

R TUTORIAL, #3: **MEASURES OF CENTER**

The (>) symbol indicates something that you will type in.

A bullet (•) indicates what the R program should output (and other comments).

TEST SCORES DATA SET

> Type: scores <- c(80, 76, 100, 83, 100)

> Type: scores

- Check to make sure you entered the values in correctly.

MEAN

- Find the mean of the scores.

> Type: mean(scores)

- You can also find the mean as follows:

> Type: sum(scores) / length(scores)

TRIMMED MEANS

> Type: mean(scores, trim=.2)

- This takes the trimmed mean of the data set after the top 20% and the bottom 20% have been deleted. Here, the '76' and one of the '100's are deleted.

MEDIAN

- Find the median of the scores.

> Type: median(scores)

MODE

- Unfortunately, ‘mode’ doesn’t give us what we want.
- The ‘table’ command gives frequencies, and it can help us find the mode:
 - > Type: table(scores)
 - Look for the data value with the highest frequency.
 - > Type: max(table(scores))
 - This indicates that the mode (100 points) has a frequency of 2.
- The longer command below actually finds the mode.
 - > Type: which(table(scores) == max(table(scores)))
 - The ‘4’ indicates that ‘100’ is the fourth-lowest unique data value, as shown in the table given by ‘table(scores)’.

MIDRANGE

- Unfortunately, there is no ‘midrange’ command.
- We will find the midrange step-by-step:
 - > Type: min(scores)
 - You should see ‘76’.
 - > Type: max(scores)
 - You should see ‘100’.
 - > Type: (76+100)/2
 - You should see ‘88’, the midrange in points.
- Here’s another way to find the midrange:
 - > Type: range(scores)
 - You should see the min and the max of the data set.
 - **WARNING:** This is not really the “range” of the data set. We will use the term “range” to refer to a measure of spread.
 - > Type: mean(range(scores))
 - You should see ‘88’, the midrange in points.

ESTIMATED MEANS FROM A FREQUENCY TABLE; WEIGHTED MEANS

- > Type: `x <- c(54.5, 64.5, 74.5, 84.5, 94.5)`
 - These are the class marks for the test scores.
- > Type: `f <- c(2, 4, 7, 10, 5)`
 - These are the frequencies.
- > Type: `data.frame(x,f)`
 - You will see a frequency table for the test scores.
- > Type: `sum(x*f) / sum(f)`
 - These gives an estimate for the mean of the test scores.
 - `sum(x*f)` is the estimated sum of the test scores using “zigzagging.”
 - `sum(f)` is N , or 28, the total number of test scores.

EXERCISE: Find the GPA in our GPA example.

COURSE GRADE

- > Type: `Exam <- c("Q1", "Q2", "Q3", "Mid", "Final")`
 - ‘Exam’ is a vector of character strings.
- > Type: `weights <- c(.1, .1, .1, .25, .45)`
 - These are the exam weights.
- > Type: `scores <- c(75, 80, 100, 95, 91)`
 - These are the exam scores. We assume the student gets a ‘91’ on the Final.
- > Type: `data.frame(Exam, weights, scores)`
 - You should see the student’s grade record in the class.
- > Type: `sum(weights*scores)`
 - You should see the student’s course average if s/he gets a ‘91’ on the Final.
 - Will the student get an ‘A’?

EXERCISE: Experiment with the ‘scores’ vector by trying different scores for the Final exam. Then, find what the student will get as the resulting class average.