Executive Summary

In my Corporate Finance class our professor stressed the need to stay up-to-date on current events. I have heard counsel from other professors many times that we should read the Wall Street Journal or consistently read business and world news. As a poor college student, I cannot afford to pay for a subscription to the Wall Street Journal, nor do I have the time to sift through pages of articles about world topics. Sometimes I just forget to read the news because I am so busy with other course work.

The "Current Events!" Excel workbook is the perfect solution to this problem. The program will send you an email every day containing the main headlines from the Wall Street Journal and the Economist. The program routes through the BYU library website to access the free version of the Wall Street Journal and the Economist. It copies the article headlines and links then it sends them in an email to the email address of your choosing. If you see an article that interests you then you can click the attached link to read the entire article for yourself. The program is automated to run at 5:00pm every day, so I will never miss a day of current events.

The "Current Events!" workbook solves the problem of cost, because it gets news headlines from BYU channels completely free of charge. It saves me time, because I do not have to go out and look for the news, the news comes right to me. Finally, on the days where I forget about current events, I will receive a nice little email at 5:00pm to remind me of what is going on in the world.

I'll never miss out on a day of current events again!

Implementation Documentation

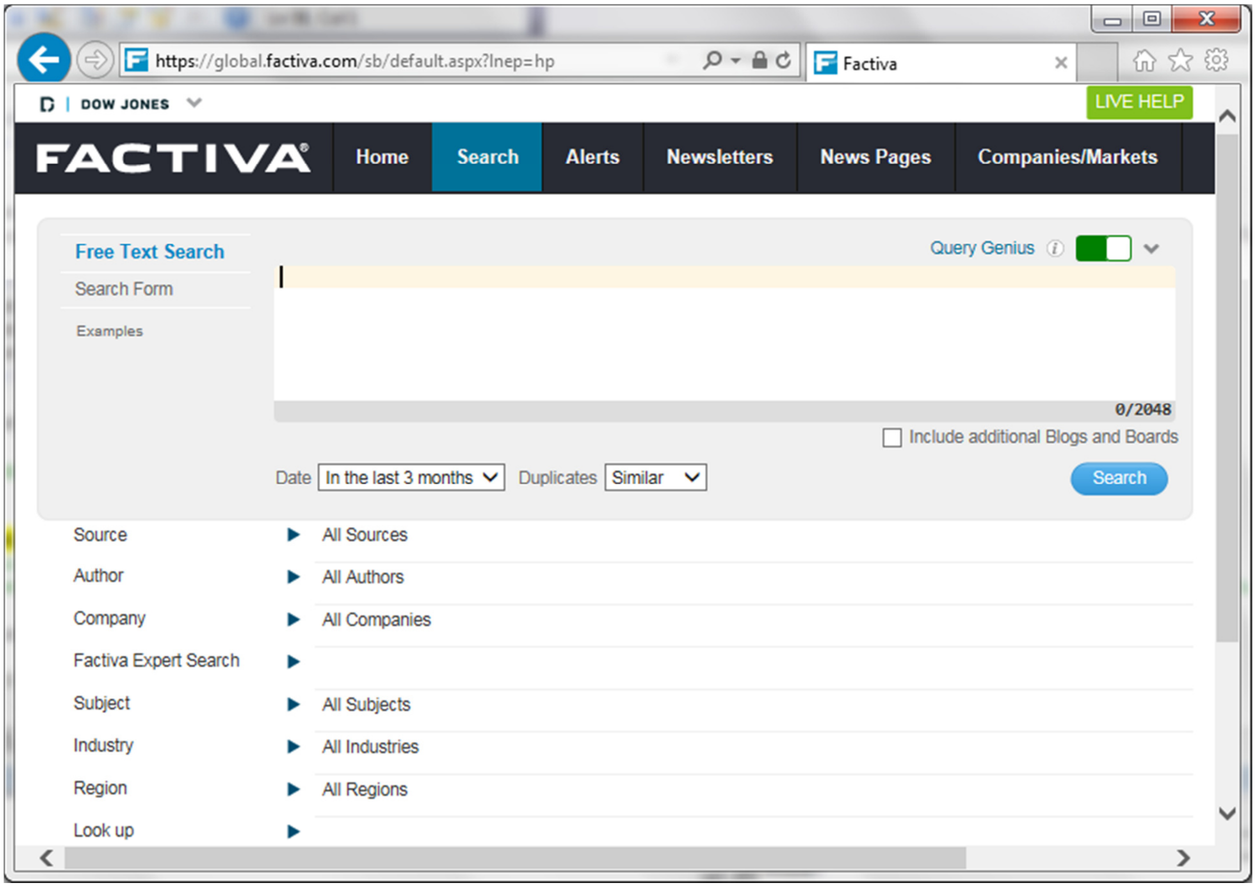
My sub procedures are separated into three different categories: (1) get the news from the Wall Street Journal, (2) get the news from the Economist, and (3) get the news by geographical region. The following paragraphs describe each process in detail.

Get News from the Wall Street Journal

To get news from the Wall Street Journal I wrote two sub procedures: WSJlogin and sendWSJ. These procedures also use Gove Allen's sendGmail function.

WSJLogin

This sub procedure uses BYU's Factiva account to get WSJ headlines. It opens Internet Explorer and navigates to the Factiva website. This website routes you to a BYU website prompting the user to enter their net ID and password. The sub procedure enters the username and password (from cell "B1" and "B2" of the excel sheet) into the appropriate boxes on the website and submits the information. This routes the user to the Factiva page which looks like this:

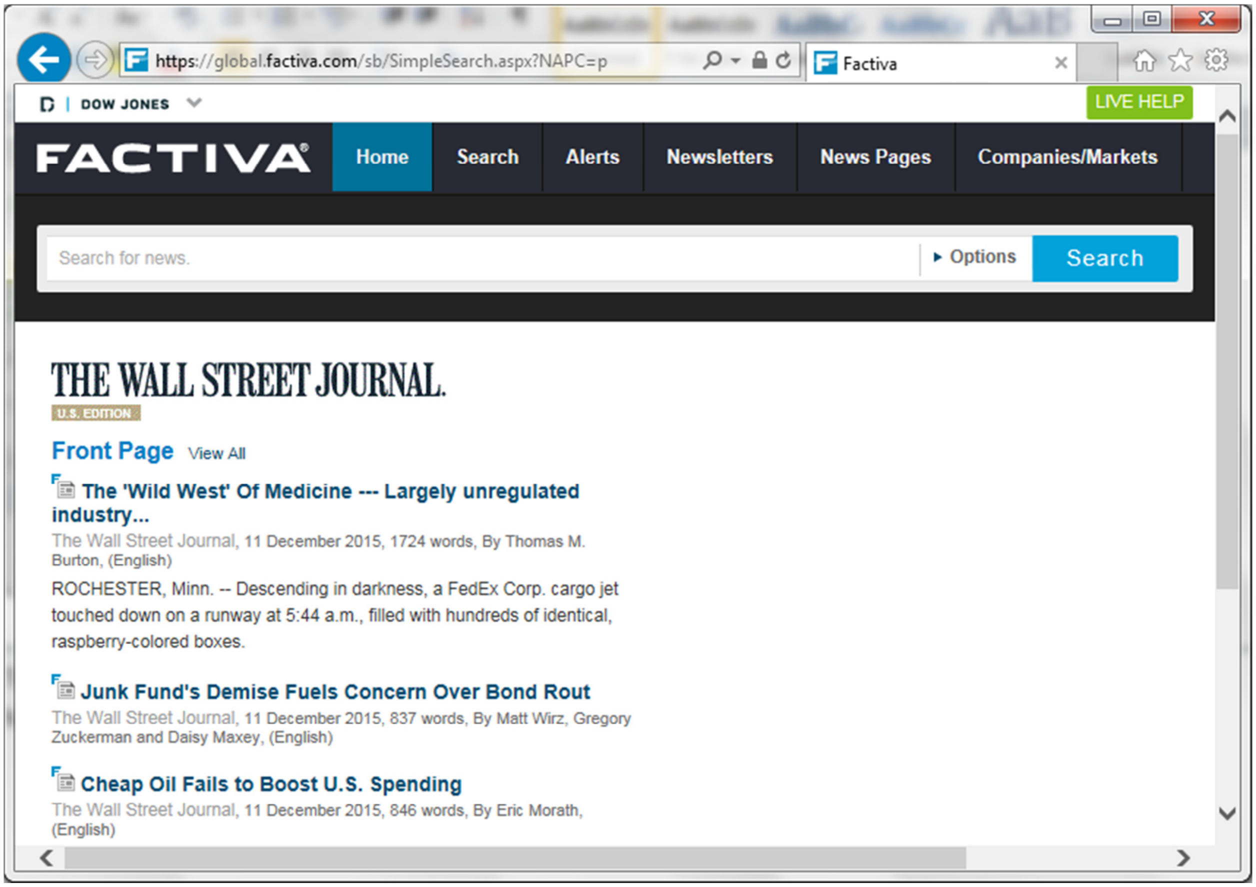


□ Include additional Blogs and Boards

Date: In the last 3 months
Duplicates: Similar

Source: All Sources
Author: All Authors
Company: All Companies
Subject: All Subjects
Industry: All Industries
Region: All Regions

To sub procedure then follows the “Home” link in the top left part of the page in order to view the WSJ headlines for current day. After following the link the page will look like this:



Front Page View All

The 'Wild West' Of Medicine --- Largely unregulated industry...
The Wall Street Journal, 11 December 2015, 1724 words, By Thomas M. Burton, (English)
ROCHESTER, Minn. -- Descending in darkness, a FedEx Corp. cargo jet touched down on a runway at 5:44 a.m., filled with hundreds of identical, raspberry-colored boxes.

Junk Fund's Demise Fuels Concern Over Bond Rout
The Wall Street Journal, 11 December 2015, 837 words, By Matt Wirz, Gregory Zuckerman and Daisy Maxey, (English)

Cheap Oil Fails to Boost U.S. Spending
The Wall Street Journal, 11 December 2015, 846 words, By Eric Morath, (English)

The subprocedure will then read the HTML on the web page to find the headlines and the links that are on the page. First, the position on the page is set to 1 (a.position = 1), so the reader can always start from the beginning of the page. Then, using a For Loop, the reader will move to the part of HTML that contains the headlines. After it gets to that point the reader will read the HTML until it gets to an href tag. The text after the href contains the link to the article so the reader copies that text and enters as the first element of an array entitled "L". The "L" array contains all of the links for the Wall Street journal. The reader then will move to the close bracket character ">", and it will read the text until the next open bracket character "<". This will read the article headline, and the sub procedure will enter this as the first element in an array entitled "HL". This array contains all of the headlines for the WSJ.

The WSJ headlines page always contains five links, so the sub procedure will perform the above task (getting the text for the link and headline then entering into their respective arrays) a total of five times. Now we have two arrays: "HL(4)", which contains the five headlines and "L(4)", which contains the five links. The sub procedure will now call the sendWSJ sub procedure to send the links via email.

sendWSJ

sendWSJ calls upon Gove Allen's sendGmail function to send an email. This function accepts values for the following variables: sendto, from, pw, subject, and body. We set values for these variables here in the sendWSJ sub procedure. Sendto, from, and pw all come from the "news" tab in excel. This enters my username and password (from cells "B3" and "B4") as the email address that will send the email. Then it will enter a destination email (cell "B5"). The destination email is currently set as my own email, so I am essentially sending an email to myself. The subject variable is set to "WSJ News Headlines".

The body of the email is more complex. The sendWSJ sub procedure organizes the headlines and links into a nice little table. It does this using a For Loop. For each element in the array the For Loop inserts the contents of "HL(4)" (the headlines) and "L(4)" (the links) in between the HTML tags for a table. Each time the loop iterates it performs the following:

1. Forms a new table row and cell by typing the tags "<tr> <td>"
2. Inserts the headline from "HL(4)"
3. Closes the first cell and forms a new cell with the text "</td> <td>"
4. Inserts the link from "L(4)"
5. Closes the table cell and the row using the text "</td> </tr>"

This loop iterates 5 times and therefore creates 5 table rows with 2 cells in each row. Finally, the code for the table rows and cells is then inserted into the beginning and ending tags to for a table: <table border = "1" style = "width:100%"> and </table>. This entire string of HTML text is assigned to the body variable.

The sendto, from, pw, subject, and body variables are then passed into the sendGmail function.

SendGmail

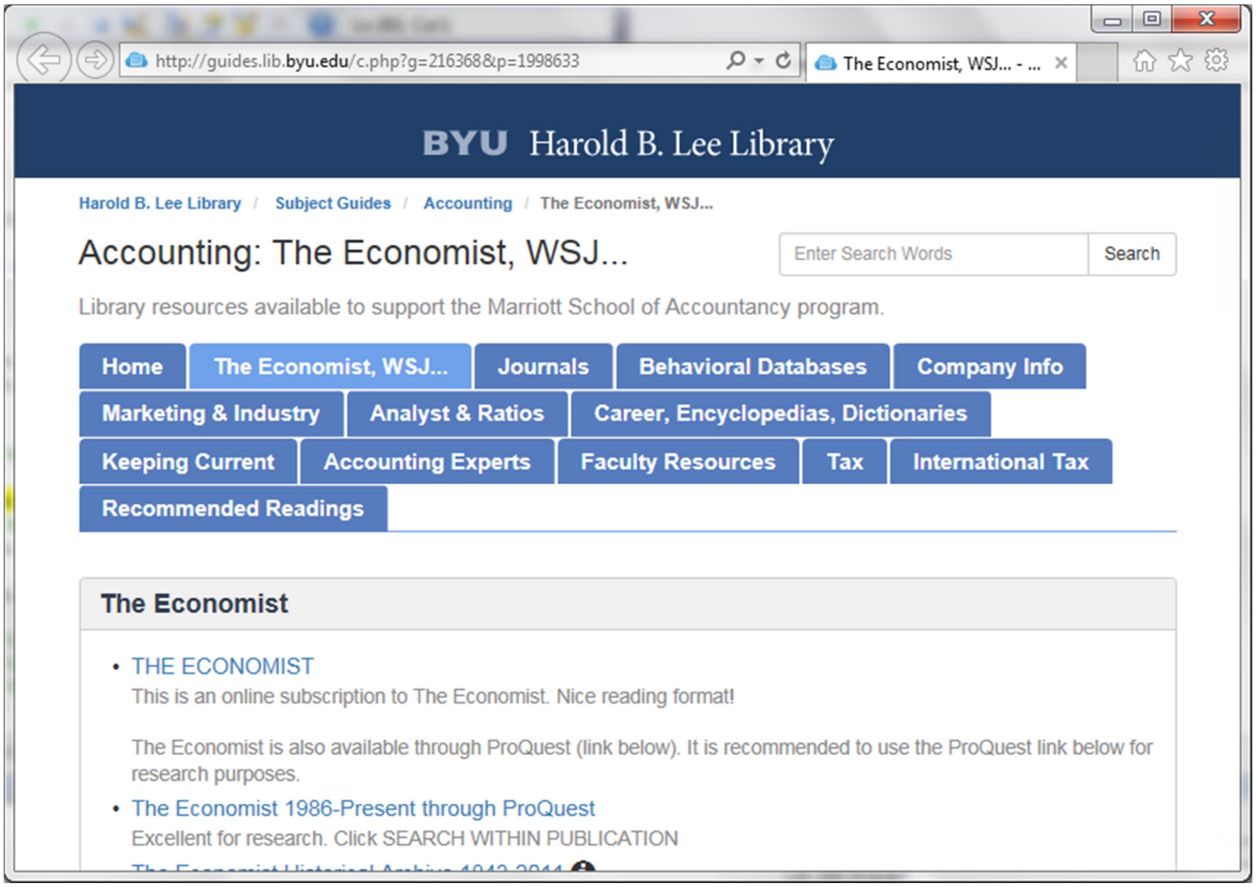
This function was written by Gove Allen. It configures the system to communicate with the Google server, and then it enters the appropriate values into each part of a Google email (Gmail). The function uses iMsg.HTMLbody for the body of the email, so the email can be send using the HTML code rather than the using only text. The end result is an email containing a neat table of WSJ headlines and links for the reader's enjoyment!

Get News from the Economist

To get news from the Economist I wrote two sub procedures: EconomistLogin, sendEcon. These procedures also use Gove Allen's sendGmail function.

EconomistLogin

This sub procedure accesses the Economist magazine that is available through the BYU library website. It first opens Internet Explorer and follows the URL for the BYU library page which links to the Economist. The page looks like this:



Accounting: The Economist, WSJ...

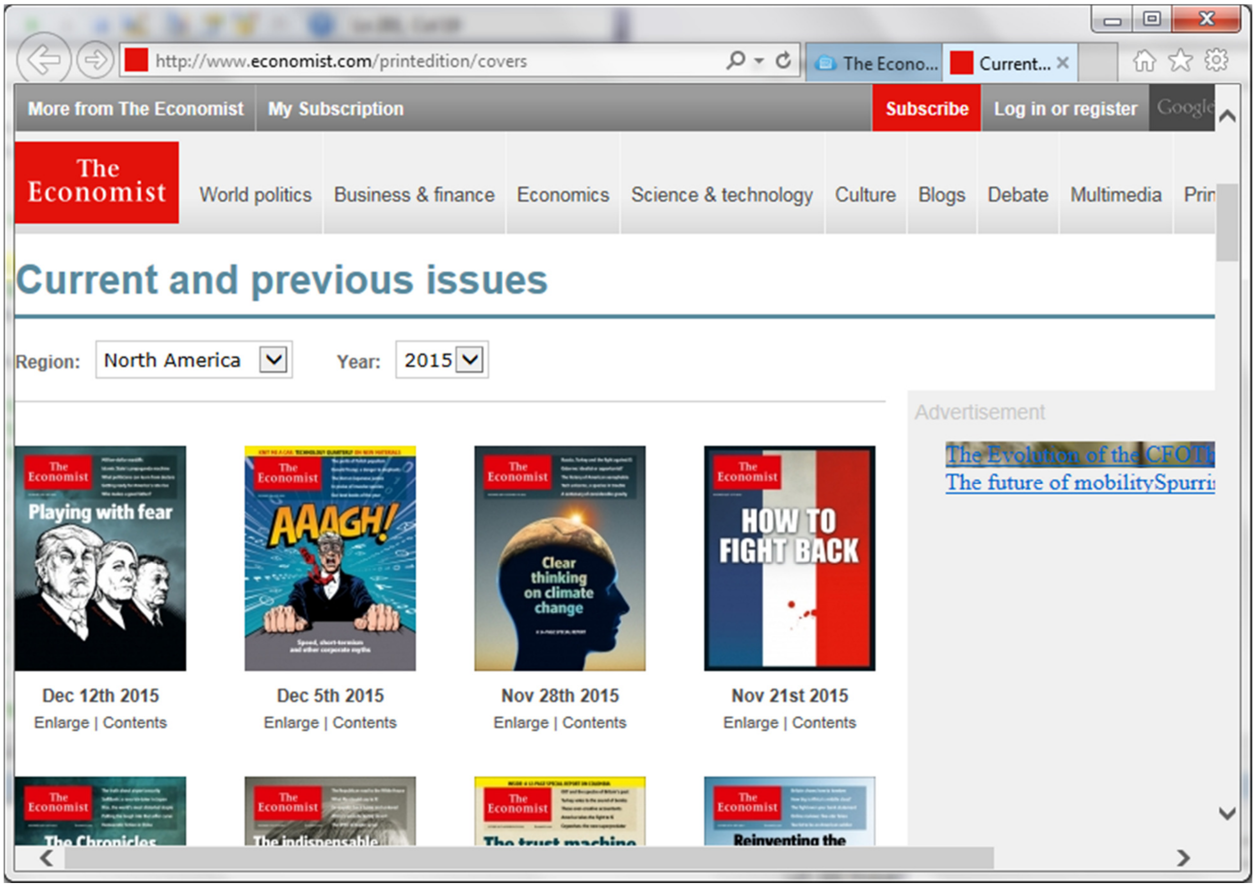
Library resources available to support the Marriott School of Accountancy program.

The Economist

- THE ECONOMIST
  This is an online subscription to The Economist. Nice reading format!

  The Economist is also available through ProQuest (link below). It is recommended to use the ProQuest link below for research purposes.
- The Economist 1986-Present through ProQuest
  Excellent for research. Click SEARCH WITHIN PUBLICATION
- The Economist Historical Archive 1843-2011

It then follows the Economist link using the agent's followlinkbytext function. It looks for the text "THE ECONOMIST" and follows that link. Upon following that link the following page appears in a new window:



Current and previous issues

Region: North America

Year: 2015

Dec 12th 2015
Enlarge | Contents

Dec 5th 2015
Enlarge | Contents

Nov 28th 2015
Enlarge | Contents

Nov 21st 2015
Enlarge | Contents

This window displays the current and previous issues of the Economist that are available. Since this web page opened in a new window, the agent is still attached to the previous window (the BYU library page). The sub procedure then attaches the agent to the new window (the Economist page) using the method "attach". This method attaches to the new page based on its URL. Now attached to the new page, the agent reads through the HTML to find the words "Politics this week". This section contains an overview of all the big news in politics this week. Once the reader has found "Politics this week" it moves back to the "href=" and then gets the text until the close of the bracket ">". This picks up the link of the "Politics this week" page and adds it as the first element in the "econL(1)" array. The econL(1) array contains the links for the economist. The reader then gets the text from the close bracket ">" to the next open bracket "<" which picks up the text "Politics this week", which is the name of the article. Then the sub procedure inserts the name of the headline into "econHL(1)" array, which contains all of the headlines for the economist.

The sub procedure then performs the same process for the "Business this week" section of the magazine. It pulls the link and the headline for this section and adds them to the "econL(1)" and "econHL(1)" arrays.

We now have two small arrays with the headlines and links from the economist. The sub procedure terminates both Internet Explorer windows, then it calls the sendEcon sub procedure to organize the email.

SendEcon

The sendEcon sub procedure acts in the same way as the sendWSJ sub procedure. It enters the appropriate values for the sendto, from, and pw variables. It enters the subject as “The Economist News Headlines”. It also enters the headlines and links into an HTML table using a For Loop. Then it passes these values into the sendGmail function to construct the email.

The sendGmail function then communicates with Google’s server to send an email containing the links for “Politics this week” and “Business this week”.

Time Delay

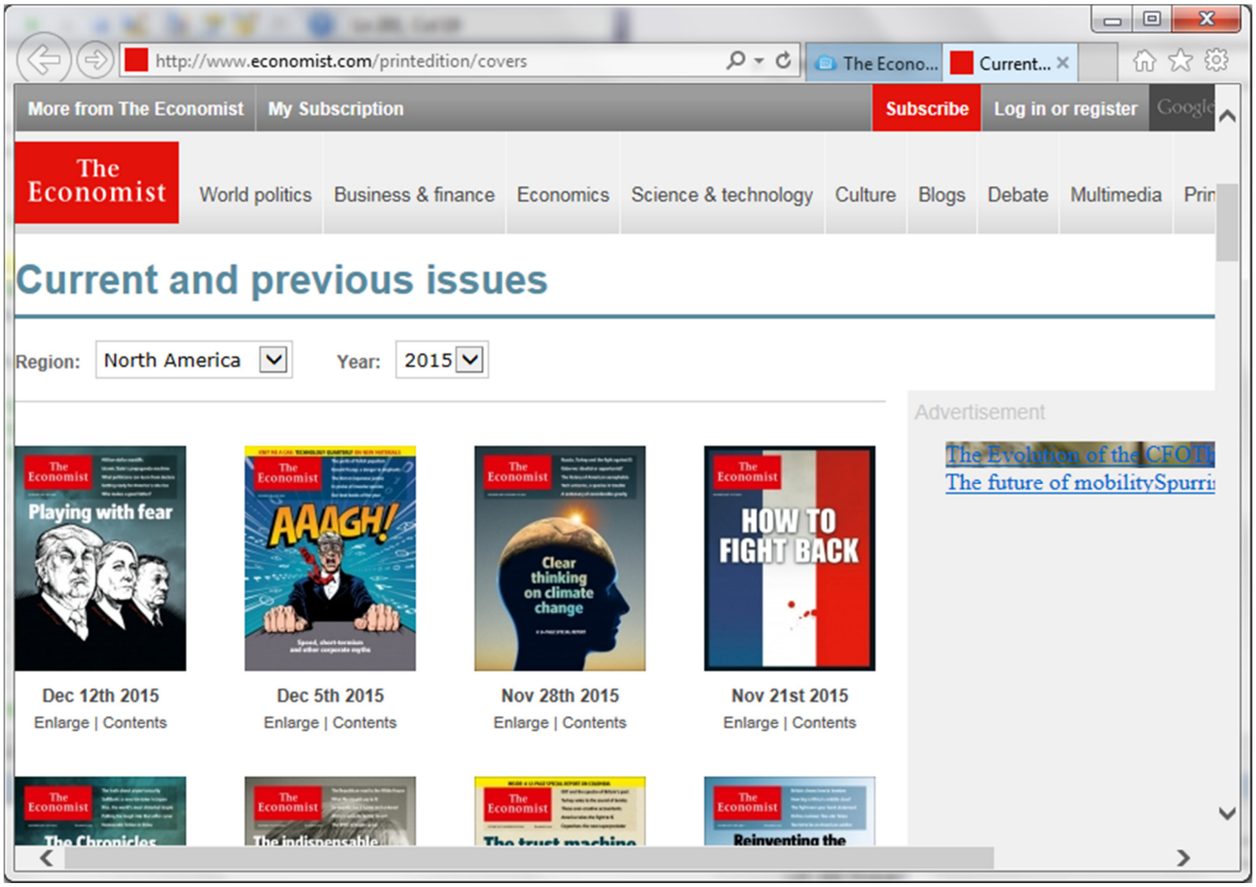
Since I want to receive news from the Wall Street Journal and the Economist every day I created a sub procedure to run the WSJ and Economist sub procedures every day. This procedure is called TimeOfDay and it is set to run the WSJ and Economist sub procedures every day at 5:00 pm.

Get News by Geographical Region

The sub procedures to sort by geographical region are all very similar. I will walk through the SortUnitedStates procedure, and this guidance can be applied to all of the other sub procedures that sort by geographical region.

SortUnitedStates

The beginning of this sub procedure is the same as the EconomistLogin procedure: the procedure opens Internet Explorer, navigates to the BYU library page, then clicks the Economist link. The page featuring current and previous articles of the Economist appears like this:



Current and previous issues

Region: North America

Year: 2015

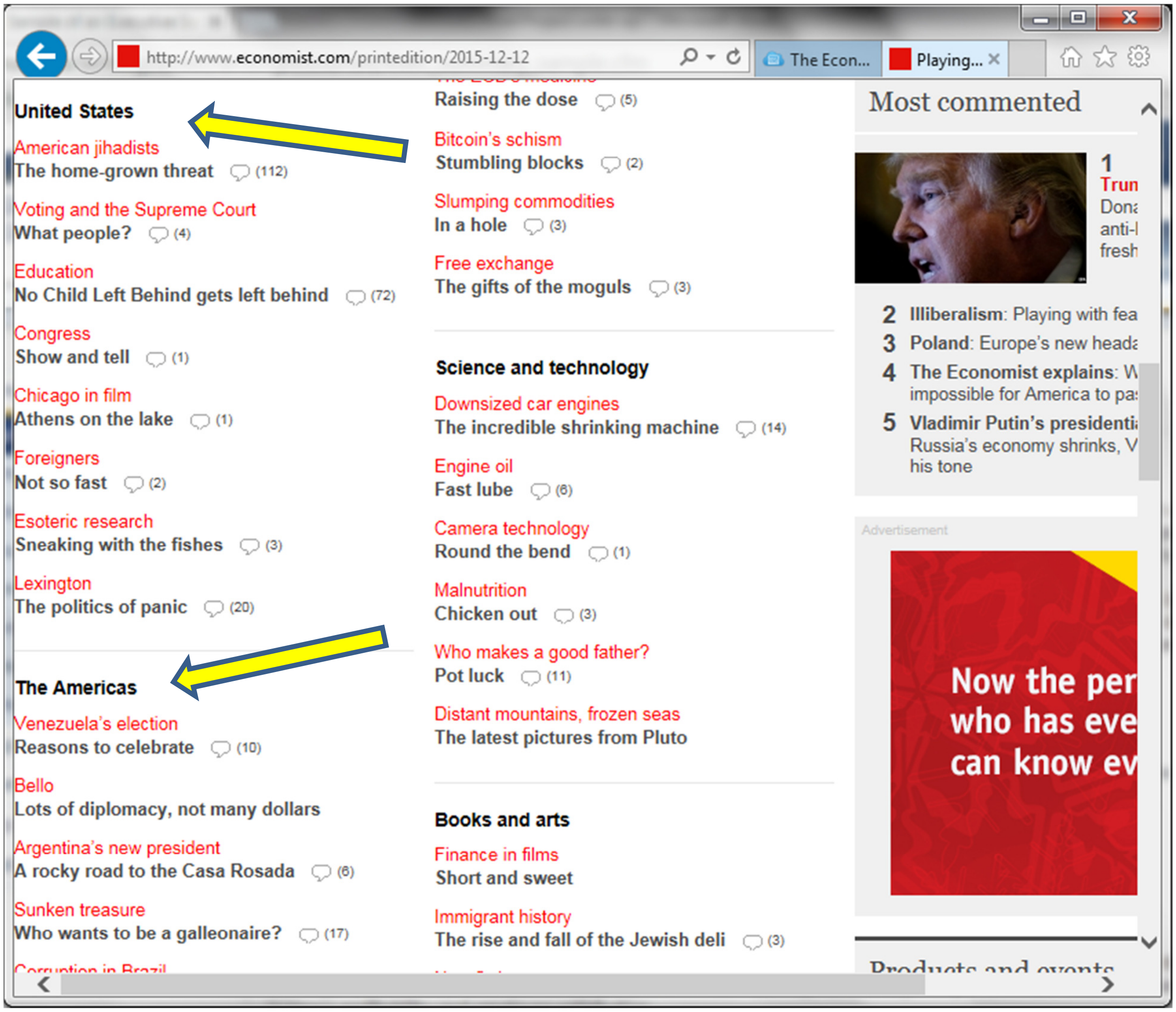
Dec 12th 2015
Enlarge | Contents

Dec 5th 2015
Enlarge | Contents

Nov 28th 2015
Enlarge | Contents

Nov 21st 2015
Enlarge | Contents

Now the sub procedure follows “Current issue” link using the followlinkbytext method. This will take the user to this week’s issue of the Economist. The Economist has articles according to 6 different world regions. The following screenshot illustrates how the United States and the Americas are separated into different sections.



United States

American jihadists
The home-grown threat (112)

Voting and the Supreme Court
What people? (4)

Education
No Child Left Behind gets left behind (72)

Congress
Show and tell (1)

Chicago in film
Athens on the lake (1)

Foreigners
Not so fast (2)

Esoteric research
Sneaking with the fishes (3)

Lexington
The politics of panic (20)

The Americas

Venezuela's election
Reasons to celebrate (10)

Bello
Lots of diplomacy, not many dollars

Argentina's new president
A rocky road to the Casa Rosada (6)

Sunken treasure
Who wants to be a galleonaire? (17)

Corruption in Brazil

Raising the dose (5)

Bitcoin's schism
Stumbling blocks (2)

Slumping commodities
In a hole (3)

Free exchange
The gifts of the moguls (3)

Science and technology

Downsized car engines
The incredible shrinking machine (14)

Engine oil
Fast lube (6)

Camera technology
Round the bend (1)

Malnutrition
Chicken out (3)

Who makes a good father?
Pot luck (11)

Distant mountains, frozen seas
The latest pictures from Pluto

Books and arts

Finance in films
Short and sweet

Immigrant history
The rise and fall of the Jewish deli (3)

Most commented

1 Trun
Doná
anti-I
fresh

2 Illiberalism: Playing with fea
3 Poland: Europe's new heada
4 The Economist explains: W
impossible for America to pa:
5 Vladimir Putin's presidenti
Russia's economy shrinks, V
his tone

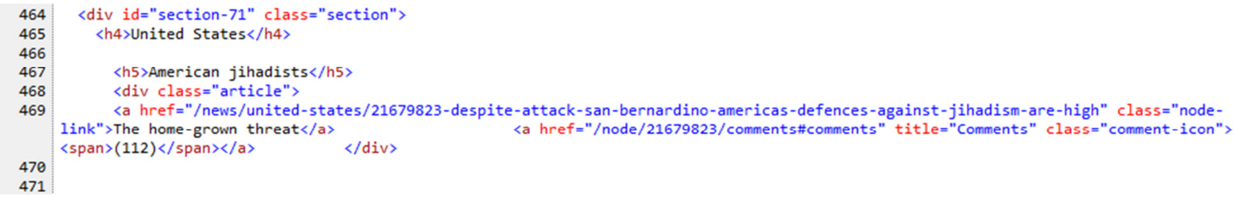
Advertisement

Now the per
who has eve
can know ev

Products and events

(If you scroll down the webpage articles are separated into the following additional regions: Asia and China, Middle East and Africa, and Europe.)

SortUnitedStates will get the headlines and the links for each of the articles under the United States heading. In the HTML code each regional section is under a different <div id>. For the United States the div id is equal to "section-71". Right below the div id there is a unique headline that is different for each region. It contains the region name (in this case "<h4>United States</h4>"). The graphic below displays the HTML for the United States' section. The first line of code shows the beginning of div section-71, and the second line of code displays the country's unique heading. The agent looks for the unique heading bracketed by "<h4>" and "</h4>".



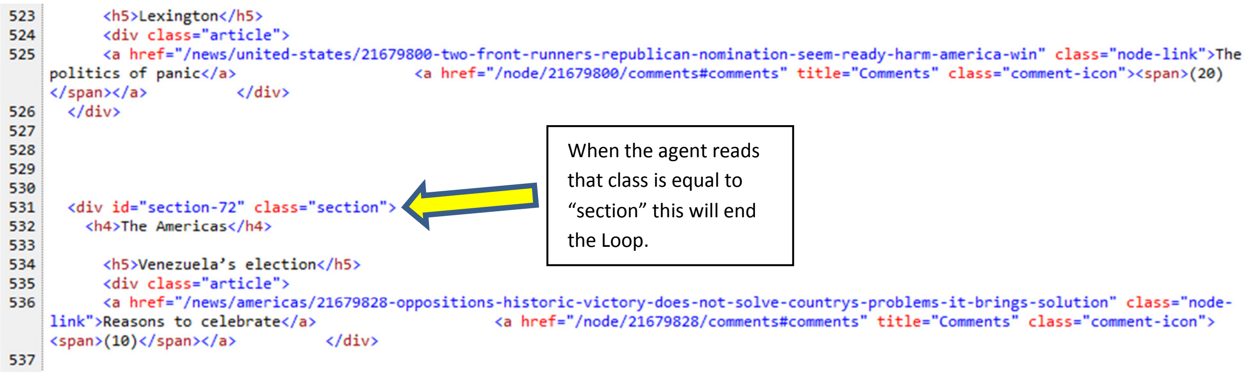
```
  <div id="section-71" class="section">
    <h4>United States</h4>

      <h5>American jihadists</h5>
      <div class="article">
      <a href="/news/united-states/21679823-despite-attack-san-bernardino-americas-defences-against-jihadism-are-high" class="node-
link">The home-grown threat</a>                    <a href="/node/21679823/comments#comments" title="Comments" class="comment-icon">
<span>(112)</span></a>          </div>


```

Now that the agent has found the United States' unique heading it can start looking for the links and the headlines for the articles. The text <div class="article"> is the flag that tells us there will be a new article. As such, the agent moves to the next iteration of the text "class=" then it reads the characters until the next close bracket ">". The procedure then sets the variable "check" equal to the string that the agent picks up. If check is equal to "article" then the reader will proceed to get the article headline and link. The reader will get the text for headline in between "<h5>" and "</h5>", the link after the href=, and the text in between the ">" and "</a>" characters. The procedure will then insert the headline into the USHL() array and the link into the USL() array.

For example, in the code above the reader will start from the line #465 reading "<h4>United States</h4>". It will then move to line #468 and set the check variable equal to the text after class= which is "article". It will now make sure that the check variable equals "article", which it does! Now that the check has passed it will get the headline and the link. It will move back to line #467 and set USHL(0) equal to "American jihadists". It will then set USL(0) equal the link after href= which is "/news/united-states/21679823-despite-attack-san-bernardino-americas-defences-against-jihadism-are-high". Then it will move to the closing bracket after the href element on line #469 to read "The home-grown threat". The procedure will finally add ": " and "The home-grown threat" to USHL(0) so now that element in the array reads "American jihadists: The home-grown threat". USHL(0) now contains the first headline under the United States section and USL(0) now contains the first link for the section.

This part of the code was written in a Do Loop, so the code will repeat these steps until it reaches a div class that does not equal "article". When the div class is equal to something else besides "article" the reader has moved out of the United States section, so the loop will end. This is how the procedure knows how many articles to record for the United States section.

The graphic below displays the final article for the United States section. The agent will get the link and the text "Lexington: politics of panic" then it will move onto the next iteration of "class=". It will read the text after "class=" which is "section". Since "section" does not equal "article" the agent will no longer continue to search for text, but will exit the Do Loop instead.



```
      <h5>Lexington</h5>
      <div class="article">
      <a href="/news/united-states/21679800-two-front-runners-republican-nomination-seem-ready-harm-america-win" class="node-link">The politics of panic</a>                    <a href="/node/21679800/comments#comments" title="Comments" class="comment-icon"><span>(20)</span></a>          </div>
  </div>




  <div id="section-72" class="section">
    <h4>The Americas</h4>

      <h5>Venezuela's election</h5>
      <div class="article">
      <a href="/news/americas/21679828-oppositions-historic-victory-does-not-solve-countrys-problems-it-brings-solution" class="node-link">Reasons to celebrate</a>                    <a href="/node/21679828/comments#comments" title="Comments" class="comment-icon"><span>(10)</span></a>          </div>

```

When the agent reads that class is equal to "section" this will end the Loop.

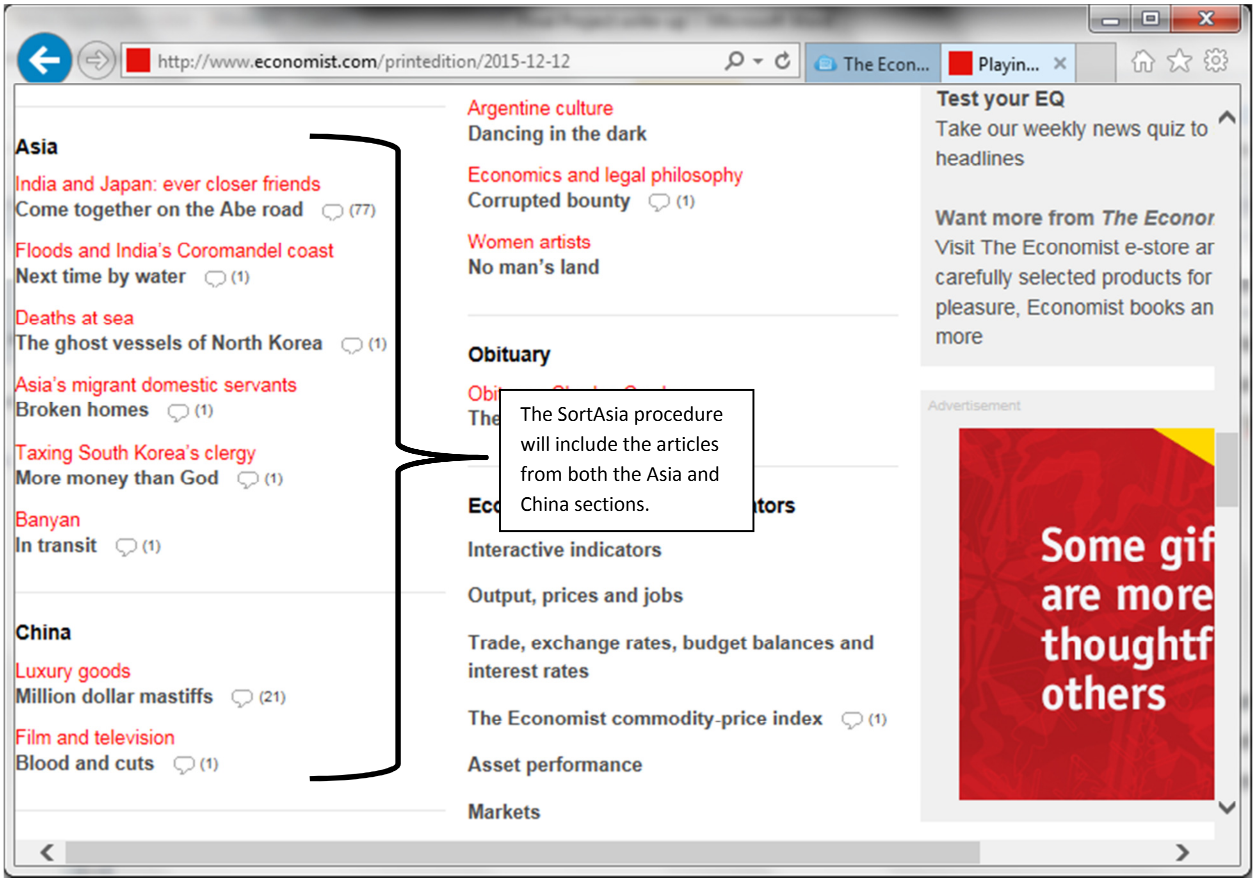
The sub procedure will then terminate Internet Explorer and set the body variable to send in the email. The Do Loop that was recording the links and the headlines has an "x" variable that counts the number of articles that have been inserted into the array. The sub procedure will insert that same number of article headlines and links into the HTML code to send the email. The HTML code creates a table out of the headlines and links and this code is set as the “body” variable. The subject variable is set as “The Economist News Headlines for United States.” Then the sub procedure inserts the values from the spreadsheet for the sendto, from, and pw varialbes.

These values are passed to the sendGmail function to be sent via email.

Sort by Other Locations

The other procedures that sort by geographical region (SortAmericas, SortAsia, SortMidEastandAfrica, and SortEurope) follow the same steps as the SortUnitedStates sub procedure, except they search for their respective regions using a different unique heading (e.g. sortAmericas will search for “<h4>The Americas</h4>” then it will being looking for headlines and links from there).

The SortAsia sub procedure includes two Do Loops instead of one because this sub procedure pulls articles from the “Asia” section and the “China” section on the Economist website (see graphic below). The sub procedure runs one Do Loop to grab all the articles from the “Asia” section. Then it runs another Do Loop to get the articles under the “China” section.



http://www.economist.com/printedition/2015-12-12

Asia

India and Japan: ever closer friends
Come together on the Abe road (77)

Floods and India’s Coromandel coast
Next time by water (1)

Deaths at sea
The ghost vessels of North Korea (1)

Asia’s migrant domestic servants
Broken homes (1)

Taxing South Korea’s clergy
More money than God (1)

Banyan
In transit (1)

China

Luxury goods
Million dollar mastiffs (21)

Film and television
Blood and cuts (1)

Argentine culture
Dancing in the dark

Economics and legal philosophy
Corrupted bounty (1)

Women artists
No man’s land

Obituary

Ec…**tors**

Interactive indicators

Output, prices and jobs

Trade, exchange rates, budget balances and interest rates

The Economist commodity-price index (1)

Asset performance

Markets

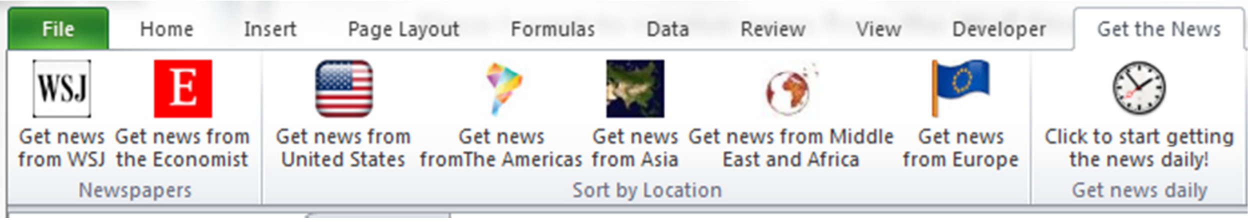
Test your EQ
Take our weekly news quiz to headlines

Want more from *The Econor*
Visit The Economist e-store ar carefully selected products for pleasure, Economist books an more

The SortAsia procedure will include the articles from both the Asia and China sections.

Ribbon Editing

The ribbon in excel has been formatted to run all of the above mentioned sub procedures. A user can click the “Get the News” tab to access the applicable buttons. The WSJ button runs the WSJlogin procedure, the Economist button runs the EconomistLogin procedure, and the different regional buttons run their respective sub procedures. Finally, the clock icon will run the TimeOfDay sub procedure to allow the user to get the news every day at 5:00 pm.



Learning and Conceptual Difficulties Encountered

I encountered a few difficulties in doing this project. The first was trying to click links on the web pages using the agent. One specific issue I had was clicking the link to get to current issues of the Economist. The links didn't have names, and it was hard to find something unique about the link I was trying to click on. At first we (Professor Allen and I) solved the problem by having the sub procedure search the whole document for the words "Current issue." It would find that tag and then click it to get to the current issue. After that we found out we could simply followlinkbytext on the words "Current issue".

A second big issue was trying to copy the links for WSJ articles. When the agent would get the link I would try to follow that link and it wouldn’t work. After scrutinizing the link I realized that the text I had copied from the source code of the website included extra characters that should not have been there. Specifically, when I used the gettext method of the agent the text would have the characters “amp;” three different times in the link. When I viewed the link in the source code those characters were not there, but when I copied the link the “amp;” characters would copy with it. After struggling with this for a while I wrote some code that would read the link until the semi-colon “;” then I would take out the three characters before the semi-colon which would delete the “amp”. Then it would read to the next semi-colon and do the same thing two more times. This would leave me in the end with a link that did not contain any “amp;” characters. This was a good solution since URLs do not usually contain semi-colons.

A third difficulty was getting the correct number of articles from the sub procedures that sorted by geographical by region. Each edition of the Economist had a different number of articles under each section. The United States section, for example, could contain anywhere from 4 to 8 articles depending on the magazine edition. I wanted the code to be able to grab all of the headlines no matter how many articles were in that section. This required a lot of trial and error, and I eventually landed on the solution that I have in my sub procedures. As I looked at the source code I realized that each div class represented either a different article or a different section. If I could check the text after the div class to make sure it represented an article rather than a section then the procedure would know when to stop checking for links and headlines. This was a good experience to understand how HTML source code works, and it was a good learning experience in “If” Statements and Loops.

From these difficulties I learned a lot about HTML source code. I spent a large chunk of my time sifting through source code to see which part of the code corresponded to what part of the website. I learned a lot about how different websites structure their HTML. I learned how to look for patterns and tip-offs that would describe the specific part of the website. Now I am no longer intimidated to look at the source code of a website. I also learned how to make some workaround solutions in getting the correct text that I wanted, and I learned how to make my code flexible to account for changing web pages.

Assistance

In my project I received all of my assistance through Gove Allen. I used his sendGmail function to send emails containing the headlines and links to the news articles.