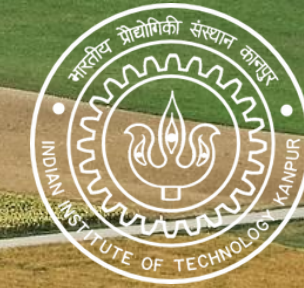


Week-01-L-03

Statistics for Agriculturists

How do we use statistics in Agriculture?

Prof. J. Ramkumar
Department of ME & Design
Indian Institute of Technology Kanpur



MedTech
IIT KANPUR



Stats, figures and data



Data
Types

Qualitative Data

- Deals with description
- Can't be measured
- Qualities of system like color, smell, appearance etc.

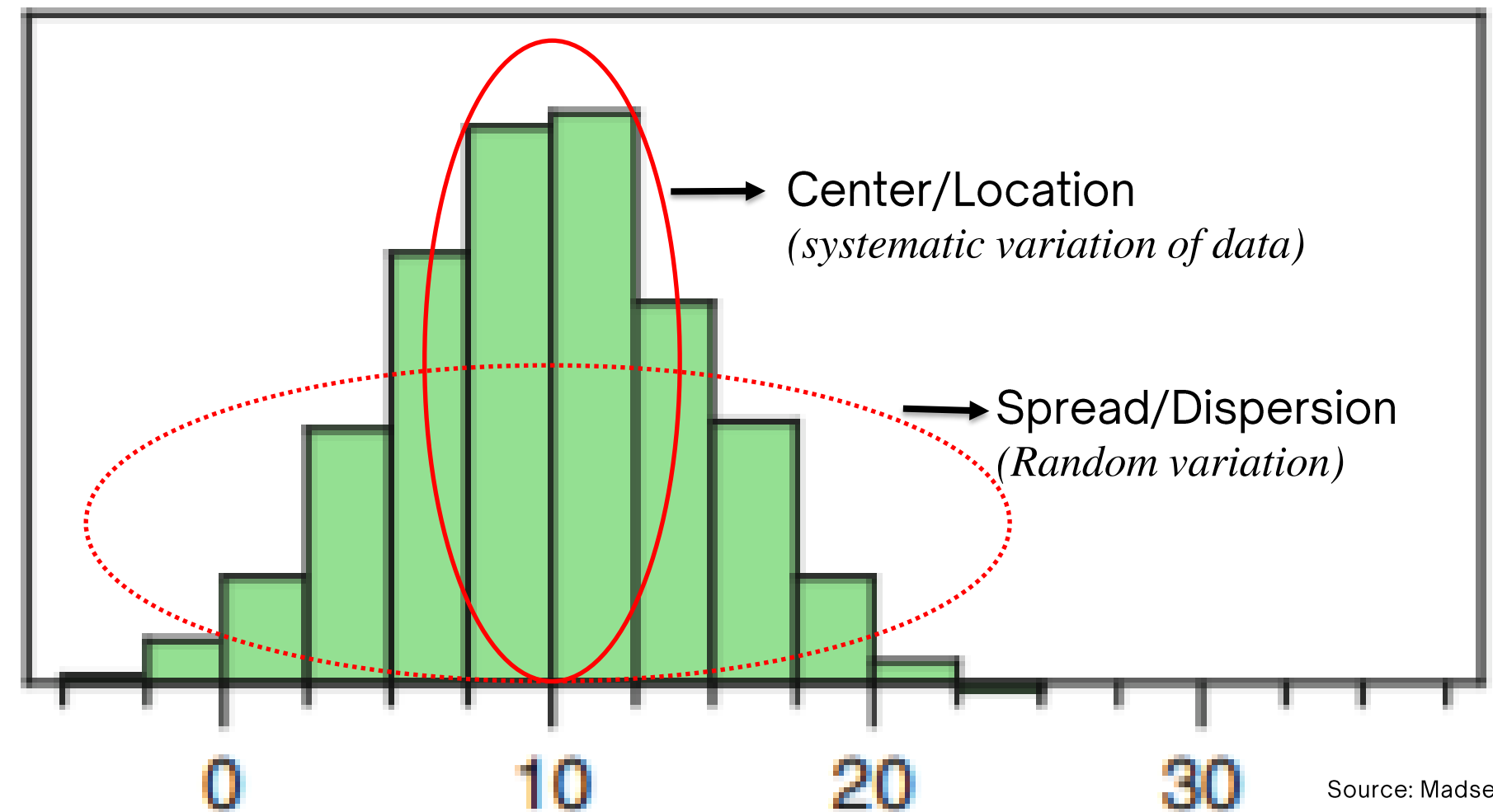
Quantitative Data

- Deals with numbers
- Can be measured
- Quantifies the system like length, temperature, weight etc.

How is Statistics used in Agriculture?



- Statistics is about describing the variation in data.
- We use word location and dispersion to characterize distribution of data.

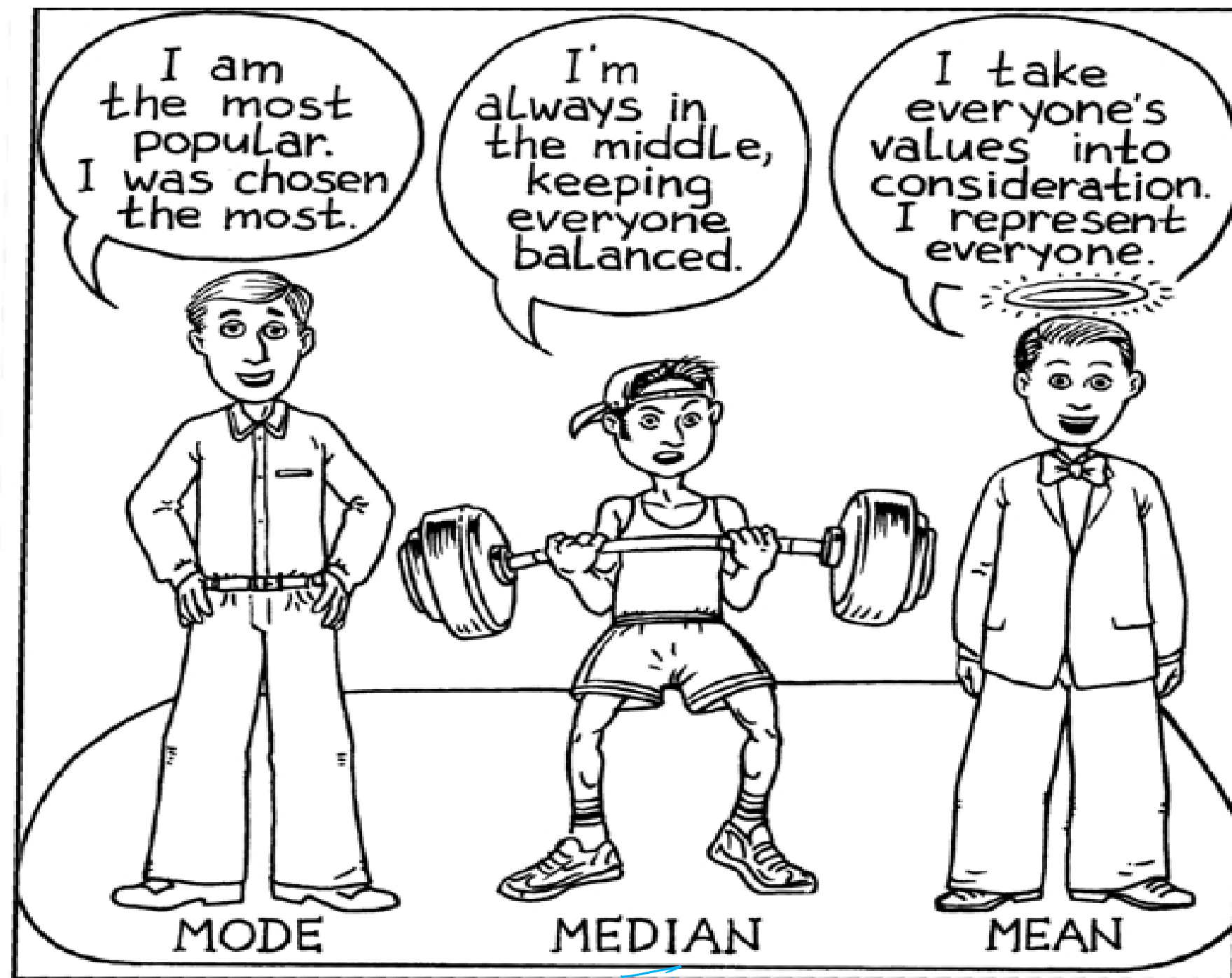


Measures of location



There are three ways to measure the location of data

- Mean
- Median
- Mode



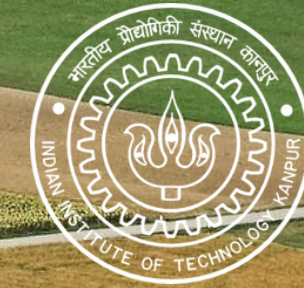
$$\begin{array}{r} 12 \qquad 18 \\ \textcircled{2} \\ \hline = \frac{32}{3} \approx 11 \end{array}$$

Week-01-L-03

Statistics for Agriculturists

How do we use statistics in Agriculture?

Prof. J. Ramkumar
Department of ME & Design
Indian Institute of Technology Kanpur



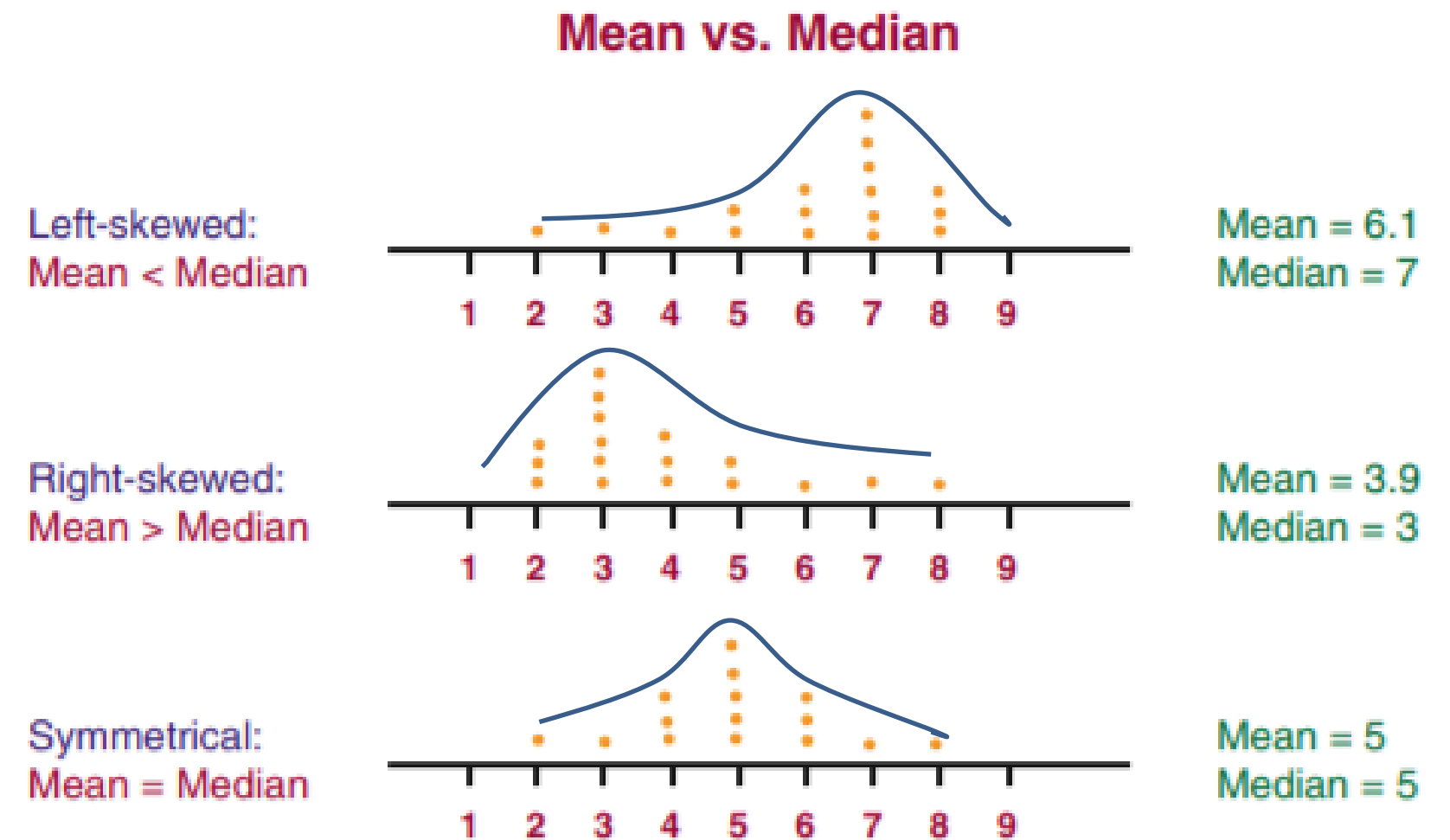
MedTech
IIT KANPUR



How to choose?



- Data is symmetrically presented choose mean or median.
-
- Data isn't symmetrically arranged or it's skewed choose median.
- As both median and mean would be different, this will have two cases right skewed or left skewed.



Measures of dispersion



There are three ways to measure the dispersion of data

- **Range**

$$= x_{max} - x_{min}$$

$\phi 50 \text{ mm}$
 $\phi 50 \pm 5 \text{ mm}$
 $\phi 45 \sim 55 \text{ mm}$

- **Standard Deviation & Variance**

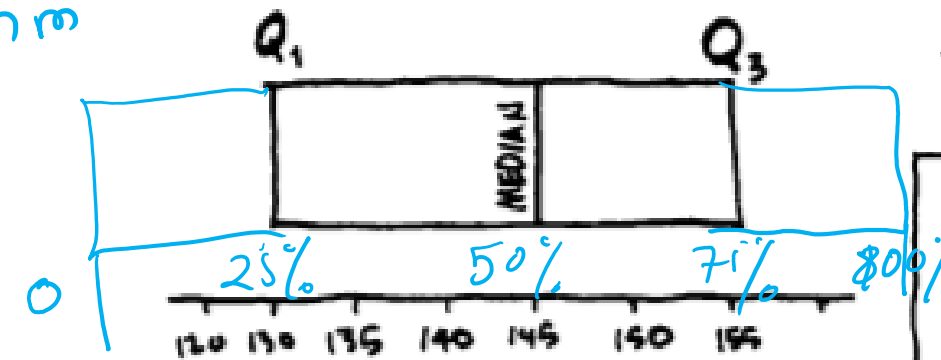
$$= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sigma$$

- **Interquartile Range**

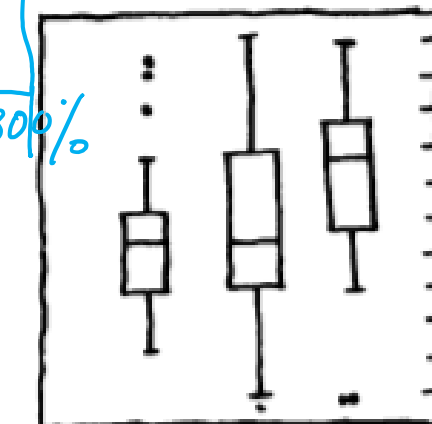
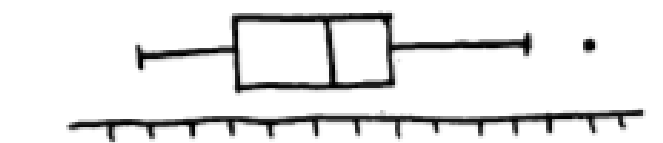
$$= Q_3 - Q_1$$

JOHN TUKEY INVENTED ANOTHER KIND OF DISPLAY TO SHOW OFF THE IQR, CALLED A **BOX AND WHISKERS PLOT**. THE BOX'S ENDS ARE THE QUANTILES Q_1 AND Q_3 . WE DRAW THE MEDIAN INSIDE THE BOX.

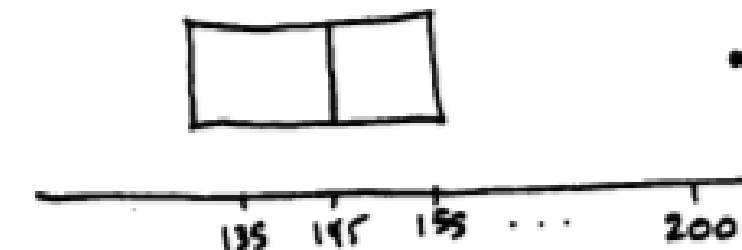
FINALLY, EXTEND "WHISKERS" OUT TO THE FARTHEST POINTS THAT ARE NOT OUTLIERS (I.E., WITHIN 1.5 IQR OF THE QUANTILES).



IF A POINT IS MORE THAN 1.5 IQR FROM AN END OF THE BOX, IT'S AN **OUTLIER**. DRAW THE OUTLIERS INDIVIDUALLY.



BOX-AND-WHISKERS PLOTS ARE ESPECIALLY GOOD FOR SHOWING OFF DIFFERENCES BETWEEN GROUPS.



How to choose?



- Distribution is symmetrical, we often use the average as a measure of location, so to supplement it standard deviation is used as a measure of dispersion.
- The standard deviation is, after all, based on the average.
- Distribution is skewed, we often use median as a measure of location, so to supplement it IQR is used as a measure of dispersion.

Relative Spread (Dispersion)

when comparing samples from several time periods

→ the average will often increase with time

→ the spread increases with an increasing average



Thank you