Chapter 8
Analyzing Quantitative Data

Sema A. Kalaian
Eastern Michigan University, USA

Rafa Kasim
Indiana Tech University, USA

ABSTRACT

The main purpose of this chapter is to present a conceptual and practical overview of some of the basic and advanced statistical tools for analyzing quantitative data. Analyzing quantitative data involves two broad analytical methods that serve two main purposes, which are descriptive and inferential statistical methods. The chapter covers both descriptive and inferential quantitative methods. It covers some of the descriptive statistical methods such as mean, median, mode, variance, standard deviation, and graphical methods (e.g., histograms). It also covers inferential statistical methods such as correlation, simple regression, multiple regression, t-test for two independent samples, t-test for two dependent samples, and analysis of variance (ANOVA).

INTRODUCTION

Conducting quality research is dependent on many significant factors and one of these factors is the researcher's ability to analyze the collected data using the most appropriate statistical methods to answer the research questions of the scientific study. Depending on the research question(s) and the nature of the collected data, the researcher might use quantitative analytical methods, qualitative analytical methods, or a combination of both of the data analytical methods. In other words, the researcher must pick the proper quantitative, qualitative, or a combination of both analytical methods (mixed methods) based on the nature and type of the data (quantitative or qualitative) to accurately answer the research questions and draw valid conclusions about the theoretical model put forward for the study.

The focus of this chapter is analyzing quantitative data that are measured using continuous interval and/or ratio scales of measurement. Analyzing qualitative data are covered in another chapter of this book. However, analyzing quantitative data involves two broad analytical methods that serve two purposes; descriptive and inferential statistical methods. This chapter covers both descriptive as well as inferential quantitative methods. Accordingly, the chapter is organized into two sections. Section I covers descriptive statistical methods and Section II covers inferential statistical methods.

DOI: 10.4018/978-1-5225-0007-0.ch008

I. DESCRIPTIVE STATISTICS

Descriptive statistics is a set of statistical methods in the statistical toolbox for describing and providing summary measures of the quantitative data. Any collected original quantitative data is usually overwhelming and uninformative, especially when the amount of the data is large. Therefore, it is necessary for the quantitative data analyst to use descriptive statistics to describe and summarize a large amount of data using simple summary measures.

For example, consider the Grade Point Average (GPA), which is a simple descriptive summary measure used to represent how well each student is performing in an educational institution such as college or high school. This single summary measure, GPA, describes the general performance (achievement) of each student across a wide range of academic courses in the academic institution. So, we need to keep in mind that important details of the data are lost (in this example, the details of each of the courses that had been taken by a particular student) when we describe a large set of data (in this example, achievement scores of each student across all the courses that had been taken) with a single summary measure or indicator (in this example, GPA).

Using descriptive statistical methods for analyzing quantitative data is the first and the most significant step in the quantitative data analysis journey. For example, it helps the researcher to be familiar with the data and discover early in the data analysis process any anomalies in the data set. In the following five subsections, the descriptive statistical methods that can be used for analyzing quantitative data are covered.

1. Describing Data Using Frequency Distributions

The *frequency distribution* for a specific variable in a data set (e.g., the age variable for a group of students) is a tabular representation of the number of instances (occurrences or frequencies) of each numerical value for a specific variable (in this case, the age variable) in the data set.

The frequency distribution is one of the simplest statistical procedures and yet it is one of the most useful tools for describing a quantitative data set.

Creating frequency distributions for each of the variables in the quantitative data set is the first step in analyzing a data set to explore and examine the numerical content and characteristics of each of the variables in the quantitative data. It helps researchers and data analysts to (1) organize and summarize the quantitative data in tabular formats, which helps the data analyst to easily examine the data set; (2) identify and assess the amount of the missing numerical values in a quantitative data set; and (3) detect one or more outliers (unusual extreme values) in the quantitative data (Field, 2009; Kalaian, 2008; Witte & Witte, 2013). For example, a value of a grand point average (GPA) of 0.1 in the frequency distribution of GPA scores, where most of the scores are above 50 on a scale from 1 to100, is considered as being an outlier score. This outlier score requires special treatment from the researcher such as the need to verify its accuracy or to exclude the student with a GPA of 0.1 from the data set.

2. Describing Data Using Graphical Representations

The *graphs (charts)* are important and useful visual representations of a quantitative data set. Generally, graphical presentations of a data set provide a more powerful, informative, and interesting message than the tabular representations. Graphs and charts provide insights and help researchers to visually (1) summarize, describe, and examine the shape of the frequency distribution of a large amount of quantitative