Data Analysis Boot Camp

Duration: 2 Days Course Code: GK2397 Delivery Method: Company Event

Overview:

Learn to communicate about data and insights to stakeholders who need to make quick business decisions to drive your organization forward.

This three-day course leverages straightforward business examples to explain practical techniques for understanding and reviewing data quality.
This course starts with an overview of data quality and data management, followed by foundational analysis and statistical techniques.
This data analysis training class is a lively blend of expert instruction combined with hands-on exercises so you can practice new skills and leave prepared to start performing practical analysis techniques. Every Data Analysis Boot Camp instructor is a veteran consultant and data guru who guides you through effective best practices and technologies for working with your data.
NB: Labs for this course are primarily in Microsoft Excel, however, students will get an opportunity to practice using R during some labs. Labs for this course can also be taught using the Python programming language for private team deliveries.

Updated Nov2026

Company Events

These events can be delivered exclusively for your company at our locations or yours, specifically for your delegates and your needs. The Company Events can be tailored or standard course deliveries.

Target Audience:

Professionals who benefit from this course include: Business Analyst and Systems Analyst Operations Researcher and Marketing Analyst Project Manager, Program Manager, Team Leader Data Modeler or Administrator Database Administrator IT Manager, Director, VP, Finance Manager, Operations Supervisor, Risk Assessment Manager Process Improvement Staff Executives exploring cost reduction and process improvement optionsSenior staff who make or recommend key business decisions

Objectives:

- In this class participants will learn how to:
- Identify opportunities, manage change and develop deep visibility into the organization
- Understand the terminology and jargon of analytics, business intelligence, and statistics
- Use a wealth of practical applications for applying data analysis capability
- Visualize both data and the results of analysis for straightforward graphical presentation to stakeholders
- Estimate more accurately while accounting for variance, error, and confidence intervals
- Create a valuable array of plots and charts to reveal hidden trends and patterns in data
- Differentiate between "signal" and "noise" in data
- Understand and leverage different distribution models, and how each applies in the real world
- Form and test hypotheses using multiple methods to define and interpret useful predictions
- Identify statistical inference and drawing conclusions about the population

Prerequisites:

Students attending this class should have a basic understanding of how data is currently used in their organization. They should also have strong Excel aptitude.

Testing and Certification

Follow-on-Courses:

No specific recommandation

Content:

Part 1: The Value and Challenges of Data-Driven Disruption

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*

Part 7: Data Visualization ; Reporting

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*

- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Part 2: Tying Data to Business Value

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields

- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- Objectives and expectations

- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Part 3: Understanding Your Data

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*

- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis

- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools

- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis

- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment

- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Part 5: Thinking Critically About Your Analysis

- Objectives and expectations
- Hurdles to becoming a data-driven organization

- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Part 8: Hands-On Introduction to R and R Studio

- Objectives and expectations
- Hurdles to becoming a data-driven organization

- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases

- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI

- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems

- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Part 4: Analyzing Data

- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Part 6: Data Analysis in the Real World

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data

- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases
- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

- Objectives and expectations
- Hurdles to becoming a data-driven organization
- Data empowerment
- Instilling data practices in the

- organization
- The CRISP-DM model of data projects
- What constitutes data-driven value
- Requirements gathering: How to approach it
- Kanban for data analysis
- Know your customers
- Stakeholder cheat sheets
- *EXERCISE: Data-driven project checklist*
- *LAB: Data analysis techniques: Aggregations*
- Data defined
- Data versus information
- Types of data
- Unstructured vs. Structured
- Time scope of data
- Sources of data
- Data in the real world
- The 3 V's of data
- Data Quality
- Cleansing
- Duplicates
- SSOT
- Field standardization
- Identify sparsely populated fields
- How to fix common issues
- *LAB: Prioritizing data quality*
- Analysis foundations
- Comparing programs and tools
- Words in English vs. data
- Concepts specific to data analysis
- Domains of data analysis
- Descriptive statistics
- Inferential statistics
- Analytical mindset
- Describing and solving problems
- Averages in data
- Mean
- Median
- Mode
- Range
- Central tendency
- Variance
- Standard deviation
- Sigma values
- Percentiles
- Demystifying statistical models
- Data analysis techniques
- *LAB: Central tendency*
- *LAB: Variability*
- *LAB: Distributions*
- *LAB: Sampling*
- *LAB: Feature engineering*
- *LAB: Univariate linear regression*
- *LAB: Prediction*
- *LAB: Multivariate linear regression*
- *LAB: Monte Carlo simulation*
- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis
- Deployment of analyses
- Best practices for BI
- Technology ecosystems
- Relational databases

- NoSQL databases
- Big data tools
- Statistical tools
- Machine learning
- Visualization and reporting tools
- Making data useable
- Best practices for data visualizations
- Visualization essentials
- Users and stakeholders
- Stakeholder cheat sheet
- Common presentation mistakes
- Goals of visualization
- Communication and narrative
- Decision enablement
- Critical characteristics
- Communicating data-driven knowledge
- Formats and presentation tools
- Design considerations
- What is R?
- *LAB: Intro to R Studio*
- *LAB: Univariate linear regression in R*
- *LAB: Multivariate linear regression in R*

Additional Information:

Digital course book provided

Further Information:

For More information, or to book your course, please call us on +1 (800) 268 7737

ilt.customerservice@globalknowledge.com

www.globalknowledge.com/en-ca/

Global Knowledge Network (Canada) Inc., 150 Knowledge Park Drive, Suite 130, Fredericton, NB, E3C 0N9