UNIVERSITY OF TARTU
Institute of Computer Science
Software Engineering Curriculum

Henri Laats

Measuring the Impact of Developer Experience Improvements on Engagement in Pipedrive

Master's Thesis (30 ECTS)

Supervisors: Dietmar Alfred Paul Kurt Pfahl, PhD
Mykhailo Dorokhov, MSc

Tartu 2025

Measuring the Impact of Developer Experience Improvements on Engagement in Pipedrive

Abstract:

Pipedrive is a customer relationship management (CRM) software company where a significant portion of the workforce consists of developers. Understanding how changes in developer experience (DX) affect overall employee engagement is key to allocating resources effectively. In this context, employee engagement refers to how motivated and enthusiastic employees are about their work and their organization, reflecting their commitment to stay and go the extra mile. Engaged employees care about performing well and contributing to the company's success, which can lead to growth and increased revenue. Pipedrive tracks overall employee sentiment through company-wide engagement surveys and collects additional feedback from developers through DX surveys. However, the relationship between improvements in DX and overall employee engagement has not been analyzed, leaving it unclear whether a better developer experience also boosts employee engagement. The aim of this thesis is to analyze data from previous surveys to understand the relationships between areas in developer experience and overall employee engagement. The thesis will also explore which aspects of DX most influence engagement scores. Various statistical methods such as correlation, regression and trend analysis are used to identify these key relationships. The findings provide insight into whether improving developer experience can influence engagement positively and describe how to identify areas for potential DX improvements that would have the biggest impact on engagement.

Keywords:

Customer relationship management, statistical analysis, correlation analysis, regression analysis, principal component analysis, employee engagement, developer experience

CERCS: P170 Computer science, numerical analysis, systems, control

Arendajate kogemuse parendamise mõju töötajate kaasatusele Pipedrive'is

Lühikokkuvõte:

Pipedrive on kliendisuhete haldamise (CRM) tarkvaraettevõte, kus oluline osa tööjõust koosneb infotehnoloogia valdkonna inimestest. Arusaam sellest, kuidas muudatused arendajakogemuses (DX) mõjutavad üldist töötajate kaasatust, on oluline, et efektiivsemalt ettevõtte ressursse jaotada. Käesoleva lõputöö kontekstis viitab töötajate kaasatus sellele, kui motiveeritud ja entusiastlikud on töötajad oma tööst ja organisatsioonist, peegeldades nende pühendumust jääda ettevõtesse pikaajaliselt ja igapäevaselt anda endast parim. Kaasatud töötajad hoolivad heade tulemuste saavutamisest ja ettevõtte edule kaasaaitamisest, mis omakorda aitab kaasa tuua ettevõtte kasvu ja suurema tulu. Pipedrive jälgib üleüldist töötajate kaasatust ettevõtteüleste kaasamisuuringute kaudu ning kogub DX-uuringute kaudu arendajatelt täiendavat tööalasast tagasisidet. Siiani ei ole Pipedrives analüüsitud DX-i ja töötajate kaasatuse vahelist seost, ehk on teadmata, kas parem arendajakogemus suurendab töötajate kaasatust. Käesoleva lõputöö eesmärk on analüüsida varasemate küsitluste andmeid, et mõista arendajakogemuse ja töötajate üldises kaasatuse vaheliseid seoseid. Samuti uuritakse lõputöös, millised arendajakogemuse aspektid mõjutavad kaasatust kõige enam. Nende peamiste seoste tuvastamiseks kasutatakse erinevaid statistilisi meetodeid nagu korrelatsioon, regressioon- ja trendianalüüs. Tulemused annavad ülevaate sellest, millisel määral arendajakogemuse parendamine võib mõjutada kaasatust positiivselt ja kirjeldab raamistiku, mille abil saab ettevõte tulevikus korrata parendamist vajavate arendajakogemuse aspektide määramist.

Võtmesõnad:

Kliendihaldustarkvara, statistiline analüüs, korrelatsioonanalüüs, regressioonanalüüs, faktoranalüüs, töötajate kaasatus, arendajakogemus

CERCS: P170 Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine

Table of Contents

1. Introduction

Pipedrive is a customer relationship management (CRM) software company that helps businesses manage their sales relationships with clients [1]. Like any organization, Pipedrive wants to attract top employees and retain them over the long term. Studies show that engaged employees tend to be more present and productive, which boosts overall performance and profitability [2]. For a company like Pipedrive, where a significant portion of employees are developers, it's crucial to understand how changes in the engineering work environment can influence engineers' engagement to the company.

To measure employee engagement and developer satisfaction, Pipedrive uses two surveys: a company-wide employee engagement survey through the Culture Amp platform [3], and a developer experience (DX) survey on GetDX platform [4] targeted only at the engineering division. The engagement survey is broader, assessing overall enthusiasm employees feel toward the organization. This includes their pride, willingness to recommend the company, present and future commitment, and motivation. The DX survey, on the other hand, is more focused, gathering feedback on key engineering performance drivers that directly impact financial results.

These surveys help Pipedrive gain a better understanding of overall company sentiment through the engagement survey, as well as more specific sentiment from engineers through the DX survey. While these surveys provide valuable insights on their own, it would also be useful to explore whether there is any relationship between the two surveys. Understanding the relationship between engineers' day-to-day work and their engagement is valuable. If a relationship exists, it would allow the company to allocate resources to the engineering aspects that most positively impact engagement. This, in turn, could lead to the benefits associated with higher engagement, such as increased productivity, reduced absenteeism, and higher profitability. Currently, Pipedrive lacks a clear understanding of whether changes in the DX survey are related to the engagement survey questions. More importantly, it is unclear whether improvements in DX translate to higher engagement scores. The goal of this thesis is to analyze existing data from previous surveys to understand the connections between DX and engagement. Specifically, it will explore whether improvements in DX survey responses lead to better engagement outcomes. By the end of the thesis, a framework exists that can be used in the future to identify which areas of developer

experience are most important to improve in order to increase future engagement. The main benefit of this thesis for Pipedrive is the ability to prioritize changes that have the most significant positive impact on employee engagement and ultimately improve organizational performance.

To clearly define the purpose of this thesis, which examines the relationship between the DX survey and the engagement survey, the following research questions have been created:

RQ1: What is the relationship between improvements in developer experience (DX), as measured through DX-specific surveys, and the company-wide engagement survey results conducted via Culture Amp?

RQ2: Which specific drivers of developer experience have had the strongest relationship with key metrics in the Culture Amp engagement survey?

This thesis consists of seven chapters. Chapter 2 defines employee engagement, discusses its benefits, and describes the factors that can influence it. Chapter 3 provides an overview of Pipedrive as the company, and describes the surveys used in this thesis. Chapter 4 covers the methodology, explaining how the analysis was conducted. Chapter 5 presents the results and key findings from the analysis, which are then discussed in chapter 6. The final chapter concludes the thesis, summarizing its purpose, methodology, and results.

For the purposes of improving clarity and refining the text of this thesis, AI tool ChatGPT [5] was used throughout the document.

2. Related Work

This chapter provides an overview of employee engagement, including its definition, benefits of increasing it, and the factors that influence it. This will help define the topic of the thesis and highlight its potential impact.

2.1 Defining Engagement

Although the term "employee engagement" is widely used, there is no universal definition. Terms like "satisfaction" and "commitment" are often used interchangeably with engagement, as they describe similar ideas of positive attitudes and behaviors in employees. This can lead to confusion about the exact meaning of the word.

One of the earliest tries to define employee engagement comes from a study by William A. Kahn [6], whose definition for it is widely cited in literature on the topic. There it is defined as the process through which individuals fully commit their physical, cognitive, and emotional selves to their work roles. In other words, engagement occurs when people actively invest and express energy across these three dimensions during work. The cognitive aspect in this case means what employees think and believe about the organization, its leaders, and the work environment. The emotional aspect is about how employees feel about these factors. The physical aspect refers to the amount of physical effort employees dedicate to their roles. Kahn suggests that an engaged employee is fully present, both psychologically and physically, when performing their role in the organization.

In the context of this thesis, employee engagement is defined as the level of enthusiasm and motivation employees feel toward their organization, reflecting their commitment to stay and put in extra effort. It includes factors such as pride, willingness to recommend the company, present and future commitment, and motivation. High engagement indicates a strong workplace culture and employees that are dedicated to and aligned with the company's values and goals. This definition is adopted because it aligns with the one used in the employee engagement survey platform, where Pipedrive conducts its engagement surveys [7].

2.2 Benefits of Higher Engagement

Employee engagement often plays an important role in a company's long-term success. When employees are engaged, they are not just fulfilling their duty, they are also emotionally and intellectually invested in helping the company achieve its goals. According to a recent Gallup meta-analysis on the relationship between engagement at work and organizational outcomes [2], the study examined the influence of employee engagement across multiple business metrics. The findings show that companies with higher engagement levels see major improvements in key performance areas. The 11 key performance areas in the Gallup's research were: customer loyalty/engagement, profitability, productivity, turnover, safety incidents (accidents), absenteeism, shrinkage (theft), patient safety incidents (mortality and falls), quality (defects), wellbeing (thriving employees), and organizational citizenship (participation). The findings clearly show that businesses with highly engaged employees experience, on average, have an 18% higher productivity and a 23% increase in profitability compared to those with low engagement [2].

In addition, engagement strongly impacts employee retention. Research shows that engaged employees are significantly less likely to leave their roles [8], and companies with strong engagement strategies can see turnover rates drop by as much as 21% [2]. This reduction results in lower recruitment costs, as it is more cost-effective to retain existing employees, while also enhancing knowledge retention across teams. Engaged employees also tend to demonstrate higher levels of creativity and problem-solving, which in turn can drive innovation and improve a company's ability to adapt to changing markets [9]. Beyond measurable results, higher engagement also encourages behaviors that strengthen workplace culture, which refers to the aspects shaping the environment, interactions, and dynamics within a workplace. Engaged employees are more likely to support their colleagues, take initiative, and contribute ideas that go beyond their immediate responsibilities [9].

2.3 Factors Influencing Employee Engagement

Employee engagement can be influenced by a variety of factors, each playing a different role in how connected and motivated employees feel at work. An important factor is the manager's role

in influencing the engagement of their team members. The Gallup report [2] mentions that well-supported and trained managers can significantly boost team engagement. Organizations with high engagement rates often invest in regular training and development opportunities for their employees. Career discussions, easy access to self-development, and clear paths for advancement have been proven to be an impactful role in maintaining engagement and loyalty [10]. Another important way companies can boost engagement is by creating an environment where employees feel safe to share their ideas without fear of negative consequences. Research [11] shows that psychological safety promotes open communication, innovation, and team performance by encouraging experimentation and risk-taking. Employees are generally more engaged when their work aligns with their personal values and when they perceive a greater purpose beyond just completing tasks [12]. Therefore, connecting daily activities to broader organizational goals helps strengthen motivation and commitment as the employees can better understand their value within the company. While there are many other factors influencing engagement, recognizing employees and rewarding them for their achievements at work is another important one, as it motivates them to continue doing quality work and keeping up with the high standard [13]. Finally, workplace conditions, access to necessary resources and overall environment play an important role in engagement, as employees are more engaged when they feel supported, have the tools they need, and can collaborate effectively with colleagues [2].

On the Culture Amp platform, where Pipedrive conducts its engagement survey, they have created a guide [7] based on their data showing what most influences employee engagement. According to their findings, the top drivers of engagement across industries are leadership, learning and development, and company performance. While factors can vary between companies, these were the trends they have determined to have remained consistent across the data over the years.

3. Background

In the introduction of this thesis, the main objective was defined as identifying relationships between the developer experience (DX) survey and the company-wide engagement survey results. This chapter provides an overview of Pipedrive as a company, along with a detailed description of the two surveys used for analysis.

3.1 Pipedrive

Pipedrive is a cloud-based software as a service (SaaS) company that provides customer relationship management (CRM) software, enabling users to track relationships with their clients [1]. Founded in 2010, Pipedrive became the fifth unicorn in Estonia, reaching a valuation of over $1 billion in 2020 when Vista Equity Partners, a global investment firm, acquired a majority share in the company [14].

Pipedrive was listed as one of the 10 Best CRM Software of 2025 by Forbes [15], an American business magazine. Currently, more than 100,000 companies across over 170 countries use Pipedrive to manage their sales activities, with the largest number of customers based in the United States [16]. The company has offices in eight countries worldwide, employs over 800 people [1] across several divisions and is regularly listed as the top ranked employer in Estonia [17]. The largest division is the engineering division, which includes around 200 employees spread across three locations: Estonia, Portugal, and the Czech Republic. This division includes different IT roles like data engineers, software engineers, DevOps engineers, and others. Each engineer works in a team that focuses on a specific area. The structure of the engineering division can be seen in Figure 1 below.

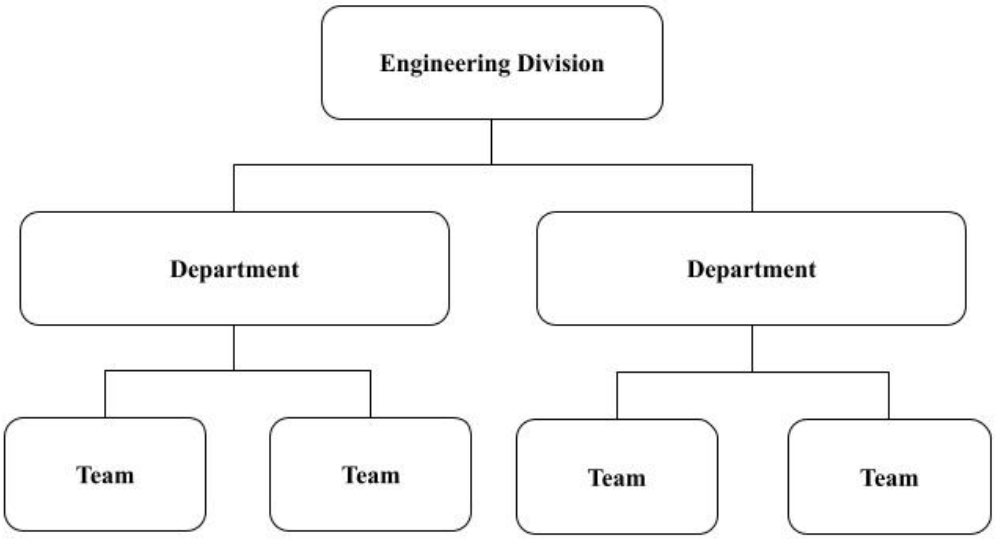

Figure 1. Engineering division structure at Pipedrive.

Teams with similar goals are grouped into departments, and all the departments together form Pipedrive's engineering division. In the context of this thesis the analysis was done on department level because the access to each teams' individual results was restricted.

3.2 Survey Instruments at Pipedrive

This sub-chapter provides an overview of the two surveys used at Pipedrive to assess employee engagement and understand how developers perceive their work environment. This chapter explains the structure of the surveys, including the questions, response scales, time frames, and other relevant details in the context of this thesis.

3.2.1 Employee Engagement Survey

The employee engagement survey is used to measure how enthusiastic and motivated employees are about their organization [7]. It helps to understand their commitment to stay with the company and put in extra effort. The survey looks at five main areas: motivation, pride, recommendations, present commitment, and future commitment. A more detailed explanation of the definition of

employee engagement can be found in chapter 2. At Pipedrive, the employee engagement survey is conducted using the Culture Amp platform, which is also used by companies like McDonald's and Canva [18]. The results from the survey help the company identify areas to improve and find ways to boost employee satisfaction and retention.

The engagement survey has been conducted on the Culture Amp platform at Pipedrive for the past two years, twice a year, for a total of four times. Additionally, a shorter version of the survey, called Pulse, has been conducted two times. The shorter surveys are used to gather quick, interim results, as the full engagement survey takes longer to complete. The Pulse surveys contain about half the number of questions compared to the full surveys and are used in this thesis to provide more data for analysis. In total, six surveys have been completed, including both the full engagement and Pulse surveys.

The survey is targeted to all Pipedrive employees, around 800 people across various divisions, including approximately 200 from the engineering division. The average participation rate has been around 80-90%, and the rate in the engineering division is similar to the company average. The participation rate for the engineering division is shown in the Figure 2 below.

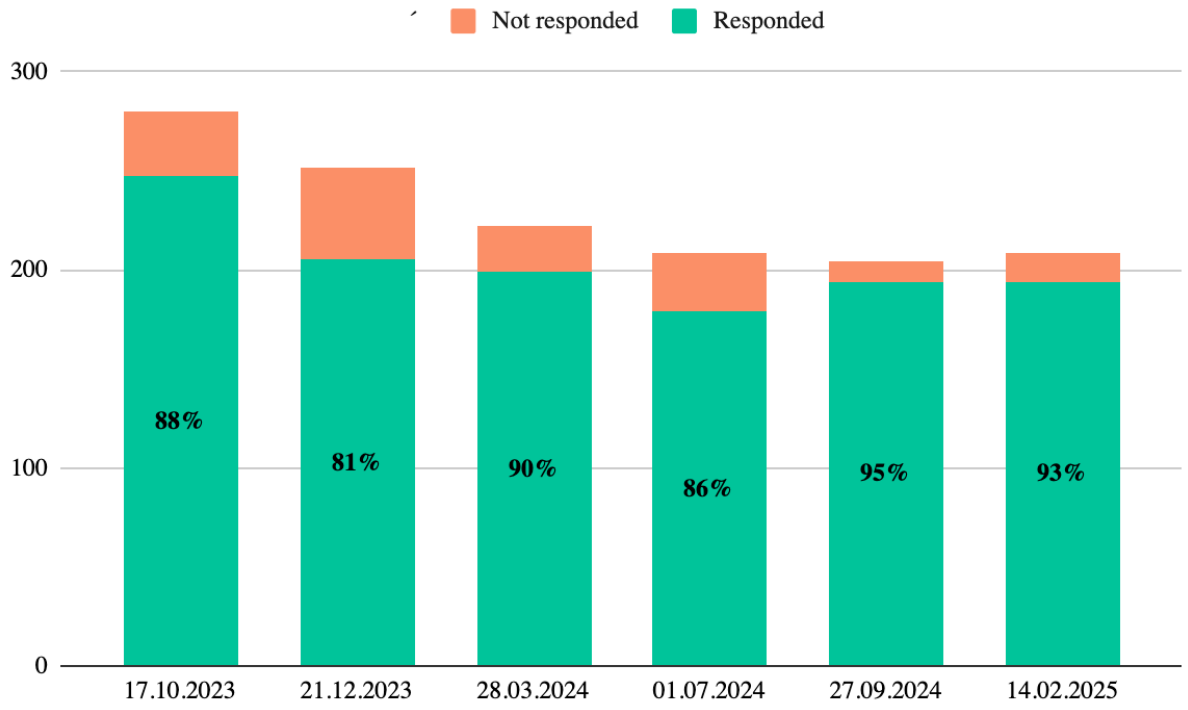

Figure 2. Engineering division engagement survey response rates.

The questions in the engagement survey are categorized into different groups, with the most important category being 'engagement.' This category targets five key areas: motivation, pride, recommendations, present commitment, and future commitment, each of which has a separate question. The overall engagement score is calculated as the mean of the positive response percentages from five questions in the engagement category, providing an overall summary of employee engagement within the company or the selected subgroup [7]. The overall engagement score for the engineering division has ranged between 53% to 77% across the six surveys, with the most recent survey showing the highest score of 77%, as shown in Figure 3.

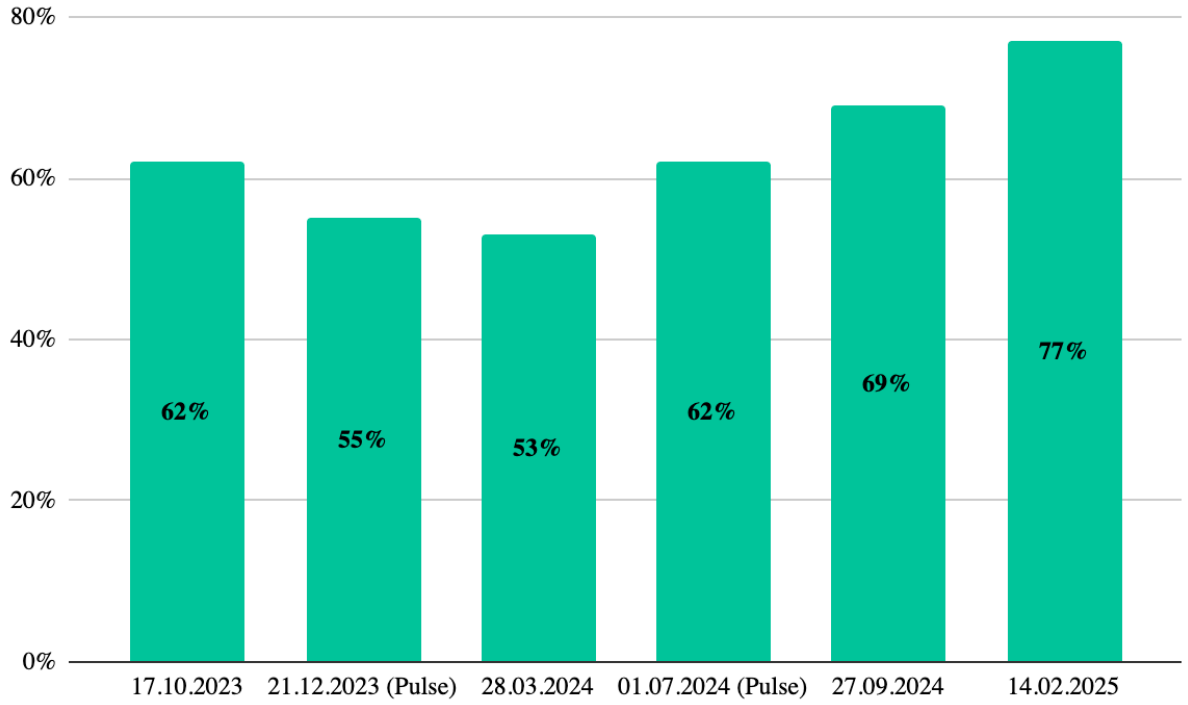

Figure 3. Overall engagement scores in the engineering division over time.

The other questions do not directly affect the engagement score as they address different categories. These categories, recommended by Culture Amp, include leadership, enablement, alignment with the company, and opportunities for development. However, each question outside the engagement category is given a correlation score with the five engagement questions [19]. Based on this, questions not from the engagement category can also be linked to engagement. The impact level is determined by analyzing how responses to engagement questions correlate with responses to other questions. People who respond most positively to engagement questions also

tend to respond more positively to other questions, and vice versa for those who respond negatively. The correlation and impact ratings are as follows [19]:

1. 0.70 – Extreme
2. 0.50 – 0.70 – Very Strong
3. 0.40 – 0.50 – Strong
4. 0.30 – 0.40 – Moderate
5. 0.20 – 0.30 – Low
6. < 0.20 – Insignificant

The number of rating questions in each survey varies, having ranged from 23 to 96. According to the Culture Amp platform, 19 main rating questions should be included in every engagement survey [20]. The company generally uses these recommended questions, but often adds additional ones to gather feedback relevant to specific situations or interests they have, such as structural changes. In Figure 4 below is shown the number of rating questions in the engagement surveys.

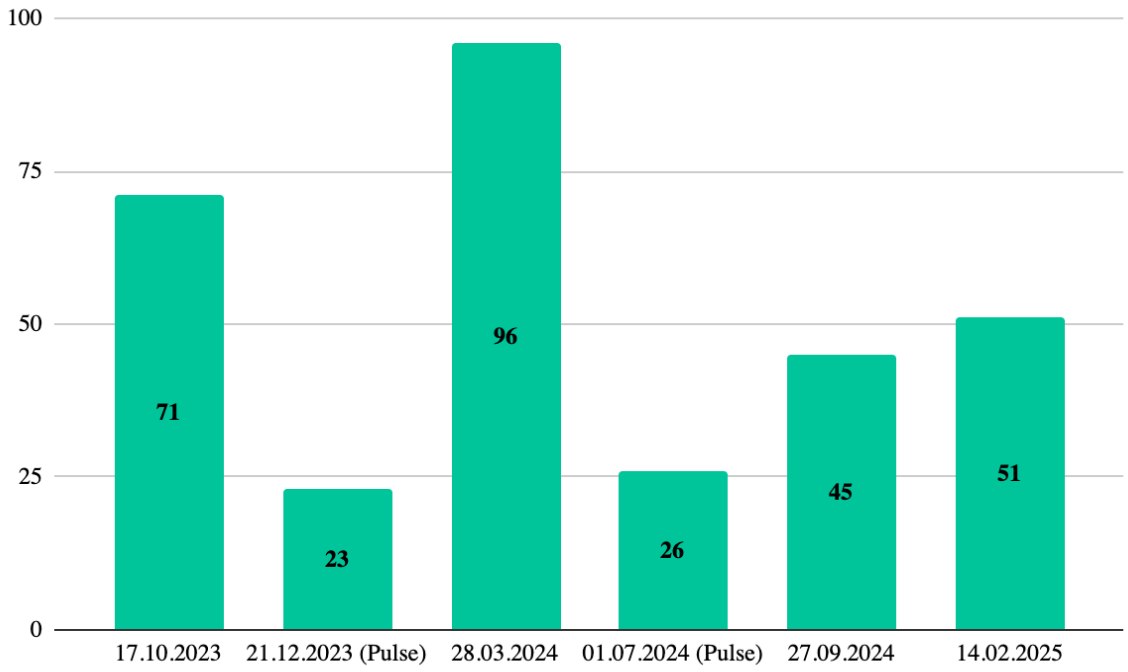

Figure 4. The number of questions in the engagement surveys.

While some questions are present in every survey there are also some which may appear only once or twice. For the purposes of this thesis, only questions that appeared in at least four out of six engagement surveys and were not missing on two consecutive surveys were included. Including questions that appear only once or twice would make it difficult to draw meaningful conclusions as there wouldn't be enough data points for analysis. The process for selecting questions is further described in chapter 4. Below are the 19 rating questions recommended by Culture Amp, along with their corresponding categories [20]:

1. "I am proud to work for [Company]" - (Engagement)
2. "I would recommend [Company] as a great place to work" - (Engagement)
3. "I rarely think about looking for a job at another company" - (Engagement)
4. "I see myself still working at [company] in two years' time" - (Engagement)
5. "[Company] motivates me to go beyond what I would in a similar role elsewhere" - (Engagement)
6. "The leaders at [company] keep people informed about what is happening" - (Leadership)
7. "My manager is a great role model for employees" - (Leadership)
8. "The leaders at [Company] have communicated a vision that motivates me" - (Leadership)
9. "I have access to the things I need to do my job well" - (Enablement)
10. "I have access to the learning and development I need to do my job well" - (Enablement)
11. "Most of the systems and processes here support us getting our work done effectively" - (Enablement)
12. "I know what I need to do to be successful in my role" - (Alignment)
13. "I receive appropriate recognition when I do good work" - (Alignment)
14. "Day-to-day decisions here demonstrate that quality and improvement are top priorities" - (Alignment)
15. "My manager (or someone in management) has shown a genuine interest in my career aspirations" - (Development)
16. "I believe there are good career opportunities for me at this company" - (Development)
17. "This is a great company for me to make a contribution to my development" - (Development)
18. "Our company effectively directs resources (funding, people, and effort) towards company goals" - (Company confidence)

19. “My company is in a position to really succeed over the next three years” - (Company confidence)

The survey questions use a 5-point Likert scale: ‘Strongly Disagree’, ‘Disagree’, ‘Neutral’, ‘Agree’, and ‘Strongly Agree’. After the surveys are completed, a report is generated that categorizes the responses into three groups: ‘Favorable’, ‘Neutral’, and ‘Unfavorable’. Responses of ‘Agree’ and ‘Strongly Agree’ are considered ‘Favorable’, ‘Neutral’ responses are categorized as ‘Neutral’, and ‘Disagree’ and ‘Strongly Disagree’ responses are classified as ‘Unfavorable’. An example of the results for a question is provided below in Figure 5.

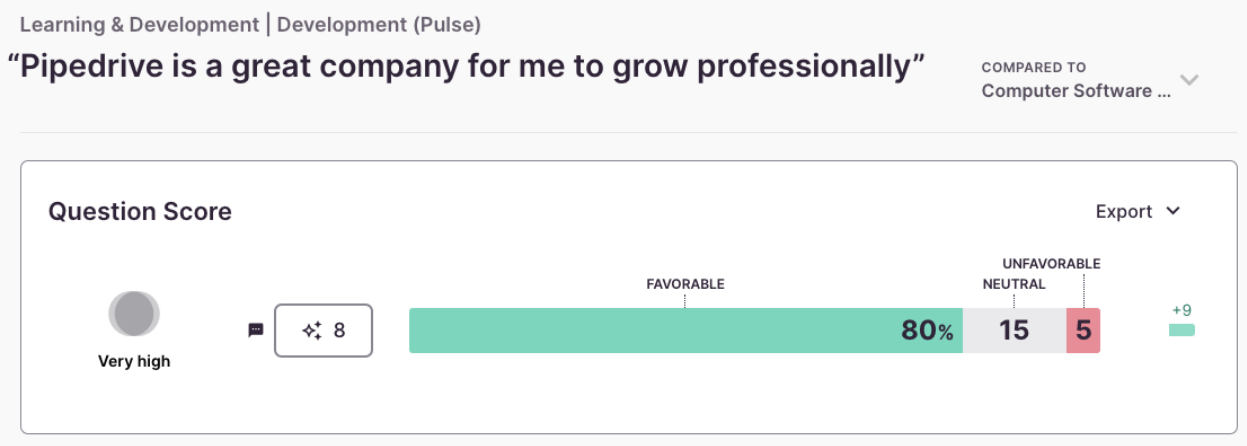

Figure 5. An example engagement survey question result.

Figure 5 shows the response for one of the questions in the engagement survey. It illustrates the correlation with engagement (marked as ‘Very High’), the number of comments, and the percentages of positive (green), neutral (gray), and negative (red) responses. Along with the rating questions, each survey included optional free-text questions that allowed employees to provide additional comments. However, these responses were not used in this thesis due to the limited number of replies, which made it difficult to draw meaningful qualitative insights.

3.2.2 Developer Experience Survey

A developer experience (DX) survey [21] is used to understand how developers perceive their work environment, covering aspects like tooling, processes, collaboration, and productivity. It’s designed to gather feedback across key dimensions such as speed, effectiveness, quality, and business impact. These surveys help teams identify friction points in the developer's day-to-day

work, for example poor documentation, slow processes, or disruptions to focus. Teams can then use that data to make informed improvements. It's a structured way to listen to developers and optimize their experience to boost productivity within the team and company.

There are many different platforms that can be used to conduct DX surveys, but the one used in Pipedrive currently is the GetDX [4]. Well-known companies using the GetDX platform include Dropbox, Ebay and others [22]. The DX survey has been conducted quarterly on the GetDX platform at Pipedrive for the past two years, for a total of eight times. This survey targets the engineering division, which has approximately 200 of Pipedrive's 800+ employees. The average response rate has consistently been around 90%, with the most recent survey reaching a 95% participation rate, as shown in Figure 6 below.

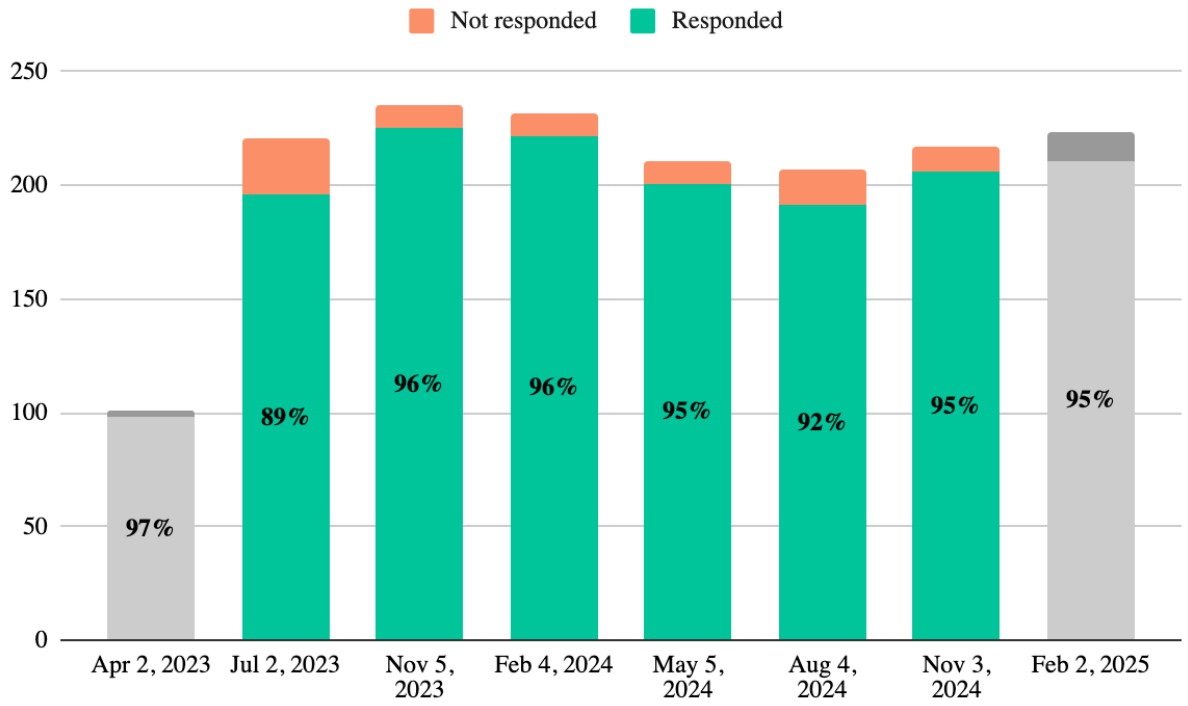

Figure 6. DX survey response rates.

Since the engagement survey has been conducted six times, this thesis uses data from the second to the second-most recent survey (highlighted in green and orange in Figure 6) when finding the relationships with the engagement survey. These time periods were selected because they align

chronologically: the DX survey is conducted first, followed by the engagement survey a month or two later. The engagement survey is described in the previous chapter.

The DX survey consists of a set of questions designed to assess the actionable areas of software delivery that are most predictive of outcomes, such as development speed, ease, and quality [21]. These questions are also referred to as "drivers," as that is the term used on the platform. The questions (“drivers”) have been selected by GetDX as the ones that affect key actionable areas, with each question impacting these areas to varying degrees [23]. However, it is not mandatory to use all 20 questions and the company using the survey can pick the ones they want. As a result, the set of questions used by Pipedrive varies slightly across surveys. Some questions appear in every survey, while others are used only once.
Each of those questions are phrased as: “Please rate the following area of your work,” followed by the specific area. Respondents rate each item using a Likert-type scale with the following options: ‘Very Good’, ‘Good’, ‘So-so’, ‘Bad’, and ‘Very Bad’. Across the eight DX surveys conducted at Pipedrive so far, all 20 questions have been used. The 20 question recommended by platform were the following:

1. Having clear team goals and direction (Clear direction)
2. Code review process (Code review)
3. Having code that is easy to understand and work with (Codebase experience)
4. Feeling connected to your colleagues (Connectedness)
5. Collaboration between different teams (Cross-team collaboration)
6. Uninterrupted time for deep work (Deep work)
7. Having good internal documentation (Documentation)
8. Ease of deploying changes (Ease of release)
9. Having efficient team processes (Efficient processes)
10. Feeling empowered to experiment and take risks (Experimentation)
11. Process for dealing with errors or incidents (Incident response)
12. Balance of tech debt versus feature work on your team (Managing tech debt)
13. Experience of being on-call (On-call experience)
14. Ease of debugging production systems (Production debugging)
15. Having realistic project timelines (Realistic timelines)

16. Speed and stability of automated tests (Test efficiency)
17. Learning and improving as a team (eg. retrospectives, post-mortems) (Learning culture)
18. Utilizing feedback from customers/users (Leveraging user feedback)
19. Having a say on team roadmap/priorities (Say on roadmap/priorities)
20. Having clear scope and requirements for tasks (Requirements quality)

Each question's responses are grouped by the platform into three categories: ‘Good’, ‘So-so’, and ‘Bad’. Specifically, responses of ‘Very Good’ and ‘Good’ are grouped as ‘Good’; ‘So-so’ responses remain in the ‘So-so’ category; and ‘Bad’ and ‘Very Bad’ are classified as ‘Bad’. The value for each category represents the percentage of the respondents that rated it in that area. Each question is then assigned a favorable score, which represents the percentage of responses that fall into the ‘Good’ category. This score can be used to assess the current state of that question by comparing it to industry benchmarks (e.g., the 50th, 75th, or 90th percentile of all companies that have responded to that question) or to previous results. An example of a question result can be seen in the Figure 7 below.

Figure 7. An example result of a DX survey question.

Respondents can also select multiple areas they would like to see improved in the survey. Therefore, for each area, in addition to the percentages of 'Good', 'So-so', and 'Bad' responses, there is also the percentage of respondents who selected that area for improvement. An example of this is shown for the whole engineering division in Figure 8 below.

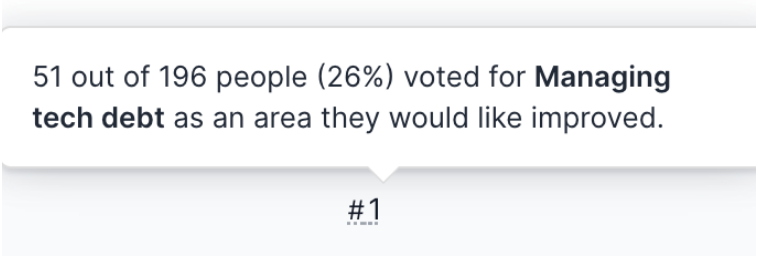

Figure 8. An example of the percentage of votes for improving an area.

Based on the votes and positive scores, each team selects key areas to focus on for improvement. This is done by identifying:

- The top 3 areas with the lowest positive scores.
- The top 3 areas with the highest votes percentage.

Each team will select at least 3 and at most 6 drivers, depending on overlap between the two groups. If all top drivers are identical across both criteria, only 3 will be selected. If there is no overlap, up to 6 may be chosen. This is how teams at Pipedrive determine which DX areas to focus on for improvement. However, due to the lack of team-level data access, the analysis is conducted at the department level, though the approach remains the same. Additionally, there may be more areas identified for improvement from this selection. To evaluate this, the P75 (75th percentile) is used, which indicates how an area compares to the value below which 75% of companies fall. If an area's value is above the P75, it is considered good. If it's below the P75, it suggests room for improvement and will be prioritized for further analysis, with the goal of identifying a relationship between that area and engagement question(s). The reasoning is that if there is a relationship then improving this area could also lead to an increase in overall engagement.

In addition to the main questions described above, some survey areas include follow-up questions designed to provide more detailed insights into potential weak points. For example, an additional question related to the area of "Uninterrupted time for deep work" is shown in Figure 9 below. For these questions, a target range is provided, and the score represents the percentage of respondents falling within that range.

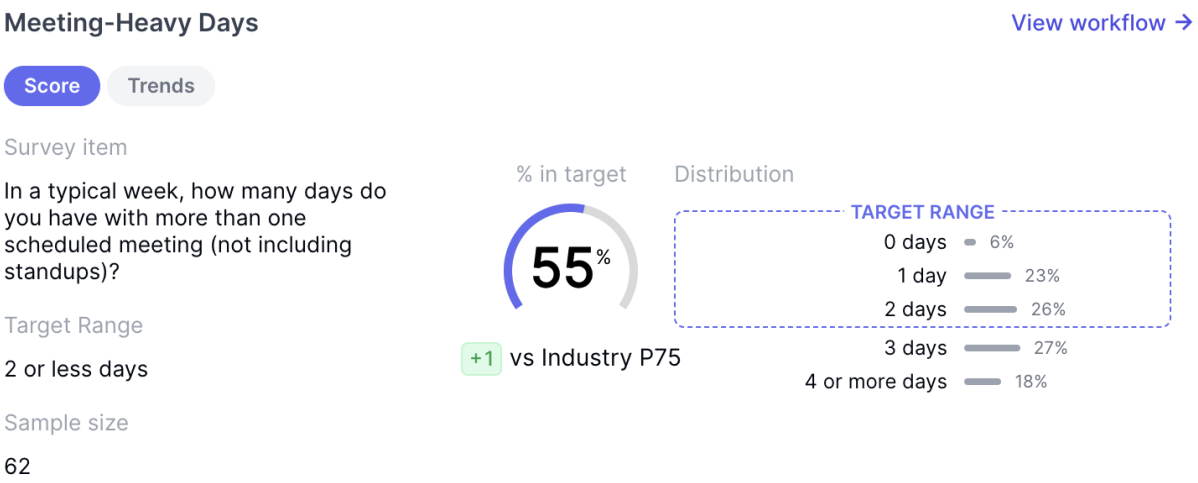

Figure 9. An example of an additional question in the DX survey.

However, these additional questions are not used in this thesis, as they are only present for some areas and vary between surveys, some appearing only once or twice. Similarly, while each question offers the option to leave comments, these are also not used due to the limited number of responses, which would make it difficult to draw meaningful qualitative insights.

4. Method

The purpose of this chapter is to explain how data from Pipedrive's surveys is used to find relationships between DX and engagement questions. The chapter describes how the data is collected, preprocessed, and analyzed to discover patterns and relationships. By the end of the thesis, the goal is to answer the research questions through the method explained in this chapter. The research question defined earlier were:

- **RQ1**: What is the relationship between improvements in developer experience (DX), as measured through DX-specific surveys, and the company-wide engagement survey results conducted via Culture Amp?
- **RQ2**: Which specific drivers of developer experience have had the strongest relationship with key metrics in the Culture Amp engagement survey?

4.1 Data Collection

Before any data could be analyzed, the first step was to collect it. Data collection can be done through various methods, such as surveys, experiments, and databases [24]. In this thesis, the data comes from surveys that were conducted over several years at Pipedrive. The two different types of surveys are described in more detail in the previous chapter. These surveys were conducted on two different platforms, so their extraction is handled slightly differently.

The DX survey has been conducted eight times so far, with a report produced for each one. Data from all the surveys was collected into an Excel spreadsheet to allow for further analysis. Before this step, it was necessary to decide which subgroups the data would be extracted for. Since some of the data extraction involved manual work, identifying the relevant subgroups in advance helped reduce unnecessary effort.

Due to limited access to data for the engagement survey, it was not possible to extract results by team. Therefore, for both surveys, subgroups were defined by department instead. As shown in Table 1, some departments were created or merged during the time the surveys were conducted.

Table 1. Engineering departments over the surveys conducted.

Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Engineering Platform	Engineering Platform	Engineering Platform	Engineering Platform	Engineering Platform	Engineering Platform
Engineering Ecosystems	Engineering Ecosystems				
Engineering Sales CRM	Engineering Sales CRM	Engineering Sales CRM	Engineering Sales CRM	Engineering Sales CRM	Engineering Sales CRM
		Engineering Excellence	Engineering Excellence	Engineering Excellence	Engineering Excellence
Infrastructure Operations	Infrastructure Operations	Infrastructure Operations	Infrastructure Operations	Infrastructure Operations	Infrastructure Operations
Data Engineering	Data Engineering	Data Engineering	Global Business Solutions	Global Business Solutions	Global Business Solutions
Engineering Business Solutions	Engineering Business Solutions	Engineering Business Solutions			

To maximize the use of the existing data, all department responses from each survey were included in the regression analysis. However, when doing correlation analysis by department, only departments that appeared in every survey were used. This approach made sure that there was enough consistent data to support more reliable conclusions.

The DX survey data was extracted in two ways. Positive scores were possible to extract using a database query, which then could be exported as a CSV file. However, data on neutral and negative scores and votes for improvement had to be collected manually, since there was no option to query nor export them. This also had to be done separately for each subgroup, which made the process time-consuming. Since the DX survey was conducted two more times than the engagement survey, all responses were collected, but only the six surveys that aligned with engagement survey dates were used for finding relationships. The surveys chosen for analysis are shown in Figure 6 in chapter 3.2.2.

The final spreadsheet of DX data had seven columns: the focus area, subgroup, date, percentages of good, neutral, and bad scores, and the percentage of votes for improvement. A few rows from this spreadsheet is shown in Table 2 below.

Table 2. An example of the data in the extracted DX spreadsheet.

Driver	Target	Date	Good	So-so	Bad	Votes as an area to improve
Codebase experience	Engineering Ecosystems	02.04.2023	56%	41%	4%	4%
Production debugging	Infrastructure Operations	02.07.2023	63%	37%	0%	17%
Test efficiency	Infrastructure Operations	02.07.2023	70%	30%	0%	13%
Realistic timelines	Engineering Business Solutions	04.02.2024	100%	0%	0%	0%
Realistic timelines	Engineering Excellence	04.02.2024	100%	0%	0%	0%
Test efficiency	Data Engineering	04.02.2024	82%	18%	0%	9%

The engagement survey had been conducted six times. For each survey, a report was generated containing all data that could be exported as a CSV file. Before exporting, the data had to be filtered within the platform based on the subgroup for which the results were needed, which in this case was by engineering departments. The relevant data was extracted from the exported CSV files into a single Excel spreadsheet. Extracting engagement data was significantly faster than for the DX survey data. An example of a few rows from the spreadsheet is shown in Table 3 below.

Table 3. An example of the data in the extracted engagement survey spreadsheet.

Question	Category	Date	Target	Favorable Score	Neutral Score	Unfavorable Score	Correlation Strength	Question's Correlation with Engagement
My manager, or someone else, has communicated some clear actions based on previous survey results	Action	01.07.2024	Infrastructure Operations	96%	0%	4%	Very Strong	0.504
I would recommend Pipedrive as a great place to work	Engagement	17.10.2023	Engineering Sales CRM	84%	13%	3%	n/a	n/a

The senior leaders at Pipedrive demonstrate that people are important to the company's success	Leadership	01.07.2024	Engineering Excellence	37%	42%	21%	Strong	0.419387755
We hold ourselves and our team members accountable for results	Teamwork & Ownership	28.03.2024	Engineering Business Solutions	90%	5%	5%	Low	0.277010738
Pipedrive is a great company for me to grow professionally	Learning & Development	28.03.2024	Engineering Business Solutions	78%	17%	5%	Extreme	0.70513757

In the resulting dataset, each row contained the question, the category it belonged to, the date of the survey, the subgroup, and the percentages of favorable, neutral, and unfavorable responses. Also the correlation with engagement was marked both by its correlation strength and the correlation itself. If a question's correlation with engagement was labeled as 'n/a', it meant the question was already part of the engagement category, making the correlation not applicable (or always 1.0). The rest of the engagement question correlation strength categories were mentioned in chapter 3.2.1.

After the extraction, the resulting spreadsheets contained data from eight DX surveys, covering 20 unique questions, and from six engagement surveys, which included 107 distinct questions. This data was then further preprocessed to ensure its suitability for detailed statistical analysis to answer the research questions. The preprocessing steps are described in the next chapter.

4.2 Data Preprocessing

After data extraction, the next necessary step was to prepare the data for analysis. Data preprocessing can involve several tasks, including converting raw data into a usable format, filtering out irrelevant data, and addressing issues such as missing values and outliers [25]. The purpose of these steps is to make the data usable for analysis. While multiple methods exist for handling these issues, the specific approaches used in this thesis are described below.

In the DX survey, a total of 20 areas, referred to as "drivers," were measured across eight surveys conducted over a period of 2 to 3 years. However, not all drivers were measured in every survey, resulting in missing data for some drivers. To address this, the first step was to filter out any drivers

that were not measured during any of the six surveys that aligned with engagement surveys. These six surveys were selected as the second to the second-most recent surveys, as shown in Figure 6. Missing data in the first and last surveys were disregarded, as these surveys were not used in the correlation or regression analysis. But in the 6 surveys between the drivers consistently available were selected. The drivers and their presence over surveys is shown in Figure 10 below.

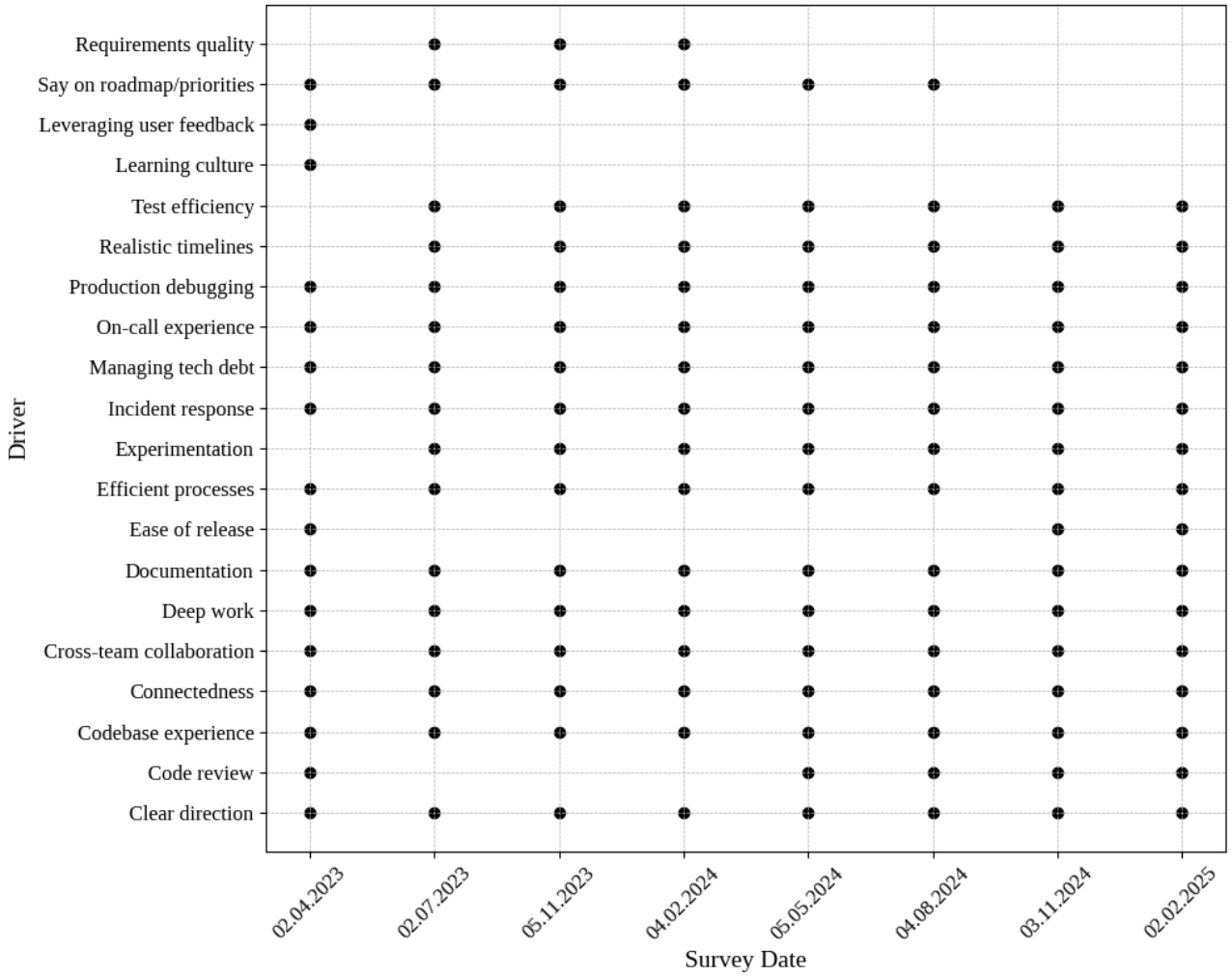

Figure 10. Presence of DX drivers across surveys.

After this filtering step, 14 drivers remained available for analysis. Although missing value imputation was considered for the drivers that were missing in one or two surveys, this approach was not used due to the already limited amount of data and using those drivers would have further decreased the reliability.

After filtering out suitable DX survey drivers, outliers caused by rounding errors were corrected. In the reports, the percentages for "good," "so-so," and "bad" responses were rounded to whole numbers, which sometimes resulted in their sum being greater or less than 100%, which is illogical. To fix this, the values were adjusted to better reflect their original relative proportions. For example, if the original report showed "good" as 34%, "so-so" as 34%, and "bad" as 33%, summing to 101%, these were proportionally adjusted to values 33.67%, 33.67%, and 32.66%, respectively, so the total equals closer to 100% while maintaining their original relative sizes.

The preprocessing of the engagement survey data followed a similar approach to that used for the DX survey data. First, only questions that appeared in at least four of the six employee engagement surveys were included. Additionally, if a question was missing from two surveys, it could not be missing from two consecutive ones, as this would make it difficult to identify trends. The threshold of being present in at least four surveys was chosen because the two Pulse surveys contained only about 25 questions, and a higher threshold would have excluded too many engagement survey questions. Most questions that were missing from one or two surveys were absent from the Pulse surveys. However, there was one exception: a question that was missing from one of the four standard engagement surveys but present in the Pulse surveys. Similar to the preprocessing of DX data, the imputation of missing values was considered for the engagement data but was abandoned due to the reliability issues. Also, the issue of rounded percentages was apparent in this engagement data and was addressed the same way by adjusting them to relative values. The total number of engagement questions that remained for further analysis was 35. Figure 11 below illustrates the presence of the selected questions across the engagement surveys. Only the selected questions are shown in the figure to keep the visualization clear and interpretable. Including all questions would have added unnecessary complexity without providing additional insight.

Figure 11. Presence of engagement survey questions across surveys.

The final step was to deal with outliers, in columns they existed. Outliers are individual data points that deviate significantly from the rest of the values in the dataset [26]. These outliers were removed using a Z-score with a threshold of 3. The Z-score is a measure that indicates how far away a data point is from the mean of the all [27]. Data points with a Z-score greater than 3 or less than -3 were considered outliers and removed. After removing the outliers, the data from the DX

and Culture Amp surveys was then ready for further analysis. These preprocessing steps ensured the data could be used in detailed statistical analysis to answer the research questions.

4.3 Correlation Analysis

In the previous chapter the process of preparing the data for analysis was described. From this chapter the analysis methods used in identifying relationships are described in more detail. The first method used in the analysis was the correlation analysis on the department level. The department used as an example throughout the rest of the methodology chapter is 'Engineering Sales CRM,' and all figures and results are based on the analysis of this department. The same approach was applied to the other departments that were present throughout all the surveys.

Correlation analysis is a method used to measure the strength of the relationship between two variables, if one even exists [28]. In this thesis, it is used to explore how responses from the DX survey are related to responses from the engagement survey. There are several correlation methods, with Pearson's being one of the most commonly used. However, to use Pearson's correlation, the data must first meet certain requirements [29]:

- **Linearity**: This assumes that the relationship between the two variables is linear. In other words, as one variable changes, the other should also change a fixed amount.
- **Normality:** The data for each variable should roughly follow a normal distribution. In other words, the majority of data points are clustered around the average and fewer at the extremes.
- **Absence of outliers:** This means that there should be no outliers in the dataset (values that differ significantly from the rest), as outliers can drastically impact the calculation of the correlation coefficient.
- **Independence:** Each data point should have a corresponding pair and the value of one variable in the pair should be independent of the other's value.
- **Scale of Measurement:** The data should be measured on an interval or ratio scale. In other words, the values should have equal gaps between them, and zero should indicate the absence of the feature being measured.

When these assumptions are met, Pearson's correlation can accurately measure the relationship between two variables. However, in this thesis, the data did not meet all the criteria outlined above, making Pearson's correlation unsuitable for the following reasons:

- **Non-Linear Relationships:** The relationship between some variables might not be linear, as other factors could also be influencing the engagement survey responses.
- **Non-Normal Distribution:** In many cases, the data was not normally distributed. The Shapiro-Wilk Test [30] was used to check if the data was normally distributed and to see if it violated this assumption.

Since the data did not meet the assumptions for Pearson's correlation, Spearman's correlation was used instead. Spearman's correlation [31] is suitable because it doesn't require the data to be normally distributed or have a linear relationship. It measures the strength of the relationship between two variables in a monotonic way. This means that as one variable increases, the other either consistently increases or decreases, but without necessarily always the same amount. Therefore, when plotted the graph will likely not be a straight line.

The correlation analysis was conducted for each individual department to see what relationship existed between each of the subgroups and whether there were some that were present in each department. In total there were 4 correlation matrices between DX and engagement survey questions, one for each department, and a total of 490 different pairs per correlation matrix, as there were 14 suitable DX questions and 35 suitable engagement survey questions.

4.4 Regression Analysis

After performing the correlation analysis, the next step in understanding the relationship between DX and engagement was to conduct a regression analysis. Regression analysis is a statistical method used to understand strength and the direction of the relationship between a dependent variable and one or more independent variables [32]. In this thesis, both simple and multivariate regression analyses were considered.

For multivariate regression analysis, it's important to have enough data, ideally there should be at least 10-15 data points for each independent variable [33]. In this case, the data at the department

level contained just under 30 data points, which was not enough to use multivariate regression with multiple independent variables. Therefore, the analysis was conducted using a single independent variable at a time.

While regression analysis shares some similarities with correlation analysis in detecting the relationship between the two variables, it also enables the prediction of a variable based on another [32]. Since the relationship between DX factors and engagement might not always be linear, both simple linear regression and non-linear regression models were used. The non-linear model was selected because it could be assumed that the relationship between DX factors and engagement was more complex. For example, some DX factors might influence engagement in ways that are not directly proportional, which non-linear models can better capture.

In this analysis, each DX variable was compared to each engagement survey variable. The R^2 value [34] was used to evaluate how well each model fit the data, as it shows how closely the data points match the predicted values. The value ranges from 0 to 1 and is calculated by comparing the explained variation to the total variation in the data. A higher R^2 indicates a better fit of the model by having a bigger proportion of the variation in the dependent variable explained by the independent variables. For example, an R^2 of 0.8 means the model explains 80% of the variation, leaving only 20% unexplained. Only pairs with an R^2 value above 0.4 were considered for further analysis. An R^2 value above 0.4 indicates a positive relationship between the DX variable and the engagement variables, while not limiting the potential results too much as could have happened with a higher value. Positive relationships meaning that as one variable increases, the other tends to increase as well.

Additionally, scatter plots with regression lines were created for each pair of variables. These plots helped visually confirm whether a relationship existed and whether it aligned with the findings from the regression analysis. The scatter plots provided insight into the strength and nature of the relationship between the DX and engagement variables, helping to validate the regression results.

4.5 Principal Component Analysis

Principal Component Analysis (PCA) is a dimensionality reduction technique that simplifies data with many variables by transforming the original variables into a smaller set of uncorrelated

components [35]. These components are ordered based on the variance they explain in the dataset. PCA is especially useful for high-dimensional data because it simplifies the analysis by reducing the number of dimensions while retaining the most significant information. In this analysis, PCA was used as an alternative to regression and correlation analysis to identify patterns in the DX questions that might impact the engagement questions.

Given the limited dataset, multivariate regression analysis wasn't reasonable to use, as discussed in the previous chapter. Although PCA ideally requires larger datasets than were available in the context of this thesis to draw reliable results [35], it was still used in this thesis. This was done for the small possibility that something could be discovered and to provide a better guide for future analysis when there is more data. The PCA was done with the restriction that only the most significant principal component would be used for finding any relationships. This component was then used in a regression analysis to explore whether multiple DX questions could together strongly influence an engagement question.

To determine how many components were needed, the first task was to explain at least 90% of the variance, a standard threshold that typically retains the majority of the data's variability. With this dataset, six components were required to reach this threshold, meaning they captured the majority of the data's variance.

The variance explained by each component was then examined. The first principal component (PC1) accounts for the largest portion of the variance. If PC1 explained less than 30% of the variance, the approach would have been reconsidered, as it would lack sufficient explanatory power due to the limited data. Next, the relationships between PC1 and the engagement questions were analyzed. As with the individual questions, regression analysis was used to calculate the R^2 value between PC1 and the engagement questions. A scatter plot was also created to visually explore the strength of these relationships that had a R^2 value above 0.5.

5. Results

In the previous chapter, the data processing steps and analysis were explained, including the methods used, the reasoning behind them, and their benefits. This chapter aims to use the results from those analyses to answer the research questions. To illustrate the process, the analysis is demonstrated using the "Engineering Sales CRM" department's results as an example for the first research question, however, the same approach was applied to other departments as well. At the end of the answer, a conclusion summarizing the overall results across all departments is given.

- **RQ1:** What is the relationship between improvements in developer experience (DX), as measured through DX-specific surveys, and the company-wide engagement survey results conducted via Culture Amp?

To answer the first research question, the first step was to identify what improvements had been made in the selected department. To do that, the lowest-scoring areas in developer experience in the first survey were identified for the given department, which in this case was “Engineering Sales CRM,” and it was determined which of those areas had improved over time. The areas with poor results for this department in the first survey totaled to eight, all of which had less than 60% positive responses and were below the industry’s 75th percentile of all companies that have responded to that question. These areas clearly had room for improvement compared to others that were already performing well. The eight areas were: documentation, codebase experience, managing tech debt, test efficiency, realistic timelines, cross-team collaboration, production debugging, and experimentation. Out of these eight, only two areas did not improve over time: production debugging and experimentation. The next step was to analyze whether these areas that showed improvement had any correlation with the engagement survey questions. To do this, the correlation matrix between the engagement survey and the developer experience survey was examined. The correlation matrix is shown in Figure 12 below.

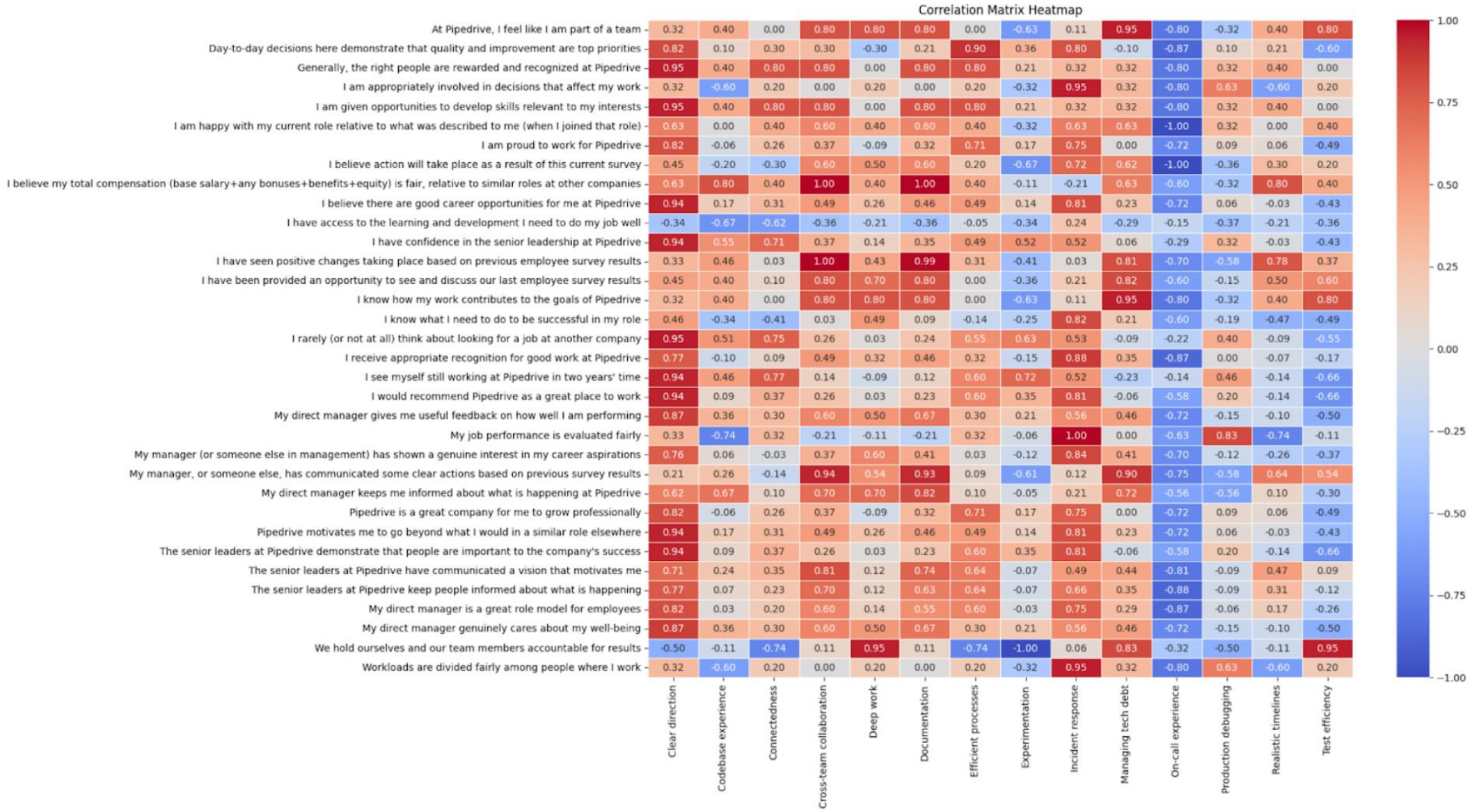

Figure 12. Correlation matrix heatmap between engagement and DX survey questions.

From the correlation matrix, the correlations between the developer experience questions and the engagement questions could be identified. On the x-axis are the DX questions and on the y-axis the engagement survey questions. The criteria for determining which correlations were strong enough to suggest potential relationships were set at 0.8 or above, due to the limited data. The results showed that each of the areas that improved in the developer experience had at least one question from the engagement survey that correlated with it. The findings were that documentation had 9 correlations, codebase experience had 1, managing tech debt had 6, test efficiency had 3, realistic timelines had 1, and cross-team collaboration had 9 correlations above 0.8 with engagement survey questions. The 0.8 correlation threshold ensured that the compared survey questions had no more than one or two trend mismatches over time. Sometimes, a trend mismatch could be explained by other factors. For example, if a DX area improved while the related engagement question got worse, it would suggest no clear link. However, if the DX driver remained a highly voted improvement area, it might indicate that the perceived improvement was limited, justifying the trend mismatch. This is why it was also important to examine the neutral, and negative scores and the votes for improvement trends over time between these strongly correlated questions. One of the trends of positive responses between a DX question and an engagement question is plotted in the Figure 13 below.

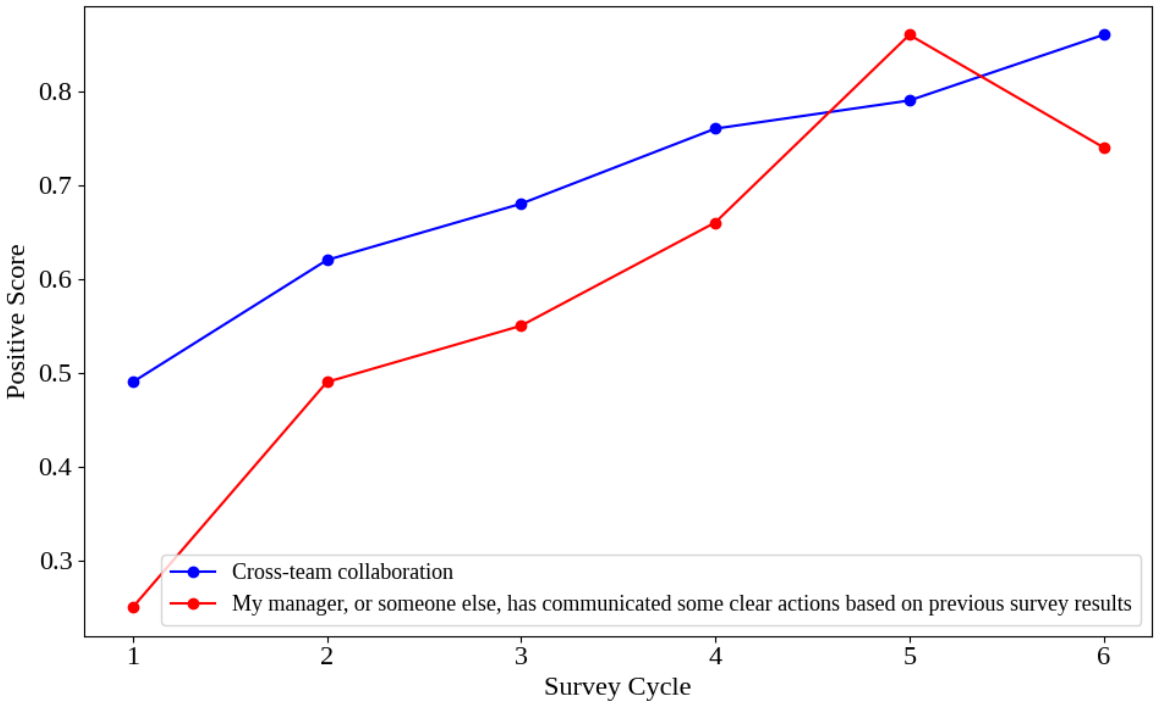

Figure 13. Line graph of a positively correlated DX and engagement question pair.

In the figure, the positive trends of two questions can be seen: one from the developer experience survey and another from the engagement survey. A mismatch between the two questions and their trends occurs on one occasion when the engagement question had a strong increase followed by a decrease, while the developer experience survey showed a more steady positive increase. These mismatches show that the relationship between DX and engagement can be more complicated, and while they are linked, other factors could be affecting the results as well. When examining the additional data on these questions, particularly the votes for improvement for the DX question, it was observed that when the engagement survey increased significantly, the votes for improvement on the DX question decreased by half. This suggests a stronger positive increase than the positive response rate might indicate, potentially explaining the mismatch. However, this could also be due to other factors, and it is possible that these two questions do not actually have a strong relationship. Therefore, it is also important to examine data from the rest of the departments to establish a stronger relationship between the two questions. In the case of 'Engineering Sales CRM,' out of the six areas that were improved and had a correlation of 0.8 or above with engagement questions, two areas, cross-team collaboration and test efficiency, were the ones that showed a correlation with engagement questions where the relationship between questions was above 0.7 in other departments as well. For the other departments, a similar analysis was conducted to determine whether improvements in DX had a related question in the engagement survey. For the four departments that were present in all the surveys, all of the areas that were scored low in the first survey and managed to increase their score had at least one question from the engagement survey with a correlation of 0.8 or above. This indicates that the answer to the first research question (RQ1) is that developer experience improvements did have relationships with engagement questions, but since improvement varied across departments, relationships also varied across departments indicating that there is not one clear relationship between DX and engagement surveys but instead individual ones per subgroup.

- **RQ2:** Which specific drivers of developer experience have had the strongest relationship with key metrics in the Culture Amp engagement survey?

To answer the second research question, the first step was to examine the correlation matrices between the developer experience survey and the engagement survey for each department, identifying relationships and strong correlations present in each department’s correlation matrix.

Unlike the first research question, which focused only on the areas that improved over time, this analysis looked at all areas of DX. The correlation matrices used for this analysis are the same as those shown for the "Engineering Sales CRM" department in Figure 12, but now all areas of developer experience are considered. This approach allowed for the identification of relationships present across all departments. In this case, the threshold for strong correlations was set at 0.7, as more data points were available, allowing the threshold to be lowered compared to the previous analysis. Setting it too high could have restricted the results too much. The total number of pairs that had strong correlation in each department was 9 and the pairs were following (on the left is the DX area and right the engagement question):

1. Connectedness - Generally, the right people are rewarded and recognized at Pipedrive
2. Connectedness - I have confidence in the senior leadership at Pipedrive
3. Cross-team collaboration - At Pipedrive, I feel like I am part of a team
4. Cross-team collaboration - My direct manager keeps me informed about what is happening at Pipedrive
5. Experimentation - I see myself still working at Pipedrive in two years' time
6. Incident response - I am appropriately involved in decisions that affect my work
7. Incident response - Day-to-day decisions here demonstrate that quality and improvement are top priorities
8. Incident response - Pipedrive is a great company for me to grow professionally
9. Test efficiency - At Pipedrive, I feel like I am part of a team

Once the most interesting pairs by correlation were identified the next step was to filter out pairs where the engagement survey question wasn't strongly correlated to engagement overall (Correlation strength column which can be seen in Table 3). In the engagement survey, only five questions were directly tied to engagement. However, all questions were given correlations from the Culture Amp platform to these five to identify indirect links. For each engagement question, the mean correlations to the engagement category were calculated separately, as correlations can vary between over time. To ensure only meaningful engagement questions were considered, filtering criteria were applied:

- A question needed a mean correlation to engagement of at least 0.4 across surveys (considered strong according to Culture Amp).

- Additionally, the correlation for that question should not fall below 0.3 in any individual survey.

This filtering resulted in the removal of one pair due to the engagement question being lowly correlated with the engagement category. This pair was:

- Cross-team collaboration - My direct manager keeps me informed about what is happening at Pipedrive

The most interesting pair out of those nine pairs was “Experimentation - I see myself still working at Pipedrive in two years' time” as the engagement question is one of the 5 questions in the engagement category and the fact that experimentation is directly linked with that question means this DX area is also directly correlated with employee engagement.

Once the most interesting pairs were identified the next thing to look for was the results from the regression analysis for each pair which was plotted as a scatter plot with a regression line and the R^2 calculated. Out of the 8 remaining pairs two pairs remained which had a R^2 value above 0.4 and showed somewhat relation that the higher the developer experience question the higher the corresponding engagement question. The remaining pairs were:

- Experimentation - I see myself still working at Pipedrive in two years' time
- Incident response - Day-to-day decisions here demonstrate that quality and improvement are top priorities

While the non-linear regression increased the R^2 values it didn’t provide any hidden relationships and therefore is not further explained. The corresponding scatter plot with linear regression line and R^2 value for the first pair can be seen below in Figure 14.

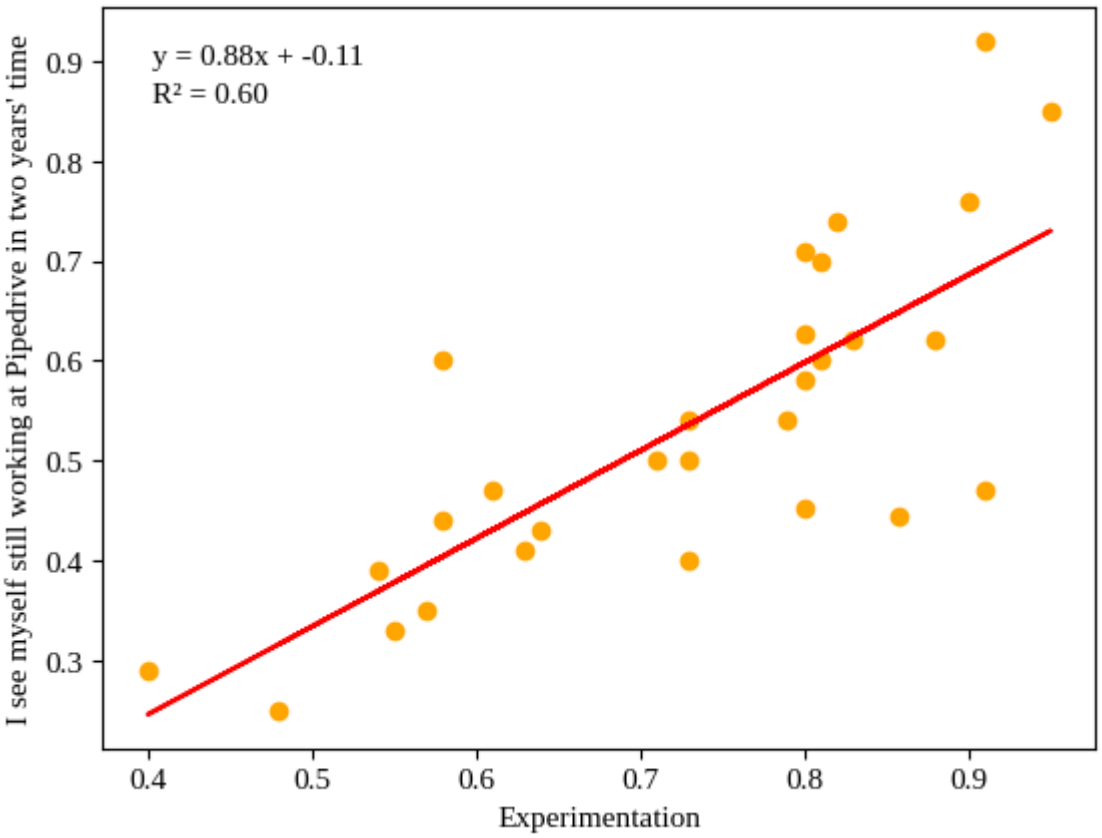

Figure 14. Scatter plot of the strongest relationship.

After analyzing the data, the most impactful DX areas appear to be experimentation and incident response. For teams or departments with low ratings in these areas, focusing on improvement could help enhance their group engagement. Experimentation stands out as an opportunity for improvement in three departments, based on the latest DX survey results where it was rated below the industry's 75th percentile. However, it may not be the primary focus area, as none of these departments ranked experimentation among their top three lowest-rated areas and only one department identified it as one of the top three areas to improve. On the other hand, incident response, received positive ratings across all departments in the most recent survey. As a result, it may not be a priority to focus on for improvement in any department.

The PCA analysis was primarily conducted for experimentation purposes, as the limited data available did not allow for highly reliable results. The first principal component (PC1) explained 45.3% of the DX data variance, meaning it accounted for nearly half of the data's variability. The contribution of each DX area to PC1 was similarly high for five areas, all with significance above

0.3, indicating they moved in similar patterns. The 5 areas with the highest contributions to PC1 were:

1. Experimentation: 0.3391
2. Managing Tech Debt: 0.3340
3. Codebase Experience: 0.3235
4. Realistic Timelines: 0.3171
5. Clear Direction: 0.3106

Regression analysis was done between the PC1 component and the engagement questions to identify potential relationships. Nine relationships had an R^2 value above 0.5 and were further explored. As described in the answer to the second research question, the same filtering of pairs was based on the importance of the engagement questions correlated with PC1. This process left five key pairs that the PC1 was related to:

1. I see myself still working at Pipedrive in two years' time - R^2=0.61
2. The senior leaders at Pipedrive have communicated a vision that motivates me - R^2=0.58
3. I would recommend Pipedrive as a great place to work - R^2=0.56
4. Pipedrive motivates me to go beyond what I would in a similar role elsewhere- R^2=0.53
5. I rarely (or not at all) think about looking for a job at another company - R^2=0.51

Four of the questions are directly related to engagement, which suggests that the top features in PC1 are strongly linked to overall employee engagement. The strongest relationship is between the engagement question, “I see myself still working at Pipedrive in two years' time” and the DX area with the highest significance is “Experimentation.” This finding is similar to the results from individual regression analysis. However, since the R^2 value remained largely unchanged, it seems that the additional features do not substantially improve the relationship. While the PCA analysis offers interesting insights into areas that could potentially enhance employee engagement, it should be interpreted with caution. The limited dataset and the risk of false results mean that these findings should not be treated as definitive.

6. Discussion

This chapter interprets the results from Chapter 5, focusing on key findings. It will also address the limitations of the study, considering factors that may have influenced the results. Finally, the chapter will reflect on the lessons learned during the research and suggest potential improvements or changes for a future study.

One of the most interesting findings was how the relationship between DX improvements and engagement varied across different departments. While some improvements in DX were strongly linked to engagement results, the strength of this relationship was different depending on the department. This is important because it shows that the link between DX and engagement isn't the same for subgroups in the company. Each department has its own pattern, influenced by things like the type of work, how teams interact, and what each department aims to achieve. For example, in the "Engineering Sales CRM" department, areas like documentation and cross-team collaboration had strong links with multiple engagement questions, but in other departments these relationships didn't exist.

The second research question aimed to find the DX areas most closely linked to engagement. Interestingly, experimentation and incident response stood out as key factors across all departments. Experimentation, especially, showed a strong correlation with the engagement survey question that was directly linked to overall engagement results. The related engagement question was "I see myself still working at Pipedrive in two years' time". This suggests that encouraging a workplace where developers have the freedom to experimentation and innovation might be connected to higher employee engagement and retention.

On the other hand, incident response stood out as an area with also consistently high correlation across all departments with another engagement survey question. But since this area is already highly rated across departments it suggests that incident response has already been well optimized and further improvements may not lead to the same level of engagement gains. Therefore, the focus could be better spent on areas like experimentation, which have more potential for future improvements and increased engagement as it had room to grow in several departments.

Another interesting, though somewhat unexpected, finding is the mismatch between some DX improvements and engagement survey results. For example, while documentation improved over time in all departments, the engagement survey responses often showed little relationship to this area or even an opposite trend in certain departments. This mismatch suggests that while DX improvements might improve internal processes, they don't always lead to higher employee engagement. This could be because either employees don't see the value in these improvements or because the documenting work itself isn't very engaging, meaning doing more documentation leads to more unfulfilling tasks and therefore to lower engagement.

The variability in these trends highlights that the relationship between DX and engagement is complex and many-sided. External factors could influence or even distort the expected impact of DX improvements on engagement. While the findings found some insights, there were also several limitations that must be acknowledged as they may have affected the reliability and generalizability of the results.

The most significant limitation was the limited amount of available data. A smaller dataset reduces statistical strength and increases the risk that observed relationships or the lack of it are due to chance rather than reflecting true underlying patterns. Additionally, the analysis relied on correlation, which cannot establish causality. This means that while some associations were identified, they should not be interpreted as one variable causing changes in another.

Another notable constraint was the access to the data. Analysis was performed at the department level due to access restrictions, even though multiple teams exist under each department. The results of individual teams were only visible to team managers, restricting deeper analysis. This limitation may have influenced potential patterns or relationships that would have been more evident at the team level. Performing the same analysis at the team level, as changes more often happen on team level, could potentially reveal stronger and more consistent relationships between DX questions and engagement questions. Despite these limitations, the study does point out some potential relationships and a framework that could be used in the future when there is more data available to repeat the process.

If this study were to be repeated, several improvements could be made to strengthen the analysis and address some of the limitations mentioned. Expanding the study to include data from all teams

in the engineering division could provide a more comprehensive understanding of the relationship between DX and engagement. Additionally, using data from a longer time period could help capture changes over time and reveal long-term trends. Incorporating qualitative data, such as comments or interviews with team leads, would allow for a deeper understanding of the context behind the quantitative findings. Gaining insight into employees' perspectives on specific DX areas, such as documentation or incident response, could offer a better understanding of why some improvements strongly correlate with engagement, while others do not.

7. Conclusion

This thesis aimed to explore the relationship between developer experience and employee engagement at Pipedrive by analyzing survey data from two sources: the engineering-focused DX surveys and the company-wide engagement surveys. The main research goals were to understand whether improvements in DX are linked to better engagement and which DX areas most relate with engagement survey results.

The findings show that while there are some links between DX areas and engagement questions, particularly at the department level, these relationships often aren't consistent or strong enough across all departments to draw broad conclusions that one area in the DX survey is strongly related with another engagement question. This suggests that the impact of DX improvements on engagement may vary by subgroup, in this case by department, with the most important developer experience areas to focus on depending on the subgroup.

From a methodological standpoint, the study mainly used correlation, regression and trend analysis, along with principal component analysis (PCA) to simplify the data and identify potential multivariate relationships. Spearman’s correlation was chosen over Pearson’s because the data was not normally distributed and was ordinal. This approach allowed for more meaningful analysis despite the small dataset and non-linear trends. While PCA helped reduce the complexity and reveal broader patterns, it didn’t uncover any hidden relationships than those found using single variable regression analysis.

The main benefit of this thesis is the creation of a repeatable framework for future analysis. By setting clear filtering criteria, data preprocessing steps, and analytical method, this study offers a structured approach that Pipedrive can use to keep evaluating the link between DX and employee engagement. As more data is collected over time, the company can apply this framework again to uncover clearer patterns.

However, the thesis has several limitations that affect its scope and applicability. The main ones are the limited number of survey rounds, the lack of access to team-level data, and the small sample sizes, which reduce the reliability of the analysis results. Additionally, survey response trends

could be influenced by external factors, such as organizational changes, team turnover, or temporary project-specific conditions, which were not considered in this analysis.

In conclusion, while this study does not establish very strong relationships between DX survey and company-wide engagement survey questions, it highlights a couple of promising areas to look into as it identifies one area for improvement across multiple departments, which has previously shown a strong correlation with employee engagement based on existing data. Rest of the results suggest that certain DX areas may be related to engagement questions at the subgroup level, and improving these areas could boost engagement in those specific teams or departments. In the future, Pipedrive could benefit from continuing tracking these relationships over time and gathering more detailed team-level data. This will help identify key factors that affect team engagement and enable better decision-making to maintain high performance and engagement.

List of References

[1] Pipedrive About page. https://www.pipedrive.com/en/about (25.04.2025).

[2] Gallup Q12 Meta-Analysis Report. *Gallup*, 2024. https://www.gallup.com/workplace/321725/gallup-q12-meta-analysis-report.aspx (25.04.2025).

[3] Culture Amp Engagement survey page. https://www.cultureamp.com/platform/engage (25.04.2025).

[4] GetDX Survey page. https://getdx.com/snapshots/ (25.04.2025).

[5] OpenAI (2025). ChatGPT 4.o. https://chat.openai.com/chat

[6] Kahn W. A. Psychological Conditions of Personal Engagement and Disengagement at Work. *The Academy of Management Journal*, 1990. https://www.jstor.org/stable/256287 (25.04.2025).

[7] Employee Engagement Guide. *Culture Amp Blog*, 2023. https://www.cultureamp.com/blog/employee-engagement-guide (25.04.2025).

[8] Harter, J. K., Schmidt, F. L., & Hayes, T. L. Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology*, 2002. https://www.researchgate.net/publication/11367971_Business-Unit-Level_Relationship_Between_Employee_Satisfaction_Employee_Engagement_and_Business_Outcomes_A_Meta-Analysis (25.04.2025).

[9] Rich B. L., Lepine J. A., Crawford E. R. Job Engagement: Antecedents and Effects on Job Performance. *Academy of Management Journal*, 2010, https://www.jstor.org/stable/25684339 (25.04.2025).

[10] Hollar, N., Kuchinka, D. G. J., & Feinberg, J. Professional development opportunities and job satisfaction: A systematic review of research. *Journal of International Management Studies*, 2022. https://www.researchgate.net/publication/361384129_PROFESSIONAL_DEVELOPMENT_OPPORTUNITIES_AND_JOB_SATISFACTION_A_SYSTEMATIC_REVIEW_OF_RESEARCH (25.04.2025).

[11] Kim, S., Lee, H., & Connerton, T. P. How psychological safety affects team performance: Mediating role of efficacy and learning behavior. *Frontiers in Psychology*, 2020. https://www.researchgate.net/publication/343197018_How_Psychological_Safety_Affects_Team_Performance_Mediating_Role_of_Efficacy_and_Learning_Behavior (25.04.2025).

[12] Alagaraja, M., & Shuck, B. Exploring organizational alignment-employee engagement linkages and impact on individual performance: A conceptual model. *Human Resource Development Review*, 2015. https://www.researchgate.net/publication/273175949_Exploring_Organizational_Alignment-Employee_Engagement_Linkages_and_Impact_on_Individual_Performance_A_Conceptual_Model (25.04.2025).

[13] Joseph, N. U. The impact of employee recognition programs on employee engagement and motivation: A case study of Southwest Airlines. 2023. https://www.researchgate.net/publication/370710004_The_Impact_of_Employee_Recognition_Programs_on_Employee_Engagement_and_Motivation_A_Case_Study_of_Southwest_Airlines (25.04.2025).

[14] Sales CRM Pipedrive announces majority investment from Vista Equity Partners. *Pipedrive Newsroom*, 2020. https://www.pipedrive.com/en/newsroom/sales-crm-pipedrive-announces-majority-investment-from-vista-equity-partners (25.04.2025).

[15] Day, B., & Williams, R. 10 Best CRM Software Of 2025. *Forbes Advisor*, 2025. https://www.forbes.com/advisor/business/software/best-crm-software/ (25.04.2025).

[16] Pipedrive reaches 100K customer milestone initiating a new growth phase. *Pipedrive Newsroom*, 2022. https://www.pipedrive.com/en/newsroom/pipedrive-reaches-100k-customer-milestone-initiating-a-new-growth-phase (25.04.2025).

[17] Adler, G. Selgusid Eesti ihaldusväärseimad tööandjad 2024. *CV Keskus*, 2024. https://www.cvkeskus.ee/karjaarikeskus/tooturu-uudised/uudised/selgusid-eesti-ihaldusvaarseimad-tooandjad-2024 (25.04.2025).

[18] Culture Amp Case Studies. https://www.cultureamp.com/case-studies (25.04.2025).

[19] Ellis, J. Impact and Driver Analysis. *Culture Amp Articles*, (updated 2025). https://support.cultureamp.com/en/articles/7048590-impact-and-driver-analysis-faqs (25.04.2025).

[20] Lall-Trail, S. 22 employee engagement survey questions you should ask. *Culture Amp Blog*, 2023. https://www.cultureamp.com/blog/employee-engagement-survey-questions (25.04.2025).

[21] Bruneaux, T. What is Developer Experience? *GetDX Blog*, 2023. https://getdx.com/blog/developer-experience/#what-is-developer-experience (25.04.2025).

[22] DX Customers and Case Studies page. https://getdx.com/customers/ (25.04.2025).

[23] The One Number You Need to Increase ROI per Engineer. *GetDX Blog*, 2024. https://getdx.com/research/the-one-number-you-need-to-increase-roi-per-engineer/ (25.04.2025).

[24] Bhandari, P. Data Collection | Definition, Methods & Examples. *Scribbr*, 2020. https://www.scribbr.com/methodology/data-collection/ (25.04.2025).

[25] Preprocessing data, *scikit-learn*. https://scikit-learn.org/1.5/modules/preprocessing.html (25.04.2025).

[26] Santos, F. Modern methods for old data: An overview of some robust methods for outliers detection with applications in osteology. *Journal of Archaeological Science*, 2020. https://www.sciencedirect.com/science/article/abs/pii/S2352409X20302145 (25.04.2025).

[27] Nevil, S. Z-Score: Meaning and Formula. *Investopedia*, 2025. https://www.investopedia.com/terms/z/zscore.asp (25.04.2025).

[28] Correlation analysis, *DataTab*. https://datatab.net/tutorial/correlation (25.04.2025).

[29] Bobbitt, Z. The Five Assumptions for Pearson Correlation. *Statology*, 2021. https://www.statology.org/pearson-correlation-assumptions/ (25.04.2025).

[30] Malato, G. An Introduction to the Shapiro-Wilk Test for Normality. *BuiltIn*, 2025. https://builtin.com/data-science/shapiro-wilk-test (25.04.2025).

[31] Frost, J. Spearman's Correlation Explained. *Statistics By Jim*, 2021. https://statisticsbyjim.com/basics/spearmans-correlation/ (25.04.2025).

[32] Beers, B. Regression: Definition, Analysis, Calculation, and Example. *Investopedia*, 2025. https://www.investopedia.com/terms/r/regression.asp (25.04.2025).

[33] Frost, J. Overfitting Regression Models: Problems, Detection, and Avoidance. *Statistics By Jim*, 2017. https://statisticsbyjim.com/regression/overfitting-regression-models/ (25.04.2025).

[34] Frost, J. How To Interpret R-squared in Regression Analysis. *Statistics By Jim*, 2018. https://statisticsbyjim.com/regression/interpret-r-squared-regression/ (25.04.2025).

[35] What is principal component analysis (PCA)? *IBM*, 2023. https://www.ibm.com/think/topics/principal-component-analysis (25.04.2025).

Appendix

I. License

I, **Henri Laats** ,
(author´s name)

1. grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the digital archives of the University of Tartu until the expiry of the term of copyright, my thesis

Measuring the Impact of Developer Experience Improvements on Engagement in Pipedrive ,
(title of thesis)

supervised by Dietmar Alfred Paul Kurt Pfahl and Mykhailo Dorokhov ;
(supervisor´s name)

2. grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright;

3. am aware of the fact that the author retains the rights specified in points 1 and 2;

4. confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Henri Laats
15/05/2025