

Day 6 Cheatsheet

Data Manipulation

Major concepts:

- Wide data - multiple columns per individual, values spread across multiple columns
- Long data - multiple rows per observation, a single column contains the values

Functions

Library/Package	Piece of code	Example of usage	What it does
dplyr	<code>separate()</code>	<code>df %>% separate(x, c("A", "B"))</code>	Separate a character column into multiple columns with a regular expression or numeric locations
dplyr	<code>unite()</code>	<code>df %>% unite("z", x:y, remove = FALSE)</code>	Unite multiple columns together into one column
tidyr	<code>pivot_longer()</code>	<code>df %>% pivot_longer(!col_to_keep, names_to = "new_col_with_labels", values_to = "new_col_with_values")</code>	Lengthens a data frame by increasing the number of rows and decreasing the number of columns.
tidyr	<code>pivot_wider()</code>	<code>df %>% pivot_wider(names_from = "col_with_names", values_from = "col_with_values")</code>	Widens a data frame by decreasing the number of rows and increasing the number of columns.
dplyr	<code>?_join()</code>	<code>inner_join(x, y)</code>	Joins data from two data frames. inner_join - only rows that match for x and y are kept. full_join - all rows of x and y are kept. left_join - all rows of x are kept even if not merged with y. right_join - all rows of y are kept even if not merged with x. anti_join - all rows from x not in y keeping just columns from x.

Library/Package	Piece of code	Example of usage	What it does
Base R	<code>duplicated()</code>	<code>duplicated(x)</code>	Determines and removes duplicate elements from <code>x</code> .
Base R	<code>t()</code>	Transpose	Returns the transpose of a matrix or data frame. If given a data frame, returns a matrix.

Data Visualization with `esquisse`

Major concepts

- The `esquisse` package can help you to test out plots and get familiar with the syntax required to make plots

Functions

Library/Package	Piece of code	Example of usage	What it does
<code>esquisse</code>	<code>esquisser()</code>	<code>esquisser(Orange)</code>	Start an <code>esquisse</code> interactive session to create a plot of the given <code>data.frame</code> or <code>tibble</code>

* This format was adapted from the cheatsheet format from AlexsLemonade.