I. Introduction

1. Organisation budgets are shrinking and non-traditional data providers and big data are putting increasing pressure on official statistics organisations to be faster and more efficient without jeopardising the core values of official statistics which are data quality and trust.

2. There is a multitude of ways to share and reuse software amongst organisations and the statistics industry employs (and has employed) a variety of different approaches.

3. This paper is focussed specifically on open source software and investigates the factors influencing the adoption of open source software within official statistics organisations.

4. The primary objective of the study was to explore open source software benefits for the statistics industry, measure the current state of adoption, and identify the motivators and amotivators affecting adoption. Secondary objectives of the investigation explore where organisations are utilising open source software and whether organisations consider open source software as a means of achieving business goals.

5. This paper presents a literature review of the factors affecting open source software adoption in organisations and the results from a survey undertaken in the fall of 2013. The paper summarises the statistical industry’s current position in regards to adopting open source software and observations regarding the existence of common or principle factors affecting adoption.
II. Review of Current Thinking

A. Open Source Software, a definition

6. In the later part of the 20th century, open systems and open standards became important parts of Information Technology. Building on that foundation, open source software has also become an important part of enterprise architecture. As the range of software applications that open source projects addressed expanded, the profile of organizations that chose to use open source software spread from academic, defense, and scientific research organisations, to commercial companies. As the 21st century began, use of open source software has spread to all geographies and industries – although some have adopted open source faster than others (Forrester, 2007).

7. Open source software (OSS) is a term used to classify software with specific characteristics of ‘openness’. The definition of open source software varies in the literature. The UK Government definition (Service, U. G. P., 2011) is that open source software is a software development and distribution model where the software license guarantees certain freedoms. These include the right to access and modify the source code and to reuse and redistribute the software without constraint or undue cost. Marsan (2012) defines open source software as software developed in a public collaborative manner, for any application, and available under a license allowing free source code accessibility, reuse, modification, and redistribution by users. Techopedia (2012) defines open source software in the following manner “open source software is software created by a community of people who are dedicated to collaborating to produce true innovation and allow the evolution of new and better software”. For this paper, the UK Government definition of open source software will be used. Focussing the definition on the end product and not the manner in which it is produced is also supported by del Bianco (2011) who observes that open source software is no longer the product of lone coders and that industrial strength open source software is often developed by organised communities and sometimes even by major software companies applying the same rigorous processes and high quality standards as commercial products.

8. Generally speaking, the strength of open source lies in its: no license costs, interoperability, easier integration and customisation, compliance with open technology and data standards and freedom from vendor lock in. Studies have shown that the benefits of open source generally materialise in the medium to long term. Furthermore, because open source software is free, there is greater flexibility in selecting the level of services or support that a customer wants to pay for, if at all. (Service, U. G. P., 2011)

B. Open Source Software, factors affecting adoption

9. There is very little information in the models, theories, and frameworks to explain the adoption of open source software in organizations. While the papers emphasize different factors, the common thread is that the adoption of OSS is not the same as adopting commercial software and a change has to happen within the organization in order to successfully implement OSS. This change involves organizational change in the form of policy, governance, staff skills/training, and so forth. These changes however relate to the successful implementation of open source software. As has been indicated by Marsan (2013), the decision whether or not to adopt the software is broader in nature and is influenced by external factors, organizational factors, and factors related to the OSS product itself and its alignment with the needs of the organization.

10. The papers all looked at OSS adoption although they were not consistent in how they approached the topic. The common thread throughout many of the papers was that OSS adoption has many similarities to software adoption in general but, it is different in a variety of ways. Additionally, OSS has a specific
culture associated with it that leads some people to be neutral, supporters, or detractors of OSS regardless of the specific situation / business case. The external context of an organization and even the organizational context will influence adoption decisions. The quality and consistency of the messages regarding OSS which are received by the IT staff and decision makers will influence their openness to considering OSS as a viable solution. High-level OSS policies and management commitment are important however individual roles within an organization will still play a role in determining the criteria which are identified and prioritized regarding adoption / non-adoption of OSS software. The industry and the national/regional market within which the firm operates affect OSS adoption rates and the rating of adoption criteria. Adoption of OSS requires an organizational change. Internal barriers must be addressed, governance introduced, risks identified and managed on a case by case basis, skills and knowledge imparted on the teams who will be affected, support for the initiative built up, purchasing policies adapted, to name a few. As indicated by Forrester (2007), once an organization has transitioned from considering OSS to adopting OSS and then to adopting OSS for a large number of business capabilities, OSS becomes an integral part of the business and is considered in parallel with other options, such as commercial software. From this perspective then, it is reasonable to consider that the considerations for adopting OSS change as the organization changes while it moves through this transition from the first OSS project to OSS being an integral part of the business.

a. The most common factors affecting adoption which have been identified in the papers are:
   i. Concerns regarding service & support
   ii. Cost / total cost of ownership
   iii. Product capabilities/maturity
   iv. Lack of technical knowledge / skills in the organization for OSS product
   v. Difficulty of adoption / integration
   vi. Viability of the open source community
   vii. Staff knowledge / skills / familiarity with how OSS is different
   viii. Concerns regarding intellectual property / licensing
   ix. Concerns regarding the security of the software

b. Secondary factors are:
   i. Fit for purpose - ability to meet business goals
   ii. Complexity - difficulty to implement or manage
   iii. Adherence to standards
   iv. Software quality – end user satisfaction
   v. Software enhancements – innovation over time
   vi. There were additional factors identified in a few papers but which were not prevalent.

III. The Study

A. Objectives of the study

11. The objective of the study was to clarify how (or if) open source software can benefit the statistics industry, measure where the industry is in regards to adopting open source software, and identify both the motivators and amotivators which are either facilitating or blocking the adoption of open source software. Secondary objectives of the study were to identify whether organisations consider open source software as a means of achieving business goals and identifying where open source software is being utilised in their operations.

12. The review of current thinking concluded that the network model of diffusion of innovation as a process was applicable and relevant to open source software adoption in organisations. As per the model, the adoption of open source software is influenced by external factors, the characteristics of the
organization, and the characteristics of the innovation, in this case, open source software. This study evaluated the characteristics of the organisation and the characteristics of the innovation. Additionally, it determined if there was a relationship between an organisation's experiences with open source software and what it identified as the motivating and amotivating factors influencing adoption.

13. The review of current thinking also highlighted that open source software adoption varied by industry. This study was limited to the official statistics industry.

14. The official statistics industry is large and varied and the concept of what constitutes an official statistics organisation is not clearly defined. National statistical organisations, central banks, and some public sector international organisations can be clearly classified and widely accepted as official statistics organisations. For the purposes of this study, the unit of analysis for the research will be ‘national and international public sector statistics organisations’.

15. This study intended to answer the following questions:
   a. Are national and international statistical agencies using open source software?
   b. Where in their operations are statistical organisations using open source software?
   c. Do statistical agencies view open source software as a means to achieving their goals?
   d. Which attributes of open source software are of most concern to statistical agencies?
   e. Which OSS adoption good practices are important to statistical agencies?
   f. Which OSS attributes are important to organisational employees & are they influenced by employee demographics?
   g. Which elements of OSS support are in highest demand?
   h. Which characteristics of OSS are most important to statistical agencies? Are they influenced by other organisational factors?

B. Respondents

16. Chief Statisticians along with participants from numerous international technical meetings during the period 2011 to 2013 were contacted for this study. Participants were from all geographic regions and consisted of both national and international public sector statistical organisations.

17. Responses to the survey were received from all major geographic regions. The responses from international organisations mostly were based in Europe and North America which explains the variance between #countries per region and #responses per region. The geographic distribution of responses is very positive. The responses largely represent organisations in upper-middle-income and high-income economies.
III. Findings

A. Geographic coverage

18. The geographic coverage achieved by the study was robust with responses from organisations in thirty-five countries and from all geographic regions of the world. Qu (2011) suggested that open source software became less attractive as countries became more developed. The comparison of open source software adoption among the respondents indicates a very high level of adoption with most respondents having either already adopted OSS and using it in production or alternatively, currently piloting OSS. In addition, the regrouping of responding countries according to World Bank Income Groups shows that almost all responses are from organisations based in either upper-middle-income or high-income economies. Whether this is unique to the statistics industry or not, this finding definitely contradicts what Qu (2011) put forward with his work. Amongst organisations in the statistics industry, countries with high levels of economic development are definitely adopters and users of open source software. What is unknown and which this study cannot comment on is the level of OSS adoption in the statistics industry amongst countries that are in other World Bank Income Groups representing lesser economies.
B. Influence of information

19. The study looked at the external information being received by and influencing employees regarding open source software (Macredie, 2011) by measuring the orientation of messages in the media (positive or negative) towards OSS and also by measuring the orientation of the organisation towards OSS. The information about OSS in the media was identified as being positive, almost without exception. The orientation of organisations towards OSS was also found to be neutral or open to OSS with a high level of agreement amongst respondents.

20. These items were not evaluated to determine if there were relationships between them and other concepts being measured however, the positive nature of the messages in the media and the positive orientation of organisations towards OSS indicates the presence of a positive and receptive environment ((Macredie, 2011), (Marsan (2012)) and provides positive signals for any industry efforts to increase OSS adoption and usage.

C. Organisation Characteristics

21. There were more than twenty responses from nine international statistical organisations and more than forty responses from thirty seven national organisations indicating a very good representation from both groups. The organisations were mostly large, as measured by number of employees, although organisations of all sizes participated in the survey.

22. The possible existence of a relationship between the size of the organisation and the organisation's position on the innovation adoption scale was explored and the findings were that organisational size was of no significance as to whether the organisation adopts OSS or not. This is inconsistent with findings from other studies ((Hammond, 2009), (Macredie, 2011), (Qu, 2011)) which cannot be explained other than to reflect on the positive factors, such as the high level of openness towards OSS and the high levels of adoption of OSS in the statistics industry. This could be an interesting area for future research.

23. One of the objectives of this study was to determine if national and international statistical organisations were using open source software. Most organisations indicated that they were using OSS in production or were in the process of piloting it. Furthermore, the correlation analysis which looked at a possible relationship between the type of organisation and the adoption of OSS found that no relationship existed meaning that both national and international organisations are actively using OSS.

24. There is a relationship however between the current position of an organisation on the innovation adoption scale and the future plans for using OSS in 24 months. This relationship is of moderate statistical strength and 23% of the variation in the variables can be explained by the relationship. This supports Forrester (2007) which states that a large majority of organisations who have considered open source software have ended up adopting it and once adopted, usage expands from simple applications to mission-critical applications and customer-facing applications.

25. That being said however, OSS adoption is not universal. There was a small minority of organisations which identified themselves as being either resistant or very resistant to OSS adoption. It is unknown if this is because of individuals in powers of authority being detractors, organisational factors, or external factors such as government policies ((Marsan, 2012), (Marsan, 2013)).
D. Important OSS Attributes

26. One of the goals of the investigation was to identify which OSS attributes are important to organisation employees and whether they are influenced by employee demographics.

27. Respondents to the survey represented a broad age group and the norm was 15-24 years of professional experience. Senior IT managers, IT technical staff, senior business managers, and project managers made up almost all of the respondents. Almost all of the respondents had been involved in OSS projects and more than half of the individuals indicated that they are using OSS in production.

28. Employee ratings of OSS attributes to identify which were the most important resulted in “Ability to use OSS without restrictions”, “Support in OSS for open standards”, and “Ability to reduce dependencies on vendors” as the most important attributes. The attributes identified as the least important were “Ability to participate and contribute to the OSS community”, “Ability to participate and contribute to OSS development”, and “Ability to create and distribute derivative works”.

29. It was found that there is a small but definite relationship between the employee role in the organisation and the ranking of OSS attributes, in fact 6% of the variation in one variable is associated with the other. This finding supports the assertion that an employee’s role influences which OSS factors are most important (del Bianco, 2011).

30. Neither the role of the employee nor the number of years of experience (Marsan, 2012) was shown to have any statistical relationship with the attributes of OSS or with the opinion of OSS as a means of achieving business goals.

E. Business Drivers

31. The individual respondents’ perspective on the business drivers for OSS was evaluated. Respondents were asked if OSS could help achieve the specified business goals and there was a high level of agreement in scoring all of the items. The item which scored the lowest on the rating scale was “acquire and retain clients/users/customers”.

32. There are no relationships of statistical significance between the respondents’ perspective on the business drivers and their years of experience, their role in the organisation, or their self-assessed position on the innovation adoption process ((del Bianco, 2011), (Marsan, 2012)).

33. There were two financial questions in the table, “lower total cost of ownership” and “lower operating costs”. The current thinking on OSS is divided on whether OSS can lower total cost of ownership. It has been stated that while OSS is free, it comes with higher operating costs than traditional software models. The literature on OSS is almost unanimous that the “operating costs” for OSS are in fact higher than they are for traditional software models ((Brandel, 2010), (Bwired, 2009)). This statement is contradicted by the responses to this question. It is possible, but unlikely, that OSS operating costs are different in the statistical industry as opposed to other industries. One possible hypothesis is that public sector statistical organisations use OSS to replace in-house development (as opposed to commercial software) and so the difference in operating costs between OSS and in-house developed software is the basis for the cost comparison. Another possible hypothesis is that public sector statistical organisations do not monitor and capture their operating costs at the granularity required in order to respond accurately to this question. This situation should be seen as a risk to the industry when considering OSS adoption as any goal of reducing costs may result in transferring highly visible and direct costs related to commercial licensing and support to hidden operating costs for OSS. Managing this risk would be an important element of any OSS adoption strategy if cost reduction was a stated goal.
34. Unlike the UK Government (Service UGP, 2011), the majority of organisations do not have policies regarding OSS software but any policies which have been established support OSS adoption. No organisation indicated that they had a policy to discourage OSS adoption, including the organisations which indicated that they were resistant or very resistant to OSS adoption. Although policies have been identified in the review of current thinking as an important element of OSS adoption, it is obvious from this study that not having policies does not prevent OSS adoption. The decision whether or not to adopt an OSS policy should therefore be considered on a case by case basis depending upon the organisational context and environment and not seen as a necessary precursor to OSS adoption.

35. There was a high level of consistency and agreement in the importance rankings given to a list of organisational goals during the current planning cycle. The top ranked goals which were all rated as somewhat important or important with a high level of agreement were “Improve integration between applications”, “Increase process-oriented statistical production (reduce silos)”, “Reduce IT costs”, “Increase industrialization of statistics”, “Increase collaboration with other statistical agencies”, “Adopt or increase use of DDI and/or SDMX standards”, “Use IT to increase innovation”, “Increase ability to meet unmet demands for IT services”, “Address IT staffing and skills shortages”, “Support regulatory requirements”, “Adopt the Generalized Statistical Business Process Model (GSBPM)”, and “Improve communication of IT value to the business”. The remaining three items in the list received much less agreement on their importance during the current planning cycle. The items “Reducing number of software vendors we work with”, “Reduce the number of in-house developed applications”, and “Expand use of open source software” were rated between Neutral and Somewhat Important but the item of note is that “Reduce IT Costs” in the above list was identified as important with a high level of agreement.

36. The observation is that there is a high level of agreement across the industry regarding current and relevant goals. These results can be restated as meaning that they are shared industry goals. These shared industry goals provide the basis for achieving the identified desire to increase collaboration which is highly relevant in the context of open source software adoption as open source software by definition is software developed in a public collaborative manner ((Marsan, 2012), (Techopedia, 2012), (Service UGP, 2011), (del Bianco, 2011)).

37. One of the research objectives is to identify where OSS is being used in statistical organisations operations. The two areas of their operations in which OSS has been adopted at the highest rate and which were identified as having the highest level of agreement amongst organisations are in dissemination/data visualisation and in software component libraries. These two areas are closely followed by Application infrastructure, Software development tools, non-mission critical statistical applications, server OS, and in statistical operations (processing/analysis). The areas where OSS is being used the least is in Statistical Operations (archiving) and in Mission-critical non-statistical applications. It is very interesting to note however that for every possible business area which was interrogated, a minimum of 7 respondents indicated that they were either piloting OSS or using OSS in production. It is clear that OSS adoption has been embraced by statistical organisations and it is being adopted across an extensive range of business areas.
38. There was a second question asking where organisations saw themselves in 24 months regarding where they would be using OSS in their operations. The four areas in which OSS was used the most as well as the four areas in which OSS was used the least mostly stayed the same. There was much less agreement in the responses to this question as opposed to the responses regarding the situation that exists today. Two items of significance regarding the future plans for OSS in organisations is the intent to increase OSS usage in every area of operations in the next 24 months. The most significant increases in OSS adoption amongst statistical agencies in the next 24 months are in Statistical operations (process/analysis), Statistical operations (Collection), Statistical operations (metadata management), and Statistical operations (quality management). This supports the claim that once an organization has transitioned from considering OSS to adopting OSS and then to adopting OSS for a large number of business capabilities, OSS becomes an integral part of the business (Forrester, 2007).

39. At a technical level, the most important role for OSS in the next 24 months which organisations had a high level of agreement was related to the implementation of web services.

40. Analysing these data discovered that there is a relationship of moderate statistical significance between the position of an organisation on the model of innovation adoption process and the current state of adoption for applications. This relationship has less practical significance than the previous one as the correlation coefficient of 0.17 indicates that approximately only 17% of the variation in one variable is associated with the other.

41. This analysis also discovered that there is a relationship of moderate statistical significance between the position of an organisation on the model of innovation adoption process and future plans to use OSS in the next 24 months. The correlation coefficient of 0.23 indicates that approximately 23% of the variation in one variable is associated with the other which is also of practical significance.

42. The existence of and similarity between these two relationships increases the confidence in the quality and accuracy of the findings. The relationships support the use of the network model of diffusion of innovation as a process when considering the adoption of OSS in an organisation ((Hof RD, 2005), (Moore, 2002), (Trott, 2005)).

I. OSS Concerns

43. The study provided a list of concerns which was prepared from the current thinking in the area of OSS adoption. The list was extensive and consisted of fifteen items. At least 66% of the respondents identified “Availability of service and support”, “Security of the software”, “Lack of relevant internal skills and knowledge in OSS development or co-development”, “Lack of relevant internal skills and knowledge in OSS operations”, “Product immaturity”, “Overall complexity and difficulty of adoption”, and “Ability of OSS to meet business goals” as primary concerns of relevance to their organisation at the current time. In addition, at least half of the respondents found “Viability of the open source communities”, “Legal issues involving intellectual property”, “Total cost of ownership”, and “Licensing models / intellectual property” as concerns. Even the final remaining items “Lack of applications”, “Lack of information on software adoption rates”, “Absence of collaboration environments for software”, and “Organizational policy regarding OSS use” were identified as a concern for over 33% of respondents.

44. Of all of the items on this list, the first two “Availability of Service and Support” and “Security” have been consistently rated as the top two concerns in all OSS studies, regardless of industry ((Hammond, 2009), (Qu, 2011)). Security was explored in the review of current thinking and the findings were that security is not an issue at all and the security of OSS software is comparable to the security of commercial software ((Forrester, 2007), (Techopedia, 2012)).
45. The study found that there was a small but definite relationship between the concerns and the position of the organisation in the innovation adoption process. This supports the idea that as an organization transitions from considering OSS to adopting OSS and then to adopting OSS for a large number of business capabilities, OSS becomes an integral part of the business (Forrester, 2007) and is influenced by change (Oram, 2011) and experience of employees with OSS (Marsan, 2012).

46. As has been clearly highlighted in earlier paragraphs, OSS adoption is healthy and well in the statistics industry. This supports findings that the public sector is one of the highest adopters of OSS ((Hammond, 2009), (Qu, 2011), (Miralles, 2006)). OSS adoption initiatives need to be aware of the concerns raised in this section of the report and manage the risks appropriately. Organisations wishing to transition along the innovation adoption process or other organisations wishing to enter into collaborative agreements or increase adoption of OSS software should also be aware of these items and manage the relevant risks as a part of the process. Evaluating themselves and their potential partners with respect to these items may help to increase the chance of success.

J. OSS Service and Support

47. Of all of the options presented to respondents regarding OSS service and support (Hammond, 2009), there is one constant, the lack of demand in the statistics industry for service providers to provide legal support for open source software. There is industry demand for all of the other OSS support and services which were tested for in the questionnaire. There were five services however that were in demand from over 80% of respondents. The high demand services are “Life cycle support”, “Integrating multiple open source software components”, “Consulting”, “Training”, and “Integrating open source and closed source software”.

48. There was no relationship found between the position of an organisation along the innovation adoption process and service & support.

K. OSS Good Practices

49. It would appear that the statistics industry loves good practices. The entire list of good practices which was prepared from the current thinking was rated by statistical organisations as somewhat important with one exception. The standout exception was “Minimising OSS development” which was rated as the least important and had the least agreement. The need to “Manage maintenance and support costs” was unanimously rated as being the most important. The list of good practices (Forrester, 2007) is repeated here for convenience:

- Managing maintenance and support costs
- Having rules governing OSS development, maintenance, security, and support
- Having guidelines and approval processes for choosing OSS products prior to adoption
- Rating and ranking OSS risks and mitigating those risks
- Monitoring the OSS community to ensure adequate development and support of components used by the organisation
- Ensuring relevant internal skills and knowledge in OSS development or operations
- Establishing a support model for the OSS product/solution
- Establishing internal development, testing, and change management processes for OSS
- Establishing guidelines and approval processes for changes to OSS components
- Establishing OSS licensing standards and perform licensing due diligence for the organization
- Having an open source strategy
- Deliver internal training regarding open source software and technologies
Minimising OSS development

50. It is possible, but considered unlikely, that all of the organisations implementing OSS have formal strategies, policies, guidelines, standards, and so forth in place for all of these items. It is not possible for this study to determine why all of these items are rated highly when they (presumably) are not being implemented by organisations. As a recommendation, initiatives to increase the adoption of OSS in the statistics industry should consider ways to assist organisations in implementing these practices at minimal effort (i.e. introduction of templates and guidelines, sharing of practices). This effort might increase the likelihood that these practices are implemented and industry needs met.

IV. Conclusion

51. The statistics industry has been proven to be a significant adopter of open source software. The shared industry goals which have been highlighted provide the perfect environment for increasing collaboration and doing more with less. There have been no significant risks or barriers identified which would prevent improvement initiatives from being launched. This study has confirmed that some of the positions put forward in the current thinking from other industries are applicable as well to the statistics industry. The study has also confirmed that the statistics industry is different and that current thinking does not always apply.

52. The study found that most statistical organisations are already using open source software. In fact, open source software is being used in all areas of statistical organisation operations with the most common areas being data dissemination/visualisation and software component libraries. After organisations have implemented open source software in production, plans to increase adoption and extend its use into other areas of the business have become a part of the subsequent planning cycle. The statistics industry was found to be consistent with other industries in identifying the top concerns with open source software, security and the availability of service and support.

53. The adoption of open source software differs from adopting commercial software and is influenced by a range of factors, both internal and external to the organisation. Successful implementation involves changes in areas such as policy, governance, and staff skills/training.

54. Organisations indicated that many best practices related to policies, governance, open source strategy, training and support were important but not being implemented. It is recommended that future strategies to increase adoption of open source software include the adoption of best practices.

55. Total cost of ownership was identified as being lower with open source software. One area of caution is that statistical organisations indicated open source software adoption will reduce operating costs. This is in contrast with current research which claims that operating costs are the same as or higher than commercial software.

56. There is a positive and receptive environment amongst statistical organisations towards open source software adoption and use as well as a high level of agreement across the industry regarding business goals. This context will support efforts to increase open source software adoption and usage (as a means of realising shared goals).
V. References

44. TECHOPEDIA. 2012. Open Source: Too Good to Be True?