DATA 607 Week 9 Assignment

Matthew Roland

2023-10-25

Article Search API

For this assignment, I will be using the NYTimes Books API to return a list of all the NYT Best Sellers, and then I will convert the resultant JSON file into an R dataframe.

##Running the API

```
##This code will generte a JSON format list containing historical data pertaining to some of the NYT be

nyt <- from JSON("https://api.nytimes.com/svc/books/v3/lists/best-sellers/history.json?api-key=PCBxRC5ID

##And now I will convert the JSON list into a dataframe
nyt <- data.frame(nyt)

kable(nyt)</pre>
```

statopyrightn <u>es</u> rd s siltislelts.desesuhtisodussiltisleltsuhtisle						results.reviews	
OKCopyi35 (c) 2023 The New York Times Company. All Rights Reserved.		The author of the	by Di- ana	Diana Ga- bal- don	0.00	Dell039917397178570, 9780399178573, 8, , Advice How-To and Miscellaneous, 97803991785,7How-To & Miscellaneous, 2016-09-04, 2016-08-20, 1, 0, 0, 0	, , ,

statuspyrigh	nutn <u>es</u> neltssultistaletts.d	l esesiphteisaaktetseihtektele	nipiiskulgsuplgskiidpes:ranks_history	results.reviews
OKCopyr3g (c) 2023 The New York Times Company. All Rights Reserved.	BLESSEHAR- OF acter THE study PA- that TRI- at- ARCHSempts to make sense of Jeffer- son's con- tra- dic- tions.	by Annette 0.00 An- Gordon- netteReed Gordond ReedPe- and ter Pe- S ter Onuf S. Onuf	Liver Mark Liver Nonfiction, Hardcover 97808 No. 464. 1280, 2016-05-01, 2016-04-16, 1, 0, 1, 0	
OKCopyr35 (c) 2023 The New York Times Company. All Rights Reserved.	56YOU The JUST co- NEED host TO of the LOSE pod- WEIGHEST "Mainte- nance Phase" examines myths about gaining and losing weight to dismantle antifat bias.	by Aubrey 0.00 Aubr@pr- Gor-don don	Bea 08070080770 06475, 0807006475, , 9780807006474, 9780807006474, 2, 6 080700 F48 3 rback Nonfiction, Combined , Print and E-Book Nonfiction, 9780 80700648 E-Book Nonfiction, 2023-01-29, 2023-01-29, 2023-01-14, 2023-01-14, 1, 1, 0, 0, 0, 0, 0, 0, 0	, , ,

```
statosyrightnesrelsselteslelts.desesiphisooliteseihitslichesphisoolitesenks_history
                                                                                                results.reviews
OKCopyr355#ASKCFARYVEEy Gary
                                        0.00
                                               Har 0 0 6 2 2 7 3 1 2 4 , 0 0 6 2 2 7 3 1 2 4 , 0 0 6 2 2 7 3 1 2 4 ,
                                                                                                , , ,
                           GaryVayn-
                                                        9780062273123, 9780062273123, 5, 6
   (c)
                  en-
   2023
                  trepreneuVayner-
                                                  006227B132ness Books , Advice How-To
   The
                           er- chuk
                  ex-
                                                        and Miscellaneous, Business, Advice,
  New
                           chuk
                                                  97800H2273h23z Miscellaneous , 2016-04-10
                  pands
   York
                                                  9780062076-1030-27, 2016-03-26, 2016-03-12
                  on
   Times
                  sub-
                                                        , 0, 1, 0, 0, 0, 0, 1, 1
   Com-
                  jects
  pany.
                  ad-
   All
                  dressed
  Rights
                  on his
  Re-
                  Inter-
   served.
                  net
                  show,
                  like
                  mar-
                  ket-
                  ing,
                  man-
                  age-
                  ment
                  and
                  social
                  me-
                  dia.
                                               Por(639)3/1692128179/3Puth591847931, 1591847931
OKCopyr3₫₺₩GIRLAOSS
                           by Sophia 0.00
                  online
                           Soph Amoruso
                                                        , 1591847931, 039916927X,
   (c)
   2023
                  fash-
                            Amoruso
                                                  1591894789B191847939, 9781591847939,
   The
                                                        9781591847939, 9781591847939,
                  ion
   New
                                                  978039793653927469274, 8, 9, 9, 8, 10,
                  re-
   York
                  tailer
                                                  9781591847939Books, Business Books,
   Times
                  traces
                                                        Business Books, Business Books,
   Com-
                  her
                                                        Business Books, Business, Business,
  pany.
                  path
                                                        Business, Business,
   All
                  to
                                                        2016-03-13, 2016-01-17, 2015-12-13,
  Rights
                                                        2015-11-15, 2014-11-09, 2016-02-27,
                  suc-
   Re-
                  cess.
                                                        2016-01-02, 2015-11-28, 2015-10-31,
                                                        2014-10-25, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  served.
                                                        0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
OKCopyr365@HMOMSOHARD Kristin 0.00
                                               Harp0628676286769X, 9780062857699, 10,
   (c)
                           Kristlinens-
                                                        Advice How-To and Miscellaneous,
  2023
                           Hensley
                                                  97800&28576,9Elow-To & Miscellaneous ,
  The
                           ley and
                                                        2019-04-21, 2019-04-06, 1, 0, 0, 1
                           and Jen
  New
                           Jen Smed-
   York
  Times
                           Smedley
  Com-
                           ley
   pany.
   All
  Rights
  Re-
   served.
```

sta tro spyri ghtn<u>es</u>nds	sulttistsheelts.d	escariphteisanlutets	ail#stcl#s	aphintesultabilites:ranks_history	results.reviews
OKCopyr356#NEV (c) 2023 The New York Times Company. All Rights Reserved.	from Mar- jory Stone- man Dou- glas High School de- scribe the Valen- tine's Day mass shoot- ing and out- line ways to pre- vent simi- lar inci- dents.	HNy David DaviHogg Hoggand and Lau- Lau-ren ren Hogg Hogg	0.00	Rank	, , ,
OKCopyr 35588 00 (c) STAR 2023 The New York Times Company. All Rights Reserved.	How	by Chris ChrisGuille- Guilleeau beau	23.00	Crows07981520. Busi- ness9780307951526	, , ,

sta tos pyr nghin esnessillessi	alts.desesuphasonhesaihaich	supidade on the transfer of the supplemental	results.review
OKCopyr 35592 0 (c) PER 2023 GAL- The LON New York Times Com- pany. All Rights Re-	by Christophen Christoeiner pher Steiner	GraMULINULL Cen- tral	, , ,
served. OKCopyr 355657 , NA (c) Chicago 2023 The New York Times Company. All Rights Reserved.	NA Stev&A 0.00N Mon- roe	ANA 07868NT3DL , 9780786867301	, NA , https: //www. nytimes. com/ 2001/ 07/29/ books/ books- in-brief- fiction- poetry- 319660. html,
OKCopyi35643OCKNA (c) OF 2023 AGES:"RO The STONE" New HIS- York TORY Times OF Com- ROCK pany. AND All ROLL' Rights Re- served.	STOKES,	ANA 06716 3068 7. , 9780671630683	NA , NA , https: //www. nytimes. com/ 1986/ 12/28/ books/ three- chord- music- in-a- three- piece- suit. html, NA

sta tu spyrig	hutn <u>es</u> ne lss.htissle lts	s.desesuphisonhesaihistich	appingsplybillerranks_history	results.reviews
OKCopyr 3	559BHE NA	NA KATNEA 0.00N	ANA 03742 NTOIDD	, NA ,
(c)	HIGH	TELTSCHER	· ,	https:
2023	ROAD		9780374217006	//www.
The	TO			nytimes.
New	CHINA:			com/
York	GEORGE			2007/
Times	BOGLE,			04/22/
Com-	THE			books/
pany.	PANCHEN			review/
All	LAMA			Stuart.t.
Rights	AND			$_{ m html},$
Re-	THE			NA
served.	FIRST			
	BRITISH			
	EX-			
	PE-			
	DI-			
	TION			
	TO			
	TI-			
	BET'			
OKCopyr 3	§ 5 ₽BIL	by Sharon 0.00	Har le7783N4217B	, , ,
(c)	DEATH	Shar Sa la	Mira	
2023		Sala	9780778314271	
The				
New				
York				
Times				
Com-				
pany.				
All				
Rights				
Re-				
served.				

sta tros pyri ght	n <u>es</u> netsultists	helts.descanin	hteisonhtetsailutstuchusuphisko	desulgabillus:ranks_history	results.reviews
2023 The	DEATHINA DO ma US in PART Via En lar tur to cri no ist hel wh she sta re- cei ing ser of tur per	atch- Ama aker Quid cto- n g- ad cns a me evel- for lp aen e crts y- g a iies dis- cbingly r- nal- d n-	a Qda ick	\$\\ \\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$\\$, , , ,
2023 The New York Times Company. All Rights Reserved.		A NA	Naoi N iA 0.00NANA Schae- fer Ri- ley	01998 N374L , 9780199873746	, NA , https: //www. nytimes. com/ 2013/ 05/26/ books/ review/ til-faith- do-us- part-by- naomi- schaefer- riley. html, NA

sta tos pyrigd	ntn <u>es</u> uds	sultisshe lts.c	l esceriphteisanhtets	zibtstch:	nipisinite on the control of the con	results.reviews
OKCopyr3g (c) 2023 The New York Times Company. All Rights Reserved.	THE SEA-SON	Two classic holiday stories — "Under the Christmas Tree" (2009) and "Midnight Confessions" (2010) — plus the novella "Backward Glance"	by Ron RobyGarr Carr	0.00	Harle 778 80 768 83 16645, 0778 316645, Mira 9780 778 316640, 9780 778 316640, 18, 9780 778 316640 Market Paperback, Mass Market Paperback, Paperback Mass-Market Fiction, Paperback Mass-Market Fiction, 2014-11-30, 2014-11-16, 2014-11-15, 2014-11-01, 0,1,0,0,0,1,0,0	, , ,
OKCopyr3g (c) 2023 The New York Times Company. All Rights Reserved.	56(B E)I IN THE USA	(1991). SORA soccer commentator describes how he embraced American popular culture while growing up in Liver- pool.	by Roger RogeBen- Ben-nett nett	0.00	Dey0062958690 58690 , 0062958690 , Street 9780062958693 , 9780062958693 , 3 , 1 0062958720bined Print and E-Book , Nonfiction, Hardcover Nonfiction , 9780062958693 Print & E-Book Nonfiction 9780062958723 Print & E-Book Nonfiction 9780062958723 Print & 2021-07-18 , 2021-07-18 , 2021-07-03 , 2021-07-03 , 1 , 1 , 0 , 0 , 0 , 0 , 0 , 0	, , ,

statoopyrigi	hin <u>es</u> ræsslitissælts.	de sesulptisionlitetsilitätielisulpisinlesulpisihlites ranks_history	results.reviev
OKCopyr38 (c) 2023 The New York Times Company. All Rights Re-	5593 —, THAT'S DE- LI- CIOUS	by Action 0.00 Abrital 9726552, 9781419726552, 9, Ac- Brontion son 97814A072655Elow-To & Miscellaneous, Bronwith 2017-10-01, 2017-09-16, 1, 0, 0, 1 son Rachel with Whar-Rachteln Whar-ton	, , ,
served. OKCopyr36 (c) 2023 The New York Times Company. All Rights Re- served.	the Horse He Rode In On: The Peo- ple V. Ken- neth Starr	NA James A 0.00NANA 06848NT3440. Carville , 9780684857343	, NA , https: //www. nytimes. com/ 1998/ 10/18/ books/ preaching- to-the- converted. html, NA
OKCopyr36 (c) 2023 The New York Times Company. All Rights Reserved.	5643 ACKThis G.U. series, , set in VOL. the 5 fu- ture, is about an on- line, multi- player game run amok. This vol- ume con- cludes the	by Hamazakii Tat-suya suya	, , ,

As we can see, this generates a rather unorganized (and incomplete) dataframe consisting of empty and nested columns (and even empty, nested columns) Of course, it is necessary to clean and tidy all of these

datapoints

```
##First, I want to remove the "results." prefix
colnames(nyt) <- sub("^results.", "", colnames(nyt))</pre>
nyt_unnest <- nyt %>% unnest_wider(c("ranks_history", "reviews", "isbns"))
#dropping the primary isbn variables, since they contain the same information contained within the nest
nyt_unnest <- subset(nyt_unnest, select = -c(primary_isbn10, primary_isbn13))</pre>
nyt_unnest <- nyt_unnest %>% unnest_longer(isbn10)
nyt_unnest <- nyt_unnest %>% unnest_longer(isbn13)
nyt unnest <- nyt unnest %>% unnest longer(rank)
nyt_unnest <- nyt_unnest %>% unnest_longer(list_name)
nyt_unnest <- nyt_unnest %>% unnest_longer(display_name)
nyt_unnest <- nyt_unnest %>% unnest_longer(published_date)
nyt unnest <- nyt unnest %>% unnest longer(bestsellers date)
nyt_unnest <- nyt_unnest %>% unnest_longer(weeks_on_list)
nyt_unnest <- nyt_unnest %>% unnest_longer(rank_last_week)
nyt_unnest <- nyt_unnest %>% unnest_longer(asterisk)
nyt_unnest <- nyt_unnest %>% unnest_longer(dagger)
glimpse(nyt_unnest)
## Rows: 7,821,209
## Columns: 26
                     <chr> "OK", "OK", "OK", "OK", "OK", "OK", "OK", "OK", "~
## $ status
## $ copyright
                     <chr> "Copyright (c) 2023 The New York Times Company.
                    <int> 35593, 35593, 35593, 35593, 35593, 35593, ~
## $ num_results
                    <chr> "\"I GIVE YOU MY BODY ...\"", "\"MOST BLESSED OF ~
## $ title
## $ description
                    <chr> "The author of the Outlander novels gives tips on~
## $ contributor
                     <chr> "by Diana Gabaldon", "by Annette Gordon-Reed and ~
## $ author
                     <chr> "Diana Gabaldon", "Annette Gordon-Reed and Peter ~
                    ## $ contributor note
                     <chr> "0.00", "0.00", "0.00", "0.00", "0.00", "0.00", "~
## $ price
                     ## $ age_group
                     <chr> "Dell", "Liveright", "Beacon", "Beacon"~
## $ publisher
## $ isbn10
                     <chr> "0399178570", "0871404427", "0807006475", "080700~
## $ isbn13
                     <chr> "9780399178573", "9780871404428", "9780807006474"~
                     ## $ rank
                     <chr> "Advice How-To and Miscellaneous", "Hardcover Non~
## $ list_name
                     <chr> "Advice, How-To & Miscellaneous", "Hardcover Nonf~
## $ display_name
                     <chr> "2016-09-04", "2016-05-01", "2023-01-29", "2023-0~
## $ published_date
                     <chr> "2016-08-20", "2016-04-16", "2023-01-14", "2023-0~
## $ bestsellers_date
## $ weeks_on_list
                     ## $ rank_last_week
                     ## $ asterisk
                     ## $ dagger
                     ## $ book_review_link
                     ## $ first chapter link
## $ sunday_review_link
```

And with that, we are left with a very, very large (albeit analyzable) unnested dataset