



Tecnológico de Monterrey

Campus Santa Fe

Desarrollo de proyectos de análisis de datos (Gpo 307)

Etapa 1. Elaboración del Project Charter

Nombre:

German Gedovius Dalloz - A01782383

Samuel Aranzola Monroy - A01786577

1.-Descripción de la situación problemática

El incremento en la oferta de alojamientos a través de plataformas como Airbnb ha generado un mercado altamente competitivo, donde los anfitriones buscan optimizar la ocupación y maximizar sus ingresos. Sin embargo, factores como la variabilidad en los precios, las diferencias entre barrios, las características del alojamiento y la demanda estacional dificultan establecer estrategias claras para determinar precios óptimos o identificar qué atributos influyen más en la rentabilidad.

Actualmente, se cuenta con un conjunto de datos que incluye precios, número de habitaciones, ubicación, servicios, tipo de propiedad y calificaciones. Sin un análisis estructurado, es complicado transformar estos datos en recomendaciones accionables. El problema central es ¿cómo aprovechar los datos disponibles para mejorar la toma de decisiones relacionadas con precios, calidad del alojamiento y competitividad en la plataforma?

2. Preguntas de negocio a resolver

- ¿Qué variables influyen más en el precio de un alojamiento?
- ¿Cómo varía el precio entre diferentes zonas o barrios de la ciudad?
- ¿Qué características incrementan la probabilidad de obtener una alta ocupación?
- ¿Cuál es el rango de precios competitivo en relación con propiedades similares?
- ¿Qué mejoras puede implementar un anfitrión para aumentar sus ingresos o su rating?

3. Objetivos del proyecto (método SMART)

Objetivo general

Aplicar análisis estadísticos y modelos predictivos para comprender los factores que afectan el precio y desempeño de las propiedades, con el fin de apoyar la toma de decisiones estratégicas.

Objetivos SMART

S – Específico: Identificar las variables más relevantes que determinan el precio y analizar patrones por zona, características y disponibilidad.

M – Medible: Evaluar el rendimiento del modelo mediante métricas como R^2 , MAE y visualización de residuales.

A – Alcanzable: Utilizar herramientas de Python (pandas, seaborn, sklearn) sobre el dataset disponible para generar modelos y visualizaciones.

R – Relevante: Proveer recomendaciones que mejoren la competitividad de los anfitriones en su estrategia de precios y calidad.

T – Temporal: Completar el análisis y entregar resultados en un plazo máximo de 2 semanas, incluyendo modelo, interpretación y presentación final.

4. Identificación de limitaciones y restricciones

Los datos pueden contener valores faltantes, duplicados o inconsistencias.

Las características registradas no siempre capturan factores externos como eventos locales, temporada o regulaciones.

La interpretación del modelo puede verse limitada por multicolinealidad entre variables.

La calidad de los resultados depende de la disponibilidad y representatividad de los datos.

No se cuenta con información histórica completa, por lo que no es posible realizar un análisis de series de tiempo profundo.

El modelo no puede garantizar causalidad, solo relaciones estadísticas.

5. Plan de trabajo con cronograma y responsabilidades

	Actividad	Responsable
semana 3	Limpieza del dataset, tratamiento de nulos, análisis exploratorio (EDA).	
semana 3	Construcción de visualizaciones principales (MAPS, distribuciones, correlaciones).	
semana 4	Selección de variables, creación del modelo predictivo y análisis de residuales.	
semana 4	Interpretación de resultados, conclusiones y recomendaciones.	
semana 4	Elaboración del reporte final y presentación.	

6. Referencias bibliográficas

Airbnb. (2024). Inside Airbnb Dataset. <https://insideairbnb.com/>

Géron, A. (2020). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media.

McKinney, W. (2022). Python for Data Analysis. O'Reilly Media.

Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.

Seaborn Documentation. (2024). <https://seaborn.pydata.org/>

Scikit-learn Documentation. (2024). <https://scikit-learn.org/>