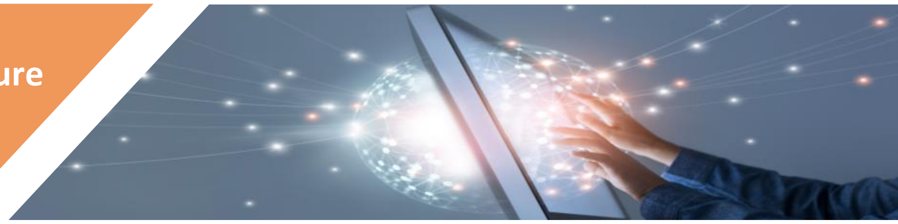
Case Studies Series

Using Nesta's Data Maturity Model to Measure Data Maturity of an Organization



Abstract

It is widely recognized that organizations should start their own data journey to leverage and take advantage of its available data. However, most of the organizations do not know how to go about it since there are many factors which affect the success of fully becoming data driven organizations. Despite their efforts, many organizations are still unable to become fully data driven. In the field of management, maturity is the factor which could improve the organization in a particular aspect. This case study looked at the possibility of using Nesta's Data maturity module to understand the maturity level of data at the Royal Institute of Management (RIM) in Bhutan. The model uses a 7 item likert scale survey questionnaire as an instrument, which includes 34 questions pertaining to different data characteristics of which each maturity level needs to be elicited. The results show that RIM scored very less in the survey with almost all data characteristics having an average score of less than 3. The score when translated to the level of maturity, using the model, shows that almost all the data characteristics were at "Nascent" level, two at "intermediate" and one at "basic level".

1. Introduction

Leveraging the use of data to optimize the different processes and systems within an organization is becoming more and more popular, especially amongst some pioneering companies such as Facebook, Google, and Amazon, with the help of technology. Technologies like self-driving cars, automated homes, augmented reality, etc. are changing the way we perceive the world. Moreover, many claim that data to be the new fuel in today's world due to its use as a resource to fuel the digital economy.[1] The success of companies like Google[2] and Facebook[3] which trade data in the form of different products and services reinforced this claim.

While these pioneering companies are creating a lot of disruptions in the market with the use of data, the majority of organizations around the world are still struggling to leverage the use of data at its basic level. Today almost everyone is talking about the importance of "data-based-decision-making" and how it has become an imperative part of business when it comes to staying relevant to technology-based data economy.[4] This has contributed to the hype that organizations would like to join the data revolution and optimize the use of data, however, they do not know how to begin the data journey.

---

[1] Mackey, S. (2019, 01 31). *Data is the new oil: Data standardization fuels digital transformation*. Retrieved from ADLIB: https://www.adlibsoftware.com/blog/2019/January/Data-is-the-New-Oil-Data-Standardization-Fuels-Digital-Transformation.aspx

[2] *How Google Innovates*. (2018, 11 06). Retrieved from Medium: https://medium.com/swlh/how-google-innovates-cf1719cda6af

[3] Monnappa, A. (2019, 07 02). *How Facebook is Using Big Data - The Good, the Bad, and the Ugly*. Retrieved from Simplelearn: https://www.simplilearn.com/how-facebook-is-using-big-data-article

[4] Rai, A. (2018, 07 04). *How is Government Leveraging Data?* Retrieved from Upgrad: https://www.upgrad.com/blog/how-is-government-leveraging-data/

According to Carretero et al., an organization needs to feed their various processes with existing data and also needs to find various new data sources that feeds into the current system. To be more competitive, one needs to find newer ways to capture, store, analyze and derive insights that would help in decision-making process within an organization. In addition to data collection and storage, and infrastructure analysis, an organization should invest sufficient resources to optimize the quality of data captured and stored to ensure knowledge derivation from the data.[5]

Highlighting the downside of moving towards a data-driven organization, a recent survey[6] that was conducted by the *NewVantage Partners* has chosen 64 c-Level Technology and Executive Representatives from different companies like Ford Motors, General Electrics, and Johnson & Johnson and found that 72% of the respondents claim that they have yet to forge the data-culture and 69% think they have not yet created a data-driven company. Another 53% think that they have not yet begun to take data as an asset and 52% admit that they are not competent in data and analytics. This result highlights the failure to move towards a data-driven organization despite their increase in investment in Big-Data and AI.[7] Hence, as one can see, moving towards becoming data-driven is not just about pouring a huge amount of investment, but there are also a lot of other factors which needs to be identified first.

So how does an organization, which does not have a capacity in terms of investment compared to corporate giants, go about moving towards becoming data-driven and avoid the pitfalls other organizations have already fallen into? To this end, there is a self-assessment framework that can help an organization define the different aspects of data in a clear and concise way. It is a tool which enables an organization to measure the current level of data-use in their institution's systems and processes, and accordingly plan a process which enables them to explore different measures to manage a move towards being a data-based organization by mitigating various risks associated with such projects.

In the Management field, maturity is a measurement of the ability of an organization for continuous improvement in a particular discipline.[8] Most maturity models measure people/culture, processes/structures, and objects/technology.[9] There are two basic ways of implementing maturity models: i) Top-Down as proposed by Becker et al. where the levels of maturity are fixed and further defined by characteristics which support the evolution from one maturity level to another[10]; and ii) Bottom-up approach proposed by Lahrmann et al. where the characteristics are determined first and clustered in different maturity levels later.[11] There are many Maturity models used for assessing data maturity of an organization. According to DataKind there are about 40-50 frameworks and theories in use, which are developed for various different purposes and in various different contexts like IBM's Big Data Framework, ODI's Open data framework and DataKind's charity sector frameworks to name a few.[12] Data

[5] Carretero, A. G., Gualo, F., Caballero, I., & Piattini, M. (2017). MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000. *Computer Standards & Interfaces, 54*, 131-151. Retrieved September 18, 2019, from https://www.sciencedirect.com/science/article/abs/pii/S0920548916301878?via%3Dihub

[6] New Vantage Partners. (2019). *Data and Innovation: How Big Data and AI are Accelerating Business Transformation*. Retrieved from Big Data and AI Executive Survey 2019: Executive Summary of Findings: http://newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf

[7] Carretero, A. G., Gualo, F., Caballero, I., & Piattini, M. (2017). MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000. *Computer Standards & Interfaces, 54*, 131-151. Retrieved September 18, 2019, from https://www.sciencedirect.com/science/article/abs/pii/S0920548916301878?via%3Dihub

[8] Almuhammadi, S. and Alsaleh, M. (2017). Information security maturity model for nist cyber security framework. *Computer Science & Information Technology*, 51-62. Retrieved from https://airccj.org/CSCP/vol7/csit76505.pdf

[9] Mettler, T. (2011). Maturity assessment models: a design science research approach. *International Journal of Society Systems Science, 3*(1/2), 213-222. doi:10.1504/IJSSS.2011.038934

[10] Becker, J., Knackstedt, R. & Pöppelbuß, J. (2009). Developing maturity models for IT management – a procedure model and its application. *Business & Information Systems Engineering, 1*(3), 213-222.

[11] Lahrmann, G., Marx, F., Mettler, T., Winter, R., Wortmann, F. (2011). Inductive design of maturity models: Applying the rasch algorithm for design science research. In Jain, H., Sinha, A.P. & Vitharana, P. (Eds.), *Service-Oriented Perspectives in Design Science Research, 6629*, 176-191. *DESRIST* 2011. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer.

[12] https://www.dataorchard.org.uk/resources/review-of-data-maturity-models

Maturity according to DataKind could be referred to readiness of an organization to take on data work of different levels of complexity to understand how value is created through the use of data (i.e. to understand the past and present and prepare for the future).

The Nesta's Data Maturity Model (to be explained in detail in section 2), created by LGA, Nesta and Porism in the UK, is used in this case study. The Nesta's model is selected because of its simplicity and conciseness compared to other models. The Nesta's module is free to use in its pilot version and it provides a self-assessment tool which has been automated. Moreover, it constitutes different factors that need to be considered for moving towards being a data-driven organization in a clear and concise way (see section 2) by responding to a series of simple questions. The tool is developed to offer a summary of its assessment and sub-divide scores by different types of survey respondents.

1.1 Objective of the Study

1) To provide an example on how an organization can use Nesta's Data Maturity Model to elicit the data maturity of their state.

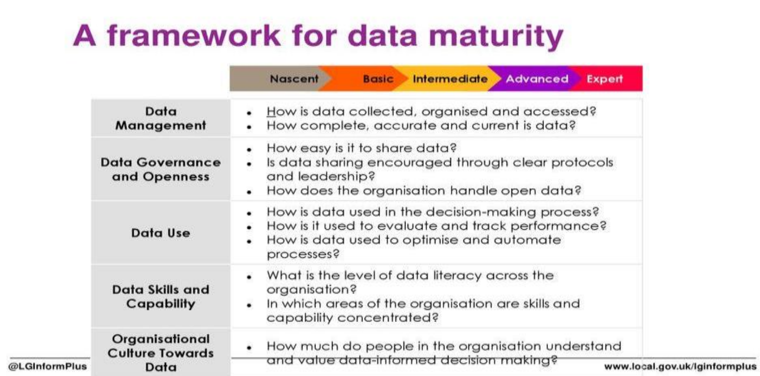
2) To compile the results generated using the model and discuss its usage in Bhutan's context.

The study is structured as follows: Section 2 discusses Nesta's Data Maturity Model in detail, followed by Section 3 which includes case-study design in terms of process, methodology and instruments used. Section 4 provides details of the analysis and results. Section 5 provides a discussion on the use of the model in Bhutan's context. Section 6 provides a conclusion and future studies.

2. Nesta's Data Maturity Model

As a self-assessment automated tool for local governments, Nesta's data maturity model is jointly-developed by LGA, Nesta and Porism in the UK[13] [14] on their own initiative based on work done in 2016. The tool is available as a free tool in its pilot phase. It uses the following framework to determine the maturity levels and characteristics as shown in Figure 1.

Figure 1: Data Maturity Framework



***Source**: www.local.gov.uk*

[13] *Data Maturity self-assessment tool for local government*. (2018). Retrieved from Local Governmnet Associaton: https://about.esd.org.uk/news/data-maturity-self-assessment-tool-local-government

[14] *LGA creates online data maturity assessment tool*. (2018). Retrieved from UKAuthority: https://www.ukauthority.com/articles/lga-creates-online-data-maturity-assessment-tool/

There are two dimensions to Nesta's data maturity framework like any other maturity models as mentioned by Becker et al. and lahrmann et al..[15] [16]

i) Maturity levels

It measures data state of an organization from the Nascent level to Expert level (also called “datavore” level) which is defined by the extent to which the data characteristics are met by the organization. The details of the levels within each data characteristic are given in Figure 2.

ii) Data characteristics

Nesta’s model looks at 5 data characteristics - Data Management, Data Governance and Openness, Data Use, Data Skills and Capability and Data Culture of the organization. These characteristics are further categorized into some sub-characteristics like data collection, organization, quality, governance, openness, decision making and performance evaluation under Data Management, Data Governance and Openness and Data Use. The data maturity of an organization is elicited through the use of a survey questionnaire that is deployed through an online automated tool. For each data characteristic and sub-characteristic, the level of maturity from nascent to datavore is assigned based on the details of how an organization is doing within each data characteristic as given in Appendix 1.

3. Case Study: Using Nesta’s Data Maturity Model at Royal Institute Management (Bhutan) to measure its Data Maturity Level

3.1 Case Study Design

In this section, details of how the case study was designed is presented to illustrate the way the level of maturity of an organization is checked through Nesta’s Data Maturity Model in contention. The following subsections provides details of the study process used, description of the organization and the sample for the study, the data instruments used and the tools used to analyze the data collected.

3.2 Process used

To achieve the objective of the case study, the process is divided into two steps:
Step 1: Deploy the questionnaires to the survey participants and analyze survey data
Step 2: Call a meeting to discuss the analysis, record the results and assign maturity to each data characteristic using Figure 2

3.3 Description of the Organization and Sample Details

RIM is a premier management institute in Bhutan mandated to train public servants in the field of public administration, finance, law and I.T.. RIM is also mandated to conduct research and acts as a back-stopping institute for the Royal Government of Bhutan (RGoB) in framing policies in various disciplines. To achieve their mandate, use of various data is required to help with the day-to-day decision-making. The institute has a total of 73 employees of

---

[15] Becker, J., Knackstedt, R. & Pöppelbuß, J. (2009). Developing maturity models for IT management – a procedure model and its application. *Business & Information Systems Engineering, 1*(3), 213-222.

[16] Lahrmann, G., Marx, F., Mettler, T., Winter, R., Wortmann, F. (2011). Inductive design of maturity models: Applying the rasch algorithm for design science research. In Jain, H., Sinha, A.P. & Vitharana, P. (Eds.), *Service-Oriented Perspectives in Design Science Research, 6629*, 176-191. *DESRIST* 2011. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer.

which only about 45 work at above supervisory level. A total of 32 staff and faculty members responded to the survey questionnaires.

3.4 Instrument, Methodology and Tool Used

To collect the data to measure the current state of data maturity at RIM, a likert scale survey questionnaire made of 34 questions/statements as defined in the Local Government online-self assessment tool[17] was deployed to the sample. They were deployed using Google forms to collect their responses on various data characteristics. Each likert question/statement provided 7 choices ranging from **Strongly Disagree, Disagree, Neutral, Don't Know, Don't Understand, Agree and Strongly Disagree in order** to elicit the maturity level in each item that represents different data characteristics. The questionnaire deployed can be viewed from

https://drive.google.com/file/d/10CmEJFxThIdoD4yw8xhEV2TW3NiHGYjc/view?usp=sharing

Quantitative method is employed to aggregate the likert responses and map them in terms of the characteristics that define different levels of maturity. Analyses are compiled using mean and mode as the central tendency and appropriate visualization of the responses is created to elicit the level using the maturity model given in Figure 2.

Tableau is used as a tool to analyze the data and create visualizations to present the findings. It was performed on a 32 Bit Windows Operating System with 4 GB RAM.

4. Analysis and Results

This section gives the details of the Data Cleaning and Transformation, Analysis and the results that are obtained.

4.1 Data Cleaning and Transformation

The data that was collected using the Google-form went through following steps to make it appropriate for analysis.

Step 1: Extracting the data collected
The data was collected using the Google form in the format as shown below in Table 1.

Table 1: Format of survey data that was collected using Google-form

ResponseID	Gender	RCSC Position level	Q1	Q2	……..	Q34
Row1 (Response 1)						
Row2 (Response 2)						
Row n (Response n)						

[17] *Data Maturity: Rate your organisation's data management skills*. (n.d.). Retrieved from Local Government Association: https://datamaturity.esd.org.uk

Step 2: Transforming the data table to make it apt for Tableau software

The format shown in Table 1 is appropriate for human but an intuitive approach is not adopted by Tableau software**.** So, the columns of the questionsID (i.e. Q1, Q2, etc) were pivoted to rows. As a result, the columns were merged and renamed "QuestionID".

Step 3: Merging Data Table (Table 1) with Meta-Data Table (Table 2)

An additional Meta-Data table as shown below in Table 2 was also created to map the questions (string form) for the QuestionID in order to make the analysis easier. Additionally, each question was labeled with its respective grouping in two levels of abstraction which represents the "Data Characteristics" as given in Nesta's Maturity Model.

Table 2: Meta-Data table created to map questions concerning the Data Characteristic

QuestionID	Question Wording	Level 1 Grouping	Level 2 Grouping

Table 1 was merged with Table 2 using the "QuestionID" as the key

Step 4: Coding likert responses to likert score between 1-5

To also have the option of looking at the survey responses as continuous values, the following coding schemes representing scores were used.

0 = Don't know, 0 = Don't Understand, 1 = Strongly Disagree, 2 = Disagree,
3 = Neutral, 4 = Agree, 5 = Strongly Agree, Yes = 5, No = 1

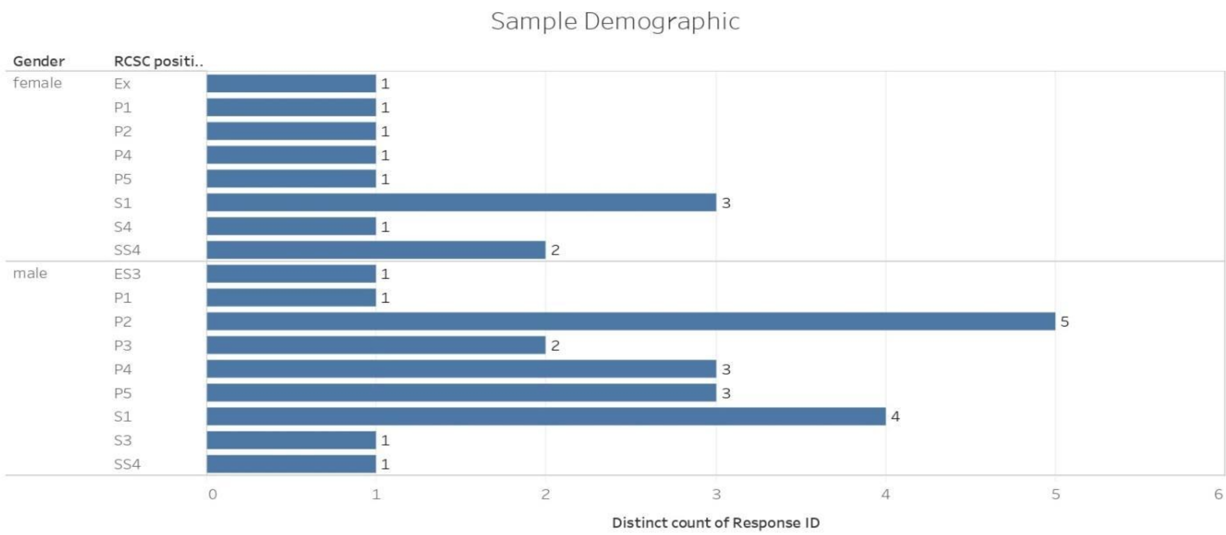
Step 5: The data which was ready for analysis was finally cleaned and transformed for analysis. The following format in Table 3 represents the cleaned and transformed data format.

Table 3: Transformed and cleaned Data format

Response ID	Gender	RCSC Level	Question ID	Response	Score	Question Wording	Level 1 Grouping	Level 2 Grouping

4.2 Analysis of Demographics

Figure 3: Analysis of Sample Demographics

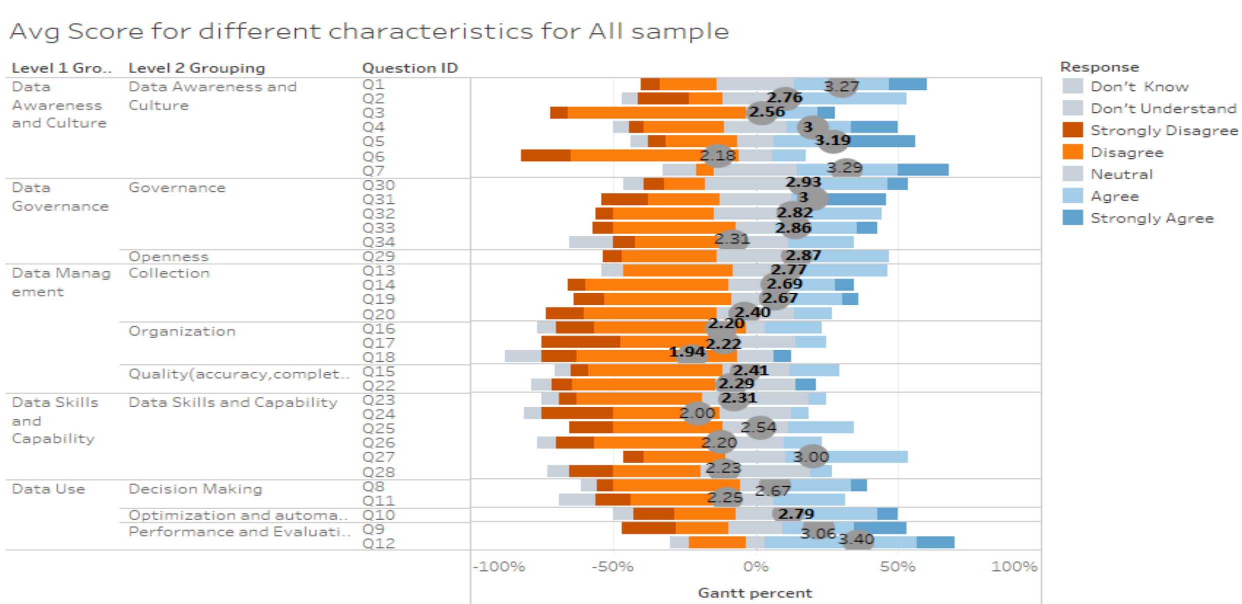


The analysis of the demographics of the sample shows that out of 32 respondents, 21 are male and 11 are female. 19 respondents are at above the P5 (RCSC civil servants rank) while 13 are at below P5. This shows that 59% of the total sample are at Operational and Strategic professional level. So, it is a good representation of decision makers within the Institute. The names of the respondents have been omitted to maintain the privacy of the respondents.

4.3 Analysis of Data Characteristics Using Diverging Stacked Bar Chart

Figure 4: Diverging stacked bar chart representing aggregated average scores

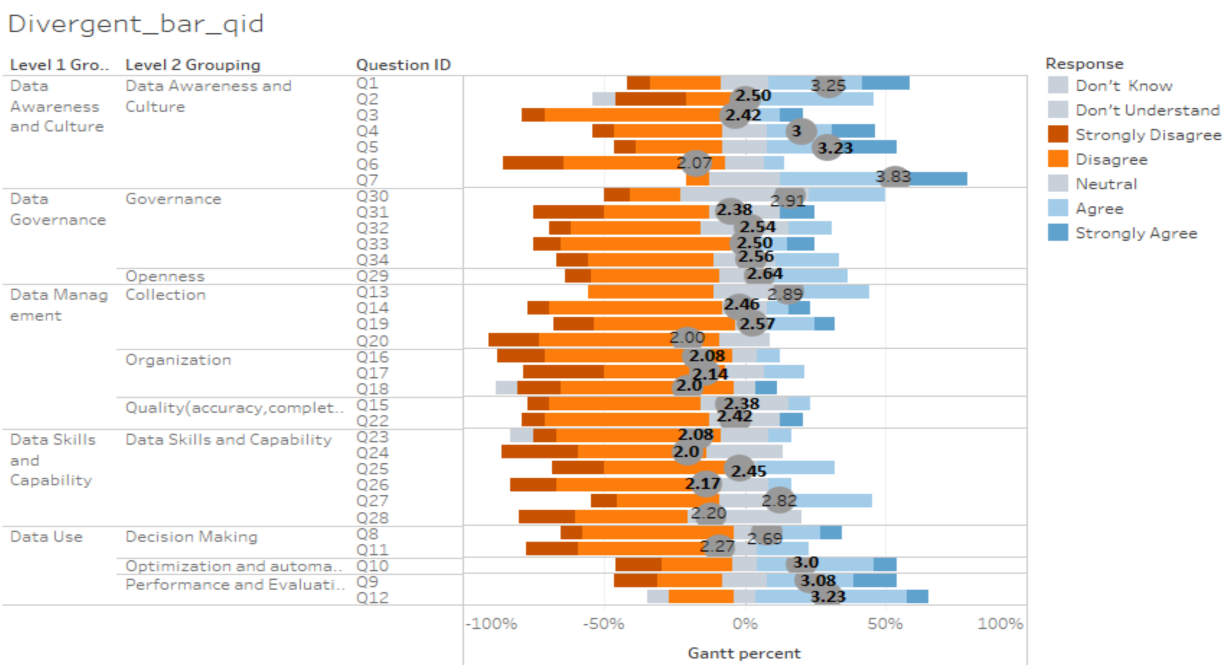
for each data characteristic for all samples



As shown in Figure 4, the use of the diverging stacked bar chart aggregating the scores of the entire sample shows that the majority of respondents respond with “Disagree”, which is represented in orange. We can see that the majority respond with “neutral” in terms of different characteristics represented in grey. Overall, an indication of “Neutral”, “Disagree”, “Strongly Disagree”, “Don’t know” and “Don’t Understand” are seen from the analysis of the entire sample. The average score represented by the number in each characteristic also represents the indication towards the negative results. The average scores for the different data characteristics range from 1.94 to 3.40. This represents that the scores are below par for almost all the data characteristics considering we code the score of 3 for the “neutral” response. There is enough evidence showing that the mean of each data characteristic is being skewed by the neutral value of 3. To further investigate the mean scores, the sample was divided into two groups using the RCSC position level. An “RCSC position level” represents the rank of an employee as categorized by the Royal Civil Service Commission of Bhutan. The position level divides the employees into groups such as “Ex” for Executives, “ES” for Specialist, “P” for Professional and “S” for Supervisor. Using the RCSC position level would help us look at the scores more objectively. Thus, we analyze the response based on two groups we create using RCSC position levels namely P5 - EX/ES and SS4-S1. The motivation to use it is to see how employees at different levels of their responsibilities have responded to the survey.

4.3.1 Analysis of Average Score of Data Characteristics of P5 to EX/ES Level

Figure 5: Score for data characteristic for employees between P5 - EX level

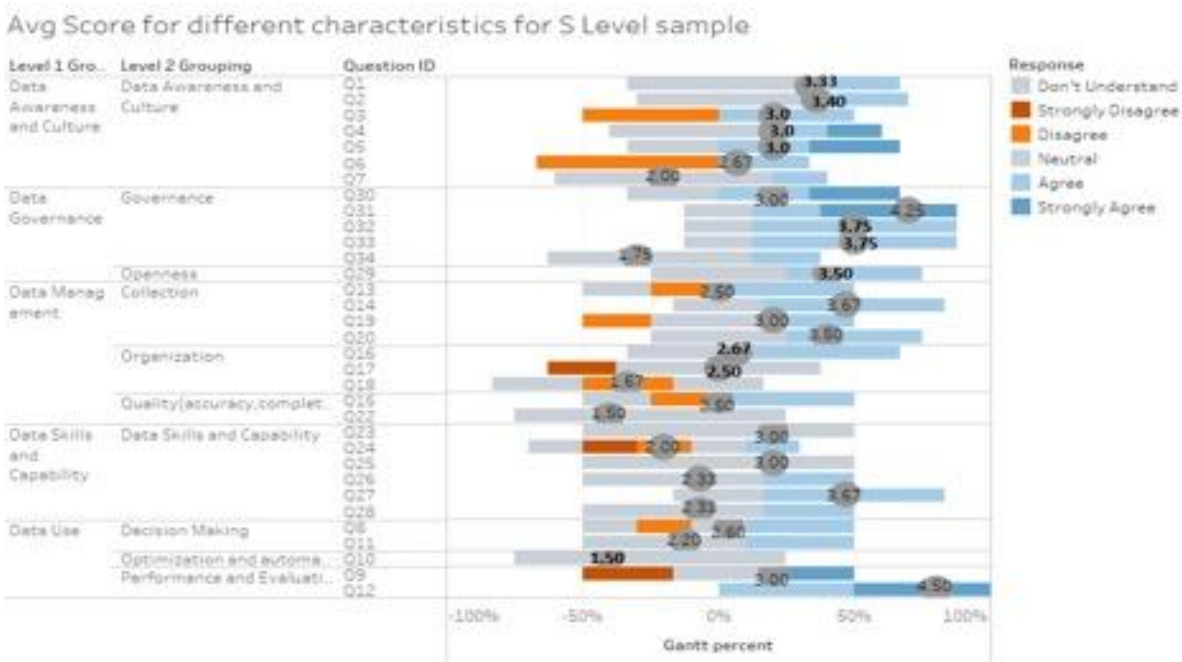


As given in **Figure 5**, the mean score does not change much when compared to the score of the entire sample. It just shows a little more inclination towards “Disagree” and “Strongly Disagree” in most of the data characteristics. However, as compared to the Figure 4, we can see that the “Don’t Know” and “Don’t Understand” responses have been reduced and the “Agree” and “Strongly Agree” have increased due to the fact that the average score of many data characteristics have been scored above 3.

4.3.2 Analysis of Average Score of Data characteristics of S-Level Employee

The Figure 6 shows the average score for different data characteristics of S-Level employees and we see why the mean for the entire population was skewed (Figure 3). As you can see from the figure, the majority of S-level employee have either responded with “neutral”, “Don't Know”, “Don’t Understand” or “Agree” or “Strongly Agree” which contradicts what P5-Ex level respond (Figure 5). The use of these responses means that the average score for almost all the data characteristics is above 3. Thus, to come to the results using these analyses, a further intimation is required with the employees to understand why the response between two groups of employees of the same organization is very different from each other.

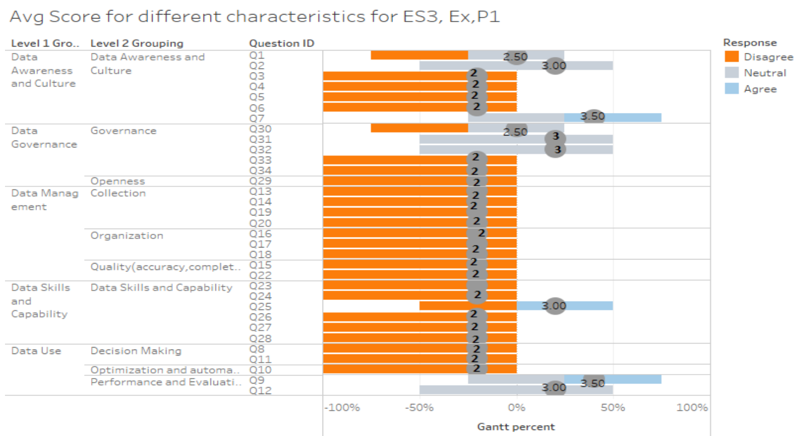
Figure 6: Score for different characteristics by employees of S-Level



4.3.3 Analysis of Average Score of Data Characteristics of RIM Management

At this point, it is important to consider a group which makes up the management of RIM. The employees at the level of P2, P1, EX and ES usually are the group who are at the managerial and executive level and the ones who make most of the decisions within the Institute. Figure7 shows that most of the employees who are within the management group have responded with “Disagree” to almost all the data characteristics with an average score of 2 which is inclined towards a negative score. They have also responded with “agree” to questionID Q7, Q25 and Q9 with an average score of 3.25 which is an indication towards a positive score.

Figure 7: Data characteristic score for Management level



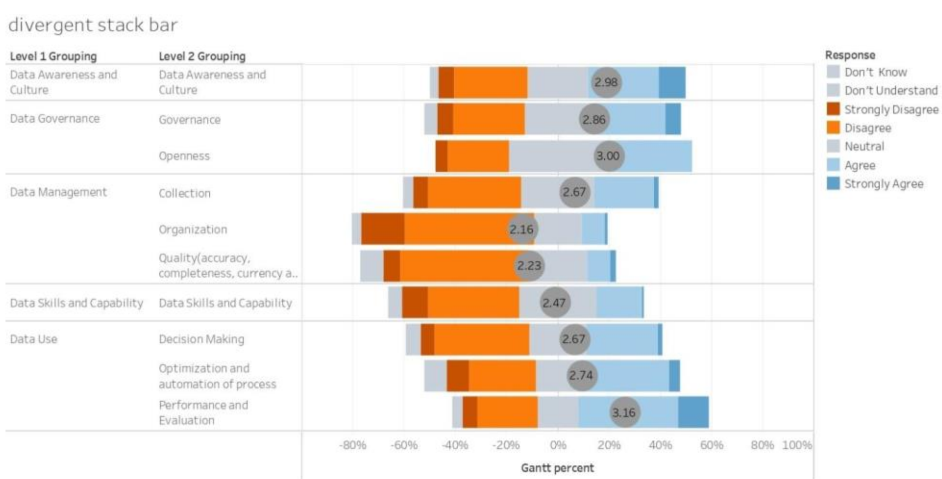
4.5 Results

This section gives the details of how the different analysis was compiled and how the maturity levels were assigned for RIM within each data characteristic.

4.5.1 Aggregation of Analyses

The analyses that were conducted in the previous sections using the diverging stacked bar chart showed that the responses in the form of average and mode (in proportion) differed between different groups (i.e. Management and P5-EX/ES, and S-level) and also contradicted with each other when viewed based on individual data characteristics. Thus, a validation of the scores needs to be performed through intimation with the respondents regarding their responses and the intention of the questionnaires by organizing a meeting to assign maturity levels. Nevertheless, it cannot be denied that the skew created by different responses is also a result of itself that shows that the knowledge about different data characteristics is very limited, thus resulting in contradictions. However, an overall mean score in terms of each grouping of data characteristics can be seen in Figure 8.

Figure 8: Scores for all the sample at each Data Characteristics grouping level



As seen in Figure 8, RIM has a score of **2.98** in **Data Awareness and Culture**, followed by **2.86** and **3.0** in **Data Governance** and **Openness respectively** which make up the Data Governance Characteristics. Within the Data Management characteristics, RIM scored **2.67** in **Data Collection**, **2.16** in **Data Organization** and **2.23** in **Data Quality (Accuracy, Completeness and Currency)**. To continue, **Data Skills and Capability** garnered a score of **2.47**. Regarding the Data Use, **Decision Making** has the score of **2.67**, **Optimization and Automation** received **2.74** and finally **Performance and Evaluation** has the score of **3.16**. Overall, only two data characteristics received the score of above or equal to 3. The scores for the rest of the data characteristics are all below 3, which give negative indication that the level of maturity is very low.

4.5.2 Data Maturity Level for Each Data Characteristic at RIM

To assign data maturity level of RIM to different data characteristics, the scores which have been obtained from the different analyses are considered, along with the definition of each maturity level as shown in **Figure 2** in Section 2.

A one-on-one intimation was scheduled with the respondents of the survey at RIM. 25 members out of a total of 32 respondents were met. The following data maturity levels and different data characteristic groupings are given as below in Table 4.

Table 4: Mapping of Data Characteristics with Maturity level for RIM

Data Characteristics Level 1	Data Characteristics Level 2	Nascent	Basic	Intermediate	Advance	Datavore
Data Use	Decision making			✔		
	Optimization and Automation Process	✔				
	Performance and Evaluation			✔		
Data Management	Collection	✓				
	Organization	✔				
	Quality (Accuracy, Currency, Completeness)		✔			
Data Capability and Skills			✔			
Data Governance	Governance	✓				
	Openness	✔				
Data Awareness and Culture			✔			

a) Data Use

In terms of the Data Use at RIM, use of data in decision making scored 2.67 which shows the respondents think that data is not used for aiding decision-making at RIM, since RIM Data analysis is requested for decision making but due to the unavailability of right data to answer the right questions and the analysis is of low quality to base a decision on. Additionally, data is the part of the decision-making process and used in most published reports. Using the model given in Figure 2, Data Use for **Decision Making** can be elicited to be at **Intermediate level** at RIM**.**

Regarding the use of data for performance and evaluation, the score given by the respondents is 3.16 which shows that most respondents believe that the performance and evaluation are usually conducted using data. Since data is used to manage the staff and services in terms of their performance but in an ad-hoc manner, the performance and evaluation can be given **intermediate level** at RIM.

Optimization and Automation of Processes scored 2.74 and since no system/processes at RIM has been automated or improved using data, it can be assigned to **Nascent Level.**

b) Data Management

Data Collection scored 2.67 from the survey and since there is evidence showing that data is collected as the by-product of operational and service delivery, and driven largely by the government requirements and KPI's, collection is still at **Nascent level** at RIM.

Organization of data at RIM scored 2.16 which gives the picture that data is not organized the way it should be. Mapping with the model in Figure 2, since data is maintained in silos and there is limited sharing between the departments, the organization is still at **Nascent level**.

Quality of data in terms of Accuracy, Completeness and Consistency scored 2.23 from the survey, and during the evaluation it was seen that the data was not checked but it could have been cleaned for basic analysis. Thus, the level assigned to it is **basic**.

c) Data Capability and Skills

Data Capability and Skills scored 2.47 from the survey, and it was elicited from the meeting in which the skills of IT system Managers and basic software like Excel are limited. As seen from the survey, most staff lack basic data literacy and skills. Hence, the Capability and Skills are assigned to the **Nascent level**.

d) Data Governance

Data Governance scored 2.86 which is higher than most of the other data characteristics. With careful analysis, most S-level employees responded with "Agree" and "Strongly Agree" (seen in Figure 6) while the P5-EX/ES level responded with "Disagree" and "Neutral" (seen in Figure 7). After the evaluation during the meeting, it was found that there was neither data policy nor the sharing policy. The department collects data in silo as the by-product of service provided and they are protected. Hence, despite scoring higher than the other data characteristics, **it was assigned to Nascent Level**.

In addition, Data Openness scored 3 from the survey but it was seen that the score was skewed with too many neutral and random responses from the survey. Since RIM does not make data available to the public nor maintain any information on how it uses data, it is assigned to **Nascent level**.

e) Data Awareness and Culture

Data awareness and culture scored 2.98 in the survey but from the analysis many responded with “neutral” which skewed the score towards the higher side. However, as seen from the meeting, the staff had very little awareness on how data could be used to improve services and outcomes. It therefore receives **Nascent Level**.

5. Discussion on the use of Nesta’s Maturity Model for Data Maturity Elicitation

Deploying Nesta’s standard questionnaire and analysis of the results shows that employee subgroup based on their RCSC level answered very differently from the questionnaire. In fact, in some cases it clearly contradicted with each other. The following were common answers given by the participants after they were asked about their experience of taking the questionnaire survey:

i) the questions were very hard to understand and included many jargons which they did not understand.

ii) some said that they just went through the survey randomly.

Accordingly, to use Nesta’s model, it is recommended to train the employees and select the sample carefully in order to get accurate data for your own cases. There is a need to consider the choice of sample to keep it objective. If the sample is not carefully chosen, the random respondents could skew the results. As long as the sample is carefully chosen, the model can be used very easily and accurately to elicit the maturity level. However, one can frame one’s own questionnaire so that the mapping based on score and definition of each maturity level can be done more comprehensively. With the standard questionnaire, mapping is not very well understood when assigning maturity.

Additionally, the use of “Don’t Know” and “Don’t Understand” as an extra option in the likert scale needs to be handled carefully. If you are going to deal with the “mean” metric, one should be careful of what score you code for these two responses. In the case study, these two responses were coded with “0” to make sure that it does not skew the mean (average) value. The mean is still skewed by the neutral response. One needs to keep in mind how one can interpret the neutral values. Does it add to the positive or the negative side?

6. Conclusion

In conclusion, an attempt to apply Nesta’s Data Maturity Model was made by deploying the standard survey questionnaires created by the Local Government Agency UK through a case-study that was deployed at the Royal Institute of Management, Bhutan. Since we could not get permission to use the automated tool online, we tried to replicate the use of model using tableau as the software to analyze the data collected. The analysis of the data collected and the results of analysis were mapped with Nesta's Data Maturity Model which measures the maturity level of data in an organization by assigning maturity level: Nascent, Basic, Intermediate, Advanced and Datavore to different Data Characteristics like Data Management, Data Use, Data Governance, Data Skills and Capability and Data Governance and Openness. From the study at RIM, we found out that leveraging data at RIM was at a very infant stage. We found that RIM was at the “Nascent Level” within Optimization and Automation of Process, Data Collection, Data Organization, Data Governance, Openness of Data, Data Awareness and Culture and Data Capability and Skills. Quality of Data at RIM in terms of Accuracy, Completeness and Currency was found to be at “Basic Level” while use of data for decision-making and use of data for performance and evaluation purpose were found to be at “Intermediate Level”.

While conducting the analysis using the survey questionnaires, regarding the use of instrument for the survey, it was noted that the sample needed to be selected properly. In addition to that, since the mapping of the questionnaire results using Nesta's model (which defines each maturity level) is not explained clearly, there are some challenges in translating scores into maturity levels. Thus, a self-designed questionnaire is recommended to use with the model. Moreover, regarding the coding of the scores, if one is using the mean metric to elicit scores, being careful about the “neutral” response and its effect on the scores is recommended.

Overall, the use of the models has helped RIM in many ways. First of all, it helped an organization like RIM to understand the different dimensions of becoming data-driven in the form of different data characteristics. Secondly, RIM has a view on how they are doing in each data characteristic dimension in the form of their maturity model. Using this module, RIM can work towards improving in each of the data characteristic by planning on how to improve each data characteristic. Future work would consist of deploying different interventions to improve each of the data characteristic and then using Nesta’s model to check the maturity levels after the interventions.

Appendix 1: Table showing data characteristics and details of each maturity level

		Nascent	Basic	Intermediate	Advanced	Datavore
Data management	**Collection**	Data collection is a by-product of operational and service delivery, and driven by central government requirements and key performance indicators.	Collection goes beyond operational use and mandatory reporting requirements but there is little strategic purpose behind collection or use.	Data is used well in operational settings and data is sometimes collected for strategic purposes but predominantly there is little strategic rationale for collection and use.	Data is used well in operational settings and other data is collected in line with broader organisational strategies and decision-making.	Data is collected extensively across all services and in-line with organisational strategy. Data can provide a holistic view but data is not collected where the immediate use is not apparent (avoiding data exhaust). Data is seen as an organisational asset.
	Organisation	Data is organised in silos with limited ability to share across the council.	Some data can be more widely published or shared and integrated manually.	Lots of data is exported and shared across the council, but mostly it requires manual integration.	Most data can be shared and integrated, some of it automatically through data warehouses or federated approaches.	A data warehouse or federated data models are used so that data is owned diffusely but can be integrated easily/ automatically. There is an information asset list or inventory which is published as metadata.
	Quality (accuracy, completeness, currency and consistency)	Data quality is patchy but is not addressed.	Data quality is low but can be addressed on an ad hoc basis when basic analysis is undertaken.	Most data that is exported from IT systems is of useable standard but errors remain and are not addressed comprehens-ively. Data quality is maintained and improved by staff involved in line of business data collection.	Data is generally of useable quality, and most staff understand the need for accuracy in inputting data.	All data is of useable quality and data quality issues understood and managed by all staff proactively. All staff take responsibility for the quality of the data they collect.
Data Governance and Openness	**Governance**	Data protection is a major reason not to share data and undertake analysis.	Information governance concerns prohibit most sharing of data for analysis purposes. Assigned senior level data owners responsible for specific data sets and accountable for	Data sharing does occur but not extensively, and there is limited consistency in decisions made about sharing. The organisation has assigned senior level data owners for specific	There are some information sharing protocols and data can be shared internally and externally to undertake analysis.	Information governance protocols based on specific use-cases have been embedded in IT systems to enable responsible data sharing. A Corporate Management team member proactively drives information /

			agreeing new uses and access to data is done on an ad hoc basis.	data sets accountable for agreeing new uses and access to data.		data integration internally and with partners to secure new insights, joined up services and savings. Information sharing and data sharing decisions are based on a balanced risk assessment that weighs privacy concerns against the risk to the organisation or individual of not sharing.
	Openness	Data is not made available to the public in machine readable formats. No public message about how the council uses data.	Data is made available on an ad hoc basis. Data is in a mixture of machine readable and non-machine readable. Public message about data use is technical/ legal in nature.	There is an ambition to make more data available and some data sets are updated at regular frequencies, but is done mostly manually. Data is mainly in machine readable formats.	There is a single open data portal and most data is machine readable. Most data has a scheduled frequency for updating, and some of this is done automatically.	There is an open data portal with multiple data sets, open by default approach and a user-friendly interface which enables basic visualisation and analysis. All open data is machine readable in standard open format and use APIs where possible. At least some data has an ODI open data certificate. A clear public message about how and why data is used.
Data Use	**Decision-making**	Rich in data, poor in intelligence. Data is not a key part of decision-making processes.	Data is used in reports but usually in a cursory way and with little reference to decisions which have to be made.	Data analysis is usually requested for decision making, but can be inadequate because analysis is not of high quality, targeted at the decision to be made or the right data is not available.	Some decisions are informed by data on both the frontline and at senior levels, but it is not consistent across the organisation.	Rich in data intelligence and insight. Data is analysed on specifically for the purposes of key decisions which have to be made, consistently across the organisation. Data is available in a timely fashion to support decision-making.
	Performance and Evaluation	Services and performance are not evaluated using the data available.	Data is used to look retrospectively at performance, often in static format such as a spreadsheet. Data offers little insight into why events or performance variations occur.	Data is sometimes used to understand why events, or levels of performance, have occurred. Performance management using data is of limited value.	Data is sometimes sought to conduct evaluations of services and interventions, but mainly on an ad hoc basis. Data can be used to usefully performance manage staff and services, and there	Data is used to support service delivery in real-time, is used to understand in granular detail issues of performance, and can be used to understand the effectiveness of services and individual interventions.

					is scope for it to trigger changes.	Relevant data is collected to monitor outcomes and historic data sets that are no longer relevant are retired.
	Optimisation and automation of processes	No processes have been automated or improved using data.	Efforts to use data to improve services tend to involve very basic analysis, and is ad hoc across the organisation.	In some services data is used as part of efforts to improve processes, but data dashboards are not routinely available and no processes have been automated.	Data dashboards are used to optimise processes. Data is used to manage services and processes and some are automated.	Data is used in real time where possible, often with APIs. Processes which require little or no human judgement have been automated and optimised using data, such as detecting fraud and error.
Data Skills Capability		Skills and capacity are limited to IT system managers and basic software use. Most staff lack basic data literacy and skills.	Some staff are able to use basic software for simple analysis. Data literacy is patchy.	Data integration and analysis can be performed by some staff, but is not highly sophisticated. Most staff have a basic level of data literacy.	Sophisticated analysis can be undertaken, but not consistently across the organisation. Some staff have good data literacy but it is not uniform.	Data analysts are highly skilled and can work with multiple software packages. Sophisticated data science can be undertaken routinely across the organisation. All staff have a level of data literacy appropriate to their role. The organisation has timely access to all its data from line of business systems whether held internally or in Cloud facilities.
Data Awareness and Culture		There is limited awareness of how data can be used to improve services and outcomes.	Data is seen as having some value in niche uses, but most staff do not routinely try to use data to help them with their work.	Data integration and analysis can be performed by some staff, but is not highly sophisticated. Most staff have a basic level of data literacy.	There are some highly data-literate staff and the culture of the organisation expects data to be used in decision-making and service delivery.	All staff see data as a tool which can support them to do their jobs better.

Source: https://media.nesta.org.uk/documents/wise_council.pdf