

# P8105 Homework I

2022-09-27

```
# read and name data
data("penguins", package = "palmerpenguins")

# load the packages necessary for submission to knit.
library(tidyverse)
```

## Problem 1

Here's a **code chunk** that creates an overall summary table, including the names of all the variables:

```
## create overall summary table
library(gtsummary)
penguins %>%
  tbl_summary(missing_text = "(Missing)", # counts missing values
              statistic = list(all_continuous() ~ "{mean} ({sd})"), # what to do with continuous variables
              label = list(sex ~ "Sex", # label variables
                            species ~ "Species",
                            island ~ "Island",
                            year ~ "Year",
                            body_mass_g ~ "Body Mass",
                            flipper_length_mm ~ "Flipper Length",
                            bill_depth_mm ~ "Bill Depth",
                            bill_length_mm ~ "Bill Length" )) %>%

  bold_labels() %>%
  italicize_levels()
```

Characteristic	N = 344
<b>Species</b>	
<i>Adelie</i>	152 (44%)
<i>Chinstrap</i>	68 (20%)
<i>Gentoo</i>	124 (36%)
<b>Island</b>	
<i>Biscoe</i>	168 (49%)
<i>Dream</i>	124 (36%)
<i>Torgersen</i>	52 (15%)
<b>Bill Length</b>	43.9 (5.5)
<i>(Missing)</i>	2
<b>Bill Depth</b>	17.15 (1.97)
<i>(Missing)</i>	2
<b>Flipper Length</b>	201 (14)
<i>(Missing)</i>	2
<b>Body Mass</b>	4,202 (802)
<i>(Missing)</i>	2
<b>Sex</b>	
<i>female</i>	165 (50%)

Characteristic	N = 344
<i>male</i>	168 (50%)
<i>(Missing)</i>	11
<b>Year</b>	
<i>2007</i>	110 (32%)
<i>2008</i>	114 (33%)
<i>2009</i>	120 (35%)

We now want to create a summary table by species:

```
# create summary table by year
penguins %>%
  tbl_summary(by = "species", # stratify by species
              missing_text = "(Missing)",
              statistic = list(all_continuous() ~ "{mean} ({sd})"),
              label = list(sex ~ "Sex", # label variables
                           species ~ "Species",
                           island ~ "Island",
                           body_mass_g ~ "Body Mass",
                           flipper_length_mm ~ "Flipper Length",
                           bill_depth_mm ~ "Bill Depth",
                           bill_length_mm ~ "Bill Length" )) %>%
  bold_labels() %>%
  italicize_levels()
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	Adelie, N = 152	Chinstrap, N = 68	Gentoo, N = 124
<b>Island</b>			
<i>Biscoe</i>	44 (29%)	0 (0%)	124 (100%)
<i>Dream</i>	56 (37%)	68 (100%)	0 (0%)
<i>Torgersen</i>	52 (34%)	0 (0%)	0 (0%)
<b>Bill Length</b>	38.8 (2.7)	48.8 (3.3)	47.5 (3.1)
<i>(Missing)</i>	1	0	1
<b>Bill Depth</b>	18.35 (1.22)	18.42 (1.14)	14.98 (0.98)
<i>(Missing)</i>	1	0	1
<b>Flipper Length</b>	190 (7)	196 (7)	217 (6)
<i>(Missing)</i>	1	0	1
<b>Body Mass</b>	3,701 (459)	3,733 (384)	5,076 (504)
<i>(Missing)</i>	1	0	1
<b>Sex</b>			
<i>female</i>	73 (50%)	34 (50%)	58 (49%)
<i>male</i>	73 (50%)	34 (50%)	61 (51%)
<i>(Missing)</i>	6	0	5
<b>year</b>			
<i>2007</i>	50 (33%)	26 (38%)	34 (27%)
<i>2008</i>	50 (33%)	18 (26%)	46 (37%)
<i>2009</i>	52 (34%)	24 (35%)	44 (35%)

## Problem 2