

# Tarea Neo4j

En este documento tendréis que elaborar un documento, a modo de tutorial, que incluya toda la información (código, capturas de pantalla, resultados). En el aula virtual subid el documento en formato PDF.

## Introducción

Yelp ayuda a las personas a encontrar negocios locales basándose en reseñas, preferencias y recomendaciones. A finales de 2018 se habían escrito más de 180 millones de reseñas en la plataforma. Desde 2013, Yelp ha organizado el reto Yelp Dataset, un concurso que anima a la gente a explorar e investigar el conjunto de datos abiertos de Yelp. A partir de la ronda 12 (realizada en 2018) del desafío, el conjunto de datos abiertos contenía:

- Más de 7 millones de reseñas más consejos
- Más de 1,5 millones de usuarios y 280.000 fotos
- Más de 188.000 negocios con 1,4 millones de atributos
- 10 áreas metropolitanas

Desde su lanzamiento, el conjunto de datos se ha hecho popular, con cientos de artículos académicos escritos utilizando este material. El conjunto de datos de Yelp representa datos reales muy bien estructurados y altamente interconectados. Es un gran escaparate para los algoritmos de grafos que también puedes descargar y explorar.

## Caso práctico

Utilizaremos un caso práctico en el que trabajamos para una empresa de información de viajes. El objetivo es encontrar buenas recomendaciones de lugares para alojarse y cosas que hacer en ciudades importantes como Las Vegas.

Para hacer esta práctica hay 2 opciones, la primera, que es la recomendable, es coger de los archivos un subconjunto de datos con 2000 elementos. Es decir, por cada JSON que generará un Label cargar únicamente en neo4j los 2000 nodos primeros. La otra sería con todos los datos, implica usar índices, neo4j-admin, apoc.periodic.iterate, etc.

## Importación de datos

Los datos se encuentran en <https://www.yelp.com/dataset>

1. Escribir el comando de importación, recomendable usar LOAD\_CSV.
2. Realizar un grafo del modelo y mostrar captura.
3. Describir los nodos y sus relaciones.

## Aplicación de viajes

En esta aplicación estamos interesados en encontrar los hoteles más importantes y los usuarios más importantes, existen una infinidad de maneras de obtenerlo pero en este caso nos basta con obtener la siguiente información:

1. Encuentra los 10 hoteles con mayor número de reviews.
2. Encuentra los 10 usuarios con un número mayor de reviews realizadas.

## Consultoría empresa de viajes

En este caso se ha detectado que los usuarios más influyentes son los que han ido al hotel Bellagio Hotel

1. Encuentra los 50 usuarios con mayor número de reviews que han hecho una review del hotel Bellagio Hotel.
2. Buscar todos los hoteles que estos usuarios han hecho una review, no vale introducir los usuarios en el Where de manera manual.
3. Obtén el hotel con mayor número de reviews de los usuarios obtenidos en el apartado 1 de esta sección

# Evaluación

Importación de datos	2
Aplicación de viajes	3
Consultoría empresa de viajes	3
Trabajar con todos los datos	2
TOTAL	10