

Data Engineering

COMP2031/8031



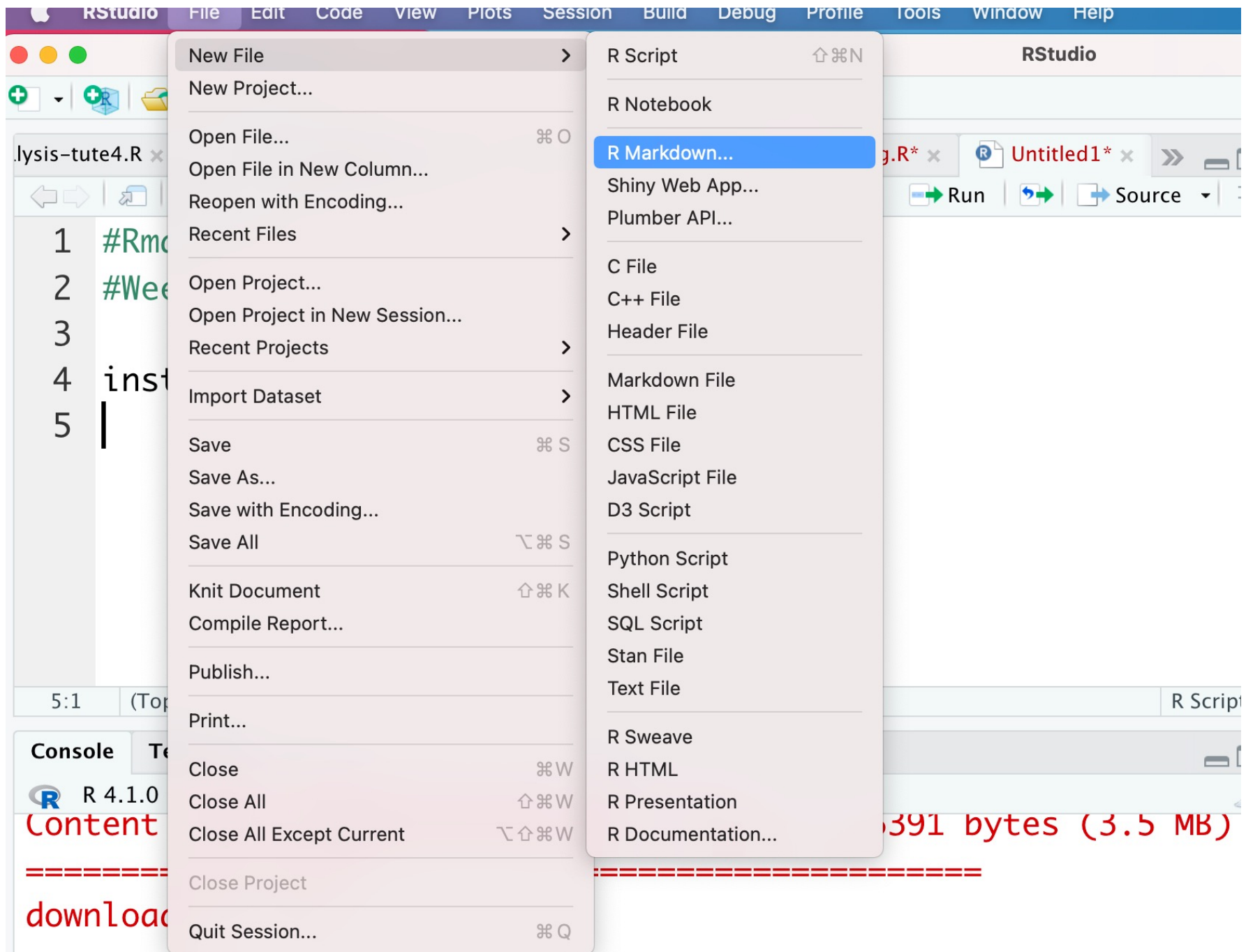
- Topic Coordinator: Dr Mehwish Nasim
- Office: 3.13 Tonsley

Flinders
UNIVERSITY




Rmarkdown

- `install.packages("rmarkdown")`



New R Markdown

 Document

 Presentation

 Shiny

 From Template

Title: myfirstRMDfile

Author: Mehwish

Default Output Format:

☒ HTML

Recommended format for authoring (you can switch to PDF or Word output anytime).

☐ PDF

PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).

☐ Word

Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

Create Empty Document

OK

Cancel

RStudio

Go to file/function | Addins

Iris-EDA-casestudy2.R x kmeans.R x topic-modeling.R* x Untitled1* x myfirstmarkdown.Rmd

Knit Run

```
1 ---
2 title: "My first R markdown report"
3 output: html_document
4 ---
```

4:4 # My first R markdown report R Markdown



My first R markdown report

- You can use regular **markdown** rules in your **R Markdown** document. Once you **knit** your document, the output will display text formatted according to the following simple rules
- Common formatting commands
 - **Italic**
 - ****Bold****
 - # Header1
 - ## Header 2
 - * Unordered list item
 - 1. Ordered list item
 - [Link](<https://www.google.com>)
 - $A = \pi r^2$

You can quickly insert chunks like these into your file with

- the keyboard shortcut **Ctrl + Alt + I** (OS X: **Cmd + Option + I**)
- the Add Chunk command in the editor toolbar

or by typing the chunk delimiters `` `` { r } `` `` and `` `` ``.

When you render your .Rmd file, R Markdown will run each code chunk and embed the results beneath the code chunk in your final report.


```
```{r}  
norm = rnorm(100, mean = 0, sd = 1)
```
```

- ```{r}
- A <- c("a", "a", "b", "b")
- B <- c(5, 10, 15, 20)
- dataframe <- data.frame(A, B)
- print(dataframe)
- ```{r}
- print(dataframe)
- ```

Examples

```
```{r} dataframe <- read.csv("~/Desktop/Code/dataframe.csv")  
```
```

For loading libraries

```
```  
{r} library(dplyr)
```
```

~/OneDrive - Flinders/DataEngineering-Material/wk6/myfirstmarkdown.html

myfirstmarkdown.html Open in Browser Find Publish

My first R markdown report

```
norm = rnorm(100, mean = 0, sd = 1)
```

```
A <- c("a", "a", "b", "b")
B <- c(5, 10, 15, 20)
dataframe <- data.frame(A, B)
print(dataframe)
```

```
##   A  B
## 1 a  5
## 2 a 10
## 3 b 15
## 4 b 20
```

```
print(dataframe)
```

```
##   A  B
## 1 a  5
## 2 a 10
## 3 b 15
## 4 b 20
```

Hiding code chunks

```
`` `{r, echo = FALSE}  
A <- c("a", "a", "b", "b")  
B <- c(5, 10, 15, 20)  
dataframe <- data.frame(A, B)  
print(dataframe)  
```
```

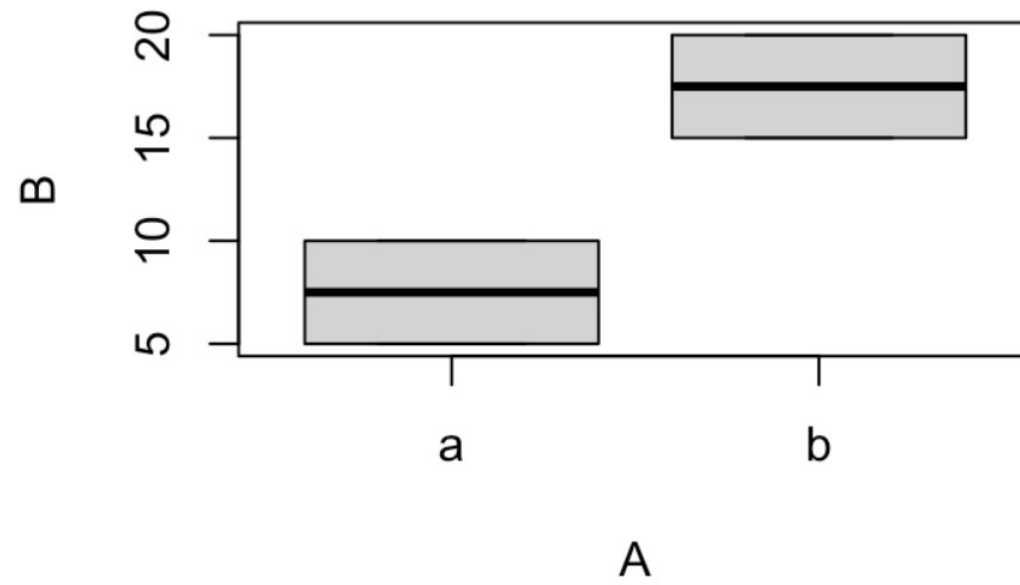


```
`` `{r, warning = FALSE}
library(dplyr)
```
```

| Rule | Example (default) | Function |
|-----------------------|-------------------|--|
| eval | eval=TRUE | Is the code run and the results included in the output? |
| include | include=TRUE | Are the code and the results included in the output? |
| echo | echo=TRUE | Is the code displayed alongside the results? |
| warning | warning=TRUE | Are warning messages displayed? |
| error | error=FALSE | Are error messages displayed? |
| message | message=TRUE | Are messages displayed? |
| tidy | tidy=FALSE | Is the code reformatted to make it look "tidy"? |
| results | results="markup" | How are results treated?
"hide" = no results
"asis" = results without formatting
"hold" = results only compiled at end of chunk (use if many commands act on one object) |
| cache | cache=FALSE | Are the results cached for future renders? |
| comment | comment="###" | What character are comments prefaced with? |
| fig.width, fig.height | fig.width=7 | What width/height (in inches) are the plots? |
| fig.align | fig.align="left" | "left" "right" "center" |

```
```{r, fig.width = 4, fig.height = 3}  
A <- c("a", "a", "b", "b")
B <- c(5, 10, 15, 20)
dataframe <- data.frame(A, B)
print(dataframe)
boxplot(B~A,data=dataframe)
```
```

```
boxplot(B~A,data=dataframe)
```



- R Markdown can print the contents of a data frame but the look is messy. Use Kable() function from the knitr package

```
library(knitr)
kable(dataframe, digits = 2)
```

| A | B |
|---|----|
| a | 5 |
| a | 10 |
| b | 15 |
| b | 20 |

For customizing your table

```
plant <- c("a", "b", "c")
temperature <- c(20, 20, 20)
growth <- c(0.65, 0.95, 0.15)
dataframe <- data.frame(plant, temperature, growth)
emphasize.italics.cols(3)    # Make the 3rd column italics
pander(dataframe)           # Create the table
```

| plant | temperature | growth |
|-------|-------------|-------------|
| a | 20 | <i>0.65</i> |
| b | 20 | <i>0.95</i> |
| c | 20 | <i>0.15</i> |

