# Smoothing, Blurring, running average

Some students have discovered an imprompty survey ([Optional survey Linear Algebra W2](https://canvas.vu.nl/courses/57731/quizzes/36797)) and provide some feedback that the explanation of the 'blur' exercise could have been more clear. Here is some more background of the exercise.

## Denoising / smoothing / blurring

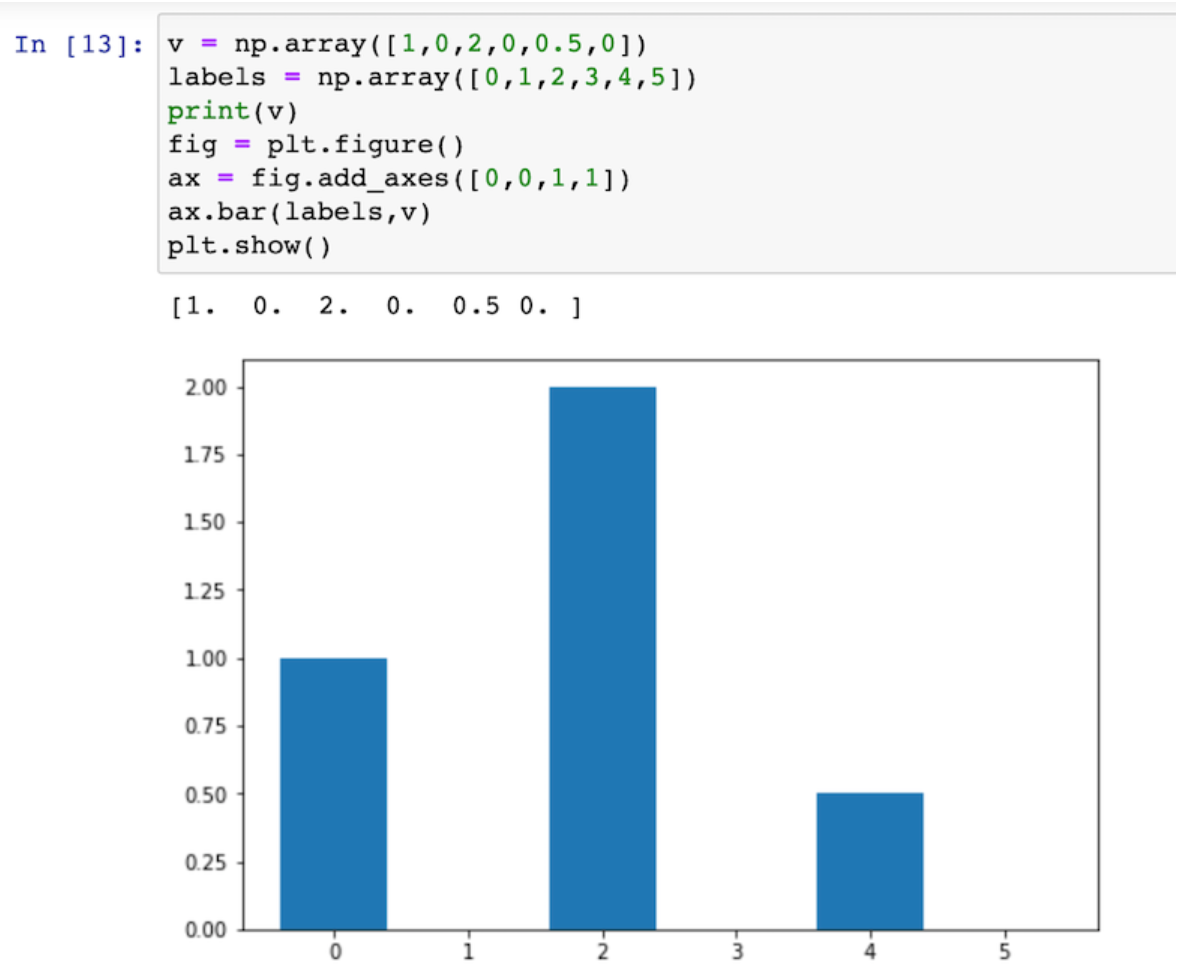
A lot of AI data is noisy. This means that the data that you are interested in, is hidden under sudden spikes. This is the case for one-dimensional data, but also two-dimensional data such as imaging. In imaging, noise occurs as single black or white pixels. The solution to this problem is to apply a filter. The filters work by blending each pixel with its neighbours: this way, single white pixels will become less noticable, while blocks of white pixels, real objects, remain visble.

The advantage of the filters is that it removes noise. A common name however is called denoising. A side-effect is that the edges of objects become smoother. Another name is therefore smooting. A disadvantage is that objects become blurred, especially if the filter is over-applied: not only the noise disappears, but all data disappears. The filters are therefore also called blurring.

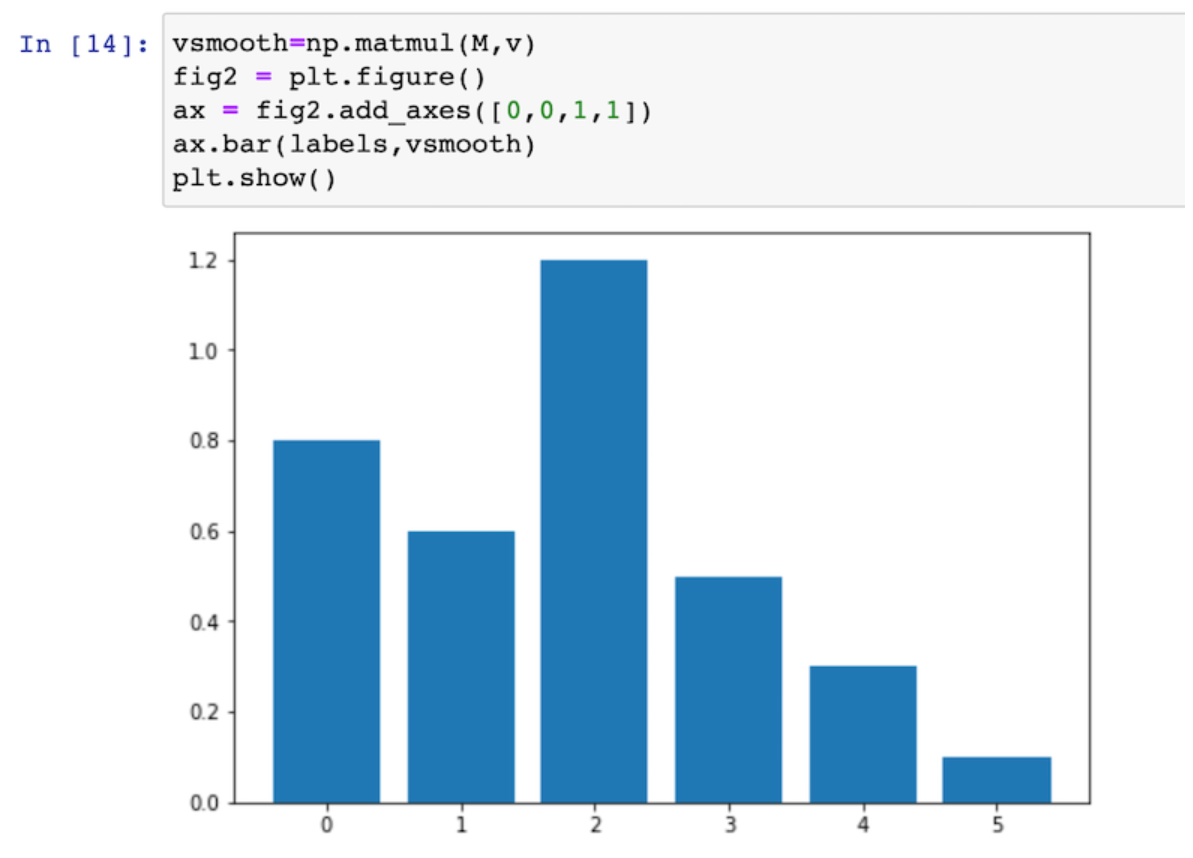
## One-dimensional data

In this course, we want to show the full range of linear algebra applications. We use 2 and 3 dimensional vectors to demonstrate basic linear algebra. but try to use larger vectors (size 6, 12, 100, 1000) to demonstrate AI application. In an ideal world we would also show picturesized vectors (a 400x300 pixel black and white image would be represented by a vector of length 120.000 ). It is harder to understand what each vector means, but that is part of the learning goals: you have to learn to recognize data that could be a vector, learn to interpret vector data, and understand vector sizes.

For the blur exercise, we use time series data: a series of values, one for each time period. We use vectors of length 12, with a value for each month. In financial data such time series are very common, e.g. the closing day value of a stock is a time series. The vectors of length 5 can be seen as a value for each day of the week. You could make a bar chart of the data. like this:



You see three spikes in the data. If you smooth these out, you see the following results



I hope you see that this chart is more 'smooth': the data of days next to each other is blended. Whether this is useful depends on the applications. Some AI algorithms work much better with denoised/smoothed data.

Note that the data in this case is represented as vectors. The blur operation that we do, is represented as a matrix. This is typical for linear algebra: data is represented by vectors, operations by matrices. Matrix vector multiplication is applying an operation on a vector. Matrix multiplication is used to combine multiple operations.

Good luck with the exercise. If you have any questions, use the discussion!

Small note: To make the code examples work, you need to import a library. Put "import matplotlib.pyplot as plt" in your notebook and run this once. Mathplotlib is a nice library for charts.