1. What is the difference between **Inferential statistics** & **Descriptive statistics**?

**Descriptive statistics** are used to describe the characteristics or features of a dataset. The term descriptive statistics can be used to describe both individual quantitative observations as well as the overall process of obtaining insights from the data.

**Descriptive statistics:**

* Describe the features of populations and/or samples.
* Organize and present data in a purely factual way.
* Present final results visually, using tables, charts, or graphs.
* Draw conclusions based on known data.
* Use measures like central tendency, distribution, and variance.

**Inferential statistics** focus on making generalizations about a larger population based on a representative sample of that population. Because inferential statistics focuses on making predictions & its results are usually in the form of a probability.

**Inferential statistics:**

* Use samples to make generalizations about larger populations.
* Help us to make estimates and predict future outcomes.
* Present final results in the form of probabilities.
* Draw conclusions that go beyond the available data.
* Use techniques like hypothesis testing, confidence intervals, and regression and correlation analysis.

1. what is the difference between **population** and **sample** in inferential statistics?

Inferential statistics let you draw conclusions about populations by using small samples. inferential statistics provide enormous benefits because typically you can’t measure an entire population.

𝐏𝐨𝐩𝐮𝐥𝐚𝐭𝐢𝐨𝐧𝐬

* The population represents the entirety of persons, units, objects & anything that is capable of being conceived, having certain properties.
* For example, the population of a country includes all people currently within that country. It’s a finite but potentially large list of members.

𝐒𝐚𝐦𝐩𝐥𝐞:

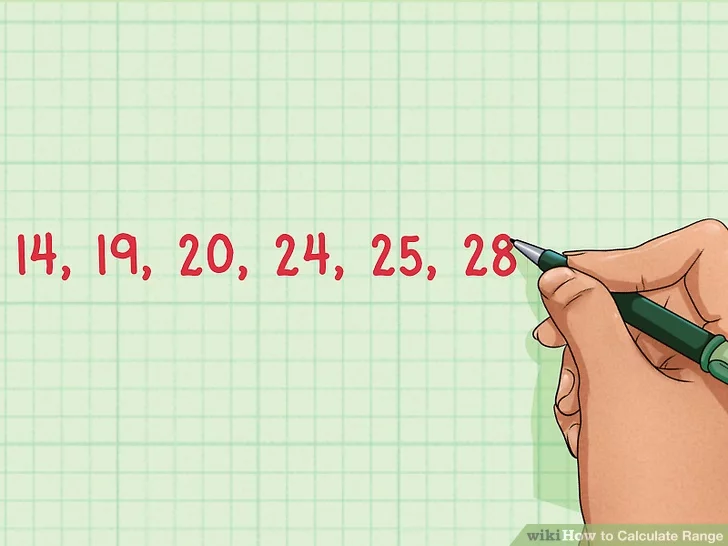
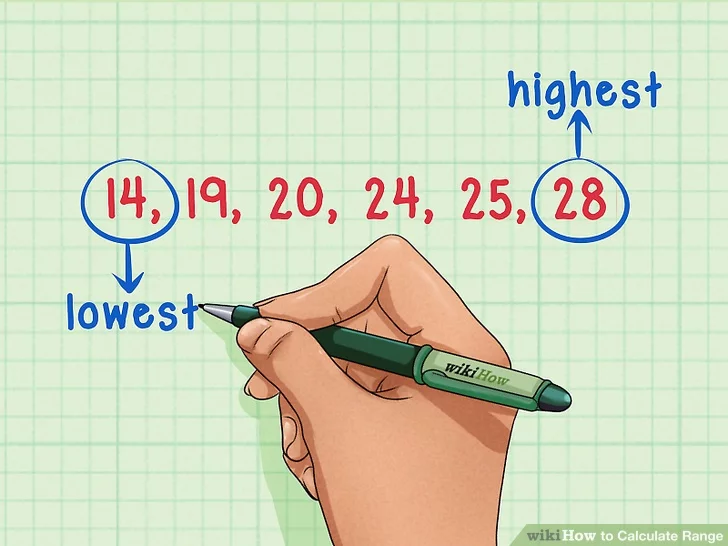
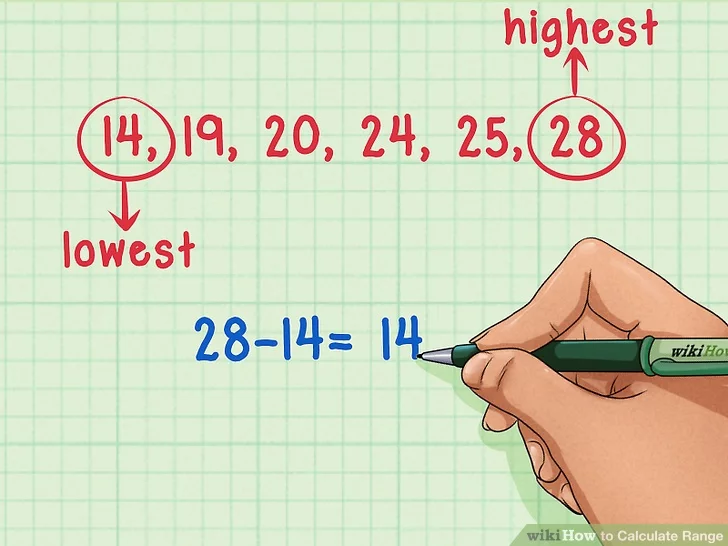
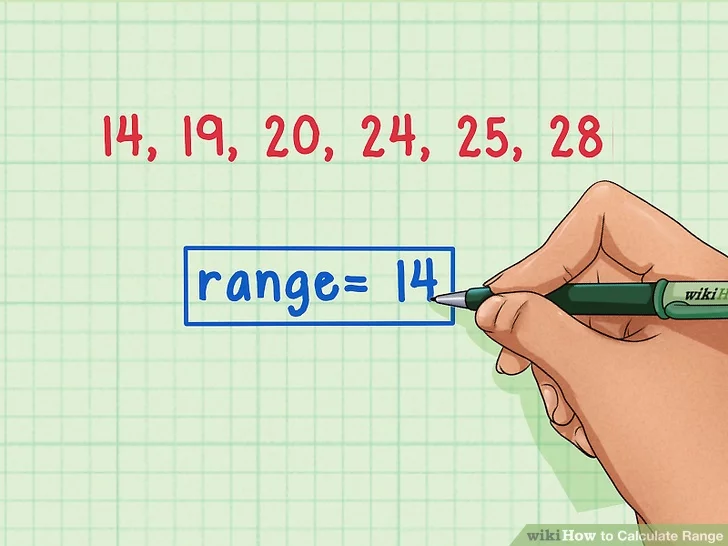
* The sample is a finite subset of the population, that is chosen by a systematic process, to find out the characteristics of the parent set.
* After drawing the sample, you can measure one or more characteristics of all items in the sample, such as height, income, temperature, opinion, etc.

1. Most common characteristics used in **descriptive statistics**?

* Measures of Central Tendency
* Measures of Dispersion
* Measures of frequency
* Measure of position

1. How to calculate **range** & **interquartile range**?

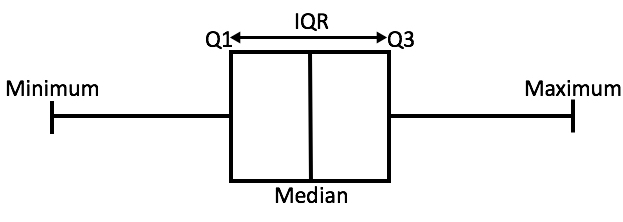
**Range:**

* The range is the difference between the smallest and highest numbers in a list or set. To find the range, first put all the numbers in order. Then subtract the lowest number from the highest. The answer will be the range of the list.

**Interquartile range:**

**IQR = Q3 – Q1.**



We can find the interquartile range or IQR in four simple steps:

* Order the data from least to greatest
* Find the median
* Calculate the median of both the lower and upper half of the data
* The IQR is the difference between the upper and lower medians.

1. How is the statistical significance of an insight assessed?

* Statistical significance is often calculated with statistical hypothesis testing, which tests the validity of a hypothesis by figuring out the probability that your results have happened by chance.
* The hypothesis is an assumption or belief about the relationship between your datasets. The result of a hypothesis test allows us to see whether this assumption holds under scrutiny or not.

**A standard hypothesis test relies on two hypotheses**.

**Null hypothesis:** The default assumption of a statistical test that you’re attempting to disprove (e.g., an increase in cost won’t affect the number of purchases).

**Alternative hypothesis:** An alternate theory that contradicts your null hypothesis (e.g., an increase in cost will reduce the number of purchases). This is the hypothesis you hope to prove.