24

DESCRIPTIVE STATISTICS

CONTENTS

Descriptive statistics are used to describe or summarize data in a meaningful way. They provide simple summaries about a sample. Descriptive statistics are broken down into measures of central tendency, measures of variability (spread) and measures of association.

MEASURES OF LOCATION

Measures of location describe the central tendency of the data. The three most common measures of location are the mean, the median, and the mode.

Mean

The **mean** of a set of numbers is their average. The mean is a measure of central location for the data. Sample mean is denoted by $\bar{x}$; population mean is denoted by the Greek letter μ.

For a sample with n observations, the sample mean is

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

☑ **EXAMPLE 24.1**

The table shows the mean annual investment in Bitcoins by 5 investors.

Investor	1	2	3	4	5
Annual Investment ($)	377	279	582	993	769

What is the mean annual investment?

SOLUTION tips

The mean annual investment is computed as

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \cdots + x_5}{5} = \frac{377 + 279 + 582 + 993 + 769}{5}$$
$$= \frac{3000}{5} = \$600$$

Median

The **median** is the value in the middle when the data are arranged from the smallest value to the largest value. With an odd number of observations, the median is the middle value. With even number of observations, the median is the average of the values for the middle two observations. Whenever a data set contains extreme values, the median is preferred as a measure of central location.

☑ **EXAMPLE 24.2**

The incomes of a sample of Hollywood actors are:

| $136,105 | $80,821 | $31,296 | $59,075 | $23,456 |
| $147,878 | $125,973 | $58,014 | $99,179 | $124,312 |

What is the median income?

SOLUTION**tips**

First arrange the data in ascending order:
23456, 31296, 58014, 59075, **80821, 99179,** 124312, 125973, 136105, 147878
Because n = 10 is even, the median is the average of the two middle values:
$$\text{Median} = \frac{80821 + 99179}{2} = \frac{180000}{2} = \$90,000$$
So, the median income is $90,000.

Mode

The **mode** is the value that occurs with highest frequency.

☑ **EXAMPLE 24.3**

The weight of a sample of students are:

| | 60 | 69 | 63 | 72 | 81 | 69 |

What is the mode?

SOLUTION**tips**

The only weight that occurs more than once is 69.
Because 69 has the highest frequency, it is the mode.

Percentiles

A **percentile** (or a **centile**) is the value below which a given percentage of observations falls. For example, the 30th percentile is the value (or score) below which 30% of the observations may be found.
To calculate the pth percentile:
- Arrange the data from in ascending order.
- Compute an index

$$i = \left(\frac{P}{100}\right)n$$

where p is the percentile and n is the number of observations.
- If i is not an integer, round up to the nearest integer. This rounded value represents the position of the *pth* percentile.
- If i is an integer, the *pth* percentile is the average of the values at positions i and $i + 1$ in the sorted data.

☑ **EXAMPLE 24.4**

The tuition fees at 10 of the world's top universities are:

| $56,070 | $48,260 | $55,120 | $48,920 | $54,710 |
| $55,200 | $46,990 | $57,870 | $46,630 | $57,720 |

Determine the 30th and 75th percentile for tuition fees.

Purchase the full book at:
     https://unimath.5profz.com/

We donate 0.5% of the book sales every year to charity, forever. When you buy **University Mathematics (I & II)** *you are helping orphans and the less privileged.*