

# Day 8 Cheatsheet

## Statistics

### Functions

| Library/Package                | Piece of code                             | Example of usage                                                 | What it does                                                                                                                                                |
|--------------------------------|-------------------------------------------|------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Base R                         | <code>cor(x, y)</code>                    | <code>cor(x, y)</code>                                           | Calculate correlation of two vectors in R.                                                                                                                  |
| <code>corrplot</code>          | <code>corrplot(cor_mat)</code>            | <code>corrplot(cor_mat, type = "upper", order = "hclust")</code> | Create a correlation matrix plot.                                                                                                                           |
| Base R                         | <code>t.test(x, y)</code>                 | <code>t.test(x, y, alternative = "two.sided")</code>             | Perform one and two sided t tests.                                                                                                                          |
| <code>broom</code>             | <code>tidy(x)</code>                      | <code>tidy(t_test_result)</code>                                 | Manipulates and tidies up wonky statistical R objects into nice data frames                                                                                 |
| Base R                         | <code>wilcox.test(x, y)</code>            | <code>wilcox.test(x, y)</code>                                   | Calculate non-parametric, Wilcoxon signed rank test, Wilcoxon rank sum test.                                                                                |
| Base R                         | <code>shapiro.test()</code>               | <code>shapiro.test(x)</code>                                     | Test for normality with Shapiro-Wilk.                                                                                                                       |
| Base R                         | <code>ks.test()</code>                    | <code>ks.test(x)</code>                                          | Test for normality with Kolmogorov-Smirnov.                                                                                                                 |
| Base R                         | <code>var.test(x, y)</code>               | <code>var.test(x, y)</code>                                      | Compare two variances with Fisher's F-test                                                                                                                  |
| Base R                         | <code>chisq.test(x, y)</code>             | <code>chisq.test(x, y)</code>                                    | Perform chi squared contingency tables and goodness of fit tests                                                                                            |
| Base R                         | <code>lm(x ~ y)</code>                    | <code>lm(x ~ y, data = df)</code>                                | Fit linear models based on a formula you provide.                                                                                                           |
| Base R <code>summary(x)</code> | <code>summary(linear_model_result)</code> | <code>summary(linear_model_result)</code>                        | Returns a summary of the values in object, including a linear model or other statistical test.                                                              |
| Base R                         | <code>glm(x ~ y)</code>                   | <code>glm(x ~ y, data = df, family = binomial())</code>          | Fit generalized linear models based on a formula you provide. Must specify the error distribution and link function using the <code>family</code> argument. |

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