

Load and Summarize

Utilizando los data sets **detallados** (no los sumariados), de listings y calendar de Airbnb, en <http://insideairbnb.com/get-the-data.html> , proceda a realizar las siguientes actividades:

1. Seleccione 3 ciudades que sean disímiles, ya sea continental o por masa cultural, ej: Latinoamérica, Norteamérica, Europa, etc.
2. Cargue los datasets de listings y calendar de las 3 ciudades en el metastore como tablas.
3. Usando Impala, cree un query que cree las tablas necesarias y las llene con los siguientes sumariados de información
 - a) Cantidad de listings por ciudad clasificados entre 5 estratos de rangos de precio
 - b) Asumiendo que la limpieza de un lugar es de 3 valores: 1 excelente, 2 moderado, 3 no tan limpio. Agrupe la cantidad de listings por ciudad por nivel de limpieza
4. Usando Hive, cree un query que cree las tablas necesarias y las llene con los siguientes datos sumariados:
 - a) Top 3 proporcional de las ciudades que tengan más experiencia como host de Airbnb
 - b) Cuáles son los top 7 amenities más comunes por ciudad

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 jueves 25 de octubre, 7pm a 9pm