5-number summary.

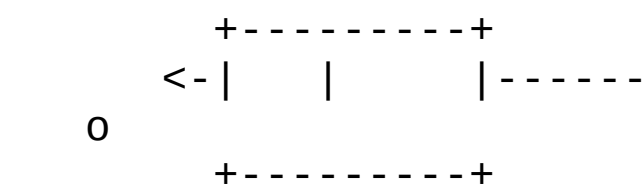
summary()

median, quartiles, extremes

```
> summary(islands)

   Min. 1st Qu.  Median    Mean 3rd Qu.
Max.
   12.0    20.5    41.0  1253.0   183.3
16990.0
```

Boxplot. box-and-whiskers



boxplot()

box: median and hinges (L,U)

(inner) fences: L – 1.5*IQR, U + 1.5*IQR

outliers: values outside fences

whiskers: arrows to most extreme values inside fences

Advantages.

Shows major features of univariate variable: location, spread, skewness, tail-length, outliers

Can see effect of transforms (graphics window)

Defines outliers

Summary resistant to outliers

Disadvantages.

Less detail than stem-and-leaf

Nitrogen example – covered up two isolated subgroups

Transformations. To make results more informative

y = g(x), e.g. y = log(x), y = sqrt(x), y = x^a

Box-Cox: y = (x^a – 1)/a

Can change origin also

Usually monotonic, 1-1

To deal with:

- asymmetry (make center clearer)
- outliers
- nonadditivity / nonadditivity
- spread dependence

Comparing batches.

displays side by side or in matrix

parallel boxplots

looking for similarities and differences (wrt center, spread, symmetry, tails, outliers, …)

boxplot display handles different sample sizes