A Work Project presented as part of the requirements for the Award of an International Masters Degree in Management from the NOVA – School of Business and Economics

DATA-DRIVEN DECISION-MAKING IN SOCIAL IMPACT SECTOR IN PORTUGAL: STATE OF THE ART, CHALLENGES, AND OPPORTUNITIES FOR IMPROVEMENTS

MARIANA MARQUES GOMES BRANDÃO BANDEIRA

26085

A Project carried out on the International Masters in Management Program, under the supervision of:

Leid Zejnilović

JANUARY 3rd, 2020

Abstract

Technological development contributed to the increasing availability of data and the capacity to process it. Data-driven decision-making became affordable and a source of competitive advantage for profit-seeking organizations. However, the third sector is falling behind in the adoption of data science. Evidence lacks in the scholarly literature of critical resources affecting social-good-oriented organizations' adoption decisions. This thesis aims to fill this gap through an empirical investigation of data science usage among Portuguese social economy entities. A mixed-method research is conducted, informed by the theoretical frameworks of technology adoption model and resource-based view of the firm. The results are discussed from the practitioners and theory contribution perspectives.

Keywords: Social-good-oriented organizations; Data-driven decision-making; Resource-Based View of the Firm; Unified Theory of Acceptance and Use of Technology

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

Table of Contents

1. Introduction

The amount of data generated every day is massive, and the tendency is not to slow down. Many industries see themselves forced to adopt data-driven strategies to remain competitive, as the benefits from Data Analytics continue to appear (Brownlow et al., 2015). Similar companies may positively differ in 5-6 percent in productivity levels when adopting data-driven decision-making (DDDM) processes (Brynjolfsson et al., 2011). Technology is a moving target and requires constant investment for one to be up to date, however, not every sector can keep up. Social-good-oriented organizations cannot afford the most recent equipment and technologies on the market (Bobsin et al., 2018) and lack capacity and appropriate skills for data analysis (Blake, 2019). Many agree that it is time for nonprofit organizations (NPOs) to shift towards data usage (McAfee et al., 2012; Fruchterman, 2016) and that the conditions have never been more appropriate (Ashby, 2019). Data science is a growing need within the third sector (Blake, 2019), however, one must assess the needs for the implementation of the technology. For the implementation of data science to occur, there are challenges necessary to overcome, which are managerial more than technical (McAfee et al., 2012). This thesis seeks to answer the following research question, '*What are the critical antecedents of the adoption of data-driven decision-making in the social impact sector in Portugal?'*. To assess the presented question, the thesis is structured as follows. Firstly, the theoretical background on the national social economy and its relationship with data science will be presented. Secondly, the research methodology and design will be described and further results' comprehensive description. Later, the discussion and suggestions on how to approach the third sector will be provided followed by the study limitations.

2. Background

2.1. Social Good Entities in Portugal

The Portuguese Law on Social Economy, *Lei das Bases da Economia Social - Law Nº 30/2013*, May 8th, defines social economy as all the economic and social activities freely carried out by entities with legal forms of i) cooperatives, ii) mutual associations, iii) mercies (misericórdias), iv) foundations, v) entities with IPSS (Private Institutions of Social Solidarity) statute, vi) associations with altruistic aims acting in cultural, recreational, sports and local development fields, vii) entities in the communitarian and self-managing subsector under the cooperative and social constitution, and viii) other entities with legal form respecting the social economy principles. The guiding principles of the social economy in Portugal are described in Law Nº 30/2013 as being the following seven: 1) the primacy of people, 2) freedom of membership, 3) democratic control, 4) conciliation of interests, 5) solidarity, 6) autonomy, and 7) surplus allocation to the social economy (Assembleia da República, 2013). Although it appears to be comprehensive, the law on social economy gives room to different interpretations, among them the question if social enterprises belong or not to the social economy. Additionally, the concept of social enterprise is not yet established in Portugal (Stoyan et al., 2014), increasing the interpretation difficulty. In Portugal, it is estimated to be over seventy thousand (70,000) entities comprising the social economy and with its vast majority, over 90%, being altruistic associations, and roughly 50% of the field of activities related to culture, communication and recreational field (INE, 2019). As concluded in the recently published *Satellite Account of Social Economy in Portugal - 2016*, in that year social economy was responsible for 3.0% of the Portuguese GVA (Gross Value Added), 5.3% of wages and total employment, and 6.1% of paid employment in the national economy (INE, 2019).

As seen, the impact of the social economy in Portugal is significant in the national economy and the national social welfare. It is of paramount importance to support these entities to do good and achieve greater social impact. One of the means to achieve this objective is the introduction of data-driven decision-making and develop competencies to test and implement data science.

2.2. Data Science for Social Good

Data science can be described as the "application of quantitative and qualitative methods to solve relevant problems and predict outcomes" (Waller et al., 2013). It is a technology that comprises multidisciplines, statistics, data management, machine learning, but also social sciences to understand the context and provide comprehension. Ultimately, the goal is to extract useful knowledge and insights from structured and unstructured data and act on it when appropriate (Dhar, 2012). Data science, or data-driven decision-making as will be used in this thesis interchangeably, adoption can increase the productivity of an organization (Brynjolfsson et al., 2011) and it is fundamental for it to remain competitive within a market (Brownlow et al., 2015).

In the corporate world, the majority of for-profit organizations have emerged in data-driven strategies, otherwise risking the businesses' rentability (Brownlow et al., 2015), working with straightforward metrics aiming profit maximization, steady IT infrastructure and experts. By contrast, in the third sector, the adoption of data science is lagging behind (McNutt, 2018), and, for changes to occur, these must be integrated with the organization's mission which guides action (Bobsin et al. 2018). Some may argue that this is due to a lack of interest from experts to invest in tailored solutions for nonprofit organizations (Jariego, 2007) with few and too expensive tools conceived to the specificities of NPOs (Bobsin et al., 2018). Others believe that there is already a significant amount of technological applications tailor-made to address NPOs' needs (McNutt, 2018), disagreeing that the greatest barrier to data science implementation is the lack of designed solutions. In the third sector, instead of productivity and competitiveness, social impact drives decisions (Bobsin et al., 2018). Therefore, NPOs must, beforehand, acknowledge the technology's potential benefits in society and ways its activities and impact may be leveraged through data science. McNutt (2018) describes data science implementation in NPOs as "the next development in nonprofit advocacy", with the potential exploration of new areas of operations and advance in organizations' missions.

One of the greatest challenges within the third sector is to measure impact (Fruchterman, 2016). NPOs struggle with what to measure and when to measure it but the theory of change and logic model support data-driven decision-making (James Bell Associates, 2018). Theory of change (Lewin - 1947) allows for problem recognition, desired outcomes, and pathways of change assessment. The logic model resorts to data collection principles to assess impact and has four components (James Bell Associates, 2018): *Inputs*, financial, material and personnel resources; *Activities*, organization's interventions targeting a social problem; *Outputs*, direct quantifiable results of an activity (e.g. presences in an event); *Outcomes*, impact or changes resulting, whether in short or long term. Each of the four components is dependent on data collection to generate results and these are dependent on the interpretation of the data collected. The process mirrors one of data science's application in the field, as data interpretation is meant to answer questions as *reporting*, *diagnosis*, *prediction*, and *recommendations* (Van Der Aalst et al., 2015). Data science could be one of the answers to the mystery of impact measurement within the third sector (James Bell Associates, 2018).

More than enabling impact measurement, data science can leverage NPOs' missions supporting decision-making and strategy (Baar et al., 2016). One of the data science's advantages is its predictive capacity, transforming insights into action (Dhar, 2012). More specifically, social-good organizations can estimate future results based on results drawn from activities' past data analysis. Allowing data-driven decisions rather than intuition-driven hence, making better decisions (McAfee et al., 2012). Moreover, organizations seeking impact can leverage decisions on databases with information on employees, clients, beneficiaries and funders (McNutt, 2018). For instance, by extracting insights from datasets on past fundraising programs, organizations can increase efficiency, attracting investors and increasing contributions (McNutt, 2018). The same would happen with clients or beneficiaries, allowing a comprehensive understanding of people's needs, enabling improvement on existing services and to better target the groups. Ultimately, acting in the appropriate time with tailored and more efficient solutions, increasing the social impact.

Nonprofit organizations have been lagging far behind in the adoption of new technologies for decades now (McNutt, 2018). And when in similar industries, third sector organizations and for-profit corporations do not have equal opportunity in accessing capital (Myser, 2016). The barriers and challenges to the implementation of new technology like data science are manifold, and different organizations will be distinctly conditioned (Eimhjellen et al., 2013).

In Portugal, the nonprofit sector has three main sources of income: earned income, private philanthropy, and government or public sector support. Earned income includes the sale of goods or services, allocating surpluses to primary activities allowing these to remain sustainable. Private philanthropy is usually comprised of individual or private institutions' donations. And lastly, the government or public sector support includes grants, contracts, and payments from government-financed social security systems (Franco et al., 2012). Despite the several sources of revenue, financial hurdles remain one of the biggest challenges of most of the organizations within the sector (Monteiro et al., 2015). Many organizations, due to lack of funding to invest in technology, work with "obsolete equipment and outdated technologies" (Bobsin et al., 2018). Funding programs, whether public or private, have rules limiting the allocation of resources to activities not considered to be primary (Bobsin et al., 2018), as technology. Moreover, funders and donors themselves offer resistance to invest in new fields of technology (West, 2019), such as data science, seeking instead for tangible results in the lasting impact their investments may generate (Fruchterman, 2016).

Despite data science's manifold benefits, one must be aware of these benefits to act upon them. Many organizations seem to fail in getting educated on the benefits before potential implementation. The perceived usefulness and applications of data science may vary among and within organizations, as well as the challenges to its implementation (Bobsin et al., 2018). Additionally, and as found on research conducted by the Data Science Portuguese Association, technology adoption levels vary among different organizational areas (DSPA, 2019). However, this relationship goes deeper than the organizational area. A study on data usage adoption concluded that there is a positive correlation between the percentage of educated people within a company's team and the likelihood of that team reporting high levels of data-driven decision-making (Brynjolfsson et al., 2016). Hence, once there is expertise within an organization, that understands and recognizes the potential value of data-driven strategies, there is a higher probability of having a more effective implementation of the technology, as it reduces personnel resistance (Brownlow et al., 2015). However, budget constraints lead to difficulties in attracting and retaining talent (Bobsin et al., 2018) and there is a general lack of statistically literate people (Ashby, 2019). In Portugal, organizations find attracting new people for their social organs a major problem (Monteiro et al., 2015) and adding a criterion of technological skills to the candidate may hinder the process. Hence, when organizations fail to see the data science usefulness, the final decision of usage is expected to be negatively affected. In particular, if they lack awareness at the top of the organization hierarchy (Bobsin et al., 2018), hence, failing to encourage investment in new technologies. Therefore, the need to educate an organization and provide guidance to implement data science seems to be of uppermost importance.

There may be several characteristics that could act as barriers to data science implementation (McNutt, 2018), leading to some resistance and initial inertia. For instance, the organization's size and structure, the larger the more expected it is to have IT capacity (Balser, 2008), and more organic and horizontal structures tend to find the adoption of new technologies easier (Bobsin et al., 2018; Eimhjellen et al., 2013). To understand the intention of an organization towards the use of data science, one should assess the current managing culture. For instance, assessing if the process of decision-making is driven by intuition or data as it is common to have managers relying on intuition over data (HBR Analytics Services, 2012). However, the phenomenon has been changing, and the adoption of data usage has begun to break the habit, reducing the predominant instinct weight on the decision-making process (Brynjolfsson et al., 2016). Moreover, the workforce profile, whether mainly represented by volunteers or employees, seems to impact the intention to adopt new technologies (Bobsin et al., 2018) as motivations and professionalism levels tend to be different, affecting project prioritization and time allocation. Hence, the cultural change must be managed effectively (McAfee et al., 2012) for one to adopt and fully benefit from data science.

It is undeniable the great potential for huge amounts of data generated within social organizations, and consequently, an increased potential for misinterpretation and misusage (Ashby, 2019). The quality and integrity of the data collected seem to act as a barrier to many organizations, whether nonprofit (Baar et al., 2016) or for-profit with already established businesses (Brownlow et al., 2015). Data quality comprises aspects such as accuracy, completeness, consistency, and currency (Scannapieco et al., 2005) amongst others. The step forward given with the adoption of data usage can lead to two steps back when data has poor quality or is misused, as it reduces the efficiency of the organizations' decisions (Baar et al., 2016). Moreover, extremely sensitive data may be held by nonprofit organizations, making data disclosure another important concern when using data for decision-making (Baar et al., 2016), slowing down the process of adopting data science. With increasing restrictions regarding data collection and usage, like the EU General Data Protection Regulation (GDPR) - Regulation (EU) 2016/679, there is greater resistance to the implementation of data-driven strategies.

Data science is now closer to being affordable for social-good-oriented organizations, not because the organizations are abounding in financial resources, but because the technology is less expensive (McAfee et al., 2012). Additionally, there is an increasing availability of data science tools to those that can act when in possession of the insights (Ashby, 2019). Data visualization tools are acquiring formats that are friendlier, more intuitive and easier to use, with datasets as inputs and a multitude of possibilities as outputs, graphics, infographics, charts, and maps (e.g. *Tableau* and *Infogram*). Moreover, there are already entities willing to help NPOs through data science, with guidance, tools, and expertise, as it is the case of The Royal Statistical Society – Statisticians for Society, a *pro bono* work that connects statisticians with charities (Ashby, 2019). Or Data Science for Social Good Solve - DSSG Solve, an online platform for social organizations to present projects in need for data science assistance, having experts as volunteers to help to scope the project and to solve the problems (DSSG, 2018). An example from Portugal is Data Science for Social Good Portuguese Foundation, an open community of data scientists aiming to match beneficiaries, that may benefit from data-driven methodologies, with voluntaries that are experts in the field of data science. What before was unintelligible it is today accessible for those that lack expertise on the field.

However, the relationships between vital resources within the third sector for data science implementation and the sector's perception and later usage intention behavioural of data science, have not yet been clarified. Therefore, this research sought to answer the following research question:

What are the critical antecedents of the adoption of data-driven decision-making in the social impact sector in Portugal?

The primary objective of this research is to assess the usage of data in the decision-making process of social-good related organizations in Portugal and assess the critical resources an organization should have to become data-driven.

Secondly, the objective is to understand the acceptance and usage of data science in social-good-oriented organizations, by looking into perceptions and future expected consequences on the usage of data science.

2.3. Theoretical Frameworks

2.3.1. Unified Theory of Acceptance and Use of Technology (UTAUT)

Hitherto, there seems to be a consensus among the authors that social-good-oriented organizations would benefit from adopting a data-driven culture. That would ideally be accompanied by an investment in education on the matter. However, such a plan only becomes viable if social organizations have the willingness to accept data science technology. This willingness can be assessed through the Unified Theory of Acceptance and Use of Technology (UTAUT), which may explain up to 70 percent of the variance in the intention of the use of a specific technology (Venkatesh et al., 2003). The theory defines four constructs that directly affect the usage intention: *performance expectancy*, *effort expectancy*, *social influence*, and *facilitating conditions*. Methodologically, the UTAUT is usually operationalized by the means of a survey. In this thesis, this framework serves to support the understanding of the social organizations' behavioural intention and attitude towards the use of data science.

2.3.2. Technologies Affordances and Constraints

UTAUT becomes a scarce source when assessing the reasons for the implementation decision, as it does not consider the perspective of potential consequences, and these may influence the final decision of adoption. Majchrzak and Markus (2012) affirm that the consequences of the uses of information systems (IS) are better drawn through the understanding of relationships between organizations and technology features. To this end, the Technology Affordances and Constraints Theory (TACT) (Majchrzak et al., 2012) complements the UTAUT, by mapping the reasons behind the decision. *Affordances* as the range of new possibilities organizations may benefit from data science and *Constraints* as the limitations.

2.3.3. Organizational Perspective

To complement UTAUT (Venkatesh et al., 2003) and TACT (Majchrzak et al., 2012), the research also resorts to the Resource-Based View (RBV) of the Firm (Wernerfelt, 1984). RBV allows mapping the critical resources and capabilities that can leverage a firm's performance. In the third sector, there are critical resources and capabilities that when leveraged affect positively organizations' performance and ultimately the magnitude of their social impact (Bacq et al., 2016). When referring to the antecedents for data science implementation, Resource-Based View can be repurposed, allowing to map critical resources and capabilities in social-good-oriented organizations to the implementation of data science and ultimately boosting social impact.

3. Methodology and Research Design

3.1. Research Methods and Data Collection

Along with the theoretical research, there was a relentless pursuit for the insights on the current status of technology within the Portuguese social economy. To this end, mixed-methods research was conducted, qualitative and quantitative research. Both the interviews and the survey were developed with the referred theories as support.

Qualitative Research

Eight in-depth semi-structured interviews were conducted among different Portuguese entities, four of the eight interviews were via call or video-call and the remaining four were in-person interviews. *Appendix 1* shows the list of organizations that participated in the research, together with the interview length. The conversations followed a general script that the author developed for the research (see *appendix 2*). To select the interviewees, the author used snowball sampling (chain referral sampling) method, with NOVA SBE Data Science Knowledge Center (DSKC) being an informant (Mack et al. 2011). The contact with all the interviewees was established through the DSCK, and they were part of the centers existing network. Snowball sampling is a type of purposive sampling (Mack et al., 2011), which implies that the entities contacted for the interviews already complied with preselected criteria. In this case, two criteria were used. Firstly, the interviewee should be a representative of an organization registered in Portugal with experience in the field of Portuguese social economy. Secondly, the organizations should have applied for the *Data for Change* program which aimed to identify organizations with problems that could potentially be solved through Data Science (DSKC 2019). The seven interviewed organizations consented the audio recording of the conversations and each interview had two recordings from different devices to reduce the risk of failure. To complement general notes taken during the conversation, the recordings were later listened thoroughly and repeatedly to extract more information applicable to the research, given its objectives. Each interview had its collection of insights that were coded and later crossed with the remaining. A single data collection of insights was later generated to find patterns and draw conclusions on the topic. The tools used to collect and code the data were, recording apps (mobile phone and computer) and Microsoft Office Word.

The first of the interviews sought to clarify already existing definitions in the social economy, such as legal formats within non-profit organizations. The referred interview was done with a senior consultant and *pro bono* legal manager, from Vieira de Almeida Associates (VdA), a Portuguese law firm, that provides legal advice to social economy entities. The other seven interviews were conducted with different entities from the Portuguese social economy. Two of the organizations have the enterprise statute, meaning that are for-profit companies, but are self-considered social enterprises as their primary focus is social impact. The remaining five interviews were conducted with non-profit associations, from different fields of action, four IPSS (Private Institutions of Social Solidarity) and one NGOD (Non-Governmental Organization for Development). As presented before, *appendix 1* provides additional information on the organization's interviewed with respective lengths.

Quantitative Research

With the insights drawn from the qualitative research, a survey was developed to reach a greater and more diverse universe of entities, as it was not restricted to *Data for Change* applicants. The survey was in Portuguese since it targeted Portuguese organizations, the final survey can be found in *appendix 3*. To guarantee the viability of the survey, and to guarantee a proper understanding, the former was pre-tested internally with 3 senior researchers, of which 2 were Portuguese native speakers and 1 foreign-language speaker. All the pre-test participants have significant experience in the use of data science in social-good-oriented organizations in Portugal. The final survey had 38 questions, 12 open answers, 18 of multiple-choice, 7 questions with five points Likert scales, and the remaining was a sort answer. The survey was divided into 9 sections, the first two sections assessed the organization's characteristics and their level of IT infrastructure. The third section directly referred to data science and aimed at understanding if organizations were familiar with the concept. The section that followed started with a comprehensive explanation of the concept expressed in an understandable language for non-professionals and included several examples to provide mental cues about data science. After the introduction in this section, the survey assessed the organization's perceptions of data science applications. The questions in the fifth section characterize the data that organizations have stored and their data collecting habits. The sixth and seventh sections focused on Venkatesh’s (2003) Unified Theory of Acceptance and Use of Technology (UTAUT) and Technologies Affordances and Constraints Theory (Majchrzak et al., 2012), their perceptions of advantages and disadvantages, affordances and constraints, as potential users of the technology. In *appendix 4* can be found the list of items used to estimate UTAUT (Venkatesh et al., 2003). The penultimate section investigates if the implementation of data science was being actively pursued, as an organizational priority or not. Finally, the last section allowed for comments on the topic and optionally share an email address for further contact on the research. The sampling method used was convenience sampling, aiming at the entire network of organizations available. The survey was sent to over 4300 addresses of social-good organizations and the data considered were collected from the 11th to the 20th of December. The reminder was sent on the 17th and 18th of December to over 4100 addresses. The author used Microsoft Office Forms to develop the survey and later downloaded the results in Microsoft Office Excel format. The sample was of 159 answers.

3.2. Data Analysis Methods

For the qualitative data analysis, the transcripts of the interviews were analysed using content analysis (Seidel, 1998). Even though there was a general script, interviews had a conversation flow. Hence, not every answer would fall on the predefined category of the question asked. To this end, the process of *noticing*, *collecting* and *thinking* (Seidel, 1998) served as a support to the qualitative data analysis. The process allowed for a translation of unique insights into structured and analysable datasets. *Noticing* was used to code the segments of the interviews, *collecting* was used to sort once coded the insights, and *thinking* was used as the process of analysing the results collected once segmented and organized.

During the interviews, the script presented in *appendix 2* served as a guide for the conversation. However, as noticing is a recursive process (Seidel, 1998), the notes that were taken during the interviews allowed for other questions generation. Each interviewed organization had its individual data collection resulting from the recording and later partial transcript. Within each collection the topics of answer were coded, for instance, the answer on what motivated *Data for Change* application (coded as *Data4Change*) would later allow an assessment on the behavioural intention of implementing data science. *Appendix 5* illustrates the defined general codes and examples transcribed from the interviews. Once labelled, the topics would fall into theoretical research categories. For instance, when a statement was coded as a *DS_new_possibility* this would fall into the Technology Affordances and Constraints Theory (Majchrzak et al., 2012) as it represents an affordance. This analysis, of attributing categories to coded statements crossed data from different data collections and resulted in a single data collection. Referring to the same example, all the statements from all organizations regarding new possibilities that data science may bring were analysed with the support of TACT (Majchrzak et al., 2012). Furthermore, the patterns were analysed within each collection and within the single data collection previously generated.

Quantitatively, once collected the survey data, the input variables were coded to facilitate the analysis. The information about the variables and the conversion from categorical to nominal scale is available in *appendices 6* to *8*. To understand the survey results and get an overview of the data, an exploratory data analysis has been conducted.

Given that part of the survey was based on the technology adoption model (TAM) (Venkatesh et al., 2012), an exploratory factor analysis (EFA) was conducted to test if the TAM constructs are identifiable and if yes, to generate the scales for these constructs. Exploratory factor analysis allowed for dimensionality reduction and the identification of four main constructs within the survey data.

From the scale questions referring to UTAUT (Venkatesh et al., 2003), twenty-one items were used to conduct the factor analysis. As a result, four factors stood out for their greater values above the boundary of one, *appendix 9* presents the Eigenvalues for all the variables used. The rotation method used was Promax as the results generated were clearer, as shown in *appendix 10*. Moreover, *appendix 11* summarizes the variable names with the questions asked and respective factor loadings. Additionally, to guarantee internal consistency, the author computed Cronbach's alpha for each factor and all the values were above the required 0.7 (Hair et al., 2014), as presented in *appendix 12*. Moreover, one of the variables, effort required for data science implementation, was dropped from the analysis since the item loaded into two factors (1 and 2).

A multivariate regression model was considered to explore associations between variables rather than causal relationships. Operationally, a nested regression was run using four blocks of variables, resulting in four linear regressions. In all four models, the dependent variable was the usage intention of data science, which is a behavioral proxy for the actual adoption of data science (Ajzen, 1991). Eighteen independent variables (see *appendix 13*) were divided into four blocks of predictors. The first block corresponds to 5 control variables (e.g. size, age group, demand matching). The second block adds 7 new variables, items that regard the resource-based view (e.g. funding availability, access to education), block three is comprised of 2 variables containing information on the data-driven decision-making culture in the organizations, and block four adds 4 variables that originate from the UTAUT framework. *Appendix 14* shows in the detail the independent variables considered of each block.

The four linear regressions were executed sequentially, starting with control variables only and adding one block of variables per regression. The EFA ran the entire sample (N=159), as all the analysed scale questions were answered. The multivariate regression analysis was run with 157 of the answers due to one non-answered question by 2 of the respondents.

4. Research Findings

4.1. Results of the Qualitative Analysis

While most of the interviewed organizations have elementary technology for their operations, there is a weak or non-existent application of data science in their operations. As stated before, all the interviewed organizations had applied to the program *Data for Change*, meaning that, there was already at least one problem recognized within the organization that could be potentially solved with data science. The motivation to apply would mainly be the lack of expertise to implement data science to work their stored data. Additionally, more than knowing what question to ask, many saw new possibilities arising with data science (affordances). For instance, data science was perceived to have great potential when it comes to targeting, through data analysis and predictive statistics. Even though data science's applications were seen mainly as efficiency magnifiers, some saw potential applications next to the beneficiaries. For instance, an AI matchmaker tool for doctors and patients on an organization's website, reducing the resistance of scheduling an appointment when searching for a doctor among dozens. Or even an app for the beneficiaries that need help to keep track of their activities with the organization and vice-versa. Moreover, organizations said that impact measurement was a difficult task that data science could help improving.

On the other hand, when asked on data science's constraints most organizations said these to be "none". However, when revising the answer, some perceived limitations arose. For instance, the risk of having technology replacing the human connection in the social sector was a big concern. GDPR was referred to as too complex and data misusage could take the organization to violate the regulation unintentionally. In the case of cultural transformation, one of the concerns referred was a long-term misperception of data science benefits consequently, increasing the resistance to the implementation. Commonly agreed was that data science implementation would have to take time and funding from other projects, however, not deviating the focus of the primary goal.

From a resource's perspective, the social-good-oriented organizations seem to have partially acquired the resources which allow data science implementation. Data, the raw material of data science, was correctly perceived as a starting point for data-driven decision-making adoption. In that regard, most organizations have programs that automatically store data and have been doing it for years, having now stored great amounts of unused data. Besides data, the mindset and willingness to adopt a more data-driven culture was constantly referred to as a resource that supports the implementation. Others referred that having a DPO (Data Protection Officer), a website to develop or even experience with former technological transformations (replacing computers by laptops) could also be a starting point of the culture shift towards data-driven organizations. Moreover, partners and reputation are perceived for some, as resources that can facilitate access to resources for data science implementation.

On the other hand, when assessing critical resources missing to implement data science, the answers were extremely similar. *Education*, *funding*, and *expertise* were the resources that the interviewees said to lack the most. Hence, these factors were identified as the principal obstacles that are preventing organizations from implementing data science.

As for challenges, GDPR and data treatment seemed to be always present. One organization showed concern in asking for more detailed information to individuals, as it may have negative impacts on their social participation. Others perceived the treatment of highly sensitive data as a challenge together with the individuals' resistance to sharing it, hence, jeopardizing the social impact.

4.2. Quantitative Analysis

4.2.1 Descriptive Statistics

The survey was sent to 4321 addresses, resulting in a sample of 159 answers. Among the respondents, 70% are non-profit associations, 6% cooperatives, 3% foundations, and the remaining are religious entities and others, mainly parish social centres. 87 of the respondent organizations have acquired the IPSS (Private Institution of Social Solidarity) statute, and 31 are NGOs (Non-Governmental Organizations), whether for development, environment or disabled persons. 16 of the 159 organizations have the cumulative statute of both, IPSS and NGO. As individual respondents, around 63% are between 36 and 55 years old, and the remaining is divided between below 36 years old (19%) or above 55 years old (18%). Roughly half (54%) of the respondents are from small organizations (less than 50 members, including employees and volunteers). The remaining are of medium size (34%) with 50 to 250 members, and 12% of larger size (8% with 251 to 1000 members and 4% with over 1000 members). Regarding the organization's structure, 46% said to have a more vertical culture (top management responsible for the decisions) and 43% horizontal (cross-hierarchical decisions), the remaining 11% claimed to have a mixture of both or another format. Only 38% of the organizations are currently meeting the demand of society regarding their primary activities and 71% said not to have the funding for data science implementation. The majority claimed to already base the process of decision-making in data (60%), however, the scenario looks different when assessing the behavioural intention and the attitude towards the use of data science in the organizations. The descriptive statistics are presented in more detail in *appendix 15*.

4.2.2 Data Analysis Results

Resulting from the exploratory factor analysis on the collected data, four main factors are identified, as shown in *appendix 10*. Internal consistency is guaranteed since the Cronbach's alpha for all the factors is above 0.7 (Hair et al., 2014).

In the case of multiple regression analysis, a single dependent variable (data science usage intention) is considered. In the model, all the blocks of independent variables (*appendix 14*) show to add statistical significance to the prediction of data science usage intention, through positive variations in R-squared value. Given the R-squared variations, the block that adds more statistical significance to the model is the second block referring to resource availability (RBV) (Wernerfelt, 1984), with $\Delta R^2=0.2209$. Regression 4 yielded the greatest R-squared *(0.519)* and hence the regression that better explains the relationship between variables and data science usage intention. The summarized information of R-squared values can be found in *appendix 16* and the results of the multivariable analysis in *Table 1*. In regression 4, usage intention of data science is positively associated with, perceived performance enhancement (coeff = 0.32; p<0.01), social influence (coeff = 0.26; p<0.01), having DDDM as an objective (coeff = 0.28; p<0.05) and available expertise for data science implementation (coeff = 0.23; p<0.05). In regressions 2 and 3, access to education affects positively data science usage intention (coeff = 0.26; p<0.05 regression 2) and (coeff = 0.21; p<0.05; regression 3). However, once block 4 is added (e.g. social influence, perceived performance augmentation), education loses statistical significance. Similarly, the respondent's age (above 65 years old) loses statistical significance in regression 4. Since in regressions 2 and 3, shows to be statistically significant by negatively affecting the intention of usage of data science (coeff = -0.99; p<0.05; regression 2) and (coeff = -1.22; p<0.01; regression 3). Unexpectedly, in none of the four regression, funding shows to have a statistical significance as a predictor of data science usage intention. Individual regressions from each block addition can be found in *appendix 17*.

Table 1. Multiple regression analysis of predictors of usage intention of data science (N=157)

Independent variables	**Regression 1**	**Regression 2**	**Regression 3**	**Regression 4**
Respondent's age				
25 - 35 years old	-0.21	-0.69**	-0.97**	-0.37
36 - 45 years old	-0.09	-0.64**	-0.78*	-0.16
46 - 55 years old	0.39	-0.21	-0.49	0.07
56 - 65 years old	0.17	-0.64*	-0.87*	-0.12
More than 65 years old	-0.52	-0.99**	-1.22***	-0.42
Organization's age				
25 - 35 years old	-0.01	-0.06	0.08	-0.13
36 - 45 years old	-0.26	-0.17	-0.02	-0.14
46 - 55 years old	-0.11	-0.10	0.01	-0.05
56 - 65 years old	0.10	0.19	0.31	0.19
More than 65 years old	-0.52	-0.75*	-0.54	-0.65*
Board's age				
36 - 45 years old	0.09	0.01	-0.02	-0.08
46 - 55 years old	-0.10	-0.11	-0.10	-0.05
56 - 65 years old	-0.14	-0.20	-0.24	-0.24
More than 65 years old	-0.33	-0.22	-0.29	-0.58
Organization's dimension				
50 - 250 members	0.12	0.04	0.01	-0.03

251 - 500 members	0.08	0.02	0.04	0.10
501 - 750 members	0.09	-0.27	-0.15	-0.01
751 - 1000 members	0.32	-0.06	0.04	0.04
More than 1000 members	-0.49	-0.68*	-0.49	-0.11
Meets demand (1, yes; 0, no)	-0.12	-0.16	-0.13	-0.02
Access to education in data science in the past 2 years (1, yes; 0, no)		0.26**	0.21**	0.03
Funding is available for data science implementation[a]		0.08	0.08	0.07
Expertise is available for data science implementation[a]		0.29***	0.32***	0.23**
Digital data collection routine[a]		-0.06	-0.08	-0.10
Internal DPO (1, exists; 0, does not exist)		-0.15	-0.13	-0.11
GDPR awareness (1, yes; 0, no)		0.48	0.31	-0.14
Willingness to collaborate/contact (1, yes; 0, no)		0.31**	0.32**	0.14
Data-driven decision-making as routine[a]			-0.21	-0.14
Data-driven decision-making as objective[a]			0.33**	0.28**
Data science usage level relative to peers[b]				0.09
Data science perceived ease of use - Factor 4 (EFA)				-0.02
Data science social influence - Factor 3 (EFA)				0.26***
Data science performance augmentation - Factor 1 (EFA)				0.32***
R^2	0.120	0.343	0.383	0.519
R^2 change		0.2209	0.0396	0.1368

*** p<0.01, ** p<0.05, * p<0.1

[a] statement scale: 1 - strongly disagree; 2 - disagree; 3 - neither disagree nor agree; 4 - agree; 5 - strongly agree

[b] peers scale: 1 - level extremely below; 2 - level below; 3 - equal level; 4 - level above; 5 - level extremely above

Table 1. Regressions (1-4) are the results from the multiple regression analysis conducted in Stata software, with usage intention of data science as the dependent variable. Each value corresponds to the coefficient of the predictor variables in the different regression. R^2 and R^2 change refer to the statistical significance and respective variation each block of independent variables has in the model.

5. Discussion

The results from the research have confirmed that data science is not a priority for third sector organizations, and hence, no active pursuit exists. Little above 10% of the organizations said that they are actively pursuing the implementation of data science. In addition, only 7% have it as an investment priority (*appendix 15*) and the general lack of resources within the sector (Monteiro et al. 2015) emphasizes this disregard. Resulting from the qualitative analysis, education and funding revealed to be key resources to make the implementation of data science possible, but the quantitative data collected does not support this view. The research interviews were conducted almost exclusively with the organization's top managers and this may have clouded the most accurate conclusions. The data analysis of over 150 respondents shows that access to education and funding available for data science are not statistically significant to predict the usage intention of data science, and 87% of the respondents (*appendix 15*) said they had no contact at all with data science education in the past 2 years. Hence, leveraging education and funding has shown not to be the most effective path to get to data science adoption.

Concluded from qualitative and quantitative data was that the presence of expertise within the organization influences the intention of usage and later adoption of data science. Expertise lacks, less than 10% (*appendix 15*) of the respondents said to have the technical resources for data science implementation. And this gap is in line with the idea that the focus of the technology industry is in the for-profit sector (Balser, 2008). As concluded from the literature, expertise within an organization leads to higher probabilities of effective implementations of technology, as it reduces personnel resistance (Brownlow et al., 2015). However, social-good organizations are incapable of attracting and retaining talent due to funding constraints (Bobsin et al., 2018). One recommendation is to find an intermediate to help to fill the gap between experts and third sector organizations otherwise, organizations may fail to recognize the problem and experts to scope the projects. For instance, the Data Science for Social Good (DSSG) Association, a community of data scientists that are aware of the third sector needs and currently available to help social organizations through data science.

Similarly to UTAUT (Venkatesh et al., 2003), the presented results show that perceived performance enhancement impacts the usage intention of data science. Since what drives the sector's decision is the social impact (Bobsin et al., 2018), organizations need to perceive the benefits applied to impact.

Hence, a suggested approach, expected to have positive effects on the adoption of data science, is the development of case studies applicable to social organizations. And even though the perceived performance enhancement may be leveraged in third sector case studies, it must be complemented with other strategies. Even after the explanation of the concept of data science with examples, only 30% of the respondents (*appendix 15*) found data science a concept easy to understand.

Organizations measure themselves with respect to their peers, and the results from the research have shown that social influence impacts the data science usage intention and later adoption decision. Yet again, in accordance with Venkatesh's constructs in UTAUT (Venkatesh et al., 2003). Therefore, another suggestion can be to use social influence as a tool, to shift organizations towards data-driven cultures. For instance, investing in the promotion of successful stories of peers that benefited from the usage of data science would increase the visibility of data science benefits. Hence, spreading the word of specific tools and techniques used to increase the organization's performance and social impact.

Lastly, in what regards policies, the Government could support the bridging between the two sides of the implementation, the experts and social entities. Through a platform allowing data science expertise to flow towards social-good organizations and helping to fix the gap, and this way facilitating the encounter of the two.

Summing up, data science is not a priority within the third sector, however, there are critical resources that can be leveraged for organizations to shift for data-driven cultures. Resulting from this research, these are expertise, performance enhancement, and social influence.

6. Limitations

The present research has several limitations. The first regards the secondary data collected for the study development. Social economy has distinct characteristics among different countries and cultures. Hence, literature and background information of different contexts was considered coherent and referred to as part of the social economy. This since in Portugal clear boundaries are missing on the national social economy concepts when compared to other nations (European Commission, 2014). The second and third limitations refer to primary data collected. To many organizations, the survey was the first introduction to data science, hence people's later perceptions of the technology may have been influenced by the provided concept and examples. In addition, the length of the survey and the mandatory nature of all questions may have caused a higher number of neutral answers to the scale questions.

7. Conclusion

The conditions may be favourable for the third sector to implement data science in its daily activities, but few are the organizations and experts prepared for its implementation. The lack of resources is critical, so the solutions must aim at what impacts the most usage intention of data science. This impact may be possible through intermediaries, experienced peers, or even the Government. Nonetheless, data science must become a priority for the third sector so it can embrace it. Otherwise, there will always be other priority activities that require the allocation of most of the resources. If it does not become a priority, the sector will be the third but also the last to embrace a data-driven decision-making culture.

8. References

Aalst, Wil Van Der, and Ernesto Damiani. 2015. "Processes Meet Big Data: Connecting Data Science with Process Science." *IEEE TRANSACTIONS ON SERVICES COMPUTING*. Vol. 1.

Ajzen, Icek. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* 50 (2): 179–211.

Ashby, Deborah. 2019. "Pigeonholes and Mustard Seeds: Growing Capacity to Use Data for Society." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 1–15.

Assembleia da República. 2013. *Lei de Bases Da Economia Social - Lei n.º 30/2013, de 8 de Maio*.

Baar, Thomas, Aikaterini Deligianni, and Christoph Johann Stettina. 2016. "Data - Driven Innovation for NGO ' S : How to Define and Mobilise the Data Revolution for Sustainable Development ?"

Bacq, Sophie, and Kimberly A Eddleston. 2016. "A Resource-Based View of Social Entrepreneurship: How Stewardship Culture Benefits Scale of Social Impact." *Journal of Business Ethics* 152.

Balser, Deborah. 2008. "Book Reviews - Nonprofits & Technology: Emerging Research for Usable Knowledge, Edited by Michael Cortes and Kevin M. Rafter. Chicago: Lyceum, 2007." 562–65.

Blake, Andrew. 2019. "Dynamics of Data Science Skills, How Can All Sectors Benefit from Data Science Talent?"

Bobsin, Debora, Maira Petrini, and Marlei Pozzebon. 2018. "The Value of Technology Affordances to Improve the Management of Nonprofit Organizations Marlei Pozzebon." *RAUSP Management Journal* 54: 14–37.

Brownlow, Josh, Mohamed Zaki, Andy Neely, and Florian Urmetzer. 2015. "Data and Analytics-Data-Driven Business Models: A Blueprint for Innovation."

Brynjolfsson, Erik, Lorin Hitt, and Heekyung Kim. 2011. "Strength in Numbers: How Does Data-Driven Decision-Making Affect Firm Performance?" *International Conference on Information Systems 2011, ICIS 2011*. Vol. 1.

Brynjolfsson, Erik, and Kristina Mcelheran. 2016. "American Economic Association The Rapid Adoption of Data-Driven Decision-Making." *The American Economic Review* 106 (5): 133–39.

Brynjolfsson, Erik, and Kristina McElheran. 2016. "The Rapid Adoption of Data-Driven Decision-Making." *American Economic Review: Papers & Proceedings*. Vol. 106.

Dhar, Vasant. 2012. "Data Science and Prediction."

DSSG. 2018. "Data Science for Social Good Solve (DSSG Solve)." 2018. https://www.solveforgood.org/.

Eimhjellen, Ivar, Dag Wollebaek, and Kristin Strømsnes. 2013. "Associations Online: Barriers for

Using Web-Based Communication in Voluntary Associations."

European Commission. 2014. "Social Enterprises and Their Eco-Systems: Developments in Europe."

Franco, Raquel Campos, S. Wojciech Sokolowski, Eileen M. H. Hairel, and Lester M. Salamon. 2012. "THE PORTUGUESE NONPROFIT SECTOR IN COMPARATIVE PERSPECTIVE."

Fruchterman, Jim. 2016. "Using Data for Action and for Impact." *Stanford Social Innovation Review*, 30–35.

Hair, Joseph F., Jr. William C. Black, Barry J. Babin, and Rolph E. Anderson. 2014. *Multivariate Data Analysis*.

HBR. 2012. "The Evolution of Decision Making: How Leading Organizations Are Adopting a Data-Driven Culture." *Harvard Business Review Analytics Services*, 2012.

INE. 2019. "Conta Satélite Da Economia Social-2016." *Destaque - Instituto Nacional de Estatística*.

James Bell Associates. 2018. "Guide to Data-Driven Decision Making: Using Data to Inform Practice and Policy Decisions in Child Welfare Organizations - U.S. Department of Health and Human Services," March.

Jariego, Isidro Maya. 2007. "Book Reviews - Nonprofits & Technology: Emerging Research for Usable Knowledge, Edited by Michael Cortes and Kevin M. Rafter. Chicago: Lyceum, 2007," 412–13.

Majchrzak, Ann, and M. Lynne Markus. 2012. "TECHNOLOGY AFFORDANCES AND CONSTRAINTS IN MANAGEMENT INFORMATION SYSTEMS (MIS)."

McAfee, Andrew, and Erik Brynjolfsson. 2012. "Big Data: The Management Revolution." *Harvard Business Review*, 2012.

McNutt, John G. 2018. "Technology and Nonprofit Organizations." In *Global Encyclopedia of Public Administration, Public Policy, and Governance*, 5904–9.

Monteiro, Elisabete, Filipe Pinto, Leonor Rodrigues, Marisa Tavares, Rosário Pereira Faria, and Rosário Silva. 2015. "Diagnóstico Das ONG Em Portugal, Fundação Calouste Gulbenkian."

Myser, Suzette. 2016. “FINANCIAL HEALTH OF NONPROFIT ORGANIZATIONS.”

Scannapieco, Monica, Paolo Missier, and Carlo Batini. 2005. “Data Quality at a Glance.”

Seidel, John V. 1998. “Qualitative Data Analysis.”

Stoyan, James, and Nick Henry. 2014. “A Map of Social Enterprises and Their Eco-Systems in Europe Country Report: Portugal.”

Venkatesh, Viswanath, Robert H Smith, Michael G Morris, Gordon B Davis, Fred D Davis, and Sam M Walton. 2003. “User Acceptance of Information Technology: Toward a Unified View.” Vol. 27.

Venkatesh, Viswanath, James Y.L. Thong, and Xin Xu. 2012. “Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology.”

Waller, Matthew A., and Stanley E. Fawcett. 2013. “Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management.” *Journal of Business Logistics* 34 (2): 77–84.

Wernerfelt, Birger. 1984. “A Resource-Based View of the Firm.” *Source: Strategic Management Journal*. Vol. 5.

West, Amy. 2019. “Data-Driven Decision Making for Not-for-Profit Organizations.” *CPA Journal*.

9. Appendices

Appendix 1: Interviewees' list with the correspondent length of the interview

#	Organization	Type of organization	Approximate Duration
1	Vieira de Almeida and Associates	Law Firm	30 minutes
2	APDP	Altruistic Association (IPSS)	30 minutes
3	Tempos Brilhantes	Altruistic Association (IPSS)	40 minutes
4	PPL Crowdfunding	Social Enterprise	30 minutes
5	WeCareOn	Social Enterprise	20 minutes
6	Entrajuda	Altruistic Association (IPSS)	50 minutes
7	Unicef Portugal	Altruistic Association (ONGD)	30 minutes
8	Cozinha Com Alma	Altruistic Association (IPSS)	60 minutes

Appendix 2: Interview general script

1. **Project introduction and recording consent**
2. **Initial questions:**
    a. Please tell about your organization's activities and mission.
3. **Asses current statues of IT and data science levels:**
    a. Does the organization currently have IT infrastructure? Hardware, software, IT team.
    b. Has the organization been having any sort of education on data science? Training, workshops.
4. **Assess methods used for impact measurement:**
    a. Does the organization have methods to measure its impact?
    b. How is it usually done?
5. **Data within the organization:**
    a. Does the organization have data stored?
    b. On whom does the organization have data? Do you consider it to be quality data?
    c. For how long has the organization been storing data?
    d. Do individuals resist to share their data?
6. **Data for Change program:**

a. What reasons lead the organization to apply for the Data for Change program?

b. Was there a pre-defined problem seeking for data science solutions?

c. What were the expectations regarding the impact of the program internal to the organization? Organizational culture, structure, management, efficiency.

d. What were the expectations regarding the impact of the program external to the organization? Impact on beneficiaries, mission.

e. Within the organization, from whom came the idea to participate in the program?

f. *(For non-winners of Data for Change)* Even though you did not win the program, do you keep seeking for data science integration within the organization?

7. **Affordances and Constraints:**

   a. What do you think data science allows the organization to do more?

   b. What do you think data science prevents the organization from doing?

8. **Perceived critical resources, present or missing, for data science implementation:**

   a. What does the organization already have to make data science implementation possible?

   b. What is missing for the organization to make data science implementation possible?

9. **Data science implementation:**

   a. Is data science implementation a priority to the organization?

   b. Are you aware of the benefits data science may bring to the organization?

   c. Is the perceived effort of data science implementation superior to data science benefits?

   d. Can you see data science applications across several areas? Besides the one looking to implement through Data for Change?

   e. What are the greatest challenges to the implementation of data science?

   f. Which of the following resources do you see as more critical to the implementation? Time, education, expertise or financial resources? Which one is lacking the most?

Appendix 3: Survey to the Portuguese social-good-oriented organizations

Questionário às Organizações Portuguesas com Impacto Social

Em conjunto com o Data Science Knowledge Center (DSKC) da Nova School of Business and Economics (Nova SBE), a estudante do Mestrado Internacional em Gestão, Mariana Bandeira, está a desenvolver uma tese que relaciona Organizações de Impacto Social portuguesas com Tecnologia.

O projeto e o questionário infra visam a compreensão do estado atual da Tecnologia Data Science (Ciência de Dados) e a sua intenção de uso futuro em prol do Impacto Social.

Os dados terão uma utilização meramente estatística para efeitos do estudo apresentado e serão tratados de forma anónima e aglomerada.

Organização

1. Qual o nome da organização a que pertence? *

Insira sua resposta

2. Qual o formato legal da organização? *

- Associação
- Cooperativa
- Fundação
- Empresa Social
- Outra

3. Possui algum dos seguintes estatutos? *

- IPSS
- ONGD
- ONGA
- ONGPD
- Mera Utilidade Pública
- Nenhum dos anteriores
- Outra

4. A qual dos seguintes intervalos pertence a sua idade? *

- Menos de 25 anos
- 25-35 anos
- 36-45 anos
- 46-55 anos
- 56-65 anos
- Mais de 65 anos
- Prefiro não responder

5. Qual o intervalo de idades mais adequado para a maioria dos membros na organização? *

- Menos de 25 anos
- 25-35 anos
- 36-45 anos
- 46-55 anos
- 56-65 anos
- Mais de 65 anos

6. Qual o intervalo de idades mais adequado para a maioria dos membros da direção na organização? *

- Menos de 25 anos
- 25-35 anos
- 36-45 anos
- 46-55 anos
- 56-65 anos
- Mais de 65 anos

7. Por quantas pessoas é constituída a organização? *

Inclui colaboradores/empregados e voluntários

- Menos de 50
- 50 - 250
- 251 - 500
- 501 - 750
- 751 - 1000
- Mais de 1000

8. Tendo em conta as suas funções principais, a organização consegue dar resposta a toda a procura/exigências/pedidos que tem? *

- Sim
- Não

9. Qual o formato mais indicado para a atual estrutura da organização? *

- Vertical (estrutura hierárquica rígida, direção toma as decisões)
- Horizontal (funcionários/voluntários têm autonomia e tomam decisões)
- Outra

Infraestrutura de IT (Information Technology)

10. Em relação a infraestrutura tecnológica, a organização tem: *

- Computadores pertencentes (à organização)
- Um programa/software especializado para um tipo de operação da organização (p.e. SAP)
- Plataforma de comunicação integrada (p.e. Microsoft Teams)
- Website
- App
- Outra

11. Em relação aos equipamentos e tecnologias *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo completamente	Discordo	Não discordo nem concordo	Concordo	Concordo completamente
O equipamento da organização é antiquado (computadores, infraestrutura IT)					
A tecnologia utilizada na organização está desatualizada					
O material tecnológico que a organização tem foi doado por terceiros					
O material que a organização tem foi comprado pela organização					

12. Tem uma equipa dedicada unicamente a IT (Information Technology)? *

- Sim
- Não

13. Na organização, a equipa de IT é: *

- Parte integrante da organização, com formação na área
- Parte integrante da organização, sem formação na área
- Contratada externamente
- Não existe equipa de IT
- Outra

14. O website/plataforma da organização: *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo Completamente	Discordo	Não discordo nem concordo	Concordo	Concordo Completamente
É apenas informativo					
Tem cookies					
Hospeda operações cruciais ao funcionamento da organização (p.e. agendamento de consultas)					
Todas as funções do site implicam intervenção humana para funcionar (zero automatismo)					
Tem funções automatizadas (p.e. compra de bilhetes)					
Tem Inteligência Artificial (p.e. chatbot)					
Tem armazenamento automático de dados (recolhe e armazena dados inseridos pelos utilizadores)					
Tem processamento de dados (p.e. estatísticas, gera resultados de acordo com os dados inseridos)					
A organização não possui website/plataforma					

Data Science (Ciência de Dados)

15. Está familiarizado com o conceito de Data Science? *

- Sim
- Não

16. O que entende por Data Science? *

Se lhe for mais conviente pode escrever por tópicos.

Insira sua resposta

Data Science e aplicações

Data Science é um conceito multi-disciplinar e é compreendido por outras tecnologias como Big Data, Inteligência Artificial (IA), Machine Learning, entre outros.
Uma das suas muitas capacidades é a previsão de resultados futuros baseando-se em dados passados.
Data Science para previsão ou estimativas envolve a análise de dados. O desenvolvimento de métodos de recolha, armazenamento e análise de dados para extrair informações úteis.
Tem como objetivo principal gerar percepções e conhecimento de qualquer tipo de dados - tanto estruturados como não estruturados.

Pedimos, por favor, que não altere a sua resposta na secção anterior.
Ao alterar estará a influenciar, de uma forma negativa, a precisão da nossa análise do sector.

Exemplo 1: Angariações de Fundos - com dados recolhidos de angariações de fundos passadas, é possível estimar tipos de doador e quantidades doadas para uma angariação de fundos futura, e com isto saber quem abordar aquando de uma nova angariação de fundos.

Exemplo 2: Chatbot - janelas pop-up de chats que simulam uma conversa humana e que tem por base um programa de computador que trabalha através de comandos de voz, conversas de texto ou ambos. Recorre a Inteligência Artificial (IA) e pode ser incorporado em aplicações de mensagens por exemplo de websites.

Exemplo 3: Churn/Desistências - com informações acerca dos colaboradores é possível prever quantos e quem são os colaboradores expectáveis de deixarem a sua participação na organização e estimar as razões. Estas informações podem ser: faltas, avaliações, feedbacks e entre outros.

17. Vê benefício no uso de Data Science para a sua organização? *

- Sim
- Não

18. Porque sim? *

Se lhe for mais conviente pode escrever por tópicos.

Insira sua resposta

19. Porque não? *

Se lhe for mais conviente pode escrever por tópicos.

Insira sua resposta

20. Quantas vezes, ao longo dos últimos 2 anos, é que a sua organização teve acesso a alguma forma de educação sobre Data Science ou Transformação Digital? (p.e. formações, palestras, workshops) *

- Nenhuma
- 1 - 3 vezes
- 4 - 6 vezes
- 7 - 10 vezes
- Mais do que 10 vezes

21. Em relação à aplicação e implementação de Data Science no dia-a-dia da organização: *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo Completamente	Discordo	Não discordo nem concordo	Concordo	Concordo Completamente
Vejo aplicação de Data Science no dia-a-dia da organização					
Vejo benefícios de Data Science no dia-a-dia da organização					
Vejo obstáculos de Data Science no dia-a-dia da organização					
A organização tem os recursos financeiros necessários para a implementação					
A organização tem os recursos técnicos/expertise necessários para a implementação					
A organização tem interesse na implementação de Data Science					

22. Quando comparando o nível de uso de Data Science com outras organizações de impacto social, a sua organização está num: *

- Nível completamente abaixo
- Nível abaixo
- Nível igual
- Nível acima
- Nível completamente acima

Recolha, Privacidade e Confidencialidade dos Dados

23. A organização está a par do Regulamento Geral sobre a Proteção de Dados (RGPD)? *

- Sim
- Não

24. A organização tem um DPO (Data Protection Officer/Encarregado de Proteção de Dados)? *

- Sim
- Não

25. Em relação à recolha de dados feita pela organização:
(Inclui recolha de dados através de qualquer formato - p.e.: softwares com histórico, questionários online/papel.) *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo Completamente	Discordo	Não discordo nem concordo	Concordo	Concordo Completamente
Existe o hábito de recolha de dados digitalmente (p.e. cloud/nuvem, ficheiros excel, software)					
Existe o hábito de recolha de dados fisicamente (p.e. escrita manual, pastas com dados passados, ficheiros impressos)					
Os dados armazenados digitalmente contém qualidade (contém informação relevante para a organização)					
Os dados armazenados fisicamente/papel têm qualidade (contém informação relevante para a organização)					
Os dados armazenados digitalmente são consistentes entre si (mesmo nível de informação)					
Os dados armazenados fisicamente/papel são consistentes entre si (mesmo nível de informação)					
A organização sabe que dados procura recolher digitalmente					

A organização sabe como retirar conhecimento dos dados recolhidos digitalmente					
A organização sabe agir de acordo com os resultados dos dados que tem armazenados digitalmente					
A organização tem como hábito tomar decisões baseadas em dados (digitalmente ou fisicamente/papel armazenados)					
A organização tem como objetivo tomar decisões baseadas em dados (digitalmente ou fisicamente/papel armazenados)					
A organização tem como hábito tomar decisões baseadas em intuição					
A organização tem como objetivo tomar decisões baseadas em intuição					

26. Como descreveria os Dados que a organização tem armazenados?

Os dados recolhidos ao longo do questionário serão utilizados apenas anonimamente. *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo Completamente	Discordo	Não discordo nem concordo	Concordo	Concordo Completamente
Os dados que a organização tem são de elevada sensibilidade					
Os dados que a organização tem são de elevada confidencialidade					
Os dados que a organização tem estão acessíveis a todos os colaboradores					
Os dados recolhidos têm todos consentimento por parte dos utilizadores					
A recolha de dados cumpre o Regulamento Geral sobre a Proteção de Dados (RGPD)					
Os utilizadores dão os seus dados pessoais sem reticência					
Os utilizadores questionam a partilha dos seus dados pessoais					

Aceitação de Data Science

27. Em relação a Data Science e ao seu impacto na organização: *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo Completamente	Discordo	Não discordo nem concordo	Concordo	Concordo Completamente
A implementação de Data Science requer mais esforço do que traz benefícios	○	○	○	○	○
Data Science aumenta o impacto social que a organização tem	○	○	○	○	○
Data Science implicaria esforço para a organização	○	○	○	○	○
Data Science é uma boa ideia para a organização	○	○	○	○	○
Data Science é algo fácil de entender	○	○	○	○	○
Data Science é algo fácil de trabalhar com	○	○	○	○	○
Data Science é um conceito intuitivo	○	○	○	○	○
Data Sience é demasiado complicado de implementar	○	○	○	○	○
A organização tenciona usar Data Science nos próximos 2 anos	○	○	○	○	○
A organização prevê o uso de Data Science nos próximos 2 anos	○	○	○	○	○
A organização planeia usar Data Science nos próximos 2 anos	○	○	○	○	○
Data Science aumenta a eficiência da organização	○	○	○	○	○
Investir em Data Science iria desviar a organização dos seus objetivos principais	○	○	○	○	○
A direção da organização implementa novas iniciativas na organização	○	○	○	○	○
Organizações semelhantes que já utilizam Data Science influenciam a nossa intenção de utilização	○	○	○	○	○

Organizações de referência para nós influenciam a nossa intenção utilização de Data Science	○	○	○	○	○
Organizações de referência para nós incentivam a nossa utilização de Data Science	○	○	○	○	○
Pessoas externas que se preocupam com a nossa organização incentivam a implementação de Data Science	○	○	○	○	○

Vantagens e Desvantagens de Data Science

28. Em relação à implementação de Data Science na organização, quais as maiores barreiras? *

Ordene do maior para o mais pequeno obstáculo

- Falta de estrutura tecnológica (computadores, website)
- Falta de pessoal qualificado em Data Science
- Dificuldades em atrair pessoas qualificadas em Data Science
- Falta de interesse por parte da direção
- Falta de tempo para novos projetos
- Dificuldades de financiamento
- Falta de apoio financeiro do Governo
- Não sabemos o que é Data Science e os benefícios que pode trazer
- Não temos interesse na implementação de Data Science

29. Por favor liste 3 ou mais vantagens que vê no uso de Data Science na sua organização. *

Se lhe for mais conveniente pode escrever por tópicos.

Insira sua resposta

30. Por favor liste 3 ou mais desvantagens que vê no uso de Data Science na sua organização. *

Se lhe for mais conveniente pode escrever por tópicos.

Insira sua resposta

31. O que acha que a sua organização já tem para tornar possível a implementação de Data Science? *

Se lhe for mais conveniente pode escrever por tópicos.

Insira sua resposta

32. O que acha que falta à sua organização para tornar possível a implementação de Data Science? *

Se lhe for mais conveniente pode escrever por tópicos.

Insira sua resposta

33. O que acha que Data Science pode fazer mais pela sua organização? *

Se lhe for mais conveniente pode escrever por tópicos.

Insira sua resposta

34. O que acha que Data Science pode impedir a sua organização de fazer? *

Se lhe for mais conveniente pode escrever por tópicos.

Insira sua resposta

Data Science

35. A organização já esteve envolvida anteriormente em algum projeto relacionado com Data Science? *

- Sim
- Não

36. No que respeita ao investimento em Data Science *

Numa escala de Discordo Completamente a Concordo Completamente selecione

	Discordo Completamente	Discordo	Não discordo nem concordo	Concordo	Concordo Completamente
O investimento em Data Science é uma prioridade da organização					
O investimento em Data Science é crucial para o futuro da organização					
O investimento em Data Science, a acontecer, será por consequência de outras decisões					
A organização procura ativamente o uso de Data Science					

Conclusão

37. Deixe-nos um comentário em relação ao tópico de Data Science no Sector de Impacto Social. *

Insira sua resposta

38. Deixe-nos o seu email se gostaria de ser contactado/a para futuras questões relacionadas com o tema.

Insira sua resposta

Appendix 4: Items used in estimating UTAUT

Performance Expectancy (Venkatesh et al. 2003)
Do you see benefits in using Data Science for your organization?
I see Data Science application in my organization's daily life
I see benefits of Data Science in the daily life of the organization
Data Science increases the social impact that the organization has
Data Science increases the efficiency of the organization
Investing in Data Science would divert the organization from its main goals
Effort Expectancy (Venkatesh et al. 2003)
I see obstacles of Data Science in the daily life of the organization
Implementing Data Science requires more effort than it brings benefits
Data Science would involve effort for the organization
Data Science is easy to understand
Data Science is easy to work with
Data Science is an intuitive concept
Data Science is too complicated to implement
Social Influence (Venkatesh et al. 2003)
Similar organizations that already use Data Science influence our intention to use Data Science
Reference organizations for us influence our intention to use Data Science
Reference organizations for us encourage our use of Data Science
Outside people who care about our organization encourage the implementation of Data Science
Facilitating Conditions (Venkatesh et al. 2003)
The organization has the necessary financial resources for the implementation of data science
The organization has the necessary technical resources/expertise for the implementation of data science
Behavioral Intention to use Data Science (Venkatesh et al. 2003)
The organization intends to use Data Science in the next 2 years
The organization foresees the use of Data Science in the next 2 years
The organization plans to use Data Science in the next 2 years
Attitude toward using Data Science (Venkatesh et al. 2003)
The organization is interested in the implementation of Data Science
Data Science is a good idea for the organization

Note: The first four categories are the constructs of the UTAUT (Venkatesh et al. 2003) that are expected to influence the last two categories, usage intention and attitude towards a technology. All the items were drawn from Venkatesh's Unified Theory of Acceptance and Use of Technology (2003).

Appendix 5: Qualitative Data Analysis (QDA) - selected codes and respective examples

Code	Example (translated)
DS_new_possibilities	"We get to know the concrete problem (…) increase the donated values"
DS_challenges	"Social barrier (…) afraid of asking more information from the users"
DS_limitations	“I think you have to be very careful in the social area (...) (as people) cannot be replaced by technology”
DM_power_knowledge	"They (top management) can't see what it (data science) is"
Data_availability	"7 years (of data stored)"
DS_access	"There is a platform with data for the professionals (…) another for the pacients with limited access to data"
Data4Change	"(Applied because) there is a lot of information but poorly worked (…) do not take the best out of what the program can give us"
IT_infrastructure	"IT (team) until now has been fully outsourced"
IT_team	“We have the (IT) infrastructure, that are computers (…), one server, a personal software and a software in the cloud”
DS_priority	"(Not something you pursue actively?) Yes, that's it"
Resources_available_to_implement	"We have the data"
Resources_lacking_to_implement	"There is no money"
DS_perceived_usefulness	"(…) automatically make classroom and teacher management more efficient"
DS_perceived_effort	"It takes a lot of willpower to implement this in the institutions"
Impact_Measurement	"We want also to measure the impact"

Note: Codes correspond to labelled topics that were referred to in the 7 interviews with organizations. Examples are segments of the conversations transcribed and translated, in the case of conversations ran in Portuguese, from the interviews’ recordings.

Appendix 6: Variables' names and respective survey questions

Variable Name	Question
Org_name	Qual o nome da organização a que pertence?
LF_org	Qual o formato legal da organização?
Statute	Possui algum dos seguintes estatutos?
Age_resp	A qual dos seguintes intervalos pertence a sua idade?
Age_org_members	Qual o intervalo de idades mais adequado para a maioria dos membros na organização?
Age_board	Qual o intervalo de idades mais adequado para a maioria dos membros da direção na organização?
Org_size	Por quantas pessoas é constituída a organização?
Demand_matching	Tendo em conta as suas funções principais, a organização consegue dar resposta a toda a procura/exigências/pedidos que tem?
Org_gov_type	Qual o formato mais indicado para a atual estrutura da organização?
IT_infrastructure	Em relação a infraestrutura tecnológica, a organização tem:
IT_equip_obsolete	O equipamento da organização é antiquado (computadores, infraestrutura IT)
Tech_obsolete	A tecnologia utilizada na organização está desatualizada
Tech_donated	O material tecnológico que a organização tem foi doado por terceiros
Tech_bought	O material que a organização tem foi comprado pela organização
Dedicated_IT_team	Tem uma equipa dedicada unicamente a IT (Information Technology)?
IT_relation	Na organização, a equipa de IT é:
Website_inf_only	É apenas informativo
Webiste_cookies	Tem cookies
Website_critical_func	Hospeda operações cruciais ao funcionamento da organização (p.e. agendamento de consultas)
Website_manual	Todas as funções do site implicam intervenção humana para funcionar (zero automatismo)
Website_automated	Tem funções automatizadas (p.e. compra de bilhetes)
Website_AI	Tem Inteligência Artificial (p.e. chatbot)
Website_storage_automated	Tem armazenamento automático de dados (recolhe e armazena dados inseridos pelos utilizadores)
Website_data_processing	Tem processamento de dados (p.e. estatísticas, gera resultados de acordo com os dados inseridos)
Org_wout_website_yn	A organização não possui website/plataforma
DS_concept_yn	Está familiarizado com o conceito de Data Science?
DS_interpretation	O que entende por Data Science?
DS_beneficial	Vê benefício no uso de Data Science para a sua organização?
Why_DS_beneficial	Porque sim?
Why_not_DS_beneficial	Porque não?
DS_ed_access	Quantas vezes, ao longo dos últimos 2 anos, é que a sua organização teve acesso a alguma forma de educação sobre Data Science ou Transformação Digital? (p.e. formações, palestras, workshops)
DS_applicable_4ops	Vejo aplicação de Data Science no dia-a-dia da organização
DS_beneficial_4ops	Vejo benefícios de Data Science no dia-a-dia da organização
DS_obstacles_4ops	Vejo obstáculos de Data Science no dia-a-dia da organização
Funds_4DS_available	A organização tem os recursos financeiros necessários para a implementação
TechRes_4DS_av	A organização tem os recursos técnicos/expertise necessários para a implementação
Interest_4DS_impl	A organização tem interesse na implementação de Data Science
DSuse_relative2peers	Quando comparando o nível de uso de Data Science com outras organizações de impacto social, a sua organização está num:
GDPR_regulation_knowledge	A organização está a par do Regulamento Geral sobre a Proteção de Dados (RGPD)?
DPO_internally	A organização tem um DPO (Data Protection Officer/Encarregado de Proteção de Dados)?
Data_collection_digital	Existe o hábito de recolha de dados digitalmente (p.e. cloud/nuvem, ficheiros excel, software)
Data_collection_physical	Existe o hábito de recolha de dados fisicamente (p.e. escrita manual, pastas com dados passados, ficheiros impressos)
DigData_quality	Os dados armazenados digitalmente contém qualidade (contém informação relevante para a organização)
PhysData_quality	Os dados armazenados fisicamente/papel têm qualidade (contém informação relevante para a organização)
DigData_consistent	Os dados armazenados digitalmente são consistentes entre si (mesmo nível de informação)
PhysData_consist	Os dados armazenados fisicamente/papel são consistentes entre si (mesmo nível de informação)
DigData_needs_understanding	A organização sabe que dados procura recolher digitalmente
DigData__processing_skills	A organização sabe como retirar conhecimento dos dados recolhidos digitalmente
DDriven_actions_knowledge	A organização sabe agir de acordo com os resultados dos dados que tem armazenados digitalmente
DDDM_routine	A organização tem como hábito tomar decisões baseadas em dados (digitalmente ou fisicamente/papel armazenados)
DDDM_as_objective	A organização tem como objetivo tomar decisões baseadas em dados (digitalmente ou fisicamente/papel armazenados)

DM_intuitive_routine	A organização tem como hábito tomar decisões baseadas em intuição
DM_intuitive_objective	A organização tem como objetivo tomar decisões baseadas em intuição
Data_sensitive	Os dados que a organização tem são de elevada sensibilidade
Data_confidential	Os dados que a organização tem são de elevada confidencialidade
Data_open_access_internal	Os dados que a organização tem estão acessíveis a todos os colaboradores
Data_collected_with_consensus	Os dados recolhidos têm todos consentimento por parte dos utilizadores
Data_collectied_GDPR	A recolha de dados cumpre o Regulamento Geral sobre a Proteção de Dados (RGPD)
DataCollected_wout_resistance	Os utilizadores dão os seus dados pessoais sem reticência
Users_questioning_data_coll	Os utilizadores questionam a partilha dos seus dados pessoais
DS_not_optimal	A implementação de Data Science requer mais esforço do que traz benefícios
DS_impact_aug	Data Science aumenta o impacto social que a organização tem
DS_effort	Data Science implicaria esforço para a organização
DS_good_idea	Data Science é uma boa ideia para a organização
DS_easy_to_understand	Data Science é algo fácil de entender
DS_easy2work	Data Science é algo fácil de trabalhar com
DS_intuitive	Data Science é um conceito intuitivo
DS_imp_complexity	Data Sience é demasiado complicado de implementar
IntUse_DS	A organização tenciona usar Data Science nos próximos 2 anos
IntUse_DS_pred	A organização prevê o uso de Data Science nos próximos 2 anos
IntUse_DS_plan	A organização planeia usar Data Science nos próximos 2 anos
DS_aug_eff	Data Science aumenta a eficiência da organização
DS_investment_focus_dev	Investir em Data Science iria desviar a organização dos seus objetivos principais
Board_likes_new_initiatives	A direção da organização implementa novas iniciativas na organização
Peer_inf_DS_decisions	Organizações semelhantes que já utilizam Data Science influenciam a nossa intenção de utilização
Peer_ref_inf_DS_decisions	Organizações de referência para nós influenciam a nossa intenção utilização de Data Science
Peer_ref_incentivize_DS_use	Organizações de referência para nós incentivam a nossa utilização de Data Science
Peer_ext_incentivize_DS_use	Pessoas externas que se preocupam com a nossa organização incentivam a implementação de Data Science
Barriers_to_implement_DS	Em relação à implementação de Data Science na organização, quais as maiores barreiras?
DS_3_adv	Por favor liste 3 ou mais vantagens que vê no uso de Data Science na sua organização.
DS_3_disadv	Por favor liste 3 ou mais desvantagens que vê no uso de Data Science na sua organização.
DS_available_resources	O que acha que a sua organização já tem para tornar possível a implementação de Data Science?
DS_missing_resources	O que acha que falta à sua organização para tornar possível a implementação de Data Science?
What_DS_can_do_4org	O que acha que Data Science pode fazer mais pela sua organização?
DS_prevents	O que acha que Data Science pode impedir a sua organização de fazer?
DS_projects_existence	A organização já esteve envolvida anteriormente em algum projeto relacionado com Data Science?
DS_investment_priority	O investimento em Data Science é uma prioridade da organização
DS_investment_vital_future	O investimento em Data Science é crucial para o futuro da organização
DS_side_effect	O investimento em Data Science, a acontecer, será por consequência de outras decisões
DS_actively_pursued	A organização procura ativamente o uso de Data Science
DS_comment	Deixe-nos um comentário em relação ao tópico de Data Science no Sector de Impacto Social.
Willingness_to_collaborate	Deixe-nos o seu email se gostaria de ser contactado/a para futuras questões relacionadas com o tema.
W2Collaborate_binary	If the email in *Willingness_to_collaborate* was provided it was considered as 1, if not 0.

Note: There are 98 variables from which 5 were generated automatically by Microsoft Office Forms that are omitted from the listing. Being these, ID number, date and time from starting to finishing the survey, email and name (no data entrances since the identification was not mandatory).

Appendix 7: Scales from survey questions and correspondence with each scale variable

Scale	Age group	Dimension (people)	Statement	Frequency	Peers
0	-	-	-	None	-
1	Less than 25 years old	Less than 50	Strongly disagree	1 - 3 times	Level extremely below
2	25 - 35 years old	50 - 250	Disagree	4 - 6 times	Level below
3	36 - 45 years old	251 - 500	Neither disagree nor agree	6 - 10 times	Equal level
4	46 - 55 years old	501 - 750	Agreee	More than 10 times	Level above
5	56 - 65 years old	751 - 1000	Strongly agree	-	Level extremely above
6	More than 65 years old	More than 1000	-	-	-

Appendix 8: Scales correspondent to each variable within the survey scale questions

Variable name	Scale type
Age_resp	Age group
Age_org_members	Age group
Age_board	Age group
Org_size	Dimension
Board_likes_new_initiatives	Statement
IT_equip_obsolete	Statement
Tech_obsolete	Statement
Tech_donated	Statement
Tech_bought	Statement
Website_inf_only	Statement
Webiste_cookies	Statement
Website_critical_func	Statement
Website_manual	Statement
Website_automated	Statement
Website_AI	Statement
Website_storage_automated	Statement
Website_data_processing	Statement
Org_wout_website_yn	Statement
Data_collectied_GDPR	Statement
DS_ed_access	Frequency
DS_applicable_4ops	Statement
DS_beneficial_4ops	Statement
DS_impact_aug	Statement
DS_aug_eff	Statement
DS_investment_focus_dev	Statement
DS_obstacles_4ops	Statement
DS_not_optimal	Statement
DS_effort	Statement
DS_easy_to_understand	Statement
DS_easy2work	Statement
DS_intuitive	Statement
DS_imp_complexity	Statement
Peer_inf_DS_decisions	Statement
Peer_ref_inf_DS_decisions	Statement
Peer_ref_incentivize_DS_use	Statement
Peer_ext_incentivize_DS_use	Statement
Funds_4DS_available	Statement
TechRes_4DS_av	Statement
IntUse_DS	Statement
IntUse_DS_pred	Statement
IntUse_DS_plan	Statement
Interest_4DS_impl	Statement
DS_good_idea	Statement
DS_investment_priority	Statement
DS_investment_vital_future	Statement
DS_side_effect	Statement
DS_actively_pursued	Statement
DSuse_relative2peers	Peers
Data_collection_digital	Statement
Data_collection_physical	Statement
DigData_quality	Statement
PhysData_quality	Statement
DigData_consistent	Statement
PhysData_consist	Statement
DigData_needs_understandi	Statement
DigData__processing_skills	Statement
DDriven_actions_knowledg	Statement
DDDM_routine	Statement
DDDM_as_objective	Statement
DM_intuitive_routine	Statement
DM_intuitive_objective	Statement
Data_sensitive	Statement
Data_confidential	Statement
Data_open_access_internal	Statement
Data_collected_with_consen	Statement
DataCollected_wout_resista	Statement
Users_questioning_data_coll	Statement

Appendix 9: Factor's Eigenvalues generated in Stata

```
Factor analysis/correlation                     Number of obs    =        159
    Method: principal factors                   Retained factors =          4
    Rotation: (unrotated)                       Number of params =         78
```

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	7.40723	5.27733	0.5461	0.5461
Factor2	2.12991	0.52756	0.1570	0.7031
Factor3	1.60234	0.11889	0.1181	0.8212
Factor4	1.48346	0.80167	0.1094	0.9306
Factor5	0.68178	0.16940	0.0503	0.9808
Factor6	0.51238	0.22103	0.0378	1.0186
Factor7	0.29135	0.11649	0.0215	1.0401
Factor8	0.17486	0.02633	0.0129	1.0530
Factor9	0.14854	0.08767	0.0109	1.0639
Factor10	0.06087	0.04749	0.0045	1.0684
Factor11	0.01338	0.00542	0.0010	1.0694
Factor12	0.00796	0.02285	0.0006	1.0700
Factor13	-0.01489	0.01763	-0.0011	1.0689
Factor14	-0.03252	0.02769	-0.0024	1.0665
Factor15	-0.06021	0.01353	-0.0044	1.0620
Factor16	-0.07375	0.02754	-0.0054	1.0566
Factor17	-0.10128	0.01830	-0.0075	1.0491
Factor18	-0.11958	0.02652	-0.0088	1.0403
Factor19	-0.14610	0.02351	-0.0108	1.0295
Factor20	-0.16961	0.06161	-0.0125	1.0170
Factor21	-0.23122	.	-0.0170	1.0000

```
LR test: independent vs. saturated: chi2(210) = 2563.97 Prob>chi2 = 0.0000
```

Note: First four values of Eigenvalue correspond to the four selected constructs, clearly standing out with values greater than one.

Appendix 10: Factor loadings generated in Stata

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Factor4	Uniqueness
DS_applica~s	0.7478				0.4388
DS_aug_eff	0.7804				0.3421
DS_benefic~s	0.9261				0.2740
DS_easy_to~d				0.7796	0.3630
DS_easy2work				0.8732	0.1908
DS_effort	0.4152	-0.3623			0.8356
DS_good_idea	0.7994				0.2501
DS_imp_com~y				-0.4340	0.6563
DS_impact_~g	0.7461				0.3924
DS_intuitive				0.5683	0.7022
DS_investm~v	-0.3732				0.8020
DS_not_opt~l	-0.4846				0.7599
DS_obstacl~s				-0.4302	0.7991
Interest_4~l	0.6087				0.4051
IntUse_DS		0.8605			0.1314
IntUse_DS_~n		0.9086			0.0682
IntUse_DS_~d		0.9218			0.0567
Peer_ext_i~e			0.6157		0.4510
Peer_inf_D~s			0.9577		0.1319
Peer_ref_i~e			0.8523		0.1858
Peer_ref_i~s			0.9492		0.1404

(blanks represent abs(loading)<.3)

Note: For clear identification of the items to consider in each of the four factors, the factor loadings below 0.3 were not exhibited.

Appendix 11: Items from UTAUT (Venkatesh et al. 2003) used in the exploratory factor analysis

Statements from UTAUT used in exploratory factor analysis (EFA)	Variable name	Loading factor
Performance Expectancy (Venkatesh et al. 2003)		
I see Data Science application in my organization's daily life	DS_applicable_4ops	0.7478
I see benefits of Data Science in the daily life of the organization	DS_beneficial_4ops	0.9261
Data Science increases the social impact that the organization has	DS_impact_aug	0.7461
Data Science increases the efficiency of the organization	DS_aug_eff	0.7804
Investing in Data Science would divert the organization from its main goals	DS_investment_focus_dev	-0.3732
Effort Expectancy (Venkatesh et al. 2003)		
I see obstacles of Data Science in the daily life of the organization	DS_obstacles_4ops	-0.4302
Implementing Data Science requires more effort than it brings benefits	DS_not_optimal	-0.4846
Data Science would involve effort for the organization	DS_effort	-
Data Science is easy to understand	DS_easy_to_understand	0.7796
Data Science is easy to work with	DS_easy2work	0.8732
Data Science is an intuitive concept	DS_intuitive	0.5683
Data Sience is too complicated to implement	DS_imp_complexity	-0.4340
Social Influence (Venkatesh et al. 2003)		
Similar organizations that already use Data Science influence our intention of use	Peer_inf_DS_decisions	0.9577
Reference organizations for us influence our intention to use Data Science	Peer_ref_inf_DS_decisions	0.9492
Reference organizations for us encourage our use of Data Science	Peer_ref_incentivize_DS_use	0.8523
Outside people who care about our organization encourage the implementation of Data Science	Peer_ext_incentivize_DS_use	0.6157
Behavioral Intention to use Data Science (Venkatesh et al. 2003)		
The organization intends to use Data Science in the next 2 years	IntUse_DS	0.8605
The organization foresees the use of Data Science in the next 2 years	IntUse_DS_pred	0.9218
The organization plans to use Data Science in the next 2 years	IntUse_DS_plan	0.9086
Attitude toward using Data Science (Venkatesh et al. 2003)		

The organization is interested in the implemention of Data Science	Interest_4DS_impl	0.6087
Data Science is a good idea for the organization	DS_good_idea	0.7994

Note: Scales are referred in *appendix 7*

Appendix 12: Scale names and respective Cronbach's alpha generated in Stata

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**
Scale name	scale_DS_augm	scale_DS_usage_intention	scale_DS_social_influence	scale_DS_easy
Cronbach's alpha	**0.8808**	**0.9744**	**0.9472**	**0.7824**

Note: required minimum value of 0.7 (Hair et al. 2014)

Appendix 13: Variables used in the multiple regression analysis and corresponding survey questions

Variable name	Questions and Factors
Age_resp	A qual dos seguintes intervalos pertence a sua idade?
Age_org_members	Qual o intervalo de idades mais adequado para a maioria dos membros na organização?
Age_board	Qual o intervalo de idades mais adequado para a maioria dos membros da direção na organização?
Org_size	Por quantas pessoas é constituída a organização?
Demand_matching	Tendo em conta as suas funções principais, a organização consegue dar resposta a toda a procura/exigências/pedidos que tem?
DS_ed_access	Quantas vezes, ao longo dos últimos 2 anos, é que a sua organização teve acesso a alguma forma de educação sobre Data Science ou Transformação Digital?
Funds_4DS_available	A organização tem os recursos financeiros necessários para a implementação
TechRes_4DS_av	A organização tem os recursos técnicos/expertise necessários para a implementação
Data_collection_digital	Existe o hábito de recolha de dados digitalmente (p.e. cloud/nuvem, ficheiros excel, software)
DPO_internally	A organização tem um DPO (Data Protection Officer/Encarregado de Proteção de Dados)?
GDPR_regulation_knowledge	A organização está a par do Regulamento Geral sobre a Proteção de Dados (RGPD)?
W2Collaborate_binary	Deixe-nos o seu email se gostaria de ser contactado/a para futuras questões relacionadas com o tema.
DDDM_routine	A organização tem como hábito tomar decisões baseadas em dados (digitalmente ou fisicamente/papel armazenados)
DDDM_as_objective	A organização tem como objetivo tomar decisões baseadas em dados (digitalmente ou fisicamente/papel armazenados)
DSuse_relative2peers	Quando comparando o nível de uso de Data Science com outras organizações de impacto social, a sua organização está num:
scale_DS_easy	Factor 4 (EFA)
scale_DS_social_influence	Factor 3 (EFA)
scale_DS_augm	Factor 1 (EFA)

Appendix 14: Blocks of independent variables used in the multiple regression model

Block 1 - Control Variables	Block 2 - Resources availibility	Block 3 - Data-driven decision-making culture	Block 4 - UTAUT
i.Age_resp*	DS_ed_access	DDDM_routine	DSuse_relative2peers
i.Age_org_members*	Funds_4DS_available	DDDM_as_objective	scale_DS_augm[a]
i.Age_board*	TechRes_4DS_av		scale_DS_social_influence[a]
i.Org_size*	Data_collection_digital		scale_DS_easy[a]
Demand_matching	DPO_internally		
	GDPR_regulation_knowledge		
	W2Collaborate_binary		

* *i* is the index that identifies the scale number. In both cases, age and size, the scales goes from 1 to 6.
[a] factors resulting from exploratory factor analysis

Appendix 15: Descriptive Statistics generated from Stata

Binary answers: yes (1), no (0)

	Proportion	Std. Err.	Logit [95% Conf.	Interval]
Demand_matching				
0	.6163522	.038564	.5379222	.6891624
1	.3836478	.038564	.3108376	.4620778
GDPR_regulation_knowledge				
0	.0377358	.0151121	.0169455	.0819087
1	.9622642	.0151121	.9180913	.9830545
DPO_internally				
0	.509434	.0396455	.4315364	.5868761
1	.490566	.0396455	.4131239	.5684636
DS_concept_yn				
0	.7484277	.0344118	.6746439	.8101867
1	.2515723	.0344118	.1898133	.3253561
W2Collaborate_binary				
0	.5471698	.0394757	.4686832	.623379
1	.4528302	.0394757	.376621	.5313168
DS_beneficial				
0	.2201258	.0328586	.1620605	.2917519
1	.7798742	.0328586	.7082481	.8379395
Dedicated_IT_team				
0	.8427673	.0288687	.7770698	.8917996
1	.1572327	.0288687	.1082004	.2229302
DS_projects_existence				
0	.9308176	.0201248	.8789017	.9614529
1	.0691824	.0201248	.0385471	.1210983

Scale variables (see appendix 7 for scale's values)

Scale questions - Organization

	Proportion	Std. Err.	Logit [95% Conf.	Interval]
Age_resp				
1	.0125786	.0088383	.0031145	.0493771
2	.1761006	.0302078	.124095	.2438336
3	.3207547	.037017	.2523949	.3977782
4	.3144654	.0368216	.2466276	.3912718
5	.1069182	.024506	.0672633	.165796
6	.0691824	.0201248	.0385471	.1210983
Age_org_members				
1	.0314465	.0138404	.0130604	.0737815
2	.1823899	.0306249	.1294473	.2507488
3	.4465409	.0394253	.3705792	.5250837
4	.2264151	.0331901	.1675721	.2985111
5	.0691824	.0201248	.0385471	.1210983
6	.0440252	.0162695	.0210109	.0899325
Age_board				
2	.0440252	.0162695	.0210109	.0899325
3	.2327044	.0335108	.1731037	.3052506
4	.3207547	.037017	.2523949	.3977782
5	.3333333	.0373848	.263973	.410748
6	.0691824	.0201248	.0385471	.1210983
Org_size				
1	.5408805	.0395198	.4624621	.6173252
2	.3396226	.0375574	.2697833	.4172116
3	.0566038	.0183261	.0295626	.1056861
4	.0062893	.0062695	.0008719	.0438868
5	.0188679	.0107901	.0060449	.0573236
6	.0377358	.0151121	.0169455	.0819087
Board_likes_new_initiatives				
1	.0251572	.0124194	.0094024	.0655636
2	.1006289	.0238579	.062332	.1584787
3	.3018868	.0364071	.2351376	.3782146
4	.4654088	.0395576	.3887407	.5437469
5	.1069182	.024506	.0672633	.165796

Scale questions - IT infrastructure

	Proportion	Std. Err.	Logit [95% Conf.	Interval]
IT_equip_obsolete				
1	.1320755	.0268506	.0874349	.1946462
2	.3018868	.0364071	.2351376	.3782146
3	.1886792	.0310284	.1348247	.2576395
4	.2955975	.0361878	.2294154	.3716634
5	.081761	.0217296	.0478652	.1362258
Tech_obsolete				
1	.1509434	.0283907	.1029603	.2159061
2	.3459119	.0377227	.2756077	.4236614
3	.2327044	.0335108	.1731037	.3052506
4	.2389937	.0338212	.1786548	.311971
5	.0314465	.0138404	.0130604	.0737815
Tech_donated				
1	.2704403	.0352263	.2066853	.3453014
2	.2138365	.0325161	.1565694	.2849725
3	.1069182	.024506	.0672633	.165796
4	.2641509	.0349641	.201044	.3386701
5	.1446541	.0278957	.0977516	.2088517
Tech_bought				
1	.1069182	.024506	.0672633	.165796
2	.1446541	.0278957	.0977516	.2088517
3	.163522	.0293303	.1134704	.2299253
4	.3207547	.037017	.2523949	.3977782
5	.2641509	.0349641	.201044	.3386701
Website_inf_only				
1	.1069182	.024506	.0672633	.165796
2	.1698113	.0297765	.118769	.2368927
3	.1069182	.024506	.0672633	.165796
4	.4213836	.0391594	.3465335	.5000299
5	.1949686	.0314188	.1402265	.2645064
Webiste_cookies				
1	.3333333	.0373848	.263973	.410748
2	.2389937	.0338212	.1786548	.311971
3	.2012579	.0317966	.1456518	.2713504
4	.1257862	.0262982	.0823308	.1874916
5	.1006289	.0238579	.062332	.1584787
Website_critical_func				
1	.4402516	.0393685	.3645495	.5188385
2	.2704403	.0352263	.2066853	.3453014
3	.1069182	.024506	.0672633	.165796
4	.1132075	.0251275	.0722423	.1730689
5	.0691824	.0201248	.0385471	.1210983
Website_manual				
1	.1257862	.0262982	.0823308	.1874916
2	.1949686	.0314188	.1402265	.2645064
3	.2201258	.0328586	.1620605	.2917519
4	.3081761	.0366183	.2408751	.3847506
5	.1509434	.0283907	.1029603	.2159061
Website_automated				
1	.5157233	.039633	.4376977	.5929898
2	.2515723	.0344118	.1898133	.3253561
3	.0943396	.0231809	.0574522	.1511136
4	.1006289	.0238579	.062332	.1584787
5	.0377358	.0151121	.0169455	.0819087

Website_AI				
1	.6540881	.0377227	.5763386	.7243923
2	.2327044	.0335108	.1731037	.3052506
3	.0943396	.0231809	.0574522	.1511136
4	.0125786	.0088383	.0031145	.0493771
5	.0062893	.0062695	.0008719	.0438868
Website_storage_automated				
1	.3773585	.0384413	.3049329	.4557079
2	.163522	.0293303	.1134704	.2299253
3	.163522	.0293303	.1134704	.2299253
4	.2138365	.0325161	.1565694	.2849725
5	.081761	.0217296	.0478652	.1362258
Website_data_processing				
1	.4465409	.0394253	.3705792	.5250837
2	.1949686	.0314188	.1402265	.2645064
3	.1194969	.0257244	.0772658	.1802999
4	.1886792	.0310284	.1348247	.2576395
5	.0503145	.0173356	.0252243	.0978559
Org_wout_website_yn				
1	.6289308	.0383116	.5506748	.7009588
2	.1572327	.0288687	.1082004	.2229302
3	.0566038	.0183261	.0295626	.1056861
4	.0628931	.019253	.034008	.1134311
5	.0943396	.0231809	.0574522	.1511136

Scale questions- UTAUT questions

	Proportion	Std. Err.	Logit [95% Conf.	Interval]
DS_applicable_4ops				
1	.1320755	.0268506	.0874349	.1946462
2	.1446541	.0278957	.0977516	.2088517
3	.1823899	.0306249	.1294473	.2507488
4	.3962264	.0387891	.3226857	.4747789
5	.1446541	.0278957	.0977516	.2088517
DS_beneficial_4ops				
1	.0880503	.0224725	.0526283	.1436973
2	.0943396	.0231809	.0574522	.1511136
3	.1572327	.0288687	.1082004	.2229302
4	.4654088	.0395576	.3887407	.5437469
5	.1949686	.0314188	.1402265	.2645064
DS_impact_aug				
1	.0125786	.0088383	.0031145	.0493771
2	.0314465	.0138404	.0130604	.0737815
3	.3710692	.0383116	.2990412	.4493252
4	.408805	.0389875	.3345846	.4874294
5	.1761006	.0302078	.124095	.2438336
DS_aug_eff				
1	.0062893	.0062695	.0008719	.0438868
2	.0566038	.0183261	.0295626	.1056861
3	.3081761	.0366183	.2408751	.3847506
4	.4591195	.0395198	.3826748	.5375379
5	.1698113	.0297765	.118769	.2368927

DS_investment_focus_dev				
1	.1132075	.0251275	.0722423	.1730689
2	.327044	.0372047	.2581767	.4042702
3	.3962264	.0387891	.3226857	.4747789
4	.1194969	.0257244	.0772658	.1802999
5	.0440252	.0162695	.0210109	.0899325
DS_obstacles_4ops				
1	.1194969	.0257244	.0772658	.1802999
2	.2389937	.0338212	.1786548	.311971
3	.2767296	.0354797	.2123434	.3519161
4	.2893082	.0359602	.2237088	.3650967
5	.0754717	.0209485	.0431691	.1286946
DS_not_optimal				
1	.0628931	.019253	.034008	.1134311
2	.2641509	.0349641	.201044	.3386701
3	.5031447	.0396518	.425387	.5807505
4	.1320755	.0268506	.0874349	.1946462
5	.0377358	.0151121	.0169455	.0819087
DS_effort				
1	.0062893	.0062695	.0008719	.0438868
2	.0188679	.0107901	.0060449	.0573236
3	.2012579	.0317966	.1456518	.2713504
4	.5283019	.039589	.450056	.6051814
5	.245283	.0341214	.1842249	.3186727
DS_easy_to_understand				
1	.0062893	.0062695	.0008719	.0438868
2	.2138365	.0325161	.1565694	.2849725
3	.4779874	.0396141	.4009085	.5561291
4	.2641509	.0349641	.201044	.3386701
5	.0377358	.0151121	.0169455	.0819087
DS_easy2work				
1	.0125786	.0088383	.0031145	.0493771
2	.1446541	.0278957	.0977516	.2088517
3	.6415094	.0380314	.5634799	.7127025
4	.163522	.0293303	.1134704	.2299253
5	.0377358	.0151121	.0169455	.0819087
DS_intuitive				
1	.0628931	.019253	.034008	.1134311
2	.2012579	.0317966	.1456518	.2713504
3	.5786164	.0391594	.4999701	.6534665
4	.1320755	.0268506	.0874349	.1946462
5	.0251572	.0124194	.0094024	.0655636
DS_imp_complexity				
1	.0440252	.0162695	.0210109	.0899325
2	.2264151	.0331901	.1675721	.2985111
3	.6289308	.0383116	.5506748	.7009588
4	.081761	.0217296	.0478652	.1362258
5	.0188679	.0107901	.0060449	.0573236
Peer_inf_DS_decisions				
1	.0691824	.0201248	.0385471	.1210983
2	.1698113	.0297765	.118769	.2368927
3	.5408805	.0395198	.4624621	.6173252
4	.2012579	.0317966	.1456518	.2713504
5	.0188679	.0107901	.0060449	.0573236
Peer_ref_inf_DS_decisions				
1	.0628931	.019253	.034008	.1134311
2	.1761006	.0302078	.124095	.2438336
3	.509434	.0396455	.4315364	.5868761
4	.2075472	.0321623	.1510996	.2781722
5	.0440252	.0162695	.0210109	.0899325

Peer_ref_incentivize_DS_use				
1	.0691824	.0201248	.0385471	.1210983
2	.1698113	.0297765	.118769	.2368927
3	.5031447	.0396518	.425387	.5807505
4	.2201258	.0328586	.1620605	.2917519
5	.0377358	.0151121	.0169455	.0819087
Peer_ext_incentivize_DS_use				
1	.1132075	.0251275	.0722423	.1730689
2	.2075472	.0321623	.1510996	.2781722
3	.4716981	.039589	.3948186	.549944
4	.1886792	.0310284	.1348247	.2576395
5	.0188679	.0107901	.0060449	.0573236
Funds_4DS_available				
1	.3459119	.0377227	.2756077	.4236614
2	.3647799	.038175	.2931627	.4429292
3	.2138365	.0325161	.1565694	.2849725
4	.0566038	.0183261	.0295626	.1056861
5	.0188679	.0107901	.0060449	.0573236
TechRes_4DS_av				
1	.3647799	.038175	.2931627	.4429292
2	.327044	.0372047	.2581767	.4042702
3	.2138365	.0325161	.1565694	.2849725
4	.0754717	.0209485	.0431691	.1286946
5	.0188679	.0107901	.0060449	.0573236
IntUse_DS				
1	.1006289	.0238579	.062332	.1584787
2	.1823899	.0306249	.1294473	.2507488
3	.5157233	.039633	.4376977	.5929898
4	.1698113	.0297765	.118769	.2368927
5	.0314465	.0138404	.0130604	.0737815
IntUse_DS_pred				
1	.1194969	.0257244	.0772658	.1802999
2	.1886792	.0310284	.1348247	.2576395
3	.4968553	.0396518	.4192495	.574613
4	.1761006	.0302078	.124095	.2438336
5	.0188679	.0107901	.0060449	.0573236
IntUse_DS_plan				
1	.1132075	.0251275	.0722423	.1730689
2	.2075472	.0321623	.1510996	.2781722
3	.490566	.0396455	.4131239	.5684636
4	.163522	.0293303	.1134704	.2299253
5	.0251572	.0124194	.0094024	.0655636
Interest_4DS_impl				
1	.0628931	.019253	.034008	.1134311
2	.1132075	.0251275	.0722423	.1730689
3	.3333333	.0373848	.263973	.410748
4	.3207547	.037017	.2523949	.3977782
5	.1698113	.0297765	.118769	.2368927
DS_good_idea				
1	.0125786	.0088383	.0031145	.0493771
2	.0377358	.0151121	.0169455	.0819087
3	.2893082	.0359602	.2237088	.3650967
4	.4842767	.039633	.4070102	.5623023
5	.1761006	.0302078	.124095	.2438336

Scale questions - Data science as priority

	Proportion	Std. Err.	Logit [95% Conf.	Interval]
DS_ed_access				
0	.866242	.0271662	.8029768	.9114331
1	.0955414	.0234607	.0581919	.1529695
2	.0318471	.0140138	.0132263	.0746988
4	.0063694	.0063491	.0008829	.0444362
DS_investment_priority				
1	.3566879	.0382301	.2851888	.4351988
2	.2802548	.035844	.2151421	.356132
3	.2929936	.0363238	.2266616	.3694635
4	.0573248	.0185525	.0299406	.1069927
5	.0127389	.0089502	.0031538	.0499944
DS_investment_vital_future				
1	.2356688	.0338721	.1753733	.3089284
2	.1910828	.0313771	.136583	.2607615
3	.343949	.0379111	.2733658	.4221686
4	.1910828	.0313771	.136583	.2607615
5	.0382166	.0153008	.0171612	.0829254
DS_side_effect				
1	.0955414	.0234607	.0581919	.1529695
2	.089172	.0227448	.0533051	.1454639
3	.4713376	.0398387	.3939866	.5500899
4	.3184713	.0371815	.2498927	.3959361
5	.0254777	.0125755	.0095217	.0663801
DS_actively_pursued				
1	.2675159	.0353284	.2036897	.342734
2	.2993631	.0365507	.2324458	.376105
3	.3184713	.0371815	.2498927	.3959361
4	.1019108	.0241446	.0631353	.1604231
5	.0127389	.0089502	.0031538	.0499944

Scale questions - Data Routines (e.g. collection, usage)

	Proportion	Std. Err.	Logit [95% Conf.	Interval]
Data_collection_digital				
1	.0880503	.0224725	.0526283	.1436973
2	.1383648	.0273827	.0925759	.2017656
3	.1006289	.0238579	.062332	.1584787
4	.5220126	.0396141	.4438709	.5990915
5	.1509434	.0283907	.1029603	.2159061
Data_collection_physical				
1	.0188679	.0107901	.0060449	.0573236
2	.0754717	.0209485	.0431691	.1286946
3	.081761	.0217296	.0478652	.1362258
4	.6100629	.03868	.5315653	.6832448
5	.2138365	.0325161	.1565694	.2849725
DigData_quality				
1	.0062893	.0062695	.0008719	.0438868
2	.0628931	.019253	.034008	.1134311
3	.1194969	.0257244	.0772658	.1802999
4	.5597484	.0393685	.4811615	.6354505
5	.2515723	.0344118	.1898133	.3253561
PhysData_quality				
1	.0188679	.0107901	.0060449	.0573236
2	.0440252	.0162695	.0210109	.0899325
3	.1006289	.0238579	.062332	.1584787
4	.5974843	.0388916	.5188896	.6713712
5	.2389937	.0338212	.1786548	.311971

DigData_consistent				
1	.0062893	.0062695	.0008719	.0438868
2	.1257862	.0262982	.0823308	.1874916
3	.2389937	.0338212	.1786548	.311971
4	.490566	.0396455	.4131239	.5684636
5	.1383648	.0273827	.0925759	.2017656
PhysData_consist				
1	.0251572	.0124194	.0094024	.0655636
2	.1006289	.0238579	.062332	.1584787
3	.2389937	.0338212	.1786548	.311971
4	.4842767	.039633	.4070102	.5623023
5	.1509434	.0283907	.1029603	.2159061
DigData_needs_understanding				
1	.0377358	.0151121	.0169455	.0819087
2	.081761	.0217296	.0478652	.1362258
3	.2075472	.0321623	.1510996	.2781722
4	.5157233	.039633	.4376977	.5929898
5	.1572327	.0288687	.1082004	.2229302
DigData__processing_skills				
1	.0377358	.0151121	.0169455	.0819087
2	.1320755	.0268506	.0874349	.1946462
3	.2830189	.0357242	.218018	.3585144
4	.427673	.0392355	.3525266	.5063117
5	.1194969	.0257244	.0772658	.1802999
DDriven_actions_knowledge				
1	.0440252	.0162695	.0210109	.0899325
2	.1320755	.0268506	.0874349	.1946462
3	.2955975	.0361878	.2294154	.3716634
4	.427673	.0392355	.3525266	.5063117
5	.1006289	.0238579	.062332	.1584787
DDDM_routine				
1	.0440252	.0162695	.0210109	.0899325
2	.1194969	.0257244	.0772658	.1802999
3	.245283	.0341214	.1842249	.3186727
4	.490566	.0396455	.4131239	.5684636
5	.1006289	.0238579	.062332	.1584787
DDDM_as_objective				
1	.0377358	.0151121	.0169455	.0819087
2	.1069182	.024506	.0672633	.165796
3	.245283	.0341214	.1842249	.3186727
4	.490566	.0396455	.4131239	.5684636
5	.1194969	.0257244	.0772658	.1802999
DM_intuitive_routine				
1	.1446541	.0278957	.0977516	.2088517
2	.2767296	.0354797	.2123434	.3519161
3	.2955975	.0361878	.2294154	.3716634
4	.2264151	.0331901	.1675721	.2985111
5	.0566038	.0183261	.0295626	.1056861
DM_intuitive_objective				
1	.2327044	.0335108	.1731037	.3052506
2	.3459119	.0377227	.2756077	.4236614
3	.2578616	.0346926	.1954199	.3320218
4	.1257862	.0262982	.0823308	.1874916
5	.0377358	.0151121	.0169455	.0819087
Data_sensitive				
1	.0314465	.0138404	.0130604	.0737815
2	.1509434	.0283907	.1029603	.2159061
3	.1823899	.0306249	.1294473	.2507488
4	.3522013	.0378806	.2814458	.4300975
5	.2830189	.0357242	.218018	.3585144

Data_confidential				
1	.0125786	.0088383	.0031145	.0493771
2	.1194969	.0257244	.0772658	.1802999
3	.1572327	.0288687	.1082004	.2229302
4	.3710692	.0383116	.2990412	.4493252
5	.3396226	.0375574	.2697833	.4172116
Data_open_access_internal				
1	.427673	.0392355	.3525266	.5063117
2	.3962264	.0387891	.3226857	.4747789
3	.081761	.0217296	.0478652	.1362258
4	.081761	.0217296	.0478652	.1362258
5	.0125786	.0088383	.0031145	.0493771
Data_collected_with_consensus				
1	.0251572	.0124194	.0094024	.0655636
2	.0566038	.0183261	.0295626	.1056861
3	.1194969	.0257244	.0772658	.1802999
4	.4465409	.0394253	.3705792	.5250837
5	.3522013	.0378806	.2814458	.4300975
DataCollected_wout_resistance				
1	.0251572	.0124194	.0094024	.0655636
2	.1069182	.024506	.0672633	.165796
3	.1823899	.0306249	.1294473	.2507488
4	.4654088	.0395576	.3887407	.5437469
5	.2201258	.0328586	.1620605	.2917519
Users_questioning_data_coll				
1	.2264151	.0331901	.1675721	.2985111
2	.3522013	.0378806	.2814458	.4300975
3	.2704403	.0352263	.2066853	.3453014
4	.1446541	.0278957	.0977516	.2088517
5	.0062893	.0062695	.0008719	.0438868

Appendix 16: Summarized results of multiple regression models generated from Stata – block residual

Block	F	Block df	Residual df	Pr > F	R2	Change in R2
1	1.55	20	136	0.0751	0.1220	
2	7.21	7	129	0.0000	0.3429	0.2209
3	3.63	2	127	0.0294	0.3825	0.0396
4	7.05	4	123	0.0000	0.5193	0.1368

Appendix 17: Linear regressions with blocks addition from multiple regression model generated from Stata

Block 1

Linear regression

Number of obs	=	157
F(19, 136)	=	.
Prob > F	=	.
R-squared	=	0.1220
Root MSE	=	.91826

scale_DS_usag~n	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_IAge_resp_2	-.2267189	.3181565	-0.71	0.477	-.8558928	.4024549
_IAge_resp_3	-.0917896	.2908707	-0.32	0.753	-.6670041	.4834249
_IAge_resp_4	.3843199	.2821648	1.36	0.175	-.1736781	.9423179
_IAge_resp_5	.1076384	.398227	0.27	0.787	-.6798797	.8951566
_IAge_resp_6	-.5313381	.4277695	-1.24	0.216	-1.377278	.3146022
_IAge_org_m_2	-.0120534	.39214	-0.03	0.976	-.7875341	.7634273
_IAge_org_m_3	-.264097	.35499	-0.74	0.458	-.9661113	.4379174
_IAge_org_m_4	-.1147271	.3628843	-0.32	0.752	-.8323527	.6028986
_IAge_org_m_5	.1158754	.4935635	0.23	0.815	-.8601765	1.091927
_IAge_org_m_6	-.5205908	.4451544	-1.17	0.244	-1.400911	.3597291
_IAge_board_3	.0886002	.3024529	0.29	0.770	-.5095189	.6867193
_IAge_board_4	-.1016752	.3203703	-0.32	0.751	-.735227	.5318766
_IAge_board_5	-.1461914	.3271285	-0.45	0.656	-.7931079	.500725
_IAge_board_6	-.3169281	.4012414	-0.79	0.431	-1.110407	.4765512
_IOrg_size_2	.0974682	.1575224	0.62	0.537	-.2140419	.4089784
_IOrg_size_3	.0850143	.3839402	0.22	0.825	-.6742507	.8442794
_IOrg_size_4	.0791122	.2722802	0.29	0.772	-.4593385	.6175629
_IOrg_size_5	.3099234	.4143679	0.75	0.456	-.5095143	1.129361
_IOrg_size_6	-.4732929	.4873737	-0.97	0.333	-1.437104	.4905181
Demand_matching	-.1322891	.1588751	-0.83	0.406	-.4464743	.1818962
_cons	3.038804	.4160994	7.30	0.000	2.215942	3.861666

Block 2 added

Linear regression			
	Number of obs	=	157
	F(26, 129)	=	.
	Prob > F	=	.
	R-squared	=	0.3429
	Root MSE	=	.81567

scale_DS_usage_intention	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_IAge_resp_2	-.692729	.3159505	-2.19	0.030	-1.317845	-.0676132
_IAge_resp_3	-.6354855	.2974468	-2.14	0.035	-1.223991	-.0469798
_IAge_resp_4	-.2121189	.300282	-0.71	0.481	-.8062342	.3819963
_IAge_resp_5	-.6427252	.3719598	-1.73	0.086	-1.378657	.0932063
_IAge_resp_6	-.9912848	.3838677	-2.58	0.011	-1.750776	-.2317932
_IAge_org_m_2	-.0567767	.3533569	-0.16	0.873	-.755902	.6423486
_IAge_org_m_3	-.1706376	.3154619	-0.54	0.589	-.7947866	.4535114
_IAge_org_m_4	-.1018906	.3181433	-0.32	0.749	-.731345	.5275637
_IAge_org_m_5	.1942983	.4553894	0.43	0.670	-.7067007	1.095297
_IAge_org_m_6	-.7515437	.4398764	-1.71	0.090	-1.62185	.1187626
_IAge_board_3	.0069354	.2980666	0.02	0.981	-.5827966	.5966674
_IAge_board_4	-.1062124	.3087222	-0.34	0.731	-.7170268	.504602
_IAge_board_5	-.2047904	.321074	-0.64	0.525	-.8400433	.4304624
_IAge_board_6	-.2231929	.4113022	-0.54	0.588	-1.036964	.5905786
_IOrg_size_2	.0410451	.1473349	0.28	0.781	-.2504607	.3325508
_IOrg_size_3	.0223468	.2934245	0.08	0.939	-.5582006	.6028943
_IOrg_size_4	-.2674546	.3197917	-0.84	0.405	-.9001704	.3652612
_IOrg_size_5	-.060353	.3907855	-0.15	0.878	-.8335316	.7128257
_IOrg_size_6	-.6760466	.3920214	-1.72	0.087	-1.451671	.0995773
Demand_matching	-.1620954	.1401522	-1.16	0.250	-.4393899	.1151992
DS_ed_access	.2570047	.1091319	2.35	0.020	.0410844	.4729249
Funds_4DS_available	.0824467	.0859494	0.96	0.339	-.0876063	.2524996
TechRes_4DS_av	.2863919	.0840529	3.41	0.001	.1200912	.4526927
Data_collection_digital	-.0606304	.0719321	-0.84	0.401	-.2029497	.0816889
DPO_internally	-.1487086	.1530917	-0.97	0.333	-.4516043	.1541871
GDPR_regulation_knowledge	.4762733	.3916579	1.22	0.226	-.2986315	1.251178
W2Collaborate_binary	.3138212	.1452379	2.16	0.033	.0264645	.6011779
_cons	2.530754	.5474306	4.62	0.000	1.447649	3.613859

Block 3 added

Linear regression			
	Number of obs	=	157
	F(28, 127)	=	.
	Prob > F	=	.
	R-squared	=	0.3825
	Root MSE	=	.79689

scale_DS_usage_intention	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_IAge_resp_2	-.9682118	.4211907	-2.30	0.023	-1.801672	-.1347514
_IAge_resp_3	-.7776951	.4018448	-1.94	0.055	-1.572873	.0174831
_IAge_resp_4	-.4874405	.4103187	-1.19	0.237	-1.299387	.3245061
_IAge_resp_5	-.8658433	.4691148	-1.85	0.067	-1.794137	.0624502
_IAge_resp_6	-1.220189	.4600539	-2.65	0.009	-2.130553	-.3098256
_IAge_org_m_2	.0789233	.297616	0.27	0.791	-.510005	.6678516
_IAge_org_m_3	-.0176492	.2712555	-0.07	0.948	-.554415	.5191166
_IAge_org_m_4	.0130441	.2743649	0.05	0.962	-.5298745	.5559627
_IAge_org_m_5	.3089828	.3901254	0.79	0.430	-.4630049	1.080971
_IAge_org_m_6	-.5433317	.3902734	-1.39	0.166	-1.315612	.228949
_IAge_board_3	-.0173749	.3045907	-0.06	0.955	-.6201049	.585355
_IAge_board_4	-.0954422	.3092493	-0.31	0.758	-.7073908	.5165065
_IAge_board_5	-.2374104	.3286522	-0.72	0.471	-.8877539	.4129331
_IAge_board_6	-.2915226	.4089034	-0.71	0.477	-1.100669	.5176235
_IOrg_size_2	.0110419	.1481756	0.07	0.941	-.2821709	.3042547
_IOrg_size_3	.0431917	.2851981	0.15	0.880	-.5211639	.6075473
_IOrg_size_4	-.151213	.3320353	-0.46	0.650	-.8082511	.505825
_IOrg_size_5	.0374105	.3462049	0.11	0.914	-.6476665	.7224875
_IOrg_size_6	-.493368	.3557584	-1.39	0.168	-1.19735	.2106136
Demand_matching	-.1279912	.135107	-0.95	0.345	-.3953436	.1393613
DS_ed_access	.2101429	.1052547	2.00	0.048	.0018629	.4184229
Funds_4DS_available	.0811765	.0855903	0.95	0.345	-.0881913	.2505443
TechRes_4DS_av	.3171574	.0849135	3.74	0.000	.1491289	.485186
Data_collection_digital	-.0820924	.0707189	-1.16	0.248	-.2220324	.0578475
DPO_internally	-.1284104	.1540651	-0.83	0.406	-.4332774	.1764567
GDPR_regulation_knowledge	.3087637	.3371445	0.92	0.361	-.3583844	.9759118
W2Collaborate_binary	.3205336	.142616	2.25	0.026	.0383222	.6027449
DDDM_routine	-.2138652	.1315992	-1.63	0.107	-.4742762	.0465458
DDDM_as_objective	.3333405	.1289016	2.59	0.011	.0782675	.5884134
_cons	2.350416	.5653663	4.16	0.000	1.231658	3.469174

Block 4 added

Linear regression			
	Number of obs	=	157
	F(32, 123)	=	.
	Prob > F	=	.
	R-squared	=	0.5193
	Root MSE	=	.71445

scale_DS_usage_intention	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
_IAge_resp_2	-.3742809	.3265717	-1.15	0.254	-1.02071	.2721477
_IAge_resp_3	-.1572703	.3091938	-0.51	0.612	-.7693003	.4547598
_IAge_resp_4	.0731539	.2930564	0.25	0.803	-.5069332	.6532409
_IAge_resp_5	-.1244037	.3957696	-0.31	0.754	-.9078053	.6589979
_IAge_resp_6	-.4243139	.438628	-0.97	0.335	-1.292551	.4439234
_IAge_org_m_2	-.1255436	.2531347	-0.50	0.621	-.6266081	.375521
_IAge_org_m_3	-.1386672	.218319	-0.64	0.527	-.5708163	.2934819
_IAge_org_m_4	-.0547298	.2245823	-0.24	0.808	-.4992767	.389817
_IAge_org_m_5	.1927974	.3404921	0.57	0.572	-.4811859	.8667806
_IAge_org_m_6	-.6532316	.3782513	-1.73	0.087	-1.401957	.0954936
_IAge_board_3	-.0807711	.3111235	-0.26	0.796	-.696621	.5350788
_IAge_board_4	-.0473816	.3188199	-0.15	0.882	-.6784661	.5837029
_IAge_board_5	-.2369805	.3382573	-0.70	0.485	-.9065401	.4325791
_IAge_board_6	-.5814089	.4121408	-1.41	0.161	-1.397216	.2343986
_IOrg_size_2	-.0254729	.1389029	-0.18	0.855	-.3004227	.2494769
_IOrg_size_3	.0988624	.3134131	0.32	0.753	-.5215197	.7192444
_IOrg_size_4	-.0050199	.3146786	-0.02	0.987	-.6279068	.617867
_IOrg_size_5	.0423957	.4110416	0.10	0.918	-.771236	.8560274
_IOrg_size_6	-.1122853	.322417	-0.35	0.728	-.75049	.5259195
Demand_matching	-.0219559	.1272311	-0.17	0.863	-.2738021	.2298903
DS_ed_access	.0337798	.0940824	0.36	0.720	-.1524505	.2200101
Funds_4DS_available	.0650002	.0908952	0.72	0.476	-.1149212	.2449215
TechRes_4DS_av	.228268	.0937821	2.43	0.016	.042632	.4139039
Data_collection_digital	-.1029313	.0628455	-1.64	0.104	-.2273301	.0214675
DPO_internally	-.1065949	.1327921	-0.80	0.424	-.3694487	.156259
GDPR_regulation_knowledge	-.1351447	.2064754	-0.65	0.514	-.54385	.2735607
W2Collaborate_binary	.1422768	.1364084	1.04	0.299	-.1277352	.4122889
DDDM_routine	-.144481	.1088057	-1.33	0.187	-.3598551	.0708932
DDDM_as_objective	.2826593	.1292255	2.19	0.031	.0268654	.5384532
DSuse_relative2peers	.0929704	.0773713	1.20	0.232	-.0601813	.2461221
scale_DS_augm	.3167403	.0990399	3.20	0.002	.1206969	.5127836
scale_DS_social_influence	.2607332	.0952441	2.74	0.007	.0722034	.449263
scale_DS_easy	-.0179805	.1273083	-0.14	0.888	-.2699794	.2340184
_cons	.5275286	.672008	0.79	0.434	-.80267	1.857727