

Letter to Editor

Investigating the Challenges of Big Data Analytics from the Viewpoints of Students in Mashhad in 2019

Elham Nazari, Marziyeh Afkanpour, Hamed Tabesh*

Department of Medical Informatics, faculty of medicine, Mashhad University of Medical Sciences, Mashhad, Iran.

ARTICLE INFO

ABSTRACT

Received 05.09.2019
 Revised 16.10.2019
 Accepted 25.11.2019
 Published 03.12.2019

Key words:

Big data;
 Challenges;
 Analyses

Introduction: Nowadays Big Data Analytics has attracted students for research due to its very high capabilities, but there are also obstacles to analyses that need to be addressed. Therefore, the purpose of this study is to investigate the viewpoints of students of different disciplines at Mashhad universities on the challenges of this analysis.

Method: This study is a cross-sectional study conducted on students of different universities and fields such as computer engineering, pharmacy, industry and biology in Mashhad, Iran. A questionnaire based on literature review in Pubmed, Google scholar, and science direct databases was designed by 10 experts from different disciplines using Delphi method. 185 students participated in the study. Students' viewpoints on the challenges were also collected. Descriptive and analytical results were reported using SPSS 21 and Maxqda software.

Results: The age range of most students was 25 - 34 years. 54.2% were female. Most of the participants in this study were students of engineering and medical informatics. Of the participants in this study, 96.4% considered big data analytics necessary, 50.6% were familiar with the benefits of analytics. Lack of awareness, inadequate management, lack of managers' knowledge, lack of expertise, and lack of priority were the most important challenges for students.

Conclusion: Despite the importance and benefits of big data analytics, challenges are a major barrier to use that need to be addressed.

Introduction

Today, big data analytics has become a topic of interest to researchers and has attracted the attention of many academic communities and has been the subject of many students' research [1]. This type of analytics have many benefits including useful information discovery, being easy to summarize and interpret, improving processes, sharing information, and thus reducing costs [2,3]. In addition to various benefits of these

analyses, there are also challenges to them that if ignored, the results will change; some of these challenges involve lack of expert knowledge, lack of knowledge of the required methods, lack of awareness, and type of data [4, 5]. Big data analytics is used in different industries and fields. In industries such as banking, medical care, communications media, insurance and transportation, as well as in the fields of biochemistry and biology, physics, military,

* Corresponding Author E-mail address: Tabeshh@mums.ac.ir

astronomy, photography, the use of macro data analytics is evident [6-8]. Due to the importance of big data analytics in various industries, and regarding the fact that students and their research are related to industry and applied research, and on the other hand, this field is in the early stages of research in Iran and there is a great lack of familiarity with its related concepts, the purpose of this study is to investigate the students' viewpoints on the challenges of big data analytics.

Method

This cross-sectional study was designed on 185 students of Ferdowsi University and Medical Sciences of Mashhad University to examine views of students of various disciplines on the challenges of big data analytics. Mashhad is the largest city in eastern Iran with a population of about three million, located on the borders with Afghanistan and Turkmenistan on the Silk Road.

Mashhad has two major universities called Ferdowsi and Medical Sciences. Students at Ferdowsi University study engineering and basic sciences, and at the University of Medical Sciences, students study the medical sciences such as biology and medicine. A questionnaire was designed to investigate the students' viewpoints in different fields of Mashhad universities on the challenges of big data analytics.

The questionnaire includes close-ended questions and one open-ended question to explore the challenges. The basic items of the questionnaire were created by searching the texts in Google Scholar and science direct databases and were designed by Delphi method with the presence of 10 experts from various fields (medical informatics, statistics, HIT and computers).

This questionnaire is designed involving 3 questions. Its questions are listed in Table 1.

Table 1. Questions of the questionnaire

Questions	Description
QK1	Is big data analytics necessary?
QK2	What are the benefits of big data analytics?
QK3	What are the challenges of big data analytics?

The validity and reliability of the questionnaire were confirmed; 10 experts confirmed the validity and the reliability was affirmed 84% by Cronbach's Alpha method. Then, sampling was done and the questionnaires were provided to 185 students who were available. Students with different disciplines were involved in the research. Students of medical science involving medicine, dentistry, pharmacy, biotechnology, toxicology, nanotechnology, biotech, nutrition, medical image, radiology, microbiology, physiology, genetics, medical informatics,

biochemistry, immunology, HIT, molecular cell, medical physics and engineering students involving mechanical engineering, natural resources, water, industry, aerospace, metallurgy, computers, civil engineering, and students of basic sciences including statistics, mathematics, physics, chemistry, participated in the study; data were collected and it was ensured that participants completed all the questionnaires. Out of 190 questionnaires, 185 were completed. Data entry and analysis were performed using SPSS21 software.

Results:

Table 2. Descriptive specifications

Variables	Sub-variables	Frequency	percentage
	18-24 year	45	26.8
	25-34 year	91	54.2
	35-44 year	30	17.9
	45-54 year	2	1.2
sex			
	Male	77	45.8
	Female	91	54.2
Field of study			
	Dentistry	3	1.8
	Pharmacology	8	4.8
	Medicine	20	11.9
	Medical Science	17	10.1
	Information Technology (IT)	11	6.5
	Computer Engineering	27	16.1
	Other Engineering	40	23.8
	Medical Informatics	36	21.4
	HIT	3	1.8
	Basic Science	3	1.8
Grade			
	B.S.	36	21.4
	M.S.	77	45.8
	Professional Ph.D.	25	14.9
	Specialized Ph.D.	30	17.9

As you can see, most of the students were in the age range 25-34. Most were women. Most were engineering, medical and medical engineering students. Most of them were postgraduate students.

162 (96.4%) believed that big data analytics was necessary for Iran.

85 (50.6%) were familiar with the benefits of big data analytics.

Big data challenges for participants are listed in the Table 3 .

Challenges of big data were coded and analyzed based on the open ended questionnaire text and the results are presented in Table 3 and Figure1 below:

Table 3. Big data challenges

NO	Name	Frequency	% percentage
1	Lack of interests	4	2.16
2	Poor data quality	6	3.24
3	Lack of prioritization	17	9.19
4	Developing country	4	2.16
5	Infrastructures	12	6.49
6	Organizations	4	2.16
7	Habituation to the old systems	2	1.08

8	Managers' lack of knowledge	17	9.19
9	Incorrect management	30	16.22
10	Political problems	8	4.32
11	Lack of independent backup and executor	1	0.54
12	Lack of awareness	46	24.86
13	Newness	1	0.54
14	High costs	7	3.78
15	Complexity	2	1.08
16	Paper-based processes	2	1.08
17	Lack of education	3	1.62
18	Lack of expertise	17	9.19
19	Lack of study	1	0.54
20	lack of time	1	0.54
	Missing	0	0.00
	Total	185	100.00

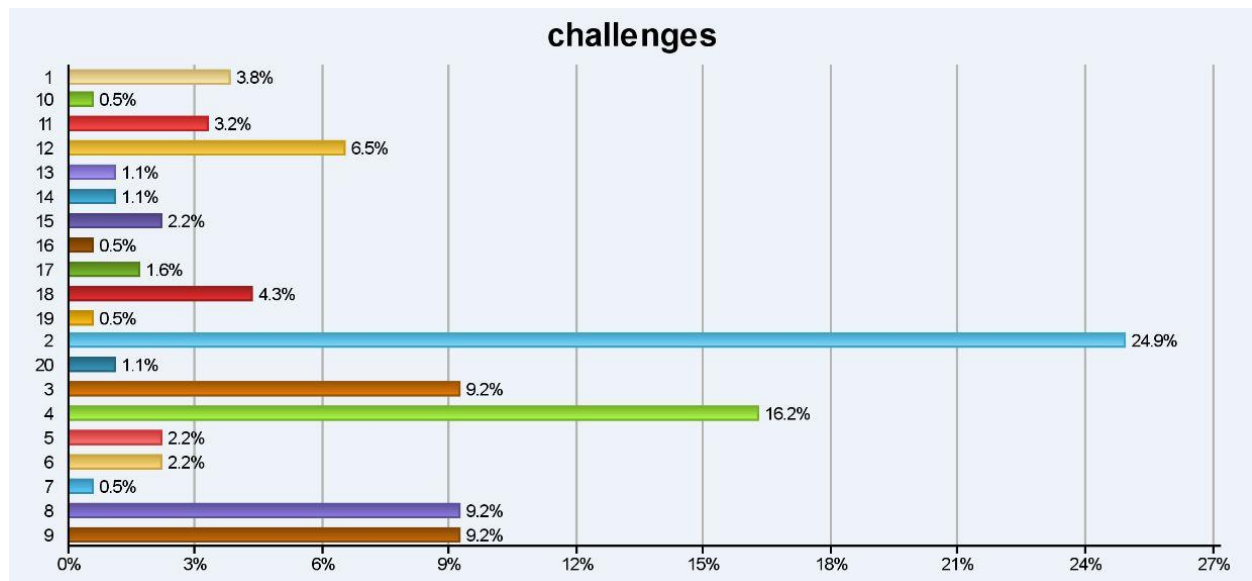


Figure 1. Frequency of challenges from the perspective of participants in the study

As you can see, lack of awareness, incorrect management, lack of managers' awareness, lack of prioritization, and habituation to the old systems are the most important challenges cited by the students.

Discussion and Conclusion

Today, with the enormous amount of data being produced, big data analytics is in the spotlight. Big data analytics has many advantages and capabilities, including the discovery of useful information. At the same time, these analyses

also present challenges that may be attributed to data quality and lack of expertise. It is essential to identify the challenges of big data analytics to make the most of its capabilities. In this study, students' viewpoints in Mashhad on the necessity and challenges of big data analytics were evaluated. Most students were in the age range 25-34. Most were female. The majority of participants in this study, considered big data analytics as necessary, half of them were familiar with the benefits of analysis. It seems necessary

to familiarize students who are in the research and learning phase with the benefits of big data analytics and the authorities need to think about it. Lack of awareness, incorrect management, lack of manager awareness, lack of prioritization, and habituation to the old system were the most significant obstacles and challenges associated with big data analytics mentioned by students. From the perspective of students, management plays an important role in encouraging these analyses and considering them for promoting organizations. It seems that holding briefings for managers and explaining the benefits of big data analytics to them and giving examples of successful projects done in other countries can provide the impetus for doing these analyses and providing the situation for them. Conducting training courses, conferences and congresses will also provide students with a familiarity to this type of analysis. In the study of 9 information systems, infrastructure, expertise and security were identified as big data challenge among employees in different companies. The challenges seem to be different from the perspectives of employees in organizations and students, and this may be related to that the staff are working in real world spaces and many students do not have work experience and their perspectives on challenges are based on studies and information from the space they are in and naturally their opinions are different from the staff's perspectives. Therefore, people's perceptions in different situations are shaped depending on the situation. It is therefore recommended that similar studies be conducted in different industries and organizations in the future. Studies have also highlighted the challenges of lack of proper infrastructure, lack of knowledge about analysis and methods, type of data and uncertainty and fuzziness of data, security, privacy, and high costs [10–14] that are worth being thought about.

Declaration

HIT: Health Information Technology

IT: Information Technology

Availability of data and methods

All data generated or analysed during this study are included in this published article.

Competing interests

The author(s) declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Authors' Contributions

Study conception and design: Elham Nazari, Hamed Tabesh;

Acquisition of data: Elham Nazari, Hamed Tabesh;

Analysis and interpretation of data: Elham Nazari, Hamed Tabesh, Marziyeh Afkanpour;

Drafting of the manuscript: Hamed Tabesh, Elham Nazari, Marziyeh Afkanpour
Critical revision: Hamed Tabesh;

Acknowledgement

The present study was carried out as a result of a research project approved by Medical Sciences University of Mashhad with the code 961731.

References

1. Jin X, Wah BW, Cheng X, Wang Y. Significance and challenges of big data research. *Big Data Research*. 2015 Jun 1;2(2):59-64.
2. Sagirolu S, Sinanc D. Big data: A review. In 2013 International Conference on Collaboration Technologies and Systems (CTS) 2013 May 20 (pp. 42-47). IEEE.
3. Srinivasan U, Arunasalam B. Leveraging big data analytics to reduce healthcare costs. *IT professional*. 2013 Nov;15(6):21-8.
4. Acharjya DP, Ahmed K. A survey on big data analytics: challenges, open research issues and tools. *International Journal of Advanced*

- Computer Science and Applications. 2016 Feb 1;7(2):511-8.
5. Alharthi A, Krotov V, Bowman M. Addressing barriers to big data. *Business Horizons*. 2017 May 1;60(3):285-92.
6. Zhang Q, Yang LT, Chen Z, Li P. A survey on deep learning for big data. *Information Fusion*. 2018;42:146-57.
7. Chen M, Mao S, Liu Y. Big data: A survey. *Mobile Networks and Applications*. 2014;19(2):171-209
8. Archana J, Anita EM. A survey of big data analytics in healthcare and government. *Procedia Computer Science*. 2015;50:408-13.
9. Raguseo E. Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*. 2018 Feb 1;38(1):187-95.
10. Jagadish H, Gehrke J, Labrinidis A, Papakonstantinou Y, Patel JM, Ramakrishnan R, et al. Big Data and its technical challenges. *Communications of the ACM*. 2014;57(7):86-94.
11. Zhang Q, Yang LT, Chen Z, Li P. A survey on deep learning for Big Data. *Information Fusion*. 2018;42:146-57.
12. Fan J, Han F, and Liu H. Challenges of Big Data analysis. *National science review*. 2014; 1(2):293-314.
13. Ristevski B, Chen M. Big Data Analytics in Medicine and Healthcare. *Journal of integrative bioinformatics*. 2018;15(3).
14. Bossé É, Solaiman B. *Information fusion and analytics for Big Data and IoT*: Artech House; 2016.