Thesis title

A thesis submitted for the degree of

Master of Business Analytics

by

Dewi Lestari Amaliah

31251587

Department of Econometrics and Business Statistics

Monash University

Australia

September 2021

Contents

Abstract

The abstract should outline the main approach and findings of the thesis and must not be more than 500 words.

Acknowledgements

I would like to thank my pet goldfish for . . .

Chapter 1

Introduction

Start with visualisation And then survey And MRP

Chapter 2

Systematic Literature Review

This study is performed using a systematic review method in which an attempt is made to collect empirical evidence explicitly and systematically using pre-specified eligibility criteria to answer a specific research question (Green et al., 2008). Systematic literature review also enable the process of finding the gap in certain field of science, in terms of what has been done and what needs to be done (Linnenluecke, Marrone, and Singh, 2020). Hence, in this case, systematic literature review could assist us to understand the common practice in MRP visualisations and how can we improve that.

According to Brown University Library (2021), the key criteria of the systematic literature review are: *"a clearly defined question with inclusion & exclusion criteria; rigorous & systematic search of the literature; critical appraisal of included studies; data extraction and management; analysis & interpretation of results; and report for publication."* Hence, to conform with these criteria, this study incorporates the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)'s checklist and flow diagram. The following subsections discuss the steps conducted following those criteria.

2.1 Literature Identification

MRP is applied in various scientific fields, ranging from social and political science to public health. Therefore, to identify the literature, this study refers to research databases instead of field-specific journals. Those databases are JSTOR, EBSCO, and PubMed. The

first two databases are chosen due to their broad range of field coverage, while the latter is chosen since MRP is sometimes also applied in the health and medical field. Choosing these databases also considers that heterogeneity of included studies is one of the important factors in a systematic literature review (Schweizer and Nair, 2017).

Further, we identify the literature using the combination of several search terms. Mostly the search term includes the term "multilevel regression", "post-stratification", "post-stratification", and "multilevel model". Our target literature is articles that are written in English. Regarding the time of publication, we exclude all of the publications before 1997 since the MRP method has not been developed in this period. Initially we included only the title/abstract when searching these databases. However, using this method limits the set of potential articles to only include those with the search term in the abstract/title. To rectify this, we also include a search with "all field" in the search criteria. Note that for EBSCO, we directly apply the search for all fields. The detailed literature identification is shown in Table 2.1.

The total number of articles from this search criteria are 327. Next, we utilize the literature manager, EndNote X9, to manage these articles and to find duplicate articles. After removing those duplicate articles, we have 212 articles to be screened in the next stage.

Table 2.1: *Detail of literature identification*

Database	Search Terms	Search Field	Inclusion	Exclusion	Number Returned
JSTOR	(multilevel regression and poststratification) OR ("post-stratification")	Abstract	Article, content I can access, English	anything before 1997	44
JSTOR	(("multilevel regression" AND ("post-stratification" OR Poststratification)) OR ("multilevel model" AND ("post-stratification" OR Poststratification)))	All field	Article, English	anything before 1997	142
EBSCO	"multilevel regression with post-stratification" OR "multilevel regression with poststratification" OR "multilevel regression and Poststratification" OR "multilevel regression and Post-stratification"	All field	Academic (Peer-Reviewed) Journals, English	anything before 1997	42
EBSCO	(multilevel regression AND post-stratification) OR (multilevel model AND post-stratification) OR (multilevel regression AND poststratification) OR (multilevel model AND poststratification)	All field	Academic (Peer-Reviewed) Journals, English	anything before 1997	45
PubMed	"multilevel regression with post-stratification" OR "multilevel regression with poststratification" OR "multilevel regression and Poststratification" OR "multilevel regression and Post-stratification"	Title/Abstract	Article, English	anything before 1997	26
PubMed	(multilevel regression AND post-stratification) OR (multilevel model AND post-stratification) OR (multilevel regression AND poststratification) OR (multilevel model AND poststratification)	All field	Article, English	anything before 1997	28

2.2 Screening and Eligibility Criteria

We screen all of the articles whether they fit the criteria to be included in the study or not. We find that 3 articles are apparently not research papers. Hence, there are 209 abstracts to be screened. To screen efficiently, we use two stages as follow.

2.2.1 Stage 1: Review of abstracts

In the first stage DA and LK independently review all articles abstracts with the following eligibility criteria:

1. The abstract should mentions analysis of data or creation of simulation data.
2. The abstract should mentions use of MRP or multilevel models to make population estimates or the use of other regression models (BART, spatial, stacked, trees) to make population estimates.

During the screening, we agreed that 61 articles meet the eligibility criteria listed above, while 104 articles do not meet the criteria. Further, we disagreed on 44 articles. Accordingly, DA and LK skim the full manuscript to decide whether the paper could be included in the next stage or not. As the result, we include 22 more articles to be reviewed in stage 2.

2.2.2 Stage 2: Full manuscript review

DA reviews the full manuscript on 83 articles meet a second, more specific criteria. The aim of this stage is to get the list of the final articles that would be included in the study. We set the criteria of inclusion as follow:

1. It should apply MRP as its method.
2. It should contain at least one plot relate to MRP findings.

During this stage, we exclude 4 articles as they do not meet the first criteria. Further, 7 articles are excluded as they do not meet the second criteria. Also, an article is not included because it is a duplicate from other article that is already in the eligible list. However, this duplicate was not detected automatically by Endnote X9. Finally, we have 71 articles to be
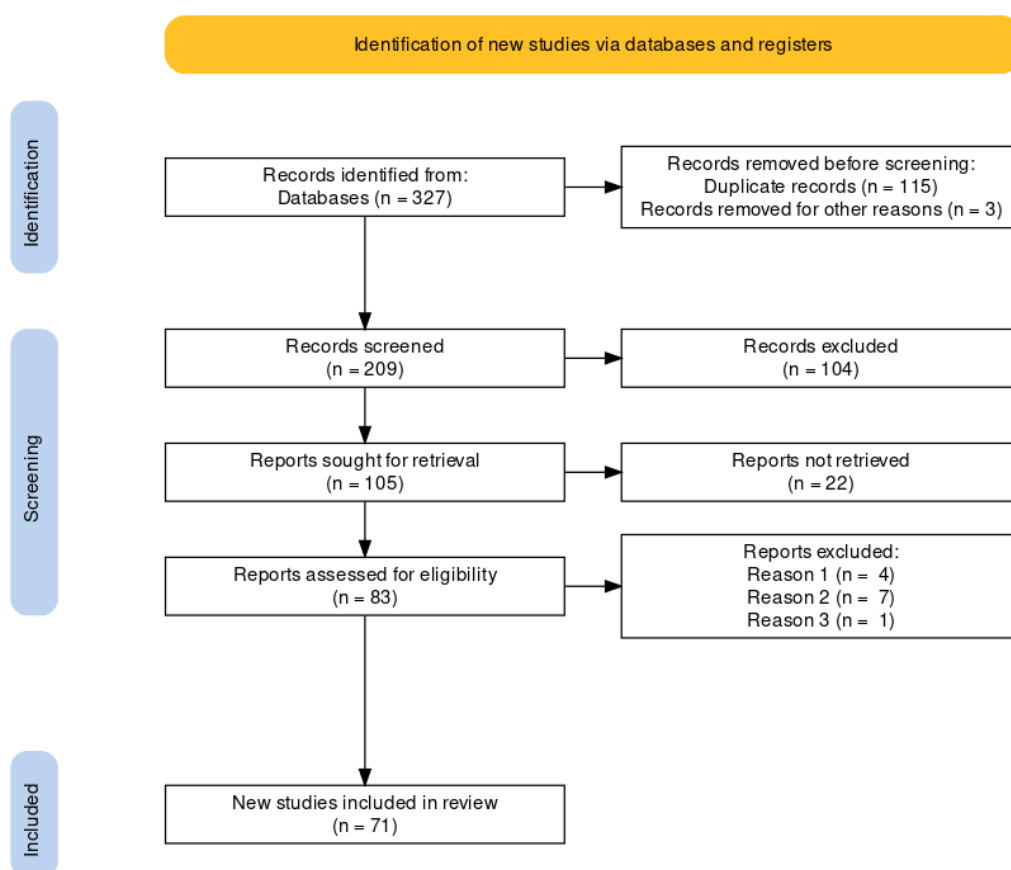
Figure 2.1: *PRISMA flow chart of this systematic literature review.*

reviewed in the next stage. Figure 2.1 displays the PRISMA flow chart of this study. This figure is generated using PRISMA2020 (Haddaway, Pritchard, and McGuinness, 2021).

2.3 Data Extraction and Analysis

We focus the data extraction on the MRP-related plot. We manually create a metadata for each plot (included in the supplementary material). We will use this metadata to analyse the current reporting practices with MRP. Besides, this metadata ensures the reproducibility of the analysis and to maintain the transparency of the systematic literature review process.

We code the plots according to their type, i.e., communication (coded to 0) and diagnostic plot (coded to 1). For diagnostic plots, we examine whether the plots compare MRP with other estimates, which are:

1. Raw (direct estimates or direct disaggregation);

2. Ground truth;

3. Weight estimation;

4. Estimates of other MRP models, for example, the paper build several MRP models from various simulation scenarios or using different covariates;

5. Estimates from another study/survey;

6. Estimates from another method, for example comparing MRP with Bayesian Additive Tress with Post-Stratification.

Plots that show a comparison of MRP with the above list would be coded to 1, otherwise coded to 0. The diagnostic plot could also display the performance of MRP using these performance criteria:

1. Bias;

2. Mean Absolute Error (MAE);

3. Mean Square Error (MSE)/ Relative Mean Square Error (RMSE);

4. Standard Error (SE);

5. Correlation.

Just like the comparison, the MRP-related plot would be reviewed whether it is contains each of those metrics (coded 1) or not (coded 0).

We also review other features of the plot using the grammar in `ggplot2` (Wickham, 2016). The common grammar used in practice allows us to understand to what extend MRP models are effectively visualised. It is worth noting that there is no specific convention or well-documented recommendation on how data should be visualised as building a graph more often involve choice or preference (Midway, 2020). For example, there is no specific convention on which variable should be put on the x and y-axis in a scatter plot, even though it has been common knowledge to put the response variable on the y-axis and the explanatory variable on the x-axis. Hence, grammar assists us in evaluating well-formed graphics (Wickham, 2010). In addition, Vanderplas, Cook, and Hofmann (2020) mention that classifying and comparing graphs according to their grammar is more robust and more elegant.

Accordingly, we examine the facet, geom, axis, color, and shape. Besides, the metadata also contains the article's author/s, article's year, article's title, and figure's number as appearing in the article. After the extraction, we analyze the data using graphical visualization with `ggplot2`. The result will be discussed in the following subsection.

2.4 Common practices in MRP visualisations

From 71 articles, we extract the data of 243 plots. 46.91 % of these plots are diagnostics plots, while the remaining are communication plots.

2.4.1 Performance metrics used in MRP

2.4.2 Common estimates to compare with MRP estimates

2.4.3 Common grammar in MRP visualisations

Plot type

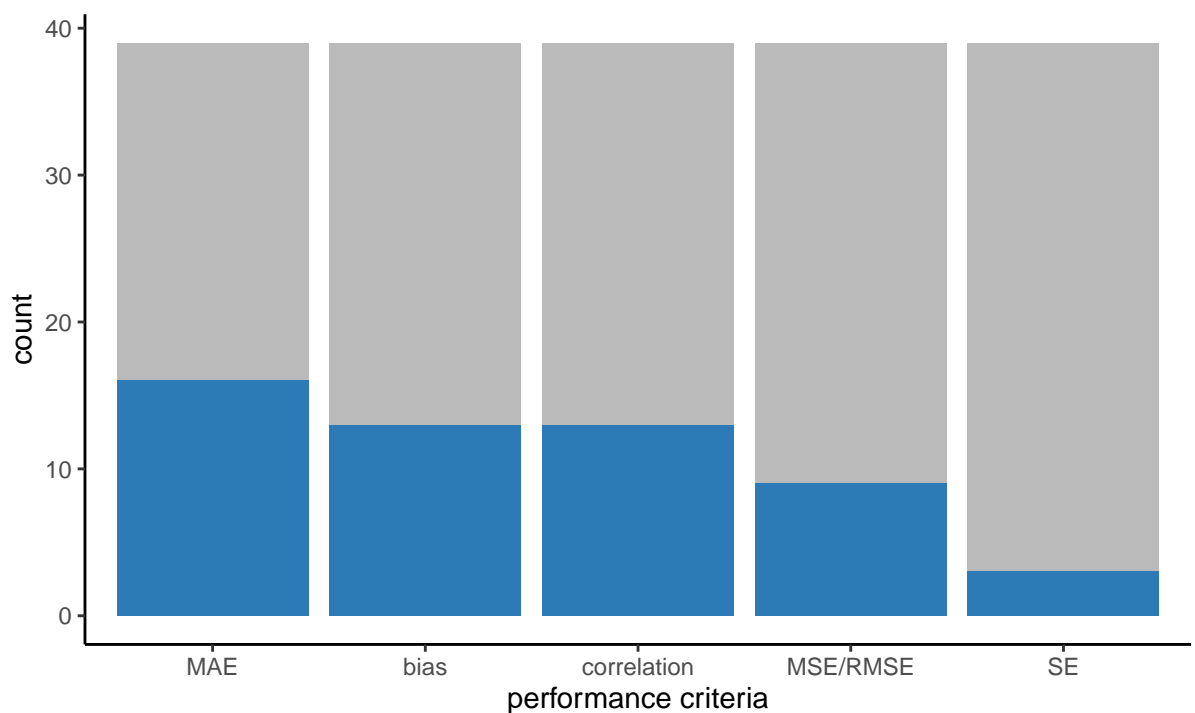Plot type, referred as `geom` in grammar of graphics represents

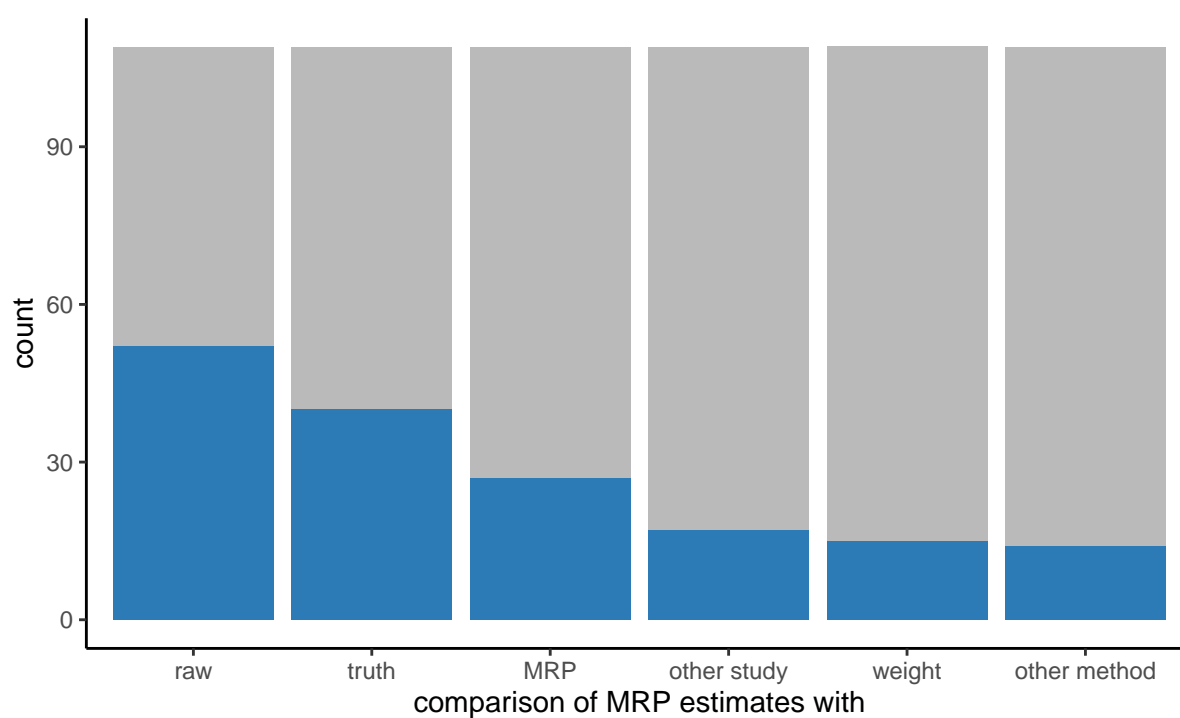**Figure 2.2:** *Metrics used as MRP performance criteria.*

Figure 2.3: *Estimates to compare with MRP.*

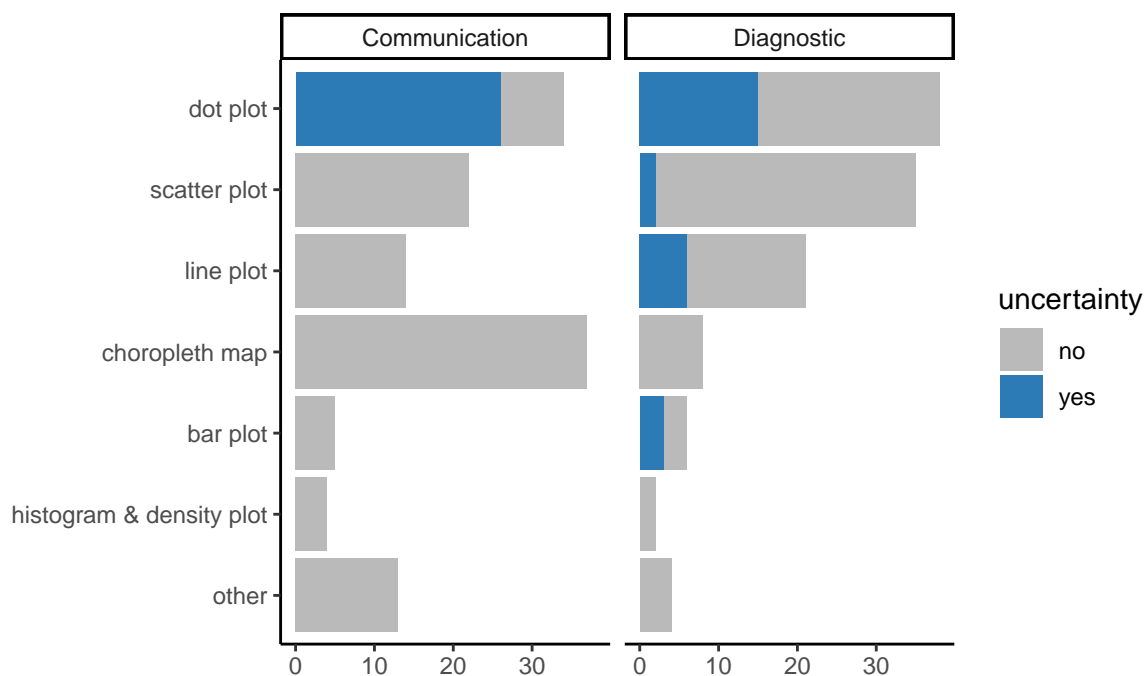**Figure 2.4:** *Common plot types used in MRP visualisations. Both communication and diagnostics plots rarely displayed uncertainty.*
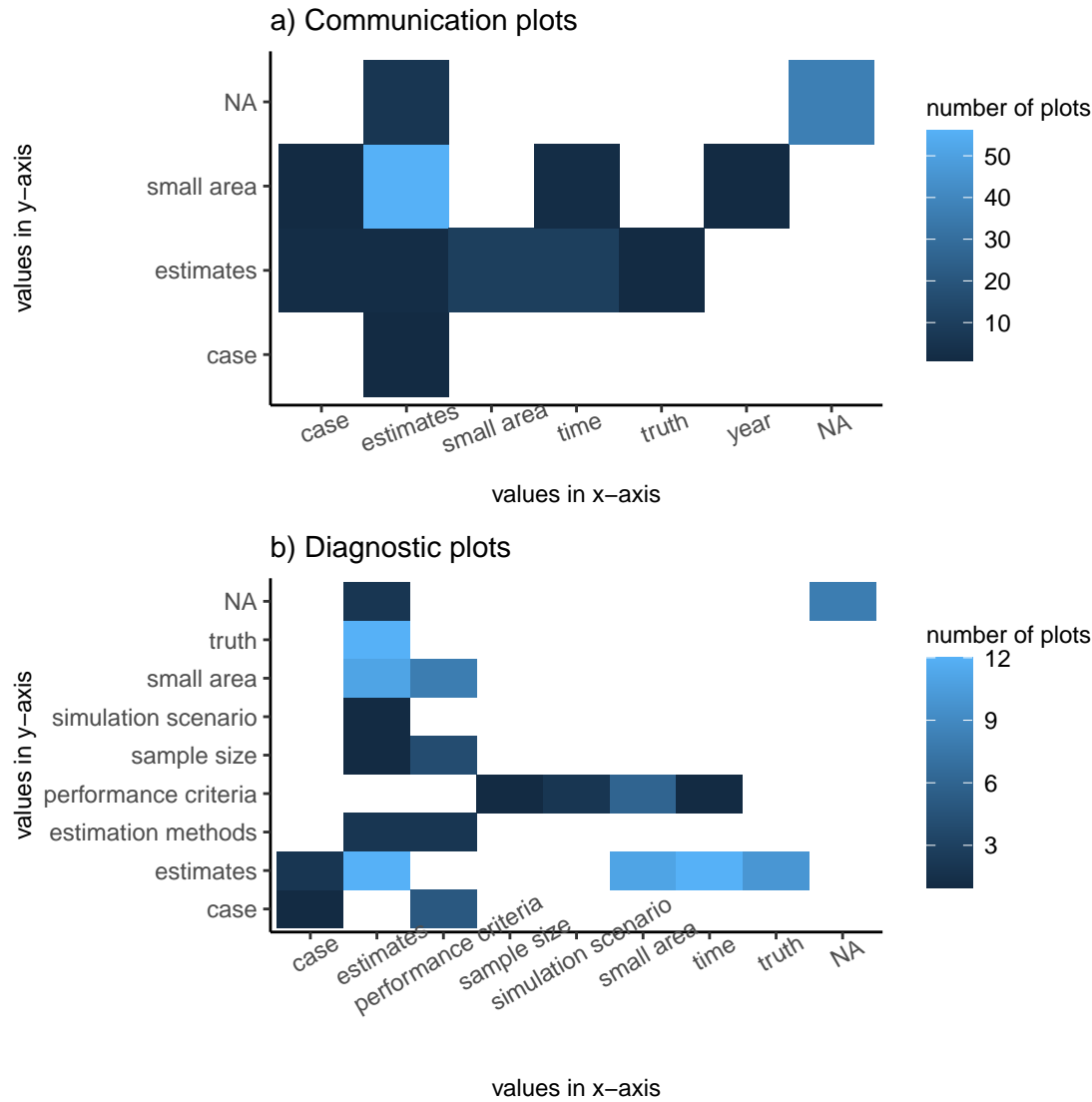
Figure 2.5: *Common values put in plots' axis. Axis in diagnostic plots more varied compared to communication plot.*

Values put in x and y-axis

Feature used

Appendix A

Additional stuff

You might put some computer output here, or maybe additional tables.

Note that line 5 must appear before your first appendix. But other appendices can just start like any other chapter.

Bibliography

Brown University Library (2021). *Scientific Literature Review Resources and Services*. `https://libguides.brown.edu/Reviews/types`.

Green, S, JP Higgins, P Alderson, M Clarke, CD Mulrow, and AD Oxman (2008). "Introduction". In: *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons, Ltd. Chap. 1, pp. 1–9. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470712184.ch1`. `https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470712184.ch1`.

Haddaway, NR, CC Pritchard, and LA McGuinness (2021). *PRISMA2020: R package and ShinyApp for producing PRISMA 2020 compliant flow diagrams (Version 0.0.2)*.

Linnenluecke, MK, M Marrone, and AK Singh (2020). Conducting systematic literature reviews and bibliometric analyses. eng. *Australian journal of management* **45**(2), 175–194.

Midway, SR (2020). Principles of Effective Data Visualization. *Patterns* **1**(9), 100141.

Schweizer, ML and R Nair (2017). A practical guide to systematic literature reviews and meta-analyses in infection prevention: Planning, challenges, and execution. eng. *American journal of infection control* **45**(11), 1292–1294.

Vanderplas, S, D Cook, and H Hofmann (2020). Testing Statistical Charts: What Makes a Good Graph? *Annual Review of Statistics and Its Application* **7**(1), 61–88.

Wickham, H (2010). A layered grammar of graphics. *Journal of Computational and Graphical Statistics* **19**(1), 3–28.

Wickham, H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. `https://ggplot2.tidyverse.org`.