Research article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

Misleading and Misreading of Graphs in Research

Supanekar S. R.*

Department of Statistics, Balasaheb Desai College, Patan, Satara, (MS) India 415206
E-mail: sup_sanjay@yahoo.com

ABSTRACT

Statistical methods plays vital role in collecting, presenting, interpreting the data from research point of view. While using statistics at every stage of the research, it is very important to take care of it. Many times improper use of statistical methods leads to an inappropriate conclusions. This not only makes wastage of time and money but also leads non scientific information to research community. Hence it is need of time to discuss and elaborate about common mistakes committed by researchers and to provide remedy for the same. Present paper explains the errors in drawing and presenting graphs with some examples. The paper will make researchers aware about the proper meaning of the graphs to draw the right conclusions of research work.

KEYWORDS: Statistical method, Graphical representation, Misleading graphs.

***Corresponding Author**

Dr. S. R. Supanekar

Department of Statistics,

Balasaheb Desai College, Patan,

Satara, (MS) India 415206

Email: sup_sanjay@yahoo.com

1. INTRODUCTION

The importance of statistics in the modern age of computers and information technology has been highlighted in all disciplines. Statistics is used in many areas such as medical sciences, engineering, insurance, marketing, social sciences, sports, agriculture, education and so on. While using statistics in all these areas, the researcher uses various methods in statistics. But many researchers are very confused about the use of these methods and their fundamental concept. Due to this confusion, the researchers cannot properly communicate what they want to say. So, they get wrong conclusions. If the researchers have a good understanding of the Statistical concepts, then they will use it fairly. If a researcher uses a wrong concept in the research paper, then it is misleading the readers of the paper. Of course this does not always happen, but it is important to take care of it.

Ethics are very important when presenting data in statistics. The collected data may be true but if it is not presenting properly, then there is chance of wrong conclusions. If information is misrepresented then people are confused. Such confusion may be sometimes due to the lack of knowledge of the researcher, and sometimes researchers deliberately do. For example, some industrial organisations highlight the product quality to show that their product is good, but at the same time they purposely hide the shortcomings in that product. It indicates what kind of modern facilities they have in their product e.g. televisions. But at the same time does not show the average consumption of electricity. As a result, not all information about these televisions is revealed to the public. As a researcher, it is very important to bring to notice the whole information to the people. Therefore it is very important to have ethics in the researchers.

Some research papers have been published on the correct or inappropriate use of the statistics. Recent trends of statistics in medical journal have been studied and discussed by Altman[1]. The common errors committed while using statistics are explained and given the guidelines to avoid them[2]. Clarity of statistical concepts is very important. The ethical challenges in statistical research[3], misusage of statistics in medical research[4], misuse and abuse of statistics in biomedical research[5] and misleading of statistical studies[6] have been discussed and elaborated by various researchers. The researcher made common mistakes in use following statistical concepts

- Purpose of survey
- Sampling method and method of data collection
- Data classification, analysis and presentation
- Graphical representation
- Use of statistical software's
- Conclusions and inferences

In this paper we focused on graphical presentation and misleading of graphs, point out their errors, and attempt to reconstruct the same graph using more accurate forms of presentation.

2. GRAPHICAL REPRESENTATION:

Graphical representation is one of the most important tools in statistics. They are used almost everywhere. Using this method, you can arrange any complicated information in a simple manner and to understand it as a common people. Therefore, graphical presentation is very popular as a statistical method. It is very important to take care when drawing a graph. An author can mislead people by manipulating the graph, or they may misrepresent because of lack of information about this method. People currently use the computer software to draw graphs. But lack of knowledge about the concept of graph and software gives inaccurate and non reliable graphs. Generally, when you draw a graph, there is a possibility of mistakes in missing information, incomplete figure, vague caption etc.

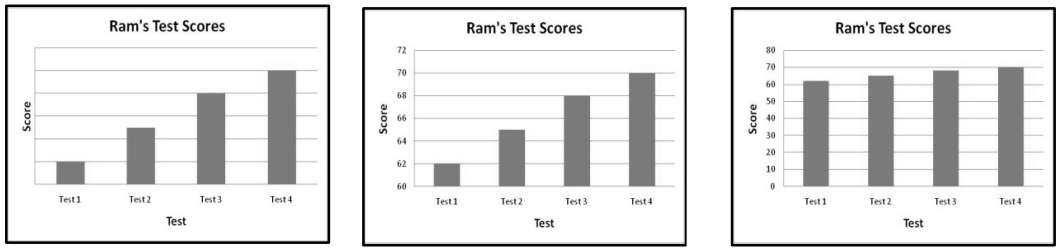
A good graph is a very powerful presentation tool. Using this tool, you can display very complicated data easily. Graph is a very effective tool for converting information into knowledge. There is a need to draw graph properly. While drawing a graph, the following thing needs to be addressed.

- Title of the graph
- Labels on both axes
- Source of the data
- Uniform size of a symbol in a graph
- Proper scale on both axis

3. EXAMPLES OF MISREADING AND MISLEADING OF GRAPHS

In this section, we have discussed how inappropriate graph misleads a researcher. How it really should be and what is the actual meaning of this graph. Some examples have been used to understand this.

i) Omitting or distorting the vertical Scale

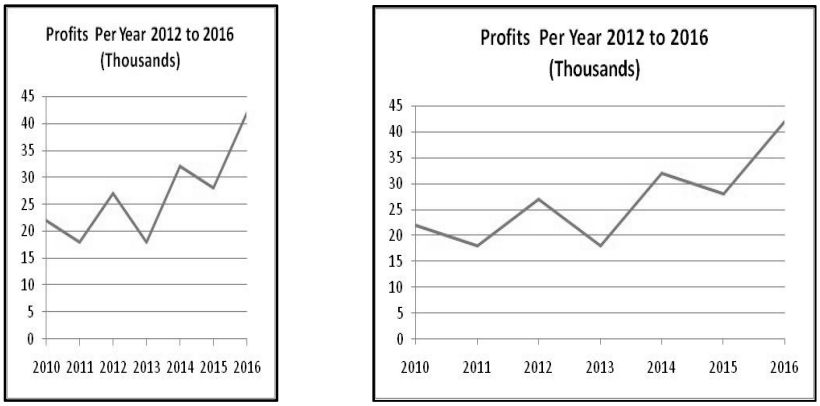


(a) **(b)** **(c)**

Fig. 1. Ram's test score using bar diagram (a) Without vertical scale, (b) Distorted vertical scale and (c) Proper vertical scale

Here first two graphs suggest that, there is a very good improvement in Ram's test score. Graph in Fig. 1(a) is without vertical scale and vertical scale in Fig. 1(b) is distorted. As a result, both graphs create a misleading impression. In fact, such type of graph has to be drawn using continuous scale and should start with zero, like graph shown in Fig. 1(c). This would show that Ram's test scores have improved, but not by as much as the first two graphs suggests. Ram's test scores have improved by a small amount over the last four tests. In this way, if we omit a scale or distorting the scale, then we get wrong conclusions.

ii) Distorting the Horizontal Scale:



(a) **(b)**

Fig. 2. Profit per year using scatter graph with (a) Compressed horizontal scale and (b) Proper horizontal scale

In the graph shown in Fig. 2(a), the horizontal scale is compressed. Actually the horizontal scale should be as shown in Fig. 2(b). As you can see that profit in Fig. 2(a), it is rising fast. But in reality as shown in Fig. 2(b), profit is not growing fast. Thus, the graph in Fig. 2(a) is misleading.

iii) Graph using insufficient sample size :

Table 1: Responses of opinion poll

	Responses	Percentage
Yes	4	80
No.	1	20
Total	5	100

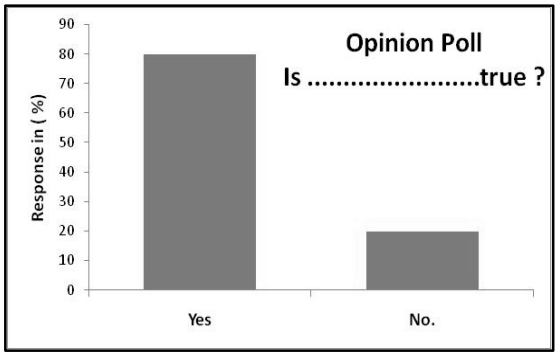
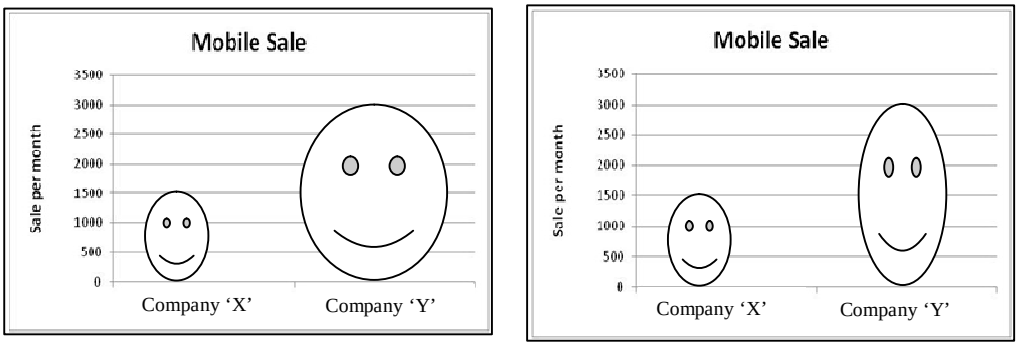


Fig. 3. Bar diagram showing opinion poll with insufficient sample size.

In Fig. 3, the opinion poll is shown in percentage. The number of people who vote 'Yes' in this graph is 80%. This number looks very big when viewing the graph. But in reality it is only four out of five people. It is wrong to make a general conclusion by only polling five people. In such surveys, sample size has immense importance. There must be enough sample size. This graph does not have enough sample size, so the results that appear in it are misleading.

iv) Distorting the Area :



(a) **(b)**

Fig. 4. Graph showing mobile sale with (a) Manipulated width of the diagram and (b) Proper width of the diagram.

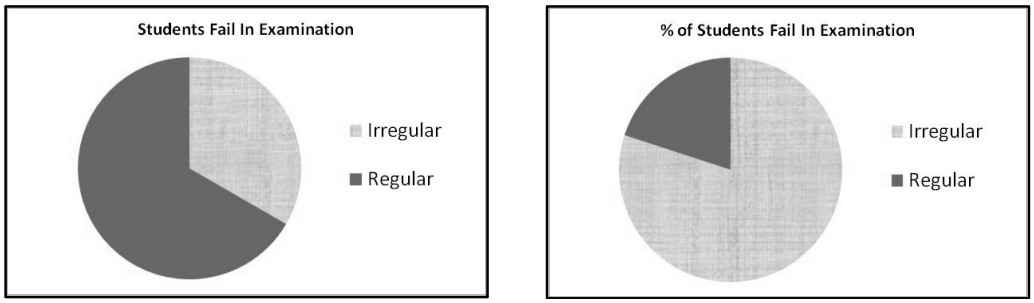
In these figures, two companies 'X' and 'Y' have shown their mobile sale for a month. The company 'X' sales 1500 mobile phones per month, while the company 'Y' sales 3,000 mobile phones per month. So the sale of 'Y' company is about twice the sales of the 'X' company. Actually

Fig. 4(a) and Fig. 4(b) have the same data, but in Fig. 4(a) the 'Y' company's diagram has shown a bigger width, so it seems that the 'Y' company is certainly more than its sales. This creates misunderstanding among people because the width of the figure is unnecessarily increased.

v) Use of Absolute Numbers Instead of Relative Percentages

Table 2: No. of student failed in examination

Students	Number of students	No. of student failed in examination	No. of student failed in examination in Percentage
Irregular in class	25	15	60
Regular in class	200	30	15



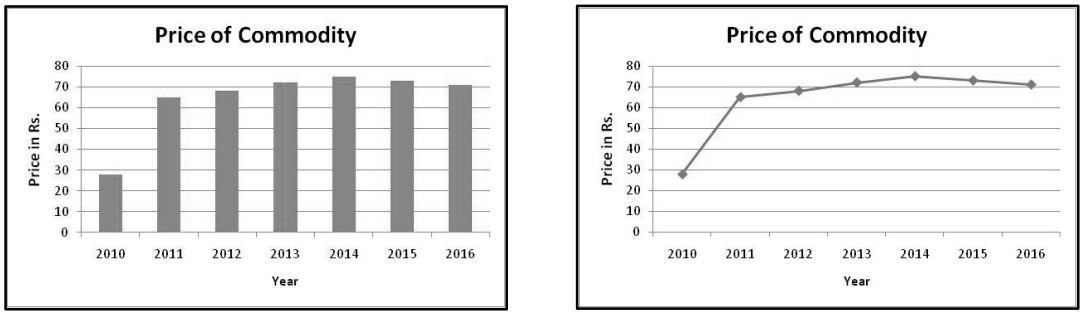
(a) **(b)**

Fig. 5. Pie chart showing (a) Incorrect presentation of fail and pass students and (b) Proper presentation of fail and pass students

Pie chart of students failing in the examination has been shown in this figure. Looking at the graph of Fig. 5(a), it seems that the failure rate is higher in regular student than an irregular student. This conclusion is surprising. Here, 30 out of 200 regular students are failed while 15 out of 25 irregular students are failing. While drawing a graph in Fig. 5(a), the total number of regular and irregular students is not considered. That is why this kind of wrong conclusion goes out. In order to draw such a graph, it is necessary to get a relative percentage rather than the absolute number. The graph in Fig. 5(b) shows the percentage of students who failed in the examination and therefore it is correct representation of the data.

vi) Selection of Proper Graph :

Researcher uses different type of graph to represent data. The proper choice of graph is depends on type of data available and the purpose to present the data. So selection of proper type of graph is very important. While comparing both of the graphs, the choice of graph in Fig. 6(b) is correct, because this graph clearly shows how commodity prices change.



(a) **(b)**

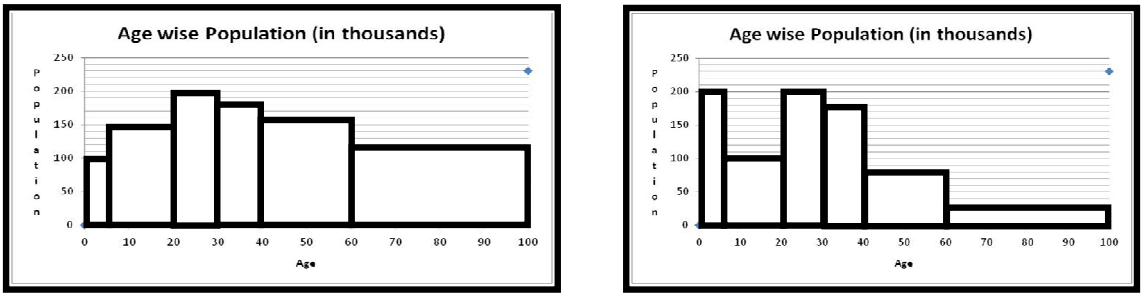
Fig. 6. (a) Improper selection of graph presenting price of commodity and (b) Proper selection of graph presenting price of commodity

vii) Improper use of Histogram :

To draw the histogram sometimes researcher not takes care of width of all classes. If data with unequal class width then it is necessary to adjust the class frequency but sometimes researcher not consider these things and draw the histogram as shown in Fig.7(a) and hence this histogram is misleading. When classes are different in the width then we need to adjust the class frequencies using different methods. One method is to use frequency density in place of class frequency. This has been shown in Fig. 7(b). When we plot the histogram one important rule is that "Area of rectangle is proportional to the frequency of the class". This rule is not used in graph shown in Fig. 7(a), so this graph is misleading.

Table 3: Age wise population

Age Group	Population (in thousands)	Frequency Density	Adjusted Frequency
00-05	100	100/ 5 = 20	200
05-20	150	150/15 = 10	100
20-30	200	200/10 = 20	200
30-40	180	180/10 = 18	180
40-60	160	160/20 = 8	80
60-100	120	120/40 = 3	30



(a) **(b)**

Fig. 7. (a) Histogram without taking care of class width and (b) Histogram with taking care of class width

4. Conclusion:

Many researchers do not use graphical representation properly. In this paper, some common mistakes made by researchers are discussed while drawing a graph. This paper highlights relevant examples of misleading and misreading of the graphs. Also we have tried to make the researchers aware of proper drawing and proper reading of the graph. As a result people can get the correct information and they will not be tricked.

REFERENCES

1. Altman D. G. Statistics in medical journals: Some recent trends. Stat Med 2000;19(23):3275–3289.
2. Good P. I, Hardin J. W. Common errors in statistics (and how to avoid them). John Wiley & Sons: New York 2003:107-125
3. Anca M., Mariana F. M., Paraschiva G. Ethical Challenges in Statistical Research. MegaByte 2011; MGBT-2011-1-ES-MD-543.
4. Ercan I., Yazıcı B., Yang Y. et al. Misusage of statistics in medical research. European Journal of General Medicine 2007; 4(3):128-134.
5. Thiese M. S., Arnold Z. C., and Walker S. D. The misuse and abuse of statistics in biomedical research. Biochem Med (Zagreb) 2015; 25(1): 5–11.
6. Wolf A. Misleading Statistical Studies. Policy Quarterly 2007; 3(4) :22-29.