Chapter 6

Tao Huang

6.1.1 R Make-like files

Example R make-like file

Sourcing functions is good practice because it is:

Reusable: It allows you to reuse functions over and over using the same code (i.e. you don't have to copy and paste the function into each new analysis script).

Easy to Maintain: It allows you to quickly fix a function that doesn't work properly - only once.

Sharable: In the same way that a library can be used by anyone, you can share your R script containing your functions with anyone, too. This is the first step towards creating an R package!

https://www.earthdatascience.org/courses/earth-analytics/multispectral-remote-sensing-data/source-function-in-R/

```
# Set working directory
setwd("/Users/owner/Desktop/EEB603/Ch6")
source("plot.R")
plot_density(1,2)
```

```
## [1] 0.5
```

6.2 Importing Locally Stored Data Sets

.csv

read.csv

```
tick<-read.csv("/Users/owner/Desktop/EEB603/Ch6/neon_CLBJ_Dermacentor.csv",sep=",",header = T)
head(tick)
```

```
##   siteID          sampleID       namedLocation.x     collectDate.y
## 1   BART BART_011.20140602 BART_011.tickPlot.tck 2014-06-02T18:32Z
## 2   BART BART_011.20150810 BART_011.tickPlot.tck 2015-08-10T21:07Z
## 3   BART BART_029.20180611 BART_029.tickPlot.tck 2018-06-11T18:03Z
## 4   BLAN BLAN_005.20160627 BLAN_005.tickPlot.tck 2016-06-27T14:01Z
## 5   BLAN BLAN_005.20180720 BLAN_005.tickPlot.tck 2018-07-20T13:49Z
## 6   BLAN BLAN_012.20160425 BLAN_012.tickPlot.tck 2016-04-25T16:55Z
##   individualCount totalSampledArea targetTaxaPresent adultCount nymphCount
## 1               1              160                 Y          1          0
## 2               1              165                 Y          0          1
```

```
## 3               1              140                 Y          1          0
## 4               1              157                 Y          1         18
## 5               1              166                 Y          1          0
## 6               1              167                 Y          1          1
##   larvaCount decimalLatitude decimalLongitude elevation       nlcdClass
## 1          0        44.04878        -71.29535     399.5     mixedForest
## 2          0        44.04878        -71.29535     399.5     mixedForest
## 3          0        44.05895        -71.28933     292.2 evergreenForest
## 4          0        39.08622        -77.97274     168.3 deciduousForest
## 5          0        39.08622        -77.97274     168.3 deciduousForest
## 6          0        39.08272        -77.95959     134.9 deciduousForest
##       density adultdensity
## 1 0.006250000  0.006250000
## 2 0.006060606  0.000000000
## 3 0.007142857  0.007142857
## 4 0.121019108  0.006369427
## 5 0.006024096  0.006024096
## 6 0.011976048  0.005988024
```

read.table

```
tick2<-read.table("/Users/owner/Desktop/EEB603/Ch6/neon_CLBJ_Dermacentor.csv",sep=",",header = T)
head(tick2)
```

```
##   siteID          sampleID       namedLocation.x     collectDate.y
## 1   BART BART_011.20140602 BART_011.tickPlot.tck 2014-06-02T18:32Z
## 2   BART BART_011.20150810 BART_011.tickPlot.tck 2015-08-10T21:07Z
## 3   BART BART_029.20180611 BART_029.tickPlot.tck 2018-06-11T18:03Z
## 4   BLAN BLAN_005.20160627 BLAN_005.tickPlot.tck 2016-06-27T14:01Z
## 5   BLAN BLAN_005.20180720 BLAN_005.tickPlot.tck 2018-07-20T13:49Z
## 6   BLAN BLAN_012.20160425 BLAN_012.tickPlot.tck 2016-04-25T16:55Z
##   individualCount totalSampledArea targetTaxaPresent adultCount nymphCount
## 1               1              160                 Y          1          0
## 2               1              165                 Y          0          1
## 3               1              140                 Y          1          0
## 4               1              157                 Y          1         18
## 5               1              166                 Y          1          0
## 6               1              167                 Y          1          1
##   larvaCount decimalLatitude decimalLongitude elevation       nlcdClass
## 1          0        44.04878        -71.29535     399.5     mixedForest
## 2          0        44.04878        -71.29535     399.5     mixedForest
## 3          0        44.05895        -71.28933     292.2 evergreenForest
## 4          0        39.08622        -77.97274     168.3 deciduousForest
## 5          0        39.08622        -77.97274     168.3 deciduousForest
## 6          0        39.08272        -77.95959     134.9 deciduousForest
##       density adultdensity
## 1 0.006250000  0.006250000
## 2 0.006060606  0.000000000
## 3 0.007142857  0.007142857
## 4 0.121019108  0.006369427
## 5 0.006024096  0.006024096
## 6 0.011976048  0.005988024
```

write.csv

.Rdata

load()

save()

6.3 Importing Data Sets from the Internet

6.3.1 Data from non-secure (http) URLs

```
tick3<-read.csv("https://raw.githubusercontent.com/th2018/EEB603_ch6/master/neon_HARV_Dermacentor.csv")
head(tick3)
```

```
##   siteID          sampleID       namedLocation.x     collectDate.y
## 1   HARV HARV_001.20170622 HARV_001.tickPlot.tck 2017-06-22T15:53Z
## 2   HARV HARV_001.20170710 HARV_001.tickPlot.tck 2017-07-10T16:30Z
## 3   HARV HARV_001.20180514 HARV_001.tickPlot.tck 2018-05-14T18:39Z
## 4   HARV HARV_001.20180606 HARV_001.tickPlot.tck 2018-06-06T17:48Z
## 5   HARV HARV_001.20180606 HARV_001.tickPlot.tck 2018-06-06T17:48Z
## 6   HARV HARV_002.20180516 HARV_002.tickPlot.tck 2018-05-16T19:26Z
##   individualCount totalSampledArea targetTaxaPresent adultCount nymphCount
## 1               1              160                 Y          1         12
## 2               1              160                 Y          1          9
## 3               1              165                 Y          1          0
## 4               1              165                 Y          3          1
## 5               2              165                 Y          3          1
## 6               1              160                 Y          3          0
##   larvaCount decimalLatitude decimalLongitude elevation       nlcdClass
## 1          0        42.42422        -72.26010     202.1 deciduousForest
## 2          0        42.42422        -72.26010     202.1 deciduousForest
## 3          0        42.42422        -72.26010     202.1 deciduousForest
## 4          1        42.42422        -72.26010     202.1 deciduousForest
## 5          1        42.42422        -72.26010     202.1 deciduousForest
## 6          0        42.47723        -72.25945     187.7 deciduousForest
##       density adultdensity
## 1 0.081250000  0.006250000
## 2 0.062500000  0.006250000
## 3 0.006060606  0.006060606
## 4 0.030303030  0.018181818
## 5 0.030303030  0.018181818
## 6 0.018750000  0.018750000
```

6.3.4 Data APIs & feeds

```
# install neonUtilities - can skip if already installed
#install.packages("neonUtilities")
# load neonUtilities
library(neonUtilities)
```

The identifier of the NEON data product: https://data.neonscience.org/data-products/explore

```
zipsByProduct(dpID="DP1.20093.001", #Chemical properties of surface water
              site="BIGC", #Upper Big Creek, CA
              startdate = "2019-01", enddate = "2019-04",
              check.size = FALSE #R would ask you to approve the file size
              #,package = "basic", avg = "all", savepath = NA,load = F
  )
```

```
## Downloading files totaling approximately 0.10178 MB
```

```
## Warning in dir.create(filepath): '/Users/owner/Desktop/EEB603/Ch6/
## filesToStack20093' already exists
```

```
## Downloading 3 files
##   |                                                                                                      |
## 3 files downloaded to /Users/owner/Desktop/EEB603/Ch6/filesToStack20093
```

6.4 Advanced Automatic Data Gathering: Web Scraping