Exploring Analytics in Software Startups

Usman Rafiq

Free University of Bozen-Bolzano, 39100, Italy

Abstract

Software startup companies are seen everywhere across the globe. These companies are driving today's economy and innovation. However, a significant number of startups fail despite the constantly evolving and widespread launch. While research provides key practices to increase the odds of success for these companies, it lacks in providing evidence on how startups reduce uncertainties by utilizing analytics in their specialized context. This is the research gap that we want to address in our study. The uncertain nature of software startups makes them good candidates to empower themselves with analytics. We employ a multi-method research design to answer the research question. In particular, we perform multi-vocal exploratory studies, multiple-case studies, and a survey study. Our preliminary results, primarily based on multi-vocal exploratory studies, show an understanding of startups regarding analytics and identify challenges they face from an analytics perspective. Currently, we are executing multiple-case studies and in-process of selecting potential cases according to the designed criteria. The future work includes completing the remaining work on multiple-case studies and corroborating the findings with a survey study.

Keywords

start-up, entrepreneurship, lean, tech firm, big data, metrics

1. Introduction

Software startups are young, yet innovative, companies that offer software-intensive products or services [1]. These companies are different from established software companies in terms of their focus on innovation and rapid growth while managing extreme uncertainty [2][3]. Such characteristics make software startups unique from other companies [4]. Despite the constantly evolving and widespread launch of these companies across the globe, many startups fail. The research reports, such as [5], indicate that the failure rate is above 90%.

To increase the odds of success of these companies, several authors, such as [3], [6], [7], and [8], propose key practices, e.g. lean startup, customer-focused development, and agile-based methodologies. For instance, in lean startup methodology [3], which is an intensively adopted practice, practitioners are encouraged to find a viable product/business model by considering the build-measure-learn loops. These fast feedback loops are executed around the idea by building a Minimum Viable Product (MVP) until the startup finds a viable business model. While Bosch et al. [6] confirm the widespread use of lean startup in software startups, Pantiuchina et al. [9] corroborate the use of agile methods to develop the product. Nevertheless, Bosch et al. [6] also highlight the difficulties of startup practitioners in adopting lean startup and tailored the

International Conference on Software Business, November 08–11, 2022, Bolzano, Italy
urafiq@unibz.it (U. Rafiq)
0000-0003-3198-851X (U. Rafiq)

original lean startup methodology by proposing *"Early Stage Software Startup Development Model"*.

In the same vein, to resolve the considerable uncertainty about the viability of the proposed business model, Eisenmann et al. [10] suggest startup companies obtain a large amount of information. Based on certain information, startups take several decisions. The impact of these decisions lays the foundation for startups' success or causes to move towards company closure [11]. Therefore, dealing with information from a variety of aspects is a significant phenomenon in the life-cycle models of software startups.

On the other hand, established software companies are seen utilizing analytics to make decisions, understand customers and determine product experiences [12] [13]. While established companies are fully harnessing the power of analytics, startups are not seen making decisions based on such information. The existing body of research on startups lacks shreds of evidence on how a software startup deals with information and applies analytics. This is the research gap that we want to address in our study. The uncertain nature of software startups makes them good candidates to empower themselves with analytics. This will lead to improving their decision-making process and offer an aerial view of their achievements and shortcomings while moving fast.

Therefore, the following Research Question (RQ) is guiding our study:

RQ: How do software startups apply analytics?

The answer to this RQ leads to several research contributions. The principle contribution is an empirically grounded framework to understand and apply analytics in software startups. In addition, the following contributions are expected from this work: providing evidence corroborating the use of analytics in the startup context, key areas of startups where analytics can be applied, challenges that startup face while applying analytics, and finally the identification of the key practices to apply analytics inside software startup companies.

The rest of the paper is organized as follows: Section 2 provides a glimpse of the overall research design. Section 3 presents the research timeline in terms of a Gantt chart. Section 4 presents preliminary results and future directions. Lastly, section 6 concludes the paper.

2. Research Design

2.1. Research Questions

Our research goal is to understand the current practices of analytics inside software startups. Therefore, we break down this research goal by formulating four further research questions. The literature review clearly shows the significant lack of studies on the understanding of analytics by software startup companies and how they consider it. Therefore, our first research question is:

RQ1: How software startups understand analytics?

Here, we focus on exploring the state of practice, i.e., the utility of analytics in software startups. In particular, the rationale is to understand how startups perceive analytics and what are the possible areas where startups think that analytics is supporting their company. This strengthens the understanding of what aspects of analytics are practiced in software startups. Additionally, it enables to make the comparison of analytics, in the context of startups, with its usage in the software engineering domain.

RQ2: What challenges software startups face when applying analytics?

We formulate this research question to better understand the pitfalls to avoid. We think that addressing this question facilitates our study to emphasize the inhibitors of analytics adoption by software startup companies. The identified challenges can serve as an alert for startups while practicing analytics.

RQ3:What benefits, related to analytics, do software startups ascertain?

We explore what benefits startups realize in employing analytics. The research outcome increases our understanding of startup awareness concerning analytics benefits and we get familiarized with what they are aware of and what additional benefits can be achieved through it. Therefore, an analysis of the perception of startups concerning the benefits of using analytics strengthens us to move towards goal-oriented analytics practices inside startup companies.

RQ4: What are the key practices to apply analytics in software startups?

Finally, based on the empirical understanding we gain from previous research questions, we aim to report key practices. These key practices can support software startup companies in realizing the power of analytics. Therefore, an overall analytics framework for software startups will be designed while moving toward the ultimate goal of this research. This framework assumes to have the potential to assist startup companies in deciding what is critical for them to monitor within the scarcity of limited resources. It will entitle startups to filter the flood of available information, formulate insights and take corrective actions accordingly.

2.2. Research Approach

The current study is driven by the goal of understanding the utilization of analytics in software startups. We employ a multi-method research design to achieve this goal. A multi-method research design employs a blend of research methods to provide more reliable and generalized research results [14]. This combination could include distinct yet complementary research methods. Therefore, given the current nature of research, which is exploratory, we perform multivocal exploratory studies, multiple-case studies, and survey research to build our understanding. Table 1 provides a glimpse of the multiple methods we use to address our research questions.

RQ ID	Description	Primary Research Method	Complementary Research Method
RQ1	How software startups understand analytics?	Multi-vocal Study	Multiple-Case Study Survey Study
RQ2	What challenges software startups face when applying analytics?	Multi-vocal Study	Multiple-Case Study Survey Study
RQ3	What benefits, related to analytics, do software startups ascertain?	Multiple-Case Study	Survey Study
RQ4	What are the key practices to apply analytics in software startups?	Multiple-Case Study	Survey Study

Table 1
Overview of Research Approach

3. Planned Timeline

The following Gantt chart depicts the research timeline and deliverables of this study.

Figure 1: Gantt Chart for Research Planning and Execution

4. Preliminary Results

The detailed state of research according to RQs is presented in the following subsection.

4.1. RQ1: How software startups understand analytics?

Like other established companies, software startups are not observed turning data to actionable insights. The existing literature also puts no light on their understanding. Therefore, as a first step, we studied what startups think about analytics and how they understand it. To answer this question, we collected text data from three popular analytics platforms that claim themselves as a good fit for startups. We selected these three platforms from the list of 12 platforms. We got a clue about these 12 platforms from Hacker News [1]. This forum is very popular among entrepreneurs and the computer science community. To triangulate our selection, we also checked these platforms on the G2 [2] website, a source to check reviews of software. Finally, we selected the documentation of three platforms that were mentioned on both these forums and also passed our inclusion criteria. Moreover, regarding the data set, we covered associated documentation of analytics platforms and experience stories of real startups about their use cases. Particularly, our data set contained nine real-world scenarios of analytics usage, described in the text with varying lengths.

Later, we analyzed the data using the conventional content analysis technique [15]. Our results show that software startups first establish the set-up of analytics before its usage. This is referred to as instrumentation in our research. During the analytics instrumentation process, startups often set up and aim to track several goals. Later, the results reflect three main scenarios where startups are applying analytics currently. According to the results, the main areas of usage include experimentation, diagnosis, and getting insights. Fig. 2 shows how software startups handle analytics.

We presented these results as a full paper at the IWSiB workshop in ICSE 2022.

4.2. RQ2: What challenges do software startups face when applying analytics?

To answer this research question, we decided to investigate failed software startup companies and understand their challenges regarding analytics. We presented these challenges in terms of analytics mistakes that startups often do. To analyze the data, we used CBInsights' [3] post-mortem reports of software startups.CBInsights is a private American firm containing the largest database of startup reports. We collected 353 post-mortems and applied our inclusion/exclusion criteria. In the end, we utilized 22 reports in our final data set. To analyze this text data, we employed thematic analysis [17].

Our results reflect 10 identified challenges, presented in terms of analytics mistakes and grouped into four top-level themes. These themes contain mistakes regarding information collection, information analysis, information communication, and lastly, information usage. The following figure Fig. 3 highlights the mistakes made by software startups when handling analytics. These mistakes show the pitfalls to avoid when dealing with information from an analytics perspective.

The results of RQ2 are published at the Evaluation and Assessment in Software Engineering (EASE) conference as a full paper in 2021.

---

[1] https://news.ycombinator.com/
[2] https://www.g2.com/
[3] https://www.cbinsights.com/research/startup-failure-post-mortem/

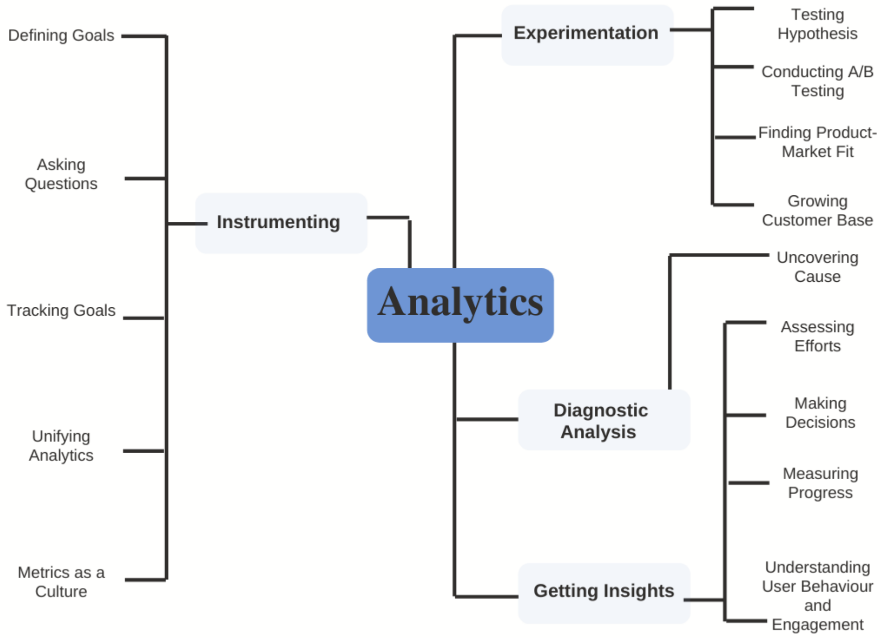

Figure 2: How Software Startups Understand Analytics (based on the results published in [16])

4.3. Future Directions

The work on RQ3 (analytics benefits for startups) and RQ4 (key analytics practices) is ongoing. We are primarily executing a multiple-case study approach to address these questions. Later, we plan to complement multiple-case studies with a survey study. Overall, these two research methods will contribute to strengthening the evidence for all RQs and help to produce generalized yet validated results.

At present, we are in process of running a multiple-case study. Our study follows a protocol that consists of semi-structured interviews with the startup founders. We developed as well as validated our protocol through pilot interviews with a software startup.

Regarding case selection, we intend to select cases based on two dimensions. The first dimension is the startup development stage. In this study, we follow the startup development stages proposed by Klotin et al.[19]. These stages include inception, stabilization, growth, and maturity. Likewise, for the other dimension, we consider the startup technology i.e. nature of the product of the startup. We propose to select startups withweb and mobile products. These are two popular market segments often covered by startup companies. This is evident from the survey studies of Pantiuchina et al. [9] and Wang et al. [20] who report that more than 50% startups are found in the web application domain, followed by startups targeting the mobile application space. We also observed a similar result while answering RQ2 and studying 353 post-mortem reports of failed startups. Along the same lines, we particularly selected mobile application based startups as the app store ecosystem has also an effect on the decisions regarding product engineering of the app [21]. Similarly, app stores, such as Apple's

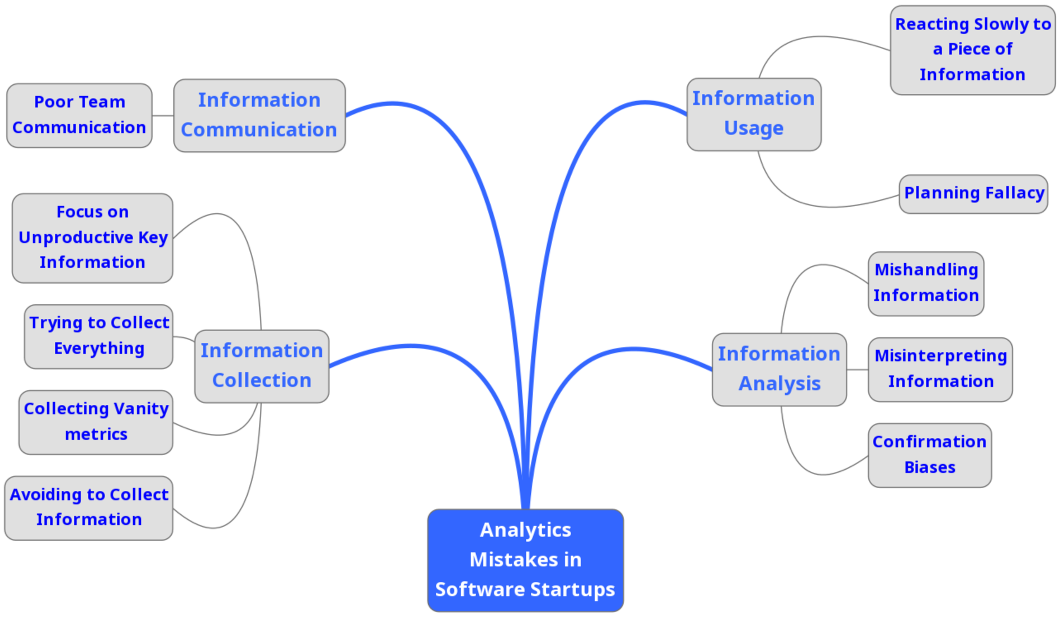

Figure 3: Analytics Mistakes Made by Failed Software Startups (published in [18])

app store and Google's Play, provide in-built analytics to measure some of the aspects of the app. Therefore, it would be interesting to relate app store effects while making a cross-case comparison. On that account, in our research, we plan to study cases in both these domains to have more generalized and strong results.

Table 2 provides an overview of our case selection dimensions and the case studies done so far.

Product Domain	**Software Startup Evolution Stages**			
	Inception	***Stabilization***	***Growth***	***Mature***
Web App	*TBD*	Case Alpha	Case Beta	TBD
Mobile App	TBD	TBD	TBD	TBD

Table 2
Overview of Case Selection Criteria. TBD Refers to Cases that Needs To Be Determined

4.3.1. Case Profiles

We conducted interviews with two startup companies up till now, aliased as Case Alpha and Case Beta to protect anonymity.

Case (Alpha) is a startup company offering cloud-based SaaS applications for software testing. The company was established in 2019 and achieved its problem-solution fit by pivoting multiple times. The startup team comprises six people including Chief Executive Officer (CEO), a developer, a designer, a marketing head, an innovation strategist, and a consultant (called an IT expert). Alpha has a B2B business model and it has four business customers currently. The company is seeking scaling now and we consider it in the second phase of startup i.e. stabilization.

The case was selected for a pilot interview, however, interesting insights were discussed during the interview and we planned to consider it as the main case. The interview was in-person with the CEO, designer, and marketing personnel. All the interview participants were holding informatics backgrounds with a professional degree in computer science/IT.

Case (Beta) is a startup company with an innovative cloud-based marketing solution for customers. It also follows the B2B business model and currently holds more than 2000 customers. The company is focusing on marketing and sales activities to boost its customer base. Therefore, we classify this company in the growth phase based on its circumstances. The company was co-founded by three members, however, currently, two are running its operations. The CEO holds a professional degree in informatics. We interviewed the CEO of the company through MS TEAMS. We slightly modified our questions after the pilot study and also eliminated a few of those, based on our prior understanding.

Regarding other cases, we plan to follow the criteria provided in Table 2. However, based on our research understanding of startups, we think that getting mature startups would not be an easy task.

5. Conclusions

Analytics plays a critical role in supporting decisions, measuring progress, and understating customer behavior. However, the existing research on startups lacks exploring analytics within this context. Therefore, a comprehensive understanding of how startups understand and apply analytics needs attention. The current research addresses this research gap and applies the multi-research method to fill the research gap. The study performs multi-vocal studies, multiple-case studies, and a survey study to develop the understanding. Currently, we are executing multiple-case studies and in-process to select potential cases according to our case selection criteria, which are twofold. The future work includes completing the multiple-case study and corroborating the findings with a survey study.

References

[1] M. Unterkalmsteiner, P. Abrahamsson, X. Wang, A. Nguyen-Duc, S. Shah, S. S. Bajwa, G. H. Baltes, K. Conboy, E. Cullina, D. Dennehy, et al., Software startups–a research agenda, e-Informatica Software Engineering Journal 10 (2016).

[2] S. Blank, The four steps to the epiphany: successful strategies for products that win, John Wiley & Sons, 2020.

[3] E. Ries, The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses, Currency, 2011.

[4] S. M. Sutton, The role of process in software start-up, IEEE software 17 (2000) 33–39.

[5] S. Genome, Global Startup Ecosystem Report 2019, Technical Report, Startup Genome, San Francisco, CA, 2019.
[6] J. Bosch, H. H. Olsson, J. Björk, J. Ljungblad, The early stage software startup development model: a framework for operationalizing lean principles in software startups, in: International Conference on Lean Enterprise Software and Systems, Springer, 2013, pp. 1–15.
[7] C. Giardino, N. Paternoster, M. Unterkalmsteiner, T. Gorschek, P. Abrahamsson, Software development in startup companies: the greenfield startup model, IEEE Transactions on Software Engineering 42 (2015) 585–604.
[8] C. Nobel, Teaching a 'lean startup' strategy, HBS Working Knowledge (2011) 1–2.
[9] J. Pantiuchina, M. Mondini, D. Khanna, X. Wang, P. Abrahamsson, Are software startups applying agile practices? the state of the practice from a large survey, in: International Conference on Agile Software Development, Springer, Cham, 2017, pp. 167–183.
[10] T. R. Eisenmann, E. Ries, S. Dillard, Hypothesis-driven entrepreneurship: The lean startup, Harvard Business School Entrepreneurial Management Case (2012).
[11] E. Klotins, M. Unterkalmsteiner, T. Gorschek, Software engineering in start-up companies: An analysis of 88 experience reports, Empirical Software Engineering (2018) 1–19. doi:10.1007/s10664-018-9620-y.
[12] P. Mikalef, I. O. Pappas, J. Krogstie, M. Giannakos, Big data analytics capabilities: a systematic literature review and research agenda, Information Systems and e-Business Management 16 (2018) 547–578.
[13] L. Guerrouj, O. Baysal, D. Lo, F. Khomh, Software analytics: challenges and opportunities, in: 2016 IEEE/ACM 38th International Conference on Software Engineering Companion (ICSE-C), IEEE, 2016, pp. 902–903.
[14] M. Wood, J. Daly, J. Miller, M. Roper, Multi-method research: An empirical investigation of object-oriented technology, Journal of systems and Software 48 (1999) 13–26.
[15] H.-F. Hsieh, S. E. Shannon, Three approaches to qualitative content analysis, Qualitative health research 15 (2005) 1277–1288.
[16] U. Rafiq, Towards understanding analytics in software startups, in: 2022 IEEE/ACM International Workshop on Software-Intensive Business (IWSiB), 2022, pp. 31–38.
[17] C. B. Seaman, Qualitative methods in empirical studies of software engineering, IEEE Transactions on software engineering 25 (1999) 557–572.
[18] U. Rafiq, J. Melegati, D. Khanna, E. Guerra, X. Wang, Analytics mistakes that derail software startups, in: Evaluation and Assessment in Software Engineering, 2021, pp. 60–69.
[19] E. Klotins, M. Unterkalmsteiner, P. Chatzipetrou, T. Gorschek, R. Prikladnicki, N. Tripathi, L. Pompermaier, A progression model of software engineering goals, challenges, and practices in start-ups, IEEE Transactions on Software Engineering (2019).
[20] X. Wang, H. Edison, S. S. Bajwa, C. Giardino, P. Abrahamsson, Key challenges in software startups across life cycle stages, in: International conference on agile software development, Springer, Cham, 2016, pp. 169–182.
[21] A. A. Al-Subaihin, F. Sarro, S. Black, L. Capra, M. Harman, App store effects on software engineering practices, IEEE Transactions on Software Engineering 47 (2021) 300–319. doi:10.1109/TSE.2019.2891715.