

# Web scraping with R rvest & RSelenium

Barret Schloerke  
Statistics PhD Candidate  
Purdue University

# About Me

- Purdue University - 3rd Year
  - Statistics PhD Candidate
  - Dr. William Cleveland and Dr. Ryan Hafen
  - [www.tessera.io](http://www.tessera.io)
- [metamarkets.com](http://metamarkets.com) - 1.5 years
  - Front end engineer - coffee script / node.js
- Iowa State University
  - B.S. in Computer Engineering
  - Research in statistical data visualization with R

# Web Scraping

- Main Goals
  - Parse information from a given website using R
  - Retrieve information from a non-“html & css only” website using R

# Software - R

# Software - R

- magrittr - forward-piping
  - [github.com/smbache/magrittr](https://github.com/smbache/magrittr)



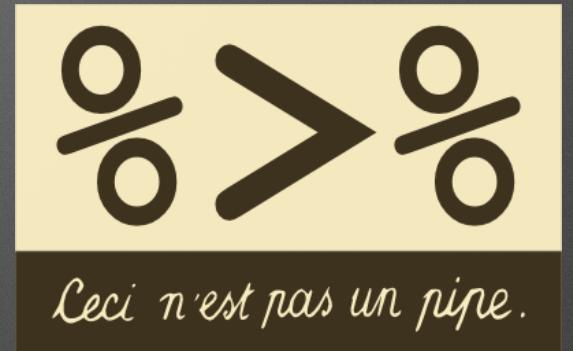
# Software - R

- magrittr - forward-piping
  - [github.com/smbache/magrittr](https://github.com/smbache/magrittr)
- rvest - web scraper
  - [github.com/hadley/rvest](https://github.com/hadley/rvest)



# Software - R

- magrittr - forward-piping
  - [github.com/smbache/magrittr](https://github.com/smbache/magrittr)
- rvest - web scraper
  - [github.com/hadley/rvest](https://github.com/hadley/rvest)
- RSelenium - headless browser
  - [github.com/ropensci/RSelenium](https://github.com/ropensci/RSelenium)



# magrittr

- “structuring sequences of data operations left-to-right (as opposed to from the inside and out),
- avoiding nested function calls,
- minimizing the need for local variables and function definitions, and
- making it easy to add steps anywhere in the sequence of operations.”

<https://github.com/smbache/magrittr>

# magrittr

- “structuring sequences of data operations left-to-right (as opposed to from the inside and out),”

# Children's Story Example

```
foo_foo <- little_bunny()
```

```
bop_on(  
  scoop_up(  
    hop_through(  
      foo_foo,  
      forest  
    ),  
    field_mouse  
  ),  
  head  
)
```

# Children's Story Example

```
foo_foo <- little_bunny()
```

```
bop_on(  
  scoop_up(  
    hop_through(  
      foo_foo,  
      forest  
    ),  
    field_mouse  
  ),  
  head  
)
```

```
foo_foo %>%  
  hop_through(forest) %>%  
  scoop_up(field_mouse) %>%  
  bop_on(head)
```

# rvest

- “It is designed to work with magrittr to make it easy to express common web scraping tasks, inspired by libraries like beautiful soup.”
- Installation
  - `install.packages("rvest")`
- Requirements
  - `url`
  - css selectors - [selectorgadget.com](http://selectorgadget.com)

# rvest - Common Functions

- `read_html()` - parse a url into html objects
- `html_nodes()` - select html objects with css selector
- `html_attr()` - retrieve html attributes from html objects
- `html_text()` - retrieve html object text

# rvest - Example

- [http://www.tripadvisor.com/Hotel\\_Review-g54448-d288616-Reviews-Spartanburg\\_Marriott-Spartanburg\\_South\\_Carolina.html](http://www.tripadvisor.com/Hotel_Review-g54448-d288616-Reviews-Spartanburg_Marriott-Spartanburg_South_Carolina.html)

Overview    Reviews (296)    Photos (57)    Location    Amenities    Q&A (2)    Room Tips (46)    Show Prices

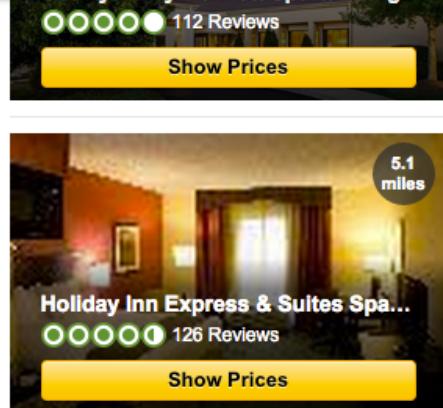
 KMS1642  
Columbia, South Carolina  
 2 reviews

**"Consistently on par .... EVERYTIME!"**  
 Reviewed 2 days ago NEW  
I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...  
More ▾  
Was this review helpful? Yes Report

 SCChad  
Florence, South Carolina  
**Level 3 Contributor**  
 12 reviews  
 3 hotel reviews  
 4 helpful votes

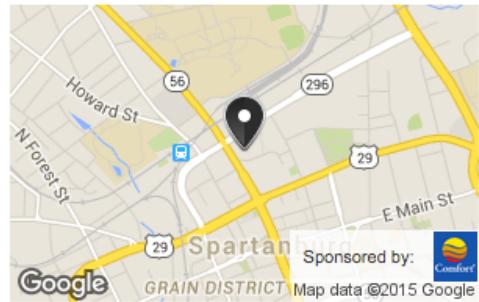
**"Nice Hotel"**  
 Reviewed 1 week ago  
I was pleasantly surprised by this hotel. It is an older hotel, but it has been kept up to date and clean. The room was modern, nice and clean, especially the bathroom. Looks like they've removed the refrigerators from the rooms on the lower levels. The lounge area is a nice place to relax or hang out, with a bar,...  
More ▾  
Was this review helpful? Yes Report

 Survey11237  
Indiana  
**Level 1 Contributor**  
**"Nice hotel"**  
 Reviewed 2 weeks ago  
Very nice. We stayed one night on our way to the beach. It's very nice, clean. Comfortable. The shower was amazing! A bit out of the way from the highway, but worth the extra few minutes drive. Would stay again if

  
112 Reviews    Show Prices  
5.1 miles  
Holiday Inn Express & Suites Spa...  
126 Reviews    Show Prices

See all Spartanburg hotels

Browse nearby  
Hotels (31) | Restaurants (288) | Things to Do (28)

  
Sponsored by:   
Map data ©2015 Google

Sponsored links \*

Explore Spartanburg  
Spartanburg Bed and Breakfast

# rvest - Example

- [http://www.tripadvisor.com/Hotel\\_Review-g54448-d288616-Reviews-Spartanburg\\_Marriott-Spartanburg\\_South\\_Carolina.html](http://www.tripadvisor.com/Hotel_Review-g54448-d288616-Reviews-Spartanburg_Marriott-Spartanburg_South_Carolina.html)

Overview    Reviews (296)    Photos (57)    Location    Amenities    Q&A (2)    Room Tips (46)    Show Prices

KMS1642  
Columbia, South Carolina  
2 reviews

**"Consistently on par .... EVERYTIME!"**  
Reviewed 2 days ago NEW

I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...

More ▾

Was this review helpful? Yes Report

SCChad  
Florence, South Carolina  
Level 3 Contributor  
12 reviews  
3 hotel reviews  
4 helpful votes

**"Nice Hotel"**  
Reviewed 1 week ago

I was pleasantly surprised by this hotel. It is an older hotel, but it has been kept up to date and clean. The room was modern, nice and clean, especially the bathroom. Looks like they've removed the refrigerators from the rooms on the lower levels. The lounge area is a nice place to relax or hang out, with a bar,...

More ▾

Was this review helpful? Yes Report

Survey11237  
Indiana  
Level 1 Contributor

**"Nice hotel"**  
Reviewed 2 weeks ago

Very nice. We stayed one night on our way to the beach. It's very nice, clean. Comfortable. The shower was amazing! A bit out of the way from the highway, but worth the extra few minutes drive. Would stay again if

112 Reviews  
Show Prices

5.1 miles  
Holiday Inn Express & Suites Spa...  
126 Reviews  
Show Prices

See all Spartanburg hotels

Browse nearby

Hotels (31) | Restaurants (288) | Things to Do (28)

Howard St  
N Forest St  
E Main St  
56  
296  
29  
Spartan  
GRAIN DISTRICT

Sponsored by: Google Map data ©2015 Google

Sponsored links \*

Explore Spartanburg

Spartanburg Bed and Breakfast

# rvest - Example

# rvest - Example

- For each review

*“Consistently on par .... EVERYTIME!”*

5 stars Reviewed 2 days ago NEW

I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...

[More ▾](#)

Was this review helpful? [Yes](#)

[Report](#)

# rvest - Example

- For each review
  - ID

*“Consistently on par .... EVERYTIME!”*

5 stars Reviewed 2 days ago NEW

I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...

[More ▾](#)

Was this review helpful?

[Report](#)

# rvest - Example

- For each review
  - ID
  - Quote

***“Consistently on par .... EVERYTIME!”***

Reviewed 2 days ago NEW

I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...

[More ▾](#)

Was this review helpful?  Yes [Report](#)

# rvest - Example

- For each review
  - ID
  - Quote
  - Rating

***“Consistently on par .... EVERYTIME!”***

5 stars Reviewed 2 days ago NEW

I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...

[More ▾](#)

Was this review helpful?  Yes [Report](#)

# rvest - Example

- For each review
  - ID
  - Quote
  - Rating
  - Review

*“Consistently on par .... EVERYTIME!”*

Reviewed 2 days ago NEW

I'm a frequent traveler and Platinum with Marriott hotels. I cover the entire state of SC for work, and this is by far my favorite hotel! What separates hotels in my opinion is the staff. You won't find better people than the Spartanburg Marriott! Every since my first stay at this hotel with ZERO status, they have consistently treated me...

More ▾

Was this review helpful? Yes

Report

# rvest - Example

```
library(rvest)

url <- "http://www.tripadvisor.com/Hotel_Review-
g54448-d288616-Reviews-Spartanburg_Marriott-
Spartanburg_South_Carolina.html"

# Grab each review
reviews <- url %>%
  read_html() %>%
  html_nodes("#REVIEWS .innerBubble")
```

# rvest - Example

```
id <- reviews %>%
  html_node(".quote a") %>%
  html_attr("id")

quote <- reviews %>%
  html_node(".quote span") %>%
  html_text()

rating <- reviews %>%
  html_node(".rating .rating_s_fill") %>%
  html_attr("alt") %>%
  gsub(" of 5 stars", "", .) %>%
  as.integer()

review <- reviews %>%
  html_node(".entry .partial_entry") %>%
  html_text()
```

# rvest - Example

```
hotel <- data.frame(id, quote, rating, review,
stringsAsFactors = FALSE)
str(hotel)

###  

'data.frame': 10 obs. of  4 variables:  

 $ id    : chr  "rn317547642" "rn314739986" "rn314556052" "rn313992798" ...  

 $ quote : chr  "Consistently on par .... EVERYTIME!" "Nice Hotel" "Nice hotel"  

 "Wofford visit" ...  

 $ rating: int  5 4 4 3 4 3 3 5 3 5  

 $ review: chr  "\nI'm a frequent traveler and Platinum with Marriott hotels. I  

cover the entire state of SC for work, and this is by far my fav"| __truncated__  

"\nI was pleasantly surprised by this hotel. It is an older hotel, but it has  

been kept up to date and clean. The room was moder"| __truncated__ "\nVery nice.  

We stayed one night on our way to the beach. It's very nice, clean. Comfortable.  

The shower was amazing! A bit out"| __truncated__ "\nWe've been to this Marriott  

several time and really enjoy the southern hospitality. This visit we were placed  

in room 909 and"| __truncated__ ...  

###
```

rvest

# rvest

- Pros
  - **very** small and efficient interface
  - one moving part

# rvest

- Pros
  - **very** small and efficient interface
  - one moving part
- Cons
  - only reads plain text
  - does not execute javascript or flash
    - ex: autonation.com car information will fail

# PhantomJS



- <http://phantomjs.org/>
- “PhantomJS is a headless WebKit scriptable with a JavaScript API. It has fast and native support for various web standards: DOM handling, CSS selector, JSON, Canvas, and SVG.”
- Executes javascript and flash
  - renders [autonation.com](http://autonation.com)

# PhantomJS - Installation

- <http://phantomjs.org/download.html>
- Mac
  - brew install node  
npm install -g phantomjs
  - building through node.js grabs the latest published version and places in correct path
  - much easier to debug installation
- (All examples done using 1.9.8)

# RSelenium

- “R Bindings for Selenium 2.0 Remote WebDriver”
- `browserName = “phantomjs”`
- `vignette(“RSelenium-headless”)`
- Retrieves web content and interacts with “browser”

# AutoNation Example

# AutoNation Example

FIND A STORE | CAREERS | CLICK TO CHAT | LAFAYETTE 47909 | ESPAÑOL | SIGN IN/REGISTER | MY ACCOUNT

AutoNation 

Keyword Search 

FIND A VEHICLE FINANCING VALUE YOUR VEHICLE SPECIAL OFFERS SERVICE & REPAIRS FIND A STORE

 What is the AutoNation Price? 

Home / Search Results  

**UPDATE RESULTS**  
78 Results Found  
Within 200 miles  
of ZIP 20874  
SEARCH AGAIN  
COMPARE UP TO 4 VEHICLES

**NARROW RESULTS**  
**Condition**  
 New  
 Used  
 Certified Pre-owned  
**Mileage**  
No Min  
No Max  
**Year**  
 2012 (1)

**Search Results** Sort by Distance - Nearest

<p>2005 Honda Element 4WD EX AT VIN: 5J6YH28655L018252 <b>IN STOCK</b>  <b>AutoNation Price</b>  \$9,293 <b>Contact Us</b> <input type="checkbox"/> Compare <input type="checkbox"/> Save</p>	<p>2006 Nissan Pathfinder SE 4WD VIN: 5N1AR18W26C662156 <b>IN STOCK</b>  <b>AutoNation Price</b>  \$9,480 <b>Contact Us</b> <input type="checkbox"/> Compare <input type="checkbox"/> Save</p>	<p>2005 Honda CR-V 4WD EX AT SE VIN: SHSRD78925U315469 <b>IN STOCK</b>  <b>AutoNation Price</b>  \$7,919 <b>Contact Us</b> <input type="checkbox"/> Compare <input type="checkbox"/> Save</p>
<p>2005 BMW X3 X3 4dr AWD 3.0i</p>	<p>2007 Volkswagen Jetta Sedan 4dr Auto</p>	<p>2004 Honda Accord Sdn EX Manual</p>



# AutoNation Example

[FIND A STORE](#) | [CAREERS](#) | [CLICK TO CHAT](#) | [LAFAYETTE 47909](#) | [ESPAÑOL](#) | [SIGN IN/REGISTER](#) | [MY ACCOUNT](#)

## AutoNation

Keyword Search 

[FIND A VEHICLE](#) [FINANCING](#) [VALUE YOUR VEHICLE](#) [SPECIAL OFFERS](#) [SERVICE & REPAIRS](#)

[FIND A STORE](#)

 What is the AutoNation Price?

Learn More 

[Home](#) / [Search Results](#)

[UPDATE RESULTS](#)

78 Results Found

Within 200 miles

of ZIP 20874

SEARCH AGAIN

COMPARE UP TO 4 VEHICLES

Search Results

Sort by

Distance - Nearest

**2005 Honda Element  
4WD EX AT**

VIN: 5J6YH28655L018252

IN STOCK 



AutoNation Price  \$9,293

Contact Us

Compare

Save

**2006 Nissan Pathfinder  
SE 4WD**

VIN: 5N1AR18W26C662156

IN STOCK 



AutoNation Price  \$9,480

Contact Us

Compare

Save

**2005 Honda CR-V 4WD  
EX AT SE**

VIN: SHSRD78925U315469

IN STOCK 



AutoNation Price  \$7,919

Contact Us

Compare

Save

NARROW RESULTS

Condition

New

Used

Certified Pre-owned

Mileage

No Min

No Max

Year

2012 (1)

**2005 BMW X3 X3 4dr  
AWD 3.0i**

**2007 Volkswagen Jetta  
Sedan 4dr Auto**

**2004 Honda Accord  
Sdn EX Manual**

 Chat Now

# AutoNation Example

# AutoNation Example

- For each search result

**2005 Honda Element  
4WD EX AT**

VIN: 5J6YH28655L018252

**IN STOCK** 



(855) 345-2224

**AutoNation Price**  \$9,293

# AutoNation Example

- For each search result
  - Name

2005 Honda Element  
4WD EX AT

VIN: 5J6YH28655L018252

IN STOCK 



(855) 345-2224

AutoNation Price 

\$9,293

# AutoNation Example

- For each search result
  - Name
  - VIN

2005 Honda Element  
4WD EX AT

VIN 5J6YH28655L018252

IN STOCK 



(855) 345-2224

AutoNation Price  \$9,293

# AutoNation Example

- For each search result
  - Name
  - VIN
  - Price

2005 Honda Element  
4WD EX AT

VIN 5J6YH28655L018252

IN STOCK ?



(855) 345-2224

AutoNation Price ? \$9,293

# AutoNation Example

```
library(RSelenium)
library(rvest)
library(magrittr)
library(stringr)

autoNationUrl <- "http://www.autonation.com/washington-
baltimore/under-10000#/used-+---+---+---+---+---+---+
+-0,10000-20874,200-+/near/1000/1"

read_html(autoNationUrl) %>%
  html_nodes("ul.search_view.search_view_2.search_view_new ")
# {xml_nodeset (0)}

# will not immediately work with rvest!
# must process website first!
```

# AutoNation Example

```
# start phantomjs
pJS <- phantom()
Sys.sleep(5) # give the binary a moment

remDr <- remoteDriver(browserName = 'phantomjs')
remDr$open()

# navigate to page and let it load completely.
# will return when loaded
remDr$navigate(autoNationUrl)

# retrieve processed page source
autoNationPageSource <- remDr$getPageSource()[[1]]

# close the PhantomJS process
remDr$close()
pJS$stop()
```

# AutoNation Example

```
# post process information
searchResults <- read_html(autoNationPageSource) %>%
  html_nodes("ul.search_view.search_view_2.search_view_new ")

searchResults
# {xml_nodeset (1)}
# [1] <ul class="search_view search_view_2 search_view_new"
data-bind="template ...
```

# AutoNation Example

```
vinNumbers <- searchResults %>%
  html_nodes("div.prod_vin.txt11 span:nth-child(2)") %>%
  html_text()

carName <- searchResults %>%
  html_nodes("div.prox_sub22 a") %>%
  html_text()

carValue <- searchResults %>%
  html_nodes("table.table-pricing-grid td:nth-child(2) strong") %>%
  html_text() %>%
  str_trim() %>%
  str_replace("\\\\$", "") %>%
  str_replace(", ", "") %>%
  as.numeric()
```

# AutoNation Example

```
carDt <- data.frame(carName = carName, carValue =  
carValue, vin = vinNumbers, stringsAsFactors=FALSE)
```

```
head(carDt)
```

```
###
```

	carName	carValue	vin
1	2005 Honda Element 4WD EX AT	9293	5J6YH28655L018252
2	2006 Nissan Pathfinder SE 4WD	9480	5N1AR18W26C662156
3	2005 Honda CR-V 4WD EX AT SE	7919	SHSRD78925U315469
4	2005 BMW X3 X3 4dr AWD 3.0i	7943	WBXPA93475WD03527
5	2007 Volkswagen Jetta Sedan 4dr...	7480	3VWEG71K17M059220
6	2004 Honda Accord Sdn EX Manual	6751	1HGCM55684A069483

```
###
```

# Cookies

```
# navigate to page and let it load completely.  
remDr$navigate("http://www.html-kit.com/tools/cookietester/")  
  
# helper function to print cookie output from html page  
print_page_source <- function() {  
  pageSource <- remDr$getPageSource()[[1]]  
  mainArea <- read_html(pageSource) %>%  
    html_nodes("div.hk7 div")  
  first <- mainArea %>%  
    html_nodes("div") %>%  
    html_text()  
  second <- mainArea %>%  
    html_nodes("dl") %>%  
    html_text()  
  print(str_c(str_c(first, collapse = "\n"), str_c(second,  
collapse = "\n"), collapse = "\n"))  
}
```

# Cookies

```
clearCookies <- remDr$findElement(using = "css selector",
"form.inl input:nth-child(2)")
clearCookies$clickElement()
remDr$screenshot(file = "1-reset_cookies.png")
print_page_source()
# [1] "Number of cookies received: 0    \n"

submitButton <- remDr$findElement(using = "css selector",
"form.zp input:nth-child(5)")
submitButton$clickElement()
remDr$screenshot(file = "2-add_cookie.png")
print_page_source()
# [1] "Number of cookies received: 1    \n1. Cookie named
\"TestCookie_Name_201510125703\""
\nTestCookie_Value_155703\n"
```

# Cookies

Before Set

This page can be used to test your browser cookies related settings.

If you've configured your browser to reject cookies, you can request this page and check if the cookie gets rejected. If you've configured your browser page to make sure that the cookies are being sent and aren't getting blocked.

Number of cookies received: 0 [Refresh](#)

This page did not receive any cookies.

Cookie name:

Cookie value:

[Set Test Cookie](#)

After Set

This page can be used to test your browser cookies related settings.

If you've configured your browser to reject cookies, you can request this page and check if the cookie gets rejected. If you've configured your browser page to make sure that the cookies are being sent and aren't getting blocked.

Number of cookies received: 1 [Refresh](#)

1. Cookie named "TestCookie\_Name\_201510125703" [Remove Cookie](#)

Cookie name:

Cookie value:

[Set Test Cookie](#)

# https and interactivity

```
remDr$setWindowSize(1800, 1800)

# navigate to page and let it load completely. will return when loaded
remDr$navigate("https://jobs.autonation.com")
remDr$screenshot(file = "0-initial_load.png")

# transitions to accounting page
# open list
remDr$findElement(using = "css selector", "div.job-category h2")$clickElement()
remDr$screenshot(file = "1-open_link.png")

# go to accounting
remDr$findElement(
  using = "css selector",
  "div.dropdowns.job-category div ul li a"
)$clickElement()
remDr$screenshot(file = "2-transition.png")

pageSource <- remDr$getPageSource()[[1]]
```

# https - load

## AutoNation. Careers

ABOUT AUTONATION  
THIRD PARTY AGENCIES

AUTONATION.COM

NEWS

CREATE A PROFILE-STORES

CREATE A PROFILE-CORPORATE

Car Sales Jobs,  
Auto Service Careers,  
and Corporate  
Opportunities at  
AutoNation



Search Jobs

Keyword

City, State, or ZIP

Radius

Search Jobs by Category

Search Jobs by Location

Search Jobs by Group

### You have the drive. We have the vehicle.

AutoNation is proud to be America's largest auto retailer representing domestic, import and luxury brands, both on and off the web. Headquartered in Fort Lauderdale, FL, and a member of the S&P 500 (NYSE AN), AutoNation employs approximately 25,000 people at over 290 store locations, representing over 290 new vehicle franchises across 15 states. We are driven to be the best, and we're always looking for passionate, motivated professionals who share the same drive to join us. Perhaps you're one of them.

Let us Search

With just 1 click and a connection with LinkedIn, you can receive job listings that best match your experience.

Jobs  AutoNation

Sign up for Job Alerts

First Name

Last Name

Phone Number (Optional)

Career Resources

What makes a great Salesperson at AutoNation

Learn about being a Service Technician at AutoNation

Opportunities for current students and recent Auto Technician Graduates

Application Tips

# https - click

## AutoNation. Careers

ABOUT AUTONATION  
THIRD PARTY AGENCIES

AUTONATION.COM

NEWS

CREATE A PROFILE-STORES

CREATE A PROFILE-CORPORATE

Car Sales Jobs,  
Auto Service Careers,  
and Corporate  
Opportunities at  
AutoNation



Search Jobs

Keyword

City, State, or ZIP

Radius

Search Jobs by Category

- Accounting Jobs
- Administrative Jobs
- Auto Body Jobs
- Auto Detail Jobs
- Auto Sales Jobs

[view all](#)

Search Jobs by Location

Search Jobs by Group

### ve the drive. We have the vehicle.

is largest auto retailer representing domestic, import and luxury brands, both on and off the web. and a member of the S&P 500 (NYSE AN), AutoNation employs approximately 25,000 people at over 290 new vehicle franchises across 15 states. We are driven to be the best, and we're always looking for passionate, motivated professionals who share the same drive to join us. Perhaps you're one of them.

Let us Search

With just 1 click and a connection with LinkedIn, you can receive job listings that best match your experience.

[Jobs In AutoNation](#)

Sign up for Job Alerts

First Name

Last Name

Phone Number (Optional)

Career Resources

- What makes a great Salesperson at AutoNation
- Learn about being a Service Technician at AutoNation
- Opportunities for current students and recent Auto Technician Graduates
- Application Tips

# https - follow link

## AutoNation. Careers

ABOUT AUTONATION    AUTONATION.COM    NEWS    CREATE A PROFILE-STORES    CREATE A PROFILE-CORPORATE  
THIRD PARTY AGENCIES

### Accounting

AutoNation - America's largest automotive retailer is looking for qualified people to join our team. If you are looking for a career that will allow you the opportunity to create results and accomplish goals, join us with a job on our Accounting team. Our ability to perform at such a high level is directly related to the efforts of our talented team of professionals, and we're looking for more great people to join us in our Accounting team. If you have the Drive, We have the Vehicle. Learn more about a job with AutoNation today.



#### Search Jobs

Keyword  City, State, or ZIP  Radius

#### Filter

State

California (1)

Florida (1)

Virginia (1)

City

Store

Position Type

#### 3 Results found for Accounting

**Accounting Associate - Mercedes-Benz of Delray**  
Delray Beach, Florida 09/21/2015

**Accounting Associate Mercedes-Benz Stevens Creek**  
San Jose, California 09/23/2015

**Accounting Associate - Honda Dulles**  
Sterling, Virginia 10/09/2015

Share: [f](#) [t](#) [in](#) [g+](#) [envelope](#)

#### Sign up for Job Alerts

First Name

Last Name

Phone Number (Optional)

Email Address\*

Interested In\*  
Search for a category, location, or category/location pair, select a term from the suggestions, and click "Add".

Category

Type to search

Location

# https - follow link

## AutoNation. Careers

ABOUT AUTONATION    AUTONATION.COM    NEWS    CREATE A PROFILE-STORES    CREATE A PROFILE-CORPORATE  
THIRD PARTY AGENCIES

### Accounting

AutoNation - America's largest automotive retailer is looking for qualified people to join our team. If you are looking for a career that will allow you the opportunity to create results and accomplish goals, join us with a job on our Accounting team. Our ability to perform at such a high level is directly related to the efforts of our talented team of professionals, and we're looking for more great people to join us in our Accounting team. If you have the Drive, We have the Vehicle. Learn more about a job with AutoNation today.



#### Search Jobs

Keyword  City, State, or ZIP  Radius

#### Filter

State

California (1)

Florida (1)

Virginia (1)

City

Store

Position Type

3 Results found for Accounting

Accounting Associate - Mercedes-Benz of Delray

Delray Beach, Florida 09/21/2015

Accounting Associate Mercedes-Benz Stevens Creek

San Jose, California 09/23/2015

Accounting Associate - Honda Dulles

Sterling, Virginia 10/09/2015

Share: [f](#) [t](#) [in](#) [g+](#) [envelope](#)

#### Sign up for Job Alerts

First Name

Last Name

Phone Number (Optional)

Email Address\*

Interested In\*  
Search for a category, location, or category/location pair, select a term from the suggestions, and click "Add".

Category

Type to search

Location

# https - Results

```
jobs <- read_html(pageSource) %>%  
  html_nodes("section#search-results-list ul li a")
```

```
jobTitle <- jobs %>%  
  html_nodes("h2") %>%  
  html_text()
```

```
jobLocation <- jobs %>%  
  html_nodes("span.job-location") %>%  
  html_text()
```

```
jobDatePosted <- jobs %>%  
  html_nodes("span.job-date-posted") %>%  
  html_text()
```

# https - Results

# Recap

- rvest
  - use as much as possible!
  - small interface
- RSelenium
  - big and clunky, but much more powerful
  - parses full websites (javascript and flash)

# Questions?