

Andres Mauricio Martinez Celis, Código: 202322624

Taller 1 MINE 4101, Universidad de los Andes

### TALLER 1

Ciudad seleccionada: New Orleans, Luisiana, Estados Unidos

#### 1. Entendimiento inicial de datos:

Las dimensiones y propiedades principales del dataset se pueden observar en la siguiente tabla:

| Propiedad                         | Valor | Descripción   |
|-----------------------------------|-------|---|
| Numero de filas<br>(registros)    | 7216  | Son la cantidad de propiedades disponibles en el Dataset                        |
| Numero de columnas                | 75    | Son las características que tienen esas propiedades                             |
| Clave Primaria<br>(identificador) | Id    | Es el identificador de la propiedad dentro del Dataset, es único por propiedad. |

En el conjunto de datos, se han identificado tres tipos principales de datos. La mayoría de las columnas se clasifican como 'Object', lo que indica una variedad de tipos de datos, incluyendo cadenas de texto, listas y otros atributos que no pueden ser categorizados en tipos de datos específicos.

Las columnas relacionadas con revisiones (como 'review\_scores\_rating') y precios (como 'Price') se representan como valores flotantes. Es importante mencionar que los precios han sido formateados previamente debido a la presencia de caracteres especiales en los datos originales. Por otro lado, columnas que involucran atributos de conteo, como el número de dormitorios, baños y otras unidades, están representadas como enteros. En resumen, los tres tipos de datos predominantes en este conjunto de datos son 'Object', valores flotantes y enteros.

Atributos más importantes elegidos:

- `neighbourhood_cleansed`: Contiene el nombre de los sectores (barrios) donde se encuentra la propiedad dentro de la ciudad, junto con otros atributos puede determinar el lugar de compra ya que es la descripción mas acercada dentro de la ciudad hablando geográficamente; La ciudad contiene 68 barrios en los cuales hay al menos una propiedad de Airbnb
- `property_type`: Contiene el tipo de propiedad que es el aribnb, esto ayuda a determinar que tipo de propiedad es mejor comprar y que necesidad se puede ver en algunos sectores.

Existen 53 tipos de propiedad, siendo la más popular Entire home teniendo una cuota de más del 39% seguidas por Entire rental unit, Entire condo y Private room in home, entre las 4 suman mas del 78% del tipo de propiedad en la ciudad.

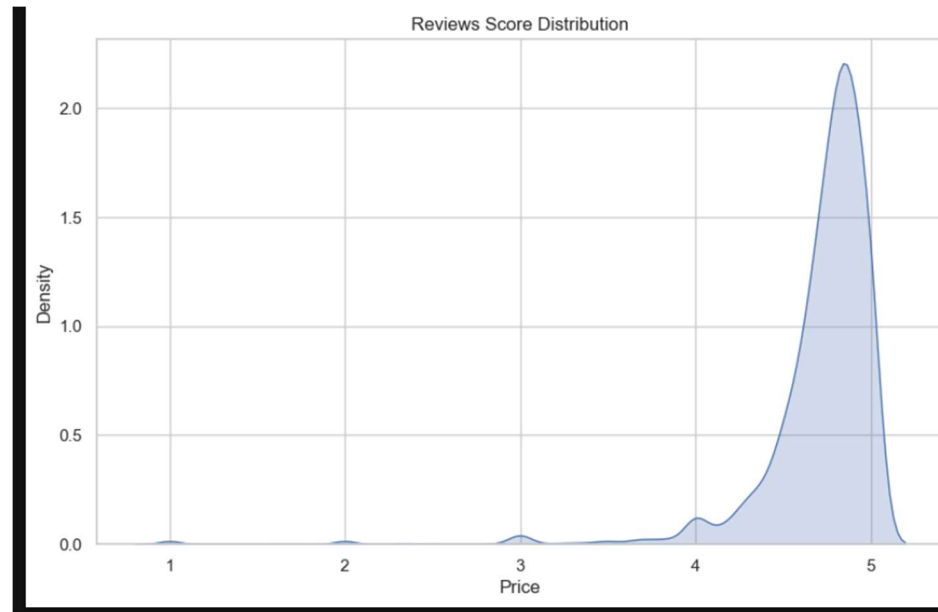
- `Price`: Es el precio por el cual se alquila la propiedad por día.

Densidad del precio (normalizada):



- `review_scores_value`: es el valor de las reviews que han hecho las personas para esa propiedad,

Densidad de los reviews:



- `availability_365`:  
Contiene la cantidad de días disponibles del Airbnb; se usa para medir que tan ocupado permanece el lugar, en el supuesto en el que los días NO disponibles sean ocupaciones de este.

## 2. Estrategia de análisis:

En el contexto de este análisis exploratorio de datos, nuestro objetivo primordial es realizar una selección de atributos relevantes dentro de nuestro conjunto de datos. Posteriormente, aplicaremos análisis bivariados para explorar las interacciones entre estos atributos identificados, centrándonos especialmente en el tipo de propiedades dentro de los diferentes barrios. Esta estrategia nos permitirá extraer información valiosa que será fundamental para definir la estrategia más adecuada. Específicamente, nos enfocaremos en establecer una relación óptima entre el precio de las propiedades y su ubicación geográfica en la plataforma Airbnb, considerando tanto el tipo de propiedad como el barrio en el que se encuentran. Mediante visualizaciones gráficas, identificaremos las ubicaciones más prometedoras para la adquisición y alquiler de propiedades según sus características específicas. Además, examinaremos los servicios más comunes ofrecidos por las propiedades seleccionadas, proporcionando así una visión integral y altamente enfocada de las oportunidades de inversión en Airbnb.