The Basics of Programming in R

Workshop 2

Aggie Math Learning Center
USU Mathematics and Statistics Department
Previous Workshop

- R and RStudio Installation
- Basic User Interface
- Basic Functions and Data Manipulation
Communication in R

How does R communicate with us?

- Error and Warning Messages
- Documentation
Error and Warning Messages

- Every package and message is different
- No ‘one size fits all’ approach
General Approach

- Error messages and warnings appear in the console.
- Tend to be different colors than the ‘standard’ output. Red is common, but it depends on your color scheme.
- An Error is a ‘No’. R will not run a code if there is an error.
- A Warning is when R can run the code, but some issue may be stopping the output from being what you expect.

```r
> plot(incom, schoolNew$sesall)
Error in plot(incom, schoolNew$sesall) : object 'incom' not found

Warning message:
Removed 1 rows containing missing values (geom_path).
```
Reading Errors and Warnings in BaseR

- Most errors and warnings have several parts
  - Error Location
  - Error Message

```
> plot(incom, schoolNew$sesall)
Error in plot(incom, schoolNew$sesall) : object 'incom' not found
```

The error happened when R tried to run this line of code. This is what stopped R from running this line of code.
There is no easy way, no workshop can go over this well - each package is written by a distinct individual, who defines the error messages themselves. However, most do contain similar information as the BaseR errors:

- Location of the Error
- Reason for the Error
  - ‘Unexpected symbol or token ‘)’ in…
  - ‘Could not find/ Not found’
  - Invalid operation
Fixing Errors and Warnings

- Go to location, analyze error
  - Typos or incorrect objects are easy.
- Errors are unique, but you can use resources like the ? and ?? commands and even Google!

There is no ‘one way’ to fix bugs. Solving such errors is the journey of any programmer in any language.
R Documentation

- [https://www.rdocumentation.org/](https://www.rdocumentation.org/)
- [https://www.r-project.org/other-docs.html](https://www.r-project.org/other-docs.html)

- These sites have the documentation of several different packages - again, the documentation is written by many different individuals, and can vary.
To test for the equality of the means of the two examples, we can use an *unpaired* $t$-test by

```r
> t.test(a, b)

Welch Two Sample t-test

data:  a and b
t = 2.2499, df = 12.037, p-value = 0.00694
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.0185260 0.7018120
sample estimates:
mean of x mean of y
80.00877 79.57875
```

which does indicate a significant difference, assuming normality. By default the R function does not assume equality of variances in the two samples (in contrast to the similar S-PLUS `t.test` function). We can use the $F$ test to test for equality in the variances, provided that the two samples are from normal populations.

```r
> var.test(a, b)

F test to compare two variances
data:  a and b
F = 0.5837, num df = 12, denom df = 7, p-value = 0.3938
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1251007 2.1052637
sample estimates:
  ratio of variances
0.5837408
```

which shows no evidence of a significant difference, and so we can use the classical $t$-test that assumes equality of the variances.

```r
> t.test(a, b, var.equal=TRUE)

Two Sample t-test
data:  a and b
t = 3.4722, df = 19, p-value = 0.002551
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.01669058 0.06734788
sample estimates:
mean of x mean of y
80.00877 79.57875
```