DGQI 2.0: Scoring Mechanism and Methodology

Section I: Pillars and Themes

DGQI 2nd edition focuses on all three pillars of data preparedness as identified under DGQI 1st edition: Data Strategy, Data Systems and Data driven Outcomes.

1. Data Strategy:

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

Under the data strategy pillar, two themes are covered within DGQI 2nd edition:

1.1. Data & Strategy Unit

As per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to Hon'ble PM dated 02nd Feb, 2021, Ministries/Departments have been 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 would be 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

As per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to Hon'ble PM dated 02nd Feb, 2021, Ministries/Departments have been 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 would be 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 would also be measured.

2. Data Systems:

After strategy is laid down, there is a role for well-defined and organized data systems encompassing various data processes such as digital data generation, ensuring data quality, data analysis to create evidence, dissemination of evidence in user-friendly manner, use of technology, data security & privacy, data management across its lifecycle and existence of capable data quality and analysis teams.

Under the data systems pillar, six themes are covered within DGQI 2nd edition:

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:

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 utilized 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 and gradual movement towards prescriptive analytics. These aspects are covered under the third pillar – 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.

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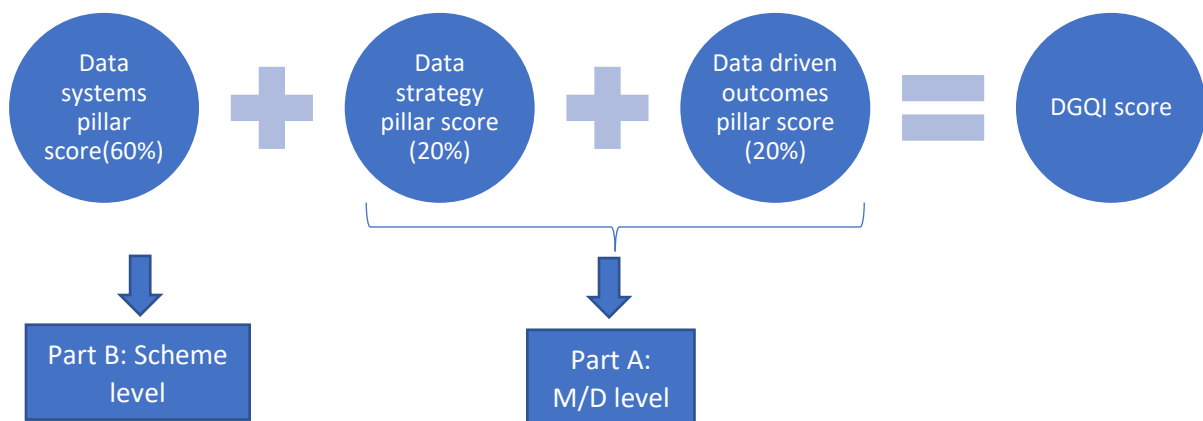
The response to each question is scored on a scale of 1 to 5, which is then aggregated using weighted averages to arrive at scores at themes, pillar and overall index level. The same is explained below in detail.

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)

2. Theme-wise weightages

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

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:

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:

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)

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.

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%

Pillar	Theme	Question No.	Question	Question Weightage within theme
		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%

Pillar	Theme	Question No.	Question	Question Weightage within theme
		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%

4. Question-wise scoring mechanism

The mechanism followed for attributing a score between 0 to 5 to response received for each question is provided below:

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 > 80% score '5', if total % is >60% and <=80% score '4', if total % is >40% to <=60% score '3', if total is >20% and <=40% score '2', if total is <=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

Pillar	Theme	Question No.	Question	Scoring mechanism
				'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

Pillar	Theme	Question No.	Question	Scoring mechanism
				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 halfyearly, '3' at Quarterly, '4' at monthly/fortnightly/weekly and '5' at daily/realtime/near realtime 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 any one 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 any one 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'.

Pillar	Theme	Question No.	Question	Scoring mechanism
		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

Pillar	Theme	Question No.	Question	Scoring mechanism
				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'.

Pillar	Theme	Question No.	Question	Scoring mechanism
				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'.

Pillar	Theme	Question No.	Question	Scoring mechanism
		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

Pillar	Theme	Question No.	Question	Scoring mechanism
				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

Pillar	Theme	Question No.	Question	Scoring mechanism
				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

Pillar	Theme	Question No.	Question	Scoring mechanism
				simple average of the three scores for each good practice will be taken to arrive at overall good practice dimension score.

5. 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.

The final share of each question/theme/pillar in DGQI is provided below:

S. No.	Pillar	Pillar Weightage	Theme	Theme wise weightage within pillar	Question No.	Question	Question Weightage within theme	Theme weightage in index	Question weightage in index
1	Data Strategy	20%	Data & Strategy Unit	50%	Part A, B1	Constitution	5%	10%	0.50%
2					Part A, B2	Head	5%		0.50%
3					Part A, B3	Verticals	10%		1.00%
4					Part A, B4	Strength	20%		2.00%
5					Part A, B5	ToR	20%		2.00%
6					Part A, B6	Review mechanisms	20%		2.00%
7					Part A, B7	Frequency of review	20%		2.00%
8			Action Plan	50%	Part A, C1	Action plan	5%	10%	0.50%
9					Part A, C2	Sections	5%		0.50%
10					Part A, C3	Schemes	5%		0.50%
11					Part A, C4	Timelines	5%		0.50%
12					Part A, C5	Responsibilities	5%		0.50%
13					Part A, C7	Compliance Scoring	75%		7.50%
14	Data Systems	60%	Data Generation	18%	Part B, A1	Requirements gathering	10%	11%	1.08%
15					Part B, A2	Collection	10%		1.08%
16					Part B, A3	Digitization	20%		2.16%
17					Part B, A4	Granularity	20%		2.16%
18					Part B, A5	Frequency	20%		2.16%
19					Part B, A6 and 7	Use of technologies in generation	20%		2.16%
20			Data Quality	18%	Part B, B1	QC mechanisms	20%	11%	2.16%
21					Part B, B2	QC automation	20%		2.16%

22					Part B, B3	Data quality assessment	40%		4.32%
23					Part B, B4	Use of mobile phones in QC	20%		2.16%
24				Data analysis, use & disseminatio n	18%	Part B, C1	Types of data analysis	15%	11%	1.62%
25						Part B, C2	Cross sectoral analysis	10%		1.08%
26						Part B, C3	Documentation of data analysis	10%		1.08%
27						Part B, C4	Use of data analysis	15%		1.62%
28						Part B, C5	Modes of dissemination	5%		0.54%
29						Part B, C6	Use of dashboards	15%		1.62%
30						Part B, C7	Data visualization types	5%		0.54%
31						Part B, C8	Data visualization on maps	5%		0.54%
32						Part B, C9 and 10	Data Accessibility for all	5%		0.54%
33						Part B, C11	Open data	5%		0.54%
34						Part B, C13	Open Data - 2	5%		0.54%
35						Part B, C12	Machine readable data	5%		0.54%
36				Use of technology	10%	Part B, D1	Linkage with PFMS	10%	6%	0.60%
37						Part B, D2	Last mile linkage of PFMS	20%		1.20%
38						Part B, D3	Linkage with other platforms	20%		1.20%
						Part B, D5	Linkage with LGD Codes	20%		1.20%
39						Part B, D4	Use of alternative data sources	10%		0.60%
40						Part B, D6	Use of emerging technologies	20%		1.20%
41				Data security & HR capacity	18%	Part B, E1	Antivirus updates	5%	11%	0.54%
42						Part B, E2	Security audits	5%		0.54%
43						Part B, E3 and 4	SSL certification	5%		0.54%

#	Domain	Weight	Sub-category	Sub-weight	Part reference	Description	%	Category weight	Final %
44					Part B, E5	Firewalls	5%		0.54%
45					Part B, E6	External communication	10%		1.08%
46					Part B, E7, 8	Personal data protection	10%		1.08%
47					Part B, E9	Personal data protection -2	10%		1.08%
48					Part B, E10	Data QC team	25%		2.70%
49					Part B, E11	Data analysis team	25%		2.70%
50			Data management	18%	Part A, D1,3,4,5	Data management architecture	25%	11%	2.70%
51					Part A, D2	Data management Compliance	10%		1.08%
52					Part B, F1 and 3	Distributed cloud storage	20%		2.16%
53					Part B, F4	Type of Cloud storage	15%		1.62%
					Part B, F2	Data sharing mechanisms	15%		1.62%
54					Part B, F5	Historical data management	15%		1.62%
55	Data driven outcomes	20%	Synergistic data use within M/D	30%	Part A, E1 and 2	Identification of data gaps	40%	6%	2.40%
56					Part A, E3	Data exchange	60%		3.60%
57			Inter-agency data collaboration	30%	Part A, F1	Collaborations	50%	6%	3.00%
58					Part A, F2	Types of collaborations	50%		3.00%
59			Prescriptive Analytics	10%	Part A, G1	Prescriptive analytics	50%	2%	1.00%
60					Part A, G2	Frequency	25%		0.50%
61					Part A, G3	Modes	25%		0.50%
62			Good Practices	30%	Part A, H	Good practices	100%	6%	6.00%

Section III: Non-applicability verification

It was important to consider the non-applicability of certain questions or sub-parts of questions to certain Ministries/Departments and/or schemes. In order to take care of the same, wherever relevant, the NA option has been included in DGQI self-assessment questionnaire this time.

As this Index is based on self-assessment, Ministries/Departments are hence recommended to use the NA option with caution and select it only if the question/sub-part of the question is actually not applicable to them, owing to the nature of the intervention and not simply because it is currently not being practiced. For instance, if there is a social sector scheme targeted at improving nutrition status of school-going children but the MIS is capturing data only at district level and not at individual/facility(school) level, this does not mean that individual or facility level data reporting is not applicable. It is applicable as it is an individual-targeted scheme, however, it is currently not being practiced. Hence, the correct option to choose in such cases would be 'No' and not 'NA'.

Ministries/Departments are hence requested to arrive at the decision to mark NA based on their sector expertise and relevance of the question to the sector or programmatic intervention, after contemplating about the relevance and utility of adopting the stated data related practice in a forward-looking manner rather than only based on past and current practices.

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:

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 halfyearly, '3' at Quarterly, '4' at monthly/fortnightly/weekly and '5' at daily/realtime/near realtime level. If question is disabled, score '0'.	**Only need to change scoring if daily/ realtime/ near realtime/ 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'.

Q.No.	Question	Scoring mechanism	Way to handle NA
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 any one responses to Q7 is 'Yes', score '5'.	**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.
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, D2	Last mile linkage with PFMS	If response is 'Yes', score '5', if 'Partially' score '3', if 'No' score '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, atleast 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.

Q.No.	Question	Scoring mechanism	Way to handle NA
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.