# 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?**

**Answer:** The **most representative value** is the **mean (average)** daily sales.

**Calculation:**

(10+12+11+15+14+13+12)/7

=787/7

=12.4 dozens

**Why:**

- The data has **no extreme outliers**, so the mean gives a good overall summary of typical sales.
- It uses **all the data points**, reflecting the general level of sales across the week.

**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?**

**ANSWER:** Mean score calculation:

= (12 + 15 + 14 + 16 + 18 + 20 + 19)/7

= 114/7

= 16.3

The mean score is about 16.3 marks.

What it tells us:

- On average, students scored a little over 16 marks on the quiz.
- This suggests the class performed fairly well overall, with most students scoring in the mid-to-high range rather than very low scores.

# Descriptive Statistics

**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?**

**ANSWER:**

**Mode calculation:**
Count each shoe size:

- 7 → 2 times
- 8 → 3 times
- 9 → 2 times
- 10 → 1 time

**Ans:**
The **mode is shoe size 8**, because it appears most frequently.

**Why this is useful:**

- It tells the store manager which shoe size is **most popular**.
- The manager can **stock more size 8 shoes** to meet customer demand and avoid running out of the most requested 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.**

**Answer:** First, list the prices in order (they already are):

$8,000, $9,500, $10,200, $11,000, $50,000

Median calculation:
There are 5 prices, so the median is the middle value:
Median = $10,200

Why the median is better than the mean:

- The $50,000 car is an extreme value (outlier) compared to the other prices.
- The mean would be pulled upward by this very expensive car and would not reflect the typical used car price.
- The median is not affected by outliers, so it gives a more accurate picture of what a "typical" used car costs.

Ans:
The median price is $10,200, and it is a better measure because it represents the typical car price without being distorted by the unusually expensive car.

# Descriptive Statistics

**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?**

**ANSWER: Range calculation:**

- Fastest time = 25 minutes
- Slowest time = 40 minutes

Range = 40 - 25 = 15 minutes

**What the range tells us:**

- The student's puzzle-solving times vary by **15 minutes** from the fastest to the slowest day.
- This indicates a **moderate amount of variation** in performance—some days the student finishes much faster than others.

**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?**

**ANSWER: Range calculation:**

- Minimum weight = 98 kg
- Maximum weight = 120 kg

Range= 120 - 98 = 22 kg

**Ans:** The **range is 22 kg**.

**How this helps the farmer:**

- It shows how much the harvest weight can **vary from week to week**.
- Knowing this variation helps the farmer **plan packaging and storage**, ensuring there are enough boxes and space even in weeks with higher yields.

# Descriptive Statistics

**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?**

**ANSWER:** Company A is more consistent.

**Why:**

- Variance measures how spread out the data is from the average.
- A lower variance means delivery delays are more tightly clustered around the mean.
- Company A’s variance (6) is much lower than Company B’s variance (15), so Company A has more predictable and consistent delivery times.

**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?**

**ANSWER:** Coin B is riskier.

**WHY:**

- Its price changes more each day (higher standard deviation).
- Bigger fluctuations mean more uncertainty for investors.
- Standard deviation measures how much the prices vary from the average.
- Coin A: SD = $30 → smaller daily fluctuations → more stable.
- Coin B: SD = $120 → larger daily fluctuations → less stable.

**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?**

**Answer: Calculate the mean**

Mean= (400+420+390+450+410) / 5

=2070/5

=414 kWh

# Descriptive Statistics

**Calculate the standard deviation**

1. Find deviations from the mean and square them:

- $(400 - 414)^2 = (-14)^2 = 196$
- $(420 - 414)^2 = 6^2 = 36$
- $(390 - 414)^2 = (-24)^2 = 576$
- $(450 - 414)^2 = 36^2 = 1296$
- $(410 - 414)^2 = (-4)^2 = 16$

**Find the average of squared deviations:**

Variance=(196+36+576+1296+16)/5

=2120/5

=424

**Standard deviation = √Variance**

SD =√424

= 20.6 kWh

**Interpretation**

- **Mean = 414 kWh** → On average, the family uses 414 kWh per month.
- **SD = 20.6 kWh** → Their monthly usage **varies moderately** around the average.
- **Conclusion:** The family's energy use is **fairly consistent**, with only small fluctuations month to month.

**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?**

**ANSWER:** MeaN

Mean={15+18+20+22+25+17+19+21)8=157/8=19.625 Mean = 19.6 points

## Median

- Arrange in order: [15, 17, 18, 19, 20, 21, 22, 25]
- 8 numbers → median = average of 4th and 5th values: (19 + 20)/2 = 19.5
  Median = 19.5 points

# Descriptive Statistics

## Mode

- All values occur **once**, so there is **no mode**.

## Range

- Range=25−15=10 points

## Standard Deviation

1.Find deviations from the mean (=19.625) and square them:

- $(15-19.625)^2 = 21.39$
- $(17-19.625)^2 = 6.89$
- $(18-19.625)^2 = 2.64$
- $(19-19.625)^2 = 0.39$
- $(20-19.625)^2 = 0.14$
- $(21-19.625)^2 = 1.89$
- $(22-19.625)^2 = 5.64$
- $(25-19.625)^2 = 28.89$

2. Sum = 21.39 + 6.89 + 2.64 + 0.39 + 0.14 + 1.89 + 5.64 + 28.89 = 67.87
3. Divide by number of data points (population SD formula) or (n-1 for sample):

Using **sample SD (n-1 = 7)**: 67.87 / 7 = 9.70

4. Take square root: $\sqrt{9.70}$ = **3.11 points**

**SD = 3.1 points**

## Insights

- Mean ≈ 19.6 → Player averages about 20 points per game.
- Median = 19.5 → Typical game score is close to the average.
- Mode: none → No score repeats frequently; scoring is varied.
- Range = 10 → Highest score is 10 points above the lowest; moderate variation.
- SD ≈ 3.1 → Most games are within ~3 points of the average → fairly consistent scorer.