J. Electrical Systems 20-7s (2024): 2184-2196

Afnan Alabduljabbar[1]
Abdulaziz Alshammari[2]

The Effect of Data Quality on Decision-Making. A Quasi Experimental Study

Abstract: -The current study aimed to investigate the relationship between data quality dimensions (completeness and timeliness) and decision-making efficiency. The researcher adopted the quasi-experimental approach to answer the research questions and hypotheses. The study participants consist of 60 subjects, distributed into 2 groups. The first group consisted of 20 sales executives from Saudi beverage manufacturing companies. The other group had 40 participants from Al-Imam university. The study experiment consisted of 4 scenarios, two for each dimension, that give the participants scenarios and ask them to make the best decision based on these data. The scenarios were applied through face-to-face meeting and time to take decisions was recorded by participants. For the first data quality dimension, completeness of data, both groups got scenarios where complete and incomplete data were offered, and they were asked to choose the best available option based on the offered data. For the second dimension, the groups were offered up-to-date data and obsolete data and were asked to choose the best decisions based on the scenarios. The results were analyzed for correlations to check If there is a correlation between the responses of the two groups. The study found strong evidence for a correlational relationship between data quality dimensions and decision-making efficiency at 0.05 and 0.01 significance level for the student and employee groups respectively. The study found that high-quality data leads to both better and faster decisions in both groups. There were no significant differences between occupation, or gender and time taken for making a decision. The current study highlights the importance of data quality dimensions. It urges organizations to use up-to-date data, and complete data sets to base their decisions.

Keywords: Data Quality, Data quality dimensions, completeness dimensions, timeliness dimension, decision, making.

I. INTRODUCTION

1.1 Background

As a consequence of the advances in information technology and data storage technologies, companies face exponential growth in terms of the amount and diversity of data they have to manage. This means that not only the volumes of data increase, but also the elements to which data and information are associated [1]. For this reason, the poor quality of data is a factor that increasingly affects the performance of the organization since it can cause faulty decisions to be taken [1].

Data is a general term that refers to some or all of the facts, letters, symbols, and numbers referring to, or describing an object, idea, situation, condition, or other knowledge-related issues [2]. It constitutes a fundamental element for the objective decision making at all levels of an organization [2]. Furthermore, for a modern organization, data constitutes one of its strategic resources [3]. Currently there is great organizational interest in achieving what has been called "knowledge management" [4]. This implies taking the data generated in business processes and converting it into information through grouping and classifying it and, subsequently, converting this information into knowledge, through processes of separation, evaluation, comparison, etc. [4]. Therefore, without the existence of high-quality data, knowledge would never be obtained. On the other hand, the use of data as a basis for decision-making has been a widely recommended practice, as opposed to developing decision-making based on intuition [5]. In fact, one of the principles of quality management is the fact-based approach to decision making, which states that effective decisions are based on the analysis of data and information [6]. However, the existence of data or the will to form decisions based on it is not enough [5], as this data has to be of adequate quality. That is to say, when a decision is based on adequate data, then it becomes favoured than other decisions which are based on intuition [5].

1.2 Data quality Definition

The precise term "data quality" implies that the data that shall be used should meet the expectations of the data users [7]. These expectations are largely fulfilled when the data is useful for the purpose that users will utilize it for, and when it is easy to be understood and interpreted, and correct [7]. To guarantee these aspects, in principle,

[1] Afnan Alabduljabbar, Al Imam Mohammad Ibn Saudi Islamic University, College of Computer& Information Sciences, 442022539@sm.imamu.edu.sa
Riyadh, Saudi Arabia
[2] Abdulaziz Alshammari, Al Imam Mohammad Ibn Saudi Islamic University, College of Computer& Information Sciences, aashammari@imamu.edu.sa
Riyadh, Saudi Arabia

an appropriate design of the database, table or list of data must be made, in order to correctly define the attributes or types of data in it; and subsequently carry out an adequate design of the data production processes, guaranteeing that the data reaches the database or data table, free of defects and with the other desired characteristics [8].

Data quality is when the data conforms to the purpose for which it was intended. Data is also considered of high quality when it accurately represents real world concepts. Data is the cornerstone of a hierarchy that is built on top of it. Based on the data, comes the information, which is data placed in context. From actionable information comes knowledge, which becomes wisdom when applied. Poor data quality will result in poor information quality, and this goes up in the hierarchy, resulting in poor business decisions [8].

From these definitions it can be deduced that data quality is a relative concept. For example, data that one consumer may consider to be of acceptable quality might be of an unacceptable quality to another consumer with more stringent usage requirements or other intended uses [9]. Therefore, when the expectations of the users regarding the data vary, the characteristics that they must consider also vary. These characteristics or qualities that data must have to be considered adequate are called data quality dimensions [5]. This means that the quality of the data is associated with a set of dimensions or attributes that define it. A fundamental objective of the definition of the quality dimensions is having a common trait and also to focus on data quality problems and opportunities for improvement [4]. Among the most important dimensions for data quality are accuracy, integrity, consistency, and coherence [7].

1.3 Implications of Poor Data Quality

Poor data quality affects business management in various ways. Obviously, a primary impact of poor data quality is its effect on decision making [10]. Broadly speaking, poor quality data implies ineffective and ultimately inefficient decision-making processes. Inefficiency materializes in the fact that when the data is faulty, it usually leads to wrong decisions. Considering that the decisions that are made in the organization are related to a multitude of elements (customers, suppliers, products, work procedures, etc.), the impact produced by inadequate decision-making can be devastating. On the other hand, the inefficiency can be due to making decisions based on correct data, but with an additional cost in time due to the delay or lack of timeliness of the data [10]. Another very costly result of poor data quality is its negative impact on the organization's customers. This can materialize in customer dissatisfaction due to incorrect names, invoices with erroneous quantities, late shipment of products etc. These difficulties will likely result in the loss of customers, high customer dissatisfaction and difficulty attracting new potential customers [11].

The costs in time and other resources that organizations dedicate to the detection and correction of errors in the data is another consequence of the low quality of the data. In some manufacturing companies, a large part of the administrative staff and even part of the staff directly related to production dedicates a non-negligible percentage of their working time to correcting errors in the data [11]. On other occasions, the company's sales or customer service department incurs various costs by having to continually make corrections to customer addresses, orders, and invoices [11]. Other negative consequences of the inadequacy of the data, perhaps with a less economic consequence, are the employees' dissatisfaction, which is caused by the fact of having to continually correct erroneous data and the very fact of working in an environment where data is unreliable [10].

The effect of poor data quality on the success of many of the new decision support applications, such as data warehouses [12], data mining and data systems, customer relationship management, is also of considerable magnitude [13]. When bad results are obtained in a data production process, it is not always advisable to apply new technologies, since their effect could be zero or even negative. This situation is due to the fact that it is not convenient to automate inefficient processes [13].Despite all the problems discussed so far, in many organizations the quality of the data is not taken into account, a situation that is motivated by various factors, such as the large amounts of data that are generated in an organization, which makes it difficult to detect errors and correct them. Process because data is, unlike other resources, intangible, which makes it difficult to be measured. In addition, there are mnay difficulties when organizations try to quantify the improvements that result from processing the low quality data through quality improvement program (Redman, 2004).

The significance of data quality has also increased due to the different costs that have been associated with poor data quality problems, such as the cost of detecting and correcting data errors, customer dissatisfaction, employee dissatisfaction, wrong decisions and failures that occur in the application of new technologies [7]. Despite the fact that, at present, some alert has been created towards quality of the data, many organizations continue their operations disregarding the problem of data quality, which is fundamentally due to the intangibility of the data, an element that makes it difficult to measure its quality, as well as the failure to evaluate the expected benefits from correcting the data faults [7].

To deal with the issue of data quality, a wide range of computer products have emerged that provide great help in this regard [14]. The functionalities of these products range from the correction and detection of errors in the databases, to the prevention of data entry errors. Despite the advantages that information technology offers in this regard, to definitely improve the quality of the data, management techniques must be used to find the causes of the problems and even to improve the quality of the data. It is necessary to focus on matters not only related to the use of modern technology to improve the quality of the data but also by detecting the causes of these errors to evade their occurrences [14]. An important part of the initiatives for the guarantee of data quality are the so-called data quality evaluation systems and methodologies or diagnostic procedures, which aims at evaluating the current state of the data with which organizations make decisions and to discover the causes of quality problems in order to take measures to prevent or reduce their occurrence [14].

II. OBJECTIVES

The current thesis aims to investigate the relationship between data quality dimensions (completeness and timeliness) and decision-making efficiency. Another subset objectives are to investigate the relationship between data quality dimensions (completeness and timeliness) and decision-making efficiency measures such as time and ability to make predictions, to provide recommendations on how to present data to better support good decision-making, and to identify whether experience in executive positions is related to better decision making or not.

III. LITERATURE REVIEW

1. Empirical literature

A. Dimensions of Data Quality

The quality of data is instrumental in the managerial decision making, especially when the intended decisions have profound consequences on the organization. Hence, probing the data obtained and making it fulfill certain features are paramount for raising organizational productivity and performance. Alshikhi and Abdullah [15] present a roadmap on the definition of quality data, quality data dimensions and the association between the quality of data and the effectiveness of the decision-making. The dimensions of data quality according to this study are: accuracy which refers to the extent to which data can represent the reality; the integrity which means the consistency of the structure of the data and the associations that exist between the attributes and entities in the data; the consistency which means the precision of the data elements; the completeness which means the existence of the necessary data; the validity which means the data values that lie withing the required limits; the timeliness which means the availability of data when needed; and the accessibility which means the comprehensibility and usability of data. The study also argues that information about quality of data can impact decision making by enabling decision makers to utilize data more efficiently and effectively.

The ETL (extract, transform, load) process is highly relevant to the data quality field. Its implementation is carried out as a fundamental and basic part of data processing, especially in storage systems. Souibgui et al. [16] present a comparative study of some commercial ETL tools to show how much these tools consider data quality dimensions. The study found that data quality dimensions include completeness, accuracy, interpretability, representational consistency, and representational consciences.

B. Data quality and Decision Making in Public Sector

Organizations seeks ways through which they can take benefit of data to enhance the process of decision making. Janssen et al. [17] aimed to identify the factors that are related to big data that influence decision making. The researcher used qualitative methodology and introduced a case study to identify the influencing factors. The organization selected for this study was The Dutch Tax Organization as it was willing to share its practices and much information was available in public about it. The researcher collected data with varying quality degrees and created a big data chain. The study found that there were many factors which contribute to the decision-making quality, they include contractual governance, relational governance, the capabilities of BDA (big data analytics), knowledge exchange, collaboration, process integration and standardization, flexible infrastructure, staff, decision making quality and data quality of the big data sources. The study also found that the combination of data source quality and the data processing ability influences decision-making quality.

There is a quest to improve decision making quality, and special attention has been paid at new innovative ideas. Alkatheeri et al. [18] aimed at revealing the direct impact of big data quality on management information systems and decision-making quality. The researcher adopted a quantitative approach through recruiting 398 employees working at Dubai local government, data was analyzed using structural equation modelling (SEM). The study found that big data quality explains thirty-eight percent of the variance in management information systems. In addition, management information systems explain forty-five percent of the variance in decision making quality which denotes that the researchers noted that big data quality influence could positively affect management information system and could also lead to positive decision making.

Data quality can be of enormous importance in education. Hoogland et al. [19] aimed to identify prerequisites for the successful implementation of data-based decision making in the educational field. The study adopted the qualitative methodology and formed focus groups that consisted mainly of experts and practitioners that are involved in the educational process. The study found that the prerequisites for successful data use in education include teacher collaboration around the use of data, data literacy and leadership. The study also discussed the importance of the presence of Data-Based Decision-Making culture within a school to be successfully implemented in the classroom.

Data quality management is very crucial for improving policymaking (at the system level), management of educational institutions and pedagogical approaches in the classroom. Agasisti and Bowers [20] aimed to outline the significance of data usage for improving policymaking, management of educational institutions and pedagogical approaches in the classroom. The study argues that the traditional data analyses have been replaced with the new, sophisticated data analytics that help in taking decisions in the educational settings.

Shankar Narayanan [21] aimed to propose a decision-support framework that permits decision-makers to gauge quality both in an objective (context-independent) and in a context-dependent manner. The study found that the quality of the decision is dependent on the data precision and the same data can be viewed with two or more different quality lenses according to the decision task that it is used for.

C. Data quality and Decision Making in Manufacturing Sector

Data- Driven Decision Making (DDDM) is about collecting data about the organization's key performance indicators (KPIs) and transforming it into actionable insights. This process is fundamental in modern business strategies. Brynjolfsson and McElheran [22] aimed to investigate the adoption, performance, and organizational integration of DDDM in the manufacturing sector. The study collected data from US Census Bureau for 2005 and 2010. The study sample consisted of 18.000 manufacturing units. The results showed that using data to drive the decision-making process was more prevalent in bigger and older manufacturing units than in small-sized, modern-established manufacturing units. Although this, it was found that the late adopters of the DDDM had high efficiency in their decision-making processes more than the earlier adopters. It was also found that performance improved after the manufacturing units adopted DDDM, but not before, which denotes a causal relationship between decision making and data and the manufacturing industry.

Decision making quality refers to a firm's capability to take efficient decisions. It is considered a vital capability for any organization that aspires to develop and better position itself in society. Ghasemaghaei [23] aimed to identify the mediating role of knowledge sharing in regard of the impact that data analytics play in quality of decision making in organizations. It also worked to discover the role that data analytics play un improving the quality of decisions in organizations. The researcher collected data from 133 top and mid-level managers of US firms. The study found that knowledge sharing mediates the impact of data analytics on firm's decisions. It also indicated that knowledge sharing does not impact the decision-making process. Also, the study found that it is necessary to pay careful attention to the data analytics-based resources to make accurate and flawless decisions.

The Internet of Things (IoT) is about the connections between different devices that can communicate with each other. It enables the generation of data that can be utilized to gain insights about what can be done in the future. Brous et al. [24] aimed to probe the potentials for adoption of new data sources for decision-making in asset management organizations. The study followed the qualitative methodology with handling multiple case study approach. The first case study addressed how IoT data can be adopted for generating actionable knowledge about the maintenance of Dutch bridges. The case study denoted that the IoT provided the authorities with high quality data that was used to get insights about the damages at the bridge parts, which were harder to reach by traditional means. The second case study addressed utilizing traditional methods for decision making instead of using data generated by IoT. The traditional methods included using cloud-based service lane technology which yielded inferior results in regard to maintenance of bridges.

2. Theoretical Literature

A. What is Data Quality

The concept of Data Quality is vast, comprising different definitions and interpretations. DQ is essentially studied in two research communities: databases and management. The first one studies DQ from a technical point of view, while the second one is also concerned with other aspects or dimensions [25] Data quality management is critical in all databases and data Warehouses and subsequently in decision making and analytical environment, because any error in data quality means that the information subsequently extracted by the business system can lead to erroneous decisions [21]. Data quality became a concern to many organizations because of its huge volume. The access to information for the aim of taking decisions is not restricted to organizational boundaries or business units. As there are huge amounts of data that increases by day, decision makers became obliged to verify the quality of this data that they are going to take upon it decisions [21]. Gauging the quality of data used to be conducted by traditional methods without paying attention to the purpose in which this data will be used in. according to [21], the perceived quality of the data is usually influenced by the decision-task, and that the same data may be adequate in a situation and inadequate in another situation. Thus, we can say the decision- task is a variable that may define the adequateness and quality of the data.

B. Challenges of Data Quality:

Obtaining high quality data in an environment where different types of data structures converge requires overcoming certain difficulties because different sources, devices and / or formats converge in a way that may raise integration problems [26]. It is currently normal to find thousands of data that are not of high quality or that don't have any structures or have partial structures. This causes problems within quality systems because most analysis systems require manual data entry, so if the information is not fully conforming to a parameter, it can be exposed to human error [25] We can also find what is called dirty data. Dirty data is a serious quality problem, especially if we are working with data concentrated in the cloud. Dirty data can force organizational structures to bear real economic costs caused by programmed automatic actions that start with invalid data. Another challenge to data quality is unstructured data [27]. This type of data is always available, but it is hard to use or adapted. It doesn't fit in a spreadsheet with rows and columns, examples include videos, pictures, etc. so, it is necessary to invest time and resources to make the data compatible with the data analytics systems.

C. Assessment of Quality of Data

Data analysis, monitoring of performance measures, prediction of market behaviour and evaluation of production process are used by organizations to produce strategic and operational decisions. For any data analysis, the data has to be of moderate or high quality to produce helpful insights. So, there have to be methodologies that can assess the quality of data. According to Janssen [17] there are some factors that affect the decision-making quality and the data quality; these factors are process transformation and integration, development of skills, retaining experience and human resources, ensuring data quality, flexible systems, collaboration, knowledge exchange, decision-maker quality, building trust and managing relationships.

IV. METHODOLOGY

A. Study subjects:

The study subjects were assigned to two groups, one control group that consists of business executives that have more than 20 years of experience in the field of beverage sales, the other group contains students from Imam Mohamed bin Saudi University. The sampling technique used in recruiting subjects was the purposive sampling technique as students who studied sales in their university syllabuses were approached. The researchers gave the study subjects an explanation for the nature of the study. The entire experiment was conducted face-to-face to provide a more flexible interaction with study subjects and to record the time used for taking the decisions.

B. Quasi-experiment Design

The current study adopted quasi-experimental design to reveal the relationship between data quality and decision-making efficiency through collecting data from participants about specific scenarios where participants are provided with different quality-level data. According to Crano & Brewer [48] quasi experimental research involves controlling the independent variable to evaluate its casual impact. In the current study, variations to the data quality were given through providing the participants with incomplete data at some items and difficult-to understand data at other items. The current study entails intentional assignment of participants into 2 groups, that is why it is quasi-experimental design and not a true experiment as, according to Crano and Brewer [27] quasi-experiments are very similar to true experiments but with the difference that a random assignment of the participants is absent.

In regard to the type of the experiment, the current study follows the comparative experiment design. According to David and Sutton [49] the comparative design is used when the study wants to identify the similarities and differences between certain groups. For conducting comparative experiments, a number of principles should be followed such as confidentiality and anonymity. Confidentiality means that people other than the study team shouldn't be able to identify any of the participants. Another principle that is followed in the current study is getting informed consent from participants. Additionally, the researcher gave the subjects a sufficient amount of time to consider whether to participate in the study or not. Throughout the study, the researcher got insured that the subjects' data is kept private. After conducting the experiment, the data was preserved in a database that is locked from third parties to ensure the privacy of the study subjects.

The study groups are two, one control group that consists of business executives in the field of sales, the other group contains students from Imam Mohamed bin Saudi University. Both groups got incomplete data to decide upon in one situation as well as complete data to decide upon in another situation. The same applies to the timeliness of data dimension. The data was for virtual company's sales in the last 12 months. Participants were asked to make estimations based on the data offered to them. An explanation of the nature of the study was demonstrated to the participants. The entire experiment was implemented face to face to be able to reach all the participants and to calculate the time needed for answering the given scenarios.

C. Study Instrument

The experiment consists of different scenarios with modified data that were provided to the study subjects to take decisions based on them. Variations of data quality were conducted to allow for the measurement of the data quality impact on decision-making efficiency. Two data quality dimensions were included in the research: completeness, and timeliness as other data quality dimensions are difficult to measure in quantitative ways. The dimensions were implemented through providing the participants with complete and incomplete data to simulate different levels of data quality. Some of variations in the data given to the participants were obsolete (timeliness). To sum up, data was manipulated for the experimental group, as to check if data quality variations would generate better or worse decisions.

According to Samitsch [25] forecasting involves the assessment of uncertainty. Some techniques may improve the judgement in forecasting such as having recent data rather than old ones. It has been proved that recent data may allow participants to get better forecasts than old data due to the change in circumstances and the new inputs. As an example, participants were asked to forecast future demand for September 2024 based on data for 2021 one time and data for 2023 another time. Forecasting the exact same value as the optimal value would result in the highest profit within the scenario of the simulated inventory management system.

D. Experiment Procedure:

There were 2 groups, a control group, and experimental groups. The two groups were given the same scenarios but with different levels of data quality. The experiment was in-person to record the time for the decision of each situation with a first page that contain an invitation to participation in the experiment, the second page contains a brief explanation about what participants have to do. Participants assumed the role of operation manager in a beverage manufacturing company and order amount of beverage bottles that can best fit the future demand. The main objective of the task is to take a decision that generate the highest profit for the company. So, each participant must order the exact number of bottles of beverages because if he requested too many bottles, then he would have extra inventory costs for storage. If the participant asked few numbers of beverage bottles, then his sales would go down, with less profits than it is possible. So, the optimal situation for the participant is to request the exact amount of beverage bottles to generate the highest possible profits for the company. To put in easy terms, the experiment is about a simple inventory management system in which participants have to have the exact number of beverages bottles they need for sales.

Participants were asked to read the scenario, analyze the given data, and think about the optimal demand for the beverage bottles. The participants' answers were recorded for analysis. At the end of the simulation, the participants stated their age, occupation, gender, and any tools that they used to determine their answers (calculator, paper, and pen, etc.). The scenarios for the data quality and decision-making items are taken from Samitch [29] with minor modifications. These scenarios were sent to 3 sales managers with more than 20 years of experience to check the best decision/ answers for each scenario according to the Saudi Food and Beverage Market.

V. RESULTS

The objective of the study is to investigate the relationship between two dimensions of data quality (completeness and timeliness) and decision-making efficiency measures such as time needed for taking a decision and one's ability to make correct decisions.

Q1. Is there a correlational relationship between quality of data and efficient decision making?

			Efficient decision making of Employees Group		
			Incorrect	Correct	Total
Quality	Low Quality Data	Count	11	29	40
		% within Quality	27.5%	72.5%	100.0%
	High Quality Data	Count	5	35	40
		% within Quality	12.5%	87.5%	100.0%

Total	Count	16	64	80
	% within Quality	20.0%	80.0%	100.0%
Pearson Chi-Square = 2.813, df=1, p-value= 0.094				

Table 1. Association between quality of data and efficient decision making among employees group

Chi-Square test result showed in table 1 indicates that Pearson Chi square statistic (2.813) suggests a potential association between data quality and decision-making efficiency among employees group. However, the p-value (0.094) is slightly higher than the typical significance level of 0.05 but it is less than 0.10. This means that while the observed data deviates somewhat from the null hypothesis of no association, the evidence is not statistically strong enough to definitively conclude a significant correlation. While the percentage of employees making correct decisions is higher with high-quality data (87.5% vs. 72.5%), the difference is not substantial because the sample size is relatively small, which could be impacting the statistical power of the test and contributing to the higher p-value. As a conclusion, the chi-square test suggests a potential link between data quality and efficient decision-making, the evidence is not statistically conclusive. Further research with a larger sample size or alternative statistical methods could provide more definitive insights into this relationship.

			Efficient decision making of Students Group		
			Incorrect	Correct	Total
Quality	Low Quality Data	Count	41	39	80
		% within Quality	51.2%	48.8%	100.0%
	High Quality Data	Count	12	48	60
		% within Quality	20.0%	80.0%	100.0%
	Total	Count	53	87	140
		% within Quality	37.9%	62.1%	100.0%
Pearson Chi-Square = 14.232, df=1, p-value<0.001					

Table 2. Association between quality of data and efficient decision making among students group

Table 2 showed that chi-square statistic (14.232) is significantly higher than the previous analysis, indicating a stronger potential association between data quality and decision-making efficiency among students group. The p-value is also much smaller than the conventional significance level of 0.05, suggesting statistically significant evidence against the null hypothesis of no association. This means we can confidently conclude that there is a significant correlational relationship between data quality and decision-making efficiency among students group. The percentage of students making correct decisions is considerably higher with high-quality data (80% vs. 48.8%). The direction of the relationship suggests that higher data quality is associated with an increase in decision-making efficiency.

Q2. Is there a correlational relationship between the quality of data and time used for taking a decision based on it?

			Time of Decision (Seconds) for Employees Group						Total
			30 Seconds	45 Seconds	60 Seconds	75 Seconds	90 Seconds	Greater than 90 Seconds	
Quality	Low Quality Data	Count	1	11	15	2	8	3	40
		% within Quality	2.5%	27.5%	37.5%	5.0%	20.0%	7.5%	100.0%
	High Quality Data	Count	3	12	10	6	5	4	40
		% within Quality	7.5%	30.0%	25.0%	15.0%	12.5%	10.0%	100.0%
	Total	Count	4	23	25	8	13	7	80
		% within Quality	5.0%	28.7%	31.3%	10.0%	16.3%	8.8%	100.0%
Pearson Chi-Square = 4.879, df=5, p-value=0.431									

Table 3. Association between the quality of data and time used for taking a decision based on it among employees group

Table 3 shows that chi-square statistic was (4.879) with p-value (0.431) which is greater than the significance level of 0.05, indicating that there is no statistically significant evidence to conclude a correlational relationship between data quality and decision-making time among employees. While the percentage of employees making quick decisions (30 seconds) tends to be higher with high-quality data (7.5% vs. 2.5%), the distribution across other time categories is not substantially different between the two quality levels. This suggests that data quality may not significantly impact decision-making time for employees.

			Time of Decision (Seconds) for Students Group						Total
			30 Seconds	45 Seconds	60 Seconds	75 Seconds	90 Seconds	Greater than 90 Seconds	
Quality	Low Quality Data	Count	4	21	20	7	21	7	80
		% within Quality	5.0%	26.3%	25.0%	8.8%	26.3%	8.8%	100.0%
	High Quality Data	Count	5	12	13	21	17	12	80
		% within Quality	6.3%	15.0%	16.3%	26.3%	21.3%	15.0%	100.0%
	Total	Count	9	33	33	28	38	19	160
		% within Quality	5.6%	20.6%	20.6%	17.5%	23.8%	11.9%	100.0%
Pearson Chi-Square = 12.787, df=5, p-value=0.025									

Table 4. Association between the quality of data and time used for taking a decision based on it among students group

The above information about the student group suggests a different picture compared to the analysis of the employee group. The chi-square statistic (12.787) is significant with p-value (0.025) less than the significance level of 0.05. This suggests statistically significant evidence for a potential association between data quality and decision-making time among students, the percentage of students making quick decisions (30 seconds) is slightly higher with high-quality data (6.3% vs. 5.0%). Moreover, the distribution across other time categories also shows differences, with high-quality data associated with a higher proportion of decisions made in the 75 second (26.3% vs. 8.8%), while in the 90 second (21.3% vs. 26.3%). This data suggests that data quality might indeed be linked to decision-making time for students, with high-quality data potentially leading to quicker and more efficient decision-making compared to low-quality data.
Conclusion:
Based on this analysis, we cannot establish a statistically significant relationship between data quality and decision-making time among students.

Q3. Is there a correlational relationship between gender and efficient decision making?

			Gender of Employees Group		Total
			Male	Female	
Efficient decision making of Employees Group	Incorrect	Count	7	1	8
		% within Efficient decision making of Employees Group	87.5%	12.5%	100.0%
	Correct	Count	10	2	12
		% within Efficient decision making of Employees Group	83.3%	16.7%	100.0%
Total		Count	17	3	20
		% within Efficient decision making of Employees Group	85.0%	15.0%	100.0%
Pearson Chi-Square = 0.065, df=1, p-value=0.798					

Table 5. Association between gender and efficient decision making among employees group

The chi-square statistic (0.065) is very low, indicating almost no difference between the observed and expected frequencies of decision-making efficiency for both genders. The p-value (0.798) is significantly higher than the conventional significance level of 0.05, further supporting the lack of evidence for a statistically significant relationship. The sample size (20) is relatively small, which could limit the statistical power of the test and contribute to the non-significant result. Based on this analysis, we cannot conclude a statistically significant association between gender and efficient decision-making among employees.

			Gender of Students Group		Total
			Male	Female	
Efficient decision making of Students Group	Incorrect	Count	9	12	21
		% within Efficient decision making of Students Group	42.9%	57.1%	100.0%
	Correct	Count	10	8	18
		% within Efficient decision making of Students Group	55.6%	44.4%	100.0%
Total		Count	19	20	39
		% within Efficient decision making of Students Group	48.7%	51.3%	100.0%
Pearson Chi-Square = 0.626, df=1, p-value=0.429					

Table 6. Association between gender and efficient decision making among students group

Similar to the employee data, the chi-square statistic (0.626) for students is quite low, indicating no significant difference in the observed and expected frequencies of decision-making efficiency according to gender. The p-value (0.429) is also greater than the conventional significance level of 0.05, further corroborating the lack of evidence for a statistically significant relationship between gender and efficient decision-making in this case. Unlike the employee data, where both genders had similar proportions of correct and incorrect decisions, the student data shows a slightly higher percentage of female students making incorrect decisions (57.1% vs. 42.9%). However, even this difference is not statistically significant due to the low chi-square value and high p-value.

Conclusion:

Consistent with the employee data, this analysis offers no compelling evidence for a statistically significant relationship between gender and efficient decision-making among the student group. While a slight imbalance in incorrect decisions favoring females was observed, it falls within the realm of chance based on the statistical test results. The results from both groups of employees and students suggest that within these contexts, gender doesn't play a significant role in overall decision-making efficiency.

Q4. Is there a correlational relationship between gender and time to take an efficient decision?

			Gender of Employees Group		Total
			Male	Female	
Time of Decision (Seconds) for Employees Group	30 Seconds	Count	0	1	1
		% within Time of Decision (Seconds) for Employees Group	0.0%	100.0%	100.0%
	45 Seconds	Count	4	0	4
		% within Time of Decision (Seconds) for Employees Group	100.0%	0.0%	100.0%
	60 Seconds	Count	7	1	8
		% within Time of Decision (Seconds) for Employees Group	87.5%	12.5%	100.0%
	75 Seconds	Count	2	0	2
		% within Time of Decision (Seconds) for Employees Group	100.0%	0.0%	100.0%
	90 Seconds	Count	2	1	3
		% within Time of Decision (Seconds) for Employees Group	66.7%	33.3%	100.0%
	Greater than 90 Seconds	Count	2	0	2
		% within Time of Decision (Seconds) for Employees Group	100.0%	0.0%	100.0%
Total		Count	17	3	20
		% within Time of Decision (Seconds) for Employees Group	85.0%	15.0%	100.0%
Pearson Chi-Square = 7.908, df=1, p-value=0.161					

Table 7. Association between gender and time to take an efficient decision ng among employees group

The chi-square statistic (7.908) suggests a potential difference between the observed and expected distributions of decision-making time across genders. However, the p-value (0.161), while slightly lower than the conventional significance level of 0.05, does not definitively reach that threshold. This leads to two possible interpretations. While females in the table appear to take longer for some time categories (e.g., 45 seconds), they also make some quicker decisions (e.g., 30 seconds). The overall difference in average decision-making time between genders might be minimal and statistically insignificant based on the p-value.

			Gender of Students Group		Total
			Male	Female	
Time of Decision (Seconds) for Students Group	30 Seconds	Count	1	0	1
		% within Time of Decision (Seconds) for Students Group	100.0%	0.0%	100.0%
	45 Seconds	Count	4	8	12
		% within Time of Decision (Seconds) for Students Group	33.3%	66.7%	100.0%
	60 Seconds	Count	7	3	10
		% within Time of Decision (Seconds) for Students Group	70.0%	30.0%	100.0%
	75 Seconds	Count	1	1	2
		% within Time of Decision (Seconds) for Students Group	50.0%	50.0%	100.0%
	90 Seconds	Count	5	6	11
		% within Time of Decision (Seconds) for Students Group	45.5%	54.5%	100.0%
	Greater than 90 Seconds	Count	1	2	3
		% within Time of Decision (Seconds) for Students Group	33.3%	66.7%	100.0%
Total		Count	19	20	39
		% within Time of Decision (Seconds) for Students Group	48.7%	51.3%	100.0%
Pearson Chi-Square = 4.335, df=1, p-value=0.502					

Table 8. Association between gender and time to take an efficient decision ng among students group

Similar to the employee data, the chi-square statistic (4.335) for students is relatively low, suggesting no significant difference in the observed and expected distributions of decision-making time across genders. The p-value (0.502) is well above the conventional significance level of 0.05, further supporting the lack of evidence for a statistically significant relationship. While there are some differences in the percentage of students making decisions within each time category, none of them are substantial or consistent across categories. Both male and female students seem to utilize the entire range of decision-making times, from quick (30 seconds) to longer (over 90 seconds).

VI. DISCUSSION

The current study aimed to identify whether data quality impacts the decision-making efficiency or not. The study found that there is a significant relationship between data quality and decision-making efficiency. The experiment yielded evidence that participants can take better decisions with high-quality data. This is consistent with Samitsch [29] which found that data quality impacts decision making. Previous research also found that the accuracy of information plays a vital role in the decision-making process [5]. The current study found that, obviously, low quality data implies ineffective and, ultimately, inefficient decision-making processes. This inefficiency materializes when the data is incomplete that makes the individual take erroneous decisions. Considering that the decisions made are related to a multitude of elements (customers, suppliers, products, work procedures) the impact produced by inadequate decision-making can also be deduced.

The process of making business decisions is driven by data. However, having a lot of data doesn't necessarily mean making a lot of good decisions. In this case, as in so many others, quality is above quantity. We know that big data impacts decision making but the current study proved that data quality, of moderate sizes and volumes, can also impact decision making. The arrival of big data caused an extraordinary technological revolution that significantly impacted all spheres of our lives in ways that we now consider common [17].

The current results found that there was no significant relationship between decision making and the time spent taking the decision. This means that the duration of time to take a decision do not correlate with the correctness or incorrectness of the decision, which reiterate the importance of data quality as having a fundamental role in the correctness/ incorrectness of the decisions being taken [16]. This means that duration of time is not related to the decision accuracy and that data quality is the only variable that may have an impact in this regard. This is contrary to what Smitsch [29] has found as the latter found a correlation between the time the individual takes to make a decision and the decision correctness or incorrectness.

The study found that decision making accuracy is not correlated to gender. The results indicate that there is almost no difference between the observed and expected frequencies of decision-making efficiency based on gender. Based on this analysis, we cannot conclude a statistically significant association between gender and efficient decision-making among employees. This is consistent with Gebre et al. [30] which found no impact for gender to decision making efficiency. This can be due largely to the bridging of the gap that used to exist between male and female in knowledge disciplines.

The current research also found no significant differences between males and females in regard of time taken for making a decision. This lack of differences among the different sexes denotes the homogeneity of the study sample and having similar characteristics in regard to taking decisions based on time constraints.

CONCLUSION

Organizations store, manage and use large volumes every day. If the data does not meet its objective, it is considered to be of poor quality. This definition of data quality implies that its meaning differs depending on the organization to which it belongs and the purpose it serves. Business leaders no longer rely on assumptions but instead use business intelligence techniques to make better decisions. This is where good data quality can allow for precision in decision making while poor data quality can bias the results of data analysis, leading organizations to base crucial decisions on incorrect forecasts. It is a common belief that when managing data quality throughout the organization, we must obtain the approval and support of decision makers.

Ensuring data quality is not easy. It is clear that accurate, up-to-date, and complete data would be desirable but, unfortunately, the real world is far from ideal. Achieving high-quality data requires having a very clear understanding of its meaning, context and intention, where ambiguities do not exist and, if possible, there are standardized definitions that can serve as a basis for future data- based decisions.

Implications

Ensuring data quality in the organization should not be considered a one-time action. Planning continuous improvement based on iterations is the most effective approach and the one that can bring the business closest to success, in terms of data quality. Having complete and up-to-date data would help organizations to take informed decisions which would promote their productivity and improve their profitability. In addition, the current study highlights the importance of the data quality dimensions. It urges organizations to use up-to-date data, and complete data sets to base their decisions.

Recommendations

Currently, computer technology is widely used with the objective of improving organizational performance in terms of data quality. Based on this organizations can develop a wide range of software solutions for decision-making purposes. In addition, organizations can build on the results of the current research to improve their productivity and performance. This necessitates preserving high quality data and investing in data analytical tools.

Limitations:

The first limitation is the lack of representativeness for the subjects enrolled in the study. Although most experimental studies use less than 100 subjects for their work, the current study enrolled students from one faculty, which limits the generalizability of its results.

Future Research:

Other dimensions of data quality can be investigated to identify its impact on decision making. In addition, utilizing similar the experiment on another set of subjects can help to give another subordination to the current results.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University for funding and supporting this work through Graduate Students Research Support Program.

REFERENCES

[1] Taleb, I., Serhani, M. A., Bouhaddioui, C., & Dssouli, R. (2021). Big data quality framework: a holistic approach to continuous quality management. Journal of Big Data, 8(1), 1-41.

[2] Fan, W., & Geerts, F. (2022). Foundations of data quality management. Springer Nature.

[3] Davoudi, S., Dooling, J. A., Glondys, B., Jones, T. D., Kadlec, L., Overgaard, S. M., ... & Wendicke, A. (2015). Data quality management model (2015 update)-retired. Journal of AHIMA, 86(10), expanded-web.

[4] Mahrinasari, M. S., Hussain, S., Yapanto, L. M., Esquivel-Infantes, S. M., Untari, D. T., Yusriadi, Y., & Diah, A. (2021). The impact of decision-making models and knowledge management practices on performance. Academy of Strategic Management Journal, 20, 1-13.

[5] Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. Information & Management, 56(6), 103135.

[6] Jeanne Ketola and Kathy Roberts. Retrieved at https://www.qualitydigest.com/april00/html/iso2k.html

[7] McGilvray, D. (2021). Executing data quality projects: Ten steps to quality data and trusted information (TM). Academic Press.

[8] Fan, W., & Geerts, F. (2022). Foundations of data quality management. Springer Nature.

[9] Cichy, C., & Rass, S. (2019). An overview of data quality frameworks. IEEE Access, 7, 24634-24648.

[10] Rego, A. L. (2020). A framework to determine the impact of poor data quality on the reliability of iot sensor-based real-time decision-making (Doctoral dissertation, University of Arkansas at Little Rock).

[11] Ghasemaghaei, M., & Calic, G. (2019). Can big data improve firm decision quality? The role of data quality and data diagnosticity. Decision Support Systems, 120, 38-49.

[12] Caro, A., Calero, C., Caballero, I., & Piattini, M. (2006). Defining a data quality model for web portals. In Web Information Systems–WISE 2006: 7th International Conference on Web Information Systems Engineering, Wuhan, China, October 23-26, 2006. Proceedings 7 (pp. 363-374). Springer Berlin Heidelberg.

[13] Olson, K. (2013). Do non-response follow-ups improve or reduce data quality?: a review of the existing literature. Journal of the Royal Statistical Society Series A: Statistics in Society, 176(1), 129-145.

[14] Cong, G., Fan, W., Geerts, F., Jia, X., & Ma, S. (2007, September). Improving data quality: Consistency and accuracy. In Proceedings of the 33rd international conference on Very large data bases (pp. 315-326).

[15] Alshikhi, O. A., & Abdullah, B. M. (2018). Information quality: definitions, measurement, dimensions, and relationship with decision making. European Journal of Business and Innovation Research, 6(5), 36-42.

[16] Souibgui, M., Atigui, F., Yahia, S. B., & Si-Said Cherfi, S. (2020). Business intelligence and analytics: On-demand ETL over document stores. In Research Challenges in Information Science: 14th International Conference, RCIS 2020, Limassol, Cyprus, September 23–25, 2020, Proceedings 14 (pp. 556-561). Springer International Publishing.

[17] Janssen, M., Van Der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. Journal of business research, 70, 338-345.

[18] Alkatheeri, Y., Ameen, A., Isaac, O., Al-Shibami, A., & Nusari, M. (2020). The mediation effect of management information systems on the relationship between big data quality and decision making quality. Test Engineering and Management, 82, 12065-12074.

[19] Hoogland, I., Schildkamp, K., Van der Kleij, F., Heitink, M., Kippers, W., Veldkamp, B., & Dijkstra, A. M. (2016). Prerequisites for data-based decision making in the classroom: Research evidence and practical illustrations. Teaching and teacher education, 60, 377-386.

[20] Agasisti, T., & Bowers, A. J. (2017). 9. Data analytics and decision making in education: Towards the educational data Scientist as a key actor in schools and higher education institutions. Handbook of contemporary education economics, 184.

[21] Shankar Narayanan, G., & Zhu, B. (2021). Enhancing decision-making with data quality metadata. Journal of Systems and Information Technology, 23(2), 199-217.

[22] Brynjolfsson, E., & McElheran, K. (2016). Data in action: Data-driven decision making in US manufacturing. US Census Bureau Center for Economic Studies Paper No. CES-WP-16-06, Rotman School of Management Working Paper, (2722502).

[23] Ghasemaghaei, M. (2019). Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency. Decision Support Systems, 120, 14-24.

[24] Brous, P., Janssen, M., Schraven, D., Spiegeler, J., & Duzgun, B. C. (2017, April). Factors Influencing Adoption of IoT for Data-driven Decision Making in Asset Management Organizations. In IoTBDS (pp. 70-79).

[25] Hassenstein, M. J., & Vanella, P. (2022). Data quality—Concepts and problems. Encyclopedia, 2(1), 498-510.

[26] Lennerholt, C., & van Laere, J. (2019). Data access and data quality challenges of self-service business intelligence. In 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. Association for Information Systems.

[27] Liu, C., Nitschke, P., Williams, S. P., & Zowghi, D. (2020). Data quality and the Internet of Things. Computing, 102(2), 573-599.

[28] Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. ACM computing surveys (CSUR), 41(3), 1-52.

[29] Samitsch, C. (2014). Data quality and its impacts on decision-making: How managers can benefit from good data. springer.

[30] Gebre, G. G., Isoda, H., Amekawa, Y., Rahut, D. B., Nomura, H., & Watanabe, T. (2022). Marketing efficiency among gender-based decision-making farm households in Southern Ethiopia. Journal of International Food & Agribusiness Marketing, 34(5), 538-563.