

Chapter 12

Data Visualization in R

S. R. Mani Sekhar

M. S. Ramaiah Institute of Technology, India

Siddesh G. M.

M. S. Ramaiah Institute of Technology, India

Sunilkumar S. Manvi

REVA University, India

ABSTRACT

Data visualization helps the users to understand the relationships and associations between information. Visualization helps in minimizing the errors generated during decision making. Different visualization methods have been developed to unlock the valuable insight. These methods have been developed on the supposition that the information to be present is free from ambiguity. This chapter provides an overview of data visualization techniques in R programming. Various methods have been discussed with supported explanation and examples which in turn helps the reader to create their own visualization method. Later, four different case studies are presented to understand the importance and use of data visualization in real-world problems.

INTRODUCTION

Data visualization is a branch of art and science (Aparicio & Costa, 2014) which deals with the various methods used to visualized data in graphical form. The goal of data visualization is to provide user with understanding of the given data. It helps the user to understand the relationship, and association between information's.

Now a day's visualization of data is far beyond the general charts and graphs using Microsoft office, the information can be represented in further refined ways such as heat maps, bar chart, pie charts, histogram, and info-graphics etc. Effective data visualization helps users to examine and analyzing the goal of data and its evidence. It also makes multifaceted data more comprehensible, manageable, and usable.

Advance analysis and visualization helps the reader to distinguish relationships between information, regardless the data size. Visualization helps in minimize the errors generate during decision making and

DOI: 10.4018/978-1-5225-4999-4.ch012

prevents in drawing wrong conclusion. The generated results can be represented in a form that is simple to understand while allowing queries and investigation.

Selection of data visualization tools depends upon the data size, data types and a way in which data to be represented. This chapter illustrates the different build-in methods used in R programming for creation of graphs and charts, in turn helps the user to visualize data effectively.

The rest of the chapter is organized as follows: Section 2, Discuss the history of Data Visualization. Section 3, tells the importance of Data Visualization. Subsequently Section 4, says the importance of Data Visualization in R. Section 5, illustrates the different methods used in R for Data Visualization. Furthermore Section 6, Discuss a Case Study on United State Air Pollution. Later Section 7 discusses the conclusion.

HISTORY OF DATA VISUALIZATION

This section provides an outline of the brief history of data visualization. As shown in Table 1 (Friendly, 2006), It consist of five columns, first and second columns displays the serial numbers and year. Column third “Named” display the historical names of that years. Column fourth shows the methods used in that year for visualizations. Thereafter column fifth displays the different examples incorporated using corresponding methods.

WHY DATA VISUALIZATION?

Visualization helps in better understand of datasets. Following below points state that why visualization is necessary:

- Easily predict the area which required more attention;
- Elucidate which feature influence the prediction;
- Provide a top to bottom overview of complex datasets;
- Enable decision makers easily;
- It helps in understanding, prediction and analysis of information;
- Provide direct interface with datasets;
- Easily recognize structure and associations in datasets;
- Help in identifying patterns in datasets;
- Foster a new business language;
- Takes less time in clarify selection and analysis of data;
- It helps in encourages appropriate clarification, selection, and association for given data.

WHY VISUALIZATION USING R

Following below points tells that why R is better than other programming language:

- It is an open source language;

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/data-visualization-in-r/210971

Related Content

Rare Association Rule Mining: An Overview

Yun Sing Koh and Nathan Rountree (2010). *Rare Association Rule Mining and Knowledge Discovery: Technologies for Infrequent and Critical Event Detection* (pp. 1-14).

www.irma-international.org/chapter/rare-association-rule-mining/36896

A Perspective on Data Mining Integration with Business Intelligence

Ana Azevedo and Manuel Filipe Santos (2011). *Knowledge Discovery Practices and Emerging Applications of Data Mining: Trends and New Domains* (pp. 109-129).

www.irma-international.org/chapter/perspective-data-mining-integration-business/46893

Measuring Diversity at a Historically Black College of Dentistry

Garnett Lee Henley, Wanda Lawrence, Candace Mitchell, Donna Henley-Jackson and Tawana Feimster (2012). *Cases on Institutional Research Systems* (pp. 212-227).

www.irma-international.org/chapter/measuring-diversity-historically-black-college/60849

Understanding Credit Card User's Behaviour: A Data Mining Approach

A. de Carvalho, A. P. Braga, S. O. Rezende, E. Martineli and T. Ludermitz (2002). *Heuristic and Optimization for Knowledge Discovery* (pp. 241-262).

www.irma-international.org/chapter/understanding-credit-card-user-behaviour/22158

Using "Blackbox" Algorithms Such AS TreeNET and Random Forests for Data-Mining and for Finding Meaningful Patterns, Relationships and Outliers in Complex Ecological Data: An Overview, an Example Using G

Erica Craig and Falk Huettmann (2009). *Intelligent Data Analysis: Developing New Methodologies Through Pattern Discovery and Recovery* (pp. 65-84).

www.irma-international.org/chapter/using-blackbox-algorithms-such-treenet/24212