Wickham, Hadley
ggplot2. Elegant graphics for data analysis. With contributions by Carson Sievert. 2nd edition. (English) Zbl 1397.62006
Use R!. Cham: Springer (ISBN 978-3-319-24275-0/pbk; 978-3-319-24277-4/ebook). xvi, 260 p. (2016).

The versatility and efficiency of ggplot have led to the development of ggplot2 and this book which overviews the standard use and presentation secrets of functions developed in the last 5 years. The additional practice exercises and the numerous pointers to other packages built around it, coupled with an increased amount of example code make it a must have for any ggplot aficionado. The book is structured in three parts: (i) an introductory overview of the geom concept and the toolbox, (ii) a detailed description of the grammar and (iii) hints for using ggplot2 for an efficient data analysis.

The first chapter commences with an introduction to ggplot2 and a quick overview of its grammar. The author also localizes ggplot2 within R and in relation to other R graphics. The second chapter comprises of the description of features of ggplot2 such as the aesthetic attributes e.g. colour, size, shape, the use of faceting and the concept of plot geoms. The third and final chapter in part 1 describes components of the toolbox such as the basic plot types, labels and annotations. Next, the author describes the collective geoms focussing on one aesthetic for multiple groups, organising different groups in various layers or overriding the default grouping. Also included are the display of maps and distributions, approaches to reveal uncertainty and to work with weighted data and hints for dealing with overplotting.

The second part of the book focuses on the grammar of ggplot2, especially on the components of the layered grammar, which are introduced through examples in chapter 4. In the next chapter the author presents the build of a plot layer-by-layer and describes in detail the use and particularities of aesthetic mappings and geoms and how to introduce position adjustments, if needed. In the sixth chapter the scales, axes and legends are discussed. The author not only shows how to modify the scales, add legends or set view limits, but also includes a set of guidelines for maximizing the use and effective information on plots. In the next chapter the art of positioning the components of a plot is described using faceting and coordinate systems. The last chapter of this part, chapter 8, describes themes and presents details on how to modify components and their elements: the plot itself, the axes, the legend, the panels or the faceting.

The third part of the book is allocated to data analysis. In the ninth chapter an introduction of spread and gather is presented through theoretical concepts and case studies such as the analysis of blood pressure. In the tenth chapter the author presents data transformations and their effect on the clarity of the plots. The filtering is presented first, including the handling of missing values. Next, the group-wise summaries and the statistical considerations behind them are overviewed. The eleventh chapter discusses the effect of adding modelling information to enhance the conclusions obtained from visualization studies. The author presents the effect of adding or removing a trend, of visualizing models or introducing model-level or coefficient-level summaries. The last chapter is dedicated to programming using ggplot2 single or multiple components and interacting with the plot functions or plot environment.

The book is written in an accessible manner and it is suitable for undergraduates, postgraduates and researchers with some R experience. All theoretical concepts are accompanied by code making it easy to learn by reproducing the examples. Most sections are followed by additional exercises which not only emphasise the core elements but also encourage the reader to step safely outside their comfort zone and take control of their plots. The structure of the book is suitable for most readers thanks to its lecture-style approach and ensures a thorough understanding of the basic and secret elements of ggplot2.

Reviewer: Irina Ioana Mohorianu (Oxford)

MSC: