Supplier Development Programme - Meet the Buyer Data Visualisation

Author:
Cameron Scott (H00337889)

Supervisors:
Dr. Nick Taylor – Heriot-Watt University (Academic Supervisor)
Jennifer Payne – Supplier Development Programme (Industrial Supervisor)

August 2021

Computer Science
School of Mathematical and Computer Sciences

Dissertation submitted as part of the requirements for the award of the degree of MSc. in Data Science (2 years)

DECLARATION

I, ...CAMERON SCOTT.....,
confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included.

***Signed**: .....Cameron Scott.....*

***Date**: .....16/08/2021.....*

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank those who have supported me throughout my time at Heriot-Watt University, Supplier Development Programme (SDP) and beyond. Firstly, I would like to thank Dr. Nick Taylor, my academic supervisor, whose support and knowledge has been invaluable for this project. I would also like to thank the whole team at SDP and particularly my industrial supervisor at SDP, Jennifer Payne, whose support and knowledge has given both myself and this project a great insight into applying my data science skills in a commercial environment. A thank you must be paid to The Data Lab who without them, this project would not have happened. Finally, I owe a thank you to my family and friends for all their support during my time at Heriot-Watt University.

ABSTRACT

This is a Data Lab internship project with a 10-week industrial placement in the Supplier Development Programme (SDP) based at South Lanarkshire Council.

This project has aimed to meet the needs of SDP by using a variety of data analysis and visualisation techniques. SDP are a partnership of local authorities, government and other public organisations who provide free support in the tendering process to Scottish-based businesses. The core of this project was to produce innovative and interactive data visualisation models based on datasets from SDP's events 'Meet the Buyer' (MTB) and 'Meet the Buyer North' (MTBN). The visualisation models aimed to ensure SDP's data was made more visible to their membership organisations, such as the Meet the Buyer Event Outcomes data, as well as raising the profile of SDP within procurement both in the public and private sector. These datasets also provided the basis for further data analysis as a review was undertaken regarding how this data has been previously collected, managed, and stored with recommendations formulated for implementation in the future. Another deliverable of this project was that by working closely with these datasets, data analysis methods and digital technologies would be utilised to identify any trends and patterns in the datasets. Overall, the project aim was to bring new data analytical thinking to SDP and provide them with sustainable improvements which would bolster productivity both internally and externally.

Table of Contents

DECLARATION i
ACKNOWLEDGEMENTS ii
ABSTRACT iii
List of Tables and Figures v
Introduction: 1
AIMS: 1
OBJECTIVES: 1
Relevance of Project 2
Literature Review 8
"Data is the new oil." 8
Drilling and Extracting Data 9
'Refining' the Data 12
The Knowledge-Based Economy 14
Knowledge Management 16
Digital & Information Technologies 19
Data Visualisation: A picture that is worth more than a thousand words 20
Data Visualisation: The good, the Bad, and the Innovative 21
Business Intelligence 27
Methodology & Implementation 29
Introduction 29
The Datasets – Meet the Buyer (MTB) & Meet the Buyer North (MTBN) 29
Reviewing Data Collection – MTB Survey Format 31
Processing and Analysing the Datasets – MTB (2017-2021), SDP & ONS data 34
Visualising the Datasets 36
The Data Model 37
The Data Dashboard 37
Power BI – Visualisation Tool 49
Discussion of Professional, Ethical, Legal, and Social Issues 51
Discussion and Evaluation 53
Introduction 53
Data Collection – New MTB Data 53
The Data Dashboard – User Experience 54

Future Work 57
Conclusion 58
Bibliography 60
APPENDICES 68
APPENDIX A: Project Requirements Analysis (MoSCoW) & Project Stakeholders 68
APPENDIX B: Risk Assessment 70
APPENDIX C: Reviewed MTB Post-Event Feedback Survey – Additional Questions 71
APPENDIX D: MTB Data Visualisation – Data Model 74
APPENDIX E: Data Dashboard User Experience Questionnaire 75

List of Tables and Figures

Figure 1.0. – Question Conceptual Framework

Figure 1.1. – Knowledge Discovery in Databases (KDD) Process

Figure 1.2. – Bad data visual #1

Figure 1.3. – Bad data visual #2

Figure 1.4. – Bad data visual #3

Figure 1.5. – Bad data visual #4

Figure 1.6. – Technology Initiatives

Figure 1.7. – MTB data visualisation (Slicer Panel)

Figure 1.8. – MTB data visualisation (Scotland Map)

Figure 1.9. – MTB data visualisation (Stacked Bar chart)

Figure 2.0. – MTB data visualisation (Text Box)

Figure 2.1. – MTB data visualisation (Tree Map)

Figure 2.2. – MTB data visualisation (Area Chart)

Figure 2.3. – MTB data visualisation (Card visuals)

Figure 2.4. – MTB data visualisation (Gauge visuals)

Figure 2.5. – MTB data visualisation (Scaffolding)

Figure 2.6. – MTB data visualisation (Info Panel)

Figure 2.7. – MTB data visualisation (Intro Page)

Figure 2.8. – MTB data visualisation (COVID-19)

Introduction:

AIMS:

- To use various data analysis techniques and methodologies to bring new analytical thinking to SDP which will help them implement sustainable and innovative improvements for future data collection, management, and analysis
- To realise business value for SDP both for internal use and external stakeholders through data analysis and visualisation of the datasets 'Meet the Buyer' (MTB) and 'Meet the Buyer North' (MTBN)

OBJECTIVES:

- Undertake a review of SDP's data collection methods (feedback surveys) and recommend any changes to be implemented to better assess the effectiveness of SDP's future activities in the procurement process
- Formulate these recommendations into an implementation plan consisting of revised feedback surveys to improve collation of consistent feedback data and a data analysis framework that allows SDP to manage, visualise and evaluate data for future activities in a sustainable manner.
- Conduct data pre-processing techniques (cleaning, transforming and reduction) on SDP's raw 'Meet the Buyer' (MTB) delegate feedback datasets
- Use data mining and machine learning algorithms on those pre-processed datasets to identify any valuable trends and patterns within the data
- Analyse the trends and patterns identified in the datasets to assess the effectiveness of SDP's existing activities within procurement both in public and private sectors and to gain insight into any market gaps and new opportunities for SDP in the procurement process as well
- Develop data visualisations to make SDP's data more visible and interactive for their membership organisations

- Embed those data visualisations into SDP's website so they can be used for communications marketing strategies and help raise the profile of SDP and procurement within both the public and private sector

Relevance of Project

In today's economy, most organisations are continuing to collect an ever-expanding amount of data and are pursuing innovative ways of obtaining and implementing value from it (International Data Group (IDG), 2016). Irrespective of the organisation size, data has now become one of the most important assets for an organisation to have in the present business world (Nair, 2019). This falls in line with the national perspective on data with the UK government releasing their 'National Data Strategy' as recently as last year outlining 'The Pillars' and 'The Missions' on how enterprise can unlock the value of data across all sectors of the economy (Department for Digital, Culture, Media & Sport, 2020). This is the position that SDP find their organisation in where data has been collected but the potential value in this asset is still to be achieved. It is also where this project can contribute by providing those innovative ways of analysing and visualising the data to derive elements of value and knowledge for SDP to implement. SDP have been collecting and managing large amounts of data from their various training and events (MTB, MTBN) they offer for small to medium sized businesses (SMEs), as well as supported businesses and other suppliers and exhibitors. In terms of the relevant data analysis that SDP have conducted however, it can be said that the potential value and knowledge residing in their datasets has not been realised thus necessitating the need for this project.

This project will also provide an insight into an initiative that most organisations see as necessary and beneficial for them, however, feel that they lack the resources to initiate and conduct it successfully. Surveys have shown that across all company sizes, more than three quarters (78%) of them feel that the collection and analysis of data has the power to be transformative for their future operations (International Data Group (IDG), 2016). In addition to this, 71% say that a data analysis initiative would create new revenue and business opportunities but only 58% believe that

their organisation would be capable of realising the potential value of a data analysis initiative (International Data Group (IDG), 2016). And in contrast to the beliefs organisations have about a data analysis initiative, the significant majority (90%) have experienced problems when trying to implement a data analytics initiative, whether it be something as simple as data storage or an area as complex as data access and analysis (International Data Group (IDG), 2016). With the expectation that the volume of data an organisation holds to increase by over 50% (International Data Group (IDG), 2016), and that in 2019, 90% of all data in the world was generated in the previous two years (Hariri, Fredericks and Bowers, 2019), this project can raise awareness for the value a data analysis initiative can have for all types of organisations and their ever-expanding stores of data.

It is clear to see from the sources considered so far that organisations are collecting data and doing so in great amounts, but it is one task to collect data, it is another to assess that the data being collected is of good quality and subsequently provide the insights (hindsights or foresights if these are preferred) that the organisation requires. In terms of SDP, utilising their collection methods to gain quality data is crucial. By using post-event feedback surveys to collect data, the vital criteria SDP aim to satisfy is that they are asking those questions – about their Meet the Buyer events – which allow them to capture the most useful insights that the organisation wants to gain. Reviewing SDP's collection and management techniques of data such as their surveys will allow this project to make recommendations to those techniques by implementing any recommendations to the existing survey format which will attempt to focus on those or other key insights and gain quality data in the process.

To shift the focus towards SDP as an organisation and the environment they themselves are conducting this data-driven initiative in. SDP are a partnership of Local Authorities, Scottish government, and other public sector bodies and are a micro team (less than 10 employees). They are also publicly funded and with such a large remit for their services, they have also found their resources diminished to a certain extent due to the austerity that has affected such wide swathes of the public

sector over previous years. The team at SDP are, however, not an isolated case when it comes to the combination of austerity-depleted resources and a data-driven project being employed. Evidence has demonstrated that organisations in the public sector have turned to data innovation to solve the issues created by the cuts to public funding, with 54% of those surveyed by Big Data LDN having turned to data and data analytics to alleviate the challenges posed by austerity in the public sector (2019, pp. 8). And the innovative methods are not just a means of overcoming spending cuts, a further 91% of respondents said that those data-driven initiatives were benefitting the departments who had instigated them (2019, pp. 6). It goes to show, not only can SDP use data as a means to conquer the impacts of austerity but also reap the benefits of such a project in the process.

Public sector organisations implementing data projects after a period of austerity such as SDP is part of a wider sense for data-driven value. The public sector is now viewing "data analytics as this catchphrase for every organisation" (Meaning and significance of data analytics for the public sector - Resources, 2020), and with the connotations of the word 'data' often being related to progress and a sort of latent intelligence where the "difficult work" can be done through data analytics without much interference from the end user, it is now seen as an accessible option for value to be gained from data-driven information at an organisational level (Beer, 2017). Translating this sense for data-driven value into facts and figures, research has shown that both tangible and intangible benefits are being appreciated from data-initiatives in the public sector (59% agreeing that this is the case). A further 30% are now understanding and visualising complex datasets and 27% are also helping their staff efficacy by using data analytics (Big Data LDN & Qlik, 2019, pp. 10). Data analytics therefore goes beyond being just a buzzword for progress with benefits ranging from visualisations of complex data to aiding decision-making in public sector departments. These types of benefits are pertinent to SDP's project and their aim to develop innovative data visualisations which could also have similar impacts for supporting understanding of and accessibility to complex data.

Research has demonstrated that data-driven projects are being invested within organisations in the public sector however, there are some who are still in the early stages and are yet to move beyond Excel where data is only understood through the format of the most widely used data analytics tool. The spreadsheet (Big Data LDN & Qlik, 2019). Nearly a third (31%) of public sector organisations are still using their raw data in spreadsheets and since Excel struggles to match the storytelling capability of other formats, it can be said that most public sector bodies are still in the infancy of their data-driven journey (Big Data LDN & Qlik, 2019, pp. 7).

This is reminiscent of SDP's data exploits where their data is perhaps not being realised to its full potential. Surveys disseminated – as part of this project's research – to SDP representatives across various Local Authorities and to other public sector organisations found that 75% felt that public sector data, such as MTB data, has been underused in terms of making it more visible to end users by showing the outcomes of MTB events. Further responses illustrated that some public sector bodies have encountered issues when incorporating a data-driven initiative citing problems with data sharing and collaboration between organisations within the public sector. One comment from an SDP representative outlined how SDP do attempt to promote the MTB data and its outcomes however data analysis, such as data visualisation, will take their data to a new level. This is reinforced by other public sector bodies who have already initiated and completed a data-driven project, responses from Scottish Futures Trust (SFT) showed that they feel the data visualisations they have implemented will benefit the efficacy of their own or other stakeholders' activities. SFT also agreed that after conducting this project, they now have the capabilities to conduct future data analysis projects more effectively. These responses underline just how valuable data analysis and visualisation can be for SDP and their MTB events and that to date, their data, and the potential benefits it could have, are lying untouched.

Now this paper could not ignore what has been the most extraordinary event to affect the planet since the world was at war (United Nations, 2021). The global pandemic has sent shockwaves throughout the current economic climate and SDP

are in no way impervious to the serious impacts it has had on organisations, discriminant of size or sector. For example, SDP's national annual in-person Meet the Buyer event was cancelled midway through 2020 with the loss of feedback data among other adverse impacts and all these events now being held virtually since then. But it is perhaps SDP's main consumer and beneficiary of their services who have been dealt the most severe blow, that being the Scottish SME. SMEs in Scotland have free and unique access to the events and activities SDP provide in the wider procurement network such as the Meet the Buyer events where businesses can attend and engage with the public sector to create new revenue and business opportunities for themselves (Why register with SDP?, 2021).

With the advent of the pandemic however, the impact on SMEs has been drastic. Scottish small businesses' confidence has dropped to a record new low as seen by The Federation of Small Businesses (FSB) – dropping by over 127 points to -154.8 in 2020 – with the overwhelmingly majority (85.7%) citing Coronavirus as being the direct cause for this drop (Federation of Small Businesses, 2021, pp. 5). Additionally, and unsurprisingly, two thirds have recorded falling turnover in 2020, nearly a quarter are thinking of making staff redundant, and in terms of the national context, more businesses in Scotland have closed than compared to the rest of the UK as a whole (54% compared to 41% respectively) (Federation of Small Businesses, 2021, pp. 4). Considering the evidence as well as other literature which shows how the businesses who are most likely to suffer from shock events are those that are small and medium in stature (Brown, Rocha and Cowling, 2020; Howell, Lerner, Nanda and Townsend, 2020; Fraser of Allander Institute (FAI), 2020). It only reinforces how SMEs, specifically those accessing SDP's services, need to be supported throughout these events and the first step in doing so is to understand how they have been affected. This understanding is where a data-driven initiative can help, data analysis can provide the framework to first collect response data from SDP's main consumers, analyse those impacted and gain insights into how Scottish SMEs are coping with the pandemic and if interaction with SDP's services are offering routes for recovery from the pandemic.

The relevance for a data-driven project that will bring new data analytical thinking to SDP has been highlighted thus far. The remainder of this paper will now concentrate on the broader field of literature around data as an asset and the processes involved in garnering that value from it in today's economy, or what is sometimes referred to as the Knowledge-Based Economy and all its driving forces and technologies. The paper will then describe the methodology implemented in this project to meet its objectives through data analysis and visualisation as well as the data science skills and software utilised in the process. The value gained for SDP from this data-driven project will also be discussed and evaluated and any recommendations also being considered and formulated for SDP's future work with data building upon and moving forward from this project.

Literature Review

"Data is the new oil."

There have been major social, economic, and political changes in the 21st century, and one of these has been the shift or what is referred to as the soft continuity from an industrial to a post-industrial economy and their main forms of capital (Brint, 2001; Dean and Kretschmer, 2007; Brinkley, 2006; Hadad, 2017). In the industrial economy, there were the traditional forms of capital such as land and labour, however in the post-industrial economy, these forms of capital although still prevalent but the resources – oil for example – that fuel this can be seen to be exhaustive, finite, and unsustainable. The resource that fuels this transitioning economy is one that can be seen to be sustainable and infinite in its generation (Is Data Really the New Oil in the 21st Century?, 2020). And that resource is data. Data has now been considered as the new oil due to it being the resource that is powering this changing economy. Hirsch (2014) summed up this analogy of the resourcefulness and value of data by saying "data is an essential resource that powers the information economy much like oil has fuelled the industrial economy" (pp. 374).

It was Clive Humby who coined the term "data is the new oil" (2006) and since then, various others who contribute hugely to the post-industrial economy have reinforced this, such as Gartner, IBM, and Microsoft to name a few (Quote: Sondergaard on Data Analytics, 2011; "Data is the New Oil" — A Ludicrous Proposition, 2016). Some of the comparisons between oil and data relate to the value they both hold and the vast quantities of them. Data is now seen as the "new raw material for business" (Data, data everywhere, 2010) where organisations see accumulating data is alike with accumulating capital and in every sector of the economy, more and more data is being collected by organisations. It is estimated that globally, 2.5 quintillion bytes of data is being generated each day, in metaphorical comparison, the U.S. alone consume over 20 million barrels of oil every

day (Is Data Really the New Oil in the 21st Century?, 2020). Internet users have also grown to 3.7 billion people since 2016 and this all contributes to the quantities of data on offer (Hariri, Fredericks and Bowers, 2019). The value from the vast quantities of data now being discovered means that a market has been driven to harness this resource, again, this is comparable with how oil is viewed as a valuable resource. To take the EU economy as one example, the market for data was expected to grow to over 700 billion euros by last year and account for 4% of overall GDP showcasing the oil-like richness data can hold (Nolin, 2019).

Another comparable aspect of both data and oil is the rawness of both resources. The true value is not just data as an entity itself but lies in how you manage and extract the value from it. Oil is much the same, as a raw natural resource, there is no value in that crude state whereas once it is mined, drilled, and extracted it can then prove to be a valuable resource. Now this involves wells, rigs and a contraption called a "Christmas Tree" (Shahul, 2018), but this is beyond the scope of this report. In terms of data as a valuable resource, it has to be exploited through digital technologies and effective methods of data collection, processing, and analysis for the value to be realised (Is Data Really "The New Oil"?, 2020). It is once the dirty data has been refined through these methods that it can finally have the desired impact for an organisation whether that be for prediction, profitability, or any other economic activity. Now the first step in gaining value from data is the collection of such a resource or in terms of the oil-data comparison narrative – to drill for the data.

Drilling and Extracting Data

Data collection and management techniques such as surveys are key to collecting data in the first place, but the quality of data is reliant upon on the survey design for collecting good insightful feedback. This is where aspects of social science can be seen to link in with the computer science and analytics side of a data-driven initiative as it is a person producing the data and not a digitised entity. When evaluating their events and activities, SDP can be seen to operating as a researcher as they want to

study the audiences of their Meet the Buyer events and gather feedback that reflects on their aims of this particular study, the population under investigation and the resources available (McColl et al., 2001). Essentially SDP want to understand how their activities within procurement both in the public and private sectors impact their consumers. All things considered, SDP's Meet the Buyer (MTB) events aim to provide new business opportunities for revenue and income to SMEs. They have to gather feedback through data collection techniques such as surveys which will then provide them with the insights on what can be improved or needs changed in their events so suppliers and buyers are both benefitting from MTB events meaning this type of event is having its desired impact on the wider procurement network – the National Procurement Framework for example (Public sector procurement - gov.scot, 2021).

The first hurdle to gaining insightful data through data collection and management techniques is good questionnaire design. The amounts of literature on good questionnaire format and design is vast so a synopsis of some of the key indicators of good survey design and the main culprits for poor surveys will be provided. Validity is one of the characteristics of a good survey, so the questions ask what was intended by the researcher and there is no ambiguity on the respondent side about the question. It also has to be reliable where responses to the same type of question yield the same answer over time, key for analysing time-series data. Survey design should also incorporate interesting and succinct questions to avoid response fatigue. Another key stage in designing a strong survey format and questions is to allow for the development of a 'Conceptual Framework' where the independent and dependent variables are identified and researched to build an idea of what needs to be included in the question (Jenn, 2006). For example, if the research question required was to gather feedback on the overall importance of Meet the Buyer events, a framework would have to be developed as detailed below:

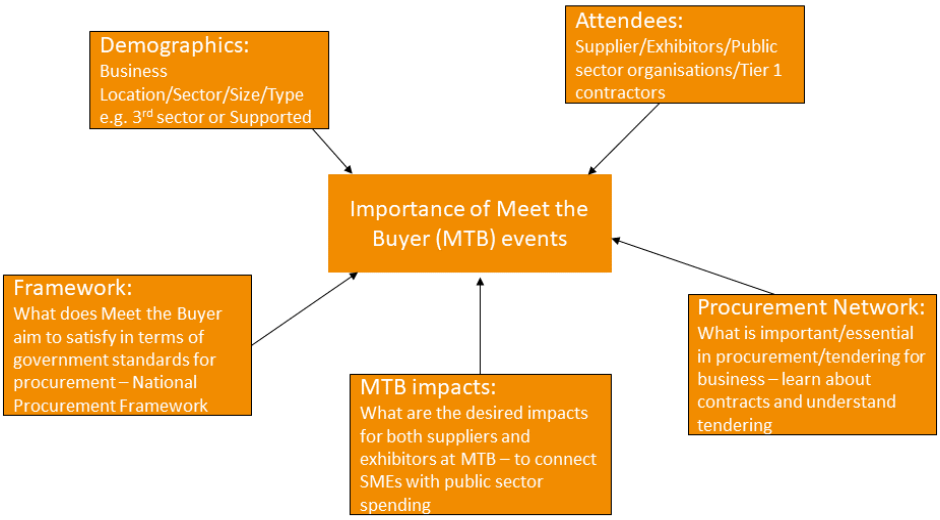


Figure 1.0. – Question Conceptual Framework

Designing a good survey format is one way to gain insightful data however there will always be culprits and responses that inject data inconsistencies into the feedback dataset. The first of these culprits being the choice of response type, either open-ended or close-ended. Close-ended responses have the potential to oversimplify a complex response but are also easier to code and analyse whereas open-ended questions can place more of a burden on a response, it will also lead to inconsistencies but can help in terms of explorative research (Synodinos, 2003). There is no ultimate right or wrong when it comes to response types, the only requirement is that responses are kept reliable so they can provide consistent data over a period of events. Another reason for data inconsistencies is when data is collected and analysed from various sources, for example, if survey datasets are being analysed over a certain period but the formats of that survey have changed over time, this can inject inconsistencies into the dataset where questions and response types were designed in different ways across a survey (Blasius, Nenadic, and Thiessen, 2012). Other measures which will lead to data inconsistency but will struggle to be avoided are those missing data values. For example, data can be inconsistent due to skip patterns (SPMD), genuine missing data (GMD) and

undetermined missing data (UMD) (Arslanturk et al., 2015). Whichever means that data is collected by, whether it is considered good quality or incomplete, the next task in the data oil analogy is cleaning and processing or 'refining' that data.

'Refining' the Data

Data cleaning is the process that ensures data collected by an organisation does not contain any erroneous pieces of data that may corrupt the value to be derived from the resource as well as restricting the final dataset from being usable when it comes to visualising the data effectively and accurately (6 Steps for data cleaning and why it matters | Geotab, 2021). The focus for data cleaning is to remove any inconsistencies but these can be inevitable when multiple data sources are being merged. Fewer corrupted datasets mean the potential value from data can be realised for all users. The analogy of 'data as the new oil' directly corresponds to data cleaning as in nearly all cases of data analysis, cleaning data or 'refining' data must be achieved before it can be used, with its compatriot requiring similar treatment before it can be turned into a product.

Real-life organisational datasets such as those collected by SDP are often seen to be "dirty" in some manner. According to the British Computer Society (BCS), as much as 30% of an organisations data has the potential to be 'dirty' and with the pervasiveness of corrupt data, this has the potential to cause issues for an organisation from misinformed decisions to poor customer experiences (Firms full of dirty data, 2008). To use the public sector as an example, the BCS saw how 81 million National Insurance numbers had been issued to a population of only 60 million, a case in point on the repercussions of poor data collection and management (Firms full of dirty data, 2008). Another example of inaccurate data that can be related to SDP's practices is the "Merge/Purge" (Hernández and Stolfo, 1995, pp.9) problem where multiple datasets that share common entities are merged to achieved Knowledge Discovery in Databases (KDD) but can result in numerous corruptions – duplicate instances is just one – which will pose a serious issue usually resulting in extending a data-driven project due to the additional data cleaning involved. SDP

also use multiple datasets with common entities across their database which must be paid close attention to due to the potential discrepancies when looking to analyse these datasets.

Assessing your data to determine its 'dirtiness' is where the data cleansing journey begins, being able to understand and visualise the format of your data is crucial before taking the plunge and processing it. The data cleansing process has to be a considered one as this is the stage where data will be taken from its raw and potentially dirty form and put into a standardised process which aims to gain the insights and knowledge in the raw resource. This process will then be implemented for future datasets so to keep erroneous data at a minimum however it is a process that data analytics will constantly be evolving. It is estimated that cleaning and processing data takes up the majority of their time for a data analyst or scientist (50% to 80%) due to the initial analysis and cleansing of data revealing further inconsistencies within the dataset that were not visible before (Chai, 2020).

In terms of this project, SDP are operating in an environment that procures huge amounts of data in all its various forms and have realised the value that can reside in cleaning and processing data due to the well-documented nature of this post-industrial resource. Conducting a data-driven initiative is also viewed to be valuable not only for SDP's current operations but will also have the potential to transform their strategies in the future allowing them to maintain a strong position in such a data-driven economy.

Now with the advent of new forms of capital such as data, a new type of economy has spawned to deal with and consolidate these commodities. Previously in this report, this has been referred to as a post-industrial economy and a sort of continuity from the industrial economy. This is not the only definition for this lucrative and fast-growing economy however, it has also been referred to as the "information economy" or the "weightless economy" (Dean and Kretschmer, 2007, pp. 1) but one other term for it that perhaps encapsulates the value data can produce is the knowledge economy.

The Knowledge-Based Economy

The knowledge economy is one where the "main propellant of competitiveness and creating wealth in the company" is knowledge (Giju, Badea, Ruiz, and Peña, 2010, pp. 28). The key factor in this concept is how to gain that knowledge to bolster productivity. This is where data, defined as a resource, can contribute to this type of economy as it is through the structure of data, that knowledge can be found. Knowledge in the knowledge-based economy as the name suggests, shows that against other forms of capital – labour, natural resources, and other types of physical capital – knowledge now takes on greater importance and is the sole factor of production for organisations. The product of knowledge is also viewed as being somehow more important than it previously was before the shift to a more knowledge-based economy (Smith, 2002).

The main factors that contribute to the knowledge economy still contain most of the elements of an economy that is considered not knowledge based. They are however, not seen to be as useful or in such abundance as before with information and digital technologies being the main contributors to the knowledge intensive sectors. A clear distinction in the knowledge-based economy than in an industrial economy is the role that information and information technologies play in supporting knowledge to be such a driving force in the economy (Brinkley, 2006). The role of information technologies is in reference to the widespread implementation of powerful and affordable computing power. The use of these systems has meant managing data – a key resource in gaining knowledge – has become commonplace throughout organisations and at the same time is one of the main driving forces behind the knowledge economy (Goodhue, Quillard and Rockart, 1988). Digital networks also provide ubiquitous access to products and services that embody knowledge and create wealth in this sort of economy (Clarke, 2001). Another factor that has emerged in a knowledge-based economy is the transition from a tacit form of knowledge to an explicit one where knowledge is "codified" (Smith, 2002, pp. 10). By codifying knowledge, it becomes less of a subjective and personal concept and one which is tangible, patentable, and implementable for an organisation to create value from.

The knowledge economy has also created innovative sectors, workers, jobs that previously did not exist or took on less significance in a pre-knowledge-based economy. Whole new knowledge intensive sectors and highly skilled knowledge workers have been established in industry with business services investing in the development and creation for opportunities that will promote the knowledge-intensive skills needed to stimulate their organisation (Hardy, Micek and Capik, 2011). In terms of what knowledge-intensive sectors and workers are defined as, they are represented by a focus on technological change and using intensive information and communications technologies (ICT) rather than investing in research and development (R&D) (Brinkley, 2006). Those who are also performing knowledge-intensive jobs can be seen to be highly skilled or "upskilled" in their roles (Science, Technology and Industry Outlook, 2001, pp.99) with firms merging ICT technology and a skilled workforce to create wealth (Brinkley, 2006). Regarding whom the sectors are which place such emphasis on knowledge-based activities, it can be said that all sectors are involved in this practice from finance and insurance companies to creative and cultural industries (Brinkley, 2006). Reports have estimated that in years to come, nearly a third of the European workforce will be employed in the production and dissemination of knowledge and many more will be needed to cope with demand across all sectors of the economy (Brinkley, 2006).

Regarding this project, SDP have found themselves operating in the knowledge-based economy which focuses its strategies on data and the information it holds, and which attempts to thrive off the valuable knowledge that data could deliver. SDP have realised that they may not be harnessing the full potential available to them in the knowledge economy and that such an initiative could present them with sustainable benefits for their operations. The new knowledge-intensive sectors and employee opportunities that have been created are also being implemented in this project as new analytical thinking merging with ICT technologies will help use the key resource of data in gaining knowledge for the organisation.

The key aspect of the knowledge-based economy is that which creates the value in this economy is infinite in supply and generation. Knowledge much like the resources

that contribute to its creation, are sustainable. It has been termed as the “ultimate economic renewable” (Brinkley, 2006, pp.5) as constant use and generation of knowledge will never diminish its supply. Given the limitless nature of knowledge in the knowledge-based economy, it does not suggest that any knowledge gained is good knowledge, there must be a degree of knowledge management for its value to be realised and implemented effectively.

Knowledge Management

Management has always existed in the economy and due to the emergence of the knowledge economy, it has continued but new forms have developed and one of which is knowledge management (KM). In essence, KM has been created in this economy as there is a need to manage knowledge in an organisation like an asset (Ghani, 2009). KM is both an asset and activity to an organisation and the data they collect and information they extract from all types of data collected are the resources that can lead to KM. There is also a hierarchical structure to KM with it involving certain activities taking place in stages when gaining value from knowledge. The first step in KM is the resource, the data, which exists in many forms often bearing no significance at all. It must then be converted so information can give meaning to the data whether it be through grouped, classified, or relational information. Information can then produce the knowledge and once that knowledge has been accepted, it is managed so the value that has been created during this process can be implemented properly. Once this has been achieved, the last step is what is referred to as wisdom where the knowledge has been managed and is benefitting the organisation in the optimal manner (Giju, Badea, Ruiz, and Peña, 2010).

The concept of knowledge is one that needs some attention paid to it as knowledge is often seen as one that is intangible and subjective and that at a management level, bears little significance. According to Nonaka and Ichijo (2007), knowledge is created through language and communication and is often seen to be a personal value and not generalisable to organisations. Nonaka and Ichijo (2007) also see from this that

there are two types of knowledge, one that is tacit and one which is explicit. The tacit knowledge is the type which represents an individual's knowledge and is therefore increasingly difficult to explain, whereas the explicit type is the form of knowledge that can be formalised into codified language with numbers, formulas, and algorithms. It is this form of "intellectual asset" (Nonaka and Ichijo, 2007, pp. 67) that can be managed and implemented at an organisational level. It is also the asset which is benefiting all sectors of the economy through digital technology developments and highly skilled workforces who are successfully transferring knowledge.

One initiative that is implementing KM in practice is knowledge discovery in databases (KDD). KDD can involve many data science methodologies such as data mining, data processing, AI, and machine learning among others but the main aim of KDD is to first extract high-level knowledge from large and unsophisticated datasets and then to formulise this knowledge so it can be displayed to the 'knowledge manager' (Fayyad, Piatetsky-Shapiro, and Smyth, 1996). The process of KDD is a multi-disciplinary activity for KM with a unique emphasis placed on identifying patterns and insights in large datasets which can then be interpreted into pieces of useful knowledge for an organisation. An overview of the steps that take place in KDD from selection to interpretation is shown in *Figure 1.1.* below:

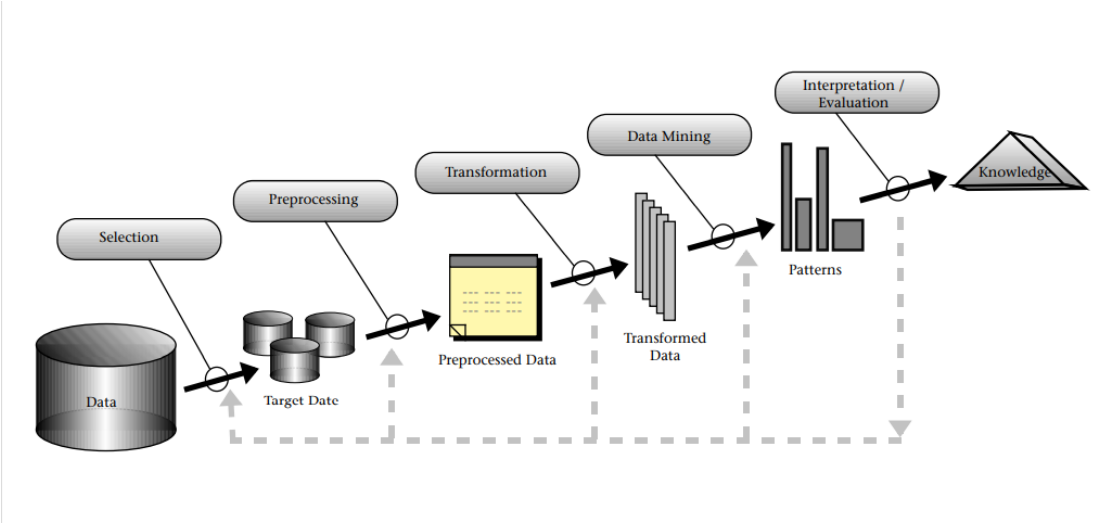


Figure 1.1. - Knowledge Discovery in Databases (KDD) Process (Fayyad, Piatetsky-Shapiro, and Smyth, 1996)

In direct comparison to this project, it can be said that SDP are conducting a knowledge management initiative with one aspect of that being KDD. Firstly, SDP are traversing their datasets to extract meaning through information and then gaining knowledge which will be related to their best practises of procurement advice and national services in the public sector (e.g. Meet the Buyer events). This initiative has also been developed to provide the organisation with actionable knowledge that can bolster and improve their productivity in a sustainable manner. Regarding how KDD is relevant to the project, one aspect of the project is selecting datasets to identify patterns within them and having the overall aim where pieces of knowledge can be interpreted from any patterns and trends discovered.

SDP are also a partnership based in the public sector which is one just sector that has not been implementing initiatives consistent with KM as proficiently as it can be said for other sectors of the economy. Even though knowledge is recognised as a valuable and strategic resource for all organisations, due to its effect on the various aspects of the economy as mentioned previously. Public sector organisations have been less inclined to discover the possible advantages that KM could offer their organisations (Arora, 2011; Moretto, Ronchi and Patrucco, 2017). One of the issues in the public sector regarding KM was the awareness of such strategies, public sector organisations may have been developing KM frameworks, but they were doing so without knowing (Cong and Pandya, 2003). This awareness issue meant that the possible benefits KM would bring were not being fully reaped. Public sector organisations do face other challenges than say a private sector one does, such as austerity, and this can hinder their ability to focus and implement a KM strategy to its full potential. There are some public sector organisations however, who are realising their collective intelligence can help to achieve and bolster their strategic intentions, one of which is SDP themselves.

KM has been seen to benefit all types of organisations however it is not an approach that is developed and implemented without a strategy or purpose behind it. In the case of this project, there is a clear vision for how valuable a KM initiative will be as "new analytical thinking with the aim of implementing sustainable, innovative

improvements" (Supplier Development Programme, 2021) is just one of the goals that SDP would seek to procure from a project of this nature. The methods behind capturing the value from this project is also planned with data collection, data analysis and data visualisation all being implemented in the project. And the value of a KM initiative does not stop once the project concludes, internal and external stakeholders will also realise the benefits it has. From the procurement advisory professionals gaining value through knowledge to businesses gaining value through contracts being satisfied and suppliers gaining with contracts being secured, this project will go far in attempting to display the contribution to the procurement strategy SDP have in Scotland.

Digital & Information Technologies

Digital and information technologies play a major part in the knowledge-based economy. Organisations are now realising the benefits and value data, in all its forms, can have. Merely collecting data is not enough, data has a value to be extracted from it. And the relevant methods for such extraction utilises digital and information technologies through processing, analysis, mining, and visualising data. With the rise of the knowledge economy, there is seen to be a revolution of digital and information technology (David and Foray, 2003).

The rise of digital and information technologies has meant that data is ubiquitous, it is constantly being created, stored, managed, retrieved, shared, and analysed. This 'revolution' has also seen an upsurge in technologies that are designed to interact with the digital and information technologies that collect and store the data itself. In the last couple of years, technologies and services that use data analytics were predicted to increase by about 36% and the global income for data to be analysed using these technologies was expected to rise by more than 60% (Hariri, Fredericks and Bowers, 2019).

Perhaps one of the biggest impacts that digital and information technologies can have in an economy where data is the resource and knowledge is the asset is that

data collected from every-day activities can now hold value. Larger and larger datasets are being stored in the knowledge economy and whereas before, manual analysis would not only be expensive but inconceivable for such copious amounts of data. Now, digital and information technologies have the capability to explore and analyse the contents of vast datasets and provide means of knowledge as an asset from it (David and Foray, 2003). Regarding how digital and information technologies will be utilised for value in this project. They will form the basis for storing the datasets in question, conducting relevant analysis, and developing visualisations, none of which would have been possible before data was considered the main resource in a knowledge-based economy that has expanded its use of digital and information technologies on a huge scale.

Data Visualisation: A picture that is worth more than a thousand words

One of the main developments that has also coincided with the revolutionary rise of digital and information technology and the ubiquity of data is the field of data visualisation. The key element to data visualisation is that they can convey the message in the dataset in the clearest and most succinct manner possible (Gromping, 2017). Data visualisation has become a huge attraction in the data science field with data visualisations being "liked and shared up to three times more than any other type of content" and the main reason for this is that we are incredibly visual learners. Nearly all the information we transmit to our brains is visual (90%) where our brains can process visual information in just 13 milliseconds meaning that data visualisation is incredibly powerful at allowing users to understand and retain huge amounts of data in a short space of time (Data visualisation: The pictures that are worth a thousand words | Longitude, 2021).

The growing attraction of data visualisation has meant that more and more software applications have been developed, each aiming to visualise data in their own unique way, from open-source tools like Python Plotly to more business orientated setups such as Power BI and Tableau (Schmidt, 2020). Data visualisation expertise is perhaps one of the most sought-after skills in the business world currently, as visual

data models can allow for the greatest exchange of knowledge across the whole organisation even when using complex datasets. Search hits on employment websites such as 'Indeed' and 'SimplyHired' showed that Microsoft with their visualisation tool Power BI topped the charts with Tableau being the next closest alternative which employers mention in their job descriptions (Diamond and Mattia, 2017). SDP have clearly recognised the power of data visualisation along with the rest of the business world as their project is aiming to develop data visual models by incorporating a data analyst position in their organisation. Allowing them the opportunity to improve on their MTB outcomes where more data can be easily accessed and understood by their stakeholders.

Data Visualisation: The good, the Bad, and the Innovative

To successfully display a data visualisation model so the information can be processed and understood is heavily reliant on using the correct visuals for the right data. The first step in creating a data visualisation is to identify the purpose of the visual model which could be the audience or users for the data visualisation, and this will dictate what data will be used in the first place. Choosing the visualisation is the next stage once the data has been selected and this decision is crucial otherwise the narrative of the data will fail to register with the audience. When selecting how to visualise data, the relationships in the datasets must be taken into account.

Drake, Pytlarz, and Patel (2018) defined the various relationships data can have when visualised, starting with 'Comparisons', this is best represented through grid visuals such as line charts for time series data or with bar charts to show a quantity comparison. Comparable data can also utilise geospatial data to compare values between different locations, for example, a choropleth map. Another data relationship is 'Distributions' which are limited in their use of effective visualisations with only histograms and box plots being best suited to this type of data relationship. 'Flows' on the other hand can take several forms when visualised effectively. They can also be seen to be most interactive data visualisations with chord charts, sunburst charts, and tree diagrams all being able to display this data

relationship. So far Drake, Pytlarz, and Patel (2018) have mostly demonstrated how numeric data relationships can be visualised but there are other types of data that need visualised, large string text data such as comments or reviews for example. 'Unstructured' data relationships like text data can be broken down into word clouds or pivot tables to convey just the highlights of such a dense data relationship.

Another crucial element to effective data visualisations as mentioned by Drake, Pytlarz, and Patel (2018) is the appropriate use of 'scaffolding'. 'Scaffolding' techniques are used to aid with the interpretation of the data visualisations and include using suitable data labels, quality labelling and correct use of axes, legends, useful metadata, and overall framing of the data visualisations. Any techniques that mitigate the chances for misinterpretation of the data visualisation are a necessity for any data-driven project otherwise the messages in the data which can be conveyed so concisely through visualisation can be lost or worst, misunderstood. Other useful 'scaffolding' techniques are by simply adding explorative text boxes to describe the data and any nuances related to it. Of course, there is also the issue of overloading the user with information and using too much 'scaffolding' which can detract from the impact of the data visualisations so a balance must be struck when including such aids (Drake, Pytlarz, and Patel, 2018).

It is hard to ignore the power of data visualisations, but it is also hard to ignore the abundance of guides, hints, tips, and principles for creating good and effective data visualisations. There are however, some organisations – some of them being major players in business and broadcasting – who still fail to comprehend how data should be visualised in order to have its desired impact and still commit some of the most common mistakes. A selection of these poorly designed visuals can be seen below:

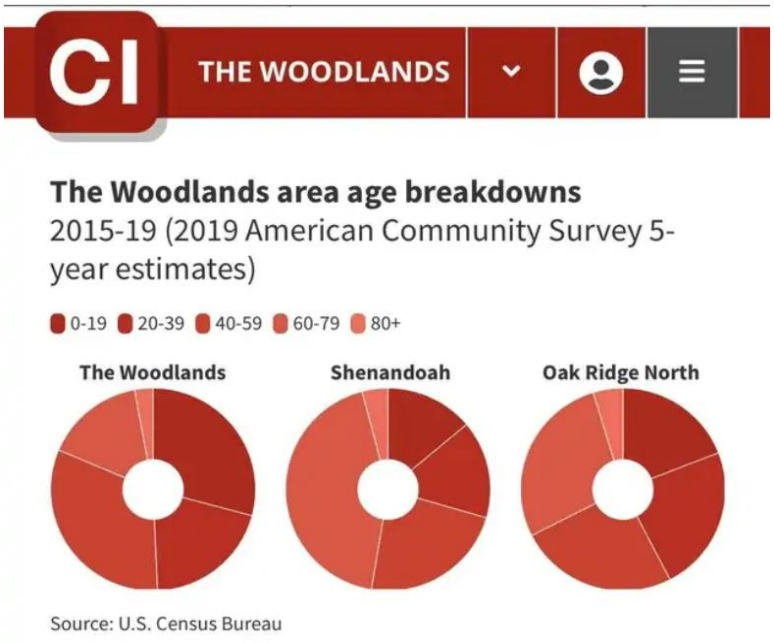


Figure 1.2. – Bad data visual #1 (Old Street Solutions, 2021)

Community Impact Newspaper released this visual using data from the U.S. Census Bureau. Using the donut graph to visualise the data does allow for the different values to be split into intervals that equal the total and the interval values are consistent, however it is evident that the colourisation does not display the data effectively. Being unable to distinguish between these sequential colour intervals results in the reader not understanding what each age bracket connotes. Using a far more divergent colour scheme in this visual would help the reader greatly.

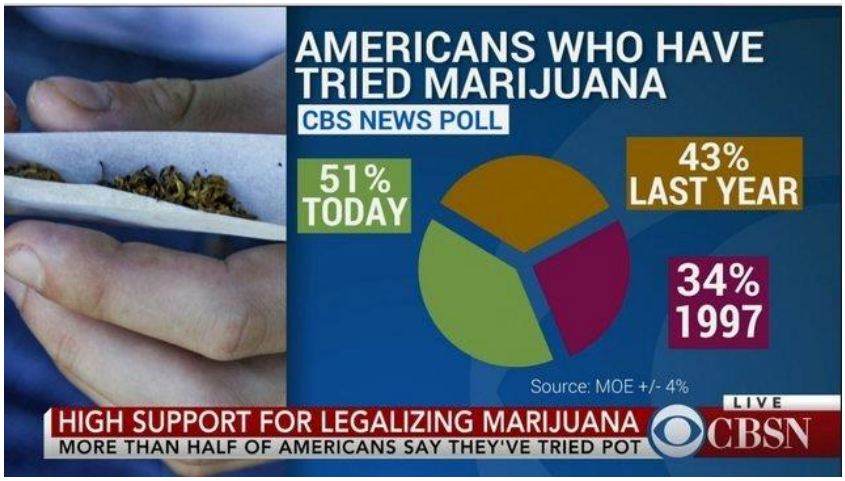


Figure 1.3. – Bad data visual #2 (tumblr, 2021)

The visual above – by a prominent news channel – is one which is using a visual for no comprehensible reason at all. The pie chart is considered an overused and often contemptuous visualisation for data (see Kozak, Hartley, Wnuk and Tartanus, 2015) however, it has no relevance in this figure. The data shown here is time-series data comparing values over various years which results in the cumulative total value not equalling 100% and illustrates how this visual has no bearing towards the percentages shown.

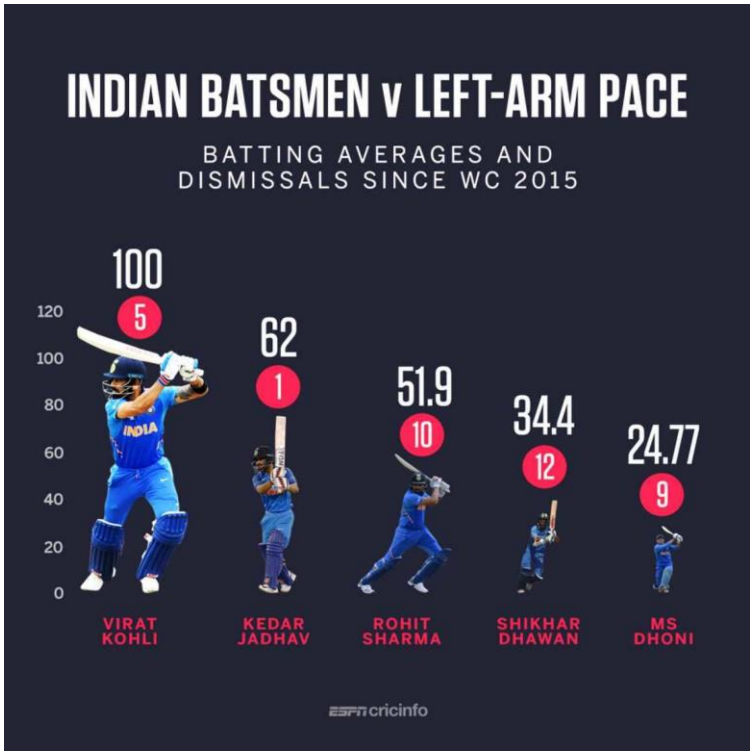


Figure 1.4. – Bad data visual #3 (tumblr, 2021)

ESPN used *Figure 1.4.* to show comparisons in this data relationship however, due to the visuals used, it struggles to convey the message effectively. Firstly, by using non-standard visuals such as the various images of the batsmen rather than a bar chart, it can hide the information in the visual as the values used on the y-axis fail to provide useful scaling to interpret the data. The link between the figures and metadata also contribute to the ambiguity within this visual, where there is no clear definition as to what the headline figures represent above the batsmen.

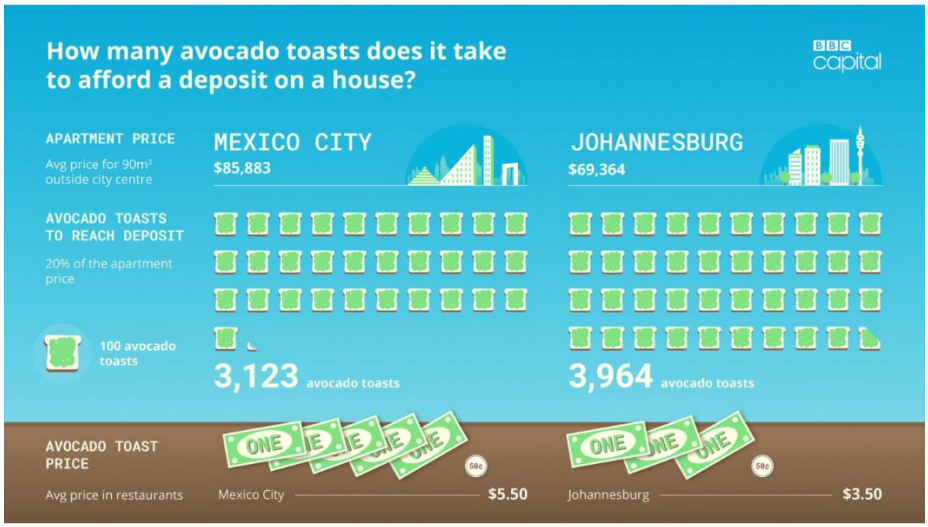


Figure 1.5. – Bad data visual #4 (BBC, 2019)

BBC used this infographic visualisation to display comparisons in a data relationship however, each visualisation only includes two cities to compare avocado toasts and house deposits. This static visualisation was repeated five times so it could include the data on all cities in question which makes it difficult for the reader to compare the data across the whole dataset when a simple bar chart would have done this.

Recent innovations in data visualisation have meant that those traditional static visualisations are not able to satisfy the data models organisations are needing visualised. A prime example of these innovative data visualisation models is what is defined as a data dashboard. To convey the development of data dashboards as the latest innovation in data visualisation models, a data visualisation below shows the data surrounding initiatives that include data dashboards:

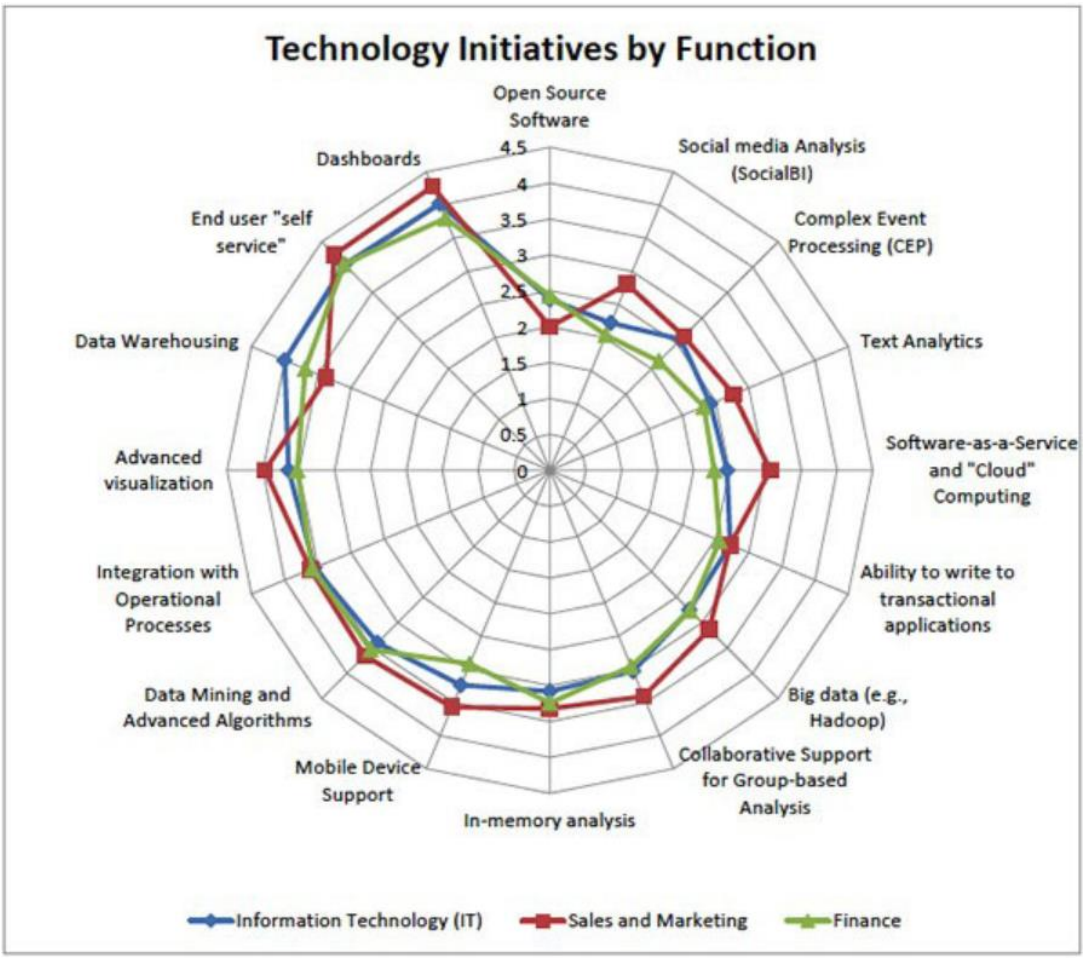


Figure 1.6. – Technology initiatives (PC Mag UK, 2019)

The dashboard interface style for data visualisation has been utilised by most of the big names in this area of expertise (see Power BI, Tableau, d3.js and Plotly) with this new tool allowing data analysts to create the data visualisations but then empower the end user with the interactive capabilities to visualise the data model in a far more accessible manner (Drake, Pytlarz, and Patel, 2018). The interactive nature is key to the benefits that the data dashboard can have for data visualisations as the consumer of the data is now able to explore far more data in an environment without the interference or skills of a data analyst required.

These new innovations in data visualisation have been implemented by several organisations so far, with one example operating alongside SDP in the procurement network in the public sector. Unsurprisingly, Microsoft have also heavily invested in

data dashboards with several versions of their investor dashboard being developed in recent years (Partner showcase | Microsoft Power BI, 2021).

The other significant example however, of a data dashboard being incorporated within an organisation's data analysis framework is that of Scottish Futures Trust (SFT). SFT developed their own data dashboard by utilising public sector data in an innovative way within their Construction Pipeline Forecast Tool (CPFT). The CPFT is a data dashboard powered by Microsoft Power BI which visualises datasets on construction projects across Scotland aiming to provide greater transparency to the data within the construction pipeline (Media - Scottish Futures Trust, 2021). SFT as a public sector organisation operate within the remit of Scottish Government, much like SDP, and are the centre for infrastructure and construction expertise for other public sector organisations. SFT's innovation came because of the Construction Industry Recovery Plan which was in response to the impacts of the COVID-19 pandemic where one of the main consumers of this data is SMEs, particularly those based in Scotland. This type of data-driven project provides great reference for SDP as it is reminiscent of their initiative using data, especially data visualisation to improve the visibility of their own event outcomes. The objectives this project aims to satisfy in terms of developing innovative data visualisation models are also exemplified in SFT's new addition to their data analysis framework as well as their aim to provide greater data accessibility, transparency, and literacy to both internal and external stakeholders.

Business Intelligence

The aim of Business Intelligence (BI) is to "improve the timeliness and quality of inputs to a decision process… with analytical tools to present complex and competitive information to planners and decision makers" (Negash and Gray, 2008, pp. 177). The process of BI also includes taking the raw resource of data with the processes involved such as data collection, storage, and analysis to provide knowledge for those planners, decision makers and ultimately stakeholders to manage (Azvine, Cui and Nauck, 2005). Organisations often implement BI by using an

architecture documenting all the activities involved when using their raw data to gain knowledge however, many of these architectures contain various components (see Baars and Kemper, 2008; Watson, 2009).

One component that all BI architectures have started using on the other hand, is that of data visualisation (Ong, Siew and Wong, 2011; Watson, 2009). The recent innovations in data visualisation tools have meant the raw data collected and stored by an organisation can be displayed to the end user – planners, decision makers, and stakeholders – in a far more accessible manner due to the amount of visual information that a person can absorb. By displaying far more information through visualisation, it provides those with the power in the business to act quicker and more competitively than if other data reporting structures were used. And with research showing the adoption of BI on the rise due to the strategic benefits it can have (Columbus, 2019; Pucci, Raes and Mareels, 2021), it is clear to see in terms of this project, how using SDP's raw data resources and visualising them can help boost productivity and better inform their stakeholders, both internally and externally.

Methodology & Implementation

Introduction

The aims of this project were to bring new analytical thinking to SDP by utilising the various and necessary data analysis techniques that a project of this nature would require. The project also aimed to realise business value for SDP as an organisation operating in the public sector and their stakeholders by analysing and visualising the datasets from their events and activities within the procurement network. The following methodology and implementation section of this paper will outline how this project developed over the course of the placement and in doing so, attempted to satisfy and capture the indicators that would result in this being a successful project. Firstly, it will focus on the raw datasets which were the subject of the project in terms of analysing and visualising them – the post-event feedback from both suppliers and exhibitors at Meet the Buyer (MTB) and Meet the Buyer North (MTBN). It will then summarise how the analysis began with recommendations to the data collection techniques before moving into the core of the project with how the data visualisation developed from an idea to implementation and then finally review the data science skills used in each aspect of the project.

The Datasets – Meet the Buyer (MTB) & Meet the Buyer North (MTBN)

The original datasets that this project was tasked with analysing and visualising contained historic feedback from post-event surveys for the two audiences that SDP provide services for in the wider procurement network – suppliers and exhibitors. Suppliers are the SMEs that SDP offer tender training to as part of their programme and are the main beneficiary of SDP's services as they are bidding for new revenue and business opportunities. It is the exhibitors or buyers that attend SDP's events who are providing the business to those suppliers, the exhibitors range from Local Authorities to corporate bodies and other support organisations.

The feedback data contained responses to questions ranging from demographic information (e.g. where the supplier's business is located; which sector the supplier's business is classified as; what organisation the exhibitor is from) to event experience (e.g. how easy was the virtual portal to use; ability to speak with suppliers in the virtual queue) as well as marketing and communications (e.g. what social media channels do you follow SDP on).

Response rates to the post-event surveys fell around the 20% mark for both audiences which can be seen to be around the average for an organisational survey (How to Increase the Response Rate on Surveys | Qualtrics, 2021). To clarify, this project did not attempt to increase any future survey response rates as this was historic data from SDP's past MTB and MTBN events which already existed within SDP's data collection framework. It did however, look to utilise additional historic datasets from other MTB events including partner and regional events such as Meet the Buyer Fife, Meet the Buyer Ayrshire, Meet the Buyer Forth Valley as well as MTB and MTBN data that transcended the original annual event data so more post-event feedback data could be analysed and visualised to provide SDP with a better picture of how their past events and activities in procurement had been interacted with. The format and design of the survey format is done by SDP themselves however the electronic dissemination and collection of the surveys to those who registered for the event is assigned by South Lanarkshire Council as part of their Consultation, Organisational Development and Equality Team.

Regarding the format the datasets were held in, this was done through Microsoft Excel with each MTB post-event feedback data file in a Microsoft Excel Worksheet. Taken into consideration the literature mentioned before around the use of the Excel spreadsheet (Big Data LDN & Qlik, 2019) and the limited reporting capabilities it has. It can be highlighted how SDP were still in the early stages of utilising the value in their data as they were yet to move beyond Excel when reporting on their data. There had been previous attempts at analysing and visualising the MTB feedback data however this was still contained in Excel and had not been maintained or innovated, possibly due to austerity-depleted resources.

The limited response rate was one factor which had to be considered when analysing the original datasets however another issue, prevalent in organisational and real-life datasets (Firms full of dirty data, 2008), is that of data inconsistencies and erroneous data. The culprits for the corrupted data in the original datasets were the response types, the reliability of the question format, and any null values. For example, in one post-event survey, the nominal business location responses included locations such as 'City of Edinburgh', 'Dumfries and Galloway', and 'Aberdeen City Council' whereas in future surveys, the response type had changed to 'Edinburgh City', 'Dumfries & Galloway', and 'Aberdeen City' which are of course the same locations however when using data analytical programming tools, these responses were classified as distinct and so resulted in data inconsistencies.

Other examples of corrupted data in relation to the reliability of question format were due to some questions changing format over subsequent surveys – a major cause for data inconsistencies (Blasius, Nenadic, and Thiessen, 2012) – such as how suppliers classified the size and type of business. In some years, suppliers were asked a dichotomous 'yes'/'no' question if their business was either an SME or 3rd sector organisation however in other surveys, suppliers were asked how many employees their business had with ordered responses ranging from 'sole trader' to 'large' business including an option for a charity business **(*see Appendix C*).** In terms of the null values in the datasets, this was down to 'missing data caused by skip patterns (*SPMD*)' where respondents who had not attended the event are not supposed to answer a question about their overall experience of the event. Other null values were caused by respondent fatigue and 'undetermined missing data (*UMD*)' where respondents would not answer a question or any of its related branching questions (Arslanturk et al., 2015, pp. 732).

Reviewing Data Collection – MTB Survey Format

The aims of this project had to be satisfied by completing the project objectives and the first of these was to undertake a review of SDP's data collection methods – post-event feedback surveys – and recommend any changes to be implemented through

formulated and revised feedback surveys so SDP could better assess the effectiveness of their future MTB events.

The first task was to sort out the inconsistencies which were causing the historic datasets to be erroneous, and this was a simple fix by implementing a standardised response type for business location so it was consistent with most of the previous questionnaires and would stop any corruption in future feedback data. On the suppliers' side, the business size and type question format were kept to the ordered responses, firstly to maintain reliability with most of the previous questionnaires' feedback but also allow for more granularity in the data analysis so a business could be viewed as an SME with data about the particular size (e.g. sole trader, micro, small, or medium) of the business being collected as well.

The next task in the early stages of the placement was to recommend additional questions to the supplier's survey that would try and gain more insightful data (responses) into how suppliers interact with and view the MTB events. Research was conducted into the wider procurement network, Scottish SMEs and their impact on the economy, Public Contracts Scotland (PCS), as well as survey 'Conceptual Frameworks' among others (Jenn, 2006). This research provided the project with knowledge about the National Procurement Framework (NPF), PCS and the contract opportunities available to suppliers through that portal, how Scottish SMEs had been impacted by the global pandemic, and an overall insight into how suppliers' responses need to be emphasised and that this could be achieved through data collection and analysis. Other considerations when designing the survey format was to avoid common errors such as double-negative statements, leading questions, ambiguous wording, and misplaced questions (Malhotra, 2006).

Some questions already included in the survey did focus on some key performance indicators (KPIs) from a supplier's perspective. One of these being whether they were 'more likely' **(*see Appendix C)*** to bid for public sector contracts after attending an MTB event – this was one of SDP's main outcomes for suppliers' post-event responses from past survey analysis. This project however, provided questions that gained additional insight into suppliers' outcomes from an MTB event. Knowledge is a key aspect to a supplier engaging with the procurement network, so questions

were asked that gained a baseline knowledge level about PCS opportunities as well as the tendering process before asking how their knowledge had been affected after attending an MTB event. Other questions added attempted to discover what the most important aspects of an MTB event to a supplier in relation to the procurement network and the NPF. Another major part of MTB event is networking between attendees, so questions were designed that firstly grouped the number of new contacts suppliers made and then aimed to discover how these contacts would be beneficial to their business. Using the current economic climate for inspiration, further questions added aimed to collect data on how Scottish SMEs who interact with SDP have been affected by the COVID pandemic and if the MTB events were providing routes for recovery to them. This type of data would provide valuable insights into comparisons with other data sources on how SMEs, particularly Scottish businesses, have been affected by COVID (Federation of Small Businesses, 2021). An exhaustive list of the questions added and edited to the survey format can be found in ***Appendix C.***

Regarding the response types for the additional questions in the survey, these were mostly ordinal in nature with a ranked or scaled order using Likert Scale. Likert Scale were used due to the easy interpretation and completion for a respondent as well as their likeliness to produce a reliable scale to the responses (Bertram, 2007). There were optional text boxes for certain questions to gain more qualitative data if the respondent preferred however these were kept to a minimum to mitigate respondent fatigue.

The review of SDP's data collection techniques was undertaken in the first week of the placement so the recommendations could be implemented for their survey that would be distributed to attendees for Meet the Buyer 2021 on 8th June.

Regarding the exhibitor perspective, no changes were needed for their survey format due to the limited questions and insights for MTB events. The emphasis was placed on how suppliers interact with SDP's events and activities in the wider procurement network.

Processing and Analysing the Datasets – MTB (2017-2021), SDP & ONS data

Once the review of SDP's data collection had been undertaken, this project had a far better insight into what data had been collected in the historic MTB feedback datasets as well as what KPIs SDP seek to analyse and visualise about their events and activities. This teamed with the additional data gathered on other MTB events which were supplementary to the original datasets (MTB Fife, MTB Ayrshire, MTB 2017, MTB 2018, MTB Forth Valley) given at the start of the placement meant that the processing (cleaning, filtering and transforming), analysing, and finally visualising of the data could begin. Ultimately, the innovative data visualisation model that this project would create and implement was intended to improve upon on the existing visualisations for SDP's membership organisations (***see Abstract***). SDP's membership organisations consist mainly of all 32 Local Authorities across Scotland and all these public sector bodies are interested in the overall outcomes from MTB events particularly the outcomes related to those Scottish SMEs who attend the events. This meant that the domain for data processing focused upon Scottish SME responses in the datasets who had attended MTB (2017-2021) events. The other reasoning behind this focus was because SDP concentrate on offering free and unique only to Scottish SMEs as their primary consumers, so it meant that the data analysis and visualisation would provide effective and useful insights for both SDP and their membership organisations.

The powerful and versatile environment offered by Jupyter Notebooks using Python 3 was utilised when processing all the datasets (see Project Jupyter, 2021; Python, 2021). Python is endowed with various libraries (pandas, numpy, matplotlib, geopandas, folium – all of which were tried and tested in this project) which aid with data cleaning, wrangling, analysing, and visualising data which was the major reason it was used in this project. The official Python docs for each library as well as other Python literature, tutorials, and forums were utilised throughout this project. Python is especially suited to Explorative Data Analysis (EDA) which made it ideal for a

project of this nature where it had to explore and visualise the datasets and make the “data speak for itself” (Bezerra et al., 2019, pp. 2).

The main Python library that was used was Python Data Analysis Library, otherwise known as pandas. Pandas is a very fast and efficient library to clean, wrangle, merge, and analyse data (pandas - Python Data Analysis Library, 2021). The main functions of pandas used in this project were the DataFrame (*df*) object which provided the ability to effectively manipulate the data. Manipulation of the data meant handling the missing data such as the null values, SPMD and UMD and ordering SDP’s ‘dirty’ datasets. Other manipulation techniques were to reshape the datasets with various indexing, selecting, and subsetting functions so only the key questions and responses from Scottish SMEs in the excel survey files were selected for analysis and visualisation. Additional data manipulation took the form of handling the afore mentioned inconsistencies in the datasets so the filtered datasets for visualisation would represent survey responses effectively and accurately. High performance merging and joining functions for datasets were also advantageous for this project as it allowed for datasets to be analysed, visualised, and compared which previously SDP would have found far more computationally intensive and time consuming (all Jupyter Notebooks and relevant code repositories can be found in the additional files).

This project did not just limit itself to data analysis and visualisation of SDP’s MTB datasets. The reporting and exporting (.csv and .pdf files) structure offered through SDP’s Content Management System (CMS) allowed this project to bring in datasets pertaining to accurate MTB attendance by region (not all attendees will complete the post-event survey) and total SME registrations by sector, business size, and region. These data sources were invaluable to this project as it meant that this type of data could be analysed and visualised in a far more accessible manner and provide a better insight to SDP’s activities in the wider procurement network to their membership organisations.

There were some limitations to the CMS as there were a select number of filters for the database however the power of pandas meant that the datasets exported could be merged, joined, filtered, and visualised in ways that SDP would not have

previously been capable of. For example, exporting datasets from the CMS related to regional Scottish SME sector and size totals could only be done one sector or business size at a time. Pandas merging functions allowed this data to be concatenated so values on each sector or business size were joint together in one dataset and could be compared and visualised in a far more accessible manner.

Another source for audited SDP data was accessed through their corporate documentation such as their annual reports. Using this data for analysis and eventually visualisation kept the data accurate and consistent across the organisation.

This project also moved beyond just analysing and visualising SDP's datasets; wider national data sources were also accessed to provide SDP with a broader sense of their progress in the procurement network. Office for National Statistics (ONS) data sources on national SME figures by business sector and size were exploited so they could provide comparable data with SDP's figures along with a deeper awareness to SDP and their members as to what progress is still available to their organisations (Small and medium enterprises in county by industry and size - Office for National Statistics, 2021).

Visualising the Datasets

Once all the relevant responses from the survey datasets, the MTB attendance data, the total SDP SME registration data, and the appropriate ONS datasets had been processed and analysed. It was time to develop the innovative data visualisations for SDP and their members which was one of the main project objectives that had to be satisfied. The final version to be implemented and delivered through SDP's website, another project objective, took the form of a data visualisation dashboard. The reasoning behind choosing the data visual model as a dashboard was to allow the data visual to be created but then empower the end user with the interactive capabilities inherent in the data dashboard so the user, in this case SDP's members,

can explore, interact, and access the data in multiple ways (Drake, Pytlarz, and Patel, 2018).

The Data Model

From the original raw datasets that had been processed and analysed, these data files now contained data on Scottish SMEs along with their MTB responses. Using data solely on Scottish SMEs meant that all 32 Local Authorities in Scotland were represented with values in the data files even if some of their values equalled zero. The data model for visualisation was therefore created based on Local Authorities and the Scottish SME data in that respective region. The critical reason for this was to allow SDP's membership organisations – who represent all 32 Local Authorities across Scotland – to access data specific to their own region in the data dashboard and see how SMEs under their authority are interacting with an organisation they are invested in.

To use Wilkinson's *'Grammar of Graphics'* to represent the data model (Wilkinson, 2012):

The data model used can be seen as using a domain *'O'* as a set of objects and a codomain *'V'* as the set of values. The function *'f'* assigns each element of *'O'* an element of *'V'*. In this case, *'O'* is represented as the domain containing all 32 Local Authority objects and *'V'* is a set of values such as MTB attendance values across all Local Authority objects. To denote a specific value, it can be represented by *'x'* as a value in the set of values *'V'* where *'x'* is an element of *'V'* $(x \in V)$, and a value for a specific object is denoted as *'X(o)'* where *'o'* is an element of the domain *'O'* $(o \in O)$ i.e., a specific value for its respective Local Authority in the domain of all 32 Local Authorities.

The Data Dashboard

The innovative data visualisation which took the form of a data dashboard was created and implemented using Power BI and the various features and tools it offers

for users to prepare and implement an interactive data visualisation, or what Power BI refer to as a 'Report'.

Dashboard Design:

Using the domain of all 32 Local Authorities as the basis for the data model, the main aspect of the data visualisation was designed so the user could interact with the dashboard by drilling down through the national data picture. This allowed the user (SDP's members) to discover more granular insights in the data such as data on Scottish SMEs' MTB attendance specific to their region in Scotland. For example, if a representative from Fife Council – a SDP member – required to find out how many SMEs had attended MTB events between 2017-2021, they would be able to interact with the dashboard and drilldown to that specific regional data.

Power BI does offer interactive capabilities as this is one of the core elements of the tool and this was utilised to create a 'slicer' which would offer the user the drilldown interaction. The 'slicer' was created using the buttons and bookmarks functions available in Power BI which would filter the hierarchy of the data model from all data on all 32 Local Authorities which is the basis for the data in the visualisation to a selected region. To make the 'slicer' more interactive and using an on-theme design to display the data, the core brand elements of SDP were implemented into the visual which meant the dashboard could come to life. That 'life' was provided courtesy of ***'10-DR'*** (ten-de-r), who is SDP's tender robot mascot which can be seen below:

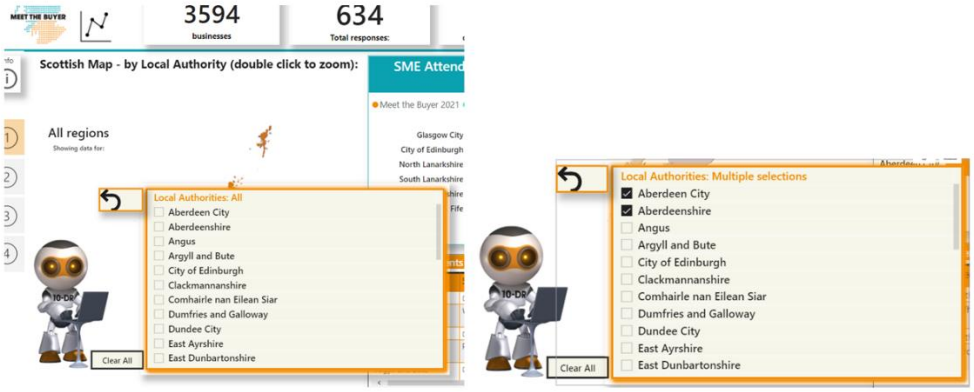


Figure 1.7. – MTB data visualisation (Slicer Panel)

The 'slicer' for users to filter the data model is provided using a panel which contains all 32 Local Authorities on it. When users are interacting with the dashboard, they simply click on ***'10-DR'*** which will bring up the slicer panel, they can then select which filter they desire for the data model, this will filter **all visuals** on the dashboard to their selected region and can then hide the panel using the 'back button' located to the top left of the panel. The multi-filter option from Power BI was also used (shown in the slicer panel on the right) so users could select multiple regions to filter the data. For example, this would be useful to a user when they wanted to filter data for the Forth Valley region, where Stirling, Clackmannanshire and Falkirk would all need to be selected. Other benefits of this option would be for partnership member organisations such as Strathclyde Partnership for Transport (SPT) where users would require up to 12 Local Authorities to be selected.

The colour scheme used in the data dashboard was also a major aspect to the visualisation and this was kept consistent with SDP's brand guide for the on-theme design. Only colours contained in SDP's branding manual were used in the visual so once it was implemented on their website, it did not look out of place with their overall brand image. The colours and hex codes used are as follows: Orange (#f28c00); Black (#1d1d1b); Red (#7d0f00); Light Blue (#0aa1b0); Grey (#6c8da2); White (#f6f6ea); Blue (#6d91cb); Green (#00f28c).

Scotland Local Authority Map:

The first data visual to be developed for the final data dashboard was the map visualisation. Using the map as a data visualisation instantly shows the reader or user which locations and values are going to form the basis for the whole data visualisation model – that being Scotland and all its 32 Local Authorities. The map went through many iterations to find a suitable format, Python packages such as folium, plotly, and matplotlib were tried and tested. After many iterations, the final version shown in the data dashboard is a combination of Python visuals and Power BI custom visuals. As Python and R developed visualisations are not supported in Power BI reports on a website (Publish to web from Power BI - Power BI, 2021), the map had to be rendered using matplotlib and then created as an .svg file using the custom Power BI visual 'Synoptic Panel' which is specifically suited to creating unique map visuals (Synoptic Designer for Power BI, 2021).

The geospatial data used to render and create the map visual was provided under the Open Government Licence (OGL) for download and use by the Spatial Hub which is part of the Improvement Service (Local Authority Boundaries - Scotland - Local Authority Boundaries - Spatial Hub Scotland, 2021). The data can be found here.

The final map visualisation in the dashboard can be viewed below:

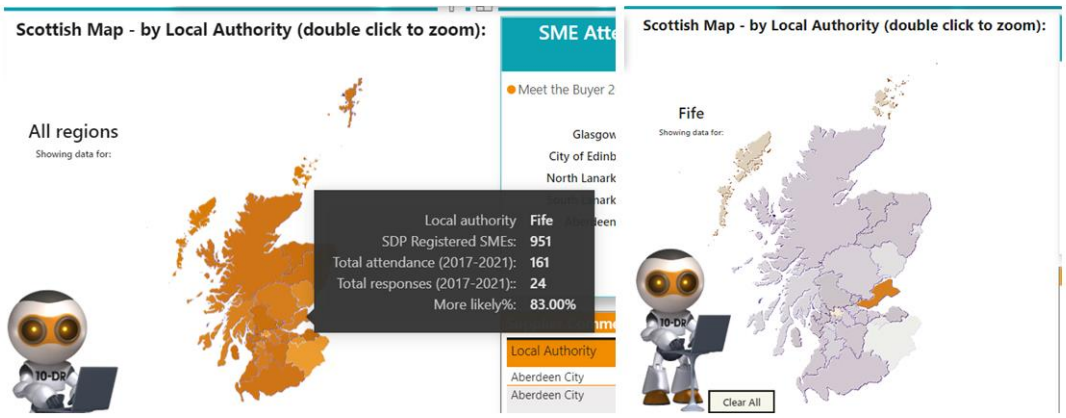


Figure 1.8. – MTB data visualisation (Scotland Map)

The Scotland map is divided into each of its 32 Local Authorities on the dashboard where each region offers interactivity for the user. On the left in the screenshot above, a user can hover over a specific region which will bring up a tooltip for that selected region displaying some of the data highlights (Local Authority title, attendance, and registration data). Once a region has been filtered using the slicer panel, this will also interact with the map visual as can be seen on the right in the screenshot above. For example, a user has filtered the data to Fife so this region will be highlighted on the map for them.

The map visualisation was also selected due to the perceived benefits geospatial data can bring to visuals and its users. This type of data has been seen to improve the insights for users due its ability to visualise location related information easily, in the case of this visual, that location information is for all Scottish 32 Local Authorities (Geospatial Data Visualization, 2021).

Stacked Bar Chart – MTB Attendance Data:

The stacked bar chart visual used data sources from SDP's CMS reporting structure to gain accurate attendance data on several past MTB events. the visual also offers for the 'Comparison' data relationship (Drake, Pytlarz, and Patel, 2018) to be compared between Local Authorities with the stacked nature allowing for more granular data comparisons with each region. The tooltips function was also utilised in this visual which meant the user could see the exact figures for that MTB event and corresponding Local Authority. The screenshot below represents the basis stacked bar chart as well as a filtered chart where the user has selected attendance data for the Forth Valley Region:

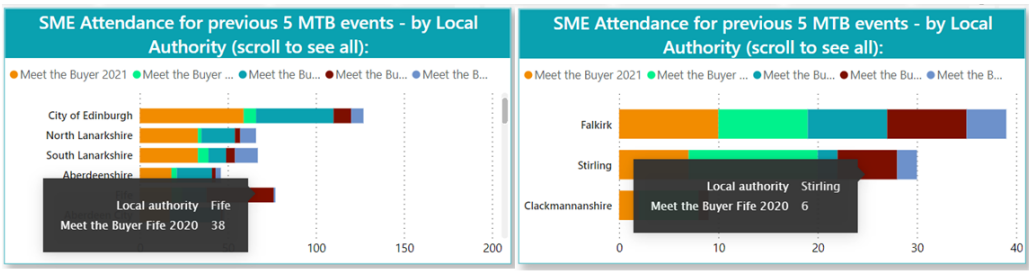


Figure 1.9. – MTB data visualisation (Stacked Bar chart)

Text Box – Supplier Comments:

This visual was used so it can represent 'Unstructured' data (Drake, Pytlarz, and Patel, 2018) and could filter a dense data relationship into just the highlights from Scottish supplier's perspective. The textual comments summarise a large amount of

information in the MTB datasets where open-ended questions were asked such as ‘Do you have any other comments?’. These sorts of questions are commonplace in organisational feedback surveys however the insights in the data they produce are rarely exploited (Decorte et al., 2019). The text box visual goes some way in trying to visualise and filter this unstructured data relationship effectively, using a scroll bar to allow the user to see all the comments:

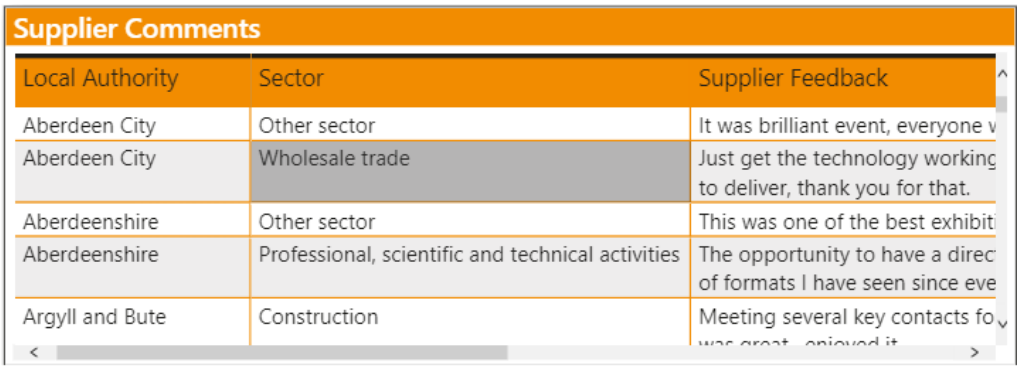


Figure 2.0. – MTB data visualisation (Text Box)

Tree Map – Sector Totals:

Figure 2.1. – MTB data visualisation (Tree Map)

The tree map was used on the dashboard to display data on all Local Authorities and the sector totals that contribute to the total SMEs registered with SDP in that corresponding region and sector (using data reported from SDP's CMS database). Using the tree map visualisation mean that all location and sector values could be displayed on one page of the dashboard as effectively as possible. It was also able to show the hierarchical 'Flow' of the data relationship (Drake, Pytlarz, and Patel, 2018) where the proportion of regions are reflected throughout the visual from the largest to the smallest Local Authority – in terms of total registered SMEs.

The flow of the data relationship between Local Authority and its sector totals could also be shown as in the filtered screenshot above where Fife has been selected by the user (using the ***10-DR*** slicer panel), this again shows the drilldown capacity of the dashboard where the proportion of the sector total is reflected in its size on the tree map visual. Tooltip functions let the user see a specific sector's exact figures if the word wrap function has cut them off. A useful function for this sector totals

visualisation is that certain sector specific organisations, SPT for example, can focus on data to their desired region and sector. In the case of SPT, this would be for Transport and Storage sectors.

Area Chart – SDP Annual Registrations:

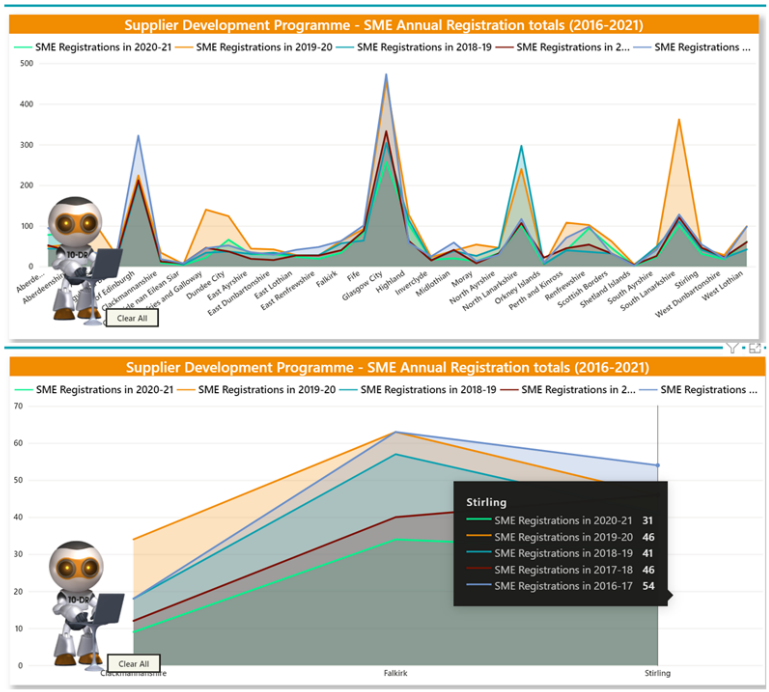


Figure 2.2. – MTB data visualisation (Area Chart)

Again, using the data sources available through the CMS data structure as well as SDP corporate documentation, such as annual reports, this visualisation aimed to provide insights into annual SME registration figures over a certain period. The area chart meant that data could be compared across all the Local Authorities. Ordinarily, line and area charts track data relationships, especially 'Comparisons' (Drake, Pytlarz, and Patel, 2018), over time however using the region values on the x-axis allowed for the data model to be visualised effectively. Users could then filter the data to certain regions and use the tooltip function to highlight the exact figures.

Cards & Gauge Visuals:

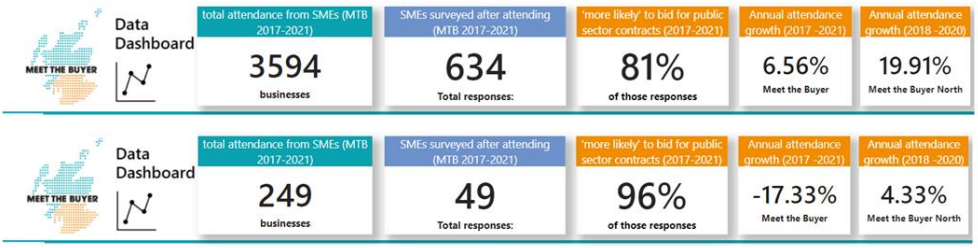


Figure 2.3. – MTB data visualisation (Card visuals)

These visualisations were used to show the headline figures from the MTB datasets and KPIs from the outcomes of MTB data analysis. 'Likelihood' to bid for public sector contracts was a question asked consistently across most of the surveys and was a main outcome that SDP analysed in historic MTB data, so this was included in these visuals. 'Annual attendance growth' was data that this project calculated using historic MTB and MTBN attendance data so this would provide SDP and their members with a better insight into how SMEs were interacting with MTB events. The above screenshot displays the basis headline figures in the dashboard as well as filtered data where a user has selected the Forth Valley region (Falkirk, Clackmannanshire and Stirling Local Authorities).

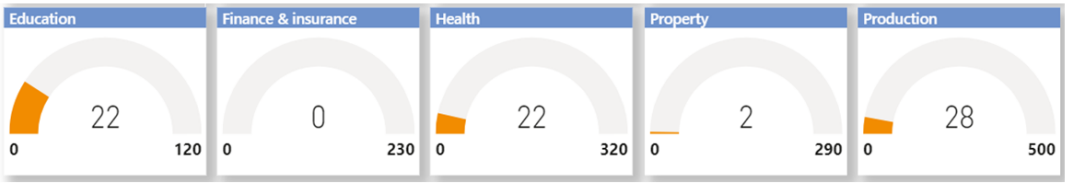


Figure 2.4. – MTB data visualisation (Gauge visuals)

The gauge visualisations were used on the last page of the dashboard for comparisons between SDP and ONS data on SME sector totals. These visuals provided the user with a hook on how much progress was being made against a target value – the ONS estimated sector total – and what progress was still be made and where to focus this progress with the fill value representing SDP's total SME count in that respective sector.

Scaffolding:

Another aspect of effective data visualisation that was focused upon in the data dashboard was the concept of 'Scaffolding' (Drake, Pytlarz, and Patel, 2018). This meant using appropriate data labels, counts on axes, formulas to display data accurately with any ambiguity from a user's perspective mitigated. Power BI does have a built-in function to automatically generate the most appropriate data counts for each visual but there were certain visuals which still needed bespoke data labelling as well as relevant metadata about the data model used:

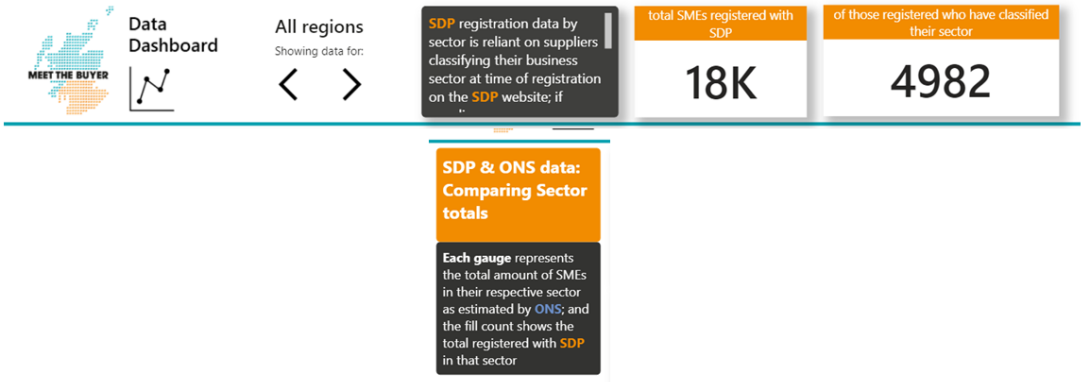


Figure 2.5. – MTB data visualisation (Scaffolding)

The screenshot above shows two aspects of 'scaffolding'; the first one clarifies how the sector totals has been visualised with the disparity between those who have classified their sector and those who have not shown in the card visuals. The one below is a text box explaining the gauge visuals and data used in the SDP and ONS comparison visual.

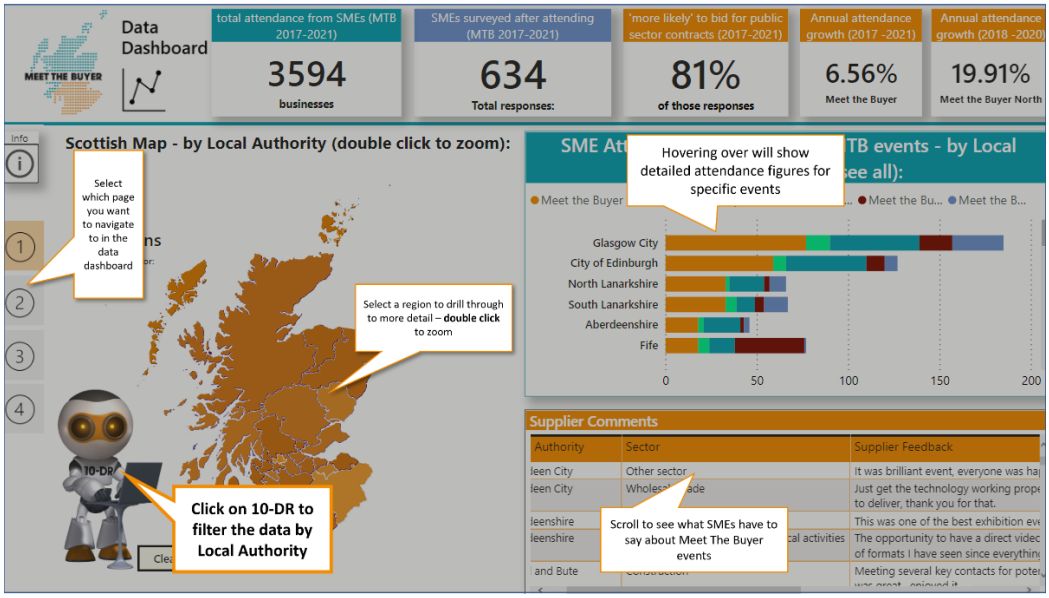


Figure 2.6. – MTB data visualisation (Info Panel)

The screenshot above displays the information panel that was incorporated in the data dashboard. The use for this is to aid navigation and interpretation for end users when first interacting with the data dashboard so any ambiguity or confusion on how it all works is mitigated.

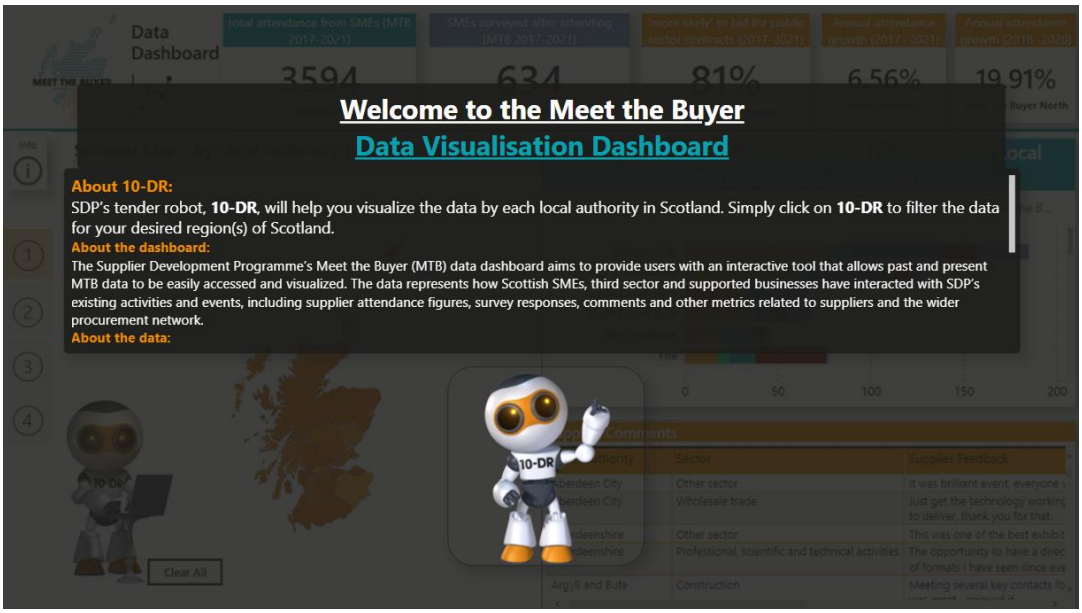


Figure 2.7. – MTB data visualisation (Intro Page)

Figure 2.7. is what users are presented with when first visiting the data dashboard. The aim of it is to clarify how users can interact with the visualisations, what datasets and sources are behind the visuals, and finally any nuances to the data that users should be aware of.

Other scaffolding features in the dashboard include the 'Clear All' button shown in *Figure 1.7.* when a user decides to clear all their filters on the data model, this can be easily achieved by clicking this button.

--- (The final MTB data dashboard can be accessed and interacted with here) ---

Power BI – Visualisation Tool

Microsoft's Power BI was chosen to create and implement the data dashboard as it was the most suitable tool to use. Power BI comprises two tools, the 'Desktop', a free local download where the data model and interactive data visual are created,

and the ‘Service’ which is accessed through a web browser where visuals prepared in the ‘Desktop’ can be viewed, managed, and edited in the workspace. The ‘Service’ also allows the user to share and collaborate on data visuals created (called ‘Reports’ in Power BI) as well as publish and embed a report such as the data dashboard into a webpage. Utilising this function meant that the MTB data dashboard created in the ‘Desktop’ could then be managed and updated from the ‘Service’ when it was published on SDP’s webpage. This crucially allowed the dashboard to stay up to date as data was maintained to keep the most relevant and recent data visualised which will continue to do so for future data on SDP’s events and activities.

Power BI is used extensively in the business world for Business Intelligence applications and highly customisable data visuals (Naneva and Stefanova, 2021) with Gartner also recognising Microsoft – its creators – as a ‘Magic Quadrant Leader’ in data analytics and Business Intelligence platforms (2020 Gartner Magic Quadrant | Power BI, 2021). Search hits on employment websites such as ‘Indeed’ and ‘SimplyHired’ have also showed that Power BI topped the charts with what data visual tool employers mention in their job descriptions (Diamond and Mattia, 2017). SFT as afore mentioned, work in a similar environment to SDP, have also implemented their own Power BI customised data visualisation.

Power BI is also an incredibly user-friendly interface and in comparison, to other prominent data visualisation tools such as Tableau, is far easier to start creating interactive data visuals. This factor was taken into consideration as this project took the form of a temporary industrial placement with the aim of developing innovative data visualisations to boost productivity for SDP so once this placement ended, the visual tool still had to be accessible to users who were not necessarily data scientists or analysts. The framework for the data visualisation was provided through Power BI for SDP so once the project had finished, the small team at SDP could develop, maintain, and update the data dashboard with ease for future data collected.

Discussion of Professional, Ethical, Legal, and Social Issues

The nature of this project meant that participants were not recruited as the data being analysed was existing data held under the auspices of SDP however, users were needed to participate in the project when responding to questionnaires about data and its uses as well as user experience of the data dashboard. These users were asked whether they agreed for their answers to be analysed for the purpose of this project before completing the survey.

Regarding the datasets that formed the basis for this project, certain actions were taken to deal with the issues relating to personal data collection. SDP when conducting their events 'Meet the Buyer' and 'Meet the Buyer North' took appropriate action upon collecting the data and agreed with the respondents that their data could be analysed for the proposed manner in this project. The datasets include details such as registration, attendance, and feedback from those businesses and organisations who are part of these events. SDP also hold these datasets securely in a local authority database and throughout this project, no data will leave SDP's servers when being analysed.

In terms of the data storage and analysis during this project. SDP will provide permission to use the datasets once I have been employed by them and I will then have access to this data for the duration of the project.

Finally, the datasets were only analysed in a manner that allowed for the creation of data visualisation models, the identification of any trends and patterns in the data, recommendations for data collection and storage for SDP in the future and any other aims in the project proposed by SDP.

This project strictly complied with the licensing terms and conditions for any of the digital technologies used to analyse the data. Any code written for the purposes of the project will be tested, documented throughout for clarity, and follow the British Computing Society code of conduct (BCS, THE CHARTERED INSTITUTE FOR IT, 2015).

All citations, software, and third-party code sources will be referenced properly as well.

Discussion and Evaluation

Introduction

As mentioned before, the aims of this study were to bring new analytical thinking to SDP and realise business value for both SDP as an organisation and their membership organisations. These aims were to be met through the project objectives which related to utilising various data science and analysis techniques including data collection, processing, and visualisation. In order to evaluate if this project and placement had met these aims and therefore be seen as a successful project, the data analysis techniques used have to be discussed and evaluated. The following chapter of this paper will look at how the review of SDP's data collection techniques was conducted, but it will also crucially evaluate the main part of this project – the user experience of the data visualisation implemented and delivered through SDP's website as this was the focus of the project conceived by SDP as well as the project title.

Data Collection – New MTB Data

After the review of SDP's data collection techniques had been undertaken and the additional questions had been formulated and implemented into the feedback survey, this new survey format was used for SDP's next national MTB event on 8th June 2021. This meant that those questions added and the responses to them would be fed back within the timeline of this placement so the insights into the data that this project had aimed to extract could be analysed. Using this data also provided the project with a chance to test the maintenance capabilities of the data dashboard as this new data would be updated to provide the data dashboard with the most recent and relevant data – essentially resulting in a test run for how the data model would be maintained for the data dashboard for future events once this project and placement with SDP had ended.

In terms of the additional questions added to the survey format, this data was analysed once the surveys had been completed and responses had been received.

Some of the insights into the data collected because of this project showed that the majority of Scottish SMEs had been adversely affected by the global pandemic – consistent with other research reviewed in this paper (Federation of Small Businesses, 2021). Further insights into COVID-19 related data found that most felt MTB had helped them find routes to recovery from the pandemic:

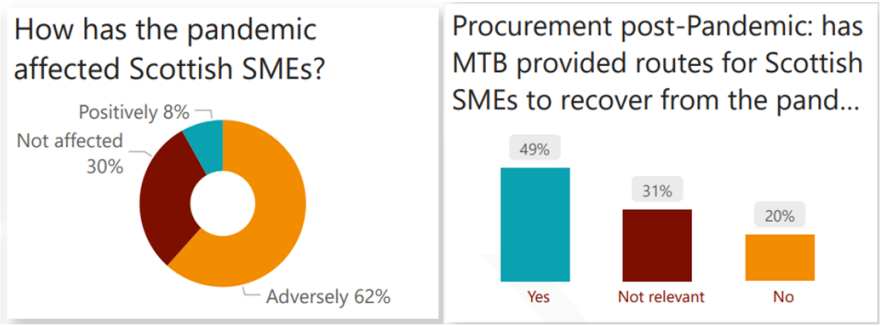


Figure 2.8. – MTB data visualisation (COVID-19)

This analysis allows SDP to compare their MTB data on SMEs with other organisations and other research which has collected similar data. Some of the analysis in the early stages as this project has only been able to collect data on the additional questions in the survey for only one MTB event. The overall aim, however, is to allow SDP to keep collecting data on these sorts of questions and track SMEs responses in future events as before this project was undertaken, SDP did not have this type of response data.

The Data Dashboard – User Experience

The most significant aspect of this project that had to be evaluated to assess the success of the project was the innovative data visualisation model developed and delivered through SDP's website. This evaluation took the form of a user experience questionnaire administered to recipients that would be the direct recipients and

benefactors of this project. These recipients were representatives from various SDP member organisations such as Fife council, SPT, South Lanarkshire Council, as well as SDP employees who tested out the data dashboard before completing the survey.

The objective of the questionnaire was to evaluate the **functionality** of the data dashboard – so the data dashboard can produce the data when requested, filtered, or sliced using PowerBI. Essentially to evaluate the ability of the dashboard to perform a task when requested by the user with no issues being encountered.

The other objective of the questions was to evaluate the response around the **usability** of the data dashboard – whether users feel the dashboard not only performs tasks without any bugs but also does it in a manner that is user-friendly when it comes to data visualisation and interface navigation.

And the final objective for evaluation of the dashboard was related to the **quality of data** that it produces – does the data produce the key insights to MTB that SDP and their members require. The questionnaire to assess user experience can be found in ***Appendix E.***

The six responses from the user experience questionnaire highlight how positive an experience it was for users as well as how the data dashboard satisfied the functionality, usability, and quality objectives that it set out to.

In terms of the **functionality** of the data dashboard – five out of six were 'very satisfied' with its ability to perform a task with no issues being encountered with the remaining response being 'somewhat satisfied'. For the **usability** question – half of respondents were 'very satisfied' with the ease of use of the interface. The other half were 'somewhat satisfied'. And nearly all were 'very satisfied' with the **quality of data** the dashboard displayed – five out of six with the remaining respondent being 'somewhat satisfied'.

Other questions asked related to the project objectives including whether the data dashboard was easy to navigate – four agreed and two strongly agreed that it was indeed easy to navigate. Half of respondents also strongly agreed that it had a user-friendly interface and had a pleasing colour scheme. Crucially, all respondents agreed that the data dashboard visualised the data effectively and accurately and their overall experience of the dashboard was either 'good' or 'excellent'.

Another significant insight from the user experience data showed that all respondents were 'very likely' to refer the dashboard to their membership organisations. This is a key response as the data dashboard was developed as part of this project's aim for SDP membership organisations to consume this data and to realise business value through the data analysis and visualisation of MTB datasets. Additional insights included how after conducting this data-driven initiative, all users strongly agreed that collection, analysis, and visualisation of data had the potential to be transformative for their future operations – consistent with research mentioned before in this report (International Data Group (IDG), 2016). The majority (four out of six) also said that the data dashboard would 'definitely' benefit their future activities within the public sector.

Finally, and perhaps most importantly, all respondents agreed that the MTB data dashboard was an innovative data visualisation model which is delivered through SDP's website – a key objective set out in this project.

Overall, the user experience of the MTB data dashboard was very positive with key responses showing that most of the necessary project objectives and requirements **(*see Appendix A)*** had been met and that this project could be deemed successful to an extent. There was an option for users to provide additional feedback and the consensus from these comments was in relation to certain visualisations and not the actual data model. For example, the Tree Map *(Figure 2.1.)* was considered a "bit busy" with other responses saying that the sector titles being cut off affected this visual and additional explanation would be useful. This sort of feedback is similar to the common pitfalls of data dashboard design (Rahman, 2017) where displaying

excessive details, specifically quantitative data, can slow down the end user's interpretation of the data visualisation. Other feedback however, complemented the SDP and ONS data visual as this would help with the immediate need to "promote" the Programme [SDP].

Future Work

Although this project did bring new analytical thinking to SDP and realised business value through data analysis and visualisation delivered by means of the MTB data dashboard, future work with SDP and their membership organisations can further benefit their use of data analysis techniques in their events and activities going forward.

One insight gained from the user experience survey that relates to SDP's potential future work is that all respondents agreed how data-driven projects will be key to collaboration with other public sector bodies. This sort of awareness is representative of a wider sense for data sharing and collaboration across the public sector with the UK Government implementing their National Data Strategy which aims for "more coordinated use of data across the wider public sector" (Department for Digital, Culture, Media & Sport, 2020). An example of this coordinated approach for SDP would be for future data-driven projects to investigate Public Contracts Scotland (PCS) contract awards data and collaborate with this organisation on their datasets in relation to how SDP's registered SMEs interact with PCS.

In terms of what this project aimed to achieve for SDP and the future data analysis work that is still to be realised. One of the project objectives set out at the start of this placement was to use data mining and machine learning algorithms on MTB datasets so any trends and patterns could be identified. This project did not achieve this objective in its timeline however as SDP keep using the revised MTB survey format implemented by this project, they can collect more consistent data over time with the potential for future data-driven projects to use more intensive data analysis techniques which can identify any trends and patterns.

Conclusion

SDP set out to bring new analytical thinking to their organisation which would help them implement sustainable and innovative improvements for future data analysis and visualisation. They also looked to realise the potential value of knowledge and business intelligence in their data using data analysis and visualisation of their Meet the Buyer datasets. In this project, objectives were devised which would meet these aims set out for SDP. These objectives meant utilising the relevant data science knowledge and techniques to firstly collect quality insightful and consistent data. It then looked to process this data and other data sources so the data analysis and crucially the data visualisation models could be developed, implemented, and delivered to SDP and their membership organisations through the format of a data dashboard on their website.

To summarise how this project achieved its aims and met those objectives, it is perhaps best done by revealing what SDP themselves thought of the project and how successful it was for their organisation – Jennifer Payne, the project's Industrial Supervisor encapsulated just how beneficial this project was in the following testimonial:

As a first-time host of a Data Lab MSc placement, the Supplier Development Programme set out to inject innovation into our existing attendance and feedback data for our largescale events, which were historically utilised in singular reporting – we wanted to do more. That is exactly what Cameron Scott brought to the virtual table! Cameron took a proactive approach to reviewing how we collect and manage data to understand our requirements and he enthusiastically identified a methodical route forward. Cameron used imagination to visualise our data on an interactive map of Scotland, systemically bringing more and more of our data sets out of easily forgotten spreadsheets into an interactive dashboard. Now embedded into our website for our public sector membership organisations to utilise on demand, the data visualisation dashboard will increase efficiency and improved workflow as our team will no longer have to knife-and-fork through the data. Cameron made a meaningful suggestion to implement an independent Office for National Statistics

(ONS) data source alongside our existing data, to give a high level benchmark on where there are estimated sectoral gaps. Given that the dashboard allows our membership organisations to drill down to a local level or scale up for a national outlook, the work completed by Cameron during this Data Lab placement and embedded within www.sdpscotland.co.uk will certainly be of use to all 32 Scottish local authorities in the period ahead. This Data Lab placement brought a timely new perspective on analytical thinking, as the Supplier Development Programme forever aims to implement sustainable, innovative improvements to bolster our productivity and outputs as a free national programme that aims to support Scottish SMEs, supported businesses and third sector organisations and contribute meaningfully to the overall procurement strategy for Scotland.

Bibliography

- 2001. Science, Technology and Industry Outlook [online]. 1st ed. OECD, p.99. Available at: http://www.oecd.org/innovation/research/1842070.pdf [Accessed 7 April 2021].
- Arora, E., 2011. KNOWLEDGE MANAGEMENT IN PUBLIC SECTOR. Journal of Arts Science & Commerce, [online] 2(1), pp.165-171. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.474.3542&rep=rep1&type=pdf [Accessed 8 April 2021].
- Arslanturk, S., Siadat, M., Ogunyemi, T., Killinger, K. and Diokno, A., 2015. Analysis of incomplete and inconsistent clinical survey data. Knowledge and Information Systems, [online] 46(3), pp.731-750. Available at: https://link.springer.com/content/pdf/10.1007/s10115-015-0850-7.pdf [Accessed 10 August 2021].
- Asian Productivity Organisation. 2020. Meaning and significance of data analytics for the public sector - Resources. [online] Available at: https://www.apo-tokyo.org/resources/articles/meaning-and-significance-of-data-analytics-for-the-public-sector/ [Accessed 9 August 2021].
- Azvine, B., Cui, Z. and Nauck, D.D., 2005. Towards real-time business intelligence. BT Technology Journal, [online] 23(3), pp.214-225. Available at: https://schep.home.xs4all.nl/bi/paper%203%20Towards%20real-time%20business%20intelligence.pdf [Accessed 13 August 2021].
- Baars, H. and Kemper, H., 2008. Management Support with Structured and Unstructured Data—An Integrated Business Intelligence Framework. Information Systems Management, [online] 25(2), pp.132-148. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.457.5544&rep=rep1&type=pdf [Accessed 13 August 2021].
- BBC, 2019. The avocado toast index: How many breakfasts to buy a house?. [image] Available at: https://www.bbc.com/worklife/article/20170530-the-avocado-toast-index-how-many-breakfasts-to-buy-a-house [Accessed 12 August 2021].
- BCS, THE CHARTERED INSTITUTE FOR IT, 2015. CODE OF CONDUCT FOR BCS MEMBERS. [online] Swindon: BCS, The Chartered Institute for IT. Available at: https://www.bcs.org/media/2211/bcs-code-of-conduct.pdf [Accessed 4 April 2021].
- Beer, D., 2017. Envisioning the power of data analytics. Information, Communication & Society, [online] 21(3), pp.465-479. Available at: https://www.tandfonline.com/doi/full/10.1080/1369118X.2017.1289232 [Accessed 9 August 2021].
- Bertram, D., 2007. Likert scales. Retrieved November, [online] 2(10). Available at: https://d1wqtxts1xzle7.cloudfront.net/34598027/topic-dane-likert-with-cover-page-v2.pdf?Expires=1628954142&Signature=ekWPAYIQ5dBW-XT9-nEVbLBI3MZKsw~bfvXvLs8N28jag7dWvhRsdAGJL971q~f5qT5LBdY5J-Ly~rxRhTbSZdA6myXvJoMVosDnqo9Nb1eWmqbMNzDPC7JgxOj3-qFsbFC9GISoUW7eNiPElP6Kd1LRVZwn0tG7Qln2EBWzXsQw0GSjNpTYe-zaTY7~5B0RuPepOOpubv3EJZR5Q6ys0f6riZZHmkx~RyjKlpbWCpQgU8N~1t~vs-

4~K3ReivG0Tn4sW6sqSn0yyGQVF0zL5QaWeiJRvdEL605h5JF-O0sfyBc5fcwJhy5upzph7ueR04bDtsmIqTKwTRx1Lvj4OQ__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA [Accessed 14 August 2021].
- Bezerra, A., Silva, I., Guedes, L., Silva, D., Leitão, G. and Saito, K., 2019. Extracting Value from Industrial Alarms and Events: A Data-Driven Approach Based on Exploratory Data Analysis. Sensors, [online] 19(12). Available at: https://www.mdpi.com/1424-8220/19/12/2772 [Accessed 11 August 2021].
- Big Data LDN & Qlik, 2019. The Public Sector Data Report 2019: HOW PUBLIC SECTOR ORGANISATIONS IN THE UK ARE USING DATA. [online] 3rd Street Group. Available at: https://bigdataldn.com/wp-content/uploads/2019/04/The-Public-Sector-Data-Report-2019.pdf [Accessed 9 August 2021].
- Blasius, J., Nenadic, O. and Thiessen, V., 2012. THE DIRTY DATA INDEX–ASSESSING THE QUALITY OF SURVEY DATA IN INTERNATIONAL COMPARISON. [online]. Available at: THE DIRTY DATA INDEX – ASSESSING THE QUALITY OF SURVEY DATA IN INTERNATIONAL COMPARISON (sa-ijas.org) [Accessed 8 August 2021].
- Brinkley, I., 2006. Defining the knowledge economy - Knowledge economy programme report. [online] London: The Work Foundation. Available at: https://www.knowledge4all.com/Temp/Files/9219fc8b-7263-416d-b3dc-a7dca118761f.pdf [Accessed 7 April 2021].
- Brint, S., 2001. Professionals and the `Knowledge Economy': Rethinking the Theory of Postindustrial Society. Current Sociology, [online] 49(4), pp.101-132. Available at: http://www.lib.csu.ru/ER/ER_Philosophy/All_Schreiber/Shreiber4/101.pdf [Accessed 7 April 2021].
- Brown, R., Rocha, A. and Cowling, M., 2020. Financing entrepreneurship in times of crisis: Exploring the impact of COVID-19 on the market for entrepreneurial finance in the United Kingdom. International Small Business Journal: Researching Entrepreneurship, [online] 38(5), pp.380-390. Available at: https://journals.sagepub.com/doi/pdf/10.1177/0266242620937464 [Accessed 10 August 2021].
- CAUSEweb. 2011. Quote: Sondergaard on Data Analytics. [online] Available at: https://www.causeweb.org/cause/resources/library/r2493 [Accessed 7 April 2021].
- Chai, C., 2020. The Importance of Data Cleaning: Three Visualization Examples. CHANCE, [online] 33(1), pp.4-9. Available at: https://www.researchgate.net/profile/Christine-Chai-2/publication/339221348_The_Importance_of_Data_Cleaning_Three_Visualization_Examples/links/5fd56fc845851553a0b1599e/The-Importance-of-Data-Cleaning-Three-Visualization-Examples.pdf [Accessed 10 August 2021].
- Clarke, T., 2001. The knowledge economy. Education+ Training [online]. Available at: https://www.emerald.com/insight/content/doi/10.1108/00400910110399184/full/html [Accessed 7 April 2021].
- Columbus, L., 2019. What Matters Most In Business Intelligence, 2019. [online] Forbes. Available at: https://www.forbes.com/sites/louiscolumbus/2019/06/09/what-matters-most-in-business-intelligence-2019/?sh=6c38b279702d [Accessed 13 August 2021].
- Cong, X. and Pandya, K., 2003. Issues of Knowledge Management in the Public Sector. Electronic Journal of Knowledge Management, [online] 1(2), pp.25-33. Available at:

https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1086.7370&rep=rep1&type=pdf [Accessed 8 April 2021].
- Data.spatialhub.scot. 2021. Local Authority Boundaries - Scotland - Local Authority Boundaries - Spatial Hub Scotland. [online] Available at: https://data.spatialhub.scot/dataset/local_authority_boundaries-is/resource/d24c5735-0f1c-4819-a6bd-dbfeb93bd8e4 [Accessed 15 August 2021].
- David, P. and Foray, D., 2003. Economic Fundamentals of the Knowledge Society. Policy Futures in Education, [online] 1(1), pp.20-49. Available at: https://journals.sagepub.com/doi/pdf/10.2304/pfie.2003.1.1.7 [Accessed 8 April 2021].
- Dean, A. and Kretschmer, M., 2007. Can Ideas be Capital? Factors of Production in the Postindustrial Economy: A Review and Critique. Academy of Management Review, [online] 32(2), pp.573-594. Available at: https://kar.kent.ac.uk/9145/1/38_Web_Version.pdf [Accessed 7 April 2021].
- Decorte, T., Malm, A., Sznitman, S., Hakkarainen, P., Barratt, M., Potter, G., Werse, B., Kamphausen, G., Lenton, S. and Asmussen Frank, V., 2019. The challenges and benefits of analyzing feedback comments in surveys: Lessons from a cross-national online survey of small-scale cannabis growers. Methodological Innovations, [online] 12(1), p.205979911982560. Available at: https://journals.sagepub.com/doi/pdf/10.1177/2059799119825606 [Accessed 12 August 2021].
- Department for Digital, Culture, Media & Sport, 2020. National Data Strategy. London: Cabinet Office. [online]. Available at: National Data Strategy - GOV.UK (www.gov.uk) [Accessed 9 August 2021].
- Diamond, M. and Mattia, A., 2017. Data visualization: An exploratory study into the software tools used by businesses. Journal of Instructional Pedagogies, [online] 18. Available at: Microsoft Word - 162386 (ed.gov) [Accessed 10 August 2021].
- Docs.microsoft.com. 2021. Publish to web from Power BI - Power BI. [online] Available at: https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-publish-to-web [Accessed 15 August 2021].
- Drake, B.M., Pytlarz, I. and Patel, M., 2018. Let me paint you a picture: Utilizing visualizations to make data more accessible. In Building capacity in institutional research and decision support in higher education [online] (pp. 81-93). Springer, Cham. Available at: Let Me Paint You a Picture: Utilizing Visualizations to Make Data More Accessible | SpringerLink [Accessed 10 August 2021].
- Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P., 1996. From data mining to knowledge discovery in databases. AI magazine, [online] 17(3), pp.37-37. Available at: https://ojs.aaai.org/index.php/aimagazine/article/view/1230 [Accessed 7 April 2021].
- Federation of Small Businesses, 2021. Finding Our Feet. [online] Federation of Small Businesses. Available at: https://www.fsb.org.uk/resource-report/finding-our-feet.html [Accessed 10 August 2021].
- Fraser of Allander Institute (FAI), 2020. The Impact of Covid-19 on Scottish Small and Medium-Sized Enterprises (SMEs): Prognosis and Policy Prescription. [online] FAI Publications. Available at: https://fraserofallander.org/the-impact-of-covid-19-on-scottish-small-and-medium-sized-enterprises-smes-prognosis-and-policy-prescription/ [Accessed 10 August 2021].

- Geotab. 2021. 6 Steps for data cleaning and why it matters | Geotab. [online] Available at: https://www.geotab.com/blog/data-cleaning/ [Accessed 10 August 2021].
- Ghani, S., 2009. Knowledge Management: Tools and Techniques. DESIDOC Journal of Library & Information Technology, [online] 29(6), pp.33-38. Available at: https://d1wqtxts1xzle7.cloudfront.net/35876680/276-680-1-PB.pdf?1418091165=&response-content-disposition=inline%3B+filename%3DKnowledge_Management_Tools_and_Technique.pdf&Expires=1617833261&Signature=fC~JqnAFE8YYv6FTDVSJYJdPV-jkDvz1pBob0gu6E4UZQQL2HRJ-egspgLgylKHRf2wwUZgYfPy2vE5XfeHiOL2uEU8p-J5jYxamVsUApWPsT82ust0eevsWhiI41ijuDDL8t5llkZvRTzCpOqX0JqlKqOX4oCo94yziJAstP3GyIrnxrhE1ecXGWJ-o5gK19i-RbLuuKvNhi-gQK3Wj5LTGesR2QR8GAGeBOz6OW-Yg-98I113wFUPOIxZIpkeixxydHxq4APbSYTmenptyZUcKSQedUwLVzJVauVMzkSHFpDruD8~2I~tr2BJ6UPYNfoCWQLbTPt-1xE0S7jcgzw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA [Accessed 7 April 2021].
- Giju, G.C., Badea, L., Ruiz, V.R.L. and Peña, D.N., 2010. Knowledge Management – the Key Resource in the Knowledge Economy. Theoretical and Applied Economics, [online] 17(6), pp.27-36. Available at: https://d1wqtxts1xzle7.cloudfront.net/43318604/473.pdf?1457025161=&response-content-disposition=inline%3B+filename%3DKnowledge_Management_the_Key_Resource_in.pdf&Expires=1617816488&Signature=EQcsQN1diZPe2EadaODdOJaJiFV4RmcNZS8tKw2L6lih2N6rdX7bd~nsYB3PyD4HuYd5s5g7B2-NyUdz4PfsQXvzmkixCKeaS42uSLhprfTtxF3dn2o0lqjPNS5ujLI-4-1nQen2s1iyQ7k68p2Chma8OXUWNnNOfJ7L1J80vFfatuAtgSDA68mRhI5004N9KtoNr5sD36j2HJSMWrot0wx2ZtTznAE3nRguVW32awEwTHbPjkuCYbsCn7jcfWa0VBa7fqtTTXID~dy1q4NHbjAH4vCxjxcNEMWmQva5KVaDIPSbteZ5lkQdu3eQUkUev8EvtlQHxVqbCzUvPSt8bQ__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA [Accessed 7 April 2021].
- Goodhue, D., Quillard, J. and Rockart, J., 1988. Managing the Data Resource: A Contingency Perspective. MIS Quarterly, [online] 12(3), p.373. Available at: https://www.jstor.org/stable/249204?seq=1 [Accessed 7 April 2021].
- Gromping, U., 2017. Data Visualisation with R: 100 Examples. Journal of Statistical Software, [online] 81(Book Review 2). Available at: https://scholar.googleusercontent.com/scholar?q=cache:L8a_k4RPs3oJ:scholar.google.com/+examples+of+bad+data+visualisation&hl=en&as_sdt=0,5 [Accessed 11 August 2021].
- Hadad, S., 2017. Knowledge economy: Characteristics and dimensions. Management dynamics in the Knowledge economy, 5(2), pp.203-225. Available at: https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=Hadad%2C+S.%2C+2017.+Knowledge+economy%3A+Characteristics+and+dimensions.+Management+dynamics+in+the+Knowledge+economy%2C+5%282%29%2C+pp.203-225.&btnG= [Accessed 7 April 2021]
- Hardy, J., Micek, G. and Capik, P., 2011. Upgrading Local Economies in Central and Eastern Europe? The Role of Business Service Foreign Direct Investment in the Knowledge Economy. European Planning Studies, [online] 19(9), pp.1581-1591.

Available at: http://shura.shu.ac.uk/5609/1/HardyMicekCapik-FullPaper-reviewed.pdf [Accessed 7 April 2021].
- Hariri, R., Fredericks, E. and Bowers, K., 2019. Uncertainty in big data analytics: survey, opportunities, and challenges. Journal of Big Data, [online] 6(1), pp.1-16. Available at: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0206-3 [Accessed 5 April 2021].
- Hernández, M. and Stolfo, S., 1995. The merge/purge problem for large databases. ACM SIGMOD Record, [online] 24(2), pp.127-138. Available at: https://link.springer.com/content/pdf/10.1023/A:1009761603038.pdf [Accessed 10 August 2021].
- Hirsch, D., 2014. The Glass House Effect: Big Data, The New Oil, and the Power of Analogy. Maine Law Review, [online] 66(2), p.374. Available at: https://digitalcommons.mainelaw.maine.edu/cgi/viewcontent.cgi?article=1088&context=mlr [Accessed 7 April 2021].
- Howell, S., Lerner, J., Nanda, R. and Townsend, R., 2020. Financial Distancing: How Venture Capital Follows the Economy Down and Curtails Innovation. SSRN Electronic Journal, [online] No. w27150. Available at: https://www.nber.org/system/files/working_papers/w27150/w27150.pdf [Accessed 10 August 2021].
- Humby, C., 2006. Data is the new oil. Proc. ANA Sr. Marketer's Summit. Evanston, IL, USA. [online]. Available at. https://pdfs.semanticscholar.org/8c46/0595fa4fb69e19e9e5afa7b9ffa31fa896f2.pdf [Accessed 5 April 2021]
- Ichijo, K. and Nonaka, I., 2007. Knowledge Creation and Management: New Challenges for Managers. [ebook] New York: Oxford University Press. Available at: http://pustaka.unp.ac.id/file/abstrak_kki/EBOOKS/KNOWLEDGE%20MANAGEMENT%20Knowledge%20Creation%20and%20Management%20New%20Challenges%20for%20Managers.pdf [Accessed 7 April 2021].
- Info.microsoft.com. 2021. 2020 Gartner Magic Quadrant | Power BI. [online] Available at: https://info.microsoft.com/ww-landing-2020-gartner-magic-quadrant-for-analytics-and-business-intelligence.html?LCID=EN-US&ls=Website [Accessed 15 August 2021].
- International Data Group (IDG), 2016. Data & Analytics Survey. [online] International Data Group, pp.1-7. Available at: https://cdn2.hubspot.net/hubfs/1624046/IDGE_Data_Analysis_2016_final.pdf [Accessed 5 April 2021].
- Itpro.co.uk. 2008. Firms full of dirty data. [online] Available at: https://www.itpro.co.uk/609057/firms-full-of-dirty-data%20 [Accessed 10 August 2021].
- Jenn, N.C., 2006. Designing a questionnaire. Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia, [online] 1(1), p.32. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4797036/ [Accessed 7 August 2021].
- Jupyter.org. 2021. Project Jupyter. [online] Available at: https://jupyter.org/ [Accessed 14 August 2021].

- Kenway Consulting. 2020. Is Data Really "The New Oil"?. [online] Available at: https://www.kenwayconsulting.com/blog/data-is-the-new-oil/ [Accessed 7 April 2021].
- Kozak, M., Hartley, J., Wnuk, A. and Tartanus, M., 2015. Multiple Pie Charts: Unreadable, Inefficient, and Over-Used. Journal of Scholarly Publishing, [online] 46(3), pp.282-289. Available at: https://www.researchgate.net/publication/277648324_Multiple_Pie_Charts_Unreadable_Inefficient_and_Over-Used [Accessed 12 August 2021].
- Longitude. 2021. Data visualisation: The pictures that are worth a thousand words | Longitude. [online] Available at: https://www.longitude.co.uk/data-visualisation-post/ [Accessed 10 August 2021].
- Malhotra, N.K., 2006. Questionnaire design and scale development. The handbook of marketing research: Uses, misuses, and future advances, [online] pp.83-94. Available at: https://www.researchgate.net/profile/Naresh-Malhotra/publication/266864633_Questionnaire_design_and_scale_development/links/566708fe08ae34c89a0220f9/Questionnaire-design-and-scale-development.pdf [Accessed 14 August 2021].
- McColl, E., Jacoby, A., Thomas, L., Soutter, J., Bamford, C., Steen, N., Thomas, R., Harvey, E., Garratt, A. and Bond, J., 2001. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. Health Technology Assessment, [online] 5(31). Available at: https://www.researchgate.net/publication/11550481_Design_and_use_of_questionnaires_a_review_of_best_practice_applicable_to_surveys_of_Health_Service_staff_and_patients [Accessed 10 August 2021].
- Medium. 2021. Geospatial Data Visualization. [online] Available at: https://neenapriyanka.medium.com/geospatial-data-visualization-b81f6e18f8fd [Accessed 15 August 2021].
- Medium. 2020. Is Data Really the New Oil in the 21st Century?. [online] Available at: https://towardsdatascience.com/is-data-really-the-new-oil-in-the-21st-century-17d014811b88 [Accessed 7 April 2021].
- Medium. 2016. "Data is the New Oil" — A Ludicrous Proposition. [online] Available at: https://medium.com/project-2030/data-is-the-new-oil-a-ludicrous-proposition-1d91bba4f294 [Accessed 7 April 2021].
- Moretto, A., Ronchi, S. and Patrucco, A., 2017. Increasing the effectiveness of procurement decisions: The value of big data in the procurement process. International Journal of RF Technologies, [online] 8(3), pp.79-103. Available at: https://www.researchgate.net/profile/Andrea-Stefano-Patrucco/publication/320120765_Increasing_the_effectiveness_of_procurement_decisions_The_value_of_big_data_in_the_procurement_process/links/59d3d802aca2721f436ce184/Increasing-the-effectiveness-of-procurement-decisions-The-value-of-big-data-in-the-procurement-process.pdf [Accessed 8 April 2021].
- Nair, B.C. 2019. Business Analytics–Leveraging The Power of Data. In MESSAGE FROM THE CONFERENCE CHAIRS [online] (pp. 9). Available at: Value-Chain-Management-of-Chilli-An-Empirical-Study.pdf (researchgate.net) [Accessed 9 August 2021].
- Naneva, V. and Stefanova, K., 2021, March. A comparative analysis of popular technologies for developing custom visuals. In AIP Conference Proceedings (Vol.

2333, No. 1, p. 070009) [online]. AIP Publishing LLC. Available at: https://aip.scitation.org/doi/pdf/10.1063/5.0042201 [Accessed 11 August 2021].
- Negash, S. and Gray, P., 2008. Business intelligence. In Handbook on decision support systems 2 [online] (pp. 175-193). Springer, Berlin, Heidelberg. Available at: https://core.ac.uk/download/pdf/301376512.pdf [Accessed 13 August 2021]
- Nolin, J., 2019. Data as oil, infrastructure or asset? Three metaphors of data as economic value. Journal of Information, Communication and Ethics in Society, [online] 18(1), pp.28-43. Available at: https://www.researchgate.net/publication/337055040_Data_as_oil_infrastructure_or_asset_Three_metaphors_of_data_as_economic_value [Accessed 7 April 2021].
- Old Street Solutions, 2021. When Data Visualization Really Isn't Useful (and When It Is). [image] Available at: https://www.oldstreetsolutions.com/good-and-bad-data-visualization [Accessed 12 August 2021].
- Ong, I., Siew, P. and Wong, S., 2011. A Five-Layered Business Intelligence Architecture. Communications of the IBIMA, [online] pp.1-11. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.642.4964&rep=rep1&type=pdf [Accessed 13 August 2021].
- Ons.gov.uk. 2021. Small and medium enterprises in county by industry and size - Office for National Statistics. [online] Available at: https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/adhocs/13286smallandmediumenterprisesincountybyindustryandsize [Accessed 14 August 2021].
- Pandas.pydata.org. 2021. pandas - Python Data Analysis Library. [online] Available at: https://pandas.pydata.org/about/ [Accessed 14 August 2021].
- PC Mag UK, 2019. The Best Data Visualization Tools for 2020. [image] Available at: https://uk.pcmag.com/cloud-services/83744/the-best-data-visualization-tools-for-2020 [Accessed 12 August 2021].
- Powerbi.microsoft.com. 2021. Partner showcase | Microsoft Power BI. [online] Available at: https://powerbi.microsoft.com/en-ca/partner-showcase/bitscape-mutual-fund-analytics-for-investment-management-organisation/ [Accessed 10 August 2021].
- Pucci, P., Raes, L. and Mareels, G., 2021. The Data Shake: Opportunities and Obstacles for Urban Policy Making [online] (p. 125). Springer Nature. Available at: https://library.oapen.org/bitstream/handle/20.500.12657/47327/9783030636937.pdf?sequence=1 [Accessed 13 August 2021].
- Python.org. 2021. Download Python. [online] Available at: https://www.python.org/downloads/ [Accessed 14 August 2021].
- Qualtrics. 2021. How to Increase the Response Rate on Surveys | Qualtrics. [online] Available at: https://www.qualtrics.com/uk/experience-management/research/improve-survey-response/?rid=ip&prevsite=en&newsite=uk&geo=GB&geomatch=uk [Accessed 13 August 2021].
- Rahman, A., 2017. Designing a Dashboard as Geo-Visual Exploration Tool for Origin-Destination Data (Master's thesis, University of Twente). Available at: rahman.pdf (utwente.nl) [Accessed 12 August 2021].
- Schmidt, J., 2020, February. Usage of Visualization Techniques in Data Science Workflows. In VISIGRAPP (3: IVAPP) [online] (pp. 309-316). Available at:

https://www.researchgate.net/profile/Johanna-Schmidt-15/publication/340047400_Usage_of_Visualization_Techniques_in_Data_Science_Workflows/links/5eaaf3cda6fdcc70509dc5ce/Usage-of-Visualization-Techniques-in-Data-Science-Workflows.pdf [Accessed 10 August 2021].

- Scottishfuturestrust.org.uk. 2021. Media - Scottish Futures Trust. [online] Available at: https://www.scottishfuturestrust.org.uk/media/online-forecast-tool-launched-to-support-investment-and-jobs-within-construction-sector [Accessed 10 August 2021].
- Scottish Government. 2021. Public sector procurement - gov.scot. [online] Available at: https://www.gov.scot/policies/public-sector-procurement/ [Accessed 10 August 2021].
- Shahul, A., 2018. Data is the New Oil. Right? - Digital Transformation. [online] Digital Transformation. Available at: https://www.ibm.com/blogs/digital-transformation/in-en/blog/data-is-the-new-oil-right/ [Accessed 7 April 2021].
- Smith, K.H., 2002. What is the'Knowledge Economy'? Knowledge intensity and distributed knowledge bases [online]. Available at: https://eprints.utas.edu.au/1235/1/2002-6.pdf [Accessed 7 April 2021].
- Synoptic.design. 2021. Synoptic Designer for Power BI. [online] Available at: https://synoptic.design/ [Accessed 15 August 2021].
- Supplier Development Programme, 2021. Supplier Development Programme Meet the Buyer Data Visualisation
- Supplier Development Programme. 2021. Why register with SDP?. [online] Available at: https://www.sdpscotland.co.uk/tender-why-register-with-sdp/#rlslider_3 [Accessed 9 August 2021].
- Synodinos, N., 2003. The “art” of questionnaire construction: some important considerations for manufacturing studies. Integrated Manufacturing Systems, [online] 14(3), pp.221-237. Available at: http://www.adolphus.me.uk/emx/surveys/synodinos_files/p221.htm [Accessed 7 August 2021].
- The Economist. 2010. Data, data everywhere. [online] Available at: https://www.economist.com/special-report/2010/02/27/data-data-everywhere [Accessed 7 April 2021].
- tumblr, 2021. Bad Visualisations. [image] Available at: https://badvisualisations.tumblr.com/ [Accessed 12 August 2021].
- United Nations. 2021. Coronavirus disease (COVID-19) | From the UN Secretary-General | United Nations. [online] United Nations. Available at: https://www.un.org/en/coronavirus/un-secretary-general [Accessed 9 August 2021].
- Watson, H., 2009. Tutorial: Business Intelligence – Past, Present, and Future. Communications of the Association for Information Systems, [online] 25(1). Available at: https://www.researchgate.net/profile/Hugh-Watson-2/publication/220889689_Business_Intelligence_Past_Present_and_Future/links/5767e55f08ae421c448dc8df/Business-Intelligence-Past-Present-and-Future.pdf [Accessed 13 August 2021].
- Wilkinson, L., 2012. The grammar of graphics. In Handbook of computational statistics [online] (pp. 375-414). Springer, Berlin, Heidelberg. Available at: https://link.springer.com/chapter/10.1007/978-3-642-21551-3_13 [Accessed 11 August 2021].

APPENDICES

APPENDIX A: Project Requirements Analysis (MoSCoW) & Project Stakeholders

Must	**Should**
- Analyse datasets 'Meet the Buyer' & 'Meet the Buyer North' - Develop data visualisation models – delivered through SDP's website – that are innovative and allow SDP's members to assess and interact with data from SDP's existing activities such as Meet the Buyer - Review data collection and management techniques (feedback surveys) currently used by SDP - Develop an implementation plan that will improve the consistency of collection of feedback data as well as providing a framework for automating the process of managing, visualising and evaluating data going forward - Perform data pre-processing techniques on the raw datasets in question so value can be extracted from them in the form of trends and patterns identified in the processed data - Showcase business value for SDP's external stakeholders (businesses/suppliers)	- Produce value for SDP to utilise in their future operations and events such as in their communications and marketing strategies to raise SDP's profile - Using the insights gained from analysing the Meet the Buyer datasets to show what opportunities/market gaps are available to SDP and their members for future activities within procurement - Demonstrate the effectiveness of the SDP programme to the overall procurement network - Introduce new analytical thinking to SDP that will develop robust and sustainable data analysis capabilities so SDP can better assess feedback data from future activities

Could	**Won't**
- Procure new businesses/suppliers and members to invest in SDP	- Benefit the overall procurement process itself as SDP develop their operations within procurement both in the public and private sector

Stakeholder	Needs
SDP (internal stakeholder)	Data analysis and visualisation of several past datasets for further insight to be gained about the environment they operate in. Recommendations and implementation of value realised from data analysis. New analytical thinking for the organisation which can improve and bolster productivity sustainably in the future.
Local authorities/Businesses/Suppliers (external stakeholders) – SDP Membership Organisations	Realise business value that SDP bring to the overall procurement strategy through their support in all aspects of tendering to businesses
Project Supervisor	Well-considered project which meets the criteria for an MSc project
Myself	Successful project which satisfies all other stakeholders

APPENDIX B: Risk Assessment

Risk	Likelihood	Impact	Contingency
Datasets insufficient for relevant data analysis	Low	High	Seek alternate datasets available. Use subset of data.
IT Equipment incapable of conducting necessary data analysis	Low	High	Notify supervisor and SDP so alternative IT equipment can be used
Licenses for software programs/products used in data analysis not financially viable to implement in a business environment (eg. Public sector)	Medium	High	Use alternative open-source and free software to conduct data analysis/develop data visualisation models
No trends or patterns of value identified within the datasets	Medium	Medium	Seek alternative insights produced by other data analysis techniques
Project falls behind planned schedule	Medium	Low	Change project plan to allocate more time to certain tasks in project
SDP's servers/CMS/website is compromised by data breaches	Low	High	Notify supervisor of impact on project and seek alternate methods to complete project

APPENDIX C: Reviewed MTB Post-Event Feedback Survey – Additional Questions

1. Before attending Meet the Buyer, how would you rate your knowledge of the tendering process?
    - Comprehensive
    - Sufficient
    - Limited
    - No knowledge at all

2. How has your knowledge of the tendering process been affected after attending Meet the Buyer?
    - Improved a great degree
    - Improved to some degree
    - Stayed the same
    - Decreased to some degree
    - Decreased a great degree
3. How confident are you about bidding for public sector contracts after attending this event?
    - Totally confident
    - Very confident
    - Neutral
    - Not very confident
    - Not confident at all

4. How important do you feel these events are for helping SMEs, charities and third sector organisations learn more about upcoming contracts and understand the tendering process?
    - 4 – Absolutely essential
    - 3 – Very important
    - 2 – Of average importance
    - 1 – Of little importance
    - 0 – Not important at all

5. How important do you feel these events are for connecting public sector buyers and decision makers to SMEs, charities, and third sector organisations?
    - 4 – Absolutely essential
    - 3 – Very important
    - 2 – Of average importance
    - 1 – Of little importance
    - 0 – Not important at all

6. Public bodies must think about how procurement will improve:
        - the economy
        - society
        - the environment

    How important do you feel Meet the Buyer is for putting into practice the National Performance Framework - part of which is encouraging better procurement practices.

    - 4 – Absolutely essential
    - 3 – Very important
    - 2 – Of average importance
    - 1 – Of little importance
    - 0 – Not important at all

7. From your experience at Meet the Buyer, do you feel Scotland's public sector is using its spending power in a way that is good for business?
    - YES
    - NO

- NOT SURE

Do you have any comments about your response: [LONG TEXT BOX]

8. How many new business contacts did you make by networking with buyers/exhibitors at Meet the Buyer?
    - None
    - Less than 5
    - Between 6-9
    - Between 10-19
    - More than 20

9. Do you believe these contacts developed have the potential to benefit your business within the next year? (i.e. Bidding for future contracts or being invited to Quick Quote.)
    - Definitely
    - Probably
    - Not sure
    - Probably not
    - Definitely not
    - NOT RELEVANT

10. Before attending Meet the Buyer, how would you rate your knowledge of the opportunities available through the Public Contracts Scotland (PCS) portal?
    - Comprehensive
    - Sufficient
    - Limited
    - No knowledge at all

11. How has your knowledge of the Public Contracts Scotland (PCS) portal been affected after attending Meet the Buyer?
    - Improved a great degree
    - Improved to some degree
    - Stayed the same
    - Decreased to some degree
    - Decreased a great degree

12. After attending Meet the Buyer, have you identified any significant barriers to you as a business in bidding for public contracts?
    - YES
    - NO
        i. If yes, can you elaborate on what these barriers are: [BIG TEXT BOX]

13. Have you had any successful bids for public contracts after participating in previous Meet the Buyer events?
    - YES
    - NO
        i. If yes, could you tell us about the experience: [TEXT BOX]

14. How has your business been impacted been affected by the COVID-19 pandemic?
    - Adversely affected
    - Not affected
    - Positively affected

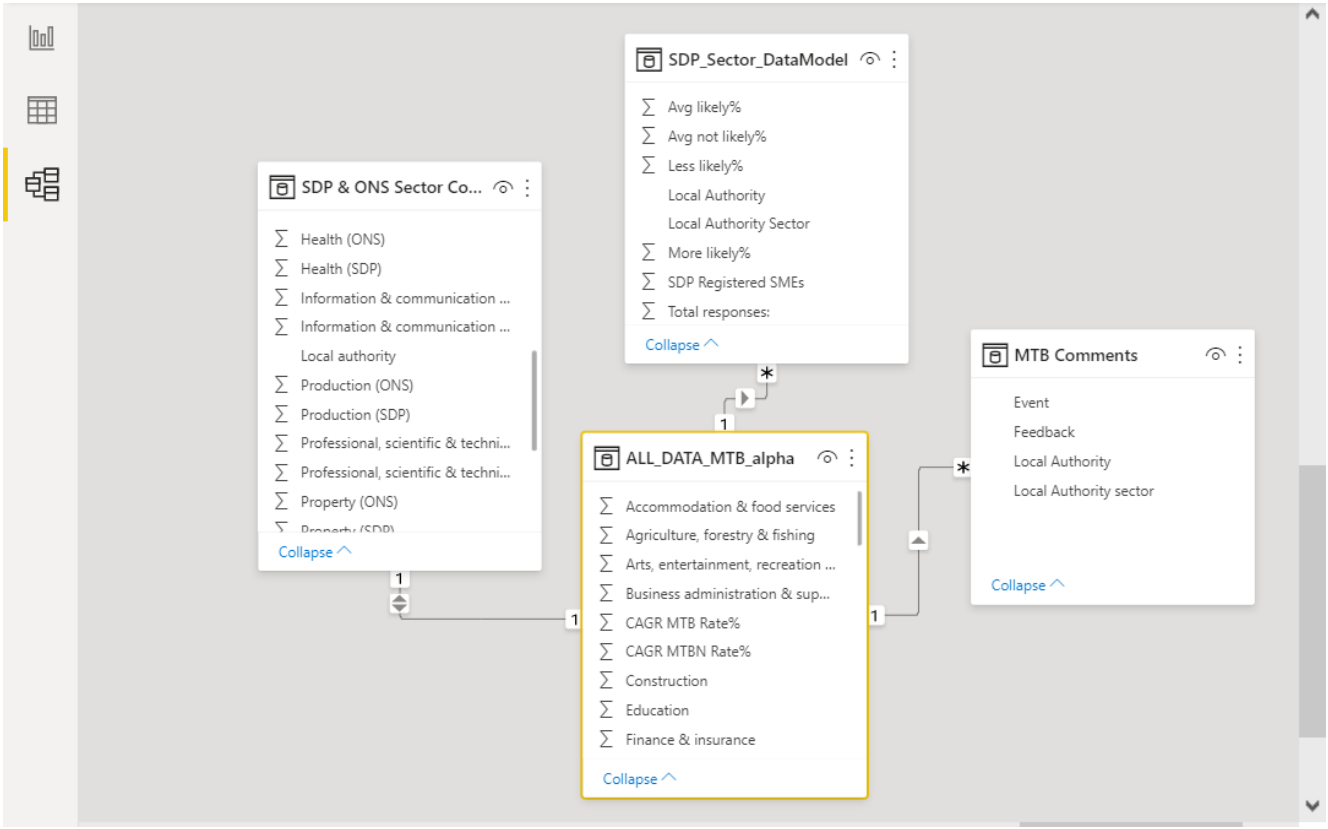
15. Do you feel the Meet the Buyer event helped move your business towards providing you with additional routes to recover from the impact of the COVID-19 pandemic?
    - YES
    - NO

- NOT RELEVANT

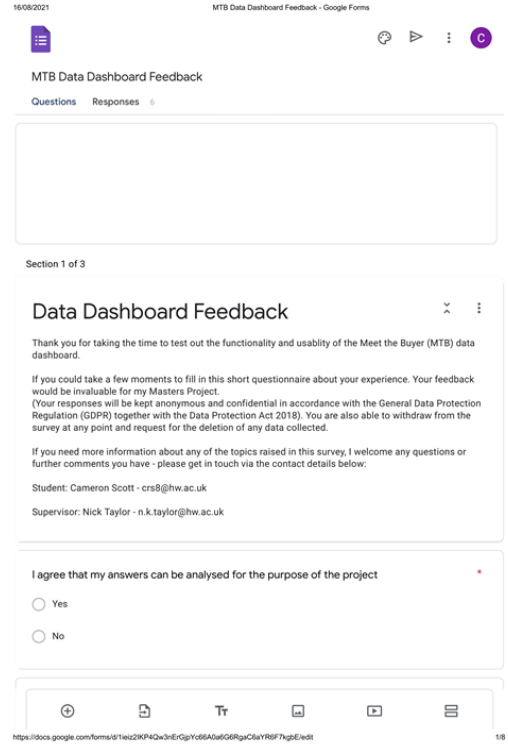
16. **How many employees do you have?** Choose one
    1 (Sole Trader)
    2-9 (Micro)
    10-49 (Small)
    50-249 (Medium)
    250+ (Large)
    I am a charity or supported business.
17. **Where is your business based? (Make this a DROP DOWN BOX)**
    - Aberdeen City
    - Aberdeenshire
    - Angus
    - Argyll and Bute
    - Clackmannanshire
    - Comhairle nan Eilean Siar
    - Dumfries and Galloway
    - Dundee City
    - East Ayrshire
    - East Dunbartonshire
    - East Lothian
    - East Renfrewshire
    - City of Edinburgh
    - Falkirk
    - Fife
    - Glasgow City
    - Highland
    - Inverclyde
    - Midlothian
    - Moray
    - North Ayrshire
    - North Lanarkshire
    - Orkney Islands
    - Perth and Kinross
    - Renfrewshire
    - Scottish Borders
    - Shetland Islands
    - South Ayrshire
    - South Lanarkshire
    - Stirling
    - West Dunbartonshire
    - West Lothian
    - Other – England
    - Other – Wales
    - Other – Northern Ireland
    - Other – Non UK

18. Are you more likely to bid for public contracts, having participated in this event?
    - Definitely
    - Very likely
    - Possibly
    - Unlikely
    - Definitely would not

APPENDIX D: MTB Data Visualisation – Data Model



APPENDIX E: Data Dashboard User Experience Questionnaire


16/08/2021 MTB Data Dashboard Feedback - Google Forms

MTB Data Dashboard Feedback

Questions Responses 6

Section 1 of 3

Data Dashboard Feedback

Thank you for taking the time to test out the functionality and usablity of the Meet the Buyer (MTB) data dashboard.

If you could take a few moments to fill in this short questionnaire about your experience. Your feedback would be invaluable for my Masters Project.
(Your responses will be kept anonymous and confidential in accordance with the General Data Protection Regulation (GDPR) together with the Data Protection Act 2018). You are also able to withdraw from the survey at any point and request for the deletion of any data collected.

If you need more information about any of the topics raised in this survey, I welcome any questions or further comments you have - please get in touch via the contact details below:

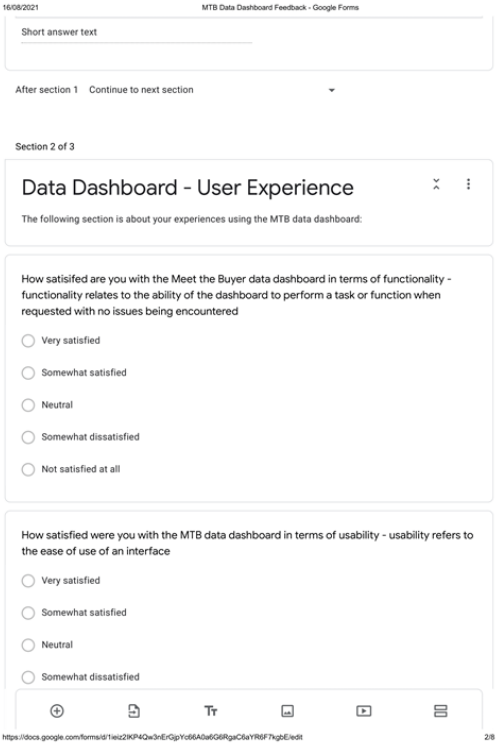
Student: Cameron Scott - crs8@hw.ac.uk

Supervisor: Nick Taylor - n.k.taylor@hw.ac.uk

I agree that my answers can be analysed for the purpose of the project *

- Yes
- No

https://docs.google.com/forms/d/1ieiz2IKP4Qw3nErGjpYc66A0a6G6RgaC6aYR6F7kgbE/edit 1/8


16/08/2021 MTB Data Dashboard Feedback - Google Forms

Short answer text

After section 1 Continue to next section

Section 2 of 3

Data Dashboard - User Experience

The following section is about your experiences using the MTB data dashboard:

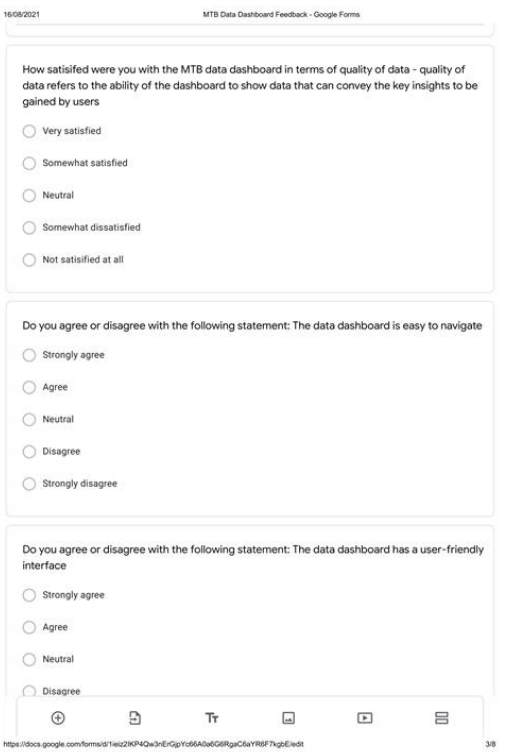
How satisifed are you with the Meet the Buyer data dashboard in terms of functionality - functionality relates to the ability of the dashboard to perform a task or function when requested with no issues being encountered

- Very satisfied
- Somewhat satisfied
- Neutral
- Somewhat dissatisfied
- Not satisfied at all

How satisfied were you with the MTB data dashboard in terms of usability - usability refers to the ease of use of an interface

- Very satisfied
- Somewhat satisfied
- Neutral
- Somewhat dissatisfied

https://docs.google.com/forms/d/1ieiz2IKP4Qw3nErGjpYc66A0a6G6RgaC6aYR6F7kgbE/edit 2/8


16/08/2021 MTB Data Dashboard Feedback - Google Forms

How satisifed were you with the MTB data dashboard in terms of quality of data - quality of data refers to the ability of the dashboard to show data that can convey the key insights to be gained by users

- Very satisfied
- Somewhat satisfied
- Neutral
- Somewhat dissatisfied
- Not satisified at all

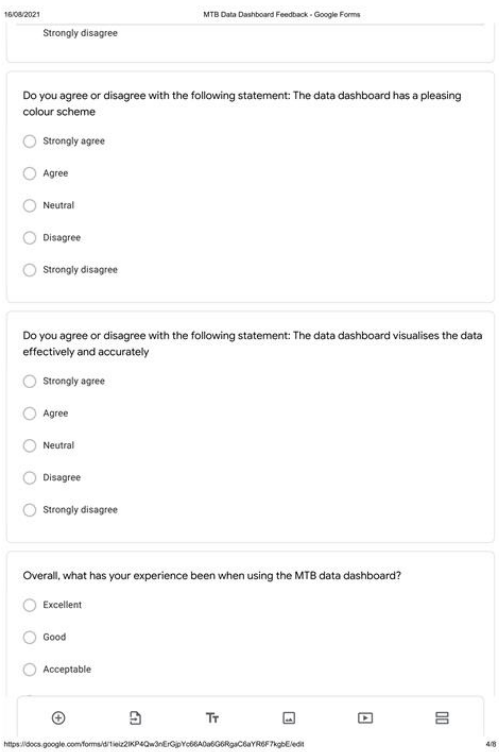
Do you agree or disagree with the following statement: The data dashboard is easy to navigate

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Do you agree or disagree with the following statement: The data dashboard has a user-friendly interface

- Strongly agree
- Agree
- Neutral
- Disagree

https://docs.google.com/forms/d/1ieiz2IKP4Qw3nErGjpYc66A0a6G6RgaC6aYR6F7kgbE/edit 3/8


16/08/2021 MTB Data Dashboard Feedback - Google Forms

Strongly disagree

Do you agree or disagree with the following statement: The data dashboard has a pleasing colour scheme

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

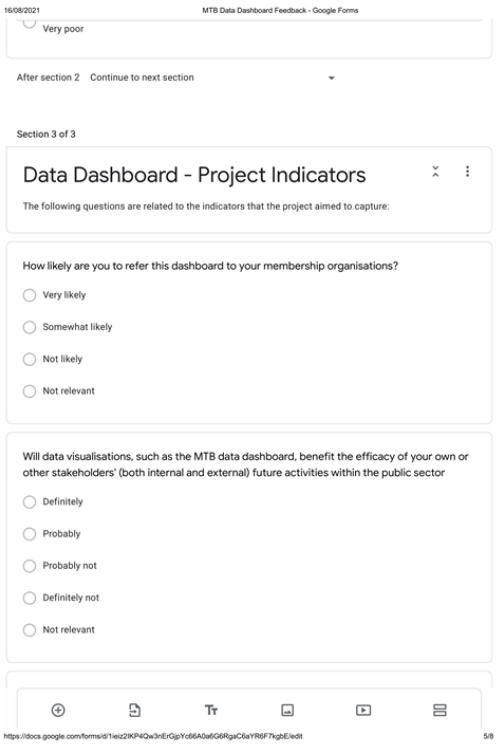
Do you agree or disagree with the following statement: The data dashboard visualises the data effectively and accurately

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Overall, what has your experience been when using the MTB data dashboard?

- Excellent
- Good
- Acceptable

https://docs.google.com/forms/d/1ieiz2IKP4Qw3nErGjpYc66A0a6G6RgaC6aYR6F7kgbE/edit 4/8



Very poor

After section 2 Continue to next section

Section 3 of 3

Data Dashboard - Project Indicators

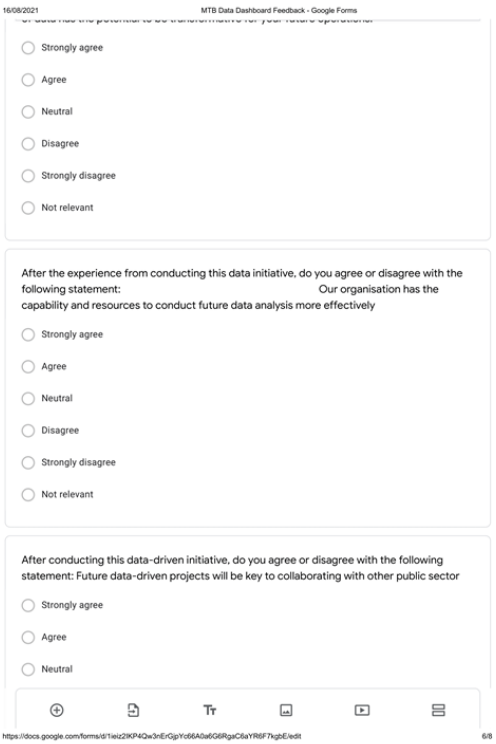
The following questions are related to the indicators that the project aimed to capture:

How likely are you to refer this dashboard to your membership organisations?

- Very likely
- Somewhat likely
- Not likely
- Not relevant

Will data visualisations, such as the MTB data dashboard, benefit the efficacy of your own or other stakeholders' (both internal and external) future activities within the public sector

- Definitely
- Probably
- Probably not
- Definitely not
- Not relevant



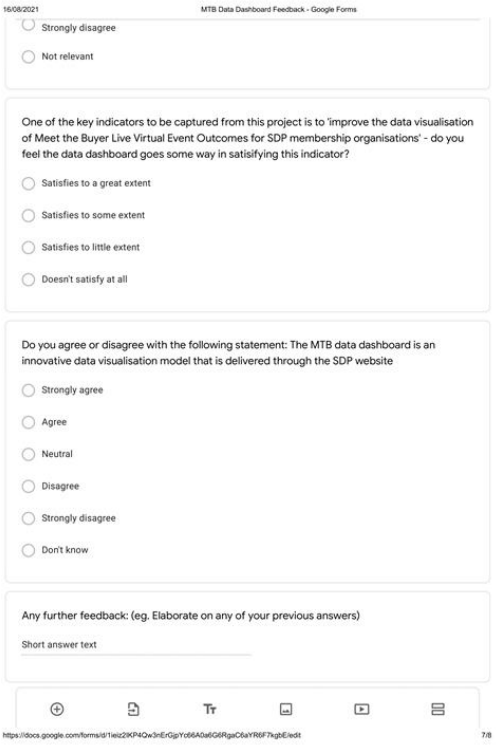
- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree
- Not relevant

After the experience from conducting this data initiative, do you agree or disagree with the following statement: Our organisation has the capability and resources to conduct future data analysis more effectively

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree
- Not relevant

After conducting this data-driven initiative, do you agree or disagree with the following statement: Future data-driven projects will be key to collaborating with other public sector

- Strongly agree
- Agree
- Neutral



- Strongly disagree
- Not relevant

One of the key indicators to be captured from this project is to 'improve the data visualisation of Meet the Buyer Live Virtual Event Outcomes for SDP membership organisations' - do you feel the data dashboard goes some way in satisifying this indicator?

- Satisfies to a great extent
- Satisfies to some extent
- Satisfies to little extent
- Doesn't satisfy at all

Do you agree or disagree with the following statement: The MTB data dashboard is an innovative data visualisation model that is delivered through the SDP website

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree
- Don't know

Any further feedback: (eg. Elaborate on any of your previous answers)

Short answer text