

Data Analytics

Workbook

Open University of the Netherlands  
Faculty of Management, Science & Technology

*Course team*

Dr S. Bromuri, project leader and *author*  
Prof. dr. ir. R.W. Helms, *author*

*Discipline direction*

-

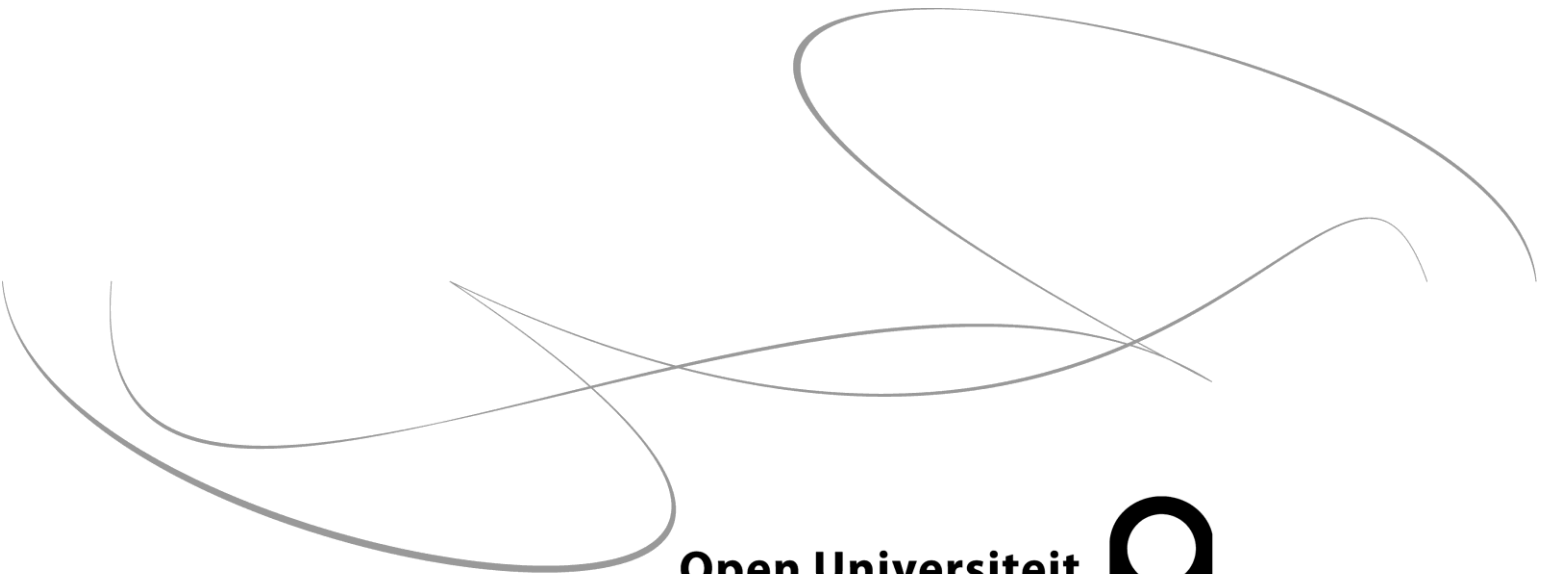
*Programme director*

Prof. dr. R.J. Kusters

*External referent*

-

# Data Analytics



**Open Universiteit**  
www.ou.nl



*Production*  
Open University of the Netherlands

*Editor*  
-

*Lay-out*  
-

*Cover*  
-

Visuele communicatie Open Universiteit

*Printing and binding*  
-

© 2016 Open University of the Netherlands

Save exceptions stated by the law no part of this publication may be reproduced in any form, by print, photoprint, microfilm or other means, including a complete or partial transcription, without the prior written permission of the publisher.

First edition: 2016

-  
ISBN XXX XX XXXXXXXX X (serie)  
ISBN XXX XX XXXXXX XX X (workbook)

Course code IM0053

## Structure of the Data Analytics course

| Part              | Block   | Learning unit   | Textbook                             | Week        | Workload (hours) | Page                       |
|-------------------|---|---|--------------------------------------|-------------|------------------|----------------------------|
| Workbook          |   | 1.1 Introduction to the course  | Workbook                             | 1           | 8                | 7                          |
|                   | 1 Data and Charting   | 1.2 Data Analytics Thinking<br>2.1 Charts in Rapid Miner  | Provost (1)<br>Work Book             | 1<br>2      | 6<br>15.5        | 1–19<br>16–25              |
|                   | 2 The Data Mining Process   | 3.1 The Data Mining Process   | Provost (2)<br>Provost (3)           | 3<br>3      | 16               | 19–71                      |
|                   | 3 Machine Learning  | 4.1 Machine Learning Part 1: Algorithms<br>4.2 Machine Learning Part 2 : Metrics                            | Provost (4)<br>Provost (8)           | 4<br>5      | 18.5<br>17.5     | 81–110<br>209–230          |
|                   | 4 Regression  | 5.1 Regression Part 1: Algorithms<br>5.2 Regression Part 2: Metrics   | Provost (5)<br>Workbook              | 6<br>7      | 9<br>10          | 111–140<br>35–38           |
|                   | 5 Clustering  | 6.1 Clustering Part 1: Algorithms<br>6.2 Clustering Part 2: Metrics   | Provost(6)<br>Provost(6)<br>Workbook | 7<br>8<br>8 | 9<br>10          | 141–161<br>161–185         |
|                   | 6 Rule Mining   | 7.1 Association Rule Mining Part 1: Algorithms<br>7.2 Association Rule Mining Part 2: Metrics               | Workbook<br>Provost (12)             | 8<br>9      | 9<br>10          | 46–53<br>291–298           |
|                   | 7 Text Mining   | 8.1 Text Mining Part 1: Algorithms<br>8.2 Text Mining Part 2: Metrics                                       | Provost (10)<br>Provost (10)         | 10<br>10    | 9<br>10          | 251–267<br>267–279         |
| Assignment bundle |   | Introduction  |                                      |             |                  | 7                          |
|                   | 1 Assignment & Forms  | Task 1: Data Mining Assignment 1<br>Task 2: Data Mining Assignment 2<br>Task 3: Webinar<br>Form 1<br>Form 2 |                                      |             | 7<br>7<br>6      | 11<br>13<br>16<br>18<br>20 |
|                   | 2 Theory & Exercises  | Data Analytics Exercises for LUs  |                                      |             |                  | 22–54                      |
| Textbook          | Data Science for Business (1 <sup>st</sup> edition), Foster Provost & Tom Fawcett (2013), Sebastopol CA: O'Reilly Media |   |                                      |             |                  |                            |
| Website           | <a href="http://www.youlearn.nl">http://www.youlearn.nl</a>   |   |                                      |             |                  |                            |