



Table-based Data Visualizations using R



PHUSE US Connect 2021

Matthew Kumar

Matthew.Kumar@Bayer.com





Agenda

// Overview

// An R workflow for table-based data visualizations

// Applications to trial data

// Extensions



Overview and Motivations

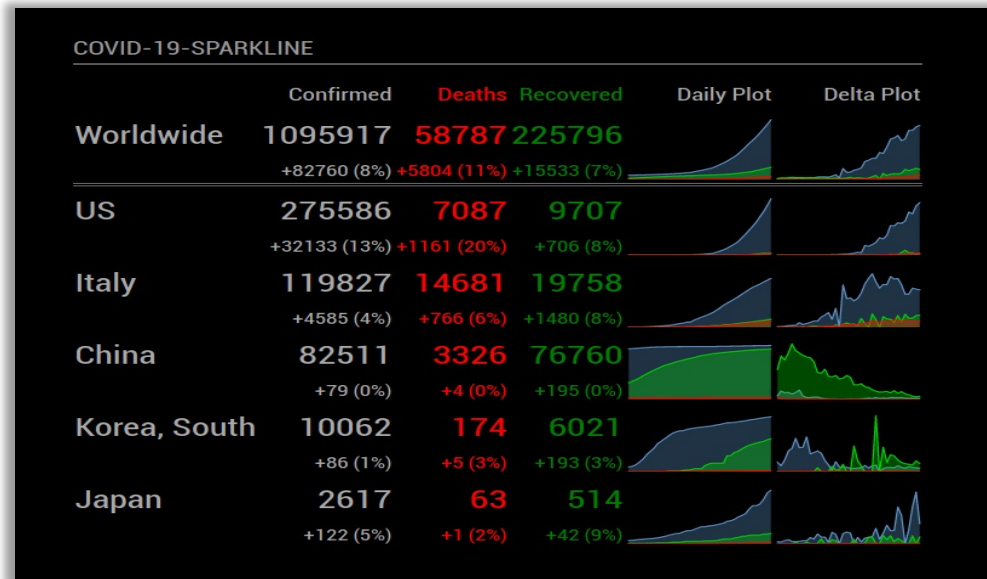
Table-based data visualizations in R



Table-based Data Visualizations

Overview

Tiny data visualizations integrated directly into a table



feed	weight	boxplot	sparkline
casein			
horsebean			
linseed			
meatmeal			
soybean			
sunflower			



Table-based Data Visualizations

Motivations

Table 1 summarizes two continuous parameters for each group

Group	Parameter	
	P1 [†]	P2 [†]
1	9.0 (3.32)	7.5 (2.03)
2	9.0 (3.32)	7.5 (2.03)
3	9.0 (3.32)	7.5 (2.03)
4	9.0 (3.32)	7.5 (2.03)
[†] Parameters are summarized by Mean (SD)		



Table-based Data Visualizations

Motivations

Figure 1 summarizes the *relationship* between parameters in a scatterplot for each group

Group	Parameter	
	P1 [†]	P2 [†]
1	9.0 (3.32)	7.5 (2.03)
2	9.0 (3.32)	7.5 (2.03)
3	9.0 (3.32)	7.5 (2.03)
4	9.0 (3.32)	7.5 (2.03)

[†] Parameters are summarized by Mean (SD)

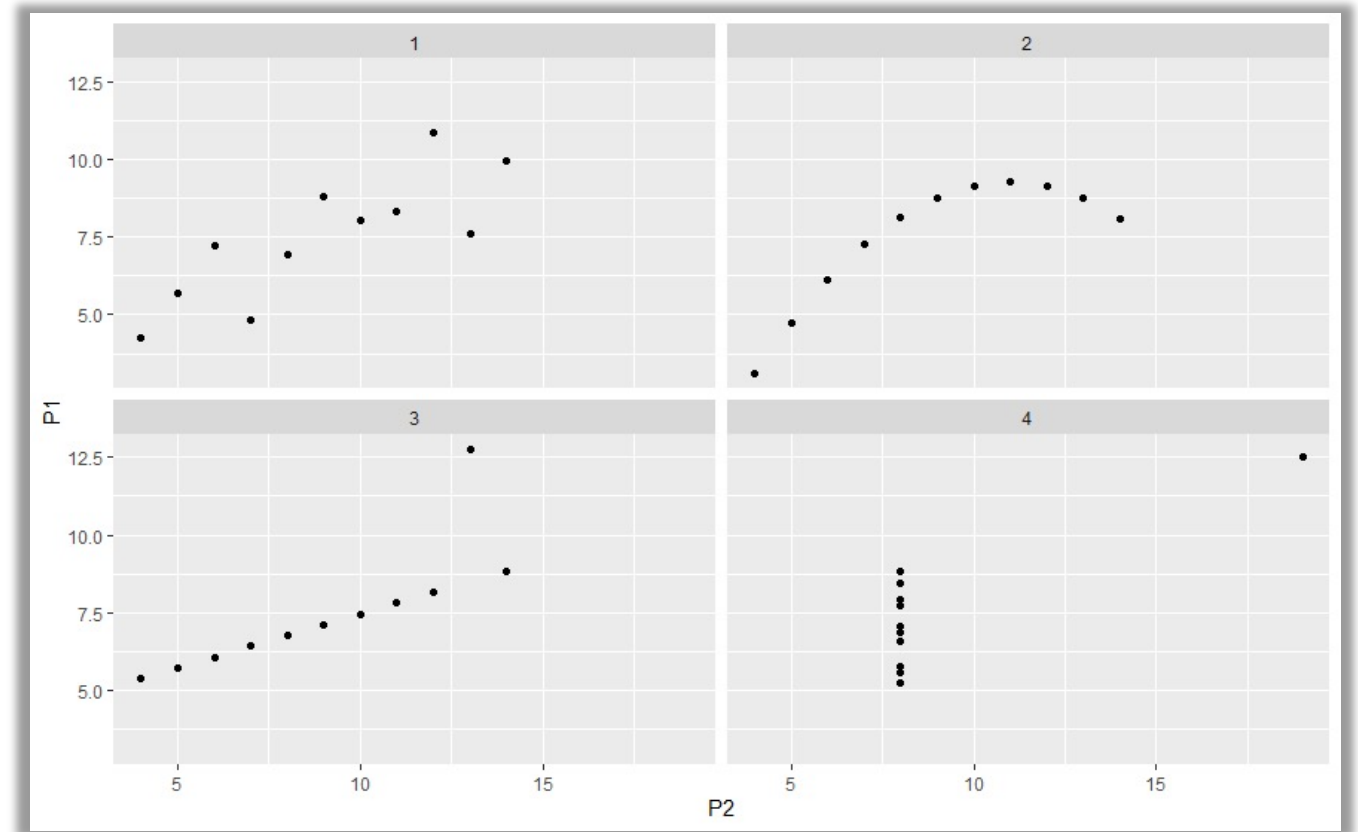




Table-based Data Visualizations

Motivations

Table 1 provides both a numerical and visual summary

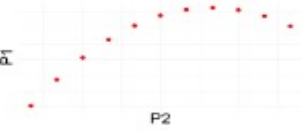
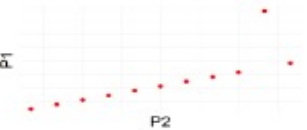
Group	Parameter		Scatter
	P1 [†]	P2 [†]	
1	9.0 (3.32)	7.5 (2.03)	
2	9.0 (3.32)	7.5 (2.03)	
3	9.0 (3.32)	7.5 (2.03)	
4	9.0 (3.32)	7.5 (2.03)	
† Parameters are summarized by Mean (SD)			



Table-based Data Visualizations

Motivations

- // Table-based data visualizations can offer a succinct, visual depiction of the data at hand
 - // Summarizing an entire distribution or temporal trend of measurements in a single cell of a table
- // They can be effective at providing context to data
 - // “Quick stories”
 - // Signal expected or unexpected findings
 - // Enable between and within comparisons between rows of a table
- // Can be applied to a multitude of functions which involve micro or macro decision processes
 - // Medical review
 - // Data quality assessments



An R workflow example

*Table-based data
visualizations in R*



R Packages

An extensive collection of contributed table and data visualization packages are freely available to R users through CRAN.



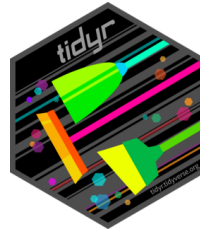


The Current R Workflow

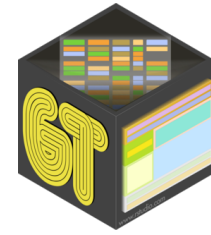
1. Create summary tabulations



2. Create inline data visualizations



3. Display both summary tabulations and data visualizations in a table





A Minimal Example

<https://github.com/phuse-org>

// Use dplyr (or other packages) to create the summary tabulation *content* to be displayed

```
# load packages
require(tidyverse)

# read in adsl
adsl <- haven::read_xpt("https://github.com/phuse-org/TestDataFactory/raw/main/Updated/TDF_ADaM/adsl.xpt")

# calculate mean and sd for years of education, for each treatment
summary_tab <- adsl %>%
  group_by(TRT01P) %>%
  summarise(avg_edu = mean(EDUCLVL), sd_edu = sd(EDUCLVL))
```

summary_tab x			
Filter			
	TRT01P Planned Treatment for Period 01	avg_edu	sd_edu
1	Placebo	12.58140	2.948440
2	Xanomeline High Dose	12.51190	2.918549
3	Xanomeline Low Dose	13.16667	4.147385





A Minimal Example

Create data visualizations

// Use ggplot2 to create the data visualization

// 'core' ggplot2 code

// A few tweaks are needed to:

// Create a visualization for each row of interest (e.g., treatment)

// Save the collection of visualizations to a data frame

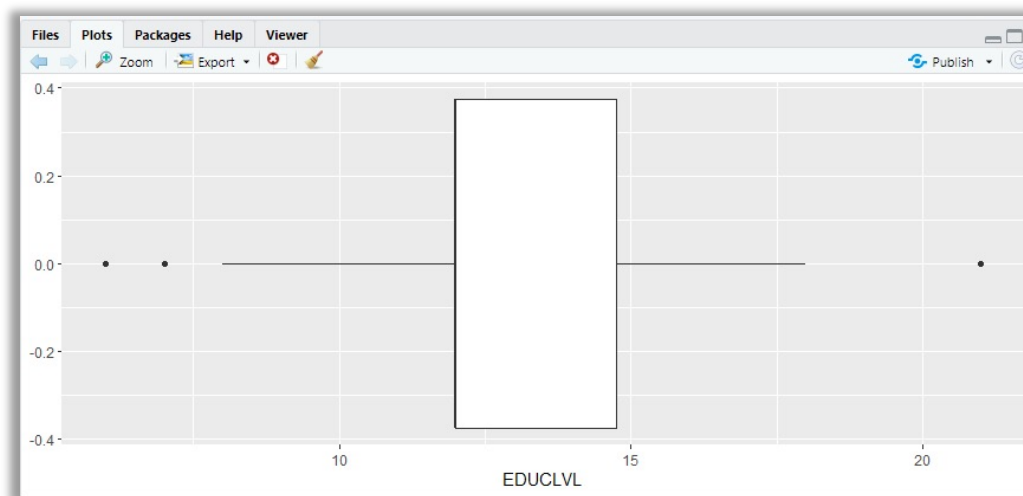
```
# create a box plot for each treatment
plot_holder <- adsl %>%
  group_by(TRT01P) %>%
  nest() %>%
  mutate(rowplot = map(data, ~ ggplot(., aes(x = EDUCLVL)) +geom_boxplot()))
```



A Minimal Example

Create data visualizations

plot_holder x			
Filter			
	TRT01P Planned Treatment for Period 01	data	rowplot
1	Placebo	135 variables	<code>list(data = list(STUDYID = c("CDISCILOT01", "CDISCILOT0...))</code>
2	Xanomeline High Dose	140 variables	<code>list(data = list(STUDYID = c("CDISCILOT01", "CDISCILOT0...))</code>
3	Xanomeline Low Dose	139 variables	<code>list(data = list(STUDYID = c("CDISCILOT01", "CDISCILOT0...))</code>





A Minimal Example

Display everything

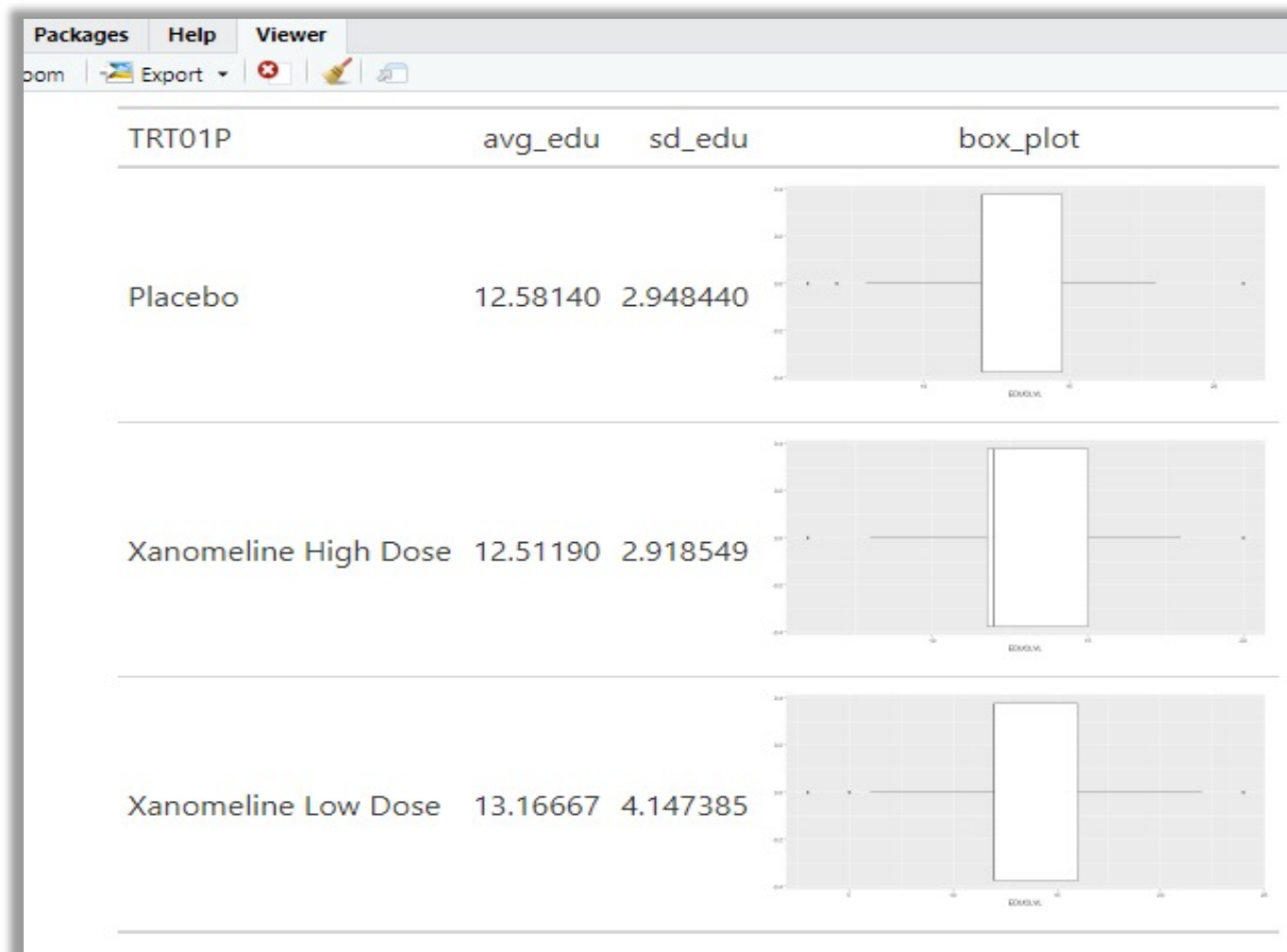
// Use gt to create a display table of the summary tabulations

// Specify the column where the data visualizations will appear

```
# display summary tabulations + graphs
summary_tab %>%
  mutate(box_plot = NA) %>%
  gt() %>%
  text_transform(
    locations = cells_body(vars(box_plot)),
    fn = function(x) {
      map(plot_holder$rowplot, ggplot_image, aspect_ratio = 2, height = 125)
    }
  )
```



A Minimal Example





A Minimal Example

Considerations

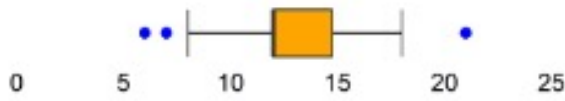
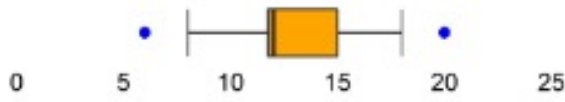
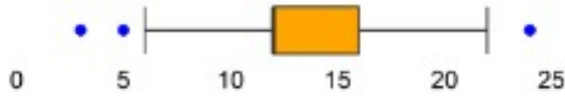
- // Modify components of the data visualization to increase clarity
 - // Grid lines, background color, axis text, point size
- // Modify components of the display table to increase clarity
 - // Height, width
- // Ensure each individual graph is on a common scale
 - // Plotting each treatment arm individually may result in different default axis scaling
- // Ensure mappings are correct
 - // Does each row in the summary table receive the correct corresponding plot?



A Minimal Example

Reworked for clarity

Table: Summary of Education by Treatment

Treatment	Education (Years) ¹	Distribution
Placebo	12.58 (2.948)	
Xanomeline High Dose	12.51 (2.919)	
Xanomeline Low Dose	13.17 (4.147)	

¹ Education (Years) is summarized by Mean (SD)



Examples and Ideas

*Table-based data
visualizations in R*



Examples

Inline plots that depict temporal *trends*

// PK

// Patient Profiles

// Exposure

// Escalation, Reduction, Modifications

// Enrollment Status

// Country, Site, Disease Populations

Table: Subject Pharmacokinetic Summary

Subject	Dose	Baseline		Follow-Up	
		Baseline	Weight [†]	Week 6	Week 12
1001-01	80mg		82.10		
1001-02	60mg		78.41		
1001-03	40mg		71.41		

[†] Weight is reported in kilograms, kg



Examples

Inline plots that depict *timing* of events

// Efficacy

// Response – swimmer plot

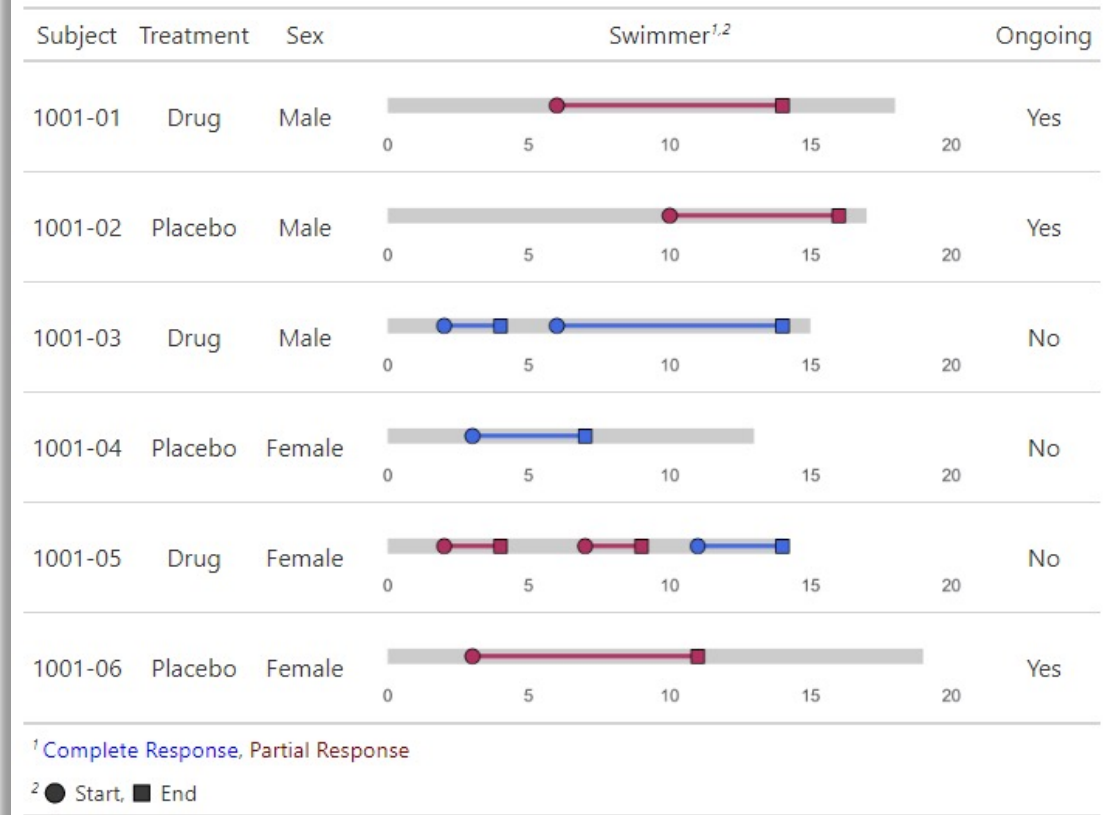
// Safety

// Special AEs of Interest

// Disposition

// Subject trajectories during the study

Table: Subject Response Summary





Examples

Inline plots that depict *statistical results*

// Efficacy

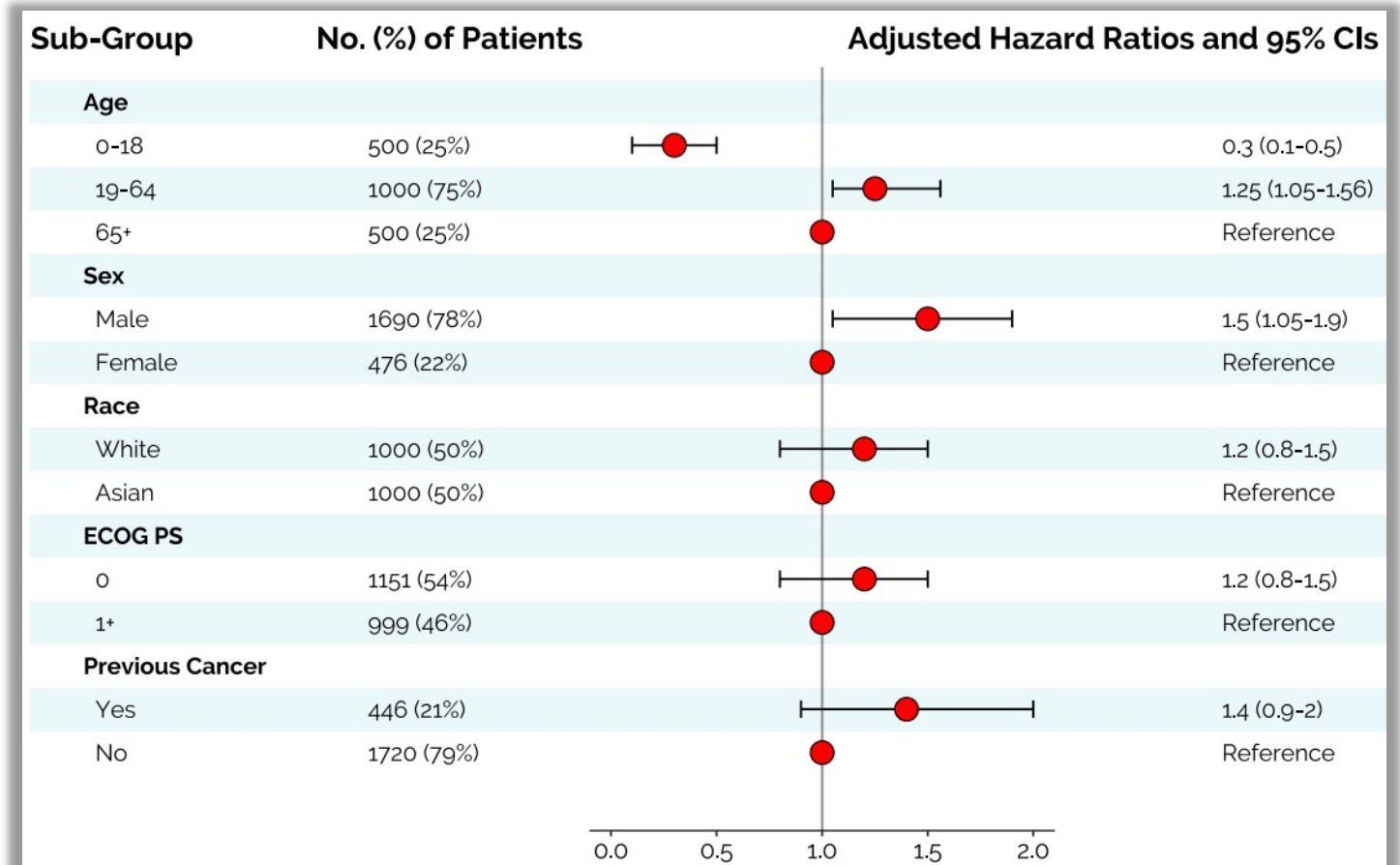
// Odds ratios, hazard ratios

// Response

// Best overall response (n, %)

// Safety

// AE incidence rates, relative risks





Extensions

*Table-based data
visualizations in R*



Extensions



- // Table-based data visualizations can be extended to both Shiny and Rmarkdown
- // An increasingly popular mode to deliver study results
 - // Web applications, dashboards
 - // Dynamic standalone documents
- // Leverage each framework's native interactivity to enhance exploration and communicate insights
 - // Value filtering, variable selection, sorting
- // Most table and visualization packages extend seamlessly to these frameworks
 - // Enhanced visualizations: contextual tool tips, animation, zoom



Extensions

Shiny, Rmarkdown



Table of Codes

Show 50 entries

Search:

	Code	Description	BarPlot
5	E324	Vaccination Type 1	
6	F435	Vaccination Type 2	
7	G435	Vaccination Type 3	
8	H234	Annual Exam (19+)	
9	I342	Annual Exam (0-18)	
10	J034	Endocrinology - Consultation	

Showing 1 to 10 of 10 entries

Previous 1 Next

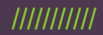
[Clear Selections](#) [Download Selections](#)

Cancer

	2014-04	2014-10	2015-04	2015-10	2016-04	2016-10	2017-04	2017-10	Trend
	60.55	60.61	59.72	62.81	60.81	60.31	62.86	61.83	
	62.87	61.34	63.19	62.26	60.36	61.91	63.76	59.04	
	44.94	47.41	45.85	45.81	48.34	48.52	47.09	44.67	
	68.35	67.29	68.78	68.32	66.78	66.96	66.18	66.13	
	69.58	70.44	70.95	70.74	69.97	69.82	70.6	70.25	
	68.46	67.16	67.87	67.47	68.58	67.52	68.97	68.51	
	65.8	58.19	65.49	57.49	62	65.46	62.88	61.06	
	31.86	31.68	30.24	33.09	34.81	33.27	32.55	30.75	
	44.35	42.57	40.04	40.1	40.72	41.53	44.13	42.51	



Thank you!



Bye-Bye

