

What Are Dichotomous Variables? (Definition & Example)

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *What Are Dichotomous Variables? (Definition & Example)*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138647>

Dichotomous variables are a type of categorical variable that can only take on one of two possible values. These values are mutually exclusive and collectively exhaustive, meaning that they cover all possible options and cannot overlap. For example, a dichotomous variable could be "gender," with the two values being "male" and "female." Another example could be "smoker," with the two values being "yes" and "no." These variables are often used in statistical analysis and research as they allow for clear and distinct categorization of data.

What Are Dichotomous Variables? (Definition & Example)

A dichotomous variable is a type of variable that only takes on two possible values.


Some examples of dichotomous variables include:

**Gender: Male or Female
Coin Flip: Heads or Tails
Property Type: Residential or Commercial
Athlete Status: Professional or Amateur
Exam Results: Pass or Fail**

These types of variables occur all the time in practice. For example, consider the following dataset that contains 10 and 4 variables:

| Athlete | Gender | Division | Average Points | Won Championship |
|---------|--------|----------|----------------|------------------|
| Arnold | Male | 1 | 13.2 | Yes |
| Bert | Male | 1 | 9.8 | No |
| Cara | Female | 2 | 15.6 | Yes |
| Derrick | Male | 2 | 22.7 | No |
| Eleanor | Female | 1 | 19.4 | No |
| Frank | Male | 2 | 7.8 | Yes |
| Greg | Male | 3 | 13.3 | No |
| Harry | Male | 3 | 6.7 | No |
| Isaiah | Male | 2 | 29.8 | No |
| Jenny | Female | 1 | 23.1 | No |

The variables gender and Won Championship are dichotomous because they can each only take on two possible values:

Dichotomous Variables

| Athlete | Gender | Division | Average Points | Won Championship |
|---------|--------|----------|----------------|------------------|
| Arnold | Male | 1 | 13.2 | Yes |
| Bert | Male | 1 | 9.8 | No |
| Cara | Female | 2 | 15.6 | Yes |
| Derrick | Male | 2 | 22.7 | No |
| Eleanor | Female | 1 | 19.4 | No |
| Frank | Male | 2 | 7.8 | Yes |
| Greg | Male | 3 | 13.3 | No |
| Harry | Male | 3 | 6.7 | No |
| Isaiah | Male | 2 | 29.8 | No |
| Jenny | Female | 1 | 23.1 | No |

However, the variables **Division** and **Average Points** are not dichotomous because they can take on multiple values.

Bonus Tip:

You can remember that dichotomous variables can only take on two values by remembering that the prefix "di" is a Greek word that means "two", "twice", or "double."

How to Create Dichotomous Variables

It's worth noting that we can create a dichotomous

variable from a continuous variable by simply separating values based on some threshold.

For example, in the previous dataset we could turn the variable Average Points into a dichotomous variable by classifying players with an average above 15 as "high scorers" and those with an average below 15 as "low scorers":

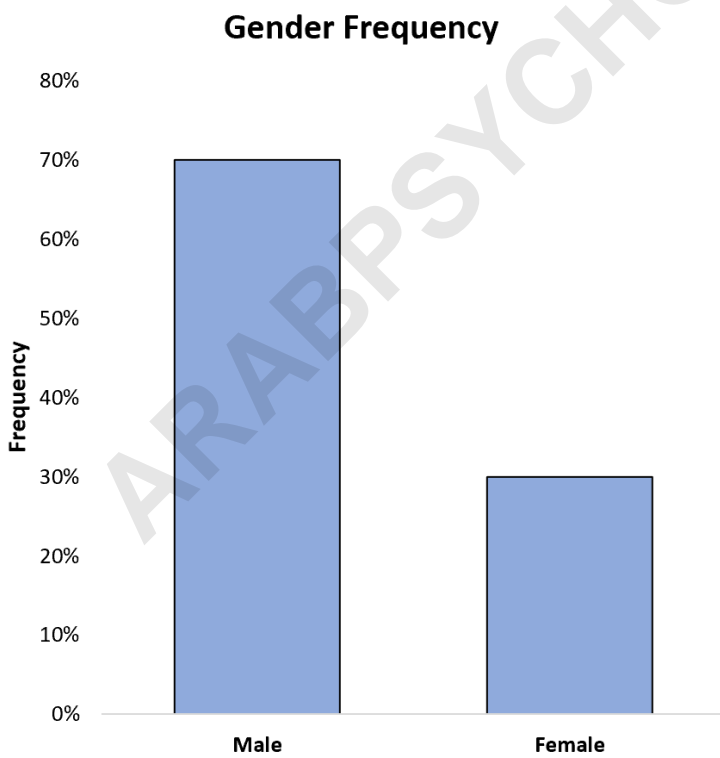
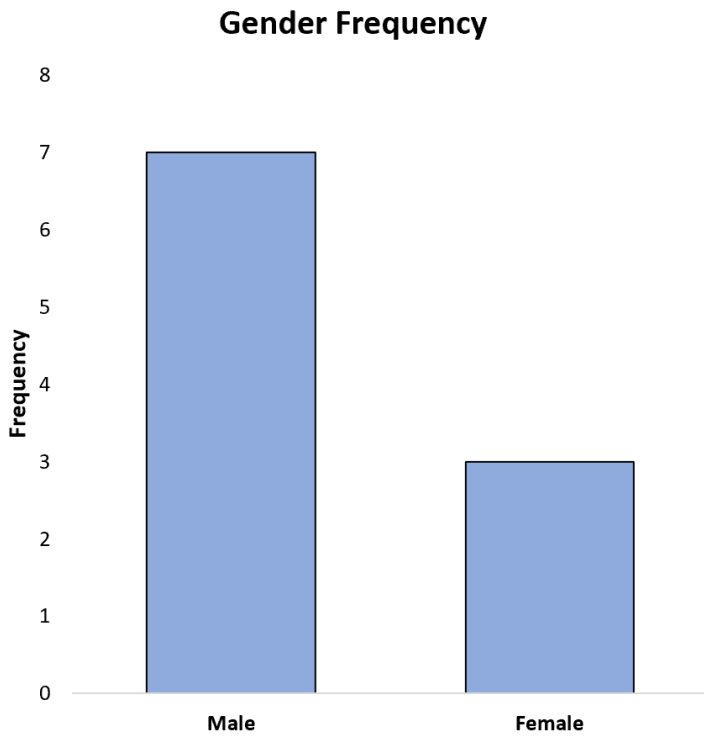
| Athlete | Gender | Division | Average Points | Won Championship |
|---------|--------|----------|----------------|------------------|
| Arnold | Male | 1 | 13.2 | Yes |
| Bert | Male | 1 | 9.8 | No |
| Cara | Female | 2 | 15.6 | Yes |
| Derrick | Male | 2 | 22.7 | No |
| Eleanor | Female | 1 | 19.4 | No |
| Frank | Male | 2 | 7.8 | Yes |
| Greg | Male | 3 | 13.3 | No |
| Harry | Male | 3 | 6.7 | No |
| Isaiah | Male | 2 | 29.8 | No |
| Jenny | Female | 1 | 23.1 | No |

| Athlete | Gender | Division | Type of Scorer | Won Championship |
|---------|--------|----------|----------------|------------------|
| Arnold | Male | 1 | Low | Yes |
| Bert | Male | 1 | Low | No |
| Cara | Female | 2 | High | Yes |
| Derrick | Male | 2 | High | No |
| Eleanor | Female | 1 | High | No |
| Frank | Male | 2 | Low | Yes |
| Greg | Male | 3 | Low | No |
| Harry | Male | 3 | Low | No |
| Isaiah | Male | 2 | High | No |
| Jenny | Female | 1 | High | No |

How to Visualize Dichotomous Variables

We typically visualize dichotomous variables by using a simple bar chart to represent the frequencies of each value it can take on.

For example, the following bar chart shows the frequencies of each gender in the previous dataset:



This allows us to easily see that 70% of the total

athletes in the dataset are male and 30% are female.

How to Analyze Dichotomous Variables

There are several ways to analyze dichotomous variables. Two of the most common ways include:

1. One proportion z-test

A determines whether or not some observed proportion is equal to a theoretical one.

For example, we might use this test to determine if the true proportion of athletes who are male in some population is equal to 50%.

2. Point-biserial correlation

is used to measure the relationship between a dichotomous variable and a continuous variable.

This type of correlation takes on a value between -1 and 1 where:

**-1 indicates a perfectly negative correlation between two variables
0 indicates no correlation between two variables
1 indicates a perfectly positive correlation**

between two variables

For example, we might calculate the point-biserial correlation between gender and average points per game to understand how strongly these two variables are related.

ARABPSYCHOLOGY.COM