Data science resources index

This document provides a selection of links to websites I have found with useful resources for various aspects of data science, organized by topic. There are many many resources not included in this document; these are merely those I have come across and used in recent years as well as a few I have found that seem useful but have not yet used. I expect to update this document periodically.

Note: have not yet gone through by data search sources or personal websites folders

Contents

Software packages

Python

1. Sean Higgins Python guide: https://github.com/skhiggins/Python_guide
    a. Style, packages, structure/organization, graphing, reproducibility
2. Python Causal inference handbook: https://github.com/matheusfacure/python-causality-handbook
3. UC Berkeley python practice: https://python.berkeley.edu/
4. D-Lab python resources: https://github.com/dlab-berkeley/python-berkeley
5. D-Lab Python fundamentals: https://github.com/dlab-berkeley/Python-Fundamentals
6. D-Lab Python data wrangling: https://github.com/dlab-berkeley/Python-Data-Wrangling
7. D-Lab intermediate python: https://github.com/dlab-berkeley/Python-Intermediate-Legacy
8. D-Lab python data visualization: https://github.com/dlab-berkeley/Python-Data-Visualization
9. D-Lab python for everything: https://github.com/dlab-berkeley/python-for-everything
10. Arthur Turrell coding for economists python: https://aeturrell.github.io/coding-for-economists/intro.html
    a. Getting started, working with data, code style, data viz, workflow, econometrics, data types, text analysis, machine learning, geospatial, math, transitioning from other languages
11. Python for Data Science: https://aeturrell.github.io/python4DS/welcome.html
    a. Basics, visualization, working with data types, programming
12. Awesome geospatial python links: https://github.com/sacridini/Awesome-Geospatial?tab=readme-ov-file#python
13. Gabor Bekes coding for data analysis with Python: https://github.com/gabors-data-analysis/da-coding-python
    a. Basics, pandas, graphs, data exploration, functions, regression, datetime, time sessions, spatial viz, cross-validation, lasso, random forest, classification
14. Gabor Bekes setting up python environment: https://github.com/gabors-data-analysis/setting-up-your-data-science-environment
15. Gabor Bekes setting up computer for python: https://gabors-data-analysis.com/howto-python/
16. Jeremy DNM python setup and quick intro
    a. https://github.com/jdnmiguel/Code-corner/blob/main/geospatial.md, https://github.com/jdnmiguel/Applied-ML/blob/main/Code/0_setup.md, https://github.com/jdnmiguel/Applied-ML/blob/main/Code/1_python.ipynb
17. Data camp data scientist with Python: https://app.datacamp.com/learn/career-tracks/data-scientist-with-python?version=5
18. NYU python data bootcamp: https://nyudatabootcamp.gitbook.io/thebook/
19. NYU intro to python for science: https://physics.nyu.edu/pine/pymanual/html/pymanMaster.html

20. Getting started with jupyter notebooks: https://www.dataquest.io/blog/jupyter-notebook-tutorial/

R

1. R introductions and setup: https://datacarpentry.github.io/r-intro-geospatial/
2. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures
3. D-Lab R push-in 4.5 hour intro: https://github.com/dlab-berkeley/R-Push-Ins
4. D-Lab R basics 1.5 hours workshop: https://github.com/dlab-berkeley/R-Haas-Workshop
5. D-Lab R basics 2 hours: https://github.com/dlab-berkeley/Fast-R
6. D-Lab R for data science: https://github.com/dlab-berkeley/R-for-Data-Science
7. D-Lab R fundamentals: https://github.com/dlab-berkeley/R-Fundamentals
8. D-Lab R causal inference: https://github.com/dlab-berkeley/R-Causal-Inference
9. D-Lab R data visualization: https://github.com/dlab-berkeley/R-Data-Visualization
10. D-Lab R data wrangling: https://github.com/dlab-berkeley/R-Data-Wrangling
11. D-Lab R functional programming: https://github.com/dlab-berkeley/R-Functional-Programming
12. D-Lab advanced data wrangling in R: https://github.com/dlab-berkeley/advanced-data-wrangling-in-R-legacy
13. D-Lab project management in R: https://github.com/dlab-berkeley/efficient-reproducible-project-management-in-R
14. D-Lab RStudio project management: https://github.com/dlab-berkeley/RStudio-Project-Management
15. DIME analytics R data visualization code and examples: https://worldbank.github.io/r-econ-visual-library/
16. Fect fixed effects package for R with pre-trends and placebo tests https://yiqingxu.org/packages/fect/fect.html
17. Sean Higgins R guide: https://github.com/skhiggins/R_guide
    a. Style, packages, structure/organization, graphing, reproducibility, version control
18. Awesome geospatial R links: https://github.com/sacridini/Awesome-Geospatial?tab=readme-ov-file#r
19. Kyle Butts Stata to R guide: https://stata2r.github.io/
20. Gabor Bekes coding for data analysis with R: https://github.com/gabors-data-analysis/da-coding-rstats
    a. Basics, data exploration, markdown, ggplot, conditionals, loops, functions, regression, datetime, timeseries, spatial viz, cross validation, lasso, random forest, classification ML
21. Gabor Bekes setting up computer for R: https://gabors-data-analysis.com/howto-r/
22. CMSA Camp data science in R applied to sports: https://www.stat.cmu.edu/cmsac/sure/2021/materials/
    a. Exploring data, data viz, clustering, supervised learning, regression, principal components analysis, nonparametric regression, splines, machine learning
23. Reproducible pipelines for social scientific research in R: https://github.com/ddekadt/MY560_reproducible_pipelines_R
24. Simon Ejdemyr R tutorials basics: https://sejdemyr.github.io/r-tutorials/basics/
25. Wickham and Grolemund R for data science: https://r4ds.had.co.nz/
26. Getting started with R markdown: https://www.rstudio.com/resources/webinars/getting-started-with-r-markdown/

27. DIME descriptive data anlsysis in R: https://osf.io/7jbfg/
28. Nick CH-K data communication with R: https://github.com/NickCH-K/DataCommSlides
29. MSU Advanced data analytics in econ in R: https://github.com/msu-econ-data-analytics/course-materials
    a. R basics, programming, data wrangling, data cleaning, data acquisition, workflows, exploratory analysis, regression, data viz, geospatial data, ML fundamentals, prediction, databases and big data, ML classification
30. Ed Rubin PhD Econometrics in R class notes: https://github.com/edrubin/EC607S24

Stata

1. D-Lab Stata fundamentals: https://github.com/dlab-berkeley/Stata-Fundamentals
2. DIME analytics Stata data visualization code and examples: https://worldbank.github.io/stata-visual-library/
3. Stata official graphics resources: https://www.stata.com/features/publication-quality-graphics/
4. Stata cheatsheets: https://geocenter.github.io/StataTraining/portfolio/01_resource/
5. DIME wiki data cleaning: https://dimewiki.worldbank.org/Data_Cleaning
6. DIME wiki data analysis: https://dimewiki.worldbank.org/Data_Analysis
7. Julien Reif Stata coding guide: https://julianreif.com/guide/
8. Nicholas Davis Stata graphs tips : https://www.nicholastdavis.com/wp-content/uploads/2018/12/davis_better_graphs.pdf
9. Asjad Naqvi Stata guide overview and links: https://github.com/asjadnaqvi/The-Stata-Guide
    a. Asjad Naqvi Stata visualizations packages: https://medium.com/the-stata-guide/welcome-to-the-stata-guide-12adf81ec3d
    b. Asjad Naqvi Stata to Latex guide: https://medium.com/the-stata-guide/the-stata-to-latex-guide-6e7ed5622856
    c. Asjad Naqvi Stata maps: https://medium.com/the-stata-guide/tagged/maps
    d. Asjad Naqvi Stata programming (Mata, MLE, regular expressinos, etc.): https://medium.com/the-stata-guide/tagged/maps
    e. Asjad Naqvi Stata Tips: https://medium.com/the-stata-guide/the-awesome-stata-tips-collection-6805afdedffa
    f. Asjad Naqvi Stata Workflow: https://medium.com/the-stata-guide/the-stata-workflow-guide-52418ce35006
    g. Asjad Naqvi Stata and Github integration: https://medium.com/the-stata-guide/stata-and-github-integration-8c87ddf9784a
10. Stata visualizations gallery and code: https://medium.com/the-stata-gallery
11. Sean Higgins Stata guide: https://github.com/skhiggins/Stata_guide
    a. Style, packages, structure/organization, reproducibility
12. Stata manual programming in Stata: https://www.stata.com/manuals14/u18.pdf
    a. Macros, program arguments, scalars and matrices, temporary objects, storing results, ado files
13. Gabor Bekes intro to data analysis in Stata: https://github.com/gabors-data-analysis/da-coding-stata
    a. Do files, prep, reshape, subsamples, graphs, data manipulation, loops, hyp testing, regression basics, visualization
14. Gabor Bekes setting up computer for Stata: https://gabors-data-analysis.com/howto-stata/
15. Data carpentry economics with Stata: https://datacarpentry.org/stata-economics/

16. Oscar Torres-Reyna getting started in data analysis with Stata: https://www.princeton.edu/~otorres/StataTutorial.pdf
17. Princeton DDS online stata tutorial : https://www.princeton.edu/~otorres/Stata/
18. Alexander Lembcke introduction to Stata: https://personal.lse.ac.uk/lembcke/ecStata/2010/MResStataNotesOct2010PartA.pdf
19. UCLA statistical consulting group Stata page: https://stats.idre.ucla.edu/stata/modules/
20. JPAL Data cleaning in Stata: https://www.povertyactionlab.org/resource/data-cleaning-and-management
21. Reghdfe 10x faster using Julia: https://www.statalist.org/forums/forum/general-stata-discussion/general/1735303-reghdfe-10x-faster
22. John Kane Making regression coef plots: https://medium.com/the-stata-gallery/making-regression-coefficient-plots-in-stata-7b100feac0cb

Git

1. D-Lab Bash-Git training: https://github.com/dlab-berkeley/Bash-Git
2. D-Lab Git fundamentals: https://github.com/dlab-berkeley/git-fundamentals
3. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures
4. GitHub howto: https://youtu.be/krDQk5ZzP8U
5. Git for economists practical guide: https://medium.com/codex/git-for-economists-a-practical-guide-7d10faf4224f

Excel

1. D-Lab Excel fundamentals: https://github.com/dlab-berkeley/Excel-Fundamentals
2. D-Lab Excel visualization: https://github.com/dlab-berkeley/visualization-in-Excel
3. D-Lab Excel basics: https://github.com/dlab-berkeley/Basics-of-Excel

Latex

1. D-Lab Latex fundamentals: https://github.com/dlab-berkeley/LaTeX-Fundamentals
2. Jorg Weber Stata table generation and intergration into Latex: https://www.jwe.cc/2012/03/stata-latex-tables-estout/
3. Kyle Butts Latex templates for articles and slides and ref responses: https://github.com/kylebutts/latex-templates
4. ESOC predodc training latex notes: https://github.com/esoclabprinceton/ESOC-Predoc-Training
5. Paul GP beamer tips: https://paulgp.github.io/2018/04/30/beamer-tips.html

SurveyCTO

1. DIME wiki SurveyCTO: https://dimewiki.worldbank.org/SurveyCTO_Additional_Topics, https://dimewiki.worldbank.org/SurveyCTO_Programming, https://dimewiki.worldbank.org/index.php?title=SurveyCTO_Coding_Practices&redirect=no

Analysis tools

Google Earth Engine

1. UC Berkeley ARE Guide: https://docs.google.com/document/d/1xoHSCS6HfNsewkhnadLXC-wgjLzwTMvvjsoAc_9iRWM/edit?tab=t.0#heading=h.tfe5u1dx8lqr
2. D-Lab Earth Engine notes: https://dlab.berkeley.edu/news/big-datasets-small-code-chunks-and-why-i-use-google-earth-engine
3. DIME wiki data management: https://dimewiki.worldbank.org/Data_Management
4. Awesome Geospatial GEE links: https://github.com/sacridini/Awesome-Geospatial?tab=readme-ov-file#google-earth-engine
5. A nice blog post walking through a protocol for land cover classification with example code by Leila Njee Bugha: https://medium.com/@dlab-berkeley/tracking-urban-expansion-through-satellite-imagery-5f51fe34e985

Machine learning

1. D-Lab machine learning working group R and python: https://github.com/dlab-berkeley/MachineLearningWG
2. Awesome geospatial deep learning links: https://github.com/sacridini/Awesome-Geospatial?tab=readme-ov-file#deep-learning
3. Gabors Data Analysis Data Exploration includes stuff on big data:
    a. Case studies with R, python, Stata code links: https://gabors-data-analysis.com/casestudies/#part-iii-prediction
    b. Overview: https://gabors-data-analysis.com/chapters/#part-iii-prediction
        i. Prediction, machine learning, lasso, big data, regression tree, random forest, boosting, classification
    c. Data and code: https://github.com/gabors-data-analysis/da_case_studies
4. Mixtape session machine learning: https://github.com/Mixtape-Sessions/Machine-Learning/
5. R
    a. Ed Rubin prediction and machine learning in econ (R) class notes: https://github.com/edrubin/EC524S24
    b. Ed Rubin PhD Econometrics class notes (R): https://github.com/edrubin/EC607S24
        i. Inference, resampling, and randomization, bootstrap
        ii. Machine learning in one lecture
    c. D-Lab machine learning in R course https://github.com/dlab-berkeley/Machine-Learning-in-R, https://github.com/dlab-berkeley/R-Machine-Learning
    d. D-Lab R deep learning: https://github.com/dlab-berkeley/R-Deep-Learning
    e. D-Lab unsupervised learning in R: https://github.com/dlab-berkeley/Unsupervised-Learning-in-R
    f. Gabor Bekes coding for data analysis with R: https://github.com/gabors-data-analysis/da-coding-rstats
        i. Basics, data exploration, markdown, ggplot, conditionals, loops, functions, regression, datetime, timeseries, spatial viz, cross validation, lasso, random forest, classification ML
    g. CMSA Camp data science in R applied to sports: https://www.stat.cmu.edu/cmsac/sure/2021/materials/

        i. Exploring data, data viz, clustering, supervised learning, regression, principal components analysis, nonparametric regression, splines, machine learning
    h. MSU Advanced data analytics in econ in R: https://github.com/msu-econ-data-analytics/course-materials
        i. R basics, programming, data wrangling, data cleaning, data acquisition, workflows, exploratory analysis, regression, data viz, geospatial data, ML fundamentals, prediction, databases and big data, ML classification
    i. Paul Goldsmith-Pinkham applied empirical methods phd course sections on ML in R: https://github.com/paulgp/applied-methods-phd
6. Python
    a. D-Lab machine learning in python course https://github.com/dlab-berkeley/python-machine-learning
    b. D-Lab python deep learning: https://github.com/dlab-berkeley/Python-Deep-Learning
    c. D-Lab python machine learning intro: https://github.com/dlab-berkeley/Python-Machine-Learning-DS-Discovery
    d. D-Lab python machine learning fundamentals: https://github.com/dlab-berkeley/Python-Machine-Learning-Fundamentals
    e. D-Lab fairness and bias in machine learning python: https://github.com/dlab-berkeley/fairML
    f. D-Lab artificial neural network machine learning fundamentals python: https://github.com/dlab-berkeley/ANN-Fundamentals
    g. Arthur Turrell machine learning in python: https://aeturrell.github.io/coding-for-economists/ml-intro.html
    h. Gabor Bekes coding for data analysis with Python: https://github.com/gabors-data-analysis/da-coding-python
        i. Basics, pandas, graphs, data exploration, functions, regression, datetime, time sessions, spatial viz, cross-validation, lasso, random forest, classification
    i. Jeremy DNM intro to ML for economists python: https://github.com/jdnmiguel/Applied-ML

Web scraping

1. Claussen and Peukert Guide to data crawling paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3403799
2. R
    a. William Marble guide to webscraping in R: https://github.com/wpmarble/webscraping
    b. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures
    c. Rochelle Terman data from the web with R: https://plsc-31101.github.io/course/collecting-data-from-the-web.html#web-apis
3. Python
    a. Sean Higgins code for webscraping in python: https://github.com/skhiggins/PythonTools
    b. D-Lab python web APIs: https://github.com/dlab-berkeley/Python-Web-APIs
    c. D-Lab python web scraping: https://github.com/dlab-berkeley/Python-Web-Scraping

d. D-Lab web scraping examples with python and raw API calls: https://github.com/dlab-berkeley/web-scraping-examples
e. D-Lab python data from web: https://github.com/dlab-berkeley/python-data-from-web

Geospatial

1. Geospatial innovation facility UC Berkeley trainings : http://gif.berkeley.edu/support/workshops.html
2. Awesome geospatial index of links, huge variety and detail: https://github.com/sacridini/Awesome-Geospatial
3. Awesome geospatial data sources links: https://github.com/sacridini/Awesome-Geospatial?tab=readme-ov-file#data-sources
4. Datasets, tips, summary of research research, tutorials, and more: https://www.spatialedge.co/
5. Introduction to GIS Programming in Python (Qiusheng Wu): https://github.com/giswqs/intro-gispro/tree/main
    a. With python code and examples
    b. https://gispro.gishub.org/
6. Geographic data science: https://pietrostefani.github.io/gds/
7. R
    a. Spatial analysis in R with sf: https://jacobgellman.github.io/files/eco-data-sci-sf.html
    b. Geocomputation with R: https://r.geocompx.org/
    c. Geospatial data carpentry UW Madison R: https://uw-madison-datascience.github.io/2019-07-11-uwmadison-dc/
    d. D-Lab R geospatial fundamentals: https://github.com/dlab-berkeley/R-Geospatial-Fundamentals
    e. D-Lab geocoding in R: https://github.com/dlab-berkeley/Geocoding-in-R
    f. D-Lab Geospatial fundamentals R sp: https://github.com/dlab-berkeley/Geospatial-Fundamentals-in-R-sp
    g. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures
    h. R as GIS for economists Taro Mieno: https://tmieno2.github.io/R-as-GIS-for-Economists-Quarto/
8. Python
    a. Using python in QGIS introduction: https://anitagraser.com/pyqgis-101-introduction-to-qgis-python-programming-for-non-programmers/
    b. D-Lab python geospatial fundamentals: https://github.com/dlab-berkeley/Python-Geospatial-Fundamentals
    c. Arthur Turrell geospatial with python: https://aeturrell.github.io/coding-for-economists/geo-intro.html

Big data

1. Raj Chetty Big Data in Economics class notes: https://opportunityinsights.org/course/
    a. Topics include equality of opportunity, education, health, the environment, and criminal justice. In the context of these topics, the course provides an introduction to basic statistical methods and data analysis techniques, including regression analysis, causal inference, quasi-experimental methods, and machine learning.

2. Gabors Data Analysis Data Exploration includes stuff on big data:
    a. Case studies with R, python, Stata code links: https://gabors-data-analysis.com/casestudies/part-I/
    b. Overview: https://gabors-data-analysis.com/chapters/
    c. Data and code: https://github.com/gabors-data-analysis/da_case_studies
3. Blumenstock UC Berkeley Big Data in Development class (Python): https://sites.ischool.berkeley.edu/bdd/
    a. A nice blog summary by Leila Njee Bugha: https://medium.com/@dlab-berkeley/using-big-data-for-development-economics-97aacf975f01
4. Data in macro development STEG lectures: https://steg.cepr.org/events/virtual-course-data-macro-development
    a. Intro dev accounting, macro data, human capital data, LSMS, ag data, multinational production data, international trade data, public finances and expenditures, geographic data and night lights, cell phone data, climate and weather data, primary data collection, gender data, time use data
5. MSU Advanced data analytics in econ in R: https://github.com/msu-econ-data-analytics/course-materials
    a. R basics, programming, data wrangling, data cleaning, data acquisition, workflows, exploratory analysis, regression, data viz, geospatial data, ML fundamentals, prediction, databases and big data, ML classification

Generative AI/LLMs

1. Generating research articles with references to published sources Storm.genie.stanford.edu
2. Lit reviews: Undermind AI (MIT)
3. D-Lab intro to LLMs for exporatory research: https://github.com/dlab-berkeley/LLMs-Exploratory-Research
4. D-Lab ChatGPT prompt engineering: https://github.com/dlab-berkeley/prompt-engineering
5. Anton Korinek Generative AI for econ research: https://www.genaiforecon.org/index.html
6. Paul GP using AI in research and teaching practical guide: https://paulgp.github.io/2024/06/24/llm_talk.html
    a. VS code and copilot, using Chat LLMs, images from midjourney, running own LLM locally

Text analysis/NLP

1. D-Lab python text analysis: https://github.com/dlab-berkeley/Python-Text-Analysis
2. D-Lab text analysis for digital humanities (python): https://github.com/dlab-berkeley/DH-Text-Analysis
3. D-Lab natural language processing part 2 python: https://github.com/dlab-berkeley/Natural-Language-Processing-Part-Two-DS-Discovery/tree/main
4. D-Lab textXD python text analysis: https://github.com/dlab-berkeley/TextXD-Python-Text-Analysis
5. D-Lab data science social justice NLP and machine learning python: https://github.com/dlab-berkeley/Data-Science-Social-Justice-2022/
6. D-Lab computational text analysis python: https://github.com/dlab-berkeley/Computational-Text-Analysis-2017
7. Arthur Turrell text analysis in python: https://aeturrell.github.io/coding-for-economists/text-intro.html

8. Data Camp NLP in Python: https://app.datacamp.com/learn/skill-tracks/natural-language-processing-in-python
9. D-Lab computation text analysis working group: https://dlabctawg.github.io/
10. Jurafsky & Martin Speech and language processing: https://web.stanford.edu/~jurafsky/slp3/
11. Introduction to Text Analysis and Social Media Research using R: https://github.com/justinchuntingho/Introduction-to-Computational-Text-Analysis-and-Social-Media-Research-using-R

Other

1. R parallel programming, docker, Google Compute Engine, Spark
    a. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures
2. Workflow and project management
    a. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures
    b. DIME Analytics coding best practices: https://dimewiki.worldbank.org/wiki/Coding_practices
    c. DIME wiki data management: https://dimewiki.worldbank.org/Data_Management
    d. Julien Reif Stata coding guide: https://julianreif.com/guide/
    e. Applied tips for applied micro RAs: https://www.dropbox.com/scl/fi/0vc9ar0gfauad8xjdanux/Applied-Tips-for-Applied-Micro-RAs.pdf?rlkey=9asv8tcqv20j67dcn6t0txpp2&e=1&dl=0
    f. ESOC predodc training Research production notes: https://github.com/esoclabprinceton/ESOC-Predoc-Training
3. Databases, SQL, Google BigQuer
    a. U of Oregon Grant McDermott data science for economists class: https://github.com/uo-ec607/lectures

Causal inference

General causal inference

1. Scott Cunningham Mixtape: https://mixtape.scunning.com/index.html
    a. https://github.com/Mixtape-Sessions
2. Chloe East Econometrics discussions: https://www.chloeneast.com/metrics-discussions.html
3. Nick Huntington-Kleim The Effect book intro to research design and causality: https://theeffectbook.net/
4. Ligon ARE 212: https://github.com/ligon/ARE212_Materials
    a. Multiple equation models, causality and correlation, identification and IV, GMM, resampling and bootstrap
5. Paul Goldsmith-Pinkham applied empirical methods phd course in R: https://github.com/paulgp/applied-methods-phd

Specific econometric methods

1. Regression discontinuity packages: https://rdpackages.github.io/replication/
2. Diff-in-diff
    a. Asjad Naqvi Diff-in-Diff guide to software packages and comparisons: https://asjadnaqvi.github.io/DiD/
        i. Stata, R, python, and Julia packages, with Stata and R code
    b. Stata HonestDiD: https://github.com/mcaceresb/stata-honestdid
    c. Pedro Sant-Anna DD checklist: https://causalinf.substack.com/p/pedros-diff-in-diff-checklist-step
3. Databases of replicable papers using different methods: https://www.dropbox.com/scl/fo/0ohwq294bknhn7hdoa7yf/ALV36_DEERdTz_XoCJPEyhc?rlkey=k9zv22k4jnco7zydfok4cvtfc&e=1&dl=0
4. Josh Blumenstock fixed effects models: http://www.jblumenstock.com/files/courses/econ174/FEModels.pdf
5. power analysis: https://stats.idre.ucla.edu/other/mult-pkg/seminars/intro-power/
6. non-standard errors: https://blogs.worldbank.org/impactevaluations/when-should-you-cluster-standard-errors-new-wisdom-econometrics-oracle, https://lost-stats.github.io/Model_Estimation/Statistical_Inference/Nonstandard_Errors/nonstandard_errors.html

Other resources

Data

1. Data in macro development STEG lectures: https://steg.cepr.org/events/virtual-course-data-macro-development
    a. Intro dev accounting, macro data, human capital data, LSMS, ag data, multinational production data, international trade data, public finances and expenditures, geographic data and night lights, cell phone data, climate and weather data, primary data collection, gender data, time use data
2. J-PAL/IPA Dataverse: https://dataverse.harvard.edu/dataverse/DFEEP?widget=dataverse%40harvard
3. J-PAL catalog of administrative datasets: https://www.povertyactionlab.org/catalog-administrative-data-sets
4. IPUMS harmonized international census data: https://www.ipums.org/projects/ipums-international
5. 3ie remote sensing data inventory: https://www.3ieimpact.org/resources/remote-sensing-inventory
6. World Bank microdata catalog: https://microdata.worldbank.org/index.php/catalog/?page=1&ps=15
7. World Bank general data catalog: https://datacatalog.worldbank.org/home
8. World Bank open data: https://data.worldbank.org/
9. World Bank climate change knowledge portal: https://climateknowledgeportal.worldbank.org/index.php/explore
10. World Bank Space2Stats: https://worldbank.github.io/DECAT_Space2Stats/readme.html
11. Ligon LSMS_Library, all LSMS-type datasets and documentation in DVC format: https://github.com/ligon/LSMS_Library
12. Harmonized LSMS-ISA (World Bank): https://github.com/lsms-worldbank/LSMS-ISA-harmonised-dataset-on-agricultural-productivity-and-welfare
13. Harmonized LSMS-HFPS: https://github.com/lsms-worldbank/HFPS_harmonization_working
14. Harmonized LSMS-ISA (EPAR): https://github.com/EvansSchoolPolicyAnalysisAndResearch/LSMS-Data-Dissemination
15. Geospatial covariate data layers (WorldPop): https://hub.worldpop.org/project/categories?id=14
    a. 100m resolution gridded data 2015-2023
16. Awesome public databases database: https://github.com/awesomedata/awesome-public-datasets?tab=readme-ov-file

Research resources

1. J-PAL resources for randomized evaluations: https://www.povertyactionlab.org/research-resources?view=toc#choose-a-view
2. J-PAL handbook on using administrative data: https://admindatahandbook.mit.edu/book/v1.1/index.html
3. Ungated research: https://ungated.research.bowdoin.edu/
    a. Access to selected papers from top econ journals
4. Connected papers see related papers by topics/papers: https://www.connectedpapers.com/

Additional resources

1. Economics conferences spreadsheet: https://docs.google.com/spreadsheets/d/1MNeXLKiwQA4MK3cZ3Hr1WWXZTReh3rKQU_yfTKnu-hg/htmlview?urp=gmail_link#
2. Econimate, videos illustrating different econ topics: https://www.youtube.com/c/econimate/playlists
3. Zotero tutorial: https://www.youtube.com/watch?v=HBynXko1wUU
4. Voxdev Where to find dev econ media: https://voxdev.org/topic/where-find-development-economics-resources-newsletters-articles-podcasts-videos
5. MIT/J-PAL Deep MicroMasters courses: https://micromasters.mit.edu/dedp
6. Arthur Turrell Research project cookie cutter code: https://github.com/aeturrell/cookiecutter-research-project