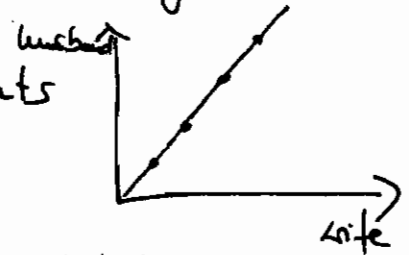


WORKSHEET - CORRELATION

1. If women always married men who were two years older than themselves, what would be the correlation between the ages of husband and wife? (Hint: Draw a scatterplot for the ages of husbands and wives when the wife is 20, 30, 40, and 50 years old.)  
Also, give the prediction line.

If a husband's age is  $y$  and the wife's age is  $x$ , the linear relationship would be

$y = x + 2$  and  $r = 1$  as all points fall on a straight line



2. For each of the following pairs of variables, would you expect a substantial negative/positive correlation, a moderate negative/positive correlation, or a small negative/positive correlation?

(a) The age of a secondhand car and its price.

moderate - substantial negative

(b) The weight of a new car and its overall miles-per-gallon rating.

moderate negative

(c) The height and the weight of a person.

moderate positive

(d) The height of a person and the height of the person's father.

moderate positive

(e) The height and the IQ of a person.

small correlation

3. Each of the following statements contains a blunder. Explain in each case what is wrong.

(a) There is a high correlation between the sex and income of American workers.

gender is a categorical variable, correlation can only be computed for quantitative variables

(b) Since student ratings of professors' teaching and colleagues' ratings of their research have a correlation of  $r = 1.21$ , the better teachers also tend to be better researchers.

we have  $-1 \leq r \leq 1$  so  $r = 1.21$  is impossible

(c) The correlation between pounds of nitrogen fertilizer applied to the field and the bushels per acre of corn harvested was  $r = 0.63$  bushels. So applying more fertilizer increases yields.

correlation does not have any units of measurements