lab2_EMF

Elizabeth Miles-Flynn

2026-01-27

Setup

```r
knitr::opts_chunk$set(echo = TRUE)

# Load necessary packages
library(here)
library(tidyverse)
library(viridis)
library("lterdatasampler")

# Rename data
crab_data <- pie_crab
```

Measures of Location and Data Spread

```r
# 1a. What is the mean fiddler crab carapace size?
mean(crab_data$size)
```

```
## [1] 14.65801
```

```r
# 1b. What is the median fiddler crab carapace size?
median(crab_data$size)
```

```
## [1] 14.435
```

```r
# 1c. What is the standard deviation in fiddler crab carapace size?
sd(crab_data$size)
```

```
## [1] 3.503304
```

Question 1a-c Answers

- **1a)** The mean fiddler crab carapace size is 14.66 millimeters.

- **1b)** The median fiddler crab carapace size is 14.44 millimeters.

- **1c)** The standard deviation in fiddler crab carapace size is 3.50 millimeters.

Creating Histograms

```
## Bin sizes with Rice Rule 2n^(1/3) = 14.63722, rounded up to 15 per bin, n = 392
min(crab_data$size)
```

```
## [1] 6.64
```

```
max(crab_data$size)
```

```
## [1] 23.43
```

```
str(crab_data$size)
```

```
##  num [1:392] 12.4 14.2 14.5 12.9 12.4 ...
```

```
2*392^(1/3)
```

```
## [1] 14.63722
```

```
# 2a. Create a figure that is a histogram of fiddler crab carapace size (across all sites)
Histogram_all_sites <- ggplot(data = crab_data, aes(x = size)) +
  geom_histogram(bins = 15, color = "black", fill = "darkorange") +
  labs(x = "Carapace Size (mm)",
       y = "Fiddler Crab Count (n = 392)",
       title = "Fiddler Crab Carapace Sizes Across All LTER Salt Marshes",
       caption = "Figure 1. All Fiddler crab body sizes (mm) recorded summer 2016
       in salt marshes from Florida to Massachusetts from the Long Term Ecological
       Research program (LTER) Network: including Plum Island Ecosystem LTER,
       Virginia Coast LTER, and NOAA's National Estuarine Research Reserve System
       (392 total observations of fiddler crabs).") +
  theme_bw() +
  theme(legend.position = "none")

Histogram_all_sites
```

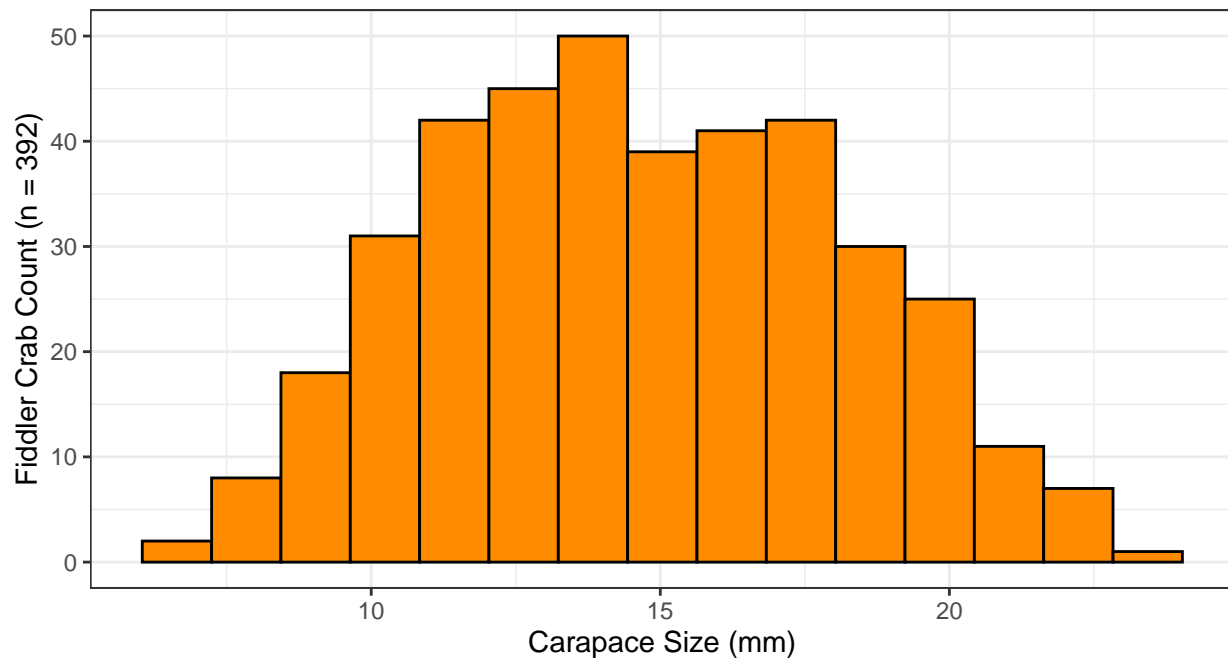# Fiddler Crab Carapace Sizes Across All LTER Salt Marshes

Figure 1. All Fiddler crab body sizes (mm) recorded summer 2016
in salt marshes from Florida to Massachusetts from the Long Term Ecological
Research program (LTER) Network: including Plum Island Ecosystem LTER,
Virginia Coast LTER, and NOAA's National Estuarine Research Reserve System
(392 total observations of fiddler crabs).

```r
# 2b. Create a figure that is a histogram of fiddler crab carapace size (each of the 13 sites)
Histogram_13_sites <- ggplot(data = crab_data, aes(x = size, fill = site)) +
  geom_histogram(bins = 15, color = "black") +
  facet_wrap(vars(site)) +
  labs(x = "Carapace Size (mm)",
       y = "Fiddler Crab Count (n = 392)",
       title = "Fiddler Crab Carapace Sizes for Each LTER Salt Marsh Site",
       caption = "Figure 2. Fiddler crab body sizes (mm) for each salt marsh site
       recorded in summer 2016 from the Long Term Ecological Research program (LTER) Network.

       Abbreviations of each site name: GTM=Guana Tolomoto Matanzas NERR
       | SI=Sapelo Island NERR | NIB=North Inlet Winyah Bay NERR | ZI=Zeke's Island NERR |
       RC=Rachel Carson NERR | VCR=Virginia Coastal Reserve LTER | DB=Delaware Bay NERR |
       JC=Jacques Cousteau NERR | CT=Sixpenny Island - Connecticut | NB=Narragansett Bay NERR |
       CC=Cape Cod | BC=Bare Cove Park | PIE=Plum Island Estuary - West Creek") +
  theme_bw() +
   theme(legend.position = "none")

Histogram_13_sites
```
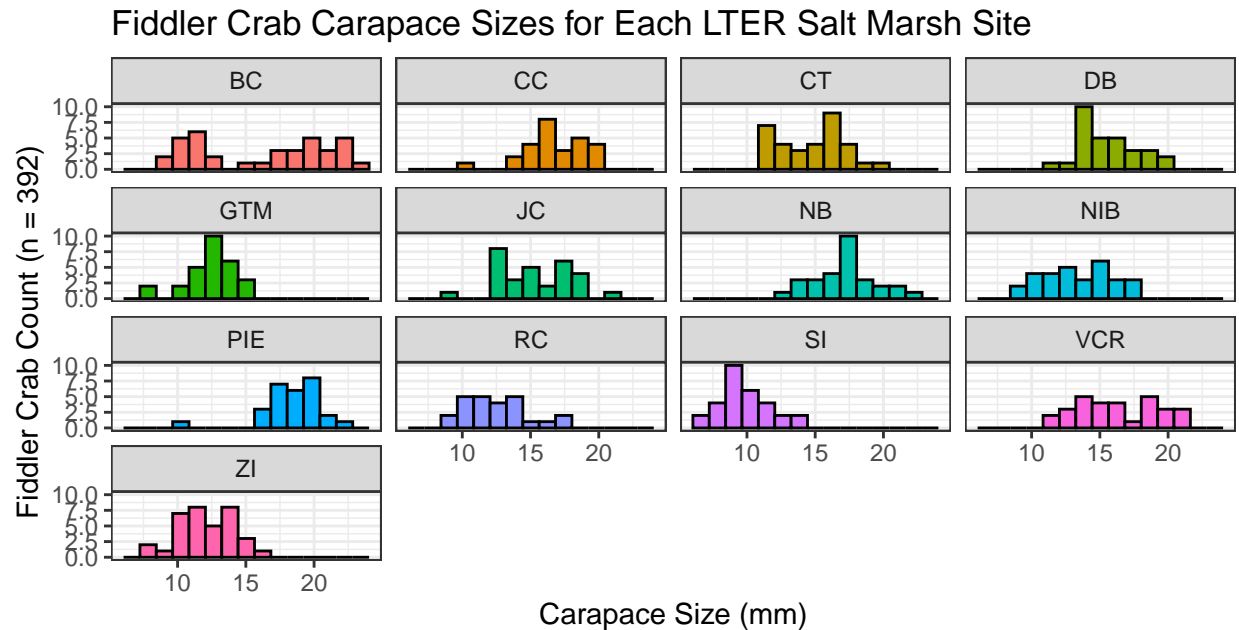
Figure 2. Fiddler crab body sizes (mm) for each salt marsh site
recorded in summer 2016 from the Long Term Ecological Research program (LTER) Network.

Abbreviations of each site name: GTM=Guana Tolomoto Matanzas NERR
| SI=Sapelo Island NERR | NIB=North Inlet Winyah Bay NERR | ZI=Zeke's Island NERR |
RC=Rachel Carson NERR | VCR=Virginia Coastal Reserve LTER | DB=Delaware Bay NERR |
JC=Jacques Cousteau NERR | CT=Sixpenny Island – Connecticut | NB=Narragansett Bay NERR |
CC=Cape Cod | BC=Bare Cove Park | PIE=Plum Island Estuary – West Creek

Question 2c Answers

- **How do the raw values of fiddler crab carapace size in your first figure compare to the mean and median values you calculated above?** The tallest bin with largest number of fiddler crabs is about less than 15 millimeters, which closely aligns and matches with the mean (14.66 mm) and median (14.44 mm) values calculated above. This is because the first figure has a binomial distribution shape that peaks around the 14.5 mm mark in fiddler crab carapace size range.

- **Based on your second figure, do any sites display strong differences in their data distributions?** Yes, there are strong differences across 13 data site distributions. For instance, four sites have symmetric, bimodal distribution (BC, CT, VCR, and ZI) with two groups either smaller or larger carapace size ranges, four sites are skewed to the left (GTM, SI, ZI, RC) favoring smaller carapace sizes, three sites are skewed to the right (NB, PIE, CC) favoring larger carapace sizes.

Data Visualization - The Golden Rules

```
Good_golden_rules <- ggplot(data = crab_data, aes(x = water_temp, y = size, fill = name)) +
  geom_boxplot(width = 2) +
  facet_wrap(vars(name)) +
  labs(x = "Water Temperature (Celcius)",
       y = "Carapace Size (mm)",
       title = "Fiddler Crab Carapace Sizes Across Different Water Temperatures for each
       LTER Salt Marsh site",
```

```
          caption = "Figure 3. Distribution of fiddler crab body sizes (mm) and their associated water
          temperatures for each salt marsh site recorded in summer 2016 from the Long Term Ecological
          Research program (LTER) Network. (Hidden full site names include Guana Tolomoto Matanzas
          NERR, North Inlet Winyah Bay NERR, Plum Island Estuary - West Creek, Sixpenny Island -
          Connecticut, and Virginia Coastal Reserve LTER)") +
    theme_bw() +
     theme(legend.position = "none")

Good_golden_rules
```
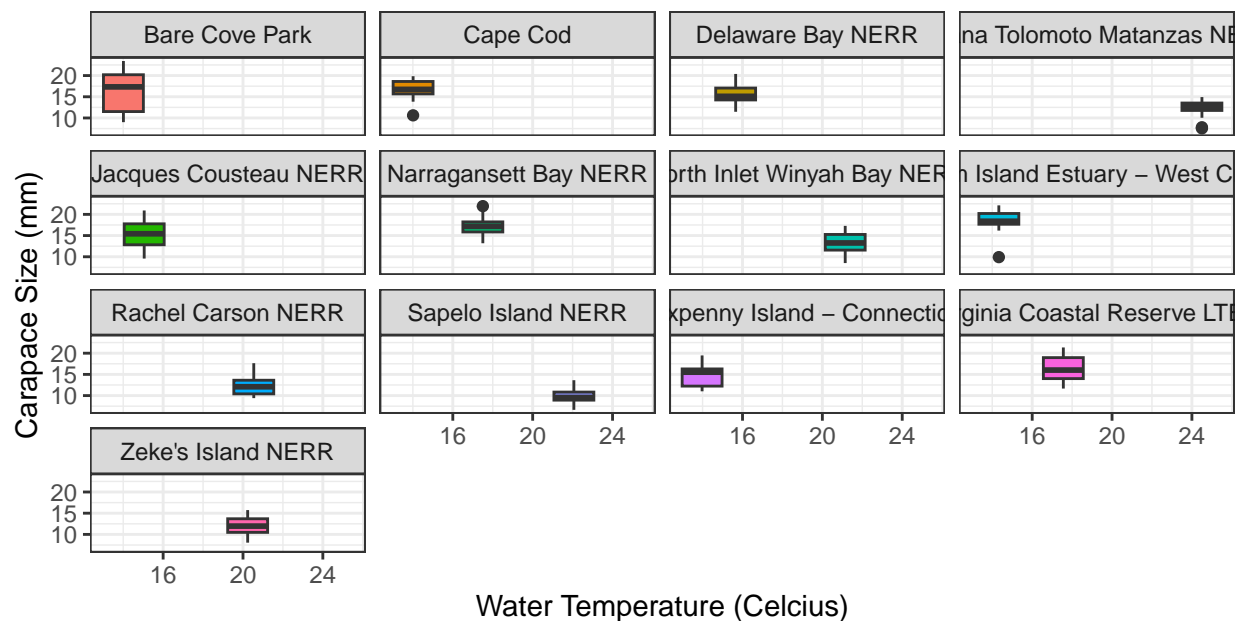

Figure 3. Distribution of fiddler crab body sizes (mm) and their associated water temperatures for each salt marsh site recorded in summer 2016 from the Long Term Ecological Research program (LTER) Network. (Hidden full site names include Guana Tolomoto Matanzas NERR, North Inlet Winyah Bay NERR, Plum Island Estuary – West Creek, Sixpenny Island – Connecticut, and Virginia Coastal Reserve LTER)

```
Bad_golden_rules <- ggplot(data = crab_data, aes(x = water_temp, y = size, group = name)) +
  geom_boxplot() +
  labs(x = "Water",
       y = "Size",
       title = "Water vs. Size",
       caption = "Figure 4. really accurate histogram.") +
    theme_bw() +
     theme(legend.position = "none")

Bad_golden_rules
```

Water vs. Size

Figure 4. really accurate histogram.
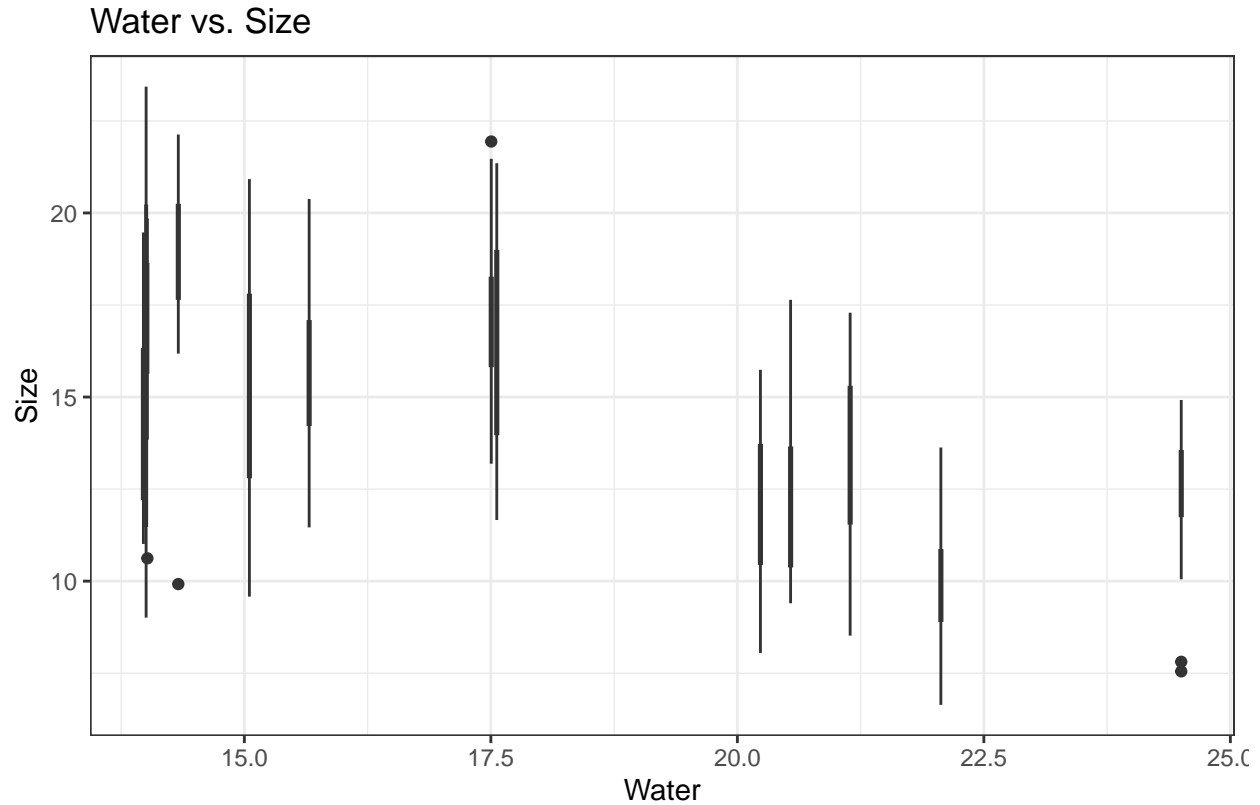
Question 3 answers **Figure 3 represents a good representation of the golden rules** because the box plot showcases **ACCURACY** with interquartile ranges (Q1, Median, Q2, range, and potential outliers) of carapace sizes, **CLARITY** with a descriptive caption, and **AESTHETICS** with separated graphs for each site with easy to read labelled site names with different colors that quickly differentiate each unique site dataset.

Figure 4 is a bad representation of the golden rules because *lacks* **ACCURACY** with overlapping sites in lower water temperature ranges and inaccurate figure caption description, **CLARITY** with very skinny width box plots where you cannot clearly see the interquartile range marks, and **AESTHETICS** with missing unique site names and colors.