A Study of the Perceptions Held by Information Technology Professionals in Relation to the Maturity, Value, and Practical Deployment of Big Data Solutions

Runion, Damon Andrick

Universidad Central de Nicaragua

January 2016

A Study of the Perceptions Held by Information Technology Professionals in Relation to the Maturity, Value, and Practical Deployment of Big Data Solutions

Damon Andrick Runion [1]

ABSTRACT

This research study investigated relationships between an information technology (IT) professional's self-assigned understanding of big data and their assessment of the maturity, value, hype, and future trends of big data. The study also examined if there was any relationship between an IT professional's understanding of big data and the position they occupy professionally. The study consisted of a twenty question survey. Research findings indicate that IT professionals are still becoming familiar with big data and related technologies. The results supported rejecting two of the five hypotheses. The study produced evidence that there is a relationship between an IT professional's level of big data understanding and their expectation that there will be an increase in technological developments related to big data in the near future.

KEYWORDS: Big Data, Analytics, Business Intelligence, Information Technology, Data Management, Management Information Systems, Information Technology Professionals.

INTRODUCTION

The explosion of data being captured and stored in information systems has created a new area of challenges and opportunities for information technology (IT) professionals. Every day

[1] Research Scholar, Universidad Central de Nicaragua, San Antonio, Texas, USA, damon.runion@gmail.com

millions upon millions of bytes of data are being collected, as related to customer transactions, social media postings, government operations, and traffic sensors. The advent of this rise in data is termed "big data" and it presents challenges from technical, managerial, and analytical perspectives. Organizations are being faced with difficult decisions related to the retention of data and how to analyze stored data to extract value. If organizations hope to obtain value from big data, they must understand the breadth and depth of big data awareness held by their IT employees.

LITERATURE REVIEW

Big data is essentially a simple set of words but it is also term that hides a surprising degree of complexity upon further inspection. Big data in the most basic sense relates to extremely large data sets. Big data also generally relates to data sets that are not able to be managed using typical relational database management systems with respect to size, quality, or other characteristics of the data. (Chen, M., Mao, S., & Liu, Y., 2014; Manyika, et al., 2011).

The definition of big data also varies based upon who you ask. Big data might be a collection of web logs to an IT system engineer or a vast trove of clinical trial data for research scientist. Since there is typically some type of business case related to big data it is necessary to examine the basic features and aspects of big data to establish a baseline of understanding (Chen, Mao, & Liu, 2014).

One the most basic frameworks used to characterize big data is that of the three "V's". The three "V's" are Volume, Variety, and Velocity. Volume corresponds to the record count and size of the data elements being stored. Variety relates to the differences in nature and qualities of data being stored. This typically calls out the types of files or formats being stored. Velocity relates to the rapid pace at which new data is being acquired. Velocity is what distinguishes big data form relational database data most distinctively. The three "V's" cause the industry standard relational database model to falter. Velocity and Volume are the main reasons that the physical limitations of RDBMS frameworks are reached and any attempt to tame these large data sets by conventional means fails. (McLellan, 2013). Generally speaking MIS and IT professionals place the "big data" label on data volume that exceeds several hundreds of terabytes up to the order of several petabytes. (Manyika, et al.,

2011). Gartner, a leading technology research firm and IT advisor sums up big data in the following manner: " 'Big data' is high -volume, -velocity and -variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision-making" (Gartner, 2013).

Big data though in and of itself does not offer much value. And it is around this concept of value that a different perspective on big data has emerged. The research firm IDC has included the idea of value in its big data definition: big data is "a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis" (Gantz & Reinsel, 2011). Big data cannot be stored and maintained simply for the sake of doing it. Jay Parikh, the Vice President of Infrastructure Engineering at Facebook makes the following observation on bigd data: "If you aren't taking advantage of the data you're collecting, then you just have a pile of data, you don't have big data" (Chen, M., Mao, S., & Liu, Y., 2014; Constine, 2012).

STATEMENT OF THE PROBLEM

There has been little research conducted related to IT professionals and big data. Specifically there have been no studies to determine if relationships exist between IT professionals and their awareness of the maturity, value, hype, future trends, and practical understanding of big data solutions. When IT executives or other leaders in an organization clearly understand the knowledge, skills, and awareness of their staff in relation specific technologies they can develop plans to proceed with the deployment of big data solutions. They can also conduct training and awareness sessions to confirm that their staff has the level of understanding needed to ensure strategic goals are being met as aligned with big data.

PURPOSE OF THE STUDY

This study was designed to evaluate perceptions held by IT professionals in regards to big data. The study focuses on an assessment of IT professionals and their self-assigned understanding of big data in relation to their perceptions of the maturity, value, hype, and future trends of big data. The study also looks at levels of big data understanding in

conjunction with occupying a managerial role. The study should prove helpful in determining linkages between training and awareness of big data and the ability of an organization to advance big data initiatives.

HYPOTHESES

A survey tool will be used to address five hypotheses aimed at answering the intended areas of focus related to big data and the beliefs or perceptions held by IT professionals. The hypotheses are as follows:

Hypothesis H_1: An IT professional's understanding of the principles of big data and related technologies is independent of his / her responsibilities in a managerial capacity.

Hypothesis H_2: An IT professional's perception of the value of big data solutions is independent of his / her understanding of big data solutions.

Hypothesis H_3: An IT professional's expectation of beneficial future development in the area of big data technologies is independent of his / her understanding of big data solutions.

Hypothesis H_4: An IT professional's perception of hype in regards to the market recognition of big data technologies is independent of his / her understanding of big data solutions.

Hypothesis H_5: An IT professional's perception of the maturity of big data and related technologies is independent of his / her understanding of big data solutions.

SIGNIFICANCE OF THE STUDY

This study will represent a new contribution to the field of IT management. It will illustrate levels of significance of critical concepts as they relate to the big data understanding of IT professionals and their perception of the innovation and value of big data solutions. The study will also help determine reasons why certain organizations have been more successful with big data projects, and move more rapidly along the path of the adoption of advanced analytic solutions. This study will also provide a baseline at a particular point in time for the pulse of IT professionals. The study is being conducted when there are a lot of advancements and innovation in the tools and techniques used in relation to big data.

ASSUMPTIONS AND LIMITATIONS

The survey is focused specifically on IT professionals who self-identified themselves as filing such a role. The survey also asks respondents to state whether or not they fill a management role. Individuals are requested to rate their understanding of big data and their agreement with certain statements related to big data technologies. As with any study there could be inflation of results, however with the completely anonymous nature of the survey individuals gain no benefit from providing such answers. A final limitation of the survey is that respondents will self-select for participation in the study. A wide distribution of messaging via LinkedIn and Twitter will be used to reach as many IT professionals as possible.

METHODOLOGY

This study was designed to evaluate perceptions held by information technology (IT) professionals in regards to big data. The study focused on an assessment of IT professionals and their self-assigned understanding of big data in relation to their perceptions of the maturity, value, hype, and future trends of big data. The study also looked at big data understanding in conjunction with IT professionals occupying a managerial role. The framework for the study was a survey instrument administered on-line.

Sample Design

The theoretical study population was all IT professionals regardless of their particular job role. The actual study population was IT professionals in the United States. Using data provided by United States Bureau of Labor Statistics, the estimated number of IT professionals in the United States is slightly less than 4 million.

Table 1: IT Professionals in the United States (adapted from Bureau of Labor Statistics)

Job title	**May 2010 employment**	**January 2015***
Computer support specialists	607,100	661,739
Computer systems analyst	544,400	604,284
Software developers, applications	520,800	593,712
Software developers, systems software	392,300	455,068
Computer programmers	363,100	384,886
Network & computer system admins	347,300	395,922
Computer & information systems managers	307,900	335,611

Information Security Analysts, Web Developers, and Computer Network Architects	302,300	335,553
Database administrators	110,800	127,974
Total	**3,496,000**	**3,894,749**

** Estimate based on annual percent growth using BLS forecasts to 2020*

For global applicability, it was assumed that in regards to a series of questions related to a particular technology family (in this case big data), that US IT professionals would hold no particularly different attitudes or preconceptions about big data than IT professionals in other parts of the world.

Survey respondents participated in the survey via a custom response form that was developed on Google Forms. Study participants received an invitation to participate in the survey via Twitter alerts and the use of LinkedIn's email tool.

Survey Design

Part one of the survey was used to gather key demographics on the respondent. Questions were asked to determine the principal IT responsibility held by the respondent. Questions were also asked to establish a baseline regarding the respondent's industry of employment, length of time with their employer, and length of time as an IT professional. Respondents were also asked if they work in a managerial role.

Part two of the survey was used to gain insight into the respondent's perception of big data as it related to their awareness of the maturity, value, hype, future trends, and understanding of big data solutions. This section consisted entirely of five-point Likert scale questions so as to pinpoint precise results from the survey audience.

Part three of the survey was used to gain insight into the respondent's interaction or observation of the deployment (or non-deployment) of big data solutions in their organization. Questions also measured the overall sense of importance of big data in their organization and who is the visionary for big data within their organization.

The entire survey consisted of twenty questions and contained no open-ended or narrative style questions. The expectation was that the survey could be completed in less than five minutes in virtually every administration.

Validity and Reliability

The survey instrument's validity was evaluated using pre-test assessments. The researcher examined similar IT-oriented survey models to determine if the technical orientation of the questions was appropriate. Since the study was not being administered to determine causality of action there was no reason to evaluate internal validity (Cooper & Schindler, 2010).

The survey instrument was however evaluated for reliability, to determine its durability of repeated results. The Cronbach Alpha was measured using the open source statistical software program PSPP and returned at a high co-efficiency of internal consistency at 0.96 which indicates strong reliability for the scaled section of the survey tool.

Variables

The scaled section of the survey consisted of six main variables to enable the evaluation of the hypotheses.

The variables are as follows: Managerial position, maturity, value, hype, future trends, and understanding of big data solutions.

Managerial position was evaluated using a simple 'yes/no' question to determine if the respondent occupies a supervisor role. This variable was assessed by question 5.

Maturity was the respondent's perception of big data as being a solution or set of technologies that are sufficiently developed to be used in a production capacity and are capable of generating value. This variable was assessed by question 6.

Value was the respondent's perception of big data solutions being worth time, effort, and resources required for deployment. This variable was assessed by question 7

Hype was the respondent's perception of the industry and other IT professionals overstating the usefulness, value, or ease of deployment of big data solutions. This variable was assessed by question 8.

Future trends was the respondent's perception of near-term future for advances in big data solutions and technologies. This variable was assessed by question 9.

Understanding was the respondent's perception of his or her knowledge of the scope and individual components of big data solutions. This variable was assessed by question 10.

Data Collection

The collection of data, via an on-line survey tool, occurred over a one week period from August 5 to August 12, 2014. The survey was available twenty-four hours a day for the respondents.

To address any concerns about their responses being used in a manner that could directly affect their employment, all respondents were advised that results were confidential. They were also informed that the Google Forms survey tool stored no personally identifiable information along with their responses. Survey results were uniquely identified only by a data-time stamp in the co-located data collection spreadsheet.

Data Analysis

Each of the five hypotheses was tested using a One-Way Analysis of Variance test in the open source software PSPP with a significance level of 0.05.

Results

There were 76 responses to the survey which represents an outstanding response rate for a survey administered solely by non-direct contact between the researcher and respondents. The number of responses exceeded a pre-study goal of 68 responses. With the overall population of IT professionals in the United States being an estimated 4 million this response rate allows a confidence interval of 90% to be established with a 9.5% +/- margin of error.

There was a relatively rapid initial response to the survey over the first two days with a few residual responses over the remaining period the survey was open.

There appear to be no indications of extreme views that might result from a non-response bias. In addition, it was assumed that the results accurately reflect the population since there was a good distribution of respondents with varying degrees of experience, fields of employment, and work role in the IT field.

Test of Hypothesis H_1

Hypothesis H_1 stated (null): An IT professional's understanding of the principles of big data and related technologies is independent of his / her responsibilities in a managerial capacity.

This hypothesis was evaluated by comparing responses to question 5 and question 6 on the survey. Question 5 was "Do you occupy a role where you manage other IT staff?" Question 6 was "How would you categorize your understanding of big data?"

Based upon survey responses the respondents were divided into two groups, those who reported that they held a managerial role and those who said they did not. The One-Way Analysis of Variance of these groups had a p-value of .155, which is greater than the established significance level of .05. Based on these findings, the null is not rejected. The data produced insufficient evidence to conclude that being a manager has any impact on an individual's understanding of big data solutions.

Table 2: One-Way Analysis of Variance – Hypothesis H1

	Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Between Groups	2.24	1	2.24	2.06	.155
Within Groups	80.12	74	1.08		
Total	82.36	75			

Table 3: Managerial Status by Big Data Understanding - Raw Data

		Big Data Understanding [1]					
		1	**2**	**3**	**4**	**5**	**Total**
Manager	Count	3.00	3.00	13.00	12.00	2.00	33.00
	Row %	9.09%	9.09%	39.39%	36.36%	6.06%	100.00%
	Column %	75.00%	33.33%	50.00%	46.15%	18.18%	43.42%
	Total %	3.95%	3.95%	17.11%	15.79%	2.63%	43.42%
Non-Manager	Count	1.00	6.00	13.00	14.00	9.00	43.00
	Row %	2.33%	13.95%	30.23%	32.56%	20.93%	100.00%
	Column %	25.00%	66.67%	50.00%	53.85%	81.82%	56.58%
	Total %	1.32%	7.89%	17.11%	18.42%	11.84%	56.58%
Total	Count	4.00	9.00	26.00	26.00	11.00	76.00
	Row %	5.26%	11.84%	34.21%	34.21%	14.47%	100.00%
	Column %	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	Total %	5.26%	11.84%	34.21%	34.21%	14.47%	100.00%

[1] *Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely*

Test of Hypothesis H_2

Hypothesis H_2 stated (null): An IT professional's perception of the value of big data solutions is independent of his / her understanding of big data solutions. This hypothesis was evaluated by comparing responses to question 6 and question 8 on the survey. Question 6 was "How

would you categorize your understanding of big data?" Question 8 was answered using a Likert scale with values starting with "Strongly Disagree" going to "Strongly Agree." The statement was: "I feel that big data solutions are worth the investment of time, resources, and money."

Based upon survey responses the respondents were divided into five groups determined by their response to Question 6 assessing their understanding of big data solutions. The One-Way Analysis of Variance of these groups had a p-value of .205, which is greater than the established significance level of .05. Based on these findings, the null is not rejected. The data produced insufficient evidence to conclude that an individual's understanding of big data solutions impacts their perception of the value of big data solutions.

Table 4: One-Way Analysis of Variance – Hypothesis H2

	Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Between Groups	3.78	4	.94	1.52	.205
Within Groups	44.01	71	.62		
Total	47.79	75			

Table 5: Big Data Understanding by Assessment of Big Data Value - Raw Data

		Value of Big Data Solutions [2]					
Big Data Understanding [1]		**SD**	**D**	**N**	**A**	**SA**	**Total**
1	Count	.00	.00	2.00	2.00	.00	4.00
	Row %	.00%	.00%	50.00%	50.00%	.00%	100.00%
	Column %	.00%	.00%	14.29%	4.76%	.00%	5.26%
	Total %	.00%	.00%	2.63%	2.63%	.00%	5.26%
2	Count	.00	.00	3.00	4.00	2.00	9.00
	Row %	.00%	.00%	33.33%	44.44%	22.22%	100.00%
	Column %	.00%	.00%	21.43%	9.52%	11.76%	11.84%
	Total %	.00%	.00%	3.95%	5.26%	2.63%	11.84%
3	Count	1.00	1.00	4.00	17.00	3.00	26.00
	Row %	3.85%	3.85%	15.38%	65.38%	11.54%	100.00%
	Column %	100.00%	50.00%	28.57%	40.48%	17.65%	34.21%
	Total %	1.32%	1.32%	5.26%	22.37%	3.95%	34.21%
4	Count	.00	1.00	4.00	14.00	7.00	26.00
	Row %	.00%	3.85%	15.38%	53.85%	26.92%	100.00%
	Column %	.00%	50.00%	28.57%	33.33%	41.18%	34.21%
	Total %	.00%	1.32%	5.26%	18.42%	9.21%	34.21%
5	Count	.00	.00	1.00	5.00	5.00	11.00

		Value of Big Data Solutions [2]					
Big Data Understanding [1]		**SD**	**D**	**N**	**A**	**SA**	**Total**
	Row %	.00%	.00%	9.09%	45.45%	45.45%	100.00%
	Column %	.00%	.00%	7.14%	11.90%	29.41%	14.47%
	Total %	.00%	.00%	1.32%	6.58%	6.58%	14.47%
Total	Count	1.00	2.00	14.00	42.00	17.00	76.00
	Row %	1.32%	2.63%	18.42%	55.26%	22.37%	100.00%
	Column %	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	Total %	1.32%	2.63%	18.42%	55.26%	22.37%	100.00%

[1] *Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely*
[2] *Value of Big Data Solutions: SD-Strongly Disagree, D-Disagree, N-Neutral, A-Agree, SA-Strongly Agree*

Test of Hypothesis H_3

Hypothesis H_3 stated (null): An IT professional's expectation of beneficial future development in the area of big data technologies is independent of his / her understanding of big data solutions. This hypothesis was evaluated by comparing responses to question 6 and question 10 on the survey. Question six is "How would you categorize your understanding of big data?" Question 10 was answered using a Likert scale with values starting with "Strongly Disagree" going to "Strongly Agree." The statement was: "I believe that there will be significant development and improvement in big data technologies over the next five years."

Based upon survey responses the respondents were divided into five groups determined by their response question number six assessing their understanding of big data solutions. The One-Way Analysis of Variance had a p-value of .001, which is less than the established significance level of .05. Based on these findings, the null is rejected. There is enough evidence to conclude that an individual's understanding of big data solutions impacts their perception of the value of big data solutions.

Since the null was rejected, the use of a Post-Hoc test following the One-Way Analysis of Variance is required to identify the groups between which there was the greatest difference.

The Post-Hoc Tukey Test was used to perform multiple comparisons of the means. Table 7 provides the results of the test. The p-value for big data understanding 1 vs. big data

understanding 4 was .003, and the p-value for big data understanding 1 vs. big data understanding 5 was .003. Since these values are below the established significance level of .05, it can be concluded that the differences between these groups are significant.

Table 6: One-Way Analysis of Variance – Hypothesis H3

	Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Between Groups	6.39	4	1.60	5.06	.001
Within Groups	22.40	71	.32		
Total	28.79	75			

Table 7: Multiple Comparisons of Means / Tukey Test

(I) Big Data Understanding [1]	**(J) Big Data Understanding [1]**	**Mean Difference (I - J)**	**Std. Error**	**Sig.**	**Lower Bound**	**Upper Bound**
1	**2**	-.94	.34	.050	-1.89	.00
	3	-.77	.30	.091	-1.61	.08
	4	-1.15	.30	.003	-2.00	-.31
	5	-1.23	.33	.003	-2.15	-.31
2	**1**	.94	.34	.050	.00	1.89
	3	.18	.22	.928	-.43	.78
	4	-.21	.22	.870	-.82	.40
	5	-.28	.25	.795	-.99	.42
3	**1**	.77	.30	.091	-.08	1.61
	2	-.18	.22	.928	-.78	.43
	4	-.38	.16	.110	-.82	.05
	5	-.46	.20	.168	-1.02	.11
4	**1**	1.15	.30	.003	.31	2.00
	2	.21	.22	.870	-.40	.82
	3	.38	.16	.110	-.05	.82
	5	-.07	.20	.996	-.64	.49
5	**1**	1.23	.33	.003	.31	2.15
	2	.28	.25	.795	-.42	.99
	3	.46	.20	.168	-.11	1.02
	4	.07	.20	.996	-.49	.64

[1] *Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely*

Table 8: Big Data Understanding by Assessment of Big Data Future Growth - Raw Data

		Future Growth of Big Data Solutions / Technology [2]					
Big Data Understanding [1]		**SD**	**D**	**N**	**A**	**SA**	**Total**
1	Count	.00	.00	2.00	2.00	.00	4.00
	Row %	.00%	.00%	50.00%	50.00%	.00%	100.00%
	Column %	.00%	.00%	40.00%	6.25%	.00%	5.26%
	Total %	.00%	.00%	2.63%	2.63%	.00%	5.26%
2	Count	.00	.00	.00	5.00	4.00	9.00
	Row %	.00%	.00%	.00%	55.56%	44.44%	100.00%
	Column %	.00%	.00%	.00%	15.63%	10.26%	11.84%
	Total %	.00%	.00%	.00%	6.58%	5.26%	11.84%
3	Count	.00	.00	2.00	15.00	9.00	26.00
	Row %	.00%	.00%	7.69%	57.69%	34.62%	100.00%
	Column %	.00%	.00%	40.00%	46.88%	23.08%	34.21%
	Total %	.00%	.00%	2.63%	19.74%	11.84%	34.21%
4	Count	.00	.00	1.00	7.00	18.00	26.00
	Row %	.00%	.00%	3.85%	26.92%	69.23%	100.00%
	Column %	.00%	.00%	20.00%	21.88%	46.15%	34.21%
	Total %	.00%	.00%	1.32%	9.21%	23.68%	34.21%
5	Count	.00	.00	.00	3.00	8.00	11.00
	Row %	.00%	.00%	.00%	27.27%	72.73%	100.00%
	Column %	.00%	.00%	.00%	9.38%	20.51%	14.47%
	Total %	.00%	.00%	.00%	3.95%	10.53%	14.47%
Total	Count	.00	.00	5.00	32.00	39.00	76.00
	Row %	.00%	.00%	6.58%	42.11%	51.32%	100.00%
	Column %	.00%	.00%	100.00%	100.00%	100.00%	100.00%
	Total %	.00%	.00%	6.58%	42.11%	51.32%	100.00%

[1] *Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely*

[2] *Future Growth of Big Data Solutions / Technology: SD-Strongly Disagree, D-Disagree, N-Neutral, A-Agree, SA-Strongly Agree*

Test of Hypothesis H_4

Hypothesis H_4 stated (null): An IT professional's perception of hype in regards to the market recognition of big data technologies is independent of his / her understanding of big data solutions. This hypothesis was evaluated by comparing responses to question 6 and question 9 on the survey. Question six is "How would you categorize your understanding of big data?" Question 9 was answered using a Likert scale with values starting with "Strongly Disagree" going to "Strongly Agree." The statement was: "I believe that there is a significant amount of hype in the marketplace in relation to big data."

Based upon survey responses the respondents were divided into five groups determined by their response question number six assessing their understanding of big data solutions. The One-Way Analysis of Variance had a p-value of .017, which is less than the established significance level of .05. Based on these findings, the null is rejected.

Since the null was rejected, the use of a Post-Hoc test following the One-Way Analysis of Variance is required to identify the groups between which there was the greatest difference.

The Post-Hoc Tukey Test was used to perform multiple comparisons of the means. Table 6 provides the results of the test. The p-value for big data understanding 3 vs. big data understanding 4 was .023. Since this value is below the established significance level of .05, it can be concluded that the differences between these groups are significant.

Table 9: One-Way Analysis of Variance – Hypothesis Four

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	6.68	4	1.67	3.25	.017
Within Groups	36.48	71	.51		
Total	43.16	75			

Table 10: Multiple Comparisons of Means / Tukey Test

(I) Big Data Understanding [1]	(J) Big Data Understanding [1]	Mean Difference (I - J)	Std. Error	Sig.	Lower Bound	Upper Bound
1	**2**	-.50	.43	.773	-1.71	.71

(I) Big Data Understanding [1]	(J) Big Data Understanding [1]	Mean Difference (I - J)	Std. Error	Sig.	Lower Bound	Upper Bound
	3	-.35	.38	.896	-1.42	.73
	4	-.96	.38	.103	-2.04	.12
	5	-.68	.42	.484	-1.85	.49
2	**1**	.50	.43	.773	-.71	1.71
	3	.15	.28	.981	-.62	.93
	4	-.46	.28	.462	-1.24	.31
	5	-.18	.32	.980	-1.08	.72
3	**1**	.35	.38	.896	-.73	1.42
	2	-.15	.28	.981	-.93	.62
	4	-.62	.20	.023	-1.17	-.06
	5	-.34	.26	.691	-1.06	.39
4	**1**	.96	.38	.103	-.12	2.04
	2	.46	.28	.462	-.31	1.24
	3	.62	.20	.023	.06	1.17
	5	.28	.26	.814	-.44	1.00
5	**1**	.68	.42	.484	-.49	1.85
	2	.18	.32	.980	-.72	1.08
	3	.34	.26	.691	-.39	1.06
	4	-.28	.26	.814	-1.00	.44

[1]Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely

Table 11: Big Data Understanding by Assessment of Hype related to Big Data - Raw Data

		Is There Significant Hype Related to Big Data [2]					
Big Data Understanding [1]		**SD**	**D**	**N**	**A**	**SA**	**Total**
1	Count	.00	.00	2.00	2.00	.00	4.00
	Row %	.00%	.00%	50.00%	50.00%	.00%	100.00%
	Column %	.00%	.00%	16.67%	5.26%	.00%	5.26%
	Total %	.00%	.00%	2.63%	2.63%	.00%	5.26%
2	Count	.00	.00	1.00	7.00	1.00	9.00
	Row %	.00%	.00%	11.11%	77.78%	11.11%	100.00%
	Column %	.00%	.00%	8.33%	18.42%	4.17%	11.84%
	Total %	.00%	.00%	1.32%	9.21%	1.32%	11.84%
3	Count	.00	1.00	6.00	15.00	4.00	26.00
	Row %	.00%	3.85%	23.08%	57.69%	15.38%	100.00%
	Column %	.00%	50.00%	50.00%	39.47%	16.67%	34.21%
	Total %	.00%	1.32%	7.89%	19.74%	5.26%	34.21%
4	Count	.00	.00	3.00	8.00	15.00	26.00
	Row %	.00%	.00%	11.54%	30.77%	57.69%	100.00%
	Column %	.00%	.00%	25.00%	21.05%	62.50%	34.21%
	Total %	.00%	.00%	3.95%	10.53%	19.74%	34.21%
5	Count	.00	1.00	.00	6.00	4.00	11.00
	Row %	.00%	9.09%	.00%	54.55%	36.36%	100.00%
	Column %	.00%	50.00%	.00%	15.79%	16.67%	14.47%
	Total %	.00%	1.32%	.00%	7.89%	5.26%	14.47%
Total	Count	.00	2.00	12.00	38.00	24.00	76.00
	Row %	.00%	2.63%	15.79%	50.00%	31.58%	100.00%
	Column %	.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	Total %	.00%	2.63%	15.79%	50.00%	31.58%	100.00%

[1] *Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely*

[2] *Significant Hype related to Big Data: SD-Strongly Disagree, D-Disagree, N-Neutral, A-Agree, SA-Strongly Agree*

Test of Hypothesis H_5

Hypothesis H_5 stated (null): An IT professional's perception of the maturity of big data and related technologies is independent of his / her understanding of big data solutions. This hypothesis was evaluated by comparing responses to question 6 and question 7 on the survey. Question six is "How would you categorize your understanding of big data?" Question 7 was answered using a Likert scale with values starting with "Strongly Disagree" going to "Strongly Agree." The statement was: "I believe that big data is a mature information technology framework."

Based upon survey responses the respondents were divided into five groups determined by their response question number six assessing their understanding of big data solutions. The One-Way Analysis of Variance had a p-value of .251, which is greater than the established significance level of .05. Based on these findings, the null is not rejected. The data produced insufficient evidence to conclude that an individual's assessment of big data being a mature information technology framework is based on their level of big data understanding.

Table 12: One-Way Analysis of Variance – Hypothesis Five

	Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Between Groups	4.36	4	1.09	1.37	.251
Within Groups	56.27	71	.79		
Total	60.63	75			

Table 13: Big Data Understanding by Assessment of the Maturity of Big Data - Raw Data

		Maturity of Big Data [2]					
Big Data Understanding [1]		**SD**	**D**	**N**	**A**	**SA**	**Total**
1	Count	.00	.00	3.00	1.00	.00	4.00
	Row %	.00%	.00%	75.00%	25.00%	.00%	100.00%
	Column %	.00%	.00%	11.11%	5.88%	.00%	5.26%
	Total %	.00%	.00%	3.95%	1.32%	.00%	5.26%
2	Count	.00	4.00	3.00	2.00	.00	9.00
	Row %	.00%	44.44%	33.33%	22.22%	.00%	100.00%
	Column %	.00%	14.81%	11.11%	11.76%	.00%	11.84%
	Total %	.00%	5.26%	3.95%	2.63%	.00%	11.84%
3	Count	1.00	14.00	7.00	4.00	.00	26.00
	Row %	3.85%	53.85%	26.92%	15.38%	.00%	100.00%
	Column %	25.00%	51.85%	25.93%	23.53%	.00%	34.21%
	Total %	1.32%	18.42%	9.21%	5.26%	.00%	34.21%
4	Count	1.00	5.00	12.00	8.00	.00	26.00
	Row %	3.85%	19.23%	46.15%	30.77%	.00%	100.00%
	Column %	25.00%	18.52%	44.44%	47.06%	.00%	34.21%
	Total %	1.32%	6.58%	15.79%	10.53%	.00%	34.21%
5	Count	2.00	4.00	2.00	2.00	1.00	11.00
	Row %	18.18%	36.36%	18.18%	18.18%	9.09%	100.00%
	Column %	50.00%	14.81%	7.41%	11.76%	100.00%	14.47%
	Total %	2.63%	5.26%	2.63%	2.63%	1.32%	14.47%
Total	Count	4.00	27.00	27.00	17.00	1.00	76.00
	Row %	5.26%	35.53%	35.53%	22.37%	1.32%	100.00%
	Column %	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	Total %	5.26%	35.53%	35.53%	22.37%	1.32%	100.00%

[1] *Big Data Understanding: 1-No Understanding, 2-Basic Understanding, 3-Familiarity, 4-Firm Understanding, 5-Understand Completely*

[2] *Maturity of Big Data: SD-Strongly Disagree, D-Disagree, N-Neutral, A-Agree, SA-Strongly Agree*

SUMMARY AND CONCLUSIONS

The study of big data related to its utilization in management contexts is a relatively new phenomenon since the creation data analysis related to big data has only been around for 10 – 15 years. As the technologies related to big data continue to evolve organically and from other existing technologies there will be specific methods to apply big data solutions to business problems. As such any research, which helps to guide decision-making in regards to identifying valid use cases and training for staff, will prove to be very beneficial.

Conclusions

The results from various statistical tests supported rejecting two of the five null hypotheses. Two nulls were rejected based on a One-Way Analysis of Variance. The first null rejected was for Hypothesis H3: An information technology professional's expectation of beneficial future development in the area of big data technologies is independent of his / her understanding of big data solutions. The second null rejected was for Hypothesis H4: An information technology professional's perception of hype in regards to the market recognition of big data technologies is independent of his / her understanding of big data solutions.

Hypothesis H_3 scored at .001 on the One-Way Analysis of Variance test. This result was well below the threshold of .05 which indicates a strong connection between big data understanding and expectation of future development in big data technologies. Follow-on use of the Tukey test confirmed this finding and identified the greatest differences occurring between the lowest level of big data understanding (1) and two highest levels of big data understanding (4 and 5).

Hypothesis H_4 scored at .017 on the One-Way Analysis of Variance test. This result placed it below the threshold of .05 which indicates a strong connection between big data understanding and recognition of hype related to big data technologies. Follow-on use of the Tukey test confirmed this finding and identified the greatest differences occurring between those responding with “Familiarity” and “Firm Understanding” as a characterization of their level of big data understanding.

Suggestions for Further Research

The findings of this study indicate that IT professionals are still becoming familiar with big data and related technologies. The study produced evidence that the more an individual knows about big data, the more they expect greater future developments and enhancements with the technology. With IT professionals expecting improvements in big data there is great optimism for the future of big data. The study also indicated a lot of future efforts are focused on big data.

Future research should focus more closely on the question of the maturity of big data solutions in an in-depth manner. Such a focused study could pinpoint if any issues exist with the tools used to build big data solutions or if there is an issue with the tools used to consume analytics derived from big data stores.

Another area of future inquiry could focus on the specific aspects of big data solutions that seem to generate the best return on investment. Enterprises could gain value by targeting efforts on big data solutions focusing on the key qualities identified in the study such as items like “gaining competitive advantage”, “customer retention”, and “product development.”

In terms of a technology-oriented study, updated research should be undertaken to see how many of the on-premise solutions identified by respondents moved to the cloud.

Finally, a research effort could be undertaken to examine the manner and methods by which big data acceptance and growth occurs in organizations. For instance, attention could be paid to the evolution of who fills the prime visionary role or what methods and techniques those individuals use to promote big data within their corresponding organizations.

References

Bureau of Labor Statistics. (2014, January 8). Computer and Information Systems Managers. Retrieved from Bureau of Labor Statistics: http://www.bls.gov/ooh/management/computer-and-information-systems-managers.htm

Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. Mobile Networks and Applications, 19(2, 171-209.

Constine, J. (2012). How Big Is Facebook's Data? 2.5 Billion Pieces Of Content And 500+ Terabytes Ingested Every Day. Retrieved from TechCrunch: http://bit.ly/1nKWiHM

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. IDC iview, 1-12.

Gartner. (2013). IT Glossary - Big Data. Retrieved from Gartner: http://gtnr.it/1rP0ANn

Manyika, J., Chui, M. B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. New York: McKinsey Global Institute.

McLellan, C. (2013). Big data: An overview. Retrieved from ZDNet: http://www.zdnet.com/big-data-an-overview_p2-7000020785