1. **Introducción**

"Inside Airbnb" es un sitio web que publica información sobre los alojamientos que se encuentran rentados bajo la plataforma de plataforma Airbnb.

Este sitio web busca mostrar el impacto que tiene esta plataforma en el mercado de la vivienda para cada ciudad.

Me decidí por seleccionar la ciudad de Estambul por ser una ciudad con mucha historia, cultura y además de su cultura gastronómica.

¿Sabías que es la única ciudad del mundo que se encuentra entre dos continentes, Asia y Europa?

**Objetivo:**

El objetivo de este análisis es ver cómo es la oferta de alojamientos en esta ciudad y ver si alguna variable influye en función de la respuesta de los huéspedes.

Para esto nos hicimos diferentes planteamientos:

1.- ¿La ubicación es importante para el huésped?

2.- ¿El tipo de habitación es importante para el huésped?

3.- ¿Un tiempo de respuesta se relaciona con más reservas?

4.- ¿Al usuario parece importarle que el host tenga la calificación de superhost?

1. **Limpieza de Datos**

Fuente: <http://insideairbnb.com/get-the-data.html>

En la web podemos encontrarnos con 7 documentos:

* listings.cvs
* listings\_details.cvs
* calendar.csv
* reviews.csv
* reviews\_details.csv
* neighbourhoods.csv
* neighbourhoods.geojson

Se utilizaron los datasets: listings.cvs y listings\_details.cvs para realizar los análisis y el modelo predictivo. Y el documento neighbourhoods.geojson para realizar los mapas.

Pasos realizados:

1. Selección de columnas:

Dado que el este dataframe listings\_details tiene 74 columnas se selccionaron las columnas que serán necesario para el análisis:

columns = ["property\_type", "accommodates", "first\_review", "review\_scores\_value", "review\_scores\_cleanliness", "review\_scores\_location", "review\_scores\_accuracy", "review\_scores\_communication", "review\_scores\_checkin", "review\_scores\_rating", "maximum\_nights", "listing\_url", "host\_is\_superhost", "host\_about", "host\_response\_time", "host\_response\_rate", "host\_listings\_count","number\_of\_reviews\_ltm","reviews\_per\_month"]

1. Merge: listings & listings\_details

Luego se procede a realizar un merge entre los listados listing y listings\_details.

df = pd.merge(listings, listings\_details[columns], on='id', how='left')

1. **Identificación de duplicados**

No se identificaron valores nulos

1. **Identificación de valores nulos**

(poner aqui el reporte de los valores nulos)

* Las columnas: **neighbourhood\_group** y **license** no contienen datos por lo que se procede a eliminarlas.
* La columna: **host\_about** se procede a eliminarla ya que contiene la descripción que hace el host para presentarse en la plataforma por lo que no se considera necesaria para el análisis.
* La columna: **name** aplicamos la función **fillna** para reemplazar los valores nulos por la palabra: “no name”.
* El resto de las columnas que presentan valores nulos no los vamos a reemplazar ya que corresponden a variables relacionadas con la reviews y se utilizarán ya que es una métrica opcional, no absoluta.

1. **Transformación de columnas**

* **host\_response\_rate**: se convierte a numérica con la función to\_numeric.
* **price\_euro**: se crea una nueva variable con la columna price convertida a euro ya que esta columna viene en moneda local (Lira turca) y así poder tener una mejor comprensión y percepción.

El tipo de cambio utilizado es de la web de la unión europea:

https://commission.europa.eu/funding-tenders/procedures-guidelines-tenders/information-contractors-and-beneficiaries/exchange-rate-inforeuro\_es

El tipo de cambio a la fecha fecha: 26/06/2023 es 1 TRY = 0.04564 EUR

Colores:

#AF1D56

#CB6CE6

#FF914D

#FFDE59

['id','host\_id','room\_type','price','minimum\_nights','number\_of\_reviews','reviews\_per\_month\_x','availability\_365','number\_of\_reviews\_ltm\_x','property\_type','review\_scores\_value','host\_is\_superhost','host\_response\_time','price\_euro','number\_of\_reviews\_ltm\_y']

'name', 'host\_id', 'host\_name', 'neighbourhood', 'latitude',

'longitude', 'room\_type', 'price', 'minimum\_nights',

'number\_of\_reviews', 'last\_review', 'reviews\_per\_month\_x',

'calculated\_host\_listings\_count', 'availability\_365',

'number\_of\_reviews\_ltm\_x', 'property\_type', 'accommodates',

'first\_review', 'review\_scores\_value', 'review\_scores\_cleanliness',

'review\_scores\_location', 'review\_scores\_accuracy',

'review\_scores\_communication', 'review\_scores\_checkin',

'review\_scores\_rating', 'maximum\_nights', 'listing\_url',

'host\_is\_superhost', 'host\_response\_time', 'host\_response\_rate',

'host\_listings\_count', 'number\_of\_reviews\_ltm\_y', 'reviews\_per\_month\_y',

'price\_euro'],

dtype='object')

-> Exordio (inicio)

╠ Saludo

╠ Me presento

╠ Doy mis credenciales

╠ Tema

╚ División de ideas (porque lo voy a hacer)

Si además conseguimos meter un dato interesante, frase célebre o una pequeña historia con intención de afectar al sentimiento del oyente, se llama Exordio complejo.

-> Narración (se suele dividir en 4 partes: narratio, divisio, confirmatio y confutatio)

╠ Organización de las ideas

╠ Transición entre las mismas

╠ Pequeños descansos conceptuales

╚ Hilo argumentativo

 -> Peroración

╠ Recapitulo las ideas

╠ Recuerdo el tema y su motivación

╠ Comento alguna anécdota, dato interesante o del estilo (intento llegar a las emociones de quien me escucha).

╠ Agradecimientos (al público, no a vuestros amigos)

╚ Despedida (no solo un "adiós", también un "estoy a vuestra disposición" o del estilo)

 Recursos:

· Entonación:  qué transmitís con vuestros sonidos (indendientemente del significado)

· Pronunciación:  cuánto cuesta entenderos

· Estructura: Qué organización tiene lo que decís

· Cohesión: Qué relación hay entre las partes de lo que decís (tanto sobre el contenido como sobre la transición)

· Pausas:  Tiempo que nos dais para procesar lo que habéis dicho y para ganar tiempo como orador/a para pensar sobre lo siguiente que vais a decir. Ayuda a evitar muletillas