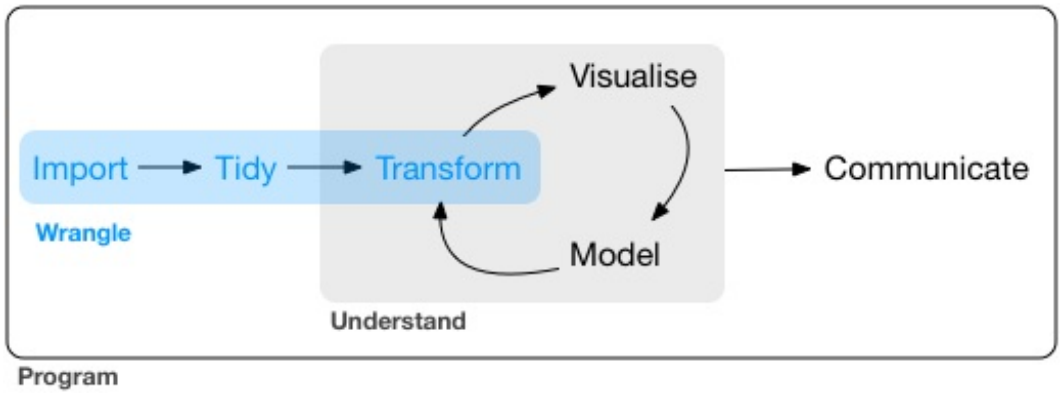
9 Introduction

In this part of the book, you'll learn about data wrangling, the art of getting your data into R in a useful form for visualisation and modelling. Data wrangling is very important: without it you can't work with your own data! There are three main parts to data wrangling:



This part of the book proceeds as follows:

- In tibbles, you'll learn about the variant of the data frame that we use in this book: the **tibble**. You'll learn what makes them different from regular data frames, and how you can construct them "by hand".

- In data import, you'll learn how to get your data from disk and into R. We'll focus on plain-text rectangular formats, but will give you pointers to packages that help with other types of data.

- In tidy data, you'll learn about tidy data, a consistent way of storing your data that makes transformation, visualisation, and modelling easier. You'll learn the underlying principles, and how to get your data into a tidy form.

Data wrangling also encompasses data transformation, which you've already learned a little about. Now we'll focus on new skills for three specific types of data you will frequently encounter in practice:

- Relational data will give you tools for working with multiple interrelated datasets.

- Strings will introduce regular expressions, a powerful tool for manipulating strings.

- Factors are how R stores categorical data. They are used when a variable

has a fixed set of possible values, or when you want to use a non-alphabetical ordering of a string.

- Dates and times will give you the key tools for working with dates and date-times.