

**Análisis de Big Data:  
EAA361A**

**Ejercicio Extra**

Profesor: Cristián Vásquez

Ayudante: Pablo González

**Recomendación de libros**

Utilizando el set de datos “Book-Crossing”, con información acerca de las calificaciones realizadas por miembros de dicha comunidad a 271.379 libros, responda las siguientes preguntas:

1. Obtenga todos los nombres de usuarios, ISBN y calificaciones de todas las reseñas escritas por gente entre 20 y 29 años que ha estado en China.
2. Utilizando su respuesta anterior, obtenga el nombre de los libros y su calificación promedio en base a dichas reseñas.
3. Obtenga los títulos, reseñas y ubicación del usuario para las novelas escritas por Natsume Sōseki o por Dazai Osamu.
4. Obtenga los 10 libros más leídos dentro de aquellos con un rating implícito.
5. **Propuesto 1:** ¿De los mejores 20 libros con más de 5 reseñas, de dónde vienen sus votantes?
6. **Propuesto 2:** De los 20 libros con mayor cantidad de reseñas, calcular cuál tiene el mayor rating promedio

*\* Datos extraídos de [www2.informatik.uni-freiburg.de/~ciegler/BX/](http://www2.informatik.uni-freiburg.de/~ciegler/BX/) con fecha 28/03/2020*