☰ **Item Navigation**

Modeling text data with a hierarchy of clusters

Hierarchical clustering refers to a class of clustering methods that seek to build a **hierarchy** of clusters, in which some clusters contain others. In this assignment, we will explore a top-down approach, recursively bipartitioning the data using k-means.

If you are using Turi Create

An IPython Notebook has been provided below to you for this assignment. This notebook contains the instructions, quiz questions and partially-completed code for you to use as well as some cells to test your code.

- Download the Wikipedia people dataset in SFrame format:

📄 **people_wiki.sframe** ZIP File		Download file ⤓

- Download the companion IPython notebook:

📄 **CLU06-NB01.ipynb** ZIP File		Download file ⤓

- Download a collection of helper functions:

📄 **em_utilities.py** ZIP File		Download file ⤓

💡 **Important reading**   ✕

This reading explains critical concepts that
- Save the files in the same directory (where you are calling IPython notebook from) and unzip the come up in the next assignment.
data file and model file.

Was this helpful? [Yes] [No]

Open the companion IPython notebook and follow the instructions in the notebook. The
~~instructions below~~ ~~do not apply~~ ~~to users of Turi Create~~