

# INTERPRETATION OF STANDARD DEVIATION AND VARIANCE

- **Standard deviation** =  $\sqrt{\text{Variance}}$
- Basically a **small SD** means that the values in statistical dataset <sup>are</sup> ~~means~~ **close to the mean of the dataset, on average** and a **large standard deviation** means that the value in the dataset are **farther away from the mean, on average**.
- In short, it measures how **concentrated the data around mean**, more concentrated → **smaller SD**.
- A **small SD** can be the goal in certain situations where results are restricted for **example in product manufacturing and quality control**. A particular type of car part that has to be 2 centimeters in diameter to fit properly had better not to have a very big standard deviation during manufacturing process. A big standard deviation would mean it will end up in trash can because they don't fit right.
- **High SD**, reflects a large amount of variation in the group. For **example**, if we look at the **salaries for everyone in a company**, including student intern to the CEO, standard deviation may be very large. On other hand, if we observe only student interns salary, standard deviation may be low/smaller.
- Outlier does affect the SD, because formula includes the mean. SD cannot be ~~0~~ negative and lowest possible value is 0 and 0 is possible only when every single entity have same number (no deviation)
- SD have the same unit as the original data.