CM4125 Extra Topic 2: Web Scraping

Lecture objectives

- Learn how to get data even when people don't want us to get it!

Web scraping

The extraction of data from a website by programmatic means

Used when...

```
* No API
* No published structure dataset
* Multiple sources to be combined
* Scheduled sampling of data at given times
```

Consider the legal barriers!

- Some sites don't want to share data so easily! (e.g. BBC, City Council, Wikipedia, etc.)

- Licensing or permission may be given by copyright, creative commons or open gov license

To scrape data, the first thing we need is a target website

Let's use this one (https://www.transport.gov.scot/publication/key-reported-road-casualties-scotland-2018/3-reported-numbers-of-accidents-table-1) as an example

This page contains the number of fatalities in road accidents in Scotland from 1970 to 2018

In Python, I can use the `requests` and the `Beautiful Soup` libraries

In [2]:
```
# Import the necessary packages
import requests
from bs4 import BeautifulSoup
```

In [3]:
```
# Specify the target

url = "https://www.transport.gov.scot/publication/key-reported-road-casualties-scotland-2018/3-reported-numbers-of-accidents-table-1"
```

In [4]:

```
# Now, we do the request
r = requests.get(url)
print(r, type(r))
```

```
<Response [200]> <class 'requests.models.Response'>
```

In [5]:

```
# We can see the content of our request

print(r.content)
```

```
b'\r\n\r\n<!DOCTYPE html>\r\n\r\n<html lang="en" class="">\r\n\r\n<head>\r\n    <meta
charset="utf-8">\r\n    <meta name="viewport" content="width=device-width, initial-sc
ale=1.0">\r\n    <link rel="stylesheet" href="/content/css/styles.css">\r\n    <scrip
t>document.documentElement.className = document.documentElement.className.split(\'no-
js\').join(\'\');</script>\r\n    <meta name="apple-mobile-web-app-status-bar-style"
content="black-translucent">\r\n    <meta name="apple-touch-fullscreen" content="YES"
>\r\n    \r\n    \r\n<title>3. Reported numbers of Accidents (Table 1) | Transport Sc
otland</title>\r\n<meta name="description" content="" />\r\n<meta property="og:title"
content="" />\r\n<meta property="og:description" content="" />\r\n\r\n\r\n<meta name
="twitter:card" content="summary_large_image"/>\r\n<meta name="twitter:site" content
="@transcotland"/>\r\n<meta name="twitter:title" content=""/>\r\n<meta name="twitter:
description" content=""/>\r\n\r\n\r\n\r\n<link rel="canonical" href="https://www.tran
sport.gov.scot/publication/key-reported-road-casualties-scotland-2018/3-reported-numb
ers-of-accidents-table-1/" />\r\n    \r\n    \r\n\r\n    <meta name="_info" content="
Transport Scotland"\r\n      data-build="300.25.0+release-300.025+2022-01-19"\r\n
data-date="2022-01-19T12:12:34.0000000Z"\r\n      data-ver="300.25.0.0"\r\n      data
-uver="7.14.0" />\r\n    <!-- favicons -->\r\n    <link rel="apple-touch-icon" sizes
="180x180" href="/content/img/icons/apple-touch-icon.png">\r\n    <link rel="icon" ty
pe="image/png" href="/content/img/icons/favicon-32x32.png" sizes="32x32">\r\n    <lin
k rel="icon" type="image/png" href="/content/img/icons/favicon-16x16.png" sizes="16x1
6">\r\n    <link rel="manifest" href="/content/img/icons/manifest.json">\r\n    <link
rel="mask-icon" href="/content/img/icons/safari-pinned-tab.svg" color="#233e99">\r\n
<meta name="msapplication-TileColor" content="#233e99">\r\n    <meta name="msapplicat
ion-TileImage" content="/content/img/icons/mstile-144x144.png">\r\n    <meta name="th
eme-color" content="#19408e">\r\n    <meta name="google-site-verification" content="e
PiQ70hyLfSbH4lb8yXaSt7BOeceeQQx_cIikqaKJAc" />\r\n    \r\n\r\n\r\n    <!--App Insight
s-->\r\n    <script>\r\n        var appInsights=window.appInsights||function(confi
g){\r\n            function i(config){t[config]=function(){var i=arguments;t.queue.pu
sh(function(){t[config].apply(t,i)})}}var t={config:config},u=document,e=window,o="sc
ript",s="AuthenticatedUserContext",h="start",c="stop",l="Track",a=l+"Event",v=l+"Pag
e",y=u.createElement(o),r,f;y.src=config.url||"https://az416426.vo.msecnd.net/scripts
```

/a/ai.0.js";u.getElementsByTagName(o)[0].parentNode.appendChild(y);try{t.cookie=u.cookie}catch(p){}for(t.queue=[],t.version="1.0",r=["Event","Exception","Metric","PageView","Trace","Dependency"];r.length;)i("track"+r.pop());return i("set"+s),i("clear"+s),i(h+a),i(c+a),i(h+v),i(c+v),i("flush"),config.disableExceptionTracking||(r="onerror",i("_"+r),f=e[r],e[r]=function(config,i,u,e,o){var s=f&&f(config,i,u,e,o);return s!==!0&&t["_"+r](config,i,u,e,o),s}),t\r\n   }({\r\n   instrumentationKey:"c80db1a7-5b42-4b8f-a7d0-08377eeafd1a"\r\n   });\r\n   \r\n   window.appInsights=appInsights;\r\n   appInsights.trackPageView();\r\n   </script>\r\n   <!--App Insights-->\r\n\r\n</head>\r\n\r\n<body class="">\r\n\r\n\r\n\r\n \r\n   <a class="btn btn-skip" tabindex="0" href="#content">Skip to main content</a>\r\n \r\n\r\n\r\n \r\n\r\n   \r\n   This site uses cookies to store information on your computer. Some of these cookies are essential to make our site work and others help us to improve by giving us some insight into how the site is being used.\r\n   \r\n   \r\n\r\n\r\n\r\n \r\n\r\n<header class="header banner">\r\n\r\n   <nav class="wrapper relative">\r\n   <a href="/" class="logo">Transport Scotland, Comhdhail alba home button</a>\r\n   </nav>\r\n\r\n   \r\n   \r\n\r\n<nav class="quick-links">\r\n   \r\n   \r\n\r\n   <ul class="quick-links__list right">\r\n   \r\n<li class="">\r\n   \r\n   <a href="/about/">About</a>\r\n   \r\n   </li>\r\n   <li class="">\r\n   \r\n   <a href="/news/">News</a>\r\n   \r\n   </li>\r\n   <li class="">\r\n   \r\n   <a href="/consultations/">Consultations</a>\r\n   \r\n   </li>\r\n   <li class="">\r\n   \r\n<a href="/media-library/">Media Library</a>\r\n   \r\n   </li>\r\n   <li class="">\r\n   \r\n<a href="/publications/">Publications</a>\r\n   \r\n   </li>\r\n   <li class="">\r\n   \r\n   <a href="/traffic-management-updates/">Traffic management updates</a>\r\n   \r\n   </li>\r\n   <li class="">\r\n   \r\n   <a href="/coronavirus-covid-19/">Coronavirus (COVID-19)</a>\r\n   \r\n   </li>\r\n   <li class="flush--right">\r\n   \r\n   <a href="/cop26/">COP26</a>\r\n   \r\n

```
</li>\r\n        \r\n        </ul>\r\n\r\n    </div>\r\n\r\n</nav>\r\n\r\n    \r\n
<a href="#search" class="mobile-search js-toggle--search" title="Search" aria-label="
Search"></a>\r\n    <a href="#navigation" class="mobile-menu js-toggle" title="Naviga
tion menu" aria-label="Navigation menu"></a>\r\n\r\n    \r\n    <div class="nav-wrapp
er">\r\n    \r\n    <div class="wrapper">\r\n        \r\n        <nav id="navigation"
class="nav-primary">\r\n            \r\n            <ul class="nav-primary__list">\r\
n                \r\n                    <li class="nav-primary__item hidden">\r\n
\r\n                        <a href="/about/" class="nav-primary__link ">About</a>\r\
n                    \r\n                    </li>\r\n                    <li class="
nav-primary__item hidden">\r\n                        \r\n                        <a
href="/news/" class="nav-primary__link ">News</a>\r\n                    \r\n
</li>\r\n                    <li class="nav-primary__item ">\r\n
\r\n                        <a href="/our-approach/" class="nav-primary__link ">Our a
pproach</a>\r\n                    \r\n                    </li>\r\n
<li class="nav-primary__item ">\r\n                        \r\n
<a href="/public-transport/" class="nav-primary__link ">Public transport</a>\r\n
\r\n                    </li>\r\n                    <li class="nav-primary__item ">\
r\n                        \r\n                        <a href="/active-travel/" clas
s="nav-primary__link ">Walking and cycling</a>\r\n                    \r\n
</li>\r\n                    <li class="nav-primary__item ">\r\n
\r\n                        <a href="/transport-network/" class="nav-primary__link ">
Transport network</a>\r\n                    \r\n                    </li>\r\n
<li class="nav-primary__item ">\r\n                        \r\n
<a href="/projects/" class="nav-primary__link ">Projects</a>\r\n                    \
r\n                    </li>\r\n                    <li class="nav-primary__item ">\
r\n                        \r\n                        <a href="/concessionary-travel
/" class="nav-primary__link ">Concessionary travel</a>\r\n                    \r\n
</li>\r\n                    <li class="nav-primary__item hidden">\r\n
\r\n                        <a href="/consultations/" class="nav-primary__link ">Cons
ultations</a>\r\n                    \r\n                    </li>\r\n
<li class="nav-primary__item hidden">\r\n                        \r\n
<a href="/media-library/" class="nav-primary__link ">Media Library</a>\r\n
\r\n                    </li>\r\n                    <li class="nav-primary__item hid
den">\r\n                        \r\n                        <a href="/publications/"
class="nav-primary__link ">Publications</a>\r\n                    \r\n
</li>\r\n                    <li class="nav-primary__item hidden">\r\n
```

\r\n   <a href="/traffic-management-updates/" class="nav-primary__link ">Traffic management updates</a>\r\n   \r\n   </li>\r\n   <li class="nav-primary__item hidden">\r\n   \r\n   <a href="/coronavirus-covid-19/" class="nav-primary__link ">Coronavirus (COVID-19)</a>\r\n   \r\n   </li>\r\n   <li class="nav-primary__item hidden">\r\n   \r\n   <a href="/cop26/" class="nav-primary__link ">COP26</a>\r\n   \r\n   </li>\r\n   \r\n   </ul>\r\n\r\n   <button class="btn-close js-toggle" data-target="#navigation">Close</button>\r\n\r\n   </nav>\r\n\r\n \r\n\r\n \r\n\r\n   \r\n   \r\n   \r\n   \r\n   \r\n   <a href="#search" class="search-button js-toggle--search" title="Search" aria-label="Search"></a>\r\n\r\n   \r\n\r\n   \r\n   \r\n   \r\n   \r\n   <form action="/search-results/" method="get">\r\n   \r\n   <label class="visually-hidden" for="query">Search</label>\r\n   <input tabindex="-1" id="query" name="q" type="search" title="search" placeholder="Search for news, projects, policies etc"/>\r\n   <button tabindex="-1" title="search-button"></button>\r\n\r\n   </form>\r\n\r\n \r\n\r\n   \r\n\r\n \r\n\r\n\r\n</header>\r\n\r\n<main id="content">\r\n\r\n   \r\n\r\n\r\n\r\n \r\n   \r\n   \r\n   \r\n   \r\n\r\n   <nav aria-label="Breadcrumb" class="">\r\n   \r\n   \r\n\r\n\r\n<ul class="breadcrumb clearfix">\r\n   \r\n   <li><a href="/">Home</a></li>\r\n   <li aria-hidden="true"></li>\r\n   <li><a href="/publications/">Publications</a></li>\r\n   <li aria-hidden="true"></li>\r\n   <li><a href="/publication/key-reported-road-casualties-scotland-2018/">Key Reported Road Casualties Scotland 2018</a></li>\r\n   <li aria-hidden="true"></li>\r\n\r\n\r\n</ul>\r\n\r\n   </nav>\r\n   \r\n   \r\n \r\n\r\n   \r\n\r\n \r\n\r\n\r\n<article class="main-wrapper soft-double--bottom clearfix">\r\n\r\n   \r\n   \r\n   \r\n   \r\n   <h1 class="article__title">3. Reported numbers of Accidents (Table 1)</h1>\r\n   \r\n \r\n   \r\n   \r\n\r\n\r\n   <div class="wrapper soft-do

uble--bottom">\r\n\r\n \r\n\r\n <div class="col xs
mall-12">\r\n\r\n \r\n\r\n\r\n <div
class="umb-grid">\r\n \r\n \r\n
<div class=\'container\'>\r\n\r\n \r\n \r\n\r\n\r\n<div class
="">\r\n \r\n \n\t\t\t\t\t\t\t\t\t\t<h2>Key Reported Road Casualties Scotland
2018</h2>\n\t\t\t\t\t\t\t\t\t\t<h3>3. Reported numbers of Accidents (<a href="#Table1">
Table 1</a>)</h3>\n 3.1 <a href="#Table1">Table 1</a> shows the downward
trend of injury road\n accidents recorded by the police. In 2018,
there were 6,412 accidents in which someone was\n killed or injured, 10 per cent f
ewer than in 2017 and the lowest number since records began. There were 149\n fata
l accidents in 2018, 9 (6%) more than in 2017. In 2018, there were 1,368 serious inju
ry accidents - a\n decrease of 9 (1%) from 2017; and 4,895 slight injury accidents
reported in 2018, 13 per cent (706) fewer than\n 2017. \n <table id="Table1" b
order="1" cellpadding="5" cellspacing="1">\n <caption>\n Table 1: Injur
y Road Accidents by Severity, 1970 \xe2\x80\x93\n 2018\n </caption>\n
<thead>\n <tr>\n <th> </th>\n <th>\n Fatal \n
</th>\n <th>\n Serious \n </th>\n <th>\n
Fatal and Serious \n </th>\n <th>\n Slight \n
</th>\n <th>\n All \n </th>\n </tr>\n </thead>\n
<tbody>\n <tr>\n <th>\n 1970 \n </th>\n <td
>\n 758 \n </td>\n <td>\n 7,860 \n
</td>\n <td>\n 8,618 \n </td>\n <td>\n
13,515 \n </td>\n <td>\n 22,133 \n </td>\n
</tr>\n <tr>\n <th>\n 1975 \n </th>\n <td>\n
699 \n </td>\n <td>\n 6,912 \n </td>\n
<td>\n 7,611 \n </td>\n <td>\n 13,041 \n
</td>\n <td>\n 20,652 \n </td>\n </tr>\n <tr>\
n <th>\n 1980 \n </th>\n <td>\n 644</
p>\n </td>\n <td>\n 7,218 \n </td>\n <td>\
n 7,862 \n </td>\n <td>\n 13,926 \n
</td>\n <td>\n 21,788 \n </td>\n </tr>\n <tr>\
n <th>\n 1985 \n </th>\n <td>\n 550</
p>\n </td>\n <td>\n 6,507 \n </td>\n <td>\
n 7,057 \n </td>\n <td>\n 13,587 \n
</td>\n <td>\n 20,644 \n </td>\n </tr>\n <tr>\
n <th>\n 1990 \n </th>\n <td>\n 491</

p>\n   </td>\n   <td>\n   5,237 \n   </td>\n   <td>\
n   5,728 \n   </td>\n   <td>\n   14,443 \n
</td>\n   <td>\n   20,171 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   1995 \n   </th>\n   <td>\n   361</
p>\n   </td>\n   <td>\n   4,071 \n   </td>\n   <td>\
n   4,432 \n   </td>\n   <td>\n   12,102 \n
</td>\n   <td>\n   16,534 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   1996 \n   </th>\n   <td>\n   316</
p>\n   </td>\n   <td>\n   3,315 \n   </td>\n   <td>\
n   3,631 \n   </td>\n   <td>\n   12,442 \n
</td>\n   <td>\n   16,073 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   1997 \n   </th>\n   <td>\n   340</
p>\n   </td>\n   <td>\n   3,312 \n   </td>\n   <td>\
n   3,652 \n   </td>\n   <td>\n   12,994 \n
</td>\n   <td>\n   16,646 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   1998 \n   </th>\n   <td>\n   339</
p>\n   </td>\n   <td>\n   3,318 \n   </td>\n   <td>\
n   3,657 \n   </td>\n   <td>\n   12,862 \n
</td>\n   <td>\n   16,519 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   1999 \n   </th>\n   <td>\n   285</
p>\n   </td>\n   <td>\n   3,209 \n   </td>\n   <td>\
n   3,494 \n   </td>\n   <td>\n   11,921 \n
</td>\n   <td>\n   15,415 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2000 \n   </th>\n   <td>\n   297</
p>\n   </td>\n   <td>\n   3,007 \n   </td>\n   <td>\
n   3,304 \n   </td>\n   <td>\n   11,828 \n
</td>\n   <td>\n   15,132 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2001 \n   </th>\n   <td>\n   309</
p>\n   </td>\n   <td>\n   2,840 \n   </td>\n   <td>\
n   3,149 \n   </td>\n   <td>\n   11,575 \n
</td>\n   <td>\n   14,724 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2002 \n   </th>\n   <td>\n   274</
p>\n   </td>\n   <td>\n   2,684 \n   </td>\n   <td>\
n   2,958 \n   </td>\n   <td>\n   11,385 \n
</td>\n   <td>\n   14,343 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2003 \n   </th>\n   <td>\n   301</

p>\n   </td>\n   <td>\n   2,495 \n   </td>\n   <td>\
n   2,796 \n   </td>\n   <td>\n   11,121 \n
</td>\n   <td>\n   13,917 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2004 \n   </th>\n   <td>\n   283</
p>\n   </td>\n   <td>\n   2,331 \n   </td>\n   <td>\
n   2,614 \n   </td>\n   <td>\n   11,305 \n
</td>\n   <td>\n   13,919 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2005 \n   </th>\n   <td>\n   264</
p>\n   </td>\n   <td>\n   2,252 \n   </td>\n   <td>\
n   2,516 \n   </td>\n   <td>\n   10,922 \n
</td>\n   <td>\n   13,438 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2006 \n   </th>\n   <td>\n   293</
p>\n   </td>\n   <td>\n   2,257 \n   </td>\n   <td>\
n   2,550 \n   </td>\n   <td>\n   10,560 \n
</td>\n   <td>\n   13,110 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2007 \n   </th>\n   <td>\n   255</
p>\n   </td>\n   <td>\n   2,049 \n   </td>\n   <td>\
n   2,304 \n   </td>\n   <td>\n   10,203 \n
</td>\n   <td>\n   12,507 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2008 \n   </th>\n   <td>\n   245</
p>\n   </td>\n   <td>\n   2,242 \n   </td>\n   <td>\
n   2,487 \n   </td>\n   <td>\n   9,672 \n
</td>\n   <td>\n   12,159 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2009 \n   </th>\n   <td>\n   196</
p>\n   </td>\n   <td>\n   1,998 \n   </td>\n   <td>\
n   2,194 \n   </td>\n   <td>\n   9,362 \n
</td>\n   <td>\n   11,556 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2010 \n   </th>\n   <td>\n   189</
p>\n   </td>\n   <td>\n   1,713 \n   </td>\n   <td>\
n   1,902 \n   </td>\n   <td>\n   8,393 \n
</td>\n   <td>\n   10,295 \n   </td>\n   </tr>\n   <tr>\
n   <th>\n   2011 \n   </th>\n   <td>\n   175</
p>\n   </td>\n   <td>\n   1,675 \n   </td>\n   <td>\
n   1,850 \n   </td>\n   <td>\n   8,135 \n
</td>\n   <td>\n   9,985 \n   </td>\n   </tr>\n   <tr>\n
<th>\n   2012 \n   </th>\n   <td>\n   162 \n

```
</td>\n        <td>\n          <p>1,736</p>\n        </td>\n        <td>\n
<p>1,898</p>\n        </td>\n        <td>\n          <p>7,879</p>\n        </td>\n
<td>\n          <p>9,777</p>\n        </td>\n      </tr>\n      <tr>\n        <th>\n
<p>2013</p>\n        </th>\n        <td>\n          <p>159</p>\n        </td>\n
<td>\n          <p>1,425</p>\n        </td>\n        <td>\n          <p>1,584</p>\n
</td>\n        <td>\n          <p>7,390</p>\n        </td>\n        <td>\n
<p>8,974</p>\n        </td>\n      </tr>\n      <tr>\n        <th>\n          <p>2014
</p>\n        </th>\n        <td>\n          <p>181</p>\n        </td>\n        <td>\
n          <p>1,488</p>\n        </td>\n        <td>\n          <p>1,669</p>\n
</td>\n        <td>\n          <p>7,164</p>\n        </td>\n        <td>\n
<p>8,833</p>\n        </td>\n      </tr>\n      <tr>\n        <th>\n          <p>2015
</p>\n        </th>\n        <td>\n          <p>157</p>\n        </td>\n        <td>\
n          <p>1,421</p>\n        </td>\n        <td>\n          <p>1,578</p>\n
</td>\n        <td>\n          <p>6,899</p>\n        </td>\n        <td>\n
<p>8,477</p>\n        </td>\n      </tr>\n      <tr>\n        <th>\n          <p>2016
</p>\n        </th>\n        <td>\n          <p>175</p>\n        </td>\n        <td>\
n          <p>1,432</p>\n        </td>\n        <td>\n          <p>1,607</p>\n
</td>\n        <td>\n          <p>6,746</p>\n        </td>\n        <td>\n
<p>8,353</p>\n        </td>\n      </tr>\n      <tr>\n        <th>\n          <p>2017
</p>\n        </th>\n        <td>\n          <p>140</p>\n        </td>\n        <td>\
n          <p>1,377</p>\n        </td>\n        <td>\n          <p>1,517</p>\n
</td>\n        <td>\n          <p>5,601</p>\n        </td>\n        <td>\n
<p>7,118</p>\n        </td>\n      </tr>\n      <tr>\n        <th>\n          <p>2018
<em>prov.</em></p>\n        </th>\n        <td>\n          <p>149</p>\n        </td>\
n        <td>\n          <p>1,368</p>\n        </td>\n        <td>\n          <p>1,51
7</p>\n        </td>\n        <td>\n          <p>4,895</p>\n        </td>\n        <t
d>\n          <p>6,412</p>\n        </td>\n      </tr>\n    </tbody>\n  </table>\n\n\
n\t\t\t\t\t\t<p>\n\t\t\t\t\t\t     <a href="/publication/key-reported-road-casualties
-scotland-2018/2-background/">Previous Page</a> | <a href="/publication/key-reported-
road-casualties-scotland-2018/4-reported-numbers-of-casualties-table-2/">Next Page</a
>\n\t\t\t\t\t\t</p>\n\n\n\t\t\t\t\t\t</p>\r\n\r\n</div>\r\n\r\n\r\n        </div>\r\n
</div>\r\n                </div>\r\n    </div>\r\n\r\n                </div>\r\n\r\n
<hr class="thick bg-brand--90 push-double--ends">\r\n\r\n            </div>\r\n\r\n
</div>\r\n\r\n        <!--Releated Content-->\r\n\r\n    </div>\r\n\r\n</article>\r\
n\r\n</main>\r\n\r\n\r\n<footer class="footer contentinfo soft-double--top">\r\n    \
r\n    <div class="wrapper">\r\n        \r\n        <div class="row">\r\n
```

To work with this request, we need the `Beautiful Soup` package

We first create a `soup` object and pass the contents of `r` to it

In [6]:
```
soup = BeautifulSoup(r.content,"html.parser")
print(soup)
```

```
<!DOCTYPE html>

<html class="" lang="en">
<head>
<meta charset="utf-8"/>
<meta content="width=device-width, initial-scale=1.0" name="viewport"/>
<link href="/content/css/styles.css" rel="stylesheet"/>
<script>document.documentElement.className = document.documentElement.className.split
('no-js').join('');</script>
<meta content="black-translucent" name="apple-mobile-web-app-status-bar-style"/>
<meta content="YES" name="apple-touch-fullscreen"/>
<title>3. Reported numbers of Accidents (Table 1) | Transport Scotland</title>
<meta content="" name="description">
<meta content="" property="og:title">
<meta content="" property="og:description">
<meta content="summary_large_image" name="twitter:card">
<meta content="@transcotland" name="twitter:site"/>
<meta content="" name="twitter:title"/>
<meta content="" name="twitter:description"/>
<link href="https://www.transport.gov.scot/publication/key-reported-road-casualties-s
cotland-2018/3-reported-numbers-of-accidents-table-1/" rel="canonical">
<meta content="Transport Scotland" data-build="300.25.0+release-300.025+2022-01-19" d
ata-date="2022-01-19T12:12:34.0000000Z" data-uver="7.14.0" data-ver="300.25.0.0" name
="_info"/>
<!-- favicons -->
<link href="/content/img/icons/apple-touch-icon.png" rel="apple-touch-icon" sizes="18
0x180"/>
<link href="/content/img/icons/favicon-32x32.png" rel="icon" sizes="32x32" type="imag
e/png"/>
<link href="/content/img/icons/favicon-16x16.png" rel="icon" sizes="16x16" type="imag
e/png"/>
<link href="/content/img/icons/manifest.json" rel="manifest"/>
```

```
<link color="#233e99" href="/content/img/icons/safari-pinned-tab.svg" rel="mask-ico
n"/>
<meta content="#233e99" name="msapplication-TileColor"/>
<meta content="/content/img/icons/mstile-144x144.png" name="msapplication-TileImage"/
>
<meta content="#19408e" name="theme-color"/>
<meta content="ePiQ70hyLfSbH4lb8yXaSt7BOeceeQQx_cIikqaKJAc" name="google-site-verific
ation">
<!--App Insights-->
<script>

        var appInsights=window.appInsights||function(config){

            function i(config){t[config]=function(){var i=arguments;t.queue.push(func
tion(){t[config].apply(t,i)})}}var t={config:config},u=document,e=window,o="script",s
="AuthenticatedUserContext",h="start",c="stop",l="Track",a=l+"Event",v=l+"Page",y=u.c
reateElement(o),r,f;y.src=config.url||"https://az416426.vo.msecnd.net/scripts/a/ai.0.
js";u.getElementsByTagName(o)[0].parentNode.appendChild(y);try{t.cookie=u.cookie}catc
h(p){}for(t.queue=[],t.version="1.0",r=["Event","Exception","Metric","PageView","Trac
e","Dependency"];r.length;)i("track"+r.pop());return i("set"+s),i("clear"+s),i(h+a),i
(c+a),i(h+v),i(c+v),i("flush"),config.disableExceptionTracking||(r="onerror",i("_"+
r),f=e[r],e[r]=function(config,i,u,e,o){var s=f&&f(config,i,u,e,o);return s!==!0&&t["
_"+r](config,i,u,e,o),s}),t

        }({

            instrumentationKey:"c80db1a7-5b42-4b8f-a7d0-08377eeafd1a"

        });



        window.appInsights=appInsights;

        appInsights.trackPageView();
```

```
    </script>
<!--App Insights-->
</meta></link></meta></meta></meta></meta></head>
<body class="">
<div class="skip">
<a class="btn btn-skip" href="#content" tabindex="0">Skip to main content</a>
</div>
<div class="banner-wrapper">
<div class="cookie-message" id="js-cookie-message" style="display:none">
<span class="center">This site uses cookies to store information on your computer. So
me of these cookies are essential to make our site work and others help us to improve
by giving us some insight into how the site is being used.</span>
<span class="right js-cookie-btn" id="close" role="button"><i class="icon icon-cance
l"></i></span>
</div>
</div>
<header class="header banner">
<nav class="wrapper relative">
<a class="logo" href="/"><span class="visually-hidden">Transport Scotland, Comhdhail
alba home button</span></a>
</nav>
<nav class="quick-links">
<div class="wrapper">
<ul class="quick-links__list right">
<li class="">
<a href="/about/">About</a>
</li>
<li class="">
<a href="/news/">News</a>
</li>
<li class="">
<a href="/consultations/">Consultations</a>
</li>
<li class="">
<a href="/media-library/">Media Library</a>
</li>
```

```
<li class="">
<a href="/publications/">Publications</a>
</li>
<li class="">
<a href="/traffic-management-updates/">Traffic management updates</a>
</li>
<li class="">
<a href="/coronavirus-covid-19/">Coronavirus (COVID-19)</a>
</li>
<li class="flush--right">
<a href="/cop26/">COP26</a>
</li>
</ul>
</div>
</nav>
<a aria-label="Search" class="mobile-search js-toggle--search" href="#search" title="
Search"></a>
<a aria-label="Navigation menu" class="mobile-menu js-toggle" href="#navigation" titl
e="Navigation menu"></a>
<div class="nav-wrapper">
<div class="wrapper">
<nav class="nav-primary" id="navigation">
<ul class="nav-primary__list">
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/about/">About</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/news/">News</a>
</li>
<li class="nav-primary__item">
<a class="nav-primary__link" href="/our-approach/">Our approach</a>
</li>
<li class="nav-primary__item">
<a class="nav-primary__link" href="/public-transport/">Public transport</a>
</li>
<li class="nav-primary__item">
```

```
<a class="nav-primary__link" href="/active-travel/">Walking and cycling</a>
</li>
<li class="nav-primary__item">
<a class="nav-primary__link" href="/transport-network/">Transport network</a>
</li>
<li class="nav-primary__item">
<a class="nav-primary__link" href="/projects/">Projects</a>
</li>
<li class="nav-primary__item">
<a class="nav-primary__link" href="/concessionary-travel/">Concessionary travel</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/consultations/">Consultations</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/media-library/">Media Library</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/publications/">Publications</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/traffic-management-updates/">Traffic management u
pdates</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/coronavirus-covid-19/">Coronavirus (COVID-19)</a>
</li>
<li class="nav-primary__item hidden">
<a class="nav-primary__link" href="/cop26/">COP26</a>
</li>
</ul>
<button class="btn-close js-toggle" data-target="#navigation">Close</button>
</nav>
</div>
</div>
<div class="relative">
```

```
<div class="wrapper relative">
<a aria-label="Search" class="search-button js-toggle--search" href="#search" title="
Search"><i class="icon icon-search-button"></i></a>
</div>
<div class="search" id="search">
<div class="wrapper relative">
<form action="/search-results/" method="get">
<label class="visually-hidden" for="query">Search</label>
<input id="query" name="q" placeholder="Search for news, projects, policies etc" tabi
ndex="-1" title="search" type="search"/>
<button tabindex="-1" title="search-button"><i class="icon icon-search"></i></button>
</form>
</div>
</div>
</div>
</header>
<main id="content">
<div class="breadcrumb-wrapper bg-brand--90">
<div class="wrapper">
<div class="clearfix breadcrumb-inner">
<nav aria-label="Breadcrumb" class="">
<ul class="breadcrumb clearfix">
<li><a href="/">Home</a></li>
<li aria-hidden="true"><i class="icon icon-angle-right"></i></li>
<li><a href="/publications/">Publications</a></li>
<li aria-hidden="true"><i class="icon icon-angle-right"></i></li>
<li><a href="/publication/key-reported-road-casualties-scotland-2018/">Key Reported R
oad Casualties Scotland 2018</a></li>
<li aria-hidden="true"><i class="icon icon-angle-right"></i></li>
</ul>
</nav>
</div>
</div>
</div>
<article class="main-wrapper soft-double--bottom clearfix">
<div class="bg-brand--90 push--bottom">
```

```
<div class="wrapper">
<div class="row">
<div class="col xsmall-12 right">
<h1 class="article__title">3. Reported numbers of Accidents (Table 1)</h1>
</div>
</div>
</div>
</div>
<div class="wrapper soft-double--bottom">
<div class="row">
<div class="col xsmall-12">
<div class="editor full-width">
<div class="umb-grid">
<div class="grid-section">
<div>
<div class="container">
<div class="">
<p>
<h2>Key Reported Road Casualties Scotland 2018</h2>
<h3>3. Reported numbers of Accidents (<a href="#Table1">Table 1</a>)</h3>
<p>3.1<em> <a href="#Table1">Table 1</a></em> shows the downward trend of injury road
    <strong>accidents</strong> recorded by the police. In 2018, there were 6,412 acci
dents in which someone was
    killed or injured, 10 per cent fewer than in 2017 and the lowest number since rec
ords began. There were 149
    fatal accidents in 2018, 9 (6%) more than in 2017. In 2018, there were 1,368 seri
ous injury accidents - a
    decrease of 9 (1%) from 2017; and 4,895 slight injury accidents reported in 2018,
13 per cent (706) fewer than
    2017.</p>
<table border="1" cellpadding="5" cellspacing="1" id="Table1">
<caption>
          Table 1: Injury Road Accidents by Severity, 1970 –
          2018
        </caption>
<thead>
```

<tr>
<th> </th>
<th>
Fatal
</th>
<th>
Serious
</th>
<th>
Fatal and Serious
</th>
<th>
Slight
</th>
<th>
All
</th>
</tr>
</thead>
<tbody>
<tr>
<th>
1970
</th>
<td>
758
</td>
<td>
7,860
</td>
<td>
8,618
</td>
<td>
13,515
</td>

<td>
22,133
</td>
</tr>
<tr>
<th>
1975
</th>
<td>
699
</td>
<td>
6,912
</td>
<td>
7,611
</td>
<td>
13,041
</td>
<td>
20,652
</td>
</tr>
<tr>
<th>
1980
</th>
<td>
644
</td>
<td>
7,218
</td>
<td>
7,862
</td>
</tr>
<tr>
<th>
1985
</th>
<td>
1,063
</td>
<td>
8,094
</td>
<td>
8,333
</td>
<td>
16,427
</td>
<td>
25,917
</td>
</tr>
<tr>
<th>
1990
</th>
<td>
1,190
</td>
<td>
8,203
</td>
<td>
8,777
</td>
<td>
16,980
</td>
<td>
26,370
</td>
</tr>
<tr>
<th>
1995
</th>
<td>
1,196
</td>
<td>
8,316
</td>
<td>
9,110
</td>
<td>
17,426
</td>
<td>
26,938
</td>
</tr>
<tr>
<th>
2000
</th>
<td>
1,248
</td>
<td>
8,424
</td>
<td>
9,443
</td>
<td>
17,867
</td>
<td>
27,539
</td>
</tr>
<tr>
<th>
2005
</th>
<td>
1,249
</td>
<td>
8,520
</td>
<td>
9,843
</td>
<td>
18,363
</td>
<td>
28,132
</td>
</tr>
<tr>
<th>
2010
</th>
<td>
1,252
</td>
<td>
8,615
</td>
<td>
10,244
</td>
<td>
18,859
</td>
<td>
28,726
</td>
</tr>
<tr>
<th>
2015
</th>
<td>
1,253
</td>
<td>
8,710
</td>
<td>
10,645
</td>
<td>
19,355
</td>
<td>
29,320
</td>
</tr>
<tr>
<th>
2020
</th>
<td>
1,256
</td>
<td>
8,805
</td>
<td>
11,046
</td>
<td>
19,851
</td>
<td>
29,914
</td>
</tr>
</tbody>
</table>
</figure>
</section>
</section>

```
</td>
<td>
<p>13,926</p>
</td>
<td>
<p>21,788</p>
</td>
</tr>
<tr>
<th>
<p>1985</p>
</th>
<td>
<p>550</p>
</td>
<td>
<p>6,507</p>
</td>
<td>
<p>7,057</p>
</td>
<td>
<p>13,587</p>
</td>
<td>
<p>20,644</p>
</td>
</tr>
<tr>
<th>
<p>1990</p>
</th>
<td>
<p>491</p>
</td>
<td>
```

5,237
</td>
<td>
5,728
</td>
<td>
14,443
</td>
<td>
20,171
</td>
</tr>
<tr>
<th>
1995
</th>
<td>
361
</td>
<td>
4,071
</td>
<td>
4,432
</td>
<td>
12,102
</td>
<td>
16,534
</td>
</tr>
<tr>
<th>
1996
</th>

<td>
316
</td>
<td>
3,315
</td>
<td>
3,631
</td>
<td>
12,442
</td>
<td>
16,073
</td>
</tr>
<tr>
<th>
1997
</th>
<td>
340
</td>
<td>
3,312
</td>
<td>
3,652
</td>
<td>
12,994
</td>
<td>
16,646
</td>
</tr>

<tr>
<th>
1998
</th>
<td>
339
</td>
<td>
3,318
</td>
<td>
3,657
</td>
<td>
12,862
</td>
<td>
16,519
</td>
</tr>
<tr>
<th>
1999
</th>
<td>
285
</td>
<td>
3,209
</td>
<td>
3,494
</td>
<td>
11,921
</td>

<td>
15,415
</td>
</tr>
<tr>
<th>
2000
</th>
<td>
297
</td>
<td>
3,007
</td>
<td>
3,304
</td>
<td>
11,828
</td>
<td>
15,132
</td>
</tr>
<tr>
<th>
2001
</th>
<td>
309
</td>
<td>
2,840
</td>
<td>
3,149

```
</td>
<td>
<p>11,575</p>
</td>
<td>
<p>14,724</p>
</td>
</tr>
<tr>
<th>
<p>2002</p>
</th>
<td>
<p>274</p>
</td>
<td>
<p>2,684</p>
</td>
<td>
<p>2,958</p>
</td>
<td>
<p>11,385</p>
</td>
<td>
<p>14,343</p>
</td>
</tr>
<tr>
<th>
<p>2003</p>
</th>
<td>
<p>301</p>
</td>
<td>
```

2,495
</td>
<td>
2,796
</td>
<td>
11,121
</td>
<td>
13,917
</td>
</tr>
<tr>
<th>
2004
</th>
<td>
283
</td>
<td>
2,331
</td>
<td>
2,614
</td>
<td>
11,305
</td>
<td>
13,919
</td>
</tr>
<tr>
<th>
2005
</th>

<td>
264
</td>
<td>
2,252
</td>
<td>
2,516
</td>
<td>
10,922
</td>
<td>
13,438
</td>
</tr>
<tr>
<th>
2006
</th>
<td>
293
</td>
<td>
2,257
</td>
<td>
2,550
</td>
<td>
10,560
</td>
<td>
13,110
</td>
</tr>

<tr>
<th>
2007
</th>
<td>
255
</td>
<td>
2,049
</td>
<td>
2,304
</td>
<td>
10,203
</td>
<td>
12,507
</td>
</tr>
<tr>
<th>
2008
</th>
<td>
245
</td>
<td>
2,242
</td>
<td>
2,487
</td>
<td>
9,672
</td>

<td>
12,159
</td>
</tr>
<tr>
<th>
2009
</th>
<td>
196
</td>
<td>
1,998
</td>
<td>
2,194
</td>
<td>
9,362
</td>
<td>
11,556
</td>
</tr>
<tr>
<th>
2010
</th>
<td>
189
</td>
<td>
1,713
</td>
<td>
1,902

```
</td>
<td>
<p>8,393</p>
</td>
<td>
<p>10,295</p>
</td>
</tr>
<tr>
<th>
<p>2011</p>
</th>
<td>
<p>175</p>
</td>
<td>
<p>1,675</p>
</td>
<td>
<p>1,850</p>
</td>
<td>
<p>8,135</p>
</td>
<td>
<p>9,985</p>
</td>
</tr>
<tr>
<th>
<p>2012</p>
</th>
<td>
<p>162</p>
</td>
<td>
```

1,736
</td>
<td>
1,898
</td>
<td>
7,879
</td>
<td>
9,777
</td>
</tr>
<tr>
<th>
2013
</th>
<td>
159
</td>
<td>
1,425
</td>
<td>
1,584
</td>
<td>
7,390
</td>
<td>
8,974
</td>
</tr>
<tr>
<th>
2014
</th>

<td>
181
</td>
<td>
1,488
</td>
<td>
1,669
</td>
<td>
7,164
</td>
<td>
8,833
</td>
</tr>
<tr>
<th>
2015
</th>
<td>
157
</td>
<td>
1,421
</td>
<td>
1,578
</td>
<td>
6,899
</td>
<td>
8,477
</td>
</tr>

<tr>
<th>
2016
</th>
<td>
175
</td>
<td>
1,432
</td>
<td>
1,607
</td>
<td>
6,746
</td>
<td>
8,353
</td>
</tr>
<tr>
<th>
2017
</th>
<td>
140
</td>
<td>
1,377
</td>
<td>
1,517
</td>
<td>
5,601
</td>

<td>
7,118
</td>
</tr>
<tr>
<th>
2018 prov.
</th>
<td>
149
</td>
<td>
1,368
</td>
<td>
1,517
</td>
<td>
4,895
</td>
<td>
6,412
</td>
</tr>
</tbody>
</table>

<a href="/publication/key-reported-road-casualties-scotland-2018/2-background/">Previ
ous Page</a> | <a href="/publication/key-reported-road-casualties-scotland-2018/4-rep
orted-numbers-of-casualties-table-2/">Next Page</a>

```
</div>
</div>
<hr class="thick bg-brand--90 push-double--ends"/>
</div>
</div>
<!--Releated Content-->
</div>
</article>
</main>
<footer class="footer contentinfo soft-double--top">
<div class="wrapper">
<div class="row">
<div class="col xsmall-12 small-6">
<ul class="footer-nav">
<li>
<a href="/contact/">Contact</a>
</li>
<li>
<a href="/freedom-of-information/">Freedom Of Information</a>
</li>
<li>
<a href="/careers/">Careers</a>
</li>
</ul>
</div>
<div class="col xsmall-12 small-6">
<ul class="social-footer">
<li><a class="instagram" href="https://www.instagram.com/transportscotland" target="_
blank" title="instagram"><i aria-label="instagram" class="icon icon-instagram" role="
img"></i></a></li>
<li><a class="flickr" href="https://www.flickr.com/photos/transportscotland" target="
_blank" title="flickr"><i aria-label="flickr" class="icon icon-flickr" role="img"></i
></a></li>
<li><a class="twitter" href="https://twitter.com/transcotland" target="_blank" title
="twitter"><i aria-label="twitter" class="icon icon-twitter" role="img"></i></a></li>
<li><a class="facebook" href="https://www.facebook.com/Transcotland" target="_blank"
```

```
title="facebook"><i aria-label="facebook" class="icon icon-facebook" role="img"></i>
</a></li>
<li><a class="youtube" href="https://www.youtube.com/user/transportscotland" target="
_blank" title="youtube"><i aria-label="youtube" class="icon icon-youtube" role="img">
</i></a></li>
<li><a class="linkedin" href="https://www.linkedin.com/company/transport-scotland" ta
rget="_blank" title="linkedin"><i aria-label="linkedin" class="icon icon-linkedin" ro
le="img"></i></a></li>
</ul>
</div>
</div>
<div class="row push--top">
<div class="legal">
<div class="col xsmall-12 small-6">
<span aria-hidden="true">©</span> Copyright Transport Scotland



                </div>
<div class="col xsmall-12 small-6">
<ul class="footer-subnav">
<li>
<a href="/accessibility/">Accessibility</a>
</li>
<li>
<a href="/website-privacy-policy/">Website privacy policy</a>
</li>
<li>
<a href="/cookie-policy/">Cookie Policy</a>
</li>
<li>
```

Keep in mind that we are looking for the *tabular data*

- Just like when you **inspect** a website:
    - Right-click, `View Page Sources`
    - `View` $\rightarrow$ `Developer` $\rightarrow$ `Inspect Elements`

In [7]:
```
# Option 1
table_v1 = soup.find('table')
print(table_v1)
```

```
<table border="1" cellpadding="5" cellspacing="1" id="Table1">
<caption>
          Table 1: Injury Road Accidents by Severity, 1970 –
          2018
        </caption>
<thead>
<tr>
<th> </th>
<th>
<p>Fatal</p>
</th>
<th>
<p>Serious</p>
</th>
<th>
<p>Fatal and Serious</p>
</th>
<th>
<p>Slight</p>
</th>
<th>
<p>All</p>
</th>
</tr>
</thead>
<tbody>
<tr>
<th>
<p>1970</p>
</th>
<td>
```

758
</td>
<td>
7,860
</td>
<td>
8,618
</td>
<td>
13,515
</td>
<td>
22,133
</td>
</tr>
<tr>
<th>
1975
</th>
<td>
699
</td>
<td>
6,912
</td>
<td>
7,611
</td>
<td>
13,041
</td>
<td>
20,652
</td>
</tr>
<tr>

<th>
1980
</th>
<td>
644
</td>
<td>
7,218
</td>
<td>
7,862
</td>
<td>
13,926
</td>
<td>
21,788
</td>
</tr>
<tr>
<th>
1985
</th>
<td>
550
</td>
<td>
6,507
</td>
<td>
7,057
</td>
<td>
13,587
</td>
<td>

20,644
</td>
</tr>
<tr>
<th>
1990
</th>
<td>
491
</td>
<td>
5,237
</td>
<td>
5,728
</td>
<td>
14,443
</td>
<td>
20,171
</td>
</tr>
<tr>
<th>
1995
</th>
<td>
361
</td>
<td>
4,071
</td>
<td>
4,432
</td>

<td>
12,102
</td>
<td>
16,534
</td>
</tr>
<tr>
<th>
1996
</th>
<td>
316
</td>
<td>
3,315
</td>
<td>
3,631
</td>
<td>
12,442
</td>
<td>
16,073
</td>
</tr>
<tr>
<th>
1997
</th>
<td>
340
</td>
<td>
3,312

```
</td>
<td>
<p>3,652</p>
</td>
<td>
<p>12,994</p>
</td>
<td>
<p>16,646</p>
</td>
</tr>
<tr>
<th>
<p>1998</p>
</th>
<td>
<p>339</p>
</td>
<td>
<p>3,318</p>
</td>
<td>
<p>3,657</p>
</td>
<td>
<p>12,862</p>
</td>
<td>
<p>16,519</p>
</td>
</tr>
<tr>
<th>
<p>1999</p>
</th>
<td>
```

285
</td>
<td>
3,209
</td>
<td>
3,494
</td>
<td>
11,921
</td>
<td>
15,415
</td>
</tr>
<tr>
<th>
2000
</th>
<td>
297
</td>
<td>
3,007
</td>
<td>
3,304
</td>
<td>
11,828
</td>
<td>
15,132
</td>
</tr>
<tr>

<th>
2001
</th>
<td>
309
</td>
<td>
2,840
</td>
<td>
3,149
</td>
<td>
11,575
</td>
<td>
14,724
</td>
</tr>
<tr>
<th>
2002
</th>
<td>
274
</td>
<td>
2,684
</td>
<td>
2,958
</td>
<td>
11,385
</td>
<td>

14,343
</td>
</tr>
<tr>
<th>
2003
</th>
<td>
301
</td>
<td>
2,495
</td>
<td>
2,796
</td>
<td>
11,121
</td>
<td>
13,917
</td>
</tr>
<tr>
<th>
2004
</th>
<td>
283
</td>
<td>
2,331
</td>
<td>
2,614
</td>

<td>
11,305
</td>
<td>
13,919
</td>
</tr>
<tr>
<th>
2005
</th>
<td>
264
</td>
<td>
2,252
</td>
<td>
2,516
</td>
<td>
10,922
</td>
<td>
13,438
</td>
</tr>
<tr>
<th>
2006
</th>
<td>
293
</td>
<td>
2,257

```
</td>
<td>
<p>2,550</p>
</td>
<td>
<p>10,560</p>
</td>
<td>
<p>13,110</p>
</td>
</tr>
<tr>
<th>
<p>2007</p>
</th>
<td>
<p>255</p>
</td>
<td>
<p>2,049</p>
</td>
<td>
<p>2,304</p>
</td>
<td>
<p>10,203</p>
</td>
<td>
<p>12,507</p>
</td>
</tr>
<tr>
<th>
<p>2008</p>
</th>
<td>
```

245
</td>
<td>
2,242
</td>
<td>
2,487
</td>
<td>
9,672
</td>
<td>
12,159
</td>
</tr>
<tr>
<th>
2009
</th>
<td>
196
</td>
<td>
1,998
</td>
<td>
2,194
</td>
<td>
9,362
</td>
<td>
11,556
</td>
</tr>
<tr>

<th>
2010
</th>
<td>
189
</td>
<td>
1,713
</td>
<td>
1,902
</td>
<td>
8,393
</td>
<td>
10,295
</td>
</tr>
<tr>
<th>
2011
</th>
<td>
175
</td>
<td>
1,675
</td>
<td>
1,850
</td>
<td>
8,135
</td>
<td>

9,985
</td>
</tr>
<tr>
<th>
2012
</th>
<td>
162
</td>
<td>
1,736
</td>
<td>
1,898
</td>
<td>
7,879
</td>
<td>
9,777
</td>
</tr>
<tr>
<th>
2013
</th>
<td>
159
</td>
<td>
1,425
</td>
<td>
1,584
</td>

<td>
7,390
</td>
<td>
8,974
</td>
</tr>
<tr>
<th>
2014
</th>
<td>
181
</td>
<td>
1,488
</td>
<td>
1,669
</td>
<td>
7,164
</td>
<td>
8,833
</td>
</tr>
<tr>
<th>
2015
</th>
<td>
157
</td>
<td>
1,421

```
</td>
<td>
<p>1,578</p>
</td>
<td>
<p>6,899</p>
</td>
<td>
<p>8,477</p>
</td>
</tr>
<tr>
<th>
<p>2016</p>
</th>
<td>
<p>175</p>
</td>
<td>
<p>1,432</p>
</td>
<td>
<p>1,607</p>
</td>
<td>
<p>6,746</p>
</td>
<td>
<p>8,353</p>
</td>
</tr>
<tr>
<th>
<p>2017</p>
</th>
<td>
```

140
</td>
<td>
1,377
</td>
<td>
1,517
</td>
<td>
5,601
</td>
<td>
7,118
</td>
</tr>
<tr>
<th>
2018 prov.
</th>
<td>
149
</td>
<td>
1,368
</td>
<td>
1,517
</td>
<td>
4,895
</td>
</tr>
</tbody>
</table>
</figure>
</body>
</html>

In [8]:

```
# Option 2
table_v2 = soup.find('table', id='Table1')
print(table_v2)
```

```
<table border="1" cellpadding="5" cellspacing="1" id="Table1">
<caption>
          Table 1: Injury Road Accidents by Severity, 1970 –
          2018
        </caption>
<thead>
<tr>
<th> </th>
<th>
<p>Fatal</p>
</th>
<th>
<p>Serious</p>
</th>
<th>
<p>Fatal and Serious</p>
</th>
<th>
<p>Slight</p>
</th>
<th>
<p>All</p>
</th>
</tr>
</thead>
<tbody>
<tr>
<th>
<p>1970</p>
</th>
<td>
```

758
</td>
<td>
7,860
</td>
<td>
8,618
</td>
<td>
13,515
</td>
<td>
22,133
</td>
</tr>
<tr>
<th>
1975
</th>
<td>
699
</td>
<td>
6,912
</td>
<td>
7,611
</td>
<td>
13,041
</td>
<td>
20,652
</td>
</tr>
<tr>

<th>
1980
</th>
<td>
644
</td>
<td>
7,218
</td>
<td>
7,862
</td>
<td>
13,926
</td>
<td>
21,788
</td>
</tr>
<tr>
<th>
1985
</th>
<td>
550
</td>
<td>
6,507
</td>
<td>
7,057
</td>
<td>
13,587
</td>
<td>

20,644
</td>
</tr>
<tr>
<th>
1990
</th>
<td>
491
</td>
<td>
5,237
</td>
<td>
5,728
</td>
<td>
14,443
</td>
<td>
20,171
</td>
</tr>
<tr>
<th>
1995
</th>
<td>
361
</td>
<td>
4,071
</td>
<td>
4,432
</td>

<td>
12,102
</td>
<td>
16,534
</td>
</tr>
<tr>
<th>
1996
</th>
<td>
316
</td>
<td>
3,315
</td>
<td>
3,631
</td>
<td>
12,442
</td>
<td>
16,073
</td>
</tr>
<tr>
<th>
1997
</th>
<td>
340
</td>
<td>
3,312

```
</td>
<td>
<p>3,652</p>
</td>
<td>
<p>12,994</p>
</td>
<td>
<p>16,646</p>
</td>
</tr>
<tr>
<th>
<p>1998</p>
</th>
<td>
<p>339</p>
</td>
<td>
<p>3,318</p>
</td>
<td>
<p>3,657</p>
</td>
<td>
<p>12,862</p>
</td>
<td>
<p>16,519</p>
</td>
</tr>
<tr>
<th>
<p>1999</p>
</th>
<td>
```

285
</td>
<td>
3,209
</td>
<td>
3,494
</td>
<td>
11,921
</td>
<td>
15,415
</td>
</tr>
<tr>
<th>
2000
</th>
<td>
297
</td>
<td>
3,007
</td>
<td>
3,304
</td>
<td>
11,828
</td>
<td>
15,132
</td>
</tr>
<tr>

<th>
2001
</th>
<td>
309
</td>
<td>
2,840
</td>
<td>
3,149
</td>
<td>
11,575
</td>
<td>
14,724
</td>
</tr>
<tr>
<th>
2002
</th>
<td>
274
</td>
<td>
2,684
</td>
<td>
2,958
</td>
<td>
11,385
</td>
<td>

14,343
</td>
</tr>
<tr>
<th>
2003
</th>
<td>
301
</td>
<td>
2,495
</td>
<td>
2,796
</td>
<td>
11,121
</td>
<td>
13,917
</td>
</tr>
<tr>
<th>
2004
</th>
<td>
283
</td>
<td>
2,331
</td>
<td>
2,614
</td>

<td>
11,305
</td>
<td>
13,919
</td>
</tr>
<tr>
<th>
2005
</th>
<td>
264
</td>
<td>
2,252
</td>
<td>
2,516
</td>
<td>
10,922
</td>
<td>
13,438
</td>
</tr>
<tr>
<th>
2006
</th>
<td>
293
</td>
<td>
2,257

```
</td>
<td>
<p>2,550</p>
</td>
<td>
<p>10,560</p>
</td>
<td>
<p>13,110</p>
</td>
</tr>
<tr>
<th>
<p>2007</p>
</th>
<td>
<p>255</p>
</td>
<td>
<p>2,049</p>
</td>
<td>
<p>2,304</p>
</td>
<td>
<p>10,203</p>
</td>
<td>
<p>12,507</p>
</td>
</tr>
<tr>
<th>
<p>2008</p>
</th>
<td>
```

245
</td>
<td>
2,242
</td>
<td>
2,487
</td>
<td>
9,672
</td>
<td>
12,159
</td>
</tr>
<tr>
<th>
2009
</th>
<td>
196
</td>
<td>
1,998
</td>
<td>
2,194
</td>
<td>
9,362
</td>
<td>
11,556
</td>
</tr>
<tr>

<th>
2010
</th>
<td>
189
</td>
<td>
1,713
</td>
<td>
1,902
</td>
<td>
8,393
</td>
<td>
10,295
</td>
</tr>
<tr>
<th>
2011
</th>
<td>
175
</td>
<td>
1,675
</td>
<td>
1,850
</td>
<td>
8,135
</td>
<td>

9,985
</td>
</tr>
<tr>
<th>
2012
</th>
<td>
162
</td>
<td>
1,736
</td>
<td>
1,898
</td>
<td>
7,879
</td>
<td>
9,777
</td>
</tr>
<tr>
<th>
2013
</th>
<td>
159
</td>
<td>
1,425
</td>
<td>
1,584
</td>

<td>
7,390
</td>
<td>
8,974
</td>
</tr>
<tr>
<th>
2014
</th>
<td>
181
</td>
<td>
1,488
</td>
<td>
1,669
</td>
<td>
7,164
</td>
<td>
8,833
</td>
</tr>
<tr>
<th>
2015
</th>
<td>
157
</td>
<td>
1,421

```
</td>
<td>
<p>1,578</p>
</td>
<td>
<p>6,899</p>
</td>
<td>
<p>8,477</p>
</td>
</tr>
<tr>
<th>
<p>2016</p>
</th>
<td>
<p>175</p>
</td>
<td>
<p>1,432</p>
</td>
<td>
<p>1,607</p>
</td>
<td>
<p>6,746</p>
</td>
<td>
<p>8,353</p>
</td>
</tr>
<tr>
<th>
<p>2017</p>
</th>
```

In [9]:
```
# You can print properties of the tables you find, such as the caption
print(table_v1.caption)
```

```
<caption>
          Table 1: Injury Road Accidents by Severity, 1970 –
          2018
        </caption>
```

In [10]:

```
# Now we can put together a command that finds the table, it's body and all entries (tr)
rows = soup.find('table').find('tbody').find_all('tr')
print(rows)
```

```
[<tr>
<th>
<p>1970</p>
</th>
<td>
<p>758</p>
</td>
<td>
<p>7,860</p>
</td>
<td>
<p>8,618</p>
</td>
<td>
<p>13,515</p>
</td>
<td>
<p>22,133</p>
</td>
</tr>, <tr>
<th>
<p>1975</p>
</th>
<td>
<p>699</p>
</td>
<td>
<p>6,912</p>
</td>
<td>
<p>7,611</p>
```

```
</td>
<td>
<p>13,041</p>
</td>
<td>
<p>20,652</p>
</td>
</tr>, <tr>
<th>
<p>1980</p>
</th>
<td>
<p>644</p>
</td>
<td>
<p>7,218</p>
</td>
<td>
<p>7,862</p>
</td>
<td>
<p>13,926</p>
</td>
<td>
<p>21,788</p>
</td>
</tr>, <tr>
<th>
<p>1985</p>
</th>
<td>
<p>550</p>
</td>
<td>
<p>6,507</p>
</td>
```

<td>
7,057
</td>
<td>
13,587
</td>
<td>
20,644
</td>
</tr>, <tr>
<th>
1990
</th>
<td>
491
</td>
<td>
5,237
</td>
<td>
5,728
</td>
<td>
14,443
</td>
<td>
20,171
</td>
</tr>, <tr>
<th>
1995
</th>
<td>
361
</td>
<td>

4,071
</td>
<td>
4,432
</td>
<td>
12,102
</td>
<td>
16,534
</td>
</tr>, <tr>
<th>
1996
</th>
<td>
316
</td>
<td>
3,315
</td>
<td>
3,631
</td>
<td>
12,442
</td>
<td>
16,073
</td>
</tr>, <tr>
<th>
1997
</th>
<td>
340

```
</td>
<td>
<p>3,312</p>
</td>
<td>
<p>3,652</p>
</td>
<td>
<p>12,994</p>
</td>
<td>
<p>16,646</p>
</td>
</tr>, <tr>
<th>
<p>1998</p>
</th>
<td>
<p>339</p>
</td>
<td>
<p>3,318</p>
</td>
<td>
<p>3,657</p>
</td>
<td>
<p>12,862</p>
</td>
<td>
<p>16,519</p>
</td>
</tr>, <tr>
<th>
<p>1999</p>
</th>
```

<td>
285
</td>
<td>
3,209
</td>
<td>
3,494
</td>
<td>
11,921
</td>
<td>
15,415
</td>
</tr>, <tr>
<th>
2000
</th>
<td>
297
</td>
<td>
3,007
</td>
<td>
3,304
</td>
<td>
11,828
</td>
<td>
15,132
</td>
</tr>, <tr>
<th>

2001
</th>
<td>
309
</td>
<td>
2,840
</td>
<td>
3,149
</td>
<td>
11,575
</td>
<td>
14,724
</td>
</tr>, <tr>
<th>
2002
</th>
<td>
274
</td>
<td>
2,684
</td>
<td>
2,958
</td>
<td>
11,385
</td>
<td>
14,343
</td>

```
</tr>, <tr>
<th>
<p>2003</p>
</th>
<td>
<p>301</p>
</td>
<td>
<p>2,495</p>
</td>
<td>
<p>2,796</p>
</td>
<td>
<p>11,121</p>
</td>
<td>
<p>13,917</p>
</td>
</tr>, <tr>
<th>
<p>2004</p>
</th>
<td>
<p>283</p>
</td>
<td>
<p>2,331</p>
</td>
<td>
<p>2,614</p>
</td>
<td>
<p>11,305</p>
</td>
<td>
```

13,919
</td>
</tr>, <tr>
<th>
2005
</th>
<td>
264
</td>
<td>
2,252
</td>
<td>
2,516
</td>
<td>
10,922
</td>
<td>
13,438
</td>
</tr>, <tr>
<th>
2006
</th>
<td>
293
</td>
<td>
2,257
</td>
<td>
2,550
</td>
<td>
10,560

```
</td>
<td>
<p>13,110</p>
</td>
</tr>, <tr>
<th>
<p>2007</p>
</th>
<td>
<p>255</p>
</td>
<td>
<p>2,049</p>
</td>
<td>
<p>2,304</p>
</td>
<td>
<p>10,203</p>
</td>
<td>
<p>12,507</p>
</td>
</tr>, <tr>
<th>
<p>2008</p>
</th>
<td>
<p>245</p>
</td>
<td>
<p>2,242</p>
</td>
<td>
<p>2,487</p>
</td>
```

<td>
9,672
</td>
<td>
12,159
</td>
</tr>, <tr>
<th>
2009
</th>
<td>
196
</td>
<td>
1,998
</td>
<td>
2,194
</td>
<td>
9,362
</td>
<td>
11,556
</td>
</tr>, <tr>
<th>
2010
</th>
<td>
189
</td>
<td>
1,713
</td>
<td>

1,902
</td>
<td>
8,393
</td>
<td>
10,295
</td>
</tr>, <tr>
<th>
2011
</th>
<td>
175
</td>
<td>
1,675
</td>
<td>
1,850
</td>
<td>
8,135
</td>
<td>
9,985
</td>
</tr>, <tr>
<th>
2012
</th>
<td>
162
</td>
<td>
1,736

```
</td>
<td>
<p>1,898</p>
</td>
<td>
<p>7,879</p>
</td>
<td>
<p>9,777</p>
</td>
</tr>, <tr>
<th>
<p>2013</p>
</th>
<td>
<p>159</p>
</td>
<td>
<p>1,425</p>
</td>
<td>
<p>1,584</p>
</td>
<td>
<p>7,390</p>
</td>
<td>
<p>8,974</p>
</td>
</tr>, <tr>
<th>
<p>2014</p>
</th>
<td>
<p>181</p>
</td>
```

<td>
1,488
</td>
<td>
1,669
</td>
<td>
7,164
</td>
<td>
8,833
</td>
</tr>, <tr>
<th>
2015
</th>
<td>
157
</td>
<td>
1,421
</td>
<td>
1,578
</td>
<td>
6,899
</td>
<td>
8,477
</td>
</tr>, <tr>
<th>
2016
</th>
<td>

175
</td>
<td>
1,432
</td>
<td>
1,607
</td>
<td>
6,746
</td>
<td>
8,353
</td>
</tr>, <tr>
<th>
2017
</th>
<td>
140
</td>
<td>
1,377
</td>
<td>
1,517
</td>
<td>
5,601
</td>
<td>
7,118
</td>
</tr>, <tr>
<th>
2018 prov.

```
</th>
<td>
<p>149</p>
</td>
<td>
<p>1,368</p>
</td>
<td>
<p>1,517</p>
</td>
```

In [11]:

```
# We will loop over the rows variable to see all text entries
fatalities = None
year = '2017' # if you want 2018, you have to put 2018 prov.
for row in rows:
    cell = row.find_all('th')
    if cell[0].find('p').get_text()==year:
        data_cells = row.find_all('td')
        fatalities = data_cells[0].find('p').get_text()
print('The number of fatalities in '+year+' was '+fatalities)
```

```
The number of fatalities in 2017 was 140
```

How would you get the number of `serious` accidents for a given year?

In [12]:
```
# We will loop over the rows variable to see all text entries
serious = None
year = '1970' # if you want 2018, you have to put 2018 prov.
for row in rows:
    cell = row.find_all('th')
    if cell[0].find('p').get_text()==year:
        data_cells = row.find_all('td')
        serious = data_cells[1].find('p').get_text()
print('The number of serious accidents in '+year+' was '+serious)
```

```
The number of serious accidents in 1970 was 7,860
```

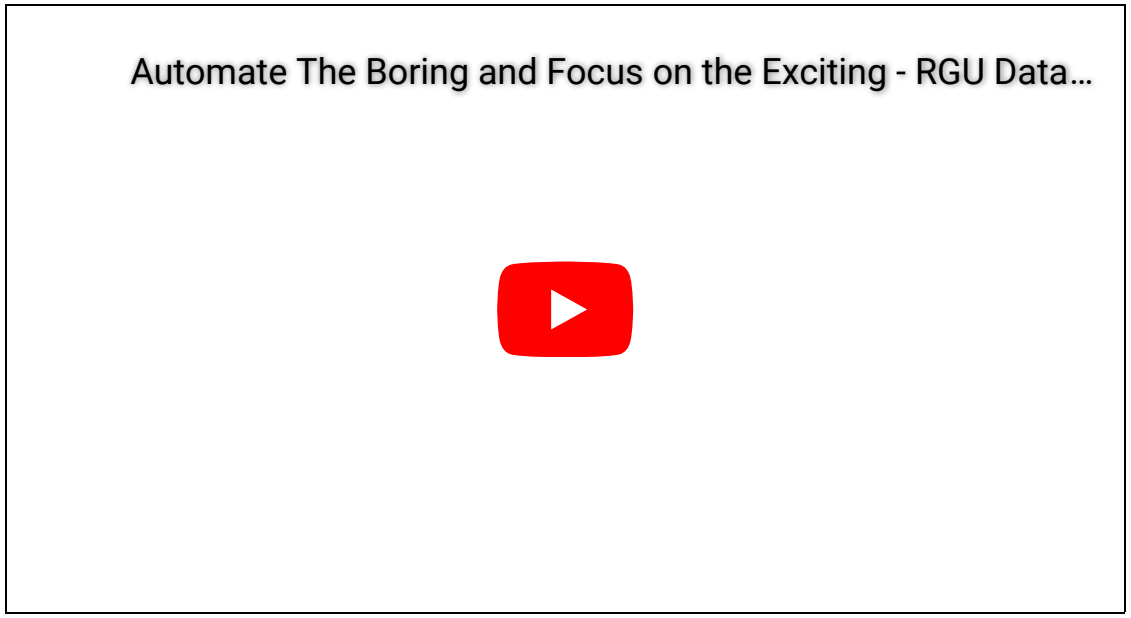
More advise on this topic

Check out other tools such as ScraPy (https://www.scrapingbee.com/blog/web-scraping-with-scrapy/) and Selenium (https://www.scrapingbee.com/blog/selenium-python/), which are much more powerful packages that allows you to navigate and scrape tables in easier ways!

Also, Arturo Regalado (a PhD student in Economics at Uni of Aberdeen) prepared the following tutorial on how he uses web scraping to develop dashboards for Oil & Gas companies

In [1]:
```
import warnings;
warnings.simplefilter('ignore')
from IPython.display import HTML
HTML('<iframe width="560" height="315" src="https://www.youtube.com/embed/6iqPwzSOkc4" t
itle="YouTube video player" frameborder="0" allow="accelerometer; autoplay; clipboard-wr
ite; encrypted-media; gyroscope; picture-in-picture" allowfullscreen></iframe>')
```

Out[1]:



Lab(s)