

Tipsheet for Data Processing Tools

1. PDF Join – allows you to merge several PDF files into one: <https://www.pdfjoin.com/>
2. Tabula – allows you to extract data from PDFs (doesn't work for scanned documents): <http://tabula.technology/>
3. CometDocs – allows you to extract data from PDFs: <http://www.cometdocs.com/>
4. ScraperWiki – allows you to extract data from PDFs and the web: <https://scraperwiki.com/>
5. Datawrapper <https://datawrapper.de/>
6. Infogram <https://infogr.am/en>
7. Tableau public <http://www.tableau.com/public/>
8. CartoDB <http://cartodb.com/>
9. TileMill <https://www.mapbox.com/tilemill/>
10. Google Spreadsheets
IMPORTHTML: <https://support.google.com/docs/answer/3093339?hl=en>
11. Chrome plugin for scraping:
<https://chrome.google.com/webstore/detail/scraper/mbigbapnjcgaffohmbkdlecacpepngjd?hl=en>
12. More complex scrapers: <https://import.io/>
13. Adobe Acrobat: <https://creative.adobe.com/products/acrobat>
14. Google Advanced search: <https://support.google.com/websearch/answer/35890?hl=en>
15. Find unprotected directories: <http://www.makeuseof.com/tag/how-to-find-unprotected-website-directories-get-interesting-files/>
16. Cool data dashboards: <https://www.silk.co/>
17. Automated data analysis: <https://www.statwing.com/>
18. Scraping for developers: <https://morph.io/>