Basic Statistical Procedures in R

How we will approach these tests

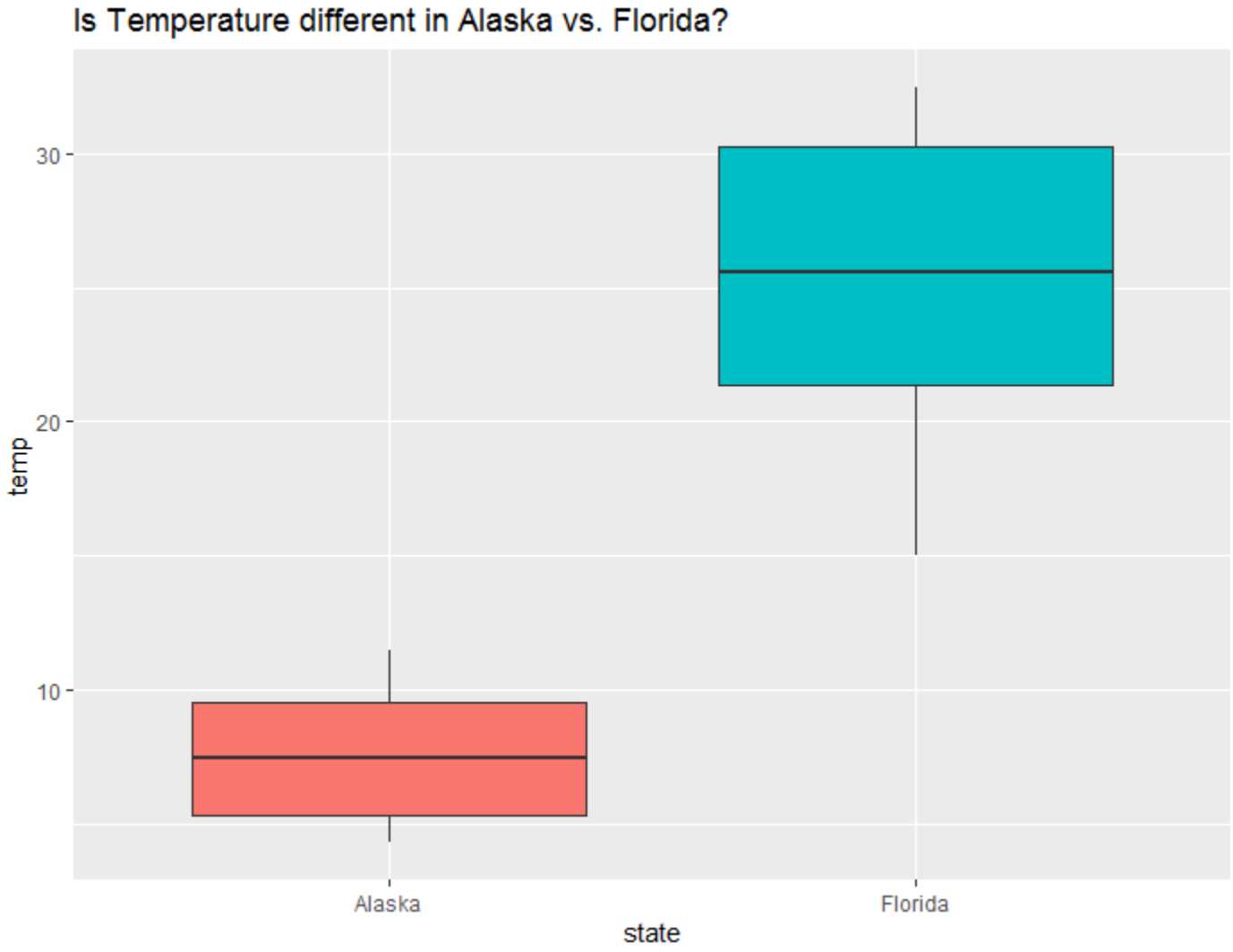
1. Decide on a test to run
2. Find the function you need for a certain test
3. Figure out what assumptions that test and function are making
   - data format: columns/rows; number vs. character; etc.
   - normality? (not always needed)
4. Decide whether your data meets those assumptions
5. Look at the function's inputs: what do you feed into it?
6. Look at the function's default values; do you want to change them?
7. **Code it!**
8. Explore the output
9. Pull out certain pieces of the output for later use

Road map for this session

1. t-test
2. ANOVA
3. linear regression
4. infinity and beyond….

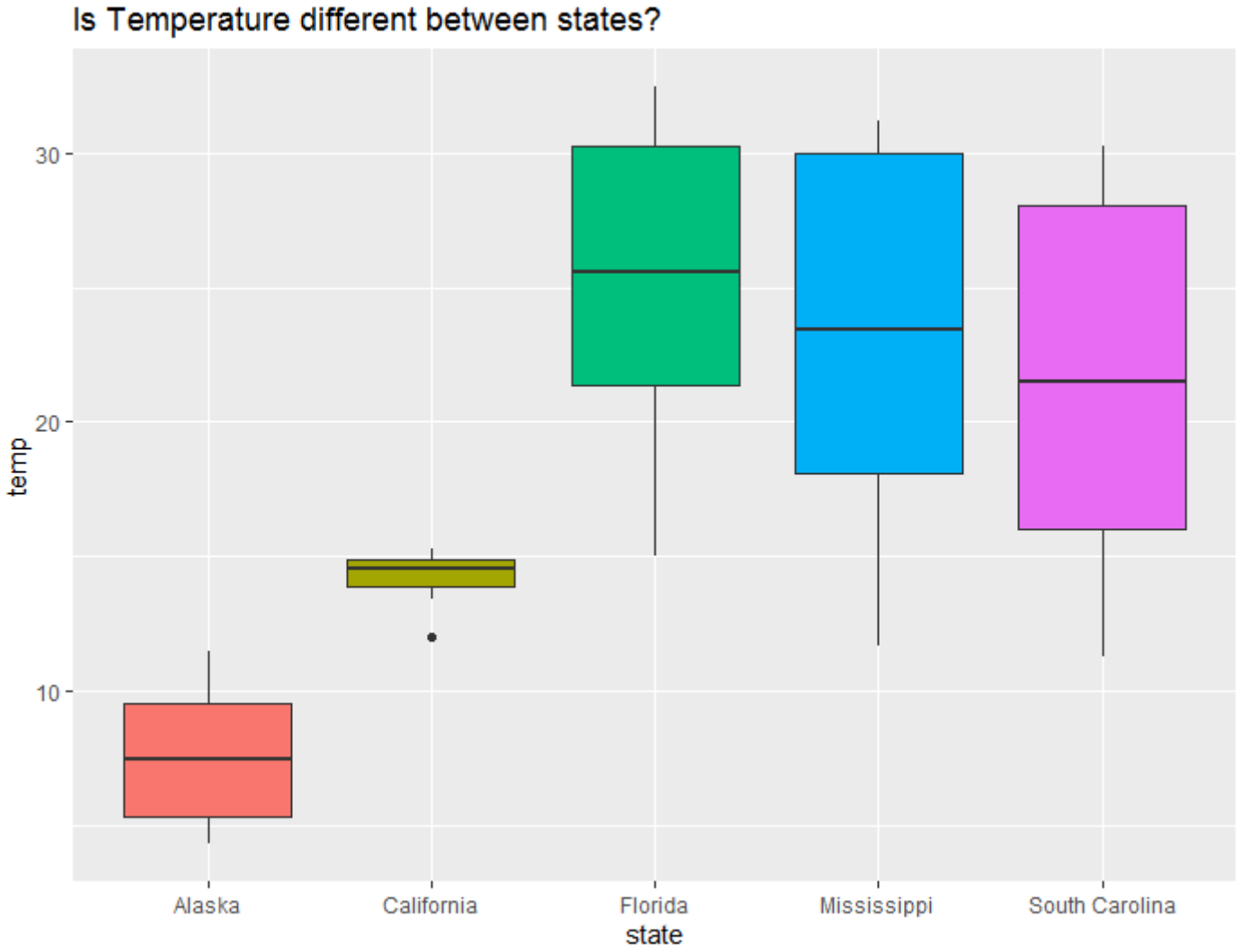
A motivating example: Temperature

1. t-test: compare means of 2 groups
2. ANOVA
3. linear regression
4. infinity and beyond….



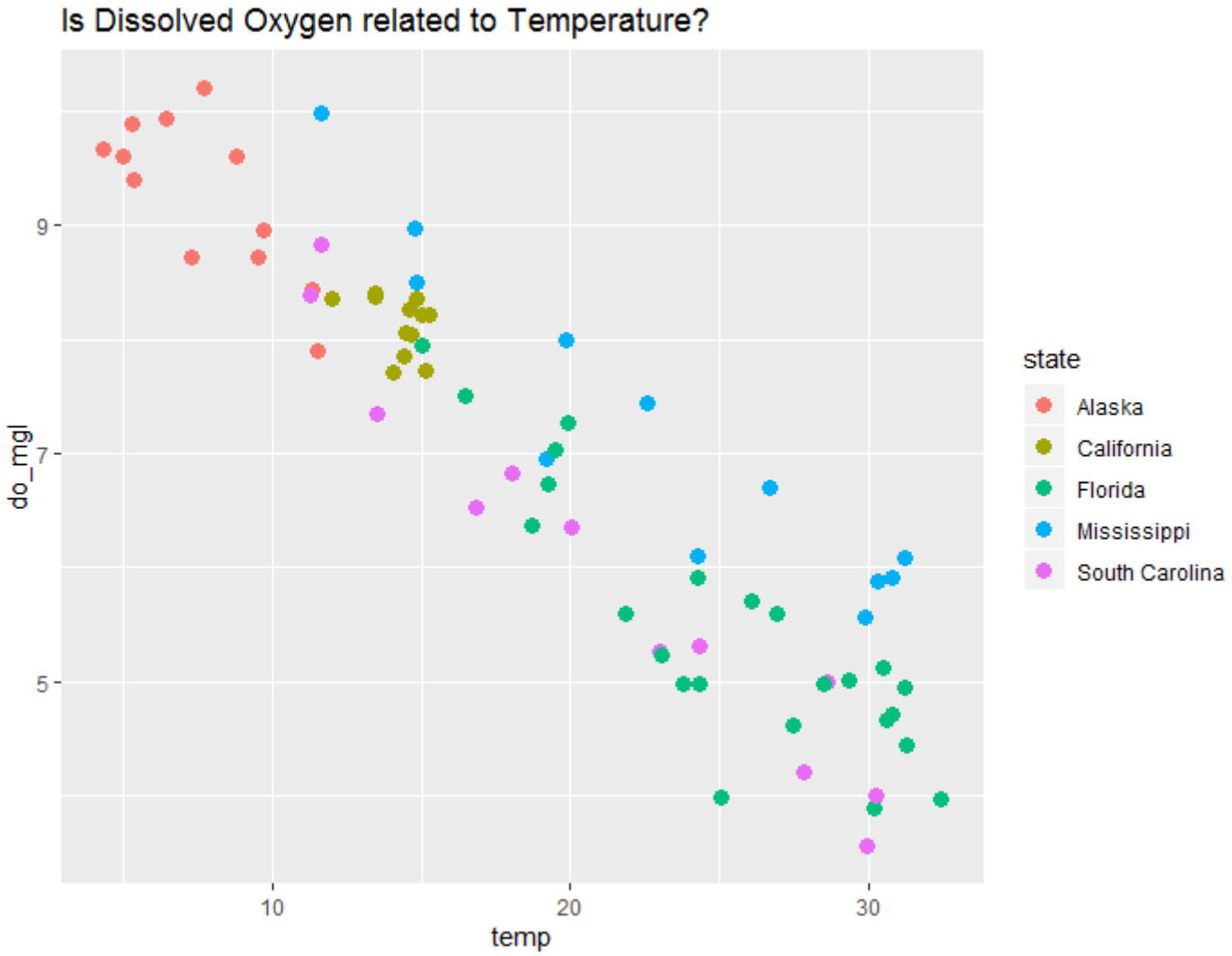
A motivating example: Temperature

1. t-test
2. ANOVA: compare means of >2 groups
3. linear regression
4. infinity and beyond….



A motivating example: Temperature/DO

1. t-test
2. ANOVA
3. linear regression: continuous relationships
4. infinity and beyond….



Deciding which function to use

- Try to see what *other* people in your situation are using:
  - CRAN task views: https://cran.r-project.org/web/views/
  - online class notes
  - blog posts
  - social media: #rstats on twitter; “Ecology in R” group on Facebook
  - scientific publications
- Look for packages and functions with good documentation
  - blog posts
  - vignettes
  - github READMEs
  - help files
- Use what *you* can understand (and test it first)

Useful functions

- Performing the tests:
  - t.test()
  - aov()
  - lm()
- Exploring test outputs:
  - summary()
  - broom::tidy()
  - broom::glance()