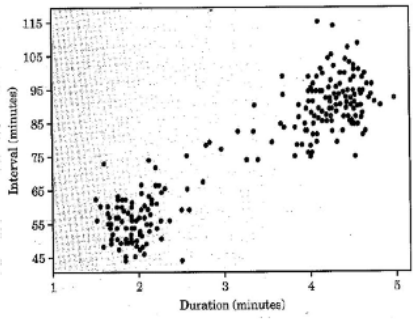
Statistical Reasoning Name:
4.1 – Scatterplots and Correlation

Read through 4.1 and answer the questions below.

A **scatterplot** shows the relationship between 2 quantitative variables measured on the same individuals. The values of one variable appears on the horizontal axis and the values of the other variable appear on the vertical axis. Each individual appears as a point in the plot.



This is a scatterplot of the interval between successive eruptions of Old Faithful. There are two groups of eruptions, one with eruptions lasting around 2 minutes and another with eruptions around 4.5 minutes.

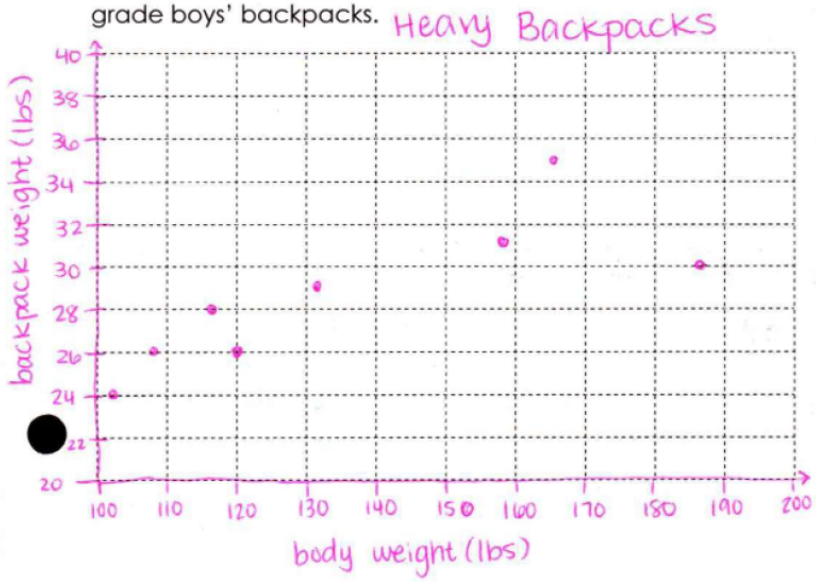
A **response variable** is a variable that measures an outcome or a result of a study.

An **explanatory variable** is a variable that ~~measures~~ we think explains or causes changes in the response variable.

The explanatory variable (if there is one) goes on which axis? x-axis

Example 4.2: Heavy Backpacks

Create a scatterplot to represent the data given to you about the weight of ninth-grade boys' backpacks.



Describe any trends that you see.

lighter kids carry lighter backpacks

as body weight ↑ backpack weight ↑

Exercise 4.1

a) explanatory: amt of time studying response: grade

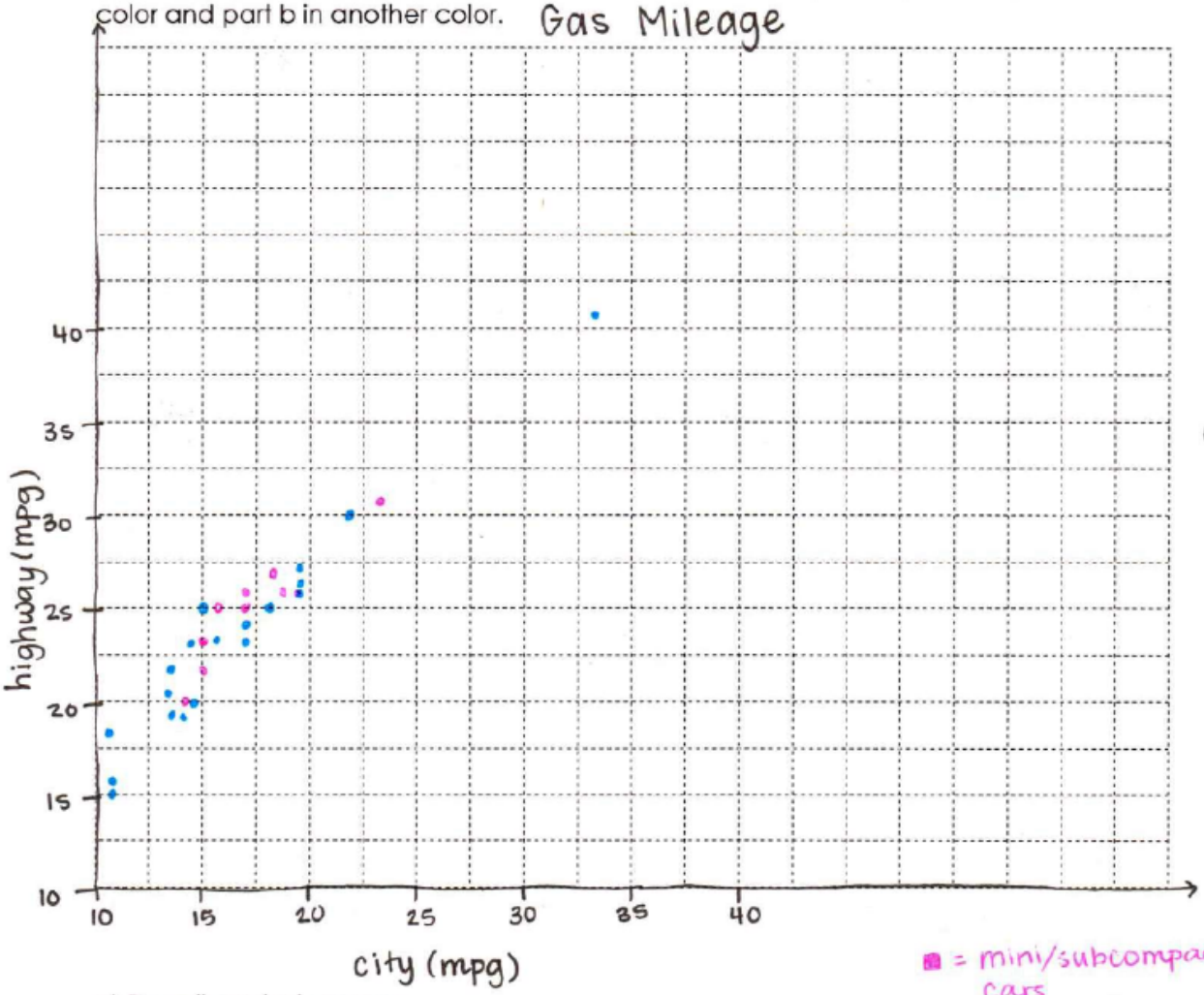
b) explore relationship

c) explanatory: amt of rain response: corn yield

d) explore relationship

Exercise 4.6

Use the graph below for parts a and b of this exercise. Graph the points of part a in one color and part b in another color.



c) Describe what you see.

- positive association (as city gas mileage increases, so does highway)
- two seater cars has one data point far from the others

Interpreting Scatterplots

Look for the overall pattern and deviations from pattern.

Describe the pattern by the direction form, and strength of the relationship.

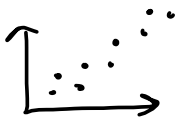
An important kind of deviation is an outliers.

What does it mean for two variables to be **positively associated?** Sketch a graph.

has a positive (upwards) slope

as x ↑, y ↑

as x ↓, y ↓



What does it mean for two variables to be **negatively associated?** Sketch a graph.

has a negative (downwards) slope

as x ↑, y ↓

as x ↓, y ↑

Correlation

The **correlation** describes the direction and strength of a straight-line (linear) relationship between two quantitative variables.

Correlation is usually written as *r*. → correlation coefficient

Correlation is always between −1 and +1. If *r* is positive, that means there is a positive association, and if *r* is negative, there is a negative association. The closer *r* is to -1 or 1, the stronger the association is, and the closer to 0 the weaker the association is.

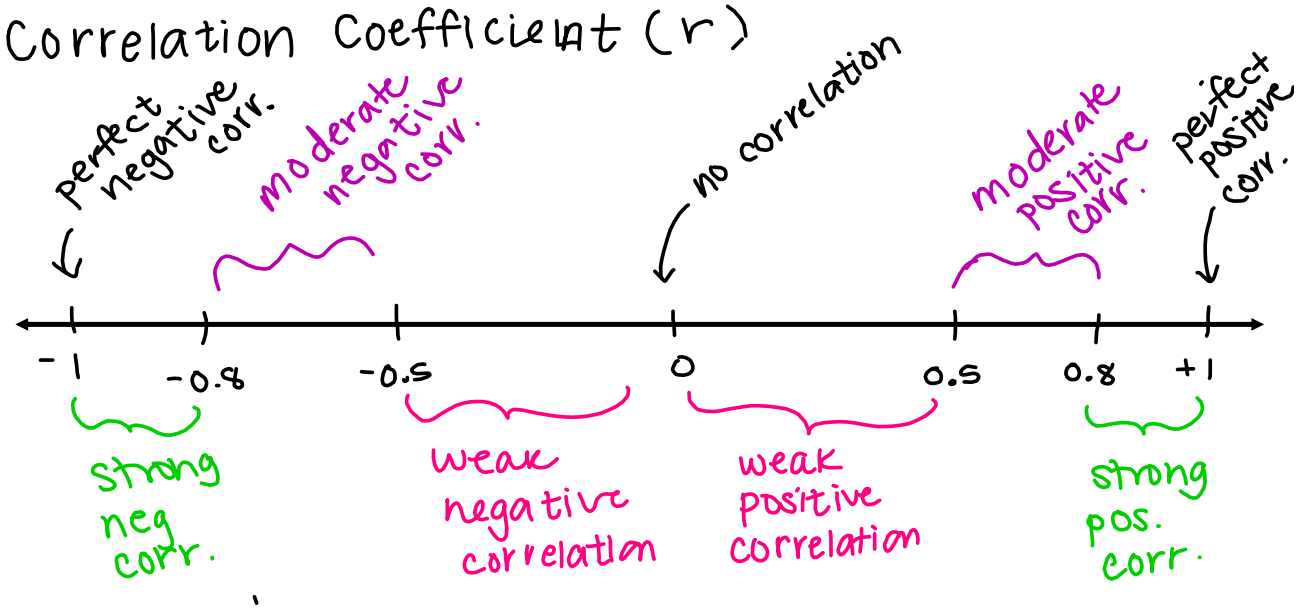
▷ Correlation does not have units, it is just a #

► Correlation does not change when we change the units of measurement.

► Correlation ignores the distinction between explanatory and response variables.

► Correlation measures the strength of only straight-line association between two variables.

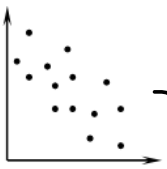
► Correlation is strongly affected by a few outlying observations (outliers).

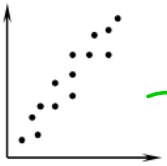

Correlation Coefficient (r)
perfect negative corr.
moderate negative corr.
no correlation
moderate positive corr.
perfect positive corr.
-1
-0.8
-0.5
0
0.5
0.8
+1
strong neg corr.
weak negative correlation
weak positive correlation
strong pos. corr.

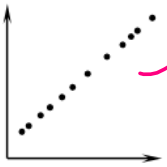
Name: _____ Date: _____

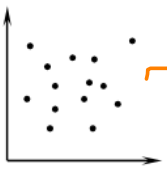
QUANTIFYING PREDICTABILITY
COMMON CORE ALGEBRA I HOMEWORK

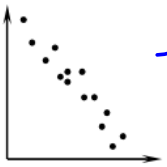
1. Below there are six scatter plots, six correlation coefficients, and six terms. Match the appropriate r-value with the scatter plot it most likely corresponds to. Then, match the term you think is most appropriate to the r-value as well (not to the graph).

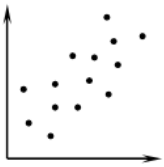
(a) 

(b) 

(c) 

(d) 

(e) 

(f) 

$r = 1.0$

$r = 0.35$

$r = -0.82$

$r = 0$

$r = -0.56$

$r = 0.93$

Weak Negative

Perfect Positive

Strong Positive

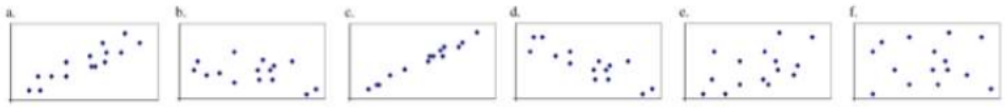
Weak Positive

Moderate Negative

No Correlation

Correlation Coefficient & Linear of Best Fit HW
Name: _____

1. List the correlation coefficients in order from least to greatest: 0.79, -0.43, -0.4, 0.82, 0.08
   -0.43, -0.4, 0.08, 0.79, 0.82
2. If the correlation coefficient is .87, is its relationship considered to be strong positive, weak positive, strong negative, weak negative, or no correlation?
   strong positive
3. If the correlation coefficient is -0.03, is its relationship considered to be strong positive, weak positive, strong negative, weak negative, or no correlation?
   no correlative or very weak neg. correlation
4. Describe the correlation between hours worked at a job and the money earned.
   positive corr. ↑ ↑
5. Describe the correlation between the age of students in the class and their height.
   no correlation
6. Describe the correlation between the value of a Honda Accord (car) and its value over time.
   negative correlation ↑ ↓
7. Match the correlation coefficients to their correct graphs. Write the letter in the space.
   -0.85 d -0.40 b 0 f 0.33 e 0.87 a 0.99 c



8. Predict the type (positive, negative, no) and strength of correlation ~~(strong, weak)~~ for the following pairs of quantities.
   a. temperature and time for a cup of coffee left on the counter. negative
   b. temperature and time for a pot of cold water placed on a hot stove to boil. positive
   c. math test score and height no correlation
   d. amount of money in the bank and the number of days since the last paycheck.
      negative correlation