Gender-based Differences in Managers' Attitude to and Use of Analytics

Chantal Suder

A dissertation submitted to the University of Dublin
in partial fulfilment of the requirements for the degree of
MSc in Management of Information Systems.

1st September 2015

Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university. I further declare that this research has been carried out in full compliance with the ethical research requirements of the School of Computer Science and Statistics.

Signed: _____

Chantal Suder

Date: _____

Permission to lend and/or copy

I agree that the School of Computer Science and Statistics, Trinity College may lend or copy this dissertation upon request.

Signed: _____

Chantal Suder

Date: _____

Acknowledgments

I would like to thank my supervisor, Dr. Frank Bannister, for his advice and guidance throughout this dissertation.

Many thanks to my parents for their support, danke/merci.

Abstract

With the on-going growth and near-ubiquity of the use of analytics within large corporations, managers are more and more likely to come into contact with analytics regularly in the workplace. As a gender gap exists between the number of male and female employees - and managers - in STEM (Science, Technology, Engineering and Mathematics) fields, it is important to understand whether differences in the use and attitude to this technology exist.

This dissertation aims to understand the differences in use of and attitude to analytics between female managers and their male counterparts. The factors considered for this study were based on those used in frameworks studying technology use and adoption such as the Technology Acceptance Model.

An online survey was conducted to collect data about these factors from a sample of managers of both genders. The data were then used to conduct quantitative analysis to compare managers' responses based on gender and, therefore, understand if there were any differences in use of or attitude towards analytics between male and female managers. Qualitative analysis of free-text responses was also used to further the understanding of male and female managers' points of view.

The results indicate that, while there are no differences between the way in which male and female managers use and approach analytics for most factors, female managers were more likely to believe that there was a difference in the attitude to analytics between managers based on gender. In addition, male managers were more likely to use certain types of analytics tools more frequently, as well as consider themselves more familiar with analytics overall. From a practical viewpoint, these findings are an important step towards dispelling stereotypes about the importance of gender in STEM fields and making sure that all employees are guaranteed the same opportunities in the workplace.

Keywords: management, big data, analytics, female managers, gender, attitude to technology, technology use, technology acceptance

Table of Contents

List of Diagrams

List of Tables

Abbreviations

BI	Business Intelligence
CSO	Central Statistics Office
EB	Exabytes
EGFSN	Expert Group on Future Skills Needs
GB	Gigabytes
GDP	Gross Domestic Product
ICT	Information and Communication Technology
IDJEI	Irish Department for Jobs, Enterprise and Innovation
IT	Information Technology
LAAM	Learning Analytics Adoption Model
MIPS	Millions of Instructions per Second
PB	Petabytes
PEOU	Perceived Ease of Use
PU	Perceived Usefulness
RFID	Radio-frequency Identification
SME	Small and Medium-sized Enterprises
STEM	Science, Technology, Engineering and Mathematics
SQL	Structured Query Language
TB	Terabytes
TAM	Technology Acceptance Model
TDWI	The Data Warehouse Institute
TOE	Technology-Organisation-Environment Framework
TRA	Theory of Reasoned Action
UTAUT	Unified Theory of Acceptance and Use of Technology
ZB	Zettabytes

Glossary

Gini	Measure of statistical inequality
Google Analytics	Website behaviour tracking and analytics software
NoSQL	Non-relational, largely distributed database system
Numpy	Python library for statistics
Pandas	Python library for data analysis
Python	Programming language
SAS	Suite of analytics and data mining software
Scipy	Python library for sciences
Scikit-learn	Python library for statistics
SPSS	Software package for data analysis
Webtrends	Website behaviour tracking and analytics software

1.   Introduction

1.1   Introduction & Background

As the processing and storage of large amounts of data becomes both faster and cheaper every year and the Information and Communication Technology (ICT) industry grows year on year (Gartner 2014), methods of data analysis that were previously only affordable for highly funded research projects or large multinational companies are becoming more widely used. This is due to several factors. One is the fact that processing power doubles about every twenty-four months, as stated by Moore's law (Moore 1965, Moore 1965, Schaller 1996). Another important factor is the steadily decreasing cost of data storage – which has been declining since the 1970s (McCallum 2002).

Number of transistors in CPU*
Log scale

MOORE'S LAW DEFINED

10^{10}

10^{8}

10^{6}

10^{4}

10^{2}

1

1960   70   80   90   2000   10 14

Source: Intel   *Central processing unit
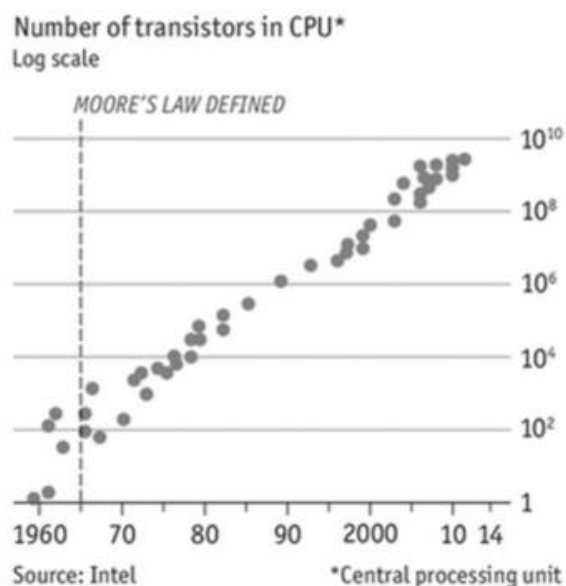
FIGURE 1.1 – Increase of Processor Performance over Time (Source: The Economist 2015) NB: A logarithmic scale is used.

The amount of data available for analysis is also growing. Hilbert et al. state that humankind was able to "store 2.9×10^{20} optimally compressed bytes, communicate almost 2×10^{21} bytes, and carry out 6.4×10^{18} instructions per second on general-purpose computers" (Hilbert et al. 2011, p. 60) in 2007, which corresponds to 290 Exabytes (EB) (or 290 million Terabytes (TB)) of data stored, and almost two Zettabytes (ZB) (or 2 billion TB) of data transferred. IBM (2013) estimates that 90% of the available data worldwide

have been gathered in the last two years. New technologies give companies the opportunity to gather various types of data - from banking transactions to website clicks and call detail records - for further analysis. Technology designed to support the processing and storage of this information is also rapidly evolving, especially since the rise of NoSQL databases, such as key-value stores popularised by Google and Amazon (Tiwari 2011), which make the analysis of large amounts of unstructured data more efficient. The most popular types of NoSQL databases include column stores such as HBase, which are optimised to store data in wide column formats, key-value stores such as Cassandra, which can be used to store unstructured data without using a schema, document databases such as MongoDB, which use semi-structured stores based on key-values of entities called documents, and graph databases such as Neo4J, which store relations between interconnected elements rather than using the traditional table format. Each of these database types enables analysts to gather and analyse data more efficiently using the data structure and database type most suited for the type of analysis being done.

These factors, as well as the buzz surrounding "Big Data" analysis (fields related to analytics and data mining of large datasets), have contributed to a rise in the use of analytics in companies worldwide. The demand for analytics has grown so much that there is now a global shortage of data analysts. Gartner, for example, claims that there is a worldwide shortage of "4.4 million IT jobs globally [needed] to support Big Data" (Pettey 2012, p.1) and the McKinsey Global Institute estimates that in the United States alone, there will be "a shortage of 140,000 to 190,000 people with analytical expertise by 2018", on top of the 1.5 million experts and managers with analytical know-how currently employed (Manyika et al. 2011, p.1). According to Davenport et al., this shortage puts "a serious constraint" on many sectors, as "demand has raced ahead of supply" (2012, p.1) in terms of qualified people with skill-sets in data storage, analysis and visualisation.

Managers, especially in departments such as marketing or finance – the analysis of whose data are among the "most frequent applications of [advanced analysis] techniques" according to Hair (2007, p. 307) - are frequently the main internal clients of analytics. They then become some of the main stakeholders of analytics. This means that managers are in a position make strategic decisions about investment into analytics teams and technologies, as well as drive a corporate culture of data-driven decision making. Since data analysis and mining are increasingly popular and play an important role in cost reduction and revenue growth in many industries today (Hair 2007), then the decisions managers make about the adoption and use of analytics within their company can have

important repercussions on the ability of a business to strive. This is especially important when analytics is one of the main drivers of competitive advantage, as is nowadays the case in many companies (Davenport 2006, Ribarsky et al. 2014, Morabito 2015).

It is, therefore, important to understand what factors influence or impact analytics adoption and use by managers, as the use of data analysis in a business can impact not only optimisation and decision-making on small projects but also corporate strategy and culture as a whole. In addition, the backgrounds of managers may be very different and social or educational barriers may exist that could impact a manager's opportunities or willingness to use analytics. Indeed, to use analytics efficiently, managers may need to understand complex statistical concepts and communicate with analytics teams effectively, so as to ensure that they are using the results of data analysis correctly.

Female managers, who are massively under-represented in ICT companies (Griffiths 2007, Tarr-Whelan 2009, Davidson et al. 2011), may be exposed to an additional gender-stereotype barrier as well as a technical and cultural one, which may be a problem if a large part of their work relies on the outputs of analytics teams.

As female managers are generally perceived as more geared toward soft skills (Wajcman 1998, Wade 2003, Roan et al. 2007), it is important to understand if there are any differences between how male and female managers use analytics, and whether gender-based skill perceptions play a role in how analytics are used or how the quality of analytics work done by members of either gender is perceived in the workplace.

In the IT industry, manager-levels jobs are still very much male-dominated and only about 16% of managers are female (Davidson et al. 2011). This inequality creates the possibility of a "men's club" attitude towards female workers, affecting their career advancement and ability to strive in their workplace (Kirchmeyer 1998). Since the increased presence of analytics will bring more female managers in contact with a male-dominated field, it is important to understand if gender has an impact on - and the extent to which it has an impact on - the use of and attitude to analytics of managers.

In essence, the democratisation of analytics brings managers into contact with more and more complex concepts to fuel more specific analytical insights. Managers will need to handle the disparity between the "softer" skills and the more technical skills required. It is then important to understand to what extent managers are familiar with advanced analytics techniques, how their understanding of analytics can help them in decision-making processes and whether the level of experience with analytics differs between male and female managers.

In the ICT industry especially, where the emphasis is on technical skills and a high gender imbalance are the norm (Davidson et al. 2011), it is important to understand whether gender has an impact on the adoption and use of analytics and how managers perceive the use of analytics so as to ensure that women have opportunities to be recognised as equally skilled. This will enable women to be able to reach management positions in data-driven corporate environments.

1.2   Academic Relevance

The current gender gap in ICT is not only large but also growing, according to Misa. He states that "the last 25 years have seen an increasing imbalance in gender in computing professions" (2011, p. xi). This is supported by Prescott, who adds that "in high-tech fields, [women] are almost twice as likely to leave their jobs" (2014, p. 7)).

As the field of analytics continues to thrive – with an estimated $16.9 billion Big Data market in 2012 (IDC 2012)) - it is important to further understanding of the gender gap in STEM (Science, Technology, Engineering and Mathematics) fields. Studying the differences in the attitude towards and use of analytics by female managers and their male counterparts is a first step towards determining whether a gender-based skills-gap or stereotypes associated with gender exist in analytics. This is, in particular, relevant to research into gender bias in management (Huyer et al. 2003, Griffiths 2007, Tarr-Whelan 2009, Haveman et al. 2009, Elacqua et al. 2009).

To help achieve this, this dissertation will aim to understand to what extent analytical outputs are being used by managers of both genders and how. An analysis of the attitude of managers toward analytics considering various factors and studying how the use of analytics by different genders is perceived will also help to achieve better insight into the challenges female managers face and the role data occupies in managers' work today. This is relevant to not only to research into career opportunities of female managers in IT, as mentioned above, but also for research into the use of analytics (LaValle et al. 2011, McAfee et al. 2012, Davenport et al. 2012, Mayer-Schönberger et al. 2013).

This contributes to research into the role of gender in the adoption of technology in the workforce and the influence gender has on the opportunities available to do so (Venkatesh et al. 2000, Ahuja 2002, Ahuja et al. 2005, Morris et al. 2005, Damanpour et al. 2009, Marlow et al. 2012).

1.3   Research Question

The primary research question is therefore:

Do female managers have a different attitude to analytics from their male counterparts?

This raises the following sub-questions:

1. Are there differences based on gender in the use of analytics by managers?
2. Do male and female managers perceive the use of analytics in the same way?
3. Is there a gender-bias in terms of perceived aptitude to use analytics?

1.4   Objectives

The aim of this dissertation is to understand if gender has an impact on the use of and attitude to analytics of managers, as they drive the adoption of analytics in companies today. Showing that analytics are omnipresent in the work of female managers is a step towards eliminating possible skewed perceptions - especially in the IT industry - of the skills of female managers and a step towards giving female managers equal opportunities for employment to male managers.

If there are large differences in how analytics are used by managers based on gender, then it will be interesting to consider what these differences are and how companies can make sure to consider the factors that are most important to drive the adoption of analytics by both genders.

1.5   Scope and Boundaries of the Study

As the extent to which managers use data depends on the availability of data sources and analytics teams, this study will be restricted to large companies - i.e. companies with more than 250 employees according to the European Commission's company size classification (2014) – who are able to invest in the required resources to maintain analytics teams and technology stacks.
This will be a cross-sectional study based on a survey of managers, studying different factors such as their use of analytics, their emotional perception of analytics and their perception of the role of gender in analytics.

1.6   Roadmap

This dissertation is organised as follows:

Chapter 1 gives an overview of the current context and the rise of analytics in the private sector. It presents the research question, gives insights into the relevance of the topic and a roadmap for the following chapters.

Chapter 2 provides the context within academic literature and a critical review of the published works relating to the gender gap in both ICT and management, the place of analytics, and "Big Data" in large companies today. This enables this dissertation to build on current knowledge on the adoption and use of technology, in particular analytics, in a corporate context and on the role of gender in individuals' attitude to technology.

Chapter 3 presents the research methodology used in this dissertation and the approach used to carry out and analyse the research while detailing the ethical considerations that this study takes into account.

Chapter 4 presents the research findings of this dissertation. It contains a detailed analysis of the gathered survey data and a presentation of the results of the research.

Chapter 5 discusses the outcomes of the research and the extent to which the research goals were achieved, as well as identifying the limitations of this study and providing insights into further research possibilities into the impact of gender on the use of technologies and into the gender gap in STEM fields.

2.   Literature Review

2.1   Introduction

The aim of this chapter is to review the literature in the field of Big Data and analytics, highlight the factors that influence the adoption and use of analytics, and understand the influence managers have on the adoption of technology. It will also highlight research done in Ireland specifically. This chapter aims to provide a better understanding of the impact gender may have on the adoption and use of analytics, and the perception of analytics in an organisational context.

This review first defines Big Data and analytics, and provides some background on the use of different types of analytics by companies. It reviews the existing models that describe which factors - including gender - influence the adoption and use of technology, with a focus on analytics in particular.

Section 2.2 provides the definition of Big Data and analytics, and insight into the importance of analytics for companies from a strategic point of view.

Section 2.3 reviews the factors that may influence the acceptance of technologies - including analytics - with an emphasis on the role of gender and the influence of managers.

Section 2.4 provides a short overview of the Irish context. It reviews studies on technology adoption in Ireland and employment statistics by gender.

Section 2.5 concludes this chapter and discusses the literature findings.

2.2   Leveraging Big Data and Analytics

'Big Data' has become a buzzword (Davenport et al. 2012b) but it, as well as what encompasses the field of analytics, is loosely defined. MIS Quarterly's editor-in-chief Paulo Goes considers Big Data to be "the new frontier in the wide spectrum of IT-enabled innovations and opportunities" (Goes 2014, p.iii) but the recency and variety of uses of the term in academic, technical and commercial contexts make it hard to pinpoint which

innovations fall under the field of Big Data and analytics. Jacobs gives the following as an "attempt at a meta-definition" of Big Data: for Jacobs, it is "data whose size forces us to look beyond the tried-and-true methods that are prevalent at that time" (Jacobs 2009, p.44) – which means that the size of 'big' data may evolve with time, as new methods of analysis of large datasets become commonplace and newer methods are developed to analyse even larger amounts of data. Big Data is 'big' not only in terms of its memory requirements when compared the storage capabilities of databases used in the past, but also in terms of the investment and technology stacks needed to process and analyse it.

A frequent definition of Big Data that highlights these challenges is its description by three properties, known as the 'three Vs' (volume, variety and velocity) or in some cases 'four Vs' when value (Hitzler et al. 2013) or veracity (Mattman 2013; Goes 2014) are added. This was first introduced by Gartner in 2001 (Laney 2001) but is now also used in academia to define Big Data (Zikopoulos et al. 2011; McAfee et al. 2012; Buhl et al. 2013).
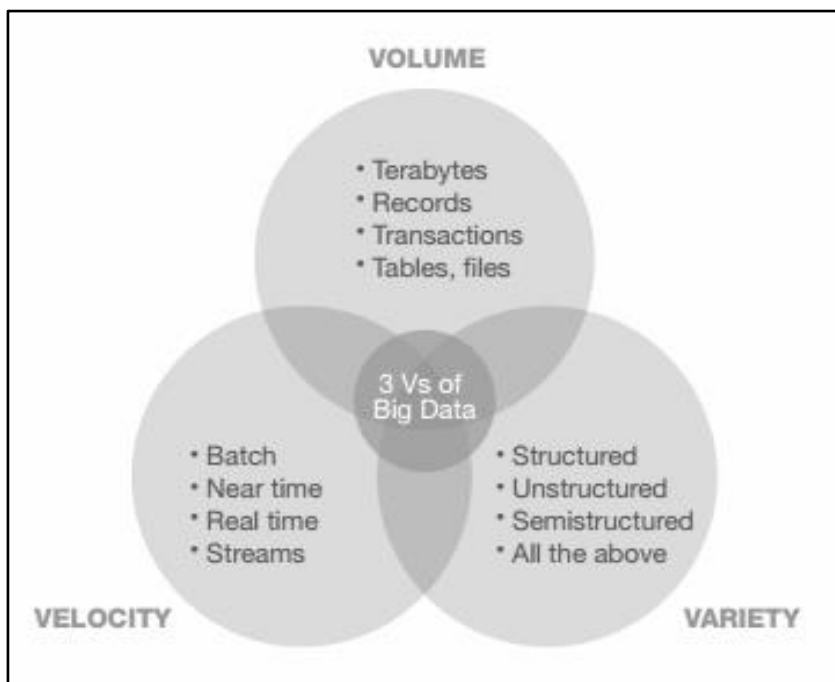
FIGURE 2.1 - The Three 'Vs'. (Russom 2011)

These 'Vs' describe Big Data using the following characteristics:

● Volume

The volume of data produced and stored worldwide is much larger than previously. It requires new structures to store and process it. Intel identifies companies 'highly

involved' with Big Data analysis as companies "working with 500 TB or more of data per week" (Intel 2012, p.4). Intel identifies the main challenges of Big Data as being data growth, infrastructure and governance (Intel 2012), as exponentially growing data entail a rapid growth of the technology stacks required to process and store this data. The growing volume of data is also highlighted by McAfee et al., who state that "[more] data now cross the internet every second than were stored in the entire internet 20 years ago" (McAfee et al. 2012, p.2). These exchanges of data are recorded and analysed, creating large data warehouses ranging from several TB to several PB. Wal-Mart, for example, accumulates an estimated 2.5 PB of data per hour from transactions and customer data (McAfee et al 2012).

- Velocity

  The speed at which data can be gathered and processed is increasing. This means more data is being transported faster. Google, for example, processes about 24 PB of data per day (Davenport et al. 2012b). 'Velocity' also means that data is being gathered and processed at growing speeds. Although a lot of data processing and analysis is still done over days, weeks or months (Russom 2011), some data are processed directly at high speeds. Data can now be gathered, stored and analysed in close to real-time, for example, to generate accurate real-time warehousing information for e-commerce or shipping companies.

- Variety

  Big Data is varied in that it encompasses a multitude of data sources, data types and methods of data collection and analysis. Intel identifies twelve top sources of data (c.f. Figure 2.2), the top two sources of data being documents and database transactions. The datasets stored then contain both structured data such as the business transactions in a database and unstructured data such as videos. It means that analysis is no longer restricted to text-based database records but can be done on images, sounds and social network relations. Documents, for example, which are cited at the top data source, can contain either structured or unstructured data in a large range of formats. New types of data warehouses, such as NoSQL databases, can be used to store and process unstructured or semi-structured data in non-tabular formats. Netflix, for example, uses an open-source NoSQL database called Apache Cassandra for its recommendation system and uses it to predict the success of certain movies or series (Netflix Tech Blog 2015).

Data Source	Total % Asked to Analyze
Documents	84%
Business transactions in database	82%
E-mail	74%
Imaging data	68%
Sensor or device data	57%
Internet search indexing	57%
Weblogs	55%
Social media	54%
Telephone conversations	52%
Videos	52%
Pictures	46%
Clickstreams	42%
Other	2%

Q: *Which of the following data sources are you being asked to analyze and glean insights from?*

FIGURE 2.2 - "Varied Data Types: Business Transactions Still Top Source" (Intel 2012)

In some cases, a fourth V is added to define Big Data. This additional V is usually either value or veracity. These are extensions and consequences of the previous three points.

● Value

Hitzler et al. argue that value is also important to the definition of Big Data, as the value companies get from the data they are storing has changed. The analysis of the data provides companies with value that was previously unavailable to fuel data-driven decision-making (Hitzler et al. 2013). Gathering more data about customers to predict or influence their behaviour – for example, by combining information about customers' past transactions, their demographics, geographical data or opinion polls - may give a company a competitive advantage over another company with no access to such data.

● Veracity

Veracity is sometimes added to the Vs that define Big Data (Davenport et al. 2012, Mattman 2013) to emphasise the raw, unaltered state – and sometimes poor quality - of the data gathered from these many sources, and the challenges that arise from this. Gathering data were it is available may give a more complete

overall picture of current and past states to companies, even if only partial information is available. NoSQL databases make it easier to store complex and evolving data models.

Analytics - or advanced analytics - is the field relevant to the study and analysis of data. The definition of the limits of what exactly the field of 'analytics' encompasses is also hard to define. A The Data Warehouse Institute (TDWI) report co-sponsored by IBM defines analytics as the field "where advanced analytics techniques operate on big datasets" (Russom 2011, p. 5) and then specifies that, in this case, advanced analytics means discovery or exploratory analysis done by an analyst (Russom 2011).

By this definition, some of the commonly used advanced analytics might not qualify as 'analytics'. Predictive modelling, for instance, is commonly used, and not exclusively exploratory in nature. This definition may, therefore, be too restrictive. On the other extreme, Rote from Teradata gives a much broader definition of analytics. He defines the field as any mathematical or computational technique that can be used to "answer questions or solve problems" (Rote 2005, p.1).

Gartner give a definition that highlights the way 'analytics' is used in both marketing and technical contexts. Their definition is:

> **"**Analytics has emerged as a catch-all term for a variety of different business intelligence (BI) - and application-related initiatives. [...] In particular, BI vendors use the 'analytics' moniker to differentiate their products from the competition. Increasingly, 'analytics' is used to describe statistical and mathematical data analysis that clusters, segments, scores and predicts what scenarios are most likely to happen. Whatever the use cases, 'analytics' has moved deeper into the business vernacular."

(Gartner IT Glossary 2013)

This definition captures both the broad sense of analytics used in sales and the more specific technical definition of analytics, as statistical methods and techniques used on large datasets.

In essence, analytics techniques are computations that leverage the processing speeds of computers to discover new information, by aggregating or manipulating datasets with mathematical and statistical techniques which could not feasibly be applied by hand.

Analytics can be split into different subcategories depending on the context. There is no one accepted definition of analytics subcategories, as the techniques and methods used for data analysis overlap with - and at times borrow from - other STEM (Science, Technology, Engineering and Mathematics) fields such as robotics and computer science. The fact that 'analytics' is used as a sales term for many different kinds of BI related products (Gartner IT Glossary 2013) further complicated the classification of analytics into clear categories, as the term is sometimes used to drive sales by leveraging the buzz surrounding Big Data and analytics.

One way to split analytics into subcategories is to consider the 'depth' of the analysis or 'how far away' from the raw data on the path to actionable information the analysis is. This usually relates to the complexity of the statistical method used.

The two frameworks below present different ways to define categories based on depth, going from simple *ad hoc* summarisation to predictive analytics. The first framework (Kotorov 2015), as shown in Figure 2.3 below, is the categorisation presented by a company called Information Builders. This reflects a view of analytics close to the one given by BI vendors. Although it includes mentions of some proprietary software, it gives a general overview of the different types of analytics by increasing complexity. While the concept of 'depth of analytics' is a consistent way to separate analytics into subcategories, the concept of 'breadth of analytics' presented here seems to relate to the amount of data generated as results of the analysis. Figure 2.3 would then suggest that data discovery usually generates fewer outputs than a dashboard of performance metrics, although the opposite is usually true. The concept of 'breadth', therefore, seems to be introduced more to suit the data visualisation, add a 'buzzy' sense of pseudo-scientific validity by showing the list of categories on a chart rather than as a list, in hopes of aiding the sales pitch of the company. The use of a chart is driven more by the need to encourage sales and surround the information by an analytics look and feel than to add insight into analytics classification.
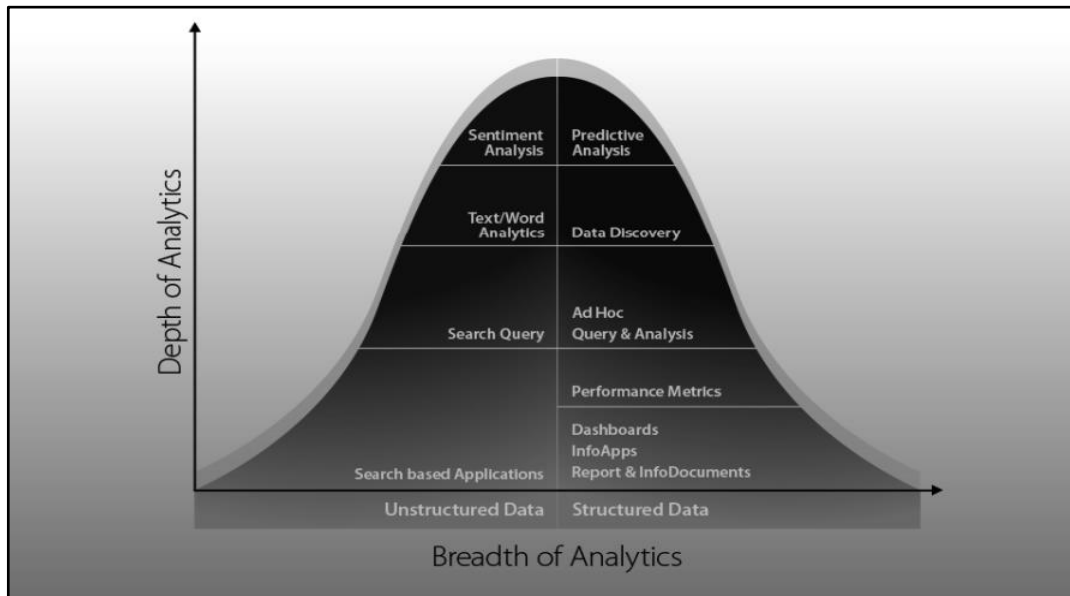
FIGURE 2.3 - Depth and Breadth of Analytics (Kotorov 2015)

Figure 2.3, while offering a categorisation of analytics by 'depth of analytics' is very BI industry and sales oriented. The use of a graph format without axes and values indicates a lack of the rigour required for a more academic definition of sub-categories of analytics.

The second framework is one proposed by the Educause Learning Initiative, which attempts to bridge the gap between the BI industry and academia. This framework's aim is to propose a common framework to create a generally accepted definition of analytics to be used in business and higher education (Van Barneveld et al. 2012).
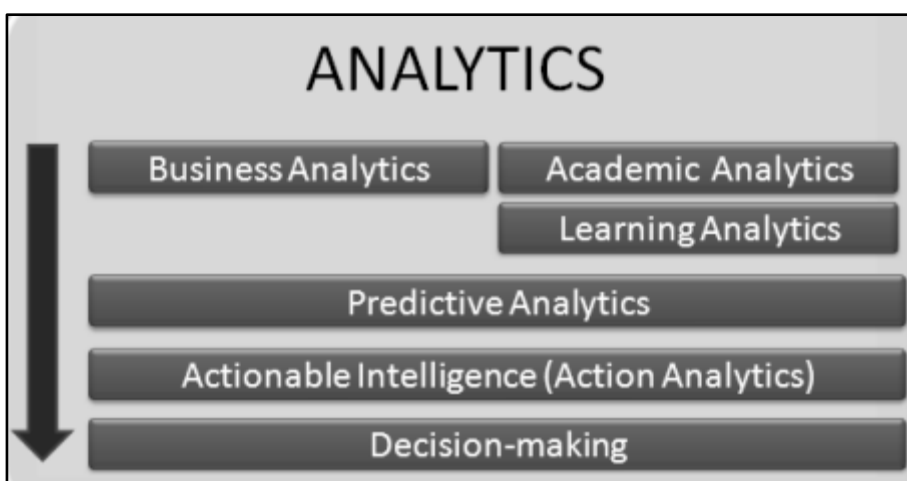
FIGURE 2.4 - Framework for Analytics in Business and Higher Education. (Van Barneveld et al. 2012)

This framework approaches the classification of analytics not in terms of the complexity of the analysis but rather in terms of the value it adds to decision-making. It classifies

analytics from less to more 'actionable'. The use of predictive analytics leads to actionable information and decision-making such as described by Van Barneveld et al. (2012) in Figure 2.4. 'Actionability' though depends on the types of decisions made and the kind of problems being solved more than on the type of analytics which makes it hard to use this as a general definition.

Although both frameworks have limitations, they each suggest that analytics can be split into one part that relates to the current state and another part which relates to the prediction of a behaviour or situation.

A better way to defined sub-categories - or "perspectives", as they are called by Chen et al. (2012, p.1182) - of analytics is perhaps to look at what kinds of questions the analysis addresses. Gartner suggests the following split – with corresponding questions relating to each subcategory:

- Descriptive analytics – what happened?
- Diagnostic analytics – why did it happen?
- Predictive analytics – what will happen?
- Prescriptive analytics – how can we make it happen? (Richardson 2013)

These categories cover all aspects of a certain situation or behaviour, going through different levels of reasoning, from reasoning about the basic *ad hoc* state of the data (answering 'what happened?' or 'what is happening?') to the analysis of past behaviour ('why did this happen in the past?', 'what factors may have contributed to this event?' or 'what factors differentiated past events?') to the prediction and influence of future events ('how likely is this event to occur?' and 'how can we make this event occur more/less?'). The complexity of data analysis techniques grows with each category, as shown in Table 2.5, as it is more complex to understand why a state occurred in the past than to state its existence, and it again is more complex to try to infer the future from the past, than it is to explain it.

Sometimes, predictive analytics and prescriptive analytics are grouped into one category, but the distinction between 'what will happen?' and 'how can we make it happen?' is important, especially in a business context. This is the difference, for example, between predicting which customers are the most likely to leave to a competitor and predicting which customers are the most likely to be saved if contacted before they leave for a competitor. From a marketing perspective, this would, for example, enable a company to

target only those customers who may still be saved instead of contacting all customers who are likely to leave to a competitor.
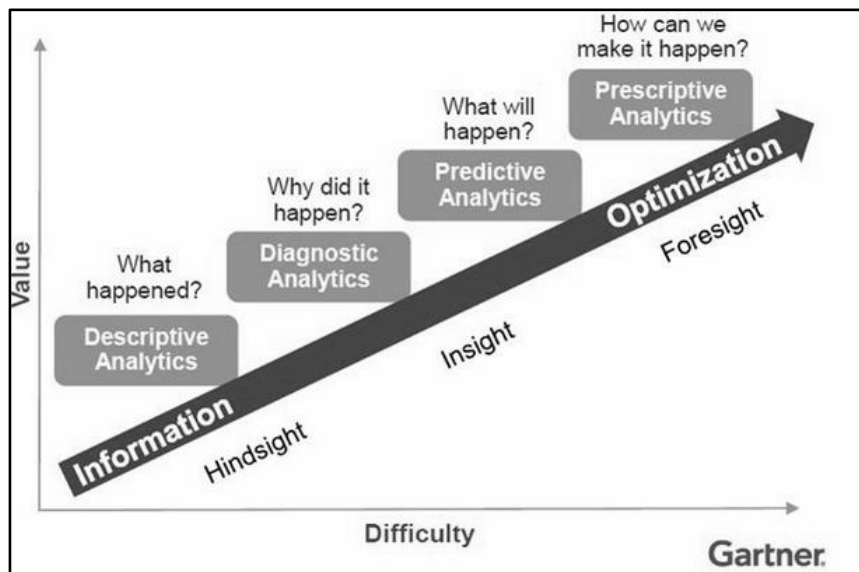
FIGURE 2.5 - Levels of Analytics (Richardson 2013)

Leveraging Big Data and its analysis can give companies a competitive advantage if they are better able to understand why events are occurring and how future events can be influenced than other companies. According to LaValle et al.,"[top-performing] companies are three times more likely [...] to [use advanced] analytics, and are two times more likely to say [it] is a competitive differentiator" (LaValle et al. 2012).

Davenport and Harris' book's title "Competing on Analytics: The New Science of Winning" (Davenport and Harris 2007) is evocative of their view of the value analytics can bring to a business. They have coined the term 'analytical competitor' to describe companies who are successfully using analytics as their main driving force for creating value and guiding strategy (Davenport and Harris 2007). Davenport highlights six different areas in which analytics will have the "most impact on decisions and actions" (Davenport 2014, p. 47):

- customer satisfaction, with the analysis of call centre records, social media content and other customer interactions,
- customer journeys, with the tracking and analysis of purchases, subscriptions and behaviours of customers,
- supply chain risk, with the use of supplier data,
- competitive intelligence, by following market trends and gathering data on the market, competitors and movements of customers across an industry,
- pricing, with the use of price optimisation analysis and software,

- discovery and experimentation, by using innovative techniques on existing data to help decision-making in different areas of a business, especially finance, marketing and business development (Davenport 2014, p.47-48).

Another area which could be added as a separate point to this list is customer loyalty, as customers are nowadays being rewarded for their loyalty to companies with loyalty programs based on the frequency of their purchases, purchasing patterns, seniority as a customer of a given company or customer segmentations created by businesses.

These advantages due to analytics, though, may only be available while a disparity exists in the extent to which analytics are used by companies. Indeed, the competitive advantage large companies attain by analysing data can only be sustained while certain companies are able to gather insights into their customers' behaviours while others are not. If the cost of analytics technology stacks continues to decrease, the demand for more analytics experts is filled, customer data continues to be traded, and analytics become even more ubiquitous, then it will be harder to sustain a competitive advantage based solely on analytics, as Davenport & al.'s term 'analytics competitor' seems to suggest.

2.3   Factors influencing Technology Adoption and Use

Although many businesses have already adopted analytics, there is a large range of stages at which companies are in its integration into their processes and decision-making behaviours. The range goes from companies relying only on basic regulatory financial reporting to businesses who compete mostly on analytics and for whom data-driven decision-making is a core part of company culture.

There are many factors that can influence the adoption and use of technology within companies, not only on an organisational level but also on an individual one. The Technology Acceptance Model (TAM) (Davis 1989, Venkatesh and Davis 2000, Venkatesh and Bala 2008) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al. 2003), based on the Theory of Reasoned Action (TRA) (Fishbein and Ajzen 1975, Ajzen and Fishbein 1980) suggest that a few main factors influence the adoption and use of technology.
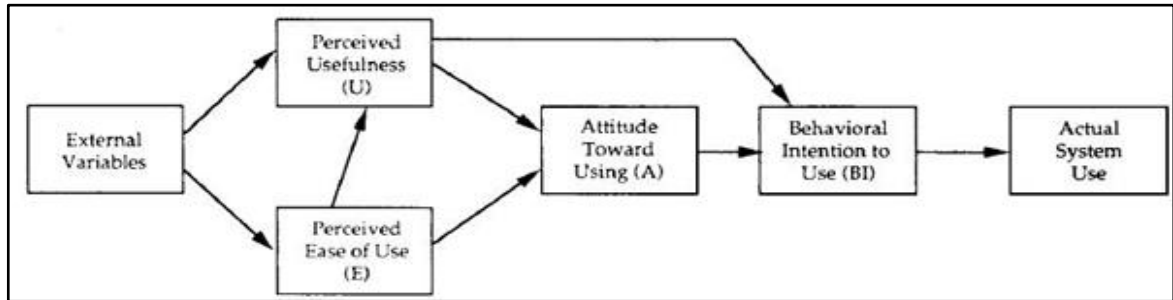
FIGURE 2.6 - Technology Acceptance Model (Davis 1989)

These main factors drive the acceptance and use of technology by an individual user. The two most important factors according to the TAM are perceived ease of use (PEOU) and perceived usefulness (PU) (Davis 1989). Perceived ease of use is defined as the "degree to which an individual believes that using the system would be free of cognitive effort" (Shroff 2011, p. 604) and perceived usefulness as the "degree to which an individual believes that using [the technology] would enhance his or her performance" (Shroff 2011, p. 603).

While these factors describe perceived characteristics of a technology which may influence its adoption, it may be hard to apply this model to product design or to training to further technology adoption, as ease of use and usefulness are already some of the main objectives considered when creating a technology. It is therefore more useful for the purpose of forging an understanding of the social and organisational context for technology adoption than for product design.

To address some of the limitations of the TAM, it has been extended and revised in further studies. One model based on the TAM is the UTAUT, which contains many of the same factors, but also includes social influence and takes experience into account.

In the revised UTAUT (Venkatesh et al. 2003), the factors are the following:

● Performance Expectancy: how well the user thinks a technology works and responds to his or her need,

● Effort Expectancy: how hard the user thinks it to use and learn how to use the technology,

● Social Influence: how much does society drive the use of this technology, how accepted is it in the rest of society and in the user's social circle,

● which are influenced by gender, age, experience and voluntariness of use, and influence the Behavioural Intention of users,

● and Facilitating Conditions which are influenced by age and experience.
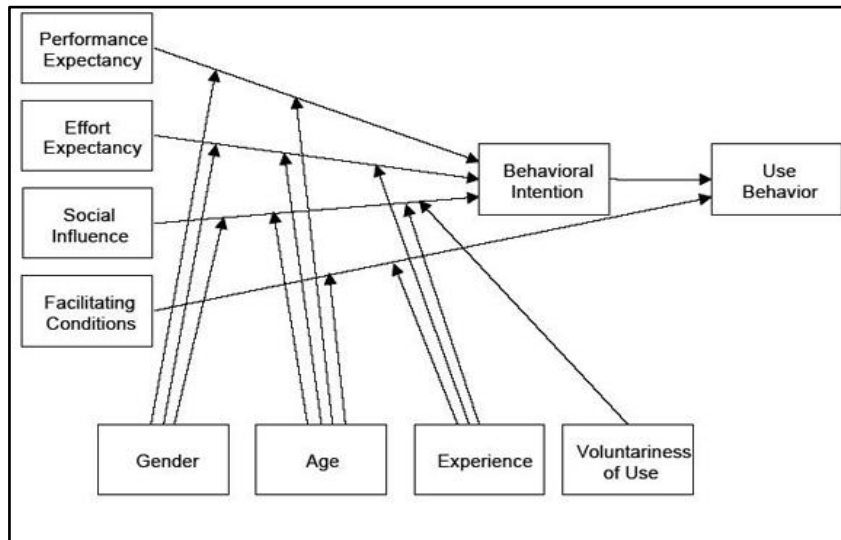
FIGURE 2.7 - UTAUT Model (Venkatesh et al. 2003)

UTAUT may be more useful than TAM for use in a business context, as strategies to encourage adoption of a technology could be tailored to potential users, based on some of their characteristics, such as age, gender and experience. The impact of these three characteristics is limited, though, as the model suggests that they are only moderating influences on other variables which then in turn influence technology acceptance. This means that, for example, gender has an impact on performance expectancy (male users tend to have higher performance expectancy (Venkatesh et al. 2003)), but not directly on technology adoption. This limits the possible applications of this model in a business context when trying to understand how to further technology acceptance and adoption.

Although there are limits to its applicability outside of academia, TAM (sometimes in combination with UTAUT) has been used, tested and extended in many studies surrounding technology acceptance (Straub et al. 1997, Hu et al. 1999, Moon et al. 2003, Legris et al. 2003, Venkatesh et al. 2008, Mariani et al. 2013) and use (Venkatesh 2000b, Gefen 2003, Hong et al. 2006). It has also been criticised for putting an emphasis on affective feelings: Yang et al. (2004) suggest that cognitive attitude (which consists of "the evaluation, judgment, reception, or perception of the object of thought based on values" (Yang et al. 2004, p. 21)) is more important in relation to IT acceptance. Yang et al. therefore suggest separating 'Attitude Toward Using' into two separate factors to highlight the difference between affective and cognitive factors, as emotions toward a technology and a more rational perception of its benefits may be different.

In an organisational context, the adoption of technology is influenced by factors more specific to a professional environment. Professionalism, specialisation and slack were found to be correlated with the adoption of technology innovations in public libraries for

example (Damanpour 1987). Tornatzky and Fleisher (1990) proposed the Technology-Organisation-Environment Framework (TOE), which explains how company demographics, external and internal factors, and the technology itself shape an organisation's technology adoption and implementation decision.
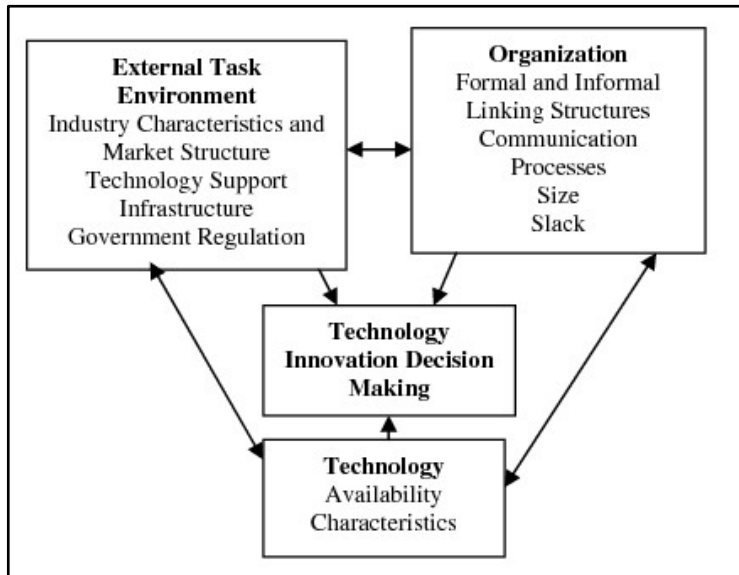
FIGURE 2.8 - Technology-organisation-environment Framework (Tornatzky and Fleisher 1990)

This is relevant to individual technology acceptance, as Frambach et al. (2002) state that the organisational adoption of information systems plays an important role in individual acceptance of technology innovations. They propose an extended framework, separating the organisational decision from the individual decision. The factors influencing adoption of technology and its continued use in on an organisational level are described by the proposed organisational innovation adoption framework (Figure 2.9).
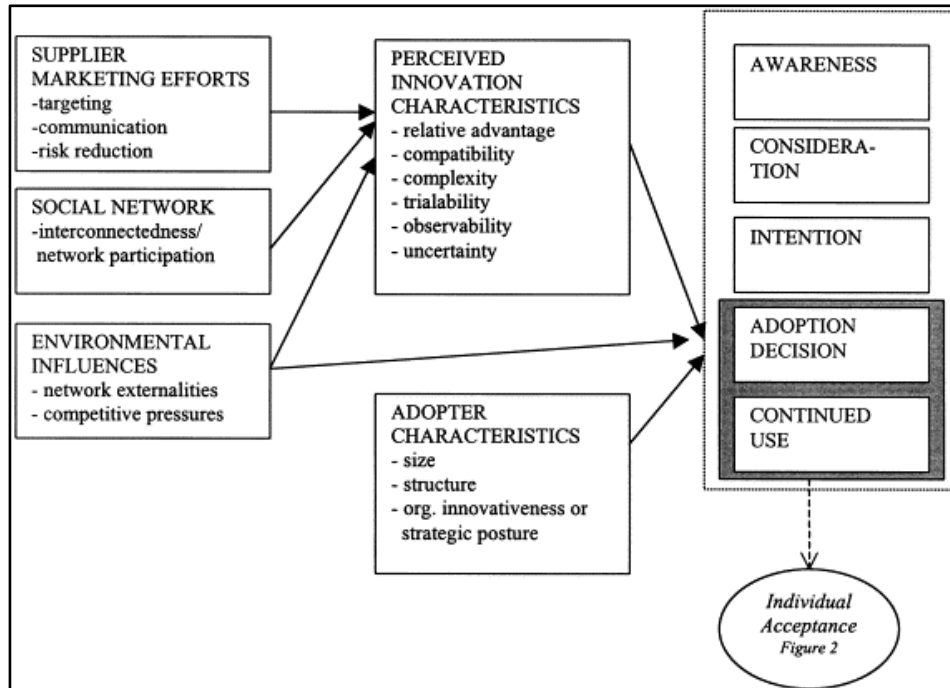
FIGURE 2.9 - "A Conceptual Framework of Organizational Innovation Adoption." (Frambach et al. 2002)

The user's decision to adopt and use a technology on an individual level (Figure 2.10) is, according to Frambach et al. (2002), influenced by the following factors:

- Organisational facilitators/internal marketing, which include training in the technology, the organisational push or support of the technology and persuasion by the other members of the organisation,
- Social Usage, which includes information, social pressure and usage outside of the organisation,
- Personal characteristics, such as demographics, experience, values and product know-how which in turn impacts the person's innovativeness and attitude to innovation.

The organisational factors, personal characteristics and social usage of the technology then shape the individual's attitude to innovation and disposition to accept such innovations, influencing the individual's acceptance of the technology (Frambach et al. 2002).
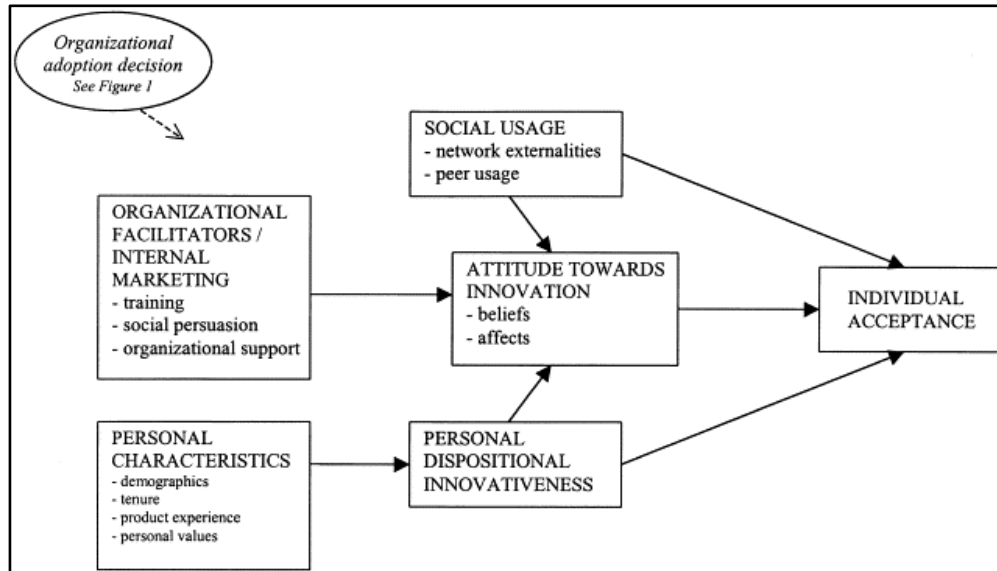
FIGURE 2.10 - "A Conceptual Framework of Individual Innovation Adoption in Organizations." (Frambach et al. 2002)

One aspect that is not mentioned in Frambach's frameworks is that within an organisation, while technology acceptance and frequency of use may depend on the individual, technology adoption can sometimes be a requirement rather than an individual's choice - as the mentions of 'training', 'social persuasion' and 'organisational support' focused on the individual seem to suggest (see Figure 2.10).

Frambach et al.'s frameworks support the view that, while many external factors influence technology adoption in organisations, the main individual factors highlighted by the TAM and UTAUT frameworks (perceived usefulness, experience for example) remain valid in an organisational context. This framework, like UTAUT, also suggests that, while demographics may influence a user's willingness to adopt a technology, they do not directly influence technology adoption but rather have a moderating effect on other factors.

According to Rapp et al. (2008), "management '[championing]' [...] technology innovation initiatives is crucial"(Rapp et al. 2008, p.7) for the successful adoption and continued use of a new technology in organisations. Especially senior management support was found to have a significant impact on technology acceptance. The impact of managers' attitudes on the acceptance and use of technology in an organisational context on individual adoption and use are highlighted by Lewis et al. (2003) as one of two main factors that can influence beliefs about technology use.

Another framework, called "diffusion of innovations", highlights that early adopters of technology may also be quite different people from those who are generally later adopters of technology (Rogers, 2010). This phenomenon exists on both an individual level and on an organisational level. For Rogers, adoption of technologies is done in five stages: knowledge, persuasion, decision, implementation, and confirmation (p. 162). It is, therefore, important to consider not only use of a technology but also frequency of use and familiarity with a technology when considering its adoption.

In an analytics context, the willingness of employees to adopt data analysis techniques is influenced by the company's culture. In particular, a company's openness to and push for the use of data-driven decision-making, its business strategy and its data quality are influences from the organisational side and the user's innovativeness and experience can be important from the individual side (Dahlan et al. 2003). This is supported by Huang et al. (2004), for the acceptance of web mining, adding to these factors the importance of business complexity and the pressure of competition on an organisational level. Huang, Liu and Chang (2012) studied the factors influencing the adoption of data mining tools on an individual level within organisations, to extend the TAM model (based on the updated version TAM3 (Venkatesh et al. 2008)) to data mining. By studying the adoption of data mining by information systems professionals, they confirmed that perceived usefulness and ease of use are main factors for the adoption of data mining tools by IT professionals. This confirms the TAM's applicability for the adoption of data mining tools by analytics professionals. Huang, Liu and Chang (2004) also found emotional motivations not to be as important as cognitive factors for the acceptance of data mining tools, supporting Yuan et al.'s (2004) critique of TAM for not separating the affective attitude of the user from his or her cognitive attitude. Proposing a version of TAM adapted to learning analytics, the Learning Analytics Adoption Model (LAAM), Ali et al. (2013) reaffirms that perceived usefulness, ease of use and prior experience are significant factors for the adoption and use of analytics.

Gender as a factor in technology adoption and use has been studied to extend the TAM in several studies. When studying emails as a technology for example, while the frequency of use of the technology was the same for both genders, women and men's perceptions of emailing differed (Gefen and Straub 1997). Men were more influenced by perceived usefulness and women by the perception of ease of use of the technology and by subjective norm (Venkatesh et al. 2000c, Venkatesh et al. 2000d, Yuen et al. 2002). There is a gender gap in the perception, usage and beliefs surrounding technology not only in a private context (Vekiri et al. 2008) but also in a professional context (Ahuja et al. 2005).

Ahuja et al. (2005, p. 427) suggest that "perceptions of the environment moderated by gender" may influence the use of technology by women, with overload and autonomy being the two main factors influenced by gender that affect IT adoption.

For the acceptance of certain technologies in a private or learning context, though, there are diverging views on whether gender has a significant impact. Whitley (1997) states that there are no significant differences between technology use by men and technology use by women. This is supported by Padilla-Melendez et al. (2013) who extend the TAM to include playfulness as a factor into the existing model. Huang et al. (2013) confirm this for technologies such as social media and video sharing, but for certain other technologies, such as Web 2.0 applications (for example, blogs and wikis) they conclude that women are more 'anxious' when using applications than men, using the UTAUT. Another study using the TAM suggests that women and men perceive e-learning differently, with women being more influenced by computer self-efficacy and ease of use of e-learning, and men by its perceived usefulness (Ong et al. 2006). The significance of gender seems to depend on the technology used and does not seem to be generalisable for all technologies.

2.4   Analytics and attitude to technology in the Irish context

The Action Plan for Jobs 2014 by the Irish Department for Jobs, Enterprise and Innovation (IDJEI) highlights Big Data and analytics as the fourth of its nine "Disruptive Reforms" (IDJEI 2014, p. 9) for job creation. It aims to make Ireland a European leader in analytics through "the launch of an Open Data initiative and further strengthening [of] enterprise engagement in the €88 million Insight Research Centre and the CeADAR Technology Centre" (IDJEI 2014, p. 10).

TABLE 2.1 - Demand for Analytics Jobs in Ireland (Forfás and EGFSN 2014)

Category		Employment Demand	% Total Employment
Deep analytical talent		3,300	0.18
Of which	*emerging analytics roles*	*1,500*	*0.08*
	established analytical roles	*1,800*	*0.10*
Big data savvy		25,780	1.38
Supporting technology professionals		6,000	0.32
Total		35,080	1.88

Forfás and the Expert Group on Future Skills Needs (EGFSN), who advise the Irish government on employment, technology and innovation, predict that the current employment demand of 35,080 analytics professionals (see Table 2.1) will rise to between 40,450 and 62,220 jobs by 2020 (Forfás and EGFSN 2014).

Although Ireland has caught up with other European countries in terms of the employment of women (Malakh-Pines et al. 2010) since the end of the marriage bar in 1973 (National Women's Council of Ireland 2015), only 29.8 per cent of information and communications sector employees are women - the lowest percentage in the tertiary sector (as defined by the Central Statistics Office (CSO)), and lower than the 46.8 per cent average over all economic sectors (CSO 2012).

Technology acceptance in an Irish context has been studied as part of a study of technology acceptance in thirty-four European countries by Kim et al. (2014), ranking twelfth out of thirty-four for its level of acceptance of science and technology. The most important factors in this model are image, knowledge, perceived benefit and perceived risk of the technology, as well as age and gender of the individual. Other factors such as gross domestic product (GDP), religiosity and post-materialism of the country were used as macro-variables. Zhu et al. (2003) and Pramatari et al. (2009) confirm that performance and effort expectancy, as well as anxiety, are important factors in technology acceptance in an Irish context, respectively for the adoption of e-business technologies and the adoption of radio-frequency identification (RFID) in Ireland. Another study, on the adoption of IT for elderly community-care in Ireland reaffirms the relevance of the TAM factors for the acceptance of devices used in the care of older adults in Ireland (Walsh et al. 2011). These studies confirm that the TAM can be applied to and is valid in an Irish context.

2.5   Literature findings

While some studies have researched the adoption of specific technologies in Ireland, there has been little to no focus on Ireland specifically in research about the role of gender in the acceptance of technology - including analytics.

Many studies on TAM and UTAUT focus solely on one country or one organisation (university, institution or company) when studying the adoption and use of a technology. This would call into question the generalisability of TAM and UTAUT, had this framework not been studied and found to be valid in different countries and contexts.

Although the TAM focuses strongly on variables which may be hard to influence in a business context (such as perceived ease of use of a technology), which may impact its usefulness as a model for technology adoption to be used in a corporate context, TAM has been studied for many technologies, including web mining, data mining tools and learning analytics.

The "Diffusion of Innovations" innovation adoption stages also highlight the importance of considering an individual's previous experience with a technology and their familiarity with it when trying to assess a user's attitude to a technology.

While these models may present some issues in terms of applicability to a business context, the variables used within them give an overview of the way in which users use technologies and what their attitude to the technology they are adopting is. This makes it a good basis for research into differences in technology use based on gender, as the variables have been previously proven to be influential in the adoption of technology. This dissertation will, therefore, use variables from the TAM, the UTAUT and the "Diffusion of Innovations" models for the analysis of differences in attitude to and use of analytics between male and female managers.

3.   Methodology

3.1   Introduction

Saunders et al. define research as "something that people undertake in order to find out things in a systematic way, thereby increasing their knowledge" (Saunders et al. 2009, p.5) and methodology as "the theory of how research should be undertaken" (Saunders et al. 2009, p.3). Methodology defines the systematic way in which the discovery of knowledge will be undertaken and there are various research philosophies and strategies that can be used to undertake research.

This chapter reviews the methodological approaches considered, describes the selected research methodology and comments on the limitations of this method. It describes the research strategy and the types of data gathered and discusses the lessons learned from the research process.

3.2   Research Methods

Saunders et al. (2009) describe research methodology as an onion, coining the term 'research onion'. This onion is composed of several layers: research philosophy, approach, strategy, choice, time-horizon, and techniques and procedures. Similar allegories have been used by others to describe the layers of research methodology. Easterby-Smith et al. (2012, p.19) for example, use the image of four tree trunk rings, with ontology at the centre, followed by rings of epistemology and methodology, and methods/techniques as the outer ring.
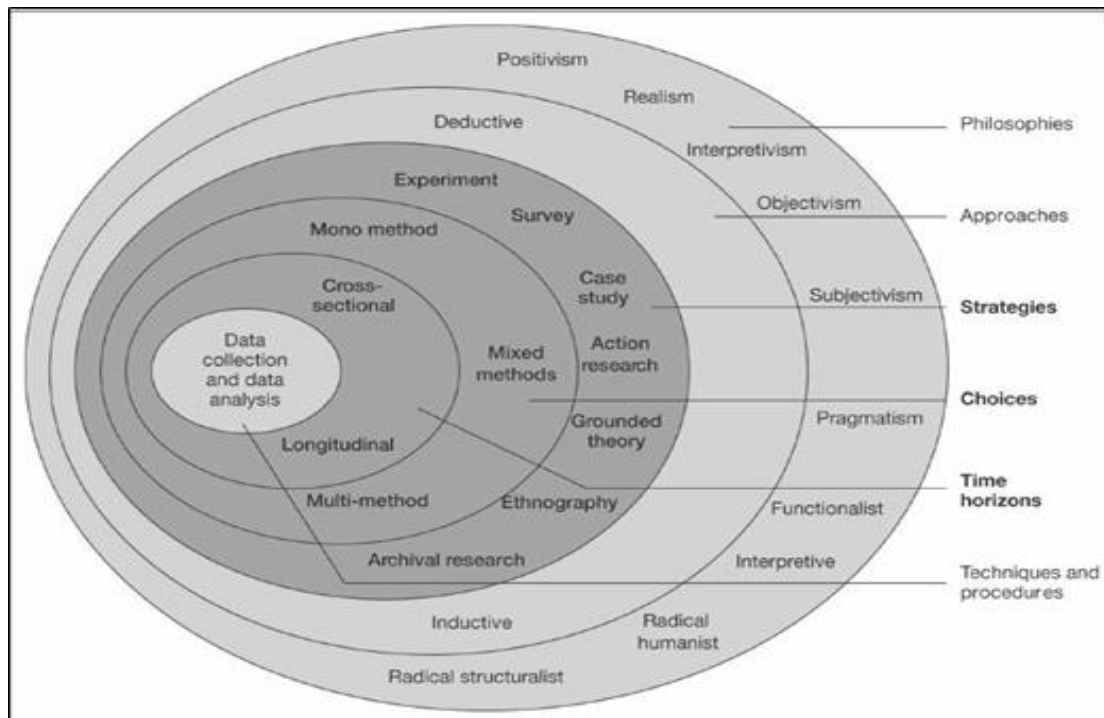
FIGURE 3.1 - The Research Onion (Saunders et al. 2009)

Saunders et al. identify four main research philosophies, as well as several minor variants of the main research philosophies, as shown above in Figure 3.1. Other authors may define the research philosophies differently and the number of philosophies considered as 'main' research philosophies as well as their definitions can vary by researcher and research field (Creswell 1994, Crossan 2003, Johnson et al. 2004, Eriksson & Kovalainen 2008, Green et al. 2012).

Saunders et al. suggest the following four main research philosophies for management research:

● **Positivism** adopts the "stance of the natural scientist" (Saunders et al. 2009, p. 113). It is the view that "properties should be measured through objective methods" (Easterby-Smith 2012, p. 22). With positivism, hypotheses are developed - usually from existing theory - to be tested on collected data, while keeping the research as independent as possible from the subject. Usually, a highly structured methodology is used to facilitate replication (Saunders et al. 2009). Kassi (2009) suggests that this research philosophy is best suited to areas of research where "the world is 'knowable'" and "[it] is knowable, fixable, provable and can be discovered and described" (p. 95).

- **Realism** suggests that reality is the truth and that it exists independently of the human mind and of human perceptions. It is similar to positivism as it also uses a scientific approach. There are two different kinds of realists: direct realists and critical realists. The difference in a business context is explained by Saunders et al. (2009) as follows: "[t]he direct realist perspective would suggest the world [...] operates, in the business context, at one level (the individual, the group or the organisation) [but] [t]he critical realist [...] would recognise the importance of multi-level study (e.g. at the level of the individual, the group and the organisation)" (p. 115).

- **Interpretivism** - also called social constructionism (Easterby-Smith 2012, p. 23) - criticised positivism by stating that the business world is too tied to social interaction, phenomena and change to have fixed laws like natural science. Interpretivists adopt an "empathetic stand" to "understand the world from their point of view" (Saunders et al. 2009, p. 116). "Reality is determined by people" (Easterby-Smith 2012, p. 23), instead of being measurable, as is the belief of positivists. Kassi (2009) claims that for this research philosophy, to the contrary of positivism, "the world is 'indefinable'" (p.96) and all knowledge gathered is relative to the researcher. The research findings are then restricted to the state of the social world they are studying as they apply to a particular point in time and context.

- **Pragmatism** adopts mixed or multiple methods trying to take into account both the objective and subjective points of view, and rally the structured scientific stance of positivism and the social interpretation stance of interpretivism to answer the research question. Belk (2007) suggests that this research philosophy is best suited to research that "does not align itself with a single system of philosophy and reality" and that "[rejects] [the] traditional dualism [between positivism and interpretivism]" by using multiple methods "either sequentially or simultaneously" (p. 199).

3.3   Research Strategy

The research philosophy of this dissertation is pragmatism as its aim is to study the behaviour of managers, the researcher being as independent as possible from the respondents, with both quantitative and qualitative research methods. The objective is to try to objectively assess the differences in attitude towards analytics of managers based

on their gender with quantitative methods, but also to supplement this with insights from qualitative analysis of text data. As the research findings are the result of testing of a hypothesis, namely the hypothesis formulated at the beginning of the study about the existence of differences in male and female managers' attitude to analytics, the research approach is deductive - starting from a hypothesis to prove or disprove it - and not inductive - in which case, the starting point would be an observation of the environment to formulate a theory. Saunders suggests that deductive reasoning may be more focused on "scientific principles", "a highly structured approach" and focuses on "explaining causal relationships between variables" based on "the collection of quantitative data" (Saunders et al. 2009, p.127).

The research strategy chosen to collect the quantitative data required to examine the correlations between the variables - including gender - is to conduct an online survey of managers. The data will then be analysed mostly using quantitative methods as the correlation between variables highlights their predictive power to explain certain studied behaviours. To add insight into the attitudes of managers towards analytics, free text data will be collected and analysed using qualitative methods. Miller et al. (1997) suggest that qualitative research is a "'window' through with [the researcher] might 'see' and comment on social issues" (p. 2) and Saunders et al. (2009) suggest that it can "aid interpretation" (p. 154). Qualitative research helps put the quantitative findings into context while making sure that the key points most important to the surveyed managers are highlighted.

The time-horizon for this study is cross-sectional and the data collection is done over a one month period.

The quantitative analysis done is focused on determining the correlation of the different studied variables to gender to understand whether the hypothesis that there are no differences between male and female managers in terms of use of analytics should be rejected. This makes it possible to identify the differences between male and female managers, to understand with what confidence the identified differences are statistically significant and, therefore, to understand if there are any differences in the way managers use analytics or in the attitude they have to analytics based on gender.

3.4   Ethical Requirements

The survey was submitted to the School of Computer Science and Statistics' Research Ethics Committee. The survey was approved by the committee, as it satisfies its

standards for ethics. Ethics are important to consider in research especially when gathering data from participants so as to ensure the safety and well-being of participants and their knowledge of the conditions surrounding their participation in a research study.
 All participants consented to their participation to the study and could withdraw from the study at any time. All participation was voluntary and all answers were collected anonymously. In additional, all questions in the survey were optional, so any question could be left blank if the respondents did not wish to answer some of the questions in the survey. All participants were required to be eighteen years of age or older. More details are available in Appendix A.

3.5   Population and Sampling

Ideally, a random sample of all managers working in organisations having adopted analytics would have been gathered and all managers would have had an equal chance to be selected for this study. This was not feasible given the complexity of such a task, as well as the financial and time constraints of this study. As there is no publically available ranking, standard or organisation to externally assess the degree of analytics adoption and amount of data available within a given company, this study is restricted to managers with experience in large companies (i.e. more than 250 employees according to the European Commission's classification (2014)), as large companies are more likely to have the infrastructure, budget and personnel resources required to maintain an analytics team and an analytics technology stack. Tavana (2014) states that the main "roadblock" (p. 108) for SMEs (small and medium-sized enterprises) is the SME "skills gap", as SMEs are less able to afford the highly competitive salaries required to maintain a highly skilled analytics workforce.

Random sampling is important to ensure the generalisability of the research findings, and non-random sampling introduces limitations to the reliability of the results, but random sampling is often difficult and, in this case, was not possible.

In addition to the difficulty of contacting a random sample of managers, financial and time constraints made it impossible to gather a pseudo-random sample of managers so snowball sampling was used. This impacts the reliability of the results of this study. For this survey, several demographics were recorded, and managers from a broad range of industries and backgrounds were contacted and responded to the survey, so as to make sure that the demographics of the sample did not affect the study to a large extent and were as close as possible to the general population.

3.6   Data Collection

An online survey was used for data collection. Surveying is "a systematic method for gathering information from (a sample of) entities for the purpose of constructing quantitative descriptors of attributes of the larger population of which the entities are members" (Growes et al. 2009, p.2). The online survey for this dissertation was created using Google Forms. Online surveys are easy to access and distribute, as only a link is required to share the survey form. They are also easier to collect records from than from paper-based surveys, as the online format means that records can be exported to a spreadsheet for analysis. Online surveys also guarantee the anonymity of respondents and more flexible survey design.

A small risk of bias exists with the use of online surveys as it limits the study to respondents who have access to computers and the internet, and are sufficiently computer-literate – although, generally speaking, computer literacy should not be an issue for analytics users.

The survey was piloted on five volunteers, three of whom were managers and two of whom were analytics experts in non-managerial positions. Piloting the survey established that the average completion time of the survey was under 10 minutes. It was also a necessary step to check if the questions and the wording of the questions were satisfactory so that respondents would understand the questions and be able to answer them quickly and without difficulty.

After some minor adjustments resulting from feedback obtained during the piloting phase, the survey was made available online for four weeks, from May 20th, 2015 to June 23rd, 2015. There were 45 respondents following the first contact reaching out to potential respondents, and 64 respondents at the end of the data collection period after following up with a reminder/thank you email to potential respondents.

The survey was used to measure the different variables relating to the research question as well as collect some demographics about the survey respondents.

For each of the variables below, one or more questions were asked in the survey. For certain variables, more than one question was used, to make sure that the way the questions were asked did not inherently bias the respondents' answers to these questions. For example, respondents were asked for their agreement or disagreement to the following statements which both assess the respondent's opinion of the influence of

gender on the ability to use analytics: "Male managers are more skilled at using analytics." and "In general, women are less good with data/maths." These questions are similar – in this example, they are sub-questions about the perception of the relation between skill and gender, so it is interesting to consider if respondents answered them differently. Multiple sub-questions were aggregated into constructs. This makes the overall construct less prone to bias introduced by the way questions are formulated in the survey.

Gender was the most important demographic to be collected, as this research focuses on the differences in attitude to analytics of managers based on gender.

The other demographics collected in this survey were age band, industry, years of managerial experience and level of management.

The use of analytics and the attitude of managers towards analytics were measured by asking questions about the following variables, as described in Table 3.1. The main variables found to influence the adoption of and attitude to technology in the Technology Acceptance Model (TAM) (Davis 1989), the United Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al. 2003) and the Diffusion of Innovations Adoption-Decision Model (Rogers 2010) were included.

TABLE 3.1 - Variables Found to Influence Attitude to Technology

Variable	Description	Used by/Found to be significant in
Perceived Usefulness	Usefulness for job speed, quality, effectiveness	Davis 1989 (TAM), Venkatesh 2003 (UTAUT)
Ease of Use	Easiness to use and adopt, easy learning curve, user-friendliness	Davis 1989 (TAM), Venkatesh 2003 (UTAUT)
Emotional Attitude	Perception of fun, positive or negative perception, emotions	Padilla-Melendez et al. 2013 (Extended TAM), Frambach et al. 2002, Yuan et al 2004
Familiarity	Experience and knowledge of the technology	Venkatesh 2003 (UTAUT), Frambach et al. 2002, Lewis et al. 2003, Rogers 2010 (Diffusion of Innovations)
Frequency of Use	Amount of use of the technology and regularity of exposure to it	Liu & Chang 2004 (Extended TAM), Rogers 2010 (Diffusion of Innovations)
Gender Attitude	Perception of a person's skills based on their gender	Gefen and Straub 1997 (Extended TAM), Ahuja et al. 2005 (Extended TAM)

Gender Attitude Perception	How people think others are judging them based on gender	Ahuja et al. 2005 (Extended TAM)

Hypothesis testing was used to understand if there were any differences between male and female managers or whether there were no differences. For each variable, the hypothesis used was that there are no differences in the use of and attitude to analytics of managers based on gender.

3.7   Data Analysis Methods

The primary method of data analysis is quantitative. As stated above, hypothesis testing to determine differences in the measured variables based on gender was the main focus of the data analysis.

A two-sample z-test was used to determine whether there are statistically significant differences in the behaviour towards, use of and attitude to analytics of male and female managers, at both a 90% confidence level and a 95% confidence level. This was done first on each construct, and then – in cases where no differences were statistically significant with at a 90% confidence level for the construct – for each sub-variable of the given construct.

The data analysis was done using a Python script. This script relies on libraries for data analysis (Numpy, Scipy, Pandas) for Python. Excerpts of the code are available in Appendix C.

The script calculated the inter-correlation of variables first. The script generated variable-by-variable heat-maps from this data using the matplotlib library.
Cronbach's alpha scores were then calculated for each construct using a custom function.

A metric, the Gini score, was used to determine which of the variables showed the most differences between the attitude of male and female managers to analytics in terms of distribution. The Gini score is a metric primarily used in economics to measure inequality, but is also used for model development and profiling exercises in consumer analytics.

The script then calculated the correlation of each variable to gender using a two-sample z-test for each of the constructs, and then, for each of the sub-variables. It also calculated

the p-value of for the corresponding z-score and determined if there were any differences at a 90% confidence and at a 95% confidence level.

Two free text fields were available in the survey for further comments of the respondents, which were then used additionally to the quantitative data to better understand some of the results of the quantitative analysis and illustrate the views of respondents on the role of gender in the use of analytics by managers. Word count frequencies were plotted in word clouds, so as to better understand the overarching themes mentioned by managers of each gender.

3.8   Limitations

As mentioned in Section 3.6, the sampling methods and size have an impact on the reliability and generalisability of the results of this dissertation. In particular, snowball sampling and smaller sample sizes make the results of quantitative statistical analysis less reliable, as the data gathered may be biased and therefore only represent a section of the population.

The data collection method, namely online surveys, can introduce also introduce a bias because the methods itself excludes people who have chosen not to answer the survey and people who do not have the access to the technology or the know-how required to complete a survey online. In the case of this dissertation, the bias due to access to technology should be minimal, as the use of analytics usually also requires access to and skills in technology, especially in large companies, and, therefore, the population surveyed would generally have the means to respond to an online survey.

A third limitation of this study is the method of data collection and analysis. Although quantitative analysis of the influence of the different studied variables gives a good overview of which variables relating to the use and attitude towards analytics are significantly correlated to gender, the free text fields were very useful in illustrating the points highlighted by the quantitative analysis further. Given more time, it would have been of interest to gather more free text data, conduct interviews with managers or facilitate a focus group to understand the attitude to analytics of individual managers better, and try to identify the underlying reasoning behind their view of the importance of gender in analytics.

3.9   Lessons learned

Although the survey ethics were approved in May, it took until the end of June to complete the survey, as it was hard to contact managers through their team members and other managers. The survey was extended by ten days from June 10th, 2015 to June 23rd, 2015 so as to allow the survey to reach a larger number of respondents and give managers enough time to respond between the reception of the survey, possibly second or third-hand from a colleague as snowball sampling was used, and the end of the survey period. Reminder emails were sent out to remind direct contacts to respond and to contact their network, but it was not possible to send out reminders to all contacted managers, as the reminder emails were not 'snowballed' like the initial survey participation email was. Managers are extremely busy and therefore time constraints impacted the participation in the survey negatively, as several respondents mentioned having delayed their response due to their workload or to a lack of time.

3.10   Conclusion

This chapter gave an overview of the different research methodologies and identified the research methodology used in this dissertation. The research philosophy used is pragmatism, using a deductive approach.

The analysis consists of quantitative analysis on data gathered through an online survey using a two-sample z-test, supplemented by qualitative analysis of free-text comments to add further insight into managers' attitude to analytics. This study is within a cross-sectional time horizon and data were collected over a four-week period.

The variables studied for this analysis were based on variables used in technology use, acceptance and adoption models such as the TAM, UTAUT and Diffusion of Innovations Adoption-Decision models. Responses to survey questions were gathered for sub-variables of each of the variables considered and aggregated into constructs.
Hypothesis testing is done first on a construct-level, and then, where no differences based on gender are highly statistically significant on construct-level, on a sub-variable level.
To better understand if male and female managers perceive any differences in their or other managers' use of analytics based on gender, the free-text responses are then analysed using qualitative methods. Word frequency analysis is used to forge a better understanding of the main themes highlighted by managers of each gender.

4.   Findings and Analysis

4.1   Introduction

The aim of this research was to determine whether male and female managers have a different attitude to analytics. The survey data is analysed primarily using quantitative methods. A total of 64 managers participated in the survey. One was discounted, as the respondent did not specify their gender. 63 respondents remained to be analysed in the findings below.
There were 28 female managers and 35 male managers in the final dataset.
For any results shown as counts by gender in the sections below, there will, therefore, be a larger total number of male respondents than female respondents. To account for the difference in sample size by gender, percentages will be used where possible.

Section 4.2 will describe the demographics recorded during the survey, and explain their correlation to gender and why no third variable impact is assumed for the rest of the chapter.

Section 4.3 explains how the recorded responses were converted to scores based on a Likert scale, and how the constructs were calculated.

Section 4.4.1 shows a preliminary analysis of the differences between the groups based on the Gini score of each studied variable by gender.

Section 4.4.2 explains the results of hypothesis testing done for each variable and explores the distribution of variables found to be statistically significantly different based on gender.

Section 4.5 investigates the overall perception of the role of gender in analytics of managers.

Section 4.6 supplements the findings with analysis of the free-text comments of respondents and studies which topics managers were most likely to mention, based on gender.

4.2   Demographics of the Surveyed Managers

Several questions in the survey were used to determine the profile of the surveyed managers. This was done to make sure that male or female managers were not over- or under-represented and that managers from varied backgrounds were surveyed, as well as provide a better overview of the surveyed managers. Understanding if any clusters of either gender existed in any demographics was important to make sure that the differences between male and female managers were gender-based and not impacted by a third variable. The most important demographic recorded was 'Gender', as this is the variable whose differences and impact on other variables are being measured here.

The other demographics measured were the following (see Appendix G for more details on the categories for each demographic variable):

- Age Band
- Industry
- Management Level
- Technology Adoption Speed
- Years of Management Experience

As the rest of the analysis will be based on the assumption that other demographics have minor to no influence on the differences between genders, it is essential to assess whether there were any differences between genders based on the demographics.

To assess if there were any differences between genders, the Gini score, which is based on the difference in two distributions, was used. The Gini score is a measure of inequality used in economics, for example, to assess levels of inequalities in different countries (e.g. differences in level of income, wealth, education). This is also used in other fields such as marketing to assess differences between distributions (e.g. how different two groups of customers are, based on different metrics). The Gini score goes from 0 (perfect equality, e.g. if the age distributions of male and female managers were exactly the same) to 1 (perfect inequality, e.g. if all male managers were over 40 and all female ones under 40). This means that the higher the Gini score is, the more different the distributions are.

TABLE 4.1 - Gini Scores and Equality

Gini Score	Equality
0	Perfect Equality

0-0.2	High Equality/Low Inequality
0.2-0.3	Moderate Equality/Moderate Inequality
0.3-0.4	Low Equality/Medium Inequality
0.4-0.5	High Inequality
>0.5	Extreme Inequality

Here, Gini is used to determine how different distributions of a demographic are when split by gender (e.g. distribution of the age of female managers vs. distribution of the age of male managers as in the examples above).

TABLE 4.2 - Gini Scores of Demographics by Gender

Variable	Gini
Management Level	0.392
Industry	0.387
Age Band	0.298
Tech Adoption Speed	0.268
Years of Management Experience	0.217

Since the scores for 'Management Level' and 'Industry' seem to reveal moderate differences in the distribution of female managers and male managers, it is important to further investigate the differences between demographics based on gender and understand whether any of the demographics are correlated significantly to 'Gender'.

To assess the correlation of each of the demographics to 'Gender', Pearson's correlation was used. This also enabled a visualisation of the inter-correlation of demographics to each other. The Pearson product-moment correlation coefficient (also known as Pearson's r or Pearson's correlation), which was used to determine the relationships between the variables - in particular with 'Gender' - two-by-two, assesses the correlation's strength.

TABLE 4.3 - Pearson's Correlation Scores

Absolute Value	Correlation
0-0.3	Low
0.4-0.5	Medium
0.6-1	High

For the purpose of visualising the inter-correlation of the demographics on the heat-map below (Figure 4.1), the absolute value of the correlation is shown, as the strength of the correlation is of importance here, and not necessarily the direction.
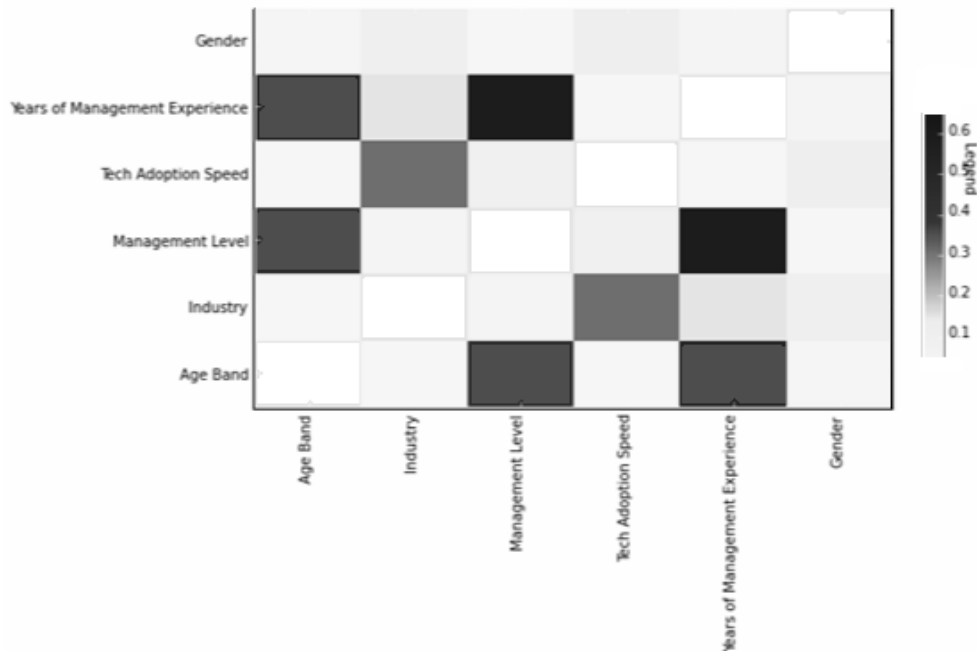
FIGURE 4.1 - Inter-correlation of Demographics

All of the demographics recorded had quite low correlations to 'Gender' using Pearson's correlation.

The demographic with the most difference based on gender was 'Industry', which also had the largest standard deviation from the mean. This may be due to the facts that the small sample size of the dataset made it harder to survey managers equally in a variety of different industries and also that different industries have different male to female employee ratios (CSO 2013).

Although gender does not seem be highly correlated to the demographics, other variables studied or the constructs based on them may be highly correlated to the demographics. Pearson's correlation was therefore also measured for the other variables used in this analysis.
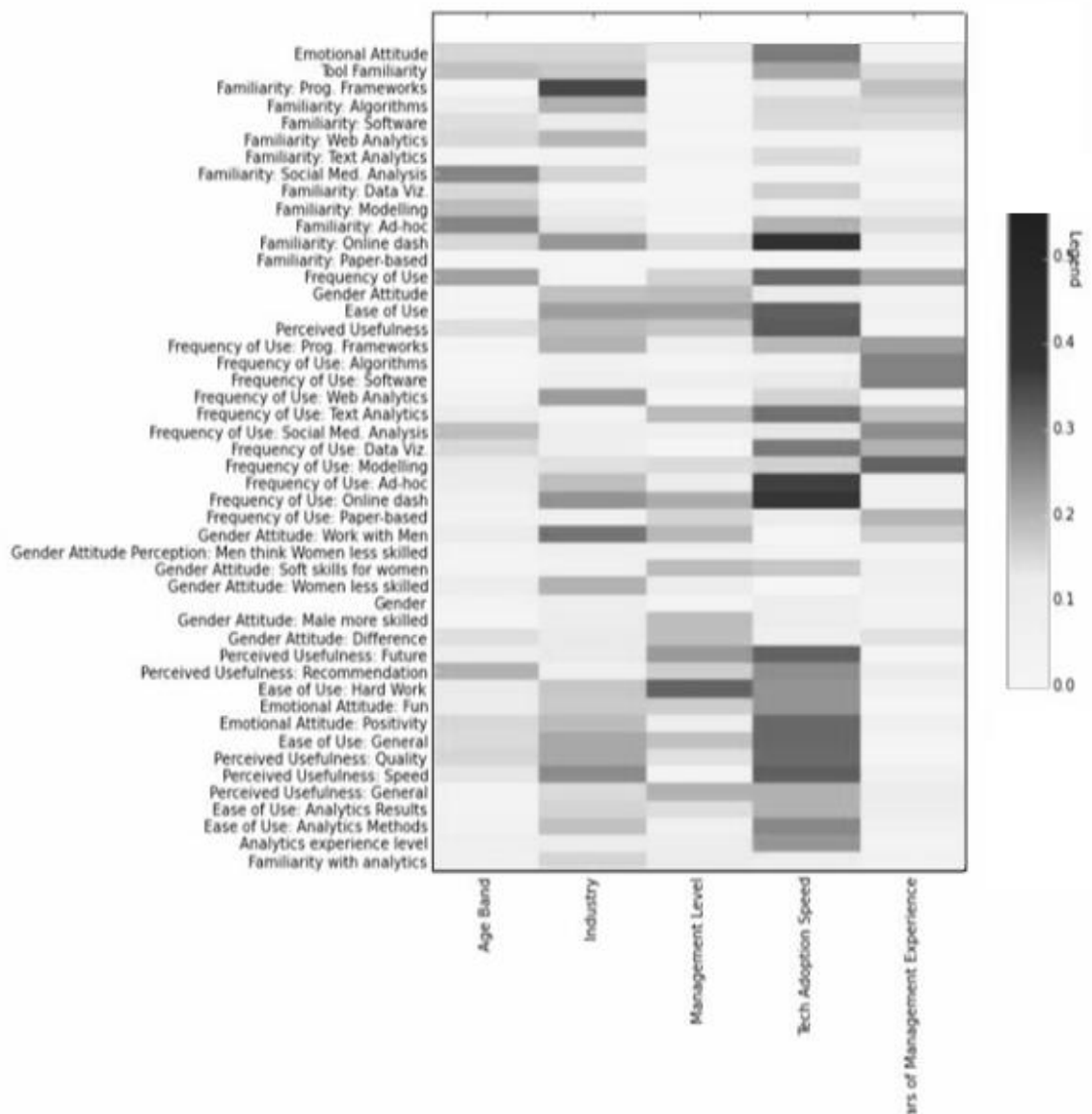
FIGURE 4.2 - Inter-correlation of Demographics with Other Variables

The only demographic that was highly correlated (|r| > 0.4) to one of the variables used in the sections of the analysis below was 'Tech Adoption Speed', which had a correlation of r = 0.419 to 'Familiarity: Online dashboards'. This may be indicative of the fact that people who are likely to adopt technology earlier than others may also be those who are more familiar with technology advances, and may therefore be more likely to be aware of the existence of online dashboards.

With the exception of the inter-correlation of these two variables, the impact of the demographics on the other variables will be considered negligible for the purposes of this analysis.

Given the size of the sample used for this study, it is not possible to accurately assess the impact of cross-correlation on gender differences (e.g. male managers under 40's

frequency of use of analytics vs. female managers over 60's frequency of use of analytics), so the rest of the analysis will focus solely on the differences between male and female managers based on the variables mentioned above.

As correlation between both gender and any of the demographics, and the other variables and the demographics is low, there are no large clusters of respondents by gender in any of the demographics. The impact of demographics as third variables will, therefore, be assumed to also be low.

4.3   Calculating Constructs and Scaling Scores

As mentioned in Section 3.4, differences in the distribution of female and male managers across several constructs will be analysed in this chapter.

The constructs were measured as follows:

● Perceived Usefulness

This construct was measured using several sub-variables that highlighted different aspects of 'usefulness' of a technology, such as, for example, improved quality of work and speed of work. The sub-variables used were perceived usefulness in general ('Perceived Usefulness: General'), usefulness in terms of accelerating the speed of outputs and making it faster to do reliable work ('Perceived Usefulness: Speed'), usefulness in terms of improved quality of work ('Perceived Usefulness: Quality'), potential usefulness for others in the same industry ('Perceived Usefulness: Recommendation') and perception of analytics as a method to be adopted further in the future ('Perceived Usefulness: Future'). The construct was calculated as an average of these individual measures and called 'Perceived Usefulness'. The scores were recorded on a Likert scale and converted as shown below.

TABLE 4.4 - Scale Item Score Conversions

Scale Item	Score
Strongly Disagree	0
Disagree	1
Do not know	2
Agree	3
Strongly agree	4

● Ease of Use

This construct was measured based on the following sub-variables: ease of use of the outputs and results of analytics ('Ease of Use: Analytics Results'), ease of use of analytics methods ('Ease of Use: Analytics Methods'), ease of use of analytics in general ('Ease of Use: General') and the perception that using analytics is hard ('Ease of Use: Hard Work'). The construct 'Ease of Use' was calculated as an average of the four sub-variables. The same scale items and score conversions as for Perceived Usefulness were used (see Table 4.4).

● Emotional Attitude

Emotional Attitude to analytics was a measure of the perception that analytics are fun ('Emotional Attitude: Fun') and that positive feeling are associated to analytics ('Emotional Attitude: Positivity'). The construct 'Emotional Attitude' was calculated as the average of the two. The same scale items and score conversion as for Perceived Usefulness were used (see Table 4.4).

● Frequency of Use

'Frequency of Use' was measured as the average of the frequencies of use of the following tools or analytics techniques: paper-based reports, online dashboards, *ad hoc* analysis, modelling/model scores, data visualisations (maps, charts, etc.), social media analysis, text analytics, web analytics (such as Google Analytics or Webtrends), software (such as SAS, SPSS), algorithms (e.g. clustering, decision trees), and programming frameworks (e.g. R, Python). The frequencies of use were measured on a range from 'Never' to 'Multiple times a day', shown below in Table 4.5.

TABLE 4.5 - Scale Item Scores for Frequency of Use

Scale Item	Score
Never, and I am not familiar with this.	0
Never, but I am familiar with this.	1
Less than once a month.	2
About once a month.	3
A few times a month.	4
About once a week.	5
Several times a week.	6
About once a day.	7
Several times a day.	8

●   Familiarity

As two options were given above for 'Never' (see Table 4.5), it was possible to assess the familiarity of managers with each tool, as it was possible to differentiate people who were not familiar with a technology ('Never, and I am not familiar with this.') from others who either did not use the technology but were familiar with it ('Never, but I am familiar with this.') or had used the technology before.

The scores given were 0 for unfamiliar and 1 for familiar with a tool. The construct 'Tool Familiarity' was calculated as the average of the familiarity of managers with each tool.

In addition to this, perceived familiarity was also recorded as 'Familiarity with analytics'. Managers selected the level of familiarity they identified with most in terms of their familiarity with analytics going from 'Not at all familiar' (0) to 'Extremely familiar' (4) - with 'Slightly familiar' (1), 'Moderately familiar' (2) and 'Very familiar' (3) as intermediate options.

'Analytics experience level' was also recorded and managers indicated how experienced they were with analytics and what kind of analytics user they felt they were – which was useful to separate technical experts from end-users of analytics. The scale used went from complete novice (0) to technical expert (4), with end-user without technical knowledge (1), end-user with a minor understanding of statistical methods (2) and analytics-savvy end-user (3) as intermediate options.

●   Gender Attitude

Gender attitude variables were different measures of the importance of gender in the interaction managers had with analytics or their perception of the skills of others in analytics based on gender. The sub-variables for 'Gender Attitude' were: the perception that there was a difference in how people used analytics based on their gender ('Gender Attitude: Difference'), that men were more skilled than women in mathematics or statistics ('Gender Attitude: Male more skilled'), that women were better suited to jobs focused on soft skills ('Gender Attitude: Soft skills for women'), that women were less skilled at analytics ('Gender Attitude: Women less skilled') and finally, the managers' opinion on whether they would rather work on analytics with a man than a woman ('Gender Attitude: Work with men').

The construct 'Gender Attitude' was calculated as an average of the sub-variables. These were measured on the Likert scale as described in Table 4.4.

●   Gender Attitude Perception

'Gender Attitude Perception' was a single variable measuring what managers perceived others' gender attitude to be, based on gender. This was measured as 'Gender Attitude Perception: Men think Women less skilled', which measured whether managers considered men to be more likely to think women were less skilled at analytics. These were measured on the Likert scale as described in Table 4.4.

4.3.1 Scaling Scores

To record the answers of the respondent's use, perception and attitude towards analytics, multi-item scales were used. The most common scale used was a Likert scale (see Table 4.4). A 'Not Applicable' option was added so as not to skew results if a certain situation did not apply to a particular respondent.

4.3.2 Rollups

To assess the reliability of the constructs described above, Cronbach's Alpha was used. This is a measure that estimates the lower bound of reliability of a multi-point scale test, by assessing the overall variance and the covariance of items two by two, and therefore ensuring the consistency of scores for rolled up variables. Any score with $\alpha > 0.7$ is considered good, with scores over 0.9 being considered excellent.

TABLE 4.6 - Cronbach's Alpha Scores

Variable	# of sub-variables	Score	Score if item deleted
Perceived Ease of Use	**4**	**0.824**	
Ease of Use: General			0.734
Ease of Use: Analytics Methods			0.775
Ease of Use: Analytics Results			0.731
Ease of Use: Hard Work			0.866
Emotional Attitude	**2**	**0.848**	
Emotional Attitude: Fun			N/A
Emotional Attitude: Positivity			N/A
Perceived Usefulness	**5**	**0.912**	
Perceived Usefulness: Future			0.922
Perceived Usefulness: General			0.891
Perceived Usefulness: Quality			0.871
Perceived Usefulness: Recommendation			0.886
Perceived Usefulness: Speed			0.893
Gender Attitude	**5**	**0.859**	
Gender Attitude: Difference			0.84
Gender Attitude: Male more skilled			0.812
Gender Attitude: Soft skills for women			0.848

Gender Attitude: Women less skilled			0.805
Gender Attitude: Work with Men			0.841
Frequency of Use	**11**	**0.843**	
Frequency of Use: Ad hoc			0.821
Frequency of Use: Algorithms			0.82
Frequency of Use: Data Visualisation			0.819
Frequency of Use: Modelling			0.816
Frequency of Use: Online dashboards			0.836
Frequency of Use: Paper-based			0.86
Frequency of Use: Prog. Frameworks			0.827
Frequency of Use: Social Med. Analysis			0.832
Frequency of Use: Software			0.817
Frequency of Use: Text Analytics			0.823
Frequency of Use: Web Analytics			0.847
Familiarity	**11**	**0.806**	
Familiarity: Ad hoc			0.795
Familiarity: Algorithms			0.747
Familiarity: Data Viz.			0.799
Familiarity: Modelling			0.794
Familiarity: Online dash			0.79
Familiarity: Paper-based			0.814
Familiarity: Prog. Frameworks			0.783
Familiarity: Social Med. Analysis			0.803
Familiarity: Software			0.752
Familiarity: Text Analytics			0.786
Familiarity: Web Analytics			0.807

All the scores for the constructs had α > 0.7, so they are reliable in terms of their validity to represent a total view of their individual components.

A further way to assess the reliability of the constructs is to ensure that each construct is highly correlated to its sub-variables. This is visible in the following heat-map, where Pearson's correlation was performed on the dataset. The inter-correlation of constructs and their sub-variables is highlighted in the grey boxes in Figure 4.3.
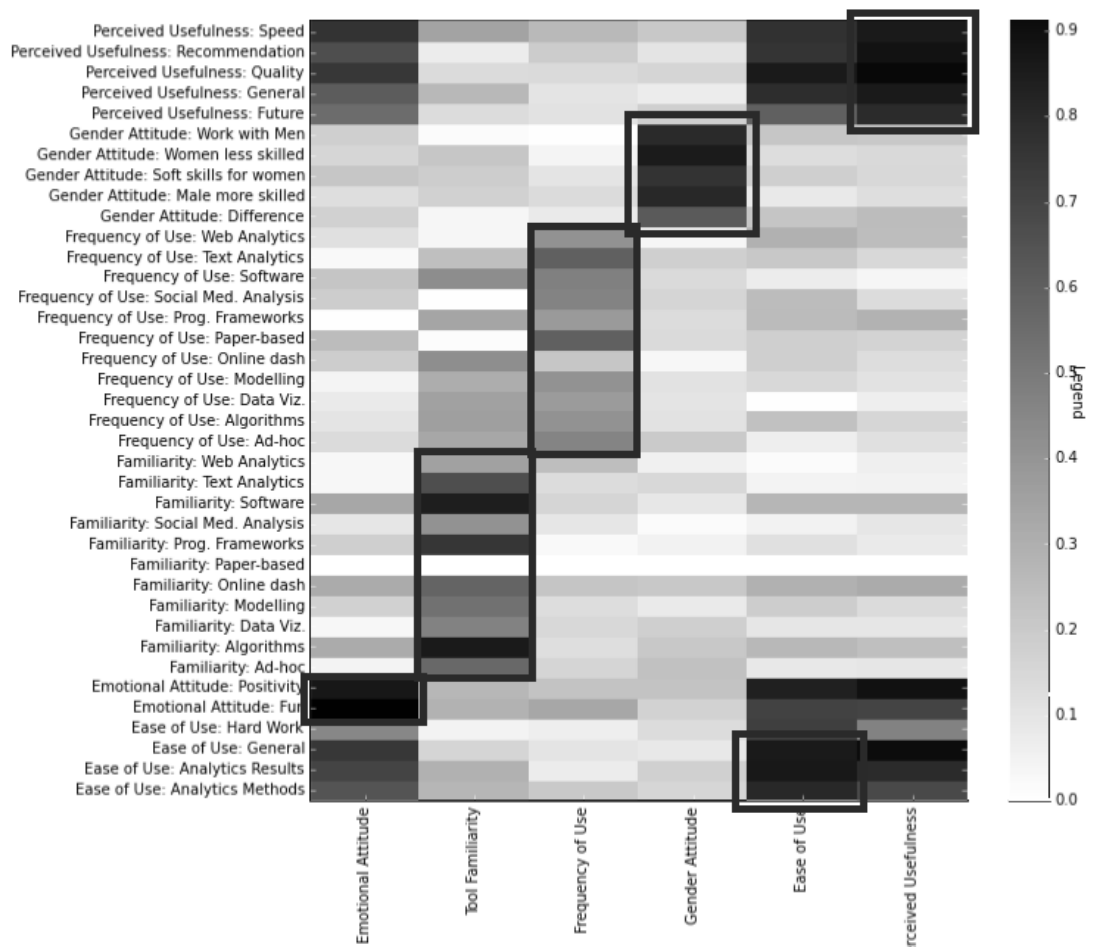
FIGURE 4.3 - Inter-correlation of Constructs and their Sub-variables

For all of the constructs and their sub-variables, there is also a strong correlation between the construct and its sub-variables. This confirms their reliability as measurements for this study.

4.4   Assessing the Difference between the Groups

4.4.1 High-level Differences in Distribution

This analysis compares two groups - male managers and female managers - to understand which variables recorded best show the difference between the groups. To do this, the Gini coefficient was used to give a first high-level understanding of which variables may be most correlated to gender and may be most different based on gender.

The top variables (constructs and sub-variables) with the most different distributions based on gender (Gini > 0.2) are shown below in Table 4.7. A full list of Gini scores can be found in Appendix D.

TABLE 4.7 - Gini Scores where Gini > 0.2

Variable	Gini
Tool Familiarity	0.601
Frequency of Use: Data Visualisations	0.436
Gender Attitude: Difference	0.395
Frequency of Use: Paper-based	0.373
Familiarity with analytics	0.344
Perceived Usefulness: Future	0.339
Frequency of Use: Social Med. Analysis	0.321
Frequency of Use: Ad hoc	0.312
Emotional Attitude	0.297
Frequency of Use: Software	0.281
Gender Attitude: Women less skilled	0.268
Analytics experience level	0.247
Emotional Attitude: Positivity	0.233
Ease of Use	0.222
Gender Attitude	0.216
Frequency of Use: Modelling	0.204

In terms of frequency distributions, there seem to be some differences in distribution between the male and female managers for four constructs ('Tool Familiarity', 'Emotional Attitude', 'Ease Of Use' and 'Gender Attitude'), as well as several sub-variables (especially 'Gender Attitude: Difference' and 'Frequency of Use: Data Visualisations').

For all of the constructs except 'Gender Attitude Perception', either the construct or one of its sub-variables appears in this list and therefore has a Gini > 0.2.

4.4.2 Hypothesis Testing

To investigate this further and assess the significance of the differences based on gender for each individual variable, hypothesis testing was done, using a two-sample z-test (with sample sizes of 35 and 28 for male managers and female managers respectively).

When there was no significant difference on construct-level, individual differences on sub-variable level were studied.

For each table shown, 'sign90' represents whether the differences between male and female managers based on the studied variable are significant at a 90% confidence level and 'sign95' represents whether the differences based on gender are significant at a 95% confidence level. The p-values calculated are two-sided, as a deviation from the mean to either side of the distribution would enable a rejection of the null hypothesis.

Perceived Usefulness

H_0: *There is no difference between how male managers and female managers perceive the usefulness of analytics.*

TABLE 4.8 - Perceived Usefulness z-test Scores

Variable	z-score	p-val	sign90	sign95
Perceived Usefulness	-1.22	0.223929	No	No
Perceived Usefulness: Future	0.17	0.863837	No	No
Perceived Usefulness: General	-0.81	0.417771	No	No
Perceived Usefulness: Quality	-1.89	0.059063	Yes	No
Perceived Usefulness: Recommendation	-1.95	0.051644	Yes	No
Perceived Usefulness: Speed	-1.13	0.258391	No	No

For the construct 'Perceived Usefulness', there were no significant differences in distribution between male and female managers at a 90% confidence level.

On the sub-variable level, for Perceived Usefulness: Quality' (usefulness for improving quality of work), as well as 'Perceived Usefulness: Recommendation' (the perception that analytics would be useful to others in a similar industry), the null hypothesis is rejected with a 90% confidence.
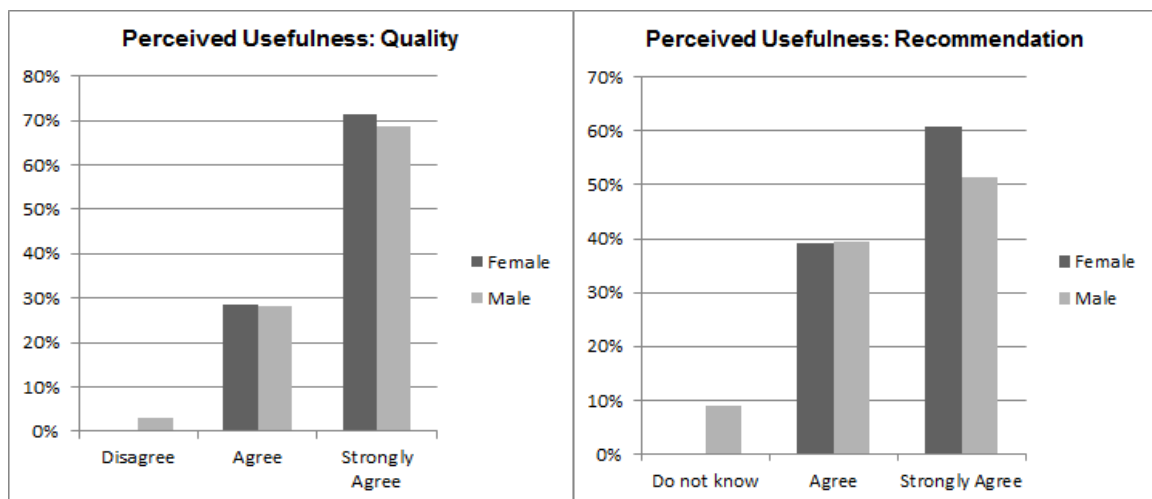

FIGURE 4.4 - Distribution of Perceived Usefulness: Quality and Perceived Usefulness: Recommendation

For both 'Perceived Usefulness: Quality' and 'Perceived Usefulness: Recommendation', female managers were, on average, more likely to agree with the statements about the usefulness of analytics. Female managers found analytics useful for improving the quality

of their work, with an average of 3.71 (Strongly Agree) vs. 3.23 (Agree) for male managers, and were more likely to agree that they would recommend it to others in the same industry with an average of 3.61 (Strongly Agree) vs. 3.17 (Agree) for male managers. 4% more female managers strongly agreed that analytics improved their quality of work than their male counterparts, and 18% more female managers strongly agreed that they would recommend analytics to others than their male counterparts. In each case, all of the female managers agreed to the statements whereas some of the male managers (3% for quality and 9% for recommendation) did not agree.

This difference in distribution is not as pronounced for the other 'Perceived Usefulness' variables, even though for all but 'Perceived Usefulness: Future', the z-score is negative, indicating that female managers were also slightly more likely to agree with other statements on perceived usefulness, albeit not differently enough to their male counterparts to be significant with a 90% confidence level.

Perceived Ease of Use

H_1: *There is no difference between how male and female managers perceive analytics ease of use.*

TABLE 4.9 - Perceived Ease of Use z-test Scores

Variable	z-score	p-val	sign90	sign95
Ease of Use	-1.06	0.288374	No	No
Ease of Use: Analytics Methods	0.15	0.877929	No	No
Ease of Use: Analytics Results	-0.93	0.35265	No	No
Ease of Use: General	-2.01	0.044873	Yes	Yes
Ease of Use: Hard Work	-0.7	0.48159	No	No

Although overall, the difference between responses from male and female managers was not significantly different with 90% confidences and H_1 cannot be rejected on construct-level, for one of the sub-variables, 'Ease of Use: General' (with the statement: "Analytics make it easier to do my job."), the hypothesis H_1 is rejected at a 95% confidence level.
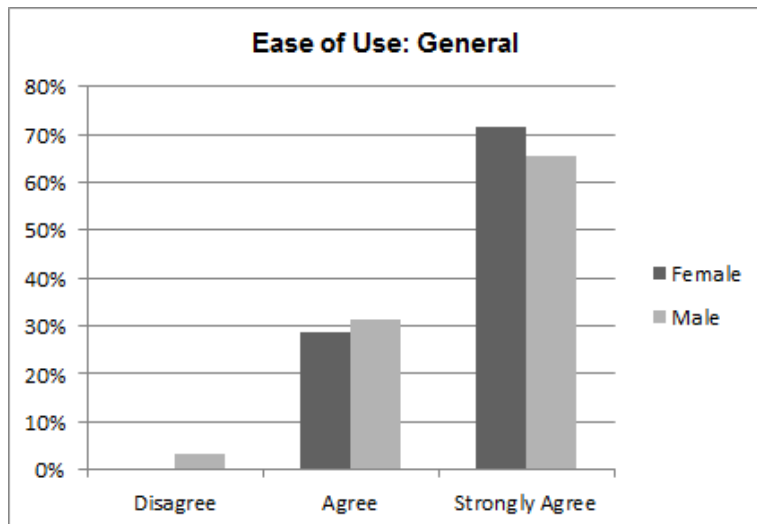
FIGURE 4.5 - Ease of Use: General

Here again, as for the 'Perceived Usefulness' variables, female managers were more likely to agree to the statement, with an average of 3.71 (Strongly Agree) vs. 3.2 (Agree) for male managers. Female managers were 9% more likely to strongly agree with the statement and 9% less likely to only agree than their male counterparts.

Frequency of Use

H_2: *There is no difference between how frequently male and female managers use analytics.*

TABLE 4.10 - Frequency of Use z-test Scores

Variable	z-score	p-val	sign90	sign95
Frequency of Use	0.07	0.943834	No	No
Frequency of Use: Ad hoc	-0.29	0.771496	No	No
Frequency of Use: Algorithms	2.03	0.042297	Yes	Yes
Frequency of Use: Data Viz.	-1.4	0.162552	No	No
Frequency of Use: Modelling	0.35	0.727199	No	No
Frequency of Use: Online dash	-0.16	0.871894	No	No
Frequency of Use: Paper-based	-0.19	0.846171	No	No
Frequency of Use: Prog. Frameworks	2.45	0.0144	Yes	Yes
Frequency of Use: Social Med. Analysis	-1.09	0.274421	No	No
Frequency of Use: Software	1.41	0.158111	No	No
Frequency of Use: Text Analytics	-0.28	0.777303	No	No
Frequency of Use: Web Analytics	0.79	0.432075	No	No

'Frequency of Use' did not show any significant differences between male and female managers at a 90% confidence level.

For 'Frequency of Use' sub-variables, the z-scores ranged from -1.09 (for 'Frequency of Use: Social Media Analysis') to 2.45 (for 'Frequency of Use: Programming Frameworks'). For two of the sub-variables, 'Frequency of Use: Algorithms' and 'Frequency of Use: Programming Frameworks', the hypothesis H_2 is rejected at a 95% confidence level. For both of these tools, male managers were more likely to use the tools more frequently than their female counterparts.

As industry was moderately correlated to 'Frequency of Use', it might be possible that this is due to the fact that there was a larger share of men working in STEM industries than there was of women in the sample (which is also the case, though to a lesser degree, in the general population (CSO 2013)).

TABLE 4.11 - Tech/Non-tech Industry by Gender

Gender	Female	Male
Non-tech industry	46%	26%
Tech industry	54%	74%

There were 20% more male managers working in technology-related industries in the sample than female managers working in technology-related industries, which may skew the differences by gender.

TABLE 4.12 - Distribution of the Variables 'Frequency of Use: Programming Frameworks' and 'Frequency of Use: Algorithms' by Industry Type by Gender

Variable	Frequency of Use: Algorithms		Frequency of Use: Prog. Frameworks	
Gender	Female	Male	Female	Male
Non-tech industry	1.17 (Never, but I am familiar with this.)	2.44 (Less than once a month)	0.54 (Never, and I am not familiar with this.)	1.78 (Less than once a month)
Tech industry	1.93 (Less than once a month)	2.84 (About once a month)	1.47 (Never, but I am familiar with this.)	2.54 (About once a month)

Indeed, across both genders, people working in technology-related industries were more likely to use analytics more frequently. For each industry-type though, 'Frequency of Use: Algorithms' and 'Frequency of Use: Programming Frameworks' are still higher for male managers than for female managers by approximately one scale item, so the difference in 'Frequency of Use: Programming Frameworks' was not due only to industry, but rather to gender.

Tool Familiarity

H_3: *There is no difference between how familiar male and female managers are with analytics tools.*

TABLE 4.13 - Tool Familiarity z-test Scores

Variable	z-score	p-val	sign90	sign95
Tool Familiarity	1.41	0.159686	No	No
Familiarity: Web Analytics	1.47	0.142213	No	No
Familiarity: Text Analytics	1.07	0.282906	No	No
Familiarity: Software	1.07	0.282906	No	No
Familiarity: Social Med. Analysis	1.02	0.308512	No	No
Familiarity: Prog. Frameworks	1.38	0.168185	No	No
Familiarity: Paper-based	0.00	1.000000	No	No
Familiarity: Online dash	0.16	0.873818	No	No
Familiarity: Modelling	-0.41	0.683862	No	No
Familiarity: Data Viz.	0.16	0.873818	No	No
Familiarity: Algorithms	1.42	0.15697	No	No
Familiarity: Ad hoc	1.02	0.308512	No	No

Even though there are significant differences in 'Frequency of Use' of analytics tools and methods, there does not seem to be a significant difference at a 90% confidence level for the construct 'Tool Familiarity'. It is not possible to reject H_3 for 'Tool Familiarity' or any of its sub-variables.

Analytics Experience Level

H_4: *There is no difference between how much analytics experience male and female managers have.*

TABLE 4.14 - Analytics Experience Level z-test Scores

Variable	z-score	p-val	sign90	sign95
Analytics experience level	1.43	0.151618	No	No

In line with the fact that no significant difference between male and female managers can be confirmed at a 90% confidence level, there is also not possible to reject the hypothesis H_4 and identify any gender-based differences at a 90% confidence for 'Analytics Experience Level'.

Familiarity with Analytics

H_5: There is no difference between how familiar male and female managers perceive analytics ease of use.

TABLE 4.15 - Familiarity with Analytics z-test Scores

Variable	z-score	p-val	sign90	sign95
Familiarity with analytics	2.18	0.029568	Yes	Yes

The 'Familiarity with Analytics' – the level of familiarity with analytics managers perceived themselves to have – was significantly different at a 95% confidence level. It is therefore possible to reject the hypothesis that there is no difference in how familiar male and female managers perceive themselves to be with analytics, with female managers perceiving themselves to be less familiar with analytics than their male counterparts.

Even if the differences between male and female managers are not significant at 90% confidence for the 'Tool Familiarity' and 'Analytics Experience Level' variables, these also showed the same trend towards male managers scoring higher than female managers on average (but to a lesser degree).
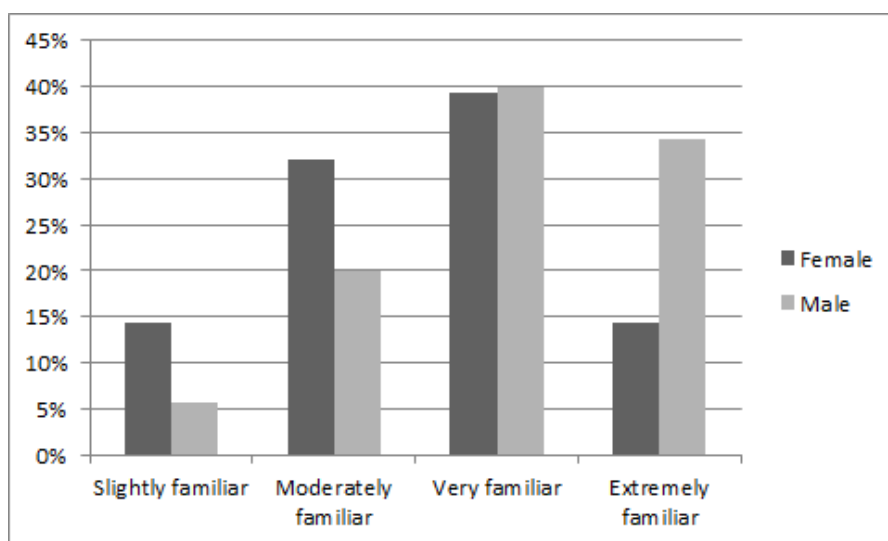

FIGURE 4.6 - Distribution of Familiarity with Analytics

Even if a similar share of the female and male managers responded that they were "Very familiar" with analytics, all other buckets show strong differences between male and female managers. Female managers are 2.5 times more likely to be only slightly familiar with analytics and 1.61 times more likely to be only moderately familiar with analytics, which in turn means that male managers are 2.4 times more likely to be extremely familiar with analytics.

One respondent mentioned that the difference in level of analytics know-how may be due to education:

> "The challenge comes from the number of female [managers] who are formally trained / educated in related fields ([computer] science, maths, statistics, data management, analytics). Therefore [there are fewer] instances of working with women in this field [...]."

This implies not only that female managers may be less likely to have received education relating to analytics but, as there are fewer women in the field, that there may be lesser visibility of equally skilled female managers. Some feel that this will rectify itself over time, as analytics become even more commonplace, and that analytics may even help to overcome perception issues:

> "Gender may have an impact in doing analytics but I think using [analytics] should help managers feel confident when [they] have numbers to back-up decisions [regardless of gender]."

Emotional Attitude

H_6: *There is no difference between the emotional attitude to analytics of male and female managers.*

TABLE 4.16 - Emotional Attitude z-test Scores

Variable	z-score	p-val	sign90	sign95
Emotional Attitude	-0.71	0.476462	No	No
Emotional Attitude: Fun	-0.45	0.655558	No	No
Emotional Attitude: Positivity	-1.67	0.095581	Yes	No

There were no significant differences between the emotional attitude of male and female managers on construct-level, so it was not possible to reject H_6 for the 'Emotional Attitude'

construct. The only sub-variable for which the difference based on gender was significant at a 90% confidence level was 'Emotional Attitude: Positivity', for which the respondents indicated their agreements to "In general, I like using analytics for my job".
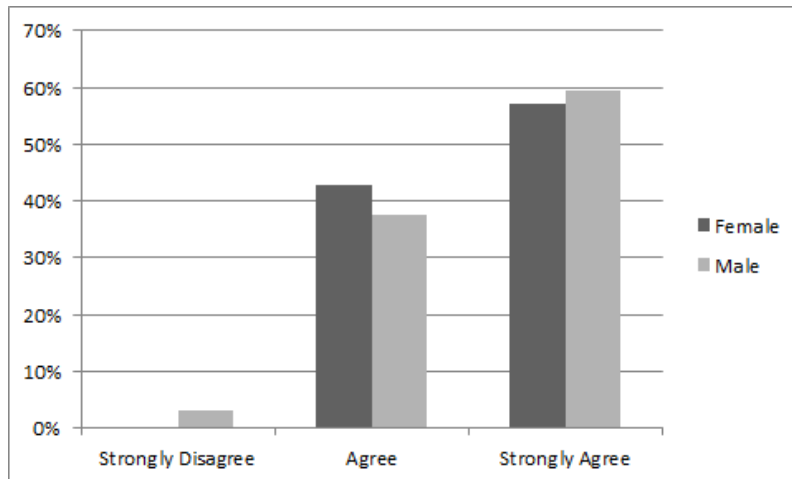

FIGURE 4.7 - Distribution of 'Emotional Attitude: Positivity'

Although the only respondent who did not agree that they liked using analytics was male and female managers had a higher average response, a larger share of male managers strongly agreed with the statement. Overall, the majority of both genders strongly agreed with the statement.

Gender Attitude

H_7: There is no difference between the perceived importance of gender in analytics of male and female managers.

TABLE 4.17 - Gender Attitude z-test Scores

Variable	z-score	p-val	sign90	sign95
Gender Attitude	0.08	0.934395	No	No
Gender Attitude: Difference	-2.94	0.00332	Yes	Yes
Gender Attitude: Male more skilled	0.33	0.740656	No	No
Gender Attitude: Soft skills for women	0.1	0.92126	No	No
Gender Attitude: Women less skilled	1.44	0.148688	No	No
Gender Attitude: Work with Men	-0.93	0.351611	No	No

The construct 'Gender Attitude' did now show any significant differences based on gender at a 90% confidence level.

At a more detailed level, 'Gender Attitude: Difference' showed significant differences. This sub-variable had the highest absolute z-score of all the studied variables and was the variable for which male managers and female managers' responses were the most different. H_7 is rejected for this variable at a 95% confidence level, although the z-scores were not high enough to do so for any of the other 'Gender Attitude' variables. The statement for this variable was "Male and female managers use analytics differently".
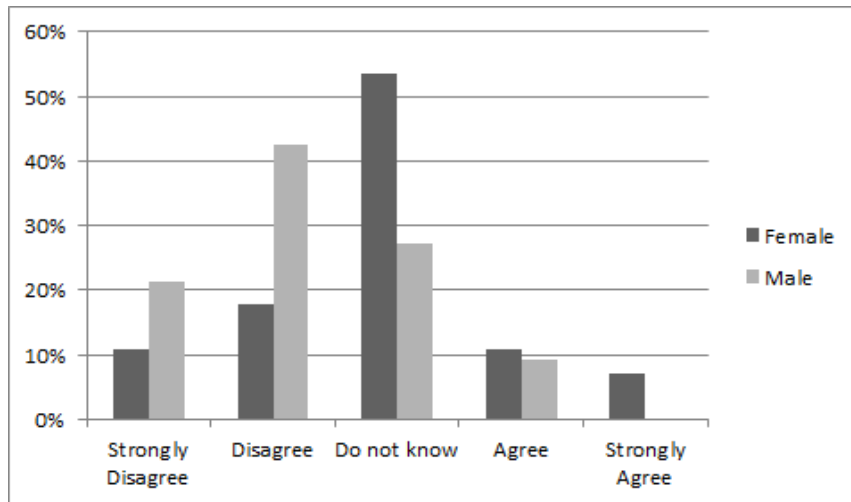

FIGURE 4.8 - Distribution of 'Gender Attitude: Difference'

The majority of female managers responded that they were unsure whether there was a difference between male and female managers, with over 50% of female respondents in the 'Do not know' bucket. The majority of male managers, to the contrary, did express an opinion on the questions and more than 60% of male managers responded either that they disagreed or strongly disagreed with the statement. It is also interesting to see that only female respondents indicated that they strongly agreed with this statement. The fact that women and men are equally skilled at using analytics was also more frequently mentioned in the free-text area comments by men than by women (see Section 4.7 below).

Gender Attitude Perception

H_8: *There is no difference between how male and female managers perceive other's attitude to the importance of gender in analytics.*

TABLE 4.18 - Gender Attitude Perception z-test Scores

Variable	z-score	p-val	sign90	sign95
Gender Attitude Perception: Men think Women less skilled	-1.3	0.194483	No	No

Although female managers were slightly more likely, on average, to agree with the statement that "Men are more likely to think that women are less good with data/mathematics.", this is not different enough to reject H_8 at a 90% confidence level.

4.5   Significant Differences Based on Gender

Overall, although there do not seem to be major differences between how male and female managers use and approach analytics on a construct-level, there seem to be differences for certain sub-variables between how men and women use and approach analytics, and how they perceive other's capabilities and use of analytics.

TABLE 4.19 - Differences Based on Gender Significant with at Least 90% Confidence

Variable	Significant at Confidence Level	Who scored higher?
Gender Attitude: Difference	95%	Women
Frequency of Use: Prog. Frameworks	95%	Men
Familiarity with analytics	95%	Men
Frequency of Use: Algorithms	95%	Men
Ease of Use: General	95%	Women
Perceived Usefulness: Recommendation	90%	Women
Perceived Usefulness: Quality	90%	Women
Emotional Attitude: Positivity	90%	Women scored higher, but more men were in 'Strongly Agree'

Contrary to the stereotype that men are more likely to think that women are not as skilled at analytics, women were more likely to think that there was a difference between how male and female managers use analytics.

Although there were no significant differences between how male and female managers use most analytics tools and methods, and how familiar they are with these, differences did exist for some metrics. Within these differences, all the variables that showed differences based on gender at 95% confidence were related the use of analytics (Ease Of Use, Gender Attitude, Frequency of Use, Familiarity) and variables with differences based on gender at 90% confidence were more centred around the emotions related to analytics (Usefulness, Emotional Attitude). It is also interesting to note that these differences were very specific: for instance, 'Gender Attitude: Difference' showed differences based off of gender with over 95% confidence, but none of the other 'Gender

Attitude' sub-variables did. In addition, none of the constructs showed enough differences based on gender to be significant at 90% confidence.

4.6   Overall Gender Attitude

Although female managers seem to be slightly more likely to agree that there is a difference between how male and female managers use analytics, it is interesting to see how managers perceive the importance of gender in analytics overall. To do this, 'Gender Attitude' was studied, regardless of the respondent's gender.

For 'Gender Attitude' and all of its sub-variables except 'Gender Attitude: Difference', the majority of respondents disagreed that gender had an impact on the level of skill that managers had in analytics or on their capability to use analytics.

For 'Gender Attitude: Difference' though, most respondents indicated that they were unsure whether there was a difference between how male and female managers used analytics or not, but did not disagree with the statement.

This may imply that while managers felt there was no difference in the skill sets of managers using analytics based on gender, they did feel that there was some difference in the way female and male managers used analytics, unrelated to skill. It is possible that managers have an overall perception that gender may play a role in the workplace, but that this is not reflected in their perception of individuals' skills in their field based on their professional experience.

4.7   Free-text Responses

In the free-text areas provided for comments during the survey, several respondents commented on their use of analytics, their perceptions of the importance of analytics from their point of view and added insight into their opinion on the role, importance and perception of gender in analytics in their professional experience.

Female managers mentioned that the use and attitude to analytics depended on the person, and felt that, while stereotypes based on gender may exist, individual skill sets were more important in the workplace:

> *"Some female managers may be good at analytics while a man may not be or the other way around. It depends on the person."*

Male managers more firmly asserted that they did not believe that gender mattered, although several male managers mentioned that they did realise that the field was still largely male-dominated. All of the comments by male managers included a statement about their disagreement with the belief that gender matters in terms of attitude to or use of analytics:

> "I don't see much gender difference in the use of analytics."

> "I don't think gender really matters as far as analytics goes but there are more men in the field than women. I find women just as competent as men."

> "Gender is not determinist of capacity, but merit is."

Several male managers mentioned that they believed that the main issue was that fewer women were educated in analytics and that, consequently, fewer female analytics specialists were employed in the workforce. They were also more likely to mention that they believed that more mix in the workforce would be beneficial for companies as it would bring more diversity and new perspectives.

FIGURE 4.9 - Comment Word Frequency Clouds by Gender

This is reflected in the word frequency clouds for each gender. Female managers focused more on topics surrounding their use of analytics ("analytics" was the most employed

word for female managers' comments) while male managers' comments were focused more on the fact that a mix was valuable and that gender was not important (the two top words were "broader" and "mix" (see Figure 4.9 above) for male managers).

Common themes across both genders were education in analytics, the perception that there was no difference in how male and female managers use analytics and the opinion that analytics was a rapidly expanding field of great strategic importance for companies. Overall, managers commented that they did not feel that female or male managers were treated or should be treated differently. A full transcript of the free text area responses can be found in Appendix E.

5.   Conclusions and Future Work

5.1   Introduction

The aim of this research was to identify whether male and female managers used analytics differently and had a different attitude towards analytics. This chapter summarises the research undertaken and makes recommendations for future research opportunities. It also discusses the limitations and generalisability of the research findings.

Section 5.2 provides a general conclusion of the study and its individual parts.

Section 5.3 discusses the generalizability of the research findings and their reliability.

Section 5.4 highlights and discusses the limitations of the study and suggests improvements for future studies.

Section 5.5 proposes opportunities for further research into differences in use and attitude to technology based on gender and into the importance of gender in different STEM (Science, Technology, Engineering and Mathematics) fields.

Section 5.6 discusses the importance of the research findings in a corporate context and highlights the possible impacts on opportunities for women in the workplace.

5.2   Conclusions Based on Findings

The literature review conducted for this study highlighted several factors that have been found to impact use and attitude to different technologies, based on pre-existing models of technology adoption and use such as the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Diffusion of Innovations Adoption-Decision model. These factors were then used to understand whether male and female managers used analytics differently and had a different attitude to analytics, as well as to forge a better understanding of managers' attitude to gender in analytics as whole.

To do this, a survey was conducted with a sample of 64 managers from various backgrounds. The collected data was analysed using quantitative methods, supplemented by qualitative analysis of free-text area responses. While there did not seem to be any significant differences in use and attitude to analytics based on gender for most variables, some minor differences did exist for particular sub-variables.

The largest difference between male and female managers was their perception of whether there is a difference between how male and female managers use analytics. In addition, male managers were slightly more likely to be familiar with certain analytics tools and to use analytics more frequently. Minor differences between male and female managers were found in their attitude to analytics: female managers were more likely to find analytics easy to use and they were slightly more likely to find analytics useful and to have a positive outlook towards the use of analytics.

Overall, male and female managers' attitudes to analytics and the importance of analytics are very similar. On average, managers of both genders do not think that there should be a difference in how employees who work with analytics should be treated based on gender. Some respondents felt that differences in familiarity with analytics may be due to education, even though previous analytics experience levels did not differ greatly based on gender.

5.3   Generalisation of Findings

Due to the financial and time constraints it was not possible to access a random sample of managers. The snowball sample of 64 managers respondents used may not accurately represent managers as a whole. This means that the generalisability of these findings may be limited.

Generalisation of the findings would, in addition, be restricted to large companies, as all respondents surveyed were from companies of more than 250 employees, so as to ensure that the resources were available to maintain analytics capabilities within the company.

While this and the limitations of this study described in 5.4 below may impact the reliability of results, this study provides insight into the differences in attitude to analytics based on

gender and better understanding of the role of gender in analytics which should be reflective of the overall population, even though some bias may exist.

5.4   Limitations of Research

Due to financial, time and resource constraints, several limitations were encountered during the course of this study.

A snowball sample of managers was used, and a medium-sized sample of managers was surveyed, which can introduce biases and limit the generalisability of the research findings. For future research within this field, a random or probability-based sampling technique should be used, with a larger sample of the population to ensure better reliability and generalisability of the results.

In addition, the qualitative research in this study was based on comments made in the survey response form. With more time, this could be supplemented by in-depth discussions such as interviews or focus groups on the topic, to forge a better understanding of the experiences of managers relating to the role and perception of gender relating to analytics in the workplace.

5.5   Future Research Opportunities

For future research, it would be interesting to compare how male and female managers approach analytics from a more qualitative perspective. It would also be interesting to look into different populations to understand cultural differences between countries, opinions within analytics teams or by industry, by company size or based on the other demographics recorded in this study.
For future quantitative research, it would be best to use a large probability sample of the population to ensure more generalisability of the research findings.

It would also be interesting to study whether analytics are an exception and if gender attitude to analytics is different than gender attitude to other technologies. Attitude to technologies based on gender could be recorded for several technologies or skills to give an overview of the fields/skill sets where differences exist (e.g. programming, robotics, design, front-end or back-end development, finance, engineering compared to analytics).

This would give an overview of the importance of gender in different STEM fields, and would enable a comparison between fields.

With further research into whether gender attitude has an impact on the share of female employees and female managers, this could help to better understand why there are not more female managers in STEM fields and how to drive more women to study and work in STEM fields - and analytics - while providing them with opportunities for professional development and growth.

If analytics is an exception and further research establishes that there are more significant differences in other fields, it would be interesting to research why this might be the case. Analytics methods, reports and tools are used in a wide range of functions across many industries, so it may be different from other technologies or functions, where the function does not communicate frequently with other business functions. In the opposite case – if analytics are not different from other STEM fields in terms of the importance of gender, studies such as this dissertation and future research can be used to dispel possible preconceptions that female managers or employees are seen as less skilled or use technologies differently. It may also be interesting to compare STEM to non-STEM fields, to see if gender plays a role in some industries more than in others.

Finally, using a longer time-horizon, the evolution of the role of gender in STEM fields could be studied in a longitudinal study, so as to better understand how gender attitude is evolving and if gender attitude is evolving differently – or at different rates - for certain fields or technologies, which could lead to better understanding of why gender gaps in certain industries are growing while they are shrinking in others.

5.6   Importance of Findings in a Corporate Context

As there are very few differences between how male and female managers use analytics, female managers or employees should be given the same opportunities to work with analytics than their male counterparts and vice versa, and gender should not be important during recruitment.

The idea that female managers are more suited to 'soft skills' or use 'soft skills' more and are less likely to use analytics is not valid, according to the findings of this study. Therefore, perceived skill sets and perceived aptitude for analytics should be based on an

individual's experience and demonstrated skill, and their gender should not play a role in the recruitment process.

Furthermore, companies should be encouraged to close the gender gap in STEM fields, as the same skill set and attitude to analytics will be available to them regardless of the gender of their employees. Giving women the same opportunities as men in the workplace is essential in making sure that corporations recruit the best employees and value both their current and prospective employees based on skill and merit.

5.7   Conclusion

The purpose of this dissertation was to determine whether female managers used analytics differently from their male counterparts.
The factors considered for this study were based on factors used in frameworks such as the TAM, UTAUT and Diffusion of Innovations Adoption-Decision models.

The hypothesis testing conducted on the data to compare managers' responses based on gender determined that there were no major differences in usage or attitude towards analytics between male and female managers.

The results indicate that, while there are no differences between the way in which male and female managers use and approach analytics for most factors, female managers were more likely to believe that there was a difference in the attitude to analytics between managers based on gender. Male managers were more likely to use programming frameworks and algorithms for analytics more frequently and they were more likely to consider themselves more familiar with analytics overall. Male managers were found to have focused their comments more on the fact that they did not believe gender to have an impact on managers' capabilities and use of analytics than female managers.

From a practical viewpoint, these findings are an important step towards dispelling stereotypes surrounding gender in STEM fields and making sure that male and female employees are guaranteed the same opportunities in the workplace. They also highlight the importance of science education for career advancement. These findings present opportunities for further research into the role of gender in the workplace, and the influence of gender on the use and adoption of technologies.

REFERENCES

Ahuja M. & Bennet Thatcher J. (2005) Moving beyond intentions and toward the theory of trying: effects of work environment and gender on post-adoption information technology use. *MIS Quarterly* **29**(3), 427-459.

Ajzen I. & Fishbein M. (1980) *Understanding Attitudes and Predicting Social Behaviour.* Prentice-Hall, Englewood Cliffs NJ USA.

Ali L., Asadi M., Gasevic D., Jovanovic J. & Hatala M. (2013) Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education* **62**, 130-148.

Belk R. (2007) *Handbook of Qualitative Research Methods in Marketing.* Edward Elgar Publishing, Cheltenham UK.

Buhl H., Roeglinger M., Moser F., Heidemann J. (2013) Big Data. *Business & Information Systems Engineering* **5**(2), 65-69.

Chen H., Chiang R. & Storey V. (2012) Business Intelligence: From Big Data to Big Impact. *MIS Quarterly* **36**(4), 1165-1188.

Creswell J. (1994) *Qualitative and quantitative approaches.* Sage Publications, New York NY USA.

Crossan F. (2004) Research philosophy: towards an understanding. *Nurse Researcher.* **11**(1), 46-55.

CSO (2013) *Women and Men in Ireland: Employment by economic sector.* CSO, Dublin. Retrieved from http://www.cso.ie/en/releasesandpublications/ep/p-wamii/womenandmeninireland2013/employmentlist/employment/#d.en.65363 on 15 April 2015.

Dahlan N., Ramayah T. & Lai Mei L. (2002) Readiness to Adopt Data Mining Technologies: An Exploratory Study of Telecommunications Employees in Malaysia. *Practical Aspects of Knowledge Management* **2569**, 75-86.

Damanpour F. (1987) The Adoption of Technological, Administrative, and Ancillary Innovations: Impact of Organizational Factors. *Journal of Management* **13**(4), 675-688.

Davenport T. & Harris J. (2007) *Competing on Analytics: The New Science of Winning.* Harvard Business Press, Boston MA USA.

Davenport T. & Patil D. (2012) Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review* **90**(10), 70-76.

Davenport T. (2006) Competing on Analytics. *Harvard Business Review* **84**(1), 98-107.

Davenport T. (2014) How strategists use "big data" to support internal business decisions, discovery and production. *Strategy & Leadership* **42**(4), 45-50.

Davenport T., Barth P. & Bean R. (2012) How 'Big Data' Is Different. *MIT Sloan Review* **54**(1), 22-24.

Davidson M. & Burke R. (eds.) (2011) *Women in Management Worldwide: Progress and Prospects.* Gower Publishing, Farnham UK.

Davis F. (1989) Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly* **13**(3), 319-340.

Easterby-Smith M., Thorpe R., Jackson P. (2012) *Management Research,* 4th edn. Sage Publications, New York NY USA.

Erikkson P. & Kovalainen A. (2008) *Qualitative Methods in Business Research.* Sage Publications, New York NY USA.

Fishbein M. & Ajzen I. (1975) *Belief, Attitude, Intention, and Behaviour: An Introduction to Theory and Research.* Addison Wesley, Reading MA USA.

Forfás & Expert Group on Future Skills Needs (2014) *Executive Summary of Assessing the Demand for Big Data and Analytic Skills.* Forfás, Dublin. Retrieved from http://www.forfas.ie/media/07052014-Assessing_the_Demand_for_Big_Data_and_Analytics_Skills-Publication.pdf on 15 April 2015.

Frambach R. & Schillewaert N. (2002) Organizational innovation adoption: a multi-level framework of determinants and opportunities for future research. *Journal of Business Research* **55**(2), 163-176.

Gartner (2013) *IT Glossary: Analytics.* Gartner, Stamford CT USA. Retrieved from http://www.gartner.com/it-glossary/analytics on 15 April 2015.

Gartner (2014) *Gartner Worldwide IT Spending Forecast.* Gartner, Stamford CT USA. Retrieved from http://www.gartner.com/technology/research/it-spending-forecast/ on 15 April 2015.

Gefen D. & Straub D. (1997) Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model. *MIS Quarterly* **21**(4), 389-400.

Gefen D., Karahanna E. & Straub D. (2003) Trust and TAM in online shopping: an integrated model. *MIS Quarterly* **27**(1), 51-90.

Goes P. (2014) Big Data and IS Research. *MIS Quarterly* **38**(3), iii-viii.

Green J., Camilli G. & Elmore P. (2002) *Handbook of Complementary Methods in Education Research.* Routledge, London UK.

Griffith M., Moore K. & Richardson H. (2007) Celebrating heterogeneity?: A survey of female ICT professionals in England. *Information, Communication & Society* **10**(3), 338-357.

Groves J., Fowler F., Couper M, Lepkowski J., Singer E., Tourangeau R. (2011) *Survey*

Metholodogy. John Wiley & Sons, Indianapolis IN USA.

Hair J. (2007) Knowledge creation in marketing: the role of predictive analytics. *European Business Review* **19**(4), 303-315.

Hilber M. & Lopez P. (2011) The World's Technological Capacity to Store, Communicate, and Compute Information. *Science* **332**(6025), 60-65.

Hitzler P. & Krzysztof J. (2013) Linked Data, Big Data and the 4th Paradigm. *Semantic Web Journal* **4**(3), 233-235.

Hong S., Thong J. & Tam K. (2006) Understanding continued information technology usage behaviour: A comparison of three models in the context of mobile internet. *Decision Support Systems* **42**(3), 1819-1834.

Hu P., Chau P., Liu Sheng O. & Yan Tam K. (1999) Examining the Technology Acceptance Model Using Physician Acceptance of Telemedicine Technology. *Journal of Management Information Systems* **16**(2), 91-112.

Huang E. & Chou T. (2004) Factors for web mining adoption of B2C firms: Taiwan experience. *Electronic Commerce Research and Applications* **3**(3), 266-276.

Huang W., Ward Hood D. & Yoo S. (2013) Gender divide and acceptance of collaborative Web 2.0 applications for learning in higher education. *The Internet and Higher Education* **16**, 57-65.

IBM (2013) *What is Big Data?* IBM, Armonk NY USA. Retrieved from http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html on 15 April 2015.

IBM (2014) *Big Data.* IBM, Armonk NY USA. Retrieved from http://www.ibm.com/big-data/us/en/ on 15 April 2015.

IDC (2012) *Worldwide Big Data Technology and Services 2012-2015 Forecast.* IDC, Framingham MA USA. Retrieved from http://www.idc.com/research/viewtoc.jsp?containerId=233485 on 15 April 2015.

Intel (2012) *Big Data Analytics: Peer Research.* Intel IT Center, Santa Clara CA USA. Retrieved from http://www.intel.ie/content/dam/www/public/us/en/documents/reports/data-insights-peer-research-report.pdf on 15 April 2015.

Irish Department for Jobs, Enterprise and Innovation (2014) *Action Plan for Jobs.* Irish Department for Jobs, Enterprise and Innovation, Dublin. Retrieved from http://www.djei.ie/publications/2014APJ.pdf on 15 April 2015.

Jacobs A. (2009) A Blind Person's Interaction with Technology. *Communications of the ACM* **52**(8), 36-44.

Johnson R. & Onwuegbuzie A. (2004) Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher* **33**(7), 14-26.

Kassi P. (2009) *Research: What, Why and How? A Treatise from Researchers to*

Researchers. AuthorHouse, Bloomington IN USA.

Kim S., Choi S. & Wang K. (2014) Individual perception vs. structural context: Searching for multilevel determinants of social acceptance of new science and technology across 34 countries. *Science and Public Policy* **41**, 44-57.

Knox K. (2004) A Researcher's Dilemma - Philosophical and Methodological Pluralism. *Electronic Journal of Business Research Methods* **2**(2), 119-128.

Kotorov R. (2015) *The New Horizon for Business Intelligence and Analytics.* Nearshore Americas, New York NY USA. Retrieved from http://www.nearshoreamericas.com/horizon-business-intelligence-analytics/ on 15 April 2015.

Laney D. (2001) *3D Data Management: Controlling Data Volume, Velocity and Variety.* Gartner Meta Group, Stamford CT USA. Retrieved from http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf on 15 April 2015.

LaValle S., Hopkins M., Lesser E. Shockley R & Kruschwitz N. (2010) Big Idea: Data & Analytics. *MIT Sloan Management Review* **52**(2), 21-31.

Legris P., Ingham J. & Collerette P. (2003) Why do people use information technology? A critical review of the technology acceptance model. *Information & Management* **40**(3), 191-204.

Lewis A. & Sambamurthy A. (2003) Sources of influence on beliefs about information technology use: an empirical study of knowledge workers. *MIS Quarterly* **27**(4), 657-678.

Lui C. & Chang D. (2012) An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management* **32**(3), 257-270.

Malakh-Pines A. & Oezbilgin M. (2010) *Handbook of Research on High-Technology Entrepreneurs.* Edward Elgar Publishing, Cheltenham UK.

Manyka J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Hung Byers A. (2011) *Big Data: The next frontier for innovation, competition and productivity.* McKinsey Global Institute, Stamford CT USA. Retrieved from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation on 15 April 2015.

Mariani M., Curcuruto M. & Gaetani I. (2013) Training opportunities, technology acceptance and job satisfaction: A study of Italian organizations. *Journal of Workplace Learning* **25**(7), 455-475.

Mattman C. (2013) Computing: A vision for data science. *Nature* **493**(7433), 472-475.

McAfee A. & Brynjolfsson E. (2012) Big Data: the management revolution. *Harvard Business Review* **90**(10), 60-66.

McCallum J. (2002) Price-Performance of Computer Technology. In *The Computer*

Engineering Handbook. (Vojin, G ed.) CRC Press, Boca Raton FL USA, pp 4-1 - 4-18.

Misa, T. (2010) *Gender Codes: Why Women Are Leaving Computing.* IEEE Computer Society, Hoboken NJ USA.

Moon J. & Kim Y. (2001) Extending the TAM for a World-Wide-Web context. *Information & Management* **38**(4), 217-230.

Moore G. (1965) Cramming more components onto integrated circuits. *Electronics Magazine* **38**(8), 4.

Morabito V. (2015) *Big Data and Analytics: Strategic and Organizational Impacts.* Springer, New York NY USA.

National Women's Council of Ireland (2015) *Our History.* National Women's Council of Ireland, Dublin. Retrieved from http://www.nwci.ie/discover/about_us/our_history/ on 15 April 2015.

Netflix Tech Blog (2015) *Cassandra.* Netflix, Los Gatos CA USA. Retrieved from http://techblog.netflix.com/search/label/Cassandra on 9 March 2015.

Ong C. & Lai J. (2006) Gender differences in perceptions and relationships among dominants of e-learning acceptance. *Computers in Human Behaviour* **22**(5), 816-829.

Padilla-Melendez A., Del Aguila-Obra A., Garrido-Moreno A. (2013) Perceived playfulness, gender differences and technology acceptance model in a blended learning scenario. *Computers & Education* **63**, 306-317.

Pettey C. (2012) *Gartner Says Big Data Creates Big Jobs.*Gartner Newsroom, Orlando FL USA. Retrieved from http://www.gartner.com/newsroom/id/2207915 on 15 April 2015.

Pramatari K & Theotokis A. (2009) Consumer acceptance of RFID-enabled services: a model of multiple attitudes, perceived system characteristics and individual traits. *European Journal of Information Systems* **18**, 541-552.

Prescott J. (2014) *Gender Considerations and Influence in the Digital Media and Gaming Industry.* IGI Global, Hersey PA USA.

Rapp A., Schillewaert N. & Wei Hao A. (2008) The Influence of Market Orientation on E-Business Innovation and Performance: The Role of the Top Management Team. *Journal of Marketing Theory and Practice* **16**(1), 7-25.

Ribarsky W., Wang D. & Dou W. (2014) Social media analytics for competitive advantage. *Computers & Graphics* **38**(1), 328-331.

Richardson J. (2013) *Gartner Business Intelligence and Analytics (slideshow).* Gartner, Stamford CT USA.

Roan A. & Whitehouse G. (2007) Women, information technology and 'waves of optimism': Australian evidence on 'mixed-skill' jobs. *New Technology, Work and Employment* **22**(1), 21-33.

Rogers E. (2010) *Diffusion of Innovations,* 4th edn. Simon and Schuster, New York NY USA.

Rote M. (2005) *Advanced Analytics: What Every Enterprise Should Know.* NCR Corporation, Duluth GA USA. Retrieved from http://www.teradata.com/Resources/Articles/Advanced-Analytics-What-Every-Enterprise-Should-Know-Mike-Rote/ on 15 April 2015.

Russom P. (2011) *TDWI Best Practices Report: Big Data Analytics.*TDWI, Treusdorf Germany. Retrieved from http://public.dhe.ibm.com/common/ssi/ecm/en/iml14293usen/IML14293USEN.PDF on 15 April 2015.

Saunders M., Lewis P. & Thornhill A. (2009) *Research Methods for Business Students*, 5th edn. Pearson Education, New York NY USA.

Schaller B. (1996) *The Origin, Nature and Implications of "Moore's Law".* Microsoft Research. Retrieved from http://research.microsoft.com/en-us/um/people/gray/moore_law.html on 15 April 2015.

Shashank T. (2011) *Professional NoSQL.* John Wiley & Sons, Indianapolis IN USA.

Shroff R., Deneen C. & Ng E. (2011) Analysis of the technology acceptance model in examining students' behavioural intention to use an eportfolio system. *Australasian Journal of Educational Technology* **27**(4), 600-618.

Simpson R. (2000) Gender mix and organisational fit: how gender imbalance at different levels of the organisation impacts on women managers. *Women in Management Review* **15**(1), 5-18.

Steward D. (1991) From Methods and Projects to Systems and Process: The Evolution of Marketing Research Techniques. *Marketing Research* **3**(3), 25-36.

Stoker J., Van der Velde M. & Lammers J. (2011) Factors Relating to Managerial Stereotypes: The Role of Gender of the Employee and the Manager and Management Gender Ratio. *Journal of Business and Psychology* **27**(1), 31-42.

Straub D., Keila M. & Brenner W. (1997) Testing the technology acceptance model across cultures: A three country study. *Information & Management* **33**(1), 1-11.

Tarr-Whelan L. (2009) *Women Lead the Way.* Berrett-Koehler Publishers, San Fransisco CA USA.

Tavana M. (2014) *Handbook of Research on Organizational Transformations through Big Data Analytics.* IGI Global, Hershey PA USA.

The Economist (2015) *Even more from Moore.* The Economist Newspaper, London UK. Retrieved from http://www.economist.com/news/business/21648683-microchip-pioneers-prediction-has-bit-more-life-left-it-ever-more-moore on 19 August 2015.

Tornatzky L. & Fleischer M. (1990) *The Processes of Technological Innovation.* Lexington

Books, Lexington MA USA.

Van Barneveld A., Arnold K & Campbell J. (2012) *Analytics in Higher Education: Establishing a Common Language.* Educause Learning Initiative, Louisville CO USA. Retrieved from http://net.educause.edu/ir/library/pdf/eli3026.pdf on 15 April 2015.

Vargo S. & Lusch R. (2004) Evolving to a New Dominant Logic for Marketing. *Journal of Marketing* **68**(1), 1-17.

Vekiri I. & Chronaki A. (2008) Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Computers & Education* **51**(3), 1392-1404.

Venkatesh V & Davis F. (2000) A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science* **46**(2), 186-204.

Venkatesh V. & Morris M. (2000) Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behaviour. *MIS Quarterly* **24**(1), 115-139.

Venkatesh V., Morris M. & Ackerman P. (2000) A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes. *Organizational Behaviour and Human Decision Processes* **83**(1), 33-60.

Venkatesh V., Morris M., Davis F. & Davis G. (2003) User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* **27**(3), 425-478.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences* **39**(2), 273–315.

Venkatsh V. (2000) Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research* **11**(4), 342-365.

Wade M. (2003) Women and Salary Negotiation: The Costs of Self-Advocacy. *Psychology of Women Quarterly* **25**(1), 65-76.

Wajcman J. (1998) *Managing Like a Man: Women and Men in Corporate Management.*Polity Press, Cambridge UK.

Walsh K. & Callan A. (2011) Perceptions, Preferences and Acceptance of Information and Communication Technologies in Older-Adult Community Care Settings in Ireland: A Case-Study and Ranked-Care Program Analysis. *Ageing International* **36**(1), 102-122.

Whitley B. (1997) Gender differences in computer-related attitudes and behavior: a meta-analysis. *Computers in Human Behavior* **13**(1), 1-22.

Yang H. & Yoo Y. (2004) It's all about attitude: revisiting the technology acceptance model. *Decision Support Systems* **38**(1), 19-31.

Yuen A. & Ma W. (2002) Gender Differences in Teacher Computer Acceptance. *Journal of Technology and Teacher Education* **10**(3). 365-382.

Zhu K., Kraemer K. & Xu S. (2003) Electronic business adoption by European firms: a cross-country assessment of the facilitators and inhibitors. *European Journal of Information Systems* **12**, 251-268.

Zikopoulos P. & Eaton C. (2011) *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data.* McGraw-Hill, New York NY USA.
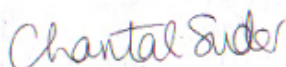
APPENDIX A: Ethics Approval

School of Computer Science and Statistics
Research Ethical Application Form

Part A

Project Title: The impact of gender on managers' use of analytics

Name of Lead Researcher (student in case of project work): Chantal Suder

Name of Supervisor: Frank Bannister

TCD E-mail: suderc@tcd.ie Contact Tel No.: 086 276 9515

Course Name and Code (if applicable): MSc. Management of Information Systems Year 2

Estimated start date of survey/research: May 22nd 2015

I confirm that I will (where relevant):

- Familiarize myself with the Data Protection Act and the College Good Research Practice guidelines http://www.tcd.ie/info_compliance/dp/legislation.php;
- Tell participants that any recordings, e.g. audio/video/photographs, will not be identifiable unless prior written permission has been given. I will obtain permission for specific reuse (in papers, talks, etc.)
- Provide participants with an information sheet (or web-page for web-based experiments) that describes the main procedures (a copy of the information sheet must be included with this application)
- Obtain informed consent for participation (a copy of the informed consent form must be included with this application)
- Should the research be observational, ask participants for their consent to be observed
- Tell participants that their participation is voluntary
- Tell participants that they may withdraw at any time and for any reason without penalty
- Give participants the option of omitting questions they do not wish to answer if a questionnaire is used
- Tell participants that their data will be treated with full confidentiality and that, if published, it will not be identified as theirs
- On request, debrief participants at the end of their participation (i.e. give them a brief explanation of the study)
- Verify that participants are 18 years or older and competent to supply consent.
- If the study involves participants viewing video displays then I will verify that they understand that if they or anyone in their family has a history of epilepsy then the participant is proceeding at their own risk
- Declare any potential conflict of interest to participants.
- Inform participants that in the extremely unlikely event that illicit activity is reported to me during the study I will be obliged to report it to appropriate authorities.
- Act in accordance with the information provided (i.e. if I tell participants I will not do something, then I will not do it).

Signed: *Chantal Suder*

Lead Researcher/student in case of project work   Date: April 5th 2015

Part B

Please answer the following questions.		Yes/No
Has this research application or any application of a similar nature connected to this research project been refused ethical approval by another review committee of the College (or at the institutions of any collaborators)?		No
Will your project involve photographing participants or electronic audio or video recordings?		No
Will your project deliberately involve misleading participants in any way?		No
Does this study contain commercially sensitive material?		No
Is there a risk of participants experiencing either physical or psychological distress or discomfort? If yes, give details on a separate sheet and state what you will tell them to do if they should experience any such problems (e.g. who they can contact for help).		No
Does your study involve any of the following?		
	Children (under 18 years of age)	No
	People with intellectual or	No

SCSS Research Ethics Application Form August 2014

School of Computer Science and Statistics
Research Ethical Application Form

Details of the Research Project Proposal must be submitted as a separate document to include the following information:

1. Title of project
2. Purpose of project including academic rationale
3. Brief description of methods and measurements to be used
4. Participants - recruitment methods, number, age, gender, exclusion/inclusion criteria, including statistical justification for numbers of participants
5. Debriefing arrangements
6. A clear concise statement of the ethical considerations raised by the project and how you intend to deal with them
7. Cite any relevant legislation relevant to the project with the method of compliance e.g. Data Protection Act etc.

Part C

I confirm that the materials I have submitted provided a complete and accurate account of the research I propose to conduct in this context, including my assessment of the ethical ramifications.

Signed: *Chantal Sivder*   Date: April 5th 2015

Lead Researcher/student in case of project work

There is an obligation on the lead researcher to bring to the attention of the SCSS Research Ethics Committee any issues with ethical implications not clearly covered above.

Part D

If external or other TCD Ethics Committee approval has been received, please complete below.

External/TCD ethical approval has been received and no further ethical approval is required from the School's Research Ethical Committee. I have attached a copy of the external ethical approval for the School's Research Unit.

Signed: ....   Date: ....
Lead Researcher/student in case of project work

Part E

If the research is proposed by an undergraduate or postgraduate student, please have the below section completed.

I confirm, as an academic supervisor of this proposed research that the documents at hand are complete (i.e. each item on the submission checklist is accounted for) and are in a form that is suitable for review by the SCSS Research Ethics Committee.

Signed: *Frank E. Bannister*   Date: 15th April 2015
Supervisor

APPENDIX B: Survey Questionnaire

Do female managers use analytics differently from their male counterparts?

*Required

INFORMATION FOR PARTICIPANTS

You are invited to participate in this research project which is being carried out by Chantal Suder as part of a dissertation in the Taught Masters Programme M.Sc. in Management of Information Systems in the School of Computer Science and Statistics, Trinity College Dublin, Ireland.

This research attempts to investigate the role of gender in the use and attitude to analytics of male and female managers.

Please note that this research is restricted to individuals with experience as managers in large companies (more than 250 employees) only.

If you are willing to participate in the survey, it will involve responding to a maximum of 13 questions, all of which are optional and it should take no more than 10 minutes to complete.

This survey is anonymous.

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

I have no conflict of interest with regard to the research topic and participants, either individually or any other level (employment or university).

This dissertation along with the gathered anonymous data may be published in Trinity College Dublin Library along with all other theses and dissertations.

If you have any questions about this survey or the research study in general, please do not hesitate to contact me at suderc@tcd.ie.

INFORMED CONSENT

RESEARCHER: Chantal Suder
CONTACT DETAILS: suderc@tcd.ie

BACKGROUND OF RESEARCH
This research attempts to investigate the role of gender in the use and attitude to analytics of male and female managers. Existing models of technology adoption and acceptance suggest that various factors can influence technology adoption and the perception of a technology by individuals and in an organisational context.
This research attempts to understand if gender is one of these factors for the adoption of analytics.
In order to carry out this research a web based survey will be conducted on managers of both genders about their use and perception of analytics.

PROCEDURES OF THIS STUDY
Participants must be 18 years of age or older.
Participants in this survey will remain anonymous.
Participation is voluntary and should take no longer than 10 minutes.
All survey questions are optional and can be skipped nevertheless the researcher would appreciate it if all questions are answered. Participants have the option to "Exit" any time before completing the questionnaire by clicking the "Exit This Survey" button, in so doing; answers will NOT be recorded.

·NOTE: As this research is restricted to individuals with experience as managers in large companies (more than 250 employees) only.

OTHER INFORMATION:
· I have no conflict of interest with regard to the research topic and with any of the participants either individually or at an organisational level
· In the extremely unlikely event that illicit activity is reported, I will be obliged to report it to appropriate authorities.
· Please do not name third parties in any open text field of the questionnaire. Any such replies will be anonymised.

PUBLICATION
At the end of the survey, individual results will be aggregated anonymously and research reported on aggregate results. The results will be used solely for a dissertation as part of the completion of a M.Sc. in Management of Information Systems course at Trinity College Dublin (TCD).

CONFIDENTIALITY
Your responses will be kept completely confidential.
If you have any questions or concerns or if you have any difficulties accessing this survey, you may contact me at: suderc@tcd.ie.

DECLARATION:
· I am 18 years or older and am competent to provide consent.
· I have read, or had read to me, a document providing information about this research and this consent form. I have had the opportunity to ask questions and all my questions have been answered to my satisfaction and understand the description of the research that is being provided to me.
· I agree that my data is used for scientific purposes and I have no objection that my data is published in scientific publications in a way that does not reveal my identity.
· I understand that if I make illicit activities known, these will be reported to appropriate authorities.
· I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights.
· I understand that I may refuse to answer any question and that I may withdraw at any time without penalty.
· I understand that my participation is fully anonymous and that no personal details about me will be recorded.
· Since this research involves viewing materials via a computer monitor I understand that if I or anyone in my family has a history of epilepsy then I am proceeding at my own risk.

By submitting this form you are indicating that you have read the description of the study, are over the age of 18, and that you agree to the terms as described.

Thank you in advance for your participation!
Chantal Suder

Do you agree to the consent information listed on this form? *

○ Yes

○ No

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

In order for everyone participating in this survey to answer this questionnaire with the same definition of analytics in mind, a definition is provided below.

Analytics is the study or analysis of data. It encompasses statistical techniques, tools and outputs that answer the following questions:

Descriptive analytics – what happened?
Diagnostic analytics – why did it happen?
Predictive analytics – what will happen?
Prescriptive analytics – how can we make it happen?

This survey is geared towards customer and business analytics used in a corporate context to impact decision-making or strategy.

1. How familiar are you with analytics?

○ Not at all familiar

○ Slightly familiar

○ Moderately familiar

○ Very familiar

○ Extremely familiar

2. What is your gender?

○ Male

○ Female

3. In general, how would you describe when you usually adopt technologies (e.g. mobile phones, computers, new software)?

○ Before everyone else

○ Among the first

○ In the early majority

○ In the late majority

○ Among the last

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

4. How would you describe yourself as an analytics user?

○ I use the results of analytics, but I do not necessarily understand the principles underlying analytics techniques.

○ I use the results of analytics and I have a basic understanding of most of the principles underlying analytics techniques, but not the mathematical/statistical details.

○ I use the results of analytics and have a good understanding of the underlying techniques and calculations used. I am technically savvy as well as an end user.

○ I am a technical expert whose role relates to analytics. I have an in-depth understanding of

the underlying statistical techniques and tools used for analytics. I am not an end user as such.

○ Other: [   ]

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

5. Please specify the frequency with which you use the following analytics tools.

	Several times a day	About once a day	Several times a week	About once a week	A few times a month	About once a month	Less than once a month	Never, but I am familiar with this.	Never and I am no familia with this.
Paper-based reports	○	○	○	○	○	○	○	○	○
Dashboards and/or online reports	○	○	○	○	○	○	○	○	○
Ad-hoc analysis	○	○	○	○	○	○	○	○	○
Modelling and/or model scores	○	○	○	○	○	○	○	○	○
Data Visualisations (e.g. maps, charts)	○	○	○	○	○	○	○	○	○
Social media analysis	○	○	○	○	○	○	○	○	○
Text analytics	○	○	○	○	○	○	○	○	○
Web analytics (e.g. Google Analytics)	○	○	○	○	○	○	○	○	○
Data mining platforms and software (e.g. SPSS, SAS)	○	○	○	○	○	○	○	○	○
Learning algorithms (e.g. regression, clustering, decision trees)	○	○	○	○	○	○	○	○	○
Programming languages and frameworks (e.g. Python, R)	○	○	○	○	○	○	○	○	○

6. Please comment below if you use any other analytics tools or methods which are not included in the options above, or if you have any comments on how you use analytics tools in your work/job.

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

7. Please specify your rating for each of the questions below.

	Strongly disagree	Disagree	Agree	Strongly agree	Do not know	Not applicable
In general, I find analytics methods easy to use and understand.	○	○	○	○	○	○
In general, I find results of data analysis easy to use and understand.	○	○	○	○	○	○
Analytics are useful in my job.	○	○	○	○	○	○
Analytics enable me to do my job faster.	○	○	○	○	○	○
Analytics improve the quality of work I do.	○	○	○	○	○	○
Analytics make it easier to do my job.	○	○	○	○	○	○
In general, I like using analytics for my job.	○	○	○	○	○	○
In general, analytics are fun.	○	○	○	○	○	○
In general, analytics are hard work.	○	○	○	○	○	○
I would recommend the use of analytics to other managers in my industry.	○	○	○	○	○	○
I intend to use analytics more frequently in the future.	○	○	○	○	○	○
Male and female managers use analytics differently.	○	○	○	○	○	○
Male managers are more skilled at using analytics.	○	○	○	○	○	○
In general, women						

are less good with data/mathematics.	○	○	○	○	○	○
Female managers are better suited to jobs focused on soft skills.	○	○	○	○	○	○
Men are more likely to think that women are less good with data/mathematics.	○	○	○	○	○	○
I would rather do work related to the use of analytics with a man than a woman.	○	○	○	○	○	○

8. Please add any further comments that you have on the use of analytics in your work/job.

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

9. How many years of managerial experience do you have?

○ Less than one year

○ 1 - 3 years

○ 4 - 6 years

○ 7 - 9 years

○ 10 or more years

10. What industry do you work in or have the most experience in?

[▼]

11. What is your level of management?

○ Front-line/operational management

○ Middle management

○ Senior management

Each question is optional. Feel free to omit a response to any question; however the researcher would be grateful if all questions are responded to.

12. What age range are you in?

○ Under 20

○ 21-30

○ 31-40
○ 41-50
○ 51-60
○ 61-70
○ Over 70
○ Would rather not say

13. Please add any additional comments on your use of analytics, your opinion on the impact of gender on analytics, or this survey below.

Thank you answering this survey! To submit, press the Submit button. If you would like to Exit without submitting, close this page or click 'Back' to go back to a previous page.

Submit

Never submit passwords through Google Forms.

Powered by

This form was created inside of Trinity College Dublin.

Report Abuse - Terms of Service - Additional Terms

APPENDIX C: Data Analysis Script (selected functions)

Gini Calculation

```python
def gini(col, df):
    # 1. using pandas and scipy libraries for data manipulation
    #get counts by gender
    gdf = df[[col, 'Gender']]
    ggdf = gdf.groupby(by = [col, 'Gender'], as_index=True).size()

    #d2 is quantile, d3 is female density
    d2 = []
    d3 = []

    #calculate density for each quantile
    for i in gdf[col].value_counts().keys():
            d2.append(i)
            #both genders in a quantile
            if len(ggdf[i]) == 2 :
                d3.append(1.000 * ggdf[i][1] /
                    (1.000 * ggdf[i][0] + ggdf[i][1]))
            else:
                #only one gender
                if len(ggdf[i]) == 1:
                    try:
                        if (ggdf[i][0] == 0):
                            d3.append(0.0)
                        else:
                            d3.append(1.0)
                    except:
                        d3.append(1.0)

    #sort quantiles by female density
    df3= pd.DataFrame(d2)
    df3['density'] = d3
    df3.columns = ['quantile', 'density']
    df3.sort(columns=['density'], ascending=[0], inplace=True)

    #change quantiles to new sorted quantiles and sort df by them
    arr = (np.array(df3['quantile'])).tolist()
    df[col] = df[col].map(lambda x: arr.index(x) if (x in arr) else x)
    df = df.sort([col], ascending=[0])

    # 2. using roc_curve and auc functions from the scikit learn library
    #calculate gini based on the array sorted by decreating female density
    fpr, tpr = roc_curve(df['Gender'], df[col])
    #get area under roc curve
    auc_size = auc(fpr, tpr)
    #calculate gini based on area
    gini = 2* (abs(0.5 - auc_size))

    return gini
```

Distributions by Gender

```python
#for each column, get variable grouped by gender
# in the following format:
# Score     Male     Female
#   1        5        15
#   2       25        21
for col in np.array(df.columns):
    if col <> 'Gender' and col <> 'ID':
        #make sure there are no nulls
        notnullT = df[df[col] <> -1]['Gender']
        notnullC = df[df[col] <> -1][col]

        #create subarray of just one var and Gender
        gdf = pd.DataFrame([notnullT, notnullC])
        gdf.columns = ['Gender', col]

        #group by gender counts, unwrap so Gender becomes the columns
        ggdf = gdf.groupby(by = [col, 'Gender'], as_index=True).size()
        ggdf2 = ggdf.unstack(level=-1)

        print ggdf2
```

Z-Score Calculation

```python
#z-score column by column
]for col in df.columns:
    #calculate the difference between the means for male (0) and female (1)
    diffmeans = (means[col][0] - means[col][1])
    #calculate the square-root of the standard deviations (0 for male, 1 for female)
    diffdevs = np.sqrt((devs[col][0] * devs[col][0] / counts[col][0]) + (devs[col][1] * devs[col][1] / counts[col][1]))
    #divide to get z-score
    zscore = diffmeans/diffdevs
    #use scipy function norm.sf to get the p-value from the z-score
    #multiply by 2 to get two-tailed value
    pval = scipy.stats.norm.sf(abs(zscore))*2
    #test of if significant at confidence levels for 90% and 95%
    sign95 = 'Yes' if pval < 0.05 else 'No'
    sign90 = 'Yes' if pval < 0.10 else 'No'
    scores.append([col, round(zscore, 2), pval, sign90, sign95])
```

Correlation Matrices using Pearson's correlation

```python
#pearson's correlation using the pandas library
corr = df.corr().fillna(0)
#use matplotlib to create a figure
plt.figure(figsize=(10,10))
#make this figure a heatmap based on the correlation matrix calculated by pandas
heatmap = plt.pcolor(corr)
#set the right axis - index and columns are the same here
plt.yticks(np.arange(0.5, len(corr.index), 1), corr.index)
plt.xticks(np.arange(0.5, len(corr.columns), 1), corr.columns)
#set the heatmap ticks and rotate them so they a legible on the x axis
locs, labels = plt.xticks()
plt.setp(labels, rotation=90)
#create a legend for the scores and put it to the right of the heatmap
cbar = plt.colorbar(heatmap)
cbar.set_label('Legend', rotation=270)
#render the correlation matrix in matplotlib
plt.show()
```

APPENDIX D: Gini Scores

Variable	Gini
Tool Familiarity	0.60
Frequency of Use: Data Viz.	0.44
Gender Attitude: Difference	0.39
Frequency of Use: Paper-based	0.37
Familiarity with analytics	0.34
Perceived Usefulness: Future	0.34
Frequency of Use: Social Med. Analysis	0.32
Frequency of Use: Ad hoc	0.31
Emotional Attitude	0.30
Frequency of Use: Software	0.28
Gender Attitude: Women less skilled	0.27
Analytics experience level	0.25
Emotional Attitude: Positivity	0.23
Ease of Use	0.22
Gender Attitude	0.22
Frequency of Use: Modelling	0.20
Ease of Use: Analytics Methods	0.20
Frequency of Use: Web Analytics	0.19
Ease of Use: Hard Work	0.19
Gender Attitude: Work with Men	0.18
Frequency of Use: Prog. Frameworks	0.17
Frequency of Use: Text Analytics	0.16
Familiarity: Prog. Frameworks	0.15
Gender Attitude: Soft skills for women	0.15

Emotional Attitude: Fun	0.14
Frequency of Use: Online dash	0.13
Familiarity: Algorithms	0.13
Familiarity: Software	0.09
Familiarity: Text Analytics	0.09
Familiarity: Web Analytics	0.07
Perceived Usefulness: General	0.06
Gender Attitude Perception: Men think Women le...	0.04
Perceived Usefulness: Recommendation	0.04
Perceived Usefulness: Quality	0.04
Familiarity: Social Med. Analysis	0.04
Familiarity: Ad hoc	0.04
Frequency of Use	0.03
Ease of Use: Analytics Results	0.03
Familiarity: Modelling	0.02
Gender Attitude: Male more skilled	0.02
Perceived Usefulness: Speed	0.01
Familiarity: Online dash	0.01
Familiarity: Data Viz.	0.01
Perceived Usefulness	0.01
Frequency of Use: Algorithms	0.01
Familiarity: Paper-based	0.00
Ease of Use: General	0.00

APPENDIX E: Free-text Response Transcripts

*Note: some responses have been edited to remove information that could make
respondents identifiable.*

Gender	Comment
Male	I think analytics are important to get good decision. I would recommend the use of analytics to other managers in my industry and all other with high level of data to know the customer, take better decision in planning budgets and sales and designing products I think analytics is a growing tool, and event is important nowadays, it will grow up in the coming years and therefore the use of it will spread. Gender may have an impact in doing analytics (maybe) but I think using it should not affect everybody feels confident when have a number where to hold decisions. I've been doing Data analysis using Microsoft SQL Server and Excel. I'm just getting in the world of R, Python and SAS and I'm about to choose R as my analytic program.
Male	In my experience females are often more analytical than males.
Male	We use Microsoft SQL Server Analysis Services for data mining and analytics.
Male	The use of "big data", the "internet of things" and how to effectively make use of disruptive technologies is becoming an increasingly important element of concern.
Male	I do not believe I've seen a different between male and female use of analytics. While there may yet still be some level of bias against women for math skills but I've never seen a man discount analysis from a woman because she was a woman. In general, as far as I've seen, MOST men and women are equally bad at analytics. They try to make the data say what they want it to say, they assume the analysis and their interpretation of the analysis is correct, they don't want to put in the effort to check and validate the work and they are both highly susceptible to confirmation bias.
Male	Gender is not determinist of capacity, but merit is. We use monitoring tools for infrastructure compute resources.
Male	In my role as a manager, the manager must be aware of all available tools. The business unit and function would probably determine the tool to be used before our team.
Male	In my experience I don't see much gender difference on the use of analytics. The challenges comes from the numbers of females who are formally trained / educated in related fields (comp science, maths, statistics, data management analytics). Therefore the instances of working with women in this field are less. However I have never seen a difference in a way men or women will work with analytics. Men and women tend to work differently in any case (my view is men are more chaotic in their approach, whereas women will be more controlled and organised).

Male	Analytics make my job easier and companies should use more analytics to drive better decision making. Companies need to invest in reporting and analytics resourcing and accept that analysis takes time and should not be rushed. I don't think gender really matters as far as analytics goes but there are more men in the field than women. I find women just as competent as men.
Male	All my analysis is done daily by automated scripts and the ad hoc analysis.
Female	I use analytics on a daily basis and find it combined with qualitative research to be the key to informing business decision making.
Female	I strongly believe that this is a perceptual issue, and there is no innate gender reason why analytics should be so male dominated. As data analysis becomes more 'socialised'- about the sharing rather than just the discovery, the more feminine (not necessarily females) skills of communication and empathy are likely to be seen are relevant and useful. However in general I feel that the area of data analytics is changing very rapidly, so a stronger inclusion of all skill/ personality types should be encouraged - and a broader gender mix. One other aspect is that when investigating data- a broader mix of interrogators will make for much more interesting lateral connections and discovery. See previous panel, however just to add- Analytics offers an enormous opportunity to understand behaviour and possibly motivations in a new and exciting way. The wider range of people involved in this - age, class, gender- whatever, the richer this will be.
Female	Analytics are provided by a specific internal team and adapted to our business for performance review.
Female	Some female managers may be good at analytics while a man may not be or the other way around. It depends on the person.

APPENDIX F: Full Correlation Matrix

Note: this correlation matrix shows absolute Pearson's correlation scores.

APPENDIX G: Categories by Demographic Variable

Demographic	Categories
Age Band	Under 20
	21-30
	31-40
	41-50
	51-60
	61-70
	Over 70
	Would rather not say
Technology Adoption Speed	Before everyone else
	Among the first
	In the early majority
	In the late majority
	Among the last
Industry	Aerospace
	Commercial retail products
	Construction
	Consultancy
	Education
	Engineering
	Financial services
	Food and drink
	Public Sector
	Software
	Telecommunications
	Internet
	Media and Entertainment
	Travel Sector
	Other
Years of Management Experience	Less than one year
	1 - 3 years
	4 - 6 years
	7 - 9 years
	10 or more years
Management Level	Front-line/lower-level management
	Middle management
	Senior management