

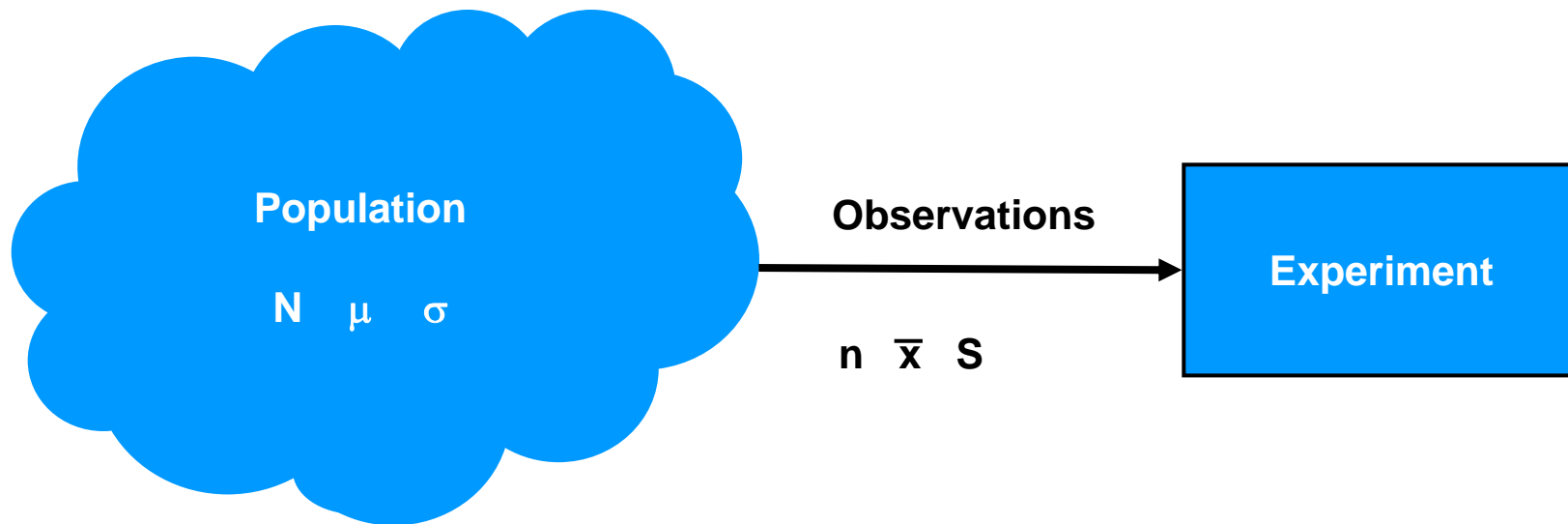
Design IV

E232 Spring 07

Class 15

Bruce McNair
bmcnair@stevens.edu

Parameter Estimation



Usually, $n \ll N$

so $\bar{X} \sim \mu$

$S \sim \sigma$

But if $\delta = |\bar{X} - \mu|$,
how large is δ likely to be?

Confidence Interval and Confidence Level

- For a population having a mean μ , the observed mean of n samples measured in one experiment is \bar{x} . The confidence interval, i.e., the region within δ of μ is:

$$\bar{x} - \delta \leq \mu \leq \bar{x} + \delta$$

- If α is the probability that the observed mean will not be within δ of μ , the confidence level is:

$$1 - \alpha = P(\bar{x} - \delta \leq \mu \leq \bar{x} + \delta)$$

Computing Confidence Interval

- Define statistic z :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

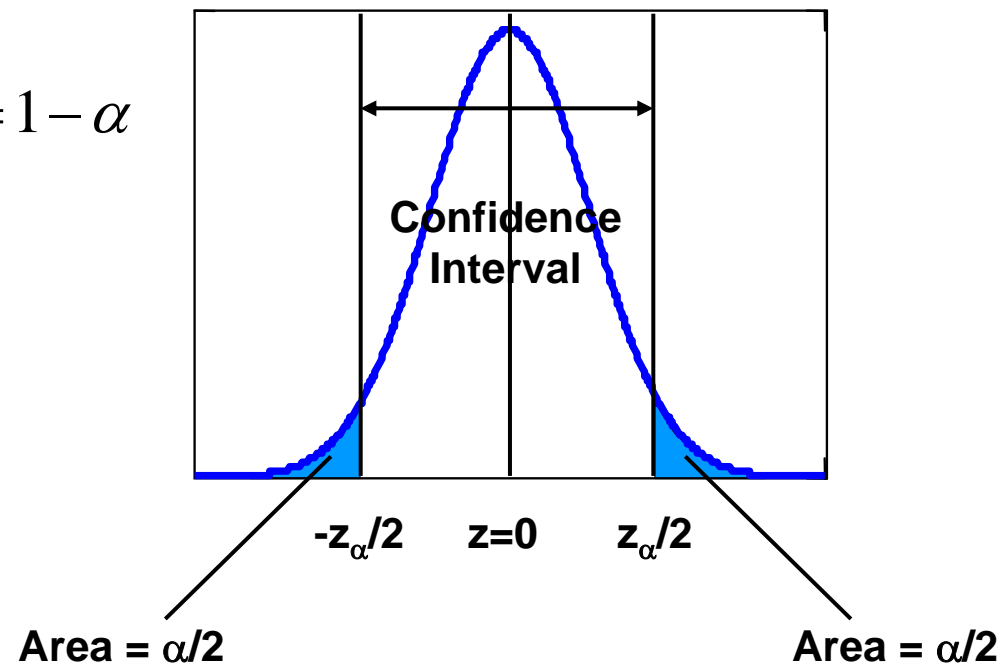
z is normally distributed with zero mean

$$P(-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \leq z_{\alpha/2}) = 1 - \alpha$$

Or,

$$\mu = \bar{x} \pm \left(z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$

with confidence $1-\alpha$



Confidence Interval Example

- We need to know the melting point of an organic compound being manufactured by a chemical process, but the results depend on the specific composition, which varies randomly. 50 samples are tested, and the average melting point is found to be 80 °C with a standard deviation of 3 °C. What is the 98% confidence interval for the average melting temperature?
 - 50 samples are enough to assume a Gaussian distribution of results.
 - 98% confidence implies $\alpha = .02$
 - We need to find $z_{\alpha/2}$ such that

$$\int_0^{z_{\alpha/2}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} = .5 - \frac{\alpha}{2} = .49$$

- Use Table 6.3 with a value .49
 $z_{\alpha/2} = 2.326$
- Find confidence interval:

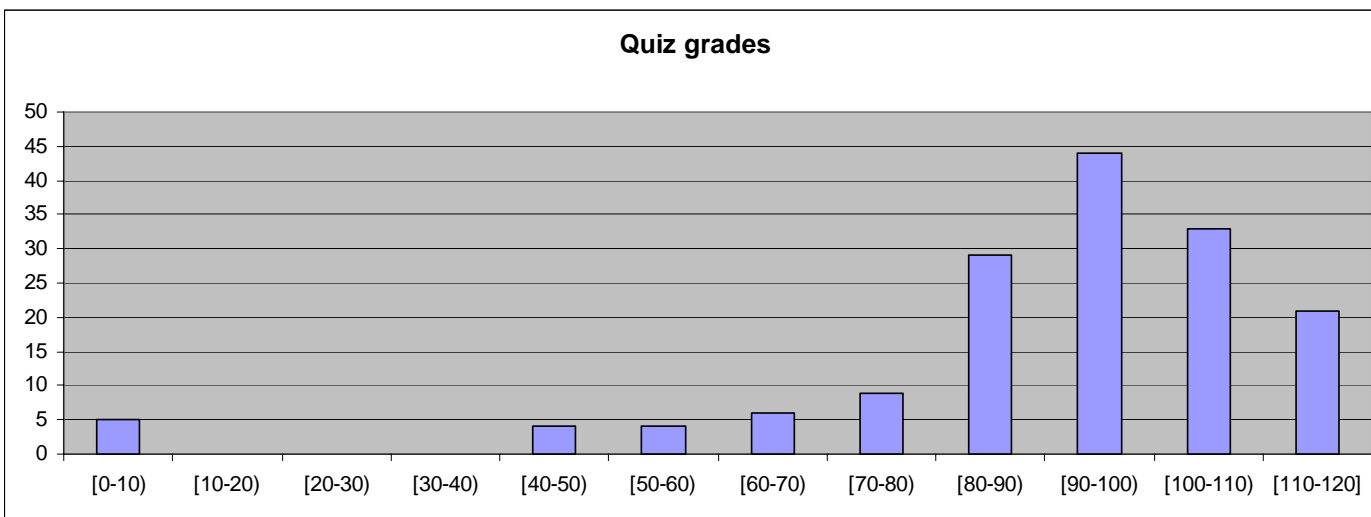
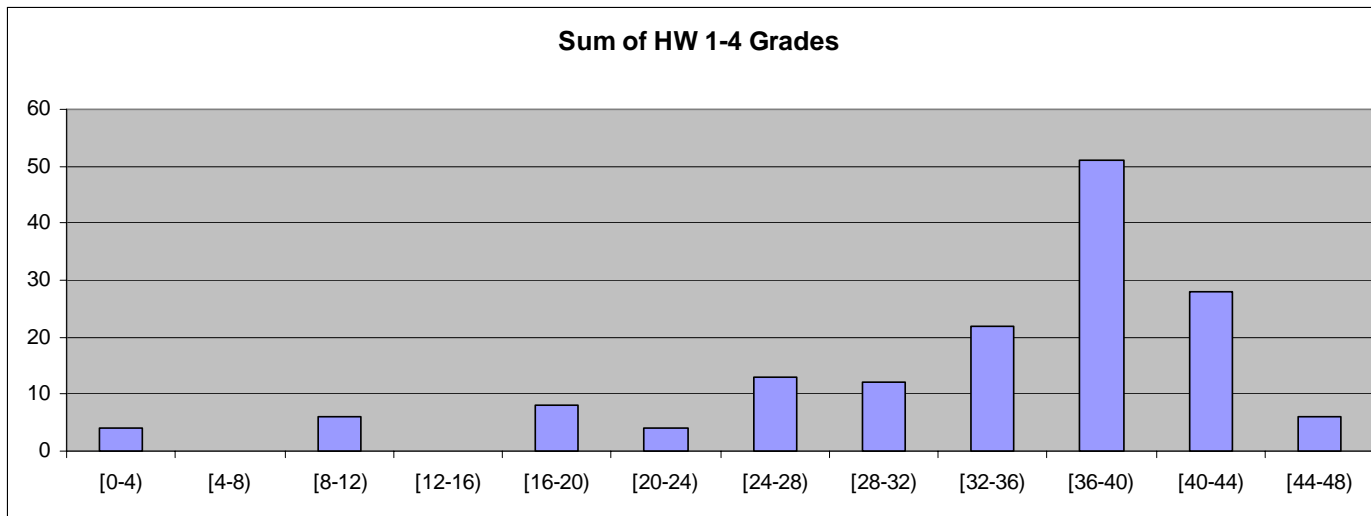
$$\bar{x} - \frac{z_{\alpha/2}S}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{z_{\alpha/2}S}{\sqrt{n}} = 80 - \frac{2.326 \cdot 3}{\sqrt{50}} \leq \mu \leq 80 + \frac{2.326 \cdot 3}{\sqrt{50}}$$
$$79.013 \leq \mu \leq 80.987$$

Today's topics

- More On Statistical Analysis of Experimental Data
 - Correlation

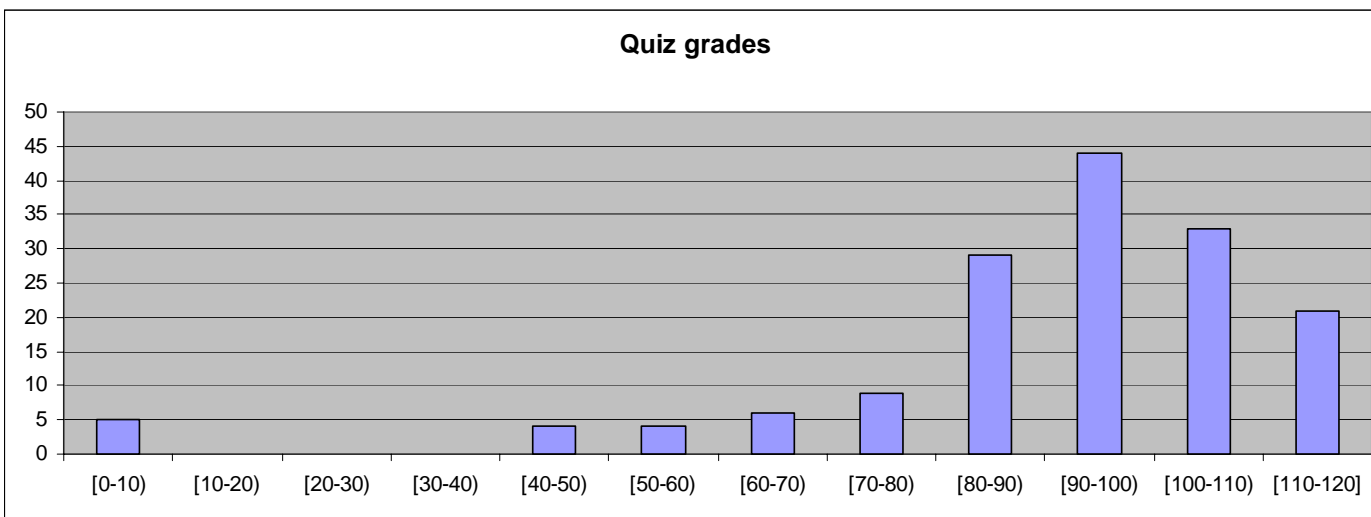
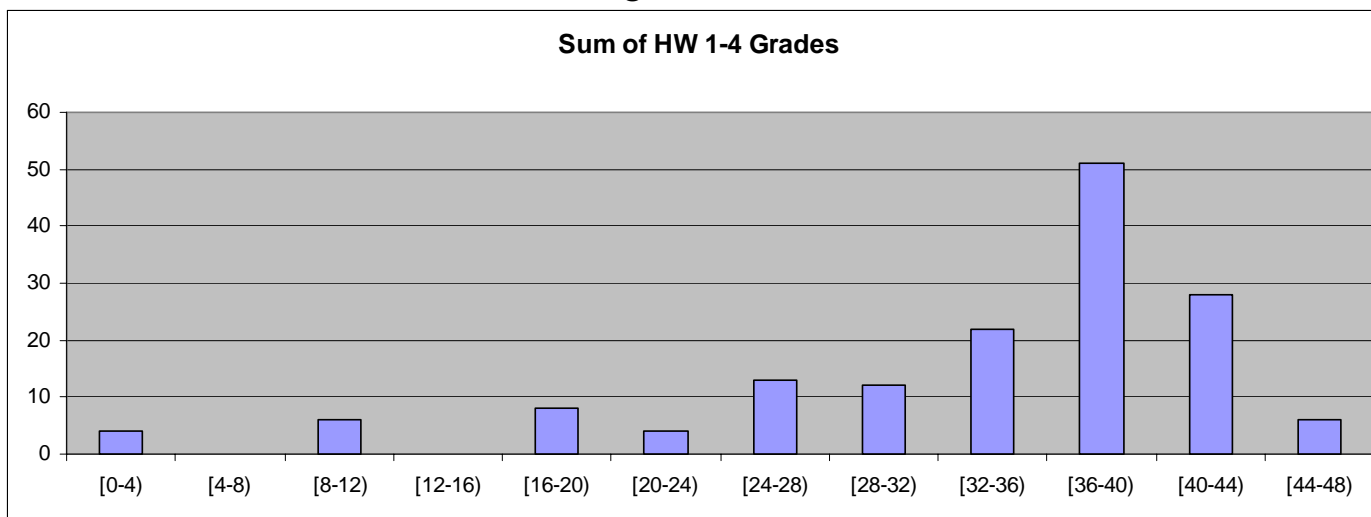
Identifying Relationships Between Data

- Consider the Homework and Quiz1 grades in E232:

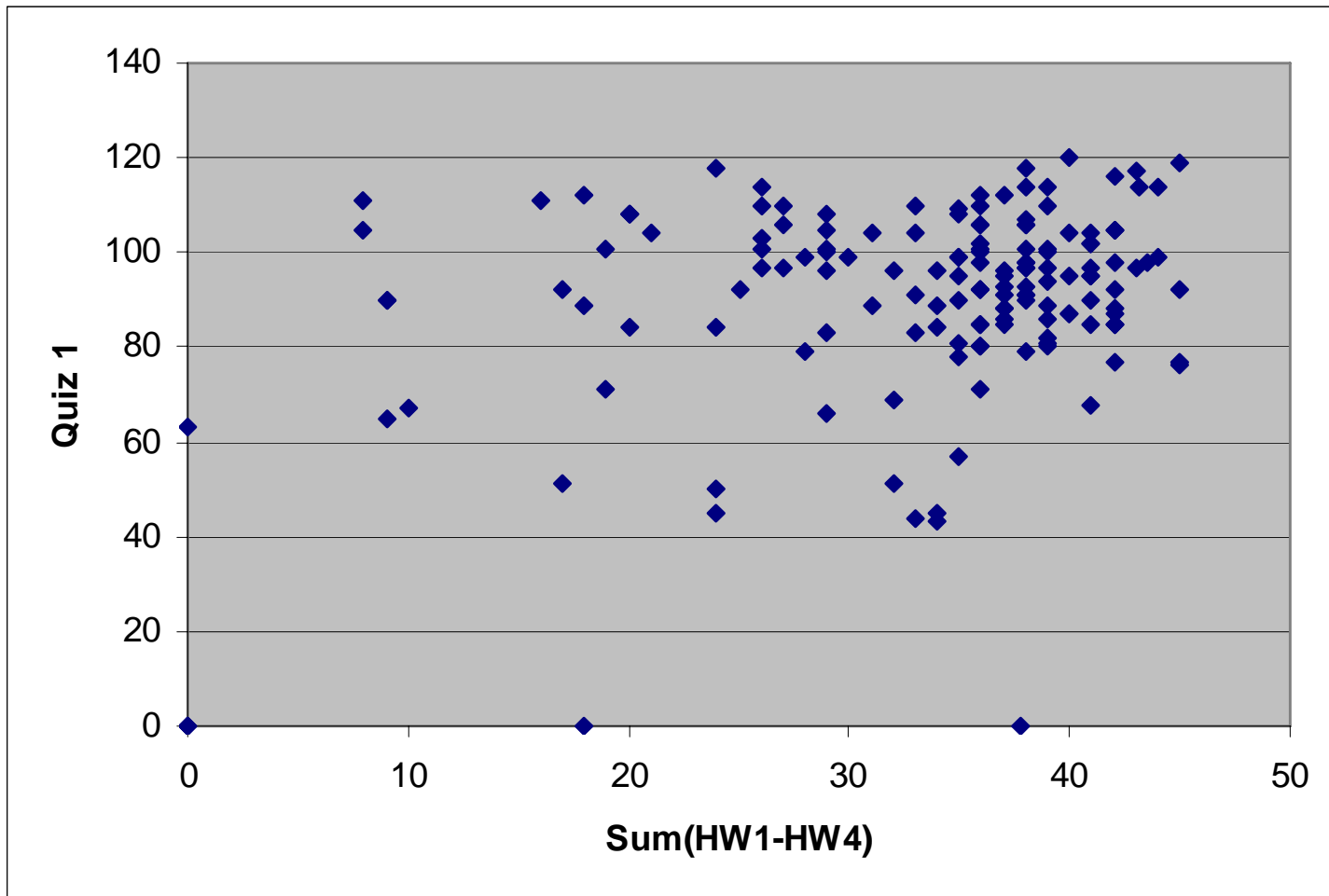


Identifying Relationships Between Data

- Consider the Homework and Quiz1 grades in E232: How are these two related?

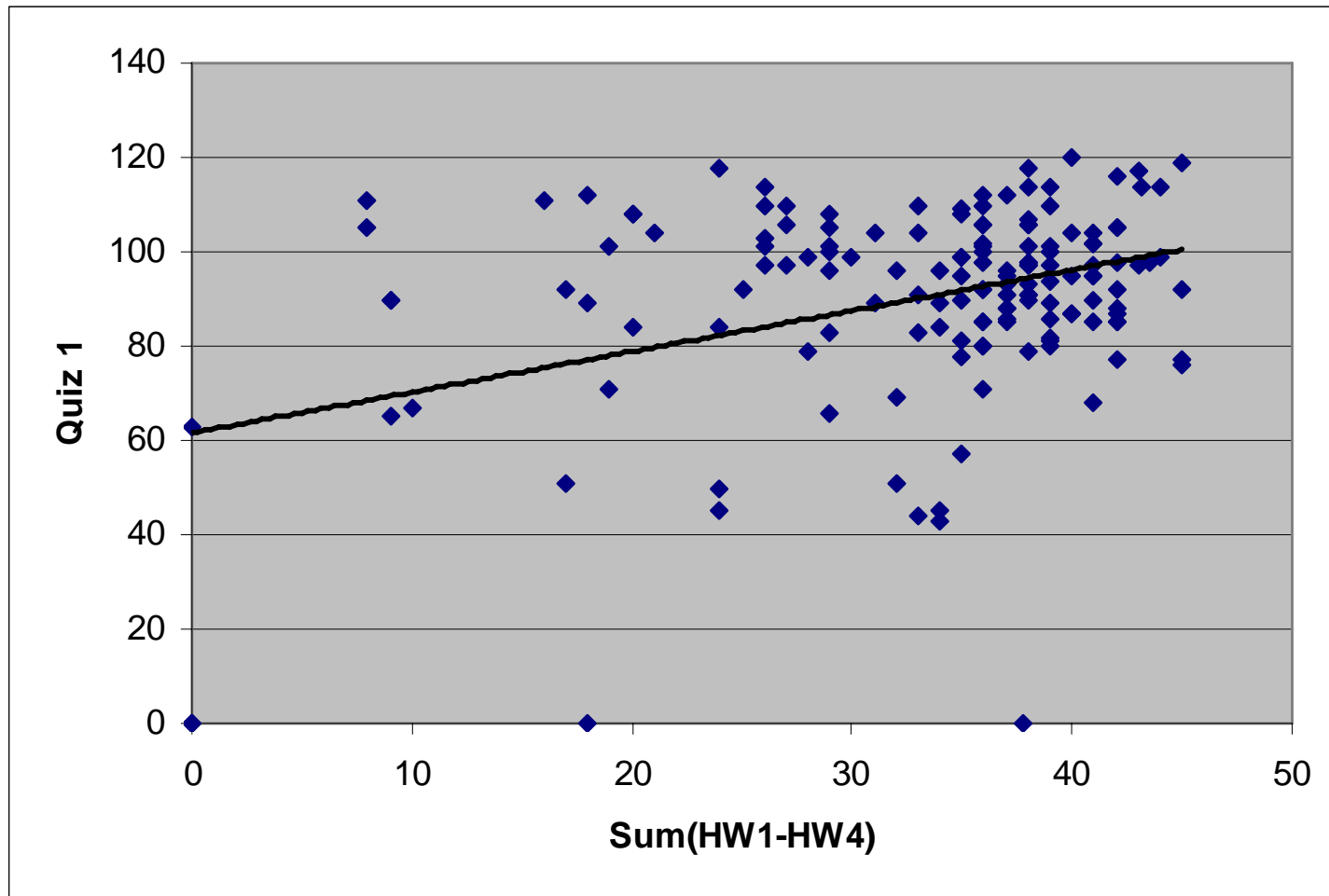


X-Y Scatter Plot



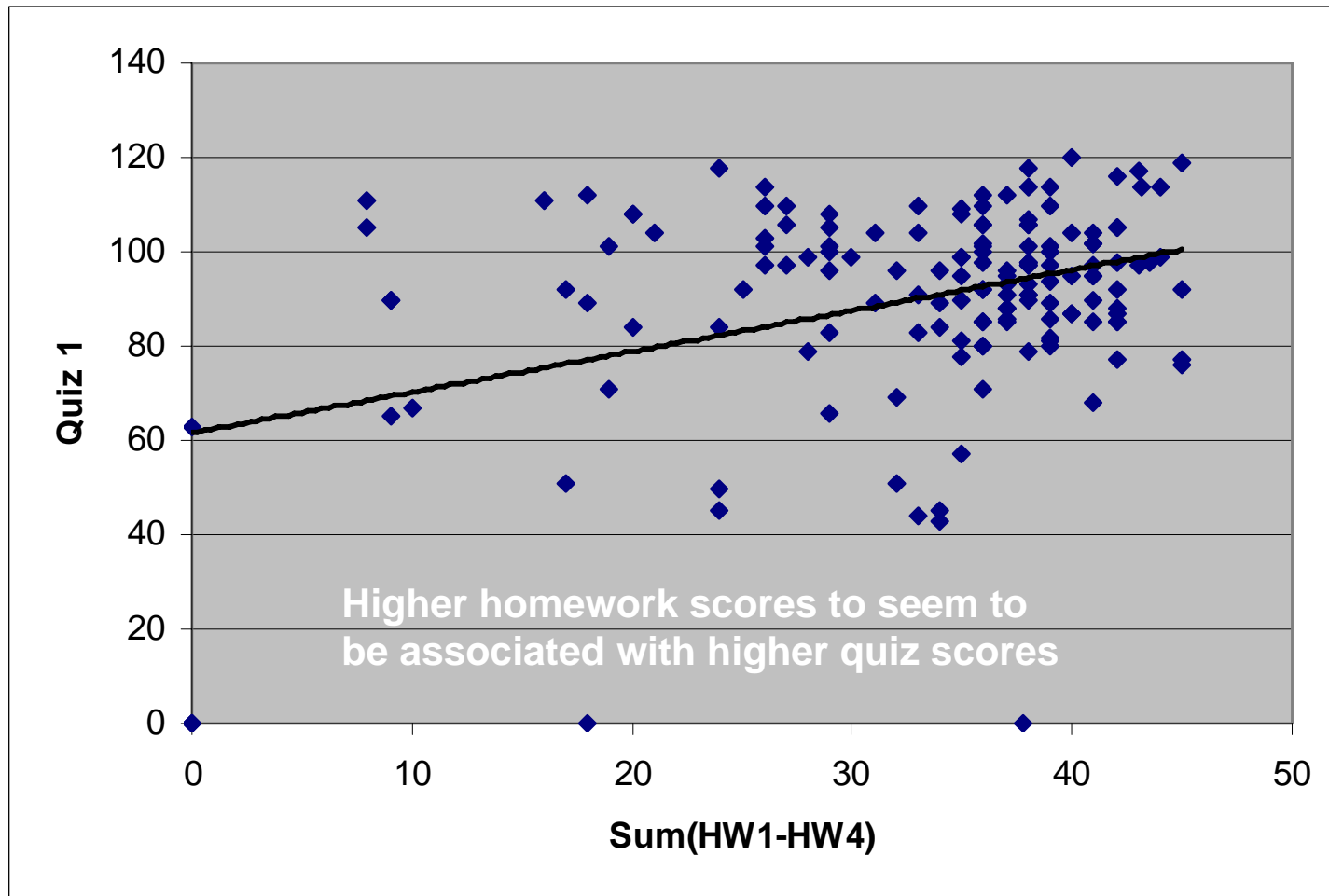
X-Y Scatter Plot

- Fit a “trendline” to available data



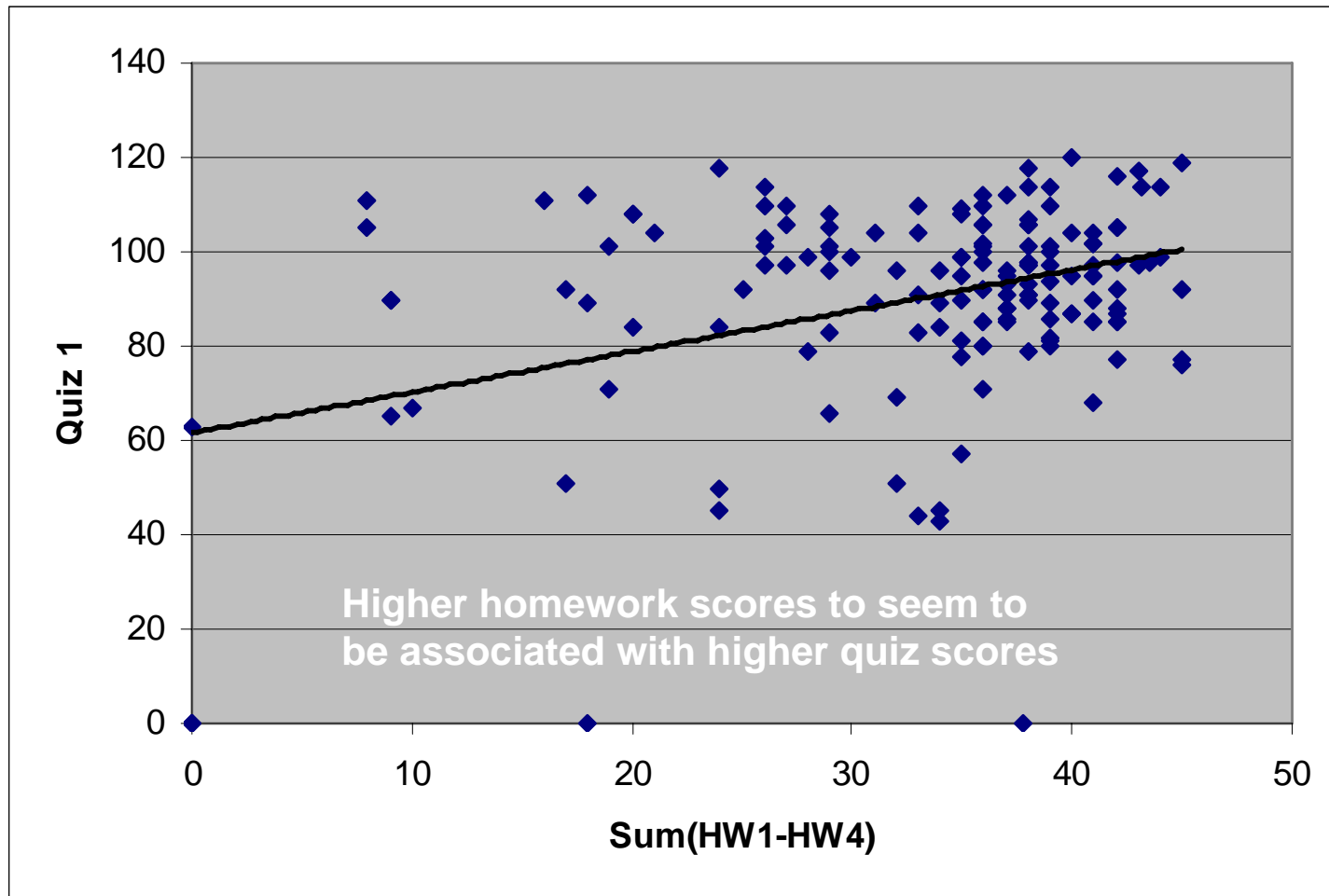
X-Y Scatter Plot

- Fit a “trendline” to available data



X-Y Scatter Plot

- Fit a “trendline” to available data – how can we compute similarity of data values?

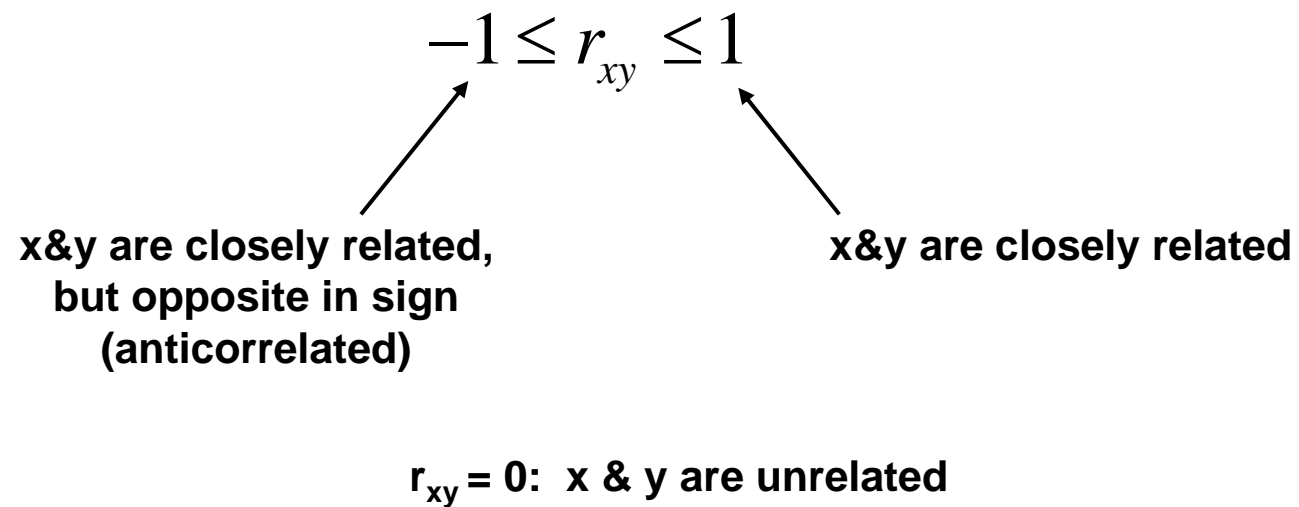


Correlation Coefficient

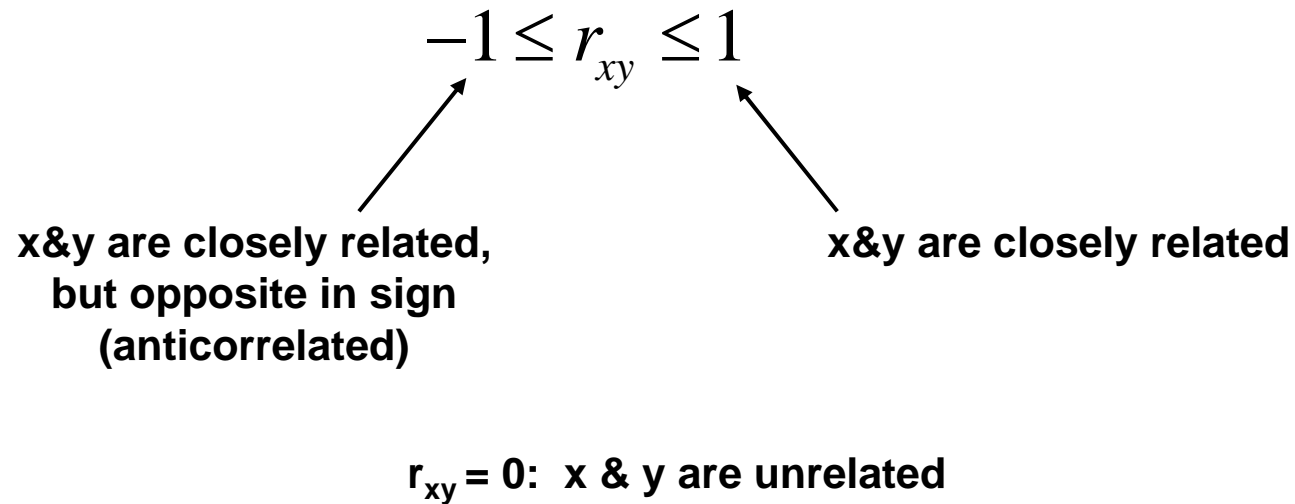
- x_i and y_i are random variables (x =HW grades, y =quiz grades)

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum (x_i - \bar{x})^2\right) \cdot \left(\sum (y_i - \bar{y})^2\right)}}$$

Characteristics of Correlation Coefficient

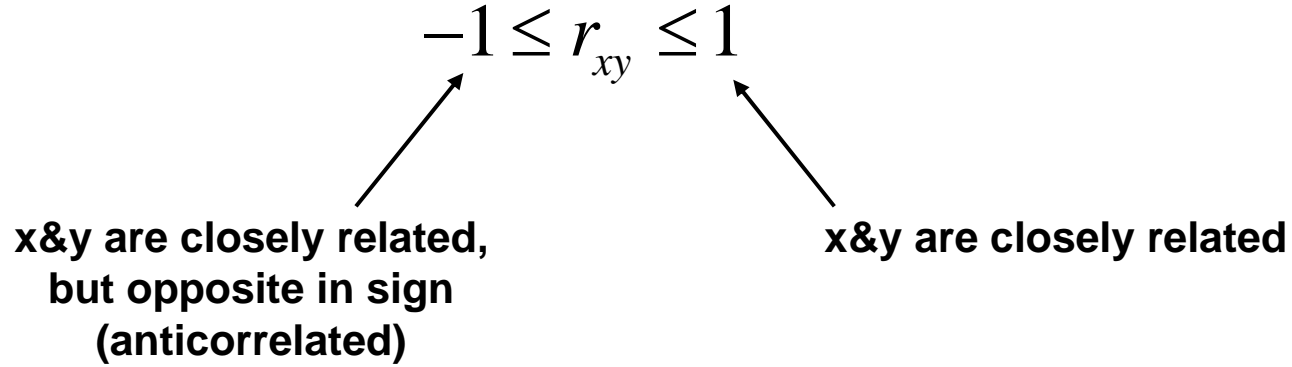


Characteristics of Correlation Coefficient



For HW1-4 and Q1 data, $r_{xy} = .38$

Characteristics of Correlation Coefficient



$r_{xy} = 0$: x & y are unrelated

For HW1-4 and Q1 data, $r_{xy} = .38$

Does this indicate significant correlation?

Next time

- Measurement sensors