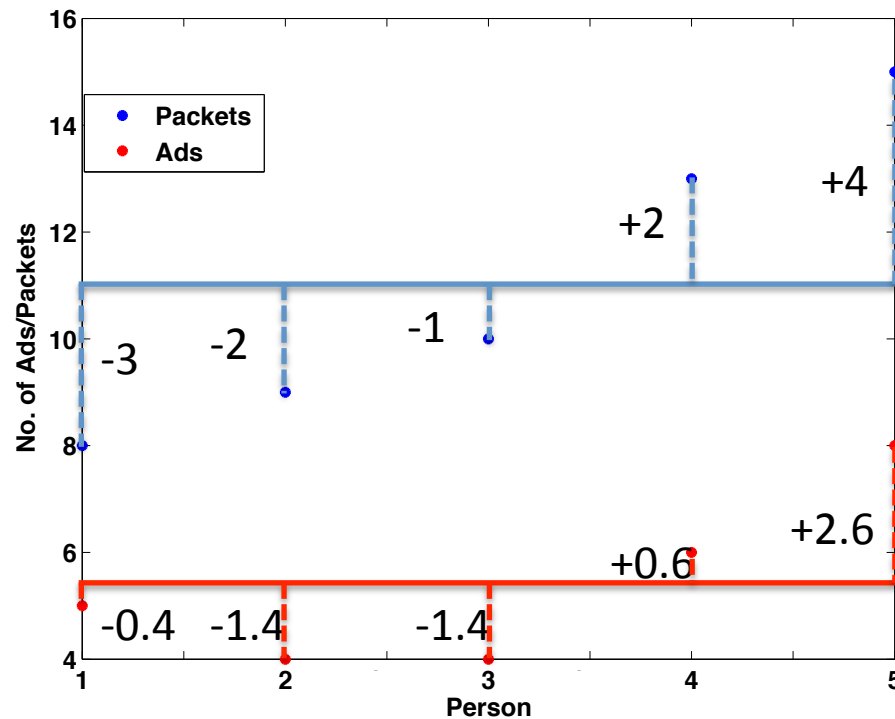# Mean, Variance, Correlation
Andy Field: Discovering Statistics Using
IBM SPSS Statistics p.262 -292

Ads and Pralinés

$$\bar{y} = \tfrac{1}{5}\sum_{i=1}^{5} y_i = 11.0$$

$$\bar{x} = \tfrac{1}{5}\sum_{i=1}^{5} x_i = 5.4$$

Participants	1	2	3	4	5
Adverts Watched	5	4	4	6	8
Packets Bought	8	9	10	13	15

Covariance:

variance

$$s_X^2 = \frac{1}{I-1}\sum_{i=1}^{I}(x_i - \bar{x})^2$$

$$\text{cov}(x,y)$$

$$= \frac{1}{I-1}\sum_{i=1}^{I}(x_i - \bar{x})(x_i - \bar{x})$$
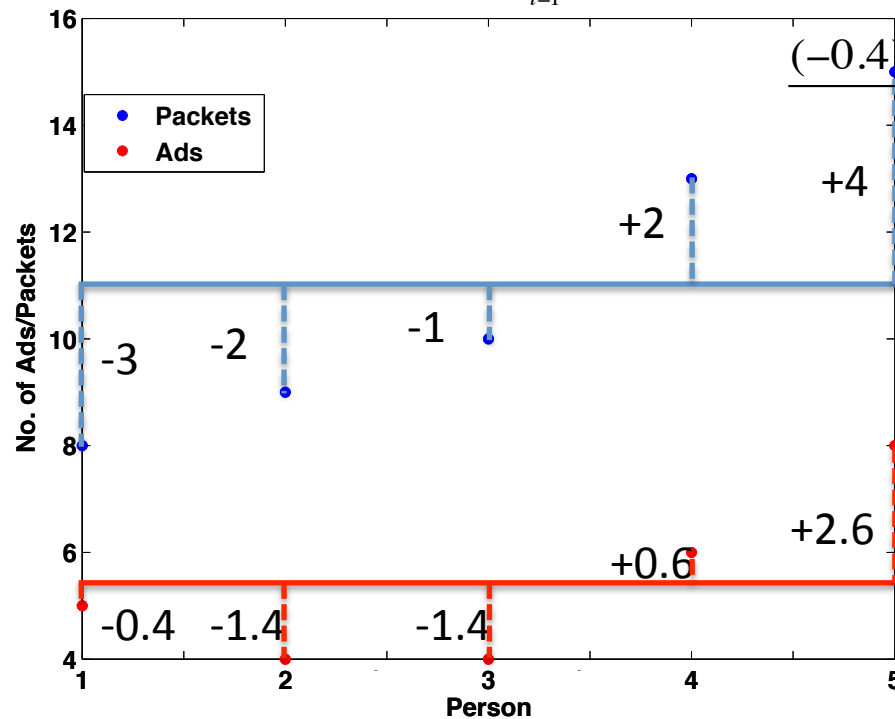
$$= \frac{1}{I-1}\sum_{i=1}^{I}(x_i - \bar{x})(y_i - \bar{y}) =$$

$$\frac{(-0.4)(-3)+(-1.4)(-2)+(-1.4)(-1)+(0.6)(2)+(2.6)(4)}{4}$$

$$= \frac{1.2+2.8+1.4+1.2+10.4}{4}$$

$$= \frac{17}{4}$$

$$= 4.25$$

Participants	1	2	3	4	5
Adverts Watched	5	4	4	6	8
Packets Bought	8	9	10	13	15

Standardization and Correlation Coefficient

Sample standard deviations:

$$s_X = \sqrt{\frac{1}{I-1}\sum_{i=1}^{I}(x_i - \overline{x})^2}$$ =1.67

$$s_Y = \sqrt{\frac{1}{I-1}\sum_{i=1}^{I}(y_i - \overline{y})^2}$$ =2.92

- $-1 \leq r \leq 1$
- r close to 1: (positive) correlation
- r close to -1: negative/anticorrelation
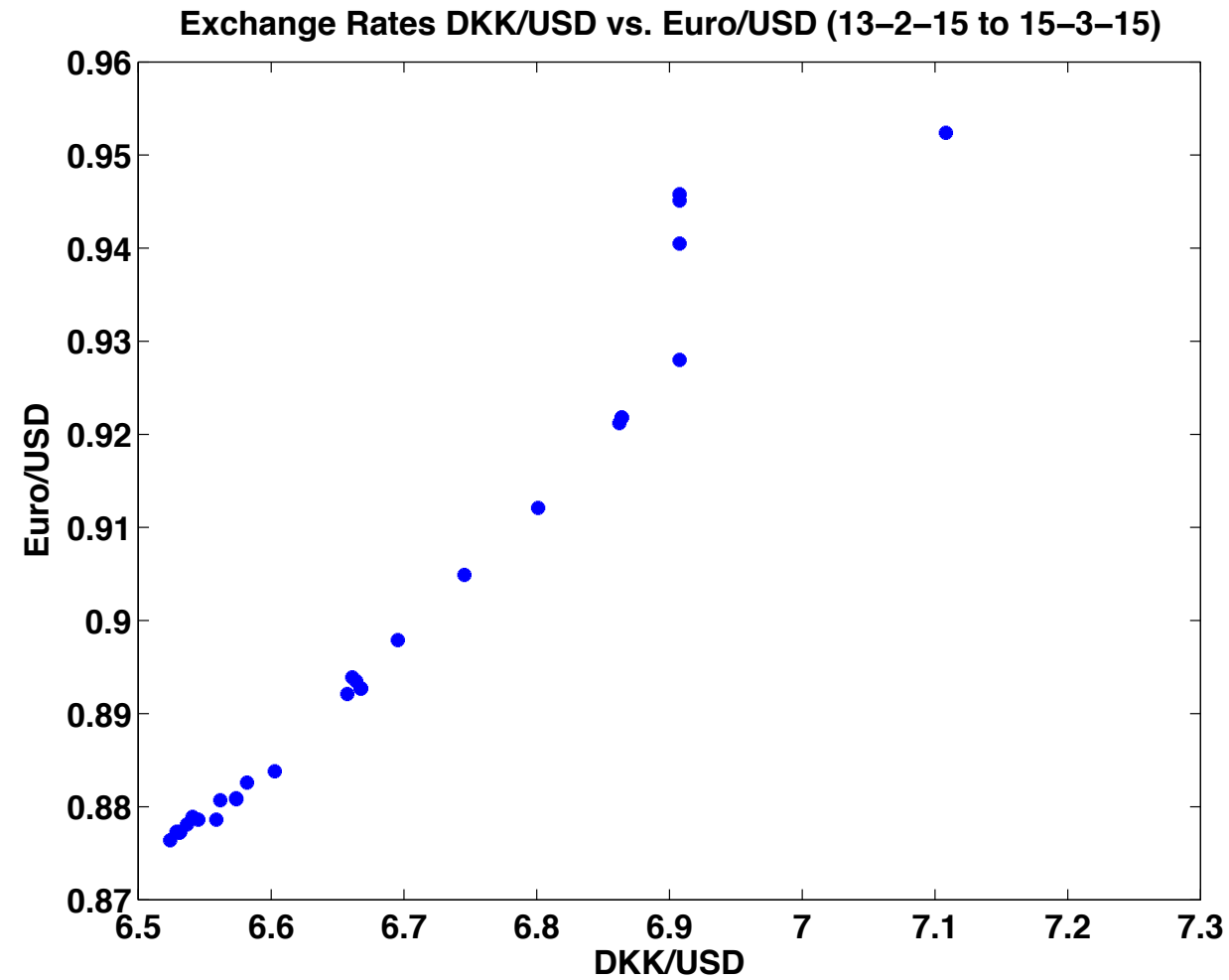- r close to 0: no correlation

Pearson correlation coefficient:

$$r = \frac{COV(x,y)}{s_X s_Y} \qquad = \frac{4.25}{1.67 \cdot 2.92} = .87$$

`r=corr(x,y)`

Measure of effect size

Exchange Rates DKK/USD vs. Euro/USD (13–2–15 to 15–3–15)

r=0.98

Math Scores and LSD Consumption
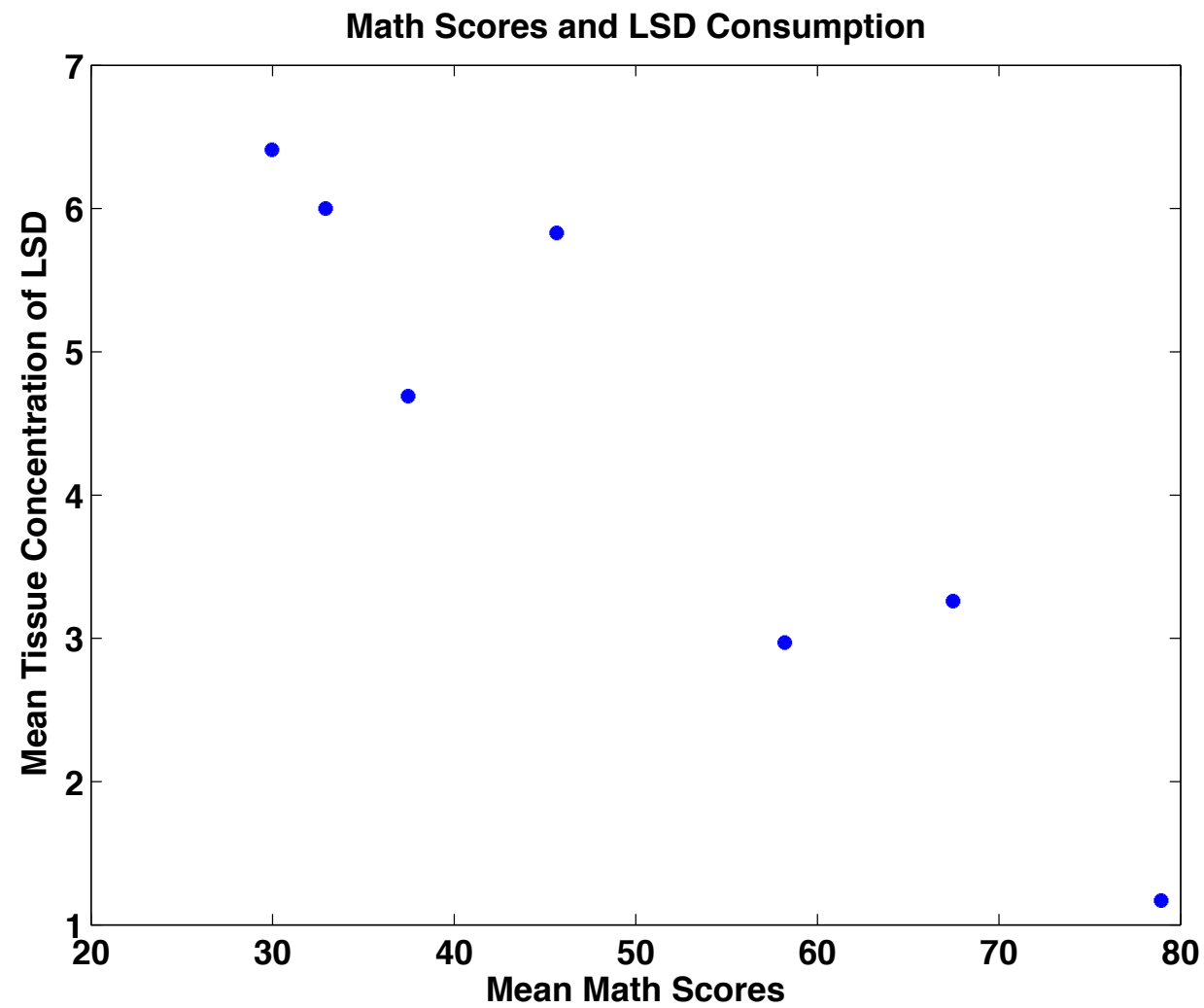
r=-0.97

Mean Monthly Temp (CPH) and Gold Prices

r=0.40