SKILLS AND KNOWLEDGE

Skills in:

Analytical Thinking
Applying Statistical Methods
Basic GIS
Basic Security
Chart/ Visualizations
Coding Languages
Communication
Critical Thinking
Data Entry (to server)
Data Manipulation
Data Structure of Organization
Design
Documentation
ETL
Multi-tasking
Operating Systems
Pattern identification/ analysis
Presenting
Prioritizing
Problem Solving
Project Management
Research Methods
Statistics (Basic)
Time Management
Writing

Knowledge of:

Business Acumen
Communication
Computer Modeling
Critical Thinking
Customer Relations
Databases
Data Discovery
Data Modeling
Data Organization
Data Quality
Data Stewardship
Data Structures
Design
Domain Knowledge
Ethics
IT
Project Management
RDBMS (SQL Server, Oracle, No SQL)
Research Methods
Software
Statistics

BEHAVIORS

A successful Data Practitioner is…

Able to manage time
Able to multi-task
Able to problem solve
Able to work independently
Collaborative
Competent
Courageous
Creative
Curious
Diligent
Effective serving customers
Effective executing work
Ethical
Focused
Inquisitive
Intellectually humble
Open to/ provides feedback
Organized
Patient
Persistent
Self-Confident

EQUIPMENT/TOOLS/SUPPLIES

Data/ Database Tools (e.g., Excel, Access, SQL Server, Oracle)
Data Mining Tools (e.g., Microsoft SQLServer Data Tools including SSRS, SSAS, SSIS)
Data Visualization Software (e.g., Tableau Software, QlikTech Qlikview, TIBCO Spotfire, Microsoft Power BI)
GIS (e.g., ArcGIS for Desktop Basic)
Mobile Devices
Online communities/ discussion groups/ forums
Open Source Tools
PowerPoint/ Prezi
Project Management Software
Python
Reporting Tools (e.g., SAP Crystal Reports, MicroStrategy, Inc.)
Statistics Packages (e.g., SPSS, R, SAS, SASJMP)
Tablet
Word Processing

FUTURE TRENDS

- Growing concern about the role of individual privacy in a world in which data is heavily collected and shared
- Growing expectation that people will use and/ or create data in their work
- Growing need for data literacy by all
- Increasing automation of the analytic process
- Increasing capacity of data to solve specific and complex problems (e.g. Genomics - 23andme)
- Increasing number of individuals with limited data analysis skills utilizing machine learning, applications, visualization tools and platforms as a means to analyze data
- Increasing speed and volume of data sources (IoT) outpaces application of the findings

INDUSTRY CONCERNS

- Costs needed to staff data practitioners
- Need to educate consumers of data to ensure they know its limitations
- Need to establish standardization of data within industries
- Ongoing necessity to question data for reliability: data quality, consistency, completeness, bias, sourcing, transparency, data security
- Possibility of AI eliminating human jobs in Data Analysis
- The need for clarity regarding marketplace and organizational strategic imperatives which drive priorities
- Too many academic programs teaching software that employers do not use

PANEL

Daniel Boisvert – Biotechnology
Biogen
Cambridge, Massachusetts

Mike DeFabbo – Education/ Nonprofit
OneGoal
Chicago, Illinois

Rebecca Hailey – Marine Science
Virginia Institute of Marine Science
Gloucester, Virginia

Paul Hansford – Business Consulting
SimpleSoft Solutions, Inc.
Dayton, Ohio

Tony Joy – Financial Services
Global Audit Management & Consulting
Urbana, Ohio

Ryan Kapaun – Law Enforcement
Eden Prairie Police Department
Eden Prairie, Minnesota

Sean Larson –
Medical Device Manufacturing
Medtronic
Minneapolis, Minnesota

Andy Ramlatchan – Health Care
Patient Advocate Foundation
Virginia Beach, Virginia

Greg Reisz – Agriculture
E-4 Crop Intelligence
Woodbine, Iowa

Joel Wright – Public Policy
Wright Consulting Services
Strawberry, Arizona

PROFILE FACILITATORS

Joseph Ippolito
Education Development Center
Cleveland, Ohio

Joyce Malyn Smith, Ed. D.
Education Development Center
Boston, Massachusetts

Profile of the Data Practitioner

Initially Developed – April 15-16, 2016

oceansofdata.org

oceansofdata@edc.org

EDC OCEANS of DATA INSTITUTE

NSF

EDC Learning transforms lives.

Learning Occupation: The Data Practitioner, in service of an organization and/or stakeholders, supports the data life cycle by collecting, transforming, and analyzing data, and communicating results in order to inform and guide decision-making.

DUTIES	TASKS										
1. INITIATES THE PROJECT	1A. Translates business problems into analytic needs.	1B. Interviews stakeholders.	1C. Refines stakeholder needs.	1D. Identifies appropriate data.	1E. Identifies whether data exists or not.	1F. Performs gap analysis of the data.	1G. Determines resource needs (e.g., SMEs, tools, timelines).	1H. Determines feasibility of analysis to be done.	1I. Creates statement of work.		
2. SOURCES THE DATA	2A. Determines data source(s).	2B. Determines target structure.	2C. Collects data.	2D. Exercises quality control (e.g., randomizes selection).	2E. Extracts data (e.g., writes SQL, API code).	2F. Cleans data (e.g., identifies outliers/errors).	2G. Tests data.	2H. Creates data dictionary.	2I. Complies with business, ethical and legal standards.		
3. TRANSFORMS THE DATA	3A. Merges data.	3B. Splits data.	3C. Derives new variables.	3D. Creates new data.	3E. Augments data.	3F. Applies meta-data.	3G. Purges data.	3H. Changes data structure.	3I. Changes data types.	3J. Normalizes data.	3K. Interpolates data.
	3L. Finalizes data dictionary.	3M. Stores data for analytics.									
4. ANALYZES THE DATA	4A. Determines what analysis to run.	4B. Applies the research method and tools.	4C. Identifies dependent and independent variables.	4D. Defines appropriate algorithms.	4E. Performs data mining.	4F. Separates any anomalies.	4G. Interprets the results.	4H. Runs additional tests as needed.	4I. Performs reasonableness tests of results.	4J. Compares results to previous findings.	4K. Confirms results.
	4L. Conducts causality testing.	4M. Creates data visualizations (e.g., dashboards, reports, charts, graphs, videos, animation).									
5. CLOSES OUT THE PROJECT	5A. Selects documentation media.	5B. Describes problem, method and analysis.	5C. Articulates conclusions.	5D. Compiles reports.	5E. Presents information to stakeholders.	5F. Integrates feedback from stakeholders.	5G. Defends analysis as needed.	5H. Reworks analysis as needed.	5I. Prepares final report.	5J. Archives work products.	5K. Communicates future processes, improvements and opportunities.
6. ENGAGES IN PROFESSIONAL DEVELOPMENT	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.		