

+

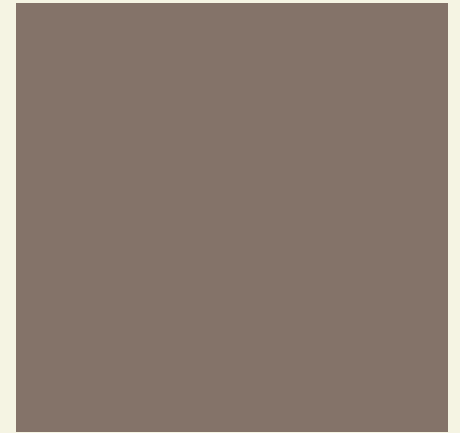
ICCPP-STATISTICS

- Statistics

Vishal Lohchab

*Scientific Assistant of
Prof. Dr. Hans-Werner Gessmann
Director ICCPP International*



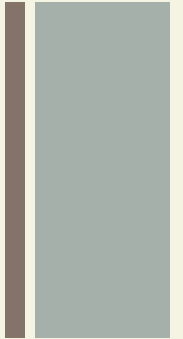


Ronald Aylmer Fisher (1890-1962)

Father of Statistics



Origin of Statistic

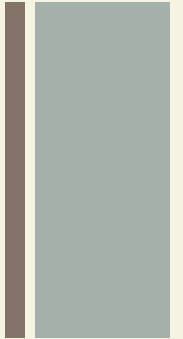


- The **word statistic** indirectly comes from the medieval Latin word *status*, for a political state although there is also a closely related word in German (*statistik*) which is also used in a political sense.

“**Statistik**” was popularized by German political scientist Gottfried Aschenwall (1719-1772) in his “Vorbereitung zur Staatswissenschaft” (1748).



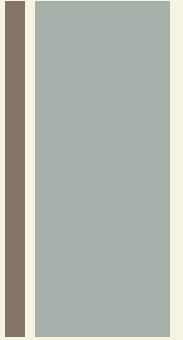
Origin of Statistic



- According to Leiden University, it's difficult to know exactly when the word ceased to have a meaning close to a “political state” and became more of a mathematical term.
- The first time the word was used in the Oxford English Dictionary is in 1770, in W. Hooper's translation of Bielfield's Elementary Universal Education: “The science, that is called statistics, teaches us what is the political arrangement of all the modern states of the known world.”



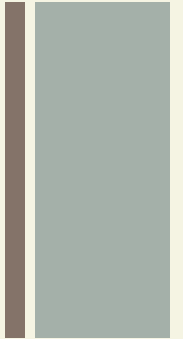
Origin of Statistic



- The Online Etymology Dictionary states that the first recorded time the word meant “numerical data collected and classified” was 1829 and the abbreviated form stats first appeared in 1961.



Origin of Statistic

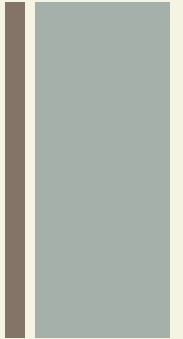


- Webster's 1828 dictionary defines statistics as: A collection of facts respecting the state of society, the condition of the people in a nation or country, their health, longevity, domestic economy, arts, property and political strength, the state of the country, &c.



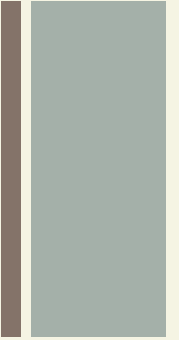
What is a Statistic?

- Statistics is the science concerned with developing and studying methods for collecting, analyzing, interpreting and presenting empirical data.
- A statistic is a piece of data from a portion of a population. It's the opposite of a parameter. A parameter is data from a census. A census surveys everyone.



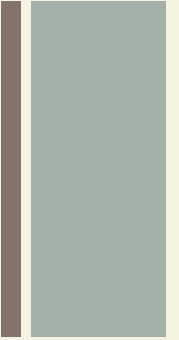
+ What is a Statistic used for?

- Statistics is a way to understand the data that is collected about us and the world.
- For example, every time you send a package through the mail, that package is tracked in a huge database.



+ What is a Statistic used for?

- The UPS database is 17 terabytes about as big as if you cataloged every book in the Library of Congress. All of that data is meaningless without a way to interpret it, which is where statistics comes in.
- Statistics is about data and variables. It's also about analyzing that data and producing some meaningful information about that data.



+ What is a Statistic: Types

1.

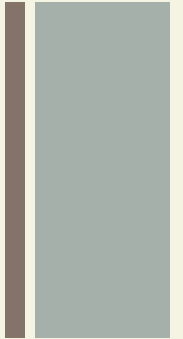
Descriptive Statistics = Describe data.

Includes sample mean or sample median.

Order statistics are a subset of descriptive statistics. They tell you something about how the data is ordered.



What is a Statistic: Types



- For example, measurements like the sample minimum. You know the order is #1. Also includes charts and graphs. Anything that describes data is descriptive statistics.

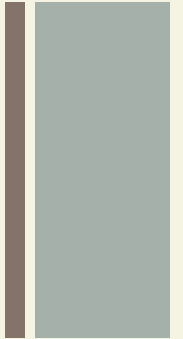
+ What is a Statistic: Types

2.

Estimators = Used to guess at a parameter. In other words, something about a population. Often taken from descriptive stats.



What is a Statistic: Types



- For example, if you know the sample mean you can use it to guess what the population mean is. Used in inferential statistics. Inferential stats is just a “best guess” about something, based on data.

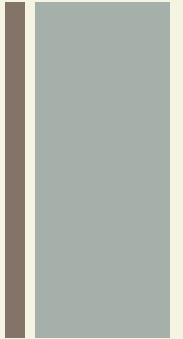
+ What is a Statistic: Types

3.

Test Statistics, which are used in null hypothesis testing. That's where you take a known fact about a population and then test that fact to see if it is true or not.



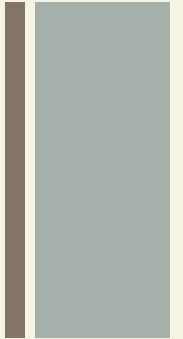
What is a Statistic: Types



- A “population” could be real people in a trial. Or it could be TVs in a factory. Which test statistic you use depends on what kind of data you have.
Examples of test stats: Chi-square.



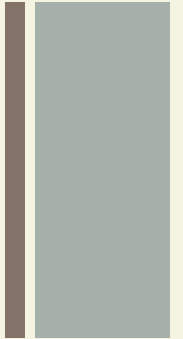
What is a Statistic: Types



- A statistic can be more than one type.
For example, the sample standard deviation can be used as a descriptive statistic to describe the standard deviation of a sample.
It can be used as an estimator: To estimate the population standard deviation.
And it can be used to test a theory (a hypothesis).



What is a Statistic? Notation



- In general, stats notation is in Roman letters, a-z. Parameters have Greek letters or uppercase Roman). If some letters look the same: look closely. For example, look for the small p and large P. Usually, if you see a large letter (i.e. P), it's a parameter. Small letters usually mean it's a stat.



What is a Statistic? Notation

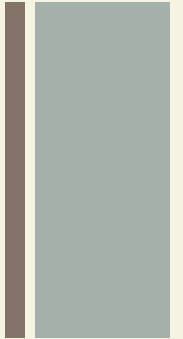
Measurement	Statistic (Roman or lowercase)	Parameter (Greek or uppercase)
Population Proportion	p	P
Data Elements	x	X
Population Mean	\bar{x}	μ
https://www.statisticshowto.com/probability-and-statistics/standard-deviation/	s	σ
Variance	s^2	σ^2
Number of elements	n	N
Correlation Coefficient	r	ρ

+ What is a Statistic: Data

- You might think that data is a list of numbers. However, in statistics, “Data” means something a little different; Data contains the who and what about something (the “something” could be anything from a book in a bookstore to a batting average to a choice about elections). Data can have numerals that have meaning.



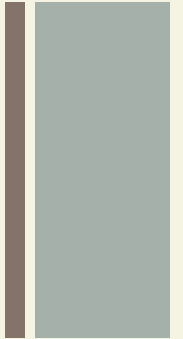
What is a Statistic: Data



- For example, 1453767142 is the ISBN for the Practically Cheating Statistics Handbook. The What as it related to ISBNs is the name of the book (The Practically Cheating statistics Handbook) and the Who as it related to book sales could be the person who ordered the book or it could be the purchase orders (as opposed to the individuals who placed those orders).



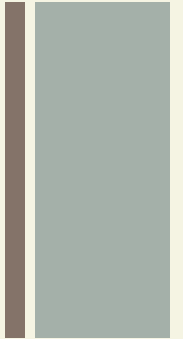
What is a Statistic: Variables



- You might be familiar with variables from algebra, like “x” or “y.” They stand for something (usually a number that you plug-in to solve an equation). In statistics, variables are broken down into two types: numerical or quantitative variables and categorical variables. Numerical variables are the variables you’re most familiar with: numbers.



What is a Statistic: Variables



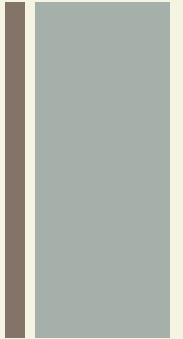
- For example, those “x” and “y” variables in algebra stand for a number.

Categorical variables are variables that aren't numbers: they are descriptive.

For example, sex (male or female), occupation, school district, state and dog breeds are all types of categorical variables.



References



Moses L E (1986). *Think and Explain with Statistics*, Addison-Wesley, ISBN 978-0-201-15619-5. pp. 1–3

Hays W L (1973). *Statistics for the Social Sciences*, Holt, Rinehart and Winston, p.xii, ISBN 978-0-03-077945-9

Moore D (1992). "Teaching Statistics as a Respectable Subject". In F. Gordon; S. Gordon (eds.). *Statistics for the Twenty-First Century*. Washington, DC: The Mathematical Association of America. pp. 14–25. ISBN 978-0-88385-078-7.

Glen S. "Welcome to Statistics How To!" From [StatisticsHowTo.com](https://www.statisticshowto.com/):
Elementary Statistics for the rest of us!
<https://www.statisticshowto.com/>