ASCENT

R Practitioners.

Training course prospectus

alex.tamparopoulos@ascent.io
ascent.io

Contents.

Introduction.

R is one of the two dominant programming languages in the field of Data Science and Analytics and is considered a vital skill in a Data Scientists toolbox.

Ongoing contributions from a large active community of users and tech corporations, lead to continuous development of new open-source packages implementing state-of-the-art methods.

Leaders in the R community, founding members of the R Consortium and with a large team of our own Data Scientists, we are considered world experts in R.

We host an annual cross-sector conference focusing on the real-world commercial use of R, as well as other regular community events designed for R users to network, collaborate, learn, and unlock business value from the effective use of the R language.

Why train with us?

Every year we train thousands of Data Scientists in the R language. Our courses are consistently rated as 92% Excellent. We guide talented individuals at every point on their journey, from beginner to mastery helping them to apply proven R-based solutions in commercial environments.

We help organisations from a range of industry sectors to build effective data capability. We have helped teams to take their learnings to commercial use cases: from optimising drug discovery and improving patient experience, to determining the optimal placement of new retail stores or fitness facilities.

We develop, produce, and deliver courses as private programmes to your organisation in virtual or in-person sessions. Our courses are taught live by one of our data experts and are designed to be fully interactive for each participant. We offer coherent learning pathways for effective upskilling, to ensure that all participants can apply what they are learning in real-world scenarios.

Ascent's R training: Benefits at a glance.

- Global training for all skill levels: Our training courses range from beginner to intermediate learning, with expert mentoring consultancy for applied learning at advanced levels.

- Flexibility of engagement: Programmes are designed to deliver a fully interactive experience for course participants, either delivered face to face on premise or via virtual classroom.
- Business-focused insights: Our instructors have extensive subject matter experience and real-world application knowledge; this means they have a unique skill set which allows them to bring theory to life.
- Beyond R: We upskill in a variety of technical and non-technical areas, including Python and Data Literacy programmes for business leaders.
- Materials, certification, and post course support: As standard, we provide comprehensive course materials, practice exercises, example code, attendance certificates and post-course email support.

R training courses.

Pinpoint the path to data-science excellence.

We assess teams and individuals and prioritise training to build internal capability, guided by Data Science Radar - a tool representing the cumulation of our experience in data science consulting and participation in the R community. Data Science Radar leverages a proprietary 'trait' model of data science skills to help organisations recruit and retain the right talent - and enable individuals to prioritise their learning and development.

The DSR traits are: Programmer, Modeller, Visualiser, Communicator, Data Wrangler and Technologist.

1

Breakthrough.

Introduction to R for Analytics, Fundamentals of Modelling in R.

2

Foundation.

Programming in R for Analysts, Introduction to Shiny, Reporting Tools in R.

3

Intermediate.

Intermediate R Programming, Databases and SQL from R, Machine Learning in R, Web Scraping and Text Analysis in R, Data Visualisation in R, Package Building for R.

4

Advanced/Mastery.

For more advanced and mastery level learning needs we take a more consultative mentoring approach to ensure applied learning to address business challenges. Our Data Scientists can help you with with real-world applied learning, use case planning, project prioritisation and mobilisation. For this level of education, we will work with you to define a program which provides real-world value and addresses your advanced learning needs.

Course title	Course duration	Level
Introduction to R for Analytics	2-day	Breakthrough
Fundamentals of Modelling in R	1-day	Breakthrough
Data Science Best Practices	0.5-day	Foundation
Programming in R for Analysts	1-day	Foundation
Introduction to Shiny	1-day	Foundation
Reporting Tools in R	1-day	Foundation
Intermediate R Programming	1-day	Intermediate
Databases and SQL from R	1-day	Intermediate
Machine Learning in R	2-day	Intermediate
Web Scraping and Text Analysis in R	1-day	Intermediate
Data Visualisation in R	1-day	Intermediate
Package Building in R	1-day	Intermediate

Level:

Breakthrough (1)

Duration:

2 days

Introduction to R for Analytics.

This two day introduction to R focuses on getting started with common data tasks. By the end of this course, attendees will be confident in how they can import data, perform common manipulation tasks and visualise data. Along the way they will be introduced to a variety of data types including dates and categorical data. This course will be taught using the popular tidyverse set of packages and is ideal for those new to R who want to quickly get started with analytics in R.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

No prior knowledge of R, or programming, is assumed.

Details.

Introduction to R and RStudio:

- The R community and resource
- Introduction to RStudio
- R packages
- Creating R scripts

Getting Data into R:

- Importing tabular and structured Excel Data
- The data frame

Data Manipulation:

- Filter and select
- Adding/updating columns
- Summarising data
- Grouping data

Working with Dates, Categories and Characters:

- Manipulating date/time format
- Working with categorical data
- Manipulating character strings

Getting Started with ggplot2:

- Creating a basic plot
- Changing the type of plot
- Controlling aesthetics

Extending ggplot2:

- Groups and Panels
- Controlling themes

R Statistics:

- The simple linear model in R
- Model diagnostics
- Updating and comparing models

Getting Data into the Correct Format:

- Changing the shape of data
- Merging/joining two datasets

Level:

Breakthrough (1)

Duration:

1 day

Fundamentals of Modelling in R.

This one day course is intended as a practical introduction to the basics of analytics. Taught using R, attendees will be introduced to ideas such as sampling, statistical testing and linear modelling. This course provides the foundations that are built upon for advanced analytic topics including machine learning.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio. The course will be taught by a Mango Solutions consultan.

Prerequisites.

It is assumed that participants have attended the Introduction to R for Analytics or similar.
It is assumed that attendees are familiar with basic data concepts such as data types, and summary statistics (pre-reading can be provided for those that require a refresher on these topics).

Details.

Sampling and Distributions:

- Populations and samples
- The Normal distribution
- Useful distributions (Poisson & Binomial)
- Confidence intervals

Statistical Testing (A/B Testing):

- Defining a statistical test
- Testing a difference in means
- Multiple testing
- Types of error

The Linear Model:

- A simple linear model
- Understanding the output
- Is it a good model?
- Multiple regression
- Picking the best model

Extending the Linear Model:

- Models for binary data
- Models for counts

Level:

Foundation (2)

Duration:

0.5 days

Data Science Best Practices.

For those who have a basic understanding of programming, to ensure that their data science projects always follow best practices in all, and to help emphasise the importance of best practices when working in a team – as we know, Data Science is a team sport. While all the Mango training courses help to instil these best practices; this course focuses on what the options are and why we need to use them in a language agnostic setting. This course outlines the key best practices associated with the 6 core traits of a data science team and project: communication, data wrangling, modelling, programming, technology, and visualisation.

This course is delivered by one of our practicing data science consultants, as a half day interactive session, supplemented by a presentation, with discussion points throughout.

Prerequisites.

Attendees are expected to have some knowledge of programming in either R or Python and have attended our Delivering Successful Analytics Projects course (or similar).

Details.

Communication:

- Collaboration (incl. multi-user version control)
- Code and process documentation
- Project framework (scoping, planning, presenting results, IDEaL framework)

Data Wrangling:

- FAIR principles
- Data governance & ethics
- Data tools

Modelling:

- Choosing the best models (what to model, performance, uncertainty, interpretability, etc.)
- Productionising models (reproducible model code, deployment and maintenance)

Programming:

- Methods of code design (TDD, pseudo code, working with legacy code)
- Development method (PP, code review, iterative approach etc.)
- Coding standards
- Testing and code coverage
- Package building

Technology:

- Architecture design (infrastructure, pipeline planning, implementation and monitoring, CI/CD, version control)
- Reproducible Projects

Visualisation:

- Designing data visualisations (choosing graph type and data, the narrative, common pitfalls, what to include)

Level:

Foundation (2)

Duration:

Programming in R for Analyst.

Once you have started to use R for common manipulation tasks you will quickly find that you want to be able to do more. This one day course introduces the topics that you will need to be familiar with in order to get the most from R. This course will go into detail on the data types in R and how to interact with them at a lower level, how you can start writing functions and, as well as introducing some further tidyverse tools for iterating over data and manipulating specific elements.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

Attendees are expected to have a basic understanding of R and analytics, and should have taken the Introduction to R for Analytics or similar.

Details.

R Data Objects: Filling in the Gaps:

- Recap of R data types
- Functions for creating vectors and matrices
- Subscripting objects
- Working with lists and data frames

Writing R Functions:

- The basic structure of an R function
- R function arguments
- Control structures (if, else)
- The basics of loops

Using the R Class System:

- What is the class system in R
- Using the class system
- Finding help for the class system

Conditional Data Manipulation:

- Testing for data types
- Filter and select by types
- Transform and summarise by type

Iteration:

- Splitting data by category
- Applying functions to subsets

Level:

Foundation (2)

Duration:

1 day

Introduction to Shiny.

As data scientists we can gain great insight from our analysis but to have impact we need to share the results of that analysis. For R users one of the simplest ways to do that is through shiny, a web development framework that allows us the power of interactive web applications combined with the power of R, all without leaving a language we are comfortable with. This one day introduction to shiny will help you to understand the building blocks of shiny and attendees will leave able to create simple applications and dashboards.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

A good working knowledge of R programming is assumed, attendees should have completed Mango's Programming in R for Analysts or similar. No prior knowledge of shiny is assumed.

Details.

The Basic Structure of a Shiny Application:

- Multi-file applications
- Single file applications
- Running a shiny application

Defining User Inputs:

- Defining an input
- Understanding input options

Rendering Output:

- Defining an output type
- Connecting inputs to outputs
- Uploading and downloading files

The Basics of Reactivity:

- Introduction to reactive programming
- Controlling input with submit
- Controlling flow with reactivity

Controlling the layout:

- Defining a page layout
- Working with tabs and pages
- Controlling css & using shinythemes
- Creating a dashboard with shinydashboard

Level:

Foundation (2)

Duration:

1 day

Reporting Tools in R.

Not only does R provide us with the tools for performing analysis but it also allows us to produced high quality documentation, meaning we can keep our reporting and analysis all in one place. This one day course introduces RMarkdown, a simple but effective way of creating documents directly from R. At the end of the course attendees will be able to generate reports in both HTML and Word or Powerpoint as well as create effective dashboards using flexdashboard.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

Attendees are expected to have a basic understanding of R programming and should have completed Mango's Programming in R for Analysts or similar.
Whilst experience of shiny would be beneficial, it is not mandatory.

Details.

Basics of R Markdown:

- Formatting text
- Including code
- Basic figures and tables
- Output options

Parameterized Reports:

- Defining report parameters
- Using parameters
- Building a report with parameters

Formatting Documents:

- Sections and Contents
- Controlling figures and tables
- Formatting Word Documents

Creating Presentations:

- Presentation styles and common formatting
- Creating Powerpoint slides

Simple Dashboards:

- Basics of flexdashboard
- Alternative layouts
- Value and gauge outputs
- Including shiny components

Level:

Intermediate (3)

Duration:

1 day

Intermediate R Programming.

As you start to write more re-usable code your code will need to be more robust. In this one day programming course we introduce some of the functionality that allows your code to be more user friendly and stand up to unexpected use cases, as well help you get started with understanding how to resolve issues. By the end of this course attendees will be familiar with extended function writing topics, how to program in the tidyverse and the basics of object orientation in R as well as being introduced to tools for debugging and profiling.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

Attendees are expected to have a good understanding of the basics of R programming and data structures and should have taken Mango's Programming in R for Analysts or similar.

Details.

Writing Robust Functions:

- Warnings, errors and messages
- Handling missing input
- Invisible return values

Programming in the tidyverse:

- Why we need tidyeval
- Capturing expressions
- Using expressions in functions

Debugging and Profiling:

- Inserting breakpoints
- Browsing the function environment
- Profiling R code
- Common efficiency gains

Getting Started with Object Oriented Programming:

- Recap of the class system
- Creating S3 objects
- Methods for S3 objects
- Methods for operations (e.g. `+`)

Level:

Intermediate (3)

Duration:

1 day

Databases and SQL from R.

As our data gets bigger, or simply shared across the business, we will typically find that it is stored in a database. To be able to get the most out of our analysis we need to be able to interact with the database from R, getting data into R for our analysis. This course will leave attendees with a basic understanding of relational databases as well as the ability to connect to a database, they will also learn basic SQL statements and tools in R for easily extracting data.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.
Please note that for this course attendees will be provided with a small SQLite database. This course cannot be taught using customer databases.

Prerequisites.

A basic working knowledge of R for data analysis is assumed, attendees should have attended Mango's Introduction to R for Analytics or similar.

Details.

Databases Overview:

- Relational Databases
- Primary and foreign keys
- Data types

Connecting to a Database:

- Creating a database connection
- Getting a table into R
- Disconnecting

Databases and dplyr:

- Connecting to a table
- Running standard dplyr commands
- Collecting data from the database

Basic SQL Statements:

- Selecting & Filtering
- Summary Statistics
- Grouping
- Joins

Creating Safe SQL Statements:

- Creating strings with glue
- Building SQL statements

Level:

Intermediate (3)

Duration:

2 days

Machine Learnig in R.

This two day course is aimed at not only teaching an understanding of some of the most common machine learning techniques, but also the approach to implementing machine learning. During this course attendees will learn how to define a problem and prepare data, the range of techniques available for solving common problems and the approaches to take to evaluate models and achieve the best results possible.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

Attendees should have a good working knowledge of R and should have completed the programming in R for analysts course.
Basic knowledge of statistics is assumed (i.e. participants should be comfortable with the contents of the Fundamentals of Modelling for Data Science).
No prior knowledge of machine learning is assumed.

Details.

Getting Data in Shape:

- Identify targets, features, and ids
- Exploratory data analysis
- Creating a model-ready dataset

Data Pre-processing:

- Introduction to the {recipes} package
- Imputing missing values
- Handling categorical data

Workflows:

- Streamlining an ML workflow
- Evaluating with ROC curves

Hyperparameter Tuning:

- Defining the hyperparameter set
- Tuning with resamples
- Finalising a model

The Machine Learning Workflow:

- Splitting data with {rsample}
- Fitting models with {parsnip}
- Evaluating models with {yardstick}

Algorithms:

- Decision Trees
- Bagging with Random Forest
- Gradient Boosted Trees

Resampling:

- Why resample?
- Cross validation and bootstrapping
- Fitting and evaluating resamples

Full Machine Learning Pipeline

Level:

Intermediate (3)

Duration:

1 day

Web Scraping and Text Analysis in R.

Many companies have a large amount of data stored as text that is not being used effectively. In this one-day course we will introduce how you can get started with analysing text data, from simple manipulation and sentiment analysis through to topic modelling. By the end of the course attendees will have a good understanding of the techniques as well as how to implement them in R.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

A good working knowledge of R programming is assumed, attendees should have complete Mango's Programming in R for Analysts or similar. Familiarity with basic analytic techniques and linear modelling is required.

Details.

Getting Data from the Web:

- Reading HTML with rvest
- Extracting HTML elements
- Extracting tables

Simple Text Manipulation:

- Tidy text format
- Tokenisation
- Removing stop words
- n-grams

Sentiment Analysis:

- Sentiment Lexicons
- Joining sentiments to documents

Word and Document Frequency:

- Term Frequency
- Inverse Document Frequency

Topic Modelling:

- Latent Dirichlet Allocation
- Topic Probabilities

Level:

Intermediate (3)

Duration:

1 day

Data Visualisation in R.

Discover and understand the principles required to create powerful visualisations which effectively and accurately communicate the stories inside data. Using the powerful and expressive ggplot2 package you will learn how to apply these principles while exploring the common pitfalls to avoid when creating your own visualisation.

This will be a hands-on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

Attendees are expected to have a good understanding of R and analytics, and should have taken or be familiar with the content of the Programming in R for Analysts course.

Details.

The Objective of a Visualisation:

- Overview
- Exploring
- Reporting

Introduction to **ggplot2**:

- Introduction
- Overview of the grammar of graphics
- A basic **ggplot2** template
- Customising your plots

A (Quick) Guide to How Humans Perceive Graphics:

- Introduction
- What do the experts say?
- The data-ink ratio
- Common perceptual biases
  - Areas
  - Colour

Mapping Data and Analysis to Graphics:

- Terminology
- Introduction
- Magnitude/Ranking
- Correlation/Relationship
- Distribution/Composition

Discussion:

- The importance of a visualisation vocabulary
- Charts we didn't talk about
- Case studies

Level:

Intermediate (3)

Duration:

Package Building for R.

Being able to build packages allows you to work more effectively and easily share code with colleagues or even the wider R community. In this course we will focus on how you can quickly get started with building packages, understand the benefits of package building best practices and be able to implement them. This includes being able to more efficiently write documentation, creating tests and understanding the benefits of version control systems and how they can enhance your package building.

This will be a hands on course taught using the RStudio IDE with exercises throughout. All attendees will need access to a computer and will need to have pre-installed a recent version of R and RStudio and will need to be able to install R packages. The course will be taught by Mango Solutions consultants.

Prerequisites.

Attendees are expected to have a strong understanding of R. Whilst not mandatory, it is recommended that attendees have completed Mango's Intermediate R programming or similar.

Details.

Working with devtools:

- Why build a package?
- The basic package structure
- Building and checking a package

Documenting Functions:

- Introduction to roxygen2
- Documenting functions
- Documenting packages

Introduction to Testing:

- Why test?
- Getting started with testthat
- Including tests in a package

Writing User Guides:

- Recap of RMarkdown
- Writing a user guide
- Including a user guide in a package

Additional Package Components:

- Including datasets in a package
- Including C++ code in a package

Best Practices for Package Building:

- Coding best practices
- Using version control
- Test coverage