

Preliminaries

```
> library(FSA)      # for Subset(), Summarize(), hist(), fact2num()
> library(plotrix) # for plotCI()

> setwd("C:/aaaWork/Web/fishR/courses/Vermont2014/CourseMaterial/") # Derek's Computer
> d <- read.csv("Data/MnFats.csv",header=TRUE)
> d <- Subset(d,sex!="UNK") # removed one unknown sex individual (for simplicity)
> str(d)

'data.frame': 304 obs. of  6 variables:
 $ unit: Factor w/ 1 level "MN-1": 1 1 1 1 1 1 1 1 1 1 ...
 $ year: int   2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 ...
 $ len : int   310 363 373 381 394 394 396 401 406 409 ...
 $ wt  : int   240 330 370 490 470 490 460 490 540 650 ...
 $ sex : Factor w/ 2 levels "F","M": 1 1 2 2 2 2 1 2 2 1 ...
 $ age : int    9 10 17 10 11 14 11 15 13 15 ...

> d <- within(d, {
  fyear <- factor(year)
  loglen <- log(len)
  logwt <- log(wt)
})
> view(d)

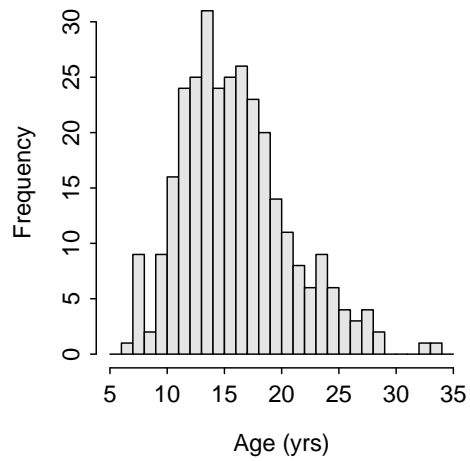
  unit year len  wt sex age logwt loglen fyear
65  MN-1 2000 498  920  F  18 6.824  6.211  2000
69  MN-1 2000 564 1720  M  16 7.450  6.335  2000
161 MN-1 2003 500 1260  M  22 7.139  6.215  2003
165 MN-1 2003 538 1170  F  16 7.065  6.288  2003
167 MN-1 2003 572 1650  M  16 7.409  6.349  2003
238 MN-1 2006 630 2130  M  21 7.664  6.446  2006
```

Simple Univariate Summaries

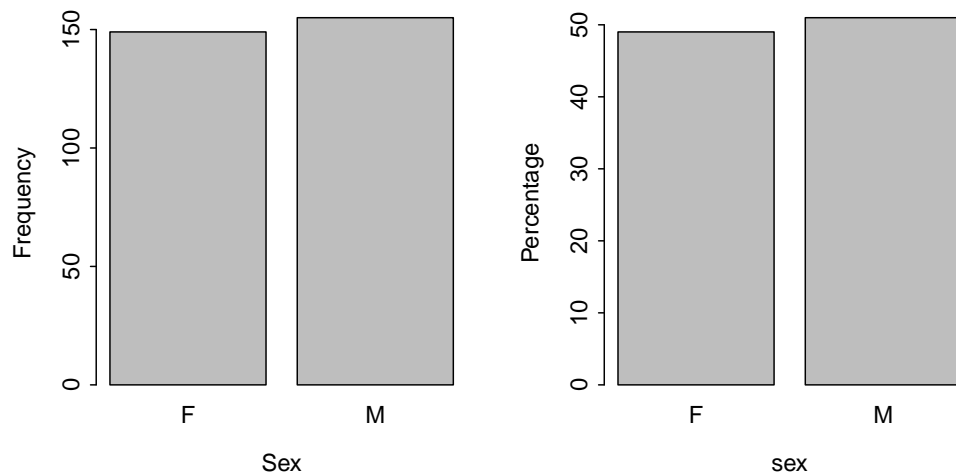
```
> Summarize(~age,data=d,digits=2)

      n      mean      sd      min      Q1      median      Q3      max percZero
304.00   15.51    4.77    6.00   12.00    15.00   18.00   33.00    0.00

> hist(~age,data=d,xlab="Age (yrs)",breaks=seq(5,35,1))
```

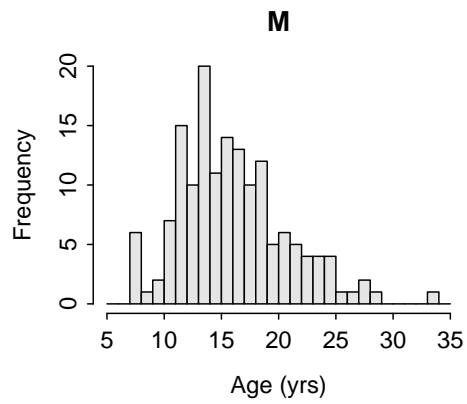
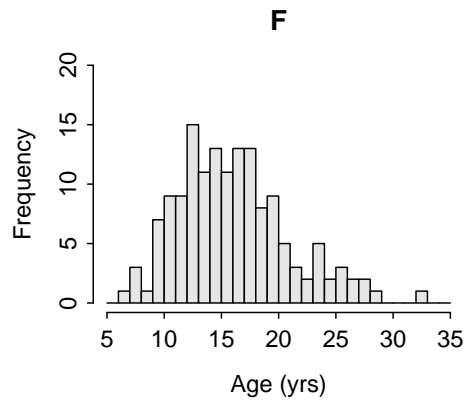


```
> ( sextbl <- xtabs(~sex,data=d) )
sex
  F  M
149 155
> prop.table(sextbl)*100
sex
  F    M
49.01 50.99
> barplot(sextbl,xlab="Sex",ylab="Frequency")           # Left
> barplot(prop.table(sextbl)*100,xlab="sex",ylab="Percentage") # Right
```

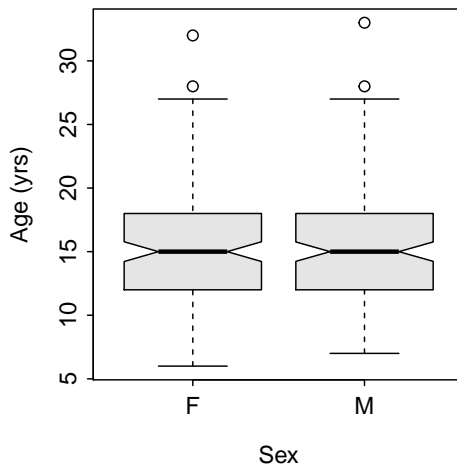


Simple Bivariate Summaries

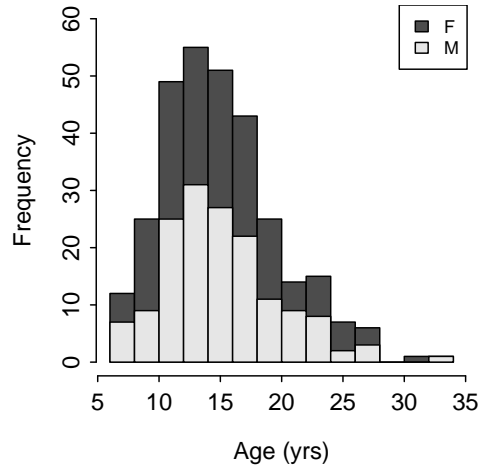
```
> Summarize(age~sex,data=d,digits=2)
  sex  n mean  sd min Q1 median Q3 max percZero
1  F 149 15.52 4.85  6 12    15 18  32         0
2  M 155 15.50 4.71  7 12    15 18  33         0
> hist(age~sex,data=d,xlab="Age (yrs)",breaks=seq(5,35,1),nrow=2,ncol=1)
```



```
> boxplot(age~sex,data=d,xlab="Sex",ylab="Age (yrs)",col="gray90",notch=TRUE)
```



```
> histStack(age~sex,data=d,xlab="Age (yrs)",breaks=seq(5,35,1),ylim=c(0,60),
  col="gray.colors",legend="topright")
```



```
> agesextbl <- xtabs(~sex+age,data=d)
> round(prop.table(agesextbl)*100,1)
```

	age																				
sex	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
F	0.3	1.0	0.3	2.3	3.0	3.0	4.9	3.6	4.3	3.6	4.3	4.3	2.6	3.0	1.6	1.0	0.7	1.6	0.7	1.0	
M	0.0	2.0	0.3	0.7	2.3	4.9	3.3	6.6	3.6	4.6	4.3	3.3	3.9	1.6	2.0	1.6	1.3	1.3	1.3	0.3	

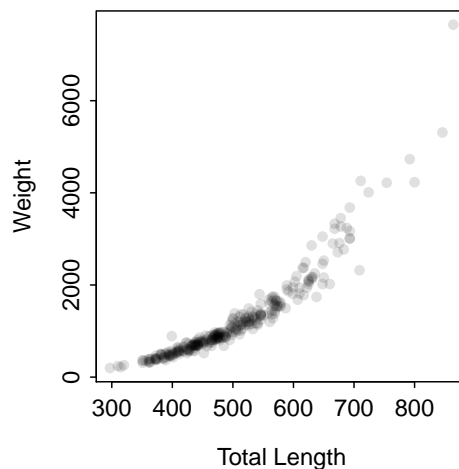
```
age
sex 26 27 28 32 33
F 0.7 0.7 0.3 0.3 0.0
M 0.3 0.7 0.3 0.0 0.3
```

```
> round(prop.table(agesextbl,margin=1)*100,1)
```

	age															
sex	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
F	0.7	2.0	0.7	4.7	6.0	6.0	10.1	7.4	8.7	7.4	8.7	8.7	5.4	6.0	3.4	2.0
M	0.0	3.9	0.6	1.3	4.5	9.7	6.5	12.9	7.1	9.0	8.4	6.5	7.7	3.2	3.9	3.2

	age								
sex	22	23	24	25	26	27	28	32	33
F	1.3	3.4	1.3	2.0	1.3	1.3	0.7	0.7	0.0
M	2.6	2.6	2.6	0.6	0.6	1.3	0.6	0.0	0.6

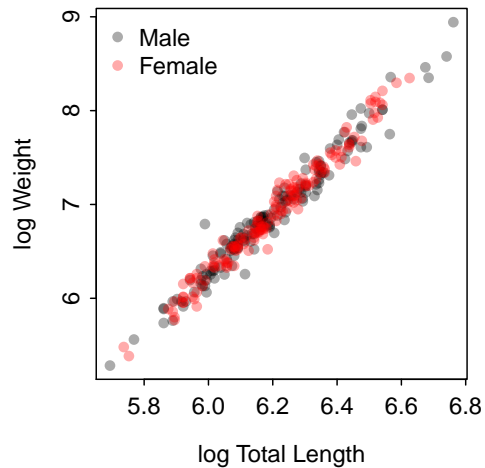
```
> plot(wt~len,data=d,xlab="Total Length",ylab="Weight",pch=16,col=rgb(0,0,0,1/8))
```



```

> colM <- rgb(0,0,0,1/3)
> colF <- rgb(1,0,0,1/3)
> plot(logwt~loglen,data=Subset(d,sex=="M"),pch=16,col=colM,
       xlab="log Total Length",ylab="log Weight")
> points(logwt~loglen,data=Subset(d,sex=="F"),pch=16,col=colF)
> legend("topleft",c("Male","Female"),pch=16,bty="n",col=c(colM,colF))

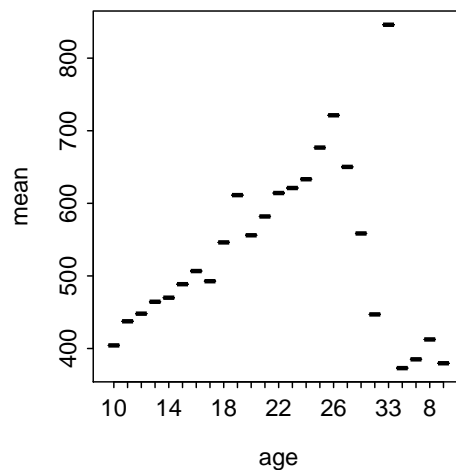
```



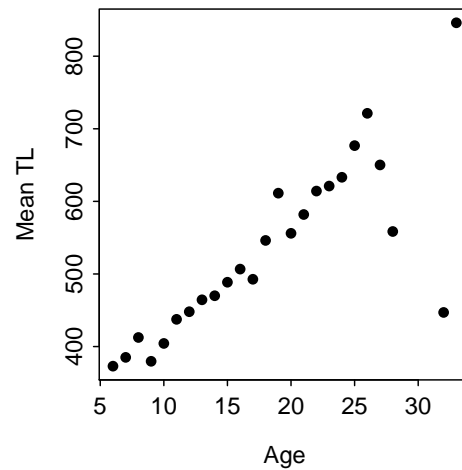
```

> lenAtAge <- Summarize(len~age,data=d,digits=1)
Warning: To continue, variable(s) on RHS of formula were converted to a factor.
> str(lenAtAge)
'data.frame': 25 obs. of 10 variables:
 $ age      : Factor w/ 25 levels "10","11","12",...: 22 23 24 25 1 2 3 4 5 6 ...
 $ n        : num  1 9 2 9 16 24 25 31 24 25 ...
 $ mean     : num  373 385 412 380 404 ...
 $ sd       : num  NA 49.8 55.9 61 48.7 46.8 54.8 56.4 34.6 57.5 ...
 $ min      : num  373 315 373 297 351 363 361 351 394 361 ...
 $ Q1       : num  373 361 393 356 375 406 419 436 462 450 ...
 $ median   : num  373 396 412 381 389 434 429 465 475 490 ...
 $ Q3       : num  373 406 432 386 421 457 495 485 488 526 ...
 $ max      : num  373 472 452 503 541 546 599 569 526 569 ...
 $ percZero: num  0 0 0 0 0 0 0 0 0 0 ...
> plot(mean~age,data=lenAtAge) # NO GOOD!!

```



```
> plot(mean~fact2num(age),data=lenAtAge,pch=16,xlab="Age",ylab="Mean TL") # GOOD!!
```



```
> lenAtAge <- within(lenAtAge, {
  LCI <- mean-1.96*sd/sqrt(n)
  UCI <- mean+1.96*sd/sqrt(n)
})
> head(lenAtAge)
```

	age	n	mean	sd	min	Q1	median	Q3	max	percZero	UCI	LCI
1	6	1	373.0	NA	373	373	373	373	373	0	NA	NA
2	7	9	385.1	49.8	315	361	396	406	472	0	417.6	352.6
3	8	2	412.5	55.9	373	393	412	432	452	0	490.0	335.0
4	9	9	379.7	61.0	297	356	381	386	503	0	419.6	339.8
5	10	16	404.3	48.7	351	375	389	421	541	0	428.2	380.4
6	11	24	437.5	46.8	363	406	434	457	546	0	456.2	418.8

```
> with(lenAtAge,plotCI(fact2num(age),mean,ui=UCI,li=LCI,
  pch=16,xlab="Age",ylab="Mean TL"))
```

