## Data Transformation with dplyr::

list(suffix for column names = ~ list(fn(.x. fn args))

A guide to 37° behaviors applied to one tibble (tb1) or data.frame **AGGREGATE** © R Data Berlin **MODIFY** COUNT action "= (ren.+sel.+transm.+ column-related mut.+arr.+fil.+ summ. row-related **COLUMNS ROWS** +gr\_by) \* (\_+\_all+\_at+\_if)+ tallv() summarize function name (part) (ungr.+tal.+t\_n+t\_frac+sl.) function name (full) • group\_by is effective on all functions but arrange, **POSITIONS VALUES GROUPINGS POSITIONS** NAMES unless using arrange\_suffix(..., .by\_group = TRUE) select\_helper = starts\_with(), ends\_with(), contains(), matches(), num\_range(), one\_of(), group\_by\* everything(), or last\_col() [from pkg tidyselect] keep keep keep keep keep keep all focal focal all all focal ungroup(` based on based on based on based on rank cut-off select test percentile cut-off transmute mutate rename arrange\* their positions on values on values on values † %>% group\_by(x) WHICH SUFFIX? %>% tallv() %>% ungroup() = filter top\_frac() slice(` top\_n() %>% count(x) specific action common action on each focal on all focal **SYNTAX:** tbl %>% verb\_suffix(., ..., .args) column columns [font style varies to tease appart placeholders from true R\_commands] tbl %>% rename(new\_name\_columnX = old\_name\_columnX) focal focal tbl %>% select(name\_columnX, new\_name\_columnY = old\_name\_columnY, select\_helper ("text")) focal = all identified by identified by tbl %>% transmute(name\_columnZ = fn\_scalar\_or\_vectorized\*(name\_columnX)) positions test on values tbl %>% mutate(name\_columnZ = fn\_scalar\_or\_vectorized (name\_columnX)) tlb %>% tallv()<sup>†</sup>: tlb %>% tallv(name numeric column, wt = weights)<sup>†</sup> tbl %>% arrange(name\_columnX, desc(name\_columnY)) tbl %>% filter(fn\_test\_vectorized(name\_columnX), fn\_test\_vectorized(name\_columnY)) if() at() tbl %>% top\_n(n = number. name\_columnX); tbl %>% top\_frac(n = fraction, name\_columnX) tbl %>% slice(row\_indices) tbl %>% summarize(name\_columnZ = fn\_scalar (name\_columnX)) ■ tbl %>% group\_by(name\_columnX, name\_columnY) %>% verb\_suffix() & tbl %>% ungroup() if the function fn does not return a scalar or a nrow-long vector, tbl %>% verb\_all(.funs = fn fn fn\_arg\_name1 = fn\_arg1, fn\_arg\_name2 = fn\_arg2) # included below use list(fn()) to nest the output \*\* for multiple columns dispatch tbl %>% verb\_at(.vars = indices\_for\_focal\_columns OR .vars = vars(select\_helper ("text")), use list(suffix\_for\_column\_names = fn) or in case of \*+ \*\*

tbl %>% verb if(.predicate = logicals for each column.