Introduction to Pandas (1.5 hours)

We’ll be using a library called pandas, which is for data manipulation, analysis and automation.

**Learning objectives:**

* Python Libraries - Install and importing Pandas
* Opening (‘read’) an excel file
* Create a dataframe (‘df’)
* Use some pandas functions e.g. head(), tail(), describe(),
* Merge two sheets,
* Save (‘write’) data back to a file
* If time, we can talk about using another library called ‘matplotlib’ to visualise our data

**Before the meeting – 5 minutes of prep:**

1. Please sign up/log in to your [Replit](https://replit.com/) account if you do not have Python installed locally already.
2. Ask perplexity.ai “I’m a beginner coder, what is a python library?” and  “What is pandas, the python library, and what can it do?”

**Resources:**

* Python files:
  + 1\_working\_with\_dataframes.py
  + 2\_multiple\_sheets.py
* Data files:
  + Holdings.csv
  + Journals.csv
  + working\_with\_de\_gruyter\_data.xlsx
* Replit <https://replit.com/>
* Importing libraries in to a repl https://docs.replit.com/replit-workspace/dependency-management
* Pandas documentation <https://pandas.pydata.org/docs/>

"Intro to Pandas" by [Sheldon Korpet](https://www.linkedin.com/in/sheldon-korpet/) is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/deed.en)