13/05/2019 Untitled

## **FINAL CODE**

In [5]: import requests from bs4 import BeautifulSoup import pandas as pd parent url = "http://www.agriculture.gov.au" request = requests.get("http://www.agriculture.gov.au/pests-diseases-weeds/plant#identify-pests-diseases") content = request.text soup = BeautifulSoup(content, "html.parser") elements = soup.find\_all(href = True) pattern = "/pests-diseases-weeds/plant/" un\_nes = ['xylella/', 'national', 'quarantine/'] urls = []for e in elements: if pattern in e['href']: if any(x in e['href'] for x in un nes): pass else: urls.append(parent\_url+ e['href']) urls = list(set(urls)) records = [] # New list to append disease name,img url, origin, identity, specimen and what can came in to AUS # particular disease for url in urls: request1 = requests.get(url) content1 = request1.text soup1 = BeautifulSoup(content1, "html.parser") dname = soup1.find all('h2')[0].text img =parent\_url + soup1.find\_all('img')[3]['src'] origin = soup1.find\_all('strong')[1].next\_sibling identity = soup1.find\_all('strong')[3].next\_sibling come\_outof\_aus = soup1.find\_all("div", {"class":"hide"})[1].text specimen = soup1.find all("div", {"class":"hide"})[2].text records.append([dname, img, origin, identity, come outof aus, specimen])

df = pd.DataFrame(records, columns=['Disease Name', 'Image Link', 'Origin', 'identification\_of\_pest', 'what can I
df.to\_csv('Agri\_scrap.csv', index = False, encoding='utf-8-sig')
df

## Out[5]:

	Disease Name	lmage Link	Origin	identification_of_pest	what can legally come into Australia	Secure any suspect specimens
0	Potato cyst nemotade	http://www.agriculture.gov.au/SiteCollectionIm	South America	Leaf wilting and discoloration, root cysts an	All Australians and international tourists hav	Remember the pest can spread easily via the mo
1	Zebra chip	http://www.agriculture.gov.au/SiteCollectionIm	Mexico	USA, Canada, Mexico, Guatemala, Honduras 	All Australians and international tourists hav	Containment is critical. This could be as eas
2	Gypsy moth	http://www.agriculture.gov.au/SiteCollectionIm	China, East Russia, Southern Europe and Northe	Males and female moths differ in sizing and c	All Australians and international tourists hav	Containment is crucial. If you suspect you've 
3	Guava (eucalyptus) rust	http://www.agriculture.gov.au/SiteCollectionIm	Brazil	Yellowish pustules on young leaves, shoots, s	All Australians and international tourists ha	Containment is critical. In as little
4	Internal and external \r\n mites of bees	http://www.agriculture.gov.au/SiteCollectionIm	Asia (tropilaelaps and varroa mites), unknown	Four pairs of legs, 0.5 mm x 1 mm (tropilaela	All Australians and international tourists hav	Containment is crucial. If you suspect you've
5	Panama disease Tropical Race 4	http://www.agriculture.gov.au/SiteCollectionIm	South East Asia	<strong>Distribution: </strong>	All Australians and international tourists ha	Containment is critical. If you suspect
6	Khapra beetle	http://www.agriculture.gov.au/SiteCollectionIm	India	Adult – tiny, dark brown, oval shaped, covere	All Australians and international tourists hav	Containment is critical. Khapra beetles do not
7	Leaf miner	http://www.agriculture.gov.au/SiteCollectionIm	South America		All Australians and international tourists hav	Containment is critical. Alert the department

13/05/2019 Untitled

	Disease Name	lmage Link	Origin	identification_of_pest	what can legally come into Australia	Secure any suspect specimens
8	Russian wheat aphid	http://www.agriculture.gov.au/SiteCollectionIm	Central Asia and southern Russia	Light green, elongated spindle-shaped aphid, a	All Australians and international tourists hav	Containment is critical. This could be as eas
9	Exotic fruit fly	http://www.agriculture.gov.au/SiteCollectionIm	Asia	Tephritids - 2.5–10 mm, colourful, pictured w	All Australians and international tourists hav	Containment is critical. Fruit fly adults are
10	Exotic bees	http://www.agriculture.gov.au/SiteCollectionIm	Varies, depending on the species	Various sizes up to 2cm, striped abdomens, of	Find out what can legally come into Australia	Containment is critical. If you work around im
11	Brown marmorated stink bug	http://www.agriculture.gov.au/SiteCollectionIm	East Asia	Shield shape, adults are mottled reddish-brow	All Australians and international tourists hav	Containment is critical. The bug is a strong
12	Ug99	http://www.agriculture.gov.au/SiteCollectionIm	Uganda	Reddish brown, powdery, oblong spore- producin	All Australians and international tourists hav	Containment is critical.If you suspect you h
13	Citrus canker	http://www.agriculture.gov.au/SiteCollectionIm	South East Asia	Forms lesions on fruit, leaves and branches, c	All Australians and international tourists ha	Containment is critical.Look out for scabby
14	Giant African snail	http://www.agriculture.gov.au/SiteCollectionIm	Eastern Africa	Large, long and narrow conical shell, eggs	All Australians and international tourists hav	Containment is critical. Giant African snails
15	Karnal bunt	http://www.agriculture.gov.au/SiteCollectionIm	India	Broken hollow grain, powdery masses of dark s	All Australians and international tourists hav	Containment is critical. This can be as easy a
16	Xylella	http://www.agriculture.gov.au/SiteCollectionIm	Americas	Scorched leaves, browning and loss of leaves,	All Australians and international tourists hav	Containment is critical. This could be as easy
17	Tramp ants	http://www.agriculture.gov.au/SiteCollectionIm	Varies, depending on the species	Classic ant body shape, under 5 mm, ability t	All Australians and international tourists hav	Containment is critical. If you suspect you've

13/05/2019 Untitled

_		Disease Name	lmage Link	Origin	identification_of_pest	what can legally come into Australia	Secure any suspect specimens
	18	Airborne phytophthora	http://www.agriculture.gov.au/SiteCollectionIm	Asia	Oomycete causing leaf spots, shoot/branch dieb	All Australians and international tourists hav	Containment is critical.It is important to not
	19	Huanglongbing	http://www.agriculture.gov.au/SiteCollectionIm	Asia and Africa	Yellowing, blotchy mottling and unseasonal le	All Australians and international tourists hav	Containment is critical. This could be as easy

In [ ]: