Data Practitioner Performance Based Rubrics & Glossary

DUTY ONE: INITIATES THE DATA

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Translates business problems into analytic process* **(Integrates Task A)**	• Communicates understanding of the problem.	• Articulates the problem after consulting with primary stakeholders.*	• Proposes a type of analytic process needed to solve the problem.	• Intuits the problem and the process needed to solve it.
Integrates stakeholders' interests into proposed analytic process **(Integrates Tasks B, C)**	• Identifies stakeholder needs by engaging them in active conversations about the proposed analytic process.	• Builds consensus among stakeholders around both the problem that has been identified and the proposed analytic process.	• Obtains stakeholder approval for revised analytic process.	• Enlists stakeholders as partners in the proposed solution.
Determines what data is needed to execute the analytic process **(Integrates Tasks D, E, F)**	• Identifies data that fits within the problem.	• Revises the analytic process after determining what data should be excluded and included.	• Aligns data to the revised analytic process.	• Anticipates how data will be sourced.
Develops statement of work* **(Integrates Tasks G,H, I)**	• Drafts a statement of work that addresses stakeholder interests/needs, includes data/analytic process and resources required to solve the problem (e.g. timeline, budget, personnel, tools, data….).	• Integrates primary stakeholder feedback into statement of work.	• Finalizes a statement of work needed to solve the problem that addresses stakeholder interests/needs, includes data/analysis processes and all resources needed to accomplish the task.	• Enhances the statement of work by drawing on long term relationships with stakeholders and organizational knowledge.

DUTY TWO: SOURCES THE DATA

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Facilitates data collection **(Integrates Tasks A, B, C, D, E)**	• Collects data, as prescribed by statement of work, in its original form. • Demonstrates familiarity with industry quality control standards.	• Recognizes problems that might occur as a result of data collected in the wrong formats. • Revises data collection to account for identified areas in which problems might occur.	• Collects all data required by the statement of work in a format that lends itself to transformation. • Exercises quality control checks of the data (e.g. conducts randomized selection* of data) to ensure that the data lends itself to transformation.	• Repurposes previously used processes to apply to new data problems. • Sets validation thresholds* for conducting quality control checks.
Prepares data for transformation **(Integrates Tasks F, G)**	• Cleans data to identify outliers* and errors.	• Streamlines analytic process to expedite cleaning of data (e.g. reuse codes previously developed, automates import/export process (ETL)*…).	• Validates data (e.g. data sources are confirmed; data is cleaned, accurate, usable and tested).	• Shares best practices (codes, formulas, ETL processes…) across organization and industry.

DUTY TWO: SOURCES THE DATA

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Produces data dictionary* **(Integrates Tasks H, 3L)**	• Creates document that includes all elements (e.g. fields, terms, processes built, workflows created, sources….) to be included in the data dictionary.	• Creates document that maps data elements within/between data systems (from source to their analytic system).	• Creates a document that describes the origins of all of the data and how it is used to solve the problem (e.g. up-to-date, accurate and detailed).	Creates a document that integrates information from historical data dictionaries to create an index of issues, problems and solutions.
Applies business, ethical and legal standards **(Integrates Tasks I)**	• Complies with business, ethical and legal standards in all aspects of work with data.	• Maintains currency with changes in business, ethical and legal standards.	• Serves as a point of reference within the organization for current business, ethical and legal standards.	• Provides expert advice on business, ethical and legal standards in all aspects of work with data (including historical context of data compliance within the organization).

DUTY THREE: TRANSFORMS THE DATA

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Prepares Data for Organization **(Integrates Tasks A, B, C, D, E, F, G)**	• Customizes validated data to align with the problem identified in the statement of work.	• Troubleshoots customized data to ensure its alignment with the problem identified in the statement of work.	• Produces a robust (e.g. standardized, includes new connections among data…) data set aligned with the problem identified in the statement of work.	• Coaches others in the preparation of robust data sets for organization.
Aligns data to structure of analysis tools **(Integrates Tasks H, I, J, K, M)**	• Applies the required structure necessary to use analysis tools.	• Troubleshoots inefficiencies of the application of the data structure (e.g. repetition of variables, errors, unanticipated incomplete data…)	• Applies revised structures to produce a comprehensive data set ready for storage.	• Shares best practices of data transformation* (codes, formulas, data governance procedures*….) across organization and industry.

DUTY FOUR: ANALYZES THE DATA

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Applies research method **(Integrates Tasks A,B,C,D)**	• Assigns the data to fit within the parameters of the concepts/ theories / hypotheses/ research questions. • Executes the analysis presented in the statement of work.	• Experiments with the analysis to identify potential best fits of the data within the parameters of the concepts/ theories / hypotheses/ research questions. • Executes the analysis presented in the statement of work.	• Tailors the data to fit within the parameters of the concepts/ theories / hypotheses/ research questions. • Executes the analysis presented in the statement of work.	• Coaches others on applying research methods.
Performs data mining* **(Integrates Tasks E,F)**	• Discovers patterns and trends in large data sets.	• Probes identified patterns/trends for new, significant trends/patterns in the data.	• Engages in an iterative process of asking new questions and identifying additional new, significant trends/patterns in the data.	• Creates code(s)/tools that facilitate data mining.

DUTY FOUR: ANALYZES THE DATA

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Interprets the results **(Integrates Tasks G,H,I,J,K,L)**	• Generates initial results by telling the data story.	• Posits initial solutions.	• Confirms results and their alignment to statement of work.	• Applies industry context to the interpretation.
Creates data visualizations (dashboards,* reports, maps, charts, graphs, videos…..) **(Integrates Tasks M)**	• Produces a visualization that communicates results.	• Produces visualization(s) that take into account the interests/concerns of stakeholder(s).	• Tailors multiple versions of visualizations to communicate the data story to stakeholder(s).	• Shares best practices (e.g. by creating templates for visualizations for various stakeholders; mentoring others….).

EDC OCEANS *of* DATA INSTITUTE

DUTY FIVE: CLOSES OUT THE PROJECT

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Communicates the results of the analysis **(Integrates Tasks A,B,C,E,F,G)**	• Presents initial deliverable(s) (e.g. report, presentation…) for feedback.	• Presents deliverable(s) that incorporate feedback from representatives of stakeholder groups.	• Defends analysis and conclusions to the stakeholders.	• Demonstrates predictive efficacy of analysis.
Documents the analysis **(Integrates Tasks H,I,J,K)**	• Reviews draft deliverable to verify its alignment to the statement of work and completeness.	• Substantiates validity of findings (evidence, multiple visualizations, well-written, detailed footnotes/ appendices linking to other data….).	• Produces final deliverable.	• Communicates future processes, improvements and opportunities.

DUTY SIX: ENGAGES IN PROFESSIONAL DEVELOPMENT

Performance Area	Level One Basic Competence	Level Two Approaching Proficiency	Level Three Proficiency	Level Four Above Proficiency
Keeps current on trends in big data analysis, best practices and emerging technologies **(Integrates Tasks A,B,C,F,H)**	This performance area does not warrant a division into levels of mastery. This does not diminish its importance. Below is a sample checklist of performance statements that describe actions data practitioners can take to apprise themselves of best practices in the field: ☐ **Maintains professional qualifications** ☐ **Stays current on emerging technologies, methods and tools** ☐ **Seeks mentors** ☐ **Attends relevant conferences and seminars** ☐ **Participates in professional organizations**			
Contributes to the development of the field **(Integrates Tasks D,E,G,I)**	This performance area does not warrant a division into levels of mastery. This does not diminish its importance. Below is a sample checklist of performance statements that describe actions data practitioners can take to apprise themselves of best practices in the field: ☐ **Shares best practices with colleagues** ☐ **Mentors others** ☐ **Suggests future projects** ☐ **Publishes articles in peer reviewed journals** ☐ **Makes presentations at relevant conferences** ☐ **Develops training materials** ☐ **Delivers workshops highlighting data analysis techniques** ☐ **Volunteers to be guest speaker at high school/ college career development events**			

EDC OCEANS *of* DATA INSTITUTE

GLOSSARY

Term	Definition
Analytic process	The application of research methodology to find the result of a given problem.
Dashboard	A data visualization tool that displays the current status of metrics and key performance indicators (KPIs) for an enterprise.
Data dictionary	A set of information describing the content, format and structure of a database, and the relationship between its elements and the elements contained in other databases or the application of those elements in an analytic process. Used to control access to, and manipulation of, the database and its data, and to provide a data model to an enterprise.
Data governance procedures	The applied policies of the availability, usability, integrity and security of the data employed in an enterprise.
Data mining	The process of analyzing, categorizing, and summarizing identified relationships from different perspectives in order to find correlations or patterns among large relational databases.
Data transformation	The modification of every point in a data set by a mathematical function, used to change data to the appropriate form for a statistical test or research method.
ETL	Acronym for Extract, Transform and Load.
Outlier	An observation point that is distant from other observations.
Randomized selection	A selection of data that is done purely by chance with no predictability.
Stakeholder	A person, organization or entity with an interest or concern in the project.
Statement of work	A document that defines project specific activities, deliverables and timelines for a project.
Validation thresholds	A minimum or maximum value established for an attribute, characteristic, or parameter that serves as a benchmark for comparison or guidance. Used to quickly spot-check data and alert an enterprise to minimize risk.

EDC OCEANS of DATA INSTITUTE