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	separate()	<pre>df %>% separate(x, c("A", "B"))</pre>	Separate a character column into multiple columns with a regular expression or numeric locations
dplyr	unite()	<pre>df %>% unite("z", x:y, remove = FALSE)</pre>	Unite multiple columns together into one column
tidyr	<pre>pivot_longer()</pre>	<pre>df %>% pivot_longer(!column_to names_to = "new_col_with_labels", values_to = "new_col_with_values")</pre>	Lengthens a data
tidyr	<pre>pivot_wider()</pre>	<pre>df %>% pivot_wider(names_from = "col_with_names", values_from = "col_with_values")</pre>	Widens a data frame by decreasing the number of rows and increasing the number of columns.
dplyr	?_join()	inner_join(x, y)	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.
Base R	<pre>duplicated()</pre>	<pre>duplicated(x)</pre>	Determines and removes duplicate elements from x.

Library/Package	Piece of code	Example of usage	What it does
Base R	t()	Transpose	Returns the transpose of a matrix or data frame. If given a data frame, returns a matrix.

Data Visualization with esquisse

Major concepts

ullet 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
esquisse	esquisser()	esquisser(Orange)	Start an esquisse interactive session to create a plot of the given data.frame or tibble

 $^{\ ^*}$ This format was adapted from the cheat sheet format from AlexsLemonade.