USING THE table1 PACKAGE IN R

Introduction: When you speak to a colleague about their "Table 1" it is typically understood that you are speaking about a table that describes the relevant characteristics of the sample being studied. Often, but not always, the sample being studied is made up of people, and the relevant descriptive characteristics about those people include sociodemographic and/or general health information. Therefore, it is important that you don't label any of your tables as "Table 1" arbitrarily. Unless you have a *really* good reason to do otherwise, your Table 1 should always be a descriptive overview of your sample.

Here is a list of other traits that should consider when creating your Table 1:

- All other formatting best practices that apply to scientific tables in general. This includes formatting requirements specific to wherever you are submitting your table (e.g., formatting requirements in the American Journal of Public Health).
- Table 1 is often, but not always, stratified into subgroups (i.e., descriptive results are presented separately for each subgroup of the study sample in a way that lends itself to between-group comparisons).
- When Table 1 is stratified into subgroups, the variable that contains the subgroups is typically the primary exposure/predictor of interest in your study.

Reference: *https://www.r4epi.com/creating-tables-with-r-and-microsoft-word.html*

Example: We have a dataset called data1, which consists of 146 observations (patients) and 5 variables (id, treat, age, sex and wt). Participants have been randomized to receive an active treatment or placebo. We will only consider three baseline characteristics: age, sex and weight.

head(data1) data set	<i>#prints the first few rows of the</i>					
	id treat	age	sex	wt		
	1 Treated	18	Female	62.6		
	2 Treated	50	Male	57.4		
	3 Treated	37	Male	104.6		
	4 Treated	25	Female	55.5		
	5 Placebo	60	Female	58.4		
	6 Treated	44	Female	41.9		

The easiest package in R to use to generate Table 1 is the **table1 package**. To install:

install.packages("table1") library(table1) Once the package is installed, it is simple to create Table 1 using one line of code:

```
table1(~age + sex + wt | treat, data=data1)
```

We are specifying that treat is our exposure variable of interest, and the variable we wish to stratify by. Age, sex and wt are the variables we would like the describe for the patients in our data set.

	Placebo (N=52)	Treated (N=94)	Overall (N=146)
Age (years)	-		
Mean (SD)	39.2 (14.2)	40.1 (13.3)	39.8 (13.6)
Median [Min, Max]	37.5 [18.0, 65.0]	39.5 [18.0, 65.0]	39.0 [18.0, 65.0]
Sex			
Female	34 (65.4%)	53 (56.4%)	87 (59.6%)
Male	18 (34.6%)	41 (43.6%)	59 (40.4%)
Weight (kg)			
Mean (SD)	68.1 (16.3)	68.3 (16.7)	68.2 (16.5)
Median [Min, Max]	66.7 [37.5, 116]	64.9 [40.0, 119]	66.2 [37.5, 119]
Missing	2 (3.8%)	3 (3.2%)	5 (3.4%)

The table generated with this line of code is below:

Note: In this simple form, table1 will generate both the mean (SD) and median (IQR) for continuous variables. You generally do not need to present both. Usually, you will pick one to present based on which summary is most appropriate (see HOW TO SUMMARIZE DATA). To do this here, you can modify the code as detailed below:

```
rndr <- function(x, name, ...) {
    if (!is.numeric(x)) return(render.categorical.default(x))
    what <- switch(name,
        age = "Median [Min, Max]",
        wt = "Mean (SD)")
    parse.abbrev.render.code(c("", what))(x)
}
table1(~ age + sex + wt | treat, data=data1, render=rndr)</pre>
```

	Placebo (N=52)	Treated (N=94)	Overall (N=146)
Age (years)			
Median [Min, Max]	37.5 [18.0, 65.0]	39.5 [18.0, 65.0]	39.0 [18.0, 65.0]
Sex			
Female	34 (65.4%)	53 (56.4%)	87 (59.6%)
Male	18 (34.6%)	41 (43.6%)	59 (40.4%)
Weight (kg)			
Mean (SD)	68.1 (16.3)	68.3 (16.7)	68.2 (16.5)

To learn more about the table1 package type:

help(table1)

Reference: https://cran.rproject.org/web/packages/table1/vignettes/table1-examples.html