Example of Output of R Markdown

No Name

R includes some features that might simplify the writing process by making it possible to gather data analysis and writing in a single document. This makes it easier to reproduce ones analysis in case the data is changed and to produce presentable papers for self-publishing. This presentation gives a brief example of some of these features.

# Introduction

There is a number of ways to incorporate R code and ordinary text into a single coherent paper. The most popular way is to write code and text in the Markdown format. Markdown is a mark-up language, meaning that code is written in an ordinary text file and then compiled into a desirable format, for example a pdf file or a Word document. This is a short example of the R Markdown features available in R using the packages rmarkdown and knitr. It gives examples of the following:

* Lists
* Figures
* Tables
* References

The data oats from the MASS package is used as example data.  
We load the dataset with the following code. In this document the R code is shown, but it is easy to hide the code and only show the results.

dat <- MASS::oats

# Example output

## Descriptive statistics

The data can be presented in table form, summarized for each variety and nitrogen combination. The following code summarizes the dataset and presents it as table using the kable function from the knitr package.

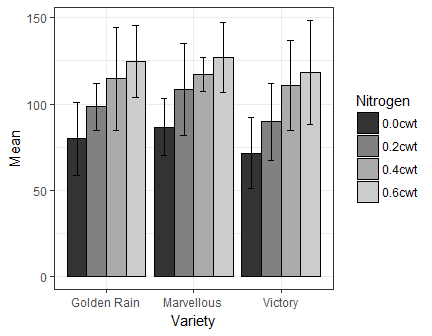
dat\_agg <- aggregate(Y ~ V + N, dat, function(x) c(mean(x), sd(x)))  
dat\_agg[,1] <- as.character(dat\_agg[,1])  
dat\_agg <- data.frame(dat\_agg[, 1:2], dat\_agg[, 3])  
names(dat\_agg) <- c("Variety", "Nitrogen", "Mean", "Std.dev")  
dat\_agg[dat\_agg$Variety == "Golden.rain", 1] <- "Golden Rain"  
dat\_agg <- dat\_agg[order(dat\_agg$Variety), ]  
  
knitr::kable(dat\_agg, digits = 2,   
 caption = "Mean and standard deviation of yield   
 per variety and nitrogen combination.",  
 row.names = F)

Mean and standard deviation of yield per variety and nitrogen combination.

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Nitrogen | Mean | Std.dev |
| Golden Rain | 0.0cwt | 80.00 | 21.00 |
| Golden Rain | 0.2cwt | 98.50 | 13.47 |
| Golden Rain | 0.4cwt | 114.67 | 29.94 |
| Golden Rain | 0.6cwt | 124.83 | 20.88 |
| Marvellous | 0.0cwt | 86.67 | 16.57 |
| Marvellous | 0.2cwt | 108.50 | 26.85 |
| Marvellous | 0.4cwt | 117.17 | 9.79 |
| Marvellous | 0.6cwt | 126.83 | 20.29 |
| Victory | 0.0cwt | 71.50 | 20.60 |
| Victory | 0.2cwt | 89.67 | 22.51 |
| Victory | 0.4cwt | 110.83 | 26.01 |
| Victory | 0.6cwt | 118.50 | 30.09 |

This can be visualized in a bar chart. The ggplot2 package is used to construct the plot.

library(ggplot2)  
barchart <- ggplot(dat\_agg, aes(Variety, Mean, fill = Nitrogen)) +   
 geom\_bar(stat = "identity", position = position\_dodge(),   
 col = "black") +  
 geom\_errorbar(aes(ymin = Mean - Std.dev,  
 ymax = Mean + Std.dev),  
 position = position\_dodge(0.9),  
 width = 0.3) +  
 theme\_bw() +  
 scale\_fill\_grey()  
print(barchart)



Mean of yield per variety and nitrogen combination. Error bars give standard deviation.

## Anova model of split-plot experiment

The data is analyzed as a mixed model using variety and nitrogen level as fixed factors, and block and interaction between block and variety as random factors. The results are given below. The model specifies a significant effect of nitrogen level, but not of variety. Once again, the kable function prints the table in a presentable form.

library(lme4, quietly = T)  
library(lmerTest, quietly = T)  
  
names(dat) <- c("Block", "Variety", "Nitrogen", "Yield")  
mod <- lmer(Yield ~ Variety + Nitrogen + (1 | Block / Variety), dat)  
  
knitr::kable(anovaTable <- anova(mod), digits = 2,   
 caption = "Anova results of analysis of yield.",  
 row.names = T)

Anova results of analysis of yield.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Sum Sq | Mean Sq | NumDF | DenDF | F.value | Pr(>F) |
| Variety | 482.91 | 241.46 | 2 | 10 | 1.49 | 0.27 |
| Nitrogen | 20020.50 | 6673.50 | 3 | 51 | 41.05 | 0.00 |

The estimated mean values of the nitrogen levels are calculated and tested for differences.

library(emmeans, quietly = T)  
  
knitr::kable(emmeansTable <- cld(emmeans(mod, "Nitrogen")),   
 digits = 2,  
 caption = "Comparison between levels of nitrogen.")

Comparison between levels of nitrogen.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nitrogen | emmean | SE | df | lower.CL | upper.CL | .group |
| 0.0cwt | 79.39 | 7.13 | 6.64 | 62.34 | 96.44 | 1 |
| 0.2cwt | 98.89 | 7.13 | 6.64 | 81.84 | 115.94 | 2 |
| 0.4cwt | 114.22 | 7.13 | 6.64 | 97.17 | 131.28 | 3 |
| 0.6cwt | 123.39 | 7.13 | 6.64 | 106.34 | 140.44 | 3 |

These tables could obviously need a little more cleaning up.

## A reference

The Markdown format can handle references in a simple way. For example, the anova model was estimated with the lmer function from the package lme4 (Bates et al. 2015).

The code to construct the reference is simply [@lme4] where the part in square brackets is linked to the full reference, which is printed in the last part of the document. In this case the reference is specified in the introduction of the .Rmd file, but references can also be stored in a separate file and handled using special software.

# References

Bates, Douglas, Martin Machler, Ben Bolker, and Steve Walker. 2015. “Fitting Linear Mixed-Effects Models Using Lme4.” *Journal of Statistical Software* 67 (1):1–48.