# pengruin bill length statistic analysis

```r
data("penguins")

bill <-
  penguins |>
  as_tibble() |>
  clean_names() |>
  mutate(state = rownames(penguins)) |>
  select(species, island, bill_length_mm, bill_depth_mm)

bill |>
  summary()
```

```
    species            island      bill_length_mm   bill_depth_mm
 Adelie   :152    Biscoe   :168   Min.   :32.10    Min.   :13.10
 Chinstrap: 68    Dream    :124   1st Qu.:39.23    1st Qu.:15.60
 Gentoo   :124    Torgersen: 52   Median :44.45    Median :17.30
                                  Mean   :43.92    Mean   :17.15
                                  3rd Qu.:48.50    3rd Qu.:18.70
                                  Max.   :59.60    Max.   :21.50
                                  NA's   :2        NA's   :2
```

```r
sample_means <- tibble(seed = c(), mean_bill_length = c(), species_ignored = c())

# calculate mean bill length excluding random species
for (i in c(1:5)) {
  set.seed(i)
  dont_get <- c(sample(x = bill$species, size = 1))
  # Add the results to the sample_means tibble
  mean_bill_length <-
    bill |>
      filter(!species %in% dont_get) |>
```

```
      summarise(mean_bill_length = mean(bill_length_mm, na.rm = TRUE)) |>
      pull(mean_bill_length)
  sample_means <- sample_means |>
    rbind(tibble(
      seed = i,
      mean_bill_length = mean_bill_length,
      species_ignored = str_c(dont_get, collapse = ", ")
    ))
}

sample_means |>
  kable(
    col.names = c("Seed", "Mean Bill Length (mm)", "Species Ignored"),
    digits = 2,
    format.args = list(big.mark = ","),
    booktabs = TRUE
  )
```

| Seed | Mean Bill Length (mm) | Species Ignored |
|------|----------------------:|-----------------|
| 1    | 42.70                 | Chinstrap       |
| 2    | 42.70                 | Chinstrap       |
| 3    | 41.91                 | Gentoo          |
| 4    | 47.98                 | Adelie          |
| 5    | 42.70                 | Chinstrap       |