ArcPy, Pandas, Beautiful Soup and Chronic Wasting Disease

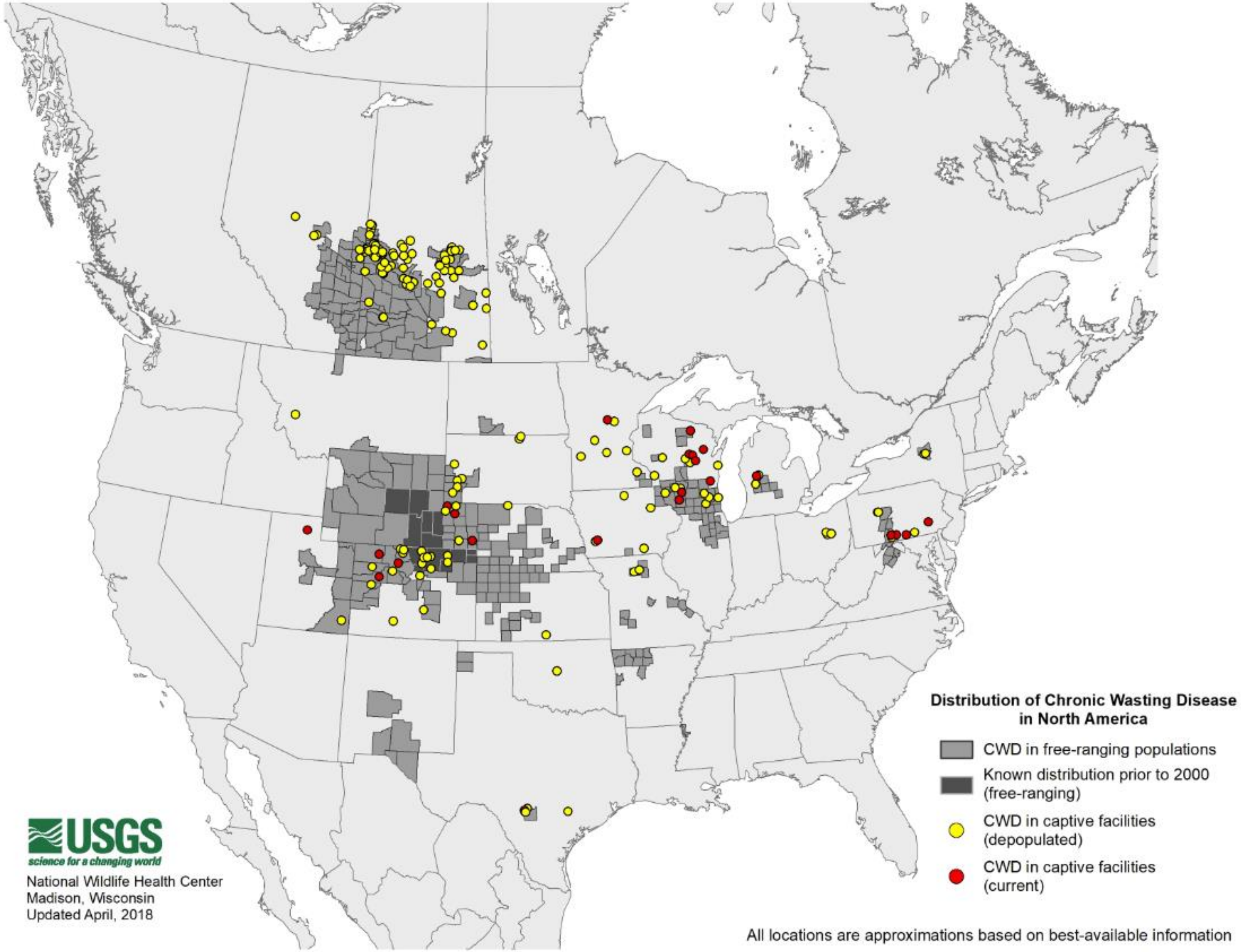
Automating the Weekly Report

Cristy G. Burch
Regional GIS Specialist

TEXAS PARKS & WILDLIFE

Life's better outside.®

Chronic Wasting Disease (CWD)


Distribution of Chronic Wasting Disease in North America
▨ CWD in free-ranging populations
■ Known distribution prior to 2000 (free-ranging)
● CWD in captive facilities (depopulated)
● CWD in captive facilities (current)
USGS
science for a changing world
National Wildlife Health Center
Madison, Wisconsin
Updated April, 2018
All locations are approximations based on best-available information

- Neurological disease in deer, elk, moose and other members of the deer family, known as “cervids”
- First case of CWD in Texas was discovered in 2012 in free-ranging mule deer in an isolated area of far West Texas

Concerns

- Decline within deer, elk or other susceptible cervid populations
- Indirect impacts on hunting, hunter participation and economic benefits derived from big game hunting
  - In Texas, hunting is a $2.2 billion economic engine, supporting many rural towns

Photo Credit: Warden Michael Hopper, Kansas Dept of Wildlife, Parks & Tourism

Management program

- Texas Parks and Wildlife Department (TPWD) and Texas Animal Health Commission (TAHC) have developed a cooperative CWD management plan to guide both agencies

TEXAS
PARKS &
WILDLIFE

Life's better outside.®

TEXAS ANIMAL HEALTH COMMISSION
Since 1893
SERVING TEXAS ANIMAL AGRICULTURE

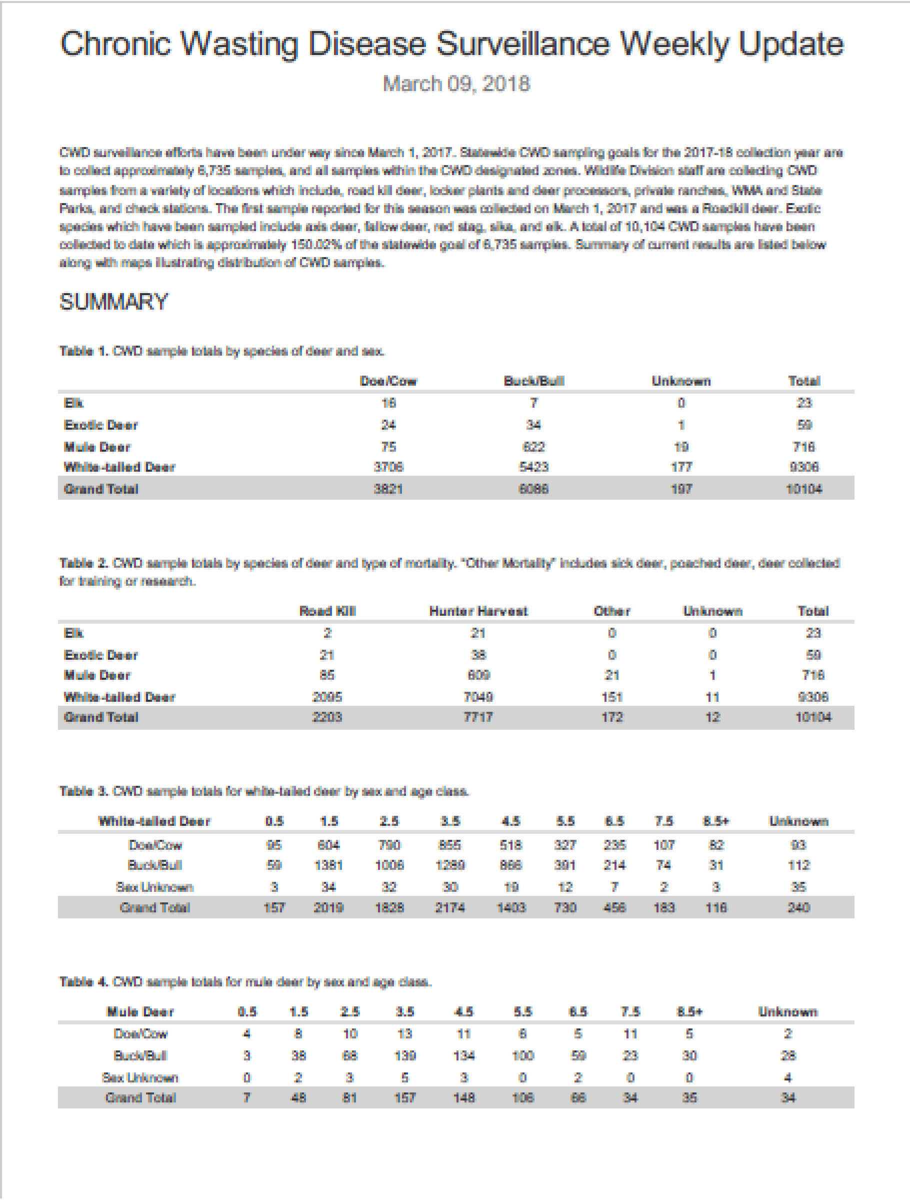
Intensive Sampling

- Wildlife Division staff collect lymph nodes from roadkill and hunter harvested animals.
- While collection takes place year round, most intensive sampling occurs during white-tailed deer season, approximately November through February each year
- The state is broken into units for wildlife management purposes; each unit has a sampling goal

Workflow

- Field data is collected with a smartphone using ArcGIS Collector
- Weekly sample status report is generated during peak sample collection (Nov-Feb)
- Report published to SharePoint for internal distribution and sent to TAHC (external distribution)
- Report consists of summary tables, charts and maps depicting sample collection status

.PDF



Chronic Wasting Disease Surveillance Weekly Update

March 09, 2018

CWD surveillance efforts have been under way since March 1, 2017. Statewide CWD sampling goals for the 2017-18 collection year are to collect approximately 6,735 samples, and all samples within the CWD designated zones. Wildlife Division staff are collecting CWD samples from a variety of locations which include, road kill deer, locker plants and deer processors, private ranches, WMA and State Parks, and check stations. The first sample reported for this season was collected on March 1, 2017 and was a Roadkill deer. Exotic species which have been sampled include axis deer, fallow deer, red stag, sika, and elk. A total of 10,104 CWD samples have been collected to date which is approximately 150.02% of the statewide goal of 6,735 samples. Summary of current results are listed below along with maps illustrating distribution of CWD samples.

SUMMARY

Table 1. CWD sample totals by species of deer and sex.

	Doe/Cow	Buck/Bull	Unknown	Total
Elk	16	7	0	23
Exotic Deer	24	34	1	59
Mule Deer	75	622	19	716
White-tailed Deer	3706	5423	177	9306
Grand Total	3821	6086	197	10104

Table 2. CWD sample totals by species of deer and type of mortality. "Other Mortality" includes sick deer, poached deer, deer collected for training or research.

	Road Kill	Hunter Harvest	Other	Unknown	Total
Elk	2	21	0	0	23
Exotic Deer	21	38	0	0	59
Mule Deer	85	609	21	1	716
White-tailed Deer	2095	7049	151	11	9306
Grand Total	2203	7717	172	12	10104

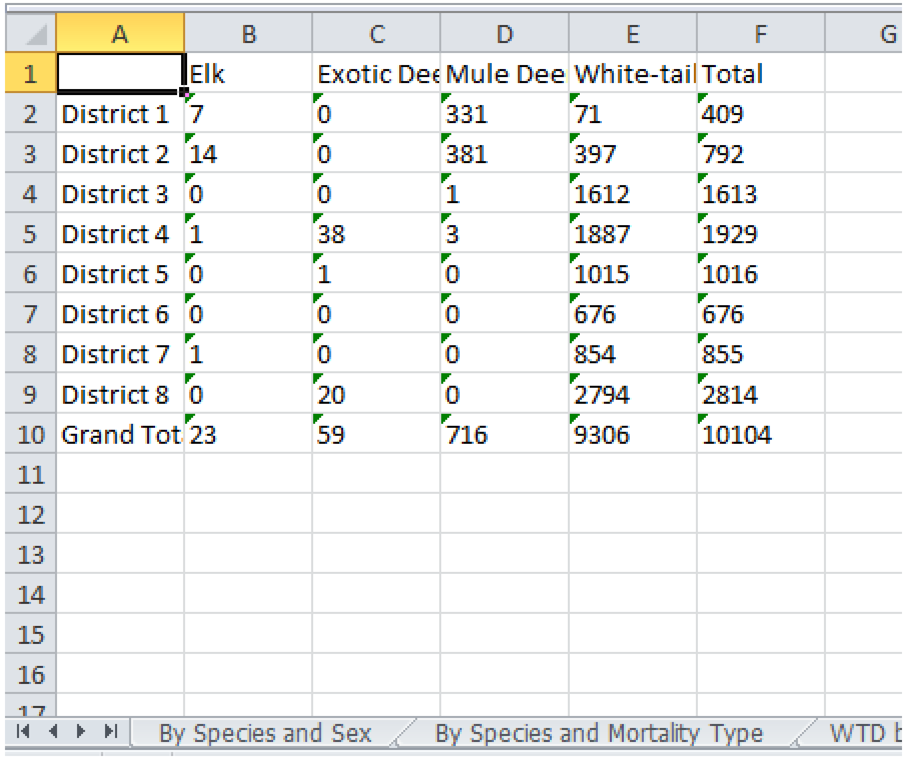
Table 3. CWD sample totals for white-tailed deer by sex and age class.

White-tailed Deer	0.5	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5+	Unknown
Doe/Cow	95	604	790	855	518	327	235	107	82	93
Buck/Bull	59	1381	1006	1289	866	391	214	74	31	112
Sex Unknown	3	34	32	30	19	12	7	2	3	35
Grand Total	157	2019	1828	2174	1403	730	456	183	116	240

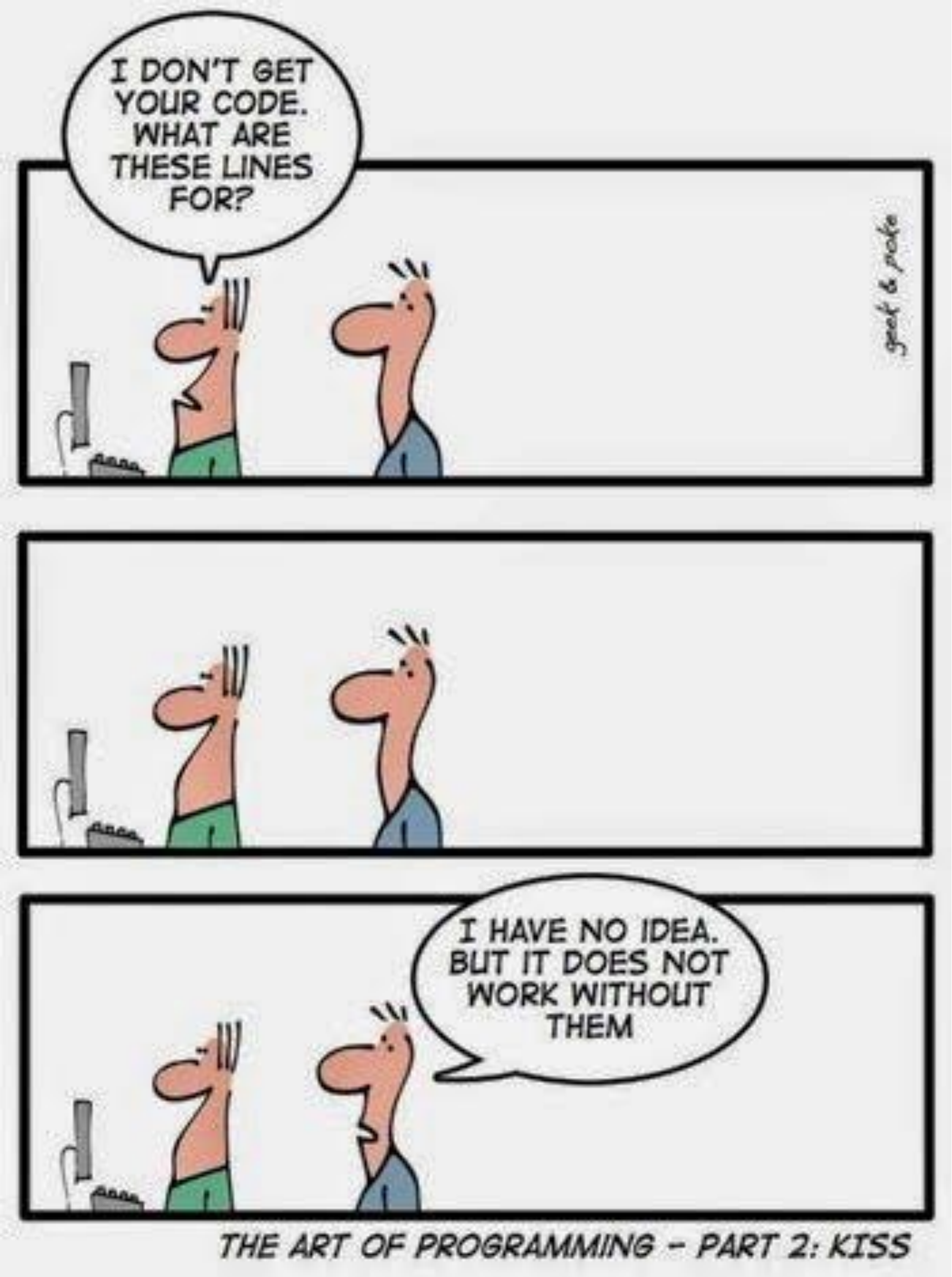
Table 4. CWD sample totals for mule deer by sex and age class.

Mule Deer	0.5	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5+	Unknown
Doe/Cow	4	8	10	13	11	6	5	11	5	2
Buck/Bull	3	38	68	139	134	100	59	23	30	28
Sex Unknown	0	2	3	5	3	0	2	0	0	4
Grand Total	7	48	81	157	148	106	66	34	35	34

.XLSX

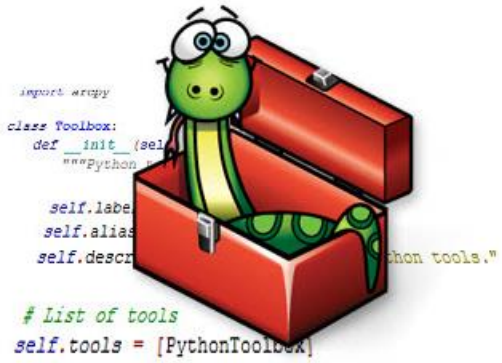


	A	B	C	D	E	F	G
1		Elk	Exotic Dee	Mule Dee	White-tail	Total	
2	District 1	7	0	331	71	409	
3	District 2	14	0	381	397	792	
4	District 3	0	0	1	1612	1613	
5	District 4	1	38	3	1887	1929	
6	District 5	0	1	0	1015	1016	
7	District 6	0	0	0	676	676	
8	District 7	1	0	0	854	855	
9	District 8	0	20	0	2794	2814	
10	Grand Tot	23	59	716	9306	10104	
11							
12							
13							
14							
15							
16							
17							


I DON'T GET YOUR CODE. WHAT ARE THESE LINES FOR?
geek & poke
I HAVE NO IDEA. BUT IT DOES NOT WORK WITHOUT THEM
THE ART OF PROGRAMMING - PART 2: KISS

ENTHOUGHT
CANOPY

+



```
import arcpy

class Toolbox:
    def __init__(sel
        """Python 

        self.labe
       self.alias
      self.descr                    thon tools."

 # List of tools
self.tools = [PythonToolbox]
```

- Install ArcGIS and 64bit background processing
- Install Canopy 64bit
- Create and paste .pth files in the appropriate locations
  - Paste zz_arcpy.pth to canopy site-packages
  - Paste zz_canopy.pth toArcGIS site-packages

There shouldn't be any reference to ArcGIS or Canopy in the System Path environment variable. If there is, remove it.

Contents of .pth files:

- #.pth file for EPD Canopy

  C:\Users\*[username]*\AppData\Local\Enthought\Canopy\User\Lib\site-packages

  C:\Users\*[username]*\AppData\Local\Enthought\Canopy\System\Lib\site-packages

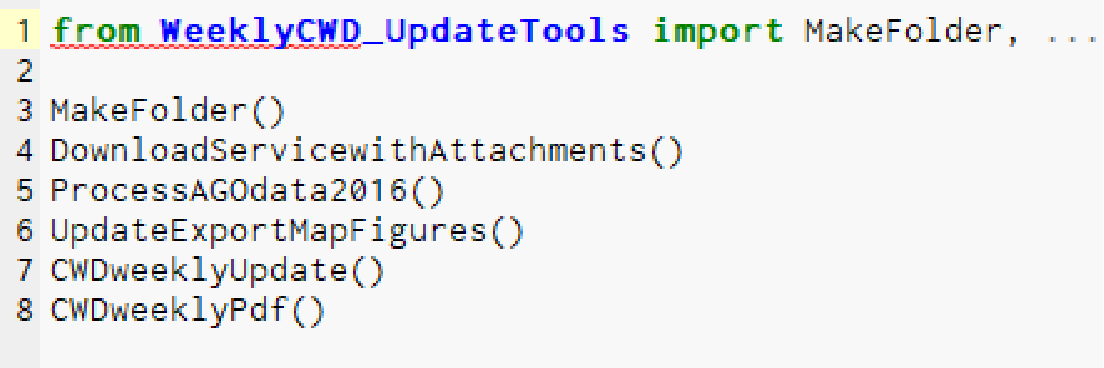
- # .pth file for arcpy

  C:\Program Files (x86)\ArcGIS\Desktop10.5\bin64

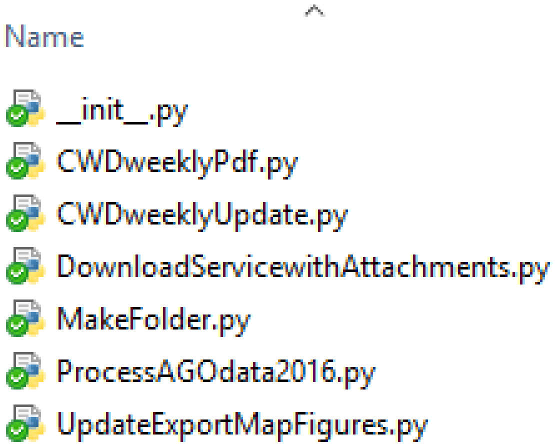
  C:\Program Files (x86)\ArcGIS\Desktop10.5\arcpy

  C:\Program Files (x86)\ArcGIS\Desktop10.5\ArcToolbox\Scripts

Six “Step” Process



```python
from WeeklyCWD_UpdateTools import MakeFolder, ...

MakeFolder()
DownloadServicewithAttachments()
ProcessAGOdata2016()
UpdateExportMapFigures()
CWDweeklyUpdate()
CWDweeklyPdf()
```

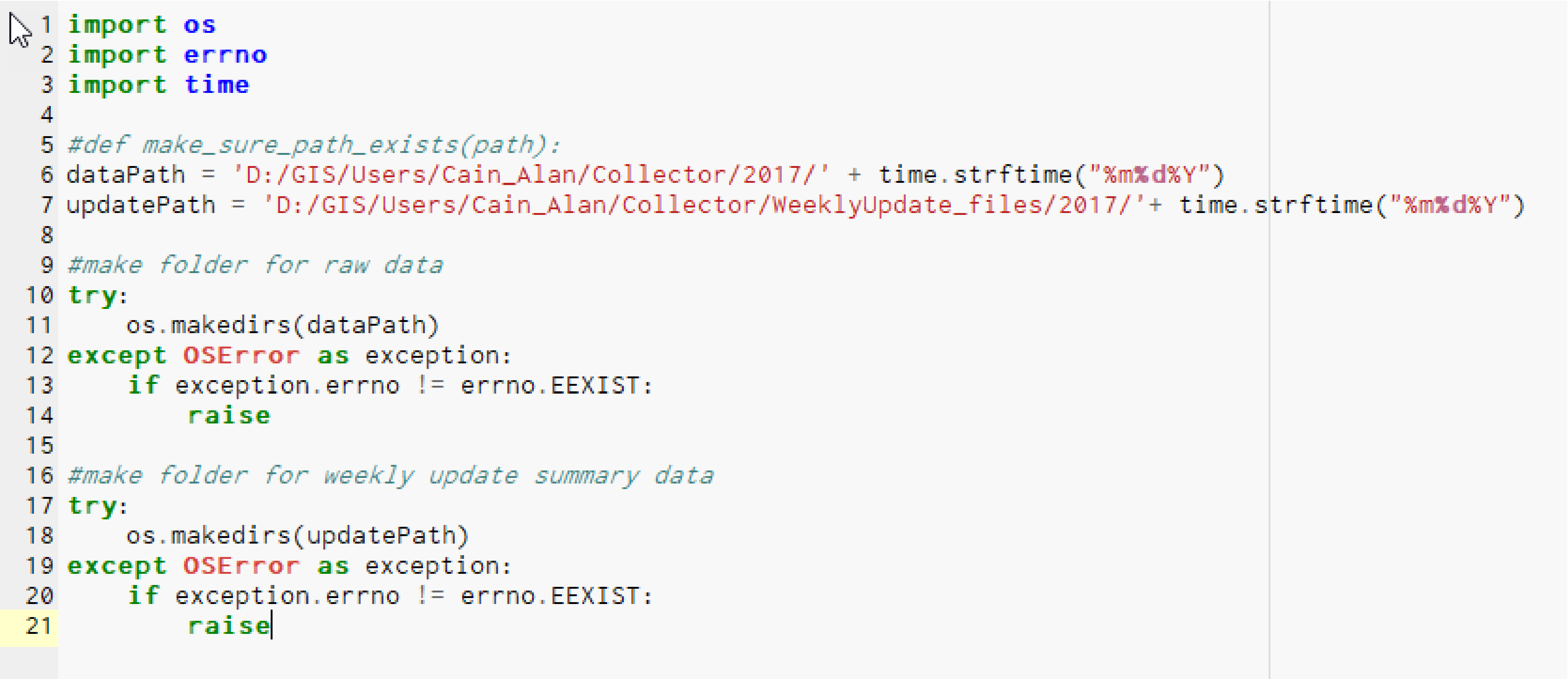


Name

init.py
CWDweeklyPdf.py
CWDweeklyUpdate.py
DownloadServicewithAttachments.py
MakeFolder.py
ProcessAGOdata2016.py
UpdateExportMapFigures.py

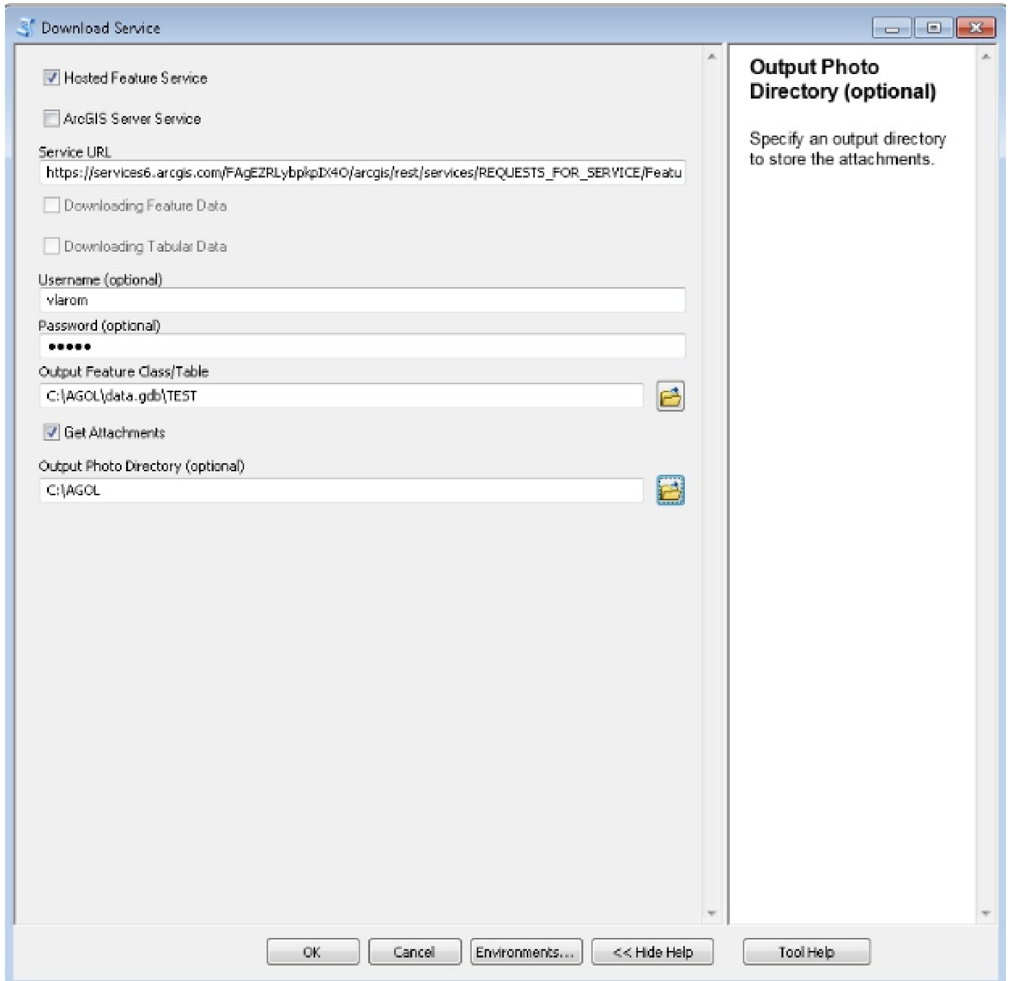
Sort of…

Step 1 – Create local folder



```
import os
import errno
import time

#def make_sure_path_exists(path):
dataPath = 'D:/GIS/Users/Cain_Alan/Collector/2017/' + time.strftime("%m%d%Y")
updatePath = 'D:/GIS/Users/Cain_Alan/Collector/WeeklyUpdate_files/2017/'+ time.strftime("%m%d%Y")

#make folder for raw data
try:
    os.makedirs(dataPath)
except OSError as exception:
    if exception.errno != errno.EEXIST:
        raise

#make folder for weekly update summary data
try:
    os.makedirs(updatePath)
except OSError as exception:
    if exception.errno != errno.EEXIST:
        raise
```

Step 2 – Download data from AGO



☑ Hosted Feature Service

□ ArcGIS Server Service

Service URL
https://services6.arcgis.com/FAgEZRLybpkpIX4O/arcgis/rest/services/REQUESTS_FOR_SERVICE/Featu

□ Downloading Feature Data

□ Downloading Tabular Data

Username (optional)
vlarom

Password (optional)
●●●●●

Output Feature Class/Table
C:\AGOL\data.gdb\TEST

☑ Get Attachments

Output Photo Directory (optional)
C:\AGOL

Output Photo Directory (optional)

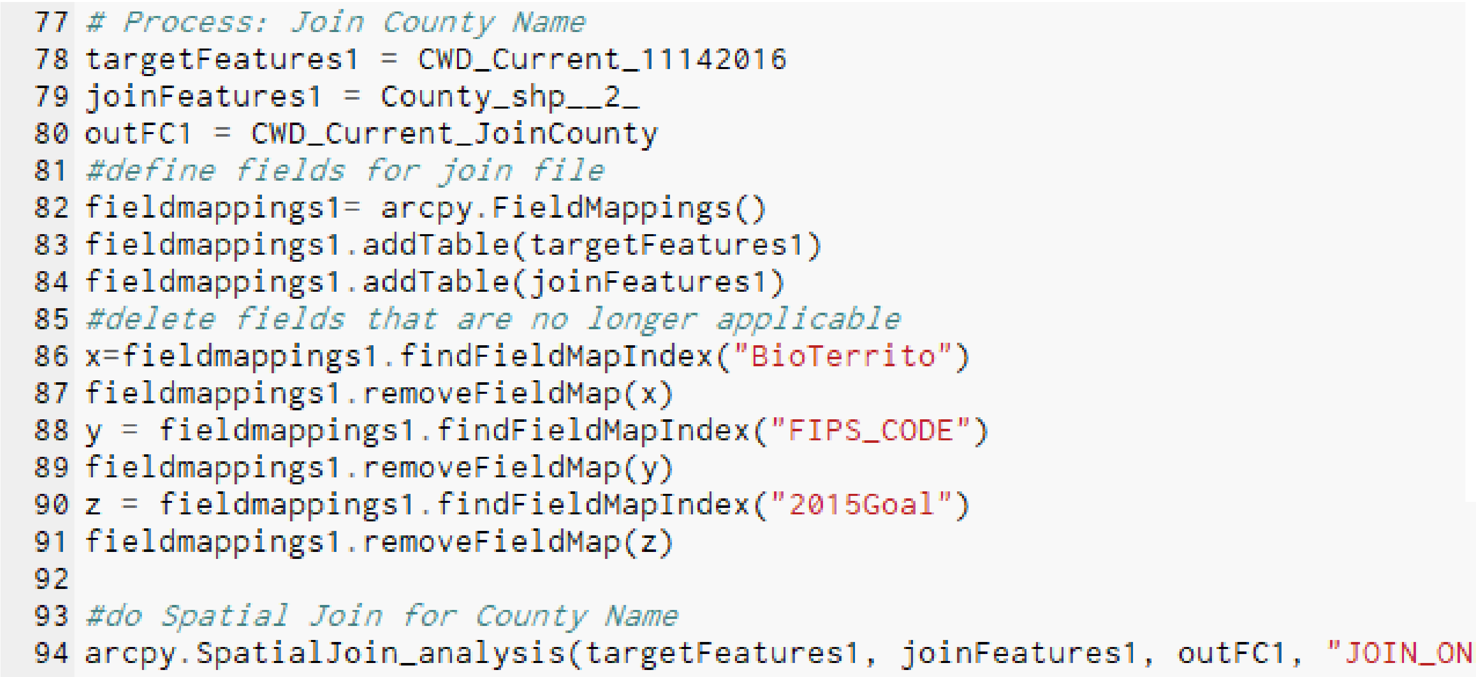
Specify an output directory to store the attachments.

- Download Service tool exported to script and adjusted to allow for input from variable
- https://community.esri.com/docs/DOC-6496-download-arcgis-online-feature-service-or-arcgis-server-featuremap-service

Step 3 – Process data with ArcPy

- Spatial joins to populate
  - County
  - Wildlife District
  - Deer Management Unit number
  - Mule Deer Management Unit name
  - CWD Zone name
  - Sample Goals
- Delete extra fields
- Export data into Excel – *this may not be necessary*

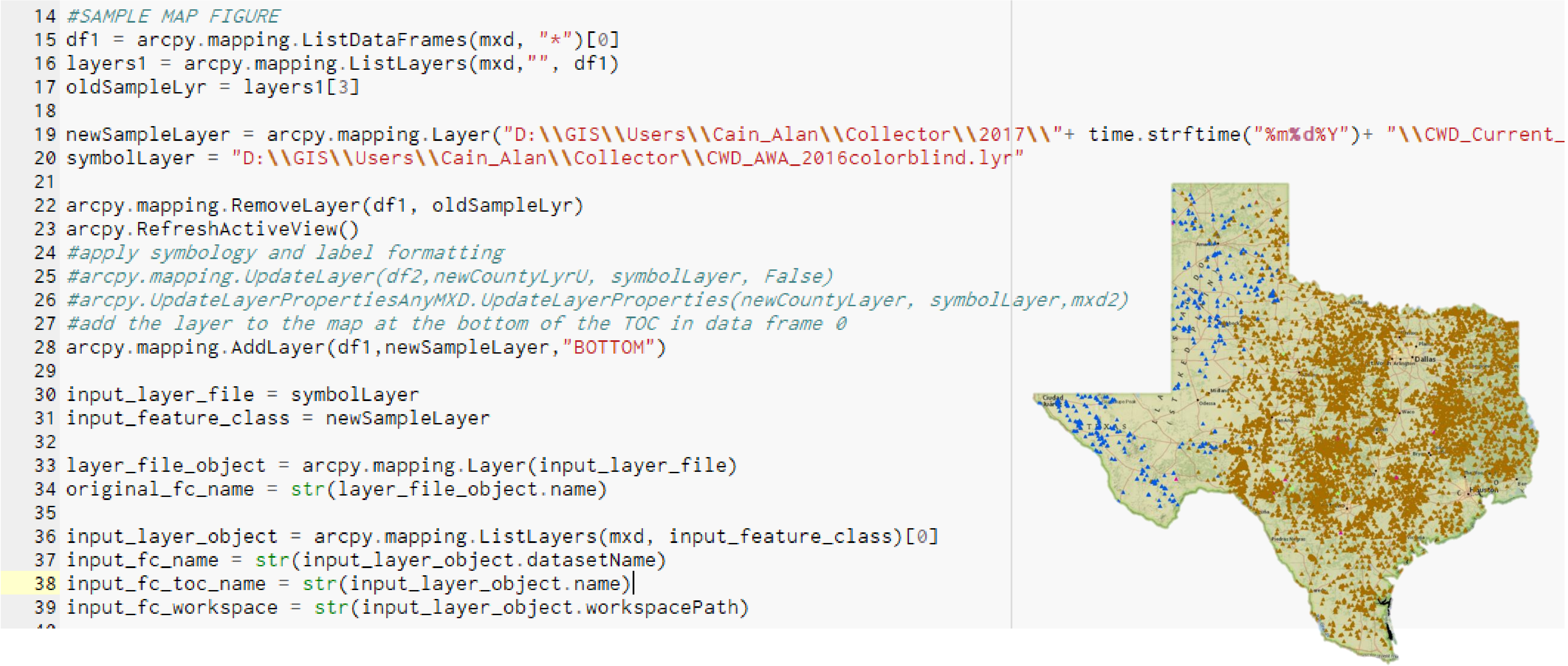
Step 3, cont. – Process data with ArcPy



```
# Process: Join County Name
targetFeatures1 = CWD_Current_11142016
joinFeatures1 = County_shp__2_
outFC1 = CWD_Current_JoinCounty
#define fields for join file
fieldmappings1= arcpy.FieldMappings()
fieldmappings1.addTable(targetFeatures1)
fieldmappings1.addTable(joinFeatures1)
#delete fields that are no longer applicable
x=fieldmappings1.findFieldMapIndex("BioTerrito")
fieldmappings1.removeFieldMap(x)
y = fieldmappings1.findFieldMapIndex("FIPS_CODE")
fieldmappings1.removeFieldMap(y)
z = fieldmappings1.findFieldMapIndex("2015Goal")
fieldmappings1.removeFieldMap(z)

#do Spatial Join for County Name
arcpy.SpatialJoin_analysis(targetFeatures1, joinFeatures1, outFC1, "JOIN_ONE_TO_ONE", "KEEP_ALL", fieldmappings1)
```

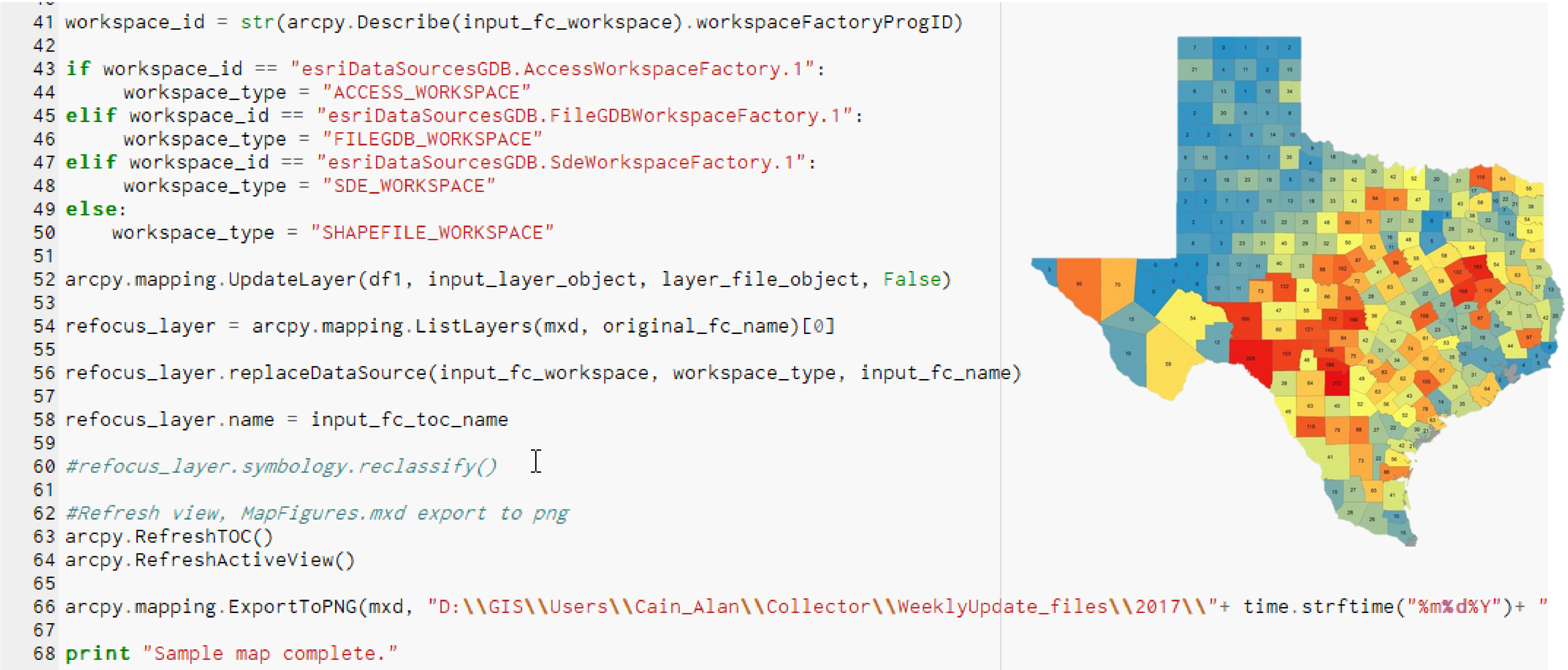
Image credit: Café Press, www.cafepress.com

Step 4 – Update Map figures with ArcPy



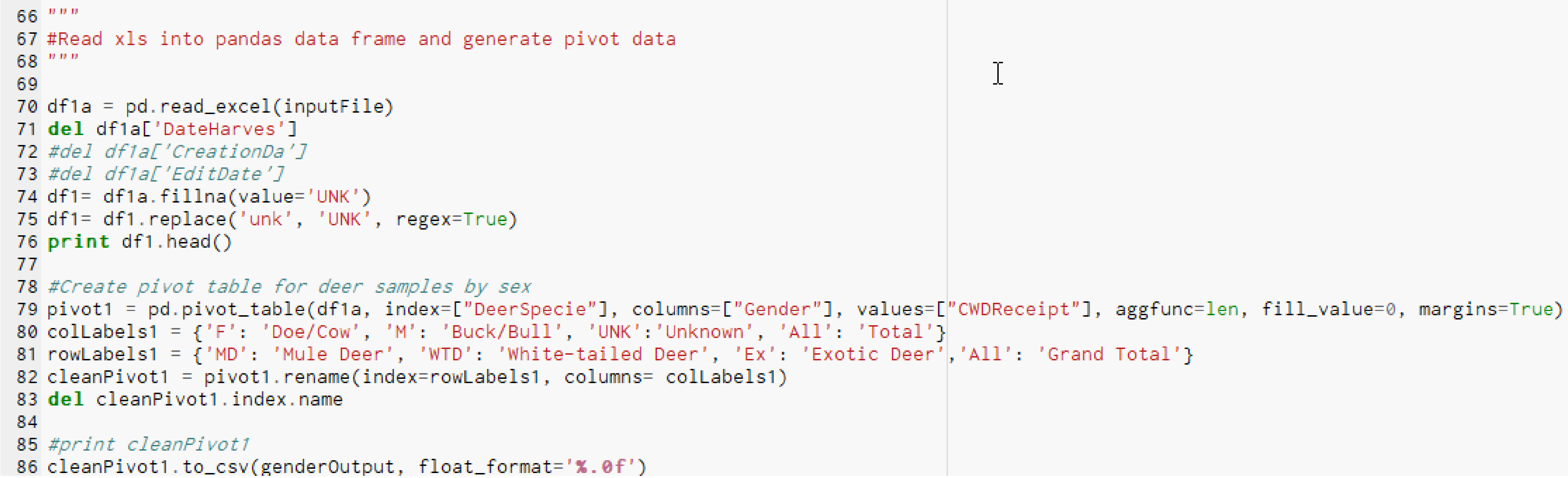
```
#SAMPLE MAP FIGURE
df1 = arcpy.mapping.ListDataFrames(mxd, "*")[0]
layers1 = arcpy.mapping.ListLayers(mxd,"", df1)
oldSampleLyr = layers1[3]

newSampleLayer = arcpy.mapping.Layer("D:\\GIS\\Users\\Cain_Alan\\Collector\\2017\\"+ time.strftime("%m%d%Y")+ "\\CWD_Current_
symbolLayer = "D:\\GIS\\Users\\Cain_Alan\\Collector\\CWD_AWA_2016colorblind.lyr"

arcpy.mapping.RemoveLayer(df1, oldSampleLyr)
arcpy.RefreshActiveView()
#apply symbology and label formatting
#arcpy.mapping.UpdateLayer(df2,newCountyLyrU, symbolLayer, False)
#arcpy.UpdateLayerPropertiesAnyMXD.UpdateLayerProperties(newCountyLayer, symbolLayer,mxd2)
#add the layer to the map at the bottom of the TOC in data frame 0
arcpy.mapping.AddLayer(df1,newSampleLayer,"BOTTOM")

input_layer_file = symbolLayer
input_feature_class = newSampleLayer

layer_file_object = arcpy.mapping.Layer(input_layer_file)
original_fc_name = str(layer_file_object.name)

input_layer_object = arcpy.mapping.ListLayers(mxd, input_feature_class)[0]
input_fc_name = str(input_layer_object.datasetName)
input_fc_toc_name = str(input_layer_object.name)
input_fc_workspace = str(input_layer_object.workspacePath)
```

Step 4, cont. – Update Map figures with ArcPy



```
workspace_id = str(arcpy.Describe(input_fc_workspace).workspaceFactoryProgID)

if workspace_id == "esriDataSourcesGDB.AccessWorkspaceFactory.1":
     workspace_type = "ACCESS_WORKSPACE"
elif workspace_id == "esriDataSourcesGDB.FileGDBWorkspaceFactory.1":
     workspace_type = "FILEGDB_WORKSPACE"
elif workspace_id == "esriDataSourcesGDB.SdeWorkspaceFactory.1":
     workspace_type = "SDE_WORKSPACE"
else:
    workspace_type = "SHAPEFILE_WORKSPACE"

arcpy.mapping.UpdateLayer(df1, input_layer_object, layer_file_object, False)

refocus_layer = arcpy.mapping.ListLayers(mxd, original_fc_name)[0]

refocus_layer.replaceDataSource(input_fc_workspace, workspace_type, input_fc_name)

refocus_layer.name = input_fc_toc_name

#refocus_layer.symbology.reclassify()

#Refresh view, MapFigures.mxd export to png
arcpy.RefreshTOC()
arcpy.RefreshActiveView()

arcpy.mapping.ExportToPNG(mxd, "D:\\GIS\\Users\\Cain_Alan\\Collector\\WeeklyUpdate_files\\2017\\"+ time.strftime("%m%d%Y")+ "

print "Sample map complete."
```

Step 5 – Summary tables with Pandas

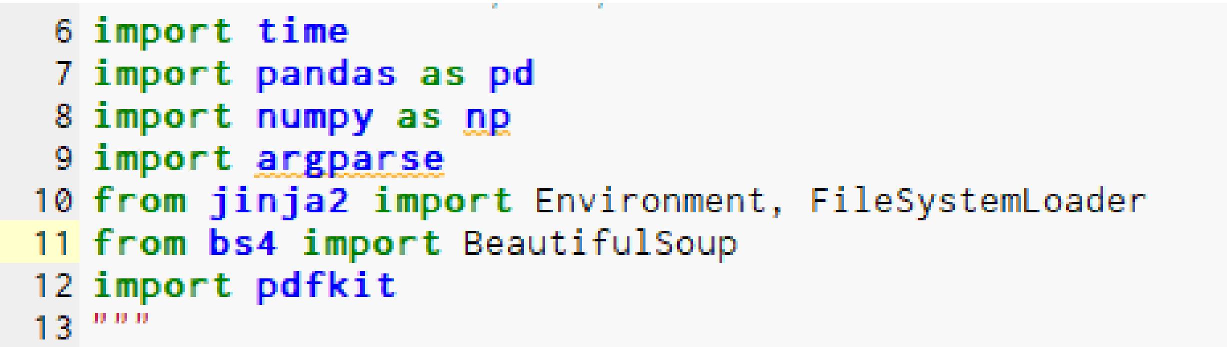


```
"""
#Read xls into pandas data frame and generate pivot data
"""

df1a = pd.read_excel(inputFile)
del df1a['DateHarves']
#del df1a['CreationDa']
#del df1a['EditDate']
df1= df1a.fillna(value='UNK')
df1= df1.replace('unk', 'UNK', regex=True)
print df1.head()

#Create pivot table for deer samples by sex
pivot1 = pd.pivot_table(df1a, index=["DeerSpecie"], columns=["Gender"], values=["CWDReceipt"], aggfunc=len, fill_value=0, margins=True)
colLabels1 = {'F': 'Doe/Cow', 'M': 'Buck/Bull', 'UNK':'Unknown', 'All': 'Total'}
rowLabels1 = {'MD': 'Mule Deer', 'WTD': 'White-tailed Deer', 'Ex': 'Exotic Deer','All': 'Grand Total'}
cleanPivot1 = pivot1.rename(index=rowLabels1, columns= colLabels1)
del cleanPivot1.index.name

#print cleanPivot1
cleanPivot1.to_csv(genderOutput, float_format='%.0f')
```

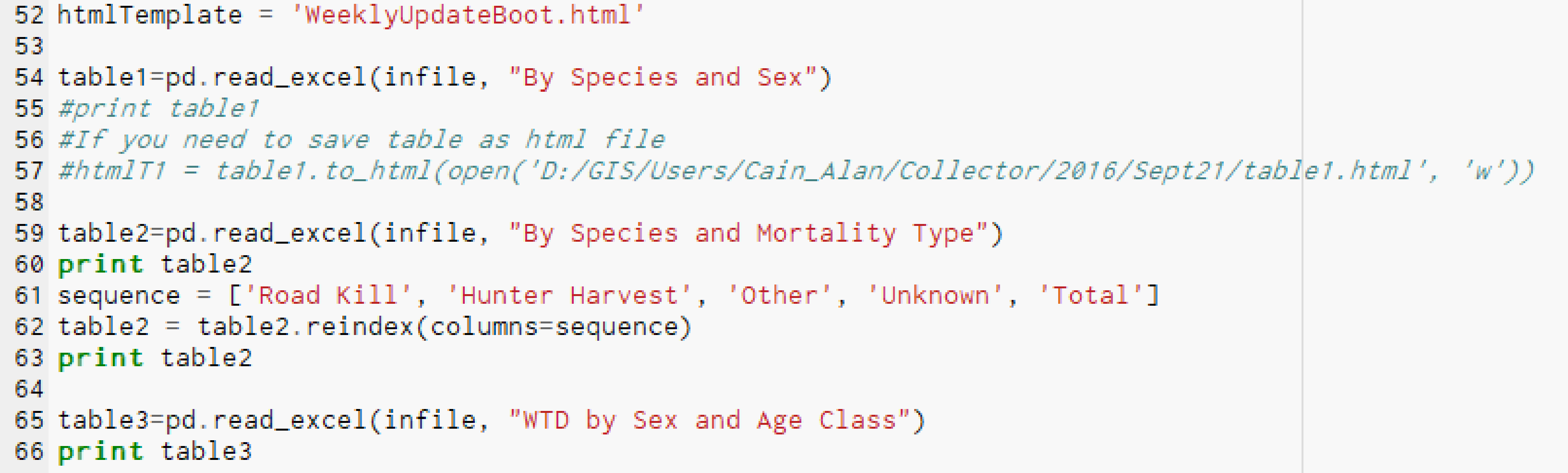
Step 5, cont. – Summary tables with Pandas

	Doe/Cow	Buck/Bull	Unknown	Total
Elk	16	7	0	23
Exotic Deer	24	34	1	59
Mule Deer	75	622	19	716
White-tailed Deer	3706	5423	177	9306
Grand Total	3821	6086	197	10104

Step 6 – Clean up using Beautiful Soup, then print

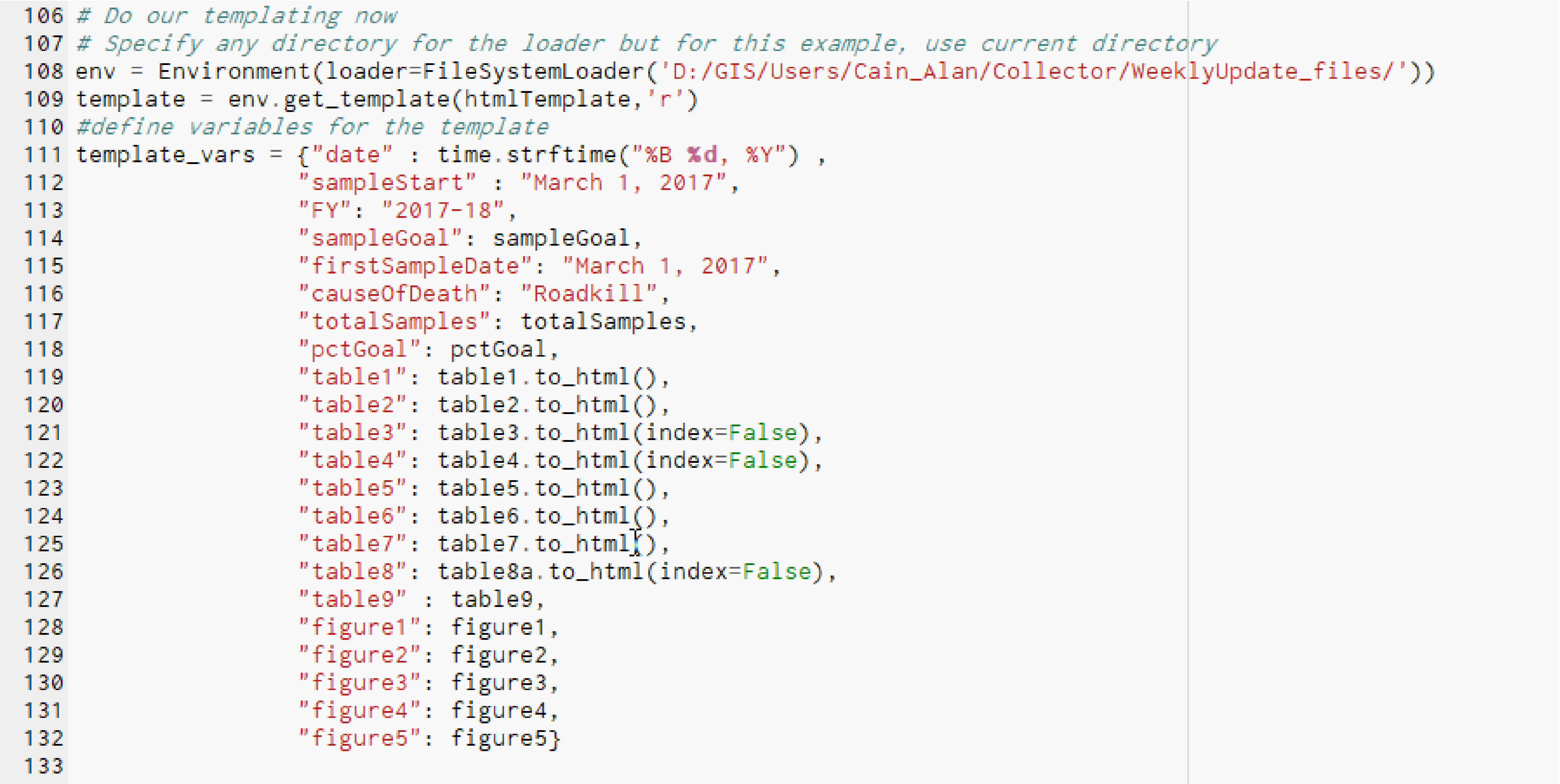


```
import time
import pandas as pd
import numpy as np
import argparse
from jinja2 import Environment, FileSystemLoader
from bs4 import BeautifulSoup
import pdfkit
"""
```



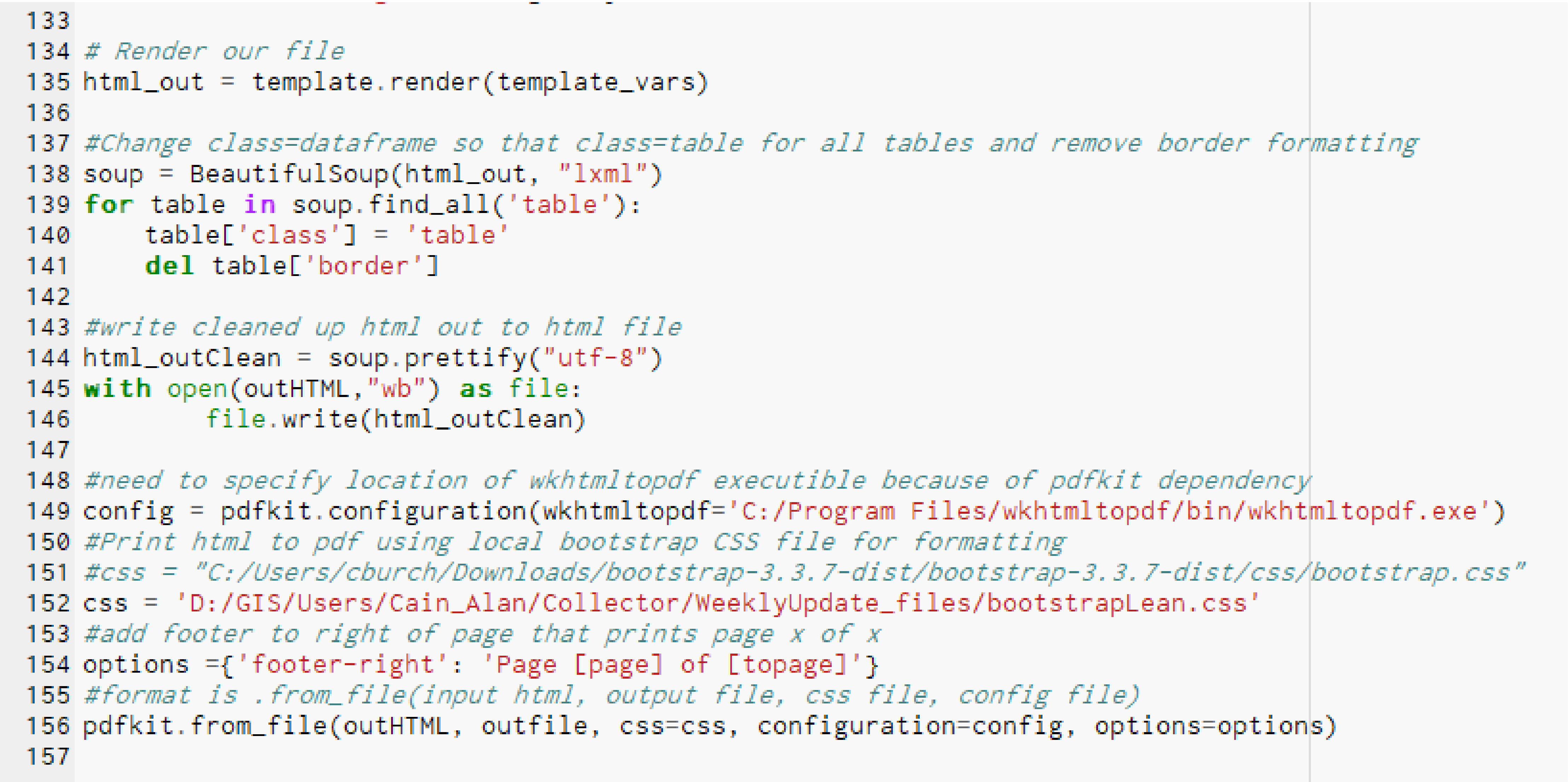
```
htmlTemplate = 'WeeklyUpdateBoot.html'

table1=pd.read_excel(infile, "By Species and Sex")
#print table1
#If you need to save table as html file
#htmlT1 = table1.to_html(open('D:/GIS/Users/Cain_Alan/Collector/2016/Sept21/table1.html', 'w'))

table2=pd.read_excel(infile, "By Species and Mortality Type")
print table2
sequence = ['Road Kill', 'Hunter Harvest', 'Other', 'Unknown', 'Total']
table2 = table2.reindex(columns=sequence)
print table2

table3=pd.read_excel(infile, "WTD by Sex and Age Class")
print table3
```

Step 6, cont. – Clean up using Beautiful Soup, then print



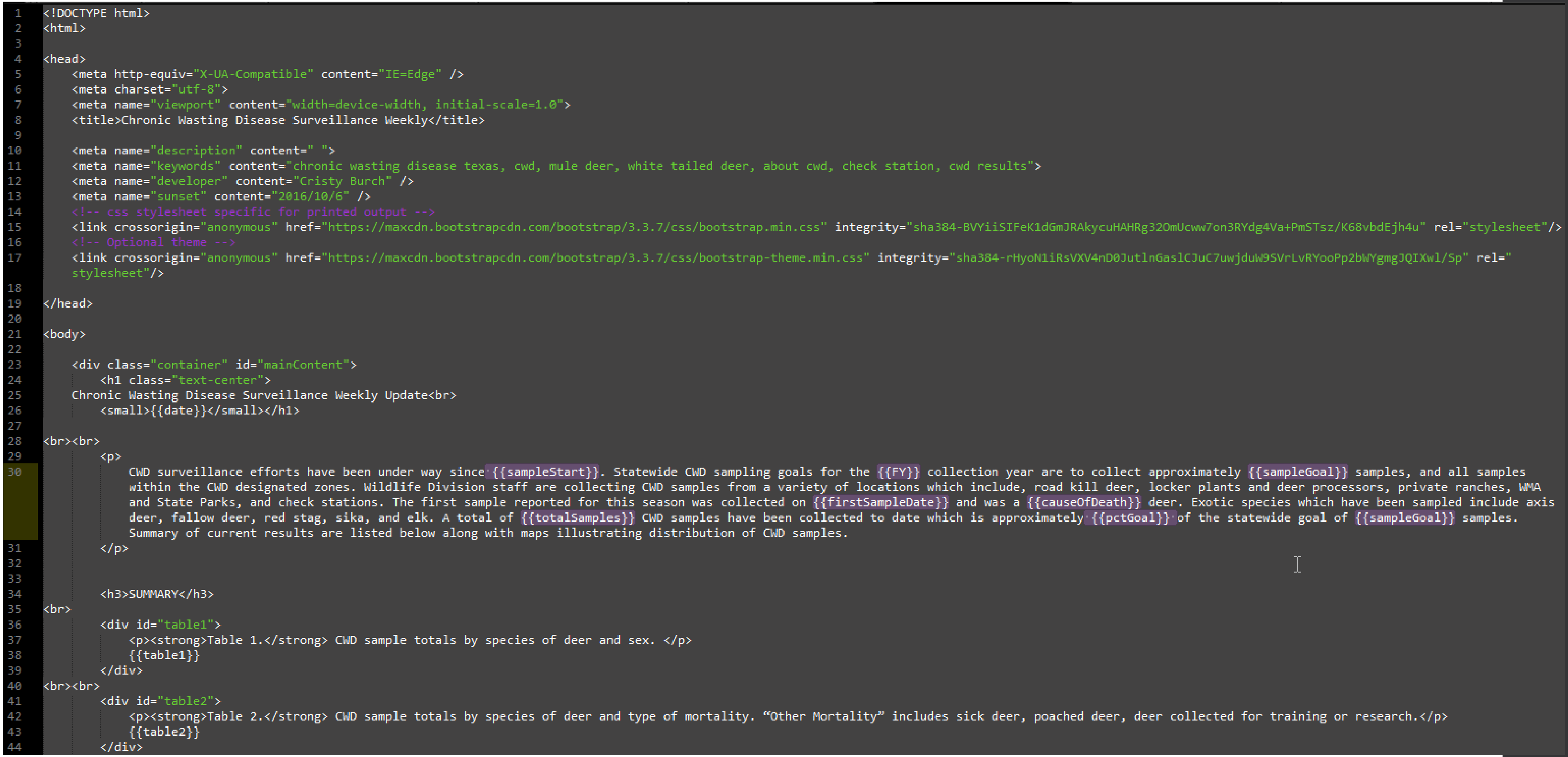
```python
# Do our templating now
# Specify any directory for the loader but for this example, use current directory
env = Environment(loader=FileSystemLoader('D:/GIS/Users/Cain_Alan/Collector/WeeklyUpdate_files/'))
template = env.get_template(htmlTemplate,'r')
#define variables for the template
template_vars = {"date" : time.strftime("%B %d, %Y") ,
                "sampleStart" : "March 1, 2017",
                "FY": "2017-18",
                "sampleGoal": sampleGoal,
                "firstSampleDate": "March 1, 2017",
                "causeOfDeath": "Roadkill",
                "totalSamples": totalSamples,
                "pctGoal": pctGoal,
                "table1": table1.to_html(),
                "table2": table2.to_html(),
                "table3": table3.to_html(index=False),
                "table4": table4.to_html(index=False),
                "table5": table5.to_html(),
                "table6": table6.to_html(),
                "table7": table7.to_html(),
                "table8": table8a.to_html(index=False),
                "table9" : table9,
                "figure1": figure1,
                "figure2": figure2,
                "figure3": figure3,
                "figure4": figure4,
                "figure5": figure5}

```

Step 6, cont. – Clean up using Beautiful Soup, then print



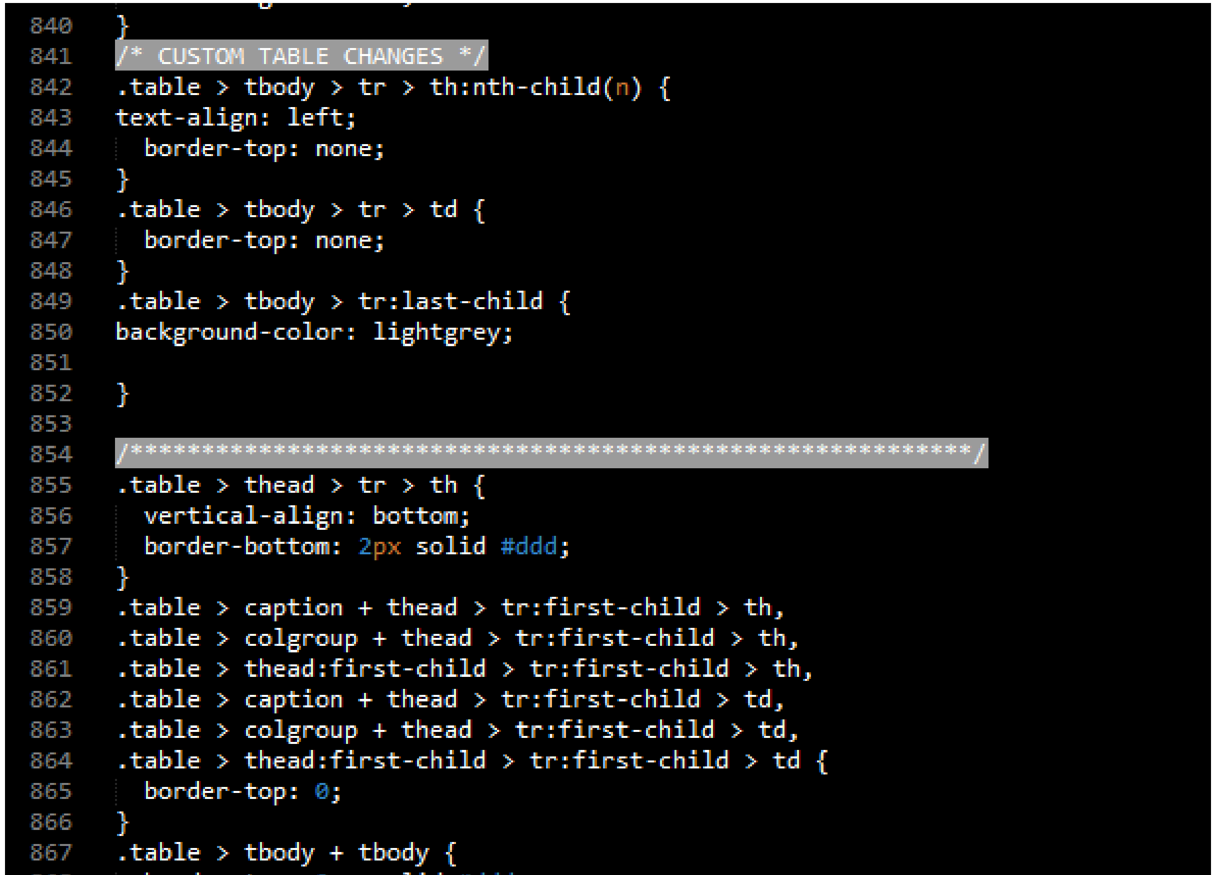
```python

# Render our file
html_out = template.render(template_vars)

#Change class=dataframe so that class=table for all tables and remove border formatting
soup = BeautifulSoup(html_out, "lxml")
for table in soup.find_all('table'):
    table['class'] = 'table'
    del table['border']

#write cleaned up html out to html file
html_outClean = soup.prettify("utf-8")
with open(outHTML,"wb") as file:
        file.write(html_outClean)

#need to specify location of wkhtmltopdf executible because of pdfkit dependency
config = pdfkit.configuration(wkhtmltopdf='C:/Program Files/wkhtmltopdf/bin/wkhtmltopdf.exe')
#Print html to pdf using local bootstrap CSS file for formatting
#css = "C:/Users/cburch/Downloads/bootstrap-3.3.7-dist/bootstrap-3.3.7-dist/css/bootstrap.css"
css = 'D:/GIS/Users/Cain_Alan/Collector/WeeklyUpdate_files/bootstrapLean.css'
#add footer to right of page that prints page x of x
options ={'footer-right': 'Page [page] of [topage]'}
#format is .from_file(input html, output file, css file, config file)
pdfkit.from_file(outHTML, outfile, css=css, configuration=config, options=options)

```

Html template – set page breaks and placeholders for variables



```html
<!DOCTYPE html>
<html>

<head>
    <meta http-equiv="X-UA-Compatible" content="IE=Edge" />
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Chronic Wasting Disease Surveillance Weekly</title>

    <meta name="description" content=" ">
    <meta name="keywords" content="chronic wasting disease texas, cwd, mule deer, white tailed deer, about cwd, check station, cwd results">
    <meta name="developer" content="Cristy Burch" />
    <meta name="sunset" content="2016/10/6" />
    <!-- css stylesheet specific for printed output -->
    <link crossorigin="anonymous" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/css/bootstrap.min.css" integrity="sha384-BVYiiSIFeK1dGmJRAkycuHAHRg32OmUcww7on3RYdg4Va+PmSTsz/K68vbdEjh4u" rel="stylesheet"/>
    <!-- Optional theme -->
    <link crossorigin="anonymous" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/css/bootstrap-theme.min.css" integrity="sha384-rHyoN1iRsVXV4nD0JutlnGaslCJuC7uwjduW9SVrLvRYooPp2bWYgmgJQIXwl/Sp" rel="
    stylesheet"/>

</head>

<body>

    <div class="container" id="mainContent">
        <h1 class="text-center">
    Chronic Wasting Disease Surveillance Weekly Update<br>
        <small>{{date}}</small></h1>

<br><br>
        <p>
            CWD surveillance efforts have been under way since {{sampleStart}}. Statewide CWD sampling goals for the {{FY}} collection year are to collect approximately {{sampleGoal}} samples, and all samples
            within the CWD designated zones. Wildlife Division staff are collecting CWD samples from a variety of locations which include, road kill deer, locker plants and deer processors, private ranches, WMA
            and State Parks, and check stations. The first sample reported for this season was collected on {{firstSampleDate}} and was a {{causeOfDeath}} deer. Exotic species which have been sampled include axis
            deer, fallow deer, red stag, sika, and elk. A total of {{totalSamples}} CWD samples have been collected to date which is approximately {{pctGoal}} of the statewide goal of {{sampleGoal}} samples.
            Summary of current results are listed below along with maps illustrating distribution of CWD samples.
        </p>


        <h3>SUMMARY</h3>
<br>
        <div id="table1">
            <p><strong>Table 1.</strong> CWD sample totals by species of deer and sex. </p>
            {{table1}}
        </div>
<br><br>
        <div id="table2">
            <p><strong>Table 2.</strong> CWD sample totals by species of deer and type of mortality. “Other Mortality” includes sick deer, poached deer, deer collected for training or research.</p>
            {{table2}}
        </div>
```

Style.css for a bootstrap style table



```
}
/* CUSTOM TABLE CHANGES */
.table > tbody > tr > th:nth-child(n) {
text-align: left;
  border-top: none;
}
.table > tbody > tr > td {
  border-top: none;
}
.table > tbody > tr:last-child {
background-color: lightgrey;

}

/******************************************************/
.table > thead > tr > th {
  vertical-align: bottom;
  border-bottom: 2px solid #ddd;
}
.table > caption + thead > tr:first-child > th,
.table > colgroup + thead > tr:first-child > th,
.table > thead:first-child > tr:first-child > th,
.table > caption + thead > tr:first-child > td,
.table > colgroup + thead > tr:first-child > td,
.table > thead:first-child > tr:first-child > td {
  border-top: 0;
}
.table > tbody + tbody {
  border-top: 2px solid #ddd;
}
.table .table {
  background-color: #fff;
}
.table-condensed > thead > tr > th,
.table-condensed > tbody > tr > th,
.table-condensed > tfoot > tr > th,
.table-condensed > thead > tr > td,
.table-condensed > tbody > tr > td,
.table-condensed > tfoot > tr > td {
  padding: 5px;
}
.table-bordered {
  border: 1px solid #ddd;
}
```

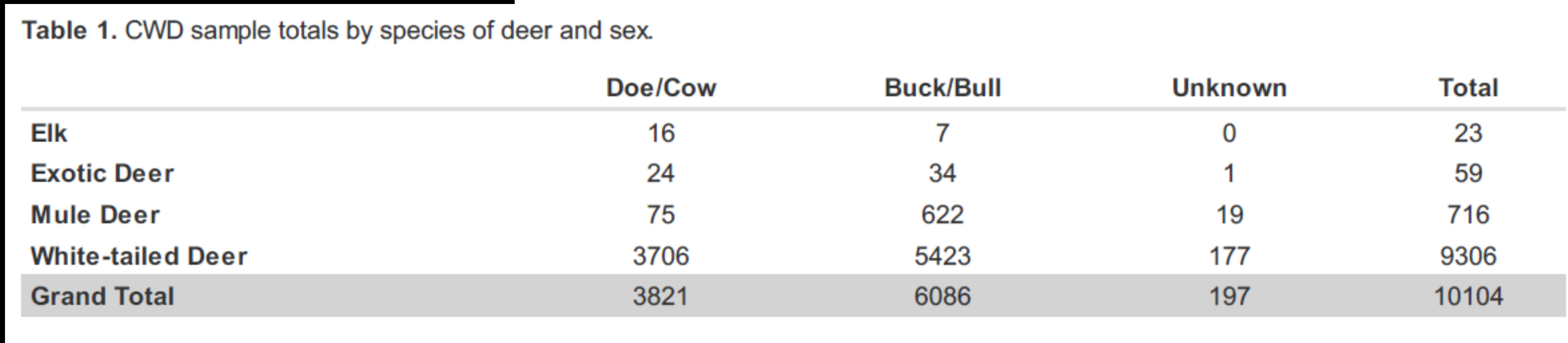
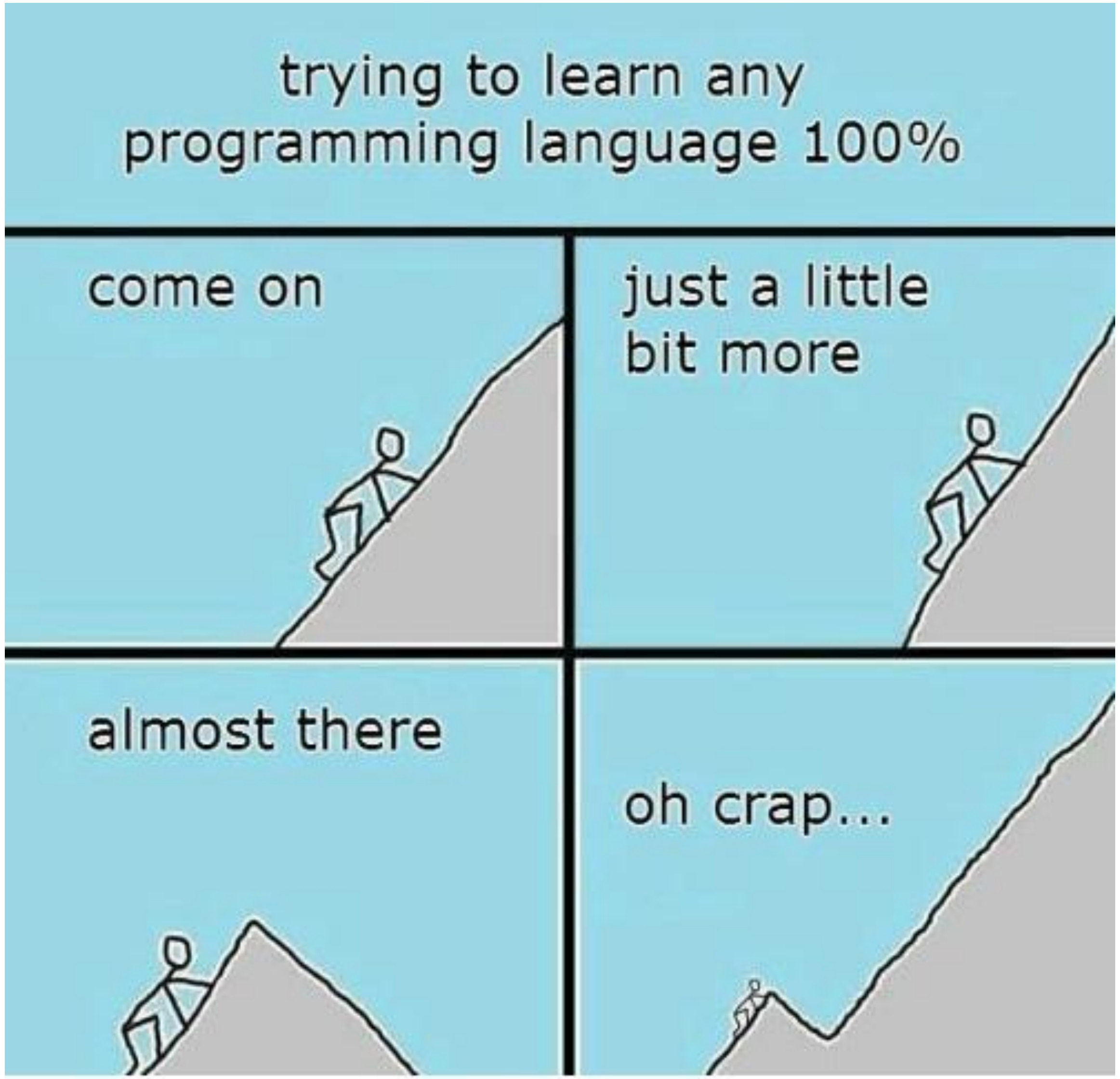


Table 1. CWD sample totals by species of deer and sex.

	Doe/Cow	Buck/Bull	Unknown	Total
Elk	16	7	0	23
Exotic Deer	24	34	1	59
Mule Deer	75	622	19	716
White-tailed Deer	3706	5423	177	9306
Grand Total	3821	6086	197	10104



Cartoon credit: http://canacopegdl.com/single.php?id=https://pics.me.me/trying-to-learn-any-programming-language-100-just-a-little-7917454.png

Cristy G. Burch

cristy.burch@tpwd.texas.gov

TEXAS
PARKS &
WILDLIFE

Life's better outside.®