



Intro



Pandas est une bibliothèque écrite pour le langage de programmation Python permettant la manipulation et l'analyse des données. Son nom est dérivé de “panel data”, une expression venant de l'économétrie.

Pandas est principalement utilisé pour l'analyse de données et permet de les importer de différents types de fichiers (CSV, SQL, Excel ou encore JSON). Il offre la possibilité d'effectuer différentes manipulations sur un dataset, de la fusion, au remodelage, en passant par la sélection, ou encore le data cleaning.

Accompagné de NumPy (manipulation de matrices ou tableaux multidimensionnels et fonctions mathématiques, base de SciPy), et Matplotlib (traçage et visualisation des données sous formes de graphiques), ils forment un trio indispensables en ce qui concerne l'analyse de données sur Python.

Demo





Ressources

- <https://pandas.pydata.org/>
- https://pandas.pydata.org/pandas-docs/stable/user_guide/10min.html
- <https://openclassrooms.com/fr/courses/4452741-decouvrez-les-librairies-python-pour-l-a-data-science/5558996-passez-de-numpy-a-pandas>
- <https://www.youtube.com/watch?v=zZkNOdBWgFQ>
- <https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-tutor2-python-pandas.pdf>
- <https://colab.research.google.com/github/boileaum/cours-python/blob/master/07-pandas.ipynb#scrollTo=TcxPPFFBWB1x>
- <https://www.javatpoint.com/pandas-vs-numpy>
- <https://github.com/CPingeon/veillepandas>