

Lecture 18, 20 October 2026

Madhavan Mukund

<https://www.cmi.ac.in/~madhavan>

Programming and Data Structures with Python

Useful libraries for Data Science

Numpy - n-dimensional homogeneous arrays



Tables (like SQL)

Col 1	Col 2	...	Col k
Type 1	Type 2	...	Type k

Not exactly
a 2D
Numpy array

Tables in Python

Pandas - built on Numpy

Columns - 1D array - Series

Table, sequence of columns - Data Frame

Spreadsheets → Text, csv file

↳ comma separate value
Pandas can import csv files - descriptive statistics

Lists vs dictionaries

Numpy - row/col indices are numbers

Table - column names