

Data Analysis Basics

There are a variety of ways data can be analyzed. Choosing appropriate methods is important. Presenting (displaying) and reporting (interpreting) data properly is also essential.

Descriptive Statistics

Descriptive statistics are used to summarize information obtained from the sample without making any direct claims about the population. Descriptive statistics are used to present the sample data in more meaningful ways, which helps us understand and interpret the data later. While descriptive statistics are meant to summarize and present survey results, you may want to point out interesting aspects or patterns in the findings, but you don't make explicit inferences or generalizations about the population yet. Common visualizations of survey results include bar charts, frequency distributions, or pie charts. Tables can also be useful for displaying descriptive data.

Inferential Statistics

Inferential statistics are used to draw conclusions (inferences or generalizations) about the population from which a sample was drawn. Statistical techniques will use confidence intervals (margins of error), regressions (predictions), or hypothesis testing (involving statistical and practical significance) to estimate something about the population based on the sample.

Statistical significance and practical significance are determined to provide evidence that the result has some importance. Statistical significance refers to the probability that observations in the sample may have occurred due to chance. Given a large enough sample, despite seemingly insubstantial results, one might still find a satisfactory level of statistical significance. Practical significance, on the other hand, looks at whether the magnitude of the observation is large enough to be considered substantial. For example, when considering the difference between the mean of two groups, you might find that a difference of 1% is statistically significant (e.g., it has only a 5% chance of occurring due to chance), but you realize that the magnitude of this difference has no practical significance (i.e., the difference is not really that different in practical terms).

Types of Data

Prior to conducting your data analysis, you need to make sure you understand the type of data you have so you can select appropriate statistical methods. For certain types of data, it is inappropriate to use some statistical analyses.

There are four basic types of data, although many statistical programs combine interval and ratio data (calling it scale data) because the statistical methods used with these types of data tend to be the same.

Nominal Data

Nominal data might best be described as categorical. These data are the most basic type of information you might collect in a survey. Rules are used to specify membership in a category. Frequency (group size, counting) and proportional information (percentages) are used to report these types of data. These are also commonly used to disaggregate data when comparing groups. However, when making group comparisons, group membership rules should make it so that

groups are mutually exclusive (i.e., no individual is a member of both groups being compared).

Ordinal Data

These data have some sense of order, but the intervals between points on these types of scales are not equidistance. For example, placement results or preference (i.e., first, second, and third) have an order, but differences between various points on the scale are not consistent (first and second choices may be close, but both might be far more preferred than anything that comes next). Computing the mean and standard deviation for ordinal data is discouraged and, in most cases, inappropriate (although some researchers regularly compute averages for results obtained from Likert scales); frequencies (mode) and proportions (percentages) are best used when describing results based on this type of data along with ranking results. When making inferences, some nonparametric statistical procedures might also be appropriate.

Scale Data (Interval and Ratio)

Scale data have all the properties of nominal and ordinal data but also have the characteristic of equal intervals; in the case of ratio-level data, they have a true zero point. This means the distance between each point on the numeric scale being used is the same regardless of where on the scale you look. For ratio-level data, this also means that comparisons can be made about differences in magnitude (e.g., twice as much). It is appropriate to calculate the mean and standard deviation of scale-level data. You can add and subtract interval-level data, but you can also multiply and divide ratio-level data. With scale data, in addition to means and standard deviations, inferential statistics can be used—including *t*-tests, correlations, and regression analysis.

Table 1

Types of Data and Their Characteristics

Type and Characteristic	Typical Applications	Scale Characteristics Possessed			
		Identity	Order	Distance	Origin
Nominal – identification or classification	Gender School number Geographical location	⊗			
Ordinal – specifies order or rank	Brand preference Placement Agreement (Likert scales)	⊗	⊗		
Interval – specifies order based on equidistant intervals (implies equal increments of measurement)	IQ, test scores Degree in F° and C° Time of day	⊗	⊗	⊗	
Ratio – interval data with a zero point denoting an absence of the characteristic being measured.	# correct, Units sold Distance, Time (amount) Height, Weight, Age Degrees in K°	⊗	⊗	⊗	⊗

Data Visualization

How you present results is important. Primarily used with descriptive statistics, tables, graphs, and charts summarize information in a readable format. These presentation methods not only organize large amounts of information, but they can also help focus readers' attention on patterns and important findings. They are often the basis from which inferential statistics are calculated. While this course does not elaborate on the data visualization theories and practices, several resources exist to help develop data visualization skills (see references for some examples).

Chapter Summary

- Descriptive statistics are used to summarize survey results.
- Inferential statistics provide evidence used to support conclusions (inferences or generalizations).
- Data obtained from a survey will fall into various data types (nominal, ordinal, interval, or ratio).
- The appropriateness of the statistical analysis used is determined by the characteristics of the data (i.e., type of data).
- In survey research, perhaps the most controversial statistical issue pertains to whether data obtained from Likert scales can be used as interval-level data (i.e., assigning numbers to responses and averaging the findings).
- Data visualization theory and practices are extremely important for presenting descriptive statistics effectively.

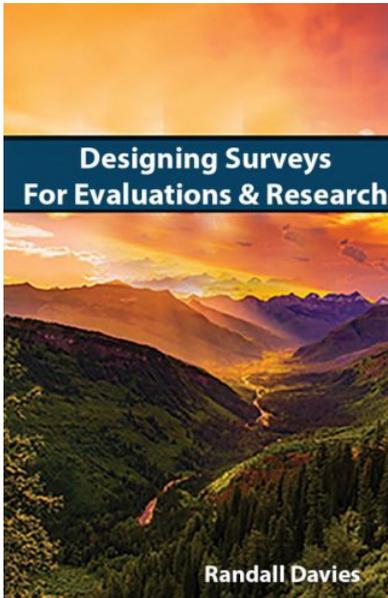
References

Evergreen, S. D. H. (2018). *Presenting data effectively*, 2nd Edition. Sage Publishing.

Evergreen, S. D. H. (2019). *Effective data visualization: The right chart for the right data*, 2nd Edition. Sage Publishing.

[Data Visualization Presentation: Choosing Charts](#)

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. Wiley Publishing.



Davies, R. S. (2020). *Designing Surveys for Evaluations and Research*. EdTech Books.

https://edtechbooks.org/designing_surveys



CC BY-NC: This work is released under a CC BY-NC license, which means that you are free to do with it as you please as long as you (1) properly attribute it and (2) do not use it for commercial gain.