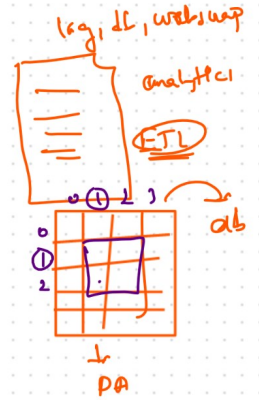


# Data Engineering Essentials

\* pandas

\* numpy



pandas: Data manipulation & analysis

Core Concepts:

\* Series: 1D labeled array, building block of Dataframe

\* Dataframe: 2D labeled data structure with columns

\* Index: labels for rows, this enables fast data lookup

[20, 10, 15, 20, ...]

[ [20, 10]  
[15, 20]

Dataframe

Student DataFrame

Column	id	name	age	school
0	1	John	28	New York
1	2	Bob	34	Paris
2	3	Peter	29	Berlin
3	4			
4				

Dataframe

	name	age	city
0	John	28	New York
1	Bob	34	Paris
2	Peter	29	Berlin

[20  
30]