Many R functions have an argument na. rm=, which can be set to be TRUE in order to remove NAs or FALSE. This is a convenient alternative to using constructs such as x[!is.na(x)].

1.3.4 Managing the work environment

If an R session runs long enough, there typically comes a point when there are more variables defined than can be remembered. The ls() function and the objects ects () function will list all the objects (variables, functions, etc.) in a given work environment. The browseEnv() function does so using web browser to show the results. The simplest usage is ls(), which shows all the objects that have been defined or loaded into your work environment. To filter the request, the pattern= argument can be used. If this argument is a quoted string then all objects with that string in their names will be listed. More complicated matching patterns are possible.

To trim down the size of the work environment the functions rm () or remove () can be used. These are used by specifying a name of the objects to be removed. The name may or may not be quoted. For example, rm ("tmp") or rm (tmp) will remove the variable tmp from the current work environment. Multiple names are possible if separated by commas, or if given in vector form, as quoted strings, to the argument list=.

1.3.5 Problems

1.14 You track your commute times for two weeks (ten days), recording the following times in minutes:

17 16 20 24 22 15 21 15 17 22

Enter these into R. Use the function max() to find the longest commute time, the function mean() to find the average, and the function min() to find the minimum.

Oops, the 24 was a mistake. It should have been 18. How can you fix this? Do so, and then find the new average.

How many times was your commute 20 minutes or more? What percent of your commutes are less than 18 minutes long?

1.15 According to *The Digital Bits* (http://www.digitalbits.com/), monthly sales (in 10,000s) of DVD players in 2003 were

JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
79	74	161	127	133	210	99	143	249	249	368	302

Enter the data into a data vector dvd. By slicing, form two data vectors: one containing the months with 31 days, the other the remaining months. Compare the means of these two data vectors.

1.16 Your cell-phone bill varies from month to month. The monthly amounts in dollars for the last year were

46 33 39 37 46 30 48 32 49 35 30 48

Enter this data into a variable called bill. Use the sum function to find the amount you spent last year on the cell phone. What is the smallest amount you spent in a month? What is the largest? How many months was the amount greater than $40? What percentage was this?

1.17 The average salary in major league baseball for the years 1990–1999 are given (in millions) by:

```
0.57 0.89 1.08 1.12 1.18 1.07 1.17 1.38 1.44 1.72
```

Use diff () to find the differences from year to year. Are there any years where the amount dropped from the previous year?

The percentage difference is the difference divided by the previous year times 100. This can be found by dividing the output of diff () by the first nine numbers (not all ten). After doing this, determine which year has the biggest percentage increase.

1.18 Define x and y with

```
> x = c(1, 3, 5, 7, 9)
> y = c(2, 3, 5, 7, 11, 13)
```

Try to guess the results of these R commands:

1. x+1
2. y*2
3. length (x) and length (y)
4. x+y (recycling)
5. sum(x>5) and sum(x[x>5])
6. sum(x>5|x<3)
7. y[3]
8. y[−3]
9. y[x] (What is NA?)
10. y[y>=7]

Remember that you access entries in a vector with [].

1.19 Consider the following “inequalities.” Can you determine how the comparisons are being done?

```
> "ABCDE" == "ABCDE"
[1] TRUE
> "ABCDE" < "ABCDEF"
[1] TRUE
> ″ABCDE" < "abcde"
[1] TRUE
> ″ZZZZZ" < "aaaaa"
[1] TRUE
> "11" < "8"
[1] TRUE
```