NITI Aayog

DEVELOPMENT MONITORING AND EVALUATION OFFICE

DATA GOVERNANCE QUALITY INDEX (DGQI) 2.0

Quarterly Summary Report (Q4 FY 2021-22)

October, 2022

This page is left intentionally blank.

Preface

The Central Government of India, through its Ministries and Departments spends an amount to the tune of more than Rs. 10 lakh Crores on various Central Sector (CS) and Centrally Sponsored Schemes (CSS). With rapidly evolving governance needs and tremendous growth in data capabilities with the advent of the Fourth Industrial Revolution technologies, it is crucial for governments to ride this transformative wave and shift to evidence-based policymaking for efficient utilization of resources to achieve intended development outcomes.

Against this backdrop, the Prime Minister's Office had directed DMEO and NIC to undertake an exercise to assess the data preparedness of Ministries and Departments (M/Ds) of the Government of India. For developing the study methodology, an in-depth literature review of various global and domestic data maturity models was undertaken. Subsequently, three pillars of data preparedness were identified, viz. (a) Data Strategy to lay down systemic guidelines, (b) Data Systems to ensure smooth processes of data generation, management and its use, and (c) Data-driven Outcomes where data is utilized and widely shared across institutions by multi-disciplinary teams to drive policymaking.

Centered around data systems, the first phase of Data Governance Quality Index (DGQI) was conceptualized and conducted in 2020-21 in self-assessment mode with 65 M/Ds covering approx. 250 Central Sector (CS)/Centrally Sponsored (CSS) schemes. The exercise showed disparities among M/Ds with regard to their data systems and highlighted huge scope for improvement.

Consequently, DGQI was envisaged to become a regular exercise for monitoring data preparedness levels and utilizing this measurement for driving specific reforms. M/Ds were encouraged to prepare a data strategy and establish a Data & Strategy Unit for improving their data maturity levels.

Following this, DGQI 2.0 was launched with enhanced horizontal and vertical scope. As part of the first round of this exercise conducted in September 2021, 74 M/Ds self-assessed their data maturity levels for about 630+ CS/CSS schemes and non-schematic interventions.

The second round of DGQI 2.0 for M/Ds to self-assess their data preparedness levels as of end of Quarter 4 of Financial Year (FY) 2021-22 (31st March, 2022) was launched in April, 2022. During this round, 74 M/Ds provided information on the twelve themes of DGQI exercise via a newly-developed, customized DGQI dashboard for 650 CS/CSS schemes and non-schematic interventions.

This report summarizes key findings based on latest DGQI scores as of Quarter 4 of FY 2021-22, and also compares movements from the previous round of the exercise (as of Quarter 2 of FY 2021-22). It identifies key areas for improvement and suggests a way forward for M/Ds.

This report is prepared with the intention to aid policy makers at the highest level to take affirmative action to improve data preparedness within and across the

Ministries/Departments of the Government of India. This is expected to result in benefits in improving resource allocation, performance and progress monitoring, and creation of feedback loops to usher in the era of data led development in India. While the exercise, so far, has primarily focused on the Central Government, it is envisaged that it will eventually be taken by the Ministries/Departments to the States in the next phase.

Acknowledgements

This report is the culmination of the second round of Data Governance Quality Index (DGQI) 2.0 exercise and is the result of the arduous efforts of a wide range of stakeholders across the Government of India. Firstly, DMEO, NITI Aayog would like to express deep gratitude to Sh. Amit Khare, Adviser to the Hon'ble Prime Minister, and his predecessor Shri Bhaskar Khulbe, for their guidance and inspiration without which, it would not have been possible to successfully pursue and complete this demanding exercise. Further, we would like to especially thank the Secretaries of all the participating Ministries and Departments who enthusiastically supported this self – assessment exercise and facilitated the process of data collection at their end. We also acknowledge the efforts of all the scheme-level nodal officers in respective Ministries and Departments who proactively provided the required information and patiently engaged with the DMEO team during the whole exercise.

We would like to extend our thanks to the outstanding National Informatics Centre (NIC-NICSI) team, specifically, Sh. Prashant Kumar Mittal, Sh. Varindra Seth, Sh. Rajiv Rathi, Ms. Garima Sogani and Ms. Shehnaz Nayeem under the leadership of Sh. Rajesh Gera, Director General, NIC and Dr. Neeta Varma, former Director General, NIC, who not only created and maintained the DGQI dashboard required in data collection and collation, but also supported the DMEO team in providing handholding support to Ministries/Departments to ensure timely data collection.

We would like to call out the DMEO DGQI core team that spearheaded the exercise - Sh. Anand Trivedi, Ms. Gunjan Saini, Ms. Aishwarya Salvi, Ms. Vaishvii Goel, Ms. Aishwarya Ravikumar, Ms. Anmol Narain and Sh. Bishun Kumar Chaurasia, -- for their outstanding efforts, diligence and dedication under the overall leadership of Dr. Sekhar Bonu, former Director General, DMEO and guidance of Sh. Shailendra Kumar Dwivedi, Joint Secretary towards producing this report on Data Governance Quality Index. This is the third report in the series of DGQI reports produced over the last two years to document the journey of the Government of India in the direction of driving better data use for policy decisions.

Antony Cyriac
Director General,
Development Monitoring & Evaluation Office,
NITI Aayog
October, 2022

Contents

List of Figures

List of Tables

1 Executive summary

Data driven decision making is being globally accepted as an effective route to increasing efficiency in public administration and improving service delivery to citizens. Availability of high quality, timely and granular administrative data is crucial for enabling decentralized evidence-based policymaking in the Government of India. While most Ministries/Departments (M/Ds) of the Government of India have developed digital Management Information Systems (MIS) and dashboards for better implementation and monitoring of their initiatives, their data quality and maturity widely vary. Data also often exists in siloes in non-interoperable formats, making it less useful for analysis purposes.

In this context, the Data Governance Quality Index (DGQI) was launched in 2020 by DMEO, NITI Aayog and NIC/NICSI under at the behest of the Prime Minister's Office with the objective of carrying out a comprehensive assessment of data preparedness of M/Ds using a standardized framework, identify areas for improvement and trigger reform actions in this domain.

As part of the exercise, three key steps to attain better data maturity levels were first identified: a) Data Strategy to lay down systemic guidelines, (b) Data Systems to ensure smooth processes of data generation, quality control, management and its use and (c) Data driven Outcomes where cross-functional data is widely shared and utilized by institutions to drive decision making. Centered around data systems, the first phase of DGQI, conducted in 2020-21, covered 65 M/Ds self-assessing data maturity of approx. 250 Central Sector (CS) / Centrally Sponsored (CSS) schemes. The information provided by M/Ds was used to generate M/D wise and scheme specific DGQI scores.

As the exercise brought out glaring disparities and opened up more questions, it was considered necessary to take the DGQI exercise forward as a regular exercise for monitoring and improving data maturity of M/Ds. M/Ds were encouraged to prepare an action plan or data strategy to lay down concrete steps and approaches to improve their data preparedness levels to be reflected in their DGQI scores. They were also advised to set up a Data & Strategy Unit as a cross functional unit with four sub-units: Monitoring, Statistics, Analytics and Technology Units to have an intersectional lens and work in close coordination with all other divisions of the M/D to foster a data driven culture in the whole of the M/D.

Towards the redefined scope and objectives, DGQI 2.0 was launched in 2021 with enhanced horizontal focus on all three pillars: Data Systems, Data Strategy and Data driven Outcomes. Additionally, the scope was vertically expanded to mandatorily include all CS/CSS schemes (except those specifically exempted) as well as non-schematic interventions (NSI) such as sectoral dashboards, SDG dashboards etc. (overall, 630+ CS/CSS schemes and NSIs) of 74 M/Ds. Based on the information reported by M/Ds using the self-assessment questionnaire between August to October 2021 (status as of end of September, 2021) draft reports were shared with M/Ds for feedback in December 2021, and finalized and circulated in February, 2021.

After the successful conclusion of the first round of DGQI 2.0, the DGQI dashboard was released for use by M/Ds in March, 2022 to update the questionnaire for the second round of DGQI 2.0 i.e., status as of end of Quarter 4 of Financial Year 2021-22. 74 M/Ds provided information on 650 CS/CSS schemes and NSIs using the dashboard between April, 2022 to June, 2022. The findings from this round are summarized in this report.

It was positive to note that average DGQI score went above from 2.85 in DGQI 2.0 Quarter 2 FY 2021-22 round to 3.20 in DGQI 2.0 Quarter 4 round. For the first time, majority of M/Ds (~35%) were in the highest quintile of scores between 4 to 5. However, there is still much scope for improvement for all M/Ds to move to the frontier scores. While the Ministry of External Affairs emerged as the top performing M/D, the Ministry of Tourism showed maximum potential for improvement. The trend followed across categories with minimum variation. However, Scientific and Infrastructure category M/Ds outperformed, while Admin and Strategic categories need to pay special attention.

Among the three pillars, there was again minimum variation, with their scores ranging between 2.99 to 3.27. It was encouraging to note that average pillar wise scores had improved by 12% and theme wise scores by 15%. Under data strategy pillar, 91% of M/Ds were found to have set up a DSU, but only half of them filled up over 80% of the posts. There is still a huge scope for DSUs to have well-defined ToRs and review mechanisms so that they can continue to function towards the overall objective of embedding data-based decision making across the Ministry/ Department. 85% of the M/Ds have also prepared action plans, but there is much that needs to be done for improving their quality and ensuring they are implemented as per set timelines. Only 10 out of 74 M/Ds were able to complete all their action points that were due by the end of the quarter.

Within data systems, 24% of the schemes were found to not have yet been digitized which is a slight improvement over 30% in the previous round. Additionally, 10% of the schemes didn't undertake any type of data quality assessment and among the remaining ones, only 60% followed all suggested protocols. Further, 23% of the schemes reported to not undertake any type of data analysis, among the remaining ones, 70% largely undertake just descriptive analysis. There is hence a need to move to more enhanced forms of data analysis as well as make active use of it for releases and budgetary decisions. Use of emerging technologies and alternative data sources was also restricted to half of the assessed schemes, though it was an improvement over 25% in the last round. Like the previous round, HR capacity was again found in the need to be ramped up as around 30% of the schemes still needed to put data quality and analysis teams in place.

Under the final pillar of data driven outcomes, there was significant potential for improvement across themes. Only 55% M/Ds reported to have set up mechanisms for exchanging data internally among divisions within the Ministry. Data collaborations with other M/Ds and agencies as well as prescriptive analytics was again rarely practiced, similar to the last round.

To summarize, there has been a significant improvement in the performance of the M/Ds from Quarter 2 to Quarter 4 of FY 21-22. However, there is still a need for continuous

focus on certain key reform areas. All M/Ds must focus on digitizing all their CS/CSS schemes and other crucial non-schematic interventions in mission mode. Wherever already digitized, M/Ds need to have well-defined data strategies to provide guidelines on standard practices to be followed for robust data generation, data quality assessment and ensuring data security and privacy. M/Ds need to specifically prioritize developing automated mechanisms for regular data exchange among scheme divisions within M/Ds as well as inter-ministerial data exchange. They need to identify use cases of cross-sectoral data for their decision making and develop mechanisms for enabling the same. Most important of these applications can be to link data of progress made by schemes on the achievement of intended outputs and outcomes with the budgeting and disbursement processes at Ministries/Departments. It is expected that DSUs at M/Ds would develop a robust mechanism with the Integrated Finance Department of the M/ Ds to enable this type of data exchange and use for policymaking.

M/Ds also need to urgently focus on capacity development of government officials especially in data management and data analytics domain to improve their data governance methods. To complement the overall capacity development efforts being undertaken as part of the DGQI exercise via multiple channels, DSU at M/Ds need to partner with private sector experts to complete their training needs assessment, develop tailored capacity development plans as per their needs and begin upskilling of their resources at all federal levels.

The findings of the second round of DGQI 2.0, as summarized in this report, would be useful for all M/Ds to understand their progress till date, further identify their areas for improvement and their prioritization, to outline action points that may need to be taken or finetuned to achieve greater data maturity. It is also envisaged that the report allows for peer learning among M/Ds as they get opportunities to learn from each other's journey where new success stories are being weaved in every round. It is anticipated that this will be reflected in further improvement in the next round of DGQI 2.0 exercise (status as of end of Quarter 2, FY 2022-23) which is planned to commence soon. This way, by collaboration among several partners, it is hoped that DGQI exercise would be used by M/Ds for fostering a culture of evidence-based policymaking with use of high-quality interconnected data and cross-functional prescriptive analytical capabilities.

2 Background

Rapid digitization and evolution of emerging technologies have transformed the nature of governance across the world. Increasing demand for transparency, accountability and public participation has modified the role of data in public policy across its lifecycle. Data preparedness allows governments to design well-designed and well-targeted policies and programmes, make mid-course corrections and evaluate impact at the end of its lifecycle to inform future decisions.

Over the last two decades, most M/Ds of the Government of India as well as State Governments have developed digital Management Information Systems (MIS) and dashboards for better implementation and monitoring of their initiatives. However, there are a lot of variations in their data granularity, frequency as well as quality. Data also often exists in siloes in non-interoperable formats, making it less useful for cross-functional analytics purposes. Though multiple policies have been rolled out to further data exchange among government agencies and various open data platforms have been developed to disseminate non-personal data for public use, compliance has not been uniform. As a consequence, even though a lot of data is generated in public administration processes, its potential in terms of using it for driving day-to-day decisions and ensuring seamless service delivery to citizens remains untapped.

Against this background, a comprehensive review of present data preparedness levels of all M/Ds was required to identify areas for improvement. In tune with this, the Data Governance Quality Index (DGQI) exercise was initiated in 2020 by the Development Monitoring and Evaluation Office (DMEO), an attached office of NITI Aayog along with the support of NIC/NICSI and concerned line M/Ds. The first phase of DGQI exercise was conducted in 2020 in self-assessment mode with 65 M/Ds covering approximately 250 Central Sector (CS) / Centrally Sponsored (CSS) schemes. The M/Ds filled up an online survey, the responses of which were used to come up with the DGQI scorecard. The exercise showed glaring disparities among M/Ds and highlighted a huge scope for improvement across the board.

Subsequently, DGQI 2.0 was launched in 2021 with enhanced horizontal (covering all three steps of data preparedness, i.e., data strategy, systems, and data-driven outcomes) and vertical scope (in terms of number of M/Ds and schemes as well as non-schematic interventions) as a regular exercise for monitoring and directing reforms in the direction. The first round of DGQI 2.0 exercise was conducted between August, 2021 to October, 2021. 74 M/Ds provided information on data maturity of 630+ CS/CSS schemes and NSIs via an online self-assessment questionnaire spanning across 12 themes within the three pillars. The responses were used to prepare DGQI scorecards and reports which were drafted in December, 2021 and finalized in February, 2022.

After the successful conclusion of the first round of DGQI 2.0, the DGQI dashboard was released for use by M/Ds in March, 2022 to update the questionnaire for the second round of DGQI 2.0 i.e., status as of end of Quarter 4 of Financial Year 2021-22. The operational approach and findings of this round have been summarized in next few sections.

3 About DGQI 2.0

3.1. Intent and Objectives

Under the overarching objective of driving continuous reforms by M/Ds in increasing their data preparedness, the intent of the DGQI is to enable M/Ds to assess themselves at various levels of data maturity on the basis of a standardized framework. This is expected to deepening of digitization in the Government of India and creation of interoperable data systems.

It is hoped that in the long run, DGQI will help in laying the foundation of more integrated data systems, for e.g., a single, online, API-integrable 'Overarching Dashboard' data system linked to MIS/dashboards of all M/Ds, ultimately leading to a state-of-the-art data-driven decision making.

In this regard, DGQI aims:

- To prepare a self-assessment diagnostic tool that will enable the M/Ds to internally contemplate the need for improving data systems.
- To enable review and assessment of data preparedness of the data/ MIS systems of the M/Ds on objective parameters of a standardized framework.
- To enable M/Ds to drive reforms in the direction and improve their data systems so as to attain frontier DGQI score of 5.0 by December 2022
- To enable peer learning by sharing best practices across M/Ds.

3.2. Guiding Principles

DGQI 2.0 is action-oriented and encourages M/Ds to act upon learnings from the findings of this exercise. In accordance with this, there are three crucial and distinguishing principles of DGQI 2.0 which need special mention.

- **Emphasis on individual 'Distance to Frontier' for each M/D, instead of comparisons**: DGQI exercise is aimed at initiating reforms to move to better data and make better use of data across M/Ds. It uses a singular framework for all to ensure some standard practices can be adopted across robust data systems which can talk to each other and can be used in synergies as and when required. It is therefore crucial for all M/Ds to move to the frontier together by bringing each other up. Hence, the main focus of DGQI 2.0 is on individual 'Distance to Frontier' for each M/D, thereby urging each M/D to improve their data preparedness level and emphasize on attaining DGQI frontier score of 5.0 rather than undertaking any other comparison.

- **Equal attention to all schemes/non-schematic interventions irrespective of budgetary outlays to drive all-inclusive growth**: Further, in order to encourage intra-ministerial dialogue on reforms, each M/D is provided with a scheme-wise score as well as an overall MD score. This aims to enable self-benchmarking of schemes within M/Ds so that they are aware of areas for improvement for each scheme and accordingly take mitigative steps. Just like M/Ds, it is equally important for all schemes' MIS and

dashboards to move to the frontier together, irrespective of their budgetary outlays. Hence, the exercise accredits equal attention to all schemes/NSIs in its methodology to drive the idea of holistic growth.

- **Due consideration to present resource constraints faced by M/Ds and non-applicability of certain features of the framework to specific schemes:** This exercise also considers the present resource constraints faced by M/Ds while laying down a way forward for them. It hence ensures that M/Ds are not unnecessarily adversely scored for features not applicable to them. It takes into consideration the non-applicability of certain questions to certain M/Ds and makes adequate provision to discount them during scoring.

3.3. Architecture

Building on the above-mentioned principles, DGQI 2.0 is designed to trigger and incentivize reforms.

It is pertinent for all M/Ds to frame a data strategy with detailed guidelines and actionables for improving data systems and its use as well as for making provisions for infrastructural, human and financial resources. As part of DGQI 2.0, M/Ds have hence been encouraged to prepare an action plan or data strategy for improving data preparedness. An indicative outline of the action plan (see Annexure 1) was prepared and shared by DMEO, NITI Aayog in 2021 to support M/Ds in the process.

It is also crucial to have an institutional mechanism within M/Ds to drive the implementation of the data strategy. As part of DGQI 2.0, M/Ds have hence been advised to set up such an institution in the form of a 'Data and Strategy Unit (DSU)', headed by a senior official directly reporting to the Secretary of the M/D. A detailed Terms of Reference (see Annexure 2) was prepared and shared by DMEO, NITI Aayog in 2021 to support M/Ds in the process.

As these reforms were launched, DGQI 2.0 assessment also focused on assessing M/Ds on these reforms.

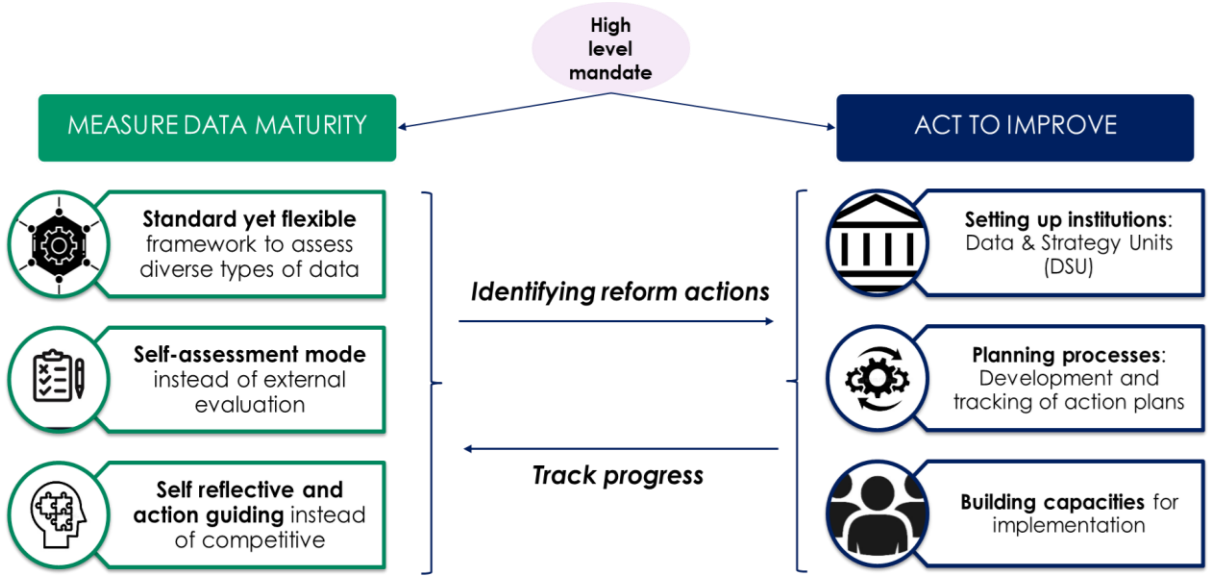


Figure 1 Data Governance Quality Index (DGQI) Architecture

3.4. Scope

As part of DGQI 2.0, 74 M/Ds have been mandated to include all CS/CSS schemes (except those exempted on rational grounds such as discontinued/to be discontinued, national security etc.) as well as non-schematic interventions (NSI) such as sectoral dashboards, SDG dashboards etc. Inclusion of NSIs over and above schemes help bring other noteworthy digital interventions under the ambit and identify additional digital practices of M/Ds which can also be benchmarked under the framework.

As part of the second round of DGQI 2.0 exercise (as of end of Quarter 4 of FY 2021-22), M/Ds provided data on 650 interventions (433 CS schemes, 110 CSS schemes and 107 NSIs), a marginal increase over the coverage in the previous round of DGQI 2.0 (as of Quarter 2 of FY 2021-22).

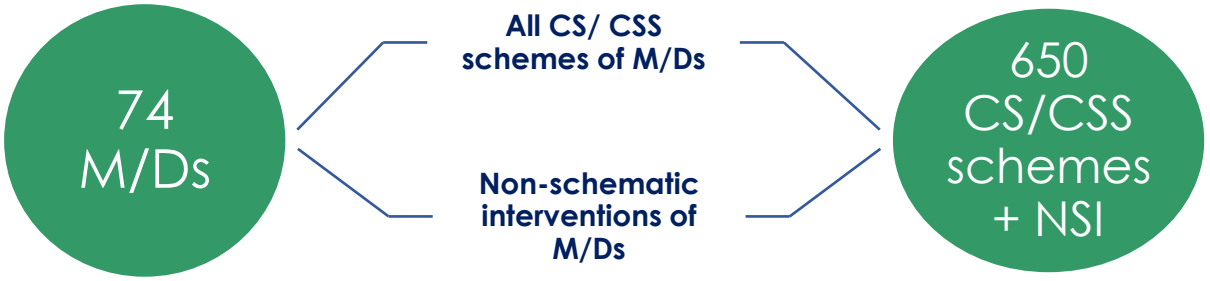


Figure 2 Vertical Scope of DGQI 2.0 Round 2

Table 1: Vertical scope expansion under DGQI 2.0

DGQI Round	No. of Ministries/ Departments	No. of CS/ CSS
DGQI 1.0	65	249

DGQI 2.0 Quarter 2 FY 21-22	74	634
DGQI 2.0 Quarter 4 FY 21-22	74	650

3.5. Operational Approach

After the successful culmination of the first round of DGQI 2.0 exercise in February 2022 when DGQI 2.0 final reports were shared with M/Ds, DGQI 2.0 dashboard was released for use by M/Ds for the second round (status as of end of Quarter 4, FY 2021-22) in March, 2022. All M/Ds were provided with separate access to fill, submit and approve the questionnaire responses on the dashboard. The DGQI 2.0 self-assessment questionnaire and methodology were also slightly modified for this round, in order to incorporate feedback received from M/Ds in the previous round, and released along with the dashboard.

During this process, JS/ Director level nodal officers nominated by each M/D assisted in coordinating and driving the entire exercise at the M/D level. Each M/D was provided with two types of user roles to access the dashboard: scheme divisions were provided with data entry user roles to enter data on the dashboard; the nodal officer or the head of the DSU of the M/D was provided with a MD Admin role to approve and finally submit the data on the dashboard. To facilitate the Ministries/ Departments in understanding the task at hand, several rounds of webinars were conducted by DMEO and NIC in April, 2022 to walk them through the dashboard and address their queries.

DMEO & NIC team also conducted regular follow-ups with the Ministries/ Departments, both telephonically and through emails, to prompt them for timely filling up of the survey and also extended support at all stages. Finally, after receiving all the requisite data (as of end of Quarter 4 of FY 2021-22) from all M/Ds by the end of June 2022, DMEO scored best practices on the dashboard and conducted data analysis. This draft report contains the findings emanating from this round of data collection and analysis.

4 Methodology

Based on a detailed assessment of several data maturity models (see Annexure 5 for details), three key pillars of data preparedness have been identified, viz., Data Strategy, Data Systems and Data Outcomes. This theory of change forms the basis for design of DGQI. All pillars are explained as mentioned below.

4.1. Data Strategy

Under the data strategy pillar, two themes are covered within DGQI 2.0: a) Data and Strategy Unit and b) Action Plan. Under Data and Strategy Unit theme, it was assessed if M/Ds have taken necessary steps to establish the DSU with adequate strength and review mechanisms. Under the Action Plan theme, it was assessed if M/Ds have developed action plans as per the outline. In addition, the compliance by M/Ds in completing the action points within the timelines set by them was also measured.

4.2. Data Systems

Under the data systems pillar, six themes are covered within DGQI 2.0: data generation (ability of M/Ds to collect and digitize data at high granularity and frequency); data quality (practices adopted by M/Ds to undertake data quality assessment of incoming data); data analysis, use & dissemination (use of collected data for analysis and decision making, open data and modes of dissemination); use of technology (use of emerging technologies and alternative data sources); data security & HR capacity (measures to ensure data security and protection of personal data and existence of data QC and analysis teams); and data management (adoption of lifecycle approach to data management).

4.3. Data driven outcomes

Under this pillar, four themes have been identified under DGQI 2.0: Synergistic data use within M/Ds (creation of better exchange systems within M/Ds to drive integrated data use); inter-agency collaboration (data-based collaborations with other agencies to drive better data-based outcomes); prescriptive analytics (creation of data culture by moving to prescriptive analytics); and good practices (good practices in using data in driving smarter, granular and quicker decisions).

4.4. Scoring

A self-assessment questionnaire has been devised around the above themes and DGQI scores are arrived at on the basis of responses filled up by M/Ds to this questionnaire. The questionnaire consists of two parts: Part A (to be filled at M/D level) and Part B (to be filled for each CS/CSS scheme/non-schematic intervention at CS/CSS/NSI level). The questionnaire can be viewed at Annexure 3.

The response to each question is scored on a scale of 0 to 5, which is then aggregated using weighted averages to arrive at scores at themes, pillar and overall index level (all scores range between 0 to 5). The data systems pillar is appropriated an overall weight of 60% as it is a major pillar where outputs of data strategy are visible which then also play a key role in the ability of M/Ds to achieve desired data driven outcomes. Remaining

40% weight is accorded to the data strategy and data driven outcomes pillar combined. This 40% is distributed equally between data strategy (20%) and data driven outcomes (20%).

Hence, overall DGQI Score = 60% *(Data systems pillar score) + 20% *(Data strategy pillar score) + 20% *(Data driven outcomes pillar score)

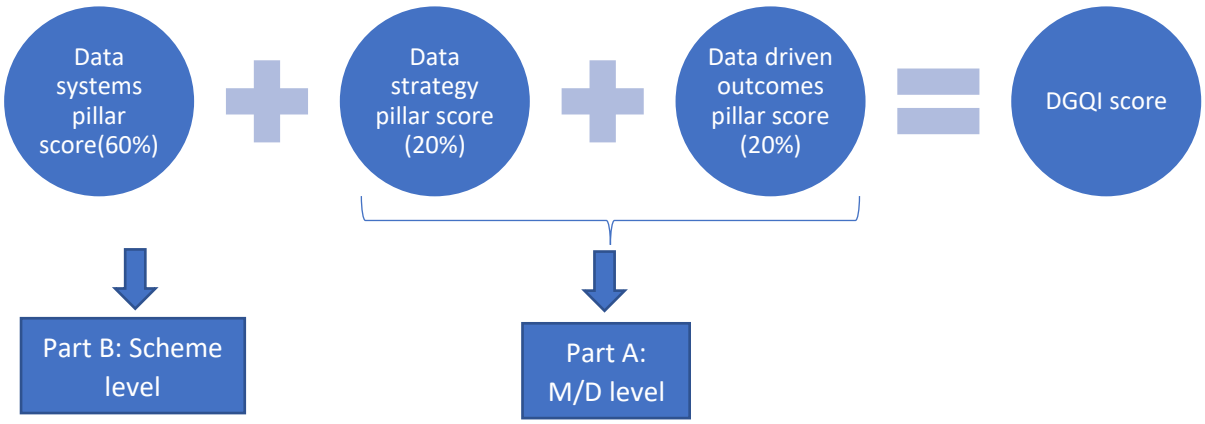


Figure 3 Calculating overall DGQI 2.0 score

Theme wise and question wise weightages can be found at Annexure 4.

4.5. Special Cases

To consider the non-applicability of certain questions or sub-parts of questions, NA option is explicitly included in the DGQI self-assessment questionnaire. For a certain question, if NA option is selected, its weight has been redistributed among other questions within the theme. However, if it is the case that only certain sub-parts (a,b,….) of a question are not applicable, a case-by-case mechanism of how they will be taken care of at the scoring stage has been devised. The same can be found at Annexure 4.

If any question is disabled based on skipping patterns, it would accordingly be given appropriate score. For example: If action plan is not formed, M/Ds would be scored zero on all other questions related to action plan that get automatically skipped.

4.6. Limitations of the Study

Exhaustiveness: As the study relies on self-assessment, it is not mandatory for all M/Ds to add all their NSIs and thus exhaustiveness on these grounds cannot be guaranteed.

First of a kind: Considering that the study is unique, some responses by M/Ds didn't follow the skip patterns correctly maybe due to lack of preparedness about the nature of the survey. In such cases, responses have been estimated based on their responses to other questions.

Diversity of the study: The self-assessment involved 630+ interventions of 74 M/Ds with extremely diverse nature of objectives and operations. Developing a general questionnaire while capturing these variations across the M/Ds also had its limitations.

5 Findings

This section presents the major findings emerging from the analysis of the information provided by 74 Ministries/Departments on 650 CS/CSS schemes/NSIs. Details on number of CS/CSS schemes/NSIs covered for each M/D can be found at Annexure 5. The section has been organized under four sub-sections: Overall performance of M/Ds, category wise performance of M/Ds, pillar wise performance and performance on few key themes.

Under category-wise performance, based on their scope and functions, 74 M/Ds have been classified into 6 categories namely: Admin, Economic, Infrastructure, Scientific, Social and Strategic. Comparisons among peer M/Ds lying within the same category have been provided to encourage peer learning.

5.1. Overall performance

Figure 4 shows the frequency distribution of M/Ds based on their scores in DGQI 1.0, DGQI 2.0 Q2 (Quarter 2 FY 2021 -22) and DGQI 2.0 Q4 (Quarter 4 FY 2021 -22) rounds.

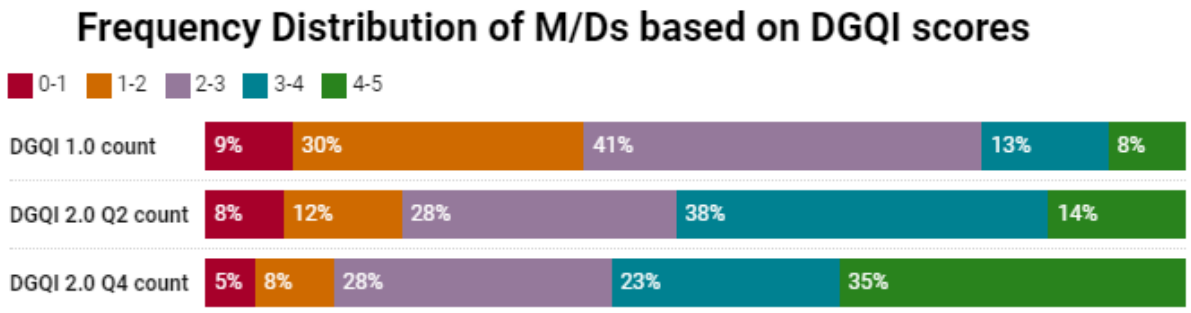


Figure 4 Frequency Distribution of 74 Ministries/Departments on basis of DGQI 2.0 Round 2 scores

As per the information submitted by the M/Ds in the latest round of DGQI exercise, it was positive to note that the majority and over one-third (26/74) of the M/Ds scored in the highest quintile i.e., between 4 and 5. However, 4 M/Ds still demonstrated a huge scope for improvement with their scores in the lowest quintile i.e., between 0 to 1.

It was also encouraging to note that the share of M/Ds in the highest quintile i.e., scores between 4 to 5 has been continuously increasing with each round of the exercise, coupled with a considerable decrease in the share of M/Ds falling in the first three quintiles i.e., with DGQI scores between 0 to 3. The same trend was also reflected on observing average DGQI scores which have seen a continuous improvement over last three rounds – from 2.50 to 3.20.

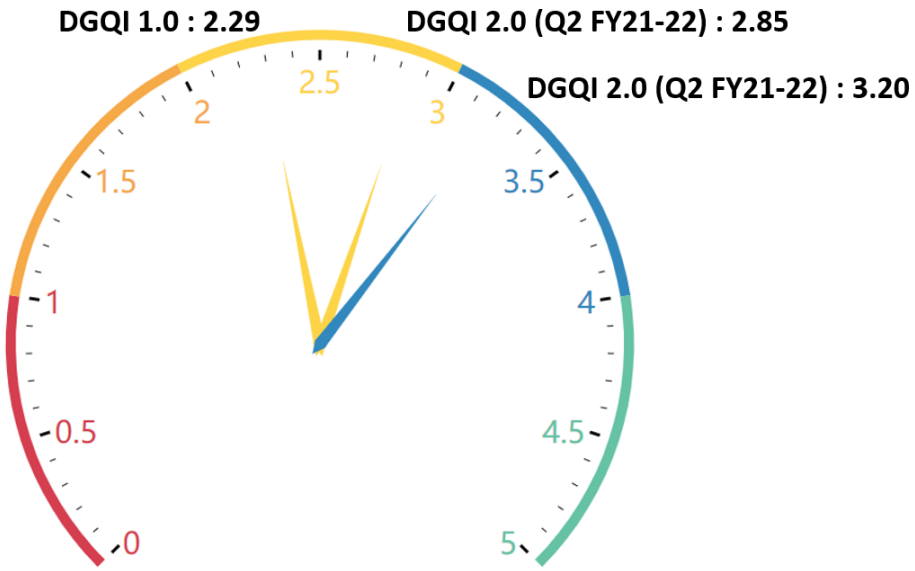


Figure 5 Time trend of average DGQI scores in last three rounds

The overall performance and relative rankings of 74 M/Ds on DGQI 2.0 in Q4 of FY 2021-22 is provided below in Table 2. The scores have been coloured with a darker gradient as the score increases i.e., the distance to frontier 5.0 score decreases. For ease of comparison, DGQI scores of all M/Ds as of Q2 FY 2021-22 is also provided in the adjacent column along with differences in scores and ranks between the two rounds.

Table 2: Overall performance of M/Ds in DGQI 2.0 Round 2

Rank (Q4 FY 21-22)	Difference in rank from previous round	Ministry/ Department	DGQI 2.0 score Round 2 (Q4 FY 21-22)	DGQI 2.0 score Round 1 (Q2 FY 21-22)	Difference in scores
1	+30	Ministry of External Affairs	4.69	3.30	+1.39
2	0	Department of Rural Development	4.65	4.37	+0.28
3	-2	Ministry of Earth Sciences	4.64	4.56	+0.08
4	0	Ministry of Tribal Affairs	4.62	4.31	+0.31
5	0	Department of School Education and Literacy	4.62	4.28	+0.34
6	+16	Department of Food and Public Distribution	4.55	3.73	+0.82
7	+34	Department of Skill Development and Entrepreneurship	4.42	2.78	+1.65
8	+12	Department of Heavy Industry	4.41	3.75	+0.66

Rank (Q4 FY 21-22)	Difference in rank from previous round	Ministry/ Department	DGQI 2.0 score Round 2 (Q4 FY 21-22)	DGQI 2.0 score Round 1 (Q2 FY 21-22)	Difference in scores
9	+5	Ministry of Petroleum and Natural Gas	4.36	3.86	+0.50
10	+17	Department of Financial Services	4.31	3.39	+0.91
11	+24	Ministry of Minority Affairs	4.29	3.15	+1.13
12	-4	Department of Science and Technology	4.26	4.08	+0.18
13	+4	Ministry of Power	4.26	3.81	+0.44
14	-7	Ministry of Food Processing Industries	4.20	4.15	+0.05
15	+8	Department of Promotion of Industry and Internal Trade	4.17	3.72	+0.46
16	-4	Department of Drinking Water and Sanitation	4.17	3.94	+0.23
17	+8	Ministry of Coal	4.16	3.54	+0.62
18	-5	Ministry of Road Transport and Highways	4.15	3.88	+0.27
19	+5	Department of Consumer Affairs	4.13	3.69	+0.43
20	-17	Department of Fertilisers	4.12	4.32	-0.21
21	+11	Ministry of Shipping	4.10	3.24	+0.86
22	-7	Ministry of Statistics and Programme Implementation	4.08	3.84	+0.24
23	-14	Department of Agricultural Research and Education	4.08	4.02	+0.06
24	-5	Department of Justice	4.05	3.77	+0.28
25	-4	Department of Public Enterprises	4.05	3.74	+0.31
26	-16	Ministry of Panchayati Raj	4.04	4.00	+0.03
27	-16	Department of Land Resources	3.96	3.98	-0.02
28	-2	Department of Space	3.92	3.49	+0.42
29	-11	Ministry of Housing and Urban Affairs	3.82	3.79	+0.04
30	-14	Ministry of Civil Aviation	3.77	3.84	-0.06
31	-2	Ministry of Railways	3.69	3.34	+0.35
32	+2	Department of Health and Family Welfare	3.55	3.20	+0.35
33	-27	Department of Biotechnology	3.54	4.21	-0.67
34	+4	Ministry of Environment, Forests and Climate Change	3.49	3.08	+0.41
35	-2	Department of Health Research	3.49	3.24	+0.24

Rank (Q4 FY 21-22)	Difference in rank from previous round	Ministry/ Department	DGQI 2.0 score Round 2 (Q4 FY 21-22)	DGQI 2.0 score Round 1 (Q2 FY 21-22)	Difference in scores
36	+8	Department of Telecommunications	3.48	2.61	+0.87
37	+3	Department of Economic Affairs	3.43	2.93	+0.50
38	-10	Department of Fisheries	3.26	3.34	-0.08
39	-3	Ministry of Electronics and Information Technology	3.22	3.15	+0.07
40	-10	Department of Pharmaceuticals	3.21	3.33	-0.12
41	+6	Department of Agriculture, Cooperation and Farmers' Welfare	3.17	2.56	+0.61
42	0	Department of Water Resources, River Development and Ganga Rejuvenation	3.04	2.62	+0.41
43	-6	Department of Animal Husbandry and Dairying	3.01	3.11	-0.10
44	+1	Ministry of Women and Child Development	2.90	2.60	+0.30
45	+18	Department of Personnel & Training	2.87	1.64	+1.23
46	+22	Ministry of Mines	2.87	1.12	+1.75
47	-8	Department of Chemicals and Petrochemicals	2.87	2.93	-0.06
48	+7	Ministry of Planning	2.86	2.16	+0.70
49	-3	Department of Defence	2.85	2.58	+0.27
50	+14	Department of Commerce	2.77	1.55	+1.23
51	+2	Department of Higher Education	2.70	2.24	+0.46
52	-4	Ministry of Steel	2.65	2.53	+0.13
53	-2	Legislative Department	2.64	2.39	+0.24
54	-5	Department of Posts	2.55	2.48	+0.07
55	-5	Ministry of Corporate Affairs	2.55	2.47	+0.08
56	-4	Ministry of Information and Broadcasting	2.38	2.34	+0.04
57	+1	Ministry of Textiles	2.32	2.06	+0.26
58	-15	Ministry of Labour and Employment	2.28	2.62	-0.34
59	-3	Ministry of New and Renewable Energy	2.26	2.16	+0.11
60	+2	Department of Social Justice and Empowerment	2.26	1.80	+0.46

Rank (Q4 FY 21-22)	Difference in rank from previous round	Ministry/ Department	DGQI 2.0 score Round 2 (Q4 FY 21-22)	DGQI 2.0 score Round 1 (Q2 FY 21-22)	Difference in scores
61	-1	Department of Scientific and Industrial Research	2.22	1.98	+0.24
62	+12	Department of Sports	2.21	0.52	+1.69
63	-6	Ministry of Micro, Small and Medium Enterprises	2.12	2.07	+0.06
64	-10	Ministry of Development of North Eastern Region	2.12	2.18	-0.06
65	-6	Department of Administrative Reforms and Public Grievances	1.94	2.02	-0.08
66	+1	Department of Youth Affairs	1.94	1.12	+0.81
67	-6	Department of Legal Affairs	1.80	1.88	-0.08
68	+1	Ministry of Home Affairs	1.76	0.97	+0.79
69	-4	Department of Defence Production	1.19	1.29	-0.10
70	-4	Ministry of Ayurveda, Yoga and Naturopathy, Unani, Siddha and Homeopathy (AYUSH)	1.19	1.22	-0.03
71	+2	Ministry of Culture	0.96	0.60	+0.36
72	-2	Department of Empowerment of Persons with Disabilities	0.86	0.87	-0.01
73	-2	Department of Ex-Servicemen Welfare	0.82	0.70	+0.12
74	-2	Ministry of Tourism	0.62	0.66	-0.04

58 M/Ds improved their DGQI scores whereas 16 M/Ds witnessed a decrease in their scores during the aforementioned period. As some M/Ds grew faster than others, it was also noted that there were several M/Ds which improved their DGQI scores, however, their relative ranking remained the same or fell down.

5.2. Category wise performance

The 74 participating M/Ds have been classified into six categories: Administrative, Economic, Infrastructure, Scientific, Social and Strategic, based on the nature of their interventions. Shifts in category wise DGQI scores may be used to identify overall category wise trends as well as variations among M/Ds within a category.

Average DGQI scores for Quarter 2 and Quarter 4 of FY 21-22 of each category are shown below in Figure 6. The average score was found to be the highest for scientific M/Ds (3.73) and lowest for admin M/Ds (2.01), as found in the previous round. It was positive to note that category wise scores have improved for all the categories, with infrastructure

category reporting the highest jump of 0.46 points and scientific category showing the least shift by 0.08 points.

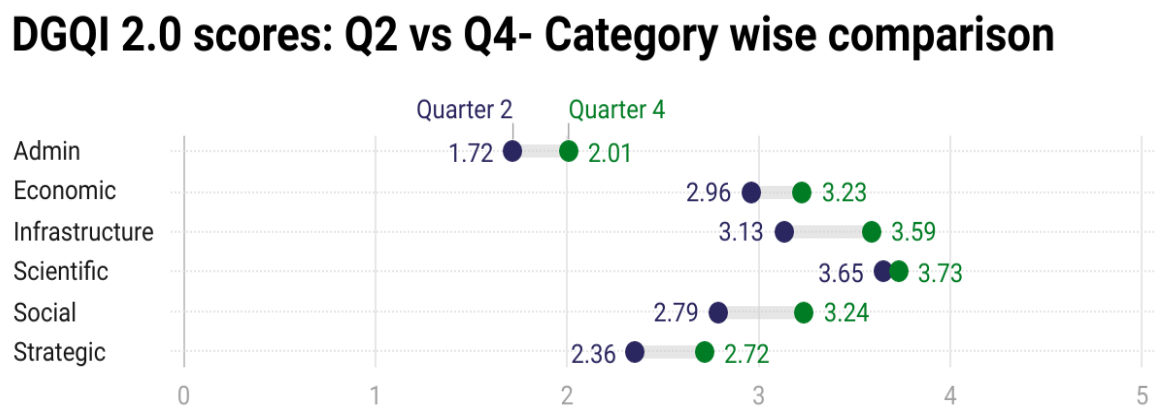


Figure 6 Category wise comparison of DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22)

Average DGQI 2.0 scores for Quarter 2 and Quarter 4 of FY 21-22 of Admin category Ministries/Departments are shown below in Figure 7. The Department of Personnel & Training recorded the highest score while The Department of Ex-Servicemen Welfare showed maximum scope for improvement. The Department of Personnel and Training also showed the highest jump of 1.23 points while the Department of Ex- Servicemen Welfare also marginally improved its score by 0.12 points. It is also pertinent to note that two M/Ds saw a reduction in their DGQI scores over the same period.

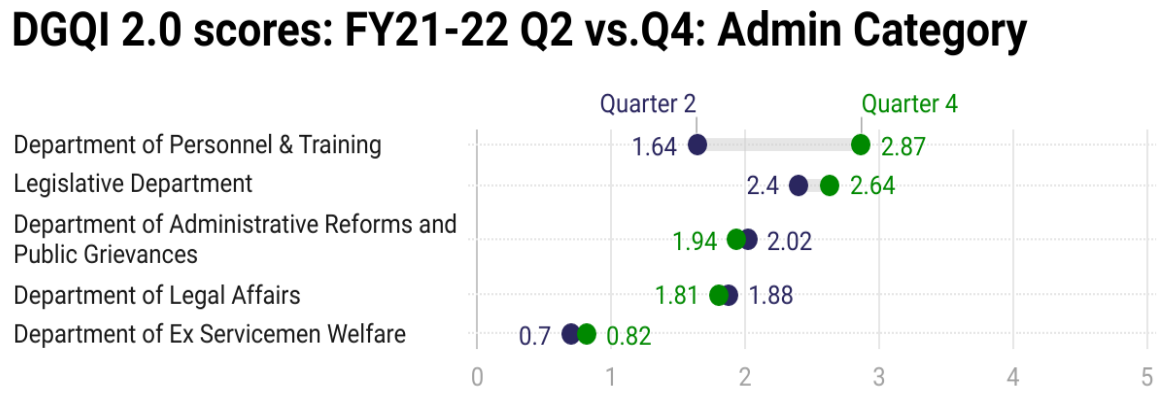


Figure 7 DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22) of M/Ds in Admin Category

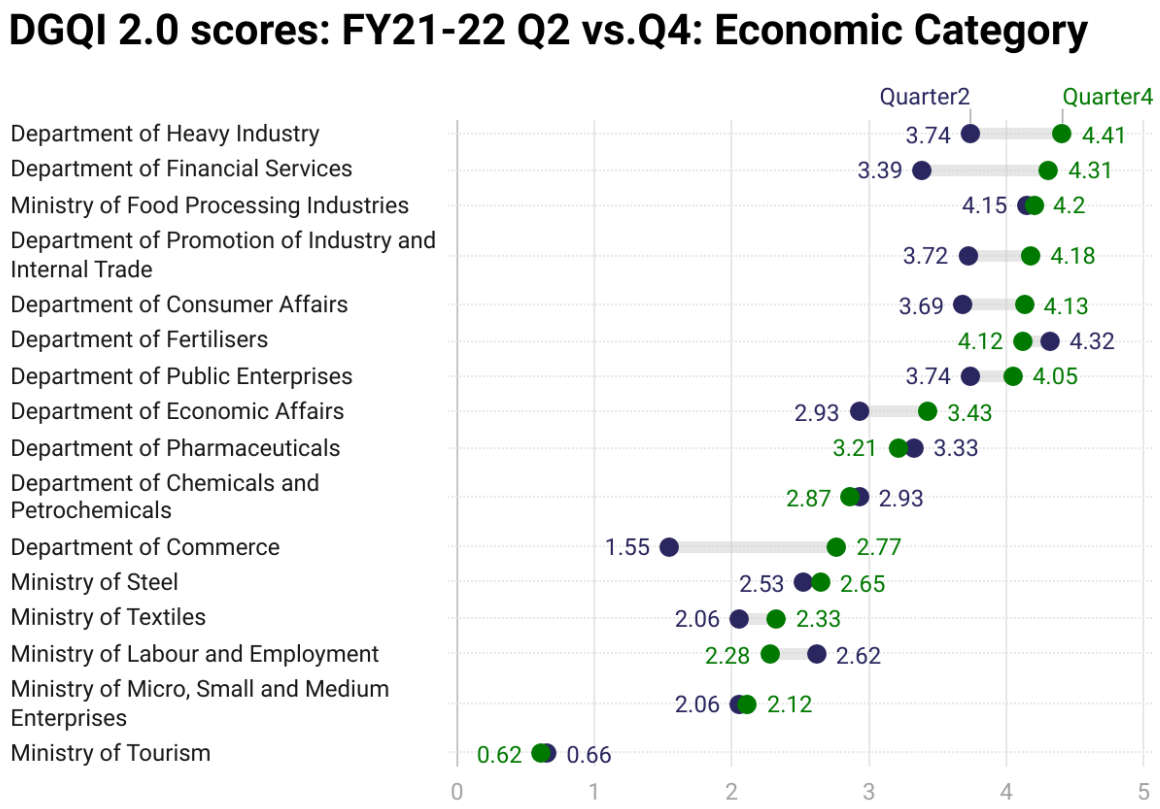


Figure 8 DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22) of M/Ds in Economic Category

Average DGQI scores for Quarter 2 and Quarter 4 of FY 21-22 of Economic category Ministries/Departments are shown above in Figure 8. The Department of Heavy Industry recorded the highest score in this category while The Ministry of Tourism showed maximum scope for improvement. Of the 16 M/Ds under this category, the Department of Commerce showed the highest jump in their DGQI scores by 1.22 points. 5 out of 16 M/Ds received lower DGQI scores in Quarter 4 as compared to Quarter 2 of the same year.

Average DGQI scores for Quarter 2 and Quarter 4 of FY 21-22 of Infrastructure category Ministries/Departments are shown below in Figure 9. The Ministry of Petroleum & Natural Gas recorded the highest score in this category while The Ministry of New & Renewable Energy showed maximum scope for improvement. Of the 13 M/Ds under this category, the Ministry of Mines reported the highest jump in their DGQI scores by 1.75 points while the Ministry of Housing and Urban Affairs recorded the smallest improvement by 0.03 points. Additionally, the Ministry of Civil Aviation observed a downfall in their DGQI scores over the same period.

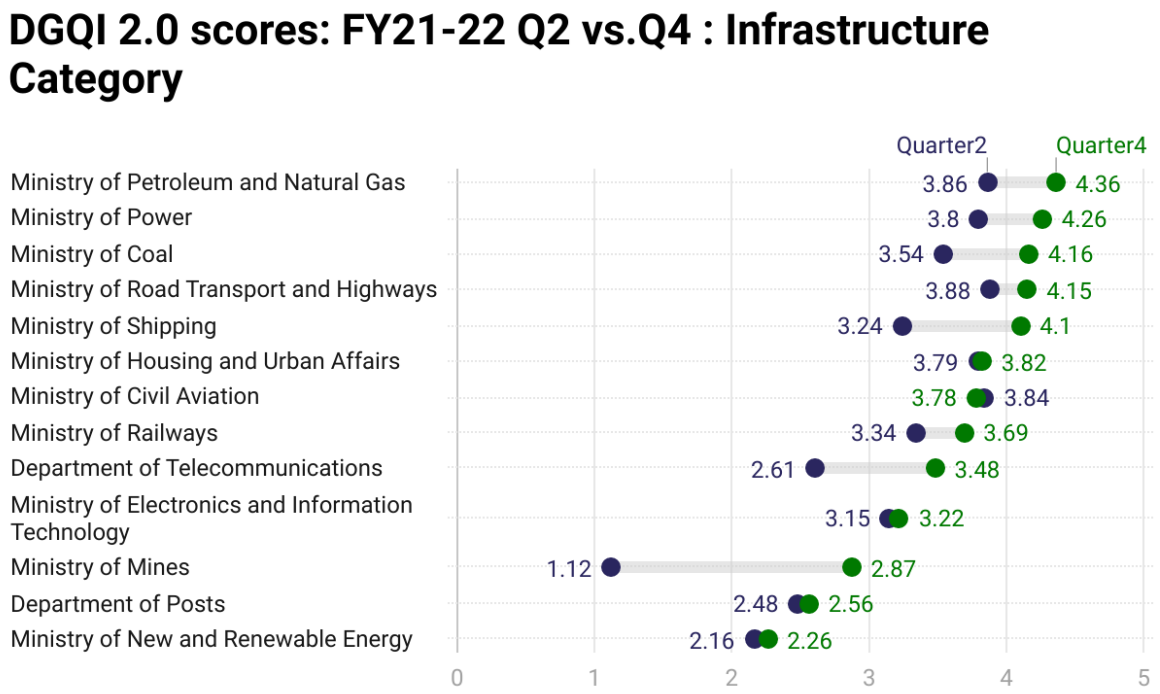


Figure 9 DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22) of M/Ds in Infrastructure Category

Average DGQI 2.0 scores for Quarter 2 and Quarter 4 of FY 21-22 of Scientific category Ministries/Departments are shown below in Figure 10. The Ministry of Earth Sciences recorded the highest score in this category while The Department of Scientific and Industrial Research showed maximum scope for improvement. Of the 7 M/Ds under this category, the Department of Space showed the highest jump of 0.43 points while the Ministry of Earth Sciences also marginally improved by 0.08 points. The Department of Biotechnology witnessed a reduction in their DGQI scores between the two rounds.

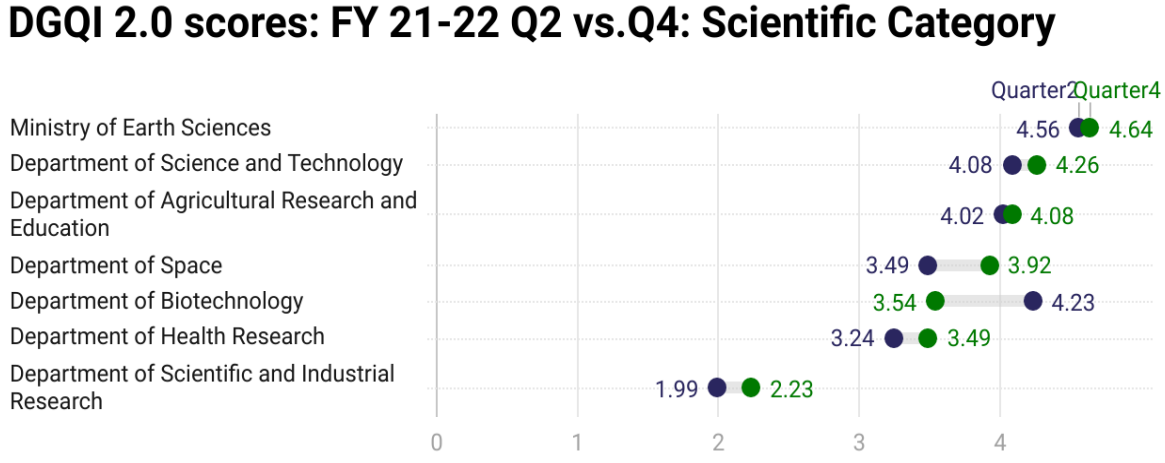


Figure 10 DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22) of M/Ds in Scientific Category

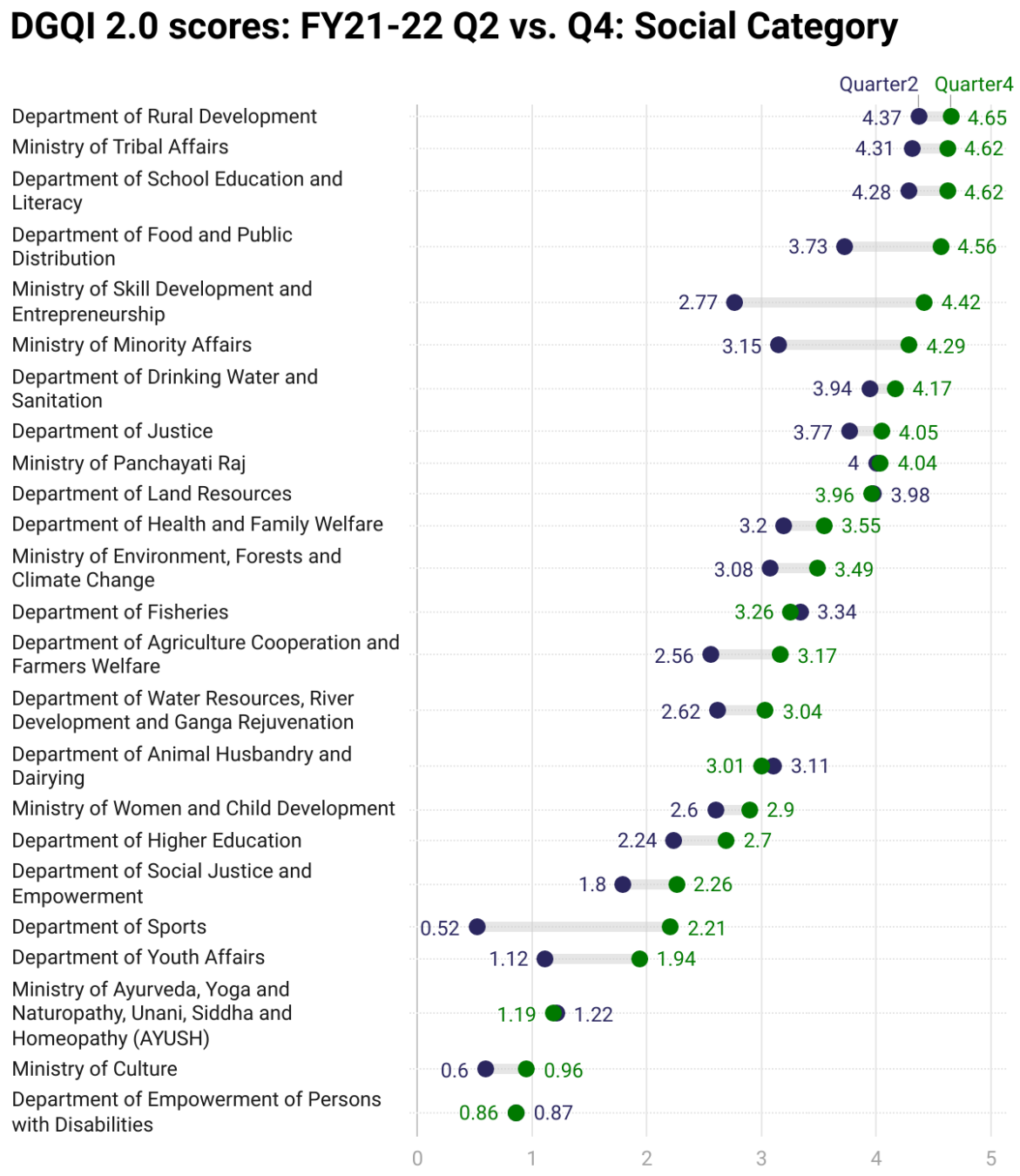


Figure 11 DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22) of M/Ds in Social Category

Average DGQI scores for Quarter 2 and Quarter 4 of FY 21-22 of social category Ministries/Departments are shown above in Figure 11. The Department of Rural Development recorded the highest score in this category while The Department of Empowerment of Disabilities showed maximum scope for improvement, similar to previous round. Of the 24 M/Ds under this category, the Department of Sports showed the highest jump of 1.69 points while the Ministry of Ayurveda, Yoga and Naturopathy, Unani, Siddha and Homeopathy (AYUSH) also reported a marginal improvement by 0.03 points. 5 out of 24 social M/Ds also witnessed a fall in their DGQI scores over these two rounds.

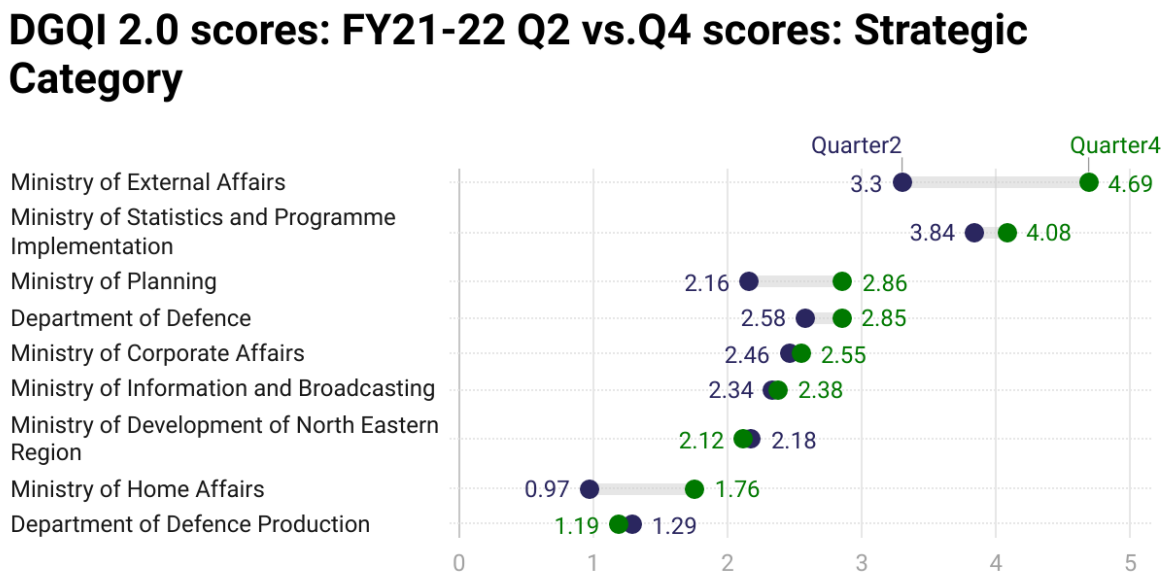


Figure 12 DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22) of M/Ds in Strategic Category

Average DGQI 2.0 scores for Quarter 2 and Quarter 4 of FY 21-22 of Strategic category Ministries/Departments are shown above in Figure 12. The Ministry of External Affairs showed the highest jump and recorded the highest score in this category while The Department of Defence Production showed maximum scope for improvement, same as in the previous round. Additionally, The Ministry of Development of North Eastern Region witnessed a fall in its DGQI scores over the same period.

It would also be worthwhile to observe category wise distribution of M/Ds based on average DGQI scores as of Quarter 4 of FY 21-22. As seen below in Figure 13, it was encouraging to note that Infrastructure, Economic and Scientific category M/Ds were heavily concentrated around scores between 3 to 5.

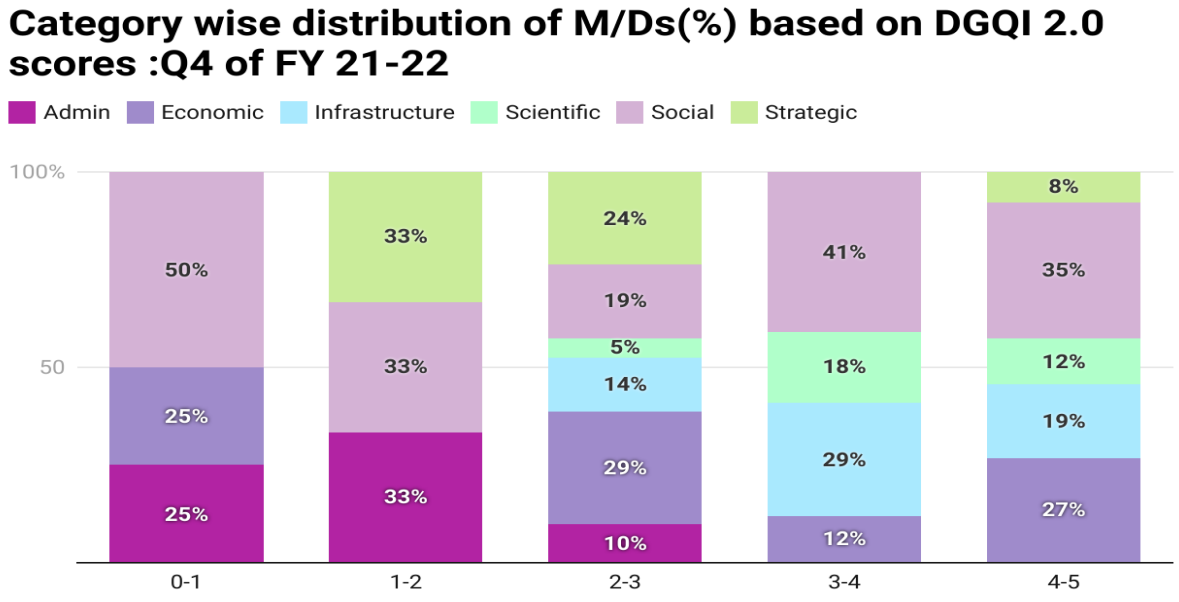


Figure 13 Category wise distribution of M/Ds based on DGQI 2.0 Q4 FY 21-22 scores

However, Admin and Strategic M/Ds had a heavier concentration towards lower end of the spectrum of scores and hence may need specific attention to improve their data preparedness levels. Further, Social category M/Ds were spread out across a range of score – while some performed very well, others demonstrated a huge scope for improvement. Similarly, few M/Ds in the Economic category scored in the lowest score quintile whereas there are other M/Ds in the same category which score much better. Hence, it was found that there is a lot of scope for peer learning within Social and Economic category M/Ds to learn from each other and improve the data preparedness of these categories at an overall level.

5.3. Pillar wise performance

As mentioned before in the methodology section, DGQI score is a weighted average of three pillar wise scores on data strategy, data systems and data driven outcomes. In this section, key findings emerging from DGQI pillar wise scores have been summarized.

As depicted below in Figure 14, M/Ds achieved highest average scores on the data systems pillar and marginally lower score on data strategy. Given that the average scores for all three pillars were found to be still lying in the average category of scores between 1.50 and 3.50, there is still a significant scope for improvement for the government to move to the frontier along all three directions. While it would be crucial to focus on developing well defined data strategies and translating them into high quality data systems, data driven outcomes need special focus to create better mechanisms for systemic data exchange and its use for prescriptive analytics, especially more so as the average scores on this pillar have been consistently lower in comparison to other two pillars.

While comparing pillar wise scores to the previous round of the exercise, it was found that the performance of M/Ds has appreciably improved across all three pillars of DGQI by an average increase of 12%. In Quarter 4, pillar wise scores were largely concentrated in the 4th quintile, which is an improvement from concentration of scores in 3rd quintile in Quarter 2. Still, as noted above, there is significant scope for M/Ds to further enhance their scores to reach the frontier DGQI score of 5.0 across all three pillars.

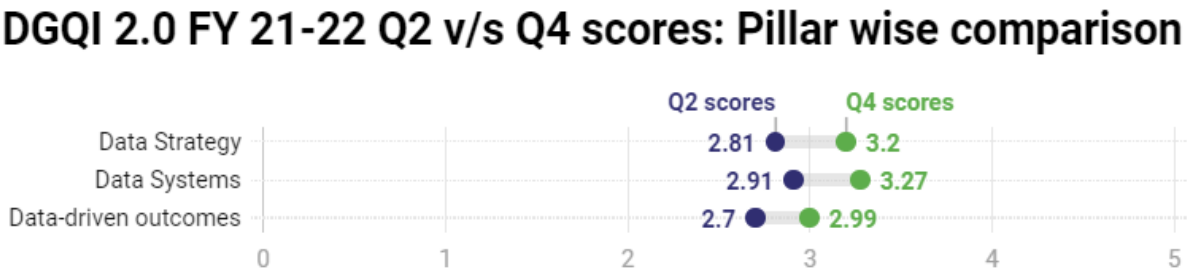


Figure 14 Pillar wise comparison of DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22)

Each DGQI pillar is further composed of several themes which measure the data maturity of M/Ds on varied important aspects.

Figure 15 shows the change in the theme-wise scores of M/Ds between Quarter 2 and Quarter 4 of FY 21-22.

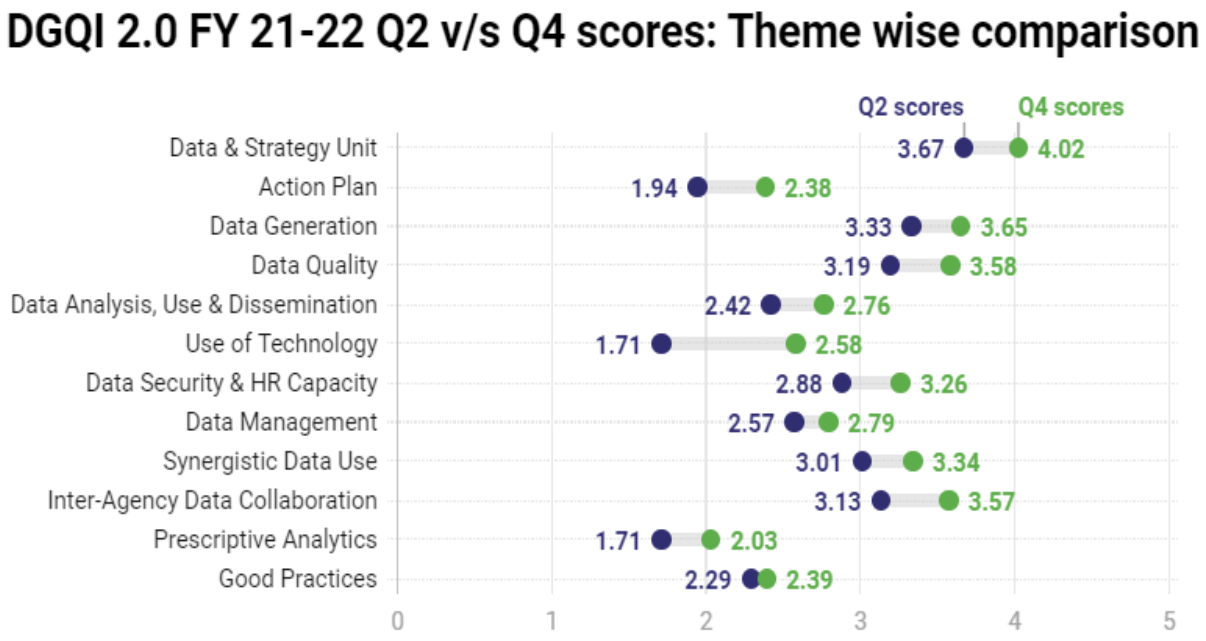


Figure 15 Theme wise comparison of DGQI 2.0 scores (Quarter 2 V/s Quarter 4 of FY 21-22)

Overall, M/Ds scored the highest on Data & Strategy Unit (DSU) theme indicating that most of them have established well functional DSUs to improve the quality of their data systems and drive better data exchange and use at M/Ds. However, M/Ds also registered third lowest score on action plan theme highlighting that there is significant scope for M/Ds to improve upon the quality of their data strategies as well as ensuring compliance against the timelines documented in their action plans. M/Ds performed next best on data generation theme, as in the previous round, again suggesting that M/Ds have digitized their data collection mechanisms. However, M/Ds also demonstrated weakest performance on prescriptive analytics theme, highlighting the need to develop necessary analytical capacities at M/Ds to rightly use the vast amount of data being collected by them in administrative processes.

In terms of temporal comparison, it was positive to note that scores across all 12 themes improved between the two rounds, with an average improvement of 15%. It was encouraging to note that M/Ds reported maximum improvement in the use of technology theme, implying improving data linkages and steadily improving use of emerging technologies by M/Ds. However, there was only a marginal improvement in good practices scores. M/Ds may hence like to focus more on documentation of their good practices in data governance domain to further peer learning. To further encourage the process, some outstanding good practices have been documented and disseminated in the form a good practices compendium by DMEO, NITI Aayog in this round.

Observing theme wise scores for each category of M/Ds revealed interesting findings.

Within the data strategy pillar, there are two themes: Data & Strategy Unit and Action Plan. With respect to the Data & Strategy Unit theme, Infrastructure category M/Ds have obtained the highest score and hence other categories can learn from this category for establishing well functional DSUs. Admin category M/Ds did not show any improvement in their DSU scores between Q2 and Q4 of FY 2021-22. Moreover, concerningly, Scientific and Economic category M/Ds observed a reduction in their average DSU scores.

On the Action Plan theme, Admin and Strategic category M/Ds again showed maximum scope for improvement, like the previous quarter. Moreover, Action Plan scores of admin and category also decreased from last round. Similarly, while Scientific category M/Ds outperformed other categories in general, this category also observed a reduction in Action Plan scores. It is hence crucial for all categories to focus on timely compliance against their action plan to continue positive movement towards the frontier.

Q2 Category wise scores of themes under Data Strategy Pillar

DGQI 2.0 Score color coding

0-1 1-2 2-3 3-4 4-5

Category	Data & Strategy Unit score	Action plan score
Admin	2.44	0.70
Economic	4.31	1.96
Infrastructure	4.68	2.23
Scientific	4.17	3.11
Social	3.53	1.92
Strategic	1.81	1.38

Figure 16 Category wise scores of themes under Data Strategy pillar (Quarter 2 of FY 21-22)

Q4 Category wise scores of themes under Data Strategy Pillar

DGQI 2.0 Score color coding

0-1 1-2 2-3 3-4 4-5

Categories	Data & Strategy Unit score	Action Plan score
Admin	2.44	0.47
Economic	4.18	2.36
Infrastructure	4.82	2.53
Scientific	4.03	2.31
Social	4.33	2.85
Strategic	2.62	2.02

Figure 17 Category wise scores of themes under Data Strategy pillar (Quarter 4 of FY 21-22)

Within data systems, M/Ds scored highest on data generation and quality, while most categories showed maximum scope for improvement in data analysis, use & dissemination theme. Infrastructure and Scientific category M/Ds again outperformed while Admin and Strategic category lagged behind across themes (Figure 19).

Temporally, all themes within data systems pillar showed improvement across all categories, which was quite appreciable. Specifically, all categories were found to be leveraging technology much more actively as compared to the last round.

Q2 Category wise scores of themes under Data Systems Pillar

DGQI 2.0 Score color coding

0-1 1-2 2-3 3-4 4-5

Category	Data generation score	Data quality score	Data analysis score	Use of technology score	Data security & HR capacity score	Data management score
Admin	2.32	2.14	1.35	0.74	1.55	2.10
Economic	3.06	2.96	2.16	1.58	2.70	2.42
Infrastructure	3.88	3.75	3.10	1.96	3.44	2.88
Scientific	4.04	4.04	3.96	2.82	4.15	3.34
Social	3.30	3.20	2.14	1.67	2.64	2.42
Strategic	2.96	2.36	2.20	0.92	2.73	2.60

Figure 18 Category wise scores of themes under Data Systems pillar (Quarter 2 of FY 21-22)

Q4 Category wise scores of themes under Data Systems Pillar

DGQI 2.0 Score color coding

0-1 1-2 2-3 3-4 4-5

Categories	Data generation score	Data quality score	Data analysis score	Use of technology score	Data security & HR capacity score	Data management score
Admin	2.94	3.56	1.90	2.27	2.46	2.22
Economic	3.39	3.28	2.51	2.56	3.04	2.47
Infrastructure	4.04	3.92	3.41	2.71	3.78	3.05
Scientific	4.26	4.27	3.94	4.03	4.53	3.15
Social	3.61	3.63	2.56	2.37	3.04	2.84
Strategic	3.49	2.92	2.53	1.99	3.14	2.81

Figure 19 Category wise scores of themes under Data Systems pillar (Quarter 4 of FY 21-22)

Under data-driven outcomes pillar, all categories witnessed an improvement under synergistic data use and inter-agency collaboration themes. However, under prescriptive analytics, Scientific category observed a reduction and Strategic category remained stagnant. Further, under good practices theme, Admin and Economic categories observed a downward trend in their performance and hence need to focus on this theme. As of Quarter 4, Infrastructure and Scientific categories were again found to be the frontrunners across themes and other categories can learn from these M/Ds. Admin and Strategic categories again showed maximum scope for improvement on all themes under this pillar.

Q2 Category wise scores of themes under Data-driven Outcomes pillar

DGQI 2.0 Score color coding

0-1 1-2 2-3 3-4 4-5

Category	Synergistic data use score	Interagency collaboration score	Prescriptive analytics score	Good practices score
Admin	0.36	1.60	0.30	0.34
Economic	3.28	2.94	1.84	2.45
Infrastructure	3.57	3.88	2.04	2.89
Scientific	3.94	4.14	2.89	2.88
Social	3.12	3.29	1.70	2.26
Strategic	2.22	2.06	0.89	1.85

Figure 20 Category wise scores of themes under Data Driven Outcomes pillar (Quarter 2 of FY 21-22)

Q4 Category wise scores of themes under Data-Driven Outcomes Pillar

DGQI 2.0 Score color coding

0-1 1-2 2-3 3-4 4-5

Categories	Synergistic data use score	Inter-Agency Collaboration score	Prescriptive Analytics score	Good Practices score
Admin	1.12	2.20	0.30	0.28
Economic	3.58	3.34	2.02	2.26
Infrastructure	4.14	4.46	2.79	3.21
Scientific	4.43	4.14	2.36	3.03
Social	3.30	3.52	2.32	2.38
Strategic	2.22	3.17	0.89	2.17

Figure 21 Category wise scores of themes under Data Driven Outcomes pillar (Quarter 4 of FY 21-22)

5.4. Performance on key themes

- **Data & Strategy Unit and Action Plan**

Figure 22 shows the current status of formation of Data & Strategy Units (DSUs) at M/Ds. It was positive to note that 91% (67/74) Ministries/Departments have already set up DSUs. Another crucial component of the data strategy pillar is the formation of detailed action plans by M/Ds to have exhaustive strategy documents to guide their data governance efforts. 85% (63/74) Ministries/Departments reported to have developed these plans.

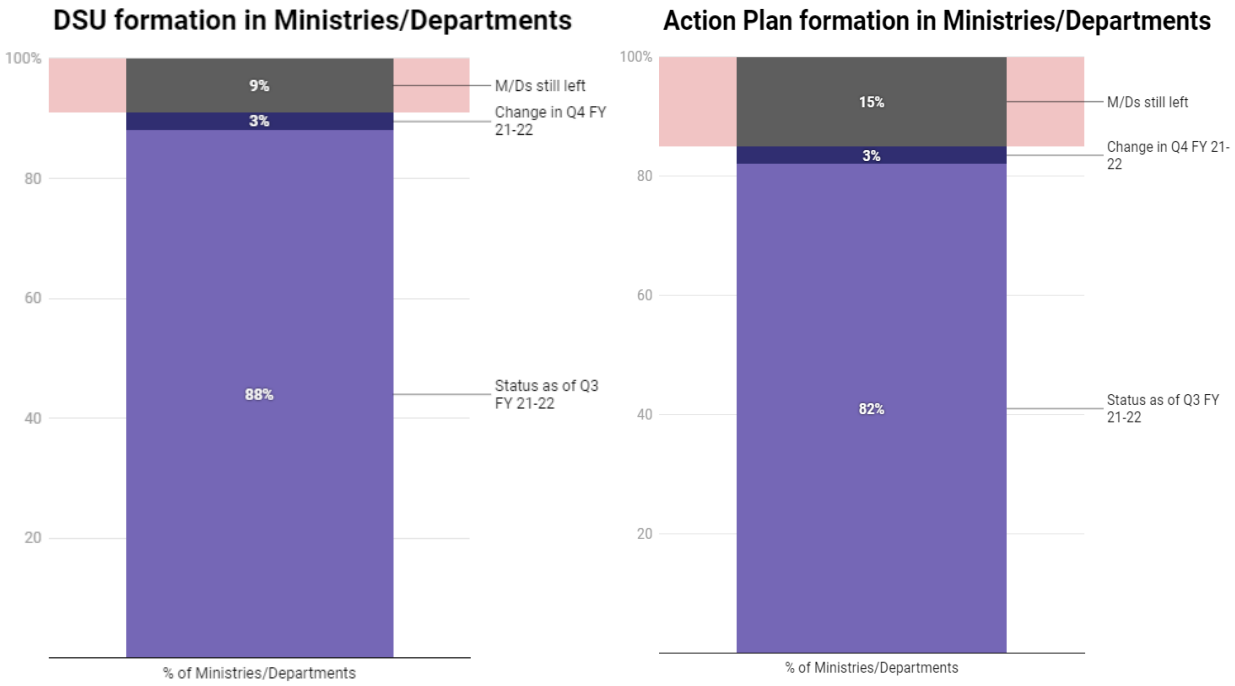


Figure 22 Status of DSU and Action Plan formation

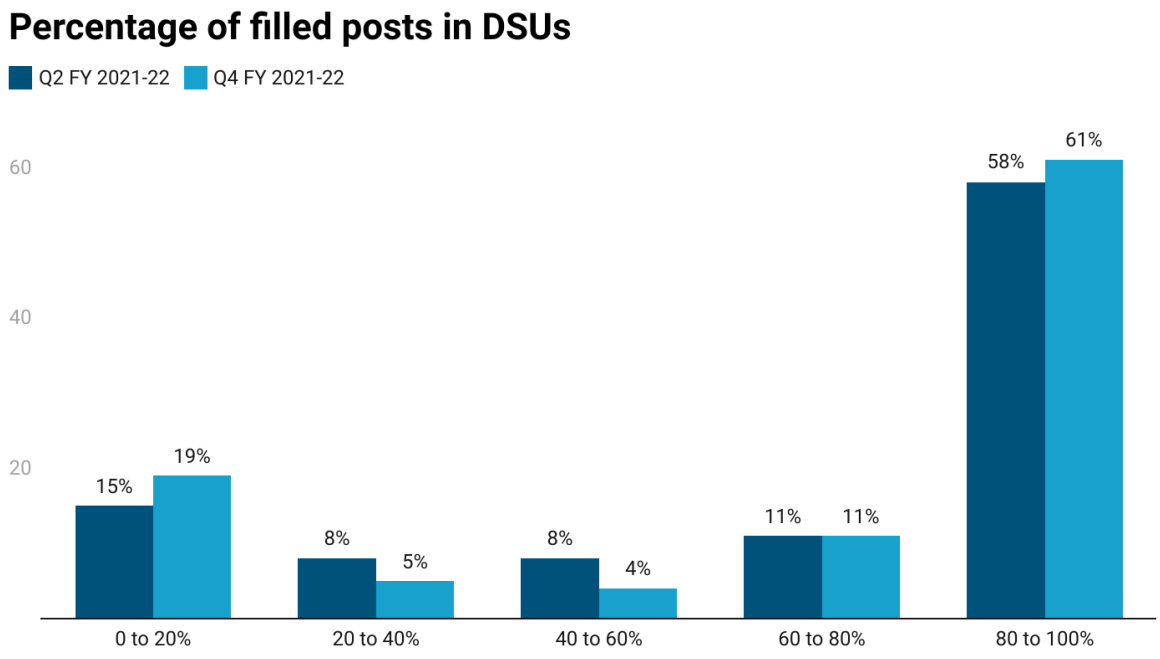


Figure 23 Percentage of filled posts in DSUs (Quarter 2 V/s Quarter 4 of FY 2021-22)

However, only over half of the DSUs were fully staffed and hence there is a lot of scope to augment the capacities of DSUs. There was also an increase in the number of M/Ds who have only filled up 0-20% of their DSU posts, implying an increase in vacant positions in DSUs (Figure 23). Moreover, only 10 out of 74 M/Ds were able to timely completely the action points that were due according to their own action plans, highlighting the need to focus on the implementation of already developed data strategies.

- **Digitization of schemes**

As of Quarter 4 FY 21-22, 76% (497/650) schemes and non-schematic interventions were reported to have been digitized (i.e., they had a Management Information System (MIS)) which was a slight improvement over last quarter (70%).

However, among digitized schemes/interventions, 24% (121/497) still don't collect data at unit level. Out of these 121 schemes/interventions, 56% (68/121) reported that they need not collect unit level data because of the nature of these schemes/interventions. However, remaining 44% (53/121) still have scope to improve their granularity up to unit level. More than 75% schemes also reported to be collecting data at real-time/near real time/daily frequency.

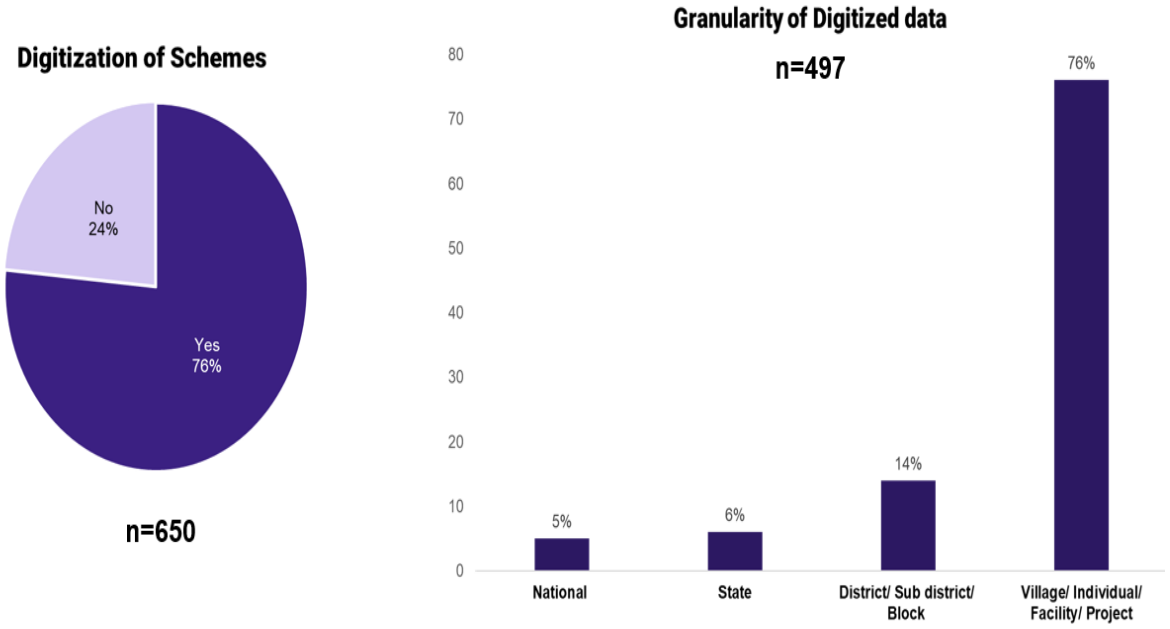


Figure 24 Status of digitization of schemes/interventions along with granularity of digitization

Even among digitized interventions, 28% of the interventions/ schemes reported to first collect data on paper which was later fed into digital systems, creating huge scope for data entry errors. This aspect saw limited change when compared to the last round, when again one-third of the interventions/ schemes first collected data on paper.

- **Data Quality Assessment**

82% schemes/interventions were reported to have pre-defined documented mechanisms to assess data quality, which was an increase of 8 percentage points over the last round (74%). It was also noted that over 10% of the interventions didn't follow any of the prescribed data quality protocols, almost same as last round. Moreover, only 60% followed all recommended protocols, which is a slight improvement from last round when the ratio was approximately 50%.

Within data quality, accurately compiling and updating metadata is one of the most crucial protocols. 74% interventions reported availability of accurate metadata, in comparison to 65% in the previous round. Data integrity measures to ensure accurate reproduction of data at various locations and points of time are also equally important. 76% interventions reported to have such protocols in place, in comparison to 71% in the previous round.

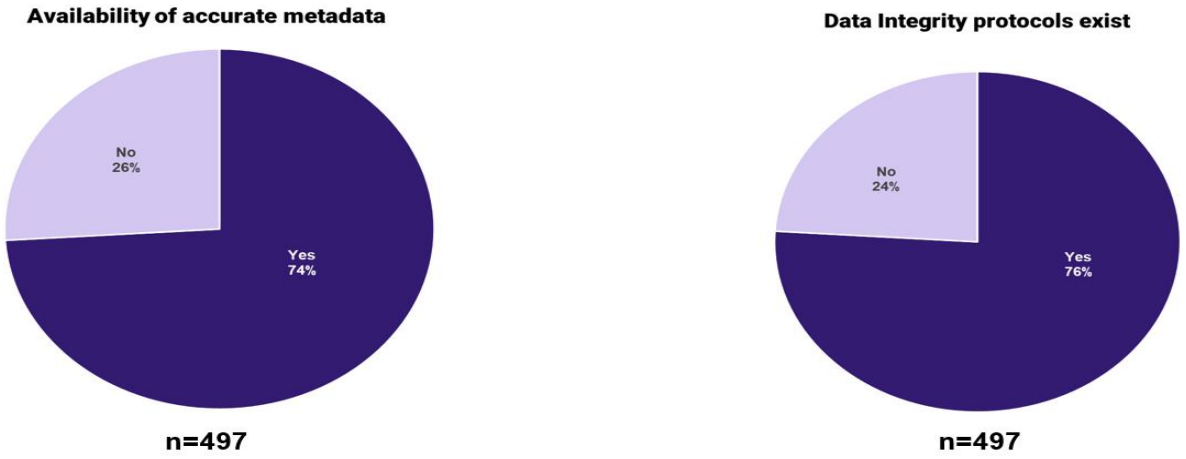


Figure 25 Status of metadata compilation and data integrity practices by M/Ds

However, only 68% schemes/interventions reported to have dedicated data quality assessment and management teams to do the needful. Remaining divisions need to focus on forming such teams to ensure high quality data can be used for policymaking purposes.

- **Data Analysis**

23% schemes reported to not undertake any type of data analysis, which slightly improved from previous round (26%), However, most M/Ds were found to be still concentrating on descriptive analytics (Figure 26), which only involves basic cross tabulation and frequency analysis. It was found that there is a need for M/Ds to graduate towards causal and predictive analytics, which was presently used for just one-third of the schemes/interventions. This is crucial to not only understand what has happened in the past, but also delineate reasons behind the same and predict future risks and opportunities. Also, dedicated data analysis teams were found to be functional only for 65% of the schemes/interventions – again highlighting a capacity gap to enable better data use, as was found in data quality assessment.

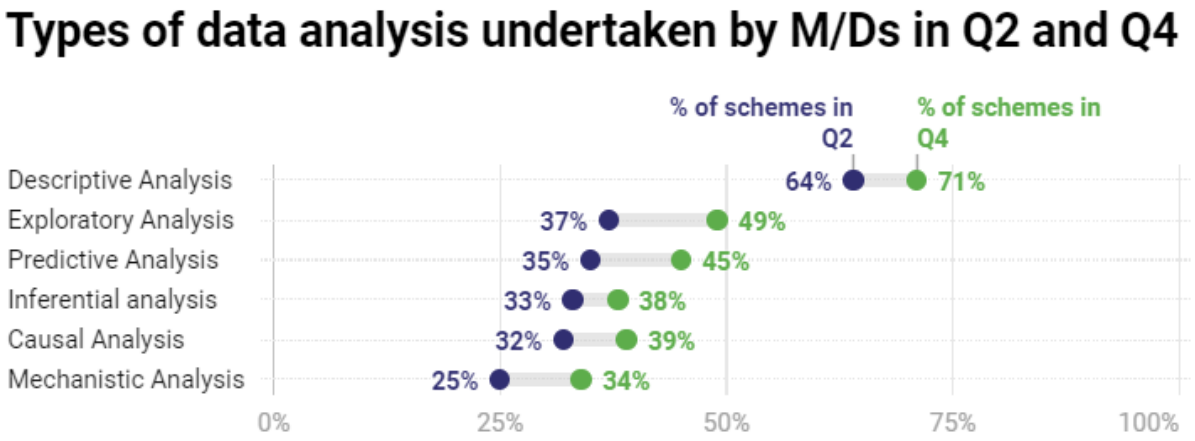


Figure 26 Types of data analysis undertaken by M/Ds (Quarter 2 V/s Quarter 4 of FY 21-22)

- **Use of Technology**

More than 40% of scheme/intervention management information systems (MIS) were found to not yet be linked with Public Financial Management System (PFMS), highlighting restrictions in using data exchange for enhanced accountability and transparency in the utilization of public expenditure. This linkage with PFMS had slightly improved by 5 percentage points in comparison to Quarter 2 of FY 21-22.

For beneficiary-oriented schemes, MIS linkages with Aadhaar, bank accounts and mobile numbers was again found to be incomplete: over 80% MIS have been linked with mobile numbers, however, only 59% have been linked with Aadhar. 226 schemes reported that they can link their MIS with all three systems, however, only 51% (116/226) have established these linkages with all three systems in Quarter 4 as compared to 39% in Quarter 2, suggesting there is still a long way to go. Moreover, there were 40 schemes which don't have linkages with any of these systems, in urgent need of action.

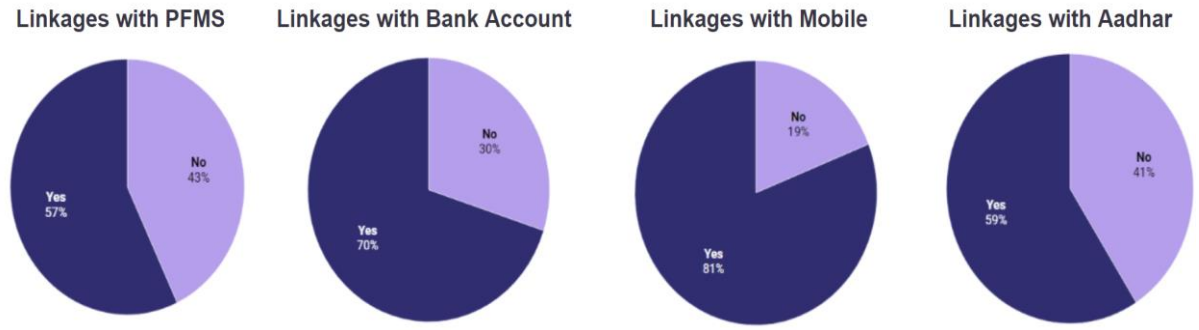


Figure 27 Status of linkages of MIS with PFMS, Aadhaar, Mobile Numbers and Bank Account Numbers

Moreover, linkages with LGD were also found to be present in over 78% (228/311) of MIS which is an appreciable improvement from 64% (185/286) in the previous round.

42% of the schemes also reported to not use any alternative source of data (even when applicable) and over 50% of the schemes didn't make use of any of the emerging technologies (even when applicable). While this figure has reduced from 75% in the

previous round, there is still a long way to go in streamlining use of technology for improving scheme implementation and monitoring.

- **Synergistic use of data within M/Ds**

As of Quarter 4 of FY 21-22, 66% (49/74) M/Ds reported to have identified data gaps as the first step to start working towards synergistic data use among their divisions. Out of these, 88% (43/49) have also developed implementation plans to overcome these gaps. However, over 18% M/Ds were found to be still struggling with developing synergistic data use systems among their scheme divisions (Figure 28).

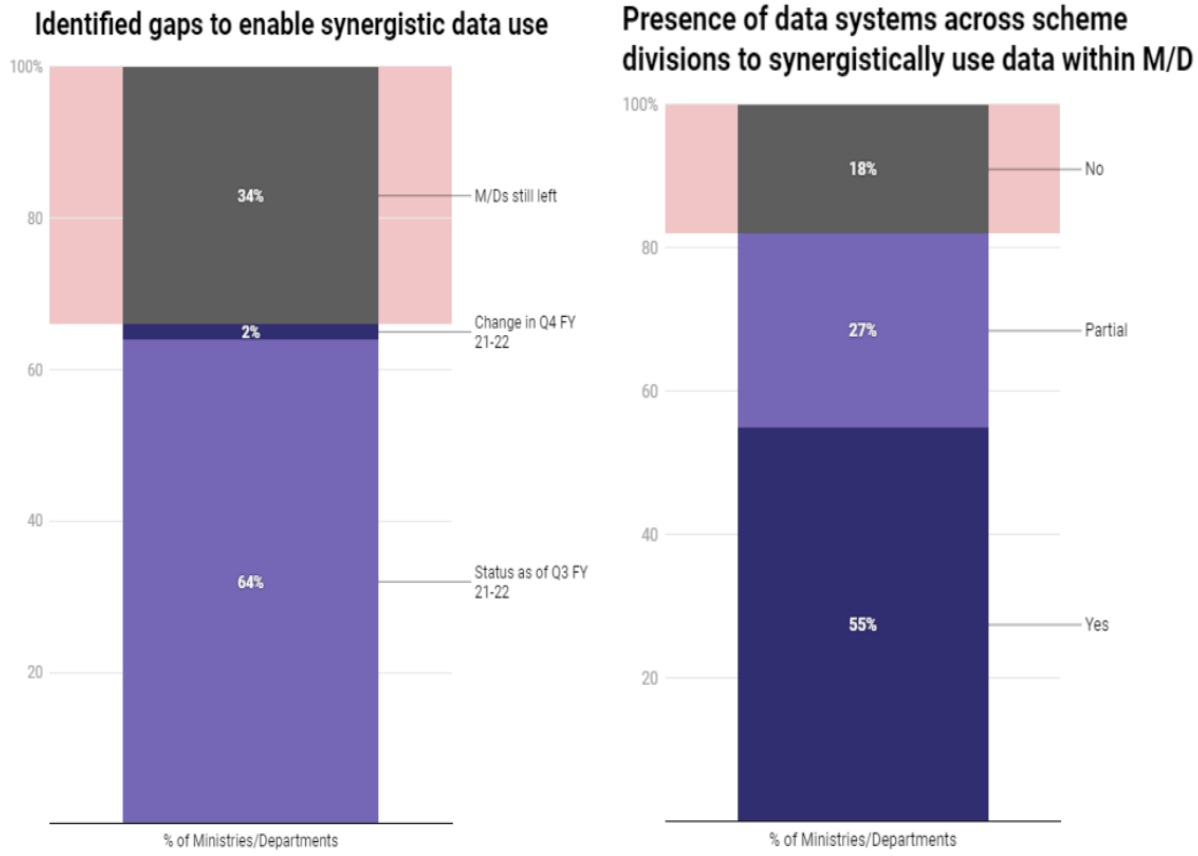


Figure 28 Status of synergistic use of data by Ministries/Departments

6 Summary of Findings

Some key findings from the current round of DGQI exercise are summarized below:

- Average DGQI scores have steadily improved from 2.29 in DGQI 1.0 to 3.20 in DGQI 2.0 2nd Round, highlighting the success of the initiative in creating increased awareness around data governance at M/Ds.
- As some M/Ds grew faster than others, it was also noted that there were several M/Ds which improved their DGQI scores, however, their relative ranking remained the same or fell down. Going forward, it would hence be crucial for M/Ds to focus on their areas of improvement and undertake reforms to improve their DGQI scores in a timely manner.
- For the first time, majority of the M/Ds (~35%) lied in the highest quintile of scores i.e., between 4 to 5, which is quite appreciable. However, there is still much scope for improvement for all M/Ds to move to the frontier as there is still another heavy concentration of M/Ds (~50%) around scores between 2 to 4.
- Similar trend was observed across categories as category wise average scores also ranged between 2 to 4. Similar to last round of the exercise, Scientific and Infrastructure category M/Ds outperformed on the Index, while Admin and Strategic M/Ds demonstrated maximum scope for improvement. Further, Social category M/Ds were spread out across a range of score – while some performed very well, others demonstrated a huge scope for improvement – suggesting a lot of scope for peer learning among these M/Ds.
- Among the three pillars, there was again minimum variation with average pillar wise scores ranging between 2.99 to 3.27. On average, pillar wise scores increased by 12%, when compared to the previous round. However, data driven outcomes was identified as the pillar in need of maximum focus as its score has been consistently lower across both rounds.
- In addition to improvements in scores at pillar level, average scores also increased across all 12 themes of DGQI with an average improvement of 15%. M/Ds performed the best on Data & Strategy Unit, suggesting that they have set up well-functional DSUs. However, they also reported a huge scope for improvement on action plan theme suggesting that while DSUs have been established, much remains to be done to ensure that they develop good quality data strategies and comply against the same. Similarly, while M/Ds performed next best on data generation theme suggesting that most interventions have been digitized, use of this data to conduct prescriptive analytics came out to be a theme with maximum scope for improvement across categories.
- As of end of Quarter 4 of FY 2021-22, 91% M/Ds reported that they have already set up DSUs and 85% have also developed action plans/data strategies. However, only half of the DSUs have been fully staffed and hence there is a lot of scope to augment the capacities of DSUs. Moreover, only 10 out of 74 M/Ds were able to timely

completely the action points that were due according to the action plans they had outlined for themselves, highlighting the need to focus on the implementation of already developed data strategies now.

- As of Quarter 4 FY 21-22, 76% (497/650) schemes and non-schematic interventions were reported to have been digitized (i.e., they had a Management Information System (MIS)) which was a slight improvement over last quarter (70%). However, remaining 24% of the interventions still relied on paper-based data and need to be urgently digitized. Even among digitized interventions, 28% of the interventions/ schemes reported to first collect data on paper which was later fed into digital systems, creating huge scope for data entry errors. While it was encouraging to note that over 75% of the digitized schemes/interventions collect data at high granularities (village/individual/facility/project) and/or high frequencies (near real time/real time/daily), a significant gap needs to be bridged by remaining schemes.
- Within data quality, 10% of the interventions were found to not follow any of the prescribed data quality protocols. Moreover, only 60% followed all recommended protocols. There is hence a strong need to focus on ensuring data quality.
- With respect to data analysis, it was found that 23% of the schemes don't undertake any type of data analysis. Remaining M/Ds mostly relied only on descriptive analysis and need to explore enhanced forms of data analysis.
- Regarding linkages of MIS with other platforms, 40% of schemes were found to not yet be linked with PFMS systems. For beneficiary-oriented schemes, only 51% schemes were found to have established linkages with all three pillars of JAM trinity: bank accounts, Aadhaar and mobile numbers; as compared to 39% in the previous round. So, substantial progress has been made on this account, but there is still capacity for development.
- Regarding use of technology, 42% of the schemes reported to not use any alternative source of data (even when applicable) and over 50% of the schemes didn't make use of any of the emerging technologies (even when applicable). While this figure was way higher at 75% in the previous round and has substantially reduced since then, there is still a long way to go in streamlining use of technology for improving scheme implementation and monitoring.
- Finally, with respect to synergistic use of data by M/Ds, which is one of the most crucial data driven outcomes, 55% of the M/Ds reported to have developed data exchange mechanisms among their divisions. However, 27% of the M/Ds were still establishing mechanisms for the same and another 18% have not yet started the process.
- M/Ds need adequate capacities to enable this transformational shift towards better data driven outcomes, however, it was noted that over one-third of the schemes/interventions don't possess data quality assessment and analysis teams, highlighting a key area of concern, which needs to be urgently addressed.

- There is hence a huge scope for improvement across the board to realise the paradigm shift in India's digital growth story.

7 Conclusion

Based on the comprehensive review of data strategies, systems and outcomes across M/Ds, it is clear that M/Ds have focused on data governance in last two years, resulting in steadily improving performance on the DGQI across the three rounds of the exercise. However, concerted efforts still need to be undertaken to move all M/Ds to frontier DGQI scores in the short run and foster a culture of evidence-based policy making in the Government of India in the long run.

Some of the immediate areas across that need urgent attention include:

- Under data strategy, majority of the M/Ds have set up Data & Strategy Units (DSU) and prepared action plans. However, going forward, it is pertinent to ensure that these **DSUs remain functional by filling up remaining vacant positions.** Secondly, it is important to focus on the quality of the action plans. Several action plans need to be made more detailed with granular milestones, timelines and responsibility mapping. It is also crucial to continue **finetuning action plans according to the findings of this report by identifying theme-wise areas for improvement and mapping their accountabilities to different units.** For instance, areas for improvement under use of technology theme may be mapped to Technology Unit. **Subsequently, timely implementation of these action plans requires constant monitoring and internal review by DSUs.**

- Under data systems, **gaps in digitization need to be urgently addressed** given the way governance needs are continuously evolving across the world. **M/Ds must focus on ensuring 100% digital administrative data at highest possible granularity and frequency** for all their interventions. During digitization, adequate focus must be paid on **reducing human interference in data collection**, ensuring regular data quality assessment, maintaining and updating metadata, having robust quality backcheck mechanisms and use technologies such as geotagging of information to **improve the quality of data right at entry stage.**

- Given that all M/Ds are mandated to use logical frameworks for their schemes at scheme design stage **as part of EFC/SFC proposals** and data sources and data points are an integral component of these frameworks, **it is crucial that M/Ds clearly lays down its adherence against various DGQI standards**: whether it has digital administrative data for the scheme or not, whether it collects data on all necessary indicators or not, the granularity and frequency of data collection, details on how they ensure quality of collected data and how is this data digitally disseminated in open domain as well as with other M/Ds.

- Data analysis and its use for policymaking is one of the most important themes within data systems to **ensure that meaningful data is collected by M/Ds which is not just used for reporting purposes but also utilized for evidence-based policymaking**. M/Ds need to increasingly focus on this theme and **graduate from data reporting to descriptive analytics and further to advanced forms of causal and predictive data analytics** on urgent basis. They should look at using

this data not only to improve implementation and monitoring but also to review and improve policies as well as **linking outcome performance data to budgetary decisions**.

- Finally, under data driven outcomes, there is huge scope for improvement across all themes. M/D DSUs must focus on **developing automated mechanisms for data exchange among divisions within the M/Ds, identify channels for data exchange with other relevant M/Ds (wherever and whenever required) as well as using this cross-sectoral data for high-end prescriptive analytics**. To enable this, M/Ds may start with compiling an exhaustive inventory of their own current databases, establishing data lakes/warehouses to allow easy exchange among these self-owned databases, identifying other datapoints and data sources they need for better policymaking and then partnering with their owner agencies to share data with each other via digital modes such as APIs, common data lakes etc.

- Throughout this cycle of data generation and use, **M/Ds need to ensure a fine balance between opening up data exchange and safeguarding personal data against data security and privacy risks.** They must ensure compliance against all data security and privacy norms for government websites as well as adopt globally accepted best practices as new developments occur in the space at a rapid pace. **Framing of data management guidelines by M/Ds would prove highly helpful in ensuring standard and robust ways of data handling across its lifecycle by all data operators including compliance against such safeguarding practices.**

- To specifically address the limited use of emerging technologies and alternative data sources, **M/Ds may like to use public-private partnerships to identify and pilot use cases for scaling up, adopt globally accepted good practices as well as build their capacities in this direction**.

- To ensure all of the above, **M/Ds need to develop necessary data analytical capacities to increase the demand for data** – to ensure that officials can understand the importance of data and have the skills to be able to utilize it for day-to-day decision making. Hence, **M/Ds must develop mechanisms to ensure that DSUs remain functional via regular review mechanisms at the highest level.** While DMEO, NITI Aayog is building a curriculum to build data governance capacities at an overall level**, M/Ds may also start developing annual training calendar on data governance and analytics for policymaking. Public Private Partnerships would be extremely crucial to devise and implement this capacity development plan**. Technology-based partnerships of M/Ds with domain leaders to enhance own data maturity levels and achieve data driven outcomes would be extremely critical, especially today, when industry is much more advanced in use of data than the government.

These measures can effectively pave the way for a culture of evidence-based policymaking in India, as envisaged under the DGQI exercise.

Annexure 1: Indicative outline of action plan to be prepared by M/Ds

Better data preparedness would help improve the monitoring and consequently the outcomes of the schemes and interventions of Ministries/Departments. An indicative outline of roadmap to improve data preparedness and improve DGQI scores is given below. Ministries /Departments may use the indicative outline to have discussions to develop a roadmap to achieve higher, systemic, ministry-wide data preparedness levels and make provisions for human and financial resources for using technology and analytics to improve service delivery.

1. **Background**

   1.1. Brief overview of the M/D's business allocation, roles and responsibilities *(in 2-3 lines)*

   1.2. Current degree of digitization of administrative data systems in the M/D *(in brief)*

      1.2.1. Scheme-wise initiatives for digitization (for CS/CSS schemes)

      1.2.2. Other initiatives for digitization (Other Central Sector Schemes (OCS)/Other MIS/Dashboards)

   1.3. Ministry/Department's reflections on DGQI Scores 2020 *(in one page)*

      1.3.1. Major takeaways from the exercise

      1.3.2. Areas for improvement based on previous performance and DGQI methodology

      1.3.3. Limitations in achieving DGQI frontier scores (Dependence on states or other executing agencies for execution, human resource/financial constraints etc.)

2. **Vision, Mission & Objectives**

   2.1. Vision statement for achieving data driven decision making within the Ministry/ Department *(in 2-3 lines)*

   2.2. Mission statement for achieving DGQI frontier scores and going beyond DGQI by 2022 *(in 4-5 lines)*

   2.3. Objectives of the roadmap to achieve DGQI frontier scores

   (In half page – explaining Ministry/Department specific goals to achieve high levels of data preparedness; for instance, ensuring end-to-end digitization for high-quality, near real-time data generation across all schemes at project/beneficiary level, ensuring user-friendly MIS and dashboard systems for all scheme and non-schematic interventions, establishing an administrative system for human capability and technological development to enable data driven policy making etc.)

3. **Strategy to achieve DGQI Frontier Scores**

   3.1. Scope of the strategy *(in 1-2 pages)*

      3.1.1. Schemes to be covered under the road map along with their contribution to overall M/D scheme budget (Encouraged to include all CS/CSS schemes

of the M/D including schemes executed by other partnering agencies/states/PSUs aligned with the M/D)

3.1.2. Non-schematic interventions to be covered under the roadmap (Other MIS/Dashboards of the M/D which are not related to schemes; for instance, a sector-level MIS/dashboard used for monitoring the overall sector performance, separate MIS/dashboards for PSUs/Other Central Expenditure/any other purposes. Kindly note that administrative interventions for digitization within the office such as E-Office is not to be included here as it is outside the scope of the strategy. This strategy aims to implement digitization to improve monitoring and accountability of government expenditure on schemes and policies.)

3.2. Overall Approach *(in 1-2 pages)*

3.2.1. Principles to be followed while developing the roadmap (For instance, accuracy in information, relevance/utility to the strategy, transparency in processes, privacy of personal information, openness in disseminating non-personal information, inclusiveness in digitization, interoperability, integration of uses, etc.)

3.2.2. Integrated approach (Outline of an integrated and well-coordinated approach to be taken by the M/D to improve digitization across the board. The approach should target end-to-end digitization of all levels of information – Scheme level MIS/Dashboards, M/D Sector level and finally linking it to digitization of necessary information needed for achieving SDG goals/national priorities relevant to the M/D. Similarly, how data collection frequency, quality and timeliness-at-entry will be ensured on the field and during subsequent stages of data flow at the district and national levels. Also, the approach should focus on across the board interventions – Capacity development at M/D, technological overhaul at M/D, coordination between various divisions of the M/D, setting up of administrative systems at M/D to lead the effort, carrot-stick approaches to improve uptake etc.)

3.3. Scheme-wise Strategy *(2-3 pages per scheme)*

3.3.1. Scheme 1 (A short assessment of current system to be provided along with areas identified for improvement. Subsequently, the strategy should entail detailed steps to improve on each theme of the DGQI as shown below.)

3.3.1.1. Data Generation Strategy (Should cover steps for identifying data requirements of the scheme to have data on all relevant inputs, outputs and outcomes of the scheme; increasing granularity (beneficiary/project level) and frequency (near real-time) of digitization using latest sources of information; use of location tracking devices for data collection; using GIS mapping/geo-coding/geo-fencing/mobile devices for data generation)

3.3.1.2. Data Quality Strategy (Should cover steps for ensuring rigorous data quality protocols for profiling/filtering incoming data, ensuring deduplication and redundancy removal within data,

enforcement of data integrity, use of metadata standards for proper classification of data; use of mobile phones or other technologies for data quality control such as multimedia evidence, telephonic surveys etc.)

3.3.1.3. Use of technology Strategy (Should cover steps for linking M/D MIS/data systems with other platforms such as PFMS for finances and JAM trinity for beneficiary-oriented schemes; use of alternative data sources to complement M/D data such as private sector or GIS data; use of emerging technologies to improve scheme processes/delivery such as Machine Learning, Artificial Intelligence, IoT etc.)

3.3.1.4. Data Analysis, Use & Dissemination Strategy (Should cover steps for improving use of data by M/D to use it for policy making purposes; dissemination of data via websites/dashboards/social media/mobile apps; user-friendly visualizations; multilingual interfaces and compatibility features for differently abled etc.)

3.3.1.5. Data Security & HR Capacity Strategy (Should cover steps for improving data security, compliance requirements and privacy; capacity development for developing data analytics capabilities in the M/D to improve use of data in policymaking etc.)

3.3.1.6. Data Management Strategy (should cover steps for managing data across various stages right from generation to its use; devising strategies for integrated data storage and data disposal; ways and means of dealing with personal data using techniques like encryption, de-identification, etc., ensuring proper data classification using good-quality meta data to enable better reporting, analytics, and use; fixing accountability for data management by fixing intra-ministry and inter-ministry data ownership and other responsibilities for dissemination and use of data)

3.3.2. Scheme 2 and so on.. (Strategy for each scheme under the purview as per section 3.1 to be framed and the strategy should entail detailed steps to improve on each theme of the DGQI as shown in section 3.3.1.)

3.4. Non-schematic Strategy *(2-3 pages per intervention)*

3.4.1. Intervention 1 (A short description of the purpose and scope of the intervention to be provided with areas identified for improvement. Subsequently, the strategy should entail detailed steps to improvise on each theme of the DGQI as shown in section 3.3.1.)

3.4.2. Intervention 2 and so on..(Strategy for each intervention under the purview as per section 3.1 to be framed and the strategy should entail detailed steps to improvise on each theme of the DGQI as shown in section 3.3.1.)

3.5. Operational Execution Plan *(After strategy is formed, execution plan to be laid down for institutional development).*

3.5.1. *Organizational Structure – Breaking the silos (To have a central unit leading the efforts to build, implement and revise the roadmap, it is recommended that a Data and Strategy Unit is established within the M/D and is placed directly under the Secretary. After setting up the unit, strategy for intra-ministerial coordination to be framed to ensure that the unit is able to work in conjunction with other scheme divisions and NIC.)*

3.5.2. *Human Resource Capacity Development (Should include steps for in-house capacity building to develop IT and data analytical capabilities, acquaint them with new tools/techniques, hire technical experts as per requirements if necessary, spread awareness about evidence-based policy making etc.)*

3.5.3. *Technological Development (Should include steps for overhaul of IT hardware and software systems in line with identified data generation, storage, management, and analytical needs including a procurement plan, development of data warehouses/ open data websites to create integrable data sources, creation of singular metadata standard/data classification norms to be followed across the M/D to create integrable datasets etc.)*

3.5.4. *Partnerships (Should include the nature of partnerships being planned with private sector or research organizations for developing capabilities, scope the possible partner landscape and areas of engagement, inter-ministerial coordination for synergies in data collection on common indicators, state-level engagements to help build adequate data systems at state level including CSS schemes)*

3.5.5. *Resource Allocation (Should include assessment of required financial resources to implement the roadmap and plans to make provisions for the same in scheme and M/D budget in the next EFC/SFC/Budget cycle; assessment of human resources to be deployed to implement the roadmap and provisions for the same; any other resources)*

3.6. Consolidated roadmap *(Consolidated plan to be provided for all schemes and interventions listed in Section 3.3. and 3.4. as well as steps to be undertaken for institutional development in Section 3.5* ***with quarterly timelines against key strategy steps****)*

Annexure 2: Detailed Terms of Reference (ToR) for Data & Strategy Unit (DSU) at M/Ds

Purpose of DSU

In order to create better mechanisms for digitization of processes related to implementation and monitoring of Central Sector/Centrally Sponsored Schemes and other non-schematic interventions of Ministries/Departments, an institutional mechanism in the form of a **"Data and Strategy Unit"** may be set up within each Ministry/Department. The Data and Strategy Unit shall drive the process of building and harnessing existing as well as augmenting the monitoring, statistical, technological and data analytics capabilities of the respective Ministry/ Department.

The key roles of the DSU shall include breaking silos within the Ministry/Department to enable creation of well-integrated monitoring and data systems while ensuring adequate focus on data quality and security and creating mechanisms for regular data analysis within the Ministry/Department to inform policy decisions. Coordinating with scheme divisions within the Ministry/Department as well as with required external partners such as States, other Ministries/Departments, research organizations, leading private players and academic institutions for taking necessary steps in the direction shall also be one of their key responsibilities.

Organization Structure of DSU

To fulfill this purpose, The Ministries/Departments can augment their present institutional setup to create DSU.

The DSU may be headed by an Additional Secretary/Joint Secretary/DDG level officer who would be directly reporting to the Secretary of the Ministry/Department. As shown below, it is proposed to have the following four verticals within the DSU

1. Monitoring Unit – For integrating siloed monitoring initiatives across the Ministry/Department
2. Statistics Unit – For identifying overall statistical needs of the Ministry/Department and ensuring coordination with necessary agencies to meet the same
3. Technology Unit – For ensuring 100% digitization and integrating siloed MIS/dashboards/data systems of the Ministry/Department
4. Data Analytics Unit – For undertaking and promoting data analysis on collected data to drive decisions

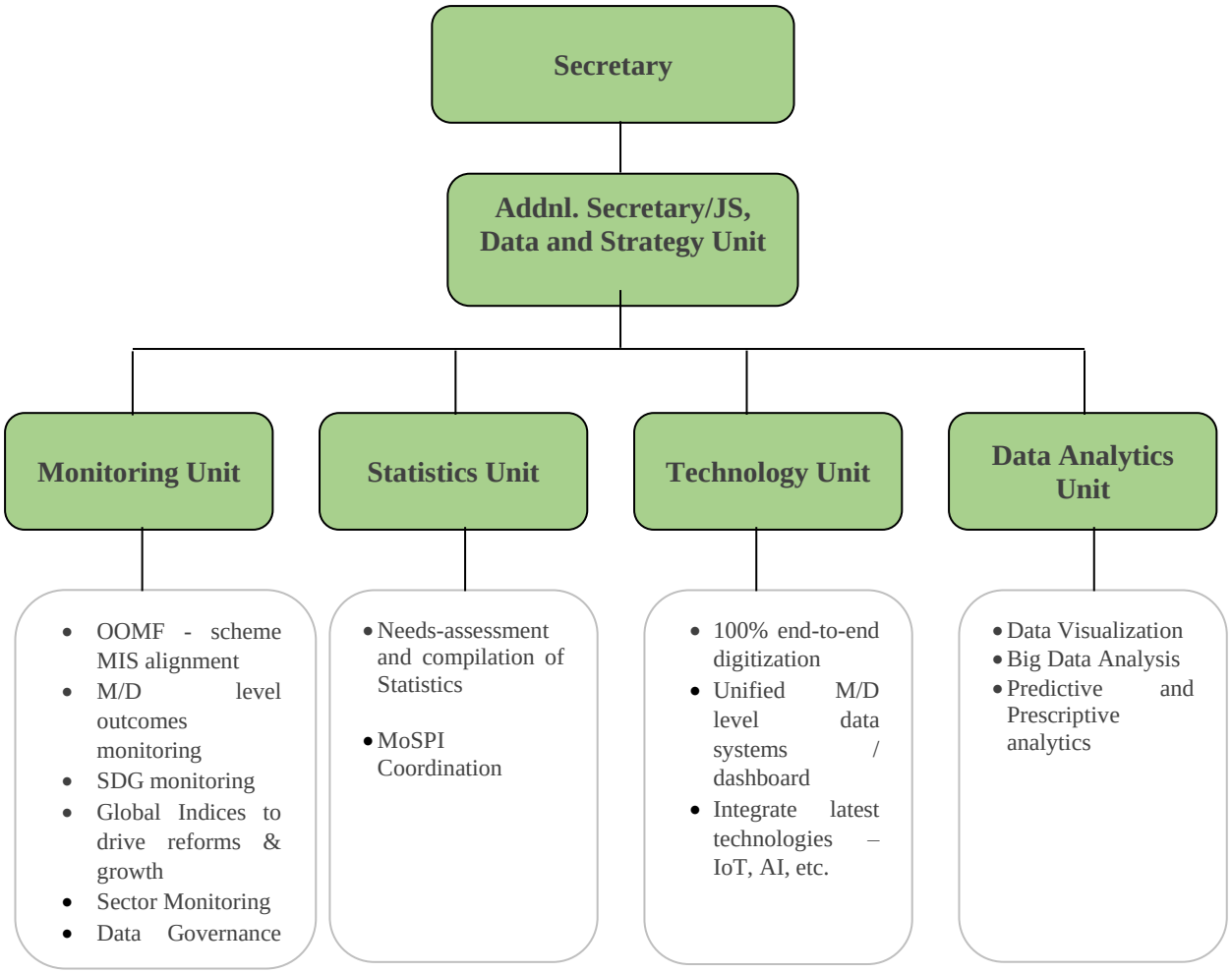


Figure 29 Structure of Data & Strategy Unit

A snapshot of the organization structure of the four verticals of the DSU has been provided below in next figure.

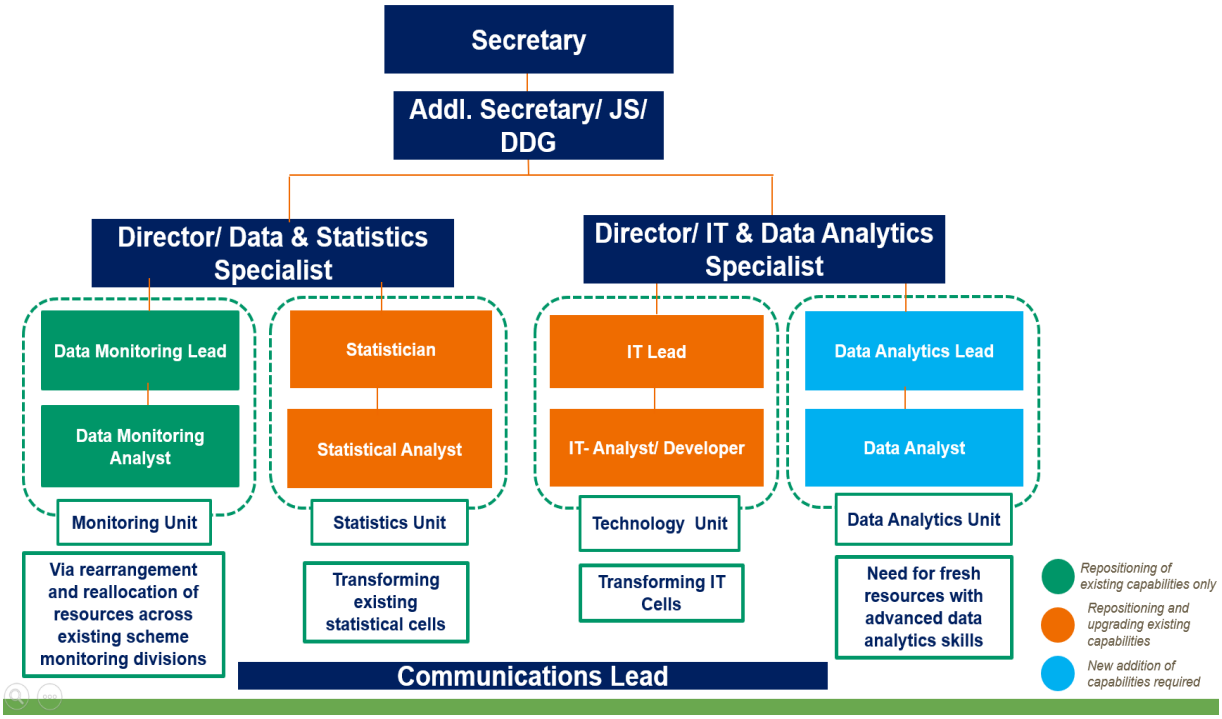


Figure 30 Organization of Data & Strategy Unit

The sub-units within the DSU should be headed by a Director level officer. In order to make the DSU lean and more responsive, it is suggested that sub-units with similar tasks may be headed by one Director level officer i.e., one Director for Monitoring and Statistics sub-units and one Director for Technology and Data Analytic sub-units.

Each sub-unit is recommended to be composed of leads who would be domain experts and provide direction to the efforts to be undertaken and skilled analysts who would have domain knowledge and be responsible for carrying out the implementation efforts.

It is envisioned and suggested that majority of these positions in these sub-units can be filled by repositioning of existing staff as explained below.

It is suggested that the Monitoring sub-unit may be almost entirely created by reallocating some of the existing staff among different scheme divisions which are already regularly monitoring scheme implementation to DSU. Similarly, current resources engaged in OOMF, Global Indices and SDG monitoring in Ministries/Departments can be integrated to form the Monitoring sub-unit of the DSU. One, this would aid in ensuring that resources from different scheme divisions work in tandem with each other, helping in breaking the siloes as they get to understand how different schemes may be monitoring similar initiatives. Two, with different types of monitoring needs (OOMF, GI, SDG etc.) being looked at by a central team, the Ministry/Department would be able to understand how these efforts can actually be synergized, reducing the administrative reporting burden on Ministries/Departments.

Similarly, Statistics sub-unit may also be largely created by reallocating roles in the existing statistical cells to the DSU. The statistical cells at present are already responsible for dissemination and compilation of Ministry/Department statistics. To be able to meet

newer expectations from the Statistics sub-unit of the DSU such as needs assessment for more statistics that need/need not to be collected and coordination with other partners (private agencies/Ministries/others) for synergistically collecting some statistics, their capabilities may require some upgradation. In order to transform existing statistical cells for this purpose, every Ministry/Department may undertake an assessment at their own level and consider the need for a few additional resources with required skills based on their present status.

The Technology sub-unit may similarly be created by transforming existing IT cells present in the Ministry/Department. Some of the resources can be re-casted to play the role of the technology arm of the DSU. However, if it is felt that there may be a need for upgradation of their capabilities in certain domains such as integration of several dashboards or creating single metadata architecture for all scheme MIS/dashboards, then the Ministry/Department can again undertake an assessment at their own level and consider the need for a few additional resources with required skills based on their present status. For e.g. If the existing IT cell in the department does not have a suitable person with required skills of IT analyst, he/she may have to be recruited as a lateral entrant.

The Data Analytics sub-unit is one sub-unit where it is believed that most Ministries/Departments may not have enough capabilities or resources at present and hence may require people with advanced data science and analytics skills to be freshly recruited to complete this missing link. This would play a crucial role in completing the vision of moving to evidence-based policymaking across Ministries/Departments.

To summarize, it is suggested that the creation of DSU may be principled on reorganization of existing organizational structure and roles of Ministries/Departments. The key idea should be to bring together existing resources with skills, experience and passion for these tasks together within the DSU to break the siloes of the present structure, and thereafter, only for required roles, recruitment may be done to fill the skill gaps, wherever necessary. The same is also depicted above in above figure where monitoring sub-unit is colored green to show that it only requires repositioning, statistics and technology sub-units are colored amber to show that they majorly need repositioning but some Ministries/Departments may also require upgradation and finally, data analytics unit is colored blue to show that it is the major place where fresh talent acquisition may be needed.

Indicative strength of DSU

This section highlights the indicative strength of manpower that may be required for the 'Data and Strategy Unit'. As already mentioned above, the unit is proposed to be headed by an Additional Secretary/ Joint Secretary level officer of the M/D, reporting directly to the Secretary. A Director level officer called Data & Statistics Specialist can head the two sub-units, Monitoring Unit and Statistics Unit. Another Director level officer called IT & Data Analytics Specialist can head the remaining two sub-units, Technology Unit and Data Analytics Unit.

For leads and analysts, indicative strength of manpower required in the Data & Strategy Unit has been arrived at in below table. The M/Ds under have been classified into three main categories: Small, Medium and Large on the basis of the number of interventions (CS schemes +CSS schemes +Non-Schematic Interventions of the M/D). M/Ds with upto 10 interventions are called small, 11-30 interventions are termed medium and M/Ds with above 30 interventions are termed large. Further, two bifurcations have been created on the basis of average outlay of M/D's interventions (for the size of interventions), depending on whether it is above or below INR 500 Crores. The same has been done keeping in mind that manpower requirements would rise in line with a greater number of interventions or increase in average outlay.

Using this classification and the following general thumb rules –

- For every **5 interventions** with average budgetary allocation **less than Rs 500 crores**, **one analyst is recommended**. M/D may hire/ allocate **one lead for every two analysts** in a sub-unit to guide and review the tasks assigned to them.
- For every **5 interventions** with average budgetary allocation **more than Rs 500 crores**, **two analysts are recommended**. M/D may hire/ allocate **one lead for every two analysts** in a sub-unit to guide and review the tasks assigned to them.

The indicative manpower strength has been arrived at in Table 1. However, it may be noted that this is only meant to act as a guidance for Ministries/Departments is by no means a mandatory requirement.

Table 3: Indicative Strength of DSU

Type of M/D	Total number of interventions (CS+CSS+NSI)	Avg outlay > = INR 500 Crores			Avg outlay < INR 500 Crores		
		Leads	Analysts	Total	Leads	Analysts	Total
Small	0-10	2	4	**6**	1	2	**3**
Medium	11-30	4	8	**12**	2	4	**6**
Large	Above 30	6	12	**18**	3	6	**9**

**Individual M/Ds may modify the numbers as per different combinations/ categories and beneficiary coverage*

Annexure 3: DGQI 2.0 Self-Assessment Questionnaire

Detailed explanations for each question may be referred to after the questionnaire.

Minor changes made in the questionnaire in this round, to address M/D concerns in previous rounds, have been highlighted in yellow.

Part -A (To be fed at Ministry/ Department Level)

A. Background Information

1.	Ministry / Department(M/D) Name:	
2.	Name of the Central Sector (CS) Schemes of the M/D:	a. b. c.
3.	Name of Centrally Sponsored Schemes (CSS) of the M/D:	a. b. c.
4.	Please enter any other non-schematic intervention (NSI) to be included for DGQI self-assessment:	a. b.
5.	Details of the nodal officer responsible for verifying authenticity of information provided in this form:	
	a. Name	
	b. Designation:	
	c. E-Mail ID:	

B. Data & Strategy Unit

1.	Has the M/D constituted a Data & Strategy Unit (DSU) as a central unit for developing data strategy? (As per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM dated 02.02.2021)	☐ Yes ☐ No
2.	*(Respond if answer to 1 is 'yes', else skip to Q1 of next section)* Who is the head of the DSU?	☐ AS and equivalent ☐ JS and equivalent ☐ Director and equivalent ☐ Below Director
3.	Please select the verticals established under the DSU of your Ministry/Department. (As per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM dated 02.02.2021)	☐ Monitoring Unit ☐ Statistics Unit ☐ Technology Unit ☐ Analytics Unit

4.	Please provide the percentage of filled posts in DSU (number of posts filled up/ number of posts created by the Ministry/Department for the DSU) in the below provided table: **Enter % of posts filled up** Monitoring Unit Statistics Unit Technology Unit Analytics Unit Total	
5.	Are the terms of reference (ToR) for all units within DSU well defined and documented by the M/D to lay down their scope of work?	☐ Yes ☐ No ☐ Partial (to be selected if ToR development is in progress)
6.	Have any regular review meeting mechanisms at the level of the head of DSU and/or the Secretary been established for regular review of the work undertaken by the DSU?	☐ Yes ☐ No
7.	*(Respond if answer to 6 is yes, else skip this question)* What is the frequency of regular review meetings/review reports?	☐ Daily ☐ Weekly ☐ Fortnightly ☐ Monthly ☐ Quarterly ☐ Annually

C. Action Plan

1.	Has the M/D framed an action plan to improve its data preparedness levels? (As per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM dated 02.02.2021)	☐ Yes ☐ No

2.	*(Respond if answer to 1 is yes, else skip to Q1 of next section)* Does the action plan have all the sections as per the outline shared with all M/Ds? (As per D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM on 02.02.2021)	☐ Yes ☐ No ☐ Partially. If partially , please specify how: _____
3.	Does the action plan include data strategy for all CS/CSS schemes of the M/D?	☐ Yes ☐ No ☐ Partially (Some schemes included) If partially, please specify which schemes are not included: _____
4.	Are clear timelines for each action point identified under the strategy?	☐ Yes ☐ No ☐ Partially (For some actions) If partially, please specify how and why: _____
5.	Are the responsibilities for each action point clearly allocated to respective divisions for ensuring accountability?	☐ Yes ☐ No ☐ Partially (For some actions). If partially, please specify how and why: _____
6.	Please upload the action plan in PDF format.	
7.	Please enter the action points in the attached excel template. Scores based on exhaustiveness (number of action points per scheme) and timely completion/compliance on the action points against the timelines set by the M/D will get auto-calculated and displayed here.	

D. Data Management

1.	Are there data management guidelines/architecture, explaining how generated data is to be processed, stored, exchanged, archived and destroyed?	☐ Yes ☐ No If yes, please briefly explain the scope implementation of these guidelines: _____
2.	*(Respond if answer to in 1 is 'yes', else skip this question)* Is there a dedicated senior-level officer responsible to check the compliance of the data management processes?	☐ Yes ☐ No
3.	Are data ownership norms clearly defined?	☐ Yes

		☐ No
4.	Is there a framework for assessing the risk and value of all the data collected by the M/D?	☐ Yes ☐ No If yes, please explain how is this done: _____
5.	Is there a framework governing the ethical use of data, including the use of predictive algorithms, machine learning etc. by the M/D?	☐ Yes ☐ No If yes, please explain how is this done: _____

Note: M/Ds may preferably fill up remaining sections of Part- A (given below from E-H) after completing Part – B of the questionnaire as these questions correspond to the third pillar of data driven outcomes.

E. Synergistic data use within the M/D

1.	Based on data analysis, has the M/D identified data gaps at M/D level that need to be plugged in from decision making/policy analysis perspectives?	☐ Yes ☐ No If yes, please specify how: _____
2.	*(Respond if answer to 1 is yes, else skip this question)* Has the M/D made any implementation plan to overcome these data gaps to aid in decision making?	☐ Yes ☐ No If yes, please specify how: _____
3.	Has the M/D created any systems for ensuring that data systems across scheme divisions are integrated so that data from different scheme divisions is shared with each other?	☐ Yes. ☐ No ☐ In progress ☐ N/A If yes or in progress, please specify how: _____ If "N/A", please provide reasons why inter schematic division data integration is not applicable: _____

F. Inter-Agency Data Collaboration

1.	Has the M/D collaborated with other agencies (other M/Ds, private agencies, research organizations etc.) for improving their data systems wherever possible?	☐ Yes ☐ No ☐In progress

2.	*(Respond if answer to 1 is yes, else skip this question)* Has the M/D undertaken any of the following steps to drive these inter-agency data collaboration initiatives?	
	☐ SoI, MoU, Partnerships with agencies	
	☐ API linking of MIS/Dashboards done to enable seamless data sharing between M/Ds	
	☐ Multiple data collection processes aimed at same target groups replaced by single synergistic process	
	☐ Integrated data storage/warehouses	
	☐ Collaboration with other M/Ds to use their data for developing own systems	
	☐ Collaboration with M/Ds to develop joint systems for data gathering/use of non-conventional data sources/emerging technologies	
	☐ Collaboration with private agencies for use of non-conventional data sources or emerging technologies	
	☐ Jointly conducting analysis using data from multiple M/Ds	
	☐ Partnerships/Collaborations for data security related measures	
	☐ Partnerships/Collaborations for capacity building of human resources	
	☐ Others - Please specify : _____	

G. Prescriptive Analytics

1.	Has the M/D gone beyond exploratory data analysis to cross-functional prescriptive analytics?	☐ Yes ☐ In Progress ☐ No If yes or in progress, please specify how: _____
2.	*(Respond if answer to 1 is yes, else skip this question)* How often is this being undertaken?	☐ Annually ☐ Quarterly ☐ Monthly
3.	*(Respond if answer to 1 is yes, else skip this question)* What is the mode in which this is being practiced? (Multiselect)	
	☐ Mechanisms for regular prescriptive data analysis reports to be prepared and shared with decision makers at the highest level have been instated	
	☐ Committee formed to hold policy review meetings/review reports at regular frequencies	
	☐ Regular policy review meetings involving all scheme divisions/sections institutionalized	
	☐ Emerging actionables are undertaken, documented and disseminated via a separate newsletter/report/document/order etc. and tracked regularly	
	☐ Others - Please specify how: _____	

H. Good Practices - Please share any three good practices of how the M/D has taken measures to strengthen data-driven decision-making (non-schematic or scheme level) within the M/D along with its positive impact.

Good Practice 1
1a. Describe the problem statement faced by the M/D. (100 words)
1b. Describe how the M/D has used and implemented data systems and analytics to address the issue to drive smart, near real-time and granular decisions (100 words).
1c. Explain the positive impact generated with supporting evidence that indicated such impact due to the solution implemented (100 words).
Good Practice 2
1a. Describe the problem statement faced by the M/D. (100 words)
1b. Describe how the M/D has used and implemented data systems and analytics to address the issue to drive smart, near real-time and granular decisions (100 words).
1c. Explain the positive impact generated with supporting evidence that indicated such impact due to the solution implemented (100 words).
Good Practice 3
1a. Describe the problem statement faced by the M/D. (100 words)
1b. Describe how the M/D has used and implemented data systems and analytics to address the issue to drive smart, near real-time and granular decisions (100 words).
1c. Explain the positive impact generated with supporting evidence that indicated such impact due to the solution implemented (100 words).

Part -B (To be fed at CS/CSS/NSI Level)

To be fed by the Ministry/Department for each CS/CSS/NSI of the Department in Q.A of Part A of the Questionnaire

A. Data Generation

1.	Are the data requirements of the scheme well defined and documented?	☐ Yes ☐ No
2.	Is data collected for all identified data requirements?	
	a. Input Data Points	☐ Yes ☐No ☐Partial
	b. Output Data Points	☐ Yes ☐No ☐ Partial
	c. Outcome Data Points	☐ Yes ☐No ☐Partial
3.	Is collected data reported digitally? (i.e., is there is a digital electronic database/MIS)?	☐ Yes ☐ No i.e. On paper only If Yes, please provide the link: _____ If credentials are required for login, please provide some username and password: User - _____ Pw - _____
4.	*(Respond if answer to 3 is 'Yes', else skip to 1 of Q1 of data quality section)* At what granularity is data reported digitally for the scheme?	
	a. At the M/D (National)	☐ Yes ☐ No ☐ NA
	b. State	☐ Yes ☐ No ☐ NA
	c. District / City	☐ Yes ☐ No ☐ NA
	d. Sub-District / Tehsil	☐ Yes ☐ No ☐ NA
	e. Block	☐ Yes ☐ No ☐ NA
	f. Village	☐ Yes ☐ No ☐ NA
	g. Individual / Household	☐ Yes ☐ No ☐ NA
	h. Facility	☐ Yes ☐ No ☐ NA
	i. Project	☐ Yes ☐ No ☐ NA
5.	At what frequency is data reported digitally for the scheme?	
	a. Realtime or near real time	☐ Yes ☐ No ☐ NA
	b. Daily	☐ Yes ☐ No ☐ NA
	c. Weekly/Fortnightly	☐ Yes ☐ No ☐ NA
	d. Monthly	☐ Yes ☐ No ☐ NA
	e. Quarterly	☐ Yes ☐ No ☐ NA
	f. Half-yearly	☐ Yes ☐ No ☐ NA
	g. Yearly	☐ Yes ☐ No ☐ NA
6.	How is this data collected at the ground level?	
	☐ Collected on paper by human resources and then fed on digital systems	
	☐ Collected using digital modes (tablets/phones etc.) by human resources	
	☐ Transactional data	

7.	*(Respond if answer to 6 is 'second/third option', else skip this question)* Are any of the following technologies used?	
	a. CAPI Surveys	☐ Yes ☐ No ☐ NA
	b. Geotagged information	☐ Yes ☐ No ☐ NA
	c. Geofenced information	☐ Yes ☐ No ☐ NA
	d. Others - Please specify which technology	_____

B. Data Quality

1.	Are there pre-defined documented mechanisms to assess quality of incoming data?	☐ Yes ☐ No
2.	How is data quality assessment done? *(If answer to 2 of data generation section is 'No', please select 'Manually')*	☐ Automatically ☐ Manually ☐ Hybrid ☐ Not done If Hybrid, please specify how: _____
3.	*(Respond if answer to 2 is not "not done", else skip to Q1 of next section)* Are following protocols followed during data quality assessment?	
	a. Incoming data is filtered/cleaned after checking for missing values, logical flaws in data, incorrect values etc.	☐ Yes ☐No
	b. Summary statistics of incoming data are generated and checked for errors/abnormalities	☐ Yes ☐No
	c. Existence and accuracy of metadata for all the scheme's data is periodically checked (Schema is well defined)	☐ Yes ☐No
	d. There is a system for identifying duplicate data and removing redundancies	☐ Yes ☐No
	e. There is a system to ensure data is accurate, consistent and traceable to origin/source, whenever it is reproduced by any agency (data integrity)	☐ Yes ☐No
4.	Are following feedback mechanisms/backchecks also leveraged for data quality control?	
	a. Social audits	☐ Yes ☐ No ☐ NA
	b. Telephonic backchecks/verification with beneficiaries	☐ Yes ☐ No ☐ NA
	c. Multimedia data – citizen voice, video, images as evidence	☐ Yes ☐ No ☐ NA
	d. Sample inspections based on data	☐ Yes ☐ No ☐ NA
	e. Third party data verification/ data audits	☐ Yes ☐ No ☐ NA

C. Data Analysis, Use and Dissemination

1.	What types of data analysis is undertaken on collected data?	
	a. Descriptive data analysis (e.g. basic cross tabulation, frequency distribution, mean, median etc.)	☐ Yes ☐ No

	b. Exploratory data analysis (e.g., correlation etc.)	☐ Yes ☐ No
	c. Inferential data analysis (Using a small sample of data to infer about a larger population)	☐ Yes ☐ No
	d. Predictive analysis (Using historical or current data to find patterns to make predictions about the future)	☐ Yes ☐ No
	e. Causal analysis (Looks at the cause and effect of relationships between variables, focused on finding the cause of a correlation)	☐ Yes ☐ No
	f. Mechanistic Analysis (Understand exact changes in variables that lead to other changes in other variables)	☐ Yes ☐ No
	g. Others - Please specify the name and the type of data analysis - _____	
2.	*(Respond if answer to any of the options in 1 is "yes", else skip to Q5)* Is cross-schematic/sectoral data also analysed, wherever needed?	☐ Yes ☐ No ☐ NA. If NA, please explain why: _____ _____
3.	How often is this data analysis well-documented (in reports/notes/publications)?	☐ Real-Time on a dashboard ☐ Quarterly ☐ Half-yearly ☐ Annually ☐ Never
4.	How often is this data analysis being used by the M/D officials for:	
	a. To re-design the schemes or activities undertaken under the scheme at the end of the tenure?	☐ Yes ☐ No
	b. To do mid-course corrections through design or implementation changes?	☐ Yes ☐ No
	c. To guide intra-scheme funding decisions like inter-state allocations, inter-component allocations, etc.?	☐ Yes ☐ No
	d. To guide inter-scheme budgetary allocations?	☐ Yes ☐ No
	e. To decide quarterly releases to implementing agencies?	☐ Yes ☐ No
	f. For fraud management and analysis	☐ Yes ☐ No
	g. Day to day delivery and monitoring of implementation/ performance of the scheme	☐ Yes ☐ No
5.	What other modes are used to disseminate the MIS/ paper-based data and related data analysis?	
	a. Dashboard	☐ Yes ☐ No ☐ NA
	b. Mobile App	☐ Yes ☐ No ☐ NA
	c. Social Media	☐ Yes ☐ No ☐ NA
	d. SMS	☐ Yes ☐ No ☐ NA
	e. Newspapers/ Magazines	☐ Yes ☐ No ☐ NA
	f. Outdoor media (signages/ billboards)	☐ Yes ☐ No ☐ NA

	g. Events	☐ Yes ☐ No ☐ NA
	h. TV/ Radio	☐ Yes ☐ No ☐ NA
	i. Others - Please mention the mode - _____	
6.	*(Respond if 'Yes' in 5a, else skip to 9)* What purposes are dashboards used for by the M/D officials?	
	a. Visual presentation of KPI/KRAs with drill-down capability to lowest level to gain total visibility	☐ Yes ☐ No
	b. Capturing trends over time and identifying preempt trends	☐ Yes ☐ No
	c. Measure efficiencies/inefficiencies in processes	☐ Yes ☐ No
	d. User friendly one stop access to multiple automated reports	☐ Yes ☐ No
7.	What types of Data Visualizations are used?	
	a. Bar chart/Histogram	☐ Yes ☐ No
	b. Pie charts	☐ Yes ☐ No
	c. Scatter plot	☐ Yes ☐ No
	d. Heat maps	☐ Yes ☐ No
	e. Treemaps	☐ Yes ☐ No
	f. Gantt chart	☐ Yes ☐No
	g. Specialized visualizations- Stripe graphics, streamgraph, etc.	☐ Yes ☐No
	h. Others - please mention data visualizations used - _____	
8.	Does the Dashboard visualize information on maps?	☐ Yes ☐ No ☐ NA
9.	*(Respond if 'Yes' in 3 of Data Generation section, else skip to Q6 of next section)* Does the MIS support multilingual features as per GIGW norms?	☐ Yes ☐ No ☐ Partially (some norms followed but not all)
10.	Does the MIS support features for differently abled as per GIGW norms?	☐ Yes ☐ No ☐ Partially (some norms followed but not all)
11.	How is the MIS data accessible for general population?	☐ Openly accessible without credentials ☐ Accessible through credentials ☐ Not accessible
12.	Is there an option on the MIS to download bulk data in excel, csv, dta files (machine readable formats)?	☐ Yes ☐ No ☐ Partial data download allowed
13.	Is the MIS data available on 'data.gov.in'?	☐ Yes ☐ No

D. Use of Technology

1.	*(Respond if 'Yes' in 3 of Data Generation section, else skip to Q4 and Q6)* Does the MIS of the scheme have linkages with PFMS?	☐ Yes ☐ No ☐ NA

2.	*(Respond if answer to 1 is yes)* Is PFMS integration completed till the field-level implementation agency?	☐ Yes ☐ No ☐ Partially
3.	Does the MIS of the scheme have linkages:	
	a. Aadhaar	☐ Yes ☐ No ☐ NA
	b. Mobile numbers	☐ Yes ☐ No ☐ NA
	c. Bank accounts	☐ Yes ☐ No ☐ NA
	d. GSTN	☐ Yes ☐ No ☐ NA
	e. Udyog Aadhaar	☐ Yes ☐ No ☐ NA
	f. Others – please specify	
4.	Does the scheme use any of the following:	
	a. Remote sensing data	☐ Yes ☐ No ☐ NA
	b. Night light data	☐ Yes ☐ No ☐ NA
	c. Social media data	☐ Yes ☐ No ☐ NA
	d. Private sector generated data	☐ Yes ☐ No ☐ NA
	e. Others – please specify	
5.	*(Respond if 'Yes' in 3 of Data Generation section, else skip to Q6)* Is the MIS compliant with Local Govt Directory (LGD)?	☐ Yes ☐ No ☐ NA
6.	Does the scheme apply/use any of the following:	
	a. Machine Learning	☐ Yes ☐ No ☐ NA
	b. Artificial Intelligence	☐ Yes ☐ No ☐ NA
	c. Blockchain	☐ Yes ☐ No ☐ NA
	d. Internet of Things (IoT)	☐ Yes ☐ No ☐ NA
	e. Big Data analytics	☐ Yes ☐ No ☐ NA
	f. Drones	☐ Yes ☐ No ☐ NA

E. Data Security and HR Capacity

1.	*(Respond if 'Yes' in 3 of Data Generation section, else skip to Q10)* Does the MIS follow regular antivirus updates?	☐ Yes ☐ No
2.	Is the MIS regularly assessed by third party auditors for the online security?	☐ Yes ☐ No
3.	Does the MIS/ website use SSL certificate?	☐ Yes ☐ No
4.	If "Yes" in previous question, is the SSL certificate at least 2048-bit SHA 256 encryption or higher?	☐ Yes ☐ No
5.	Does the MIS use firewall to secure access to data?	☐ Yes ☐ No
6.	All external communication/ 3rd party integration/ API integration for the MIS is done through encrypted channel?	☐ Yes ☐ No ☐ No external communication established
7.	What measures are undertaken to secure sensitive/personally identifiable information? (Multiselect)	
	☐ Single-factor/ Multi-factor authentication	
	☐ Access control list is maintained	
	☐ Data is encrypted	
	☐ Data is anonymized	

	☐ No such data	
8.	(If anonymization is selected in previous question) How do you protect de-identified data from re-identification risks?	
	☐ No efforts made	
	☐ Tighter security for databases that store anonymized information	
	☐ Implementation of Differential Privacy	
	☐ Generation of Synthetic Data that exhibits the statistical properties of the raw data, without allowing real individuals to be identified	
	☐ Others - provide details – _____	
9.	*(Respond if answer to 7 is any option other than "no such data", else skip to Q10)* Is permission taken from user to collect, store and use their personal data?	☐ Yes ☐ No
10.	Is there a dedicated data quality assessment and management team for the scheme?	☐ Yes ☐ No
11.	Is there a dedicated data analysis team for the scheme?	☐ Yes ☐ No

F. Data Management

Respond if you have answered 'Yes' in 3 of Data generation section, else skip this section.		
1.	Where is MIS data stored?	☐ On separate servers for different schemes (distributed storage) ☐ On central server which is used for all schemes
2.	(*Respond if first option is selected in 1, else skip this question)* Are there mechanisms in place which can enable data sharing with other scheme divisions?	☐ Yes ☐ No If yes, please explain how: _____
3.	How is MIS data stored?	☐ Physical servers ☐ Cloud Storage ☐ Hybrid servers ☐ Others
4.	*(Respond if "Cloud Storage is selected in 3, else skip this question)* Which cloud service is being used?	☐NIC/ Gov cloud- Meghraj ☐ Cloud Services directly from CSP (Cloud Service Provider) (empanelled by MeiTY) / Cloud services through System Integrators (SI) after Standardisation Testing and Quality Certification / Cloud services through Managed Service Provider (MSP) after Standardisation Testing and Quality Certification

		☐ Cloud Services from other CSPs which are not empanelled / from other MSPs or SIs which don't have Standardisation Testing and Quality Certification
5.	How is historical MIS data managed?	☐ Data is not backed up (i.e. it is destroyed) ☐ Data is backed up and data is archived ☐ Data history is well maintained including retention, destruction, and audit trail details

Detailed explanations to questions of DGQI Self-Assessment Questionnaire for Ministries/Departments of GoI (2021-22)

Part	Section	Question	Explanation
A	A	1	M/D name would be automatically filled up when the M/D logins using their credentials.
A	A	2	A pre-populated list of CS schemes of the M/D would be visible here.
A	A	3	A pre-populated list of CS schemes of the M/D would be visible here.
A	A	4	M/Ds to enter any other non-schematic intervention such as sector dashboards, sector level MIS, any other MIS/dashboards etc. that they would like to include for DGQI assessment using the self-assessment questionnaire.
A	A	5	M/Ds to enter details of DGQI nodal officer. He/she would be assumed to have verified the correctness and authenticity of the information filled in this self-assessment form.
A	B	1	Constitution refers to establishing the unit, hence, even if its staffing is ongoing, M/Ds can select 'yes' if they have established the admin structure of the unit and some members have been assigned to it.
A	B	4	M/Ds to undertake calculations at their end based on how many posts they have proposed to create for the DSU based on their requirement and how many of these posts have been filled up. Total will be auto-calculated using values entered in the table.
A	B	5	Documentation of terms of reference here refers to the documentation of detailed objectives, roles and responsibilities of the DSU specific to the M/D. Indicative ToR for guiding M/Ds was shared by DMEO earlier.
A	B	6	M/Ds to select yes if guidelines for a standard system for regularly scheduling review meetings (via OM etc.) has been issued.
A	C	1	M/Ds to select yes if they have completed preparation of exhaustive action plan to improve data preparedness levels of the M/D.
A	C	2	Action plan is to have 3 sections with all sub-sections: 1. Background, 2. Vision, Mission & Objectives, 3. Strategy – Scope, Overall approach, scheme wise strategy, non-schematic strategy, operational execution plan.
A	C	3	M/Ds to select yes if the action plan has separate action points for all CS/CSS schemes of the M/D (as per the list on this portal).
A	C	4	M/Ds to select yes if every action point has a corresponding mm/yy timeline by which it is aimed to be completed, clearly documented in the action plan.
A	C	5	M/Ds to select yes if every action point is mapped to unit/personnel within the M/D by whom it is expected to be completed, clearly documented in the action plan.

Part	Section	Question	Explanation
A	D	1	Data management guidelines/architecture explains how data is to be managed across its lifecycle, i.e., how is it to be collected, stored, processed? How will it be exchanged? What will be done with historical data?
A	D	3	Data ownership norms would define who would be the owner of data when data is shared with other divisions or M/Ds or in public.
A	D	4	Understanding the value of the data collected by the M/D from utility perspectives and comparing it to the associated data security and privacy risks to ensure there is a balance between the two.
A	D	5	Data ethics refers to systemizing, defending, and recommending concepts of right and wrong conduct in relation to data, particularly personal data. With use of machine learning and predictive algorithms, it becomes even more important to protect sensitive data.
A	E	1	Data gaps refer to data that is required by the M/D from decision making point of view, however, for some reasons, such data is not available with the M/D. M/Ds to select yes if they have identified such data gaps based on analysis of their current data.
A	E	2	After identification of data gaps, M/Ds must take reform actions to develop data capture mechanisms/exchange mechanisms to fill up data gaps. M/Ds to select yes if they have started planning these actions.
A	E	3	For schemes with similar target groups, data collection can be done together rather than separately. This is an example of integrated data systems for collection. Similarly, if one scheme is collecting data on some indicator which is required by another division on its portal, it should be able to get this data from the scheme division via suitable exchange systems. M/Ds to select yes if this is possible currently.
A	F	1	If M/Ds are collecting similar data or running similar interventions, data collaborations can be undertaken. If private sector has some useful data (let's say e-commerce or traffic data), data collaborations can be undertaken by M/Ds.
A	F	2	If some data collaboration has been undertaken, M/Ds to select how it has been done from the given options.
A	G	1	Prescriptive analytics is the final stage of analytical capabilities. While predictive analytics answers what, when and why something will happen, prescriptive analytics builds on this further by specifying what present actions need to be undertaken to achieve the predictions and how will these decisions affect /impact other outcomes. Therefore, it helps in taking advantage of a future opportunity or mitigating future risks. It can also improve the accuracy of predictions by continuously taking in new data to re-predict and re-prescribe.
A	G	2	M/Ds to select the frequency of prescriptive analytics.

Part	Section	Question	Explanation
A	G	3	M/Ds to select the modes/mechanisms by which they have institutionalized prescriptive analytics, to ensure it is continuously undertaken to inform policymaking, and not just undertaken on random basis.
A	H	1,2,3	M/Ds to enter good practices of how they have used data for policymaking and/or set up systems for institutionalizing data driven policymaking.
B	A	1	Data requirements refer to various input, output, and outcome data points/indicators that need to be monitored. They must be clearly documented for each scheme. M/Ds to select yes if this is done.
B	A	2	After gathering of data requirements, scheme division to select the indicators for which it is collecting data also. For e.g.: If scheme has multiple outcome indicators documented but the division is collecting data on only some of them due to various reasons, it must select Partial.
B	A	3	After data is collected, it must be collated and reported via paper or digitally through a MIS. Scheme divisions to accordingly choose Yes/No. Regarding credentials, this is optional. However, M/Ds are encouraged to create dummy login credentials for DMEO with view-only rights. This shall stay confidential and not be used outside the government for unintended purposes.
B	A	4	Scheme division to select all the granularities at which data is reported on the MIS. For e.g.: if a scheme MIS has district, state as well as national level data, scheme division to select all three options.
B	A	5	Scheme division to select the frequency at which data is updated on the MIS.
B	A	6	This question is to essentially understand if the data reported by the M/D on the MIS is "collected" by humans or is it transaction-based collection. If it is collected by human resources, is it directly collected using digital tablets/mobiles etc. or is it the case that it is first collected on paper and then fed on computers by someone else.
B	A	7	If data is collected using digital modes or it is transactional in nature, use of survey tools and/or geotagging can improve the data reliability. Scheme division to select yes if the same is done.
B	A	7a	Computer-assisted personal interviewing (CAPI) refers to survey data collection by an in-person interviewer (i.e., face-to-face interviewing) who uses a computer to administer the questionnaire to the respondent and captures the answers onto the computer.
B	A	7b	A geotagged photograph is a photograph which is associated with a geographic position by geotagging. Usually this is done by assigning at least a latitude and longitude to the image, and optionally altitude, compass bearing and other fields may also be included.

Part	Section	Question	Explanation
B	A	7c	Geofencing is a location-based service which triggers some pre-programmed action like a survey when a mobile device or RFID tag enters or exits a virtual geographical boundary.
B	B	1	Data quality protocols and mechanisms should be clearly documented by the scheme division. Scheme division to select yes if the same is done.
B	B	2	Data quality assessment of collected data against data quality protocols can be undertaken automatically by advanced digital systems, manually or using a hybrid of both manual and automated systems.
B	B	3	Question to assess which protocols are included and followed by the scheme division in its data quality assessment.
B	B	3a	This is the first step of data quality where collected data is cleaned by checking missing values, incorrect responses etc.
B	B	3b	Next step is to generate summary statistics of data (like mean, median, trends etc.) to check for outliers
B	B	3c	Another important step is to ensure metadata is properly defined. Metadata is data about data – containing details on variables covered in the data, their number of observations, summary statistics, units etc. This must also be regularly updated if new data is collected.
B	B	3d	Next important protocol is to check collected data for duplicate values (this duplication may be in old data or new data) and remove any such redundancies
B	B	3e	Finally, ensuring data integrity. This means that if collected data is being reflected anywhere (on the MIS, on any other portal etc.), it must be ensured that the accurate and recent most value is reflected everywhere. It should not be the case that at one place, data is updated as of last month, but at other portal, it is updated as of last year or showing inaccurate value due to some error.
B	B	4	Apart from data qual assessment, backchecks may be deployed to further improve data quality and increase its reliability.
B	B	4a	Social audit is a form of citizen participation that focuses on government performance and accountability. If social audits are being used to improve scheme data, select yes.
B	B	4b	If telephonic backchecks are undertaken based on collected data to verify that data is correctly collected, select yes. E.g.: Based on PDS beneficiary data available on MIS, random sample of ppl are contacted on phone to validate data entries made on MIS.
B	B	4c	If there are provisions for citizens to submit multimedia evidence which is then used to improve the quality of data, select yes. For e.g.: People submitting photos of quality of roads built near their locations and this feedback data being used to reflect the quality of roads on MIS.
B	B	4d	Based on reported data on MIS, random inspections are made by MD officials to verify data on ground.
B	B	4e	Getting data on MIS verified/audited by third parties.

Part	Section	Question	Explanation
B	C	1	Scheme divisions to select all types of data analysis undertaken by them.
B	C	2	Apart from scheme data, if data from other schemes or sector level data is also used to complement scheme data for analysis purpose, select yes.
B	C	3	Data analysis must be documented in some manner. Select the frequency at which this is done.
B	C	4	M/Ds to select the uses for which data analysis is done.
B	C	5	Select different modes used for disseminating data and its analysis.
B	C	5a	DB is essentially a tool to display key KPIs from data and important analytics through interesting visualizations.
B	C	5b	Mobile apps can be used to share data with citizens and interact with them.
B	C	5c	Social media outlets can be used to share data with citizens and interact with them.
B	C	5d	SMS are often used to send details to users/beneficiaries with respect to the scheme activities.
B	C	5e	Such mass communication methods may also be used to share data with citizens.
B	C	6	M/Ds to select the purposes for which dashboard are being used by them.
B	C	9	This is important to ensure data is accessible to all.
B	C	10	This is important to ensure data is accessible to all.
B	C	11	All MIS may not be in public domain. Hence, scheme divisions to enter details on how can public in general can access MIS data.
B	C	12	Scheme divisions to check if there is an option to download all MIS data in machine readable formats by users on the MIS and accordingly select.
B	C	13	As per NDSAP, all non-personal data should be available on data.gov.in to facilitate easy access to all govt. data at one place. Scheme division to select yes if non-personal data of their MIS is available on this platform.
B	D	1	MIS linkage with PFMS means that latest status of funds being routed through PFMS should be linked with MIS of the scheme.
B	D	2	The field-level implementation agency is the last agency to which funds are to flow. For e.g.: if PFMS integration is done till state implementing agency level but fund flow below states is not PFMS integrated for a scheme where projects are implemented by city level agencies, integration is not completed till last mile.
B	D	3a	Applicable for beneficiary-oriented schemes
B	D	3b	Applicable for beneficiary-oriented schemes
B	D	3c	Applicable for beneficiary-oriented schemes
B	D	3d	Applicable for industry/firm oriented schemes
B	D	3e	Applicable for industry/firm oriented schemes
B	D	4a	Remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted day-time radiation at a distance

Part	Section	Question	Explanation
			(typically from satellite or aircraft). Special cameras collect remotely sensed images, which help researchers "sense" things about the Earth. For e.g.: large forest fires can be mapped from space, tracking clouds to help predict the weather or watching erupting volcanoes, and help watching for dust storms, tracking the growth of a city etc.
B	D	4b	Night-light data is basically the data of night-time lights emanating from the earth captured by satellites from outer space. These sources include moonlight, light directly emitted by a source (e.g., buildings and transport), and light reflected by the ground. It has several use cases - aid in disaster mitigation, estimating economic activity etc.
B	D	4c	This data is collected from social media networks to see how people are engaging on specific topics of interest. Scheme divisions may use the same to check for behavior change etc.
B	D	4d	Scheme may use data generated by private sector also as per requirement. For e.g.: mobility data from private cab aggregators, economic activity data from e-commerce websites etc.
B	D	5	Unique LGD codes have been created for each state, distt, sub-distt, block, village and local body by GoI. All MIS must use the same codes so that data on different platforms is easily integrable.
B	D	6a	Machine learning gives computers the ability to learn and predict from data without being explicitly programmed. E.g.: predicting the probability that individuals commit crimes, targeting hygiene inspections by data-mining online restaurant reviews or estimating poverty levels based on satellite imagery.
B	D	6b	AI refers to intelligence demonstrated by machines and can have several use cases in governance and delivery of schemes. e.g.: Monitoring social media for public feedback on policies, monitoring social media to identify emergency situations, Anticipating Road maintenance requirements, Providing personalized education to students etc.
B	D	6c	Blockchain refers to having distributed ledgers or blocks of transactional data that are linked together. Using this structure, govt. can offer services with improved data security. For e.g.: electronic health records, e-registries etc.
B	D	6d	IoT refers to network of objects embedded with sensors and technologies for collecting and exchanging data over Internet. e.g.: IoT to measure air quality, IoT to monitor power consumption i.e., smart metering etc.
B	D	6e	The use of advanced analytic techniques against very large, diverse data sets that include structured, semi-structured and unstructured data, from different sources.
B	D	6f	Drones can be used for monitoring of various sectors like agri, infra projects, commerce, logistics etc.
B	E	3	An SSL certificate is a digital certificate that authenticates a website's identity and enables an encrypted connection. SSL stands for Secure Sockets Layer, a security protocol that

Part	Section	Question	Explanation
			creates an encrypted link between a web server and a web browser.
B	E	5	a firewall refers to a network device which blocks certain kinds of network traffic, forming a barrier between a trusted and an untrusted network.
B	E	7	Sensitive/PII contains personal information of individuals, firms etc. which are not freely accessible to all.
B	E	7a	Single-Factor Authentication (SFA) is a method of logging users by having them present only one way of verifying their identity (usually, username and password). Multi-factor authentication uses more than one way – such as OTP, Captcha etc.
B	E	7b	List of users of MIS along with details of which user has access to which type of data is regularly maintained.
B	E	7c	Encryption refers to conversion of data from readable format to encoded format. Encrypted data can only be read and processed after its decrypted by recipient if they have the codes.
B	E	7d	Data anonymization refers to the process by which personal data is altered in a way that the data subject can no longer be identified directly by data user.
B	E	8	With advancements in machine learning and big data analytics, it is becoming increasingly easier to de-identify anonymized data using indirect means. Hence, it is important to protect personal data from re-identification risks.
B	E	8b	Includes provisions for mandatory audit trails, controlled access, only central server logins allowed etc.
B	E	8c	Sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals.
B	E	9	Before using and putting personal data in public domain such as photographs, names, other details of individuals or firms, their consent must be asked for and documented.
B	F	2	Before using and putting personal data in public domain such as photographs, names, other details of individuals or firms, their consent must be asked for and documented.
B	F	3	Data may be stored on physical servers or cloud servers. Cloud servers offer better disaster recovery.
B	F	4	Select the cloud server used by the scheme MIS.
B	F	5	Historical data refers to data corresponding to previous time periods which may not be actively used at present.

Annexure 4: Methodology in detail

Minor changes made in the methodology in this round, based on M/D's concerns in previous rounds, have been highlighted in yellow.

While developing the methodology for DGQI 1.0, DMEO had reviewed several existing frameworks for assessing data preparedness of organizations. Four data maturity models spanning both private and public context were shortlisted for a detailed study based on their relevance, exhaustiveness and representativeness: US Federal Government Data Maturity Model, Data Governance Maturity Model (IBM), Data Maturity Assessment Framework (SCM) and Data Maturity Management Model (CMMI). Based on the assessment of these models, three key pillars of data preparedness were identified viz., Data Strategy, Data Systems and Data Outcomes and this theory of change formed the basis for design of DGQI.

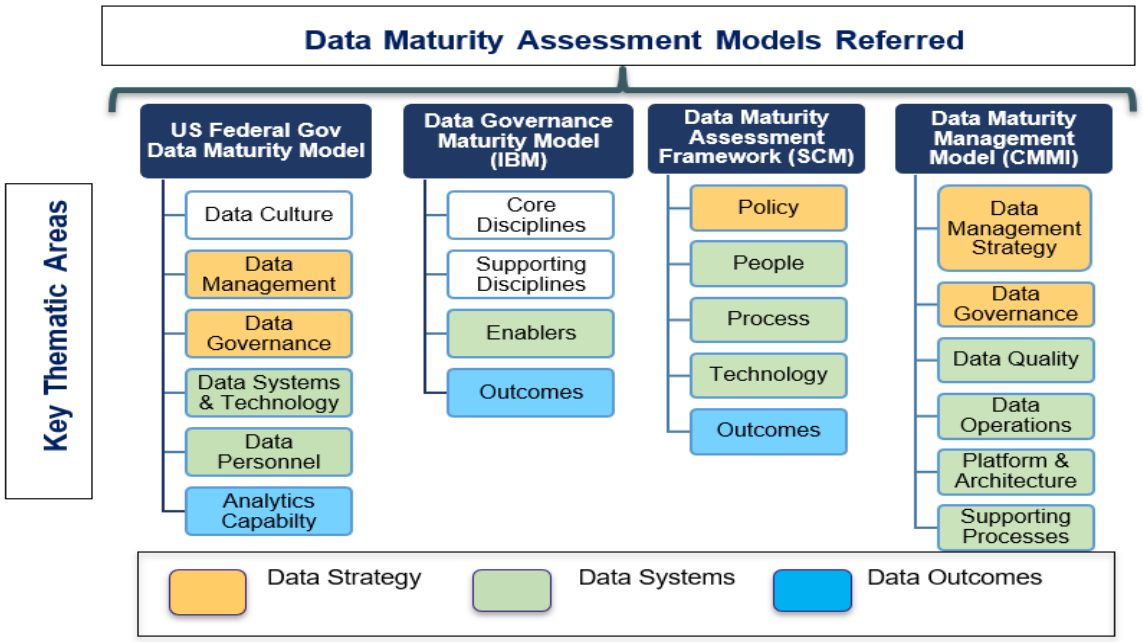


Figure 31 Reference Data Maturity Models

First of all, data strategy is required to lay down systemic guidelines for data governance by organisations.

Next, there is a role for well-defined and organised data systems encompassing various data processes such as data generation, ensuring data quality, use of technology, data analysis to create evidence, dissemination of evidence in user-friendly manner and existence of capable data management teams. Data systems are to be supported by enablers such as adequate financial allocation, correct placement of data management teams to ensure coordination with decision makers and configuration management to take care of other technical support.

The first and the second pillar work in conjunction with each other to enable the third pillar of data-driven outcomes. However, the existence of data strategies and systems alone cannot ensure that data is converted to information and is actually utilised as

evidence to guide decisions. The same has to be fostered within institutions through a step-by-step approach. This would involve integrated data use facilitated by exchange of data among various agencies, development of strong data analytical capabilities within Ministries/Departments and finally well-articulated data use plans. These aspects hence get covered under the third pillar – data-driven outcomes.

While DGQI 1.0 covered only data systems, DGQI 2.0 covered all three pillars. Hence, the corresponding themes from these four data maturity models mapped to the two new pillars were used this time to develop questions under these pillars in DGQI 2.0. All pillars are explained in detail below:

Data Strategy:

Under the data strategy pillar, two themes are covered within DGQI 2.0:

1.1. Data & Strategy Unit

Ministries/Departments were advised to set up a Data & Strategy Unit (DSU) as a central unit to steer the development and implementation of an action plan or data strategy to improve their data preparedness levels in general and reach DGQI 5.0 scores by December, 2022 in specific.

Within this theme, it was assessed if the Ministries/Departments have taken necessary steps in this direction to establish the necessary arrangements that are required for the development and maintenance of a robust data strategy.

1.2. Action Plan

Ministries/Departments were also advised to develop an action plan or data strategy with clear actionables, definite timelines and responsibilities to improve their data preparedness levels in general and reach DGQI 5.0 scores by December, 2022 in specific. An indicative outline of the action plan was also shared for reference and guidance.

Within this theme, it was assessed if the Ministries/Departments have developed action plans as per the outline. In addition, the compliance by the Ministries/Departments in completing the action points within the timelines set by themselves were measured.

2. Data Systems:

Under the data systems pillar, six themes are covered within DGQI 2.0:

2.1. Data generation

This theme measures the ability of Ministries/Departments to collect and report generate data on inputs, outputs and outcomes of their schemes. It covers areas related to the granularity and frequency of digitization and also covers if new approaches like CAPI surveys, GIS mapping, transactional data collection etc. is used to improve quality of generated data.

2.2. Data quality

This theme measures whether Ministries/Departments undertake data quality assessment procedures to evaluate the quality of incoming data and make suitable corrections. Key areas included under the theme pertain to data quality assessment,

automation of data quality assessment and use of new feedback and backcheck mechanisms to further validate data quality.

2.3. Data analysis, use & dissemination

This theme measures the ways in which collected data is analyzed and used by Ministries/Departments for evidence creation and decision making. Use of dashboards and other modes of dissemination are also included within this theme. Key areas also include ensuring accessibility of data, machine readability of data and open data systems for wider dissemination.

2.4. Use of technology

This theme covers linkage of Ministries/Departments' portals with other platforms like PFMS, JAM, GSTN, Udyog Aadhaar, LGD etc. wherever applicable. Use of alternative data sources outside the government like remote sensing data, social media data etc. to improve data robustness and use of emerging technologies in scheme monitoring are other key areas.

2.5. Data security & HR capacity

This theme measures the capacity of Ministries/Departments to ensure data security and privacy related concerns of their data systems. It also covers questions on human resource capacity of data quality and analysis teams for various schemes of Ministries/Departments.

2.6. Data management

This theme covers areas related to data management across its lifecycle i.e., guidelines for data management, data storage and historical data management.

3. Data driven outcomes:

Under this pillar, four themes have been identified under DGQI:

3.1. Synergistic data use within Ministries/Departments

This theme covers how Ministries/Departments have identified gaps in their data based on their existing data and begun to take steps internally to create better exchange systems to drive integrated data use.

3.2. Inter-agency data collaboration

This theme covers how Ministries/Departments have undertaken data-based collaborations with other agencies to drive better data-based outcomes and creating a rich data culture in the organization.

3.3. Prescriptive analytics

This theme covers how Ministries/Departments are trying to create a data culture by moving to prescriptive analytics and developing mechanisms for institutionalizing it in the long run.

3.4. Good practices

This theme highlights good practices adopted by Ministries/Departments in using data in driving smarter, granular and quicker decisions for informing policy along with its

quantified impact. It is expected to help unlock hidden potential by opening doors for cross-learning from challenges faced and solutions devised by peers.

4. Weightages & Scoring

DGQI scores are arrived at on the basis of responses filled up by Ministries/Departments to the self-assessment questionnaire. The self-assessment questionnaire consists of two parts: Part A (to be filled at M/D level) and Part B (to be filled for each CS/CSS scheme/non-schematic intervention at CS/CSS/NSI level). The above-mentioned pillars and themes have been covered within these two parts of the questionnaire.

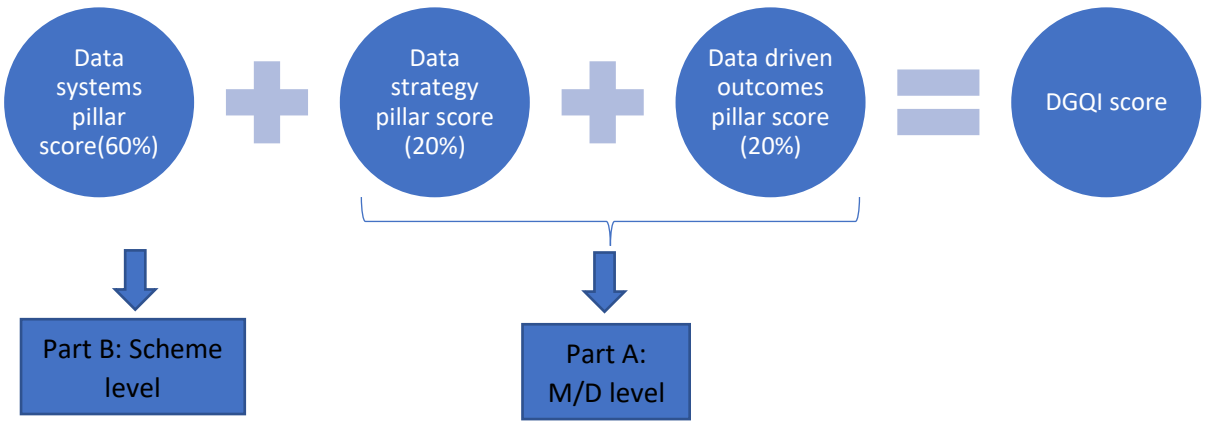
The response to each question is scored on a scale of 1 to 5, which is then aggregated using weighted averages to arrive at scores at themes, pillar and overall index level. The same is explained below in detail.

4.1. Pillar-wise weightages

The data systems pillar would be appropriated an overall weight of 60% as it is a major pillar where outputs of data strategy are visible which then also play a key role in the ability of Ministries/ Departments to achieve desired data driven outcomes. The data systems pillar scores would be based on scheme-level information provided in Part B of the self-assessment questionnaire.

Remaining 40% weight would be appropriated to the data strategy and data driven outcomes pillar combined. This 50% has been distributed equally between data strategy (20%) and data driven outcomes (20%). The scores on these two pillars would be based on M/D level information provided in Part A of the self-assessment questionnaire.

Hence, overall DGQI Score = 60% *(Data systems pillar score) + 20% *(Data strategy pillar score) + 20% *(Data driven outcomes pillar score)



4.2. Theme-wise weightages

To arrive at each pillar score, each theme was further assigned a weight, as shown below:

- **Data systems pillar:**

Table 4: Theme wise weightages within data systems pillar

Pillar	Theme	Theme weightage within data systems pillar
Data Systems	Data Generation	18%
	Data Quality	18%
	Data analysis, use & dissemination	18%
	Use of technology	10%
	Data security & HR capacity	18%
	Data management	18%
	All themes	100%

Apart from use of technology, all remaining themes within data systems were decided to be allocated equal weightage as all these systems were found to be equally important for ensuring robust data systems. Use of technology was allocated a weight of 10% to promote the use of emerging technologies across data systems.

Hence,

Data systems score = 18% * (Data generation score) + 18% *(Data quality score) + 18% * (Data analysis, use & dissemination score) + 10% * (Use of technology score) + 18% *(Data security & HR capacity score) + 18% * (Data management score)

As explained above, data systems pillar scores would be based on scheme-level information provided in Part B of the self-assessment questionnaire. Hence, for each scheme (filled up in Part B), a data systems score would be generated using the above formula. Then, a simple average of these scheme-level scores would be calculated to arrive at a combined data systems score.

- **Data strategy pillar:**

Table 5: Theme wise weightages within data strategy pillar

Pillar	Theme	Theme weightage within data strategy pillar
Data Strategy	Data & Strategy Unit	50%
	Action Plan	50%

Both DSU and action plan were appropriated equal weightages as both were found to be equally important components of data strategy pillar.

Hence,

Data strategy score = 50% * (Data & Strategy Unit score) + 50% * (Action plan score)

- **Data driven outcomes pillar:**

Table 6: Theme wise weightages within data driven outcomes pillar

Pillar	Theme	Theme weightage within data driven outcomes pillar
Data driven outcomes	Synergistic data use within M/D	30%
	Inter-agency data collaboration	30%
	Prescriptive Analytics	10%
	Good practices	30%

Driving intra-ministerial and inter-ministerial integrated use of data is one of the key outcomes for fostering a data culture. Similarly, good practices offer huge scope for peer learning. Hence, these three themes have been given the highest equal weightages.

Hence,

Data driven outcomes score = 30% * (Synergistic data use within M/D score) + 30% * (Inter-agency data collaboration score) + 10% * (Prescriptive Analytics score) + 30% * (Good practices score)

4.3. Question wise weightages

As each theme had multiple questions within it as a part of the self-assessment questionnaire, each question was also accorded an appropriate weightage within the theme. The same is tabulated below for all questions.

Table 7: Question wise weightages within each theme

Pillar	Theme	Question No.	Question	Question Weightage within theme
Data Strategy	Data & Strategy Unit	Part A, B1	Constitution	5%
		Part A, B2	Head	5%
		Part A, B3	Verticals	10%
		Part A, B4	Strength	20%
		Part A, B5	ToR	20%
		Part A, B6	Review mechanisms	20%
		Part A, B7	Frequency of review	20%
	Action Plan	Part A, C1	Action plan	5%
		Part A, C2	Sections	5%
		Part A, C3	Schemes	5%
		Part A, C4	Timelines	5%
		Part A, C5	Responsibilities	5%
		Part A, C7	Compliance Scoring	75%
Data Systems	Data Generation	Part B, A1	Requirements gathering	10%
		Part B, A2	Collection	10%
		Part B, A3	Digitization	20%
		Part B, A4	Granularity	20%

		Part B, A5	Frequency	20%
		Part B, A6 and 7	Use of technologies in generation	20%
	Data Quality	Part B, B1	QC mechanisms	20%
		Part B, B2	QC automation	20%
		Part B, B3	Data quality assessment	40%
		Part B, B4	Use of mobile phones in QC	20%
	Data analysis, use & dissemination	Part B, C1	Types of data analysis	15%
		Part B, C2	Cross sectoral analysis	10%
		Part B, C3	Documentation of data analysis	10%
		Part B, C4	Use of data analysis	15%
		Part B, C5	Modes of dissemination	5%
		Part B, C6	Use of dashboards	15%
		Part B, C7	Data visualization	5%
		Part B, C8	Data visualization on maps	5%
		Part B, C9 and 10	Data Accessibility for all	5%
		Part B, C11 and C13	Open data	5%
		Part B, C13	Open data – 2	5%
		Part B, C12	Machine readable data	5%
	Use of technology	Part B, D1	Linkage with PFMS	10%
		Part B, D2	Last mile linkage of PFMS	20%
		Part B, D3	Linkage with other platforms	20%
		Part B, D5	Linkage with LGD Codes	20%
		Part B, D4	Use of alternative data sources	10%
		Part B, D6	Use of emerging technologies	20%
	Data security & HR capacity	Part B, E1	Antivirus updates	5%
		Part B, E2	Security audits	5%
		Part B, E3 and 4	SSL certification	5%
		Part B, E5	Firewalls	5%

		Part B, E6	External communication	10%
		Part B, E7, 8	Personal data protection	10%
		Part B, E9	Personal data protection -2	10%
		Part B, E10	Data QC team	25%
		Part B, E11	Data analysis team	25%
	Data management	Part A, D1,3,4,5	Data management architecture	25%
		Part A, D2	Data management Compliance	10%
		Part B, F1 and 3	Distributed cloud storage	20%
		Part B, F4	Type of cloud storage	15%
		Part B, F2	Data sharing mechanisms	15%
		Part B, F5	Historical data management	15%
Data driven outcomes	Synergistic data use within M/D	Part A, E1 and 2	Identification of data gaps	40%
		Part A, E3	Data exchange	60%
	Inter-agency data collaboration	Part A, F1	Collaborations	50%
		Part A, F2	Types of collaborations	50%
	Prescriptive Analytics	Part A, G1	Prescriptive analytics	50%
		Part A, G2	Frequency	25%
		Part A, G3	Modes	25%
	Good Practices	Part A, H	Good practices	100%

5. Question Wise Scoring Mechanism

Table 8: Scoring mechanism

Pillar	Theme	Question No.	Question	Scoring mechanism
Data Strategy	Data & Strategy Unit	Part A, B1	Constitution	If response is 'Yes' score '5', if 'No' score '0'.
		Part A, B2	Head	If response if 'AS' or 'JS' score '5', if response is 'Director' score '3', if response is 'Below Director' score '0'. If question is disabled, score '0'.

		Part A, B3	Verticals	If all four verticals/sub-units are selected score '5', if three/two verticals are selected score '3', if only one is selected score '1', else score '0'. If question is disabled, score '0'.
		Part A, B4	Strength	If total % is above 80% score '5', if total % is between 60% to 80% score '4', if total % is between 40% to 60% score '3', if total is between 20% to 40% score '2', if total is below 20% score '0'. If question is disabled, score '0'.
		Part A, B5	ToR	If response is 'Yes' score '5', if 'Partial' score '3', if 'No' score '0'. If question is disabled, score '0'.
		Part A, B6	Review mechanisms	If response is 'Yes' score '5', if 'No' score '0'. If question is disabled, score '0'.
		Part A, B7	Frequency of review	If response is 'daily'/'weekly'/'fortnightly/monthly' score '5', if response is 'quarterly' score '3', if response is 'annually' score '1'. If question is disabled, score '0'.
	Action Plan	Part A, C1	Action plan	If response is 'Yes' score '5', if 'No' score '0'.
		Part A, C2	Sections	If response is 'Yes' score '5', if 'Partial' score '3', if 'No' score '0'. If question is disabled, score '0'.
		Part A, C3	Schemes	If response is 'Yes' score '5', if 'Partial' score '3', if 'No' score '0'. If question is disabled, score '0'.
		Part A, C4	Timelines	If response is 'Yes' score '5', if 'Partial' score '3', if 'No' score '0'. If question is disabled, score '0'.
		Part A, C5	Responsibilities	If response is 'Yes' score '5', if 'Partial' score '3', if 'No' score '0'. If question is disabled, score '0'.
		Part A, C7	Compliance Scoring	If timely compliance is above 80% score '5', if between 60% to 80% score '4', if between 40% to 60% score '3', if between 20% to 40% score '2', if below 20%, score '0'. If question is disabled, score '0'. If none of the action points are due when scores are being calculated, a standard score of 1 is given (as no timelines are due, it suggests action plan is not detailed and granular enough - hence low score).

Data Systems	Data Generation	Part B, A1	Requirements gathering	If response is 'Yes' score '5', if 'No' score '0'.
		Part B, A2	Collection	If response to all three parts is 'Yes', score 5. If response to two is 'Yes' and one is 'Partial', score 4. If response to two is 'Yes' and one is 'No', score 3. If response to one part is 'Yes' and two is partial, score '3'. If response to one is 'Yes', one is 'Partial', one is 'No', score 2. If response to one part is 'Yes' and two is 'No', score 1. If response to two is 'Partial' and one is 'no', then score '2'. If response to one part is 'partial' and two is 'no', score '1'. If response to all three parts is 'Partial', score 3. If response to all three parts is 'No', score 0.
		Part B, A3	Digitization	If response is 'Yes' score '5', if 'No' score '0'.
		Part B, A4	Granularity	Lowest level of granularity to be used - '1' at national level, '3' at State level, '4' at district/sub-district/block level and '5' at village/individual/facility/ project level. If question is disabled, score '0'.
		Part B, A5	Frequency	Lowest level of frequency to be used - '1' at Yearly, '2' at half-yearly, '3' at Quarterly, '4' at monthly/fortnightly/weekly and '5' at daily/real time/near real time level. If question is disabled, score '0'.
		Part B, A6 and 7	Use of technologies in generation	If Q6 is disabled, score '0'. If response to Q6 is 'Option 1' score '1'. If response to Q6 is 'Option2' or 'Option 3', then use Q7 responses to score. If none of the responses to Q7 is 'Yes', score '3'. If anyone responses to Q7 is 'Yes', score '5'.
	Data Quality	Part B, B1	QC mechanisms documentation	If response is 'Yes' score '5', if 'No' score '0'.
		Part B, B2	QC automation	If response is 'Not done' score '0', if 'Manually' score '2', if 'Hybrid' score '3', if 'Automatically' score '5'.

		Part B, B3	Data quality assessment	If no response is 'Yes', score '0'. If anyone response is 'Yes' score '1', if any two responses are 'Yes' score '2', if any three responses are Yes, score '3', if any four responses are 'Yes' score '4', if all responses are yes, score '5'. If question is disabled, score '0'.
		Part B, B4	Use of mobile phones in QC	If no response is 'Yes', score '0'. If one or two responses are 'Yes' score '3'. if three or more responses are 'Yes' score '5'. If question is disabled, score '0'.
		Part B, C1	Types of data analysis	If no response is 'Yes' score '0'. If any 1/6 option is selected then score '1', if 2/6 options are selected then score '2'. If 3/6 options are selected then score '3'. If 4 or more options are selected then score '5'.
		Part B, C2	Cross sectoral analysis	If response is 'Yes' score '5'. If response is 'No' score '0'. If question is disabled, score '0'.
		Part B, C3	Documentation of data analysis	If response is 'Never' score '0'. If 'Annually', score '2'. If 'Half-yearly', score '3'. If 'Quarterly', score '4'. If 'Real time on a dashboard', score '5'. If question is disabled, score '0'.
		Part B, C4	Use of data analysis	If question is disabled, score '0'. If response is 'no' for all the sub-categories, score '0'. If response yes for 1-2/7 sub-categories, score '1'. If response is yes for 3-4/7 sub-categories, score 3'. If response is yes for 5-7/7 sub-categories, score '5'.
		Part B, C5	Modes of dissemination	If response is 'No' for all sub-categories, score '0'. If response is 'Yes' for 1-2 sub-categories, score '1'. If response is 'Yes' for 3-4 sub-categories, score '3'. If response is 'Yes' for 5-7 sub-categories, score '5'.
		Part B, C6	Use of dashboards	If question is disabled, score '0'. If response is 'No' for all sub-categories, score '0'. If response is 'Yes' for 1-2 sub-categories, score '3'. If response is 'Yes' for 3-4 sub-categories, score '5'.

		Part B, C7	Data visualization types	If question is disabled, score '0'. If response is 'No' for all sub-categories, score '0'. If response is 'Yes' for 1-2 sub-categories, score '1'. If response is 'Yes' for 3-4 sub-categories, score '3'. If response is 'Yes' for 5-7 sub-categories, score '5'.
		Part B, C8	Data visualization on maps	If question is disabled, score '0'. If Yes, score '5', else '0'.
		Part B, C9 and 10	Data Accessibility for all	If question is disabled, score '0'. If response to both Q9 and Q10 is 'No', score '0'. If response to one is 'No' and one is 'partially' score '1'. If response to both are 'partially', score '3'. If response to one is 'Yes' and one is 'partially', score 4. If both are 'yes' score '5'.
		Part B, C11	Open data	If question is disabled, score '0'. If response is 'Not accessible', score '0'. If response is 'Accessible through credentials', score '3'. If response is 'Openly accessible', score '5'.
		Part B, C13	Open data	If question is disabled, score '0'. If response is 'Yes', score 5, else '0'.
		Part B, C12	Machine readable data	If question is disabled, score '0'. If response is 'Yes', score '5', if 'Partially' score '3', if 'No' score '0'.
	Use of technology	Part B, D1	Linkage with PFMS	If response is 'Yes' score '5', else '0'. If question is disabled, score '0'.
		Part B, D2	Last mile linkage of PFMS	If previous question was disabled, this will also be disabled and scored '0'. If 'yes' is selected in previous question, score based on response provided to this question. If 'no' is selected as a response here, score '0', if 'Partially', score '3', if 'Yes', score '5'. If 'no' is selected in previous question, this question will be disabled and scored '0'. If 'NA' is selected in previous question, this question will be disabled and scored '5'.

		Part B, D3	Linkage with other platforms	If no option selected, score '0'. If one option is yes, score '3'. If more than two options are selected, score '5'. If question is disabled, score '0'.
		Part B, D5	Linkage with LGD Codes	If response is 'Yes' score '5', else '0'. If question is disabled, score '0'.
		Part B, D4	Use of alternative data sources	If no option selected, score '0'. If any one option is yes, score '5'.
		Part B, D6	Use of emerging technologies	If no option selected, score '0'. If any one option is yes, score '5'.
	Data security & HR capacity	Part B, E1	Antivirus updates	If response is 'Yes', score 5, else '0'. If question is disabled, score '0'.
		Part B, E2	Security audits	If response is 'Yes', score 5, else '0'. If question is disabled, score '0'.
		Part B, E3 and 4	SSL certification	If response to Q3 is 'No', score '0'. If response to Q3 is 'Yes', use responses for q4 to score further. If response to Q4 is 'No', score '3'. If response to Q4 is also 'Yes', score '5'. If Q3 is disabled, score '0'.
		Part B, E5	Firewalls	If response is 'Yes', score 5, else '0'. If question is disabled, score '0'.
		Part B, E6	External communication	If response is 'Yes' or 'No external communication established', score 5, else '0'. If question is disabled, score '0'.
		Part B, E7, 8	Personal data protection	If question is disabled due to no MIS, score '0'. First check if response is 'No such data', score '5'. If this option is not selected, check which of remaining four options are selected. If only first/second option is selected, score '1'. If both first and second option are selected (but not third and fourth) score '2'. If third option is selected (but fourth is not), score '3'. If fourth option is selected (either along with other options or only fourth option is selected) and 'No efforts made'/'others' selected in Q8, score '4'. Further, if fourth option is selected (either along with other options or only fourth option is selected) and any other option selected in Q8, score '5'.

		Part B, E9	Personal data protection - 2	If question is disabled due to no MIS, score '0'. If question was disabled due to 'No such data' response in Q7, score '5'. If response is 'Yes', score 5, else '0'.
		Part B, E10	Data QC team	If response is 'Yes', score 5, else '0'.
		Part B, E11	Data analysis team	If response is 'Yes', score 5, else '0'.
	Data management	Part A, D1,3,4,5	Data management architecture	If response to all four questions is 'No', score '0'. If response to only one question is 'Yes', score '1', if response to only any two questions is 'Yes', score '2'. If response to any three is 'Yes', score '4'. If response to all is 'Yes', score '5'.
		Part A, D2	Data management Compliance	If response is 'Yes', score 5, else '0'. If question is disabled, score '0'.
		Part B, F1 and 3	Distributed cloud storage	If 'Separate servers' in Q1 and 'physical servers' or 'hybrid servers' in Q3, score '0'. If 'Central server' in Q1 and 'physical servers' or 'hybrid servers' in Q3, score '1'. If 'Separate servers' in Q1 and 'cloud servers' in Q3, score '3'. If 'Central server' in Q1 and 'cloud servers' in Q3, score '5'.
		Part B, F2	Data sharing mechanisms	If 'yes', score '5', else score '0'.
		Part B, F4	Type of cloud storage	If first or second option is selected, score '5', else score '0'.
		Part B, F5	Historical data management	Is response is "Data is not backed up", score '0'. If response is "Data is backed up and data is archived", score '3'. If response is "Data history is well maintained including retention, destruction, and audit trail details", score '5'. If question is disabled, score '0'.
Data driven outcomes	Synergistic data use within M/D	Part A, E1 and 2	Identification of data gaps	Is response to Q1 is "No", score 0. If response to Q1 is "Yes", use Q2 to score further. If response to Q2 is "No", score '3', If response to Q2 is also 'Yes', score '5'.

		Part A, E3	Data exchange	If response is "No", score 0. If response is "In- progress", score 3. If response is "Yes", score 5. If "Yes" is the response, its veracity will be validated from the subjective descriptions and hence responding to descriptions is also important.
	Inter-agency data collaboration	Part A, F1	Collaborations	If response is "No", score 0. If response is "In- progress", score 3. If response is "Yes", score 5. If "Yes" is the response, its veracity will be validated from the subjective descriptions and hence responding to descriptions is also important.
		Part A, F2	Types of collaborations	If none of the options are selected, score '0'. If one to five options selected, score '3'. If more than five options are selected, score '5'. If question is disabled, score '0'.
	Prescriptive Analytics	Part A, G1	Prescriptive analytics	If response is "No", score 0. If response is "In- progress", score 3. If response is "Yes", score 5. If "Yes" is the response, its veracity will be validated from the subjective descriptions and hence responding to descriptions is also important.
		Part A, G2	Frequency	If response is "Annually", score 3. If response is "Quarterly/Monthly" score 5. If question is disabled, score '0'.
		Part A, G3	Modes	If none of the options are selected, score '0'. If any one option is selected (other than "others"), score '5'. If question is disabled, score '0'.
	Good Practices	Part A, H	Good practices	Each good practice will be assessed on 3 parameters – relevance of practice to DGQI exercise and objectives (40%), exhaustiveness of the case study (30%) and impact of the intervention (30%).” Then, a simple average of the three scores for each good practice will be taken to arrive at overall good practice dimension score.

6. Summary

This way, in order to arrive at DGQI scores, a three-tiered weighted average process is used:

(a) First, weighted average of question wise scores within each theme.
(b) Second, weighted average of theme wise scores within each pillar. Within this step, for data systems pillar, initially, data systems scores are calculated for each scheme separately. To aggregate the same into a single score at M/D level, a simple average of these scheme level scores is calculated to arrive data systems pillar score.
(c) Third, weighted average of pillar wise scores to arrive at final DGQI score for the M/D.

7. Special Cases

For scoring purposes, for a certain question, NA option is selected, to not penalize any entity for any requirement that is not applicable for them, its weight will be redistributed among other questions within the theme. However, if it is the case that only certain sub-parts (a,b,....) of a question are not applicable, a case-by-case mechanism of how they will be taken care of in at the scoring stage has been devised in the following manner:

Table 9: NA scoring mechanism

Q.No.	Question	Scoring mechanism	Way to handle NA
Part B, A4	Granularity	Lowest level of granularity to be used - '1' at national level, '3' at State level, '4' at district/sub-district/block level and '5' at village/individual/facility/ project level. If question is disabled, score '0'.	**Scoring should not be changed since it's a range. Any one of the options from village/individual/facility/project must be applicable for all schemes.** Hence, if the scheme is collecting data at any level not equivalent to these four options, scores should be decreased the way they have been done.
Part B, A5	Frequency	Lowest level of frequency to be used - '1' at Yearly, '2' at half-yearly, '3' at Quarterly, '4' at monthly/fortnightly/weekly and '5' at daily/real time/near real time level. If question is disabled, score '0'.	**Only need to change scoring if daily/ real time/ near real time/ monthly/fortnightly/weekly - all of these options are not applicable** (infra schemes with long gestation periods). In this case, quarterly to be scored as '5', half yearly as '3' and yearly as '1'.
Part B, A6 and 7	Use of technologies in generation	If Q6 is disabled, score '0'. If response to Q6 is 'Option 1' score '1'. If response to Q6 is 'Option2' or 'Option 3', then use Q7 responses to score. If none of the responses to Q7 is 'Yes', score '3'. If anyone	**Only need to change scoring if none of the options in Q7 are applicable**, otherwise scheme already gets full score. In this case, scoring will be done only on basis of 6 - 1 if first option is selected, 5 if other two options are selected.

		responses to Q7 is 'Yes', score '5'.	
Part B, B4	Use of mobile phones in QC	If no response is 'Yes', score '0'. If one or two responses are 'Yes' score '3'. if three or more responses are 'Yes' score '5'. If question is disabled, score '0'.	The options are such that if one is applicable, all others would also be applicable. **Hence, only need to change scoring if none of the options are applicable** (research/defence schemes). If this is the case, weights to be redistributed within data quality theme.
Part B, C2	Cross sectoral analysis	If response is 'Yes' score '5'. If response is 'No' score '0'. If question is disabled, score '0'.	If it is NA, its weight will be redistributed within data analysis theme.
Part B, C5	Modes of dissemination	If response is 'No' for all sub-categories, score '0'. If response is 'Yes' for 1-2 sub-categories, score '1'. If response is 'Yes' for 3-4 sub-categories, score '3'. If response is 'Yes' for 5-7 sub-categories, score '5'.	It should not be the case that all options are NA- If scheme enters so, it would be scored '0' as a disincentive for entering wrong responses. Otherwise, NA treated as yes and accordingly scored as per the method.
Part B, C8	Data visualization on maps	If question is disabled, score '0'. If Yes, score '5', else '0'.	If it is NA, its weight will be redistributed within data analysis theme.
Part B, D1	Linkage with PFMS	If response is 'Yes' score '5', else '0'. If question is disabled, score '0'.	If it is NA, its weight will be redistributed within use of technology theme.
Part B, D3	Linkage with other platforms	If no option selected, score '0'. If one option is yes, score '3'. If more than two options are selected, score '5'. If question is disabled, score '0'.	If every option is NA - weight to be redistributed. Otherwise, ranges are defined in a manner that schemes get correct scores. If any one option is yes, by virtue of options, at least two become applicable.
Part B, D4	Use of alternative data sources	If no option selected, score '0'. If any one option is yes, score '5'.	If every option is NA – weight to be redistributed. Otherwise, ranges are defined in a manner that schemes get full scores.
Part B, D5	Linkage with LGD Codes	If response is 'Yes' score '5', else '0'. If question is disabled, score '0'.	If it is NA, its weight will be redistributed within use of technology theme.
Part B, D6	Use of emerging technologies	If no option selected, score '0'. If any one option is yes, score '5'.	If every option is NA - weight to be redistributed. Otherwise, ranges are defined in a manner that schemes get full scores.

Part A, E3	Data exchange	If response is "No", score 0. If response is "In- progress", score 3. If response is "Yes", score 5. If "Yes" is the response, its veracity will be validated from the subjective descriptions and hence responding to descriptions is also important.	If it is NA, its weight will be redistributed within synergistic data use within M/D theme.

Annexure 5: M/D wise Number of CS/CSS/NSIs covered under DGQI 2.0 Round 2 (Quarter 4 of FY 2021-22)

Category	Ministry Department Name	CS	CSS	NSI	Total
Admin	Department of Administrative Reforms and Public Grievances	1			1
Admin	Department of Ex-Servicemen Welfare			3	3
Admin	Department of Legal Affairs			1	1
Admin	Department of Personnel & Training	5			5
Admin	Legislative Department			1	1
Economic	Department of Chemicals and Petrochemicals	2			2
Economic	Department of Commerce	16			16
Economic	Department of Consumer Affairs	7			7
Economic	Department of Economic Affairs	4		2	6
Economic	Department of Fertilisers	2			2
Economic	Department of Financial Services	13		10	23
Economic	Ministry of Heavy Industry	2		1	3
Economic	Department of Pharmaceuticals	5			5
Economic	Department of Promotion of Industry and Internal Trade	8		3	11
Economic	Department of Public Enterprises	2		1	3
Economic	Ministry of Food Processing Industries	2	1		3
Economic	Ministry of Labour and Employment	14	1	5	20
Economic	Ministry of Micro, Small and Medium Enterprises	24		3	27
Economic	Ministry of Steel	2		2	4
Economic	Ministry of Textiles	23			23
Economic	Ministry of Tourism	7	1		8
Infrastructure	Department of Posts	4			4
Infrastructure	Department of Telecommunications	2			2
Infrastructure	Ministry of Civil Aviation	1		5	6
Infrastructure	Ministry of Coal	3			3
Infrastructure	Ministry of Electronics and Information Technology	10			10
Infrastructure	Ministry of Housing and Urban Affairs	3	4	2	9
Infrastructure	Ministry of Mines			1	1
Infrastructure	Ministry of New and Renewable Energy	8			8
Infrastructure	Ministry of Petroleum and Natural Gas	9			9
Infrastructure	Ministry of Power	2	1	1	4
Infrastructure	Ministry of Railways	2	1	1	4
Infrastructure	Ministry of Road Transport and Highways	2		8	10
Infrastructure	Ministry of Shipping	6			6
Scientific	Department of Agricultural Research and Education	28			28
Scientific	Department of Biotechnology	2			2

Category	Ministry Department Name	CS	CSS	NSI	Total
Scientific	Department of Health Research	7		3	10
Scientific	Department of Science and Technology	4		4	8
Scientific	Department of Scientific and Industrial Research	2		1	3
Scientific	Department of Space	4			4
Scientific	Ministry of Earth Sciences	5			5
Social	Department of Agriculture, Cooperation and Farmers' Welfare	17	16		33
Social	Department of Animal Husbandry and Dairying		8		8
Social	Department of Drinking Water and Sanitation		2		2
Social	Department of Empowerment of Persons with Disabilities	7			7
Social	Department of Fisheries		2		2
Social	Department of Food and Public Distribution	8	1	3	12
Social	Department of Health and Family Welfare	5	10	1	16
Social	Department of Higher Education	32	1		33
Social	Department of Justice	2	3	1	6
Social	Department of Land Resources	1	1		2
Social	Department of Rural Development	1	7		8
Social	Department of School Education and Literacy	1	2		3
Social	Ministry of Skill Development and Entrepreneurship	1	4		5
Social	Department of Social Justice and Empowerment	13	11		24
Social	Department of Sports	6			6
Social	Department of Water Resources, River Development and Ganga Rejuvenation	8	6	3	17
Social	Department of Youth Affairs	7		2	9
Social	Ministry of Ayurveda, Yoga and Naturopathy, Unani, Siddha and Homeopathy (AYUSH)	13	1		14
Social	Ministry of Culture	6			6
Social	Ministry of Environment, Forests and Climate Change	11	3	7	21
Social	Ministry of Minority Affairs	17	2		19
Social	Ministry of Panchayati Raj	1	3		4
Social	Ministry of Tribal Affairs	6	5	8	19
Social	Ministry of Women and Child Development	1	11	2	14
Strategic	Department of Defence			3	3
Strategic	Department of Defence Production			8	8
Strategic	Ministry of Corporate Affairs	3			3

Category	Ministry Department Name	CS	CSS	NSI	Total
Strategic	Ministry of Development of North Eastern Region	8	1	1	10
Strategic	Ministry of External Affairs	1		10	11
Strategic	Ministry of Home Affairs	6	1		7
Strategic	Ministry of Information and Broadcasting	4			4
Strategic	Ministry of Planning	2			2
Strategic	Ministry of Statistics and Programme Implementation	2			2
Total	**All M/Ds**	**433**	**110**	**107**	**650**