

# Course Outline

## Data Analysis Boot Camp Course DABC: 3 days Instructor Led

### About this course

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. Throughout the course, you learn to communicate about data and insights to stakeholders who need to make quick business decisions to drive your organization forward.

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.

*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.*

### Audience profile

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.

Professionals who benefit from this course include:

- Business Analyst
- Systems Analyst
- Operations Researcher
- Marketing Analyst
- Project Manager
- Program Manager
- Team Leader
- Data Modeler or Administrator
- Database Administrator
- IT Manager, Director, VP
- Finance Manager, Director, VP
- Operations Supervisor, Manager, Director, VP
- Risk Assessment Manager
- Process Improvement Staff
- Executives exploring cost reduction and process improvement options
- Senior staff who make or recommend key business decisions

# Course Outline

## At course completion

After completing this course, students will be able to:

- Identify opportunities, manage change and develop deep visibility into your 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 your 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 your data
- Differentiate between "signal" and "noise" in your 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

## Course Outline

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

1. Objectives and expectations
2. Hurdles to becoming a data-driven organization
3. Data empowerment
4. Instilling data practices in the organization
5. The CRISP-DM model of data projects

### Part 2: Tying Data to Business Value

1. What constitutes data-driven value
2. Requirements gathering: How to approach it
3. Kanban for data analysis
4. Know your customers
5. Stakeholder cheat sheets
- **EXERCISE:** *Data-driven project checklist*
- **LAB:** *Data analysis techniques: Aggregations*

### Part 3: Understanding Your Data

1. Data defined
2. Data versus information
3. Types of data
  1. Unstructured vs. Structured
  2. Time scope of data
  3. Sources of data
4. Data in the real world
5. The 3 V's of data
6. Data Quality

## Course Outline

1. Cleansing
  2. Duplicates
  3. SSOT
  4. Field standardization
  5. Identify sparsely populated fields
  6. How to fix common issues
- **LAB:** *Prioritizing data quality*

### Part 4: Analyzing Data

1. Analysis foundations
  1. Comparing programs and tools
  2. Words in English vs. data
  3. Concepts specific to data analysis
  4. Domains of data analysis
  5. Descriptive statistics
  6. Inferential statistics
  7. Analytical mindset
  8. Describing and solving problems
2. Averages in data
  1. Mean
  2. Median
  3. Mode
  4. Range
3. Central tendency
  1. Variance
  2. Standard deviation
  3. Sigma values
  4. Percentiles
4. Demystifying statistical models
5. 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*

### Part 5: Thinking Critically About Your Analysis

1. Descriptive analysis
2. Diagnostic analysis
3. Predictive analysis
4. Prescriptive analysis

# Course Outline

## Part 6: Data Analysis in the Real World

1. Deployment of analyses
2. Best practices for BI
3. Technology ecosystems
  1. Relational databases
  2. NoSQL databases
  3. Big data tools
  4. Statistical tools
  5. Machine learning
  6. Visualization and reporting tools
4. Making data useable

## Part 7: Data Visualization & Reporting

1. Best practices for data visualizations
  1. Visualization essentials
  2. Users and stakeholders
  3. Stakeholder cheat sheet
2. Common presentation mistakes
3. Goals of visualization
  1. Communication and narrative
  2. Decision enablement
  3. Critical characteristics
4. Communicating data-driven knowledge
  1. Formats and presentation tools
  2. Design considerations

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

1. What is R?
  - **LAB:** *Intro to R Studio*
  - **LAB:** *Univariate linear regression in R*
  - **LAB:** *Multivariate linear regression in R*