**5-number summary.**

summary()

median, quartiles, extremes

> summary(islands)

<table>
<thead>
<tr>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.0</td>
<td>20.5</td>
<td>41.0</td>
<td>1253.0</td>
<td>183.3</td>
<td>16990.0</td>
</tr>
</tbody>
</table>

**Boxplot. box-and-whiskers**

```
+---------+
|<--|      |------>      |
| o |         | o |
| o |
+---------+
```

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

\[ y = g(x), \text{ e.g. } y = \log(x), \ y = \sqrt{x}, \ y = x^a \]

Box-Cox: \[ y = (x^a - 1)/a \]

Can change origin also

Usually monotonic, 1-1

To deal with:

asymmetry (make center clearer)

outliers

nonadditivty / nonadditivty

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