

Definitions

Data = *Information*

Numbers, Text, Images, Videos, Signals, etc..

Variable - *characteristic being measured*

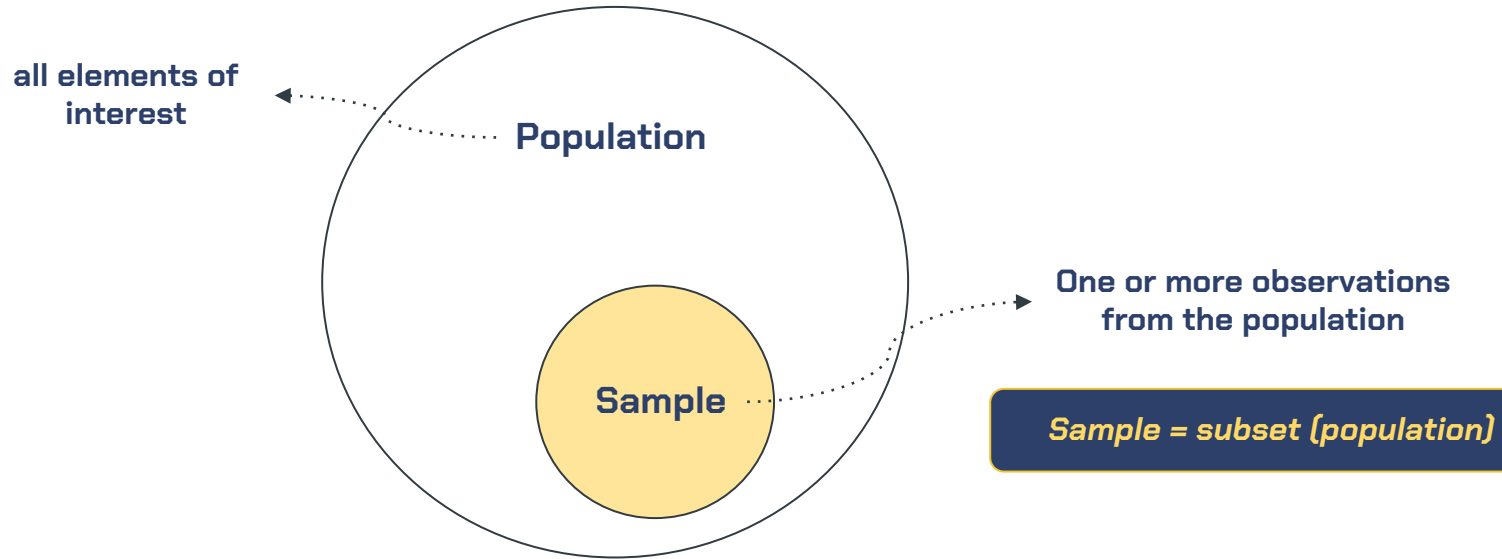
▶ **Height** - 6 ft

▶ **“varies”** in a population

Observation - *set of values related to a variable*



Population and sample



Sampling

“Random sampling” is the method to collect a sample representative of the population.

The Literary Digest
NEW YORK OCTOBER 31, 1936

Topics of the day

LONDON, 1,293,669; ROOSEVELT, 972,897
Final Returns in The Digest's Poll of Ten Million Voters

Well, the great battle of the ballots in the poll of ten million voters, scattered throughout the forty-eight States of the Union, is now finished, and in the table below we record the figures received up to the hour of going to press.

These figures are exactly as received from more than one in every five voters polled in our country—they are neither weighted, adjusted nor interpreted.

Never before in an experience covering more than a quarter of a century in taking polls have we received so many different varieties of criticism—praise from many; condemnation from many others—and yet it has been just of the same type that has come to us every time a Poll has been taken in all these years.

A telegram from a newspaper in California asks: “Is it true that Mr. Beard has purchased The Literary Digest?” A telephone message only the day before these lines were written: “Has the Repub-

lican National Committee purchased The Literary Digest?” And all types and varieties, including: “Have the Jews purchased The Literary Digest?” “Is the Pope of Rome a stockholder of The Literary Digest?” And so it goes—all equally absurd and amusing. We could add more to this list, and yet all of these questions in recent days are but repetitions of what we have been experiencing all down the years from the very first Poll.

Problem.—Now, are the figures in this Poll correct? In answer to this question we will simply refer to a telegram we sent to a young man in Massachusetts the other day in answer to his challenge to us to wager \$100,000 on the accuracy of our Poll. We wired him as follows:

“For nearly a quarter century, we have been taking Polls of the voters in the forty-eight States, and especially in Presidential years, and we have always merely mailed the ballots, counted and recorded those

returned and let the people of the Nation draw their conclusions as to our accuracy. So far, we have been right in every Poll. Will we be right in the current Poll? That, as Mrs. Roosevelt said concerning the President's reelection, is in the ‘lap of the gods.’

“We never make any claims before election but we respectfully refer you to the opinion of one of the most quoted citizens to-day, the Hon. James A. Farley, Chairman of the Democratic National Committee. This is what Mr. Farley said October 14, 1936:

“Any sane person can not escape the implication of such a gigantic sampling of popular opinion as is embodied in The Literary Digest straw vote. I consider this conclusive evidence as to the desire of the people of this country for a change in the National Government. The Literary Digest poll is an achievement of no little magnitude. It is a Poll fairly and correctly conducted.”

In studying the table of the voters from

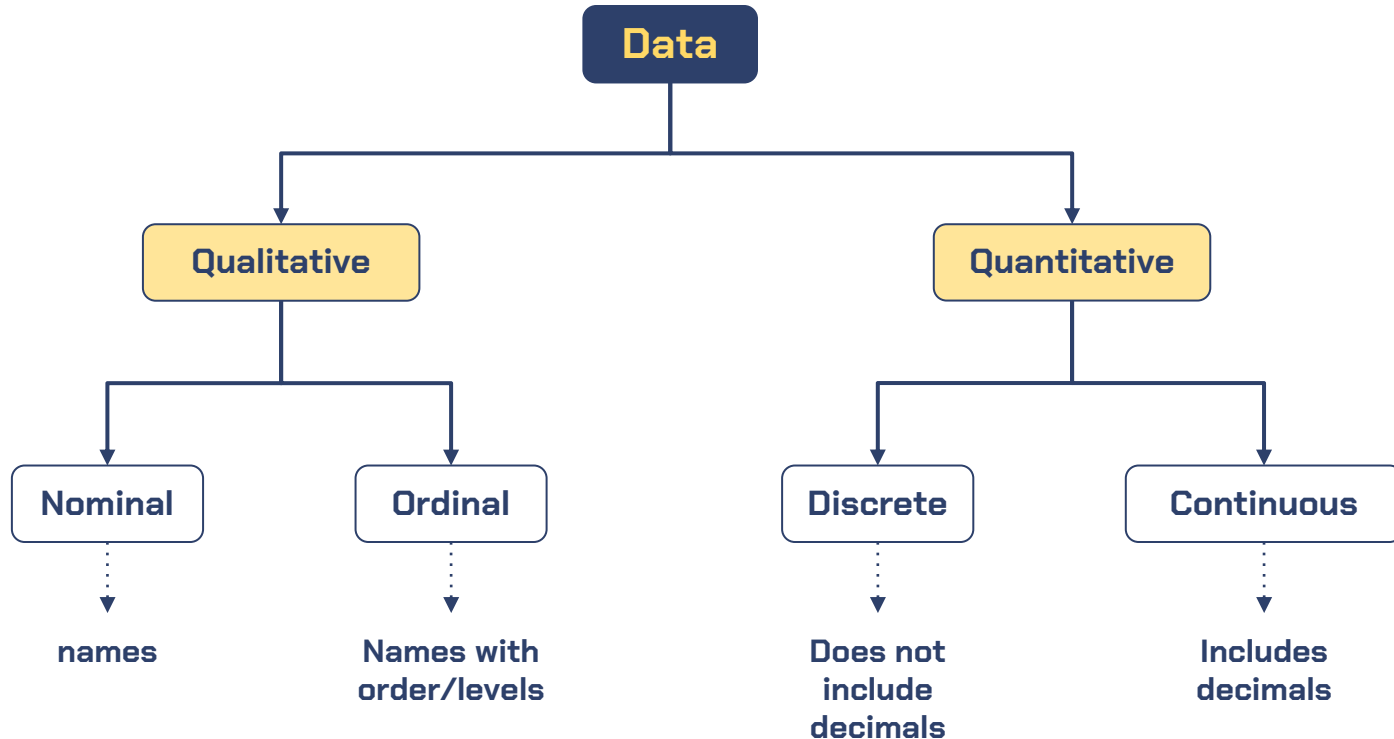
The statistics and the material in this article are the property of Funk & Wagnalls Company and have been copyrighted by it; neither the whole nor any part thereof may be reprinted or published without the special permission of the copyright owner.

Other methods:

- Available sample
- Volunteer sample
- Quota sample
- Referral sample
- Stratified sample



Types of data - based on the information type



Types of data - based on collection

	2015	2016	2017	2018	2019	2020
Employees	50	55	70	90	160	500
Profit	\$1000	\$6000	\$15,000	\$30,000	\$80,000	\$100,000
Sales	400	800	1000	1500	4000	10000

Time series data

.....▶ Data collected over **several time periods**

	2015	2016	2017	2018	2019	2020
Employees	50	55	70	90	160	500

Cross-sectional data

↓
Data collected from **several entities** at the **same time**

	2015
Employees	50
Profit	\$1000
Sales	400



Types of data - based on source

Experimental

Introduce a **stimulus** and study the **effects**

Data is collected for controlled variables (*stimulus*) and variable of interest (*effects*)

Observational

No control over any variables

Example: **Surveys**

