

R Formula Cheatsheet

Main Terms and Special Operators

- `+` : add a variable — $\sim x_1 + x_2 \rightarrow$ include x_1 and x_2
- `-` : remove a variable — $\sim x_1 + x_2 - x_1 \rightarrow$ only x_2
- `:` : interaction only — $\sim x_1:x_2 \rightarrow$ include only $x_1 \times x_2$
- `*` : main effects + interaction — $\sim x_1 * x_2 \rightarrow$ includes x_1 , x_2 , and $x_1:x_2$
- `^` : include all interactions up to given order — $\sim (x_1 + x_2 + x_3)^2$
- `I()` : treat expression as-is — $\sim I(x_1^2) \rightarrow$ squared term
- `.` : all other variables — $y \sim . \rightarrow$ all columns except y
- `0` or `-1` : remove intercept — $\sim 0 + x_1 + x_2$

Common Formula Examples

- **Only main effects**
 $\sim x_1 + x_2 + x_3$
- **Main effects and interactions**
 $\sim x_1 * x_2 \#$ same as $\sim x_1 + x_2 + x_1:x_2$
- **Only the interaction**
 $\sim x_1:x_2$
- **Quadratic term**
 $\sim x_1 + I(x_1^2)$
- **All main effects and all two-way interactions**
 $\sim (x_1 + x_2 + x_3)^2$
- **Up to three-way interactions**
 $\sim x_1 * x_2 * x_3$
- **No intercept**
 $\sim 0 + x_1 + x_2$
- **Use all variables automatically**
 $\sim .$

Tip: Use `model.matrix(~ formula, data)` to see how variables are expanded in design matrices.