Week 03

Contents

- **Measures of the Center of the Data**
- **Skewness and the Mean, Median, and Mode**
- **Measures of the Spread of the Data, Descriptive Statistics [comparing two data sets using Z].**

Measures of the Center of the Data

The "center" of a data set is also a way of describing location. The two most widely used measures of the "center" of the data are the mean (average) and the median. The median is generally a better measure of the center when there are extreme values or outliers because it is not affected by the precise numerical values of the outliers. The mean is the most common measure of the center.

Application of Measure of center of the data in computer science

Data Analysis and Interpretation

These measures provide a concise way to summarize large datasets, making it easier to identify trends and patterns**.**

Performance Evaluation

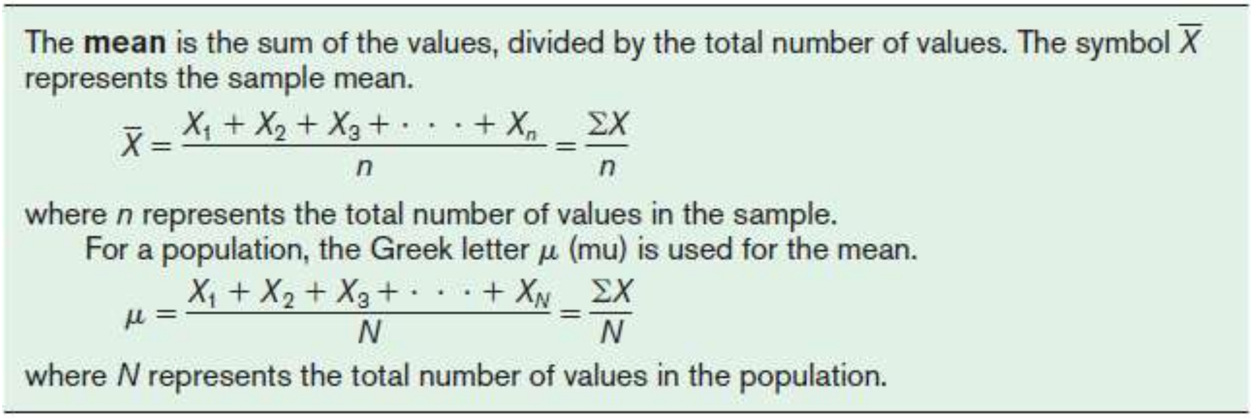
They are essential for evaluating the performance of algorithms, systems, and networks.

Decision Making

Measures of central tendency help in making informed decisions based on data analysis

The Mean

The mean, also known as the arithmetic average, is found by adding the values of the data and dividing by the total number of values.



The **mean** is the sum of the values, divided by the total number of values. The symbol $\overline{X}$ represents the sample mean.

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \cdot \cdot \cdot + X_n}{n} = \frac{\Sigma X}{n}$$

where n represents the total number of values in the sample.

For a population, the Greek letter μ (mu) is used for the mean.

$$\mu = \frac{X_1 + X_2 + X_3 + \cdot \cdot \cdot + X_N}{N} = \frac{\Sigma X}{N}$$

where N represents the total number of values in the population.

Application

Performance Analysis: Calculating the average execution time of algorithms or programs.
Network Analysis: Determining the average network latency or data transfer rates.
Data Analysis: Finding the average user activity on a website or application.

Question

An algorithm is run 5 times, and its execution times (in milliseconds) are: 23, 25, 21, 28, and 23. What is the mean execution time of the algorithm?

Solution:

Sum of execution times: 23 + 25 + 21 + 28 + 23 = 120

Number of runs: 5

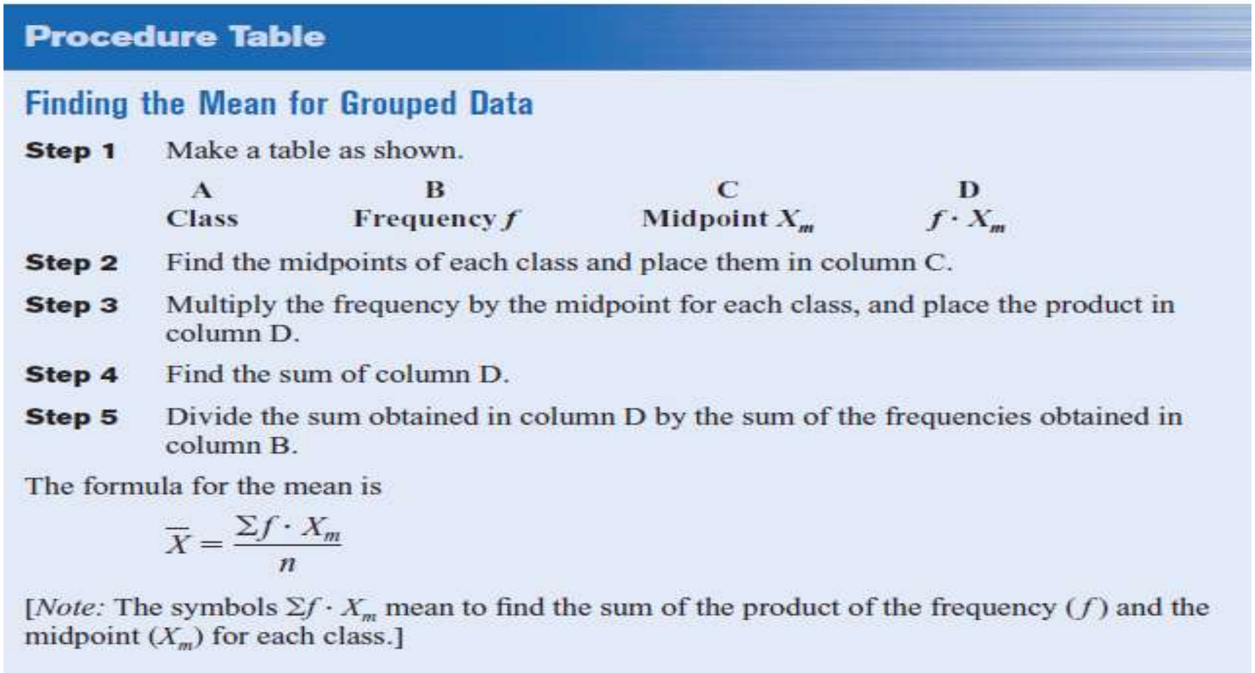
Mean execution time: 120 / 5 = 24 milliseconds.

Practice Question

Q1:The data show the number of patients in a sample of six hospitals who acquired an infection while hospitalized. Find the mean.

110 76 29 38 105 31

Q2:Find the mean for the daily vehicle pass charge for five U.S. National Parks. The costs are $25, $15, $15, $20, and $15.



Procedure Table

Finding the Mean for Grouped Data

Step 1 Make a table as shown.

A	B	C	D
Class	Frequency f	Midpoint X_m	$f \cdot X_m$

Step 2 Find the midpoints of each class and place them in column C.

Step 3 Multiply the frequency by the midpoint for each class, and place the product in column D.

Step 4 Find the sum of column D.

Step 5 Divide the sum obtained in column D by the sum of the frequencies obtained in column B.

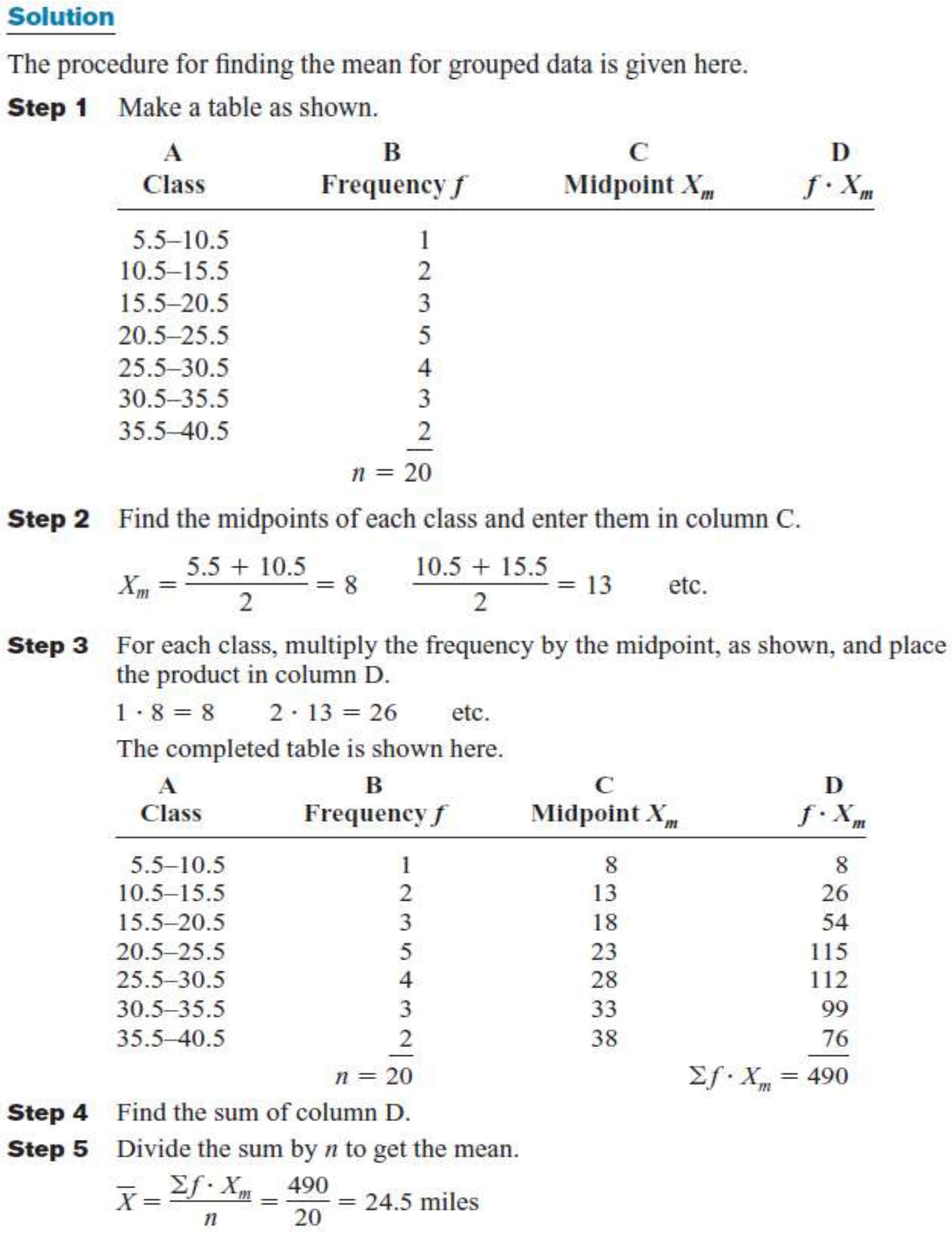
The formula for the mean is

$$\overline{X} = \frac{\Sigma f \cdot X_m}{n}$$

[*Note:* The symbols $\Sigma f \cdot X_m$ mean to find the sum of the product of the frequency (f) and the midpoint (X_m) for each class.]

Question

The data represent the number of miles run during one week for a sample of 20 runners.



Solution

The procedure for finding the mean for grouped data is given here.

Step 1 Make a table as shown.

A Class	B Frequency f	C Midpoint X_m	D $f \cdot X_m$
5.5–10.5	1		
10.5–15.5	2		
15.5–20.5	3		
20.5–25.5	5		
25.5–30.5	4		
30.5–35.5	3		
35.5–40.5	2		
	$n = 20$		

Step 2 Find the midpoints of each class and enter them in column C.

$$X_m = \frac{5.5 + 10.5}{2} = 8 \qquad \frac{10.5 + 15.5}{2} = 13 \qquad \text{etc.}$$

Step 3 For each class, multiply the frequency by the midpoint, as shown, and place the product in column D.

$1 \cdot 8 = 8 \qquad 2 \cdot 13 = 26 \qquad$ etc.

The completed table is shown here.

A Class	B Frequency f	C Midpoint X_m	D $f \cdot X_m$
5.5–10.5	1	8	8
10.5–15.5	2	13	26
15.5–20.5	3	18	54
20.5–25.5	5	23	115
25.5–30.5	4	28	112
30.5–35.5	3	33	99
35.5–40.5	2	38	76
	$n = 20$		$\Sigma f \cdot X_m = 490$

Step 4 Find the sum of column D.

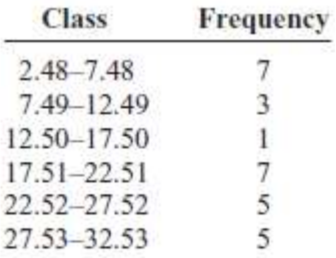
Step 5 Divide the sum by n to get the mean.

$$\overline{X} = \frac{\Sigma f \cdot X_m}{n} = \frac{490}{20} = 24.5 \text{ miles}$$

Practice Question:

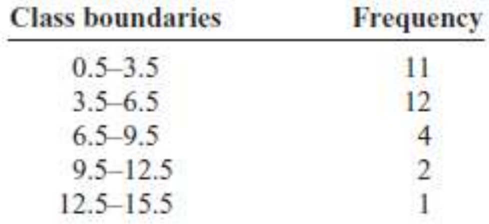
Find the mean

Q1:The hourly compensation costs (in U.S. dollars) for production workers in selected countries are represented below.



Class	Frequency
2.48–7.48	7
7.49–12.49	3
12.50–17.50	1
17.51–22.51	7
22.52–27.52	5
27.53–32.53	5

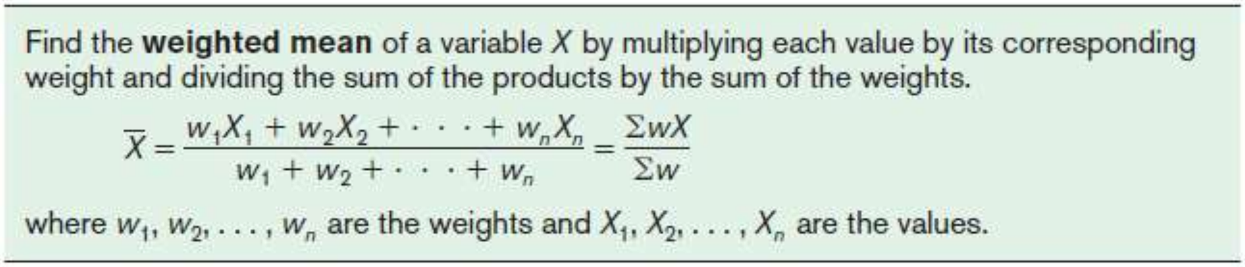
Q2: A random sample of bonuses (in millions of dollars) paid by large companies to their executives is shown:



Class boundaries	Frequency
0.5–3.5	11
3.5–6.5	12
6.5–9.5	4
9.5–12.5	2
12.5–15.5	1

Weighted Mean

Sometimes, you must find the mean of a data set in which not all values are equally represented. The type of mean that considers an additional factor is called the weighted mean, and it is used when the values are not all equally represented.



Find the **weighted mean** of a variable *X* by multiplying each value by its corresponding weight and dividing the sum of the products by the sum of the weights.

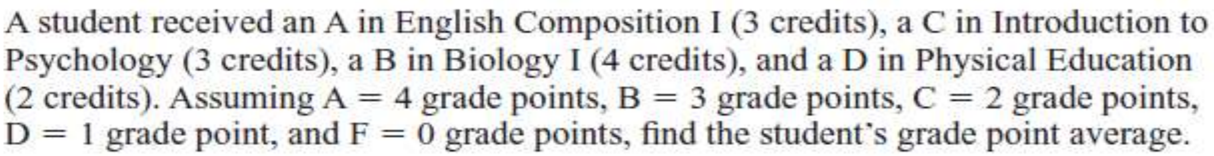
$$\overline{X} = \frac{w_1X_1 + w_2X_2 + \cdot \cdot \cdot + w_nX_n}{w_1 + w_2 + \cdot \cdot \cdot + w_n} = \frac{\Sigma wX}{\Sigma w}$$

where $w_1, w_2, \ldots, w_n$ are the weights and $X_1, X_2, \ldots, X_n$ are the values.

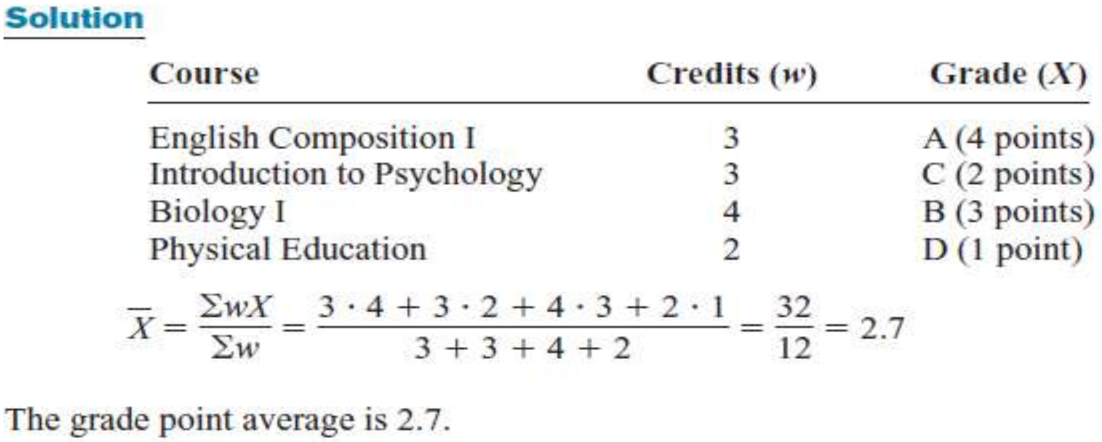
Key Points:

- Weighted means are used when different data points contribute differently to the overall average.
- The weights must sum to 1 (or 100%).
- In computer science, weighted means are useful for combining various performance metrics, calculating grades, and analyzing data with varying levels of importance.

Question



A student received an A in English Composition I (3 credits), a C in Introduction to Psychology (3 credits), a B in Biology I (4 credits), and a D in Physical Education (2 credits). Assuming A = 4 grade points, B = 3 grade points, C = 2 grade points, D = 1 grade point, and F = 0 grade points, find the student's grade point average.



Solution

Course	Credits (w)	Grade (X)
English Composition I	3	A (4 points)
Introduction to Psychology	3	C (2 points)
Biology I	4	B (3 points)
Physical Education	2	D (1 point)

$$\overline{X} = \frac{\Sigma wX}{\Sigma w} = \frac{3 \cdot 4 + 3 \cdot 2 + 4 \cdot 3 + 2 \cdot 1}{3 + 3 + 4 + 2} = \frac{32}{12} = 2.7$$

The grade point average is 2.7.

Practice Question

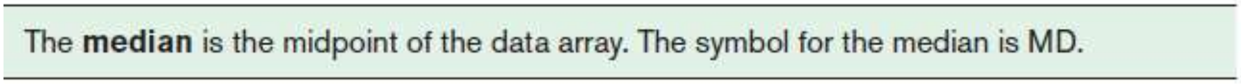
Q1: An instructor grades exams, 20%; term paper, 30%; final exam, 50%. A student had grades of 83, 72, and 90, respectively, for exams, term paper, and final exam. Find the student's final average. Use the weighted mean.

Q2:A grade point average is a weighted average that gives greater weight to courses that earn more credits. Hailey's grade points are 4.0 in Chemistry, which is worth 4 credits, 3.5 in English, which is worth 3 credits, and 3.7 in Physics, which is worth 2 credits. What is Hailey's grade point average?

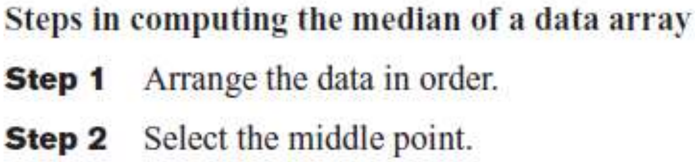
Q3: At a language school, each student is given a score that measures his or her fluency. The fluency score is a weighted average that is determined by rating the student on a scale of 0 to 10 in three categories: grammar, vocabulary, and pronunciation. Grammar counts for 40% of the score, vocabulary counts for 25%, and pronunciation counts for 35%. Thomas gets ratings of 8 for grammar, 6 for vocabulary, and 5 for pronunciation. What is his fluency score?

The Median

The median is the halfway point in a data set. Before you can find this point, the data must be arranged in order. When the data set is ordered, it is called a data array. The median either will be a specific value in the data set or will fall between two values.



The **median** is the midpoint of the data array. The symbol for the median is MD.



Steps in computing the median of a data array

Step 1 Arrange the data in order.

Step 2 Select the middle point.

Application

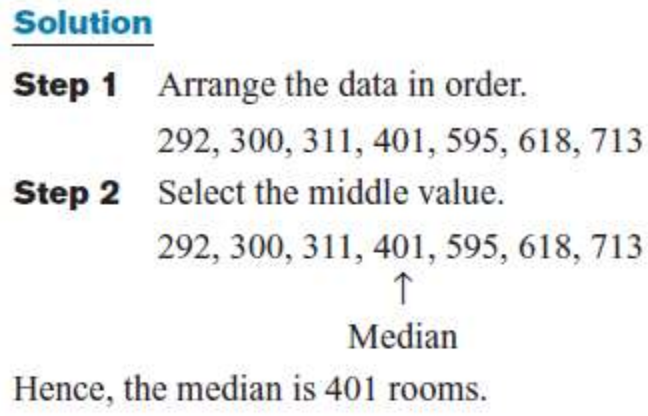
Robust Statistics: Less affected by outliers than the mean, making it useful for analyzing data with extreme values.

Data Analysis: Finding the median response time of a server, which is less skewed by occasional slow responses.

Image Processing: Used in median filtering to reduce noise in images.

Question 1(For Odd number of values)

The number of rooms in the seven hotels in downtown Pittsburgh is 713, 300, 618, 595, 311, 401, and 292. Find the median.



Solution

Step 1 Arrange the data in order.

292, 300, 311, 401, 595, 618, 713

Step 2 Select the middle value.

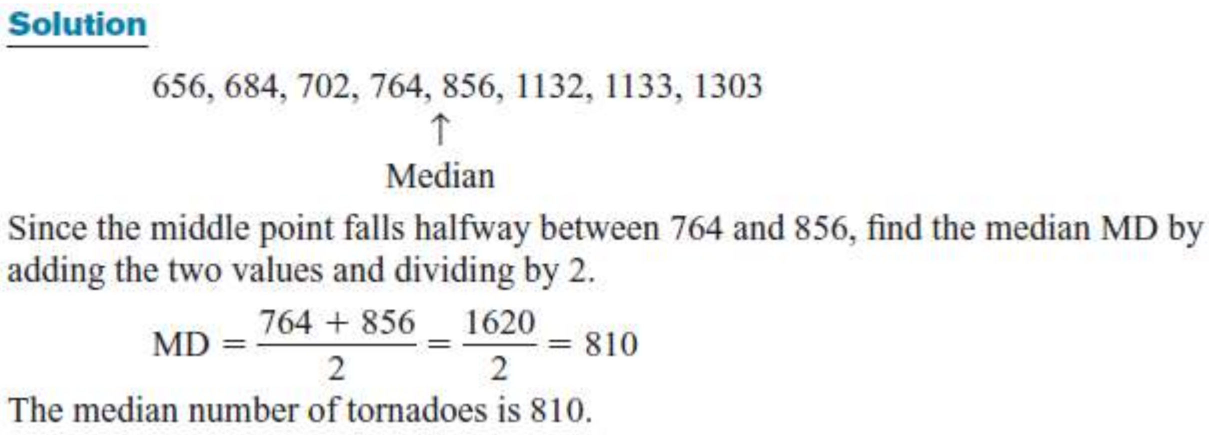
$$\begin{array}{c} 292, 300, 311, \underset{\overset{\uparrow}{\text{Median}}}{401}, 595, 618, 713 \end{array}$$

Hence, the median is 401 rooms.

Question 2 (For Even number of values)

The number of tornadoes that have occurred in the United States over an 8-year period follows. Find the median.

684, 764, 656, 702, 856, 1133, 1132, 1303



Solution

656, 684, 702, 764, 856, 1132, 1133, 1303
↑
Median

Since the middle point falls halfway between 764 and 856, find the median MD by adding the two values and dividing by 2.

$$\text{MD} = \frac{764 + 856}{2} = \frac{1620}{2} = 810$$

The median number of tornadoes is 810.

Practice Question

Q1:The number of tornadoes that have occurred in the United States over an 8-year period follows. Find the median.

684, 764, 656, 702, 856, 1133, 1132, 1303

Q2: The number of children with asthma during a specific year in seven local districts is shown. Find the median.

253, 125, 328, 417, 201, 70, 90

Q3: Six customers purchased these numbers of magazines: 1, 7, 3, 2, 3, 4. Find the median.

The Mode

The third measure of average is called the mode. The mode is the value that occurs most often in the data set. It is sometimes said to be the most typical case.

Uses:

Data Analysis: Identifying the most common user behavior or the most frequent error code.

Network Analysis: Determining the most common network packet size.

Database Management: Finding the most frequent query or data access pattern.

Analyzing categorical data.

Question

A software application logs error codes encountered during execution. The logged error codes are: 404, 500, 404, 200, 404, 503, 500, 404, 200, 404.What is the mode of the error codes?

Solution:

Count the occurrences of each error code:

200,200,404,404,404,404,404,500,500,503

The error code 404 appears most frequently (5 times).

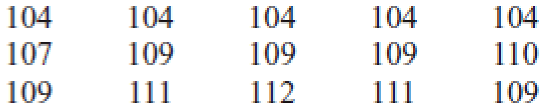
The mode is 404.

Practice Question

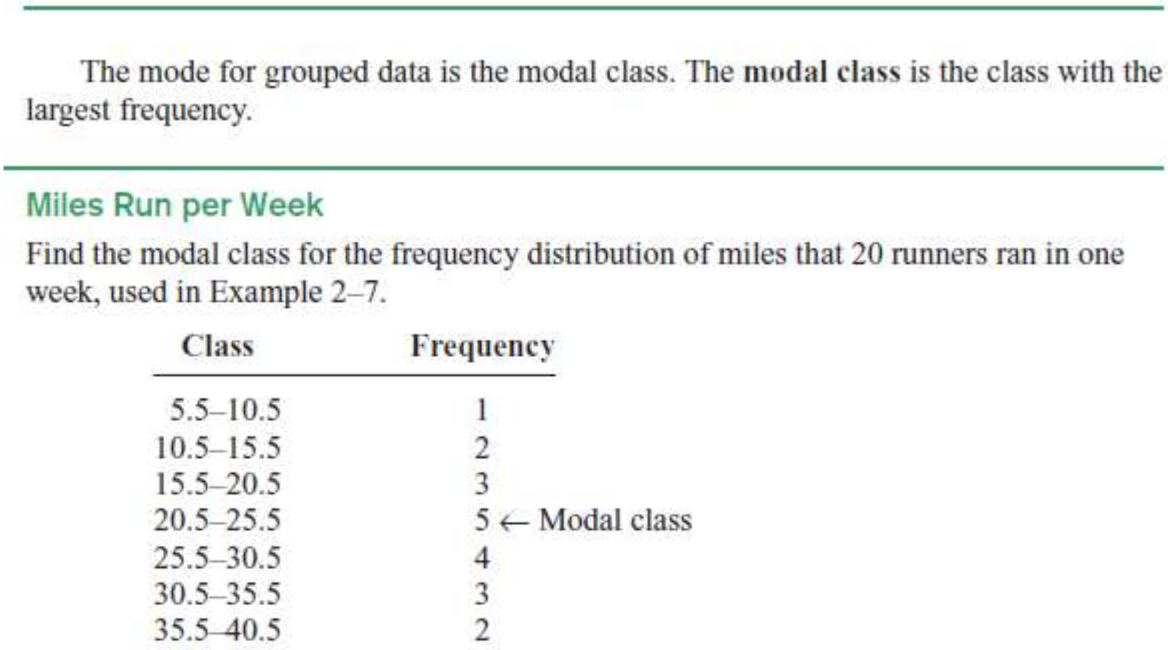
Q1:Find the mode for the number of branches that six banks have.

401, 344, 209, 201, 227, 353

Q2:The data show the number of licensed nuclear reactors in the United States for a recent 15-year period. Find the mode.



104	104	104	104	104
107	109	109	109	110
109	111	112	111	109



The mode for grouped data is the modal class. The **modal class** is the class with the largest frequency.

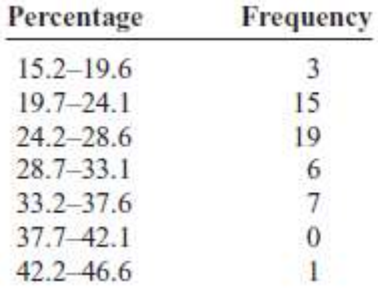
Miles Run per Week

Find the modal class for the frequency distribution of miles that 20 runners ran in one week, used in Example 2–7.

Class	Frequency
5.5–10.5	1
10.5–15.5	2
15.5–20.5	3
20.5–25.5	5 ← Modal class
25.5–30.5	4
30.5–35.5	3
35.5–40.5	2

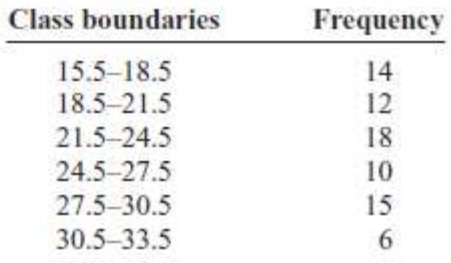
Practice Question

Q1:Below are the percentages of the population over 25 years of age who have completed 4 years of college or more for the 50 states and the District of Columbia. Find the mean and modal class



Percentage	Frequency
15.2–19.6	3
19.7–24.1	15
24.2–28.6	19
28.7–33.1	6
33.2–37.6	7
37.7–42.1	0
42.2–46.6	1

Q2:This frequency distribution represents the data obtained from a sample of 75 copying machine service technicians. The values represent the days between service calls for various copying machines. Find the mean and modal class



Class boundaries	Frequency
15.5–18.5	14
18.5–21.5	12
21.5–24.5	18
24.5–27.5	10
27.5–30.5	15
30.5–33.5	6

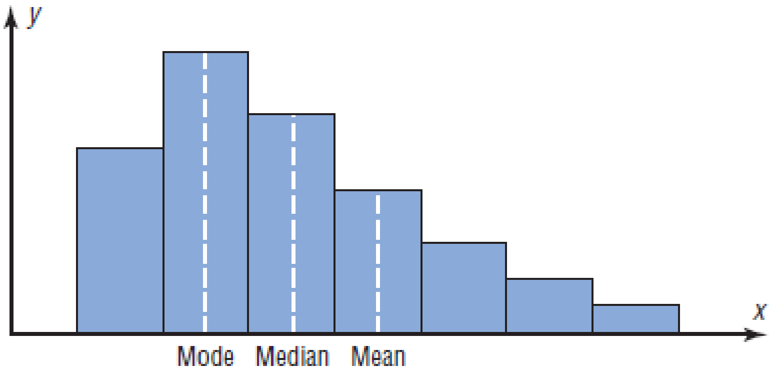
Skewness and the Mean, Median, and Mode

are at the center of the distribution. Examples of symmetric distributions are IQ scores and heights of adult male. Frequency distributions can assume many shapes. The three most important shapes are positively Skewed, symmetric, and negatively skewed.

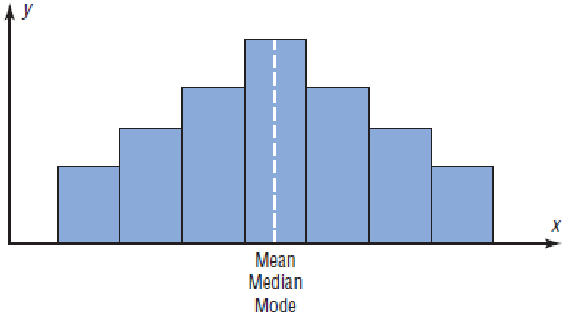
In a **positively skewed or right-skewed distribution**, the majority of the data values fall to the left of the mean and cluster at the lower end of the distribution; the “tail” is to the right. Also, the mean is to the right of the median, and the mode is to the left of the median.

In a **symmetric distribution**, the data values are evenly distributed on both sides of the mean. In addition, when the distribution is unimodal, the mean, median, and mode are the same.

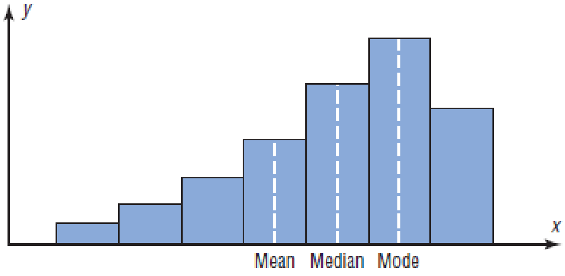
When the majority of the data values fall to the right of the mean and cluster at the upper end of the distribution, with the tail to the left, the distribution is said to be **negatively skewed or left-skewed.**



(a) Positively skewed or right-skewed



(b) Symmetric



(c) Negatively skewed or left-skewed

Measures of Variation

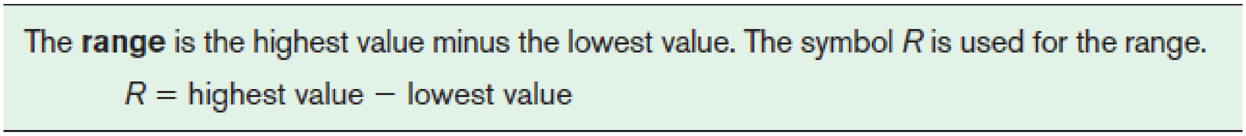
For the spread or variability of a data set, three measures are commonly used: *range, variance,* and *standard deviation which* are fundamental tools in computer science, particularly in data analysis, machine learning, and performance evaluation.

Key Measures and Their Relevance:

- **Standard Deviation:** A widely used measure of the spread of data around the mean. It's valuable for identifying outliers and understanding data distribution.
- **Variance:** The square of the standard deviation, providing a measure of the overall spread of data. It's often used in statistical calculations and model evaluation.
- **Range:** The difference between the maximum and minimum values, providing a simple measure of data spread.

Range

The range is the simplest of the three measures and is defined now.



The **range** is the highest value minus the lowest value. The symbol R is used for the range.

$$R = \text{highest value} - \text{lowest value}$$

Question

you're analyzing the response times (in milliseconds) of a web server to a series of requests. You collect the following data: 120ms, 85ms, 210ms, 150ms, 95ms.

Solution:

To find the range, we need the maximum and minimum values

Maximum value: 210ms

Minimum value: 85ms

Range = Maximum value - Minimum value

Range = 210ms - 85ms = 125ms

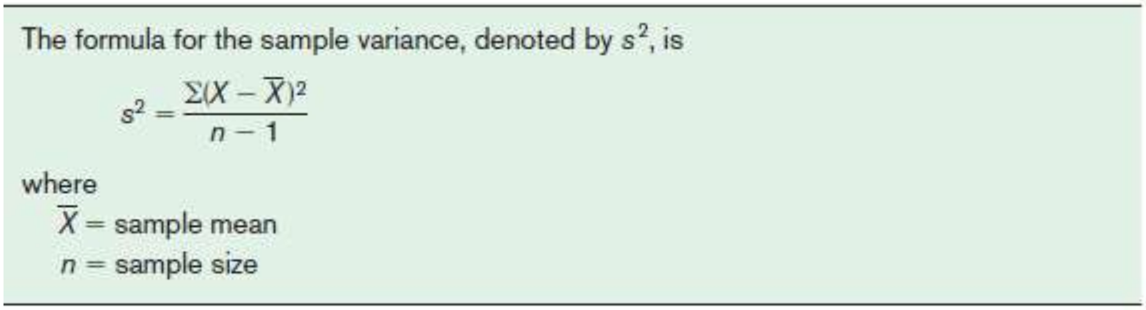
Interpretation

The range of 125ms tells you the spread of the response times. A larger range indicates more variability in the server's performance.

<u>Population Variance and Standard Deviation</u>

Before the variance and standard deviation are defined formally, the computational procedure will be shown, since the definition is derived from the procedure.

Rounding Rule for the Standard Deviation The rounding rule for the standard deviation is the same as that for the mean. The final answer should be rounded to one more decimal place than that of the original data.



The formula for the sample variance, denoted by s^2, is

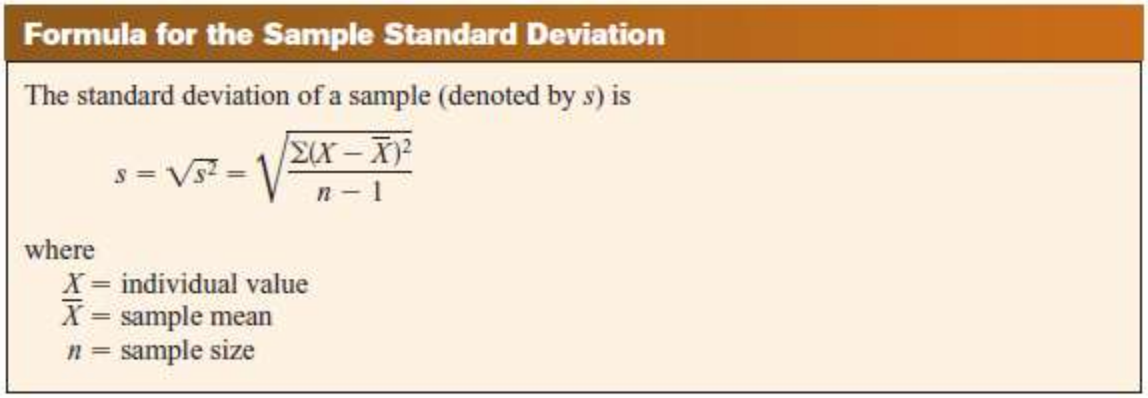
$$s^2 = \frac{\Sigma(X - \overline{X})^2}{n - 1}$$

where

$\overline{X}$ = sample mean

n = sample size

To find the standard deviation of a sample, you must take the square root of the sample variance, which was found by using the preceding formula.



Formula for the Sample Standard Deviation

The standard deviation of a sample (denoted by s) is

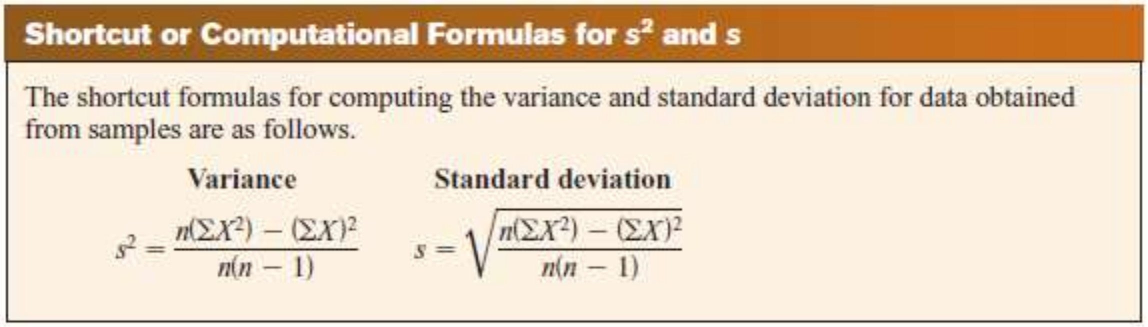
$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma(X - \overline{X})^2}{n - 1}}$$

where

X = individual value

$\overline{X}$ = sample mean

n = sample size



Shortcut or Computational Formulas for s^2 and s

The shortcut formulas for computing the variance and standard deviation for data obtained from samples are as follows.

Variance

$$s^2 = \frac{n(\Sigma X^2) - (\Sigma X)^2}{n(n-1)}$$

Standard deviation

$$s = \sqrt{\frac{n(\Sigma X^2) - (\Sigma X)^2}{n(n-1)}}$$

Exercises:

Find the range, variance, and standard deviation

Q1:Police Calls in Schools The number of incidents in which police were needed for a sample of 10 schools in Allegheny County is 7, 37, 3, 8, 48, 11, 6, 0, 10, 3.

Q2: Cigarette Taxes The increases (in cents) in cigarette taxes for 7 states in a 6-month period are: 60, 20, 40, 40, 45, 12, 34.

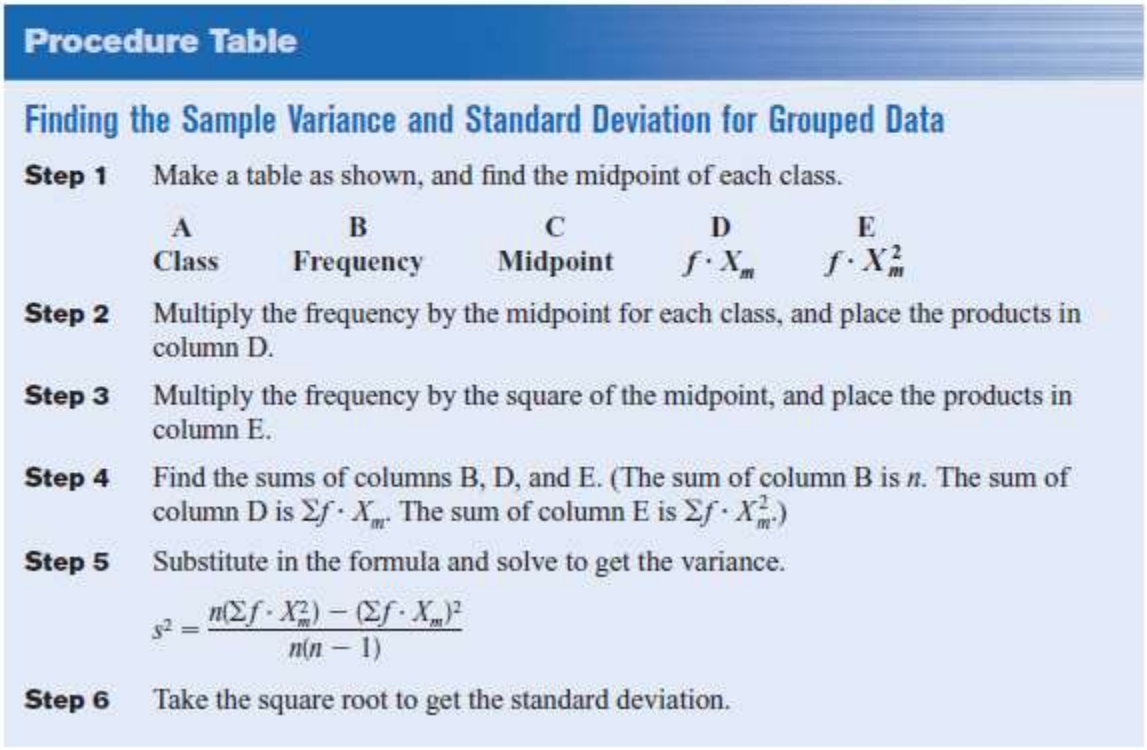
Q3: Stories in the Tallest Buildings The number of stories in the 7 tallest buildings for two different cities is listed below. Which set of data is more variable?

Houston: 75, 71, 64, 56, 53, 55, 47.

Pittsburgh: 64, 54, 40, 32, 46, 44, 42.

Variance and Standard Deviation for Grouped Data

The procedure for finding the variance and standard deviation for grouped data is similar to that for finding the mean for grouped data, and it uses the midpoints of each class.



Procedure Table

Finding the Sample Variance and Standard Deviation for Grouped Data

Step 1 Make a table as shown, and find the midpoint of each class.

A	B	C	D	E
Class	Frequency	Midpoint	$f \cdot X_m$	$f \cdot X_m^2$

Step 2 Multiply the frequency by the midpoint for each class, and place the products in column D.

Step 3 Multiply the frequency by the square of the midpoint, and place the products in column E.

Step 4 Find the sums of columns B, D, and E. (The sum of column B is n. The sum of column D is $\Sigma f \cdot X_m$. The sum of column E is $\Sigma f \cdot X_m^2$.)

Step 5 Substitute in the formula and solve to get the variance.

$$s^2 = \frac{n(\Sigma f \cdot X_m^2) - (\Sigma f \cdot X_m)^2}{n(n-1)}$$

Step 6 Take the square root to get the standard deviation.

Miles Run per Week

Find the variance and the standard deviation for the frequency distribution of the data in Example 2–7. The data represent the number of miles that 20 runners ran during one week.

Class	Frequency	Midpoint
5.5–10.5	1	8
10.5–15.5	2	13
15.5–20.5	3	18
20.5–25.5	5	23
25.5–30.5	4	28
30.5–35.5	3	33
35.5–40.5	2	38

Solution

A	B	C	D	E
Class	Frequency	Midpoint	$f \cdot X_m$	$f \cdot X_m^2$
5.5–10.5	1	8	8	64
10.5–15.5	2	13	26	338
15.5–20.5	3	18	54	972
20.5–25.5	5	23	115	2,645
25.5–30.5	4	28	112	3,136
30.5–35.5	3	33	99	3,267
35.5–40.5	2	38	76	2,888
	$n = 20$		$\Sigma f \cdot X_m = 490$	$\Sigma f \cdot X_m^2 = 13{,}310$

Substitute in the formula and solve for s^2 to get the variance.

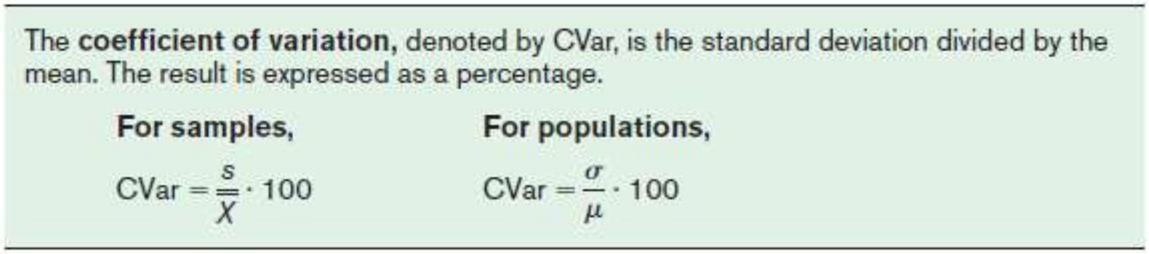
$$\begin{aligned} s^2 &= \frac{n(\Sigma f \cdot X_m^2) - (\Sigma f \cdot X_m)^2}{n(n-1)} \\ &= \frac{20(13{,}310) - 490^2}{20(20-1)} \\ &= \frac{266{,}200 - 240{,}100}{20(19)} \\ &= \frac{26{,}100}{380} \\ &= 68.7 \end{aligned}$$

Take the square root to get the standard deviation.

$$s = \sqrt{68.7} = 8.3$$

Coefficient of Variation

The Coefficient of Variation (CV) is a relative measure of dispersion. It tells us how much variability exists in relation to the mean of the dataset. Unlike standard deviation, which is an absolute measure, the CV allows us to compare variability between datasets with different units or vastly different means.



The **coefficient of variation,** denoted by CVar, is the standard deviation divided by the mean. The result is expressed as a percentage.

For samples,

$$\text{CVar} = \frac{s}{\overline{X}} \cdot 100$$

For populations,

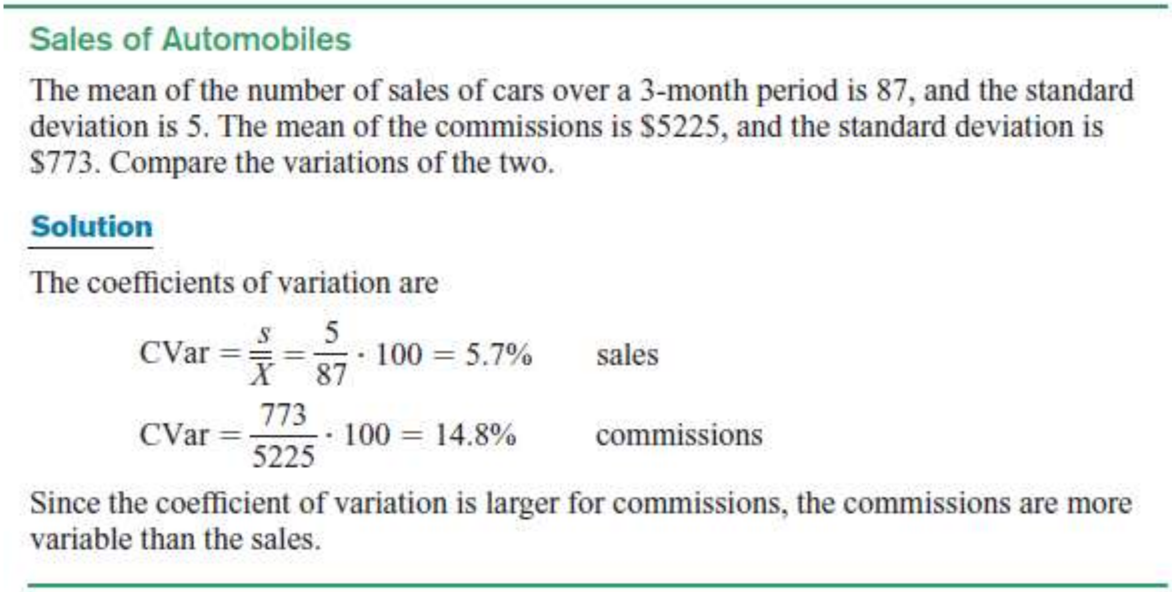
$$\text{CVar} = \frac{\sigma}{\mu} \cdot 100$$

Why is CV Important in Computer Science?

In computer science, especially in fields like:

- **Machine learning**: CV helps compare variability in model accuracy across different datasets.
- **Algorithm analysis**: It can compare execution time variability of different algorithms.
- **Performance testing**: CV measures consistency in metrics like response time, throughput, and memory usage.

For example, if two sorting algorithms have similar average execution times but different variabilities, CV can highlight which one performs more consistently.



Sales of Automobiles

The mean of the number of sales of cars over a 3-month period is 87, and the standard deviation is 5. The mean of the commissions is $5225, and the standard deviation is $773. Compare the variations of the two.

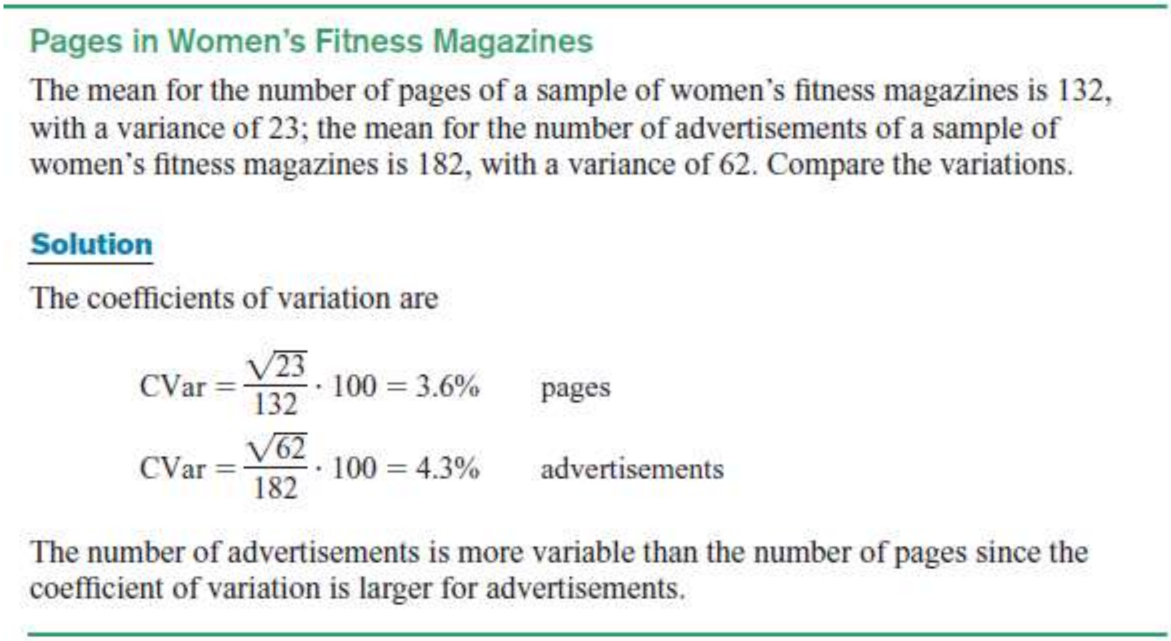
Solution

The coefficients of variation are

$$\text{CVar} = \frac{s}{\overline{X}} = \frac{5}{87} \cdot 100 = 5.7\% \qquad \text{sales}$$

$$\text{CVar} = \frac{773}{5225} \cdot 100 = 14.8\% \qquad \text{commissions}$$

Since the coefficient of variation is larger for commissions, the commissions are more variable than the sales.



Pages in Women's Fitness Magazines

The mean for the number of pages of a sample of women's fitness magazines is 132, with a variance of 23; the mean for the number of advertisements of a sample of women's fitness magazines is 182, with a variance of 62. Compare the variations.

Solution

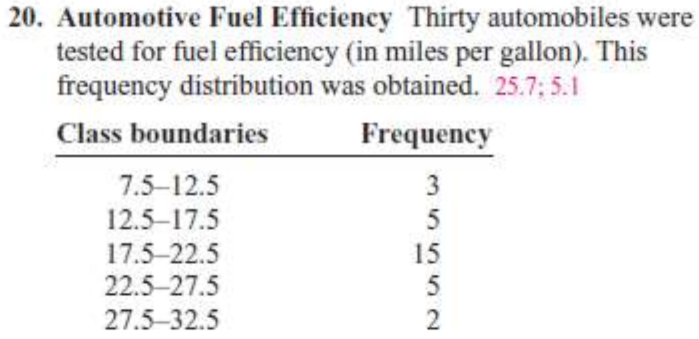
The coefficients of variation are

$$\text{CVar} = \frac{\sqrt{23}}{132} \cdot 100 = 3.6\% \qquad \text{pages}$$

$$\text{CVar} = \frac{\sqrt{62}}{182} \cdot 100 = 4.3\% \qquad \text{advertisements}$$

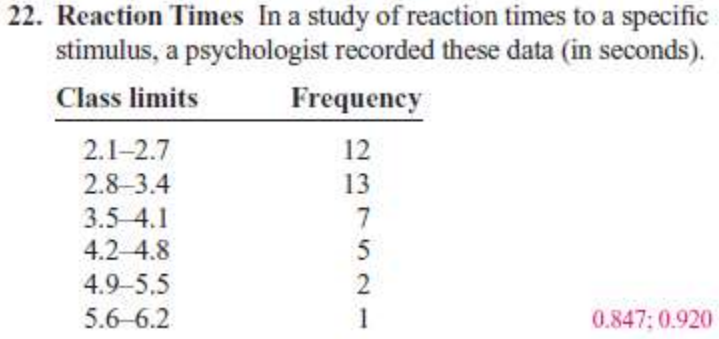
The number of advertisements is more variable than the number of pages since the coefficient of variation is larger for advertisements.

Exercises: Find the variance and standard deviation



20. Automotive Fuel Efficiency Thirty automobiles were tested for fuel efficiency (in miles per gallon). This frequency distribution was obtained. 25.7; 5.1

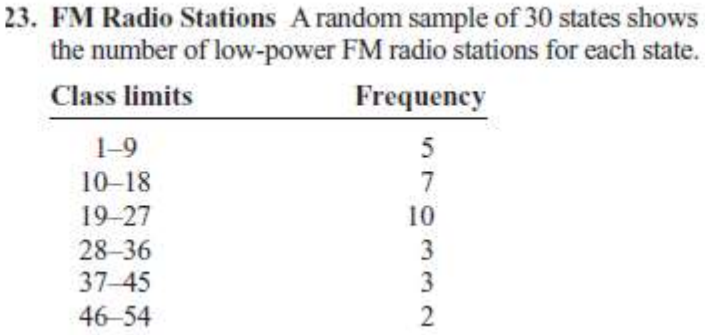
Class boundaries	Frequency
7.5–12.5	3
12.5–17.5	5
17.5–22.5	15
22.5–27.5	5
27.5–32.5	2



22. Reaction Times In a study of reaction times to a specific stimulus, a psychologist recorded these data (in seconds).

Class limits	Frequency
2.1–2.7	12
2.8–3.4	13
3.5–4.1	7
4.2–4.8	5
4.9–5.5	2
5.6–6.2	1

0.847; 0.920



23. FM Radio Stations A random sample of 30 states shows the number of low-power FM radio stations for each state.

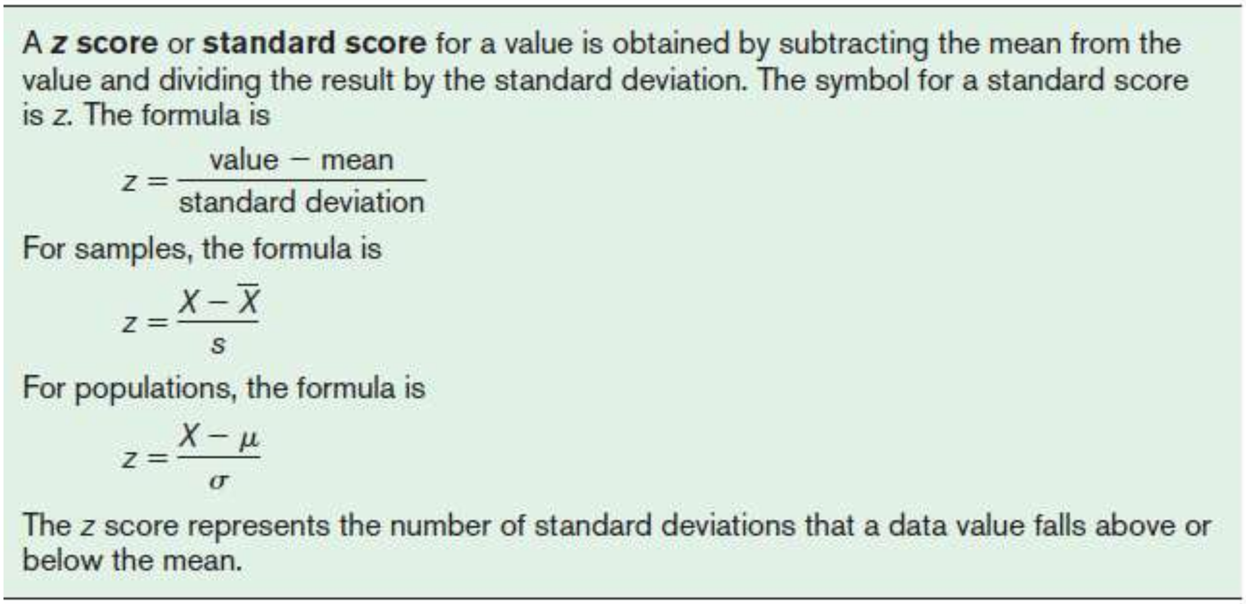
Class limits	Frequency
1–9	5
10–18	7
19–27	10
28–36	3
37–45	3
46–54	2

Measures of Position

In addition to measures of central tendency and measures of variation, there are measures of position or location. These measures include standard scores, percentiles, deciles, and quartiles. They are used to locate the relative position of a data value in the data set.

Standard Scores

A standard score or *z* score tells how many standard deviations a data value is above or below the mean for a specific distribution of values. If a standard score is zero, then the data value is the same as the mean.



A ***z* score** or **standard score** for a value is obtained by subtracting the mean from the value and dividing the result by the standard deviation. The symbol for a standard score is *z*. The formula is

$$z = \frac{\text{value} - \text{mean}}{\text{standard deviation}}$$

For samples, the formula is

$$z = \frac{X - \overline{X}}{s}$$

For populations, the formula is

$$z = \frac{X - \mu}{\sigma}$$

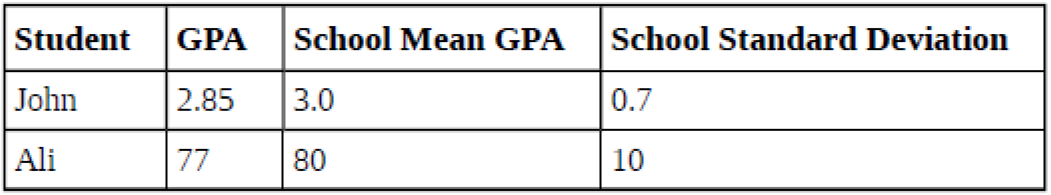
The *z* score represents the number of standard deviations that a data value falls above or below the mean.

Application

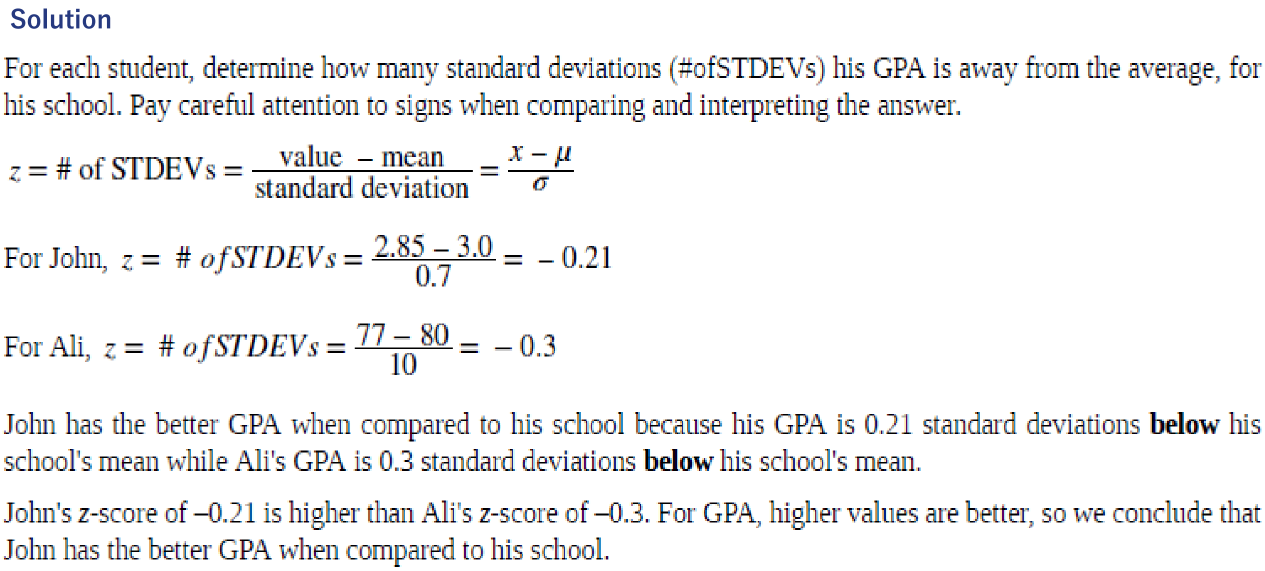
The Z-score is a versatile tool in computer science for data preprocessing, outlier and anomaly detection, and statistical analysis, helping to gain valuable insights from data and build robust systems

Question

Two students, John and Ali, from different high schools, wanted to find out who had the highest GPA when compared to his school. Which student had the highest GPA when compared to his school?



Student	GPA	School Mean GPA	School Standard Deviation
John	2.85	3.0	0.7
Ali	77	80	10



Solution

For each student, determine how many standard deviations (#ofSTDEVs) his GPA is away from the average, for his school. Pay careful attention to signs when comparing and interpreting the answer.

$$z = \text{\# of STDEVs} = \frac{\text{value } - \text{mean}}{\text{standard deviation}} = \frac{x-\mu}{\sigma}$$

For John, $z = \ \# \, of STDEVs = \frac{2.85-3.0}{0.7} = \ -0.21$

For Ali, $z = \ \# \, of STDEVs = \frac{77-80}{10} = \ -0.3$

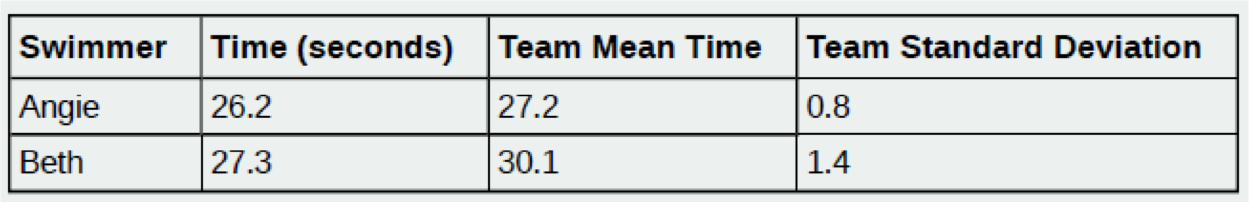
John has the better GPA when compared to his school because his GPA is 0.21 standard deviations **below** his school's mean while Ali's GPA is 0.3 standard deviations **below** his school's mean.

John's z-score of –0.21 is higher than Ali's z-score of –0.3. For GPA, higher values are better, so we conclude that John has the better GPA when compared to his school.

Practice Question

Q1

Two swimmers, Angie and Beth, from different teams, wanted to find out who had the fastest time for the 50 meter freestyle when compared to her team. Which swimmer had the fastest time when compared to her team?



Swimmer	Time (seconds)	Team Mean Time	Team Standard Deviation
Angie	26.2	27.2	0.8
Beth	27.3	30.1	1.4

Q2

A music school has budgeted to purchase three musical instruments. They plan to purchase a piano costing \$3,000, a guitar costing \$550, and a drum set costing \$600. The mean cost for a piano is \$4,000 with a standard deviation of \$2,500.The mean cost for a guitar is \$500 with a standard deviation of \$200. The mean cost for drums is \$700 with a standard deviation of \$100. Which cost is the lowest, when compared to other instruments of the same type? Which cost is the highest when compared to other instruments of the same type. Justify your answer.