Annex A

Statement of Requirements (SOR)
for the Canada Revenue Agency (CRA)

Business Intelligence Solution

This page is intentionally left blank

Table of contents

General information (I)

The Canada Revenue Agency (CRA) is a large government agency responsible for the administration of tax programs as well as the delivery of economic and social benefits. It also administers certain provincial and territorial tax programs. The CRA has the authority to enter into new partnerships with the provinces, territories, and other government bodies to administer non-harmonized taxes and other services upon request and on a cost-recovery basis.

The CRA promotes compliance with Canada's tax legislation and regulations and plays an important role in the economic and social well-being of Canadians. The CRA is committed to working closely with stakeholders, providing excellent service to clients, and ensuring responsible enforcement of legislation.

More information on the CRA can be obtained from our Web site: www.cra-arc.gc.ca

High-Level View of SSC/CRA Computing Infrastructure

Figure 1
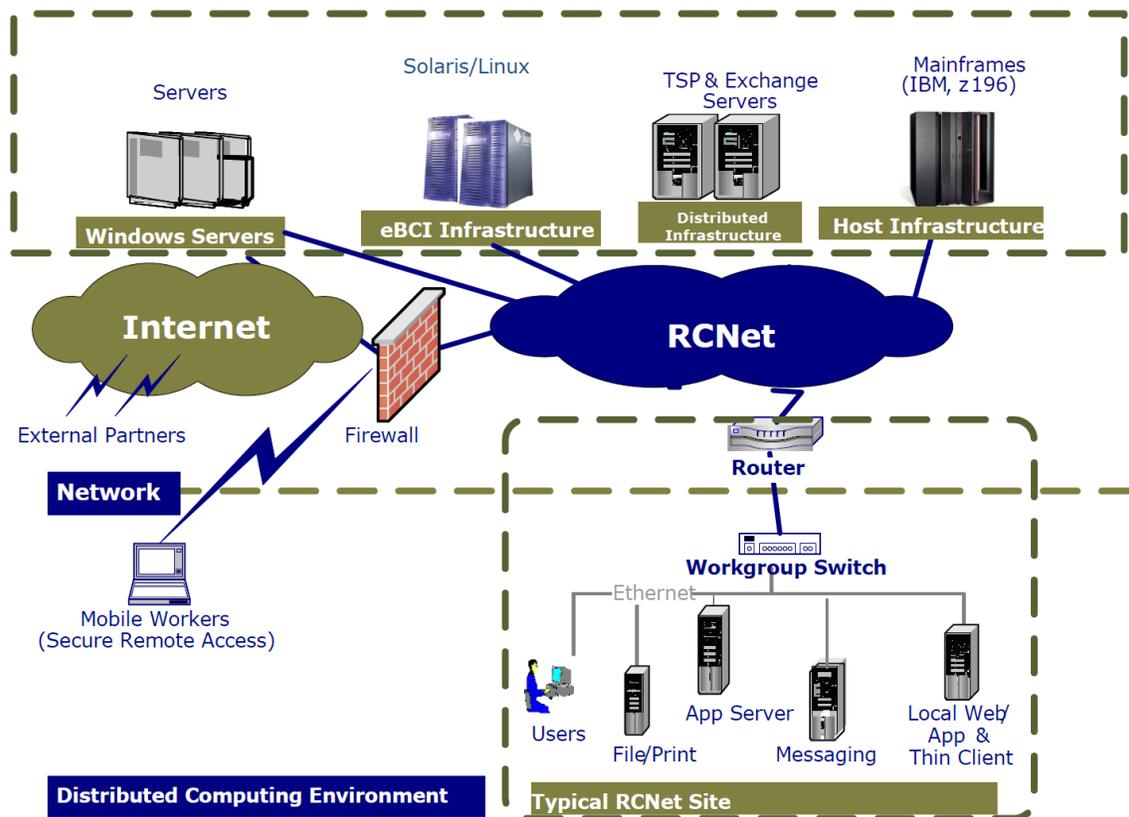
Distributed Computing Environment

The CRA's Distributed Computing Environment (DCE) is comprised of three national platforms, distributed and centralized, encompassing approximately:
1. 50,000 Desktop computers,
2. 25,000 Laptop/notebook/tablet computers, and
3. 3,000 BlackBerry devices.

The DCE is a client/server based infrastructure that consists of Microsoft Windows based servers and end-user computing devices with Microsoft Windows Active Directory (AD) providing back-end directory services.

There are approximately 400 sites across Canada supported by the DCE. These sites will vary in size from a handful of users to thousands in a single building. Bandwidth at these sites varies. A typical distributed site is comprised of one or more file and print servers, access to local or centralized Microsoft Exchange mail services, an AD domain controller, and a number of locally network desktops.

The CRA has also implemented the Centralized Technology Platform (CTP) using Citrix XenApp 6.5, which consists of central servers located in the National Capital Region hosting a variety of applications and services for of end-users. These applications and services include, but are not limited to specific line-of-business applications along with base productivity applications such as Microsoft Office, a TN3270 emulator (Attachmate), and basic file and print services. In addition, the CRA utilizes Microsoft App-V application virtualization to enhance application access and management within the CTP farm.

Secure Remote Access (SRA) users, who are not on the corporate network, can connect to the DCE via Virtual Private Networks (VPNs) through public Internet Service Providers (ISPs). The SRA platform is a subset of the DCE and is also based on Microsoft Windows Server and Windows Client operating systems.
The following bullets will highlight the key software installed within the CRA DCE:
- MS Windows Server 2008 64-bit / MS Windows Server 2008 R2 64-bit / MS Windows Server 2012 R2 64-bit / MS Windows Server 2016 64-bit
- VMWare vSphere 5.5;
- Citrix XenApp 6.5;
- Windows 7 SP1 32-bit Enterprise;
- Windows 8.1 64-Bit Enterprise;
- Windows 10 64-Bit Enterprise;
- MS Exchange 2010;
- MS Office 2010 SP2 Standard, MS Office 2010 SP2 Professional, MS Office 2013 ProPlus SP1;
- Entrust Certificate Services;
- McAfee Security Suite;
- Microsoft Internet Explorer Browser.

The underlying hardware for the Distributed Computing Environment consists of servers and end-user devices based on x86and x64 Advanced Micro Devices (AMD) and Intel processor architecture using multi-core and/or multi-processor technology.

Network Environment

Shared Services Canada (SSC) operates a Wide Area Network (WAN), referred to as RCNet, on behalf of the CRA that extends to approximately 400 sites across Canada. SSC installs multi-protocol routers in each building to interconnect user Local Area Network (LAN) segments and to provide access to the WAN. The majority of the buildings are interconnected via 1.5Mbps or higher MultiProtocol Label Switching (MPLS) circuits with various network-based Quality of Service (QoS) configurations. Internet Protocol Security (IPsec) VPN over Internet as a backup circuit is deployed at most of these sites. At certain remote locations, IPsec VPN over Internet (Digital Subscriber Line (DSL), Cable, and Satellite) is used for primary WAN access.

Electronic Business Computing Infrastructure (eBCI) Environment

The eBCI platform is a service-centric computing infrastructure designed to host and support the CRA applications from Unit Testing through Production. It is comprised of a multitude of infrastructure components and services including server and storage hardware, Web server, application integration server, messaging, database connectivity, security, directory, application testing and migration. This platform supports a set of technology standards based on Java component architecture.

Other highlights of this computing infrastructure include:

- Tier one hardware deployed for reliability;
- Maximized utilization, resiliency and flexibility through the use of virtualization technologies;
- High availability design with load-balancing and redundancy across 2 data centres, supported 7/24;
- Supports 3-tier architecture using Enterprise Java Bean (EJB) technology, integrates with existing mainframe and distributed components and services;
- Monitored and managed infrastructure based on the ITIL best practices.

The basic platform standards are the following:

- Hardware: x86 servers, SPARC Based Servers;
- Virtualization: VMWare ESX 5.5, RHEL KVM and Solaris Zones (containers);
- OS Standard: RedHat Enterprise Linux 6.x and 7.x, Oracle/Sun Solaris v.10;
- Web Server : Apache 2.2.x;
- Java Application Platform: Oracle Weblogic 11g and 12c.

Project Scope

The requirement is for Canada Revenue Agency (CRA) to acquire Commercial-Off-The-Shelf (COTS) Business Intelligence Software Solution (s) and related services to provide business intelligence functionality, including, but not limited to data mining, advanced data integration, advanced data visualization, unstructured data parsing and manipulation, statistical analysis, prescriptive analytics and predictive modelling capabilities.

Mandatory requirements (M)

Functional requirements

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
Common Mandatory Functional Requirements (Apply to Data Visualization, Unstructured Data and Advanced Analytics components)				
CR-1-M	The bidder **must** include **with the bid**, a list of all functional requirements which rely on the use of Adobe Flash or Shockwave player or their components. The bidder **must** describe the dependency related to each requirement.			
CR-2-M	The proposed solution **must** include all end-user interfaces in English and French.			
CR-3-M	The proposed solution **must** include online help in English.			
CR-4-M	The proposed solution **must** support either of the following character sets/encodings through import, export, user interfaces, reports and manual input: UTF-8 **OR** ISO 8859			
CR-5-M	The proposed solution **must** provide user access to functionality using an intuitive desktop Windows style Graphical User Interface (GUI).			
CR-6-M	The proposed solution **must** allow users the choice of input tools to be used independently and interchangeably (e.g., use of keyboard without mouse and vice versa).			
CR-7-M	The proposed solution **must** provide users with copy/paste functionality.			
CR-8-M	The proposed solution **must** allow a user to change (add, remove, modify) data or dataset(s) in an existing data visualization to form a new visualization.			
CR-9-M	The proposed solution **must** include an "undo" feature.			
CR-10-M	The proposed solution **must** have the ability to internally create business metadata for all BI objects in English.			
CR-11-M	The proposed solution **must** provide the functionality to save non-proprietary metadata content and export to IBM Infosphere Information Governance Catalog v11.5.			
CR-12-M	The proposed solution **must** provide the functionality to export data directly into the following formats: a) TXT b) XLS c) XLSX d) XML e) CSV f) HTML (including embedded graphics)			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Common Mandatory Functional Requirements (Apply to Data Visualization, Unstructured Data and Advanced Analytics components)
CR-13-M	The proposed solution **must** automatically notify, <u>at time of import</u> when imported data exceeds configured limit(s).			
CR-14-M	The proposed solution **must** provide the functionality to create, update, save and share output templates.			
CR-15-M	The proposed solution **must** support collaboration between multiple users within the tool allowing users to share and collaborate on coding and templates.			
CR-16-M	The proposed solution **must** provide the functionality for canceling a query in progress.			
CR-17-M	The proposed solution **must** provide the functionality to monitor job(s) in progress and create alerts based on user-defined parameters.			
CR-18-M	The proposed solution **must** allow the authorized users to drill up, drill down and drill across from one level of hierarchical data structures to another.			
CR-19-M	The proposed solution **must** allow the authorized users to utilize hierarchical data structures (temporal and spatial) and conversion of data into hierarchical structures.			
CR-20-M	The proposed solution **must** be able to import files and documents in the following formats: a) XLS b) XLSX c) CSV d) XML e) RTF f) PDF g) DOC h) DOCX i) PS j) PPT k) PPTX l) TXT m) HTML			
CR-21-M	The proposed solution **must** allow authorized users to read data from and write data to the following databases: a) IBM-DB2/LUW v9.7, v10.5 and newer b) Oracle 11g and newer c) IBM Netezza v7.2 and newer d) MS Access 2010 and newer e) MS SQL Server 2008 R2 and newer f) MySQL v5.6 and newer g) PostgreSQL v9.3, v9.4 and newer			
CR-22-M	The proposed solution **must** provide the functionality to analyze data and produce results within the same tool.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Common Mandatory Functional Requirements				
(Apply to Data Visualization, Unstructured Data and Advanced Analytics components)				
CR-23-M	The proposed solution **must** provide the functionality to identify trends and patterns in time and space (e.g., how data changes by province).			
CR-24-M	The proposed solution **must** provide the user a description of the methodology used in a given procedure.			
CR-25-M	All components including, but not limited to software components, drivers, plugins required for full functionality of the proposed solution assessed against Mandatory and Rated requirements **must** be included at no additional charges to CRA.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Mandatory Functional Requirements Unique to Data Visualization				
DV-1-M	The proposed solution **must** provide the functionality to change data labels within the visualization.			
DV-2-M	The proposed solution **must** be able to process large datasets (more than a billion records) without interruption.			
DV-3-M	The proposed solution **must** provide the functionality to export data directly into the following formats: a) DOC b) DOCX c) PDF d) RTF e) JPEG f) JPG g) BMP			
DV-4-M	The proposed solution **must** allow end-users to change visualization parameters without performing a new query.			
DV-5-M	The proposed solution **must** support dynamic data content (i.e. displayed data is updated in visualization/report reflecting changes of data in the input source(s)).			
DV-6-M	The proposed solution **must** provide the functionality to assemble graphics, images and text suitable for viewing on mobile devices.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
Mandatory Functional Requirements Unique to Data Visualization				
DV-7-M	The proposed solution must have the ability to simultaneously import a minimum of 3 different formats listed in **CR-21-M**, **DV-16-M OR DV-8-R** to create a visualization output. *In order to meet this requirement, the proposed solution must be able to import **either** three different formats listed in one of **CR-21-M**, **DV-16-M** or **DV-8-R** OR three different formats that are combination of one or more formats listed in **CR-21-M**, **DV-16-M OR DV-8-R**. For example, solution that simultaneously imports one format listed in each (**CR-21-M**, **DV-16-M** and **DV-8-R**) meets the requirement equally to the solution that simultaneously imports three different formats listed in **CR-21-M**.			
DV-8-M	The proposed solution **must** provide pre-built, user-customisable visualization templates, dashboard templates and report templates in English and French.			
DV-9-M	The proposed solution **must** provide users the capability to build bilingual reports.			
DV-10-M	The proposed solution **mus**t include print preview and print formatting functions.			
DV-11-M	The proposed solution **must** allow users to annotate or add comments to analysis and models and to position objects as required.			
DV-12-M	The proposed solution **must** provide the ability to add CRA branding (CRA logo) to reports.			
DV-13-M	The proposed solution **must** provide the functionality to display on a monitor the microdata that generates the final data visualization.			
DV-14-M	The proposed solution **must** provide the functionality for development of multi-colour data visualizations.			
DV-15-M	The proposed solution **must** allow users to define and use filters and prompts (e.g., pre-defined filter that prompts users to select a range).			
DV-16-M	The proposed solution **must** support the input in the following formats: a) JSON b) SHP			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
Mandatory Functional Requirements Unique to Unstructured Data				
UD-1-M	The proposed solution **must** provide multiple languages recognition capability supporting recognition of English and French at a minimum.			
UD-2-M	The proposed solution **must** include a built-in dictionary for English and French at a minimum.			
UD-3-M	The proposed solution **must** allow users to parse data based on user-specified criteria.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Mandatory Functional Requirements Unique to Unstructured Data				
UD-4-M	The proposed solution **must** allow users to save text analysis results directly into the following common industry standard file formats**:** a) HTML (including embedded graphics) b) PDF c) DOC d) DOCX e) RTF f) BMP g) JPEG h) JPG i) PNG j) TIFF			
UD-5-M	The proposed solution **must** provide the functionality to view unstructured input and the processed results.			
UD-6-M	The proposed solution **must** provide named entity recognition functionality allowing relationship, fact, and event extraction (e.g., people, places, dates, companies).			
UD-7-M	The proposed solution **must** provide the functionality to identify trends, most common words, subjects and topics.			
UD-8-M	The proposed solution **must** provide the functionality to identify the attitudinal value (sentiment, opinion and mood) of a document or record of text.			
UD-9-M	The proposed solution **must** provide data transformation functionality allowing to transform unstructured data into structured data and save the results.			
UD-10-M	The proposed solution **must** support collaboration between multiple users within the tool allowing users to define the field and structure, to create, edit, manage, share and personalize internal custom subject-related dictionaries and acronyms collaboratively, simultaneously and off-line.			
UD-11-M	The proposed solution **must** provide the functionality to organize, sort, categorize and assign priorities to unstructured data based on user-specified criteria and keywords (e.g., questionnaire responses).			
UD-12-M	The proposed solution **must** provide co-reference resolution functionality allowing identification of noun phrases and other terms that refer to the same subject or object.			
UD-13-M	The proposed solution **must** allow power users to review, revise, develop and modify the generated program code (which must be accessible and editable) for advanced functions.			
UD-14-M	The proposed solution **must** provide the functionality for identification of synonyms and typographic errors.			
UD-15-M	The proposed solution **must** provide the functionality for addressing missing values based on user-defined rules.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
	Mandatory Functional Requirements Unique to Unstructured Data			
UD-16-M	The proposed solution **must** provide users the capability to perform the following matching capabilities: a) Deterministic matching b) Probabilistic matching c) Fuzzy logic matching d) Algorithmic matching e) Similarity matching f) Stratifications and summarizations			
UD-17-M	The proposed solution **must** provide the functionality to apply the following unstructured data manipulation techniques: a) Document clustering b) Document classification c) Entity analysis & extraction d) Text mining			
UD-18-M	The proposed solution **must** support the input in the following formats: a) ASCII (mainframe flat files) b) EBCDIC (mainframe flat files)			
UD-19-M	The proposed solution **must** provide the functionality to perform data profiling based on the following parameters: a) conformity to field characterization(s) b) identification of empty field(s) c) identification of number of categories d) identification of extreme value(s) (high / low ranges) e) identification of distribution through histograms			
UD-20-M	The proposed solution **must** allow users to create unions, intersections, inner and outer joins.			
UD-21-M	The proposed solution **must** perform segmentation of text (word, sentence, topic/theme) and provide the user with descriptive measures of the segmentation performed.			
UD-22-M	The proposed solution **must** be capable of coding (pre- and post-collected coding of unstructured text) as well as facilitate the creation of a term matrix.			
UD-23-M	The proposed solution **must** provide users the capability to measure correlation based on word/sentence/topic.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
	Mandatory Functional Requirements Unique to Advanced Analytics			
AA-1-M	The proposed solution **must** be able to process large datasets (more than a billion records) without interruption.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Mandatory Functional Requirements Unique to Advanced Analytics
AA-2-M	The proposed solution **must** provide the functionality to export all report results directly into the following common industry standard formats**:** a) HTML (including embedded graphics) b) PDF c) DOC d) DOCX e) RTF f) BMP g) JPEG h) JPG i) PNG j) TIFF			
AA-3-M	The proposed solution **must** support dynamic data content (i.e. displayed data is updated in visualization/report reflecting changes of data in the input source(s)).			
AA-4-M	The proposed solution **must** provide the functionality to assemble graphics, images, and text suitable for viewing on mobile devices.			
AA-5-M	The proposed solution **must** provide pre-built, user-customisable visualization templates, dashboard templates and report templates in English and French.			
AA-6-M	The proposed solution **must** provide data transformation functionality allowing to transform unstructured data into structured data and save the results.			
AA-7-M	The proposed solution **must** provide the functionality to build bilingual reports.			
AA-8-M	The proposed solution **mus**t include print preview and print formatting functions.			
AA-9-M	The proposed solution **must** allow users to annotate or add comments to analysis and models and to position objects as required.			
AA-10-M	The proposed solution **must** provide the ability to add CRA branding (CRA logo) to the output.			
AA-11-M	The proposed solution **must** support collaboration between multiple users within the tool allowing users to define the field and structure, to create, edit, manage, share and personalize internal custom subject-related dictionaries and acronyms collaboratively, simultaneously and off-line.			
AA-12-M	The proposed solution **must** provide the functionality to perform the following basic arithmetic functions: a) add b) subtract c) multiply d) divide e) exponents			
AA-13-M	The proposed solution **must** allow power users to review, revise, develop and modify the generated program code (which must be accessible and editable) for advanced functions.			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
Mandatory Functional Requirements Unique to Advanced Analytics
AA-14-M	The proposed solution **must** allow users to create unions, intersections, inner and outer joins.			
AA-15-M	The proposed solution **must** provide the functionality for addressing missing values based on user-defined rules.			
AA-16-M	The proposed solution **must** allow users to run multiple tasks concurrently (e.g., user can conduct data analysis while a data extraction or other analysis program is running).			
AA-17-M	The proposed solution **must** provide performance evaluation functionality that allows the evaluation of multiple models and the results to be graphically displayed.			
AA-18-M	The proposed solution **must** provide the functionality to perform the following statistical analyses and techniques: a) Panel Data Analysis b) Survey Data Analysis c) Multiple Imputation d) Treatment Effect Models			
AA-19-M	The proposed solution **must** provide users the capability to perform the following analyses: a) Hypothesis testing for risk (to include parametric, non-parametric, estimate powers, sample size calculations, adjust test statistics for multiple comparisons) b) Budget and financial analysis c) Behavioural (nudge) analysis			
AA-20-M	The proposed solution **must** provide users the capability allowing application of the following Association Techniques: a) Apriori b) Carma c) Sequence			
AA-21-M	The proposed solution **must** provide the functionality allowing application of the following Classification Techniques (with machine learning support): a) Regression model (e.g., Generalized Linear Model, Average Marginal Effects for nonlinear regression models (e.g., logistic), Logistic regression, Poisson regression) b) Neural Networks c) Decision trees (e.g., C5.0, Quest, CHAID, C&R) d) Discriminant Analysis e) Support Vector Machines f) Naïve Bayes			
AA-22-M	The proposed solution **must** provide the functionality to perform Dimensionality Reduction using Principal Component Analysis (PCA) technique.			
AA-23-M	The proposed solution **must** provide the functionality to perform the following Regression Model Checking tests: a) Heteroscedasticity test b) Multicollinearity test c) Autocorrelation test			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
Mandatory Functional Requirements Unique to Advanced Analytics
AA-24-M	The proposed solution **must** provide the functionality to perform the following Time Series Model analyses: a) Timeline Analysis b) Linear and non-linear forecasting			
AA-25-M	The proposed solution **must** provide users the capability to perform <u>balanced</u> and <u>unbalanced</u> Panel Data Analysis using the following techniques: a) Fixed effects analysis b) Random effects analysis			
AA-26-M	The proposed solution **must** provide users the capability to perform the following Statistical Analysis (inferential statistics): a) F test for ANOVA b) One and Two sample test (Z and t tests) c) Fisher's test (for categorical variables) d) Survival Analysis e) Multivariate Analysis (e.g., multiple regression, MANOVA, canonical, conjoint, structural equation, multidimensional, correspondence and classifiers) f) Factor Analysis g) Non-parametric analysis h) Calculations for descriptive statistics using survey and frequency weights and summarizing the data into useful information including mean, median, mode, min and max, skewness and kurtosis, range, covariance, correlation, standard deviation, mean square error)			
AA-27-M	The proposed solution **must** provide the functionality to apply the following Visual Analytics methodologies and techniques: a) Correlation matrix b) Cluster Analysis c) Timeline Analysis d) Time series analysis e) Linear forecasting f) Non-linear forecasting g) Choropleth maps h) Street-level mapping i) Contour plots j) Map layering k) Mapping with shape files and point data l) Entity network analysis and visualization			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Mandatory Functional Requirements Unique to Advanced Analytics				
AA-28-M	The proposed solution **must** provide users the capability to perform the Record Linkage (RL) tasks below. For each, the user must have the ability to set blocks or pockets, set thresholds, use frequency weights, specify rules and weights, determine the linkage strategy, and output all linked datasets. a) Deterministic matching b) Algorithmic matching c) Fuzzy logic matching d) Probabilistic matching e) Similarity matching			
AA-29-M	The proposed solution **must** provide the functionality to perform the following query operations: a) Records selection b) Merging c) Sorting d) Aggregation e) Balancing f) Append g) Distinct h) Derive i) Filter j) Filler k) Reclassify l) Bin m) Partition n) Restructure o) Transpose p) Field reorder q) String search r) Numeric search			

Req. No.	Requirement Description	Compliant		Reference
		Yes	**No**	
Mandatory Functional Requirements Unique to Advanced Analytics

Req. No.	Requirement Description	Yes	No	Reference
AA-30-M	The proposed solution **must** provide data simulation functionality to generate random data from the following distributions: a) Bernoulli b) Beta c) Binomial d) Cauchy e) Chi-squared f) Exponential g) F h) Gamma i) Geometric j) Hypergeometric k) Log-normal l) Logistic m) Multinomial n) Negative binomial o) Normal p) Poisson q) Students t r) Uniform s) Weibull			
AA-31-M	The proposed solution **must** provide the following data sampling functionally: a) Simple Random b) Stratified c) Weighted d) Cluster e) Systematic f) First N g) Rare event sampling			
AA-32-M	The proposed solution **must** provide the functionality to generate random numbers from any distribution having a cumulative distribution function using: a) Inverse transform method b) Monte Carlo method			
AA-33-M	The proposed solution **must** provide the functionality to cast (convert) the following types of data from one type to another: a) numeric b) character c) date d) time e) alpha f) alpha-numeric			

Req. No.	Requirement Description	Compliant		Reference
		Yes	No	
Mandatory Functional Requirements Unique to Advanced Analytics			
AA-34-M	The proposed solution **must** allow users to define and use filters and prompts (i.e., pre-defined filter that prompts users to select a range).			
AA-35-M	The proposed solution **must** support the input in the following formats: a) JSON b) SHP			
AA-36-M	The proposed solution **must** provide the functionality to perform data profiling based on the following parameters: a) Conformity to field characterization(s) b) Identification of empty field(s) c) Identification of number of categories d) Identification of extreme value(s) (high / low ranges) e) Identification of distribution through histograms			
AA-37-M	The proposed solution **must** be capable of predictive analytics (hierarchical/quadratic/mixed linear, ridge regression, splines, cox-proportional, ARIMA).			

Non-Functional Requirements

Technical Requirements

Req. No.	Mandatory Technical Requirements	Compliant		Reference
		Yes	**No**	
TR-1-M	If the proposed solution includes Adobe Flash or Shockwave players or components (included or embedded), the bidder **must** provide the CRA with the capability to remove and/or upgrade the impacted components of the proposed solution on or before January 2020 without impact to functionality required for meeting any of the "Mandatory" or "Rated" requirements; at no additional cost to the CRA.			
TR-2-M	If the proposed solution includes Adobe Flash or Shockwave players or components (included or embedded), the vendor **must** allow the CRA to disable Adobe Flash or Shockwave players and components in such a way that they cannot be run by a user or program/script and without impact to functionality required for meeting any of the "Mandatory" Requirements.			
TR-3-M	The proposed solution **must** be able to run on existing SSC/CRA Computing Infrastructure *(Figure 1)* without customization of the infrastructure.			
TR-4-M	The proposed solution, if server infrastructure is required, **must** operate in **at least one** of the following two operating system environments: Windows Server 2012 R2 – all versions for x64 microprocessor architectures, and any subsequent releases of Windows Server, **OR** Red Hat Enterprise Linux (RHEL) v6.x or v7.x server for x64 microprocessor architectures, and any subsequent releases of RHEL			
TR-5-M	The proposed solution **must** be fully compatible with 64 bit binaries on computer systems based on x86-64 architecture.			
TR-6-M	The proposed solution, if server infrastructure is required, **must** support running in the VMWare virtualized environment.			
TR-7-M	The proposed solution **must** integrate with existing authentication and authorization solutions in use by CRA (CA Siteminder & CA Identity Manager)			
TR-8-M	The proposed solution **must** support existing server and application monitoring systems in use by CRA either through SNMP traps, log monitoring or monitoring agent integration.			
TR-9-M	The proposed solution **must** be scalable, allowing for and making use of the addition of server computing capacity such as additional CPUs and memory.			
TR-10-M	The bidder **must** provide a description of the hardware requirements to support the proposed solution (i.e., Server and Workstation configuration, topology).			
TR-11-M	The proposed solution **must** be supported on the client Windows 10 platform.			

Req. No.	Mandatory Technical Requirements	Compliant		Reference
		Yes	**No**	
TR-12-M	The proposed solution **must** be commercial-off-the-shelf (COTS), freely available for purchase, have a published product definition and pricing structure, and have an ongoing funded development and support investment behind it. If the solution consists of multiple, independent products, each proposed product must be COTS as defined above. ALPHA or BETA versions of the product do not qualify as COTS.			
TR-13-M	The proposed solution **must** be entirely contained within the CRA environment and not depend on or require data of any kind to be sent outside the CRA environment for processing. No "service" or "service bureau" based solutions will be acceptable.			
TR-14-M	The proposed solution **must** not require any form of external connection to activate any component of the product. Requiring a serial number or other form of license activation, string or code, is acceptable, but the solution must not require any form of on-going electronic or Internet-enabled validation or check for license compliance.			

Security Requirements

Req. No.	Mandatory Security Requirements	Compliant		Reference
		Yes	**No**	
SR-1-M	The provider of the proposed solution agrees to supply documentation and evidence to support a Security Assessment (such as a Threat Risk Assessment), if requested to do so by the CRA.			
SR-2-M	The proposed solution **must** not interfere with the operation of any Anti-Virus, Anti-Malware or Host Intrusion Detection systems on a host computer.			
SR-3-M	The proposed solution **must** support the enforcement of a role-based access control policy over defined subjects and objects.			
SR-4-M	The proposed solution **must** provide authorized users the ability to perform, at minimum, the following user management functions: a) Create, read, update, and delete roles; b) Assign specific features and functions to specific user roles; c) Assign, revoke, and update a user group to one or more roles; d) Allow all users to have a default role.			
SR-5-M	The proposed solution **must** automatically terminate/disable temporary and emergency accounts after the CRA predetermined period or it is able to leverage similar functionality within the organization identity and access management solution.			
SR-6-M	The proposed solution **must** automatically disable inactive user accounts after a CRA predetermined inactive time period (e.g. 30 days) or is able to leverage similar functionality within the organization identity and access management solution.			

Req. No.	Mandatory Security Requirements	Compliant		Reference
		Yes	No	

SR-7-M	The proposed solution **must** leverage organization identity and access management solution or **must**:
	   a) enforce a limit of consecutive invalid login attempts (e.g. 5 attempts) by a user during a CRA defined period (e.g.10 minutes);
	   b) automatically lock the account/node until released by an administrator when the maximum number of unsuccessful attempts is exceeded.
	This is applicable regardless of whether the login occurs via a local or network connection.			

SR-8-M	The proposed solution **must** provide the CRA the ability to define, collect and store audit records associated with any user operations performed within the proposed solution, including at a minimum:
	   a) Successful and unsuccessful attempts to access, modify, or delete security objects (Security objects include audit data, system configuration files and file or users' formal access permissions.);
	   b) Successful and unsuccessful logon attempts;
	   c) Privileged activities or other system level access.			

SR-9-M	The proposed solution **must** include the ability to audit the execution of privileged functions and provide the ability to define specific events which generate audit entries. These events include, but are not limited to:
	   a) Successful and unsuccessful attempts to access, modify, or delete security objects (Security objects include audit data, system configuration files and file or users' formal access permissions.);
	   b) Successful and unsuccessful logon attempts;
	   c) Privileged activities or other system level access.			

SR-10-M	The proposed solution **must** automatically audit the following account management events, or it is able to leverage similar functionality within the organization identity and access management solution.
	   a) account creation;
	   b) account modification;
	   c) account suspension;
	   d) account termination;
	   e) account deletion;
	   f) notifies CRA predefined individuals on any of the above listed events.			

SR-11-M	The proposed solution **must** provide audit record generation that contains information to, at a minimum establish:
	   a) what type of event occurred;
	   b) when (date and time) the event occurred;
	   c) where the event occurred;
	   d) the source of the event;
	   e) the outcome (success or failure) of the event;
	   f) and the identity of any individuals or subjects associated with the event.			

SR-12-M	The proposed solution **must** allow CRA to define the time source or use internal system clock to generate time stamps for audit records.			

Req. No.	Mandatory Security Requirements	Compliant		Reference
		Yes	No	
SR-13-M	The proposed solution **must** produce audit logs composed of audit records in a human-readable format.			
SR-14-M	The proposed solution must protect audit information and audit tools from unauthorized access, modification, and deletion.			
SR-15-M	The proposed solution **must** alert CRA designated officials in the event of an audit processing failure event.			
SR-16-M	The proposed solution, if using an internal database, **must** provide a warning when allocated audit record storage volume reaches CRA predefined threshold values.			
SR-17-M	The proposed solution **must** uniquely identify and authenticate organizational users of the solution (or processes acting on behalf of organizational users) or leverage similar functionality within the organization's identity and access management solution.			
SR-18-M	For password-based authentication the proposed solution **must** leverage the organization identity and access management solution or **must**: a) Enforces minimum password complexity of a mix of upper-case letters, lower-case letters, numbers, and special characters, including minimum of 8 characters; b) Enforces at least a 5 changed characters when new passwords are created; c) Enforces password minimum and maximum lifetime restrictions of 90 days; d) Prohibits password reuse for 10 generations; e) Encrypts passwords in storage and in transmission using Communications Security Establishment (CSE) approved cryptographic algorithms as described in ITSP.40.111 (https://www.cse-cst.gc.ca/en/node/1831/html/26515).			
SR-19-M	The proposed solution **must** obscure feedback of authentication information during the authentication process to protect the information from possible exploitation/use by unauthorized individuals.			
SR-20-M	The proposed solution, if it includes a web console, **must** be accessible to users via a secure web interface (i.e. HyperText Transfer Protocol Secure (HTTPS).) using Communications Security Establishment (CSE) approved protocols (e.g. TLS 1.2) and cryptographic algorithms specified in ITSP.40.111 (https://www.cse-cst.gc.ca/en/node/1831/html/26515).			
SR-21-M	The proposed solution, if it includes a web console **must** provide mechanisms to protect the authenticity of communications sessions including but not limited to: a) generating a unique, randomly generated session identifier for each session; b) session ID must be at least 128 bits; c) recognizes only session identifiers that are system generated; d) invalidates session identifiers upon user logout, session timeout or other session termination.			
SR-22-M	The proposed solution **must** employ cryptographic mechanisms for protection of data in motion that have been approved by CSE and validated by the Cryptographic Algorithm Validation Program (CAVP), and are specified in ITSP.40. 111 (https://www.cse-cst.gc.ca/en/node/1831/html/26515).			

Req. No.	Mandatory Security Requirements	Compliant		Reference
		Yes	No	
SR-23-M	The proposed solution **must** support the use of secure protocols for user authentication (e.g. Lightweight Directory Access Protocol - LDAPs) and employ mechanisms for authentication to a cryptographic module that meet the requirements of applicable CSE ITSP.40.111 guidance for such authentication (https://www.cse-cst.gc.ca/en/node/1831/html/26515).			
SR-24-M	The proposed solution **must** not use or contain unencrypted static authenticators embedded in applications.			
SR-25-M	The proposed solution **must** enforce authentication and approved authorizations for any logical access to information and system resources in accordance with the access control policies defined by the CRA **AND** perform authorization checks before performing any action that creates, views, updates, transmits or deletes data.			
SR-26-M	The proposed solution **must** not release information outside the CRA established system boundary defined by the project requirements. Any release of information **must** be authorized by CRA.			
SR-27-M	The proposed solution **must** prevent non-privileged users from executing privileged functions to include disabling, circumventing, or altering implemented security safeguards/countermeasures.			
SR-28-M	The proposed solution **must** prevent the presentation of information system management-related functionality at an interface for general (i.e. non-privileged) users.			
SR-29-M	The proposed solution **must** provide a readily observable logout capability whenever authentication is used to gain access to web pages.			
SR-30-M	The proposed solution **must** support commercially available Vulnerability Assessment scan tools for web application components.			

References

Req. No.	Mandatory Reference Requirements	Compliant		Reference
		Yes	**No**	
RR-1-M	The bidder **must** provide with the bid two (2) corporate references where the proposed solution has been successfully implemented (integrated whole or in part) for similar sites that have been using the proposed solution, (i.e., installed and operational for at least six (6) months prior to the issuance date of this RFP). In order to demonstrate, each reference must include: a) the business name and location; b) the contact's full name, position title, phone number and e-mail address; c) date or date range from-to (YYYY-MM-DD) of the implementation. ** References may be verified.*			
RR-2-M	The bidder **must** provide with the bid two (2) corporate references where the proposed training has been successfully provided. In order to demonstrate, each reference must include: a) the business name and location; b) the contact's full name, position title, phone number and e-mail address; c) date or date range from-to (YYYY-MM-DD) of the training. ** References may be verified.*			

Services

Req. No.	Mandatory Services Requirements	Compliant		Reference
		Yes	No	
MS-1-M	The bidder **must** offer in class AND online training in both English and French.			
MS-2-M	The bidder **must** offer training for the end user(s), which will be exercised at the CRA's option during the life of the contract.			
MS-3-M	The bidder **must** offer training for the administrator(s), to be exercised at the CRA's option during the life of the contract.			
MS-4-M	The bidder **must** provide a description of all management consulting services available for the implementation of the proposed solution.			
MS-5-M	The bidder **must** offer the option for the services of an on-site technical software support person on a per diem basis as and when requested by the CRA, who will provide assistance with design, installation, configuration, problem resolution and ongoing maintenance of the solution, system performance tuning, and the implementation of new software components and versions. ** Per diem rates must be provided with the bid (for reference purpose only.)*			

Rated requirements (R)

Functional requirements

For each point-rated functional requirement that the vendor has identified as requiring Adobe Flash or Shockwave players or components and by CRA disabling Adobe Flash or Shockwave players and components, said capability is no longer available to users as described in the bidder's submission to CR-1-M, the vendor will receive zero points.

Req. No.	Requirement Description	Point Value	Reference
	Common Rated Functional Requirements (Apply to Data Visualization, Unstructured Data and Advanced Analytics components)		
CR-1-R	The proposed solution provides online help in French.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-2-R	The proposed solution includes context sensitive help in English and French.	**Maximum 8 points** 4 points: for context sensitive help in English 4 points: for context sensitive help in French	
CR-3-R	The proposed solution includes a drag and drop interface.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-4-R	The proposed solution provides point and click functionality.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-5-R	The proposed solution provides standard tools for navigation, including: a) scroll bars b) next and previous c) select a page d) hot keys	**Maximum 12 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points	
CR-6-R	The proposed solution provides users the ability to customize the onscreen layout (e.g., size of notes, icons, and changing the canvas).	**Maximum 4 points** YES: 4 points *No partial points will be awarded*	
CR-7-R	The proposed solution provides users the functionality to email the output results directly from within the solution.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
CR-8-R	The proposed solution provides multi-screen (2 files viewed concurrently on dual monitors) and multi-panel capabilities (side-by-side view of 2 files on one monitor).	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
CR-9-R	The proposed solution **does** not impose any limits on the size of imported data.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-10-R	The proposed solution includes the capability to: a) Schedule jobs to run during off peak hours or repetitively at intervals (daily, weekly, monthly) b) Revise, reschedule or delete scheduled jobs c) Notify users when the job is completed and if an error or unexpected interruption occurs d) Notify through e-mail when the job is completed and if an error or unexpected interruption occurs	**Maximum 12 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points	

Req. No.	Requirement Description	Point Value	Reference
Common Rated Functional Requirements (Apply to Data Visualization, Unstructured Data and Advanced Analytics components)		
CR-11-R	The proposed solution provides users the functionality to embed the following components within a report or dashboard: a) files b) hyperlinks c) comments	**Maximum 9 points** a) = 3 points b) = 3 points c) = 3 points	
CR-12-R	The proposed solution provides users with an auto-save feature.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-13-R	The proposed solution allows users to apply naming standards on metadata items.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
CR-14-R	The proposed solution provides users the functionality to print and save queries, code and logs.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-15-R	The proposed solution provides users the functionality to combine data from multiple file formats and databases within a single job.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-16-R	The proposed solution provides users the functionality to export data directly into the following formats: a) ACCDA b) ACCDR c) ACCDT d) MDB	**Maximum 8 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points	
CR-17-R	The proposed solution includes versioning and rollback features for users	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-18-R	The proposed solution allows users to import from the following common industry standard formats: a) BMP b) PNG c) JPEG d) JPG e) GIF f) TIFF	**Maximum 12 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points e) = 2 points f) = 2 points	
CR-19-R	The proposed solution does not impose any limits on the file size of the outputs.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
CR-20-R	All browser based interfaces of the proposed solution conform to Web Accessibility Guidelines (WCAG) 2.0 at level AA.	**Maximum 15 points** YES: 15 points *No partial points will be awarded*	
	Maximum Points Available	**212 points**	
	Minimum Threshold Requirement (60%)	**127 points**	
Bidder's Points		

Req. No.	Requirement Description	Point Value	Reference
Rated Functional Requirements Unique to Data Visualization			
DV-1-R	The proposed solution allows users to share interactive visualizations with various levels of aggregation.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
DV-2-R	The proposed solution provides users the functionality to perform side-by-side / comparative analysis.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
DV-3-R	The proposed solution provides users the functionality to perform embedded row and column operations.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
DV-4-R	The proposed solution allows users to review and modify the existing code for development of custom visualizations.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
DV-5-R	The proposed solution provides the functionality to export data directly into the following common industry standard formats: a) JSON b) SHP c) PPT d) PPTX e) PS f) GIF g) PNG h) TIFF i) DBF	**Maximum 9 points** a) = 1 point b) = 1 point c) = 1 point d) = 1 point e) = 1 point f) = 1 point g) = 1 point h) = 1 point i) = 1 point	
DV-6-R	The proposed solution provides users the functionality to create dynamic and animated visualizations.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
DV-7-R	The proposed solution provides the ability for users to create the following advanced visualizations: a) Heat maps b) Clusters c) Geospatial maps d) 2 and 3 dimensional charts and graphs e) Word clouds f) Infographics g) Decomposition trees h) Temporal analytics i) Spatio-temporal analytics j) Parallel coordinate plots k) Network diagrams	**Maximum 33 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points e) = 3 points f) = 3 points g) = 3 points h) = 3 points i) = 3 points j) = 3 points k) = 3 points	

Req. No.	Requirement Description	Point Value	Reference
Rated Functional Requirements Unique to Data Visualization			
DV-8-R	The proposed solution supports import from the following file formats: a) MXD b) MXT c) LYR d) PMF e) PRJ f) SHX g) DBF	**Maximum 7 points** a) = 1 point b) = 1 point c) = 1 point d) = 1 point e) = 1 point f) = 1 point g) = 1 point	
DV-9-R	The proposed solution provides the functionality to allow users to import and link to animations and audio and video clips in the following formats: a) GIF b) FLASH c) MPEG d) MP4 e) AVI f) MOV g) MP3 h) WAV i) WMV j) FLV k) AAC l) WMA	**Maximum 18 points** * 2 points per format up to maximum of **18 points**	
Maximum Points Available		**123 points**	
Minimum Threshold Requirement (60%)		**74 points**	
Bidder's Points			

Req. No.	Requirement Description	Point Value	Reference
Rated Functional Requirements Unique to Unstructured Data			
UD-1-R	The proposed solution includes multiple languages recognition capability allowing recognition of other languages in addition to English and French.	**Maximum 10 points** *2 points per language in addition to English and French, up to maximum of **10 points**	
UD-2-R	The proposed solution includes a built-in dictionary for other languages in addition to English and French.	**Maximum 10 points** *2 points per built-in dictionary in addition to English and French, up to max. of 10 points	
UD-3-R	The proposed solution provides the functionality to allow users to capture and report on utilization statistics, provide metrics on: a) How long it took to run b) How many records were processed c) How many records were skipped d) Number of time outs	**Maximum 12 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points	

Req. No.	Requirement Description	Point Value	Reference
	Rated Functional Requirements Unique to Unstructured Data		
UD-4-R	The proposed solution provides users with a web crawling functionality.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
UD-5-R	The proposed solution provides the functionality to exclude characters from a search.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
UD-6-R	The proposed solution provides the functionality to save text analysis results directly into the following common industry standard file formats: a) PPT b) PPTX c) PS d) GIF	**Maximum 8 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points	
UD-7-R	The proposed solution includes the capability for users to create the following advanced visualizations: a) Word clouds b) Charts and graphs c) Document cluster maps	**Maximum 9 points** a) = 3 points b) = 3 points c) = 3 points	
UD-8-R	The proposed solution supports import in .TEX format.	**Maximum 3 points** YES: 3 points *No partial points will be awarded*	
Maximum Points Available		**68 points**	
Minimum Threshold Requirement (60%)		**41 points**	
Bidder's Points			

Req. No.	Requirement Description	Point Value	Reference
	Rated Functional Requirements Unique to Advanced Analytics		
AA-1-R	The proposed solution provides the functionality to capture and report on utilization statistics to provide metrics on: a) How long it took to run b) How many records processed c) How many records skipped d) Number of time outs	**Maximum 12 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points	
AA-2-R	The proposed solution provides a web crawling functionality.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
AA-3-R	The proposed solution provides the functionality to identify duplicates and notify the users for appropriate actions.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
AA-4-R	The proposed solution allows users to review data at each stage of data manipulation.	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
AA-5-R	The proposed solution provides the functionality to perform side-by-side / comparative analysis.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	

Req. No.	Requirement Description	Point Value	Reference
	Rated Functional Requirements Unique to Advanced Analytics		
AA-6-R	The proposed solution provides the functionality to create digital surveys with the following features: a) Online data collection and analysis b) Skip logic c) Option to output printed surveys d) Integrate with distribution data e) Response tracking (compare with survey distribution list to remind subjects to respond)	**Maximum 10 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points e) = 2 points	
AA-7-R	The proposed solution provides the functionality for users to perform embedded row and column operations.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
AA-8-R	The proposed solution provides a flexible GUI that does not impose any limits within the solution on the number of the following elements in a data query: a) Table joins b) Keys (combinations of fields to create a unique identifier) c) Selected variables d) Selected records e) Selection criteria f) Sorting levels (number of variables on which to sort) g) Renamed variables h) Calculated variables i) Levels of aggregation (grouping levels) ***Bidder to provide the limit(s) with the bid in the event that any limit(s) is/are imposed.***	**Maximum 18 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points e) = 2 points f) = 2 points g) = 2 points h) = 2 points i) = 2 points *No points will be provided if limits are imposed by the solution.*	
AA-9-R	The proposed solution provides the functionality to export all report results directly into the following common industry standard formats**:** a) PPT b) PPTX c) PS d) GIF	**Maximum 8 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points	
AA-10-R	The proposed solution provide the functionality to perform the following analyses: a) Item response theory analysis and modelling b) Forensic analysis c) Entity (graph) analytics d) Risk analysis	**Maximum 12 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points	
AA-11-R	The proposed solution supports import from the following formats: a) MXD b) MXT c) LYR d) PMF e) PRJ f) SHX g) DBF	**Maximum 14 points** a) = 2 points b) = 2 points c) = 2 points d) = 2 points e) = 2 points f) = 2 points g) = 2 points	

Req. No.	Requirement Description	Point Value	Reference
Rated Functional Requirements Unique to Advanced Analytics			
AA-12-R	The proposed solution supports the import directly from the following BI Tools proprietary data file formats: a) SAS 9.x and future releases b) SPSS Modeler 14.2 and above c) SAP Lumira d) SAP Data services e) Cognos Impromptu f) Microsoft SQL Server Analysis Services (SSAS)	**Maximum 18 points** a) = 3 points b) = 3 points c) = 3 points d) = 3 points e) = 3 points f) = 3 points	
AA-13-R	The proposed solution provides users the capability to perform the following Cluster Analyses: a) K-means b) Kohonen c) Minimax d) Sequence Clustering e) MARKOV Chain Analysis	**Maximum 25 points** a) = 5 points b) = 5 points c) = 5 points d) = 5 points e) = 5 points	
Maximum Points Available		**161 points**	
Minimum Threshold Requirement (60%)		**97 points**	
Bidder's Points			

Non-Functional Requirements

Rated Security Requirements

Req. No.	Requirement Description	Point Value	Reference
Rated Security Requirements			
SR-1-R	The proposed solution **should** support commercially available Vulnerability Assessment scan tools for internal database(s).	**Maximum 12 points** YES: 12 points *No partial points will be awarded*	
SR-2-R	The proposed solution **should** enable the CRA to display a custom preapproved system use notification message or banner before granting access to the system that provides privacy and security notices in accordance with TBS Policy on the Use of Electronic Networks.	**Maximum 5 points** YES: 5 points *No partial points will be awarded*	
SR-3-R	The proposed solution **should** support two-factor authentication with policy control so that two-factor authentication may be applied to some users but not others.	**Maximum 8 points** YES: 8 points *No partial points will be awarded*	
SR-4-R	The proposed solution **should** provide the CRA the ability to define, collect and store audit records associated with any user operations where CRA defined identifiable information (e.g. PII) is presented to the user (e.g. Print, display).	**Maximum 15 points** YES: 15 points *No partial points will be awarded*	

Appendix A – Glossary

Term	Definition
Algorithmic matching	A statistical method used to solve graph matching problems in graph theory.
Alpha version	An alpha version is the first stage of product or software development. Alpha versions of computer software are in the early testing phase.
Analysis - Behavioural (nudge)	A concept in behavioural science, political theory and economics that proposes the use of positive reinforcement and indirect suggestions to influence the motives, incentives and decision making of groups and individuals and achieve unforced or voluntary compliance.
Analysis - Budget and financial	The financial planning techniques that help business personnel in the decision-making process.
Analysis - Entity (graph) analytics	An incremental context accumulator for detecting like and related entities across large, sparse, and disparate collections of data, including both new and old data, small and big data environments, to perform analytics on events, people, things, transactions, and relationships.
Analysis - Forensic data	A branch of digital forensics that examines structured data to discover and analyse patterns of fraudulent activity in order to identify incidents of financial crime.
Analysis - Item response theory analysis and modelling	A paradigm for the design, analysis, and scoring of tests, questionnaires, and similar instruments measuring abilities, attitudes, or other variables.
Analysis - Risk analysis	The process of defining and analyzing the dangers to individuals, businesses and government agencies posed by potential natural and human-caused adverse events.
Analytics - Hierarchical predictive	Encompasses a variety of statistical techniques from predictive modeling, machine learning, and data mining that analyze current and historical facts to make predictions about future or otherwise unknown events.
Analytics - ARIMA predictive	Autoregressive Integrated Moving Average (ARIMA) models are fitted to time series data either to better understand the data, or predict future points in the series (forecasting).
Analytics – Cox-proportional predictive	A method for investigating the effect of several variables upon the time a specified event takes to happen.
Analytics – Mixed linear predictive	An analytic tool used to not only perform a regression analysis on genotypic data while correcting for cryptic relatedness and pedigree structure, but also provide an estimation of random effects related to genotypic data.
Analytics – Quadratic predictive	Predictive analytics is the branch of advanced analytics used to analyze data and make predictions about unknown future events using various techniques including data mining, statistical modeling, machine learning, artificial intelligence, etc. The analytical precision of a functional set can be extended from linear to quadratic using the products of the original functions.
Analytics – Ridge regression predictive	A predictive technique designed to fit a multiple regression model when the independent variables exhibit multicollinearity.
Analytics – Splines predictive	A non-parametric technique that builds flexible models by fitting piecewise linear regressions.
Animation	A dynamic visual medium produced from static drawings, models, or objects posed in a series of incremental movements that are then rapidly sequenced to give the illusion of lifelike motion * Note: this definition does not include video, which is the result of differences in the images within individual video frames, and is not created by the display application.
Append query operation	An operation which creates a new query that contains all rows from a first query followed by all rows from a second query.
Application	A computer program designed to perform a group of coordinated functions, tasks, or activities for the benefit of the user.
Apriori association technique	A technique which identifies the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database.

Term	Definition
ASCII	American Standard Code for Information Interchange format.
Autocorrelation test	A test for similarity between observations as a function of the time lag between them.
Auto-save	A function in computer applications or programs which saves an opened document automatically, helping to reduce the risk or impact of data loss in case of a crash or freeze.
Balancing query operation	A function of many relational database management systems used to determine the most efficient way to execute a given query and is based on a trade-off between the amount of time required to find the best query path and the quality of the choice. Database management systems vary in the way that they balance quality and time efficiencies.
Bernoulli	The probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability $q = 1 - p$.
Beta version	A beta version is a nearly complete prototype of a product or software with some feature or specific part of the plugin currently in testing mode. Beta versions can cause issues if not used properly.
BI object	A product of business intelligence set of tools including reports, dashboards, visualizations, etc.
Bilingual	A reference to material being in both English and French.
Bin query operation	Evaluated using per-bin bitvectors returned by FastBit, FastBit queries, and a baseline sequential scan method.
Binomial	An algebraic expression of the sum or the difference of two terms.
BMP	Bitmap Image file.
Capability	The ability to perform or achieve certain actions or outcomes.
Carma association technique	A technique using an efficient two-pass method for finding sequences.
Cauchy	The distribution of the x-intercept of a ray issuing from (x 0 , γ) (x_{0},γ) with a uniformly distributed angle. It is also the distribution of the ratio of two independent normally distributed random variables if the denominator distribution has mean zero.
Character sets	A defined list of characters recognized by the computer hardware and software.
Chi-squared	Any statistical hypothesis test wherein the sampling distribution of the test statistic is a chi-squared distribution when the null hypothesis is true.
Choropleth maps	A map that uses differences in shading, coloring, or the placing of symbols within predefined areas to indicate the average values of a property or quantity in those areas.
Classification technique – Average Marginal Effects for nonlinear regression models (e.g., logistic)	The average of the instantaneous rates of change for all variables in the model.
Classification technique – Decision trees (e.g., C5.0, Quest, CHAID, C&R)	A decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm.
Classification technique – Discriminant Analysis	A statistical analysis to predict a categorical dependent variable (called a grouping variable) by one or more continuous or binary independent variables (called predictor variables).
Classification technique – Generalized Linear Model	A flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution.
Classification technique – Logistic regression	A regression model where the dependent variable (DV) is categorical.
Classification technique – Naïve Bayes	A family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features.
Classification technique – Neural Networks	A learning algorithm that is inspired by the structure and functional aspects of biological neural networks.
Classification technique – Poisson regression	A regression technique that assumes the response variable Y has a Poisson distribution, and assumes the logarithm of its expected value can be modeled by a linear combination of unknown parameters.

Term	Definition
Classification technique – Regression model	A statistical process for estimating the relationships among variables.
Classification technique – Support Vector Machines	A discriminative classifier formally defined by a separating hyperplane.
Cluster Analyses – K-means	A method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining.
Cluster Analyses – Kohonen	A type of neural network that perform clustering, also known as a knet or a self-organizing map.
Cluster Analyses – MARKOV Chain Analysis	A stochastic process satisfies the Markov property if one can make predictions for the future of the process based solely on its present state just as well as one could knowing the process's full history, hence independently from such history.
Cluster Analyses – Minimax	A decision rule used in decision theory, game theory, statistics and philosophy for minimizing the possible loss for a worst case (maximum loss) scenario.
Cluster Analyses – Sequence Clustering	The algorithm finds the most common sequences, and performs clustering to find sequences that are similar.
Cluster Analysis	The task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters).
Cluster data sampling functionally	A sampling plan used when mutually homogeneous yet internally heterogeneous groupings are evident in a statistical population.
Code	A collection of computer instructions, possibly with comments, written using a human-readable programming language.
Collaboration	The ability to communicate and work jointly with other users **within** the solution.
Conformity to field characterization(s)	A collection of various techniques used in the field help to characterize the optical performance of point- and line-focus optical concentrators.
Contour plots	A graphical technique for representing a 3-dimensional surface by plotting constant z slices, called contours, on a 2-dimensional format.
Co-reference resolution functionality	The retrieving and clustering of different mentions referring to the same entity.
Correlation matrix	The matrix of Pearson product-moment correlation coefficients between each of the random variables in a random vector.
COTS	Commercial-Off-The-Shelf. * *ALPHA or BETA versions of the product **DO NOT** qualify as COTS.*
CRA	Canada Revenue Agency.
CSV	Character Separated Value file format.
CTP	Centralized Technology Platform.
Dashboard	A visual display of the most important information, consolidated and arranged in an easy-to-read format to support business with meaning and useful data; for example, by metadata item status.
Data export	An automated or semi-automated output of data sets between different software applications.
Data extraction	The process of retrieving data out of data sources for further data processing or data storage.
Data import	An automated or semi-automated input of data sets between different software applications.
Data parsing	The process of analyzing text made of a sequence of tokens to determine its grammatical structure.
Data profiling	The process of examining data available from an existing information source (e.g., a database or a file) and collecting statistics or informative summaries about that data.
Data sampling	A statistical analysis technique used to select, manipulate and analyze a subset of data points in order to identify patterns and trends in the larger data set being examined.
Data transformation functionality	The functionality allowing conversion of data or information from one format to another, from the format of a source system into the required format.

Term	Definition
Data visualization	The presentation of requested data in a clear and effective pictorial or graphical manner.
Dataset	A collection of related sets of information that is composed of separate elements but can be manipulated as a unit by a computer.
DCE	Distributed Computing Environment.
Derive query operation	An operation to create data values using an algorithm that has as inputs only other data within the database.
Deterministic matching	The comparison of unique identifiers for each record to determine a match or an exact comparison is used between fields.
Digital surveys feature – Integrate with distribution data	The capability to associate each survey response with other information about the survey recipient.
Digital surveys feature – Online data collection and analysis	The capability to collect survey data from online surveys with analytic capabilities.
Digital surveys feature – Option to output printed surveys	The capability to send survey results to another tool via exporting.
Digital surveys feature – Response tracking	The capability to track who has and has not taken a particular survey.
Digital surveys feature – Skip logic	The capability to allow users to skip to a future question or page within the survey even if they do not select a response.
Dimensionality Reduction	The process of reducing the number of random variables under consideration, via obtaining a set of principal variables.
Directly Export	The capability to export from within the solution interface, without requiring the user to manually modify, or manually import from, another tool. This may be achieved through the use of a supplied metadata integration bridge.
Directly Import	The capability to import from within the solution interface, without requiring the user to manually modify, or manually export from, another tool. This may be achieved through the use of a supplied metadata integration bridge.
Distinct query operation	An operation to identify unique records.
DOC	Microsoft Word Document file.
DOCX	Microsoft Word Open XML Format Document file.
Drag and drop	The act of selecting an element in a Graphical User Interface with a pointing device and dragging it from its present location to a different location.
Drill Down	A method of moving from summary information to detailed data.
Drill Through	An action that uses the current context of a report to access information in other cubes or reports.
Drill Up	A method of moving from detailed data to summary information.
Dynamic data content	A content that constantly or regularly changes based on user interactions, timing and other parameters that determine what content is delivered to the user.
eBCI	Electronic Business Computing Infrastructure.
EJB	Enterprise Java Bean.
Entities	The individuals, organizations, and events that are the subjects of an investigative analysis effort.
Entity network analysis and visualization	The capability to visualise relations between entities such as persons or organizations within the data (co-occurrences of named entities).
Exponential	The process of finding the equation of the exponential function that fits best for a set of data.
F test	A statistical test comparing models to identify which model best fits the population from which the data were sampled.
Field reorder query operation	An operation to change the order of table columns.
Filler query operation	An operation to replicate values across an array of cells.
Filter query operation	An operation to set conditions so that only certain data is displayed or retained.
Firewall	A network security system that monitors and controls the incoming and outgoing network traffic based on predetermined security rules.

Term	Definition
First N data sampling functionally	The selection of the first N records from a population to be used as a sample.
Fixed effects analysis	An estimator for the coefficients in the regression model. Impose time independent effects for each entity that are possibly correlated with the regressors.
Function	A named section of a program that performs a specific task.
Functionality	The sum or any aspect of what a software application can do for a user.
Fuzzy logic matching	A technique used in computer-assisted translation as a special case of record linkage. Works with matches that may be less than 100% perfect when finding correspondences between segments of a text and entries in a database of previous translations.
Gamma	A two-parameter family of continuous probability distributions (exponential and chi-squared).
Geometric	The probability distribution of either (a) the number X of Bernoulli trials needed to get one success, supported on the set { 1, 2, 3, …} or (b) the number $Y = -X - 1$ of failures before the first success, supported on the set { 0, 1, 2, 3, ... }.
GIF	Graphic Interchange Format.
GUI	Graphical User Interface, a type of user interface that allows users to interact with electronic devices through graphical icons and visual indicators.
Heteroscedasticity test	A test of the circumstance in which the variability of a variable is unequal across the range of values of a second variable that predicts it.
Hot key	A key or a combination of keys providing quick access to a particular function within a program.
HTML	HyperText Markup Language.
Hypergeometric	A discrete probability distribution that describes the probability of k successes in n draws, without replacement, from a finite population of size N that contains exactly K successes, wherein each draw is either a success or a failure.
Hyperlink	A reference to information that the reader can directly follow either by clicking, tapping, or hovering.
Hypothesis testing for risk - adjust test statistics for multiple comparisons	An adjustment to the significance threshold to account for the expectation that one or more of larger scores would be expected to be observed in the null distribution when performing multiple tests (e.g. the Bonferroni adjustment).
Hypothesis testing for risk - estimate powers	The power of a binary hypothesis test is the probability that the test correctly rejects the null hypothesis (H0) when the alternative hypothesis (H1) is true.
Hypothesis testing for risk - non-parametric	A statistic not based on parameterized families of probability distributions. Include both descriptive and inferential statistics. Make no assumptions about the probability distributions of the variables being assessed.
Hypothesis testing for risk - parametric	A statistic based on the assumption that sample data comes from a population that follows a probability distribution based on a fixed set of parameters.
Hypothesis testing for risk - sample size calculations	The act of choosing the number of observations or replicates to include in a statistical sample.
IBM Infosphere Information Governance Catalog v11.5	A catalog within the IBM Infosphere software used to govern information assets through the development of a governance catalog of categories and terms. This catalog documents information assets and governance policies and rules that implement the high-level strategy and objectives of a governance program.
Identification of distribution through histograms	A histogram is used to profile the underlying frequency distribution of a set of continuous data.
Identification of empty field(s)	An ability to identify fields that contain no data or are 'blank'.
Identification of extreme value(s) (high / low ranges)	An ability to identify the highest and lowest points of a function.
Identification of number of categories	An ability to identify how many different individual categories exist within a function.
Infrastructure	The collection of hardware, software, networks, data centers, facilities and related equipment used to develop, test, operate, monitor, manage and/or support information technology services.

Term	Definition
Inner join	A relational database operation which selects rows from two tables such that the value in one column of the first table also appears in a certain column of the second table.
Input	The signal(s) or data received by the information processing system as a result of the communication between an information processing system, and the outside world.
Input tool	A device that feeds data into a computer, such as a keyboard or mouse.
Interactive visualization	A visualisation technology that enables the exploration of data via the manipulation of chart images, with the color, brightness, size, shape and motion of visual objects representing aspects of the dataset being analyzed.
Intuitive desktop Windows style Graphical User Interface (GUI)	A user friendly user interface that allows users to interact with electronic devices through graphical icons and visual indicators when users understand its behavior and effect without use of reason, experimentation, assistance, or special training.
Inverse transform method	A basic method for pseudo-random number sampling.
IP	Internet Protocol.
IPsec	Internet Protocol Security.
ISO 8859	International Organization for Standardization format.
ITIL	Information Technology Infrastructure Library.
Job	A group of homogeneous, concurrent tasks related by similarity of functions.
JPEG / JPG	Joint Photographic Experts Group file formats.
JSON	JavaScript Object Notation format.
LAN	Local Area Network.
Linear forecasting	The capability to impose a line of best fit to time series historical data.
Linux	A Unix-like computer operating system.
Logistic	A continuous probability distribution. Its cumulative distribution function is the logistic function, which appears in logistic regression and feedforward neural networks.
Log-normal	A continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then $Y = \ln(X)$ has a normal distribution.
Mainframe	A large high-speed computer, supporting numerous workstations or peripherals.
Map layering	The differentiation of different layers from a map such as streams and lakes, terrain, roads, political boundaries, parcels, building footprints, utility lines, and orthophoto imagery.
Mapping with shapefiles and point data	The use of geospatial vector data and point coordinates to associate database records with geographically oriented information.
Merging query operation	An operation to join multiple tables and insert, update, or delete values to agree with a specified source table.
Metadata	The information that defines and describes the characteristics of other data, used to improve both business and technical understanding of data, and the data-related processes. This can be referred to as "data about data."
Methodology	A system of methods used in a particular area of activity.
Microdata	A collection of quantitative and/or qualitative information on the characteristics of units of a population, such as individuals, households, or establishments, collected by a census, survey, or experiment.
Missing value	An instance wherein no data is present for the variable in question.
Mobile device	A mobile computing device. Typical examples include smartphones, tablets and laptops.
Monte Carlo method	A broad class of computational algorithms that rely on repeated random sampling to obtain numerical results. Their essential idea is using randomness to solve problems that might be deterministic in principle.
Multicollinearity test	A test for the phenomenon in which two or more predictor variables in a multiple regression model are highly correlated, meaning that one can be linearly predicted from the others with a substantial degree of accuracy.

Term	Definition
Multinomial	A generalization of the binomial distribution. For n independent trials each of which leads to a success for exactly one of k categories, with each category having a given fixed success probability, the multinomial distribution gives the probability of any particular combination of numbers of successes for the various categories.
Multi-panel capability	The capability to display 2 different files side-by-side on one monitor.
Multiple Imputation	A statistical technique for analyzing incomplete data sets, that is, data sets for which some entries are missing. Application of the technique requires three steps: imputation, analysis and pooling.
Multi-screen capability	The capability for 2 different files to be viewed concurrently on dual monitors.
Named entity recognition	A subtask of information extraction that seeks to locate and classify named entities in text into pre-defined categories such as the names of persons, organizations, locations, expressions of times, quantities, monetary values, percentages, etc.
Naming standard	A set of rules for choosing the character sequence to be used for identifiers which denote variables, types, functions, and other entities in source code and documentation.
Negative binomial	A discrete probability distribution of the number of successes in a sequence of independent and identically distributed Bernoulli trials before a specified (non-random) number of failures (denoted r) occurs.
Network	A group of computer systems and other computing hardware devices that are linked together through communication channels to facilitate communication and resource-sharing among a wide range of users.
Non-linear forecasting	A forecasting technique that generates a line (typically a curve) as if every value of Y was a random variable.
Non-proprietary	A generic, not registered or protected as a trademark or brand name.
Normal distribution	A continuous probability distribution.
Notepad	A basic text editor program for Microsoft Windows which enables computer users to create documents.
Numeric search query operation	An operation to locate specified numeric values within a data source.
Object	A particular instance of a class where the object can be a combination of variables, functions, and data structures.
Outer join	A relational database operation which returns all rows from the left table (table1) and all rows from the right table (table2).
Output	The signal(s) or data sent by the information processing system as a result of the communication between an information processing system, and the outside world.
Panel Data Analysis	A statistical method to analyze two dimensional (typically cross sectional and longitudinal) data.
Parameter	A numerical or other measurable factor forming one of a set that defines a system or sets the conditions of its operation.
Partition query operation	An operation to divide large tables into multiple smaller parts, either horizontally, vertically, or by range.
PDF	Portable Document Format.
Platform	A major piece of software- such as an operating system, an operating environment, or a database - under which various smaller application programs can be designed to run.
PNG	Portable Network Graphics format.
Point and click functionality	The functionality where an action is selected by placing a cursor over its depiction on the display using a pointing device, and is then initiated by clicking.
Poisson	A discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event.
PPT / PPTX	Microsoft PowerPoint Presentation formats.
Principal Component Analysis (PCA) technique	A statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly

Term	Definition
	uncorrelated variables called principal components (or sometimes, principal modes of variation).
Print preview function	A software function that allows a user to see what a document will look like if printed.
Probabilistic matching	A comparison of several field values between two records where each field is assigned a weight that indicates how closely the two field values match. The sum of the individual field's weights indicates the likelihood of a match between two records.
Prompt	A character or message provided by an operating system or program to indicate that it is ready to accept input.
PS	PostScript format.
Query	A request for data or information from a database table or combination of tables.
Random effects analysis	An analysis of hierarchical or panel data when one assumes no fixed effects (it allows for individual effects).
Rare event sampling data sampling functionally	A group of computer simulation methods intended to selectively sample 'special' regions of the dynamic space of systems which are unlikely to visit those special regions through brute-force simulation.
RCNet	The Wide Area Network (WAN) operated by Shared Services Canada (SSC) on behalf of the Canada Revenue Agency (CRA).
Reclassify query operation	An operation to modify data attributes to redefine the type of data (e.g. date, text, numeric) or to reassign observations to new sub-populations based on new criteria.
Record linkage (RL) task	An analytical task for finding records in a data set that refer to the same entity across different data sources (e.g., data files, books, websites, and databases).
Records selection query operation	An operation to select records based on specified conditions.
Rollback	An operation which returns the database to some previous state.
Router	A networking device that forwards data packets between computer networks.
RTF	Rich Text Format.
Scroll bar	A vertical or horizontal bar commonly located on the far right or bottom of a window that allows user to move the window viewing area up, down, left, or right.
Sequence association technique	A special form of argument association that applies to character, array, and sequence structure arguments.
SHP	Shapefile format.
Similarity matching	The recognition of similar individuals based on the information known about them. If two entities (products, services, companies) are similar in some way they share other characteristics as well.
Simple Random data sampling functionally	The taking of a subset of a statistical population in which each member of the subset has an equal probability of being chosen.
Solaris	A Unix operating system.
Solution	The complete set of software components and configurations that are proposed as a response to this RFP. Any items that are not proposed and priced explicitly within the response will not be regarded as part of the solution.
Sorting query operation	An operation to arrange records in an ordered sequence based on specified criteria.
SSC	Shared Services Canada.
Statistical analysis - Canonical analysis	A method for exploring the relationships between two multivariate sets of variables (vectors), all measured on the same individual.
Statistical analysis - Conjoint analysis	A survey-based statistical technique used in market research that helps determine how people value different attributes (feature, function, benefits) that make up an individual product or service.
Statistical analysis - Correspondence analysis	A statistical technique that provides a graphical representation of cross tabulations (which are also known as cross tabs, or contingency tables).
Statistical analysis - F test for ANOVA	A test in one-way analysis of variance that is used to assess whether the expected values of a quantitative variable within several pre-defined groups differ from each other.
Statistical analysis - Factor Analysis	A process in which the values of observed data are expressed as functions of a number of possible causes in order to find which are the most important.

Term	Definition
Statistical analysis - Fisher's test (for categorical variables)	A statistical test used when you have two nominal variables in the analysis of contingency tables and categorical data.
Statistical analysis - MANOVA	An ANOVA with several dependent variables where ANOVA tests for the difference in means between two or more groups, while MANOVA tests for the difference in two or more vectors of means.
Statistical analysis - Multidimensional analysis	A data analysis process that groups data into two categories: data dimensions and measurements.
Statistical analysis - Multiple Regression	A regression model where multiple dependent variables or functions are linear combinations of the parameters.
Statistical analysis - Multivariate Analysis	Any statistical technique used to analyze data that arises from more than one variable. This essentially models reality where each situation, product, or decision involves more than a single variable.
Statistical analysis - Non-parametric analysis	Analysis that does not assume that sample data comes from a population that follows a probability distribution based on a fixed set of parameters.
Statistical analysis - One and Two sample test (Z and t tests)	A t-test is used for testing the mean of one population against a standard or comparing the means of two populations if you do not know the populations' standard deviation and when you have a limited sample (n < 30). A z-test is used if the populations' standard deviation is known.
Statistical analysis - Structural equation analysis	A multivariate statistical analysis technique that is used to analyze structural relationships. This technique is the combination of factor analysis and multiple regression analysis, and it is used to analyze the structural relationship between measured variables and latent constructs.
Statistical analysis - Survival Analysis	A set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest.
Statistical classifier	An attribute or algorithm that identifies the sub-populations to which a given observation belongs.
Stratified data sampling functionally	The process of dividing members of the population into homogeneous subgroups before sampling.
Street-level mapping	The use of street map and address data sources to associate database records with geographic coordinates, civic addresses, municipal zoning, electoral districts, and other layers of geographically oriented information.
String search query operation	An operation to locate specified text values within a data source.
Structured Data	The information with a high degree of organization, such that inclusion in a relational database is seamless and readily searchable by simple, straightforward search engine algorithms or other search operations.
Survey Data Analysis	A process that involves data validation, response partitioning, coding, standard analysis, and ordinal and nominal data analysis.
Systematic data sampling functionally	A type of probability sampling method in which sample members from a larger population are selected according to a random starting point and a fixed periodic interval.
Task	A request for data or information from a database table or combination of tables.
Technical Authority	The representative of the CRA for whom the Work is being carried out under the Contract and is responsible for all matters concerning the technical content of the Work under the Contract.
TEX	LaTex typesetting file format.
TIF / TIFF	Tagged Image File Format.
Time series analysis	Analysis of a consistent population over multiple time intervals in order to observe meaningful changes over time.
Timeline Analysis	A tool used in a variety of investigation types and is often used to answer questions about when a computer is used or what events occurred before or after a given event.
Transpose query operation	An operation to restructure a table by converting data values into data categories or vice versa.

Term	Definition
Treatment Effect Models	The causal effect of a binary (0–1) variable on an outcome variable of scientific or policy interest.
Trend	A pattern of gradual change in a condition, output, or process, or an average or general tendency of a series of data points to move in a certain direction over time.
TXT	TextFile.
Undo function	A function performed to reverse the action of an earlier action.
Uniform	A distribution that has constant probability.
Union	A combination of the result sets of 2 or more SELECT statements.
Unstructured data	The information that either does not have a pre-defined data model or is not organized in a pre-defined manner.
URL	Universal Resource Locator.
UTF-8	Unicode Transformation Format.
Vendor	The successful bidder after the selection process is complete.
Versioning	The process of assigning either unique version names or unique version numbers to unique states.
VLAN	Virtual Local Area Network.
VPN	Virtual Private Network.
WAN	Wide Area Network.
Web crawling capability	The ability to browse the World Wide Web in a methodical, automated manner.
Weibull	A distribution type used in reliability and life data analysis due to its versatility and can be used to model a variety of life behaviors.
Weighted data sampling functionally	A methodology used to adjust the results of a study to bring them more in line with what is known about a population.
Wizard	A help feature of a software package that automates complex tasks by asking the user a series of easy-to-answer questions.
WWW	World Wide Web.
XLS / XLSX	Microsoft Excel Spreadsheet formats.
XML	eXtensible Markup Language format.