

## *Practical 2*

### *Jumping Rivers*

#### *A suggestion*

This is just a suggestion to try and help get you into “good habits” early. If I was sitting to take this practical now, I would start a new R script file. That way all of the work that you have done associated with today’s course is in one place, and the code for each of the practicals is separate from one another. This might feel a bit tedious right now, but as the amount of code you write and number of projects you take part in increases it will pay off to have a structured workflow.

#### *Tibbles*

For this set of questions we will use the movies data from the IMDB database.

1. Use the `head()` function to inspect the top of the data, this can help give you a feel for what the data looks like and what variables are contained within the data
2. How many films and how many variables are in this data set?
3. Recall that if I want to pull out a single column, or variable, from a data frame we can use `$`. To extract the titles from this data set we write `movies$title`.<sup>1</sup> Using `mean()` and `median()` calculate these summary statistics for the film lengths
4. What year is the oldest film in the data set from?
5. How long is the longest film?
6. What is the standard deviation of the film ratings?
7. Try running the following code

```
table(movies$action)
```

What do you think is happening?

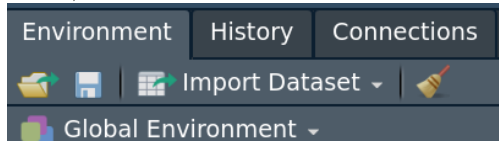
#### *Loading data from a CSV file*

To give you some practice at reading in your own data, we’re going to get you to read it in. The function

```
library("jrIntroduction")
get_csv_movies_file()
```

<sup>1</sup> If you can’t remember what the names of the columns are, you can use `colnames(movies)` to find out.

will copy a file called `movies.csv` into your current working directory. You can now import the data set using the *Import Dataset* button in RStudio, under the Environment tab



This will generate R code that you can reuse.

### *Solutions*

Solutions to the practical questions are contained within the package

```
vignette("solutions2", package = "jrIntroduction")
```