
| | |
|---|---|
| Project acronym: | **EDSA** |
| Project full name: | **European Data Science Academy** |
| Grant agreement no: | **643937** |

# D3.4 Report on the delivery of video-lectures, webinars and face-to-face trainings

| | |
|---|---|
| Deliverable Editor: | **Angi Voss (Fraunhofer)** |
| Other contributors: | **OU, SOTON, JSI, ODI, TU/e, Persontyle, KTH** |
| | **Inna Koval (JSI)** |
| Deliverable Reviewers: | **Joos Buijs (TU/e)** |
| Deliverable due date: | **31/01/2018** |
| Submission date: | **12/01/2018** |
| Distribution level: | **P** |
| Version: | **FINAL** |

This document is part of a research project funded
by the Horizon 2020 Framework Programme of the European Union

## Change Log

| Version | Date | Amended by | Changes |
|---|---|---|---|
| 0.1 | 18/10/2017 | Angi Voss | Initial version |
| 0.2 | 15/12/2017 | Angi Voss | All partners' contributions integrated |
| 0.3 | 09/01/2018 | Angi Voss | Changes due to internal reviews |
| 0.4 | 12/01/2018 | Elena Simperl | Scientific review |
| 1.0 | 12/01/2018 | Aneta Tumilowicz | Final QA |

# Table of Contents



# List of Tables

## List of Figures

# 1. Executive Summary

This deliverable reports on the courses delivered by the EDSA partners. As a follow-up to D3.1 and D3.2 it covers the second half of the project. Recommendations based on learning analytics can be found in the parallel deliverable D3.5.

EDSA has more than reached the promised size of the audience with its video lectures, distance learning and face-to-face learning courses.

In the first half of the project the partners produced materials for EDSA from their courses. In the second half, they reintegrated materials produced for EDSA in their courses and started to diversify: from F2F to online courses, from academic to professional courses, from national to international delivery, from courses without exams to certifications.

Deep learning has turned out as a hot topic. It dominates the most popular videolectures, was included in the revised EDSA curriculum and is covered by three partners in their courses. Together with machine learning it is also the topic of the EDSA bootcamps, a marketing action to promote the EDSA label.

# 2. About this report

The objectives of WP3 are to

1. Deploy the materials developed in WP2 in courses for different target groups and in different environments: webinars, videolectures and face-to-face training.
2. Gather data about the effectiveness of learning in the courses.
3. Analyse this data and obtain indications of how to improve the content or form of deployment.

WP3 has produced two series of deliverables. While the first series refers to objective 1, the second one addresses objectives 2. Objective 3 is addressed by both series of deliverables.

1. D3.1 (month 6), D3.2 (month 18) and D3.4 (month 36) give an overview of all courses delivered, while recommendations for the learning resources and future delivery focus on face-to-face courses.
2. D3.3 (month 18) and D3.5 (month 36) contain the results of learning analytics for the online courses and corresponding recommendations.

Recommendations on the core curriculum as a whole was also obtained from the industrial board in WP1.

This deliverable D3.4 is the last of the first series. It reports on data science courses that were delivered by the partners in Task 3.1 (videolectures), Task 3.2 (MOOCS, e-learning, webinars) and Task 3.3 (face-to-face training) since D3.2 in month 18.

Feedback from the courses delivered was used to improve and diversify these courses. The feedback was obtained either through forms from the participants of face-to-face courses or through automated learning analytics of online courses. The evaluation of online courses and their delivery based on learning analytics can be found in the parallel deliverable D3.5.

# 3. Delivered courses

The technical annex of the project proposed target figures for the audience to be reached by the delivery of different formats: eLearning (passive) corresponds to the videolectures of T3.1, eLearning (active) corresponds to the MOOCs and e-learning courses of T3.2, while face-to-face trainings are covered by Task 3.3. The next table shows how we reached our targets in the different formats. For videolectures the number promised was 200 000. In fact, videos in the data science category were viewed 271.333 times from 2015-2017, and 2 632 318 times since the start of the platform. In the category of MOOCs and e-learning we attracted in total 115 954 participants, which is more than twice the promised number of 50 000. In the category of face-to-face courses we reached 4 782 participants, which is almost twice the promised 2 500.

**Table 1: Target audience figures**

| **Activity** | **eLearning (passive) / Videolectures (views counted)** | **eLearning (active) / MOOCs (registered students counted)** | **Face-to-face-training (participants counted)** |
|---|---|---|---|
| Audience reported in D3.1, month 6 | 1 005 for ESWC 2015 | 24 558 | 356 |
| Total audience reported in D3.2, month 18 | 31 733 for ESWC 2015 | 50 089 | 2 018 |
| Audience reported since month 18 | 59596 (number of views of data science videolectures from M18 to M36 years) | 65 865 | 2 764 |
| Total audience reached | 271 333 (number of views of data science videolectures from 2015-2017 years) 2 632 318 (number of views in Data Science category) | 115 954 | 4 782 |
| **Overall target** | **200 000** | **50 000** | **2 500** |

The next subsections are devoted to the courses given in the different formats of Table 1. Where possible we provide the course characteristics suggested in WP2. This can be title, length, stage, sector, target group, experience, and level of the course:

Stages: The core curriculum in WP2 distinguishes four so-called stages, where the courses reported here can address multiple stages:

- Foundations
- Storage & processing
- Analysis
- Interpretation & use

Target groups: We distinguish four roles of data scientists. A single course can address multiple roles in its target group:

- BE: managers, product developers and business experts
- DM: data-skilled persons: data managers, curators and data engineers
- DA: data analysts
- IT: system architects and application developers

Experience: We consider two levels of experience in the target group.

- S: Students
- P: Practitioners

This essentially corresponds to the distinction between commercial and academic courses.

Level: We consider courses at three levels:

- 1: Basic
- 2: Advanced
- 3: Expert

Sectors: We use the same sectors as in the online survey in WP1 and in D3.2.

### *3.1 Videolectures at JSI*

Task 3.1 of WP3 is dedicated to content delivery through videolectures. VideoLectures.NET is a free and open access educational videolectures repository providing lectures from known scholars and scientists at various events, such as conferences, summer schools, workshops and science promotional events. Below we present work performed in the context of EDSA training delivery through VideoLectures.Net portal.

**Recommended Lectures at VideoLectures.NET portal**

The EDSA project continues to collate videos on topics related to data science in order to further data science learners' knowledge in the respective fields. Lectures are recommended on a periodic basis and distributed through http://edsa-project.eu/video-lectures. The lectures are selected by the EDSA project partners based on their relevance to data science education and training around the European Union, and to the level of expertise of the speakers.

**Introducing a Data Science Category at VideoLectures.NET portal**

As a part of Task 3.1 (Content delivery through Videolectures), JSI introduced a new category at VideoLectures.NET portal. The "Data Science" category incorporates a number of relevant topics. The number of videos per category is constantly growing and is shown in parentheses below.

Algorithms (6), Algorithms and Data Structures (32), Artificial Intelligence (468), Big Data (489), Bioinformatics (227), Compressed Sensing (27), Computational Linguistics (113), Computer Vision (647), Crowdsourcing (29), Databases (71), Data Mining (772), Data Modeling (23), Data Visualisation (73), DBMS (1), Decision Support (60), Econometrics (4), Image Analysis (103), Information Extraction (86), Information Retrieval (216), Internet, World Wide Web (177), Java (4), Knowledge Extraction (259), Machine Learning (3618), Machine Translation (41), Multilingual Information Access (168),

Multimedia Search (10), Natural Language Processing (246), Network Analysis (351), Optimization Methods (132), Pattern Recognition (107), PHP (4), Python (22), Semantic Computing (14), Semantic Search (56), Semantic Web (1600), Sensor Networks (10), Social Computing (27), Social Media (226), Social Sciences Methodology and Statistics (6), Speech Analysis (62), Statistics (95), Streaming Data (3), Text Mining (168), Visual Computing (6), Web Mining (179), Web Search (99).

Categories stay the same from previous mapping to the stages of our data science curriculum (Table 2).

**Table 2*:* Matching categories to the stages of data science**

| Stage | Categories |
|---|---|
| Foundations | Econometrics, Statistics, Social_Sciences_Methodology_and_Statistics, Data Mining, Information Extraction, Knowledge Extraction, Machine Learning, Semantic Web, Text Mining |
| Storage and processing | Big Data, Computer Vision, Crowdsourcing, Information Extraction, Information Retrieval, Knowledge Extraction, Machine Learning, Natural Language Processing, Semantic Computing, Visual Computing, Streaming Data, Text Mining, Web Mining |
| Analysis | Compressed Sensing, Computer Vision, Data Mining, Decision Support, Image Analysis, Information Extraction, Information Retrieval, Knowledge Extraction, Machine Learning, Natural Language Processing, Network Analysis, Pattern Recognition, Semantic Computing, Semantic Search, Text Mining, Visual Computing, Web Mining, Web Search, Multimedia Search, Sensor Networks |
| Interpretation and use | Bioinformatics, Computational Linguistics, Data Visualisation, Image Analysis, Internet, World Wide Web, Multilingual Information Access, Multimedia Search, Sensor Networks, Semantic Search, Social Media, Web Search |

In total over 11,500 lectures and tutorials related to data science can be found at Videolectures.NET which were recorded before the start of the project.

The next figure displays the distribution of videolectures by data science categories.

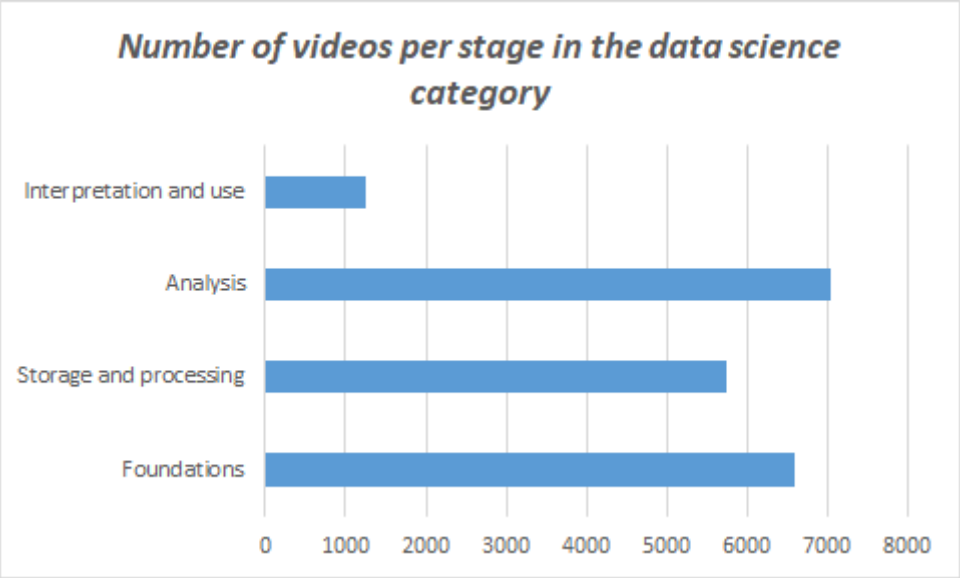


**Figure 1: Number of videos per stage in the data science category**

**Training Delivery through Data Science Lectures at VideoLectures.NET portal**

Among the videos in the data science category, over 1200 have been published since the beginning of the EDSA project (in years 2015-2017). The number of views of all videos in the data science category is 2 632 318. The next table presents the top-ten videos that were published in 2015-2017. A longer list is given in the appendix. Some videos created before EDSA have been viewed more often. The appendix shows the top videos in the data science category including these earlier ones.

**Table 3: Top ten Data Science Videolectures published in 2015-2017**

| Date | Title | Views |
|---|---|---|
| 7/28/2015 | Deep Reinforcement Learning | 10267 |
| 5/27/2016 | Incorporating Structure in Deep Learning | 1763 |
| 5/27/2016 | Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding | 1447 |
| 7/28/2015 | Basics of Computational Reinforcement Learning | 1426 |
| 5/27/2016 | Deep Robotic Learning | 1315 |
| 2/10/2016 | Fast R-CNN | 1259 |
| 12/5/2015 | Two high stakes challenges in machine learning | 1135 |
| 8/31/2016 | Is Deep Learning the New 42? | 730 |
| 2/10/2016 | Multi-Task Recurrent Neural Network for Immediacy Prediction | 713 |
| 5/27/2016 | Convergent Learning: Do different neural networks learn the same representations? | 712 |

### *3.2 Webinars, MOOCs, Self-Learning Courses*

Each partner selected the most suitable platform for online learning. After the first MOOC in EDSA was delivered on Coursera, partners switched to FutureLearn. It appears to be a suitable platform and also was recommended as a European platform. The results of learning analytics carried out for the MOOCs of EDSA can be found in the parallel deliverable D3.5.

Compared to the previous D3.2 report, ODI and OU, who had formerly focused on F2F courses, have now included MOOCS or online courses. In total, 65 865 participants were reached in the second half of the project, and 115 954 in total. Persontyle is about to launch two MOOCs on FutureLearn at the end of the project.

#### 3.2.1 MOOCs at Southampton

In line with the revisions to the EDSA curriculum made in D2.2, the University of Southampton implemented the first FutureLearn MOOC within the EDSA project, releasing the course "Introduction to Linked Data and the Semantic Web" in April 2016 as the delivery mechanism for the relevant module in Version 2 of the EDSA curriculum presented in D2.2 (M18). FutureLearn was chosen for the MOOC due to the availability of existing resources from the EUCLID project, and as test case for the FutureLearn platform in the project. This has helped us to establish a number of guidelines for developing a MOOC on this platform and allowed us to prepare learning analytics for future MOOCs based on the data received.

**Table 4: MOOCs by Southampton**

| | |
|---|---|
| Title | Introduction to Linked Data and the Semantic Web |
| Course characteristics | Online course, interactive exercises, quizzes, community discussions, lecturer feedback |
| Stage | Storage & processing |
| Target group | DM: Database managers/administrators |
| Experience | Practitioners |
| Level | Basic |
| Length | 3 weeks |
| Delivery | Online (FutureLearn MOOC) |
| Start date of delivery | 11/04/2016 |
| Number of registered participants since D3.2 | 4 715 |

### 3.2.2 MOOCs at TU/e

The first MOOC produced by TU/e in the beginning of EDSA was "Process Mining: Data Science in Action"[1]. It concentrated on the theoretical side of process mining and ran on the Coursera platform. In the meantime two further MOOCs were produced within the EDSA project:

- "Introduction to Process Mining with ProM"[2]
- "Process mining in Healthcare"[3]

The first of them focuses on the free and open-source process mining tool ProM. The course gives hands-on instructions how to use the tool, providing ample data sets to the students. The course also explains how to convert custom data to be used by the tool.

The second new MOOC focuses on a specific application area where TU/e has a strong expertise: healthcare. This MOOC is created in collaboration with other researchers in the area in order to present both the theories (although in less detail than in the other two TU/e MOOCs), as well as ample examples of process mining applications in the healthcare domain.

All three MOOCS complement each other and address different target groups. Furthermore, the MOOCs can be followed in every order desired, with little overlap and redundancy. All three MOOCs use the open source process mining software ProM, and provide ample data sets, both artificial and real. The two new MOOCs run on the FutureLearn platform, which the project decided to prefer as a European provider of e-learning platforms. The number of participants since July 2016 reached by the three MOOCS is 60 950.

**Table 5*:* MOOCs by TU/e**

| Title | Process Mining: Data Science in Action | Introduction to Process Mining with ProM | Process Mining in Healthcare |
|---|---|---|---|
| Course characteristics | | | |
| Stage | Analysis | Analysis | Analysis |
| Level | Basic | Basic | Basic |
| Length | 8 weeks | 4 weeks | 4 weeks |
| Platform | Coursera | FutureLearn | FutureLearn |
| Delivery | | | |

[1] http://courses.edsa-project.eu/course/view.php?id=2

[2] http://courses.edsa-project.eu/course/view.php?id=34

[3] http://courses.edsa-project.eu/course/view.php?id=71

| | | | |
|---|---|---|---|
| Start date of delivery | 07/10/2015 | 07/07/2016,<br>8 sessions run so far | 07/08/2017,<br>2 sessions run so far |
| Number of registered participants since D3.2 | 46.244 | 12.799 | 1.907 |

Note that the self-study module on "Process Mining" hosted on the EDSA portal is only a small part of the Coursera MOOC reported here. The self-study course insists on abstract ideas, introducing the purposes and explaining in an abstract way the methods and the theory of process mining.

### 3.2.3 MOOCs at OU

Following the feedback from the M18 review of the project, our focus was shifted from producing bespoke courses, to the aggregation and curation of a wider variety of courses that address different audiences of various levels of expertise in data science. These courses are produced not only by members of the EDSA team, but also from organisations external to the project (Higher Education institutes and other). Particular focus was given on the aggregation of MOOCs, as these are usually of high quality, they are open to the public and attract the attention of large numbers of learners.
As an indication of the different types of courses that EDSA is looking at, two MOOCs offered by the OU shall be mentioned here:

- Learn to Code for Data Analysis
- Smart Cities

As these MOOCs have not been produced by members of the EDSA team, their numbers of learners are not eligible towards the KPIs of the project.
The course entitled "Learn to Code for Data Analysis" is offered on the FutureLearn platform as a MOOC[4] and also on the OU's OpenLearn repository as an open course without a fixed start and end date.[5] This is a basic level course, targeting learners that are unfamiliar or have very little experience with programming. The course teaches learners how to write their own computer programs, how to access open data, clean it and analyse it, as well as how to produce visualisations. Learners also learn how to write up and share analyses of datasets, either privately or publicly. Learners look at real datasets from the World Health Organisation, the World Bank and other organisations and are encouraged to discuss these datasets and their analyses with their fellow learners, in order to build a community of researchers around these and other datasets. The course was selected due to its potential to introduce essential data science topics to an audience that does not necessarily have a computer science background.

The course entitled "Smart Cities" is available as a FutureLearn MOOC.[6] This course is again at a basic level, but targets a different audience, namely learners with an interest in the application of new technologies to city challenges and smart city solutions. The course helps learners to navigate their own path through the complex landscape of smart cities, hear from smart city innovators and entrepreneurs, city leaders, communities and business, as well as connect with other learners from around the world

[4] http://courses.edsa-project.eu/course/view.php?id=64
[5] http://courses.edsa-project.eu/course/view.php?id=65
[6] http://courses.edsa-project.eu/course/view.php?id=51

to reflect on issues facing smart cities of different sizes and situations. Like in the "Learn to Code" MOOC, the community building element is a core objective of this course. Besides this critical element, the "Smart Cities" MOOC has been selected for inclusion in the EDSA courses portal as it addresses a new emerging domain that has exciting novel applications for data scientists.

### 3.2.4 E-learning at ODI

After receiving the feedback from the M18 review of the project, the ODI have focused less on face-to-face training and more on creating online content. Taking a blended approach to training increases the geographical reach and provides the participants with an on demand service. The ODI's courses differ in lengths, which gives prospective participants the option to choose a format that suits their learning style best. Participants who wish to learn more and learn at their own pace are encouraged to go to the EDSA courses portal to try the technical MOOCs available.

At ODI we constantly monitor engagement and feedback of their practical courses in order to adapt the content to better serve users' needs, including the adaptation of materials that were delivered as part of EDSA. By combining this with the demand analysis for skills that are produced via the dashboard, we are able to constantly evaluate the relevance of their content.

When running the initial pilot of the course "Finding Stories in Open Data" in early 2017, some participants had problems with the pace of the course, as they were at different levels of ability, and as a consequence needed more or less time. This was counterbalanced by offering an on-demand service through eLearning. This enables participants of varying levels of ability to undertake modules that are of specific interest or relevance, whilst at the same time managing the pace at which they journey through the content. Since the feedback has shown more engagement and interest from participants in eLearning, face-to-face delivery was abandoned.

**Table 6*:* E-learning courses at ODI**

| | |
|---|---|
| Title | Finding Stories in Open Data |
| Course Details | e-learning |
| Stage | Interpretation & use |
| Sector | Media, IT, Humanities, |
| Target group | BE<br>Journalists, project managers, developers, writers, artists, producers, communications managers, press relations, presenters and consultants. |
| Experience | Students or practitioners. |
| Level | Basic |
| Length | On-line |
| Delivery since D3.2 | Yes |

| | |
|---|---|
| Start date of delivery | On-going |
| Number of participants | 200+ |

### *3.3 Face-to-Face Trainings*

The face-to-face training courses delivered reached 2 764 persons in the reporting period and a total of 4 782 against the planned target of 2 500. Part of the courses are offered to paying customers, typically practitioners, the others are given to students at the university. The number of courses for practitioners has dropped, since partners shifted efforts to other tasks or switched to online courses and MOOCs.

Feedback for commercial face-to-face courses are obtained from registration and feedback forms. As the partners use different forms, the data presented here varies with the partner. For academic courses such data is typically not collected.

A special marketing action to promote EDSA is a series of bootcamps. A bootcamp consists of two F2F days with presentations and many practical exercises in machine learning and deep learning. The target audience are practitioners. The bootcamps are conducted by Persontyle and take place at the location of different partners: so far in the UK and Germany, next in Sweden and Slovenia.

#### 3.3.1 Courses at Fraunhofer IAIS

At Fraunhofer IAIS courses for data scientists have been offered since 2014. Step-by-step a comprehensive data scientist training programme was established, with contributions from several other institutes in Fraunhofer's Big Data Alliance (see D3.2). New courses developed in the second half of the project are:

- Deep learning (2 days): This course uses TensorFlow to induce convolutional neural networks, recurrent neural networks, sequence-to-sequence modelling for applications in text and image processing
- Smart data and big data for industry 4.0 (4 days): This course gives an introduction to engineering data-driven business models, data-driven optimization of industrial production & smart sensors, sensor data acquisition and process modelling, and knowledge discovery in industrial data.
- Certified data analyst (4 days): This course leads to the certificate "data analyst". In combination with the certificate "data scientist basic level" participants can reach the second level of our certification program for professional data scientists. The course is an extension of our popular course on "Basic Data Analytics".

In general, we offer a course twice a year, adding further deliveries as demand increases and capacities permit. In the first half of EDSA the most popular courses were "Basic Data Analytics", "Big Data Analytics" and "Big Data Architecture". Now the course for the basic level certificate, which also targets project leaders and business experts, is at the top with demand surpassing our capacities. Also, the new course on "deep learning" is in high demand. We are reacting to this demand by integrating deep learning into a new certificate course on machine learning.

The next figure shows the number of participants reached with our top three courses, all of which were launched before EDSA, the certificate course "Data Scientist Basic Level", and the three new courses introduced in the reporting period.

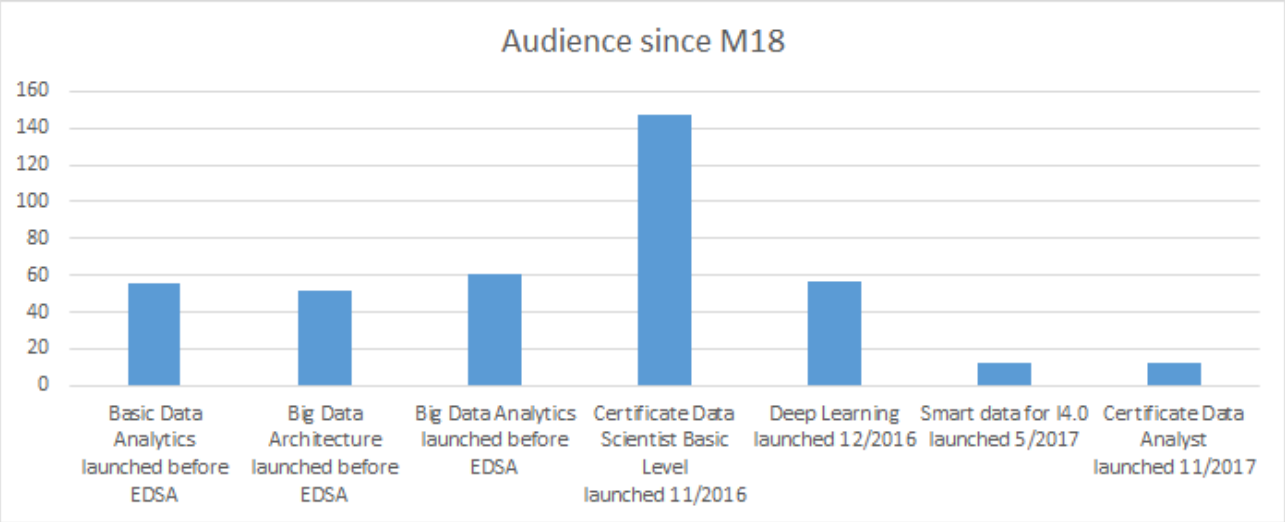


**Figure 2: Audience of most popular and new courses since M18**

The next table contains the details for these courses.

**Table 7: Popular and new F2F-courses at Fraunhofer**

| Title | Certificate Data Scientist Basic Level | Certificate Data Analyst | Basic Data Analytics | Big Data Architecture | Big Data Analytics | Deep Learning | Smart data for I4.0 |
|---|---|---|---|---|---|---|---|
| Course characteristics | | | | | | | |
| Stage | Storage & processing, Analysis, Interpretation & use | Analysis | Analysis | Storage & processing | Storage & processing, Analysis | Analysis | Storage & processing, Analysis |
| Sector | | | | | | | Manufacturing |
| Target group | | Data analysts | Data analysts | System engineers | System engineers | Data analysts | |
| Experience | Practitioners | Practitioners | Practitioners | Practitioners | Practitioners | Practitioners | Practitioners |
| Level | Basic | Advanced | Advanced | Advanced | Advanced | Advanced | Advanced |
| Length (in days) | 5 | 4 | 3 | 2 | 2 | 2 | 4 |
| Delivery since D3.2 | | | | | | | |
| Start date of delivery | 7/11/16 | | 24/10/16 | 24/10/16 | 27/10/16 | | |
| | 14/11/16 | | 30/11/16 | 28/11/16 | 5/12/16 | 7/12/16 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 16/01/17 | | 20/02/17 | | | 8/02/17 | |
| | 13/02/17 | | 13/03/17 | | 22/03/17 | 28/03/17 | |
| | 15/05/17 | | | | | | 17/05/17 |
| | 19/06/17 | | | 13/06/17 | | | |
| | 25/09/17 | | | | | | 28/08/17 |
| | 09/10/17 | | 16/10/17 | 17/10/17 | 19/10/17 | 4/10/17 | |
| | 20/11/17 | 27/11/17 | | 28/11/17 | 30/11/17 | 8/11/17 | |
| Number of participants since D3.2 | 147 | 12 | 56 | 52 | 61 | 57 | 12 |

The remaining courses at our institute, summarized in the next table, have attracted significantly less participants. The total number of participants reached with all courses since D3.2 is 435. The courses on big data architectures, big data analytics and social media analytics correspond to the syllabus of the EDSA curriculum.

**Table 8: Other F2F-courses at Fraunhofer**

| Course | Deliveries since D3.2 | Participants |
|---|---|---|
| Big Data Business Potentials | 1/02/17 and 5/12/17 | 13 |
| Visual Analytics | 2/11/16 and 7/11/17 | 16 |
| Social Media Analytics | 23/11/16 | 6 |
| Linked enterprise data integration | 9/12/16 | 3 |

“Social Media Analytics” and “Smart Data for Industry 4.0” are the only sector-specific courses. The course for the basic level certificate and the course on big data potentials are the only basic level courses. Advanced level courses focus on analysis, storage and processing and correspondingly target practicing data analysts and big data systems engineers. The number of participants is limited to 10, 12 or 15, depending on the degree of interaction desired in the course. The courses are either open and take place at our campus or they are commissioned by a company and delivered at their premises. The participants counted above went to campus courses; on-premise courses had approximately the same number of participants.

The percentage of female participants lies around 16%, with a peak in the basic level certificate course. This has not changed since D3.2. As shown in the next figure, most participants come from the public sector, with about half of them from other Fraunhofer Institutes. Other top sectors are telecommunications, data and information services, professional services, construction and engineering, finance, insurance and real estate.

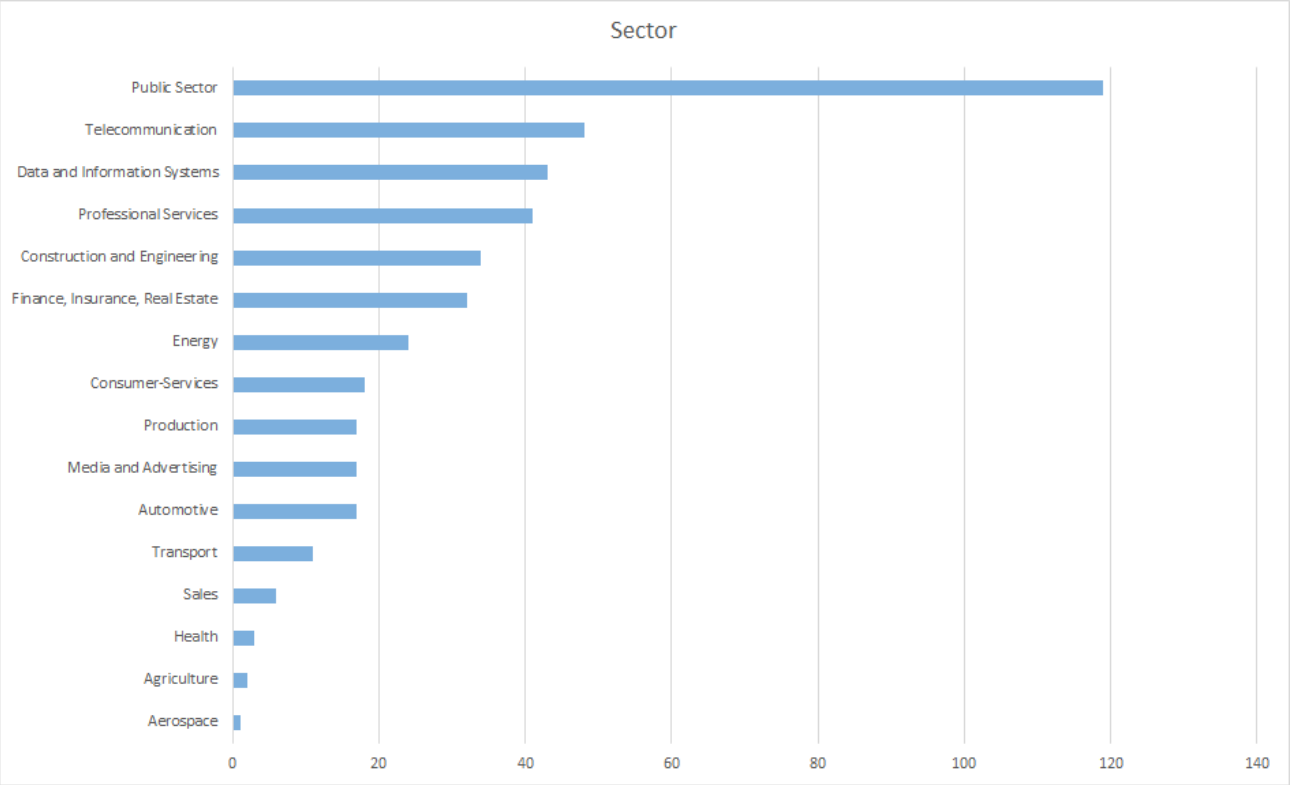


**Figure 3: Sectors of the participants**

The basic level certificate course, which wants to create a common understanding between different data science practitioners, targets a heterogeneous group of practitioners. As shown in the next figure, most participants work as data analysts, many are decision makers or heads of projects, but there are also business developers, data managers and software engineers. In contrast, the deep learning course reaches mostly data analysts, but, interestingly, also decision makers and heads of projects and business developers.

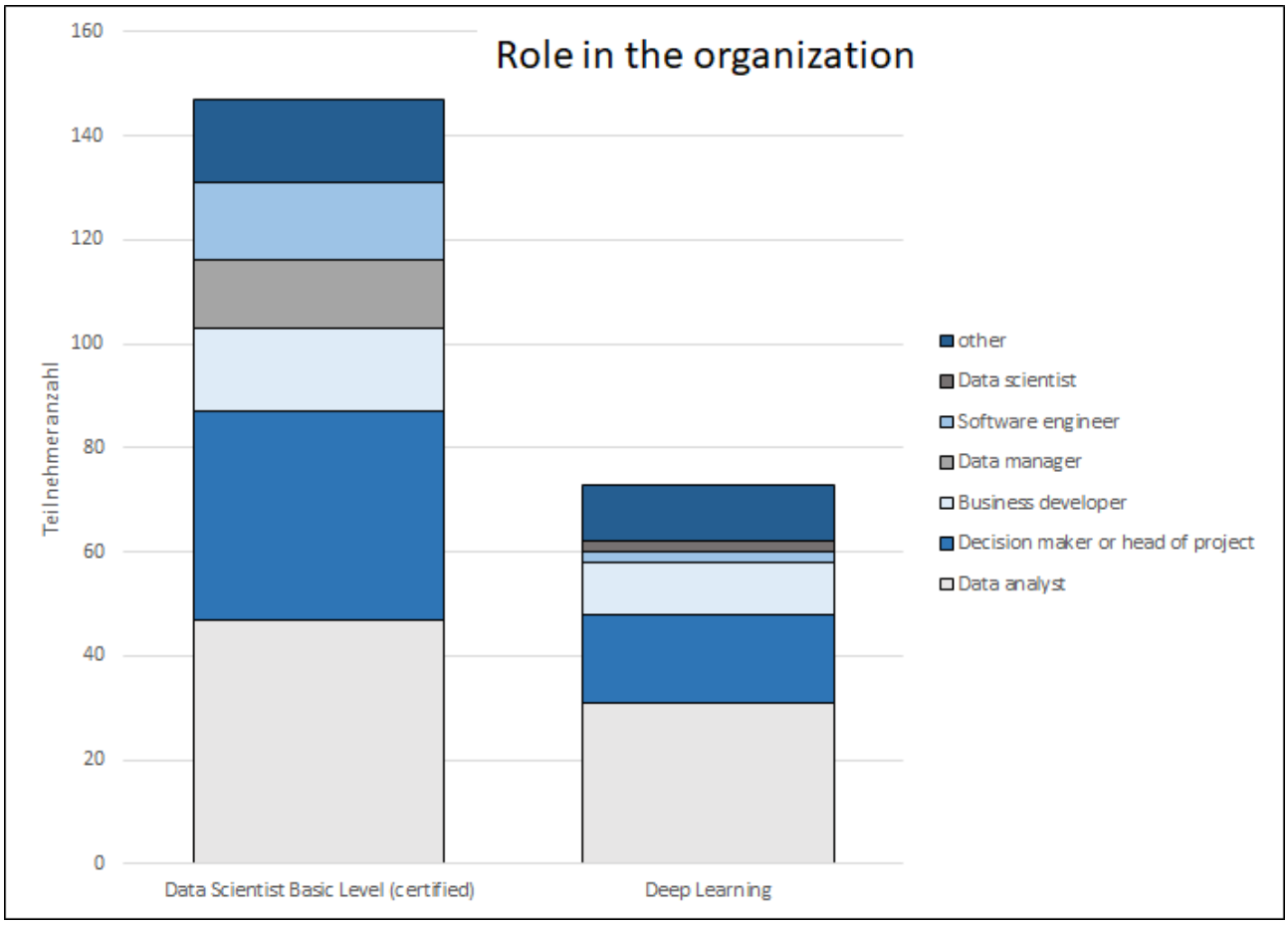


**Figure 4: Roles of the participants**

**Further plans**

Our certification program for data scientists will be extended further in 2018. In addition to the "Basic Level Certificate" and the "Certified Data Analyst" there will be a new certificate "Specialist for Machine Learning". The exam will test theoretical knowledge and practical skills, which will be provided by a group of new blended courses. The mandatory course includes deep learning and other current methods of machine learning. It is followed by optional application-specific courses. All courses start face-to-face and end with a practical phase, where the participants analyse big data sets in a distributed machine learning platform. The new courses will contain material from existing courses on deep learning and social media analytics.

### 3.3.2 Courses at Persontyle

Machine learning and deep learning are not only the top skills this year but, based on hiring demand and the potential for salary growth, it's poised to be the Europe's top skill in the future as well. To fill the demand and talent gap across Europe, we developed a dedicated workshop and a bootcamp:

- Deep learning workshop
- EDSA bootcamp

They were conducted six times with a total of 179 participants from across the Europe. Both are based on the feedback received from industry and EDSA demand analysis.

The deep learning workshops are designed to practically learn everything a practitioner needs to design, train, and integrate neural network-powered artificial intelligence into applications with widely used open-source frameworks. The workshops were targeted for researchers, developers, hackers, postgraduate students, data scientists, quants, or data analysts that already know about machine learning and have experience in programming.

The primary aim of the EDSA bootcamps is to address a skills gap, providing the training and practical understanding of machine learning necessary to meet the demand for specialist roles in Europe. The two days of the bootcamp are designed to get industry professionals and researchers started on their machine learning journey. Practical Machine Learning Bootcamp covers conceptual and applied foundations of the subject. Topics covered include machine learning theory, types of learning, techniques, models and methods. Labs are developed to practically learn how to use the R programming language and packages for applying the main concepts and techniques of Machine Learning.

**Table 9: Courses at Persontyle**

| Title | Deep Learning Workshops | Machine Learning Bootcamp for Data Scientists |
|---|---|---|
| Course characteristics | | |
| Stage | Analysis | Analysis |
| Target group | Developers, Data Scientists, Analysts and Technical Architects | Developers, Data Scientists, Analysts and Technical Architects |
| Experience | Practitioners | Practitioners |
| Level | Advanced | Advanced |
| Length | 1-2 days | 2 days |
| Delivery since D3.2 | | |
| Start date of delivery | Sep 15-16, 2016, London, UK<br>Jan 24-25, 2017, London, UK<br>06 Mar 2017, Copenhagen Denmark<br>16-17 May 2017 Munich, Germany | June 5-6, 2017, London, UK<br>August 16-17, 2017, Germany |
| Number of participants | 152 | 27 |

Participants attended our workshops and bootcamps from all over Europe, bringing with them a range of experience in Machine Learning and Data Science – from absolute beginners to experienced

practitioners. Learners reported that the blend of practical tasks and theory was well-balanced, and that they appreciated the chance to attempt a variety of methods in R.

When asked, all said they intended to put the new knowledge gained from the bootcamp to practical use, and would recommended EDSA's training to others.

> "*Even though I had some knowledge about machine learning, the course gave me a deeper understanding about it and taught me some pitfalls that I need to be aware of.*"
>
> *"My expectations were overreached. I wish more bootcamps, more online courses in this quality"*

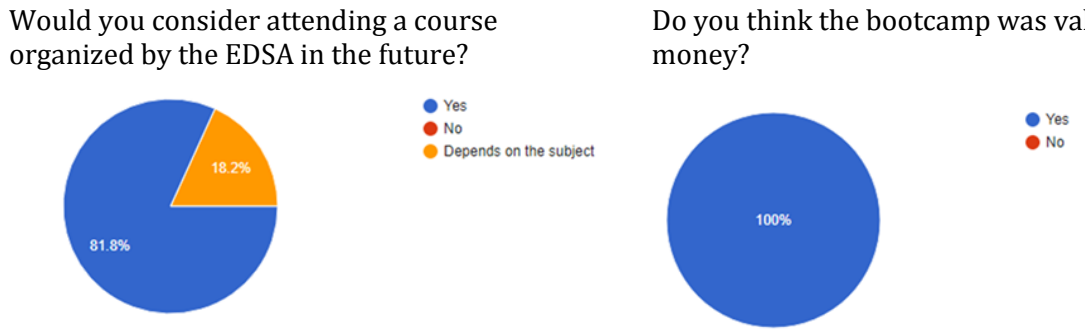


**Figure 5: Feedback from participants of the bootcamp**

**Further plans**

We will continue to deliver these bootcamps across major cities of EU with the support of industry and university partners, starting with Slovenia and Sweden. Also, these F2F instructor-led workshops and the feedback of the attendees have helped us in designing two new EDSA MOOCs.

The demand analysis and the feedback obtained from the F2F instructor-led workshops and the bootcamp have helped us in designing two new EDSA MOOCs in partnership with Future Learn:

- Advanced Machine Learning
- Deep Learning for Data Scientists

The MOOCs cover advanced Machine Learning respectively Deep Learning techniques used in data science, and the roles they play within a large-scale data projects. We look in at a large number of advanced supervised, unsupervised, semi-supervised and reinforcement learning algorithms, examining how these are related to the simpler techniques they are built from and how models generated by them can be tuned and evaluated. We also look at feature engineering, dealing with missing data and analysis of sufficiency of data. The MOOCs are balanced between theory and practice, and participants will gain experience applying all algorithms covered.
Both MOOCs will be ready for delivery at the end of the project.

### 3.3.3 Courses at ODI

As explained above the ODI have focused less on face-to-face training and more on creating online content, as a reaction to the M18 project review. The course "Finding Stories in Open Data" was turned into an e-learning course and no longer delivered face-to-face. "Open Data in practice", another course designed for face-to-face training, was not delivered in the reporting period.

A third course is "Open Data in Practice". The ODI are constantly updating its content to ensure its market relevance, and one way in which we have maintained Open Data in Practice's relevance is by

testing the content on a wider audience. The ODI have achieved this by extending its reach to beyond European boundaries. From exploiting the content produced in the project, we have been able to run courses internationally. The ODI have run instances of this course in 2017 with an audience of 50 practitioners trained. As the feedback for this course is unknown and to ensure our recommendations for ongoing improvements are validated, the ODI will need to run one more instance within the same environment.

**Table 10: Face-to-face courses at ODI**

| | |
|---|---|
| Title | Open Data in Practice |
| Course Details | Face to Face |
| Stage | |
| Sector | Applicable across all public, private and third sectors. |
| Target group | All roles:<br>Managers, technologists, all data roles, directors and those working in knowledge and transparency. |
| Experience | Practitioners |
| Level | Advanced |
| Length | 3 days |
| Delivery since D3.2 | Yes |
| Start date of delivery | May 2017 |
| Number of participants | 50 |

### 3.3.4 Courses at Southampton

In addition to the MOOC "Introduction to Linked Data and the Semantic Web" mentioned above, the University of Southampton began delivering a face-to-face Data Science MSc course in September 2015. Two new and central modules in this course are related to corresponding modules in the EDSA curriculum: Foundations of Data Science and Data Visualisation. Each module was implemented based on the curricula designed for EDSA.

In 2015, 15 students were registered on the MSc programme itself, followed by 31 in 2016 and 51 in 2017. The two new modules were also open to students of similar courses including computer science and artificial intelligence. The total number of participants reached in these courses is 230.

**Table 11: Courses at Southampton**

| Title | Foundations of Data Science | Data Visualisation |
|---|---|---|
| Course characteristics | | |
| Stage | Foundations | Interpretation & use |
| Experience | Students/Graduates | Students/Graduates |
| Level | Basic | Basic |
| Length | 1 semester (3.5 months) | 1 semester (3.5 months) |
| Delivery since D3.2 | | |
| Start date of delivery | 01/10/2016<br>28/09/2017 | 25/01/2017<br>29/01/2018 |
| Number of participants | 2016/17 58<br>2017/18 71 | 2016/17 29<br>2017/18 72 |

Based on our experiences of delivering both modules, we have revised the syllabus for "Foundations of Data Science" as part of D2.2. This is based on the experiences of the module teachers after the first instance of delivery, with a particular need to reduce the number of topics and highlight more particular aspects of the data science pipeline.

**Further plans**

As a further variation of their postgraduate courses, Southampton are currently preparing a number of Continuing Professional Development (CPD) courses in data science, as outlined in their exploitation plan in D5.3. The first, 'Fundamentals of Data Science' online CPD course went live in Autumn 2016, and has been followed with data science for marketing with further specialised courses planned in finance and healthcare. Each course will run for four weeks, will be offered online through the Canvas platform, and will be targeted towards professionals seeking to upskill in order to meet the skills requirements of data science jobs.
Southampton are also involved in the preparation of the ESWC Summer School and as with the OU will use this as an opportunity to test elements of the data science curriculum.

### 3.3.5 Courses at JSI

JSI continued to provide periodical internal trainings. In the reporting period this amounted to 100 participants. They can be seen as a source of topics and material for future EDSA modules.

**Table 12: Courses at JSI**

| Title | Workflow at Videolectures | Big Data in finance | New tools for data analytics | Job analysis and training analysis in the area of Data Science | Big data skills for official statistics |
|---|---|---|---|---|---|
| Course characteristics | | | | | |
| Stage | Analysis | Analysis | Analysis | Analysis | Analysis |
| Sector | | | | | |
| Target group | IT+DA: computer scientists, data analysts | IT+DA: computer scientists, data analysts | IT+DA: computer scientists, data analysts | IT+DA: computer scientists, data analysts | IT+DA: statisticians, computer scientists, data analysts, |
| Experience | Student | Student | Student | Student | Student |
| Level | Basic | Basic | Basic | Basic | Basic |
| Length | 1 day | 1 day | 1 day | 1 day | 1 day |
| Delivery since D3.2 | | | | | |
| Date of delivery | 14.12.2016 | 12.10.2016 | 19.10.2016 | 9.11.2016 | 04.19.17 |
| Number of participants | 20 | 20 | 20 | 20 | 20 |

Our experience shows that having internal trainings is a useful way to detect and cover the gap of formal education. In particular, the new and emergent technologies and methods presented at the internal trainings can be exploited as sources for extension of EDSA training materials, PhD and Master educational programs (in particular, at Jožef Stefan International Postgraduate School).

### 3.3.6 Courses at KTH

KTH provides courses that are part of the master programs “Software Engineering of Distributed Systems” at KTH and “Cloud Computing and Services “and “Data Science” at EIT Digital. The reported courses are parts of the “Distributed Computing” and “Data Intensive Computing” modules in the EDSA curricula. The courses are also popular among students outside the programs with 30-150 registered participants. All courses (except “Scalable Machine Learning and Deep learning” course) are video recorded and put on EDSA portal and available via YouTube. The number of participants since July 2016 is 547. The next table is split into two parts for readability purposes.

**Table 13: Courses at KTH**

| Title | Distributed Systems, Part 1 | Distributed Artificial Intelligence and Intelligent Agents | Programming Web Services | Distributed Systems, Part 2 |
|---|---|---|---|---|
| Course characteristics | | | | |
| Stage | Storage & processing | Storage & processing | Storage & processing | Storage & processing |
| Target group | IT:<br>Students of KTH and EIT Digital Master programs, IT prof | IT:<br>Students of KTH and EIT Digital Master programs, IT prof | IT:<br>Students of KTH and EIT Digital Master programs, IT prof | IT:<br>Students of KTH and EIT Digital Master programs, IT prof |
| Experience | Students | Students | Students | Students |
| Level | Master/advanced | Master/advanced | Master/advanced | Master/advanced |
| Length | 1.5 months | 1.5 months | 1.5 months | 1.5 months |
| Delivery since D3.2 | | | | |
| Start date of delivery | 20/01/2017 | 30/19/2016<br>30/10/2017 | 20/01/2017 | 25/03/2017 |
| Number of participants | 2017: 45 | 2016: 52<br>2017:53 | 2017: 44 | 2017: 42 |

| Title | Data Intensive Computing Foundations | Scalable Machine Learning and Deep learning | Advanced Topics in Distributed Computing |
|---|---|---|---|
| Course characteristics | | | |
| Stage | Storage & processing | Storage & processing | Storage & processing |
| Target group | IT:<br>Students of KTH and EIT Digital Master programs, IT prof | IT:<br>Students of KTH and EIT Digital Master programs, IT prof | IT:<br>Students of KTH and EIT Digital Master programs, IT prof |
| Experience | Students | Students | Students |
| Level | Master/advanced | Master/advanced | Master/advanced |

| | | | |
|---|---|---|---|
| Length | 1.5 months | 1.5 months | 1.5 months |
| Delivery since D3.2 | | | |
| Start date of delivery | 01/09/2017 | 30/10/2017 | 01/09/2016 |
| Number of participants | 2017: 150 | 2017: 130 | 2016: 31 |

Our experience shows that clustering related courses into an educational module is positive for the learners. Therefore we updated the EDSA course curricula so that all the above mentioned courses are included either into module "Distributed Computing" or "Data Intensive Computing".

After the review meeting we were also focused on aggregation and curation of external MOOCs.

### 3.3.7 Courses at TU/e

As a University, TU/e provides courses to students, and has a broad and increasing variety of courses related to data science. TU/e held the following data science courses at the Master level. The total number of participants is 1223, a significant increase since the previous reporting period. The table is split into two parts for readability purposes.

**Table 14: Courses at TU/e**

| Title | Advanced process mining | Advanced Data Analysis | Web information retrieval and data mining | Introduction to process mining | Visualization |
|---|---|---|---|---|---|
| Course characteristics | | | | | |
| Stage | Analysis | Analysis | Foundations | Analysis | Interpretation & use |
| Experience | Students | Students | Students | Students | Students |
| Level | Advanced | Advanced | Basic | Basic | Basic |
| Length | 17.5 days in 2.5 months | 17.5 days in 2.5 months | 17.5 days in 2.5 months | 17.5 days in 2.5 months | 17.5 days in 2.5 months |
| Delivery since D3.2 | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Start date of delivery | 16/11/16 | 05/09/16 | 05/09/16 | 05/09/16 | 16/11/16 |
| Number of participants | 135 | 114 | 105 | 136 | 231 |

| Title | Statistics for big data | Foundations of data mining | Principles of data protection | Big data and experiments for urban analysis | Data engineering |
|---|---|---|---|---|---|
| Course characteristics | | | | | |
| Stage | Foundations | Foundations | Storage & processing | Analysis | Storage & processing |
| Sector | | | | | |
| Target group | | | | | |
| Experience | Students | Students | Students | Students | Students |
| Level | Basic | Basic | Basic | Basic | Basic |
| Length | 17.5 days in 2.5 months | 17.5 days in 2.5 months | 17.5 days in 2.5 months | 35 days in 5 months | 17.5 days in 2.5 months |
| Delivery since D3.2 | | | | | |
| Start date of delivery | 24-4-2017 | 6-2-2017 | 5-9-2016 | 6-2-2017 | 24-4-2017 |
| Number of participants | 65 | 108 | 191 | 31 | 107 |

Among the 1223 participants, 573 were not Dutch, 312 were female (about 26%). As usual in this field, there was a majority of male participants. The proportion varied among topics, and for instance "advanced data analysis" almost had a 50-50 ratio of female participants. No clear pattern emerges from the graphs below, showing the proportion of female participants in the total of participants per course.

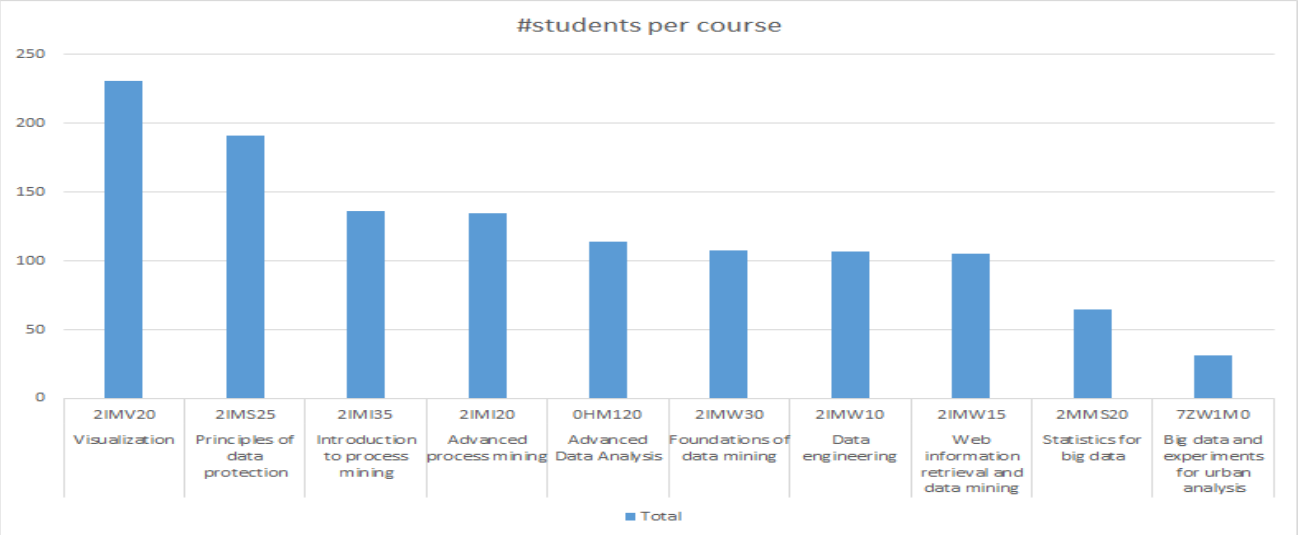


**Figure 6: Number of students per course at TU/e**

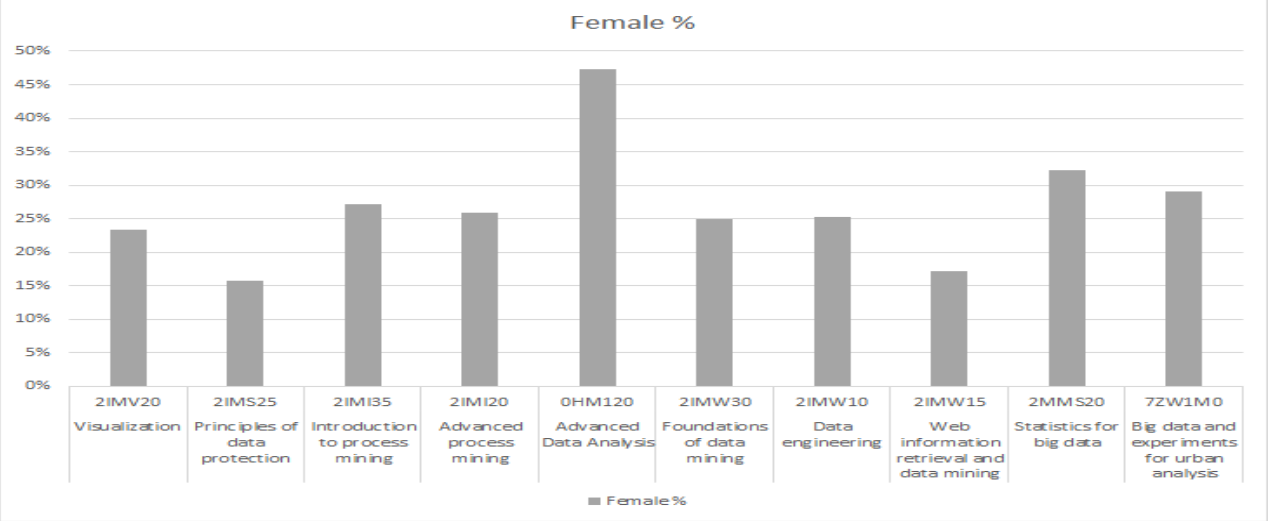


**Figure 7: Proportion of female participants**

The graph below shows the proportion of participants from other countries than the Netherlands. We can note that "introduction to process mining" had a large majority (of 71%) of foreign students. On average 47% of the students in the courses was not Dutch.

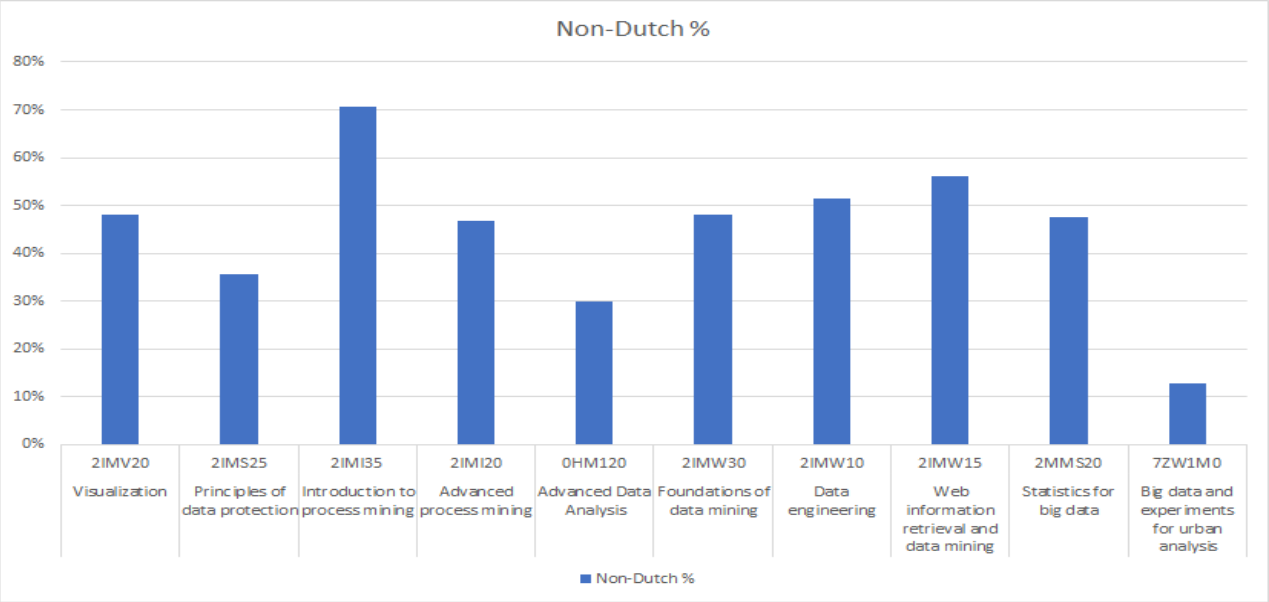


**Figure 8: Proportion of non-Dutch participants**

## 4. Summary and Conclusions

The preceding section presented a range of courses delivered by the partners in the second half of the project. Both, MOOCs/e-learning and F2F courses have reached far more participants than promised.

11 500 lectures and tutorials have accumulated in the video portal in the data science category, which was created during EDSA but includes videos produced before EDSA. Six different MOOCs and one e-learning course were offered in the second half of the project and 37 different F2F courses were delivered. In the second half of the project, FutureLearn became the platform for EDSA MOOCs,

F2F courses divide into academic and professional courses. As the consortium is dominated by universities, there are clearly more academic courses (24:13). In the academic courses there is a focus on basic courses (15:9). Nine academic courses focus on storage & processing, nine on analysis, four on foundations, and two on interpretation and use. In contrast to the academic courses, the courses for practitioners are mostly at the advanced level (11:2) and, like the academic courses, mostly sector-independent (11:2). Seven focus on data analysis, two on storage and processing, one on interpretation and use.

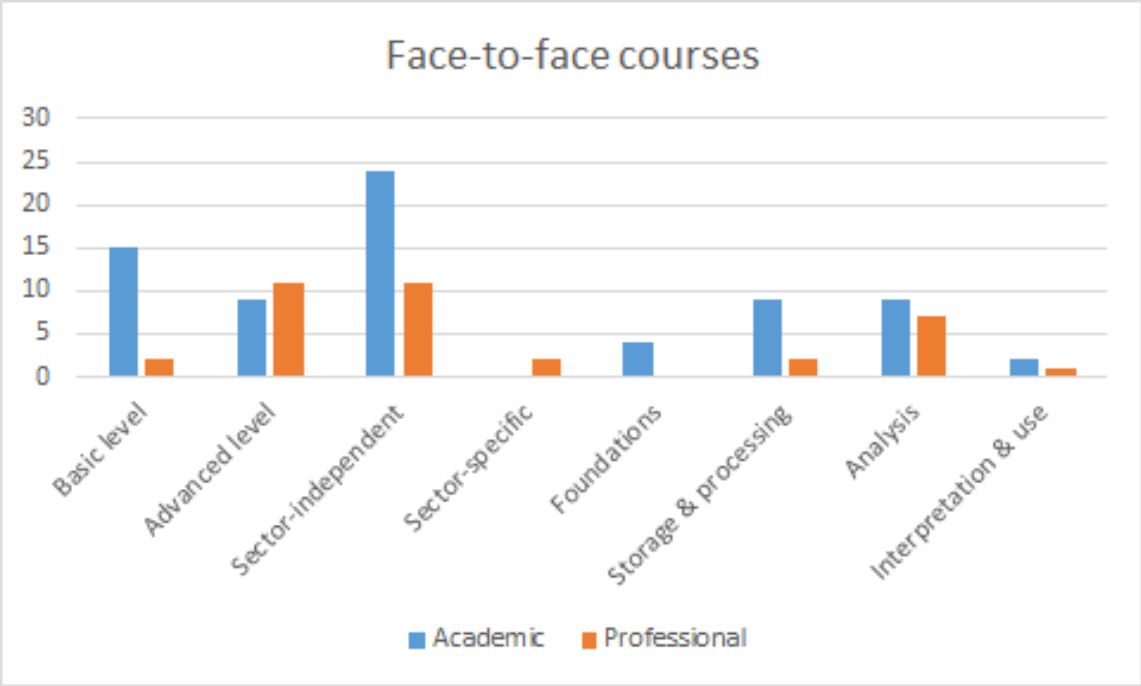


**Figure 9: Overview of face-to-face courses**

In the first half of the project feedback from course deliveries served to improve the curriculum and generate learning material in WP2. In the second half of the project, as a reaction to the project review, efforts were shifted to the collection and curation of courses produced by external organizations, thus extending the scope of EDSA. At the same time EDSA partners started to diversify their offers: from F2F to online courses, from academic to professional courses, from national to international delivery, from courses without exams to courses with certificates. This was guided by the EDSA demand analysis and feedback obtained from the participants. Reengineering courses was facilitated by reintegrating material produced for EDSA.

Judging from the videolectures recorded during the project, deep learning has been by far the most popular topic (see Table 3). Deep learning entered the EDSA curriculum in the first revision. Fraunhofer, Persontyle and KTH offer courses on this topic, and a MOOC is just being developed by Persontyle. Deep learning, together with other machine learning methods, was also the topic of the EDSA bootcamps. These two day events to promote EDSA already took place in two different countries and are planned for two more already. In the next years it will be interesting to observe whether machine learning and deep learning will continue to be seen as part of the data science profession or whether we will see another profession of "machine learning specialists" emerge.

# 5. Appendix 1: Videolectures

The next table presents the top videolectures that were published in 2015-2017.

**Table 15: Top Data Science Videolectures published in 2015-2017**

| Date | Title | URL | Views |
|---|---|---|---|
| 7/28/2015 | Deep Reinforcement Learning | http://videolectures.net/rldm2015_silver_reinforcement_learning | 10267 |
| 5/27/2016 | Incorporating Structure in Deep Learning | http://videolectures.net/iclr2016_urtasun_incoporating_structure | 1763 |
| 5/27/2016 | Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding | http://videolectures.net/iclr2016_han_deep_compression | 1447 |
| 7/28/2015 | Basics of Computational Reinforcement Learning | http://videolectures.net/rldm2015_littman_computational_reinforcement | 1426 |
| 5/27/2016 | Deep Robotic Learning | http://videolectures.net/iclr2016_levine_deep_learning | 1315 |
| 2/10/2016 | Fast R-CNN | http://videolectures.net/iccv2015_girshick_fast_r_cnn | 1259 |
| 12/5/2015 | Two high stakes challenges in machine learning | http://videolectures.net/icml2015_bottou_machine_learning | 1135 |
| 8/31/2016 | Is Deep Learning the New 42? | http://videolectures.net/kdd2016_broder_deep_learning | 730 |
| 2/10/2016 | Multi-Task Recurrent Neural Network for Immediacy Prediction | http://videolectures.net/iccv2015_chu_neural_network | 713 |
| 5/27/2016 | Convergent Learning: Do different neural networks learn the same representations? | http://videolectures.net/iclr2016_yosinski_convergent_learning | 712 |

| | | | |
|---|---|---|---|
| 12/5/2015 | Natural Language Understanding: Foundations and State-of-the-Art | http://videolectures.net/icml2015_liang_language_understanding | 692 |
| 5/27/2016 | Opening | http://videolectures.net/iclr2016_larochelle_opening | 620 |
| 5/27/2016 | Neural Programmer-Interpreters | http://videolectures.net/iclr2016_reed_neural_programmer | 571 |
| 5/27/2016 | Should Model Architecture Reflect Linguistic Structure? | http://videolectures.net/iclr2016_dyer_model_architecture | 562 |
| 2/24/2016 | It's Learning All the Way Down | http://videolectures.net/iccv2015_lecun_learning | 548 |
| 5/27/2016 | Beyond Backpropagation: Uncertainty Propagation | http://videolectures.net/iclr2016_lawrence_beyond_backpropagation | 523 |
| 2/17/2015 | Introduction to Hidden Markov Models | http://videolectures.net/mlpmsummerschool2014_artes_rodriguez_models | 493 |
| 8/31/2016 | Learning to learn and compositionality with deep recurrent neural networks | http://videolectures.net/kdd2016_de_freitas_recurrent_neural | 480 |
| 7/28/2015 | Quickly Learning to Make Good Decisions | http://videolectures.net/rldm2015_brunskill_good_decisions | 468 |
| 10/24/2016 | XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks | http://videolectures.net/eccv2016_rastegari_neural_networks | 424 |
| 5/27/2016 | Guaranteed Non-convex Learning Algorithms through Tensor Factorization | http://videolectures.net/iclr2016_anandkumar_nonconvex_learning | 422 |
| 2/23/2016 | Convex Optimization with Abstract Linear Operators | http://videolectures.net/iccv2015_boyd_convex_optimization | 414 |
| 10/24/2016 | Colorful Image Colorization | http://videolectures.net/eccv2016_zhang_image_colorization | 404 |

| | | | |
|---|---|---|---|
| 12/5/2015 | Bayesian Time Series Modeling: Structured Representations for Scalability | http://videolectures.net/icml2015_fox_structured_representations | 401 |
| 2/10/2016 | Deep Neural Decision Forests | http://videolectures.net/iccv2015_kontschieder_decision_forests | 376 |
| 12/5/2015 | Advances in Structured Prediction | http://videolectures.net/icml2015_daume_structured_prediction | 366 |
| 2/17/2015 | Machine learning for brain imaging | http://videolectures.net/mlpmsummerschool2014_varoquaux_brain_imaging | 359 |
| 7/28/2015 | Natural RLDM: Optimal and Subptimal Control in Brain and Behavior | http://videolectures.net/rldm2015_daw_brain_and_behavior | 353 |
| 2/10/2016 | Human Parsing With Contextualized Convolutional Neural Network | http://videolectures.net/iccv2015_liang_human_parsing | 337 |
| 5/27/2016 | Net2Net: Accelerating Learning via Knowledge Transfer | http://videolectures.net/iclr2016_chen_net2net | 336 |

Most popular video lecture in the data science category since the first publications in 2007.

**Table 16:** ***Top Data Science Videolectures Viewing Statistics published in 2007-2017***

| Date | Title | URL | Views |
|---|---|---|---|
| 7/2/2007 | Basics of probability and statistics | http://videolectures.net/bootcamp07_keller_bss | 70741 |
| 2/25/2007 | Machine Learning, Probability and Graphical Models | http://videolectures.net/mlss06tw_roweis_mlpgm | 43519 |
| 11/2/2009 | Markov Chain Monte Carlo | http://videolectures.net/mlss09uk_murray_mcmc | 43462 |
| 2/25/2007 | Gaussian Process Basics | http://videolectures.net/gpip06_mackay_gpb | 42390 |
| 11/2/2009 | Topic Models | http://videolectures.net/mlss09uk_blei_tm | 39460 |
| 9/15/2009 | A tutorial on Deep Learning | http://videolectures.net/jul09_hinton_deeplearn | 36013 |
| 3/13/2008 | Monte Carlo Simulation for Statistical Inference, Model Selection and Decision Making | http://videolectures.net/mlss08au_freitas_asm | 28686 |
| 2/5/2008 | Introduction to Support Vector Machines | http://videolectures.net/epsrcws08_campbell_isvm | 24096 |
| 2/25/2007 | Semisupervised Learning Approaches | http://videolectures.net/mlas06_mitchell_sla | 17273 |
| 2/25/2007 | Dirichlet Processes, Chinese Restaurant Processes, and all that | http://videolectures.net/icml05_jordan_dpcrp | 16598 |
| 8/5/2010 | Introduction to Machine Learning | http://videolectures.net/bootcamp2010_murray_iml | 16175 |
| 3/12/2009 | Challenges in Building Large-Scale Information Retrieval Systems | http://videolectures.net/wsdm09_dean_cblirs | 15390 |

| | | | |
|---|---|---|---|
| 11/2/2009 | Information Theory | http://videolectures.net/mlss09uk_mackay_it | 14660 |
| 7/2/2007 | Introduction to Machine Learning | http://videolectures.net/bootcamp07_guyon_itml | 14451 |
| 7/4/2012 | Big-Data Tutorial | http://videolectures.net/eswc2012_grobelnik_big_data | 14132 |
| 11/2/2009 | Deep Belief Networks | http://videolectures.net/mlss09uk_hinton_dbn | 13592 |
| 7/30/2009 | An Overview of Compressed Sensing and Sparse Signal Recovery via L1 Minimization | http://videolectures.net/mlss09us_candes_ocsssrl1m | 13081 |
| 2/25/2007 | Generative Models for Visual Objects and Object Recognition via Bayesian Inference | http://videolectures.net/mlas06_li_gmvoo | 12914 |
| 2/25/2007 | Text Classification | http://videolectures.net/mlas06_cohen_tc | 11848 |
| 11/2/2009 | Particle Filters | http://videolectures.net/mlss09uk_godsill_pf | 11675 |
| 9/9/2011 | Social Media Analytics | http://videolectures.net/single_leskovec_social | 11506 |
| 7/28/2015 | Deep Reinforcement Learning | http://videolectures.net/rldm2015_silver_reinforcement_learning | 10267 |
| 8/20/2007 | Introduction to bioinformatics | http://videolectures.net/mlss07_gunnar_intbio | 10224 |
| 11/24/2008 | How to Publish Linked Data on the Web | http://videolectures.net/iswc08_heath_hpldw | 9926 |
| 1/12/2011 | Optimization Algorithms in Machine Learning | http://videolectures.net/nips2010_wright_oaml | 9746 |
| 2/25/2007 | A short Tutorial on Semantic Web | http://videolectures.net/training06_sure_stsw | 9715 |

| | | | |
|---|---|---|---|
| 1/12/2011 | How to Grow a Mind: Statistics, Structure and Abstraction | http://videolectures.net/nips2010_tenenbaum_hgm | 8898 |
| 12/20/2008 | Matplotlib | http://videolectures.net/mloss08_hunter_mat | 8636 |
| 2/25/2007 | UÄenje povzemanja besedil s pretvorbo v semantiÄno mreÅ¾o | http://videolectures.net/single_leskovec_diploma | 8556 |
| 9/3/2010 | Introduction to Statistics | http://videolectures.net/cernstudentsummerschool09_cowan_is | 8512 |