Getting Started with R – Part I

Perry Rivera | Rivera Consulting | perry-rivera@alumni.calpoly.edu

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.

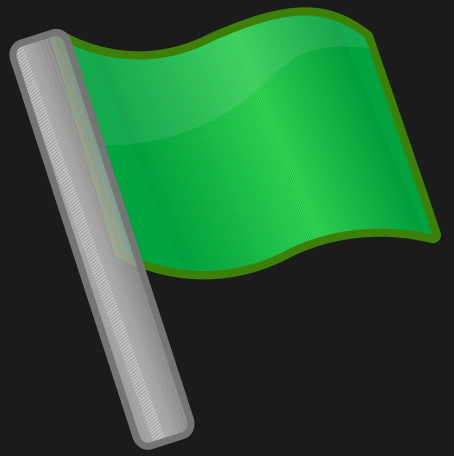
Observations

- 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

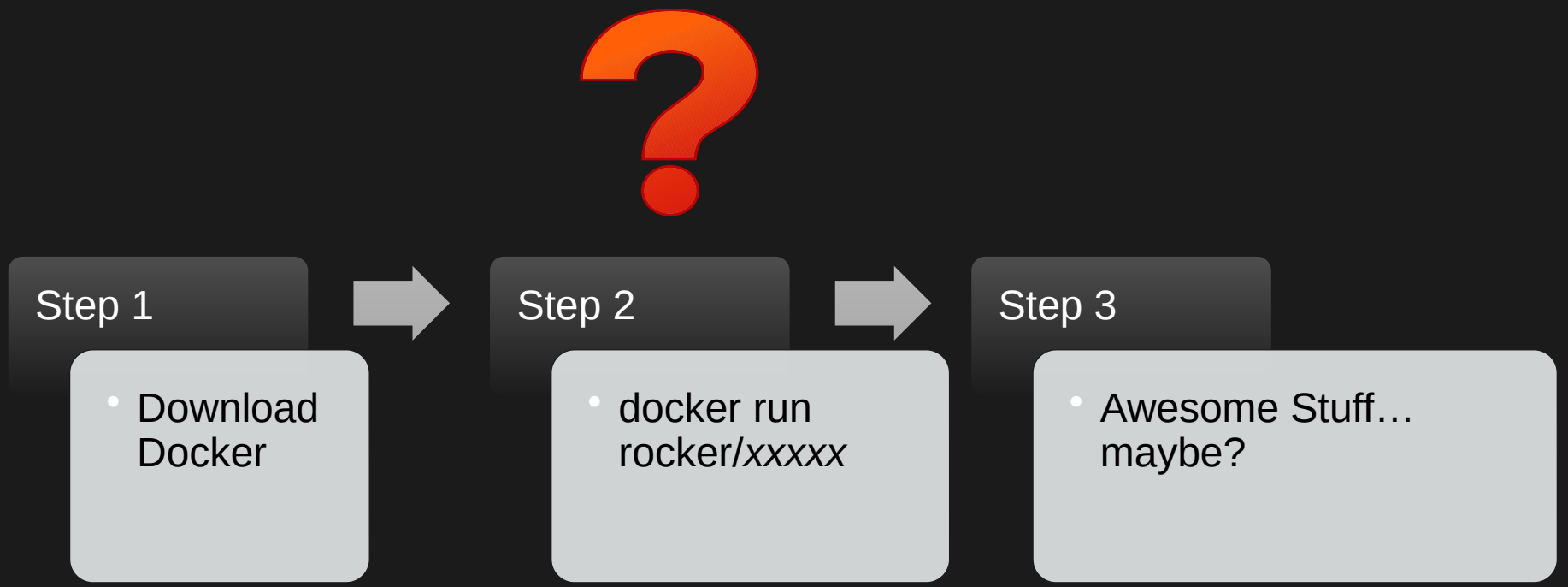
Overview (Revised)

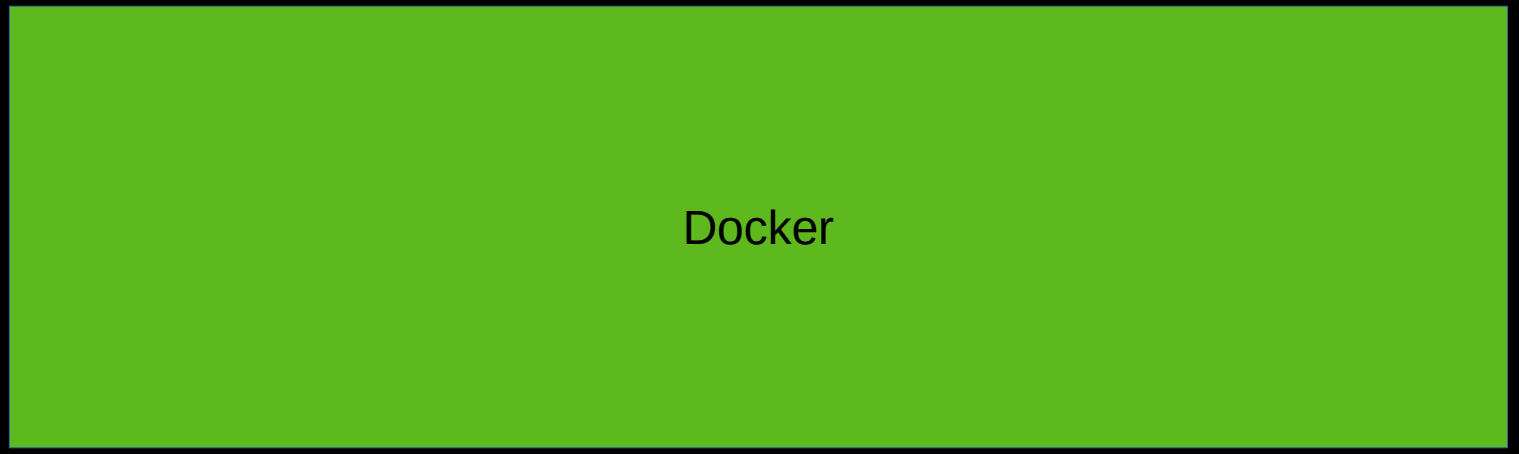
- Download R / Desktop
- Download Rstudio
  - Download tidyverse package
- Compile everything.

- Download Docker
- docker run *something something*
- Done!

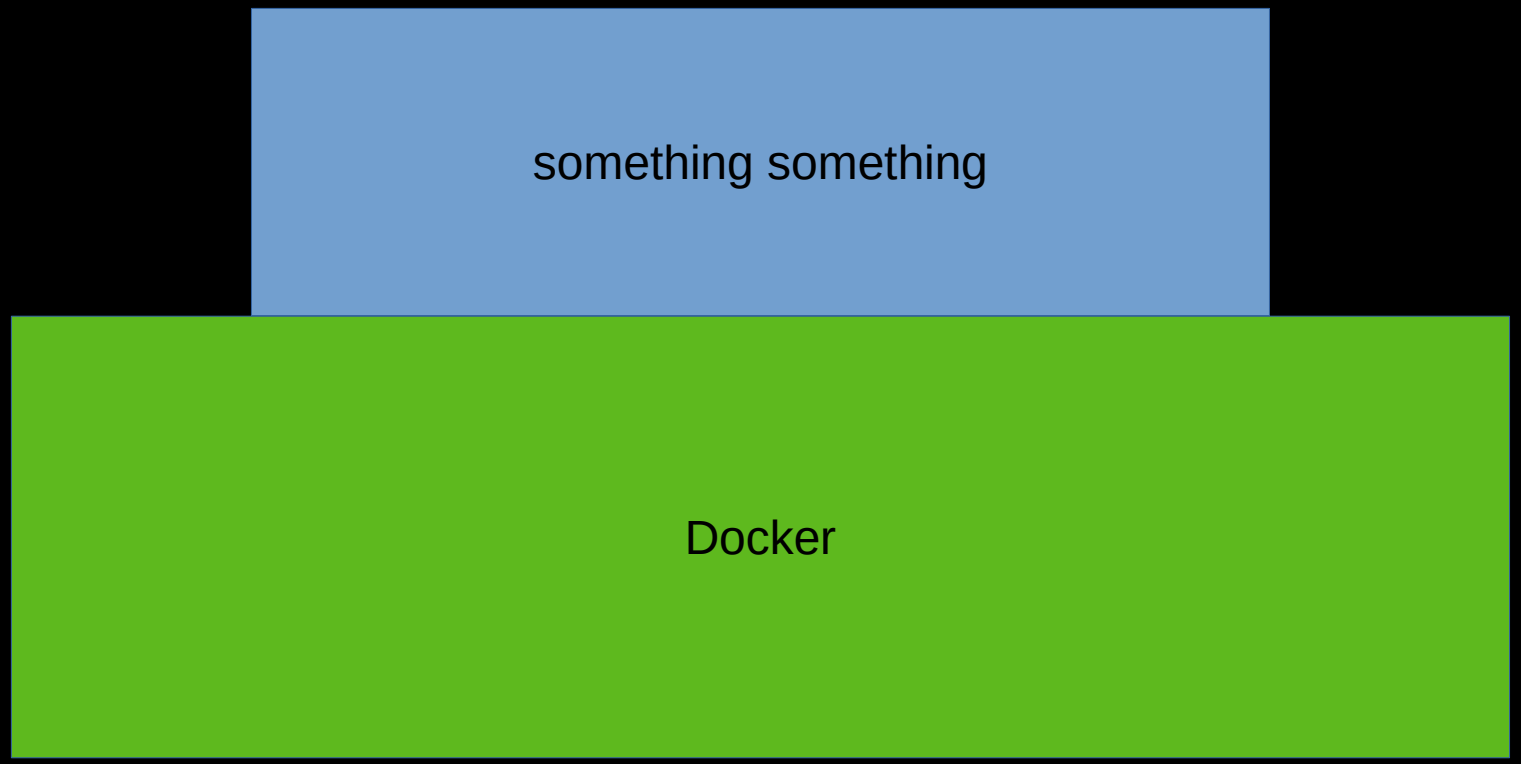


Procedure

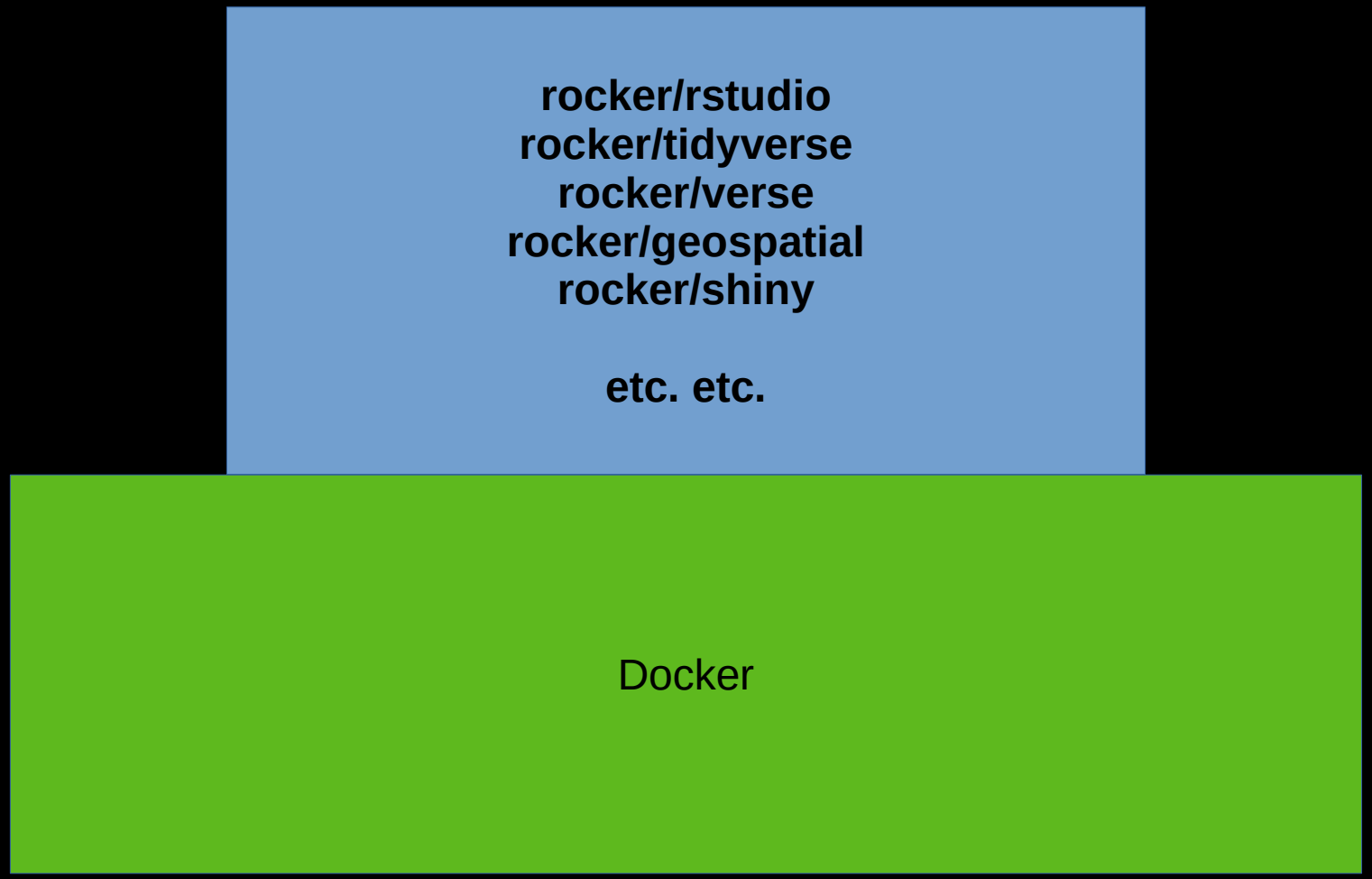

Step 1
Download Docker
Step 2
docker run rocker/xxxxx
Step 3
Awesome Stuff… maybe?



Overview



Overview


rocker/rstudio
rocker/tidyverse
rocker/verse
rocker/geospatial
rocker/shiny
etc. etc.
Docker

Docker at 50,000 Feet

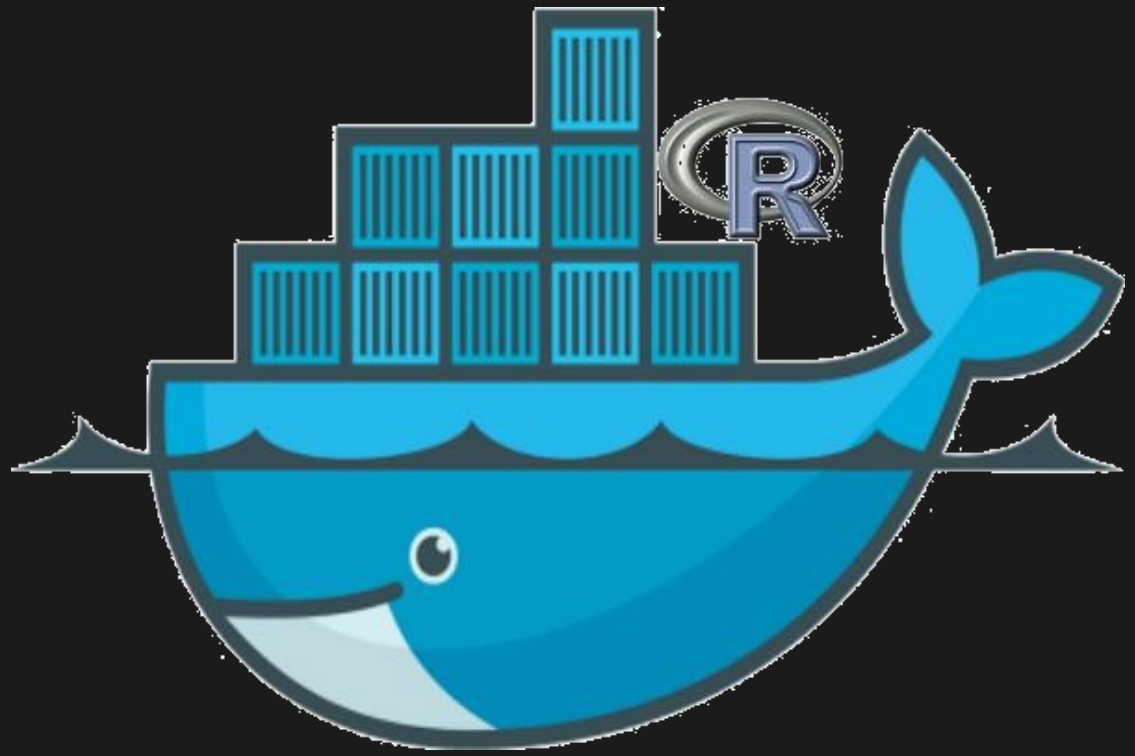
- Runs software packages called containers
- Developers can package up applications, including libraries and dependencies into a container
- We'll be using the docker run command to facilitate pulling and running of images from the Internet

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

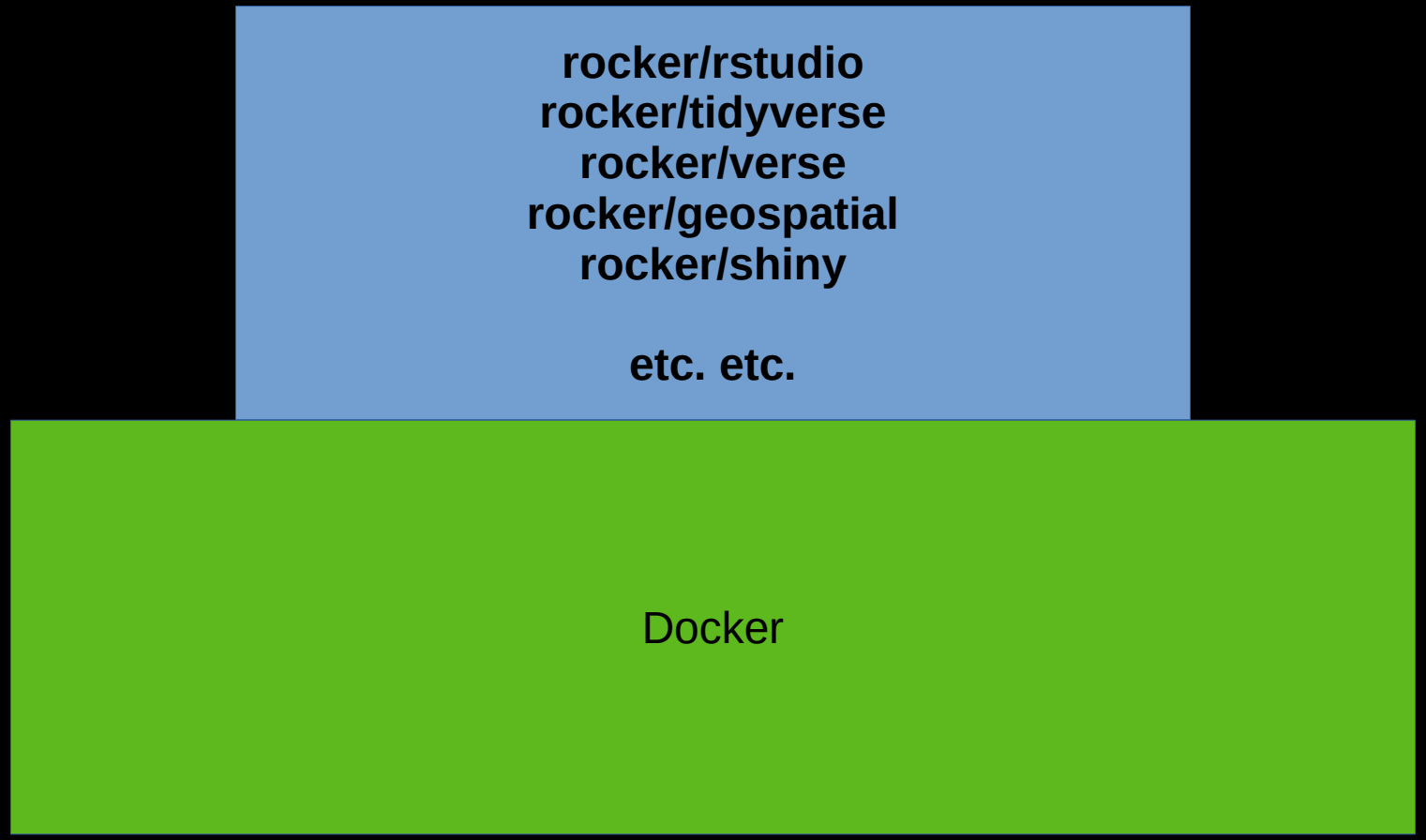
Post-configuration Docker (Part 1 of 2)

- 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?


rocker/rstudio
rocker/tidyverse
rocker/verse
rocker/geospatial
rocker/shiny
etc. etc.
Docker

Uh Oh...There’s a Problem!

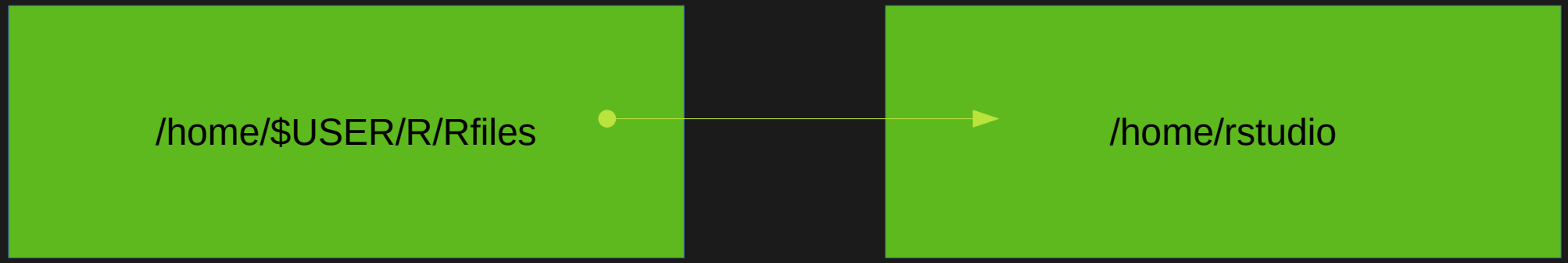
- Without the -v switch, can’t save to the host operating system.
- So... use the -v switch

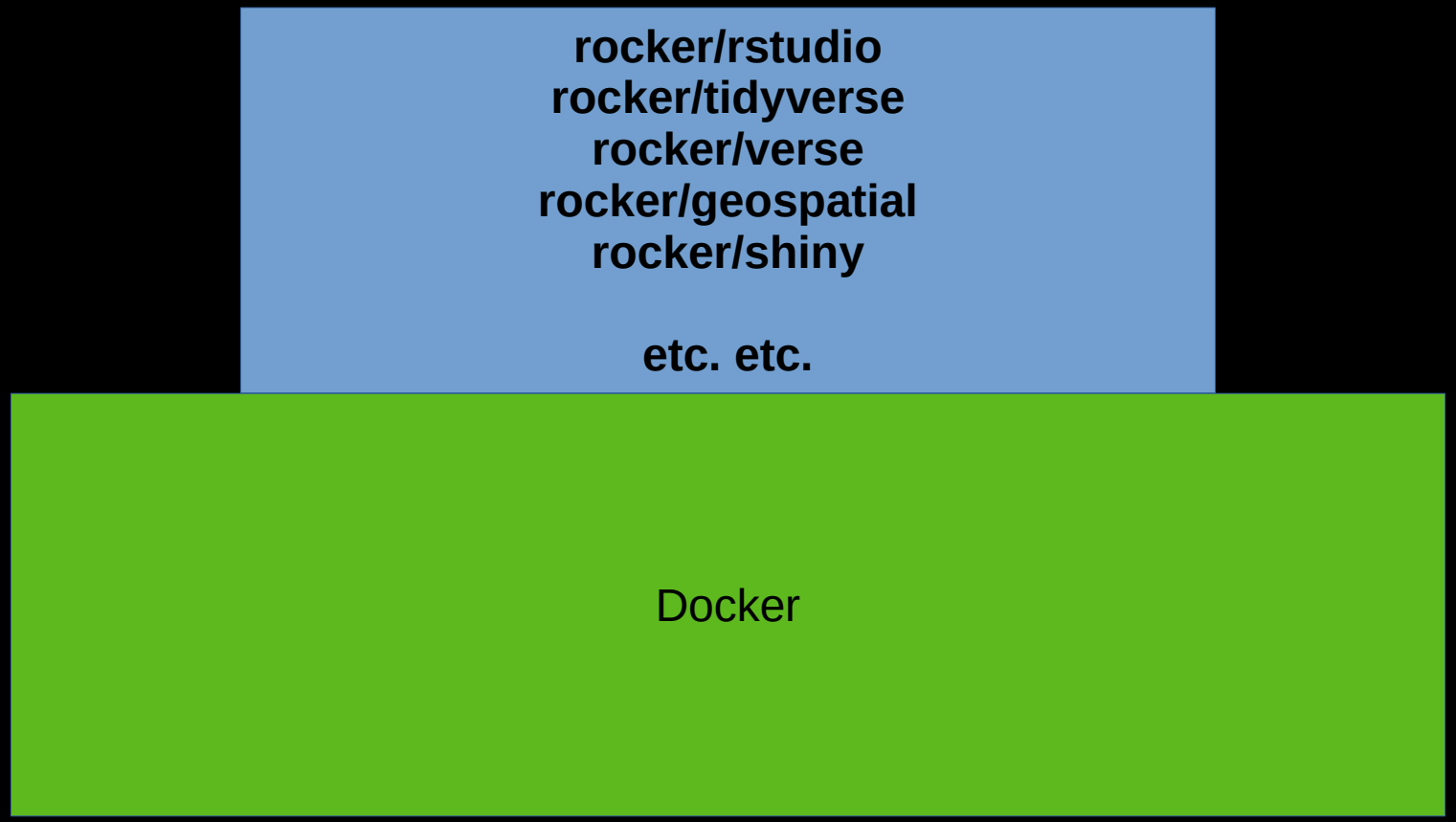
Setup Considerations...

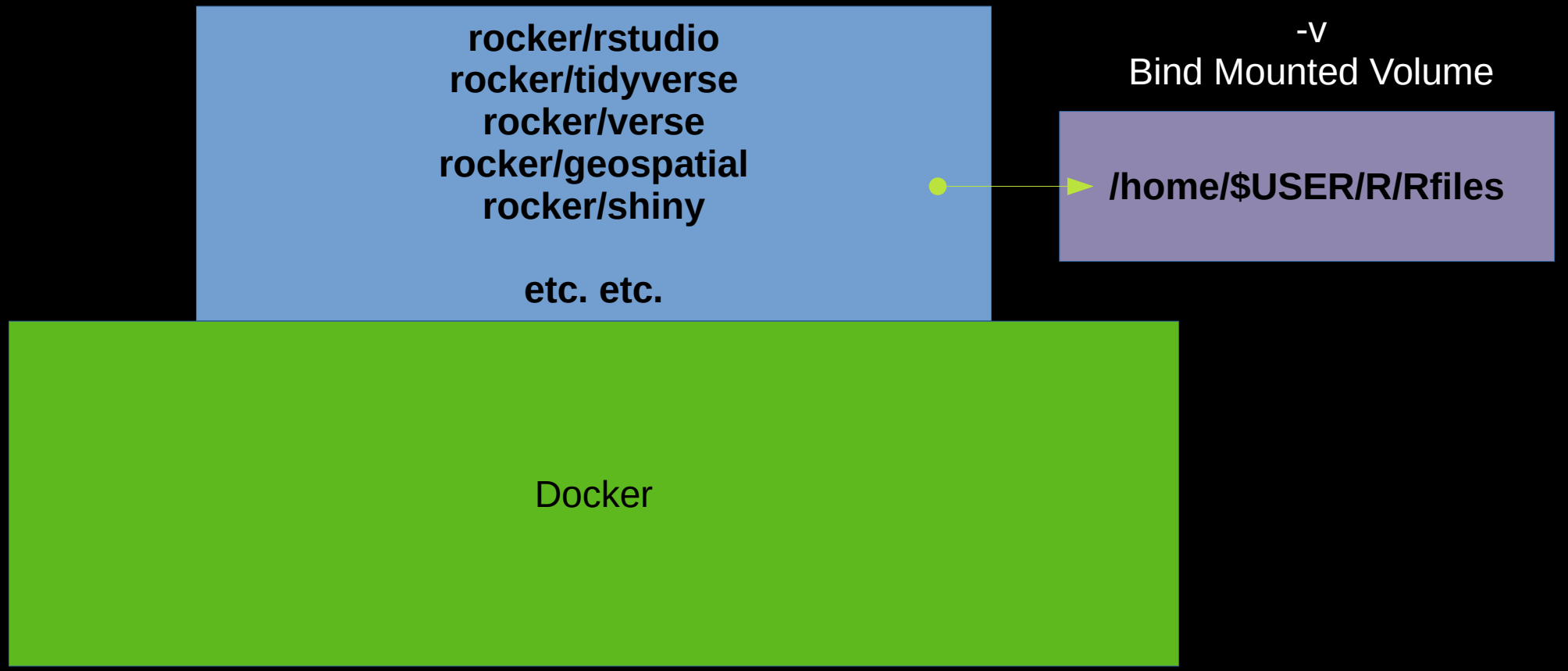
- Be sure to mkdir a directory, similar to:
  - /home/$USER/R/Rfiles
  - **Back that directory tree up somewhere!**
- So... the revised command is now...
  - docker run -e PASSWORD=SHHH -v /home/$USER/R/Rfiles:/home/rstudio --rm -p 8787:8787 rocker/tidyverse

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



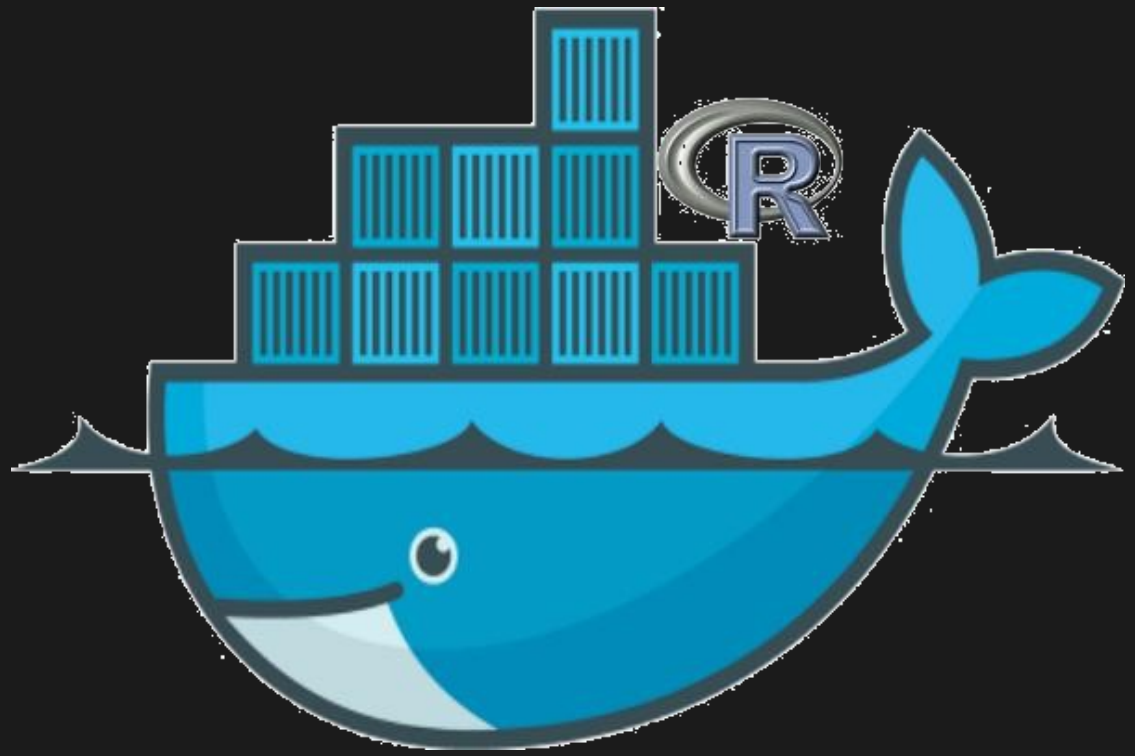

rocker/rstudio
rocker/tidyverse
rocker/verse
rocker/geospatial
rocker/shiny
etc. etc.
Docker


rocker/rstudio
rocker/tidyverse
rocker/verse
rocker/geospatial
rocker/shiny
etc. etc.
-v
Bind Mounted Volume
/home/$USER/R/Rfiles
Docker

What Next?

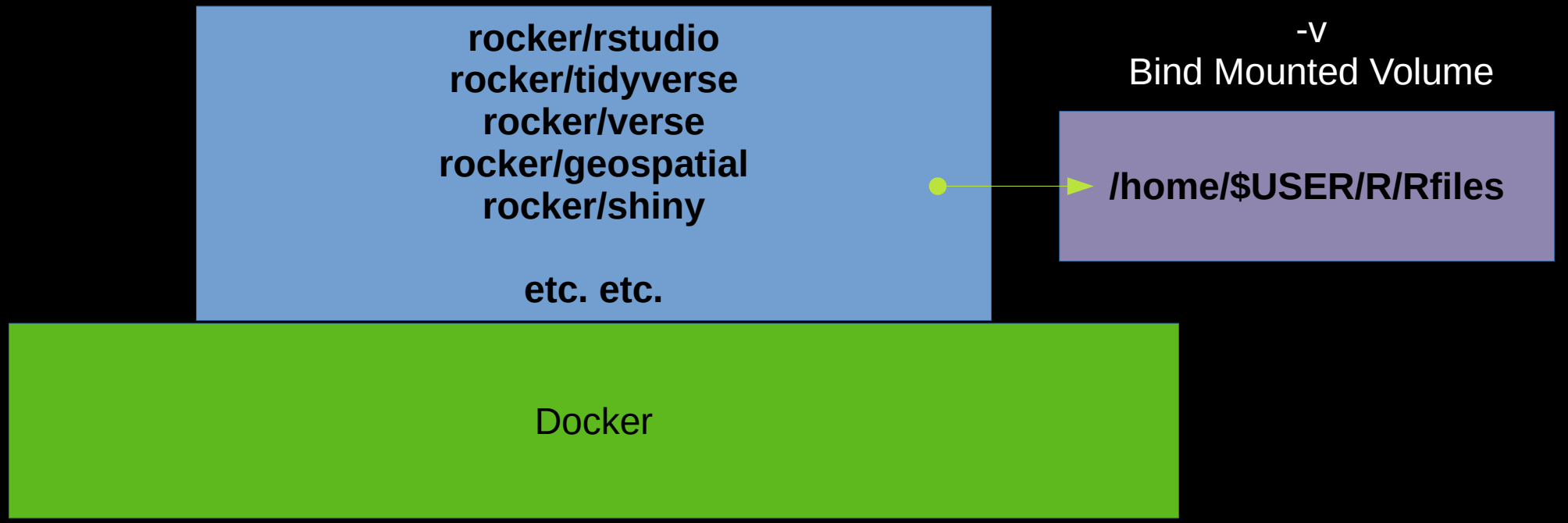
Jupyter at 50,000 Feet

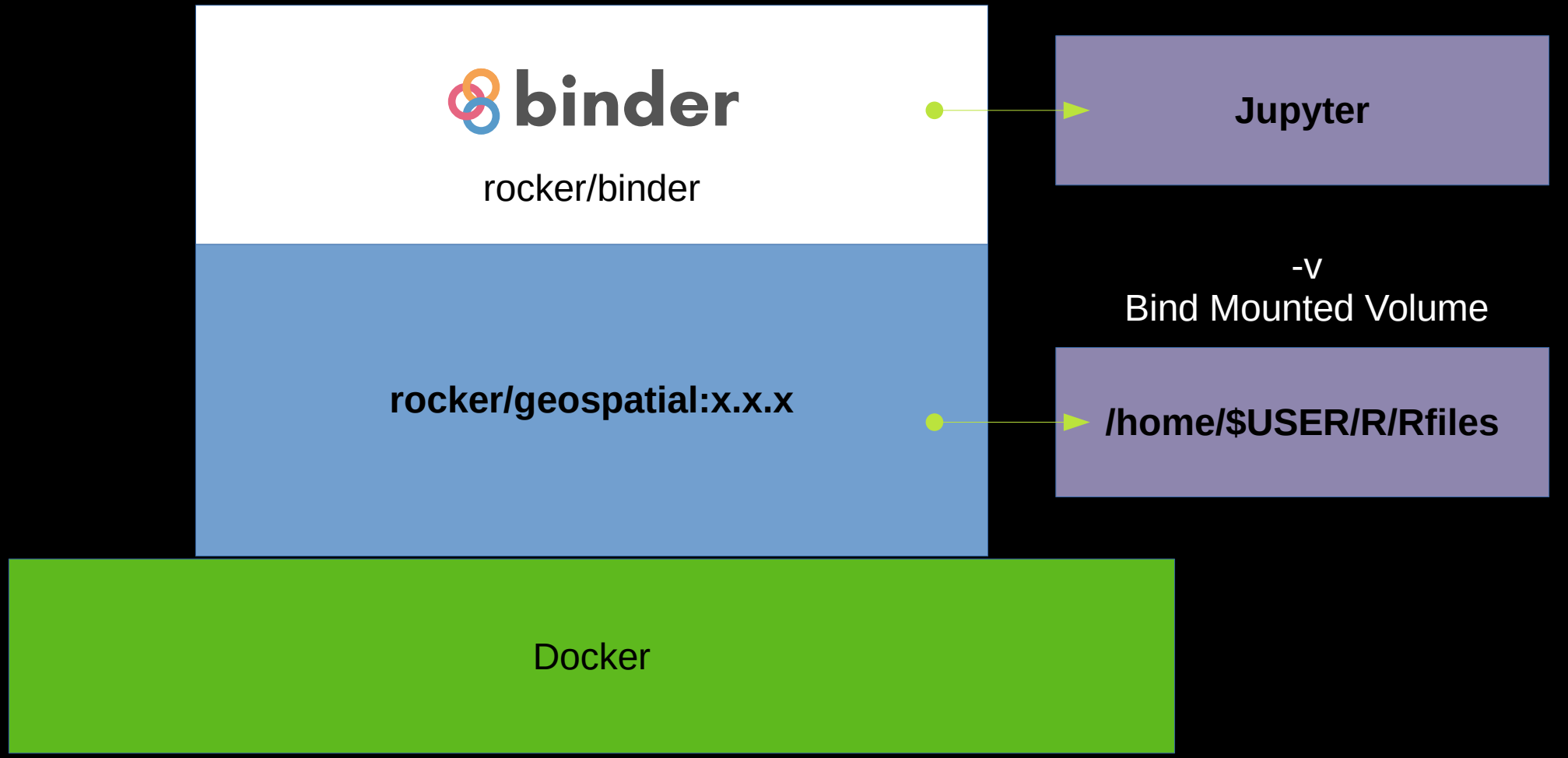
- Web application
- Allows you to create notebooks that contain live code



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


rocker/rstudio
rocker/tidyverse
rocker/verse
rocker/geospatial
rocker/shiny
etc. etc.
-v
Bind Mounted Volume
/home/$USER/R/Rfiles
Docker


binder
rocker/binder
Jupyter
-v
Bind Mounted Volume
rocker/geospatial:x.x.x
/home/$USER/R/Rfiles
Docker

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?

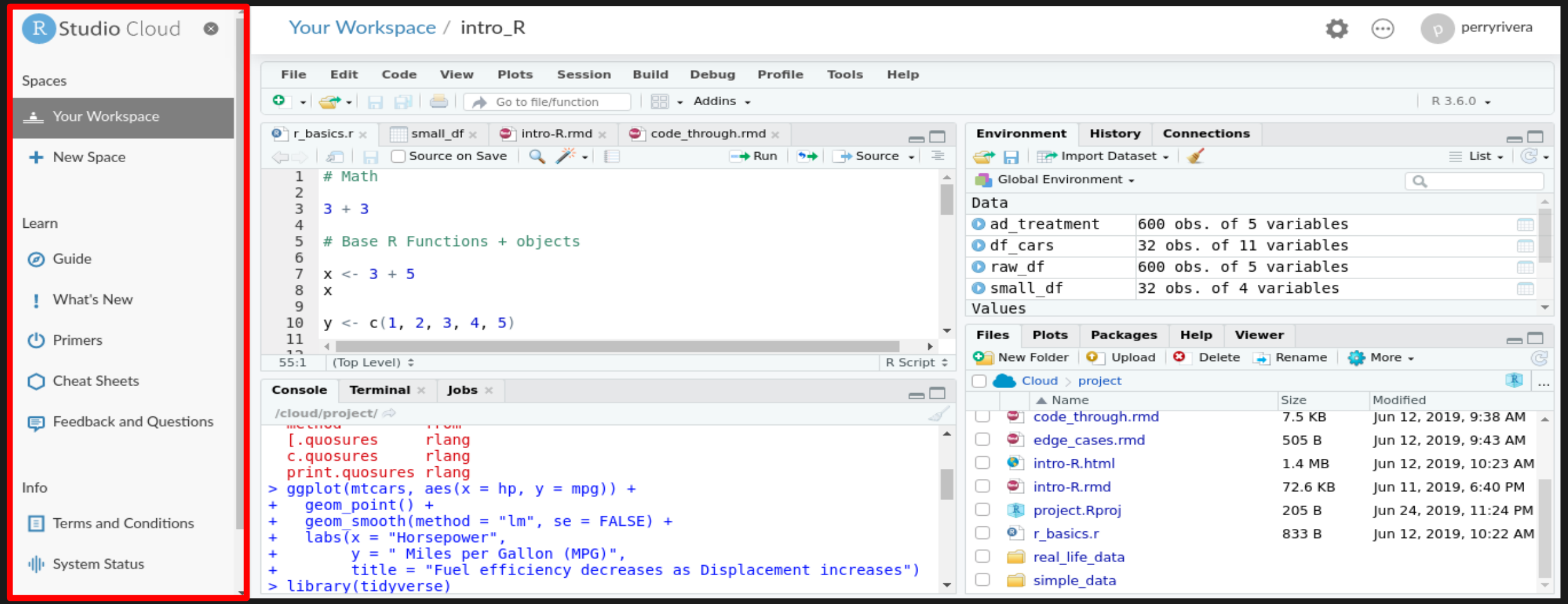
RStudio Cloud at 50,000 Feet

- 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

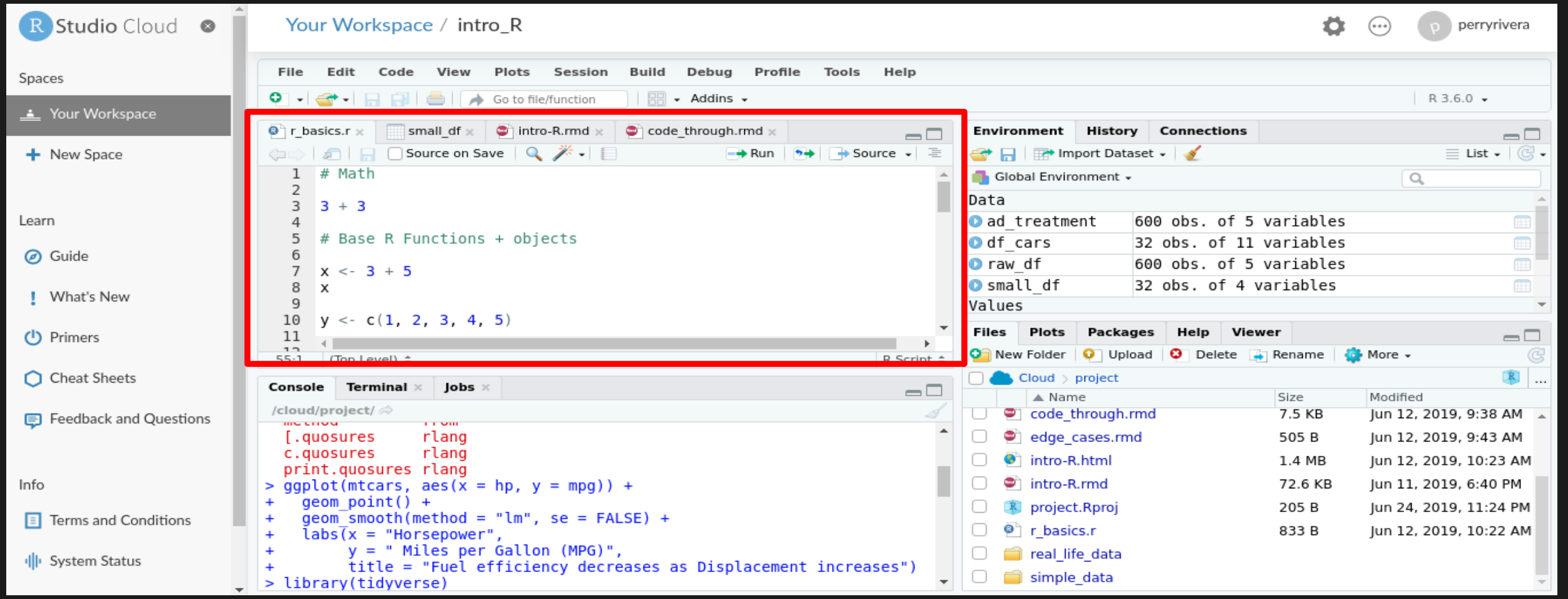
To the Cloud...

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

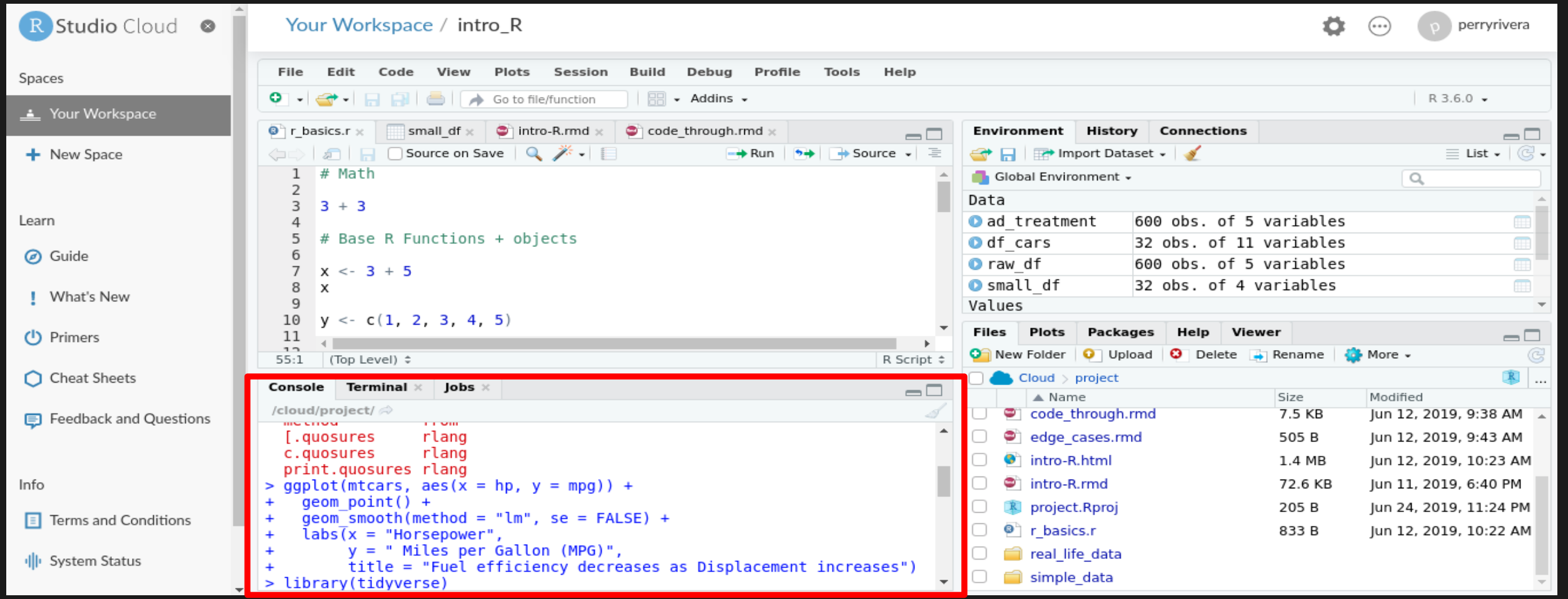
What It Looks Like



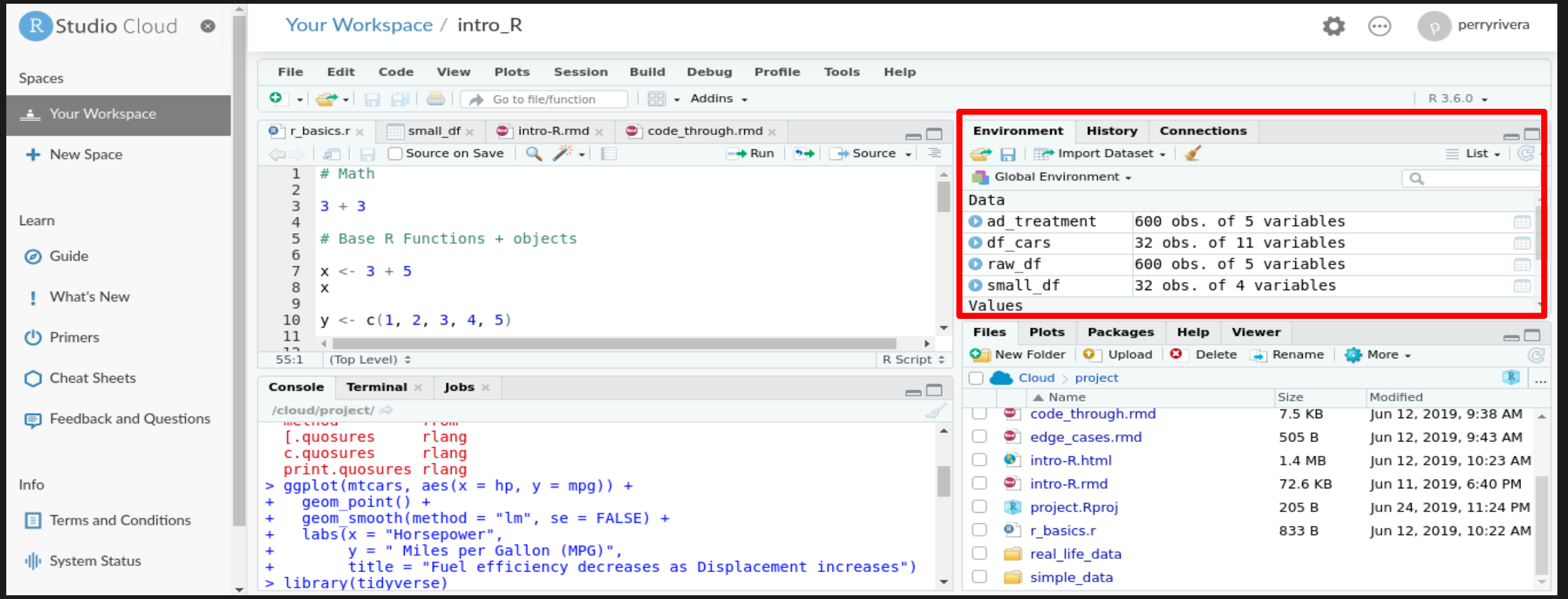
What It Looks Like



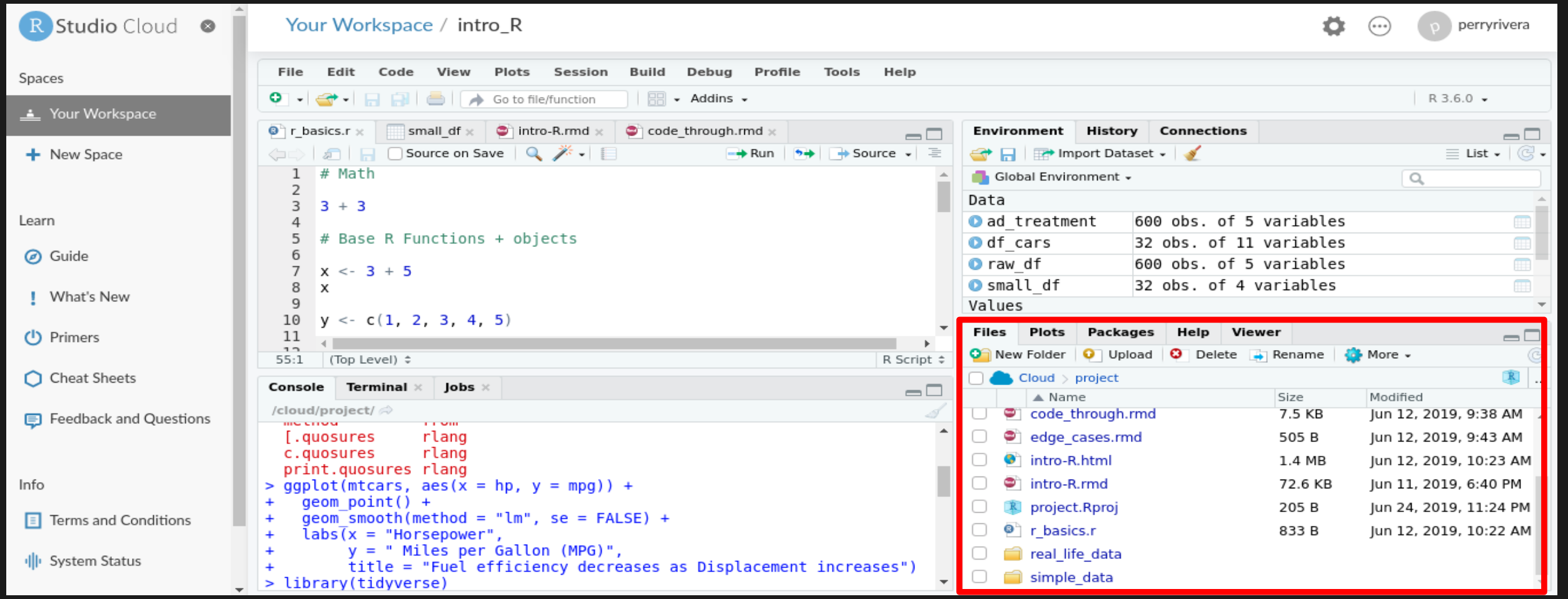
What It Looks Like



What It Looks Like



What It Looks Like



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	https://www.rocker-project.org/
Rocker Tidyverse on Github	Rocker/Tidyverse Image Information	https://hub.docker.com/r/rocker/tidyverse/
Rocker Binder	Rocker Binder	https://github.com/rocker-org/binder
RStudio Cloud	Hosted Rstudio Cloud Service	https://rstudio.cloud/

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
- LinkedIn: @pvrconsulting

- Slides:
  - https://tinyurl.com/yxe3q23y
- Github:
  - https://github.com/perryrivera/r_development_presentation

Thank You!

Perry Rivera | Rivera Consulting | perry-rivera@alumni.calpoly.edu