Statistical Thinking for Industrial Problem Solving

Course outline

Statistical Thinking for Industrial Problem Solving is a free, online course available to anyone interested in building practical skills in using data to solve problems better. The course is comprised of seven modules, totaling about 30 hours of self-paced learning. Each module includes short instructional videos, JMP demonstrations, questions and exercises. Learn more and enroll at jmp.com/statisticalthinking.

The topics covered in each module are outlined below.

**Module 1: Statistical Thinking and Problem Solving**
Learn how to map a process, define and scope your project, and determine the data you need to solve your problem.

<table>
<thead>
<tr>
<th>Statistical Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>• What Is Statistical Thinking</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem Solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Overview of Problem Solving</td>
</tr>
<tr>
<td>• Statistical Problem Solving</td>
</tr>
<tr>
<td>• Types of Problems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Defining the Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Defining the Problem</td>
</tr>
<tr>
<td>• Goals and Key Process Indicators</td>
</tr>
<tr>
<td>• The White Polymer Case Study</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Defining the Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>• What Is a Process?</td>
</tr>
<tr>
<td>• Developing a SIPOC Map</td>
</tr>
<tr>
<td>• Developing an Input/Output Process Map</td>
</tr>
<tr>
<td>• Top-Down and Deployment Flowcharts</td>
</tr>
</tbody>
</table>

**Module 2: Exploratory Data Analysis**
Learn how to describe data with graphics and use interactive visualizations to find and communicate the story in your data.

<table>
<thead>
<tr>
<th>Describing Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Introduction to Descriptive Statistics</td>
</tr>
<tr>
<td>• Types of Data</td>
</tr>
<tr>
<td>• Histograms</td>
</tr>
<tr>
<td>• Measures of Central Tendency and Location</td>
</tr>
<tr>
<td>• Measures of Spread – Range and Interquartile Range</td>
</tr>
<tr>
<td>• Measures of Spread – Variance and Standard Deviation</td>
</tr>
<tr>
<td>• Visualizing Continuous Data</td>
</tr>
<tr>
<td>• Describing Categorical Data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probability Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Introduction to Probability Concepts</td>
</tr>
<tr>
<td>• Samples and Populations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Identifying Potential Root Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Tools for Identifying Potential Causes</td>
</tr>
<tr>
<td>• Brainstorming</td>
</tr>
<tr>
<td>• Multi-Voting</td>
</tr>
<tr>
<td>• Using Affinity Diagrams</td>
</tr>
<tr>
<td>• Cause-and-Effect Diagrams</td>
</tr>
<tr>
<td>• The 5 Whys</td>
</tr>
<tr>
<td>• Cause-and-Effect Matrices</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compiling and Collecting Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Data Collection for Problem Solving</td>
</tr>
<tr>
<td>• Types of Data</td>
</tr>
<tr>
<td>• Operational Definitions</td>
</tr>
<tr>
<td>• Data Collection Strategies</td>
</tr>
<tr>
<td>• Importing Data for Analysis</td>
</tr>
</tbody>
</table>

**Exploratory Data Analysis for Problem Solving**

<table>
<thead>
<tr>
<th>Understanding the Normal Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Checking for Normality</td>
</tr>
<tr>
<td>• The Central Limit Theorem</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exploratory Data Analysis for Problem Solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Introduction to Exploratory Data Analysis</td>
</tr>
<tr>
<td>• Exploring Continuous Data: Enhanced Tools</td>
</tr>
<tr>
<td>• Pareto Plots</td>
</tr>
<tr>
<td>• Packed Bar Charts and Data Filtering</td>
</tr>
<tr>
<td>• Tree Maps and Mosaic Plots</td>
</tr>
<tr>
<td>• Using Trellis Plots and Overlay Variables</td>
</tr>
<tr>
<td>• Bubble Plots and Heat Maps</td>
</tr>
<tr>
<td>• Summary of Exploratory Data Analysis Tools</td>
</tr>
</tbody>
</table>
Module 2: Exploratory Data Analysis (continued)

Communicating With Data
- Introduction to Communicating With Data
- Creating Effective Visualizations
- Evaluating the Effectiveness of a Visualization
- Designing an Effective Visualization
- Communicating Visually With Animation
- Designing for Your Audience
- Understanding Your Target Audience
- Designing Visualizations for Communication
- Designing Visualizations: The Do's and Don'ts

Data Preparation for Analysis
- Data Tables Essentials
- Common Data Quality Issues
- Identifying Issues in the Data Table
- Identifying Issues One Variable at a Time
- Restructuring Data for Analysis
- Combining Data
- Deriving New Variables
- Working With Dates

Saving and Sharing Results
- Introduction to Saving and Sharing Results
- Saving and Sharing Results in JMP
- Saving and Sharing Results Outside of JMP
- Deciding Which Format to Use

Module 3: Quality Methods

Learn about tools to quantify, control and reduce variation in your product, service or process.

Statistical Process Control
- Introduction to Control Charts
- Individual and Moving Range Charts
- Common Cause Versus Special Cause Variation
- Testing for Special Causes
- X-bar and R, and X-bar and S Charts
- Rational Subgrouping
- 3-Way Control Charts
- Control Charts With Phases

Process Capability
- The Voice of the Customer
- Process Capability Indices
- Short- and Long-Term Estimates of Capability
- Understanding Capability for Process Improvement

- Estimating Process Capability: An Example
- Calculating Capability for Nonnormal Data
- Estimating Process Capability for Many Variables
- Identifying Poorly Performing Processes
- A View From Industry

Measurement System Studies
- What Is a Measurement Systems Analysis (MSA)?
- Language and Terminology
- Designing a Measurement System Study
- Designing and Conducting an MSA
- Analyzing an MSA
- Studying Measurement System Accuracy
- Improving the Measurement Process

Module 4: Decision Making With Data

Learn to draw inferences from data, construct statistical intervals, perform hypothesis tests, and understand the relationship between sample size and power.

Estimation
- Introduction to Statistical Inference
- What Is a Confidence Interval?
- Estimating a Mean
- Visualizing Sampling Variation
- Constructing Confidence Intervals
- Understanding the Confidence Level and Alpha Risk
- Prediction Intervals
- Tolerance Intervals
- Comparing Interval Estimates

Foundations in Statistical Testing
- Introduction to Statistical Testing
- Statistical Decision Making
- Understanding the Null and Alternative Hypotheses
- Sampling Distribution Under the Null
- The p-Value and Statistical Significance
Module 4: Decision Making with Data (continued)

Hypothesis Testing for Continuous Data
• Conducting a One-Sample t Test
• Understanding p-Values and t Ratios
• Equivalence Testing
• Comparing Two Means
• Unequal Variances Tests
• Paired Observations
• One-Way ANOVA (Analysis of Variance)
• Multiple Comparisons
• Statistical Versus Practical Significance

Sample Size and Power
• Introduction to Sample Size and Power
• Sample Size for a Confidence Interval for the Mean
• Outcomes of Statistical Tests
• Statistical Power
• Exploring Sample Size and Power
• Calculating the Sample Size for One-Sample t Tests
• Calculating the Sample Size for Two-Sample t Tests and ANOVA

Module 5: Correlation and Regression
Learn how to study the linear association between pairs of variables, and how to fit and interpret linear and logistic regression models.

Correlation
• What Is Correlation?
• Interpreting Correlation

Simple Linear Regression
• Introduction to Regression Analysis
• The Simple Linear Regression Model
• The Method of Least Squares
• Visualizing the Method of Least Squares
• Regression Model Assumptions
• Interpreting Regression Results
• Fitting a Model With Curvature

Multiple Linear Regression
• What Is Multiple Linear Regression?
• Fitting the Multiple Linear Regression Model

• Interpreting Results in Explanatory Modeling
• Residual Analysis and Outliers
• Multiple Linear Regression With Categorical Predictors
• Multiple Linear Regression With Interactions
• Variable Selection
• Multicollinearity

Introduction to Logistic Regression
• What Is Logistic Regression?
• The Simple Logistic Model
• Simple Logistic Regression Example
• Interpreting Logistic Regression Results
• Multiple Logistic Regression
• Logistic Regression With Interactions
• Common Issues

Module 6: Design of Experiments
Learn the language of design of experiments (DOE) and see how to design, conduct and analyze an experiment in JMP.

Introduction to DOE
• What Is DOE?
• Conducting Ad Hoc and One-Factor-at-a-Time (OFAT) Experiments
• Why Use DOE?
• Terminology of DOE
• Types of Experimental Designs

Factorial Experiments
• Designing Factorial Experiments
• Analyzing a Replicated Full Factorial
• Analyzing an Unreplicated Full Factorial

Screening Experiments
• Screening for Important Effects
• A Look at Fractional Factorial Designs
• Custom Screening Designs

Response Surface Experiments
• Introduction to Response Surface Designs
• Analyzing Response Surface Experiments
• Creating Custom Response Surface Designs
• Sequential Experimentation

DOE Guidelines
• Introduction to DOE Guidelines
• Defining the Problem and the Objectives
• Identifying the Responses
• Identifying the Factors and Factor Levels
• Identifying Restrictions and Constraints
• Preparing to Conduct the Experiment
• Case Study

(Module 7 on back)
Module 7: Predictive Modeling and Text Mining
Learn how to identify possible relationships, build predictive models and derive value from free-form text.

Essentials of Predictive Modeling
- Introduction to Predictive Modeling
- Overfitting and Model Validation
- Assessing Model Performance: Prediction Models
- Assessing Model Performance: Classification Models
- Receiver-Operating Characteristic (ROC) Curves

Decision Trees
- Introduction to Decision Trees
- Classification Trees
- Regression Trees
- Decision Trees With Validation
- Random (Bootstrap) Forests

Neural Networks
- What Is a Neural Network?
- Interpreting Neural Networks
- Predictive Modeling With Neural Networks

Generalized Regression
- Introduction to Generalized Regression
- Fitting Models Using Maximum Likelihood
- Introduction to Penalized Regression

Model Comparison and Selection
- Comparing Predictive Models

Introduction to Text Mining
- Introduction to Text Mining
- Processing Text Data
- Curating the Term List
- Visualizing and Exploring Text Data
- Analyzing (Mining) Text Data

jmp.com/statisticalthinking