Caltech Center for Technology & Management Education

Post Graduate Program in
Data Science

Understand **Generative AI** and its potential

Leverage **Caltech CTME's Academic Excellence**

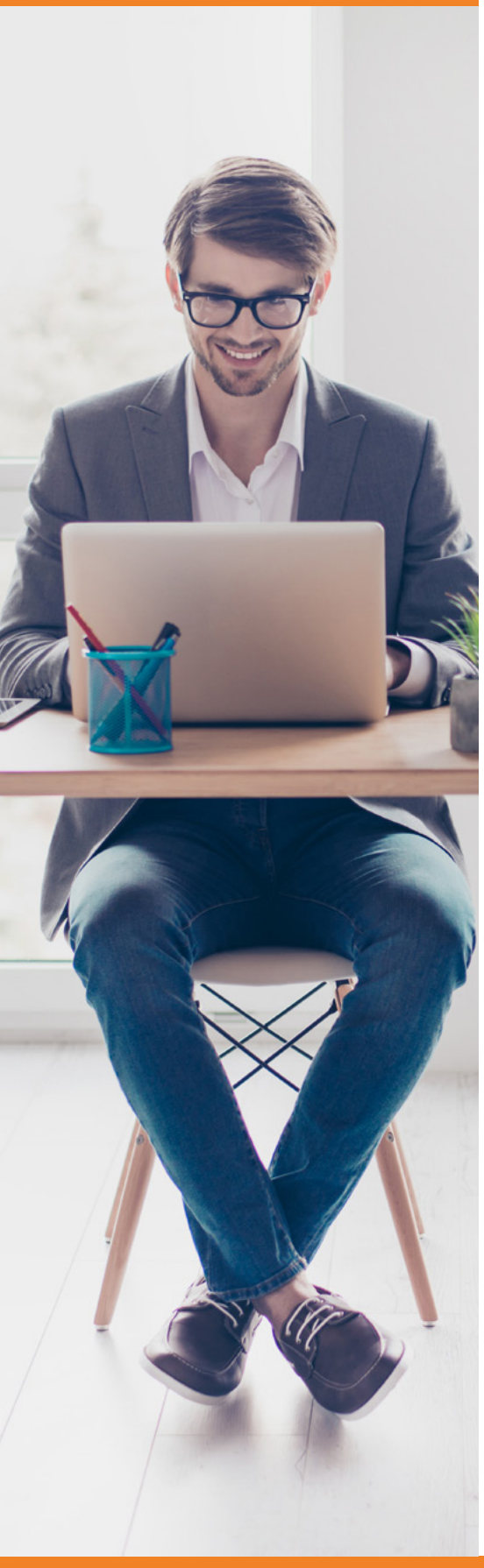
Exposure to prominent tools:

Bard | DALL·E 2 | ChatGPT | Midjourney | and more

Powered by simplilearn

Table of Contents

About the Program



Accelerate your career with the acclaimed Post Graduate Program in Data Science, delivered in collaboration with IBM. This program features the perfect mix of theory, casestudies and extensive hands-on practice to master the data science concepts and tools. It provides comprehensive training on data science, leveraging Caltech CTME's academic excellence and IBM's industry prowess.

Designed for working professionals, this post graduate program provides a deep dive into data science through a blend of online self-paced videos, live virtual classes, hands-on projects, and integrated labs, with mentorship sessions to provide a high-engagement learning experience. This program offers in-depth exposure to various tools and technologies to prepare you for an exciting career in data science.

- Excel
- MySQL
- python
- Machine Learning
- tableau
- Generative AI
- Prompt Engineering
- Explainable AI
- ChatGPT
- Power BI

Key Features of the Program

- Earn a program completion certificate from Caltech CTME
- Earn up to 14 CEUs from Caltech CTME
- Dedicated courses and live classes on generative AI, prompt engineering, and much more
- Gain exposure to ChatGPT, OpenAI, Dall-E, Midjourney and other prominent tools
- Masterclasses delivered by distinguished Caltech instructors and IBM experts
- Gain Caltech CTME Circle membership
- Online convocation with the Caltech CTME Executive Director
- Curriculum delivered in live online sessions by industry experts
- IBM certificates for IBM courses
- Access to hackathons and Ask Me Anything sessions from IBM
- 25+ hands-on projects with seamless access to integrated labs
- Industry-relevant capstone projects in 3 domains
- Simplilearn Career Service to help you get noticed by top hiring companies

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 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.

Eligibility Criteria

For admission to this program, candidates should have:

- A bachelor's degree with an average score of at least 50 percent
- Prior knowledge or experience in programming and mathematics
- 2+ years of formal work experience (preferred)

Application Process

Candidates can apply to the Caltech Data Science Program in 3 simple steps:

STEP 1

Submit an Application

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

STEP 2

Application Review

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

STEP 3

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

We have a team of dedicated admissions counselors 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 your questions

Who is this Program Ideal for?

This program caters to working professionals from a variety of industries and backgrounds. The diversity of our students adds richness to class discussions and interactions. A role in data science requires an amalgam of experience, data science knowledge, and the use of the correct tools and technologies. It is a solid career choice for professionals seeking to transition into the data science field. Aspiring professionals of any educational background with an analytical bent of mind are most suited to pursue this Post Graduate Program in Data Science.

Professionals keen to develop expertise in data science, with the objective of:

- Enhancing effectiveness in their current role
- Transitioning to data science roles in their organization
- Seeking to advance their career in the industry
- Giving shape to entrepreneurial aspirations

Program Outcomes

Apply effective prompt engineering techniques to improve the performance and control the behavior of generative AI models

Learn about prompt engineering, explainable AI, conversational AI, large language models, ChatGPT, and much more

Understand generative AI, its landscape, and practical applications

Grasp fundamentals of Excel analytics functions and conditional formatting

Understand the tools and techniques used in business analysis planning and monitoring

Gain an in-depth understanding of data structure and data manipulation

Understand and use linear and non-linear regression models and classification techniques for data analysis

Obtain a comprehensive knowledge of supervised and unsupervised learning models such as linear regression, logistic regression, clustering, dimensionality reduction, K-NN, and pipelines

Perform scientific and technical computing using the SciPy package and its sub-packages such as Integrate, Optimize, Statistics, IO, and Weave

Master the concepts of recommendation engines and time series modeling

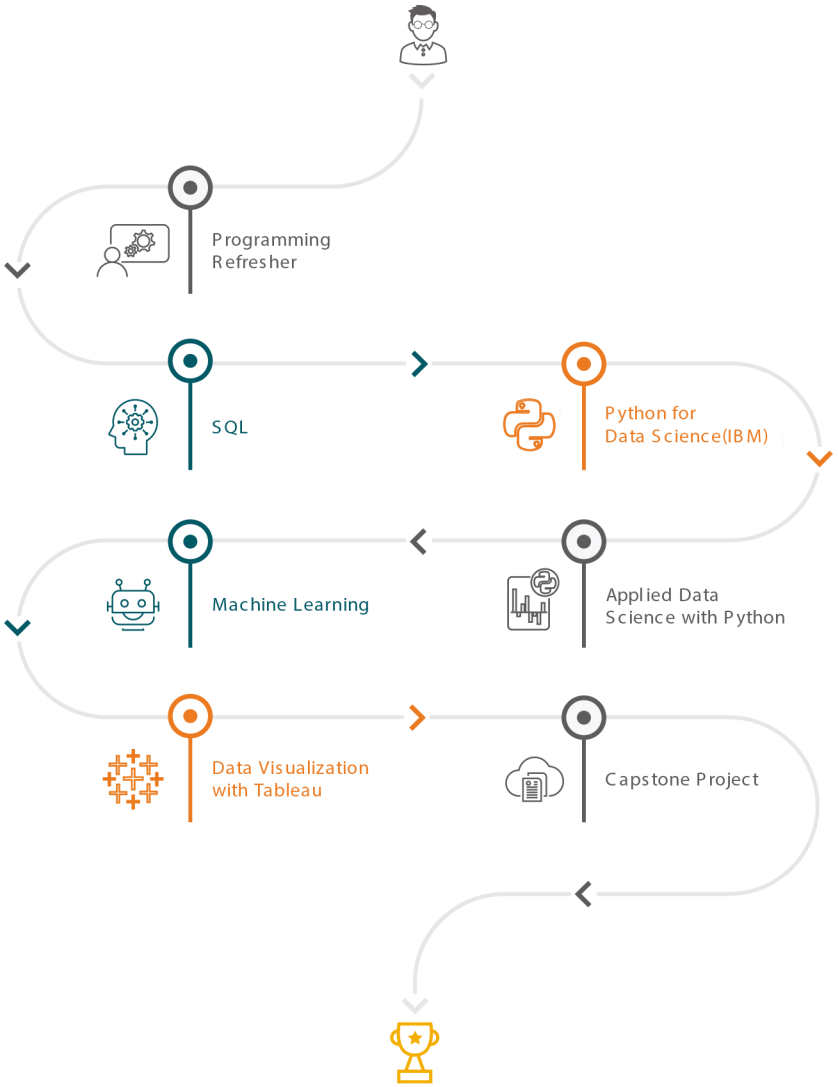
Gain practical mastery over principles, algorithms, and applications of machine learning

Gain expertise in mathematical computing using the NumPy and scikit-Learn package

Learn to visualize data using Tableau and Power BI and become proficient in building interactive dashboards

Learning Path

Core Courses



Electives

- Business Analytics with Excel
- Data Storytelling using Power BI
- Essentials of Generative AI, Prompt Engineering & ChatGPT
- R Programming for Data Science (IBM)
- Academic Masterclass
- Industry Masterclass

STEP

1

2

3

4

5

6

7

Programming Refresher

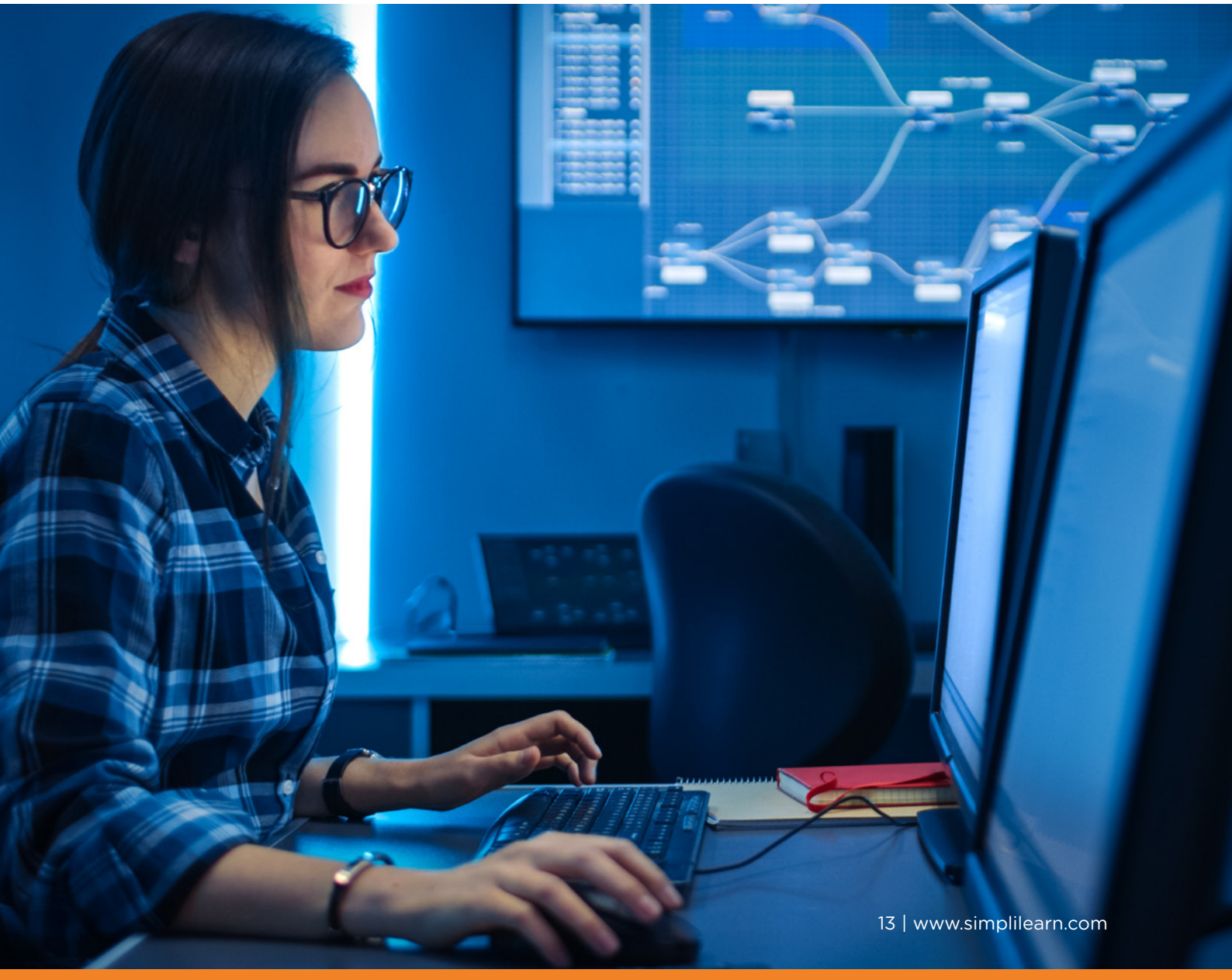
In this course, you will develop essential Python skills that are a crucial foundation for your progress throughout the program.

Learning Outcomes:

- Acquire proficiency in both procedural and object-oriented programming
- Recognize the advantages and benefits of using Python as a programming language
- Install Python and its integrated development environment
- Become familiar with Jupyter Notebook and its practical applications
- Implement Python identifiers, indentations, and comments effectively
- Understand Python's data types, operators, and string functions
- Learn about the various types of loops in Python
- Explore the concept of variable scope within functions
- Explain the principles and characteristics of object-oriented programming
- Describe methods, attributes, and access modifiers in Python
- Gain a solid understanding of multi-threading

Topics Covered

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



STEP

1

2

3

4

5

6

7

SQL

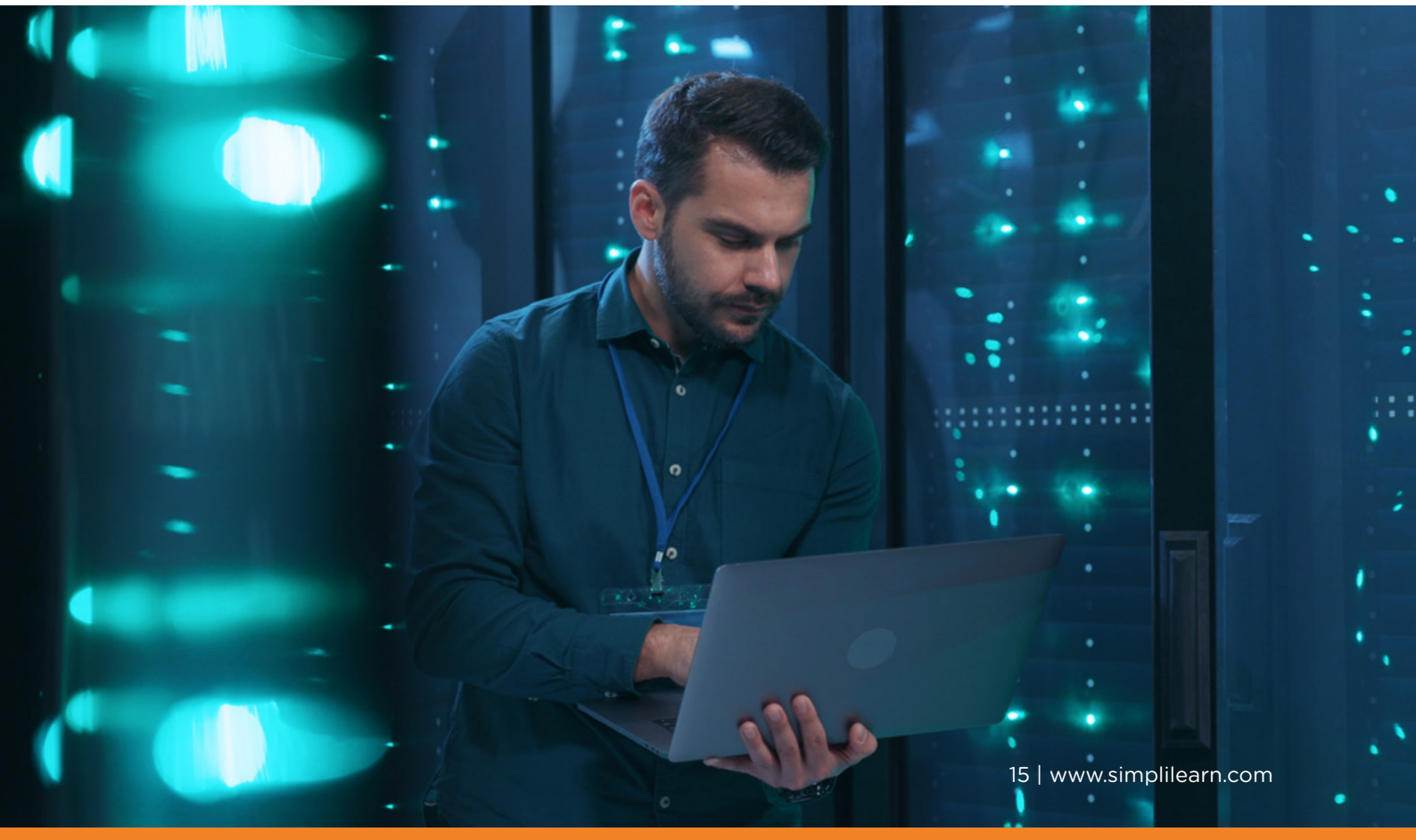
Enroll in this course to acquire essential knowledge for effectively working with SQL databases and leveraging them in your applications. Throughout the course, you will gain a solid understanding of SQL statements, conditional statements, commands, joins, subqueries, and various functions, empowering you to manage your SQL database for scalable growth.

Learning Outcomes:

- Develop a comprehensive understanding of databases and their relationships
- Learn how to use common query tools and work with SQL commands
- Master transactions, table creation, and views for efficient database management
- Comprehend and execute stored procedures to perform complex operations
- Acquire expertise in various SQL lessons, including filtering, ordering, aliasing, aggregate commands, grouping, conditional statements, joins, subqueries, views, and indexing
- Explore different SQL functions such as string, mathematical, date and time, and pattern-matching functions
- Understand user access control functions to ensure database security

Topics Covered

- SQL Statements
- Restore and Back-up
- Selection Commands - Filtering & Ordering
- Aggregate Commands
- Group By Commands
- Conditional Statements
- Joins
- String Functions
- Mathematical Functions
- Date and Time Functions
- Pattern (String) Matching
- User Access Control Functions



STEP

1

2

3

4

5

6

7

Python for Data Science (IBM)

The course provides the necessary skills to leverage Python for data science. By the course's conclusion, participants will be proficient in writing Python scripts and performing hands-on data analysis using a Jupyter-based lab environment.

Learning Outcomes:

- Create your first Python program using variables, strings, functions, loops, and conditions
- Understand and apply concepts related to lists, sets, dictionaries, conditions, branching, objects, and classes in Python
- Utilize the pandas library to load, manipulate, and save data, and read and write files in Python

Topics Covered

- Python Basics
- Python Data Structures
- Python Programming Fundamentals
- Working with Data in Python
- Working with NumPy Arrays

STEP

1

2

3

4

5

6

7

Applied Data Science with Python

This course comprehensively explores data science essentials, covering data preparation, model building, and evaluation. Participants will delve into fundamental Python concepts, including strings, Lambda functions, and lists. Additionally, they will gain proficiency in essential topics such as NumPy, linear algebra, and statistical concepts, which include measures of central tendency and dispersion, skewness, covariance, and correlation. The course also encompasses hypothesis testing techniques such as Z-test, T-test, ANOVA, and data manipulation using pandas. Participants will enhance their data visualization skills using popular libraries like Matplotlib, Seaborn, Plotly, and Bokeh.

Learning Outcomes:

- Explain the fundamentals of data science and its practical applications
- Explore the processes of data preparation, model building, and evaluation
- Apply Python concepts related to strings, Lambda functions, and lists comprehensively
- Develop a strong understanding of NumPy and its applications, including array indexing and slicing techniques
- Apply principles of linear algebra in data analysis, including its application in calculus
- Calculate measures of central tendency and dispersion in data
- Gain a clear understanding of statistical concepts like skewness, covariance, and correlation
- Describe the null hypothesis and alternative hypothesis in hypothesis testing

- Examine different hypothesis tests, including Z-test, T-test, and ANOVA
- Understand the concept of ANOVA for statistical analysis
- Work with pandas' two primary data structures: Series and DataFrame.
- Utilize pandas for data loading, indexing, reindexing, and data merging
- Prepare, format, normalize, and standardize data using data binning techniques
- Create effective visualizations using Matplotlib, Seaborn, Plotly, and Bokeh

Topics Covered

- Introduction to Data Science
- Essentials of Python Programming
- NumPy
- Linear Algebra
- Statistics Fundamentals
- Probability Distributions
- Advanced Statistics
- Working with pandas
- Data Analysis
- Data Wrangling
- Data Visualization
- End-to-End Statistics Applications in Python

STEP

1

2

3

4

5

6

7

Machine Learning

This course comprehensively explores diverse machine learning types and their practical applications. Participants will gain insights into the machine learning pipeline, delving into supervised learning, regression models, and classification algorithms. Additionally, the course covers unsupervised learning, clustering techniques, and ensemble modeling. Participants will also evaluate well-known machine learning frameworks like TensorFlow and Keras and get hands-on experience building a recommendation engine using PyTorch.

Learning Outcomes:

- Examine various types of machine learning and understand their unique characteristics
- Analyze the machine learning pipeline and gain a comprehensive understanding of key operations involved in machine learning operations (MLOps)
- Explore supervised learning and its wide range of applications
- Understand the concepts of overfitting and underfitting and learn techniques to detect and prevent them
- Analyze different regression models and identify their suitability for specific scenarios
- Identify linearity between variables and create correlation maps
- List various types of classification algorithms and comprehend their specific applications
- Master various types of unsupervised learning methods and determine their appropriate use
- Gain a deep understanding of different clustering techniques within unsupervised learning

- Examine different ensemble modeling techniques, such as bagging, boosting, and stacking
- Evaluate and compare different machine learning frameworks, including TensorFlow and Keras
- Build a recommendation engine using PyTorch

Topics Covered

- Machine Learning
- Supervised Learning
- Regression and its Applications
- Classification and its Applications
- Unsupervised Learning
- Ensemble Learning
- Recommendation Systems

STEP

1

2

3

4

5

6

7

Data Visualization with Tableau

This Tableau course gives participants a comprehensive understanding of creating captivating visualizations, optimizing data organization, and designing informative charts and dashboards to facilitate informed business decisions. Throughout the program, you will delve into data visualization concepts, gain proficiency in crafting diverse combo charts and stories, and learn to work with filters, parameters, and sets. Additionally, you will master the art of constructing interactive dashboards.

Learning Outcomes:

- Acquire expertise in various visualization techniques, such as heat maps, treemaps, waterfall charts, and Pareto charts
- Understand the importance of metadata and its application in Tableau
- Skillfully work with filters, parameters, and sets to manipulate data effectively
- Become proficient in utilizing special field types and Tableau-generated fields and creating and employing parameters
- Learn how to construct different charts, interactive dashboards, and captivating story interfaces and how to share insights effectively
- Gain proficiency in data blending, creating data extracts, and efficiently organizing and formatting data
- Master various calculations, including arithmetic, logical, table, and level of detail (LOD)

Topics Covered

- Introduction to Data Visualization
- Overview of Tableau
- Tableau Workspace and Different Chart Types
- Creating Charts and Data Preparation Techniques
- Effective Data Preparation Techniques
- Working with Filters and Analytics in Tableau
- Building Dashboards in Tableau

STEP

1

2

3

4

5

6

7

Capstone Project

The data science capstone project provides a valuable opportunity to apply the skills you have acquired during this program. With the guidance of dedicated mentors, you will address a real-world data science problem aligned with industry standards. This comprehensive project covers data processing, model building, and reporting business results and insights. It is the final step in your learning journey and enables you to demonstrate your data science expertise to potential employers.

Key learning objectives:

- Data Processing: Utilizing various techniques to transform raw data into meaningful insights
- Model Building: Employing techniques such as regression and decision trees to create accurate and intelligent machine learning models capable of making predictions
- Python or SAS: Developing your model and conducting a complete model-building exercise, including data splitting, testing, and validating data using the k-fold cross-validation process
- Model Fine-tuning: Applying various techniques to enhance the model's accuracy and selecting the best-performing champion model
- Dashboarding and Result Presentation: Using Tableau to create a dashboard with meaningful insights to present your final results

Through this project, you can showcase your practical data science skills and reinforce your understanding of the entire data science process. As a result, you will be well-prepared to impress potential employers with your expertise.

Electives

Business Analytics with Excel

Enroll in this course to gain practical, data-driven decision-making skills by mastering data analysis and statistics. Leveraging the widely-used office tool Excel, you will acquire the expertise to perform sophisticated data analytics, empowering you to make informed business decisions confidently.

Learning Outcomes:

- Understand the importance of business analytics and its role in various industries
- Grasp the fundamentals of Excel analytics functions and conditional formatting
- Learn how to analyze complex data sets efficiently using pivot tables and slicers
- Solve stochastic and deterministic analytical problems using Excel's powerful tools, including Scenario Manager, Solver, and Goal Seek
- Apply statistical tools and concepts like moving averages, hypothesis testing, ANOVA, and regression to data sets using Excel
- Effectively represent your findings using charts and dashboards
- Get introduced to the latest Microsoft analytics tools and technologies

Topics Covered

- Introduction to CBAP certification
- Introduction to Business Analytics
- Formatting and Conditional Formatting in Excel
- Important Functions in Excel
- Analyzing Data with Pivot Tables
- Dashboarding
- Business Analytics with Excel
- Data Analysis Using Statistics

By completing this course, you will gain valuable data analysis and statistical skills using Excel, preparing you to make data-driven decisions that can significantly impact your professional endeavors.

Electives

Data Storytelling using PowerBI

Microsoft Power BI offers robust tools to analyze data and extract valuable business insights through interactive dashboards. This comprehensive training course on Power BI empowers you to fully harness its potential, enabling you to solve business challenges and improve operations effectively. Throughout the course, you'll learn to expertly develop dashboards from published reports, utilize Quick Insights to discover valuable patterns rapidly, and adopt practical approaches for various tasks performed within Power BI, from data gathering to in-depth analysis. Additionally, the course provides helpful troubleshooting techniques to address various issues that may arise while using Power BI.

Learning Outcomes:

- Create dynamic dashboards from published reports, enhancing data visualization and interactivity
- Rapidly generate visuals and dashboards with quick insights to gain valuable insights from your data
- Utilize natural language in the Q&A feature to generate visuals for actionable insights
- Create and manage data alerts to stay informed of important changes in your data
- Learn best practices for report layout and data visualization to maximize the impact of your reports
- Understand when to use specific charts or graphs based on the questions you're addressing or the story you're presenting
- Incorporate shapes into your reports to design and emphasize key elements to create narratives

- Integrate custom visuals into your reports and dashboards to tailor them to your specific needs
- Share reports and dashboards effectively, understanding the pros and cons of different sharing methods
- Complete a comprehensive Power BI data analysis and visualization project from start to finish
- Enhance team collaboration using Microsoft Teams to facilitate smooth communication and sharing

Topics Covered

- Data Retrieval and Preparation Techniques for Efficient Analysis
- Developing Proficiency in Data Management
- Generating Interactive Reports and Dashboards
- Tips and Tricks for Efficient Power BI Usage

By completing this course, you will gain a strong command of Power BI, enabling you to analyze data effectively, create compelling visualizations, and derive valuable insights to make informed business decisions.

Electives

Essentials of Generative AI, Prompt Engineering & ChatGPT

In this course, participants will comprehensively study generative AI models, particularly ChatGPT. The curriculum covers essential principles of generative AI, prompt engineering, explainable AI, conversational AI, ChatGPT, and other large language models.

Learning Outcomes:

- Build strong foundations in generative AI models, understanding their core principles and various types
- Recognize the importance of explainable AI and distinguish between different approaches to achieve transparency in AI systems
- Apply effective prompt engineering techniques to optimize performance and control the behavior of generative AI models
- Develop a comprehensive understanding of ChatGPT, exploring its operational mechanisms, notable features, and limitations
- Explore diverse applications and scenarios where ChatGPT can be effectively utilized
- Master fine-tuning techniques to personalize and enhance ChatGPT models
- Address the ethical challenges of generative AI models, ensuring responsible data usage, mitigating bias, and preventing misuse
- Gain insights into the transformative potential of generative AI across various industries and examine prominent generative AI tools
- Dive into the future of generative AI, identifying challenges and the necessary steps to unlock its full potential

Topics Covered

- Introduction to Generative AI and its Landscape
- Explainable AI: Understanding the Importance of Transparency
- Conversational AI: Applications and Advancements
- The Art of Prompt Engineering
- Designing and Generating Effective Prompts for AI models
- Exploring Large Language Models
- In-depth Analysis of ChatGPT and its Applications
- Fine-tuning Techniques for Personalization
- Ethical Considerations in Generative AI Models: Responsible Data Usage and Privacy
- Exploring the Future of Generative AI
- Leveraging AI Technologies for Innovation

Upon completing this course, you will possess comprehensive knowledge of generative AI, focusing on ChatGPT, enabling you to apply these cutting-edge technologies responsibly and creatively across various domains.

Electives

R Programming for Data Science

R programming is a vital tool for data analysis and is essential for aspiring data science professionals. This course teaches you how to write R code, explore R's data structures, and create custom functions. By the end of the course, you will be well-prepared to embark on your first data analysis project.

Learning Outcomes:

- Learn about fundamental concepts such as math, variables, strings, vectors, factors, and vector operations in R
- Gain essential knowledge about arrays and matrices, lists, and data frames
- Explore conditions and loops, functions in R, objects, classes, and debugging
- Master the art of accurately reading and handling text, CSV, and Excel files
- Learn how to save and write data objects in R
- Understand and effectively work with strings and dates in R

Topics Covered:

- R Basics
- Data Structures in R
- R Programming Fundamentals
- Working with Data in R
- Strings and Dates in R

By completing this course, you will acquire the necessary skills in R programming to kickstart your data analysis journey and prepare yourself for success in data science.

Electives

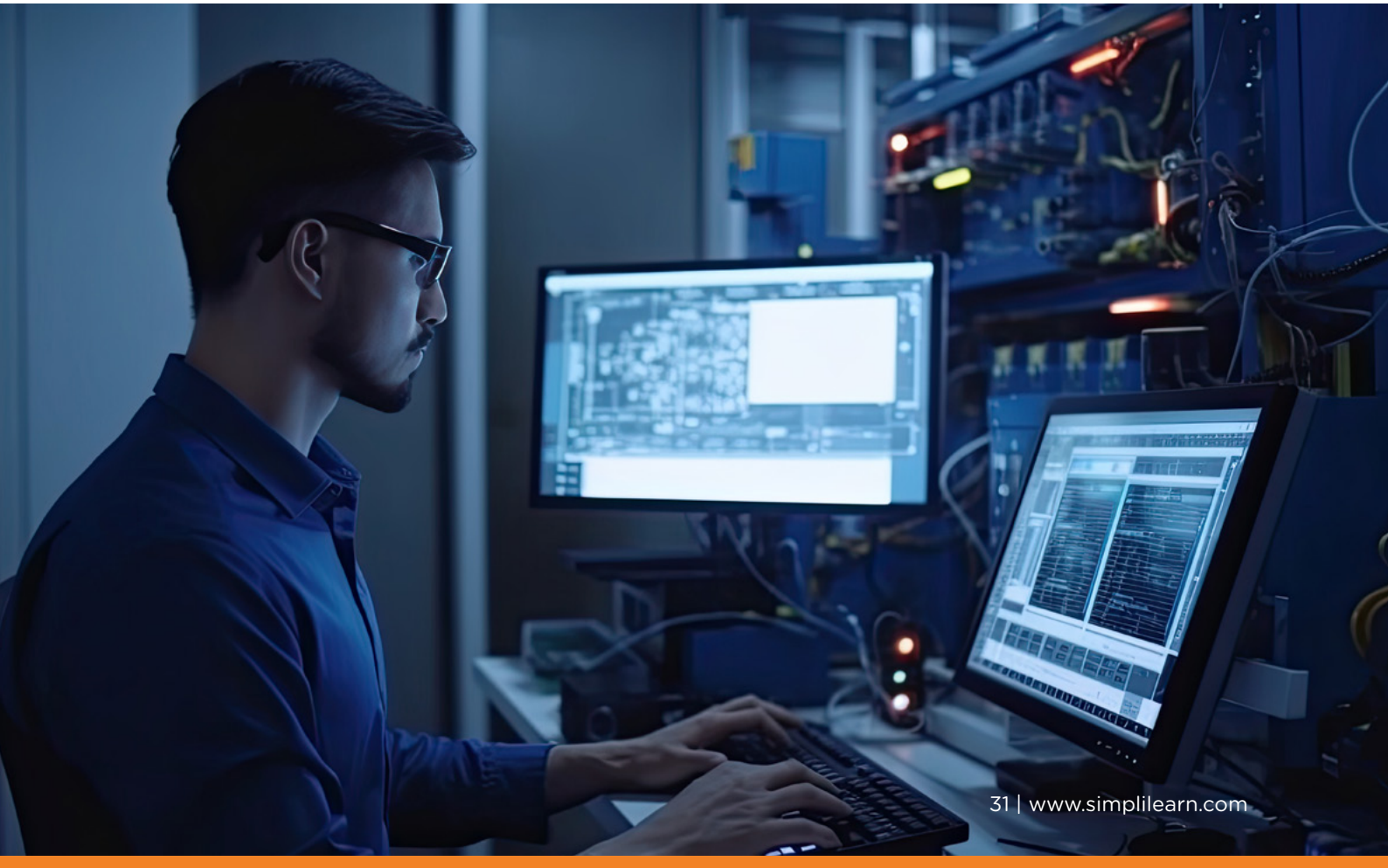
Academic Masterclass

Attend this live online Masterclass conducted by Caltech CTME instructors and get insights about advancements in data science.

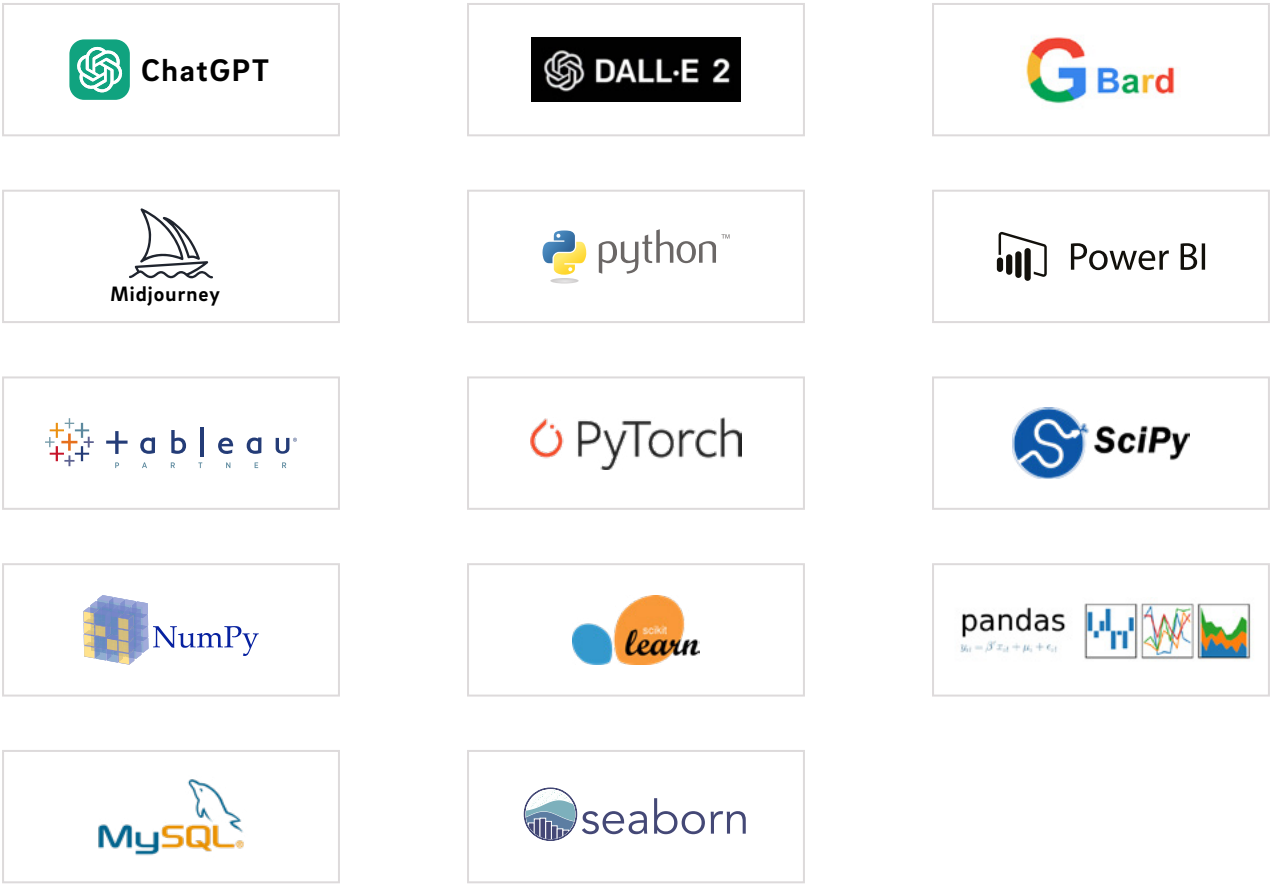
Electives

Industry Masterclass

Attend this live interactive industry masterclass to gain insights about the latest advancements in data science.



Tools Covered



Projects

Sales Analysis

Utilize Python to analyze a clothing company's sales data for the fourth quarter across Australian states to help the company make data-driven decisions for the coming year.

Employee performance Analysis

Create ML programs to understand different factors affecting employee turnover. Use clustering, the SMOTDE technique, and the K-fold cross-validation model to analyze performance and suggest employee retention strategies.

Classification of Songs

Perform exploratory data analysis and cluster analysis to create personalized song lists and an efficient recommendation system.

Interactive Sales Dashboard

Create an interactive sales dashboard for an apparel OEM in Tableau for the sales department to use for ad-hoc analysis and reporting.

Crime Analysis with Tableau Dashboard

Prepare a dashboard to keep the police department and the city updated on the statistics of crime events. You must create a dashboard/story using Tableau.

Marketing Strategies with Exploratory Data Analysis

Perform exploratory data analysis and hypothesis testing to better understand the various factors contributing to customer acquisition for a marketing department.

Ecommerce App with Python

Develop an ecommerce app on the Python platform that can categorize, add or remove items from the cart and support different payment options.

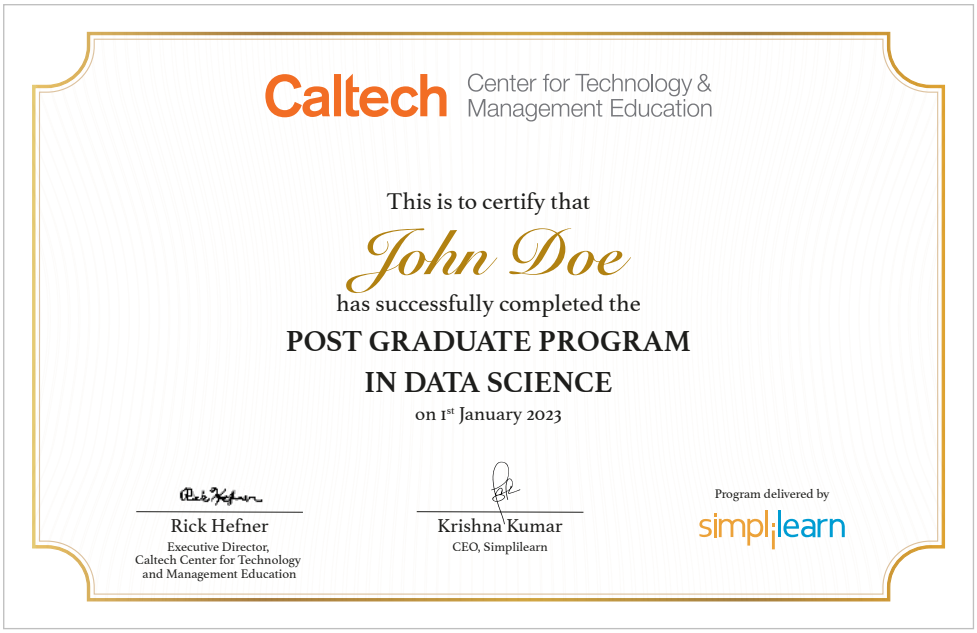
Credit Card Fraud Analysis

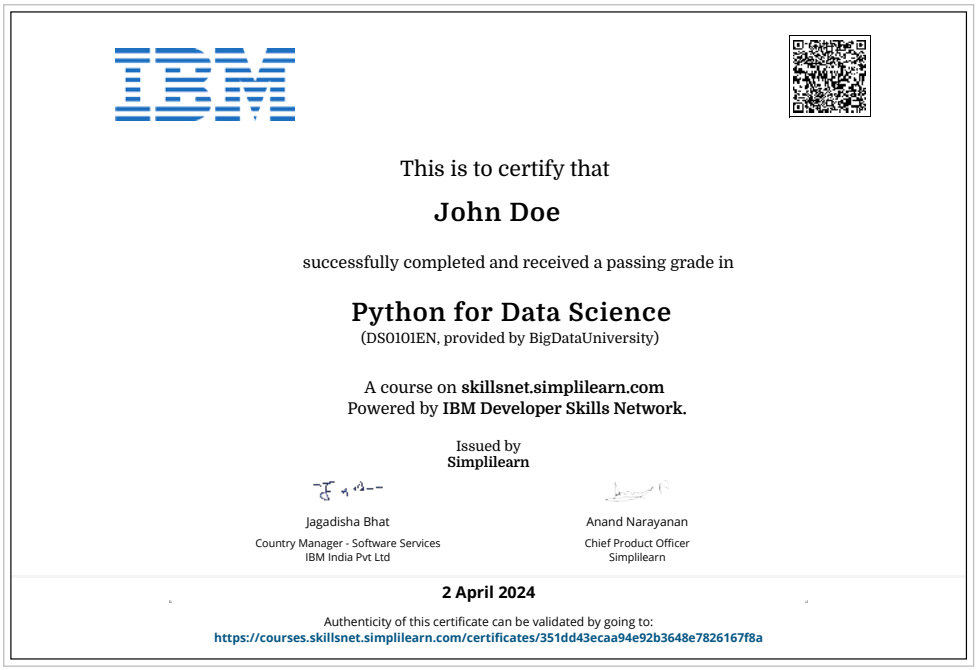
Utilize data science and machine learning methodologies to identify fraudulent credit card transactions so that end customers are not charged for items they did not purchase.

Weather Prediction

Create a classification model using ten years of rainfall data to predict the weather for different locations across Australia.

Certificates





Upon successful completion of the Caltech Data Science program, you will receive a certificate of completion from Caltech CTME. You will also receive IBM certificates for all IBM courses along with certificates from Simplilearn for the courses completed in the learning path. These certificates will testify to your skills as a data science expert.

Advisory Board Member

Rick Hefner

Program Director, Caltech Center for Technology & Management Education

rhefner@caltech.edu

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.

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