

Data Science for Everyone – Intro to Coding with R

Dr. Ab Mosca (they/them)

Slides based off slides courtesy of Jordan Crouser (<https://jcrouser.github.io/>)

Plan for Today

- Practice coding with the R programming language

R

R Studio

R Markdown

- R is a programming language
- R Studio is an interface for coding
- R Markdown is a type of file that allows you to intersperse writing, code, and output

R Studio

- Launch R Studio:
- <https://posit.cloud/>

The screenshot displays the Posit Cloud web interface. The top navigation bar includes the Posit Cloud logo, a user profile icon for 'Ab Mosca', and tabs for 'Your Workspace', 'Content', 'Usage', and 'About'. The left sidebar contains sections for 'Spaces' (with 'Your Workspace' selected), 'Learn' (with links to Guide, What's New, Primers, and Cheat Sheets), 'Help' (with links to Current System Status and Posit Community), and 'Info' (with links to Plans & Pricing and Terms and Conditions). The main content area, titled 'Your Content (0)', shows a list of items: 'Your Content', 'Archive', and 'Trash'. A 'New Project' button is visible in the top right of the content area. The bottom of the interface features a dark blue footer with the Posit Cloud logo, links for 'Terms' and 'Status', social media icons, and the copyright notice '© 2022 Posit Software, PBC'.

R Studio

- Move to your personal space

The screenshot displays the Posit Cloud web interface. On the left sidebar, under the 'Spaces' section, the 'Your Workspace' is listed, and the 'ab' space (owned by Ab Mosca) is highlighted with a red rectangular box. The main content area shows 'Your Content' for the selected space, which is currently empty (0 items). Below the content list are filters for 'TYPE', 'ACCESS', and 'SORT', along with a search bar. The footer of the interface includes the Posit Cloud logo, links for 'Terms' and 'Status', social media icons, and the copyright notice '© 2022 Posit Software, PBC'.

R Studio

- Start a new project

The screenshot displays the Posit Cloud web interface. On the left is a sidebar with navigation links: Spaces (Your Workspace, ab Ab Mosca, + New Space), Learn (Guide, What's New, Primers, Cheat Sheets), Help (Current System Status, Posit Community), and Info (Plans & Pricing, Terms and Conditions). The main content area is titled 'ab Ab Mosca' and has tabs for Content, Data, Members, Usage, and About. Under the 'Content' tab, there's a list of items: All Content, Your Content, Templates, Archive, and Trash. A 'New List' button is also present. Below this list is a message: 'This space can have up to 9 more projects or outputs. To create additional content, upgrade your account.' with an 'Upgrade Account' button. The 'All Content' section shows one item, 'Untitled Project', which is an RStudio Project, owned by 'Ab Mosca', set to 'Private', and created on 'Aug 7, 2023 10:39 AM'. A 'New Project' button with a dropdown arrow is highlighted with a red rectangle in the top right of the content area. The footer includes the Posit Cloud logo, social media icons, and the copyright notice '© 2022 Posit Software, PBC'.

R Studio

- Start a new project

The screenshot displays the Posit Cloud web interface. On the left is a sidebar with navigation links: Spaces (Your Workspace, ab Ab Mosca, New Space), Learn (Guide, What's New, Primers, Cheat Sheets), Help (Current System Status, Posit Community), and Info (Plans & Pricing, Terms and Conditions). The main content area shows the 'ab Ab Mosca' workspace with tabs for Content, Data, Members, Usage, and About. Under the 'Content' tab, there's a list of items: All Content, Your Content, Templates, Archive, Trash, and New List. A message states: 'This space can have up to 9 more projects or outputs. To create additional content, upgrade your account.' with an 'Upgrade Account' button. A 'New Project' dropdown menu is open on the right, listing options: New Project from Template, New RStudio Project (highlighted with a red box), New Jupyter Project, and New Project from Git Repository. The main content area also shows an 'Untitled Project' card for an RStudio Project, owned by 'Ab Mosca', set to 'Private', and created on 'Aug 7, 2023 10:39 AM'. The footer contains the Posit Cloud logo, social media icons, and the copyright notice '© 2022 Posit Software, PBC'.

R Studio

- New project

The screenshot displays the RStudio Cloud web interface. On the left is a sidebar with navigation links: Spaces (Your Workspace, ab Ab Mosca, + New Space), Learn (Guide, What's New, Primers, Cheat Sheets), Help (Current System Status, Posit Community), and Info (Plans & Pricing, Terms and Conditions). The main area is titled 'Ab / Untitled Project' with a subtitle 'Click to name your project'. It features a menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help) and a toolbar with icons for file operations and a 'Go to file/function' search bar. The 'Console' tab is active, showing the R version 4.3.1 (2023-06-16) and the standard R startup messages. The 'Environment' tab on the right shows an empty global environment. At the bottom right, a file explorer shows the project directory containing '.Rhistory' and 'project.Rproj' files.

posit Cloud

Ab / Untitled Project Click to name your project

R 4.3.1

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Console Terminal Background Jobs

R 4.3.1 · /cloud/project/

R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Environment History Connections Tutorial

Import Dataset 126 MiB

R Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

Folder Blank File Upload Delete Rename

Cloud project

	Name	Size	Modified
	..		
<input type="checkbox"/>	.Rhistory	0 B	Sep 18, 2023, 10:24 AM
<input type="checkbox"/>	project.Rproj	205 B	Sep 18, 2023, 10:25 AM

R Studio

- Console

The screenshot displays the RStudio Cloud web interface. The top bar shows the user 'Ab / Untitled Project' and the R version 'R 4.3.1'. The left sidebar contains navigation options like 'Spaces', 'Your Workspace', and 'Learn'. The main area is divided into several panes: 'Console' (highlighted with a red border), 'Terminal', 'Background Jobs', 'Environment', 'History', 'Connections', 'Tutorial', 'Files', 'Plots', 'Packages', 'Help', 'Viewer', and 'Presentation'. The 'Console' pane shows the R startup message and a prompt '> |'. The 'Environment' pane shows 'Environment is empty'. The 'Files' pane shows a list of files in the 'project' directory.

Console Output:

```
R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

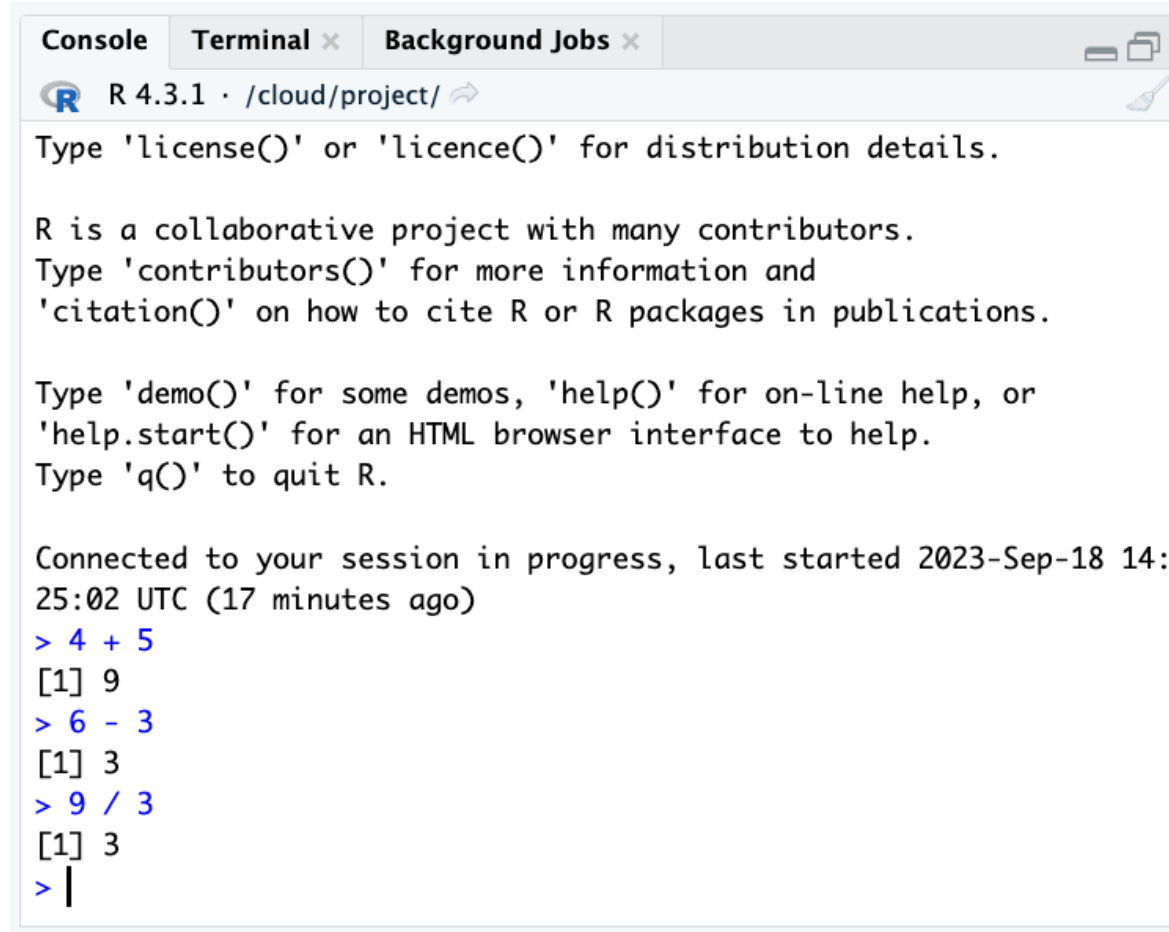
> |
```

Files List:

Name	Size	Modified
..		
.Rhistory	0 B	Sep 18, 2023, 10:24 AM
project.Rproj	205 B	Sep 18, 2023, 10:25 AM

R Studio

- Console



```
Console Terminal x Background Jobs x
R 4.3.1 · /cloud/project/
Type 'license()' or 'licence()' for distribution details.

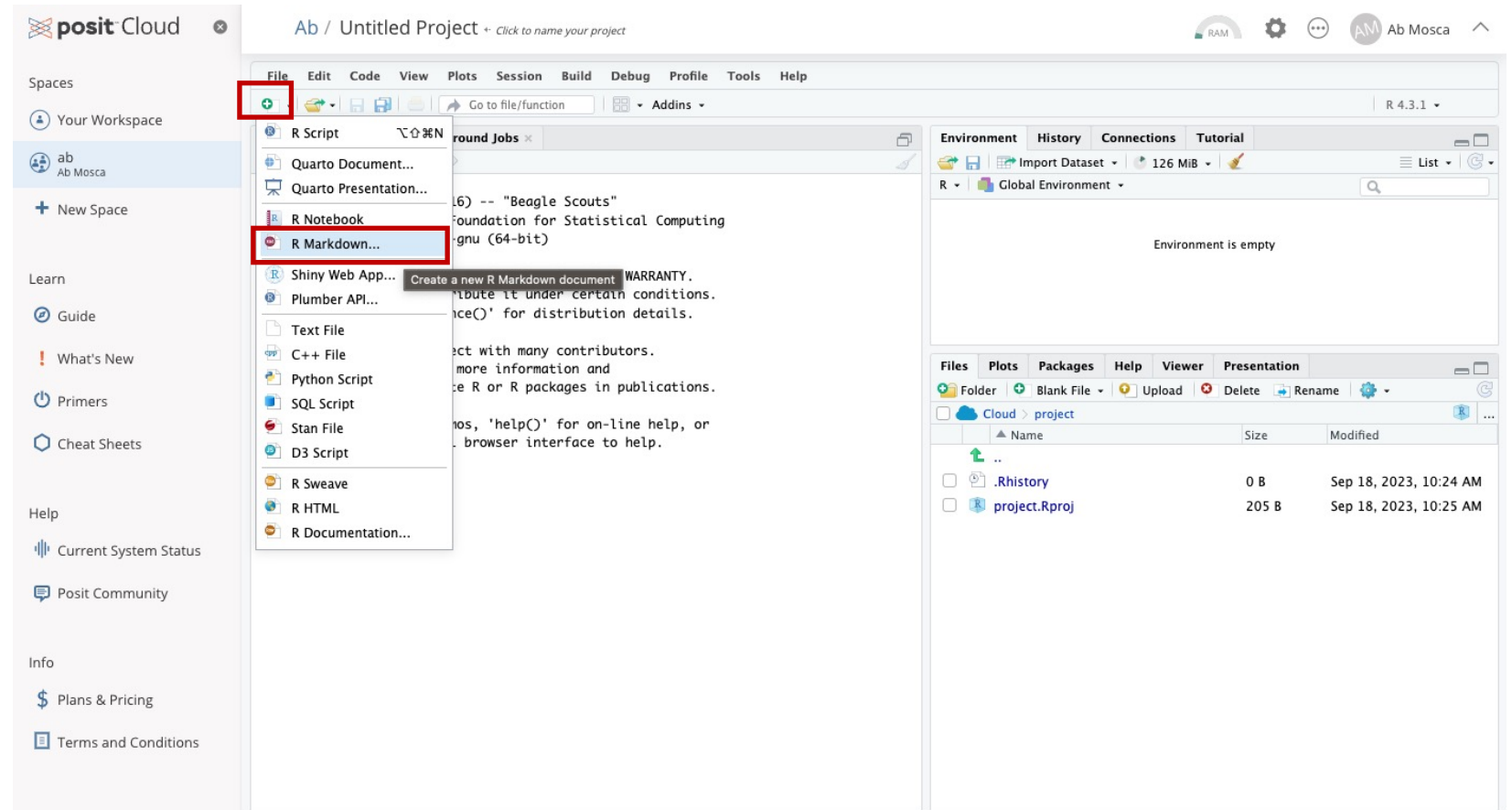
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Connected to your session in progress, last started 2023-Sep-18 14:
25:02 UTC (17 minutes ago)
> 4 + 5
[1] 9
> 6 - 3
[1] 3
> 9 / 3
[1] 3
> |
```

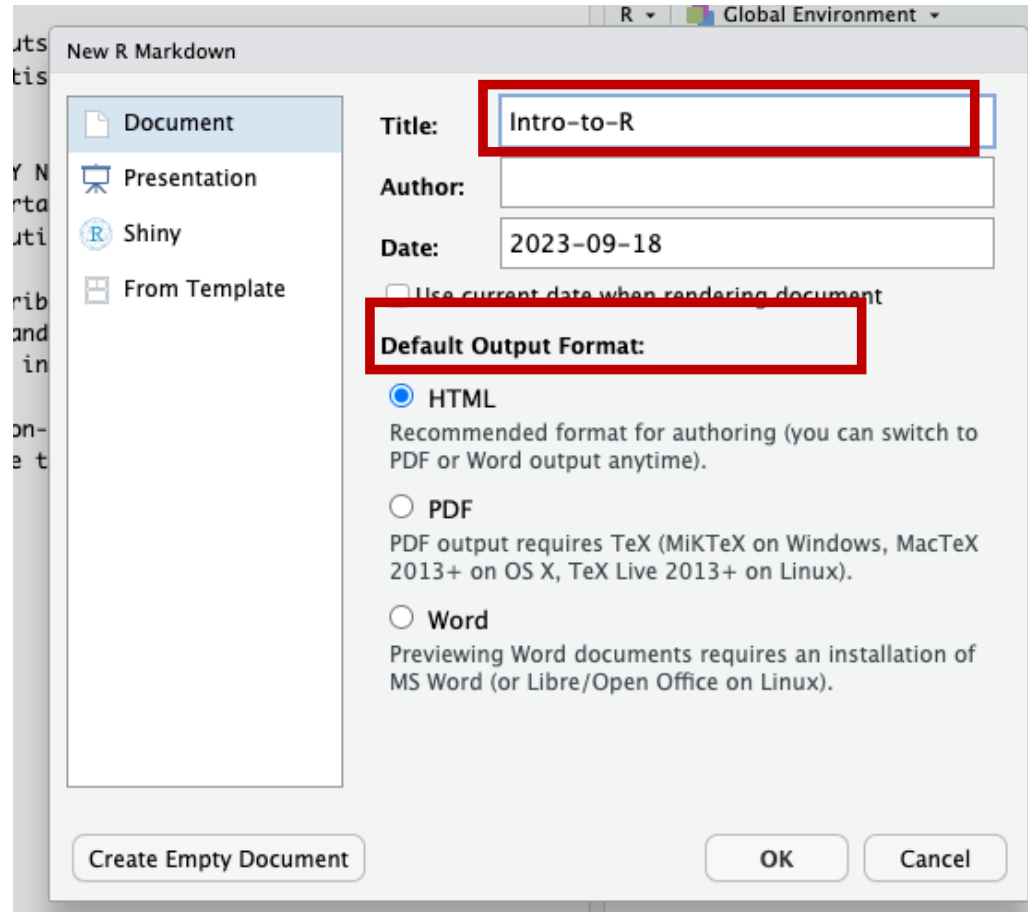
R Studio

- New file



R Studio

- New file



R Studio

- New file

The screenshot displays the RStudio web interface within a Posit Cloud environment. The top bar shows the user 'Ab Mosca' and the project name 'Untitled Project'. The left sidebar contains navigation links for Spaces, Your Workspace, a New Space button, Learn (Guide, What's New, Primers, Cheat Sheets), Help (Current System Status, Posit Community), and Info (Plans & Pricing, Terms and Conditions).

The main editor window shows a new file 'Untitled1' with the following content:

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML,
14 PDF, and MS Word documents. For more details on using R Markdown see
15 <http://rmarkdown.rstudio.com>.
16
17 When you click the Knit button a document will be generated that includes both content
```

The bottom console window shows the R version and system information:

```
R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
```

The right sidebar contains the Environment pane (showing 'Environment is empty'), the Files pane (showing a file explorer with '..' and 'project.Rproj'), and the Plots, Packages, Help, Viewer, and Presentation panes.

R Studio

- New file

The screenshot displays the RStudio Cloud web interface. On the left sidebar, the 'Spaces' section shows 'Your Workspace' and a user profile 'ab Ab Mosca'. Below this is a 'Learn' section with links to 'Guide', 'What's New', 'Primers', and 'Cheat Sheets'. The 'Help' section includes 'Current System Status', 'Posit Community', and 'Info' with links to 'Plans & Pricing' and 'Terms and Conditions'.

The main editor area is titled 'Ab / Untitled Project' with a subtitle 'Click to name your project'. It features a menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help) and a toolbar with icons for file operations and running code. A red box highlights the 'New File' icon (a document with a plus sign) in the toolbar. The editor window shows a new file named 'Untitled1' containing R Markdown code:

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML,
14 PDF, and MS Word documents. For more details on using R Markdown see
15 <http://rmarkdown.rstudio.com>.
16
17 When you click the Knit button a document will be generated that includes both content
```

Below the editor is a 'Console' tab showing the R version and system information:

```
R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
```

On the right side, the 'Environment' pane shows 'R 4.3.1' and 'Global Environment'. The 'Files' pane shows a list of files in the 'project' directory:

Name	Size	Modified
..		
.Rhistory	0 B	Sep 18, 2023, 10:24 AM
project.Rproj	205 B	Sep 18, 2023, 10:25 AM

R Studio

- New file - SAVE

The screenshot displays the RStudio IDE interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The toolbar below the menu bar contains icons for creating a new file, saving, opening, and running code. The 'Save' icon (a floppy disk) is highlighted with a red rectangle. The main editor window shows a new R Markdown file titled 'Untitled1.Rmd' with the following content:

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 {r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML,
14 PDF, and MS Word documents. For more details on using R Markdown see
15 <http://rmarkdown.rstudio.com>.
16
17 When you click the **Knit** button a document will be generated that includes both content
```

The bottom pane shows the Console with the R version 4.3.1 startup message:

```
R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
```

The right-hand side of the interface shows the Environment pane (empty), the Files pane (listing the project directory and files like .Rhistory and project.Rproj), and the Plots, Packages, Help, Viewer, and Presentation panes.

R Studio

- New file - saved

The screenshot displays the RStudio interface within the Posit Cloud environment. The main editor window shows a new R Markdown file titled "Intro-to-R.Rmd", which is highlighted with a red box. The file content includes a YAML header with title, output, and date, followed by R code chunks for setup and a cars dataset. The right-hand pane shows the "Files" tab, also highlighted with a red box, listing the project files: ".Rhistory", "Intro-to-R.html", "Intro-to-R.Rmd", and "project.Rproj". The "Intro-to-R.Rmd" file is selected. The bottom pane shows the console output for the R session.

Files | Plots | Packages | Help | Viewer | Presentation

Name	Size	Modified
..		
.Rhistory	0 B	Sep 18, 2023, 10:24 AM
Intro-to-R.html	631.2 KB	Sep 18, 2023, 10:28 AM
Intro-to-R.Rmd	817 B	Sep 18, 2023, 10:27 AM
project.Rproj	205 B	Sep 18, 2023, 10:25 AM

Console | Terminal | Background Jobs

R 4.3.1 · /cloud/project/

R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

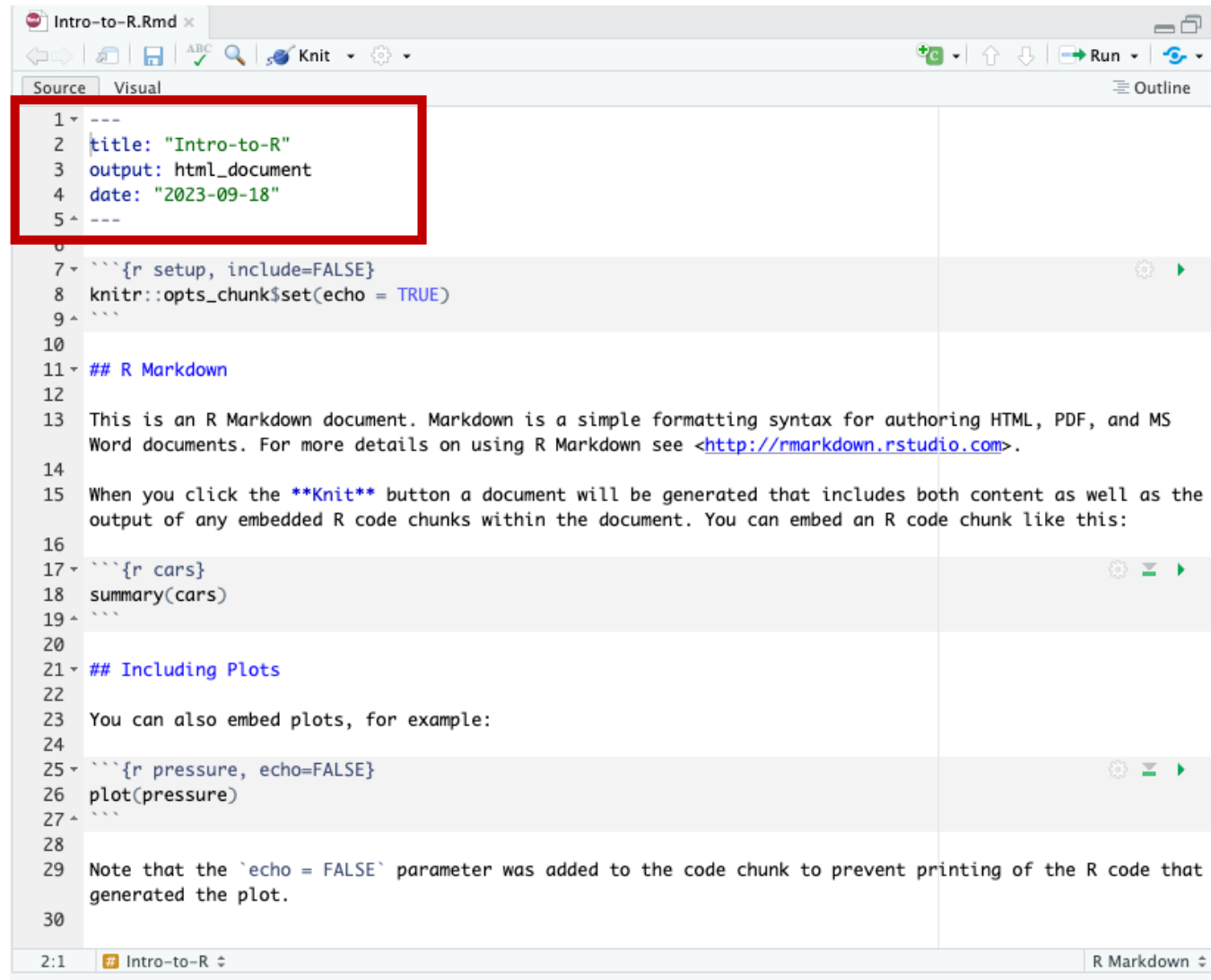
R Studio

- New file - saved

The screenshot displays the Posit Cloud web interface. On the left is a sidebar with navigation links: Spaces (Your Workspace, ab Ab Mosca, New Space), Learn (Guide, What's New, Primers, Cheat Sheets), Help (Current System Status, Posit Community), and Info (Plans & Pricing, Terms and Conditions). The main header shows the user 'ab Ab Mosca' with tabs for Content, Data, Members, Usage, and About. The 'Content' tab is active, showing 'All Content (1)'. A list contains one item, 'Intro-to-R', which is an RStudio Project owned by 'Ab Mosca', set to 'Private', and created on 'Sep 18, 2023 10:24 AM'. A notification box on the left states: 'This space can have up to 9 more projects or outputs. To create additional content, upgrade your account. Upgrade Account'. The footer includes the Posit Cloud logo, links for Terms and Status, social media icons, and the copyright notice '© 2022 Posit Software, PBC'.

R Markdown

- Meta data



The screenshot shows the RStudio interface with an R Markdown document titled "Intro-to-R.Rmd". The document is displayed in the "Source" view. A red rectangular box highlights the meta data section at the top of the document, which includes the following code:

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
```

Below the meta data, the document content includes an R code chunk for setting up the environment, a section titled "## R Markdown" with explanatory text, another R code chunk for summarizing cars, a section titled "## Including Plots" with explanatory text, and a final R code chunk for plotting pressure data. The status bar at the bottom indicates the current position is at line 2, column 1.

R Markdown

- Code chunks

The screenshot shows an R Markdown document titled "Intro-to-R.Rmd" in the RStudio editor. The document is divided into sections by headers. Three code chunks are highlighted with red boxes, and callouts explain their syntax:

- Chunk 1 (lines 6-9):** A setup chunk. The callout "type of code" points to the opening triple backticks. The callout "short name" points to the `{r}` identifier. The callout "other options" points to the `include=FALSE` and `knitr::opts_chunk$set(echo = TRUE)` options.
- Chunk 2 (lines 11-14):** A text chunk. The callout "type of code" points to the opening triple backticks. The callout "short name" points to the `{r}` identifier.
- Chunk 3 (lines 16-19):** A code chunk. The callout "type of code" points to the opening triple backticks. The callout "short name" points to the `{r}` identifier.
- Chunk 4 (lines 21-24):** A code chunk. The callout "type of code" points to the opening triple backticks. The callout "short name" points to the `{r}` identifier. The callout "other options" points to the `echo=FALSE` option.

The document content includes a title, output format, date, a setup chunk, a text chunk, a code chunk, a section header "## Including Plots", a text chunk, another code chunk, and a concluding note.

R Markdown

- Code chunks

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS
14 Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
15
16 When you click the Knit button a document will be generated that includes both content as well as the
17 output of any embedded R code chunks within the document. You can embed an R code chunk like this:
18
19 ```{r cars}
20 summary(cars)
21 ```
22
23 ## Including Plots
24
25 You can also embed plots, for example:
26
27 ```{r pressure, echo=FALSE}
28 plot(pressure)
29 ```
30
31 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that
32 generated the plot.
```

run
chunk

run all
preceding
chunks

R Markdown

- Text

The screenshot shows an R Markdown document titled "Intro-to-R.Rmd" in the RStudio editor. The document is in "Source" view. The code is as follows:

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5
6 # Setup
7 knitr::opts_chunk$set(include=FALSE)
8 knitr::opts_chunk$set(echo = TRUE)
9
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
14
15 Whenever you click the Knit button 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:
16
17 ```{r cars}
18 summary(cars)
19 ```
20
21 ## Including Plots
22
23 You can also embed plots, for example:
24
25 ```{r pressure, echo=FALSE}
26 plot(pressure)
27 ```
28
29 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```

Callouts explain the following syntax:

- `##` == heading
- `**` == bold
- `<>` == link
- ``` == code

R Markdown

- Knit

output
report

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS
14 Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
15
16 When you click the **Knit** button a document will be generated that includes both content as well as the
17 output of any embedded R code chunks within the document. You can embed an R code chunk like this:
18
19 ```{r cars}
20 summary(cars)
21 ```
22
23 ## Including Plots
24
25 You can also embed plots, for example:
26
27 ```{r pressure, echo=FALSE}
28 plot(pressure)
29 ```
30
31 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that
32 generated the plot.
```

R Markdown

- Knit

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS
14 Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
15
16 When you click the Knit button a document will be generated that includes both content as well as the
17 output of any embedded R code chunks within the document. You can embed an R code chunk like this:
18
19 ```{r cars}
20 summary(cars)
21 ```
22
23 ## Including Plots
24
25 You can also embed plots, for example:
26
27 ```{r pressure, echo=FALSE}
28 plot(pressure)
29 ```
30
31 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that
32 generated the plot.
```

[a2725c493c49447e91580193bcead4fa.app.posit.cloud/?view=rmarkdown](#)
Intro-to-R.html Open in Browser Find

Intro-to-R

2023-09-18

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

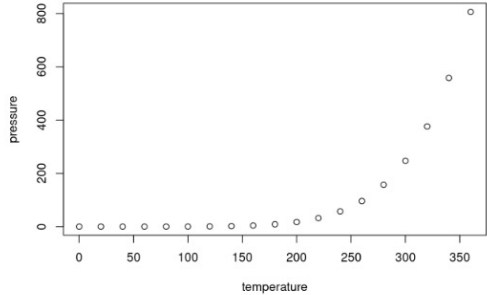
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
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

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

R Markdown

- Knit

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS
14 Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
15
16 When you click the Knit button a document will be generated that includes both content as well as the
17 output of any embedded R code chunks within the document. You can embed an R code chunk like this:
18
19 ```{r cars}
20 summary(cars)
21 ```
22
23 ## Including Plots
24
25 You can also embed plots, for example:
26
27 ```{r pressure, echo=FALSE}
28 plot(pressure)
29 ```
30
31 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that
32 generated the plot.
```

Intro-to-R.html

Open in Browser

Find

Intro-to-R

2023-09-18

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

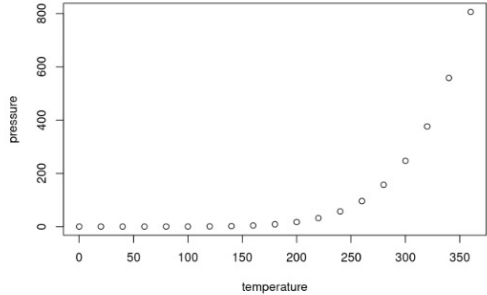
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

	speed	dist
## Min.	4.0	2.00
## 1st Qu.	12.0	26.00
## Median	15.0	36.00
## Mean	15.4	42.98
## 3rd Qu.	19.0	56.00
## Max.	25.0	120.00

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

R Markdown

- Knit

```
1 ---
2 title: "Intro-to-R"
3 output: html_document
4 date: "2023-09-18"
5 ---
6
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
10
11 ## R Markdown
12
13 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS
14 Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.
15
16 When you click the Knit button a document will be generated that includes both content as well as the
17 output of any embedded R code chunks within the document. You can embed an R code chunk like this:
18
19 ```{r cars}
20 summary(cars)
21 ```
22
23 ## Including Plots
24
25 You can also embed plots, for example:
26
27 ```{r pressure, echo=FALSE}
28 plot(pressure)
29 ```
30
31 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that
32 generated the plot.
```

[a2725c493c49447e91580193bcead4fa.app.posit.cloud/?view=rmarkdown](#)
Intro-to-R.html | Open in Browser | Find

Intro-to-R

2023-09-18

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

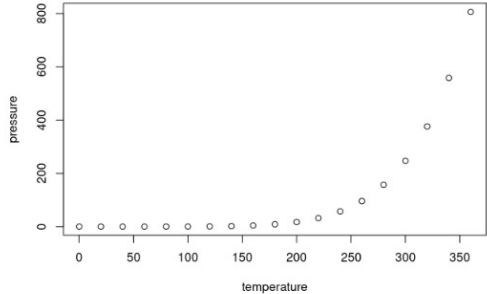
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
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

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

R

- Packages

```
7 {r load-packages, message = FALSE, eval=TRUE}  
8 library(tidyverse)  
9 {  
0
```

R

- Data

The screenshot shows an RStudio window with two tabs: 'Intro-to-R.Rmd' and 'R data sets'. The main editor displays the output of the `data()` function for the 'datasets' package, listing 50 datasets in two columns. The datasets include various biological, economic, and social data sets. The console at the bottom shows the command `> data()` being entered.

```
Intro-to-R.Rmd x R data sets x

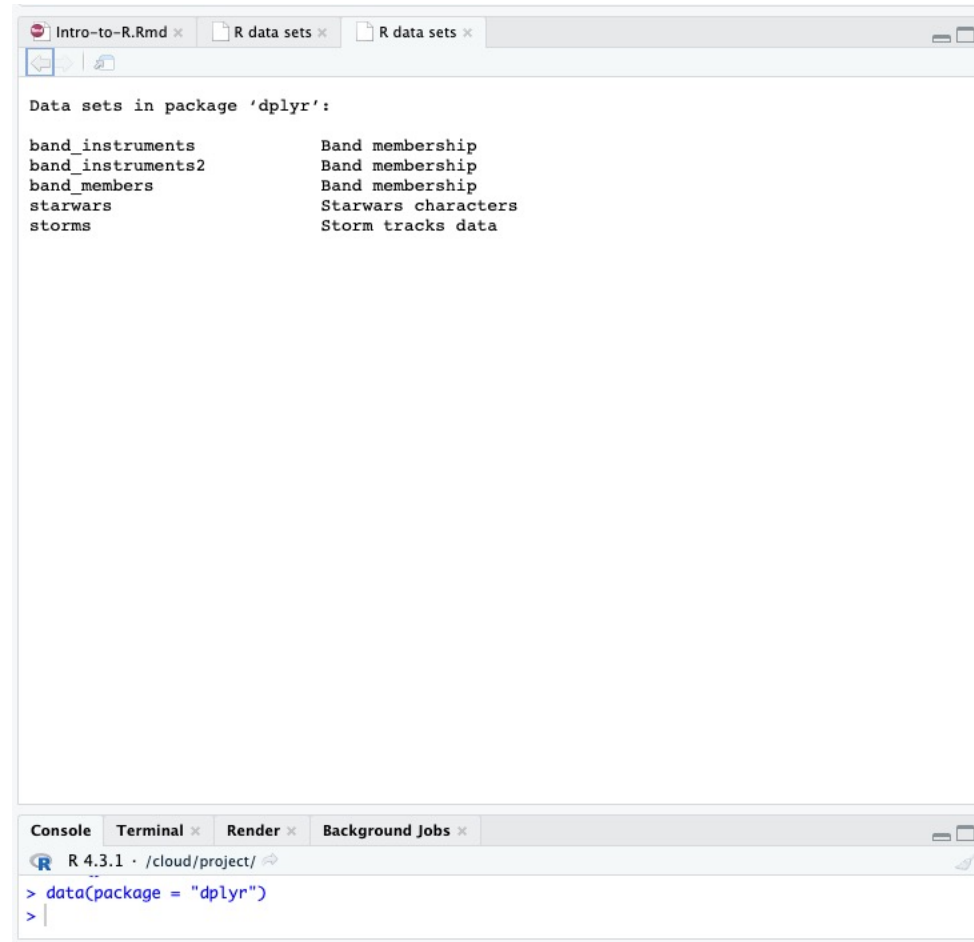
Data sets in package 'datasets':

AirPassengers      Monthly Airline Passenger Numbers 1949-1960
BJsales            Sales Data with Leading Indicator
BJsales.lead (BJsales) Sales Data with Leading Indicator
BOD                Biochemical Oxygen Demand
CO2                Carbon Dioxide Uptake in Grass Plants
ChickWeight        Weight versus age of chicks on different diets
DNase              Elisa assay of DNase
EuStockMarkets     Daily Closing Prices of Major European Stock Indices,
                  1991-1998
Formaldehyde       Determination of Formaldehyde
HairEyeColor       Hair and Eye Color of Statistics Students
Harman23.cor       Harman Example 2.3
Harman74.cor       Harman Example 7.4
Indometh            Pharmacokinetics of Indomethacin
InsectSprays       Effectiveness of Insect Sprays
JohnsonJohnson    Quarterly Earnings per Johnson & Johnson Share
LakeHuron          Level of Lake Huron 1875-1972
LifeCycleSavings   Intercountry Life-Cycle Savings Data
Loblolly           Growth of Loblolly pine trees
Nile               Flow of the River Nile
Orange            Growth of Orange Trees
OrchardSprays      Potency of Orchard Sprays
PlantGrowth        Results from an Experiment on Plant Growth
Puromycin          Reaction Velocity of an Enzymatic Reaction
Seatbelts          Road Casualties in Great Britain 1969-84
Theoph             Pharmacokinetics of Theophylline
Titanic            Survival of passengers on the Titanic
ToothGrowth        The Effect of Vitamin C on Tooth Growth in Guinea Pigs
UCBAdmissions      Student Admissions at UC Berkeley
UKDriverDeaths     Road Casualties in Great Britain 1969-84
UKGas              UK Quarterly Gas Consumption
USAccDeaths        Accidental Deaths in the US 1973-1978
USArrests          Violent Crime Rates by US State
USJudgeRatings     Annual Ratings of State Judges in the US Superior

Console Terminal x Render x Background Jobs x
R 4.3.1 · /cloud/project/
> data()
>
```

R

- Data



The screenshot shows an RStudio interface with three tabs at the top: 'Intro-to-R.Rmd', 'R data sets', and 'R data sets'. The main editor window displays the following text:

```
Data sets in package 'dplyr':
```

band_instruments	Band membership
band_instruments2	Band membership
band_members	Band membership
starwars	Starwars characters
storms	Storm tracks data

At the bottom, the console window shows the R prompt and the command:

```
> data(package = "dplyr")  
> |
```

R

- Data – reading a CSV

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code. A red box highlights the chunk starting at line 32:

```
32 {r data}
33 magazines <- read_csv("http://gatonweb.uky.edu/sheather/book/docs/datasets/magazines.csv")
34
```
- Environment Pane:** Shows the variable `magazines` with 204 observations and 5 variables.
- Files Pane:** Shows the project structure with files like `.Rhistory`, `Intro-to-R.html`, `Intro-to-R.Rmd`, and `project.Rproj`.
- Console:** Shows the output of the `read_csv` function, including column specifications and data types.

Two callouts are present:

- A yellow callout pointing to `magazines <-` with the text "variable assignment".
- A yellow callout pointing to `read_csv` with the text "function".

Console Output:

```
R 4.3.1 - /cloud/project/
i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
> magazines <- read_csv("http://gatonweb.uky.edu/sheather/book/docs/datasets/magazines.csv")
Rows: 204 Columns: 5— Column specification
Delimiter: ","
chr (1): Magazine
dbl (4): AdRevenue, AdPages, SubRevenue, NewsRevenue
i Use 'spec()' to retrieve the full column specification for this data.
i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
>
```

R

- Data – reading a CSV

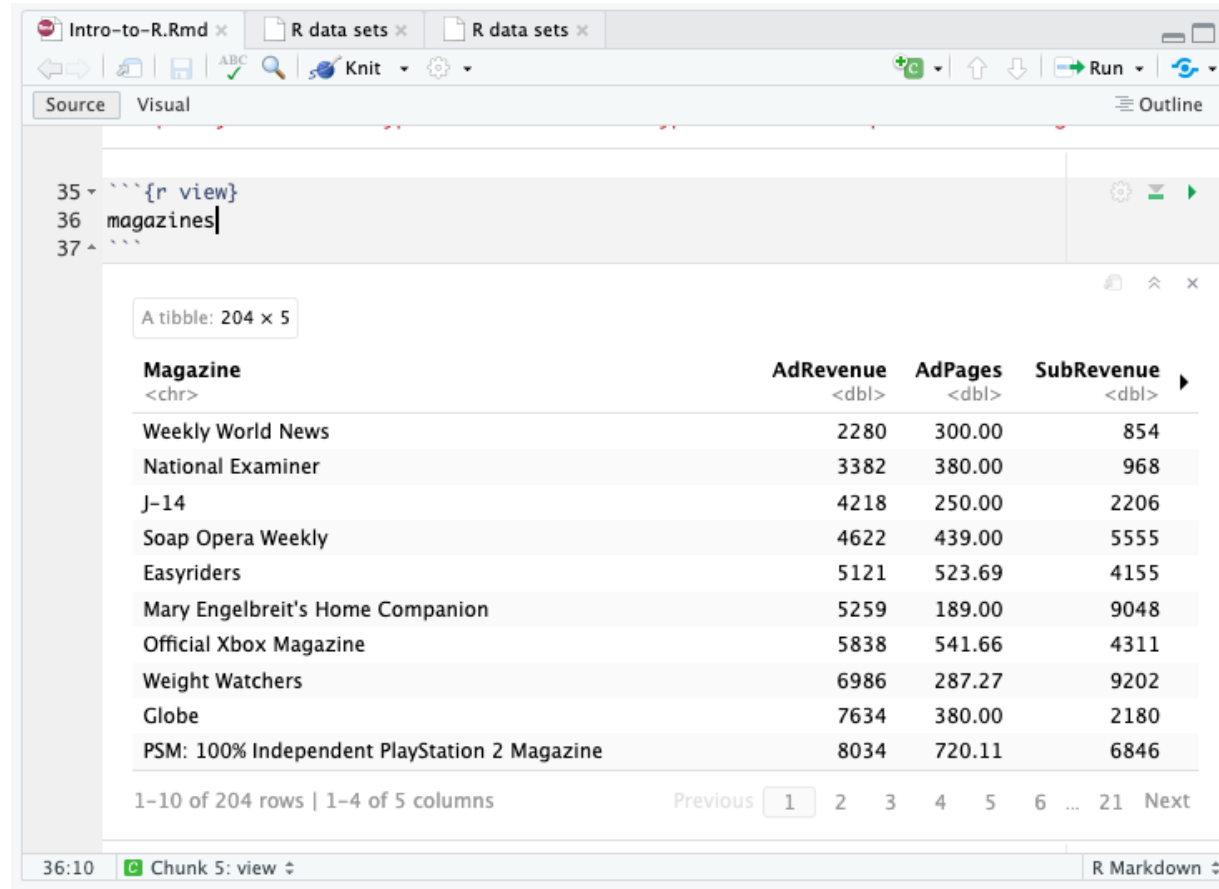
The screenshot displays the RStudio interface with several components highlighted by red boxes and annotated with callouts:

- Source Editor:** A red box highlights the code chunk containing the `read_csv` function call. A callout labeled "function" points to the `read_csv` function name.
- Environment Pane:** A red box highlights the "Data" section, showing the variable `magazines` with 204 observations and 5 variables. A callout labeled "variable assignment" points to the `magazines` variable name.
- Environment Pane:** A red box highlights the "Global Environment" section, showing the variable `magazines` with 204 observations and 5 variables. A callout labeled "saved objects" points to the `magazines` variable name.
- Environment Pane:** A red box highlights the "clear" button, which is used to remove objects from the environment. A callout labeled "clear" points to the button.

The console output at the bottom shows the execution of the code, including the column specification and the resulting data structure:

```
R 4.3.1 - /cloud/project/
i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
> magazines <- read_csv("http://gattoweb.uky.edu/sheather/book/docs/datasets/magazines.csv")
Rows: 204 Columns: 5— Column specification
Delimiter: ","
chr (1): Magazine
dbl (4): AdRevenue, AdPages, SubRevenue, NewsRevenue
i Use 'spec()' to retrieve the full column specification for this data.
i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
>
```

- Data – viewing



Intro-to-R.Rmd x R data sets x R data sets x

Source Visual Outline

```
35 {r view}  
36 magazines  
37 }
```

A tibble: 204 × 5

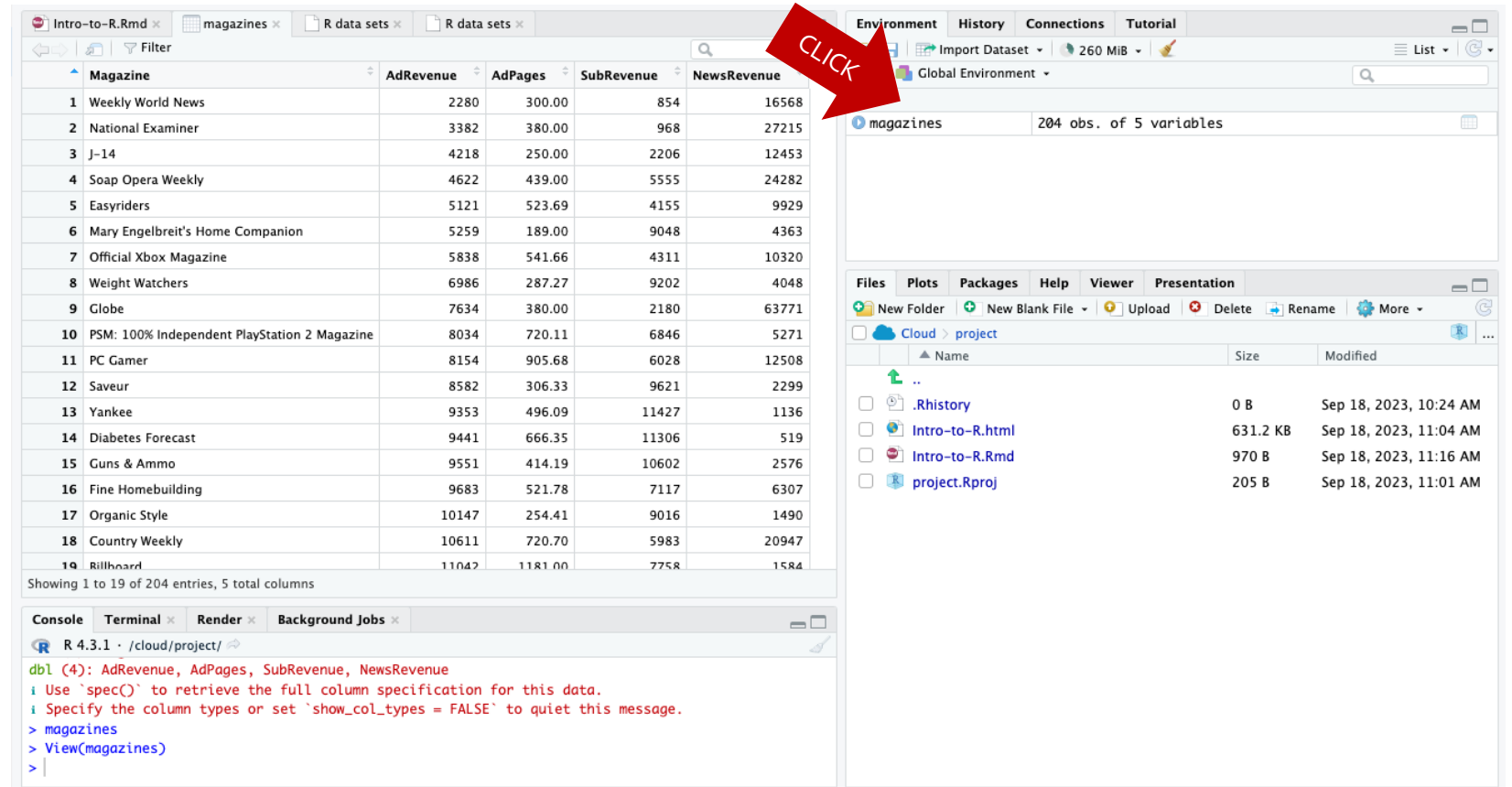
Magazine <chr>	AdRevenue <dbl>	AdPages <dbl>	SubRevenue <dbl>
Weekly World News	2280	300.00	854
National Examiner	3382	380.00	968
J-14	4218	250.00	2206
Soap Opera Weekly	4622	439.00	5555
Easyriders	5121	523.69	4155
Mary Engelbreit's Home Companion	5259	189.00	9048
Official Xbox Magazine	5838	541.66	4311
Weight Watchers	6986	287.27	9202
Globe	7634	380.00	2180
PSM: 100% Independent PlayStation 2 Magazine	8034	720.11	6846

1-10 of 204 rows | 1-4 of 5 columns Previous 1 2 3 4 5 6 ... 21 Next

36:10 Chunk 5: view R Markdown

R

- Data – viewing



CLICK

	Magazine	AdRevenue	AdPages	SubRevenue	NewsRevenue
1	Weekly World News	2280	300.00	854	16568
2	National Examiner	3382	380.00	968	27215
3	J-14	4218	250.00	2206	12453
4	Soap Opera Weekly	4622	439.00	5555	24282
5	Easyriders	5121	523.69	4155	9929
6	Mary Engelbreit's Home Companion	5259	189.00	9048	4363
7	Official Xbox Magazine	5838	541.66	4311	10320
8	Weight Watchers	6986	287.27	9202	4048
9	Globe	7634	380.00	2180	63771
10	PSM: 100% Independent PlayStation 2 Magazine	8034	720.11	6846	5271
11	PC Gamer	8154	905.68	6028	12508
12	Saveur	8582	306.33	9621	2299
13	Yankee	9353	496.09	11427	1136
14	Diabetes Forecast	9441	666.35	11306	519
15	Guns & Ammo	9551	414.19	10602	2576
16	Fine Homebuilding	9683	521.78	7117	6307
17	Organic Style	10147	254.41	9016	1490
18	Country Weekly	10611	720.70	5983	20947
19	Rolling Stone	11042	1181.00	7758	1584

Showing 1 to 19 of 204 entries, 5 total columns

```
R 4.3.1 · /cloud/project/  
dbl (4): AdRevenue, AdPages, SubRevenue, NewsRevenue  
i Use 'spec()' to retrieve the full column specification for this data.  
i Specify the column types or set 'show_col_types = FALSE' to quiet this message.  
> magazines  
> View(magazines)  
>
```


- Data – viewing

```
38  
39 ~ ``{r view-some}  
40 glimpse(magazines)|  
41 ~ ``
```

Rows: 204
Columns: 5
\$ Magazine <chr> "Weekly World News", "National Examiner", "J-14", "Soap Opera Wee...
\$ AdRevenue <dbl> 2280, 3382, 4218, 4622, 5121, 5259, 5838, 6986, 7634, 8034, 8154,...
\$ AdPages <dbl> 300.00, 380.00, 250.00, 439.00, 523.69, 189.00, 541.66, 287.27, 3...
\$ SubRevenue <dbl> 854, 968, 2206, 5555, 4155, 9048, 4311, 9202, 2180, 6846, 6028, 9...
\$ NewsRevenue <dbl> 16568, 27215, 12453, 24282, 9929, 4363, 10320, 4048, 63771, 5271,...

- Data – viewing

```
38  
39 ~ ``{r view-some}  
40 glimpse(magazines)|  
41 ~ ``
```

Rows: 204
Columns: 5

\$ Magazine	<chr>	"Weekly World News", "National Examiner", "J-14", "Soap Opera Wee...
\$ AdRevenue	<dbl>	2280, 3382, 4218, 4622, 5121, 5259, 5838, 6986, 7634, 8034, 8154,...
\$ AdPages	<dbl>	300.00, 380.00, 250.00, 439.00, 523.69, 189.00, 541.66, 287.27, 3...
\$ SubRevenue	<dbl>	854, 968, 2206, 5555, 4155, 9048, 4311, 9202, 2180, 6846, 6028, 9...
\$ NewsRevenue	<dbl>	16568, 27215, 12453, 24282, 9929, 4363, 10320, 4048, 63771, 5271,...

- Data – viewing

```
38  
39 ~ ``{r view-some}  
40 glimpse(magazines)|  
41 ~ ``
```

Rows: 204
Columns: 5

\$ Magazine	<chr>	Weekly World News", "National Examiner", "J-14", "Soap Opera Wee...
\$ AdRevenue	<dbl>	2280, 3382, 4218, 4622, 5121, 5259, 5838, 6986, 7634, 8034, 8154,...
\$ AdPages	<dbl>	300.00, 380.00, 250.00, 439.00, 523.69, 189.00, 541.66, 287.27, 3...
\$ SubRevenue	<dbl>	854, 968, 2206, 5555, 4155, 9048, 4311, 9202, 2180, 6846, 6028, 9...
\$ NewsRevenue	<dbl>	15568, 27215, 12453, 24282, 9929, 4363, 10320, 4048, 63771, 5271,...

- Data – viewing

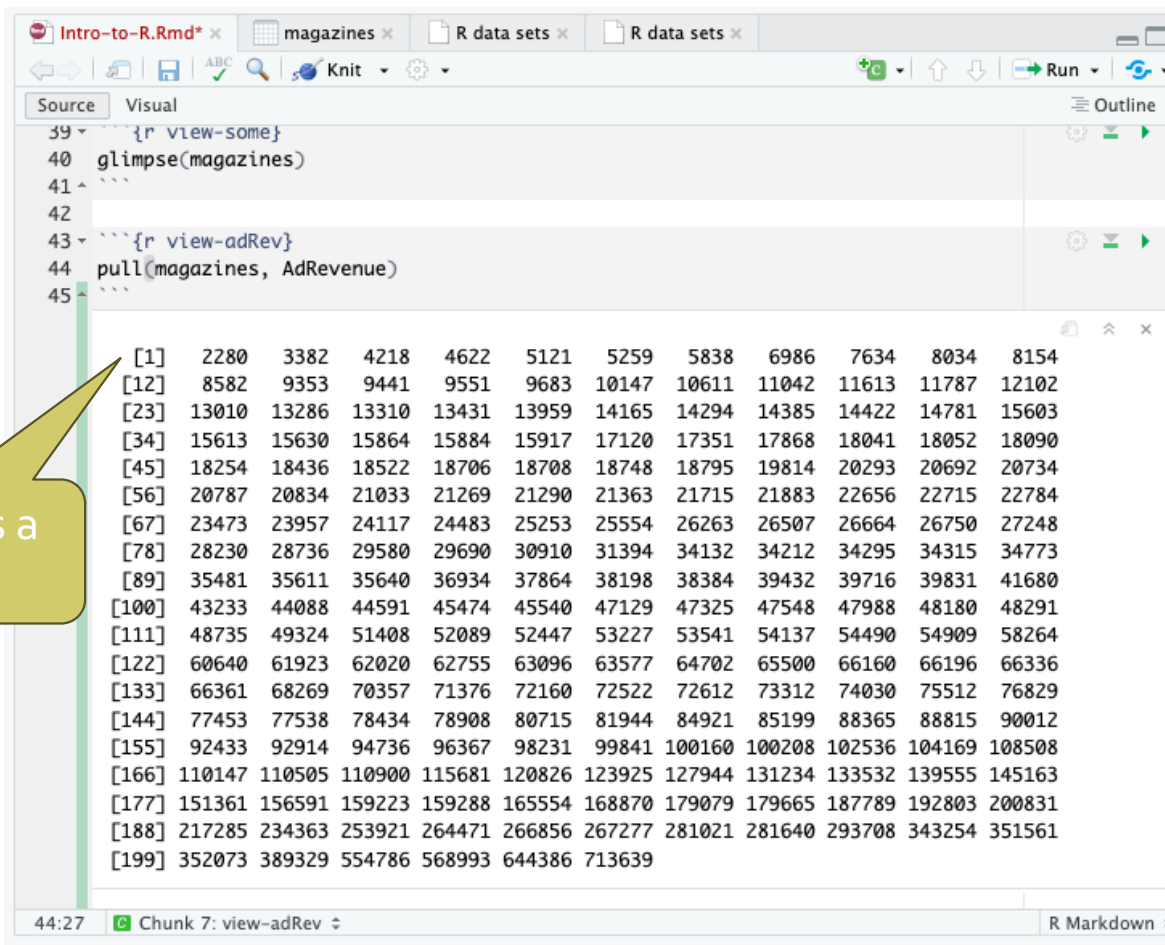
```
38  
39 ~ ``{r view-some}  
40 glimpse(magazines)|  
41 ~ ``
```

Rows: 204
Columns: 5
\$ Magazine <chr> "Weekly World News", "National Examiner", "J-14", "Soap Opera Wee...
\$ AdRevenue <dbl> 2280, 3382, 4218, 4622, 5121, 5259, 5838, 6986, 7634, 8034, 8154,...
\$ AdPages <dbl> 300.00, 380.00, 250.00, 439.00, 523.69, 189.00, 541.66, 287.27, 3...
\$ SubRevenue <dbl> 854, 968, 2206, 5555, 4155, 9048, 4311, 9202, 2180, 6846, 6028, 9...
\$ NewsRevenue <dbl> 16568, 27215, 12453, 24282, 9929, 4363, 10320, 4048, 63771, 5271,...

R

- Data – single column

column is a
vector



```
Intro-to-R.Rmd x magazines x R data sets x R data sets x
Source Visual
39 {r view-some}
40 glimpse(magazines)
41
42
43 {r view-adRev}
44 pull(magazines, AdRevenue)
45
```

[1]	2280	3382	4218	4622	5121	5259	5838	6986	7634	8034	8154
[12]	8582	9353	9441	9551	9683	10147	10611	11042	11613	11787	12102
[23]	13010	13286	13310	13431	13959	14165	14294	14385	14422	14781	15603
[34]	15613	15630	15864	15884	15917	17120	17351	17868	18041	18052	18090
[45]	18254	18436	18522	18706	18708	18748	18795	19814	20293	20692	20734
[56]	20787	20834	21033	21269	21290	21363	21715	21883	22656	22715	22784
[67]	23473	23957	24117	24483	25253	25554	26263	26507	26664	26750	27248
[78]	28230	28736	29580	29690	30910	31394	34132	34212	34295	34315	34773
[89]	35481	35611	35640	36934	37864	38198	38384	39432	39716	39831	41680
[100]	43233	44088	44591	45474	45540	47129	47325	47548	47988	48180	48291
[111]	48735	49324	51408	52089	52447	53227	53541	54137	54490	54909	58264
[122]	60640	61923	62020	62755	63096	63577	64702	65500	66160	66196	66336
[133]	66361	68269	70357	71376	72160	72522	72612	73312	74030	75512	76829
[144]	77453	77538	78434	78908	80715	81944	84921	85199	88365	88815	90012
[155]	92433	92914	94736	96367	98231	99841	100160	100208	102536	104169	108508
[166]	110147	110505	110900	115681	120826	123925	127944	131234	133532	139555	145163
[177]	151361	156591	159223	159288	165554	168870	179079	179665	187789	192803	200831
[188]	217285	234363	253921	264471	266856	267277	281021	281640	293708	343254	351561
[199]	352073	389329	554786	568993	644386	713639					

44:27 Chunk 7: view-adRev R Markdown

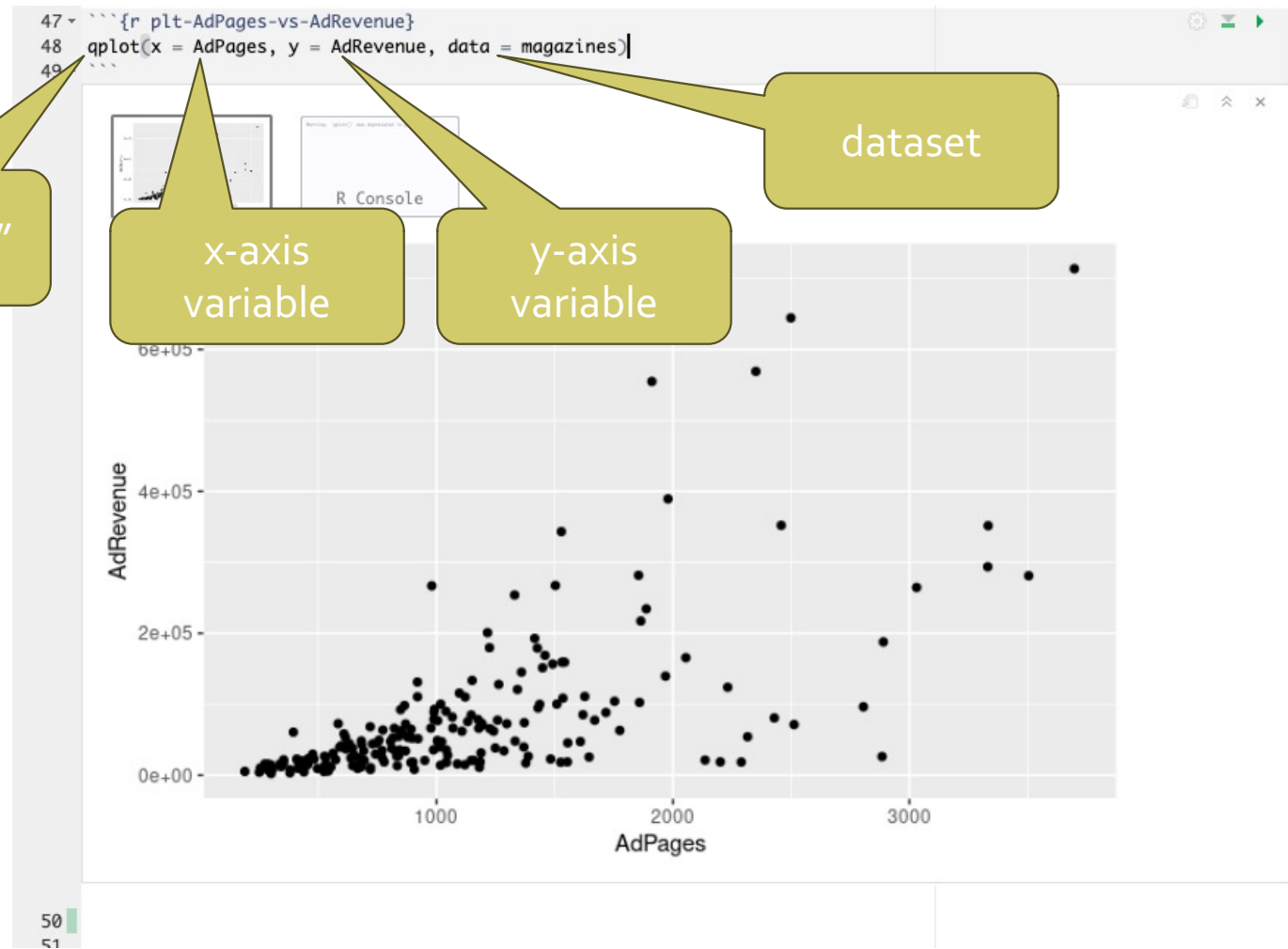
- Visualization

"quick plot"

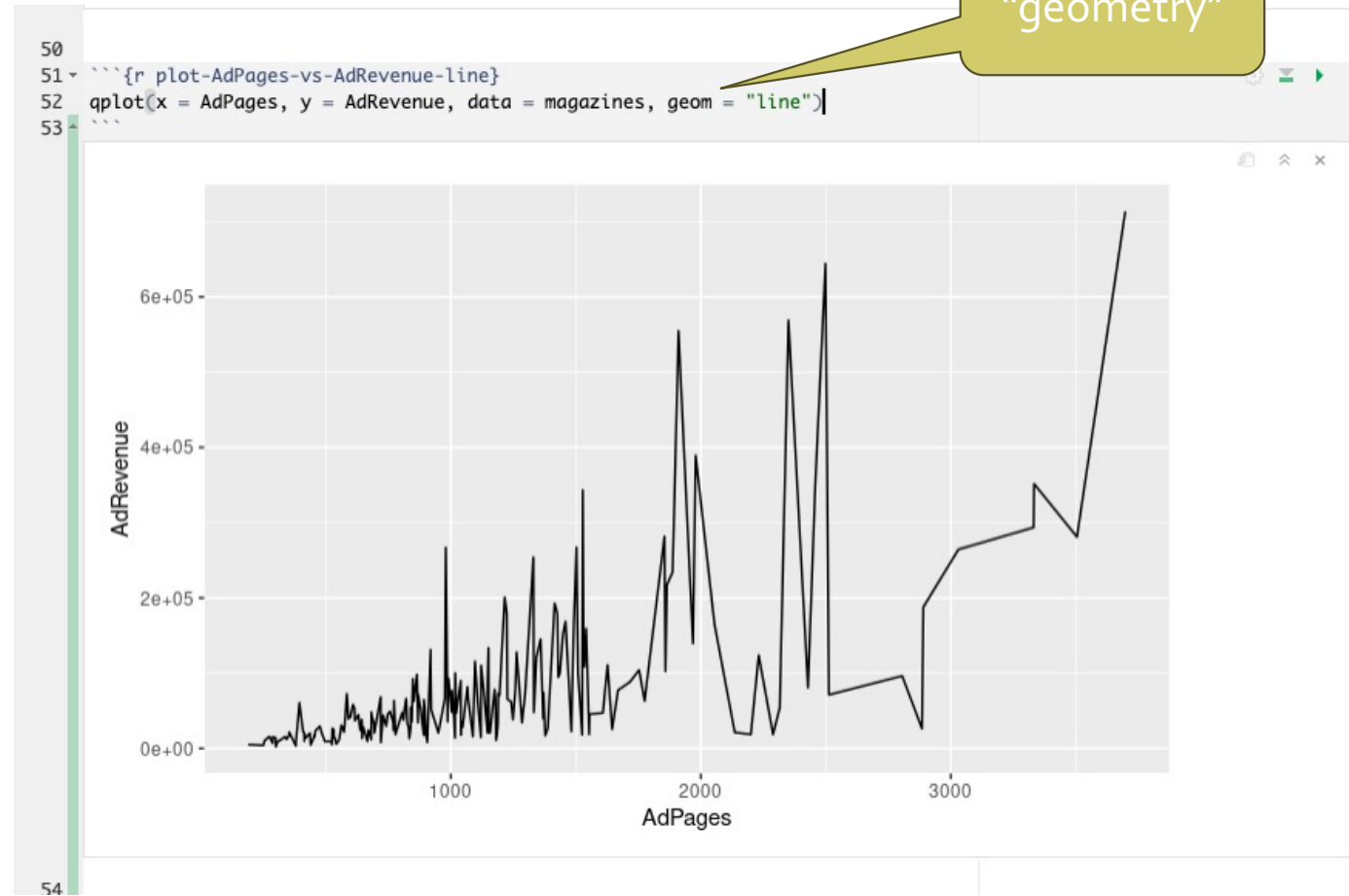
x-axis
variable

y-axis
variable

dataset



- Visualization



R

- Documentation

The screenshot displays the RStudio environment with several components:

- Source Editor:** Contains R code for creating a scatter plot and a line plot. The line plot code is highlighted with a red box.
- Environment:** Shows the 'magazines' data frame with 204 observations and 5 variables.
- Plots:** Displays two plots: a scatter plot of 'AdRevenue' vs 'AdPages' and a line plot of 'AdRevenue' vs 'AdPages'.
- R Documentation:** A window showing the documentation for the `qplot` function, which is highlighted with a red box. The documentation includes a description, usage, and arguments.

Source Editor Code:

```
50  
51 {r plot-AdPages-vs-AdRevenue-line}  
52 qplot(x = AdPages, y = AdRevenue, data = magazines, geom = "line")  
53  
54  
55 {r plot-help, tidy = FALSE}  
56 ?qplot  
57  
58  
59  
56:7 Chunk 10: plot-help
```

R Documentation for `qplot`:

Quick plot

Description

`qplot()` is now deprecated in order to encourage the users to learn `ggplot()` as it makes it easier to create complex graphics.

Usage

```
qplot(  
  x,  
  y,  
  ...,  
  data,  
  facets = NULL,  
  margins = FALSE,  
  geom = "auto",  
  xlim = c(NA, NA),  
  ylim = c(NA, NA),  
  log = "",  
  main = NULL,
```


- Adding Variables

```
58  
59 ~ ``{r calc-total-revenue-save}  
60 magazines <- magazines %>%  
61   mutate(TotalRevenue = AdRevenue + SubRevenue + NewsRevenue)  
62 ~``  
63
```

"pipe"

- %>% strings functions together
- Takes output of pervious expression and sends it as first parameter to next function in the chain
- Ex. $x \%>\% f(y)$ is equivalent to $f(x, y)$

R

- Adding Variables

```
58  
59 {r calc-total-revenue-save}  
60 magazines <- magazines %>%  
61   mutate(TotalRevenue = AdRevenue + SubRevenue + NewsRevenue)  
62  
63
```

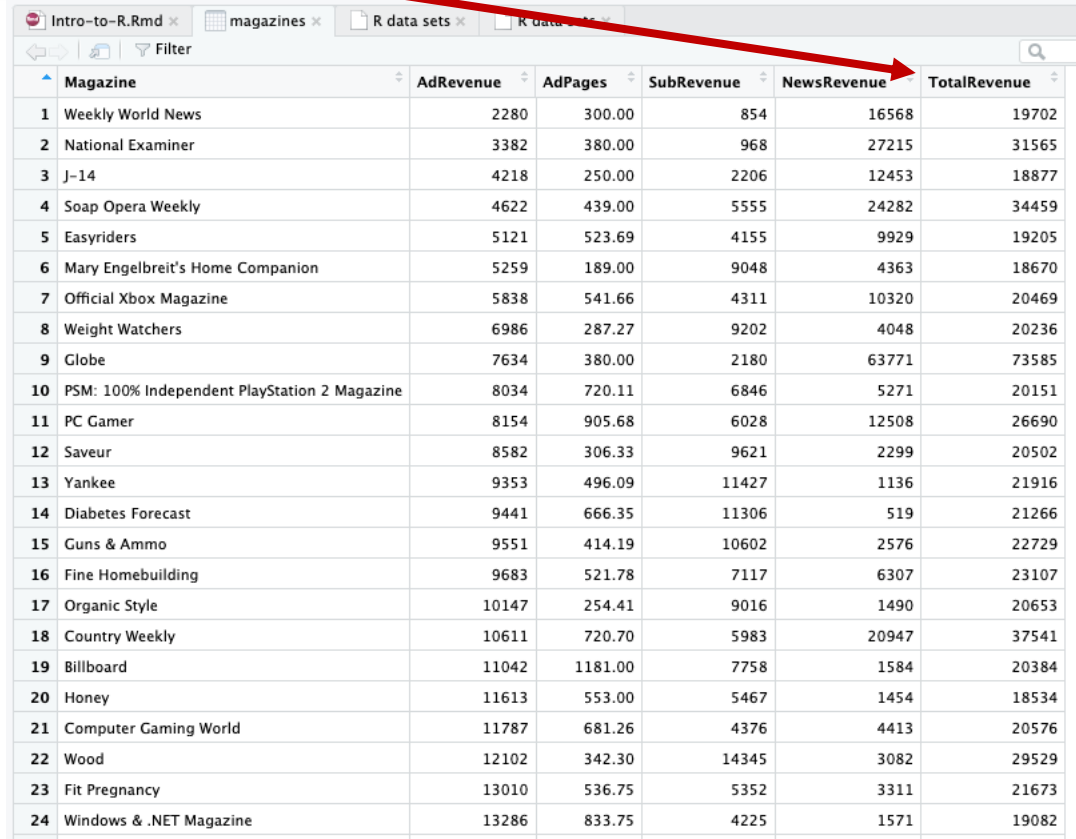
change
magazines by ...

adding new var

set magazines to
new dataset

- Adding Variables

```
58  
59 {r calc-total-revenue-save}  
60 magazines <- magazines %>%  
61 mutate(TotalRevenue = AdRevenue + SubRevenue + NewsRevenue)  
62  
63
```



	Magazine	AdRevenue	AdPages	SubRevenue	NewsRevenue	TotalRevenue
1	Weekly World News	2280	300.00	854	16568	19702
2	National Examiner	3382	380.00	968	27215	31565
3	J-14	4218	250.00	2206	12453	18877
4	Soap Opera Weekly	4622	439.00	5555	24282	34459
5	Easyriders	5121	523.69	4155	9929	19205
6	Mary Engelbreit's Home Companion	5259	189.00	9048	4363	18670
7	Official Xbox Magazine	5838	541.66	4311	10320	20469
8	Weight Watchers	6986	287.27	9202	4048	20236
9	Globe	7634	380.00	2180	63771	73585
10	PSM: 100% Independent PlayStation 2 Magazine	8034	720.11	6846	5271	20151
11	PC Gamer	8154	905.68	6028	12508	26690
12	Saveur	8582	306.33	9621	2299	20502
13	Yankee	9353	496.09	11427	1136	21916
14	Diabetes Forecast	9441	666.35	11306	519	21266
15	Guns & Ammo	9551	414.19	10602	2576	22729
16	Fine Homebuilding	9683	521.78	7117	6307	23107
17	Organic Style	10147	254.41	9016	1490	20653
18	Country Weekly	10611	720.70	5983	20947	37541
19	Billboard	11042	1181.00	7758	1584	20384
20	Honey	11613	553.00	5467	1454	18534
21	Computer Gaming World	11787	681.26	4376	4413	20576
22	Wood	12102	342.30	14345	3082	29529
23	Fit Pregnancy	13010	536.75	5352	3311	21673
24	Windows & .NET Magazine	13286	833.75	4225	1571	19082

R

- Adding Variables

```
67  
68 {r calc-adrev-per-page}  
69 magazines <- magazines %>%  
70   mutate(PerPageRevenue = AdRevenue / AdPages)  
71  
72  
73
```

What does this do?

R

- Adding Variables

```
67  
68 {r calc-adrev-per-page}  
69 magazines <- magazines %>%  
70   mutate(PerPageRevenue = AdRevenue / AdPages)  
71  
72  
73
```

What does this do?

- Adding Variables

What did this do?

```
73 {r newsstand-vs-subscription}  
74 magazines <- magazines %>%  
75   mutate(OneOff = NewsRevenue > SubRevenue)  
76  
77 pull(magazines, OneOff)  
78
```

```
[1] TRUE TRUE TRUE TRUE TRUE FALSE TRUE FALSE TRUE FALSE TRUE FALSE FALSE FALSE FALSE FALSE  
[18] TRUE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE  
[35] FALSE FALSE FALSE FALSE TRUE FALSE TRUE FALSE FALSE FALSE FALSE FALSE TRUE TRUE FALSE  
[52] FALSE FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE TRUE TRUE  
[69] FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE TRUE  
[86] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE TRUE FALSE  
[103] FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE TRUE FALSE FALSE FALSE FALSE TRUE  
[120] FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE  
[137] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE  
[154] FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE  
[171] FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE TRUE FALSE FALSE  
[188] FALSE FALSE FALSE TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

79

R

- Try on your own!
- Lab is on the course website under “Labs” tab (<https://amoscao1.github.io/Intro-DS-F23/>)
- Go to the “More Practice” section of the lab for today and complete as much as you can
- Ask questions when they come up!