Numerical Measures

mean:

```
duration = faithful$eruptions # the eruption durations
mean(duration)
                       # apply the mean function
duration = faithful$eruptions # the eruption durations
median(duration)
                        # apply the median function
[1]4
duration = faithful$eruptions # the eruption durations
                        # apply the quantile function
quantile(duration)
 0% 25% 50% 75% 100%
1.6000\,2.1627\,4.0000\,4.4543\,5.1000
Percentile
duration = faithful$eruptions # the eruption durations
quantile(duration, c(.32, .57, .98))
 32% 57% 98%
2.3952 4.1330 4.9330
Range
duration = faithful$eruptions # the eruption durations
max(duration) – min(duration) # apply the max and min functions
```

```
[1] 3.5
```

Interquartile Range

duration = faithful\$eruptions # the eruption durations

IQR(duration) # apply the IQR function

[1] 2.2915

duration = faithful\$eruptions # the eruption durations
boxplot(duration, horizontal=TRUE) # horizontal box plot

duration = faithful\$eruptions # the eruption durations

var(duration) # apply the var function

[1] 1.3027

duration = faithful\$eruptions # the eruption durations

sd(duration) # apply the sd function

[1] 1.1414

duration = faithful\$eruptions # eruption durations

waiting = faithful\$waiting # the waiting period

cov(duration, waiting) # apply the cov function

[1] 13.978

duration = faithful\$eruptions # eruption durations

```
waiting = faithful$waiting # the waiting period
cor(duration, waiting) # apply the cor function
[1] 0.90081
```

Central Moment

library(e1071) # load e1071

duration = faithful\$eruptions # eruption durations

moment(duration, order=3, center=TRUE)

[1]-0.6149

Skewness

library(e1071) # load e1071

duration = faithful\$eruptions # eruption durations

skewness(duration) # apply the skewness function

[1]-0.41355

Kurtosis

library(e1071) # load e1071

duration = faithful\$eruptions # eruption durations

kurtosis(duration) # apply the kurtosis function

[1]-1.5116

Box Plot

The **box plot** of an observation variable is a graphical representation based on its quartiles, as well as its smallest and largest values. It

attempts to provide a visual shape of the data distribution. You can observe the 4 quartiles and dark vertical line represents the median. Find the shorter and longer distances for bunched data and spread data.

