The Methods, Benefits and Problems of The Interpretation of Data

Mohaiminul Islam
School of Intelligent Technology and Engineering
Chongqing University of Science and Technology
Chongqing, China
Email: mohaiminul100@hotmail.com

Anower Hossain
School of Computer Science and Engineering
Chongqing University of Technology
Chongqing, China
Email: anowercse33@gmail.com

Rakibul Hasan
School of Intelligent Technology and Engineering
Chongqing University of Science and Technology
Chongqing, China
Email: rakibulhasan47374@gmail.com

Abu Rayhan Soton
School of Intelligent Technology and Engineering
Chongqing University of Science and Technology
Chongqing, China
Email: aburaihansoton@gmail.com

Abstract— **Data analysis and interpretation have now taken center stage with the advent of the digital age and the sheer amount of data can be frightening. In fact, a Digital Universe study found that the total data supply in 2012 was 2.8 trillion gigabytes! Based on that amount of data alone, it is clear the calling card of any successful enterprise in today's global world will be the ability to analyze complex data, produce actionable insights and adapt to new market needs all at the speed of thought. Business dashboards are the digital age tools for big data. Capable of displaying key performance indicators (KPIs) for both quantitative and qualitative data analyses, they are ideal for making the fast-paced and data-driven market decisions that push today's industry leaders to sustainable success. Through the art of streamlined visual communication, data dashboards permit businesses to engage in real-time and informed decision making, and are key instruments in data interpretation. This research article based on methods, benefits and problems of the interpretation of data.**

Keywords—**Interpretation, Data, Visualization, Statistics, Analysis**

I. INTRODUCTION

Data interpretation refers to the implementation of processes through which data is reviewed for the purpose of arriving at an informed conclusion. The interpretation of data assigns a meaning to the information analyzed and determines its signification and implications[1].

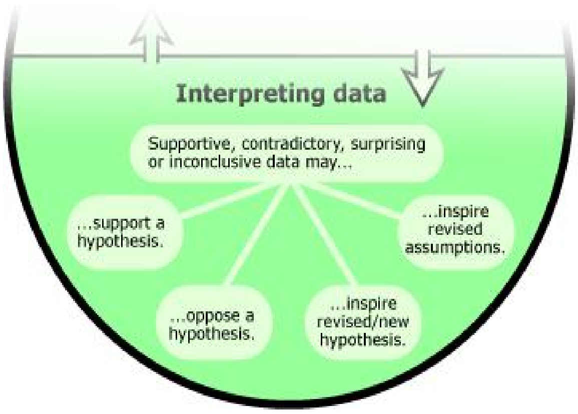

Fig. 1. Interpreting

The importance of data interpretation is evident and this is why it needs to be done properly. Data is very likely to arrive from multiple sources and has a tendency to enter the analysis process with haphazard ordering. Data analysis tends to be extremely subjective. That is to say, the nature and goal of interpretation will vary from business to business, likely correlating to the type of data being analyzed[2]. While there are several different types of processes that are implemented based on individual data nature, the two broadest and most common categories are "quantitative analysis" and "qualitative analysis".

Yet, before any serious data interpretation inquiry can begin, it should be understood that visual presentations of data findings are irrelevant unless a sound decision is made regarding scales of measurement. Before any serious data analysis can begin, the scale of measurement must be decided for the data as this will have a long-term impact on data interpretation ROI. The varying scales include:

- Nominal Scale: non-numeric categories that cannot be ranked or compared quantitatively. Variables are exclusive and exhaustive.
- Ordinal Scale: exclusive categories that are exclusive and exhaustive but with a logical order. Quality ratings and agreement ratings are examples of ordinal scales (i.e., good, very good, fair, etc., OR agree, strongly agree, disagree, etc.).
- Interval: a measurement scale where data is grouped into categories with orderly and equal distances between the categories. There is always an arbitrary zero point.
- Ratio: contains features of all three.

For a more in-depth review of scales of measurement, read our article on data analysis questions. Once scales of measurement have been selected, it is time to select which of the two broad interpretation processes will best suit your data needs[3]. Let's take a closer look at those specific data interpretation methods and possible data interpretation problems.

II. Process if Interpret Data

When interpreting data, an analyst must try to discern the differences between correlation, causation and coincidences, as well as many other bias – but he also has to consider all the factors involved that may have led to a result. There are various data interpretation methods one can use.

The interpretation of data is designed to help people make sense of numerical data that has been collected, analyzed and presented. Having a baseline method (or methods) for interpreting data will provide your analyst teams a structure and consistent foundation. Indeed, if several departments have different approaches to interpret the same data, while sharing the same goals, some mismatched objectives can result[4]. Disparate methods will lead to duplicated efforts, inconsistent solutions, wasted energy and inevitably – time and money. In this part, we will look at the two main methods of interpretation of data: with a qualitative and a quantitative analysis.

III. Qualitative Data Interpretation

Qualitative data analysis can be summed up in one word – categorical. With qualitative analysis, data is not described through numerical values or patterns, but through the use of descriptive context (i.e., text). Typically, narrative data is gathered by employing a wide variety of person-to-person techniques. These techniques include:

- Observations: detailing behavioral patterns that occur within an observation group. These patterns could be the amount of time spent in an activity, the type of activity and the method of communication employed.
- Documents: much like how patterns of behavior can be observed, different types of documentation resources can be coded and divided based on the type of material they contain.
- Interviews: one of the best collection methods for narrative data. Enquiry responses can be grouped by theme, topic or category. The interview approach allows for highly-focused data segmentation.

A key difference between qualitative and quantitative analysis is clearly noticeable in the interpretation stage. Qualitative data, as it is widely open to interpretation, must be “coded” so as to facilitate the grouping and labeling of data into identifiable themes. As person-to-person data collection techniques can often result in disputes pertaining to proper analysis, qualitative data analysis is often summarized through three basic principles: notice things, collect things, think about things[5].

If quantitative data interpretation could be summed up in one word (and it really can’t) that word would be “numerical.” There are few certainties when it comes to data analysis, but you can be sure that if the research you are engaging in has no numbers involved, it is not quantitative research. Quantitative analysis refers to a set of processes by which numerical data is analyzed. More often than not, it involves the use of statistical modeling such as standard deviation, mean and median. Let’s quickly review the most common statistical terms:

- Mean: a mean represents a numerical average for a set of responses. When dealing with a data set (or multiple data sets), a mean will represent a central value of a specific set of numbers. It is the sum of the values divided by the number of values within the data set. Other terms that can be used to describe the concept are arithmetic mean, average and mathematical expectation.
- Standard deviation: this is another statistical term commonly appearing in quantitative analysis. Standard deviation reveals the distribution of the responses around the mean. It describes the degree of consistency within the responses; together with the mean, it provides insight into data sets.
- Frequency distribution: this is a measurement gauging the rate of a response appearance within a data set. When using a survey, for example, frequency distribution has the capability of determining the number of times a specific ordinal scale response appears (i.e., agree, strongly agree, disagree, etc.). Frequency distribution is extremely keen in determining the degree of consensus among data points[6].

Typically, quantitative data is measured by visually presenting correlation tests between two or more variables of significance. Different processes can be used together or separately, and comparisons can be made to ultimately arrive at a conclusion. Other signature interpretation processes of quantitative data include:

- Regression analysis
- Cohort analysis
- Predictive and prescriptive analysis

IV. Importance of Data Interpretation

The purpose of collection and interpretation is to acquire useful and usable information and to make the most informed decisions possible. From businesses, to newlyweds researching their first home, data collection and interpretation provides limitless benefits for a wide range of institutions and individuals.

Data analysis and interpretation, regardless of method and qualitative/quantitative status, may include the following characteristics:

- Data identification and explanation
- Comparing and contrasting of data
- Identification of data outlines
- Future predictions

Data analysis and interpretation, in the end, helps improve processes and identify problems. It is difficult to grow and make dependable improvements without, at the very least, minimal data collection and interpretation. What is the key word? Dependable. Vague ideas regarding performance enhancement exist within all institutions and industries. Yet, without proper research and analysis, an idea is likely to remain in a stagnant state forever (i.e., minimal growth). So… what are a few of the business benefits of digital age data analysis and interpretation? Let’s take a look!

A. Informed decision-making

A decision is only as good as the knowledge that formed it. Informed data decision making has the potential to set industry leaders apart from the rest of the market pack. Studies have shown that companies in the top third of their industries are, on average, 5% more productive and 6% more profitable when implementing informed data decision-making processes. Most decisive actions will arise only after a problem has been identified or a goal defined.

Data analysis should include identification, thesis development and data collection followed by data communication[7].

If institutions only follow that simple order, one that we should all be familiar with from grade school science fairs, then they will be able to solve issues as they emerge in real time. Informed decision making has a tendency to be cyclical. This means there is really no end, and eventually, new questions and conditions arise within the process that need to be studied further. The monitoring of data results will inevitably return the process to the start with new data and sights.

B. Anticipating needs with trends identification

Data insights provide knowledge, and knowledge is power. The insights obtained from market and consumer data analyses have the ability to set trends for peers within similar market segments. A perfect example of how data analysis can impact trend prediction can be evidenced in the music identification application, Shazam. The application allows users to upload an audio clip of a song they like, but can't seem to identify. Users make 15 million song identifications a day. With this data, Shazam has been instrumental in predicting future popular artists.

When industry trends are identified, they can then serve a greater industry purpose. For example, the insights from Shazam's monitoring benefits not only Shazam in understanding how to meet consumer needs, but it grants music executives and record label companies an insight into the pop-culture scene of the day. Data gathering and interpretation processes can allow for industry-wide climate prediction and result in greater revenue streams across the market. For this reason, all institutions should follow the basic data cycle of collection, interpretation, decision making and monitoring[8].

C. Cost efficiency

Proper implementation of data analysis processes can provide businesses with profound cost advantages within their industries. A recent data study performed by Deloitte vividly demonstrates this in finding that data analysis ROI is driven by efficient cost reductions. Often, this benefit is overlooked because making money is typically viewed as "sexier" than saving money. Yet, sound data analyses have the ability to alert management to cost-reduction opportunities without any significant exertion of effort on the part of human capital.

A great example of the potential for cost efficiency through data analysis is Intel. Prior to 2012, Intel would conduct over 19,000 manufacturing function tests on their chips before they could be deemed acceptable for release. To cut costs and reduce test time, Intel implemented predictive data analyses. By using historic and current data, Intel now avoids testing each chip 19,000 times by focusing on specific and individual chip tests. After its implementation in 2012, Intel saved over $3 million in manufacturing costs. Cost reduction may not be as "sexy" as data profit, but as Intel proves, it is a benefit of data analysis that should not be neglected.

D. Clear foresight

Companies that collect and analyze their data gain better knowledge about themselves, their processes and performance. They can identify performance challenges when they arise and take action to overcome them[9]. Data interpretation through visual representations lets them process their findings faster and make better-informed decisions on the future of the company.

V. Data Interpretation Problems

The oft-repeated mantra of those who fear data advancements in the digital age is "big data equals big trouble." While that statement is not accurate, it is safe to say that certain data interpretation problems or "pitfalls" exist and can occur when analyzing data, especially at the speed of thought. Let's identify three of the most common data misinterpretation risks and shed some light on how they can be avoided.

Fig. 2. Interpretation Problems

A. Correlation mistaken for causation

Our first misinterpretation of data refers to the tendency of data analysts to mix the cause of a phenomenon with correlation. It is the assumption that because two actions occurred together, one caused the other. This is not accurate as actions can occur together absent a cause and effect relationship.

- Digital age example: assuming that increased revenue is the result of increased social media followers… there might a definitive correlation between the two, especially with today's muti-channel purchasing experiences. But, that does not mean an increase in followers is the direct cause of increased revenue. There could be both a common cause or an indirect causality.
- Remedy: attempt to eliminate the variable you believe to be causing the phenomenon.

B. Confirmation bias

Our second data interpretation problem occurs when you have a theory or hypothesis in mind, but are intent on only discovering data patterns that provide support, while rejecting those that do not.

- Digital age example: your boss asks you to analyze the success of a recent muti-platform social media marketing campaign. While analyzing the potential data variables from the campaign (one that you ran and believe performed well), you see that the share rate for Facebook posts were great, while the share rate for Twitter Tweets were not. Using only the Facebook posts to prove your hypothesis that the campaign was successful would be a perfect manifestation of confirmation bias[10].

- Remedy: as this pitfall is often based on subjective desires, one remedy would be to analyze data with a team of objective individuals. If this is not possible, another solution is to resist the urge to make a conclusion before data exploration has been completed. Remember to always try to disprove a hypothesis, not prove it.

C. Irrelevant data

The third and final data misinterpretation pitfall is especially important in the digital age. As large data is no longer centrally stored, and as it continues to be analyzed at the speed of thought, it is inevitable that analysts will focus on data that is irrelevant to the problem they are trying to correct.

- Digital age example: in attempting to gauge the success of an email lead generation campaign, you notice that the number of homepage views directly resulting from the campaign increased, but the number of monthly newsletter subscribers did not. Based on the number of homepage views, you decide the campaign was a success when really it generated zero leads.
- Remedy: proactively and clearly frame any data analysis variables and KPIs prior to engaging in a data review. If the metric you are using to measure the success of a lead generation campaign is newsletter subscribers, there is no need to review the number of homepage visits. Be sure to focus on the data variable that answers your question or solves your problem and not on irrelevant data.

VI. Interpretation of Data: The Use of Dashboards Bridging The Gap

As we have seen, quantitative and qualitative methods are distinct types of data analyses. Both offer a varying degree of return on investment (ROI) regarding data investigation, testing and decision-making. Because of their differences, it is important to understand how dashboards can be implemented to bridge the quantitative and qualitative information gap. Here are a few of the ways:

A. Connecting and blending data

With today's pace of innovation, it is no longer feasible (nor desirable) to have bulk data centrally located. As businesses continue to globalize and borders continue to dissolve, it will become increasingly important for businesses to possess the capability to run diverse data analyses absent the limitations of location. Data dashboards decentralize data without compromising on the necessary speed of thought while blending both quantitative and qualitative data. Whether you want to measure customer trends or organizational performance, you now have the capability to do both without the need for a singular selection.

B. Mobile Data

Related to the notion of "connected and blended data" is that of mobile data. In today's digital world, employees are spending less time at their desks and simultaneously increasing production. This is made possible by the fact that mobile solutions for analytical tools are no longer standalone. Today, mobile analysis applications seamlessly integrate with everyday business tools. In turn, both quantitative and qualitative data are now available on demand where they're needed, when they're needed and how they're needed.

C. Visualization

Data dashboards are merging the data gap between qualitative and quantitative methods of interpretation of data, through the science of visualization. Dashboard solutions come "out of the box" well-equipped to create easy-to-understand data demonstrations. Modern online data visualization tools provide a variety of color and filter patterns, encourage user interaction and are engineered to help enhance future trend predictability. All of these visual characteristics make for an easy transition among data methods – you only need to find the right types of data visualization to tell your data story the best way possible[11].

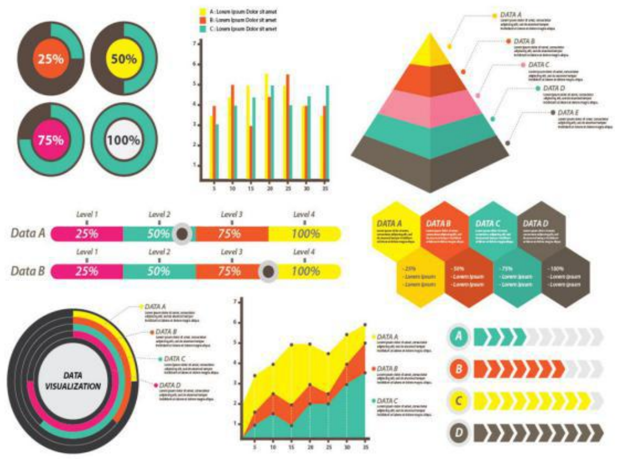

Fig. 3. Pic Chart (Visualize Report)

To give you an idea of how a market research dashboard fulfills the need of bridging quantitative and qualitative analysis, and helps in understanding how to interpret data in research thanks to visualization, have a look at the following one. It brings together both qualitative and quantitative data knowledgeably analyzed and visualizes it in a meaningful way that everyone can understand, thus empowering any viewer to interpret it[12].

VII. Data Interpretation Methods

Data analysis and interpretation are critical to develop sound conclusions and make better informed decisions. As we have seen all along this article, there is an art and science to the interpretation of data. Hereafter is a list-summary of how to interpret data and some tips:

- Collect your data and make it as clean as possible.
- Choose the type of analysis to perform: qualitative or quantitative, and apply the methods respectively to each.
- Qualitative analysis: observe, document and interview notice, collect and think about things.
- Quantitative analysis: you lead a research with a lot of numerical data to be analyzed through various statistical methods – mean, standard deviation or frequency distribution for instance[13].
- Take a step back: and think about your data from various perspectives, and what it means for various participants or actors of the project.
- Reflect on your own thinking and reasoning: and be aware of the many pitfalls data analysis and interpretation carries. Correlation versus causation, subjective bias, false information and inaccurate data, etc.

Conclusion

Data interpretation refers to the process of critiquing and determining the significance of important information, such as survey results, experimental findings, observations or narrative reports. Interpreting data is an important critical thinking skill that helps you comprehend text books, graphs and tables. Researchers use a similar but more meticulous process to gather, analyze and interpret data. Experimental scientists base their interpretations largely on objective data and statistical calculations. Social scientists interpret the results of written reports that are rich in descriptive detail but may be devoid of mathematical calculations.The importance of data interpretation is undeniable. Dashboards not only bridge the information gap between traditional data interpretation methods and technology, but they can help remedy and prevent the major pitfalls of interpretation. As a digital age solution, they combine the best of the past and the present to allow for informed decision making with maximum data interpretation ROI.

Acknowledgement

This work was supported by Chongqing science and technology innovation leading talent project.

References

[1] Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber. Bigtable: A Distributed Storage System for Structured Data. In OSDI, pages205–218, 2006.

[2] Rajeev Gupta, Himanshu Gupta, and Mukesh Mohania, "Cloud Computing and Big Data Analytics: What Is New from Databases Perspective?". S. Srinivasa and V. Bhatnagar (Eds.): BDA 2012, LNCS 7678, pp. Springer-Verlag Berlin Heidelberg 42–61, 2012.

[3] Curino, C., Jones, E.P.C., Popa, R.A., Malviya, N., Wu, E., Madden, S., Balakrishnan, H., Zeldovich, N.: Realtional Cloud: A Database-as-a-Service for the Cloud. In: Proceedings of Conference on Innovative Data Systems Research, CIDR- 2011.

[4] Alberto Ferandez, Sara del R, Victoria opez, Abdullah Bawakid, Maria J. del Jesus, Jose M. Benitez, and Francisco Herrera. "Big Data with Cloud Computing: an insight on the computingenvironment, MapReduce, and programming frameworks". doi: 10.1002/widm.1134. WIREs Data Mining Knowl Discov, 4:380–409, 2014.

[5] Lu, Huang, Ting-tin Hu, and Hai-shan Chen. "Research on Hadoop Cloud Computing Model and its Applications.". Hangzhou, China: 2012, pp. 59 – 63, 21-24 Oct. 2012.

[6] Wie, Jiang , Ravi V.T, and Agrawal G. "A Map-Reduce System with an Alternate API for Multi-core Environments.". Melbourne, VIC: 2010, pp. 84-93, 17-20 May. 2010

[7] K, Chitharanjan, and Kala Karun A. "A review on hadoop - HDFS infrastructure exten-sions.". JeJu Island: 2013, pp. 132-137, 11-12 Apr. 2013.

[8] F.C.P, Muhtaroglu, Demir S, Obali M, and Girgin C. "Business model canvas perspective on big data applications." Big Data, 2013 IEEE International Conference, Silicon Valley, CA, Oct 6-9, p. 32 – 37, 2013.

[9] Castelino, C., Gandhi, D., Narula, H. G., & Chokshi, N. H. (2014). Integration of Big Data and Cloud Computing. International Journal of Engineering Trends and Technology (IJETT), 100-102.

[10] Chandrashekar, R., Kala, M., & Mane, D. (2015). Integration of Big Data in Cloud compu-ting environments for enhanced data processing capabilities. International Journal of Engi-neering Research and General Science, 240- 245.

[11] James Kobielus, I., & Bob Marcus, E. S. (2014). Deploying Big Data Analytics Applica-tions to the Cloud: Roadmap for Success. Cloud Standards Customer Council

[12] M. Herland, T. M. Khoshgoftaar and R. Wald,A review of data miningusing big data in health informatics, Journal of Big Data, 1(2) (2014),pp. 1-35.

[13] X. Y.Chen and Z. G.Jin,Research on key technology and applicationsfor internet of things, Physics Procedia, 33, (2012), pp. 561-56