data.frames II

Preliminaries

Load Necessary Packages

```
> library(FSA)    # for headtail(), filterD()
> library(dplyr)  # for select(), mutate(), filter()
```

Loading Data from External CSV File

```
> # Set your working directory to where your external data files (and scripts) are located.
> setwd("C:/aaaWork/Web/GitHub/RcourseNunavut2016/Handouts")
> dSC <- read.csv("SawyerCo_reduced.csv")
> str(dSC)
'data.frame':   42682 obs. of  11 variables:
 $ waterbody: Factor w/ 11 levels "BLACK DAN LAKE",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ year     : int  2012 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...
 $ mon      : Factor w/ 7 levels "Apr","Aug","Jul",..: 7 7 7 7 7 7 7 7 7 7 ...
 $ gear     : Factor w/ 7 levels "BACKPACK SHOCKER",..: 2 2 2 2 2 2 2 2 2 2 ...
 $ species  : Factor w/ 15 levels "Black Crappie",..: 13 13 13 13 13 13 13 13 13 13 ...
 $ len      : int  191 196 198 211 218 251 277 312 208 208 ...
 $ weight   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ sex      : Factor w/ 4 levels "","F","M","U": 1 1 1 1 1 1 1 1 1 1 ...
 $ age      : int  NA NA NA NA NA NA NA NA NA NA ...
 $ age_strux: Factor w/ 4 levels "","OTOLITH","SCALE",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ lennote  : Factor w/ 2 levels "Expanded length",..: 2 2 2 2 2 2 2 2 2 2 ...
```

```
> headtail(dSC,n=6)
           waterbody year mon         gear species len weight sex age age_strux         lennote
1     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 191     NA      NA           Observed length
2     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 196     NA      NA           Observed length
3     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 198     NA      NA           Observed length
4     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 211     NA      NA           Observed length
5     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 218     NA      NA           Observed length
6     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 251     NA      NA           Observed length
42677      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA           Observed length
42678      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA           Observed length
42679      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA           Observed length
42680      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA           Observed length
42681      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA           Observed length
42682      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA           Observed length
```

Select or Remove Variables from a data.frame

```
> dSC <- select(dSC,-age_strux,-lennote)
> headtail(dSC)
           waterbody year mon         gear species len weight sex age
1     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 191     NA      NA
2     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 196     NA      NA
3     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 198     NA      NA
42680      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA
42681      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA
42682      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA
```

Add Variables to a data.frame

```
> dSC <- mutate(dSC,len.in=len/25.4)
> headtail(dSC)
           waterbody year mon         gear species len weight sex age    len.in
1     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 191     NA      NA  7.519685
2     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 196     NA      NA  7.716535
3     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 198     NA      NA  7.795276
42680      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748
42681      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748
42682      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748
```

```
> dSC <- mutate(dSC,lcat10=lencat(len,w=10))
> headtail(dSC)
           waterbody year mon         gear species len weight sex age    len.in lcat10
1     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 191     NA      NA  7.519685    190
2     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 196     NA      NA  7.716535    190
3     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 198     NA      NA  7.795276    190
42680      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748    350
42681      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748    350
42682      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748    350
```

```
> dSC <- mutate(dSC,newspec=mapvalues(species,from=c("Walleye","Muskellunge"),to=c("WAE","MUE")))
> headtail(dSC)
           waterbody year mon         gear species len weight sex age    len.in lcat10 newspec
1     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 191     NA      NA  7.519685    190     WAE
2     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 196     NA      NA  7.716535    190     WAE
3     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 198     NA      NA  7.795276    190     WAE
42680      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748    350     WAE
42681      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748    350     WAE
42682      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA 14.015748    350     WAE
```

```
> dSC <- select(dSC,-(len.in:newspec))
> headtail(dSC,n=2)
           waterbody year mon         gear species len weight sex age
1     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 191     NA      NA
2     BLACK DAN LAKE 2012 Sep BOOM SHOCKER Walleye 196     NA      NA
42681      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA
42682      SAND LAKE 2013 Oct BOOM SHOCKER Walleye 356     NA      NA
```

Filter or Subset a data.frame

Conditionally Select Individuals

```
> levels(dSC$waterbody)
 [1] "BLACK DAN LAKE"  "CHIPPEWA RIVER"  "CONNORS LAKE"    "GRINDSTONE LAKE" "HUNTER LAKE"
 [6] "LAKE CHETAC"     "LAKE CHIPPEWA"   "MOSQUITO BROOK"  "NAMEKAGON RIVER" "NELSON LAKE"
[11] "SAND LAKE"
```

```
> dSC_chip <- filter(dSC,waterbody=="CHIPPEWA RIVER")
> levels(dSC_chip$waterbody)
 [1] "BLACK DAN LAKE"  "CHIPPEWA RIVER"  "CONNORS LAKE"    "GRINDSTONE LAKE" "HUNTER LAKE"
 [6] "LAKE CHETAC"     "LAKE CHIPPEWA"   "MOSQUITO BROOK"  "NAMEKAGON RIVER" "NELSON LAKE"
[11] "SAND LAKE"
```

```
> dSC_chip <- filterD(dSC,waterbody=="CHIPPEWA RIVER")
> levels(dSC_chip$waterbody)
[1] "CHIPPEWA RIVER"
```

```
> dSC_rivers <- filterD(dSC,waterbody %in% c("CHIPPEWA RIVER","NAMEKAGON RIVER"))
> levels(dSC_rivers$waterbody)
[1] "CHIPPEWA RIVER"  "NAMEKAGON RIVER"
```

```
> levels(dSC_rivers$gear)
[1] "BOTTOM GILL NET"   "MINI BOOM SHOCKER" "STREAM SHOCKER"
```

```
> dSC_rivers <- filterD(dSC_rivers,gear!="BOTTOM GILL NET")
> levels(dSC_rivers$gear)
[1] "MINI BOOM SHOCKER" "STREAM SHOCKER"
```

```
> levels(dSC_rivers$species)
 [1] "Black Crappie"      "Bluegill"           "Brook Trout"        "Brown Trout"
 [5] "Lake Sturgeon"      "Largemouth Bass"    "Muskellunge"        "Northern Pike"
 [9] "Rock Bass"          "Shorthead Redhorse" "Smallmouth Bass"    "Walleye"
[13] "Yellow Perch"
```

```
> dSC_rivers_trout <- filterD(dSC_rivers,species %in% c("Brook Trout","Brown Trot"))
> levels(dSC_rivers_trout$species)
[1] "Brook Trout"
```

```
> dSC_rivers_trout <- filterD(dSC_rivers,species %in% c("Brook Trout","Brown Trout"))
> levels(dSC_rivers_trout$species)
[1] "Brook Trout" "Brown Trout"
```

```
> tmp <- filterD(dSC_rivers_trout,len>=100 & len<200)
> headtail(tmp)
          waterbody year mon           gear     species len weight sex age
1   NAMEKAGON RIVER 2010 Sep STREAM SHOCKER Brook Trout 157     NA      NA
2   NAMEKAGON RIVER 2010 Aug STREAM SHOCKER Brown Trout 109     NA      NA
3   NAMEKAGON RIVER 2010 Aug STREAM SHOCKER Brown Trout 150     NA      NA
636 NAMEKAGON RIVER 2014 Jul STREAM SHOCKER Brown Trout 178     NA      NA
637 NAMEKAGON RIVER 2014 Jul STREAM SHOCKER Brown Trout 160     NA      NA
638 NAMEKAGON RIVER 2014 Jul STREAM SHOCKER Brown Trout 191     NA      NA
```

```
> tmp <- filterD(dSC_rivers_trout,len<100 | len>200)
> headtail(tmp)
           waterbody year mon              gear     species len weight sex age
1    NAMEKAGON RIVER 2010 Aug MINI BOOM SHOCKER Brown Trout 279    250      NA
2    NAMEKAGON RIVER 2010 Aug MINI BOOM SHOCKER Brown Trout 533   1800      NA
3    NAMEKAGON RIVER 2010 Aug    STREAM SHOCKER Brown Trout  71     NA      NA
1478 NAMEKAGON RIVER 2014 Jul    STREAM SHOCKER Brown Trout  89     NA      NA
1479 NAMEKAGON RIVER 2014 Jul    STREAM SHOCKER Brown Trout  89     NA      NA
1480 NAMEKAGON RIVER 2014 Jul    STREAM SHOCKER Brown Trout  89     NA      NA
```

```
> ( bigBrowns <- filterD(dSC_rivers_trout,species=="Brown Trout",mon %in% c("Jul","Aug"),len>=500) )
        waterbody year mon              gear     species len weight sex age
1 NAMEKAGON RIVER 2010 Aug MINI BOOM SHOCKER Brown Trout 533   1800      NA
2 NAMEKAGON RIVER 2010 Aug    STREAM SHOCKER Brown Trout 508     NA      NA
3 NAMEKAGON RIVER 2011 Aug    STREAM SHOCKER Brown Trout 521     NA      NA
4 NAMEKAGON RIVER 2011 Aug    STREAM SHOCKER Brown Trout 536     NA      NA
5 NAMEKAGON RIVER 2012 Jul    STREAM SHOCKER Brown Trout 511     NA      NA
6 NAMEKAGON RIVER 2012 Jul    STREAM SHOCKER Brown Trout 516     NA      NA
```

```
> ( bigBrowns_nowt <- filterD(bigBrowns,is.na(weight)) )
        waterbody year mon           gear     species len weight sex age
1 NAMEKAGON RIVER 2010 Aug STREAM SHOCKER Brown Trout 508     NA      NA
2 NAMEKAGON RIVER 2011 Aug STREAM SHOCKER Brown Trout 521     NA      NA
3 NAMEKAGON RIVER 2011 Aug STREAM SHOCKER Brown Trout 536     NA      NA
4 NAMEKAGON RIVER 2012 Jul STREAM SHOCKER Brown Trout 511     NA      NA
5 NAMEKAGON RIVER 2012 Jul STREAM SHOCKER Brown Trout 516     NA      NA
```

```
> ( bigBrowns_wghd <- filterD(bigBrowns,!is.na(weight)) )
        waterbody year mon              gear     species len weight sex age
1 NAMEKAGON RIVER 2010 Aug MINI BOOM SHOCKER Brown Trout 533   1800      NA
```

Remove individuals

```
> which(bigBrowns$len==536)
[1] 4
```

```
> bigBrowns[4,]
        waterbody year mon           gear     species len weight sex age
4 NAMEKAGON RIVER 2011 Aug STREAM SHOCKER Brown Trout 536     NA      NA
```

```
> ( bigBrowns2 <- bigBrowns[-4,] )
        waterbody year mon              gear     species len weight sex age
1 NAMEKAGON RIVER 2010 Aug MINI BOOM SHOCKER Brown Trout 533   1800      NA
2 NAMEKAGON RIVER 2010 Aug    STREAM SHOCKER Brown Trout 508     NA      NA
3 NAMEKAGON RIVER 2011 Aug    STREAM SHOCKER Brown Trout 521     NA      NA
5 NAMEKAGON RIVER 2012 Jul    STREAM SHOCKER Brown Trout 511     NA      NA
6 NAMEKAGON RIVER 2012 Jul    STREAM SHOCKER Brown Trout 516     NA      NA
```