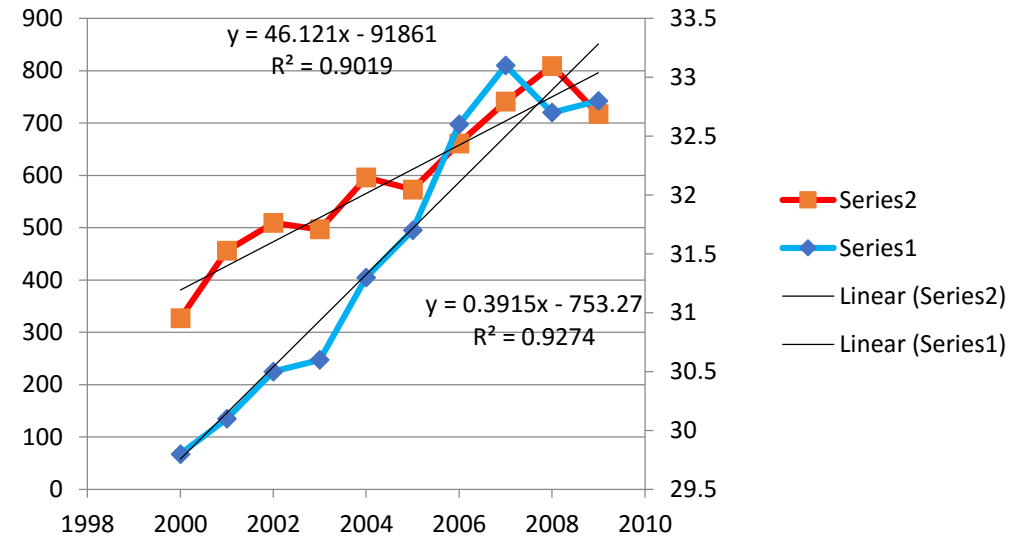
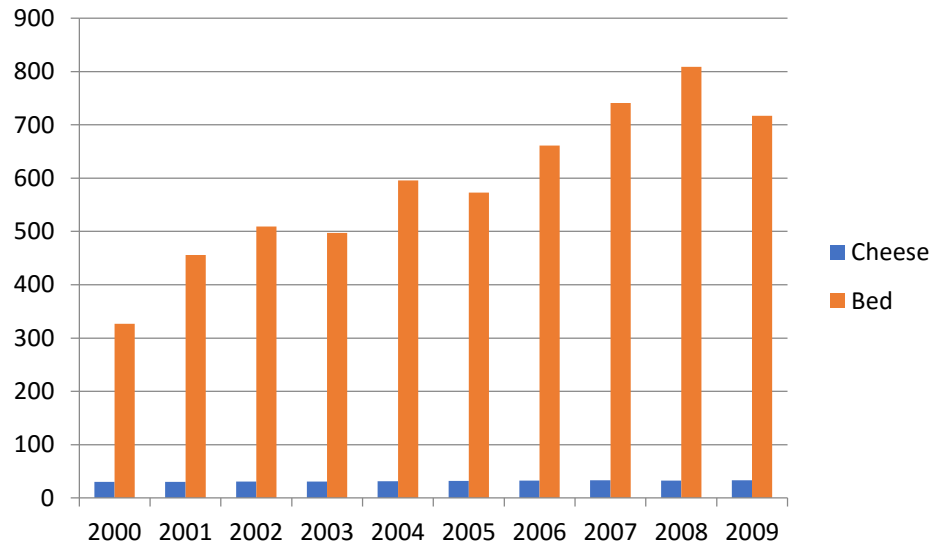
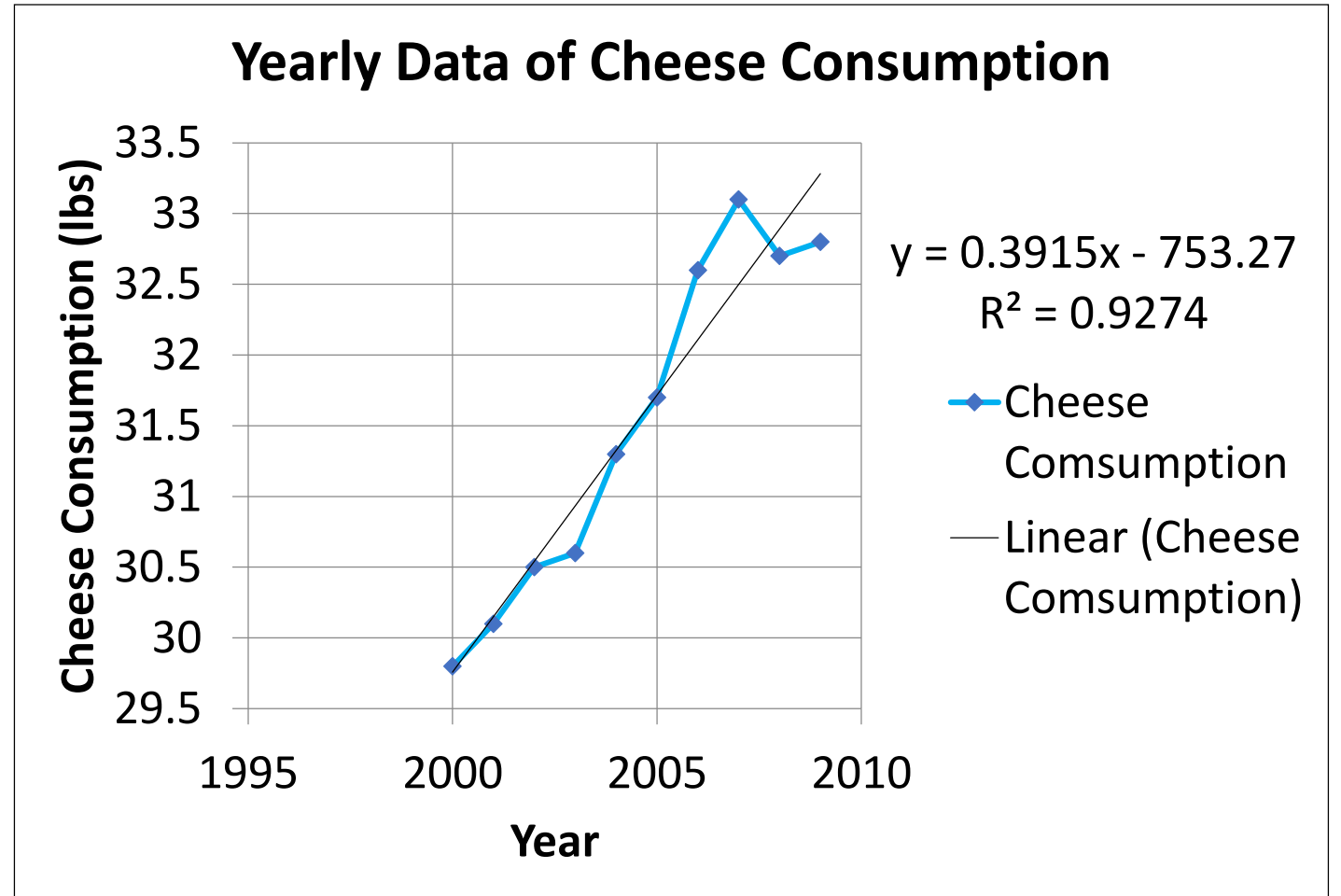
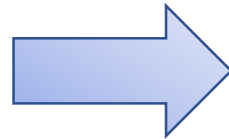


# Graphical Relationships Handout



# Analyzing Data Visually – Recognize Patterns

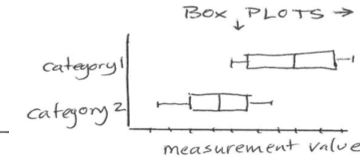
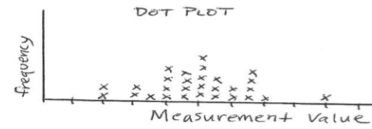
Year	Consumption of Cheese (lbs)
2000	29.8
2001	30.1
2002	30.5
2003	30.6
2004	31.3
2005	31.7
2006	32.6
2007	33.1
2008	32.7
2009	32.8



# Graph Types – Depends on Question(s)

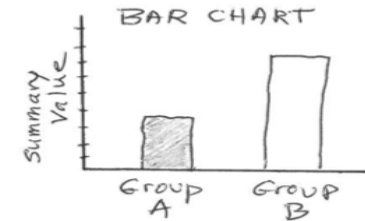
1. **Variability** within a group?

→ frequency plot, PBC

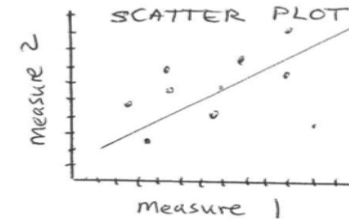


2. **Compare** groups?

→ 2 frequency plots **or** 1 bar graph

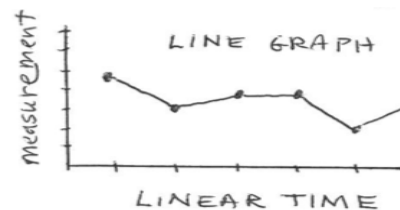


3. **Correlation** between two variables? →  
scatter plot



4. **Change** through **time**?

→ line graph

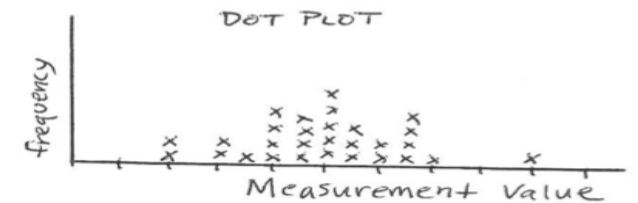
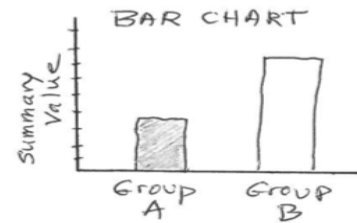


# Question Development: Variability

What is the measure of variability within a group?

*Sample Questions*

1. How deep are the lakes in Connecticut?
2. How many red hawks are spotted at Sleeping Giant every month?
3. What are teachers paid in CT?
4. How old are deer in CT when hunted?



## Graph Types:

Frequency Plots = Dot, Histogram, Box, Bar Plots

Process Behavior Chart - PBC

# Question Development: Comparison

How do two/more groups compare using a single variable or measure?

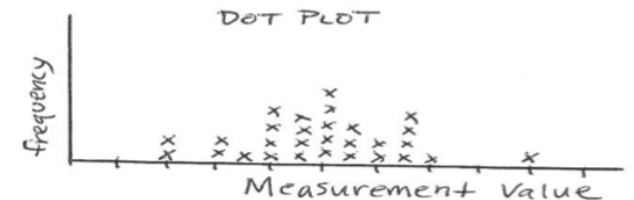
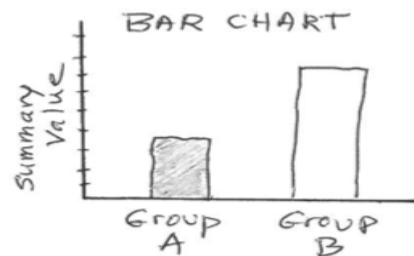
## *Sample Questions*

1. Do Cats and Dogs have the same average body temperature?
2. Which of the two car designs are consistently the fastest?

## **Graph Types:**

One Variable – Bar Graph

Multiple Variables – Frequency Plot



# Question Development: Numerical Correlations

What is the mathematical model or relationship between variables?

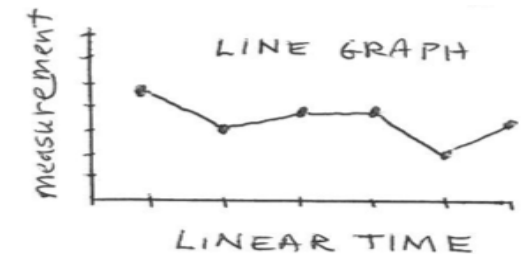
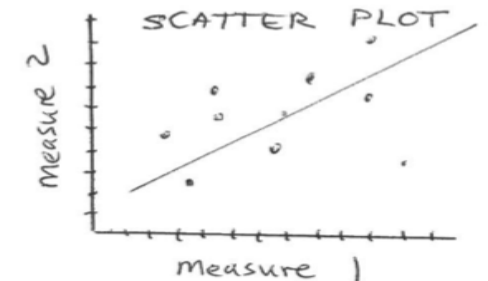
*Sample Questions*

1. How does force vary with mass?
2. What is the variance of rainfall in Connecticut in 2010?

## **Graph Types:**

Correlation between variables – Scatter Plot

Relationship between time and variable – Line Graph



# Question Development: Total/Proportionality

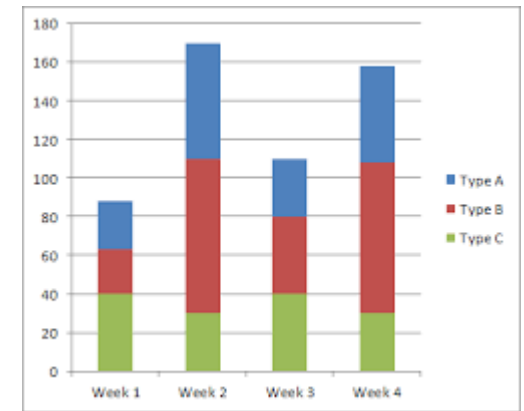
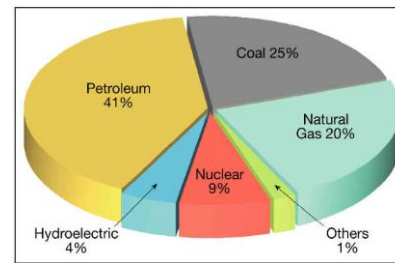
What is the percentage/amount/total of variable?

## *Sample Questions*

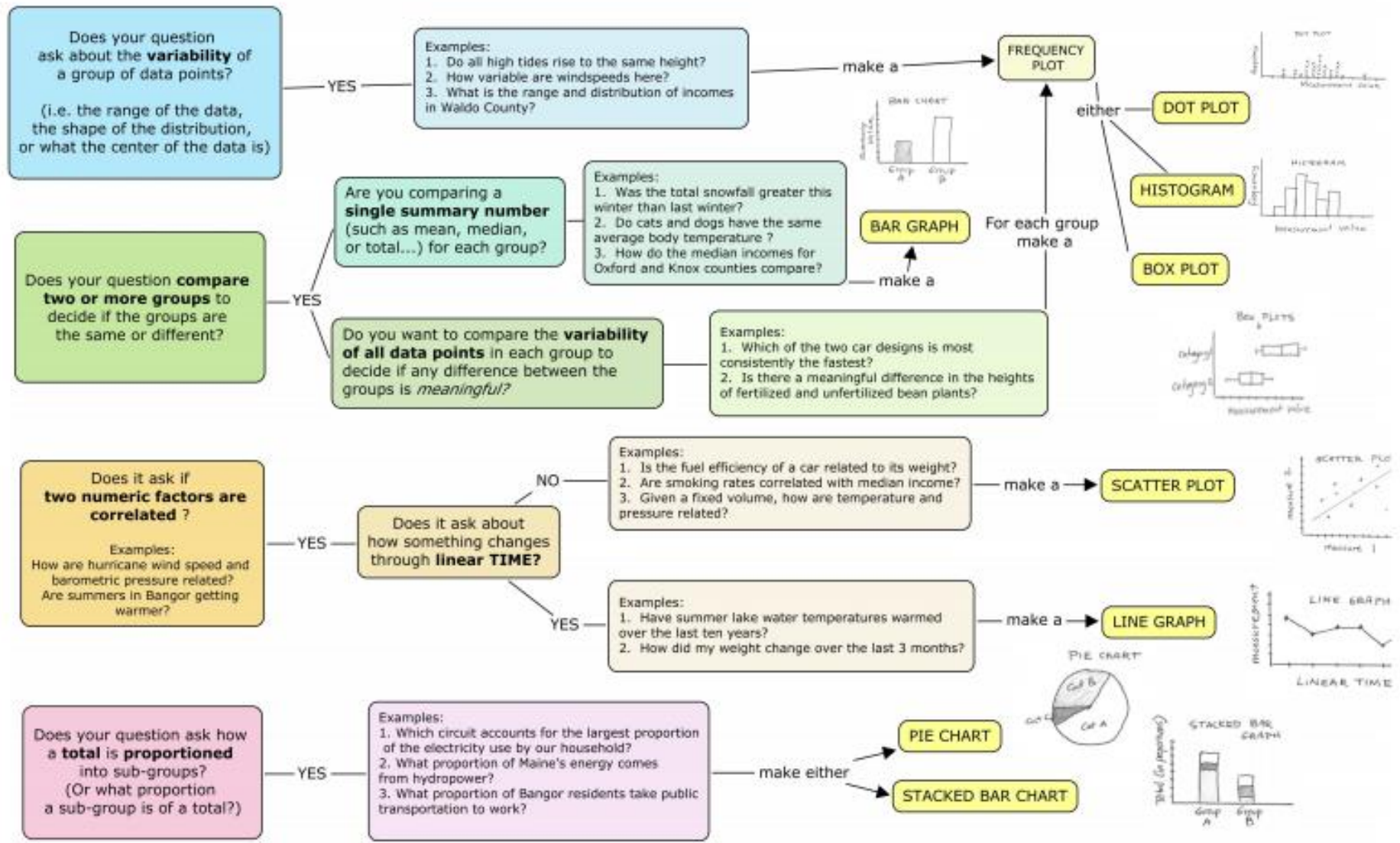
1. What consumes the most amount of energy in your household?
2. What portions of corn production is used for animal consumption?

## **Graph Types:**

Pie Chart, Stacked Bar Graph



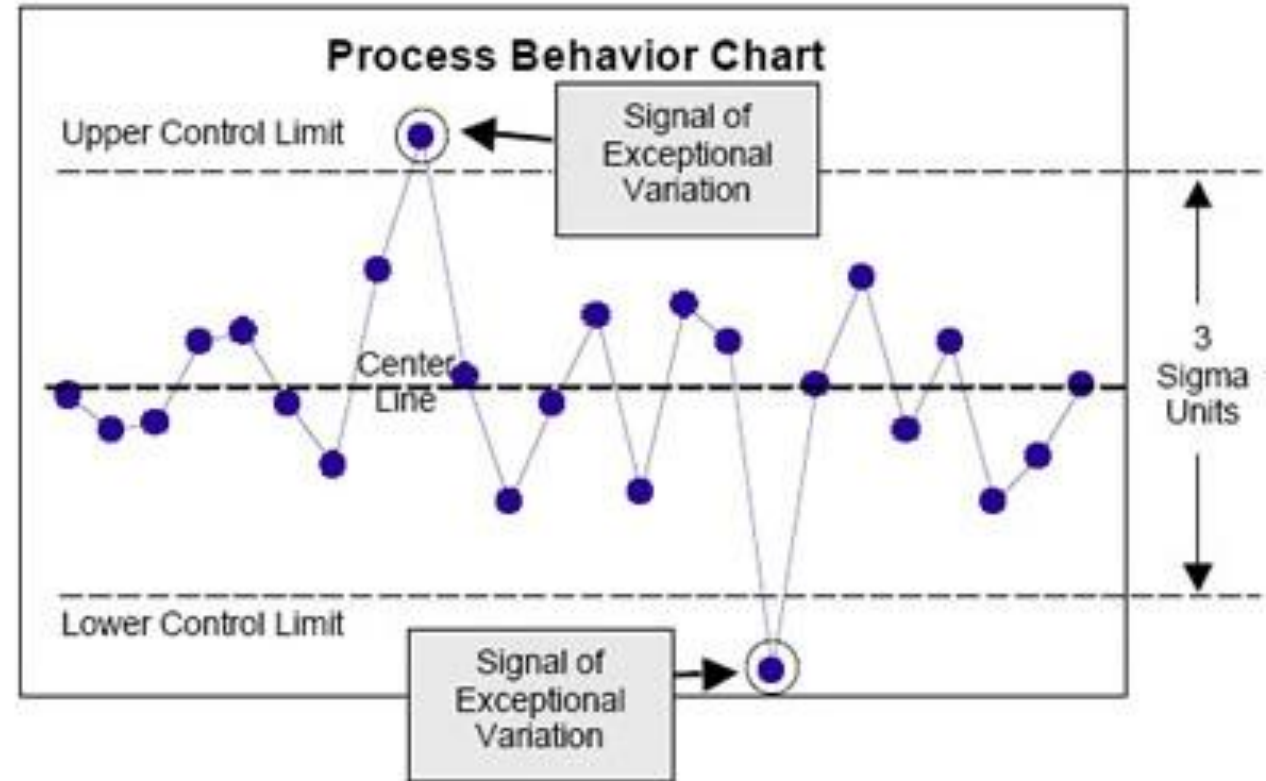
What question would you like to explore? Write it at the top of the page.





# Variation - Process Behavior Chart (PBC)

- Dr. Donald Wheeler
- Accounts for natural variable within systems – distinguishes from signals and noise
  - Noise = natural variable in systems
  - Signal = data outside the upper/lower limits of noise
- Analyzes variances in systems for
  - Trends
  - Signals
  - Extremes

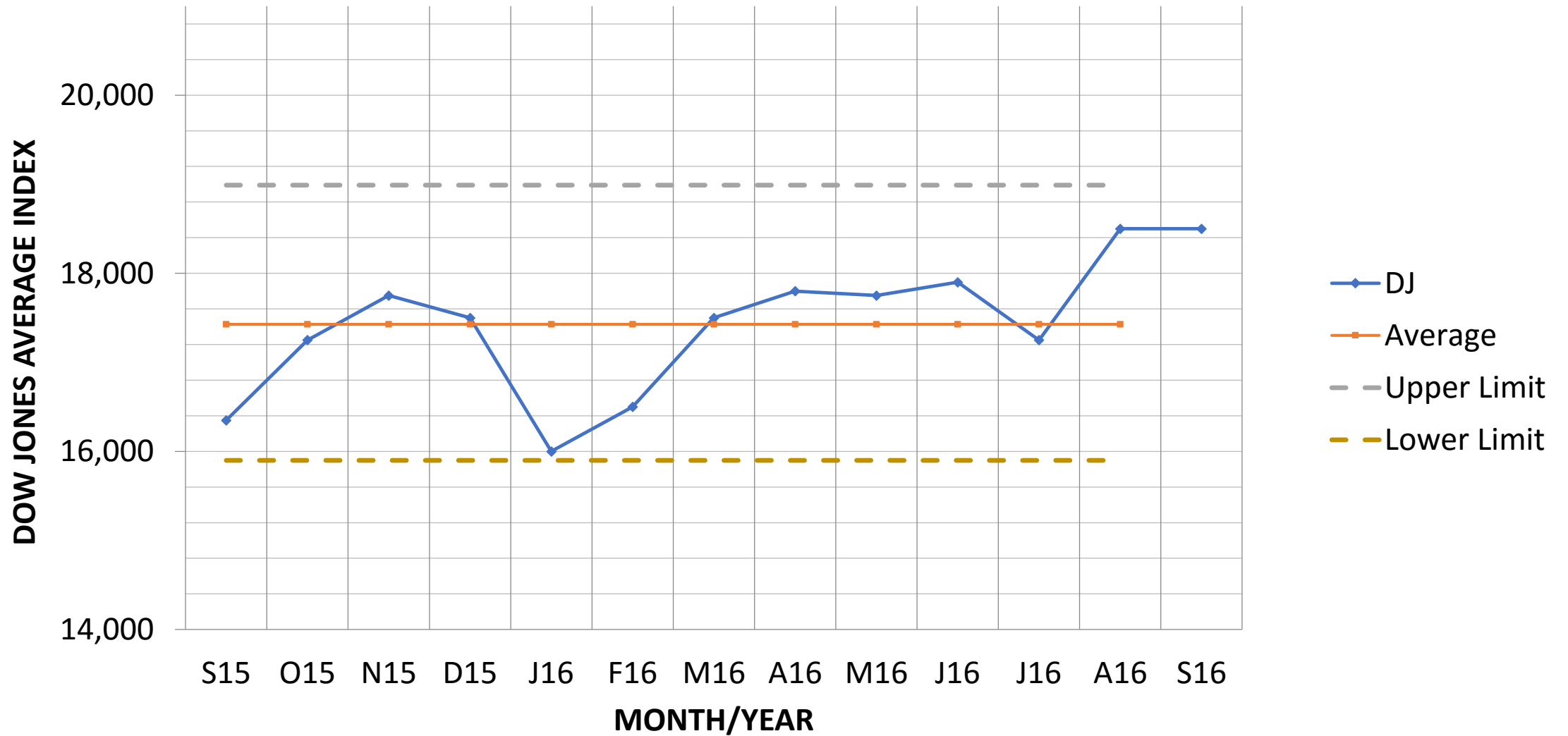


# DOW JONES INDU AVERAGE INDEX



**WOW! LOOK AT JAN 2016, THERE WAS A BIG DIP IN DATA!  
Is the variance within the natural limits of the system?**

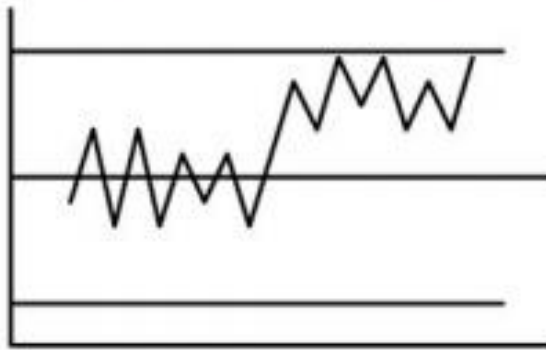
# PBC -Individual Values: Index Dow Jones



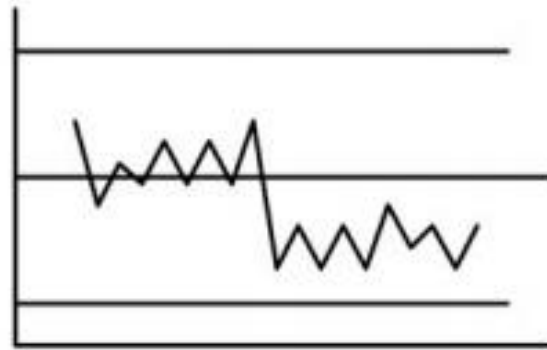
**There are no signals. The noise is within the natural variance of the system.**

## Process Behavior Chart Patterns

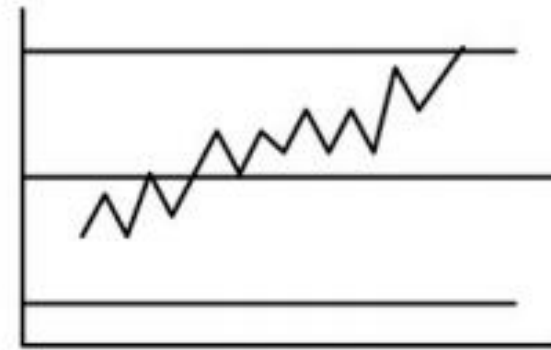
When you see certain patterns you should look for some systematic explanation



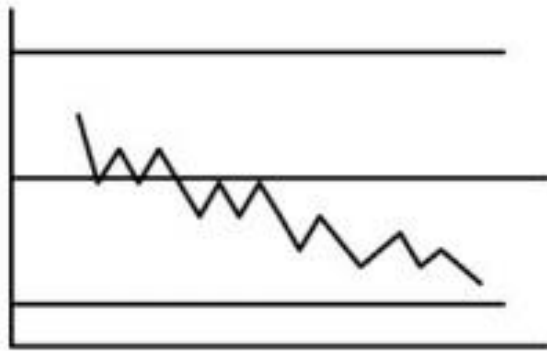
Upward Shift Pattern



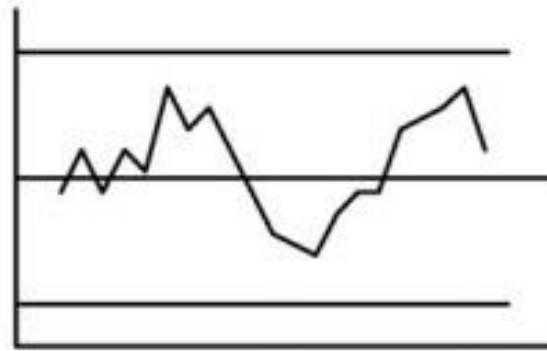
Downward Shift Pattern



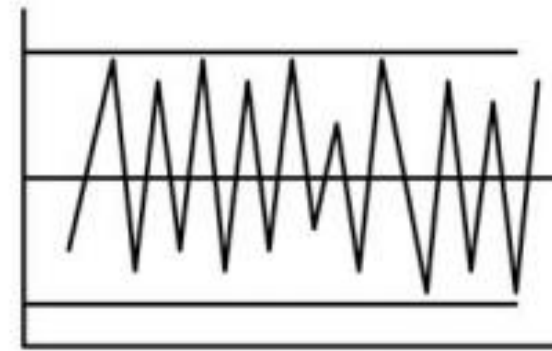
Upward Trend Pattern



Downward Trend Pattern

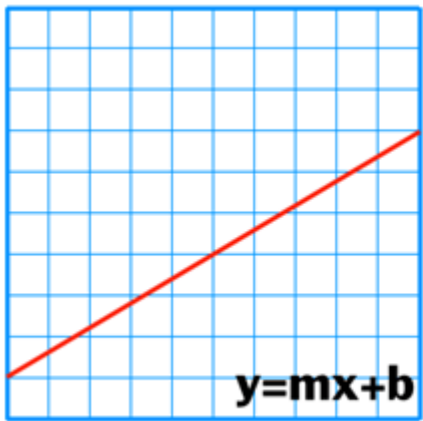


Cycle Pattern

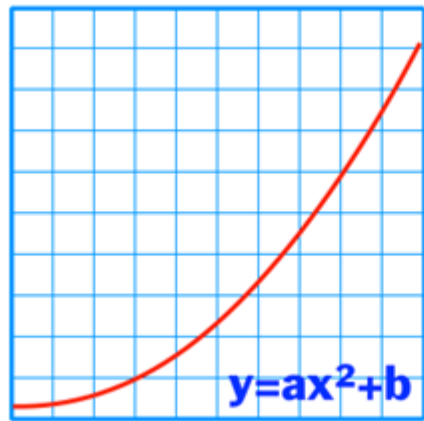


Systematic Pattern

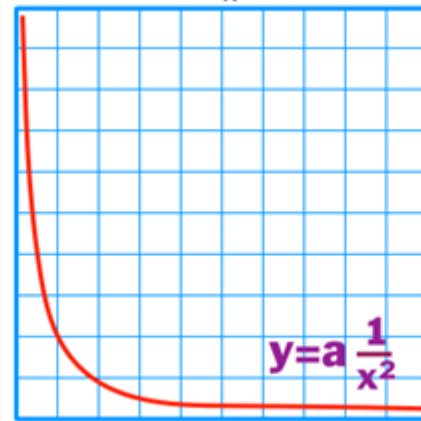
# Which mathematical model does your data fit?



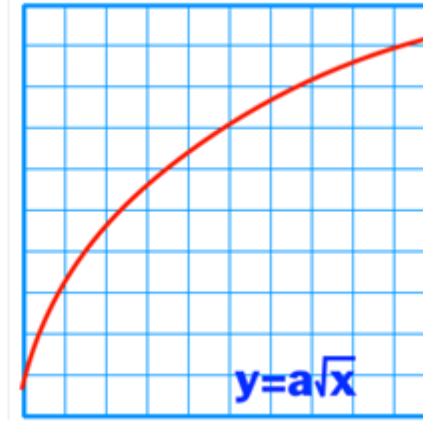
Linear



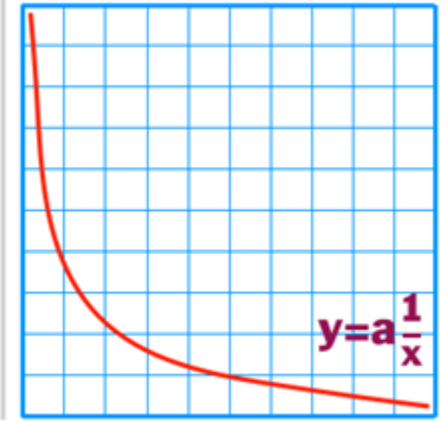
Power



Inverse  
Square



Square  
Root



Inverse

Shape of the plot gives information about the mathematical relationship between the variables - predictability

# Linear Model

## Equation

$$y = mx + b$$

y = dependent variable  
x = independent variable  
m = slope – rate of change  
b = y intercept – initial conditions

## Shape

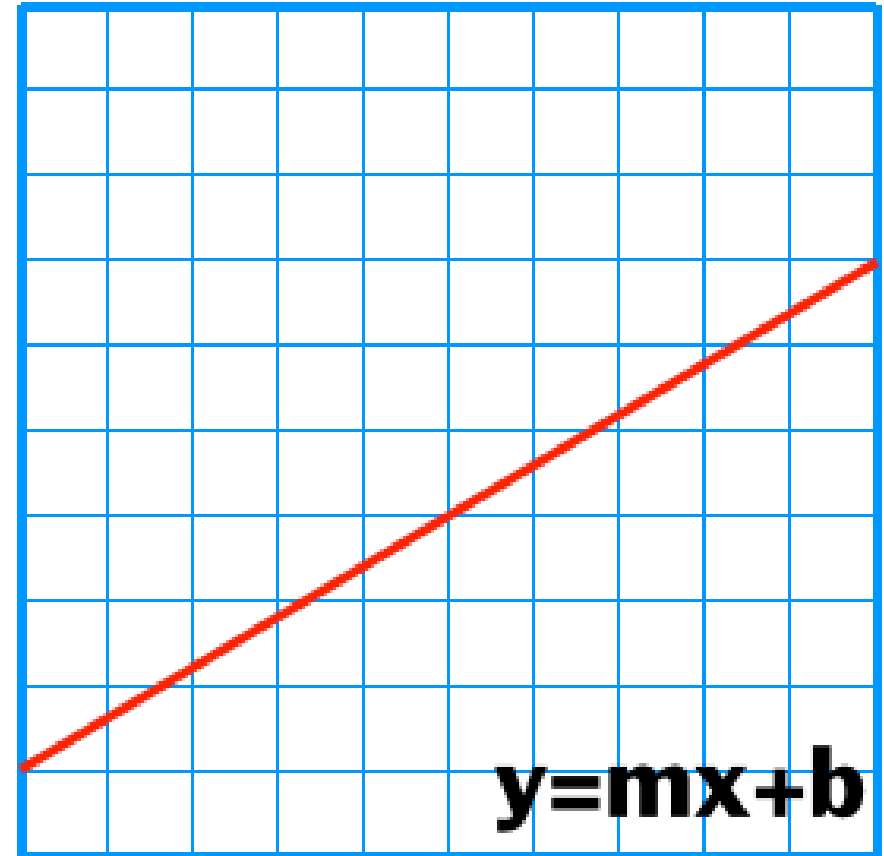
Straight Line

## Meaning: Proportional

Direct Relationship, as x increases, y increases variable; constant rate of change

# Linear

$$y = mx + b$$



# Power Model

## Equation

$$y = ax^c + b$$

y = dependent variable  
x = independent variable  
a = coefficient – scaling  
c = power  
B = y intercept – initial conditions

## Shape

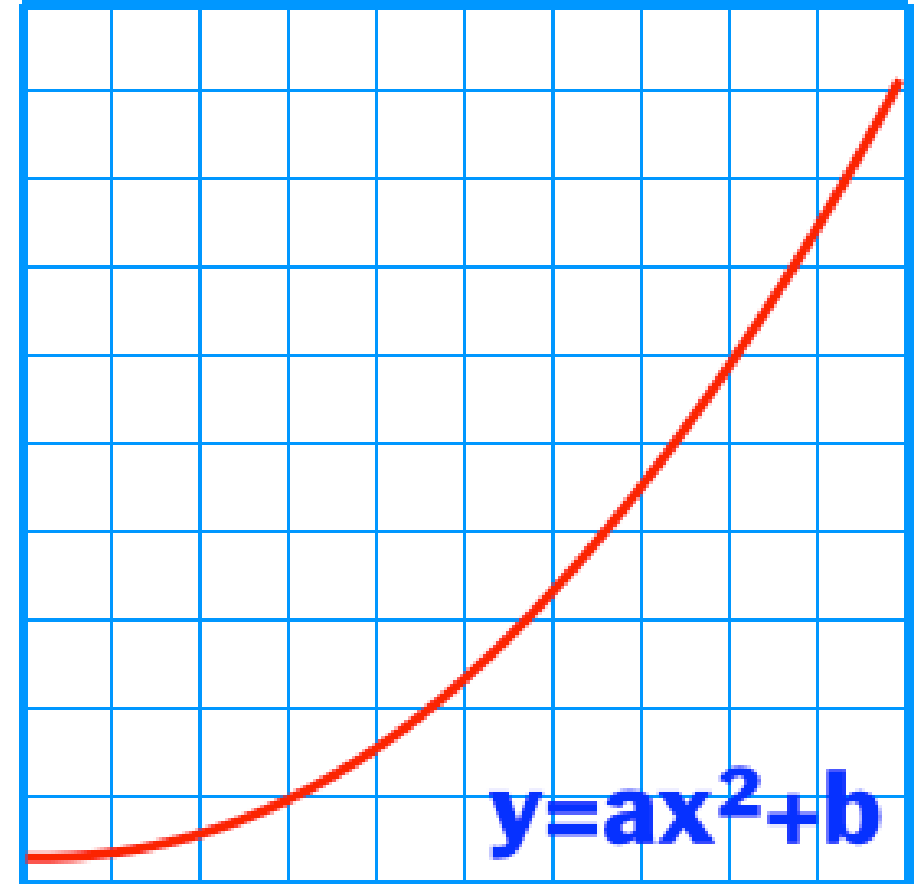
Curved line

## Meaning: Proportional

Power Relationship, as x increases, y increases as a square; does not increase at a constant rate; changes the rate it increases

**Power**

$$y = ax^c + b$$



# Inverse Square Model

## Equation

$$y = a \frac{1}{x^2}$$

y = dependent variable  
x = independent variable  
a = coefficient – scale  
c = power

## Shape

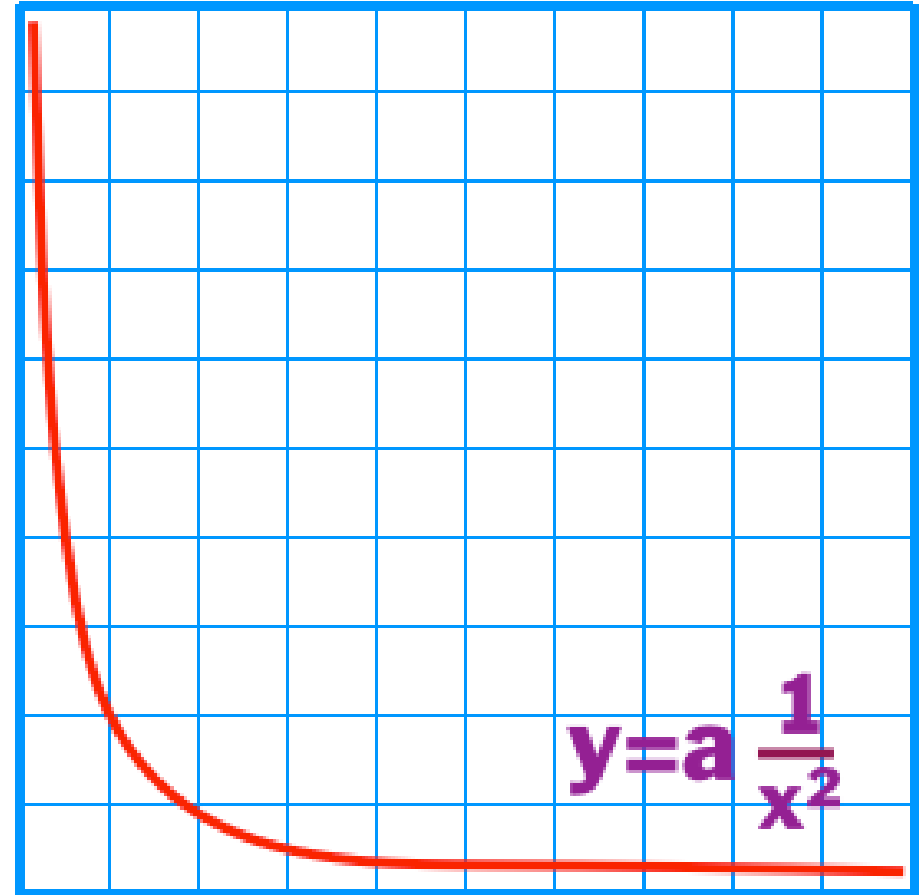
Steep Drop off; downward curve

## Meaning: Proportional

Inverse square Relationship, as x increases, y decreases as a square; factor quickly dissipates; only visible at close ranges

## Inverse

$$y = a \frac{1}{x^c}$$





# Square Root Model

## Equation

$$y = a\sqrt{x}$$

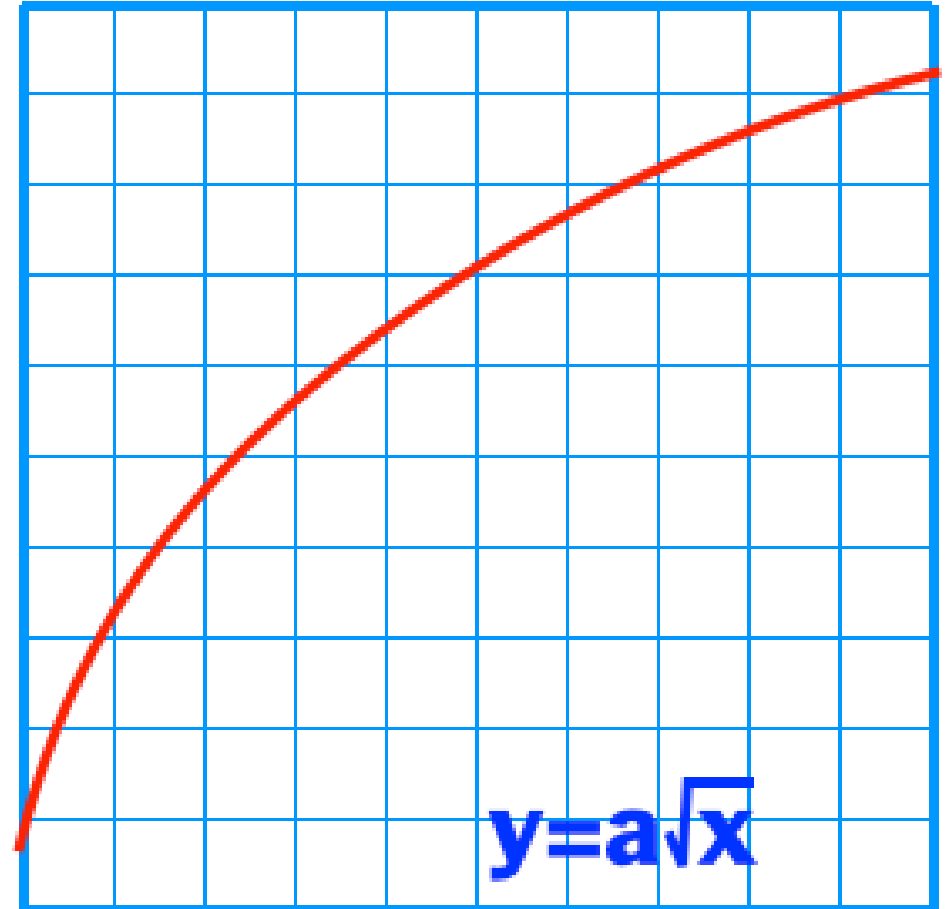
y = dependent variable  
x = independent variable  
a = coefficient – scale

## Shape

Curve that plateaus over a long interval

## Meaning: Proportional

Power Relationship, as x increases, y decreases as a square; increases quickly and then effect goes away



# Inverse Model

## Equation

$$y = a \frac{1}{x}$$

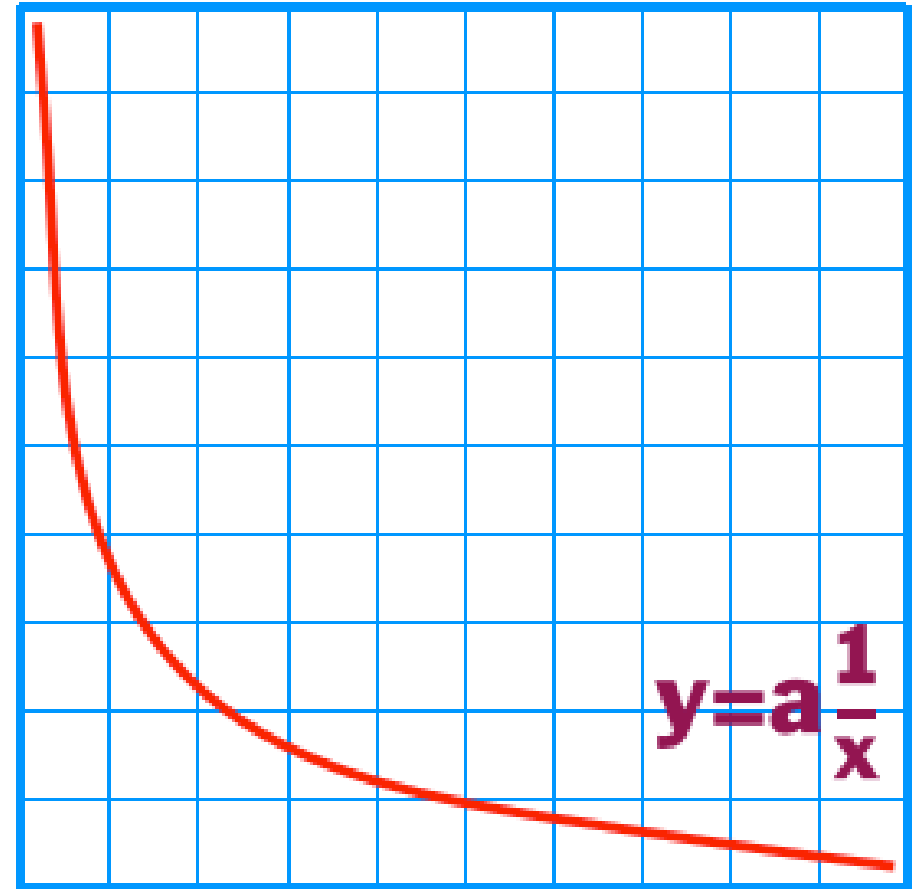
y = dependent variable  
x = independent variable  
a = coefficient – scale

## Shape

Curve that plateaus over a long interval

## Meaning: Proportional

Power Relationship, as x increases, y decreases as a square; increases quickly and then effect goes away



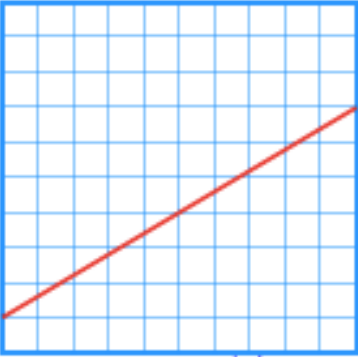
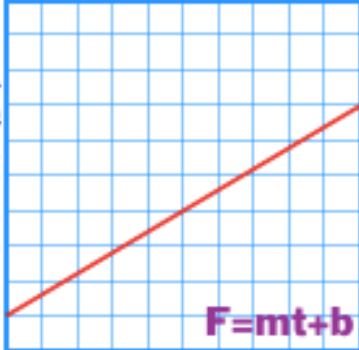
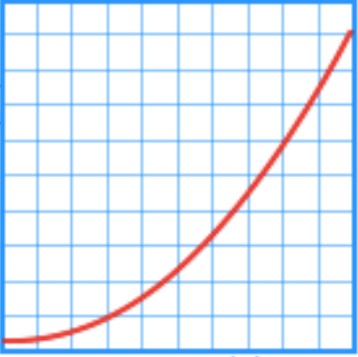
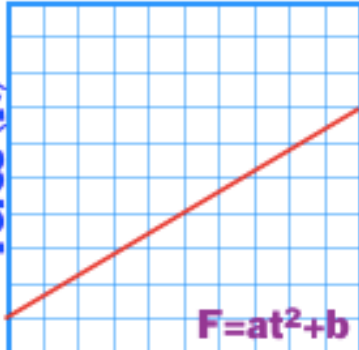
# Linearization

- Vary the power of the variables to generate a linear trend
- Generation of an equation

Ex: Kepler's Third Law

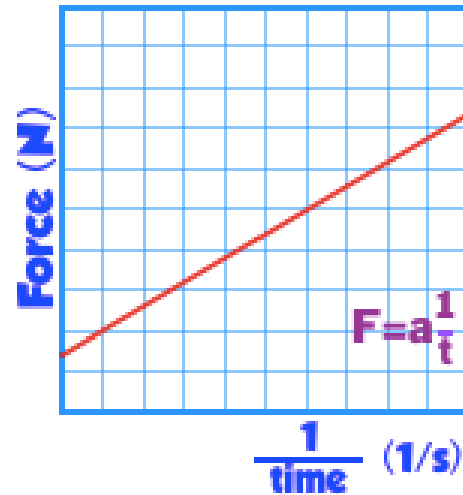
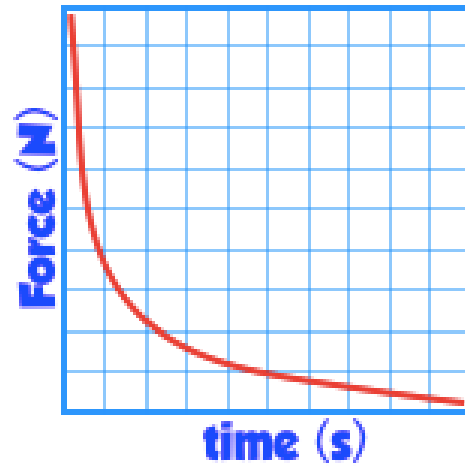
What is the relationship between the average orbital distance ( $r$ ) of the planets and their average orbital period ( $T$ )?

# Example Relationships

What the graph of raw data looks like	What the linearized data looks like & its equation.	Description in words
Linear		
 <p data-bbox="196 499 242 678">Force (N)</p> <p data-bbox="369 778 509 821">time (s)</p>	 <p data-bbox="797 499 843 678">Force (N)</p> <p data-bbox="963 778 1103 821">time (s)</p> <p data-bbox="1031 728 1184 771"><math>F=mt+b</math></p>	<p data-bbox="1403 578 2000 664">Force is directly proportional to the time.</p>
Power		
 <p data-bbox="196 978 242 1156">Force (N)</p> <p data-bbox="369 1256 509 1299">time (s)</p>	 <p data-bbox="797 978 843 1156">Force (N)</p> <p data-bbox="937 1249 1116 1292">time<sup>2</sup> (s<sup>2</sup>)</p> <p data-bbox="1031 1206 1184 1249"><math>F=at^2+b</math></p>	<p data-bbox="1403 1056 1949 1142">Force is proportional to time squared.</p>

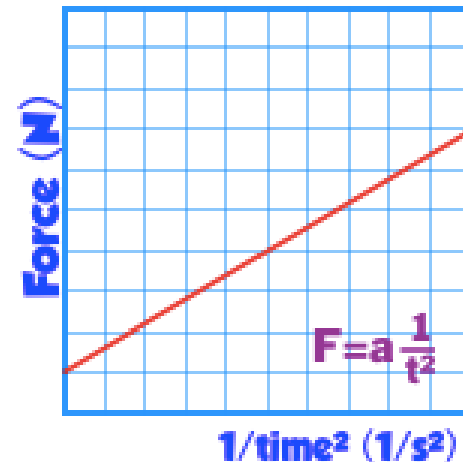
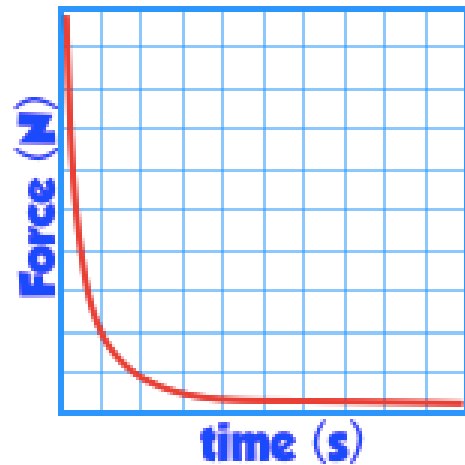
# Example Relationships

Inverse



Force is proportional to the inverse of time.

Inverse squared



Force is proportional to the inverse square of time.

# Correlation Data

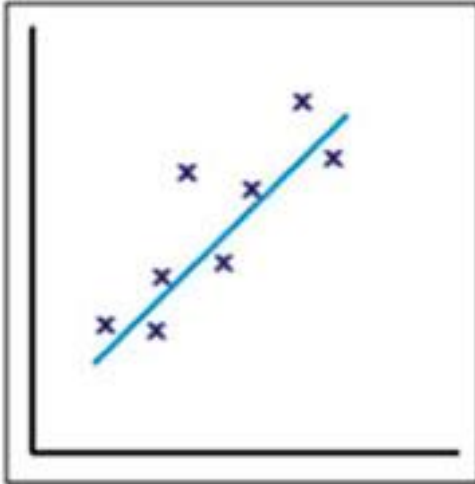


- Determines the relationship between two variables
- Determines the predictability of two variables
- $R^2$  on excel Plot measures the percent of verifiability in y in relation to x & y.

Correlation does not determine causality

# Correlation Examples

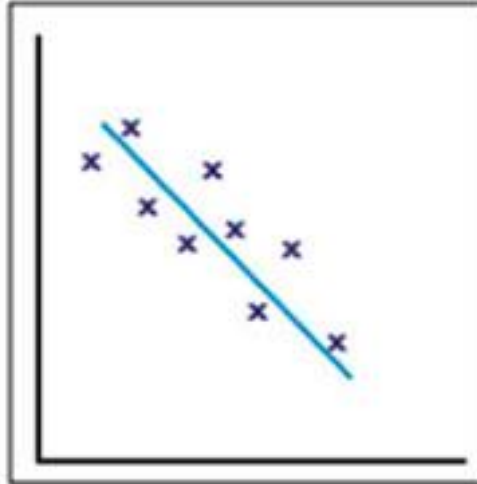
**Positive correlation**



The points lie close to a straight line, which has a positive gradient.

This shows that as one variable **increases** the other **increases**.

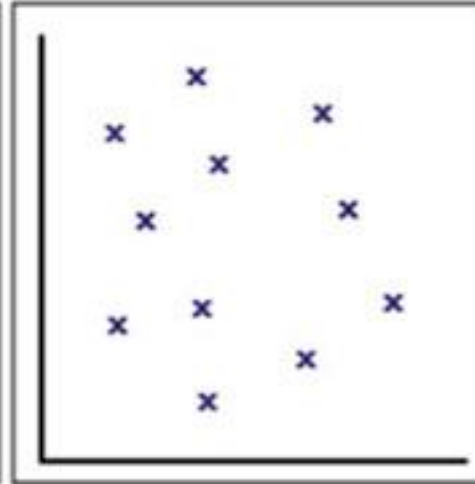
**Negative correlation**



The points lie close to a straight line, which has a negative gradient.

This shows that as one variable **increases**, the other **decreases**.

**No correlation**

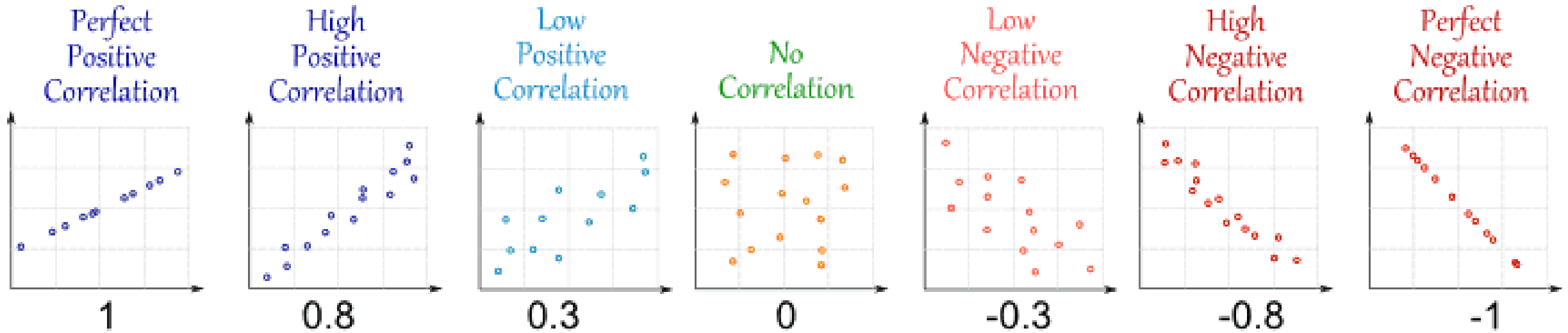


There is no pattern to the points.

This shows that there is **no connection** between the two variables.

Visuals to match their data with the most appropriate correlation

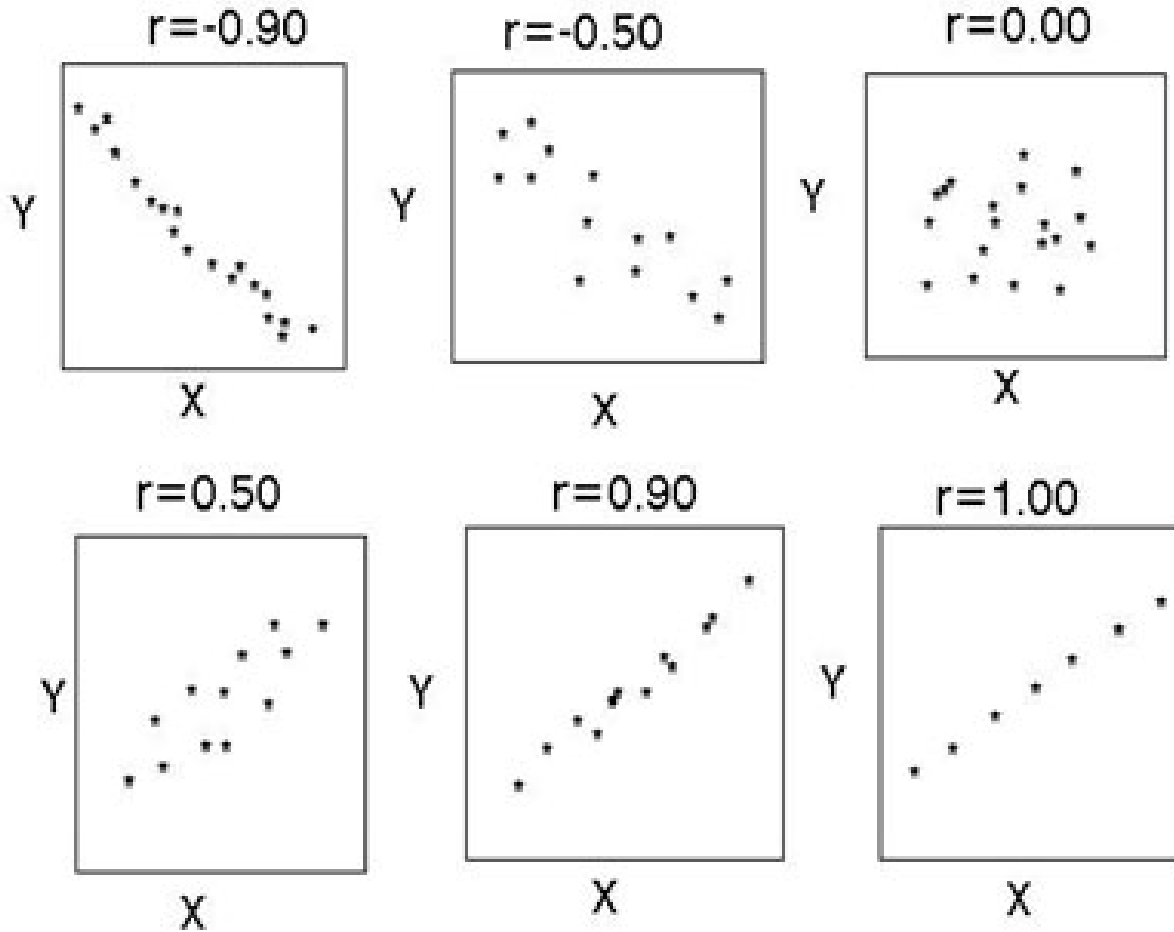
# Correlation & R<sup>2</sup> Examples



- Positive correlation refers to an upward trend
- Negative correlation refers to a downward trend
- 0 correlation shows no relation between variables



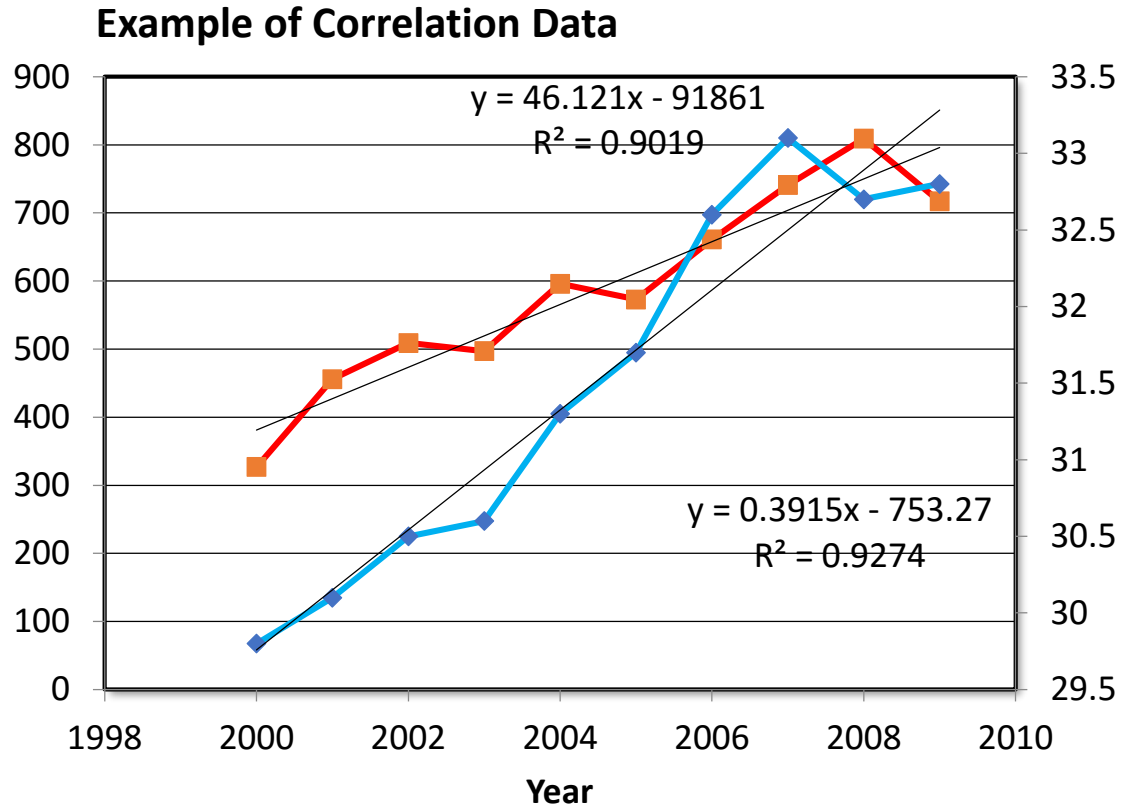
# Correlation Examples



Using Excel:

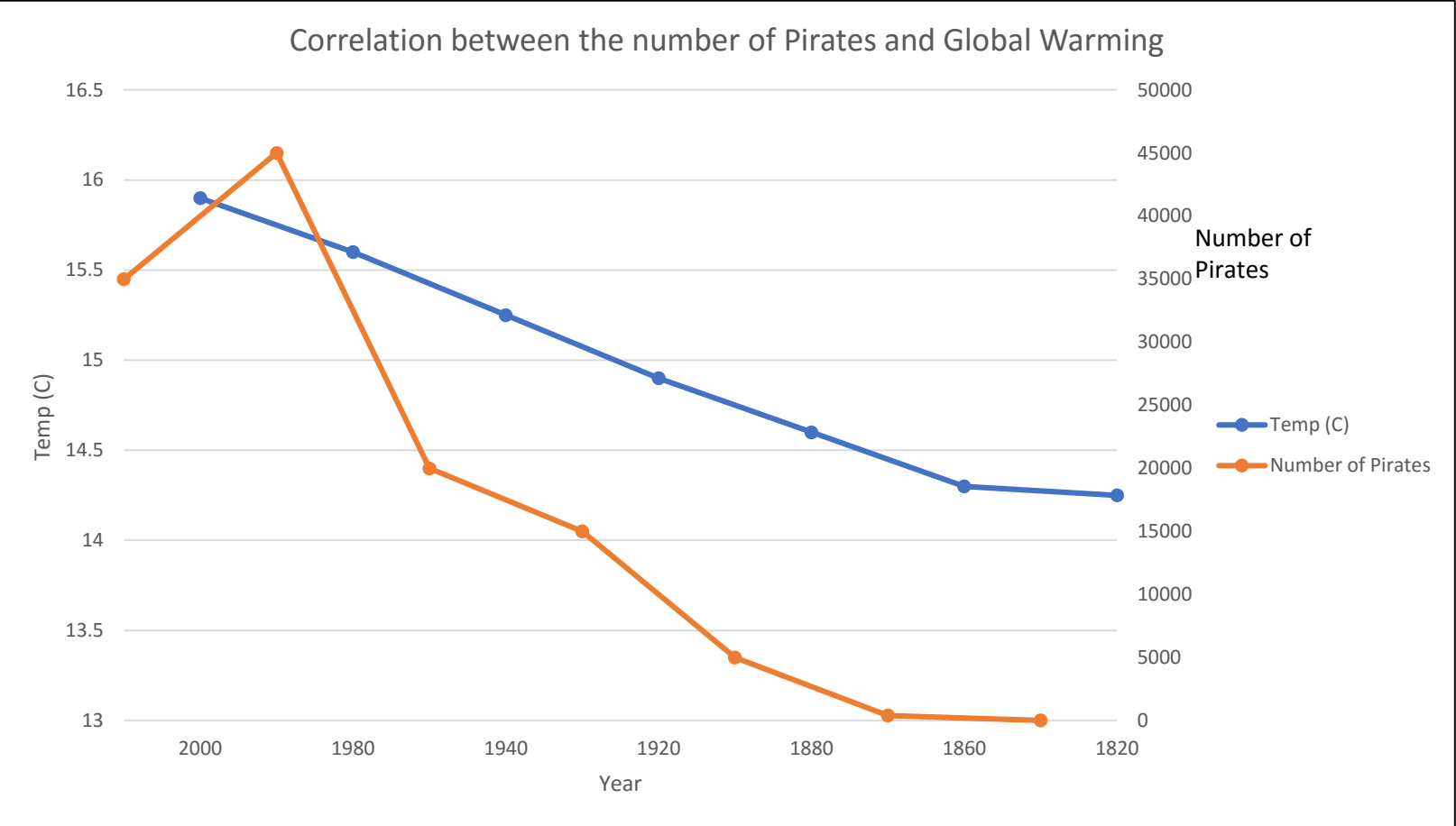
- Right click on data series
- Select Add Trendline
- Select Display equation
- Select Show  $R^2$
- CORREL = 0.941

# Correlation Between Two Data Sets



- Individual correlation values are different than the correlation between two data sets
- Use CORREL Function in excel to find correlation between two data sets
- CORREL = 0.941

# Correlations between two data sets



R = -0.93

Correlation and Causality

**Pirate shortage is the cause of global warming**

# Citations

- Wayne, Tony. "Linearizing data." *Linearizing data*. N.p., n.d. Web. 09 Mar. 2017.
- "Practice 4". *Participatoryscience.org*. N.p., 2017. Web. 8 Mar. 2017.
- McLeod, S. A. (2008). Correlation. Retrieved from [www.simplypsychology.org/correlation.html](http://www.simplypsychology.org/correlation.html)
- "The Scatter Plot & Linear Regression." *CQE Academy*. N.p., n.d. Web. 08 Mar. 2017.