QUALITY ASSURANCE WITH REFERENCE TO DATA ANALYTICS

Muhammad Aamir Awan, Fahad Rasool
amirprince114@gmail.com

Abstract: Software quality is explicit property which determines what sort of standards software ought to have. In production of any software quality is a key factor which have to be accomplished. Numerous inquires about have been finished with respect to software quality. These Software Products progressively perplexing algorithms to operate and then for ensuring their quality, one needs further developed strategies for handling these Software Products. According to our survey, Data Analytical Software Products require more work because of their Products progressively perplexing nature. One of the potential reasons can be the volume and assortment of Data. On the other hand, this research emphasizes on testifying the Data Analytical Software Products which have many problems because testing of these Software Products requires real time data. however, each time the testing of these Product Items depends either on dummy information or reproductions and these Product failed when they work on real data. For making this softwares work good before and after the deployment phase we need to apply better software quality standards.

I.   Introduction:

Quality is significant for the accomplishment of software. Quality has the little effect in fruitful and bombed software ventures. Nature of Software is the explanation of achievement and decrease of a product related association. Numerous inquires about have been finished with respect to quality administration, estimation and confirmation. In this paper we will discuss the quality assurance problems of data analytics softwares and we will also define best rehearses for testing of product. Data analytics has a contribution in almost every industry, especially for making life comfortable and for the advancement of our society, whether it is health sector or smart homes, data analytics has a huge impact on human race [3]. We should examine the inner structure and formation of information driven or information Investigative virtual products in light of the fact that before discussing Quality Assurance we should think about the structure, advancement and reason for those frameworks. Principle objectives of data analytics are to peruse the information, comprehend the information and afterward discover patterns from it and above all to anticipate about the topic. These expectations can be highlighting a greater significance of subject that is being watched, and results acquiring from these investigations can be prompting new hypotheses and advancements, that is the reason these diagnostic software analytics are getting commonality step by step.

II.   Background

Today internet is available at every place especially in industrial sector and it has a huge impact on the production of data, previously data used to be stored on hard drives now tendency of storing data has moved towards digitization because production rate of data in streams is increasing every second. Especially in the industrial field where data is in huge amount and to handle and manage this huge data we need some special product [4].

Many methods have been proposed by the developers nin the field of research by multiple means of research methodology [5]. However the technology is going on a very fast pace, the Data gathering apps have introduced people with the big data and these applications generate huge amount of data that can be estimated as Terabytes, in a half day worldwide. Health sector also needs a

quality assurance mechanism for medical figures, such as ex-rays. On an average a normal size hospital generates about 1 million figures per year [6]. To get some better results we have to intellect with some better

We will have a literature review, because research in this domain is not yet properly done so we will have to conduct a research survey too. The research survey will consist of questionnaire related to complexities and practices used in industrial sector, this survey will be conducted from industrial sector engineers developers and the students who have at least a bachelor degree in computer science because they are the one who will be working on analytical softwares.

III.   LITERATURE REVIEW

Quality Assurance for data needs a huge amount of knowledge and it also needs a lot of effort because gathering such huge amount of data is not an easy task, but getting the useful data from the raw data is the difficult task, because results from data can be ambiguous many times. That is why we have to involve the people who are working on the quality assurance in the developing phases of the software. We have to make a quality assurance team who is responsible of the validity of the results produced by them. Adding Quality Assurance peoples in developing team can be beneficial because they will find and resolve quality related issues at earliest stages and this approach can save so much time for us, as a output, money and time can be saved Quality Assurance team must make plans for testing; such as test cases etc. This approach gives a clear view to Quality assurance team and with this strategy we can see a positive movement in our project testing side [8].

Testing in real time data is complex and money wasting technique, that's why to resemble devices is a better option and this technique is successful so far, but in case of data analytics we need real data, like in case of traffic analytics we cannot visualize it, though we can but it is not

a successful process. Recently Automated vehicle crashes has occurred and this minor incident has caused company so much loss. In result, for data analytics we need real time data so that beneficial output can be generated [9].

These techniques are very necessary because in many of the systems data sources are too much and data is in huge amount to check it manually. For a simple query for example how many bikes crossed this signal at 8:00 pm yesterday? Now this seems very simple in first sight but when we will be running this query manually it will take many hours may be a day or two in traditional systems, that's why we need to use especial systems and we need to concentrate on quality more [10].

Many developers have highlighted the quality assurance as a autonomous field and it is also mentioned that software quality assurance is introduced after hardware quality assurance and similarities between these two domains are also processed. Manual methods are not enough to do the necessary advancements; there need to do some changes [11].

IV.   METHODOLOGY

This research is dependent on the basis of literature review and the surveys conducted on the industrial sectors workers and students who are engaged in software development processes. This paper was written because the excess availability of the literature was not enough [12].

V.   Survey Responses

When have taken multiple surveys in these context that how much quality assurance on data analytics is important. Some of the people think that its not important but some of them think it is. We have done the surveys with the people who are engineers in the industry or the students who are having the background of Computer Science and have a know

how about the quality assurance. So we can get better results in our surveys. The graph summary in Fig. 1 showed that almost 60% are students, 30% are developers and remaining 10% are business analysts. We have taken people from every profession.
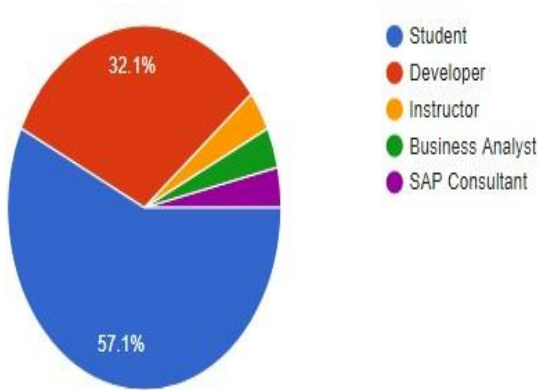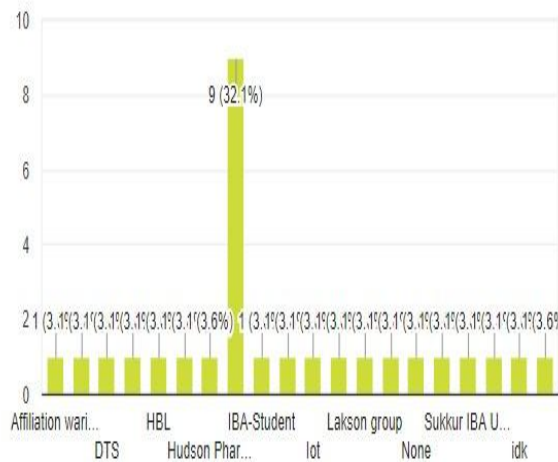

Fig. 1. Professional

Fig. 2. Affiliations with companies.

We also wanted to observe the affiliation of the participants that from which company they are or from which university they have studied to make sure about the productivity of participants by analyzing their background shown in fig 2. Developers from some reputable organizations were included like HBL, Lakson Group IBA Sukkur students. We can observe in this fig the

affiliation is on the horizontal line and and no of developers on the vertical end.
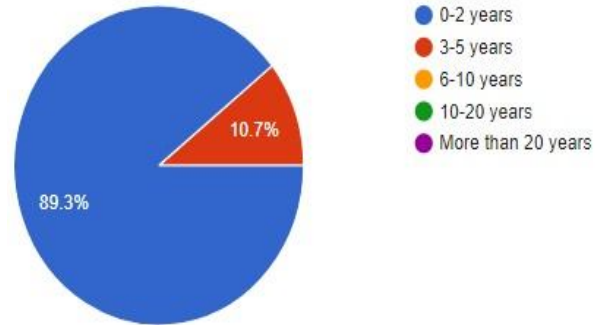

Fig. 3. Experience of developers in years.

Now that we have taken certains developers from these organizations we will see how much experience they have In the field of the computer science and how much they know about the quality assurance or how much work they have done in the field of quality assurance.

Despite about their experience in the field. Fig 3 shows the people who have upto 2 years of experience are almost 89% and people with a more experienced background are 10.7% respectively. We tried to reach people with more experienced background but we only got 2 to 3 responses.

Fig. 4. According to you, is quality assurance important?

As we have discussed earlier that do people think software quality assurance is important for any software or not. So in these surveys we have gathered developers from the multiple categories as we have discussed earlier, but in majority we have

the people of the final year students or the employees from the companies which are 96.4% and they think quality assurance is the key component in the software product. And only 3.6% thinks that its not important factor for any software product shown in fig 4.
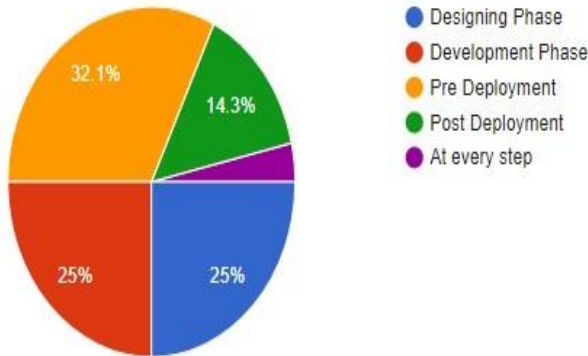

Fig. 5. At which stage of software development cycle, quality assurance group should interfere?

As we have seen the earlier results that majority is in the favor of doing quality assurance on the product but in what phase we should do this. So we ask the Q/A sessions with developers about this shown in fig 5. 32.1% of them said that we have to do it before the deployment. 14.3% said we should do it after the deployment phase. 25% said we should apply quality assurance in the designing phase and 25% said we should apply it in the development phase.
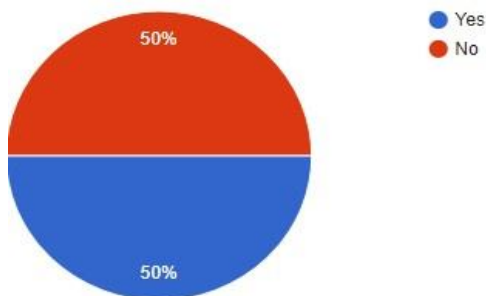

Fig.
6. Have you ever worked for the development of data analytical software.

We have asked the developers that if they have worked in the field of data analytics so we get the result as 50% overall shown in fig 6. 50% of them have worked in the field and have observed that quality assurance is really important for the field of data analytics.
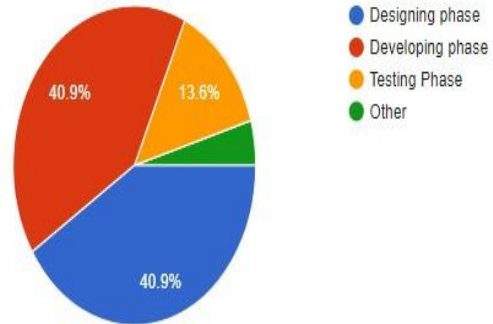

Fig. 7. In which phase these complexities occur in data analytics?

We asked a question just to examine that in which phase developers face the complexities occur in the analytical software. So 40.9% said it occurs in the development phase. And 2/5% said it comes in the designing phase. So the developers have to get ready in the initial stages of the development and designing of the softwares in fig 7 and should plan accordingly for the future work.

VI. Conclusion:

Quality assurance is the essential part of any software development model when it comes to the data analytics it is very hard to obtain results without the quality assurance because making of the data analytics softwares are more tough then the traditional softwares so we have to apply the quality assurance factor to the data analytics software. Some apply them in the development phase some of them in the initial stages and some apply after the deployment in the testing phases.

REFERENCE

[1]   M. Bruneforth, Martin and I. V. Mullis, Quality Assurance in Data Collection, vol. 10, M. O. Martin and I. V. Mullils, Eds., Boston: Center

for the Study of Testing, Evaluation, and Educational Policy, Boston College, 1996.

[2]   L. Stahl, "Quality Assurance Project Plan for Data Analysis Activities for the National Study of Chemical Residues in Lake Fish Tissue," U.S. Environmental Protection Agency, Office of Science and Technology, Washington, D.C. 20460, 2007.

[3]   F. Diko, Z. Alzoabi and m. Alnoukari, "Enhancing Education Quality Assurance Using Data Mining," 2016.

[4]   "Quality Assurance for Analytics: 4 Steps to Avoid Big Headaches," 22

August   2017. [Online].
Available:

https://www.qualitylogic.com/2017/08/22/quality-assurance-foranalytics/. [Accessed 04 April 2018].

[5]   H. Foidl and M. Felderer, "Data Science Challenges to Improve Quality Assurance of Internet of Things Applications," in International Symposium on Leveraging Applications of Formal Methods, 2016.

[6]   W. Raghupath and V. Raghupathi, "Big data analytics in healthcare: promise and potential," Health Information Science and System, p. 10, 2014.

[7]   R. Per and M. Höst, "Guidelines for conducting and reporting case study research in software engineering," Download PDF, April 2009.

[8]   V. S.Moustakis and L. Tsironis, "Knowledge Quality Assurance in Medical Data Mining," in Proceedings of International Conference on Information Quality Management., Chania, Greece, 1996.

[9]   C. Tao and J. Gao, "Quality Assurance for Big Data Application– Issues, Challenges, and Needs," National Natural Science China, p. 7, 2009.

[10]   S. Farooqui and W. Mahmood, "A survey of Pakistan's SQA Paractices: a Comparative Study," in 29th International Business Information Management Association Conference, Vienna, Austria, 2017.

[11]   F. Lambert, "Tesla Autopilot confuses markings toward barrier in recreation of fatal Model X crash at exact same location," 3 April 2018.   [Online].   Available: https://electrek.co/2018/04/03/tesla-autopilot-crashbarrier-markings-fatal-model-x-accident/. [Accessed 4 April 2018].

[12]   SeattleDataGuy, "Data Quality Is Not as Sexy As Data Science," 15

September   2017. [Online].
Available:

https://medium.com/@SeattleDataGuy/good-data-quality-is-key-forgreat-data-science-and-analytics-ccfa18d0fff8. [Accessed 4 April 2018].

[13]   F. J. Buckley and R. Poston, "Software Quality Assurance," IEEE Transactions on Software Engineering, pp. 36-41, 1984.

[14]   R. M. Groves, F. J. Flower Jr., M. P. Couper, J. M. Lepkowski, E. Singer and R. Tourangeau, Survey methodology, 2 ed., John Wiley & Sons, 2009, p. 488.