* [Using a Function to Transform a Feature](#gjdgxs)
* [MinMax Scaling a Feature](#30j0zll)
* [Dealing with Dates](#1fob9te)
* [Clustering and Visualizing Clustering](#3znysh7)

# CSE 450 - Machine Learning & Data Mining

## Module 02 — More Hints

Don't Sell Yourself Short

Coding and Data Science is hard. Most of what I'm going to put here, you could look up on your own from official docs, tutorials, and Stack Overflow questions.

As we proceed throughout the semester, you'll get fewer hints and be expected to look more stuff up on your own.

If you haven't tried independent research yet, you might want to try that for a few hours, then come back here.

## Using a Function to Transform a Feature

Sometimes you want to do a complex transformation where you create a new feature based on the data in one or more existing features, or where you need to use boolean logic in your transformation.

One way to accomplish this is with apply() plus a function or lambda, as shown in the following colab notebook:



## MinMax Scaling a Feature

If we need to do some kind of distance-based algorithm with these features, it would help to transform them to the same scale first.

Min-max scaling will make the values span the range 0.0 - 1.0, as shown in the following colab notebook:



## Dealing with Dates

Sometimes we need to calculate the span between two dates using different units. Or, we might need to to do more complex calculations on a date using boolean logic. The following colab notebook shows some examples of that:



## Clustering and Visualizing Clustering

If you're struggling to use the elbow method to determine the right value for the "k" in k-means, or not clear on how the different ways to visualize clusters, this colab notebook should help:

