PANDAS

Pandas

- Python library to work with CSV tables

- Kaggle course:
  https://www.kaggle.com/learn/pandas

Workflow

- Create a panda object
- Read the file
  - read_csv
- Vector & Dictionary
  - Index by position
    - iloc, loc
  - Index by column name

- Operations
  - Select tuples (rows) based on attribute value
  - Create new column
  - Column statistics
  - GroupBy-Aggregate
  - Sort on columns
  - Missing values (isnull)
  - Column renaming
  - Join two pandas

Panda Programming in Python Notebooks

- Function calls for every operation

- Create new pandas from the input

- Imperative
  - Write code that tells what to do

- Interactive
  - Get the output of every operation (cell)
  - Visualization
  - Debugging