Python Lab
Pandas - III

Proteomics Informatics, Spring 2014
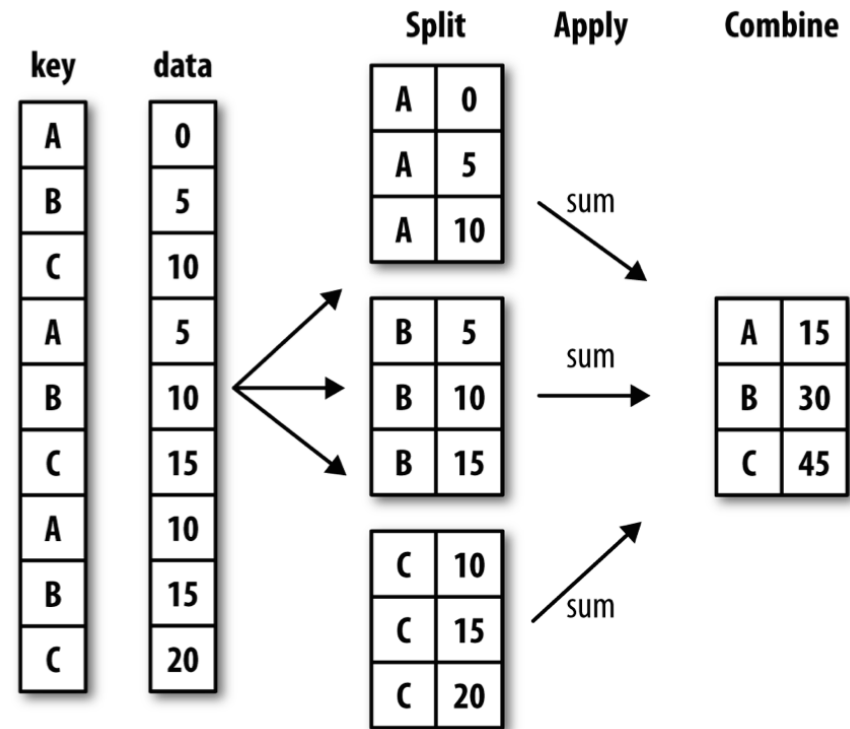
Week 11

11th Apr, 2014

Himanshu.Grover@nyumc.org

Data Aggregation, Transformations & Other Manipulations

- Part of most data analysis workflows

- Categorize data into groups, according to keys

- Apply an operation to each group. For ex.
  - within group means/StDev etc.
  - within group normalizations, or other transformations etc.
  - Group-wise modeling (regression etc.)

Pandas' **"groupby"** Mechanics: Split-Apply-Combine

- Split Pandas objects into pieces (**groupby**)

- Apply:
  - Compute group statistics
  - Apply functions group-wise to columns (**aggregate**, **transform**)

- Other, more generic manipulations (**apply**)

From "Python for Data Analysis", Chapter 9