Dplyr Lecture Comments

2017-08-10 – Damian

* slide 10, In class exercise: select -- this would make a good live coding exercise, watching the instructor do the task while they follow along.
* slides 12-14, boolean logic: I think it would be better to demonstrate the conditionals using filter examples--it would be more concrete to students and keep the main focus on the 5 verbs. Running multiple filter() examples would also make a good live coding exercise or in-class quiz if you ask them to predict results
* slides 16-19. Introducing the pipe operator can be tricky. Walking the students through a sequence without pipes first would make the pipe operator more concrete and demonstrate the benefits. Also good live coding exercises so you can show the results are equivalent. E.g:
  1. Start with creating a sequence with intermediate objects

df2 <- filter(df1, condition1)

df2 <- filter(df2, condidtion2)

* 1. Then do the same sequence with pipes instead of objects (or do this last)

filter(df1, condition1) %>% filter(condidtion2)

* 1. Then nest the functions instead of creating objects (or do this 2nd)

filter(filter(df1, condition1), condidtion2)

* 1. Introduce the explicit dot notation
* slide 32-34. this plot is fairly complex -- asking students to master dplyr verbs and elaborate on ggplot may be too much to information to cover. The review could be useful but they'll probably need time to do the plots or time for a refresher. Perhaps it would be better to have them practice using the dplyr verbs to output data frames without visualizing, or simpler plots
* slide 36. Joins. The venn diagrams are intuitive for a lot of people, but they leave out a lot of detail about what happens to each row and column, which will take time explain. I suggest using the Rstudio data wrangling cheatsheets diagrams also (or instead). 