Public Service Commission of Canada
Commission de la fonction publique du Canada

Formative Evaluation of the Implementation of the Public Service Commission Data Management Strategy

Canada

Table of Contents

1.0 Purpose

1. This report provides the results of the formative evaluation of the implementation of the Public Service Commission Data Management Strategy conducted by the Internal Audit and Evaluation Directorate.

2.0 Background

2. The volume of data collected, used and stored by all levels of government, private sector entities and Canadian citizens is growing exponentially. This is creating new opportunities for the federal public service to improve data management to support better-informed decision making. These opportunities come with the responsibility and expectation to use data wisely and manage them as a strategic asset. Proper data management also has the potential to promote innovation and improve how goods and services are provided to Canadians.

3. Good data management is essential for the Public Service Commission to fulfill its mandate, including the requirement to report annually to Parliament on the state of staffing throughout the federal public service. Data are also used to support the development and implementation of new policy directions, the provision of advice to departments and agencies, and the oversight of staffing activities across the system. In particular, the New Direction in Staffing, which came into effect on April 1, 2016, has increased the need for robust data management practices and processes to support decisions related to policy adjustments, system-wide staffing improvements, and effective oversight. Finally, data are also used to support evaluation work, research, and to forecast recruitment requirements – most recently through the macro-simulation model that forecasts the number of staffing appointments required for the federal public service to attain its objective of hiring 5,000 persons with disabilities over the next 5 years. The Public Service Commission responds to internal and external requests for data that it collects as part of its mandate.

4. There are numerous datasets divided among multiple databases managed by the 4 sectors of the organization. The data stored in these datasets are used to support various decision-making and reporting needs. The key datasets are as follows:

- Public Service Resourcing System (PSRS)
- Priority Information Management System (PIMS)
- Job-Based Analytical Information System (JAIS)

- Staffing and Non-partisanship Survey
- Employment Equity Databank (EEDB)
- Test Scoring and Results Reporting System (TSRR)
- Political Activities Self-Assessment Questionnaire
- Duty to Accommodate Assessment System

5. Data stewards, dataset managers and data users throughout the organization all share in the responsibility to ensure the consistency, quality, privacy and ethical use of data.

6. In April 2018, the responsibilities of the Chief Data Officer were assigned to the Vice-President, Oversight and Investigations Sector. Within the Data Services and Analysis Directorate, the Office of Data Management supports the Chief Data Officer with data management practices, awareness, learning, change management and communication. In addition the Chief Information Officer and the staff of the Information Technology Services Directorate, Corporate Affairs Sector, work with the Chief Data Officer and business owners to provide access to information technology systems and tools required to manage and use data and information.

7. In June 2018, the Executive Management Committee approved the Public Service Commission Data Management Strategy. The objective is to standardize data handling, to provide better services and products of value, and to support learning and capability building. The Data Management Strategy is defined by the Public Service Commission as "the process of planning or creating strategies/plans for handling the data created, stored, managed and processed by an organization"[1] and includes 3 pillars:

- **Data Management.** The goal is to establish an organizational standard for how data are handled, including governance, quality, security, privacy and risk management.
- **Data Infrastructure.** The aim is to establish modern data infrastructure and supporting practices.

[1] Public Service Commission Data Management Strategy: Components and Enablers. p. 1 4 – Formative Evaluation of the Implementation of the Public Service Commission Data Management Strategy

- **Data Users.** The goal is to support data users by leveraging the data infrastructure renewal to create modern channels and products, such as automated reports and open datasets. Another goal is to help provide value added products, services, learning and capacity building in the areas of data management and data analysis.

A focus on data management

In May 2018, the President sent a message to all staff to recognize Open Government Week:

"We continue to work on related initiatives such as our data management strategy, which will help us lay the foundation for Open government and strengthen our approach to data management. From policy making to oversight, to informing and improving our services, the need for quality data in support of evidence based decision making is central to our work."

Public Service Commission Data Vision

"The right data, the right way, in the right hands, at the right time"

---

8. The implementation of the Data Management Strategy relies on multiple stakeholders[2]. The main stakeholders are:

- The **Chief Data Officer** is responsible for implementing the Data Management Strategy and for driving the cultural change needed to manage data as a corporate asset. The Chief Data Officer's main goal is to ensure consistency in the quality, methodology and interpretation of data to uphold the integrity of the overall public service staffing system.

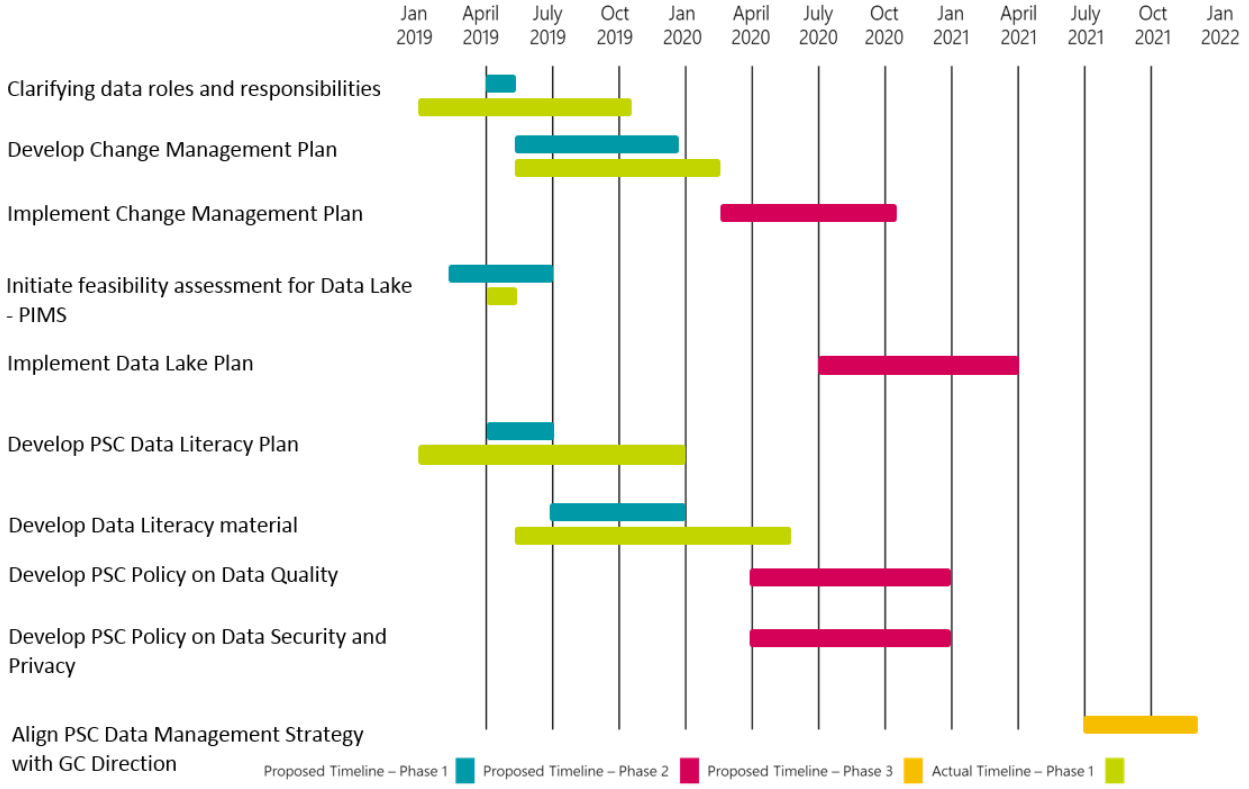
[2] These roles and responsibilities are from the document *Roles and Responsibilities with Respect to Data at the Public Service Commission* . The document was presented to the Executive Management Committee on August 28, 2019.

- The **Data Services and Analysis Directorate**, along with its Office of Data Management and Data Reporting, plays a key role in supporting the Chief Data Officer and providing data analysis support to data stewards and data users as well as to external stakeholders by fulfilling external requests for data.
- The **Chief Information Officer and Information Management Senior Officer** is responsible for the development and maintenance of the technical infrastructure relied on by the various business lines and ensures that data and information assets are created, stored, retained and disposed of in accordance with established IM guidelines and policies set by the Treasury Board Secretariat.
- **Data stewards** (the vice-presidents of each sector) and individual dataset managers are responsible for data requirements and specifications from a business perspective. They must ensure the security, privacy, quality and ethical use of the data they produce.
- **Dataset managers** are responsible for Public Service Commission datasets or are the business owners of a system that generates data (e.g., the Public Service Resourcing System, the Priority Information Management System).
- **Data users** are employees who use data in their day-to-day work. Each user has a business level responsibility for the maintenance, security and confidentiality of data assets. Their role is to understand the data collection methods, including the data retention and disposition periods, to perform data analysis, and to develop reporting methods and standards related to the use of the data that they hold.

9. Since June 2018, the Office of Data Management has been leading the implementation of the Data Management Strategy. This work is carried out by 4 full time equivalent (FTE) employees, who are supported by colleagues from the Oversight and Investigations Sector and from the Information Technology Services Directorate.

10. On January 20, 2020, the first full year of implementation was nearing completion. Data management roles and responsibilities had been documented and approved, and a feasibility study had been conducted on the implementation of a data lake (data repository) using the data of the Priority Information Management System. Since then, work has begun on initiatives to be delivered over the next 2 fiscal years. These include the development and implementation of standardized data practices, a data literacy plan, a data approval and dissemination protocol, a change management plan, and performance measures to track progress over time. The associated timelines are presented in Table 1 below.

Table 1 – Data Management Strategy – Key Timelines



Text version

Table 1 – Data Management Strategy – Key Timelines

Activities	**Proposed timeline – Phase 1**	**Proposed timeline – Phase 2**	**Proposed timeline – Phase 3**	**Actual timeline – Phase 1**
Clarify data roles and responsibilities	April 2019 – May 2019			January 2019 – November 2019
Develop Change Management Plan	May 2019 – December 2019			May 2019 – March 2020

Activities	**Proposed timeline – Phase 1**	**Proposed timeline – Phase 2**	**Proposed timeline – Phase 3**	**Actual timeline – Phase 1**
Implement Change Management Plan		March 2020 – November 2020		
Initiate feasibility assessment for Data Lake – PIMS (Priority Information Management System)	March 2019 – July 2019			April 2019 – May 2019
Implement Data Lake Plan		July 2020 – April 2021		
Develop Data Literacy Plan	April 2019 – July 2019			January 2019 – January 2020
Develop data literacy material	July 2019 – January 2020			May 2019 – June 2020
Develop PSC Policy on Data Quality		April 2020 – January 2021		
Develop PSC Policy on Data Security and Privacy		April 2020 – January 2021		

Activities	Proposed timeline – Phase 1	Proposed timeline – Phase 2	Proposed timeline – Phase 3	Actual timeline – Phase 1
Align PSC Data Management Strategy with GC direction				July 2021 – November 2021

11. The federal government has also made progress in the area of data management. In January 2018, the Clerk of the Privy Council asked that a data strategy be developed to position the public service to use data more strategically in order to provide the best possible advice to ministers. The Privy Council Office released the *Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service* in November 2018 The report asked departments and agencies to implement an appropriate data strategy for their line of business by September 2019 and to develop clear roles and responsibilities regarding data. The Treasury Board Secretariat has also released a new policy and directive on service and digital to communicate how departments and agencies should manage information and data. These new policy instruments will come into effect on April 1, 2020.

3.0 Objective, scope and methodology

Objective and scope

12. The evaluation objective was to assess and report on the implementation of the Public Service Commission Data Management Strategy one year into implementation.

13. This evaluation was included in the Public Service Commission Internal Audit and Evaluation Plan 2019-22 and covers the period from January 2018 to January 2020. It was carried out in accordance with the Treasury Board Policy on Results and assessed the relevance, effectiveness and efficiency of the implementation of the Public Service Commission Data Management Strategy.

14. The evaluation included a general review of the following datasets:

- Public Service Recruitment System (PSRS)

- Priority Information Management System (PIMS)
- Job-Based Analytical Information System (JAIS)
- Employment Equity Databank (EEDB)
- Test Scoring and Results Reporting System (TSRR)

Methodology

15. The methodology included multiple lines of evidence derived from the logic model (Annex 1) and evaluation matrix (Annex 2). The following methods were used to answer the evaluation questions:

- **Document and literature review.** This included reviewing documents from governance committees (i.e., Executive Management Committee, Data and Open Government Advisory Board, Data Lake Steering Committee, etc.), strategic and tactical federal public service and departmental policy and guidance documents (i.e., the Privy Council Office Data Strategy Roadmap, the Public Service Commission Data Management Strategy, data projects, data quality evaluations, minutes of meetings, communication plans, planning documents, surveys, etc.). External key documents were reviewed with varying purposes and levels of specificity (e.g., similar data management documents in other federal departments or other jurisdictions, and / or publications on data management). The literature review consisted of, among other things, documents that provide further insights and perspectives on relevant data management issues, such as reference books.
- **Interviews.** Interviews were conducted with various internal stakeholders to obtain further insights into data management practices and to assess the level of the organization's data management maturity. Among others, we met with employees from the Data Services and Analysis Directorate (including the Office of Data Management), Information Technology Services Directorate, and dataset managers from across the organization who are responsible for managing key datasets and systems.

- **Surveys.** Two surveys were administered as part of this evaluation. One survey was sent to data users and a separate survey was sent to dataset managers. The surveys were administered to obtain information and comments on various data related subjects including data access, roles and responsibilities, and data quality. The surveys combined open-ended and close-ended questions and were administered between November and December 2019. The data user survey was sent to 927 employees and had a completion rate of 36 percent (337 respondents). From that number, 202 respondents identified themselves as data users. The dataset manager survey was sent to 28 dataset managers identified by their sectors. We had a response and completion rate of 29 percent (8 respondents).
- **Data maturity model.** The evaluation methodology included the use of the Capability Maturity Model Integration Institute, an ISACA Enterprise's Data Management Maturity model. This model was used in the development of criteria and tools to conduct the work. The Data Management Maturity model provides standard sets of practices to evaluate organizational capabilities. The Data Management Maturity assessment will be presented and addressed separately from the evaluation.

4.0 Relevance

Alignment with government priorities

16. In 2018 the federal government set out to improve how data was managed and used across the system. In November 2018, the Privy Council Office published the Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service. The federal Data Strategy Roadmap contains 21 recommendations that should be implemented over the short to medium term by departments and agencies to ensure that their data management initiatives are relevant, scaled and adapted to specific business requirements. Of the 21 recommendations, 6 were deemed priorities for the short term. Table 2 highlights these 6 recommendations and the actions taken as of January 2020.

Table 2 – Actions to address the 6 most important recommendations of the roadmap

Federal Data Strategy Roadmap six most important recommendations	**Actions to address them**
By September 2019, all departments, agencies or portfolios have a data strategy in place appropriate to their line of business.	In June 2018, the Executive Management Committee approved the Data Management Strategy.
Provide greater clarity on who is in charge of data within individual organizations and for the government as a whole.	In 2018, the Vice-President, Oversight and Investigations was nominated as the Chief Data Officer. In addition, a document titled, Public Service Commission Roles and Responsibilities with Respect to Data was submitted to the Privy Council Office as required. This document identifies the roles and responsibilities for various data stakeholders.
Improve and develop overall standards and guidelines that govern how departments access, collect, use, safeguard and share data, and a clear process for developing and refining these over time.	The Data Management Strategy includes initiatives to improve available guidelines. These include the development of a dissemination protocol and a new data quality framework.
Clarify the governance around data to ensure that the Government of Canada manages	A data governance structure was developed, which documents the roles and responsibilities of the various data management stakeholders.

Federal Data Strategy Roadmap six most important recommendations	**Actions to address them**
valuable data assets for the public good.	
Improve recruitment and professional development practices to ensure that we have the skilled people we need to do data work in a digital environment.	The Office of Data Management is currently developing a data literacy plan and training modules for employees in accordance with planned actions in the Data Management Strategy. Data Services and Analysis Directorate also developed competencies for data analysts/scientists (EC-02 to EC-07)
Ensure we have the right information technology environment that allows skilled professionals to use the disruptive technologies that will support the ambitious agenda outlined in this report.	There is on-going collaboration occurring between Information Technology Services Directorate and Data Services and Analysis Directorate through the data governance model (Data & Open Government Advisory Board, Information Management /information Technology Committee, Executive Management Committee, etc.) and the IT prioritization process. This approach is continuously reviewed to ensure that the IT environment supports the successful implementation of the Data Management Strategy.

17. In January 2020, the Public Service Commission was on track to meet the expectations of the federal Data Strategy Roadmap. The Chief Data Officer has put a structure in place to help continually implement actions in keeping with federal public service goals. While the Data Management Strategy and related data management initiatives are aligned with federal government priorities, in light of rapidly changing data management practices and the upcoming Treasury Board digital and service policies, the Public Service Commission must continuously monitor the execution of its

strategy to maintain alignment with overall government priorities in this area. This is particularly important since under the new policy, it is the Chief Data Officer's responsibility to lead the data management functions.

5.0 Data governance

18. Governance was identified as a key element to support the proper management of data and information assets. The Public Service Commission approach to data management is decentralized: Business owners, supported by the Information Technology Services Directorate, are responsible for managing the systems they use to administer policies, programs and services. This means that data stewards and dataset managers play a very important role in organizational data management.

19. In the Data Management Strategy, the Public Service Commission defines data governance as "a partnership between business owners, data experts and technology stewards." Over the last two fiscal years, governance has been changed in a manner that supports data management practices. The Information Management/Information Technology Committee (IM/ITC) and the Integration Committee (IC) began holding joint meetings in September 2017 to provide guidance and oversee the implementation of data related initiatives. The objective was to strengthen partnerships between business owners, data experts and information technology staff. The Data and Open Government Advisory Board was also created to provide advice, guidance and recommendations to the Chief Data Officer, the departmental Open Government Champion as well to the Information Management Senior Officer (IMSO) regarding the development of the Data Management Strategy and the Open Government agenda. Now that these initiatives are underway, the Data Services and Analysis Directorate is in the process of shutting down the Data and Open Government Advisory Board and leveraging the joint meetings of the IC and IM/ITC.

20. The Data Management Strategy and the Public Service Commission Roles and Responsibilities with Respect to Data document were developed and approved through established governance processes and submitted to the Privy Council Office in accordance with the Federal Data Strategy Roadmap expectations.

21. The evaluation found areas where data governance could be improved, including collaboration and stakeholder relationship management. A number of interviewees told us that they were not concerned with the objectives of the Data Management Strategy, an important finding given the decentralized approach to managing datasets and

systems. Furthermore, some interviewees had questions regarding the Data Lake project and the level of consultation that had taken place over the past 2 fiscal years. It wasn't clear why the Data Lake had not been continued or funded in the IM/IT Plan 2019-20 and had been withdrawn for consideration during the development of the IM/IT Plan 2020-21, given that the project is a priority for the organization and the Data Management Strategy.

22. Dataset managers interviewed noted that the current data governance model does not allow for sufficient participation in the project design phase of initiatives. Similar concerns were raised regarding decision making for the Data Management Strategy implementation: There is a perceived lack of consultation, except for when items are ready to be presented to the main governance committees. Several interviewees indicated that data management decisions seem to be made in silos without appropriate consultation. In this regard, there may be work to be done to increase awareness of data management decision-making processes in order to strengthen collaboration among key stakeholders.

23. In addition, collaboration could be improved between Information Technology Services Directorate and Data Services and Analysis Directorate to support data initiatives. The two most frequent examples provided by interviewees were the Data Lake project and the general state of collaboration between the two groups. During interviews with Information Technology Services Directorate staff, Data Services and Analysis Directorate staff, and dataset managers, the evaluation noted a general lack of awareness of the realities and complexities of the work of colleagues. This lack of a common understanding has led to increasing levels of frustration, creating false expectations, and decreasing collaboration in some instances. Essentially, interviewees identified an opportunity to improve collaboration to support the successful implementation of all the elements of the Data Management Strategy.

24. Other departments have put various initiatives in place to manage data related issues from an organizational perspective. In January 2020, the Privy Council Office released a document titled: Organizational Data Strategies – A Review that outlined some good practices related to data governance. These included the following: creation of a departmental working group on data to support decision-making processes (Transport Canada) and the establishment of communities of practice to support better data governance (Canadian Nuclear Safety Commission, Global Affair Canada, Public Safety Canada). The implementation of some of these practices would be important for the Public Service Commission to consider given the decentralized management of

systems, the important role that data stewards and dataset managers have, and the need for the Chief Data Officer to establish and integrate common operating principles for how data are managed.

25. **Recommendation 1:** It is recommended that the Chief Data Officer, in collaboration with the Chief Information Officer and Data Stewards, review data governance to ensure it supports internal collaboration in the implementation of the Data Management Strategy resulting in an organization-focused perspective on data management.

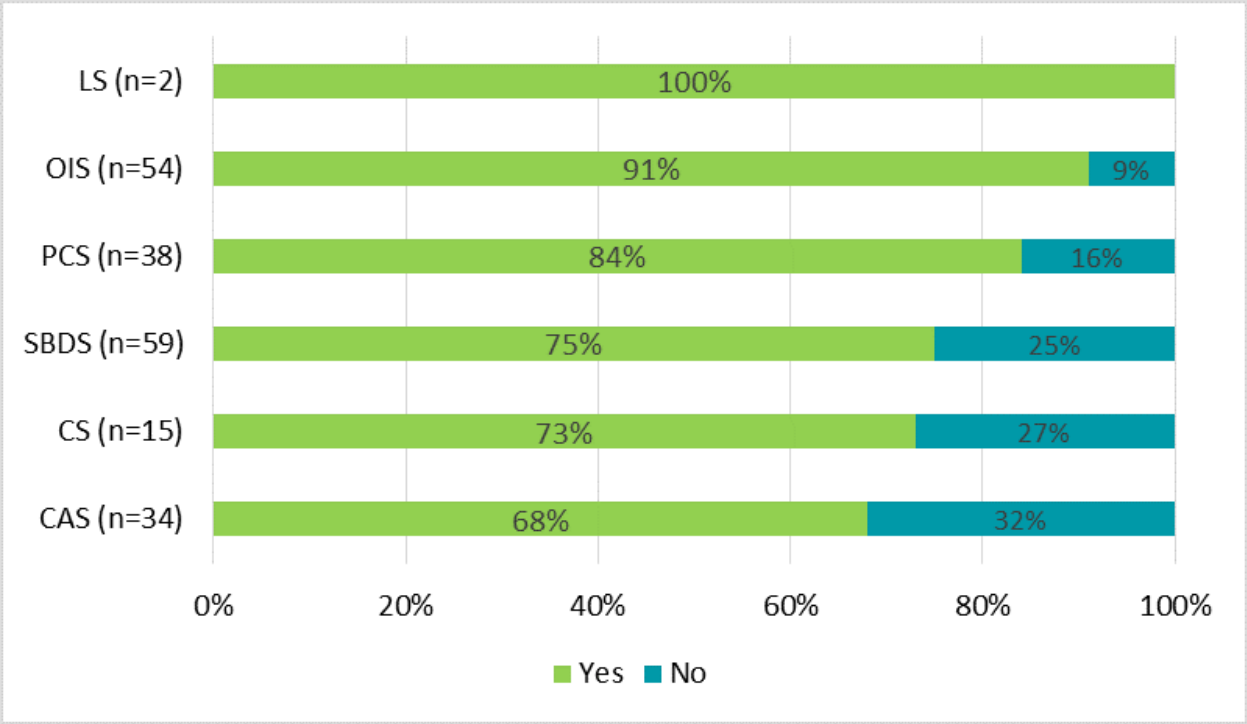
5.1 Data management

Roles and responsibilities of the Chief Data Officer

26. The Office of Data Management is focused on raising awareness of the Chief Data Officer and the role of the position. The Chief Data Officer is leading a cultural change within the organization that will enhance the management of data as a valuable corporate asset. Some of the Chief Data Officer's key responsibilities include ensuring consistency in data quality, developing and improving methodologies, and creating synergies across the organization to advance the implementation of the Data Management Strategy. The Chief Data Officer also has a key role in ensuring the consistency of data published by the Public Service Commission and reviewing data sent to external stakeholders related to the public service wide staffing system.

27. In November 2018, the Data Services and Analysis Directorate sent a survey to employees to develop a baseline assessment of the state of data management at the Public Service Commission. Overall, 54% of respondents were aware that the organization has a Chief Data Officer. The more recent results from data users surveyed for this evaluation demonstrate that awareness building is having a positive impact: 80 percent were aware that the Public Service Commission has a Chief Data Officer. The level of knowledge varied across the organization (See Exhibit A).

Exhibit A – Employees' awareness of a Chief Data Officer (n=number of respondents)



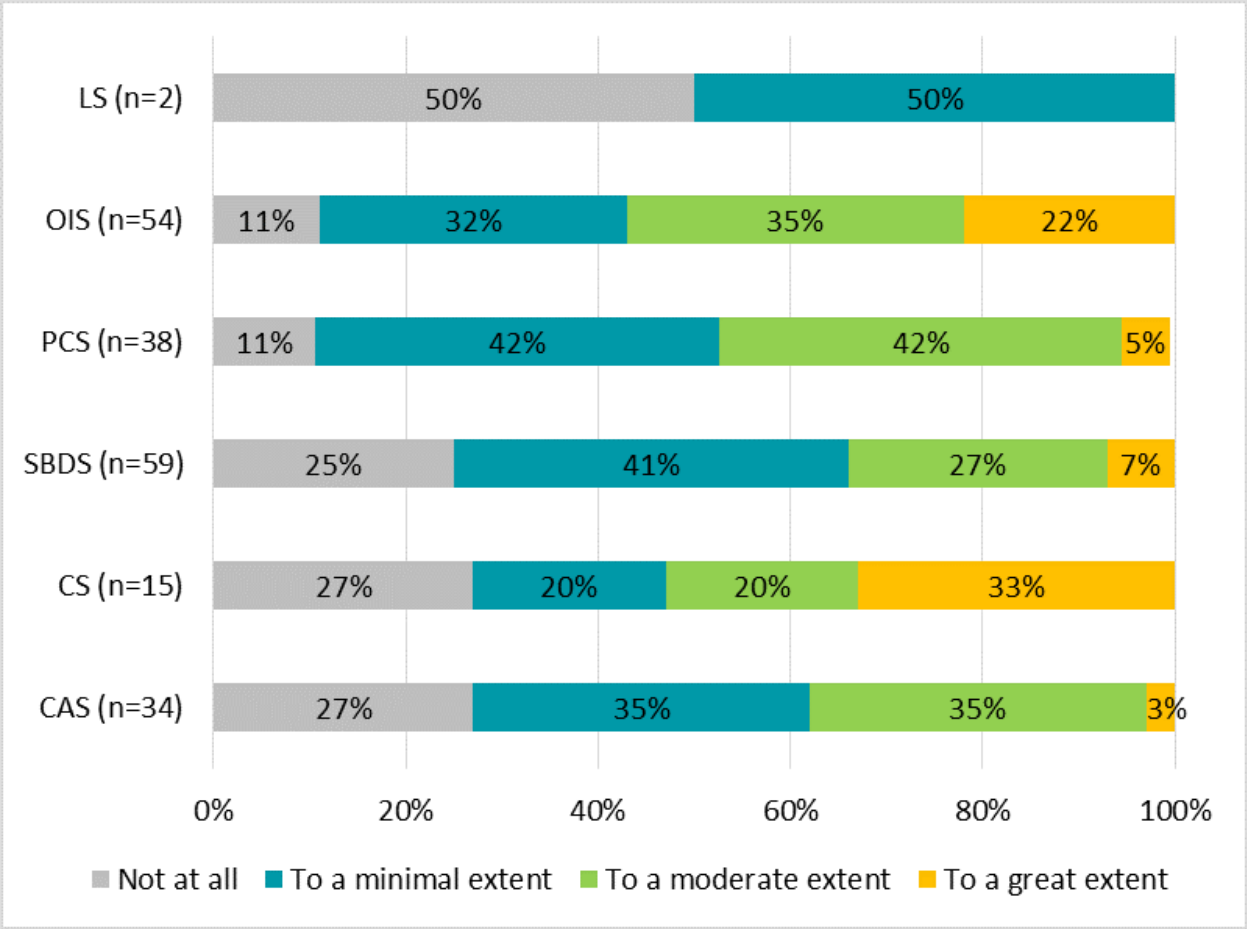
Text version

Employees' awareness of a Chief Data Officer (n=number of respondents)

Category	CAS (n=34)	Corp Sec (n=15)	SBDS (n=59)	P&C (n=38)	OIS (n=54)	Legal (n=2)
Yes	68%	73%	75%	84%	91%	100%
No	32%	27%	25%	16%	9%	0

28. In November 2018, the Data Services and Analysis Directorate also asked employees how familiar they were with the roles and responsibilities of the Chief Data Officer. Among employees who knew there was a Chief Data Officer, 40% knew of the position to a moderate or great extent. The evaluation survey found a similar level of awareness. Overall, 45% of data user respondents knew the roles and responsibilities of the Chief Data Officer to a moderate or great extent. Exhibit B provides a breakdown by sector.

Exhibit B – Data users' knowledge of the Chief Data Office roles and responsibilities (n=number of respondents)



Text version

Data users' knowledge of the Chief Data Officer's roles and responsibilities (n = number of respondents)

Category	**CAS (n=34)**	**Corp Sec (n=15)**	**SBDS (n=59)**	**P&C (n=38)**	**OIS (n=54)**	**Legal (n=2)**
Not at all	27%	27%	25%	11%	11%	50%
To a minimal extent	35%	20%	41%	42%	32%	50%

Category	CAS (n=34)	Corp Sec (n=15)	SBDS (n=59)	P&C (n=38)	OIS (n=54)	Legal (n=2)
To a moderate extent	35%	20%	27%	42%	35%	0
To a great extent	3%	33%	7%	5%	22%	0

29. Our survey results indicate that the Office of Data Management should continue awareness building efforts, such as the interview with the Chief Data Officer that was published in the December 2019 edition of Update, the Public Service Commission's online newsletter. This will help normalize the understanding of the roles and responsibilities of the Chief Data Officer across the organization and lead to better understanding of processes and procedures related to analyzing and publishing data.

Roles and responsibilities regarding data management and use

Dataset managers

30. Dataset managers[3] have an important role related to managing data that they are responsible for. Their general roles and responsibilities are as follows:

- **Data quality management.** Manage, control and monitor quality throughout the data lifecycle by ensuring that data are fit for their intended use and easy to access.
- **Policy compliance.** Ensure that data are managed and protected in accordance with policies, principles and standards, verify adherence to the Public Service Commission Data Approval and Dissemination Protocol (including the Chief Data Officer review process) and governance (where appropriate) before data are released.

[3] These are the responsibilities of the data stewards, identified as the Vice-Presidents. These responsibilities are delegated to those who manage day to-day activities, i.e., the dataset managers.

- **Process.** Define the business requirements and validation rules for the specific sector's data to allow data users to effectively manage and implement data quality measures that support reporting.
- **Support.** Help establish data definitions and business rules.
- **Consultation and collaboration.** Consult on data related issues and opportunities with other stakeholders such as the Chief Data Officer, the Chief Information Officer, Access to Information and Privacy and/or Legal Services.
- **Guidance.** Encourage discussion and debate on data.
- **Openness and transparency.** Actively work with the Open Government Secretariat to publish program specific open datasets in collaboration with the Information Management office and Access to Information and Privacy office.

31. Dataset managers were sent a survey asking whether they were aware of and understood their responsibilities and were confident in their ability to apply their knowledge to fulfill their roles and responsibilities. From a statistical point of view, the number of respondents is too low to be representative, but the results provide certain insights that are useful for this evaluation. Dataset managers' awareness and understanding of their roles and responsibilities could be improved. More than half of the 8 respondents indicated that they were not aware of and/or did not understand their responsibilities.

32. Dataset managers rated their confidence in their ability to fulfill their roles and responsibilities higher than their awareness and understanding of these aspects. A majority of respondents reported that they were confident to a moderate or a great extent that they could fulfill their tasks and none indicated that they were not confident not at all.

33. This elevated confidence in ability combined with less positive results regarding understanding of the dataset manager roles and responsibilities approved by the organization was also reflected in interviews and comments from survey respondents. The evaluation found that dataset managers approach their roles and responsibilities pragmatically. In interviews and comments, most reported knowing how to achieve results by following informal processes implemented over the years. In addition, there was a perceived lack of clarity regarding the expectations related to the tasks outlined in the document on Public Service Commission data management roles and responsibilities. Some expressed a need for more support and documentation to better

understand what is expected of them. In the absence of this clarity they continue to manage their datasets as has been done in the past.

34. Furthermore, most dataset managers interviewed stated that they had not seen any documentation outlining their roles and responsibilities. When informed of the documented roles and responsibilities, some dataset managers did not accept that all the tasks were applicable, either because of their limited access to data systems or because of their teams' limited skillsets.

35. In this environment, the work performed by dataset managers varies according to each dataset, which partially explains why it may be difficult for them to identify with generic definitions of tasks. For example, some dataset managers or their teams have the ability to extract, interpret and analyze data, while others rely on employees from Data Services and Analysis Directorate and Information Technology Services Directorate employees for those tasks. These differences arise from the historical approach of managing each dataset independently rather than using an organization wide approach to data management and use, which is one of the Data Management Strategy's objectives.

36. Dataset managers had concerns regarding how data are extracted from systems and how they are interpreted by business owners and clients to support decision making or reporting. Similarly, some dataset managers informed us that they were not familiar with all the data they held, nor with the types of data that could be reported on. Dataset managers are not always aware of the impact that changes to systems will have on generating performance data to support reporting.

37. There are concerns regarding the interpretation of data, more specifically how data are used to represent certain observations in Public Service Commission reports. Data managers and IT specialists informed us that work needs to be done to enhance data literacy. This would promote better use of data, data analysis, and reporting to support better decision making. Currently, the Office of Data Management is creating training and learning modules for data users in order to better equip them to fulfill their responsibilities.

Data users

38. Data users play an important role in generating, analyzing, using and reporting on data. The general roles and responsibilities for data users are:

- ensuring the ethical use of data (by protecting personal information and preserving confidentiality)
- maintaining an objective use of data to support policies, programs and oversight activities
- helping identify and correct input/output errors in data
- confirming that data / information products are aligned with previously released products and messages
- liaising with subject matter experts to ensure that proper methodology is applied and confidentiality is preserved
- leveraging data to answer business questions or resolve business issues
- using data to tell stories to stakeholders and clients
- maintaining and further developing data skills / competencies
- raising questions and fostering discussions on data

39. The evaluation survey asked data users the reasons why they required this information to perform their work. The top three uses of data at the Public Service Commission were deemed to be data analysis and reporting for internal and external clients, supporting senior management decision making, and establishing trends and forecasts. The results are outlined in Table 3.

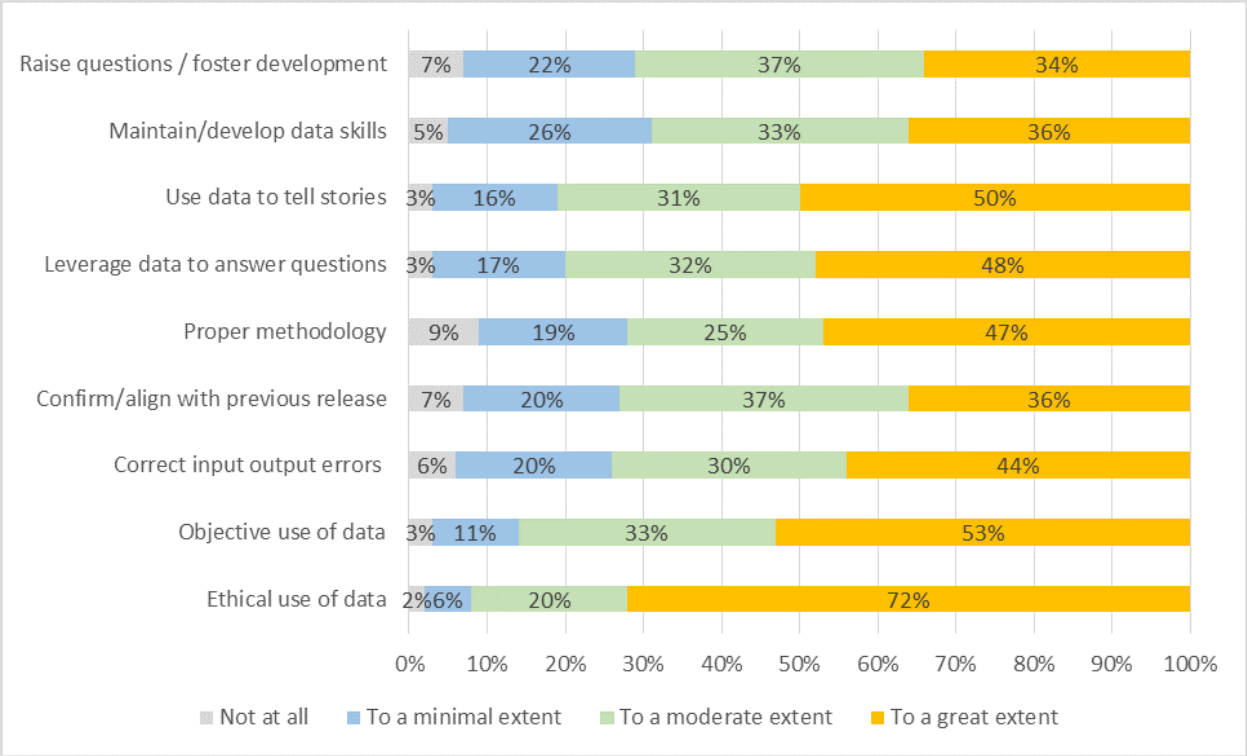
Table 3 – Employee use of data (202 respondents)

Use of data at work	**#**	**CS**	**PCS**	**OIS**	**CAS**	**SBDS**	**LS**
Data analysis and reporting for internal or external clients	123	12	25	34	19	33	0
Supporting senior management decision making	93	9	17	26	14	26	1
Trends and forecasting analysis	68	5	10	18	12	23	0
Administration (access to forms, emails, agendas, etc.)	67	4	15	15	11	21	1

Use of data at work	#	CS	PCS	OIS	CAS	SBDS	LS
Reporting to the President (reports, briefing notes, speaking points)	63	5	12	20	11	14	1
Collect and/or analyze information on candidates and/or employees	61	5	10	14	8	23	1
Program/system improvement/evaluation	59	5	10	15	11	17	1
Planning activities (oversight projects, outreach, tests sessions, etc.)	55	5	10	17	7	16	0
Creation/update of datasets/databases	50	2	9	14	8	16	1
Reports on the staffing system (job evaluations, appointments, trends, etc.)	49	4	10	14	9	12	0
Performance measurement	46	5	6	15	5	14	1
Audit and investigation	40	4	9	13	7	7	0
Finances (billing, budget, pay problems, etc.)	29	3	2	6	3	13	1
IT-related system solutions (development, testing environment, support)	28	3	2	6	3	13	1

40. The evaluation survey sent to data users asked whether they were aware of their responsibilities, understood them, and were confident in their ability to apply their knowledge to fulfill their roles and responsibilities. A majority of data users responded that they were aware of and understood their roles and responsibilities and that they were confident in their ability to perform their functions to a moderate or great extent. Exhibit C highlights data user awareness.

Exhibit C – Data users' awareness of their roles and responsibilities (202 respondents)



Text version

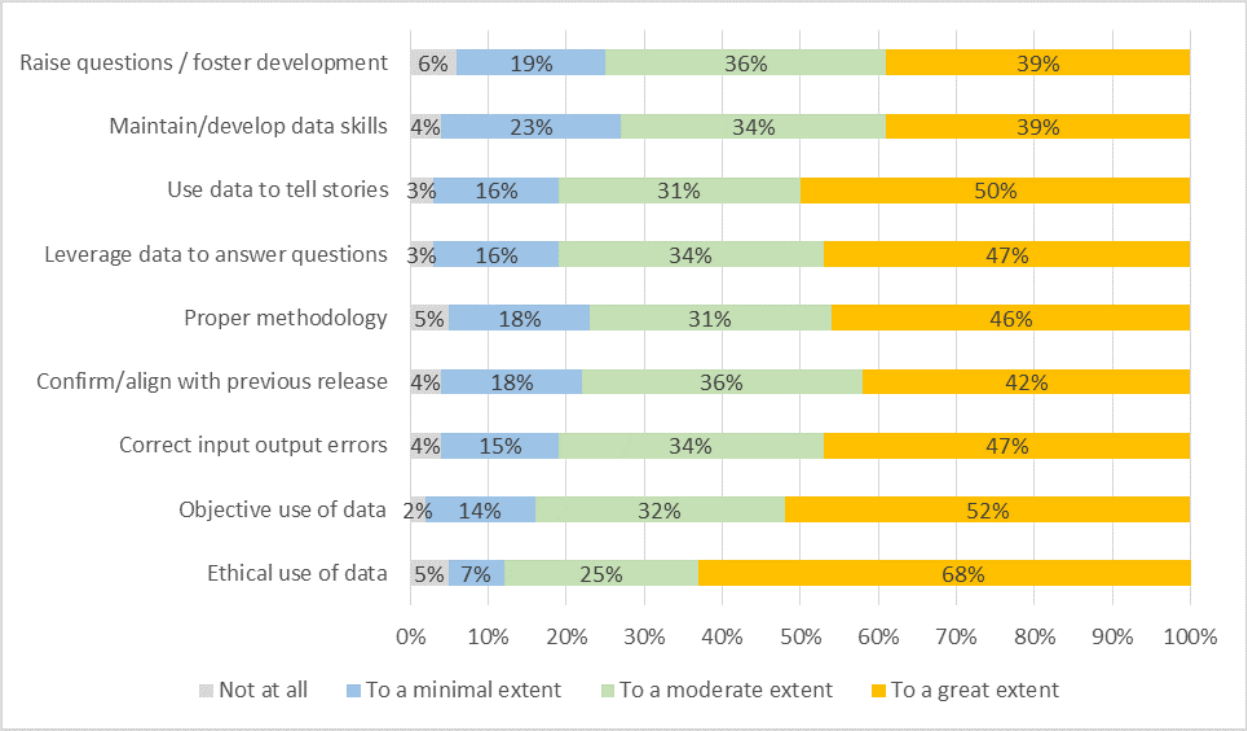
Data users' awareness of their roles and responsibilities (202 respondents)

Category	Not at all	To a minimal extent	To a moderate extent	To a great extent
Ethical use of data	2%	6%	20%	72%
Objective use of data	3%	11%	33%	53%
Correct input/output errors	6%	20%	30%	44%
Confirm/align with previous release	7%	20%	37%	36%

Category	Not at all	To a minimal extent	To a moderate extent	To a great extent
Proper methodology	9%	19%	25%	47%
leverage data to answer questions	3%	17%	32%	48%
Use data to tell stories	3%	16%	31%	50%
Maintain / develop data skills	5%	26%	33%	36%
Raise questions / foster development	7%	22%	37%	34%

41. The majority of data users responded that they had a very good awareness of their roles and responsibilities regarding data usage. Between 71 and 92 percent responded to a moderate or great extent for all questions. A majority of data users also considered having a good understanding of their roles and responsibilities regarding data usage (See Exhibit D). Between 75 and 92 percent of respondents answered to a moderate or a great extent to all questions.

Exhibit D – Data users' understanding of their roles and responsibilities (202 Respondents)



Text version

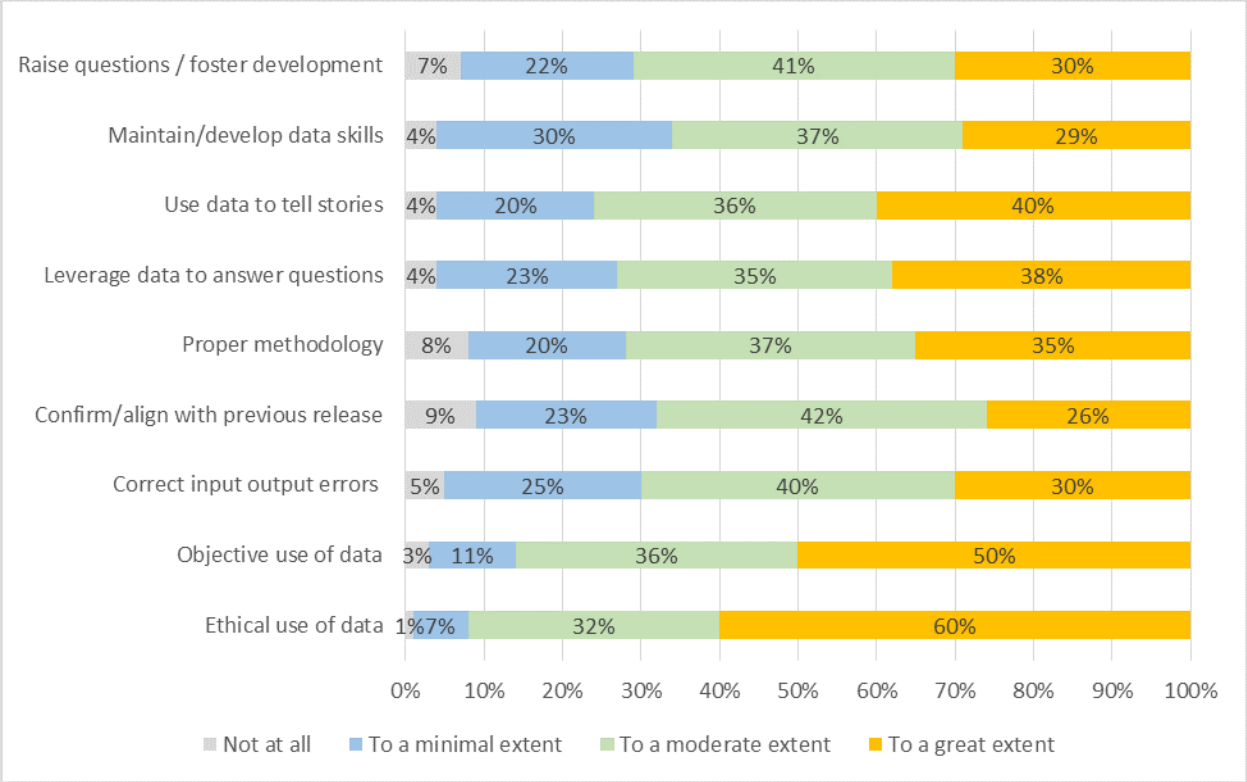
Data users' understanding of their roles and responsibilities (202 respondents)

Category	Not at all	To a minimal extent	To a moderate extent	To a great extent
Ethical use of data	5%	7%	25%	68%
Objective use of data	2%	14%	32%	52%
Correct input/output errors	4%	15%	34%	47%
Confirm/align with previous release	4%	18%	36%	42%

Category	Not at all	To a minimal extent	To a moderate extent	To a great extent
Proper methodology	5%	18%	31%	46%
leverage data to answer questions	3%	16%	34%	47%
Use data to tell stories	3%	16%	31%	50%
Maintain / develop data skills	4%	23%	34%	39%
Raise questions / foster development	6%	19%	36%	39%

42. Data users also indicated a high level of confidence in their ability to execute their roles and responsibilities. The results were very positive (see Exhibit E below).

Exhibit E – Data users' confidence in performing their roles (202 respondents)



Text version

Data users' confidence in performing their roles (202 respondents)

Category	Not at all	To a minimal extent	To a moderate extent	To a great extent
Ethical use of data	1%	7%	32%	60%
Objective use of data	3%	11%	36%	50%
Correct input/output errors	5%	25%	40%	30%
Confirm/align with previous release	9%	23%	42%	26%

Category	Not at all	To a minimal extent	To a moderate extent	To a great extent
Proper methodology	8%	20%	37%	35%
leverage data to answer questions	4%	23%	35%	38%
Use data to tell stories	4%	20%	36%	40%
Maintain / develop data skills	4%	30%	37%	29%
Raise questions / foster development	7%	22%	41%	30%

43. Those who responded “To a minimal extent” or “Not at all” provided comments to support their answers. These respondents identified certain challenges regarding their understanding of and ability to fulfill their roles and responsibilities. They reported lacking information to determine a clear direction, including on data processing and data sharing rules. Others expressed a desire to obtain more training and guidance on their responsibilities, particularly as they relate to analyzing data.

44. Overall, data users rated very positively their awareness and understanding of their roles and responsibilities with regard to using data and were confident in their ability to apply their knowledge.

5.2 Data quality, security and privacy

Data quality

45. The Data Management Strategy highlights the importance of having quality data to support informed decisions and reporting. The Office of Data Management plans to develop various initiatives in 2020 and 2021 to drive the improvement of data quality, including a policy on data quality and training.

46. In 2012, the Data Quality Management Framework was established in preparation for the BI-EDWs. The framework includes information on accuracy, relevance, timeliness, accessibility, interpretability, coherence, privacy and security. It was developed to identify and address data quality gaps prior to loading the data into the EDW. The framework has not been consistently applied by dataset managers and some of them are not aware of its existence.

47. While the Public Service Commission has a data quality management framework, most systems do not have specific data quality assurance processes or documented procedures in place to identify and address data discrepancies or errors. In the absence of documented data quality assurance practices, a number of dataset managers and data users have been performing certain data quality tasks. All dataset managers interviewed had informal quality assurance practices in place to manage their datasets. These practices included comparing data extracts with information from other datasets and performing manual verification processes to validate the data they manage with that of other systems. For example, the Data Services and Analysis Directorate validates the figures produced and the interpretation of the data, and ensures that no personal information is shared. This is time consuming work, and there is limited documentation on the various approaches used to manage dataset quality and mitigate data quality risks.

48. Some of the systems currently in use present data extraction challenges. This is the case for systems that were not built to generate data reports for analytical purposes as well as those that were created by merging multiple systems. For example, the evaluation team was informed that the volume of data in the Test Scoring and Results Reporting System (TSRR), has made it difficult for employees in the past to address data errors and inconsistencies in the system and verify whether data are accurate and up to date for reporting purposes. Senior management from the Personnel Psychology Centre stated that once the Apollo system is up and running, these types of data reporting issues will be resolved. Also, regarding the current recruitment platform, the Public Service Resourcing System, which was created by merging multiple systems (including Publiservice), interviewees questioned whether data were properly disposed of prior to this merger. Senior management responsible for the Public Service Recruitment System indicated that a Data Retention and Disposition project was approved in 2017. The project included a five-year plan to dispose of information no longer required to be retained by phasing out applications transferred to the system. Nevertheless, employees using data generated from legacy systems must ensure that the data are reliable and

must understand the limitations prior to producing analytical documents or reports. This can help mitigate the risk to data quality from the lack of proper data disposition and cleansing in the past.

49. As of January 2020, systems including the Job-Based Analytical Information System, the Test Scoring and Results Reporting System, and the Public Service Recruitment System are currently being upgraded or are the subject of plans for replacement with systems that have more advanced functions to support the generation of data for reporting purposes. As such, consideration should be given to developing and documenting data quality assurance procedures up front to support improved data management practices once new or upgraded solutions are in place.

Data cleansing practices

Data cleansing: Data cleansing or scrubbing transforms data to make it conform to data standards and domain rules. Cleansing include detecting and correcting data errors to bring the quality of data to an acceptable level (DAMA – DMBOK, Data Management Body of Knowledge (2nd edition), p. 471).

50. There has been a lack of consistent data cleansing and data disposition. Up until now the approach has been to maintain all production system data, including data that no longer have business value. Very little work has been done to implement controls to prevent data entry errors, correct data in source systems, or improve business processes regarding the collection, use and disposition of data. The evaluation team was made aware of some data cleansing work underway, but this mostly concerns post-extraction data that is cleaned for analytical purposes.

51. A significant debt has accumulated in this area following decades of very limited data cleansing and inconsistent data disposition practices. Over time, some data have been merged from one system into another without proper cleansing. This presents challenges in developing new systems or upgrading current systems. For example, the Application Rationalization Project experienced delays because for reasons of data quality, more time was needed to migrate data from legacy systems into newly developed systems because of data quality.

Data dictionary and metadata

Data dictionary: A data dictionary describes data in business terms and includes other information needed to use the data (e.g., data types, details of structure, security,

restrictions). (DAMA – DMBOK, Data Management Body of Knowledge (2nd edition), p. 402).

Metadata: Metadata describes what data an organization has, what it represents, how it is classified, where it came from, how it moves within the organization, how it evolves through use, who can and cannot use it, and whether it is of high quality. (DAMA – DMBOK, Data Management Body of Knowledge (2nd edition), p. 27).

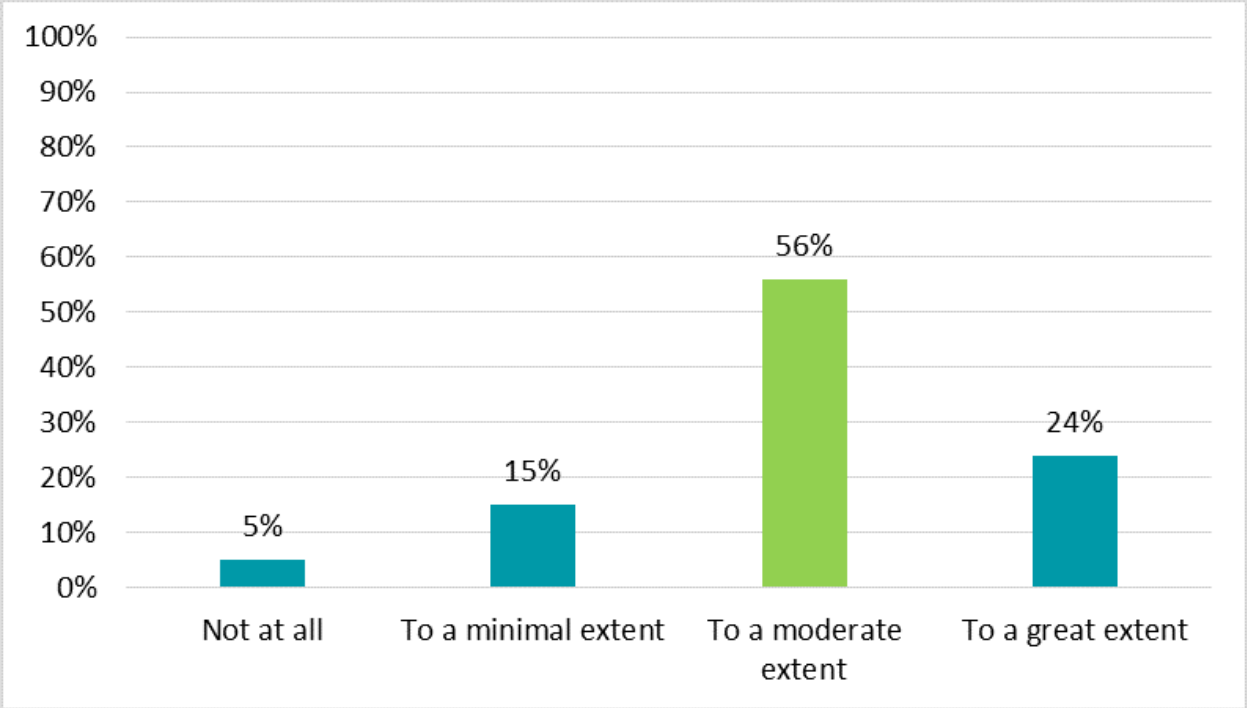
52. Dataset managers identified a lack of proper data management documentation and processes. As of January 2020, there is no standardized data dictionary and metadata. Metadata do not appear to code as the same fields between different data systems and databases. This presents challenges in analyzing data and developing reports. According to literature, inconsistencies with respect to the construct of metadata impact the ability of data users and programs to extract accurate and consistent information on variables used for reporting and decision-making purposes.

53. The lack of consistency regarding data cleansing, disposition and proper documentation means that extra attention must be paid to understanding data, their original purpose for collection and what employees are trying to use them for. Corporate knowledge is often shared between analysts, a practice that increases the risk of data being used for incorrectly. The lack of proper data documentation also creates situations where analysts may draw erroneous conclusions based on the data used to perform their analysis. In interviews, some stakeholders raised the necessity of improving employee data literacy to increase their ability to interpret, analyze and report on data.

Confidence in data quality

54. The evaluation asked data users about their confidence in the quality, privacy and ethical use of data. The survey found that only 24%of data users had a great amount of confidence in our data, with 56% responding that they were confident to a moderate extent. This shows that while confidence was generally good, there were some doubts among data users regarding the overall quality of Public Service Commission data (see Exhibit F below). This result is interesting in light of the interviewees' comments concerning weaknesses in data quality administration across the organization.

Exhibit F – Data users' confidence in data quality (202 respondents)



Text version

Data users' confidence in data quality (202 respondents)

Confidence in data quality	**Percentage of respondents**
Not at all	5%
To a minimal extent	15%
To a moderate extent	56%
To a great extent	24%

55. While the level of confidence in data quality was relatively high, the 20% of respondents who answered "To a minimal extent" or "Not at all" provided the following feedback outlining their concerns:

- Lack of clearly defined methodology to treat data

- Lack of documentation concerning the approaches to be used to clean data or check for data integrity issues
- Lack of metadata and data dictionary
- Unreliable data in systems used for analytical purposes, such as the federal public service pay system
- Low level of data literacy, which contributes to data sometimes being used beyond their intended purpose, leading to interpretation inconsistencies

56. One of the reasons for the high level of confidence in data quality provided to the evaluation team was the excellent work performed by the Data Services and Analysis Directorate's reporting and analytics teams and other data specialists within the organization. In interviews and survey comments, evaluation participants stated that the Data Services and Analysis Directorate data analysts provide necessary support to data users and help mitigate some of the weaknesses observed regarding data cleansing, metadata and associated processes. Essentially, confidence in data quality is high because the Data Services and Analysis Directorate devotes a great deal of time ensuring that data used and published are accurate.

57. The data quality issues observed in this evaluation had been identified in previous studies and reviews. A number of data quality assessments were performed between 2013 and 2016 on key datasets. The recommendations made in these reports highlight issues similar to those expressed by dataset managers and data users during this evaluation. Data quality has been a topic of discussion over the past few years; steps to address the accuracy and coherence of our data are identified in the Data Management Strategy.

Privacy, security and ethical use of data

58. The ethical use of data is another key element of the federal Data Strategy Roadmap. Roles and Responsibilities with respect to data at the Public Service Commission defines the ethical use of data as "protecting personal information and preserving confidentiality throughout the data lifecycle and being respectful of consent of individuals, the organizational mandate and the intended use of data." Ensuring employees use data ethically is a core principle of the Data Management Strategy. This is important because the Public Service Commission manages databases that include the personal information of employees across the federal public service and candidates that apply through recruitment campaigns, such as the Post-Secondary Recruitment Program and Federal Student Work Experience Program.

59. The Public Service Commission policy suite in this area has been in place for a decade. The Privacy Policy Model and Privacy Governance Model were approved in 2010 and the Security Policy was approved in 2011. The evaluation found that, in addition to being somewhat dated, existing policies related to data privacy and security were not being consistently applied across the organization. None of the dataset managers interviewed mentioned using these documents and none were even aware of their existence. The Access to Information and Privacy (ATIP) office is currently developing a Privacy Management Policy for the organization that will replace the above policies.

60. All the dataset managers interviewed had established informal protocols to mitigate risks related to the protection of personal information and privacy when data from databases are extracted or analyzed. Formal guidance in this area has existed for more than 10 years, but there is a lack of awareness of these policies and concerns regarding the management of privacy and security risks. For example, data managers for one dataset explained that they require signed consent in order to process data requests, regardless of whether it is for internal or external stakeholders, and that this rule is rigorously enforced. Since 2017, the Open Government Secretariat has established, documented and consistently applied a formal process to protect confidentiality when releasing data or information assets on the Open Government Portal. However, there is a lack of documentation of this procedure and it is not consistently applied across the Public Service Commission. Analysts also informed the evaluation team that access to systems and files is not always properly controlled. This is a risk to the organization, since very sensitive information, such as the social insurance numbers of all public service employees, is available in our systems.

Privacy impact assessments

61. In 2019–20, the Executive Management Committee approved a privacy impact assessment (PIA) five-year work plan to help address the knowledge gap related to potential privacy risks to programs and activities at the Public Service Commission. Privacy impact assessments are conducted to document risks and develop appropriate protocols and standard operational procedures with respect to the handling of personal information.

62. Some of the datasets reviewed as part of this evaluation have subject to privacy impact assessments. For example, in the 2015 assessment of the Public Service Recruitment System, it was identified that the system lacked the technical capability required for the disposition (destruction) of personal information following the

mandatory five-year retention period. A decision has been made to carry out the Government of Canada Jobs Transformation Project to find a replacement solution for the system. The solution should have the functionalities required for proper management of personal information in accordance with the privacy impact assessment recommendations. It should be noted that this problem is not limited to the Public Service Recruitment System: According to interviewees from Information Management, managing data lifecycles across the Public Service Commission datasets is an ongoing challenge. Due to the lack of technical functionalities in certain systems, information and date may not be properly disposed of in accordance with identified retention periods. The lack of procedures to dispose of personal information within the specified periods may present a risk to the organization with respect to data privacy and personal information.

63. **Recommendation 2.** It is recommended that the Chief Data Officer, in collaboration with the Chief Information Officer and Data Stewards review, update, and document data quality policies and procedures to enable proper data quality assurance, data disposition, and data privacy practices across the Public Service Commission.

5.3 Data infrastructure

64. The current information technology environment is comprised of a mix of new technologies and legacy systems. There is a significant amount of aging information technology infrastructure that contains an important amount of data used on a daily basis. The age of some of the systems presents challenges and risks to the implementation of the Data Management Strategy and must be taken into account when considering some of the limitations that may prevent the intended outcomes from being fully achieved. It is also important to note that improving data infrastructure will entail collaboration between the Information Technology Services Directorate and Shared Services Canada in many cases.

Data repository

65. One of the key objectives of the Data Management Strategy is to develop modern data infrastructure and supporting practices that strengthen the connection between data sources and data users. The Data Management Strategy underlines the need for new infrastructure solutions to improve data storage, integration, accessibility, reporting and dissemination across all databases and systems. One important component is the proposed establishment of a data lake (data repository). The following business needs for establishing a data repository were identified:

- A storage solution for legacy data from key program areas that are set to undergo significant renewal (the Priority Information Management System, the Public Service Recruitment System, etc.)
- An alternative to using a live system to retrieve data, which can result in overloading applications, necessitate system reboots and lead to query timeouts
- A solution to the lack of point-in-time traceability, which reduces the analytical capacity of business areas;
- Increased accessibility of data to support the growing demand for more complex data analytics and research
- The protection of privacy and the security of personal information during the development of the infrastructure.

66. The Data Lake project was approved by the Executive Management Committee in November 2017 and was identified as a key component of the Data Management Strategy. It was planned that the Data Lake solution would be in place by July 2020. This was to be followed by a gradual integration of data and the development of supporting policies until 2022. A steering committee was established in fall 2018 to facilitate a multisector approach to the development of this data infrastructure and to provide advice, feedback and oversight of the project scope, schedule, cost, issues and risk management.

67. Time and resources have been invested in the Data Lake project. March 31, 2019, saw the successful completion of the Phase 1 feasibility assessment of the establishment of a data lake at the Public Service Commission using the Priority Information Management System. This preparatory work included the identification of file formats, the identification and mapping of data, a data dictionary, and the issuance of recommendations. Following Phase 1, approval was obtained to use the Priority Information Management System as a prototype to prepare the implementation of a full data lake. If this was successful, the Office of Data Management was planning to continue incorporating other datasets into the Data Lake, one at the time.

68. The evaluation found that, in January 2020, the Data Lake project was not on track to meet its original timelines. Furthermore, the future of the project was uncertain as the Data Lake proposal was withdrawn from the IM/IT Plan 2020-21 deliberations in November 2019. The evaluation team was told that a new approach was being developed to break the Data Lake concept into various components to differentiate between the information technology system requirements, the software tool

requirements (including the procurement of data visualization tools) and the upgrades to various datasets (e.g., the Job-Based Analytical Information System). In recent months, the Data Services and Analysis Directorate has placed a focus on working with the Information Technology Services Directorate to procure data visualization tools (i.e., Microsoft Power BI) to support improvements to data analytics in the short term while planning continues to address the other components of the Data Lake concept.

69. The Data Lake project is the most recent attempt to build a repository that stores data and makes them available to data users. The Public Service Commission has spent considerable time and resources over the past decade to build a centralized, standardized and integrated data repository to support analytical and reporting functions, but not achieved the level of success originally envisioned.

70. In 2008, the Business Intelligence and Enterprise Data Warehouse (BI-EDW) project was proposed as a centralized, structured, standardized and integrated data repository to support analytical and reporting functions across the Public Service Commission. The project had ambitious objectives, including integrating data and making them accessible, increasing efficiency of data analysis, facilitating reporting and, eventually, incorporating metadata and business intelligence tools to address complex issues. Ultimately, the system was to provide automatic processes to produce reports, find new business opportunities, support decision making, improve business outcomes, enhance analytical capabilities, and increase transparency. The BI-EDW was established gradually, starting with the integration of the Public Service Recruitment System data in 2011. In 2014, a portal (Ingenium) was implemented as a means to share a variety of information presented through different tools such as dashboards, data cubes, reports and stored processes.

71. While the BI-EDW was used to analyze the Public Service Recruitment System data, it never served its ultimate purpose as a fully useable data repository. Few employees used the reporting tools due to the non intuitive interface, which was too complex for the average user. Building the solution also required resources to automate processes that were not available at the time. Therefore, only the Public Service Recruitment System data were integrated into the BI EDW. An attempt was made in 2015 to integrate data from the Priority Information Management System and the Job-Based Analytical Information System through the Hiring and Staffing Project. However, the project was cancelled in November 2016 following a recommendation from the Information Technology Services Directorate, with alternative solutions being put in place. A staffing dashboard refreshed monthly was built in Excel, and made Ingenium's content

redundant. Maintaining Ingenium was time consuming for the Data Services and Analysis Directorate, and the number of user visits was very low. An impact analysis by the directorate showed that they would be able to support an increase in requests for data. Ingenium was decommissioned on March 31, 2017.

72. The evaluation found that while Phase 1 of the project was well managed, some lessons learned from the BI-EDW were not fully incorporated into the planning for the overall Data Lake project. This caused a great deal of uncertainty for stakeholders, especially concerning the project management practices of documenting and communicating estimated project costs and time frames and securing the support needed. Some stakeholders were also concerned that the expertise required to develop a data lake was not well understood by the project team, which led to a lack of a common understanding of what the Data Lake was trying to achieve.

73. The evaluation found a lack of buy-in and understanding from some stakeholders toward the Data Lake project. No consensus was reached on the definition of a data lake. Interviewees expressed concerns about the usefulness of a data lake without processes in place to ensure proper data disposition and cleansing. The main concerns highlighted were data errors, dataset inconsistencies and employees' current level of data literacy. An example provided was the extensive work needed to migrate data from the Test Scoring and Results Reporting System to Apollo as part of the Application Rationalization project, which has been in progress for over 7 years. In interviews for the evaluation, some stakeholders suggested, as a first step, standardizing metadata and ensuring that the information on data is complete and that processes are well established to manage the data lifecycle.

74. The BI-EDW and the Data Lake project are 2 examples of the challenges that the Public Service Commission has faced in developing data infrastructure to provide self-service data to users. Given the past work performed in this regard, it is very important to fully understand what data architecture is needed to enable data users to fulfill their mandates.

75. The approved roles and responsibilities for data management at the Public Service Commission do not clearly assign responsibility for identifying data architecture business needs.

76. **Recommendation 3.** It is recommended that the Chief Information Officer, in collaboration with the Chief Data Officer assess what data architecture and data repositories are required for the Public Service Commission. The initiatives to achieve

the requirement should be costed with established timelines, approved and monitored through governance.

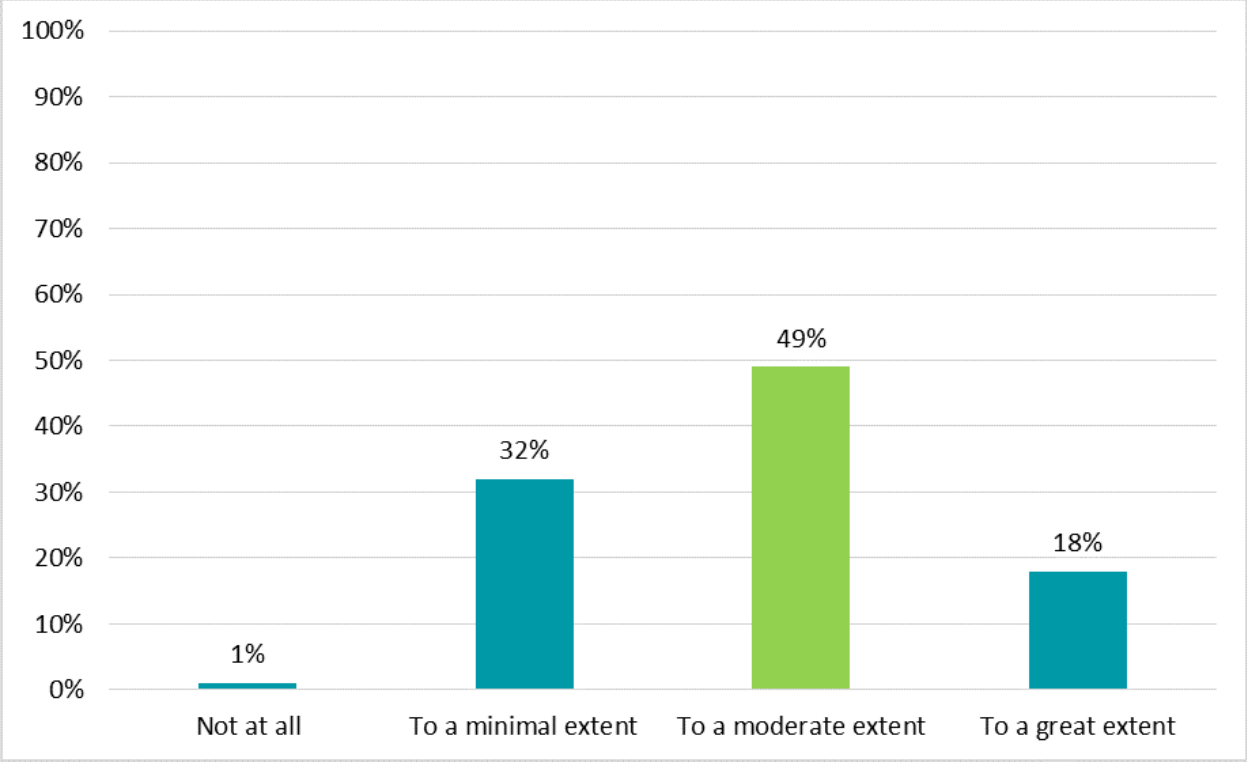
5.4 Data use

Data access and sharing

77. A key objective of the Data Management Strategy is to promote a self-service data culture by improving data access and sharing tools. This objective was to be achieved through various initiatives, such as the implementation of the data lake, the creation of a corporate data assets inventory in 2020, and the integration of datasets into the Open Government Portal.

78. The evaluation team surveyed data users to find out if they were aware of the data available to them. A total of 67%of respondents indicated that they were aware a moderate or great extent (see Exhibit G below). In addition, almost 32%were aware to a minimal extent.

Exhibit G – Data users' awareness of available data (202 respondents)



Text version

Data users' awareness of available data (202 respondents)

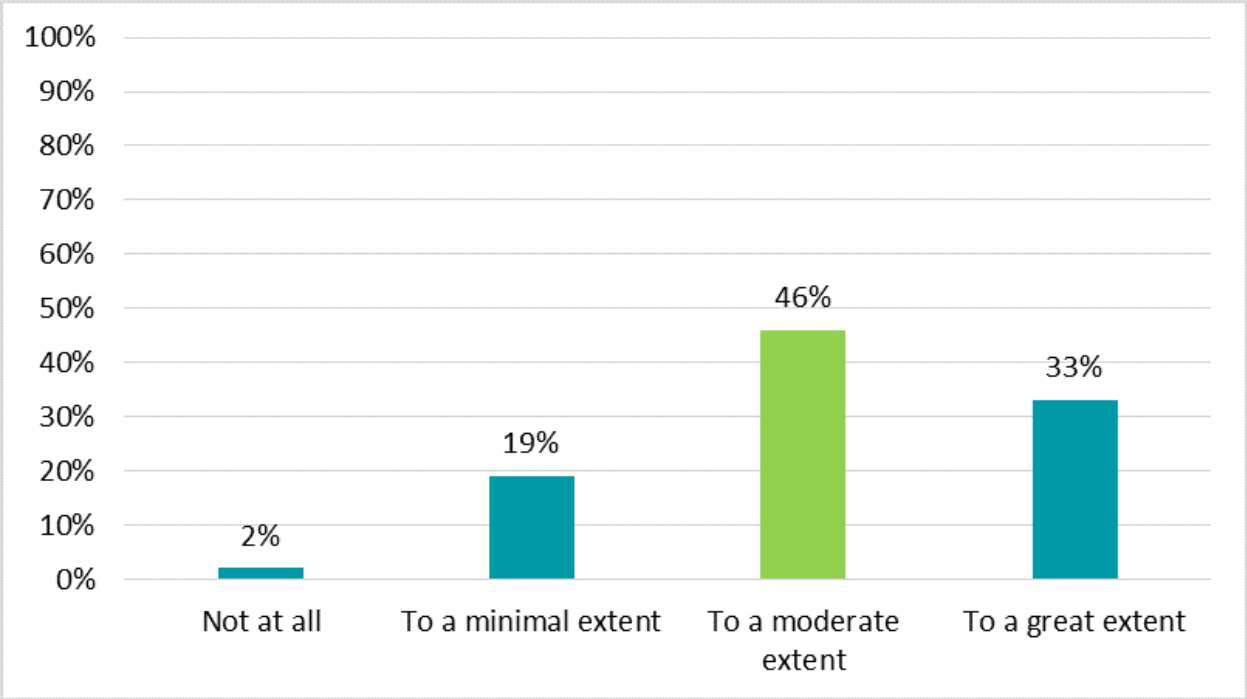
Awareness of data available within the PSC	**Percentage of respondents**
Not at all	1%
To a minimal extent	32%
To a moderate extent	49%
To a great extent	18%

79. Those that reported being aware to a moderate or minimal extent provided comments to explain their perspectives. It was noted that employees have a relatively sound understanding of the data available to them to perform their day to day roles and responsibilities. However, their knowledge is often limited to their sphere of activities, mostly due to a lack of a proper data catalogue of datasets for users to consult. The evaluation found that it is difficult to fully understand the extent of all Public Service Commission data holdings.

80. The evaluation found that data assets are not well known across the department. For example, when the evaluation team asked sector representatives to provide a list of the datasets they manage, some had challenges identifying their sector's datasets and who managed them. It should be noted that in 2016 and 2017, a centralized list of the Public Service Commission data holdings was developed. An inventory was created that included details on each dataset and key information to make these important sources of data easier to use and thereby support data analytics and reporting. However, the inventory document was not maintained. A separate list of data sources was compiled in response to the Directive on Open Government, which requires departments and agencies to produce an inventory of their data holdings. This inventory is maintained and updated as new open data assets are identified. The inventory available on the Open Government portal is updated annually, but is not well known across the organization and does not contain complete information.

81. Data users were also asked whether they had access to the data they needed to perform their jobs and, if so, they were able to access the data in a timely manner. A total of 79% reported having access to the data they need to do their jobs to a great or moderate extent (see Exhibit H below) and 74%were able to access the data in a timely manner to a great or moderate extent (see Exhibit I below).

Exhibit H – Data users' access to the data they need to perform their job (202 respondents)



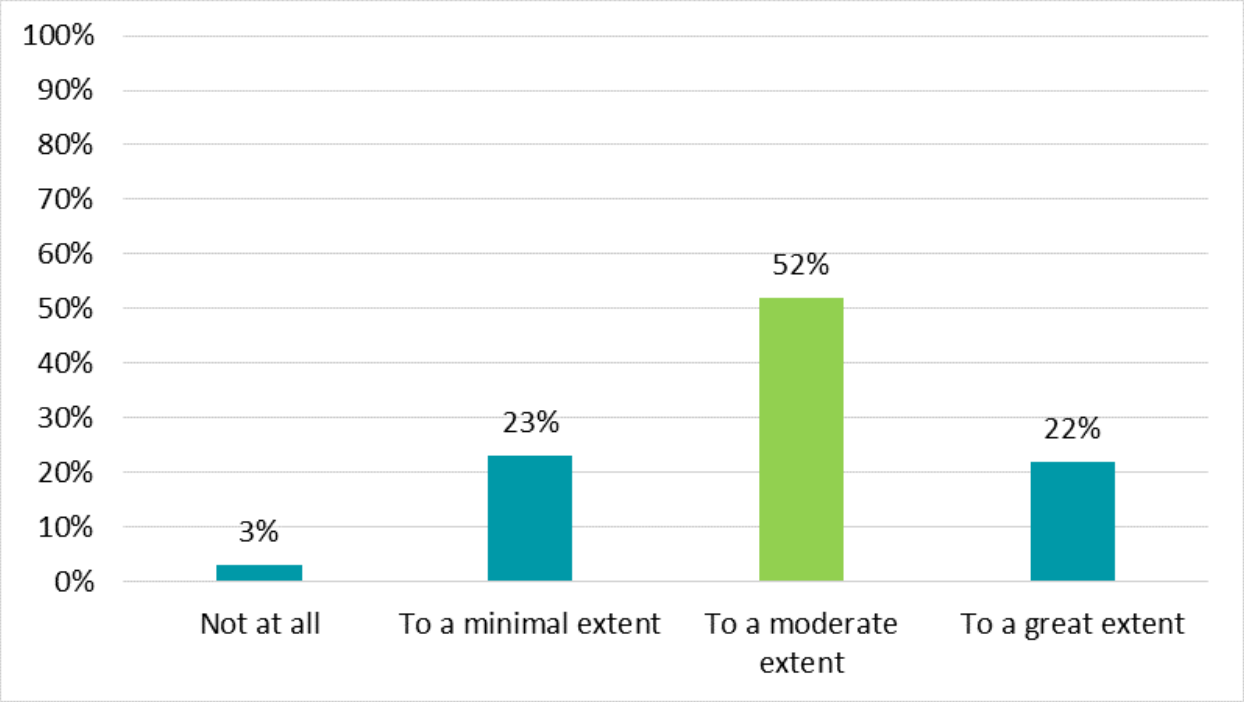
Text version

Data users' access to the data they need to perform their job (202 respondents)

Access to the data you need to do your job	**Percentage of respondents**
Not at all	2%
To a minimal extent	19%
To a moderate extent	46%

Access to the data you need to do your job	Percentage of respondents
To a great extent	33%

Exhibit I – Data users' access to the data they need in a timely manner (202 respondents)



Text version

Data users' access to the data they need in a timely manner (202 respondents)

Access to the data they need in a timely manner	Percentage of respondents
Not at all	3%
To a minimal extent	23%
To a moderate extent	52%

Access to the data they need in a timely manner	Percentage of respondents
To a great extent	22%

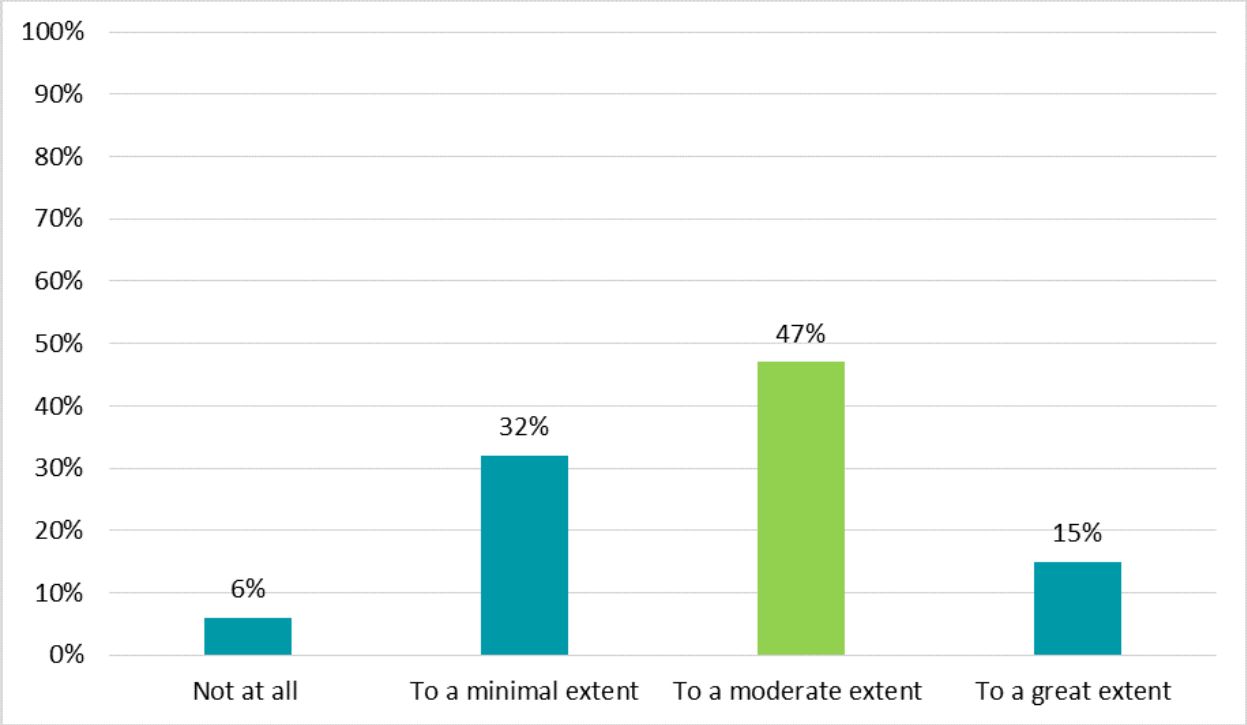
82. The 21% and 26% of respondents who answered "To a minimal extent" or "Not at all" to each question raised a number of concerns. The most common issue was the dependency on the Data Services and Analysis Directorate to obtain data from systems that they manage (e.g., the Job-Based Analytical Information System). The provision of quality data was not the issue; rather, in some instances, it was delays in fulfilling requests. This problem is a function of the priorities of the Data Services and Analysis Directorate at the time, priorities that are not well understood by data requesters. For example, the directorate prioritizes data requests related to the annual report and those intended to support the President in discussions with the Deputy Heads regarding the Appointment Delegation and Accountability Instrument. The capacity to support various stakeholders therefore depends on workloads.

83. In recent years, the Open Government Secretariat has worked to support better access to data for internal and external stakeholders. The Public Service Commission has shared approximately 97 datasets on the Open Government Portal. Even if the information published is limited, the is a significant improvement that has the potential to reduce the number of ad hoc requests sent to the Data Services and Analysis Directorate. In January 2020, the Office of Data Management developed and released a YouTube video to inform stakeholders how to use data available on the Open Government portal, an important initiative that may help data users access and better understand the data that they require. In addition to this video, the Public Service Commission has also created an interactive visualization tool to accompany its data on the Staffing and Non-Partisanship Survey. Finally, it should be noted that each dataset on the portal has accompanying supporting documentation.

Tools to use and manage data

84. Data users were asked whether they had the appropriate tools to manage data. Overall, 62%indicated that they had access to data analysis tools to a great or moderate extent (see Exhibit J below). Therefore, 38% of the data users had access to data analysis tools to a minimal extent or not at all.

Exhibit J – Data users' access to data analysis tools (202 respondents)



Text version

Data users' access to data analytics tools (202 respondents)

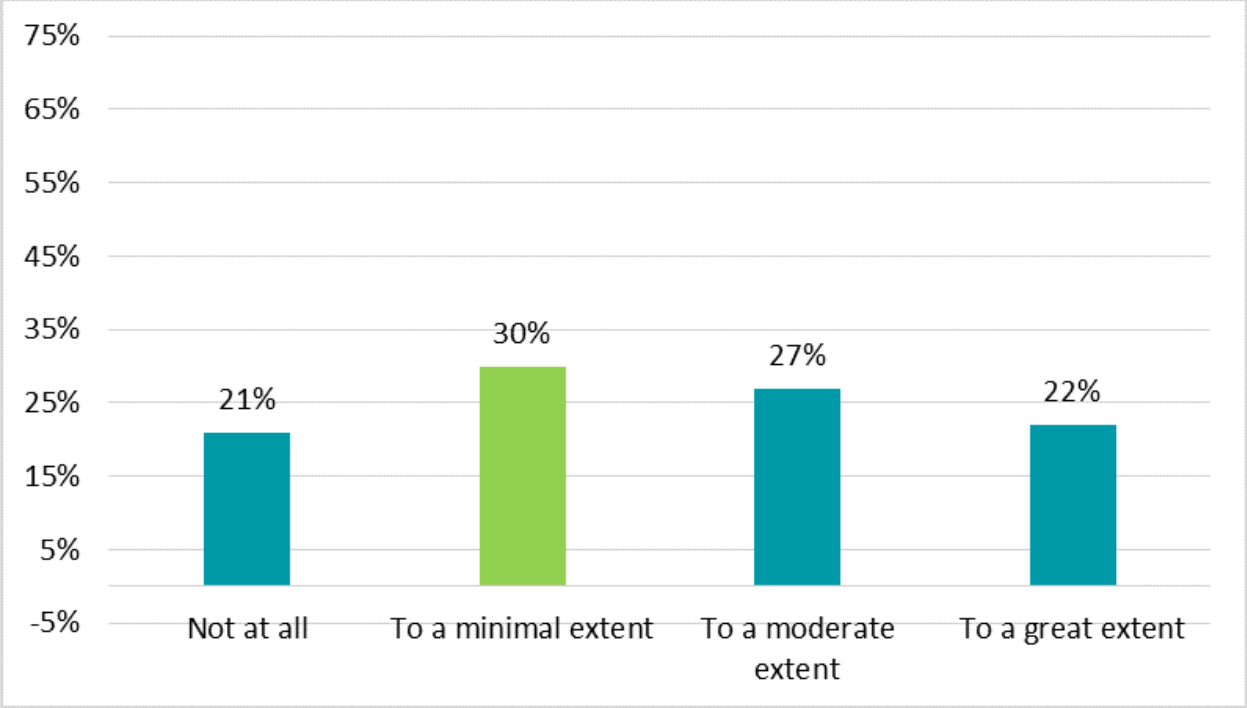
Access to data analytics tools	**Percentage of respondents**
Not at all	6%
To a minimal extent	32%
To a moderate extent	47%
To a great extent	15%

85. These results are similar to those from the Data Services and Analysis Directorate baseline survey conducted in November 2018, to which 28.8% of data users answered that they had the tools to use data to a large or very large extent, 35.6% to a moderate extent, 30.5% to a small extent and 5.1% not at all.

86. The approximately 38% of respondents that answered "To a minimal extent" or "Not at all" provided comments to justify their concerns. Many respondents commented that Excel was sufficient to meet their needs, while others wanted more sophisticated tools and support to enhance their capability to analyze data. Some respondents identified a need for tools such as Microsoft Power BI, R, Python and a new version of SPSS. Some users commented on the lack of support for more advanced data analysis tools. The comments demonstrate that not all data user groups have the same tools: For example, some data users requested Microsoft Power BI, while others reported using it and that it effectively met their needs.

87. The evaluation survey also asked data users the extent to which they used data analysis tools, data visualization tools and data science tools. The results are outlined in Exhibits K, L, and M. Users were more likely to use data analytics and visualization tools than data science tools, which are primarily used by Data Services and Analysis Directorate data specialists.

Exhibit K – Use of data analytics tools (202 respondents)

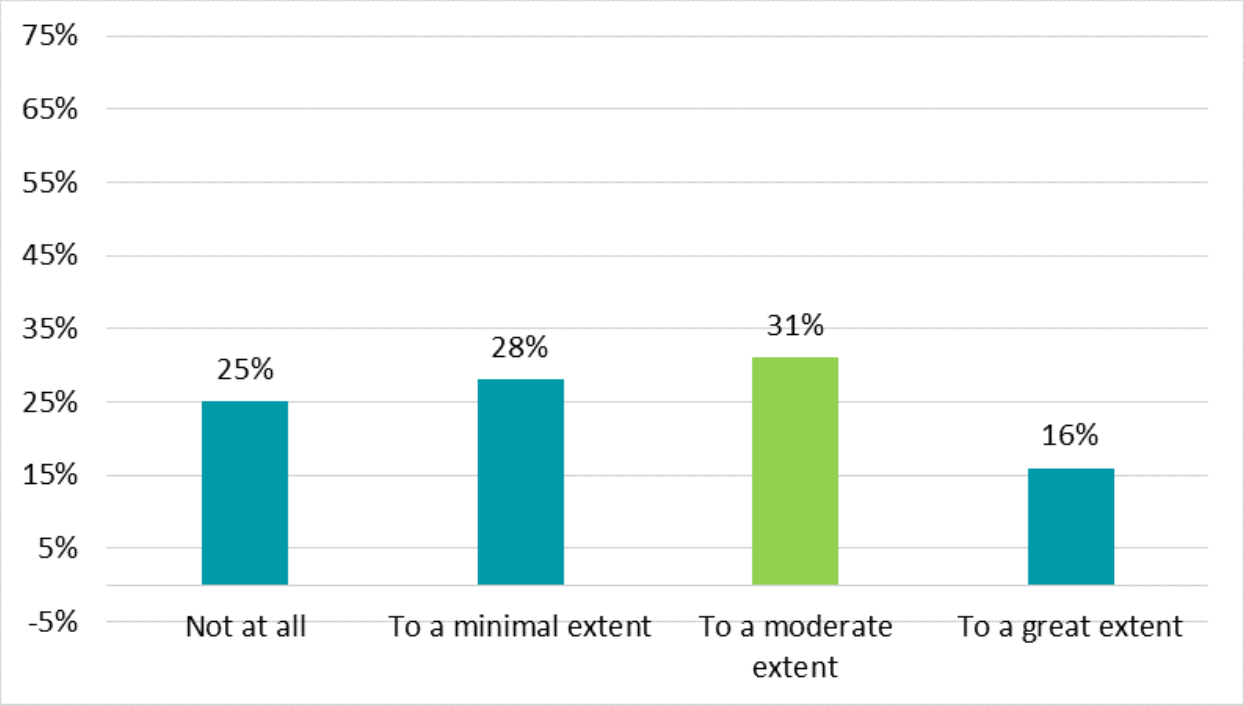


Text version

Use of data analytics tools (202 respondents)

Extent of use of data analytics tools	**Percentage of respondents**
Not at all	21%
To a minimal extent	30%
To a moderate extent	27%
To a great extent	22%

Exhibit L – Use of data visualization tools (202 respondents)

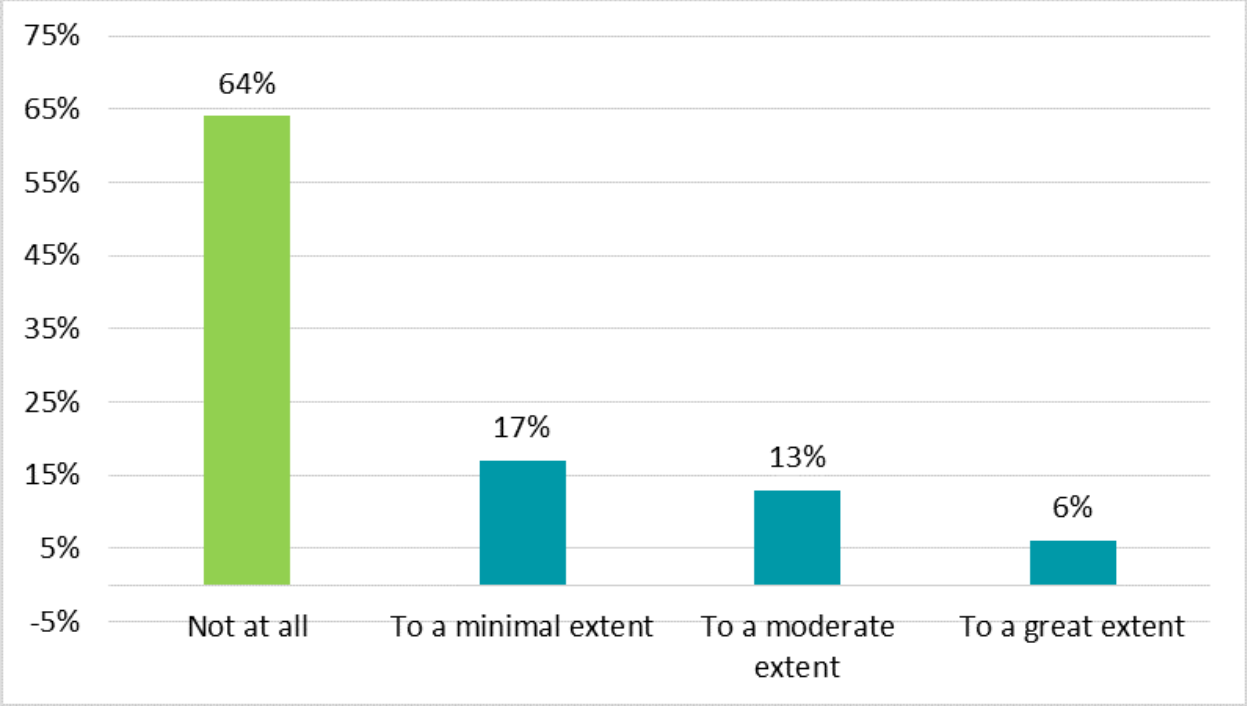


Text version

Use of data visualization tools (202 respondents)

Extent of use of data visualization tools	**Percentage of respondents**
Not at all	25%
To a minimal extent	28%
To a moderate extent	31%
To a great extent	16%

Exhibit M – Use of data science tools (202 respondents)



Text version

Use of data science tools (202 respondents)

Extent of use of data science tools	**Percentage of respondents**
Not at all	64%
To a minimal extent	17%
To a moderate extent	13%
To a great extent	6%

88. The evaluation found that the Data Services and Analysis Directorate and the Information Technology Services Directorate were well versed with data tools. However, such expertise in other sectors varies and is often held by only a handful of people. The evaluation found that due to relatively high employee turnover in the Information Technology Services Directorate in the past, it has been difficult for the directorate to support the implementation of a number of tools requested by data users. Given the limited resources available to support data tools, it may be useful for the organization to consider how expertise on these tools could be shared throughout the organization.

89. A good practice identified by this evaluation is the Office of the Comptroller General's use of Microsoft Power BI to enhance data analysis and reporting to Chief Audit Executives on their performance and the overall performance of the internal audit function across the federal government. The office had recognized the need for a tool to more efficiently analyze and report on data regarding the state of the internal audit function. They invested resources and made conscious hiring decisions to bring in data experts to lead these efforts, establishing a dedicated team. The results were presented to Chief Audit Executives in fall 2019 and to Internal Audit Committee members from across the federal government at the Departmental Audit Committee Symposium in November 2019.

90. The Office of the Comptroller General's model is applicable to the Public Service Commission. Given that the Public Service Commission maintains data used to report on the health of staffing across the federal public service and provides these data to

Deputy Heads through the Public Service Commission Staffing Dashboard and during Appointment Delegation and Accountability Instrument discussions, there are many lessons that could be learned from the office's experience.

91. **Recommendation 4.** It is recommended that the Chief Information Officer, in collaboration with the Chief Data Officer, develop a strategy to identify and obtain the required data tools for users that support improved methods to analyze and report on key datasets. The strategy should be costed and approved through governance.

Change management, data literacy and training plans to support employees

92. The implementation of the Data Management Strategy requires a culture change among employees and throughout the organization. This is particularly important in light of the planned shift to self-serve data that data users can use to perform analysis. While it is not expected that every employee will become a data scientist, the Office of Data Management is working to develop a strategy to increase employee knowledge and expertise to help them become comfortable using data. To do so, they have developed a change management plan that, once approved, will support the planned shift in how data are accessed and used.

93. In January 2020, the Office of Data Management was in the process of developing the change management plan for approval. The work is progressing and will be informed by the results of this evaluation and other sources, such as the recently published Privy Council Office assessment of organizational data strategies (January 2020). The Office of Data Management is aware of and considering the years of informal data management practices and processes that have fostered habits that will take time to change.

94. The Office of Data Management is also developing an approach to enhance data literacy across the organization. They have conducted surveys and obtained further information on data management practices and are using this information to develop an appropriate data literacy plan. The baseline survey administered in November 2018 identified a need to raise awareness among employees of the Chief Data Officer and to explain the main roles and responsibilities of the position. Ultimately, the implementation of the data literacy plan will enhance dataset managers' and data users' ability to understand and interpret data and optimize their use while supporting an organizational culture that values data as an asset.

95. Furthermore, training modules are being developed to help data users increase their knowledge and understanding of data use, data governance, sources of available data (e.g., staffing data) and data quality. The plan is to leverage the work being done by various centres of expertise, such as the Canada School of Public Service and Statistics Canada. In January 2020 progress was being made, with a number of modules still under development. Once approved, the data literacy and training plan is expected to show results over a number of years as the shift in data culture is felt across the Public Service Commission.

96. In summary, the Office of Data Management has demonstrated a sound understanding of the environment in which it is developing the change management strategy, data literacy plan and training plan to support the implementation of the Data Management Strategy. The implementation of these important elements of the Data Strategy will help employees to better understand how they can use data in their work and to generate higher quality information for decision making. This work aligns with the federal Data Strategy Roadmap, which recommends that organizations "assess the current state of data literacy as well as skills and competencies

6.0 Conclusion and recommendations

97. In conclusion, 2019–20 marked the first full year of implementation of the Data Management Strategy. The Public Service Commission's vision aligns with the overall federal government vision of improving data management to support evidence based decision making and reporting. The Public Service Commission is on track with respect to implementing the recommendations made in the federal Data Strategy Roadmap and has made improving data management a key priority that supports the achievement of the organization's mandate.

98. The Public Service Commission has taken important steps to improve data management. A Chief Data Officer was appointed in April 2018 and is supported by the Office of Data Management and the Data Services and Analysis Directorate reporting team. The Public Service Commission Data Management Strategy was approved in June 2018 and sets out the vision and actions planned to achieve it. Data governance was established to provide support and oversight for data management activities, including joint IM/IT Committee and Integration Committee meetings that bring data, business and information technology staff together to advance various initiatives in support of the Data Management Strategy. Work has also started to develop change management

and data literacy plans to build capacity to drive the necessary culture change and build the awareness required to fully implement the Data Management Strategy.

99. Furthermore, the work performed by the Open Government Secretariat has improved access to data for internal and external stakeholders. The Public Service Commission has shared approximately 97 datasets on the Open Government Portal. This is a significant improvement that has the potential to increase data users' access to a greater amount of data to enable better data analysis to support improved decision making.

100. While progress has been made, there are opportunities to improve data governance to promote collaboration between stakeholders as well as project management practices, data quality assurance, data disposition and cleansing, and the documentation of dataset management practices. It is also important for the Public Service Commission to determine the organization's needs in modern data infrastructure in order to help realize the vision of having the right data, the right way, in the right hands, at the right time.

101. The following recommendations are designed to improve the implementation of the Data Management Strategy at the Public Service Commission.

102. **Recommendation 1.** It is recommended that the Chief Data Officer, in collaboration with the Chief Information Officer and Data Stewards, review data governance to ensure it supports internal collaboration in the implementation of the Data Management Strategy resulting in an organization-focused perspective on data management.

103. **Recommendation 2.** It is recommended that the Chief Data Officer, in collaboration with the Chief Information Officer and Data Stewards review, update, and document data quality policies and procedures to enable proper data quality assurance, data disposition, and data privacy practices across the Public Service Commission.

104. **Recommendation 3.** It is recommended that the Chief Information Officer, in collaboration with the Chief Data Officer assess what data architecture and data repositories are required for the Public Service Commission. The initiatives to achieve the requirement should be costed with established timelines, approved and monitored through governance.

105. **Recommendation 4.** It is recommended that the Chief Information Officer, in collaboration with the Chief Data Officer, develop a strategy to identify and obtain the

required data tools for users that support improved methods to analyze and report on key datasets. The strategy should be costed and approved through governance.

7.0 Management response and action plan

Recommendation	**Management response and planned actions**	**Office of primary interest (OPI)**	**Proposed completion date**
1. It is recommended that the Chief Data Officer (CDO), in collaboration with the Chief Information Officer (CIO) and Data Stewards, review data governance to ensure it supports internal collaboration in the implementation of the Data Management Strategy resulting in an organization-focused perspective on data management.	Management agrees with this recommendation. The CDO, CIO and Data Stewards will leverage the results of the Public Service Commission (PSC) governance review (in progress at the time of writing). This will allow for the holistic integration of data into PSC's governance. This will best position the organization-focused implementation of the Data Management Strategy. **Planned actions** - The planned actions will be on-going and integrated into the renewed PSC governance structure once approved by December 30, 2020. - Proactively present Data Management Strategy	OPI : CDO OPI: CDO in collaboration with the CIO, the Office of Data Management	December 30, 2020 December 30, 2020

Recommendation	Management response and planned actions	Office of primary interest (OPI)	Proposed completion date
	initiatives, at key points, from conception to completion, through governance to facilitate horizontal collaboration and understanding of progress and direction. - Develop a framework for planning and managing change resulting from the Data Management Strategy. - Create a PSC Data Change Agents Network to help promote and communicate the Strategy.	(ODM) and Data Services and Analysis Directorate (DSAD) OPI: CDO in collaboration with ODM OPI: CDO in collaboration with ODM.	September 30, 2020 January 31, 2021
2. It is recommended that the Chief Data Officer, in collaboration with the Chief Information Officer and Data Stewards review, update, and document data quality policies and procedures to enable proper data quality assurance, data disposition, and data	Management agrees with this recommendation. The governance resulting from item #1 will be the formal governance used to ensure policy instruments are developed, maintained and implemented. These include: **Planned actions**	Data quality	September 30, 2020

Recommendation	Management response and planned actions	Office of primary interest (OPI)	Proposed completion date
privacy practices across the PSC.	Data quality policy instruments to enable proper data quality assurance. - Approve the data literacy plan. - Partner with Statistics Canada and the Canada School of Public Service to develop data quality material. - Develop a policy on data quality. - Implement competencies for data analysts / scientists. - Acquire right talent and skills to ensure proper data quality practices. Data disposition policy instruments - Review and update data disposition policy instruments in line with the findings in this report. Data privacy practices (led by the executive lead of privacy)	OPI: CDO in collaboration with the ODM OPI: CDO in collaboration with DSAD OPI: CIO OPI: ATIP Office	March 31, 2021 June 30, 2021 March 31, 2021 December 31, 2020

Recommendation	Management response and planned actions	Office of primary interest (OPI)	Proposed completion date
	- The ATIP office will update and communicate the PSC Privacy Management Framework (including its related policy instruments). - The ATIP office, in collaboration with DSAD, will complete a Privacy Impact Assessment (PIA) on the PSC's Data and Analytics Services.	OPI: ATIP Office in collaboration with DSAD	March 31, 2021 March 31, 2021 September 30, 2020
3. It is recommended that the Chief Information Officer, in collaboration with the Chief Data Officer assess what data architecture and data repositories are required for the PSC. The initiatives to achieve the requirement should be costed with established timelines, approved and monitored through governance.	Management agrees with this recommendation. - The CIO will lead, in collaboration with the CDO and Data Stewards, the assessment of the data architecture and associated repositories that will be required for the PSC in light of the findings in this report. The resulting proposed initiatives will be costed, with established timelines, and submitted through governance (see item #1). Approved initiatives will be	OPI: CIO in collaboration with the CDO and Data Stewarts.	March 31, 2021

Recommendation	Management response and planned actions	Office of primary interest (OPI)	Proposed completion date
	monitored through governance (see item #1) accordingly.		
4. It is recommended that the Chief Information Officer, in collaboration with the Chief Data Officer, develop a strategy to identify and obtain the required data tools for users that support improved methods to analyze and report on key datasets. The strategy should be costed and approved through governance.	Management agrees with this recommendation. - The CIO will lead, in collaboration with the CDO, a review of the data related business objectives of each sector area and develop a strategy to identify and obtain organization-wide tools, training and processes to support the various requirements amongst sectors in light of the findings within this report. The resulting strategy will be costed and submitted through governance (see item #1). As well, approved initiatives will be monitored through governance (see item #1) accordingly.	OPI: CIO in collaboration with the CDO	March 31, 2021

Annex 1: Data Management Strategy logic model

Input	Activities	Output	Immediate outcomes	Intermediate outcomes	Ultimate outcomes
Tools/ Infrastructure solutions **Funding** **Personnel**	Establish a governance for data management	• A data-focused governance • The Office of Data Management	Internal stakeholders understand their roles and responsibilities	The PSC produces reliable data	Internal and external stakeholders better leverage PSC's data The PSC produces high quality data supporting public service hiring and non-partisanship
	Integrate data	• A data storage and integration solution (data lake) • Data architecture	Increased integration of data		

Input	Activities	Output	Immediate outcomes	Intermediate outcomes	Ultimate outcomes
	Develop policies for data management practices	• Guidelines and procedures for quality control, security, privacy, and ethical use of data • Data retention and disposition standards	The Public Service Commission (¨SC) employees are aware of the data management policies		
	Share data of business value	• A plan to share data of business value • Data of business value	The Office of Data Management understands what data of business value are needed by the organization	Data users have access to the data they need	

Input	Activities	Output	Immediate outcomes	Intermediate outcomes	Ultimate outcomes
	Develop and communicate the new vision for managing and using data	• Change management plan	Increased awareness and buy-in of PSC data strategy by employees	The PSC has skilled people to use, analyze and interpret data	
	Build capacity	• Data literacy plan • Training • Tools	Increased level of data literacy amongst PSC's employees		

Key logic model assumptions and risks:

Governance. The Data Management Strategy is championed by the Chief Data Officer and supported by the Office of Data Management.

Assumption. Data are more likely to support evidence-based decisions with the proper Data Management Strategy in place.

Risks. The key risks identified to date are:

- Roles and responsibilities of the Data Management Strategy may not be well understood across the Public Service Commission
- The Public Service Commission may not have a proper data storage and integration solution;
- The Public Service Commission may not have the proper policies to support quality control, security, privacy, and ethical use of data;
- The objectives of the Data Management Strategy may not be well communicated;

- Public Service Commission's employees may not have the proper level of data literacy to leverage data;
- A lack of buy-in of Public Service Commission employees may result in lack of collaboration, low proficiency and under-utilisation of data.

Annex 2: Evaluation matrix

Evaluation questions	Indicators	Methods
RELEVANCE		
Alignment		
1.1 To what extent is the Data Management Strategy supporting the PSC (Public Service Commission) strategic direction?	**1.1.1** The extent to which the Strategy includes best practices to ensure alignment with PSC's data needs	• Document review • Interviews • Literature review • Questionnaires • Maturity model
1.2 To what extent is the Data Management Strategy supporting the PSC (Public Service Commission) strategic direction?	**1.2.1** To what extent is the Data Management Strategy aligned with the requirements of the government? **1.2.2** The degree of alignment of the Strategy with Treasury Board Secretariat Policy and Directive on Service and	• Document review • Interviews • Literature review

	Digital	
EFFECTIVENESS AND EFFICIENCY		
Outputs		
2.1 To what extent the data storage and integration solutions are expected to be implemented in a timely fashion?	**2.1.1** The expected time required to implement a data repository ready to be used by stakeholder	• Document review • Literature review • Interviews • Maturity model
2.2 To what extent is the PSC developing an effective approach to share data of business value?	**2.2.1** The extent to which the PSC has identified data of business value	• Document review • Literature • review • Interviews • Maturity model
2.3 To what extent is the Office of Data Management implementing an efficient change management plan to promote a new data management culture?	**2.3.1** The extent to which the Office of Data Management has defined objectives for its change management plan **2.3.2** The extent to which the change management plan supports the	• Document review • Literature review • Interviews • Maturity model

	realization of the Strategy	
2.4 To what extent is the Office of Data Management implementing an appropriate data literacy plan?	**2.4.1** The extent to which the Office of Data Management knows the level of data literacy of the PSC's employees **2.4.2** The extent to which the Office of Data Management has defined the level of data literacy required for PSC's employees **2.4.3** The extent to which stakeholders are confident they have the ability to implement the practices and processes being developed for data management	• Document review • Literature review • Interviews • Maturity model
2.5 To what extent is the PSC having the appropriate tools to manage and interpret data?	**2.5.1** The extent to which the PSC is having the appropriate tools to manage and interpret data	• Document review • Literature review • Interviews • Questionnaires

2.6 Are there better alternatives, ways or lessons learned (for example, design changes, cost) for achieving results?	**2.6.1** Evidence of alternative models for data management in other federal departments, national or international jurisdictions	• Document review • Literature review • Interviews • Maturity model
Immediate outcomes		
3.1 To what extent is the PSC having an effective governance for the management of data	**3.1.1** The extent to which stakeholders are aware of their roles and responsibilities regarding data management **3.1.2** The extent to which stakeholders consider that their roles and responsibilities regarding data management are clear **3.1.3** The extent to which stakeholders consider they have the capacity to fulfill their roles and responsibilities	• Document review • Interviews • Literature review • Questionnaires • Maturity model

	regarding data management	
Intermediate outcomes		
4.1 To what extent stakeholders have access to the data they need?	**4.1.1** The extent to which the stakeholders have access to the data they need **4.1.2** The extent to which stakeholders are confident the Strategy will help them obtaining the data they need in a timely fashion	• Document review • Interviews • Literature review • Questionnaires • Review of administrative and performance data
4.2 To what extent is the PSC managing data in a manner that produces accurate data and maintains confidence	**4.2.1** The extent the PSC is moving towards the creation of policies for data quality **4.2.2** The extent stakeholders are confident data management ensures data quality, privacy and ethical usage of data **4.2.3**	• Document review • Interviews • Literature review • Questionnaires • Review of administrative and performance data • Maturity model

	The extent to which stakeholders have confidence in PSC's data	