BASİC STAGES AND TYPES OF DATA ANALYSİS

Mahmudova S.
Nakhchivan State University
Ajdarova N.
Nakhchivan University
https://doi.org/10.5281/zenodo.10724527

Abstract

The article conducted an examination of different types of data and data analysis and their main stages. Data analysis is the process of researching, manipulating and modeling data to uncover useful concepts and insights that can be used to make business decisions. There are different types of information analysis that organizations can use for their own purposes and techniques. This includes analytics and generalized descriptive analytics. Each type serves a different purpose and offers unique information concepts. Furthermore, data analysis is usually divided into stages such as data collection, data processing, data analysis and information visualization. Understanding the basic stages of these types and data analytics can help better leverage businesses and gain competitive advantage.

Keywords: data, data analysis, data preparing, data modelling, data visualization

Data, that is, data, is information that has been transformed into an efficient form as a result of information processing. In computing, data can be in the form of text, documents, images, audio and video. According to today's computers and transmission media, any type of data has been converted into binary digital form. Data is stored in binary digital form both in cloud technology and on carriers. There are different types of data in statistics, there are data collected, analyzed, interpreted and presented. Data are individual pieces of different information recorded and used for the purpose of the analysis process. The two processes of data analysis are interpretation and presentation. Statistics are the result of data analysis. Data classification and data management are important processes that have a large number of labels and tags to determine the integrity and confidentiality of data. Data are divided into four categories in statistics. Nominal data, ordinal data, discrete data, continuous data. [1]

Qualitative data, also known as categorical data, describe data that fit into categories. Qualitative data is not digital. Categorical information such as a person's gender, hometown, etc. It includes categorical variables that describe some characteristics. Categorical dimensions are defined in terms of natural language specification, but not in terms of numbers. [1]

Sometimes categorical data can store numerical values, but that value has no mathematical meaning. Examples of categorical data are date of birth, favorite sport, school zip code. Here, date of birth and school zip code have a quantitative value, but no numeric meaning.

Nominal data is one type of qualitative data that helps to label variables without giving a numerical value. Nominal data is also called nominal scale. It cannot be ordered or measured. However, some data can be both qualitative and quantitative. Examples of nominal data are letters, symbols, words, severity, etc. we can show. It is checked using the clustering method with nominal data. In this method, the data is grouped into categories and then the frequency or percentage of the data can be calculated. This data is presented visually using pie charts. [2]

An ordinal data variable is a data type that follows a natural order. An important feature of nominal data is that the difference between data values is not determined. This variable includes more surveys, finance, economics, questionnaires, etc. belongs to. Ordinal data is usually represented using a column chart. This data is explored and interpreted using many visualization tools. Data can be expressed using tables, where each row in the table shows a different category.

Quantitative data is also known as numerical data that represents a numerical value. Numerical data gives information about the quantity of a certain thing. Some examples of numerical data are height, length, measurement, weight, etc. Quantitative data can be classified into two different types based on the data set. Two different classifications of numerical data are discrete data and continuous data.

Discrete data can only take discrete values. Discrete data contains only a finite number of possible values. These values cannot be meaningfully divided. Here you can count things with whole numbers. Example: the number of students in the class.

Continuous data is data that can be calculated. It has an infinite number of possible values that can be chosen within a given particular range. Example: temperature range.

The need for data around the world is increasing day by day and as the world becomes more data driven, analytics through data analysis is becoming an essential component and aspect of businesses large and small. Data helps companies make better business decisions and solve problem situations. With the help of data and database, it is possible to analyze the data recorded in the past and the data used in the present, and draw conclusions suitable for the future. [3]

Businesses based on data and data-based solutions make decisions based on data and analytics. Note that in today's data-driven world, the potential of data is not being used to its full potential, and this is a problem that most businesses face. Data analysis, including

conducting market and product research, identifying trends or other insights, is a critical aspect of and should be applied to help all organizations determine the impact on their business. There are five main stages in the data analytics process. Data preparation, data analysts, divide most of their time between data preparation and modeling tasks. Inaccurate information can lead to unreliable reports, resulting in a major negative impact on business decisions, such as loss of revenue, negative impact on business development, etc. can lead to

Data mining is the process of collecting, processing, organizing, and organizing data that can be used to analyze a dataset or build a machine learning model. This process involves searching, cleaning, and editing each data item in the database. The following steps are typically used in the data preparation process. [5]

Data collection: Data is collected from various sources to create a data set. These resources can include databases, APIs, web pages, and manually entered data.

Data cleaning: Data cleaning steps are used to detect and correct missing or incorrect information in the data. These steps include editing the data format, correcting the data, and correcting incorrect data entries.

Data preprocessing: In this step, the data must be fitted to machine learning models. These steps include selecting data properties, normalizing properties, and transforming data. [5]

Data partitioning: Data can be partitioned into training, validation and test datasets. This makes it easy to classify data that can be used to train, tune, and test the model. Data preparation is an important step in data analysis and can directly affect the results. Proper data preparation is essential for more accurate results and better performing machine learning models.[5]

Data Modeling: Data modeling refers to the description of data according to various structures and models. This allows data to be better managed, analyzed and used for specific purposes. [4]

Data modeling is mainly achieved through two activities.

Data documentation: This process involves organizing the data into groups such as grids or stacks, typically for visual representation. This process is essential for finding and gathering information necessary to perform data operations.

Creating a data model: This process refers to describing the data according to various structures and models. This makes data analysis and information disclosure much easier. Various technologies and programming languages are used to create this model. Data modeling is used by researchers, analysts, engineers and users from many different fields. Creating an optimal data model for the field of work facilitates the management of data in this field and accelerates analysis.

References:

1. Karimov S.Q. Information systems and data market. - Baku: Elm, 1999. - 300 p.

2. Karimov S.Q. information systems. Bakr: Elm, 2008.- 650 p

3. . Kopnees B.8., fapeeB A.O., Bacrums C.B., Pa x B.B. Ea:rr laurrrrx. l{Hre,l;rercrya:nxag o6pa6orxa iruQopuausr. M.: HoLralr. 2000.-352 c.

4. . Kepuuoa C.f Mera.lrassue a nHQcplrauzoHHLrx cHcreMa\ iil{H$opMarJr{orrHr,re rexnor'ror uH. Mocrga,,\l 5, 200-1. c. -l 7-42

5. Korneev N.V., Minitaeva A.M. Human-machine interface as the most important component of the architecture of promising ergatic control systems // Scientific notes of the RGSU. 2012. No. 7(107). pp. 93–97.