

**Autoencoders****More Dimension Reduction****Clustering**

✓ **Video:** Clustering, K-Means  
3 min

✓ **Reading:** Further Reading:  
Clustering  
11 min

📋 **Quiz:** Unsupervised  
Learning  
5 questions

**Handling Missing Data****EOW (End Of Week)**

# Further Reading: Clustering

## K-Means

Here is the H2O documentation on k-means:

<http://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/k-means.html>

Here is the Wikipedia article on k-means:

[https://en.wikipedia.org/wiki/K-means\\_clustering](https://en.wikipedia.org/wiki/K-means_clustering)

This recent article shows some problems with k-means and alternatives (the Wikipedia article also covered them):

<https://www.datascience.com/blog/k-means-alternatives>

## Curse of dimensionality

This is an important topic to understand, as most interesting data sets will typically have quite a few dimensions, and some might have dimensions numbering in the tens of thousands or higher.

[https://en.wikipedia.org/wiki/Curse\\_of\\_dimensionality](https://en.wikipedia.org/wiki/Curse_of_dimensionality)

The answer [here](#) argues that k-means is still fine with high-dimensions:

✓ Complete

Go to next item