Tidyverse Cheat Sheet

|  |  |  |  |
| --- | --- | --- | --- |
| **I want to …** | | **Example** | |
| View/Read in data | Read in a CSV, TSV, or other Delimited file  Make a tibble from scratch or with files  View dataframe in a new tab  Get a glimpse of a datatable | **read\_csv**(“filepath”)  **read\_tsv**(“filepath”)  **read\_delim**(“filepath ”, “Single character delim”)  **tibble**(variable 1, variable 2, …)  **view**(data)  **glimpse**(data) | |
| Restructure Tibble | Apply functions to a subset of columns within a dataset, creating a new dataframe.  Merge multiple columns together by pasting strings together  Separate single column into multiple columns  Stack data frames on top of each other by column  Stack data frames side by side of each other  by row  Mutate joining columns from y to x, matching rows based on keys. Inner includes all rows  in x and y, left includes all rows in x, right all rows in y, full all rows in x or y.  Sorting tibble rows by column values  Pivot data from wide to long  Pivot data from long to wide | **summarise**(data, function, grouping structure)  **unite**(data, vector of column names, joining character, TRUE to remove used columns)  **separate**(data, column, into, separator, remove)  **bind\_rows**(list or vector of what to bind)  **bind\_cols**(list of vector of what to bind)  **inner\_join**(x, y, by)  **left\_join**(x, y, by)  **right\_join**(x, y, by)  **full\_join**(x, y, by)  **arrange**(data, column)  **pivot\_longer**(data, columns to pivot, names\_to = “NAME”, values\_to = “VALUE”)  **pivot\_wider**(data, columns to pivot, names\_to = “NAME”, values\_to = “VALUE”) | |
| Grab Something from a tibble | Subset rows using column values  Subset columns using their names and data types.  Group by one or more variables.  Remove grouping on one or more variables.  Select variables that match a pattern by prefix, suffix, literal string, regular expression, numerical range.  Compare/Compute the values behind/lagged or ahead/leading of the current value | **filter**(data, by)  **select**(data, vector of columns)  **group\_by**(data, variables or computations to group by)  **ungroup**()  **starts\_with**(“match”)  **ends\_with**(“match”)  **contains**(“match”)  **matches**(“match”)  **num\_range**(prefix that starts the numeric range, range)  **lag**(vector of values)  **lead**(vector of values) | |
| Change Elements of a Tibble | Add new variables and preserves existing ones  Add new variables and drop existing ones.  Rename columns  Replace numeric values based on their position or their name, and character or factor values only by their name  Clean up the names of a tibble | **mutate**(data, what to change)  **transmutate**(data, what to change)  **rename**(data, new name = old name)  **rename\_with**(data, new name = old name, function)  **recode**(vector, key)  **clean\_names**(data) | |
| Alter a string in a tibble | Join Multiple strings into a single string  Remove matched patterns in a string  Replace matched patterns in a string  Alter case of a string  Keep strings matching a pattern  Count the number of matches of a string | **str\_c**(“string1”, “string2”, sep)  **str\_remove**(string, pattern)  **str\_remove\_all**(string, pattern)  **str\_replace**(string, pattern, replacement)  **str\_replace\_all**(string, pattern, replacement)  **str\_replace\_na**(string, pattern, replacement)  **str\_to\_upper**(string)  **str\_to\_lower**(string)  **str\_subset**(string, pattern)  **str\_which**(string, pattern)  **str\_detect**(string, pattern)  **str\_count**(string, pattern) |  |
|  |
| Other Useful Functions | Do one thing if stated condition is met, do a different thing if not.  Vectorise multiple if\_else statements  Apply a function to each element of a list or vector | **if\_else**(condition, if true, if false)  **case\_when**(statement 1, statement 2, ... )  **map**(list or vector, function, additional arguments) | |