

D.A.T.A.

Data Analytics Technician Advancement

	2015 Compression Planning Skills/Abilities													
	1a. Business Understanding	1b. Data Understanding	1c. Data Preparation	1d. Modeling	1e. Evaluation	1f. Deployment	2. Basic profiling skills and ability to gather data	3. Basic Data preparation and blending	4. Prepare data with appropriate process/best practice for study being performed	5. Think critically and logically to solve problems.	6. Awareness of how to use data to solve problems within study design.	7. Understand data structure and what question is being asked.	8. Summarized data (stats and review).	9. Ability to use tools that are constant changing including scripting and notebooks.
1A. Translates business problems into analytic needs.	X											X		
1B. Interviews stakeholders.	X													
1C. Refines stakeholder needs.	X											X		
1D. Identifies appropriate data.		X				X					X			
1E. Identifies whether data exists or not.		X				X					X			
1F. Performs gap analysis of the data.		X									X			
1G. Determines resource needs (e.g., SMEs, tools, timelines).	X													
1H. Determines feasibility of analysis to be done.												X		
1I. Creates statement of work.												X		
2A. Determines data source(s).		X	X			X								
2B. Determines target structure.		X												
2C. Collects data.		X				X		X						
2D. Exercises quality control (e.g., randomizes selection).		X				X		X						
2E. Extracts data (e.g., writes SQL, API code).		X				X	X	X						
2F. Cleans data (e.g., identifies outliers/errors).		X				X	X	X						
2G. Tests data.		X				X		X						
2H. Creates data dictionary.		X				X		X						
2I. Complies with business, ethical and legal standards.		X				X		X						
3A. Merges data.		X				X	X	X						
3B. Splits data.		X				X	X	X						
3C. Derives new variables.		X				X		X						
3D. Creates new data.		X				X		X						
3E. Augments data.		X				X		X						
3F. Applies meta- data.		X				X		X						
3G. Purges data.		X				X	X	X						
3H. Changes data structure.		X				X	X	X						
3I. Changes data types.		X				X	X	X						
3J. Normalizes data.		X				X	X	X						
3K. Interpolates data.		X				X		X						
3L. Finalizes data dictionary.		X				X	X	X						
3M. Stores data for analytics.		X				X	X	X						
4A. Determines what analysis to run.			X								X	X		
4B. Applies the research method and tools.			X								X			
4C. Identifies dependent and independent variables.			X								X	X	X	
4D. Defines appropriate algorithms.			X								X	X		
4E. Performs data mining.		X	X								X			
4F. Separates any anomalies.		X									X			
4G. Interprets the results.			X								X	X		
4H. Runs additional tests as needed.			X											
4I. Performs reasonableness tests of results.			X											
4J. Compares results to previous findings.			X											
4K. Confirms results.			X											
4L. Conducts causality testing.			X											
4M. Creates data visualizations (e.g., dashboards, reports, charts, graphs, videos, animation).						X								
5A. Selects documentation media.			X			X								
5B. Describes problem, method and analysis.			X			X						X		
5C. Articulates conclusions.			X			X						X		
5D. Compiles reports.			X			X						X		
5E. Presents information to stakeholders.			X			X								
5F. Integrates feedback from stakeholders.			X			X								
5G. Defends analysis as needed.			X			X								
5H. Reworks analysis as needed.			X			X								
5I. Prepares final report.			X			X								
5J. Archives work products.			X			X								
5K. Communicates future processes, improvements and opportunities.			X			X								
6A. Maintains professional qualifications.														
6B. Stays current on emerging technologies, methods and tools.														
6C. Seeks out mentors.														
6D. Shares best practices.														
6E. Contributes new knowledge to the field.														
6F. Attends relevant conferences and seminars.														
6G. Mentors others.														
6H. Participates in professional organizations.														
6I. Suggests future projects.														

EDC Data Practitioner Tasks



This material is based upon work supported by the National Science Foundation under Grant No. 1700454.