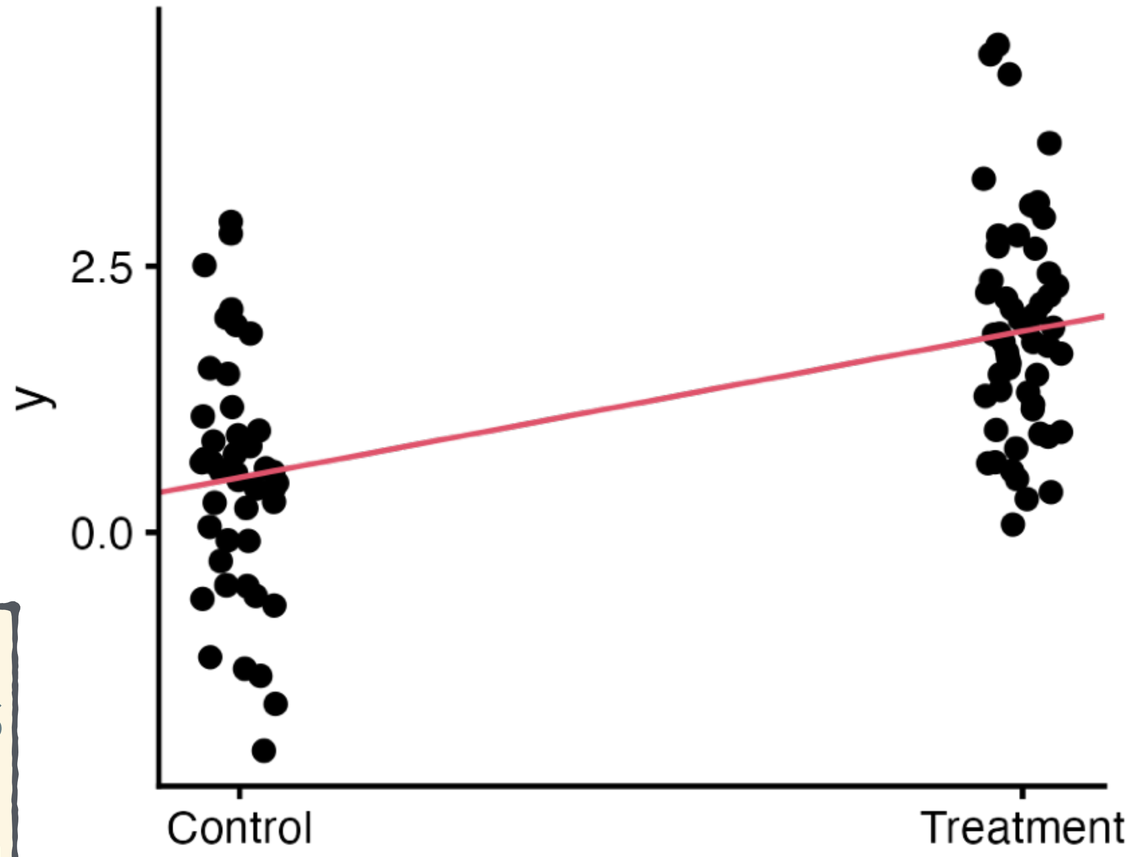
Binary predictor, same model

Control-treatment, two categories...

$$y_i \sim N(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta x_i$$

- x_i: 0 for control, 1 for treatment
- α: the intercept is the **mean of the Control** group
- β: the slope is the **difference in the means across the groups**

```
> precis(fit, prob = 0.95)
      mean   sd 2.5% 97.5%
a     0.52 0.16 0.20  0.82
b     1.37 0.22 0.96  1.79
sigma 1.08 0.08 0.94  1.23
```



How about more categories?

- There are a few ways of modeling predictors with many categories:
  - **Contrasts**: Each category is compared to a **baseline**, and the coefficients are comparisons between baseline and levels
  - **One-hot**: coefficients are means of each level of the predictor
  - **Residuals**: an overall mean is measured, and coefficients are differences between each level and the global mean

```
> x = sample(LETTERS[1:3], 9, replace = TRUE)
> y = 1 + ifelse(x == "A", 0,
           ifelse(x == "B", 1, 2)) + rnorm(9)
> df = tibble(y, x)
> df
# A tibble: 9 × 2
          y x
      <dbl> <chr>
1   1.25    B
2   0.870   A
3  -0.00180 A
4   1.18    B
5   2.03    C
6   2.60    B
7   2.55    B
8   3.92    C
9   4.66    B
```