

Add and Rename Variables

Derek H. Ogle, Northland College

2-Mar-2015

Preliminaries

```
> library(fishWiDNR) # for setDBClasses(), changeDBNames()
> library(dplyr)      # for filter(), select(), mutate(), rename()
> library(lubridate)  # for month()
> library(FSA)        # for capFirst(), expandCounts()

> setwd("C:/aaaWork/Web/fishR/Courses/WiDNR_Statewide_2015/Day1_IntroR_FMDData")
> d <- read.csv("FMDB_Sawyer_MultiYr_APEX.csv", stringsAsFactors=FALSE, na.strings=c("-", "NA", ""))
> d <- setDBClasses(d, type="RDNR")
> d <- expandCounts(d, ~Number.of.Fish, ~Length.or.Lower.Length.IN+Length.Upper.IN, new.name="Len")
> d1 <- filter(d, Species=="LAKE STURGEON", Waterbody.Name=="BARKER LAKE", !is.na(Weight.Pounds))
> d1 <- select(d1, Species, Survey.Year, Survey.Begin.Date, Len, Weight.Pounds)
> head(d1, n=4)
```

	Species	Survey.Year	Survey.Begin.Date	Len	Weight.Pounds
1	LAKE STURGEON	2012	2012-08-02	59.6	48.00
2	LAKE STURGEON	2013	2013-08-06	58.4	48.70
3	LAKE STURGEON	2013	2013-08-06	60.1	53.79
4	LAKE STURGEON	2013	2013-08-06	60.5	51.59

Variable Additions

Simple Mutations

```
> tmp <- mutate(d1, loglen=log(Len), logwt=log(Weight.Pounds))
> head(tmp, n=4)
```

	Species	Survey.Year	Survey.Begin.Date	Len	Weight.Pounds	loglen	logwt
1	LAKE STURGEON	2012	2012-08-02	59.6	48.00	4.087656	3.871201
2	LAKE STURGEON	2013	2013-08-06	58.4	48.70	4.067316	3.885679
3	LAKE STURGEON	2013	2013-08-06	60.1	53.79	4.096010	3.985088
4	LAKE STURGEON	2013	2013-08-06	60.5	51.59	4.102643	3.943328

Simple Special Purpose Mutations

```
> tmp <- mutate(d1, mon1=month(Survey.Begin.Date),
                 mon2=month(Survey.Begin.Date, label=TRUE))
> head(tmp, n=4)
```

	Species	Survey.Year	Survey.Begin.Date	Len	Weight.Pounds	mon1	mon2
1	LAKE STURGEON	2012	2012-08-02	59.6	48.00	8	Aug
2	LAKE STURGEON	2013	2013-08-06	58.4	48.70	8	Aug
3	LAKE STURGEON	2013	2013-08-06	60.1	53.79	8	Aug
4	LAKE STURGEON	2013	2013-08-06	60.5	51.59	8	Aug

```
> tmp <- mutate(d1,Species1=capFirst(Species),
                Species2=capFirst(Species,which="first"))
> head(tmp,n=4)
```

	Species	Survey.Year	Survey.Begin.Date	Len	Weight.Pounds	Species1	Species2
1	LAKE STURGEON	2012	2012-08-02	59.6	48.00	Lake Sturgeon	Lake sturgeon
2	LAKE STURGEON	2013	2013-08-06	58.4	48.70	Lake Sturgeon	Lake sturgeon
3	LAKE STURGEON	2013	2013-08-06	60.1	53.79	Lake Sturgeon	Lake sturgeon
4	LAKE STURGEON	2013	2013-08-06	60.5	51.59	Lake Sturgeon	Lake sturgeon

Length Category Mutations

```
> tmp <- mutate(d1,lcat2=lencat(Len,w=2),
                lcat2a=lencat(Len,w=2,as.fact=TRUE),
                lcatA=lencat(Len,breaks=c(46,54,56,58,70)),
                lcatB=lencat(Len,breaks=c(small=0,medium=50,large=60,very_large=70),use.names=TRUE) )
> head(tmp)
```

	Species	Survey.Year	Survey.Begin.Date	Len	Weight.Pounds	lcat2	lcat2a	lcatA	lcatB
1	LAKE STURGEON	2012	2012-08-02	59.6	48.00	58	58	58	medium
2	LAKE STURGEON	2013	2013-08-06	58.4	48.70	58	58	58	medium
3	LAKE STURGEON	2013	2013-08-06	60.1	53.79	60	60	58	large
4	LAKE STURGEON	2013	2013-08-06	60.5	51.59	60	60	58	large
5	LAKE STURGEON	2013	2013-08-06	62.2	56.88	62	62	58	large
6	LAKE STURGEON	2013	2013-08-06	66.3	67.24	66	66	58	large

```
> xtabs(~lcat2,data=tmp)
```

```
lcat2
46 54 56 58 60 62 66
 1  6  1  6  7  2  2
```

```
> xtabs(~lcat2a,data=tmp)
```

```
lcat2a
46 48 50 52 54 56 58 60 62 64 66
 1  0  0  0  6  1  6  7  2  0  2
```

```
> xtabs(~lcatA,data=tmp)
```

```
lcatA
46 54 56 58
 1  6  1 17
```

```
> xtabs(~lcatB,data=tmp)
```

```
lcatB
      small      medium      large very_large
      1         13         11          0
```

Rename Variables

```
> tmp <- rename(d1,year=Survey.Year,wt=Weight.Pounds)
> head(tmp)
```

	Species	year	Survey.Begin.Date	Len	wt
1	LAKE STURGEON	2012	2012-08-02	59.6	48.00
2	LAKE STURGEON	2013	2013-08-06	58.4	48.70
3	LAKE STURGEON	2013	2013-08-06	60.1	53.79
4	LAKE STURGEON	2013	2013-08-06	60.5	51.59
5	LAKE STURGEON	2013	2013-08-06	62.2	56.88
6	LAKE STURGEON	2013	2013-08-06	66.3	67.24

```
> tmp <- changeDBNames(d1)
> names(tmp)
```

```
[1] "srvy_begin" "species"    "year"      "Len"       "wt_lbs"
```

```
> tmp <- changeDBNames(tmp,from="R",to="RDNR")
> names(tmp)
```

```
[1] "Species"          "Survey.Year"      "Survey.Begin.Date" "Len"
[5] "Weight.Pounds"
```

```
> tmp <- changeDBNames(tmp,from="RDNR",to="DNR")
> write.csv(tmp,"LKS_Barker.csv",row.names=FALSE)
```

Application Assignment

Create a script that performs the following tasks:

1. Load and prepare (set classes, expand counts, examine structure) your FM data in R (**HINT:** *use all or some of your scripts from previous application assignments*).
2. Rename two or more variables that have names that annoy you (or change all names according to the definitions in `changeDBNames()`).
3. Create a new variable that has the species names with only the first letters capitalized.
4. Create a new variable that has the water body names with only the first letters capitalized.
5. Create a new variable that is the length in mm computed from the length in inches (even though this is already in the FM database).
6. Isolate a game species from a waterbody (and possibly a gear) of interest to you.
7. Create a new variable that contains evenly-spaced length categories that are appropriate for your species. Construct a frequency table of that variable.
8. Create a new variable that contains length categories that could be defined as “not of interest”, “marginally interesting”, “preferred”, and “very interesting” to anglers for your species. Construct a frequency table of that variable.
9. (*Time Permitting*) Create a new variable that contains the Gabelhouse length categories (“stock”, “quality”, etc.) for your species (**HINT:** *use, for example, `psdVal("Largemouth Bass",units="in")` to find Gablehouse lengths for a particular species*).

Save your script!