Exploring the R gtsummary Package to Create Professional-Quality Descriptive Tables for Academic Publications

Weisi Chen

2025-02-09

Contents

### **Install and read in R packages needed**

library(NHANES)
library(gtsummary)
library(gt)
library(dplyr)
library(purrr)

### **Read in the demo data**

data <- NHANES::NHANES

### **Example basic table**

Table 1: Table 1: Sociodemographic Characteristics of Patients With and Without Diabetes in the Demo Dataset

| **Characteristic** | **Overall** N = 9,858*1* | **No** N = 9,098*1* | **Yes** N = 760*1* | **p-value***2* |
| --- | --- | --- | --- | --- |
| **Gender** |  |  |  | 0.064 |
|     female | 4,949 (50%) | 4,592 (50%) | 357 (47%) |  |
|     male | 4,909 (50%) | 4,506 (50%) | 403 (53%) |  |
| **Age** | 37 (22) | 35 (22) | 59 (15) | <0.001 |
| **Age group** |  |  |  | <0.001 |
|      0-9 | 1,254 (13%) | 1,254 (14%) | 0 (0%) |  |
|      10-19 | 1,371 (14%) | 1,354 (15%) | 17 (2.5%) |  |
|      20-29 | 1,356 (14%) | 1,344 (15%) | 12 (1.7%) |  |
|      30-39 | 1,338 (14%) | 1,295 (15%) | 43 (6.2%) |  |
|      40-49 | 1,398 (15%) | 1,302 (15%) | 96 (14%) |  |
|      50-59 | 1,304 (14%) | 1,126 (13%) | 178 (26%) |  |
|      60-69 | 917 (9.6%) | 713 (8.1%) | 204 (30%) |  |
|      70+ | 587 (6.2%) | 447 (5.1%) | 140 (20%) |  |
|     Unknown | 333 | 263 | 70 |  |
| **Ethnicity** |  |  |  | <0.001 |
|     Black | 1,184 (12%) | 1,053 (12%) | 131 (17%) |  |
|     Hispanic | 602 (6.1%) | 555 (6.1%) | 47 (6.2%) |  |
|     Mexican | 991 (10%) | 925 (10%) | 66 (8.7%) |  |
|     White | 6,290 (64%) | 5,840 (64%) | 450 (59%) |  |
|     Other | 791 (8.0%) | 725 (8.0%) | 66 (8.7%) |  |
| **BMI group** |  |  |  | <0.001 |
|     12.0\_18.5 | 1,277 (13%) | 1,274 (14%) | 3 (0.4%) |  |
|     18.5\_to\_24.9 | 2,908 (30%) | 2,797 (32%) | 111 (15%) |  |
|     25.0\_to\_29.9 | 2,664 (28%) | 2,461 (28%) | 203 (27%) |  |
|     30.0\_plus | 2,749 (29%) | 2,321 (26%) | 428 (57%) |  |
|     Unknown | 260 | 245 | 15 |  |
| **Education** |  |  |  | <0.001 |
|     8th Grade | 451 (6.2%) | 351 (5.4%) | 100 (13%) |  |
|     9 - 11th Grade | 886 (12%) | 781 (12%) | 105 (14%) |  |
|     High School | 1,517 (21%) | 1,352 (21%) | 165 (22%) |  |
|     Some College | 2,267 (31%) | 2,039 (31%) | 228 (31%) |  |
|     College Grad | 2,098 (29%) | 1,954 (30%) | 144 (19%) |  |
|     Unknown | 2,639 | 2,621 | 18 |  |
| **MaritalStatus** |  |  |  | <0.001 |
|     Divorced | 705 (9.8%) | 605 (9.3%) | 100 (13%) |  |
|     LivePartner | 560 (7.7%) | 531 (8.2%) | 29 (3.9%) |  |
|     Married | 3,945 (55%) | 3,519 (54%) | 426 (57%) |  |
|     NeverMarried | 1,380 (19%) | 1,313 (20%) | 67 (9.0%) |  |
|     Separated | 183 (2.5%) | 159 (2.5%) | 24 (3.2%) |  |
|     Widowed | 456 (6.3%) | 361 (5.6%) | 95 (13%) |  |
|     Unknown | 2,629 | 2,610 | 19 |  |
| **Household income** |  |  |  | <0.001 |
|      0-4999 | 182 (2.0%) | 169 (2.0%) | 13 (1.9%) |  |
|      5000-9999 | 250 (2.8%) | 223 (2.7%) | 27 (3.9%) |  |
|     10000-14999 | 537 (5.9%) | 472 (5.6%) | 65 (9.3%) |  |
|     15000-19999 | 515 (5.7%) | 461 (5.5%) | 54 (7.8%) |  |
|     20000-24999 | 605 (6.7%) | 546 (6.5%) | 59 (8.5%) |  |
|     25000-34999 | 945 (10%) | 865 (10%) | 80 (11%) |  |
|     35000-44999 | 851 (9.4%) | 765 (9.1%) | 86 (12%) |  |
|     45000-54999 | 776 (8.6%) | 726 (8.7%) | 50 (7.2%) |  |
|     55000-64999 | 615 (6.8%) | 571 (6.8%) | 44 (6.3%) |  |
|     65000-74999 | 521 (5.7%) | 476 (5.7%) | 45 (6.5%) |  |
|     75000-99999 | 1,066 (12%) | 999 (12%) | 67 (9.6%) |  |
|     more 99999 | 2,200 (24%) | 2,094 (25%) | 106 (15%) |  |
|     Unknown | 795 | 731 | 64 |  |
| **Employment status** |  |  |  | <0.001 |
|     Looking | 311 (4.0%) | 288 (4.1%) | 23 (3.1%) |  |
|     NotWorking | 2,845 (37%) | 2,418 (34%) | 427 (57%) |  |
|     Working | 4,613 (59%) | 4,319 (61%) | 294 (40%) |  |
|     Unknown | 2,089 | 2,073 | 16 |  |
| *1*n (%); Mean (SD) |
| *2*Pearson's Chi-squared test; Wilcoxon rank sum test |

### **Customize the table’s appearance**

* **Move the total column** to the far-right end of the table for improved readability.
* **Remove the ‘N = xxxx’** from the header to streamline the table’s appearance.
* **Add a “Total (denominator)” row** at the top of the table for better context and clarity.
* **Avoid decimal places** for both numbers and percentages for a cleaner presentation.
* **Include additional summary statistics** for continuous variables, such as mean (SD), median (IQR), and range, to provide a more comprehensive summary.
* **Customize the footnotes**

Table 1: **Table 1: Sociodemographic Characteristics of Patients With and Without Diabetes in the Demo Dataset**

|  | **With Diabetes** | **Without Diabetes** | **Total** | **p-value***1* |
| --- | --- | --- | --- | --- |
| **Total (column denominator)** | 760 (100%) | 9,098 (100%) | 9,858 (100%) |  |
| **Gender*2*** |  |  |  | 0.064 |
|     female | 357 (47%) | 4,592 (50%) | 4,949 (50%) |  |
|     male | 403 (53%) | 4,506 (50%) | 4,909 (50%) |  |
| **Age*3*** |  |  |  | <0.001 |
|     Mean, (SD) | 59, (15) | 35, (22) | 37, (22) |  |
|     Median, (IQR) | 61, (51, 70) | 34, (17, 52) | 37, (18, 54) |  |
|     Range | 11, 80 | 1, 80 | 1, 80 |  |
| **Age group** |  |  |  | <0.001 |
|      0-9 | 0 (0%) | 1,254 (14%) | 1,254 (13%) |  |
|      10-19 | 17 (2%) | 1,354 (15%) | 1,371 (14%) |  |
|      20-29 | 12 (2%) | 1,344 (15%) | 1,356 (14%) |  |
|      30-39 | 43 (6%) | 1,295 (15%) | 1,338 (14%) |  |
|      40-49 | 96 (14%) | 1,302 (15%) | 1,398 (15%) |  |
|      50-59 | 178 (26%) | 1,126 (13%) | 1,304 (14%) |  |
|      60-69 | 204 (30%) | 713 (8%) | 917 (10%) |  |
|      70+ | 140 (20%) | 447 (5%) | 587 (6%) |  |
| **Ethnicity** |  |  |  | <0.001 |
|     Black | 131 (17%) | 1,053 (12%) | 1,184 (12%) |  |
|     Hispanic | 47 (6%) | 555 (6%) | 602 (6%) |  |
|     Mexican | 66 (9%) | 925 (10%) | 991 (10%) |  |
|     White | 450 (59%) | 5,840 (64%) | 6,290 (64%) |  |
|     Other | 66 (9%) | 725 (8%) | 791 (8%) |  |
| **BMI group** |  |  |  | <0.001 |
|     12.0\_18.5 | 3 (0%) | 1,274 (14%) | 1,277 (13%) |  |
|     18.5\_to\_24.9 | 111 (15%) | 2,797 (32%) | 2,908 (30%) |  |
|     25.0\_to\_29.9 | 203 (27%) | 2,461 (28%) | 2,664 (28%) |  |
|     30.0\_plus | 428 (57%) | 2,321 (26%) | 2,749 (29%) |  |
| **Education** |  |  |  | <0.001 |
|     8th Grade | 100 (13%) | 351 (5%) | 451 (6%) |  |
|     9 - 11th Grade | 105 (14%) | 781 (12%) | 886 (12%) |  |
|     High School | 165 (22%) | 1,352 (21%) | 1,517 (21%) |  |
|     Some College | 228 (31%) | 2,039 (31%) | 2,267 (31%) |  |
|     College Grad | 144 (19%) | 1,954 (30%) | 2,098 (29%) |  |
| **MaritalStatus** |  |  |  | <0.001 |
|     Divorced | 100 (13%) | 605 (9%) | 705 (10%) |  |
|     LivePartner | 29 (4%) | 531 (8%) | 560 (8%) |  |
|     Married | 426 (57%) | 3,519 (54%) | 3,945 (55%) |  |
|     NeverMarried | 67 (9%) | 1,313 (20%) | 1,380 (19%) |  |
|     Separated | 24 (3%) | 159 (2%) | 183 (3%) |  |
|     Widowed | 95 (13%) | 361 (6%) | 456 (6%) |  |
| **Household income** |  |  |  | <0.001 |
|      0-4999 | 13 (2%) | 169 (2%) | 182 (2%) |  |
|      5000-9999 | 27 (4%) | 223 (3%) | 250 (3%) |  |
|     10000-14999 | 65 (9%) | 472 (6%) | 537 (6%) |  |
|     15000-19999 | 54 (8%) | 461 (6%) | 515 (6%) |  |
|     20000-24999 | 59 (8%) | 546 (7%) | 605 (7%) |  |
|     25000-34999 | 80 (11%) | 865 (10%) | 945 (10%) |  |
|     35000-44999 | 86 (12%) | 765 (9%) | 851 (9%) |  |
|     45000-54999 | 50 (7%) | 726 (9%) | 776 (9%) |  |
|     55000-64999 | 44 (6%) | 571 (7%) | 615 (7%) |  |
|     65000-74999 | 45 (6%) | 476 (6%) | 521 (6%) |  |
|     75000-99999 | 67 (10%) | 999 (12%) | 1,066 (12%) |  |
|     more 99999 | 106 (15%) | 2,094 (25%) | 2,200 (24%) |  |
| **Employment status** |  |  |  | <0.001 |
|     Looking | 23 (3%) | 288 (4%) | 311 (4%) |  |
|     NotWorking | 427 (57%) | 2,418 (34%) | 2,845 (37%) |  |
|     Working | 294 (40%) | 4,319 (61%) | 4,613 (59%) |  |
| *1*Pearson's Chi-squared test; Wilcoxon rank sum test |
| *2*This is a sample footnote 1. |
| *3*This is a sample footnote 2. |

### **Customize the table’s appearance II**

* **Separate the Number and Percentage Columns**: Split the n (count) and p (percentage) values into two separate columns in the table.
* **Right-align the Number and Percentage Columns**: Apply cell\_text(align = “right”) to these columns.
* **Label the Columns as n and col%**
* **Add Colors**: Apply cell\_fill() for background colors and/or cell\_text() for text colors to enhance readability.

Table 1: **Table 1: Sociodemographic Characteristics of Patients With and Without Diabetes in the Demo Dataset**

|  | **With Diabetes** | **Without Diabetes** | **Total** |
| --- | --- | --- | --- |
|  | **n** | **col%** | **n** | **col%** | **n** | **col%** |
| **Total (column denominator)** | 760 | (100%) | 9,098 | (100%) | 9,858 | (100%) |
| **Gender*1*** |  |  |  |  |  |  |
|     female | 357 | (47%) | 4,592 | (50%) | 4,949 | (50%) |
|     male | 403 | (53%) | 4,506 | (50%) | 4,909 | (50%) |
| **Age*2*** |  |  |  |  |  |  |
|     Mean (SD) | 59 (15) |  | 35 (22) |  | 37 (22) |  |
|     Median (Q1, Q3) | 61 (51, 70) |  | 34 (17, 52) |  | 37 (18, 54) |  |
|     Range | 11, 80 |  | 1, 80 |  | 1, 80 |  |
| **Age group** |  |  |  |  |  |  |
|      0-9 | 0 | (0%) | 1,254 | (14%) | 1,254 | (13%) |
|      10-19 | 17 | (2%) | 1,354 | (15%) | 1,371 | (14%) |
|      20-29 | 12 | (2%) | 1,344 | (15%) | 1,356 | (14%) |
|      30-39 | 43 | (6%) | 1,295 | (15%) | 1,338 | (14%) |
|      40-49 | 96 | (14%) | 1,302 | (15%) | 1,398 | (15%) |
|      50-59 | 178 | (26%) | 1,126 | (13%) | 1,304 | (14%) |
|      60-69 | 204 | (30%) | 713 | (8%) | 917 | (10%) |
|      70+ | 140 | (20%) | 447 | (5%) | 587 | (6%) |
| **Ethnicity** |  |  |  |  |  |  |
|     Black | 131 | (17%) | 1,053 | (12%) | 1,184 | (12%) |
|     Hispanic | 47 | (6%) | 555 | (6%) | 602 | (6%) |
|     Mexican | 66 | (9%) | 925 | (10%) | 991 | (10%) |
|     White | 450 | (59%) | 5,840 | (64%) | 6,290 | (64%) |
|     Other | 66 | (9%) | 725 | (8%) | 791 | (8%) |
| **BMI group** |  |  |  |  |  |  |
|     12.0\_18.5 | 3 | (0%) | 1,274 | (14%) | 1,277 | (13%) |
|     18.5\_to\_24.9 | 111 | (15%) | 2,797 | (32%) | 2,908 | (30%) |
|     25.0\_to\_29.9 | 203 | (27%) | 2,461 | (28%) | 2,664 | (28%) |
|     30.0\_plus | 428 | (57%) | 2,321 | (26%) | 2,749 | (29%) |
| **Education** |  |  |  |  |  |  |
|     8th Grade | 100 | (13%) | 351 | (5%) | 451 | (6%) |
|     9 - 11th Grade | 105 | (14%) | 781 | (12%) | 886 | (12%) |
|     High School | 165 | (22%) | 1,352 | (21%) | 1,517 | (21%) |
|     Some College | 228 | (31%) | 2,039 | (31%) | 2,267 | (31%) |
|     College Grad | 144 | (19%) | 1,954 | (30%) | 2,098 | (29%) |
| **MaritalStatus** |  |  |  |  |  |  |
|     Divorced | 100 | (13%) | 605 | (9%) | 705 | (10%) |
|     LivePartner | 29 | (4%) | 531 | (8%) | 560 | (8%) |
|     Married | 426 | (57%) | 3,519 | (54%) | 3,945 | (55%) |
|     NeverMarried | 67 | (9%) | 1,313 | (20%) | 1,380 | (19%) |
|     Separated | 24 | (3%) | 159 | (2%) | 183 | (3%) |
|     Widowed | 95 | (13%) | 361 | (6%) | 456 | (6%) |
| **Household income** |  |  |  |  |  |  |
|      0-4999 | 13 | (2%) | 169 | (2%) | 182 | (2%) |
|      5000-9999 | 27 | (4%) | 223 | (3%) | 250 | (3%) |
|     10000-14999 | 65 | (9%) | 472 | (6%) | 537 | (6%) |
|     15000-19999 | 54 | (8%) | 461 | (6%) | 515 | (6%) |
|     20000-24999 | 59 | (8%) | 546 | (7%) | 605 | (7%) |
|     25000-34999 | 80 | (11%) | 865 | (10%) | 945 | (10%) |
|     35000-44999 | 86 | (12%) | 765 | (9%) | 851 | (9%) |
|     45000-54999 | 50 | (7%) | 726 | (9%) | 776 | (9%) |
|     55000-64999 | 44 | (6%) | 571 | (7%) | 615 | (7%) |
|     65000-74999 | 45 | (6%) | 476 | (6%) | 521 | (6%) |
|     75000-99999 | 67 | (10%) | 999 | (12%) | 1,066 | (12%) |
|     more 99999 | 106 | (15%) | 2,094 | (25%) | 2,200 | (24%) |
| **Employment status** |  |  |  |  |  |  |
|     Looking | 23 | (3%) | 288 | (4%) | 311 | (4%) |
|     NotWorking | 427 | (57%) | 2,418 | (34%) | 2,845 | (37%) |
|     Working | 294 | (40%) | 4,319 | (61%) | 4,613 | (59%) |
| *1*This is a sample footnote 1. |
| *2*This is a sample footnote 2. |

## **Code Available**

### Code to produce the example basic table

data %>%
 # Remove missing data in the Diabetes variable for simplicity
 filter(!is.na(Diabetes)) %>%
 # Select relevant variables
 select(Gender, Age, AgeDecade, Race1, BMI\_WHO, Education, MaritalStatus, HHIncome, Work, Diabetes) %>%
 # Create a summary table by Diabetes group
 tbl\_summary(
 by = Diabetes,
 statistic = list(
 all\_continuous() ~ "{mean} ({sd})",
 all\_categorical() ~ "{n} ({p}%)"
 ),
 label = list(
 AgeDecade = "Age group",
 Race1 = "Ethnicity",
 BMI\_WHO = "BMI group",
 HHIncome = "Household income",
 Work = "Employment status"
 )
 ) %>%
 add\_overall() %>%
 add\_p() %>% # Test for differences between groups
 bold\_labels() %>%
 modify\_header(label = "\*\*Characteristic\*\*") %>% # Update column header
 as\_gt() %>%
 gt::tab\_header(
 "Table 1: Sociodemographic Characteristics of Patients With and Without Diabetes in the Demo Dataset")

### Code to produce the customized table I

data %>%
 # Remove missing data in the Diabetes variable for simplicity
 filter(!is.na(Diabetes)) %>%

 # Format the Diabetes variable
 mutate(
 Diabetes = case\_when(
 Diabetes == "Yes" ~ "With Diabetes",
 Diabetes == "No" ~ "Without Diabetes"
 ),
 Diabetes = factor(Diabetes, levels = c("With Diabetes", "Without Diabetes"))
 ) %>%

 # Add total number
 mutate(total = TRUE) %>%

 # Select relevant variables
 select(
 total, Gender, Age, AgeDecade, Race1, BMI\_WHO, Education,
 MaritalStatus, HHIncome, Work, Diabetes
 ) %>%

 # Create a summary table by Diabetes group
 tbl\_summary(
 by = Diabetes,
 type = all\_continuous() ~ "continuous2",
 statistic = list(
 # Include additional summary statistics for continuous variables
 all\_continuous() ~ c("{mean}, ({sd})",
 "{median}, ({p25}, {p75})",
 "{min}, {max}"),
 all\_categorical() ~ "{n} ({p}%)"
 ),
 label = list(
 total = "Total (column denominator)",
 AgeDecade = "Age group",
 Race1 = "Ethnicity",
 BMI\_WHO = "BMI group",
 HHIncome = "Household income",
 Work = "Employment status"
 ),
 missing = "no",

 # Remove decimal places for all numbers and percentages
 digits = list(
 all\_continuous() ~ c(0, 0),
 all\_categorical() ~ c(0, 0)
 )
 ) %>%

 # Add total column
 add\_overall() %>%

 # Move the total column to the far end of the table
 modify\_table\_body(
 ~ .x %>%
 dplyr::relocate(stat\_0, .after = stat\_2) %>%

 # Change label name
 dplyr::mutate(
 label = ifelse(label == "Median, (Q1, Q3)", "Median, (IQR)", label)
 ) %>%
 dplyr::mutate(
 label = ifelse(label == "Min, Max", "Range", label)
 )
 ) %>%

 # Modify the header
 modify\_header(
 update = list(
 all\_stat\_cols(TRUE) ~ "\*\*{level}\*\*",
 label = "",
 stat\_0 = "\*\*Total\*\*",
 stat\_1 = "\*\*{level}\*\*",
 stat\_2 = "\*\*{level}\*\*"
 )
 ) %>%

 # Test for differences between groups
 add\_p() %>%

 # Bold labels for readability
 bold\_labels() %>%

 # Modify footnotes
 modify\_footnote(
 c(all\_stat\_cols()) ~ NA
 ) %>%

 # Add more footnotes to specific rows
 modify\_table\_styling(
 columns = label,
 row = label == list("Gender"),
 footnote = "This is a sample footnote 1."
 ) %>%
 modify\_table\_styling(
 columns = label,
 row = label == list("Age"),
 footnote = "This is a sample footnote 2."
 ) %>%

 # Convert to gt table
 as\_gt() %>%

 # Add table header with title
 gt::tab\_header(
 title = md("\*\*Table 1: Sociodemographic Characteristics of Patients With and Without Diabetes in the Demo Dataset\*\*")
 ) %>%

 # Prevent footnotes from being split across multiple lines
 tab\_options(footnotes.multiline = FALSE)

### Code to produce the customized table II

tab <- c("{n}", "({p}%)") %>%
 map(
 ~data %>%
 # Remove missing data in the Diabetes variable for simplicity
 filter(!is.na(Diabetes)) %>%

 # Format the Diabetes variable
 mutate(
 Diabetes = case\_when(
 Diabetes == "Yes" ~ "With Diabetes",
 Diabetes == "No" ~ "Without Diabetes"
 ),
 Diabetes = factor(Diabetes, levels = c("With Diabetes", "Without Diabetes"))
 ) %>%

 # Add total number
 mutate(total = TRUE) %>%

 # Select relevant variables
 select(
 total, Gender, Age, AgeDecade, Race1, BMI\_WHO, Education,
 MaritalStatus, HHIncome, Work, Diabetes
 ) %>%

 # Create a summary table by Diabetes group
 tbl\_summary(
 by = Diabetes,
 type = all\_continuous() ~ "continuous2",
 statistic = list(
 # Include additional summary statistics for continuous variables
 all\_continuous() ~ c("{mean} ({sd})",
 "{median} ({p25}, {p75})",
 "{min}, {max}"),
 all\_categorical() ~ .x
 ),
 label = list(
 total = "Total (column denominator)",
 AgeDecade = "Age group",
 Race1 = "Ethnicity",
 BMI\_WHO = "BMI group",
 HHIncome = "Household income",
 Work = "Employment status"
 ),
 missing = "no",

 # Remove decimal places for all numbers and percentages
 digits = list(
 all\_continuous() ~ c(0, 0),
 all\_categorical() ~ c(0, 0)
 )
 ) %>%

 # Add total column
 add\_overall() %>%

 # Bold labels for readability
 bold\_labels()) %>%
 tbl\_merge() %>%
 modify\_spanning\_header(everything()~NA) %>%

 # Re-arrange the number and percentage columns
 modify\_table\_body(
 ~ .x %>%
 dplyr::relocate(stat\_1\_2, .after=stat\_1\_1) %>%
 dplyr::relocate(stat\_2\_2, .after=stat\_2\_1) %>%
 dplyr::relocate(stat\_0\_1, .after=stat\_2\_2) %>%
 dplyr::relocate(stat\_0\_2, .after=stat\_0\_1)
 %>%
 # Change label name
 dplyr::mutate(
 label = ifelse(label == "Median, (Q1, Q3)", "Median, (IQR)", label)
 ) %>%
 dplyr::mutate(
 label = ifelse(label == "Min, Max", "Range", label)
 ) %>%

 # Remove the summary statistics for the continuous variable in the % column
 dplyr::mutate(
 stat\_0\_2 = ifelse(label == "Mean (SD)", "",stat\_0\_2 ),
 stat\_0\_2 = ifelse(label == "Median (Q1, Q3)", "",stat\_0\_2 ),
 stat\_0\_2 = ifelse(label == "Range", "",stat\_0\_2 ),
 stat\_1\_2 = ifelse(label == "Mean (SD)", "",stat\_1\_2 ),
 stat\_1\_2 = ifelse(label == "Median (Q1, Q3)", "",stat\_1\_2 ),
 stat\_1\_2 = ifelse(label == "Range", "",stat\_1\_2 ),
 stat\_2\_2 = ifelse(label == "Mean (SD)", "",stat\_2\_2 ),
 stat\_2\_2 = ifelse(label == "Median (Q1, Q3)", "",stat\_2\_2 ),
 stat\_2\_2 = ifelse(label == "Range", "",stat\_2\_2 ),
 )
 ) %>%

 # Modify the header
 modify\_header(
 update = list(
 all\_stat\_cols(TRUE) ~ "\*\*{level}\*\*",
 label = "",
 stat\_0\_1 = "\*\*n\*\*",
 stat\_0\_2 = "\*\*col%\*\*",
 stat\_1\_1 = "\*\*n\*\*",
 stat\_1\_2 = "\*\*col%\*\*",
 stat\_2\_1 = "\*\*n\*\*",
 stat\_2\_2 = "\*\*col%\*\*"
 )
 ) %>%

 # Modify spanning header
 modify\_spanning\_header(
 stat\_1\_1:stat\_1\_2 ~ "\*\*With Diabetes\*\*",
 stat\_2\_1:stat\_2\_2 ~ "\*\*Without Diabetes\*\*",
 stat\_0\_1:stat\_0\_2 ~ "\*\*Total\*\*",
 ) %>%

 # Modify footnotes
 modify\_footnote(
 c(all\_stat\_cols()) ~ NA
 ) %>%

 # Add more footnotes to specific rows
 modify\_table\_styling(
 columns = label,
 row = label == list("Gender"),
 footnote = "This is a sample footnote 1."
 ) %>%
 modify\_table\_styling(
 columns = label,
 row = label == list("Age"),
 footnote = "This is a sample footnote 2."
 ) %>%

 # Convert to gt table
 as\_gt() %>%

 # Add table header with title
 gt::tab\_header(
 title = md("\*\*Table 1: Sociodemographic Characteristics of Patients With and Without Diabetes in the Demo Dataset\*\*")
 ) %>%

 # Prevent footnotes from being split across multiple lines
 tab\_options(footnotes.multiline = FALSE) %>%

 # Right-align all columns except the label column
 tab\_style(
 style = cell\_text(align = "center"),
 locations = cells\_column\_labels(
 columns = everything()
 )
 ) %>%
 tab\_style(
 style = cell\_text(align = "right"),
 locations = cells\_body(
 columns = !label
 )
 )

# Adding some colors to the tables
tab %>%
 tab\_style(
 style = cell\_fill(color = "#E8E4E6"), # Apply the background color
 locations = cells\_body(
 rows = seq(2, nrow(tab$`\_data`), by = 2) # Select every second row (alternating)
 )
 ) %>%
 tab\_style(
 style = cell\_fill(color = "#DAE9F7"),
 locations = cells\_column\_labels()
 ) %>%
 tab\_style(
 style = cell\_fill(color = "#DAE9F7"),
 locations = cells\_column\_spanners()
 )