

## SEGMENTATION AND PREDICTIVE MODELLING WITH R

**Programme Title:** Segmentation and Prediction Modelling with R.

**Overview:** This course will teach the learner how to use R to develop segmentation and prediction models using key data mining and statistical techniques.

**Who is this programme for:** This programme is for anyone who is already familiar with R and basic statistical analysis, but who now wants the knowledge of how to apply segmentation and prediction modelling within their industry.

**Key Topics Covered:**

- CRISP-DM Process:
  - o CRISP-DM is a commonly used data mining process, which covers all topics of a data mining project including (i) Business Understanding (ii) Data Understanding, (iii) Data Preparation, (iv) Statistical Analysis, (v) Evaluation and (vi) Deployment.
- Segmentation Modelling:
  - o Application of key segmentation techniques including k-means and hierarchical clustering. Learners will be taught all aspects of segmentation modelling including how to (i) prepare the data, (ii) apply the techniques, (iii) assess quality of segmentation models, (iv) profile segments and (v) monitor accuracy of segmentation models.
- Prediction Modelling:
  - o Application of key data mining and statistical techniques including classification trees, regression and neural networks. Learners will be taught all aspects of prediction modelling including how to (i) prepare the data, (ii) apply the techniques, (iii) assess the quality of prediction models and (iv) monitor the accuracy of prediction models.

**Duration:**

- Day 1: Segmentation Modelling with R.
- Day 2: Prediction Modelling with R.

**Teaching Approach:** the tutor will use a highly practical approach to deliver this course. Learners will have access to R/RStudio on their own laptops/PCs, providing them with the immediate opportunity to implement and reinforce the material presented throughout the course.

**Prerequisites:** Learners are expected to be familiar with R and basic statistical analysis, to the point where they should be able to carry out a basic statistical technique such as linear regression in R.

**Equipment/Software Needed:** Each learner will require their own laptop/PC with both R and RStudio installed. Laptops/PCs will also require internet access throughout the course.

**Tutor:**

**Dr. Denise Earle**

**Academic Profile:**

In September 2010, I completed a PhD in Statistics at Maynooth University. My thesis focussed on the use of statistical techniques to improve data visualisation. Prior to my PhD, I completed an MA in Mathematics and a BA in Mathematical Studies and Statistics.

## **Professional Profile:**

Since September 2017, I have been a lecturer at the Institute of Technology Carlow. The majority of the courses I teach are related to data analysis, with key courses focussing on the application of statistical techniques to digital marketing data and sports performance data.

Prior to my position with IT Carlow, I spent five years working as a Data Scientist for Paddy Power Betfair. During my time here, I developed several data models including segmentation models using techniques like k-means clustering and prediction models using techniques like classification trees. These models formed the basis of many marketing strategies implemented by the various marketing teams.