

Data Analyse met R

Doelgroep Cursus Data Analyse met R

De cursus Data Analyse met [R](#) is bestemd voor Big Data analisten en wetenschappers die R willen gebruiken om hun data te analyseren en voor het maken van statistische analyses.

Voorkennis Data Analyse met R

Ervaring met [programmeren](#) is bevorderlijk voor een goede begripsvorming maar is niet vereist.

Uitvoering Training Data Analyse met R

De theorie wordt behandeld aan de hand van presentaties en voorbeelden. De concepten worden toegelicht met demo's. Daarna is er tijd om er zelf mee te oefenen. R-Studio wordt gebruikt als ontwikkelomgeving. De cursustijden zijn van 9.30 tot 16.30

Certificering Cursus Data Analyse met R

De deelnemers krijgen na het goed doorlopen van de cursus een officieel certificaat R Programmeren.

Duur: 4 dagen

Prijs: € 2650

[Open Rooster](#)


Data Analysis with R



Inhoud Data Analyse met R

In de cursus Data Analyse met R leert u programmeren in de taal R en hoe u R kunt gebruiken voor data analyse en visualisatie. R is uitgegroeid tot een standaard platform voor data analyse en het maken van grafieken en kan een enorme reeks statistische procedures uitvoeren. In de cursus Data Analyse met R wordt gebruikt gemaakt van een reeks samenhangende R packages die bekend staan onder de naam de tidyverse. Deze packages hebben een onderliggende design filosofie, grammatica en data structuren gemeen en zijn speciaal geschikt voor data science.

R Intro

De cursus Data Analyse met R gaat van start met de installatie van R en de R Studio ontwikkelomgeving. Ook wordt ingegaan op de basis syntax van R en de installatie van R packages.

Plotting in R

Vervolgens leert hoe u met het ggplot2 package door middel van grafieken snel inzicht kunt krijgen in de data. Hierbij komen de verschillende plot types, themes en layouts aan de orde.

Transformaties

Dan is het tijd voor het dplyr package waarmee veel voorkomende data transformatie problemen zoals filteren, sorteren, sommeren en groeperen, kunnen worden opgelost.

Data Cleaning

Ook het presenteren van data met het rmarkdown package wordt behandeld. Evenals het in een nette vorm brengen van ruwe data met het tidyr package, waarbij kolommen variabelen en rijen observaties worden.

Date en Times

In veel data sets komen tijd reeksen voor. Op de verwerking daarvan wordt ingegaan met het lubridate package dat beschikt over vele handige functies voor de verwerking van datums en tijd.

Data Import

Onderdeel van het programma van de cursus is ook het importeren van data uit csv files en file formaten van andere statistische pakketten zoals SPSS of SAS. Het lezen uit en schrijven naar databases passeert ook de revue.

Statistische Analyses

Tenslotte behandelt de cursus Data Analyse met R statistische analyse modellen zoals lineaire en niet-lineaire modellen, variable transformaties en regressies. Dit alles wordt ondersteund met veel voorbeelden uit de praktijk en kan ook toegepast worden op cases die door de cursisten worden meegenomen.

Modules Data Analyse met R

Module 1 : Intro R	Module 2 : Graphics and Plots	Module 3 : Transformations
Overview of R History of R Installing R The R Community R Development R Studio R Console R Style Using R Packages Cheatsheets R Syntax R Objects	ggplot2 Graphics Devices and Colors High-Level Graphics Functions Low-Level Graphics Functions Graphical Parameters Controlling the Layout Changing Plot Types Quick Plots and Basic Control Aesthetics Changing Plot Types Labels Themes and Layout	dplyr R Functions Functions for Numeric Data Scoping Rules mutate arrange group by summarize select filter joining dataframe
Module 4 : Presentation	Module 5 : Data Cleaning	Module 6 : Date Times
rmarkdown Reproducible research Reporting Sharing results Repetitive Tasks Family of apply Functions apply Function lapply Function sapply Function tapply Function	tidyr spread gather separate unite Logical Data Missing Data Character Data Duplicate Values NA's	Time and Date Variables lubridate Setting a datetime Getting values from a datetime strftime Command strptime Command as.Date function Datetimes Calculations difftime Command Time Series Analysis
Module 7 : Data Import	Module 8 : Linear Models	Module 8 : Non-Linear Models
R Datasets Data.Frames Importing CSV Files Import from Text Files Import from Excel Import from Spss or SAS Connecting to a database Connecting to a cluster Databases and ODBC dbplyr	What is a model? Statistical Models in R How to evaluate a model? How to use a model? Simple Linear Models logistic regression linear regression R squared p values confidence intervals	Decision Trees random forest boosting overfitting Optional material : Interactive dashboards with Shiny Web Scraping Writing packages Spark Functional programming