

## Training: CompTIA CompTIA Data+ Prep Course



### TRAINING GOALS:

The CompTIA Data+ certification is a foundational-level certification designed for professionals with 18-to-24 months of hands-on experience in that includes exposure to databases and analytical tools, a basic understanding of statistics, and familiarity with data virtualization.

This course can benefit you in two ways. If you intend to pass the CompTIA Data+ (DA0-002)(DA0-002) exam to receive a Data+ certification, this course can be a significant part of your preparation. However, certification is not the only key to professional success in the field of IT support. Today's job market demands individuals have demonstrable skills, and the information and activities in this course can help you build your skill set so that you can confidently perform your duties in any entry-level IT support role.

Upon course completion, you will be able to:

- Translate business requirements in support of data-driven decisions
- Acquire data
- Prepare data
- Transform data
- Create appropriate reports and visualizations
- Apply basic statistical methods and analyze complex data sets
- Adhere to governance and quality standards throughout the entire data life cycle

### Skills you'll learn

- Build a foundation in modern data concepts so you can confidently navigate analytics tools and data sources in any organization.
- Transform, clean, and organize raw data to make it reliable and useful for analysis.
- Apply statistical methods to extract trends, uncover insights, and support business decisions.
- Use data visualizations and dashboards to translate complex results into presentations anyone can understand.

- Maintain data quality, ensure compliance, and protect sensitive information in alignment with industry standards.
- Solve data problems, validate reports, and resolve analytics issues as they arise on the job.

#### Job roles that benefit from Data+ skills

- Junior Data Analyst
- Business Intelligence (BI) Analyst
- Marketing Analyst
- Operations Analyst
- Data Reporting Specialist
- Financial Analyst
- Business Analyst
- Junior Data Scientist

This is the perfect course for people who want to go beyond simple Excel tables and learn professional analysis, data visualization, and working with databases to support management decisions.

*Each participant in an authorized training CompTIA Data+ Prep Course held in Compendium CE will receive a free DA0-002 CompTIA Data+ Certification Exam vouchers.*

#### CONSPECT:

- Summarizing Database Concepts
  - Describe Different Data Structures
    - Structured Data versus Unstructured Data
    - Structured Data
    - Unstructured Data
    - Semi-Structured Data
  - Identify Relational and Non-Relational Databases
    - Relational Databases
    - Relational Databases
    - ACID Principles
    - Non-Relational Databases
    - Key Differences between Relational and Non-Relational Databases

- Live Lab: Explore the Lab Environment
- Live Lab: Navigate Database Design
- Describe the Way We Use Tables, Primary Keys, and Normalization
  - Normalization
  - Relationships in Data
  - Primary Keys and Indexing
  - Types of Relationships
  - Identifying Relationships in Data
  - Relating Table Records to Each Other
  - Lab: Database Design in Microsoft Access
  - Lab: Tables and Relationships
  - Referential Integrity
  - Denormalized Data
  - Activity: Recognize Data Design Standards
- Differentiate Field Data Types
  - Introduction to Field Data Types
  - Field Data Types
  - String Data Types
  - Date Data Type
  - Numeric Data Types
  - Currency Data Type
  - Boolean Data Type
  - Spatial Data
  - Large Objects
  - Data Type Conversion
  - Live Lab: Implement Data Type Selection and Conversion
  - Applied Live Lab: Demonstrate Data Structure Types and SQL Statements
  - Universal Identifiers
- Comparing and Contrasting Different Data Systems
  - Describe Types of Data Processing and Storage Systems
    - Types of Data Processing
    - Source Systems
    - Data Warehouses and Data Marts
    - Schemas and Data Modeling
    - Data Lakes and Lakehouses
  - Identify Types of Infrastructure

- Cloud vs. On-Premise
- Platform as a Service (PaaS)
- Software as a Service (SaaS)
- Infrastructure as a Service (IaaS)
- Explain How Data Changes
  - Overview of Slowly Changing Dimensions
  - The Impact of Slowly Changing Dimensions
  - Transaction Control Language (TCL)
  - Committing Data
  - Rolling Back Data
- Recognizing AI's Impact on Data Projects
  - Define Artificial Intelligence Concepts
    - Generative AI
    - Structure an AI Prompt
    - Deep Learning
    - Natural Language Processing
    - Introduction to Natural Language Processing
    - Preparing Data for NLP
  - Describe Machine Learning
    - Machine Learning Concepts
    - Introduction to Machine Learning
    - Supervised Learning
    - Unsupervised Learning
    - Machine Learning Techniques
    - Data Analysis and Machine Learning
  - Employ Robotic Process Automation
    - Robotic Process Automation (RPA)
    - Automating Reporting with RPA
    - Build a Basic RPA to Deliver Custom Messages
- Comparing Languages and Tools for Data
  - Discuss Coding Environments and Languages
    - Understanding Software Development
    - What is Code?
    - Types of Coding and Software
    - Text Editor - Notepad vs Notepad++
    - Managing Database Systems

- Structured Query Language (SQL)
- Comparing SQL Statements
- Coding R for Statistical Analysis
- Coding Python for Statistics
- Using Python for Data Analytics
- Coding in Python with VS Code
- Identify Markup Languages
  - HyperText Markup Language (HTML)
  - Extensible Markup Language (XML)
  - JavaScript Object Notation (JSON)
  - Activity: Recognize Languages Used for Data
- Identify Common Data Analytics Tools
  - Business Intelligence Software
  - Statistical Tools and Programming Languages
  - Database Management Tools
- Using Data Acquisition Methods
  - Identify Different File Formats
    - Delimited Files
    - Flat Files
    - File Extensions
    - Live Lab: Prepare Data in Different File Formats
  - Work with Public Data
    - Public, Generated, and Synthetic Data
    - Finding Public and Publicly Available Data
    - Creating a Data Set from Census Data
    - Live Lab: Appraise Public Data
  - Collect Data
    - Building a Survey and Collecting Data
    - Research Questions
    - Data Sources and Collection Methods
    - Sampling and Observations
    - Creating Data Collections with Surveys
    - Activity: Use and Collect Survey Data
  - Integrate Data Into Tables
    - Data Definition Language (DDL)
    - Creating and Altering Databases

- Creating Tables
- Altering Tables
- Deleting Fields and Tables
- Creating Views
- Identify Types of Data Systems and Software
  - Application Programming Interface (API)
  - Live Lab: Leverage a Public API
  - Web Services
  - Web Scraping
  - Web Scraping with Python
  - Live Lab: Perform Web Scraping
  - Machine Data
  - Files and Log Data
- Explain the Processes of Extracting, Transforming, and Loading Data
  - Extracting, Loading, and Transforming Data
  - Extracting Data
  - Transforming Data
  - Loading Data
  - Extract, Load, Transform (ELT)
- Applying Quality Control to Data
  - Recall Characteristics, Rules, and Metrics of Data Quality
    - Reasons to Check Data Quality
    - Understanding Quality
    - Rules and Metrics for Data Quality
  - Perform Methods of Data Validation
    - Data Validation Methods
    - Automated Validation
    - Lab: Ensure Valid Input
    - Data Verification Methods
    - Activity: Validate Data
- Profiling and Cleansing Data
  - Profile Data
    - Steps of Data Profiling
    - Data Profiling Tools and Techniques
    - Live Lab: Profile Data Sets
    - Live Lab: Profile a Large Data Set

- Address Redundant, Duplicated, and Unnecessary Data
  - Redundant Data
  - Duplicated Data
  - Removing Duplicate Records
  - Identifying Duplicates in Large Data Sets
  - Unnecessary Fields
  - Lab: Addressing Redundant and Duplicated Data
- Address Missing and Null Values
  - Causes of Null Values
  - Filtering Null Values
  - Applying Logic to Missing Values
  - Live Lab: Address Missing Values
- Address Invalid Data
  - Correcting or Removing Invalid Data
  - Identifying Invalid Data
  - Removing and Replacing Invalid Data
  - Lab: Use Data to Inform Product Decisions
- Convert Data to Meet Specifications
  - Data That Does Not Meet Specifications
  - Converting Data Types
  - Applied Live Lab: Prepare Data for Use
- Executing Data Manipulation Techniques
  - Work with the Whole Data Set
    - Sorting Data
    - Filtering Data
    - Filtering Data on Dashboards and Visuals
    - Live Lab: Filter Dashboard Data
    - Aggregation and the Basic Types of Aggregate Functions
    - Add and Average
    - Lab: Use Arithmetic Operators
    - Reduction in Data Sets
  - Use Functions to Manipulate Data
    - Using Functions to Manipulate Data
    - Cleaning String Functions
    - Text Functions
    - Creating Dates with Functions

- Date and Time Functions
- Lab: Use Advanced Date and Time Functions
- Lab: Use Advanced Date and Time Functions on a Project Timeline
- Lab: Use Advanced Date and Time Functions in Project Planning
- Creating Date Tables
- Combining Data Fields
- Concatenating in Excel
- Parsing Strings for Information
- Correcting String Data
- Standardizing Strings with Functions
- Create and Recode Variables
  - Derived Variables
  - Recoding Numerical and Categorical Data for Grouping and Standardization
  - Recoding Text Values to Numbers
  - Activity: Recode Numerical and Categorical Data
  - Logical Functions and Conditional Data
  - Logical Functions
  - The IF Function
  - Lab: Use IF Functions
  - Lab: Perform Logical Operations on Payroll Data
  - Lab: County Fair
  - Lab: Perform Logical Operations in Formulas
  - Lab: Use Data to Determine Shipping Costs
  - Imputing Values
- Building Queries to Model Data
  - Write Queries on Tables
    - Common SQL Commands
    - Querying Data
    - Lab: Select All Records in a Table
    - Lab: Select Records with a Single Condition
    - Lab: Select Records with Multiple Conditions
    - Types of Joins
    - The Impact of Each Join to Data
    - Discovering How Joins Impact Data Results
    - Using the WHERE Statement to Filter
    - Lab: Microsoft Access Queries

- Grouping and Aggregating Data
- Live Lab: Implement Queries and Join Types
- Transpose and Append Data
  - Transposing Data and Appending Data
  - Exploding Data
  - Appending Data
  - Merging Data
  - Merging Data to Create a Data Set
  - Live Lab: Recode Data
- Update, Insert, and Delete Data
  - Updating Data
  - Inserting Data
  - Deleting Data
  - Lab: Delete a Record from a Table
  - Live Lab: Execute Data Manipulation Language (DML)
- Preparing for Data Analysis
  - Recognize Characteristics of Data
    - Quantitative and Qualitative Data
    - Activity: Compare Qualitative and Quantitative Data
    - Nominal and Ordinal Data
  - Perform Basic Statistical Methods of Analysis
    - Exploratory Data Analysis
    - Methods of Basic Statistics
    - Performance Analysis
    - Gap Analysis
    - Trend Analysis
- Applying Descriptive Statistical Methods
  - Use Measures of Central Tendency
    - Overview of the Measures of Central Tendency
    - Central Tendency
      - Mean
      - Median
      - Mode
    - Lab: Using the Measures of Central Tendency
    - Live Lab: Calculate Measures of Central Tendency
  - Use Measures of Dispersion

- Distribution of a Data Set
- The Range of Data
- Standard Deviation
- Z-Scores
- Lab: Using Measures of Variability
- Live Lab: Calculate Measures of Variability
- Use Frequency and Percentages
  - Frequency
  - Percentage Difference
  - Percentage Change
  - Calculating Percentages
  - Activity: Use Frequencies and Percentages
- Illustrating Different Statistical Methods
  - Explain the Importance of Statistical Tests
    - Understanding the Importance of Statistical Tests
    - Samples
    - Confidence Intervals
    - P-Values
  - Utilize Tests and Methods to Determine Relationships Between Variables
    - Hypothesis Testing
    - T-Tests
    - Chi-Square
    - Chi-Square Tests
    - Activity: Analyze Student Data
    - Simple Linear Regression
    - Correlation
    - Use Excel to Apply Statistical Methods
    - Applied Live Lab: Analyze Data
- Summarizing Business Requirements in a Report Format
  - Determine Communication Approaches
    - Providing Access for an Internal versus External Audience
    - Controlling the Level of Detail Based on Consumer Types
    - Technical versus Non-Technical Audience
    - Sensitive versus Non-Sensitive Information
    - Follow Practices for Accessibility
  - Describe Data Source Considerations

- Accessing Source Data and Creating Reports
- Documenting the Source Data
- Determining Access to Data
- Developing Views of the Data
- Prepare for the Delivery of Reports and Dashboards
  - Determining How Visuals Will Be Viewed or Consumed
  - Determining How Data Will Be Delivered
  - Frequency of Reporting
  - Recurring Reports
- Using the Appropriate Type of Visualization
  - Use Basic Visuals
    - Visualize Data with the Right Charts
    - Pivots
    - Introduction to Pivot Tables
    - Lab: Use Pivot Tables to Analyze Fees and Cost
    - Advanced Pivot Table Features
    - Lab: Create and Modify Pivot Tables
    - Lab: Use Pivot Tables to Analyze Project Data
    - Creating and Manipulating Pivot Charts
    - Lab: Create and Modify Pivot Charts
    - Pie Charts and Treemaps
    - Column and Bar Charts
    - Line Graphs
    - Chart Types
    - Lab: Use Microsoft Excel Charts to Analyze Data
    - Lab: Display Data in Charts
    - Live Lab: Build Visuals to Make an Impact
  - Build Advanced Visuals
    - Stacked Charts
    - Building and Reading Stacked Charts
    - Line Graphs with Multiple Lines
    - Combination Charts
    - Dual Axis Charts
    - Scatter Plots and Bubble Charts
    - Histograms
    - Waterfall Charts

- Selecting Advanced Chart Types
- Advanced Chart Formatting
- Lab: Create and Modify Advanced Charts
- Build Maps with Geographical Data
  - Preparing Geo Fields for Mapping
  - Geographic Maps
  - Activity: Use Geographical Data
  - Live Lab: Apply Mapping Techniques with Geographical Data
- Use Visuals to Find or Tell a Story
  - Heat Maps
  - Infographics
  - Live Lab: Create Visuals to Tell a Story
- Designing Components for Reports and Dashboards
  - Choose Report and Dashboard Visuals
    - Selecting Different Visualization Layouts
    - Visualization Layouts
    - Mockup and Wireframing for Design
    - Types of Visuals
    - Types of Dashboard Navigation
  - Select Design Elements
    - Using Appropriate Design Components for Reports and Dashboards
    - Branding Guidelines
    - Appropriate Color Schemes
    - Appropriate Fonts and Layout
    - Labels
    - Live Lab: Apply Dashboard Design Tools to Enhance Visualizations
  - Use Standard and Written Reporting Elements
    - Standard Information and Formatting Elements
    - Best Practices for Using Reports and Dashboards
    - Activity: Utilize Standard Elements for Reports and Dashboards
    - Narrative
    - Other Supporting Materials
    - Applied Live Lab: Visualize Data
- Preparing for the Delivery and Consumption of Reports
  - Describe How Updates and Timing Affect Reporting
    - Static versus Dynamic Reports

- Point-In-Time Reporting
- Real-Time Reporting
- Using System Functions for Versioning
- Differentiate Between Types of Reports
  - Operational and Compliance Reports
  - Tactical and Research-Driven Reporting
  - Ad-Hoc Reporting
  - Live Lab: Build an Ad-Hoc Report
  - Self-Service Reporting
  - Activity: Determine Appropriate Report Types
- Prepare for Self-Service Capabilities
  - Deployment Considerations
  - Filtering Techniques for User Experience
  - Deploy to Production
- Summarizing the Importance of Data Governance
  - Maintain Compliance through Data Governance
    - The Importance of Data Governance
    - Data Governance
    - Roles Within a Data Governance Team
    - Jurisdiction Requirements
    - Regulations and Compliance
    - Data Classifications
  - Explain Access and Security Requirements
    - Data Use Agreements
    - Data Retention and Destruction Policies
    - Data Processing
    - Data Transmission
    - Data Encryption
    - Encrypting Data in Transit
    - Encrypting Data at Rest
  - Employ Privacy and Protection Practices
    - The Challenges of Data Discovery
    - Privacy and Security in Data
    - De-Identification
    - Masking Values
    - Pseudonymization

- Anonymization
- Data Masking
- Data Breaches
- Data Access and Role-Based Permissions
- Saving Data Files and Storage Types
- Live Lab: De-Identify Records
- Explaining Data Management Concepts
  - Explain the Basics of Data Management
    - Understanding Data Management
    - Master Data
    - Master Data Management Software
    - Reasons for Master Data Management
    - Using Warehouses for Data Management
  - Describe Data Management Processes
    - Field Standardization
    - Data Fields and Attributes
    - Data Integrity
    - Data Dictionary
    - Integration
    - Data Lineage
    - Entity Relationship Models
- Troubleshooting Issues and Measuring Performance
  - Managing Changes to Data, Databases, and Software
    - Developing a Change Management Process
    - Designing for Data Integrity
    - Data Validation
    - Troubleshooting Filtered Data
    - Common Issues
    - Addressing Issues Through Testing
    - Data Health Checks
    - Data Refresh Cycles and Archiving
  - Use Common Techniques for Query Optimization
    - Query Optimization
    - Activity: Analyze Queries to Increase Performance
    - Filtering with Parameters
    - Indexing Data

- Activity: Create Indexes to Increase Performance
- Temporary Tables
- Subquerying and Subsets of Information
- Applied Live Lab: Build Queries and Transform Data
- Monitor Database Performance
  - Resolving Locks
  - Data Corruption
  - Investigating Database Corruption
  - Slow Downs and Refresh Issues
  - Performance Tuning
  - Enable Logging and Monitoring Alerts
  - Live Lab: Ensure Data Integrity with Constraints
- Explain User Permissions and Backup Best Practices
  - The Principle of Least Privilege
  - Authentication
  - Identity and Access Management
  - Identity Access Management
  - Scheduling Backups
  - Replication and Snapshots
  - Disaster Recovery Documentation

## REQUIREMENTS:

Recommended experience: 18–24 months in a data analyst or similar job role, with exposure to databases, analytical tools, basic statistics, and data visualization 3–4 years in IT, inclusive of 2+ years hands-on cybersecurity; Security+, CySA+, PenTest+, or equivalent recommended.

## Difficulty level



## CERTIFICATE:

The participants will obtain certificates signed by CompTIA (course completion). This course will help prepare you for the CompTIA Data+ certification exam, which is available through the Pearson VUE test centers.

*Each participant in an authorized training CompTIA Data+ Prep Course held in Compendium CE will receive a free DA0-002 CompTIA Data+ Certification Exam vouchers.*

## TRAINER:

Authorized CompTIA Trainer