



# Getting Started with R – Part I

Who am I?

---

# About Perry

---

- Motivates, mobilizes, and connects cross-functional teams with technical solutions and support
- Provides customer-focused Computer Professional services with Data Science / Systems Engineering experience in commercial and non-profit industries.
- Delivers system, network, and security support in a wide variety of business and home environments.
- Partners with clients for training and end-developer support efforts, especially in the areas of configuration management, operating system integration.

How am I going about learning  
about R?

---

# Overview

- Download R / Desktop
- Download Rstudio
  - Download tidyverse package
- Compile everything.



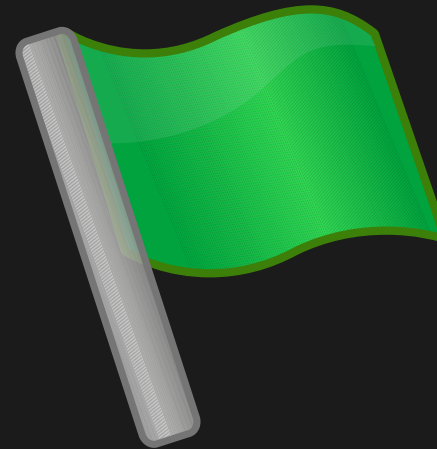
# Observations

---

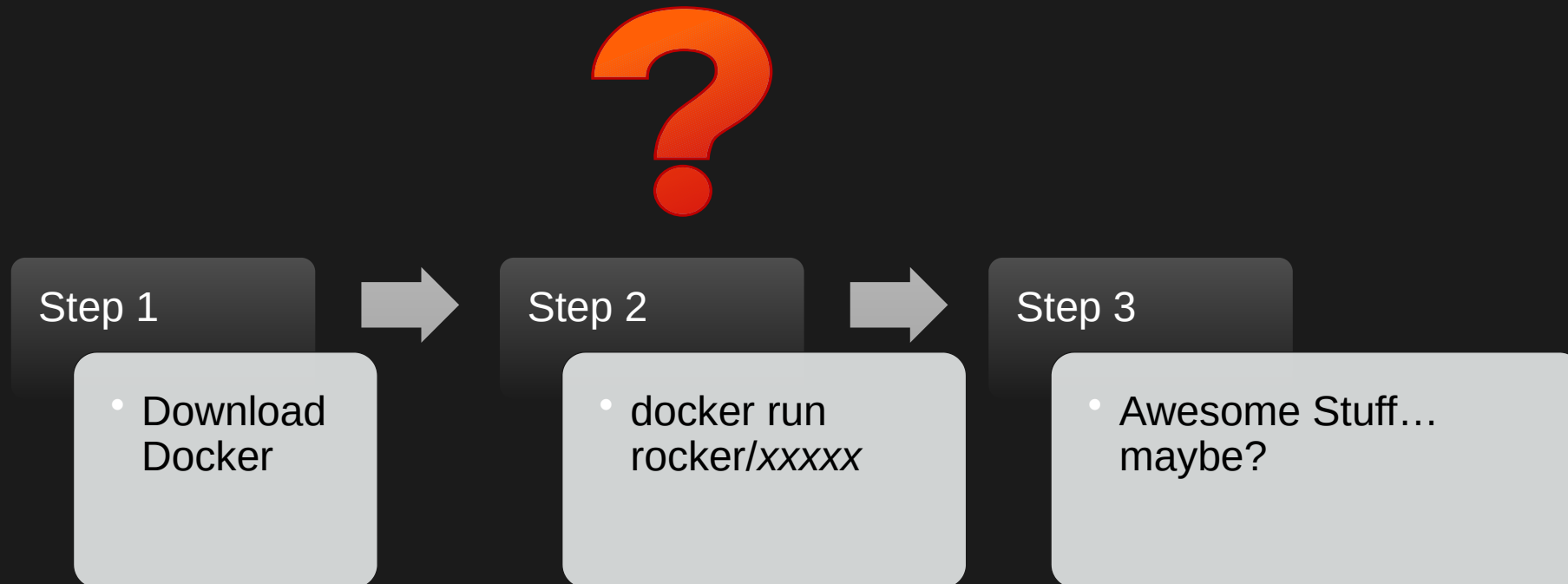
- Long compile times
- Aging laptop kept shutting off because Linux kept overclocking during compilation and system kept overheating
- Sooooo many dependencies and compilation issues/errors

# Overview (Revised)

- Download R / Desktop
  - Download Rstudio
    - Download tidyverse package
  - Compile everything.
- Download Docker
  - docker run *something something*
  - Done!



# Procedure







Docker



something something

Docker

A diagram illustrating the relationship between Docker and R packages. A large green rectangle at the bottom is labeled 'Docker'. Above it, a smaller blue rectangle contains a list of R packages: rocker/rstudio, rocker/tidyverse, rocker/verse, rocker/geospatial, and rocker/shiny, followed by 'etc. etc.'.

**rocker/rstudio**  
**rocker/tidyverse**  
**rocker/verse**  
**rocker/geospatial**  
**rocker/shiny**

**etc. etc.**

Docker

# Docker at 50,000 Feet

- Runs software packages called containers
- Developers can package up applications, including libraries and dependencies into a container
- We'll be using the `docker run` command to facilitate pulling and running of images from the Internet



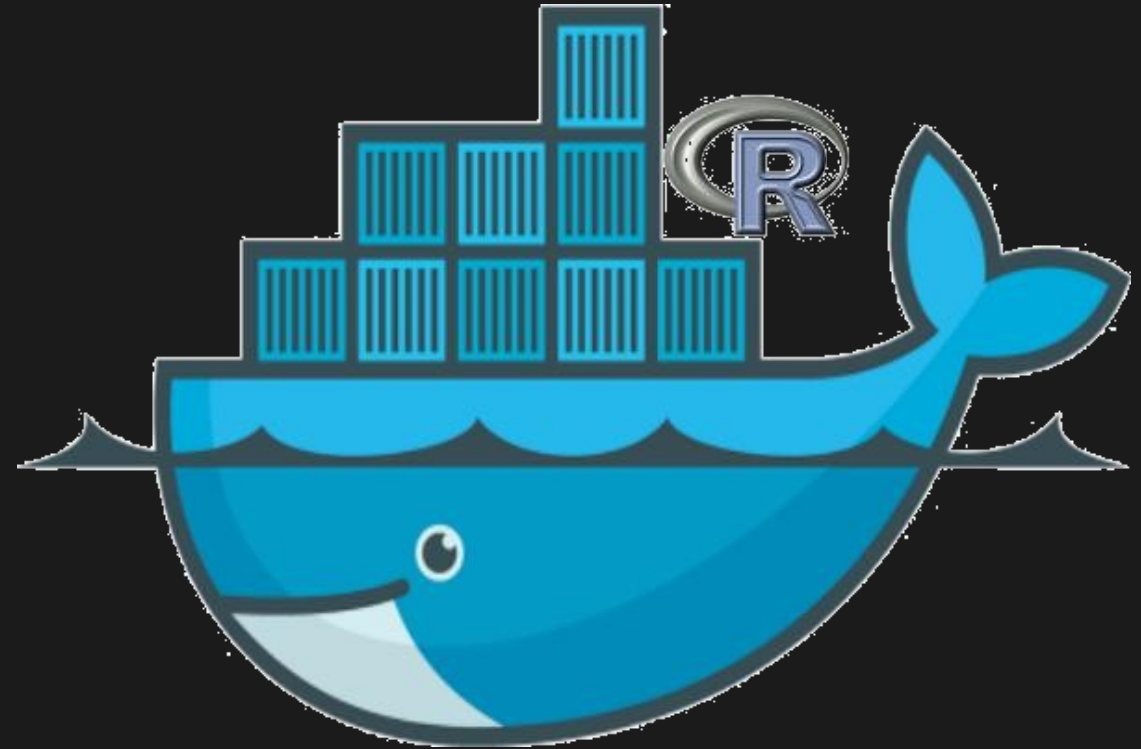
# R at 50,000 Feet

- Programming language for statistical computing
- Created by Ross Ihaka and Robert Gentleman (R's in their first names!)
- Interpreted language
- Used for modeling and analysis



# Rocker at 50,000 Feet

- R on Docker = Rocker
- <https://www.rocker-project.org/images/>
- R with libraries and dependencies in nifty containers
- TidyVerse is a collection of packages that facilitates our visualizations



# RStudio at 50,000 Feet

- An open-source, free integrated development environment (IDE) for R on Docker = Rocker
- Software Engineer JJ Allaire, Founder
  - Created ColdFusion language
  - LoseIt! - Weight Tracking App



# Some Images in the Rocker Universe...

- <https://www.rocker-project.org/images/>

Image	Comments
r-ver	• Minimal
rstudio	• Adds rstudio. Mid-sized container. Sans tidyverse
tidyverse	• Adds tidyverse. Great mid-sized container. Includes ggplot2 (data visualization stuff)
verse	• Adds TeX and publishing
geospatial	• Adds geospatial libraries. Large image



# How do I download Docker?

---

# Setting Up Docker (Part 1 of 2)

- <https://docs.docker.com/>
- Left Navigation Menu > Get Docker > Docker CE > Linux (or MacOS or ...?)
- For Fedora:
  - <https://docs.docker.com/install/linux/docker-ce/fedora/>
  - `sudo dnf -y install dnf-plugins-core`
  - `sudo dnf config-manager --add-repo`  
`https://download.docker.com/linux/fedora/docker-ce.repo`
  - `sudo dnf install docker-ce docker-ce-cli containerd.io`
  - `dnf list docker-ce --showduplicates | sort -r`

# Setting Up Docker (Part 2 of 2)

- For Fedora:
  - Enable docker daemon, start it, and add your username to sudo'ers
    - `sudo systemctl enable docker`
    - `sudo systemctl start docker`
    - `sudo groupadd docker`
    - `sudo usermod -aG docker $USER`
- Test basic functionality:
  - `docker run hello-world`

# Setting Up Rocker

- Pick an image:
  - Did you want r-ver, rstudio, tidyverse, etc.?
- Invoke docker (replace rstudio with the desired image):
  - `docker run -e PASSWORD=SHHH --rm -p 8787:8787 rocker/tidyverse`
    - Replace SHHH with your own password.  
Note: don't use rstudio for PASSWORD
    - --rm provides clean-up of the container when you're done.
    - -p is the port number
    - Last argument (rocker/tidyverse, etc.) is the desired image

# Post-configuration Docker (Part 1 of 2)

---

- Open up a browser
- Browse to:
  - localhost:8787

# Post-configuration Docker (Part 2 of 2)

- ~~In the console window, type:~~
  - ~~– `docker pull rocker/tidyverse`~~
- In rstudio:
  - `library(tidyverse)`
  - `library(devtools)`
  - `install.packages(ggfortify) [or install_github('sinhrks/ggfortify')]`
  - `install.packages("zoo")`
  - `library(ggfortify)`
  - `autoplot(AirPassengers)`

How Do I Save My Stuff?

---

A diagram illustrating the relationship between Docker and rocker images. A large green rectangle at the bottom is labeled 'Docker'. Above it, a smaller blue rectangle contains a list of rocker images: rocker/rstudio, rocker/tidyverse, rocker/verse, rocker/geospatial, and rocker/shiny, followed by 'etc. etc.'.

**rocker/rstudio**  
**rocker/tidyverse**  
**rocker/verse**  
**rocker/geospatial**  
**rocker/shiny**

**etc. etc.**

Docker



# Uh Oh...There's a Problem!

---

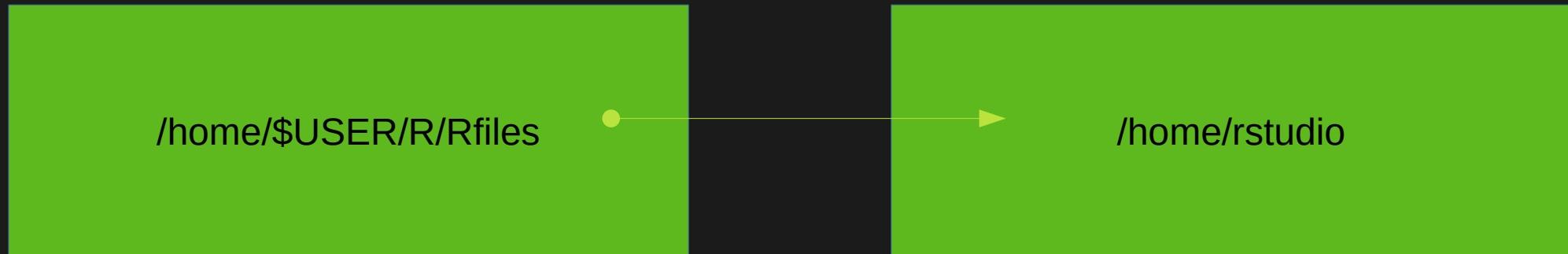
- Without the -v switch, can't save to the host operating system.
- So... use the -v switch

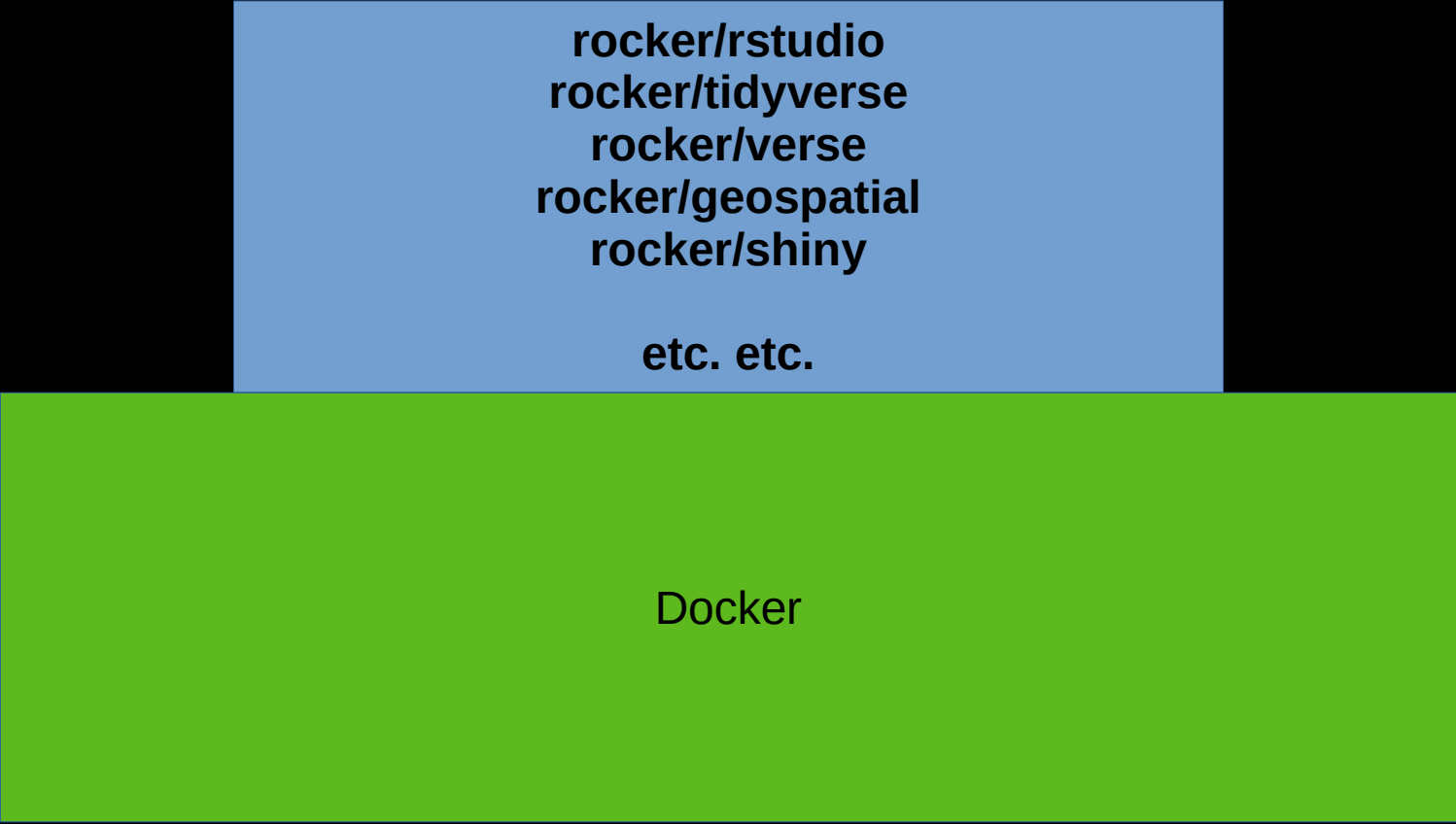
# Setup Considerations...

- Be sure to mkdir a directory, similar to:
  - /home/\$USER/R/Rfiles
  - **Back that directory tree up somewhere!**
- So... the revised command is now...
  - docker run -e PASSWORD=SHHH -v /home/\$USER/R/Rfiles:/home/rstudio --rm -p 8787:8787 rocker/tidyverse

# What the heck is /home/rstudio??

- When the developers created this general-use image, they made the home directory of your Rocker image /home/rstudio.
- So... let docker know where the source files come from using the docker run -v switch
- The /home/\$USER/R/Rfiles host directory gets bind mounted to /home/rstudio in the container

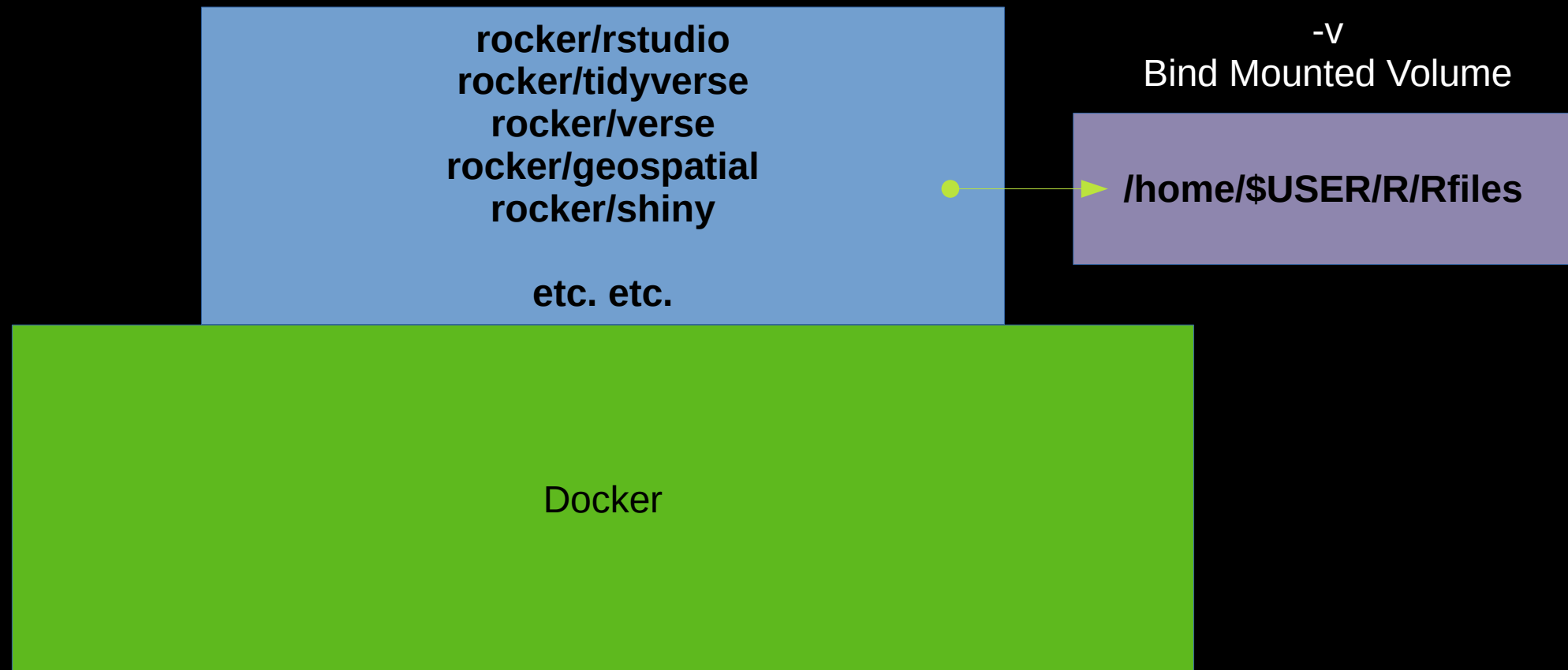




**rocker/rstudio**  
**rocker/tidyverse**  
**rocker/verse**  
**rocker/geospatial**  
**rocker/shiny**

**etc. etc.**

Docker

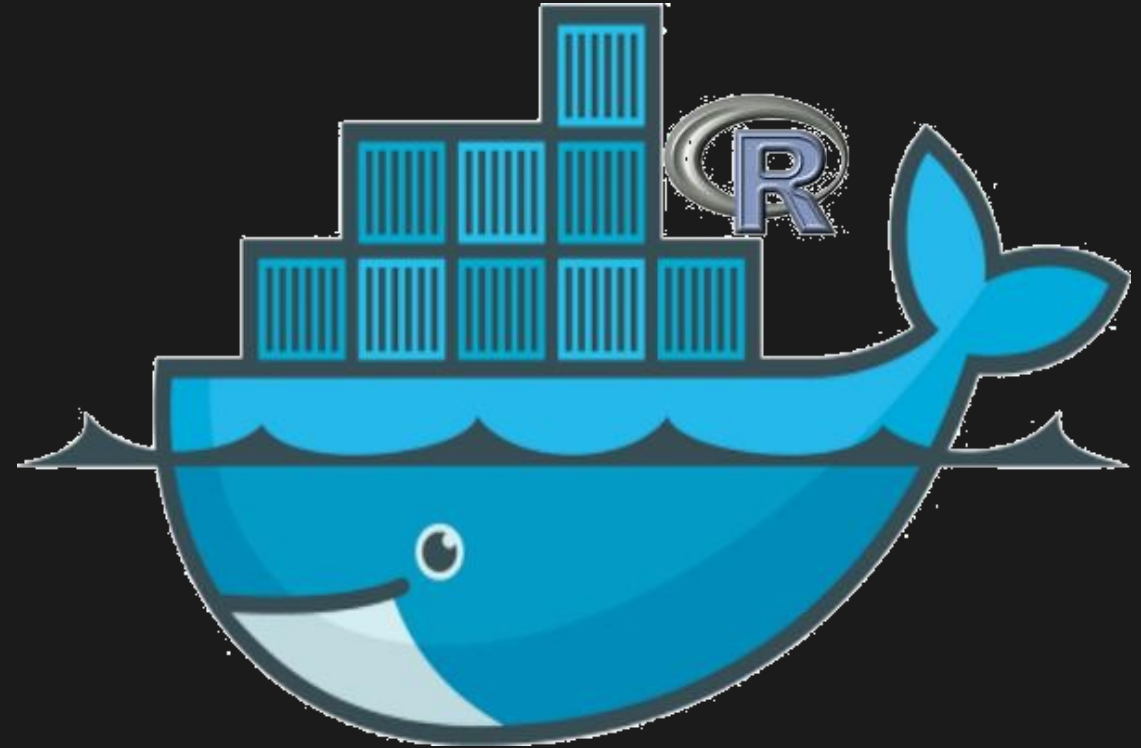


What Next?

---

# Jupyter at 50,000 Feet

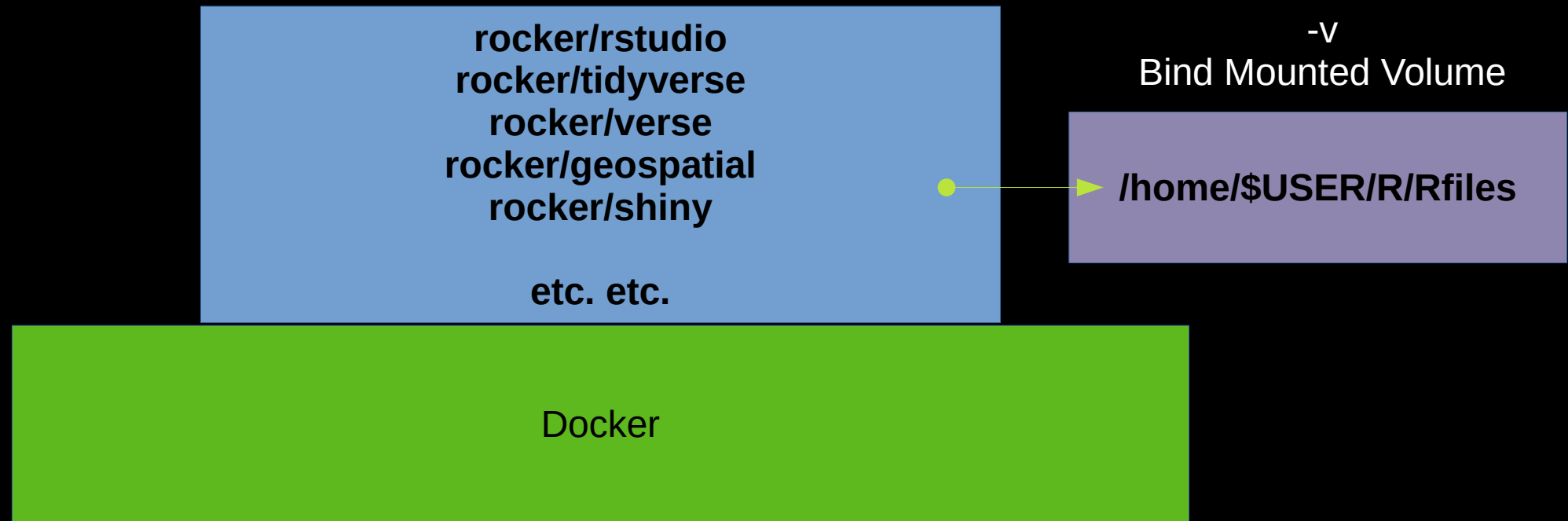
- Web application
- Allows you to create notebooks that contain live code

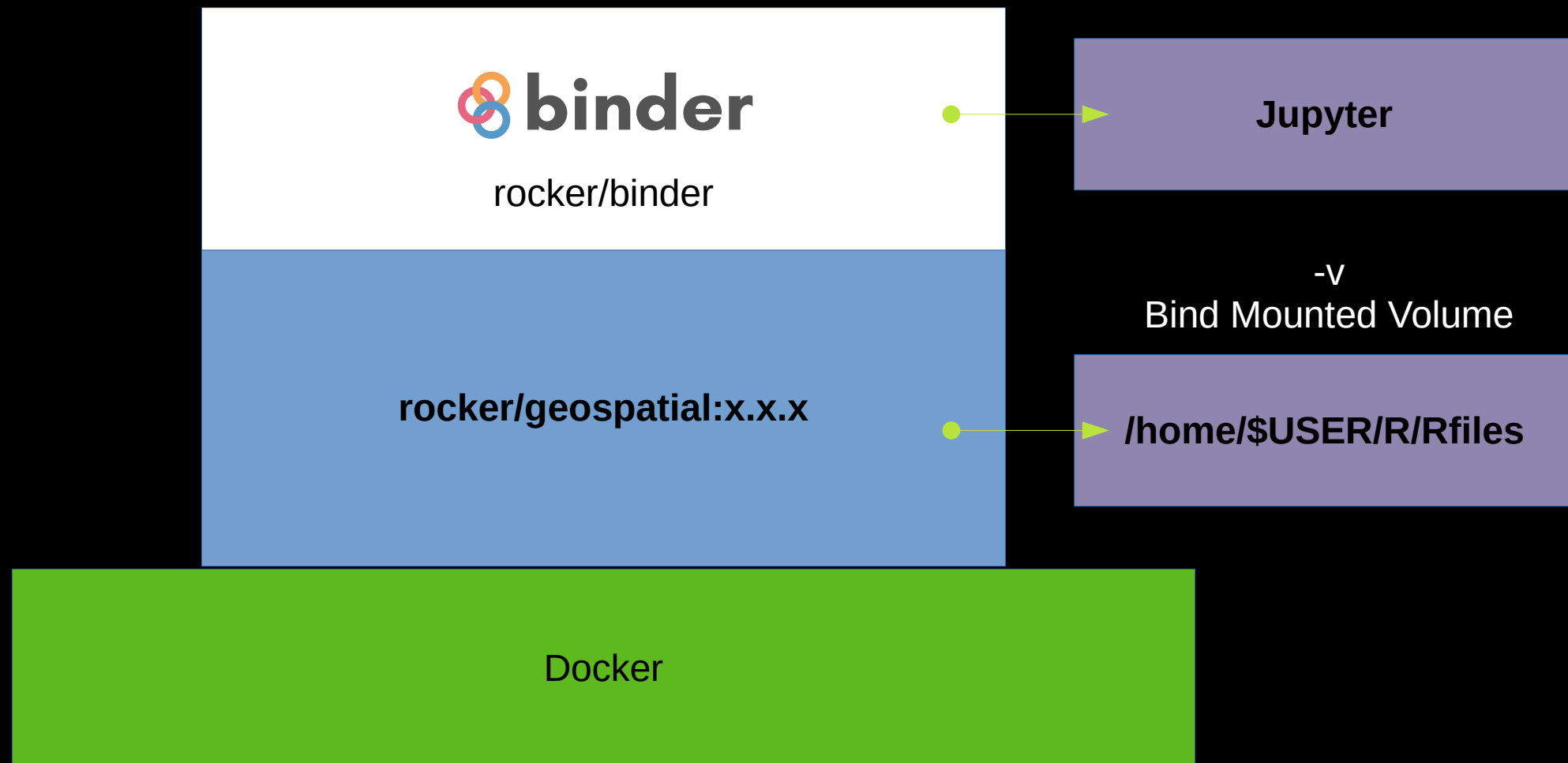


# New Problem!

- How do we use R from Jupyter Notebook?
- Solution:
  - <https://github.com/rocker-org/binder>
- Binder facilitates docker'd (is that a word?) R within Jupyter
- Fix the problem by:
  - `docker pull rocker/binder`
  - `docker run -ti --rm -e PASSWORD=SHHHHH`  
`-v /home/$USER/R/Rfiles:/home/rstudio -p 8888:8888 rocker/binder`
    - New switch: `-ti` pseudo-tty and interactive shell







# Caveat Emptor!

- When you run the command:
  - `docker run -ti --rm -e PASSWORD=SHHHHH -v /home/$USER/R/Rfiles:/home/rstudio -p 8888:8888 rocker/binder`
  - Be sure to copy and paste that token generated in the URL you paste into the browser.

Live Demo Maybe?

---

Is There Another Way?

---

# RStudio Cloud at 50,000 Feet

- Allows you to explore data and the RStudio environment on a hosted environment
- Nothing to install.  
You just need a browser!
- As an alpha release,  
currently free to use
- Your data lives somewhere on the Amazon US-East cloud



# To the Cloud...

---

- <https://rstudio.cloud/>
- Upper Right-hand Corner → Sign Up
  - Use your GitHub account or Google Account

# What It Looks Like

The screenshot displays the RStudio Cloud interface. On the left, a sidebar contains navigation links: Spaces, Your Workspace, New Space, Learn (Guide, What's New, Primers, Cheat Sheets), Feedback and Questions, and Info (Terms and Conditions, System Status). The main workspace area is titled 'Your Workspace / intro\_R' and 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. The code editor shows a script with R code for basic arithmetic and a ggplot2 plot. The console at the bottom displays the output of the code, including the creation of a ggplot object. On the right, a panel shows the Environment (Global Environment) with a list of data objects: ad\_treatment (600 obs. of 5 variables), df\_cars (32 obs. of 11 variables), raw\_df (600 obs. of 5 variables), and small\_df (32 obs. of 4 variables). Below this, a Files panel shows a directory listing of the project, including files like code\_through.rmd, edge\_cases.rmd, intro-R.html, intro-R.rmd, project.Rproj, r\_basics.r, real\_life\_data, and simple\_data.

**Spaces**

- Your Workspace
- New Space

**Learn**

- Guide
- What's New
- Primers
- Cheat Sheets

**Feedback and Questions**

**Info**

- Terms and Conditions
- System Status

**Your Workspace / intro\_R**

**File Edit Code View Plots Session Build Debug Profile Tools Help**

Go to file/function

Addins

R 3.6.0

**Environment History Connections**

Global Environment

**Data**

Object	Description
ad_treatment	600 obs. of 5 variables
df_cars	32 obs. of 11 variables
raw_df	600 obs. of 5 variables
small_df	32 obs. of 4 variables

**Values**

**Files Plots Packages Help Viewer**

New Folder Upload Delete Rename More

Cloud > project

Name	Size	Modified
code_through.rmd	7.5 KB	Jun 12, 2019, 9:38 AM
edge_cases.rmd	505 B	Jun 12, 2019, 9:43 AM
intro-R.html	1.4 MB	Jun 12, 2019, 10:23 AM
intro-R.rmd	72.6 KB	Jun 11, 2019, 6:40 PM
project.Rproj	205 B	Jun 24, 2019, 11:24 PM
r_basics.r	833 B	Jun 12, 2019, 10:22 AM
real_life_data		
simple_data		

```
1 # Math
2
3 3 + 3
4
5 # Base R Functions + objects
6
7 x <- 3 + 5
8 x
9
10 y <- c(1, 2, 3, 4, 5)
11
```

55:1 (Top Level) R Script

**Console Terminal Jobs**

```
/cloud/project/
method from
[.quosures rlang
c.quosures rlang
print.quosures rlang
> ggplot(mtcars, aes(x = hp, y = mpg)) +
+   geom_point() +
+   geom_smooth(method = "lm", se = FALSE) +
+   labs(x = "Horsepower",
+        y = "Miles per Gallon (MPG)",
+        title = "Fuel efficiency decreases as Displacement increases")
> library(tidyverse)
```



# What It Looks Like

The screenshot displays the RStudio Cloud interface for a workspace named 'intro\_R'. The interface is divided into several panels:

- Left Panel:** Contains navigation links for 'Spaces' (Your Workspace, New Space), 'Learn' (Guide, What's New, Primers, Cheat Sheets, Feedback and Questions), and 'Info' (Terms and Conditions, System Status).
- Top Panel:** Shows the workspace name 'intro\_R' and the R version 'R 3.6.0'.
- Code Editor:** Displays a script file 'r\_basics.r' with the following code:

```
1 # Math
2
3 3 + 3
4
5 # Base R Functions + objects
6
7 x <- 3 + 5
8 x
9
10 y <- c(1, 2, 3, 4, 5)
11
```
- Console:** Shows the output of the code execution, including the creation of a ggplot object and the loading of the tidyverse library.

```
/cloud/project/
method from
[.quosures rlang
c.quosures rlang
print.quosures rlang
> ggplot(mtcars, aes(x = hp, y = mpg)) +
+   geom_point() +
+   geom_smooth(method = "lm", se = FALSE) +
+   labs(x = "Horsepower",
+        y = "Miles per Gallon (MPG)",
+        title = "Fuel efficiency decreases as Displacement increases")
> library(tidyverse)
```
- Environment Panel:** Lists the data objects in the global environment:

Data	Observations	Variables
ad_treatment	600 obs.	5 variables
df_cars	32 obs.	11 variables
raw_df	600 obs.	5 variables
small_df	32 obs.	4 variables
- Files Panel:** Shows the file structure of the project, including files like 'code\_through.rmd', 'edge\_cases.rmd', 'intro-R.html', 'intro-R.rmd', 'project.Rproj', 'r\_basics.r', 'real\_life\_data', and 'simple\_data'.

# What It Looks Like

The screenshot displays the RStudio Cloud interface. The top bar shows 'Your Workspace / intro\_R' and the R version 'R 3.6.0'. The left sidebar contains navigation options: Spaces (Your Workspace, New Space), Learn (Guide, What's New, Primers, Cheat Sheets, Feedback and Questions), and Info (Terms and Conditions, System Status).

The main editor area shows a script with the following code:

```
1 # Math
2
3 3 + 3
4
5 # Base R Functions + objects
6
7 x <- 3 + 5
8 x
9
10 y <- c(1, 2, 3, 4, 5)
11
```

The bottom console area, highlighted with a red box, shows the output of the code execution:

```
/cloud/project/
> method
[.quosures rlang
c.quosures rlang
print.quosures rlang
> ggplot(mtcars, aes(x = hp, y = mpg)) +
+   geom_point() +
+   geom_smooth(method = "lm", se = FALSE) +
+   labs(x = "Horsepower",
+        y = "Miles per Gallon (MPG)",
+        title = "Fuel efficiency decreases as Displacement increases")
> library(tidyverse)
```

The right sidebar contains the Environment, History, and Connections panels. The Environment panel shows the Global Environment with the following data:

Data	Size
ad_treatment	600 obs. of 5 variables
df_cars	32 obs. of 11 variables
raw_df	600 obs. of 5 variables
small_df	32 obs. of 4 variables

The Files panel shows the project structure:

Name	Size	Modified
code_through.rmd	7.5 KB	Jun 12, 2019, 9:38 AM
edge_cases.rmd	505 B	Jun 12, 2019, 9:43 AM
intro-R.html	1.4 MB	Jun 12, 2019, 10:23 AM
intro-R.rmd	72.6 KB	Jun 11, 2019, 6:40 PM
project.Rproj	205 B	Jun 24, 2019, 11:24 PM
r_basics.r	833 B	Jun 12, 2019, 10:22 AM
real_life_data		
simple_data		

# What It Looks Like

The screenshot displays the RStudio Cloud interface for a workspace named 'intro\_R'. The interface is divided into several panels:

- Left Panel (Spaces):** Contains options for 'Your Workspace', 'New Space', 'Learn' (Guide, What's New, Primers, Cheat Sheets), and 'Feedback and Questions'.
- Top Panel (Menu):** Includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The R version is 3.6.0.
- Code Editor:** Shows a script with R code for a linear model and a ggplot2 plot. The code includes comments and variable assignments.
- Environment Panel (Highlighted with a red box):** Displays the 'Global Environment' with a list of objects: 'ad\_treatment' (600 obs. of 5 variables), 'df\_cars' (32 obs. of 11 variables), 'raw\_df' (600 obs. of 5 variables), and 'small\_df' (32 obs. of 4 variables).
- Console:** Shows the output of the R code, including the execution of 'ggplot()' and 'library(tidyverse)'.
- Files Panel:** Lists files in the 'project' directory, including 'code\_through.rmd', 'edge\_cases.rmd', 'intro-R.html', 'intro-R.rmd', 'project.Rproj', 'r\_basics.r', 'real\_life\_data', and 'simple\_data'.

# What It Looks Like

The screenshot displays the RStudio Cloud interface for a workspace named 'intro\_R'. The interface is divided into several panels:

- Left Panel (Spaces):** Contains links for 'Your Workspace', 'New Space', 'Learn' (Guide, What's New, Primers, Cheat Sheets), and 'Feedback and Questions'.
- Top Panel (Menu):** Includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a 'Go to file/function' search bar.
- Code Editor:** Shows a script with R code for basic arithmetic and a ggplot2 visualization of the 'mtcars' dataset. The code includes comments and variable assignments.
- Environment Panel:** Displays the 'Global Environment' with a list of data objects: 'ad\_treatment' (600 obs. of 5 variables), 'df\_cars' (32 obs. of 11 variables), 'raw\_df' (600 obs. of 5 variables), and 'small\_df' (32 obs. of 4 variables).
- Files Panel (highlighted with a red box):** Shows a file browser for the 'project' directory. It lists files and folders with their sizes and modification dates.
- Console:** Displays the output of the R code, including the execution of 'ggplot()' and 'library(tidyverse)'.

**Files Panel Data:**

Name	Size	Modified
code_through.rmd	7.5 KB	Jun 12, 2019, 9:38 AM
edge_cases.rmd	505 B	Jun 12, 2019, 9:43 AM
intro-R.html	1.4 MB	Jun 12, 2019, 10:23 AM
intro-R.rmd	72.6 KB	Jun 11, 2019, 6:40 PM
project.Rproj	205 B	Jun 24, 2019, 11:24 PM
r_basics.r	833 B	Jun 12, 2019, 10:22 AM
real_life_data		
simple_data		

# Explore!

- Check Out Left Hand Side:
  - Guide Section
  - Primer Section
  - Cheat Sheets
    - Python with R: Reticulate
    - Deep Learning with Keras
    - More!
- Watch the Mock Video:  
<https://resources.rstudio.com/tidyverse/a-gentle-introduction-to-tidy-statistics-in-r>

# What Resources Are Available?

---

# Websites

Site	Description	Link
Rocker Project	Rocker Image Information	<a href="https://www.rocker-project.org/">https://www.rocker-project.org/</a>
Rocker Tidyverse on Github	Rocker/Tidyverse Image Information	<a href="https://hub.docker.com/r/rocker/tidyverse/">https://hub.docker.com/r/rocker/tidyverse/</a>
Rocker Binder	Rocker Binder	<a href="https://github.com/rocker-org/binder">https://github.com/rocker-org/binder</a>
RStudio Cloud	Hosted Rstudio Cloud Service	<a href="https://rstudio.cloud/">https://rstudio.cloud/</a>

Questions???

---



# Recap

---

- Reviewed installation considerations for the R Studio environment
- Discussed Docker and why it can be useful for quick setup
- Discovered Rocker Projects (tidyverse, binder) that include ready-to-go containers to meet immediate needs
- Learned about the Hosted RStudio Cloud environment

# Perry Rivera

- E-mail: [perry-rivera@alumni.calpoly.edu](mailto:perry-rivera@alumni.calpoly.edu)
- LinkedIn: [@pvrconsulting](#)
- Slides:
  - <https://tinyurl.com/yxe3q23y>
- Github:
  - [https://github.com/perryrivera/r\\_development\\_presentation](https://github.com/perryrivera/r_development_presentation)





# Thank You!