings, class)) %>%
  showT()
```

De forma a perceber melhor a base dados, o *Airbnb* disponibiliza de um "dicionário de dados"² que explica o significado de cada uma das variáveis:

- **id**: Número que representa um identificador único do anúncio;
- **name**: Título do anúncio;
- **host_id**: Identificador único da conta do hospedeiro;
- **host_name**: Nome da conta do hospedeiro;
- **host_since**: Normalmente este campo inclui apenas o primeiro nome ou nome da instituição hospedeira
- **neighbourhood_group**: Este campo encontra-se vazio e não inclui descrição no dicionário;
- **neighbourhood**: Embora este campo não inclua descrição no dicionário, nesta base de dados este campo descreve bairros de São Francisco como definido pelo Departamento de Planeamento da cidade ³.
- **latitude/longitude**: Coordenadas geográficas do alojamento;
- **room_type**: Tipo de alojamento, Entre "Quarto privado", "Quarto partilhado", "Quarto de hotel", e "Casa/Apartamento inteiro";

¹ About Inside Airbnb. insideairbnb.com/about.html. Acessado: 24/11/2022

	type
id	numeric
name	character
host_id	integer
host_name	character
neighbourhood_group	logical
neighbourhood	character
latitude	numeric
longitude	numeric
room_type	character
price	integer
minimum_nights	integer
number_of_reviews	integer
last_review	character
reviews_per_month	numeric
calculated_host_listings_count	integer
availability_365	integer
number_of_reviews_ltm	integer
license	character

² tinyurl.com/DataDictAirbnb. Accessed: 24/11/2022

³ Os bairros de São Francisco não contém fronteiras oficiais e dependem da fonte(tldrify.com/19p8), logo a definição das fronteiras definidas pelo Airbnb tiveram de ser determinadas. Mais à frente será demonstrado as fronteiras