

NHISS Example

Contents

1	Baseline characteristic table	5
1.1	Baseline characteristic table	6
1.2	Baseline characteristic table_total	7
2	Multiple imputation	9
2.1	The number of missing values	9
2.2	Imputation for missing values	10
3	Cross-references	13
3.1	Chapters and sub-chapters	13
3.2	Captioned figures and tables	13
4	Parts	17
5	Footnotes and citations	19
5.1	Footnotes	19
5.2	Citations	19
6	Blocks	21
6.1	Equations	21
6.2	Theorems and proofs	21
6.3	Callout blocks	21

7	Sharing your book	23
7.1	Publishing	23
7.2	404 pages	23
7.3	Metadata for sharing	23

Chapter 1

Baseline characteristic table

Baseline tables show the characteristics of research subjects included in a study. A table characterizing baseline characteristics is so important that it's typically the first table that appears in any observational epidemiology (or clinical trial) manuscript, so it's commonly referred to as a "Table 1". The "Table 1" contain information about the mean and standard deviation(or median and IQR) for continue/scale variable, and proportion for categorical variable.

Baseline characteristic table should be created before imputaion, matching, or weighting.

Using data **final_db**
Outcome variable : **HTN**
Follow-up period : **DATEDIFF**
Exposure variable : **DM**
Covariates : **Age, Sex, SES, Region, BMI, CCI, Comorbidities(Dyslipidemia, Ischemic heart disease)**

```
## load library
library(moonBook)
library(dplyr)
```

```
## load data
final_db <- read.csv('Data/final_db.csv', header=T)
```

```
## formula
formula.bc <- formula(DM ~ HTN + DATEDIFF + AGE + SEX + SES + REGION + BMI + CCI + DYS + IHD)
```



```
##      - 0          2340 (99.3%)          116 (98.3%)
##      - 1           16 ( 0.7%)           2 ( 1.7%)
## -----
```

1.2 Baseline characteristic table_total

```
tot1 <- final_db %>% mutate(tmp=1)
tot2 <- final_db %>% mutate(tmp=2)
tot3 <- rbind(tot1,tot2)
```

```
mytable(tmp ~ HTN + DATEDIFF + AGE + SEX + SES + REGION + BMI + CCI + DYS + IHD, data=tot3, method="glm")
```

```
##
##      Descriptive Statistics by 'tmp'
## -----
##              1              2              p
##              (N=2474)      (N=2474)
## -----
## HTN                                     1.000
##   - 0          2284 (92.3%)          2284 (92.3%)
##   - 1           190 ( 7.7%)           190 ( 7.7%)
## DATEDIFF 1656.0 [811.0;2458.0] 1656.0 [811.0;2458.0] 1.000
## AGE      36.0 [22.0;50.0]      36.0 [22.0;50.0] 1.000
## SEX
##   - 1          1240 (50.1%)          1240 (50.1%)
##   - 2          1234 (49.9%)          1234 (49.9%)
## SES
##   - 1           697 (29.3%)          697 (29.3%)
##   - 2           743 (31.3%)          743 (31.3%)
##   - 3           935 (39.4%)          935 (39.4%)
## REGION
##   - 1          1218 (49.5%)          1218 (49.5%)
##   - 2           514 (20.9%)          514 (20.9%)
##   - 3           729 (29.6%)          729 (29.6%)
## BMI      23.2 [21.0;25.3]      23.2 [21.0;25.3] 1.000
## CCI
##   - 0          1890 (76.4%)          1890 (76.4%)
##   - 1           456 (18.4%)          456 (18.4%)
##   - 2           128 ( 5.2%)          128 ( 5.2%)
## DYS
##   - 0          2385 (96.4%)          2385 (96.4%)
##   - 1           89 ( 3.6%)           89 ( 3.6%)
## IHD
##                                     1.000
```

##	- 0	2456 (99.3%)	2456 (99.3%)
##	- 1	18 (0.7%)	18 (0.7%)
##	-----		

Chapter 2

Multiple imputation

Multiple imputation is a general approach to the problem of missing data. It aims to allow for the uncertainty about the missing data by creating several different plausible imputed data sets and appropriately combining results obtained from each of them.

Multiple imputation using chained equations (MICE) were performed to generate 10 imputed datasets. For the imputation model, predictive mean matching was used for continuous data and logistic regression was used for binary data.

Using data **final_db**
Outcome variable : **HTN**
Follow-up period : **DATEDIFF**
Exposure variable : **DM**
Covariates : **Age, Sex, SES, Region, BMI, CCI, Comorbidities(Dyslipidemia, Ischemic heart disease)**

```
## load library
library(mice)
library(dplyr)

## load data
final_db <- read.csv('Data/final_db.csv', header=T)
```

2.1 The number of missing values

```
na_count <- function(data){
  num.na <- colSums(is.na(data))
}
```

```

per.na <- paste0(round(colSums(is.na(data))/nrow(data) *100,2),"%")

return(data.frame(missing=paste0(num.na,"(",per.na,")"),row.names = names(num.na)))
}

na_count(final_db)

```

```

##                missing
## RN_INDI          0(0%)
## DM               0(0%)
## INDEX_DT         0(0%)
## HTN              0(0%)
## FU_DT            0(0%)
## AGE              0(0%)
## SEX              0(0%)
## SES              99(4%)
## REGION           13(0.53%)
## BMI             1565(63.26%)
## CCI              0(0%)
## DYS              0(0%)
## IHD              0(0%)
## DATEDIFF         0(0%)

```

- Use **mice()** function in **mice** package to deal with missing data.
 - m=10 refers to the number of imputed datasets. Five is the default value.
 - Extract imputed data sets using **compleate()** function

2.2 Imputation for missing values

```

## Exclude subject ID, index date before imputation
dat_mice <- final_db %>% select(-RN_INDI, -INDEX_DT, -FU_DT)
dat_imp <- mice(dat_mice, m=10, seed=1)

```

```

##
## iter imp variable
## 1 1 SES REGION BMI
## 1 2 SES REGION BMI
## 1 3 SES REGION BMI
## 1 4 SES REGION BMI
## 1 5 SES REGION BMI

```

```
## 1 6 SES REGION BMI
## 1 7 SES REGION BMI
## 1 8 SES REGION BMI
## 1 9 SES REGION BMI
## 1 10 SES REGION BMI
## 2 1 SES REGION BMI
## 2 2 SES REGION BMI
## 2 3 SES REGION BMI
## 2 4 SES REGION BMI
## 2 5 SES REGION BMI
## 2 6 SES REGION BMI
## 2 7 SES REGION BMI
## 2 8 SES REGION BMI
## 2 9 SES REGION BMI
## 2 10 SES REGION BMI
## 3 1 SES REGION BMI
## 3 2 SES REGION BMI
## 3 3 SES REGION BMI
## 3 4 SES REGION BMI
## 3 5 SES REGION BMI
## 3 6 SES REGION BMI
## 3 7 SES REGION BMI
## 3 8 SES REGION BMI
## 3 9 SES REGION BMI
## 3 10 SES REGION BMI
## 4 1 SES REGION BMI
## 4 2 SES REGION BMI
## 4 3 SES REGION BMI
## 4 4 SES REGION BMI
## 4 5 SES REGION BMI
## 4 6 SES REGION BMI
## 4 7 SES REGION BMI
## 4 8 SES REGION BMI
## 4 9 SES REGION BMI
## 4 10 SES REGION BMI
## 5 1 SES REGION BMI
## 5 2 SES REGION BMI
## 5 3 SES REGION BMI
## 5 4 SES REGION BMI
## 5 5 SES REGION BMI
## 5 6 SES REGION BMI
## 5 7 SES REGION BMI
## 5 8 SES REGION BMI
## 5 9 SES REGION BMI
## 5 10 SES REGION BMI
```

```
## Create 10 imputed data
for (i in 1:dat_imp$m){
  z <- assign(paste0('dat_imp',i),complete(dat_imp,i))
  assign(paste0('dat_imp',i),cbind(z,final_db %>% select(RN_INDI)))
}

## list of 10 imputed data
dat_imp_list <- list(dat_imp1,dat_imp2,dat_imp3,dat_imp4,dat_imp5,dat_imp6,dat_imp7,dat_imp8,dat_imp9,dat_imp10)

## Save multiple imputation result
save(dat_imp,file="Data/dat_imp.RData")
## Save list for imputed data
save(dat_imp_list,file="Data/dat_imp_list.RData")
```

Chapter 3

Cross-references

Cross-references make it easier for your readers to find and link to elements in your book.

3.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
 - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world = # Hello world {#hello-world}`.
 - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter 3.
 - If you prefer text as the link instead of a numbered reference use: any text you want can go here.

3.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure 3.1.

```
par(mar = c(4, 4, .1, .1))  
plot(pressure, type = 'b', pch = 19)
```

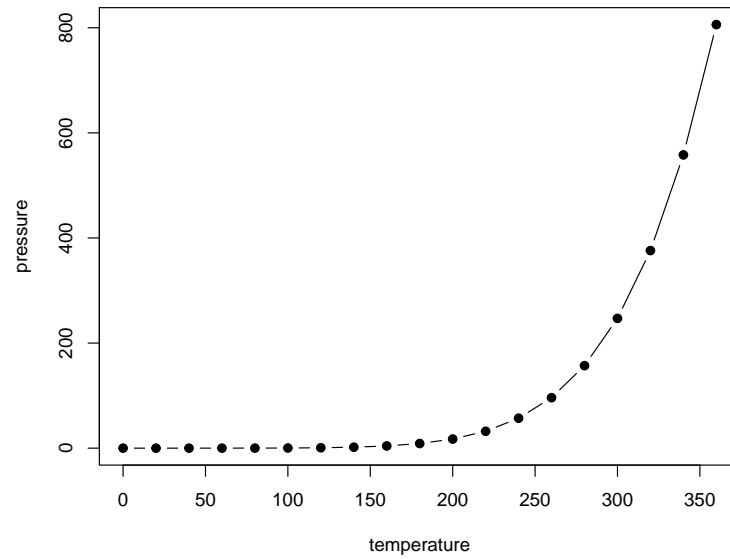


Figure 3.1: Here is a nice figure!

Don't miss Table 3.1.

```
knitr::kable(  
  head(pressure, 10), caption = 'Here is a nice table!',  
  booktabs = TRUE  
)
```

Table 3.1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000

Chapter 4

Parts

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

Chapter 5

Footnotes and citations

5.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one ¹.

5.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package [Xie, 2023] (check out the last code chunk in `index.Rmd` to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** [Xie, 2015] (this citation was added manually in an external file `book.bib`). Note that the `.bib` files need to be listed in the `index.Rmd` with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: <https://rstudio.github.io/visual-markdown-editing/#/citations>

¹This is a footnote.

Chapter 6

Blocks

6.1 Equations

Here is an equation.

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (6.1)$$

You may refer to using `\@ref{eq:binom}`, like see Equation (6.1).

6.2 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref{thm:tri}`, for example, check out this smart theorem 6.1.

Theorem 6.1. *For a right triangle, if c denotes the length of the hypotenuse and a and b denote the lengths of the **other** two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here <https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html>.

6.3 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: <https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html>

Chapter 7

Sharing your book

7.1 Publishing

HTML books can be published online, see: <https://bookdown.org/yihui/bookdown/publishing.html>

7.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

7.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `gitbook` uses the same social sharing data across all chapters in your book—all links shared will look the same.

Specify your book's source repository on GitHub using the `edit` key under the configuration options in the `_output.yml` file, which allows users to suggest an edit by linking to a chapter's source file.

Read more about the features of this output format here:

<https://pkgs.rstudio.com/bookdown/reference/gitbook.html>

Or use:

```
?bookdown::gitbook
```


Bibliography

Yihui Xie. *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition, 2015. URL <http://yihui.org/knitr/>. ISBN 978-1498716963.

Yihui Xie. *bookdown: Authoring Books and Technical Documents with R Markdown*, 2023. URL <https://CRAN.R-project.org/package=bookdown>. R package version 0.32.