1 Numpy

I had these notes with me on Friday to talk about Numpy. Look at those for the content.

2 Project 2 Prep

Today, in lab, we will begin writing the class that will read and store data.

- Not all data is numeric. Consider the example from last week, in which one of the columns contained the name of a country. The other two columns had numeric data. We will draw a logical distinction between raw and numeric data. In our class, we will store data in both raw form and numeric form.

- Indexing can get complicated. If only a subset of our columns are numeric then the numbering of the columns is different between the raw and numeric data sets. So I propose a design in which we select columns by using the column header (e.g. “GDP per capita”). Then, to map the header string to the column index we use two dictionaries – one of the raw data and one for the numeric data.

- In the example from last week, when we transformed the data from data coordinates to view coordinates, we normalized each column separately. This makes sense when the columns contain different kinds of information (such as life expectancy and GDP). But what if there are multiple columns with related information (e.g. height at age 10 and height at age 20). If we normalize those columns separately, then it no longer makes sense to compare their values to each other. In this case, we may want to normalize the columns together (e.g. subtract the overall min and divide by the overall extent).

3 Moving on

We move on to the notes associated with lecture 7, the 3D viewing pipeline.