

UNIVERSIDAD DE LOS ANDES

Ciencia de datos aplicada.

Informe Taller 1

1. **Entendimiento inicial de los datos** Se determinó si los datos cumplirán con los requisitos mínimos para ser trabajados, para esto se hizo un conteo de los campos que estaban nulos, se transformó el costo de los alquileres a número ya que estaban en texto. Ya que, el requerimiento de negocio estaba orientado a las inversiones, se escogió como primera variable, el precio (por noche), la ubicación, el tipo de inmueble, las clasificaciones de los usuarios y por último los servicios adicionales.
2. **Estrategia de análisis:** Para este análisis se empieza por determinar cuál es la zona con más inmuebles en Airbnb, luego se compara el costo de las zonas, y se escogen las zonas con más inmuebles y que tengan el mayor costo de servicio. Además, se determina cuáles son los servicios que los clientes más utilizan, Además de obtener una comparación por las calificaciones dejadas por los clientes
3. Desarrollo de la estrategia: Mediante la hipótesis nula, se intenta determina si las propiedades y los tipos de propiedades tienen algún tipo de correlación, el resultado de la hipótesis es que esta correlación no se cumple.

Conclusión:

Aunque en un inicio al realizar un análisis por la desviación estándar, se ve que este es un negocio con un nivel de incertidumbre grande, ya que los precios son muy variados, sin embargo, al analizar zona a zona esta desviación disminuye lo que baja la volatilidad de este negocio. Es importante tener en cuenta la ubicación del inmueble y los servicios adicionales que se prestan.