

A Minimal YSU Thesis Example

by

Your Name Here

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Mathematics

YOUNGSTOWN STATE UNIVERSITY

May, 20XX

A Minimal YSU Thesis Example

Your Name Here

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Your Name Here, Student

Date

Approvals:

Thesis Advisor's Name, Degree, Thesis Advisor

Date

Committee Member's Name, Degree, Committee Member

Date

Committee Member's Name, Degree, Committee Member

Date

Salvatore A. Sanders, PhD, Dean, College of Graduate Studies

Date

This page must be deleted if you have not filed for a copyright at

copyright.gov

©

Your Name Here

20XX

ABSTRACT

An abstract is a concise account of the thesis or dissertation. An abstract is required for all documents. Format the paragraphs with the same layout used in the document. All lines on the abstract page are double-spaced. Abstract should not exceed one double-spaced page.

ACKNOWLEDGEMENTS

If there is anybody that you would like to acknowledge, then acknowledge those person(s) here. This page may be omitted by commenting out the acknowledgements section in the `template.tex` file.

Contents

1	R Markdown Basics	1
1.1	Lists	1
1.2	Line breaks	2
1.3	R chunks	2
1.4	Inline code	3
1.5	Including plots	3
1.6	Loading and exploring data	4
1.7	Additional resources	9
2	Mathematics and Science	10
2.1	Math	10
2.2	Chemistry 101: Symbols	11
2.2.1	Typesetting reactions	11
2.2.2	Other examples of reactions	11
2.3	Physics	12
2.4	Biology	12
3	Graphics, References, and Labels	12
3.1	Figures	12
3.2	Footnotes and Endnotes	16
3.3	Bibliographies	17
	Conclusion	18
A	The First Appendix	19
B	Another Appendix, if you want one	23
	References	23

List of Tables

1	Max Delays by Airline	8
---	---------------------------------	---

List of Figures

1	YSU logo	13
2	Mean Delays by Airline	14
3	Subdiv. graph	16
4	A Larger Figure, Flipped Upside Down	16

1 R Markdown Basics

Here is a brief introduction into using *R Markdown*. *Markdown* is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. *R Markdown* provides the flexibility of *Markdown* with the implementation of **R** input and output. For more details on using *R Markdown* see <https://rmarkdown.rstudio.com>.

Be careful with your spacing in *Markdown* documents. While whitespace largely is ignored, it does at times give *Markdown* signals as to how to proceed. As a habit, try to keep everything left aligned whenever possible, especially as you type a new paragraph. In other words, there is no need to indent basic text in the Rmd document (in fact, it might cause your text to do funny things if you do).

1.1 Lists

It's easy to create a list. It can be unordered like

- Item 1
- Item 2

or it can be ordered like

1. Item 1
2. Item 2

Notice that I intentionally mislabeled Item 2 as number 4. *Markdown* automatically figures this out! You can put any numbers in the list and it will create the list. Check it out below.

To create a sublist, just indent the values a bit (at least four spaces or a tab). (Here's one case where indentation is key!)

1. Item 1

2. Item 2
3. Item 3
 - Item 3a
 - Item 3b

1.2 Line breaks

Make sure to add white space between lines if you'd like to start a new paragraph. Look at what happens below in the outputted document if you don't:

Here is the first sentence. Here is another sentence. Here is the last sentence to end the paragraph. This should be a new paragraph.

Now for the correct way:

Here is the first sentence. Here is another sentence. Here is the last sentence to end the paragraph.

This should be a new paragraph.

1.3 R chunks

When you click the **Knit** button above a document will be generated that includes both content as well as the output of any embedded **R** code chunks within the document. You can embed an **R** code chunk like this (`cars` is a built-in **R** dataset):

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean    : 42.98
```

```
## 3rd Qu.:19.0    3rd Qu.: 56.00  
## Max.      :25.0    Max.      :120.00
```

1.4 Inline code

If you'd like to put the results of your analysis directly into your discussion, add inline code like this:

The `cos` of 2π is 1.

Another example would be the direct calculation of the standard deviation:

The standard deviation of `speed` in `cars` is 5.2876444.

One last neat feature is the use of the `ifelse` conditional statement which can be used to output text depending on the result of an **R** calculation:

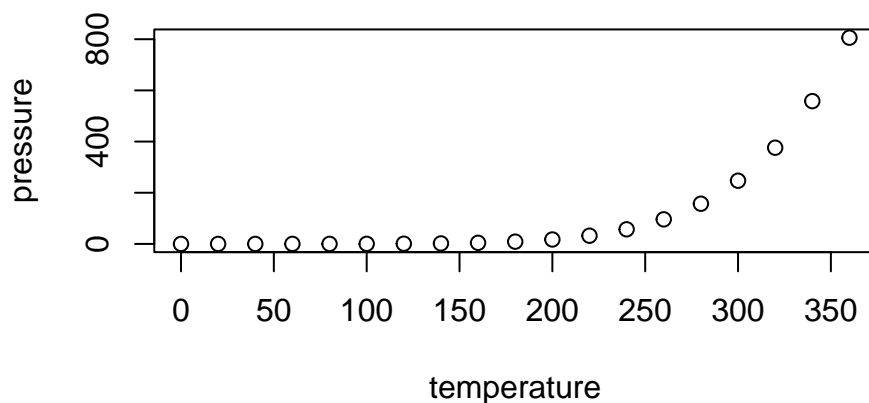
The standard deviation is less than 6.

Note the use of `>` here, which signifies a quotation environment that will be indented.

As you see with `2π` above, mathematics can be added by surrounding the mathematical text with dollar signs. More examples of this are in Mathematics and Science if you uncomment the code in Math.

1.5 Including plots

You can also embed plots. For example, here is a way to use the base **R** graphics package to produce a plot using the built-in `pressure` dataset:



Note that the `echo=FALSE` parameter was added to the code chunk to prevent printing of the **R** code that generated the plot. There are plenty of other ways to add chunk options (like `fig.height` and `fig.width` in the chunk above). More information is available at <https://yihui.org/knitr/options/>.

Another useful chunk option is the setting of `cache=TRUE` as you see here. If document rendering becomes time consuming due to long computations or plots that are expensive to generate you can use knitr caching to improve performance. Later in this file, you'll see a way to reference plots created in **R** or external figures.

1.6 Loading and exploring data

Included in this template is a file called `flights.csv`. This file includes a subset of the larger dataset of information about all flights that departed from Seattle and Portland in 2014. More information about this dataset and its **R** package is available at <https://github.com/ismayc/pnwflights14>. This subset includes only Portland flights and only rows that were complete with no missing values. Merges were also done with the `airports` and `airlines` data sets in the `pnwflights14` package to get more descriptive airport and airline names.

We can load in this data set using the following commands:

```
# flights.csv is in the data directory
flights_path <- here::here("data", "flights.csv")
# string columns will be read in as strings and not factors now
flights <- read.csv(flights_path, stringsAsFactors = FALSE)
```

The data is now stored in the data frame called `flights` in **R**. To get a better feel for the variables included in this dataset we can use a variety of functions. Here we can see the dimensions (rows by columns) and also the names of the columns.

```
dim(flights)
```

```
## [1] 12649    16
```

```
names(flights)
```

```
## [1] "month"      "day"        "dep_time"   "dep_delay"
## [5] "arr_time"   "arr_delay"  "carrier"    "tailnum"
## [9] "flight"     "dest"       "air_time"   "distance"
## [13] "hour"       "minute"     "carrier_name" "dest_name"
```

Another good idea is to take a look at the dataset in table form. With this dataset having more than 20,000 rows, we won't explicitly show the results of the command here. I recommend you enter the command into the Console *after* you have run the **R** chunks above to load the data into **R**.

```
View(flights)
```

While not required, it is highly recommended you use the `dplyr` package to manipulate and summarize your data set as needed. It uses a syntax that is easy to understand using

chaining operations. Below I've created a few examples of using `dplyr` to get information about the Portland flights in 2014. You will also see the use of the `ggplot2` package, which produces beautiful, high-quality academic visuals.

We begin by checking to ensure that needed packages are installed and then we load them into our current working environment:

```
# List of packages required for this analysis
pkg <- c("dplyr", "ggplot2", "knitr", "bookdown")
# Check if packages are not installed and assign the
# names of the packages not installed to the variable new.pkg
new.pkg <- pkg[!(pkg %in% installed.packages())]
# If there are any packages in the list that aren't installed,
# install them
if (length(new.pkg)) {
  install.packages(new.pkg, repos = "https://cran.rstudio.com")
}
# Load packages
library(thesisdown)
library(dplyr)
library(ggplot2)
library(knitr)
```

The example we show here does the following:

- Selects only the `carrier_name` and `arr_delay` from the `flights` dataset and then assigns this subset to a new variable called `flights2`.
- Using `flights2`, we determine the largest arrival delay for each of the carriers.

```
flights2 <- flights %>%  
  select(carrier_name, arr_delay)  
max_delays <- flights2 %>%  
  group_by(carrier_name) %>%  
  summarize(max_arr_delay = max(arr_delay, na.rm = TRUE))
```

A useful function in the `knitr` package for making nice tables in *R Markdown* is called `kable`. It is much easier to use than manually entering values into a table by copying and pasting values into Excel or LaTeX. This again goes to show how nice reproducible documents can be! (Note the use of `results="asis"`, which will produce the table instead of the code to create the table.) The `caption.short` argument is used to include a shorter title to appear in the List of Tables.

```
kable(max_delays,  
  col.names = c("Airline", "Max Arrival Delay"),  
  caption = "Maximum Delays by Airline",  
  caption.short = "Max Delays by Airline",  
  longtable = TRUE,  
  booktabs = TRUE  
)
```

Table 1: Maximum Delays by Airline

Airline	Max Arrival Delay
Alaska Airlines Inc.	338
American Airlines Inc.	1539
Delta Air Lines Inc.	371
Frontier Airlines Inc.	166
Hawaiian Airlines Inc.	116
JetBlue Airways	256
SkyWest Airlines Inc.	321
Southwest Airlines Co.	315
United Air Lines Inc.	319
US Airways Inc.	347
Virgin America	366

The last two options make the table a little easier-to-read.

We can further look into the properties of the largest value here for American Airlines Inc. To do so, we can isolate the row corresponding to the arrival delay of 1539 minutes for American in our original `flights` dataset.

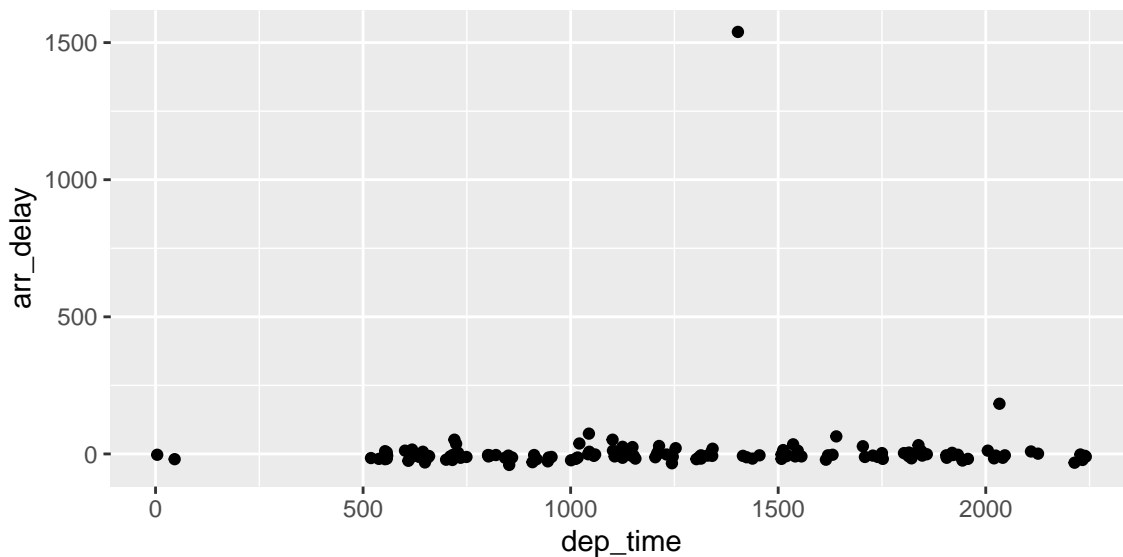
```
flights %>%
  filter(
    arr_delay == 1539,
    carrier_name == "American Airlines Inc."
  ) %>%
  select(-c(
    month, day, carrier, dest_name, hour,
```

```
minute, carrier_name, arr_delay
))
```

```
##   dep_time dep_delay arr_time tailnum flight dest air_time distance
## 1    1403      1553     1934  N595AA   1568  DFW       182       1616
```

We see that the flight occurred on March 3rd and departed a little after 2 PM on its way to Dallas/Fort Worth. Lastly, we show how we can visualize the arrival delay of all departing flights from Portland on March 3rd against time of departure.

```
flights %>%
  filter(month == 3, day == 3) %>%
  ggplot(aes(x = dep_time, y = arr_delay)) +
  geom_point()
```



1.7 Additional resources

- *Markdown Cheatsheet* - <https://github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet>

- *R Markdown*
 - Reference Guide - <https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf>
 - Cheatsheet - <https://github.com/rstudio/cheatsheets/raw/master/rmarkdown-2.0.pdf>
- *RStudio IDE*
 - Cheatsheet - <https://github.com/rstudio/cheatsheets/raw/master/rstudio-ide.pdf>
 - Official website - <https://rstudio.com/products/rstudio/>
- Introduction to dplyr - <https://cran.rstudio.com/web/packages/dplyr/vignettes/dplyr.html>
- ggplot2
 - Documentation - <https://ggplot2.tidyverse.org/>
 - Cheatsheet - <https://github.com/rstudio/cheatsheets/raw/master/data-visualization-2.1.pdf>

2 Mathematics and Science

2.1 Math

\TeX is the best way to typeset mathematics. Donald Knuth designed \TeX when he got frustrated at how long it was taking the typesetters to finish his book, which contained a lot of mathematics. One nice feature of *R Markdown* is its ability to read LaTeX code directly.

If you are doing a thesis that will involve lots of math, you will want to read the following section which has been commented out. If you're not going to use math, skip over or delete this next commented section.

2.2 Chemistry 101: Symbols

Chemical formulas will look best if they are not italicized. Get around math mode's automatic italicizing in LaTeX by using the argument `$\mathrm{formula here}$` , with your formula inside the curly brackets. (Notice the use of the backticks here which enclose text that acts as code.)

So, $\text{Fe}_2^{2+}\text{Cr}_2\text{O}_4$ is written `$\mathrm{Fe_2^{2+}Cr_2O_4}$` .

Exponent or Superscript: O^-

Subscript: CH_4

To stack numbers or letters as in Fe_2^{2+} , the subscript is defined first, and then the superscript is defined.

Bullet: $\text{CuCl} \bullet 7\text{H}_2\text{O}$

Delta: Δ

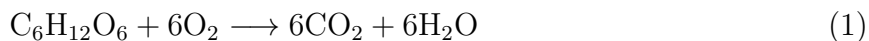
Reaction Arrows: \longrightarrow or $\xrightarrow{\text{solution}}$

Resonance Arrows: \leftrightarrow

Reversible Reaction Arrows: \rightleftharpoons

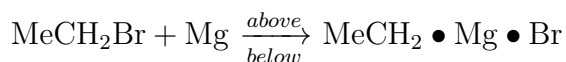
2.2.1 Typesetting reactions

You may wish to put your reaction in an equation environment, which means that LaTeX will place the reaction where it fits and will number the equations for you.



We can reference this combustion of glucose reaction via Equation (1).

2.2.2 Other examples of reactions



2.3 Physics

Many of the symbols you will need can be found on the math page <https://web.reed.edu/cis/help/latex/math.html> and the Comprehensive LaTeX Symbol Guide (<https://mirror.utexas.edu/ctan/info/symbols/comprehensive/symbols-letter.pdf>).

2.4 Biology

You will probably find the resources at <https://www.lecb.ncifcrf.gov/~toms/latex.html> helpful, particularly the links to bst files for various journals. You may also be interested in TeXShade for nucleotide typesetting (<https://homepages.uni-tuebingen.de/beitz/txe.html>). Be sure to read the next section on graphics and tables.

3 Graphics, References, and Labels

3.1 Figures

If your thesis has a lot of figures, *R Markdown* might behave better for you than that other word processor. One perk is that it will automatically number the figures accordingly in each section. You'll also be able to create a label for each figure, add a caption, and then reference the figure in a way similar to what we saw with tables earlier. If you label your figures, you can move the figures around and *R Markdown* will automatically adjust the numbering for you. No need for you to remember! So that you don't have to get too far into LaTeX to do this, a couple **R** functions have been created for you to assist. You'll see their use below.

In the **R** chunk below, we will load in a picture stored as `YSU.png` in our main directory. We then give it the caption of "YSU logo", the label of "ysulogo", and specify that this is a figure. Make note of the different **R** chunk options that are given in the R Markdown file (not shown in the knitted document).

```
include_graphics(path = "figure/YSU.png")
```



Figure 1: YSU logo

Here is a reference to the YSU logo: Figure 1. Note the use of the `fig:` code here. By naming the **R** chunk that contains the figure, we can then reference that figure later as done in the first sentence here. We can also specify the caption for the figure via the R chunk option `fig.cap`.

Below we will investigate how to save the output of an **R** plot and label it in a way similar to that done above. Recall the `flights` dataset from Chapter 1. (Note that we've shown a different way to reference a section here.) We will next explore a bar graph with the mean flight departure delays by airline from Portland for 2014.

```
mean_delay_by_carrier <- flights %>%  
  group_by(carrier) %>%  
  summarize(mean_dep_delay = mean(dep_delay))  
ggplot(mean_delay_by_carrier, aes(x = carrier, y = mean_dep_delay)) +  
  geom_bar(position = "identity", stat = "identity", fill = "red")
```

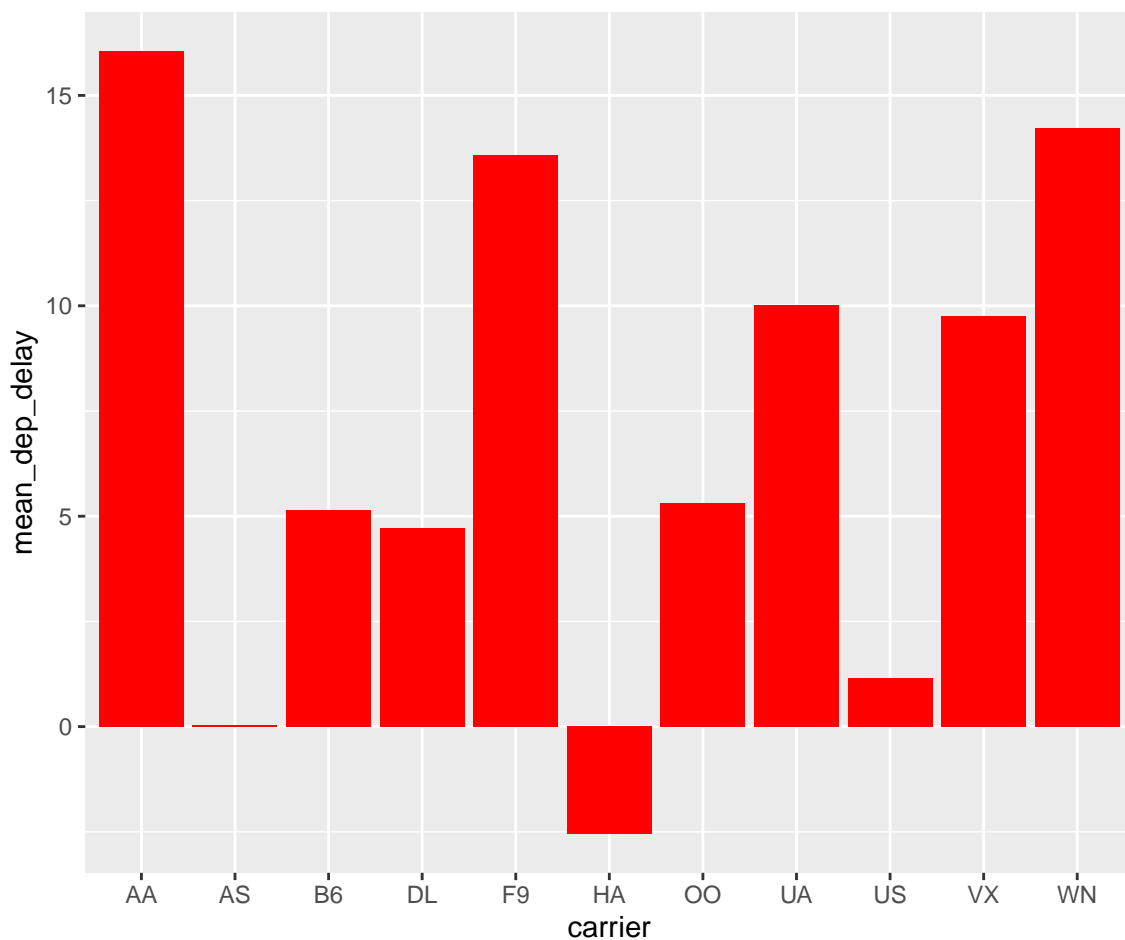


Figure 2: Mean Delays by Airline

Here is a reference to this image: Figure 2.

A table linking these carrier codes to airline names is available at <https://github.com/ismayc/pnwflights14/blob/master/data/airlines.csv>.

Next, we will explore the use of the `out.extra` chunk option, which can be used to shrink or expand an image loaded from a file by specifying "`scale=`". Here we use the mathematical graph stored in the “subdivision.pdf” file.

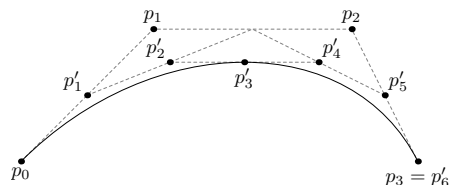


Figure 3: Subdiv. graph

Here is a reference to this image: Figure 3. Note that `echo=FALSE` is specified so that the **R** code is hidden in the document.

More Figure Stuff

Lastly, we will explore how to rotate and enlarge figures using the `out.extra` chunk option. (Currently this only works in the PDF version of the book.)

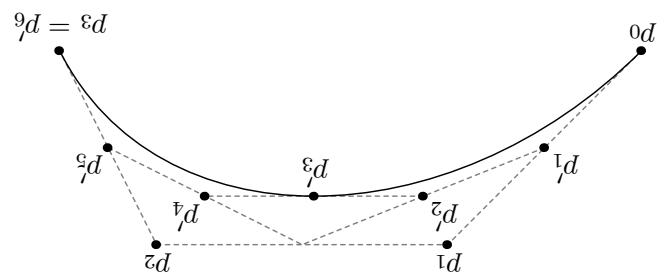


Figure 4: A Larger Figure, Flipped Upside Down

As another example, here is a reference: Figure 4.

3.2 Footnotes and Endnotes

You might want to footnote something.¹ The footnote will be in a smaller font and placed appropriately. Endnotes work in much the same way.

¹footnote text

3.3 Bibliographies

Of course you will need to cite things, and you will probably accumulate an armful of sources. There are a variety of tools available for creating a bibliography database (stored with the .bib extension). In addition to BibTeX suggested below, you may want to consider using the free and easy-to-use tool called Zotero. The librarians at Reed College have created Zotero documentation at <https://libguides.reed.edu/citation/zotero>. In addition, a tutorial is available from Middlebury College at <https://sites.middlebury.edu/zoteromiddlebury/>.

R Markdown uses *pandoc* (<https://pandoc.org/>) to build its bibliographies. One nice caveat of this is that you won't have to do a second compile to load in references as standard LaTeX requires. To cite references in your thesis (after creating your bibliography database), place the reference name inside square brackets and precede it by the “at” symbol. For example, here's a reference to a book about worrying: (Molina & Borkovec, 1994). This `Molina1994` entry appears in a file called `thesis.bib` in the `bib` folder. This bibliography database file was created by a program called BibTeX. You can call this file something else if you like (look at the YAML header in the main .Rmd file) and, by default, is to placed in the `bib` folder.

For more information about BibTeX and bibliographies, see the CUS site at Reed College (<https://web.reed.edu/cis/help/latex/index.html>)². There are three pages on this topic: *bibtex* (which talks about using BibTeX, at <https://web.reed.edu/cis/help/latex/bibtex.html>), *bibtexstyles* (about how to find and use the bibliography style that best suits your needs, at <https://web.reed.edu/cis/help/latex/bibtexstyles.html>) and *bibman* (which covers how to make and maintain a bibliography by hand, without BibTeX, at <https://web.reed.edu/cis/help/latex/bibman.html>). The last page will not be useful unless you have only a few sources.

If you look at the YAML header at the top of the main .Rmd file you can see that we

²Reed College (2007)

can specify the style of the bibliography by referencing the appropriate csl file. You can download a variety of different style files at <https://www.zotero.org/styles>. Make sure to download the file into the csl folder.

Tips for Bibliographies

- Like with thesis formatting, the sooner you start compiling your bibliography for something as large as thesis, the better. Typing in source after source is mind-numbing enough; do you really want to do it for hours on end in late April? Think of it as procrastination.
- The cite key (a citation’s label) needs to be unique from the other entries.
- When you have more than one author or editor, you need to separate each author’s name by the word “and” e.g. `Author = {Noble, Sam and Youngberg, Jessica},.`
- Bibliographies made using BibTeX (whether manually or using a manager) accept LaTeX markup, so you can italicize and add symbols as necessary.
- To force capitalization in an article title or where all lowercase is generally used, bracket the capital letter in curly braces.
- You can add a YSU Thesis citation³ option. The best way to do this is to use the `phdthesis` type of citation, and use the optional “type” field to enter “YSU thesis” or “Undergraduate thesis.”

Conclusion

If we don’t want Conclusion to have a section number next to it, we can add the `{-}` attribute.

More info

And here’s some other random info: the first paragraph after a section head *shouldn’t be* indented, because indents are to tell the reader that you’re starting a new paragraph. Since that’s obvious after a chapter or section title, proper typesetting doesn’t add an indent there.

³Noble (2002)

A The First Appendix

This first appendix includes all of the R chunks of code that were hidden throughout the document (using the `include = FALSE` chunk tag) to help with readability and/or setup.

In the main Rmd file

```
# This chunk ensures that the thesisdown package is
# installed and loaded. This thesisdown package includes
# the template files for the thesis.
if (!require(remotes)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("remotes", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste('You need to run install.packages("remotes")',
            "first in the Console.")
    )
  }
}

if (!require(thesisdown)) {
  if (params$`Install needed packages for {thesisdown}`) {
    remotes::install_github("ismayc/thesisdown")
  } else {
    stop(
      paste(
        "You need to run",
        'remotes::install_github("ismayc/thesisdown")',
        "first in the Console."
      )
    )
  }
}
```

```

    )
  )
}
}
library(thesisdown)
# Set how wide the R output will go
options(width = 70)

```

In Section 3:

```

# This chunk ensures that the thesisdown package is
# installed and loaded. This thesisdown package includes
# the template files for the thesis and also two functions
# used for labeling and referencing
if (!require(remotes)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("remotes", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("remotes")',
        "first in the Console."
      )
    )
  }
}

if (!require(dplyr)) {
  if (params$`Install needed packages for {thesisdown}`) {

```

```

    install.packages("dplyr", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("dplyr")',
        "first in the Console."
      )
    )
  }
}

if (!require(ggplot2)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("ggplot2", repos = "https://cran.rstudio.com")
  } else {
    stop(
      paste(
        'You need to run install.packages("ggplot2")',
        "first in the Console."
      )
    )
  }
}

if (!require(bookdown)) {
  if (params$`Install needed packages for {thesisdown}`) {
    install.packages("bookdown", repos = "https://cran.rstudio.com")
  } else {
    stop(

```

```

    paste(
      'You need to run install.packages("bookdown")',
      "first in the Console."
    )
  )
}
}

if (!require(thesisdown)) {
  if (params$`Install needed packages for {thesisdown}`) {
    remotes::install_github("ismayc/thesisdown")
  } else {
    stop(
      paste(
        "You need to run",
        'remotes::install_github("ismayc/thesisdown")',
        "first in the Console."
      )
    )
  }
}

library(thesisdown)
library(dplyr)
library(ggplot2)
library(knitr)

flights <- read.csv("data/flights.csv", stringsAsFactors = FALSE)

```

B Another Appendix, if you want one

References

- Angel, E. (2000). *Interactive computer graphics : A top-down approach with OpenGL*. Boston, MA: Addison Wesley Longman.
- Angel, E. (2001a). *Batch-file computer graphics : A bottom-up approach with QuickTime*. Boston, MA: Wesley Addison Longman.
- Angel, E. (2001b). *Test second book by angel*. Boston, MA: Wesley Addison Longman.
- Molina, S. T., & Borkovec, T. D. (1994). The Penn State worry questionnaire: Psychometric properties and associated characteristics. In G. C. L. Davey & F. Tallis (Eds.), *Worrying: Perspectives on theory, assessment and treatment* (pp. 265–283). New York: Wiley.
- Noble, S. G. (2002). *Turning images into simple line-art* (Undergraduate thesis). Reed College.
- Reed College. (2007). LaTeX your document. Retrieved from <https://web.reed.edu/cis/help/LaTeX/index.html>