Impact of adopting best practices of agile methodologies (Scrum) in success of Big Data projects: Pakistan”

NAME: Agha Hassan Afzal Khan **ENROLLMENT:** 03-298171-003

SEMESTER: Spring-2018, MSPM - III **PROGRAM:** MS - Project Management

DEPARTMENT OF MANAGEMENT SCIENCES

BAHRIA UNIVERSITY LAHORE CAMPUS

RESEARCH THESIS

Principal Supervisor: Dr. Shahid Iqbal

COPYRIGHT

SUBMISSION FORM OF THESIS

Candidate Name: Agha Hassan Afzal Khan

I have submitted proposal of thesis for examination for the degree of MSPM.

Thesis Titled: "Impact of adopting best practices of agile methodologies (Scrum) in success of Big Data projects: Pakistan"

Candidate Signature: _____ **Date:** _____

Certificate of Principal Supervisor

I, Dr. Shahid Iqbal being the Principal Supervisor for the above student, certify that proposal is in a form suitable for examination and the candidate has perused his course in accordance with the Rules of the University.

Signature: **Date:** _____ _____

Recommendation of Examination

I recommend that the proposal be examined.

Principal Supervisor: _____ **Date:** _____

Not Recommendation of Examination

I recommend that the proposal may not be examined.

Principal Supervisor: _____ **Date:** _____

Statement by the Head Faculty/Department

I support the submission of the thesis of the above-named student for examination under the University Rules for higher degrees.

Signature: _____ **Date:** _____

BAHRIA UNIVERSITY, LAHORE

Statement by the Head Faculty/Department I support the submission of the thesis of the above-named student for examination under the University Rules for higher degrees.

Signature: _____ **Date:** _____

BAHRIA UNIVERSITY, LAHORE

APPROVAL SHEET SUBMISSION FOR HIGHER RESEARCH DEGREE THESIS

Candidate Name:	**Agha Hassan Afzal Khan**
Discipline:	Project Management
Faculty/Department:	Department of Management Sciences

I hereby certify that the above candidate's work, including the thesis, has been completed to my satisfaction and that the thesis is in a format and of an editorial standard recognized by the faculty/department as appropriate for examination.

Signature:

Principal Supervisor: **Dr. Shahid Iqbal**

Dated:

The undersigned certify that a pre-completion seminar, and overview and synthesis of major findings of the thesis, and the research is of a standard and extent appropriate for submission.

I have checked the candidate's thesis and its scope, format, editorial standards are recognized by the faculty/ department as appropriate.

Dean/Head of Faculty/Department

Dated:

DECLARATION OF AUTHENTICATION

I, Agha Hassan Afzal Khan, Student in the Department of Management Sciences, Bahria University, Lahore Campus, certify that the research work presented in the thesis is to the best of my knowledge, my own. All resources used, and help received in the preparation of this dissertation have been acknowledged. I hereby declare that I have not submitted this material, either in whole or in the part, for any other degree at this or other institution.

Name: **Agha Hassan Afzal Khan**

Signature: _____

ACKNOWLEDGMENTS

Thanks to Allah Almighty, the author of knowledge and wisdom, for his countless blessings. He has blessed me with the company of people who are unique in their own way and have been of immense help to me throughout this dissertation. The success and outcome of this dissertation required a lot of guidance.

I am highly indebted to my supervisor, Dr. Shahid Iqbal, for his continuous support and regular supervision for this research.

I am grateful for the sincere efforts of all my teachers of MS Project Management who taught and helped me to reach at this stage including (Mr. Shehzad, Mr. Ahsen Maqbool, Dr. Shahid, Mr. Shawaz Baluch, and Mr. Faisal Shehzad).

I am obliged to my fellows and my family for supporting me all the way. Without their support, this could not be possible.

DEDICATION

I would like to dedicate this hard work to the two most important people in my life:

- First, my best friend, my mother who has supported me throughout my life and has always trusted me to perform my best in every phase of my life.
- Second person I would dedicate this work is my father who has always prayed to Allah almighty for my success.

ABSTRACT

The main purpose of carrying out research on this topic is to identify the factors of an effective project management methodology that cause impact on project success in software houses of Pakistan where Big Data is involved. This research was dependent on the analysis of two independent variables (IV) which also can be counted as important features in success of big data projects i.e. agile methodologies & scrum practices and tailored scrum.

This research demonstrates the capacity in which above mentioned variables are significant for project success.

This research adopted quantitative research method. A total of 74 questionnaires were distributed, considering our boundaries to resources of project teams in companies such as software houses with a title of technical leads, project manager, senior manager, director and senior management. The response rate was 85% which is 63 in number.

Data analysis was carried out considering all mentioned independent variables and their influence distresses the delivery of successful project in software houses of Pakistan. The conclusions, significant findings and limitations are also discussed in this study.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

TABLE OF ABBREVIATIONS

ICT	Integration of Advanced Information and Communication Technology
ITES	Information Technology Enabled Service
BPO	Business Process Outsourcing
CEO	Chief Executive Officer
3D	Three Dimensional
TRG	The Resource Group
PMO	Project Management Office
IT	Information Technology
PMI	Project Management Institute
IPMA	International Project Management Association
CSF	Critical Success Factor
PM	Project Management
SPSS	Statistical Package for the Social Sciences
AMSP	Agile methodologies and Scrum practices
TS	Tailored Scrum
PS	Project Success
BDM	Big Data Manager
XP	Extreme Programming
IV	Independent Variable
DV	Dependent Variable
EFA	Exploratory Factor Analysis
KMO	Kaiser-Meyer-Olkin Test

ANOVA	Analysis of Variance
API	Application Program Interface

CHAPTER 1: INTRODUCTION

Deliberate the world devoid of data storage; every operation you make which involves very important data or data of students in university or data of employee in organization or anything in your life which can be documented is just vanished. Organizations would as a result lose the ability to extract valuable information and knowledge, perform detailed analyses, as well as provide new opportunities and advantages. “Anything ranging from customer names and addresses, to products available, to purchases made, to employees hired, etc. has become essential for day-to-day continuity”. (Gunasekaran, Yusuf, Adeleye, & Papadopoulos, 2017)

1.1 Research Background

In a quickly evolving Information technology and software development field, Business Intelligence methods and ICT solutions must become more agile to keep with the race (Sachan, 2017). The new world generates a confounding volume of data which require capacity to store, broadcast and compute this information remains to grow exponentially, with one estimate suggesting that the installed capacity to store information would reach 2.5 zettabytes (2.5 x 1,021 bytes) in 2012. International Data Corporation (IDC) research suggests that the world’s digital information is doubling every two years and will increase by fifty times between 2011 and 2020. The commercial sector is gradually manipulating these vast quantities of data in a variety of ways, from sophisticated market analysis that allows precisely targeted advertising, to the real-time analysis of financial trends for investment decisions and ultimately to completely new business models. (Couch & Robins, 2013)

The term ‘Big Data’ is used to describe these very large datasets and ‘Big Data analytics’ to refer to the process of seeking insights by combining and examining them. Against the backdrop of this huge growth in storage capacity, data collection and computation power, Big Data is arguably one of the most important global trends of the coming decade. (Couch & Robins, 2013)

Data Science and big data are evolving discipline that syndicates expertise across a range of spheres, including software development, data management and statistics. Data science and big data projects typically have an objective to identify relationships and causal associations, classify and predict events, identify patterns and anomalies, and infer probabilities, interest and sentiment.

Big Data is a related field, often thought of as a subgroup of data science, in that data science smears to large and small data sets and covers the end-to-end process of collecting, analyzing and communicating the results of the analysis.

1.2 Motivation

"With the growing ability to collect, store and analyze an ever-growing variety of data that is being generated with increasing frequency, the field of data science is increasing rapidly. As a new field, much has been written about the use of data science and algorithms that can generate useful results. In fact, many in the field, such as Chen, believe that data science research needs to continue to focus on analytics. Unfortunately, less has been written about how a group could best work together to execute a data science project. For example, in the field of data science, there is no known "best" process to do a data science project". (Franková, Drahošová, & Balco, 2016)
Having a well-defined repeatable practice of project management can help data science teams across a range of challenges, including understanding who needs to be included as a stakeholder in the process, selecting an appropriate data architecture / technical infrastructure, determining the appropriate analytical techniques and validating the results. Without a well-defined process, these tasks would still likely get addressed, but the team might forget a step or not learn from their own experience. The individual attributes characteristic of Big Data suggests the definition "Big Data are large-volume, high speed and widely diverse information assets that require new forms of processing to allow improved decision making, greater insight into the discovery and optimization of processes"
The aim of this work is to document agile approach to project management and suggests ways of using it in projects related to Big Data management. Project management can be understood as "the planning, delegation, monitoring and control of all aspects of the project and encourage the participants to achieve the project objectives within the expected performance parameters set for time, cost, quality, scope, benefits and risks". Similarly, standard ISO 1006, defines project management as "the planning, organizing, monitoring, managing and reporting of all aspects of the project and the motivation of all project participants to achieve the project objectives." In principle, it is the implementation and control of the actions that are necessary during project

implementation in order to achieve the project objectives within predefined limits (Gunasekaran, Yusuf, Adeleye, & Papadopoulos, 2017)

1.3 Research Scope

This proposition is focused to the ICT and Software companies which mainly work on big data projects in Pakistan that want to adopt Agile Methodologies (Scrum), Big Data Managers (BDMs) and Project Managers in such companies who wants to implement Scrum, and Big Data engineers who are involved in the delivery of successful and unsuccessful projects with scrum. This report also a preliminary point for research of Agile Methodologies (Scrum) in context of Pakistan.

1.4 Problem Statement

There are several companies (known companies) in Pakistan, are working on Big Data with association of US based companies like Facebook, Amazon, Ebay and others. Teradata, Northbay Solutions, Arbisoft, Infotech, NexDegree, I2c inc., Arpatech, Analytics, Red Buffer, Contour Softwares, Softronic Systems, TRG (Afiniti) and Speridian Technologies are known companies which are working on Data Science projects in Pakistan. And we can analyze the more companies are coming and investing on Big Data and Data Sciences projects in Pakistan. (Industries, 2017) Projects of Big Data are not running in Pakistan only, but also, they are running in worldwide. The objective of the thesis which is significant influences in the management of Big Data projects was to regulate whether the management of Big Data projects can apply the value of the agile manifesto or if at all agile approach is applicable in managing these projects (Franková, Drahošová, & Balco, 2016).

1.5 Research Questions

1. "Is it possible to use an agile approach to managing Big Data projects"?
2. What are the chances of project success after implementing Agile?

1.6 Research Objective

Below is the sum of major objectives for this research:

1. To identify agile approach is suitable for the management of Big Data Projects.
2. To show a project success after implementing Agile.

1.7 Gap Analysis

The following are the key gaps that exist in context of research in this area:

1. There is very less work performed in this area of research in Pakistan the only work performed in context of technology.
2. It is analysed that in Pakistani culture we do not follow Agile Methodologies (Scrum) in terms of Big Data Projects, due to which projects delivers late, hence are unable to produce profitable results.

1.8 Structure of Thesis

This thesis is organised in the subsequent means: this first chapter delivers roughly background to the research approved out and described main motivations of doing this research and deliberated problem statement in such way which communicate and make interaction with the reader.

The second chapter delivers a detailed overview of literature presented on different sources. It mainly focuses on project management tools i.e. agile methodologies and scrum. It also deliberates many methodologies which measures the efficiency of agile methodologies in terms of software houses and ICT companies.

Third chapter tells us about the research methodologies which we adopted and discussed why we adopted. Also discussed the population and sample size.

Forth chapter is completely on the testable hypothesis which we concluded from the literature review. Likewise, analysis of conducted surveys question by question and also relationship between independent variables and dependent variable. The result of the analysis also shown in the same chapter.

Fifth chapter is on conclusion and results which we measure from the survey and also describes research finding, practical implications and further recommendations.

And in the end, we have references and appendix with questionnaire which were conducted to proof the hypothesis.

CHAPTER 2: LITERATURE REVIEW

Data is the building block upon which any organization thrives. While there has been some research on the challenges of doing Big Data projects, this has focused on the technical challenges in executing the projects. For example: Finance focused on storage and data management challenges. Similarly, Katal discussed the challenges caused by the rapid growth in the data that is stored and analyzed, which is outstripping the computing resources and tools available to analyze the data. Unfortunately, there has not been much written on how to ensure data teams operate effectively and efficiently. (Gunasekaran, Yusuf, Adeleye, & Papadopoulos, 2017)

2.1 Organization Performances

It is observed that organization consists of several departments such as marketing, finance, project management and many others. The results or the outcome that is obtained from these departments have a direct or indirect impact on the organization performances. (Selig, 2015)

Indicated that if the organization department is not working well there is a high probability that key objectives or goals would not be achieved, and the organization performances will be impacted. So, throughout the globe, there is a number of an organization that has investigated a large number of their budget percentages to improve the performances and outcome supporting the argument from (Selig, 2015).

2.2 Big Data processes

Nada and Ahmed deliberated, enormous of data have become accessible on hand to decisions fabricators. Big Data mentions to data sets which are not only big but high in variety, velocity and veracity, which makes them difficult to handle using traditional tools and techniques. Due to the rapid growth of such data, solutions need to be studied and provided to handle and extract value and knowledge from these datasets (Elragal & Elgendy, 2014). With the increase in storage capabilities and methods of data collection, huge amounts of data have become easily available. Every second, more and more data are being created and needs to be stored and analyzed to extract value. Furthermore, data has become cheaper to store, so organizations need to get as much value as possible from the huge amounts of stored data. (Elragal & Elgendy, 2014)

The size, variety, and express change of such data require a new type of big data analytics, as well as different storage and analysis methods. (Cleland, 2004). Henceforth, speeding up company's capability to generate in-sights and business value from the data which is agile and developed or gathered by the mixed team of business and IT experts. (Sachan, 2017)

The one stated by Sachan for getting the data agile is Data Lake. A data lake is a repository for all structured and unstructured business information collected from the myriad systems located in different business units and functions in a company. It would include current and archived data and, using APIs, could be enriched with information from external providers such as Face-book or Twitter. Data in Data Lake remain in their original formats. Data Lake is compatible with standard data-discovery tools. To fully utilize the data lake, it must be integrated within a company's existing data architecture and serves as the primary source of information. The biggest benefit of Data Lake is that data stored there is reliable which can be accessed anytime for analysis by broad range of applications. (Sachan, 2017)

Sachan also explained the Mckinsey's points for adopting agility. (Sachan, 2017)

- Use pilot projects to test out the unique features and ability.
- Empower agile teams
- Updated IT infrastructure
- Enhanced and effective communications amongst all business unit
- Devising KPIs for company's internal growth and performance review

2.3 Project Management

"Project Management, is application of knowledge, skills, tools, and techniques to project activities to meet the project requirements" (PMI).

Many have attempted to define project management in very early periods. "Project Management is the application of tools and techniques to direct the use of diverse resources toward the

accomplishment of a unique, complex, one-time task within time, cost and quality constraints" (Oisen, 1971).

Project management is the process of planning, scheduling, and controlling of project activities to meet project objectives (Lewis, 1995)

Project management is the art and science of converting vision into reality (Turner, 1996). This description does not include the criteria to measure project management.

2.4 Agile approach to software development

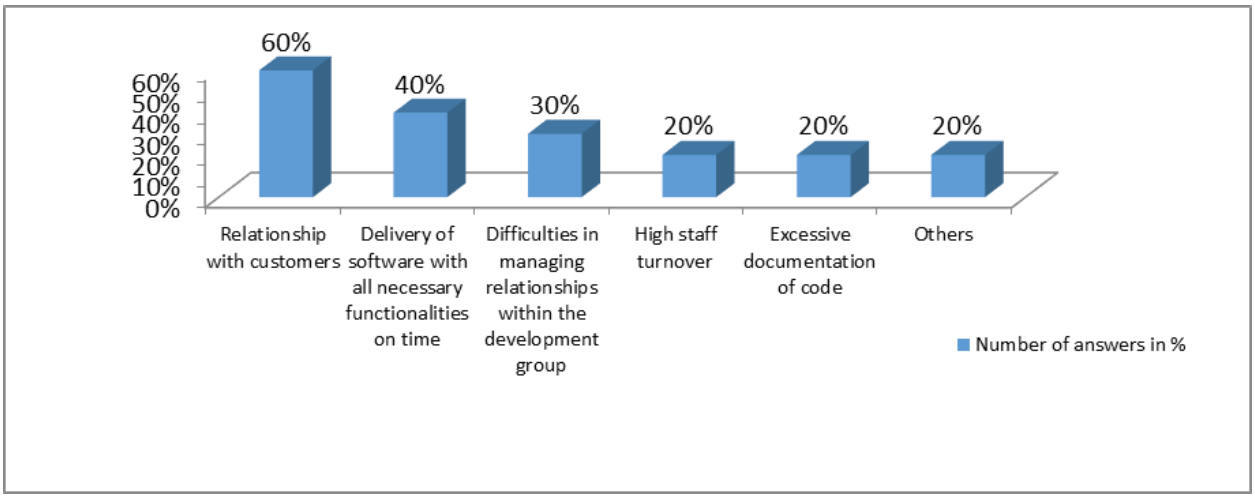
For addition to agile software development, this software development management approach is gradually focusing on an aggressive approach. The words Ophthalmic became one of the modern concepts of the eighth century. When we came to the popularity of aggressive thinking and the ultimate scrum framework, we were in terms of searching for English through the search engine in the thorough Google search engine.

What 'agile' refers to? The clearing of this term is not in the development environment. The beginning of using agility terminology in manufacturing "to successfully sell in low delivery, short delivery times and adaptation for user needs that provide increased prices for high-quality products. The customer has been praised as "capacity." It is stated that 'agility is the freedom to adopt a key priority method for software development and the methods of fulfilling the needs of a particular project', the success of aggressive development practice is the latest research version of the epilepsy. On the one hand, the business software in the field is dealing with a supplier, confirmed by the state's next chapter of the sponsor. More than 3,500 people had said, primarily evaluated% of the software using 88 prompt developments in the field of software development. The annual growth rate of 2% over the last two years also reflects the increasing popularity of aggregate thinking. (Franková, Drahošová, & Balco, 2016)

2.4.1 Agile Methods

There are instinctive methods "(without any defects) working software creation efficient software are trying to focus on the main goal of development 'efficient software development means that we have' software ' For a given situation, and requirements will be found in different varieties. The

term 'sufficient' is not necessarily necessary for something to grow in progress like 'fixed' and is sufficient for a given repeat.". (Franková, Drahošová, & Balco, 2016)

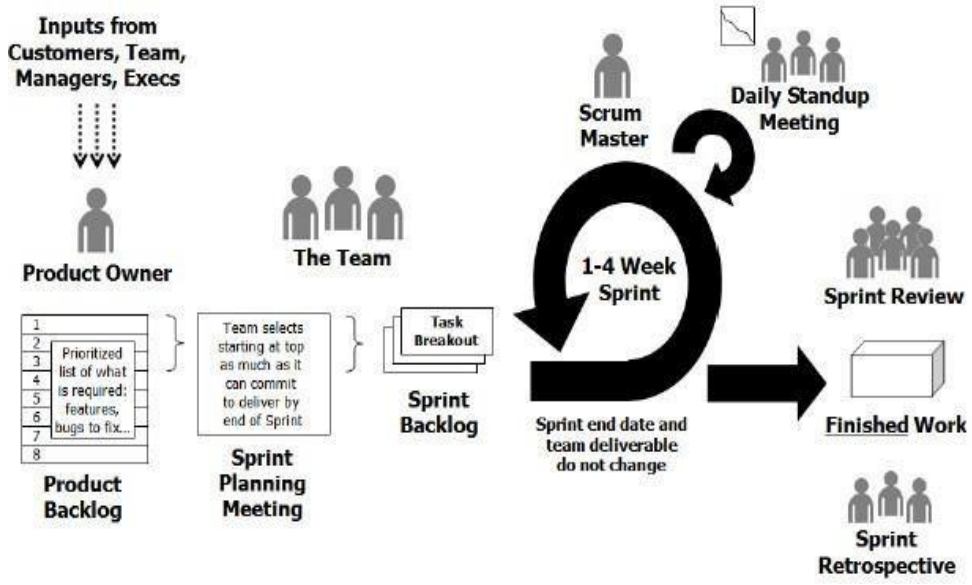


"The transition to agile methods during the software development mitigated these problems".

2.5 Scrum practices

There are multiple methodologies which are implemented in various software houses but we have found most productive and easiest methodology are: Agile data warehousing and scrum. Scrum is consisted of product backlog, features and its related user stories, sprint backlog, sprint and daily scrum (Sachan, 2017).

We have also observed in different context, as said by Schwaber, there are list of methodologies and practices to use scrum and agile methodologies and he also said that we can divide it into eight different areas: The ScrumMaster, Product owner, Product Backlog, Scrum team, Sprints, daily Scrum meetings, Sprint planning meeting, Sprint review meetings and Sprint retrospective. Sutherland and Schwaber provide the following graphical summary which gives the complete process of scrum practices used in software houses.



An agile BI solution utilizing **Scrum process** (Sachan, 2017) Schwaber

2.5.1 ScrumMaster

The ScrumMaster is the being accountable for the Scrum process, its correct execution, and the extension of its benefits. In our culture, we consider the scrum master as a leader but Schwaber et al. also point it out: "The ScrumMaster is not the front-runner of the team because the Scrum squad is self-organizing. Instead, the ScrumMaster acts as a facilitator for issues resolution and communication, rather than as a manager controlling the team." Thus, he is accountable for eliminating any barriers between the development teams, the Product Owner and customers so that the team can directly drive development in their comfort zone. He is also answerable for presentation the Product Owner how to use Scrum to maximize the project return on investment (ROI) and meet the project's objectives.

2.5.2 Product Backlog

The Product Backlog is nothing but the list of tasks which needs to be further prioritized according to the project requirement with respect to estimated time to execute them into product objective. We usually estimate in days but in beauty of scrum we can estimates in hours, which means every

task is associated with its estimated time and story point. Story points are just numbers in terms of task complexities.

2.5.3 Sprint

Usually sprints are measured in period of time, when agile team is focusing on the tasks of the sprints commitment. While in this period, the agile team have full authority over their actions, there is not any external influence from the product owner, or anybody else is allowed.

Each Sprint consists of two fundamentals, the Sprint objectives and the Sprint Backlog. Sprint objectives are most important for every scrum team member and are meant to be top priority items of the product backlog. This is the actual objective which needs to be met through implementation of the product backlog. Once you create and define the complete sprint goal, the agile team discuss initially with the product owner to achieve the required goal with in the time.

Each sprint lasts from one week till four weeks, depends on the length of the product backlog.

2.5.4 Product Owner

The Product Owner is the person which manage and control the Product Backlog and requirements. It is his or her responsibility to maintain the product backlog in such manner that everyone can see the list of features, stories and tasks. Actual responsible of the delivery of the successful project is the Product Owner. In other words, he or she accountable to manage the requirements and make it understand then convey it to the team accordingly.

This also eliminates contradictory lists being created and ensures that the teams know who to listen to. To the successful product owner, all stakeholders of the project must listen him or her and respect his or her decisions. His actions and planned decisions are always visible on the product backlog. As soon as the product backlog is created, the product owner starts working with the developers to put the estimation against each task. Once it is completed then scrum team members

checks and picks the tasks related to them and product owner ensures each task must be completed according to the estimation and supervised the team accordingly.

2.5.5 Scrum team

To deliver the successful product with scrum it is deeply reliant to see how each team is constructed. As suggested by Sachan in 2017, that ideal team to deliver the successful product it should be around 7-9 in big data which include: product owner and Scrum master as well (Sachan, 2017). Scrum team is responsible and work on the product backlog. These are also main stakeholders of that particular project.

Collaboration between the scrum team is very important, after their opinions developers work on the product backlog. Developers can also put their opinions to do their work efficiently. If there is a gap between collaboration and communication, it may also fall into delivery of unsuccessful project. It is important to keep the project in-lined through their communication

2.5.6 Daily Scrum meetings

Daily Scrum Meeting is usually called by different names: Stand up meeting, short meeting and daily status meeting depends on the organizational culture. As per scrum alliance it is necessary to have daily scrum meeting in success of delivery of successful project. In this meeting, each scrum team member gives status report to the product owner and the challenges faced by him or her doing this. Daily scrum meeting time could be the start of the shift or the end of the shift. If it is start of the shift then developer, SQA and other will let everyone what have been done previous day and what needs to be done by today and how he or she will do and what are the risks associated to that tasks and ripple effect. And if it is in the end of shift then he or she is responsible to convey the status report of all his work done and challenges associated to achieving that task and what further need to be done by tomorrow and how he or she will do it.

By the help of daily scrum meeting, it not only improves the communication but also eliminates other meetings and improves the overall product knowledge. To conduct the daily scrum meeting it is the responsibility of the SCRUM MASTER to coordinate every member of the team.

Daily scrum meeting also encourages to everyone in front of all and improves everyone project knowledge.

2.5.7 Sprint planning meeting

The most important meeting of every project manager which last 1 day usually 8 hours before the starts of the sprint. During this meeting, the product owner usually the project manager presents the product backlog with highest priority items to the execution, and quality and control team. Now this is the meeting where product owner and developers sit together and estimates the sprint tasks and see how much objectives of the product backlog they will achieve during the upcoming sprint. In this meeting, suggestions are open from everyone to achieve maximum product backlog. After that team gives commitment and acknowledgement to the product owner in completion of objective. This could be done by estimating the tasks and gives story points to each story and task accordingly.

2.5.8 Sprint review meeting

This meeting conducted at the end of every Sprint, usually lasts 4 hours where team gives demo to the product owner and the client (major stakeholder), what it was able to accomplish during the sprint. In this meeting, only completed and main functionality of the project can be demonstrated.

There was no experimental (i.e. subjective or systematic) indication on the Scrum acceptance challenges associated to the Sprint review.

2.5.9 Sprint retrospective meeting

A Sprint retrospective meeting is also conducted end of every sprint with the agenda: What has been accomplished? What challenges faced during the sprint? How we can make it better in achieving that goal. This meeting last usually 3 hours which could also determines what might make next Sprint more productive and enjoyable.

2.6 The Scrum processes

Scrum is referred by the iterative and incremental software development technique, which is further motivated by Product backlog list. As mentioned earlier product backlog list contains all active project requirements. The product owner is the responsible of the managing product backlog. Product owner is the only person which is authorized to change and prioritize the requirements accordingly.

Scrum life cycle of development works between the iteration of work and sprints which are usually seven to 30 days in length. As discussed in above, each sprint starts with creating product backlog which is further and deeply discussed in sprint planning meetings, usually consist of scrum team includes product owner and scrum master. Sprint Backlog created from the Product Backlog, each requirement is visible and defined as story. And each story further decomposed into tasks and each task will be further estimated between 1 to 8 hours.

Each functionality is defined into product backlog but non-functional requirements are not mentioned into product backlog but the Sprint backlog in the form of unpatented tasks. As per Frankova and Balco, he explained that agile team is self-managing, cross functional and self-organizing, and they are responsible for reckoning out how to execute each task of product backlog into the achieving of the goal by the end of the sprint (Franková, Drahošová, & Balco, 2016)

The Project Manager is in charge for handling the Scrum development so that it turns in an organization's culture and still distributes the predictable assistances, and for guaranteeing that

everybody tails Scrum rules and methods. The Project Manager leads a 10 to 15 -minute daily scrum meeting which could be the start of the shift or the end of the shift. If it is start of the shift then developer, SQA and other will let everyone what have been done previous day and what needs to be done by today and how he or she will do and what are the risks associated to that tasks and ripple effect. And if it is in the end of shift then he or she is responsible to convey the status report of all his work done and challenges associated to achieving that task and what further need to be done by tomorrow and how he or she will do it. It is also the responsibility of project manager to ensure whether it should be start of the day or end of the day.

A Sprint retrospective meeting is also conducted end of every sprint with the agenda: What has been accomplished? What challenges faced during the sprint? How we can make it better in achieving that goals. Sprint retrospective meeting in command to guarantee continuous improvement.

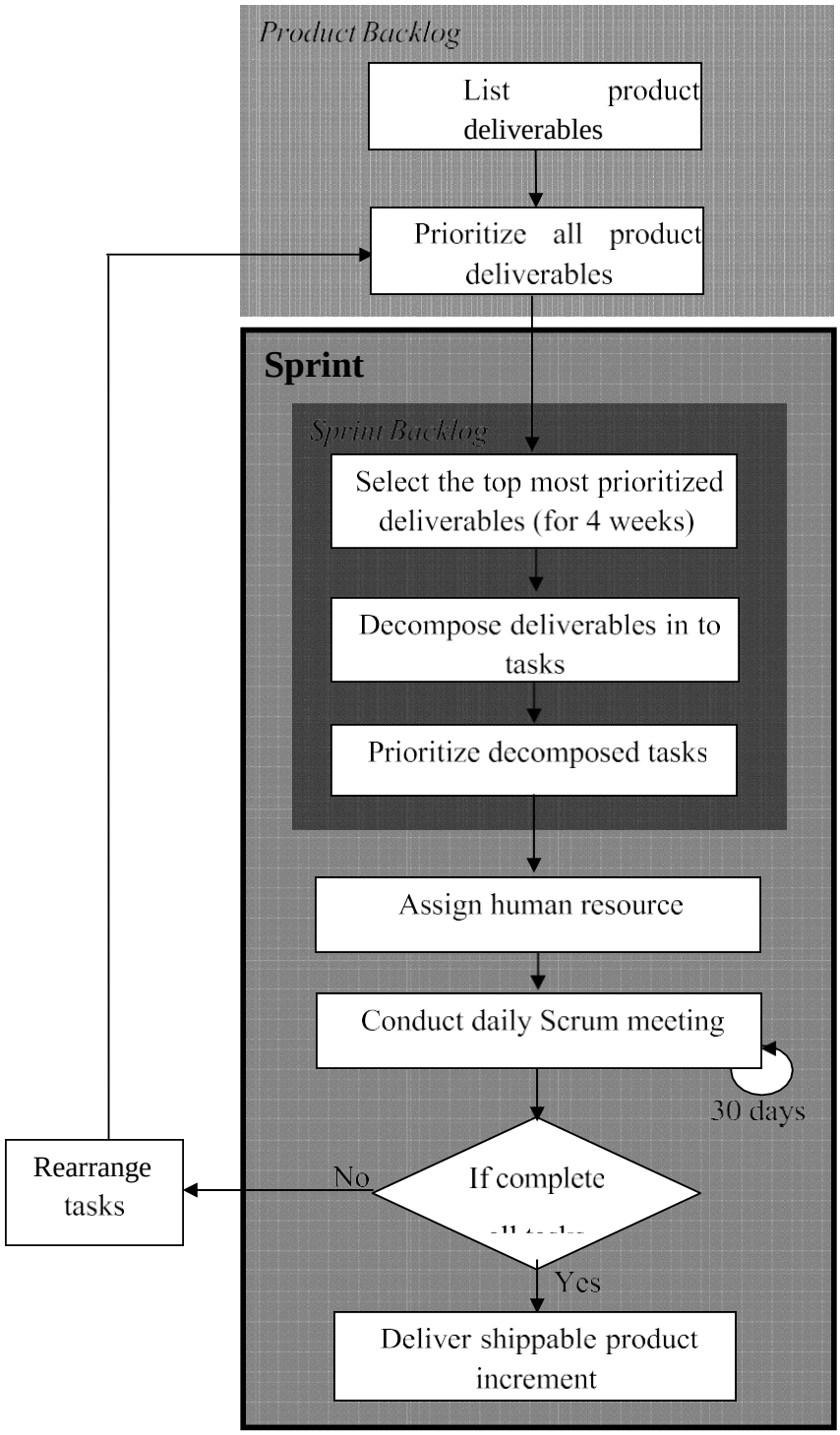


Figure - Scrum Flow 1 (Franková, Drahošová, & Balco, 2016)

2.7 Effectiveness of Agile Methodology (Scrum)

There are many case studies which prove that Scrum is a productive and effective framework to manage ICT and big data development teams under the different contexts. Similarly, this paper also ensures the effectiveness of agile methodology (Scrum) in big data companies of Pakistan.

Janoff and Rising originate that teams which are in between 9-12 (includes product owner and project manager) can be adoptable and flexible in applying an suitable variant of Scrum. Therefore, they suggested, it is difficult for the large and complex team structures to adopt Scrum and scrum is also not approaching the same. According to the Scrum alliance it is recommended to use the small teams not more than 10 members in a single team. (Industries, 2017) . "Crowston and Shamshurin have supported out a longitudinal ethnographic case study in industry settings and thoroughly shadowed how Scrum can be used in multi-project environment and what challenges to be aware of during the adoption of Scrum in such environments". "There are ten challenges identified in their work. (Saltz, Shamshurin, & Crowston, 2017)

- "Engaging an over stress on the Scrum method and practices
- The Project Manager considerate only about the entities and communications (and disregarding the process)
- A absence of clear organization opportunities and actions
- Too much maintenance and bug fixing discouragement the team throughput and self-esteem
- Appropriate Scrum and short repetitions into examination intensive teamwork
- Over specialization discouragement collaboration
- Over distinctiveness
- Promising to too much
- Trouble in tracking growth and by means of the results of the tracing

- Management meddlesome too much”

2.8 Challenges in Scrum

Although many case studies have proven that scrum software development has been an effective framework for organizing under several contexts, only a limited number of students who point out the scrums of challenges. The second biggest problem in the “too many meetings” was taught especially with the scrum that revealed a study conducted on Microsoft. One of the root roles claimed that inefficient meetings which are not effective and worthy, especially when unimportant a project manager, is not focused on, is enough to run the meeting in a timely manner. There is a problem of another problem so that all team members are getting together on a fixed time without failing every day and time.

Another big challenge with the scrum is that when a scrum team has challenge, they are more likely to change scrumming ways instead of changing themselves. When the supply of team spirits is difficult to supply, for example, the team may have decided to increase the sprint period, which they have ever been out of time. But then the team is ever estimating and doing a better job of managing them time will be able to know. Thus, coaching and assistance can be an experienced project manager, without the organization just scaring its own weaknesses and lengths of scary mortality, and skeleton in a small image, weakening the original benefits.

The most difficult part of implementing the scrum in a regular environment is a good development of a good product. Put one time and effort in a stored place and it reduces the amount involved in the development of the user's story for each stored item. So, when adopting an aggressive method, collaboration with the collaborative, cross-functional style is one of the key issues of cultural change from isolated expert work. Generally, the second is the fault according to given that a practice is encouraged or just because the scrum is not particularly needed. For example, the scrum product owner does not need to set up a long-term strategy for this product, nor more experienced

engineers need engineers to take advice on complex technical issues. Scrum drops people involved in making it the right decision. In most cases, both of these behaviors (many with others) are well-informed.

As a try to micromanage. Processors are astonishing to see using Agile processes. However, processes like scrum and XP. Etc. developers often communicate with their managers for short periods. In Microsoft the fear of micromanagement is seen as one of the reasons why daily Stand up meetings can be ineffective. (Saltz, Shamshurin, & Crowston, 2017)

2.9 The definition of the Big Data

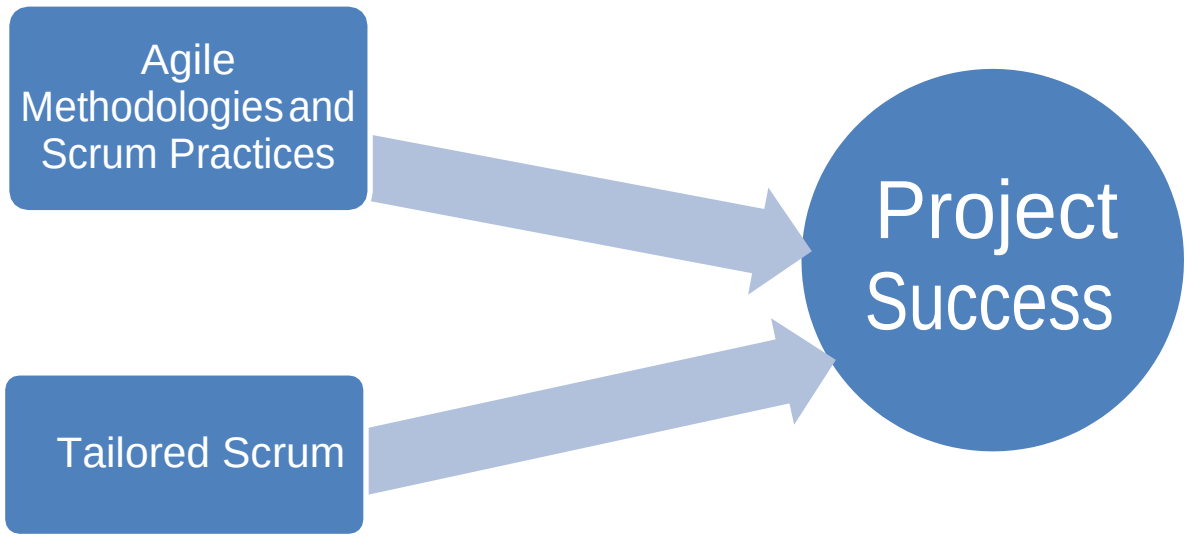
Big Data Term We are usually hard to portray a large, diverse and complicated structure and background of the background, analysis and other processes and results that can introduce as 'audios files on a large scale'. Can do It is not possible to determine the size limit for bug data clearly. Big included 'data files' capabilities used to handle their data with a size exceedingly, used and managed to use software tool in the acceptable total time. Large size scales are constantly increasing the target of a few dozen terabytes to many petabytes of data in one data set.

Individual entity Big Data definition shows 'Big data are required for new decision-making, better decision-making processes, high-speed and widely-diversified information processing processes, allowing maximum insights to be discovered. And (mathematics) process'. The first part of its formation shows that the data from the three key components of the data is from the traditional analysis: volume, speed, and varieties. The analyst Doug Laney 18 shows its cool 3-dimensional framework for the definition of Big Data, often the '3VS' model, which is recently called "5VS Model" of Capacity and Extension. . The first character of this name is derived from 'V' from the components of the model (speed, speed, variety and current). Let's explain more detail which is whether it is suitable for adoption of a large data view for analytical analyzes to determine the components of a common test-made framework.. (Franková, Drahošová, & Balco, 2016) (Gunasekaran, Yusuf, Adeleye, & Papadopoulos, 2017) (Sachan, 2017)

2.10 Theoretical Framework

To check out the best practices, everyone must understand what Big data project is and its aspects, foundations, related terminologies and frameworks.

Following is the key theoretical framework that will be utilized in the context of this research:



2.11 Variables Description

Independent variables presented in hypothetical framework are adapted from present research of Dai and Wells (2004) on project management office and its features. All two variables in context of deliver successful projects are briefly described below:

2.11.1 Agile Methodologies and Scrum Practices

A project manager can develop certain set of standards and adopt as per the culture to deliver the successful project. Agile provides you such guidelines which helps the agile team and scrum team to not to become a barrier in way of creativity of a team. But helps them to perform their associated tasks and stories according to plan. It should provide assistance in regard to basic project management practices such as development of proposal, charter, scope, project plans and provide assistance in risk assessment and documentation.

2.11.2 Tailored Scrum

Tailoring scrum as per our requirements helps the project manager in successfully deliver of sprint. Through tailored scrum we are obliged to not to deliver retrospective every end of sprint but deliver the successful tasks which needs to be completed by the end of Sprint. As in the end, client is more concerned with the output of the sprint. There are certain times, tasks are not completed on time due to some associated risks with the help of scrum we planned and move uncompleted tasks before the end of current sprint to the next sprint.

2.12 Relationship

Based on the theoretical framework it is analyzed that we conducted two factor that are defined as dependent and independent. However, during the analysis Agile methodologies & Scrum Practices and Tailored Scrum are considered under the umbrella of independent and on the other side the Project Success will be considered in the context of dependent variable. It is analyzed that the entire variable that are driven from the context of this research are adapted from (Sachan, 2017). Through the execution of this research, we will be focusing that on finding the negative or positive relationship between the dependent and independent variable.

2.13 Summary

Hence, we can conclude our literature review in one paragraph that: agile software methodologies have emerged over the years and the position that scrum holds among agile methodologies. Scrum helps software development organizations patterns and encourages them to cast the methodologies

according to their organizational culture. We all know that scrum and agile both are specially designing for those projects where we have high numbers of change requests. There are other methodologies like XP, incremental and in Big Data projects we can see there is rapid changes of business requirement which leads to deliver unsuccessful projects in terms of SCOPE, TIME, COST and QUALITY. “The practices to follow in order to use SCRUM, it is divided into different areas like, Scrum Master (Project Manager), product backlog, sprints, product owner, daily scrum meetings, sprint planning meetings, sprint reviews and sprint retrospectives”. By using all these practices, we could see the impact has been changed to successful projects. It is also observed that too much meetings also lead to delays which is further carried through tailoring scrum according to the project delivery. With these IVs (Agile methodologies & Scrum Practices and Tailored Scrum) we can derive relationship between its dependent Variable (Success project). To prove this we have choose different companies which mainly focused on Big Data projects, with the sample size of 60 which is 1.5% of total population that is approximately 4000.

CHAPTER 3: RESEARCH DESIGN & METHODOLOGY

3.1 RESEARCH DESIGN

Designing research helps the researchers to plan a process to work and manage the research in appropriate manner.

3.2 Population and Sample Size

It is observed that there are several Software houses and ICT companies that are working on Big Data in Pakistan. (Quora, 2017) However, it is quite complex to identify the population so based on the existing data, it is observed that around about the population of 4,000 is considered (Industries, 2017). If the population of software houses employees are around about 4,000 and the sample size for survey is 1.85% of population i.e. 74.

For equitable data collection against above mention sample size, only employees with designation of BDM and above from the technical teams who have direct dealings with clients and more accountable in terms of successful delivery. A total of 74 questionnaires were dispersed among different software houses all over Pakistan with help of their relevant HR departments, out of which 63 questionnaires were acknowledged back which shows a response rate of 85%. Out of 63 respondents, only 60 respondents were able to fill the questionnaire completely whereas 3 respondents were unable to complete the survey. Therefore, sample size of 60 respondents was taken into account of data analysis.

3.3 Data Collection

In the context of this research, both the primary and secondary was considered. However, in primary research we have used the questionnaire they will directly help us in analyzing the Impacts of agile methodologies scrum and the key problem that people working in Software houses and ICT industry of Pakistan are facing in terms of project outcomes. On another side we have also considered journal and articles that are published and focus on the area of Scrum & agile methodologies and its different sort of impact.

3.4 Tool for Data Collection

A survey method is used for primary data collection. Questionnaire was dispersed electronically via online Google form and physically to target respondents. For better understanding, required instructions to fill survey and brief detail regarding research were mentioned on the questionnaire.

3.5 Research Instrument

The main objective was to get a questionnaire that defines the impacts of all independent variables on dependent variable in context of delivery of successful projects. For this purpose, the items of the instrument were adapted from existing researches in context of Scrum and agile methodologies in Big Data.

Once questionnaire was designed, it was sent to PMO department of known Big Data Company "North bay Solutions" to check whether the questionnaire justifies the idea of this research. Director and Manager of PMO gave positive response.

3.6 Scale for Measurement

A five-point-Likert scale was used for this study. "Likert scale is one of the most commonly used scale in surveys" (Sekaran, 2006).

3.7 Research Approach

Research approach can be defined as procedure employed by the researcher for analysis and representation of collected data.

For this study, quantitative research was used.

CHAPTER 4: RESULTS AND DISCUSSIONS

Statistical data analysis was performed to check the relationship between independent and dependent variables. Abbreviations used for all variables related to this study as follows:

Abbreviations	Variables
AMSP	Agile Methodologies and Scrum Practices
TS	Tailored Scrum
IV	Independent Variable
DV	Dependent Variable

4.1 Demographic Data Analysis

Data was gathered from technical team resources working in various Software houses throughout the Pakistan. Suggested sample size to be considered for analysis was minimum 63. After removing incomplete surveys, 60 surveys were considered acceptable for analysis. Below are the demographic stats for gathered data.

4.1.1 Respondents: Category Gender

Table 4.1 and figure 4.1 represent the information about respondents from perspective of their gender. Below stats show that 17 respondents were female and 43 were male which is approximately 28% and 70% respectively.

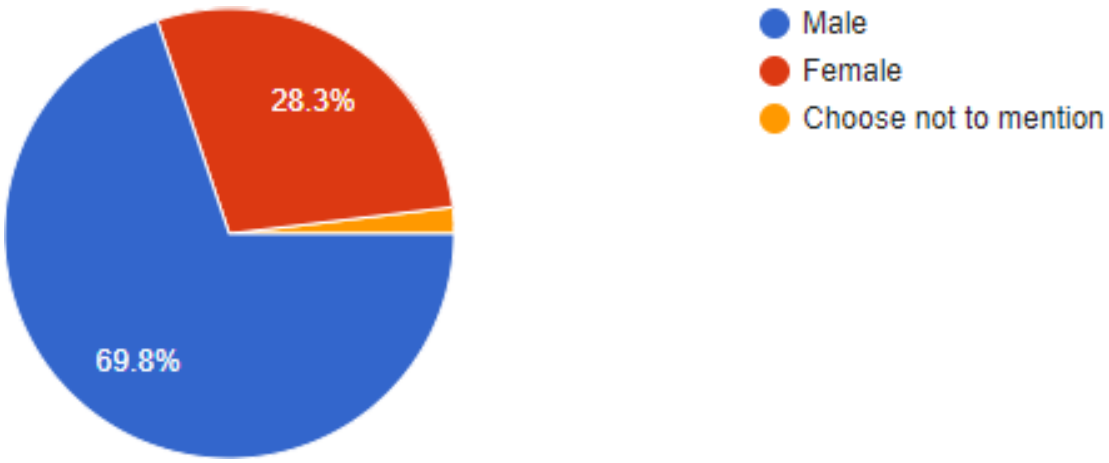
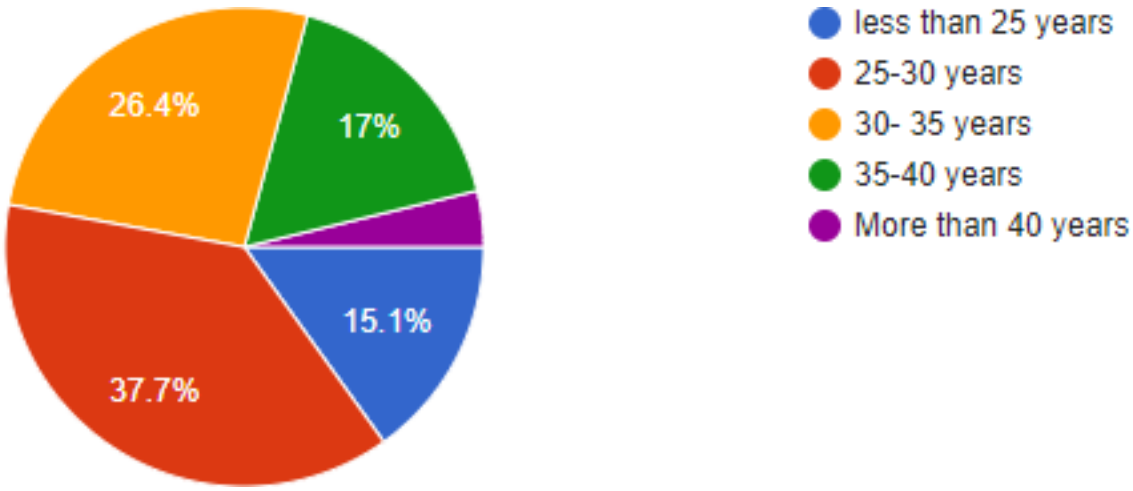


Figure 4.1 Respondent: Category Gender

4.1.2 Respondent: Category Age

Figure 4.2 represent the information about respondents from perspective of their age. Below stats show that 46 respondents were from age group 25 years and above 8 were from age group less than 25 years, which is approximately 85% and 15% respectively of the total sample.

Figure 4.2 Respondent: Category Age



4.1.3 Respondent: Category Qualification

Figure 4.3 represent the information about respondents from perspective of their qualification. Below stats show that respondents with bachelor's and master's degree represent 53% and 45 % are of the total sample.

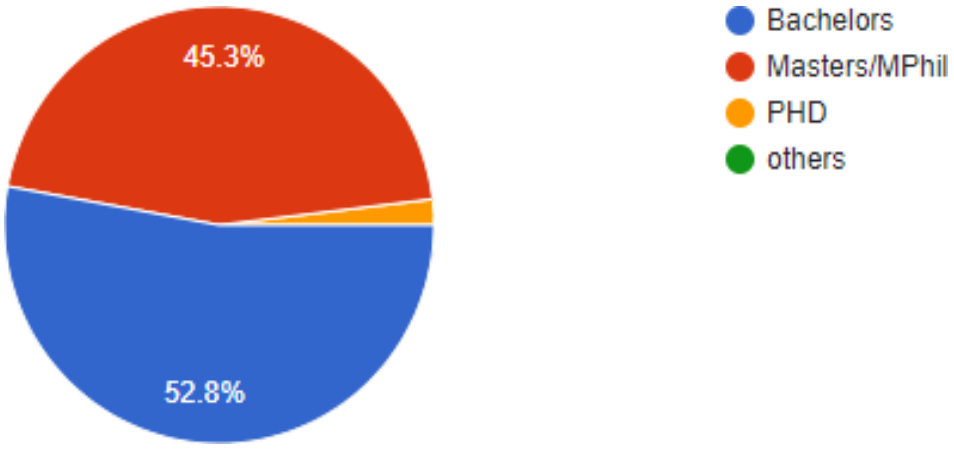


Figure 4.3 Respondent: Category Qualification

4.1.4 Respondent: Category Experience of in Software house

Figure 4.4 represent the information about respondents from perspective of their job experience. Below stats show that approximately 85% of population of population have more than 2 years of experience whereas 15% have less than 2 years of experience in field of development in software houses

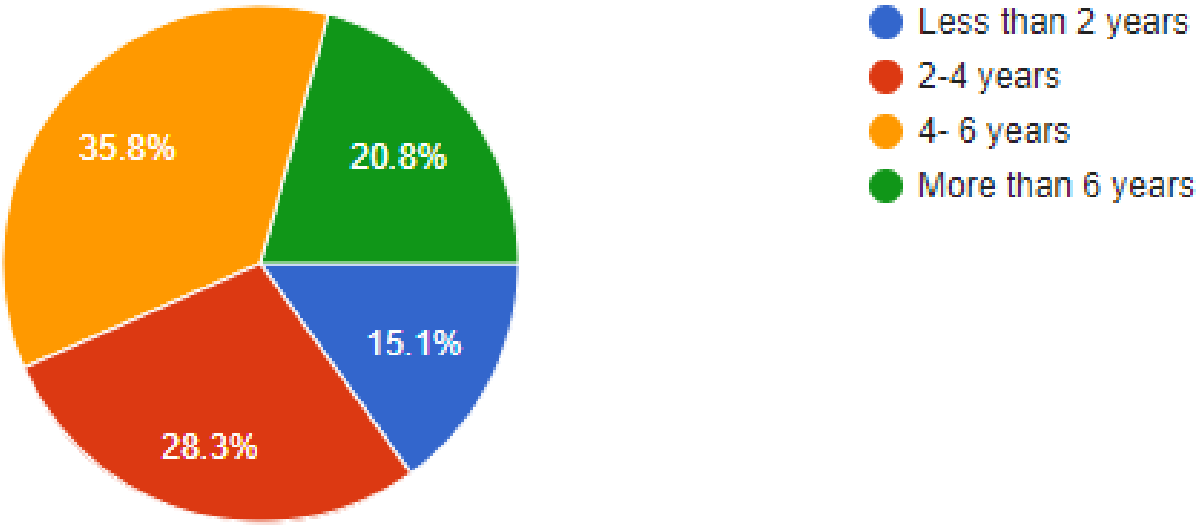


Figure 4.4 Respondent: Category Experience

4.1.5 Respondent: Category Job Title

Figure 4.5 represent the information about respondents from perspective of their job title. Below stats show that highest percentage of respondents have a job title of manager in their respective organizations.

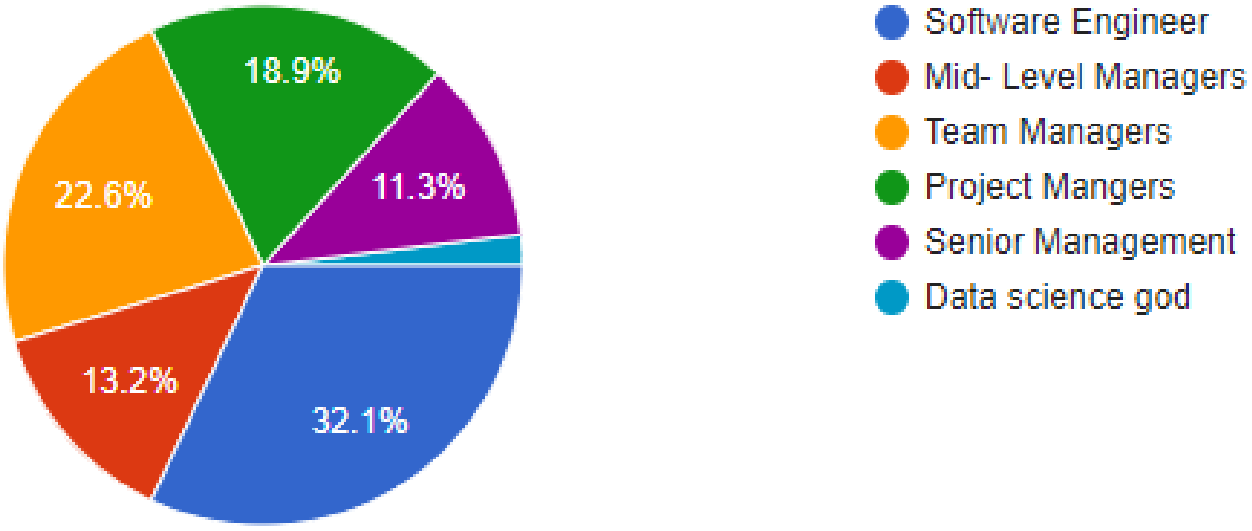


Figure 4.5 Respondent: Category Job Title

4.1.6 Respondent: Category OF BIG DATA experienced

Table 4.6 and figure 4.6 represent the information received from respondents against the existence of PMO in their respective organizations. 83% respondents say that they have a fully functioning PMO in their organizations where as 16% say that they have a kid of department performing PMO roles.

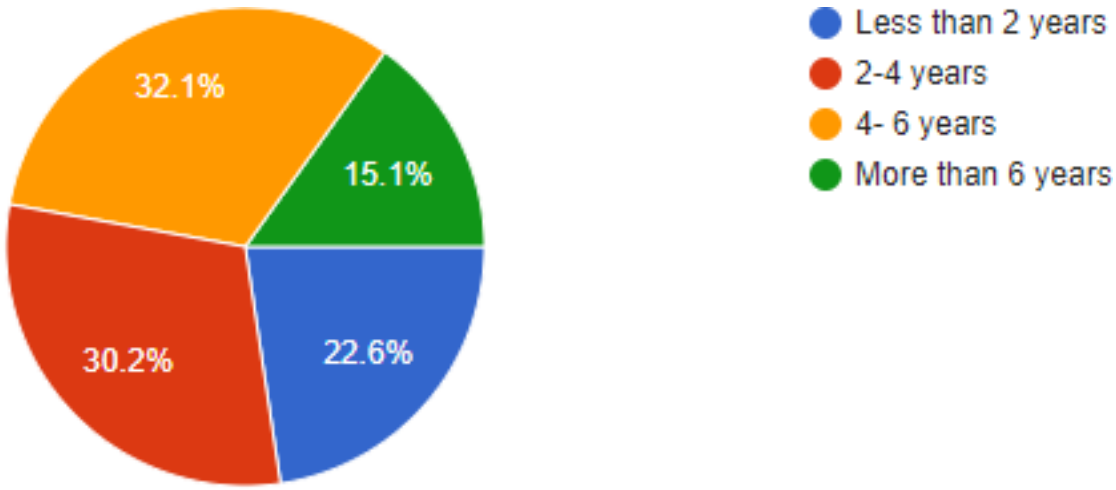


Figure 4.6 Respondent: Category Big data experienced

4.2 Reliability Analysis

Reliability of collected data was checked by applying Cronbach's Co-efficient Alpha. Reliability test was applied on all items against each variable. The closer the value of Cronbach's Alpha is to 1, the higher is the reliability. Reliability test for this study was performed on every variable independently. The threshold set for Cronbach's alpha is considered > 0.6.

4.2.1 Agile methodologies and Scrum Practices

After reliability analysis of independent variable project management standards and practices, the Cronbach's alpha value achieved is 0.748, so it can be said that data is reliable.

Table 4.1

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.812	.818	7

Reliability Statistics of ASMP

4.2.2 Tailored Scrum

After reliability analysis of independent variable tailored scrum, the Cronbach's alpha value achieved is 0.654, so it can be said that data is reliable.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.645	.650	2

Table 4.2 Reliability Statistics of TS

4.3 Factor Analysis

Factor analysis is performed to check the validity of data. As the original study was conducted in a different environment under different conditions, so the check of the same study can be validated in our environment an exploratory factor analysis (EFA) has been applied using SPSS. The find out the results of EFA, the sampling adequacy known as Kaiser-Meyer-Olkin and Barlett's Test was used as instrument.

4.3.1 KMO and Bartlett's Test

KMO & Bartlett's value ranges between 0 to 1 and to consider a data as valid, a value > 0.6 is acceptable. Data provided by respondents (N=60) was used for this test. Test was run with help of SPSS and adequacy value of 0.634 was achieved with a p-value < 0.05. Hence data was acceptable for further analysis.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.634
Bartlett's Test of Sphericity	Approx. Chi-Square	315.064
	Df	55
	Sig.	.000

Table 4.3 KMO & Bartlett's Test

4.4 Item Frequencies

Frequency analysis has been performed against each item in the survey. The items of the research tool were divided into 5 parts, one for each variable.

4.4.1 Agile Methodologies and scrum Practices

Below are the frequency stats for all items related to independent variable Agile Methodologies and scrum Practices

4.4.1.1 *Daily Scrum*

The results show that majority of respondents (73%) have indicated that project managers usually plays its role in regard to daily scrum.

AMSP1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	No	11	17.5	17.5	27.0
	Yes	46	73.0	73.0	100.0
	Total	63	100.0	100.0	

Table 4.4 AMSP1

4.4.1.2 *Duration of dailyscrum*

The results show that majority of respondents (49%) have indicated that project managers usually conducting scrum for less than 20 minutes

AMSP2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	less than 15 minutes	3	4.8	4.8	14.3
	Less than 20 Minutes	31	49.2	49.2	63.5
	Less than 30 Minutes	23	36.5	36.5	100.0
	Total	63	100.0	100.0	

Table 4.5 AMSP2

4.4.1.3 Scrum team

The results show that majority of respondents (41%) have indicated that Scrum team must is 5-9.

AMSP3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	10-14	26	41.3	41.3	50.8
	5-9	15	23.8	23.8	74.6
	less than 5	7	11.1	11.1	85.7
	More than 14	9	14.3	14.3	100.0
	Total	63	100.0	100.0	

Table 4.6 AMSP 3

4.4.1.4 Demo the deliverable at the end of each Sprint

The results show that majority of respondents (59%) have indicated that conducted at the end of each sprint.

AMSP4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	No	9	14.3	14.3	23.8
	Sometime yes	11	17.5	17.5	41.3
	Yes	37	58.7	58.7	100.0
	Total	63	100.0	100.0	

Table 4.7 AMSP 4

4.4.1.5 Recent Sprint

The results show that majority of respondents (48%) have indicated that their recent sprint was between 2-4 weeks.

AMSP5

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	1 Week	11	17.5	17.5	27.0
	2 weeks	16	25.4	25.4	52.4
	2-4 weeks	30	47.6	47.6	100.0
	Total	63	100.0	100.0	

Table 4.8 AMSP 5

4.4.1.6 *Retrospective meeting*

The results show that majority of respondents (41%) have indicated that project success always depends on retrospective meeting.

AMSP6

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	Less than 1 hour	26	41.3	41.3	50.8
	less than 20 minutes	13	20.6	20.6	71.4
	less than 30 Minutes	18	28.6	28.6	100.0
	Total	63	100.0	100.0	

Table 4.9 AMSP 6

4.4.1.7 *Total duration of the project*

The results show that majority of respondents (40%) have indicated that project success depends on total duration of the project

AMSP7

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	2 -3 Months	11	17.5	17.5	27.0
	3- 6 Months	25	39.7	39.7	66.7
	6- 12 Months	12	19.0	19.0	85.7
	Less than 2 months	5	7.9	7.9	93.7
	More than a year	4	6.3	6.3	100.0
	Total	63	100.0	100.0	

Table 4.10 AMSP 7

4.4.1.8 *Sprints in projects*

The results show that majority of respondents (44%) have indicated that each Big data project have 8-16 sprints

AMSP8

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	16-32 Sprints	12	19.0	19.0	28.6
	4-8 Sprints	10	15.9	15.9	44.4
	8-16 Sprints	28	44.4	44.4	88.9
	Less than 4 Sprints	3	4.8	4.8	93.7
	More than 32 Sprints	4	6.3	6.3	100.0
	Total	63	100.0	100.0	

4.4.1.9 *Product Backlog items did you achieve?*

The results show that majority of respondents (35%) have indicated that through agile methodologies they have met 60 to 80% deliverables.

AMSP9

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	20-40%	12	19.0	19.0	28.6
	40-60%	17	27.0	27.0	55.6
	60-80%	22	34.9	34.9	90.5
	80-100%	2	3.2	3.2	93.7
	Less than 20%	4	6.3	6.3	100.0
	Total	63	100.0	100.0	

Table 4.12 AMSP 9

4.4.1.10 *Effectiveness of sprint meetings*

The results show that majority of respondents (71%) have indicated that project success depends on sprint meetings

AMSP10

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	Maybe	10	15.9	15.9	25.4
	No	2	3.2	3.2	28.6
	Yes	45	71.4	71.4	100.0
	Total	63	100.0	100.0	

Table 4.13 AMSP 10

4.4.1.11 *Scrum help in managing development work*

The results show that majority of respondents (43%) have indicated that Scrum help in managing development work which leads to project success.

AMSP11

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		6	9.5	9.5	9.5
	Low level (Less than 25%)	5	7.9	7.9	17.5
	Mid level (Less than 50%)	18	28.6	28.6	46.0
	Upper level (75-100%)	7	11.1	11.1	57.1
	Upper- Mid level (Less than 75	27	42.9	42.9	100.0
	Total	63	100.0	100.0	

Table 4.14 AMSP 11

4.4.2 Tailored Scrum

Below are the frequencies for all items related to this variable.

4.4.2.1 Included in Sprint meeting

The results show that majority of respondents (66%) have indicated that including sprint meetings leads them in project success.

TS1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		14	19.7	19.7	19.7
	Maybe	3	4.2	4.2	23.9
	No	7	9.9	9.9	33.8
	Yes	47	66.2	66.2	100.0
	Total	71	100.0	100.0	

Table 3.15 TS 1

4.4.2.2 Accept change requests during the sprints

The results show that majority of respondents (52%) have indicated they accept change requests during sprints which still helps them in achieving goals.

TS2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		14	19.7	19.7	19.7
	Maybe	10	14.1	14.1	33.8
	No	10	14.1	14.1	47.9
	Yes	37	52.1	52.1	100.0
	Total	71	100.0	100.0	

Table 3.16 TS2

4.4.2.3 Cater change requests during the running sprints?

The results show that majority of respondents (53%) have indicated that they cater change

Requests during the running sprints.

TS3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		14	19.7	19.7	19.7
	Maybe	9	12.7	12.7	32.4
	No	10	14.1	14.1	46.5
	Yes	38	53.5	53.5	100.0
	Total	71	100.0	100.0	

Table 3.17 TS3

4.4.2.4 Change requests are being accepted

The results show that majority of respondents (38%) have indicated that change requests are being accepted.

TS4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		14	19.7	19.7	19.7
	(Less than 25%)	3	4.2	4.2	23.9
	(Less than 50%)	21	29.6	29.6	53.5
	(Less than 75%)	27	38.0	38.0	91.5
	between 75 -100%	6	8.5	8.5	100.0
	Total	71	100.0	100.0	

Table 4.18 TS4

4.5 Correlation Analysis

The possible values for correlation range from -1 to +1 where below 0 values indicate a negative relationship or inversely relationship between variables and above 0 values indicate a positive or direct relationship between variables. A correlation analysis was performed on all variables and below table shows their relationship with each other.

Correlations

		AVGAMSP	AVGTS
AVGAMSP	Pearson Correlation	1	.400**
	Sig. (2-tailed)		.002
	N	60	60
AVGTS	Pearson Correlation	.400**	1
	Sig. (2-tailed)	.002	
	N	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.159 Correlation Analysis

4.5.1 Correlation between agile methodologies & Scrum practices and project success:

As statistical results show that correlation value between IV and DV is 0.792, which is statistically significant at 0.01. Hence, it can be concluded that there is a positive relationship between Project success (DV) and agile methodologies and scrum practices (IV).

4.5.2 Correlation between Tailored Scrum and project success:

As statistical results show that correlation value between ASMP (IV) and PP (DV) is 0.295, which is statistically significant at 0.01. Hence, it can be concluded that there is a positive relationship between Project Performance (DV) and Project historical archives management (IV).

4.6 Regression Analysis

Regression is a statistical analysis to investigate the possible relationships between variables. By applying regression, the investigator tries to learn about the effect of one variable upon the other variable. For exploration of this this relationship, data on all variables is assembled and regression analysis is applied to estimate the quantitative effect of a variable upon the target variable. Similarly, a simple linear regression was applied with help of SPSS and results are shown below:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.604[a]	.365	.342	.54965	.365	16.351	2	57	.000	1.822

a. Predictors: (Constant), AVGAMSP, AVGTS

b. Dependent Variable: AVGDV

Table 4.16 Model Summary

The above table for model summary expresses that R=0.604, R^2=0.365, adjusted R^2=0.342 and significance at .000 level. Putting it in more understandable terms, we can say that there is a 34% variation in Project Success (DV) due to all independent variables.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.880	2	4.940	16.351	.000[b]
	Residual	17.220	57	.302		
	Total	27.100	59			

a. Dependent Variable: AVGDV

b. Predictors: (Constant), AVGAMSP, AVGTS

Table 4.21 ANOVA

The results generated in Table 4.39 by ANOVA test shows that F-test is significant at .000 level. In other words, it means that all IV's (ASMP and TS) collectively have a significant relationship with Project Performance (DV).

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.197	.276		.711	.480		
	AVGTS	.748	.217	.396	3.441	.001	.840	1.190
	AVGAMSP	.398	.142	.324	2.811	.007	.840	1.190

a. Dependent Variable: AVGDV

Table 4.22 Coefficients

The value of β-coefficients generated in table 4.40 shows that if there is an average 1-unit change in IV agile methodologies and scrum practices, there will be 0.324-unit change in DV project performance at .007 significance level.

If there is an average 1-unit change in IV tailored scrum, it will cause 0.396-unit change in DV project performance at .001 significance level.

Below is the correlation matrix against each question.

Correlation Matrix

		AMSP3	AMSP4	AMSP5	AMSP7	AMSP8	AMSP9	AMSP10	AMSP11	TS2	TS3
AMSP3	Correlation	1.000	.474	.035	.708	.565	.578	.237	.557	-.020	.230
AMSP4	Correlation	.474	1.000	.254	.402	.395	-.018	.272	.258	-.380	-.114
AMSP5	Correlation	.035	.254	1.000	.245	.506	.006	-.243	.363	-.580	-.095
AMSP7	Correlation	.708	.402	.245	1.000	.632	.444	.016	.564	-.186	.354
AMSP8	Correlation	.565	.395	.506	.632	1.000	.300	-.060	.538	-.222	.347
AMSP9	Correlation	.578	-.018	.006	.444	.300	1.000	-.133	.402	.127	.521
AMSP10	Correlation	.237	.272	-.243	.016	-.060	-.133	1.000	-.218	-.110	-.232
AMSP11	Correlation	.557	.258	.363	.564	.538	.402	-.218	1.000	-.183	.170
TS2	Correlation	-.020	-.380	-.580	-.186	-.222	.127	-.110	-.183	1.000	.482
TS3	Correlation	.230	-.114	-.095	.354	.347	.521	-.232	.170	.482	1.000

Table 3.23 Correlation Matrix

CHAPTER 5: CONCLUSION

All of our research is concluded with the final result and limitations of the study and recommendations accordingly. Our knowledge is not in the current literature duties to adopt measures in the context of the scrum in Pakistan. This research is working on big data for some of the difference was held to fill, Pakistan Software Houses and ICT industry. This research study is designed to find these factors to develop a survey that influence the effectiveness of a project using the amount of ways. Results for sufficient data were collected from six companies of different sizes and analysis of experimental data was provided enough information to get. A sole purpose of this study was to analyze the effects of agile methodologies/ways to the success of Big Data projects in Pakistan industry. There are three main objectives of this study include the following listed below:

- Agile is the best method for big data Project Management.
- To implement the fire process for the project's success.

5.1 Research Findings

In accordance with determinations of this study, some significant aspects of agile methodologies (Scrum) were thoroughly analysed and deliberated. The research instrument was circulated randomly to a targeted population resources occupied on projects and project teams to get their insight regarding the impact of agile methodologies (Scrum) on their delivery of successful projects

Majority respondents were on a manager's position with an experience of 6-10 years. Almost 82% respondents agreed to have an agile methodologies scrum helps them to deliver successful projects whereas 17% indicated to have another entity of agile methodology XP.

Results of linear regression (Table 4.40) prove that correlations of all independent variables (IVs) with dependent variable (DV) were positive.

The standards and practices variable (ASMP) showed the highest variation (R^2) against DV project performance. Sachan (2017) reported a similar pattern regarding impact of impact of agile

methodologies in field of data science. From these results, we can conclude that organizations that pay more devotion towards implementation of agile methodology with Scrum standards and practices have higher odds to deliver successful projects.

The next variable to show high variation is Tailored Scrum (TS) with an R^2 value of 0.654. So, we can conclude that according to targeted population emphasises that tailored scrum also helps them to deliver the successful project according to their culture. As we can see in results as well, that through tailoring the scrum helps them achieving their goal.

Out of the targeted population, almost 88% respondents have agreed that projects have met their objectives more efficiently after implementation of agile methodologies and almost 75% for the targeted population have agreed that projects were more successful over a period of time after implementing agile methodology- SCRUM.

5.2 Practical Implications

The findings of this study are of very importance and have important practical implications in a way that they can used to improve the existing software houses and IT companies of Pakistan. They also provided guidelines of agile methodology scrum to the existing project managers to be more agile in their role.

This study might help the Software houses and IT companies in Pakistan to understand the importance of agile methodologies and its value in terms of delivery of successful projects.

5.3 Recommendations

Based on the findings of this study following endorsements can be made for existing agile project managers and the ones to be formed yet:

- This study has produced a strong evidence that agile methodologies and Scrum practices have a high correlation with delivery of successful projects.
- There is a need to pay some attention towards tailoring scrum as it helps to adopt the best practices of agile and scrum according to their culture which helps in delivering successful projects of Big Data in Pakistan.

- When implementing agile or scrum at once it is not possible to have fruitful results but timely it will be more helps in industry like data science and Big Data in Pakistan. That is the reason to start with tailoring scrum according to organization culture.

5.4 Research Limitations

There are a few limitations or barriers that were face during this study, which are discussed below:

- Due to limited resources, time and existing research work done in the field of big data and IT companies in Pakistan, it required more time to find related material and integrate the data to perform this study.
- It also acquires lot of effort to persuade companies and respondents to fill complete the research tool because of the general idea that might have to provide any confidential information or any kind of sensitive data.
- Due to shortage of time, sample size was kept small. A larger sample size might provide some more insight into current implementation of software houses and IT companies of Pakistan and their efforts to deliver successful projects.
- The other extreme is that it was for, by performing pilot testing and detailed research/studies failed to do early stage testing later, instead of completely projects questions referring to the completed sprints was reradiating, because there are many people started with scrum but couldn't make it successful. And it took time for the survey as they did not complete a project using scrum.
- Lastly, sample size is a bit small but it was accepted that the population is not large. If more, then the results may be more representative of the industry in Pakistan.

REFERENCES

Armour, F. J., Espinosa, A. J., & Kaisler, S. H. (2017). Introduction to Big Data and Analytics:. *Introduction to Big Data and Analytics: Concepts, Techniques, Methods, and Applications Minitrack*, 3.

Braganza, A., Moro, R., Ali, M., Nepelski, D., & Brooks, L. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Resource management in big data initiatives: Processes and dynamic capabilities*, 10.

Cleland, D. (2004). *Field Guid to Project Management 2nd Edition.*

Couch, N., & Robins, B. (2013). BIG DATA FOR DEFENCE and Security. *big-data-for-defence-and-security-report-final-130924210732-phpapp01*, 06.

Dai , C. X., & Wells, W. (2004). An exploration of project management office features and their relationship to project performance. *International Journal of Project Management 22 (2004) 523–532*, 9.

Desouza, K. C., & Evaristo, J. R. (2006). Project management offices: A case of. *International Journal of Information Management*, 414–423.

Elragal, A., & Elgendy, N. (2014). Big Data Analytics: A Literature Review Paper. *Research Gate*, (p. 15).

Franková, M., Drahošová, M., & Balco, P. (2016). Agile project management approach and its use in big data. *Agile project management approach and its use in big data*, 8.

Gunasekaran, A., Yusuf, Y. Y., Adeleye, E. O., & Papadopoulos, T. (2017). Agile manufacturing practices: the role of big data and business analytics with multiple case studies. *Agile*

manufacturing practices: the role of big data and business analytics with multiple case studies, 14.

Han, S., Meng, W., & Cheng Li. (2017). An Agile Confidential Transmission Strategy Combining Big Data Driven Cluster and OBF. *An Agile Confidential Transmission Strategy Combining Big Data Driven Cluster and OBF*, 12.

Industries, W. B. (2017). *http://pas.org.pk/6375-2/*. Retrieved from Pakistan Advisor Society: http://pas.org.pk/

IPMA. (2006). *IPMA Competence Baseline (ICB) V 3.0.*

Ishwarappa, A. J. (2015). A brief introduction on Big data 5Vs characteristics and hadoop technology. *Elsevier*, 6.

Kerzner, H. (2009). *Project Management: A Systems Approch to Planning and Control 10th Edition.*

Lewis. (1995).

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design.* Routledge.

Oisen, R. (1971). Can project management be defined? *Project Management Quarterly, 1971, 2(1), 12±14.*

PMI. (n.d.). *Project Management Body of Knowlege (PMBOK).*

Quora. (2017). *List of companies in Pakistan*. Retrieved from Quora: https://www.quora.com/Which-companies-in-Pakistan-work-on-data-analytics

Rauf, W. (2015, 25 7). Role of Computing Organizational Justice in Reducing the Conflict between the Employees and Its Impact on the Job Performance. *Journal of Information Technology and Management Sciences, 54*(60), 54-60.

Sachan, S. (2017). Best practices for adopting agility in Business Intelli-gence. *Best practices for adopting agility in Business Intelli-gence*, 47.

Saltz, J. S., Shamshurin, I., & Crowston, K. (2017). Comparing Data Science Project Management Methodologies via a. *Comparing Data Science Project Management Methodologies via a*, 10.

Sanchez, O. P., & Terlizzi, M. A. (2017). Cost and time project management success factors for information systems development projects. *International Journal of Project Management, 35*, 1608-1626.

Selig, G. J. (2015). *Implementing Effective IT Governance and IT Management.* Van Haren.

Stanberry, L. (2018). *Critical Success Factors for Large and Distributed Agile Software Development Projects Using Scrum in U.S.-Based Global Companies.* USA: ProQuest.

Turner, J. (1996). Editorial: International Project Management Association global qualifcation, certifcation and accreditation. *International Journal of Project Management, 1996, 14(1), 1±6.*

Unhelkar, B. (2017). Big Data Framework for Agile Business (BDFAB) As a Basis for Developing Holistic Strategies in Big Data Adoption. *Big Data Framework for Agile Business (BDFAB) As a Basis for Developing Holistic Strategies in Big Data Adoption*, 11.

APPENDIX A – QUESTIONNAIRE

The following is the questionnaire distributed among Big Data professionals.

5.5 Impacts of agile methodology (Scrum) in success of Big Data projects.

This questionnaire aims to study the critical success factors of projects which use Scrum in success of big data projects

The data from this survey will only be used for academic research, for the Thesis requirement of the MS in Project Management, at the Bahria University Lahore Campus. Your responses will be kept strictly confidential. It is guaranteed that the information provided by you will not be disclosed to any third party.

If you have any suggestions, or would like clarification about any of the questions or how the data will be used, please do not hesitate to contact me or my Thesis advisor (contact details are below).

Thank you very much in advance for your time and cooperation.

Agha Hassan Afzal Khan
MS PM - 2018.
Agha.hassan.khan@live.com

Advisor:
Dr. Shahid Iqbal, PHD
diq@bahria.edu.pk

5.6 Section 1: Company Information

1. Is your organization involved in Big Data Projects?
2. How many people are employed at organization?
3. How long has Scrum been used in your organization
4. Are you based in Pakistan?

5.7 Section 2: Demographic Information

1. What is your gender?
2. What is your age?
3. What is your total work experience in Big Data?
4. What is your job designation?
5. What is your highest qualification?
6. How many years of work experience do you have in software development?
7. What is your total work experience in Big Data?
8. For how long you have been practiced agile methodology (Scrum)?

5.8 Section 3A: Information on adherence of Scrum principles in a successful project

Glossary:
Please answer the following questions with regard to a **successful** project completed recently, which you were part of.

Glossary:
Successful project - Completed within allocated Time frame, with allocated Budget while meeting Quality objectives.

If you have not completed any project yet, think of the most recent Sprint which was successful.
Glossary:

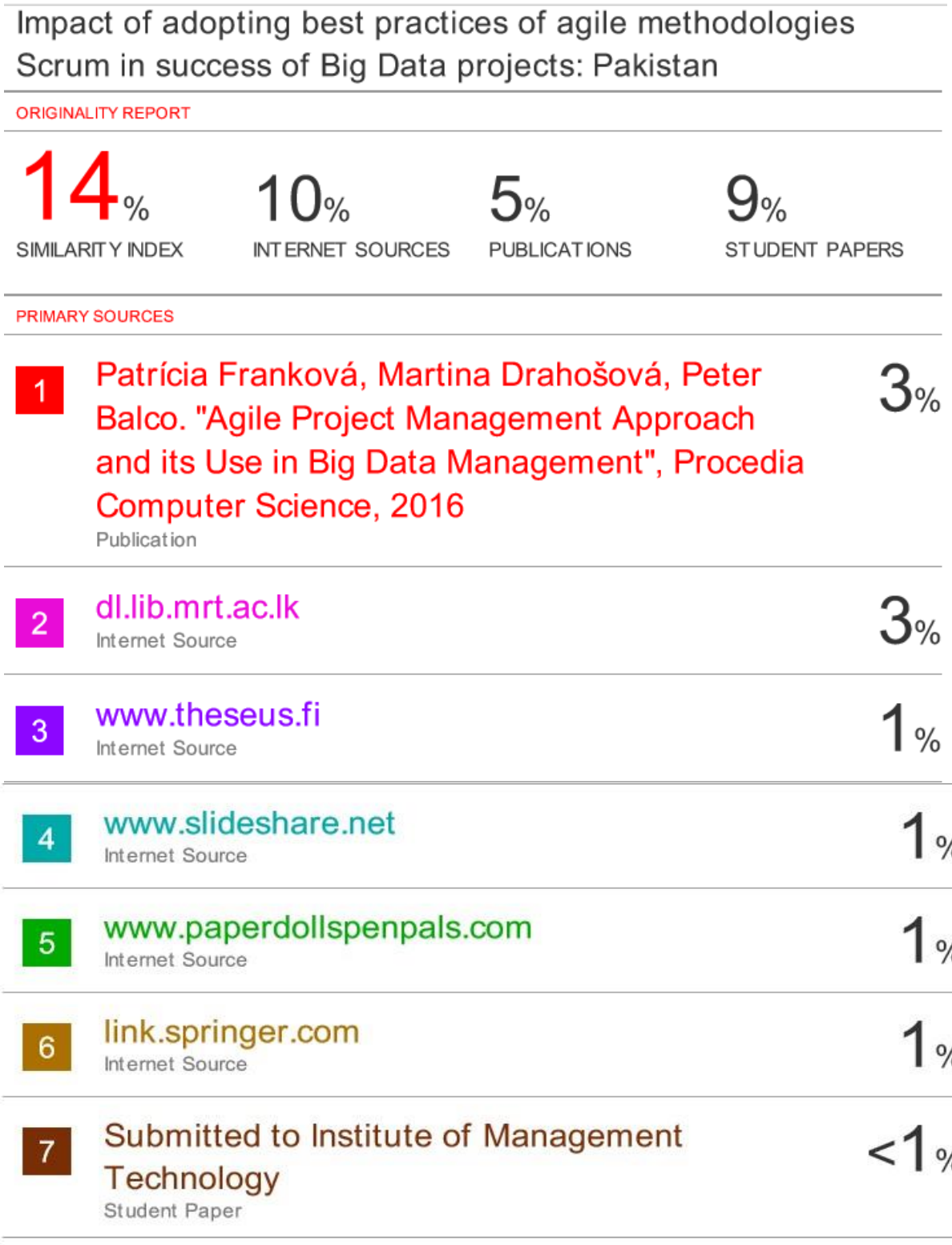
Successful Sprint - Completed all the tasks in the Sprint and delivered the shippable product increment.

1. Did you have a daily Scrum?
2. How long did the daily Scrum take?
3. How many members were present in the Scrum team? (Including the Product Owner(s), Project Manager and BDM)
4. Did you demo the deliverable at the end of each Sprint?
5. How long was the most recent Sprint?
6. How long was the Retrospective meeting?
7. What was the total duration of the project?
8. How many Sprints did the project have?
9. What percentage of Product Backlog items did you achieve?
10. Did you think that daily Scrum meetings were effective in making your project successful?
11. In your opinion, to what level did Scrum help in managing development work?
12. Are you included in Sprint meeting?
13. Do you accept change requests during the sprints?
14. Do you cater change requests during the running sprints?
15. How often change requests are being accepted?
16. To what level do you agree with the following statements?

1-Strongly disagree 2- Disagree 3- Somewhat disagree 4 – Undecided 5- Somewhat agree 6 – Agree 7 – Strongly agree

		1	2	3	4	5	6	7
1	Management was extremely supportive of Scrum:							
2	The team attitude towards Scrum was very hostile:							
3	The Product Owner's involvement in the project was complete and adequate:							
4	The management constantly attempted to change the priorities while in a Sprint:							
5	The team was never disturbed while in a Sprint:							
6	The Project Manager influenced the team members in making decisions:							
7	The tools used to manage the work (e.g. ScrumWorks) were excellent:							
8	The level of communication among team members was very poor:							
9	The quality objectives were significantly exceeded:							
10	The team was unable to self-manage itself to any extent:							
11	The actual budget and the estimated budget was the same:							
12	The actual time frame significantly overran the estimated time frame:							
13	Early and continuous delivery of valuable software were offered:							
14	Changing requirements, even late in development were accepted:							
15	The team self-policed and reinforced use of process and rules:							
16	The organization was able to comply with Scrum rules:							
17	The Project Manager was effective in getting process followed:							
18	The team was self-managing:							
19	Surprises didn't occur:							
20	The team was cross-functional:							
21	The team and Product Owner collaborated and worked closely together:							
22	Customer’s feedback on demos were given priority:							
23	Team members were dedicated and honored commitments:							

24	The team effectively acted upon indicators in Sprint Burndown:							
25	The team effectively managed conflict within team:							
26	The team improved internal development processes:							
27	The team worked to improve itself and its processes:							
28	The team adequately managed dependencies:							



Impact of adopting best practices of agile methodologies Scrum in success of Big Data projects: Pakistan

ORIGINALITY REPORT

14%	10%	5%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Patrícia Franková, Martina Drahošová, Peter Balco. "Agile Project Management Approach and its Use in Big Data Management", Procedia Computer Science, 2016 Publication	3%
2	dl.lib.mrt.ac.lk Internet Source	3%
3	www.theseus.fi Internet Source	1%
4	www.slideshare.net Internet Source	1%
5	www.paperdollspenpals.com Internet Source	1%
6	link.springer.com Internet Source	1%
7	Submitted to Institute of Management Technology Student Paper	<1%

9	Submitted to Napier University Student Paper	<1%
10	Submitted to University of Hertfordshire Student Paper	<1%
11	Submitted to University of West London Student Paper	<1%
12	ir.dut.ac.za Internet Source	<1%
13	sites.tufts.edu Internet Source	<1%
14	www.mportal.com Internet Source	<1%
15	www.icobm.com.pk Internet Source	<1%
16	Jeffrey S. Saltz, Ivan Shamshurin. "Exploring the process of doing data science via an ethnographic study of a media advertising company", 2015 IEEE International Conference on Big Data (Big Data), 2015 Publication	<1%
17	Submitted to Universiti Teknologi MARA Student Paper	<1%

18	Submitted to Rivers State University of Science & Technology Student Paper	<1%
19	article.sapub.org Internet Source	<1%
20	eprints.utcc.ac.th Internet Source	<1%
21	documents.mx Internet Source	<1%
22	Mohd Bakhari, Zawiyah M.. "Chapter 15 The Contribution of Technological Factors on Knowledge Sharing Quality among Government Officers in Malaysia", InTech, 2010 Publication	<1%

Exclude quotes	Off	Exclude matches	< 10 words
Exclude bibliography	Off		