Thesis title

A thesis submitted for the degree of

Doctor of Philosophy

by

Jane Doe

B.Sc. (Hons), University of Sydney

School of Mathematics and Statistics

University of Sydney

Australia

July 2018

Contents

Acknowledgements

Generally anyone that helped you in your research journey should be acknowledged here.

Large credit of this thesis template goes to Rob J Hyndman. His original template can be found at his [github repositary](.).

Statement of Originality

This is to certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

This statement is not needed for Honours Thesis.

Jane Doe

Preface

The material in Chapter 1 has been submitted to the journal *Journal of Impossible Results* for possible publication.

The contribution in Chapter 2 of this thesis was presented in the International Symposium on Nonsense held in Dublin, Ireland, in July 2015.

This section is not necessary for Honours Thesis.

Abstract

The abstract is a summary of the whole thesis. Typically this section would have about 250-350 words. Check the rules of your School or University.

Chapter 1

Introduction

This is where you introduce the main ideas of your thesis, and an overview of the context and background.

Later chapters should be divided into coherent pieces describing your analysis. The final chapter should provide some concluding remarks, discussion, ideas for future research, and so on. Appendixes can contain additional material that don't fit into any chapters, but that you want to put on record. For example, additional tables, output, etc.

1.1 Rmarkdown

In this template, the rest of the chapter shows how to use Rmarkdown. The big advantage of using Rmarkdown is that it allows you to include your R code directly into your thesis, to ensure there are no errors in copying and pasting, and that everything is reproducible. It also helps you stay better organized.

For details on using *R Markdown* see http://rmarkdown.rstudio.com.

1.2 Data

Included in this template is a file called `sales.csv`. This contains quarterly data on Sales and Advertising budget for a small company over the period 1981–2005. It also contains

the GDP (gross domestic product) over the same period. All series have been adjusted for inflation. We can load in this data set using the following command:

```r
sales <- ts(read.csv("data/sales.csv")[,-1], start=1981, frequency=4)
```

Any data you use in your thesis can go into the data directory. The data should be in exactly the format you obtained it. Do no editing or manipulation of the data outside of R. Any data munging should be scripted in R and form part of your thesis files (possibly hidden in the output).

1.3 Figures

Figure 1.1 shows time plots of the data we just loaded. Notice how figure captions and references work. Chunk names can be used as figure labels with `fig:` prefixed. Never manually type figure numbers, as they can change when you add or delete figures. This way, the figure numbering is always correct.
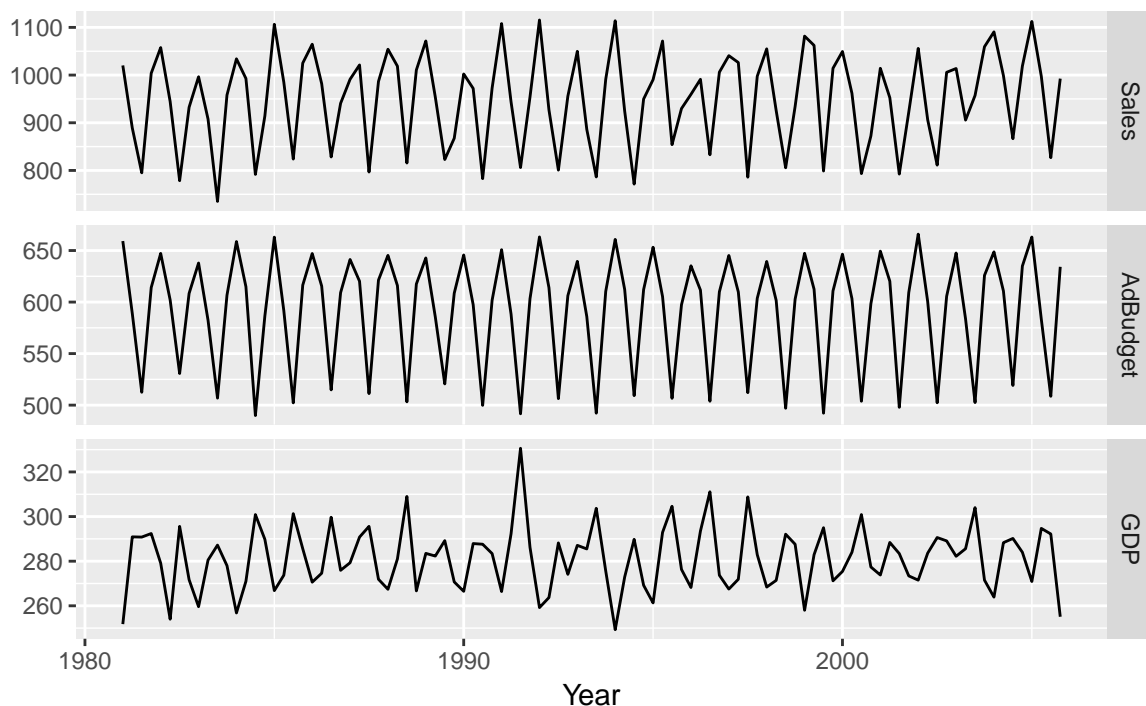

Figure 1.1: *Quarterly sales, advertising and GDP data.*

1.4 Results from analyses

We can fit a dynamic regression model to the sales data.

If y_t denotes the sales in quarter t, x_t denotes the corresponding advertising budget and z_t denotes the GDP, then the resulting model is:

$$y_t - y_{t-4} = \beta(x_t - x_{t-4}) + \gamma(z_t - z_{t-4}) + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-4} + \varepsilon_t \tag{1.1}$$

where $\beta = 2.29$, $\gamma = 0.99$, $\theta_1 = 0.20$, and $\Theta_1 = -0.90$.

1.5 Tables

Let's assume future advertising spend and GDP are at the current levels. Then forecasts for the next year are given in Table 1.1.

Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
1000.6	948.0	1053.3	920.2	1081.1
1013.3	959.6	1066.9	931.2	1095.3
1076.8	1023.1	1130.4	994.7	1158.8
1003.3	949.6	1057.0	921.2	1085.4

Table 1.1: *Forecasts for the next year assuming Advertising budget and GDP are unchanged.*

Again, notice the use of labels and references to automatically generate table numbers. In this case, we need to generate the label ourselves.

The `knitLatex` package is useful for generating tables from R output. Other packages can do similar things including the `kable` function in `knitr` which is somewhat simpler but you have less control over the result. If you use `knitLatex` to generate tables, don't forget to include `results="asis"` in the chunk settings.

Chapter 2

Literature Review

This chapter contains a summary of the context in which your research is set.

Imagine you are writing for your fellow PhD students. Topics that are well-known to them do not have to be included here. But things that they may not know about should be included.

Resist the temptation to discuss everything you've read in the last few years. And you are not writing a textbook either. This chapter is meant to provide the background necessary to understand the material in subsequent chapters. Stick to that.

You will need to organize the literature review around themes, and within each theme provide a story explaining the development of ideas to date. In each theme, you should get to the point where your ideas will fit in. But leave your ideas to later chapters. This way it is clear what has been done beforehand, and what new contributions you are making to the research field.

All citations should be done using markdown notation as shown below. This way, your bibliography will be compiled automatically and correctly.

2.1 Exponential Smoothing

Exponential smoothing was originally developed in the late 1950s (Brown, 1959, 1963; Holt, 1957; Winters, 1960). Because of their computational simplicity and interpretability, they became widely used in practice.

Empirical studies by Makridakis and Hibon (1979) and Makridakis et al. (1982) found little difference in forecast accuracy between exponential smoothing and ARIMA models. This made the family of exponential smoothing procedures an attractive proposition (see Chatfield et al., 2001).

The methods were less popular in academic circles until Ord, Koehler, and Snyder (1997) introduced a state space formulation of some of the methods, which was extended in Hyndman et al. (2002) to cover the full range of exponential smoothing methods.

Appendix A

Additional stuff

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

Note that \appendix must appear before your first appendix although this is not required for other appendices.

Appendix B

Linear Mixed Model

The extension to the linear models is the linear mixed models. A general form for the linear mixed model is given by

$$y = X\tau + Zu + e \tag{B.1}$$

where y is the $n \times 1$ vector of observations on the trait (or response) of interest, τ is a $p \times 1$ vector of fixed effects, u is a $q \times 1$ vector of random effects, X and Z are associated (known) design matrices specifying the factors and covariates (explanatory variables) with corresponding fixed and random effects vector respectively, and e is the $n \times 1$ vector of errors. We assume that rank of X is $p_0 < p$ (i.e. non-full rank).

To complete the specification, we assume that the joint distribution of (u, e) is

$$\begin{bmatrix} u \\ e \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \right)$$

where $G(\kappa_G)$ and $R(\kappa_R)$ are covariance matrices which depend on vectors κ_G and κ_R respectively; and θ is the global scale parameter. We let $\kappa = (\kappa_G^\top, \kappa_R^\top)^\top$ denote the complete vectors of variance parameters.

It follows that $y \sim N(X\tau, V)$ where $V = ZGZ^\top + R$ where $V = V(\kappa)$.

Bibliography

Brown, RG (1959). *Statistical forecasting for inventory control*. McGraw-Hill, New York.

Brown, RG (1963). *Smoothing, forecasting and prediction of discrete time series*. Englewood Cliffs, New Jersey: Prentice Hall.

Chatfield, C, AB Koehler, JK Ord, and RD Snyder (2001). A new look at models for exponential smoothing. *The Statistician* **50**, 147–159.

Holt, CE (1957). *Forecasting trends and seasonals by exponentially weighted averages*. O.N.R. Memorandum 52/1957. Carnegie Institute of Technology.

Hyndman, RJ, AB Koehler, RD Snyder, and S Grose (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* **18**(3), 439–454.

Makridakis, S, A Anderson, R Carbone, R Fildes, M Hibon, RLJ Newton, E Parzen, and R Winkler (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting* **1**, 111–153.

Makridakis, S and M Hibon (1979). Accuracy of forecasting: an empirical investigation (with discussion). *Journal of Royal Statistical Society (A)* **142**, 97–145.

Ord, JK, AB Koehler, and RD Snyder (1997). Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of American Statistical Association* **92**, 1621–1629.

Winters, PR (1960). Forecasting sales by exponentially weighted moving averages. *Management Science* **6**, 324–342.