Statistics for Social Research I

Derrick Snow

Omega Graduate School

February 17, 2024

Professor

Dr. Taladay

Assignment

### *Developmental Readings*

Review Assignment #3, the course essential elements, assigned readings, and recommended readings to identify selections of books and scholarly articles to identify and select developmental reading sources and entries.

* Refer to the “[Student Guide to Developmental Readings](https://drive.google.com/file/d/161V_FaYR2BnNGCSFUlWPjUSIQzcH04Hq/view?usp=share_link)” for updated information on sample comments, rubrics, and key definitions related to developmental readings.

**Source One:** Terrell, S. R. (2021). *Statistics Translated: A Step-by-Step Guide to Analyzing and Interpreting Data*. Guilford Publications.

**Comment 1:**

**Quote/Paraphrase:** “Nominal data, sometimes called *categorical* or *discrete*, are data that are measured in categories” (Terrell, 2021, p. 39).

**Essential Element:** This comment relates to the essential element of Measurement Levels.

**Additive/Variant Analysis:** This is an additive reading, where the nominal level of measurement is a qualitative grouping in a variable, such as sex or gender grouped by male or female. With nominal data, one could report the counts and percentages that fall within the categories.

**Contextualization:**  Nominal data has limited statistical operations that could be performed on the data. However, they are still extremely useful variables to use in research, such as demographic information on the participants, such as variables like (does one’s gender affect y, does one’s race influence y). It is also useful to display nominal data in frequency tables.

**Comment 2:**

**Quote/Paraphrase:** “The second type of data is called *ordinal* data or, as it is sometimes called, *rank* data” (Terrell, 2021, p. 39).

**Essential Element:** This quote refers to the essential element of Measurement Levels.

**Additive/Variant Analysis:** This comment is additive to my understanding of ordinal variables. There are qualitative responses that have a logical order to them, such as the socio-economic class of the lower class, working class, middle class, and upper class. Like nominal data, there are a limited number of statistical operations that could be performed on ordinal data. However, it is possible to average ordinal data if responses are converted into numerals, such as the Linkert Scale.

**Contextualization:** Ordinal data can be useful for researchers who would like to classify their sample that have logical order without asking a specific value to a question. Such as asking a participant to select their income range, i.e., <$10,000; $10,000 - $20,000, etc., or if the researcher asks for the participant’s actual annual income the researcher could group that data into ranges and thus ordinal data.

**Comment 3:**

**Quote/Paraphrase:** “Interval data are the first of two types of data that are called *quantitative* or *continuous*. By this, we mean that a data value can fall anywhere on a number line within the range of a given dataset” (Terrell, 2021, p. 40).

**Essential Element:** This comment refers to the course essential element of Measurement Levels.

**Additive/Variant Analysis:** This information about the level of measurement of interval data is additive to my understanding of interval data. As the quote states, the values of interval data could lay anywhere on a number line, including negative numbers.

**Contextualization:** Mathematical and statistical operations can be used on interval data; thus, they could be statistically analyzed to evaluate differences or relationships. For example, we can measure the average annual temperature in different regions of the world over time. The poles where it is cold may have negative average temperatures, and along the equator where it is warm will have warmer positive temperatures.

**Comment 4:**

**Quote/Paraphrase:** “Ratio data are also classified as quantitative or continuous data. Ratio data differ from interval data because they do have an absolute zero point and the various points on the scale can be used to make comparisons between one another” (Terrell, 2021, p. 41).

**Essential Element:** Thiscomment refers to the course essential element of Measurement Levels.

**Additive/Variant Analysis:** This comment about the ratio level of measurement is additive to my current understanding of ratio-level variables. Whereas interval data and ratio data can have a wide range of mathematical and statistical operations. The difference between interval and ratio data is that in ratio data zero has an absolute value of zero or none. For example, one’s annual income can have any value from zero to anywhere above, one cannot have a negative income.

**Contextualization:** Interval and ratio data are quantitative data, where any number of statistical analyses could be done. In social research, the use of ratio data is used, such as income, years of education, age, scores on an exam, etc. In many cases, quantitative data is grouped in a category called interval-ratio data or level of measurement. Social statistics in large part deals with ratio data to perform descriptive statistics as well as correlation and test statistics.

**Source Two:** Healey, J. F. (2016). *The Essentials of Statistics: A Tool for Social Research*. Cengage Learning.

**Comment 5:**

**Quote/Paraphrase:** “In the language of science, the causes are independent variables and the effects are or results are dependent variables” (Healey, 2016, p. 4).

**Essential Element:** This comment refers to the course's essential element of Dependent and Independent Variable.

**Additive/Variant Analysis:** This comment is additive to my understanding of independent and dependent variables. However, I take issue with the author’s use of “causes” where in much of social science research we may not be able to identify the “cause,” but we can measure the influence or impact independent variables have on dependent variables.

**Contextualization:** The variable types of independent variables and dependent variables are important to understanding where in one research a variable may considered an independent variable and in another study that same variable may be a dependent variable. Independent variables (x) are the variables that are measured to have an impact or an affect on the dependent variable (y). Often symbolized by x--> y. We can measure if years of education (x) have an impact on annual income (y).

**Comment 6:**

**Quote/Paraphrase:** “…some commonly used techniques for presenting research results, including tables, percentages, rates, and graphs. These univariate descriptive statistics are not mathematically complex (although they are not as simple as they might seem at first glance), but they can be extremely useful tools for organizing and analyzing results and communicating conclusions” (Healey, 2016, p. 24).

**Essential Element:** This comment refers to the course's essential element of Descriptive Statistics.

**Additive/Variant Analysis:** This comment is additive to my understanding of descriptive statistics. Descriptive statistics are statistics that are performed on a variable to describe the variable in statistical terms. The descriptive statistics that are used on a variable are dependent on the level of measurement of the variable.

**Contextualization:** Descriptive statistics are particularly important to the researcher, not only to describe the findings of the research but also to gain an understanding of the responses to the variables to know how to manage them in your research. The quote identifies tables, percentages, and graphs as descriptive statistics, along with those measures of central tendency (mean, median, mode) and measure of dispersion (variation and standard deviation), along with minimum, maximum, and range are descriptive statistics. By performing descriptive statistics on data, one could find data entry errors or outliners, and the researcher must make choices in how to handle these abnormalities.

**Source Three:** Siedlecki, S. L. (2020). Understanding Descriptive Research Designs and Methods. *Clinical Nurse Specialist*, *34*(1), 8–12. <https://doi.org/10.1097/NUR.0000000000000493>

**Comment 7:**

**Quote/Paraphrase:** “The purpose of descriptive studies is to describe individuals, events, or conditions by studying them as they are in nature. The researcher does not manipulate any of the variables but rather only describes the sample and/or the variables” (Siedlecki, 2020, p. 8).

**Essential Element:** This comment is associated with the essential element of Descriptive Statistics.

**Additive/Variant Analysis:** This quote is additive to my understanding of descriptive statistics, where the researcher is using statistical process in describing the variables in their study to gain a sense of the responses and the demographic information of their sample. This is the first step before moving into any other statistical analysis.

**Contextualization:** The author of this article is suggesting that descriptive statistics could be the essence of a publication, which is true especially in healthcare where the first step in solving a problem is understanding the problem and the use of descriptive statistics will go far in laying out the magnitude of a problem. The use of descriptive statistics is important to researchers in understanding the data with which they are working. Just looking at a spreadsheet full of numbers does not help us understand the data, we need to summarize the data to make sense out of it. Descriptive statistics provide us with the summarization needed to gain an understanding of data we are working with in the research.

**Comment 8:**

**Quote/Paraphrase:** “It is important, when defining variables, that the operational definition (measurement) is congruent with the conceptual definition for each variable” (Siedlecki, 2020, p. 9).

**Essential Element:** This comment is associated with the essential element of Dependent and Independent Variable.

**Additive/Variant Analysis:** This concept of defining variables and operationalizing variables is additive to my understanding of appropriately handling variables in research. Operationalizing means that the researcher needs to clearly define how they are using a variable, such as education level, which could be measured in a few ways such as the number of years attended school, grade level, or class standing. Thus, the researcher needs to define/operationalize how they are defining educational level.

**Contextualization:** As the author states it is important for the researcher to clearly define the variables they are working with and how it is being measured from the raw data they have collected. In social research, these operationalized definitions are key to having your readers understand your research since social variables can be measured in numerous ways. In some of the research I have come across, I was not entirely clear on how the researchers were understanding the variables, they have just taken it for granted, but it points out that we as researchers need to dedicate some space in our writings to clearly operationalize our variables.

**Source Four:** Kumar, G., Adams, A., Hererra, M., Rojas, E. R., Singh, V., Sakhuja, A., Meersman, M., Dalton, D., Kethireddy, S., Nanchal, R., & Guddati, A. K. (2021). Predictors and outcomes of healthcare-associated infections in COVID-19 patients. *International Journal of Infectious Diseases*, *104*, 287–292. <https://doi.org/10.1016/j.ijid.2020.11.135>

**Comment 9:**

**Quote/Paraphrase:** “Categorical data were described using frequency count and percentages. Medians and interquartile ranges were used for continuous variables as they were not normally distributed” (Kumar, et al, 2021, p. 104).

**Essential Element:** This comment is associated with the course’s essential element ofDescriptive Statistics.

**Additive/Variant Analysis:** This concept in this quote is additive to my understanding of the use of descriptive statistics. The author used frequencies and relative frequencies as descriptive statistics for nominal data and median and interquartile ranges as descriptive statistics for interval-ratio variables.

**Contextualization:** Descriptive statistics inform us about the responses and the layout of the data, so we can gain an understanding of the responses in a summarized way using descriptive statistics. The author states that the interval-ratio data was not normally distributed, the only way to find this out is by using descriptive statistics or a histogram, in which the researcher chose to use the median and interquartile range rather than the mean and standard deviation precisely because they used descriptive statistics to identify that the data was not normally distributed and thus used the appropriate descriptive statistics to describe the continuous variable.

**Source Five:** Kowal, S., Ng, C. D., Schuldt, R., Sheinson, D., Jinnett, K., & Basu, A. (2023). Estimating the US Baseline Distribution of Health Inequalities Across Race, Ethnicity, and Geography for Equity-Informative Cost-Effectiveness Analysis. *Value in Health*, *26*(10), 1485–1493. <https://doi.org/10.1016/j.jval.2023.06.015>

**Comment 10:**

**Quote/Paraphrase:** “The 5-year ACS data were used to obtain reliable estimates of disability and sociodemographic variables for all counties within the United States, including the least populated” (Kowal et al, 2023, p. 1486).

**Essential Element:** This comment refers to the course essential element of Measurement Levels.

**Additive/Variant Analysis:** This comment is additive to my understanding of demographic data as nominal data. The research used the 5-year estimate of the American Community Survey from the U.S. Census Bureau to obtain the number of people disabled and other social, demographic, and economic information by U.S. counties.

**Contextualization:** Demographic information is important in social research because demographic and other socioeconomic information informs us of the population that we are working within our research. Understanding the information on participants in our study ties us to the social in social research, we are studying social phenomena actual things that impact the lives of people. Even though we classify these demographic characteristics of people as nominal variables in our data set, these cases represent the actual lives of people.

**Source Six:** Schwartz, B. S., Kolak, M., Pollak, J. S., Poulsen, M. N., Bandeen-Roche, K., Moon, K. A., DeWalle, J., Siegel, K. R., Mercado, C. I., Imperatore, G., & Hirsch, A. G. (2022). Associations of Four Indexes of Social Determinants of Health and Two Community Typologies with New Onset Type 2 Diabetes Across a Diverse Geography in Pennsylvania. *PLOS ONE*, *17*(9), e0274758. <https://doi.org/10.1371/journal.pone.0274758>

**Comment 11:**

**Quote/Paraphrase:** “Univariate distributions for study variables were first examined. Key study variables were then described within community definitions of interest” (Schwartz, et al, 2022, p. 7).

**Essential Element:** This comment refers to the course's essential element of Descriptive Statistics.

**Additive/Variant Analysis:** This quote is additive to my understanding of the use of descriptive statistics in a research study. The researcher performed descriptive statistics on the variables and from that information was able to use to further describe the variables in the operationalization of the variables.

**Contextualization:** Conducting research is a large process of collecting data then analyzing the data, but before the process of analyzing data the researcher must gain a sense of the data, they have collected to understand the dynamics of the data and to know how to continue with analyzing the data. This is done through the process of descriptive statistics. Performing descriptive statistics not only informs the researcher but also informs the readers of the data used in the research. Readers of research are not only interested in the ultimate findings of the research but also what to know about the data that was collected, as they may also gain insight into the sample being studied.

**Source Seven:** Spoer, B. R., Chen, A. S., Lampe, T. M., Nelson, I. S., Vierse, A., Zazanis, N. V., Kim, B., Thorpe, L. E., Subramanian, S. V., & Gourevitch, M. N. (2023). Validation of a Geospatial Aggregation Method for Congressional Districts and Other US Administrative Geographies. *SSM - Population Health*, *24*, 101511. <https://doi.org/10.1016/j.ssmph.2023.101511>

**Comment 12:**

**Quote/Paraphrase: “**We perform sensitivity analyses to validate the aggregation method and optimize the rigor of our CD estimates. We calculate summary measures (mean, median, standard deviation (SD) for derived estimates” (Spoer, et al, 2023, p. 3).

**Essential Element:** This comment refers to the course's essential element of Descriptive Statistics.

**Additive/Variant Analysis:** This comment is additive to my understanding of descriptive statistics where some common descriptive statistics are mean, median, and standard deviation. Having the mean and median will inform us of the skewness of the data, and the standard deviation informs us of the dispersion of the data.

**Contextualization:** Descriptive statistics are useful to help understand the data, but they also give us the means to present the data in an understandable way to our readers. They are used to inform the readers of the highlights of the research by providing summary data in descriptive statistics so our readers will be able to understand our research better, this is a primary goal of research to inform people of a new understanding of a social phenomenon.

**Works Cited**

Healey, J. F. (2016). *The Essentials of Statistics: A Tool for Social Research*. Cengage Learning.

Kowal, S., Ng, C. D., Schuldt, R., Sheinson, D., Jinnett, K., & Basu, A. (2023). Estimating the US Baseline Distribution of Health Inequalities Across Race, Ethnicity, and Geography for Equity-Informative Cost-Effectiveness Analysis. *Value in Health*, *26*(10), 1485–1493. <https://doi.org/10.1016/j.jval.2023.06.015>

Kumar, G., Adams, A., Hererra, M., Rojas, E. R., Singh, V., Sakhuja, A., Meersman, M., Dalton, D., Kethireddy, S., Nanchal, R., & Guddati, A. K. (2021). Predictors and outcomes of healthcare-associated infections in COVID-19 patients. *International Journal of Infectious Diseases*, *104*, 287–292. <https://doi.org/10.1016/j.ijid.2020.11.135>

Terrell, S. R. (2021). *Statistics Translated: A Step-by-Step Guide to Analyzing and Interpreting Data*. Guilford Publications.

Schwartz, B. S., Kolak, M., Pollak, J. S., Poulsen, M. N., Bandeen-Roche, K., Moon, K. A., DeWalle, J., Siegel, K. R., Mercado, C. I., Imperatore, G., & Hirsch, A. G. (2022). Associations of Four Indexes of Social Determinants of Health and Two Community Typologies with New Onset Type 2 Diabetes Across a Diverse Geography in Pennsylvania. *PLOS ONE*, *17*(9), e0274758. <https://doi.org/10.1371/journal.pone.0274758>

Siedlecki, S. L. (2020). Understanding Descriptive Research Designs and Methods. *Clinical Nurse Specialist*, *34*(1), 8–12. <https://doi.org/10.1097/NUR.0000000000000493>

Spoer, B. R., Chen, A. S., Lampe, T. M., Nelson, I. S., Vierse, A., Zazanis, N. V., Kim, B., Thorpe, L. E., Subramanian, S. V., & Gourevitch, M. N. (2023). Validation of a Geospatial Aggregation Method for Congressional Districts and Other US Administrative Geographies. *SSM - Population Health*, *24*, 101511. <https://doi.org/10.1016/j.ssmph.2023.101511>