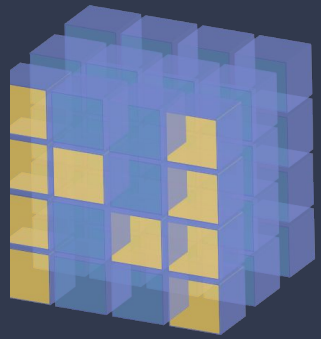
Python Libraries

Session -3
Lean-In (ML Circle)



1. ***NumPy** is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.*
2. **In Numpy, number of dimensions of the array is called rank of the array.**

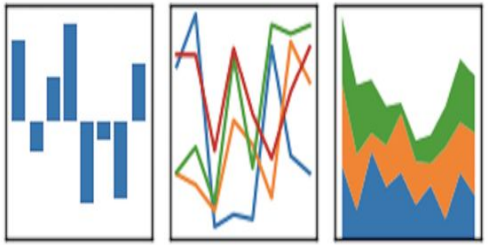
Difference: List and Array

- The main difference between a list and an array is the functions that you can perform to them.
- It does take an extra step to use arrays because they have to be declared while lists don't because they are part of Python's syntax
- For arithmetic functions to list, list converted to array

Ex: you can divide an array by 3, and each number in the array will be divided by 3 and the result will be printed if you request it. If you try to divide a list by 3, Python will tell you that it can't be done

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



- ***Pandas*** is the most popular python library that is used for data analysis.
- We can analyse data in Pandas with:
  1. Series
  2. DataFrames

1. Series

a = pd.Series(Data, index = Index)

Data can be :

- Scalar Value
- Dictionary
- Ndarray

2. DataFrames

DataFrames is two-dimensional(2-D) data structure defined in pandas which consists of rows and columns.

a = pd.DataFrame(Data)

Data can be:

1. One or more ***dictionaries***
2. One or more ***Series***
3. ***2D-numpy Ndarray***

- Matplotlib is a multi-platform data visualization library built on NumPy arrays.

1. Line Plot
2. Bar Graph
3. Histogram
4. Scatter Plot