# Basic R and how to read in data

This guide is partly based on online material from Amy Willis, Kiirsti Owen and Amelia McNamara, and the book "R for Data Science" by Hadley Wickham and Garrett Grolemund. Thank you amazing R community!

**R as a calculator**

In the Console window below, type: 2+2 and press Enter Also try:

```
2^5
```

```
## [1] 32
```

```
3/10
```

```
## [1] 0.3
```

```
(3+5)^2
```

```
## [1] 64
```

```
sqrt(4)
```

```
## [1] 2
```

Tip: To run a line (or multiple lines) of code from a script without typing them into the Console, select the line(s) you want to run and press Ctrl+Enter (Command+Enter on a Mac)

**Objects**

R stores data as objects. You create new objects when you assign a value to them using "<-":

```
x <- 3  # Check the "Environment" window!
```

Tip: use the R studio shortcut Alt+ - (Alt and the minus sign) to easily create the assignment symbol <-

```
y <- 6
x+y
```

```
## [1] 9
```

Tip: R is case sensitive so if you've defined your object as x, it will not recognise (capital) X. Similarly, the function for square root is sqrt, R will give you an error if you try to use Sqrt.

## Packages

Packages extend the functionality of base R. They are distributed via CRAN: the Comprehensive R Archive Network

To install a package, use: install.packages("packagename") You then need to load it, using library(packagename)

We will be using a collection of packages called the Tidyverse:

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

When you load the tidyverse, you'll see a message about conflicts. As there is an (increasingly) large number of packages in R, it is possible to have functions with the same name in more than one package. The message tells you that packages dplyr and stats both have a function called filter and the one that will be used is the one from dplyr. It is the one that was loaded last.

If you want to use a function from a particular package, you need to include packagename:: before the name of the function.

In this example, you can use stats::filter() instead of just filter() to use filter from the stats package.

```r
find("filter") # this shows you the packages a function belongs to, in order of priority
```

```
## [1] "package:dplyr" "package:stats"
```

The tidyverse packages we will be using mostly in this course are readr (for reading in data), dplyr (for transforming data) and ggplot2 (for plotting).

## Functions

When using the Tidyverse, you can call functions in two ways:

```r
sqrt(4) # base R
```

```
## [1] 2
```

```r
4 %>%
  sqrt  # "pipe" operator (you can read is as "and then...")
```

```
## [1] 2
```

Tip: use the R Studio shortcut Ctrl + Sft + M to create the pipe operator %>%

Tip: If you are not sure what a function does, type ?functionname in the Console, e.g. ?sqrt

**Reading in data**

Before we read in our data, let's consider where we have saved our data file. Since we want our code to be reusable (by us and other people), the last thing we want is to include the location of the file in our code, something like:

"C:/dimitra/data/datafile.csv"

The above would only work for me, and only for the particular computer where folder "dimitra" contains a folder called "data".

To avoid these issues, we need to do two things:

1. Use R projects. (I hope you are doing that already!) Save the data and R markdown file inside the R project. Exactly where you save your code doesn't matter, you just need to note the location of your data with respect to the .Rproj file.
2. Use the R package "here". "Here" points to the location of the .Rproj file (which is the working directory for your project), so you just need to add "here" in front of the relative path to your data file.

For example, if your data file (a comma-separated value (csv) file) was saved inside a "data" directory, you would say:

library(here)

fev_data <- read_csv(here("data/fev.csv"))

To read in a file that is saved in the same directory as the .Rproj file:

```
library(here)
```

```
## here() starts at C:/Users/charl/OneDrive/Documentos/University of Aberdeen/Introduction to Health Da
```

```
fev_data <- read_csv(here("fev.csv"))
```

```
## Rows: 654 Columns: 7
```

```
## -- Column specification ---------------------------------------------------
## Delimiter: ","
## dbl (7): seqnbr, subjid, age, fev, height, sex, smoke
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

(Remember to install the "here" package the first time.)

–> How would you use read_csv with the pipe operator?

```
"fev.csv" %>% read.csv()
```

```
##   seqnbr subjid age   fev height sex smoke
## 1      1    301   9 1.708   57.0   0     0
## 2      2    451   8 1.724   67.5   0     0
## 3      3    501   7 1.720   54.5   0     0
## 4      4    642   9 1.558   53.0   1     0
```

```
## 5      5    901  9 1.895  57.0  1    0
## 6      6   1701  8 2.336  61.0  0    0
## 7      7   1752  6 1.919  58.0  0    0
## 8      8   1753  6 1.415  56.0  0    0
## 9      9   1901  8 1.987  58.5  0    0
## 10    10   1951  9 1.942  60.0  0    0
## 11    11   1952  6 1.602  53.0  0    0
## 12    12   2001  8 1.735  54.0  1    0
## 13    13   2101  8 2.193  58.5  0    0
## 14    14   2401  8 2.118  60.5  1    0
## 15    15   3102  8 2.258  58.0  1    0
## 16    16   3142  7 1.932  53.0  1    0
## 17    17   3541  5 1.472  50.0  1    0
## 18    18   3551  6 1.878  53.0  0    0
## 19    19   4201  9 2.352  59.0  1    0
## 20    20   4301  9 2.604  61.5  1    0
## 21    21   4351  5 1.400  49.0  0    0
## 22    22   5151  5 1.256  52.5  0    0
## 23    23   5152  4 0.839  48.0  0    0
## 24    24   5201  7 2.578  62.5  1    0
## 25    25   5601  9 2.988  65.0  0    0
## 26    26   5642  3 1.404  51.5  1    0
## 27    27   5702  9 2.348  60.0  1    0
## 28    28   6042  5 1.755  52.0  1    0
## 29    29   6101  8 2.980  60.0  0    0
## 30    30   6801  9 2.100  60.0  0    0
## 31    31   6851  5 1.282  49.0  0    0
## 32    32   7201  9 3.000  65.5  1    0
## 33    33   7251  8 2.673  60.0  0    0
## 34    34   7252  7 2.093  57.5  0    0
## 35    35   7253  5 1.612  52.0  0    0
## 36    36   8501  8 2.175  59.0  0    0
## 37    37   8801  9 2.725  59.0  1    0
## 38    38   9101  8 2.071  55.0  1    0
## 39    39   9201  8 1.547  57.0  1    0
## 40    40   9301  8 2.004  57.0  1    0
## 41    41   9501  9 3.135  60.0  0    0
## 42    42   9801  8 2.420  59.0  1    0
## 43    43  10041  5 1.776  51.0  1    0
## 44    44  10401  8 1.931  57.0  0    0
## 45    45  10451  5 1.343  50.0  0    0
## 46    46  10601  9 2.076  57.0  0    0
## 47    47  10701  7 1.624  54.0  1    0
## 48    48  10751  8 1.344  52.5  0    0
## 49    49  10841  6 1.650  55.0  1    0
## 50    50  11201  8 2.732  60.5  1    0
## 51    51  11241  5 2.017  54.5  1    0
## 52    52  11501  9 2.797  61.5  0    0
## 53    53  12001  9 3.556  62.0  1    0
## 54    54  12201  8 1.703  54.5  1    0
## 55    55  12241  6 1.634  54.0  1    0
## 56    56  12402  9 2.570  57.0  1    0
## 57    57  13351  9 3.016  62.5  0    0
## 58    58  13701  7 2.419  60.0  0    0
```

```
## 59          59    13751    4 1.569    50.0    0        0
## 60          60    14051    8 1.698    57.5    0        0
## 61          61    14101    8 2.123    60.0    1        0
## 62          62    14201    8 2.481    60.0    0        0
## 63          63    14251    6 1.481    51.0    0        0
## 64          64    14252    4 1.577    49.0    0        0
## 65          65    14501    8 1.940    59.0    1        0
## 66          66    14541    6 1.747    57.5    1        0
## 67          67    14601    9 2.069    58.0    1        0
## 68          68    14651    7 1.631    55.5    0        0
## 69          69    15153    5 1.536    52.0    0        0
## 70          70    15201    9 2.560    60.5    0        0
## 71          71    15241    8 1.962    57.0    1        0
## 72          72    15301    8 2.531    58.0    0        0
## 73          73    15401    9 2.715    60.0    1        0
## 74          74    15443    9 2.457    59.0    1        0
## 75          75    15601    9 2.090    59.5    1        0
## 76          76    15701    7 1.789    56.0    1        0
## 77          77    15741    5 1.858    53.0    1        0
## 78          78    15842    5 1.452    51.0    1        0
## 79          79    16101    9 3.842    69.0    1        0
## 80          80    16151    6 1.719    53.0    0        0
## 81          81    16251    7 2.111    57.0    0        0
## 82          82    16252    6 1.695    53.0    0        0
## 83          83    16501    8 2.211    63.0    1        0
## 84          84    16601    8 1.794    54.5    1        0
## 85          85    16651    7 1.917    58.0    0        0
## 86          86    18101    8 2.144    63.0    0        0
## 87          87    18501    7 1.253    52.0    1        0
## 88          88    18843    9 2.659    61.5    1        0
## 89          89    18844    5 1.580    52.5    1        0
## 90          90    19401    9 2.126    62.0    1        0
## 91          91    20152    9 3.029    61.5    0        0
## 92          92    20801    9 2.964    64.5    1        0
## 93          93    20841    7 1.611    57.5    1        0
## 94          94    21552    8 2.215    60.0    0        0
## 95          95    22201    8 2.388    60.0    0        0
## 96          96    22901    9 2.196    61.0    1        0
## 97          97    23101    9 1.751    58.0    1        0
## 98          98    23441    9 2.165    61.5    1        0
## 99          99    23442    7 1.682    55.0    1        0
## 100        100    23641    8 1.523    55.0    1        0
## 101        101    23653    8 1.292    52.0    0        0
## 102        102    23801    7 1.649    54.0    1        0
## 103        103    23802    9 2.588    63.0    1        0
## 104        104    23841    4 0.796    47.0    1        0
## 105        105    23902    9 2.574    60.5    0        0
## 106        106    23941    6 1.979    56.0    1        0
## 107        107    24401    8 2.354    58.5    1        0
## 108        108    24843    6 1.718    55.0    1        0
## 109        109    24851    7 1.742    58.5    0        0
## 110        110    25001    7 1.603    51.0    0        0
## 111        111    25052    8 2.639    59.5    0        0
## 112        112    25201    7 1.829    54.0    0        0
```

```
## 113   113   27001   7 2.084   58.0   1   0
## 114   114   27401   7 2.220   58.0   1   0
## 115   115   27952   7 1.473   52.5   0   0
## 116   116   28452   8 2.341   60.5   0   0
## 117   117   28501   7 1.698   54.5   0   0
## 118   118   28551   5 1.196   46.5   0   0
## 119   119   28801   8 1.872   56.5   0   0
## 120   120   29001   7 2.219   55.0   1   0
## 121   121   29101   9 2.420   57.0   1   0
## 122   122   29151   7 1.827   54.5   0   0
## 123   123   29401   7 1.461   54.0   0   0
## 124   124   29443   6 1.338   53.0   1   0
## 125   125   29601   8 2.090   57.0   1   0
## 126   126   29701   8 1.697   59.0   0   0
## 127   127   29741   8 1.562   55.0   1   0
## 128   128   29901   9 2.040   55.5   0   0
## 129   129   29954   7 1.609   51.5   0   0
## 130   130   30001   8 2.458   61.0   0   0
## 131   131   30043   9 2.650   63.5   1   0
## 132   132   31001   8 1.429   57.5   1   0
## 133   133   31501   8 1.675   53.0   1   0
## 134   134   31551   9 1.947   56.5   0   0
## 135   135   32501   8 2.069   54.0   1   0
## 136   136   32542   6 1.572   52.0   1   0
## 137   137   32742   6 1.348   53.0   1   0
## 138   138   32751   8 2.288   61.5   0   0
## 139   139   33301   9 1.773   58.5   1   0
## 140   140   33351   5 0.791   52.0   0   0
## 141   141   33501   7 1.905   58.0   1   0
## 142   142   33601   9 2.463   61.0   0   0
## 143   143   33641   6 1.431   51.0   1   0
## 144   144   33801   9 2.631   62.0   0   0
## 145   145   34101   9 3.114   64.5   1   0
## 146   146   34301   9 2.135   58.5   1   0
## 147   147   34341   6 1.527   52.5   1   0
## 148   148   35102   8 2.293   58.0   0   0
## 149   149   35103   9 3.042   66.0   0   0
## 150   150   35601   8 2.927   63.5   1   0
## 151   151   37201   8 2.665   64.0   0   0
## 152   152   37301   9 2.301   58.5   1   0
## 153   153   37901   9 2.460   64.0   1   0
## 154   154   38051   9 2.592   60.5   0   0
## 155   155   38152   7 1.750   55.0   0   0
## 156   156   38241   8 1.759   53.0   1   0
## 157   157   38242   6 1.536   48.0   1   0
## 158   158   38801   9 2.259   58.5   0   0
## 159   159   39001   9 2.048   64.5   0   0
## 160   160   39101   9 2.571   60.5   1   0
## 161   161   39141   7 2.046   56.0   1   0
## 162   162   39201   8 1.780   58.5   0   0
## 163   163   39251   5 1.552   54.0   0   0
## 164   164   39301   8 1.953   58.0   0   0
## 165   165   39401   9 2.893   64.5   1   0
## 166   166   39741   6 1.713   50.5   1   0
```

```
## 167   167  39801   9 2.851   60.0   0     0
## 168   168  39841   6 1.624   51.5   1     0
## 169   169  39901   8 2.631   59.0   1     0
## 170   170  39941   5 1.819   53.0   1     0
## 171   171  40001   7 1.658   53.0   1     0
## 172   172  40501   7 2.158   53.5   1     0
## 173   173  40541   4 1.789   52.0   1     0
## 174   174  42101   9 3.004   64.0   0     0
## 175   175  42201   8 2.503   63.0   1     0
## 176   176  42501   9 1.933   58.0   0     0
## 177   177  42552   9 2.091   58.5   0     0
## 178   178  42901   9 2.316   59.5   0     0
## 179   179  42941   5 1.704   51.0   0     0
## 180   180  43201   9 1.606   57.5   0     0
## 181   181  43242   7 1.165   47.0   1     0
## 182   182  43651   6 2.102   55.5   0     0
## 183   183  43901   9 2.320   57.0   0     0
## 184   184  44201   9 2.230   61.0   1     0
## 185   185  44301   9 1.716   55.5   1     0
## 186   186  44501   7 1.790   53.5   1     0
## 187   187  44551   5 1.146   50.0   0     0
## 188   188  44701   8 2.187   61.5   0     0
## 189   189  45001   9 2.717   61.5   1     0
## 190   190  45041   7 1.796   55.0   1     0
## 191   191  45241   9 1.953   58.0   1     1
## 192   192  45251   8 1.335   56.5   0     0
## 193   193  45301   9 2.119   57.0   1     0
## 194   194  45641   6 1.666   52.0   1     0
## 195   195  45642   6 1.826   52.5   1     0
## 196   196  45653   8 2.709   62.5   0     0
## 197   197  46101   9 2.871   65.0   1     0
## 198   198  46151   5 1.092   50.0   0     0
## 199   199  46641   6 2.262   57.5   1     0
## 200   200  46642   6 2.104   56.5   1     0
## 201   201  47001   9 2.166   57.5   0     0
## 202   202  47053   7 1.690   54.0   0     0
## 203   203  47242   9 2.973   59.5   1     0
## 204   204  47301   8 2.145   59.5   0     0
## 205   205  47341   5 1.971   58.0   1     0
## 206   206  47552   7 2.095   57.0   0     0
## 207   207  48052   6 1.697   55.0   0     0
## 208   208  48101   9 2.455   60.0   0     0
## 209   209  48401   7 1.920   56.5   1     0
## 210   210  48402   9 2.164   60.0   1     0
## 211   211  48901   9 2.130   59.0   0     0
## 212   212  49301   8 2.993   63.0   0     0
## 213   213  49401   9 2.529   59.0   0     0
## 214   214  49501   7 1.726   53.0   0     0
## 215   215  49542   9 2.442   61.5   0     0
## 216   216  49551   4 1.102   48.0   0     0
## 217   217  49701   9 2.056   63.0   0     0
## 218   218  49741   5 1.808   55.5   1     0
## 219   219  49751   8 2.305   64.5   0     0
## 220   220  50501   9 1.969   59.0   0     0
```

```
## 221        221   50701    8 1.556    58.5     0        0
## 222        222   50951    3 1.072    46.0     0        0
## 223        223   51241    9 2.042    62.0     1        0
## 224        224   51301    8 1.512    53.0     0        0
## 225        225   51341    6 1.423    49.5     1        0
## 226        226   51501    9 3.681    68.0     1        0
## 227        227   51542    8 1.991    59.5     1        0
## 228        228   52101    8 1.897    55.5     1        0
## 229        229   53601    7 1.370    55.0     0        0
## 230        230   53651    6 1.338    51.5     0        0
## 231        231   54201    8 2.016    56.0     1        0
## 232        232   54701    9 2.639    63.0     0        0
## 233        233   54751    4 1.389    48.0     0        0
## 234        234   54941    7 1.612    56.5     1        0
## 235        235   54952    8 2.135    59.0     0        0
## 236        236   55645    8 2.681    60.5     1        0
## 237        237   55652    9 3.223    65.0     0        0
## 238        238   57651    6 1.796    55.0     0        0
## 239        239   57901    8 2.010    55.0     1        0
## 240        240   58341    6 1.523    51.0     0        0
## 241        241   58601    8 1.744    52.5     1        0
## 242        242   58602    9 2.485    64.0     0        0
## 243        243   60051    8 2.335    59.0     0        0
## 244        244   60251    7 1.415    53.5     0        0
## 245        245   60801    9 2.076    60.5     1        0
## 246        246   61101    8 2.435    59.5     1        0
## 247        247   61601    7 1.728    56.5     0        0
## 248        248   61801    9 2.850    63.0     0        0
## 249        249   61901    8 1.844    56.5     0        0
## 250        250   62301    9 1.754    61.5     0        0
## 251        251   62351    6 1.343    52.0     0        0
## 252        252   62641    8 2.303    57.0     1        0
## 253        253   62702    9 2.246    63.5     1        0
## 254        254   63102    8 2.476    63.0     0        0
## 255        255   63201    9 3.239    65.0     1        0
## 256        256   63941    9 2.457    61.5     1        0
## 257        257   64101    8 2.382    62.0     0        0
## 258        258   64151    7 1.640    55.0     0        0
## 259        259   64152    5 1.589    51.0     0        0
## 260        260   64201    7 2.056    54.0     1        0
## 261        261   70001    8 2.226    57.0     1        0
## 262        262   70401    9 1.886    56.0     0        0
## 263        263   71101    9 2.833    61.5     1        0
## 264        264   71241    6 1.715    53.0     1        0
## 265        265   71401    8 2.631    59.0     1        0
## 266        266   71444    7 2.550    56.0     1        0
## 267        267   71851    9 1.912    59.0     0        0
## 268        268   72001    7 1.877    52.5     0        0
## 269        269   72051    7 1.935    52.5     0        0
## 270        270   72052    5 1.539    50.0     0        0
## 271        271   72501    9 2.803    59.5     1        0
## 272        272   73001    9 2.923    64.0     1        0
## 273        273   73151    8 2.358    61.0     0        0
## 274        274   73342    8 2.094    57.5     1        0
```

```
## 275   275   74201    9 1.855   60.0   1   0
## 276   276   74241    6 1.535   55.0   0   0
## 277   277   74401    7 2.135   56.0   1   0
## 278   278   74441    5 1.930   51.0   1   0
## 279   279   74601    9 2.182   59.5   0   0
## 280   280   74641    5 1.359   50.5   1   0
## 281   281   74652    7 2.002   57.5   0   0
## 282   282   74941    6 1.699   54.0   1   0
## 283   283   75701    8 2.500   57.0   1   0
## 284   284   75751    7 2.366   58.0   0   0
## 285   285   75901    8 2.069   60.0   0   0
## 286   286   75951    4 1.418   49.0   0   0
## 287   287   76501    8 2.333   57.0   0   0
## 288   288   76541    5 1.514   52.0   1   0
## 289   289   76751    8 1.758   52.0   0   0
## 290   290   77141    7 2.535   59.5   1   0
## 291   291   77901    7 2.564   58.0   0   0
## 292   292   78301    9 2.487   64.0   0   0
## 293   293   78352    9 1.591   57.0   0   0
## 294   294   80001    8 1.624   53.0   1   0
## 295   295   80301    9 2.798   62.0   1   0
## 296   296   80341    6 1.691   53.0   1   0
## 297   297   80601    8 1.999   56.5   0   0
## 298   298   80801    9 1.869   57.0   1   0
## 299   299   80841    4 1.004   48.0   1   0
## 300   300   81241    6 1.427   49.5   1   0
## 301   301   81401    7 1.826   51.0   1   0
## 302   302   81451    9 2.688   59.5   0   0
## 303   303   81501    8 1.657   56.0   1   0
## 304   304   81751    6 1.672   54.0   0   0
## 305   305   82701    8 2.015   57.5   0   0
## 306   306   83801    7 2.371   55.5   0   0
## 307   307   83841    5 2.115   50.0   1   0
## 308   308   83901    8 2.328   60.0   0   0
## 309   309   83952    7 1.495   57.0   0   0
## 310   310     201   11 2.884   69.0   1   0
## 311   311     202   10 2.328   64.0   1   0
## 312   312     341   14 3.381   63.0   1   0
## 313   313     351   11 2.170   58.0   0   0
## 314   314     401   11 3.470   66.5   1   0
## 315   315     551   12 3.058   60.5   0   0
## 316   316     601   10 1.811   57.0   1   0
## 317   317     641   11 2.524   64.0   1   0
## 318   318    1751   10 2.642   61.0   0   0
## 319   319    2041   14 3.741   68.5   1   0
## 320   320    2042   13 4.336   69.5   1   0
## 321   321    2142   14 4.842   72.0   1   0
## 322   322    2143   12 4.550   71.0   1   0
## 323   323    2451   12 2.841   63.0   0   0
## 324   324    2801   10 3.166   61.5   0   0
## 325   325    2851   13 3.816   63.5   0   0
## 326   326    3141   10 2.561   62.0   1   0
## 327   327    3501   11 3.654   65.0   0   0
## 328   328    3901   10 2.481   61.0   1   0
```

```
## 329   329    4001  11 2.665   63.0  0       0
## 330   330    4341  10 3.203   66.0  1       0
## 331   331    4901  13 3.549   68.0  1       0
## 332   332    4952  14 2.236   66.0  0       1
## 333   333    5001  11 3.222   72.0  1       0
## 334   334    5101  10 3.111   66.0  1       0
## 335   335    5251  11 3.490   67.0  0       0
## 336   336    5351  13 3.147   64.0  0       0
## 337   337    5352  10 2.520   60.5  0       0
## 338   338    5641  10 2.292   63.0  1       0
## 339   339    5701  12 2.889   64.0  0       0
## 340   340    6001  10 2.246   60.5  1       0
## 341   341    6041  10 1.937   62.0  1       0
## 342   342    6145  10 2.646   60.0  1       0
## 343   343    6201  11 2.957   64.5  1       0
## 344   344    6401  11 4.007   67.0  1       0
## 345   345    7101  11 2.386   61.5  0       0
## 346   346    7241  10 3.251   66.0  1       0
## 347   347    7901  11 2.762   60.0  0       0
## 348   348    8301  11 3.011   64.0  0       0
## 349   349    8541  13 4.305   68.5  1       0
## 350   350    8542  13 3.906   67.0  1       0
## 351   351    8842  11 3.583   67.0  1       0
## 352   352    8901  11 3.236   66.0  0       0
## 353   353    9141  14 3.436   62.5  1       0
## 354   354    9142  11 3.058   61.0  1       0
## 355   355    9502  10 3.007   62.0  1       0
## 356   356    9802  10 3.489   66.5  1       0
## 357   357   10001  10 2.864   60.0  0       0
## 358   358   10053  14 3.428   64.0  0       1
## 359   359   10054  13 2.819   62.0  0       0
## 360   360   10501  10 2.250   58.0  0       0
## 361   361   10642  14 4.683   68.5  1       0
## 362   362   10801  10 2.352   61.5  1       0
## 363   363   10901  11 3.108   64.5  1       0
## 364   364   10942  13 3.994   67.0  1       0
## 365   365   11101  12 4.393   68.5  1       0
## 366   366   11151  13 3.208   61.0  0       1
## 367   367   11301  10 2.592   65.0  1       0
## 368   368   11341  13 3.193   70.0  1       0
## 369   369   11601  11 1.694   60.0  1       1
## 370   370   11642  14 3.957   72.0  1       1
## 371   371   11901  11 2.346   59.0  0       0
## 372   372   11942  13 4.789   69.0  1       1
## 373   373   12501  11 3.515   67.5  1       0
## 374   374   13301  11 2.754   65.5  0       0
## 375   375   14001  10 2.720   65.5  1       0
## 376   376   14143  11 2.463   64.5  1       0
## 377   377   14401  11 2.633   62.0  0       0
## 378   378   15101  10 3.048   65.5  0       0
## 379   379   15141  11 3.111   67.5  1       0
## 380   380   15152  13 3.745   68.0  0       0
## 381   381   15302  12 2.384   63.5  0       1
## 382   382   15342  10 2.094   58.5  1       0
```

```
## 383   383   15451   10 3.183   65.5   0   0
## 384   384   15751   14 3.074   65.0   0   1
## 385   385   15801   11 3.977   70.5   1   0
## 386   386   15841   10 3.354   63.0   1   0
## 387   387   16201   11 3.411   63.5   0   0
## 388   388   16551   10 2.387   66.0   0   1
## 389   389   16901   11 3.171   63.0   0   0
## 390   390   16941   13 3.887   67.5   1   0
## 391   391   16951   13 2.646   61.5   0   0
## 392   392   16952   10 2.504   60.0   0   0
## 393   393   17301   11 3.587   64.5   1   0
## 394   394   17401   11 3.845   68.5   1   0
## 395   395   17801   12 2.971   64.5   1   0
## 396   396   17851   10 2.891   61.0   0   0
## 397   397   18502   10 1.823   57.0   0   0
## 398   398   18541   11 2.417   62.5   1   0
## 399   399   18551   10 2.175   58.0   0   0
## 400   400   18801   11 2.735   62.5   0   0
## 401   401   18841   14 4.273   72.5   1   0
## 402   402   18842   13 2.976   65.5   1   0
## 403   403   19601   12 3.835   69.5   0   1
## 404   404   19901   11 4.065   66.5   1   0
## 405   405   20101   11 2.318   59.0   0   0
## 406   406   21301   11 3.596   68.0   1   0
## 407   407   21352   14 3.395   67.0   0   0
## 408   408   21353   12 2.751   63.0   0   0
## 409   409   21501   10 2.673   64.5   0   0
## 410   410   22253   12 2.556   62.0   0   0
## 411   411   23151   11 2.542   62.0   0   0
## 412   412   23401   10 2.608   66.0   1   0
## 413   413   23601   11 2.354   62.0   0   0
## 414   414   23651   13 2.599   62.5   0   1
## 415   415   23652   10 1.458   57.0   0   0
## 416   416   23901   10 3.795   68.5   1   0
## 417   417   24201   11 2.491   59.0   0   0
## 418   418   24251   13 3.060   61.5   0   0
## 419   419   24501   10 2.545   65.0   1   0
## 420   420   24543   11 2.993   66.5   1   0
## 421   421   24601   10 3.305   65.0   0   0
## 422   422   24642   13 4.756   68.0   1   1
## 423   423   24701   11 3.774   67.0   0   0
## 424   424   24741   10 2.855   64.5   1   0
## 425   425   24801   11 2.988   70.0   1   0
## 426   426   25041   11 2.498   60.0   1   0
## 427   427   25051   14 3.169   64.0   0   0
## 428   428   25501   11 2.887   62.5   1   0
## 429   429   25551   13 2.704   61.0   0   0
## 430   430   25901   11 3.515   64.0   0   0
## 431   431   27441   11 3.425   65.5   1   0
## 432   432   27701   10 2.287   61.0   0   0
## 433   433   27751   13 2.434   65.4   0   0
## 434   434   27901   10 2.365   63.5   0   0
## 435   435   27951   13 3.086   67.5   0   1
## 436   436   28401   10 2.696   66.0   1   0
```

```
## 437    437   29952   12 2.868    62.0    0       0
## 438    438   29953   10 2.813    61.5    0       0
## 439    439   30042   14 4.309    69.0    1       1
## 440    440   30051   12 3.255    66.0    0       0
## 441    441   30052   10 3.413    66.0    0       1
## 442    442   30401   11 4.593    69.0    1       0
## 443    443   31042   14 4.111    71.0    1       0
## 444    444   31201   12 1.916    60.5    1       0
## 445    445   31242   10 1.858    58.0    1       0
## 446    446   31502   10 2.975    63.0    0       1
## 447    447   31901   10 3.350    69.0    1       0
## 448    448   32541   10 2.901    59.5    1       0
## 449    449   32701   12 2.241    64.0    1       0
## 450    450   32741   13 4.225    74.0    1       0
## 451    451   33001   11 3.223    64.5    0       0
## 452    452   33041   12 5.224    70.0    1       0
## 453    453   33502   11 4.073    67.0    1       0
## 454    454   33541   12 4.080    64.5    1       0
## 455    455   33701   11 2.606    65.0    0       0
## 456    456   34201   11 3.169    62.5    0       1
## 457    457   34243   12 4.411    68.0    1       0
## 458    458   35101   12 3.791    68.5    1       0
## 459    459   35901   13 3.089    67.5    1       0
## 460    460   35941   11 2.465    60.0    1       0
## 461    461   36001   12 3.343    68.0    1       1
## 462    462   36041   10 3.200    65.0    1       0
## 463    463   36101   12 2.913    64.0    1       0
## 464    464   37241   13 4.877    73.0    1       0
## 465    465   37351   10 2.358    59.0    0       0
## 466    466   37401   12 3.279    70.5    1       0
## 467    467   37701   10 2.581    66.0    1       0
## 468    468   37702   12 2.347    61.5    0       0
## 469    469   37951   10 2.691    67.0    0       0
## 470    470   38001   11 2.827    62.5    0       0
## 471    471   38201   10 1.873    52.5    1       0
## 472    472   39041   12 3.751    72.0    1       1
## 473    473   39052   14 2.538    71.0    0       0
## 474    474   39601   10 2.758    65.5    1       0
## 475    475   39701   10 3.050    60.0    0       0
## 476    476   40051   12 3.079    60.0    0       0
## 477    477   40301   10 2.201    60.5    1       0
## 478    478   41601   10 1.858    59.0    1       0
## 479    479   42151   13 2.216    68.0    0       1
## 480    480   42251   12 3.403    62.0    0       0
## 481    481   42301   12 3.501    64.5    0       0
## 482    482   42502   11 2.578    63.0    0       0
## 483    483   42551   13 3.078    66.0    0       1
## 484    484   43002   12 3.186    67.0    0       1
## 485    485   43241   10 1.665    57.0    1       0
## 486    486   43251   11 2.081    63.0    0       0
## 487    487   43301   11 2.974    62.0    0       0
## 488    488   43351   13 3.297    65.0    0       1
## 489    489   43601   12 4.073    68.5    1       0
## 490    490   43641   13 4.448    69.0    1       0
```

```
## 491    491   44242  13 3.984    71.0    1       0
## 492    492   44302  10 2.250    58.0    0       0
## 493    493   44351  12 2.752    63.5    0       0
## 494    494   44601  12 2.304    66.5    1       1
## 495    495   44742  14 3.680    67.0    1       0
## 496    496   45201  11 3.102    64.0    0       1
## 497    497   45601  10 2.862    61.0    0       0
## 498    498   45651  13 2.677    67.0    0       1
## 499    499   45652  11 3.023    67.5    0       0
## 500    500   46301  11 3.681    68.0    0       0
## 501    501   46353  13 3.255    66.5    0       0
## 502    502   46601  12 3.692    67.0    1       0
## 503    503   46651  10 2.356    60.5    0       0
## 504    504   46901  10 4.591    70.0    1       0
## 505    505   46951  12 3.082    63.5    0       0
## 506    506   47051  13 3.297    65.0    0       1
## 507    507   47052  11 3.258    63.0    0       0
## 508    508   47801  10 2.216    61.0    1       0
## 509    509   47841  11 3.247    65.5    1       0
## 510    510   48001  11 4.324    67.5    1       0
## 511    511   48142  11 2.362    61.0    0       0
## 512    512   48951  11 2.563    63.0    0       0
## 513    513   49201  11 3.206    63.5    1       0
## 514    514   49341  14 3.585    70.0    1       0
## 515    515   49342  12 4.720    71.5    1       0
## 516    516   49451  13 3.331    65.5    0       0
## 517    517   49541  13 5.083    74.0    1       0
## 518    518   50301  10 3.498    68.0    1       1
## 519    519   50551  12 2.417    61.0    0       0
## 520    520   50702  10 2.364    61.0    1       0
## 521    521   50901  10 2.341    61.0    1       0
## 522    522   51101  12 2.759    61.5    0       1
## 523    523   51201  11 2.953    67.0    0       1
## 524    524   51302  12 3.231    63.0    1       0
## 525    525   51541  11 3.078    67.5    1       0
## 526    526   51601  11 3.369    70.5    1       0
## 527    527   51641  12 3.529    70.5    1       0
## 528    528   52401  12 2.866    62.0    0       0
## 529    529   52452  14 2.891    62.0    0       0
## 530    530   52801  11 3.022    61.5    0       0
## 531    531   53101  10 3.127    62.0    1       0
## 532    532   53151  11 2.866    60.5    0       0
## 533    533   53301  12 2.605    62.5    0       0
## 534    534   54251  13 3.056    63.0    0       0
## 535    535   54901  12 2.569    63.0    0       0
## 536    536   54951  11 2.501    62.0    0       0
## 537    537   55601  11 3.320    65.5    1       0
## 538    538   55901  11 2.123    65.0    1       0
## 539    539   55941  14 3.780    70.0    1       0
## 540    540   56601  11 3.847    66.0    1       0
## 541    541   56651  13 3.785    63.0    0       1
## 542    542   56901  12 3.924    68.0    1       0
## 543    543   57001  10 2.132    59.0    1       0
## 544    544   57042  12 2.752    68.5    1       0
```

```
## 545  545  57053  13 2.449  63.0  0  0
## 546  546  57601  10 3.456  63.0  1  0
## 547  547  58301  10 3.073  66.0  0  0
## 548  548  58603  10 2.688  62.0  0  0
## 549  549  59901  10 3.329  68.0  1  0
## 550  550  59941  14 4.271  72.5  1  0
## 551  551  60001  12 3.530  64.0  1  0
## 552  552  60201  11 2.928  65.5  1  0
## 553  553  60851  11 2.689  61.5  0  0
## 554  554  61301  12 2.332  57.0  1  0
## 555  555  61351  14 2.934  64.0  0  0
## 556  556  61941  14 2.276  66.0  1  1
## 557  557  62601  10 3.110  64.5  1  0
## 558  558  62701  11 2.894  67.0  1  0
## 559  559  62801  11 4.637  72.0  1  1
## 560  560  62802  10 2.435  65.0  0  0
## 561  561  63101  10 2.838  63.0  0  0
## 562  562  63252  12 3.035  62.0  0  0
## 563  563  63901  12 4.831  71.0  1  0
## 564  564  64241  11 2.812  61.0  1  0
## 565  565  70051  12 2.714  65.5  0  0
## 566  566  70052  10 3.086  62.0  0  0
## 567  567  70201  12 3.519  65.5  0  0
## 568  568  70441  13 4.232  70.5  1  0
## 569  569  70442  10 2.770  62.0  1  0
## 570  570  71151  12 3.341  65.5  0  0
## 571  571  71201  10 3.090  65.0  1  0
## 572  572  71443  13 2.531  61.0  1  0
## 573  573  71801  12 2.822  69.5  1  0
## 574  574  72552  10 3.038  65.0  0  1
## 575  575  72901  12 2.935  65.5  1  0
## 576  576  73101  10 2.568  63.5  0  0
## 577  577  73301  11 2.387  60.5  1  0
## 578  578  73341  12 2.499  65.0  1  0
## 579  579  73701  11 4.130  67.0  1  0
## 580  580  74651  12 3.001  63.5  0  0
## 581  581  74701  10 3.132  59.5  0  0
## 582  582  74751  13 3.577  63.5  0  0
## 583  583  74752  12 3.222  61.0  0  0
## 584  584  74901  11 3.280  66.0  1  0
## 585  585  75801  11 2.659  64.0  1  0
## 586  586  76601  11 2.822  62.0  0  0
## 587  587  76701  11 2.140  60.5  0  0
## 588  588  77101  12 4.203  71.0  1  0
## 589  589  77152  14 2.997  64.5  0  0
## 590  590  77501  11 3.120  61.0  0  1
## 591  591  77601  11 2.562  62.5  0  0
## 592  592  80051  12 3.082  64.5  0  0
## 593  593  80641  14 3.806  68.0  1  0
## 594  594  80642  11 3.339  68.5  1  1
## 595  595  80651  13 3.152  62.0  0  1
## 596  596  81201  11 2.458  60.0  0  0
## 597  597  81541  10 2.391  59.5  1  0
## 598  598  81551  13 3.141  61.0  0  0
```

```
## 599   599   81552   12 2.579    63.0    0       0
## 600   600   81701   11 3.104    67.5    0       1
## 601   601   81741   13 4.045    69.0    1       1
## 602   602   82743   14 4.763    68.0    1       1
## 603   603   82744   10 2.100    58.0    1       0
## 604   604   83951   11 3.069    65.0    0       1
## 605   605   90001   11 2.785    69.0    1       0
## 606   606     441   15 4.284    70.0    1       0
## 607   607    4041   15 4.506    71.0    1       1
## 608   608    4051   18 2.906    66.0    0       0
## 609   609    6144   19 5.102    72.0    1       0
## 610   610    6252   19 3.519    66.0    0       1
## 611   611    6441   16 3.688    68.0    1       1
## 612   612    7142   17 4.429    70.0    1       0
## 613   613    8841   15 4.279    67.5    1       0
## 614   614   10941   15 4.500    70.0    1       0
## 615   615   15351   15 2.635    64.0    0       0
## 616   616   15752   15 2.679    66.0    0       1
## 617   617   20151   15 2.198    62.0    0       1
## 618   618   21351   19 3.345    65.5    0       1
## 619   619   22251   18 3.082    64.5    0       0
## 620   620   22252   16 3.387    66.5    0       0
## 621   621   24541   17 3.082    67.0    1       1
## 622   622   24552   16 2.903    63.0    0       1
## 623   623   24553   15 3.004    64.0    0       1
## 624   624   25941   15 5.793    69.0    1       0
## 625   625   30041   15 3.985    71.0    1       0
## 626   626   30441   18 4.220    68.0    1       0
## 627   627   30442   17 4.724    70.5    1       0
## 628   628   33741   15 3.731    67.0    1       0
## 629   629   34241   17 3.406    69.0    1       1
## 630   630   37251   17 3.500    62.0    0       0
## 631   631   37252   16 3.674    67.5    0       0
## 632   632   37441   17 5.633    73.0    1       0
## 633   633   37451   15 3.122    64.0    0       1
## 634   634   39051   15 3.330    68.5    0       1
## 635   635   40351   16 2.608    62.0    0       1
## 636   636   44241   16 3.645    73.5    1       0
## 637   637   47241   15 3.799    66.5    1       1
## 638   638   48141   18 4.086    67.0    1       1
## 639   639   48152   15 2.887    63.0    0       0
## 640   640   48441   16 4.070    69.5    1       1
## 641   641   52841   17 3.960    70.0    1       0
## 642   642   52842   16 4.299    66.0    1       0
## 643   643   55951   16 2.981    66.0    0       0
## 644   644   57052   15 2.264    63.0    0       1
## 645   645   59944   18 4.404    70.5    1       1
## 646   646   61951   15 2.278    60.0    0       1
## 647   647   63241   16 4.504    72.0    1       0
## 648   648   71141   17 5.638    70.0    1       0
## 649   649   71142   16 4.872    72.0    1       1
## 650   650   73041   16 4.270    67.0    1       1
## 651   651   73042   15 3.727    68.0    1       1
## 652   652   73751   18 2.853    60.0    0       0
```

```
## 653      653  75852  16 2.795   63.0   0      1
## 654      654  77151  15 3.211   66.5   0      0
```

Look at the top few rows of the data:

```
head(fev_data)
```

```
## # A tibble: 6 x 7
##    seqnbr subjid   age   fev height   sex smoke
##     <dbl>  <dbl> <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1       1    301     9  1.71     57     0     0
## 2       2    451     8  1.72   67.5     0     0
## 3       3    501     7  1.72   54.5     0     0
## 4       4    642     9  1.56     53     1     0
## 5       5    901     9  1.90     57     1     0
## 6       6   1701     8  2.34     61     0     0
```

fev_data is a tibble - this is a tidyverse structure similar to a data frame (from base R) but with some differences:

- default printing is shorter
- tells you the column types (character, double, etc.)
- doesn't change the types of inputs

Tip: if your data is in a Microsoft Excel spreadsheet, you will need a different package to read it in, such as readxl. So you'll need:

install.packages("readxl")

library(readxl)

excel_data <- read_xlsx(filename, sheet = 1) #(to read the first sheet)

–> How would you read in a text file? (Check the data import cheat sheet!) There is a text file in your dataset so you can practice: psa.txt

```
library(readr)
psa <- read_csv("psa.txt")
```

```
## Rows: 50 Columns: 1
## -- Column specification ---------------------------------------------------
## Delimiter: ","
## chr (1): ptid   nadirpsa pretxpsa  ps      bss    grade      age    obstime...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
View(psa)
```

–> Have a look at the "Useful arguments" section of the data import cheat sheet. Use a few of them when you read in fev.csv and look at the data, is that what you expected?

–> Apply the summary function to a tibble. What does it do?

**Operating on data: columns**

Individual columns are identified using the $ symbol:

```
head(fev_data$fev)
```

```
## [1] 1.708 1.724 1.720 1.558 1.895 2.336
```

```
summary(fev_data$fev)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.791   1.981   2.547   2.637   3.119   5.793
```

```
length(fev_data$fev)
```

```
## [1] 654
```

Other useful functions for tibbles and data frames:

```
names(fev_data)
```

```
## [1] "seqnbr" "subjid" "age"    "fev"    "height" "sex"    "smoke"
```

```
dim(fev_data)
```

```
## [1] 654   7
```

Other useful functions for columns:

```
max(fev_data$fev)
```

```
## [1] 5.793
```

```
mean(fev_data$fev)
```

```
## [1] 2.63678
```

```
sd(fev_data$fev)
```

```
## [1] 0.8670591
```