Introduction to Data Science
GIRI NARASIMHAN, SCIS, FIU

Momentos

- Slides and Audio online

- You need to register
  - ❑ Go to https://fiu.momentos.life
  - ❑ If you don't already have an account
    - ▪ Click on "Sign up"
    - ▪ Follow instructions & use referral code: **9PQG2X**
  - ❑ If you have an account, "Add Course" with code **9PQG2X**
  - ❑ Verify account using link sent to email

Languages for Data Science

- Python & Perl
- R
- MATLAB
- SQL
- Ruby

Two kinds of data

- Structured
  - Lists
    - Arrays, Tables and Spreadsheets
  - Strings
  - Matrices
    - Images
  - Maps & Associations
    - (Key, Value) Pairs
  - Time Series & Trajectories
    - Audio, Video
- Unstructured

Basics of Python

Python …

▶ Helps to perform actions on data

▶ Provides features to do this conveniently

▶ Acts as a glue language to tie together code in multiple languages

▶ Helps to quickly prototype an idea

▶ Not well suited for concurrent programs, multi-threaded applications

▶ Has a lot of specialized libraries for data science (e.g., NumPy, SciPy)

▶ Provides **DataFrame** to provide structured data (e.g., pandas)

▶ Provides plotting tools (e.g., matplotlib)

▶ Provides notebook facilities (e.g., iPython, JuPyter)