Exploratory Statistical Data Analysis With R Software (ESDAR)
Swayam Prabha

Lecture 23

Variation in Data and Measurement of Variability

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Slides can be downloaded from

http://home.iitk.ac.in/~shalab/sp

1

Measures of Variation (or Dispersion)

Measures of central tendency gives an idea about the location where most of the data is concentrated.

Two different data sets may have same arithmetic mean but they may have different concentrations around mean.

Measures of Variation (or Dispersion)

Example: The temperature of three cities in degree centigrade on 6 days are recorded as follows:

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
City 1	0	0	0	0	0	0
City 2	-15	-15	-15	15	15	15
City 3	11	9	10	8	12	10

Arithmetic mean of the city 1 = 0

Arithmetic mean of the city 2 = 0

Arithmetic mean of the city 3 = 10

Measures of Variation (or Dispersion)

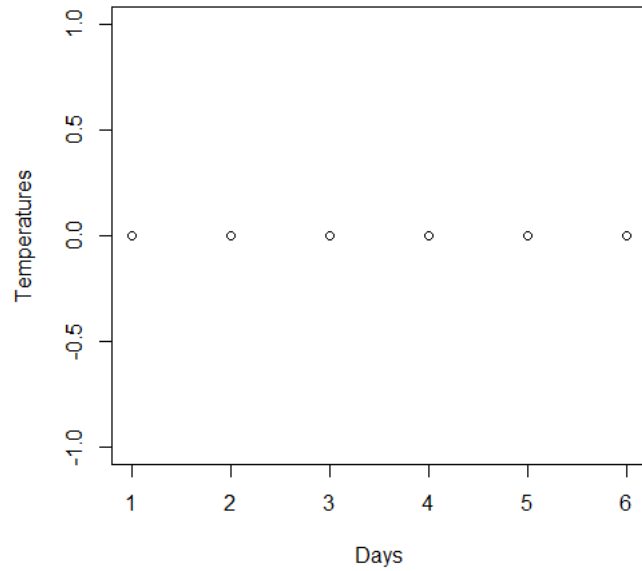**Mean Temperatures in City 1 and City 2 are the same as 0 but does this makes any sense?**

City 2 has two extreme temperatures on -15 and 15.

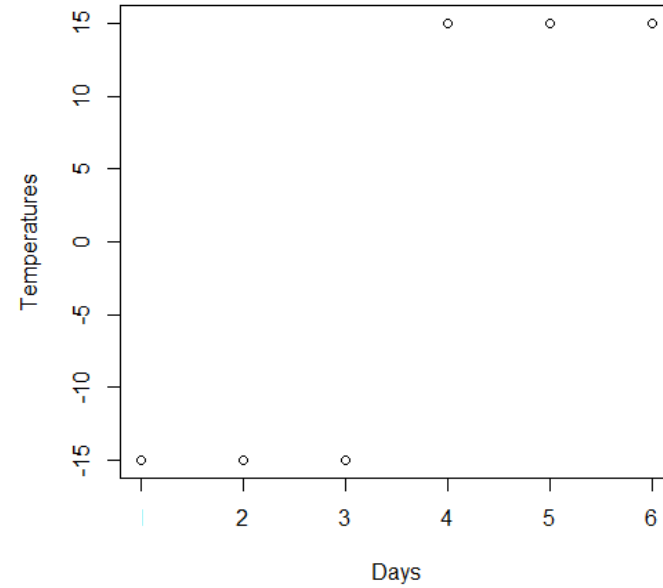**City 3 has variation among the values. Do you think, is it more reliable temperature?**

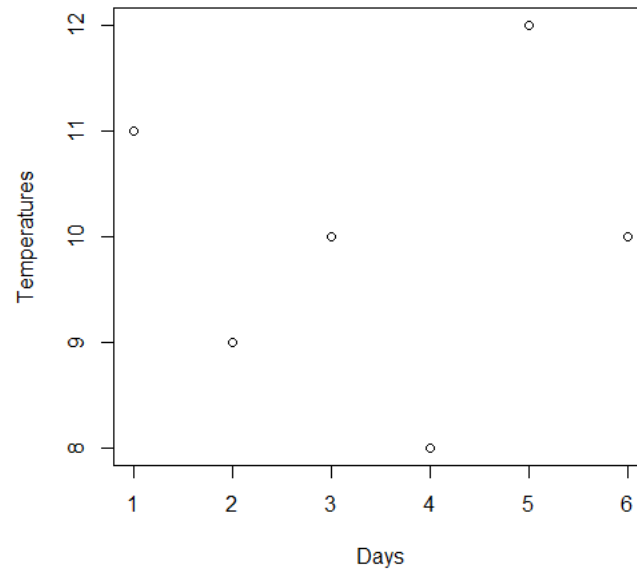**Let us have a graphical look.**

Measures of Variation (or Dispersion)

Measures of Variation (or Dispersion)

Location measures are not enough to describe the behaviour of data.

Concentration or dispersion of observations around any particular value is another property to characterize the data.

How to capture this variation?

Various statistical measures of variation or dispersion are available.
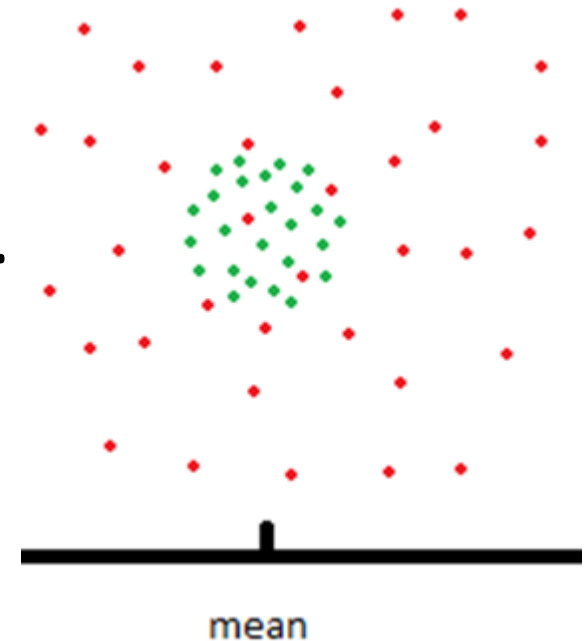
Measures of Variation (or Dispersion)

Two data sets – Green and red dots,

Same mean of red and green colour data points.

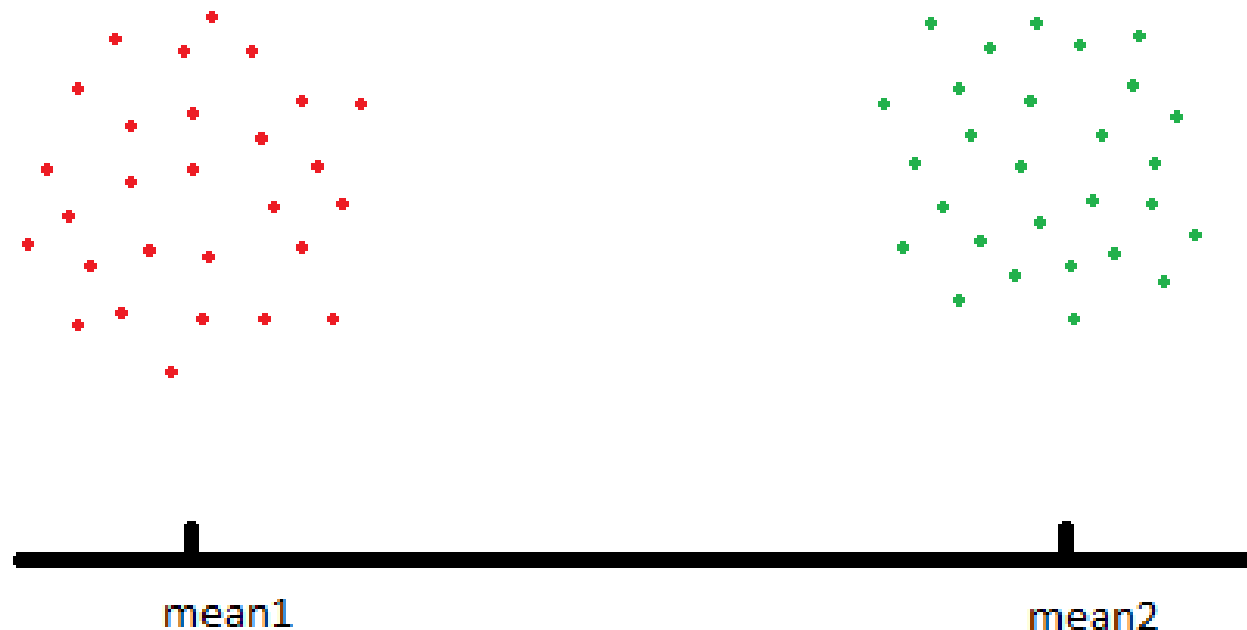**Whose variation is more?**

Which data is more dispersed?

Which data is more concentrated around mean?

mean

Measures of Variation (or Dispersion)

Two data sets – Green and red dots.

Same variability but different means.

Spread and scatterdness of data can be measured around any point but the mean value is more preffered.