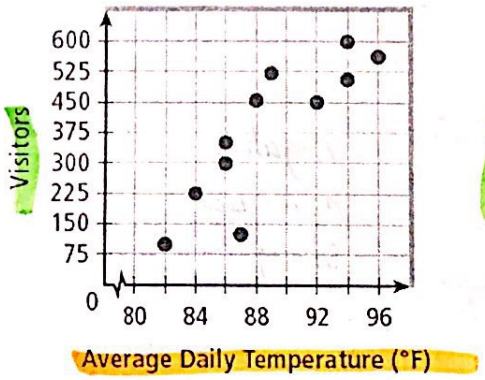
Name: 6th period Date: _____ Hour: _____

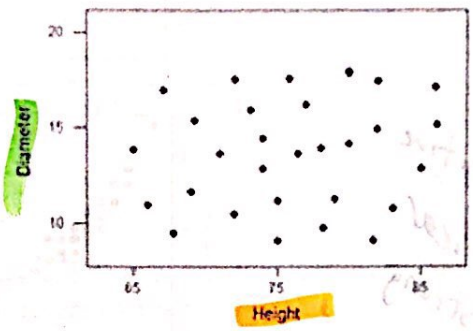
**Unit 2 Day 6 and 7: Scatterplots Vocabulary**

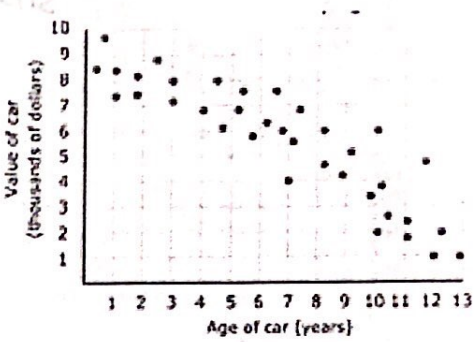
*Focus Question: How do I explain the relationships shown in a scatterplot like a statistician?*

A. Correlation

1. The first thing statisticians look for when examining a scatterplot is a **positive or negative correlation** (relationship). The correlation is the "answer" to the title of the graph. Some graphs have no correlation. We always read a graph from left to right !!!!!



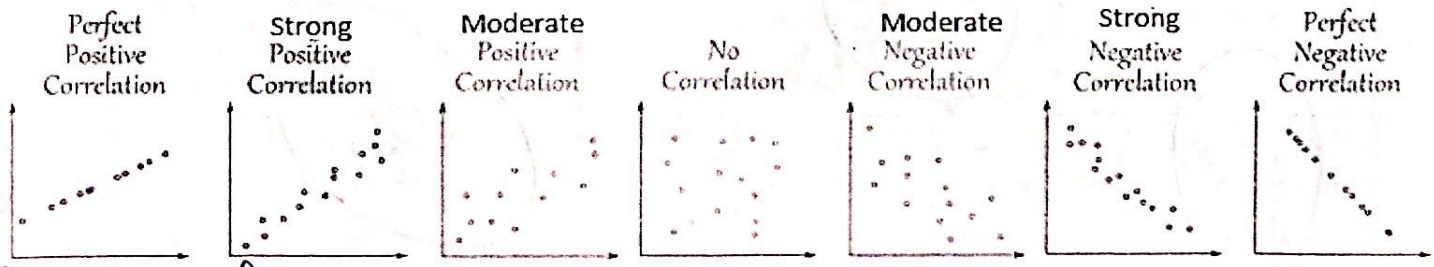

For each graph above give a good title and then describe its correlation by filling in the blanks.

Graph A: As the Ave. Temp. (IV) increases, the Visitors (DV) goes up. Therefore this graph has positive correlation.

Graph B: As the Height increases, the Diameter goes everywhere. Therefore this graph has no correlation.

Graph C: As the Age of Car increases, the Value goes down. Therefore this graph has negative correlation.

2. They also give the strength of the relationship



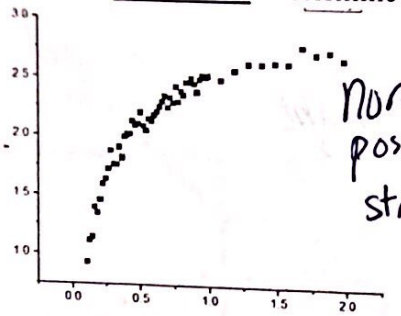

Perfect : form a straight line

Strong : dots are tight, but not a straight line

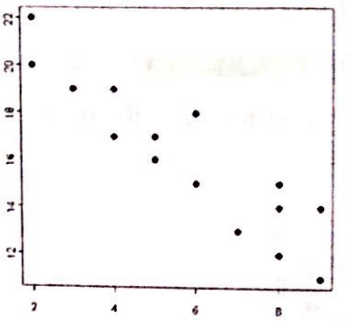
Moderate : dots are spaced, but can still see pattern

Weak to No Correlation: they're scattered, no pattern

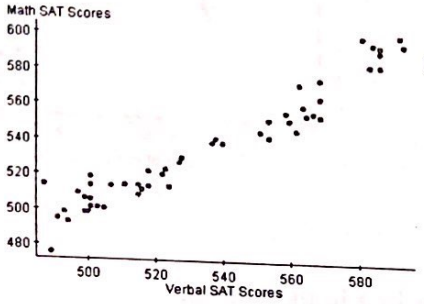
3. Once they are sure that a correlation exists, statisticians determine if the correlation is **linear or non-linear**. Determine if each graph is linear or non-linear. Explain your choice.



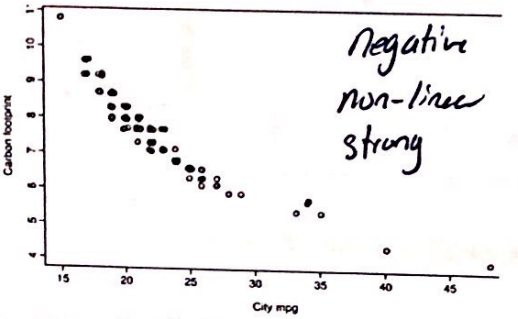
non-linear
positive
strong



linear
negative
moderate




positive
linear
strong



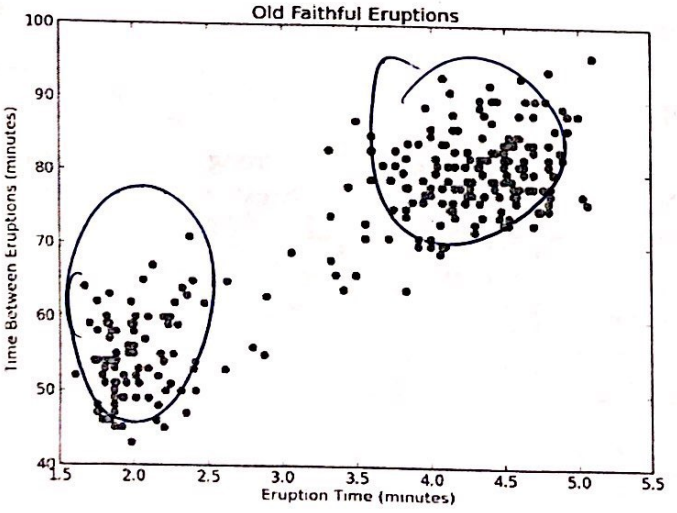

negative
non-linear
strong

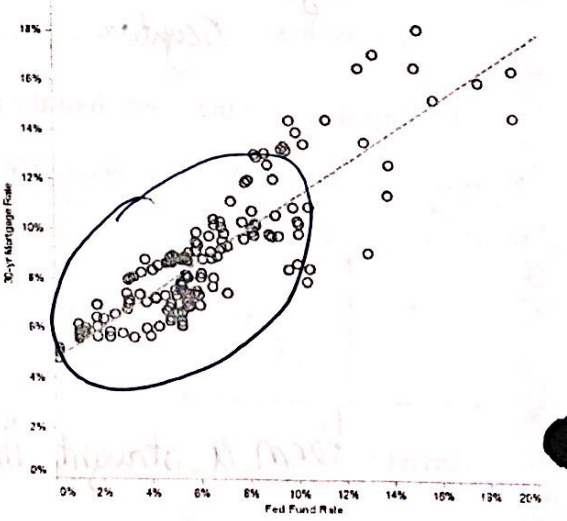
**This one can be VERY tricky: non-linear means** it curves.

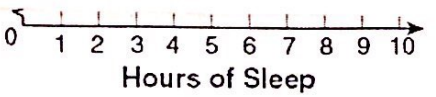
**EVERYTHING else that has a correlation is** linear.

B. Other features of a scatter plot

1. **Clusters** (large groups **in relation to all the points**): Circle the clusters on each scatterplot.

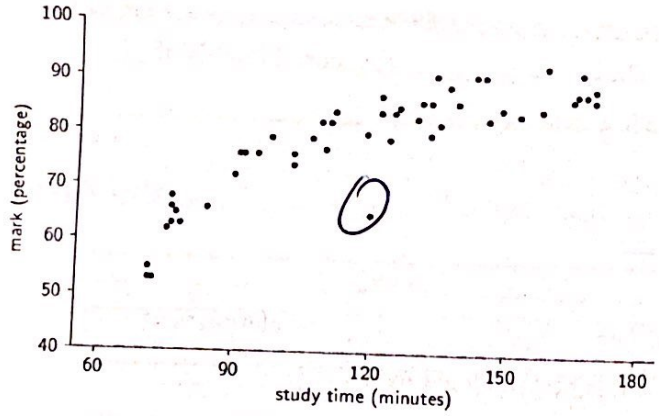


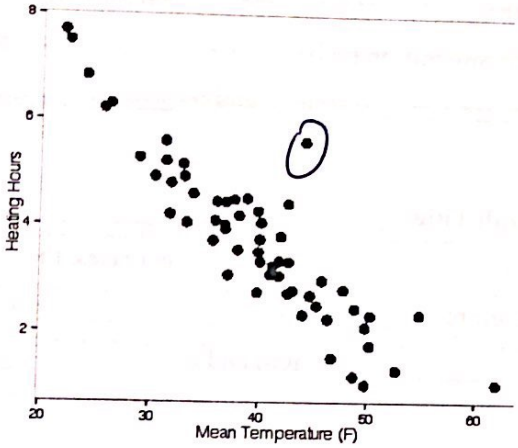



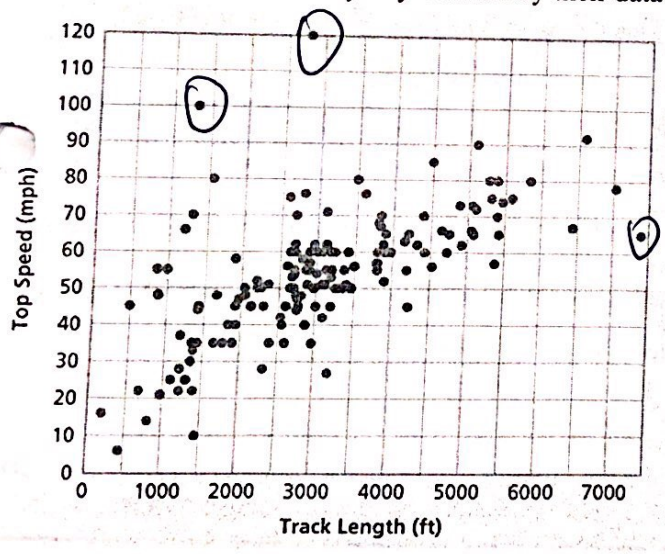
Hour: _____

2. **Outliers** (points that **don't fit the trend** of the rest of the data) Identify any outliers by their data point.



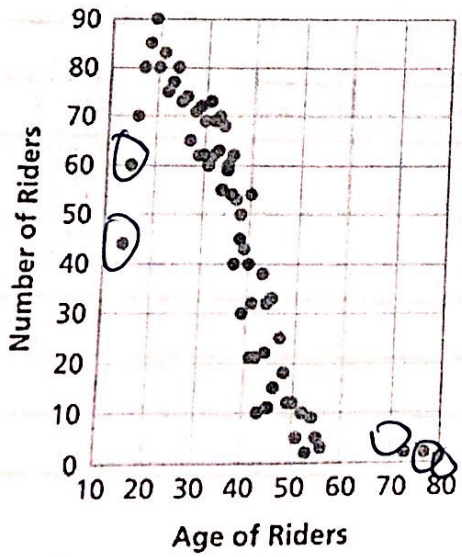




**C.** Putting it all together practice. For each scatter plot below, describe the correlation. Then circle any clusters and identify any outliers by their data point.




As the Track Length increases, the Top Speed goes up. Therefore this graph has positive correlation. The positive correlation is of moderate strength and is linear because it is best modeled by a line.

Outlier(s): (2750, 120) 3




As the Age of Riders increases, the # of Riders goes down. Therefore this graph has negative correlation. The negative correlation is of strong strength and is non-linear because it is best modeled by a curve.

Outlier(s): 5

_____ Hour: _____

**Graph Title:** _____

**Graph Title:** _____

**D.** Continuous Vs. Discrete Data

Sometimes the data you are studying is discrete (countable values only). Two example of discrete data are "number of people" or "age in months."

1. Give two other examples of a discrete variable.

Boxes sold

Books

Sometimes the data is continuous. Continuous means that ALL values of the variable are possible including fractions, decimals, pi, etc. For instance, if time was only measured in seconds, all 5 swimmers at the right would tie!

In this case we show a smooth, connected line or curve.

Besides "time," two other examples of continuous data are "height" and "weight."

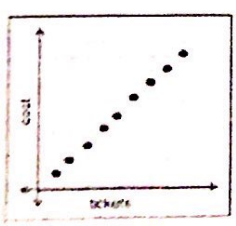
2. Give two other examples of a continuous variable.

Distance speed

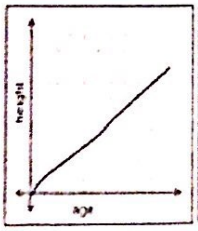
2020 US Olympic Swimming Trials Qualifiers : Al

| Time | Name | LSC | Event |
|---|---|---|---|
| 21.04 | Dressel, Caeleb | FL | 50 FR |
| 21.62 | Andrew, Michael | SI | 50 FR |
| 21.87 | Adrian, Nathan | PC | 50 FR |
| 21.87 | Held, Ryan | MR | 50 FR |
| 21.95 | Chadwick, Michael | NC | 50 FR |

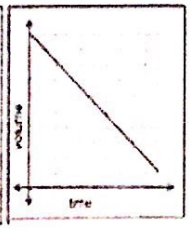
3. Tell whether each graph below is continuous or discrete.



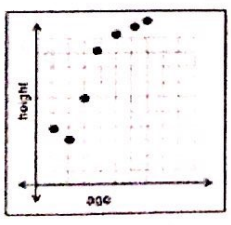

Discrete



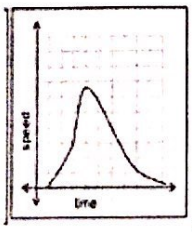

continuous



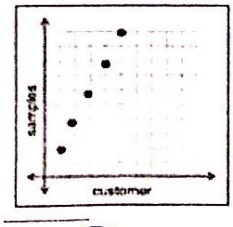

C




D




C




D