Example 15
Mann-Whitney U
Comparing sample medians,
Nonparametric
Setup

Mussels were collected from an area that had been heavily harvested during the 1950s and a nature reserve where mussels were never harvested. The company responsible for the over-harvesting has claimed that the mussel populations have recovered. The state wildlife agency thinks the mussels from the harvested area are abnormally small (Units are cm). Do the harvested area and the refuge differ in the size of the mussels they contain?

Data (cm)

Harvested	Refuge
6.0	9.3
6.1	8.1
8.0	9.4
5.3	10.2
5.3	8.0
5.2	10.4
4.4	10.0
4.4	9.0
8.1	10.0
4.3	9.4
4.4	10.0
5.3	10.2
5.2	10.4
5.1	11.0
7.1	8.2
7.4	10.4
6.2	9.2
5.2	8.1
6.3	9.2
5.1	13.0

Example 15
Mann-Whitney U
Comparing sample medians,
Nonparametric
Solution

1. State your question: Are mussels from the harvested area significantly smaller than the mussels from the refuge?
    a. Is it a good scientific question? Definable, measurable, and controllable
    b. Identify your population: The size of mussels
    c. Identify your dependent variable: Size
    d. Identify your independent variable: Location (Harvested vs. refuge area)
2. State your hypothesis set
    a. Verbal hypothesis: The mussels from the harvested area are significantly smaller than those from the refuge.
    b. Statistical hypothesis (H_O, H_A). (No parameters)
        H_O: The size of mussels does not differ between the two sites.
        H_A: The mussels from the harvested area are significantly smaller than those from the refuge.
    c. Is your hypothesis set exhaustive? No, truncated
    d. Is your hypothesis set exclusive? Yes
3. State your significance level: α=0.05
4. Select the appropriate test.
    a. Variable scales.
        i. Dependent variable: Ratio
            - Converted: Ratio → Ordinal
        ii. Independent variable: Nominal
            - Converted or transformed? No
    b. What information is given or available
        i. Sample data
    c. Number of samples: 2
    d. Are the data paired or unpaired? Unpaired
    e. What aspect of the variable do you want to compare?
        ~~i. Central tendency - means~~
        ii. Central tendency - medians
    ~~f. State the test to be used: t-test (Two-sample)~~
        ~~i. Are the assumptions of the test met? No~~
            - ~~Random sample – Assumed~~
            - ~~Independent samples – Assumed~~
            - ~~Normally distributed populations – Tested – Failed~~
            - ~~Equal variances – Tested – Passed~~

Shapiro-Wilk normality test

data: MusselsA$Size
W = 0.891, p-value = 0.02811

Shapiro-Wilk normality test

data: MusselsB$Size
W = 0.9099, p-value = 0.06359

F test to compare two variances

data: Mussels$Size by Mussels$Location
F = 0.9748, num df = 19, denom df = 19, p-value = 0.9562
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
0.3858334 2.4627573
sample estimates:
ratio of variances
0.9747893

g. State the test to be used: Mann-Whitney U
   i. Are the assumptions of the test met? Yes
      - Random sample – Assumed
      - Independent samples – Assumed
      - Distributions of similar shape -- Graphed

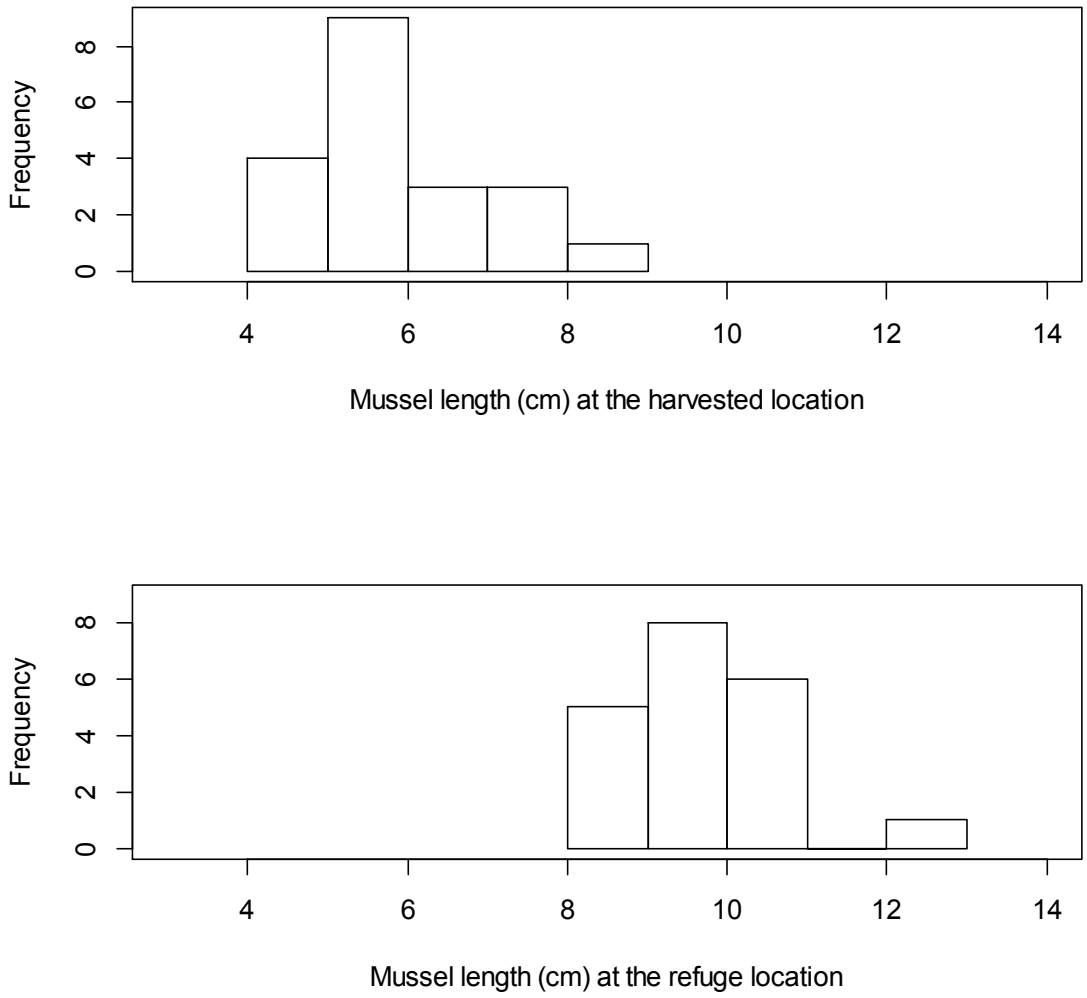

Figure 1. The distribution of mussel length (cm) at harvested and refuge locations.

5. Conduct your sampling
   Sample data provided above
6. Graph the data
   See above
7. Summarize the data.
   $M_{Harvested}$ = 5.3 cm

   M_{refuge} = 9.7 cm

8. Calculate your test statistic.

Wilcoxon rank sum test with continuity correction

data: Dependent by Independent
W = 2.5, p-value = 4.748e-08
alternative hypothesis: true location shift is less than 0

Warning message:
In wilcox.test.default(x = c(6, 6.1, 8, 5.3, 5.3, 5.2, 4.4, 4.4, :
cannot compute exact p-value with ties

9. Retain or reject your null hypothesis based on your test statistic.
Since the calculated p-value(4.75-08) is less than the significance level (0.05) and the median of the harvested location (5.3 cm) is less than the median of the refuge location (9.7 cm) we would reject our null hypothesis and retain our alternate hypothesis.

10. Interpret the results in biological terms.
The mussels at the harvested location are significantly smaller than those at the refuge location ($W=2.5$, $n=20,20$, $p<0.001$).