Open Data

by [Santiago Mota](https://www.linkedin.com/in/santiagomota/)

2024-07-15

Este fichero es copia de uno alojado en Github, en este [link](https://github.com/santiagomota/Open_Data) y que se actualiza periódicamente.

## Fuentes de datos abiertos y APIs

* [CRAN Task View OpenData](https://github.com/ropensci/opendata)
* [Datos en paquetes de R](http://stat.ethz.ch/R-manual/R-patched/library/datasets/html/00Index.html)
* [Kaggle datasets](https://www.kaggle.com/datasets)
* [UCI Machine Learning Repository](http://archive.ics.uci.edu/ml/)
* [DH Network](http://opendhn.dhnetwork.opendata.arcgis.com/)
* [Helsinki Open Data](http://www.hri.fi/en/)
* [Datasets de Quandl](https://www.quandl.com/search?query=)
* Amazon AWS: [este](http://aws.amazon.com/es/datasets/) y [este](https://aws.amazon.com/es/public-data-sets/)
* [Gobierno Estados Unidos](http://www.data.gov/)
* [Datos abiertos de la Unión Europea](https://data.europa.eu/es)
* [Canada Open Government Portal](https://open.canada.ca/data/en/dataset?q=education)
* [UK Open Data](https://data.gov.uk/search)
* [API de GitHub](https://developer.github.com/v3/)
* [API de Facebook](https://developers.facebook.com/docs/graph-api)
* [Blog. 100 recursos sobre Big Data y Data Science](https://www.todobi.com/mas-de-100-recursos-sobre-big-data-y/)
* [NASDAQ](https://indexes.nasdaqomx.com/Index/History/NQASPA8600AUD)
* [Google finanzas](http://www.google.com/finance/)
* [CaixaBank Research](https://www.caixabankresearch.com/es)
* [Satélite Landsat](https://aws.amazon.com/public-data-sets/landsat/)
* [OCDE](https://data.oecd.org/)
* [Open data EMT](http://opendata.emtmadrid.es/)
* [Datos abiertos del gobierno de España](http://datos.gob.es/)
* [Datos abiertos del Ayuntamiento de Madrid](http://datos.madrid.es/)
* [Datos abiertos de la Generalitat de Cataluña](http://dadesobertes.gencat.cat/es/)
* [Datos abiertos Junta de Andalucía](http://www.juntadeandalucia.es/datosabiertos/portal.html)
* [Datos abiertos de Santander](http://datos.santander.es/)
* [Natural Earth](http://www.naturalearthdata.com/)
* [Fuentes de datos espaciales (Diva-GIS)](https://diva-gis.org/)
* [Opendata del CERN](http://opendata.cern.ch/)
* [Paquete de R ‘datasets’](http://stat.ethz.ch/R-manual/R-patched/library/datasets/html/00Index.html)
* [46 museos y bibliotecas que han digitalizado todo su conocimiento y lo ofrecen gratis en internet](http://www.xataka.com/otros/46-museos-y-bibliotecas-que-han-digitalizado-todo-su-conocimiento-humano)
* [Infraestructura de Datos Espaciales de España](https://idee.es/web/idee/inicio)
* [Infraestructura de Datos Espaciales de la Comunidad de Madrid](http://www.madrid.org/cartografia/idem/html/web/index.htm)
* [Microsoft Cognitive Services](https://www.microsoft.com/cognitive-services/)
* [Google Cloud Vision API](https://cloud.google.com/vision/)
* [Análisis de 1.100 millones de trayectos de taxis y uber en NYC](https://github.com/toddwschneider/nyc-taxi-data)
* [European Data Portal](http://www.europeandataportal.eu/)
* [Propublica](https://www.propublica.org/data/)
* [NOAA. Agencia de meteo. USA.](http://www.nesdis.noaa.gov/index.html)
* [Datosclima. Base de datos meteo](http://datosclima.es/Aemet2013/DescargaDatos.html)
* [Dirección General de Tráfico (DGT)](https://sedeapl.dgt.gob.es/WEB_IEST_CONSULTA/inicio.faces)
* [National Historical Geographic Information System (NHGIS)](https://www.nhgis.org/)
* [Datos de todos los vuelos en USA entre 1987 y 2008 (datos originales)](http://stat-computing.org/dataexpo/2009/the-data.html)
* [Datos de todos los vuelos en USA entre 1987 y 2008 (otra fuente y ejemplos de uso en H2O). 120G](https://github.com/h2oai/h2o-2/wiki/Hacking-Airline-DataSet-with-H2O)
* [Open Data Inception. 1.600 portales abiertos](http://wwwhatsnew.com/2016/03/19/open-data-inception-recopilacion-de-1600-portales-de-datos-abiertos/?utm_content=buffer4e4d4&utm_medium=social&utm_source=linkedin.com&utm_campaign=buffer)
* [ImageNet database](http://www.image-net.org/)
* [API TomTom. Tráfico en ciudades](http://developer.tomtom.com/products/onlinenavigation/onlinetraffic/onlinetrafficflow)
* [Mapas de Open Street Maps](http://download.geofabrik.de/)
* [European Data Portal](https://www.europeandataportal.eu/)
* [20 Awesome Websites For Collecting Big Data](https://datafloq.com/read/20-awesome-websites-for-collecting-big-data/2737?utm_source=Datafloq%20newsletter&utm_campaign=979b1fada5-EMAIL_CAMPAIGN_2017_03_13&utm_medium=email&utm_term=0_655692fdfd-979b1fada5-90449429)
* [Climate Data Online](https://www.ncdc.noaa.gov/cdo-web/)
* [Una recopilación de APIs públicas](https://github.com/toddmotto/public-apis)
* [Una recopilación de datasets públicos](https://github.com/caesar0301/awesome-public-datasets)
* [Recopilación de datasets de BigML](https://blog.bigml.com/list-of-public-data-sources-fit-for-machine-learning/)
* [Datasets de ejemplo de IBM Watson Analytics](https://www.ibm.com/communities/analytics/watson-analytics-blog/guide-to-sample-datasets/)
* [Some datasets for teaching data science](https://simplystatistics.org/posts/2018-01-22-the-dslabs-package-provides-datasets-for-teaching-data-science/)
* [More datasets for teaching data science: The expanded dslabs package](https://simplystatistics.org/posts/2019-07-19-more-datasets-for-teaching-data-science-the-expanded-dslabs-package/)
* [El planeta Tierra en AWS](https://aws.amazon.com/es/earth/)
* [Tráfico en el Reino Unido](https://webarchive.nationalarchives.gov.uk/ukgwa/*/http://www.dft.gov.uk/traffic-counts/)
* [European Banking Authority (EBA)](https://www.eba.europa.eu/risk-and-data-analysis)
* World Bank Open Data [1](https://data.worldbank.org/) y [2](https://datacatalog.worldbank.org/)
* [Fondo Monetario Internacional](http://www.imf.org/en/data)
* [Lista de algunos datatsets dentro de paquetes de R](https://vincentarelbundock.github.io/Rdatasets/datasets.html)
* [openflights.org/](https://openflights.org/)
* [30 Amazing (And Free) Big Data And AI Public Data Sources For 2018](https://www.linkedin.com/pulse/30-amazing-free-big-data-ai-public-sources-2018-bernard-marr/?trackingId=nkTXcNLieYPDBqZuB3KIsw%3D%3D&lipi=urn%3Ali%3Apage%3Ad_flagship3_feed%3B9KuSD9KfQ6ie%2BALso3gwvw%3D%3D&licu=urn%3Ali%3Acontrol%3Ad_flagship3_feed-object)
* [Awesome Public Datasets 1](https://github.com/dipanjanS/awesome-public-datasets)
* [Awesome Public Datasets 2](https://github.com/awesomedata/awesome-public-datasets)
* [25 Open Datasets for Deep Learning Every Data Scientist Must Work With](https://www.analyticsvidhya.com/blog/2018/03/comprehensive-collection-deep-learning-datasets/)
* [NOAA Daily Global Historical Climatology Network - Kaggle dataset](https://www.kaggle.com/noaa/ghcn-d)
* [Crimen en UK](https://data.police.uk/)
* [Datos abiertos Ayuntamiento de Valencia](https://www.valencia.es/cas/ayuntamiento/gobierno-abierto)
* [Microsoft Research Open Data](https://msropendata.com/)
* [Kaggle Weekly Kernels Award Winner Announcements](https://www.kaggle.com/general/37924#post354114)
* [Open Data Renfe](http://data.renfe.com/)
* [Open Data Barometer](https://opendatabarometer.org/?_year=2017&indicator=ODB)
* [CIS. Centro de Investigaciones Sociológicas](https://www.cis.es/inicio)
* [Fivethirtyeight](https://data.fivethirtyeight.com/)
* [Reddit datasets](https://www.reddit.com/r/datasets/)
* [Data World](https://data.world/)
* [The world’s economic database](https://db.nomics.world/)
* [25 Open Datasets for Deep Learning Every Data Scientist Must Work With](https://www.analyticsvidhya.com/blog/2018/03/comprehensive-collection-deep-learning-datasets/?utm_source=linkedin.com&utm_medium=social)
* [Paquete pasra acceder al API del Instituto de Canario de Estadística](https://github.com/rOpenSpain/istacbaser)
* [Datos estadísticos DGT](https://sedeapl.dgt.gob.es/WEB_IEST_CONSULTA/)
* [Yelp Dataset](https://www.yelp.com/dataset)
* [Awesome Sentinel. Copernicus Sentinel Satellites resources](https://github.com/Fernerkundung/awesome-sentinel)
* [Datos Abiertos del Consorcio Regional de Transportes de Madrid](https://datos.crtm.es/)
* [Idealista ux&tech](https://www.idealista.com/labs/blog/)
* [Ayuntamiento de Madrid. Censo de locales, sus actividades y terrazas de hostelería y restauración](https://datos.gob.es/es/catalogo/l01280796-censo-de-locales-sus-actividades-y-terrazas-de-hosteleria-y-restauracion-historico1)
* [Agencia Tributaria. Estadísticas](https://www.agenciatributaria.es/AEAT.internet/Inicio/La_Agencia_Tributaria/Memorias_y_estadisticas_tributarias/Estadisticas/Estadisticas.shtml)
* [UK Office for National Statistics](https://www.ons.gov.uk/)
* [UK Open Geography Portal](https://geoportal.statistics.gov.uk/)
* [Open Trade Statistics](https://tradestatistics.io/)
* Facebook Neural-Code-Search-Evaluation-Dataset [dataset]](<https://github.com/facebookresearch/Neural-Code-Search-Evaluation-Dataset>) y [noticia](https://venturebeat.com/2019/10/03/facebook-open-sources-data-set-for-code-search-ai-benchmark/)
* [Ultimos datos de Open Street Map. Spain](https://download.geofabrik.de/europe/spain.html)
* [NLP Datasets](https://github.com/niderhoff/nlp-datasets/blob/master/README.md)
* [United Nations World Urbanization Prospects](https://population.un.org/wup/)
* [GHSL - Global Human Settlement Layer](https://ghslsys.jrc.ec.europa.eu/)
* [Socioeconomic Data and Applications Center (sedac)](https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/sets/browse)
* [Cómo los datos abiertos pueden ayudar en la crisis de los refugiados](https://datos.gob.es/es/blog/como-los-datos-abiertos-pueden-ayudar-en-la-crisis-de-los-refugiados?utm_source=newsletter&utm_medium=email&utm_campaign=Datos-en-tiempo-real-open-access-y-mucho-ms-en-datosgobes)
* [Worldpop - Open Spatial Demographic Data](https://www.worldpop.org/) y [Worldpop Hub](https://hub.worldpop.org/)
* [Center for Disease Control](https://wonder.cdc.gov/)
* [NASA](https://nssdc.gsfc.nasa.gov/)
* [World Economic Forum](https://www.weforum.org/publications/)
* [Universidad de Harvard](https://dataverse.harvard.edu/)
* MIT [1](http://web.mit.edu/towtank/www/vivdr/datasets.html) y [2](https://ocw.mit.edu/courses/sloan-school-of-management/15-097-prediction-machine-learning-and-statistics-spring-2012/datasets/)
* [Pew Research Center](https://www.pewresearch.org/download-datasets/)
* [Gapminder](https://www.gapminder.org/data/)
* [Reddit datasets](https://www.reddit.com/r/datasets/)
* [Center for Applied Internet Data Analysis](https://www.caida.org/data/overview/)
* [Google Books Ngram Viewe](http://storage.googleapis.com/books/ngrams/books/datasetsv2.html)
* [Free GIS Data](http://freegisdata.rtwilson.com/)
* [Google Public Data](https://www.google.com/publicdata/directory)
* [Google Datset Search](https://datasetsearch.research.google.com/)
* [Open Data Sources Database](https://anthonyhuntley.com/data-science-databases/#DataSourceDatabase)
* [Síntesis de Indicadores e Informes Macroeconómicos](https://portal.mineco.gob.es/es-es/economiayempresa/EconomiaInformesMacro/Paginas/EconomiaInformesMacro.aspx)
* [Tematicas.org Recopilación de series e índices](https://tematicas.org/)
* [Amazon product data 2014](http://jmcauley.ucsd.edu/data/amazon/)
* [Amazon product data 2018](https://nijianmo.github.io/amazon/index.html)
* [The Big Bad NLP Database](https://datasets.quantumstat.com/)
* [openaddresses](https://openaddresses.io/)
* [Awesome Geospatial](https://github.com/sacridini/Awesome-Geospatial)
* [Satellite imagery datasets containing ships](https://github.com/jasonmanesis/Satellite-Imagery-Datasets-Containing-Ships)
* [AI for Copernicus - a data repository by CALLISTO](https://github.com/Agri-Hub/Callisto-Dataset-Collection)
* [ESA WorldCover 2021. Global land cover product at 10 m for 2021 based on Sentinel-1 and 2 data](https://worldcover2021.esa.int/)
* [Data Derived from OpenStreetMap for Download](https://osmdata.openstreetmap.de/)
* [25 Satellite Maps To See Earth in New Ways](https://gisgeography.com/satellite-maps/)
* [Armed Conflict Location & Event Data Project (ACLED)](https://acleddata.com/)
* [Sentinel Hub NoR Sponsored Accounts and Data Collections](https://www.sentinel-hub.com/Network-of-Resources/)
* [rspatialdata is a collection of data sources and tutorials on visualising spatial data using R](https://rspatialdata.github.io/)
* [awesome-gee-community-datasets](https://github.com/samapriya/awesome-gee-community-datasets)
* [gee-community-catalog](https://gee-community-catalog.org/)
* [The World Bank Open Knowledge Repository](https://openknowledge.worldbank.org)
* [The Government Finance Database](https://willamette.edu/mba/research-impact/public-datasets/index.html)
* [Microsoft - A Planetary Computer for a Sustainable Future](https://planetarycomputer.microsoft.com/)
* HREA: High Resolution Electricity Access. [Universidad de Michigan](https://hrea.isr.umich.edu/index.html) y [Microsoft](https://planetarycomputer.microsoft.com/dataset/hrea#overview)
* [geoBoundaries](https://www.geoboundaries.org/)
* [FAO Map Catalog](https://data.apps.fao.org/map/catalog/srv/eng/catalog.search?id=12691#/metadata/9c35ba10-5649-41c8-bdfc-eb78e9e65654)
* [IPUMS GIS Boundary Files](https://international.ipums.org/international/gis.shtml)
* [geodata.state.gov](https://geodata.state.gov/geonetwork/srv/spa/catalog.search#/home)
* Natural Earth Vector. [Github](https://github.com/nvkelso/natural-earth-vector) y [Web](https://www.naturalearthdata.com/)
* [Geoportal Registradores](https://geoportal.registradores.org/)
* [ESA OpenSR - Robust, accountable super-resolution for Sentinel-2 and beyond](https://isp.uv.es/opensr/)
* [SEN2NAIP - Remote sensing dataset designed to support conventional and reference-based SR model training](https://huggingface.co/datasets/isp-uv-es/SEN2NAIP)
* [M3LEO: A Multi-Modal Multi-Label Earth Observation Dataset](https://huggingface.co/M3LEO)
* [SARDet-100K: large-scale multi-class SAR object detection dataset](https://eod-grss-ieee.com/dataset-detail/U1dJZE1BY1RwclAvOFFJQmlKR1Btdz09)
* [Sen2Like](https://docs.openeo.cloud/usecases/ard/sen2like/#_1-sen2like-for-rgb)
* [Google Patents Public Data](https://console.cloud.google.com/marketplace/product/google_patents_public_datasets/google-patents-public-data)
* [Digital Earth Africa (DE Africa) Map](https://www.digitalearthafrica.org/platform-resources/platform)
* [idealista18 - 2018 real estate listings in Spain. 3 cities](https://github.com/paezha/idealista18)
* [British Ordnance Survey Data Hub](https://osdatahub.os.uk/)
* [Freshwater Ecoregions of the World](https://www.worldwildlife.org/pages/freshwater-ecoregions-of-the-world--2)
* [FAO’s Global Information System on Water and Agriculture](https://www.fao.org/aquastat/en/geospatial-information/wapor)
* [Data on CO2 and Greenhouse Gas Emissions by Our World in Data](https://github.com/owid/co2-data/tree/master)
* [Open Africa dataset](https://open.africa/dataset)
* [OSM Landuse](https://osmlanduse.org/)
* [One versus One - European football statistics](https://one-versus-one.com/en)
* [StatsBomb sports data](https://statsbomb.com/what-we-do/hub/free-data/)
* [Data Kicks](https://data-kicks.com/index.php/blog/)
* [FBREF - Estadísticas e Historia del Fútbol](https://fbref.com/es/)
* [Understat](https://understat.com/)
* [Geospatial Data Catalogs](https://github.com/opengeos/geospatial-data-catalogs)
* [AWS Open Data](https://github.com/opengeos/aws-open-data)
* [AWS Open Data Geo](https://github.com/opengeos/aws-open-data-geo)
* [AWS Open Data SpatioTemporal Asset Catalog (STAC)](https://github.com/opengeos/aws-open-data-stac)
* [Google Earth Engine Catalog](https://github.com/opengeos/Earth-Engine-Catalog)
* [NASA Common Metadata Repository (CMR) SpatioTemporal Asset Catalog (STAC)](https://github.com/opengeos/aws-open-data-stac)
* [STAC Index SpatioTemporal Asset Catalog (STAC)](https://github.com/opengeos/stac-index-catalogs)
* Maxar Open Data: [Aquí](https://github.com/opengeos/maxar-open-data) y [aquí](https://radiantearth.github.io/stac-browser/#/external/maxar-opendata.s3.amazonaws.com/events/catalog.json?.language=es)
* [OpenGEOS data](https://github.com/opengeos/data)
* [Hugging Face Datasets](https://huggingface.co/datasets)
* [UC Irvine Machine Learning Repository](https://archive.ics.uci.edu/datasets)
* [Mendeley Data](https://data.mendeley.com/)
* [Nature Scientific Data](https://www.nature.com/sdata/)
* [AWS Datasets](https://registry.opendata.aws/)
* [Data.World Datasets](https://data.world/datasets/data)
* [USGS 3DEP LiDAR Point Clouds](https://registry.opendata.aws/usgs-lidar/)
* [Datos de la Eurocopa 2024](https://github.com/Jelagmil/Euro2024_data)

## Otras referencias interesantes

* [Data Science Blogs](https://github.com/rushter/data-science-blogs)
* [Chuleta general de R](https://cran.r-project.org/doc/contrib/Baggott-refcard-v2.pdf)
* [R Learning Path: From beginner to expert in R in 7 steps](http://www.kdnuggets.com/2016/03/datacamp-r-learning-path-7-steps.html)
* [Rstudio cheatsheets](https://www.rstudio.com/resources/cheatsheets/?utm_content=buffer1b56a&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)
* [R Markdown cheatsheet](https://raw.githubusercontent.com/rstudio/cheatsheets/main/rmarkdown.pdf)
* [R Markdown referencia](https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf)
* [A dive into R Markdown](http://cfss.uchicago.edu/program_rmarkdown.html)
* [Información de Rmarkdown en R Studio](http://rmarkdown.rstudio.com/)
* [Template para documentos científicos con Rmarkdown](http://www.petrkeil.com/?p=2401)
* [Formatos a medida para R Markdown](http://www.r-bloggers.com/r-markdown-custom-formats/)
* [blogdown: Creating Websites with R Markdown](https://bookdown.org/yihui/blogdown/)
* [Utilizando Sweave y Knitr](https://support.posit.co/hc/en-us/articles/200552056-Using-Sweave-and-knitr)
* [Pandoc User’s Guide](http://pandoc.org/MANUAL.html#templates)
* [Soporte técnico de RStudio](https://support.posit.co/hc/en-us)
* [100 Active Blogs on Analytics, Big Data, Data Mining, Data Science, Machine Learning](http://www.kdnuggets.com/2016/03/100-active-blogs-analytics-big-data-science-machine-learning.html#.VvqjkSV5Tio.linkedin)
* [Plataforma H2O](https://github.com/h2oai)
* [Computer vision](https://github.com/kjw0612/awesome-deep-vision)
* Pautas para dar formato al código programando en R: [Google](https://google.github.io/styleguide/Rguide.xml), [Hadley Wickham (RStudio)](http://adv-r.had.co.nz/Style.html) y [Coding Club](https://ourcodingclub.github.io/2017/04/25/etiquette.html#syntax)
* [The State of Naming Conventions in R](https://journal.r-project.org/archive/2012-2/RJournal_2012-2_Baaaath.pdf)
* [R Code – Best practices](https://www.r-bloggers.com/r-code-best-practices/)
* [R Coding Style Guide](https://irudnyts.github.io//r-coding-style-guide/)
* [Naming files](https://speakerd.s3.amazonaws.com/presentations/5e4b07f0d9a94f8e9a29b902bad6ed0b/naming-slides.pdf)
* [Documentacion de R](https://www.rdocumentation.org/)
* [Chuleta de expresiones regulares](https://github.com/rstudio/cheatsheets/blob/main/regex.pdf)
* [Regular Expressions Every R programmer Should Know](https://www.r-bloggers.com/regular-expressions-every-r-programmer-should-know/)
* [Regular Expression Language - Quick Reference](https://docs.microsoft.com/en-us/dotnet/standard/base-types/regular-expression-language-quick-reference)
* [Dealing with Regular Expressions](http://uc-r.github.io/regex)
* [Writing an R package from scratch](https://hilaryparker.com/2014/04/29/writing-an-r-package-from-scratch/)
* [16 Cursos](https://www.analyticsvidhya.com/blog/2016/10/16-new-must-watch-tutorials-courses-on-machine-learning/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+AnalyticsVidhya+%28Analytics+Vidhya%29)
* [Galerias de graficos](http://www.r-graph-gallery.com/)
* [Curso Caltech. Learning from data](https://work.caltech.edu/telecourse.html)
* [Usar git](https://try.github.io/levels/1/challenges/1)
* [Blogs con github](http://jmcglone.com/guides/github-pages/) y [Blogs con github y RStudio](http://andysouth.github.io/blog-setup/)
* [Ejemplos de Shiny](http://zevross.com/blog/2016/04/19/r-powered-web-applications-with-shiny-a-tutorial-and-cheat-sheet-with-40-example-apps/)
* [UK government using R to modernize reporting of official statistics](https://www.r-bloggers.com/uk-government-using-r-to-modernize-reporting-of-official-statistics/)
* [Ggplot](http://socviz.co/)
* Sistemas de Coordenadas. [Aqui](https://rspatial.org/spatial/rst/6-crs.html) y [aqui](https://www.nceas.ucsb.edu/~frazier/RSpatialGuides/OverviewCoordinateReferenceSystems.pdf)
* [Codificación de caracteres](https://www.joelonsoftware.com/2003/10/08/the-absolute-minimum-every-software-developer-absolutely-positively-must-know-about-unicode-and-character-sets-no-excuses/)
* [Tutorials for learning R](https://www.r-bloggers.com/how-to-learn-r-2/)
* [Awesome R](https://github.com/qinwf/awesome-R)
* [R Data Science Tutorials](https://github.com/ujjwalkarn/DataScienceR)
* [useR! Machine Learning Tutorial](https://github.com/ledell/useR-machine-learning-tutorial)
* [Computerworld - Paquetes de R interesantes](https://www.computerworld.com/article/1375862/great-r-packages-for-data-import-wrangling-visualization.html)
* [Otra lista de recursos variados en Github](https://github.com/Shujian2015/FreeML)
* [Tipos de licencias de software](https://choosealicense.com/licenses/)
* [Glosario de Machine Learning de Google](https://developers.google.com/machine-learning/glossary/)
* [Google Rules of Machine Learning: Best Practices for ML Engineering](http://martin.zinkevich.org/rules_of_ml/rules_of_ml.pdf)
* Statistical Learning de Stanford with R [Curso](https://online.stanford.edu/courses/sohs-ystatslearning-statistical-learning-r), [Libro](https://hastie.su.domains/ElemStatLearn/), [Código](https://github.com/khanhnamle1994/statistical-learning) y [Transparencias](https://github.com/khanhnamle1994/statistical-learning/tree/master/Lecture-Slides)
* [An Introduction to Statistical Learning - Web R & Python](https://www.statlearning.com/)
* [100 Free Tutorials for Learning R](https://www.listendata.com/p/r-programming-tutorials.html)
* [Google’s best practices in machine learning](https://developers.google.com/machine-learning/guides/rules-of-ml/)
* [Web Scraping TripAdvisor, Text Mining and Sentiment Analysis for Hotel Reviews](https://towardsdatascience.com/scraping-tripadvisor-text-mining-and-sentiment-analysis-for-hotel-reviews-cc4e20aef333)
* [Common Probability Distributions: The Data Scientist’s Crib Sheet](https://blog.cloudera.com/blog/2015/12/common-probability-distributions-the-data-scientists-crib-sheet/?utm_content=buffer49e9f&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer)
* [RDocumentation](https://www.rdocumentation.org/)
* [ArcGIS to R spatial cheat sheet](http://www.seascapemodels.org/data/ArcGIS_to_R_Spatial_CheatSheet.pdf)
* [Hand-curated, high quality resources for doing data journalism with R](https://rddj.info/)
* [RMarkdown Driven Development (RmdDD)](https://emilyriederer.netlify.app/post/rmarkdown-driven-development/)
* [htmlwidgets for R - gallery](http://gallery.htmlwidgets.org/)
* [Free R Reading Material](https://committedtotape.shinyapps.io/freeR/)
* [The Chartmaker Directory](chartmaker.visualisingdata.com)
* [Kaggle Winning Solutions](http://kagglesolutions.com/)
* [Simplifying the ROC and AUC metrics](https://towardsdatascience.com/understanding-the-roc-and-auc-curves-a05b68550b69)
* [Feature Engineering for Machine Learning](https://trainindata.medium.com/feature-engineering-for-machine-learning-a-comprehensive-overview-a7ad04c896f8)
* [Awesome Data Science](https://github.com/academic/awesome-datascience)
* [Data Scientist Roadmap](https://github.com/MrMimic/data-scientist-roadmap)
* [Data Science Collected Resources](https://github.com/tirthajyoti/Data-science-best-resources)
* [Data Science Cheatsheets](https://github.com/FavioVazquez/ds-cheatsheets)
* [Data Science Resources](https://github.com/jonathan-bower/DataScienceResources)
* [R for Water Resources Data Science](https://www.r4wrds.com/)
* [GIS and mapping](https://nowosad.github.io/SIGR2021/workshop1/workshop1_jn.html#1)
* [R package primer](https://kbroman.org/pkg_primer/)
* [Using Geospatial Data in R](https://www.mzes.uni-mannheim.de/socialsciencedatalab/article/geospatial-data/)
* [Cómo crear una API en Python](https://anderfernandez.com/blog/como-crear-api-en-python/)
* [Cursos para aprender más sobre R](https://datos.gob.es/es/noticia/cursos-para-aprender-mas-sobre-r)
* [Estadística con R](https://www.cienciadedatos.net/estadistica-con-r.html)
* [Practical Introduction to Web Scraping in R](https://blog.rsquaredacademy.com/web-scraping/)
* [R for Water Resources Data Science](https://www.r4wrds.com/)
* [A ggplot2 Tutorial for Beautiful Plotting in R](https://cedricscherer.netlify.app/2019/08/05/a-ggplot2-tutorial-for-beautiful-plotting-in-r/)
* [Bivariate Choropleth Maps: A How-to Guide](https://www.joshuastevens.net/cartography/make-a-bivariate-choropleth-map/)
* [Tipos de licencias open data (minicurso de data.europa.edu)](https://data.europa.eu/en/academy/open-data-licensing)
* Legalidad Web sraping: [Is Web Scraping Legal? : The Definitive Guide (2024 update)](https://prowebscraper.com/blog/is-web-scraping-legal/) y [Web Scraping: ¿legal o ilegal?](https://ecija.com/web-scraping-legal-ilegal/)
* [Information is Beautiful Awards](https://www.informationisbeautifulawards.com/)
* [Financial-Times / chart-doctor](https://github.com/Financial-Times/chart-doctor/tree/main/visual-vocabulary)
* [The Data Visualisation Catalogue](https://datavizcatalogue.com/)
* [From Data to Viz](https://www.data-to-viz.com/)
* [EUMETSAT science studies](https://www.eumetsat.int/science-studies)
* [overpass turbo - Herremaienta de filtrado para OSM](https://overpass-turbo.eu/)
* [Remote Sensing for OSINT](https://bellingcat.github.io/RS4OSINT/)
* [Investigative Journalism with Satellite Images](https://bourgoing.com/en/linvestigation-par-satellite/)
* [CAMIS - A PHUSE DVOST Working Group](https://psiaims.github.io/CAMIS/). The repository below provides examples of statistical methodology in different software and languages, along with a comparison of the results obtained and description of any discrepancies.
* [Microsoft Presidio - Data Protection and De-identification SDK](https://microsoft.github.io/presidio/)
* [GIS formats](https://atlas.co/formats/)
* [Data Viz Catalogue](https://graphica.app/catalogue)
* [Dataviz Project](https://datavizproject.com/)
* [Periodic Table Of Visualization Methods](https://www.visual-literacy.org/periodic_table/periodic_table.html)
* [The TimeViz Browser 2.0](https://browser.timeviz.net/)
* [The R Graph Gallery](https://r-graph-gallery.com/)
* [HOT - Drone Tasking Manager](https://github.com/hotosm/Drone-TM)
* [Remote sensing image retrieval](https://github.com/IBM/remote-sensing-image-retrieval)
* [BigEarthNet A Large-Scale Sentinel Benchmark Archive](https://bigearth.net/)
* Global Fishing Watch. AI and satellite imagery to reveal the expanding footprint of human activity at sea. [Post](https://globalfishingwatch.org/press-release/new-research-harnesses-ai-and-satellite-imagery-to-reveal-the-expanding-footprint-of-human-activity-at-sea/?utm_source=GFW+subscribers&utm_campaign=9363c93195-EMAIL_CAMPAIGN_JAN_2024_CURRENT_ENGLISH&utm_medium=email&utm_term=0_-9363c93195-%5BLIST_EMAIL_ID%5D). [Github](https://github.com/GlobalFishingWatch/paper-industrial-activity/tree/main). [Train data](https://figshare.com/articles/journal_contribution/Satellite_mapping_reveals_extensive_industrial_activity_at_sea_-_training_data/24309469). [Analysis data](https://figshare.com/articles/journal_contribution/Satellite_mapping_reveals_extensive_industrial_activity_at_sea_-_analysis_data/24309475) and [Vessel detection from Sentinel-1 SAR](https://globalfishingwatch.org/data-download/datasets/public-sar-vessel-detections:v20231026)
* [rseek.org - rstats search engine](https://rseek.org/)

## Libros

* [R Markdown Cookbook](https://bookdown.org/yihui/rmarkdown-cookbook/)
* [R Markdown: The Definitive Guide](https://bookdown.org/yihui/rmarkdown/)
* [Todos los libros en bookdown](https://bookdown.org/home/archive/)
* [R intro](https://cran.r-project.org/doc/manuals/R-intro.pdf)
* [R for everyone](https://www.jaredlander.com/r-for-everyone/)
* [R in action](https://www.manning.com/books/r-in-action-second-edition)
* [R Programming for Data Science. Roger D. Peng.](https://leanpub.com/rprogramming)
* [R para principiantes](https://cran.r-project.org/doc/contrib/rdebuts_es.pdf)
* [Introducción a R](https://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf)
* [R para profesionales de los datos: una introducción](https://www.datanalytics.com/libro_r/)
* [10 Free Must-Read Books for Machine Learning and Data Science](https://www.kdnuggets.com/2017/04/10-free-must-read-books-machine-learning-data-science.html?utm_content=bufferc386f&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)
* Introduction to Data Science [Libro](https://rafalab.github.io/dsbook/) y [Código](https://github.com/rafalab/dsbook)
* Fundamentals of Data Visualization [Libro](https://clauswilke.com/dataviz/) y [Código](https://github.com/clauswilke/dataviz)
* Data Science Live Book [Libro](https://livebook.datascienceheroes.com/) y [Código](https://github.com/pablo14/data-science-live-book)
* R for Statistical Learning [Libro](https://daviddalpiaz.github.io/r4sl/) y [Código](https://github.com/daviddalpiaz/r4sl)
* Applied Statistics with R [Libro](https://daviddalpiaz.github.io/appliedstats/) y [Código](https://github.com/daviddalpiaz/appliedstats)
* Geocomputation with R [Libro](https://geocompr.robinlovelace.net/) y [Código](https://github.com/Robinlovelace/geocompr/)
* [Handling Strings with R](http://www.gastonsanchez.com/r4strings/)
* [Text Mining with R](https://www.tidytextmining.com/)
* [Efficient R programming](https://csgillespie.github.io/efficientR/)
* [BBC Visual and Data Journalism cookbook for R graphics](https://bbc.github.io/rcookbook/)
* [Databases using R by RStudio](https://db.rstudio.com/getting-started/)
* [Interpretable Machine Learning](https://christophm.github.io/interpretable-ml-book/)
* [Forecasting: Principles and Practice](https://otexts.com/fpp3/)
* [Bioinformática Estadística. Análisis estadístico de datos Ómicos](https://www.uv.es/ayala/docencia/tami/tami13.pdf)
* [Estadística básica](https://www.uv.es/ayala/docencia/nmr/nmr13.pdf)
* [Probabilidad básica](https://www.uv.es/ayala/docencia/probabilidad/prob.pdf)
* [Estadística básica](https://www.uv.es/ayala/docencia/nmr/nmr13.pdf)
* [An Introduction to Spatial Data Analysis and Visualisation in R](https://www.spatialanalysisonline.com/An%20Introduction%20to%20Spatial%20Data%20Analysis%20in%20R.pdf)
* [Exploratory Data Analysis with R - Roger D. Peng](https://bookdown.org/rdpeng/exdata/)
* [What They Forgot to Teach You About R](https://whattheyforgot.org/)
* Mastering Apache Spark with R [Libro](https://therinspark.com/intro.html) y [Código](https://github.com/r-spark/the-r-in-spark)
* [Hands-On Machine Learning with R](https://bradleyboehmke.github.io/HOML/)
* [Hands-On Programming with R](https://rstudio-education.github.io/hopr/)
* [The 20 Best Data Science Books Available online in 2020](https://www.ubuntupit.com/best-data-science-books-available-online/)
* [Creating APIs in R with Plumber](https://www.rplumber.io/docs/index.html)
* [Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny](http://www.paulamoraga.com/book-geospatial/)
* [Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA](https://becarioprecario.bitbucket.io/spde-gitbook/)
* [Estilometría, análisis de textos en R para filólogos](http://www.aic.uva.es/cuentapalabras/presentacion.html)
* [Engineering Production-Grade Shiny Apps](https://engineering-shiny.org/)
* [Bayesian inference with INLA](https://becarioprecario.bitbucket.io/inla-gitbook/index.html)
* [Happy Git and GitHub for the useR](https://happygitwithr.com/)
* R for Data Science. [Inglés](https://r4ds.hadley.nz/) y [Castellano](https://es.r4ds.hadley.nz/)
* [Data Visualization with R](https://rkabacoff.github.io/datavis/)
* [Big Book of R](https://www.bigbookofr.com/index.html)
* [JavaScript for R](https://book.javascript-for-r.com/)
* [Hands-On Data Visualization](https://handsondataviz.org/)
* [Fundamentals of Data Visualization](https://github.com/clauswilke/dataviz)
* [Supervised Machine Learning for Text Analysis in R](https://smltar.com/)
* [Modern R with the tidyverse](https://b-rodrigues.github.io/modern_R/)
* [Spatial Microsimulation with R](https://spatial-microsim-book.robinlovelace.net/index.html)
* [R Advanced Spatial Lessons](https://bbest.github.io/R-adv-spatial-lessons/)
* [R Packages](https://r-pkgs.org/)
* [R Graphics Cookbook](https://r-graphics.org/index.html)
* [Advanced R](https://adv-r.hadley.nz/index.html)
* [rstudio4edu](https://rstudio4edu.github.io/rstudio4edu-book/)
* [R for Health Data Science](https://argoshare.is.ed.ac.uk/healthyr_book/)
* [Linear Algebra for Data Science](https://shainarace.github.io/LinearAlgebra/index.html)
* [Aprendizaje Estadístico con R](https://rubenfcasal.github.io/aprendizaje_estadistico/index.html)
* [Simulación Estadística con R](https://rubenfcasal.github.io/simbook/)
* [R para profesionales de los datos: una introducción](https://datanalytics.com/libro_r/)
* [Aprendiendo R sin morir en el intento](https://aprendiendo-r-intro.netlify.app/)
* [Interpretable Machine Learning](https://christophm.github.io/interpretable-ml-book/)
* [The caret Package](http://topepo.github.io/caret/index.html)
* [Elegant and informative maps with tmap](https://r-tmap.github.io/tmap-book/)
* [Spatial Modelling for Data Scientists](https://gdsl-ul.github.io/san/)
* [Command Line Basics for R Users](https://bash-intro.rsquaredacademy.com/)
* [Engineering Production-Grade Shiny Apps](https://engineering-shiny.org/)
* [Outstanding User Interfaces with Shiny](https://unleash-shiny.rinterface.com/)
* [The Shiny AWS Book](https://business-science.github.io/shiny-production-with-aws-book/)
* [Twitter for Scientists](https://t4scientists.com/)
* [Spatial Data Science](https://keen-swartz-3146c4.netlify.app/)
* [Handbook of Graphs and Networks in People Analytics](https://ona-book.org/)
* [Officeverse R & Office](https://ardata-fr.github.io/officeverse/index.html)
* [Spatial Analysis With R](http://gis.humboldt.edu/OLM/r/Spatial%20Analysis%20With%20R.pdf)
* [R Programming for Data Science](https://www.cs.upc.edu/~robert/teaching/estadistica/rprogramming.pdf)
* [The R Book](https://www.cs.upc.edu/~robert/teaching/estadistica/TheRBook.pdf)
* [R Notes for Professionals](https://books.goalkicker.com/RBook/)
* [An Introduction to Spatial Data Analysis and Visualisation in R](https://www.spatialanalysisonline.com/An%20Introduction%20to%20Spatial%20Data%20Analysis%20in%20R.pdf)
* [Think Bayes 2e](https://github.com/AllenDowney/ThinkBayes2)
* [Quantitative Politics with R](http://qpolr.com/)
* [Data Science in Education Using R](https://datascienceineducation.com/)
* [Data Skills for Reproducible Science](https://psyteachr.github.io/msc-data-skills/)
* [Handbook of Regression Modeling in People Analytics](https://peopleanalytics-regression-book.org/)
* [Handbook of Graphs and Networks in People Analytics With Examples in R and Python](https://ona-book.org/)
* [Tidy Finance with R](https://tidy-finance.org/)
* [Predictive Soil Mapping with R](https://soilmapper.org/)
* [Deep Learning](https://srdas.github.io/DLBook/)
* [Technical Foundations of Informatics](https://info201.github.io/)
* [YaRrr! The Pirate’s Guide to R](https://bookdown.org/ndphillips/YaRrr/)
* [An R companion to Statistics: data analysis and modelling](https://mspeekenbrink.github.io/sdam-r-companion/index.html)
* [Learning Statistics with R](https://learningstatisticswithr.com/)
* [Happy Git and GitHub for the useR](https://happygitwithr.com/index.html)
* [Introduction to Econometrics with R](https://www.econometrics-with-r.org/)
* [R4JournalismBook](https://smach.github.io/R4JournalismBook/)
* [Libro Vivo de Ciencia de Datos](https://librovivodecienciadedatos.ai/)
* [Introduction to Probability for Data Science](https://probability4datascience.com/index.html)
* [Spatial Data Science with applications in R](https://r-spatial.org/book/)
* [R for Data Analysis](https://trevorfrench.github.io/R-for-Data-Analysis/)
* [Data Analysis and Prediction Algorithms with R](http://rafalab.dfci.harvard.edu/dsbook/)
* [Análisis de datos y algoritmos de predicción con R](http://rafalab.dfci.harvard.edu/dslibro/)
* [The Art of Data Science](https://bookdown.org/rdpeng/artofdatascience/)
* [R for data science: tidyverse and beyond](https://bookdown.org/Maxine/r4ds/)
* [Statistical Inference via Data Science](https://moderndive.com/index.html)
* [Introduction to urban accessibility: a practical guide in R](https://github.com/ipeaGIT/intro_access_book)
* [Tidy Finance](https://www.tidy-finance.org/)
* [Deep Learning and Scientific Computing with R torch](https://skeydan.github.io/Deep-Learning-and-Scientific-Computing-with-R-torch/)
* [Econometrics with the Tidyverse](https://colleen.quarto.pub/the-tidy-econometrics-workbook/)
* [Building reproducible analytical pipelines with R](https://raps-with-r.dev/)
* [Satellite Image Time Series Analysis on Earth Observation Data Cubes](https://e-sensing.github.io/sitsbook/)
* [Spatial Statistics for Data Science: Theory and Practice with R](https://www.paulamoraga.com/book-spatial/index.html)
* [Geographic Data Science with R: Visualizing and Analyzing Environmental Change](https://bookdown.org/mcwimberly/gdswr-book/)
* [Apache Arrow R Cookbook](https://arrow.apache.org/cookbook/r/)
* [An Introduction to R and Python For Data Analysis: A Side By Side Approach](https://randpythonbook.netlify.app/)
* [Using the flextable R package](https://ardata-fr.github.io/flextable-book/index.html)
* [Spatial sampling with R](https://dickbrus.github.io/SpatialSamplingwithR/)
* [Interactive web-based data visualization with R, plotly, and shiny](https://plotly-r.com/)
* [Modern Data Science with R](https://mdsr-book.github.io/mdsr3e/)
* [Quarto for Scientists](https://qmd4sci.njtierney.com/)
* [Introduction to Environmental Data Science](https://bookdown.org/igisc/EnvDataSci/)
* [Manual de R](https://fhernanb.github.io/Manual-de-R/)
* [Modern R with the tidyverse](https://modern-rstats.eu/)
* [Tidy Modeling with R](https://www.tmwr.org/)
* [Data Analysis in Medicine and Health using R](https://bookdown.org/drki_musa/dataanalysis/)
* [Crime by the Numbers: A Criminologist’s Guide to R](https://crimebythenumbers.com/)
* [R Tutorial: Mastering Data Analysis and More](https://pyoflife.com/r-programming-made-easy/)
* [Data transformation with R](https://pyoflife.com/data-transformation-with-r-pdf/)
* [Data Visualization and Exploration with R](https://pyoflife.com/data-visualization-and-exploration-with-r-pdf/)
* [Introduction to Probability and Statistics Using R](https://pyoflife.com/introduction-to-probability-and-statistics-using-r/)
* [Fundamentos de ciencia de datos con R](https://cdr-book.github.io/)
* [Financial Risk Modelling and Portfolio Optimization with R](https://pyoflife.com/financial-risk-modelling-and-portfolio-optimization-with-r/)
* [R in Finance: Introduction to R and Its Applications in Finance](https://pyoflife.com/r-in-finance-introduction-to-r-and-its-applications-in-finance/)
* [Deep R Programming](https://deepr.gagolewski.com/)
* [Beyond Multiple Linear Regression](https://bookdown.org/roback/bookdown-BeyondMLR/)
* [Analyzing US Census Data: Methods, Maps, and Models in R](https://walker-data.com/census-r/index.html)
* [Psychometrics in Exercises using R and RStudio](https://bookdown.org/annabrown/psychometricsR/)
* [Introduction to NFL Analytics with R](https://bradcongelio.com/nfl-analytics-with-r-book/)
* [Feature Engineering and Selection: A Practical Approach for Predictive Models](https://bookdown.org/max/FES/)
* [Applied Machine Learning Using mlr3 in R](https://mlr3book.mlr-org.com/)
* [Behavior Analysis with Machine Learning Using R](https://enriquegit.github.io/behavior-free/index.html)
* [Data Analytics: A Small Data Approach](https://dataanalyticsbook.info/)
* [R for Social Network Analysis](https://schochastics.github.io/R4SNA/)
* [Cookbook Polars for R](https://ddotta.github.io/cookbook-rpolars/)
* [DIY API with Make and {plumber}](https://www.andrewheiss.com/blog/2024/01/12/diy-api-plumber-quarto-ojs/_book/)
* [ggplot2: Elegant Graphics for Data Analysis (3e)](https://ggplot2-book.org/)
* [R for the Rest of Us: A Statistics-Free Introduction](https://book.rfortherestofus.com/)
* [Un Recorrido por los Métodos Cuantitativos en Ciencias Sociales a bordo de R](https://estadisticacienciassocialesr.rbind.io/)
* [What They Forgot to Teach You About R](https://rstats.wtf/)
* [Tidyverse Skills for Data Science](https://jhudatascience.org/tidyversecourse/)
* [A Business Analyst’s Introduction to Business Analytics](https://www.causact.com/#welcome)
* [Telling Stories with Data](https://tellingstorieswithdata.com/)
* [The tidyverse style guide](https://style.tidyverse.org/)
* [Empirical Research in Accounting: Tools and Methods](https://iangow.github.io/far_book/)
* [Data Visualization with R](https://datavizs23.classes.andrewheiss.com/example/)
* [Introduction to MySQL with R](https://programminghistorian.org/en/lessons/getting-started-with-mysql-using-r)
* [Physical Oceanography Distributed Active Archive Center totorials](https://podaac.github.io/tutorials/)
* [The SAR Handbook](https://gis1.servirglobal.net/TrainingMaterials/SAR/SARHB_FullRes.pdf)
* [Advanced International Trade in R](https://pacha.dev/advanced-international-trade-in-r/)
* [The Orange Book of Machine Learning](https://github.com/Carl-McBride-Ellis/TOBoML)
* [Data Management in Large-Scale Education Research](https://datamgmtinedresearch.com/)