PharmaSUG
NORTH CAROLINA • OCT. 21, 2022

Clinical Tables With the Latest

in Tplyr

Mike Stackhouse
Jessica Higgins, PhD

atorus
Your Complete Data Solution

Why'd we make Tplyr?

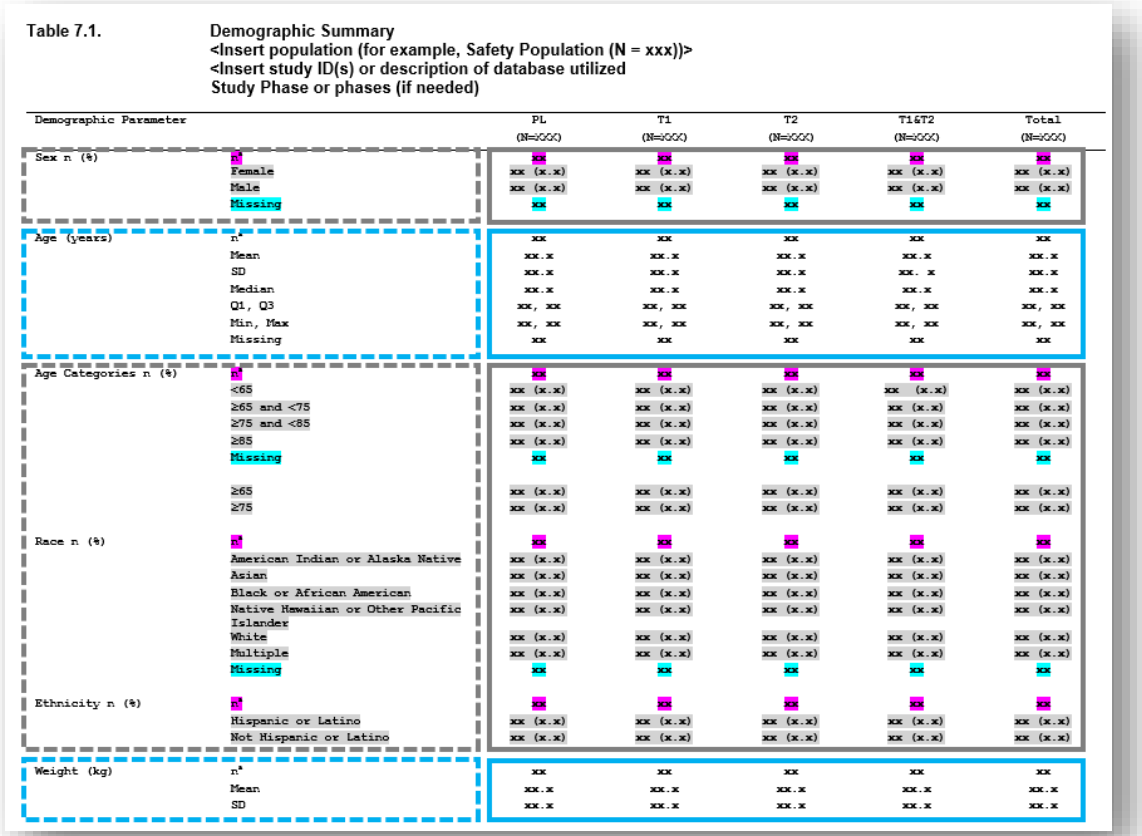


Table 7.1. Demographic Summary
<Insert population (for example, Safety Population (N = xxx))>
<Insert study ID(s) or description of database utilized
Study Phase or phases (if needed)

Demographic Parameter		PL (N=XXX)	T1 (N=XXX)	T2 (N=XXX)	T1&T2 (N=XXX)	Total (N=XXX)
Sex n (%)	n*	xx	xx	xx	xx	xx
	Female	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Male	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Age (years)	n*	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx. x	xx.x
	Median	xx.x	xx.x	xx.x	xx.x	xx.x
	Q1, Q3	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Missing	xx	xx	xx	xx	xx
Age Categories n (%)	n*	xx	xx	xx	xx	xx
	<65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥65 and <75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75 and <85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
	≥65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Race n (%)	n*	xx	xx	xx	xx	xx
	American Indian or Alaska Native	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Asian	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Black or African American	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Native Hawaiian or Other Pacific Islander	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	White	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Multiple	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Ethnicity n (%)	n*	xx	xx	xx	xx	xx
	Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Not Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Weight (kg)	n*	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx.x	xx.x

Why'd we make Tplyr?

1. Summary tables are highly repetitive, so code should be highly reusable

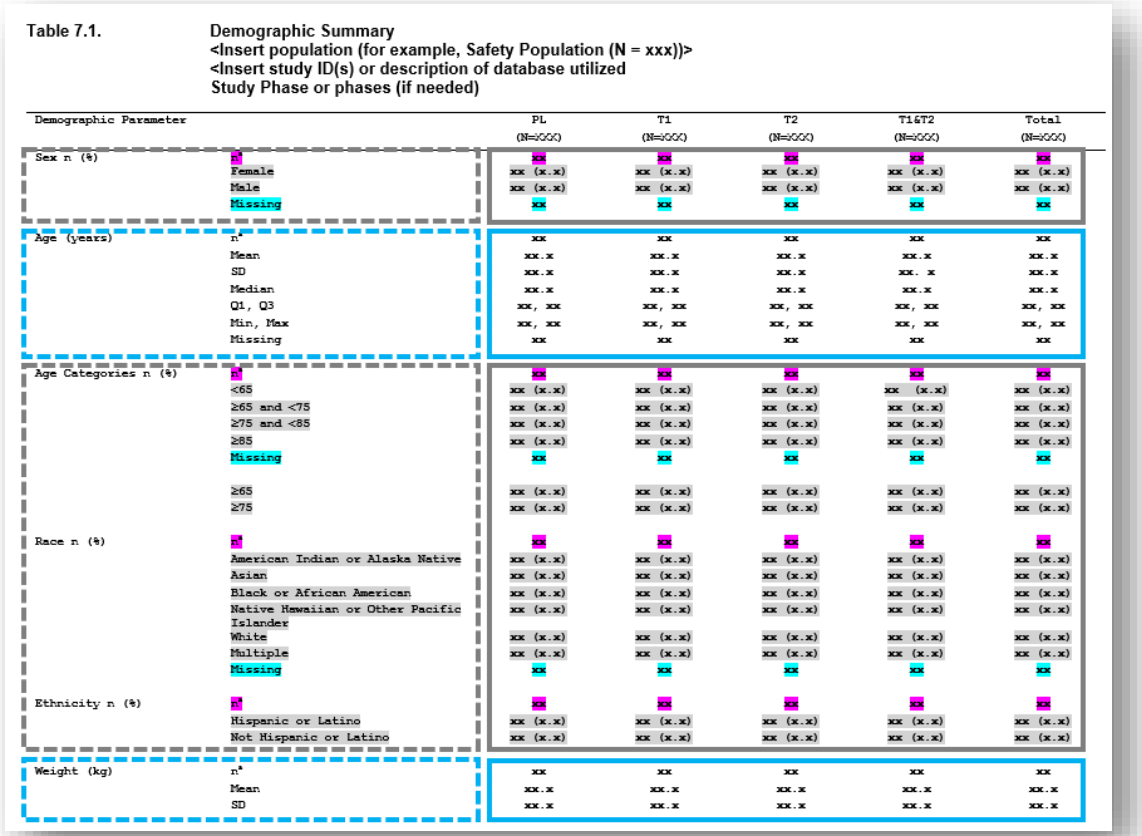


Table 7.1. Demographic Summary
<Insert population (for example, Safety Population (N = xxx))>
<Insert study ID(s) or description of database utilized
Study Phase or phases (if needed)

Demographic Parameter		PL (N=XXX)	T1 (N=XXX)	T2 (N=XXX)	T1&T2 (N=XXX)	Total (N=XXX)
Sex n (%)	n*	xx	xx	xx	xx	xx
	Female	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Male	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Age (years)	n*	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx. x	xx.x
	Median	xx.x	xx.x	xx.x	xx.x	xx.x
	Q1, Q3	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Missing	xx	xx	xx	xx	xx
Age Categories n (%)	n*	xx	xx	xx	xx	xx
	<65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥65 and <75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75 and <85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
	≥65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Race n (%)	n*	xx	xx	xx	xx	xx
	American Indian or Alaska Native	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Asian	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Black or African American	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Native Hawaiian or Other Pacific Islander	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	White	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Multiple	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Ethnicity n (%)	n*	xx	xx	xx	xx	xx
	Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Not Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Weight (kg)	n*	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx.x	xx.x

Why'd we make Tplyr?

1. Summary tables are highly repetitive, so code should be highly reusable
2. Text formatting of numbers can be very tedious

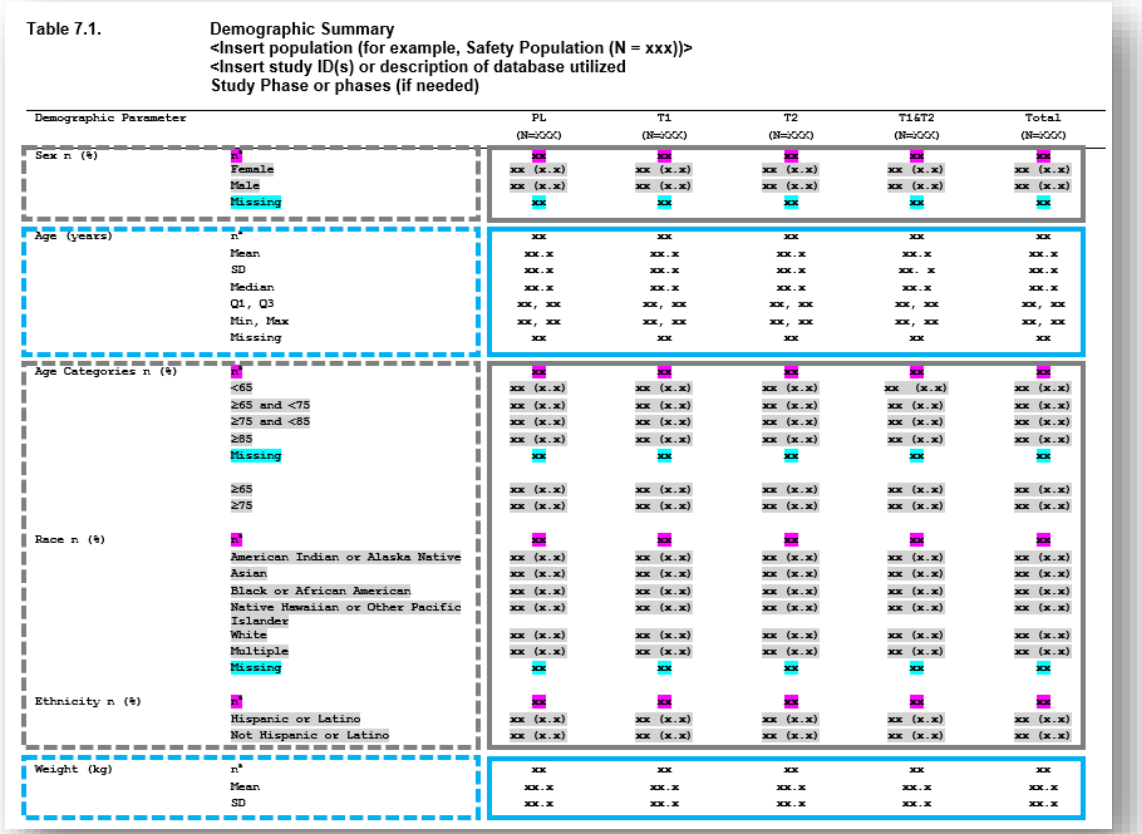


Table 7.1. Demographic Summary
<Insert population (for example, Safety Population (N = xxx))>
<Insert study ID(s) or description of database utilized
Study Phase or phases (if needed)

Demographic Parameter		PL (N=XXX)	T1 (N=XXX)	T2 (N=XXX)	T1&T2 (N=XXX)	Total (N=XXX)
Sex n (%)	n	xx	xx	xx	xx	xx
	Female	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Male	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Age (years)	n	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx. x	xx.x
	Median	xx.x	xx.x	xx.x	xx.x	xx.x
	Q1, Q3	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Missing	xx	xx	xx	xx	xx
Age Categories n (%)	n	xx	xx	xx	xx	xx
	<65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥65 and <75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75 and <85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
	≥65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Race n (%)	n	xx	xx	xx	xx	xx
	American Indian or Alaska Native	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Asian	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Black or African American	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Native Hawaiian or Other Pacific Islander	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	White	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Multiple	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Ethnicity n (%)	n	xx	xx	xx	xx	xx
	Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Not Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Weight (kg)	n	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx.x	xx.x

Why'd we make Tplyr?

1. Summary tables are highly repetitive, so code should be highly reusable
2. Text formatting of numbers can be very tedious
3. Bridging into interactive tables requires metadata for traceability

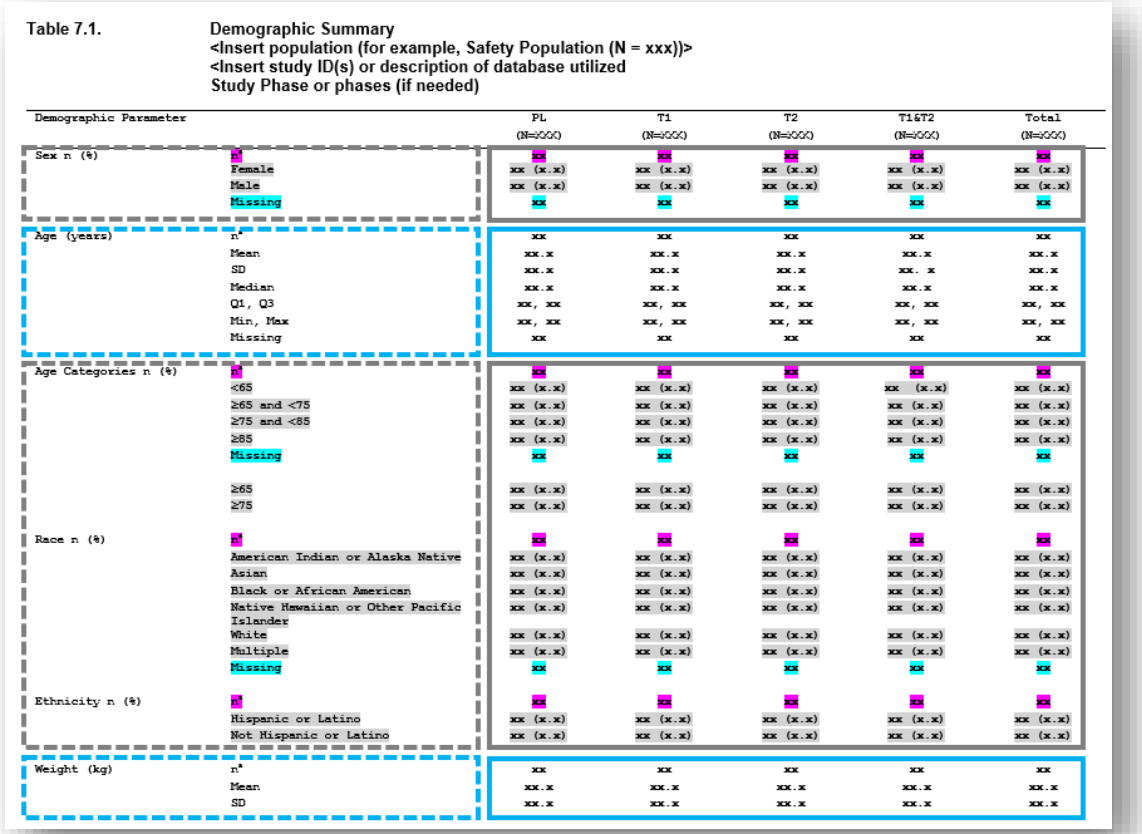


Table 7.1. Demographic Summary
<Insert population (for example, Safety Population (N = xxx))>
<Insert study ID(s) or description of database utilized
Study Phase or phases (if needed)

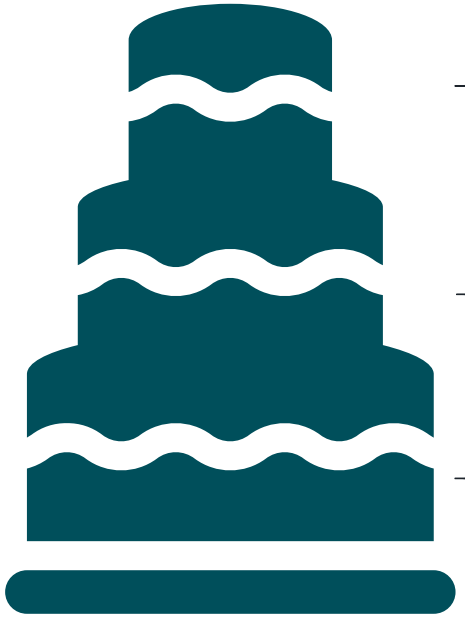
Demographic Parameter		PL (N=XXX)	T1 (N=XXX)	T2 (N=XXX)	T1&T2 (N=XXX)	Total (N=XXX)
Sex n (%)	n	xx	xx	xx	xx	xx
	Female	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Male	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Age (years)	n	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx. x	xx.x
	Median	xx.x	xx.x	xx.x	xx.x	xx.x
	Q1, Q3	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
	Missing	xx	xx	xx	xx	xx
Age Categories n (%)	n	xx	xx	xx	xx	xx
	<65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥65 and <75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75 and <85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥85	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
	≥65	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	≥75	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Race n (%)	n	xx	xx	xx	xx	xx
	American Indian or Alaska Native	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Asian	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Black or African American	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Native Hawaiian or Other Pacific Islander	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	White	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Multiple	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Missing	xx	xx	xx	xx	xx
Ethnicity n (%)	n	xx	xx	xx	xx	xx
	Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
	Not Hispanic or Latino	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)	xx (x.x)
Weight (kg)	n	xx	xx	xx	xx	xx
	Mean	xx.x	xx.x	xx.x	xx.x	xx.x
	SD	xx.x	xx.x	xx.x	xx.x	xx.x

What is Tplyr?

- R Package released in 2020
- A grammar of clinical summary tables
- Design by describing the output
- In goes data, out goes presentation ready data.frame

```
tplyr_table(adsl, TRT01P) %>%
   add_layer(
    group_count(AGEGR1, by = "Age Group n (%)") %>%
      set_format_strings(f_str("xx (xx.x%)", n, pct)) %>%
   add_layer(
    group_desc(AGE, by = "Age (years)") %>%
      set_format_strings(
        "n" = f_str("xx", n),
        "Mean (SD)" = f_str("xx.x (xx.xx)", mean, sd)
      ) %>%
  ) %>%
  build()
```

Think Cake!



		Placebo	Xanomeline Low Dose	Xanomeline High Dose
Age Group n (%)	<65	14 (16.3%)	8 (9.5%)	11 (13.1%)
	65-80	42 (48.8%)	47 (56.0%)	55 (65.5%)
	>80	30 (34.9%)	29 (34.5%)	18 (21.4%)
Age (years)	n	86	84	84
	Mean (SD)	75.2 (8.59)	75.7 (8.29)	74.4 (7.89)
	Median	76	77.5	76
	Q1, Q3	69.2, 81.8	71.0, 82.0	70.8, 80.0
	Min, Max	52, 89	51, 88	56, 88
	Missing	0	0	0
Race n (%)	American Indian or Alaskan Native	0 (0.0%)	0 (0.0%)	1 (1.2%)
	Black or African American	8 (9.3%)	6 (7.1%)	9 (10.7%)
	White	78 (90.7%)	78 (92.9%)	74 (88.1%)

New Features in Tplyr 1.0.0

- Traceability metadata framework
- Re-usable layer templates
- External precision data
- Descriptive statistics as columns
- Apply string formatting outside of Tplyr tables
- Hyphenated string wrapping

Tplyr's Metadata Concept

PharmaSUG NORTH CAROLINA • OCT. 21, 2022

		Placebo	Xanomeline Low Dose	Xanomeline High Dose
Age Group n (%)	<65	14 (16.3%)	8 (9.5%)	11 (13.1%)
	65-80	42 (48.8%)	47 (56.0%)	55 (65.5%)
	>80	30 (34.9%)	29 (34.5%)	18 (21.4%)
Age (years)	n	86	84	84
	Mean (SD)	75.2 (8.59)	75.7 (8.29)	74.4 (7.89)
	Median	76	77.5	76
	Q1, Q3	69.2, 81.8	71.0, 82.0	70.8, 80.0
	Min, Max	52, 89	51, 88	56, 88
	Missing	0	0	0
Race n (%)	American Indian or Alaskan Native	0 (0.0%)	0 (0.0%)	1 (1.2%)
	Black or African American	8 (9.3%)	6 (7.1%)	9 (10.7%)
	White	78 (90.7%)	78 (92.9%)	74 (88.1%)

What data produced this result?

atorus
Your Complete Data Solution

Tplyr's Metadata Concept

		Placebo	Xanomeline Low Dose	Xanomeline High Dose
Age Group n (%)	<65	14 (16.3%)	8 (9.5%)	11 (13.1%)
	65-80	42 (48.8%)	47 (56.0%)	55 (65.5%)
	>80	30 (34.9%)	29 (34.5%)	18 (21.4%)
Age (years)	n	86	84	84
	Mean (SD)	75.2 (8.59)	75.7 (8.29)	74.4 (7.89)
	Median	76	77.5	76
	Q1, Q3	69.2, 81.8	71.0, 82.0	70.8, 80.0
	Min, Max	52, 89	51, 88	56, 88
	Missing	0	0	0
Race n (%)	American Indian or Alaskan Native	0 (0.0%)	0 (0.0%)	1 (1.2%)
	Black or African American	8 (9.3%)	6 (7.1%)	9 (10.7%)
	White	78 (90.7%)	78 (92.9%)	74 (88.1%)

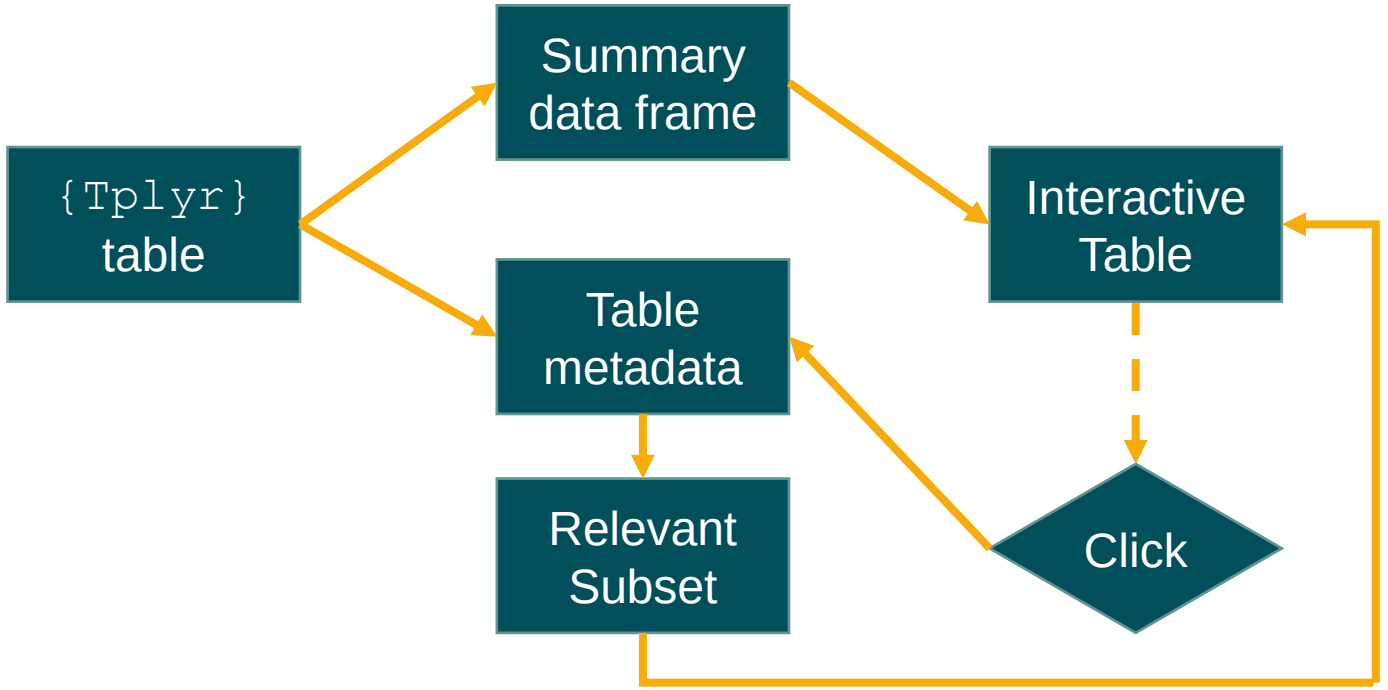
What data produced this result?

What Do We Mean By "Metadata"

metadata

- What **variables** were necessary to derive this result?
- What **filters** were necessary to derive this result?

Using Tplyr's Metadata



Building Metadata

```
t <- tplyr_table(adsl, TRT01P) %>%
  add_layer(
    group_count(AGEGR1, by = "Age Group n (%)")
  ) %>%
  add_layer(
    group_desc(AGE, by = "Age (years)")
  ) %>%
  add_layer(
    group_count(RACE, by = "Race n (%)")
  )

t %>%
  build(metadata = TRUE) %>%
  apply_row_masks()
```

row_id			Placebo
c1_1	Age Group n (%)	<65	14 (16.3%)
c2_1		65-80	42 (48.8%)
c3_1		>80	30 (34.9%)
d1_2	Age (years)	n	86
d2_2		Mean (SD)	75.2 (8.59)
d3_2		Median	76
d4_2		Q1, Q3	69.2, 81.8
d5_2		Min, Max	52, 89
d6_2		Missing	0
c1_3	Race n (%)	American Indian or Alaskan Native	0 (0.0%)
c2_3		Black or African American	8 (9.3%)
c3_3		White	78 (90.7%)

Building Metadata

PharmaSUG NORTH CAROLINA • OCT. 21, 2022

row_id			Placebo
c1_1	Age Group n (%)	<65	14 (16.3%)
c2_1		65-80	42 (48.8%)
c3_1		>80	30 (34.9%)
d1_2	Age (years)	n	86
d2_2		Mean (SD)	75.2 (8.59)
d3_2		Median	76
d4_2		Q1, Q3	69.2, 81.8
d5_2		Min, Max	52, 89
d6_2		Missing	0
c1_3	Race n (%)	American Indian or Alaskan Native	0 (0.0%)
c2_3		Black or African American	8 (9.3%)
c3_3		White	78 (90.7%)

Output Data Frame

```
# A tibble: 12 × 4
   row_id row_label1      row_label2                        Placebo
   <chr>  <chr>           <chr>                             <list>
 1 c1_1   Age Group n (%) <65                               <tplyr_mt>
 2 c2_1   Age Group n (%) 65-80                             <tplyr_mt>
 3 c3_1   Age Group n (%) >80                               <tplyr_mt>
 4 d1_2   Age (years)     n                                 <tplyr_mt>
 5 d2_2   Age (years)     Mean (SD)                         <tplyr_mt>
 6 d3_2   Age (years)     Median                            <tplyr_mt>
 7 d4_2   Age (years)     Q1, Q3                            <tplyr_mt>
 8 d5_2   Age (years)     Min, Max                          <tplyr_mt>
 9 d6_2   Age (years)     Missing                           <tplyr_mt>
10 c1_3   Race n (%)      American Indian or Alaskan Native<tplyr_mt>
11 c2_3   Race n (%)      Black or African American         <tplyr_mt>
12 c3_3   Race n (%)      White                             <tplyr_mt>
```

Tplyr Metadata

Building Metadata

row_id			Placebo
c1_1	Age Group n (%)	<65	14 (16.3%)
c2_1		65-80	42 (48.8%)
c3_1		>80	30 (34.9%)
d1_2	Age (years)	n	86
d2_2		Mean (SD)	75.2 (8.59)
d3_2		Median	76
d4_2		Q1, Q3	69.2, 81.8
d5_2		Min, Max	52, 89
d6_2		Missing	0
c1_3	Race n (%)	American Indian or Alaskan Native	0 (0.0%)
c2_3		Black or African American	8 (9.3%)
c3_3		White	78 (90.7%)

Output Data Frame

```
# A tibble: 12 × 4
   row_id row_label1      row_label2                        Placebo
   <chr>  <chr>           <chr>                             <list>
 1 c1_1   Age Group n (%) <65                               <tplyr_mt>
 2 c2_1   Age Group n (%) 65-80                             <tplyr_mt>
 3 c3_1   Age Group n (%) >80                               <tplyr_mt>
 4 d1_2   Age (years)     n                                 <tplyr_mt>
 5 d2_2   Age (years)     Mean (SD)                         <tplyr_mt>
 6 d3_2   Age (years)     Median                            <tplyr_mt>
 7 d4_2   Age (years)     Q1, Q3                            <tplyr_mt>
 8 d5_2   Age (years)     Min, Max                          <tplyr_mt>
 9 d6_2   Age (years)     Missing                           <tplyr_mt>
10 c1_3   Race n (%)      American Indian or Alaskan Native<tplyr_mt>
11 c2_3   Race n (%)      Black or African American         <tplyr_mt>
12 c3_3   Race n (%)      White                             <tplyr_mt>
```

Tplyr Metadata

Extracting Metadata


```
get_meta_result(t, 'c2_3', 'Placebo')
```

Tplyr table *row_id* *column name*

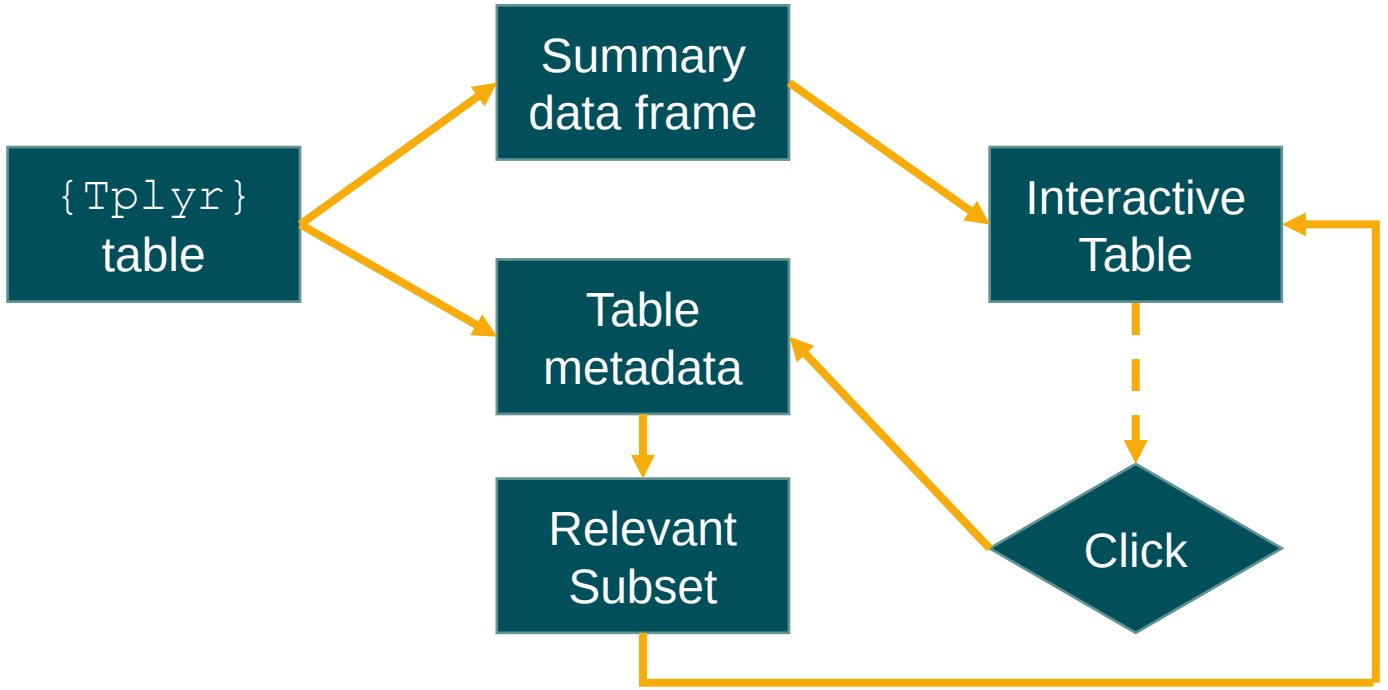
```
tplyr_meta: 2 names, 4 filters
Names:
    TRT01P, RACE
Filters:
    TRT01P == c("Placebo"), TRUE, TRUE, RACE == c("BLACK OR AFRICAN AMERICAN")
```

Extracting Metadata


```
get_meta_subset(t, 'c2_3', 'Placebo', add_cols=vars(USUBJID))
```

```
# A tibble: 8 × 3
  USUBJID     TRT01P  RACE
  <chr>       <chr>   <chr>
1 01-701-1203 Placebo BLACK OR AFRICAN AMERICAN
2 01-701-1363 Placebo BLACK OR AFRICAN AMERICAN
3 01-705-1282 Placebo BLACK OR AFRICAN AMERICAN
4 01-706-1041 Placebo BLACK OR AFRICAN AMERICAN
5 01-708-1286 Placebo BLACK OR AFRICAN AMERICAN
6 01-708-1296 Placebo BLACK OR AFRICAN AMERICAN
7 01-708-1378 Placebo BLACK OR AFRICAN AMERICAN
8 01-711-1036 Placebo BLACK OR AFRICAN AMERICAN
```

Using Tplyr's Metadata



Using Tplyr's Metadata

App: https://tinyurl.com/Tplyr-shiny-demo

Code: https://tinyurl.com/Tplyr-shiny-code

Expanding Tplyr's Metadata

PharmaSUG
NORTH CAROLINA • OCT. 21, 2022

Function	Description
`tplyr_meta()`	Create a tplyr_meta object
`add_variables()`	Add variables to a tplyr_meta object
`add_filters()`	Add filters to a tplyr_meta object
`append_metadata()`	Append a Tplyr table's metadata object

Expanding Tplyr's Metadata

PharmaSUG
NORTH CAROLINA • OCT. 21, 2022

App: https://tinyurl.com/Tplyr-efficacy-demo

Code: https://tinyurl.com/Tplyr-eff-shiny-code

atorus
Your Complete Data Solution

Questions?

PharmaSUG
NORTH CAROLINA • OCT. 21, 2022

atorus
Your Complete Data Solution