

# More R Markdown

*Andrew Ba Tran*

## Contents

This is from the sixth chapter of [learn.r-journalism.com](http://learn.r-journalism.com).

Let's create some R Markdown files.

Make sure your working directory is set.

If you're not working with the `learn-chapter-6-master` folder you downloaded with **usethis**, download your files to a folder called *data*.

If you get lost, the `.Rmd` files can be found in the lesson repo.

We'll start out by generating a report with Boston city payroll data.

## Datatables

1. Create a new R Markdown file and call it **Chunk 1**.
  - Leave author blank for these exercises

The top of your file (currently called *Untitled 1*) should look like this:

```
---
title: "Chunk 1"
output: html_document
---
```

and then that will be followed by the dummy code.

Delete everything beneath the YAML code.

Replace it with this code:

```
```{r loading}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")
```

Let's look at the data in R Markdown with a package called [DT]
(https://rstudio.github.io/DT/) that uses the Datatables [jquery library]
(https://datatables.net/).

```{r display_data}
library(DT)
datatable(payroll)
```
```

Save the file as `01_chunk.Rmd` and click the knit button.

## Chunk 1

```
# load packages
library(tidyverse)

## -- Attaching packages -- tidyverse 1.2.1 --
-

## ✓ ggplot2 3.0.0      ✓ purrr 0.2.5
## ✓ tibble 1.4.2       ✓ dplyr 0.7.6
## ✓ tidyr 0.8.1        ✓ stringr 1.3.1
## ✓ readr 1.1.1        ✓ forcats 0.3.0

## -- Conflicts -- tidyverse_conflicts() --
-
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()     masks stats::lag()

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")

## Warning: Missing column names filled in: 'X15' [15], 'X16' [16]

## Parsed with column specification:
## cols(
##   NAME = col_character(),
##   TITLE = col_character(),
##   DEPARTMENT = col_character(),
##   REGULAR = col_character(),
##   RETRO = col_character(),
##   OTHER = col_character(),
##   OVERTIME = col_character(),
##   INJURED = col_character(),
##   DETAIL = col_character(),
##   QUINN = col_character(),
##   `TOTAL EARNINGS` = col_character(),
##   Community = col_character(),
##   ZIPCode = col_integer(),
##   State = col_character(),
##   X15 = col_character(),
##   X16 = col_character()
## )

Let's look at the data in R Markdown with a package called DT that uses the Datatables jquery library.

library(DT)
datatable(payroll)

## Warning in instance$preRenderHook(instance): It seems your data is too
## big for client-side DataTables. You may consider server-side processing:
## https://rstudio.github.io/DT/server.html

Show 10 entries Search:
```

|   | NAME                | TITLE          | DEPARTMENT               | REGULAR   | RETRO | OTHER     | OVERTIME | INJURED | DETAIL   | QUINN    | TOTAL EARNINGS | Community | ZIPCode |     |
|---|---------------------|----------------|--------------------------|-----------|-------|-----------|----------|---------|----------|----------|----------------|-----------|---------|-----|
| 1 | Darosa, Baltazar    | Police Officer | Boston Police Department | \$66,829  | \$0   | \$209,237 | \$10,473 | \$0     | \$629    | \$6,724  | \$293,892      | Brockton  | 2302    | Mat |
| 2 | Creaven, Jacqueline | Police         | Boston Police            | \$104,661 | \$0   | \$79,248  | \$47,572 | \$0     | \$33,918 | \$23,721 | \$289,121      | Boston    | 2132    | Mat |

Yikes, okay, that's way too much.

## Hide warnings, messages

We can hide those console messages adding `warning=F` and `message=F` by the R code chunk labels.

Create a new R Markdown file and call it **Chunk 2**.

Type the code in below.

The new code can be found on lines 6 and 16.

```

---
title: "Chunk 2"
output: html_document
---

```{r loading, warning=F, message=F}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")
```

Let's look at the data in R Markdown with a new package called [DT]
(https://rstudio.github.io/DT/) that uses the Datatables [jquery library]
(https://datatables.net/).

```{r display_data, warning=F}
library(DT)
datatable(payroll)
```

```

Save the file as 02\_chunk.Rmd and click the knit button.

## Chunk 2

```

# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")

```

Let's look at the data in R Markdown with a new package called `DT` that uses the Datatables [jquery library](https://datatables.net/).

```

library(DT)
datatable(payroll)

```

Show 10 entries

Search:

|   | NAME                  | TITLE                          | DEPARTMENT               | REGULAR   | RETRO | OTHER     | OVERTIME | INJURED | DETAIL   | QUINN    |
|---|-----------------------|--------------------------------|--------------------------|-----------|-------|-----------|----------|---------|----------|----------|
| 1 | Darosa, Baltazar      | Police Officer                 | Boston Police Department | \$66,829  | \$0   | \$209,237 | \$10,473 | \$0     | \$629    | \$6,724  |
| 2 | Creaven, Jacqueline D | Police Lieutenant              | Boston Police Department | \$104,661 | \$0   | \$79,248  | \$47,572 | \$0     | \$33,918 | \$23,721 |
| 3 | Cawley, Stephen C     | Police Lieutenant              | Boston Police Department | \$44,632  | \$0   | \$216,037 | \$10,158 | \$0     | \$4,046  | \$6,031  |
| 4 | Danilecki, John H     | Police Captain                 | Boston Police Department | \$121,244 | \$0   | \$15,663  | \$71,198 | \$0     | \$41,943 | \$15,156 |
| 5 | Kervin, Timothy M.    | Police Lieutenant/Hdq Dispatch | Boston Police Department | \$104,588 | \$0   | \$13,494  | \$88,906 | \$0     | \$41,840 | \$12,988 |

Now that's much more readable and gets to the data quicker.

## Hide code

If the person you're sharing this with has no interest in the code and only the quick results, use `echo=F` to hide the chunk of code and just display the output. It's on line 8.

We'll also narrow down the variables selected so the table isn't way too wide.

Create a new R Markdown file and call it **Chunk 3**.

Type the code in below.

The new code can be found on 8 and 17.

```
---
title: "Chunk 3"
output: html_document
---

# Boston employee pay in 2014

```{r loading, warning=F, message=F, echo=F}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")
payroll_total <- select(payroll, NAME, TITLE, DEPARTMENT, TOTAL.EARNINGS)
```

```{r display_data, warning=F, message=F, echo=F}
library(DT)
datatable(payroll_total)
```
```

Save the file as 03\_chunk.Rmd and click the knit button.

## Chunk 3

### Boston employee pay in 2014

Show 10 entries

Search:

|    | NAME                  | TITLE                          | DEPARTMENT               | TOTAL EARNINGS |
|----|-----------------------|--------------------------------|--------------------------|----------------|
| 1  | Darosa, Baltazar      | Police Officer                 | Boston Police Department | \$293,892      |
| 2  | Creaven, Jacqueline D | Police Lieutenant              | Boston Police Department | \$289,121      |
| 3  | Cawley, Stephen C     | Police Lieutenant              | Boston Police Department | \$280,904      |
| 4  | Danilecki, John H     | Police Captain                 | Boston Police Department | \$265,203      |
| 5  | Kervin, Timothy M.    | Police Lieutenant/Hdq Dispatch | Boston Police Department | \$261,815      |
| 6  | Hosein, Haseeb        | Police Lieutenant              | Boston Police Department | \$255,020      |
| 7  | Lee, Thomas F         | Police Captain/DDC             | Boston Police Department | \$252,543      |
| 8  | Assad, Mark L         | Police Sergeant (Det)          | Boston Police Department | \$247,964      |
| 9  | Eversley, Eric V      | Police Lieutenant (Det)        | Boston Police Department | \$247,201      |
| 10 | Crossen, Patrick J    | Police Captain/DDC             | Boston Police Department | \$245,734      |

Showing 1 to 10 of 5,173 entries

Previous **1** 2 3 4 5 ... 518 Next

#### Inline R code

Embed lines of R code within the markdown narrative with

Create a new R Markdown file and call it **Chunk 4**.

Type the code in below.

The new code can be found on line 29 and 31.

```

---
title: "Chunk 4"
output: html_document
---

```{r loading, warning=F, message=F, echo=F}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")

# Cleaning up column names
colnames(payroll) <- make.names(colnames(payroll))

# Cleaning out dollar signs and commas so it'll convert to numbers correctly
payroll$TOTAL.EARNINGS <- gsub("\\$", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- gsub(",", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- as.numeric(payroll$TOTAL.EARNINGS)

# Narrowing down the scope of the data
payroll_total <- select(payroll, NAME, TITLE, DEPARTMENT, TOTAL.EARNINGS)

most_pay <- payroll_total %>%
  arrange(desc(TOTAL.EARNINGS)) %>%
  head(1)
```

The Boston city employee who was paid the most in 2014 was a `r most_pay$TITLE` at `r
most_pay$DEPARTMENT`.

This person made $`r prettyNum(most_pay$TOTAL.EARNINGS,big.mark=",",scientific=FALSE)`.

```{r display_data, warning=F, message=F, echo=F}
library(DT)
datatable(payroll_total)
```

```

Save the file as 04\_chunk.Rmd and click the knit button.

## Chunk 4

The Boston city employee who was paid the most in 2014 was a Police Officer at Boston Police Department.

This person made \$293,892.

Show  entries

Search:

|   | NAME                  | TITLE                          | DEPARTMENT               | TOTAL.EARNINGS |
|---|-----------------------|--------------------------------|--------------------------|----------------|
| 1 | Darosa, Baltazar      | Police Officer                 | Boston Police Department | 293892         |
| 2 | Creaven, Jacqueline D | Police Lieutenant              | Boston Police Department | 289121         |
| 3 | Cawley, Stephen C     | Police Lieutenant              | Boston Police Department | 280904         |
| 4 | Danilecki, John H     | Police Captain                 | Boston Police Department | 265203         |
| 5 | Kervin, Timothy M.    | Police Lieutenant/Hdq Dispatch | Boston Police Department | 261815         |
| 6 | Hosein, Haseeb        | Police Lieutenant              | Boston Police Department | 255020         |

This type of self-generating analysis is important because if you get the next year of payroll data, running

this report will sub in the new city employee who makes the most money automatically.

### **Pretty tables**

Make pretty tables with the **knitr** package and the **kable()** function.

Create a new R Markdown file and call it **Chunk 5**.

Type the code in below.

The new code can be found all the way down on line 60 and 61.

```

---
title: "Chunk 5"
output: html_document
---

# Departments with the highest average pay

```{r loading, warning=F, message=F, echo=F}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")

```{r cleaning_data, warning=F, echo=F}

colnames(payroll) <- make.names(colnames(payroll))

payroll$REGULAR <- gsub("\\$", "", payroll$REGULAR)
payroll$REGULAR <- gsub(",", "", payroll$REGULAR)
payroll$REGULAR <- as.numeric(payroll$REGULAR)
payroll$RETRO <- gsub("\\$", "", payroll$RETRO)
payroll$RETRO <- gsub(",", "", payroll$RETRO)
payroll$RETRO <- as.numeric(payroll$RETRO)
payroll$OTHER <- gsub("\\$", "", payroll$OTHER)
payroll$OTHER <- gsub(",", "", payroll$OTHER)
payroll$OTHER <- as.numeric(payroll$OTHER)
payroll$OTHER <- gsub("\\$", "", payroll$OTHER)
payroll$OTHER <- gsub(",", "", payroll$OTHER)
payroll$OTHER <- as.numeric(payroll$OTHER)
payroll$OVERTIME <- gsub("\\$", "", payroll$OVERTIME)
payroll$OVERTIME <- gsub(",", "", payroll$OVERTIME)
payroll$OVERTIME <- as.numeric(payroll$OVERTIME)
payroll$INJURED <- gsub("\\$", "", payroll$INJURED)
payroll$INJURED <- gsub(",", "", payroll$INJURED)
payroll$INJURED <- as.numeric(payroll$INJURED)
payroll$DETAIL <- gsub("\\$", "", payroll$DETAIL)
payroll$DETAIL <- gsub(",", "", payroll$DETAIL)
payroll$DETAIL <- as.numeric(payroll$DETAIL)
payroll$QUINN <- gsub("\\$", "", payroll$QUINN)
payroll$QUINN <- gsub(",", "", payroll$QUINN)
payroll$QUINN <- as.numeric(payroll$QUINN)
payroll$TOTAL.EARNINGS <- gsub("\\$", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- gsub(",", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- as.numeric(payroll$TOTAL.EARNINGS)

```{r analysis, warning=F, message=F, echo=F}
top5 <- payroll %>%
group_by(DEPARTMENT) %>%
summarize(Average.Earnings=mean(TOTAL.EARNINGS, na.rm=T)) %>%
arrange(desc(Average.Earnings)) %>%
head(5)

```

```{r table, warning=F, echo=F}
library(knitr)
kable(top5)
```

```

Save the file as 05\_chunk.Rmd and click the knit button.



## Chunk 5

### Departments with the highest average pay

| DEPARTMENT                 | Average.Earnings |
|----------------------------|------------------|
| Boston Police Department   | 136454.9         |
| Mayor's Office-Public Info | 133708.0         |
| Law Department             | 133662.2         |
| Environment Department     | 131208.0         |
| Mayor's Office             | 127376.4         |

#### Change theme and style

Change the appearance and style of the HTML document by changing the theme up top.

Options from the Bootswatch theme library includes:

- `default`
- `cerulean`
- `journal`
- `cosmo`

highlights (for the code syntax)

- `tango`
- `pygments`
- `kate`

Create a new R Markdown file and call it **Chunk 6**.

Type the code in below.

The new code is at the top in the YAML section.

```

---
title: "Chunk 6"
author: "Andrew"
date: "7/23/2018"
output:
  html_document:
    theme: united
    highlight: espresso
---

```{r loading, warning=F, message=F, echo=F}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")

```{r cleaning data, warning=F, echo=F}
colnames(payroll) <- make.names(colnames(payroll))

payroll$REGULAR <- gsub("\\$", "", payroll$REGULAR)
payroll$REGULAR <- gsub(",", "", payroll$REGULAR)
payroll$REGULAR <- as.numeric(payroll$REGULAR)
payroll$RETRO <- gsub("\\$", "", payroll$RETRO)
payroll$RETRO <- gsub(",", "", payroll$RETRO)
payroll$RETRO <- as.numeric(payroll$RETRO)
payroll$OTHER <- gsub("\\$", "", payroll$OTHER)
payroll$OTHER <- gsub(",", "", payroll$OTHER)
payroll$OTHER <- as.numeric(payroll$OTHER)
payroll$OTHER <- gsub("\\$", "", payroll$OTHER)
payroll$OTHER <- gsub(",", "", payroll$OTHER)
payroll$OTHER <- as.numeric(payroll$OTHER)
payroll$OVERTIME <- gsub("\\$", "", payroll$OVERTIME)
payroll$OVERTIME <- gsub(",", "", payroll$OVERTIME)
payroll$OVERTIME <- as.numeric(payroll$OVERTIME)
payroll$INJURED <- gsub("\\$", "", payroll$INJURED)
payroll$INJURED <- gsub(",", "", payroll$INJURED)
payroll$INJURED <- as.numeric(payroll$INJURED)
payroll$DETAIL <- gsub("\\$", "", payroll$DETAIL)
payroll$DETAIL <- gsub(",", "", payroll$DETAIL)
payroll$DETAIL <- as.numeric(payroll$DETAIL)
payroll$QUINN <- gsub("\\$", "", payroll$QUINN)
payroll$QUINN <- gsub(",", "", payroll$QUINN)
payroll$QUINN <- as.numeric(payroll$QUINN)
payroll$TOTAL.EARNINGS <- gsub("\\$", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- gsub(",", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- as.numeric(payroll$TOTAL.EARNINGS)

```{r analysis, warning=F, message=F}
top5 <- payroll %>%
group_by(DEPARTMENT) %>%
summarize(Average.Earnings=mean(TOTAL.EARNINGS, na.rm=T)) %>%
arrange(desc(Average.Earnings)) %>%
  head(5)

---

```{r table, warning=F, echo=F}
library(knitr)
kable(top5)
```

```

Save the file as 06\_chunk.Rmd and click the knit button.

# Chunk 6

Andrew

7/23/2018

## Departments with the highest average pay

```
top5 <- payroll %>%
  group_by(DEPARTMENT) %>%
  summarize(Average.Earnings=mean(TOTAL.EARNINGS, na.rm=T)) %>%
  arrange(desc(Average.Earnings)) %>%
  head(5)
```

| DEPARTMENT                 | Average.Earnings |
|----------------------------|------------------|
| Boston Police Department   | 136454.9         |
| Mayor's Office-Public Info | 133708.0         |
| Law Department             | 133662.2         |
| Environment Department     | 131208.0         |
| Mayor's Office             | 127376.4         |

### Table of contents

Add a floating table of contents by changing `html_document` to `toc: true` and `toc_float: true`.

Create a new R Markdown file and call it **Chunk 7**.

Type the code in below.

The new code is at the top in the YAML section.

```

---
title: "Chunk 7"
author: "Andrew"
date: "3/10/2018"
output:
  html_document:
    toc: true
    toc_float: true
---

# Boston employee pay in 2014

```{r loading, warning=F, message=F, echo=F}
# load packages
library(tidyverse)

# Loading the Boston city payroll
payroll <- read_csv("data/bostonpayroll2013.csv")

colnames(payroll) <- make.names(colnames(payroll))

payroll_total <- select(payroll, NAME, TITLE, DEPARTMENT, TOTAL.EARNINGS)

```{r display_data, warning=F, message=F, echo=F}
library(DT)
datatable(payroll_total)
```

# Departments with the highest average pay

```{r cleaning_data, warning=F, echo=F}
payroll$REGULAR <- gsub("\\$", "", payroll$REGULAR)
payroll$REGULAR <- gsub(",", "", payroll$REGULAR)
payroll$REGULAR <- as.numeric(payroll$REGULAR)
payroll$RETRO <- gsub("\\$", "", payroll$RETRO)
payroll$RETRO <- gsub(",", "", payroll$RETRO)
payroll$RETRO <- as.numeric(payroll$RETRO)
payroll$OTHER <- gsub("\\$", "", payroll$OTHER)
payroll$OTHER <- gsub(",", "", payroll$OTHER)
payroll$OTHER <- as.numeric(payroll$OTHER)
payroll$OTHER <- gsub("\\$", "", payroll$OTHER)
payroll$OTHER <- gsub(",", "", payroll$OTHER)
payroll$OTHER <- as.numeric(payroll$OTHER)
payroll$OVERTIME <- gsub("\\$", "", payroll$OVERTIME)
payroll$OVERTIME <- gsub(",", "", payroll$OVERTIME)
payroll$OVERTIME <- as.numeric(payroll$OVERTIME)
payroll$INJURED <- gsub("\\$", "", payroll$INJURED)
payroll$INJURED <- gsub(",", "", payroll$INJURED)
payroll$INJURED <- as.numeric(payroll$INJURED)
payroll$DETAIL <- gsub("\\$", "", payroll$DETAIL)
payroll$DETAIL <- gsub(",", "", payroll$DETAIL)
payroll$DETAIL <- as.numeric(payroll$DETAIL)
payroll$QUINN <- gsub("\\$", "", payroll$QUINN)
payroll$QUINN <- gsub(",", "", payroll$QUINN)
payroll$QUINN <- as.numeric(payroll$QUINN)
payroll$TOTAL.EARNINGS <- gsub("\\$", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- gsub(",", "", payroll$TOTAL.EARNINGS)
payroll$TOTAL.EARNINGS <- as.numeric(payroll$TOTAL.EARNINGS)

```

```{r analysis, warning=F, message=F, echo=F}
top5 <- payroll %>%
group_by(DEPARTMENT) %>%
summarize(Average.Earnings=mean(TOTAL.EARNINGS, na.rm=T)) %>%
arrange(desc(Average.Earnings)) %>%
head(5)

```

```{r table, warning=F, echo=F}
library(knitr)
kable(top5)
```

```

Save the file as 07\_chunk.Rmd and click the knit button.

Boston employee pay in 2014

Departments with the highest  
average pay

## Chunk 7

Andrew

3/10/2018

### Boston employee pay in 2014

Show 10 entries

Search:

|   | NAME                  | TITLE                          | DEPARTMENT               | TOTAL.EARNINGS |
|---|-----------------------|--------------------------------|--------------------------|----------------|
| 1 | Darosa, Baltazar      | Police Officer                 | Boston Police Department | \$293,892      |
| 2 | Creaven, Jacqueline D | Police Lieutenant              | Boston Police Department | \$289,121      |
| 3 | Cawley, Stephen C     | Police Lieutenant              | Boston Police Department | \$280,904      |
| 4 | Danilecki, John H     | Police Captain                 | Boston Police Department | \$265,203      |
| 5 | Kervin, Timothy M.    | Police Lieutenant/Hdq Dispatch | Boston Police Department | \$261,815      |

#### Next steps?

Exporting as a PDF will require LaTeX installed first \* Get it from [latex-project.org](http://latex-project.org) or MacTex

Check out all the features of R Markdown at RStudio

**Publish your results to Github pages**