Studying Students' Knowledge of the Benefits, Challenges, and Applications of Big Data Analytics In Healthcare

Elham Nazari
Mashhad University of Medical Sciences
Maryam Edalati Khodabandeh
Mashhad University of Medical Sciences
Ali Dadashi
Mashhad University of Medical Sciences
Marjan Rasoulian
Mashhad University of Medical Sciences
hamed tabesh (tah2584@gmail.com)
Mashhad University of Medical Sciences

Research

Keywords: Big Data, Benefits, Challenges, Analysis

Posted Date: February 4th, 2020

DOI: https://doi.org/10.21203/rs.2.22540/v1

Abstract

Introdution

Today, with the advent of technologies and the production of huge amounts of data, Big Data analytics has received much attention especially in healthcare. Understanding this field and recognizing its benefits, applications and challenges provide useful background for conducting efficient research. Therefore, the purpose of this study was to evaluate the students' familiarity from different universities of Mashhad with the benefits, applications and challenges of Big Data analysis.

Method

This is a cross-sectional study that was conducted on students of Medical Engineering, Medical Informatics, Medical Records and Health Information Management in Mashhad-Iran. A questionnaire was designed based on literature review in pubmed, google scholar, science direct and EMBASE databases, using Delphi method and presence of 10 experts from different fields of study. The designed questionnaire evaluated the opinion of students regarding benefits, challenges and applications of Big Data analytics. 200 students participated in the study and completed the designed questionnaire. Participants' opinions were evaluated descriptively and analytically.

Result

Most students were between 20 and 30 years old. 63% of them were male and 43.5% had no work experience. Current and previous field of study of most of the students were HIT, HIM, and Medical Records. Most of the participants in this study were undergraduates. 61.5% were economically active, 54.5% were exposed to Big Data. The mean scores of participants in benefits, applications, and challenges section were 3.71, 3.68, and 3.71, respectively, and process management was significant in different age groups (p=0.046), information, modeling, research, and health informatics across different fields of studies were significant (p=0.015, 0.033, 0.001, 0.024) Information and research were significantly different between groups (p=0.043 and 0.019), research in groups with / without economic activity was significant (p= 0.017) and information in exposure / non exposure to Big Data groups was significant (p=0.02).

Conclusion

Despite the importance and benefits of Big Data analytics, students' lack of familiarity with the necessity and importance of these analytics in industries and research is significant. The field of study and level of study do not appear to have an effect on the degree of knowledge of individuals regarding Big Data analysis. The design of technical training courses in this field may increase the level of knowledge of individuals regarding Big Data analysis.

Introduction

Today, with the advent of various technologies, a huge amount of data that is known as Big Data in being generated esspeciall in health care. Big data analytics has become a hot topic and has been the focus of many academic communities and the subject of many students research. This type of data has features such as high volume and diversity and due to these features, they cannot be managed and analyzed using conventional hardware and software [1, 2]. Analytics for analyzing Big Data are known as Big Data Analytics and have many benefits including useful data patterns discovery and important features extraction [3]. These analyzes have many applications in various medical and insurance industries [4]. In addition to the many benefits of these analytics, there are challenges that if ignored, the results will change, such as a lack of expert staff, lack of familiarity with the tools and methods required, data type, security issues, budget and etc. [5, 6]. Understanding the benefits, challenges, and applications of this area can be helpful in conducting useful and efficient research. Due to the importance of Big Data analysis in various industries and the fact that students and their research are related to industry and applied research, this field in Iran is in the early stages of research and unfamiliarity with the concepts is severely felt. The purpose of this study is to investigate students' familiarity with the different Benefits, applications, and challenges of Big Data.

Method

This cross-sectional study was designed on 200 students of Ferdowsi University and Mashhad University of Medical Sciences. Mashhad is the largest city in eastern Iran with a population of about three million, located on the border with Afghanistan and Turkmenistan on the Silk Road. Mashhad has two major universities, Ferdowsi and Medical Sciences, which students in engineering and basic sciences study at Ferdowsi University and students in medical sciences such as medical Records, Health Information Management and Medical Informatics study at Mashhad University of Medical Sciences. A questionnaire was designed to assess the level of knowledge of students in Mashhad universities about the benefits, applications and challenges of Big Data analysis.The questionnaire contains close-end questions with a five-point Likert scale. The basic items of the questionnaire were based on literature searches in Google Scholar, Science Direct, Google Scholar and EMBASE databases and were designed and validated by Delphi method with the participation of 10 experts from various fields (Medical Informatics, Biostatistics, HIT and Computer Sciences). The questionnaire was designed in the form of 3 general items of benefits, applications and challenges. Benefits included information with 5 questions, modeling with 3 questions, data with 5 questions, process management with 6 questions. Application questions consisted of health service delivery with 17 questions, research with 4 questions, health information with 16 questions, essential medicine with 15 questions, health financial with one question, leadership and governance with 6 questions. Challenge questions included 9 questions. The questions are listed in Table 1:

Table 1

Questions

Items	Questions	Category	Subcategory	1	2	3	4	5
Advantages	In your opinion, which advantages are related with Big Data analysis?	Information	Generating new knowledge					
			Sharing information					
			Displaying and summarizing information					
			Extracting information and delivery for better results					
			Using meaningful information					
		Modeling	Predicting disease epidemics					
			Increasing confidence					
			Discovering and exploring behavioral pattern or activities					
		Data	Decreasing ambiguity					
			Increasing reliability					
			Reducing uncertainty					
			Improving data quality					
			Managing massive volumes of data					
		Process management	Improving clinical trial quality					

Items	Questions	Category	Subcategory	1	2	3	4	5
			Improving operational efficiencies					
			Interpreting easiness					
			Improving entity detection					
			Managing communications that are seemingly unrelated					
			Improving the ability of intelligent systems					
Applications	In your opinion, Which applications are related with Big Data analysis?	1.Health Service Delivery	Disease screening					
			public health					
			Disease earlier diagnosis					
			Patient-centered services					
			Therapeutic approaches improvement					
			Surgery					
			Rehabilitation					
			Clinical operations analysis					
			Primary care					
			Readmissions management					
			Health care delivery					
			Disease management					
			Cause of disease detection					

Items	Questions	Category	Subcategory	1	2	3	4	5
			Decompensation management					
			Blood transfusion management					
			Triage management					
			Health care data management					
		2.Research	Prediction					
			Disease pattern analysis					
			Side effects discovery					
			Research& development& Innovation					
		3.Health Information	personalized medicine					
			PHR(Personal Health Record) and HER					
			EBM (Evidence Base Medicine)					
			Patient monitoring					
			Web and social media					
			IOT(Internet Of Things)					
			Semantic standards					
			Biometric					
			Patient profile analytics					
			CPOE (computerized physician order entry)					

Items	Questions	Category	Subcategory	1	2	3	4	5
			Health informatics					
			Coding management					
			IT infrastructure management					
			Quality measurement					
			Bioinformatics and genetics					
			Comorbidity Discovery, Adverse events Discovery					
		4.(Essential) Medicines	Diagnosis					
			Precision medicine					
			CDSS(Clinical Decision Support System)					
			Sensor processing					
			RFID(Radio-Frequency identification)					
			Signal processing					
			Drug discovery & clinical Research					
			Vision augment					
			GPS(Global Positioning System)					
			Telemedicine, E-health, Remote healthcare system					
			Mobile health					
			Information Support					
			Image processing					

Items	Questions	Category	Subcategory	1	2	3	4	5
			BCI(Brain Computer Interface) and smart home					
			Recommender systems					
		5.Health Financing	Cost Reduction & Insurance service					
		6.Leadership and Governance	R & D in medications					
			Hospital quality monitoring					
			Resource management					
			Resource management					
			Operational management					
			Business and organizational and Strategic management					
Challenges	In your opinion, what challenges is there in big data analysis	---	Lack of knowledge about appropriate for the purpose					
			Lack of IT infrastructure					
			Lack of expertise about appropriate tools and algorithms					
			Variable and scalable data					
			Lack of data quality					
			Data uncertainty and missing data					

Items	Questions	Category	Subcategory	1	2	3	4	5
			Unstructured data					
			Security and privacy issue					
			High cost					

The validity and reliability of the questionnaire were confirmed by the presence of 10 validity experts and the reliability was confirmed by Alpha Cronbach's 92.1%. The questionnaires were then distributed to 200 students. Students in Medical Engineering, Medical Informatics, Medical Records, and Health Information Management participated in the study. Data were collected to ensure that participants answered all the questions. 200 questionnaires were completed. Data entry and analysis were performed using EXCEL (v. 2007) and SPSS (v. 21 software).

Results

For this study, 200 students participated and the results are as follows.

Table 2
Individual characteristics of the participants

Variables	Items	Frequency (percentage) of student)n = 200)
Age	< 20 year	22(11%)
	20–30 year	113(56.5%)
	30–40 year	46(23%)
	> 40 year	19(9.5%)
Gender	Male	126(63%)
	Female	73(36.5%)
	Missing	1(0.5%)
Field of study	Medical Engineering	70(35%)
	MI	43(21.5%)
	HIT	82(97.5%)
	missing	5(2.5%)
Degree	BA	77(38.5%)
	MA	73(36.5%)
	Professional doctorate	43(21.5%)
	Missing	7(3.5%)0
Prior field	HIT,HIM,Medical record	55(27.5%)
	MI	12(6%)
	C-E-M*	33(16.5%)
Work experience	0 year	87(43.5%)
	1–5 year	62(31%)
	5–10	24(12%)
	> 10	27(13.5%
Activity	Yes	123(61.5%)
	No	70(35%)
	Missing	7(3.5%)
Exposure	Yes	81(40.5%)

Variables	Items	Frequency (percentage) of student)n = 200)
	No	109(54.5%)
	missing	10(5%)

*Computer,Electronic,Mathematic

Most students were between 20 and 30 years old. 63% of them were male and 43.5% had no work experience. Current and previous field of study of most of the students were HIT, HIM, and Medical Records. Most of the participants in this study were undergraduates. 61.5% were economically active. 54.5% were exposed to Big Data. The mean scores of participants in benefits, applications, and challenges section were 3.71, 3.68, and 3.71, respectively (saschallenge, sasadvantage and sasapplication). Examination of saschallenge, sasadvantage, and sasapplication by variables of age, gender, field of study, Prior field, work experience, with / without activity, exposure / non-exposure to Big Data can be seen below:

Table 3
Comparison of mean of saschallenge, sasadvantage and sasapplication across different age groups

		n	Mean ± SD(n)
Questions	Age	N	
Advantages	20 > year	22	.6986 ± .11620
	20–30 year	113	.7522±..12519
	30–40 year	46	.7574 ± .12829.
	> 40 year	19	.7252±..12159
	Total	200	.7449±..12508
Applications	< 20 year	22	.6989±..12051
	20–30 year	113	.7413±..13019
	30–40 year	46	.7528 ± .11257.
	> 40 year	19	.7147±..12070
	Total	200	.7368±..12462
Challenges	< 20 year	22	.6869±..15257
	20–30 year	113	.7392±..16566
	30–40 year	46	.7744±..15188
	> 40 year	19	.6982±..18948
	Total	200	.7377±..16466

One-way Anova test was used to compare mean of saschallenge, sasadvantage and sasapplication in different age groups with no significant difference in different age groups in these factors. Pvalue was 0.228, 0.317, 0.139 respectively.

Table 4
Comparison of mean of saschallenge, sasadvantage and sasapplication in different gender groups

		N	Mean ± SD(n)
	Gender	N	
Advantages	Male	126	.7454 ± .11719
	Female	73	.7471 ± .13709
Applications	Male	126	.7329 ± .13281
	Female	73	.7446 ± .11009
Challenges	Male	126	.7383 ± .17504
	Female	73	.7370 ± .14741

Independent t-test was used to compare the mean of saschallenge, sasadvantage and sasapplication in different gender groups with no significant difference in different age groups in these factors.

Table 5
Comparison of average of saschallenge, sasadvantage, and sasapplication across different fields of study

		n	Mean ± SD(n)
	field	N	
advantages	Medical engineering	70	.7302 ± .13611
	MI	43	.7760 ± .12040
	HIT	82	.7488 ± .11194
	Total	195	.7481 ± .12348
applications	Medical engineering	70	.7236 ± .11333
	MI	43	.7778 ± .13359
	HIT	82	.7337 ± .12527
	Total	195	.7398 ± .12416
challenges	Medical engineering	70	.7140 ± .15265
	MI	43	.8114 ± .16246
	HIT	82	.7293 ± .16021
	Total	195	.7419 ± .16167

One-way Anova test was used to compare the mean of saschallenge, sasadvantage and sasapplication in different fields, but the mean of sasapplication and sasadvantage were not significant.

The mean of saschallenge was significant in different disciplines. The mean of saschallenge in medical informatics was higher than other majors (Fig. 1).

Table 6
Comparison of the mean of saschallenge, sasadvantage, and sasapplication between different levels of study

		n	Mean ± SD(n)
	Degree	N	
Advantages	BSC	77	.7270 ± .12249
	MSC	73	.7521 ± .13956
	PHD	43	.7718 ± 08582
	Total	193	.7465 ± .12313
Applications	BSC	77	.7249 ± .13235
	MSC	73	.7415 ± .12568
	PHD	43	.7602 ± .09684
	Total	193	.7390 ± .12285
Challenges	BSC	77	.6987 ± .16116
	MSC	73	.7461 ± .17661
	PHD	43	.7953 ± .12388
	Total	193	.7382 ± .16345

One-way Anova test was used to compare the mean of saschallenge, sasadvantage, and sasapplication at different levels of study that the mean of sasapplication, sasadvantage, and saschallenge were not significant. Pvalues were 0.142, 0.313, 0.006 respectively.

Table 7
Comparison of the mean of saschallenge, sasadvantage, and sasapplication between previous fields of study

		n	Mean ± SD(n)
	Perior field	N	
Advantages	HIT	55	.7678 ± .11516
	MI	12	.7675 ± .06710
	Engineering,electronic,math	33	.7652 ± .16349
	Total	100	.7669 ± .12796
Applications	HIT	55	.7503 ± .11978
	MI	12	.7893 ± .08649
	Engineering,electronic,math	33	.7548 ± .12074
	Total	100	.7564 ± .11628
Challenges	HIT	55	.7693 ± .16136
	MI	12	.7889 ± .13283
	Engineering,electronic,math	33	.7946 ± .16150
	Total	100	.7800 ± .15728

The one-way Anova test was used to compare the mean of saschallenge, sasadvantage and sasapplication between the previous fields of study, but the mean of sasapplication, sasadvantage and saschallenge were not significant.

Table 8
Comparison of the mean of saschallenge, sasadvantage, and sasapplication between different work experiences

		n	Mean ± SD(n)
	Work experience		
Advantages	0 year	87	.7459 ± .11910
	1–5 year	62	.7620 ± .13001
	5–10	24	.7154 ± 14901
	> 10	27	.7290 ± .10848
	Total	200	.7449 ± .12508
Applications	0 year	87	.7426 ± .12193
	1–5 year	62	.7441 ± .13534
	5–10	24	.7185 ± .10836
	> 10	27	.7176 ± .12415
	Total	200	.7368 ± .12462
Challenges	0 year	87	.7367 ± .16638
	1–5 year	62	.7559 ± .15951
	5–10	24	.7167 ± .16671
	> 10	27	.7177 ± .17332
	Total	200	.7377 ± .16466

One-way Anova test was used to compare the mean of saschallenge, sasadvantage and sasapplication between different work experiences that the mean of sasapplication, sasadvantage and saschallenge were not significant. Pvalues were 0.404, 0.673, 0.673 respectively.

Table 9
Comparison of the mean of saschallenge, sasadvantage, and sasapplication in groups with / without economic activity

	Activity	N	Mean ± SD(n)
Advantages	Yes	123	.7521 ± .12231.
	No	70	.7403 ± .13160.
Applications	Yes	123	.7454 ± .12092.
	No	70	.7185 ± .13357.
Challenges	Yes	123	.7478±..17636
	No	70	.7251 ± .14521.

Independent t-test was used to compare the mean of saschallenge, sasadvantage and sasapplication in the groups with / without economic activity in these factors. Pvalues were 0.532, 0.155, 0.361 respectively.

Table 10
Comparison of the mean of saschallenge, sasadvantage, and sasapplication in groups with / without exposure to Big Data

		n	Mean ± SD(n)
	Exposure	N	
Advantages	Yes	81	.7619 ± .11752
	No	109	.7359 ± .13009
Applications	Yes	81	.7561 ± .11112
	No	109	.7239 ± .13370
Challenges	Yes	81	.7627 ± .15108
	No	109	.7252 ± .16977

Independent t-test was used to compare the mean of rasadvantage, saschallenge and sasapplication in the groups with / without exposure to Big Data that there is no significant difference between the groups with / without exposure to Big Data in these factors. Pvalues were 0.157, 0.08, 0.116 respectively.

In order to examine the sasadvantage, saschallenge and sasapplication sub-domains, the previous analyzes of each sub-domain are repeated in terms of variables such as age, gender, field of study,

degree, and so on.

Table 11
Comparison of the mean of sasadvantage, saschallenge and sasapplication domains by age

		n	
	Age	N	Mean ± SD(n)
Information	< 20 year	22	.7491 ± .14458
	20–30 year	113	.7692 ± .17013
	30–40 year	46	.7843 ± .15966
	> 40 year	19	.7789 ± .12534
	Total	200	.7714 ± .16057
Modeling	< 20 year	22	.7364 ± .15324
	20–30 year	113	.7611 ± .18425
	30–40 year	46	.7754 ± .18979
	> 40 year	19	.7719 ± .16226
	Total	200	.7627 ± .17954
Data	< 20 year	22	.6545 ± .14790
	20–30 year	113	.7384 ± .15538
	30–40 year	46	.7252 ± .14910
	> 40 year	19	.7137 ± .17802
	Total	200	.7238 ± .15637
Process_Managment	< 20 year	22	.6742 ± .17516
	20–30 year	113	.7451 ± .14399
	30–40 year	46	.7529 ± .14633
	> 40 year	19	.6667 ± .20458
	Total	200	.7317 ± .15655
Health_Sevice_Delivery	< 20 year	22	.6786 ± .13032
	20–30 year	113	.7358 ± .14164
	30–40 year	46	.7453 ± .11181
	> 40 year	19	.7009 ± .11773
	Total	200	.7284 ± .13269

Research	< 20 year	22	.6909 ± 19557
	20–30 year	113	.7996 ± 19451
	30–40 year	46	.8087 ± .20718
	> 40 year	19	.7421 ± .20430
	Total	200	.7843 ± .20054
Health_Information	< 20 year	22	.7182 ± .13040
	20–30 year	113	.7357 ± .15005
	30–40 year	46	.7405 ± .13875
	> 40 year	19	.7092 ± .16398
	Total	200	.7324 ± .14611
Essential_Medicines	< 20 year	22	.6982 ± .11377
	20–30 year	113	.7371 ± .15626
	30–40 year	46	.7591 ± .13224
	> 40 year	19	.7039 ± .13988
	Total	200	.7347 ± .14565
Health_Financing	< 20 year	22	.7091 ± .24477
	20–30 year	113	.7469 ± .21384
	30–40 year	46	.7478 ± .20842
	> 40 year	19	.7789 ± .22992
	Total	200	.7460 ± .21661
Leadership_Governance	< 20 year	22	.7106 ± .13584
	20–30 year	113	.7428 ± .16770
	30–40 year	46	.7551 ± .16810
	> 40 year	19	.7667 ± .14741
	Total	200	.7443 ± .16227

One-way Anova test was used to compare the mean of sasadvantage, saschallenge and sasapplication domains by age groups that process management became significant. Pvalues were 0.855, 0.861, 0.145, 0.046, 0.172, 0.072, 0.831, 0.315, 0.784, 0.680, respectively.

Table 12
Mean comparison of sasadvantage, saschallenge and sasapplication domains by gender

		n	Mean ± SD(n)
	Gender	N	
Information	Male	126	.7679 ± .15380
	Female	73	.7759 ± .17317
Modeling	Male	126	.7566 ± .17391
	Female	73	.7735 ± .19076
Data	Male	126	.7168 ± .15714
	Female	73	.7370 ± .15605
process_managment	Male	126	.7447 ± .14168
	Female	73	.7183 ± .16207
health_sevice_delivery	Male	126	.7252 ± .13799
	Female	73	.7357 ± .12353
Research	Male	126	.7794 ± .21441
	Female	73	.7932 ± .17664
Health_information	Male	126	.7268 ± .15116
	Female	73	.7426 ± .13836
Essential_medicines	Male	126	.7328 ± .15952
	Female	73	.7394 ± .15952
Health_financing	Male	126	.7317 ± .22650
	Female	73	.7699 ± .19908
Leadership_governance	Male	126	.7405 ± .16874
	Female	73	.7516 ± .15245

Independent t-test was used to compare the mean of sasadvantage, saschallenge and sasapplication in gender groups with no significant difference in gender in these factors. Pvalues were 0.738, 0.525, 0.383, 0.230, 0.592, 0.642, 0.463, 0.761, 0.234, respectively.

Table 13
Mean comparison of sasadvantage, saschallenge and sasapplication domains by different fields of study

		n	Mean ± SD(n)
	Field of study	N	
Information	Medical engineering	70	.7354 ± .18137
	MI	43	.8242 ± .14688
	HIT	82	.7780 ± .14113
	Total	195	.7729 ± .16058
Modeling	Medical engineering	70	.7238 ± .20968
	MI	43	.8124 ± .15753
	HIT	82	.7715 ± .15539
	Total	195	.7634 ± .17949
Data	Medical engineering	70	.7211 ± .15692
	MI	43	.7433 ± .13972
	HIT	82	.7224 ± .16066
	Total	195	.7266 ± .15441
Process_Managment	Medical engineering	70	.7367 ± .15739
	MI	43	.7450 ± .15514
	HIT	82	.7350 ± .13465
	Total	195	.7378 ± .14699
Health_Sevice_Delivery	Medical engineering	70	.7227 ± .12854
	MI	43	.7502 ± .13100
	HIT	82	.7261 ± .13873
	Total	195	.7302 ± .13320
Research	Medical engineering	70	.7321 ± .17796
	MI	43	.8802 ± .16873
	HIT	82	.7854 ± .21907
	Total	195	.7872 ± .20119
Health_Information	Medical engineering	70	.7212 ± .13062

	MI	43	.7887 ± .14366
	HIT	82	.7216 ± .14812
	Total	195	.7363 ± .14310
Essential_Medicines	Medical engineering	70	.7181 ± .13169
	MI	43	.7758 ± .14452
	HIT	82	.7379 ± .15266
	Total	195	.7391 ± .14449
Health_Financing	Medical engineering	70	.7457 ± 19537
	MI	43	.7349 ± .22560
	HIT	82	.7512 ± .22566
	Total	195	.7456 ± .21423
Leadership_Governance	Medical engineering	70	.7367 ± .14625
	MI	43	.7713 ± .19032
	HIT	82	.7394±..16188
	Total	195	.7455 ± .16304

One-way Anova test was used to compare the mean of sasadvantage, saschallenge, and sasapplication domains by field of study, that the mean of sasadvantage, saschallenge, and saschallenge in information, modeling, research, and health informatics were significant. Pvalues were 0.015, 0.033, 0.726, 0.935, 0.532, 0.001, 0.024, 0.119, 0.922 and 0.500 respectively (Fig. 2 and Fig. 3).

Table 14
Mean comparison of sasadvantage, saschallenge and sasapplication domains by different levels of study

		n	Mean ± SD(n)
		N	
Information	BSC	77	.7356 ± .17902
	MSC	73	.7863 ± .15581
	PHD	43	.8047 ± .12443
	Total	193	.7702 ± .16131
Modeling	BSC	77	.7359 ± .18589
	MSC	73	.7553 ± .19470
	PHD	43	.8155 ± 13161
	Total	193	.7610 ± 18059
Data	BSC	77	.7122 ± .16066
	MSC	73	.7332 ± .16547
	PHD	43	.7386 ± .11787
	Total	193	.7260 ± .15380
Process_Managment	BSC	77	.7277 ± .13050
	MSC	73	.7379 ± .18116
	PHD	43	.7504 ± .10193
	Total	193	.7366 ± .14627
Health_Sevice_Delivery	BSC	77	.7178 ± .13130
	MSC	73	.7289 ± .14290
	PHD	43	.7505 ± .10839
	Total	193	.7293 ± .13117
Research	BSC	77	.7487 ± .20822
	MSC	73	.7849 ± .21192
	PHD	43	.8558 ± .14809
	Total	193	.7863 ± .20112
Health_Information	BSC	77	.7237 ± .15430

	MSC	73	.7408 ± 14471
	PHD	43	.7462 ± .12148
	Total	193	.7352 ± .14352
Essential_Medicines	BSC	77	.7186 ± .15838
	MSC	73	.7450 ± .14472
	PHD	43	.7606 ± .11002
	Total	193	.7380 ± .14393
Health_Financing	BSC	77	.7377 ± .22771
	MSC	73	.7479 ± .22367
	PHD	43	.7581 ± .17759
	Total	193	.7461 ± .21505
Leadership_Governance	BSC	77	.7455 ± .15284
	MSC	73	.7406 ± .17080
	PHD	43	.7605 ± .15106
	Total	193	.7470 ± .15886

One-way Anova test was used to compare the mean of sasadvantage, saschallenge and sasapplication domains by different levels of study that the mean of sasadvantage, saschallenge and sasapplication in information and research were significant that was more significant in PhD level. Pvalues were 0.043, 0.064, 0.589, 0.717, 0.427, 0.019., 0.654, 0.269, 0.880, 0.807.

Table 15
Mean comparison of sasadvantage, saschallenge and sasapplication domains by different previous fields of study

		n	Mean ± SD(n)
		N	
Information	HIT	55	.8196 ± .12290
	MI	12	.8033 ± .13694
	C-E-M	33	.7600 ± .19183
	Total	100	.7980 ± .15153
Modeling	HIT	55	.7782 ± .16079
	MI	12	.8444 ± .12818
	C-E-M	33	.7980 ± .19825
	Total	100	.7927 ± .17053
Data	HIT	55	.7484 ± .13612
	MI	12	.7133 ± .12630
	C-E-M	33	.7442 ± .18599
	Total	100	.7428 ± .15226
Process_Managment	HIT	55	.7358 ± .15567
	MI	12	.7444 ± .10856
	C-E-M	33	.7707 ± .18004
	Total	100	.7483 ± .15894
Health_Sevice_Delivery	HIT	55	.7435 ± .12499
	MI	12	.7578 ± .11098
	C-E-M	33	.7390 ± .15264
	Total	100	.7438 ± .13211
Research	HIT	55	.8091 ± .21860
	MI	12	.9333 ± .07177
	C-E-M	33	.8167 ± .16802
	Total	100	.8265 ± .19325
Health_Information	HIT	55	.7389 ± .14249

	MI	12	.7885 ± .10091
	C-E-M	33	.7583 ± .14340
	Total	100	.7512 ± .13829
Essential_Medicines	HIT	55	.7556 ± .13039
	MI	12	.7978 ± .10511
	C-E-M	33	.7503 ± .13676
	Total	100	.7589 ± .12946
Health_Financing	HIT	55	.7964 ± .19048
	MI	12	.7333 ± .19695
	C-E-M	33	.7455 ± .23061
	Total	100	.7720 ± .20503
Leadership_Governance	HIT	55	.7394 ± .16840
	MI	12	.7722 ± .11962
	C-E-M	33	.7616 ± .18373
	Total	100	.7507 ± .16774

One-way Anova test was used to compare the mean of sasadvantage, saschallenge and sasapplication domains by different previous fields of study that the mean of sasadvantage, saschallenge and sasapplication was not significant. Pvalues were 0.202, 0.469, 0.772, 0.610, 0.916, 0.122, 0.501, 0.537, 0.420 and 0.749 respectively.

Table 16
Mean comparison of sasadvantage, saschallenge and sasapplication domains by experience

		n	Mean ± SD(n)
	Workexperience	N	
Information	0 year	87	.7674 ± .15341
	1–5 year	62	.7923 ± .17650
	5–10 year	24	.7283 ± .15999
	> 10 year	27	.7748 ± .14471
	Total	200	.7714 ± .16057
Modeling	0 year	87	.7678 ± .16824
	1–5 year	62	.7785 ± .18098
	5–10 year	24	.6944 ± .23003
	> 10 year	27	.7704 ± .15616
	Total	200	.7627 ± .17954
Data	0 year	87	.7223 ± .15424
	1–5 year	62	.7497 ± .14771
	5–10 year	24	.6917 ± .18062
	> 10 year	27	.6978 ± .15858
	Total	200	.7238 ± .15637
Process_managment	0 year	87	.7368 ± .15668
	1–5 year	62	.7387 ± .14897
	5–10 year	24	.7347 ± .14955
	> 10 year	27	.6963 ± .18171
	Total	200	.7317 ± .15655
Health_sevice_delivery	0 year	87	.7348 ± .14072
	1–5 year	62	.7454 ± .13775
	5–10 year	24	.7049 ± .11289
	> 10 year	27	.6893 ± .10273
	Total	200	.7284 ± .13269

Research	0 year	87	.7810 ± .19681
	1–5 year	62	.8081 ± .19210
	5–10 year	24	.7437 ± .22904
	> 10 year	27	.7759 ± .20911
	Total	200	.7843 ± .20054
Health_information	0 year	87	.7385 ± .14954
	1–5 year	62	.7403 ± .14303
	5–10 year	24	.7156 ± .13054
	> 10 year	27	.7093 ± .15893
	Total	200	.7324 ± .14611
Essential_medicines	0 year	87	.7352 ± .14816
	1–5 year	62	.7452 ± .14842
	5–10 year	24	.7206 ± .13025
	> 10 year	27	.7220 ± .14984
	Total	200	.7347 ± .14565
Health_financing	0 year	87	.7770 ± .19630
	1–5 year	62	.6968 ± .22830
	5–10 year	24	.7583 ± .26361
	> 10 year	27	.7481 ± .19684
	Total	200	.7460 ± .21661
Leadership_governance	0 year	87	.7625 ± .13826
	1–5 year	62	.7129 ± .18782
	5–10 year	24	.7361 ± .16794
	> 10 year	27	.7654 ± .16316
	Total	200	.7443 ± .16227

One-way Anova test was used to compare the mean of sasadvantage, saschallenge and sasapplication domains by experience that the mean of sasapplication, sasadvantage, and saschallenge were not significant. Pvalues were 0.419, 0.255, 0.327, 0.661, 0.231, 0.592, 0.725, 0.863, 0.167, 0.270 respectively.

Table 17
Mean comparison of sasadvantage, saschallenge and sasapplication domains by economical activity

	Activity	N	Mean ± SD(n)
Information	yes	123	.7776 ± .16034
	no	70	.7657 ± .16447
Modeling	yes	123	.7691 ± .17850
	no	70	.7533 ± .18970
Data	yes	123	.7389 ± .15444
	no	70	.7091 ± .15941
Process_managment	yes	123	.7333 ± .15611
	no	70	.7386 ± .14319
Health_sevice_delivery	yes	123	.7296 ± .13578
	no	70	.7217 ± .13088
Research	yes	123	.8089 ± .19110
	no	70	.7364 ± .21752
Health_information	yes	123	.7419 ± .14312
	no	70	.7112 ± .15432
Essential_medicines	yes	123	.7457 ± .13364
	no	70	.7168 ± .16645
Health_financing	yes	123	.7463 ± .22914
	no	70	.7314 ± .19821
Leadership_governance	yes	123	.7561 ± .15770
	no	70	.7195 ± .17450

Independent t-test was used to compare the mean of sasadvantage, saschallenge and sasapplication by economical activity that there was a significant difference in different groups in research. Pvalues were 0.625, 0.565, 0.205, 0.693, 0.818, 0.017, 0.167, 0.761, 0.188, 0.649 and 0.133. respectively.

Table 18
Mean comparison of sasadvantage, saschallenge and sasapplication domains by exposure / non-exposure to Big Data

	Exposure	N	Mean ± SD(n)
Information	yes	81	.7970 ± .14571
	no	109	.7545 ± .16905
Modeling	yes	81	.7835 ± .17177
	no	109	.7468 ± .18851
Data	yes	81	.7358 ± .15135
	no	109	.7196 ± .16271
Process_managment	yes	81	.7436 ± .15755
	no	109	.7284 ± .14613
Health_sevice_delivery	yes	81	.7413 ± .13150
	no	109	.7182 ± .13423
Research	yes	81	.8142 ± .19577
	no	109	.7624 ± .20940
Health_information	yes	81	.7475 ± .13899
	no	109	.7218 ± .15380
Essential_medicines	yes	81	.7567 ± .12935
	no	109	.7231 ± .15477
Health_financing	yes	81	.7704 ± .20028
	no	109	.7248 ± .23060
Leadership_governance	yes	81	.7778 ± .15330
	no	109	.7217 ± .16957

Independent t-test was used to compare the mean sasadvantage, saschallenge and sasapplication by exposure / non-exposure to the Big Data that there was a significant difference in groups in information. Pvalues were 0.071, 0.169,0.486, 0.085, 0.494, 0.236, 0.114, 0.156, 0.020, 0.761, 0.188, 0.649, 0.133 respectively. The mean of sasinformation was higher among those exposed to the Big Data than those not exposed to the Big Data.

Discussion

Today, with the advent of technologies and the production of huge amounts of data, Big Data analytics has received much attention especially in health care. Understanding this field and recognizing its benefits, applications and challenges provide useful background for conducting efficient research. Therefore, the purpose of this study was to evaluate the students' familiarity from different universities of Mashhad with the benefits, applications and challenges of Big Data analysis. Most students were between 20 and 30 years old. Most of them were male and had no work experience. Current and previous field of study of most of the students were HIT, HIM, and Medical Records. Most of the participants in this study were undergraduates. Most of them were economically active and were exposed to Big Data. The mean scores of participants in benefits, applications, and challenges section were 3.71, 3.68, and 3.71, respectively. Considering that the participants in this study are students from the most prestigious universities in the country and have done some Big Data research, it is assumed that Mashhad students have a better level of knowledge in the field of Big Data analysis. Yet there should be more opportunities for students, even organizations' staff to get to know the field more. Training in this field is essential for many disciplines, also conferences could be effective in introducing this field. Students can also provide more familiarity and usage of functional analytics by conducting new researches in this field. In the section of challenges, benefits and applications analytics, process management was significant in different age groups, research, modeling and information and health informatics across different fields of studies were significant. Information and research were significantly different between different levels of studies. Research in groups with / without economic activity was significant and information in exposure / non exposure to Big Data groups was significant.

Conclusion

Despite the importance and benefits of Big Data analytics, students' lack of familiarity with the necessity and importance of these analytics in industries and research is significant. The field of study and level of study do not appear to have an effect on the degree of knowledge of individuals regarding Big Data analysis. In future studies, it is suggested that students, practitioners, and other disciplines in different cities and countries evaluate the specific benefits and applications of Big Data analytics and compare the results. Because it will be possible to study in different places and different perspectives. In other businesses, checking their familiarity with Big Data analytics can be helpful in applying management and advertising policies. Big data analytics can play a constructive role in all industries, and today it is widespread in most industries and businesses. Because of the growing trend of data generation, Big Data analytics will become a necessity for all industries and areas in the coming years.

Abbreviations

MI
Medical Informatics
HIM
Health Information Management

Declarations

Availibility 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: Maryam Edalati Khodabandeh, Marjan Rasoulian;

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

Drafting of the manuscript: Hamed Tabesh, Elham Nazari Ali Dadashi ;

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 980102.

References

1. Manogaran G, Lopez D, Thota C, Abbas KM, Pyne S, Sundarasekar R. Big data analytics in healthcare Internet of Things. InInnovative healthcare systems for the 21st century 2017 (pp. 263-284). Springer, Cham.
2. Belle A, Thiagarajan R, Soroushmehr SM, Navidi F, Beard DA, Najarian K. Big data analytics in healthcare. BioMed research international. 2015;2015. Häyrinen K, Saranto K, Nykänen P.
3. Srinivasan U, Arunasalam B. Leveraging big data analytics to reduce healthcare costs. IT professional. 2013 Nov;15(6):21-8.
4. Archenaa J, Anita EM. A survey of big data analytics in healthcare and government. Procedia Computer Science. 2015 Jan 1;50:408-13.

5. 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.
6. Alharthi A, Krotov V, Bowman M. Addressing barriers to big data. Business Horizons. 2017 May 1;60(3):285-92.

Figures

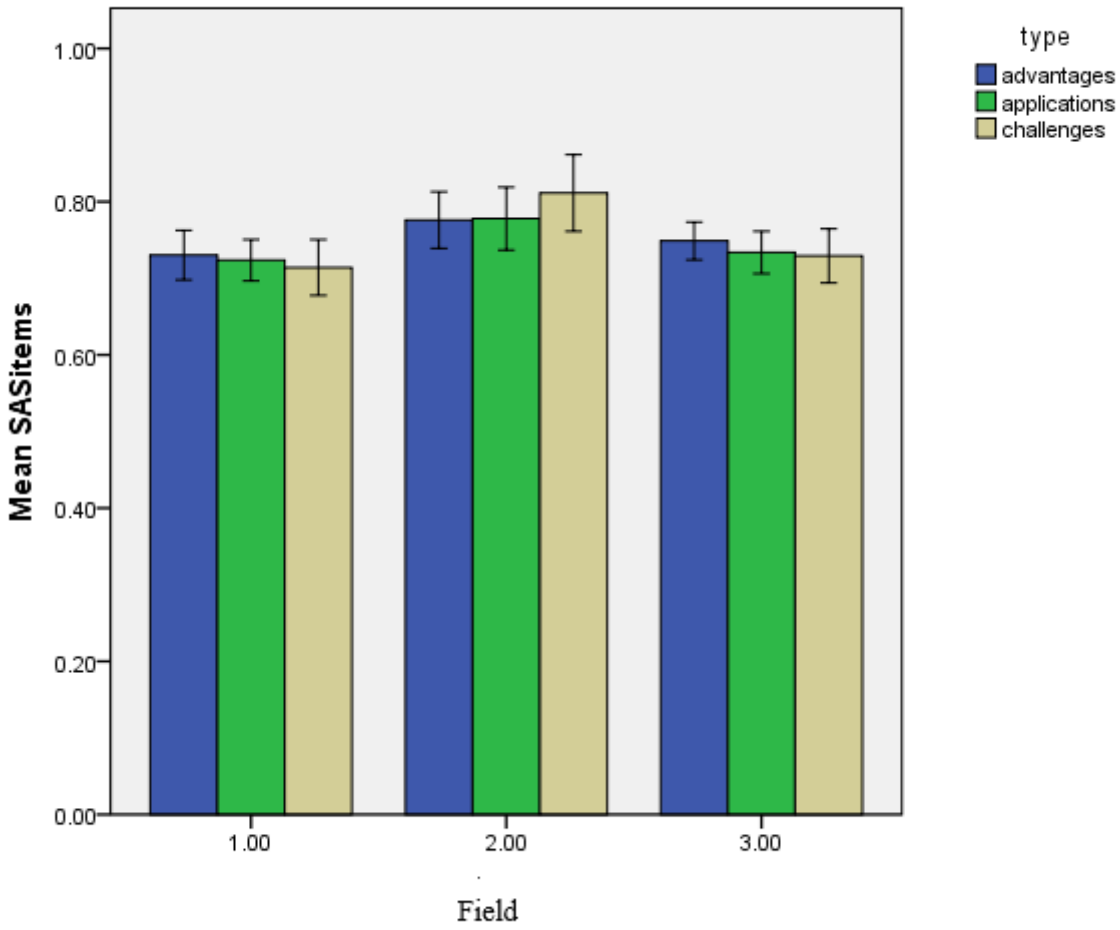

Figure 1

Mean of Benefits, applications and challenges in terms of the different fields of study

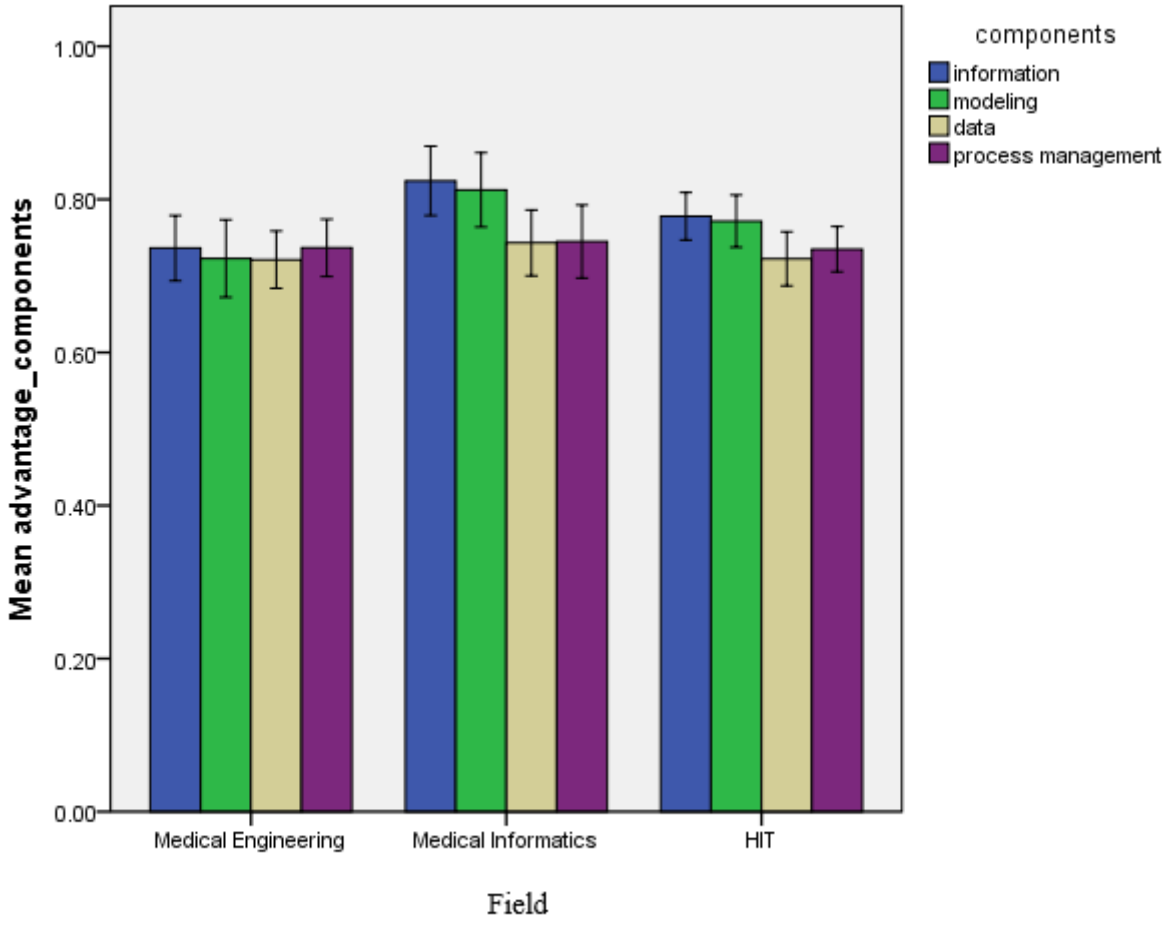

Figure 2

Average of the components of Benefits by field of study

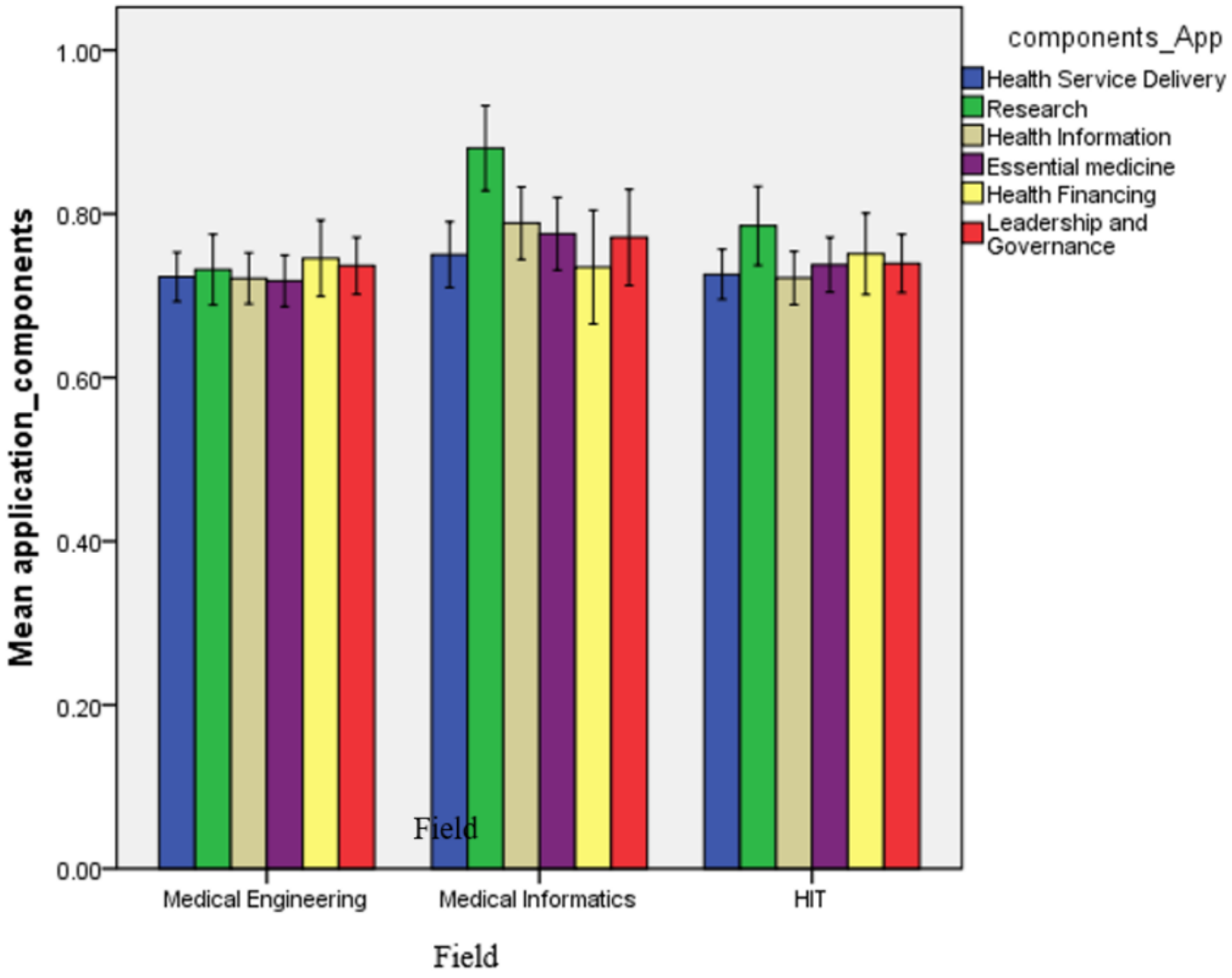

Figure 3

Average of the components of Application by field of study