GUIDE *for* Data Worker Internships

EDC Education Development Center

NSF

COLUMBUS STATE COMMUNITY COLLEGE

ACKNOWLEDGEMENTS

The development of this guide benefited from the insight and contributions of local industry representatives. These individuals participated in meetings, phone conferences, and/or work sessions aimed at identifying local industry needs and expectations for middle-skilled data workers.

Alicia Croft, Columbus State Community College
Dannee Cunningham, Nationwide Insurance
Mark Gorman, The Gorman Group
David Judson, Cardinal Health
Jeff Kanel, Centric Consulting
Josh Mills, Nationwide Insurance
Krish Rajanam, Nationwide Insurance
Andy Ramlatchan, NASA Langley Research Center
Muhammed Salley Sinare, Private Investigator
Leslie Smith, Nationwide Insurance
Chris Wilson, Colorado State Patrol
Joel Wright, Wright Consulting
Wayne Zhao, Nationwide Insurance

Columbus State Community College staff provided insights into the college program and institution-wide internship policies and procedures:

Gloria Rogiers, Chair, Department of Computer Science
Andrew Kerr, Instructor, Department of Mathematics
Stephanie Schuler, Project Manager, Grants Office
Britney Williams-Ward, Instructor, Information Systems Technology

Education Development Center (EDC) senior staff associated with EDC's Oceans of Data Institute (oceansofdata.org) facilitated discussions with Columbus, Ohio, employers who contributed to the guide's content. In addition, the guide integrates and adapts occupational profiles, performance-based rubrics, and other relevant materials developed from EDC's Creating Pathways for Big Data Careers project (NSF-ATE # 1501927):

Joseph Ippolito – Senior Project Director, Education Development Center
Joyce Malyn-Smith – Distinguished Scholar, Education Development Center

The National Science Foundation provided funds that supported the work leading to the development of this guide through the Data Analytics Technician Advancement (DATA) project (NSF-ATE # 1700454).

CONTENTS

INTRODUCTION

In our technology-driven world, our access to and use of data is transforming the workplace. The use of data as decision-drivers has expanded into every scientific, industry, and business enterprise. Real-time data are generated at dizzying velocities and the variety of data being generated is expanding as new sensors and technologies are deployed. The amount and the speed by which data is being generated is creating an immediate, pressing demand for skilled data workers[1] in all industry sectors. What we have seen so far is just the tip of the iceberg. Analyzing data, spotting patterns, and extracting useful information have become gateway skills to full participation in the workforce and civic engagement of the 21st century. Individuals skilled in the variety of tasks and duties required to acquire, aggregate, clean, organize, and analyze these massive data sets are now in high demand.

Real time data are generated at dizzying velocity and the variety of data that are generated are expanding as new sensors and technologies are deployed. This is creating an immediate, high demand for skilled data workers in all industry sectors. What we have seen so far is just the tip of the iceberg.

To speed and ease the transition from education to employment in data fields, many community colleges are establishing data internships. Internships provide students with immediate opportunities to apply their data skills and knowledge to the tasks and problems challenging data workers in today's workplaces. Internships benefit both students and employers. They provide students with opportunities to work on data teams, to learn to solve real-world data problems found in local industries, and to develop new data skills working in industry sectors that interest them. Interns learn about the world of work while also developing employability skills and learning, first-hand, how to fit in and succeed in various workplace cultures. Employers benefit from interns as well. Without making long-term commitments, employers are able to assess interns as potential employees, evaluating their technical skills, employability, and soft skills such as their ability to solve problems, learn on their feet, and adapt to new situations. Additionally, the ongoing communication between employers and faculty provides employers with opportunities to help shape curriculum to ensure that graduates are able to meet the skill demands of their workplaces. Colleges benefit from the close connections with business partners, deepening their understanding of the specific skills needed by local employers. This close relationship provides colleges with the information needed to keep their curriculum current with industry's expectations, thereby adding value to their data programs.

1 In this guide, *data worker* covers a variety of job titles and is used interchangeably with *data practitioner*.

While this guide is primarily aimed at employers of data workers, the resources that are included can be used to facilitate greater dialogue between those employers and college faculty, resulting in richer learning opportunities for students.

The guide is organized into five sections and an appendix. Each section provides guidance and resources to help design and implement an internship in data work. Development of the resources has been led by the Oceans of Data Institute (ODI) at Education Development Center (EDC) over the course of multiple projects funded by the National Science Foundation. To learn more about EDC's Oceans of Data Institute visit oceansofdata.org.

For questions regarding ODI's body of work and to connect to other colleges and employers involved in training the next generation of data workers, contact the following EDC staff:

Joe Ippolito jippolito@edc.org

Joyce Malyn-Smith jmsmith@edc.org

SECTION 1 – Case Study: Nationwide's Internships for Future Data Workers

A Rationale for Providing Internships

Nationwide, a Fortune 100 company based in Columbus, Ohio, is one of the largest and strongest diversified insurance and financial services organizations in the United States. The company has partnered with Columbus State Community College for five years, providing students enrolled in data science and data analysis courses with the opportunity to participate in summer internships provided by the company.

Like most companies, Nationwide has a continuous need to recruit new talent. According to LinkedIn, between 2012 and 2017, the number of job openings for positions requiring data analytical skills increased by 650%. Beginning in 2015, Nationwide turned to internships as an approach to identify skilled data workers to fill its workforce needs. Since then, internships have proven to be the company's major entry level pipeline for talent.

At the same time, Nationwide has discovered that interns add significant value to organizational success. Interns are assigned to existing project work where their contributions have had immediate impact. In some cases, interns bring new ideas and methods to the table whose applications reach further than the boundaries of their project. For example, one of Nationwide's interns piloted a technique that provided a more transparent means of interpreting computational business models. The technique, dubbed "Locally Interpretable Model Agnostic Explanations," was favorably received by Nationwide's business partners and has now been adopted within other modeling efforts.

What Does an Internship Look Like at Nationwide?

Nationwide offers two kinds of internships within the Predictive Analytics Team, based within the Enterprise Analytics Office. One internship focuses on predictive analytics, which requires a deep technical background in quantitative fields, such as statistics, engineering, economics, and related areas. The other deals with the company's model scoring infrastructure and requires interns to have a background in computer science. Both internships are offered for 12 weeks during the summer and are viewed by the department's program director as "12-week mutual interview processes." Nationwide seeks to provide work experiences that engage interns and limit their time spent on routine administrative tasks. Every effort is made to match interns to projects that meet their interests.

All interns begin their experience by participating in a two-week onboarding experience. During this time, interns receive a general orientation to Nationwide's mission, history, and corporate culture. Each intern is assigned an intern manager who becomes their primary point of contact during their internship. The intern manager provides a technical overview of projects, reviews company and project expectations, outlines specific intern responsibilities and the means of evaluating the performance of those responsibilities, and explains how interns and intern managers work together.

Sample Projects Involving Interns

1) One recent project that engaged an intern involved assessing the performance of a new third-party vendor. Nationwide engaged the vendor to evaluate the state of the company's brand. The vendor was charged with administering surveys aimed at quantifying Nationwide's brand health by developing metrics for things such as brand awareness, brand consideration, and social media buzz. The project presented two difficulties. First, some of Nationwide's business partners expressed a natural skepticism because the vendor was new and unknown. Second, the data collected were not presented initially in a way that was easily consumable. Data from Nationwide and competitors were captured monthly. The intern involved in the project was tasked with developing an automated reporting system that would update new data as they were collected and would then update the report on a regular schedule. At the same time, the intern was expected to work with subject matter experts to help tell a compelling story that better communicated the impact of the collected data. The intern was able to prototype a new reporting system that added three-month forecasts, which proved to be an effective tool for interpreting market data. The intern also assuaged the concerns of Nationwide's business partners by providing them weekly updates, thereby allowing them to see how the vendor's reports were evolving.

2) A second project involved two interns who worked together to create the System Issue Tracker for the Customer Service & Billing (CSB) Business Service Area (BSA). The CSB team was building a spreadsheet to compare system downtime for service representatives to tickets opened for the purpose of tracking issues and determining the primary cause of the system being down. The report was taking weeks to create and was often too late to detect real-time issues. The interns used Tableau to build a scorecard that automated the merging of IEX and ServiceNow tickets, allowing near real-time management of system downtime for commercial line and personal line service representatives, saving the CSB team 80 hours a month in effort.

3) A third project had an intern work on deploying Liquibase into Nationwide's databases in Commercial Lines Transformation. Liquibase allows the versioning of database changes, creating a clear audit trail, consistent across multiple environments, and an easy rollback capability. Working closely with a data analyst lead, the intern's main responsibilities involved creating the Python scripting that enabled the synchronization across environments. The intern also tested the various components. The implementation was a success, and the team has already recognized the benefits.

Nationwide attempts to put its interns through as much of the analytic process as possible, in part, to determine the degree to which they can fulfill the duties of mid-level data workers and succeed as full-time employees. Internships require interns to call upon their skills and knowledge to transform data, to analyze data, and to close out their projects. When possible, interns are included in meetings with stakeholders and in activities involved in initiating projects. At the conclusion of their internships, Nationwide expects its interns to create a compelling presentation that includes possible future directions for the work, which may involve a continuation of the intern's work by that intern's manager or another member of their team.

Evaluating Interns

Nationwide evaluates its interns by assessing their technical skills, business acumen, and communication skills. To rate intern performance, intern managers grade a list of these competencies using a scale of 1–5. They also provide general commentary on the job done by the intern. Evaluations are done as a collaboration between the intern manager and his or her supervisor, which additionally provides a means of giving the intern manager experience conducting performance reviews.

SECTION 2 – Making the Decision to Sponsor Internships

Companies that choose to sponsor internships for college students commit to making a significant investment of time and resources. They make that investment because they conclude that doing so provides real benefits to their business success. As companies consider whether or not to initiate an internship program, it may help them to learn what has prompted others to do so. Following is a sampling of relevant resources that are readily available online:

1) **"14 Benefits of Starting an Internship Program for Your Company"** (Chegg Internships) – Identifies factors that demonstrate how internships address a company's self-interests: https://www.internships.com/employer/resources/setup/benefits

2) **"The Benefits of Hiring an Intern: What Interns Bring to the Table"** (Penny Loretto, The Balance Careers) – Provides a rationale for companies to sponsor internships: https://www.thebalancecareers.com/why-hire-interns-1986579

3) **"Developing a Successful Internship Program"** (Penny Loretto, The Balance Careers) – Offers a general overview of what employers can do to create an environment supporting a successful internship program: https://www.thebalancecareers.com/developing-a-successful-internship-program-1986608

4) **"Legal Guidelines for Hiring Paid or Unpaid Internships"** (WayUp) – Provides answers to frequently asked questions regarding some of the legal considerations involved in the decision to sponsor interns: https://www.wayup.com/employers/hiring-intern-law/

Questions such as the following provide a good starting point for company leaders to discuss whether or not to offer internships:

1) In what ways can an internship program benefit our company?

2) What value can interns add to our company's operations?

3) What are our immediate data-related workforce needs and how will an internship program help us meet those needs?

4) How can internships help our company address our projected future data-related workforce needs?

5) How can an internship program help us build stronger relationships with local colleges and the rest of the community?

6) What are some of the potential challenges we should consider in preparing for an internship program? (Examples include the amount of time needed to supervise an intern, which takes away from the internship manager's assigned tasks, and potential liabilities.)

SECTION 3 – Identifying Intern Responsibilities

Ideally, internships provide important work-oriented learning and job experience for interns while at the same time contributing to the success of their business sponsors. To ensure this happens, it is essential that sponsors clearly define what they expect interns to do and to communicate that to both interns and the college faculty charged with instructing them. Company leaders can begin defining the responsibilities of interns by considering questions such as the following:

1) Which of our current projects lend themselves to accepting interns?
2) What value would interns add to those projects?
3) What would an intern be responsible for doing on those projects?
4) What skills and knowledge would an intern need in order to perform those responsibilities effectively?
5) Are there job tasks that an intern can perform while being supervised? If so, what are they?
6) Are there job tasks that an intern can perform unsupervised? If so, what are they?

Answering questions such as these can be facilitated by using a tool that provides employers and educators with a shared language for describing the responsibilities of data workers. The matrix of work activities that comprise the Profile of a Data Practitioner (see pages 22 through 25) has been used successfully across the country

for precisely this purpose. The profile was developed by a panel of expert data workers who were charged with identifying the "middle" data and big data skills needed in industry. These middle skills are meant to reflect the work responsibilities a member of a data team might be expected to perform under the supervision of the data team leader or data scientist. The panel of experts represented 10 different industry sectors: biotechnology, education, marine science, business consulting, financial services, law enforcement, medical device manufacturing, health care, agriculture, and public policy. The profile was validated through a national survey of nearly 100 expert data workers.

Data practitioner may be a new title to employers. However, a worker defined by the data profile as one who "in service of the organization and/or stakeholders supports the data life cycle by collecting, transforming, and analyzing data, and communicating results in order to inform decision-making" is likely very familiar. The data practitioner may be more commonly referred to as a *data analyst, business intelligence analyst, financial analyst,* or *systems analyst*. Regardless of the title, the chart of work activities presented by the profile should provide a reasonable approximation of what any employer's middle-skill data workers are expected to do. As a result, employers can begin the process of defining an internship by reviewing the chart and deciding which of the listed work activities a prospective intern might be expected to do.

Profile Components

The profile essentially captures and organizes the universe of possible work responsibilities of a middle-skilled data worker. The screenshots that follow identify the profile's key components. A full-size version of the profile matrix can be found on page 23.

The ***occupational definition*** serves as the foundation of the profile.

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.		

The ***duties*** represent major work responsibilities. They are listed in the far-left column.

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.		

Copyright ©2016 by Education Development Center, Inc.

All rights reserved.

These six duties are subdivided into 64 related minor work responsibilities or ***tasks***. It is unlikely that any one individual would be responsible for performing ALL of these activities. For an intern, an even more limited of number of tasks are likely to be realistic or appropriate for the assignment.

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.		

In addition to the chart, the back page of the profile lists the following:

Skills: The abilities, or ways of doing things, necessary to perform the data practitioner tasks effectively

Knowledge: The content areas and the information an effective data practitioner needs to know

Behaviors: The character traits, dispositions, or habits that are embodied by an effective data practitioner

Employers will recognize that few, if any, of the data practitioners in their businesses are expected to perform the full range of duties and tasks listed in the profile. Furthermore, employers will know that a number of the activities demand an advanced skill level not likely to be found in an intern. Similarly, faculty will note that graduates of their programs will not have had an opportunity to practice every listed task. As both parties consider the design of internships, it may be helpful for employers and college representatives to review the chart together and identify together those areas where the overlap exists between course/program content and tasks appropriate for interns.

For added guidance in selecting tasks suitable for a data worker internship, employers and college representatives can also consult the "heat map" version of the profile chart. The heat map on page 24 is a document displaying the feedback from 10 employers of data workers who were asked to look over the matrix and to identify which tasks they considered "suitable for an intern working under supervision." Each task has been shaded to designate the frequency by which they were identified by employers as appropriate for an intern. The darker the shade of red, the more frequently the task was identified by survey respondents. Employers and the college may want to begin the design of an internship by incorporating several, closely related tasks that have been most frequently identified.

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

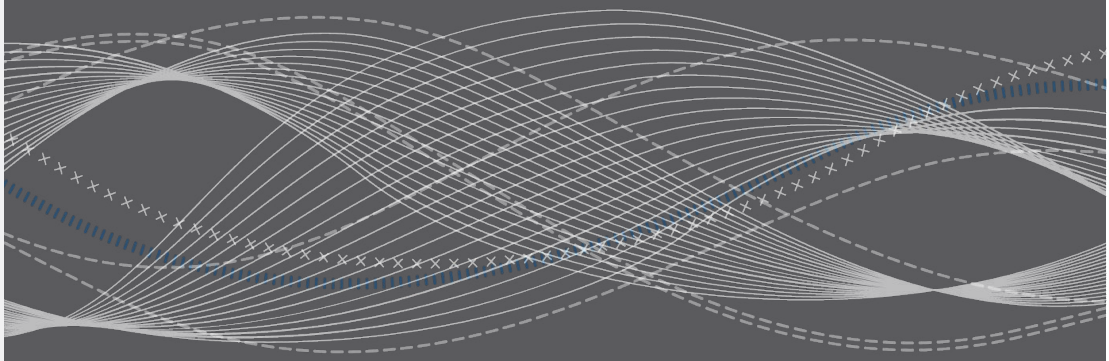
Joel Wright – Public Policy
Wright Consulting Services
Strawberry, Arizona

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.		

SUITABILITY OF DATA PRACTITIONER TASKS FOR AN INTERNSHIP

Less Suitable----------More Suitable

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.		

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

SECTION 4 – Evaluating Intern Performance

The means by which an intern's performance is assessed will depend upon the purpose of the evaluation and the amount of detail desired to meet that purpose. Employers and educators will want to choose and agree upon a tool that best communicates the extent to which an intern meets workplace expectations of data worker skills:

1) What will be evaluated?
2) Will soft skills (workplace behaviors and dispositions) be evaluated as well as technical skills?
3) How will performance be measured?
4) How frequently will interns be evaluated?
5) Who will be responsible for evaluating interns?
6) How much time should be devoted to evaluating interns?
7) To what extent will an intern be involved in his or her evaluation?

Included in this guide are three sample evaluation tools that can be used to assess the performance of the work responsibilities presented in the data practitioner profile. In addition, a sample tool is included to assess the extent to which interns possess the soft skills (employability skills) deemed essential to job success.

Evaluation tools aid in student learning in various ways. The sample evaluation tools included in this guide help to clarify workplace standards by providing specific language describing what employers expect to observe in their data workers. By illustrating what competence looks like on the job, the evaluation tools enable students to identify and to set concrete goals to be met during their internship experience. The specificity of the language found in the profile and the sample tools assist interns to articulate questions about work assignments in discussions with supervisors and to self-assess gaps in skills. These evaluation tools serve as a bridge connecting the internship experience to the college curriculum, providing material for conversations with faculty and classmates about work experiences, successes, and challenges. Lastly, interns can use the language provided on evaluation tools to build strong resumes, to describe what they know and are able to do during job interviews, and to serve as a framework for building their portfolios.

Evaluating Technical Skills

Each of the following examples is aligned to the Profile of a Data Practitioner. Evaluation Tool A – Ten Point Scale and Evaluation Tool B – Performance Area Checklist draw upon a collection of performance-based rubrics developed by the expert data workers who collaborated with EDC to develop the Profile. The rubrics provide examples of what the work responsibilities of a data practitioner look like when performed at four different levels of proficiency. They are built from simple, contextualized examples of a data practitioner's work activities, then scaffolded to illustrate increasing skill mastery. Employers and college representatives can either use one of the tools presented here or use these examples as a starting point for designing a different means of evaluation.

Before selecting one of these tools, it would be helpful to review *Reading the Rubrics*, found on page 29, to fully understand their structure and their definition of worker proficiency. The key terms involved with each of the tools follow:

- **Performance area** – What is being assessed
- **Performance level** – How well the task(s) is performed
- **Performance statement** – A guide to aid judgment about the level of performance

Reading the Rubrics

The rubrics for each duty are organized by following the format below:

DUTY				
Performance Areas and Associated Tasks from the Profile ↓	**Level 1 Basic Competence**	**Level 2 Approaching Proficiency**	**Level 3 Proficiency**	**Level 4 Above Proficiency**
	Performance Statements	*Performance Statements*	*Performance Statements*	*Performance Statements*

The far left column on the rubric charts lists the ***performance area*** for the duty being analyzed. Performance areas integrate one or more work responsibilities identified on the Profile of a Data Practitioner.

The columns to the right of the performance area present four ***performance levels*** that show incremental levels of effectiveness. These levels range from basic competence to approaching proficiency to proficiency to above proficiency, with the following definitions:

- **Basic competence** – Represents the performance of someone who meets minimal job requirements. One would see the presence of work skills that still require further development.
- **Approaching proficiency** – Indicates a more highly developed skill level and a more mature integration of work experience.
- **Proficiency** – Provides examples of skill development expected of an effective data practitioner. This is someone who is able to integrate and to apply skills and knowledge to solve complex problems.
- **Above proficiency** – Includes statements that indicate exemplary job-related skills. Individuals at this level provide role models to other data practitioners.

Each performance level includes several ***performance statements***. The performance statements illustrate levels of effectiveness within the performance area. They are, in effect, performance indicators. When used to assess performance, performance statements presume mastery of the performance levels to the left. For example, within a line of a rubric, being proficient at a task presumes the intern is able to do the activities described at the basic competence and approaching proficiency levels.

Examples of Technical Evaluation Tools

Following is a sample page of each evaluation tool for technical skills based upon the Profile of a Data Practitioner. Full versions of each evaluation tool are in the appendix.

Evaluation Tool A – Ten Point Scale

Tool A targets the rubric *performance areas* for evaluation. It uses a numerical scale ranging from 1 to 10 to signify intern performance. The numerical scale is distributed across all four levels of proficiency.

Evaluation Tool A – Ten Point Scale

Evaluation by performance area/Degrees of proficiency/Assessment scale of 1 (low basic competence) to 10 (above proficiency).
Place number representing intern performance at the end of each row.

DUTY 1: INITIATES THE DATA

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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.	e.g. 2
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.	e.g. 7
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.	e.g. 5
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.	e.g. NA

Advantages to Using Tool A

- It provides detailed guidance for making judgments about performance by offering examples of different levels of proficiency.
- By providing the complete rubric, it enables employers, faculty, and interns to envision how basic proficiency scales up to expertise.
- By using rubric performance areas to organize evaluation of the intern, it aligns his or her evaluation to the tasks or clusters of tasks that employers have identified as job activities that they assess.
- It provides room for employers to indicate gradations of achievement of each proficiency level.

Evaluation Tool B – Performance Area Checklist

Tool B also focuses evaluation on the rubric *performance areas*, but it eliminates all of the performance statements within the rubric. The rubric performance levels alone (basic competence, approaching proficiency, proficiency, and above proficiency) are used to measure intern performance.

Evaluation Tool B – Performance Area Checklist

Evaluation by performance area/General categories of proficiency/Assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 1: INITIATES THE DATA

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
	Performance of Intern			
Translates business problems into analytic process* *(Integrates Task A)*		X		
Integrates stakeholders' interests into proposed analytic process *(Integrates Tasks B, C)*	X			
Determines what data is needed to execute the analytic process *(Integrates Tasks D, E, F)*	X			
Develops statement of work* *(Integrates Tasks G, H, I)*	NA			

Advantages to Using Tool B

- By eliminating the performance statements, it simplifies assessment, making it less burdensome for supervisors to complete.
- It provides flexibility for supervisors to exercise their judgement within a performance level.
- By using rubric performance areas to organize the evaluation of the intern, it aligns his or her evaluation to the tasks or clusters of tasks that employers have identified as job activities that they assess.

Technical Skills Evaluation Tool C – Task Checklist

Tool C focuses evaluation on the tasks of a data practitioner as presented in the Profile of a Data Practitioner. It eliminates all of the performance statements within the rubrics. The rubric performance levels alone (basic competence, approaching proficiency, proficiency and above proficiency) are used to measure intern performance.

Evaluation Tool C – Task Checklist

Evaluation by performance area/General categories of proficiency/Assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 1: INITIATES THE DATA

Profile Task	N/A	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
		Performance of Intern			
1A. Translates business problems into analytic needs.	X				
1B. Interviews stakeholders.	X				
1C. Refines stakeholder needs.					
1D. Identifies appropriate data.		X		X	
1E. Identifies whether data exists or not.			X		
1F. Performs gap analysis of the data.			X		
1G. Determines resource needs (e.g., SMEs, tools, timeliness).			X		
1H. Determines feasibility of analysis to be done.	X				
1I. Creates statement of work.	X				

KEY

LEVEL 1 – Performs as a student with no/limited job experience.

LEVEL 2 – Meets employer expectation for an entry level worker.

LEVEL 3 – Meets employer expectation for a proficient worker.

LEVEL 4 – Exceeds employer expectation for a proficient worker.

Advantages to Using Tool C

- It measures performance according to individual tasks, rather than grouping tasks in clusters that might not be relevant to the particular internship.
- It may provide a more accurate description of the responsibilities involved in a particular internship.
- It provides flexibility for supervisors to exercise their judgement within a performance level.

Evaluating Soft (Employability) Skills

There are numerous compilations of soft skills that aim to identify the essential behaviors and dispositions that lead to workplace success. The State of Ohio has established its own list of 14 Job Readiness Criteria. Included here is a chart that lists the OhioMeansJobs-Readiness Criteria. The column on the left provides criteria definitions. The next column to the right lists performance indicators that offer a more detailed description of what employers expect from their workers. The performance of each criterion is evaluated by marking either *outstanding*, *satisfactory*, *needs improvement* or *not applicable*. Employers and the college faculty may want to include some or all of these as part of an evaluation of an intern. The full version of this tool is in the appendix.

OhioMeansJobs-Readiness Criteria

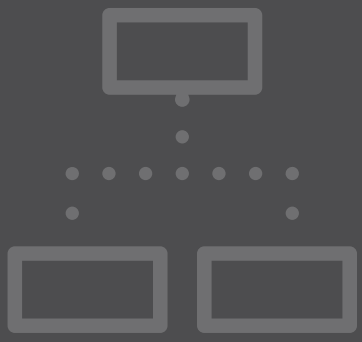
The State of Ohio's Job Readiness Criteria are listed in the far left column. The column to its right provides indicators drawn from the state's performance rubrics that can be used to guide assessment.

O = Outstanding S = Satisfactory N = Needs improvement N/A = Not applicable

Job Readiness Criteria	Performance Indicators	Performance of Intern			
		O	S	N	N/A
1. **Reliability** The intern has integrity and responsibility in professional settings.	• Manages commitments. • Has sense of right and wrong. • Adheres to standard operating procedures and organizational principles.		X		
2. **Work Ethic** The intern has effective work habits, personal accountability and a determination to succeed.	• Completes assignments on or ahead of schedule. • Manages a positive attitude and disposition. • Respectful to those in authority. • Seeks out and takes advantage opportunities for continued growth.		X		
3. **Punctuality** The intern arrives to commitments on time.	• Arrives ahead or on time prepared for assigned tasks.	X			
4. **Discipline** The intern abides by guidelines, demonstrates self-control, stays on task.	• Follows rules and regulations. • Stays on task and minimizes distractions. • Serves as an example of personal behavior.		X		
5. **Teamwork/ Collaboration** The intern builds collaborative relationships with others and can work as part of a team.	• Accepts responsibility for assignments and contributes to team projects. • Demonstrates willingness to contribute to all aspects of team projects. • Encourages all members to use personal strengths to achieve a common goal.			X	
6. **Professionalism** The intern demonstrates honesty. S/he dresses and acts appropriately and responsibly. S/he learns from mistakes.	• Maintains an appropriate appearance. • Demonstrates ethical behavior. • Resolves the needs of customers in a timely and professional manner. • Builds and maintains cooperative and respectful relationships with others.		X		

Source: Ohio Department of Education. (n.d.).The OhioMeansJobs-Readiness Seal rubric. http://education.ohio.gov/Topics/New-Skills-for-Youth/SuccessBound/OhioMeansJobs-Readiness-Seal
Used and adapted by permission, 2021.

SECTION 5 – Structuring the Internship



In addition to identifying what an intern does, companies need to determine what steps will be taken to situate interns within the normal flow of business. Questions such as the following should guide employers' thinking:

1) How will interns be supervised?
2) What current employee(s) will be responsible for supervising interns?
3) What will be involved with the supervision of interns?
4) What will be the procedures for communicating between supervisors, interns, and the college?
5) When will interns be expected to work (e.g., what months, days of the week, hours of the day)?
6) How will interns be compensated?
7) What kind of orientation to the company and to the specific assignments will be necessary?
8) Who will be responsible for providing interns with an orientation?

Intern Supervisors

A key to hosting successful internships lies in the selection of employees who will serve as supervisors. Employees who supervise interns should do the following:

- Have a strong understanding of the project and the requisite responsibilities that will be assigned to the intern
- Take the time to familiarize the intern with the company
- Make themselves available for regular interactions to answer questions and to provide feedback
- Establish scheduled times to conduct more formal evaluations of the intern's performance
- Maintain communication with college staff as needed and appropriate

In addition to monitoring technical performance, supervisors of interns typically take on the role of a coach or mentor. They help acclimate interns to the work environment and model professional behavior. One university provides these tips to intern supervisors:

- **Communicate** – Give constructive feedback on a regular basis.
- **Be prepared** – Have tasks ready. When an intern is successful with one project, give him/her another task that uses the same skills on a larger scale or adds new skills.
- **Include the intern** – Make the intern feel a part of the group. Take him/her to meetings/lunches.
- **Be sensitive** – Interns do not always possess the everyday knowledge that you take for granted, such as office etiquette and appropriate dress. Provide coaching on these topics when needed.
- **Be interested** – Listen to the intern's ideas. He/she wants to contribute to the organization.[2]

Orientation

Providing interns with an orientation is an important means to communicate company expectations and to help them acclimate themselves to the company's culture. Ideally, the individual(s) involved in supervising an intern should be involved in providing the orientation. While a company will want to tailor an orientation to address its particular environment, typically topics include the following:

2 Methodist University. (n.d.). *Employer guide to structuring a successful internship program* (p. 11). Retrieved from https://www.methodist.edu/wp-content/uploads/2019/09/employer_guide.pdf

- Company mission
- Company history
- Key services or products
- Intern work responsibilities and their connection to company mission
- Work schedule
- Dress code
- Reporting responsibilities – To whom, by what method, and how frequently
- Who to contact for help
- Policies regarding absenteeism
- Safety procedures

The University of Richmond has a very extensive list of topics that can be addressed during an intern's orientation. It can be accessed at https://careerservices.richmond.edu/employers/post-job-internship/internships/pdfs/orientation-sample.pdf


1037.06 1051.32 1037.04 1051.32
Open
High
Low
Close
Feb March April May June July Aug Sep Oct Nov Dec

APPENDIX

Evaluation Tool A – Ten Point Scale

Evaluation Tool B – Performance Area Checklist

Evaluation Tool C – Task Checklist

OhioMeansJobs-Readiness Criteria

Data Practitioner Rubrics – Glossary

Evaluation Tool A – Ten Point Scale

Evaluation by performance area/degrees of proficiency/assessment scale of 1 (low basic competence) to 10 (above proficiency).
Place number representing intern performance at the end of each row.

DUTY 1: INITIATES THE DATA

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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.	

Evaluation Tool A – Ten Point Scale

Evaluation by performance area/degrees of proficiency/assessment scale of 1 (low basic competence) to 10 (above proficiency).
Place number representing intern performance at the end of each row.

DUTY 2: SOURCES THE DATA

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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.	
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.	

Duty 2: Sources the Data (cont.)

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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).	

Evaluation Tool A – Ten Point Scale

Evaluation by performance area/degrees of proficiency/assessment scale of 1 (low basic competence) to 10 (above proficiency).
Place number representing intern performance at the end of each row.

DUTY 3: TRANSFORMS THE DATA

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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.	

Evaluation Tool A – Ten Point Scale

Evaluation by performance area/degrees of proficiency/assessment scale of 1 (low basic competence) to 10 (above proficiency).
Place number representing intern performance at the end of each row.

DUTY 4: ANALYZES THE DATA

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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.	
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....).	

Evaluation Tool A – Ten Point Scale

Evaluation by performance area/degrees of proficiency/assessment scale of 1 (low basic competence) to 10 (above proficiency).
Place number representing intern performance at the end of each row.

DUTY 5: CLOSES OUT THE PROJECT

Performance Area	Level 1 Basic Competence 1----------→3	Level 2 Approaching Proficiency 4----------→6	Level 3 Proficiency 7----------→9	Level 4 Above Proficiency 10	Performance of Intern
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.	

Evaluation Tool A – Ten Point Scale

DUTY 6: ENGAGES IN PROFESSIONAL DEVELOPMENT

The performance statements listed for professional development do not lend themselves to being divided into levels. They are best assessed as either *done* or *not done*.

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency	Performance of Intern
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				Check each one performed.
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				Check each one performed.

Evaluation Tool B – Performance Area Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 1: INITIATES THE DATA

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
	Performance of Intern			
Translates business problems into analytic process* *(Integrates Task A)*				
Integrates stakeholders' interests into proposed analytic process *(Integrates Tasks B, C)*				
Determines what data is needed to execute the analytic process *(Integrates Tasks D, E, F)*				
Develops statement of work* *(Integrates Tasks G, H, I)*				

Evaluation Tool B – Performance Area Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 2: SOURCES THE DATA

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
	Performance of Intern			
Facilitates data collection *(Integrates Tasks A, B, C, D, E)*				
Prepares data for transformation *(Integrates Tasks F, G)*				
Produces data dictionary* *(Integrates Tasks H, 3L)*				
Applies business, ethical and legal standards *(Integrates Tasks I)*				

Evaluation Tool B – Performance Area Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 3: TRANSFORMS THE DATA

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
	Performance of Intern			
Prepares Data for Organization *(Integrates Tasks A, B, C, D, E, F, G)*				
Aligns data to structure of analysis tools *(Integrates Tasks H, I, J, K, M)*				

Evaluation Tool B – Performance Area Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 4: ANALYZES THE DATA

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
	Performance of Intern			
Applies research method *(Integrates Tasks A, B, C, D)*				
Performs data mining* *(Integrates Tasks E, F)*				
Interprets the results *(Integrates Tasks G, H, I, J, K, L)*				
Creates data visualizations (dashboards,* reports, maps, charts, graphs, videos.....) *(Integrates Tasks M)*				

Evaluation Tool B – Performance Area Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 5: CLOSES OUT THE PROJECT

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
	Performance of Intern			
Communicates the results of the analysis *(Integrates Tasks A, B, C, E, F, G)*				
Documents the analysis *(Integrates Tasks H, I, J, K)*				

Evaluation Tool B – Performance Area Checklist

Duty 6: Engages in Professional Development

The performance statements listed for professional development do not lend themselves to being divided into levels. They are best assessed as either *done* or *not done*.

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency	Performance of Intern
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				Check each one performed.
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				Check each one performed.

Evaluation Tool C – Task Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 1: INITIATES THE DATA

Profile Task	N/A	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
		Performance of Intern			
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, timeliness).					
1H. Determines feasibility of analysis to be done.					
1I. Creates statement of work.					

KEY

LEVEL 1 – Performs as a student with no/limited job experience.

LEVEL 2 – Meets employer expectation for an entry level worker.

LEVEL 3 – Meets employer expectation for a proficient worker.

LEVEL 4 – Exceeds employer expectation for a proficient worker.

Evaluation Tool C – Task Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 2: SOURCES THE DATA

Profile Task	N/A	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
		Performance of Intern			
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.					

KEY

LEVEL 1 – Performs as a student with no/limited job experience.

LEVEL 2 – Meets employer expectation for an entry level worker.

LEVEL 3 – Meets employer expectation for a proficient worker.

LEVEL 4 – Exceeds employer expectation for a proficient worker.

Evaluation Tool C – Task Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 3: TRANSFORMS THE DATA

Profile Task	N/A	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
		Performance of Intern			
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.					
3L. Finalizes data dictionary.					
3M. Stores data for analytics.					

KEY

LEVEL 1 – Performs as a student with no/limited job experience.

LEVEL 2 – Meets employer expectation for an entry level worker.

LEVEL 3 – Meets employer expectation for a proficient worker.

LEVEL 4 – Exceeds employer expectation for a proficient worker.

Evaluation Tool C – Task Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 4: ANALYZES THE DATA

Profile Task	N/A	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
		Performance of Intern			
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).					

KEY

LEVEL 1 – Performs as a student with no/limited job experience.

LEVEL 2 – Meets employer expectation for an entry level worker.

LEVEL 3 – Meets employer expectation for a proficient worker.

LEVEL 4 – Exceeds employer expectation for a proficient worker.

Evaluation Tool C – Task Checklist

Evaluation by performance area/general categories of proficiency/assessment scale of basic competence to above proficiency.
Place "X" in cell that best represents intern's performance.

DUTY 5: CLOSES OUT THE PROJECT

Profile Task	N/A	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency
		Performance of Intern			
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.					

KEY

LEVEL 1 – Performs as a student with no/limited job experience.

LEVEL 2 – Meets employer expectation for an entry level worker.

LEVEL 3 – Meets employer expectation for a proficient worker.

LEVEL 4 – Exceeds employer expectation for a proficient worker.

Evaluation Tool C – Task Checklist

DUTY 6: ENGAGES IN PROFESSIONAL DEVELOPMENT

The performance statements listed for professional development do not lend themselves to being divided into levels. They are best assessed as either *done* or *not done*.

Performance Area	Level 1 Basic Competence	Level 2 Approaching Proficiency	Level 3 Proficiency	Level 4 Above Proficiency	Performance of Intern
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				Check each one performed.
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				Check each one performed.

OhioMeansJobs-Readiness Criteria

The State of Ohio's Job Readiness Criteria are listed in the far left column. The column to its right provides indicators drawn from the state's performance rubrics that can be used to guide assessment.

O = Outstanding **S = Satisfactory** **N = Needs improvement** **N/A = Not applicable**

Job Readiness Criteria	Performance Indicators	Performance of Intern			
		O	S	N	N/A
1. **Reliability** The intern has integrity and responsibility in professional settings.	• Manages commitments. • Has sense of right and wrong. • Adheres to standard operating procedures and organizational principles.				
2. **Work Ethic** The intern has effective work habits, personal accountability and a determination to succeed.	• Completes assignments on or ahead of schedule. • Manages a positive attitude and disposition. • Respectful to those in authority. • Seeks out and takes advantage opportunities for continued growth.				
3. **Punctuality** The intern arrives to commitments on time.	• Arrives ahead or on time prepared for assigned tasks.				
4. **Discipline** The intern abides by guidelines, demonstrates self-control, stays on task.	• Follows rules and regulations. • Stays on task and minimizes distractions. • Serves as an example of personal behavior.				
5. **Teamwork/ Collaboration** The intern builds collaborative relationships with others and can work as part of a team.	• Accepts responsibility for assignments and contributes to team projects. • Demonstrates willingness to contribute to all aspects of team projects. • Encourages all members to use personal strengths to achieve a common goal.				
6. **Professionalism** The intern demonstrates honesty. S/he dresses and acts appropriately and responsibly. S/he learns from mistakes.	• Maintains an appropriate appearance. • Demonstrates ethical behavior. • Resolves the needs of customers in a timely and professional manner. • Builds and maintains cooperative and respectful relationships with others.				

OhioMeansJobs-Readiness Criteria (cont.)

Job Readiness Criteria	Performance Indicators	Performance of Intern			
		O	S	N	N/A
7. Learning Agility The intern desires to continuously learn new information and skills.	• Seeks out opportunities to expand knowledge. • Synthesizes knowledge and experiences and applies them in various situations. • Shares insights gained and engages with others.				
8. Critical Thinking/ Problem-Solving The intern exercises strong decision-making skills, analyzes issues effectively and thinks creatively to overcome problems.	• Identifies and examines problems, considers risks and proposes solutions. • Anticipates potential issues. • Navigates challenging situations independently.				
9. Leadership The intern leverages the strengths of others to achieve common goals. S/he coaches and motivates peers and can prioritize and delegate work.	• Demonstrates ability to lead regardless of position. • Capable of prioritizing and delegating work. • Promotes creativity and energy among team through persuasion and influence. • Achieves high levels of performance from team by coaching and encouraging individuals.				
10. Creativity/ Innovation The intern is original and inventive. S/he communicates new ideas to others, drawing on knowledge from different fields to find solutions.	• Generates and communicates original ideas. • Demonstrates ability to think differently. • Sought out for ability to create energy and draw out ideas from others.				
11. Oral and Written Communication The intern articulates thoughts and ideas clearly and effectively in written and oral forms.	• **Oral:** • Speaks clearly and expressively without hesitation. • Tone and volume are appropriate and consistent. • Extremely well poised and engages and captivates audience. • **Written:** • Skillfully communicates meaning with clarity and fluency, and is error-free. • Adheres to conventions of selected writing style (e.g., APA, Chicago, MLA). • Writing is consistently detailed and precise.				

OhioMeansJobs-Readiness Criteria (cont.)

Job Readiness Criteria	Performance Indicators	Performance of Intern			
		O	S	N	N/A
12. Digital Technology The intern has an in-depth understanding of emerging technology and leverages technology to solve problems, complete tasks and accomplish goals.	• Is an expert in the use of current digital technologies and can select appropriate technology for the setting. • Uses existing and emerging technologies to solve problems, complete tasks and accomplish goals. • Can teach others to use existing and emerging technologies. • Is proficient in the use of current digital technologies, may need assistance to select appropriate technology for the setting. • Can demonstrate existing and emerging technologies to others.				
13. Global/ intercultural fluency The intern values, respects and learns from diverse groups of people.	• Values and respects people from all cultures. • Seeks out opportunities to facilitate growth in cultural diversity. • Openly accepts another's culture as a part of his/ her working relationship. • Accepts another's culture as a part of his or her working relationship. Participates in cultural awareness conversations.				
14. Career management The intern is a self-advocate. S/he articulates strengths, knowledge and experiences, relevant to success in a job or postsecondary education.	• Has established goals aligned to chosen career pathway. • Can articulate long-term plans and steps to achieve them. • Utilizes strengths and manages weaknesses. • Seeks out and takes advantage of opportunities for self-development. • Mentors others on career planning.				

Source: Ohio Department of Education. (n.d.).The OhioMeansJobs-Readiness Seal rubric. http://education.ohio.gov/Topics/New-Skills-for-Youth/SuccessBound/OhioMeansJobs-Readiness-Seal
Used and adapted by permission, 2021.

Data Practitioner Rubrics – 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 Education Development Center

EDC 43 Foundry Avenue, Waltham, MA 02453
E-mail: contact@edc.org | Phone: 617-969-7100
Boston | Chicago | New York | Washington, D.C.