Data Wrangling basics via the dplyr package

STAT 234

Suppose we have the following data set named `df`. The first two columns are numeric while the third column is categorical.

x1	x2	cat1
1	3	Yes
7	NA	Yes
4	2	No

Create New Variables with `mutate()` (Perhaps with `if_else()` or `case_when()`)

Choose Rows to Keep with `filter()`

Choose Columns to Keep with `select()`

x1	x2	cat1
1	3	Yes
7	NA	Yes
4	2	No

Order/Sort Your Data Set with `arrange()`

Obtain Numerical Summaries with `summarise()`

Obtain Numerical Summaries by Group with `group_by()` and `summarise()`