# Revolutionize Clinical Trial Data Exploration with *teal*

Dony Unardi

BBSW 2023
Embracing Change and Shaping the Future

# Agenda

1. Background
2. Introduction to teal framework
3. Using teal as Data Scientist
4. Anatomy of teal and key features
5. teal-verse product map
6. Collaboration

# Product Development Data Sciences

**Main task: clinical trials reporting**

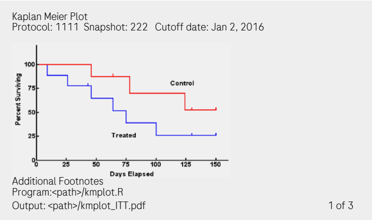
Summarising safety and efficacy data

Providing an accurate picture of trial outcomes

Managing data collection across international sites

# Improve efficiency in the way we work

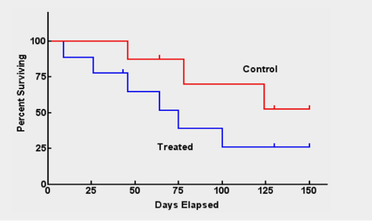
Analyzing clinical trial data requires multiple ways of presenting and interacting with our data



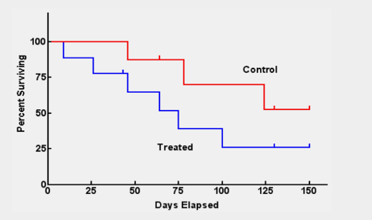
Kaplan Meier Plot
Protocol: 1111  Snapshot: 222  Cutoff date: Jan 2, 2016

Additional Footnotes
Program:<path>/kmplot.R
Output: <path>/kmplot_ITT.pdf
1 of 3

Per-SAP static output
on **OS**



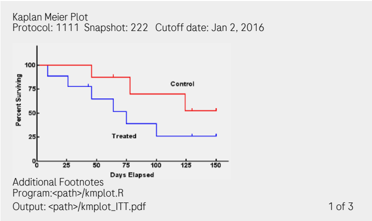
Ad-hoc analysis
on **PFS**



Ad-hoc analysis on **PFS**
in **subpopulation**

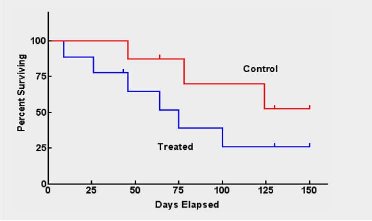
# Improve efficiency in the way we work

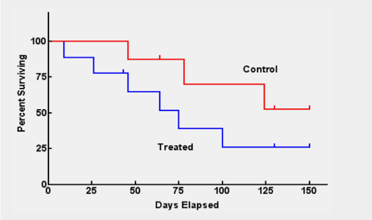
Analyzing clinical trial data requires multiple ways of presenting and interacting with our data



Kaplan Meier Plot
Protocol: 1111 Snapshot: 222 Cutoff date: Jan 2, 2016

Additional Footnotes
Program:<path>/kmplot.R
Output: <path>/kmplot_ITT.pdf
1 of 3

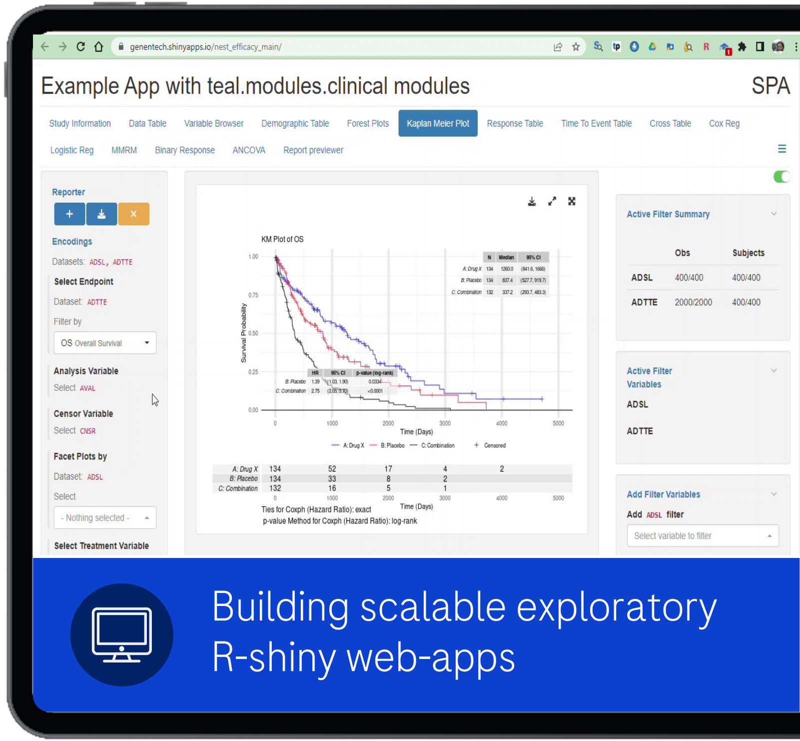




Per-SAP static output
on **OS**

Ad-hoc analysis
on **PFS**

Ad-hoc analysis on **PFS**
in **subpopulation**



Example App with teal.modules.clinical modules

**Encodings**

Datasets: ADSL, ADTTE

**Select Endpoint**

Dataset: ADTTE

Filter by: OS Overall Survival

**Analysis Variable**

Select AVAL

**Censor Variable**

Select CNSR

**Facet Plots by**

Dataset: ADSL

Select: - Nothing selected -

**Select Treatment Variable**

KM Plot of OS

| | N | Median | 95% CI |
|---|---|---|---|
| A: Drug X | 134 | 1260.0 | (841.6, 1666) |
| B: Placebo | 134 | 837.4 | (527.7, 919.7) |
| C: Combination | 132 | 337.2 | (293.7, 483.3) |

| | HR | 95% CI | p-value (log-rank) |
|---|---|---|---|
| B: Placebo | 1.39 | (1.03, 1.90) | 0.0334 |
| C: Combination | 2.75 | (2.05, 3.70) | <0.0001 |

| | 0 | 1000 | 2000 | 3000 | 4000 | 5000 |
|---|---|---|---|---|---|---|
| A: Drug X | 134 | 52 | 17 | 4 | 2 | |
| B: Placebo | 134 | 33 | 8 | 2 | | |
| C: Combination | 132 | 16 | 5 | 1 | | |

Time (Days)

Ties for Coxph (Hazard Ratio): exact
p-value Method for Coxph (Hazard Ratio): log-rank

**Active Filter Summary**

| | Obs | Subjects |
|---|---|---|
| ADSL | 400/400 | 400/400 |
| ADTTE | 2000/2000 | 400/400 |

**Active Filter Variables**

ADSL

ADTTE

**Add Filter Variables**

Add ADSL filter

Select variable to filter

Building scalable exploratory R-shiny web-apps

# What is {teal}?

A Rshiny-based interactive data exploration framework

Modularized and standardized building blocks

Collection of specialized R packages

Streamlines creation of web-apps that offers:
- Dynamic filtering facility
- Code reproducibility
- Reporting engine
- Many data summarization and visualizations

# How does {teal} work?

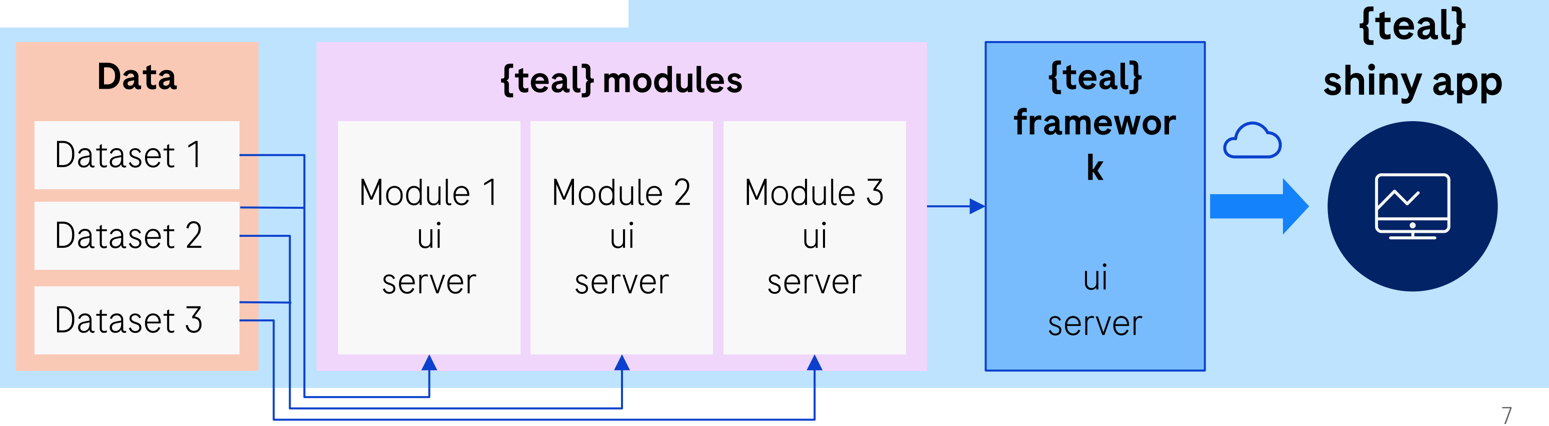
## Traditional shiny app

**Data**
Dataset 1
Dataset 2
Dataset 3

**Shiny**
ui
server

**{teal} framework** with modularized components

- Data agnostic
- Flexible
- Cumulative
- Collaborative
- Crowdsourcing



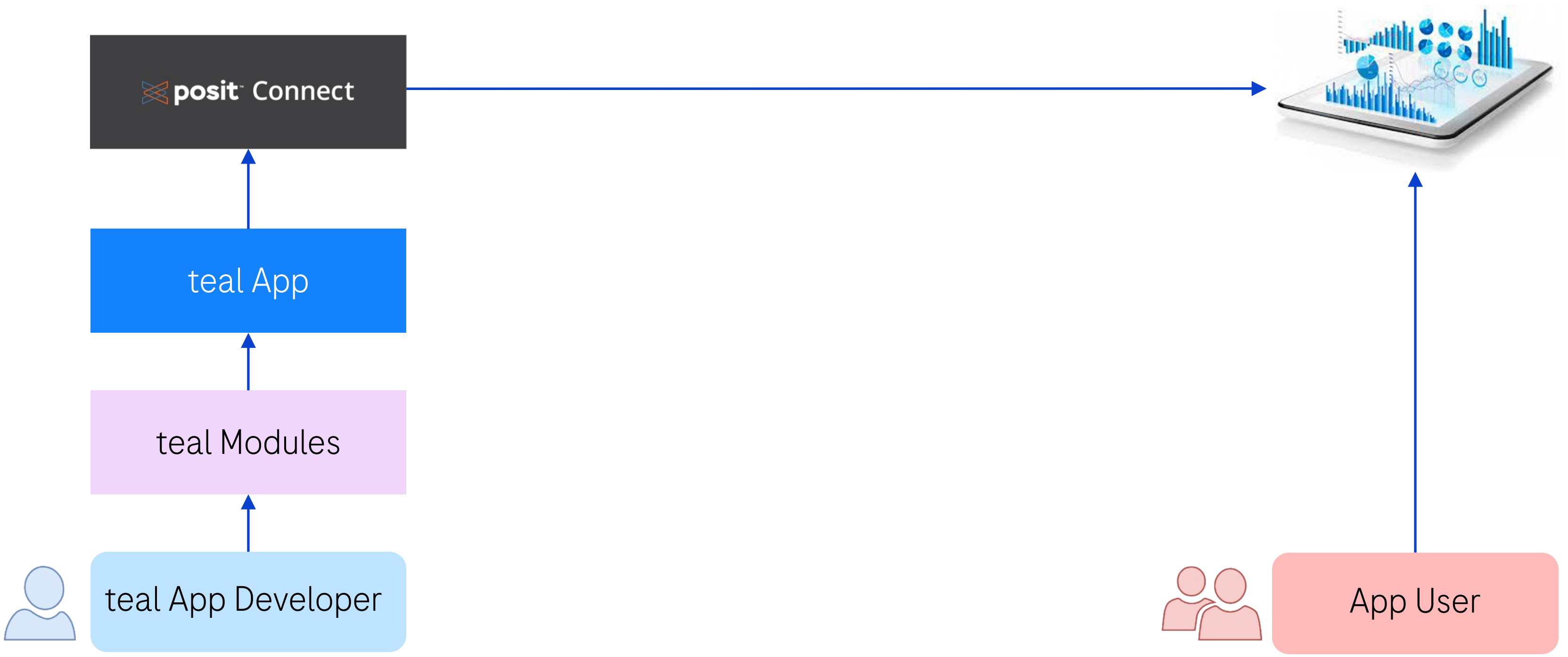
**Data**
Dataset 1
Dataset 2
Dataset 3

**{teal} modules**
Module 1 ui server
Module 2 ui server
Module 3 ui server

**{teal} framework**
ui
server

**{teal} shiny app**

# Using {teal} as a data scientist I



posit Connect

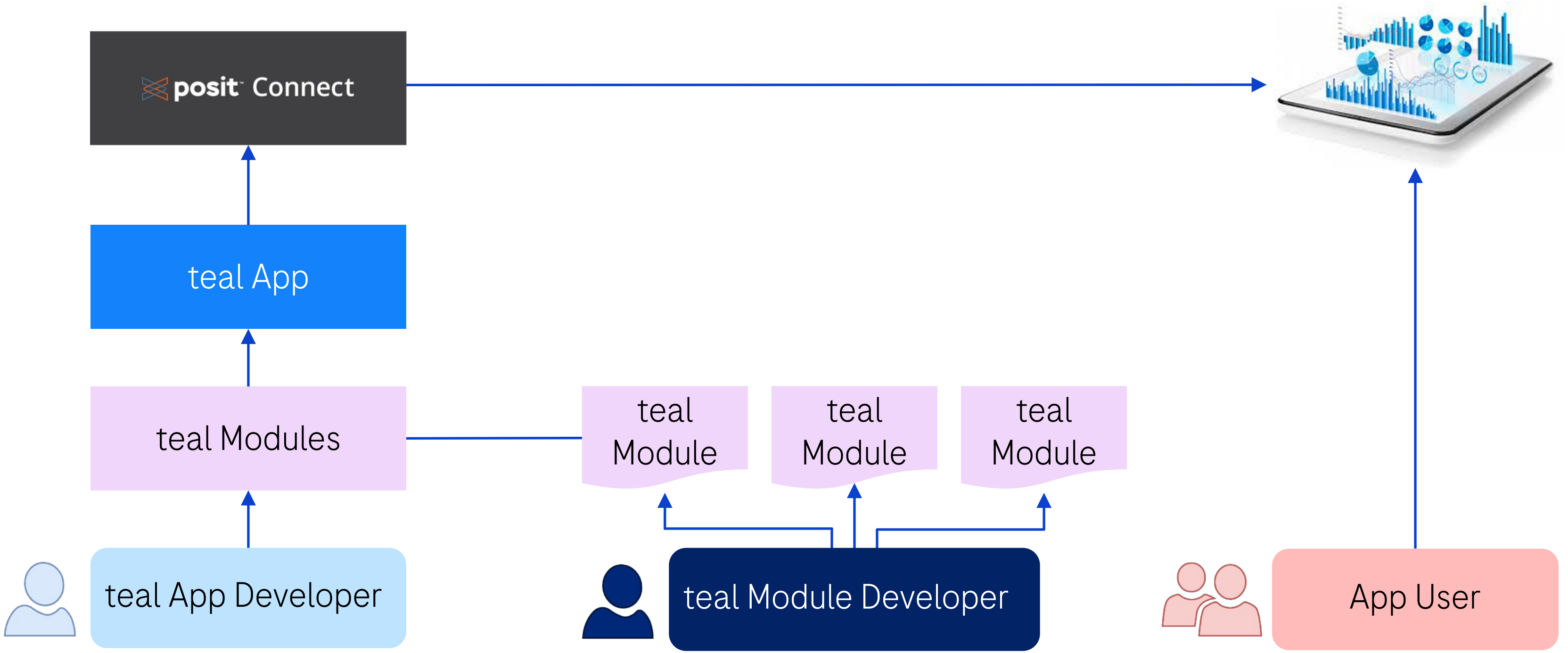
teal App

teal Modules

teal App Developer

App User

# Using {teal} as a data scientist II



posit Connect

teal App

teal Modules

teal Module

teal Module

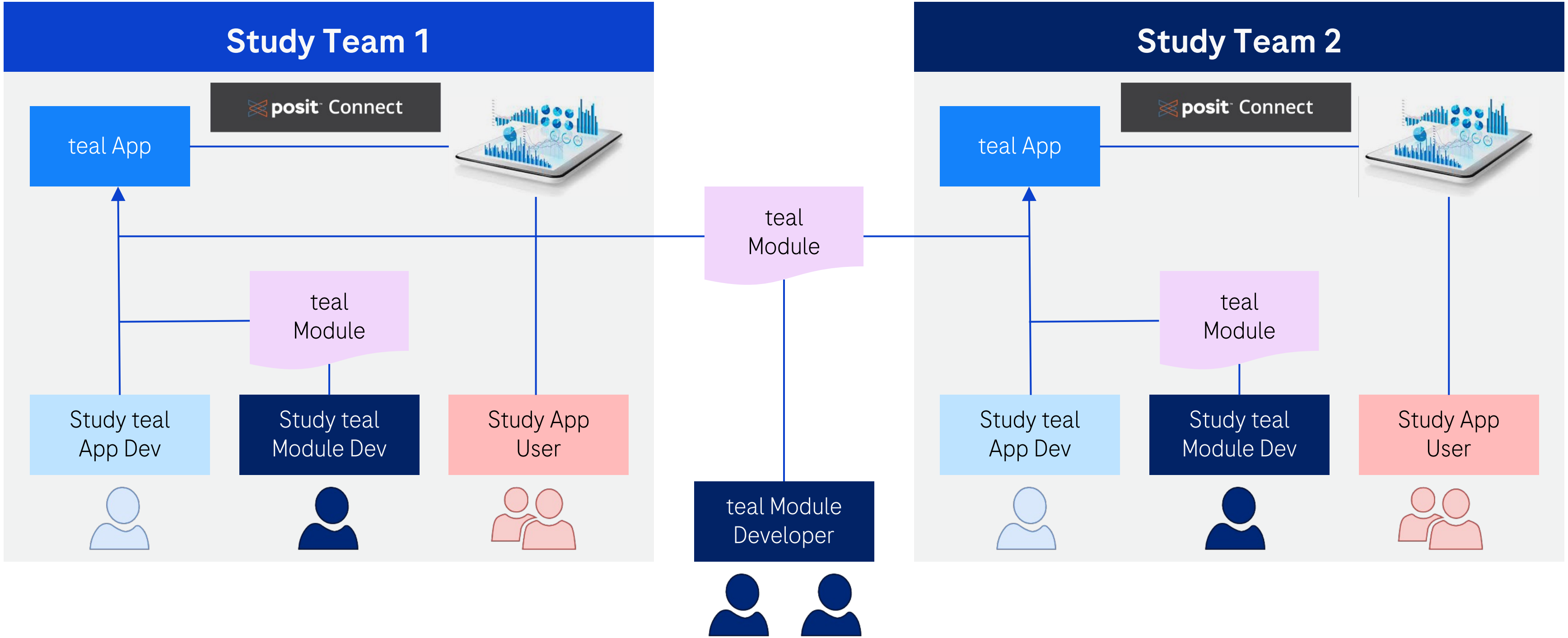
teal Module

teal App Developer

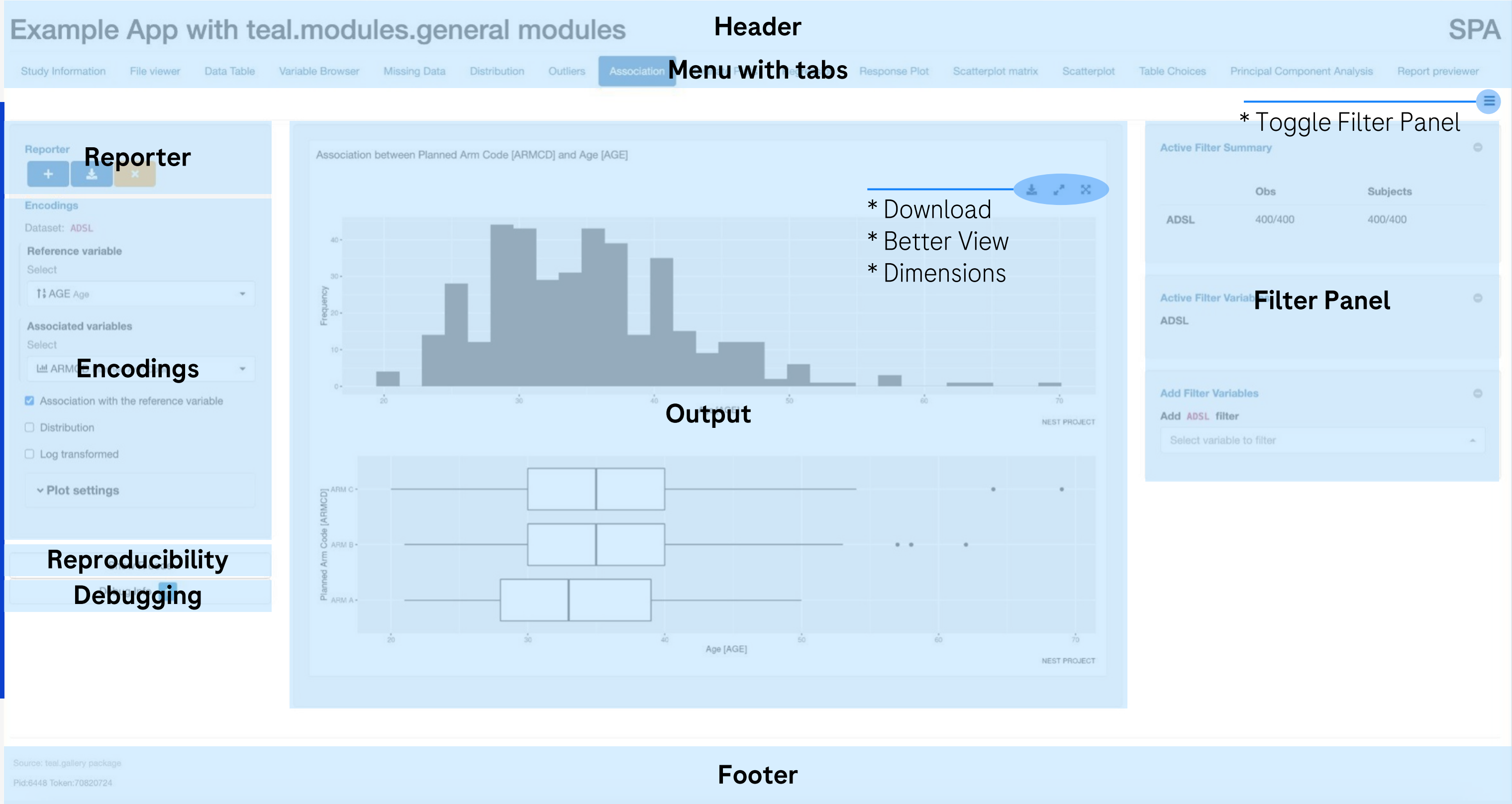
teal Module Developer

App User

# Using {teal} as a data scientist III



# Anatomy of a teal app



Example App with teal.modules.general modules

Dataset: ADSL

Reference variable: AGE Age

Associated variables: ARMCD

☑ Association with the reference variable

☐ Distribution

☐ Log transformed

Association between Planned Arm Code [ARMCD] and Age [AGE]

| | Obs | Subjects |
|---|---|---|
| ADSL | 400/400 | 400/400 |

Source: teal.gallery package

Pid:6448 Token:70820724

**Header**

**Menu with tabs**

* Toggle Filter Panel

**Reporter**

* Download
* Better View
* Dimensions

**Filter Panel**

**Encodings**

**Output**

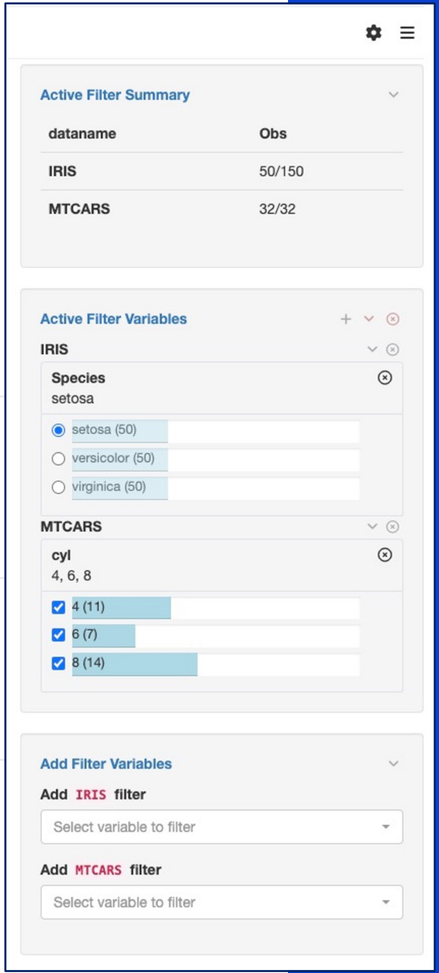
**Reproducibility**

**Debugging**

**Footer**

# Key Features:
# Filter Panel

- Ability to subset your dataset
- Out of the box from teal framework
- Active filter summary
- Pre-defined filter state



**Active Filter Summary**

| dataname | Obs |
|---|---|
| IRIS | 50/150 |
| MTCARS | 32/32 |

**Active Filter Variables**

**IRIS**

**Species**
setosa

● setosa (50)
○ versicolor (50)
○ virginica (50)

**MTCARS**

**cyl**
4, 6, 8

☑ 4 (11)
☑ 6 (7)
☑ 8 (14)

**Add Filter Variables**

**Add `IRIS` filter**

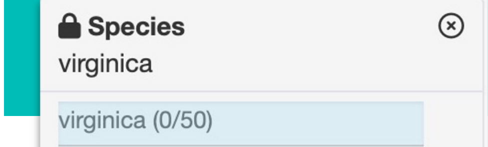
Select variable to filter

**Add `MTCARS` filter**

Select variable to filter

# New Filter Panel Features!
# Introducing filter state behavior

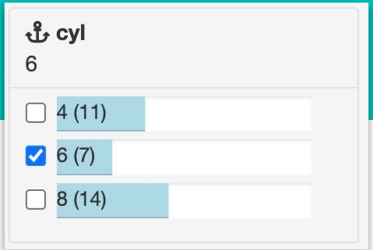
**`fixed selection`**


🔒 **Species**
virginica

virginica (0/50)

```
teal_slice(
  dataname = "iris",
  varname = "Species",
  id = "filter1",
  selected = "virginica",
  fixed = TRUE
)
```

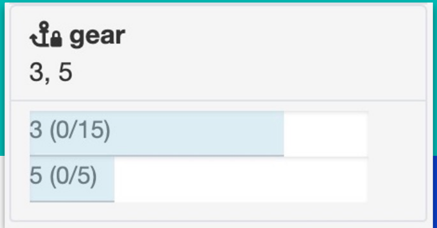
**`anchored filter`**


⚓ **cyl**
6

☐ 4 (11)
☑ 6 (7)
☐ 8 (14)

```
teal_slice(
  dataname = "mtcar",
  varname = "cyl",
  id = "filter2",
  selected = 6,
  anchored = TRUE
)
```

**`fixed AND anchored`**

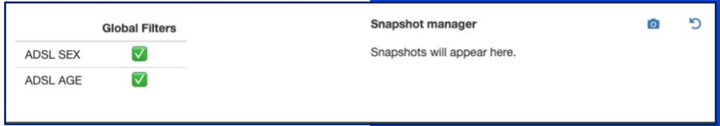

⚓🔒 **gear**
3, 5

3 (0/15)
5 (0/5)

```
teal_slice(
  dataname = "mtcar",
  varname = "gear",
  id = "filter3",
  selected = c(3,5),
  anchored = TRUE,
  fixed = TRUE
)
```

# New Filter Panel Features!
# Snapshot Manager

- Restore initial state of filter state
- Take a snapshot of all filter states
- Restore filter states from a snapshot
- Share snapshot with other users (upcoming!)

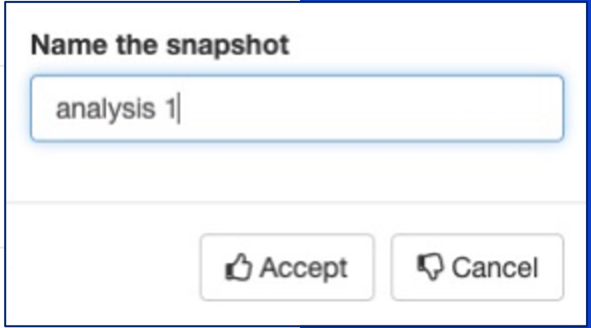


**Global Filters**

ADSL SEX ✓
ADSL AGE ✓

**Snapshot manager**

Snapshots will appear here.



**Name the snapshot**

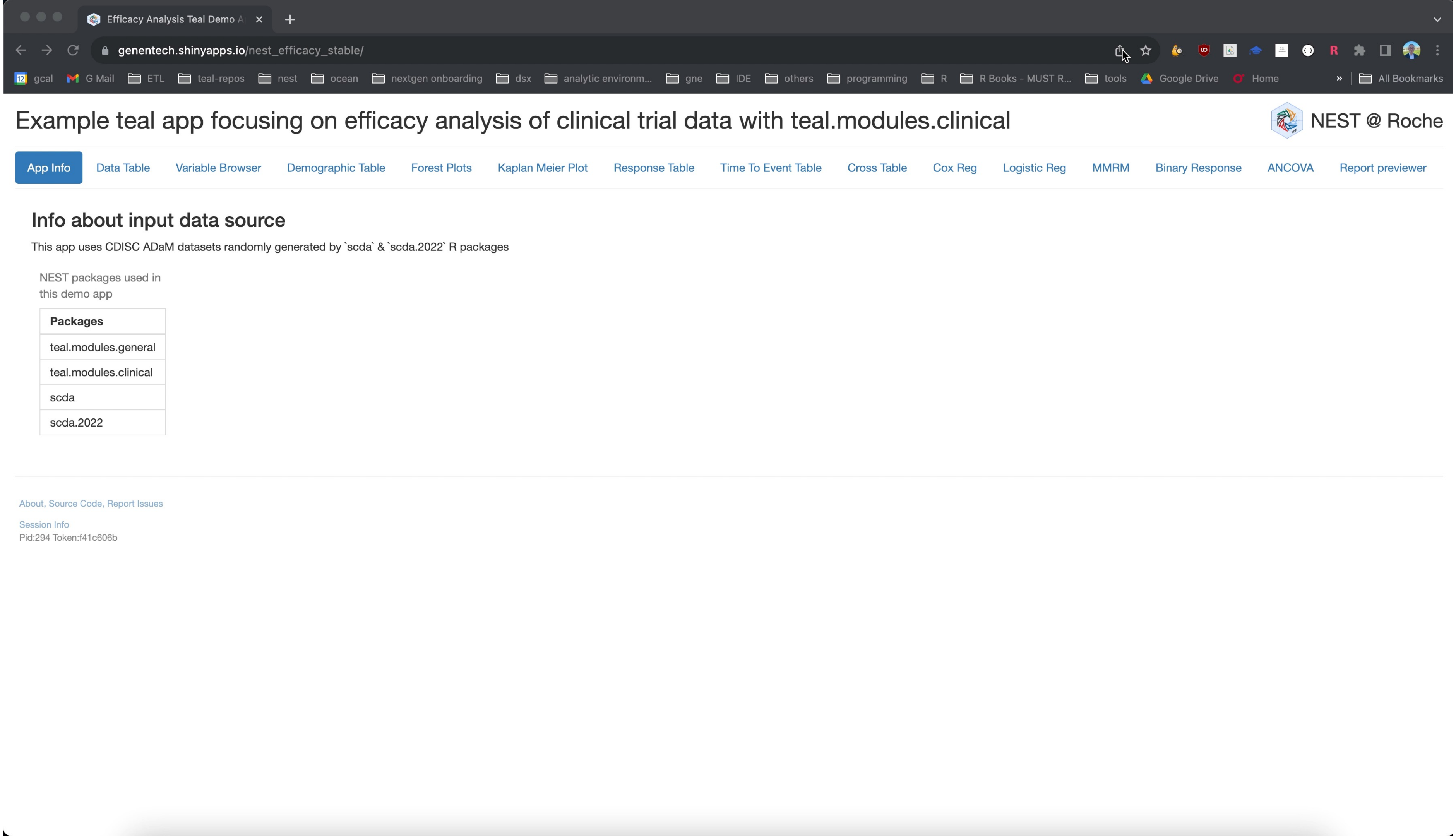
analysis 1

**Global Filters**

ADSL SEX ✓
ADSL AGE ✓

**Snapshot manager**

analysis 1



# Example teal app focusing on efficacy analysis of clinical trial data with teal.modules.clinical

## Info about input data source

This app uses CDISC ADaM datasets randomly generated by `scda` & `scda.2022` R packages

NEST packages used in this demo app

| Packages |
|---|
| teal.modules.general |
| teal.modules.clinical |
| scda |
| scda.2022 |

About, Source Code, Report Issues

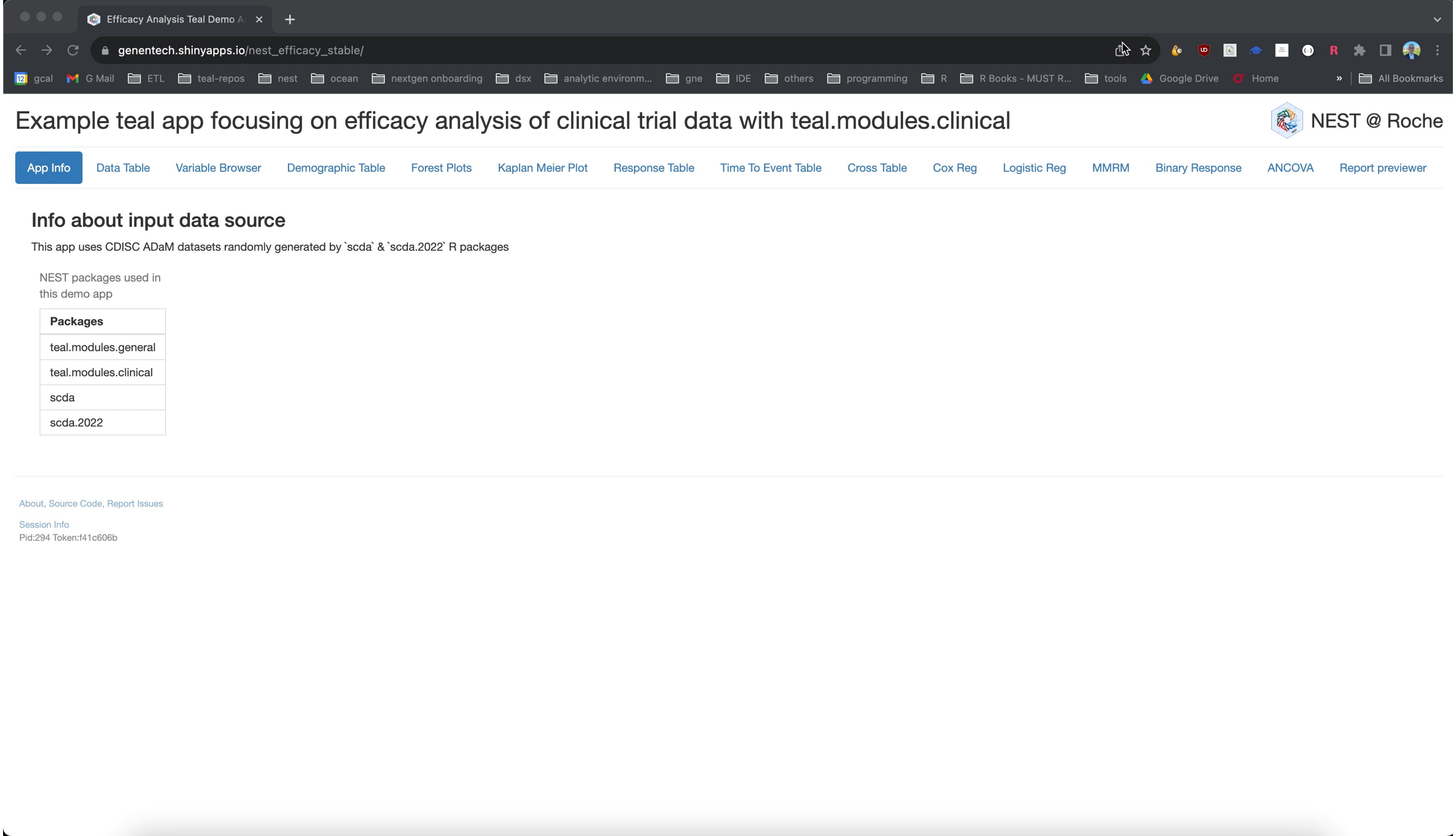
Session Info

Pid:294 Token:f41c606b

# Key Features:
# Show R Code

- Reproducible code when you see the desired output in analysis
- Include any filters added
- Include library ( ) calls

```
## and might have omitted this step for some reason. Please reach
## out to the app developer for details.

# ADSL MD5 hash at the time of analysis: 01965f31841821489767446baaff0ad2
# ADTTE MD5 hash at the time of analysis: af41fb660390618275fbcba67c82a563


ADTTE <- dplyr::inner_join(x = ADTTE, y = ADSL[, c("STUDYID", "USUBJID"), drop = FALSE], by = c("STUDYID", "US



ANL_1 <- ADTTE %>% dplyr::select(STUDYID, USUBJID, PARAMCD, AVAL, CNSR, AVALU)
ANL_2 <- ADSL %>% dplyr::select(STUDYID, USUBJID, ARM, STRATA1, AGEGR1)
ANL_3 <- ADTTE %>%
  dplyr::filter(PARAMCD == "OS") %>%
  dplyr::select(STUDYID, USUBJID, PARAMCD)
ANL <- ANL_1
ANL <- dplyr::inner_join(ANL, ANL_2, by = c("STUDYID", "USUBJID"))
ANL <- dplyr::inner_join(ANL, ANL_3, by = c("STUDYID", "USUBJID", "PARAMCD"))
ANL <- ANL %>% formatters::var_relabel(AVAL = "Analysis Value", CNSR = "Censor", ARM = "Description of Planne

anl <- ANL %>%
  dplyr::filter(ARM %in% c("A: Drug X", "B: Placebo", "C: Combination")) %>%
  dplyr::mutate(ARM = stats::relevel(ARM, ref = "A: Drug X")) %>%
  dplyr::mutate(ARM = droplevels(ARM)) %>%
  dplyr::mutate(is_event = CNSR == 0)
variables <- list(tte = "AVAL", is_event = "is_event", arm = "ARM", strat = "STRATA1")
grid::grid.newpage()
lyt <- grid::grid.layout(nrow = nlevels(ANL$AGEGR1), ncol = 1) %>%
  grid::viewport(layout = .) %>%
  grid::pushViewport()
result <- mapply(df = split(anl, f = anl$AGEGR1), nrow = seq_along(levels(anl$AGEGR1)), FUN = function(df_i, 
  if (nrow(df_i) == 0) {
    grid::grid.text("No data found for a given facet value.", x = 0.5, y = 0.5, vp = grid::viewport(layout.po
  } else {
    g_km(df = df_i, variables = variables, font_size = 8L, xlab = paste0("Time", " (", gsub("(^|[[:space:]])(
      conf_level = 0.95,
      pval_method = "log-rank", ties = "exact"
    ), ci_ribbon = FALSE, vp = grid::viewport(layout.pos.row = nrow_i, layout.pos.col = 1), draw = TRUE)
  }
}, SIMPLIFY = FALSE)
km_grobs <- tern::stack_grobs(grobs = result)
km_grobs
```



# Example teal app focusing on efficacy analysis of clinical trial data with teal.modules.clinical

## Info about input data source

This app uses CDISC ADaM datasets randomly generated by `scda` & `scda.2022` R packages

NEST packages used in this demo app

| Packages |
|---|
| teal.modules.general |
| teal.modules.clinical |
| scda |
| scda.2022 |

About, Source Code, Report Issues

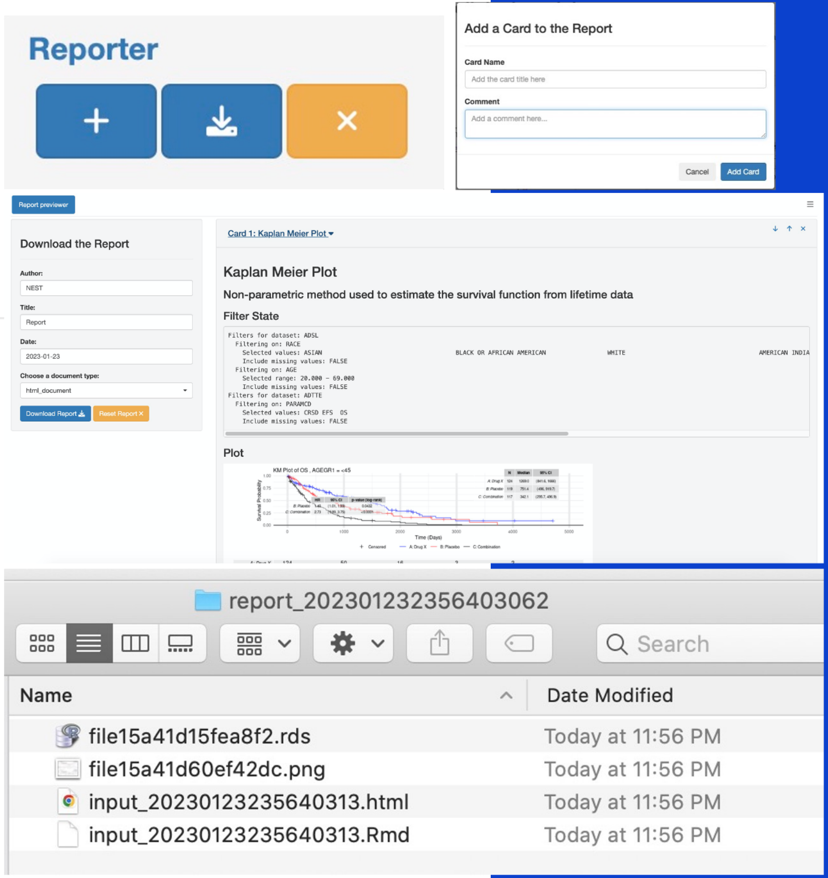
Session Info
Pid:294 Token:f41c606b

# Key Features:
# teal Reporter

Create a report for your analysis

Zip file that contains:

- .Rmd
- .rds
- .png
- And the requested report type (html, pdf, or ppt)



Add a Card to the Report

**Card Name**

Add the card title here

**Comment**

Add a comment here...

Download the Report

**Author:** NEST

**Title:** Report

**Date:** 2023-01-23

**Choose a document type:** html_document

Card 1: Kaplan Meier Plot ▾

Kaplan Meier Plot

Non-parametric method used to estimate the survival function from lifetime data

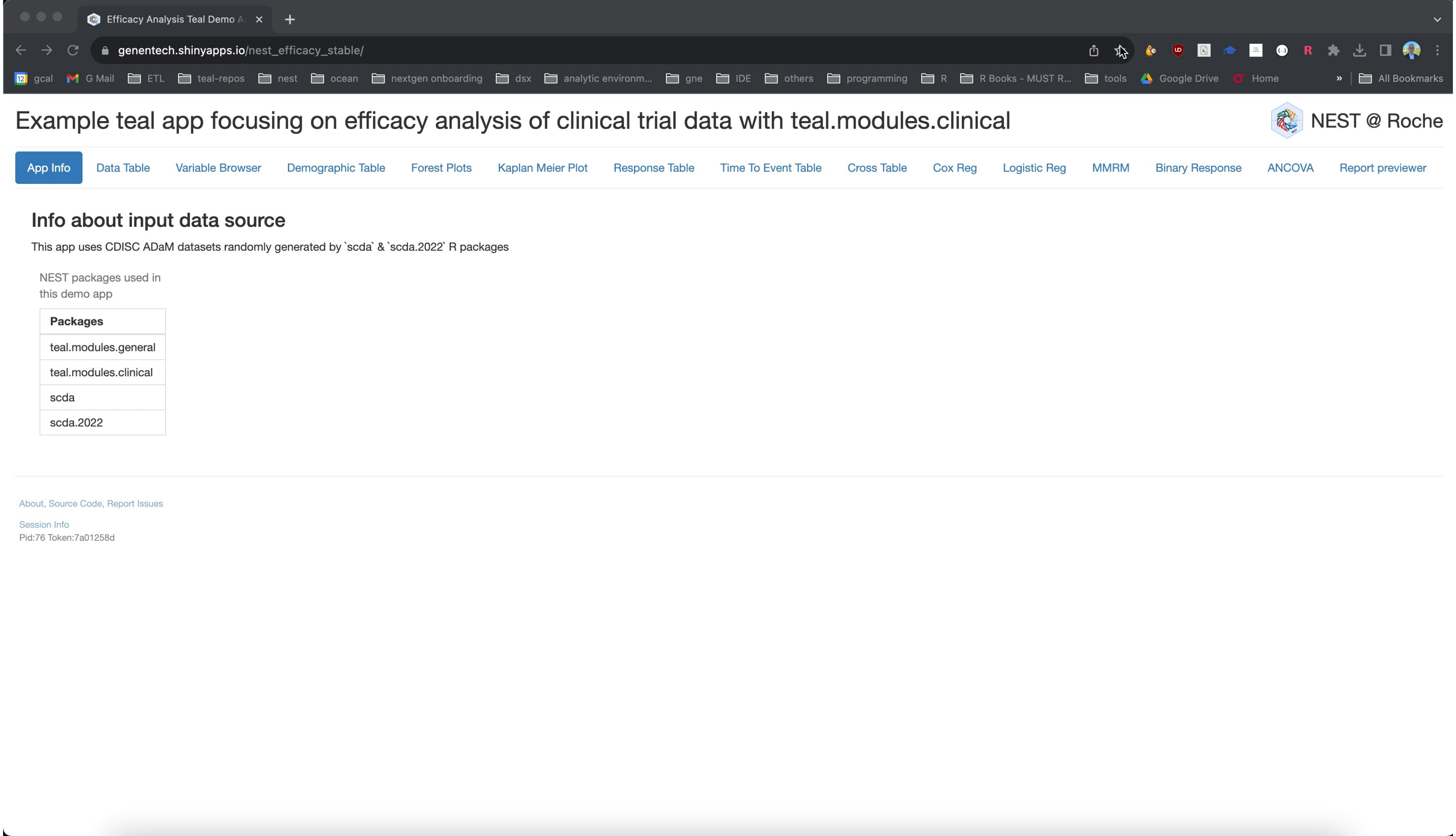
Filter State

```
Filters for dataset: ADSL
  Filtering on: RACE
    Selected values: ASIAN                         BLACK OR AFRICAN AMERICAN         WHITE                     AMERICAN INDIA
    Include missing values: FALSE
  Filtering on: AGE
    Selected range: 20.000 - 69.000
    Include missing values: FALSE
Filters for dataset: ADTTE
  Filtering on: PARAMCD
    Selected values: CRSD EFS  OS
    Include missing values: FALSE
```

Plot

report_202301232356403062

| Name | Date Modified |
|---|---|
| file15a41d15fea8f2.rds | Today at 11:56 PM |
| file15a41d60ef42dc.png | Today at 11:56 PM |
| input_20230123235640313.html | Today at 11:56 PM |
| input_20230123235640313.Rmd | Today at 11:56 PM |



# Example teal app focusing on efficacy analysis of clinical trial data with teal.modules.clinical

## Info about input data source

This app uses CDISC ADaM datasets randomly generated by `scda` & `scda.2022` R packages

NEST packages used in this demo app

| Packages |
|---|
| teal.modules.general |
| teal.modules.clinical |
| scda |
| scda.2022 |

About, Source Code, Report Issues

Session Info

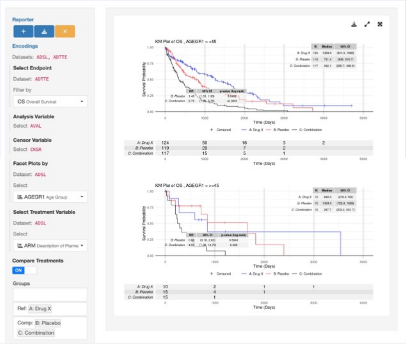
Pid:76 Token:7a01258d

# Key Features:
# Curated sets of teal modules for targeted purposes

Actively developed and maintained centrally by a dedicated team

Purposefully designed, e.g. for clinical trials reporting or biomarker analysis

50+ common analysis modules available for use



Reporter

Encodings

Datasets: ADSL, ADTTE

**Select Endpoint**

Dataset: ADTTE

Filter by: OS Overall Survival

**Analysis Variable**

Select AVAL

**Censor Variable**

Select CNSR

**Facet Plots by**

Dataset: ADSL

Select: AGEGR1 Age Group

**Select Treatment Variable**

Dataset: ADSL

Select: ARM Description of Planne

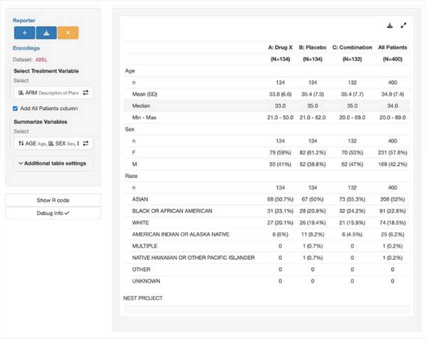
**Compare Treatments**

ON

Groups

Ref: A: Drug X

Comp: B: Placebo<br>C: Combination



Reporter

Encodings

Dataset: ADSL

**Select Treatment Variable**

Select: ARM Description of Plan

☑ Add All Patients column

**Summarize Variables**

Select: AGE Age, SEX Sex,

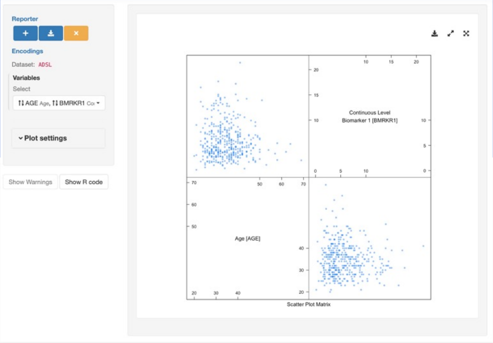
Additional table settings

Show R code

Debug Info ✓

| | A: Drug X<br>(N=134) | B: Placebo<br>(N=134) | C: Combination<br>(N=132) | All Patients<br>(N=400) |
|---|---|---|---|---|
| Age | | | | |
| n | 134 | 134 | 132 | 400 |
| Mean (SD) | 33.8 (6.6) | 35.4 (7.9) | 35.4 (7.7) | 34.9 (7.4) |
| Median | 33.0 | 35.0 | 35.0 | 34.0 |
| Min - Max | 21.0 - 50.0 | 21.0 - 62.0 | 20.0 - 69.0 | 20.0 - 69.0 |
| Sex | | | | |
| n | 134 | 134 | 132 | 400 |
| F | 79 (59%) | 82 (61.2%) | 70 (53%) | 231 (57.8%) |
| M | 55 (41%) | 52 (38.8%) | 62 (47%) | 169 (42.2%) |
| Race | | | | |
| n | 134 | 134 | 132 | 400 |
| ASIAN | 68 (50.7%) | 67 (50%) | 73 (55.3%) | 208 (52%) |
| BLACK OR AFRICAN AMERICAN | 31 (23.1%) | 28 (20.9%) | 32 (24.2%) | 91 (22.8%) |
| WHITE | 27 (20.1%) | 26 (19.4%) | 21 (15.9%) | 74 (18.5%) |
| AMERICAN INDIAN OR ALASKA NATIVE | 8 (6%) | 11 (8.2%) | 6 (4.5%) | 25 (6.2%) |
| MULTIPLE | 0 | 1 (0.7%) | 0 | 1 (0.2%) |
| NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER | 0 | 1 (0.7%) | 0 | 1 (0.2%) |
| OTHER | 0 | 0 | 0 | 0 |
| UNKNOWN | 0 | 0 | 0 | 0 |

NEST PROJECT



Reporter

Encodings

Dataset: ADSL

**Variables**

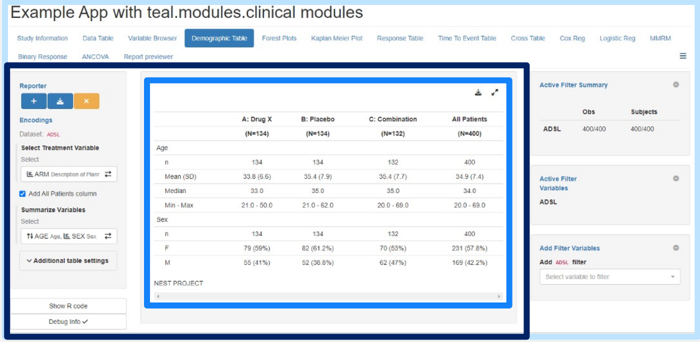
Select: AGE Age, BMRKR1 Co

Plot settings

Show Warnings

Show R code

# {teal} Universe Products Map



Example App with teal.modules.clinical modules

Dataset: ADSL

Select Treatment Variable
Select: ARM Description of Planr

☑ Add All Patients column

Summarize Variables
Select: AGE Age, SEX Sex

| | A: Drug X<br>(N=134) | B: Placebo<br>(N=134) | C: Combination<br>(N=132) | All Patients<br>(N=400) |
|---|---|---|---|---|
| Age | | | | |
| n | 134 | 134 | 132 | 400 |
| Mean (SD) | 33.8 (6.6) | 35.4 (7.9) | 35.4 (7.7) | 34.9 (7.4) |
| Median | 33.0 | 35.0 | 35.0 | 34.0 |
| Min - Max | 21.0 - 50.0 | 21.0 - 62.0 | 20.0 - 69.0 | 20.0 - 69.0 |
| Sex | | | | |
| n | 134 | 134 | 132 | 400 |
| F | 79 (59%) | 82 (61.2%) | 70 (53%) | 231 (57.8%) |
| M | 55 (41%) | 52 (38.8%) | 62 (47%) | 169 (42.2%) |

NEST PROJECT

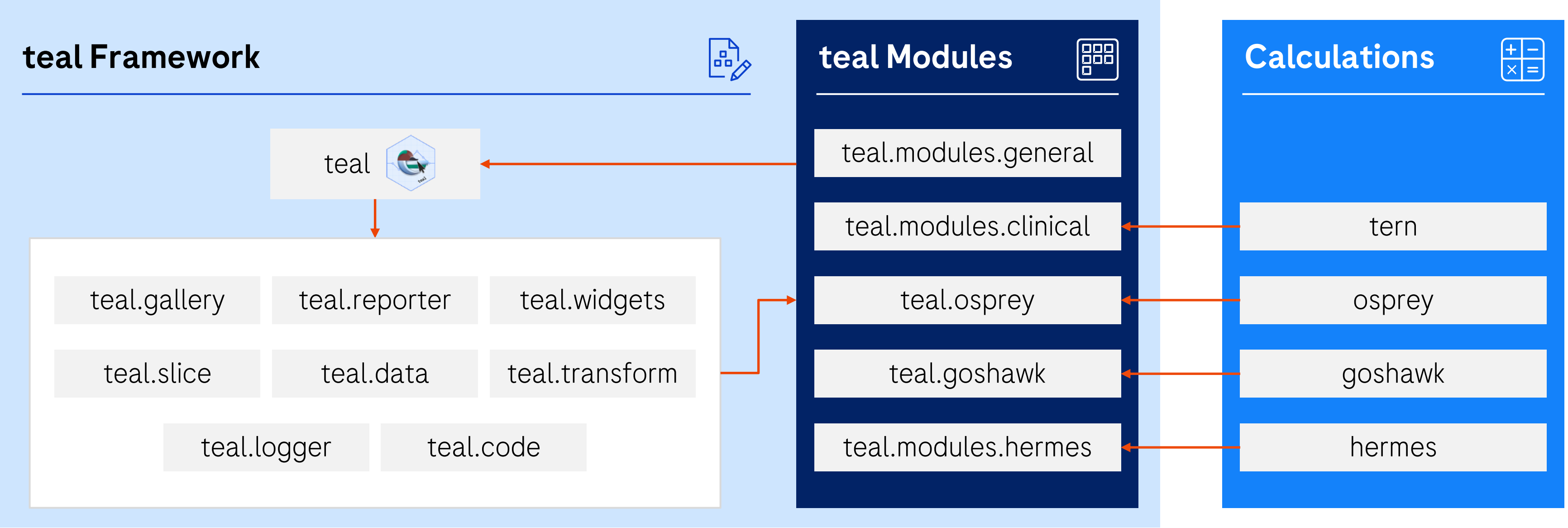
Active Filter Summary

| | Obs | Subjects |
|---|---|---|
| ADSL | 400/400 | 400/400 |

Active Filter Variables
ADSL

Add Filter Variables
Add ADSL filter
Select variable to filter

Roche



## teal Framework

teal

- teal.gallery
- teal.reporter
- teal.widgets
- teal.slice
- teal.data
- teal.transform
- teal.logger
- teal.code

## teal Modules

- teal.modules.general
- teal.modules.clinical
- teal.osprey
- teal.goshawk
- teal.modules.hermes

## Calculations

- tern
- osprey
- goshawk
- hermes

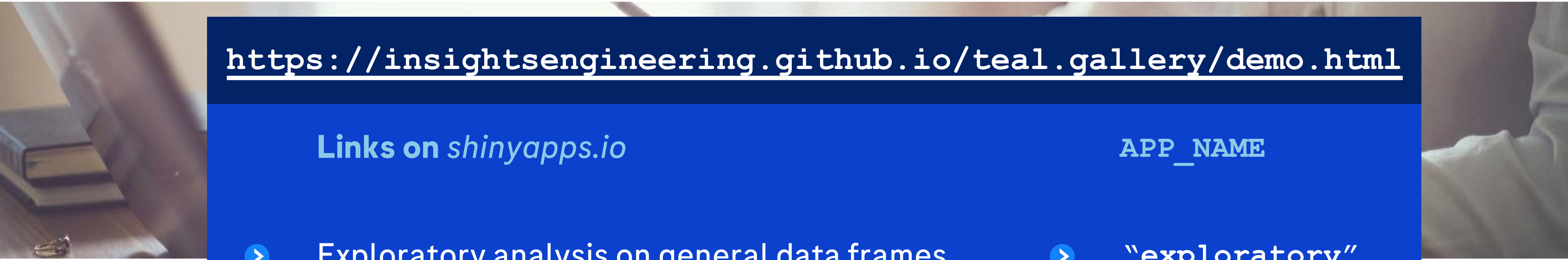
## teal Framework R Packages

- teal: shiny-based interactive exploration framework for analyzing data.
- teal.gallery: gallery of sample teal apps.
- teal.widgets: shiny components used within teal.
- teal.reporter: allows teal applications to generate reports.
- teal.slice: provides a filtering panel to allow subset of data.
- teal.data: creating and loading the data needed for teal applications.
- teal.code: handles reproducibility of outputs.
- teal.transform: standardizes extracting and merging data
- teal.logger: standardizes logging within teal framework.

## teal Modules R Packages

- teal.modules.general: general analysis modules for exploring any data types
- teal.modules.clinical: modules for analyzing CDISC data and clinical trial reporting with tern R package
- teal.osprey: modules for analyzing and reporting early-phase clinical trial data with osprey R package
- teal.goshawk: modules for analyzing and visualizing biomarker data with goshawk R package
- teal.modules.hermes: modules for analyzing and visualizing RNAseq data with hermes R package

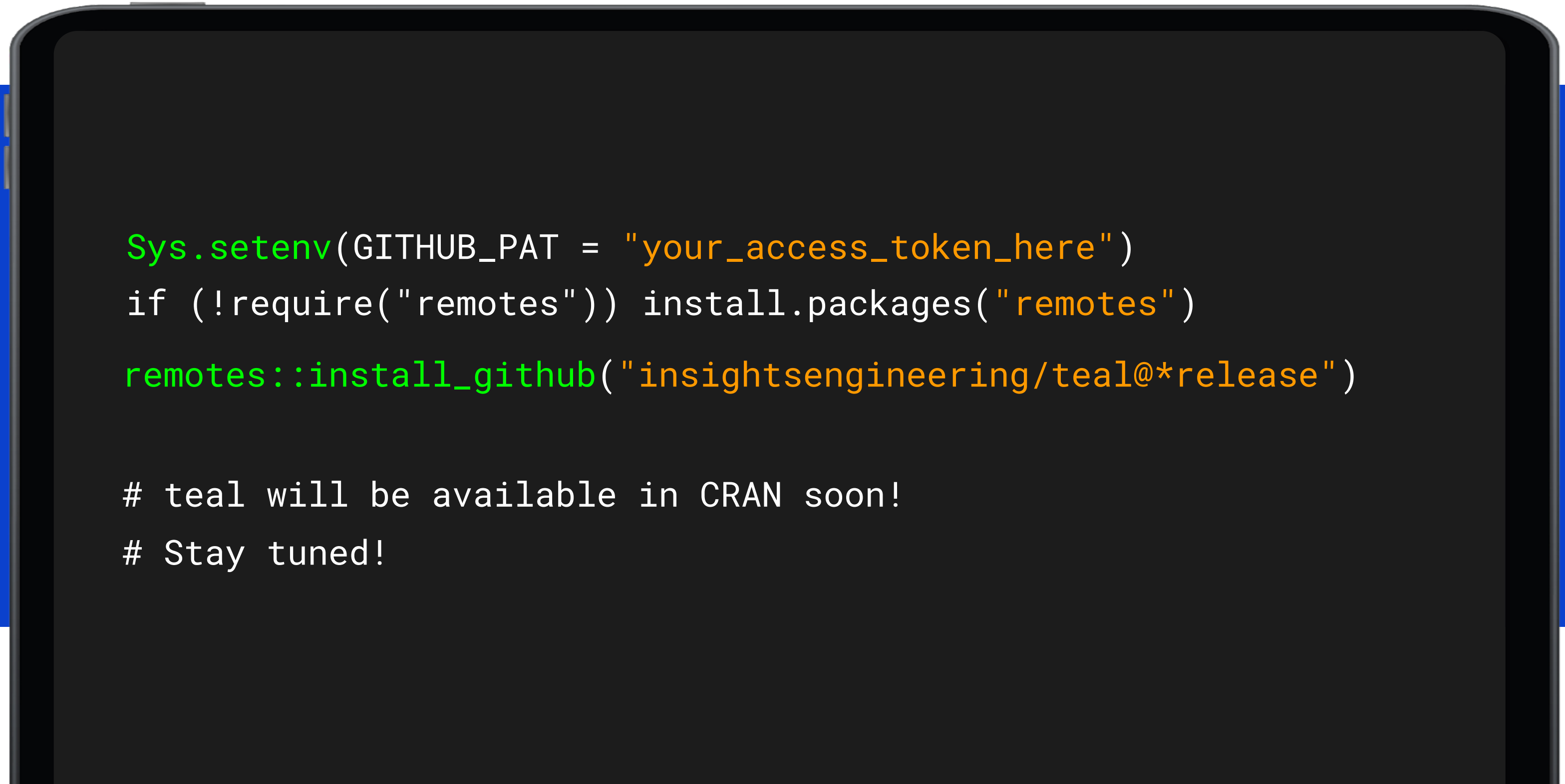
# Demo Apps via {teal} Gallery



`https://insightsengineering.github.io/teal.gallery/demo.html`

| **Links on** ***shinyapps.io*** | `APP_NAME` |
|---|---|
| Exploratory analysis on general data frames | `"exploratory"` |
| Safety analysis on clinical trial data | `"safety"` |
| Efficacy analysis on clinical trial data | `"efficacy"` |

# Installation



```r
Sys.setenv(GITHUB_PAT = "your_access_token_here")
if (!require("remotes")) install.packages("remotes")
remotes::install_github("insightsengineering/teal@*release")

# teal will be available in CRAN soon!
# Stay tuned!
```

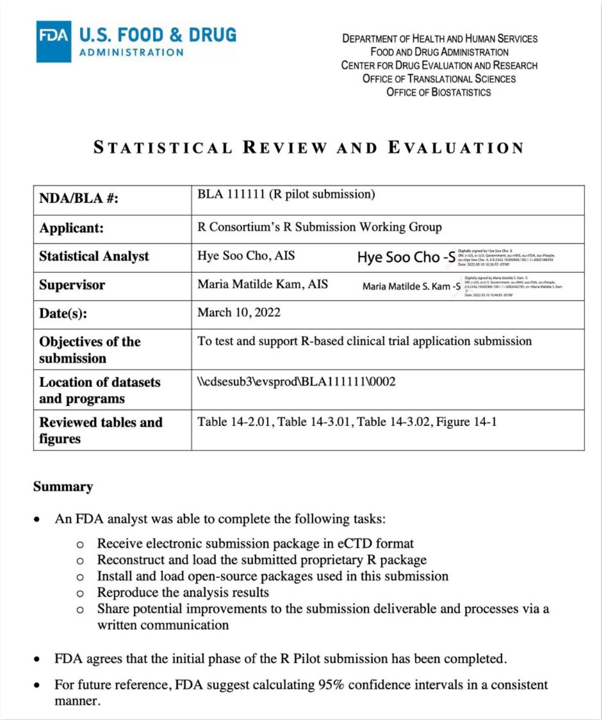
# Technology Shift across Industry and HA

**R submissions working group**: Streamline R based FDA submission - Collaboration by ~20 industry companies and FDA.

- Great engagement from FDA to enable R based submissions
    - FDA provided formal response letter in public (link)
- Sponsored by R Consortium

**Pharmaverse**: Co-create R based tools for clinical trial reporting - collaboration being driven by 5 industry companies, with a growing community of >1,000 people from across ~50 companies.

- Similar effort in co-creating tools for advanced statistical methodologies: American Statistical Association Biopharmaceutical Section Software Engineering Working Group



FDA U.S. FOOD & DRUG ADMINISTRATION

DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF TRANSLATIONAL SCIENCES
OFFICE OF BIOSTATISTICS

STATISTICAL REVIEW AND EVALUATION

| | |
|---|---|
| **NDA/BLA #:** | BLA 111111 (R pilot submission) |
| **Applicant:** | R Consortium's R Submission Working Group |
| **Statistical Analyst** | Hye Soo Cho, AIS Hye Soo Cho -S |
| **Supervisor** | Maria Matilde Kam, AIS Maria Matilde S. Kam -S |
| **Date(s):** | March 10, 2022 |
| **Objectives of the submission** | To test and support R-based clinical trial application submission |
| **Location of datasets and programs** | \\cdsesub3\evsprod\BLA111111\0002 |
| **Reviewed tables and figures** | Table 14-2.01, Table 14-3.01, Table 14-3.02, Figure 14-1 |

**Summary**

- An FDA analyst was able to complete the following tasks:
    - Receive electronic submission package in eCTD format
    - Reconstruct and load the submitted proprietary R package
    - Install and load open-source packages used in this submission
    - Reproduce the analysis results
    - Share potential improvements to the submission deliverable and processes via a written communication
- FDA agrees that the initial phase of the R Pilot submission has been completed.
- For future reference, FDA suggest calculating 95% confidence intervals in a consistent manner.

# R submission Pilots to FDA - *By the R consortium R submission Working Group*

Presenter(s): Ning Leng, Roche-Genentech , Eric Nantz, Eli Lilly, on behalf of the R submission WG

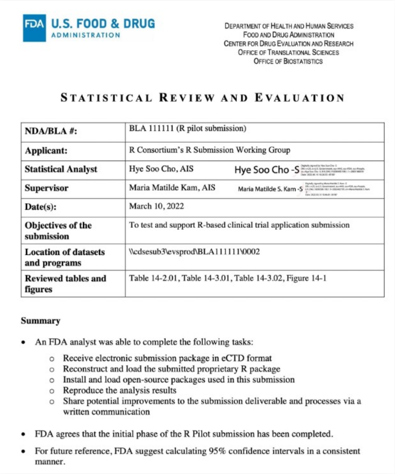
## R Submission Working Group:

A cross industry collaboration to improve open-source language usage in the regulatory setting

| | **Pilot 1 (finished)** | **Pilot 2 (submitted)** | **Pilot 3 (submitted)** | **Pilot 4** |
|---|---|---|---|---|
| **Main Purpose** | TLF | Shiny | ADaM | Shiny in Container |
| **Operational feasibilities tested** | | | | |
| Submission of Analytical codes | .r files | | .r files | |
| Submission of proprietary R package | pkglite | golem + pkglite | Github repo | |
| Dependency management | mran snapshot | renv | renv | Container, WebAssembly |

## Pilot 1 Submission to FDA:

Successful open pilot that showcased R based FDA submission



U.S. FOOD & DRUG ADMINISTRATION

DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF TRANSLATIONAL SCIENCES
OFFICE OF BIOSTATISTICS

STATISTICAL REVIEW AND EVALUATION

| | |
|---|---|
| **NDA/BLA #:** | BLA 111111 (R pilot submission) |
| **Applicant:** | R Consortium's R Submission Working Group |
| **Statistical Analyst** | Hye Soo Cho, AIS Hye Soo Cho -S |
| **Supervisor** | Maria Matilde Kam, AIS Maria Matilde S. Kam -S |
| **Date(s):** | March 10, 2022 |
| **Objectives of the submission** | To test and support R-based clinical trial application submission |
| **Location of datasets and programs** | \\cdsesub3\evsprod\BLA111111\0002 |
| **Reviewed tables and figures** | Table 14-2.01, Table 14-3.01, Table 14-3.02, Figure 14-1 |

**Summary**

- An FDA analyst was able to complete the following tasks:
  - Receive electronic submission package in eCTD format
  - Reconstruct and load the submitted proprietary R package
  - Install and load open-source packages used in this submission
  - Reproduce the analysis results
  - Share potential improvements to the submission deliverable and processes via a written communication
- FDA agrees that the initial phase of the R Pilot submission has been completed.
- For future reference, FDA suggest calculating 95% confidence intervals in a consistent manner.

## Key Links

- Webpage: https://rconsortium.github.io/submissions-wg/
- Pilot 1: https://github.com/RConsortium/submissions-pilot1-to-fda
- Pilot 2: https://github.com/RConsortium/submissions-pilot2-to-fda
- Pilot 3: https://github.com/RConsortium/submissions-pilot3-adam-to-fda

All materials are open to the public
Anyone is welcome to join the working groups! (info@r-consortium.org)

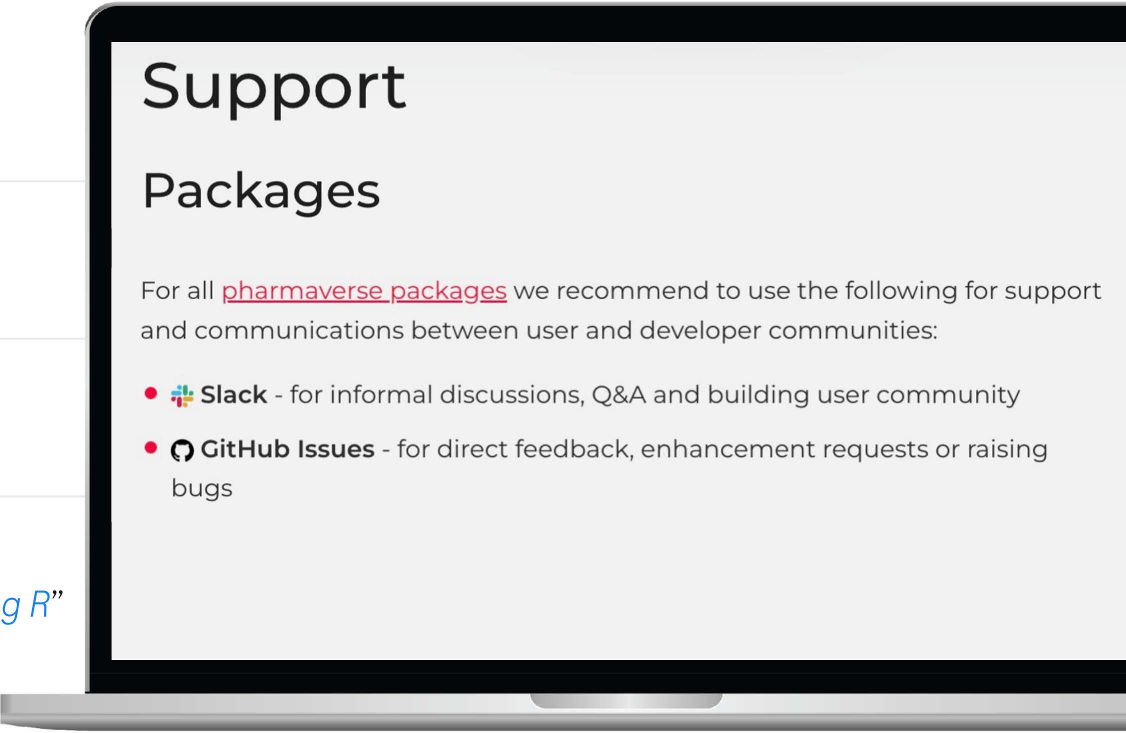
# Additional Resources

{teal} is part of pharmaverse:
https://pharmaverse.org/

More information about support:
https://pharmaverse.org/support/

Slack channel *#teal* under
pharmaverse workspace

Upcoming course on Coursera
"*Hands On Clinical Reporting Using R*"
in Q4 2023



# Support

## Packages

For all pharmaverse packages we recommend to use the following for support and communications between user and developer communities:

- **Slack** - for informal discussions, Q&A and building user community
- **GitHub Issues** - for direct feedback, enhancement requests or raising bugs

# Collaborating on {teal}

We are looking for collaborators to develop this framework further!

---

If you're an individual, please contribute on GitHub and join us via pharmaverse Slack *#teal* channel

---

If you're an organization wanting to adopt {teal}
and co-develop, please get in touch with our Product Owner chendi.liao@roche.com

*Doing now what patients need next*