Correlation

- A correlation examines the relationship between two measured variables.

– No manipulation by the experimenter/just observed. – E.g., Look at relationship between height and weight.

- You can correlate any two variables as long as they are numerical (no nominal variables)
- Is there a relationship between the height and weight of the students in this room?

– Of course! Taller students tend to weigh more.

- Two aspects of the relationship two aspects of the relationship: Strength and Direction.
- The relationship between any two variables is rarely a perfect correlation.
- Perfect correlation: +1.00 OR –1.00 – strongest possible relationship – Tough to find.
- No correlation: 0.00 (no relationship). – e.g., height and social security.

- The value of correlation (i.e., correlation coefficient) does not depend on the specific measurement units used; for example,

– the correlation between height and weight will be identical regardless of whether inches and pounds, or centimeters and kilograms are used as measurement units.

- Pearson correlation -Parametric
- Spearman correlation-Non-Parametric

- The Pearson correlation evaluates the linear relationship between two continuous variables.
- A relationship is linear when a change in one variable is associated with a proportional change in the other variable.
- For example, you might use a Pearson correlation to evaluate whether increases in temperature at your production facility are associated with decreasing thickness of your chocolate coating.

- The Spearman correlation evaluates the monotonic relationship between two continuous or ordinal variables.
- In a monotonic relationship, the variables tend to change together, but not necessarily at a constant rate.
- The Spearman correlation coefficient is based on the ranked values for each variable rather than the raw data.
- Spearman correlation is often used to evaluate relationships involving ordinal variables.
- For example, you might use a Spearman correlation to evaluate whether the order in which employees complete a test exercise is related to the number of months they have been employed.

- r = the Pearson coefficient
- r measures the amount that the two variables (X and Y) vary together (i.e., covary) taking into account how much they vary apart
- Pearson's r is the most common correlation coefficient

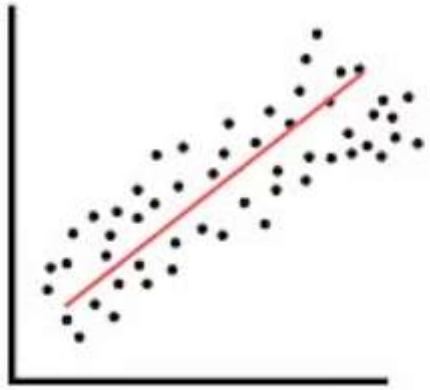
- What is the correlation between study time and test score:

X (hours)	Y (score)
0	30
10	90
4	30
8	60
8	90

- The correlation coefficient (r) represents the linear relationship between two variables.
- If the correlation coefficient is squared, then the resulting value (r2 , the coefficient of determination) will represent the proportion of common variation in the two variables (i.e., the "strength" or "magnitude" of the relationship).
- Correlation coefficients cannot simply be averaged

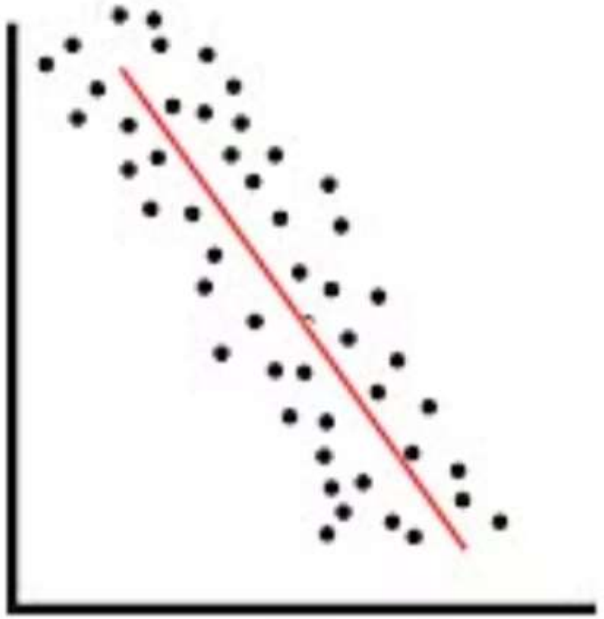
- There are two directions in which this change might occur, producing what are known as **positive** correlations and **negative** correlations.

- POSITIVE CORRELATION occurs when one variable increases as the other increases.
- For instance: l the further you walk, the more money you collect for charity;
- the more papers you have to deliver, the longer it takes you.



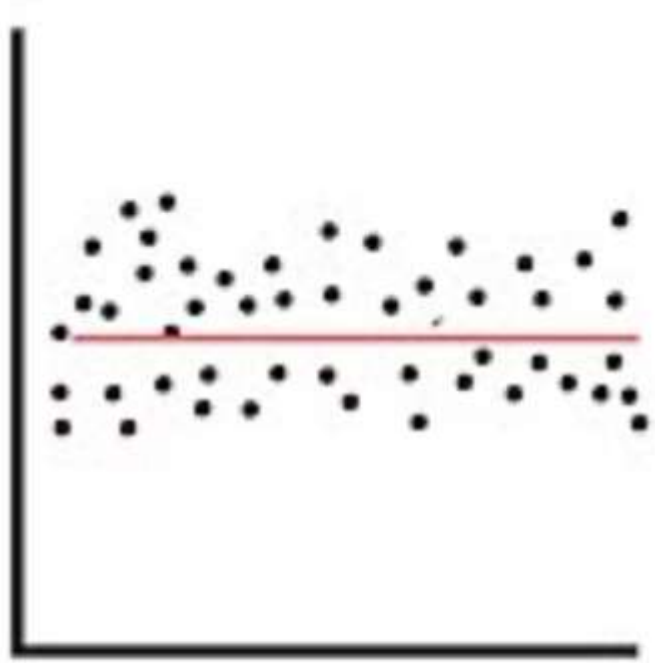
Positive Correlation

- **Negative Correlation**: A negative correlation is just the opposite. The relationship line has a negative slope, and the variables change in opposite directions, i.e., one variable decreases while the other increases.



Negative Correlation

- No Correlation: No correlation simply means that the variables behave very differently and thus, have no linear relationship.



No Correlation