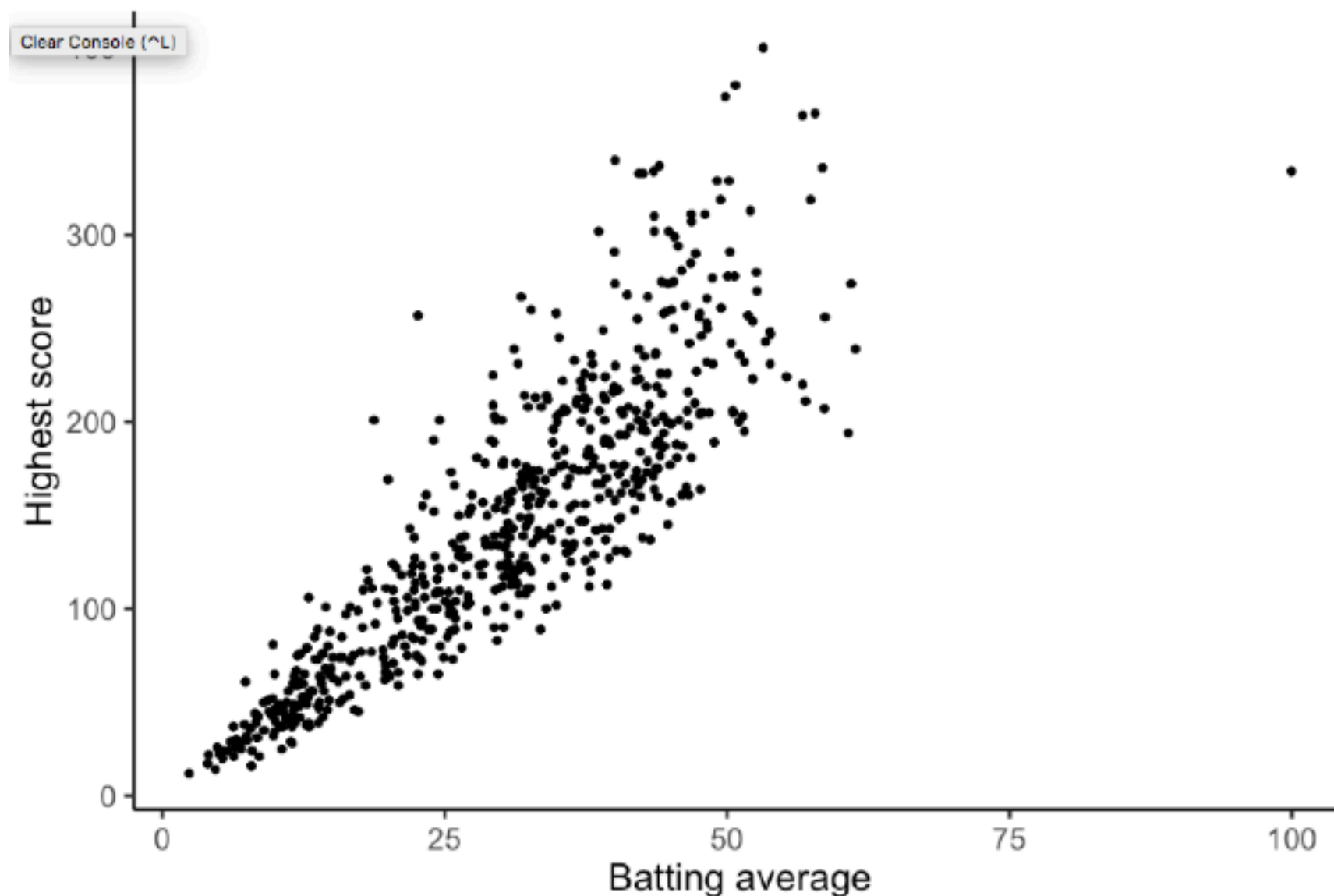


Let's Take A Step Back First

- ▶ What if we want to predict the **highest number of runs** by **batting average**?



Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

- ▶ β_0 and β_1 are both real numbers, to be estimated from the data
- ▶ ϵ_i are the random errors