# getting the (wrong) picture from the data – Correlation, causality

## Data

Researchers measured the foot size (in cm) of a large number of students. They also gave them a dictation and counted the number of "mistakes" made by each. Data were collected on all classes in an elementary school. Here are the collected data :

| feet size | 17.5 | 17.5 | 17.5 | 17.5 | 18 | 18 | 18 | 18 | 18.5 | 18.5 | 18.5 | 19 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| number of mistakes | 15 | 18 | 19 | 20 | 16 | 17 | 18 | 19 | 14 | 16 | 17 | 15 | 16 | 13 |

| 20 | 20 | 20.5 | 20.5 | 20.5 | 20.5 | 21 | 21 | 21 | 21 | 21.5 | 21.5 | 21.5 | 22 | 22 | 22 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 14 | 15 | 12 | 13 | 14 | 15 | 10 | 11 | 13 | 15 | 10 | 12 | 13 | 8 | 10 | 11 | 12 | 8 |

| 23 | 23 | 23.5 | 23.5 | 23.5 | 23.5 | 24 | 24 | 24 | 24.5 | 24.5 | 24.5 | 24.5 | 25 | 25 | 25 | 25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 10 | 7 | 8 | 9 | 11 | 6 | 8 | 9 | 6 | 7 | 8 | 10 | 4 | 6 | 7 | 8 |

| 25.5 | 25.5 | 26 | 26 | 26 | 26.5 | 26.5 | 26.5 | 27 | 27 | 27 | 27 | 27.5 | 27.5 | 28 | 28 | 28 | 28 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 6 | 4 | 5 | 7 | 3 | 4 | 5 | 2 | 3 | 4 | 7 | 2 | 3 | 0 | 1 | 2 | 4 |

| 28.5 | 28.5 | 29 | 29 | 29 |
|---|---|---|---|---|
| 0 | 2 | 0 | 1 | 2 |

## Steps

— Propose a graph to represent the above data

— At the same time, note the important points in your approach to building the graph

— Why did you make that graph ?

— What can you calculate to make a summary of the variables ?

— What can be said about the size of students' feet and the number of mistakes they make in dictation ?

— What do you deduce from this graph ? Does it correspond to your intuition ?

— Is there a correlation between the two quantities ? positive or negative ? Is there a causal link ?

— What is happening ?