

Nithish Hulikanthe Math

---

**Multi-Objective Success Factor  
Analysis for IT based Startups**

---





FAKULTÄT FÜR  
INFORMATIK

FACULTY OF COMPUTER SCIENCE

Master Thesis

## Multi-Objective Success Factor Analysis for IT based Startups

Author: Nithish Hulikanthe Math

Professor: Prof. Dr. Sanaz Mostaghim

Supervisor and Examiner: Prof. Dr. Sanaz Mostaghim

Examiner: Prof. Dr. André Presse,  
University of Bolzano, Italy

Winter term 2016/2017

**Nithish Hulikanthe Math:** *Multi-Objective Success Factor Analysis for IT based Startups*  
Otto-von-Guericke-Universität  
Magdeburg, 2016/2017.

---

# Abstract

In this thesis, data set consisting of information regarding various phases of the new firm creation process of a number of IT based start-up companies are subjected to the application of the modern method of computational science, multi-objective optimization. The dataset with a number of variables is sorted based on certain objectives to identify successful start-ups. These start-ups are analyzed to determine the success factors responsible for the success of the companies. The identified success factors are further analyzed to measure the performance of one objective over the other. The analysis resulted in interesting facts from the most successful companies compared to the least successful companies.



---

# Acknowledgement

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Sanaz Mostaghim for providing me with this incredible opportunity to conduct my Master thesis at the department of Intelligent Systems, Otto von Guericke University Magdeburg. Also, I thank Prof. Sanaz for being my supervisor and mentor through out the thesis work guiding and supporting me with invaluable insights in achieving the results.

I would also like to thank my parents, Aruna Himachal and Himachal H M for the financial and emotional support during the entire study program. And last but not the least, I would like to thank my brother and my friends for being there for me at every stage of my studies.





---

# Contents

<b>List of Figures</b>	<b>VII</b>
<b>List of Tables</b>	<b>IX</b>
<b>List of Acronyms</b>	<b>XI</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. Proposed Approach</b>	<b>3</b>
2.1. Independent Variables . . . . .	3
2.2. Performance Measures . . . . .	4
<b>3. PSED Dataset</b>	<b>7</b>
3.1. PSED OUTCOME . . . . .	8
3.1.1. PSED Screener . . . . .	10
3.1.2. STATUS OF BUSINESS . . . . .	10
3.1.3. TYPE AND LOCATION OF BUSINESS . . . . .	10
3.1.4. LEGAL FORM OF BUSINESS . . . . .	11
3.1.5. START-UP ACTIVITIES . . . . .	11
3.1.6. START-UP FINANCES . . . . .	12
3.1.7. ATTITUDES TOWARD COMPETITION . . . . .	12
3.1.8. OWNERS, KEY NON-OWNERS, AND HELPERS . . . . .	13
3.1.9. OWNER DEMOGRAPHICS . . . . .	13
3.1.10. COMMUNITY RESOURCES . . . . .	13
3.1.11. START-UP INVESTMENTS, DEBTS AND NET WORTH . . . . .	14
3.1.12. LEGAL ENTITY START-UP INVESTMENTS, NET WORTH . . . . .	14

3.1.13. MARKET AND COMPETITION . . . . .	15
<b>4. Multi – Criteria Optimization</b>	<b>17</b>
4.1. Computational science and its potential . . . . .	17
4.2. Multi-criteria optimization . . . . .	17
4.3. Multi-Objective Optimization Problem . . . . .	18
4.3.1. The concept of Domination . . . . .	19
4.3.2. Pareto Optimality and Pareto Optimal Fronts . . . . .	19
4.4. Non-Dominated Sorting . . . . .	20
4.4.1. Non-Dominated Sorting Algorithm . . . . .	21
<b>5. Analysis</b>	<b>25</b>
5.1. Software used . . . . .	25
5.2. Method used for Analysis . . . . .	28
5.2.1. Intra front analysis . . . . .	28
5.2.2. Inter Front Analysis . . . . .	35
<b>6. Conclution and Future Work</b>	<b>43</b>
6.1. Conclution . . . . .	43
6.2. Future Work . . . . .	44
<b>Appendices</b>	<b>45</b>
<b>A. PSED Dataset Variables</b>	<b>47</b>
<b>B. Variable values</b>	<b>51</b>
<b>Bibliography</b>	<b>70</b>

---

# List of Figures

2.1. Success Factors [Curtin, 2012] . . . . .	4
2.2. Analytical Model : Performance Measures . . . . .	4
4.1. Non-Dominated Sorting Example . . . . .	20
4.2. First set of solutions . . . . .	22
4.3. Second set of solutions . . . . .	23
4.4. Non-Dominated Sorting of Start-up Companies . . . . .	23
5.1. Pareto Fronts Representation . . . . .	26
5.2. Pareto Front 10 . . . . .	27
5.3. Pareto Optimal Front . . . . .	27
5.4. Cases considered for Analysis . . . . .	29



---

# List of Tables

5.1. Front 1 vs 17 case 1 Analysis . . . . .	36
5.2. Front 1 vs 17 case 2 Analysis . . . . .	38
5.3. Front 1 vs 17 case 3 Analysis . . . . .	40
A.1. Start-up Variables and Labels from PSED dataset . . . . .	49
B.1. Start-up Variables and Value ranges from PSED dataset . . . . .	67



---

# List of Acronyms

PSED - Panel Study Of Entrepreneurial Dynamics

SPSS - Statistical Package for the Social Sciences

OR - Operation Research

SAS - Statistical Analysis System





---

# 1. Introduction

From the Eric Ries's book "The Lean Startup," it is described that a start-up is an organization with a group of individuals aimed at building a new product or service under extreme uncertainty [Ries, 2011].

Nascent Entrepreneurship is the trend in recent times resulting in an increase in Start-up companies around the world. In search of new economic growth, Start-ups have received significant interest from policy makers. To illustrate, Mr. Narendra Modi, Prime Minister of India, launched an initiative named Start-up India in January 2016 which is a government program which aims at supporting start-ups with economic growth and development thereby boosting digital entrepreneurship [GovtIndia, 2016].

In this research, the real-time data set from Panel Study Of Entrepreneurial Dynamics program sponsored by the University of Michigan is analyzed to identify some success factors with which a start-up can be termed successful. This data set is comprising of data in the form of questionnaires which are interviews conducted with some start-ups, and the resulting answers are formulated into a large data file categorized into different sections. With this, the primary intention of this research would be that for a new start-up, one can be more prepared before starting a new firm keeping in mind the success factors identified in this thesis. Here, a computational method called multi-objective optimization has been used to study the start-up factors of 649 companies with 86 factors from the Panel Study of Entrepreneurial Dynamics program. A non-dominated sorting algorithm is used to implement the multi-objective optimization problem to sort the data of the companies based on the objectives defined.

This study is structured as follows. With this Introduction constituting the first chapter, in the second chapter, the approach for this thesis is proposed. In the third section, the data set which is used in the research from PSED is explained in detail; this chapter is the foundation for the entire research

as the dataset described is the sole data utilized for the complete research. Then in the fourth chapter, the computational science method: Multi-objective optimization is introduced followed by the non-dominated sorting algorithm used in this study to sort all the data for analysis. In chapter five, Analysis of the data set is conducted based on Multi-criteria optimization<sup>1</sup> and obtained results from the non-dominated sorting algorithm. Research is concluded with an outlook on future work in chapter six.

---

<sup>1</sup>Multi-objective optimization and Multi-criteria optimization both are used invariably

---

## 2. Proposed Approach

In this research, the Multi-criteria optimization phenomenon is used to analyze a real-time data set. Both Multi-criteria Optimization and the real-time dataset used are explained in separate chapters. As the name indicates, in multi-criteria optimization, more than one criteria is used to solve a problem which can be minimization or maximization problem[Deb, 2001]. And optimization is the process of finding one or more feasible solutions which correspond to extreme values of one or more criteria[Deb, 2001]. Therefore, the primary objectives used in this study for maximization are: Number of months a company survived and Number of employees in a company at the end of the survival period. Here we further build our model to understand the approach of this study in a more precise way.

### 2.1. Independent Variables

The variables or the factors from the companies used in this analysis are categorized as in 2.1.

Independent Variables or Success Factors	Variables
PSED SCREENER	QS7, INCOME, QFF3, QFF4, QFF5, QFF6
STATUS OF BUSINESS	AA5a, AA5b, AA6a, AA6b AA7, AA10, AE13, FE13, BA37
TYPE AND LOCATION OF BUSINESS	AB10
LEGAL FORM OF BUSINESS	AC1
START-UP ACTIVITIES	AD1, AD6, AD9, AD11, AD13, AD16, AD20, AD22, AD24, AD26
START-UP FINANCES	AE3, AE5, AE7, AE9, AE11, AE24
ATTITUDES TOWARD COMPETITION	AF1, AF2, AF3, AF4, AF5, AF6, AF8, AF9, AF10
OWNERS, KEY NON-OWNERS, AND HELPERS	AG1
OWNER DEMOGRAPHICS	AH6_1, AH15_1, AH19_1, AH25_1, AH27_1
COMMUNITY RESOURCES	AP2, AP3, AP7, AP8
START-UP INVESTMENTS, DEBTS AND NET WORTH	AQ4_1, AQ5_1, AQ6_1, AQ7_1, AQ8_1, AQ9_1, A Q10_1, AQ12x_1, AQ12_1, AQ13_1, AQ12x, AQ1 3
LEGAL ENTITY START-UP INVESTMENTS	AR1, AR4, AR32, AR22
MARKET AND COMPETITION	AS1, AS2, AS3, AS6, AS9

Figure 2.1.: Success Factors [Curtin, 2012]

The model is prepared based on the data set from The Panel Study of Entrepreneurial Dynamics (PSED) research program by classifying the data into thirteen independent variables where the real-time data or factors in the form of questionnaires are further grouped into these independent variables. More on this in the next chapter. The labels of these variables are mentioned in appendix A.1. The factors grouped into these independent variables are carefully selected keeping in mind the scope of this research. All the factors or variables are studied individually from the PSED code book.

## 2.2. Performance Measures

Factor	Variable
Growth	Number of Employees
Survival	Months of Survival

Figure 2.2.: Analytical Model : Performance Measures

Measuring performance of any organization is important as it reflects the progress of a company [Taticchi, 2010]. Performance becomes an important

parameter in measuring the success of a start-up. Hence assuming that the performance measures are to be maximized, two performance measures are identified, and objectives for multi-objective optimization problem are formed.

For a company to be considered as successful, performance on long-term survival, job creation, value and other parameters should be optimized [Neely, 2002]. Based on the goals and primary objectives of a company, the decision regarding which parameter to focus on at the beginning can be decided. The primary objectives envisaged for this study are based on the performance measures indicated in figure 2.2. The variables or the factors of start-ups are deeply analyzed based on these criteria or objectives.

So, as indicated in Chapter Introduction, Performance measures data of 649 companies are sorted with the independent variables from figure 2.1 using the non-dominated sorting algorithm as described in Chapter 4 algorithm 1. Then with the achieved sorted data in different sets, the results are analyzed. A total of forty sets are achieved with all the details of six hundred and forty nine companies distributed in them.

In the next chapter, PSED dataset and the independent variables are discussed in detail.



---

## 3. PSED Dataset

Ewing Marion Kauffman Foundation, being a primary sponsor of the Panel Study of Entrepreneurial Dynamics (PSED) research program, had the sole purpose of distributing knowledge and critical factors in the entrepreneurial process [Reynolds and Curtin, 2009]. The program focused on studying the situations and circumstances of the people who decided to start a new venture. Situations could be weather quitting a job or passion for solving a problem and other instances which led to starting the new venture. The program also tried to understand the factors which were for starting the new venture and which factors acted as barriers in establishing a new firm [Reynolds and Curtin, 2009].

PSED takes into account each and every step right from idea generation phase to subsequent stages until the creation of new company [Reynolds and Curtin, 2009]. These actions may include the creation of a business plan, marketing strategy, financial modeling, and others. However, PSED does not concentrate much on the early years of the start-up. I.e. the survival phase in the initial years of the start-up [Reynolds and Curtin, 2009]. A nascent enterprise is created if all the challenges and barriers are overcome. It is during the early years of the start-up that the discontinuation rate of start-ups is high [Reynolds and Curtin, 2009].

With the successful start of this project, Kauffman Foundation provided funds for continuation of initial panel interviews and further research [Reynolds and Curtin, 2009]. It also decided to support second-panel projects. The data set obtained from these projects were used for analysis and also, students and scholars used the dataset to conduct research and write their dissertations, books, and articles [Reynolds and Curtin, 2009]. Other studies followed by in other countries following the same design and procedure as that of PSED resulting in further projects similar to PSED [Reynolds and Curtin, 2009]. Now countries like China, Sweden, Norway,

Austria, Canada, Netherlands have similar projects which give more opportunities to students and scholars to conduct research all over the world [Reynolds and Curtin, 2009]. It is clear that PSED data is aimed at providing greater opportunities for entrepreneurial and research scientists in the years to come [Reynolds and Curtin, 2009].

## 3.1. PSED OUTCOME

Two projects namely Panel Study of Entrepreneurial Development I and II were conducted in U.S to study the business creation process of new firms [Reynolds and Curtin, 2012]. First, the projects began with the screening of samples of adults located in U.S to check if the people were still active in the business creation process [Reynolds and Curtin, 2012]. With the selected candidates from the selection, further interviews were conducted to record the efforts and their outcome in the firm creation process [Reynolds and Curtin, 2012]. It is the answers of these interviews which are converted into large sets of data files which is used in this research. The data files could be accessed using SPSS software.

The results of conducting such interviews were two large-scale, multi-wave data sets [Reynolds and Curtin, 2012]. PSED I produced 1261 cases with four waves of questionnaire and screener data amounting to 6000 variables [Reynolds and Curtin, 2012]. Out the 1261 cases, eight hundred and thirty were nascent entrepreneurs. And PSED II produced 1214 cases with six waves of questionnaire and screener data resulting in eight thousand variables [Reynolds and Curtin, 2012]. In PSED II, all were nascent entrepreneurs. It is a very tedious task to process and analyze such large amounts of data to any professional analyst. Hence, a consolidated data set is provided which contains 2,024 cases and one hundred and twenty-seven standardized variables [Reynolds and Curtin, 2012]. It is also reported in the outcome document that after six years of entry in the start-up process, forty-eight percent have quit and thirty percent have reported first profits [Reynolds and Curtin, 2012].

Panel Study Of Entrepreneurial Dynamics II is a program coordinated by Survey Research Centre at the University of Michigan under the supervision of co-principle investigator Paul Reynolds and Richard Curtin [Curtin, 2012]. The design of screener and questionnaire of PSED were a group of researchers



working at the University [Curtin, 2012]. As said earlier panel participants were identified before the launch of their new business based on screener questions and these companies were tracked during every stage from launch to birth or death of a start-up [Curtin, 2012].

In this research, 649 cases have been selected with a range of companies which can be categorized into number of months a company survived and number of employees a company had until it quit or survived. Since the dataset included too many businesses that survived for a very minimum period or had very fewer employees, these cases have not been considered to draw a meaningful analysis using multi-objective optimization. Also, for the purpose of Multi-criteria optimization (which will be explained in detail in a later chapter) eighty-six variables are considered from the dataset which can be grouped under thirteen Independent variables or success factors of our research.

From the PSED code book [Curtin, 2012], The thirteen independent variables are:

- PSED SCREENER
- STATUS OF BUSINESS
- TYPE AND LOCATION OF BUSINESS
- LEGAL FORM OF BUSINESS
- START-UP ACTIVITIES
- START-UP FINANCES
- ATTITUDES TOWARD COMPETITION
- OWNERS, KEY NON-OWNERS, AND HELPERS
- OWNER DEMOGRAPHICS
- COMMUNITY RESOURCES
- START-UP INVESTMENTS, DEBTS AND NET WORTH
- LEGAL ENTITY START-UP INVESTMENTS
- MARKET AND COMPETITION

So here, from the questionnaire in the form of variables from the PSED data set, the factors mentioned above are being measured through a total

of eighty-six variables. Refer A.1 for the details of the variables used in this research. To provide a better understanding, each of the factors is introduced to justify its use and explain how they are measured through the variables. All the variable information discussed here are based on the PSED code book [Curtin, 2012].

### **3.1.1. PSED Screener**

The PSED Screener variable gives information of the business such as education of the owner, income of the owner before and after deciding to start a new business, does the owner owns part or all of the new business, did the business make progress in terms of revenue in the first six months of its initiation and were the salaries included in the calculation of the expenses.

### **3.1.2. STATUS OF BUSINESS**

It is important to consider the status of the business as an independent variable for our research because it gives some useful insights about the firm during the early stages of the enterprise. The status of a business is an independent variable with variables involving questionnaire from the PSED dataset which gives information such as opportunities and major problems faced during the start of the business, was there already an idea or the decision to initiate a trade was the first step to start the new firm and whether the company received income from sales during the first six months of the new business.

### **3.1.3. TYPE AND LOCATION OF BUSINESS**

The type and location of a business play a crucial factor for the firm. Here the variables corresponding to the type and location are in the form of questions asked to the business owners about the type of their activities which can be individual, independent, franchise, take-over business. All possible values of all the variables are presented in B.1. Location for some businesses plays a vital role as it defines the market of a product or

service. Hence it is important to consider these variables as the measures of success for a company

### **3.1.4. LEGAL FORM OF BUSINESS**

It is important to see if the legal form of business has been formally established by registering with the appropriate government agency and also sometimes it is interesting to see if a business is carrying liability insurance or plans to take it in the future or if its not relevant for the firm.

### **3.1.5. START-UP ACTIVITIES**

In this Independent variable, information related to Start-up activities is obtained. Information such as:

- Is there a business plan developed during the first six months of the start-up
- Stage of product development
- Whether promotional activities have been started
- Has the business developed any proprietary technology, processes, or procedures that no other company can use
- Has application for patent submitted
- Are major items relevant to business purchased, leased or rented
- Have there been discussions with potential customers
- Is information regarding competitors obtained?
- Are Market opportunities defined and
- Are Financial projections developed

These variables from the PSED code book [Curtin, 2012] are imperative to analyze the start-up performance of a company and Start-up activities are an important success factor for this research.

### 3.1.6. START-UP FINANCES

It is evident that start-ups rely heavily on external funding. It may be through investors or public financing [Cable, 2010, Oranburg, 2016]. During initial years, it is common for any new prospective start-up owner to use his or her personal savings to fund the business [Cole, 2009, Oranburg, 2016]. In some cases, the funds during the gestation phase of a firm may be from friends and family members or other relatives [Alden, 2011, Oranburg, 2016]. Hence, it is important to study financial activities of a start-up. Variables used in this study as described in PSED code book [Curtin, 2012] relating to finance are as follows:

- Have the start-ups received external funding through financial institution or other people
- credit with a supplier established
- Are there any employees working for pay during the first six months of the firm creation process
- How many employees or managers work for more than 35 hours a week
- Is there a bank account opened exclusively for this business in the first six months

### 3.1.7. ATTITUDES TOWARD COMPETITION

In this research, the variables involved in defining this success factor are:

- Are lower prices important for a business to be an effective competitor
- Quality products or services are important for a business to be an active competitor
- Serving those missed by others is important for a business to be an effective competitor
- Being first to market a new product or service is important for a business to be an effective competitor
- Doing a better job of marketing and promotions
- A superior location and customer convenience

- Technical or scientific expertise of the start-up team is important
- Developing new or advanced product technology or process technology for creating goods or services is important

### **3.1.8. OWNERS, KEY NON-OWNERS, AND HELPERS**

It is a well-known fact that creation of a new business involves the contribution of many individuals. Here one variable is considered which depicts if the firm is self-owned, shared with a spouse or shared with other people.

### **3.1.9. OWNER DEMOGRAPHICS**

Owner demographics helps in understanding the characteristics of the holder, and thereby it can be considered as a very important success factor in analyzing the start-up performance of business. Here, following information from the variables is considered for the research:

- Highest level of education of the owner
- Regarding work activity is the owner working for others for pay
- What is the owner's primary role in the new business
- Does the owner have provided training with business related tasks or skills
- Is the owner provided with physical resources, use of land, space, buildings or equipment

### **3.1.10. COMMUNITY RESOURCES**

In some way, the attitude of the community plays a role in the success of a firm. Following are the information extracted from the variables of Community resources:

- Do the social norms and culture of a society emphasize self-sufficiency, autonomy, and personal initiative
- Do the social norms and culture of a community encourage entrepreneurial risk-taking.
- Do state and local governments in a community provide excellent support for those starting new businesses
- Do Bankers and other investors in a community go out of their way to help new businesses get started.

### **3.1.11. START-UP INVESTMENTS, DEBTS AND NET WORTH**

In this independent variable, information regarding how business was funded, were there any loans made to start a business is obtained. The information from variables is as follows:

- Funds invested from personal savings
- Funds as loans from family
- Funds as loans from friends
- Funds from credit card loans
- Personal bank loans
- Funds from asset bank loan
- Funds from other sources
- Total amount of funding
- Total amount of funding as loans
- And total amount funding and funding as loans from all owners of a company

### **3.1.12. LEGAL ENTITY START-UP INVESTMENTS, NET WORTH**

From this independent variable, the information regarding financial support, which bank account were the funds accommodated, the net worth

of the company is obtained from a legal perspective. So here the collected information is:

- Whether business received formal financial support
- Amount invested after registration
- Whether all the funds were in new business bank account
- Total market value of the business if sold

### 3.1.13. MARKET AND COMPETITION

There were few questions in the questionnaire which were developed keeping both market and competition in mind. Here variables under this category are very interesting for the analysis as it gives an overall picture of the mindset of the start-up team on their market and competition. The information obtained from the variables is:

- Whether all/some or none of the potential customers consider the product/service as new or unfamiliar
- Whether other businesses offer same or similar product/service to the potential customers
- Were the technology/procedures required for this product or service available more than a year ago
- Whether the new business is hi-tech
- Expectation from owners regarding the percentage of National customers(Here it is U.S) within two to three years of operation.

Apart from the variables mentioned above, a measure of these factors is also done through thirteen variables which provide information regarding the status of each company. By an interval of three months, thirteen variables are divided from zero months a company survived to sixty months.

Above questionnaires are answered by all the companies considered in this research and are used in determining the start-up performance of enterprises using multi-criteria optimization.

It is important to note that not all companies answered all the questions asked. Many businesses chose to ignore some of the questions and analysis is conducted with the available data.



---

## 4. Multi – Objective Optimization

### 4.1. Computational science and its potential

Now in this chapter let us understand how computational science, multi-criteria optimization, in particular, can help to find the companies which may be considered as successful in terms of objectives defined in chapter 2.

Computational Science is a very dynamically evolving area of research which mainly focusses on building mathematical models and quantitative analysis techniques for problem-solving in various scientific disciplines [Maxville, 2013]. The evolution of computational science has led researchers and scientists to perform complex computational tasks on large data sets [Peng, 2011]. With the availability of large public datasets, one can perform computational techniques instead of using traditional techniques and contribute positively in the scientific study fields [Peng, 2011]. Computational science will contribute significantly in the upcoming future advances in the field of science and technology. In this study, out of many computational techniques, we introduce only the concept multi-criteria optimization which is relevant to this study. Here we familiarize ourselves with Pareto-optimality which is the base of our research.

### 4.2. Multi-Objective optimization

Optimization can be defined as the task of finding one or more feasible solutions by minimizing or maximizing one or more objectives of a mathematically formulated problem [Deb, 2001].

In every area of research, problems with multiple objectives are common and finding solutions to it has been challenging to researchers for a long time.

Despite having various approaches and techniques in Operation Research (OR) and other disciplines to handle such problems, the complexity of their solutions gives an opportunity for other alternate solutions [Coello et al., 2007]. One such alternative is the use of multi-criteria optimization.

### 4.3. Multi-Objective Optimization Problem

A general multi-criteria optimization problem which is to be minimized or maximized consists of a number of objectives and is associated with a number of inequality and equality constraints which any feasible solution must satisfy [Deb, 2001]. Mathematically from [Deb, 2001], the problem can be written as follows:

$$\begin{aligned} & \text{Maximize } f_1(x), f_2(x), \dots, f_N(x) \\ & \text{Subject to } x \in S \end{aligned} \tag{4.1}$$

The solution  $x$  is a  $p$  dimensional vector having  $p$  decision variables and  $S$  is the feasible region.

Solutions to a multi-criteria optimization problem are expressed in terms of non-dominated points. For a maximization problem, if a vector  $x(1)$  is greater than vector  $x(2)$  i.e.  $(x(1) > x(2))$ . If at least one value of  $x(1)$  is greater than  $x(2)$  and no value of  $x(2)$  is greater than  $x(1)$ , we say  $x(1)$  dominates  $x(2)$ . Also, any such member of the problem which is non-dominated belongs to the non-dominated set of solutions, also known as Pareto-optimal solutions. In the sub-section Pareto-optimality, it is further illustrated with an example. Mathematically, an optimization algorithm should be terminated once a Pareto-optimal solution is reached. In Practice, it is desirable to find all the Pareto-optimal solutions as the suitability of the solution depend on a number of factors like problem environment, designer's choice.

In this research, it is very important for the analysis to find not one Pareto optimal set of companies but many such sets called Pareto fronts which are sorted based on above said principle. Algorithm and detailed explanation are given in the further sections.

### 4.3.1. The concept of Domination

For a maximization problem, a solution  $x(1)$  is said to dominate solution  $x(2)$ , if it satisfies both the conditions stated in equation 4.2 [Deb, 2001].

1. *Solution  $x(1)$  is no worse than solution  $x(2)$  in all objectives, or  $f_j(x(1)) \geq f_j(x(2))$  for all  $j = 1, 2, 3 \dots M$ .*
  2. *Solution  $x(1)$ ; is strictly better than  $x(2)$  in atleast one objective, or  $f_j(x(1)) > f_j(x(2))$  for atleast one  $j \in \{1, 2, \dots M\}$ .*
- (4.2)

In multi-criteria optimization, there need not be only one optimal solution but a set of equally optimal values that are identified using the above equation. The objective value of each of these solutions is not dominated by any other solution in the feasible objective region [Deb, 2001].

### 4.3.2. Pareto Optimality and Pareto Optimal Fronts

For a maximization problem, the entire feasible region contains solutions. In those solutions, the set of solutions which are not dominated by any other member in the feasible region is called Pareto optimal set or Global Pareto optimal set [Deb, 2001].

After finding the Pareto-optimal front, other consecutive fronts can be found by eliminating the members of the already found Pareto-optimal front in each iteration.

It is tough to make decisions in multi-criteria optimization as there are some solutions in the Pareto-optimal set, where none of the solutions can be said to be better than any other solutions of the front. It is therefore required to make a trade-off between solutions by compromising with one objective to gain in other objectives.

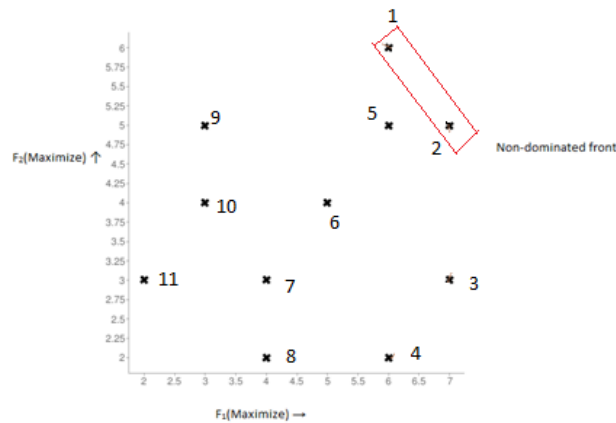


Figure 4.1.: Non-Dominated Sorting Example

From the example, one can clearly see that the members of the non-dominated front are not dominated by any other members. This is a maximization problem, and cases 1,2 are the best solutions having the highest values regarding their objectives compared to all other members.

And the next front consists of case 5,3 followed by case 9,6,4 in the next front, case 10, 7 in the next front and 11,8 in the last front.

A similar approach is used in this study to sort all the companies into various fronts as in the above-mentioned example. Later analyses of the best companies are carried out on their variables or success factors.

## 4.4. Non-Dominated Sorting

Using the theory explained in the section concept of domination, the non-dominated sorting algorithm used for this study is explained: Here the objectives to be maximized are a number of employees in a company and number of months a company survived, with the available data for sixty months.

The condition for domination is as follows:

$$\begin{aligned}
 & \text{if } ((a(1) > b(1)) \wedge (a(2) \geq b(2))) \vee ((a(1) \geq b(1)) \wedge (a(2) > b(2))) \\
 & \quad a \text{ dominates } b \\
 & \text{elseif } ((b(1) > a(1)) \wedge (b(2) \geq a(2))) \vee ((b(1) \geq a(1)) \wedge (b(2) > a(2))) \\
 & \quad b \text{ dominates } a
 \end{aligned} \tag{4.3}$$

a and b being two different companies and 1 and 2 are the objectives to be maximized.

#### 4.4.1. Non-Dominated Sorting Algorithm

The Non-dominated Sorting algorithm as described in the book Multi-Objective Optimization [Deb, 2001] is as follows:

---

**Algorithm 1** Non-Dominated Sorting Algorithm [Deb, 2001]

---

Input: Excel file with all the variables and objectives of all the companies

1: **procedure**

- 2: Set all non-dominated sets  $P_j$ , ( $j=1,2,\dots$ ) as empty sets. Set non-domination level counter  $j = 1$ .
- 3: Initialize  $P' = \{1\}$ . Set solution counter  $i = 2$ .
- 4: Compare solution  $i$  with  $j$  from  $P'$  for domination using equation 4.3.
- 5: If  $i$  dominates  $j$ , delete the  $j^{\text{th}}$  member from  $P'$ . If  $j < |P'|$ , increment  $j$  by one and go to step 4. Otherwise go to step 6. Alternatively, if  $j^{\text{th}}$  member of  $P'$  dominates  $i$ , increment  $i$  by one and go to step 3.
- 6: Insert  $i$  in  $P'$ . If  $i < N$ , increment  $i$  by one and go to step 3. Otherwise, stop and declare  $P'$  as the non-dominated set of Population  $P$ .
- 7: Update  $P_j = P'$  and  $P = P'$ .
- 8: If  $P \neq 0$ , increment  $j$  by one and go to step 3. Otherwise, stop and declare all non-dominated sets  $P_j$ , for  $i = 1,2,\dots,j$ .

Output : Complete Excel file sorted according to equation 4.3 in 40 fronts

---

With the data of six hundred and forty nine companies, the non-domination level counter of non-dominated empty sets is set to one.

In step three of the algorithm, the first set of solution is initialized with company one, solution counter is set to company two. Matlab program starts to compare company one with the company two based on the objectives defined and using equation 4.3. If company two dominates company one, delete company one from the first non-domination set and insert company two to this set. Now increment the solution counter to compare company three with the already existing company two in the non-domination set. If this company does not dominate company two and company two also does not dominate company three, add company three to the first non-domination set and increase the solution counter to check if the company replaces an existing company from the set or gets added to the set. This process is continued with all the six hundred and forty-nine companies, and the first non-domination set is declared.

The first set of solutions which are optimal solutions is represented as in 4.2.

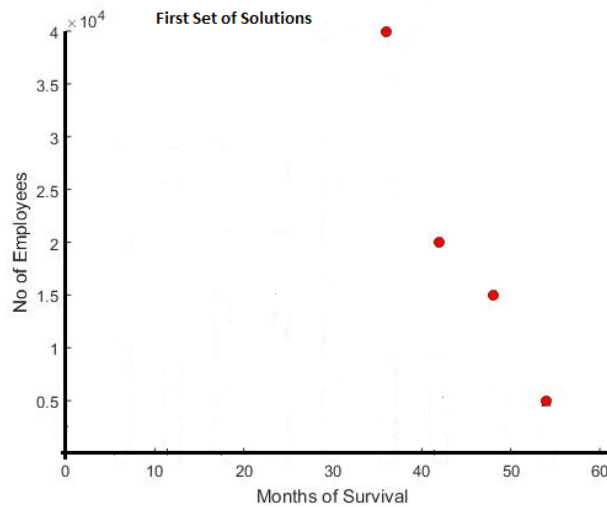


Figure 4.2.: First set of solutions

Now, the population is updated without the members of the first non-domination level, and the steps from step three of the algorithm are repeated to get the next set of non-dominated members, which is as follows:

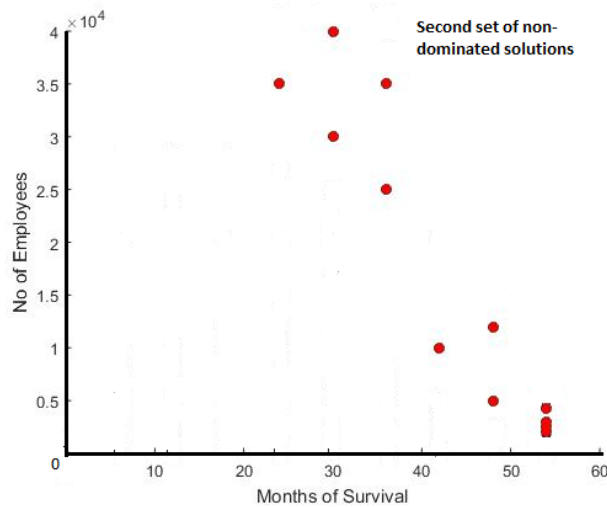


Figure 4.3.: Second set of solutions

Again, the population is updated, and the process is repeated until all the members are sorted to all the non-domination levels.

Using this algorithm, the data file with information from six hundred and forty-nine companies are sorted in 40 fronts.

And access to each company's variables within each Pareto front is achieved.

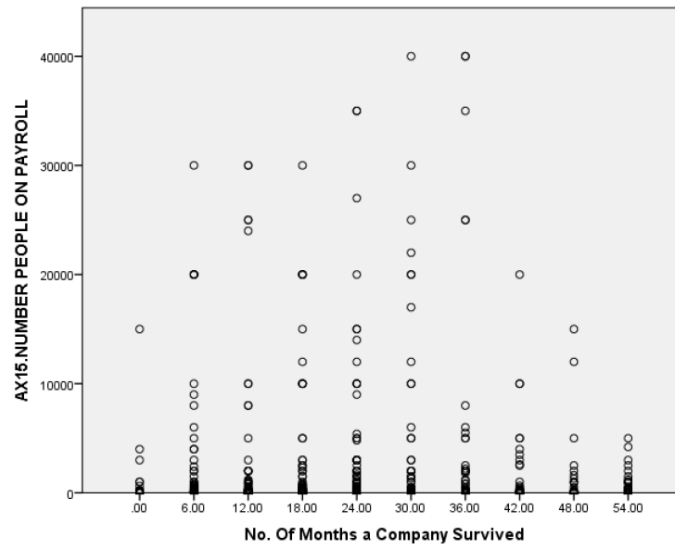


Figure 4.4.: Non-Dominated Sorting of Start-up Companies

Figure 4.4 shows all the companies plotted with respect to their objectives.

In the chapter Analysis, we can see Pareto-Optimal fronts and other Pareto fronts with their detailed analysis.



---

## 5. Analysis

So far in this study, we introduced the objectives and the independent variables for the multi-criteria optimization problem in chapter 2, followed by the non-dominated sorting algorithm used in this study to determine the Pareto fronts.

Now, in this chapter, the analysis of the acquired sorted Pareto-optimal front is done to measure the success factors which contributed to the success of the companies in this front compared to the not so successful companies in the later fronts. First, we will describe the software used to conduct the research, and then we will describe the method of analysis succeeded by Results.

### 5.1. Software used

For this study, the software used is SPSS and Matlab.

SPSS – Statistical Package for Social Sciences is a software package used for statistical analysis which was extensively used in the field of social sciences. With its versatile capabilities, its usage is extended in various other fields used by market researchers, scientists, data miners and many others. Complex statistical analysis on quantitative data sets can be performed using SPSS software package [Morrison, 1999]. This software allows users to create their own data, modify them and analyze these data, and also allows representing the analysis in the form of graphs and graphical representations [Morrison, 1999]. SPSS can be compared with other statistical software such as SAS, Stata, and S-Plus [Morrison, 1999].

In this study, the software is used to read the original data file from PSED which is explained in detail in the chapter PSED data set. The original file with 2000 variables was carefully analyzed to select the most useful variables for the research. With that being done, SPSS tool is further used to summarize three

companies from each front to conduct an inter-front and intra-front analysis which will be explained more in detail in this chapter.

Matlab - The selected variables from the PSED dataset in SPSS is then exported into an Excel file for use in Matlab for implementing the non-dominated sorting algorithm.

The Excel data file is loaded into Matlab, and the non-dominated sorting algorithm from chapter 4 is implemented to get all the companies sorted in different Pareto-fronts for the Objective functions defined in chapter two.

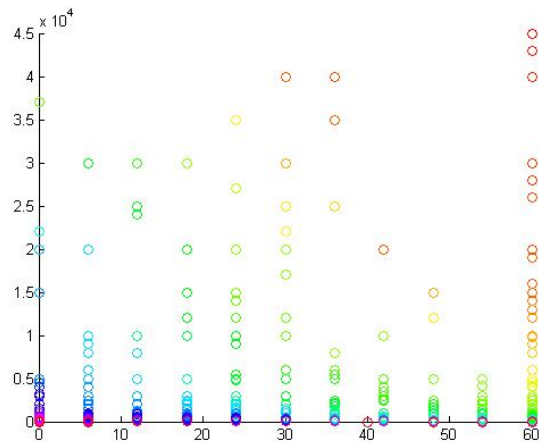


Figure 5.1.: Pareto Fronts Representation

In Matlab, using HSV (hue, saturation, value) function, all the fronts are distributed over RGB color range. So here front 1 starts from Red color and goes on until the last front is displayed in blue color. This view is just an example to display all the fronts at once.

On the other hand, in Matlab, the code was written in such a way that when executed, an input field would pop-up on the screen asking the user to enter the front number, and the result would be an image indicating members of that particular front. For example, when number 10 was entered, a picture with members of front ten would be displayed:

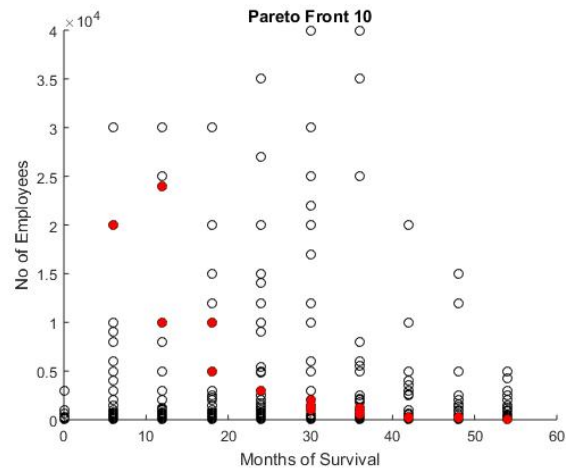


Figure 5.2.: Pareto Front 10

The circles in red are the members of front 10.

And when the input was number one, we get the Pareto Optimal Front.

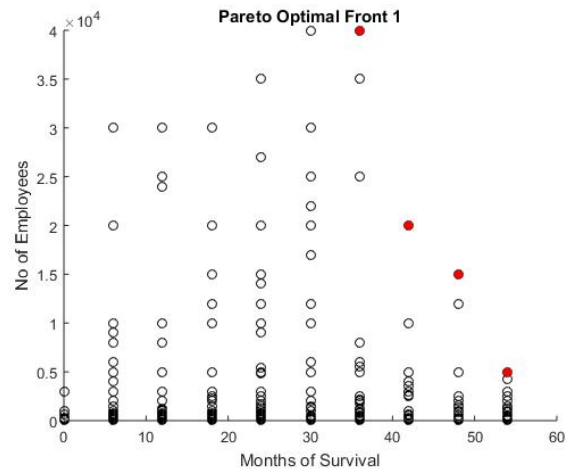


Figure 5.3.: Pareto Optimal Front

## 5.2. Method used for Analysis

Two approaches are used for in-depth analysis of fronts 1 and front 17 as one being the optimal front other being the front with very poor performing companies. And front 17 is considered as not much of a difference is seen from front 17 to front 40. I.e. all the members or the companies in these fronts are not very successful, and hence it makes sense to analyze success factors for front one compared to front 17. But as a future work, if one would like to analyze details from other fronts, one can do it as all the information pertaining to other members of other fronts can be retrieved quickly. The two approaches are:

1. Intra front analysis
2. Inter front analysis

### 5.2.1. Intra front analysis

In this form of analysis, three cases from front one and front 17 are analysed, three cases are selected such that a company each is chosen for its best objective, i.e. one company with the maximum number of employees in that front, one company with the utmost months it survived compared to other companies in that front and one company which is fairly good but not the best with respect to both the objectives.

For example, from front 17, one can see the selected three cases as in the figure 5.4. Similarly, three cases are selected for front one.

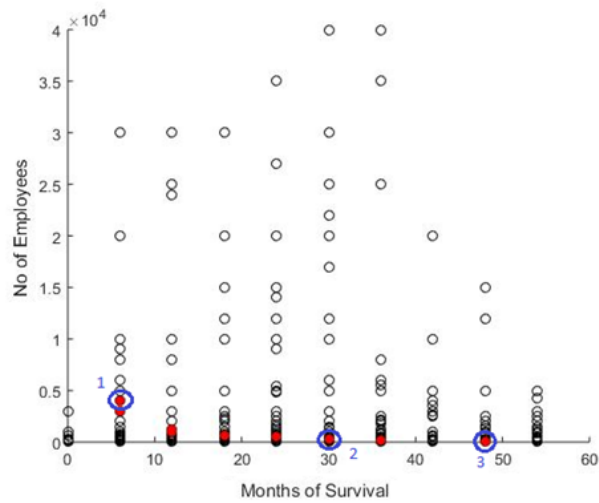


Figure 5.4.: Cases considered for Analysis

Now, in Intra-front analysis, the analysis is done in three steps:

1. Common variable of all the three cases are identified within front one
2. Unique features of each instance is outlined
3. Summary is drawn from the analysis

A similar structure is followed for front seventeen.

### Common variables from all the three cases in Front one

1. Owners of all three cases are post graduates
2. All Business being owned by single Owner
3. Independent Business creation type
4. Agree that Quality of the Product is very important to be an effective competitor
5. Serving those missed by others is Important
6. The social norms and culture of the community encourage entrepreneurial risk-taking.

### **Unique features of case 2**

1. Household Income was 100,000\$ greater than other two cases
2. The business received income for more than six months from the first twelve months whereas the other two cases failed to do so.
3. Revenue is greater than expenses. Other two cases did not have revenue greater than expenses.
4. Case two and three devoted more than 160 hours in the first twelve Months for the Business
5. Ph.D. or equivalent degree whereas other two cases had Master's degree
6. No Business offer same product or service.

### **Unique features of case 1**

1. Started promotional activities in the first six months whereas case two and three did not start.
2. First and third case started collecting information about their competitors in the first six months.
3. Market opportunities defined in the first six months for case one and three.
4. Case one agreed to have their product's price to be lower than the competitors for the success of the Business, but other two cases disagreed with this.
5. Superior Location is very important whereas other cases did not agree.
6. Some of the potential customers considered the product or service new and unfamiliar
7. No Business offer same product or service
8. 35 % National Customers expected

### Unique features of case 3

1. Financial Projections were already developed in the first six Months whereas case one considered it to be important in the future, case two considered it to be not relevant at this stage.
2. In terms of current work activity, not working for others for pay whereas other two cases worked for others.
3. Business received financial support after being registered as a legal entity.
4. Many Businesses offer same product or Service.
5. 90% National customers expected.

### Summary

It is interesting to see that all the three owners in front one are post graduates and social norms and culture of the community encourage entrepreneurial risk-taking. So, it can be said that social norms and the level of education can play a very crucial role for start-ups to be at their best as it influences the attitude and mind set of owners.

Apart from this, case two with its higher financial potential and the owner holding a Ph.D. degree and considering a product or service which is not being offered by any other company made good sales and was able to generate revenue greater than expenses. With these factors, it was able to be a successful start-up although not being the best with respect to any of the objectives.

Case one started promotional activities very soon and considered it is important that product's price should be lower than competitors and considered superior location very important for their business. Hence, it can be seen that this company had the attitude to grow faster and expenses were made on superior location and revenue generated initially was low, so although, it being a successful company, it could not survive to a greater time compared to case two and case three.

Case three gave priority and focused on financial projections at a very early stage, so it could be said that this could be the main reason for it to have a greater survival rate but many other businesses offer same service or product, and therefore growth of the company with respect to its employees was not the best compared to other cases.

**Common variables from all three cases in front 17**

1. All Business being owned by single Owner
2. Independent business creation type
3. The business received income through sales in the first six months.
4. Only one Physical location
5. Major items purchased in the first six months
6. Agree that Quality of the Product is very important to be an effective competitor
7. Serving those missed by others is Important
8. All agree being first to market a new product or service is important
9. All agree it's important to do better marketing job
10. All agree that developing new or advanced product technology or process technology for creating goods or services is important for this new business to be an effective competitor
11. All agree on bankers in their community help in new business start
12. Technology available more than one year ago

**Unique features of case 1**

1. Total household income: \$25000 to \$29000
2. Opportunities that led to business: High demand for products/business
3. Major problems starting the business: Acquiring other capital/Money; Financing
4. Preparation of business plan is considered not relevant whereas case two and case three started preparing the business models in the first six months
5. Product development was completed and ready to be sold in the first six months whereas case two and case three had prototype tested with customers



6. Promotional efforts considered not relevant whereas other two cases started promotional activities in the first six months
7. Collecting information about the competitors not relevant whereas other two cases collected
8. Not working for others for pay whereas others were working for others
9. Funds from personal saving: 1500\$
10. Total funding from all owners: 1680\$
11. Business is not hi-tech whereas it is high-tech for other two cases

### **Unique features of case 2**

1. Case one and three high school complete whereas case 2 attended some college.
2. Total household income: \$60000 to \$74000
3. Opportunities that led to business: Help others/ help community
4. Major problems starting the business: Regulations / Zoning
5. Financial projections developed by case two in the first six months whereas financial projections yet to be developed by other two cases.
6. Funds from personal savings: 2000\$
7. Total funding from all owners: 62000\$
8. Many businesses offering same product/services whereas very few offering same product or service for case one and three

### **Unique features of case 3**

1. Total household income: \$35000 to \$39000
2. Opportunities that led to business: Work experience/ Knowledge of field
3. Major problems starting the business: Acquiring contacts/connections
4. Decision to start the business came first whereas for case one and two got the idea to start

5. Market opportunities is not defined in the first six months whereas it was defined by case one and two
6. Disagree that lower prices are important for the business to succeed whereas other two cases agree.
7. Agree that Government supported for the start of new business whereas other two did not agree
8. Funds from personal saving: 3000\$
9. Funds from personal bank loan: 1000\$
10. Total funding from all owners: 4000\$

### **Summary**

From all the three cases, it can be said that the factor "major problems faced starting a business" could be the reason for all the three cases failing at an early stage. Financing being the major challenge for case one, Regulations for case two and acquiring contacts for case three.

Case one had the product ready to be sold to customers within the first six months but interestingly considered preparing a business plan, starting promotional activities and collecting information regarding competitors as not relevant. These might be the factors which were shadowed by case one which led to early failure of the start-up. Also, it should be noted that owner's education is high school complete.

Case two surprisingly had fewer reasons to fail, in fact, it is hard to point out what caused this start-up to fail apart from the barrier of Regulations and Zoning to start the business. Having completed some college, with a good intention to help the community and with great house hold income, case two decided to start this start-up, but unfortunately, one other reason for this start-up to fail could be it considered the product/service which was being offered by many other businesses.

For case three, it is worth noting that decision to start the business was the motivation to start rather the idea which would solve a problem should have been the motivation. Additionally, factors like unable to acquire contacts, market opportunities not defined might be the critical factors responsible for the start-up to fail.

### 5.2.2. Inter Front Analysis

In this analysis technique, the cases of Front 1 are compared with cases of Front 17 to deduce relation between defined objective functions and the success factors. So here, we compare each case of front one with its identical case of front 17. And later we conclude with the effect of the success factors of each of the comparisons in achieving a certain objective.

#### **Similarities between Front 1(all cases) and Front 17 (all cases)**

1. All business being owned by single owner
2. Independent business creation type
3. Agree that Quality of the Product is very important to be an effective competitor
4. Serving those missed by others is Important

#### **Summary**

Here we can see that three cases in front 1 compared to three cases in front 17 have above-stated factors in common. As expected, there are no much of similarities.

The one important difference in front one from front 17 is that all the cases in front 1 were post graduates where as in front 17, three cases were school completed (2 cases) and some college (one case).

Also in all the three cases in front one, social norms and culture of the society encourage entrepreneurial risks whereas this was not the case in front 17.

#### **Analysis of Front 1 case 1 (Survived 36 Months with 40,000 employees) and front 17 case 1(Survived 6 Months with 4,000 employees)**

Here the main reasons or factors which can be drawn to differentiate the progress of front 1 case 1 compared to front 17 case 1 are:

Table 5.1.: Front 1 vs 17 case 1 Analysis

<b>Variables</b>	<b>Front 1 case 1</b>	<b>Front 17 case 1</b>
Education	Post Graduate	Completed School
Income from sales	No income in first six months but received income and maintained consistency after six months	Received some income in the first six months but failed to survive after that
Revenue	Revenue greater than expenses	Expenses greater than revenue
Opportunities led to business	New Technology, Product/Service	Extra time available and taking advantage of the opportunity
Business plan	Started in the first six months	Considered irrelevant
Competitors	Collected information about competitors in the first six months	No information collected about competitors
Market Opportunities	Defined in the first six months	Not defined in the first six months
Working for others	Worked for others for pay during the first six months	Did not work for others
Superior location	Important	Not Important

From the table 5.1, it can be easily depicted that Front 1 case 1 were always ahead with their plans and executions whereas Font 17 case 1 contradicts Front 1 in all the stated variables.

Here, one can declare the success factors for case one of Pareto-optimal front to be:

1. Education
2. Income from sales
3. Revenue
4. New technology as the Opportunity which led to business
5. Business plan
6. collecting information from competitors
7. Market Opportunities
8. Location
9. Promotional activities
10. No business offer same product

All these factors with a positive and steady growth of some factors is an indication for a successful start-up.

#### **Common variables between Front 1 case 1 and Front 17 case 1**

1. Both Independent Creation type
2. Both started promotional efforts in the first six months
3. Both agree that lower prices are important for the business to be important competitor
4. Both strongly agree that quality of products or services are important for the business to be successful.
5. Both agree serving missed by others is important
6. Both agree it's important to do better marketing job

Table 5.2.: Front 1 vs 17 case 2 Analysis

Variables	Front 1 case 2	Front 17 case 2
Education	Post Graduate	Some college
Revenue	Revenue greater than expenses	Expenses greater than revenue
Opportunities led to business	Able to buy property, Passion for business	Help others, help community
Competitors	Not relevant	Information about competitors collected
Market Opportunities	Not relevant	Defined in the first six months
Amount invested for loan or savings	10,000\$	62000\$
Business offering	No business offer same business	Other business offer same business
New Business	Not hi-tech	Hi-tech

- Both agree The technical and scientific expertise of the start-up team is important for this new business to be an effective competitor

It can be estimated that to achieve comparatively high growth (number of employees), one should take care of the factors mentioned above.

Here, the success of a company keeping one objective in mind is analyzed. Concentrating on promotional efforts, considering lower prices than other competitors, serving those customers who were missed by others, and doing better marketing job, one can expand the company by the number of employees which is one of the objectives.

**Analysis of Front 1 case 2 (Survived 48 Months with 15,000 employees) and front 17 case 2(Survived 30 Months with 260 employees)**

Here, from table 5.2 it's important to see that clearly the reason why case 2 deserves to be in front 1 is the amount invested from personal savings is low compared to the one from front 17 and business is not high tech and is not being offered by others which are contrary to case 2 front 17.

It could also be the reason for its stable performance in terms of objectives. Did not survive the highest number of months and did not have the greatest number of employees. A trade-off between both the objectives can be imagined here.

Hence, the success factors for case two of the Pareto-optimal front can be deduced as:

1. Education
2. Passion for business and able to buy a property
3. Revenue
4. Amount invested from loan or savings to be minimal
5. No business offering the same business
6. Not a hi-tech business
7. Devoting more than 160 hours per month

#### **Similarities between Front 1 case 2 and Front 17 case 2**

1. Both received Income from sales in the first six months.
2. Both agree that quality of products or services are important for the business to be successful.
3. Both agree serving missed by others is important
4. Both agree it's important to do better marketing job
5. Both agree The technical and scientific expertise of the start-up team is important for this new business to be an effective competitor
6. Both Technology available one year ago

#### **Summary**

From the above similarities, it can be said that the approach is very calculative and risk is at minimal. So for any company to gradually progress, it is advisable to consider a technology which is already in the market, do good marketing job and see to it that income is generated through sales from the start.

Table 5.3.: Front 1 vs 17 case 3 Analysis

<b>Variables</b>	<b>Front 1 case 3</b>	<b>Front 17 case 3</b>
Education	Post Graduate	High School complete
Opportunities led to business	To make money	Work experience
Income from sales	Started receiving income after six months	Received income in the first 6 months but failed in the second and third year
Promotional efforts	Started after six months	Started in the first six months
Market Opportunities	Defined in the first six months	Not defined in the first six months
Financial Projections	Developed in the first six months	Not developed
Superior location	Not important	Important
Development of advanced technology	Not Important	Important
Financial support	Received after business registration	3000\$ invested from personal savings before registration
Product offered by others	Many	Few

**Analysis of Front 1 case 3 (Survived 54 Months with 5,000 employees) and front 17 case 3(Survived 48 Months with 45 employees)**

Here again, from table 5.3, some of the variables from front 1 case 3 are more favoring and practical compared to Front 17 case 3.

The success factors for case three of the Pareto-optimal front can be deduced as:

1. Education
2. Income from sales
3. promotional efforts
4. Market Opportunities
5. Financial Projections
6. Financial support



7. Devoted more than 160 hours per month

### **Similarities between Font 1 case 3 and Font 17 case 3**

1. Independent Business Type
2. Both collected information about competitors
3. Both Disagree that lower prices are important for the business to be important competitor
4. Both agree that quality of products or services are important for the business to be successful.
5. Both agree serving missed by others is important
6. Both agree being first to market a new product or service is important
7. Both agree it's important to do better marketing
8. Technology required for this business available more than one year ago for both the businesses
9. Business considered Hi-Tech

### **Summary**

For a company to survive for a longer time with comparatively smaller employees, it is important to consider a business with technology being available more than one year ago, do better marketing job, be informed about competitors, have lower prices for the product/ Service and to serve the market which has not been considered by others.



---

## 6. Conclusion and Future Work

### 6.1. Conclusion

The objective of this study to identify the successful start-ups from a real-time data set and identifying and analyzing the success factors of these start-ups was achieved through the application of Multi-criteria optimization supported by the non-dominated sorting algorithm.

Two objectives were defined for the multi-criteria optimization problem, and all the start-up companies with their relevant factors or variables were sorted based on the defined criteria using a non-dominated sorting algorithm. Pareto optimal front with the most successful companies was generated followed by all other members in different sets or fronts.

Intra and Inter front analysis were conducted to identify the success factors and summaries are written depicting how success factors helped companies in Pareto-optimal front to be successful compared to companies in front number seventeen. Also, factors which are responsible for companies to fail in front seventeen were identified.

The Inter-front analysis also tries to identify certain success factors with which a start-up can achieve success with respect to one objective compared to other and vice-versa.

## 6.2. Future Work

This study contributed to generalized analysis giving the opportunity for conducting specific analysis in the future such as how a success factor affects other factors, how changing certain parameters of a success factor can make a difference in the overall performance of a start-up.

Other methods of computational science can be used to predict the success of a new start-up with its variables based existing data.

Prediction of a start-up's success can be revolutionary in this area of research.

# Appendices



---

## A. PSED Dataset Variables

Information here is retrieved from PSED codebook [Curtin, 2012].

<b>Variable Name</b>	<b>Variable Label</b>
QS7	EDUCATION
INCOME	INCOME SUMMARY
QFF3	OWN ALL OR PART OF NEW BUS
QFF4	BUS RECEIVED MONEY/INCOME
QFF5	MONTHLY REVENUE > EXPENSES
QFF6	SALARY INCLUDED IN CALC OF EXPENSES
AA5a	AA5.OPPORTUNITIES LED TO BUSINESS (1)
AA5b	AA5.OPPORTUNITIES LED TO BUSINESS (2)
AA6a	AA6.MAJOR PROBLEMS STARTING BUS (1)
AA6b	AA6.MAJOR PROBLEMS STARTING BUS (2)
AA7	AA7.CAME 1ST: BUS IDEA OR DECIS TO START
AA10	AA10.BUSINESS CREATION TYPE
AE13	AE13.BUS RECEIVED INCOME FROM SALES
FE13	FE13.BUS RECEIVED INCOME FROM SALES
BA37	BA37.PAST 12 MO WORKED ON BUS 160 HOURS
AB10	AB10.BUSINESS LOCATION: ONE/SEVERAL/NONE
AC1	AC1.CURRENT LEGAL FORM OF BUSINESS
AD1	AD1.PREPARATION OF BUS PLAN STARTED
AD6	AD6.STAGE OF PRODUCT DEVELOPMENT
AD9	AD9.PROMOTIONAL EFFORTS STARTED
AD11	AD11.BUSINESS DEVELOPED PROPRIETARY TECH
AD13	AD13.APPLICATION FOR PATENT SUBMITTED
AD16	AD16.MAJOR ITEMS PURCHASED/LEASED/RENTED
AD20	AD20.DISCUSSION W/POTENTIAL CUSTOMERS
AD22	AD22.COLLECT INFO ABOUT COMPETITORS
AD24	AD24.MARKET OPPORTUNITIES DEFINED

## A. PSED Dataset Variables

---

AD26	AD26.FINANCIAL PROJECTIONS DEVELOPED
AE3	AE3.FIRST FUNDING FROM FINAN INST/PEOPLE
AE5	AE5.CREDIT W/SUPPLIER BEEN ESTABLISHED
AE7	AE7.EMPLOYEES/MANAGERS/SUBCON HIRED
AE9	AE9.PEOPLE WORKING AT LEAST 35 HRS/WK
AE11	AE11.BANK ACCOUNT OPENED FOR BUSINESS
AE24	AE24.BUS HAS PHONE/INTERNET CONTACT INFO
AF1	AF1.IMPORTANT: LOWER PRICES
AF2	AF2.IMPORTANT: QUAL PRODUCT/SERVICES
AF3	AF3.IMPORTANT: SERVING MISSED BY OTHERS
AF4	AF4.IMPORTANT: BEING 1ST MARKET NEW PROD
AF5	AF5.IMPORTANT: DOING BETTER MKTING JOB
AF6	AF6.IMPORTANT: SUPERIOR LOCATION
AF8	AF8.IMPORTANT: TECH/SCIENTIFIC EXPERTISE
AF9	AF9.IMPORTANT: DEVELOP NEW PROD TECH
AF10	AF10.IMPORTANT: DEVELOP PATENT/COPYRT
AG1	AG1.WHO WILL OWN NEW BUSINESS
AH6_1	AH6_1.HIGHEST LEVEL OF EDUCATION
AH15_1	AH15_1.WORKING FOR OTHERS FOR PAY
AH19_1	AH19_1.PRIMARY ROLE IN THE NEW BUS
AH25_1	AH25_1.PROVIDED TRAINING
AH27_1	AH27_1.PROVIDED PHYSICAL RESOURCES
AP2	AP2.SOC NORMS: EMPHASIZE SELF SUFFIENCY
AP3	AP3.SOC NORMS: ENCOURAGE ENTREPR RISKS
AP7	AP7.GOV'T SUPPORT FOR STARTING NEW BUS
AP8	AP8.BANKERS HELP NEW BUS STARTED
AQ4_1	AQ4_1.FUNDS: PERSONAL SAVINGS
AQ5_1	AQ5_1.FUNDS: LOANS FROM FAMILY
AQ6_1	AQ6_1.FUNDS: LOANS FROM FRIENDS
AQ7_1	AQ7_1.FUNDS: CREDIT CARD LOANS
AQ8_1	AQ8_1.FUNDS: PERSONAL BANK LOAN
AQ9_1	AQ9_1.FUNDS: ASSET BACK LOAN
AQ10_1	AQ10_1.FUNDS: OTHER SOURCES
AQ12x_1	AQ12x_1.TOTAL AMOUNT OF FUNDING
AQ12_1	AQ12_1.TOTAL FUNDING AMOUNT CORRECT
AQ13_1	AQ13_1.AMOUNT OF FUNDING LOANED
AQ12x	AQ12x.TOTAL FUNDING AMOUNT ALL OWNERS



---

AQ13	AQ13.AMOUNT OF FUNDING LOANED ALL OWNERS
AR1	AR1.BUS RECVD FORMAL FINANCIAL SUPPORT
AR4	AR4.AMT INVESTED AFTER REGISTERED
AR32	AR32.ALL FUNDS IN NEW BUS BANK ACCOUNT
AS1	AS1.PRODUCT/SERVICE NEW OR UNFAMILIAR
AS2	AS2.OTHER BUS OFFERING SAME PRODUCT
AS3	AS3.TECH AVAILABLE MORE THAN 1 YEAR AGO
AS6	AS6.NEW BUS HI-TECH
AS9	AS9.PERCENT OF CUSTOMERS NATIONAL
SUST_000	00 MTH AFTER ENTRY:START-UP STATUS
SUST_006	06 MTH AFTER ENTRY:START-UP STATUS
SUST_012	12 MTH AFTER ENTRY:START-UP STATUS
SUST_018	18 MTH AFTER ENTRY:START-UP STATUS
SUST_024	24 MTH AFTER ENTRY:START-UP STATUS
SUST_072	72 MTH AFTER ENTRY:START-UP STATUS
SUST_001	01 MTH AFTER ENTRY:START-UP STATUS
SUST_030	30 MTH AFTER ENTRY:START-UP STATUS
SUST_036	36 MTH AFTER ENTRY:START-UP STATUS
SUST_042	42 MTH AFTER ENTRY:START-UP STATUS
SUST_048	48 MTH AFTER ENTRY:START-UP STATUS
SUST_054	54 MTH AFTER ENTRY:START-UP STATUS
SUST_060	60 MTH AFTER ENTRY:START-UP STATUS
AR22	AR22.TOTAL MARKET VALUE OF BUS IF SOLD
Months	No. Of Months a Company Survived
AX15	AX15.NUMBER PEOPLE ON PAYROLL

Table A.1.: Start-up Variables and Labels from PSED dataset



---

## B. Variable values

Information here is retrieved from PSED codebook [Curtin, 2012].

<b>Value</b>	<b>Label</b>
QS7 1	Eighth grade or less
2	High school incomplete
3	High school complete
4	Some college
5	Associates degree
6	Bachelors degree
7	Postgraduate degree
99a	Refused
INCOME 1	Under \$15,000
2	\$15,000-\$24,999
3	\$25,000-\$29,999
4	\$30,000-\$34,999
5	\$35,000-\$39,999
6	\$40,000-\$49,999
7	\$50,000-\$59,999
8	\$60,000-\$74,999
9	\$75,000-\$99,999
10	\$100,000 or more
21	Less than \$40,000
22	\$40,000 or more
99a	Refused
QFF3 1	All
2	Part
3	None
98a	DK
99a	Refused
QFF4 1	Yes

*B. Variable values*

---

2	No
98a	DK
99a	Refused
QFF5 1	Revenue greater than expenses
2	Expenses greater than revenue
98a	DK
99a	Refused
QFF6 1	Yes
2	No
98a	DK
99a	Refused
AA5a 10	Low overhead
11	Low cost property; have property
12	Low cost supplies or services
13	Tax write-off
14	Low start-up costs
19	Other cost references
20	Have resources: saved up to do it
21	Have cash backing; have large investors
22	Loan or grant
23	Sold home, property or business
29	Other current financial resources
30	Income; to make money
31	Extra income
32	Need to supplement income
33	Investment
34	Financial independence; job security
35	Income for educational expenses
36	Income for retirement
37	To leave business/money to children
39	Other income references
40	Good business idea
41	Take advantage of opportunity
42	High demand for products/business
43	Market opportunity; untapped market
44	New technology/product/service
45	Good product; faith in product

---

46	Expansion of old/current business
47	Vast resources or material
48	Able to buy building, property, business
49	Other business opportunity references
50	Be own boss; tired of working for others
51	Flexibility; more free time; set hours
52	Stay home with children; work from home
53	Potential more money working for self
54	Cannot find employment; lost job
55	Further career; career change
56	Retired
59	Other employment references
60	Work experience; knowledge of field
61	Have formal training/education in field
62	Enjoy work, have passion for it
63	Have talent in field, area of expertise
64	Change in personal situation
65	Inheritance
69	Other personal references
70	Just decided to do it; boredom
71	Self-fulfillment; always wanted to do
72	Timing is right; time in life; extra time
73	Easy; does not require a lot
79	Other lifestyle references
80	Had mentor; discussed with a mentor
81	Encouraged to start business, industry
82	Encouraged to start business, community
83	Approached to start business, net/phone
84	Mentors available and willing
85	Business partner's influence
86	Encouraged
89	Other mentor references
90	Can do better than the competition
91	Help others; help community
92	Aid in economy; economic development
93	Good location; easily accessible
98a	DK

*B. Variable values*

---

99a	NA
AA5b 10	Low overhead
11	Low cost property; have property
12	Low cost supplies or services
13	Tax write-off
14	Low start-up costs
19	Other cost references
20	Have resources: saved up to do it
21	Have cash backing; have large investors
22	Loan or grant
23	Sold home, property or business
29	Other current financial resources
30	Income; to make money
31	Extra income
32	Need to supplement income
33	Investment
34	Financial independence; job security
35	Income for educational expenses
36	Income for retirement
37	To leave business/money to children
39	Other income references
40	Good business idea
41	Take advantage of opportunity
42	High demand for products/business
43	Market opportunity; untapped market
44	New technology/product/service
45	Good product; faith in product
46	Expansion of old/current business
47	Vast resources or material
48	Able to buy building, property, business
49	Other business opportunity references
50	Be own boss; tired of working for others
51	Flexibility; more free time; set hours
52	Stay home with children; work from home
53	Potential more money working for self
54	Cannot find employment; lost job
55	Further career; career change

---

56	Retired
59	Other employment references
60	Work experience; knowledge of field
61	Have formal training/education in field
62	Enjoy work, have passion for it
63	Have talent in field, area of expertise
64	Change in personal situation
65	Inheritance
69	Other personal references
70	Just decided to do it; boredom
71	Self-fulfillment; always wanted to do
72	Timing is right; time in life; extra time
73	Easy; does not require a lot
79	Other lifestyle references
80	Had mentor; discussed with a mentor
81	Encouraged to start business, industry
82	Encouraged to start business, community
83	Approached to start business, net/phone
84	Mentors available and willing
85	Business partner's influence
86	Encouraged
89	Other mentor references
90	Can do better than the competition
91	Help others; help community
92	Aid in economy; economic development
93	Good location; easily accessible
98a	DK
99a	NA
AA6a 10	Acquiring information on business costs
11	Cost of location; rent, lease, etc.
12	Costs of equipment
13	Costs of services
14	Supplies
15	Transportation
16	Taxes; insurance
17	Start-up costs
19	Other cost references

---

*B. Variable values*

---

20	Acquiring information on financing
21	Acquiring new owner capital
22	Acquiring new bank loan
23	Acquiring other capital/money; financing
24	Interest rates
29	Other capital/financing references
30	Research information on laws/regulations
31	Registration; licensing
32	Regulations; zoning
39	Other gov't/legal barrier references
40	Researching competitors
41	Price competition
42	Market competition
43	Competition
44	Economic Conditions
49	Other market/competitors references
50	Researching labor markets/wages
51	Hiring labor
52	Training labor
53	Market wages; pay scales
59	Other labor references
60	Researching the targeted market
61	Product marketing
62	Customer marketing/ sales
63	Advertising; marketing
69	Other marketing references
70	Researching product
71	Product/service development
72	Product/service distribution
79	Other product/service development
80	Acquiring information business plans
81	Developing a business plan
82	Forecasting future costs
83	Scheduling/time management
84	Accounting
85	Acquiring location
86	Acquiring supplies



---

87	Receiving timely payments
88	Acquiring contacts/connections
89	Other business decision references
90	Acquiring experience/education
91	Lack of motivation
92	Disability
93	Discrimination; race, age, gender
94	Family obligations
95	Other personal references
96	None
98a	DK
99a	NA
AA6b 10	Acquiring information on business costs
11	Cost of location; rent, lease, etc.
12	Costs of equipment
13	Costs of services
14	Supplies
15	Transportation
16	Taxes; insurance
17	Start-up costs – NFS
19	Other cost references
20	Acquiring information on financing
21	Acquiring new owner capital
22	Acquiring new bank loan
23	Acquiring other capital/money; financing
24	Interest rates
29	Other capital/financing references
30	Research information on laws/regulations
31	Registration; licensing
32	Regulations; zoning
39	Other gov't/legal barrier references
40	Researching competitors
41	Price competition
42	Market competition
43	Competition
44	Economic Conditions
49	Other market/competitors references

---

*B. Variable values*

---

50	Researching labor markets/wages
51	Hiring labor
52	Training labor
53	Market wages; pay scales
59	Other labor references
60	Researching the targeted market
61	Product marketing
62	Customer marketing/ sales
63	Advertising; marketin
69	Other marketing references
70	Researching product
71	Product/service development
72	Product/service distribution
79	Other product/service development
80	Acquiring information business plans
81	Developing a business plan
82	Forecasting future costs
83	Scheduling/time management
84	Accounting
85	Acquiring location
86	Acquiring supplies
87	Receiving timely payments
88	Acquiring contacts/connections
89	Other business decision references
90	Acquiring experience/education
91	Lack of motivation
92	Disability
93	Discrimination; race, age, gender
94	Family obligations
95	Other personal references
96	None
98a	DK
99a	NA
AA7 1	Idea First
2	Decision First
3	Both
8a	DK

---

9a	NA
AA10 1	Independent
2	Takeover
3	Franchise
4	Marketing Initiative
5	Sponsored New Business
98a	DK
99a	NA
AE13 1	Yes
5	No
8a	DK
9a	NA
FE13 1	Yes
5	No
8a	DK
9a	NA
BA37 1	Yes
5	No
8a	DK
9a	NA
AB10 1	One Physical Location
5	Several Physical Locations
6	No Specific Location
8a	DK
9a	NA
AC1 1	Sole Proprietorship
2	General Partnership
3	Limited Partnership
4	Limited Liability Corporation or LLC
5	Sub Chapter S Corporation
6	General Corporation
96	Not Yet Determined
98a	DK
99a	NA
AD1 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant

---

*B. Variable values*

---

8a	DK
9a	NA
AD6 1	Completed and ready for sale/delivery
2	Prototype/procedure tested with customers
3	Model/procedure is being developed
4	Still in the idea stage; no work done yet
8a	DK
9a	NA
AD9 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AD11 1 Yes	
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AD13 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AD16 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AD20 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AD22 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK

---

9a	NA
AD24 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AD26 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AE3 1	Yes
5	No
8a	DK
9a	NA
AE5 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AE7 1	Yes
2	Not Yet, Will in Future
5	No, Not Relevant
8a	DK
9a	NA
AE9 999998a	DK
999999a	NA
AE11 1	Yes
2	No, not yet; will in future
5	No, not relevant
6	Using exist account
8a	DK
9a	NA
AE24 1	Phone
2	Internet
3	Both
6	Neither

---

*B. Variable values*

---

8a	DK
9a	NA
AF1 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF2 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF3 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF4 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF5 1	Strongly Agree
2	Agree
3	Neither

---

4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF6 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF8 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF9 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA
AF10 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
6	Not Relevant
8a	DK
9a	NA

---

B. Variable values

---

AG1 1	Self Only
2	Self and Spouse
3	Self and Other
AH6_1 1	Up to eighth grade
2	Some high school
3	High school degree
4	Technical or vocational degree
5	Some college
6	Community college degree
7	Bachelors degree
8	Some graduate training
9	Masters degree
10	Law, MD, PHD, EDD, Degree
98a	DK
99a	NA
AH15_1 1	Yes
5	No
8a	DK
9a	NA
AH19_1 1	General management; everything
2	Sales/marketing/customer service
3	Finance/accounting
4	Technical/research/science/engineering
5	Manufacturing/operations
6	Administration/human resource management
8a	DK
9a	NA
AH25_1 1	Yes
5	No
8a	DK
9a	NA
AH27_1; 1	Yes
5	No
8a	DK
9a	NA
AP2 1	Strongly Agree
2	Agree



---

3	Neither
4	Disagree
5	Strongly Disagree
8a	DK
9a	NA
AP3 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
8a	DK
9a	NA
AP7 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
8a	DK
9a	NA
AP8 1	Strongly Agree
2	Agree
3	Neither
4	Disagree
5	Strongly Disagree
8a	DK
9a	NA
AQ4_1 999999998a	DK
999999999a	NA
AQ5_1 999999998a	DK
999999999a	NA
AQ6_1 999999998a	DK
999999999a	NA
AQ7_1 999999998a	DK
999999999a	NA
AQ8_1 999999998a	DK
999999999a	NA
AQ9_1 999999998a	DK

*B. Variable values*

---

999999999a	NA
AQ10_1 999999998a	DK
999999999a	NA
AQ12x_1 999999999a	NA
AQ12_1 999999998	DK
999999999	NA
AQ13_1 999999998a	DK
999999999a	NA
AQ13 999999999a	NA
AR1 1	Yes
5	No
8a	DK
9a	NA
AR4 999999998a	DK
999999999a	NA
AR32 1	Yes
5	No
8a	DK
9a	NA
AS1 1	All
5	Some
6	None
8a	DK
9a	NA
AS2 1	Many
5	Few
6	No Other
8a	DK
9a	NA
AS3 1	Yes
5	No
8a	DK
9a	NA
AS6 1	Yes
5	No
8a	DK
9a	NA

---

AS9 998a	DK
999a	NA
SUST_000 to SUST_60 1.00	NEW FRM:PROFIT
2.00	ACTIVE SU
3.00	QUIT
9.00a	MISSING DATA
AR22 999999998a	DK
999999999a	NA
AX15 999998a	DK
999999a	NA
a - missing value	

Table B.1.: Start-up Variables and Value ranges from PSED dataset



---

# Bibliography

- [Alden, 2011] Alden, E. (2011). Primum non nocere: The impact of dodd-frank on silicon valley.
- [Cable, 2010] Cable, A. J. (2010). Fending for themselves: Why securities regulations should encourage angel groups.
- [Coello et al., 2007] Coello, C. A. C., Lamont, G. B., Van Veldhuizen, D. A., et al. (2007). *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer.
- [Cole, 2009] Cole, C. W. (2009). Financing an entrepreneurial venture: Navigating the maze of corporate, securities, and tax law. *UMKC L. Rev.*, 78:473.
- [Curtin, 2012] Curtin, R. (2012). Panel study of entrepreneurial dynamics ii: Codebook.
- [Deb, 2001] Deb, K. (2001). Multi-objective optimization using evolutionary algorithms, 2001. *Chichester, John-Wiley*.
- [GovtIndia, 2016] GovtIndia (2016). Start-up India start-up india govt. project. <http://www.startupindia.gov.in/>. Accessed: 2017-02-27.
- [Maxville, 2013] Maxville, V. (2013). Introducing: Computational science. *Procedia Computer Science*, 18:1456–1465.
- [Morrison, 1999] Morrison, E. (1999). Introduction to spss. *Center for Social Science Computation and Research, Universidad de Washington*.
- [Neely, 2002] Neely, A. (2002). *Business performance measurement: theory and practice*. Cambridge University Press.
- [Oranburg, 2016] Oranburg, S. C. (2016). Start-up financing. *Browser Download This Paper*.

- [Peng, 2011] Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060):1226–1227.
- [Reynolds and Curtin, 2009] Reynolds, P. D. and Curtin, R. T. (2009). *New firm creation in the United States: Initial explorations with the PSED II data set*, volume 23. Springer Science & Business Media.
- [Reynolds and Curtin, 2012] Reynolds, P. D. and Curtin, R. T. (2012). PSED i, ii harmonized transitions, outcomes data set.
- [Ries, 2011] Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Business.
- [Taticchi, 2010] Taticchi, P. (2010). *Business performance measurement and management*. Springer.

# Declaration

I, Nithish Hulikanthe Math declare that this Master Thesis is a work composed solely by myself and I confirm with the best of my knowledge that apart from the indicated resources, no other materials are used in the completion of this work.

Nithish Hulikanthe Math

Magdeburg, March 6, 2017