Empirical applications of Big Data techniques on start-up companies in Catalonia

Group: GCE01

Authors: Rui Bao, Manel Castro, Victor Lluva and Ji Qi
Tutor: Carolina Luis Bassa
Academic year: 2017-2018

Universitat Pompeu Fabra Barcelona

Index

1. Abstract

Recent improvements in techniques for analysis and usage of Big Data allow us to extract key conclusions from a vast amount of data. Applying this knowledge in the business world has proven to be beneficial in some cases for company performance indicators and profitability, while being a waste of resources in others. In this study we analyse a sample of 182 start-up companies in Catalonia by classifying them into three company types based on their business model and dependence on what we define as Big Data Techniques. We present the situation of this businesses by analysing the characteristics of each of the three company types. A final section concludes with our results and their implications, as well as flaws and further research possibilities.

Keywords: Big Data, Business classification, Business model, Business Intelligence and analytics, Catalonia, Start-up Business.

2. Theoretical framework

a. Business Intelligence

Using data about business performance to improve revenue results can be traced back to 1865, when Richard Miller Devens described the success of the banker Henry Furnese, in his work "Cyclopaedia of Commercial and Business Anecdotes" (1865). This was the first time the term Business Intelligence (BI) was mentioned. However, it was not until almost a century later that the modern concept of "Business Intelligence System" would be proposed: "[...] an automatic system is needed which can accept information in its original form, disseminate the data promptly to the proper places and furnish information on demand" (Luhn, 1958).

Around the decade of the 1990's, the increasing awareness of this concept and the dramatic drop in the cost of computational technologies lead to its popularity among firms. This first phase is widely considered as the Business Intelligence 1.0, where data was produced and stored in commercial relational database management systems (RDBMS). Once stored it could be managed and organized in a presentable way. Starting in the early 2000, HTTP-based Web 1.0 systems characteristic of Business Intelligence 2.0 allowed for the collection of vast amounts of detailed and IP-specific user search and interaction logs, which opened a new world of unique analytical and research opportunities. During recent years, in the era of Business Intelligence 3.0, the number of mobile phones and tablets surpasses the number of laptops and PCs (The Economist, 2011). This new era promises more ways than ever before to obtain and analyse information. As Chen and his colleagues point out: "[...] The ability of such mobile and Internet-enabled devices to support highly mobile, location-aware, person-centered, and context-relevant operations and transactions will continue to offer unique research challenges and opportunities throughout the 2010s" (Chen et al., 2012).

The increasing volume and variety of data generation characteristic of BI 2.0 and BI 3.0, measured in the scale of terabytes or even exabytes, and its consequential technological requirements, have led to the recent emergence of the Big Data era, unveiling new exciting applications in many spheres of society (Chen et al., 2012). In the field of E-Commerce and Market Intelligence, BI and Big Data analytics are shown to be high return investments to analyse data and learn useful insights to support in the decision processes of the company, leading to changes in organizational, management and production settings in order to increase performance on many levels such as: productivity (Tambe and Hitt (2012), Bloom et al. (2012)), improve sales forecast results (Bajari et. al. 2012), design personalized promotions (McAfee and Brynjolfsson, 2015), capturing the long tail market (Anderson 2014), ultimately translating to increased profits and survival rates (Farah 2017).

b. Big Data

Digital information, including unstructured and multi-structured data, is the composition of Big Data. It is usually derived from intersections between people and machines (Arthur 2013).The key characteristic of "big data" is "big", which means, compared with traditional analytics, " the volume of stored data exceeds human analytic capacity and pushes the boundaries of currently-available computing power" (Gutmann, Merchant and Roberts, 2018). The increasingly economical data-intensive approach makes it possible for the formation of modern Big Data. The defining characteristics of Big Data are summarised in the three V's of Doug Laney (Lanley 2001): volume, variety and velocity.

Regarding to Volume, our fast-track life is transformed in real time into data. In business analysis, the data you have right now will be replaced soon by new incoming data in several minutes, or even seconds. The variety refers to the multiplicity of data sources. With our smartphones, laptops and other mobile devices, for example, with IPad, we use social networks to connect with our friends (Social media data), glance over web pages to search for what we are interested in (Web activities tracking), take photos and share our locations with GPS when we travel (Location data). Vast data from various dimensions make huge volume and high variety the two main characteristics of modern Big data. Moreover, in order to extract the insights and conclusions from this data, the requirements for high velocity in dealing with such great amount of information are becoming higher and higher.

During the Big Data Innovation Summit in 2013, IBM proposed a fourth dimension: Veracity. When the sources of data are multiple, it is likely that the veracity of data sometimes can be doubtful. The quality of the data becomes an essential factor when dealing with large data bases. Inaccurate or incomplete data would eventually lead to making the wrong decisions.

With the rapidly developing technology, the cost of producing, storing and sharing of data is much lower than ever before, Big Data turns into a more powerful tool to explore business profitability and business niches. Collecting and analysing a large amount of data is not a new activity in business analysis. Aggregate data were extensively used in studies since nineteenth century. Nevertheless, the modern Big Data with four "Vs" creates a revolution. We explain this with the following two instances.

Firstly, in the paper "First-Degree Price Discrimination Using Big Data", it shows that large database on individual behaviour provides sufficient information to estimate individuals' reservation values and makes profitable personalize pricing available in practice, that means, firms can extract more profits by using first-degree price discrimination and set a different price schedule to each individual with the help of Big Data. Siller found that the feasible personalized pricing generates about 12.2% of increase in profits when web browsing behaviours are employed to predict individuals' reservation values while only 0.8% with demographics.

Secondly, in the aspect of marketing, as Lisa Arthur wrote in her book: "Big data marketing centers on one thing and one thing only: driving value by engaging customer more effectively". Data-driven marketing techniques not only engage current customers, but also capture new customers, which leads to higher market shares for companies. What's more, using data helps market decisions to be more accurate and faster, as a result, generates higher business benefits and reduces costs. (Figure 1)

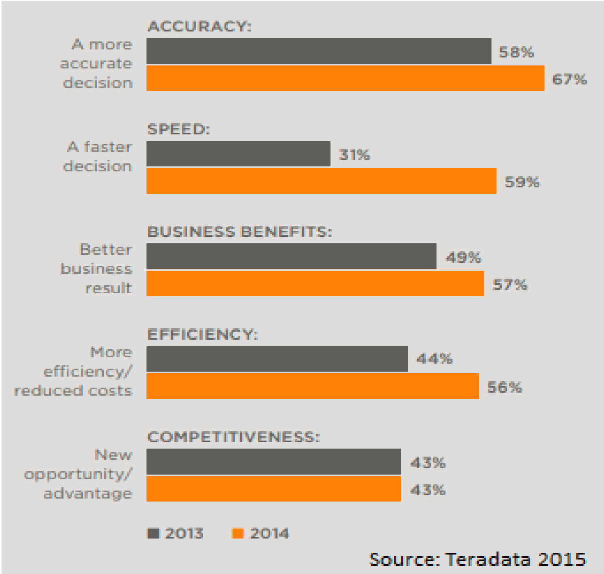

Figure1 Benefits of using data in making decisions

Already in 2014, Gartner's Hype Cycle for emerging technologies (Figure 2) placed Big Data on its way down the peak of inflated expectations, with an estimated time of reaching the plateau of productivity of 5 to 10 years. Four years later, it would be highly engaging to observe the results of companies implementing this technology in their operations.

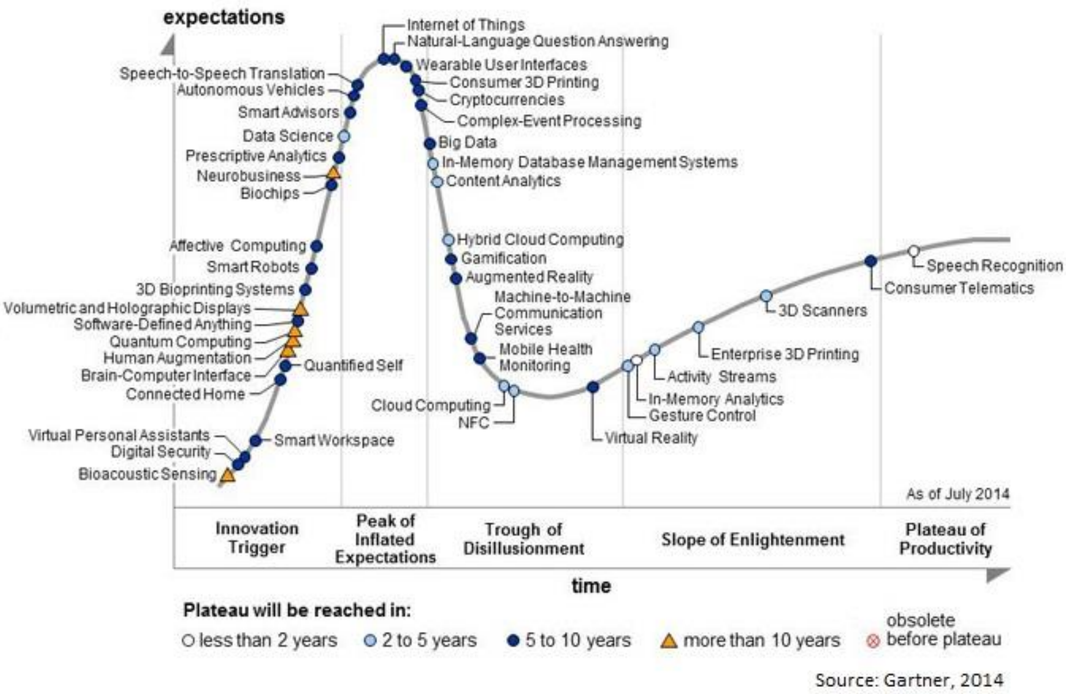

Figure 2. Gartner's Hype Cycle for emerging technologies 2014

3. Big Data in Catalonia

a. Area of study

Our goal in this study is to classify and analyse an extensive list of startup businesses in Catalonia. For this purpose, first we need to have a clear definition of the word “start-up”. According to Merriam-Webster, “start-up” means “the act or an instance of setting in operation or motion” or “a fledgling business enterprise.” Since there is no official or technical definition of startup businesses, empirically, we choose those companies that have less than 10 years since their foundation, less than 50 employees and less than 20 million of capital (equity). These three criterias imply that these companies are in the first stage of their operations. We focus on these startup businesses for several reasons. Big Data techniques, which implies Business Intelligence and Analytics techniques, have been playing an increasingly important role in business for the last 10 years. New companies created during this period may have already started using some of these techniques to gain insights from their business activities and improve their performance. Moreover, startups companies are engines for job-creation, innovation and economic growth (Y. Lowrey, 2009).

The reasons we choose Catalonia as our target of study are that we have better access to the relevant data, and better understanding of its economic and business environment. It is also worthy to mention that Barcelona was recently classified as one of the most dynamic technology-driven cities in Europe (Vanguardia 2017, Invest Europe 2017).

b. Definition of Big Data Techniques

In our attempt to define what we refer to in this study as “Big Data Techniques”, we encounter the disagreement that exists on the definition of Big Data, as reviewed in Ward & Barker (2013). We attempted to make our own definition of Big Data using the 4 V’s theory (Gartner 2001 & IBM 2009). When you are treating with Big Data you have to go away from standard database and algorithms designs because your data is really complex to be treated due to the volume and velocity that it’s collected, the variety is needed to be treated and the veracity level at which it has to be assured. What really differentiates Big Data is the complexity in solving the 4 V's problem, which requires to be in the border line of the top technology. Most of the times this leads to the need of customized databases and algorithms in order to adapt to the specific conditions of each case. For example the way Google Maps have tackled its big data problem, the car optimization of GPS routes using the available and real time data obtained from the cars in circulation is really different from what Twitter is doing, attempting to comprehend on an aggregate level the feelings of its users about a specific topic. Both cases require really different types of databases and algorithms, to such extent that the techniques used have to be customized for that specific dataset. The constant development of this techniques implies a ever-changing scenario of new and more capable algorithms and methods. This doesn't exclude the fact that in some cases Big Data firms live behind top Big Data firms and copy their procedures (or the procedures made by a Phd in Computer Sciences), these techniques are so closer to the edge of technology that also have to be recognized as Big Data Techniques, and in these cases always require up to a certain level of customization.

For the purposes of this study and for the present state of the existing technology, the definition Big Data techniques would include: schema-free data storage techniques such as NoSQL databases, extended RDBMS and MapReduce or YARN methods and machine learning techniques. But, what will happen once these cutting-edge technologies become the norm? Data generation is becoming faster and faster over time and, probably, someday these techniques will become the "standard methods" in order to treat data. The question is, is this still going to be Big Data? The answer is no. At this moment the requirements for storing and analysing data would be much higher than what current technologies are made for, so "Big Data" term has to catch up to the new technologies. In essence, Big Data is the need to keep pace with border line technologies due to the increasing demands in the treatment of the data because of volume, variety and velocity and guaranteeing the veracity of the data.

c. Classification

In order to study the how the different Big Data Emerging techniques have been applied in

Catalonia, we focus on the companies that operate with a strongly data-driven business model.

We extract a list of startup businesses founded in Catalonia that claim to engage in Business Intelligence and Analytics techniques. The database we use is from the Catalonia Trade & Investment webpage (Barcelona & Catalonia Startup Hub, retrieved on June 2018). This public agency aims to promote Catalan startup companies by attracting foreign investment and offering a range of financial and advisory services. The webpage contains a platform with a large collection of Catalan startup companies, and introductory information about each one. With this information we construct a database of 182 companies.

An initial inspection reveals that the big data based startup environment in Catalonia is rich and diverse. In the last few years, dozens of companies have been created with the aim of implementing, each in its sector, the potential of big data management that so much interest has generated. In order to describe this diverse environment, we face the problem of market atomization. Each company dedicates its business to a different sector, and adds value in its own unique and innovative way. In this sense, we believe that a rigorous taxonomy based on a sectoral analysis would not be the appropriate way to present this information. First because the extent of the different sectors would require us have each activity sector concisely and clearly defined, in such way that each sector would unambiguously represent the business activity. We were unable to find such a satisfactory sectoral definition and defining them ourselves exceeds the purposes of this study. Second, even if the different sectors are eventually well defined, the classification would result in a large amount of very specific activity sectors forming a network derived from more general sector types. Moreover, we would see some sectors with very few companies. These problems lead us to conclude that a sectoral classification would poorly satisfy our objective of describing the startup environment in Catalonia and its future development possibilities.

In order to overcome these difficulties, we formulate the following classification method to sort our list of 182 companies in the following three main groups.

The first group would be composed by companies that specialize in the management of big data. In other words, these are companies whose activity is only big data analysis and big data technologies development. Business that need big data analysis on their business activity usually subcontract these services to this first group of companies, since these have much more advanced techniques and specialized personal. Among these companies we could find *Kernel Analytics*, which works basically as a consulting company providing B2B services. Another example would be *Atomial*, which focuses on the very act of centralizing and filtering previously distributed information.

The second group would be composed by companies that have a business model that it's not big data by itself but use and depend on big data technology. Here the key term is "dependent", the company has an independent activity that needs big data to survive. In contrast with Type 1 firms, Type 2 have their own business and product, which happens to strictly and crucially depend on big data techniques to be innovative and competitive. This kind of companies doesn't need to have a internal department of big data, they can externalise the activity to other companies from group one. Examples of that would be *IBT Index*, a company that measures effort based on hundreds of data measures, or *Counterest* that uses big data to count people in massive public events.

The third group is a kind of companies that use big data for improve its company performance, but its business activity doesn't depend on the big data techniques, so they could survive without it. In other words, they do not depend on Big Data techniques to function, neither to be competitive, nor to be innovative. But the product or service they offer strongly suggests that big data techniques are being used to add value to their business performance. Normally this kind of companies don't use big data regularly but intermittently. The companies that can belong to this group are companies that use marketing analytics tools like Google Trends or Semrush.

The essence of this classification aims to determine how important big data is for the business activity of each company, the first group uses big data as the source of value in the business model, while the third one just uses big data to add some value into their main activity. All three groups can be related with big data, although it's the first one that has the closest perspective of what big data really is. It is also important to highlight the difference between group one and groups two and three, that is while the first group is a big data provider, the second and third group are big data users. This effectively solves the problems arising from the sectoral classification and allows us to meaningfully describe the different applications of emerging technologies. Table 1 summarizes our classification.

Company types	Criteria		
	Uses Big Data techniques	**Depends** on Big Data techniques	**Provides** services to other firms
Group 1	Yes	Yes	Yes
Group 2	Yes	Yes	No
Group 3	Yes	No	No

Table 1. Different criteria of the company groups

After our classification, we found that the vast majority of companies in our list belong to this third group since their activities are not actually dependent on big data technologies. One reason for this situation can be derived from the signalling theory. Companies claiming to involve Big Data techniques in their decision processes can give a better image of themselves by signalling commitment and willingness to engage in such emerging technologies, thus potentially attracting more investments and improving brand image. Startup companies would gain a significant advantage by soaking themselves with any techniques even remotely related to Business Intelligence or Big Data. The credibility of such claims would depend on a number of factors, which include transparency or the tangible results of such investments, but fundamentally the degree of knowledge both managers and investors have about this emerging technology. When classifying the companies, we often encounter this problem. Most group 3 companies claim to be using Big Data techniques, but from the data we have available, we cannot know for certain what are the exact applications they have in their business activities, so we basically assume that they use data-driven tools for marketing purposes and customer analysis.

One could think that Facebook, as a famous company, belongs to the first group, since its activity is basically big data and data treatment, and it's true that maybe big data and data obtaining is its main source of revenue. However, it is not its main business activity nor its main source of value, the business is not focused on big data, but rather on social relations on internet, they could subcontract its big data activity and the business could still exist. Thus Facebook belongs to group two, since their business activity strictly depends on big data. For example, Facebook needs Big Data techniques to properly recommend new friends among the vast quantity of registered users in the platform, without which, if it was uneasy to find a friend, Facebook could never be what it is today. However, this reasoning is recognizably a little subjective, and this is due to the term "depend". Whether a business depends on big data is really difficult to say, it could only be determined by analysing sufficiently deep into each case and retrieving the necessary information, but the time and effort this task requires exceeds the limits of this study, so instead the analysis is done via reasoning over the data we are able to access. For each company, we infer whether it depends or not on Big Data Techniques based on its activities. One could argue that Facebook would survive without a properly friend recommendation feature, thus without dependence on Big Data techniques. In that case such techniques would only add value to the company, so Facebook would belong to group three. This example illustrates how relative our classification can be, especially among groups two and three. The fact that new companies (not so focused on big data) are starting to implement this technology into their business plans indicates that it is gaining momentum. This is a good thing from the perspective of its proponents, but it can also have bad consequences.

More specifically, the technology may still be not mature enough as to be adopted by the general business environment. Working with big data technology requires specific and highly technical human capital that is nowadays scarce on the labour market. For this reason, companies may be assuming risks when trying to implement those technologies. McKinsey & Company (Fleming et al., 2018; Chui et al., 2018) has a series of very recent work regarding this issue for multinational companies, and how companies investing in advanced analytic techniques should be aware of the real impact they may have.

d. Data Analysis

1. Data Descriptives

	N	Minimum	Maximum	Mean	Std. Deviation
Type	181	1	3	1.91	.777
Number of Employees	150	5.00	50.00	14.0000	14.01102
Capital Raised	150	250000.00	200000000.0	3098333.333	16598897.65
Age	174	1.00	10.00	4.8563	2.21582
Target of the Company	155	.00	1.00	.6323	.48375
Valid N (listwise)	118				

Table 2. Variable general descriptives

The descriptives table (Table 2) will provide us a brief insight of the variables we have about our sample of data-driven startups. From top to down, our variables are as follows:

1. Type: Our classification model. This is a categorical variable that takes the integer values from 1 to 3, each option for a type of company in our classification.
2. Number of employees: Number of people reported to be working in the company. We will use this variable to proxy the size of the company.
3. Capital raised: The equity of each firm at the moment of us accessing to the database.
4. Age: Years passed since the foundation of the company
5. Target: Binary variable indicating 0 for companies that have a B2B business model and 1 for companies serving both to businesses and customers.

The number of companies we find in each type is fairly equal, however we have more type 1 companies than any other type. In general, companies have a mean of 14 employees, although the standard deviation is greater than the mean, which means that there is a great polarization between companies. The average capital is a little bit higher than 3 million €, but as we saw with the number of employees, the high standard deviation reveals that there are many companies with few capital and some companies that pool enormous quantities of capital. Startups in this sector are 4.85 years old on average, with a significantly lower standard deviation than in other cases (relative to the average).

Finally, there are more companies focused on business only than on business and consumers. Nevertheless, we should be cautious, as the standard deviation here is also high (almost as high as the average).

2. Correlations:

Creating a correlation table will show us the interactions between variables and, after interpreting them, we may be able to extract clear conclusions. Before creating this table we have standardized all variables. Having said that, it is clear that there are some significant correlations:

		Zscore(Type)	Zscore: Number of Employees	Zscore: Capital Raised	Zscore(Age)	Zscore: Target of the Company
Zscore(Type)	Pearson Correlation	1	-.102	.073	-.154*	-.199*
	Sig. (2-tailed)		.214	.378	.043	.013
	N	181	149	149	173	155
Zscore: Number of Employees	Pearson Correlation	-.102	1	.322**	.163	.139
	Sig. (2-tailed)	.214		.000	.051	.110
	N	149	150	137	145	134
Zscore: Capital Raised	Pearson Correlation	.073	.322**	1	-.023	.153
	Sig. (2-tailed)	.378	.000		.785	.079
	N	149	137	150	145	133
Zscore(Age)	Pearson Correlation	-.154*	.163	-.023	1	.082
	Sig. (2-tailed)	.043	.051	.785		.322
	N	173	145	145	174	149
Zscore: Target of the Company	Pearson Correlation	-.199*	.139	.153	.082	1
	Sig. (2-tailed)	.013	.110	.079	.322	
	N	155	134	133	149	155

*. Correlation is significant at the 0.05 level (2-tailed).

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

Table 3. Variables correlation panel

Although we have included many variables in our analysis, not many are correlated significantly with other variables. Nevertheless, there is one that is significantly correlated and that is size and capital raised, a result to be expected from the start. Other weaker but still significant correlations are Age with Type and Target with Type. First of all, type of company and age seem to be correlated. Type 1 companies are more likely to be older than type 2 and type 3 ones. Also related with the type of company there is the target of the company. Type 1 companies are more likely to be B2B enterprises. In other words, they are specialized in providing services to other companies. This is not surprising though, as we defined type 1 companies as having precisely these characteristics.

The following tables are descriptive statistics of the variables: size, capital, target and age controlled by the variable Type of company. We aim to determine whether there are differences in these variables depending on the type of company. Each variable has been segmented to include only values of one type of company.

3. Size (In employees):

	N	Minimum	Maximum	Mean	Std. Deviation
Size1	55	5.00	50.00	16.4545	15.08227
Size2	54	5.00	50.00	12.3148	12.12019
Size3	40	5.00	50.00	13.1250	14.83618
Valid N (listwise)	0				

Table 4. Number of employees by company type

Table 4 indicates the number of employees for companies of each type (i.e.: Size 1 = number of employees in companies of type 1). All three types of companies have different average sizes in employees. Type1 companies tend to be bigger than other types of companies. Nevertheless it is important to note that the standard deviations are very large (as large as the averages, or even larger) and that this tells that we should be careful in taking conclusions. If we were to perform a statistical test to know if there are significant differences between the averages the results would be obviously inconclusive (meaning that there are not significant differences).

4. Capital:

	N	Minimum	Maximum	Mean	Std. Deviation
Capital1	50	250000.00	20000000.00	2590000.000	4761066.593
Capital2	59	250000.00	20000000.00	1652542.373	3694128.890
Capital3	39	250000.00	5000000.00	961538.4615	1236158.179
Valid N (listwise)	0				

Table 5. Amount of capital raised by company type.

Table 5 shows the amount of capital different companies in each type have been able to raise over the years. This is one of the variables where all standard deviations are bigger than the averages. On the other hand, there are notorious differences in the averages, showing that type 1 companies have larger raised capitals than type 2 companies, and that both have raised on average more capital than type 3 companies. As a graphic example: type 1 companies have an average of 2.6 million € and type 3 companies of nearly 1 million€.

5. Age:

	N	Minimum	Maximum	Mean	Std. Deviation
Age1	58	1.00	10.00	5.5690	2.34043
Age2	69	1.00	8.00	4.3478	1.79727
Age3	46	1.00	10.00	4.7609	2.43297
Valid N (listwise)	0				

Table 6. Age by company type

Table 6 shows the years passed since the foundation of the company disaggregated by company type. In contrast with other variables, the standard deviations of age are lower than its averages. Type 1 companies are apparently older than other types of companies. Between type 2 and type 3 companies it appears to be no much difference, and statistically we could expect some significant difference between the averages.

6. Target:

	N	Minimum	Maximum	Mean	Std. Deviation
Target1	59	.00	1.00	.7288	.44839
Target2	60	.00	1.00	.6333	.48596
Target3	36	.00	1.00	.4722	.50631
Valid N (listwise)	0				

Table 7. Market target by company types

Finally, table 7 shows the market orientation of different companies in each company type. At a first sight, there appear to be differences in the target of companies. Nevertheless, it is important to note that there may be collinearity between the variable TYPE and this variable, because the first was designed taking the last one into account. In any case, type 3 companies are more focused in the consumer while the other two types (1 & 2) do the same but on other businesses.

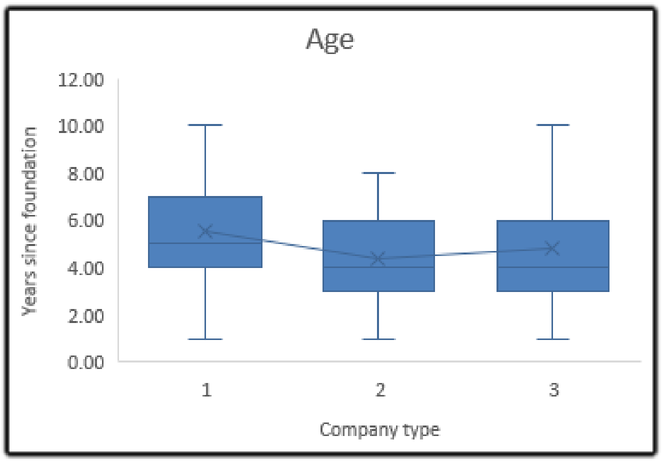

Figure 3. Years since foundation by company type

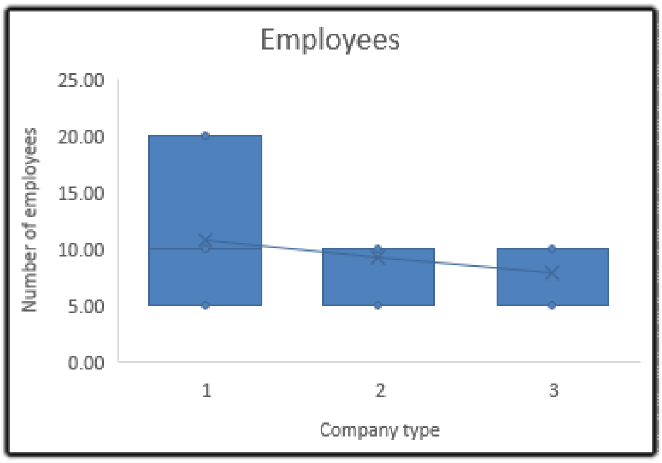

Figure 4. Number of Employees by Company Type

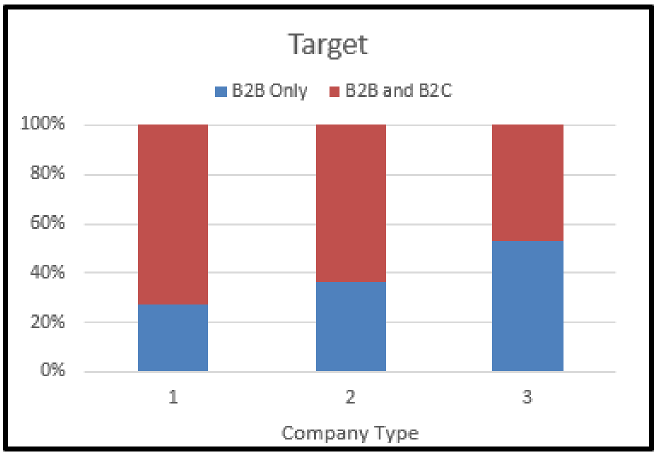

Figure 5. Market Target by Company Type

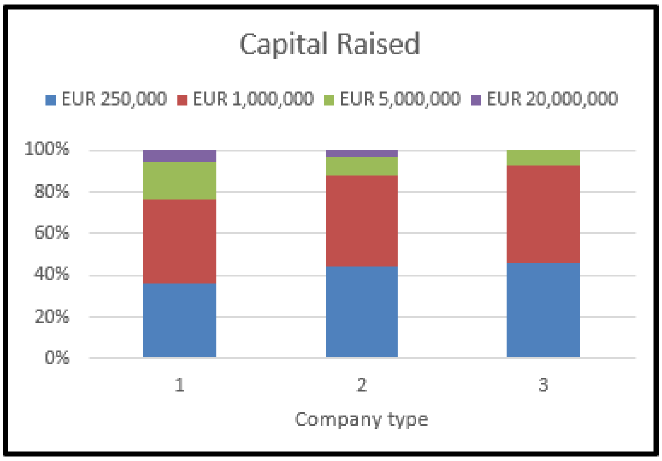

Figure 6. Total capital raised by Company Type

The situation of these companies is clearly summarized in the figures 3, 4, 5, and 6. We plot the different variables (Age, Size, Target, Capital) by company size. The box and whiskers plot for Age and Employees tries to give an indication of the strong variation our data has. The blue “x” mark in each box indicates the mean of each variable, while the blue line inside each box indicates the median. Piled bar charts for Target and Capital try to convey the percentage of each category a variable in each company type.

We can infer a series of insights from these figures. First, although not significantly, the average age of Type 1 companies is slightly higher than the other two types. This suggests that Type 1 companies were created before and was the first market demand to be satisfied. Further evolution and demand could have given place to the other types. Second, on average there are more people are employed in Type 1 startups than in other types.

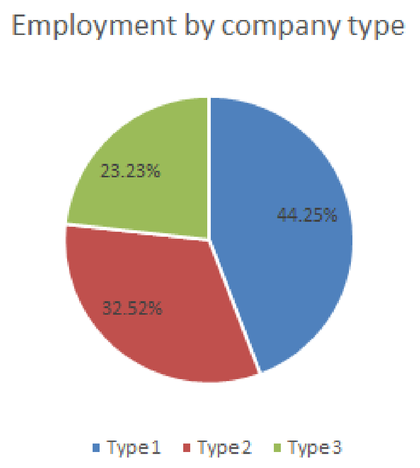

Figure 7. Percentage of people employed by company type

Figure 7 shows that on aggregate Type 1 companies create the most employment, followed by Type 2. The previous data suggests that the activities carried out in Type 1 companies are more labour-intensive than the others, requiring on average and in total more workers. Thirdly, regarding the market orientation, we observe that a surprising 72.41% of Type 1 companies provide third-party services not only to other businesses, but also to consumers. In fact, we see that Type 3 companies are 1,9 times more likely to work with businesses only than Type 1 companies (52.28% for Type 3 compared to 27.59% for Type 1). This result is probably due to the great demand there is for the consulting services of Type 1 firms, both from businesses and consumers, so that Type 1 companies have expanded many of their services to serve this new market niche. This would imply that Type 2 companies also have a significant penetration providing services to the customer segment.

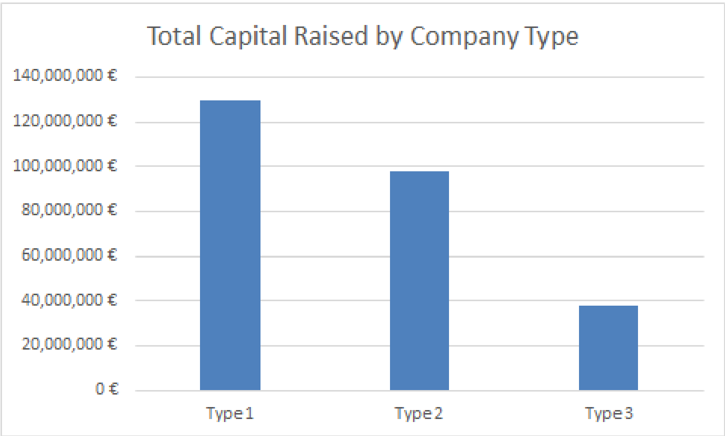

Figure 8. Total Capital Raised by company type

Regarding the capital raised, Figure 6 shows that Type 2 and type 3 companies have approximately the same capabilities for attracting investment. The composition of the capital raised suggests that Type 1 companies are the most attractive investment, followed by Type 2. On the aggregate, Figure 8 show that Type 1 companies have raised substantially more capital than all other types (129,500,000€). Its success is followed by Type 2 (97,500,000€), and Type 3 in a smaller scale (37,500,000€).

4. Conclusions

a. Results and implications

Along the years, companies rise and fall. Some achieve success and grow exponentially while others stagnate. During our project it was clear that the majority of companies in this industry end up stagnating. The important thing for us and we think that also for future research is why do they stagnate.

Our empirical analysis has found that type 3 companies are more likely to be newer (which indicates a trend in the industry) and also are more likely to be focused on both business and consumer. Finally, of all companies, type 3 companies are the statistically smallest ones. In other words, type 3 companies are getting more common in the big data industry, and they are more consumer focused than before. In contrast, Type 1 companies seem to be the most successful ones in terms of functioning years, market potential, employment creation and investment attraction. Further policies in this area may need to focus on this type of companies as they are gaining increasingly more importance. We found that type 2 companies tend to evolve related to other sectors different to big data, meaning that these companies do not focus their force in evolving into the big data but more about connecting big data to different sectors. It also appears to be the case that some type 2 companies appear with the objective in mind to ending up selling data from their specific sector, because it's also a way to obtain benefits apart from the main activity.

Another interesting finding in our scale is that type 1 companies seem to have market dominance trends or oligopolistic behaviour (few companies tend to dominate the market). This can be due to them taking advantage of the specialization and in this sector they can take a huge amount of scale economics from specializing. But still there is a lot of work left to do in order to corroborate this theory. We encourage further research to expand the study to different geographical areas, with more variables and different approaches.

We recognize our study has a series of flaws. First, there may be many causal relationships between these characteristics and other variables that escape from our analysis. We must interpret the correlations cautiously and with common sense. The main problem listing startups is that you never know when the startups are actually active. A second problem is that most of the times the startups have a webpage and apparently is active but some indicators reveal signs of inactivity, for example when the level of turnover remains unknown for a long time, or if the size of the company remains low over the time. This can be due to principal reasons. On one hand, the startup is working in a passive or seasonal way, normally the second job of a computer engineer, and it is waiting for an eventful project. On the other hand, the business has become inactive recently but the webpage remains active because the domain is paid for a longer period.

b. Future research

Having analysed the big data sector in Catalonia has provided us with a rich understanding of the industry. Nevertheless, our model can also be applied to other geographic regions, which would provide important insights and validity comparisons with our study. A disaggregated sectorial study can be performed just as it has been done for Catalonia. Also, our classification could be replicated to include not only startups but also economy wide companies of different sectors and sizes, to get a more comprehensive picture of the state of the art of Big Data applications. A more complete database would also be generally more accessible when studying with larger companies.

In order to improve the quality of our data a qualitative approach could be done. Our classification model relies heavily on the strategic information we can infer from the activities of companies. This is not necessarily true and may vary on subjective views. Further research needs to address this issue by studying in more depth each company. This could be achieved with interviews, surveys and conversations with management roles or large samples of consumers. Companies can be asked about financial data, feelings of the market, expectations, past experiences, among others. This research approach would provide much richer information of this market, and would avoid considering inactive companies with active websites.

5.Bibliography

Anderson, C. 2004. *The Long Tail*. WIRED Magazine (12:10) Available at (http://www.wired.com/wired/archive/12.10/tail.html).

Arthur L. Big data marketing: engage your customers more effectively and drive value[M]. John Wiley & Sons, 2013

Barcelona & Catalonia Startup Hub. List of startups. (n.d.). Retrieved June 12, 2018, from http://startupshub.catalonia.com/list-of-startups

Bajari, P., Chernozhukov, V., Hortaçsu, A., & Suzuki, J. (2012). The Impact of Big Data on Firm Performance - An Empirical Investigation. IDC Government Insights, GI237315(October), 1–12. https://doi.org/10.11871/j.issn.1674-9480.2014.03.002

Bloom, N., Sadun R., and Reenen J.V.. 2012. "Americans Do IT Better: US Multinationals and the Productivity Miracle." American Economic Review, 102 (1): 167-201. DOI: 10.1257/aer.102.1.167

IBM What is big data? Bringing big data to the enterprise. http://www-01.ibm.com/software/data/bigdata/, July 2013.

L. Douglas. 3d data management: Controlling data volume, velocity and variety. Gartner. Retrieved, 6, 2001.

Catalunya una region lider en Start-ups en Europa. (February 23rd, 2017). In Spanish. Retrieved From: http://www.lavanguardia.com/monograficos/catalunya-impulsa/catalunya-una-region-lider-en-start-ups-en-europa

Chen, H., Chiang R. H. L., & Storey, V. C. Business Intelligence and Analytics: From Big Data to Big Impact, MIS Quarterly, Special Issue: Business Intelligence Research, 36(4), pp. 1165 – 1188, December 2012.

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chun, R., Nel, P., Malhotra, S. (April 2018). Notes from the AI frontier: Applications and value of deep learning. *Mckinsey & Global Institute discussion papers.* Retrieved June 12, 2018 from: https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning

Economist, The. 2011. "Beyond the PC," Special Report on Personal Technology, October 8 (http://www.economist.com/node/21531109).

Farah, B. (2017). Profitability and Big Data. Journal of Management Policy and Practice, 18(2), 47–53.Retrieved from https://search.proquest.com/docview/1949456868/fulltextPDF/BFAAF42D8B27483BPQ/21?accountid=13631

Fleming, O., Fountaine, T., Henke, N., Saleh, T. (May 2018). Ten red flags signaling your analytics program will fail. *Mckinsey & Company discussion papers.* Retrieved June 3, 2018 from: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ten-red-flags-signaling-your-analytics-program-will-fail

Gartner's 2014 Hype Cycle for Emerging Technologies Maps the Journey to Digital Business. (n.d.). Retrieved June 12, 2018, from https://www.gartner.com/newsroom/id/2819918

Gutmann M P, Merchant E K, Roberts E. "Big Data" in Economic History[J]. The journal of economic history, 2018, 78(1): 268-299.

IEEE Engineering Management Review, *41*(2), 14–22. https://doi.org/10.1109/EMR.2013.2259978

Infographic: The Four V's of Big Data | IBM Big Data & Analytics Hub. (n.d.). Retrieved June 12, 2018, from http://www.ibmbigdatahub.com/infographic/four-vs-big-data

Invest Europe (May 7th, 2017), The Acceleration Point: Why Now is the Time for European Venture Capital. Retrieved June 5, 2018 from https://www.investeurope.eu/media/664430/InvestEurope_Acceleration-Point_VCReport_05072017.pdf

Jelinek, M., & Bergey, P. (2013). Innovation as the strategic driver of sustainability: Big data knowledge for profit and survival.

Laney, D (2001). 3D Data management: Controlling Data Volume, Velocity and variety. Application Delivery Strategies. Metagroup. File 949. Available at https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Lowrey, Y. (2009). Startup business characteristics and dynamics: A data analysis of the Kauffman Firm Survey. *SSRN Electronic Journal*, (350), 59. https://doi.org/10.2139/ssrn.1496545

Luhn, H. P. (1958). A Business Intelligence System. IBM Journal of Research and Development, 2(4), 314–319. https://doi.org/10.1147/rd.24.0314

McAfee, A. & Brynjolfsson, E. (2015). *Big Data:The Management Revolution*. Harvard Business Review, October 2012, 1–10.

Richard Miller Devens, *Cyclopaedia of Commercial and Business Anecdotes: Comprising Interesting Reminiscences and Facts, Remarkable Traits and Humors ... of Merchants, Traders, Bankers ... Etc. in All Ages and Countries*, D. Appleton, 1868.

Shiller B R. First-degree price discrimination using big data[J]. 2014. Available at Ben Shiller's webpage at http://benjaminshiller.com/images/First_Degree_PD_Using_Big_Data_Jan_18,_2014.pdf

Tambe, Prasanna and Lorin Hitt, "The productivity of Information Technology Investments: New Evidence from IT labor data," Information Systems Research, 2012, 23(3-part-1), 599-617.

Terradata. (2015). *Progressing Toward True Individualization*. *Forbes insight*.

The Four Essential V's for a Big Data Analytics Platform - Dataconomy. (n.d.). Retrieved June 12, 2018, from: http://dataconomy.com/2015/06/the-four-essentials-vs-for-a-big-data-analytics-platform/

Ward, J. S., & Barker, A. (2013). Undefined By Data: A Survey of Big Data Definitions. Retrieved from http://arxiv.org/abs/1309.5821

Appendix 1. List of selected startup companies from the Barcelona & Catalonia Startup Hub (startupshub.catalonia.comlist-of-startups)

Type	Company	Size (employees)	Capital	Web	Headquarters	Year of foundation	Tags	Target
2	123compare.me	< 5	< 250k	123compare.me	Barcelona	2017	Ecommerce & Trade, Saas, Software licence	B2B
2	4dlife	< 10	< 1 million	www.4d-life.com	Barcelona	2010	Consultancy & Agency, Ecommerce & Trade, Marketplace, R+D+I Services, Saas	B2B and B2C
1	8wires	< 5	< 250k	8wires.io	Barcelona	N.A.	Consultancy & Agency, Saas	B2B
2	9ineSports	< 20	< 1 million	www.9inesports.com	Barcelona	2016	Advertising, Ecommerce & Trade, Freemium	B2B and B2C
1	Ad-Pure	< 5	< 1 million	www.ad-pure.com	Barcelona	2014	Software Licence, subscription	B2B
3	AeInnova	< 10	< 1 million	www.aeinnova.com	Terrassa	2014	Devlopment and Manufacturing	B2B
2	AVUXI	< 5	< 1 million	www.avuxi.com	Barcelona	N.A.	Saas, Subscription	B2B
3	Adictik	< 10	< 250k	www.adictik.com	Barcelona	2013	Advertising, Consultancy & Agency	B2B
3	Adver2play	< 5	< 250k	www.adver2play.com	Barcelona	2017	Advertising, Subscription	B2B
1	Aistech Space	N.A.	N.A.	www.aistechspace.com	Castelldefels	2015	Aerospace	B2B
1	Aniling	< 10	< 1 million	www.aniling.com	Badalona	2014	R+D+I Services, Other types of license	B2B and B2C
1	Apimatica	< 5	< 250k	www.apimatica.com	Sant Cugat del Vallès	2008	Consultancy & Agency	B2B
2	App&Town	< 10	< 250k	www.appandtown.com	Cerdanyola del Vallès	2012	Marketplace, Saas, Software Licence	B2B and B2C
1	Ascidea	< 10	< 1 million	www.ascidea.com	Sant Cugat del Vallès	2011	R+D+I Services, Saas, Subscription	B2B and B2C
2	Atlantis Fleet	< 10	< 1 million	atlantis-technology.com	Barcelona	2014	Freemium, Saas, Subscription	B2B and B2C
1	Atomian	< 20	< 1 million	www.atomian.com	Barcelona	2014	Software license	B2B
2	Autodiscovery	< 5	< 250k	www.butlerscientifics.com aboutus	Barcelona	2014	Saas, Biopharma	B2B
2	BAGmovies	N.A.	< 250k	www.isnotTV.com	Barcelona	2014	Advertising, Software License, Mediatech	B2B and B2C
1	Bcare	< 5	< 250k	www.bettercare.es	Barcelona	2010	R+D+I Services, Sofware Licence, Medtech	B2B
2	Bet4talent	< 5	< 1 million	www.bet4talent.com	Barcelona	2014	Subscription, Human tech	B2B
2	BiBold	N.A.	< 250k	www.bibold.com	Barcelona	2014	Consultancy & Agency, Mobile Software	B2B
2	Billage	< 10	< 1 million	www.billage.com	Barcelona	2013	Subscription, Fintech, Mediatech, Mobile Software	B2B
2	BillyMobile	N.A.	< 250k	www.billymob.com	Barcelona	2015	Advertising, Consultancy & Agency	B2B and B2C
3	Bioprognos	N.A.	N.A.	www.bioprognos.com	Barcelona	2016	Biotech & Pharma	B2B

1	Bismart	< 50	< 5 million	www.bismart.com	Barcelona	2009	Consultancy & Agency, Systems Integration, IoT	B2B
1	Biwel	< 5	< 1 million	www.biwel.cat	Santa Cristina d'Aro	2014	Wearables, HelthTech, IoT, Wellness	B2B and B2C
1	Boardfy			www.boardfy.comenhome	BCN and Vigo	2016		
2	Boira Digital	< 10	< 1 million	www.boiradigital.com	Barcelona	2016	Blockchain, Mobile Software, Saas	B2B
2	Buy Yourself	< 5	< 250k	www.buyyourself.io	Barcelona	2015	Fashion & Design, Marketplace	B2B and B2C
2	Buytrendy	< 5	< 250k	www.buytrendyapp.com	Barcelona	2016	Ecommerce & Trade, Logistic Tech	B2C
2	CIGO!	< 20	< 1 million	www.smart-cigo.com	Barcelona	2010	Business	B2B
2	Crono Analytics	N.A.	N.A.	www.businessintelligence.es	Barcelona	N.A.	Productivity Tech, Industry 4.0	B2B
1	Capit Mobile	< 5	< 250k	www.capitmobile.com	Barcelona	N.A.	Mobile Software, Productivity Tech, Consultancy	B2B
1	Carritus	< 10	< 1 million	www.carritus.com	Barcelona	2009	Foodtech & Drinks, Ecommerce	B2C
3	Catuav	N.A.	N.A.	www.catuav.com	Moià	2009	Film & Video, Hardware, Drones (UAV)	B2B and B2C
2	CeliCity	< 5	< 250k	www.celicity.com	Barcelona	2014	Adtech & Marketing, Foodtech, Marketplace	B2B and B2C
2	Clinical Document Engineering	< 5	< 250k	www.clinicaldocumentengineering.com	Barcelona	2015	Cloud Computing, Healthtech, Software	B2B and B2C
1	Codiwise	N.A.	N.A.	www.codiwise.com	Salt	2013	Adtech & Marketing, Mobile Software	B2B and B2C
2	ConnecThink	< 5	< 250k	connecthink.eu	Barcelona	2016	AI, Cloud Computing, Software, Freemium	B2B and B2C
2	Cooncert.com	< 5	< 1 million	www.cooncert.com	Cabrils	2014	Ecommerce, Music, Sharing Economy	B2B and B2C
2	Counterest	< 10	Less htan 1 million	counterest.net	Barcelona	2013	AI, IoT, Smart Cities, Saas	B2B
2	Coverfy	< 50	< 5 million	www.coverfy.com	Barcelona	2016	Fintech, Traveltech, Social Economy	B2C
2	Cuinee	< 5	< 250k	www.cuinee.com	Sant Pere de Torelló	2016	Foodtech, Logistic Tech, Social Network	B2C
2	Curler	N.A.	N.A.	www.curlerapp.com	Barcelona	2015	Human Tech & Jobs, Mobile Software, Wellness	B2C
2	Currency Alliance	< 5	< 1 million	www.currencyalliance.com	Barcelona	2016	Adtech & Marketing, Smartcities, Blockchain	B2B
1	Dunforce	< 10	< 250k	www.dunforce.com	Barcelona	2016	Fintech, Insurtech, Saas	B2B
1	Dapcom-Data Services	< 5	N.A.	www.dapcom.es	Castelldefels	2013	Consultancy & Agency	B2B
2	Itech515	< 10	< 250k	www.itech515.com	Tarragona	2013	IoT, Consultancy & Agency, Software	B2B
2	Datalong 16	N.A.	N.A.	www.datalong16.com	Olivella	2014	Wearables, Logistic Tech, Smart Cities	B2B and B2C
1	Datumize	< 20	< 5 million	www.datumize.com	Viladecans	2014	Software, IoT, Data capture	B2B

2	Delectatech	< 5	< 250k	www.delectatech.com	Sant Cugat del Vallès	2014	Foodtech, IA, Productivity Tech, Advertising	B2B
3	Disjob On Line	< 5	< 1 million	www.disjob.com	Barceleona	2013	Cloud Computing, Human Tech, Software, Consultancy	B2B
3	Doctux	< 5	< 250k	www.doctux.com	Barcelona	2015	AI, Healthtech, Wellness	B2B and B2C
3	E Sonde	< 5	< 250k	www.esonde.com	Molins de Rei	2009	Cloud Computing, IoT, Logistic Tech, Software	B2B
2	Edenway	< 20	< 5 million	www.edenwaygroup.com	Barcelona	2011	IoT, Smart Cities, Consultancy	B2B
3	El comprador	< 20	N.A.	www.elcomprador.es	Altafulla	2012	Foodtech, Logistic Tech, Ecommerce	B2C
3	Enide	< 10	N.A.	www.enide.eu	Sant Andreu de la Barca	2011	Logistic Tech, Consultancy, Software	B2B
3	Entretenemos	< 10	< 1 million	www.entretenemos.com	Sant Hilari Sacalm	2016	Traveltech, Leisure, Ecommerce	B2B and B2C
2	Epinium	< 5	< 1 million	www.epinium.com	Mataró	2015	AI, Freemium, Adtech and Marketing	B2B
1	Expert Ymaging	N.A.	N.A.	www.ymaging.com	Barcelona	2012	Hardware & Wearables, Robotics, IoT	B2B
2	Findthatlead	< 10	< 1 million	www.findthatlead.com	Barcelona	2013	Human tech, Productivity Tech	B2B
2	FitLab	N.A.	N.A.	www.healthsportlab.com	Cerdanyola del Vallès	2013	Healthtech, Mobile Software, Sports	B2C
1	Foot Analytics	< 5	N.A.	www.footanalytics.com	Barcelona	2013	Robotics, Smart Cities, Internet	B2B
2	GLOBE MOD	N.A.	N.A.	www.globmod.com	Barcelona	2012	Healthtech, Biotech & Pharma, Wellness	B2B
2	Gowlook	< 20	< 250k	www.gowlook.com	Barcelona	2015	Internet, Mediatech, AI	B2B and B2C
1	HPCNow!	< 5	< 250k	www.hpcnow.com	Barcelona	2012	Biotech & Pharma ,Hardware, Systems Integration	B2B and B2C
2	Happyclick	< 5	N.A.	www.happyclick.info	Torrefarrera	2013	Cloud Computing, Systems Integration	B2B
3	Hibox	< 20	< 1 million	www.hibox.co	Barcelona	2016	Cloud Computing, Human Tech, Software	B2B
2	Holded	< 10	< 5 million	www.holded.com	Barcelona	2015	Consultancy & Agency, Fintech, Cloud Computing	B2B and B2C
2	Homeaway	N.A.	N.A.	www.homeaway.es	Barcelona	2015	Real State, Social Network, Leisure	B2C
2	Hutoma Artificial Intelligence	N.A.	< 5 million	www.hutoma.com	Barcelona	2016	Cloud Computing, Human Tech, IA	B2B
2	IBP Index	< 5	< 250k	www.ibpindex.com	Barcelona	2015	Sports, Data Analytics, Wellness, Mobile Software	B2C
3	IdFinance	< 50	< 200 million	www.idfinance.com	Barcelona	2015	Fintech & Insurtech, Online Lending	B2C
3	Imersivo	< 10	< 1 million	www.imersivo.com	Barcelona	2013	Ecommerce & Trade, Fashion & Design, Mobile Software	B2B
2	Imotion	< 5	< 250k	www.imotionretail.com	Castelldefels	2015	AI, Robotics, Smart Cities, Consultancy	B2B
1	Innovartium	< 5	N.A.	www.innovartium.com	Terrassa	2013	Productivity Tech, Consultancy & Agency	B2B
3	Intranetum	N.A.	< 1 million	www.intranetum.com	Manresa	2015	Productivity Tech, Data Structuring, Sharing tool	B2B

3	Iomando	< 5	< 250k	www.iomando.com	Barcelona	2011	Hardware & Wearables, IoT, Mobile Software	B2B and B2C
2	JobisJob	< 50	< 1 million	www.jobisjob.com	Sant Cugat del Vallès	2010	Human Tech, Advertising	B2B
3	KangooSave	< 5	< 1 million	www.kangoosave.com	Castelldefels	2009	Cloud Computing, Human Tech, Advertising	B2B and B2C
1	Kernel Analytics	< 50	< 20 million	kernel-analytics.com	Barcelona	2013	AI, Cloud Computing, Systems Integration, Consultancy & Agency	B2B
2	Kompyte	< 20	< 1 million	www.kompyte.com	Barcelona	2014	AI, Cloud Computing, Tracking Software	B2B
1	Konodrac	< 5	N.A.	www.konodrac.com	Barcelona	2012	Security & Cibersecurity, Software	B2B
2	Leadratings	N.A.	N.A.	www.lead-ratings.com	Cabrils	2014	Cloud Computing, Consultancy & Agency, Ratings analysis	B2B
3	Libranda	< 20	N.A.	www.libranda.com	Barcelona	2010	Ecommerce, Edtech, Distribution	B2C
2	Loyal Guru	< 5	< 1 million	www.loyal.guru	Barcelona	2013	Adtech & Marketing, Mobile Software, Productivity Tech	B2B
1	Magnis Commodities	< 20	< 1 million	www.magnuscmd.com	Sant Cugat del Vallès	2014	Blockchain, Fintech & Insurtech, Consultancy & Agency	B2B
2	Manzaning	< 10	< 1 million	www.manzaning.com	Barcelona	2016	Foodtech, Logistic Tech, Ecommerce & Trade	B2B and B2C
3	MC Health Tech	< 5	< 1 million	www.physiumsystem.com	Barcelona	2012	Medtech, Robotics, Health Tech	B2B
1	MOCA platform	< 5	< 5 million	www.mocaplatform.com	Sant Cugat del Vallès	2010	Adtech & Marketing, Mobile Software, IA	B2B
2	MR.NOOW	< 20	< 1 million	www.mrnoow.com	Barcelona	2011	Foodtech, Systems Integration	B2B and B2C
2	Made of Genes	< 20	< 1 million	www.madeofgenes.com	Esplugues de Llobregat	2015	Biotech & Pharma, Healthtech, Wellness	B2B and B2C
2	Meetmaps	< 5	< 250k	www.meetmaps.com	Barcelona	2015	Mobile Software, Augmented Reality, Internet	B2B
1	Mefio	< 5	< 250k	www.mefio.es	Barcelona	2009	Social Network, IoT	B2B and B2C
1	MindtheByte	< 20	< 1 million	www.mindthebyte.com	Barcelona	2011	Biotech & Pharma, Healthtech, Saas	B2B
1	Mint Labs	< 20	< 5 million	www.mint-labs.com	Barcelona	2013	Biotech & Pharma, Healthech, Cloud Computing	B2B
1	ModpoW	< 10	< 250k	www.modpow.es	Gavà	2010	Agritech, Greentech, IoT, Development & Manufacturing	B2B and B2C
3	Monkingme	< 10	< 1 million	www.monkingme.com	Barcelona	2014	Advertising, Mobile Software, Music	B2C
1	Mostrarium	Les than 5	< 250k	www.mostrarium.comes	Sant Pol de Mar	2012	HealthTech, IoT, Smart Cities, Mobile Software	B2B and B2C
1	NECADA	< 5	< 250k	www.polyhedra.tech	Barcelona	2017	Greentech, Smart Cities	B2B and B2C
1	NPAW	< 20	< 5 million	www.nicepeopleatwork.com	Barcelona	2008	Mediatech, Mobile Software, Film & Video	B2B
1	NUBALIA CLOUD COMPUTING	< 20	N.A.	www.nubalia.com	Barcelona	2011	Cloud Computing, Services	B2B

1	Nektria	< 10	< 5 million	www.nektria.es	Barcelona	2012	AI, Logistic Tech, Saas, Logistic Tech	B2B
2	Neuroelectrics	< 50	< 20 million	www.neuroelectrics	Barcelona	2011	Hardware, Healthech, Cloud computing	B2B
2	Nnegix	< 5	< 1 million	www.nnergix.com	Sant Cugat del Vallès	2013	Greentech, Energy Generatuin	B2B
2	Nubart	N.A.	< 250k	www.nubart.eu	Barcelona	2013	Mediatech & Content, Mobile Software	B2B
2	Ofertia	< 50	< 20 million	www.ofertia.com	Barcelona	2011	Adtech & Marketing, Mediatech, Advertising	B2B and B2C
2	Ontrace	< 5	< 1 million	www.on-trace.com	Mataró	2011	Hardware, Smart Cities, IoT	B2B
2	Online Booking Apartments	N.A.	< 250k	www.chictravelling.com	Sabadell	2012	Traveltech, Leisure, Ecommerce	B2C
1	Onna	< 20	< 1 million	www.onna.io	Barcleona	N.A.	Productivity Tech, Data Structuring, Software	B2B
1	OpenSeneca	< 50	< 1 million	www.civiciti.com	Barcelona	2016	IoT, Consultancy & Agency, Public Systems	B2B and B2C
1	Optimus Price	< 5	< 250k	www.optimusprice.ai	Vilanova I la Geltrú	2015	AI, Saas, Ecommerce, Dynamic Pricing	B2B
2	Payxpert	< 20	< 1 million	www.payxpert.es	Barcelona	2016	Fintech & Insurtech, Telecom	B2B
3	Psious	< 50	< 5 million	www.psious.com	Barcelona	2013	Healthtech, Virtual Reality, Wellness	B2B
1	Quantion	< 50	< 5 million	www.quantion.com	Barcelona	2014	IoT, Productivity Tech, Consultancy & Agency	B2B
3	Qustodio	< 20	< 5 million	www.qustodio.com	Barcelona	2012	Security & Cibersecurity, Software, Freemium	B2B and B2C
2	REMEMORI	< 5	< 250k	www.rememori.com	Sabadell	2012	Legaltech, Social Network, Cemetery Services	B2B and B2C
2	Radarprice	< 5	< 250k	www.radarprice.com	Barcelona	2011	Advertising, Ecommerce, Dynamic Pricing, Open Source	B2C
2	ReamMe	N.A.	N.A.	www.visual-tagging.com	Cerdanyola del Vallès	2012	Film & Video, Photo, Productivity Tech	B2B
2	Rich Audience	< 10	< 5 million	www.richaudience.com	Sabadell	2016	Adtech & Marketing, Advertising	B2B
1	RocketROI	< 50	< 5 million	www.rocketroi.com	Barcelona	2013	AI, Adtech & Marketing, Advertising, Consultancy & Agency	B2B
1	RunOpinion	< 5	< 250k	www.runopinion.com	Vilafranca del Penedès	2016	Adtech & Marketing, Sports, Wellness, Freemium	B2B and B2C
1	Runnersquare	< 5	< 1 million	www.runnersquare.com	Cabrera de Mar	2009	Social Network, Wellness, Ecommerce	B2C
1	SGClima	< 10	< 1 million	www.indoorclima.com	Sant Cugat del Vallès	2011	Greentech, Systems Integration, Augmented Reality, Software	B2B and B2C
1	Social & Loyal	< 20	< 250k	www.socialandloyal.com	Barcelona	2016	Gaming, Social Network	B2B
2	Scipedia	< 5	< 250k	www.scipedia.com	Barcelona	2015	AI, Social Netwrok, Advertising, Consultancy & Agency	B2B and B2C
1	Sekg	< 10	< 250k	www.sekg.net	Barcelona	2015	Hardware, Gaming, Mediatech	B2B and B2C
1	Sequentia Biotech	< 10	< 250k	www.sequentialbiotech.com	Barcelona	2013	Biotech & Pharma, Cloud Computing, Greentech, Consultancy & Agency	B2B and B2C

3	Sharify	< 5	< 250k	www.sharifyapp.com	Barcelona	2017	Social Network, Travel Tech, Leisure, Sharing Economy	B2B and B2C
2	Seekr	< 5	< 250k	www.sheekr.com	Barcelona	2015	Fashion & Design, Mobile Software, Social Network	B2B and B2C
1	ShuttleCloud	< 20	N.A.	www.shuttlecloud.com	Sabadell	2011	Cloud Computing, Mobile Software	B2B
1	Sincomis	N.A.	N.A.	www.sincomis.com	Barcelona	N.A.	Fintech & Insurtech, Mobile Software	Customer
1	Sisu Labs	< 10	N.A.	www.sisu-lab.com	Cornellà de Llobregat	2013	Data Analysis, Consultancy & Agency	B2B
1	Skoolpoint	< 5	< 250k	www.skoolpoint.com	Tarragona	2014	Edtech, Mobile Software, Smart Cities, Fremium	B2B and B2C
1	Smadex	< 50	< 1 million	www.smadex.com	Barcelona	2010	Adtech & Marketing, AI, Mobile Software, Advertising	B2B
2	Smart Engineering	< 10	< 250k	www.smartengineeringbcn.com	Barcelona	2013	Circular Economy, Software, R+D+I services, Consultancy & Agency	B2B
2	Smartive	< 20	< 1 million	www.smartive.eu	Sabadell	2013	Cloud Computing, Mobile Software, Development & Manufacturing	B2B
2	Social Coin	< 20	< 1 million	www.thesocialcoin.com	Barcelona	2013	AI, Mediatech & Content, Social Network	B2B
2	Social Diabetes	< 5	< 1 million	www.socialdiabetes.com	Barcelona	2012	Healthtech, Medtech, Mobile Software, IoT	B2B and B2C
2	Social Elephants	< 5	< 250k	www.socialelephants.com	Barcelona	2012	Saas, Fintech, Data structuring, Public Relations	B2B
2	Socialwibox	< 10	< 1 million	www.socialwibox.com	Mataró	2014	Adtech & Marketing, Mobile Software, Data analytics	B2B
3	Sphera Global Health Care	< 50	< 1 million	www.spheraglobalhealthcare.com	Barcelona	2012	Hardware, Healthtech, Medtech, Wellness, Martket Research	B2B and B2C
1	Squirro Software	< 50	< 20 million	www.squirro.com	Barcelona	2015	AI, Cloud Computing, Software,	B2B
3	Stayforlong	< 5	< 1 million	es.stayforlong.com	Barcelona	2015	Traveltech, Leisure, Ecommerce, Consultancy & Agency	B2C
3	Streeters	< 5	< 250k	www.streeters.es	Barcelona	2016	Fashion & Design, Systems Integration	B2C
3	TICOMBO	N.A.	N.A.	www.ticombo.com	Barcelona	N.A.	Sports, Consumer purchases, Ticket Sales	B2C
1	Talaia	< 20	< 1 million	www.talaia.io	Barcelona	2013	Cloud Computing, Security & Cibersecurity, Telecom, Network Operations	B2B
3	Talent Clue	< 50	< 5 million	www.sntalent.com	Barcelona	2009	Human Tech, Mobile Software, Telecom	B2B and B2C
3	Tamic	< 5	< 250k	www.tamic.es	Vilafranca del Penedès	2014	Hardware & Wearables, IoT, Agritech, Advertising, Consultancy & Agency	B2B
2	Tarifico Technologies	N.A.	N.A.	www.tarifico.es	Salt	2013	Telecom, Consultancy & Agency	B2C
3	Tekstum	< 5	< 250k	www.tekstum.com	Barcelona	2014	AI, Mediatech, Traveltech, Social Media data	B2B
3	Testamenta	< 10	< 250k	www.testamenta.es	Sabadell	2012	Adtech & Marketing, Legaltech, IoT	B2B and B2C
1	The Blue Dots	< 5	< 250k	www.thebluedots.io	Barcelona	N.A.	Hardware, Agritech, Saas, Data analytics	B2B
2	Tiendeo Web Marketing	More than 50	< 1 million	www.tiendeo.com	Barcelona	2011	Adtech & Marketing, Mediatech, Ecommerce	B2C

3	TimTul	< 5	< 250k	www.timtul.com	Barcelona	2015	Adtech & Marketing, Productivity Tech, Freemium	B2B
1	Tracktio	< 20	< 1 million	www.tracktio.com	Barcelona	2015	Hardware, IoT, Software, Smart Cities, Asset Management, Human Tech	B2B
1	Transmural Biotech	< 10	< 5 million	www.transmuralbiotech.com	Barcelona	2009	AI, Biotech & Pharma, Cloud Computing, Healthech	B2B
1	Transversal Business International	N.A.	< 250k	www.transversalbusiness.com	Barcelona	2011	Greentech, IoT, Advertising, Consultancy & Agency	B2B and B2C
3	Trustivity	< 5	< 250k	www.trustivity.com	La Roca del Vallès	2010	Ecommerce, Rating analysis, Customer Satisfaction analysis	B2B and B2C
3	TuManitas	N.A.	< 250k	www.tumanitas.com	Barcelona	2009	Human Tech, Marketplace	Bsuiness and B2C
2	Udobu	N.A.	N.A.	www.udobu.com	Barcelona	2011	Adtech & Marketing, Mobile Software, Leisure	B2B
1	VVreasy	< 50	< 1 million	www.vreasy.com	Barcelona	2015	Cloud Computing, Real State, TravelTech, Sharing Economy	B2B
3	Vudoir hub	< 5	< 1 million	www.vudoir.com	Barcelona	2015	Fashion & Design, Mobile Software, Social Network	B2B, B2C
3	WSN	< 5	< 1 million	www.worldsatnet.com	Gavà	2014	Cloud Computing, Telecom, Smart Cities, Drones (UAV)	B2B and B2C
1	WebSAys	< 10	< 1 million	www.websays.com	Barcelona	2011	Adtech & Marketing, Social Network	B2B
3	Whimed	N.A.	< 250k	www.whimed.com	Barcelona	2014	Fashipn & Design, Ecommerce, Marketplace	B2B and B2C
1	Wide Eyes Technologies	< 20	< 1 million	www.wide-eyes.it	Barcelona	2013	AI, Fashion & Design, Adtech & Marketing	B2B
1	Wordsensing	More than 50	< 20 million	www.worldsensing.com	Barcelona	2008	Telecom, Smart Cities, IoT, Software	B2B
3	YUMEHUB	< 5	< 1 million	www.yumehub.com	Barcelona	2016	AI, Fashion & Design, Fremium, Software	B2B and B2C
1	ZeedSecurity	< 10	N.A.	www.zeedsecurity.com	Sant Cugat del Vallès	2011	Mobile Software, Security & Cibersecurity	B2B
3	Zhilabs	< 50	< 250k	www.zhilabs.com	Barcelona	2008	Customer Service, Data analysis	B2B
3	barcelonackecin	N.A.	N.A.	www.barcelonacheckin.com	Barcelona	2015	Real State, Traveltech, Social Network	Costumer
3	bonoom	< 10	< 1 million	www.bonoom.com	Barcelona	2016	Cloud Computing, Healthtech, Medtech	Customer
1	globalbook	< 10	< 1 million	www.globalbook.info	Barcelona	2012	Logistic Tech, Publishing	B2B
1	inAtlas	< 10	< 1 million	www.inatlas.com	Barcelona	2010	Adtech & Marketing, Consultancy & Agency	B2B
3	invertiaWeb	< 5	< 1 million	www.invertiaweb.com	Torroella de Montgrí	2013	Adtech & Marketing, IoT, Productivity Tech	B2B and B2C
3	pa-community.com	< 5	< 1 million	www.pa-community.com	Barcelona	2015	Mediatech, Social Network, Leisure	B2B and B2C
3	smArDS	< 5	< 250k	www.smards.net	Barcelona	2015	AI, Mediatech, Telecom, Augmented Reality	B2B

3	wwwowww	< 5	< 250k	www.wwwowww.me	Cerdanyola del Vallès	2013	AI, Mobile Software, Social Network, Advertising, Marketplace	B2B and B2C
3	zinkgame	< 5	< 250k	www.zinkmediastudio.com	Barcelona	2015	Gaming, Mobile Software, Gaming, Freemium	B2B and B2C
3	zyrcle	< 5	< 1 million	www.zyrcle.com	Matadepera	2012	Mobile Software, Productivity Tech, Social Network, Advertising	B2B and B2C

A short description of the activities of each company is accessible on the Barcelona & Catalonia Startup Hub (http://startupshub.catalonia.com/list-of-startups). If the reader is interested, we have compiled the descriptions of the companies we used in our sample in one single document that can be accessed publicly in the following site.

https://drive.google.com/file/d/13wpwo7MFEGkLYdArBJaHBxy97ulGTty8/view?usp=sharing