

DESCRIPTIVE STATISTICS

Descriptive Statistics: Summarize and Organize Characteristics of a data. A data set is a collection of responses or observations from a Sample or entire population.

In quantitative research, after collecting data, the first step of statistical analysis is to describe characteristics of the response such as the average of one variable (e.g., age), or the relation between two variables (e.g., age and Creativity).

The Next Step is inferential statistics, which help you decide whether your data conforms or refutes your hypothesis and whether it is generalizable to a large population.

≡ TABLE OF CONTENTS

1. Types of descriptive Statistics
2. Frequency distribution
3. Measures of central tendency
4. Measures of Variability
5. Univariate descriptive Statistics
6. Bivariate descriptive Statistics

TYPES OF DESCRIPTIVE STATISTICS

There are 3 main types of descriptive Statistics

The Distribution Concerns the frequency of each Value

The Central tendency Concerns the Average of the Values

The Variability or dispersion Concerns how Spread out the Values are

Research example
you want to study the popularity of different leisure activities by gender. You distribute a Survey and ask participants how many times they did each of the following in the Past year

- * Go to a library
- * Watch a movie at a theatre
- * Visit a National park

your data set is the collection of responses to the Survey. Now you can use descriptive statistics to find out the overall frequency of each activity (distribution), the average for each activity (central tendency) and the spread (variability)

Gender	Number
Male	280

Frequency distribution

A data set is made up of a distribution of values, or scores. in table or graphs. You can summarize the frequency of every possible value of a variable in numbers or percentages. This is called frequency distribution.

Simple frequency distribution table

Grouped frequency distribution table

For the Variable of gender, you list all possible answers, on the left hand column, you count the number or percentage of responses for each answer and display it on the right hand column.

Gender	Male
Number	182

Gender	Female
Number	235

Gender	Other
Number	27

From this table, you can see that more women than men or people with another identity took part in the study.

Measures of Central Tendencies

* Measures of Central Tendency estimate the center, or average, of a data set. The mean, median and mode are 3 ways of finding the average.

Mean : The Mean M is the most commonly used method for finding the average.

To find the mean, simply add up all responses values and divide the sum by the total number of responses. The total number of responses or observations is called N .

Mean number of library visits

Data Set

Sum of all values : 15, 3, 12, 0, 24, 3

$$\text{Sum} = 15 + 3 + 12 + 0 + 24 + 3 = 57$$

$$N = 6$$

$$\text{Mean} = \text{Sum} / N = \boxed{57 / 6 = 9.5}$$

Median

The Median is the value that's exactly in the middle of a data set.

To find the Median, Order each response

Value from the Smallest to the biggest.

The Median is the number in the middle. If there are two in the middle, find their mean

Ordered data set

0, 3, 3, 12, 15, 24, 27

$$\text{Median} = 12$$

Mode

The mode is simply the most popular or the most frequent response value.

To find the mode order your data set from the lowest to highest and find the response

Ordered Data Set

Mode: 0, 3, 3, 12, 15, 24

$M = 3$ is the most occurred

Measures of Variability

Measures of variability give you a sense of how spread out the response values are. The standard deviation and variance each reflect

Range: The range gives you an idea of how far apart the most extreme responses scores are. To find the range, simply subtract the lower value from the higher value.

Data Set: 0, 3, 3, 12, 15, 24

$$\text{Range} : 24 - 0 = 24$$

Standard Deviation:

The Standard deviation is the average amount of variability in your data set. It tells you on average how far each score lies from the mean.

There are Six Steps for finding the Standard deviation

1. List each Score and find their mean

2. Subtract the mean from each Score to get the Deviations from the mean

3. Square each of these deviations

4. Add up all of the Squared deviations

5. Divide the sum of the Squared deviations by $N - 1$

Raw data : 15

Deviation from : $15 - 9.5 = 9.5$

Squared : $3^2 = 30.25$

R : 3

D : $3 - 9.5 = -6.5$

S : 42.25

R : 0

D : $0 - 9.5 = -9.5$

S : 90.25

R : 24

D : $24 - 9.5 = 14.5$

S : 210.25

R : 3

D : $3 - 9.5 = -6.5$

S : 42.25

Raw of Data : $M = 9.5$

Deviation : Sum = 0

Squared : Sum of Squares = 421.5

$$\text{Step 5: } 421.5 / 5 = 84.3$$

$$6: \sqrt{84.3} = 9.18$$

9.18 average, each Score deviates
from the mean by 9.18 points

Variance

The Variance is the average of Squared
deviations of Spread in the Data Set

To find the Variance, Simple Square the
Standard Deviation. The Symbol of Variance
is s^2 .

Data Set : 15, 3, 12, 6, 24, 3

$$S = 9.18$$

$$s^2 = 84.3$$

Univariate descriptive Stat

Univariate descriptive statistics focus on
Only One Variable at a time

It's important to examine data from each variable separately using multiple measures. Using measures of distribution, central tendency and spread, programs like SPSS and Excel can be used to easily calculate these.

Visited library : 6

Mean : 9.5

Median : 7.5

Mode : ~~84~~ 33

Variance

Standard deviation : 9.18

Variance : 84.3

Range : 24

If you were to only consider the mean as a measure of central tendency, your impression of the middle of the data can be skewed.