



# Notes from rstudio::conf2020

---

Isabelle Valette

03-03-2020

Hilton San Francisco Union Square

# rstudio::conf 2020 schedule

WEDNESDAY, JANUARY 28

## Breakfast

8:00a-9:00a

CONTINENTAL BALLROOM & GOLDEN GATE

## Keynote 1

9:00a-10:00a

GRAND BALLROOM

Welcome to rstudio::conf 2020

J.J. Allaire, RStudio

## Keynote 2

10:00a-11:00a

GRAND BALLROOM

Data, visualization, and designing with AI

Fernando Viegas & Martin Wattenberg, Google

## Session 1

11:30a-1:00p

GRAND BALLROOM B

Deploying End-To-End Data Science with shiny, plumber, and rproj

Aleks Gold, RStudio

We're hitting R a million times a day so we made a talk about it

Massimo M., Mobile & D. Jacqueline Nolis, Nolis, LLC

Growth Hacking with R - Product Analytics at Scale using R and RStudio

Andrew Mongomery, Salesforce

Practical Plumber Patterns

James Blair, RStudio

GRAND BALLROOM

Education

IMPERIAL BALLROOM

Meet the RDA

Lucas Cope, RDA

Data Science Education in 2022

Carin Maier & Gregor Gorjanc

How Vibrant Emotional Health Connected Slided Data Sources and Streamlined Reporting Using R

David Littman, DataCamp

Building a free data science pipeline for the FT with RStudio Connect

George Kotsiopoulos, Financial Times

How to win an AI Hackathon, without using AI

Com Slepnev, Jumping Rivers

Case Study

IMPERIAL BALLROOM

Professional Case Studies

Kate Hoswill, RStudio

Data Science Education in 2022

Carin Maier & Gregor Gorjanc

How Vibrant Emotional Health Connected Slided Data Sources and Streamlined Reporting Using R

David Littman, DataCamp

Building a free data science pipeline for the FT with RStudio Connect

George Kotsiopoulos, Financial Times

How to win an AI Hackathon, without using AI

Com Slepnev, Jumping Rivers

Programming

PLAZA ROOM

Simplified Data Quality Monitoring of Dynamic Longitudinal Data: A Functional Programming Approach

David Gutierrez, RStudio

vctr: Creating custom vector classes with the vctr package

Jesse Soule, Loyola Marymount University

Asynchronous Programming in R

Matthew Chang, RStudio

Asura Pipelines and GitHub Actions

Jim Hester, RStudio

# The more the merrier :)

Keynote 3  
9:00a-10:00a

Object of new disclosure is not substitutable  
Jerry Bryant, RStudio

Session 4  
10:30a-12:00p

Visualization  
GRAND BALLROOM B

The Glamour of Graphics  
William Cleveland, University of Pennsylvania

3D ggplot with raphader  
D. T. Wickham, Wick, Inventive for Defense Analysis

Designing Effective Visualizations  
Miriah Meyer, University of Utah

Third-Party Tools 2020  
Tibbles.org, Hadley Wickham

Communication  
GRAND BALLROOM

Creating and Publishing Reports with R Markdown  
R. Michael Weyand, RStudio

Do I repeat myself? Stop yourself! Repeated reporting in the universe  
Mike Lawrence, Shiny Developer

New Features in R shiny 1.7  
New Features in shiny 1.7  
New Features in shiny 1.7  
One R Markdown Document, Fourteen Demos  
Yihui Xie, RStudio

Workflow  
IMPERIAL BALLROOM

RMarkdown Driven Development  
Emily Reddick, Captain One

Reactive Components to R  
Kevin Ushey, RStudio

rstudio::ui: Sneak Preview  
Jonathan McPherson, RStudio

Using Jupyter with RStudio Server Pro  
Karl Fauschuk, RStudio

Medicine  
PLAZA ROOM

Building a Medical Device with R  
Austin Miller, RStudio

Developing an web-based clinical decision support application for patient translation management  
Justin Juskevich, Mayo Clinic

Predicting Platelet Blood Bag Demand to Reduce Inventory Waste at the Stanford Blood Center  
CJ O'Donnell, Stanford Blood Center

Shiny New Things: Using R to Bridge the Gap in EMR Reporting  
Mr. Brendon Graham, Children's Hospital of Philadelphia

Modeling  
GRAND BALLROOM B

MLops for R with Azure Machine Learning  
David Gómez, Microsoft Research

Tidy Tip: Tuning Techniques  
Moritz Krause, RStudio

Neural Networks for Longitudinal Data Analysis  
Dr. Syndeika Watson, Komodo Data Insights, Dick Neigus

Stochastic Block Models with R: Statistically rigorous clustering with marginal code  
Ricardo Fraire, Universidad de Valencia

ggridges2  
GRAND BALLROOM A

Best practices for programming with ggridges2  
David Gómez, Microsoft Research

Spruce up your ggridges2 visualizations with formatted text  
Claus Wilke, University of Texas at Austin

The little package that could: taking visualizations to the next level with the scales package  
Claus Wilke, University of Texas at Austin

Extending your ability to extend ggridges2  
Pharrell Williams, University of Valencia

Organizational Thinking  
IMPERIAL BALLROOM

Small Team, Big Value: Using R to Design Visualizations and Build Data Specs  
Alice Schep, Outlier AI

Unidirectional are real  
Dr. Priscila Sobera, Moffit Cancer Center

Data Science in Meatspace  
BenJaboula Government, Poetry Unlimited, Inc.

Value in Data Science Beyond Metrics in Production  
Esteban Arrieta de la Rubia

Programming  
PLAZA ROOM

Autodriving: UI-less development: Building an R API for building UI-less Specs  
Alice Schep, Outlier AI

Bringing the gap between SQL and R: Introducing queryable and tidyquery  
BenJaboula Government

Use columns in data.table: Reducing the cognitive and computational burden when working with complex ryan.scorrett, Ohio State University

Advances in Shiny  
Loren Henry, RStudio

Panel  
GRAND BALLROOM B

Career Advice for Data Scientists  
Sarah E. Johnson, DataCamp

Systematic Writers, Data  
Giovanni De Querida, Alveo Inclusive and R-Ladies

David Keyes, If for the Rest of Us

Lightning Talks  
GRAND BALLROOM A

Colin Rundall, University of Nottingham

Any Good? Benji Art, DataCamp

Rebecca Butler, UC Berkeley

Me, You, RStudio, and the rest of us  
Mike Weyand, RStudio

Ali Mirza Ortíz Montes, COABIO

Carsten Stach, University of Göttingen

Anna Domínguez, R

Theresa Ayres, Merlin School

Audrey Tang, California Polytechnic State University

Katherine Simmon, Northwestern University

Amelia Gottlieb, RStudio

# What to choose





# Agenda

---

- Travel 8000 km?
  - How: Funding & Practical info
  - Why go to Rstudio::conf?
- Applied ML Workshop
- Rstudio::conf Highlights
  - Keynotes
  - Interesting packages

# How and why to travel 8000 km?

---

# Funding and Practical tips



Get a scholarship



Academia gets a discounts



Don't miss the early bird



Present what you do



Be creative – Airbnb is great



Come early and stay longer



Community

# The Workshop

---

Day 1 and Day 2  
Applied Machine Learning

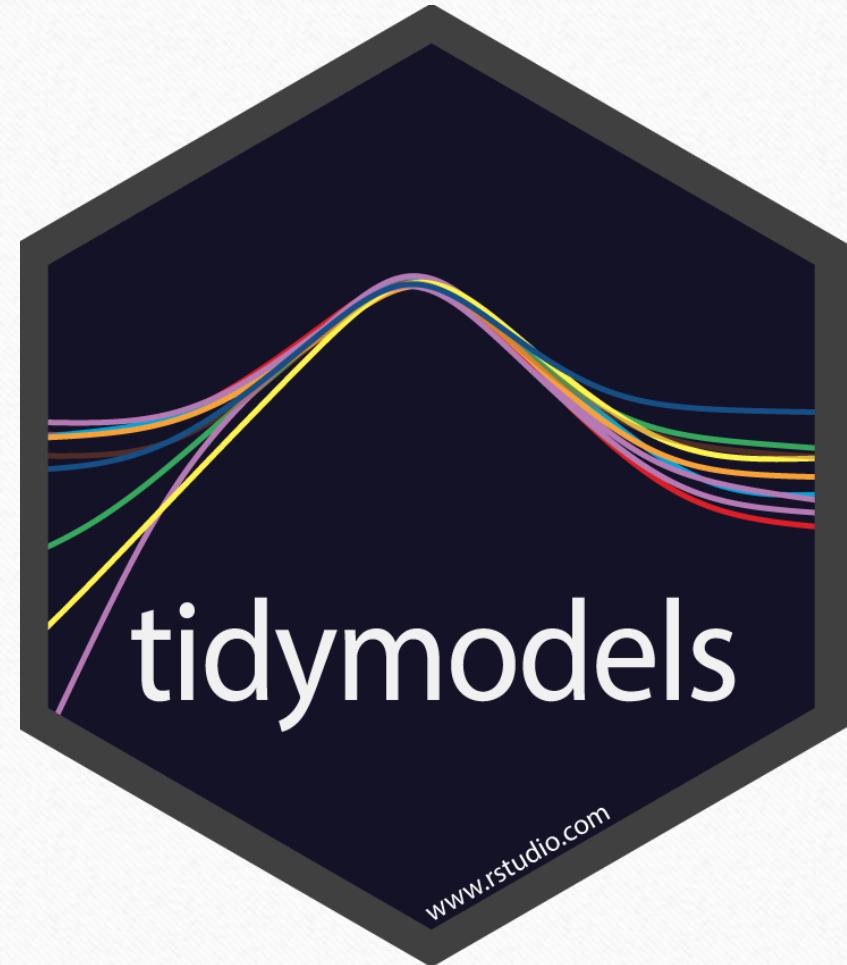
# Workshop - Applied Machine Learning

---

- Instructors Max Kuhn and Davis Vaughan
- It is massive with at least 200 attendees
- Covered ML workflow and code for regression and classification models (process wise)
- See RMarkdown files for more details



# ML Framework

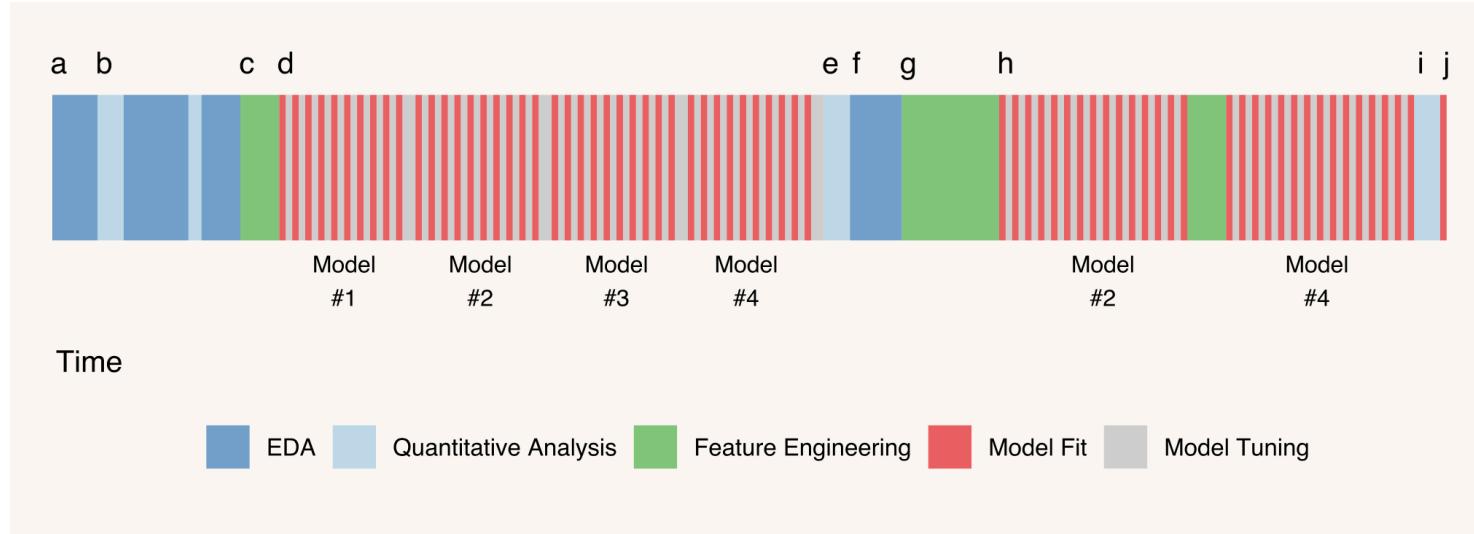




DEMO

## What did I get from the workshop?

- Useful wrapper functions
- More maintainable code
- Simplified code and simplified workflow with a module based approach
- Integration of all packages
- Possibilities for global model evaluations
- Connection possibilities to powerful engines
- Preview of coming attraction: future integration of mining tools with text mining tools thanks to Julia Silge



## Book Signing

- RStudio gave more than 1000 books for free in 2 hours
- All books were signed by the authors with personal comments
- Got all of them ...except Hadley's Advanced R

# Feature Engineering and Selection

A Practical Approach for Predictive Models

ISA,  
THANKS FOR COMING  
TO THE WORKSHOP!

MAX

# The Conference

---

Day 3 and Day 4

Why travel 8000 km?

I am here

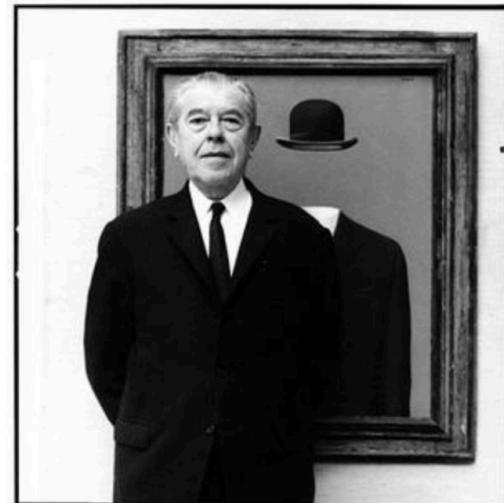


# Art Quiz: How are R and René Magritte related?

---



HINT: library(magrittr)



René Magritte



# Keynotes

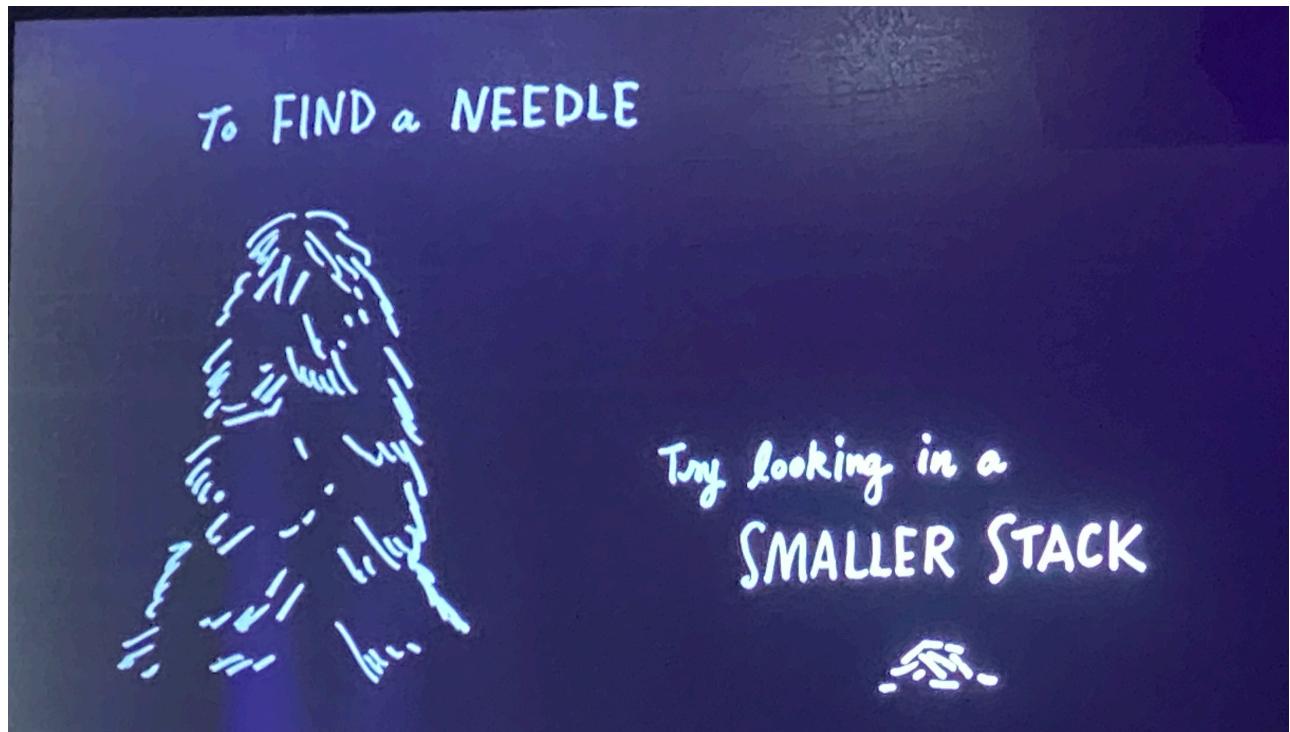
---

JJ Allaire and  
news from  
RStudio



RStudio reborn as a  
Public Benefit Corporation  
– commits itself  
to open source,  
not shareholders

# Debugging with Jenny Brian





lifecycle	deprecated	vs	lifecycle	superseded
This function is on its way out in the near future (<2 years)			Alternative is easier to use/more powerful/faster	
You'll be warned when you use it.			Learn new approach when you have some free time	
Goal is to go through questioning or retired phase first.			Not going anywhere, but will only critical bug fixes	
e.g. dplyr::tbl_df(), dplyr::do()			e.g. spread(), gather()	

- Experimental
- Superseded
- Deprecated

1. Where are we today?
2. Tidy eval post-mortem
3. Lifecycle updates

{ } var }

# The Learning Tracks

---

# R packages for documents, reports, CVs webpages

---





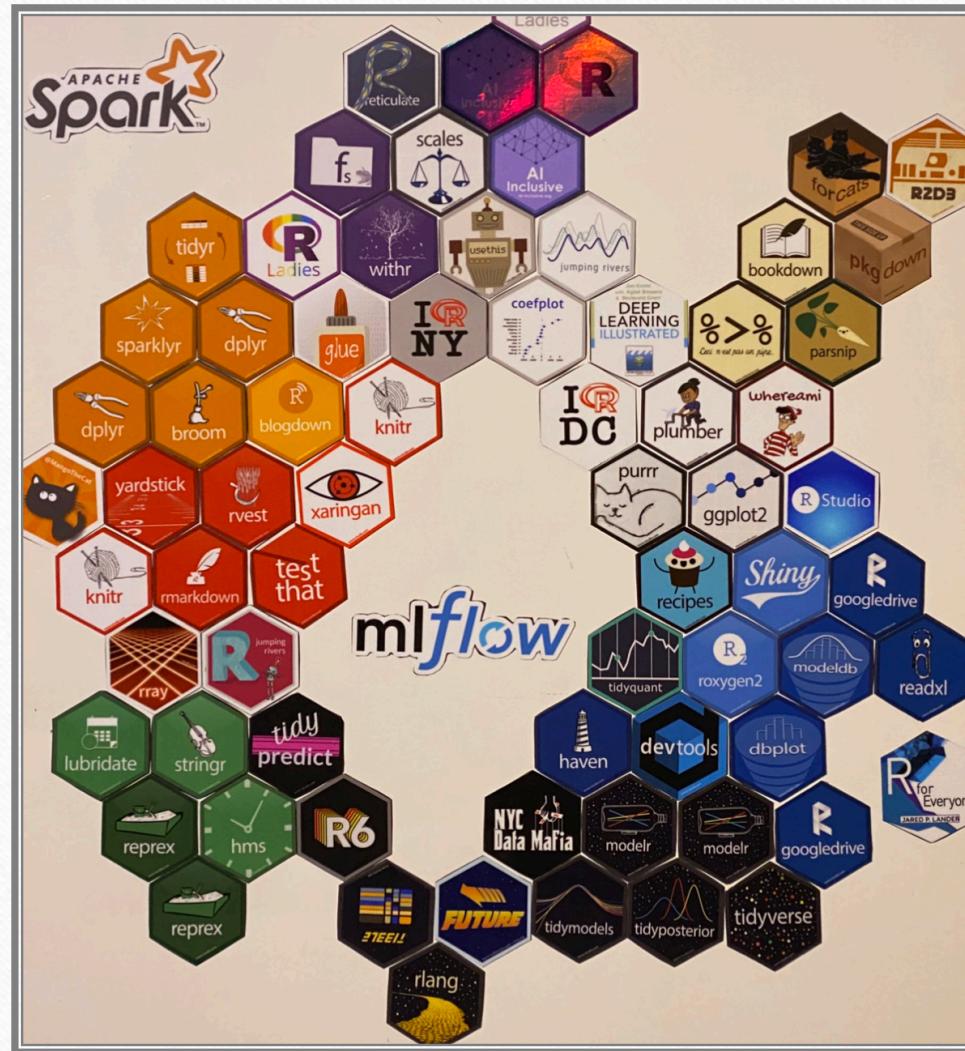
DEMO

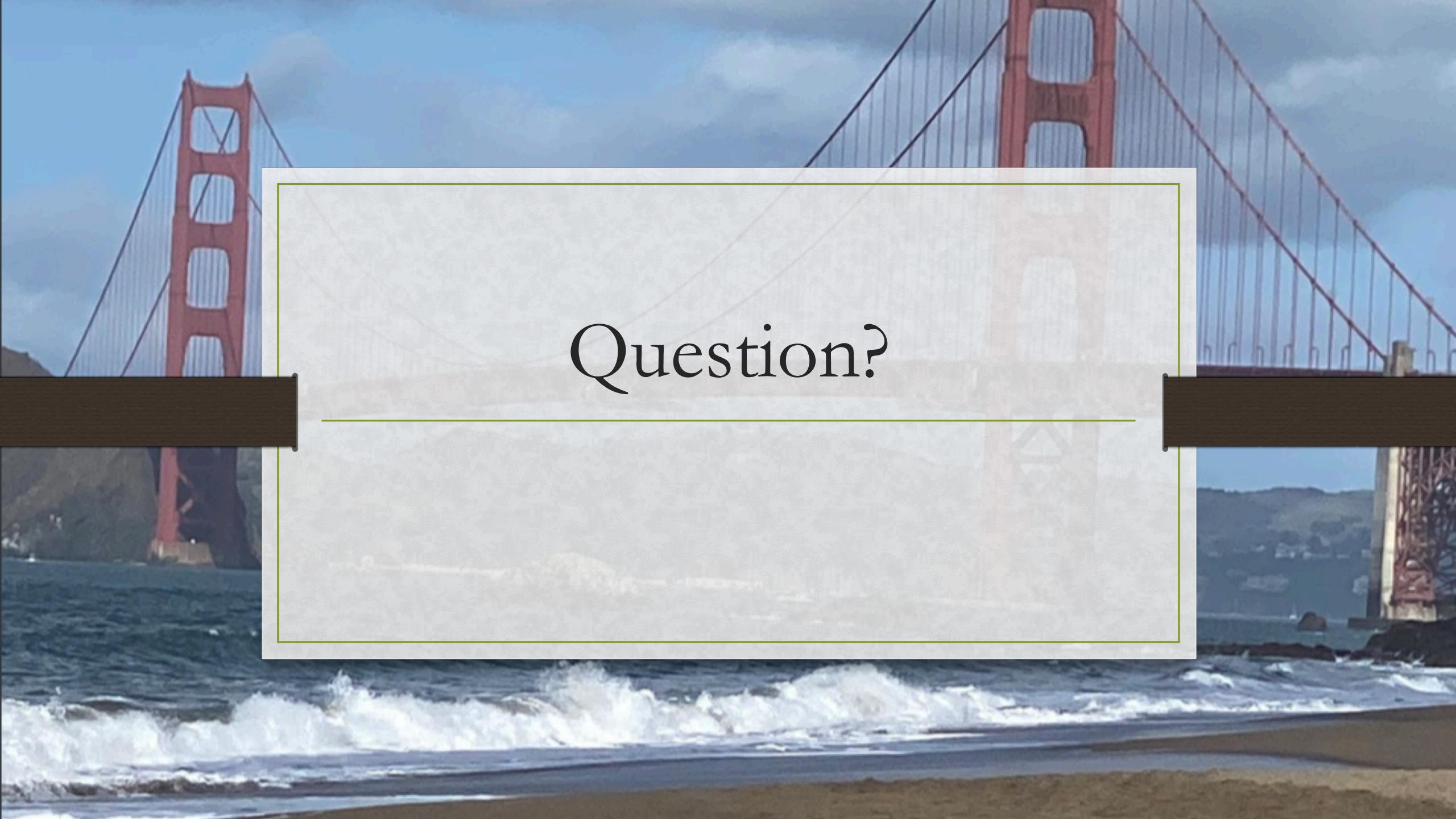
# Building 3D Visualization in R



# Ressources

- Workshops: <https://github.com/rstudio-conf-2020>
    - Keynotes: <https://resources.rstudio.com/rstudio-conf-2020>
    - GitHub repo for Rladies Oslo: [https://github.com/rldadies/meetup-presentations\\_oslo](https://github.com/rldadies/meetup-presentations_oslo)
    - GitHub repo for webpage: <https://github.com/ivalette/ivalette.github.io>





Question?

---