International Journal of Research (IJR)

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 9 Issue 03
March 2022

The Potential of Implementing Data Science in Yemeni Private Universities (IBB City)

Dr. Hamzah Abdulrahman Al habri[1], Safaa abdullah Abduljabar[2]

[1]*Assistant Professor in Information system management , AL Jazeera University,* **(h.habri2021@gmail.com)**

[2]*A researcher in university of science and technology, (* **abdulsfa726@gmail.com)**

Abstract: *data science and analysis are very important concepts in any organization and they need to be managed properly. This study aims to reveal the potential of implementing data science in Yemeni private universities. The study was performed in four private universities namely UST,* Al Jazeera, al Quran and I M U in Ibb city. *Employees from top level were selected to expose their views regarding data science and its implementation. Quantitative method was the appropriate technique for this study and the questionnaires were employed for data collection. Furthermore, data was statistically analyzed. Based on the empirical finding it was clear there is a potential for implementing data science in universities and it was highly recommended that universities should give a serious thinking in applying machine learning and install new technologies to evaluate data science and analysis. This study will give an attempt to create awareness of the importance of data science in universities and other organizations.*

Keywords: data science, data analysis, machine learning

I. Introduction

Data is a crucial resource for any organization and it needs to be managed judiciously. It assists in decision making, problem solving, risk assessment and others. To gain the most benefit of data some approaches and related tools should be taking into consideration statistics, programming, database management. Theses approaches have to be deliberated scientifically. Extracting data from different resources in several patterns is a big task, hence data analysis and science have been emerged widely, bringing numerous ideas, machine learning and all required strategies in concern with gaining the advantages of data.

Currently data science is a significant part of all types of organization for many reasons like recognizing the clients with their desire sand behaviors, explaining the requirements and ideas powerfully and that help in making effective decision making and solving complex issues. Many benefits organizations can acquire from data science and universities have a tendency to to take this field seriously for more understanding the activities in all areas and managing these activities successfully.

Data science is increasing rapidly in all aspects taking several methods such as machine learning, pattern recognitions, visualization, modeling and others. However this

Copyright © authors 2022

study is giving attempt to investigate the potential of implementation of data science in Yemeni private universities.

II. Statement of the problem

Data science aims to reach clues that give a visualization of the workflow mechanism. However, private universities in Yemen are not aware of the importance of data science and analysis and due to the current trends in digital perspectives the workflow without data handling may leads to a failure to take the right decisions and failure to predict the future. consequently, this failure rises incorrect plans and strategies and lack of knowing strengths and weakness spots. From the point of view the problem statement rises the following question:

Is there a potential of implementing data science and analysis in Yemeni private universities?

III. Objectives of the study

1. To study whether the data science is given priority in Yemeni private universities
2. To study whether the technologies tools are adequate for data science and analysis in Yemeni private universities
3. To analyze the potential of implementing data science and analysis Yemeni private universities

IV. Hypotheses of the study

To provide solution to this research paper, the hypotheses that formulated are:

$\boldsymbol{H_0}$**:** Implementation of data science is not given priority in private universities in Yemen.

$\boldsymbol{H_1}$**:** Implementation of data science is given priority in private universities in Yemen.

V. Review of literature

In the current era data science has become an important technique in all sides of life specially in organizations and according to the book of Alex Galea (2018)[1], it has been revealed the instructions and functions of some tools specially Jupter as a magic function to be used in data analysis and science. visualization such as histogram, violin plot, scatter plot and others are explained widely to facilitate data science when in use Advanced machine learning as well as Data cleaning has been illustrated practically as once of the main concepts that evaluate data science. The author also

Copyright © authors 2022

shows how predictive models can be compile with machine learning strategy, and explain some terminology such as supervised learning, classification, regression which consider as crucial concepts in data science. one of the most important fact in this book is how the author has explained some libraries that used in python language such as pandas,scikit_learn. scraping web page tables are also clearly described in an interactive way using different methods like HTTP,GET,HTML and others.

Considering the study of Il-Yeol Song &Yongjun Zhu (2017)[2], the societies have been rapidly going through because of data science. It is a current trend in digital aspects that effect positively the future jobs, student careers and others. The core of this digital transformation is data science which leads to the perception of big data. The study discloses perspectives on iSchools' opportunities and suggestions in data science education. The study suggested iSchools to empower their students with "information computing" to solve problems and create values, information, and knowledge using the right technologies. It was also suggested the three foci of user-based, tool-based, and application based to make the difference between data science and computer science. The study introduced Data Science Education Framework (DSEF) with building blocks that consists of three pillars of data science people, technology, and data.

VI. Research design and methodology

This study reveals the potential of implementing data science in Yemeni private universities. **Quantitative method** was the appropriate technique for this study, the questionnaire was fully involved to collect the required data and it was sent to participants that were randomly selected. The responses were affirmative to fulfil the purpose of the study. **The questionnaire** was distributed to employees in higher level in the universities. **The population** of this study consists of employees in private universities, namely UST, Al Jazeera, al Quran and I M U universities in Ibb city, Yemen. In obtaining the **sample size** of the population, employees from top level were selected to response. **30 respondents** in the four universities have exposed their opinion regarding the data science and its possibility for implementation.

VII. Data presentation and analysis

Technology usage: Respondents were asked if the university use technology tools for data science and analysis with two respond options (1=Yes, 2=No).

Copyright © authors 2022

Are there any technology tools used for data science and analysis?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	22	73.3	73.3	73.3
	no	8	26.7	26.7	100.0
	Total	30	100.0	100.0	
Table1: usage of technology tools for data science and analysis					

From the above table, it can be seen that, 22 respondents in the universities said that there are technology tools use for data science and analysis while 8 respondents there are not.

Technology tools: Respondents were asked to comment on the technology tools are used for data science and analysis. They were further provided with the following options to choose from (Excel, SQL, SPSS, Python, R,BI Power, others)

what are the technology tools are used for data science and analysis ?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	EXCEL	27	90.0	90.0	90.0
	SPSS	3	10.0	10.0	100.0
	Total	30	100.0	100.0	
Table2: the availability of technology tools for data science and analysis					

From the above frequency distribution in table2, it can be seen that the most preferred technology tools for data science is Excel with percent of 90 out of whole respondents followed by SPSS with only 10 percent.

The potential of implementing data science and analysis: employees were offered to comment on the potential of implementing data science and analysis, using the options (5= strongly agree, 4= agree, 3= neutral, 2= disagree, 1= strongly disagree)

	N	Mean	Std. Deviation
university has a clear plan to have many questions that can be solved for all resources	30	1.9333	.86834
univesirty works regulary on data definiation and integration	30	4.03	.718
univesirty works regulary on data quality	30	2.33	.711

Copyright © authors 2022

university give a serious thinking on selecting key variables and correlation between them	30	2.3333	.71116
university give a serious thinking to clarify workflows of data	30	2.3333	.71116
university has a plan to build right models to performs data analysis	30	2.3333	.71116
university give a serious thinking in applying machine learning	30	2.3333	.71116
university gives a serous tninking in installing new technologies for data sience purpose	30	2.5333	.97320
Valid N (listwise)	30		

Table3: Descriptive Statistics of potential of implementing data science and analysis

From table3 it can be seen that the highest potential point regarding data science is univesirty works regulary on data definiation and integration with a mean of 4.03 and Std. Deviation of 0.7 followed by university gives a serous tninking in installing new technologies for data sience purpose with a mean of 2.5 and Std. Deviation of 0.9. On the other hand, the other points in the potential of implementing data science stated at low level that indicates applying machine learning, building models, installing new technologies for data science and analysis.

VIII. Hypotheses testing

$\boldsymbol{H_0}$**:** Implementation of data science is not given priority in private universities in Yemen.

$\boldsymbol{H_1}$**:** Implementation of data science is given priority in private universities in Yemen.

	t	df	Sig. (2-tailed)	Mean	Std. Deviation	95% Confidence Interval of the Difference	
						Lower	Upper
dat_si	33.750	29	.000	2.583	.41924	2.4268	2.7399

Table4: one sample test for hypotheses

Copyright © authors 2022

Interpretation: From table4 the hypothesis of the Implementation of data science is given priority in private universities in Yemen can be interpreted as following:

Observed proportion: mean:2.583, Std. Deviation:0.41, P=0.000. Hence P value is less than 0.05, hence, it can be concluded that there is a potential of implementing data science in private universities in Yemen, consequently "Implementation of data science is not given priority in private universities in Yemen." is proved.

IX. The empirical findings

The study findings have been written as the results that was gathered from the presented original data and analyzed.

1- Regarding the potential of implementing data science, it was found that Yemeni private universities give priority for implementation of data science and implementation based on significance of level that indicate to 0.000.
2- Considering to the technologies that are used for data science and analysis, it was found only Excel and SPSS only the software that used in data science and analysis with 90% of respondents for Excel and 10% of respondents for SPSS while other technologies like Python or R languages are not used at all.
3- In concern with data definition and integration it was found that universities work regularly on that with a mean rank of 4.3.
4- Regarding machine learning, it was found that universities do not give serious thinking in applying and practicing machine learning in data science with a mean rank of 2.3.
5- In concern with building models to perform data science and analysis, it was found that Yemeni private companies do not give full thinking to evaluate theses models with a mean rank of 2.3.

X. Recommendation and Conclusion

This study dealt with the potential of implementing data science and analysis in Yemeni private companies. Based on the study finding the authors can give recommendations which could clear the awareness of the importance of data science in Yemeni universities. Hence it is recommended that university should shift-focus from only using computer science to implementing data science and analysis to facilitate the sources and decisions making. Furthermore, Yemeni universities should give deep thinking in applying machine learning, building logic models to perform data science and analysis. Moreover, it is highly recommended that universities should improve the new technologies that improve the quality of data

Copyright © authors 2022

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Vol. 9 Issue 03
March 2022

science and analysis such as Python an R languages as well as SQL for retrieving the required data. Finally, data science plays an important role in managing the sources effectively, predicting intelligently, helping in making right decisions, solving the problems, assessing the risks and many benefits. Consequently, it can be concluded that data science and analysis have to be takin into consideration by universities and other organizations.

XI. References

1) Alex Galea (2018),Beginning Data Science with Python and Jupyter,pact published by Packt publishing Ltd 35livry street ,Birmingham B3 2PB,UK.
2) Il-Yeol Song &Yongjun Zhu (2017). Big Data and Data Science: Opportunities and Challenges of iSchools, Vol. 2 No. 3, pp 1–18.
3) Storey, V., & Song, I.-Y. (2017). Big data technologies and management: What conceptual modeling can do? Data & Knowledge Engineering, 108, 50–67.
4) B. Wieder, M. L. Ossimitz, P. Chamoni, "The Impact of Business Intelligence Tools on Performance: A User Satisfaction Paradox?", International Journal of Economic Sciences and Applied Research (2012); 5:3. 7-32.
5) Alex Galea(2018),Beginning Dada Science With Python and Jupyter,Packt Publishing Ltd .
6) Ashely Morrison(2019), Importance of Data Science, ZaronTech

Copyright © authors 2022