

Open Data

by [Santiago Mota](#)

2024-07-14

Este fichero es copia de uno alojado en Github, en este [link](#) y que se actualiza periódicamente.

Fuentes de datos abiertos y APIs

- [CRAN Task View OpenData](#)
- [Datos en paquetes de R](#)
- [Kaggle datasets](#)
- [UCI Machine Learning Repository](#)
- [DH Network](#)
- [Helsinki Open Data](#)
- [Datasets de Quandl](#)
- [Amazon AWS: este y este](#)
- [Gobierno Estados Unidos](#)
- [Datos abiertos de la Unión Europea](#)
- [Canada Open Government Portal](#)
- [UK Open Data](#)
- [API de GitHub](#)
- [API de Facebook](#)
- [Blog. 100 recursos sobre Big Data y Data Science](#)
- [NASDAQ](#)
- [Google finanzas](#)
- [CaixaBank Research](#)
- [Satélite Landsat](#)
- [OCDE](#)
- [Open data EMT](#)
- [Datos abiertos del gobierno de España](#)
- [Datos abiertos del Ayuntamiento de Madrid](#)
- [Datos abiertos de la Generalitat de Cataluña](#)
- [Datos abiertos Junta de Andalucía](#)
- [Datos abiertos de Santander](#)
- [Natural Earth](#)
- [Fuentes de datos espaciales \(Diva-GIS\)](#)
- [Opendata del CERN](#)
- [Paquete de R ‘datasets’](#)
- [46 museos y bibliotecas que han digitalizado todo su conocimiento y lo ofrecen gratis en internet](#)
- [Infraestructura de Datos Espaciales de España](#)
- [Infraestructura de Datos Espaciales de la Comunidad de Madrid](#)
- [Microsoft Cognitive Services](#)
- [Google Cloud Vision API](#)
- [Análisis de 1.100 millones de trayectos de taxis y uber en NYC](#)
- [European Data Portal](#)
- [Propublica](#)
- [NOAA. Agencia de meteo. USA.](#)

- Datosclima. Base de datos meteo
- Dirección General de Tráfico (DGT)
- National Historical Geographic Information System (NHGIS)
- Datos de todos los vuelos en USA entre 1987 y 2008 (datos originales)
- Datos de todos los vuelos en USA entre 1987 y 2008 (otra fuente y ejemplos de uso en H2O). 120G
- Open Data Inception. 1.600 portales abiertos
- ImageNet database
- API TomTom. Tráfico en ciudades
- Mapas de Open Street Maps
- European Data Portal
- 20 Awesome Websites For Collecting Big Data
- Climate Data Online
- Una recopilación de APIs públicas
- Una recopilación de datasets públicos
- Recopilación de datasets de BigML
- Datasets de ejemplo de IBM Watson Analytics
- Some datasets for teaching data science
- More datasets for teaching data science: The expanded dslabs package
- El planeta Tierra en AWS
- Tráfico en el Reino Unido
- European Banking Authority (EBA)
- World Bank Open Data 1 y 2
- Fondo Monetario Internacional
- Lista de algunos datasets dentro de paquetes de R
- openflights.org/
- 30 Amazing (And Free) Big Data And AI Public Data Sources For 2018
- Awesome Public Datasets 1
- Awesome Public Datasets 2
- 25 Open Datasets for Deep Learning Every Data Scientist Must Work With
- NOAA Daily Global Historical Climatology Network - Kaggle dataset
- Crimen en UK
- Datos abiertos Ayuntamiento de Valencia
- Microsoft Research Open Data
- Kaggle Weekly Kernels Award Winner Announcements
- Open Data Renfe
- Open Data Barometer
- CIS. Centro de Investigaciones Sociológicas
- Fivethirtyeight
- Reddit datasets
- Data World
- The world's economic database
- 25 Open Datasets for Deep Learning Every Data Scientist Must Work With
- Paquete para acceder al API del Instituto de Canarias de Estadística
- Datos estadísticos DGT
- Yelp Dataset
- Awesome Sentinel. Copernicus Sentinel Satellites resources
- Datos Abiertos del Consorcio Regional de Transportes de Madrid
- Idealista ux&tech
- Ayuntamiento de Madrid. Censo de locales, sus actividades y terrazas de hostelería y restauración
- Agencia Tributaria. Estadísticas
- UK Office for National Statistics
- UK Open Geography Portal
- Open Trade Statistics
- Facebook Neural-Code-Search-Evaluation-Dataset [dataset]](<https://github.com/facebookresearch/>

- [Neural-Code-Search-Evaluation-Dataset](#)) y [noticia](#)
- [Ultimos datos de Open Street Map. Spain](#)
- [NLP Datasets](#)
- [United Nations World Urbanization Prospects](#)
- [GHSL - Global Human Settlement Layer](#)
- [Socioeconomic Data and Applications Center \(sedac\)](#)
- [Cómo los datos abiertos pueden ayudar en la crisis de los refugiados](#)
- [Worldpop - Open Spatial Demographic Data](#) y [Worldpop Hub](#)
- [Center for Disease Control](#)
- [NASA](#)
- [World Economic Forum](#)
- [Universidad de Harvard](#)
- [MIT 1 y 2](#)
- [Pew Research Center](#)
- [Gapminder](#)
- [Reddit datasets](#)
- [Center for Applied Internet Data Analysis](#)
- [Google Books Ngram Viewe](#)
- [Free GIS Data](#)
- [Google Public Data](#)
- [Google Datset Search](#)
- [Open Data Sources Database](#)
- [Síntesis de Indicadores e Informes Macroeconómicos](#)
- [Tematicas.org Recopilación de series e índices](#)
- [Amazon product data 2014](#)
- [Amazon product data 2018](#)
- [The Big Bad NLP Database](#)
- [openaddresses](#)
- [Awesome Geospatial](#)
- [Satellite imagery datasets containing ships](#)
- [AI for Copernicus - a data repository by CALLISTO](#)
- [ESA WorldCover 2021. Global land cover product at 10 m for 2021 based on Sentinel-1 and 2 data](#)
- [Data Derived from OpenStreetMap for Download](#)
- [25 Satellite Maps To See Earth in New Ways](#)
- [Armed Conflict Location & Event Data Project \(ACLED\)](#)
- [Sentinel Hub NoR Sponsored Accounts and Data Collections](#)
- [rsatialdata is a collection of data sources and tutorials on visualising spatial data using R](#)
- [awesome-gee-community-datasets](#)
- [gee-community-catalog](#)
- [The World Bank Open Knowledge Repository](#)
- [The Government Finance Database](#)
- [Microsoft - A Planetary Computer for a Sustainable Future](#)
- [HREA: High Resolution Electricity Access. Universidad de Michigan y Microsoft](#)
- [geoBoundaries](#)
- [FAO Map Catalog](#)
- [IPUMS GIS Boundary Files](#)
- [geodata.state.gov](#)
- [Natural Earth Vector. Github y Web](#)
- [Geoportal Registradores](#)
- [ESA OpenSR - Robust, accountable super-resolution for Sentinel-2 and beyond](#)
- [SEN2NAIP - Remote sensing dataset designed to support conventional and reference-based SR model training](#)
- [M3LEO: A Multi-Modal Multi-Label Earth Observation Dataset](#)
- [SARDet-100K: large-scale multi-class SAR object detection dataset](#)

- [Sen2Like](#)
- [Google Patents Public Data](#)
- [Digital Earth Africa \(DE Africa\) Map](#)
- [idealista18](#) - 2018 real estate listings in Spain. 3 cities
- [British Ordnance Survey Data Hub](#)
- [Freshwater Ecoregions of the World](#)
- [FAO's Global Information System on Water and Agriculture](#)
- [Data on CO2 and Greenhouse Gas Emissions by Our World in Data](#)
- [Open Africa dataset](#)
- [OSM Landuse](#)
- [One versus One](#) - European football statistics
- [StatsBomb sports data](#)
- [Data Kicks](#)
- [FBREF](#) - Estadísticas e Historia del Fútbol
- [Understat](#)
- [Geospatial Data Catalogs](#)
- [AWS Open Data](#)
- [AWS Open Data Geo](#)
- [AWS Open Data SpatioTemporal Asset Catalog \(STAC\)](#)
- [Google Earth Engine Catalog](#)
- [NASA Common Metadata Repository \(CMR\) SpatioTemporal Asset Catalog \(STAC\)](#)
- [STAC Index SpatioTemporal Asset Catalog \(STAC\)](#)
- [Maxar Open Data](#)
- [OpenGEOS data](#)
- [Hugging Face Datasets](#)
- [UC Irvine Machine Learning Repository](#)
- [Mendeley Data](#)
- [Nature Scientific Data](#)
- [AWS Datasets](#)
- [Data.World Datasets](#)

Otras referencias interesantes

- [Data Science Blogs](#)
- [Chuleta general de R](#)
- [R Learning Path: From beginner to expert in R in 7 steps](#)
- [Rstudio cheatsheets](#)
- [R Markdown cheatsheet](#)
- [R Markdown referencia](#)
- [A dive into R Markdown](#)
- [Información de Rmarkdown en R Studio](#)
- [Template para documentos científicos con Rmarkdown](#)
- [Formatos a medida para R Markdown](#)
- [blogdown: Creating Websites with R Markdown](#)
- [Utilizando Sweave y Knitr](#)
- [Pandoc User's Guide](#)
- [Soporte técnico de RStudio](#)
- [100 Active Blogs on Analytics, Big Data, Data Mining, Data Science, Machine Learning](#)
- [Plataforma H2O](#)
- [Computer vision](#)
- [Pautas para dar formato al código programando en R: Google, Hadley Wickham \(RStudio\) y Coding Club](#)
- [The State of Naming Conventions in R](#)
- [R Code – Best practices](#)

- [R Coding Style Guide](#)
- [Naming files](#)
- [Documentacion de R](#)
- [Chuleta de expresiones regulares](#)
- [Regular Expressions Every R programmer Should Know](#)
- [Regular Expression Language - Quick Reference](#)
- [Dealing with Regular Expressions](#)
- [Writing an R package from scratch](#)
- [16 Cursos](#)
- [Galerias de graficos](#)
- [Curso Caltech. Learning from data](#)
- [Usar git](#)
- [Blogs con github](#) y [Blogs con github y RStudio](#)
- [Ejemplos de Shiny](#)
- [UK government using R to modernize reporting of official statistics](#)
- [Ggplot](#)
- [Sistemas de Coordenadas. Aqui y aqui](#)
- [Codificación de caracteres](#)
- [Tutorials for learning R](#)
- [Awesome R](#)
- [R Data Science Tutorials](#)
- [useR! Machine Learning Tutorial](#)
- [Computerworld - Paquetes de R interesantes](#)
- [Otra lista de recursos variados en Github](#)
- [Tipos de licencias de software](#)
- [Glosario de Machine Learning de Google](#)
- [Google Rules of Machine Learning: Best Practices for ML Engineering](#)
- [Statistical Learning de Stanford with R](#) [Curso](#), [Libro](#), [Código](#) y [Transparencias](#)
- [An Introduction to Statistical Learning - Web R & Python](#)
- [100 Free Tutorials for Learning R](#)
- [Google's best practices in machine learning](#)
- [Web Scraping TripAdvisor, Text Mining and Sentiment Analysis for Hotel Reviews](#)
- [Common Probability Distributions: The Data Scientist's Crib Sheet](#)
- [RDocumentation](#)
- [ArcGIS to R spatial cheat sheet](#)
- [Hand-curated, high quality resources for doing data journalism with R](#)
- [RMarkdown Driven Development \(RmdDD\)](#)
- [htmlwidgets for R - gallery](#)
- [Free R Reading Material](#)
- [The Chartmaker Directory](#)
- [Kaggle Winning Solutions](#)
- [Simplifying the ROC and AUC metrics](#)
- [Feature Engineering for Machine Learning](#)
- [Awesome Data Science](#)
- [Data Scientist Roadmap](#)
- [Data Science Collected Resources](#)
- [Data Science Cheatsheets](#)
- [Data Science Resources](#)
- [R for Water Resources Data Science](#)
- [GIS and mapping](#)
- [R package primer](#)
- [Using Geospatial Data in R](#)
- [Cómo crear una API en Python](#)
- [Cursos para aprender más sobre R](#)

- [Estadística con R](#)
- [Practical Introduction to Web Scraping in R](#)
- [R for Water Resources Data Science](#)
- [A ggplot2 Tutorial for Beautiful Plotting in R](#)
- [Bivariate Choropleth Maps: A How-to Guide](#)
- [Tipos de licencias open data \(minicurso de data.europa.edu\)](#)
- [Legalidad Web scraping: Is Web Scraping Legal? : The Definitive Guide \(2024 update\)](#) y [Web Scraping: ¿legal o ilegal?](#)
- [Information is Beautiful Awards](#)
- [Financial-Times / chart-doctor](#)
- [The Data Visualisation Catalogue](#)
- [From Data to Viz](#)
- [EUMETSAT science studies](#)
- [overpass turbo - Herremainta de filtrado para OSM](#)
- [Remote Sensing for OSINT](#)
- [Investigative Journalism with Satellite Images](#)
- [CAMIS - A PHUSE DVOST Working Group](#). 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](#)
- [GIS formats](#)
- [Data Viz Catalogue](#)
- [Dataviz Project](#)
- [Periodic Table Of Visualization Methods](#)
- [The TimeViz Browser 2.0](#)
- [The R Graph Gallery](#)
- [HOT - Drone Tasking Manager](#)
- [Remote sensing image retrieval](#)
- [BigEarthNet A Large-Scale Sentinel Benchmark Archive](#)
- [Global Fishing Watch](#). AI and satellite imagery to reveal the expanding footprint of human activity at sea. [Post](#). [Github](#). [Train data](#). [Analysis data](#) and [Vessel detection from Sentinel-1 SAR](#)
- [rseek.org - rstats search engine](#)

Libros

- [R Markdown Cookbook](#)
- [R Markdown: The Definitive Guide](#)
- [Todos los libros en bookdown](#)
- [R intro](#)
- [R for everyone](#)
- [R in action](#)
- [R Programming for Data Science](#). Roger D. Peng.
- [R para principiantes](#)
- [Introducción a R](#)
- [R para profesionales de los datos: una introducción](#)
- [10 Free Must-Read Books for Machine Learning and Data Science](#)
- [Introduction to Data Science Libro y Código](#)
- [Fundamentals of Data Visualization Libro y Código](#)
- [Data Science Live Book Libro y Código](#)
- [R for Statistical Learning Libro y Código](#)
- [Applied Statistics with R Libro y Código](#)
- [Geocomputation with R Libro y Código](#)
- [Handling Strings with R](#)
- [Text Mining with R](#)

- [Efficient R programming](#)
- [BBC Visual and Data Journalism cookbook for R graphics](#)
- [Databases using R by RStudio](#)
- [Interpretable Machine Learning](#)
- [Forecasting: Principles and Practice](#)
- [Bioinformática Estadística. Análisis estadístico de datos Ómicos](#)
- [Estadística básica](#)
- [Probabilidad básica](#)
- [Estadística básica](#)
- [An Introduction to Spatial Data Analysis and Visualisation in R](#)
- [Exploratory Data Analysis with R - Roger D. Peng](#)
- [What They Forgot to Teach You About R](#)
- [Mastering Apache Spark with R Libro y Código](#)
- [Hands-On Machine Learning with R](#)
- [Hands-On Programming with R](#)
- [The 20 Best Data Science Books Available online in 2020](#)
- [Creating APIs in R with Plumber](#)
- [Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny](#)
- [Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA](#)
- [Estilometría, análisis de textos en R para filólogos](#)
- [Engineering Production-Grade Shiny Apps](#)
- [Bayesian inference with INLA](#)
- [Happy Git and GitHub for the useR](#)
- [R for Data Science. Inglés y Castellano](#)
- [Data Visualization with R](#)
- [Big Book of R](#)
- [JavaScript for R](#)
- [Hands-On Data Visualization](#)
- [Fundamentals of Data Visualization](#)
- [Supervised Machine Learning for Text Analysis in R](#)
- [Modern R with the tidyverse](#)
- [Spatial Microsimulation with R](#)
- [R Advanced Spatial Lessons](#)
- [R Packages](#)
- [R Graphics Cookbook](#)
- [Advanced R](#)
- [rstudio4edu](#)
- [R for Health Data Science](#)
- [Linear Algebra for Data Science](#)
- [Aprendizaje Estadístico con R](#)
- [Simulación Estadística con R](#)
- [R para profesionales de los datos: una introducción](#)
- [Aprendiendo R sin morir en el intento](#)
- [Interpretable Machine Learning](#)
- [The caret Package](#)
- [Elegant and informative maps with tmap](#)
- [Spatial Modelling for Data Scientists](#)
- [Command Line Basics for R Users](#)
- [Engineering Production-Grade Shiny Apps](#)
- [Outstanding User Interfaces with Shiny](#)
- [The Shiny AWS Book](#)
- [Twitter for Scientists](#)
- [Spatial Data Science](#)
- [Handbook of Graphs and Networks in People Analytics](#)

- [Officeverse R & Office](#)
- [Spatial Analysis With R](#)
- [R Programming for Data Science](#)
- [The R Book](#)
- [R Notes for Professionals](#)
- [An Introduction to Spatial Data Analysis and Visualisation in R](#)
- [Think Bayes 2e](#)
- [Quantitative Politics with R](#)
- [Data Science in Education Using R](#)
- [Data Skills for Reproducible Science](#)
- [Handbook of Regression Modeling in People Analytics](#)
- [Handbook of Graphs and Networks in People Analytics With Examples in R and Python](#)
- [Tidy Finance with R](#)
- [Predictive Soil Mapping with R](#)
- [Deep Learning](#)
- [Technical Foundations of Informatics](#)
- [YaRrr! The Pirate's Guide to R](#)
- [An R companion to Statistics: data analysis and modelling](#)
- [Learning Statistics with R](#)
- [Happy Git and GitHub for the useR](#)
- [Introduction to Econometrics with R](#)
- [R4JournalismBook](#)
- [Libro Vivo de Ciencia de Datos](#)
- [Introduction to Probability for Data Science](#)
- [Spatial Data Science with applications in R](#)
- [R for Data Analysis](#)
- [Data Analysis and Prediction Algorithms with R](#)
- [Análisis de datos y algoritmos de predicción con R](#)
- [The Art of Data Science](#)
- [R for data science: tidyverse and beyond](#)
- [Statistical Inference via Data Science](#)
- [Introduction to urban accessibility: a practical guide in R](#)
- [Tidy Finance](#)
- [Deep Learning and Scientific Computing with R torch](#)
- [Econometrics with the Tidyverse](#)
- [Building reproducible analytical pipelines with R](#)
- [Satellite Image Time Series Analysis on Earth Observation Data Cubes](#)
- [Spatial Statistics for Data Science: Theory and Practice with R](#)
- [Geographic Data Science with R: Visualizing and Analyzing Environmental Change](#)
- [Apache Arrow R Cookbook](#)
- [An Introduction to R and Python For Data Analysis: A Side By Side Approach](#)
- [Using the flextable R package](#)
- [Spatial sampling with R](#)
- [Interactive web-based data visualization with R, plotly, and shiny](#)
- [Modern Data Science with R](#)
- [Quarto for Scientists](#)
- [Introduction to Environmental Data Science](#)
- [Manual de R](#)
- [Modern R with the tidyverse](#)
- [Tidy Modeling with R](#)
- [Data Analysis in Medicine and Health using R](#)
- [Crime by the Numbers: A Criminologist's Guide to R](#)
- [R Tutorial: Mastering Data Analysis and More](#)
- [Data transformation with R](#)

- [Data Visualization and Exploration with R](#)
- [Introduction to Probability and Statistics Using R](#)
- [Fundamentos de ciencia de datos con R](#)
- [Financial Risk Modelling and Portfolio Optimization with R](#)
- [R in Finance: Introduction to R and Its Applications in Finance](#)
- [Deep R Programming](#)
- [Beyond Multiple Linear Regression](#)
- [Analyzing US Census Data: Methods, Maps, and Models in R](#)
- [Psychometrics in Exercises using R and RStudio](#)
- [Introduction to NFL Analytics with R](#)
- [Feature Engineering and Selection: A Practical Approach for Predictive Models](#)
- [Applied Machine Learning Using mlr3 in R](#)
- [Behavior Analysis with Machine Learning Using R](#)
- [Data Analytics: A Small Data Approach](#)
- [R for Social Network Analysis](#)
- [Cookbook Polars for R](#)
- [DIY API with Make and {plumber}](#)
- [ggplot2: Elegant Graphics for Data Analysis \(3e\)](#)
- [R for the Rest of Us: A Statistics-Free Introduction](#)
- [Un Recorrido por los Métodos Cuantitativos en Ciencias Sociales a bordo de R](#)
- [What They Forgot to Teach You About R](#)
- [Tidyverse Skills for Data Science](#)
- [A Business Analyst's Introduction to Business Analytics](#)
- [Telling Stories with Data](#)
- [The tidyverse style guide](#)
- [Empirical Research in Accounting: Tools and Methods](#)
- [Data Visualization with R](#)
- [Introduction to MySQL with R](#)
- [Physical Oceanography Distributed Active Archive Center tutorials](#)
- [The SAR Handbook](#)
- [Advanced International Trade in R](#)
- [The Orange Book of Machine Learning](#)