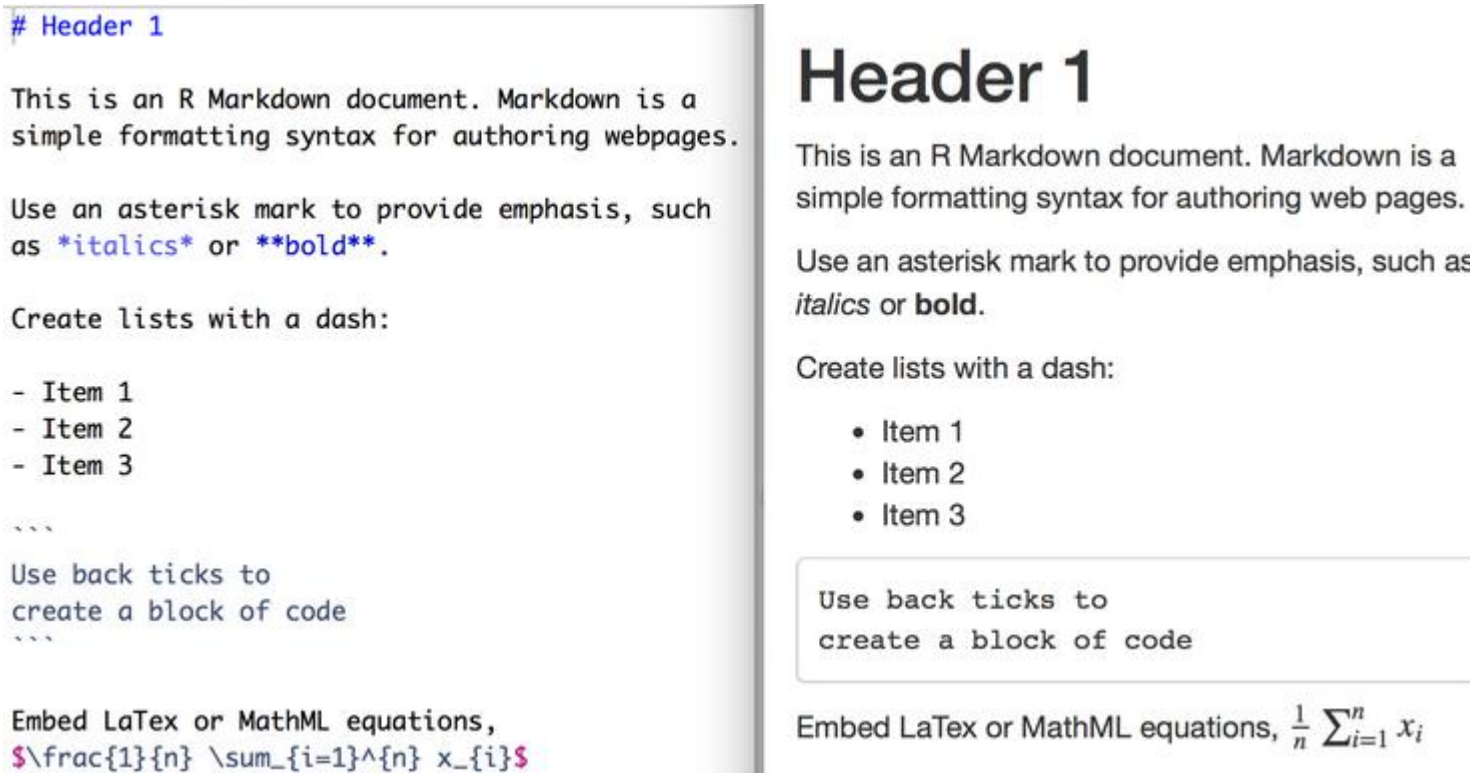
R Markdown: The Basics

Andrew Muehleisen

In this class, we will frequently use the ‘R Markdown’ file format.

For example, you will submit problem sets as .Rmd files.



```
# Header 1

This is an R Markdown document. Markdown is a
simple formatting syntax for authoring webpages.

Use an asterisk mark to provide emphasis, such
as *italics* or **bold**.

Create lists with a dash:

- Item 1
- Item 2
- Item 3

```
Use back ticks to
create a block of code
```

Embed LaTex or MathML equations,
$\frac{1}{n} \sum_{i=1}^{n} x_{i}$
```

Header 1

This is an R Markdown document. Markdown is a simple formatting syntax for authoring web pages.

Use an asterisk mark to provide emphasis, such as *italics* or **bold**.

Create lists with a dash:

- Item 1
- Item 2
- Item 3

```
Use back ticks to
create a block of code
```

Embed LaTex or MathML equations, $\frac{1}{n} \sum_{i=1}^{n} x_i$

What is Markdown?

Markdown is a simple language for converting plain text into formatted, rich text (as in a Word document).

Why are we using R Markdown?

In a single document, R Markdown lets us have both…

1. **Formatted**, *stylized* text

2. Embedded R code that runs and renders with the document

```
print("R Markdown is awesome!")
```

Example:

```
---
title: "Problem Set 2"
output:
  html_document: default
  html_notebook: default
  pdf_document: default
  word_document: default
---

### Name:

About this format: This problem set is saved as an [R Markdown](http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code. When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file). To add a new R chunk click the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

**To submit: Please rename this file LASTNAME_ProblemSet2 and upload both the .Rmd and the final .html file to the assignments folder on Canvas**


### I. CO2 Trends
First read in the data and get it ready to use
You can find the data and read more about it here:
https://climate.nasa.gov/vital-signs/carbon-dioxide/
In particular, the metadata is at the top of the txt file if you click on the Download Data button

To get you started here's the code to read in the data and give it better column names

```{r}
# library data.table allows you to read files directly from the internet
# install.packages("data.table")
library(data.table)

# Let's skip down to the data, which starts at line 72
CO2data <- fread('ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_mm_mlo.txt', skip = 71)

# give descriptive but short names
names(CO2data) <- c("year", "month", "decimalDate", "averageCO2", "interpolatedCO2", "trendCO2", "numberDays")


```
```

Example:

```
---
title: "Problem Set 2"
output:
  html_document: default
  html_notebook: default
  pdf_document: default
  word_document: default
---

### Name:

About this format: This problem set is saved as an [R Markdown](http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code. When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file). To add a new R chunk click the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

**To submit: Please rename this file LASTNAME_ProblemSet2 and upload both the .Rmd and the final .html file to the assignments folder on Canvas**


### I. CO2 Trends
First read in the data and get it ready to use
You can find the data and read more about it here:
https://climate.nasa.gov/vital-signs/carbon-dioxide/
In particular, the metadata is at the top of the txt file if you click on the Download Data button

To get you started here's the code to read in the data and give it better column names

```{r}
# library data.table allows you to read files directly from the internet
# install.packages("data.table")
library(data.table)

# Let's skip down to the data, which starts at line 72
CO2data <- fread('ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_mm_mlo.txt', skip = 71)

# give descriptive but short names
names(CO2data) <- c("year", "month", "decimalDate", "averageCO2", "interpolatedCO2", "trendCO2", "numberDays")


```
```

Problem Set 2

Name:

About this format: This problem set is saved as an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code. When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file). To add a new R chunk click the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

To submit: Please rename this file LASTNAME_ProblemSet2 and upload both the .Rmd and the final .html file to the assignments folder on Canvas

I. CO2 Trends

First read in the data and get it ready to use You can find the data and read more about it here: https://climate.nasa.gov/vital-signs/carbon-dioxide/ In particular, the metadata is at the top of the txt file if you click on the Download Data button

To get you started here's the code to read in the data and give it better column names

```
# library data.table allows you to read files directly from the intern
et
# install.packages("data.table")
library(data.table)

# Let's skip down to the data, which starts at line 72
CO2data <- fread('ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_mm_
mlo.txt', skip = 71)
```

```
## Warning in fread("ftp://aftp.cmdl.noaa.gov/products/trends/co2/
## co2_mm_mlo.txt", : Detected 4 column names but the data has 7 colum
ns (i.e.
## invalid file). Added 3 extra default column names at the end.
```

```
# give descriptive but short names
names(CO2data) <- c("year", "month", "decimalDate", "averageCO2", "int
erpolatedCO2", "trendCO2", "numberDays")
```

Getting started

For all problem sets, we have provided the initial Markdown scaffolding.

You will:

1. Open problem sets in R Studio

2. Fill in the necessary text or code

3. Create the associated HTML document, and submit both documents on Canvas

Getting started: Headers

To format text in a markdown document, we use symbols alongside the text.

For example, we use # to indicate a header. Multiple #s can be used to create a hierarchy of headers.

```
# This is a top-level header

Some text

## This is a sub-header

Some more text

### This is a lower-level header

An example
```

This is a top-level header

Some text

This is a sub-header

Some more text

This is a lower-level header

An example

Getting started: Emphasis

To create emphases in text, like bold or italics, we surround our text with asterisks **‘*’**

Italics = *text*

Bold text = **text**

```
I really want to emphasize this *word*.

Perhaps I want to emphasize this **word** in bold.

Why not emphasize our ***word*** using both?
```

I really want to emphasize this *word*.

Perhaps I want to emphasize this **word** in bold.

Why not emphasize our ***word*** using both?

Getting started: Lists

Lists are very easy. Just use numbers or dashes.

```
1. The first entry
2. The second entry
3. The final entry

- Point 1
- Point 2
- Point 3
```

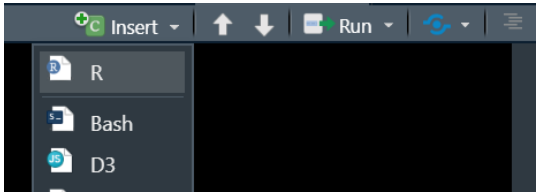
1. The first entry
2. The second entry
3. The final entry

- Point 1
- Point 2
- Point 3

Getting started: Code Chunks

Finally, using R Markdown, we can embed code into the resulting document. We do so by delineating chunks of code with backticks.

Type, or use R Studio




```
## Embedded code

Using three backticks, we can run code:

```{r}
data(cars)
summary(cars)
```
```

Embedded code

Using three backticks, we can run code:

```
data(cars)
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
```

Getting started: Code Chunks

Note that at the beginning of the code chunk, we specify ‘r’ to indicate that the code will be an R expression. We can also add helpful options here.

“**echo = FALSE**”: hide code, just show output

````
## Embedded code

Using three backticks, we can run code:

```{r, echo = FALSE}
data(cars)
summary(cars)
```
````

Embedded code

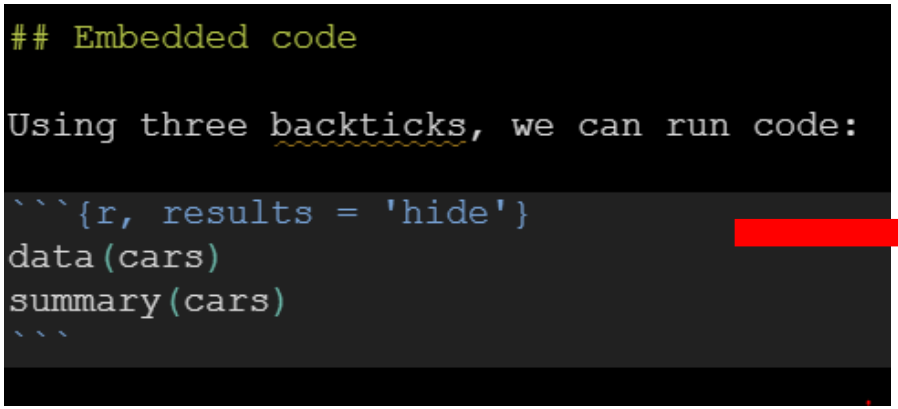
Using three backticks, we can run code:

```
##      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
```

Getting started: Code Chunks

Note that at the beginning of the code chunk, we specify ‘r’ to indicate that the code will be an R expression. We can also add helpful options here.

“**results = ‘hide’**”: show code, but not output



Embedded code

Using three backticks, we can run code:

```
data(cars)
summary(cars)
```

Getting started: knit

Finally, we generate the HTML document output by clicking on the ‘knit’ button in our script editor.