Chapter6

March 2, 2022

1 Chapter 6 - Data Sourcing via Web

1.1 Part 1 - Objects in BeautifulSoup

[1]:
```
import sys
print(sys.version)
```

```
3.8.8 (default, Apr 13 2021, 15:08:03) [MSC v.1916 64 bit (AMD64)]
```

[2]:
```
from bs4 import BeautifulSoup
```

1.1.1 BeautifulSoup objects

[3]:
```
our_html_document = '''
<html><head><title>IoT Articles</title></head>
<body>
<p class='title'><b>2018 Trends: Best New IoT Device Ideas for Data Scientists␣
 ↪and Engineers</b></p>

<p class='description'>It's almost 2018 and IoT is on the cusp of an explosive␣
 ↪expansion. In this article, I offer you a listing of new IoT device ideas␣
 ↪that you can use...
<br>
<br>
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this␣
 ↪article, I offer you a listing of new IoT device ideas that you can use to␣
 ↪get practice in designing your first IoT applications.
<h1>Looking Back at My Coolest IoT Find in 2017</h1>
Before going into detail about best new IoT device ideas, here's the backstory.␣
 ↪<span style="text-decoration: underline;"><strong><a href="http://bit.ly/
 ↪LPlNDJj">Last month Ericsson Digital invited me</a></strong></span> to tour␣
 ↪the Ericsson Studio in Kista, Sweden. Up until that visit, <a href="http://
 ↪www.data-mania.com/blog/m2m-vs-iot/">IoT</a> had been largely theoretical to␣
 ↪me. Of course, I know the usual mumbo-jumbo about wearables and␣
 ↪IoT-connected fitness trackers. That stuff is all well and good, but it's␣
 ↪somewhat old hat - plus I am not sure we are really benefiting so much from␣
 ↪those, so I'm not that impressed.
```

```
It wasn't until I got to the Ericsson Studio that I became extremely impressed␣
 ↪by how far IoT has really come. Relying on the promise of the 5g network␣
 ↪expansion, IoT-powered smart devices are on the cusp of an explosive growth␣
 ↪in adoption. It was Ericsson's Smart Car that sent me reeling:<a href="bit.
 ↪ly/LPlNDJj"><img class="aligncenter size-full wp-image-3802" src="http://www.
 ↪data-mania.com/blog/wp-content/uploads/2017/12/new-IoT-device-ideas.jpg"␣
 ↪alt="Get your new iot device ideas here" width="1024" height="683" /></a>

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform␣
 ↪that manages services for the Smart Cars to which it's connected. The Volvo␣
 ↪pictured above acts as a drop-off location for groceries that have been␣
 ↪ordered by its owner.

To understand how it works, imagine you're pulling your normal 9-to-5 and you␣
 ↪know you need to grab some groceries on your way home. Well, since you're␣
 ↪smart you've used Ericsson IoT platform to connect your car to the local␣
 ↪grocery delivery service (<a href="http://mat.se/">Mat.se</a>), so all you␣
 ↪need to do is open the Mat.se app and make your usual order. Mat.se␣
 ↪automatically handles the payment, grocery selection, delivery, and delivery␣
 ↪scheduling. Since your car is IoT-enabled, Mat.se issues its trusted␣
 ↪delivery agent a 1-time token to use for opening your car in order to place␣
 ↪your groceries in your car for you at 4:40 pm (just before you get off from␣
 ↪work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio,␣
 ↪make sure to go <span style="text-decoration: underline;"><strong><a␣
 ↪href="http://bit.ly/LPlNDJj">watch the videos on this page</a></strong></
 ↪span>.
<h1>Future Trends for IoT in 2018</h1>
New IoT device ideas won't do you much good unless you at least know the basic␣
 ↪technology trends that are set to impact IoT over the next year(s). These␣
 ↪include:
<ol>
        <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are␣
 ↪embedded within IoT devices spin off machine-generated data like it's going␣
 ↪out of style. For IoT to function, the platform must be solidly engineered␣
 ↪to handle big data. Be assured, that requires some serious data engineering.
 ↪</li>
        <li><strong>Machine Learning</strong> Data Science: While a lot of IoT␣
 ↪devices are still operated according to rules-based decision criteria, the␣
 ↪age of artificial intelligence is upon us. IoT will increasingly depend on␣
 ↪machine learning algorithms to control device operations so that devices are␣
 ↪able to autonomously respond to a complex set of overlapping stimuli.</li>
        <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT␣
 ↪networks must be secure. Blockchain technology is primed to meet the␣
 ↪security demands that come along with building and expanding the IoT.</li>
```

```
</ol>
<h1>Best New IoT Device Ideas</h1>
This listing of new IoT device ideas has been sub-divided according to the main␣
 ↪technology upon which the IoT devices are built. Below I'm providing a list␣
 ↪of new IoT device ideas, but for detailed instructions on how to build these␣
 ↪IoT applications, I recommend the <a href="https://click.linksynergy.com/
 ↪deeplink?id=*JDLXjeE*wk&amp;mid=39197&amp;murl=https%3A%2F%2Fwww.udemy.
 ↪com%2Ftopic%2Finternet-of-things%2F%3Fsort%3Dhighest-rated">IoT courses on␣
 ↪Udemy</a> (ß Please note: if you purchase a Udemy course through this link,␣
 ↪I may receive a small commission), or courses that are available at <a␣
 ↪href="http://www.skyfilabs.com/iot-online-courses">SkyFi</a> and <a␣
 ↪href="https://www.coursera.org/specializations/iot">Coursera</a>.
<h2>Raspberry Pi IoT Ideas</h2>
Using Raspberry Pi as open-source hardware, you can build IoT applications that␣
 ↪offer any one of the following benefits:
<ol>
        <li>Enable built-in sensing to build a weather station that measures␣
 ↪ambient temperature and humidity</li>
        <li>Build a system that detects discrepancies in electrical readings␣
 ↪to identify electricity theft</li>
        <li>Use IoT to build a Servo that is controlled by motion detection␣
 ↪readings</li>
        <li>Build a smart control switch that operates devices based on␣
 ↪external stimuli. Use this for home automation.</li>
        <li>Build a music playing application that enables music for each room␣
 ↪in your house</li>
        <li>Implement biometrics on IoT-connected devices</li>
</ol>
<h2>Arduino IoT Ideas</h2>
There are a number of new IoT device ideas that deploy Arduino as a␣
 ↪microcontroller. These include:
<ol>
        <li>Integrate Arduino with Android to build a remote-control RGB LED␣
 ↪device.</li>
        <li>Connect PIR sensors across the IoT to implement a smart building.</
 ↪li>
        <li>Build a temperature and sunlight sensor system to remotely monitor␣
 ↪and control the conditions of your garden.</li>
        <li>Deploy Arduino and IoT to automate your neighborhood streetlights.
 ↪</li>
        <li>Build a smart irrigation system based on IoT-connected temperature␣
 ↪and moisture sensors built-in to your agricultural plants.</li>
</ol>
```

```
[caption id="attachment_3807" align="aligncenter" width="300"]<a href="bit.ly/
 ↪LPlNDJj"><img class="wp-image-3807 size-medium" src="http://www.data-mania.
 ↪com/blog/wp-content/uploads/2017/12/IMG_3058-300x295.jpg" alt="" width="300"␣
 ↪height="295" /></a> An IoT Chatbot Tree at the Ericsson Studio[/caption]
<h2>Wireless (GSM) IoT Ideas</h2>
Several new IoT device ideas are developed around the GSM wireless network.␣
 ↪Those are:
<ol>
        <li>Monitor soil moisture to automate agricultural irrigation cycles.</
 ↪li>
        <li>Automate and control the conditions of a greenhouse.</li>
        <li>Enable bio-metrics to build a smart security system for your home␣
 ↪or office building</li>
        <li>Build an autonomously operating fitness application that␣
 ↪automatically makes recommendations based on motion detection and heart rate␣
 ↪sensors that are embedded on wearable fitness trackers.</li>
        <li>Build a healthcare monitoring system that tracks, informs, and␣
 ↪automatically alerts healthcare providers based on sensor readings that␣
 ↪describe a patients vital statistics (like temperature, pulse, blood␣
 ↪pressure, etc).</li>
</ol>
<h2>IoT Automation Ideas</h2>
Almost all new IoT device ideas offer automation benefits, but to outline a few␣
 ↪more ideas:
<ol>
        <li>Build an IoT device that automatically locates and reports the␣
 ↪closest nearby parking spot.</li>
        <li>Build a motion detection system that automatically issues emails␣
 ↪or sms messages to alert home owners of a likely home invasion.</li>
        <li>Use temperature sensors connected across the IoT to automatically␣
 ↪alert you if your home windows or doors have been left open.</li>
        <li>Use bio-metric sensors to build a smart system that automate␣
 ↪security for your home or office building</li>
</ol>
To learn more about IoT and what's happening on the leading edge, be sure to␣
 ↪pop over to Ericsson's Studio Tour recap and <span style="text-decoration:␣
 ↪underline;"><strong><a href="http://bit.ly/LPlNDJj">watch these videos</a></
 ↪strong></span>.

<em>(I captured some of this content on behalf of DevMode Strategies during an␣
 ↪invite-only tour of the Ericsson Studio in Kista. Rest assure, the text␣
 ↪and opinions are my own</em>)
<p class='description'>...</p>
'''
```

[4]:
```
our_soup_object = BeautifulSoup(our_html_document, 'html.parser')
print(our_soup_object)
```

```
<html><head><title>IoT Articles</title></head>
<body>
<p class="title"><b>2018 Trends: Best New IoT Device Ideas for Data Scientists
and Engineers</b></p>
<p class="description">It's almost 2018 and IoT is on the cusp of an explosive
expansion. In this article, I offer you a listing of new IoT device ideas that
you can use…
<br/>
<br/>
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use to get
practice in designing your first IoT applications.
<h1>Looking Back at My Coolest IoT Find in 2017</h1>
Before going into detail about best new IoT device ideas, here's the backstory.
<span style="text-decoration: underline;"><strong><a
href="http://bit.ly/LPlNDJj">Last month Ericsson Digital invited
me</a></strong></span> to tour the Ericsson Studio in Kista, Sweden. Up until
that visit, <a href="http://www.data-mania.com/blog/m2m-vs-iot/">IoT</a> had
been largely theoretical to me. Of course, I know the usual mumbo-jumbo about
wearables and IoT-connected fitness trackers. That stuff is all well and good,
but it's somewhat old hat - plus I am not sure we are really benefiting so much
from those, so I'm not that impressed.

It wasn't until I got to the Ericsson Studio that I became extremely impressed
by how far IoT has really come. Relying on the promise of the 5g network
expansion, IoT-powered smart devices are on the cusp of an explosive growth in
adoption. It was Ericsson's Smart Car that sent me reeling:<a
href="bit.ly/LPlNDJj"><img alt="Get your new iot device ideas here"
class="aligncenter size-full wp-image-3802" height="683" src="http://www.data-
mania.com/blog/wp-content/uploads/2017/12/new-IoT-device-ideas.jpg"
width="1024"/></a>

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform
that manages services for the Smart Cars to which it's connected. The Volvo
pictured above acts as a drop-off location for groceries that have been ordered
by its owner.

To understand how it works, imagine you're pulling your normal 9-to-5 and you
know you need to grab some groceries on your way home. Well, since you're smart
you've used Ericsson IoT platform to connect your car to the local grocery
delivery service (<a href="http://mat.se/">Mat.se</a>), so all you need to do is
open the Mat.se app and make your usual order. Mat.se automatically handles the
payment, grocery selection, delivery, and delivery scheduling. Since your car is
IoT-enabled, Mat.se issues its trusted delivery agent a 1-time token to use for
```

opening your car in order to place your groceries in your car for you at 4:40 pm (just before you get off from work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio, make sure to go watch the videos on this page.
<h1>Future Trends for IoT in 2018</h1>
New IoT device ideas won't do you much good unless you at least know the basic technology trends that are set to impact IoT over the next year(s). These include:

Big Data & Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it's going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.
Machine Learning Data Science: While a lot of IoT devices are still operated according to rules-based decision criteria, the age of artificial intelligence is upon us. IoT will increasingly depend on machine learning algorithms to control device operations so that devices are able to autonomously respond to a complex set of overlapping stimuli.
Blockchain-Enabled Security: Above all else, IoT networks must be secure. Blockchain technology is primed to meet the security demands that come along with building and expanding the IoT.

<h1>Best New IoT Device Ideas</h1>
This listing of new IoT device ideas has been sub-divided according to the main technology upon which the IoT devices are built. Below I'm providing a list of new IoT device ideas, but for detailed instructions on how to build these IoT applications, I recommend the IoT courses on Udemy (ß Please note: if you purchase a Udemy course through this link, I may receive a small commission), or courses that are available at SkyFi and Coursera.
<h2>Raspberry Pi IoT Ideas</h2>
Using Raspberry Pi as open-source hardware, you can build IoT applications that offer any one of the following benefits:

Enable built-in sensing to build a weather station that measures ambient temperature and humidity
Build a system that detects discrepancies in electrical readings to identify electricity theft
Use IoT to build a Servo that is controlled by motion detection readings
Build a smart control switch that operates devices based on external stimuli. Use this for home automation.
Build a music playing application that enables music for each room in your

```
house</li>
<li>Implement biometrics on IoT-connected devices</li>
</ol>
<h2>Arduino IoT Ideas</h2>
There are a number of new IoT device ideas that deploy Arduino as a
microcontroller. These include:
<ol>
<li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>
<li>Connect PIR sensors across the IoT to implement a smart building.</li>
<li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>
<li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
<li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>
</ol>
[caption id="attachment_3807" align="aligncenter" width="300"]<a
href="bit.ly/LPlNDJj"><img alt="" class="wp-image-3807 size-medium" height="295"
src="http://www.data-mania.com/blog/wp-
content/uploads/2017/12/IMG_3058-300x295.jpg" width="300"/></a> An IoT Chatbot
Tree at the Ericsson Studio[/caption]
<h2>Wireless (GSM) IoT Ideas</h2>
Several new IoT device ideas are developed around the GSM wireless network.
Those are:
<ol>
<li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
<li>Automate and control the conditions of a greenhouse.</li>
<li>Enable bio-metrics to build a smart security system for your home or office
building</li>
<li>Build an autonomously operating fitness application that automatically makes
recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>
<li>Build a healthcare monitoring system that tracks, informs, and automatically
alerts healthcare providers based on sensor readings that describe a patients
vital statistics (like temperature, pulse, blood pressure, etc).</li>
</ol>
<h2>IoT Automation Ideas</h2>
Almost all new IoT device ideas offer automation benefits, but to outline a few
more ideas:
<ol>
<li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>
<li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>
<li>Use temperature sensors connected across the IoT to automatically alert you
if your home windows or doors have been left open.</li>
<li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>
```

```
</ol>
To learn more about IoT and what's happening on the leading edge, be sure to pop
over to Ericsson's Studio Tour recap and <span style="text-decoration:
underline;"><strong><a href="http://bit.ly/LPlNDJj">watch these
videos</a></strong></span>.

<em>(I captured some of this content on behalf of DevMode Strategies during an
invite-only tour of the Ericsson Studio in Kista. Rest assure, the text
and opinions are my own</em>)
<p class="description">…</p>
</p></body></html>
```

[11]:
```
print(our_soup_object.prettify()[0:300])
```

```
<html>
 <head>
  <title>
   IoT Articles
  </title>
 </head>
 <body>
  <p class="title">
   <b>
    2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers
   </b>
  </p>
  <p class="description">
   It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article,
```

1.1.2 Tag objects

Tag names

[5]:
```
soup_object = BeautifulSoup('<h1 attribute_1 = "Heading Level 1"">Future Trends␣
↪for IoT in 2018</h1>', "lxml")

tag = soup_object.h1
type(tag)
```

[5]:
```
bs4.element.Tag
```

[6]:
```
print(tag)
```

```
<h1 attribute_1="Heading Level 1">Future Trends for IoT in 2018</h1>
```

[7]:
```
tag.name
```

[7]:
```
'h1'
```

[8]:
```
tag.name = 'heading 1'
tag
```

[8]:
```
<heading 1 attribute_1="Heading Level 1">Future Trends for IoT in 2018</heading
1>
```

[9]:
```
tag.name
```

[9]:
```
'heading 1'
```

Tag attributes

[10]:
```
soup_object = BeautifulSoup('<h1 attribute_1 = "Heading Level 1"">Future Trends␣
↪for IoT in 2018</h1>', "lxml")
tag = soup_object.h1
tag
```

[10]:
```
<h1 attribute_1="Heading Level 1">Future Trends for IoT in 2018</h1>
```

[11]:
```
tag['attribute_1']
```

[11]:
```
'Heading Level 1'
```

[12]:
```
tag.attrs
```

[12]:
```
{'attribute_1': 'Heading Level 1'}
```

[13]:
```
tag['attribute_2'] = 'Heading Level 1*'
tag.attrs
```

[13]:
```
{'attribute_1': 'Heading Level 1', 'attribute_2': 'Heading Level 1*'}
```

[14]:
```
tag
```

[14]:
```
<h1 attribute_1="Heading Level 1" attribute_2="Heading Level 1*">Future Trends
for IoT in 2018</h1>
```

[15]:
```
del tag['attribute_2']
tag
```

[15]:
```
<h1 attribute_1="Heading Level 1">Future Trends for IoT in 2018</h1>
```

[16]:
```
del tag['attribute_1']
tag.attrs
```

[16]:
```
{}
```

Navigating a parse tree using tags

[17]:
```
# First we will recreate our original parse tree.
our_html_document = '''
<html><head><title>IoT Articles</title></head>
<body>
<p class='title'><b>2018 Trends: Best New IoT Device Ideas for Data Scientists␣
 ↪and Engineers</b></p>

<p class='description'>It's almost 2018 and IoT is on the cusp of an explosive␣
 ↪expansion. In this article, I offer you a listing of new IoT device ideas␣
 ↪that you can use...
<br>
<br>
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this␣
 ↪article, I offer you a listing of new IoT device ideas that you can use to␣
 ↪get practice in designing your first IoT applications.
<h1>Looking Back at My Coolest IoT Find in 2017</h1>
Before going into detail about best new IoT device ideas, here's the backstory.␣
 ↪<span style="text-decoration: underline;"><strong><a href="http://bit.ly/
 ↪LPlNDJj">Last month Ericsson Digital invited me</a></strong></span> to tour␣
 ↪the Ericsson Studio in Kista, Sweden. Up until that visit, <a href="http://
 ↪www.data-mania.com/blog/m2m-vs-iot/">IoT</a> had been largely theoretical to␣
 ↪me. Of course, I know the usual mumbo-jumbo about wearables and␣
 ↪IoT-connected fitness trackers. That stuff is all well and good, but it's␣
 ↪somewhat old hat - plus I am not sure we are really benefiting so much from␣
 ↪those, so I'm not that impressed.

It wasn't until I got to the Ericsson Studio that I became extremely impressed␣
 ↪by how far IoT has really come. Relying on the promise of the 5g network␣
 ↪expansion, IoT-powered smart devices are on the cusp of an explosive growth␣
 ↪in adoption. It was Ericsson's Smart Car that sent me reeling:<a href="bit.
 ↪ly/LPlNDJj"><img class="aligncenter size-full wp-image-3802" src="http://www.
 ↪data-mania.com/blog/wp-content/uploads/2017/12/new-IoT-device-ideas.jpg"␣
 ↪alt="Get your new iot device ideas here" width="1024" height="683" /></a>

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform␣
 ↪that manages services for the Smart Cars to which it's connected. The Volvo␣
 ↪pictured above acts as a drop-off location for groceries that have been␣
 ↪ordered by its owner.
```

```
To understand how it works, imagine you're pulling your normal 9-to-5 and you␣
 ↪know you need to grab some groceries on your way home. Well, since you're␣
 ↪smart you've used Ericsson IoT platform to connect your car to the local␣
 ↪grocery delivery service (<a href="http://mat.se/">Mat.se</a>), so all you␣
 ↪need to do is open the Mat.se app and make your usual order. Mat.se␣
 ↪automatically handles the payment, grocery selection, delivery, and delivery␣
 ↪scheduling. Since your car is IoT-enabled, Mat.se issues its trusted␣
 ↪delivery agent a 1-time token to use for opening your car in order to place␣
 ↪your groceries in your car for you at 4:40 pm (just before you get off from␣
 ↪work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio,␣
 ↪make sure to go <span style="text-decoration: underline;"><strong><a␣
 ↪href="http://bit.ly/LPlNDJj">watch the videos on this page</a></strong></
 ↪span>.
<h1>Future Trends for IoT in 2018</h1>
New IoT device ideas won't do you much good unless you at least know the basic␣
 ↪technology trends that are set to impact IoT over the next year(s). These␣
 ↪include:
<ol>
        <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are␣
 ↪embedded within IoT devices spin off machine-generated data like it's going␣
 ↪out of style. For IoT to function, the platform must be solidly engineered␣
 ↪to handle big data. Be assured, that requires some serious data engineering.
 ↪</li>
        <li><strong>Machine Learning</strong> Data Science: While a lot of IoT␣
 ↪devices are still operated according to rules-based decision criteria, the␣
 ↪age of artificial intelligence is upon us. IoT will increasingly depend on␣
 ↪machine learning algorithms to control device operations so that devices are␣
 ↪able to autonomously respond to a complex set of overlapping stimuli.</li>
        <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT␣
 ↪networks must be secure. Blockchain technology is primed to meet the␣
 ↪security demands that come along with building and expanding the IoT.</li>
</ol>
<h1>Best New IoT Device Ideas</h1>
This listing of new IoT device ideas has been sub-divided according to the main␣
 ↪technology upon which the IoT devices are built. Below I'm providing a list␣
 ↪of new IoT device ideas, but for detailed instructions on how to build these␣
 ↪IoT applications, I recommend the <a href="https://click.linksynergy.com/
 ↪deeplink?id=*JDLXjeE*wk&amp;mid=39197&amp;murl=https%3A%2F%2Fwww.udemy.
 ↪com%2Ftopic%2Finternet-of-things%2F%3Fsort%3Dhighest-rated">IoT courses on␣
 ↪Udemy</a> (ß Please note: if you purchase a Udemy course through this link,␣
 ↪I may receive a small commission), or courses that are available at <a␣
 ↪href="http://www.skyfilabs.com/iot-online-courses">SkyFi</a> and <a␣
 ↪href="https://www.coursera.org/specializations/iot">Coursera</a>.
<h2>Raspberry Pi IoT Ideas</h2>
```

```
Using Raspberry Pi as open-source hardware, you can build IoT applications that␣
 ↪offer any one of the following benefits:
<ol>
        <li>Enable built-in sensing to build a weather station that measures␣
 ↪ambient temperature and humidity</li>
        <li>Build a system that detects discrepancies in electrical readings␣
 ↪to identify electricity theft</li>
        <li>Use IoT to build a Servo that is controlled by motion detection␣
 ↪readings</li>
        <li>Build a smart control switch that operates devices based on␣
 ↪external stimuli. Use this for home automation.</li>
        <li>Build a music playing application that enables music for each room␣
 ↪in your house</li>
        <li>Implement biometrics on IoT-connected devices</li>
</ol>
<h2>Arduino IoT Ideas</h2>
There are a number of new IoT device ideas that deploy Arduino as a␣
 ↪microcontroller. These include:
<ol>
        <li>Integrate Arduino with Android to build a remote-control RGB LED␣
 ↪device.</li>
        <li>Connect PIR sensors across the IoT to implement a smart building.</
 ↪li>
        <li>Build a temperature and sunlight sensor system to remotely monitor␣
 ↪and control the conditions of your garden.</li>
        <li>Deploy Arduino and IoT to automate your neighborhood streetlights.
 ↪</li>
        <li>Build a smart irrigation system based on IoT-connected temperature␣
 ↪and moisture sensors built-in to your agricultural plants.</li>
</ol>
[caption id="attachment_3807" align="aligncenter" width="300"]<a href="bit.ly/
 ↪LPlNDJj"><img class="wp-image-3807 size-medium" src="http://www.data-mania.
 ↪com/blog/wp-content/uploads/2017/12/IMG_3058-300x295.jpg" alt="" width="300"␣
 ↪height="295" /></a> An IoT Chatbot Tree at the Ericsson Studio[/caption]
<h2>Wireless (GSM) IoT Ideas</h2>
Several new IoT device ideas are developed around the GSM wireless network.␣
 ↪Those are:
<ol>
        <li>Monitor soil moisture to automate agricultural irrigation cycles.</
 ↪li>
        <li>Automate and control the conditions of a greenhouse.</li>
        <li>Enable bio-metrics to build a smart security system for your home␣
 ↪or office building</li>
        <li>Build an autonomously operating fitness application that␣
 ↪automatically makes recommendations based on motion detection and heart rate␣
 ↪sensors that are embedded on wearable fitness trackers.</li>
```

```
        <li>Build a healthcare monitoring system that tracks, informs, and␣
 ↪automatically alerts healthcare providers based on sensor readings that␣
 ↪describe a patients vital statistics (like temperature, pulse, blood␣
 ↪pressure, etc).</li>
</ol>
<h2>IoT Automation Ideas</h2>
Almost all new IoT device ideas offer automation benefits, but to outline a few␣
 ↪more ideas:
<ol>
        <li>Build an IoT device that automatically locates and reports the␣
 ↪closest nearby parking spot.</li>
        <li>Build a motion detection system that automatically issues emails␣
 ↪or sms messages to alert home owners of a likely home invasion.</li>
        <li>Use temperature sensors connected across the IoT to automatically␣
 ↪alert you if your home windows or doors have been left open.</li>
        <li>Use bio-metric sensors to build a smart system that automate␣
 ↪security for your home or office building</li>
</ol>
To learn more about IoT and what's happening on the leading edge, be sure to␣
 ↪pop over to Ericsson's Studio Tour recap and <span style="text-decoration:␣
 ↪underline;"><strong><a href="http://bit.ly/LPlNDJj">watch these videos</a></
 ↪strong></span>.

<em>(I captured some of this content on behalf of DevMode Strategies during an␣
 ↪invite-only tour of the Ericsson Studio in Kista. Rest assure, the text␣
 ↪and opinions are my own</em>)
<p class='description'>...</p>
'''
our_soup_object = BeautifulSoup(our_html_document, 'html.parser')
```

[22]:
```
our_soup_object.head
```

[22]:
```
<head><title>IoT Articles</title></head>
```

[21]:
```
our_soup_object.title
```

[21]:
```
<title>IoT Articles</title>
```

[20]:
```
our_soup_object.body.b
```

[20]:
```
<b>2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers</b>
```

[23]:
```
our_soup_object.body
```

[23]:
```
<body>
<p class="title"><b>2018 Trends: Best New IoT Device Ideas for Data Scientists
and Engineers</b></p>
```

<p class="description">It’s almost 2018 and IoT is on the cusp of an explosive expansion. In this article, I offer you a listing of new IoT device ideas that you can use…

It’s almost 2018 and IoT is on the cusp of an explosive expansion. In this article, I offer you a listing of new IoT device ideas that you can use to get practice in designing your first IoT applications.
<h1>Looking Back at My Coolest IoT Find in 2017</h1>
Before going into detail about best new IoT device ideas, here’s the backstory. Last month Ericsson Digital invited me to tour the Ericsson Studio in Kista, Sweden. Up until that visit, IoT had been largely theoretical to me. Of course, I know the usual mumbo-jumbo about wearables and IoT-connected fitness trackers. That stuff is all well and good, but it’s somewhat old hat - plus I am not sure we are really benefiting so much from those, so I’m not that impressed.

It wasn’t until I got to the Ericsson Studio that I became extremely impressed by how far IoT has really come. Relying on the promise of the 5g network expansion, IoT-powered smart devices are on the cusp of an explosive growth in adoption. It was Ericsson’s Smart Car that sent me reeling:

This car is connected to Ericsson’s Connected Vehicle Cloud, an IoT platform that manages services for the Smart Cars to which it’s connected. The Volvo pictured above acts as a drop-off location for groceries that have been ordered by its owner.

To understand how it works, imagine you’re pulling your normal 9-to-5 and you know you need to grab some groceries on your way home. Well, since you’re smart you’ve used Ericsson IoT platform to connect your car to the local grocery delivery service (Mat.se), so all you need to do is open the Mat.se app and make your usual order. Mat.se automatically handles the payment, grocery selection, delivery, and delivery scheduling. Since your car is IoT-enabled, Mat.se issues its trusted delivery agent a 1-time token to use for opening your car in order to place your groceries in your car for you at 4:40 pm (just before you get off from work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio, make sure to go watch the videos on this page.
<h1>Future Trends for IoT in 2018</h1>

New IoT device ideas won’t do you much good unless you at least know the basic technology trends that are set to impact IoT over the next year(s). These include:

Big Data & Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it’s going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.
Machine Learning Data Science: While a lot of IoT devices are still operated according to rules-based decision criteria, the age of artificial intelligence is upon us. IoT will increasingly depend on machine learning algorithms to control device operations so that devices are able to autonomously respond to a complex set of overlapping stimuli.
Blockchain-Enabled Security: Above all else, IoT networks must be secure. Blockchain technology is primed to meet the security demands that come along with building and expanding the IoT.

<h1>Best New IoT Device Ideas</h1>
This listing of new IoT device ideas has been sub-divided according to the main technology upon which the IoT devices are built. Below I’m providing a list of new IoT device ideas, but for detailed instructions on how to build these IoT applications, I recommend the IoT courses on Udemy (ß Please note: if you purchase a Udemy course through this link, I may receive a small commission), or courses that are available at SkyFi and Coursera.
<h2>Raspberry Pi IoT Ideas</h2>
Using Raspberry Pi as open-source hardware, you can build IoT applications that offer any one of the following benefits:

Enable built-in sensing to build a weather station that measures ambient temperature and humidity
Build a system that detects discrepancies in electrical readings to identify electricity theft
Use IoT to build a Servo that is controlled by motion detection readings
Build a smart control switch that operates devices based on external stimuli. Use this for home automation.
Build a music playing application that enables music for each room in your house
Implement biometrics on IoT-connected devices

<h2>Arduino IoT Ideas</h2>
There are a number of new IoT device ideas that deploy Arduino as a microcontroller. These include:

Integrate Arduino with Android to build a remote-control RGB LED device.
Connect PIR sensors across the IoT to implement a smart building.
Build a temperature and sunlight sensor system to remotely monitor and control the conditions of your garden.
Deploy Arduino and IoT to automate your neighborhood streetlights.
Build a smart irrigation system based on IoT-connected temperature and moisture sensors built-in to your agricultural plants.

[caption id="attachment_3807" align="aligncenter" width="300"] An IoT Chatbot Tree at the Ericsson Studio[/caption]
<h2>Wireless (GSM) IoT Ideas</h2>
Several new IoT device ideas are developed around the GSM wireless network. Those are:

Monitor soil moisture to automate agricultural irrigation cycles.
Automate and control the conditions of a greenhouse.
Enable bio-metrics to build a smart security system for your home or office building
Build an autonomously operating fitness application that automatically makes recommendations based on motion detection and heart rate sensors that are embedded on wearable fitness trackers.
Build a healthcare monitoring system that tracks, informs, and automatically alerts healthcare providers based on sensor readings that describe a patients vital statistics (like temperature, pulse, blood pressure, etc).

<h2>IoT Automation Ideas</h2>
Almost all new IoT device ideas offer automation benefits, but to outline a few more ideas:

Build an IoT device that automatically locates and reports the closest nearby parking spot.
Build a motion detection system that automatically issues emails or sms messages to alert home owners of a likely home invasion.
Use temperature sensors connected across the IoT to automatically alert you if your home windows or doors have been left open.
Use bio-metric sensors to build a smart system that automate security for your home or office building

To learn more about IoT and what’s happening on the leading edge, be sure to pop over to Ericsson’s Studio Tour recap and watch these videos.

```
<em>(I captured some of this content on behalf of DevMode Strategies during an
invite-only tour of the Ericsson Studio in Kista. Rest assure, the text
and opinions are my own</em>)
<p class="description">…</p>
</p></body>
```

[24]:
```
our_soup_object.li
```

[24]:
```
<li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are embedded
within IoT devices spin off machine-generated data like it's going out of style.
For IoT to function, the platform must be solidly engineered to handle big data.
Be assured, that requires some serious data engineering.</li>
```

[26]:
```
our_soup_object.a
```

[26]:
```
<a href="http://bit.ly/LPlNDJj">Last month Ericsson Digital invited me</a>
```

1.2 Part 2 - NavigatableString Objects

[27]:
```
soup_object = BeautifulSoup('<h1 attribute_1 = "Heading Level 1"">Future Trends␣
↪in IoT in 2018</h1>', "lxml")

tag = soup_object.h1

type(tag)
```

[27]:
```
bs4.element.Tag
```

[28]:
```
tag.name
```

[28]:
```
'h1'
```

[29]:
```
tag.string
```

[29]:
```
'Future Trends in IoT in 2018'
```

[30]:
```
type(tag.string)
```

[30]:
```
bs4.element.NavigableString
```

[31]:
```
our_navigatable_string = tag.string
our_navigatable_string
```

[31]:
```
'Future Trends in IoT in 2018'
```

[32]:
```
our_navigatable_string.replace_with('NaN')
tag.string
```

[32]:
```
'NaN'
```

Utilizing NavigatableString objects

[34]:
```
our_html_document = '''
<html><head><title>IoT Articles</title></head>
<body>
<p class='title'><b>2018 Trends: Best New IoT Device Ideas for Data Scientists␣
 ↪and Engineers</b></p>

<p class='description'>It's almost 2018 and IoT is on the cusp of an explosive␣
 ↪expansion. In this article, I offer you a listing of new IoT device ideas␣
 ↪that you can use...
<br>
<br>
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this␣
 ↪article, I offer you a listing of new IoT device ideas that you can use to␣
 ↪get practice in designing your first IoT applications.
<h1>Looking Back at My Coolest IoT Find in 2017</h1>
Before going into detail about best new IoT device ideas, here's the backstory.␣
 ↪<span style="text-decoration: underline;"><strong><a href="http://bit.ly/
 ↪LPlNDJj">Last month Ericsson Digital invited me</a></strong></span> to tour␣
 ↪the Ericsson Studio in Kista, Sweden. Up until that visit, <a href="http://
 ↪www.data-mania.com/blog/m2m-vs-iot/">IoT</a> had been largely theoretical to␣
 ↪me. Of course, I know the usual mumbo-jumbo about wearables and␣
 ↪IoT-connected fitness trackers. That stuff is all well and good, but it's␣
 ↪somewhat old hat - plus I am not sure we are really benefiting so much from␣
 ↪those, so I'm not that impressed.

It wasn't until I got to the Ericsson Studio that I became extremely impressed␣
 ↪by how far IoT has really come. Relying on the promise of the 5g network␣
 ↪expansion, IoT-powered smart devices are on the cusp of an explosive growth␣
 ↪in adoption. It was Ericsson's Smart Car that sent me reeling:<a href="bit.
 ↪ly/LPlNDJj"><img class="aligncenter size-full wp-image-3802" src="http://www.
 ↪data-mania.com/blog/wp-content/uploads/2017/12/new-IoT-device-ideas.jpg"␣
 ↪alt="Get your new iot device ideas here" width="1024" height="683" /></a>

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform␣
 ↪that manages services for the Smart Cars to which it's connected. The Volvo␣
 ↪pictured above acts as a drop-off location for groceries that have been␣
 ↪ordered by its owner.
```

```
To understand how it works, imagine you're pulling your normal 9-to-5 and you␣
 ↪know you need to grab some groceries on your way home. Well, since you're␣
 ↪smart you've used Ericsson IoT platform to connect your car to the local␣
 ↪grocery delivery service (<a href="http://mat.se/">Mat.se</a>), so all you␣
 ↪need to do is open the Mat.se app and make your usual order. Mat.se␣
 ↪automatically handles the payment, grocery selection, delivery, and delivery␣
 ↪scheduling. Since your car is IoT-enabled, Mat.se issues its trusted␣
 ↪delivery agent a 1-time token to use for opening your car in order to place␣
 ↪your groceries in your car for you at 4:40 pm (just before you get off from␣
 ↪work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio,␣
 ↪make sure to go <span style="text-decoration: underline;"><strong><a␣
 ↪href="http://bit.ly/LPlNDJj">watch the videos on this page</a></strong></
 ↪span>.
<h1>Future Trends for IoT in 2018</h1>
New IoT device ideas won't do you much good unless you at least know the basic␣
 ↪technology trends that are set to impact IoT over the next year(s). These␣
 ↪include:
<ol>
        <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are␣
 ↪embedded within IoT devices spin off machine-generated data like it's going␣
 ↪out of style. For IoT to function, the platform must be solidly engineered␣
 ↪to handle big data. Be assured, that requires some serious data engineering.
 ↪</li>
        <li><strong>Machine Learning</strong> Data Science: While a lot of IoT␣
 ↪devices are still operated according to rules-based decision criteria, the␣
 ↪age of artificial intelligence is upon us. IoT will increasingly depend on␣
 ↪machine learning algorithms to control device operations so that devices are␣
 ↪able to autonomously respond to a complex set of overlapping stimuli.</li>
        <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT␣
 ↪networks must be secure. Blockchain technology is primed to meet the␣
 ↪security demands that come along with building and expanding the IoT.</li>
</ol>
<h1>Best New IoT Device Ideas</h1>
This listing of new IoT device ideas has been sub-divided according to the main␣
 ↪technology upon which the IoT devices are built. Below I'm providing a list␣
 ↪of new IoT device ideas, but for detailed instructions on how to build these␣
 ↪IoT applications, I recommend the <a href="https://click.linksynergy.com/
 ↪deeplink?id=*JDLXjeE*wk&amp;mid=39197&amp;murl=https%3A%2F%2Fwww.udemy.
 ↪com%2Ftopic%2Finternet-of-things%2F%3Fsort%3Dhighest-rated">IoT courses on␣
 ↪Udemy</a> (ß Please note: if you purchase a Udemy course through this link,␣
 ↪I may receive a small commission), or courses that are available at <a␣
 ↪href="http://www.skyfilabs.com/iot-online-courses">SkyFi</a> and <a␣
 ↪href="https://www.coursera.org/specializations/iot">Coursera</a>.
<h2>Raspberry Pi IoT Ideas</h2>
```

```
Using Raspberry Pi as open-source hardware, you can build IoT applications that␣
 ↪offer any one of the following benefits:
<ol>
        <li>Enable built-in sensing to build a weather station that measures␣
 ↪ambient temperature and humidity</li>
        <li>Build a system that detects discrepancies in electrical readings␣
 ↪to identify electricity theft</li>
        <li>Use IoT to build a Servo that is controlled by motion detection␣
 ↪readings</li>
        <li>Build a smart control switch that operates devices based on␣
 ↪external stimuli. Use this for home automation.</li>
        <li>Build a music playing application that enables music for each room␣
 ↪in your house</li>
        <li>Implement biometrics on IoT-connected devices</li>
</ol>
<h2>Arduino IoT Ideas</h2>
There are a number of new IoT device ideas that deploy Arduino as a␣
 ↪microcontroller. These include:
<ol>
        <li>Integrate Arduino with Android to build a remote-control RGB LED␣
 ↪device.</li>
        <li>Connect PIR sensors across the IoT to implement a smart building.</
 ↪li>
        <li>Build a temperature and sunlight sensor system to remotely monitor␣
 ↪and control the conditions of your garden.</li>
        <li>Deploy Arduino and IoT to automate your neighborhood streetlights.
 ↪</li>
        <li>Build a smart irrigation system based on IoT-connected temperature␣
 ↪and moisture sensors built-in to your agricultural plants.</li>
</ol>
[caption id="attachment_3807" align="aligncenter" width="300"]<a href="bit.ly/
 ↪LPlNDJj"><img class="wp-image-3807 size-medium" src="http://www.data-mania.
 ↪com/blog/wp-content/uploads/2017/12/IMG_3058-300x295.jpg" alt="" width="300"␣
 ↪height="295" /></a> An IoT Chatbot Tree at the Ericsson Studio[/caption]
<h2>Wireless (GSM) IoT Ideas</h2>
Several new IoT device ideas are developed around the GSM wireless network.␣
 ↪Those are:
<ol>
        <li>Monitor soil moisture to automate agricultural irrigation cycles.</
 ↪li>
        <li>Automate and control the conditions of a greenhouse.</li>
        <li>Enable bio-metrics to build a smart security system for your home␣
 ↪or office building</li>
        <li>Build an autonomously operating fitness application that␣
 ↪automatically makes recommendations based on motion detection and heart rate␣
 ↪sensors that are embedded on wearable fitness trackers.</li>
```

```
        <li>Build a healthcare monitoring system that tracks, informs, and␣
 ↪automatically alerts healthcare providers based on sensor readings that␣
 ↪describe a patients vital statistics (like temperature, pulse, blood␣
 ↪pressure, etc).</li>
</ol>
<h2>IoT Automation Ideas</h2>
Almost all new IoT device ideas offer automation benefits, but to outline a few␣
 ↪more ideas:
<ol>
        <li>Build an IoT device that automatically locates and reports the␣
 ↪closest nearby parking spot.</li>
        <li>Build a motion detection system that automatically issues emails␣
 ↪or sms messages to alert home owners of a likely home invasion.</li>
        <li>Use temperature sensors connected across the IoT to automatically␣
 ↪alert you if your home windows or doors have been left open.</li>
        <li>Use bio-metric sensors to build a smart system that automate␣
 ↪security for your home or office building</li>
</ol>
To learn more about IoT and what's happening on the leading edge, be sure to␣
 ↪pop over to Ericsson's Studio Tour recap and <span style="text-decoration:␣
 ↪underline;"><strong><a href="http://bit.ly/LPlNDJj">watch these videos</a></
 ↪strong></span>.

<em>(I captured some of this content on behalf of DevMode Strategies during an␣
 ↪invite-only tour of the Ericsson Studio in Kista. Rest assure, the text␣
 ↪and opinions are my own</em>)
<p class='description'>...</p>
'''
our_soup_object = BeautifulSoup(our_html_document, 'html.parser')
```

[35]:
```
for string in our_soup_object.stripped_strings:
    print(repr(string))
```

```
'IoT Articles'
'2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers'
'It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use…'
'It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use to get
practice in designing your first IoT applications.'
'Looking Back at My Coolest IoT Find in 2017'
'Before going into detail about best new IoT device ideas, here's the
backstory.'
'Last month Ericsson Digital invited me'
'to tour the Ericsson Studio in Kista, Sweden. Up until that visit,'
'IoT'
'had been largely theoretical to me. Of course, I know the usual mumbo-jumbo
```

about wearables and IoT-connected fitness trackers. That stuff is all well and good, but it’s somewhat old hat - plus I am not sure we are really benefiting so much from those, so I’m not that impressed.\n\nIt wasn’t until I got to the Ericsson Studio that I became extremely impressed by how far IoT has really come. Relying on the promise of the 5g network expansion, IoT-powered smart devices are on the cusp of an explosive growth in adoption. It was Ericsson’s Smart Car that sent me reeling:'
'This car is connected to Ericsson’s Connected Vehicle Cloud, an IoT platform that manages services for the Smart Cars to which it’s connected. The Volvo pictured above acts as a drop-off location for groceries that have been ordered by its owner.\n\nTo understand how it works, imagine you’re pulling your normal 9-to-5 and you know you need to grab some groceries on your way home. Well, since you’re smart you’ve used Ericsson IoT platform to connect your car to the local grocery delivery service ('
'Mat.se'
'), so all you need to do is open the Mat.se app and make your usual order. Mat.se automatically handles the payment, grocery selection, delivery, and delivery scheduling. Since your car is IoT-enabled, Mat.se issues its trusted delivery agent a 1-time token to use for opening your car in order to place your groceries in your car for you at 4:40 pm (just before you get off from work).\n\nTo watch some of the amazing IoT device demos I witnessed at Ericsson Studio, make sure to go'
'watch the videos on this page'
'.'
'Future Trends for IoT in 2018'
'New IoT device ideas won’t do you much good unless you at least know the basic technology trends that are set to impact IoT over the next year(s). These include:'
'Big Data'
'& Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it’s going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.'
'Machine Learning'
'Data Science: While a lot of IoT devices are still operated according to rules-based decision criteria, the age of artificial intelligence is upon us. IoT will increasingly depend on machine learning algorithms to control device operations so that devices are able to autonomously respond to a complex set of overlapping stimuli.'
'Blockchain'
'-Enabled Security: Above all else, IoT networks must be secure. Blockchain technology is primed to meet the security demands that come along with building and expanding the IoT.'
'Best New IoT Device Ideas'
'This listing of new IoT device ideas has been sub-divided according to the main technology upon which the IoT devices are built. Below I’m providing a list of new IoT device ideas, but for detailed instructions on how to build these IoT applications, I recommend the'

'IoT courses on Udemy'
'(ß Please note: if you purchase a Udemy course through this link, I may receive a small commission), or courses that are available at'
'SkyFi'
'and'
'Coursera'
'.'
'Raspberry Pi IoT Ideas'
'Using Raspberry Pi as open-source hardware, you can build IoT applications that offer any one of the following benefits:'
'Enable built-in sensing to build a weather station that measures ambient temperature and humidity'
'Build a system that detects discrepancies in electrical readings to identify electricity theft'
'Use IoT to build a Servo that is controlled by motion detection readings'
'Build a smart control switch that operates devices based on external stimuli. Use this for home automation.'
'Build a music playing application that enables music for each room in your house'
'Implement biometrics on IoT-connected devices'
'Arduino IoT Ideas'
'There are a number of new IoT device ideas that deploy Arduino as a microcontroller. These include:'
'Integrate Arduino with Android to build a remote-control RGB LED device.'
'Connect PIR sensors across the IoT to implement a smart building.'
'Build a temperature and sunlight sensor system to remotely monitor and control the conditions of your garden.'
'Deploy Arduino and IoT to automate your neighborhood streetlights.'
'Build a smart irrigation system based on IoT-connected temperature and moisture sensors built-in to your agricultural plants.'
'[caption id="attachment_3807" align="aligncenter" width="300"]'
'An IoT Chatbot Tree at the Ericsson Studio[/caption]'
'Wireless (GSM) IoT Ideas'
'Several new IoT device ideas are developed around the GSM wireless network. Those are:'
'Monitor soil moisture to automate agricultural irrigation cycles.'
'Automate and control the conditions of a greenhouse.'
'Enable bio-metrics to build a smart security system for your home or office building'
'Build an autonomously operating fitness application that automatically makes recommendations based on motion detection and heart rate sensors that are embedded on wearable fitness trackers.'
'Build a healthcare monitoring system that tracks, informs, and automatically alerts healthcare providers based on sensor readings that describe a patients vital statistics (like temperature, pulse, blood pressure, etc).'
'IoT Automation Ideas'
'Almost all new IoT device ideas offer automation benefits, but to outline a few more ideas:'

```
'Build an IoT device that automatically locates and reports the closest nearby
parking spot.'
'Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.'
'Use temperature sensors connected across the IoT to automatically alert you if
your home windows or doors have been left open.'
'Use bio-metric sensors to build a smart system that automate security for your
home or office building'
'To learn more about IoT and what's happening on the leading edge, be sure to
pop over to Ericsson's Studio Tour recap and'
'watch these videos'
'.'
'(I captured some of this content on behalf of DevMode Strategies during an
invite-only tour of the Ericsson Studio in Kista. Rest assure, the text
and\xa0opinions are my own'
')'
'…'
```

[36]:
```
first_link= our_soup_object.a
print(first_link)
```

```
<a href="http://bit.ly/LPlNDJj">Last month Ericsson Digital invited me</a>
```

[37]:
```
first_link.parent
```

[37]:
```
<strong><a href="http://bit.ly/LPlNDJj">Last month Ericsson Digital invited
me</a></strong>
```

[38]:
```
first_link.string
```

[38]:
```
'Last month Ericsson Digital invited me'
```

[39]:
```
first_link.string.parent
```

[39]:
```
<a href="http://bit.ly/LPlNDJj">Last month Ericsson Digital invited me</a>
```

1.3 Segment 3 - Data parsing

[41]:
```
import urllib
import urllib.request
with urllib.request.urlopen('https://raw.githubusercontent.com/BigDataGal/
 ↪Data-Mania-Demos/master/IoT-2018.html') as response:
    html = response.read()
```

[42]:
```
soup = BeautifulSoup(html, "lxml")
type(soup)
```

[42]:
```
bs4.BeautifulSoup
```

1.3.1 Parsing your data

[44]:
```
print(soup.prettify()[0:100])
```

```
<html>
 <head>
  <title>
   IoT Articles
  </title>
 </head>
 <body>
  <p class="title">
   <b>
```

1.3.2 Getting data from a parse tree

[46]:
```
text_only = soup.get_text()
print(text_only)
```

```
IoT Articles

2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use…


It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use to get
practice in designing your first IoT applications.
Looking Back at My Coolest IoT Find in 2017
Before going into detail about best new IoT device ideas, here's the backstory.
Last month Ericsson Digital invited me to tour the Ericsson Studio in Kista,
Sweden. Up until that visit, IoT had been largely theoretical to me. Of course,
I know the usual mumbo-jumbo about wearables and IoT-connected fitness trackers.
That stuff is all well and good, but it's somewhat old hat - plus I am not sure
we are really benefiting so much from those, so I'm not that impressed.

It wasn't until I got to the Ericsson Studio that I became extremely impressed
by how far IoT has really come. Relying on the promise of the 5g network
expansion, IoT-powered smart devices are on the cusp of an explosive growth in
adoption. It was Ericsson's Smart Car that sent me reeling:

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform
that manages services for the Smart Cars to which it's connected. The Volvo
pictured above acts as a drop-off location for groceries that have been ordered
by its owner.

To understand how it works, imagine you're pulling your normal 9-to-5 and you
```

know you need to grab some groceries on your way home. Well, since you're smart you've used Ericsson IoT platform to connect your car to the local grocery delivery service (Mat.se), so all you need to do is open the Mat.se app and make your usual order. Mat.se automatically handles the payment, grocery selection, delivery, and delivery scheduling. Since your car is IoT-enabled, Mat.se issues its trusted delivery agent a 1-time token to use for opening your car in order to place your groceries in your car for you at 4:40 pm (just before you get off from work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio, make sure to go watch the videos on this page.

Future Trends for IoT in 2018

New IoT device ideas won't do you much good unless you at least know the basic technology trends that are set to impact IoT over the next year(s). These include:

Big Data & Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it's going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.

Machine Learning Data Science: While a lot of IoT devices are still operated according to rules-based decision criteria, the age of artificial intelligence is upon us. IoT will increasingly depend on machine learning algorithms to control device operations so that devices are able to autonomously respond to a complex set of overlapping stimuli.

Blockchain-Enabled Security: Above all else, IoT networks must be secure. Blockchain technology is primed to meet the security demands that come along with building and expanding the IoT.

Best New IoT Device Ideas

This listing of new IoT device ideas has been sub-divided according to the main technology upon which the IoT devices are built. Below I'm providing a list of new IoT device ideas, but for detailed instructions on how to build these IoT applications, I recommend the IoT courses on Udemy (ß Please note: if you purchase a Udemy course through this link, I may receive a small commission), or courses that are available at SkyFi and Coursera.

Raspberry Pi IoT Ideas

Using Raspberry Pi as open-source hardware, you can build IoT applications that offer any one of the following benefits:

Enable built-in sensing to build a weather station that measures ambient temperature and humidity

Build a system that detects discrepancies in electrical readings to identify electricity theft

Use IoT to build a Servo that is controlled by motion detection readings

Build a smart control switch that operates devices based on external stimuli. Use this for home automation.

Build a music playing application that enables music for each room in your house

Implement biometrics on IoT-connected devices

Arduino IoT Ideas
There are a number of new IoT device ideas that deploy Arduino as a microcontroller. These include:

Integrate Arduino with Android to build a remote-control RGB LED device.
Connect PIR sensors across the IoT to implement a smart building.
Build a temperature and sunlight sensor system to remotely monitor and control the conditions of your garden.
Deploy Arduino and IoT to automate your neighborhood streetlights.
Build a smart irrigation system based on IoT-connected temperature and moisture sensors built-in to your agricultural plants.

[caption id="attachment_3807" align="aligncenter" width="300"] An IoT Chatbot Tree at the Ericsson Studio[/caption]
Wireless (GSM) IoT Ideas
Several new IoT device ideas are developed around the GSM wireless network. Those are:

Monitor soil moisture to automate agricultural irrigation cycles.
Automate and control the conditions of a greenhouse.
Enable bio-metrics to build a smart security system for your home or office building
Build an autonomously operating fitness application that automatically makes recommendations based on motion detection and heart rate sensors that are embedded on wearable fitness trackers.
Build a healthcare monitoring system that tracks, informs, and automatically alerts healthcare providers based on sensor readings that describe a patients vital statistics (like temperature, pulse, blood pressure, etc).

IoT Automation Ideas
Almost all new IoT device ideas offer automation benefits, but to outline a few more ideas:

Build an IoT device that automatically locates and reports the closest nearby parking spot.
Build a motion detection system that automatically issues emails or sms messages to alert home owners of a likely home invasion.
Use temperature sensors connected across the IoT to automatically alert you if your home windows or doors have been left open.
Use bio-metric sensors to build a smart system that automate security for your home or office building

To learn more about IoT and what's happening on the leading edge, be sure to pop over to Ericsson's Studio Tour recap and watch these videos.

(I captured some of this content on behalf of DevMode Strategies during an

invite-only tour of the Ericsson Studio in Kista. Rest assure, the text
and opinions are my own)
…

1.3.3 Searching and retrieving data from a parse tree

Retrieving tags by filtering with name arguments

[48]:
```
soup.find_all("li")
```

[48]:
```
[<li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are embedded
within IoT devices spin off machine-generated data like it's going out of style.
For IoT to function, the platform must be solidly engineered to handle big data.
Be assured, that requires some serious data engineering.</li>,
 <li><strong>Machine Learning</strong> Data Science: While a lot of IoT devices
are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine
learning algorithms to control device operations so that devices are able to
autonomously respond to a complex set of overlapping stimuli.</li>,
 <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands
that come along with building and expanding the IoT.</li>,
 <li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>,
 <li>Build a system that detects discrepancies in electrical readings to
identify electricity theft</li>,
 <li>Use IoT to build a Servo that is controlled by motion detection
readings</li>,
 <li>Build a smart control switch that operates devices based on external
stimuli. Use this for home automation.</li>,
 <li>Build a music playing application that enables music for each room in your
house</li>,
 <li>Implement biometrics on IoT-connected devices</li>,
 <li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>,
 <li>Connect PIR sensors across the IoT to implement a smart building.</li>,
 <li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>,
 <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>,
 <li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>,
 <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>,
 <li>Automate and control the conditions of a greenhouse.</li>,
 <li>Enable bio-metrics to build a smart security system for your home or office
building</li>,
 <li>Build an autonomously operating fitness application that automatically
makes recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>,
 <li>Build a healthcare monitoring system that tracks, informs, and
```

```
automatically alerts healthcare providers based on sensor readings that describe
a patients vital statistics (like temperature, pulse, blood pressure,
etc).</li>,
 <li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>,
 <li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>,
 <li>Use temperature sensors connected across the IoT to automatically alert you
if your home windows or doors have been left open.</li>,
 <li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>]
```

Retrieving tags by filtering with keyword arguments

[50]:
```
soup.find_all(id="link 7")
```

[50]:
```
[<a class="preview" href="http://www.skyfilabs.com/iot-online-courses" id="link
7">SkyFi</a>]
```

Retrieving tags by filtering with string arguments

[51]:
```
soup.find_all('ol')
```

[51]:
```
[<ol>
 <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are embedded
within IoT devices spin off machine-generated data like it's going out of style.
For IoT to function, the platform must be solidly engineered to handle big data.
Be assured, that requires some serious data engineering.</li>
 <li><strong>Machine Learning</strong> Data Science: While a lot of IoT devices
are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine
learning algorithms to control device operations so that devices are able to
autonomously respond to a complex set of overlapping stimuli.</li>
 <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands
that come along with building and expanding the IoT.</li>
 </ol>,
 <ol>
 <li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>
 <li>Build a system that detects discrepancies in electrical readings to
identify electricity theft</li>
 <li>Use IoT to build a Servo that is controlled by motion detection
readings</li>
 <li>Build a smart control switch that operates devices based on external
stimuli. Use this for home automation.</li>
 <li>Build a music playing application that enables music for each room in your
house</li>
 <li>Implement biometrics on IoT-connected devices</li>
```

```
 </ol>,
 <ol>
 <li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>
 <li>Connect PIR sensors across the IoT to implement a smart building.</li>
 <li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>
 <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
 <li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>
 </ol>,
 <ol>
 <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
 <li>Automate and control the conditions of a greenhouse.</li>
 <li>Enable bio-metrics to build a smart security system for your home or office
building</li>
 <li>Build an autonomously operating fitness application that automatically
makes recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>
 <li>Build a healthcare monitoring system that tracks, informs, and
automatically alerts healthcare providers based on sensor readings that describe
a patients vital statistics (like temperature, pulse, blood pressure, etc).</li>
 </ol>,
 <ol>
 <li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>
 <li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>
 <li>Use temperature sensors connected across the IoT to automatically alert you
if your home windows or doors have been left open.</li>
 <li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>
 </ol>]
```

Retrieving tags by filtering with list objects

[52]:
```
soup.find_all(['ol', 'b'])
```

[52]:
```
[<b>2018 Trends: Best New IoT Device Ideas for Data Scientists and
Engineers</b>,
 <ol>
 <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are embedded
within IoT devices spin off machine-generated data like it's going out of style.
For IoT to function, the platform must be solidly engineered to handle big data.
Be assured, that requires some serious data engineering.</li>
 <li><strong>Machine Learning</strong> Data Science: While a lot of IoT devices
are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine
```

```
learning algorithms to control device operations so that devices are able to
autonomously respond to a complex set of overlapping stimuli.</li>
 <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands
that come along with building and expanding the IoT.</li>
 </ol>,
 <ol>
 <li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>
 <li>Build a system that detects discrepancies in electrical readings to
identify electricity theft</li>
 <li>Use IoT to build a Servo that is controlled by motion detection
readings</li>
 <li>Build a smart control switch that operates devices based on external
stimuli. Use this for home automation.</li>
 <li>Build a music playing application that enables music for each room in your
house</li>
 <li>Implement biometrics on IoT-connected devices</li>
 </ol>,
 <ol>
 <li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>
 <li>Connect PIR sensors across the IoT to implement a smart building.</li>
 <li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>
 <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
 <li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>
 </ol>,
 <ol>
 <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
 <li>Automate and control the conditions of a greenhouse.</li>
 <li>Enable bio-metrics to build a smart security system for your home or office
building</li>
 <li>Build an autonomously operating fitness application that automatically
makes recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>
 <li>Build a healthcare monitoring system that tracks, informs, and
automatically alerts healthcare providers based on sensor readings that describe
a patients vital statistics (like temperature, pulse, blood pressure, etc).</li>
 </ol>,
 <ol>
 <li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>
 <li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>
 <li>Use temperature sensors connected across the IoT to automatically alert you
```

```
if your home windows or doors have been left open.</li>
 <li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>
 </ol>]
```

Retrieving tags by filtering with regular expressions

[54]:
```
import re
t = re.compile("t")
for tag in soup.find_all(t):
    print(tag.name)
```

```
html
title
strong
strong
strong
strong
strong
strong
```

[55]:
```
with urllib.request.urlopen('https://raw.githubusercontent.com/BigDataGal/
↪Data-Mania-Demos/master/IoT-2018.html') as response:
    html = response.read()
```

[56]:
```
soup = BeautifulSoup(html, "lxml")
type(soup)
```

[56]:
```
bs4.BeautifulSoup
```

1.3.4 Parsing your data

[57]:
```
print(soup.prettify()[0:100])
```

```
<html>
 <head>
  <title>
   IoT Articles
  </title>
 </head>
 <body>
  <p class="title">
   <b>
```

1.3.5 Getting data from a parse tree

[58]:
```
text_only = soup.get_text()
print(text_only)
```

```
IoT Articles

2018 Trends: Best New IoT Device Ideas for Data Scientists and Engineers
It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use…


It's almost 2018 and IoT is on the cusp of an explosive expansion. In this
article, I offer you a listing of new IoT device ideas that you can use to get
practice in designing your first IoT applications.
Looking Back at My Coolest IoT Find in 2017
Before going into detail about best new IoT device ideas, here's the backstory.
Last month Ericsson Digital invited me to tour the Ericsson Studio in Kista,
Sweden. Up until that visit, IoT had been largely theoretical to me. Of course,
I know the usual mumbo-jumbo about wearables and IoT-connected fitness trackers.
That stuff is all well and good, but it's somewhat old hat - plus I am not sure
we are really benefiting so much from those, so I'm not that impressed.

It wasn't until I got to the Ericsson Studio that I became extremely impressed
by how far IoT has really come. Relying on the promise of the 5g network
expansion, IoT-powered smart devices are on the cusp of an explosive growth in
adoption. It was Ericsson's Smart Car that sent me reeling:

This car is connected to Ericsson's Connected Vehicle Cloud, an IoT platform
that manages services for the Smart Cars to which it's connected. The Volvo
pictured above acts as a drop-off location for groceries that have been ordered
by its owner.

To understand how it works, imagine you're pulling your normal 9-to-5 and you
know you need to grab some groceries on your way home. Well, since you're smart
you've used Ericsson IoT platform to connect your car to the local grocery
delivery service (Mat.se), so all you need to do is open the Mat.se app and make
your usual order. Mat.se automatically handles the payment, grocery selection,
delivery, and delivery scheduling. Since your car is IoT-enabled, Mat.se issues
its trusted delivery agent a 1-time token to use for opening your car in order
to place your groceries in your car for you at 4:40 pm (just before you get off
from work).

To watch some of the amazing IoT device demos I witnessed at Ericsson Studio,
make sure to go watch the videos on this page.
Future Trends for IoT in 2018
New IoT device ideas won't do you much good unless you at least know the basic
technology trends that are set to impact IoT over the next year(s). These
```

include:

Big Data & Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it's going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.
Machine Learning Data Science: While a lot of IoT devices are still operated according to rules-based decision criteria, the age of artificial intelligence is upon us. IoT will increasingly depend on machine learning algorithms to control device operations so that devices are able to autonomously respond to a complex set of overlapping stimuli.
Blockchain-Enabled Security: Above all else, IoT networks must be secure. Blockchain technology is primed to meet the security demands that come along with building and expanding the IoT.

Best New IoT Device Ideas
This listing of new IoT device ideas has been sub-divided according to the main technology upon which the IoT devices are built. Below I'm providing a list of new IoT device ideas, but for detailed instructions on how to build these IoT applications, I recommend the IoT courses on Udemy (ß Please note: if you purchase a Udemy course through this link, I may receive a small commission), or courses that are available at SkyFi and Coursera.
Raspberry Pi IoT Ideas
Using Raspberry Pi as open-source hardware, you can build IoT applications that offer any one of the following benefits:

Enable built-in sensing to build a weather station that measures ambient temperature and humidity
Build a system that detects discrepancies in electrical readings to identify electricity theft
Use IoT to build a Servo that is controlled by motion detection readings
Build a smart control switch that operates devices based on external stimuli. Use this for home automation.
Build a music playing application that enables music for each room in your house
Implement biometrics on IoT-connected devices

Arduino IoT Ideas
There are a number of new IoT device ideas that deploy Arduino as a microcontroller. These include:

Integrate Arduino with Android to build a remote-control RGB LED device.
Connect PIR sensors across the IoT to implement a smart building.
Build a temperature and sunlight sensor system to remotely monitor and control the conditions of your garden.
Deploy Arduino and IoT to automate your neighborhood streetlights.
Build a smart irrigation system based on IoT-connected temperature and moisture sensors built-in to your agricultural plants.

[caption id="attachment_3807" align="aligncenter" width="300"] An IoT Chatbot Tree at the Ericsson Studio[/caption]
Wireless (GSM) IoT Ideas
Several new IoT device ideas are developed around the GSM wireless network. Those are:

Monitor soil moisture to automate agricultural irrigation cycles.
Automate and control the conditions of a greenhouse.
Enable bio-metrics to build a smart security system for your home or office building
Build an autonomously operating fitness application that automatically makes recommendations based on motion detection and heart rate sensors that are embedded on wearable fitness trackers.
Build a healthcare monitoring system that tracks, informs, and automatically alerts healthcare providers based on sensor readings that describe a patients vital statistics (like temperature, pulse, blood pressure, etc).

IoT Automation Ideas
Almost all new IoT device ideas offer automation benefits, but to outline a few more ideas:

Build an IoT device that automatically locates and reports the closest nearby parking spot.
Build a motion detection system that automatically issues emails or sms messages to alert home owners of a likely home invasion.
Use temperature sensors connected across the IoT to automatically alert you if your home windows or doors have been left open.
Use bio-metric sensors to build a smart system that automate security for your home or office building

To learn more about IoT and what's happening on the leading edge, be sure to pop over to Ericsson's Studio Tour recap and watch these videos.

(I captured some of this content on behalf of DevMode Strategies during an invite-only tour of the Ericsson Studio in Kista. Rest assure, the text and opinions are my own)
…

1.3.6 Searching and retrieving data from a parse tree

Retrieving tags by filtering with name arguments

[60]:
```
soup.find_all("li")
```

[60]: [
Big Data & Data Engineering: Sensors that are embedded within IoT devices spin off machine-generated data like it's going out of style. For IoT to function, the platform must be solidly engineered to handle big data. Be assured, that requires some serious data engineering.,
Machine Learning Data Science: While a lot of IoT devices

```
are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine
learning algorithms to control device operations so that devices are able to
autonomously respond to a complex set of overlapping stimuli.</li>,
 <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands
that come along with building and expanding the IoT.</li>,
 <li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>,
 <li>Build a system that detects discrepancies in electrical readings to
identify electricity theft</li>,
 <li>Use IoT to build a Servo that is controlled by motion detection
readings</li>,
 <li>Build a smart control switch that operates devices based on external
stimuli. Use this for home automation.</li>,
 <li>Build a music playing application that enables music for each room in your
house</li>,
 <li>Implement biometrics on IoT-connected devices</li>,
 <li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>,
 <li>Connect PIR sensors across the IoT to implement a smart building.</li>,
 <li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>,
 <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>,
 <li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>,
 <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>,
 <li>Automate and control the conditions of a greenhouse.</li>,
 <li>Enable bio-metrics to build a smart security system for your home or office
building</li>,
 <li>Build an autonomously operating fitness application that automatically
makes recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>,
 <li>Build a healthcare monitoring system that tracks, informs, and
automatically alerts healthcare providers based on sensor readings that describe
a patients vital statistics (like temperature, pulse, blood pressure,
etc).</li>,
 <li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>,
 <li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>,
 <li>Use temperature sensors connected across the IoT to automatically alert you
if your home windows or doors have been left open.</li>,
 <li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>]
```

Retrieving tags by filtering with keyword arguments

[62]:
```
soup.find_all(id="link 7")
```

[62]:
```
[<a class="preview" href="http://www.skyfilabs.com/iot-online-courses" id="link
7">SkyFi</a>]
```

Retrieving tags by filtering with string arguments

[64]:
```
soup.find_all('ol')
```

[64]:
```
[<ol>
 <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are embedded
within IoT devices spin off machine-generated data like it's going out of style.
For IoT to function, the platform must be solidly engineered to handle big data.
Be assured, that requires some serious data engineering.</li>
 <li><strong>Machine Learning</strong> Data Science: While a lot of IoT devices
are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine
learning algorithms to control device operations so that devices are able to
autonomously respond to a complex set of overlapping stimuli.</li>
 <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands
that come along with building and expanding the IoT.</li>
 </ol>,
 <ol>
 <li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>
 <li>Build a system that detects discrepancies in electrical readings to
identify electricity theft</li>
 <li>Use IoT to build a Servo that is controlled by motion detection
readings</li>
 <li>Build a smart control switch that operates devices based on external
stimuli. Use this for home automation.</li>
 <li>Build a music playing application that enables music for each room in your
house</li>
 <li>Implement biometrics on IoT-connected devices</li>
 </ol>,
 <ol>
 <li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>
 <li>Connect PIR sensors across the IoT to implement a smart building.</li>
 <li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>
 <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
 <li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>
 </ol>,
 <ol>
 <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
```

```
 <li>Automate and control the conditions of a greenhouse.</li>
 <li>Enable bio-metrics to build a smart security system for your home or office
building</li>
 <li>Build an autonomously operating fitness application that automatically
makes recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>
 <li>Build a healthcare monitoring system that tracks, informs, and
automatically alerts healthcare providers based on sensor readings that describe
a patients vital statistics (like temperature, pulse, blood pressure, etc).</li>
 </ol>,
 <ol>
 <li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>
 <li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>
 <li>Use temperature sensors connected across the IoT to automatically alert you
if your home windows or doors have been left open.</li>
 <li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>
 </ol>]
```

Retrieving tags by filtering with list objects

[65]:
```
soup.find_all(['ol', 'b'])
```

[65]:
```
[<b>2018 Trends: Best New IoT Device Ideas for Data Scientists and
Engineers</b>,
 <ol>
 <li><strong>Big Data</strong> &amp;amp; Data Engineering: Sensors that are embedded
within IoT devices spin off machine-generated data like it’s going out of style.
For IoT to function, the platform must be solidly engineered to handle big data.
Be assured, that requires some serious data engineering.</li>
 <li><strong>Machine Learning</strong> Data Science: While a lot of IoT devices
are still operated according to rules-based decision criteria, the age of
artificial intelligence is upon us. IoT will increasingly depend on machine
learning algorithms to control device operations so that devices are able to
autonomously respond to a complex set of overlapping stimuli.</li>
 <li><strong>Blockchain</strong>-Enabled Security: Above all else, IoT networks
must be secure. Blockchain technology is primed to meet the security demands
that come along with building and expanding the IoT.</li>
 </ol>,
 <ol>
 <li>Enable built-in sensing to build a weather station that measures ambient
temperature and humidity</li>
 <li>Build a system that detects discrepancies in electrical readings to
identify electricity theft</li>
 <li>Use IoT to build a Servo that is controlled by motion detection
readings</li>
```

```
 <li>Build a smart control switch that operates devices based on external
stimuli. Use this for home automation.</li>
 <li>Build a music playing application that enables music for each room in your
house</li>
 <li>Implement biometrics on IoT-connected devices</li>
 </ol>,
 <ol>
 <li>Integrate Arduino with Android to build a remote-control RGB LED
device.</li>
 <li>Connect PIR sensors across the IoT to implement a smart building.</li>
 <li>Build a temperature and sunlight sensor system to remotely monitor and
control the conditions of your garden.</li>
 <li>Deploy Arduino and IoT to automate your neighborhood streetlights.</li>
 <li>Build a smart irrigation system based on IoT-connected temperature and
moisture sensors built-in to your agricultural plants.</li>
 </ol>,
 <ol>
 <li>Monitor soil moisture to automate agricultural irrigation cycles.</li>
 <li>Automate and control the conditions of a greenhouse.</li>
 <li>Enable bio-metrics to build a smart security system for your home or office
building</li>
 <li>Build an autonomously operating fitness application that automatically
makes recommendations based on motion detection and heart rate sensors that are
embedded on wearable fitness trackers.</li>
 <li>Build a healthcare monitoring system that tracks, informs, and
automatically alerts healthcare providers based on sensor readings that describe
a patients vital statistics (like temperature, pulse, blood pressure, etc).</li>
 </ol>,
 <ol>
 <li>Build an IoT device that automatically locates and reports the closest
nearby parking spot.</li>
 <li>Build a motion detection system that automatically issues emails or sms
messages to alert home owners of a likely home invasion.</li>
 <li>Use temperature sensors connected across the IoT to automatically alert you
if your home windows or doors have been left open.</li>
 <li>Use bio-metric sensors to build a smart system that automate security for
your home or office building</li>
 </ol>]
```

[66]:
```
##### Retrieving tags by filtering with regular expressions
```

[67]:
```
t = re.compile("t")
for tag in soup.find_all(t):
    print(tag.name)
```

```
html
title
strong
```

```
strong
strong
strong
strong
strong
```

Retrieving tags by filtering with a Boolean value

[69]:
```
for tag in soup.find_all(True):
    print(tag.name)
```

```
html
head
title
body
p
b
p
br
br
h1
span
strong
a
a
a
img
a
span
strong
a
h1
ol
li
strong
li
strong
li
strong
h1
a
a
a
h2
ol
li
li
li
li
```

```
li
li
h2
ol
li
li
li
li
li
a
img
h2
ol
li
li
li
li
li
h2
ol
li
li
li
li
span
strong
a
em
p
```

Retrieving weblinks by filtering with string objects

[70]:
```
for link in soup.find_all('a'):
    print(link.get('href'))
```

```
http://bit.ly/LPlNDJj
http://www.data-mania.com/blog/m2m-vs-iot/
bit.ly/LPlNDJj
http://mat.se/
http://bit.ly/LPlNDJj
https://click.linksynergy.com/deeplink?id=*JDLXjeE*wk&mid=39197&murl=https%3A%2F
%2Fwww.udemy.com%2Ftopic%2Finternet-of-things%2F%3Fsort%3Dhighest-rated
http://www.skyfilabs.com/iot-online-courses
https://www.coursera.org/specializations/iot
bit.ly/LPlNDJj
http://bit.ly/LPlNDJj
```

Retrieving strings by filtering with regular expressions

[72]:
```
soup.find_all(string=re.compile("data"))
```

[72]:
```
[' & Data Engineering: Sensors that are embedded within IoT devices spin off
machine-generated data like it's going out of style. For IoT to function, the
platform must be solidly engineered to handle big data. Be assured, that
requires some serious data engineering.']
```

1.4 Segment 4 - Web scraping

[73]:
```
from IPython.display import HTML
```

[74]:
```
r = urllib.request.urlopen('https://analytics.usa.gov/').read()
soup = BeautifulSoup(r, "lxml")
type(soup)
```

[74]:
```
bs4.BeautifulSoup
```

[75]:
```
print(soup.prettify()[:100])
```

```
<!DOCTYPE html>
<html lang="en">
 <!-- Initalize title and data source variables -->
 <head>
  <!--
```

[76]:
```
for link in soup.find_all('a'):
    print(link.get('href'))
```

```
/
#explanation
https://analytics.usa.gov/data/
https://open.gsa.gov/api/dap/
data/
#top-pages-realtime
#top-pages-7-days
#top-pages-30-days
https://analytics.usa.gov/data/live/all-pages-realtime.csv
https://analytics.usa.gov/data/live/all-domains-30-days.csv
https://www.digitalgov.gov/services/dap/
https://www.digitalgov.gov/services/dap/common-questions-about-dap-faq/#part-4
https://support.google.com/analytics/answer/2763052?hl=en
https://analytics.usa.gov/data/live/second-level-domains.csv
https://analytics.usa.gov/data/live/sites.csv
mailto:dap@gsa.gov
https://analytics.usa.gov/data/
https://open.gsa.gov/api/dap/
mailto:dap@gsa.gov
```

```
https://github.com/GSA/analytics.usa.gov/issues
https://github.com/GSA/analytics.usa.gov
https://github.com/18F/analytics-reporter
http://www.gsa.gov/
https://www.digital.gov/guides/dap/
https://cloud.gov/
```

[77]:
```
print(soup.get_text())
```

```
analytics.usa.gov | The US government's web traffic.
```

```
analytics.usa.gov
```

About this site
Data | API

Select an agency

All Participating Websites
Agency for International Development
Department of Agriculture
Department of Commerce
Department of Defense
Department of Education
Department of Energy
Department of Health and Human Services
Department of Homeland Security
Department of Housing and Urban Development
Department of Justice
Department of Labor
Department of State
Department of Transportation
Department of Veterans Affairs
Department of the Interior
Department of the Treasury
Environmental Protection Agency
Executive Office of the President
General Services Administration
National Aeronautics and Space Administration
National Archives and Records Administration
National Science Foundation
Nuclear Regulatory Commission
Office of Personnel Management
Postal Service
Small Business Administration
Social Security Administration

…

people on government websites now

Visits Today
Eastern Time

Visits in the Past 90 Days

There were … visits over the past 90 days.

Devices

Based on rough network segmentation data, we estimate that less than 5% of all traffic across all agencies comes from US federal government networks.

Much more detailed data is available in downloadable CSV and JSON. This includes data on combined browser and OS usage.

Browsers

Internet Explorer

Operating Systems

Windows

Visitor Locations Right Now

Cities

Countries

United States & Territories

International

Top Pages

Now
7 Days
30 Days

People on a single, specific page now. We only count pages with at least 10 people on the page.

Download the full dataset.

Visits over the last week to domains, including traffic to all pages within that domain.

Visits over the last month to domains, including traffic to all pages within that domain. We only count pages with at least 1,000 visits in the last month.

Download the full dataset.

Top Downloads

Total file downloads yesterday on government domains.

About this Site

These data provide a window into how people are interacting with the government online.

The data come from a unified Google Analytics account for U.S. federal government agencies known as the Digital Analytics Program.

This program helps government agencies understand how people find, access, and use government services online. The program does not track individuals,

and anonymizes the IP addresses of visitors.

Not every government website is represented in these data.

Currently, the Digital Analytics Program collects web traffic from around 400 executive branch government domains,
across about 5,700 total websites,
including every cabinet department.
We continue to pursue and add more sites frequently; to add your site, email the Digital Analytics Program.

Download the data

You can download the data here. Available in JSON and CSV format.
Additionally, you can access data via our API project (currently in Beta).

A note on sampling

Due to varying Google Analytics API sampling thresholds and the sheer volume of data in this project, some non-realtime reports may be subject to sampling.
The data are intended to represent trends and numbers may not be precise.

Have a question or problem?

Get in touch.

Suggest a feature or report an issue

View our code on GitHub

View our code for the data on GitHub

```
Analytics.usa.gov is a project of GSA's Digital Analytics Program.
This website is hosted on cloud.gov.
```

[78]:
```
print(soup.prettify()[0:1000])
```

```
<!DOCTYPE html>
<html lang="en">
 <!-- Initalize title and data source variables -->
 <head>
  <!--

    Hi! Welcome to our source code.

    This dashboard uses data from the Digital Analytics Program, a US
    government team inside the General Services Administration.


    For a detailed tech breakdown of how 18F and friends built this site:

    https://18f.gsa.gov/2015/03/19/how-we-built-analytics-usa-gov/


    This is a fully open source project, and your contributions are welcome.

    Frontend static site: https://github.com/18F/analytics.usa.gov
    Backend data reporting: https://github.com/18F/analytics-reporter

    -->
  <meta charset="utf-8"/>
  <meta content="IE=Edge" http-equiv="X-UA-Compatible"/>
  <meta content="NjbZn6hQe7OwV-nTsa6nLmtrOUcSGPRyFjxm5zkmCcg" name="google-site-
verification"/>
  <link href="/css/vendor/css/uswds.v0.9.6.css" rel="stylesheet"/>
  <link href="/css/public_analytics.css" rel="stylesheet"/>
  <link href="/images/analytics-favicon.ico" rel="ic
```

[79]:
```
for link in soup.findAll('a', attrs={'href': re.compile("^http")}):
    print(link)
type(link)
```

```
<a href="https://analytics.usa.gov/data/">Data</a>
<a href="https://open.gsa.gov/api/dap/" rel="noopener" target="_blank">API</a>
<a href="https://analytics.usa.gov/data/live/all-pages-realtime.csv">Download
the full dataset.</a>
<a href="https://analytics.usa.gov/data/live/all-domains-30-days.csv">Download
the full dataset.</a>
<a class="external-link" href="https://www.digitalgov.gov/services/dap/">Digital
Analytics Program</a>
<a class="external-link" href="https://www.digitalgov.gov/services/dap/common-
questions-about-dap-faq/#part-4">does not track individuals</a>
<a class="external-link"
href="https://support.google.com/analytics/answer/2763052?hl=en">anonymizes the
IP addresses</a>
<a class="external-link" href="https://analytics.usa.gov/data/live/second-level-
domains.csv">400 executive branch government domains</a>
<a class="external-link"
href="https://analytics.usa.gov/data/live/sites.csv">about 5,700 total
websites</a>
<a href="https://analytics.usa.gov/data/">download the data here.</a>
<a href="https://open.gsa.gov/api/dap/" rel="noopener" target="_blank"> API
project</a>
<a class="usa-button usa-button-secondary-inverse"
href="https://github.com/GSA/analytics.usa.gov/issues">
<img alt="Github Icon" class="github-icon" src="/images/github-logo-white.svg"/>
                  Suggest a feature or report an issue
            </a>
<a href="https://github.com/GSA/analytics.usa.gov">
<img alt="Github Icon" class="github-icon" src="/images/github-logo.svg"/>
              View our code on GitHub</a>
<a href="https://github.com/18F/analytics-reporter">
<img alt="Github Icon" class="github-icon" src="/images/github-logo.svg"/>
              View our code for the data on GitHub</a>
<a href="http://www.gsa.gov/">
<img alt="GSA" src="/images/gsa-logo.svg"/>
</a>
<a href="https://www.digital.gov/guides/dap/">Digital Analytics Program</a>
<a href="https://cloud.gov/">cloud.gov</a>
```

[79]:
```
bs4.element.Tag
```

[80]:
```
file = open("parsed_data.txt", "w")
for link in soup.findAll('a', attrs={'href': re.compile("^http")}):
    soup_link = str(link)
    print(soup_link)
```

```
    file.write(soup_link)
file.flush()
file.close()
```

```
<a href="https://analytics.usa.gov/data/">Data</a>
<a href="https://open.gsa.gov/api/dap/" rel="noopener" target="_blank">API</a>
<a href="https://analytics.usa.gov/data/live/all-pages-realtime.csv">Download
the full dataset.</a>
<a href="https://analytics.usa.gov/data/live/all-domains-30-days.csv">Download
the full dataset.</a>
<a class="external-link" href="https://www.digitalgov.gov/services/dap/">Digital
Analytics Program</a>
<a class="external-link" href="https://www.digitalgov.gov/services/dap/common-
questions-about-dap-faq/#part-4">does not track individuals</a>
<a class="external-link"
href="https://support.google.com/analytics/answer/2763052?hl=en">anonymizes the
IP addresses</a>
<a class="external-link" href="https://analytics.usa.gov/data/live/second-level-
domains.csv">400 executive branch government domains</a>
<a class="external-link"
href="https://analytics.usa.gov/data/live/sites.csv">about 5,700 total
websites</a>
<a href="https://analytics.usa.gov/data/">download the data here.</a>
<a href="https://open.gsa.gov/api/dap/" rel="noopener" target="_blank"> API
project</a>
<a class="usa-button usa-button-secondary-inverse"
href="https://github.com/GSA/analytics.usa.gov/issues">
<img alt="Github Icon" class="github-icon" src="/images/github-logo-white.svg"/>
                  Suggest a feature or report an issue
            </a>
<a href="https://github.com/GSA/analytics.usa.gov">
<img alt="Github Icon" class="github-icon" src="/images/github-logo.svg"/>
              View our code on GitHub</a>
<a href="https://github.com/18F/analytics-reporter">
<img alt="Github Icon" class="github-icon" src="/images/github-logo.svg"/>
              View our code for the data on GitHub</a>
<a href="http://www.gsa.gov/">
<img alt="GSA" src="/images/gsa-logo.svg"/>
</a>
<a href="https://www.digital.gov/guides/dap/">Digital Analytics Program</a>
<a href="https://cloud.gov/">cloud.gov</a>
```

[81]:
```
%pwd
```

[81]:
```
'C:\\Users\\aadar\\Documents\\TERM2\\BDM 1034 - Application Design for Big
Data\\Week6\\Assignment'
```

1.5 Segment 5 - Introduction to NLP

[83]:
```
text = "On Wednesday, the Association for Computing Machinery, the world's␣
↪largest society of computing professionals, announced that Hinton, LeCun and␣
↪Bengio had won this year's Turing Award for their work on neural networks.␣
↪The Turing Award, which was introduced in 1966, is often called the Nobel␣
↪Prize of computing, and it includes a $1 million prize, which the three␣
↪scientists will share."
```

[85]:
```
import nltk
nltk.download('punkt')
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\aadar\AppData\Roaming\nltk_data…
[nltk_data]   Package punkt is already up-to-date!
```

[85]:
```
True
```

Sentence Tokenizer

[87]:
```
from nltk.tokenize import sent_tokenize
sent_tk = sent_tokenize(text)
print("Sentence tokenizing the text: \n")
print(sent_tk)
```

```
Sentence tokenizing the text:

['On Wednesday, the Association for Computing Machinery, the world's largest
society of computing professionals, announced that Hinton, LeCun and Bengio had
won this year's Turing Award for their work on neural networks.', 'The Turing
Award, which was introduced in 1966, is often called the Nobel Prize of
computing, and it includes a $1 million prize, which the three scientists will
share.']
```

1.5.1 Word Tokenizer

[89]:
```
from nltk.tokenize import word_tokenize
word_tk = word_tokenize(text)
print("Word tokenizing the text: \n")
print(word_tk)
```

```
Word tokenizing the text:

['On', 'Wednesday', ',', 'the', 'Association', 'for', 'Computing', 'Machinery',
',', 'the', 'world', ''', 's', 'largest', 'society', 'of', 'computing',
'professionals', ',', 'announced', 'that', 'Hinton', ',', 'LeCun', 'and',
'Bengio', 'had', 'won', 'this', 'year', ''', 's', 'Turing', 'Award', 'for',
'their', 'work', 'on', 'neural', 'networks', '.', 'The', 'Turing', 'Award', ',',
'which', 'was', 'introduced', 'in', '1966', ',', 'is', 'often', 'called', 'the',
'Nobel', 'Prize', 'of', 'computing', ',', 'and', 'it', 'includes', 'a', '$',
```

```
'1', 'million', 'prize', ',', 'which', 'the', 'three', 'scientists', 'will',
'share', '.']
```

1.5.2 Removing stop words

[91]:
```
nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\aadar\AppData\Roaming\nltk_data…
[nltk_data]   Package stopwords is already up-to-date!
```

[91]:
```
True
```

[92]:
```
from nltk.corpus import stopwords

sw = set(stopwords.words("english"))
print("Stop words in English language are: \n")
print(sw)
```

```
Stop words in English language are:

{'were', 'once', 'did', 'can', 't', 'having', 'own', 'hadn', 'just', "aren't",
'in', 'am', "mustn't", 'himself', 'have', 'wouldn', 'won', 'for', "don't", 've',
'but', 'my', 'm', "mightn't", 'here', "should've", 'shouldn', "won't", 'been',
"you'd", 'an', 'now', 'needn', 'below', 'out', 'yours', 'by', 'herself', 'more',
'aren', 'about', 'll', 'should', 'which', 'doesn', 'your', 'both', 'how',
'yourself', 'her', 'we', 'they', 'this', 'and', "doesn't", 'mustn', 'most',
"needn't", 'our', 'haven', 'same', "hasn't", 'being', 'his', 'of', 'are', 'a',
'she', 'me', 'from', "couldn't", "you'll", 'him', 'while', 'y', 'so', 'had',
"isn't", 'ours', 'it', 're', 'shan', 'into', 'these', 'weren', "you're",
'because', 'over', 'or', 'o', 'he', 'theirs', 'on', 'up', "haven't", 'with',
'be', 'ma', 'some', 'only', 'when', 'ain', "hadn't", 'does', 'mightn', 'again',
'what', 'yourselves', 'above', 'very', 'itself', 'no', 'off', "weren't", 'at',
'if', 's', 'do', "it's", 'i', "wouldn't", 'down', 'doing', 'there', 'their',
'the', 'd', 'further', 'through', 'is', 'other', 'ourselves', 'them', 'after',
'to', 'than', 'any', 'until', 'you', 'don', 'that', 'few', 'during', 'where',
"that'll", 'has', 'themselves', 'why', 'all', 'was', 'between', 'didn',
"didn't", 'hers', 'its', 'who', 'before', 'under', 'each', 'such', 'as',
'those', 'too', 'wasn', "wasn't", 'hasn', "shan't", "she's", 'against', 'isn',
"shouldn't", 'myself', "you've", 'nor', 'not', 'then', 'whom', 'will', 'couldn'}
```

[93]:
```
filtered_words = [w for w in word_tk if not w in sw]

print("The text after removing stop words \n")
print(filtered_words)
```

```
The text after removing stop words

['On', 'Wednesday', ',', 'Association', 'Computing', 'Machinery', ',', 'world',
```

```
'’', 'largest', 'society', 'computing', 'professionals', ',', 'announced',
'Hinton', ',', 'LeCun', 'Bengio', 'year', '’', 'Turing', 'Award', 'work',
'neural', 'networks', '.', 'The', 'Turing', 'Award', ',', 'introduced', '1966',
',', 'often', 'called', 'Nobel', 'Prize', 'computing', ',', 'includes', '$',
'1', 'million', 'prize', ',', 'three', 'scientists', 'share', '.']
```

Stemming

[95]:
```
from nltk.stem import PorterStemmer
from nltk.tokenize import sent_tokenize, word_tokenize

port_stem = PorterStemmer()
```

[96]:
```
stemmed_words = []

for w in filtered_words:
    stemmed_words.append(port_stem.stem(w))

print("Filtered Sentence: \n", filtered_words, "\n")
print("Stemmed Sentence: \n", stemmed_words)
```

```
Filtered Sentence:
 ['On', 'Wednesday', ',', 'Association', 'Computing', 'Machinery', ',', 'world',
'’', 'largest', 'society', 'computing', 'professionals', ',', 'announced',
'Hinton', ',', 'LeCun', 'Bengio', 'year', '’', 'Turing', 'Award', 'work',
'neural', 'networks', '.', 'The', 'Turing', 'Award', ',', 'introduced', '1966',
',', 'often', 'called', 'Nobel', 'Prize', 'computing', ',', 'includes', '$',
'1', 'million', 'prize', ',', 'three', 'scientists', 'share', '.']

Stemmed Sentence:
 ['on', 'wednesday', ',', 'associ', 'comput', 'machineri', ',', 'world', '’',
'largest', 'societi', 'comput', 'profession', ',', 'announc', 'hinton', ',',
'lecun', 'bengio', 'year', '’', 'ture', 'award', 'work', 'neural', 'network',
'.', 'the', 'ture', 'award', ',', 'introduc', '1966', ',', 'often', 'call',
'nobel', 'prize', 'comput', ',', 'includ', '$', '1', 'million', 'prize', ',',
'three', 'scientist', 'share', '.']
```

2 Lemmatizing

[98]:
```
nltk.download('wordnet')
```

```
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\aadar\AppData\Roaming\nltk_data…
[nltk_data]   Package wordnet is already up-to-date!
```

[98]: True

[99]:
```
from nltk.stem.wordnet import WordNetLemmatizer

lem = WordNetLemmatizer()

from nltk.stem.porter import PorterStemmer
stem = PorterStemmer()

lemm_words = []

for i in range(len(filtered_words)):
    lemm_words.append(lem.lemmatize(filtered_words[i]))

print(lemm_words)
```

```
['On', 'Wednesday', ',', 'Association', 'Computing', 'Machinery', ',', 'world',
'’', 'largest', 'society', 'computing', 'professional', ',', 'announced',
'Hinton', ',', 'LeCun', 'Bengio', 'year', '’', 'Turing', 'Award', 'work',
'neural', 'network', '.', 'The', 'Turing', 'Award', ',', 'introduced', '1966',
',', 'often', 'called', 'Nobel', 'Prize', 'computing', ',', 'includes', '$',
'1', 'million', 'prize', ',', 'three', 'scientist', 'share', '.']
```

Parts of Speech Tagging

[101]:
```
nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     C:\Users\aadar\AppData\Roaming\nltk_data…
[nltk_data]   Package averaged_perceptron_tagger is already up-to-
[nltk_data]       date!
```

[101]:
```
True
```

[102]:
```
from nltk import pos_tag
pos_tagged_words = pos_tag(word_tk)

print(pos_tagged_words)
```

```
[('On', 'IN'), ('Wednesday', 'NNP'), (',', ','), ('the', 'DT'), ('Association',
'NNP'), ('for', 'IN'), ('Computing', 'VBG'), ('Machinery', 'NNP'), (',', ','),
('the', 'DT'), ('world', 'NN'), ('’', 'NNP'), ('s', 'RB'), ('largest', 'JJS'),
('society', 'NN'), ('of', 'IN'), ('computing', 'VBG'), ('professionals', 'NNS'),
(',', ','), ('announced', 'VBD'), ('that', 'IN'), ('Hinton', 'NNP'), (',', ','),
('LeCun', 'NNP'), ('and', 'CC'), ('Bengio', 'NNP'), ('had', 'VBD'), ('won',
'VBN'), ('this', 'DT'), ('year', 'NN'), ('’', 'VBZ'), ('s', 'JJ'), ('Turing',
'NNP'), ('Award', 'NNP'), ('for', 'IN'), ('their', 'PRP$'), ('work', 'NN'),
('on', 'IN'), ('neural', 'JJ'), ('networks', 'NNS'), ('.', '.'), ('The', 'DT'),
('Turing', 'NNP'), ('Award', 'NNP'), (',', ','), ('which', 'WDT'), ('was',
'VBD'), ('introduced', 'VBN'), ('in', 'IN'), ('1966', 'CD'), (',', ','), ('is',
'VBZ'), ('often', 'RB'), ('called', 'VBN'), ('the', 'DT'), ('Nobel', 'NNP'),
```

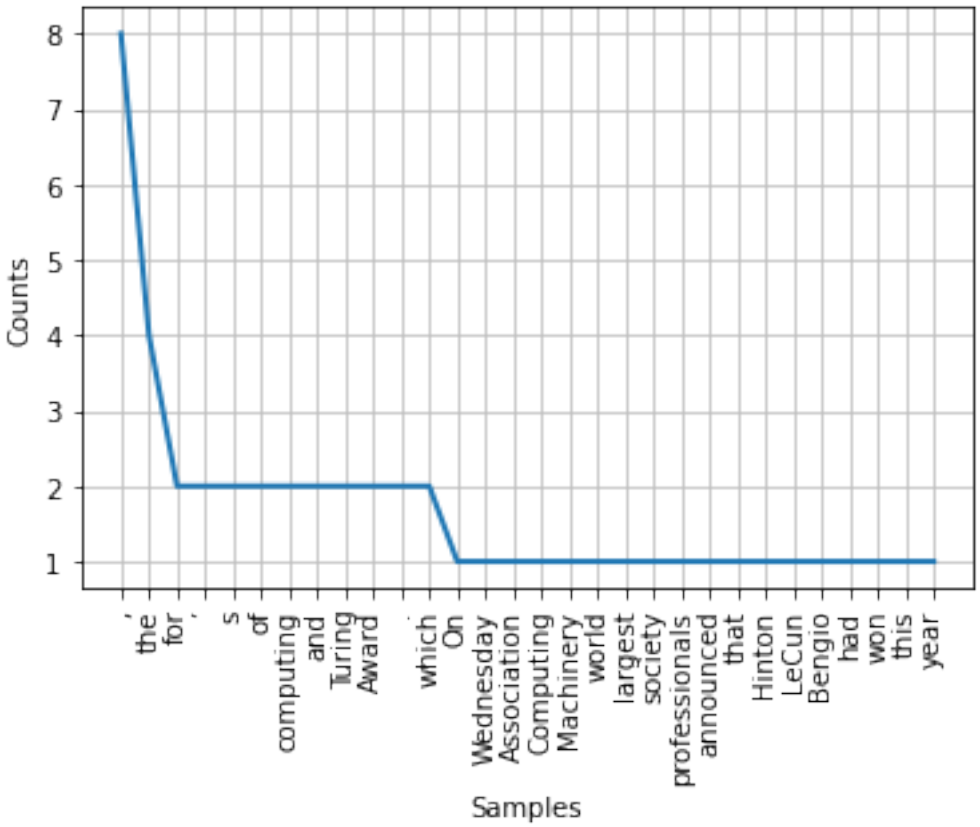
```
('Prize', 'NNP'), ('of', 'IN'), ('computing', 'NN'), (',', ','), ('and', 'CC'),
('it', 'PRP'), ('includes', 'VBZ'), ('a', 'DT'), ('$', '$'), ('1', 'CD'),
('million', 'CD'), ('prize', 'NN'), (',', ','), ('which', 'WDT'), ('the', 'DT'),
('three', 'CD'), ('scientists', 'NNS'), ('will', 'MD'), ('share', 'NN'), ('.',
'.')]
```

Frequency Distribution Plots

[104]:
```
from nltk.probability import FreqDist
fd = FreqDist(word_tk)
print(fd)
```

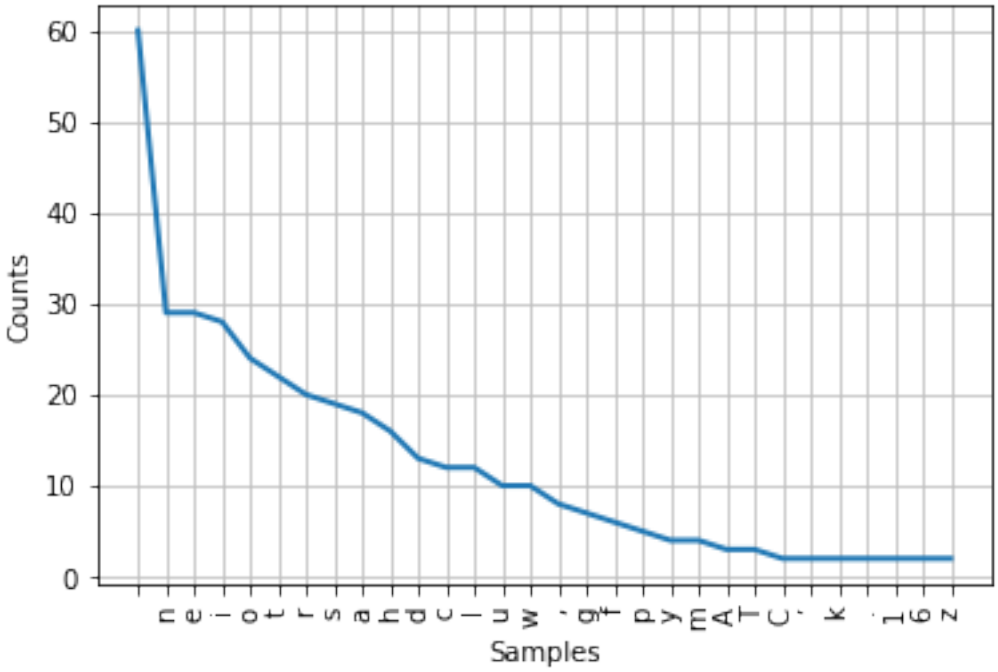
```
<FreqDist with 56 samples and 76 outcomes>
```

[105]:
```
import matplotlib.pyplot as plt
fd.plot(30, cumulative=False)
plt.show()
```



[106]:
```
fd_alpha = FreqDist(text)
print(fd_alpha)
fd_alpha.plot(30, cumulative=False)
```

```
<FreqDist with 41 samples and 387 outcomes>
```



[106]: <AxesSubplot:xlabel='Samples', ylabel='Counts'>

[ ]: