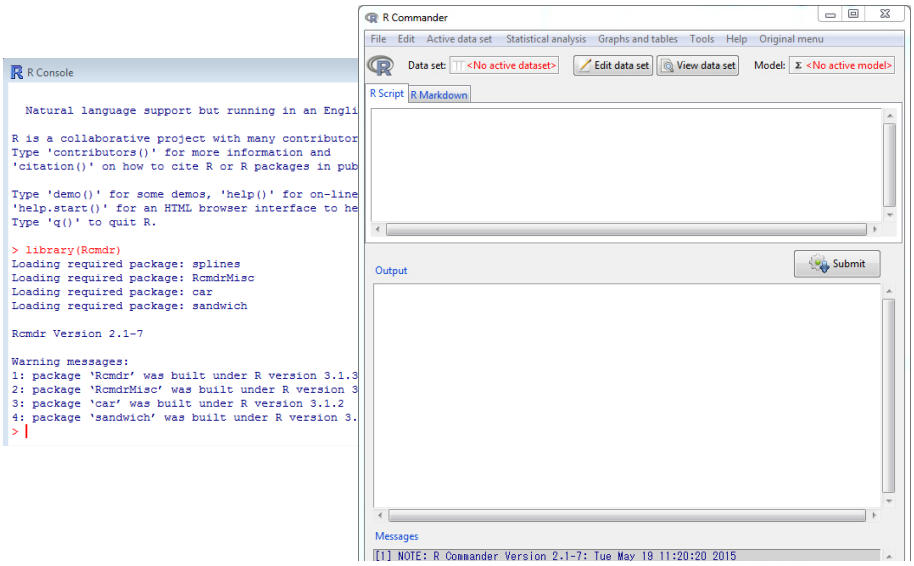
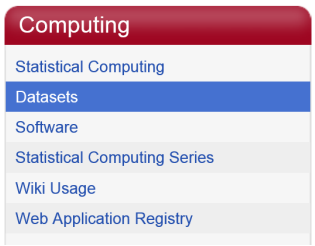
EZR as a R commander plugin



1

Dataset: Titanic

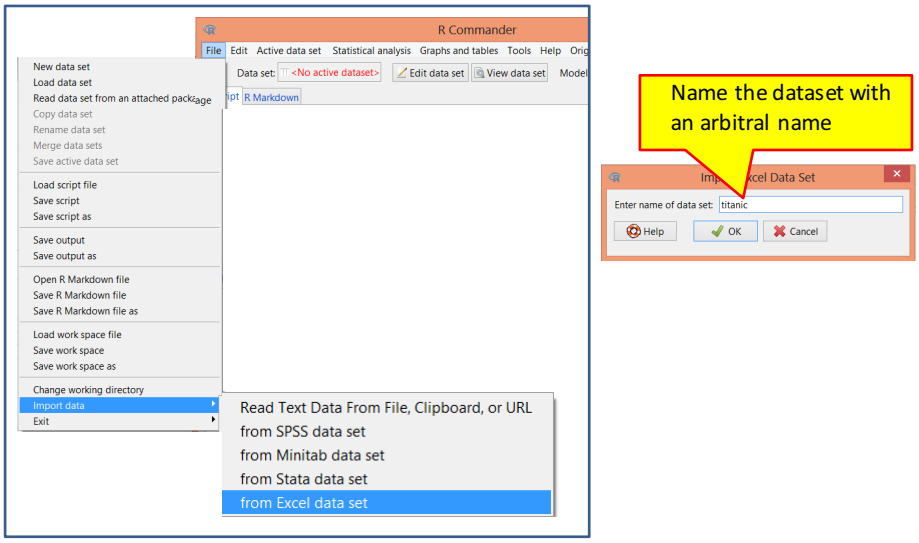
- Go to http://biostat.mc.vanderbilt.edu/wiki
- Click on Dataset



- Select titanic.xls

Data for Titanic passengers				
titanic.html	titanic.sav	titanic.sdd	NA	titanic.txt
NA	titanic2.sav	titanic2.sdd	NA	NA
NA	titanic3.sav	titanic3.sdd	titanic3.xls	titanic3.csv

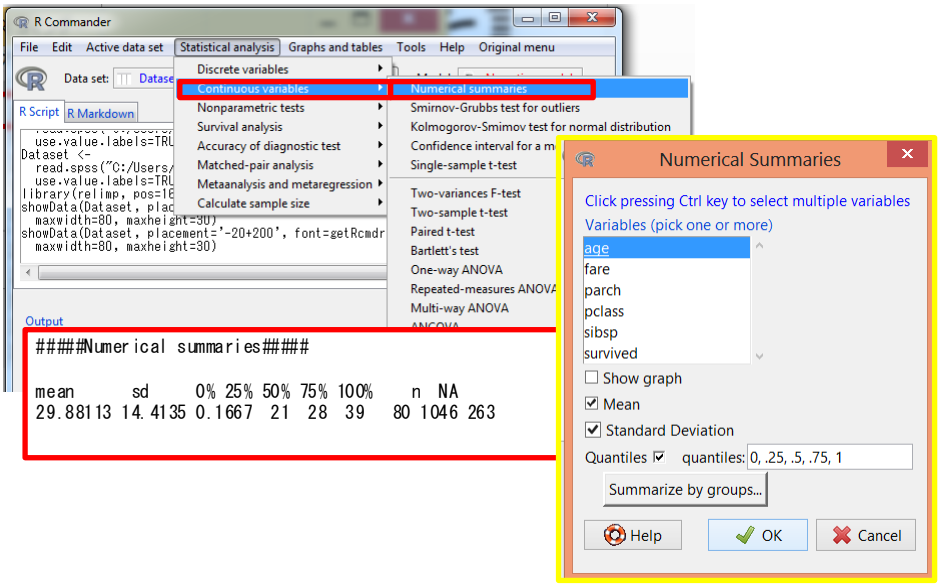
Reading data into Rcmdr-EZR



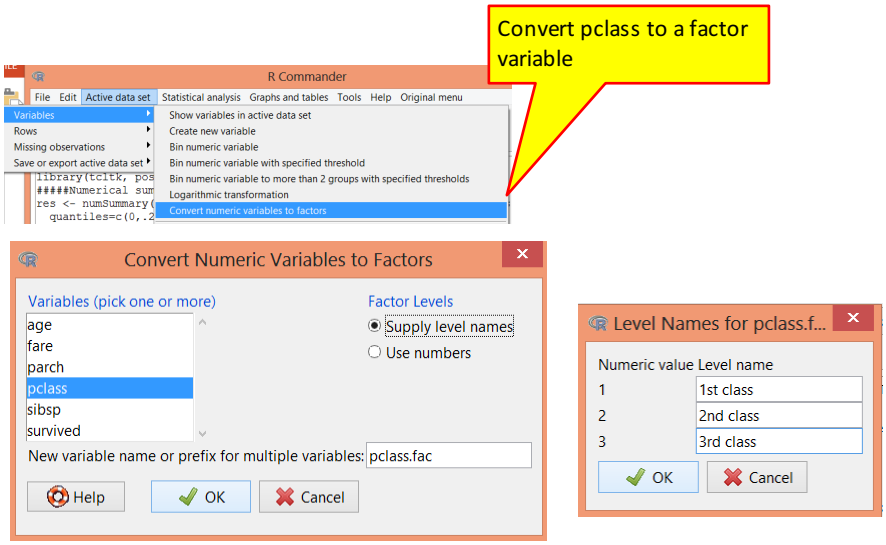
Passenger Characteristics

	Died (N=809)	Survived (N=500)
Age、Mean (SD)	30.5 (13.9)	28.9 (15.1)
Fare、Mean (SD)	23.3 (34.1)	49.4 (68.6)
Passenger Class, N（%）		
1st	123 (15.2%)	200 (40%)
2nd	158 (19.5%)	119 (23.8%)
3rd	528 (65.3%)	189 (36.2%)

Descriptive Statistics with EZR(R)

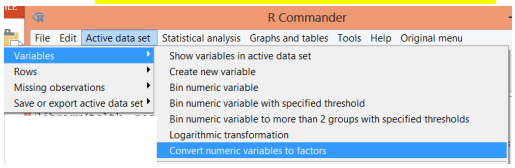


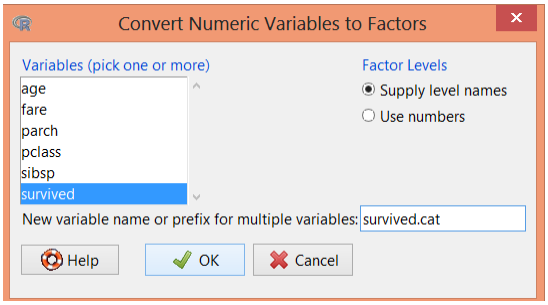
Describing data by groups (pclass)

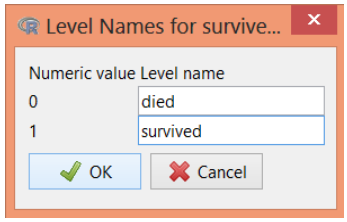


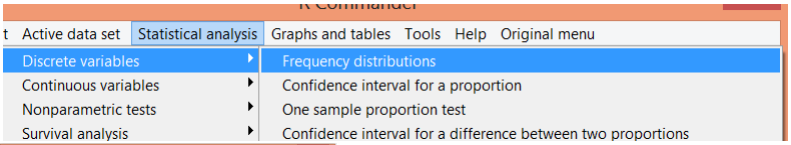
Describing data by groups (pclass)

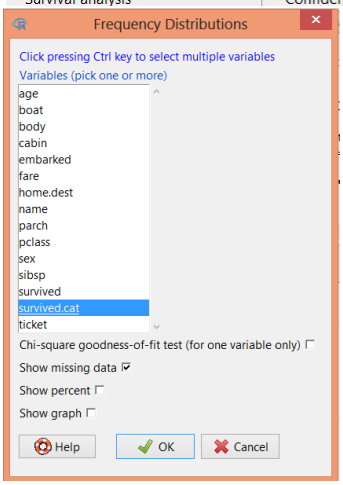
Convert "survived" to a factor variable





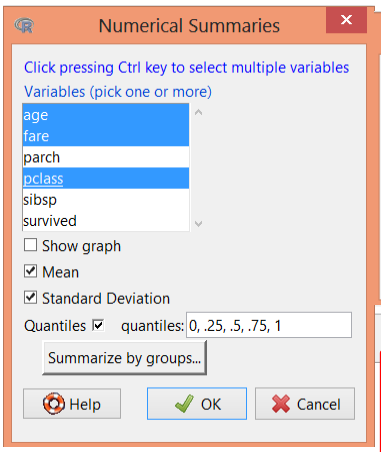


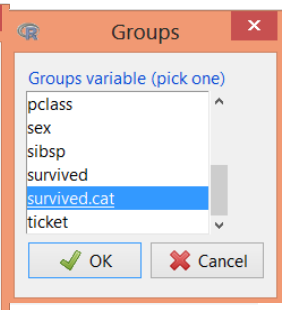





```
> #####Frequency Distributions#####

> (.Table <- table(titanic$survived.cat,
exclude=NULL))
+   # counts for survived.cat

   died survived     <NA>
    809      500        0

> round(100*.Table/sum(.Table), 2) #
percentages for   survived.cat

   died survived     <NA>
   61.8     38.2      0.0
```

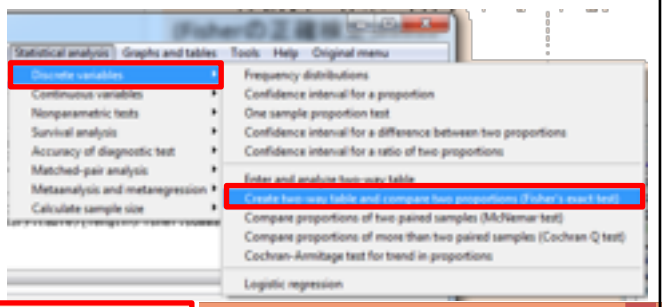
Describing data by groups (survived)



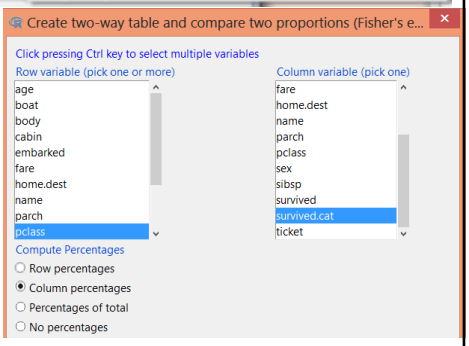



```
Variable: age
              mean        sd      0% 25% 50% 75% 100%    n  NA
died      30.54537 13.92254 0.3333  21  28  39   74 619 190
survived  28.91823 15.06148 0.1667  20  28  38   80 427  73

Variable: fare
              mean       sd 0%      25%  50%    75%      100%   n NA
died      23.35383 34.1451  0  7.8542 10.5 26.00 263.0000 808  1
survived  49.36118 68.6488  0 11.2146 26.0 57.75 512.3292 500  0

Variable: pclass
              mean        sd 0% 25% 50% 75% 100%   n NA
died      2.500618 0.7448251  1   2   3   3    3 809  0
survived  1.962000 0.8729720  1   1   2   3    3 500  0
```

Cross-tabulation table




```
> .Table
      survived.cat
pclass died survived
     1  123      200
     2  158      119
     3  528      181

> colPercents(.Table) # 列のパーセント表示
      survived.cat
pclass  died survived
  1     15.2     40.0
  2     19.5     23.8
  3     65.3     36.2
  Total 100.0    100.0
  Count 809.0    500.0
```

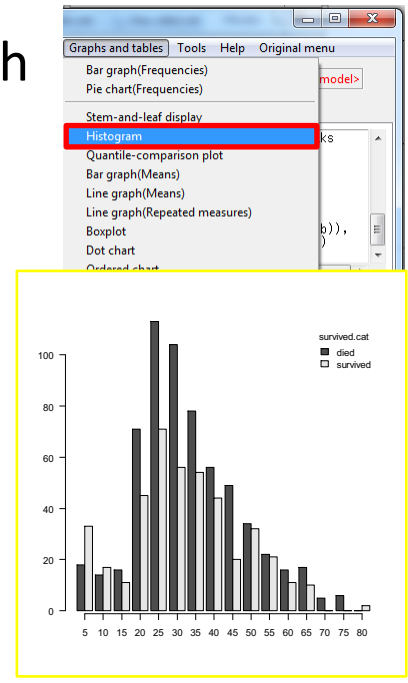


Passenger Characteristics

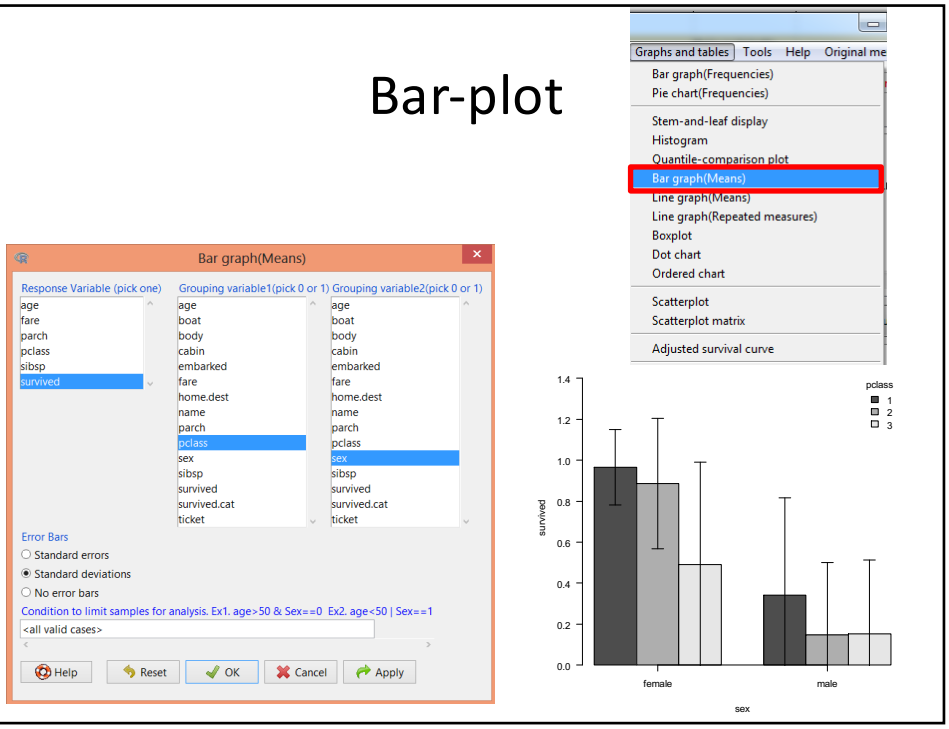
	Died (N=809)	Survived (N=500)
Age、Mean (SD)	30.5 (13.9)	28.9 (15.1)
Fare、Mean (SD)	23.3 (34.1)	49.4 (68.6)
Passenger Class, N（%）		
1st	123 (15.2%)	200 (40%)
2nd	158 (19.5%)	119 (23.8%)
3rd	528 (65.3%)	189 (36.2%)

Is there a better way to show the data?

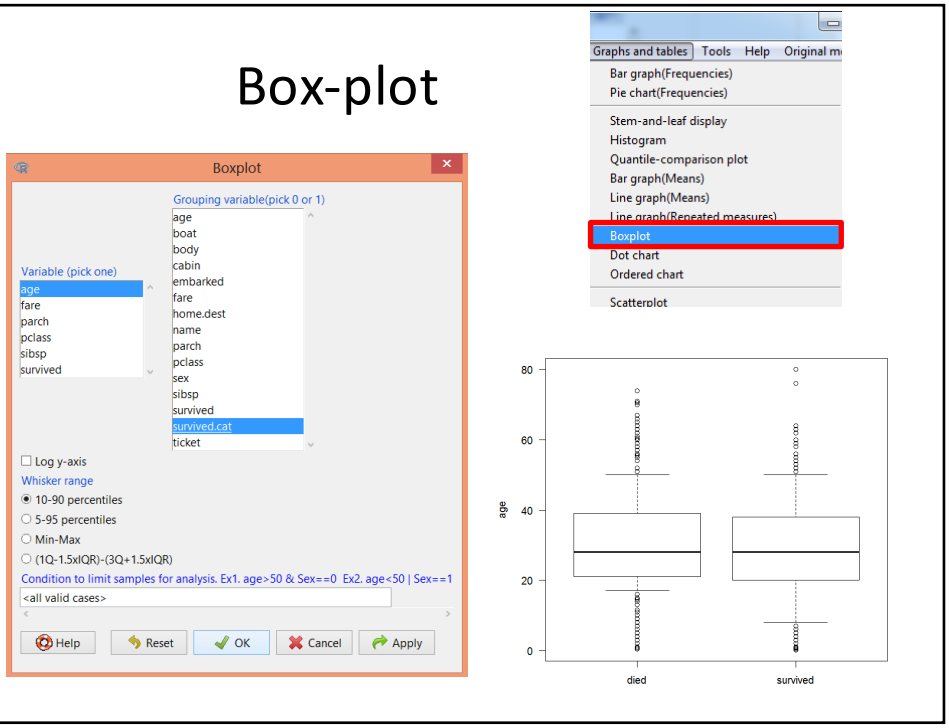
Graph



Bar-plot



Box-plot



Scatter-plot