

# Descriptive Statistics

## Q1: Understanding Central Tendency (Easy)

A bakery tracks the daily sales of muffins (in dozens) over a week: [10, 12, 11, 15, 14, 13, 12]. What is the most representative value of their weekly sales, and why?

- The mean is the most representative value for this data set because the sales data is balanced and has no extreme outliers.
- $\text{Mean} = (10+12+11+15+14+13+12) / 7$
- $\text{Mean} = 87/7$
- $\text{Mean} = 12.42$

## Q2: Mean in Real Life (Easy)

A teacher records the marks of her students in a short quiz: [12, 15, 14, 16, 18, 20, 19]. What is the mean score, and what does it tell us about the class's performance?

- $\text{Mean} = (12+15+14+16+18+20+19) / 7$
- $\text{Mean} = (114) / 7$
- $\text{Mean} = 16.28$
- This Mean score provides the average performance of the class which is approx 16.28 out of 20. This suggest that the class performed well.

## Q3: Mode in Real Life (Easy)

A store records the shoe sizes sold in one day: [7, 8, 9, 8, 8, 10, 7, 9]. What is the mode, and why is this information useful for the store manager?

- Mode is 8 as it is the most frequently occurring number in data. This information is useful to the store manager as it gives an idea of the most popular shoe size sold and it would help store managers to make decisions about inventory management, ensuring that they stock more of this size.

## Q4: Median in Real Life (Medium)

A car dealer notes the prices of used cars: [\$8,000, \$9,500, \$10,200, \$11,000, \$50,000]. Why is the median a better measure than the mean in this case? Calculate the median

- Median = \$8,000, \$9,500, \$10,200, \$11,000, \$50,000
- Median = \$10,200
- (Since there are 5 data points in odd number hence middle value is median)
- The median is a better measure than the mean because the mean is affected by the extreme outlier (\$50,000) which makes median a more representative measure of central tendency.

## Q5: Dispersion Introduction (Medium)

A student times how long it takes to finish a puzzle each day: [25, 30, 27, 35, 40]. What does the range tell us about the variation in the student's puzzle-solving time?

## Descriptive Statistics

- Range = Maximum value - Minimum Value
- Range = 40 - 25
- Range = 15 minutes
- This range tells us the total spread in the student puzzle solving times is 15 minutes. This explains the difference between the longest and shortest times taken to complete the puzzle.

### Q6: Range in Action (Medium)

A farmer records the weekly weight of harvested apples (kg): [100, 105, 98, 110, 120]. Find the range. How can this help the farmer in planning his packaging?

- Range = Maximum Value - Minimum Value
- Range = 120 - 98
- Range = 22
- This range can help farmer to determine differences in weekly weight of harvested apples and this can help farmer to plan better packaging of material inventory and logistics to accommodate largest and smallest weekly yields.

### Q7: Variance for Decision-Making (Medium)

Two delivery companies track delivery delays (in minutes). Company A: variance = 6  
Company B: variance = 15 Which company is more consistent, and why?

- Company A is more consistent. Variance measures how spread out a set of data points are from their mean. A lower variance means that the data are clustered more closely around the mean meaning less variability and more consistency, Company A has a variance of 6 which is lower than company B's variance of 15 which means that company A is more consistent than B.

### Q8: Standard Deviation in Context (Hard)

A finance student compares the daily price fluctuations of two cryptocurrencies.

- Coin A: standard deviation = \$30
- Coin B: standard deviation = \$120

Which coin is riskier to invest in, and why?

- Standard deviation measures the variability of a set of values around the mean. A higher standard deviation of price fluctuations indicates greater volatility and higher risk. A lower standard deviation indicates lower volatility and less risk.
- Standard deviation of coin B is \$120, which is higher than the standard deviation of coin A which is \$30. That indicates that coin B is more riskier to invest as it has a higher standard deviation which indicates greater price volatility and higher risk.

# Descriptive Statistics

## Q9: Combining Measures (Hard)

A family records their monthly electricity usage (in kWh): [400, 420, 390, 450, 410].

Find the mean and standard deviation. What do these values together tell you about the family's energy use pattern?

- Mean = Sum of all the numbers / Total numbers
  - Mean =  $[400+420+390+450+410] / 5$
  - Mean =  $2070 / 5$
  - Mean = 414
  
  - Standard deviation =
  - $400 - 414 = -14$
  - $420 - 414 = 6$
  - $390 - 414 = -24$
  - $450 - 414 = 36$
  - $410 - 414 = -4$
  
  - $-14^2 = 196$
  - $6^2 = 36$
  - $-24^2 = 576$
  - $36^2 = 1296$
  - $-4^2 = 16$
  - Sum of squares:  $196+36+576+1296+16 = 2120$
  
  - $= 2120/5-1$
  - $=2120/4$
  - $=530$
  - 23.02
  - Sample standard deviation = 23.02 KWH
- The mean of 414KWH represents the family average monthly electricity usage and the standard deviation (23.02kWh) measures how much their usage varies from that average.

## Q10: Practical Application (Hard)

A basketball player's points in 8 games are recorded: [15, 18, 20, 22, 25, 17, 19, 21]. Find the mean, median, mode, range, and standard deviation. What insights can these measures provide about the player's scoring performance?

- 15, 18, 20, 22, 25, 17, 19, 21
  
- Mean = Sum of all the numbers / Total numbers
- $= 15+18+20+22+25+17+19+21 / 8$
- $= 157 / 8$
- Mean = 19.625

## Descriptive Statistics

- Median - 15, 17, 18, 19, 20, 21, 22, 25
- As it is in even number,
- $19+20 / 2$
- $39/2$
- Median = 19.5
  
- Mode = most frequently occurring no. in data
- Mode = Provided data set has no mode
  
- Now, lets find range
- Range = max value - min value
- $= 25-15$
- Range = 10
  
- Standard Deviation, 15, 17, 18, 19, 20, 21, 22, 25
  
- $15 - 19.625 = -4.625$
- $17 - 19.625 = -2.625$
- $18 - 19.625 = -1.625$
- $19 - 19.625 = -0.625$
- $20 - 19.625 = 0.375$
- $21 - 19.625 = 1.375$
- $22 - 19.625 = 2.375$
- $25 - 19.625 = 5.375$
  
- $(-4.625)^2 = 21.390625$
- $(-2.625)^2 = 6.890625$
- $(-1.625)^2 = 2.640625$
- $(-0.625)^2 = 0.390625$
- $(0.375)^2 = 0.140625$
- $(1.375)^2 = 1.890625$
- $(2.375)^2 = 5.640625$
- $(5.375)^2 = 28.890625$
  
- ---
- $67.875$
- $=67.875 / 8-1$
- $=67.875 / 7$
- Variance = 9.69642857
- Standard deviation = square root of 9.69642857
- Standard deviation = 3.11390889