DA102: Data Analytics - Basic with Python

You Will Learn

- A brief history of machine learning and advanced pattern recognition (APR)
- Regression, statistics, and first principles techniques for prediction
- Single-In Single-Out (SISO), Multi-In Single-Out (MISO), and Multi-in Multi-Out (MIMO) modeling definitions and techniques
- Scatter Plots, Time Series diagrams, and Histograms and how each are utilized
- The differences between dependent and independent variables and how each are used
- How to train, validate, and test data sets for optimal performance
- Identification of distributions for training data selection and residual analysis
- Data cleansing, historical data analysis, quantity and quality of data
- How to use a bell curve to determine if a residual is zero-centered and balanced

The objectives of any Asset Monitoring & Diagnostics (M&D) program include minimizing the opportunity for catastrophic failure, maximizing machine and cycle efficiency, and eliminating time-based preventive maintenance tasks. Benefits of each of these three objectives rely understanding the principles behind applications found in M&D programs as well as developing their own applications to meet specific M&D program needs.

Data Analytics - Basic with Python will help you understand the analytics tools you use beyond the user manual. You will learn how to be more effective and efficient with your daily tasks, as well as new concepts and techniques to boost your M&D program value. We'll examine the basics utilizing Excel and Python, starting with an introduction to data analytics techniques and model validation. To help you develop retention and long-term recall of the course material, over 25% of class time is spent on hands-on exercises using visual association tools to break down complex topics. This course prepares you to identify and execute proper analytical techniques to help you develop these in-demand skills.

DATA AS SIMPLE AS BLACK & WHITE

DA102 Section Descriptions

DA102.1: Analytics Toolbox Orientation
Learn how to use practitioners' tool for industrial asset analytics, and why a good practitioner needs more than basic APR tools to be successful. Topics include the history of machine learning and APR, regression techniques, statistical techniques, and first principles techniques.

DA102.2: Explore Data Analysis Part 1
Analyze different variable relationships using time-series, x-y scatter, correlation analysis, and 3D scatter plots, including dependent-independent variables as well as "influence" and "predicted" variables. Topic include analytics capabilities, the need for specific sensors to detect failure modes, and avoided costs versus direct savings. Exercises include scatter plot analysis, time-series analysis, and histogram analysis.

DA102.3: Explore Data Analysis Part 2
Continue learning how to analyze different variable relationships using more complex modeling techniques combined with data requirements and data fidelity. Topics include data set types and purposes, data historian considerations, and model and data relationships. Exercises include exploratory data analysis and advanced scatter plot analysis.

DA102.4: Probability Distributions
Learn what a probability distribution is and how it relates to a frequency distribution, including discrete (binomial, discrete, and Poisson distributions) and continuous (normal, continuous uniform, log-normal, and exponential) probability distributions and when each are used. Topics include regression analysis, clustering analysis, vector-based analysis, and supervised and unsupervised learning.

DA102.5: Data Selection
Learn how to scope assets for criticality and available data, model design criteria, and how to tell if program capabilities are being oversold. Additionally, attendees learn proper data selection, cleansing, and proper quantity and quality. Topics include model scoping, model coverage, model deployment consideration, data selection, and model documentation. Exercises include regression model analysis and APR model - regression model comparison.

DA102.6: Residual Analysis
Understand which statistical and graphical data analytics techniques when evaluating regression models. Utilizing practical examples, determine residual quality and model performance by comparing predicted and actual variable data. Topics include APR model analysis, model quality and performance, and graphical analysis. Exercises include model regression analysis, predicted vs. actual comparison, and graphical analysis.

DA102.7: Model Performance Analysis
Learn how to analyze residual quality using statistics, statistical analysis, trend analysis, and model validation and testing. Topics include metrics vs. measures, and prediction uncertainty analysis.

Who Should Attend
- Data Analysts
- Asset Monitoring & Diagnostics Program Managers
- Directors or VPs of Data Analytics for manufacturing and utilities
- Information Technology professionals
- Systems Administrators who are responsible for M&D program implementation