

# Open Data

by [Santiago Mota](#)

2022-02-24

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](#)
- [Revolution Analytics datasets \(Microsoft R\)](#)
- [UCI Machine Learning Repository](#)
- [DH Network](#)
- [9 Datasets para sistemas de recomendación](#)
- [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](#)
- [Recopilatorio de datasets en Github](#)
- [API de GitHub](#)
- [API de Facebook](#)
- [Blog. 100 recursos sobre Big Data y Data Science](#)
- [NASDAQ](#)
- [Google finanzas](#)
- [CaixaBank Research](#)
- [Satélite Landsat](#)
- [OCDE](#)
- [19 Free Public Data Sets For Your First Data Science Project](#)
- [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](#)
- [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
- Microsoft Project Oxford
- 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
- Conferencia internacional de datos abiertos en Madrid
- Open Data Inception. 1.600 portales abiertos
- Yahoo Flickr Creative Commons 100 Million (YFCC100m) dataset
- ImageNet database
- Deep Learning datasets
- 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
- 70 amazing and free data sources for data visualization
- Some datasets for teaching data science
- 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
- 10.000 airports, train stations and ferry terminals spanning the globe
- 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 Canario de Estadística
- Datos estadísticos DGT

- [Yelp Dataset](#)
- [Awesome Sentinel](#). Copernicus Sentinel Satellites resources
- [Datos Abiertos del Consorcio Regional de Transportes de Madrid](#)
- [Idealista labs](#)
- [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](#)
- [Github semantic code search](#) [page](#) 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](#)
- [Open Spatial Demographic Data](#)
- [Center for Disease Control](#)
- [NASA](#)
- [World Economic Forum](#)
- [Universidad de Harvard](#)
- [MIT 1 y 2](#)
- [Indiana University Click Dataset](#)
- [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](#)
- [Base de datos de Series de Indicadores de Coyuntura Económica \(BDSICE\)](#)
- [Tematicas.org](#) Recopilación de series e índices
- [Amazon product data 2014](#)
- [Amazon product data 2018](#)
- [The Big Bad NLP Database](#)
- [openaddresses](#)
- [Awesome Geospatia](#)

## Otras referencias interesantes

- [Data Science Blogs](#)
- [Chuleta general de R](#)
- [R Learning Path: From beginner to expert in R in 7 steps](#)
- [Tutorial de R Markdown](#)
- [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 y Hadley Wickham \(RStudio\)](#)
- [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](#)
- [Aprender ciencia de datos. Fuentes para Python y R](#)
- [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](#)
- [Great R packages for data import, wrangling and visualization](#)
- [Ggplot](#)
- [Sistemas de Coordenadas. Aqui y aqui](#)
- [Codificación de caracteres](#)
- [Tutorials for learning R](#)
- [Awesome R](#)
- [R Data Science Tutorials](#)
- [useR! Machine Learning Tutorial](#)
- [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 Curso, Libro, Código y Transparencias](#)
- [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 \(RmdDD\)](#)
- [Is Web Scraping Legal? : The Definitive Guide \[2019 update\]](#)
- [Web Scraping: ¿legal o ilegal? \(2017\)](#)
- [United Nations World Urbanization Prospects](#)

- [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](#)

## Libros

- [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 great books about R](#)
- [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](#)
- [Análisis de datos con R](#)
- [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 R programmers](#)
- [Spatial Data Science](#)