

UNITED STATES ▼

FEATURE

These R packages import sports, weather, stock data and more

By Sharon Machlis

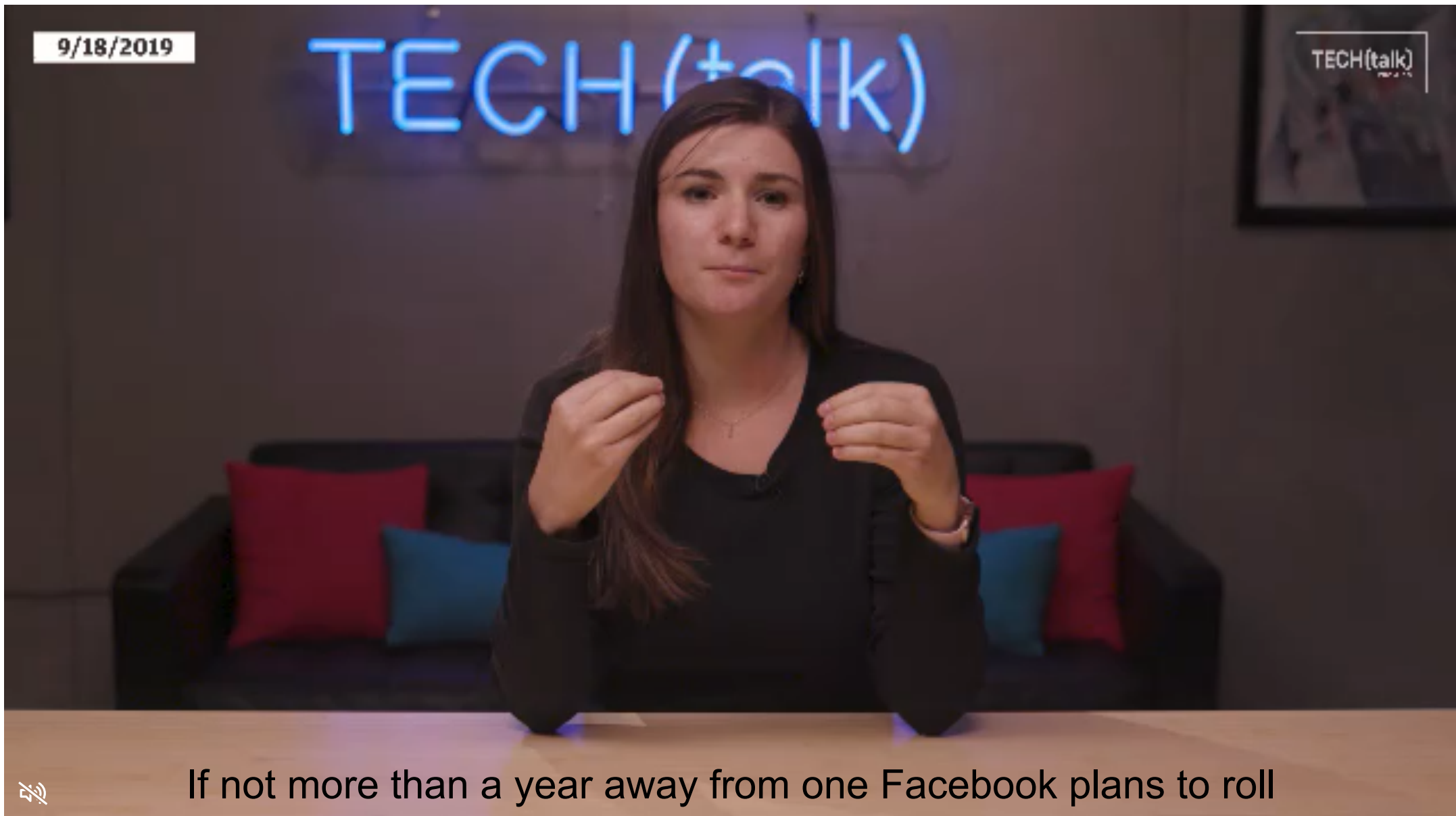
Executive Editor, Data & Analytics, Computerworld Mar 7, 2018 12:38 pm PST

There are lots of good reasons you might want to analyze public data, from detecting salary trends in government data to uncovering insights about a potential investment (or your favorite sports team).

But before you can run analyses and visualize trends, you need to have the data. The packages listed below make it easy to find economic, sports, weather, political and other publicly available data and import it directly into R -- in a format that's ready for you to work your analytics magic.

[Learn more about R with our Advanced Beginner's Guide to R]

Packages that are on CRAN can be installed on your system by using the R command `install.packages("packageName")` -- you only need to run this once. GitHub packages are best installed with the devtools package -- install that once with `install.packages("devtools")` and then use that to install packages from GitHub using the format `devtools::install_github("repositoryName/packageName")`. Once installed, you can load a package into your working session once each session using the format `library("packageName")`.



Some of the sample code below comes from package documentation or blog posts by package authors. For more information about a package, you can run `help(package="packageName")` in R to get info on functions included in the package and, if available, links to package vignettes (R-speak for additional documentation). To see sample code for a particular function, try `example(topic="functionName", package="packageName")` or simply `?functionName` for all available help about a function including any sample code (not all documentation includes samples).

For more on best R packages, see [Great R Packages for data import, wrangling and visualization](#).

Search the table below

R packages to import public data

PACKAGE	CATEGORY	DESCRIPTION	SAMPLE CODE	MORE INFO
blscrapeR	Economics, Government	For specific information about U.S. salaries and employment info, the Bureau of Labor Statistics offers a wealth of data available via this new package. blsAPI package is another option. CRAN.	<code>bls_api(c("LEU0254530800", "LEU0254530600"), startyear = 2000, endyear = 2015)</code>	Package vignettes
quantmod	Finance, Government	This package is designed for financial modelling but also has functions to easily pull data from Google Finance, Yahoo Finance and the St. Louis Federal Reserve (FRED). CRAN.	<code>getSymbols("DEXJPUS",src="FRED")</code>	Intro on getting data
Bureau of Economic Analysis	Economics, Government	Maintained by Andera Batch at BEA, this taps into the bureau's API to download data sets. CRAN.	<code>beaSpecs <- list('UserID' = beaKey , 'Method' = 'GetData', 'datasetname' = 'NIPA', 'TableName' = 'T20305', 'Frequency' = 'Q', 'Year' = 'X',</code>	See the GitHub repo, including info about recent project

			<code>'ResultFormat' = 'json'); beaPayload <- beaGet(beaSpecs);</code>	changes
edgarWebR	Finance, Government	This package is designed to let you search and download data from the U.S. Securities and Exchange Commission, including corporate and mutual-fund financial filings. CRAN.	<code>getSymbols("DEXJPUS",src="FRED")</code>	See the package vignette
FredR	Finance, Government	If you're interested just in Fed data, FredR can access data from the Federal Reserve Economic Data API, including 240,000 US and international data sets from 77 sources. Free API key needed. GitHub.	<code>fred <- FredR(api.key) fred\$series.search("GDP") gdp <- fred\$series.observations(series_id = 'GDPC1')</code>	Project's GitHub page
tidycensus	Government, GIS	This package downloads data from the U.S. 10-year census and American Community Survey in R-ready format. In addition, you can import data <i>and</i> geospatial files for easy mapping. Free API key needed. (For Canadian Census data, there's a completely separate <i>cancensus</i> package). CRAN.	<code>get_acs(state = "CA", county = "Orange", geography = "tract", variables = "B19013_001", geometry = TRUE)</code>	Project documentation page
censusapi	Government	This aims to be complete and offer data from all the bureau's APIs, not just the decennial Census and ACS. API key required. GitHub.	<code>mydata <- getCensus(name="acs5", vintage=2014, key=mycensuskey, vars=c("NAME", "B01001_001E", "B19013_001E"), region="congressional district:*", regionin="state:36")</code>	See the Computerworld tutorial or package's GitHub repo.
ipumsr	Government	For those doing heavy-duty Census analysis with raw IPUMS microdata, this package makes it easy to import that data into R. CRAN.	<code>read_ipums_micro()</code>	See the package vignettes, including the

				intro vignette.
eu.us.opendata	Government	This joint project of the U.S. and European Union governments aims to make it easy to import and compare US and EU data sets. Free U.S. Bureau of Economic Analysis API key needed. GitHub.	<pre>getRel('gross domestic product', lucky = T, beaKey = myKey) getRel('gdp', lucky = T, beaKey = myKey)</pre>	See the project's README file for installation and usage instructions.
RSocrata	Government	Pull data from any municipality that uses the Socrata data platform. Created by the City of Chicago data team. CRAN.	<pre>mydata <- read.socrata("https://data.cityofchicago.org/ Transportation/Towed- Vehicles/ygr5-vcbg")</pre>	RSocrata GitHub repo
forbesListR	Misc	A bit of a niche offering, this taps into lists maintained by Forbes including largest private companies, top business schools and top venture capitalists. GitHub.	<pre>#top venture capitalists 2012-2016 mydata <- get_years_forbes_list_data(years = 2012:2016, list_name = "Top VCs")</pre>	See the project GitHub page. You may need to manually load the tidyr package for code to work.
pollstR	Politics	This package pulls political polling data from the Huffington Post Pollster API. CRAN.	<pre>slug <- "donald-trump-favorable- rating" polls <- pollster_charts_polls(slug) [["content"]]</pre>	See the example vignette
data.world	Multi	R package to use data sets from data.world. API key needed. CRAN.	See examples in GitHub repo.	See the Quick Start vignette
Lahman	Sports	R interface for the famed Lahman baseball database. CRAN.	<pre>batavg <- battingStats()</pre>	See this blog post on using R

				for baseball research
rnoaa	Weather	Tap into numerous National Oceanic and Atmospheric Administration APIs, including climate, tornadoes and the Climate Prediction Center. NOAA API key needed.	<code>options(noaakey = "yourAPIkey") storms14 <- storm_data(year = 2014)</code>	See this rOpenSci tutorial.
rtweet	Social Media	Tap into Twitter's REST and stream APIs with R. API key needed. CRAN.	<code>search_tweets("#rstats", n = 18000, include_rts = FALSE)</code>	See the introductory vignette.

Sharon Machlis is Executive Editor, Data & Analytics at IDG, where she works on data analysis and in-house editor tools in addition to writing and editing. Her book *Practical R for Mass Communication and Journalism* was published in December 2018.

Follow    

Copyright © 2018 IDG Communications, Inc.

Where does this document go — OneDrive for Business or SharePoint?

SPONSORED STORIES



Free trial Watson Knowledge Catalogue to index and catalog data assets

IBM



Don't Turn Off Your Computer Without Doing This First

The Review Experts



[Free Trial] Empower your business with App Connect

IBM

Where the World's Billionaires Live

Mansion Global

Ireland: Everyone Is Going Crazy Over This Cheap WiFi Booster

Wise Web Savings

Los Angeles' Luxury Home Market at a Glance: See What's New

Mansion Global



iPhone 11: What is Deep Fusion and how does it work?

Homepage



The major hybrid cloud options compared: AWS Outposts vs Azure Stack vs Google Anthos

Homepage



The best UK universities for computer science and IT degrees

Homepage



(2019 Apple) #1 Computer Protection Now Free For Mac Users Only
My Antivirus Review



[Download] Latest insights on the increasingly complex role of the CMO
IBM

SHOP TECH PRODUCTS AT AMAZON

SPONSORED LINKS

dtSearch® instantly searches terabytes of files, emails, databases, web data. See site for hundreds of reviews; enterprise & developer evaluations

It's a Race to Connect the Globe. 451 Research and Iron Mountain explore this Hybrid IT trend.

Online Master of Science in Information Systems at Northwestern University

Copyright © 2019 IDG Communications, Inc.