

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ADMINISTRAÇÃO DE EMPRESAS DE SÃO PAULO

Federico Venni

Practical applications of Scrum: an empirical analysis of
Product Backlog management practices

SÃO PAULO
2024

Federico Venni

Practical applications of Scrum: an empirical analysis of
Product Backlog management practices

Thesis presented to Escola de
Administração de Empresas de São Paulo
of Fundação Getúlio Vargas, as a
requirement to obtain the title of Master
in International Management (MPGI).

Knowledge field: International Management

Thesis Advisor: Prof. Dr. André Cherubini

SÃO PAULO

2024

Dados Internacionais de Catalogação na Publicação (CIP)
Ficha catalográfica elaborada pelo Sistema de Bibliotecas/FGV

Venni, Federico.

Practical applications of Scrum : an empirical analysis of Product Backlog management practices / Federico Venni. - 2024.

f. 93

Orientador: André Cherubini Alves.

Dissertação (mestrado profissional MPGI) – Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo.

1. Administração de produtos. 2. Scrum (Desenvolvimento de software). 3. Desenvolvimento ágil de software. 4. Satisfação do consumidor. I. Alves, André Cherubini. II. Dissertação (mestrado profissional MPGI) – Escola de Administração de Empresas de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 658.62

Federico Venni

Practical applications of Scrum: an empirical analysis of
Product Backlog management practices

Thesis presented to Escola de Administração
de Empresas de São Paulo of Fundação
Getulio Vargas, as a requirement to obtain
the title of Master in International
Management (MPGI)

Knowledge Field: International Management

Approval Date: 6/12/2024

Committee members:

Prof. Dr. André Cherubini
FGV-EAESP

Prof. Dr. Fábio Botelho Josgrilberg
FGV-EAESP

Prof. Dr. Fernando Ferreira
University of Southern Indiana

ABSTRACT

In the ever-evolving and competitive business landscape, organizations are increasingly adopting Agile methodologies to enhance flexibility, foster collaboration, and deliver value efficiently. Among these methodologies, the Scrum framework has emerged as a prominent approach, providing structured yet adaptable tools to address the complexities of project management. Central to the Scrum framework is the Product Backlog, a dynamic and prioritized list of work items that guides the development process. Despite its critical role, practical insights into the daily management and optimization of the Product Backlog remain limited in the academic discourse. This thesis aims to bridge this gap by exploring the methodologies and practices employed by professionals in managing the Product Backlog effectively. A thorough literature review lays the groundwork by examining existing theoretical models and identifying key practices outlined in Agile and Scrum principles. Building on this foundation, the study employs an empirical approach through a structured questionnaire disseminated among Scrum practitioners across various industries. The collected data underwent rigorous statistical analysis, including descriptive statistics, factor analysis, multiple regression, and k-means cluster analysis. These methods provided a multifaceted view of the prevailing practices, highlighting how factors such as team collaboration, experience level, and the utilization of specific estimation and prioritization techniques influence the overall satisfaction and effectiveness of Product Backlog management. The findings underscore the importance of collaborative story creation, regular and precise estimation of backlog items, and a continuous prioritization process to enhance the efficiency and alignment of development efforts with stakeholder expectations. Furthermore, the cluster analysis revealed distinct approaches to backlog management based on the practitioners' experience levels, offering nuanced insights into how seasoned and novice Scrum users differ in their application of these practices. This research contributes to the existing body of knowledge by not only delineating the practical aspects of Product Backlog management but also offering actionable recommendations for organizations aiming

to refine their Scrum processes. By understanding these practices, companies can better tailor their Agile implementations to improve project outcomes, foster team satisfaction, and ultimately drive greater value for their stakeholders.

KEY WORDS: Product Backlog, Collaboration, SCRUM, Satisfaction

RESUMO

No cenário empresarial em constante evolução e altamente competitivo, as organizações estão adotando cada vez mais metodologias ágeis para aumentar a flexibilidade, fomentar a colaboração e entregar valor de forma eficiente. Entre essas metodologias, o framework Scrum se destaca como uma abordagem proeminente, oferecendo ferramentas estruturadas, mas adaptáveis, para lidar com as complexidades da gestão de projetos. No centro do framework Scrum está o Product Backlog, uma lista dinâmica e priorizada de itens de trabalho que orienta o processo de desenvolvimento. Apesar de seu papel crítico, as percepções práticas sobre a gestão e otimização diária do Product Backlog permanecem limitadas no discurso acadêmico. Esta tese busca preencher essa lacuna ao explorar as metodologias e práticas empregadas por profissionais na gestão eficaz do Product Backlog. Uma revisão bibliográfica abrangente estabelece a base ao examinar modelos teóricos existentes e identificar práticas-chave descritas nos princípios Ágeis e do Scrum. Com base nesse fundamento, o estudo adota uma abordagem empírica por meio de um questionário estruturado distribuído entre praticantes de Scrum de diversos setores. Os dados coletados foram submetidos a uma análise estatística rigorosa, incluindo estatísticas descritivas, análise fatorial, regressão múltipla e análise de cluster pelo método k-means. Esses métodos forneceram uma visão multifacetada das práticas predominantes, destacando como fatores como colaboração em equipe, nível de experiência e a utilização de técnicas específicas de estimativa e priorização influenciam a satisfação e a eficácia geral na gestão do Product Backlog. Os resultados enfatizam a importância da criação colaborativa de histórias, da estimativa regular e precisa dos itens do backlog e de um processo contínuo de priorização para aumentar a eficiência e o alinhamento dos esforços de desenvolvimento com as expectativas das partes interessadas. Além disso, a análise de clusters revelou abordagens distintas para a gestão do backlog com base nos níveis de experiência dos praticantes, oferecendo insights detalhados sobre como usuários experientes e iniciantes do Scrum diferem na aplicação dessas práticas. Esta pesquisa contribui para o corpo de

conhecimento existente ao não apenas delinear os aspectos práticos da gestão do Product Backlog, mas também ao oferecer recomendações acionáveis para organizações que buscam aprimorar seus processos Scrum. Compreendendo essas práticas, as empresas podem adaptar melhor suas implementações Ágeis para melhorar os resultados dos projetos, promover a satisfação das equipes e, em última instância, gerar maior valor para seus stakeholders.

PALAVRAS CHAVE: Product Backlog, colaboração, SCRUM, satisfação.

Table of Contents

| | |
|---|-----------|
| 1 - Introduction | 11 |
| 2 - Literature Review | 13 |
| 2.1 Agile | 13 |
| 2.2 Scrum Framework | 16 |
| 2.2.1 Product Backlog composition | 19 |
| 2.2.2 Backlog Refinement | 22 |
| 2.2.3 Items Estimation | 23 |
| 2.2.4 Product Backlog Prioritization | 24 |
| 2.3 Research questions | 26 |
| 3 - Methodology | 28 |
| 3.1 Questionnaire | 29 |
| 3.1.1 Section One: Respondent Background and Experience | 29 |
| 3.1.2 Section Two: Product Backlog Composition | 30 |
| 3.1.3 Section Three: Stories | 30 |
| 3.1.4 Section Four: Estimates | 31 |
| 3.1.5 Section Five: Prioritization | 31 |
| 3.1.6 Section Six: General Product Backlog Management Practices | 32 |
| 3.1.7 Questionnaire's dissemination | 32 |
| 3.2 Statistical Analysis | 33 |
| 4 - Results | 37 |
| 4.1 Descriptive Analysis | 37 |
| 4.1.1 Respondents' information | 37 |
| 4.1.2 Product Backlog Management descriptive analysis | 44 |
| 4.2 Factor Analysis | 53 |
| 4.3 Multiple Regression | 58 |
| 4.4 Cluster Analysis | 68 |
| 5 - Managerial implications and conclusions | 76 |
| 5.1 Limitations of the study and directions for future research | 80 |
| 6 – Appendix | 82 |
| 7 - References | 89 |

List of Figures and Tables

| | |
|--|----|
| Table 1; Number of questions for each section of the questionnaire | 29 |
| Table 2; Likert Scales | 44 |
| Table 3; SPSS; Analysis of questions about Stories | 46 |
| Table 4; SPSS; Analysis of questions about Estimates | 48 |
| Table 5; SPSS; Analysis of questions about Prioritization | 50 |
| Table 6; Analysis about General Product Backlog Management practices | 52 |
| Table 7; SPSS; KMO and Bartlett's Test of Sphericity | 54 |
| Table 8; SPSS; Extracted Factors | 56 |
| Table 9; SPSS; Summary of the Regression Model | 60 |
| Table 10; SPSS; ANOVA | 61 |
| Table 11; SPSS; Multiple Regression Coefficients and Collinearity Statistics for the Independent Variables | 63 |
| Table 12; SPSS; Cluster Analysis ANOVA Table | 75 |
| | |
| Figure 1; SPSS; Industry Representation in Survey Respondents | 38 |
| Figure 2; SPSS; Firm's dimension | 38 |
| Figure 3; SPSS; Scrum Team dimension | 39 |
| Figure 4; SPSS; Respondents' experience with Scrum | 40 |
| Figure 5; SPSS; Respondents' product/project purpose | 41 |
| Figure 6; SPSS; "How are Product Backlog items classified/represented?" percentage responses | 42 |
| Figure 7; SPSS; "Which fields do you use in your PB?" percentage responses | 43 |
| Figure 8; SPSS; User Stories' syntax usage | 45 |
| Figure 9; SPSS; Number of items estimated | 46 |
| Figure 10; SPSS; Number of items prioritized | 49 |
| Figure 11; SPSS; Scree Plot | 55 |
| Figure 12; SPSS; Normal P-P Plot | 61 |
| Figure 13; SPSS; Scatterplot | 62 |

1 - Introduction

In the contemporary business environment, companies are operating within a highly competitive landscape that is characterized by rapid change and evolving stakeholder expectations. It is precisely to address these challenges that the necessity has arisen to implement project management tools that are flexible and oriented towards the continuous satisfaction of business stakeholders. First, the Agile methodology and subsequently the Scrum framework emerged in the 1990s in response to the aforementioned challenges, offering a viable alternative to traditional project management models. At the outset, these alternative methodologies were primarily adopted within the IT and software development sectors. In recent years, however, the effectiveness and adaptability of these methodologies in different contexts has led to their implementation in a wide variety of sectors. In the context of the Scrum framework, the Product Backlog, which serves to aggregate all the requisite elements for implementation, is of particular significance. Despite the growing prevalence of Scrum in a diverse array of sectors and contexts, there remains a scarcity of literature that provides comprehensive insights into the daily management of product backlogs in practice. The primary source is the Scrum Guide, authored by Schwaber and Sutherland, which delineates the fundamental values and principles that must be adhered to in order to effectively implement the framework they have developed. In their introduction to the guide, the authors set forth their overarching intention in writing it as follows: The Scrum Framework is intentionally incomplete, defining only the parts necessary to implement Scrum theory. Rather than providing detailed instructions, the rules of Scrum guide the relationships and interactions of those who use it. As a result, the Scrum Guide and most of the articles that can be found in academic and scientific literature are not intended to define in practical terms the practices to be implemented in order to use Scrum effectively. Rather, it only lays down basic pillars that can be adapted to a variety of contexts. For this reason, a literature review highlights a lack of sources that actually describe how the Scrum Framework,

and in particular the Product Backlog, is managed in practical terms by companies. Therefore, the aim of this research is to investigate, by means of a specific and targeted questionnaire, which practices are most used by companies in applying the Scrum framework and, in particular, the tool known as the "Product Backlog". The research project consists of an initial literature review concluded with the three research questions that this project aims to address, a chapter dedicated to the implemented methodology and a final section dedicated to the analysis of the results and their managerial implications. The three research questions are the following:

- I. Are there specific approaches to implementing key Product Backlog management practices that are more commonly utilized than others?
- II. What are the key factors that most significantly influence satisfaction levels in Product Backlog management? Additionally, which Product Backlog management approaches should companies adopt to maximize satisfaction?
- III. Does the experience of professionals with Scrum influence their approach to managing the Product Backlog?

2 - Literature Review

2.1 Agile

"Agile" is an approach to product development characterized by its propensity to prioritize customer satisfaction, flexibility and teamwork (Beck et al., 2001) and is a reaction to traditional project management methods which have often been criticized for their rigidity and lack of change (Highsmith, 2002). Agile encourages the so-called "iterative development" which consists in dividing projects into smaller portions or "iterations" (Schwaber & Sutherland, 2020). These iterations, by giving teams the opportunity to receive feedback and change the course of the project frequently, aim to produce a potentially shippable product increment (Schwaber & Beedle, 2002). As already mentioned, the fundamental components of Agile Project Management are based on managing the effects of complexity and uncertainty in a project, on the realization that there should be a considerable reduction in the time between planning and execution, and on the fact that planning an action does not necessarily provide all the details that emerge during the implementation phase (Cockburn, 2006).

The origins of Agile date back to the early 1990s, a period of rapid technological development and considerable growth in demand for product delivery software (Highsmith, 2002). In those years, software development was proving to be too dynamic and fast for conventional project management methodologies such as the "Waterfall" model and, given the linear and sequential nature of these traditional approaches, it was very difficult to make changes once the project was at an advanced stage (Beck, 1999). The creation of the Agile Manifesto in 2001 marked a considerable turning point in the history of Agile. Indeed, seventeen software developers gathered in Utah and created a manifesto containing twelve key principles and four fundamental values on which Agile practices should be based (Beck et al., 2001). The four core values are the following:

1. Individuals and interactions over processes and tools

2. Working software over comprehensive documentation
3. Customer collaboration over contract negotiations
4. Responding to change over following a rigid plan

These values emphasize the importance of human elements in project management, the need for functional products, the value of customer input, and the flexibility to adapt to changes.

The 12 fundamental principles described in the Agile Manifesto provide a more accurate and detailed framework for Agile and they include:

1. "Satisfy Customers Through Early and Continuous Delivery". The main objective is to consistently provide customers with valuable software, ensuring that their needs are satisfied (Beck et al., 2001).
2. "Welcome Changing Requirements Even Late in the Project". Agile methods leverage change to the benefit of the client's competitive edge (Highsmith, 2002).
3. "Deliver Value Frequently". Regular delivery of functional software facilitates early feedback and modifications (Schwaber & Beedle, 2002).
4. "Break the Silos of Your Project". This principle guarantees continuous communication and alignment amongst all parties involved thanks to the team's cross-functionality (Cockburn, 2006).
5. "Build Projects Around Motivated Individuals". It's critical to create the conditions and resources necessary for team members to flourish (Cockburn & Highsmith, 2001).
6. "The Most Effective Way of Communication is Face-to-face". Indeed, this way of communicating minimizes miscommunication and expedites decision-making (Beck et al., 2001).
7. "Working Software is the Primary Measure of Progress". This statement is rather straightforward as it emphasizes that the most significant indicator of progress

- in a project is the delivery of functional, usable software, not the amount of documentation, time spent on meetings, or adherence to a strict project plan.
8. "Maintain a Sustainable Working Pace". Sustainable development is facilitated by agile processes as the group should be able to work at a steady pace indefinitely (Highsmith, 2002).
 9. "Continuous Excellence Enhances Agility". Therefore, agility is increased by constant focus on technical excellence and good design while it cannot be maintained without high standards in technical and design practices (Cockburn, 2006).
 10. "Simplicity, the art of maximizing the amount of work not done, is essential". Project management teams should focus on what is necessary and eliminate waste as much as possible (Beck et al., 2001).
 11. "Self-organizing Teams Generate Most Value". Indeed, empowered teams are more likely to create innovative solutions (Cockburn & Highsmith, 2001).
 12. "Regularly Reflect and Adjust Your Way of Work to Boost Effectiveness". Considering ways to improve effectiveness on a regular basis and then fine-tuning and modifying behavior accordingly is an essential component of Agile (Highsmith, 2002).

The aforementioned Agile principles and values can be implemented through several frameworks and, among some of the most popular, there are Scrum, Kanban, Extreme Programming, and Lean Software Development. Scrum is a framework that breaks work up into fixed-length iterations called "sprints", outcome of which should always be a potentially shippable product increment. The Product Owner, Scrum Master, and Development Team are essential roles in Scrum, and among the crucial elements of the Scrum process we find daily stand-up meetings, sprint planning, sprint review and sprint retrospective (Schwaber & Sutherland, 2020). Kanban emphasizes limiting work-in-progress and visualizing the workflow to increase productivity (Nowak & Kowalski, 2022). Tasks are moved across a Kanban board's columns, which represent the various

stages of the workflow, as they are completed (Anderson, 2010). Through techniques like pair programming, test-driven development, continuous integration, and frequent releases, Extreme Programming prioritizes both technical excellence and customer satisfaction (Beck, 1999). Finally, lean principles concentrate on reducing waste, integrating quality, and delivering results quickly (Womack, J. P., & Jones, D. T., 1996).

Agile practices offer numerous benefits that are described thoroughly in literature. First off, teams can react swiftly to modifications in requirements, technology, and market conditions thanks to Agile's iterative nature (Schwaber & Sutherland, 2017; Serrador & Pinto, 2015). According to Moe, Dingsøy, and Dybå (2010), there is an improvement in comprehension and purpose alignment as a result of the enhanced continuous interaction and communication between team members and stakeholders. Additionally, Agile approaches guarantee that the finished product satisfies the needs and expectations of the clients by incorporating them in the development process (Beck et al., 2001; Williams & Cockburn, 2003). As a result, the risk of project failure declines as regular delivery of functional software provides early value to customers (Livermore, 2007; Pichler, 2010). Lastly, Agile practices like continuous integration and regular retrospectives help maintain high quality and encourage continuous improvement (Fowler & Highsmith, 2001; Derby, Larsen, & Schwaber, 2006). Clearly, with Agile practices come not only benefits but also challenges, such as cultural resistance faced by companies in which traditional methodologies are deeply rooted or the misinterpretation of Agile principles, which can lead to unsatisfactory results (Conboy, 2009; Dingsøy, Nerur, Balijepally, & Moe, 2012).

2.2 Scrum Framework

Scrum is a widely used framework within Agile, designed to enhance productivity, foster collaboration, and deliver high-quality products (Sundararajan & Subramanian, 2022). Initially conceptualized in the early 1990s by Ken Schwaber and Jeff Sutherland, Scrum provides a framework that enables teams to address complex problems while productively and creatively delivering products of the highest possible value (Moe,

Dingsøy, & Dybå, 2010). Research has highlighted Scrum's effectiveness in improving project outcomes, particularly in complex and dynamic environments where flexibility and iterative development are crucial (Serrador & Pinto, 2015).

The Scrum framework structures development work into iterative cycles, known as "sprints", that typically last two to four weeks. As already mentioned, each sprint aims at delivering a potentially shippable product increment, ensuring continuous improvement and regular delivery of value (Rubin, 2012). The Scrum framework includes several key components and roles designed to ensure effective project management and product development (Moe, Dingsøy, & Dybå, 2010). These key pillars it is composed of are the Scrum Team, Scrum Artifacts and Scrum Events. The developers, the Product Owner, and the Scrum Master make up the Scrum team, and their goal is to collaborate to accomplish the product goal (Rubin, 2012). Two essential qualities that the Scrum Team must have are cross-functionality and self-management, according to the Scrum Guide (Schwaber & Sutherland, 2020). This implies that the tasks to be completed are divided internally in a flexible way and that every team member needs to possess the skills required to add value in each Sprint (Gren, Torkar, & Feldt, 2015). In addition, a Scrum Team should be composed of a maximum of ten people to ensure the best possible productivity and communication is accomplished (Rubin, 2012). The Scrum Master has the key task of ensuring that each Scrum Team member applies the Scrum framework as defined in the Scrum Guide (Schwaber & Sutherland, 2020). Hence, he could be described as a supervisor who makes sure the Scrum process is run properly. The Product Owner, on the other hand, is responsible for maximizing the value of the product resulting from the work of the Scrum Team (Kuusinen, Mikkonen, & Pakarinen, 2016). Furthermore, he is responsible for the effective management of the Product Backlog through developing and communicating transparently the Product Goal, creating Product Backlog items, ordering these items through the prioritization process and ensuring that the Product

Backlog is visible and clear to all team members (Kuusinen et al., 2016). Finally, Developers are defined as all team members who work on the continuous development of the product who do not hold the role of Scrum Master or Product Owner. "Artifacts" represent the second pillar of Scrum framework and are defined in the Scrum Guide as a representation of work or value designed to maximize the transparency of key information (Schwaber & Sutherland, 2020). The three main elements that compose this category are the Product Backlog, Sprint Backlog and Increment (Gren et al., 2015). The Product Backlog is a dynamic, ordered list of what is needed to make a product improvement and it evolves as the Scrum Team receives additional information and feedback from stakeholders (Zhao & Li, 2023). Product Backlog Refinement activities aim to break down the items in the Product Backlog into smaller work units and reorder these items through a prioritization process (Strode, 2016). The Sprint Backlog is a tool consisting of the Sprint Goal, the Product Backlog items selected for a specific Sprint by means of Backlog refinement and a plan to achieve the Increments (Rubin, 2012). Finally, Increments are defined as concrete product improvements towards the achievement of the Product Goal and each of them is a potentially shippable, functional version of the product that adds value and builds on previous work (Rubin, 2012).

The Scrum Events listed in Schwaber and Sutherland's Scrum Guide (2020) are divided into Sprints, Sprint Planning, Daily Scrum, Sprint Review and Sprint Retrospective. Sprints are defined as events of fixed length, generally lasting between two and four weeks, in which all the work necessary to achieve the Product Goal takes place (Moe, Šmite, & Šāblis, 2019). According to several studies, sprints have demonstrated to improve team productivity and reduce the risk of project failure because, unlike traditional waterfall development models, teams can correct errors early and adapt to changes (Rubin, 2012). Sprint Planning is an event that starts the Sprint by defining the work to be done in it through a decision-making process involving the entire Scrum

Team (Schwaber & Sutherland, 2020). Academic literature points out that sprint planning is divided into two distinct phases. The first, primarily driven by the product owner, involves deciding which items contained in the product backlog will be implemented in a given sprint. The second phase involves the developers and consists of deciding how those items will be implemented through work planning activities (Rubin, 2012). Research shows that focusing on the Sprint Planning phase significantly reduces the risks associated with excessive workload and improves overall team productivity (Sharma & Gupta, 2023). The Daily Scrum is defined as a short, daily meeting where the team synchronizes activities and plans for the next 24 hours, ensuring alignment and quick problem resolution (Schwaber & Sutherland, 2020). According to Moe et al. (2019), Daily Scrum promotes transparency and individual accountability as team members regularly share their progress and openly address any problems. This meeting also helps keep the focus on the Sprint goals, reducing the risk of unnecessary deviations (Santos et al., 2020). The Sprint Review is a meeting that takes place at the end of each sprint to review the work completed and gather useful feedback, fostering continuous improvement and stakeholder engagement (Strode, 2016). Finally, the Sprint Retrospective is a meeting for the team to reflect on the sprint, identify what went well, and pinpoint areas for improvement, which is essential for team learning and process enhancement (Kuusinen et al., 2016; Strode, 2016). A study by Kuusinen et al. (2016) showed that the effectiveness of retrospectives depends on the team's ability to openly address issues and implement proposed changes. Without these characteristics, retrospectives risk becoming a formal exercise with no real impact.

2.2.1 Product Backlog composition

The product backlog can consist mainly of 'Items' and 'Fields'. The "items" are the individual units of work or requirements to be implemented during the development of a product, while the "fields" are the attributes or characteristics that describe each individual backlog item (Rubin, 2012). Various types of items can be found in literature

such as user stories, bugs, tasks, themes, epics, spikes and enablers (Schwaber & Sutherland, 2020). According to Rubin (2012), every item needs to have a precise description, acceptance criteria, and an estimate of the work needed to finish it. Items may vary in size and detail since higher priority items are typically smaller and more detailed, ready to be included in an upcoming Sprint, while lower priority items may be larger and less defined, waiting to be refined in the future (Ramesh et al., 2010). Scrum backlog items are frequently formatted as "user stories", which offer a straightforward but effective means of capturing functional requirements as seen from the viewpoint of the user (Cohn, 2008). In order to ensure that the focus stays on providing value to the user and that the requirements are understandable, a typical user story format is as follows: "As a [user], I want [function] so that [benefit]" (Cohn, 2008). User stories should be small enough to be completed within a single sprint and detailed enough to provide clear guidance to the development team (Schwaber & Sutherland, 2020). Moreover, they should include acceptance criteria that define the conditions under which the story is considered complete such that all team members have a shared understanding of what needs to be done and how success will be measured (Cohn, 2008). "Bugs" and "technical tasks" are also essential components of the Product Backlog. Indeed, bugs represent defects or issues in the product that need to be fixed, while technical tasks involve work required to support development, such as setting up environments, refactoring code, or conducting research (Rubin, 2012). On the other hand, research tasks involve exploratory work needed to gather information, validate assumptions, or investigate potential solutions (Gren, Torkar, & Feldt, 2015). These tasks are particularly important in complex projects where uncertainty is high and, according to Rubin (2012), they should be time-boxed and have clear objectives to ensure they provide value without consuming excessive resources. An "epic" is a large, high-level piece of work that is too big to be completed in a single work iteration or sprint (Cohn, 2006). They stand for important characteristics or features of the product and they offer a means of capturing broad

requirements without first becoming trapped in specifics (Rubin, 2012). "Themes" are broad, overarching goals or high-level strategic objectives that direct the course of the product (Kuusinen, Mikkonen, & Pakarinen, 2016). They are extensive and cover several connected epics and user stories and, although they are not immediately applicable, they greatly aid in high-level work organization and prioritization (Kuusinen et al., 2016). "Spikes" are time-limited research projects designed to find answers, acquire data, or lower project uncertainty (Schwaber & Sutherland, 2020). Deliverable product features are not directly produced by them but, by using spikes, teams can address ambiguities and arrive at well-informed decisions without having to commit to a particular implementation up front (Rubin, 2012). Moreover, spikes are essential for both technical exploration and risk management. Finally, tasks related to technical infrastructure or architectural design that facilitate feature development are known as "enablers" (Schwaber & Sutherland, 2020). These items are crucial in making sure the technical foundation is stable and scalable, which helps with the delivery of user stories and epics (Kuusinen et al., 2016). As far as "fields" are concerned, the most commonly used in Scrum are ID, Title, Description, Priority, Assigned Sprint, Status, Size, Acceptance Criteria, Assignee and Dependencies between Items. The "ID" is a unique identifier that allows an item to be tracked within the product backlog (Rubin, 2012). The "Title" consists of a short title that briefly represents an item and is useful for quickly identifying an item (Cohn, 2008). The "Description" explains what needs to be implemented to complete the item to which it is associated and must be detailed enough to be understood by the development team and provide sufficient context for the work to be done (Ramesh, Cao, & Baskerville, 2010). The "Priority" is usually set by the Product Owner and allows the Scrum Team to understand which items need to be implemented most urgently. An "Assigned Sprint" indicates which Sprint should be implemented in a particular Product Backlog Item. The "Status" can be distinguished as "To Do", "In Progress" or "Done" and provides information about the progress of work on a particular Item (Rubin, 2012). The "size" of an item is obtained through an

estimation process and determines the time and effort required to complete it (Cohn, 2008). An "Acceptance Criteria" defines the conditions that must be met for a specific backlog item to be considered completed (Moe et al., 2010). An "Assignee" identifies the specific person within the Scrum Team who is responsible for the implementation of a Product Backlog Item. The "assignee" may be subject to change over time, depending on the priorities or capabilities of the team (Rubin, 2012). Finally, "dependencies" represent the interdependencies between backlog items. Indeed, some items may depend on the completion of others before they can be started or completed. Identifying and managing dependencies is crucial to avoid blockages and ensure smooth development (Moe, Dingsøy, & Dybå, 2010).

2.2.2 Backlog Refinement

The process of reviewing and updating the Product Backlog by the development team and the product owner is called "backlog refinement" (or "grooming"). This task entails re-estimating effort, adding details to backlog items, breaking down large items into smaller tasks, and re-prioritizing items in light of newly acquired information or input. Frequent backlog refinement meetings guarantee that items are prepared for the next sprints and help maintain the backlog updated (Schwaber & Sutherland, 2020). Indeed, empirical studies have shown that teams that regularly perform refinement have a better understanding of project priorities and greater accuracy in effort estimates, leading to greater efficiency and fewer delays in Sprints (Ramesh et al., 2010; Schön et al., 2017). Moreover, academic literature indicates that refinement helps prevent the product backlog from becoming too large or disorganized by ensuring that only the most valuable items remain active (Kropp et al., 2014). The main objective of a Product Backlog refinement is to ensure that the items in the Product Backlog are in the so-called "ready state". This condition is reached when the Development team is confident that the items are clear enough for stakeholders to understand what they need and small enough to be completed in a single sprint. The interaction between the Development Team, stakeholders, and the Product Owner is

of paramount importance in this activity (Cohn, 2005; Schwaber & Sutherland, 2020). Before a Product Backlog item achieves the "ready state", it typically undergoes three refinement meetings. Initially, the team makes an educated guess on how big the item would be when a stakeholder presents an idea. Assigning "story points" to the item is the second step, which is also done quickly and imperfectly (Grenning, 2002). The method that is frequently employed is "magic estimation", also known as silent estimation, which involves estimating without having prolonged, in-depth conversations about the item (Kua, 2013). "Planning poker", a commonly used technique for item estimation, is the last step before an item is deemed ready by a development team. However, the fact that it is considered rather time-consuming makes its use feasible only for those items the team is convinced to implement and considers as valuable (Cohn, 2005).

2.2.3 Items Estimation

Although the originators of Scrum, Ken Schwaber and Jeff Sutherland, emphasized the simplicity of the framework and did not initially place great emphasis on detailed estimates, the value of estimates for better planning and expectation management emerged over time. Estimation techniques help teams in understanding the effort required for each backlog item, enabling better planning and resource allocation. Estimation of Product Backlog elements is typically based on relative estimation methods, such as "story points", rather than absolute estimates in hours or days (Cohn, 2006). Indeed, the literature shows that relative estimates of items tend to be more accurate and flexible than absolute estimates, especially in contexts characterized by uncertainty such as agile projects (Moe et al., 2010). Several estimation techniques are widely used in Scrum but, in this specific literature review, only the mostly used are going to be mentioned as the main goal of this study is more related to how estimates are done more than what are the specific techniques used. However, the main estimation techniques are the following: Planning Poker, Work System Snapshots and T-shirt sizing (Cohn, 2005). "Planning Poker" is a consensus-based estimation

technique that involves team members independently estimating the effort required for a backlog item using cards. Each card has a number representing the effort estimate, team members reveal their cards simultaneously, and if there is a significant discrepancy, they discuss their estimates and rationale (Cohn, 2005). The main advantages of "Planning Poker" are based on the fact that it promotes equal participation, mitigates the influence of dominant voices, and leverages the collective intelligence of the team (Mahnič & Hovelja, 2012; Babiker, Mahmoud, & Abdalrahman, 2018). The second technique we want to analyze is known as "Work system snapshots" and provides a structured approach to understanding the broader context of user stories and refining them more effectively. This technique captures elements such as participants, processes, and technologies involved in a work system and, by analyzing these elements, teams can produce higher quality and more focused user stories, leading to better estimates and prioritization (Cohn, 2005). This method helps ensure that all relevant aspects of a user story are considered, reducing the likelihood of overlooked dependencies or requirements (Bolloju et al., 2018). Finally, "T-Shirt Sizing" is a quick and straightforward estimation technique where backlog items are categorized into different sizes (e.g., XS, S, M, L, XL) based on their perceived effort. This method provides a relative estimation scale that is easy to understand and communicate (Cohn, 2005). While it is less precise than other techniques, it is useful for initial estimation and prioritization during backlog refinement sessions (Schwaber & Sutherland, 2020).

2.2.4 Product Backlog Prioritization

Prioritizing the Product Backlog is essential for ensuring that the most valuable and critical work is completed first (Kumar & Patel, 2022). Ramesh et al. (2010) argue that effective priority management reduces the risks associated with non-delivery of critical functionality and improves customer satisfaction. The Scrum Guide does not define in detail how to apply prioritization to the items in the Product Backlog, but it merely states that the Product Owner is responsible for the orderly maintenance of the Product

Backlog. However, various techniques can be used, each offering different benefits depending on the project's context. During the past two decades the main techniques have been modified and adapted to specific contexts thus generating a vast variety of different techniques but, in this literature review, we are going to describe only the main and purest methodologies. Firstly, the "MoSCoW" method categorizes requirements into four groups: "must have", "should have", "could have", and "won't have". This method provides a straightforward way to prioritize features based on their importance and urgency. "Must-have" items are critical to the product's success and must be included in the release. "Should-have" items are important but not essential, "could-have" items are desirable but not necessary and "won't-have" items are least important and can be deferred or omitted (Ashmore & Runyan, 2014). "Kano analysis" helps identify features that are essential, those that increase user satisfaction, and those that delight users. Features are categorized into "basic needs", "performance needs", and "excitement needs". "Basic needs" are essential for the product to function, "performance needs" enhance user satisfaction, and "excitement needs" provide unexpected delight to users. By understanding these categories, teams can prioritize features that provide the most significant impact on user satisfaction (Cohn, 2008). The "Analytical Hierarchical Process" is a structured technique that uses pairwise comparisons to rank items based on various criteria such as business value, risk, and cost. This approach is particularly useful for complex projects with many decision-makers as it provides a systematic approach to decision-making by breaking down complex problems into simpler comparisons, making it easier to prioritize items objectively (Hudaib et al., 2018). The "Big Wall Method" is a visual approach where all backlog items are placed on a large board and categorized based on their size and value. This helps teams quickly identify high-value, high-effort tasks and prioritize them accordingly. The visual representation makes it easier to see the overall structure of the backlog and facilitates discussions about prioritization and dependencies (Lacey, 2015).

Prioritizing product backlog items is a process that involves two main difficulties. The first is balancing the technical needs of the team with meeting customer requirements. In fact, the literature points out that business needs sometimes conflict with technical constraints, making it difficult to make balanced prioritization decisions (Moe et al., 2010). The second difficulty is the lack of complete information about backlog items, which can lead to incorrect or inaccurate value estimates (Racheva et al., 2010).

2.3 Research questions

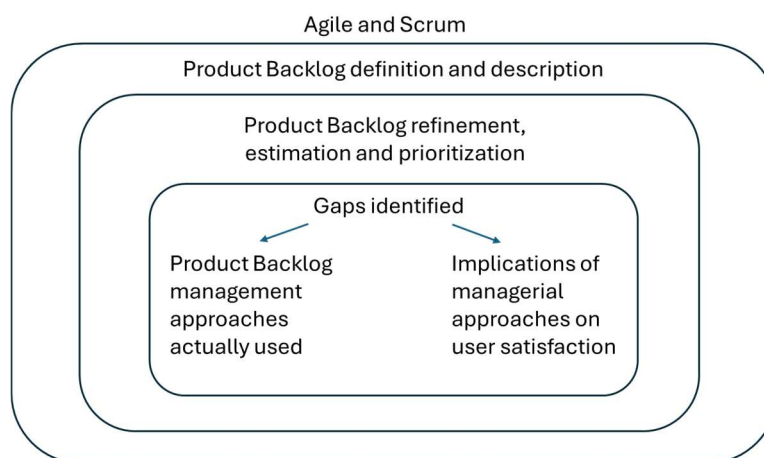


Figure 1; Literature review framework and gaps identified

Following this analysis of the literature, we have gained a deeper understanding of many aspects related to Product Backlog management practices. This initial analysis has undoubtedly highlighted the high degree of complexity and the various ways in which the Product Backlog can be managed. However, what is not easily apparent from the literature and studies conducted in past years is how professionals who operate within the agile framework actually implement the practices described in theory. The primary objective of this study, therefore, is to understand, through methodologies that will be described in the following chapters, from a more practical perspective, how the Product Backlog is actually managed by professionals using the Scrum framework. Specifically, this study aims to answer the following research questions:

- I. Are there specific approaches to implementing key Product Backlog management practices that are more commonly utilized than others?

- II. What are the key factors that most significantly influence satisfaction levels in Product Backlog management? Additionally, which Product Backlog management approaches should companies adopt to maximize satisfaction?
- III. Does the experience of professionals with Scrum influence their approach to managing the Product Backlog?

3 - Methodology

The purpose of this chapter is to clarify the approach and techniques used to answer the research questions listed in the final part of the literature review. The research adopts a mixed-method approach to explore practical applications of Scrum, focusing on Product Backlog management practices. The study begins with a comprehensive literature review to establish a theoretical foundation and identify gaps in existing research. A structured questionnaire was designed and disseminated through LinkedIn, targeting Scrum practitioners, including Product Owners and Scrum Masters, across various industries. The collected data, comprising responses from 216 participants, underwent rigorous statistical analyses, including descriptive statistics, factor analysis, multiple regression, and k-means cluster analysis. These methodologies enabled the identification of prevailing practices, their impact on satisfaction levels, and distinctions in management approaches based on experience with Scrum, offering actionable insights for both academia and practitioners.

Firstly, it was necessary to create a questionnaire in order to obtain the information required to further investigate the subjects covered in this study, as the existing data was insufficient. This was accomplished through the use of Qualtrics, a well-known cloud-based software platform that enables users to create, distribute, and analyze surveys and other forms of feedback. Subsequent to the dissemination of the questionnaire and the aggregation of responses, the analysis proceeded on SPSS through three distinct methodologies. First, a descriptive analysis aimed at capturing the main tendencies in the responses given by respondents. Second, a factor analysis was implemented in order to obtain a reduced number of factors that encapsulated as many original variables as possible. Next, a multiple regression was performed with the aim of understanding the extent to which previously obtained factors influenced the degree of satisfaction felt by respondents in managing the Product Backlog. Finally, a cluster analysis was implemented aimed at understanding whether it is possible to

identify different styles of Product Backlog management based on respondents' years of experience in implementing the Scrum framework.

3.1 Questionnaire

As previously stated, the questionnaire was constructed using Qualtrics due to the advanced customization it allows, its ease of use, and its direct integration with well-known statistical analysis software (i.e., SPSS). The questionnaire comprised a total of 55 questions, distributed across six sections according to the topics investigated and the subjects explored in the literature review. The structure of the questionnaire is described in the following table.

| Questionnaire structure | |
|--|-----------------------|
| Section | # of questions |
| 1) Respondents' information | 6 |
| 2) Product Backlog composition | 2 |
| 3) Stories | 9 |
| 4) Estimates | 13 |
| 5) Prioritization | 8 |
| 6) General Product Backlog management | 17 |

Table 1; Number of questions for each section of the questionnaire

3.1.1 Section One: Respondent Background and Experience

The first section of the questionnaire aimed to gather useful information about respondents. It thus consisted of questions regarding the sector in which the respondents operated, the number of employees in the company in which they worked, the role held within their team, and the number of members that composed their team. Finally, the nature of the project or product being considered, which may have had an internal or external purpose to the company, and the respondent's years of experience in using the Scrum framework were investigated. In order to avoid confusion in the answers, which would have led to less reliability and consistency in the data collected, the respondent was asked to answer the questions by considering a specific product or project.

3.1.2 Section Two: Product Backlog Composition

In the initial chapter, the literature review enabled us to ascertain the theoretical constituents of the product backlog. Therefore, the second block of questions in the questionnaire consisted of two questions designed to understand the actual composition of the Product Backlog. Specifically, the first question focused on how the items within the Product Backlog are classified and the following 8 different options were given: themes, epics, user stories, tasks, spikes, bugs, enablers, and others. The second question, on the other hand, was more focused on what fields are used in the construction of the Product Backlog. Following the analysis carried out in the literature review, the options available to the respondent in this case were 11: ID, title, description, priority, assigned sprint, status, size, acceptance criteria, assignee, dependencies between items, and others. In both questions, the option of giving more than one answer was provided due to the possibility to include a large number of items and fields in the Product Backlog.

3.1.3 Section Three: Stories

The third section of the questionnaire consisted of 9 questions and was focused on the use, creation and management of stories, following the topics explained in the literature review. In order not to exclude from the questionnaire those respondents who do not use the typical User Story syntax, it is clarified that the term "Stories" means the smallest unit of work and not specifically User Stories. Therefore, after an introductory question asking for the number of items in the Product Backlog described by the typical User Story syntax (i.e. As I..., I would like to..., so that...), there were 8 sentences whose possible answers followed the logic of a Likert scale. The extremes of these scales ranged from "Strongly Agree" to "Strongly Disagree" and included 5 other intermediate responses. These statements focused in particular on the people responsible for creating the stories and breaking them down into smaller tasks. The aim here was also to investigate deeper the degree of collaboration used in the management of stories.

3.1.4 Section Four: Estimates

The fourth block of the questionnaire consisted of 13 questions related to the estimation of items within the Product Backlog. These questions were constructed on the basis of the arguments raised in the literature review about how estimates are generated. Again, the section began with a question focused on the number of items that have undergone an estimation process within the respondent's Product Backlog. This was followed by 12 statements, the answers to which followed once again a Likert scale as described above. The following aspects were investigated at this stage: who is responsible for creating the estimates, how often the items are estimated and how these estimates are made. Thus, the answer to these questions aimed to understand which approaches are most commonly used by respondents in managing and creating estimates.

3.1.5 Section Five: Prioritization

The fifth section of the questionnaire consisted of 8 questions focused on the process of item prioritization which was deepened through the review of scientific and academic literature. Since it is widely believed that item prioritization is widely used in the Scrum framework but no clear evidence of this phenomenon can be found in the literature, this part of the questionnaire started with a question aimed at understanding how many items within the Product Backlog are generally subject to a prioritization process. Those who gave a negative answer to this question would have been excluded from the following questions about the prioritization process and would have gone directly to the next stage of the questionnaire. As in the previous sections of the questionnaire, there were 7 statements whose answers follow a Likert scale. The purpose of these statements was to specifically investigate the frequency with which the process of item prioritization takes place and who is responsible for managing this process.

3.1.6 Section Six: General Product Backlog Management Practices

The last block of the questionnaire consisted of 17 questions, also based on a Likert scale as described above, which aimed to investigate additional practices implemented by the respondents in the management of the Product Backlog. Following the analysis carried out in the literature review section, the focus here was mostly on the structure of sprints, the degree of accessibility of the Product Backlog and the information it contains, its customization according to customer characteristics, and the general difficulties experienced by respondents in managing the Product Backlog. Finally, this section concluded with an open-ended question asking respondents what they would like to change in the management of their Product Backlog, in order to obtain more specific and qualitative responses on this topic.

3.1.7 Questionnaire's dissemination

Once constructed, the questionnaire was distributed via LinkedIn, a well-known professional social networking platform used to create connections between individuals and companies. The decision to use this platform was based on the numerous advantages offered by LinkedIn. Firstly, its ability to allow users to easily reach a highly targeted audience with skills relevant to the research in question. Secondly, the significant presence of groups that make it possible to quickly reach a large number of professionals who are specialized and interested in the topic of this thesis. Finally, the presence of a private chat, which made it possible to have in-depth conversations on the topics covered in the questionnaire. The target respondents were primarily Scrum Masters and Product Owners, as they are generally the ones within a Scrum Team who have a higher level of experience in Product Backlog Management issues. These professionals were reached through a connection request, followed at least 72 hours after its confirmation by a private chat message aimed at involving the individual in the collection of questionnaire responses. On the other hand, no targets were set in terms of geography, age or gender, as these were considered to be parameters with little influence for the purposes of the research.

Initially, the questionnaire underwent rigorous testing to ensure that the questions were easily understandable and clear to the respondents. Indeed, "pre-testing" is a crucial and fundamental step that significantly impacts the quality of the questionnaire. During this phase, the survey was administered to a small sample of participants, typically 15 to 30 respondents are considered sufficient for this purpose (Malhotra, 2006). It was essential that these participants closely resembled the main target groups in terms of background, educational levels, interests, attitudes, and familiarity with the subject matter. This similarity ensured that various aspects of the questionnaire, such as wording, question flow and sequence, format, and content, could be effectively evaluated and refined (Malhotra, 2006).

Therefore, the pre-testing phase of the questionnaire was conducted following these outlined procedures. Specifically, a few days after sending connection requests on LinkedIn to 20 professionals identified using keywords such as "scrum master," "product owner," or "scrum certified" in the search bar, the questionnaire was sent to these individuals via the platform's private messaging section. This targeted approach ensured that the responses were relevant and came from individuals with appropriate knowledge and experience, thereby enhancing the reliability of the data (Leonardi et al., 2013). Upon collecting 15 responses and receiving positive feedback confirming the well-structured nature of the questionnaire and the clarity of the questions posed, the next phase commenced: the actual collection of responses. Using the same procedure to connect with the aforementioned target group of professionals, 216 complete responses were gathered, which were subsequently subjected to statistical analysis.

3.2 Statistical Analysis

Following the response collection phase and before the actual analysis of data, the data cleaning phase commenced. This phase was crucial to ensure that any errors in the dataset did not compromise the subsequent statistical analyses and their reliability. Data cleaning was divided into three main parts: checking for errors, finding the error

in the data file, and correcting the error (Pallant, 2020). When checking for errors, the primary focus was on identifying values that fall outside the possible range for a given variable. This was conducted using the tools provided by SPSS, one particularly suited for checking categorical variables and another for continuous variables (Pallant, 2020). Specifically, there were two fundamental parts of the output that must be checked: the maximum and minimum values and the number of valid and missing cases. First, the maximum and minimum values must fall within the expected range of values for each question in the questionnaire. Second, a significant number of missing cases would indicate a lack of data that could affect subsequent statistical analyses. In this instance, the review of the data and the application of the data-cleaning phase derived from the questionnaire revealed no notable defects or issues, thereby allowing the progression to the statistical analysis phase (Pallant, 2020).

The analysis was conducted through four main methodologies:

1. **Descriptive Analysis:** This involved summarizing the basic features of the data, providing simple summaries about the sample and the measures. Descriptive statistics are crucial for understanding the general tendencies and patterns in the data (Field, 2013). In this study, the main trends regarding the characteristics of the respondents were firstly highlighted by means of explanatory graphs. Secondly, the mean, median, mode and standard deviation were calculated for each question in the questionnaire in order to get an initial idea of what were the main approaches that were being implemented in Product Backlog Management.
2. **Factor Analysis:** This technique is used to identify underlying relationships between variables by grouping them into factors. Factor analysis reduces the number of variables by combining those that are correlated, thus simplifying the data structure without losing significant information (Hair, Black, Babin, & Anderson, 2019). In this case, a Principal Component Analysis was carried out,

from which 16 factors were extracted. Finally, through a qualitative analysis of the extracted factors and the variables contained in them, they were given a name that summarized the main information contained in each of them. This step proved to be very useful as it greatly simplified the interpretation of the results obtained from the following multiple regression.

3. Multiple Regression: This statistical technique is used to understand the relationship between one dependent variable and two or more independent variables. Multiple regression analysis helps in identifying which factors have the most significant impact on the dependent variable (Hair et al., 2019). In this study, multiple regression was used to understand which of the extracted factors had the greatest and most significant impact on the respondents' level of satisfaction with Product Backlog management (i.e. the dependent variable). The independent variables used corresponded to the 16 factors extracted from the Factor Analysis explained above. This helped to significantly reduce multicollinearity, scale down the dataset and improve the interpretability of the results. After using the SPSS tool to carry out the multiple regression, output tables were obtained that allowed a positive judgement to be made about the regression. In particular, considerable attention was paid to the R^2 , Adjusted R^2 and Standard Error values, as they are fundamental to understanding the goodness of the regression model. Also of particular importance was the analysis of the ANOVA table to understand the variance explained by the model and the residual variance, i.e. the part of the variance that the model cannot explain, as well as the F-test and the associated p-value. Another aspect of fundamental importance for a positive assessment of the multiple regression model is the absence of multicollinearity between the predictors. Multicollinearity occurs when two or more independent variables in a regression model are highly correlated, which can negatively affect the reliability of the regression coefficients. In this study, the absence of multicollinearity between

the independent variables was verified by observing two values related to this aspect, tolerance and VIF. Finally, to assess the assumption of homoscedasticity and normality of the residuals of the regression model, the scatterplot and the normal P-P plot were observed before analyzing the results. The results obtained here have strong managerial implications as they enable those utilizing Scrum to identify which practices have the greatest impact on satisfaction in product backlog management.

4. **Cluster Analysis:** Cluster analysis helps in identifying distinct groups within the data, which can provide insights into different segments of the population (Kaufman, 2009). In this study, a k-means cluster analysis was used, which is an unsupervised machine learning technique used to group a set of data into groups or "clusters" based on the similarity of their characteristics. The number of clusters, 'K', is defined by the user before running the algorithm and each data item is assigned to the cluster with the closest "centroid", which represents the midpoint of all data in the cluster. In this research, the cluster analysis was based on two clusters that differed from each other based on the experience of using the Scrum framework. Specifically, one cluster was identified as the group of respondents with limited experience and another with extensive experience in using Scrum. Based on this distinction, four conclusions were drawn regarding the implementation of specific practices in Product Backlog Management.

4 - Results

In this section of the research, the results obtained in each of the 4 different parts of the statistical analysis will be analyzed in detail. From the descriptive analysis we expect to understand what the most common ways are of managing the Product Backlog, broken down into the different sections of the questionnaire: composition of the Product Backlog, implementation of stories, use of estimates, the prioritization process and other general practices in Product Backlog management. From the results of the factor analysis, we expect to understand which factors were extracted, what they mean for the research purposes and how they can be useful for further analysis. From the multiple regression section and its results, we expect to understand which practices in Product Backlog Management lead to an actual and significant increase in respondents' satisfaction. Finally, from the results of the cluster analysis, we will be able to understand whether the years of experience in using the Scrum framework are a reliable variable for distinguishing different ways of managing the Product Backlog.

4.1 Descriptive Analysis

4.1.1 Respondents' information

The descriptive analysis represents the initial stage of statistical examination conducted in this study. This analysis was primarily undertaken to identify and highlight the presence of general tendencies for each question within the questionnaire (Creswell & Creswell, 2017). Specifically, the first part of the descriptive analysis aims to understand the distinctive characteristics of the respondents who contributed to the data collection (Field, 2013). These characteristics include the sector in which the respondents operate, the number of employees (indicative of the company size) in the organization where they work, the number of individuals in their Scrum team, the years of experience they possess in using Scrum, and whether the project/product they reference in their subsequent responses has an internal/company-focused purpose or an external/client-focused purpose.

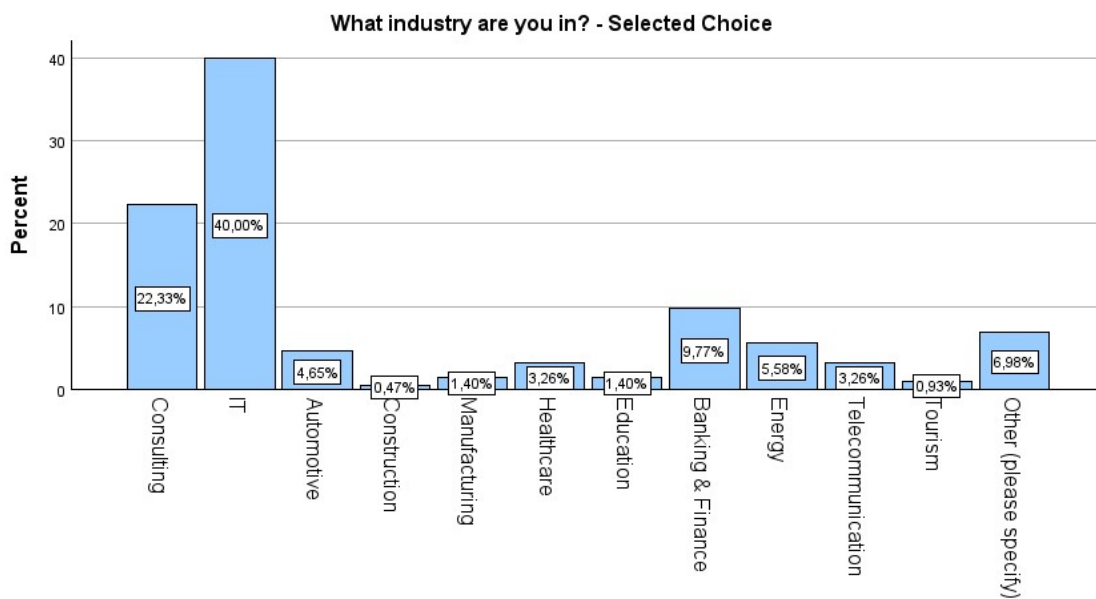


Figure 2; SPSS; Industry Representation in Survey Respondents

As observed in the analysis of Figure 1, the industries represented by the respondents are primarily two: IT, with a respondent percentage of 39,7%, and consulting, which accounts for nearly 24% of the respondents. These findings are not particularly surprising, as the Scrum methodology is predominantly used in the IT field and often finds useful applications in consulting.

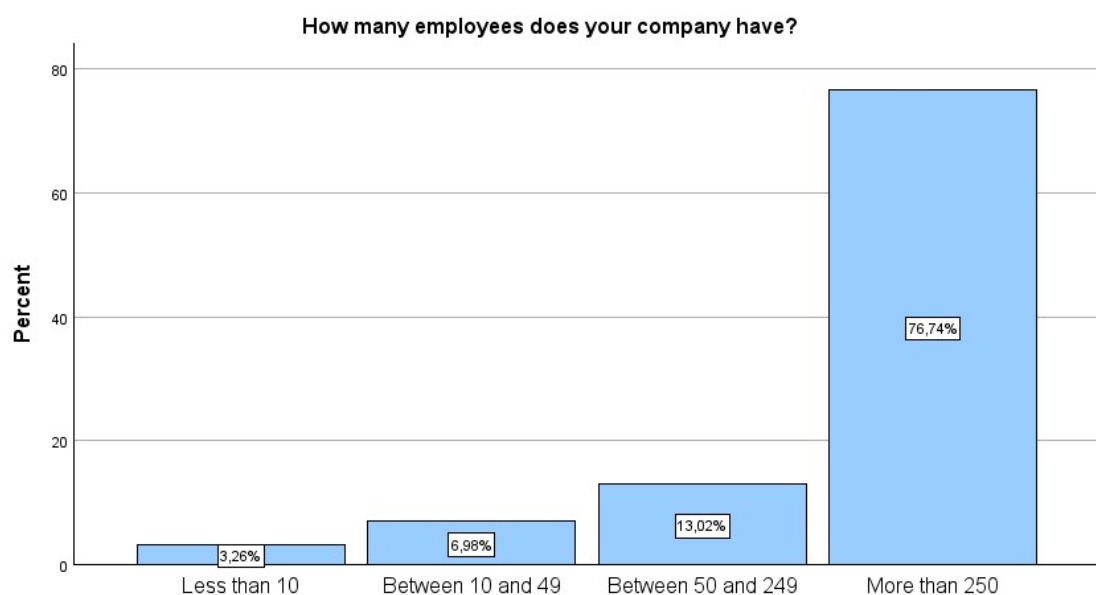


Figure 3; SPSS; Firm's dimension

The bar chart represented in Figure 2, illustrates the distribution of company sizes

based on the number of employees, as reported by the respondents. The data shows that a significant majority of respondents (76,74%) work in companies with more than 250 employees, indicating a strong presence of large enterprises in the sample. In contrast, smaller companies are less represented: 13,02% of respondents are from companies with 50 to 249 employees, 6,98% are from companies with 10 to 49 employees, and only 3,26% are from companies with less than 10 employees. This distribution suggests that the study's findings may predominantly reflect the practices and conditions of larger organizations, a trend that highlights the prevalence of Scrum and Agile methodologies in larger companies due to their complex project management needs (Schwaber & Sutherland, 2017).

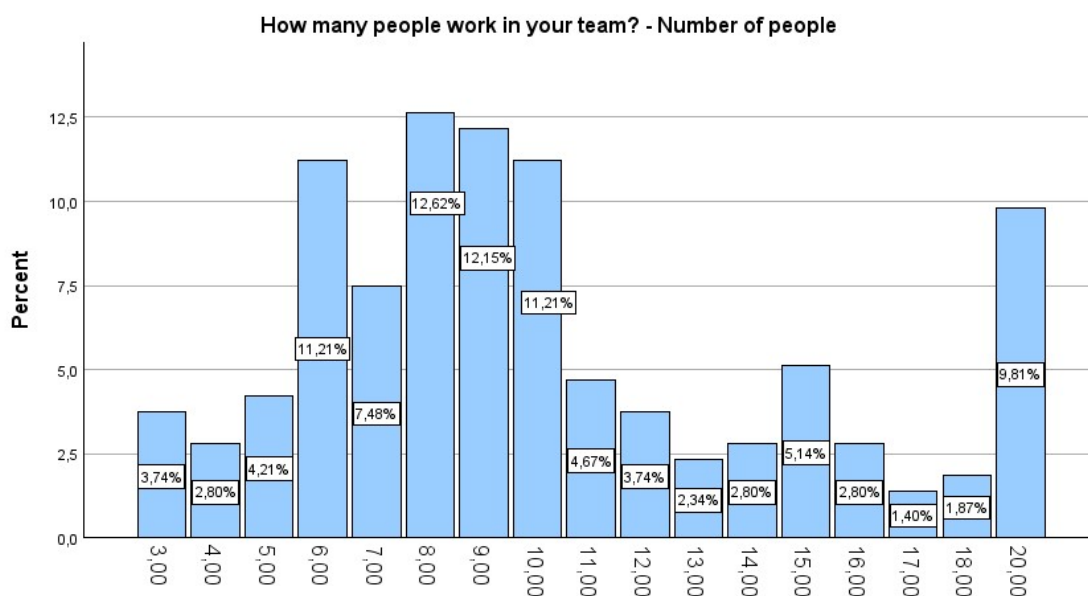


Figure 4; SPSS; Scrum Team dimension

Figure 3 illustrates the distribution of team sizes, measured by the percentage of respondents working in teams of various sizes. The data indicates that the most common team sizes range between 6 to 10 members, with the highest concentration at 8 members (12,62%), followed by 6 and 9 members (both around 12%). Smaller teams, such as those with 3 or 4 members, are less common, as are significantly larger teams, like those with 20 members (9,81%). This pattern is consistent with best practices in Agile methodologies, where teams are often recommended to have

between 5 to 9 members to maintain a significant degree of efficiency and active collaboration (Schwaber & Sutherland, 2020).

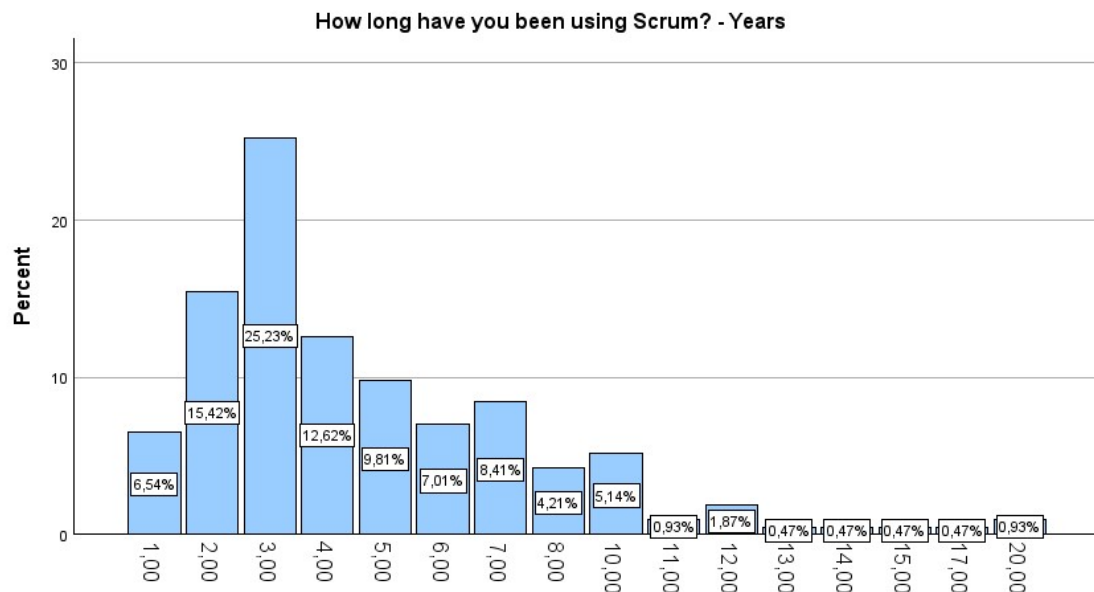


Figure 5; SPSS; Respondents' experience with Scrum

The bar chart represented in Figure 4, shows the distribution of respondents based on the number of years they have been using Scrum. Most of the participants have two to four years of experience, with the largest group (25,23%) having been using Scrum for three years. A smaller but statistically significant percentage of participants reported having used Scrum for only one year (6,54%) and for five to six years (roughly 7-9%). The data shows that most Scrum practitioners are relatively new to the process; only a small percentage of respondents has more than ten years of experience. As companies continue to realize the benefits of Scrum in managing complicated projects, this pattern reflects the growing adoption of Agile methodologies in recent years (Schwaber & Sutherland, 2020).

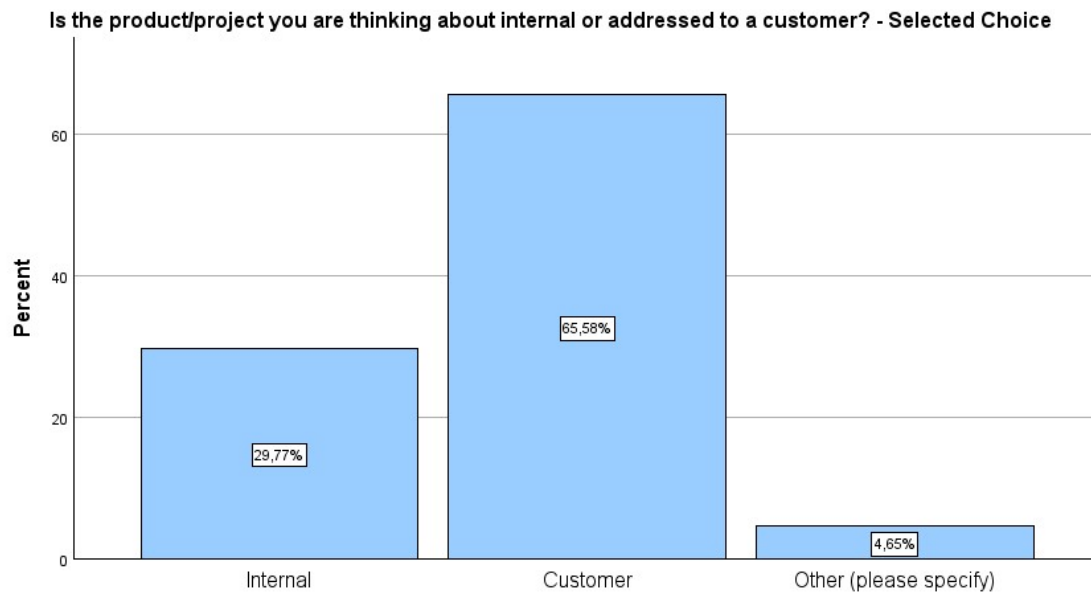


Figure 6; SPSS; Respondents' product/project purpose

Figure 5 illustrates the focus of the projects or products the respondents are involved in, categorizing them as either internal or customer-facing. Most of the participants (65,58%) stated that their projects are directed towards a client. On the other hand, 29,77% of the participants stated that they were engaged in internal projects, which are probably meant to enhance internal procedures or systems in their companies. Just 4,65% of respondents chose "Other," suggesting that their projects don't quite fit into either category. This distribution highlights the widespread application of Scrum in projects with a customer focus, demonstrating its efficacy in handling demands from clients and raising customer satisfaction levels (Schwaber & Sutherland, 2020).

After identifying the primary characteristics of the respondents, the descriptive analysis proceeded by calculating key statistical measures such as means, medians, modes, and standard deviations for all the questions regarding the Product Backlog management practices. The "means" provide an average value, the "medians" point to the middle value, the "modes" highlight the most frequently occurring responses, and the "standard deviations" measure the spread or dispersion of the data for each question. Together, these statistics offer a comprehensive view of the central tendencies and variability within the data, allowing for a deeper understanding of the

responses collected. To conduct this analysis, the "frequencies" function from the descriptive statistics section of the SPSS statistical software was employed.

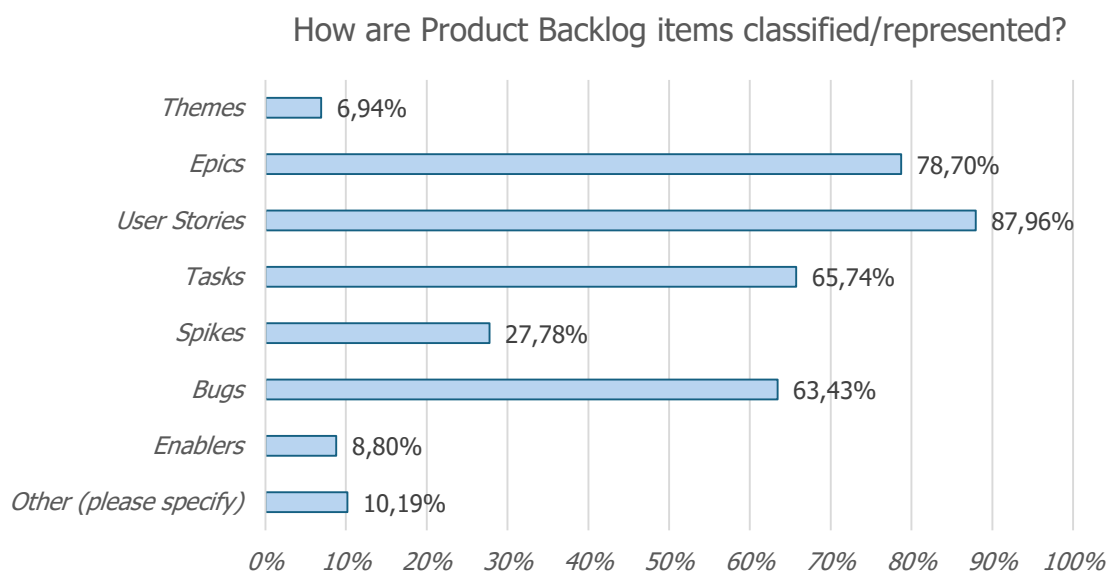


Figure 7; SPSS; "How are Product Backlog items classified/represented?" percentage responses

We shall commence by undertaking an analysis of the results derived from the questions included in the section of the questionnaire pertaining to the composition of the product backlog. As already explained in the methodology, this part of the questionnaire consisted of two questions. The first focused on the way "items" are classified within the Product Backlog, while the second focused more on the "fields" used by the respondents within the Product Backlog. As shown in Figure 6, the most recurring item classifications appear to be User Stories (87,96%) and Epics (78,70%). User Stories are the most popular way to manage requirements, and their high percentage indicates that much of the Product Backlog is focused on describing functionality from the user's perspective. Epics can be defined as larger stories and their high usage underlines the tendency of Scrum teams to focus on larger goals. Although to a lesser extent, Tasks and Bugs are also heavily used by respondents with percentages of around 66% and 64% respectively. This data shows that technical problem solving and the focus on technical tasks are quite important in Scrum. Spikes, Enablers and Themes seem to have little room in the respondents' Product Backlog, given the low usage rates found in the analysis. Finally, it is worth noting that the most

recurring elements in the "Other" section are "Features", "Initiatives" and "Subtasks". These are elements that are not recognized in the literature and show how the Scrum framework is flexible and modifiable according to the needs of its users.

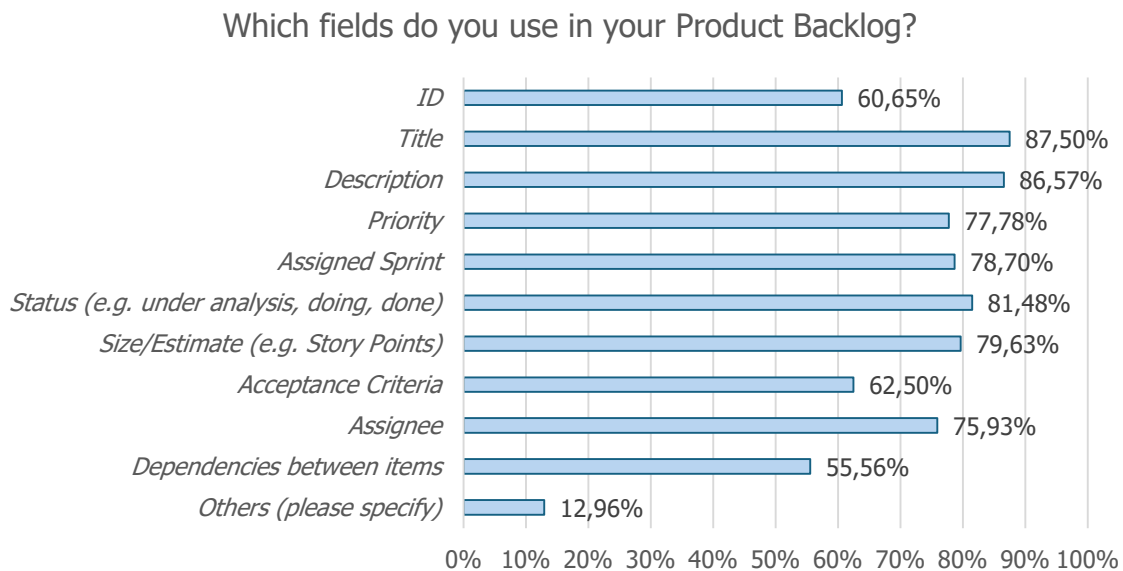


Figure 8; SPSS; "Which fields do you use in your PB?" percentage responses

Figure 7 shows the percentage of "fields" used by respondents in their Product Backlogs. In contrast with the figure analyzed above regarding item representation, there do not appear to be any fields that are utilized to a significantly lesser extent. However, fields can be divided into two macro categories based on their percentage of usage. The first category consists of the most frequently used fields, such as Title (87,50%), Description (86,57%), Status (81,48%), Size (79,63%), Assigned Sprint (78,70%), Priority (77,78%) and Assignee (75,93%). Title and Description are thus the most used fields, which is not surprising as they provide the basic information and detailed description of each Product Backlog item. Status is a crucial field for tracking the progress of each item (e.g. "under analysis", "in progress", "closed") and the percentage obtained highlights the importance of tracking and workflow management. Size and Priority show that many teams use measures such as 'story points' to estimate the workload and complexity of each task, to better manage the backlog and plan the work to be done. Assigned Sprint and Assignee are additional widely used fields, reflecting how elements are assigned to specific sprints and team members, which is

critical for iterative planning and delivery in the Scrum context. Acceptance Criteria, ID, and Dependencies represent the second macro group of fields used. In the specific case of ID, which is intended to uniquely identify Product Backlog items, it is used less than it might be expected. This might be due to the fact that the previously mentioned Title and Description are more intuitive elements to use and are therefore preferred to ID. Finally, in the response "Others" are grouped numbers of different elements such as "Labels", "Subtasks" or "Notes" together emphasizing once again the flexibility and adaptability of the Product Backlog structure according to the needs of those who use it.

4.1.2 Product Backlog Management descriptive analysis

In order to identify reliable concentrations of responses in the questionnaire questions focusing on Product Backlog management practices, a distinction was made based on the number answers that could be given. In particular, there were 3 questions that were answered using a 5-point Likert scale, while the remaining 43 questions were answered using a 7-point scale, both of which are described in Table 2 below. For the analysis of the 5-point Likert scale questions, graphs are used to facilitate the interpretation of the answers collected. For the analysis of the questions with a 7-point scale, a table was created provided with four fundamental measures: mean, median, mode and standard deviation.

| 7-Points Likert Scale | | 5-Points Likert Scale | |
|-----------------------------------|---|------------------------------|---|
| <i>Strongly Disagree</i> | 1 | <i>None of them</i> | 1 |
| <i>Disagree</i> | 2 | <i>Some of them</i> | 2 |
| <i>Somewhat Disagree</i> | 3 | <i>About half of them</i> | 3 |
| <i>Neither Agree nor Disagree</i> | 4 | <i>Most of them</i> | 4 |
| <i>Somewhat Agree</i> | 5 | <i>All of them</i> | 5 |
| <i>Agree</i> | 6 | | |
| <i>Strongly Agree</i> | 7 | | |

Table 2; Likert Scales

As an initial step, it was of particular importance to analyze the value of the standard deviation which, if less than 1.5, permitted the determination of the presence of a significant consensus on a given answer.

Descriptive Analysis: Stories

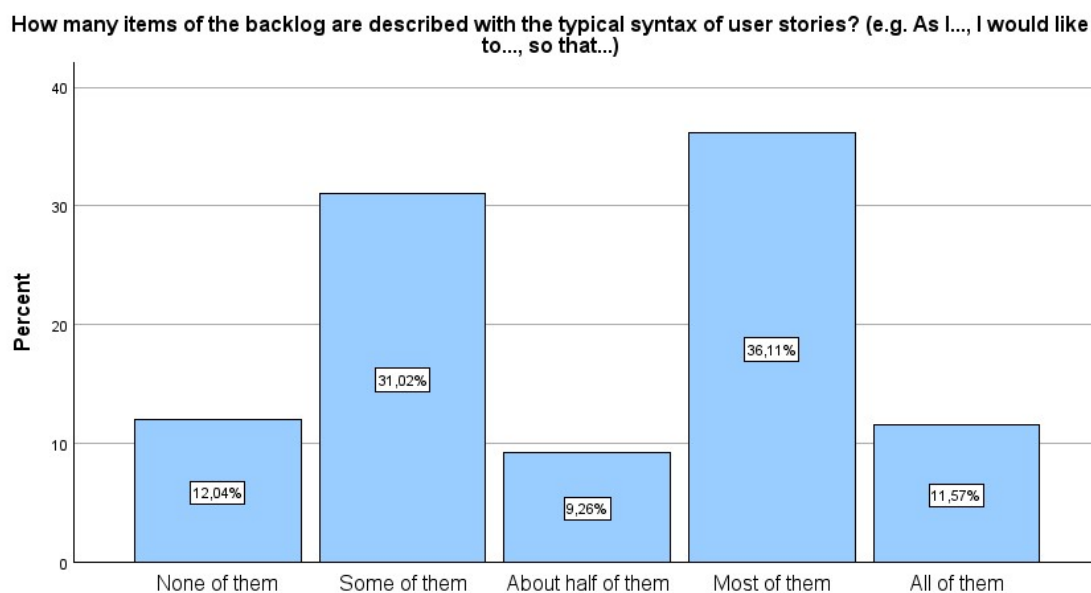


Figure 9; SPSS; User Stories' syntax usage

The histogram above shows the distribution of responses to the first question in the Stories section of the questionnaire. As can be seen, it is not possible to identify any concentration in the answers, as they are divided in mostly constant percentages between contrasting answers such as "None of them" and "All of them" (12%) and "Some of them" and "Most of them" (31-36%). It thus appears that the use of the typical syntax of user stories is evenly distributed across all the sample.

| Questions about Stories | Mean | Median | Mode | Std. Deviation |
|---|-------------|---------------|-------------|-----------------------|
| Stories are created by the product owner alone. | 3,00 | 2,00 | 2 | 1,748 |
| Stories are created collaboratively at the team level. | 5,42 | 6,00 | 6 | 1,417 |
| Stories are created by the customer. | 2,78 | 2,00 | 1 | 1,736 |
| Stories are created by stakeholders external to the Scrum team. | 2,64 | 2,00 | 1 | 1,722 |
| Stories have a size that permit them to be implanted in a single sprint. | 5,52 | 6,00 | 6 | 1,475 |
| Stories are broken down into tasks when there is enough information to do it. | 5,17 | 6,00 | 6 | 1,464 |
| Stories are broken down into tasks during Sprint Planning. | 4,63 | 5,00 | 6 | 1,616 |
| We do not break down Stories into tasks. | 3,02 | 2,00 | 2 | 1,825 |

Table 3; SPSS; Analysis of questions about Stories

Table 3 shows the mean, median, mode and standard deviation for each 7-point Likert scale question in the Stories section of the questionnaire. As explained above, the analysis of the questions where the standard deviation of the answers was lower than 1,5 is particularly relevant because it indicates a significant homogeneity of the answers collected and therefore makes it possible to highlight the existence of a concentration in the answers. The questions that met this requirement were the followings:

- “Stories are created collaboratively at the team level”
- “Stories have a size that permit them to be implanted in a single sprint”
- “Stories are broken down into tasks when there is enough information to do it”

Through the analysis of the mean, median and mode values of these three sentences and their answers some conclusions can be drawn. In fact, the data shows that there is a general agreement between respondents to create stories through collaboration at team level, that stories are created in such a way that they have a suitable size so that they can be implemented in a single sprint and that they are reduced to “tasks” if the information possessed by the team allows it.

Descriptive Analysis: Estimates

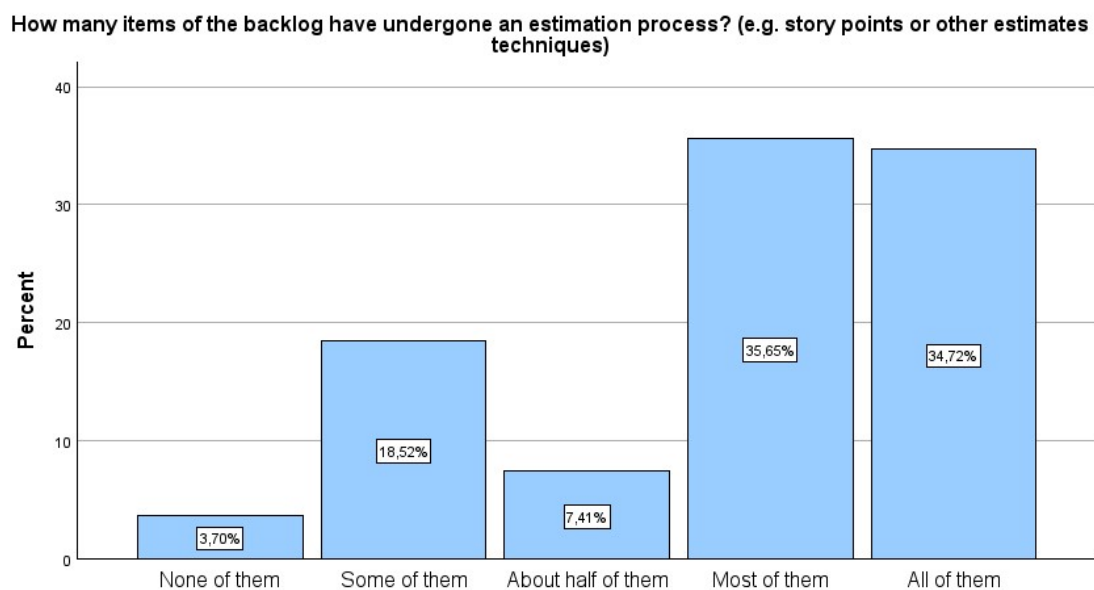


Figure 10; SPSS; Number of items estimated

The graph above shows the percentage distribution of the answers to the first question of the section of the questionnaire concerning the use of Estimates. In contrast to what was said about User Stories, in this case we can identify a rather significant convergence among the responses. In particular, 35,65% of the respondents stated that they apply an estimation process to most of the items contained in their Product Backlog while 34,72% apply this process to all items. These data underline how the practice of estimating items within the Product Backlog is particularly used probably due to the fact that it enables teams to manage tasks more precisely and plan more accurately the usage of resources during the project (Fernandez-Diego et al., 2020; Breyter, 2022).

| <i>Questions about estimates</i> | <i>Mean</i> | <i>Median</i> | <i>Mode</i> | <i>Std. Deviation</i> |
|--|--------------------|----------------------|--------------------|------------------------------|
| We continuously estimate/re-estimate the items of the backlog. | 4,64 | 5,00 | 6 | 1,596 |
| We estimate/re-estimate the items of the backlog during dedicated meetings. | 5,54 | 6,00 | 6 | 1,305 |
| At the task level, we estimate in hours or days. | 3,58 | 4,00 | 6 | 2,124 |
| Estimates are done by the product owner alone. | 1,66 | 1,00 | 1 | 1,137 |
| Estimates are done by the development team alone. | 4,14 | 5,00 | 6 | 1,974 |
| Estimates are done at the scrum team level. | 5,47 | 6,00 | 6 | 1,524 |
| Estimates are done by the customer. | 1,55 | 1,00 | 1 | 1,025 |
| Estimates are imposed to the scrum team. | 1,81 | 1,00 | 1 | 1,274 |
| External experts are involved in the estimate process. | 2,94 | 2,00 | 1 | 1,864 |
| Historical data are used in the estimate process. | 5,26 | 5,00 | 5 | 1,298 |
| We estimate by leveraging relative estimates, but we convert relative estimates into absolute estimates (hours, days etc.) because of reporting needs. | 3,31 | 3,00 | 1 | 1,855 |

| | | | | |
|--|------|------|---|-------|
| It is difficult to provide accurate estimates. | 4,74 | 5,00 | 5 | 1,567 |
|--|------|------|---|-------|

Table 4; SPSS; Analysis of questions about Estimates

Table 4 shows the mean, median, fashion and standard deviation for each question on the 7-point Likert scale concerning the Estimates. In this case, the questions with a standard deviation of less than 1,5 were the following:

- "We estimate/re-estimate the items of the backlog during dedicated meetings."
- "Estimates are done by the product owner alone."
- "Estimates are done by the customer."
- "Estimates are imposed to the scrum team."
- "Historical data are used in the estimate process."

The statistical measures collected from these questions show that users of the Scrum framework prefer to apply the estimation process to Product Backlog items during specific meetings, that estimates are not made by one individual alone, and that they are not imposed to the Scrum team by external parties such as customers. Finally, the practice of using historical data as a means of making the item estimation process easier and more accurate is particularly common among respondents. These results are consistent with what Breyter (2022) states about the fact that the estimation process needs to be performed collaboratively and during specific meetings in order for it to improve the alignment among team members.

Descriptive Analysis: Prioritization

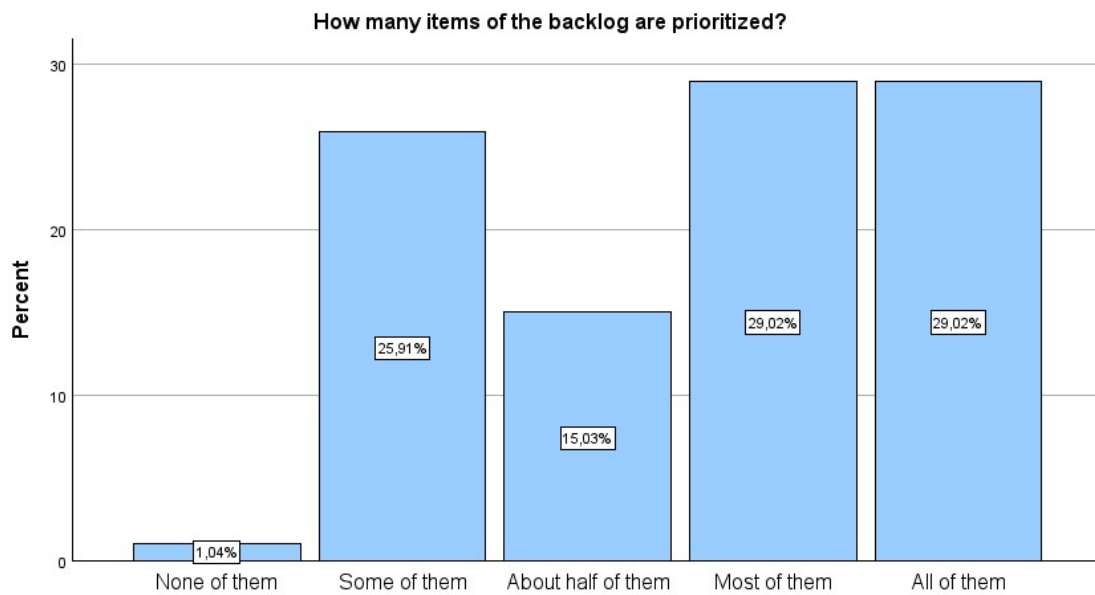


Figure 11; SPSS; Number of items prioritized

The histogram above shows the percentage distribution of the answers to the first question of the section of the questionnaire concerning the prioritization process. Although it is clear that at least some items in the Product Backlog are subject to a prioritization process, with only 1% of respondents not prioritizing any items, we cannot say with certainty what the specific concentration of responses is in this area. In fact, all 4 of the remaining responses were between 15 and 29%. This could be due to the fact that the prioritization process takes a considerable amount of time and does not always lead to desirable and satisfactory results for all members of the Scrum team. In fact, it often leads to lengthy discussions and disagreements, especially in contexts where the needs of various stakeholders change frequently (Hoy & Xu, 2023).

| Questions about Prioritization | Mean | Median | Mode | Std. Deviation |
|--|-------------|---------------|-------------|-----------------------|
| The backlog is continuously prioritized. | 5,39 | 6,00 | 6 | 1,421 |
| The backlog is prioritized during specific events. | 5,12 | 5,00 | 6 | 1,458 |
| Priorities are defined in a collaborative way. | 5,23 | 6,00 | 6 | 1,480 |
| Priorities are imposed to the Scrum team. | 3,45 | 3,00 | 2 | 1,827 |
| Priorities are defined by the customer. | 3,71 | 4,00 | 5 | 1,812 |

| | | | | |
|--|------|------|---|-------|
| Priorities are influenced by internal constraints (e.g. resource availability) | 4,88 | 5,00 | 5 | 1,355 |
| It is difficult to define priorities. | 3,73 | 4,00 | 2 | 1,621 |

Table 5; SPSS; Analysis of questions about Prioritization

Table 5 shows the mean, median, mode and standard deviation of the 7-point Likert scale questions regarding the prioritization process of the Product Backlog items. The questions that showed significant convergence in the answers, i.e. those with a small standard deviation, were as follows:

- “The Product Backlog is continuously prioritized.”
- “The backlog is prioritized during specific events.”
- “Priorities are defined in a collaborative way.”
- “Priorities are influenced by internal constraints (e.g. resource availability)”

From the analysis of the results above, some interesting conclusions can be drawn about the prioritization process. Indeed, the responses suggest that this process is continuous and occurs at specific events through collaboration between team members. Indeed, academic studies indicate that events such as sprint planning, backlog refinement, and retrospective sprints are designed to ensure that prioritization is not a one-time event, but rather an ongoing and collaborative process involving all team members (Dingsøy et al., 2012; Fernandez-Diego et al., 2020). In addition, team involvement and collaboration during these events serves to ensure that all necessary skills are considered when deciding on the order of items. Finally, the results demonstrate that the process of item prioritization is often influenced by internal constraints that, to some extent, guide the organization of sprints and the order in which items are implemented. In fact, Fernandez-Diego et al. (2020) find that team capacity, available time, and technical complexity significantly influence the decision of what to implement first.

| <i>Questions about general PB management practices</i> | <i>Mean</i> | <i>Median</i> | <i>Mode</i> | <i>Std. Deviation</i> |
|---|--------------------|----------------------|--------------------|------------------------------|
| A sprint goals is defined. | 5,59 | 6,00 | 6 | 1,501 |
| The length of Sprints is fixed. | 6,24 | 7,00 | 7 | 1,124 |
| All the information of the Product Backlog is accessible to the Scrum team. | 6,27 | 6,50 | 7 | 1,021 |
| Only the Product Owner can modify the backlog. | 3,57 | 3,00 | 2 | 1,836 |
| All the components of the Scrum team can modify the backlog. | 4,15 | 5,00 | 6 | 1,832 |
| The structure of the Product Backlog is defined by the company, the team cannot modify it. | 2,70 | 2,00 | 2 | 1,737 |
| The Product Backlog is customized in order to take into account the customer's preferences. | 4,47 | 5,00 | 5 | 1,646 |
| The Product Backlog is customized in order to take into account the customer's competencies. | 3,67 | 4,00 | 4 | 1,547 |
| The Product Backlog structure does not change according to the customer's preferences and competences. | 3,83 | 4,00 | 2 | 1,789 |
| There is a mismatch between the priorities desired by the customer and the actual implementation of our work. | 3,46 | 3,00 | 2 | 1,476 |
| Once the team is committed to specific items during a Sprint, no changes are allowed. | 3,56 | 3,00 | 2 | 1,758 |
| We leave some spare capacity to our team in order to embrace urgent requests. | 4,87 | 5,00 | 6 | 1,610 |
| It is very difficult to maintain the backlog up to date. | 3,45 | 3,00 | 2 | 1,608 |
| It is difficult to have items with the proper level of detail. | 3,98 | 4,00 | 5 | 1,672 |
| The management of the backlog is too time consuming. | 3,31 | 3,00 | 2 | 1,539 |

| | | | | |
|---|------|------|---|-------|
| I am satisfied with the way the backlog is managed. | 5,05 | 5,00 | 6 | 1,437 |
|---|------|------|---|-------|

Table 6; Analysis about General Product Backlog Management practices

Table 6 shows the mean, median, mode and standard deviation data for the 7-point Likert scale questions on common practices used to manage the Product Backlog. These common practices include topics such as the structure of sprints, the adaptation of the backlog to internal and external requirements, the accessibility of the Product Backlog, and the difficulty encountered in various aspects of its management. Among these questions, the ones that showed a significant convergence in the answers were the followings:

- "The length of Sprints is fixed."
- "All the information of the Product Backlog is accessible to the Scrum team."
- "There is a mismatch between the priorities desired by the customer and the actual implementation of our work."
- "I am satisfied with the way the backlog is managed."

From the data collected it appears that there is considerable agreement among the respondents that the length of the sprints is fixed, that all information in the Product Backlog is accessible to the Scrum Team, that there is no particular mismatch between the priorities desired by the customer and those actually implemented by the team, and that the respondents are quite satisfied with the way their Product Backlog is managed. These conclusions indicate a general lack of flexibility in the management of the Product Backlog on the part of the respondents, who prefer to keep the length of the sprints fixed, as indicated by Schwaber and Sutherland in their Scrum Guide (2020), and do not adapt the prioritization process to the customer's wishes. Finally, it appears that there is generally a high degree of transparency in the management of the Product Backlog towards the members of the entire team given that the majority of respondents stated that each of them can have complete access to the information contained in the Product Backlog.

4.2 Factor Analysis

Factor Analysis is a statistical tool used to identify underlying relationships between variables by grouping them into factors (Tabachnick & Fidell, 2013). This technique reduces the number of variables by combining those that are correlated, thus simplifying the data structure without losing significant information (Hair, Black, Babin, & Anderson, 2019). It starts with a large number of variables and searches for a way to "reduce" or summarize the data using a smaller number of components. To accomplish this, it searches for "clumps" or groups within a set of variables' intercorrelations (Pallant, 2020). There are several uses for this method such as creating and assessing tests and scales or implementing a subsequent multiple regression as it was done in this research (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

There are three main steps involved in conducting a factor analysis. First, it is necessary to assess the suitability of the data for such an analysis. The two main issues to consider regard the size of the sample and the strength of the relationship among variables. As far as the sample size is concerned, there is little agreement among researchers even if they all agree that the larger the better given the fact that the correlation coefficients between the variables are less consistent in small samples and tend to differ between samples. Hence, small data set factors don't generalize as effectively as larger sample size factors do. Tabachnick and Fidell (2001) suggest that "having at least 300 cases for factor analysis is comforting" but they also admit that if the solutions contain multiple high loading marker variables, then smaller samples (such as 150 cases) might be adequate. According to Stevens (1996), the suggested sample size has dropped over time as more studies on the subject have been carried out. In light of these considerations, the goal of 190–230 replies was established to guarantee a sufficient number of cases for carrying out significant statistical analyses. As for the second issue to be addressed, which is related to the strength of the intercorrelations among the items, Tabachnick and Fidell recommend an inspection of the

"correlation matrix" for evidence of at least some coefficients above the threshold value of 0.3. Indeed, if only few correlations above this level are found then factor analysis might not be appropriate. Moreover, SPSS provides researchers with two statistical indicators that help assess the factorability of the data: the Bartlett's test of sphericity (Bartlett, 1954), and the Kaiser-Meyer-Olkin measure of sampling adequacy (Kaiser, 1970, 1974). For the factor analysis to be considered appropriate, the Bartlett's test should be significant (i.e. $p\text{-value} < 0.05$) while the KMO index, which ranges between 0 and 1, should assume a value greater than 0.6. As indicated in Table 7 below, in this research the Bartlett's Test was significant, underlining the appropriateness of the variables for factor analysis (Bartlett, 1954), and the Kaiser-Meyer-Olkin measure was equal to 0,656 which confirms the data suitability for the implementation of a Factor Analysis.

| | |
|---|-----------------------------|
| <i>Kaiser-Meyer-Olkin Measure of Sampling</i> | 0,656 |
| <i>Bartlett's Test of Sphericity</i> | Approx. Chi-Square 2491,986 |
| | df 1035 |
| | Sig. <,001 |

Table 7; SPSS; KMO and Bartlett's Test of Sphericity

Factor extraction is the second critical step in performing a factor analysis. Finding the fewest components that can accurately capture the interactions between a group of variables is the goal of this method. The necessity for a straightforward solution with the fewest possible elements and the requirement to account for the greatest percentage of variance in the original data set must be balanced in this phase. The Kaiser criterion and the Scree test are the two main methods provided by the statistical program SPSS for figuring out how many components need to be extracted. The Kaiser criterion states that only components that explain a significant portion of the total variance and have eigenvalues greater than one should be kept (Kaiser, 1960). Plotting the eigenvalues and locating the point at which the plot levels off, a sign that the following elements add very little to the variance explanation, is the Scree test (Cattell, 1966). In this research the methodology employed for factor extraction utilized both of the aforementioned tools. This combined approach resulted in a final number of 16

factors extracted, which allowed maximum simplification of the model while still representing a substantial portion of the variance in the original variables.

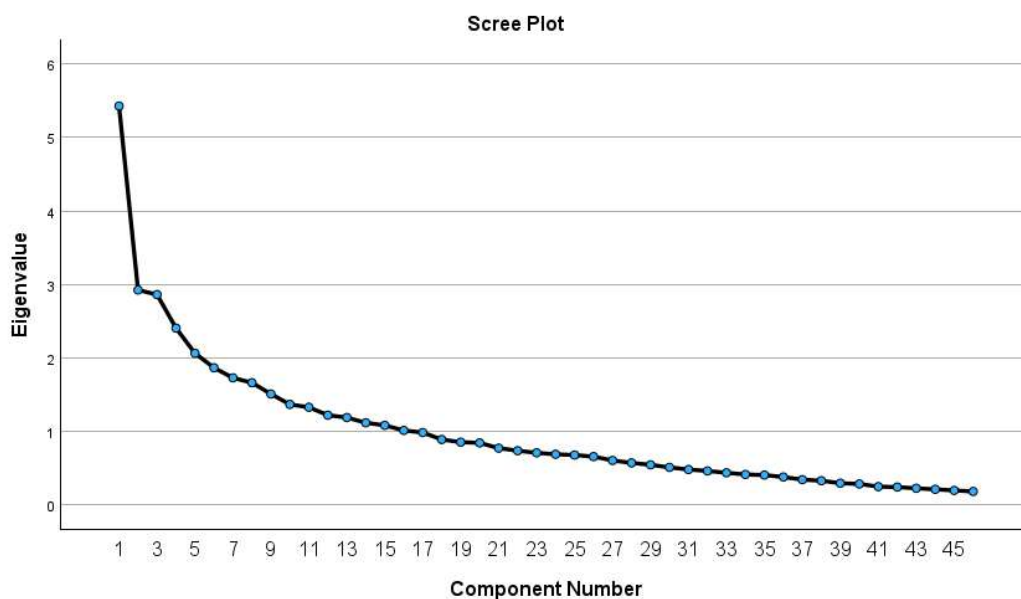


Figure 12; SPSS; Scree Plot

The final fundamental step in factor analysis is the analysis and interpretation of the extracted factors. Indeed, it is extremely important to understand which variables of the original data set are represented by each extracted factor. For this purpose, SPSS provides as output a table called the "Rotated Component Matrix". A key aspect of this table concerns "factorial loadings", i.e. how much each variable in the dataset is related to the extracted factor. The closer a factorial load is to a value of 1, the greater the correlation of the original variable with the extracted factor that contains it. If the factorial load has a positive sign, the correlation with the factor is positive, otherwise it is negative. All 16 extracted factors, the variables they represent, and the corresponding loading factors are shown in Table 8 below.

| Factor Name | Variable Name | Loading Factor |
|-----------------------------|---|----------------|
| FAC1_DIFF | § It is difficult to have items with the proper level of detail | 0.805 |
| | § It is very difficult to maintain the backlog up to date | 0.779 |
| | § The management of the backlog is too time consuming | 0.724 |
| | § It is difficult to define priorities | 0.557 |
| | § It is difficult to provide accurate estimates | 0.494 |
| FAC2_ESTIM_ORIGIN | § Estimates are done by the customer | 0.663 |
| | § Estimates are imposed to the scrum team | 0.627 |
| FAC3_PB_MODIFY | § All the components of the Scrum team can modify the backlog | 0.842 |
| | § Only the Product Owner can modify the backlog | 0.833 |
| FAC4_PB_CUSTOMIZE | § The Product backlog is customized in order to take into account the customer's preferences | 0.847 |
| | § The Product backlog structure does not change according to the customer's preferences and competences | (0.695) |
| | § The Product backlog is customized in order to take into account the customer's competencies | 0.643 |
| FAC5_FLEX | § We leave some spare capacity to our team in order to embrace urgent requests | 0.643 |
| | § How many items of the backlog are described with the typical syntax of user stories? (e.g. As I..., I would like to..., so that...) | 0.640 |
| FAC6_STORIES_EXT_CREATE | § Stories are created by the customer | 0.814 |
| | § Stories are created by stakeholders external to the Scrum team | 0.812 |
| FAC7_PRIO_CREATE_FLEX | § Priorities are imposed to the Scrum team | 0.743 |
| | § Priorities are defined by the customer | 0.524 |
| FAC8_STORIES_DIVISION | § We do not break down Stories into tasks | (0.762) |
| | § Stories are broken down into tasks during Sprint Planning | 0.671 |
| | § Stories are broken down into tasks when there is enough information to do it | 0.638 |
| FAC9_SPGOAL_PBACCESS | § A sprint goals is defined | 0.538 |
| | § All the information of the product backlog is accessible to the Scrum team | 0.498 |
| | § The length of Sprints is fixed | 0.495 |
| FAC10_PRIO_ORIGIN | § The backlog is prioritized during specific events (e.g. backlog refinement meetings) | 0.755 |
| | § Priorities are defined in a collaborative way | 0.620 |
| FAC11_NUM_ITEM_ESTIM | § How many items of the backlog have undergone an estimation process? (e.g. story points or other estimates techniques) | 0.741 |
| | § How many items of the backlog are prioritized? | 0.684 |
| FAC12_PRIO_ALIGN | § There is a mismatch between the priorities desired by the customer and the actual implementation of our work | 0.655 |
| FAC13_STORIES_CREATE_COLLAB | § Stories are created by the product owner alone | (0.730) |
| | § Stories are created collaboratively at the team level | 0.456 |
| FAC14_ESTIM_FREQ | § We continuously estimate/re-estimate the items of the backlog | 0.788 |
| | § We estimate/re-estimate the items of the backlog during dedicated meetings | 0.643 |
| FAC15_ESTIM_CREATE_DEVTEAM | § Estimates are done by the development team alone | 0.802 |
| FAC16_PRIO_INFLUENCE_CONSTR | § Priorities are influenced by internal constraints (e.g. resource availability) | 0.783 |

Table 8; SPSS; Extracted Factors

The first extracted factor, 'FAC1_DIFF', shows a significant correlation with the questionnaire variables related to the respondents' difficulty in managing the Product Backlog. FAC2_ESTIM_ORIGIN', the second factor, describes the variables related to the external parties involved in the estimation process of the product backlog items. The third factor extracted, called 'FAC3_PB_MODIFY', has a strong correlation with the

variables describing the level of accessibility of the Product Backlog and the possibility of modifying it, while the fourth factor, 'FAC4_PB_CUSTOMIZE', refers to the tendency to customize the structure of the Product Backlog according to the customer's requirements. The variables related to the flexibility implemented in the management of the Product Backlog are represented by the fifth factor extracted, named 'FAC5_FLEX'. The sixth factor, 'FAC6_STORIES_EXT_CREATE', summarizes the variables of the questionnaire that refer to the creation of stories by parties outside the Scrum team. The variables related to the flexibility in defining priorities are represented by the seventh extracted factor, called 'FAC7_PRIO_CREATE_FLEX'. The eighth factor, 'FAC8_STORIES_DIVISION', refers to all the variables describing the respondents' tendency to divide stories into tasks. The variables related to the definition of 'Sprint Goals' and the accessibility of the Product Backlog by the entire Scrum Team are described by the ninth extracted factor, called 'FAC9_SPGOAL_PBACCESS'. The tenth extracted factor, 'FAC10_PRIO_ORIGIN', is significantly correlated with the variables concerning the moments when the prioritization process is implemented. The eleventh extracted factor, 'FAC11_NUM_ITEM_ESTIM', describes the number of Product Backlog Items that are subject to an estimation process. The twelfth, fifteenth and sixteenth factors are related to a single variable of the questionnaire and describe, respectively, the alignment of the priorities set by the Scrum team and those preferred by the customer, the definition of estimates by the developers alone and, finally, the influence of internal constraints on the definition of priorities. The thirteenth factor extracted, 'FAC13_STORIES_CREATE_COLLAB', represents the variables related to the degree of collaboration used in the creation of stories. Finally, the fourteenth factor, called 'FAC14_ESTIM_FREQ', represents the frequency used by the respondents in defining the estimates of the Product Backlog items.

4.3 Multiple Regression

Multiple regression is a statistical technique used to analyze the relationship between a dependent variable and two or more independent variables (Montgomery, Peck, & Vining, 2012). More specifically, the aim of multiple regression is to predict the value of a dependent variable based on the values of the independent variables, thereby estimating the contribution of each predictor in the context of all the other variables included in the model (Kutner, Nachtsheim, & Neter, 2004). In this study, multiple regression is of particular importance as it is the tool by which the third research question, "What are the key factors that most significantly influence satisfaction levels in Product Backlog management", is to be answered. The first step consisted in determining the dependent and independent variables. The dependent variable, i.e. the one to be predicted, corresponds to the statement in the questionnaire "I am satisfied with the way the Product Backlog is managed". On the other hand, the independent variables correspond to the 16 factors extracted by means of factor analysis, which have already been examined above. Using the factors extracted from a factor analysis as independent variables in a multiple regression analysis, instead of the original variables, has several advantages. First, there is a clear reduction in multicollinearity between the independent variables (Hair, Black, Babin, & Anderson, 2010). As explained in the methodology section, multicollinearity occurs when the independent variables are highly correlated with each other, which can distort the regression coefficients, worsen the quality of the model, and make it difficult to determine the individual effect of each predictor (Tabachnick & Fidell, 2013). Since Factor Analysis creates factors that are orthogonal to each other, i.e. uncorrelated, the use of these factors as predictors allows the problem of multicollinearity to be mitigated, thus improving the interpretation of the model (Field, 2018). Another advantage lies in the reduction of model complexity and information synthesis (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Indeed, instead of including a large number of variables that may be redundant, factors extracted by Factor Analysis

represent linear combinations of the original variables that summarize the main information (Jolliffe, 2002). Finally, using extracted factors increases the parsimony of the model, making it simpler and more compact (Kline, 2015). A model with a high degree of parsimony is preferable as it considerably simplifies its interpretation and reduces the risk of including unnecessary variables that could compromise the overall significance of the model (Hair et al., 2010).

Before analyzing the results obtained from the regression model and interpreting its significance, it was of fundamental importance to judge the reliability and goodness of the model through a series of statistical data provided by SPSS. These comprised the observation of the R^2 and standard error of estimate values, the ANOVA table in order to ascertain the significance of the model, and finally, the scatterplot and normal P-P plot graphs that provide useful information on the homoscedasticity and normality of the model's residuals, respectively. Table 9 shows a summary of the main statistics for assessing the goodness of fit of a multiple regression model. "R" is the multiple correlation coefficient and measures the strength of the linear relationship between the observed values of the dependent variable and the values predicted by the model. Its value can vary from 0 to 1, and the closer it is to the upper threshold, the better the model's ability to explain the variability of the dependent variable (Holmes & Rinaman, 2015). The coefficient of determination, or " R^2 ", shows how much of the variance of the dependent variable can be accounted for by the independent variables in the model. Once more, the regression model's predictive power increases as its value approaches one (Springer, 2020). A modified version of R^2 that accounts for the number of independent variables in the model and the total number of observations is called "Adjusted R^2 ". It is primarily used to compare various models and penalize the inclusion of an excessive number of independent variables in the regression model. Contrary to R^2 , the Adjusted R^2 does not inevitably rise as more independent variables are added to the model unless those additions significantly enhance its predictive power (Lee et al., 2016; Karch, 2021). Finally, the standard error of estimate depicted

in the figure measures the precision and accuracy of the predictions by representing the standard deviation of the model's residuals. In this instance, the lower the value the better the accuracy of the model (Skiera et al., 2022). In this specific case, the values observed in the table below demonstrate the excellent predictive ability of the regression model. In fact, on the one hand, the R^2 value (0,625) shows the ability of the model to explain a significant portion of variance of the dependent variable and, on the other hand, the standard error of the estimate (0,923) proves the fact that the estimates made by the model have a low degree of error.

| <i>R</i> | <i>R Square</i> | <i>Adjusted R Square</i> | <i>Std. Error of the Estimate</i> |
|----------|-----------------|--------------------------|-----------------------------------|
| 0,79 | 0,625 | 0,587 | 0,923 |

Table 9; SPSS; Summary of the Regression Model

Table 10 represents the ANOVA table obtained on SPSS and presents some important data on the regression model of this study. The ANOVA table consists of 4 main elements and its usefulness lies in determining whether there is a significant relationship between the independent variables and the dependent variable in the overall model. The table assesses, through the value of the Sum of Squares of the regression and the residuals, whether the variance explained by the model is significantly greater than the unexplained variance (Cleophas & Zwinderman, 2021). The value of the Mean Square of the regression represents the average portion of variance explained by each independent variable while the Mean Square of the residuals explains the average portion of variance unexplained. The value of F relates the variance explained and unexplained by the regression model and, the higher its value, the better the model. Finally, the p-value is associated with the value of F and, in order for the variance explained by the model to be significantly greater than the unexplained variance, it must assume a value less than 0.05 (Pardo, 2023). In this case, the ANOVA table demonstrates the statistical significance of the overall

regression model through an optimal p-value and an explained variance of the dependent variable considerably greater than the residual one.

| ANOVA | | | | |
|------------|----------------|-------------|--------|-------|
| | Sum of Squares | Mean Square | F | Sig. |
| Regression | 224,388 | 14,024 | 16,446 | <,001 |
| Residual | 134,733 | 0,853 | | |
| Total | 359,121 | | | |

Table 10; SPSS; ANOVA

Following the analysis of the summary table of the regression model and the ANOVA table, it is important to observe the Normal P-P Plot below. In fact, this type of plot is used to assess the normality assumption of the residuals in a regression model. If the points on the graph roughly follow a diagonal line, it can be concluded that the residuals are normally distributed and therefore meet the assumption of normality (Das & Imon, 2015). The P-P Plot showed in Figure 12 suggests that the normality assumption of the residuals is fulfilled, as the points follow the diagonal line uniformly.

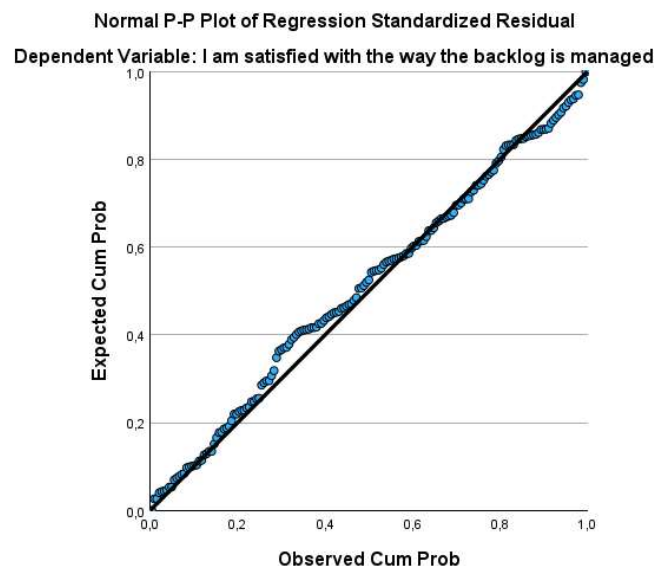


Figure 13; SPSS; Normal P-P Plot

Figure 13 shows the Scatterplot of "standardized residuals" against the "standardized predicted value" used to verify that the variance of the residuals was constant. The variance of residuals was not entirely consistent across all predicted values, indicating a moderate violation of homoscedasticity. However, this was deemed an acceptable deviation.

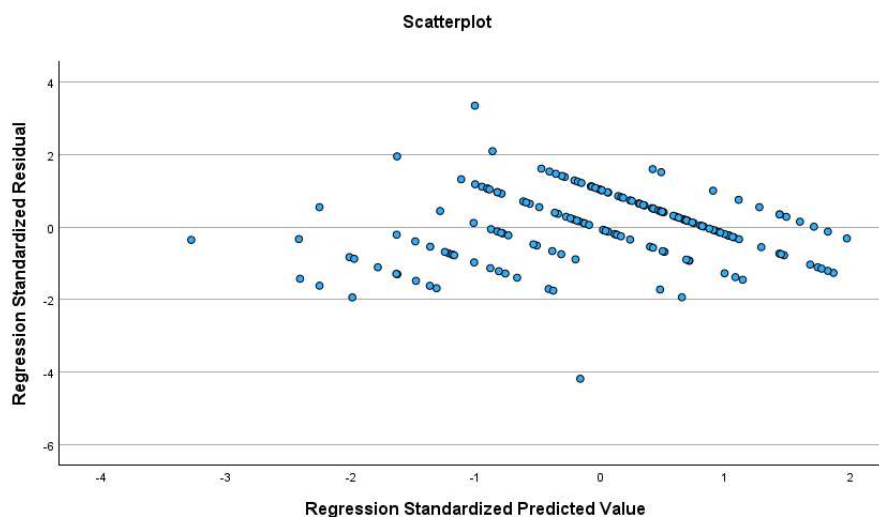


Figure 14; SPSS; Scatterplot

Having validated the statistical significance and predictive capacity of the regression model in estimating the level of satisfaction with the management of the product backlog, it is now appropriate to undertake a detailed examination of the results obtained. To this end, it is particularly important to examine Table 11 as it shows the results of the regression coefficients for the independent variables included in the model, together with the significance and multicollinearity statistics. The interpretation of each element is essential to understand the importance and influence of each independent variable in the regression model. More specifically, the table shows the names of the independent variables (which are none other than the factors previously extracted from the factor analysis), data on the unstandardized and standardized coefficients, and data on collinearity. The unstandardized coefficients indicate the absolute impact of each independent variable on the dependent variable. Specifically, the unstandardized beta indicates the expected change in the dependent variable for each unit increase in the independent variable, holding the other variables in the model constant. The sign of these data provides useful information on the correlation between the independent and dependent variables. The correlation may be positive or negative, with positive indicating a direct relationship between the variables and negative indicating an inverse relationship. The standardized coefficients (Beta) indicate the effect of the independent variables on the dependent variable in terms of

standard deviation units, allowing a direct comparison between predictors (Kraha et al., 2012). The “t-statistic” is used to test the null hypothesis that the regression coefficient is equal to zero. In addition, the “Sig” column indicates the p-value associated with the t-test, and if its value is less than 0.05, then it means that the coefficient under consideration is statistically significant, and thus the effect of the independent variable on the dependent variable is not random (Das, 2019). Finally, the two columns “Tolerance” and “VIF” are useful for assessing the presence of multicollinearity between the independent variables. In particular, to ensure that multicollinearity problems are not encountered, the value corresponding to Tolerance must be as close as possible to 1 while VIF must be less than 10 (Das, 2019). In this research, the Tolerance and VIF values for all independent variables included in the model totally rule out the presence of multicollinearity problems.

| Independent Variables | Unstandardized Coefficients | | Standardized Coefficients | | | Collinearity Statistics | |
|-----------------------------|-----------------------------|------------|---------------------------|--------|-------|-------------------------|-------|
| | Unstandardized Beta | Std. Error | Beta | t | Sig. | Tolerance | VIF |
| (Constant) | 5,029 | 0,071 | | 71,249 | <,001 | | |
| FAC1_DIFF | -0,635 | 0,069 | -0,448 | -9,154 | <,001 | 0,99 | 1,01 |
| FAC2_ESTIM_ORIGIN | 0,119 | 0,077 | 0,076 | 1,557 | 0,122 | 0,985 | 1,016 |
| FAC3_PB_MODIFY | 0,102 | 0,072 | 0,069 | 1,404 | 0,162 | 0,986 | 1,014 |
| FAC4_PB_CUSTOMIZE | -0,041 | 0,072 | -0,028 | -0,577 | 0,565 | 0,977 | 1,023 |
| FAC5_FLEX | 0,757 | 0,071 | 0,521 | 10,616 | <,001 | 0,986 | 1,014 |
| FAC6_STORIES_EXT_CREATE | 0,097 | 0,072 | 0,066 | 1,347 | 0,18 | 0,984 | 1,017 |
| FAC7_PRIO_CREATE_FLEX | -0,015 | 0,068 | -0,011 | -0,217 | 0,828 | 0,993 | 1,007 |
| FAC8_STORIES_DIVISION | -0,224 | 0,069 | -0,158 | -3,227 | 0,002 | 0,991 | 1,009 |
| FAC9_SPGOAL_PBACCESS | 0,245 | 0,077 | 0,156 | 3,166 | 0,002 | 0,974 | 1,027 |
| FAC10_PRIO_ORIGIN | 0,14 | 0,071 | 0,097 | 1,981 | 0,049 | 0,986 | 1,014 |
| FAC11_NUM_ITEM_ESTIM | -0,011 | 0,076 | -0,007 | -0,151 | 0,88 | 0,971 | 1,03 |
| FAC12_PRIO_ALIGN | -0,241 | 0,072 | -0,164 | -3,352 | 0,001 | 0,988 | 1,012 |
| FAC13_STORIES_CREATE_COLLAB | 0,306 | 0,071 | 0,21 | 4,296 | <,001 | 0,994 | 1,006 |
| FAC14_ESTIM_FREQ | -0,127 | 0,069 | -0,09 | -1,852 | 0,066 | 0,997 | 1,003 |
| FAC15_ESTIM_CREATE_DEVTEAM | -0,182 | 0,071 | -0,124 | -2,541 | 0,012 | 0,991 | 1,009 |
| FAC16_PRIO_INFLUENCE_CONSTR | 0,016 | 0,07 | 0,012 | 0,237 | 0,813 | 0,986 | 1,014 |

Table 11; SPSS; Multiple Regression Coefficients and Collinearity Statistics for the Independent Variables

Once we have established the predictive power of the regression model using the above tools, we move on to the stage of analyzing the results. Firstly, it is important to highlight those independent variables among the 16 used that are statistically significant, i.e. the ones that have the ability to positively or negatively influence the dependent variable. To do this, it is useful to look at Table 11 and, in particular, the unstandardized beta and the p-value for each independent variable. In fact, a variable

with a p-value of less than 0.05 is defined as statistically significant and therefore able to predict, at least partially, the value of the dependent variable. The value of the unstandardized beta indicates the strength and direction of the correlation between the independent and dependent variables (Amrhein, Greenland, & McShane, 2019). In the table above, the significant independent variables are underlined in blue with their respective Beta and p-values. In light of the aforementioned independent variables, which correspond to the factors extracted from the factor analysis, it is possible to reach the results and conclusions of the regression model. For this purpose, it is also advisable to consult Table 8 in the Factor Analysis section in order to gain clarity regarding the variables included in the extracted factors.

FAC1_DIFF

The first factor, concerning the difficulty in managing the Product Backlog, has an extremely low p-value ($<0,001$), which makes it statistically significant. As can be seen in Table 11, the Unstandardized Beta takes on a rather high negative value (-0,635), indicating a significant negative correlation between the independent variable and the level of satisfaction in managing the Product Backlog. This means that, logically, the greater the difficulty that the respondents have in managing the Product Backlog, the lower the level of satisfaction observed. Decomposing the FAC1_DIFF into its original component variables, we find that the concept of "difficulty" is articulated in different contexts, such as the difficulty in obtaining items with the right level of detail, in keeping the Product Backlog up to date, in defining priorities and estimates.

FAC5_FLEX

The fifth factor extracted includes variables related to the degree of flexibility implemented in the management of the Product Backlog, with a particular focus on the respondents' tendency to leave spare capacity to deal with emergency situations and the number of items described using the typical User Stories syntax. It shows a high statistical significance (p-value $<0,001$) and a very high correlation with the

dependent variable (0,757), indicating that the greater the number of items described with the syntax of User Stories and the capacity left for emergency situations, the greater the degree of satisfaction perceived by respondents in the management of the Product Backlog. This is also stated by Pecchia et al. (2016) when they argue that when stories are structured according to specific and structured methods, there is a significant improvement in the understanding of project progress requirements by the entire team and, as a result, greater satisfaction in managing the product backlog. In addition, effective backlog management that incorporates contingency margins reduces the risk of overload and increases team satisfaction and productivity by allowing better management of priorities and dealing with unforeseen contingencies without compromising the main project goals (Charikar et al., 2022).

FAC8_STORIES_DIVISION

This factor is particularly related to questions about respondents' tendency to divide stories into tasks, either generally or in specific situations such as sprint planning, or simply when there is enough information to do so. The eighth factor in Table 11 shows high significance (p-value equal to 0,002) and a negative correlation with the dependent variable (-0,224). Thus, the results seem to indicate that the greater the tendency of the respondents to divide the stories into tasks, the lower the level of satisfaction with the management of the Product Backlog. This could be due to the fact that dividing stories into tasks increases the overall complexity and density of the Product Backlog, which also negatively affects the simplicity of its management. This can make it difficult to visualize and manage long-term goals, reducing clarity and increasing the likelihood of errors or forgetfulness (Müter et al., 2019). In addition, the excessive segmentation of stories into smaller tasks makes prioritization of items much more difficult and therefore risks increasing the time required for prioritization excessively (Edison et al., 2021). Finally, an overly detailed backlog can limit the team's ability to adapt quickly to change. In fact, if each story is broken down into micro-

tasks, it may take longer and be more laborious to implement the necessary changes (Edison et al., 2021).

FAC9_SPGOAL_PBACCESS

The ninth independent variable used in the regression contains information about the tendency to define Sprint Goals, to set Sprint lengths and to make the information contained in the Product Backlog available to all Scrum Team members. It is, as the p-value of 0,002 shows, statistically significant and positively correlated with the level of satisfaction with the management of the Product Backlog (0,245). This shows how a higher level of transparency and accessibility of the Product Backlog and a tendency to set sprint goals and their length increase satisfaction with the management of the Product Backlog. These findings are consistent with some of the core principles of the Scrum framework, such as transparency and shared understanding of goals. In fact, the shared understanding of what needs to be done that is created among team members when the Product Backlog is visible and accessible to all reduces uncertainty, misunderstandings and misaligned expectations, improving collaboration and satisfaction in managing the Backlog (Kautz et al., 2014). Finally, clarity about sprint goals helps the team know exactly what to expect and where to focus, facilitating overall management and strategic alignment (Fowler, 2019).

FAC10_PRIO_ORIGIN

This independent variable is correlated with the tendency to perform the prioritization process of the items in specific moments and through a joint collaboration of all Scrum team members. The variable is statistically significant and positively correlated, as shown by the p-value of 0,049, although the correlation is not particularly strong in this case (unstandardized beta of 0,14). This suggests that satisfaction with Product Backlog management increases moderately as the tendency to prioritize items collaboratively and at specific times such as refinement meetings increases. Again, the increase in satisfaction may be due to the fact that collaborative prioritization increases

transparency and drastically reduces the likelihood of disagreement on priorities within the team, thus improving the alignment of members on project goals (Abrahams, 2022).

FAC12_PRIO_ALIGN

The twelfth factor used in the regression corresponds to the following statement from the questionnaire: "There is a mismatch between the priorities desired by the client and the actual implementation of our work". It has a p-value of 0,001, making it statistically significant, and a negative correlation of 0,241. This leads to the conclusion that the greater the mismatch between the priorities of the items requested by the customer and those actually implemented by the team, the lower the level of satisfaction with the management of the Product Backlog. In fact, a large mismatch not only leads to lower customer satisfaction but is also a symptom of poor communication within the team, misinterpretation of requirements and poor expectation management (Jaakkola et al., 2015). These are all factors that we have already seen to be in marked contrast to an increase in satisfaction with the management of the Product Backlog.

FAC13_STORIES_CREATE_COLLAB

This independent variable is significantly correlated with the extent of collaboration between Scrum Team members in the process of story creation. In particular, there is a significant negative correlation between this factor and the variable "Stories are created by the Product Owner alone," and a positive correlation with the original variable "Stories are created collaboratively at the team level." In turn, the extracted factor (i.e. independent variable) is statistically significant (p-value less than 0,001) and positively correlated with the dependent variable of the regression. This shows once again that a higher degree of cooperation in the implementation of the practices of Product Backlog Management significantly increases the observed level of satisfaction. In general, the degree of transparency and sense of ownership exhibited

by a team with respect to all members is positively correlated with the level of satisfaction experienced by the team.

FAC15_ESTIM_CREATE_DEVTEAM

The fifteenth and final significant factor within the regression model of this study is closely associated with the inclination to exclusively delegate the management of the estimation process of items to the development team. This approach effectively excludes the Scrum Master and Product Owner from the process. This independent variable is statistically significant (p-value of 0,012) and negatively correlated with the level of satisfaction with the management of the Product Backlog (-0,182). This provides further evidence that the management of an aspect of the Product Backlog by a single member of the Scrum Team is an ineffective approach, as it results in higher levels of dissatisfaction. In addition to excluding some members from a fundamental step of the project, it increases the likelihood of misunderstandings and deviations from the objectives, which ultimately leads to a deterioration in the quality of the work done (Edison et al., 2021).

In summary, the regression model enabled the identification of eight factors that influence the degree of satisfaction experienced by Scrum users in the management of the product backlog. These eight factors can be summarized in six main areas: the perceived difficulty in managing the product backlog, the degree of flexibility implemented, the division of stories into smaller tasks, the degree of transparency and accessibility used, the mismatch between the priorities implemented by the Scrum team and those desired by the customer, and finally, the degree of collaboration used in the management of the product backlog. In subsequent chapters, this study will examine the managerial implications of these findings.

4.4 Cluster Analysis

A k-means cluster analysis was conducted using the IBM SPSS Statistics program to investigate the utilization of distinct methodologies in product backlog management

based on varying degrees of experience. K-means cluster analysis is one of the most well-known clustering algorithms and is utilized to assign observations into k clusters based on certain specified characteristics; some of these characteristics include similarity among observations within the same cluster and distinction among observations between different clusters (IBM, 2022).

The k-means cluster analysis was identified as appropriate for this study due to its efficiency in treating large datasets by producing distinct and non-overlapping clusters (Ahamad & Bharti, 2021). This was achieved by fixing a priori the number of clusters, K , based on both theoretical considerations and an exploratory analysis. The algorithm repeatedly assigned all respondents to the nearest centroid, recalculating the centroid based on the mean position of all points in the cluster, and repeated that process until there was no further reassignment (convergence). By conducting the initial exploratory analysis for several values of k , the optimal number of clusters was set at 2. The process was guided by within-cluster sum of squares (WCSS) evaluation and the elbow method which shows WCSS versus the number of clusters, and it identifies the point at which a sharp decrease in the rate of decrease occurs as being the optimal k (Ketchen & Shook, 1996).

For the quality and reliability assessment of the clustering solution, a one-way ANOVA was conducted in order to compare the mean scores of the factors extracted across the clusters. The results acquired are listed below in Table 12; specifically, for all factors, the F values and related p -values point out that there exist significant differences across clusters. As the differences reached statistical significance, it can be assumed that clusters will be quite well-differentiated on the grounds of factor scores, thus also establishing a given clustering solution's validity (Hair et al., 2014).

Interpreting the clusters that came out was based on the examination of the mean factor scores of the variables under each of the clusters. These centroid scores made it possible to describe typical features and practices of Product Backlog management

within the respective clusters and allowed recoding them to years of experience of the respondents. Table 12 shows the results of the k-means cluster analysis carried out in this study. Firstly, the questionnaire questions are shown, which divide the respondents into clusters based on their years of experience in using the Scrum framework. Next, the mean squares between and within clusters, F and significance (i.e. p-value) for each question and the final cluster centers for the first and second clusters are shown. The Mean Square between clusters represents the variance between clusters, i.e. how much the means of the different groups differ from each other and is calculated as the sum of the squares of the differences between clusters divided by the degrees of freedom between clusters (Fränti et al., 2006). The intra-cluster mean square represents the variance within clusters, i.e. how spread out the data are within each cluster compared to the cluster mean (Sreevalsan-Nair, 2021). The value shown in the 'F' column is the split between the "between" cluster mean square and the "within" cluster mean square. A high F value indicates that there is much more variability between clusters than within clusters, suggesting a clear distinction between them, i.e. the data are well separated in the identified clusters (Breunig et al., 2000). The p-value indicates the statistical significance of the F-value and measures the probability that the observed differences between clusters are due to chance rather than actual differences between clusters (Fränti et al., 2006). In general, a p-value of less than 0.05 indicates that the differences between clusters are statistically significant. Finally, the table shows the final cluster centers (also called cluster centroids or cluster centers of gravity), which represent the average coordinates of each cluster across all variables used in the cluster analysis. Each final center is a combination of the averages of the variables for all elements included in the cluster and, in other words, describes the typical characteristics of an average cluster member. Finally, it should be noted that the final centers shown in the table correspond to the range of answers (7-point Likert scale) available to the respondents. Thus, values close to 1 indicate disagreement with the statement, while values close

to 7 indicate a high level of agreement with the question. It should be stressed that those questions in the questionnaire that did not have a p-value that would indicate reliable statistical significance have been excluded from Table 12.

As already mentioned, the number of clusters obtained was 2. Since they are based on the variable expressing the number of years of experience in using the Scrum framework, cluster 1 and cluster 2 represent a group of respondents with a low level of experience and a group with a higher level of experience, respectively. By analyzing the data obtained, it was possible to identify the existence of specific approaches to product backlog management based on the experience gained in this area. In particular, the following conclusions can be drawn:

- *"Individuals belonging to the cluster characterized by a higher level of experience implement Product Backlog management practices with a superior degree of flexibility"*

Looking at the table below, it can be seen that the respondents belonging to the cluster with a higher level of experience in using the Scrum framework adopt a more flexible approach to managing the Product Backlog than their colleagues with less experience. This can be deduced from several statements and questions in the questionnaire and the corresponding final cluster centers. Firstly, it can be seen that in all the statements where it was asked whether a certain practice is carried out by a single person in a specific role or through the collaboration of the whole team, those with less experience tend to follow the schemes and rules imposed by the Scrum framework more rigidly. For example, it seems that less experienced respondents tend to assign the tasks of creating stories, estimating items or modifying the structure of the Product Backlog exclusively to the Product Owner more than more experienced respondents. In fact, for the latter, there seems to be less rigidity in defining specific tasks for each team member and more of a collaborative approach throughout the whole Scrum team. This could be explained by the fact that an individual with less

experience and familiarity in the use of Scrum tends to rely on specific figures within the team with a higher level of experience (Dingsøyr et al., 2012).

Furthermore, another element suggesting that a higher degree of experience also corresponds to a higher level of flexibility concerns the use of specific techniques in the implementation of practices such as the estimation process. In fact, the more experienced cluster tends to answer more negatively than less experienced respondents to those questions that impose a strict application of specific techniques and schemes such as estimating in days or hours or converting estimates from relative to absolute. This is probably due to the fact that more experienced respondents do not approach every project in the same way but use a degree of flexibility that allows them to adapt the management of the Product Backlog to each specific case (Fernández-Diego et al., 2020). Finally, the higher level of flexibility implemented by the more experienced respondents in the use of Scrum is also demonstrated by the usage of additional sources in the estimation process of the items. In fact, it seems that the more experienced respondents rely more on the use of historical data to make estimates than their less experienced colleagues. This could be due to the fact that experienced professionals have many projects behind them from which to take inspiration, unlike those who have only been using Scrum for a short time and therefore have to rely only on current sources (Duszkiewicz et al., 2022).

- *"Users belonging to the more experienced cluster have a lower tendency to rely on individuals outside the Scrum team compared to the group of users belonging to the less experienced cluster"*

Within the questionnaire there are questions about the imposition of certain structures and practices on the Scrum Team by external parties (i.e. customers, company, stakeholders, etc.). From Table 12, it can be seen that those with less experience tend to be more reliant on and subject to these external impositions, thus limiting their freedom in managing the Product Backlog. For

example, it appears that less experienced respondents respond more positively than more experienced ones to the statement that the Product Backlog structure is determined by the company and that the Scrum team does not have the power to change it. In addition, less experienced respondents are more likely to say that the priorities of the items are imposed on the Scrum team or determined by the customers. This reduced freedom in the management of the Product Backlog on the part of the less experienced respondents may be motivated by the fact that, with a limited degree of familiarity in the field, they tend to rely more on externally imposed practices and make fewer decisions that deviate from pre-established patterns (Dingsøy et al., 2012; Hummel & Epp, 2015). On the other hand, the experience gained by Cluster 2 respondents allows them to manage the Product Backlog more actively and to take more responsibility (Hoda et al., 2017; Edison et al., 2021).

- *"Individuals belonging to the more experienced cluster perceive a lower degree of difficulty in managing the Product Backlog compared to their colleagues belonging to the less experienced cluster"*

As might be expected, this cluster analysis demonstrates the fact that a higher level of experience corresponds to a lower level of perceived difficulty in managing the Product Backlog. This is demonstrated by the questions in the lower part of the table, which explore the respondents' difficulty in keeping the Product Backlog up to date, having items with the right level of detail, prioritizing between items and the fact that managing the Product Backlog is time consuming. As already mentioned, more experienced respondents tend to respond more negatively to these statements than their less experienced colleagues.

- *"Users belonging to the more experienced cluster possess a greater tendency to prioritize items (even continuously) and set Sprint Goals compared to individuals belonging to the less experienced cluster"*

The analysis of the table below shows that the more experienced respondents are more likely to prioritize items within the Product Backlog and to set Sprint Goals. In fact, when asked "How many items in the Product Backlog are prioritized?", the more experienced cluster responds on average with the Likert scale option "Most of them", while the less experienced cluster responds with "About half of them". This tendency of the more experienced respondents may be due to several reasons, such as a better understanding of the importance of priorities and better experience in adapting to change. In fact, more experienced users know that the prioritization of items in the Product Backlog is critical to the success of the project. They have directly experienced the consequences of poor priority management and understand how allocating resources and time to the most important items can make a difference to achieving business goals. Regarding the tendency of more experienced respondents to set more sprint goals, this may be due to the fact that a high level of experience with a large number of projects leads to a greater awareness of the importance of focusing on a specific goal to be achieved through a sprint (Edison et al., 2021).

| <i>Questions</i> | <i>Mean Square (between clusters)</i> | <i>Mean Square (within clusters)</i> | <i>F</i> | <i>Sig.</i> | <i>Cluster 1 - Less experience</i> | <i>Cluster 2 - More experience</i> |
|--|---|--|----------|-------------|--|--|
| Stories are created by the product owner alone | 75,153 | 2,714 | 27,696 | <,001 | 4 | 3 |
| Stories are created collaboratively at the team level | 70,911 | 1,68 | 42,218 | <,001 | 5 | 6 |
| Stories are created by stakeholders external to the Scrum team | 33,216 | 2,82 | 11,778 | <,001 | 3 | 2 |
| Stories have a size that permit them to be implanted in a single sprint | 71,205 | 1,844 | 38,605 | <,001 | 5 | 6 |
| At the task level, we estimate in hours or days | 147,214 | 3,781 | 38,931 | <,001 | 5 | 3 |
| Estimates are done by the product owner alone | 25,93 | 1,167 | 22,221 | <,001 | 2 | 1 |
| Estimates are done at the scrum team level | 10,604 | 0,939 | 4,653 | 0,032 | 5 | 6 |
| Estimates are done by the customer | 22,968 | 1,48 | 24,462 | <,001 | 2 | 1 |
| Estimates are imposed to the scrum team | 29,468 | 1,593 | 19,912 | <,001 | 2 | 2 |
| Historical data are used in the estimate process | 19,729 | 2,829 | 12,389 | <,001 | 5 | 6 |
| We estimate by leveraging relative estimates but we convert relative estimates into absolute estimates (hours, days etc.) because of reporting needs | 123,627 | 1,319 | 43,692 | <,001 | 4 | 3 |
| How many items of the backlog are prioritized? | 11,967 | 1,691 | 9,076 | 0,003 | 3 | 4 |
| The backlog is continuously prioritized | 67,871 | 2,114 | 40,132 | <,001 | 5 | 6 |
| Priorities are defined in a collaborative way | 17,355 | 3,008 | 8,211 | 0,005 | 5 | 5 |
| Priorities are imposed to the Scrum team | 69,214 | 3,111 | 23,01 | <,001 | 4 | 3 |
| Priorities are defined by the customer | 37,562 | 2,39 | 12,074 | <,001 | 4 | 3 |
| It is difficult to define priorities | 50,015 | 2,029 | 20,926 | <,001 | 4 | 3 |
| A sprint goals is defined | 46,892 | 1,207 | 23,11 | <,001 | 5 | 6 |
| The length of Sprints is fixed | 12,339 | 0,868 | 10,223 | 0,002 | 6 | 6 |
| All the information of the product backlog is accessible to the Scrum team | 36,01 | 2,689 | 41,497 | <,001 | 6 | 7 |
| Only the Product Owner can modify the backlog | 139,621 | 2,814 | 51,926 | <,001 | 5 | 3 |
| All the components of the Scrum team can modify the backlog | 111,378 | 2,444 | 39,575 | <,001 | 3 | 5 |
| The structure of the product backlog is defined by the company, the team cannot modify it | 117,313 | 2,361 | 47,994 | <,001 | 4 | 2 |
| There is a mismatch between the priorities desired by the customer and the actual implementation of our work | 55,283 | 1,908 | 28,974 | <,001 | 4 | 3 |
| It is very difficult to maintain the backlog up to date | 70,759 | 2,236 | 31,644 | <,001 | 4 | 3 |
| It is difficult to have items with the proper level of detail | 47,912 | 2,567 | 18,668 | <,001 | 5 | 4 |
| The management of the backlog is too time consuming | 57,962 | 2,085 | 27,802 | <,001 | 4 | 3 |

Table 12; SPSS; Cluster Analysis ANOVA Table

5 - Managerial implications and conclusions

This section of the study discusses in detail the main findings, their contribution to the current knowledge in the field of Agile, and how they can influence the way Scrum users work. The main purpose of this research was to test Scrum's theoretical principles regarding product backlog management and to understand how much the daily implementation of Scrum deviates from them. The achievement of this primary goal was contingent upon the accomplishment of 3 further objectives.

The first one consisted in determining if there are common patterns in the implementation of specific product backlog management practices. In this regard, the study found that not all aspects described by the Scrum principles are handled in the same way by all users. This demonstrates that the Scrum Principles are intentionally generic and incomplete, as their originators intended to provide general tools for users to adapt the framework to any context. However, in some areas of product backlog management, the results showed general agreement among respondents. Indeed, they agreed that the creation of stories, estimates, and priorities should be done through collaboration and involvement of all members of the Scrum team. This leads to numerous benefits, such as considering a variety of perspectives, aligning each team member with project goals, and understanding the product backlog management methods used reducing the risk of misunderstandings and disputes within the team. Moreover, a considerable degree of transparency in the Product Backlog leads to greater alignment and awareness among team members about the product development and goals, which are fundamental principles determining the effectiveness of the Scrum framework. The responses also indicated a common need for continuity in the estimation and prioritization processes, as the Product Backlog is not a rigid and static tool but needs to be constantly updated and refined in an iterative way. It is therefore of the utmost importance that companies cultivate a collaborative environment within the Scrum team, facilitate regular refinement of the product

backlog, and refrain from the adoption of specific practices by individuals that may disadvantage others.

The second aim of this study was to understand the factors that influence the perceived level of satisfaction with product backlog management. As demonstrated by the regression model, these factors can be summarized into six main areas: perceived difficulty, flexibility used in managing the product backlog, the division of stories into smaller tasks, accessibility and transparency of the product backlog content, alignment between the team's work and the priorities expressed by the customer, and collaboration implemented in the processes of prioritization, estimation and story creation. As for difficulty, it could be reduced by setting clearer guidelines for creating stories, which would help overcoming the issues expressed by respondents in defining the right level of detail in product backlog items. Moreover, companies could reduce difficulty by continually updating the product backlog to prevent it from becoming outdated and misaligned with customer needs. Furthermore, considering techniques such as the "20/80 Rule" might be helpful to ensure that Product Backlog management does not become overly time-consuming. Indeed, it could assist Scrum Teams in focusing on the truly urgent items and those that provide the highest value. As for a higher level of flexibility, an increase of which has a positive impact on satisfaction, it could be achieved in several ways. First, an ongoing backlog "refinement" process would allow teams to adjust and modify backlog items as new information and priorities emerge during the course of the project. Second, greater flexibility can be achieved by adopting an approach that encourages collaboration between team members and less use of specific roles in managing practices related to the Product Backlog. In addition to this, flexibility can be promoted by creating more flexible stories that can adapt to changes as the project progresses. This can be done by using more elastic acceptance criteria and by avoiding splitting stories into overly detailed and rigid elements. Finally, leaving spare capacity in Sprints for dealing with emergencies would allow the team to react with greater agility to unforeseen events, avoiding work

overload and improving priority management based on the urgency of new needs. Finally, the fact that greater flexibility in product backlog management was shown in this study to increase user satisfaction is consistent with discussions in the literature about the "framework trap". This is described in numerous studies as a trap in which Scrum teams apply the framework in a rigid and mechanical way, following rules without understanding the underlying agile principles such as empiricism and continuous adaptation (Wolpers, 2023). Furthermore, academic research has pointed out that uncritical adoption of Scrum can lead to limited effectiveness if the framework is treated as a set of rules to be followed rather than a flexible and adaptable approach (Richter & Siemon, 2019). For these reasons, it can be suggested that organizations using Scrum implement practices that increase the degree of flexibility, such as those described in this section. As for the third factor that influences satisfaction, the division of stories into smaller tasks leads to an excessive level of complexity in product backlog management, making it time-consuming without adding real value. To improve satisfaction, companies should ensure that product backlog items are as simple as possible. Furthermore, accessibility and transparency of the product backlog are crucial factors for satisfaction. For this reason, managers should implement practices and tools that promote transparency and facilitate alignment, such as adopting a visible and accessible product backlog management platform, holding regular alignment meetings on the content of the product backlog, and creating clear and centralized documentation. This study demonstrated that the mismatch between customer's desired priorities and the actual implementation of the work by the Scrum Team can negatively impact satisfaction. To avoid this, organizations should ensure high levels of communication between the Scrum team and the customer, ensuring their goals and expectations are fully understood and applied. Finally, a collaborative approach to managing the processes of item prioritization, stories, and estimates can positively impact satisfaction. To this end, encouraging the involvement of the entire team in backlog refinement meetings, implementing more collaborative estimation techniques

such as "Planning Poker", and minimizing instances where specific individuals take responsibility for specific aspects of the product backlog are essential.

The third and last objective of this research was to investigate the changes in product backlog management with different levels of experience in using Scrum. The results showed that individuals that have been using Scrum for a longer period of time tend to implement an approach to Product Backlog management characterized by higher flexibility, a lower tendency to rely on people outside the scrum team, a lower perception of difficulty in managing the Product Backlog, and a greater tendency to frequently prioritize items and set goals. The increased flexibility in product backlog management can be attributed to several reasons. First, experienced Scrum users feel more confident in applying the framework flexibly due to their deeper understanding of its principles and practices. Second, experienced users are able to adapt to internal and external demands on the Scrum team, allowing them to change priorities dynamically according to changes in the business environment or resources. Third, experienced team members have a greater ability to make decisions based on strategic priorities, leading to a more responsive and less rigid approach to product backlog management. However, there is a reverse relationship between experience and the tendency to rely on external parties in product backlog management practices. Less experienced individuals have greater uncertainty in managing the product backlog, which may lead them to prefer more external imposition of rules and less decision-making autonomy. This implies that Scrum teams composed of more experienced members require less supervision or external input, allowing organizations to focus on lighter and decentralized governance. However, this improved operational efficiency and decision-making autonomy must be accompanied by clear communication to avoid misalignment with broader strategic goals. Lastly, the tendency shown by individuals belonging to the highly experienced cluster to frequently prioritize items and set sprint goals might be due to their ability to adapt to customer feedback or product requirements, their awareness of the importance of setting well-defined sprint goals,

and the development of a more rigorous iterative discipline. The analysis of the managerial implications of the results obtained from the regression model showed that there are several factors that increase the level of satisfaction in managing the product backlog. Furthermore, the cluster analysis showed that some of these factors are adopted to a greater extent by individuals belonging to the more experienced cluster. These include the degree of flexibility used in product backlog management, the lower perceived difficulty and the continuous and frequent prioritization of items. Therefore, It can be concluded that companies could increase the level of team satisfaction in product backlog management by including more experienced individuals in their Scrum teams. In fact, experienced members could direct the work of less experienced users towards an approach to product backlog management that would increase the overall level of satisfaction.

5.1 Limitations of the study and directions for future research

The work carried out in this thesis has three main limitations that could be overcome by future research. The first is indicated by the size and representativeness of the sample used. In fact, the research is based on a questionnaire distributed via LinkedIn, which collected 216 responses from professionals using Scrum. However, most of the respondents come from large companies in the IT or consulting sector (about 75% in companies with more than 250 employees). This distribution may limit the generalizability of the results to other sectors or to small and medium-sized companies, as product backlog management practices may vary depending on the business context. Future research could expand the sample of respondents to include professionals from companies of different sizes and sectors to explore if and how product backlog management differs in different contexts. The second limitation of this thesis is the limited analysis of external variables that might influence the way the product backlog is managed. In fact, the thesis focuses more on internal variables rather than external factors such as corporate culture or the market context in which the Scrum user company operates. This could be an interesting point for future

research, which could investigate more the influence of external factors on the management of the product backlog and the success of Scrum practices. Finally, the last limitation of this research paper is the limited time period in which the data collection took place, which reduces the ability to observe the evolution of Product Backlog management practices over time. It would be appropriate for future research to collect data over a longer time period in order to identify any changes in product backlog management over time.

6 – Appendix

The following section reports the questionnaire administered to Scrum Masters, Product Owners and developers across the world divided into thematic areas.

Block 1: General information about respondents

Q1) What industry are you in?

- Consulting
- IT
- Automotive
- Food
- Agriculture
- Construction
- Manufacturing
- Retail
- Healthcare
- Education
- Banking & Finance
- Energy
- Telecommunication
- Tourism
- Other (Please specify)

Q2) How many employees does your company have?

- Less than 10
- Between 10 and 49
- Between 49 and 249
- More than 250

“To answer the following questions, think about a specific product or project in the development of which you are currently involved.”

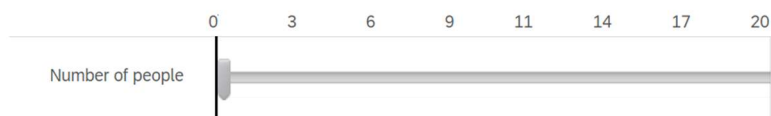
Q3) What is your role? You can select more than one role

- Product Owner
- Scrum Master
- Development Team
- Other (Please specify)

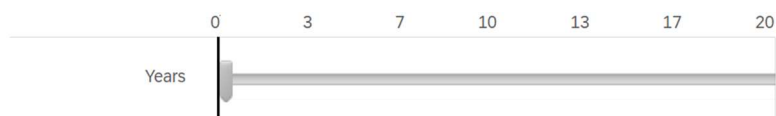
Q4) Is the product/project you are thinking about internal or addressed to a customer?

- Internal
- Customer
- Other (Please specify)

Q5) How many people work in your Scrum Team?



Q6) How long have you been using Scrum?



Block 2: Product Backlog Composition

Q1) How are Backlog Items classified/represented? You can select more than one answer

- Themes
- Epics
- User Stories

- Tasks
- Spikes
- Bugs
- Enablers
- Other (Please specify)

Q2) Which fields do you use in your Product Backlog? You can select more than one answer

- ID
- Title
- Description
- Priority
- Assigned Sprint
- Status
- Size/Estimate
- Acceptance Criteria
- Assignee
- Dependencies between items
- Other (Please specify)

Block 3: Stories

Q1) How many items of the backlog are described with the typical syntax of user stories?

- None of them
- Some of them
- About half of them
- Most of them
- All of them

"From now on we will use the term "Story" to indicate the smallest unit of work, typically represented by a functionality, in the Product Backlog."

Q2) Please, rate your level of agreement/disagreement to the following statements based on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree):

- "Stories are created by the product owner alone"
- "Stories are created collaboratively at the team level"
- "Stories are created by the customer"
- "Stories are created by stakeholders external to the Scrum team"
- "Stories have a size that permit them to be implanted in a single sprint"
- "Stories are broken down into tasks when there is enough information to do it"
- "Stories are broken down into tasks during Sprint Planning"
- "We do not break down Stories into tasks"

Block 4: Estimates

Q1) How many items of the backlog have undergone an estimation process?

- None of them
- Some of them
- About half of them
- Most of them
- All of them

Q2) Please, rate your level of agreement/disagreement to the following statements based on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree):

- "We continuously estimate/re-estimate the items of the backlog"
- "We estimate/re-estimate the items of the backlog during dedicated meetings"
- "At the task level, we estimate in hours or days"
- "Estimates are done by the product owner alone"

- "Estimates are done by the development team alone"
- "Estimates are done at the scrum team level"
- "Estimates are done by the customer"
- "Estimates are imposed to the scrum team"
- "External experts are involved in the estimate process"
- "Historical data are used in the estimate process"
- "We estimate by leveraging relative estimates but we convert relative estimates into absolute estimates (hours, days etc.) because of reporting needs"
- "It is difficult to provide accurate estimates"

Block 5: Prioritization

Q1) How many items of the backlog are prioritized?

- None of them
- Some of them
- About half of them
- Most of them
- All of them

Q2) Please, rate your level of agreement/disagreement to the following statements based on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree):

- "The backlog is continuously prioritized"
- "The backlog is prioritized during specific events (e.g. backlog refinement meetings)"
- "Priorities are defined in a collaborative way"
- "Priorities are imposed to the Scrum team"
- "Priorities are defined by the customer"
- "Priorities are influenced by internal constraints (e.g. resource availability)"
- "It is difficult to define priorities"

Block 6: General Product Backlog Management

Q1) Please, rate your level of agreement/disagreement to the following statements based on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree):

- "A sprint goals is defined"
- "The length of Sprints is fixed"
- "All the information of the product backlog is accessible to the Scrum team"
- "Only the Product Owner can modify the backlog"
- "All the components of the Scrum team can modify the backlog"
- "The structure of the product backlog is defined by the company; the team cannot modify it"
- "The Product backlog is customized in order to take into account the customer's preferences"
- "The Product backlog is customized in order to take into account the customer's competencies"
- "The Product backlog structure does not change according to the customer's preferences and competences"
- "There is a mismatch between the priorities desired by the customer and the actual implementation of our work"
- "Once the team is committed to specific items during a Sprint, no changes are allowed (that is, the team implements the items on which they are committed)"
- "We leave some spare capacity to our team in order to embrace urgent requests"
- "It is very difficult to maintain the backlog up to date"
- "It is difficult to have items with the proper level of detail"
- "The management of the backlog is too time consuming"
- "I am satisfied with the way the backlog is managed"

Q2) What would you like to improve in the way your Product Backlog is managed?

7 - References

- Abrahams, A. S. (2022). Collaborative approaches in Scrum teams: A pathway to effective backlog management. *Journal of Software Engineering Research and Development*, 35(2), 98-112.
- Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, 567(7748), 305-307.
- Anderson, D. J. (2010). *Kanban: Successful evolutionary change for your technology business*. Blue Hole Press.
- Ashmore, S., & Runyan, K. (2014). *Introduction to agile methods*. Addison-Wesley.
- Babiker, E., Mahmoud, H., & Abdalrahman, E. (2018). Planning Poker: A collaborative estimation method for software projects. *International Journal of Computer Applications*, 179(18), 13-17.
- Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 16(2), 296-298.
- Beck, K. (1999). *Extreme programming explained: Embrace change*. Addison-Wesley.
- Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., ... & Thomas, D. (2001). *Manifesto for agile software development*. Agile Alliance. <https://agilemanifesto.org/>
- Breyter, R. (2022). *Agile estimation techniques: Mastering the complexity of modern software development*. Springer.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276.

Charikar, M., Girard, L., & Li, J. (2022). The role of flexible backlog management in Agile: Reducing overload and improving satisfaction. *Agile Project Management Journal*, 48(3), 213-230.

Cleophas, T. J., & Zwinderman, A. H. (2021). *Statistical methods in clinical research* (8th ed.). Springer.

Cohn, M. (2005). *Agile estimating and planning*. Prentice Hall.

Cohn, M. (2006). *User stories applied: For agile software development*. Addison-Wesley.

Cockburn, A. (2006). *Agile software development: The cooperative game* (2nd ed.). Addison-Wesley.

Cockburn, A., & Highsmith, J. (2001). Agile software development: The people factor. *Computer*, 34(11), 131-133.

Conboy, K. (2009). Agility from first principles: Reconstructing the concept of agility in information systems development. *Information Systems Research*, 20(3), 329-354.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Derby, E., Larsen, D., & Schwaber, K. (2006). *Agile retrospectives: Making good teams great*. Pragmatic Bookshelf.

Dingsøy, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software*, 85(6), 1213-1221.

Fowler, M. (2019). *Refactoring: Improving the design of existing code* (2nd ed.). Addison-Wesley.

Fowler, M., & Highsmith, J. (2001). The agile manifesto. *Software Development*, 9(8), 28-35.

- Grenning, J. (2002). Planning poker or how to avoid analysis paralysis while release planning. XProgramming.com.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Highsmith, J. (2002). *Agile software development ecosystems*. Addison-Wesley.
- Hudaib, A., HajKhalil, A., Hudaib, M., & Al-Najjar, M. (2018). AHP-based requirements prioritization for software product management. *International Journal of Computer Applications*, 180(38), 10-16.
- Kaufman, L. (2009). *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons.
- Kropp, M., Meier, A., Anslow, C., & Biddle, R. (2014). Agile practices, collaboration, and experience: A qualitative study on agile software development. *Agile Conference*, 1-10.
- Kua, P. (2013). *The retrospective handbook: A guide for agile teams*. Leanpub.
- Lacey, M. (2015). *The scrum field guide: Agile advice for your first year and beyond* (2nd ed.). Addison-Wesley.
- Leonardi, P. M., Huysman, M., & Steinfield, C. (2013). Enterprise social media: Definition, history, and prospects for the study of social technologies in organizations. *Journal of Computer-Mediated Communication*, 19(1), 1-19.
- Livermore, J. A. (2007). Factors that impact implementing an agile software development methodology. *Software Engineering Research and Practice*, 7, 482-487.
- Mahnič, V., & Hovelja, T. (2012). On using planning poker for estimating user stories. *Journal of Systems and Software*, 85(9), 2086-2095.
- Malhotra, N. K. (2006). *Marketing research: An applied orientation* (5th ed.). Pearson Education.

Moe, N. B., Dingsøy, T., & Dybå, T. (2010). A teamwork model for understanding an agile team: A case study of a Scrum project. *Information and Software Technology*, 52(5), 480-491.

Moe, N. B., Šmite, D., & Šāblis, A. (2019). From agile to DevOps: Is it all about automation? *IEEE Software*, 36(2), 17-23.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS (7th ed.)*. Routledge.

Pichler, R. (2010). *Agile product management with Scrum: Creating products that customers love*. Addison-Wesley.

Racheva, Z., Daneva, M., Sikkel, K., & Herrmann, A. (2010). Do we know enough about prioritization in agile requirements engineering? *Requirements Engineering*, 15(2), 161-186.

Ramesh, B., Cao, L., & Baskerville, R. (2010). Agile requirements engineering practices and challenges: An empirical study. *Information Systems Journal*, 20(5), 449-480.

Rubin, K. S. (2012). *Essential Scrum: A practical guide to the most popular agile process*. Addison-Wesley.

Santos, D. R., Santos, G. F., Oliveira, M., & Mendonça, M. (2020). Lessons learned in agile retrospectives: A qualitative study. *Information and Software Technology*, 125, 106376.

Schön, E.-M., Escalona, M. J., Thomaschewski, J., & Winter, D. (2017). Agile process models for small teams in the context of software-based innovation projects. *International Journal of Human Capital and Information Technology Professionals*, 8(2), 38-60.

Schwaber, K., & Beedle, M. (2002). *Agile software development with Scrum*. Prentice Hall.

Schwaber, K., & Sutherland, J. (2020). *The Scrum guide: The definitive guide to Scrum*. Scrum.org.

Serrador, P., & Pinto, J. K. (2015). Does agile work? A quantitative analysis of agile project success. *International Journal of Project Management*, 33(5), 1040-1051.

Strode, D. (2016). Using the Scrum retrospective to support regulated software development. *ACM SIGSOFT Software Engineering Notes*, 41(1), 1-7.

Williams, L., & Cockburn, A. (2003). Agile software development: It's about feedback and change. *Computer*, 36(6), 39-43.

Womack, J. P., & Jones, D. T. (1996). *Lean thinking: Banish waste and create wealth in your corporation*. Simon & Schuster.

Sundararajan, R., & Subramanian, S. (2022). Risks and challenges of Scrum: A systematic literature review.

Zhao, Y., & Li, X. (2023). Product backlog optimization technique in agile software development using clustering algorithm.

Kumar, A., & Patel, R. (2022). A deep dive into product backlog prioritization technique in Scrum. *2022 IEEE International Conference on Software Engineering Advances (ICSEA)*, 56–65.

Sharma, P., & Gupta, N. (2023). Scrum project management challenges and solutions: Systematic literature review.

Nowak, A., & Kowalski, M. (2022). Scrum, Kanban or a mix of both? A systematic literature review. *Annals of Computer Science and Information Systems*, 30, 154–162.