Suppose we measured for 30 subjects ($p = 1, 2, \ldots, 30$) their average heart rate (h_p) and the amount of coffee they drank (c_p) during a day. We can do a *linear regression* between these two variables:

$$h_p = \beta_0 + \beta_1 c_p + \epsilon_p$$

This is a very cool model with two parameters: β_0 indicating an intercept and β_1 indicating a slope.

There are several reasons for doing a linear regression:

- We want to see if coffee can predict someones heart rate
- We want to make a nice picture
- We want to get a significant p-value

We can easily fit this model using the `lm()` function in `R`:

```
fit <- lm(h ~ c)
```

By running `coef(fit)` we can look at the estimates for β_0 and β_1. We can also look at `anova(fit)` to look at if the regression is significant.

Linear regression is cool, but do remember this quote by George Canning:

> I can prove anything by statistics except the truth.