Regex to scrape phone numbers and emails from your clipboard

In [ ]:
```python
# Code will scrape the phone and email from https://www.phila.gov/departments/
department-of-revenue/about/contact-us/
import re, pyperclip

# Create a regex for phone numbers
phoneRegex = re.compile(r'''
# types of numbers 415-555-5555, (610) 555-5555, 555,0000, 555-0000 ext. 12345
 x12345

(((\d\d\d)|(\(\d\d\d\)))?     # area code (optional)
(\s|-)                       # first separator
\d\d\d                       # first 3 digits
-                            # separator
\d\d\d\d                     # last 4 digits
(((ext(\.)?\s)|x)             # extension word-part (optional)
(\d{2,5}))?                  # extension number-part (optional)
)
''', re.VERBOSE)

# Create a regex for email addresses
emailRegex = re.compile(r'''

# some.+_thing@(\d{2,5}))?.com

[a-zA-Z0-9_.+]+         # name part
@                       # @ symbol
[a-zA-Z0-9_.+]+         # domain name part

''', re.VERBOSE)
```

In [7]:
```python
# Get the text off the clipboard
text = pyperclip.paste()

# Extract the email/phone from this text
extractedPhone = phoneRegex.findall(text)
extractedEmail = emailRegex.findall(text)

allPhoneNumbers = []
for phoneNumber in extractedPhone:
    allPhoneNumbers.append(phoneNumber[0])

allEmail = []
for email in extractedEmail:
    allEmail.append(email)

print(extractedPhone)
print(extractedEmail)
```

```
[('(215) 686-6600', '(215)', '', '(215)', ' ', '', '', '', '', ''), ('(215) 6
86-6831', '(215)', '', '(215)', ' ', '', '', '', '', ''), ('(877) 309-3710',
'(877)', '', '(877)', ' ', '', '', '', '', ''), ('(215) 686-6442', '(215)',
'', '(215)', ' ', '', '', '', '', ''), ('(215) 686-4343', '(215)', '', '(21
5)', ' ', '', '', '', '', ''), ('(215) 686-4336', '(215)', '', '(215)', ' ',
'', '', '', '', ''), ('(877) 309-3709', '(877)', '', '(877)', ' ', '', '',
'', '', ''), ('(215) 685-6300', '(215)', '', '(215)', ' ', '', '', '', '',
''), ('(215) 685-6300', '(215)', '', '(215)', ' ', '', '', '', '', ''), ('(21
5) 686-6565', '(215)', '', '(215)', ' ', '', '', '', '', ''), ('(215) 686-657
4', '(215)', '', '(215)', ' ', '', '', '', '', ''), ('(215) 686-6575', '(21
5)', '', '(215)', ' ', '', '', '', '', ''), ('(215) 686-6578', '(215)', '',
'(215)', ' ', '', '', '', '', ''), ('(215) 686-2670', '(215)', '', '(215)', '
', '', '', '', '', ''), ('(215) 686-6648', '(215)', '', '(215)', ' ', '', '',
'', '', '')]
['revenue@phila.gov', 'vicki.riley@phila.gov', 'revenue@phila.gov', 'wrbhelpd
esk@phila.gov', 'tax.clearance@phila.gov', 'refundunit@phila.gov', 'agency.re
ceivables@phila.gov', 'egovservices@phila.gov', 'amountdue@phila.gov', 'water
amountdue@phila.gov']
```

In [12]:
```python
# shows first 10 phone numbers
allPhoneNumbers[:10]
```

Out[12]:
```
['(215) 686-6600',
 '(215) 686-6831',
 '(877) 309-3710',
 '(215) 686-6442',
 '(215) 686-4343',
 '(215) 686-4336',
 '(877) 309-3709',
 '(215) 685-6300',
 '(215) 685-6300',
 '(215) 686-6565']
```

In [13]:
```python
# shows first 10 email addresses
allEmail[:10]
```

Out[13]:
```
['revenue@phila.gov',
 'vicki.riley@phila.gov',
 'revenue@phila.gov',
 'wrbhelpdesk@phila.gov',
 'tax.clearance@phila.gov',
 'refundunit@phila.gov',
 'agency.receivables@phila.gov',
 'egovservices@phila.gov',
 'amountdue@phila.gov',
 'wateramountdue@phila.gov']
```

In [34]:
```python
# copies results onto your computer's clipboard
results = '\n'.join(allPhoneNumbers) + '\n' + '\n'.join(extractedEmail)
pyperclip.copy(results)
```