

# Data Mining and Big Data Analytics Using SPSS



Business Continuity  
Institute  
Corporate Partnership

An **informattech** Training Course  
all copyrights reserved

Transforming  
Business *for* Good



# Course content

## Why Attend

The amount of data collected in recent years has risen exponentially, and this has created the collection of datasets on customer interactions, sales, logistics, and production processes. Large retailer companies handle up to ten million transactions and five thousand items per second. This also enabled companies to introduce analysis of this data and achieve reduction in their lead times, reduce overheads, optimize workflows and increase customer satisfaction, while at the same time increasing their profits. There are many software packages for Big Data mining and Big Data Analytics, however many of these tools are mainly programming languages, and although free, they require a knowledge of computer programming.

As the data rises, the number of people who can perform the Data Mining and Big Data analytics should also rise, to handle the tide, and by using one of easiest to handle, big data analytics software is SPSS.

## This training course will feature:

- Use of SPSS software packages
- Data mining and visualization with SPSS
- Data analytics of Big Data with the use of SPSS packages
- Advanced uses of SPSS and its interoperability with other software
- Using the Big Data analytics for the benefits of the companies

## What are the Goals?

By the end of this training course, participants will be able to:

- Understand the power of Big Data and Big Data Analytics
- Harness the benefits of graphical interface of SPSS for Big Data Analytics
- Learn the concepts of data analysis and visualization
- Apply the drag and drop functionality for data visualization
- Acquire the knowledge of advanced analysis of big data, like sentiment analysis in SPSS

## Who is this Training Course for?

This training course is designed for all people involved in data mining and data analysis, as well as planning and forecasting, and especially the ones who are inclined to the use of graphical software interface, and their busy schedules are not allowing them to learn or apply programming languages like JAVA, R or Python. This course is for those who do want to harness the power of Big Data but are not programmers themselves.

This training course is suitable to a wide range of professionals but will greatly benefit:

- CTOs, CIOs and Engineers
- Data Scientists, Data Analysts
- Statisticians and technology personnel
- Marketing and research specialists



# Course content

## Course Outline

### Day One: Introduction to SPSS

- Getting to know SPSS
- Starting SPSS and Working with data files
- Descriptive Statistics
- Frequencies (categorical variables)
- Central tendency, standard deviations, and range (continuous variables)
- Getting help and running tutorials

### Day Two: Manipulating the Data and Initial Data Visualization

- Using graphs to describe and explore the data
- Histograms
- Bar graphs
- Boxplots
- Line graphs
- Calculating total scale scores
- Transforming variables
- Recode procedure
- Compute procedure
- Select cases procedure
- Split file procedure
- Reliability analysis using coefficient alpha (also known as Cronbach's alpha)

### Day Three: Inferential Statistics

- T-test, one sample t-test, independent and dependent sample t-test
- Analysis of variance (ANOVA)
- Correlation and Pearson r correlation coefficient
- Linear regression-simple and multiple linear regression
- The chi-square goodness of fit and test of independence procedures
- Autocorrelation and time series analysis

### Day Four: Advanced Statistics with SPSS Modeller

- SPSS Modeler and its main elements
- Clustering analysis, using k-means
- Association rules and apriori algorithm
- Logistics regression
- Forecasting using the Autoregressive Integrated Moving Average (ARIMA) model
- Building a decision tree model
- Sensitivity analysis of texts in SPSS

### Day Five: The Integration of SPSS with Python and R

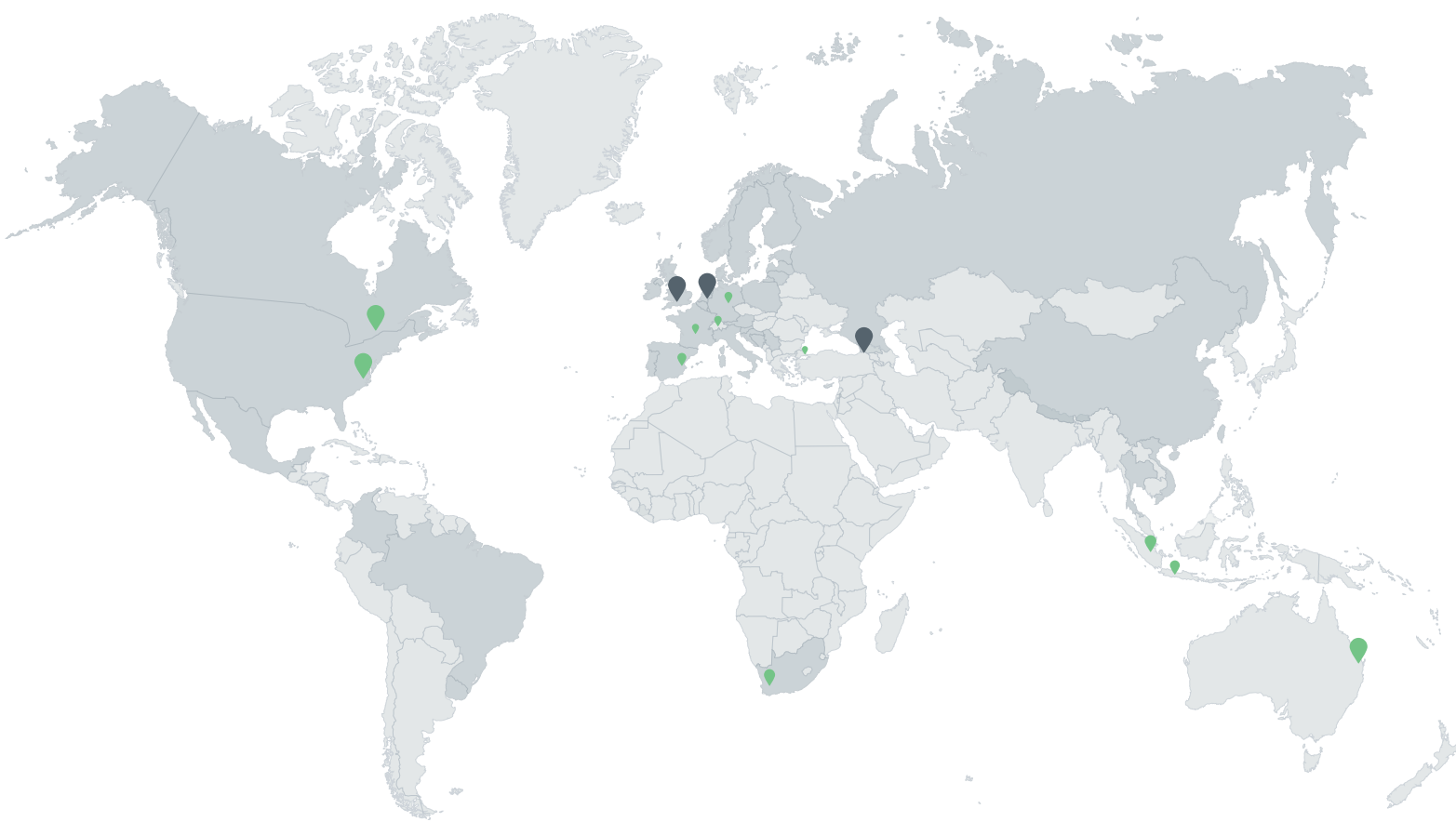
- Big Data Analytics programs
- Brief introduction to R
- SPSS Modeler-R integration
- Introduction to Python
- SPSS Modeler Python Integration
- Putting it all together

## informatech GLOBAL LEADERSHIP CONSULTANTS

**informatech** is a global leadership consultancy that aligns people, purpose & strategy - driving socially responsible transformation in global organisations. Our international network includes 215 partners, consultants, and coaches in 24 countries throughout Europe, North America, Latin America, Asia, and Africa. We transform leaders, align teams and create resolve and passion to win. Typical interventions are to accelerate performance, execute strategy and embed capability and change.



For more information about **informatech** GLOBAL LEADERSHIP CONSULTANTS or to get in touch, visit us at: [www.informatech.co.uk](http://www.informatech.co.uk)



Copyright © 2022 **informatech** All rights reserved.

Transforming  
Business *for* Good

**informatech** GLOBAL  
LEADERSHIP  
CONSULTANTS

[www.informatech.co.uk](http://www.informatech.co.uk)

+44 33 000 111 90

Performance  
Consulting

Leadership  
Development

Team & Exec  
Coaching