Caltech Center for Technology & Management Education

Data Analytics Bootcamp

Powered by simplilearn

Table of Contents

About the Program

Over the past decade, the data analytics field has witnessed a meteoric rise, transforming industries across the spectrum. The proliferation of digital data coupled with technological advancements has propelled this growth. The global market size for data analytics is burgeoning. According to Grand View Research reports, the global data analytics market size was valued at USD 116.4 billion in 2020 and is expected to expand at a CAGR of 24.9% from 2021 to 2028. This growth is driven by a surge in demand for real-time data analysis, the need to extract meaningful insights from big data, and the rising adoption of cloud-based analytics solutions across industries.

The career prospects in data analytics are promising. The U.S. Bureau of Labor Statistics projects a significant increase in demand for data analysts and related roles in the coming years. The data analyst job market is growing at a rate faster than the average for other occupations, offering competitive salaries and diverse career paths.

Take the next leap in your career by joining the Data Analytics Bootcamp, delivered in collaboration with Caltech CTME. This Data Analytics Bootcamp is a comprehensive program that will help learners fast-track their careers in the field of data analytics. This bootcamp will help learners master the in-demand skills needed for a successful career in the data analytics field. It covers key tools and concepts like Excel, SQL, programming fundamentals, mathematical and statistical computing, data manipulation, ETL, and data visualization with Tableau. Learn from experienced instructors through live interactive sessions and practical training through hands-on projects in this program.

Designed for professionals, you can complete this Data Analytics Bootcamp according to your schedule and build your skills through continuous learning sessions. With a mix of theory and practical sessions, this bootcamp is a great place to start your data analytics journey.

In the end, you will get the opportunity to leverage the skills and knowledge you gained in the bootcamp by working on a real-world capstone project to showcase your technical expertise to prospective employers.

Key Features of the Program

Data Analytics Bootcamp Certificate from Caltech CTME

Dedicated course with live sessions on generative AI and prompt engineering

Up to 13 CEUs from Caltech CTME

Practical exposure to ChatGPT, DALL-E, Midjourney, and other prominent tools

Caltech CTME Circle membership

23 Hands-on projects and 3 capstone projects with seamless access to integrated labs

Online Convocation by Caltech CTME Executive Director

Simplilearn Career Service to help you get noticed by top hiring companies

Masterclasses delivered by distinguished Caltech instructors

About Caltech CTME

Founded in 1891, Caltech is a world-renowned science and engineering institute that marshals some of the world's brightest minds and most innovative tools to address fundamental scientific questions and pressing societal challenges. Caltech prizes excellence and ambition. The contributions of Caltech's faculty and alumni have earned national and international recognition, including 38 Nobel Prizes and nearly 60 National Medals of Science. The Institute manages the Jet Propulsion Laboratory (JPL) for NASA.

CTME is embedded in Caltech's Division of Engineering and Applied Science. Caltech CTME has a unique role to play in applying the capabilities of scientists and engineers to the challenges of today's technology-driven businesses. Caltech CTME applies executive education and professional development directly to real-world problems. Caltech CTME experts teach the tools and perspectives that elevate careers and help companies achieve their goals.

About Simplilearn

Simplilearn is the world's #1 online bootcamp provider, enabling learners around the globe with rigorous and highly specialized training offered in partnership with world-renowned universities and leading corporations. We focus on emerging technologies and skills transforming the global economy, such as artificial intelligence, data science, cloud computing, programming, and more. Our hands-on and immersive training includes live virtual classes, integrated labs and projects, 24x7 support, and a collaborative learning environment. Over two million professionals and 2000 corporate training organizations across 150 countries have harnessed our award-winning programs to achieve their career and business goals.

Eligibility Criteria

For admission to this Data Analytics Bootcamp, candidates:

Are not required to have prior work experience

Must be at least 18 years old and have a high school diploma or equivalent

Don't need prior coding experience or technical know-how

Application Process

Candidates can apply to this bootcamp in 3 simple steps:

Submit an Application

Complete the application and include a brief statement of purpose. The latter informs our admissions counselors why you're interested and qualified for the bootcamp.

Application Review

A panel of admissions counselors will review your application and statement of purpose to determine whether you qualify for acceptance.

Admission

An offer of admission will be made to qualified candidates. You can accept this offer by paying the program fee.

Talk to an Admissions Counselor

Our dedicated admissions counselors are here to help guide you in the application process and related matters.

They are available to:

- Address questions related to the application
- Assist with financial aid (if required)
- Help you better understand the program and answer other questions

Who is this Bootcamp Ideal for?

This bootcamp caters to working professionals from a variety of industries and backgrounds. Aspiring professionals of any educational background with an analytical bent of mind are most suited to pursue this Data Analytics Bootcamp. We welcome professionals keen to develop expertise in data analytics, with the objective of:

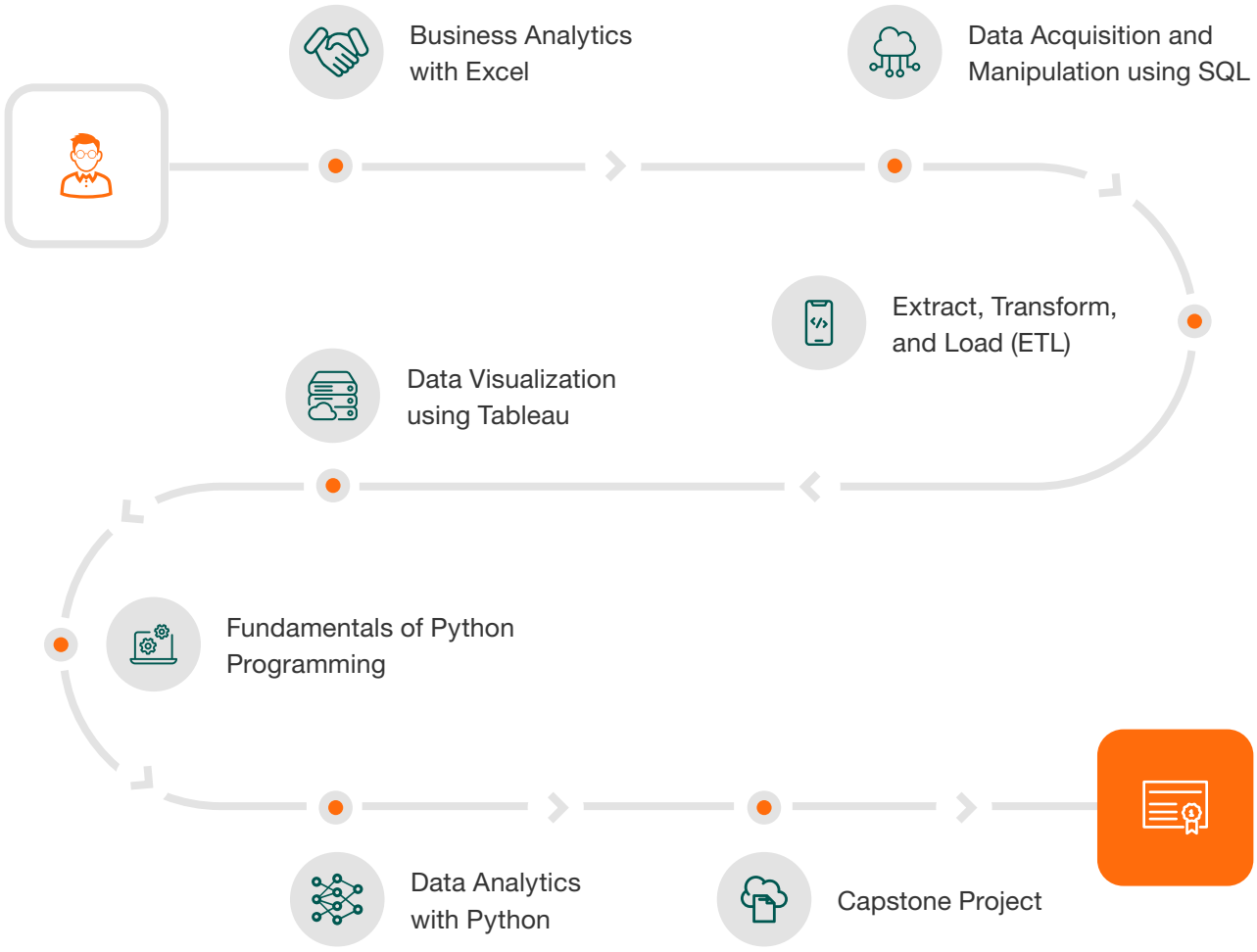
- Enhancing effectiveness in their current role
- Transitioning to data analytics roles in their organization
- Advancing their career in the data analytics field
- Giving shape to entrepreneurial aspirations

Program Outcomes

- Gaining proficiency in Excel for data analysis and decision-making
- Utilizing Excel functions and tools for business analytics
- Gaining proficiency in SQL for database management and querying
- Understanding database structures and relationships
- Mastering ETL processes for data extraction, transformation, and loading
- Implementing ETL tools and techniques for data integration
- Leveraging Tableau to create insightful data visualizations
- Building interactive dashboards for effective data presentation
- Understanding basic Python concepts and syntax
- Developing the ability to write and execute Python programs for data manipulation
- Applying Python libraries (pandas, NumPy, Matplotlib) for data analytics
- Performing statistical analysis and data manipulation with Python
- Developing practical skills through real-world data analysis projects
- Solving complex problems with data analytics tools and techniques
- Understanding AI fundamentals and generative AI models
- Mastering prompt engineering and understanding ChatGPT mechanisms
- Exploring generative AI's applications in various data analytics scenarios.
- Using generative AI to uncover insights and anomalies in data
- Developing proficiency in Power BI for data storytelling and visualization
- Creating interactive dashboards and reports for effective data communication
- Leveraging R programming for statistical analysis and data manipulation
- Using R for regression, classification, and data visualization
- Understanding ethical considerations in data analytics and AI
- Implementing data privacy, security, and ethical decision-making in data-related scenarios

Learning Path

Core Topics



Electives

- Essentials of Generative AI, Prompt Engineering, and ChatGPT
- Applications of Generative AI in Data Analytics
- Data Storytelling using Power BI
- Data Analytics with R Programming
- Data Ethics

Learning Path Details

Step 1

Business Analytics with Excel

Become proficient in business analytics with foundational statistics and techniques, an initial step in the Data Analytics Bootcamp.

Learning Outcomes

- Comprehend the significance of business analytics in the industry
- Master Excel analytics functions and conditional formatting
- Analyze intricate data sets using pivot tables and slicers
- Apply statistical tools (e.g., moving averages, hypothesis testing, ANOVA, regression) in Excel
- Visualize insights using charts and dashboards

Topics Covered

- Introduction to Business Analytics
- Conditional Formatting and Essential Functions
- Analysis with Pivot Tables
- Dashboard Creation
- Excel in Business Analytics
- Statistical Data Analysis
- Power BI

Step 2

Data Acquisition and Manipulation using SQL

Acquire essential skills to begin working with SQL databases. Learn about fundamental SQL concepts, including statements, conditional queries, commands, joins, subqueries, and diverse functions crucial for effective SQL database management and scalability.

Learning Outcomes

- Understand database structures and relationships
- Utilize common query tools and execute SQL commands
- Master transactions, table creation, and views
- Execute and comprehend stored procedures

Topics Covered

- Fundamental SQL Statements
- Database Restoration and Back-up
- Filtering with Selection Commands
- Ordering in Selection Commands
- Alias Implementation
- Application of Aggregate Commands
- Utilizing Group By Commands
- Conditional Statement Implementation
- Understanding Joins
- Execution of Subqueries
- Usage of Views and Index
- Implementing String Functions
- Mathematical Functions Application
- Date and Time Function Usage
- Utilizing Pattern (String) Matching
- Implementing User Access Control Functions

Step 3

Extract, Transform and Load (ETL)

Embark on a comprehensive exploration of ETL essentials crucial for robust data analysis. Learn about systematically extracting structured/unstructured data, rule formulation, and utilizing diverse ETL tools like Nifi and Talend. Develop repository data loading expertise, distinguish between batch and real-time ETL, apply real-time processing and optimize with parallel techniques, and efficiently manage ETL workflows for seamless data analysis.

Learning Outcomes:

- Define ETL's role in data management and analysis, emphasizing its significance in integration & transformation
- Identify and access diverse data sources (databases, APIs), understanding source connections and configurations
- Extract structured/unstructured data employing incremental and full extraction techniques
- Perform data cleansing, validation, normalization, and aggregation to handle quality issues effectively
- Create transformation rules and mappings for varied data types, managing schema changes
- Work with ETL platforms like Apache Nifi and Talend, using specific features
- Load transformed data efficiently into target repositories, optimizing loading for performance
- Differentiate between batch and real-time ETL and implement real-time processing tools
- Develop strategies for data quality issues and implement logging and monitoring mechanisms
- Optimize ETL processes for enhanced speed and efficiency and learn parallel processing techniques
- Implement automation to schedule and monitor ETL jobs for seamless execution
- Set up systems to monitor ETL process health and learn maintenance and version control best practices

Topics Covered

- ETL Fundamentals
- Data Source Identification
- Data Extraction
- Data Transformation
- Data Mapping and Conversion
- ETL Tools and Technologies
- Data Loading
- Batch and Real-time ETL
- Error Handling and Logging
- Performance Optimization
- Automation
- Monitoring and Maintenance

Step 4

Data Visualization using Tableau

This Tableau course will give a comprehensive understanding of crafting impactful visualizations, efficient data organization, and designing informative charts and dashboards to empower better business decisions. Explore data visualization concepts and diverse chart creation, and master constructing interactive dashboards by working with filters, parameters, and sets. Learn how to use various visualization techniques like heat maps, treemaps, and Pareto charts.

Learning Outcomes

- Acquire expertise in visualization techniques, including heat maps, treemaps, waterfalls, and Pareto charts
- Understand metadata's significance and its application in Tableau
- Manipulate data effectively using filters, parameters, and sets
- Master specific field types, Tableau-generated fields, and creating and utilizing parameters
- Learn how to create various charts, interactive dashboards, captivating story interfaces, and effective data-sharing methods
- Gain proficiency in data blending, creating extracts, and organizing/formatting data effectively
- Master various calculations, including arithmetic, logical, table, and level of detail (LOD)

Topics Covered

- Data Visualization
- Introduction to Tableau
- Tableau Workspace and Chart Types
- Chart Creation and Data Preparation
- Preparation Techniques
- Filters and Analytics in Tableau
- Tableau Dashboards

Step 5

Fundamentals of Python Programming

This course focuses on building fundamental Python skills crucial for your progression in the bootcamp.

Learning Outcomes

- Attain proficiency in procedural and object-oriented programming
- Understand Python's benefits and practical applications
- Set up Python and Jupyter Notebook and grasp their utility
- Implement Python basics, including identifiers, indentations, and comments
- Master Python data types, operators, and string functions
- Explore Python loop types and variable scope within functions
- Grasp object-oriented programming principles and attributes in Python
- Gain insight into multi-threading in Python

Topics Covered

- Programming Fundamentals
- Introduction to Python
- Python Data Types and Operators
- Conditional Statements and Loops in Python
- Python Functions
- Object-Oriented Programming in Python
- Threading in Python

Step 6

Data Analytics with Python

Learning Outcomes

- Develop proficiency in retrieving and managing data through Application Programming Interfaces (APIs) for analytical purposes
- Utilize SQL and Python for robust database management, ensuring seamless data handling and analysis
- Comprehend foundational statistical principles and theories critical for data analysis
- Apply various probability distributions as analytical tools for insightful data exploration and interpretation
- Employing sophisticated statistical methodologies for in-depth data analysis and interpretation
- Get practical experience in testing hypotheses and drawing conclusions from data samples
- Choose and utilize correlation and regression analysis techniques effectively
- Analyze and interpret time-based data patterns and trends
- Conduct ANOVA and Chi-Square tests for varied data sets and deriving meaningful insights

Topics Covered

- Data Acquisition with APIs
- Database Management with SQL and Python
- Statistics Fundamentals
- Probability Distributions
- Advanced Statistics
- Statistical Analysis with Python
- Hypothesis Testing
- Correlation and Regression Analysis
- Time Series Analysis
- ANOVA and Chi-Square Tests

Step 7

Capstone Project

The capstone project culminates your immersive learning journey, delving into business analytics, data manipulation, visualization, and programming. Apply your new skills in Excel, Python, SQL, Tableau, and more to a real-world project by solving data-related challenges. Showcase your proficiency in ETL processes, visualization techniques, and ethical considerations. Choose between elective courses to further enrich your project, exploring generative AI, Power BI storytelling, R programming, or diving deeper into data ethics. This comprehensive project allows you to showcase expertise, creativity, and strategic thinking in solving complex data analytics problems.

Electives

Essentials of Generative AI, Prompt Engineering & ChatGPT

This course offers an in-depth exploration of generative AI models, particularly focusing on ChatGPT. Participants will acquire comprehensive knowledge of foundational generative AI principles, including prompt engineering, explainable AI, conversational AI, ChatGPT, and other extensive language models.

Learning Outcomes

- Build a solid understanding of diverse generative AI models, encompassing their core principles and variations
- Grasp the importance of explainable AI, exploring different strategies for ensuring transparency in AI systems
- Apply effective prompt engineering techniques to enhance performance and manage generative AI model behavior
- Develop a deep comprehension of ChatGPT, investigating its operational mechanisms, notable features, and limitations
- Explore various applications and contexts where ChatGPT can be utilized effectively
- Acquire familiarity with fine-tuning methods to personalize and refine ChatGPT models
- Recognize the ethical complexities of generative AI models, ensuring responsible data usage, bias mitigation, and misuse prevention
- Gain insights into the transformative potential of generative AI across industries while evaluating leading generative AI tools

Topics Covered

- Overview of Generative AI and its Landscape
- Understanding Explainable AI
- Exploring Conversational AI
- Fundamentals of Prompt Engineering
- Designing and Creating Effective Prompts
- Extensive Language Models
- ChatGPT: Applications and Functionality
- Fine-tuning Strategies for ChatGPT
- Ethical Considerations in Generative AI Models
- Responsible Data Usage and Privacy Measures
- AI Technologies Driving Innovation

Applications of Generative AI in Data Analytics

Uncover generative AI's crucial role in data analytics by understanding how it generates synthetic data, enhances visualizations, and discovers hidden insights. Leverage its potential for ETL optimization, accurate forecasting, and addressing integration challenges. Explore future trends and ethical considerations, gaining a holistic view of generative AI's impact. The curriculum spans data augmentation, visualization, and forecasting, offering a comprehensive understanding of how generative AI intersects with data analytics.

Learning Outcomes

- Understand the fundamental concepts and importance of generative AI within modern data analytics
- Employ generative AI techniques to address data scarcity and generate synthetic data for testing models, privacy-preserving analysis, and data augmentation
- Utilize generative AI tools to create tailored, interactive, and accessible data visualizations
- Apply generative AI to discover hidden patterns, identify anomalies, and extract insights from data exploration
- Optimize ETL processes and data pipelines by leveraging AI for automated data cleaning, intelligent transformation, and real-time analysis
- Enhance data modeling and forecasting accuracy by implementing generative AI to build predictive models and simulate different scenarios
- Recognize and address challenges associated with integrating generative AI into data projects, considering computational constraints, ethical implications, and model robustness
- Explore future trends and evolving directions in generative AI concerning quantum computing, transfer learning, multi-modal data analysis, and

Topics Covered

- Overview of Generative AI in Data Analytics
- Data Augmentation using Generative AI
- Generative AI for Tailored Data Visualization
- Generative AI in Data Exploration
- AI-Optimized ETL Processes
- Generative AI in Data Modeling and Forecasting
- Challenges in Integrating Generative AI
- Future Directions in Generative AI

Data Storytelling using Power BI

Opt for this elective course to master insightful data analysis and interactive dashboard creation. Learn how to harness Power BI's potential to address business challenges and streamline operations. The course delves into crafting dashboards from published reports, extracting valuable insights via Quick Insights, and practical methodologies across Power BI functions, spanning data collection to analysis. Additionally, it offers troubleshooting strategies for various Power BI issues.

Learning Outcomes

- Craft dynamic dashboards from reports, boosting data visualization and interactivity
- Generate visuals and dashboards using Quick Insights to extract key insights from data
- Utilize natural language in Q&A for actionable visual generation
- Create and manage data alerts for timely data updates
- Master report layout best practices and data visualization for impactful reports
- Optimize chart selection based on contextual questions or narratives
- Incorporate shapes for emphasis and narrative enhancement in reports
- Integrate custom visuals tailored to specific requirements
- Explore different methods to share reports and dashboards efficiently
- Complete end-to-end Power BI projects for comprehensive data analysis and visualization
- Foster team collaboration through Microsoft Teams for seamless communication and sharing

Topics Covered

- Efficient Data Retrieval and Preparation Techniques
- Proficient Data Management
- Interactive Report and Dashboard Creation
- Efficiency-enhancing Power BI Tips and Tricks

Data Analytics with R

R is a robust language for data science and analytics, boasting a vibrant community despite its steep learning curve. Its burgeoning popularity makes it the preferred choice for organizations seeking analytical prowess for competitive advantages

Learning Outcomes

- Establish a foundational understanding of business and data analytics
- Set up R, RStudio, and R Workspace and explore various R packages
- Gain proficiency in R programming and executing diverse statements in R
- Comprehend R data structures and data import/ export functions
- Implement Apply and DPLYR functions
- Leverage R graphics for effective data visualization
- Understand essential statistical concepts
- Perform hypothesis testing to enable informed business decision-making
- Use linear and non-linear regression models, along with classification techniques for data analysis
- Understand and apply association rules and the Apriori algorithm
- Master clustering methods like K-means, DBSCAN, and hierarchical clustering

Topics Covered

- Introduction to Business Analytics
- Introduction to R Programming
- Data Structures
- Data Visualization
- Statistics for Data Science I
- Statistics for Data Science II
- Regression Analysis
- Classification
- Clustering
- Association

Data Ethics

This course provides a comprehensive understanding of the ethical considerations and responsibilities in the realm of data analytics. Participants will delve into the multifaceted landscape of data ethics, exploring its relevance, legal frameworks, privacy concerns, security protocols, biases, decision-making frameworks, visualization practices, and the societal impacts of data analytics.

Learning Outcomes:

- Define the concept of data ethics in the context of data analytics
- Recognize the relevance and significance of data ethics in the field of data analytics
- Familiarize yourself with data protection laws and regulations like GDPR and CCPA
- Understand the legal obligations concerning data handling and privacy
- Comprehend the principles of data privacy and the significance of informed consent
- Analyze the ethical implications linked to data collection, storage, and sharing
- Recognize the significance of maintaining data security and integrity
- Learn techniques for safeguarding data from unauthorized access and breaches
- Explain the importance of transparency within data analytics processes
- Explore mechanisms for ensuring accountability in data analysis procedures
- Understand the concepts of bias and fairness in the context of data analytics
- Discuss strategies for identifying and mitigating bias within data and algorithms.
- Apply ethical decision-making frameworks to scenarios within data analytics
- Utilize tools to address the ethical dilemmas encountered in real-world situations
- Implement ethical practices in data visualization and communication
- Discuss responsible methods for presenting data insights to avoid misleading interpretations
- Investigate ethical challenges during development and deployment of ML and AI models
- Examine fairness, transparency, and accountability in algorithmic decision-making
- Explore the social and ethical implications of data analytics on society
- Evaluate issues such as discrimination, surveillance, and profiling stemming from data analytics

Topics Covered

- Introduction to Data Ethics
- Legal and Regulatory Frameworks
- Data Privacy and Consent
- Data Handling and Security
- Transparency and Accountability
- Bias and Fairness in Data Analysis
- Ethical Decision-Making
- Data Visualization and Communication
- Ethics in Machine Learning and AI
- Social and Ethical Impacts

Skills Covered

- Generative AI
- Data Ethics
- Data Analytics
- SQL
- ETL
- Visualization
- Python
- Statistics
- Statistical Analysis using Excel
- Data Analytics using R and Python
- Linear and Logistic Regression
- Clustering using K-means
- Supervised Learning
- Data Visualization with Tableau and Power BI

Tools Covered

SQL

Excel

Python

Scikit-learn

NumPy

Pandas

Tableau

R

Power BI

Certificate

Upon completing the Caltech Data Analytics Bootcamp, you will receive a Certificate of Completion from Caltech CTME. This certificate will testify to your skills as a data analytics professional.

Program Advisors

Rick Hefner

Executive Director,
Caltech Center for Technology & Management Education

Rick Hefner, PhD, specializes in systems development and maintenance; project management; Lean Six Sigma; process improvement, technology transfer; and risk management. His experience spans over 35 years. Dr. Hefner recently served as Director of Process Management at Northrop Grumman Corporation, where he managed corporate process initiatives related to Lean Six Sigma and program management. Previous positions at Northrop Grumman (formerly TRW) included managing technology process initiatives and helping to establish the corporate engineering and program management processes. Previously, at Aerospace Corporation, Dr. Hefner was the Director of their Software Development department. He served as an engineer, technical specialist, project manager, and section manager. Dr. Hefner has also worked with companies in the communications, electronics, and health sciences industries, including Applied Physics Laboratory, Ares Management, Boeing, DRS Technologies, Herbalife, Honeywell, Jet Propulsion Laboratory, John Deere, L-3 WESCAM, Maytag, Motorola, Pacific Bell, Raytheon, Schlumberger, Southern California Edison, St. Jude Medical, Toshiba, U.S. Navy, and Xerox. Dr. Hefner is credited with over 200 publications and presentations. He earned his PhD from the University of California, Los Angeles, in applied dynamic systems control. He received his MS and BS from Purdue University in interdisciplinary engineering

Caltech Center for Technology & Management Education

simplilearn

USA

Simplilearn Americas, Inc.

201 Spear Street, Suite 1100, San Francisco,

CA 94105, United States

Phone No: +1-844-532-7688

INDIA

Simplilearn Solutions Pvt Ltd.

53/1 C, Manoj Arcade, 24th Main, Harlkunte

2nd Sector, HSR Layout

Bangalore - 560102

Call us at: 1800-212-7688

www.simplilearn.com