



# Data Science Course



## Lecture 5



# Overview

---

- Join / Merge / Concatenate
- Data Aggregation
- Correlation
- Activity!

# Join / Merge / Concatenate

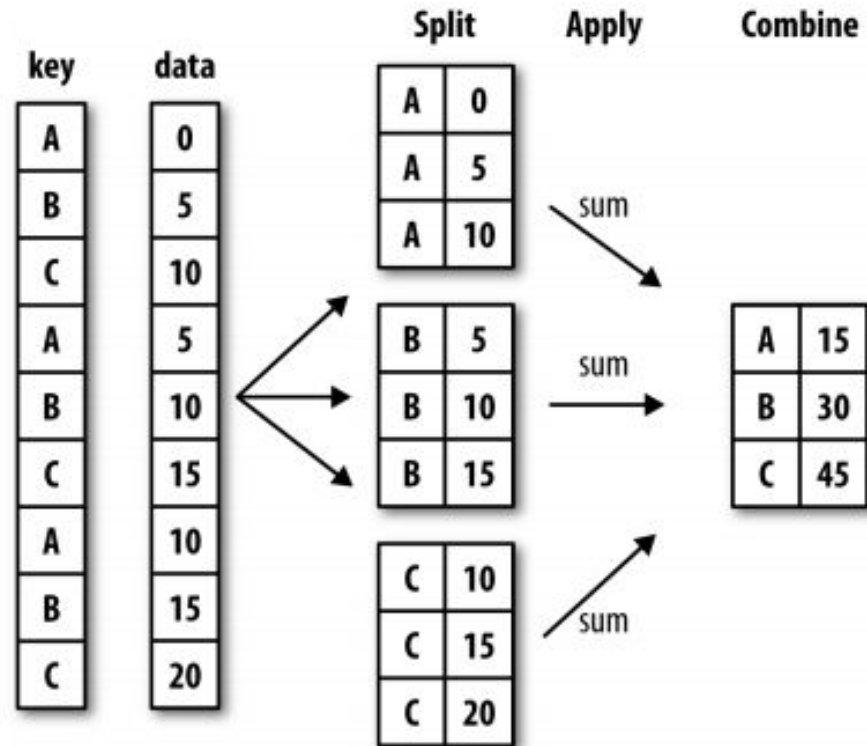
---

- Go to notebook!

# Data Aggregation

---

- GroupBy
- Optimized groupBy methods.



Function name	Description
count	Number of non-NA values in the group
sum	Sum of non-NA values
mean	Mean of non-NA values
median	Arithmetic median of non-NA values
std, var	Unbiased ( $n - 1$ denominator) standard deviation and variance
min, max	Minimum and maximum of non-NA values
prod	Product of non-NA values
first, last	First and last non-NA values

# Correlation

---

