

# Load Data from CSV

*Derek H. Ogle, Northland College*

*2-Mar-2015*

## Preliminaries

```
> library(fishWiDNR) # for setDBClasses()
```

## Loading Data and Initial Examination

```
> setwd("C:/aaaWork/Web/fishR/Courses/WiDNR_Statewide_2015/Day1_IntroR_FMDData")
> d <- read.csv("FMDB_Sawyer.csv",stringsAsFactors=FALSE)
> d <- setDBClasses(d,type="RDNR")
> str(d)
'data.frame': 21358 obs. of 54 variables:
 $ County          : Factor w/ 1 level "SAWYER": 1 1 1 1 1 1 1 1 1 1 ...
 $ Waterbody.Name  : Factor w/ 55 levels "ASHEGON LAKE",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ WBIC            : int 2448800 2448800 2448800 2448800 2448800 2448800 2448800 2448800 2448800 2448800 ...
 $ Survey.Year     : int 2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
 $ Station.Name    : chr "ASHEGON LAKE_GENERAL LAKE STATION" "ASHEGON LAKE_GENERAL LAKE STATION" "ASHEGON LAKE_GENERAL LAKE STATION" ...
 $ Swims.Station.Id : int 10005674 10005674 10005674 10005674 10005674 10005674 10005674 10005674 10005674 10005674 ...
 $ Site.Seq.No     : int 108967 108967 108967 108967 108967 108967 108967 108967 108967 108967 ...
 $ Srvy.Seq.No     : int 56064296 56064296 56064296 56064296 56064296 56064296 56064296 56064296 56064296 56064296 ...
 $ Survey.Begin.Date : POSIXct, format: "2010-04-02" "2010-04-02" "2010-04-02" ...
 $ Survey.End.Date  : POSIXct, format: "2010-04-04" "2010-04-04" "2010-04-04" ...
 $ Survey.Status    : Factor w/ 3 levels "DATA ENTRY COMPLETE",...: 2 2 2 2 2 2 2 2 2 2 ...
 $ Data.Entry.Name  : chr "warwir" "warwir" "warwir" "warwir" ...
 $ Entry.Date       : POSIXct, format: "2010-07-01" "2010-07-01" "2010-07-01" ...
 $ Visit.Fish.Seq.No : int 683069 683069 683069 683069 683069 683069 683069 683069 683069 683069 ...
 $ Visit.Type       : Factor w/ 2 levels "ELECTROFISHING",...: 2 2 2 2 2 2 2 2 2 2 ...
 $ Gear             : Factor w/ 6 levels "BACKPACK SHOCKER",...: 4 4 4 4 4 4 4 4 4 4 ...
 $ Sample.Date      : POSIXct, format: "2010-04-03" "2010-04-03" "2010-04-03" ...
 $ Substation.Name  : chr "" "" "" "" ...
 $ Target.Species    : Factor w/ 12 levels "ALL SPECIES",...: 6 6 6 6 6 6 6 6 6 6 ...
 $ Fish.Data.Seq.No  : int 8837020 8837021 8837022 8837023 8837030 8837039 8837040 8837041 8837042 8837043 ...
 $ Net.Number       : chr "1" "1" "1" "1" ...
 $ Species.Code      : Factor w/ 64 levels "A03","A03J","A05J",...: 61 61 61 61 61 61 61 61 61 61 ...
 $ Species          : Factor w/ 64 levels "AMERICAN BROOK LAMPREY (AMMOCOETE)",...: 58 58 58 58 58 58 58 58 58 58 ...
 $ Number.of.Fish    : int 1 1 1 1 1 1 1 2 1 1 ...
 $ Length.or.Lower.Length.IN : num 20 21 21.5 23 12 12.5 14.5 15 15.5 16 ...
 $ Length.Upper.IN   : num 20.4 21.4 21.9 23.4 12.4 12.9 14.9 15.4 15.9 16.4 ...
 $ Length.or.Lower.Length.MM : num 508 533 546 584 305 ...
 $ Length.Upper.MM   : num 518 544 556 594 315 ...
 $ Weight.Pounds     : num NA NA NA NA NA NA NA NA NA NA ...
 $ Weight.Grams      : num NA NA NA NA NA NA NA NA NA NA ...
 $ Gender            : Factor w/ 4 levels "", "F", "M", "U": 2 2 2 2 4 3 3 3 3 3 ...
 $ Disease           : Factor w/ 0 levels: NA NA NA NA NA NA NA NA NA NA ...
 $ Injury.Type       : Factor w/ 2 levels "", "DEAD": 1 1 1 1 1 1 1 1 1 1 ...
 $ Age..observed.annuli. : logi NA NA NA NA NA NA ...
 $ Edge.Counted.Desc : Factor w/ 0 levels: NA NA NA NA NA NA NA NA NA NA ...
 $ Age.Structure     : Factor w/ 3 levels "", "SCALE", "SPINE": 1 1 1 1 1 1 1 1 1 1 ...
 $ Mark.Given        : Factor w/ 8 levels "", "AN", "LP", "LV",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ Mark.Found        : Factor w/ 10 levels "", "AN", "BC", "LP",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ Second.Mark.Found  : Factor w/ 2 levels "", "PIT": 1 1 1 1 1 1 1 1 1 1 ...
 $ Tag.Number.Given  : chr "" "" "" "" ...
```

```

$ Second.Tag.Number.Given : chr NA NA NA NA ...
$ Tag.Number.Found : chr "" "" "" "" ...
$ Second.Tag.Number.Found : chr NA NA NA NA ...
$ YOY : Factor w/ 2 levels "", "Y": 1 1 1 1 1 1 1 1 1 ...
$ Entry.Date.1 : POSIXct, format: "2010-07-01" "2010-07-01" "2010-07-01" ...
$ Last.Update.Date : POSIXct, format: NA NA NA ...
$ Data.Ent.Name : chr "prattf" "prattf" "prattf" "prattf" ...
$ Last.Update.Name : chr "" "" "" "" ...
$ Invalid.Species : chr NA NA NA NA ...
$ Non.Standard.Bin : chr NA NA NA NA ...
$ Length.Unit.Error : chr NA NA NA NA ...
$ Length.Outside.Range : chr NA NA NA NA ...
$ Count.Outside.Range : chr NA NA NA NA ...
$ Status.Code : chr NA NA NA NA ...

```

```
> head(d) # also can use tail(d)
```

	County	Waterbody.Name	WBIC	Survey.Year		Station.Name	Swims.Station.Id
1	SAWYER	ASHEGON LAKE	2448800	2010	ASHEGON LAKE_GENERAL	LAKE STATION	10005674
2	SAWYER	ASHEGON LAKE	2448800	2010	ASHEGON LAKE_GENERAL	LAKE STATION	10005674
3	SAWYER	ASHEGON LAKE	2448800	2010	ASHEGON LAKE_GENERAL	LAKE STATION	10005674
4	SAWYER	ASHEGON LAKE	2448800	2010	ASHEGON LAKE_GENERAL	LAKE STATION	10005674
5	SAWYER	ASHEGON LAKE	2448800	2010	ASHEGON LAKE_GENERAL	LAKE STATION	10005674
6	SAWYER	ASHEGON LAKE	2448800	2010	ASHEGON LAKE_GENERAL	LAKE STATION	10005674

	Site.Seq.No	Srvy.Seq.No	Survey.Begin.Date	Survey.End.Date		Survey.Status
1	108967	56064296	2010-04-02	2010-04-04	DATA ENTRY COMPLETE AND PROOFED	
2	108967	56064296	2010-04-02	2010-04-04	DATA ENTRY COMPLETE AND PROOFED	
3	108967	56064296	2010-04-02	2010-04-04	DATA ENTRY COMPLETE AND PROOFED	
4	108967	56064296	2010-04-02	2010-04-04	DATA ENTRY COMPLETE AND PROOFED	
5	108967	56064296	2010-04-02	2010-04-04	DATA ENTRY COMPLETE AND PROOFED	
6	108967	56064296	2010-04-02	2010-04-04	DATA ENTRY COMPLETE AND PROOFED	

	Data.Entry.Name	Entry.Date	Visit.Fish.Seq.No	Visit.Type	Gear	Sample.Date	Substation.Name
1	warwir	2010-07-01	683069	NETTING FYKE NET		2010-04-03	
2	warwir	2010-07-01	683069	NETTING FYKE NET		2010-04-03	
3	warwir	2010-07-01	683069	NETTING FYKE NET		2010-04-03	
4	warwir	2010-07-01	683069	NETTING FYKE NET		2010-04-03	
5	warwir	2010-07-01	683069	NETTING FYKE NET		2010-04-03	
6	warwir	2010-07-01	683069	NETTING FYKE NET		2010-04-03	

	Target.Species	Fish.Data.Seq.No	Net.Number	Species.Code	Species	Number.of.Fish
1	GAMEFISH SPECIES	8837020	1	X22	WALLEYE	1
2	GAMEFISH SPECIES	8837021	1	X22	WALLEYE	1
3	GAMEFISH SPECIES	8837022	1	X22	WALLEYE	1
4	GAMEFISH SPECIES	8837023	1	X22	WALLEYE	1
5	GAMEFISH SPECIES	8837030	1	X22	WALLEYE	1
6	GAMEFISH SPECIES	8837039	1	X22	WALLEYE	1

	Length.or.Lower.Length.IN	Length.Upper.IN	Length.or.Lower.Length.MM	Length.Upper.MM	Weight.Pounds
1	20.0	20.4	508.0	518.2	NA
2	21.0	21.4	533.4	543.6	NA
3	21.5	21.9	546.1	556.3	NA
4	23.0	23.4	584.2	594.4	NA
5	12.0	12.4	304.8	315.0	NA
6	12.5	12.9	317.5	327.7	NA

	Weight.Grams	Gender	Disease	Injury.Type	Age..observed.annuli.	Edge.Counted.Desc	Age.Structure
1	NA	F	<NA>			NA	<NA>
2	NA	F	<NA>			NA	<NA>
3	NA	F	<NA>			NA	<NA>
4	NA	F	<NA>			NA	<NA>
5	NA	U	<NA>			NA	<NA>
6	NA	M	<NA>			NA	<NA>

	Mark.Given	Mark.Found	Second.Mark.Found	Tag.Number.Given	Second.Tag.Number.Given	Tag.Number.Found
1						<NA>

```

2                                     <NA>
3                                     <NA>
4                                     <NA>
5                                     <NA>
6                                     <NA>
  Second.Tag.Number.Found YOY Entry.Date.1 Last.Update.Date Data.Ent.Name Last.Update.Name
1                      <NA>    2010-07-01          <NA>      prattf
2                      <NA>    2010-07-01          <NA>      prattf
3                      <NA>    2010-07-01          <NA>      prattf
4                      <NA>    2010-07-01          <NA>      prattf
5                      <NA>    2010-07-01          <NA>      prattf
6                      <NA>    2010-07-01          <NA>      prattf
  Invalid.Species Non.Standard.Bin Length.Unit.Error Length.Outside.Range Count.Outside.Range
1             <NA>             <NA>             <NA>             <NA>             <NA>
2             <NA>             <NA>             <NA>             <NA>             <NA>
3             <NA>             <NA>             <NA>             <NA>             <NA>
4             <NA>             <NA>             <NA>             <NA>             <NA>
5             <NA>             <NA>             <NA>             <NA>             <NA>
6             <NA>             <NA>             <NA>             <NA>             <NA>
  Status.Code
1          <NA>
2          <NA>
3          <NA>
4          <NA>
5          <NA>
6          <NA>

> nrow(d)
[1] 21358

```

## Expanding Counts of Fish to Individual Measurements

Example portion of the Sawyer Co. FM database with “tallied” fish lengths.

Fish.Data.Seq.No	Species	Number.of.Fish	Length.or.Lower.Length.IN	Length.Upper.IN
8837020	WALLEYE	1	20.0	20.4
8837040	WALLEYE	1	14.5	14.9
8837041	WALLEYE	2	15.0	15.4
8837042	WALLEYE	1	15.5	15.9
8837562	WALLEYE	2	14.5	14.9
8837563	WALLEYE	1	15.5	15.9
8837564	WALLEYE	1	17.0	17.4

Example expansion but keeping the length bins.

Fish.Data.Seq.No	Species	Length.or.Lower.Length.IN	Length.Upper.IN
8837020	WALLEYE	20.0	20.4
8837040	WALLEYE	14.5	14.9
8837042	WALLEYE	15.5	15.9
8837563	WALLEYE	15.5	15.9
8837564	WALLEYE	17.0	17.4
8837041	WALLEYE	15.0	15.4
8837041	WALLEYE	15.0	15.4
8837562	WALLEYE	14.5	14.9
8837562	WALLEYE	14.5	14.9

Example expansion but adding random digit from within the length bin.

Fish.Data.Seq.No	Species	Length.or.Lower.Length.IN	Length.Upper.IN	Len	lennote
8837020	WALLEYE	20.0	20.4	20.3	Expanded length
8837040	WALLEYE	14.5	14.9	14.7	Expanded length
8837042	WALLEYE	15.5	15.9	15.5	Expanded length
8837563	WALLEYE	15.5	15.9	15.5	Expanded length
8837564	WALLEYE	17.0	17.4	17.4	Expanded length
8837041	WALLEYE	15.0	15.4	15.3	Expanded length
8837041	WALLEYE	15.0	15.4	15.2	Expanded length
8837562	WALLEYE	14.5	14.9	14.5	Expanded length
8837562	WALLEYE	14.5	14.9	14.7	Expanded length

```
> ## without random digits
```

```
> d1 <- expandCounts(d,~Number.of.Fish)
```

Some rows (13884, 20543) had zero counts in Number.of.Fish.

17430 rows had an individual measurement.

3926 rows with multiple measurements were expanded to 36156 rows of individual measurements.

```
> ## with random digits
```

```
> d1 <- expandCounts(d,~Number.of.Fish,~Length.or.Lower.Length.IN+Length.Upper.IN,new.name="Len")
```

Some rows (13884, 20543) had zero counts in Number.of.Fish.

17430 rows had an individual measurement.

3926 rows with multiple measurements were expanded to 36156 rows of individual measurements.

```
> nrow(d1)
```

```
[1] 53588
```

```
> ## sum of Number.of.Fish variable (note from above that 2 rows had zero fish)
```

```
> sum(d$Number.of.Fish)
```

```
[1] 53586
```

## Individual Variables

> d1\$Length.or.Lower.Length.IN

```
[1] 5.0 9.8 10.2 10.3 12.0 12.2 12.5 4.1 4.4 4.6 5.0 5.3 6.2 8.4 9.2 11.8 12.2 12.3 9.0
[20] 11.7 1.8 2.0 2.2 2.3 2.7 2.8 3.2 3.7 4.0 4.1 4.2 4.9 5.6 5.9 6.4 6.6 7.5 7.6
[39] 7.9 8.0 8.2 8.8 9.0 3.0 3.2 3.4 3.6 3.8 4.9 5.1 5.4 5.6 5.7 6.0 6.7 7.0 7.3
[58] 7.4 9.0 NA NA NA NA NA NA NA NA NA NA NA NA NA 3.2 3.4 3.5 3.6
[77] 3.9 4.0 4.2 6.4 6.5 6.7 5.3 5.7 6.0 6.2 6.3 6.5 6.7 6.8 7.0 8.1 8.3 8.6 8.7
[96] 9.1 3.0 3.6 4.1 4.5 5.6 4.0 4.5 5.5 6.2 6.7 8.5 5.1 6.0 6.2 2.0 8.3 8.6 8.7
[115] 8.8 9.0 9.1 9.2 9.5 9.7 9.8 9.8 9.9 10.0 10.1 10.2 10.3 10.3 10.3 10.6 10.9 11.2 11.9
[134] 12.6 12.7 13.0 13.0 13.0 13.3 13.5 14.0 14.0 14.0 14.0 14.5 14.5 14.5 15.0 15.0 15.0 15.5 15.5
[153] 16.0 16.0 18.5 15.4 15.0 17.3 58.0 61.5 59.7 62.5 55.7 56.4 55.6 55.0 60.9 58.4 46.8 66.3 66.3
[172] 55.7 55.1 62.2 54.1 60.5 60.1 21.0 9.0 10.2 11.2 11.7 11.8 12.1 12.4 8.8 9.5 9.8 6.0 5.8
[191] 6.1 7.3 10.2 10.4 9.5 10.1 6.0 6.3 6.8 7.4 6.6 7.0 8.2 6.5 7.0 7.2 7.3 6.3 6.5
[210] 5.9 6.4 8.4 6.5 9.2 7.8 6.6 7.0 6.5 6.5 6.8 4.9 7.7 7.0 6.5 6.5 6.9 6.6 4.9
[229] 7.7 7.2 9.0 15.5 7.1 7.0 6.4 6.9 6.2 4.9 6.1 6.9 15.1 10.5 6.4 12.0 4.9 9.6 7.1
[248] 7.2 6.6 7.1 7.2 4.6 5.4 6.7 11.1 6.8 6.8 7.0 9.0 8.4 4.1 6.8 6.2 8.7 8.0 7.0
[267] 6.4 29.0 9.1 23.2 20.4 17.4 21.9 20.0 20.1 19.7 17.2 18.1 20.3 16.6 15.5 10.1 16.5 13.1 13.2
[286] 8.3 6.7 6.7 6.5 6.5 6.9 4.8 6.0 6.7 6.5 6.9 4.4 4.8 6.3 6.7 7.0 6.8 8.9 6.4
```

> d1\$Species

```
[1] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[5] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[9] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[13] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[17] BLACK CRAPPIE          BLACK CRAPPIE          YELLOW PERCH           BLACK CRAPPIE
[21] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[25] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[29] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[33] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[37] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[41] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[45] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[49] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[53] BROOK TROUT            BROOK TROUT            BROOK TROUT            BROOK TROUT
[57] BROOK TROUT            BROOK TROUT            BROOK TROUT            WHITE SUCKER
[61] LARGESCALE STONEROLLER WHITE SUCKER           PUMPKINSEED           GOLDEN SHINER
[65] GOLDEN SHINER          GOLDEN SHINER          ROCK BASS              BLACK CRAPPIE
[69] WHITE SUCKER           WHITE SUCKER           BLACK CRAPPIE          LARGEMOUTH BASS
[73] BLUEGILL               BLUEGILL               BLUEGILL               BLUEGILL
[77] BLUEGILL               BLUEGILL               BLUEGILL               BLUEGILL
[81] BLUEGILL               BLUEGILL               BLACK CRAPPIE          BLACK CRAPPIE
[85] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[89] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[93] BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE          BLACK CRAPPIE
[97] YELLOW PERCH           YELLOW PERCH           YELLOW PERCH           YELLOW PERCH
[101] YELLOW PERCH           PUMPKINSEED           PUMPKINSEED           PUMPKINSEED
[105] PUMPKINSEED           PUMPKINSEED           YELLOW BULLHEAD        GOLDEN SHINER
[109] GOLDEN SHINER          GOLDEN SHINER          TADPOLE MADTOM         WALLEYE
[113] WALLEYE                WALLEYE                WALLEYE                WALLEYE
[117] WALLEYE                WALLEYE                WALLEYE                WALLEYE
```

64 Levels: AMERICAN BROOK LAMPREY (AMMOCOETE) BLACK BULLHEAD BLACK CRAPPIE ... YELLOW PERCH

# Application Assignment

Create a script that performs the following tasks:

1. Load your FM data into R.
2. Set the classes on your data.frame.
3. How many total fish in your data.frame?
4. Expand the counts to individual fish (assign to a new data.frame).
5. How many variables (columns) and individual fish (rows) in your new data.frame?
6. Display all expanded length measurements.
7. Show all “values” for two other variables of your choice.

**Save your script!**