

SCATTER PLOTS AND LINES OF BEST FIT



An effective way to see a relationship in data is to display the information as a scatter plot.

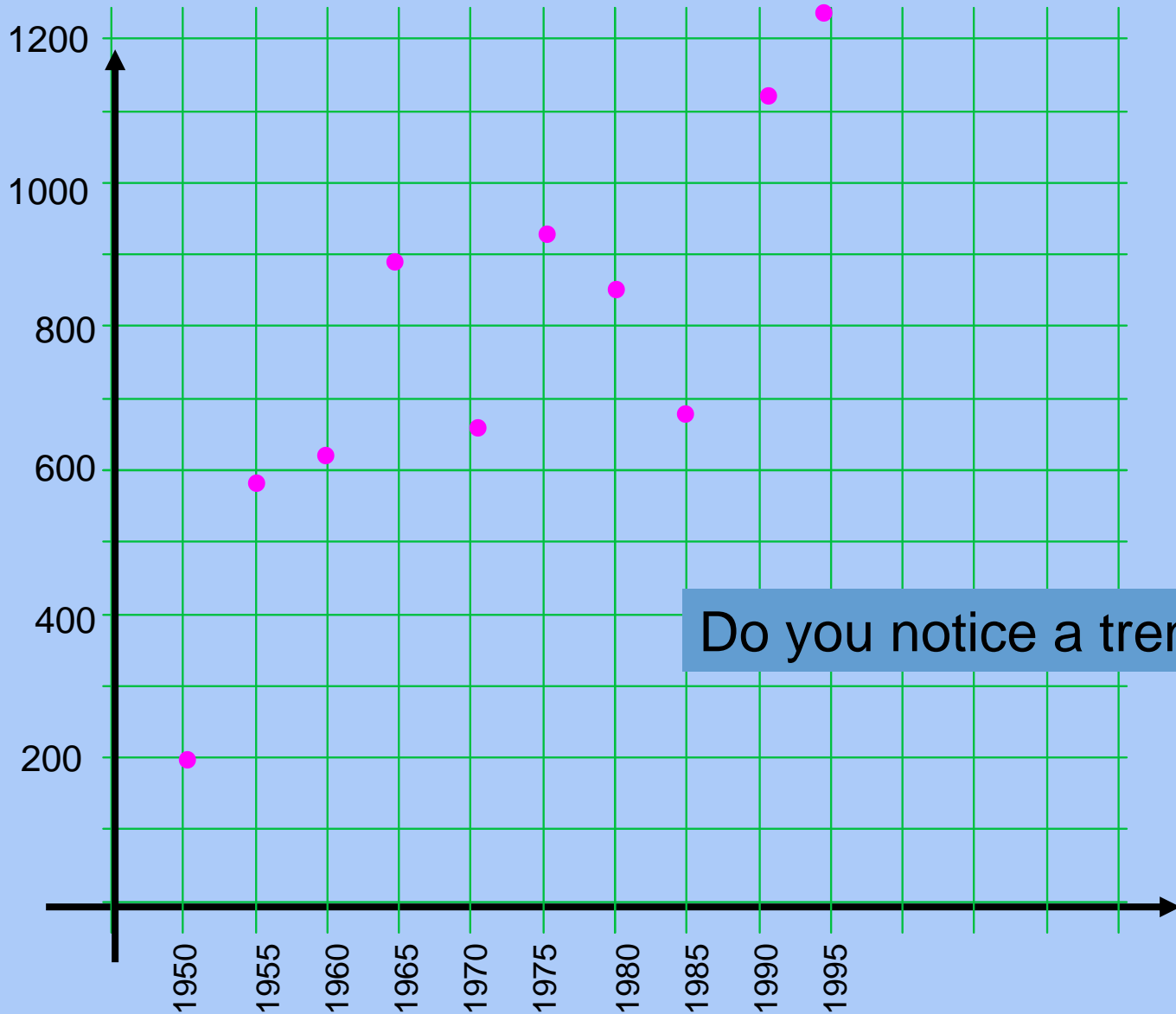
It shows how two variables relate to each other by showing how closely the data points fit to a line.

The following table presents information on tornado occurrences. Make a scatter plot for the table.

Year	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995
# of Tornadoes	201	593	616	897	654	919	866	684	1133	1234



Year	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995
# of Tornadoes	201	593	616	897	654	919	866	684	1133	1234



Do you notice a trend?



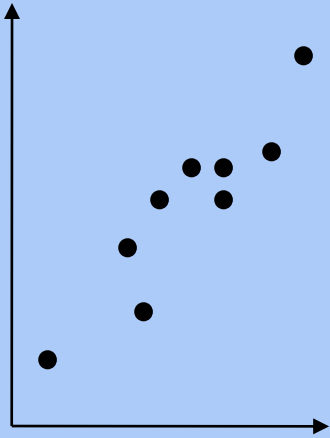
Scatter plots provide a convenient way to determine whether a correlation exists between two variables.

A positive correlation occurs when both variables increase.

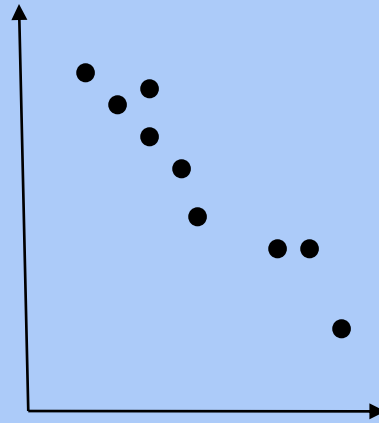
A negative correlation occurs when one variable increases and the other variable decreases.

If the data points are randomly scattered there is little or no correlation.

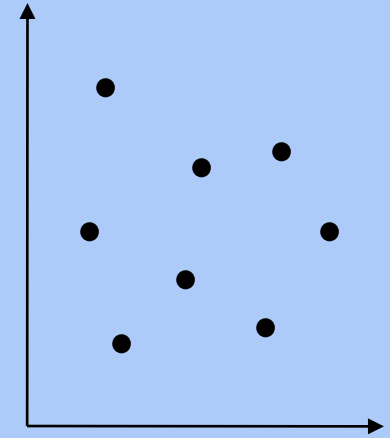




Positive
correlation



Negative
correlation



little or no
correlation

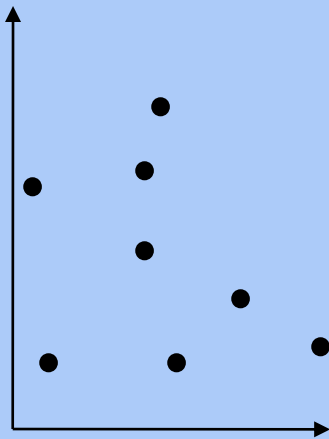


Example 1:

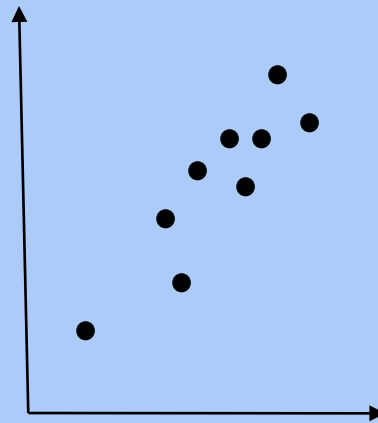
The scatter plots of data relate characteristics of children from 0 to 18 years old.

Match each scatter plot with the appropriate variables studied.

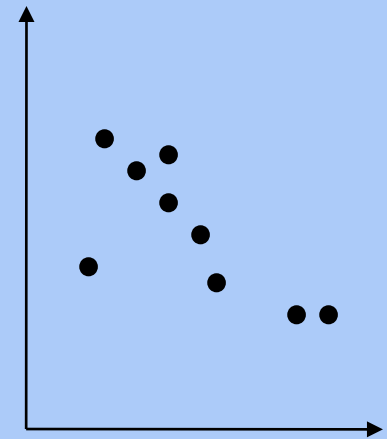
1. age and height
2. age and eye color
3. age and time needed to run a certain distance



no correlation
between **2** age
and eye color



as your age increases
your height **1** also increases



as your age
increases **3** the time
will decrease



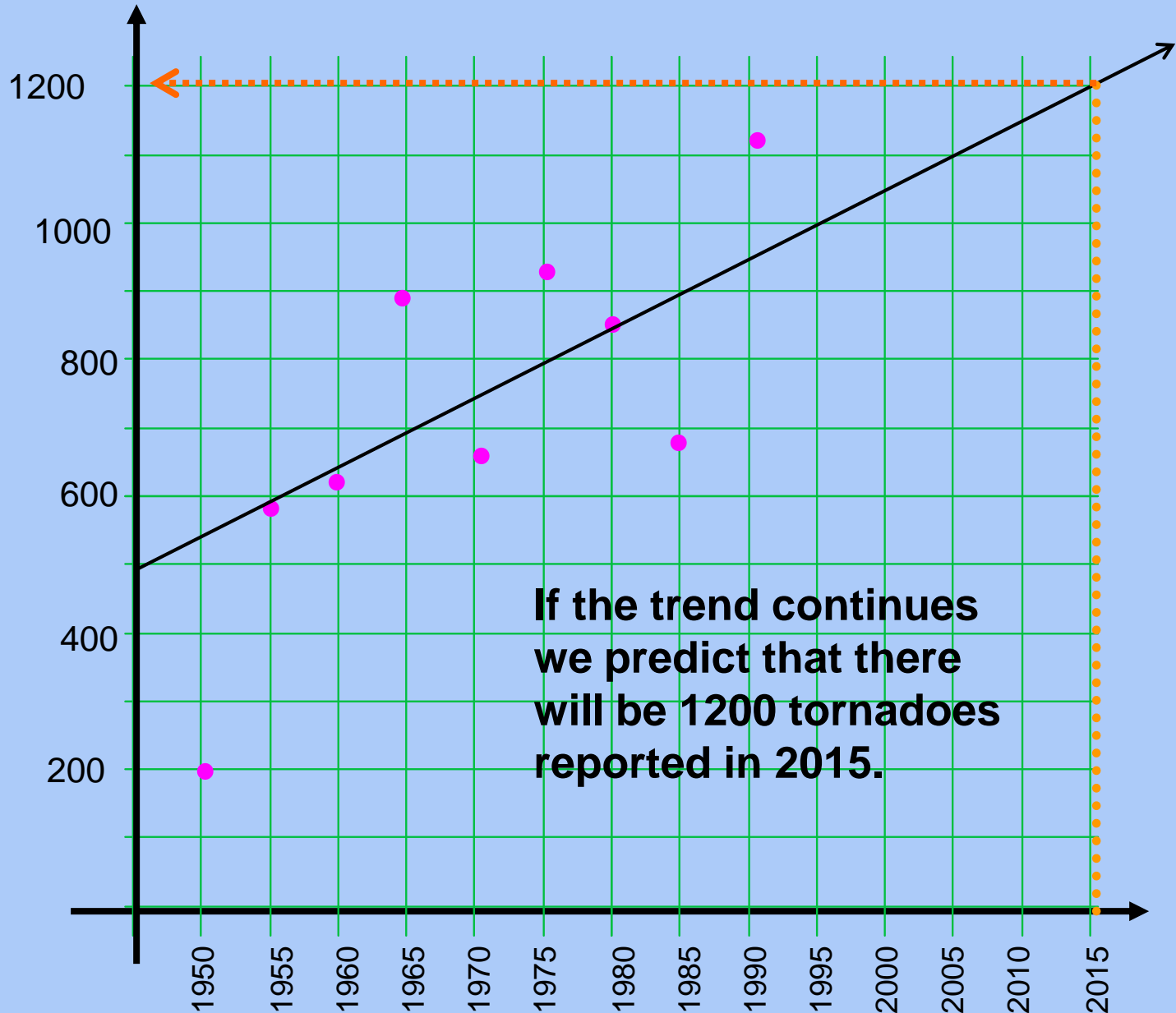
Sometimes points on a scatter plot are represented by a trend line or a line of best fit.

You can study the line to see how the data behaves. You may have a basis to predict what the data might be for values not given.

Example 2: Find the line of best fit for the scatter plot you made on the first page. To fit the line to the points, choose your line so that it best matches the overall trend. The line ***does not have to*** pass through any of the points.



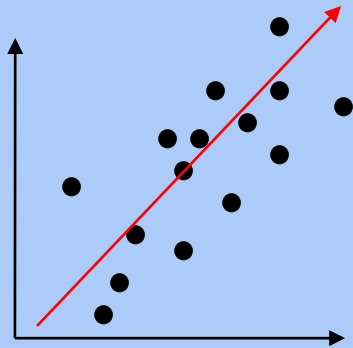
Use the line of best fit to predict how many tornadoes may be reported in the United States in 2015 if the trend continues.



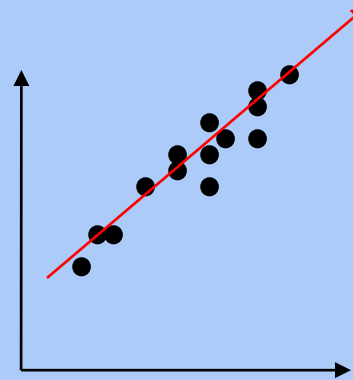
**If the trend continues
we predict that there
will be 1200 tornadoes
reported in 2015.**



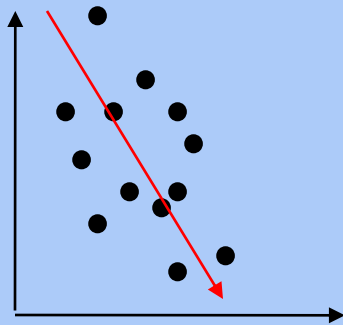
If the data points are close to the line of best fit, it is said to have a strong correlation.



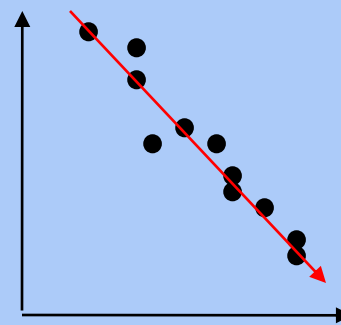
weak positive



strong positive



weak negative



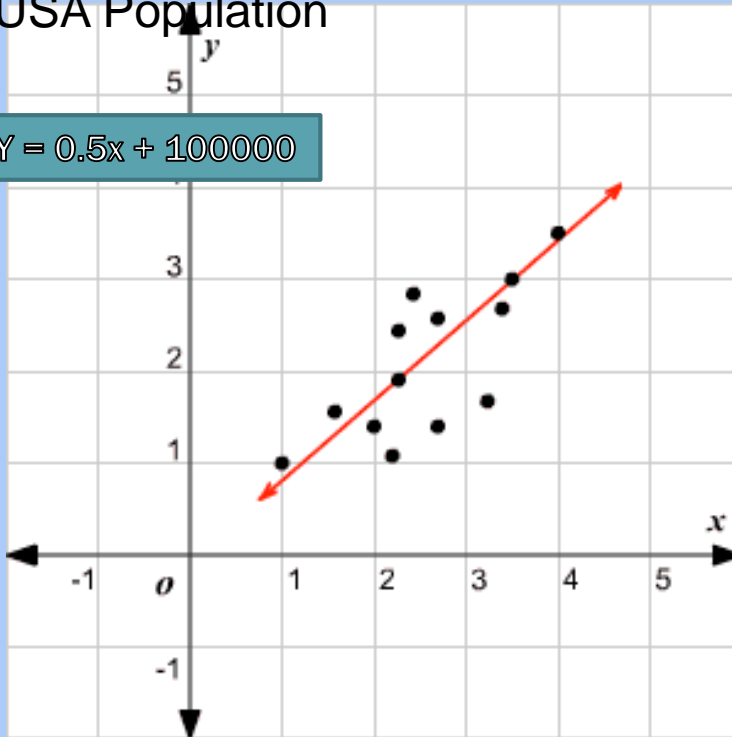
strong negative



TIP Math

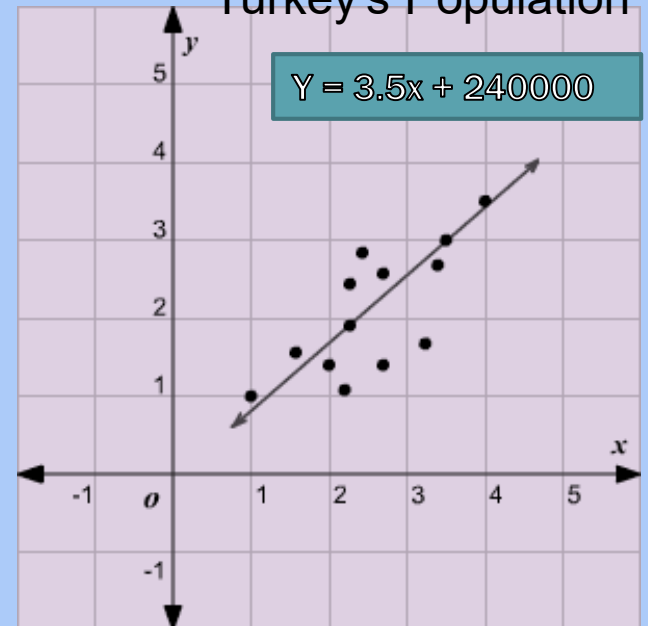
USA Population

$$Y = 0.5x + 100000$$



Turkey's Population

$$Y = 3.5x + 240000$$



<http://data.worldbank.org/indicator/SP.POP.TOTL?page=2>

