pandas

Kristo Raun

Data Engineering 2022

UNIVERSITY OF TARTU

What is pandas?

⬥ pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.
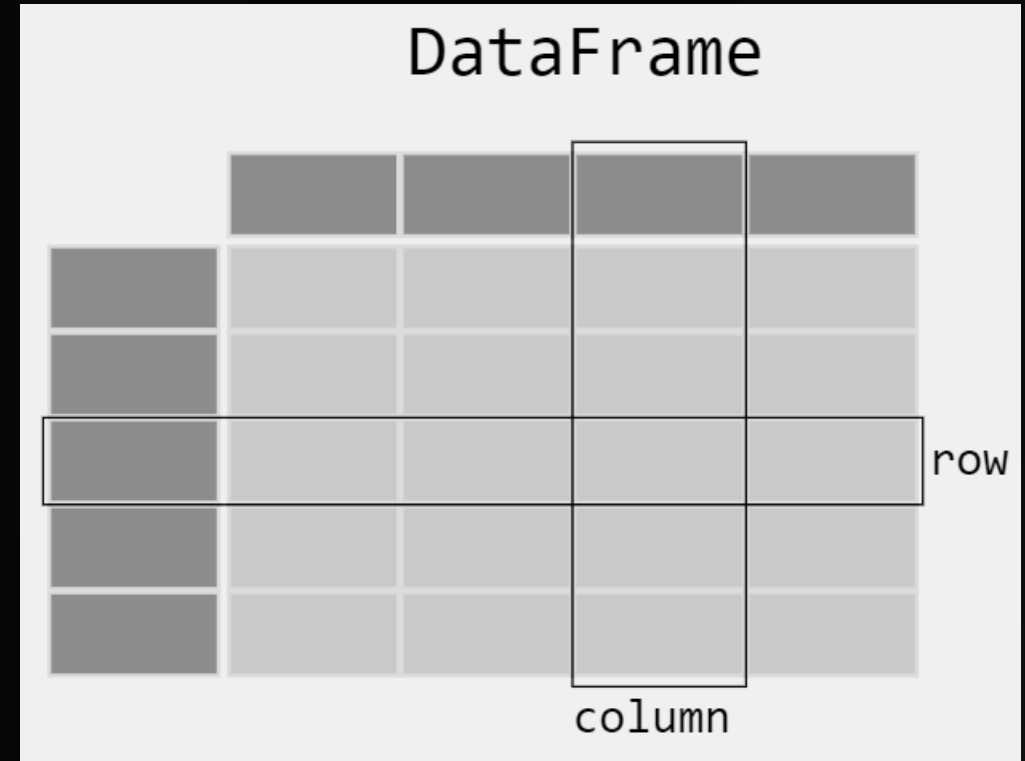
Why pandas?

Why pandas?

◈ Data analysis – what is this data about?

Why pandas?

- Data analysis – what is this data about?

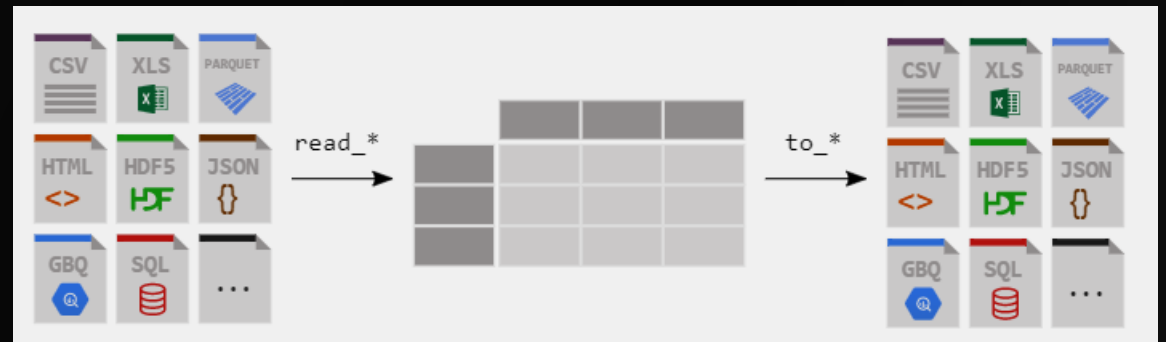- Data transformation – what does this data need to look like?

About

◇ DataFrame – handling tabular data

◇ Source: https://pandas.pydata.org/docs/getting_started/index.html

About

◈ DataFrame – handling tabular data

◈ Read … to

About

- DataFrame – handling tabular data

- Read … to

- In memory analytics

◈ Source: https://pandas.pydata.org/docs/getting_started/index.html

Final words…

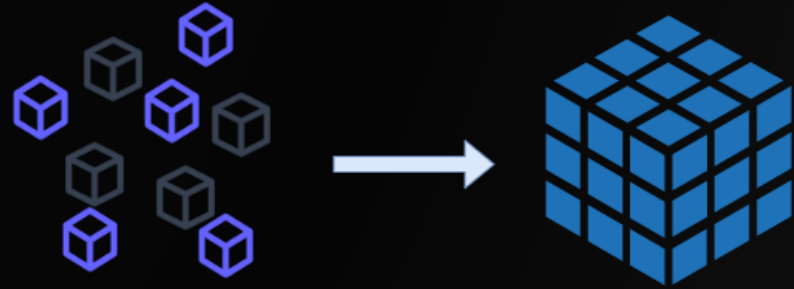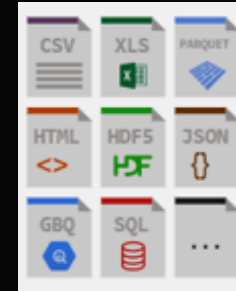# Final words…

◈ pandas is great for small to medium data
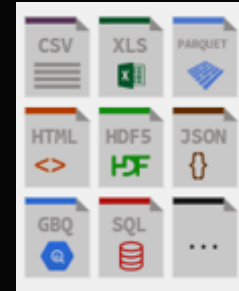
Final words...

◈ pandas is great for small to medium data

◈ pandas allows for almost all conceivable data manipulation techniques

Final words…

- pandas is great for small to medium data

- pandas allows for almost all conceivable data manipulation techniques

- pandas is NOT usable for truly big data