

# 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 genereate a JSON format list containing historical data pertaining to some of the NYT be  
nyt <- fromJSON("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)
```

status	copyright	res	title	author	description	rank	rank_history	reviews
OK	Copyright	3550B	The	by Diana	0.00	Dell	0399103991078570 , 9780399178573 , 8 ,	, , ,
(c)	GIVE	au-	Di- Ga-				, Advice How-To and Miscellaneous,	
2023	YOU	thor	ana bal-			9780399178573	How-To & Miscellaneous ,	
The	MY	of the	Ga- don				2016-09-04 , 2016-08-20 , 1 , 0 , 0 , 0	
New	BODY	Out-	bal-					
York	... ”	lander	don					
Times		novels						
Com-		gives						
pany.		tips						
All		on						
Rights		writ-						
Re-		ing						
served.		sex						
		scenes,						
		draw-						
		ing on						
		exam-						
		ples						
		from						
		the						
		books.						

status	copyrights	titles	descriptions	results	source	titles	status	titles	ratings	reviews
OK	Copyright	3550	OSTA	by Annette	0.00	Live	0871408727	04427 , 9780871404428 , 16 , , ,		
(c)		BLESSED	Har-	An- Gordon-				, Hardcover Nonfiction, Hardcover		
2023		OF	acter	netteReed			9780871404428	, 2016-05-01 , 2016-04-16 , 1		
The		THE	study	Gordon				, 0 , 1 , 0		
New		PA-	that	ReedPe-						
York		TRI-	at-	and ter						
Times		ARCHS	empts	Pe- S						
Com-			to	ter Onuf						
pany.			make	S.						
All			sense	Onuf						
Rights			of							
Re-			Jeffer-							
served.			son's							
			con-							
			tra-							
			dic-							
			tions.							
OK	Copyright	3550	YOU The	by Aubrey	0.00	Bea	0807006475	006475 , 0807006475 , , ,		
(c)		JUST	co-	Aubrey				, 9780807006474 , 9780807006474 , 2 , 6		
2023		NEED	host	Gor-			0807006475	Paperback Nonfiction , Combined		
The		TO	of the	don				, Print and E-Book Nonfiction,		
New		LOSE	pod-				9780807006474	Nonfiction , Combined		
York		WEIGHT	has				9780807006474	E-Book Nonfiction ,		
Times			"Main-					2023-01-29 , 2023-01-29 , 2023-01-14 ,		
Com-			te-					2023-01-14 , 1 , 1 , 0 , 0 , 0 , 0 , 0 , 0		
pany.			nance							
All			Phase"							
Rights			exam-							
Re-			ines							
served.			myths							
			about							
			gain-							
			ing							
			and							
			losing							
			weight							
			to dis-							
			man-							
			tle							
			anti-							
			fat							
			bias.							

[illegible]



status	copyright	rights_reserved	results.description	results.sources	results.channels	results.pubs	results.urls	results.ranks_history	results.reviews
OK	Copyright	355930	(c) PER 2023 GAL- The LON New York Times Com- pany. All Rights Re- served.	by Christopher Steiner pher Steiner	0.00	Grand	NULL	NULL	, , ,
OK	Copyright	355937	(c) Chicago 2023 The New York Times Com- pany. All Rights Re- served.	NA Steve Mon- roe	NA 0.00	NA	0786867301	0786867301	, NA , https: //www. nytimes. com/ 2001/ 07/29/ books/ books- in-brief- fiction- poetry- 319660. html, NA , NA , https: //www. nytimes. com/ 1986/ 12/28/ books/ three- chord- music- in-a- three- piece- suit. html, NA
OK	Copyright	355938	(c) OF 2023 AGES:'ROLLING The STONE" New HIS- York TORY Times OF Com- ROCK pany. AND All ROLL' Rights Re- served.	NA GEOFFREY STOKES, KEN TUCKER' 'ED WARD	NA 0.00	NA	0671630683	0671630683	





status	copyrights	titles	descriptors	results	details	links	ratings	reviews	history	results	reviews
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								
			“Mid-								
			night								
			Con-								
			fes-								
			sions”								
			(2010)								
			—								
			plus								
			the								
			novella								
			“Back-								
			ward								
			Glance”								
			(1991).								
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								
			“Mid-								
			night								
			Con-								
			fes-								
			sions”								
			(2010)								
			—								
			plus								
			the								
			novella								
			“Back-								
			ward								
			Glance”								
			(1991).								
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								
			“Mid-								
			night								
			Con-								
			fes-								
			sions”								
			(2010)								
			—								
			plus								
			the								
			novella								
			“Back-								
			ward								
			Glance”								
			(1991).								
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								
			“Mid-								
			night								
			Con-								
			fes-								
			sions”								
			(2010)								
			—								
			plus								
			the								
			novella								
			“Back-								
			ward								
			Glance”								
			(1991).								
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								
			“Mid-								
			night								
			Con-								
			fes-								
			sions”								
			(2010)								
			—								
			plus								
			the								
			novella								
			“Back-								
			ward								
			Glance”								
			(1991).								
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								
			“Mid-								
			night								
			Con-								
			fes-								
			sions”								
			(2010)								
			—								
			plus								
			the								
			novella								
			“Back-								
			ward								
			Glance”								
			(1991).								
OK	Copyright	3550BIS	Two	by Ron	0.00	Harle	0778316645 , 0778316645 ,	, , ,			
	(c)	THE	classic	Roby Carr		Mira	9780778316640 , 9780778316640 , 18 ,				
	2023	SEA-	holi-	Carr		9780778316640	Market Paperback , Mass				
	The	SON	day				Market Paperback , Paperback				
	New		sto-				Mass-Market Fiction, Paperback				
	York		ries				Mass-Market Fiction, 2014-11-30 ,				
	Times		—				2014-11-16 , 2014-11-15 , 2014-11-01 ,				
	Com-		“Un-				0 , 1 , 0 , 0 , 0 , 1 , 0 , 0				
	pany.		der								
	All		the								
	Rights		Christ-								
	Re-		mas								
	served.		Tree”								
			(2009)								
			and								



status	copyright	times.results	title	description	results	notes	links	books	reviews
OKCopyrigh	35593	(c) —,	2023 THAT'S	The DE-	New LI-	York CIOUS	Times Com-pany.	All Rights Re-served.	by Action 0.00 Abramowitz, Advice How-To and Miscellaneous, 9781419726552 , 9781419726552 , 2017-10-01 , 2017-09-16 , 1 , 0 , 0 , 1
OKCopyrigh	35593	(c) the	2023 Horse	The He	New Rode	York In	Times On:	Com- The	NA James NA 0.00 NANA 0684857340 , NA , https://www.nytimes.com/1998/10/18/books/preaching-to-the-converted.html, NA
OKCopyrigh	35593	(c) G.U.	2023 ,	The VOL.	New 5	York fu-	Times ture,	Com- is	by Hamazaki H.99 TONYOPOLL , , ,

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))

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))

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
## $ status      <chr> "OK", "OK", "OK", "OK", "OK", "OK", "OK", "OK", "~
## $ copyright   <chr> "Copyright (c) 2023 The New York Times Company. ~
## $ num_results <int> 35593, 35593, 35593, 35593, 35593, 35593, 35593, ~
## $ title       <chr> "\"I GIVE YOU MY BODY ...\"", "\"MOST BLESSED OF ~
## $ 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> "", "", "", "", "", "", "", "", "", "", "", "~
## $ price       <chr> "0.00", "0.00", "0.00", "0.00", "0.00", "0.00", "~
## $ age_group   <chr> "", "", "", "", "", "", "", "", "", "", "", "~
## $ publisher   <chr> "Dell", "Liveright", "Beacon", "Beacon", "Beacon"~
## $ isbn10      <chr> "0399178570", "0871404427", "0807006475", "080700~
## $ isbn13      <chr> "9780399178573", "9780871404428", "9780807006474"~
## $ rank        <int> 8, 16, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, ~
## $ list_name   <chr> "Advice How-To and Miscellaneous", "Hardcover Non~
## $ display_name <chr> "Advice, How-To & Miscellaneous", "Hardcover Nonf~
## $ published_date <chr> "2016-09-04", "2016-05-01", "2023-01-29", "2023-0~
## $ bestsellers_date <chr> "2016-08-20", "2016-04-16", "2023-01-14", "2023-0~
## $ weeks_on_list <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ rank_last_week <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ asterisk     <int> 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ dagger       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ book_review_link <chr> "", "", "", "", "", "", "", "", "", "", "", "~
## $ first_chapter_link <chr> "", "", "", "", "", "", "", "", "", "", "", "~
## $ sunday_review_link <chr> "", "", "", "", "", "", "", "", "", "", "", "~
```

```
## $ article_chapter_link <chr> "", "", "", "", "", "", "", "", "", "", "", "", "", ""~  
nyt_unnest[nyt_unnest == ""] <- NA
```

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