

Statistics



Statistics, as a branch of mathematics, is the science of collecting, organizing and interpreting numerical facts. It involves presenting complicated masses of information in simplified yet detailed form. Statistics involve presenting summarized results of phenomenon taken from real life happening. The misuse of statistics has led to many to believe that the subject is hard which is not necessarily true.

The greatest numbers of consumers of statistics are the laymen who may not understand the terminologies used. Yet much of the statistical results in print today use these terminologies without offering clear expectations. It is therefore necessary to emphasize the interpretations drawn from statistical results than just presenting mere figures. This will remove the difficulties that are experienced in understanding statistics. That statistics is important in everyday living is an evident.

The application of statistics in research has resulted to great inventions and discoveries. Statistical results have enabled researchers to identify causal relationships between various factors. In planning, statistics have been used to forecast phenomenon and therefore to plan the available resources well. The study of the population has enabled governments to plan for their people. In the economic front, statistics have been used to measure economic growth and to understanding how an economy is affected by various factors.

In technology, statistics have been used to improve the quality of products. In manufacturing, statistics have been used to check the quality of the raw materials as well as that of the finished products. In the agricultural sector,

great improvements have been done based on statistical data that is collected and analyzed. Statistics have also been used in weather forecasting.

Practically all fields use statistics at one time or another. Statistics field is broadly divided into two classes; mathematical statistics and applied statistics. Mathematical statistics are concerned with the study of underlying mathematical theories while applied statistics is involved with the application of the statistical principles to real life situations. Applied statistics is further divided into descriptive and inferential statistics.

Inferential statistics involve the study of a sample and generalizing the results to the whole population. Descriptive statistics are involved with the representation of results from a studied group. Several ways of representing data have been developed such as diagrams, ratios, percentages among others. The choice of method is guided by the data itself. Mostly the data is represented in numerical or pictorial form. The usefulness of statistical results lies in the ability of the end users to draw meaningful conclusions from them.

Correct interpretations lead to a greater appreciation of the data presented and hence application of the findings. It is important that those who present these results provide as much information, not leaving the consumer guessing. The results should not just be numbers without explanations as to what they measure, how they are computed and what they mean. Other important elements to provide when giving statistical data are dates and the population studied. In addition it should be made easy for the reader to

follow these results by presenting the findings in an orderly manner with clarity. We shall critically evaluate how statistical results have been presented against the expected standards.

We shall evaluate whether the reports below have been complicated by the writers of the results or the results are just hard to understand. We shall begin with the report by Kelley and Twin, staff at CNN Money. The heading given is appropriate since it clearly informs the reader on what their article is about. The time frame is clear because it gives the results of the final day in a week hence the results are both weekly and a day's stocks performance.

Descriptive statistics are used and they indicate actual findings as calculated in a week's trading. Indices have been used to present the results. It has made easy for readers to follow by the inclusion of numerical figures. A rise in the indices is given in numerical form with the final value given.

It is easy to compute the initial value for further exploration of the data. This enables readers to follow up to determine the accuracy of the results. All the percentages in the report were found to be correct. Further, the presentation is also very orderly.

It starts with the overall performance of the stocks, goes to expectations in the following week and then explores individual company's analysis. The presentation is brief yet thorough; it does not leave unanswered questions. If there is an increase the value of the increment is stated. The second pack of statistical results is from an opinion poll carried out by Opinion Research Corporation.

The results are inferential statistics since they are a generalization of how the entire republican electorate may vote in general elections. A sample must have been studied. In these results some important information is missing. The sample size picked is not specified yet this is important. It enables the reader to quickly know the scope of the study. Further, if the percentages for favorability of the candidates are added, up they do not total 100% which leaves the reader wondering what the remaining percentage represented.

It is hard for readers to know what the missing percentage represents thereby raising more questions. If the rest did not have an opinion on the matter, this should be stated like so but not leaving some details hanging. (http://www.cnn.com/2007/POLITICS/03/12/giuliani_race/index.html).

A statistical term has been used which may not be clearly understood by laymen.

It would have been better understood if instead of just saying a sampling error of plus or minus 5% was allowed, the writers said there was only a five percent chance that the results could be wrong. That way the reader will find the results more meaningful. They are not required to have prior knowledge of what sampling error is and the effect it has on the results. In addition the results are not orderly; it is hard to compare one candidate with the others. A candidate's results are mentioned at the beginning, the other candidate's result somewhere in the middle of the text and another's at the end.

The accuracy of the results cannot be ascertained; no figures have been given as to how many voted for whom. It would have been better if a diagram like a chart was used to represent the data. This would be quick to understand for all readers since interpreting a chart does not require any statistical knowledge. The mode of data collection should have been included in the introductory part and the actual dates when the survey was done.

The data given in the results is, therefore, too scanty to be appreciated by readers. The third report by Isidore posted on CNN Money website on March 30 2007 is well presented. The results are inferential since a sample is studied and the results applied to the whole population. The heading almost totally explains what the data is about. It is brief yet informative. All the numerical figures posted are accurate.

Of importance to note here is explanation of the meaning of individual rates. The PCE deflator measures prices paid by customers for goods other than food and energy while the national savings rate compares consumers after tax income to their spending. This explanation is important since it gives the reader a clue as to its usefulness. We are, however, not told how many people are in the sample. It is always important to provide the data on sample size and if possible the sampling techniques used. This report has not used complex statistical terms therefore is easily understood by everyone.

The report is easy to follow and comparisons are possible since figures are given. The forecasted figures are also indicated making it easier for the readers to know the margin by which the forecasts were missed. An

example is given to help readers grasp what the national savings rate implies. (<http://money.cnn.com/2007/103/30/news/economy/income-spending/index.htm?posiversion=2007052410>).

In the houses sales report, all the calculations are confirmed to be accurate. It is also indicated that the results are for the month of June. There is provided much information which minimizes chances of more questions.

The scope of the study is clearly stated. New houses are separated from existing houses which increases the readers' ability to differentiate. The forecasts have been indicated to enable the reader know by how much actual results were off target. In presenting results, it is important to explain factors that caused the results. In the report, the low prices of houses were brought about by an influx of new houses. The report includes limitations of the study done.

The counting of sales at the beginning of the sales contract may render the results useless because not all contracts lead to actual selling. It is important therefore that results weaknesses are identified. These checks against making decisions while shortcomings are not included. The results suggest an annual decline in new home sales by 22.3% and supportive data is lacking to authenticate this.

The term median has been used and it may not be clear to non statisticians. The report is not orderly presented it leaves some details hanging, goes over them again in another perceptive. Such presentations easily loose the reader

as they try to comprehend what is going on. The report is a descriptive because it displays the outcome of the units counted.

The results do not come from the analysis of a sample but they are the results of a complete count of all the houses sold. The results are not used to make inferences about the whole population. The report on technology firms gaining value at opening is yet another report with a well stated heading. It provides the reader with an idea of what they are about to read. Most of the data is represented in charts which give quick understanding of the trend.

The type of statistics used in this case is the descriptive statistics. Indices have been used to represent more findings. The target population is singled out, the US market on a particular day. We've not been told the initial and final value of the day's activities.

In spite of good pictorial representation using charts, more information is needed on the indices. To new readers, it may not make sense to read of a 0.3% rise in composite index. They may want to know what those are.

They may also want to know what implications result from the rise or decline of the index value. There is no data to support either the indices accuracy or that of individual firms. The writers need to provide more information to allow for more in depth understanding of the results. The opinion poll report should allow comparisons, an element that is largely absent. It would be more appropriate to rank from the most favorite candidate.

In such a survey that is of great public interest, the writer should provide as much information but be brief. The data could also better be explained by trend lines to show the fluctuations of a candidate's favorability. The breakdown given by the report by Kelly should have been employed for the opinion poll's case. This could be achieved by breaking down results for individual candidates and plotting against time would have better comparisons. This would make it easy for readers to evaluate each candidate's performance with respect to time. The element lacking in Kelly's report and is present in income and spending report is the explanation of the indices.

The latter report enlists commodities included in the computations and an example is given to explain. The former report only states the increase or decrease in indices which may not be appreciated by new readers. It's imperative that explanations be offered where necessary to maintain the reader's interest. A mere presentation of charts in the technology firms report is not enough.

A brief explanation should be given as to the reasons that may have led to particular results. The causes of the decline are important considerations that should be made. Information about other non technological firms could also be mentioned and their performance indicated for readers to gauge whether the gain was significant. It is only one report that has included limitations and this is worrying. The weaknesses should be pointed out so that decisions may not be made on insufficient data.

In all statistical presentations of data, the writer should provide as much information as possible. There should also be a distinction of opinion from fact by first presenting facts. The diagrams used should be clearly labeled and any inferences drawn should be well explained. The statistical terms used should be explained to help non- statistician consumers understand the results. Of great emphasis is the need to present the results in an orderly manner.

Most writers rush over documents and end up presenting sketchy numerical figures that do not have much meaning. The writers are in fact the ones who make statistics hard to understand. They should therefore make sure that they do thorough work to present interesting results. The motivation should be to present clear and detailed information as opposed to brief and confusing results. There should be no hurry while summarizing statistical data because we risk complicating matters and not conveying the right information.