Sharks and Ice Cream

Interpreting Correlations & Determining Cause and Effect Relationships

Grades: 6-8

Subject: Correlation / Cause and Effect

Overview of the lesson plan:

Students will use graphs and data from the "Let's Get Healthy" fair to identify comparative strengths of correlations and how correlations correspond to cause and effect. They will identify graphs depicting comparative strengths, analyze data from the "Let's Get Healthy" fair, and determine correlations don't always show a cause and effect.

Suggested Time: 45 minutes

Objectives:

- Students will identify graph depicting comparative strengths of correlations
- Students will analyze and interpret real data from the "Let's Get Healthy!" fair
- Students will evaluate scenarios to determine if there is a direct cause and effect correlation

Materials:

- Six scatterplot graphs: Copied and cut into 6 separate graphs and distributed to groups of 4-6 students
- Scatterplot graphs and key for teacher
- One set of "Let's Get Healthy" graphs depicting data from the fair for students to analyze and interpret
- Cause and Effect Scenarios

Activities and Procedures:

Activity 1: Identify strength of correlations

Opener: Ask students about their favorite sports athletes and teams and how they track how well their favorite athletes or team are performing. These questions will allow you to discuss the use of statistics in sports. Focus on how statistics come from raw data and another way to visualize raw data is by using scatter plots. Students will read and interpret scatter plots while participating in activities one

and two. Pass out to each small group of 4-6 students the six scatter plot cards. Ask students match the graphs to the following correlations:

- Strong/positive correlation
- Weak/positive correlation
- Strong/negative correlation
- Weak/negative correlation
- Moderate correlation
- No correlation

Discuss with students how they determined which graphs depicted which correlation. Then use the y axis and x axis descriptions found in *Scatter Plot Graphs and Key for Teacher* to connect the graphs to subject matter relating to the *Let's Get Healthy!* stations.

Teacher Note: Using plant analogies might another way to help students visualize what the data on each graph is illustrating. For instance a plant's rate of growth will have a strong/positive correlation to the amount of nutrients in its soil while the direction of the sunlight might have weak/positive correlation on the plant's rate of growth. A strong/negative rate of growth will occur by over watering a plant while irregular watering could contribute to a moderate/negative correlation. The graph depicting no correlation could explain by plants being planted in pots of different colors has no correlation to the plant's rate of growth. The horizontal zero correlation graph listed on the teacher's key might be an example of a plant's rate of growth has no correlation to the number of chairs around a dining table.

Activity 2: Gallery-walk with graphs of "Let's Get Healthy" fair data
Create gallery exhibits by attaching each *Let's Get Healthy!* graph to blank poster sized butcher paper along different walls in the classroom.

Then have students travel in their small groups to each exhibit. When visiting an exhibit, students should analyze and discuss the graph's data, and then record on the blank poster-sized paper at least one observation and one question they have from the information they can interpret from each exhibit. *Note to teacher: By giving each group a different colored marker to record their responses, you can track each group's contributions at each exhibit*.

After all the groups have visited all the exhibits, lead the students in a discussion by sharing with them their questions and comments.

> *IE – "The blue group was wondering why blood glucose correlated to BMI and how it would correlate to the amount of sleep someone gets."*
>
> *IE – "The orange group thinks that it is obvious that the percentage of fat in one's diet would correlate directly with an individual's BMI."*

This discussion should include some of the inaccurate or false connections students have made regarding the data allowing for a review of the following vocabulary and terminology:

correlation is a relationship between two quantities, such that when one changes, the other does.

- If the quantities simultaneously increase or decrease in value, a positive correlation exists.
- If one increases as the other decreases, a negative correlation exists.

A ***cause*** is something that makes something else happen. Out of two events, it is the event that happens first. To determine the cause, ask the question "Why Did it Happen?"

An ***effect*** is what happens as a result of the cause. Of two related events, it's the one that happens second or last. To determine the effect, ask the question "What Happened?"

Activity 3: Evaluating Correlation Scenarios

After reviewing ***correlation*** and ***cause and effect*** with students, pass out the following *scenarios worksheet* to them in their small groups, and have them discuss if the correlations in each scenario have a direct cause and effect. After students share their explanations with the class, go over the explanations provided in the *scenario key*.

Extension activity: Have students in their small groups revisit "gallery exhibits" and identify correlations fallacies recorded early during the *gallery walk*.

Standards

Academic Content Standards - Oregon Department of Education (Grade 8)

Health Skills: Demonstrate ability to use health skills, to obtain and interpret health information, to manage personal behaviors and to advocate for healthy and safety issues.

- HE.08.HS.03 Analyze influences on health and well-being (e.g., culture, family, media, technology, peers, body image, emotions, and physical and social environments)

Historical Skills: Analyze cause and effect relationships, including multiple causalities.

- SS.08.HS.02 Distinguish between cause and effect relationships and events that happen or occur concurrently or sequentially.

Scientific Inquiry: Analyzing Data and Interpreting Results: Analyze scientific information to develop and present conclusions.

- SC.08.SI.04 Summarize and analyze data including possible sources of error. Explain results and offer reasonable and accurate interpretations and implications.

Social Science Analysis: Identify and analyze an issue

- SS.08.SA.04 Examine the various characteristics, causes, and effects of an event, issue, or problem.

National Science and Education Standards (Grades 5-8)

Science as Inquiry - Content Standard A

- Develop descriptions, explanations, predictions, and models using evidence.
- Think critically and logically to make the relationships between evidence and explanations.
- Recognize and analyze alternative explanations and predictions.
- Communicate scientific procedures and explanations.

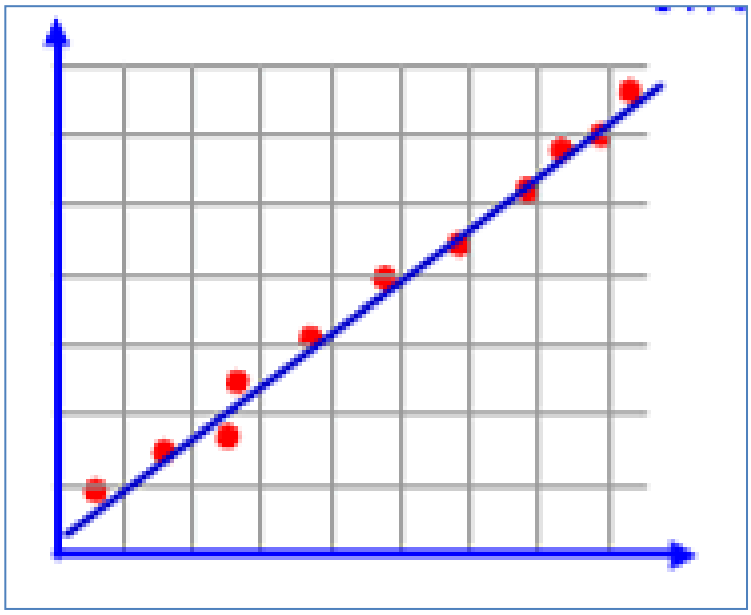
Science in Personal and Social Perspectives **-** Content Standard F: PERSONAL HEALTH

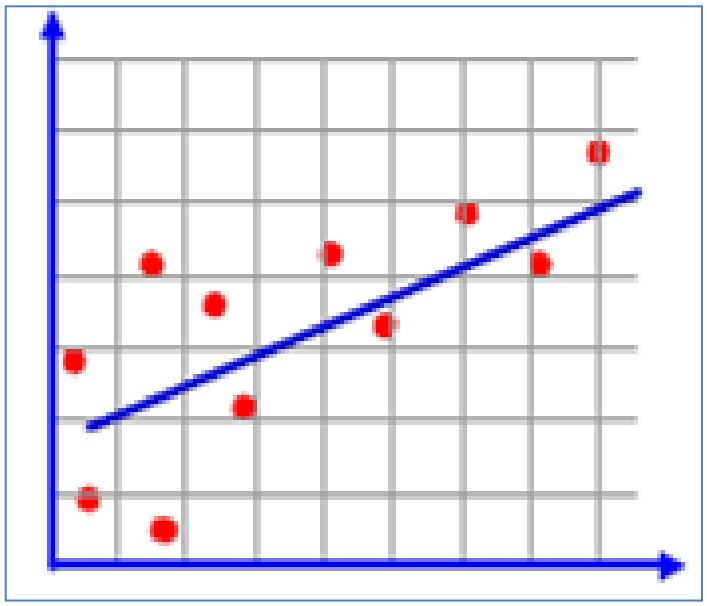
- Regular exercise is important to the maintenance and improvement of health.
- The use of tobacco increases the risk of illness.
- Food provides energy and nutrients for growth and development.

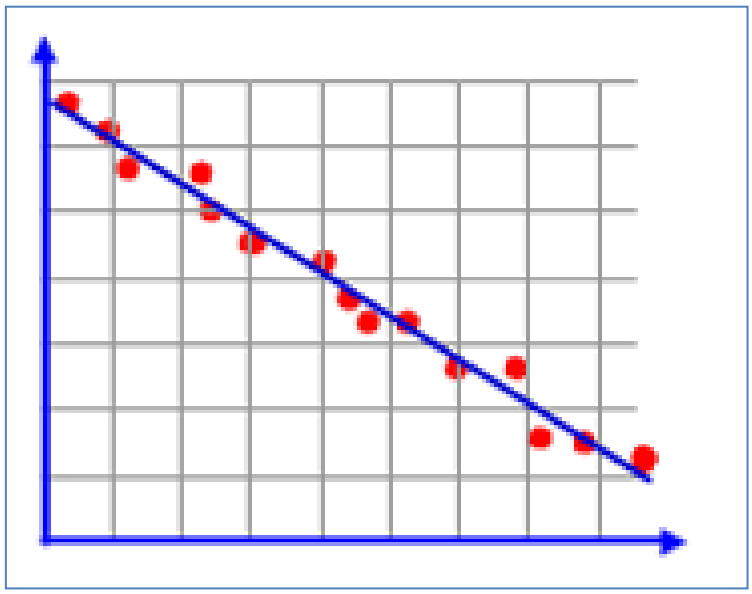
History and Nature of Science - Content Standard G

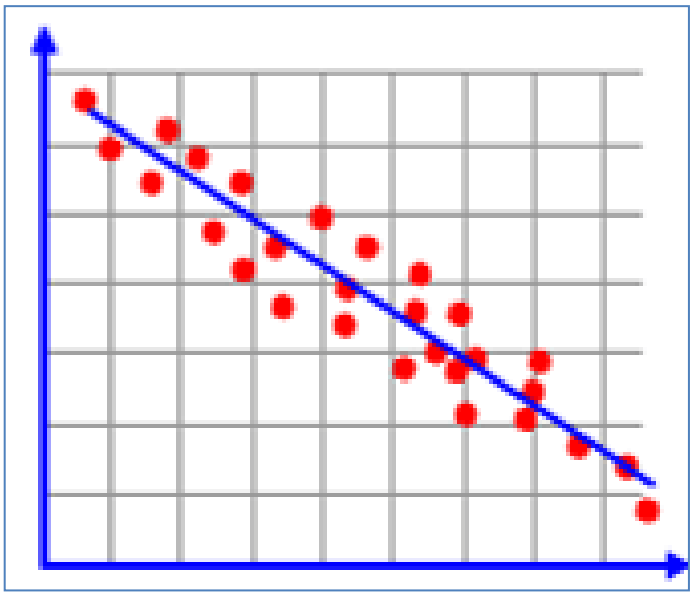
- Science As A Human Endeavor: Science is very much a human endeavor, and the work of science relies on basic human qualities, such as reasoning, insight, energy, skill, and creativity—as well as on scientific habits of mind, such as intellectual honesty, tolerance of ambiguity, skepticism, and openness to new ideas.
- Nature Of Science:
    2. In areas where active research is being pursued and in which there is not a great deal of experimental or observational evidence and understanding, it is normal for scientists to differ with one another about the interpretation of the evidence or theory being considered.
    3. It is part of scientific inquiry to evaluate the results of scientific investigations, experiments, observations, theoretical models, and the explanations proposed by other scientists. Evaluation includes reviewing the experimental procedures, examining the evidence, identifying faulty reasoning, pointing out statements that go beyond the evidence, and suggesting alternative explanations for the same observations. Although scientists may disagree about explanations of phenomena, about interpretations of data, or about the value of rival theories, they do agree that questioning, response to criticism, and open communication are integral to the process of science. As scientific knowledge evolves, major disagreements are eventually resolved through such interactions between scientists.

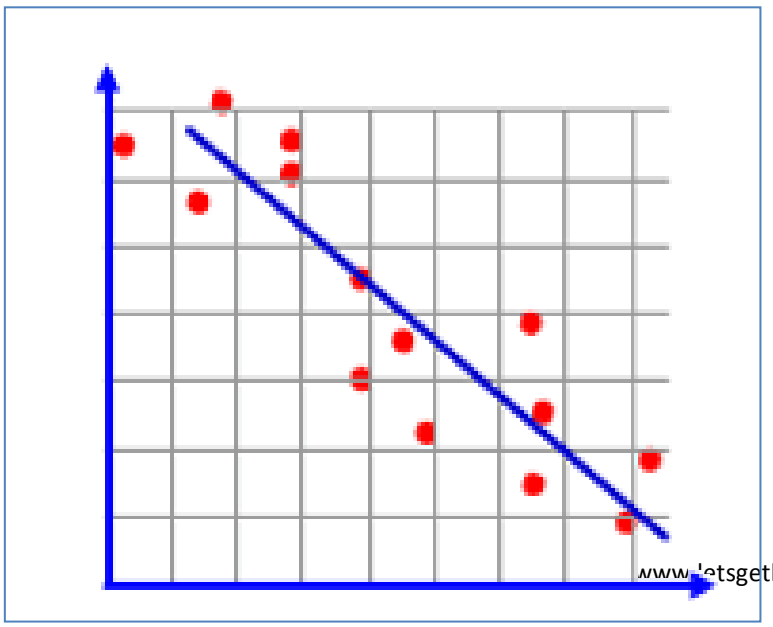
Handout for Activity 1 - *Directions: Cut out graphs and distribute to individual small groups of 4-6 students*

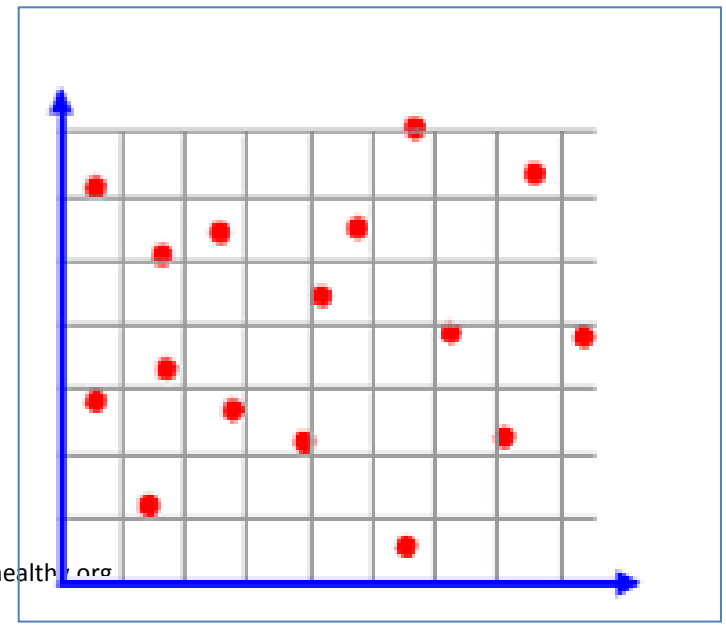












SCATTER PLOT GRAPHS AND KEY FOR TEACHER

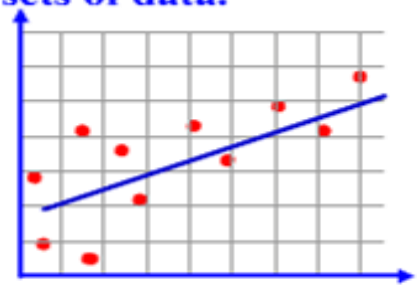
SCATTERPLOTS & CORRELATION

Correlation - indicates a relationship (connection) between two sets of data.

Strong positive correlation

Figure 1: It is easy to see that as you increase one variable, the other increases as well. A good example of this would be weight BMI because BMI is calculated based on weight. So as weight goes up, so should BMI. The data points stay close to the line. This is the kind of "tight" data scientists want.

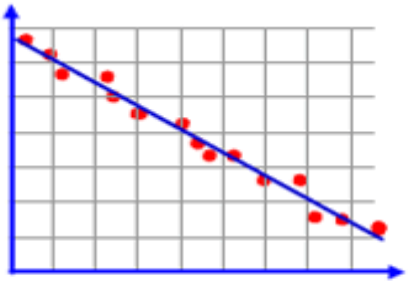
Y axis = Body Mass Index
X axis = Weight



Weak positive correlation

Figure 2: While this is a positive correlation, data points do not land exactly on the line. This means that these variables are correlated, but there are other things that may contribute. For example, there may be a correlation between students how often students drink soda and eat pizza but other things may contribute as well – such as not being able to eat dairy, not liking tomato sauce or students preferring milk with their pizza.

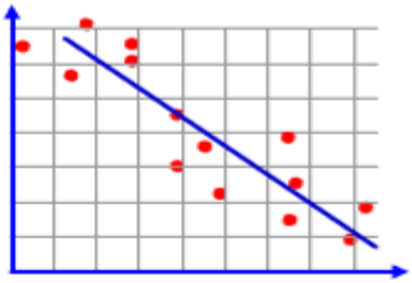
Y axis = Eating pizza
X axis = Drinking soda



Strong negative correlation

Figure 3: This is a strong negative correlation. That means when one variable goes up, the other goes down. A good example would be the longer a student stays awake, the less time that is available for a student to sleep since there are only 24 hours in a day. This is the kind of "tight" data scientist want.

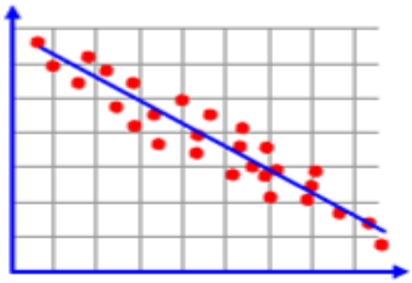
Y axis = Amount of sleep possible
X axis = Awake time



Weak negative correlation

Figure 4: As one variable goes up, the other tends to go down. There are a lot of other issues that may affect results, for example wearing a higher SPF of sunscreen reduces the severity of sunburn; however other factors influence results like skin type (some people burn more easily than others), amount of time a person spends in the sun or the climate where they live (Alaska vs. equator).

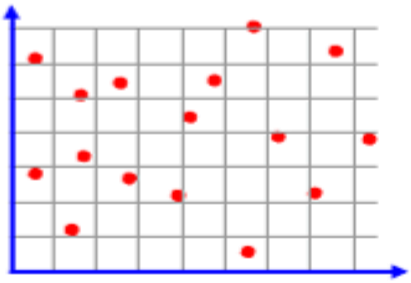
Y axis = Severity of sunburn
X axis = SPF of sunscreen



Moderate negative correlation

Figure 5: This is a moderate negative correlation meaning as one variable goes up; the other tends to go down. The data points typically follow the same pattern, but it is not a perfect line. A good example would be the longer a student is actively awake, the less sleep they will get. While this is mostly true, other factors might influence results. For instance, students could be in bed reading before they fall asleep so the number of hours slept may vary.

Y axis = number of hours sleeping
X axis = awake time (active)

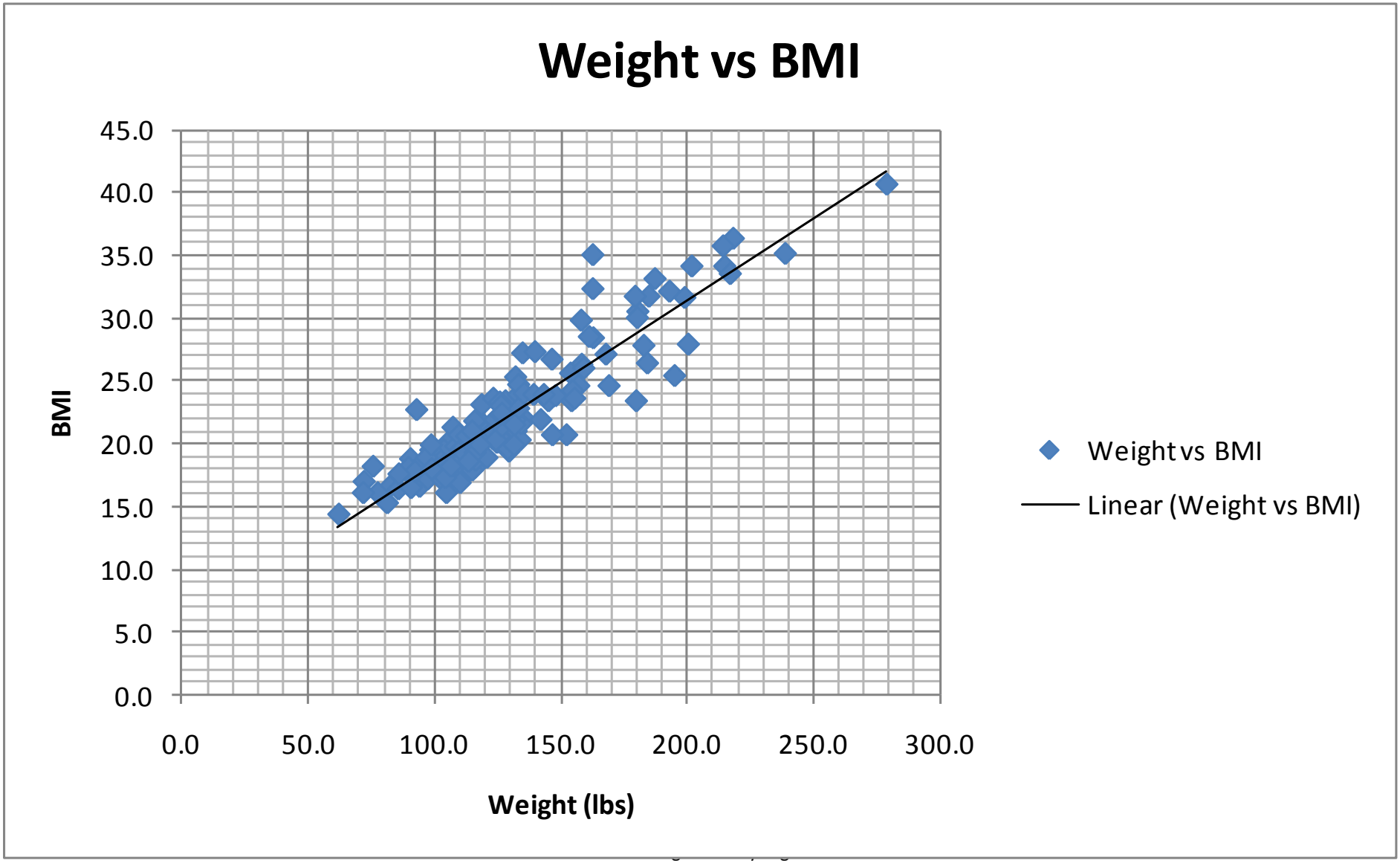


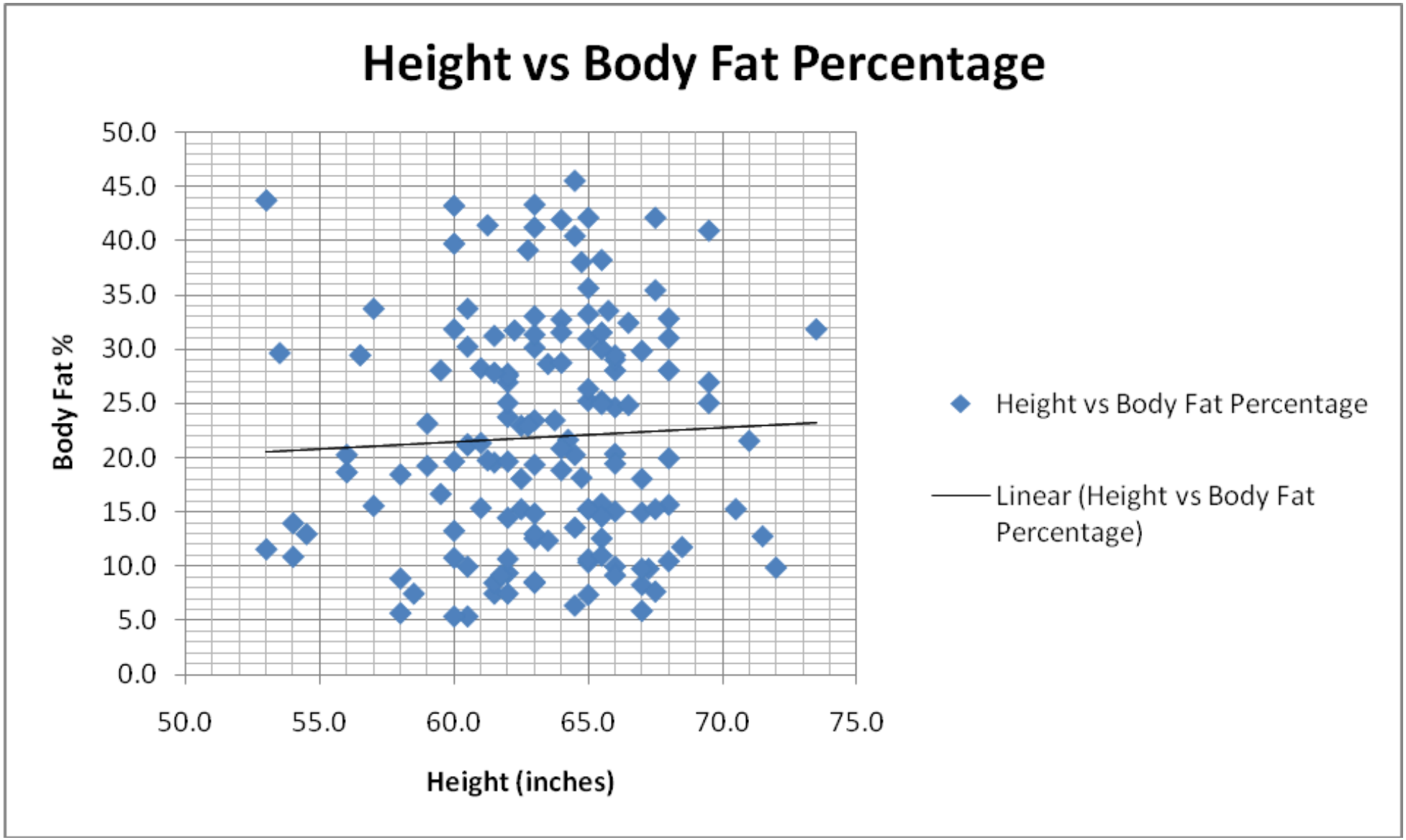
No correlation

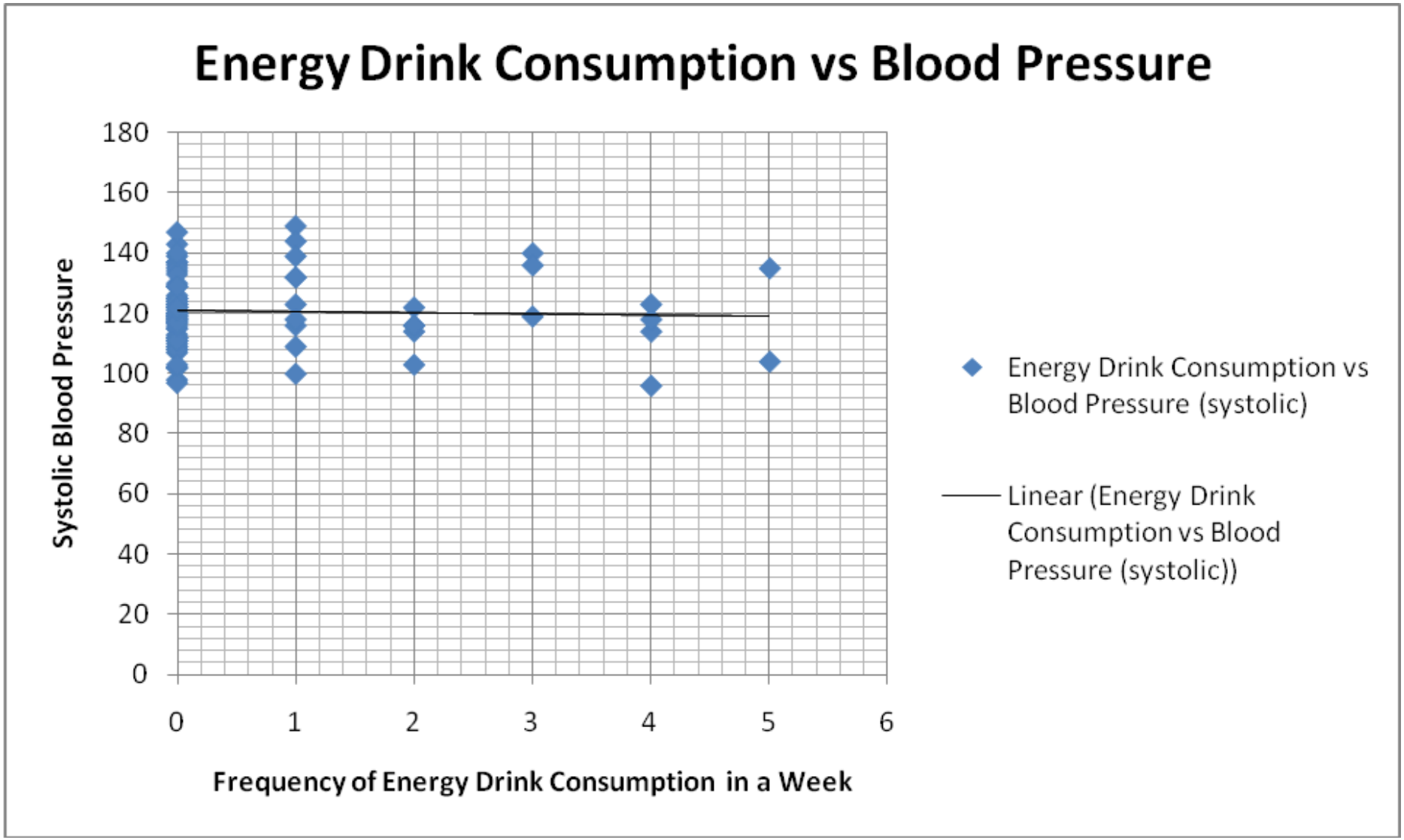
Figure 6: This scatter plot shows no correlation. Increasing one variable has no predictable effect on the other. An example would be wearing sunscreen more often does not predict a person's height. Since a short or tall person may or may not wear sunscreen.

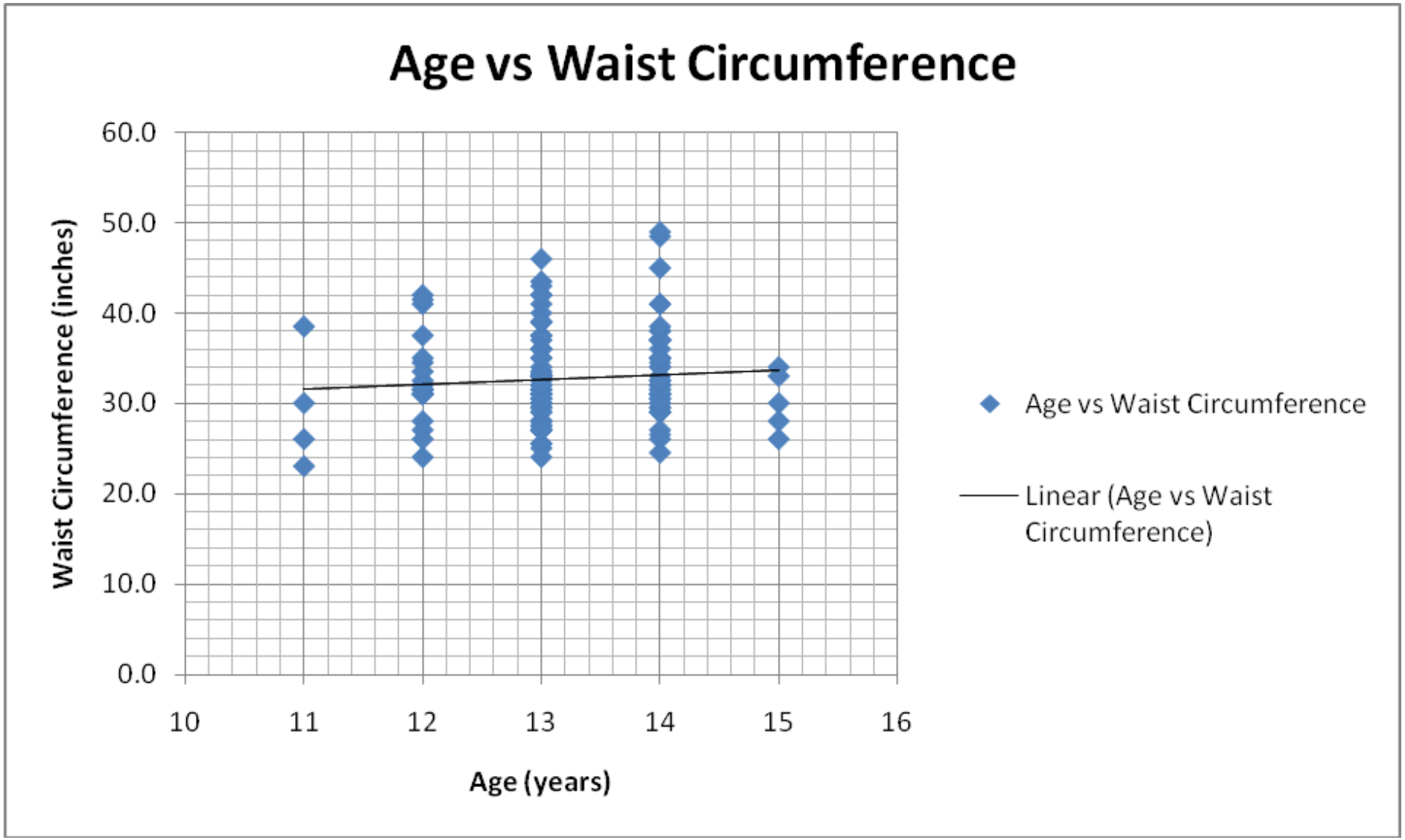
Y axis = height of a person
X axis = frequency of wearing sunscreen

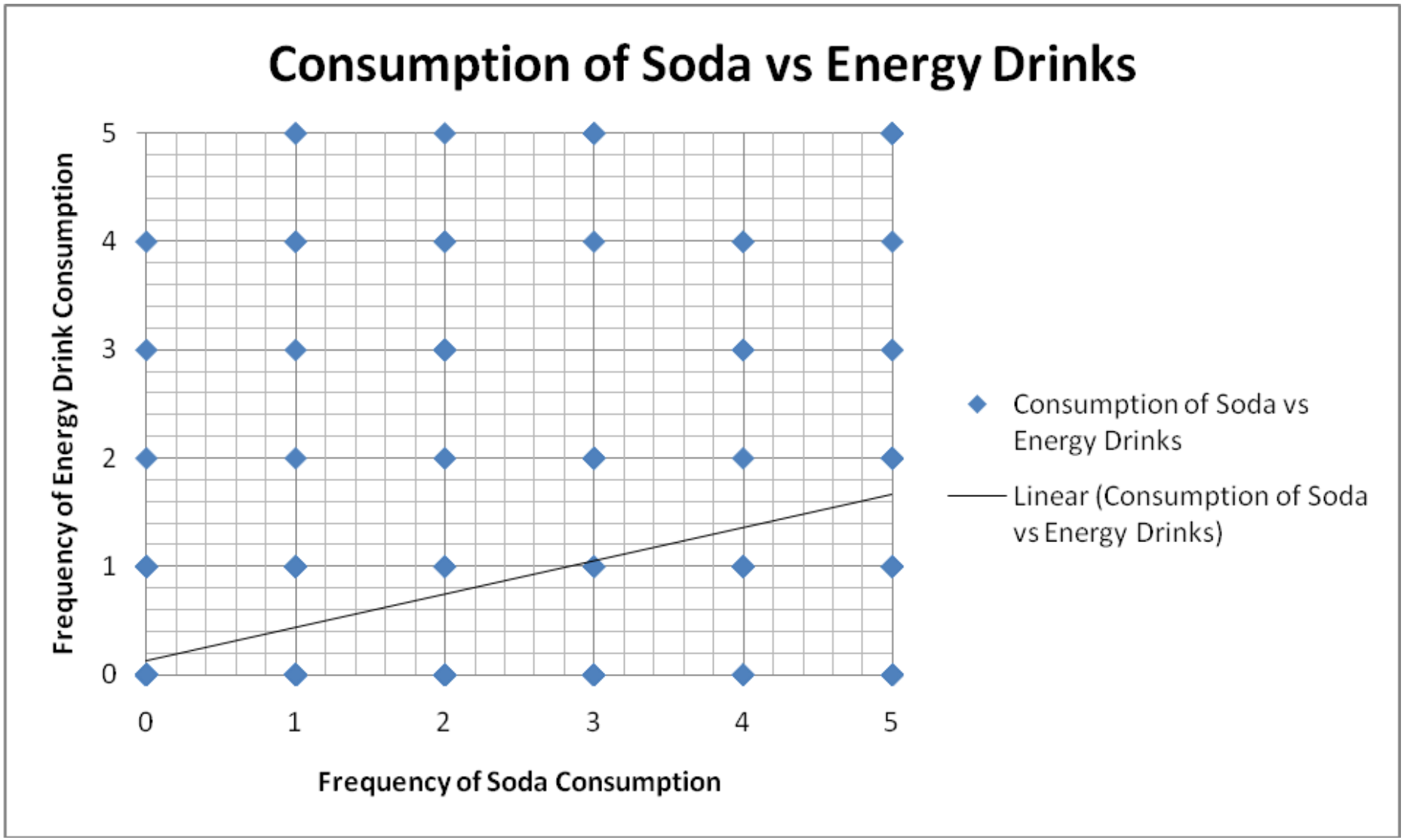
Handouts for Activity 2 - *Directions: Cut out graphs and distribute to individual small groups of 4-6 students*

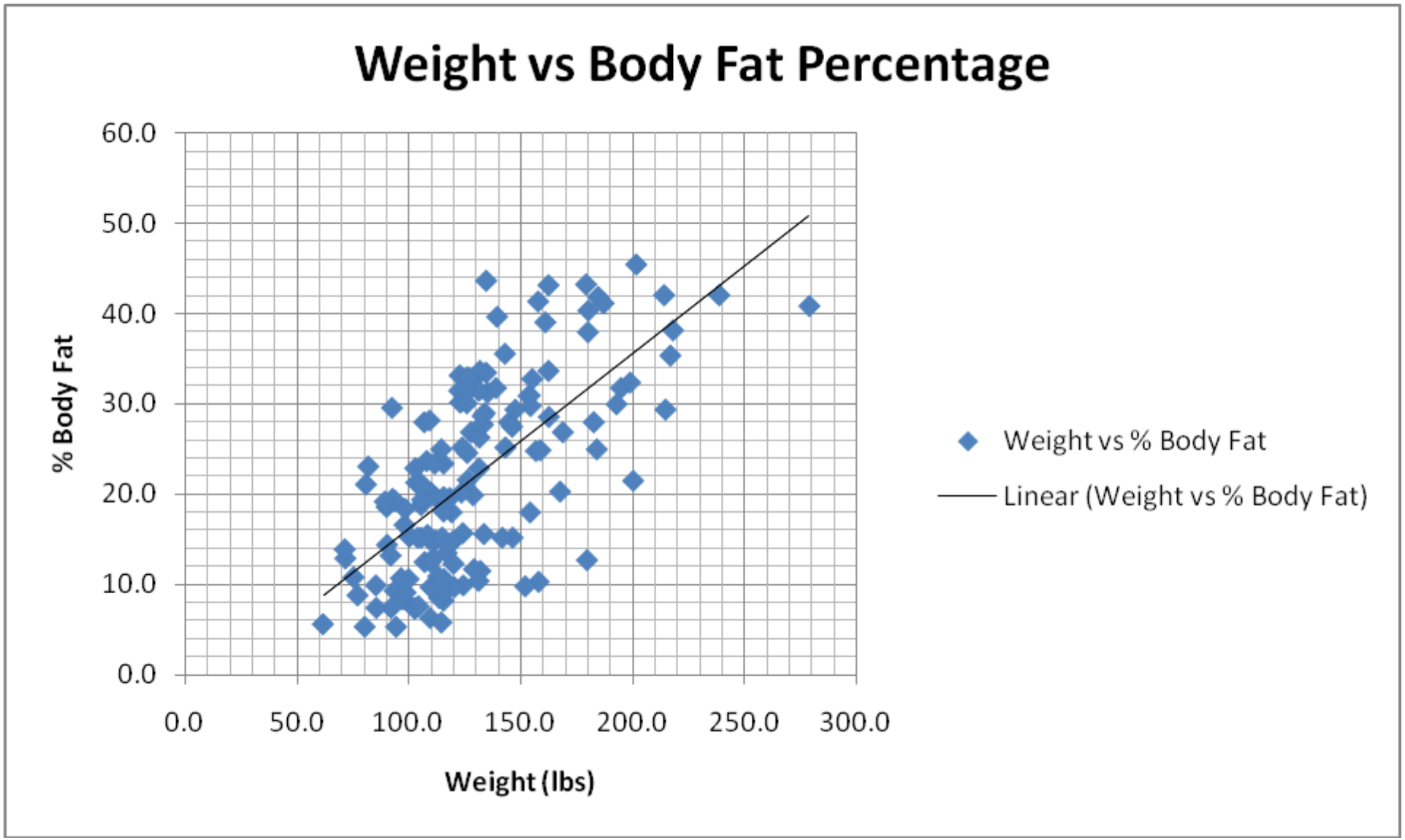



Height vs Body Fat Percentage
Body Fat %
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
50.0
55.0
60.0
65.0
70.0
75.0
Height (inches)
Height vs Body Fat Percentage
Linear (Height vs Body Fat Percentage)

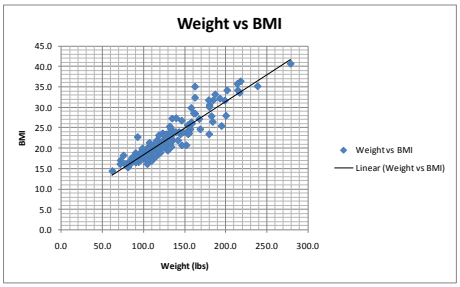

Energy Drink Consumption vs Blood Pressure
Systolic Blood Pressure
180
160
140
120
100
80
60
40
20
0
0
1
2
3
4
5
6
Frequency of Energy Drink Consumption in a Week
Energy Drink Consumption vs Blood Pressure (systolic)
Linear (Energy Drink Consumption vs Blood Pressure (systolic))


Age vs Waist Circumference
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Waist Circumference (inches)
10
11
12
13
14
15
16
Age (years)
Age vs Waist Circumference
Linear (Age vs Waist Circumference)

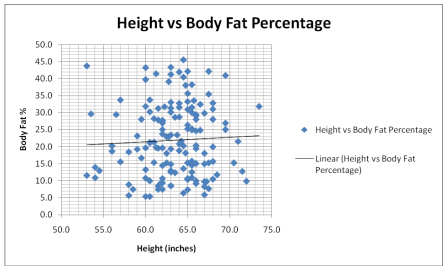

Consumption of Soda vs Energy Drinks
Frequency of Energy Drink Consumption
0
1
2
3
4
5
Frequency of Soda Consumption
Consumption of Soda vs Energy Drinks
Linear (Consumption of Soda vs Energy Drinks)


Weight vs Body Fat Percentage
60.0
50.0
40.0
30.0
20.0
10.0
0.0
% Body Fat
0.0
50.0
100.0
150.0
200.0
250.0
300.0
Weight (lbs)
Weight vs % Body Fat
Linear (Weight vs % Body Fat)

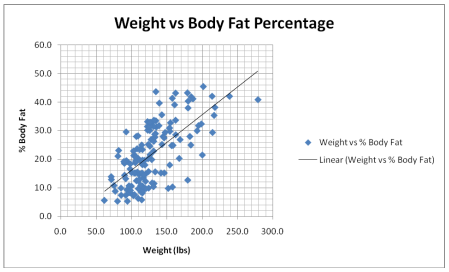
LET'S GET HEALTHY! GRAPHS KEY



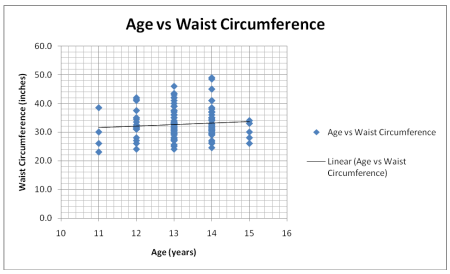
BMI is calculated based on weight so as weight goes up, so should BMI. This is a strong positive correlation between these two factors.



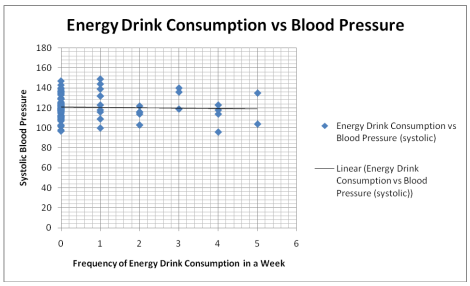
Being tall doesn't necessarily mean you'll have a high or low percentage of body fat. As a result, there is a weak positive correlation, at best, between these factors.



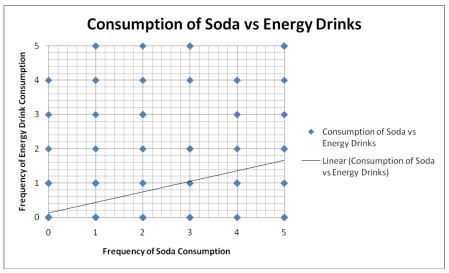
Increased weight doesn't necessarily mean you'll have a high body fat percentage because your height isn't taken into account. As a result, there is a moderate positive correlation between these factors.



There is a weak positive correlation between age and waist circumference. Just because a person is getting older and they are growing, it doesn't mean that their waist will increase dramatically.



There is a no correlation between frequency of energy consumption and blood pressure. A person could have drank 6 energy drinks last Monday, but had their blood pressure measured on Friday. Or a person could have drank one energy drink just before being measured. Or no energy drinks but just came in from recess! A lot of variables could explain these results.



This graph is unclear because each dot could potentially represent one person, six people or 60 people. There are also no units on the axes because the variables indicate categories. A better way to illustrate this type of data would be to use different bar graphs or pie charts to compare soda and energy drink consumption.

Scenario 1: As ice cream sales increase, the rate shark attacks increases sharply. Therefore, increase in ice cream sales causes shark attacks since consuming ice cream makes people tastier to sharks.

Scenario 2: There is strong correlation between the number of cavities in elementary school children and their vocabulary size. Therefore, the more words children know, the higher number of cavities they will have. Or if students consume more sugar, they will gain a larger vocabulary.

Scenario 3: A 2003 study identified Tucson and Phoenix as the first and third cities with the highest occurrence of asthma sufferers. Therefore, living in Tucson and Phoenix will cause residence to suffer from asthma.

Scenario 4: Over the course of several weeks the needles from the pine trees along the Wombat River fell into the water. Shortly thereafter, many dead fish washed up on the riverbanks. When the EPA investigated, the owners of the Wombat River Chemical Company claimed that is it was obvious that the pine needles had killed the fish.

Scenario 5: Nations that add fluoride to their water have a higher cancer rate than those that don't.

SCENARIO KEY

Scenario 1: As ice cream sales increase, the rate shark attacks increases sharply. Therefore, increase in ice cream sales causes shark attacks since consuming ice cream makes people tastier to sharks.

Explanation: The aforementioned example fails to recognize the importance of time in relationship to ice cream sales. Ice cream is sold during the summer months at a much greater rate, and it is during the summer months that people are more likely to engage in activities involving water, such as swimming. The increased shark attacks are simply caused by more exposure to water based activities, not ice cream.

Scenario 2: There is strong correlation exists between the number of cavities in elementary school children and their vocabulary size. Therefore, the more words children know, the higher number of cavities they will have. Or if students consume more sugar, they will gain a larger vocabulary.

Explanation: No one advocates eating more candy to increase knowledge. These variables are both tied with age. As students get older their vocabulary increases, and as they get older the number of cavities they have will increase.

Scenario 3: A 2003 study identified Tucson and Phoenix as the first and third cities with the highest occurrence of asthma sufferers. Therefore, living in Tucson and Phoenix will cause residence to suffer from asthma.

Explanation: Why do the two major cities in Arizona seem to be the worst places for asthmatics? The answer is, they aren't. This is a relationship of cause vs. result. People who have asthma are more likely to move to Arizona due to its warm, dry desert air in an effort to lessen their asthmatic symptoms. Living in Tucson and Phoenix doesn't cause asthma. Instead having asthma causes people to want to live in Tucson and Phoenix.

Scenario 4: Over the course of several weeks the needles from the pine trees along the Wombat River fell into the water. Shortly thereafter, many dead fish washed up on the riverbanks. When the EPA investigated, the owners of the Wombat River Chemical Company claimed that is it was obvious that the pine needles had killed the fish.

Explanation: This false conclusion is made when a common cause is ignored. A conclusion is made without considering the possibility that a third factor is causing both the pine trees to lose their needles and the fish to die. Many local environmentalists later claimed that it was chemical plant's toxic wastes being released into the river that caused both the trees and the fish to die. Thus the pine needles had no real effect on the fish.

Scenario 5: Nations that add fluoride to their water have a higher cancer rate than those that don't.

Explanation: This is an example of two correlations being somewhat related in some way but that one correlation does not necessarily cause the other. In this scenario nations that add fluoride to their water are generally wealthier and more health-conscious than nations who don't add fluoride. Since these nations are wealthier and more health conscious, they have populations that live longer. Thus some of these citizens live long enough to develop cancer, which is, to a large extent, a disease of old age.

Online Resources

http://www.mathematics.com.au/images/600x480/4329.gif

http://www.mathhelpforum.com/math-help/basic-statistics-probability/57041-correlation-r-horizontal-points.html