

TALLER 1

El objetivo de este taller es evaluar las capacidades del estudiante para realizar análisis de datos utilizando técnicas estadísticas y de visualización, que le permitan entender un conjunto de datos, descubrir *insights* y sugerir accionables al negocio.

Contexto de negocio: Demanda de ocupación hotelera.

Como consultor científico de datos especializado en el sector hotelero, su objetivo es ayudar a cadenas de hoteles a optimizar sus estrategias de ocupación. Con el aumento de la competencia y la variabilidad en la demanda, cada vez resulta más complejo tomar decisiones acertadas respecto a la política de precios, la gestión de cancelaciones y la anticipación de las reservas para garantizar ingresos sostenibles en el tiempo. Una vez que un cliente realiza una reserva, entender los factores que determinan la duración de la estadía, la probabilidad de cancelación y la influencia de características como el tipo de hotel, la composición del grupo de huéspedes o el canal de reservación es vital para el éxito del negocio.

Para este fin se le ha proporcionado un dataset, disponible [aquí](#), que incluye información tanto de un hotel de ciudad como de uno tipo resort. En él se encuentran atributos como la fecha de la reserva, duración de la estadía, número de adultos, niños y bebés, disponibilidad de parqueaderos, estado de la reserva, entre otros. El diccionario detallado de los datos lo puede consultar [aquí](#).

Mediante el uso de técnicas estadísticas y de visualización de datos se espera que usted esté en capacidad de descubrir patrones, tendencias y oportunidades que puedan ser de interés para la cadena hotelera en la formulación de sus estrategias.

ACTIVIDADES Y ENTREGABLES

1. **[20%] Entendimiento inicial de datos:** Generar un breve reporte de entendimiento inicial de datos en donde destaque las dimensiones del dataset, los tipos de datos que contiene y el top 5 de los atributos que considera más importantes para el

análisis. Por cada uno de los 5 atributos incluya algunos elementos básicos de su comportamiento o distribución (análisis univariado).

2. **[15%] Estrategia de análisis:** Describir de manera concreta en uno o dos párrafos la estrategia de análisis y las técnicas que utilizará, con su respectiva justificación, para dar respuesta al requerimiento de establecer las mejores estrategias para reducir cancelaciones y mejorar la ocupación hotelera. La estrategia debe contemplar desde el cálculo de estadísticos básicos hasta el uso de técnicas de visualización que involucren varios atributos.
3. **[40%] Desarrollo de la estrategia:** Implementar la estrategia previamente definida en donde se evidencie claramente cualquier tipo de procesamiento de datos que haya tenido que realizar, las técnicas estadísticas o de visualización de datos utilizadas y los *insights* extraídos tras la interpretación de los resultados.
4. **[25%] Generación de resultados:** Redactar un informe ejecutivo o una presentación corta en donde recomiende a la cadena hotelera un conjunto de estrategias basadas en factores como duración de la estadía, anticipación de la reserva (*lead time*), tipo de hotel, tipo de habitación, entre otros de su elección, que permitan reducir cancelaciones y mejorar la ocupación. Recuerde que todas estas decisiones deben estar basadas en datos.

INSTRUCCIONES DE ENTREGA

- El taller debe desarrollarse en parejas.
- Todos los recursos generados deben entregarse mediante un repositorio de GitHub. **El repositorio debe ser público.**
- Documente en el archivo Readme información relevante como integrantes del equipo, objetivo, alcance, conclusiones (*insights*), instrucciones de ejecución y dependencias.
- El repositorio debe ser auto-contenido, es decir, debe incluir de una forma clara y concreta la solución a cada uno de los puntos del taller (implementación + interpretación).
- El o los notebooks generados deben poderse ejecutar secuencialmente sin ningún error. En el caso de haber creado varios notebooks, se debe clarificar el orden en que deben ser ejecutados.
- Se penalizará a los estudiantes que no cumplan con estos criterios de entrega.