## Instituto Tecnológico de Costa Rica, Escuela de Computación, Bases de Datos II Prof. Rodrigo Núñez Núñez, Proyecto II, Big Data, parte IV, 10%

## Discovery

Utilizando los data sets *detallados* (no los sumarizados), de listings y calendar de AirBnb, en <a href="http://insideairbnb.com/get-the-data.html">http://insideairbnb.com/get-the-data.html</a> , los mismos que usó para la parte II, proceda a realizar las siguientes actividades:

- 1. Utilizando ya sea Tableau o PowerBI (trial version), conéctelo a MemSQL debidamente cargado con los datos consolidados de la parte III de éste mismo proyecto.
- 2. Proceda a realizar un análisis que le permita descubrir las siguientes incognitvas:
  - a) Cuáles son las diferencias más sustanciales de los modelos de casas, habitaciones o apartamentos entre las diferentes ciudades
  - b) ¿Existe alguna relación entre la distribución de una propiedad y la recurrencia de su reservación?
  - c) ¿Existe alguna relación que mejore o desmejore el booking cuando las propiedades tienen cocina, parqueo o ambas?
  - d) ¿En qué estaciones del año para cada ciudad es cuando más bajan las reservaciones? ¿Sabiendo esto, esos bookings mínimos estarán relacionados a algunos de los amenities en particular?
  - e) ¿Las calificaciones bajas de limpieza, han incidido negativamente en la decisión de bookear la propiedad o no ha sido tan relevante?
  - f) ¿Las calificaciones de limpieza, estarán influenciadas positiva o negativamente por la influencia de la temperatura, la lluvia o las nevadas?

## **Otros aspectos**

- El trabajo es individual
- Cualquier mínima sospecha de copia anulará el trabajo de los involucrados.
- La revisión será con cita en el lab H del Tec
- Fecha de revisión martes 13 de noviembre, 5pm a 7pm. Se revisará la parte III y parte IV juntas.