

Government of Canada Data Competency Framework Version 2 Draft for Discussion

December 2019

GC Data Competency Framework
V.2 Draft December 2019

Table of Contents

GC Competency Framework	1
Data Concepts and Culture	
Introduction to Data	2
Data Culture.....	3
Data Ethics	4
Evidence Based Decision Making.....	5
Data Collection and Management	
Data Stewardship (Collect, Store, Protect)	6
Data Governance	7
Metadata Creation and Use.....	8
Cleansing Data	9
Tools and processing	10
Database and Systems Architecture	
Database Principles.....	11
Stability and Storage	12
Create and Test New Systems	13
Deploy and Maintain Systems	14
Extract, Transform, Load (ETL).....	15
Analytics and Evaluation	
Tools.....	16
Identifying Questions Using Data	17
Data Analysis.....	18
Advanced Analytics.....	19
Visualization.....	20
Storytelling.....	21
Evaluation Decisions Based on Data	22
Annex A Generic Proficiency Level Descriptors.....	23
Annex B ESDC Data Personas (Draft)	24
Annex C Business and Policy Analyst Journey Map.....	25
Annex D Business and Policy Analyst Proficiency Mapping.....	26
Annex E Mapping Example – Data Ethics	27

Government of Canada Data Competency Framework Context

A data competency framework and competency mapping was developed to establish a common understanding core, leadership and niche data competencies across the Government of Canada. It is intended to create a baseline for individual and organizational competency assessments and recruitment initiatives, and to guide the development of learning pathways and resources that will provide foundational data literacy to all public service employees, as well as continuous development pathways for data-driven roles.

In 2018, the Government of Canada published 'A Data Strategy Roadmap for the Federal Public Service' as a guide to help departments and agencies think about how they could unlock the value of their data and move towards a data driven business model.

Developments in technology have created the conditions for exponential growth in the fields of data science, data management and artificial intelligence. For federal departments and agencies to deliver modern, inclusive, digital services to Canadians, they must increase their ability to read, understand and use data before it can be used as a strategic asset.

"The volume of data that governments, business and Canadians produce is growing exponentially, animated by digital technologies. Organizations are changing their business models, building new expertise and devising new ways of managing and unlocking the value of their data. Governments need to evolve rapidly to keep up." – A Data Strategy Roadmap for the Federal Public Service

Widespread data literacy is critical to enable a data-driven culture that is open, shares by default and realizes business value from data holdings. A data literate workforce is one that can use data to improve the quality of decisions and services; knows when and how to question validity of data; read, comprehend and use statistical information ethically; explore data to generate new ideas, methods and understanding; and represent information visually and communicate it to an audience.

The Data Competency Framework was developed in response to Data Strategy Roadmap Recommendation #9 - *Assess the current state of data literacy as well as skills and competencies required* and Recommendation #10 - *Renew HR strategies with targeted training and development, including engagement with colleges, universities, and unions.*

Government of Canada Data Competency Framework Introduction

This document is the second version of a Government of Canada (GC) data competency framework and was created through a working group of departments. It builds on:

- a literature review of data competencies framework conducted by Statistics Canada
- the Strategies and Best Practices for Data Literacy Education report released by Dalhousie University
- a model adapting the above for the GC by Shared Services Canada
- data community calls convened by Environment and Climate Change Canada, now convened by the Digital Academy, Canada School of Public Service (CSPS)

The competencies and knowledge/tasks are derived primarily from the Strategies and Best Practices for Data Literacy Education (Ridsdale, 2015 - Appendix 1), along with input from multiple federal departments. Competency definitions were derived from a variety of sources including a Data Science Competency Framework (Data to Decisions, CRC, 2017), Termium Plus, internet searches and the Digital Academy, CSPS. Version 2.0 will be shared more broadly in January 2020 for further feedback, and changes are anticipated.

In addition to the competency matrix and supporting pages, appendices are included with proficiency levels, examples of data personas and persona competency mapping. These are intended as supports as departments and agencies consider the data competencies required across functions and communities.

GC Data Competency Framework v.2

DRAFT

Data Concepts and Culture

Introduction to Data

Data Culture

Data Ethics

Evidence Based Decision Making

Data Collection and Management

Data Stewardship
(Collect, Store, Protect)

Data Governance

Metadata Creation and Use

Cleansing Data

Tools and processes

Database and Systems Architecture

Database Principles

Stability and Storage

Create and Test systems

Deploy and Maintain Systems

Extract, Transform, Load

Analytics and Evaluation

Tools

Identifying Questions Using Data

Data Analysis

Advanced Analytics

Visualization

Storytelling

Evaluating Decisions Based on Data

Data Concepts and Culture

Introduction to Data

Competency Description - Introduction to Data

Foundational learning on data topics for all public servants

Knowledge and Tasks

- ▶ Knowledge and understanding of data (e.g. what is data? Types of data, open data, big data, data trends, business value, opportunities with data, risks of failing to leverage data)
- ▶ Knowledge of Government of Canada Data Strategy Roadmap
- ▶ Data concepts (categories, formats, transformation, data life cycle)
- ▶ Organizational data roles, policies, standards and processes
- ▶ Knowledge of basic statistical concepts
- ▶ Importance of data management (Reference Data Management, Master Data Management)

Data Concepts and Culture

Data Culture

Competency Description - Data Culture

Attitudes and behaviors that support a data-driven organization.

Knowledge and Tasks

- ▶ Recognizes importance of data (e.g. data quality, access to data and data literacy)
- ▶ Supports critical use of data for learning, research and decision-making
- ▶ Champion for data culture
- ▶ Data privacy/security/individual's role to protect information

Data Concepts and Culture

Data Ethics

Competency Description - Data Ethics

Ethical issues relative to the sharing and usage of data

Knowledge and Tasks

- ▶ Recognizes the legal and ethical issues associated with data. Knowledge of principles of data ethics (unfair discrimination, reinforcing human biases, lack of transparency)
- ▶ Applies and works in an ethical manner
- ▶ Shares data legally and ethically

Data Concepts and Culture

Evidence Based Decision Making

Competency Description - Evidence Based Decision Making

Evidence based decision making 'helps people make well informed decisions about policies, programmes and projects by putting the best available evidence from research at the heart of policy development and implementation.' (Davies, 2004: 3).

Knowledge and Tasks

- ▶ Aware of high level issues and challenges associated with data
- ▶ Thinks critically when working with data
- ▶ Formulates appropriate business questions
- ▶ Locates applicable datasets
- ▶ Decides on measurement priorities
- ▶ Prioritizes information garnered from data
- ▶ Converts data into actionable information
- ▶ Weighs the merit and impacts of possible solutions/decisions
- ▶ Implements decisions/solutions

Data Collection and Management

Data Stewardship (Collect, Store, Protect)

Competency Description - Data Stewardship (Collect, Store, Protect)

Data Stewardship concerns the process(es) of effectively managing access, quality, security and accessibility of an organization's data assets.

Knowledge and Tasks

- ▶ Performs data exploration
- ▶ Identifies useful data
- ▶ Collects data
- ▶ Organizes data
- ▶ Critically assesses source data and datasets for quality and trustworthiness
- ▶ Curates data
- ▶ Assesses requirements for preservation
- ▶ Preserves data

Data Collection and Management

Data Governance

Competency Description - Data Governance

Data governance (DG) is the overall management of the availability, usability, integrity and security of data used in an enterprise.

Knowledge and Tasks

- ▶ Assesses data organization requirements
- ▶ Assesses data curation requirements (e.g. retention schedule, storage, accessibility, sharing requirements, etc)
- ▶ Assesses data security requirements (e.g. restricted access, protected drives, etc.)
- ▶ Assesses methods and platforms for sharing data
- ▶ Assess methods and tools for data preservation
- ▶ Develops/has knowledge of organizational policies, processes and standards for data management/stewardship
- ▶ Depersonalization of data for sharing
- ▶ Risk management

Data Collection and Management

Metadata Creation and Use

Competency Description - Metadata Creation and Use

The creation and assignment of searchable key attributes for an organization's data assets, in order to effectively organize and access data assets.

Knowledge and Tasks

- ▶ Creates metadata descriptors
- ▶ Assigns appropriate metadata descriptors to original data sets

Data Collection and Management

Cleansing Data

Competency Description - Cleansing Data

Data cleansing is the process of altering data in a given dataset to make sure that it is accurate and correct. This includes identifying and correcting inaccurate, incomplete, or irrelevant data and replacing, modifying or deleting them.

Knowledge and Tasks

- ▶ Assess methods to cleanse data
- ▶ Identifies outliers and anomalies
- ▶ Knowledge of different data types and conversion methods
- ▶ Cleans data

Data Collection and Management Tools and Processes

Competency Description - Tools and Processes

Software, tools and processes used to collect, organize, store and manage data.

Knowledge and Tasks

- ➊ Data organization methods and tools
- ➋ Data collection methods and tools
- ➌ Implements and adheres to data policies, processes, procedures and standards

Database and Systems Architecture

Database Principles

Competency Description - Database Principles

Principles of database design in order to maximize usability, integrity, security, scalability and accessibility of data.

Knowledge and Tasks

- ▶ Knowledge of relational database principles
- ▶ System integration and inter-operability
- ▶ Experience in relational data modelling

Database and Systems Architecture

Stability and Storage

Competency Description - Stability and Storage

Knowledge of data storage and distribution systems and relevant software tools.

Knowledge and Tasks

- ▶ Knowledge of enterprise data warehouse structures
- ▶ Ensures storage size is in line with organizational needs and industry best practices
- ▶ Knowledge of cloud storage options and impact on data

Database and Systems Architecture

Create and Test Systems

Competency Description - Create and Test Systems

Creating software components to collect, manage and analyse data. Testing and configuring database and data processing systems.

Knowledge and Tasks

- ▶ Assess system requirements (e.g. storage, retention, schedule)
- ▶ Assess data security requirements
- ▶ Develop, implement, test and troubleshoot system workflows

Database and Systems Architecture

Deploy and Maintain

Competency Description - Deploy and Maintain Systems

Installing, testing and configuring scalable database and data processing systems

Knowledge and Tasks

- ▶ Assesses new tools for use internally
- ▶ Knowledge of existing infrastructure, architecture and data technologies to deploy systems, techniques and technologies in a way that is sustainable and can be maintained over time
- ▶ Ability to identify problem points in system deployment and develop solutions
- ▶ Knowledge of best practices to minimize system down time

Database and Systems Architecture

Extract, Transform, Load (ETL)

Competency Description - Extract, Transform, Load (ETL)

The process by which data is extracted from data sources that are not optimized for analytics, moved to a central host, and optimized for analytics.

Knowledge and Tasks

- ▶ Knowledge of ETL tools and practices
- ▶ Tools and techniques to script ETL functions

Analytics and Evaluation Tools

Competency Description - Analytics and Evaluation Tools

Commercial and open-sourced data analysis and exploration software tools and programming languages, including their advantages and disadvantages, constraints, applicability and best practices.

Knowledge and Tasks

- ▶ Knowledge of data analysis tools and techniques
- ▶ Awareness of industry best practices
- ▶ Selects appropriate data analysis tool or technique to solve business problem
- ▶ Applies data analysis tools and techniques

Analytics and Evaluation

Identifying Questions Using Data

Competency Description - Identifying Questions Using Data

The process of interpreting data and analytic output to identify problems and research questions, draw conclusions and support data driven decision making.

Knowledge and Tasks

- ▶ Uses data to identify problems or questions in practical situations (e.g. workplace efficiency)
- ▶ Uses data to identify higher level problems or research questions (e.g. policy, environment, scientific experimentation, marketing, economics, etc.)

Analytics and Evaluation

Data Analysis

Competency Description - Data Analysis

Accessing, manipulating, querying and analysing data using software and tools, applying statistical analysis techniques and presenting the output of data analysis to stakeholders and decision makers.

Knowledge and Tasks

- ▶ Develops analysis plans
- ▶ Applies analysis methods and tools
- ▶ Conducts exploratory analysis
- ▶ Evaluate results of analysis
- ▶ Compares results of analysis with other findings

Analytics and Evaluation

Advanced Analytics

Competency Description - Advanced Analytics

Advanced analytics tools, techniques and tools, including machine learning, artificial intelligence, predictive modelling, etc.

Knowledge and Tasks

- ▶ Knowledge of advanced analytics techniques (e.g. machine learning)
- ▶ Utilizes advanced tools and techniques (e.g. data mining, machine learning) to perform data exploration
- ▶ Knowledge of best-practice advanced modelling techniques (predictive modelling, advanced clustering, association rules)
- ▶ Ability to build accurate, valid and efficient modelling solutions using a variety of complex data types

Analytics and Evaluation Visualization

Competency Description - Visualization

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

Knowledge and Tasks

- ▶ Reads and understands charts, tables and graphs
- ▶ Creates meaningful tables to organize and visually present data
- ▶ Creates meaningful graphical representations of data
- ▶ Evaluates effectiveness of graphical representations
- ▶ Critically assesses graphical representations for accuracy and misrepresentation of data

Analytics and Evaluation Storytelling

Competency Description - Storytelling

The process of translating data analyses into layman's terms and setting the data within a storyline to influence a business decision or action.

Knowledge and Tasks

- ▶ Assesses the desired outcome(s) for presenting the data
- ▶ Assesses audience needs and familiarity with subject(s)
- ▶ Plans the appropriate meeting or presentation type
- ▶ Utilizes meaningful tables and visualizations to communicate data
- ▶ Presents arguments and/or outcomes clearly and coherently
- ▶ Knowledge of widely accepted citation methods
- ▶ Creates correct citations

Analytics and Evaluation

Evaluating Decisions Based on Data

Competency Description - Evaluating Decisions Based on Data

Ability to critically assess the efficacy of policies, programs or decisions using research, evaluation and data science/analytics techniques.

Knowledge and Tasks

- ▶ Identifies key take-away points from charts, tables and graphs, and integrates this with other important information
- ▶ Collects follow up data to assess effectiveness of decisions or solutions based on data
- ▶ Conducts analysis of follow up data
- ▶ Compares results of analysis with other findings
- ▶ Evaluates decisions or solutions based on data
- ▶ Retains original conclusions or decisions, or implements new decisions/solutions

Annex A

Generic Proficiency Level Descriptors

Proficiency Level Descriptors			
Level 1 Novice	Level 2 Apprentice	Level 3 Practitioner	Level 4 Expert
Awareness of data-related roles and practices in own organization	Ability to analyze business problems to determine solvability using data	Uses data and data practices to inform/improve policy/program development or operational activities	Leads and directs strategic and technical projects to maximize value from data assets and sources
Knowledge of ethical considerations and own organizations' policies and practices for collection and sharing of data	Consistently demonstrates appropriate behaviours in support of a healthy data culture	Can instruct others in data-related area of expertise (analytics, governance, devops, etc.)	Champions healthy data culture through developing people, enabling tools and fostering innovation to derive business value from data
Demonstrates appropriate behaviours in support of a healthy data culture	Understands own role within own organizations' data strategy	Is an active champion of healthy data culture	Can address complex ethical issues with data
	Can perform basic data tasks in area of expertise to derive value from data assets (locate, analyze, transform, visualize, etc.)	Continuously learns and develops new skills	Creates an environment of continuous learning

Annex B

DATA PERSONA



BUSINESS/POLICY ANALYST

What I do

- ⇒ Recommend improvements to systems, processes, policies, etc.
- ⇒ Use existing research, reports or processed data

What I need

- ⇒ Access to processed data
- ⇒ Basic analytics tools



FRONTLINE/OPERATIONS

What I do

- ⇒ Support the delivery of programs & services
- ⇒ Collect capture and/or create data
- ⇒ Interact directly with clients and their files

What I need

- ⇒ Access to data that helps me do my job
- ⇒ Mechanism to report inaccuracies in data



DATA STEWARD

What I do

- ⇒ Define business terms
- ⇒ Document and maintain metadata
- ⇒ Manage access to analytics environments/platforms
- ⇒ Identify, analyze & resolve data quality issues
- ⇒ Act as a liaison between business and technology

What I need

- ⇒ Access to inventory of departmental data assets
- ⇒ Mechanism to manage data issues
- ⇒ Data policies/standards



ANALYTICS CONSUMER

What I do

- ⇒ Apply analytical insights to business operations
- ⇒ Understand & interpret data findings
- ⇒ Communicate findings & results from analysis to stakeholders

What I need

- ⇒ Access to data relevant to my work
- ⇒ Clearly defined business questions



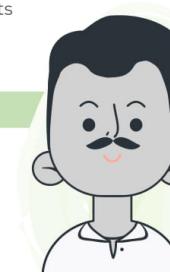
APPS DEVELOPER

What I do

- ⇒ Build/maintain analytics ecosystems, platforms & environments.
- ⇒ Embed/operationalize analytics into existing platforms

What I need

- ⇒ Technical Infrastructure
- ⇒ Testing Environments
- ⇒ Access to the business to translate problems into solutions
- ⇒ Clear business requirements



DATA PROVISIONER

What I do

- ⇒ Extract, link, cleanse and prepare data for use
- ⇒ Provide data in response to business needs

What I need

- ⇒ Access to raw data
- ⇒ Full suite of data integration tools
- ⇒ Data quality tools



DATA SCIENTIST

What I do

- ⇒ Use advanced analytics and modelling techniques to turn raw data into actionable insights
- ⇒ Develop data driven products and solutions

What I need

- ⇒ Full access to authorized raw data
- ⇒ Extensive advanced tools including open source



CITIZEN DATA SCIENTIST

What I do

- ⇒ Interpret raw data, statistics or reports and translate them into action
- ⇒ Leverage data & analytics to answer business questions
- ⇒ Recommend or build the right analytical methods to solve business problems

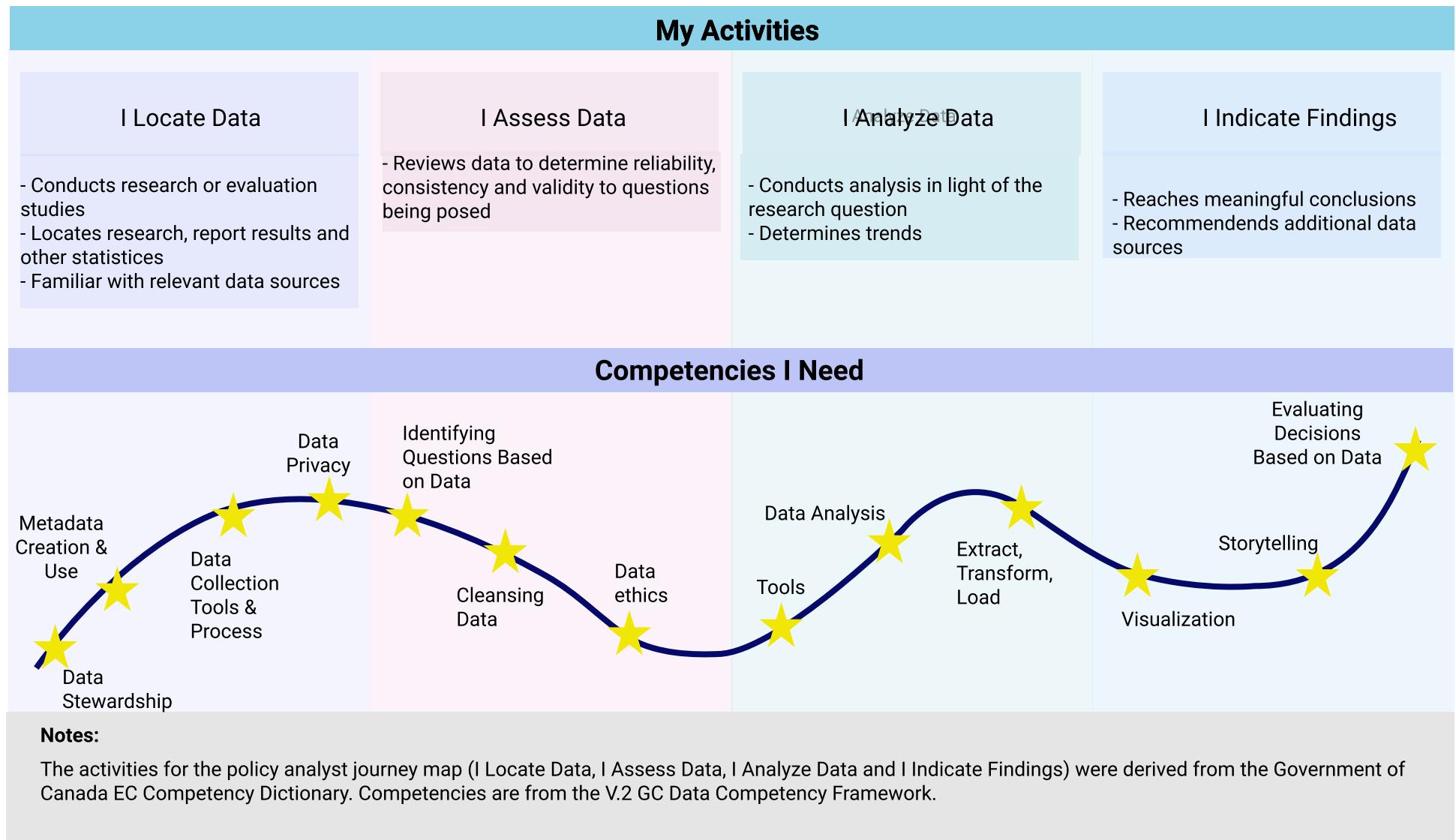
What I need

- ⇒ Access to authorized raw data
- ⇒ Advanced analytical tools

Annex C Journey Map Example

Business/Policy Analyst Journey Map

This journey map uses work description competencies for an EC policy analyst and maps them against the data competencies for the Business/Policy Analyst persona.



Annex D

Business and Policy Analyst Proficiency Mapping

This example illustrates one way of mapping the data competencies with the proficiency levels in Annex A for a particular persona.

Business/Policy Analyst	Competency	Proficiency Level
	Data Concepts	Level 3
	Data Ethics	Level 3
	Data Culture	Level 3
	Evidence Based Decision Making	Level 4
	Data Stewardship	Level 1
	Governance	Level 2
	Metadata creation and use	Level 2
	Database principles	Level 1
	Tools	Level 2
	Questions (data interpretation)	Level 3
	Data Analysis	Level 3
	Advanced Analytics	Level 1
	Visualization	Level 3
	Storytelling	Level 2-3
	Evaluating Decisions	Level 3-4

Annex E

Mapping Example

Example: Organizational Ideal

Data Ethics

