

# Clinical Tables in Python with Great Tables

`github:philbowsher`

Thank you

Richard Iannone, Michael Chow

# Slides and Examples

- [github.com/philbousher/Clinical-Tables-in-Python-with-Great-Tables](https://github.com/philbousher/Clinical-Tables-in-Python-with-Great-Tables)

# But Why Another Table Package?

- [GT R Package](#)
- [Clinical-Tables-in-R-with-gt](#)
- Open Source Backbone
- Open Source Next-Gen SCEs
- Multilingual Programming
- Language-agnostic
- [Open-Source-in-New-Drug-Applications-NDAs-FDA](#)



Viewing labelled data

| View(flights) |        |             |       | View(flights_labelled) |        |             |       |
|---------------|--------|-------------|-------|------------------------|--------|-------------|-------|
| Flight        | Origin | Destination | Time  | Flight                 | Origin | Destination | Time  |
| 1             | SEA    | SEA         | 12:00 | 1                      | SEA    | SEA         | 12:00 |
| 2             | SEA    | SEA         | 12:00 | 2                      | SEA    | SEA         | 12:00 |
| 3             | SEA    | SEA         | 12:00 | 3                      | SEA    | SEA         | 12:00 |
| 4             | SEA    | SEA         | 12:00 | 4                      | SEA    | SEA         | 12:00 |
| 5             | SEA    | SEA         | 12:00 | 5                      | SEA    | SEA         | 12:00 |
| 6             | SEA    | SEA         | 12:00 | 6                      | SEA    | SEA         | 12:00 |
| 7             | SEA    | SEA         | 12:00 | 7                      | SEA    | SEA         | 12:00 |
| 8             | SEA    | SEA         | 12:00 | 8                      | SEA    | SEA         | 12:00 |
| 9             | SEA    | SEA         | 12:00 | 9                      | SEA    | SEA         | 12:00 |
| 10            | SEA    | SEA         | 12:00 | 10                     | SEA    | SEA         | 12:00 |

# And in Python?

- FDA Briefing Document NDA# 212833"Source: FDA Review staff analysis using Applicant-submitted data adsl.xpt, adlbce.xpt, adlbhb.xpt, adcm.xpt, adexsum.xpt in Python (Ver. 3) "
- 2020 Genentech, Inc. Source: adae.xpt; Software: Python



Yan Holtz

@R\_Graph\_Gallery

There are many [#rstats](#) packages allowing to create perfect tables:

➔ DT, gt, gtExtras, kable, kableExtra, formattable, reactable....

But in [#python](#), nothing or almost 🤔🤔🤔 ?? Am I missing something?

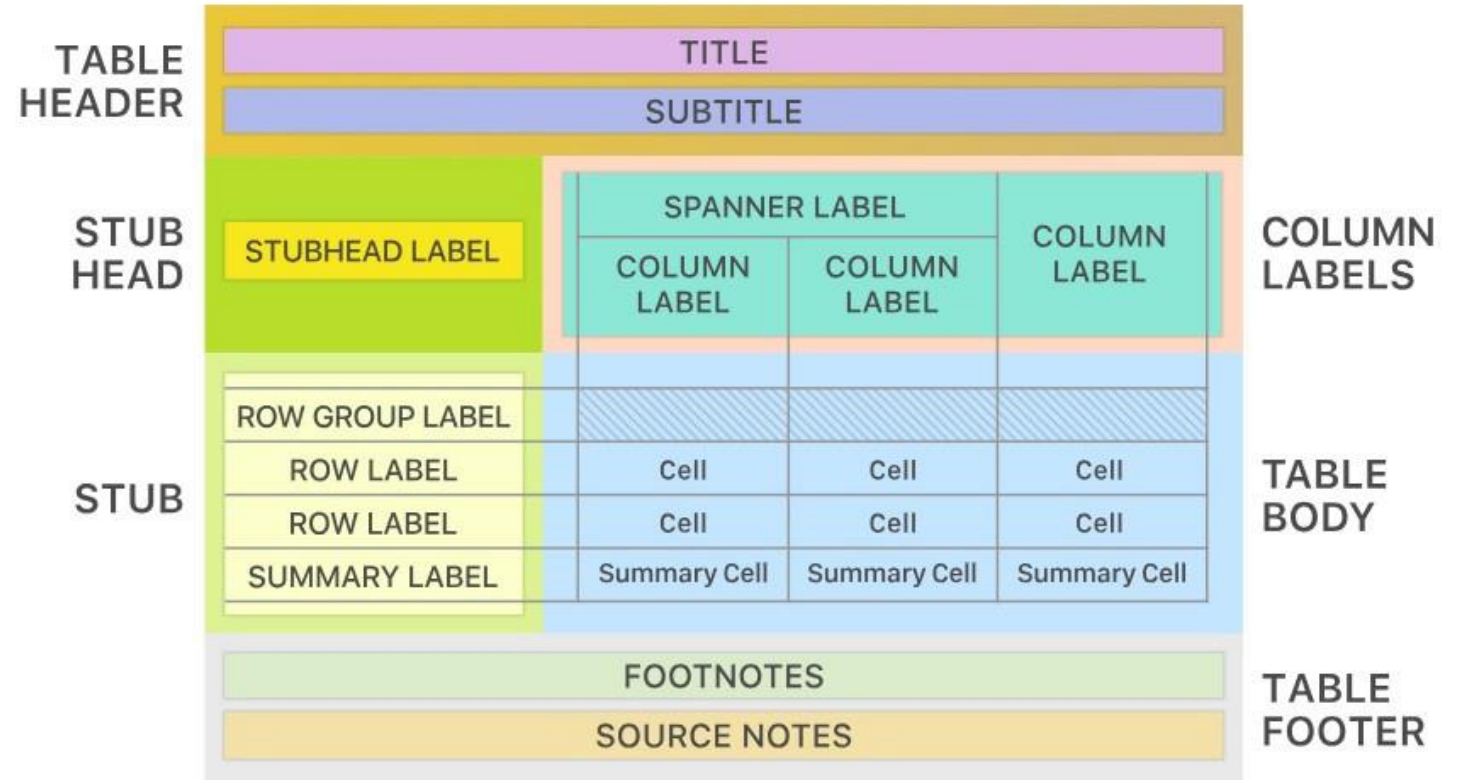
### 2019 Women's World Cup Predictions

Soccer Power Index (SPI) ratings and chances of advancing for every team

| TEAM               | GROUP | Team Rating |      |      | Chance of Finishing Group Stage In ... |           |           | Knockout Stage Chances |                 |                 |            |               |
|--------------------|-------|-------------|------|------|----------------------------------------|-----------|-----------|------------------------|-----------------|-----------------|------------|---------------|
|                    |       | SPI         | OFF. | DEF. | 1ST PLACE                              | 2ND PLACE | 3RD PLACE | MAKE ROUND OF 16       | MAKE QTR-FINALS | MAKE SEMIFINALS | MAKE FINAL | WIN WORLD CUP |
| USA 6 pts.         | F     | 98.3        | 5.5  | 0.6  | 83%                                    | 17%       | —         | ✓                      | 78%             | 47%             | 35%        | 24%           |
| France 6 pts.      | A     | 96.3        | 4.3  | 0.5  | >99%                                   | <1%       | <1%       | ✓                      | 78%             | 42%             | 30%        | 19%           |
| Germany 6 pts.     | B     | 93.8        | 4.8  | 0.7  | 98%                                    | 2%        | —         | ✓                      | 89%             | 48%             | 28%        | 12%           |
| Canada 6 pts.      | E     | 93.5        | 3.7  | 0.6  | 39%                                    | 61%       | —         | ✓                      | 59%             | 36%             | 20%        | 9%            |
| England 6 pts.     | D     | 91.9        | 3.5  | 0.6  | 71%                                    | 29%       | —         | ✓                      | 69%             | 43%             | 16%        | 8%            |
| Netherlands 6 pts. | E     | 92.7        | 3.9  | 0.7  | 61%                                    | 39%       | —         | ✓                      | 59%             | 37%             | 19%        | 8%            |
| Australia 3 pts.   | C     | 92.8        | 4.2  | 0.9  | 13%                                    | 54%       | 34%       | >99%                   | 54%             | 26%             | 10%        | 5%            |

# Design Philosophy

## The Components of a Table



# Some Feedback

 **Hacker News** [new](#) | [past](#) | [comments](#) | [ask](#) | [show](#) | [jobs](#) | [submit](#)

▲ The design philosophy of Great Tables ([posit-dev.github.io](#))  
504 points by [randyzwitch](#) 7 months ago | [hide](#) | [past](#) | [favorite](#) | 92 comments

▲ [paddy\\_m](#) 7 months ago | [next](#) [-]  
Great tables has done some really nice work on python/jupyter tables. It looks like they are almost building a "grammar of tables" similar to a grammar of graphics. More projects should write about their philosophy and aims like this.

▲ [benterix](#) 7 months ago | [prev](#) | [next](#) [-]  
This guy deserves some prize for 1) great work, 2) attention to detail, 3) in-depth research, 4) excellent presentation of his work (sparing the usual questions like "but what is it, really?", "how do I start?", "can you provide some examples?").

▲ [bigger\\_cheese](#) 7 months ago | [prev](#) | [next](#) [-]  
I Use SAS reports in my job pretty heavily I'm keen to find alternatives - this looks pretty promising.

\*Cropped

# Tables with Python

Tables with  
Python...



Tables with  
Python...



# Basic Example

```
import polars as pl
from great_tables import GT, md, html
from great_tables.data import islands
islands_mini = (
    pl.from_pandas(islands).sort("size", descending=True)
    .head(10)
)
gt_table_final = (
    GT(islands_mini)
    .tab_header(
        title="Large Landmasses of the World",
        subtitle="The top ten largest are presented"
    )
    .tab_stub(rowname_col="name")
    .tab_source_note(source_note="Source: The World Almanac and Book of
Fac
    .tab_source_note(
        source_note=md("Reference: McNeil, D. R. (1977) *Interactive Data
A
    )
    .tab_stubhead(label="landmass")
    .fmt_integer(columns="size")
)
gt_table_final
gt_table_final.as_raw_html()
```



# Large Landmasses of the World

The top ten largest are presented

| landmass      | size   |
|---------------|--------|
| Asia          | 16,988 |
| Africa        | 11,506 |
| North America | 9,390  |
| South America | 6,795  |
| Antarctica    | 5,500  |
| Europe        | 3,745  |
| Australia     | 2,968  |
| Greenland     | 840    |
| New Guinea    | 306    |
| Borneo        | 280    |

Source: The World Almanac and Book of Facts, 1975, page 406.

Reference: McNeil, D. R. (1977) *Interactive Data Analysis*. Wiley.

# Tables in Clinical Study Reports

- E3 Reporting
- Title & footnotes are integral parts of (TFGs) in (CSRs) & (CRT)
- Titles & footnotes are created or typed in a requirement document such as (SAP)
- R Submissions
- Many orgs are now creating Standard & Custom TLGs with R (Open Source)

# OpenFDA Example

```
# Create and configure the table using GreatTable
gt_table_final = (
  GT(top10_df).tab_header(
    title="Top 10 Adverse Events for Tylenol by Gender",
    subtitle="Most frequent adverse event reports for male and female
    patients retrieved from OpenFDA"
  ).tab_stub(
    rowname_col="Adverse Event"
  ).tab_stubhead(
    label="Patient Reaction"
  ).tab_source_note(
    source_note=md("Data obtained from the OpenFDA database
    of 1000 data points.")
  )
)
gt_table_final.as_raw_html()
```

# Basic Example

## Top 10 Adverse Events for Tylenol by Gender

Most frequent adverse event reports for male and female patients retrieved from OpenFDA

| Patient Reaction | Female | Male | Unknown | Total |
|------------------|--------|------|---------|-------|
| Headache         | 62     | 13   | 6       | 81    |
| Fatigue          | 49     | 21   | 10      | 80    |
| Pain             | 56     | 9    | 3       | 68    |
| Nausea           | 50     | 10   | 4       | 64    |
| Diarrhoea        | 45     | 17   | 0       | 62    |
| Drug ineffective | 46     | 13   | 2       | 61    |
| Dyspnoea         | 35     | 21   | 3       | 59    |
| Arthralgia       | 40     | 11   | 1       | 52    |
| Dizziness        | 40     | 9    | 3       | 52    |
| Asthenia         | 36     | 7    | 3       | 46    |

Data obtained from the OpenFDA database of 1000 data points.

# elmer-assistant

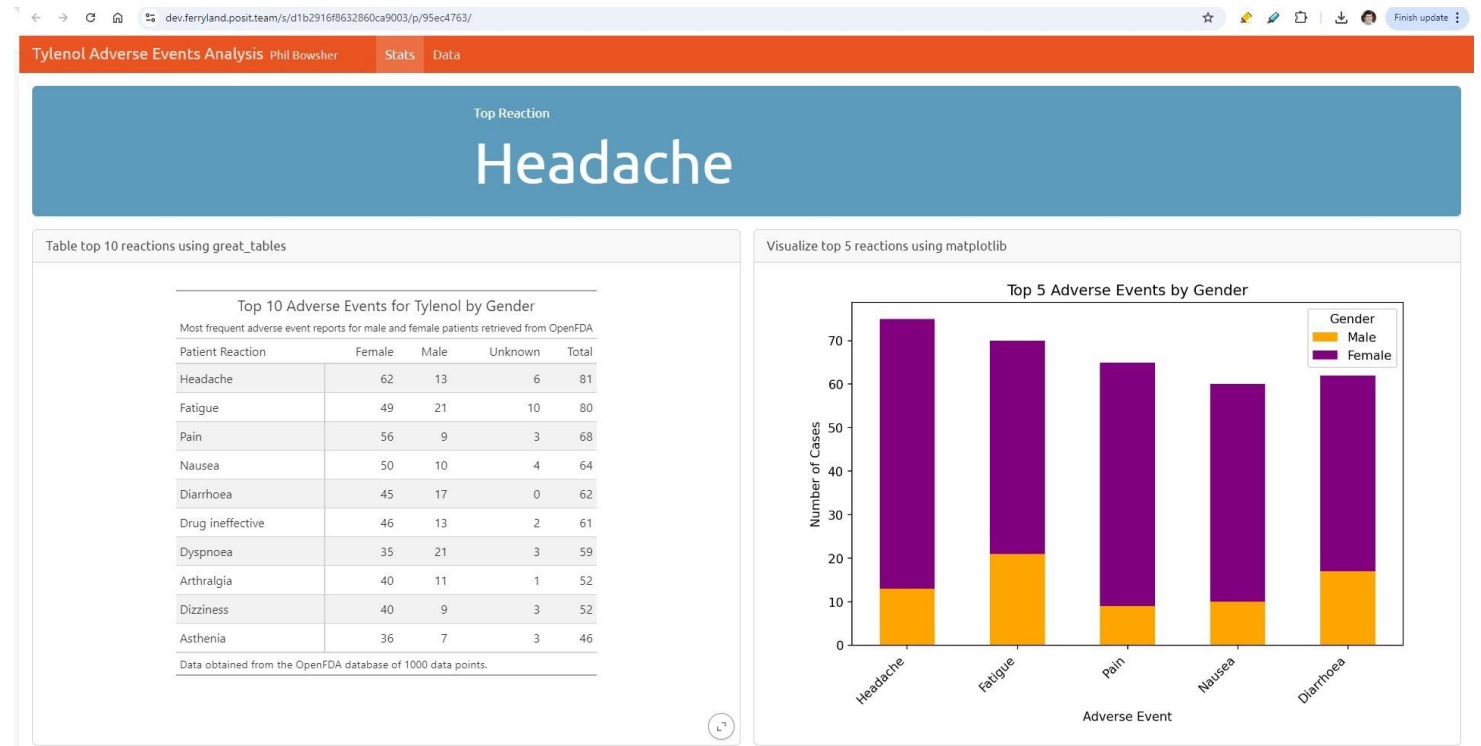


👋 Hi, I'm **Elmer Assistant**! I'm here to answer questions about `elmer` and `shinychat`, or to generate code for you.

Enter a message...



# Moderate Complexity Quarto Dashboard



# Great Tables Data Prep

- [Pandas](#)
- [Polars](#)

# SDTM/ADaM Pilot Project

- [Atorus Research & Agustin Calatroni](#)
- [Pilot Submission Package](#)
- Xanomeline Transdermal
- Therapeutic System (TTS) in Patients with Mild to Moderate Alzheimer's Disease

```

gt_table_final = (
    GT(tbl)
    .tab_header(
        title="Table 14.2.01",
        subtitle="Summary of Demographic and Baseline Characteristics"
    )
    .sub_missing(missing_text="")
    .cols_hide(columns=cs.ends_with("_pct"))
    .cols_label(
        placebo=md(f"Placebo \n(N={placebo_n})"),
        xanomeline_ld=md(f"Xanomeline \nLow Dose \n(N={xanomeline_ld_n})"),
        xanomeline_hd=md(f"Xanomeline \nHigh Dose \n(N={xanomeline_hd_n})"),
        total=md(f"Total \n(N={total_n})"),
        p="p-value[1]"
    )
    .tab_source_note(
        source_note=f"Program Source: 14-2.01 Executed: (Draft)
        {datetime.now().date().isoformat()}"
    )
)
gt_table_final

```

Table 14.2.01

Summary of Demographic and Baseline Characteristics

| category      | label     | Placebo<br>(N=86) | Xanomeline<br>Low Dose<br>(N=84) | Xanomeline<br>High Dose<br>(N=84) | Total<br>(N=254) | p-value[1] |
|---------------|-----------|-------------------|----------------------------------|-----------------------------------|------------------|------------|
| Age (y)       | n         | 86.0              | 84.0                             | 84.0                              | 254              | 0.5934     |
|               | Mean      | 75.2              | 75.7                             | 74.4                              | 75.1             |            |
|               | SD        | 8.59              | 8.29                             | 7.89                              | 8.25             |            |
|               | Median    | 76.0              | 77.5                             | 76.0                              | 77.0             |            |
|               | Min       | 52.0              | 51.0                             | 56.0                              | 51.0             |            |
|               | Max       | 89.0              | 88.0                             | 88.0                              | 89.0             |            |
|               |           |                   |                                  |                                   |                  |            |
|               | <65 yrs   | 14.0 (16%)        | 8.00 (10%)                       | 11.0 (13%)                        | 33.0 (13%)       | 0.1439     |
|               | 65-80 yrs | 42.0 (49%)        | 47.0 (56%)                       | 55.0 (65%)                        | 144 (57%)        |            |
|               | >80 yrs   | 30.0 (35%)        | 29.0 (35%)                       | 18.0 (21%)                        | 77.0 (30%)       |            |
| Sex           | n         | 86.0              | 84.0                             | 84.0                              | 254              | 0.1409     |
|               | Male      | 33.0 (38%)        | 34.0 (40%)                       | 44.0 (52%)                        | 111 (44%)        |            |
|               | Female    | 53.0 (62%)        | 50.0 (60%)                       | 40.0 (48%)                        | 143 (56%)        |            |
| Race (Origin) | n         | 86.0              | 84.0                             | 84.0                              | 254              | 0.6477     |
|               | Caucasian | 75.0 (87%)        | 72.0 (86%)                       | 71.0 (85%)                        | 218 (86%)        |            |

```

gt_table_final = (
  GT(tbl)
  .tab_header(
    title="Table 14.3.13",
    subtitle="CIBIC+ - Categorical Analysis - LOCF"
  )
  .sub_missing(missing_text="")
  .cols_hide(columns=cs.ends_with("_pct"))
  .cols_label(
    category="",
    label="Assessment",
    placebo=md(f"Placebo \n(N={placebo_n})"),
    xanomeline_ld=md(f"Xanomeline \nLow Dose \n(N={xanomeline_ld_n})"),
    xanomeline_hd=md(f"Xanomeline \nHigh Dose \n(N={xanomeline_hd_n})"),
    p="p-value"
  )
  .tab_source_note(
    source_note="Overall comparison of treatments using CMH test (Pearson Chi-Square), controlling
    for site group."
  )
  .tab_source_note(
    source_note=f"{datetime.now().date().isoformat()}"
  )
)

gt_table_final

```

Table 14.3.13

CIBIC+ - Categorical Analysis - LOCF

| Assessment |                      | Placebo<br>(N=79) | Xanomeline<br>Low Dose<br>(N=81) | Xanomeline<br>High Dose<br>(N=74) | p-value |
|------------|----------------------|-------------------|----------------------------------|-----------------------------------|---------|
| Week 8     | n                    | 77                | 81                               | 73                                | 0.2727  |
|            | Marked improvement   | 0                 | 0                                | 0                                 |         |
|            | Moderate improvement | 1 (1%)            | 2 (2%)                           | 1 (1%)                            |         |
|            | Minimal improvement  | 19 (25%)          | 16 (20%)                         | 13 (18%)                          |         |
|            | No Change            | 45 (58%)          | 48 (59%)                         | 38 (52%)                          |         |
|            | Minimal worsening    | 10 (13%)          | 14 (17%)                         | 20 (27%)                          |         |
|            | Moderate worsening   | 2 (3%)            | 1 (1%)                           | 1 (1%)                            |         |
|            | Marked worsening     | 0                 | 0                                | 0                                 |         |
| Week 16    | n                    | 79                | 81                               | 74                                | 0.4003  |
|            | Marked improvement   | 0                 | 0                                | 0                                 |         |
|            | Moderate improvement | 0                 | 3 (4%)                           | 2 (3%)                            |         |
|            | Minimal improvement  | 12 (15%)          | 12 (15%)                         | 13 (18%)                          |         |
|            | No Change            | 41 (52%)          | 46 (57%)                         | 39 (53%)                          |         |
|            | Minimal worsening    | 25 (32%)          | 19 (23%)                         | 20 (27%)                          |         |
|            | Moderate worsening   | 1 (1%)            | 1 (1%)                           | 0                                 |         |

```

gt_table_final = (
  GT(tbl)
  .tab_header(
    title="Table 14-5.01",
    subtitle="Incidence of Treatment Emergent Adverse Events by Treatment
    Group",
  )
  .sub_missing(missing_text="")
  .cols_hide(columns=cs.ends_with("_pct"))
  .cols_label(
    label="Label",
    placebo_n="n(%)",
    xanomeline_ld_n="n(%)",
    xanomeline_hd_n="n(%)",
    placebo_ae="[AEs]",
    xanomeline_ld_ae="[AEs]",
xanomeline_hd_ae="[AEs]",
    fisher_placebo_v_xanomeline_ld=md("Placebo \nvs. \nLow Dose"),
    fisher_placebo_v_xanomeline_hd=md("Placebo \nvs. \nHigh Dose")
  )
  .tab_spanner(
    label=md(f"Placebo \n(N={placebo_n})"),
    columns=cs.starts_with("placebo")
  )
  .tab_spanner(
    label=md(f"Xanomeline \nLow Dose \n(N={xanomeline_ld_n})"),
    columns=cs.starts_with("xanomeline_ld")
  )
  .tab_source_note(source_note = "Note: Treatment emergent events are defined as events which start on or
  .tab_source_note(source_note = "Note: Adverse events are coded using MedDRA.")
  .tab_source_note(source_note = "Note: Percentages are based on the number of subjects in the safety pop
  .tab_source_note(source_note = "Note: P-values are based on Fisher's Exact test for the comparison of p
  .tab_source_note(source_note = "Note: The column [AE] represents the total number of times an event26 was

```

Table 14-5.01

Incidence of Treatment Emergent Adverse Events by Treatment Group

| Label                               | Placebo<br>(N=86) |       | Xanomeline<br>Low Dose<br>(N=84) |       | Xanomeline<br>High Dose<br>(N=84) |       | Fisher's Exact<br>p-values    |                                |
|-------------------------------------|-------------------|-------|----------------------------------|-------|-----------------------------------|-------|-------------------------------|--------------------------------|
|                                     | n(%)              | [AEs] | n(%)                             | [AEs] | n(%)                              | [AEs] | Placebo<br>vs.<br>Low<br>Dose | Placebo<br>vs.<br>High<br>Dose |
| ANY BODY SYSTEM                     | 65<br>(76%)       | [281] | 77<br>(92%)                      | [412] | 76<br>(90%)                       | [433] | 0.007                         | 0.014                          |
| CARDIAC DISORDERS                   | 12<br>(14%)       | [26]  | 13<br>(16%)                      | [30]  | 15<br>(18%)                       | [30]  | 0.831                         | 0.534                          |
| SINUS BRADYCARDIA                   | 2 (2%)            | [2]   | 7 (8%)                           | [10]  | 8 (10%)                           | [12]  | 0.097                         | 0.056                          |
| MYOCARDIAL INFARCTION               | 4 (5%)            | [4]   | 2 (2%)                           | [4]   | 4 (5%)                            | [8]   | 0.682                         | 0.999                          |
| ATRIAL FIBRILLATION                 | 1 (1%)            | [1]   | 1 (1%)                           | [1]   | 3 (4%)                            | [5]   | 0.999                         | 0.365                          |
| ATRIAL FLUTTER                      | 0                 |       | 1 (1%)                           | [1]   | 1 (1%)                            | [2]   | 0.494                         | 0.494                          |
| CARDIAC DISORDER                    | 0                 |       | 0                                |       | 1 (1%)                            | [1]   |                               | 0.494                          |
| SUPRAVENTRICULAR EXTRASYSTOLES      | 1 (1%)            | [2]   | 1 (1%)                           | [2]   | 1 (1%)                            | [1]   | 0.999                         | 0.999                          |
| VENTRICULAR EXTRASYSTOLES           | 0                 |       | 2 (2%)                           | [4]   | 1 (1%)                            | [1]   | 0.243                         | 0.494                          |
| ATRIAL HYPERTROPHY                  | 1 (1%)            | [2]   | 0                                |       | 0                                 |       | 0.999                         | 0.999                          |
| ATRIOVENTRICULAR BLOCK FIRST DEGREE | 1 (1%)            | [1]   | 1 (1%)                           | [1]   | 0                                 |       | 0.999                         | 0.999                          |

# Pharma Work on Top of Packages

- [Teal](#)
- [gtsummary](#)
- [Admiral](#)



`{gtsummary}`:  
Tables for Research and  
Regulatory Submissions

PySummaries

---

---

---

# pysummaries by Otto Fajardo @ Genentech

```
summary_table = get_table_summary(df, strata='group', backend='native')  
display(summary_table)
```

|                  | Control<br>(N=50)  | Experimental<br>(N=51) | Overall<br>(N=101) |
|------------------|--------------------|------------------------|--------------------|
| <b>gender</b>    |                    |                        |                    |
| Female           | 15 (30.0 %)        | 22 (43.1 %)            | 37 (36.6 %)        |
| Male             | 34 (68.0 %)        | 29 (56.9 %)            | 63 (62.4 %)        |
| Missing          | 1 (2.0 %)          | 0 (0%)                 | 1 (1.0 %)          |
| <b>age</b>       |                    |                        |                    |
| Mean (SD)        | 53.9 (17.1)        | 49.3 (17.6)            | 51.5 (17.4)        |
| Median [Q1 ; Q3] | 58.0 [46.0 ; 68.0] | 49.0 [33.0 ; 63.5]     | 51.5 [40.5 ; 68.0] |
| Min ; Max        | 20.0 ; 78.0        | 20.0 ; 78.0            | 20.0 ; 78.0        |
| Missing          | 1 (2.0 %)          | 0 (0.0 %)              | 1 (1.0 %)          |
| <b>region</b>    |                    |                        |                    |
| East             | 7 (14.0 %)         | 6 (11.8 %)             | 13 (12.9 %)        |
| Missing          | 1 (2.0 %)          | 0 (0%)                 | 1 (1.0 %)          |
| North            | 11 (22.0 %)        | 11 (21.6 %)            | 22 (21.8 %)        |
| South            | 8 (16.0 %)         | 15 (29.4 %)            | 23 (22.8 %)        |
| West             | 23 (46.0 %)        | 19 (37.3 %)            | 42 (41.6 %)        |

# Interactive TLGs, ARDs & Shiny for Python

- [Pilot 4](#)
- [WebAssembly, WebR & Shinylive](#)
- [Analysis Results Dataset in gtsummary](#)
- [Shiny for Python, Teal?](#)

Great Tables

# Collaboration

R in  
Pharma

R  
Submissions  
Working  
Group

Great Tables

# Questions

[phil@rstudio.com](mailto:phil@rstudio.com)

