

Data Wrangling in the tidyverse

Tools Used in the Workshop

- `read_csv()` # load a csv formatted file into R memory
- `str()` # check structure of the object and information about the class, length and content of each column
- `View()` # invoke a spreadsheet-style data viewer
- `select()` # select columns of a data frame
- `filter()` # allows you to select a subset of rows in a data frame
- `%>%` # pipes to select and filter at the same time
- `mutate()` # create new columns based on the values in existing columns
- `head()` # shows the first 6 rows
- `group_by()` # split the data into groups, apply some analysis to each group, and then combine the results.
- `summarize()` # collapses each group into a single-row summary of that group
- `mean()` # calculate the mean value of a vector
- `!is.na()` # test if there are no missing values
- `print()` # print values to the console
- `min()` # return the minimum value of a vector
- `arrange()` # arrange rows by variables
- `desc()` # transform a vector into a format that will be sorted in descending order
- `count()` # counts the total number of records for each category
- `case_when()` # allows you to vectorize multiple `if_else()` statements
- `left_join()` # return all rows from x, and all columns from x and y
- `right_join()` # return all rows from y, and all columns from x and y
- `inner_join()` # return all rows from x where there are matching values in y
- `full_join()` # return all rows from x and all rows from y
- `spread()` # reshape a data frame by a key-value pair across multiple columns
- `gather()` # reshape a data frame by collapsing into a key-value pair
- `n_distinct()` # get a count of unique values
- `wday()` # grabs the weekday component of a date-time
- `write_csv()` # save to a csv formatted file
- `ymd()` # transforms dates stored in character and numeric vectors to dates