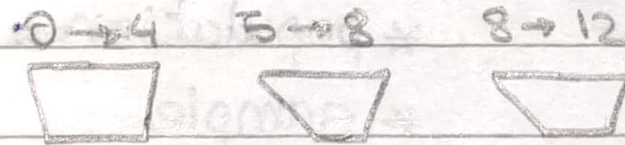Units of spread   how far points from each obter

range — interquartile range — variance — Standard Deviatation

Histograms   → Box Plot

you use binning to organize   15/5/8/8/15/22

0→4   5→8   8→12   7/6/13/13/2/12/16

5-number summary   to calculate range & IQRange   2 3 3 3 4 6

max - min - 1st quartile - 2nd " - 3rd "   ⑦ 8 9 12 15 15

Q_1 \ Q_2(median)/ Q_3 22

medians of quartiles

IQRange = $Q_3 - Q_1$ Quora

University of Texas

Standard Deviation   far each point from mean

① find #mean $\bar{x}$

② calculate $x_i - x$ $\frac{1}{n} \sum\limits_{i=1}^{n} (x_i - x)^2$ variance

③ " $(x_1 - x)^2$

$\sqrt{\text{variance}} =$ standard deviation

high standard deviation comes with high risk   single no. that represents

money / economy   spread of data
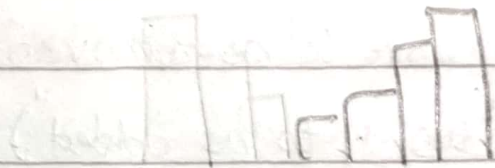
Histograms   Left skewed   symmetric   right skewed   Shape
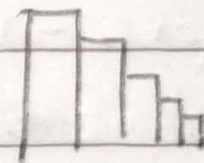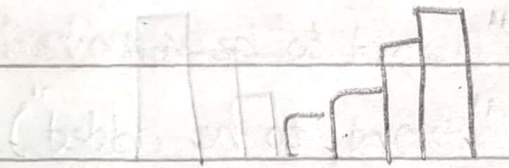
mean =

median =

mode

Histograms

Left skewed = symmetric   right skewed   **Shape**
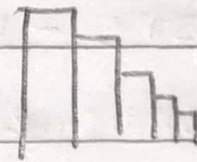
mean =

median =

mode

Outliers   far point from set   use 5-no. summary *dangerous*

① Plot Data

② Note Outliers   Kolmogorov-Smirnov

③ Bell Shaped → mean & SDev   Normal Quantile

④ Skewed → 5-no. summary

Inferential Statistics

drawing conclusions about a population based on data collected from a sample of individuals from that population.

Descriptive

describe data that is available

* population
* sample
* statistic
* parameter

Normal distribution

Sometimes a dataset exhibits a particular shape that is evenly distributed around the mean. Such a distribution is called a normal distribution. It can also be called a Gaussian distribution or a bell curve. Although exam grades are not always distributed in this way, the phrase "grading on a curve" comes from the practice of assigning grades based on a normally distributed bell curve. Figure 3 shov

distributed bell curve. Figure 3 shov

Characteristics of the normal distribution

The mathematical equation for the normal distribution may seem daunting, but the distribution is defined by only two parameters: the mean (μ) and the standard deviation (σ).

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

The mathematical form of the normal distribution

The mean is the center of the distribution. Because the normal distribution is symmetrical about the mean, the median and mean have the same value in an ideal dataset. The standard deviation provides a measure of variability, or spread, within a dataset. For a normal distribution, t