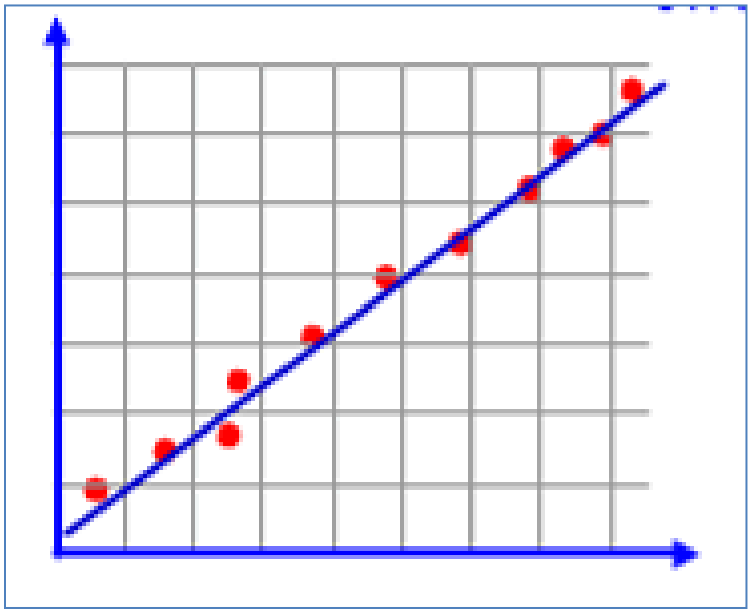
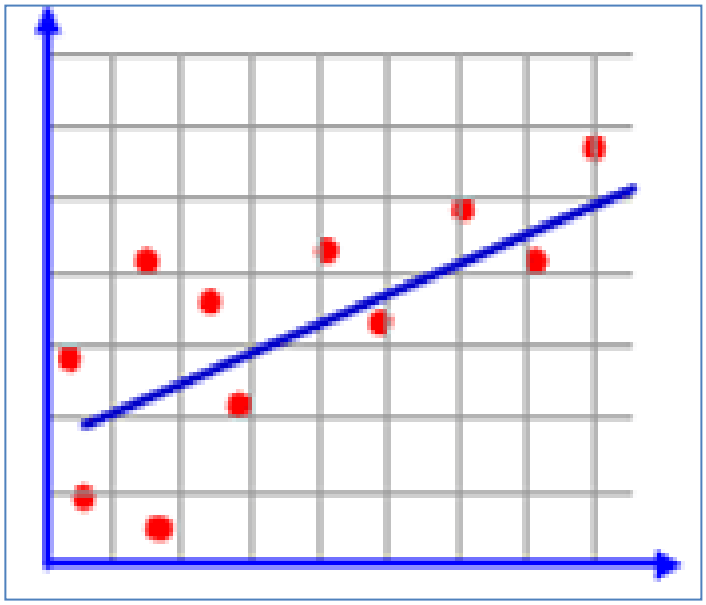
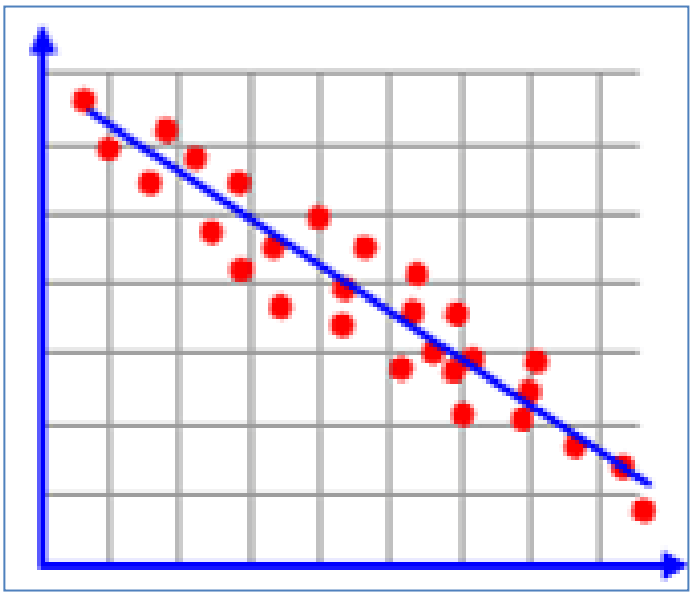
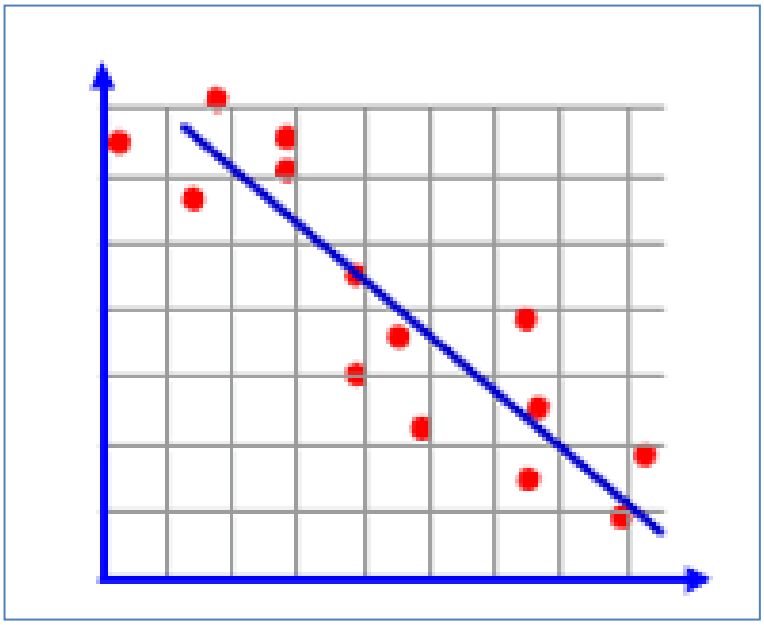
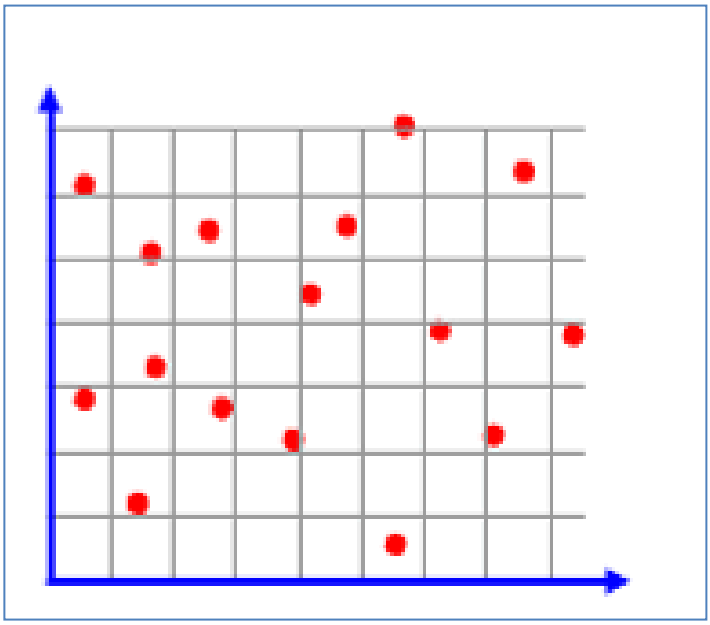
Handout for Activity 1 - *Directions: Cut out graphs and distribute to individual small groups of 4-6 students*







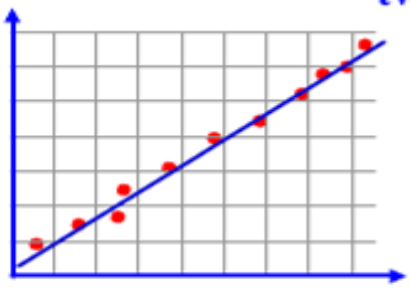




SCATTER PLOT GRAPHS AND KEY FOR TEACHER

SCATTERPLOTS & CORRELATION

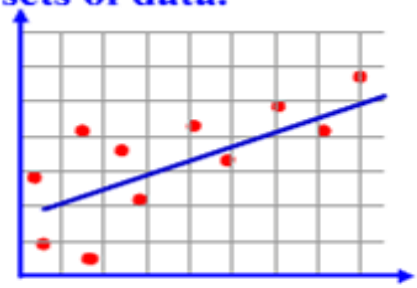
Correlation - indicates a relationship (connection) between two sets of data.



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.

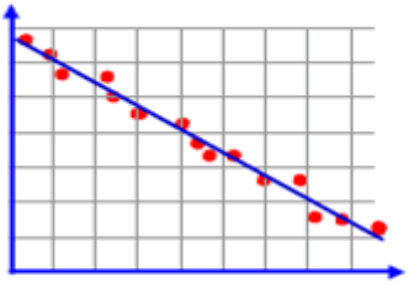
Y axis = Body Mass Index
X axis = Weight



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.

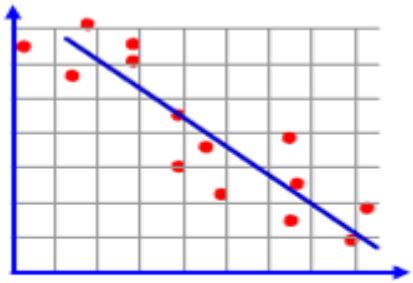
Y axis = Eating pizza
X axis = Drinking soda



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.

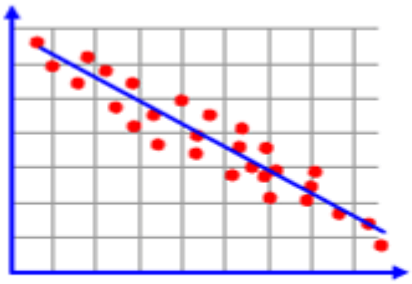
Y axis = Amount of sleep possible
X axis = Awake time



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).

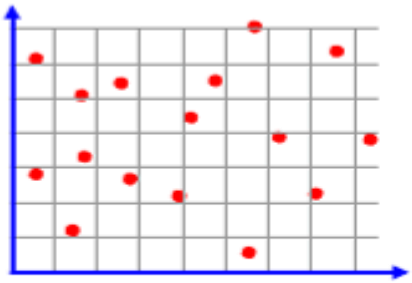
Y axis = Severity of sunburn
X axis = SPF of sunscreen



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