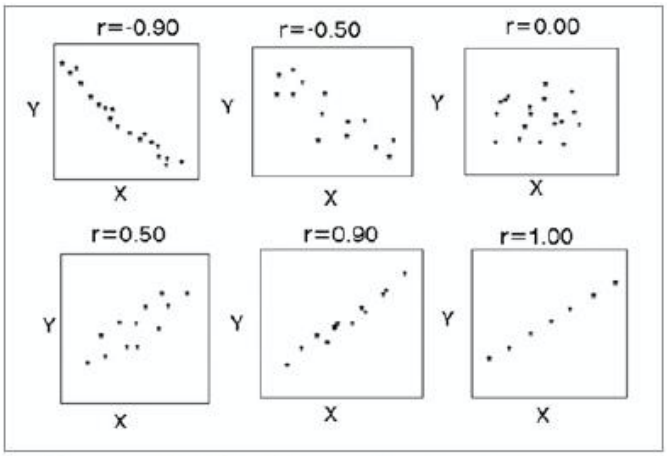
Correlation analysis
a statistical method used to evaluate the strength of relationship between two quantitative variables.

Pearson's correlation coefficient

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

- $-1 \leq r \leq 1$
- $r > 0$: Positive correlation (if x increases, so does y.)
- $r < 0$: Negative correlation (if x increases, y decreases.)
- The closer to either 1 or -1 the value r is, the more correlated x and y are.



Plot the graph and calculate the correlation coefficient.

beers	5	2	9	8	3	7	3	5	3	5
bal	0.10	0.03	0.19	0.12	0.04	0.095	0.07	0.06	0.02	0.05

linear regression is a linear approach to modeling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables)