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?

Project Overview

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

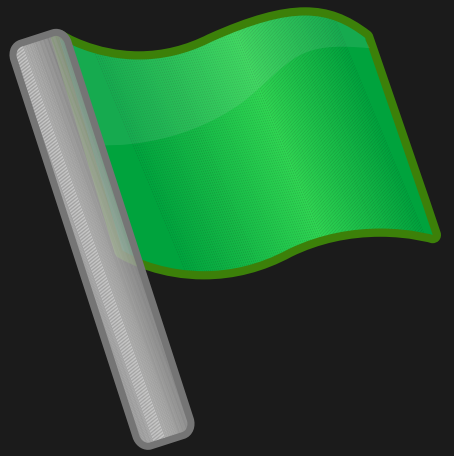
Observations

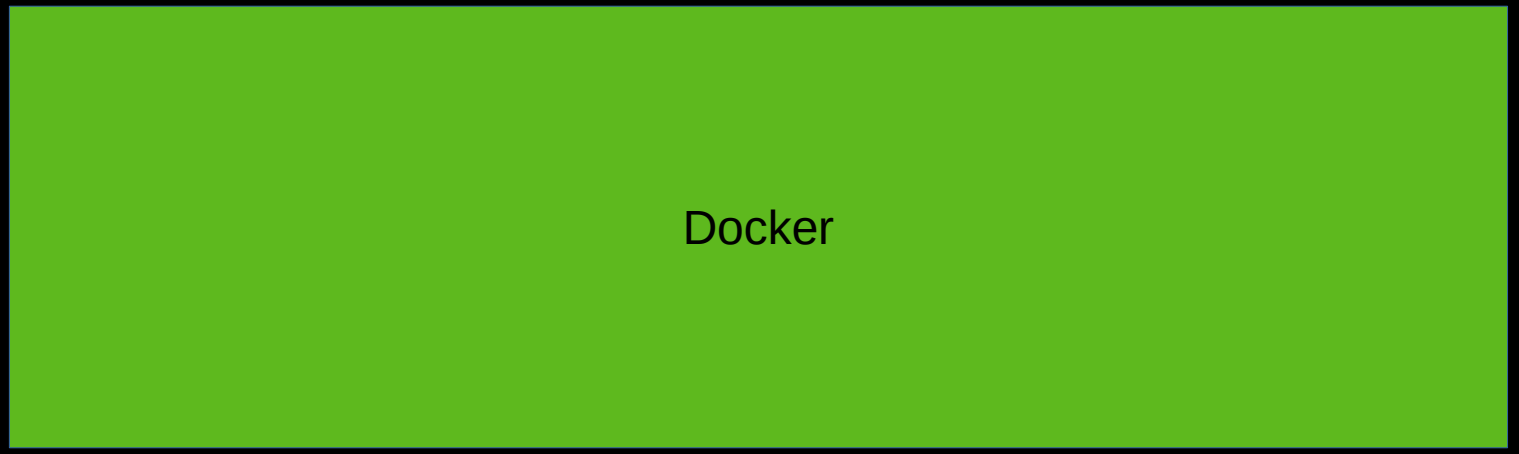
- Long compile times
- Aging laptop kept shutting off because Linux kept overclocking during compilation and system kept overheating
- Sooooo many dependencies and compilation issues/errors

Project Overview (Revised)

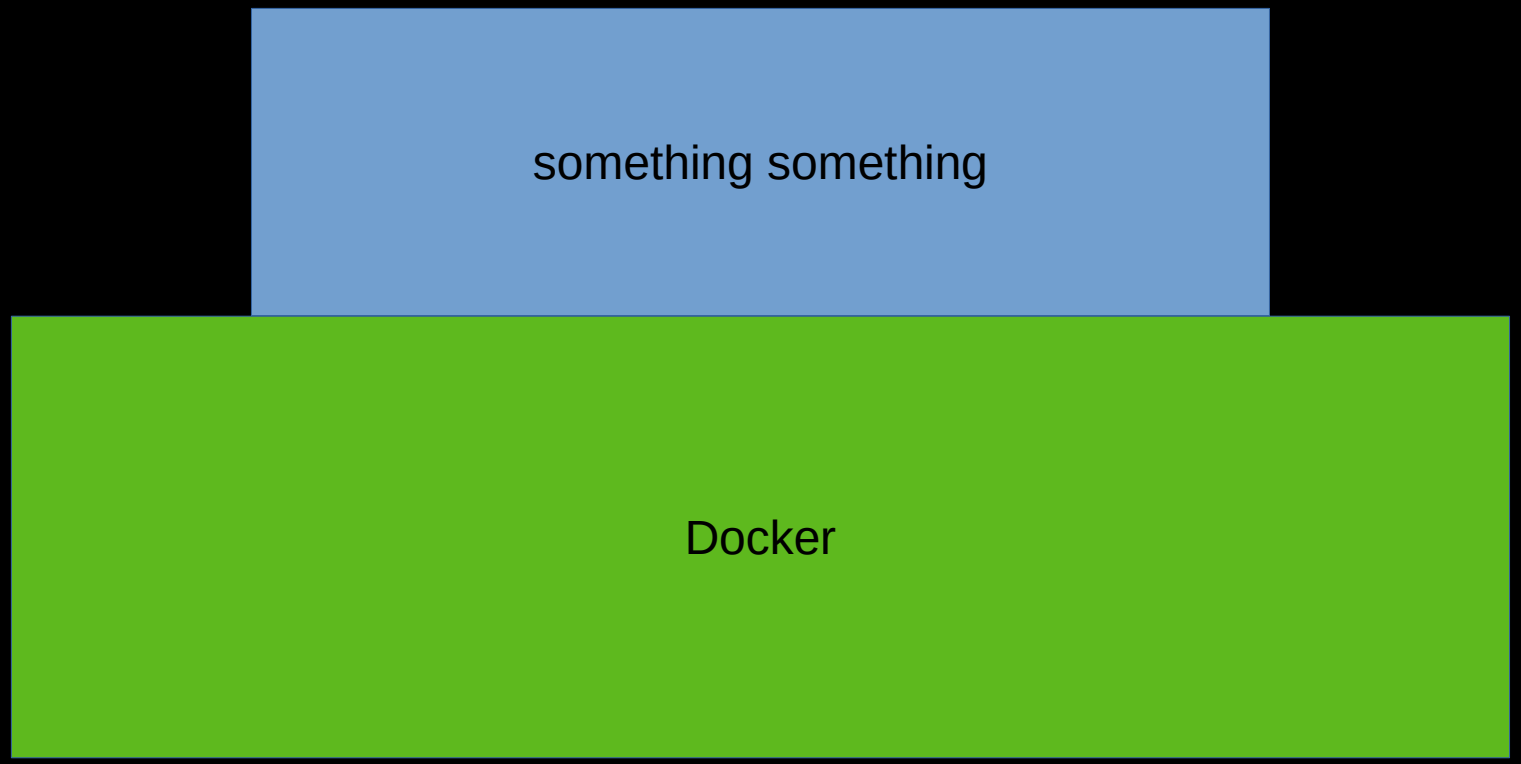
- Download R
- Download Rstudio
  - Download tidyverse package
- Compile everything.

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

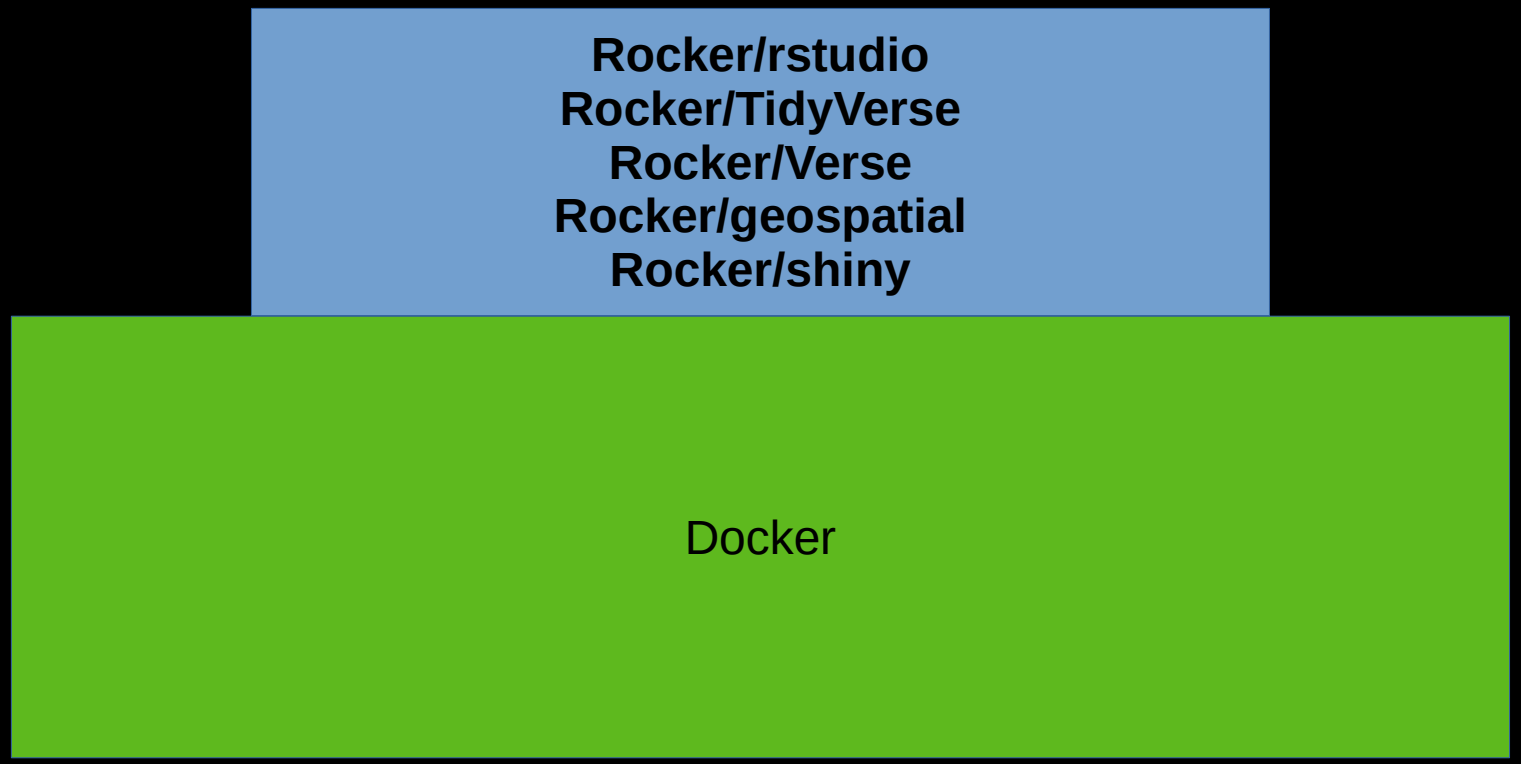




Overview



Overview

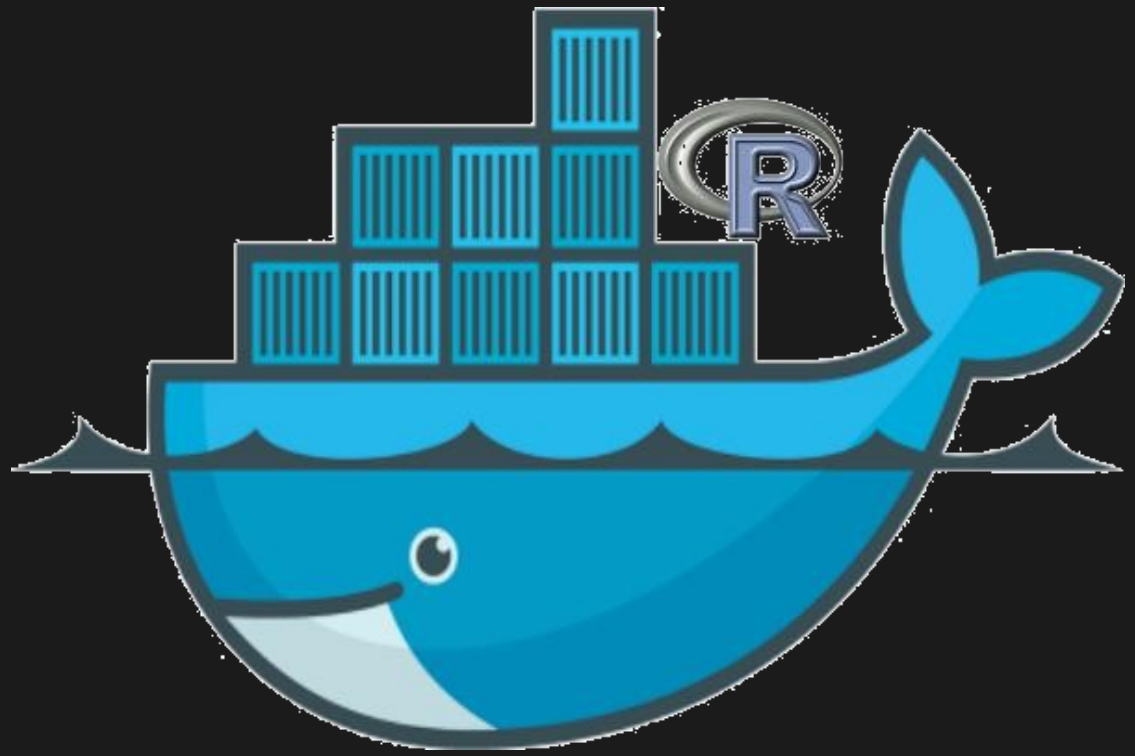

Rocker/rstudio
Rocker/TidyVerse
Rocker/Verse
Rocker/geospatial
Rocker/shiny
Docker

Docker at 50,000 Feet

- Runs software packages called containers
- Developers can package up applications, including libraries and dependencies into a container

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



Some R images in the Rocker Universe

Image	Comments
r-ver	• Minimal
rstudio	• Great mid-sized container, sans tidyverse
tidyverse	• Great mid-sized container
verse	• TeX and publishing
geospatial	• Large image

How do I download Docker?

Procedure

Step 1

- Download Docker

Step 2

- docker run rocker/rstudio

Step 3

- docker run rocker/tidyverse

Step 4

- Awesome Stuff… maybe?

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 -v /home/$USER/R/Rfiles:/home/rstudio --rm -p 8787:8787 rocker/rstudio
    - Replace SHHH with your own password.
      Note: don’t use rstudio for PASSWORD
    - -v is where your files live in the host operating system
    - --rm provides clean-up of the container when you’re done.
    - -p is the port number
    - Last argument (rocker/rstudio, etc.) is the desired image

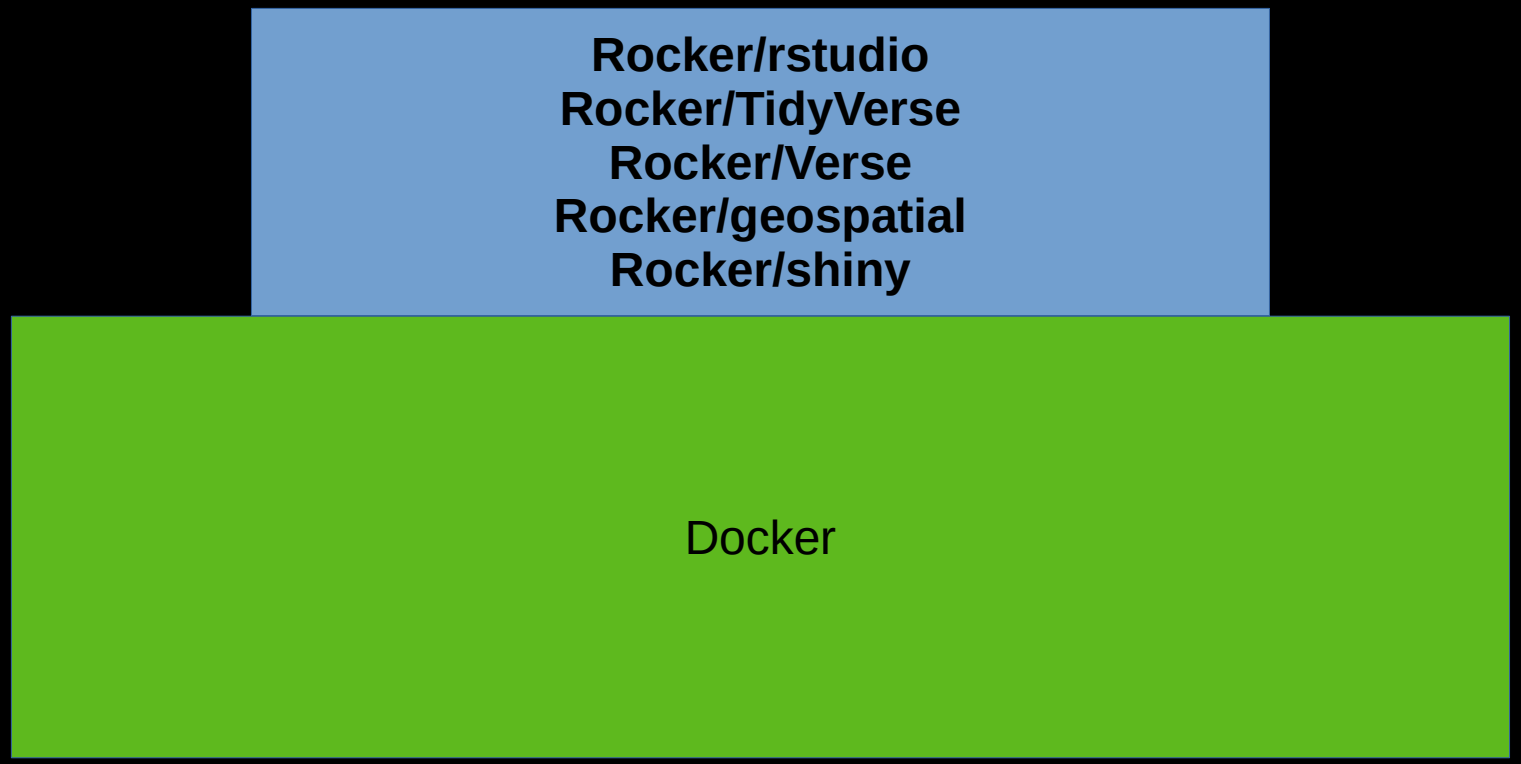
Post-configuration Docker (Part 1 of 2)

- Open up a browser
- Browse to:
  - localhost:8787
- ~~Set a working directory: setwd("/home/privera/R/Rfiles/Rfiles")~~

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)
- ~~Set a working directory: setwd("/home/privera/R/Rfiles/Rfiles")~~

How Do I Save My Stuff?

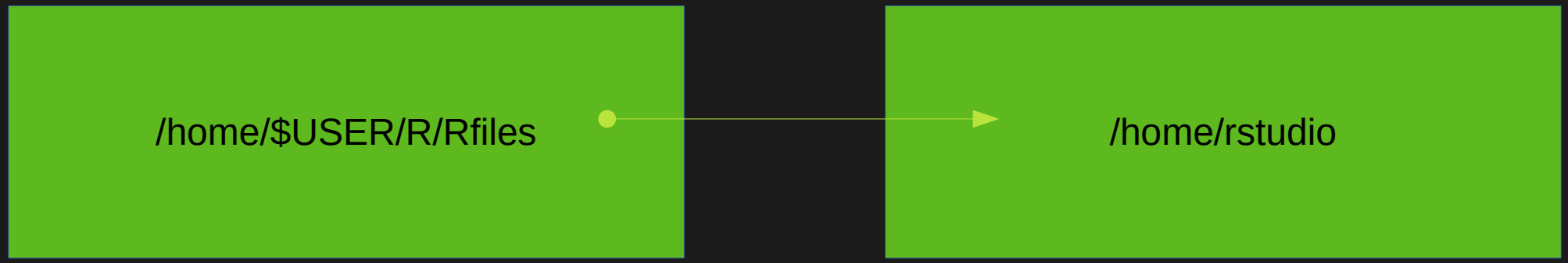

Rocker/rstudio
Rocker/TidyVerse
Rocker/Verse
Rocker/geospatial
Rocker/shiny
Docker

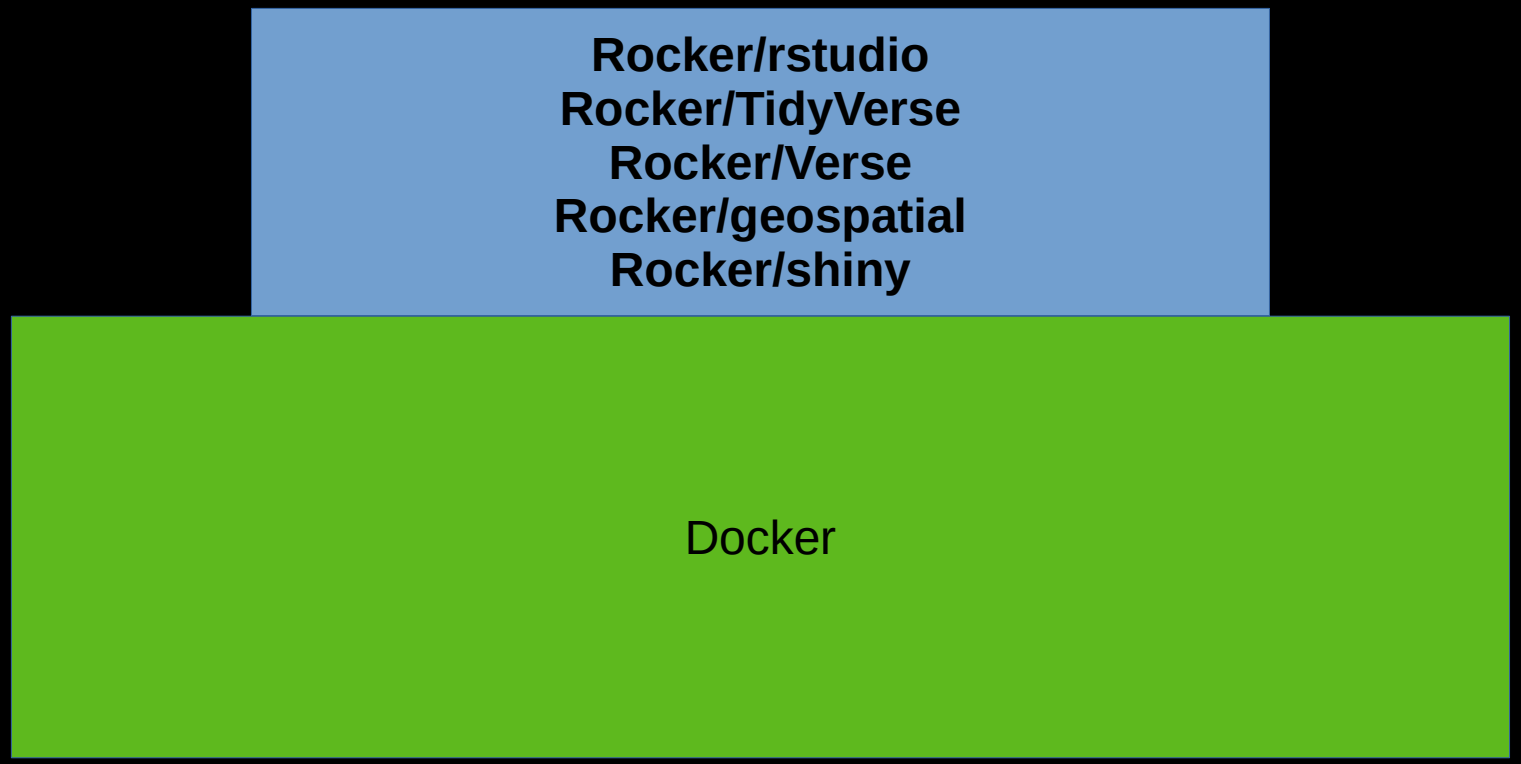
Uh Oh...There's a Problem!

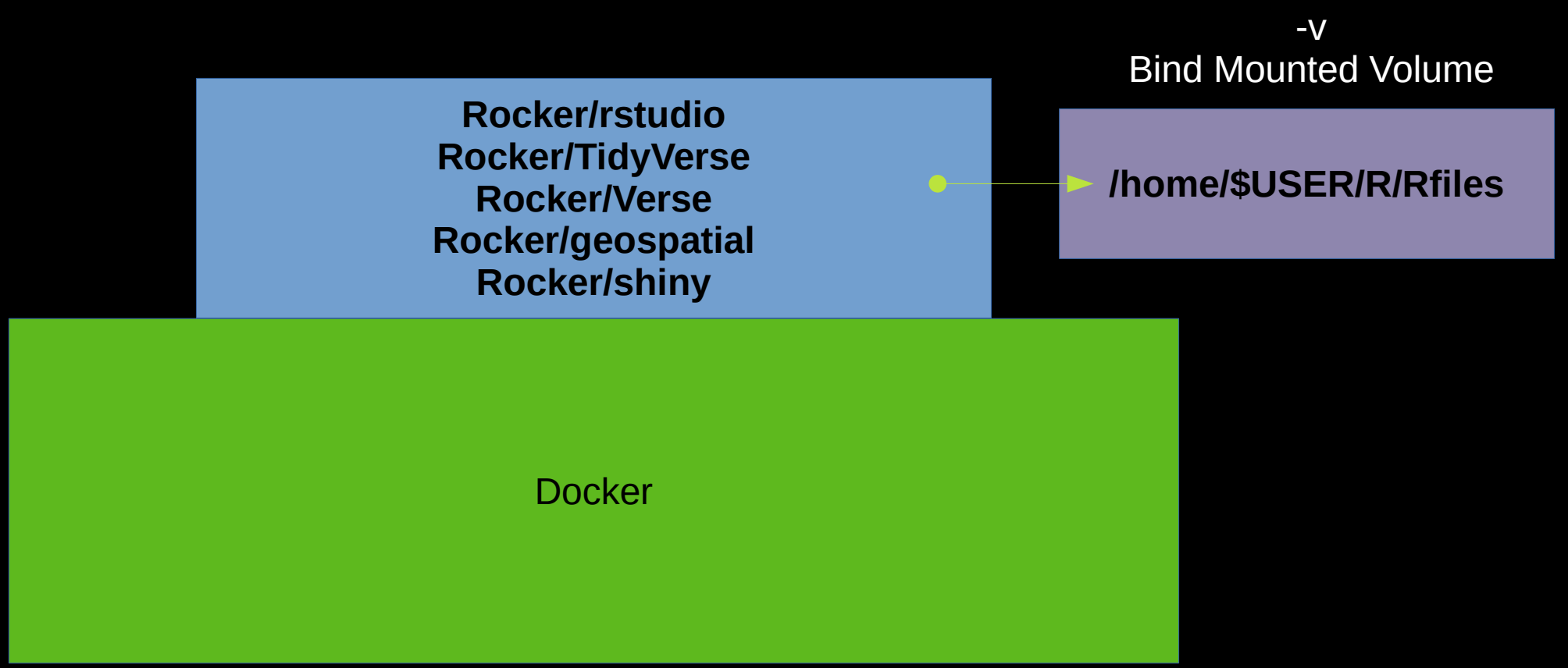
- Without the -v switch, can't save to the host operating system.
- So, use the -v switch
- Be sure to mkdir a directory, similar to:
  - /home/$USER/R/Rfiles
  - **Back that directory tree up somewhere!**
- So...for the command:
  - docker run -e PASSWORD=SHHH -v /home/$USER/R/Rfiles:/home/rstudio --rm -p 8787:8787 rocker/rstudio
  - The /home/$USER/R/Rfiles directory gets bind mounted to /home/rstudio in the container

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



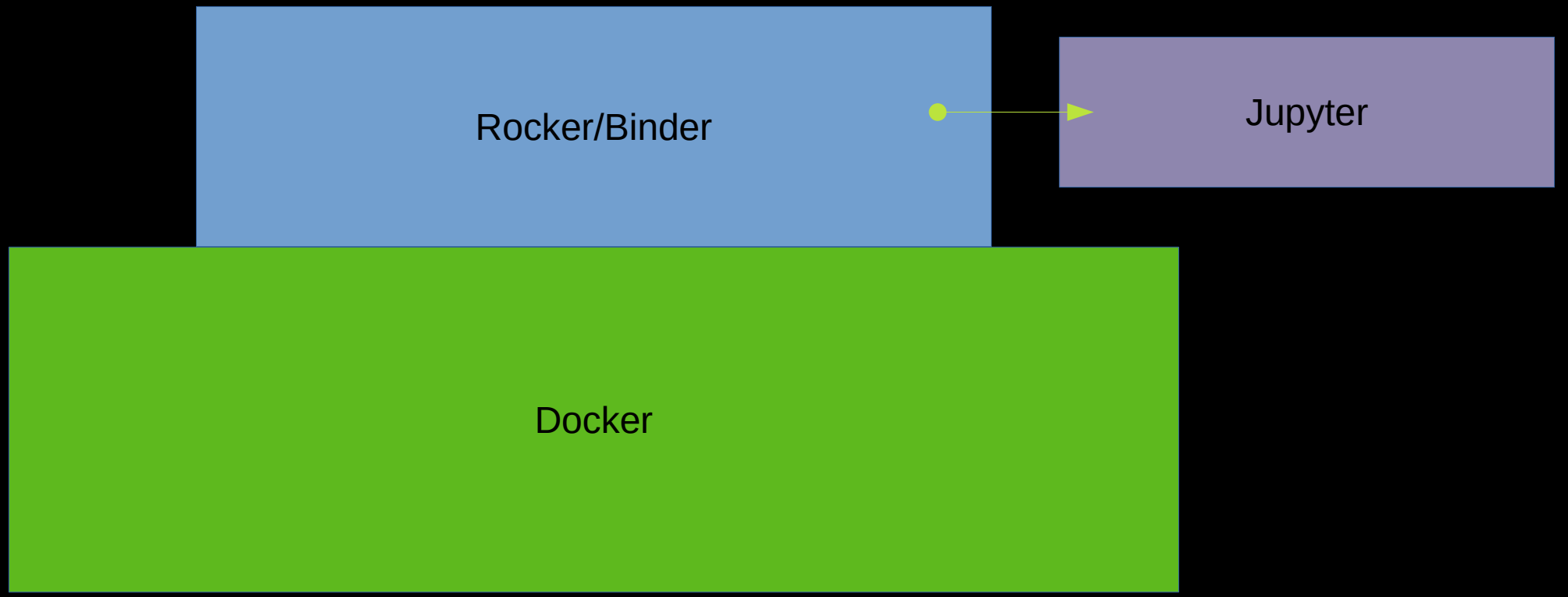

Rocker/rstudio
Rocker/TidyVerse
Rocker/Verse
Rocker/geospatial
Rocker/shiny
Docker


-v
Bind Mounted Volume
Rocker/rstudio
Rocker/TidyVerse
Rocker/Verse
Rocker/geospatial
Rocker/shiny
/home/$USER/R/Rfiles
Docker

What Next?

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



Overview

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 include that token in the URL you paste into the browser.

Live Demo Maybe?

What resources are available?

Materials (Part 1 of 2)

Materials (detailed list)	Description	Link
R for Data Science Book	Hadley Wickham Reference	https://r4ds.had.co.nz/
Hands-On Programming with R	Garrett Grolemund R-specific Learn By Doing Book	http://shop.oreilly.com/product/0636920028574.do
Various R related cheatsheets	R cheatsheets	https://www.rstudio.com/resources/cheatsheets/

Materials (Part 2 of 2)

Materials (detailed list)	Description	Link
Datacamp Subscription	Datacamp Training Courses	https://www.datacamp.com/home
ggplot2 Book	Hadley Wickham ggplot Reference	https://www.amazon.com/dp/0387981403/ref=cm_sw_su_dp?tag=ggplot2-20

Websites (1 of 1)

Materials (detailed list)	Description	Link
Sharpsight Tutorial	Beginner Tips from Sharpsight Labs	https://www.sharpsightlabs.com/blog/ggplot2-tutorial/
Sharpsight Tutorial	Histogram and Density Plot	https://www.sharpsightlabs.com/blog/ggplot-histogram/
R-Statistics Tutorial	Complete ggplot2 Tutorial (3 parts!)	http://r-statistics.co/Complete-Ggplot2-Tutorial-Part1-With-R-Code.html

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

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