

## "Extracción y Análisis de Información de Hoteles en la Costa Dorada utilizando Web Scraping en Booking.com"

### Introducción:

El turismo es una de las industrias más importantes del mundo y España es uno de los destinos turísticos más populares en Europa. En particular, la Costa Dorada, ubicada en la provincia de Tarragona, es conocida por sus hermosas playas y su clima mediterráneo.

En este trabajo final, se presentará una base de datos creada a través de la aplicación de web scraping con Selenium en la página web de Booking, con el objetivo de buscar hoteles disponibles en la Costa Dorada para una familia de dos adultos y dos niños durante una semana en el mes de julio de 2023.

El web scraping es una técnica utilizada para recopilar datos de sitios web de forma automatizada. En el caso de la página web Booking, esta es una plataforma que ofrece información sobre alojamientos y servicios turísticos en todo el mundo, lo que la convierte en una fuente de datos valiosa para aquellos que desean obtener información sobre la industria del turismo.

Al realizar un dataset a partir del web scraping de Booking, es posible obtener información detallada sobre los alojamientos, servicios, precios y disponibilidad en diferentes destinos turísticos, lo que puede resultar muy útil para diversas aplicaciones, como la planificación de viajes, la investigación de mercado, el análisis de la industria turística y mucho más.

Además, con la ayuda de herramientas de análisis de datos, es posible extraer información valiosa de los datos recopilados, como tendencias de precios, popularidad de destinos turísticos, preferencias de los usuarios y más. Todo esto puede ser de gran utilidad para tomar

decisiones informadas en el sector turístico o en cualquier otra industria relacionada.

En resumen, realizar un dataset a partir del web scraping de Booking puede ser una tarea interesante y valiosa para obtener información relevante y útil sobre la industria turística y otros campos relacionados, lo que puede ayudar a tomar decisiones informadas y estratégicas en el futuro.

### **Presentación del Dataset y objetivos:**

El **objetivo principal** de crear un dataset a través de web scraping de la web de Booking.com es recopilar información relevante y útil que pueda ser utilizada para realizar análisis y tomar decisiones informadas en el ámbito turístico.

*Algunos posibles objetivos específicos incluyen:*

- Obtener información sobre los precios, ubicaciones, servicios y características de los alojamientos disponibles en diferentes destinos turísticos de la Costa Dorada.
- Identificar tendencias en los precios de alojamiento .
- Analizar la satisfacción de los clientes y las opiniones de los usuarios a través de sus valoraciones y comentarios sobre los alojamientos.
- Representar gráficamente las variables analizadas.
- Realizar análisis estadísticos más complejos, como regresiones y modelos de predicción. Estos análisis podrían ayudar a identificar patrones y tendencias en los datos, lo que podría ser útil para predecir el comportamiento futuro del dataset.

Una vez que se ha obtenido el dataset, el siguiente paso es realizar un análisis exploratorio de datos para comprender mejor la estructura y las características de los datos recopilados. Esto puede incluir la identificación de valores atípicos, la eliminación de datos faltantes y la exploración de relaciones entre diferentes variables.

Después del análisis exploratorio de datos, se puede realizar un preprocesado de los datos para prepararlos para su posterior análisis. Esto puede incluir la normalización de los datos, la selección de características relevantes, la eliminación de ruido y la limpieza de los datos. El objetivo final es tener un dataset limpio y estructurado que sea fácilmente utilizable para llevar a cabo análisis y tomar decisiones informadas en el ámbito turístico.

### **Definición de las variables:**

- Nombre del hotel: Se refiere al nombre del establecimiento donde se desea alojarse.
- Sitio: Hace referencia al lugar donde se desea alojarse. Esta variable es importante porque permite a los clientes buscar hoteles en una ubicación específica. Los clientes pueden buscar hoteles en una ciudad o en una región determinada, o pueden especificar un punto de interés cercano al hotel.
- Tipo de alojamiento: Se refiere al tipo de habitación o alojamiento que se desea reservar. Los clientes pueden seleccionar entre diferentes tipos de alojamiento, como habitaciones individuales, dobles, suites, apartamentos, etc. Esta variable es importante porque los diferentes tipos de alojamiento tienen diferentes características y precios, lo que puede influir en la elección del cliente.
- Tipo de cancelación: Se refiere a las condiciones de cancelación de la reserva. Algunos clientes pueden querer reservar un hotel con una política de cancelación flexible, mientras que otros pueden estar dispuestos a arriesgarse con una política de cancelación más restrictiva si esto significa conseguir un precio más bajo. Esta variable es importante porque permite a los clientes seleccionar una política de cancelación que se adapte a sus necesidades y preferencias.
- Precio: Se refiere al precio de la reserva. Esta variable es importante porque el precio es un factor determinante en la elección del hotel por parte del cliente. Los clientes pueden seleccionar diferentes rangos de precios para ajustarse a su

presupuesto y preferencias. Además, el precio puede estar influenciado por otras variables, como el tipo de alojamiento, la política de cancelación, la ubicación del hotel, etc.

- Valoración: valoración del 1 al 10 que le han otorgado los huéspedes el alojamiento.
- Habitaciones: tipo de habitaciones del alojamiento.