GUÍA DE LECTURA: Clustering

1. ¿Cuál es el objetivo de un algoritmo de clustering?
2. ¿Qué función tienen las medidas de distancia? ¿Qué diferencia existe entre Euclidiana, Manhattan y Mikowsky?
3. ¿Qué tareas transformación son requeridas para aplicar clustering?
4. ¿Qué ocurre con las variables categóricas?
5. Para los algoritmos de clustering presentados en clase (k-mean y jerárquicos), describir los siguientes aspectos:
   1. La forma de los clusters que pueden ser determinados (Cap. 7 Han)
   2. Parámetros de entrada
   3. Limitaciones
6. ¿Utilidad tiene un dendrograma? ¿Cómo se construye?
7. ¿Para qué sirven las estrategias de linkage? ¿Qué tipos existen? ¿Qué limitaciones pueden tener?
8. ¿Cómo se interpretan los centroides en un ajuste de k-means?

Bibliografía sugerida:

* Han, J., Kamber, M., & Pei, J. (2012). Data mining concepts and techniques third edition. University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University. Cap. 7
* Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining* (Vol. 4). John Wiley & Sons. **Cap 10** [ [pdf](https://drive.google.com/file/d/0B5RFUdXfMb1iZXFhQlpqV014OUk/view?usp=sharing&resourcekey=0-2WUqBmWAWgz06vYAjSHLFw) ]