



Quantitative Data Analysis and Representation

Anupama Pal

School of Electronics Engineering, KIIT University, India

Abstract:

Data analysis is a process to inspect, clean, transform and model to retrieve important information, to support in decision making and to suggest conclusions. This paper focuses on quantitative data analysis and on the different methods to represent them in real life. Therefore, one can gain knowledge and implement it in their workplace.

Keywords: Data analysis, data representation, quantitative methods.

I. INTRODUCTION

Data analysis is a way in which one can obtain real market data from various fields as per requirement. This raw data is then converted into useful information to get a hassle free decision from the users. The collected data is analysed to solve different questions, refute theories or test thesis. There are number of iterative phases of data analysis. A data can be both numerical and categorical. The input data is an analysis directly from the customers who will be using the finished products. The type of entity to collect the required data is referred as an experimental unit. Some particular variables can even be specified and obtained from a population. This paper focuses on data analysis in varied scenarios. The required data can be obtained by information technology personnel in an organization. It can also be collected using sensors in the environment like recording equipments, traffic cameras, satellites and so on. Reading different documents, interviewing or downloading from online sources can even help us in this process. Moreover, this paper provides proper examples of the different methods in order to represent a clear view. To analyse the obtained data they are organized into rows and columns of a table format like within a statistical software or spreadsheet. After arranging the data, duplicate errors might exist. If there is problem in entering and storing data then data cleaning is very necessary. There are different kinds of data cleaning. To correct the mistyped words textual data spellcheckers are used. After all these procedures the data gets ready to be analysed.

II. BACKGROUND

In different domains the data analysis is referred with different names. It has number of approaches in case of science and various businesses. Data mining is a type of data analysis which is done to model or discover knowledge. Business intelligence means to aggregate and focus on the business information. In the statistical field data analysis is divided into three types. They are exploratory data analysis (EDA), descriptive statistics and confirmatory data analysis (CDA). Finding new features in the data is done by EDA whereas confirming existing thesis is focused by CDA. Predictive analysis helps in forecasting from statistical models whereas the text analytics uses statistics and structural techniques to get information from the textual sources. These are all different types of data analysis. Data integration is a forerunner of it and is directly related to data dissemination and visualization.

III. DATA ANALYSIS

The evolution of computer has increased the efficiency of *data analysis*. In this era it plays a really very important role for all types of companies and start-ups. The statistics helps government to plan new things for the betterment of the society. While the data analysis here helps in proper decision making. The RDB (relational database) and Sequel Query Language (SQL) help in collecting and analysing data. The process is easy, cheap and fast and therefore helps to do a lot with the data. The initialization of data warehouse on cloud has decreased the cost as well as made the work with big data hassle free [1]. The job of a data analyst is data collection and to produce better business decisions by analysing them. In order to analyse data first one should extract them. There are two ways of data collection, primary data collection methods and secondary data collection methods. The primary data collection is of two types' quantitative methods and qualitative methods.

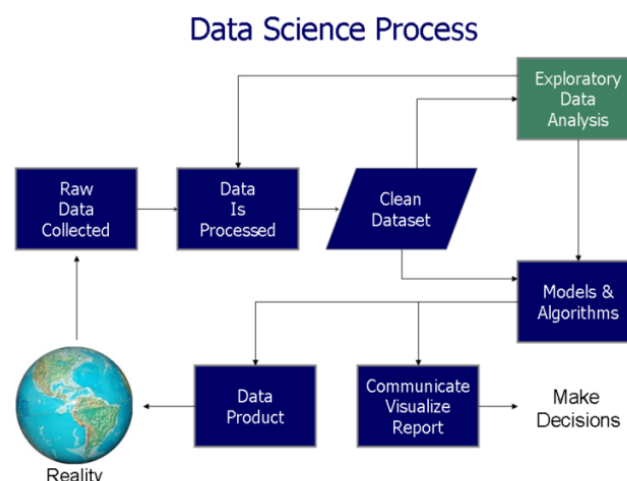


FIGURE.1. DATA ANALYSIS FLOW CHART

IV. QUALITATIVE DATA ANALYSIS

Qualitative data analysis helps to find relationships among different sets of data. There are tools like spreadsheet, statistical packages, MAXQDA, ATLAS. ti, Dedoose, QDA MINER and NVivo which supports data analysis. This kind of data analysis is done by different methods: Content analysis is a way to categorize verbal or behavioural data. The content is analysed as descriptive level and

interpretative level. In case of narrative analysis stories collected during an interview or said by people is presented in form of various context. Discourse analysis is a process to analyse a verbal interaction along with different types written texts. Framework analysis involves different stages like familiarization, to identify a thematic framework, coding, charting and finally mapping and interpretation. Grounded theory begins with an examination to frame a general statement over a hypothesis. An analyst then checks another case if it fits the statement. If it fits then the case is selected else another case is selected and same process continues until a case gets selected. To analyse a data an analyst can approach in two ways. One is inductive approach in which data are grouped into emergent framework and then analysed to get the link between them. The other one is deductive approach as the data is used to form group using the research questions and used especially when the resources and time is limited [2]. The qualitative data can be presented in presentations, handouts, slide documents, reports and so on. The qualitative data can be presented in presentations, handouts, slide documents, reports and so on.

V. QUANTITATIVE DATA ANALYSIS

Quantitative data collection focuses on surveys, statistics and measurement, pre-existing data and questionnaires. In this method, thorough study of research problems is necessary to control and address the accumulation of data or facts [3]. The frequency distribution table represents the frequency of different outcomes may be within a particular interval of time. The percentage distribution table is a way to express the frequency distribution. This is an easier way to compare various sets of data. Before analysing data it is necessary to consider the question scales and to compare the present result with previous one. Then the open and closed questions are encoded. Coding frames are formed. Now these coded data are scanned or need to be entered manually. In this case scanning is quicker and avoids any human error. Packages like SPSS Data Entry or Pinpoint, Snap decreases the scope of errors in the coded data entry. After the data entry it is cleaned in order to get rid of duplicate variables [4]. The mathematical ways of data collection are Mean, the sum of the data by number of data. It tells about the centre of distribution. Median, when a set of numbers is arranged in ascending order and there appears two middle numbers then the number that is in halfway is calculated. Mode is the numerical value which appears number of times in a set.

	Age	GPA	Gender	Hours	
1	Dick	20	1.9	M	1
2	Edward	19	1.5	M	1
3	Emmett	20	2.1	M	2
4	Lauren	20	2.4	F	3
5	Mike	19	2.75	M	4
6	Benjie	18	3	M	4
7	Joe	19	2.85	M	5
8	Larry	17	2.75	M	5
9	Rose	18	3.3	F	5
10	Bob	18	3.1	M	6
11	Kate	19	3.4	F	7
12	Sally	21	4	F	8
13	Sylvia	23	3.9	F	8
Sum	251	36.95			59
Mean	19.308	2.8423			4.5385
Variance	2.3974	0.5437			5.6026
Std Dev	1.5484	0.7374			2.367
Median	19	2.85			5

FIGURE. 2. MEAN, MEDIAN AND MODE

Some other useful numerical methods are Range, the difference between two extreme values of a set of data. Variance describes the variability of distribution. Standard Deviation is the index of variance and it tells the spread of the distribution. Higher the index value the data is more dispersed [5]. The correlation and regression is very important too. The correlation coefficient has no meaning unless it has direct connection and be symmetrical to the distributions of the variables. On the other hand, regression model is used to find the values in y-axis depending on the values in x-axis.

A. PRESENTATION OF FINDINGS

Frequency distribution table with quantitative data:

- Fasting blood glucose level in diabetics at the time of diagnosis

Fasting glucose level	No of diabetics		
	Male	Female	Total
120-129	8	4	12
130-139	4	4	8
140-149	6	4	10
150-159	5	5	10
160-169	9	6	15
170-179	9	9	18
180-189	3	2	5
	44	34	78

FIGURE.3. TABLE REPRESENTATION

One of the most effective ways to represent quantitative data is in this format. The table should be self-explanatory. The readers should easily understand the differences and the present trend.

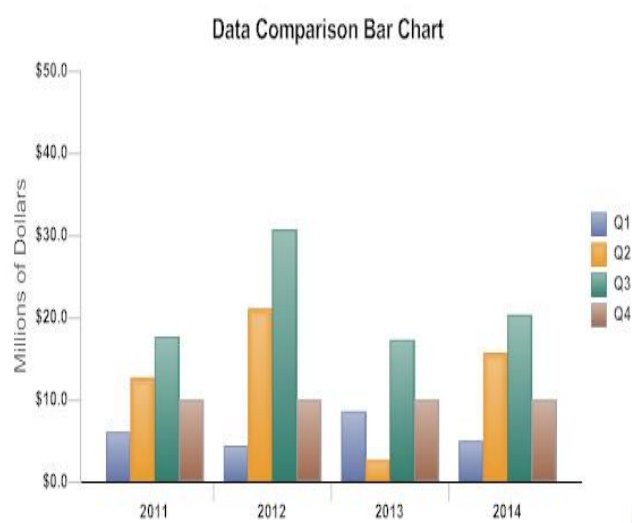


FIGURE.4. BAR GRAPH REPRESENTATION

This type of representation is friendly to access and allows easy comparison among sets of data. The only flaw here is it does not reflect the trend.

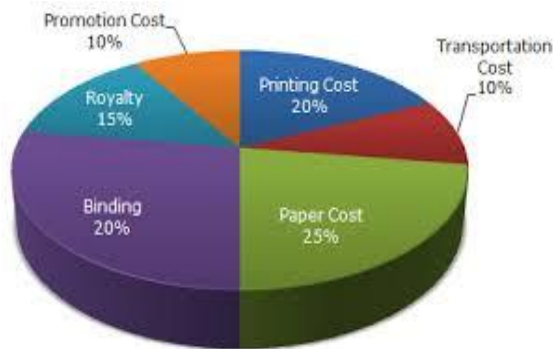


FIGURE. 5. CIRCLE GRAPH REPRESENTATION

The circle graph shows values in percentages and divides the total thing into parts. It fails to tell about the current trend.

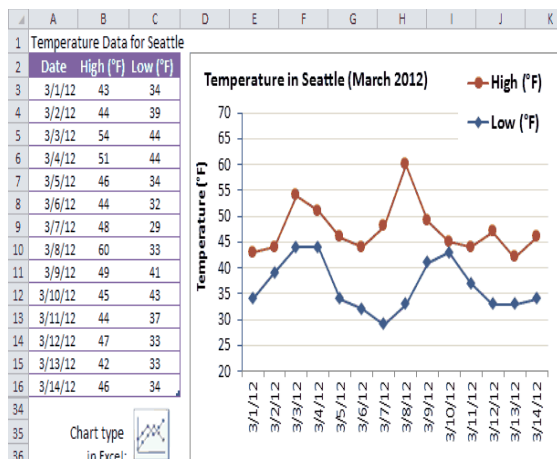
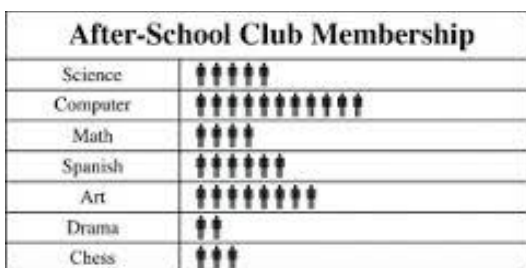


FIGURE.6. LINE GRAPH REPRESENTATION

This kind of graphs shows change with time and also gives an idea about trends. Various types of data cannot be used or compared in these cases.



Key: Each stands for 5 members

FIGURE. 7. PICTOGRAPH

This represents data using symbol or pictures so it is visually appealing and compares different sets of data.

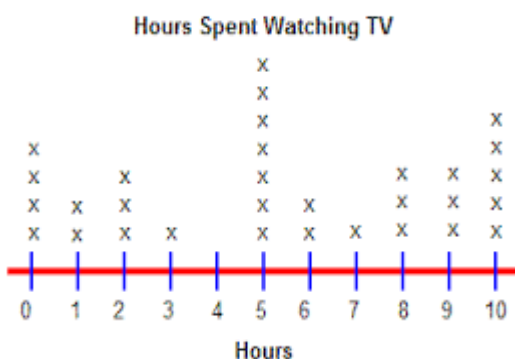


FIGURE.8. LINE PLOT GRAPH REPRESENTATION

Organizing data is easy using this type of graph. But this representation is not good to compare different categories of data.



FIGURE.9. VENN DIAGRAM

Here, circles symbolize the link among sets of data. Though it does not show trends, it is good in to compare and contrast [6]. Once the analysing and representing part is over it is the time to form the final report. The content of the quantitative research report may vary depending on the type of projects. The report must contain executive summary, background or introduction, methods, findings and ultimately conclusion and appendices.

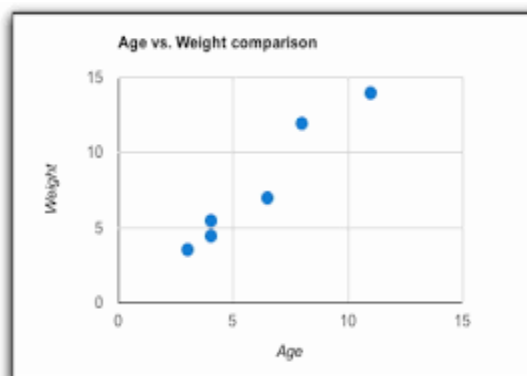


FIGURE.10. SCATTER DIAGRAM

This graph shows the relationship between two continuous variables. Number of sets of paired data is plotted and the pattern which is formed by them indicates the relation.

VI CONCLUSION AND FUTURE SCOPES

In order to represent the data it is necessary to use effective data presentation tools. Moreover it is necessary to work on a single idea at a time and in arranged manner. The variables should be well-explained. Growth of a country depends on business analytic. In this modern era, business analysis helps to take the future of a country to a different height but for this type of analysis data is handled. So data security is one of the major factors. Data analytic deals with these security issues. Presently the overall market is facing crisis of efficient and well-experienced data analysts. Many business organizations are opting this in order to understand the needs of their customers. The traditional methods are going obsolete as easy access of internet provides with many important data to work with. Moreover, the analysts not only help to interpret the collected data but also help in decision making. The way data analysis is expanding it shows that huge scope is available in upcoming few years.



FIGURE.11. GROWTH IN USE OF DATA ANALYSIS

This paper presents the different way one can extract and analyse the quantitative data from the raw informations. It even points out the different methods to represent them and these informations can help people to understand one of the trending term 'Data Analysis' and the future of the market.

VII. REFERENCES

- [1] Masashi Miyazaki. "*A Brief History of Data Analysis*" on 2015.
- [2] TilahunNigatu (MPH) M&E and Research Manager African Medical & Research Foundation. "Qualitative Data Analysis" on March 2009.
- [3] USC Libraries Research Guide by Babbie, Earl R. *The Practice of Social Research*. 12th ed. Belmont, CA: Wadsworth Cengage, 2010; Muijs, Daniel. *Doing Quantitative Research in Education with SPSS*. 2nd edition. London: SAGE Publications, 2010
- [4] "Analysing and reporting quantitative data" by Kirklees Council.
- [5] Asma Muhamad and FarhanaBintiYaakub. "Quantitative Data Analysis" on 2013.
- [6] Data Analysis by S.Hallum published on 2010.

VIII. ACKNOWLEDGEMENT

I extend my gratefulness to KIIT University for allowing me to use their resources throughout the research.