

REPORT OF ACTIVITIES

Statistical Methodology
and Computing Service

1 September 2017 to 31 August 2018



2017-2018

Statistical Methodology and Computing Service
Report of Activities
1 September 2017 to 31 August 2018

Contents

Introduction.....	4
News	5
Organization	5
Personnel.....	5
Statistical Software.....	5
Active watch for tenders	5
People	6
Management and Administration	6
Consultants.....	6
Associated Consultants	7
Professors	7
Consultants at other institutes.....	7
Researchers and Research Assistants.....	7
Interns	8
Workshop Instructors.....	8
Training	9
Consulting.....	12
Servers and Statistical Software	14
Statistical Computing Server	14
Ucilla and UDS	14
Available Software.....	14
Usage Statistics.....	15
Help for students.....	15
Larger Projects	16
Survey Support	16
General interest Research Projects	16
Research Project - Work and Organizational Psychology (2016-2018)	16
Research Project - Erasmus+OLS (2015-2021)	16
Research Project - Macular degeneration (2017-2019)	16
Research project - First spin-off (2018-...)	17
Research project - Micro-array analysis for GSK (2018)	17
Course Module - FOREM (2011-...)	17
Course Modules – BioWin (2010-2018)	17
University Certificate - Junior Data Analyst (2016-...)	17
Collaboration with the LSBA.....	17
Your Project?	17

Contents

Appendix	18
List of Trainings and Seminars 2017-2018.....	18
Trainings	18
Publications and Presentations by SMCS members	20
Publications	20
Seminars and Presentations.....	20
Collaborations, Response to Tenders and Supervisions.....	21
Collaborations and Response to Tenders.....	21
Supervision of Master and PhD thesis.....	21

Introduction

Introduction

The SMCS continues its activity as a diversified service provider in the area of practical statistics. On September 2017, it has hired 1,5 FTE to develop the survey support and to help the whole team in general in all the SMCS activities and in particular, to monitor the research project opportunities (projects linked to statistics). This has lead us to answer to different calls and to develop collaborations with external partners as well as inside UCLouvain. The 5.5 to 6.5 full time equivalents of consulting staff and a network of more than 20 other specialists assist UCLouvain researchers and external clients in data collection, processing, analysis, and interpretation.

The current report gives an overview of the evolution of the activities in consulting, training and computing over the past three academic years focusing on 2017/18.

The consulting activity has been stable at around 420 projects per year. The training activity continued to grow to 57 courses, representing 168.5 days of training. These courses, focusing on practical statistics and the use of statistical software, were attended by about 880 participants for a total of about 2254 person days. The statistical computing server had about 52 registered users and the software-sharing server 50 users of statistical tools.

Since the current workload started to exceed the staff capacity, the management committee supported the appointment of an extra 0.5 FTE starting in September 2018. This will allow maintaining an efficient support to the university and the community for a medium term horizon.

Since the evolution of a more data-driven research society will lead many research institutes to upgrade their data collection and analysis capabilities, we foresee a need to extend our capabilities in the area of data technology and science. This issue was raised at the management committee meeting in spring and we will need to include it in our long-term plan such that we can provide appropriate guidance, advice, and assistance.

Following the recommendations of the users committee, the SMCS continues working on its visibility towards a community of potential users which continuously renews itself as doctoral researchers graduate and new candidates are accepted.

With respect to larger projects, the support of the language testing and training of the Erasmus+ Online Linguistic Support (OLS) continues and will last until 2020/21. The SMCS also continues working with the Croix-Rouge de Belgique on the quality assurance of blood products.

We have launched the second edition of the new certificate "Junior Data Analyst" which took place in fall 2017. Our collaboration with the Walloon career center, FOREM, continues. Our intervention consists of an intensive course in biostatistics for people with degrees or experience in the biosciences who need additional qualifications.

Two people of our team have been involved in the realization of videos in the framework of a MOOC called CRITX. It concerns critical thinking and will be open to all the study programs? For more details, please see at <https://mooc-francophone.com/cours/mooc-developper-sa-pensee-critique/>.

Our main challenge in the coming years will be handling growth. Additional demand will come because of better visibility and the increase in data driven research. The group and in particular its expert network will become bigger and maintaining good social integration and communication in the evolving team will require close attention.

For 2018/19 we are considering several new initiatives. Ideas include a summer school for researchers. Should we do this? Or do you have better ideas?

As usual, your input, assistance, support, and curiosity will be greatly appreciated.

Enjoy reading, Christian Ritter - President of the management committee.

News

News

Organization

The management committee (comité de gestion) met twice under the new ROI. We decided to switch from three meetings per year to only two and to give them specific themes. The fall meeting focuses on long term planning whereas the spring meeting concentrates on finances, project accounting, and short term planning. A continuing problem is uniting all members on these two dates. Since the new ROI specifies that some of the members of the management committee are institutional (presidents of reference institutes), reconciling the respective agendas has become more difficult and we may have to find a different way of working.

With respect to the users committee, the new format of an afternoon with public presentations of case studies organized on the same day as the official meeting restricted to the official representatives was very well-received and continued in 2018.

Personnel

Arnaud Polaris has left UCLouvain.

The SMCS hired one consultant, Vincent Bremhorst, who will start working for the team in September 2018. V. Bremhorst will help on general projects and trainings, with a specific attention for Bayesian statistics.

The SMCS could count on the help of David Stern, for projects with analysis of micro-array data, from April to July 2018.

Statistical Software

The Big5-Stat initiative is still operated by the central computing service (SGSI) and the SMCS only maintains the role of advisor and tester.

The SGSI has made the transition from Inuvika to UDS during the year. This new tool provides softwares through a kind of remote desktop connection and is very useful for softwares we do not use often or with shared costs. During the year, extensive testing was

implemented and this tool was considered good enough to be on production the year after.

Active watch for tenders

Since September 2017, we have actively monitored the different research project opportunities and this has lead us to answered to new kinds of calls : a training course in the field of Data Scientist (60 days) for Bruxelles formation, in collaboration with private partners (not obtained) and 2 collaborative research projects combining SMCS knowledge and expertise from UCLouvain professors. Those two projects have been obtained in the following topics: training evaluation for Bruxelles formation and creation of gender indicator for le Service Public Regional de Bruxelles. This will give us the opportunity to develop new fields of expertise and new collaborations with external partners as well as inside UCLouvain.

People

People

Management and Administration

Director: Nathalie Lefèvre

President of Management Committee: Christian Ritter

Administrative coordinator: Nancy Guillaume

Communications and Quality Manager: Marie-Pascale Cosse

Administrative support: Tatiana Regout

Consultants

Aurélie Bertrand (UCL/IMMAQ/SMCS – 0.5 FTE - since September 2017)

Main areas of activity: Humanities Sector (SSH), Medical Sector (SSS), specific projects in collaboration with the Institute of Statistics, Biostatistics and Actuarial Sciences



Doctorat en Sciences – Statistique (UCL, 2017)

Master en statistiques, orientation générale, à finalité spécialisée (UCL, 2011)

Bachelier en Ingénieur de gestion (UCL, 2009)

Céline Bugli (UCL/IMMAQ/SMCS – 0.8 FTE)

Main areas of activity: Medical Sector (SSS), Science and Technology Sector (SST), services for external clients.



Doctorat en Sciences – Statistique (UCL, 2006)

Master (DEA) en statistique (UCL, 2003)

Ingénieur Civil en Mathématiques Appliquées (UCL, 2001)

Lieven Desmet (UCL/IMMAQ/SMCS – 0.75 FTE)

Main area of activity: Health Sciences Sector (SSS).



Doctor of Sciences, orientation statistics (KUL, 2009)

Master in Mathematics (VUB 2004)

Master in Applied computer sciences (VUB, 1999)

Alain Guillet UCL/IMMAQ/SMCS – 1 FTE)

Main areas of activity: Medical Sector (SSS), and statistical computing.



Master professionnel Sciences, Santé, Technologie, mention Mathématiques et Applications, spécialité Ingénierie Mathématique (Université de Nantes, 2005)
DEA en informatique (Université de Nantes, 2004)

Séverine Guisset UCL/IMMAQ/SMCS – 1 FTE - since September 2017)

Main areas of activity: Humanities Sector (SSH), Medical Sector (SSS), survey support, services for external clients.



Master en statistique (UCL, 2017)

Master in European Studies (King's College London, 2007)

Diplôme de l'Institut d'Etudes Politiques de Strasbourg (Université Robert Schuman, 2006)

People

Nathalie Lefèvre (UCL/IMMAQ/SMCS – 0.5 FTE and UCL/IPSY – 0.5 FTE)

Main areas of activity: Humanities Sector (SSH) and University Administration.



Master en Statistique (UCL, 2007)

Doctorat en Sciences psychologiques (UCL, 2003)

Licence en Psychologie et en Sciences de l'Education, orientation psychologie expérimentale (UCL, 1992).

Arnaud Pollaris (UCL/IMMAQ/SMCS – 1 FTE – until February 2018)

Main areas of activity: Humanities Sector (SSH).



Master en Statistique (UCL, 2011).

Master en Sciences psychologiques, à finalité spécialisée, (UCL, 2009).

Catherine Rasse (UCL/IMMAQ/SMCS – 0.8 FTE)

Main areas of activity: Science and Technology Sector (SST).



Master (DEA) en Statistique (Gembloux, 2007).

Ingénieur agronome, orientation Génie rural (Gembloux, 2001).

David Stern (UCL/IMMAQ/SMCS – 1 FTE – April 2018 – July 2018)

Main areas of activity: DMLA project – OMIC data



Master en Biochimie et Biologie moléculaire et cellulaire (ULG, 2011)

Bachelier en Biologie (ULG, 2009)

Associated Consultants

Associated consultants are professors or consultants who collaborate with the SMCS on trainings and consulting projects but who are not formally SMCS staff.

Professors

Bernadette Govaerts (UCL/IMMAQ/ISBA)

Main areas of activity: Industrial statistics, bioinformatics.

Catherine Legrand (UCL/IMMAQ/ISBA)

Main areas of activity: biostatistics.

Christian Ritter (UCL/IMMAQ/LSBA)

Main areas of activity: Industrial statistics, statistical computing.

Consultants at other institutes

Jonathan Dedonder (UCL/IACCHOS)

Sébastien Fontenay (UCL/IMMAQ/IRES)

Nathalie Lefèvre (UCL/IPSY)

Jérôme Ambroise (IREC)

Alix Collard and Aline van Maanen (Institut Roi Albert II - Cancérologie et Hématologie)

Researchers and Research Assistants

Some researchers and doctoral students of the ISBA participate in the consulting and training activities of the SMCS. This gives them an opportunity to gain valuable experience in practical statistics.

People

Interns

David Stern, Augustus-October 2018

Workshop Instructors

The following list gives the names of the people who contributed to the training and consulting activities of the SMCS. This includes SMCS and associated consultants, other members of the ISBA, visitors, and external trainers.

Aurélie Bertrand (SMCS), Céline Bugli (SMCS), Vincent Bremhorst (ISBA), Joris Chau (ISBA), Mickaël De Backer (ISBA), Sébastien de Valeriola (ISBA), Jonathan Dedonder (IACCHOS), Sébastien Fontenay (IRES), Thomas François (ILC), Bernadette Govaerts (ISBA), Oswaldo Gressani (ISBA), Alain Guillet (SMCS), Séverine Guisset (SMCS), Alexandre Jacquemain (ISBA), Jonathan Jones (CORE), Eric Lecoutre (WeLoveDataScience!), Nathalie Lefèvre (SMCS), Rebecca Marion (ISBA), Manon Martin (ISBA), Gaëtan Montero Redondo (CORE), Gilles Mordant (ISBA), Hélène Morsomme (ISBA), Arnaud Pollaris (SMCS), Catherine Rasse (SMCS), Christian Ritter (ISBA), Xavier Rouby (Pfease), Marco Saerens (LRIM), Isabelle Thomas (CORE), Nathan Uyttendaele (ISBA), Patrick Watrin (ILC), Sven Wauters (WeLoveDataScience!), SAS Institute trainers.

Training

Training

Training is one of our main missions. In general, we focus on hands-on training in practical statistics and the use of statistical software in ways that are not covered by regular university courses. Five of the training sessions took place on the Woluwe campus. One training planned in English had to be organized in French at the request of the participants. For very advanced trainings, participants generally prefer that they be given in their mother tongue.

Eleven new training programs have been developed.

The SMCS offers the following types of training:

- **Regular SMCS trainings:** 2-3 day courses (once or twice per year) on commonly used statistical and associated software at UCLouvain (SPSS, R, SAS, SAS Enterprise Guide, JMP, Stata, NVivo, Excel, Limesurvey).

The SMCS offers many trainings about R, some as an initiation to the R language and others more advanced. Some trainings are organized to respond to users' requests. Some trainings organized to respond to user requests thereafter become Regular SMCS trainings. For example, the R training, "*Pratique de la statistique avec R*".

- **Customized trainings for specific audiences** (UCLouvain and external). Examples: "*Introduction aux modèles multi-niveaux avec SPSS*" or "*Introduction to the basic concepts of Bayesian statistics*" (organized in collaboration with IPSY but open to everyone), "*Créez votre liste de randomisation avec le logiciel JMP*" (in collaboration with IREC but open to everybody).
- **Courses sponsored by the Walloon government agencies in the context of BioWin**, a project to enhance the competitiveness of health-related companies in the Walloon region. Courses offered within this framework are advertised by BioWin and course fees are partially subsidized until end of 2018.

- **In partnership: courses taught in collaboration with SAS.** Some annual or special courses qualify for sponsoring from SAS. The course "*Formation SAS: SAS advanced programming*" belongs to this category.
- **University certificate.** In collaboration with the LSBA, the SMCS can also be involved in and/or coordinate university certificates. Example: *the second edition of the certificate "Junior Data Analyst"* - Teaching this course implied working with several external experts active in the data science business and we learned a lot about current practices.
- **Seminars and Lunch sessions (Midis du SMCS):** short (1-3h) informal demonstrations of statistical methods or software of particular practical interest.

Table 1 shows that the SMCS organized 57 courses in the academic year 2017-2018 for 168.5 teaching days. This represents an increase of 33 days of training due to, among others, the organization of new trainings with the R and SAS languages and on specialized subject like Text Mining, Life Trajectory Mining, Bayesian statistics.

Table 1. Global statistics on SMCS trainings.

	2015/16	2016/17	2017/18
Courses	38	40	57
Course Days	108.5	135	168.5
Participants UCL	448	537	660
Person-Days* UCL	826	1199	1517
Participants Ext	117	128	224
Person-Days* Ext	660.5	680	737.5

*Total effective number of training days: for instance, 9 people attending 2 training days is equivalent to 18 person-days.

Table 2 shows the number of participants from the three sectors.

Training

Table 2. Number of participants per sector

	2015/16	2016/17	2017/18
SSH	147	173	277
SSS	77	122	130
SST	65	122	90
Students	159	120	159
Central services			4

The observed fluctuations between years and sectors reflect natural variation and the organization of special events. For example, the spike for SST in 2016/17 corresponds to a special session “*La statistique, outil du chercheur et de l'ingénieur*” organized in April 2017. Compared to previous years, there is an increase in all sectors except for SST where the training mentioned above was not organized this year.

Table 3 (SSH) shows an exceptional increase in the number of SSH participants. The statistics training program reaches the entire sector. The high participation rate of IPSY-PSP stems from the fact that the IPSY Institute finances the organization by the SMCS of specific trainings. The high participation of IACS-SPLE-ESPO and IMMAQ are explained by specific seminars or training requested by these entities. The concerning trainings are often organized in collaboration with associated consultants coming from these entities. However, these trainings remain open to all.

Table 3. Origin of participants from the humanities sector (SSH)

	2015/16	2016/17	2017/18
SSH	239	214	321
IACS-SPLE-ESPO	53	61	87
ILC	13	7	19
LOURIM-LSM	14	16	15
IMMAQ	16	8	40
INCAL-FIAL	7	5	10
IPSY-PSP	42	71	98
JURI-DRT	1	5	3
RSCS-TECO	1		5
Students	92	41	44

Table 4 (SSS) shows that also in the health sciences sector courses were followed by participants from a wide range of institutes. Many clinical researchers find it difficult to attend longer trainings due to their

appointment schedules at the university clinics. We have tried to adapt the training format and have also increased the number of trainings offered in Woluwé. This probably explains the higher number of SSS participants. The organization of new trainings with the R and SAS languages or seminars on specialized subject were also greatly appreciated by SSS participants and especially in the IRSS institute.

Table 4. Origin of participants from the health sciences sector (SSS)

	2015/16	2016/17	2017/18
SSS	82	133	155
ASS	1	1	
DDUV-Ludwig Inst.	11	9	13
IONS	14	21	23
IREC-MEDE-FSM	22	36	40
IRSS-FSP	11	28	42
LDRI-FASB	15	20	10
Clinics	3	7	2
Students	5	11	25

Table 5 (SST) shows a stabilization compared to last year. As explained earlier, the spike in attendance in 2016/17 corresponds to a special session more specifically dedicated to engineers.

Table 5. Origin of participants from the science and technology sector (SST)

	2015/16	2016/17	2017/18
SST	127	190	180
AGRO	1	7	9
ELI	39	39	32
ICTEAM-EPL	3	9	5
IMCN		2	1
IMMC	4	32	7
IRMP	4	9	5
LIBST	8	16	20
LOCI	5	6	7
SC	1	2	4
Students	62	68	90

The SMCS is also regularly invited to sessions organized by other entities, for example, a seminar requested by IREC, “*The curse of multiple comparisons in statistics: why too many questions may cause lack of significance?*”, November 27 (C. Bugli and L. Desmet).

Training

These presentations are not included in the training statistics.

Overall, the training activity of the SMCS occupies about 3.3 full time staff equivalents (FTE). 1.8 FTE of this effort is directly oriented towards the UCLouvain research and teaching units, the remainder towards external clients, such as the FOREM and some private companies and public administrations.

The trainings given to external clients fulfill three objectives. First, they are a part of our mission of service to society. Second, they allow us to develop new course modules for internal use, such as the FOREM training which allowed us to bring together an entire course portfolio in the area of bioinformatics. These course modules can then be offered to members of the life and health sciences groups in SST, SSH or SSS. Third, but not least, they allow us to generate financial resources to have enough funds for our staff members not financed by UCLouvain. We of course have to be cautious in the choice and management of such trainings to make sure there is a positive return (financial and others).

Finally, members of the SMCS occasionally participate in sessions of the statistical consulting course (LSTAT3220) where they tell the students about their career path and their way of working.

Consulting

Consulting

Consulting is another of our main missions. We help researchers, administrators, and external clients in collecting meaningful data and in interpreting them.

Whenever possible, we provide methodological support to enable our clients to solve their problems by themselves. Our support can just be a piece of advice or short-term assistance. This type of consultancy is pre-paid by the financial contribution of the sectors up to 1 FTE and requires about 1.5 FTE.

On occasion, the complexity or the amount of the statistical work requires that the SMCS consultant spend a larger amount of time on a project. The consultancy thereby evolves into a "*collaboration*". This requires additional funding by the client. Moreover, if the input of the statistical consultant yields new insight, common publishing rules require that he/she should become a co-author of publications based on the work. This type of support is increasing and corresponds to more or less 750 hours.

Table 6. Origin of consulting projects

Client split	2015/16	2016/17	2017/18
Research*	294	384	401
Master's projects*	73		
Educational project			2
Admin. services	10	5	8
External clients	10	11	9
Total	387	400	420

*In 2016-2017, the SMCS has supported the master's projects based on its organizational rules more systematically. The SMCS has tried as much as possible to meet the students and the promotor at the same time. Therefore, there is no distinction anymore between research projects and Master's ones. Indeed, these two types of projects are part of the UCLouvain research at large.

From 1 September 2017 to 31 August 2018, the SMCS worked on 420 projects of which 411 came from UCLouvain. As expected, most of the projects are linked to university research. Some larger projects involving the university administration or the schools or institutes and

external clients are presented in the Larger Projects section.

The following tables present the breakdown of the project origins in the three university sectors. They clearly show that the SMCS services reach almost all areas of research at UCLouvain. For the second year we include in our statistics a cross-disciplinary support to each sector. This consists in statistical computing support for all the sector community. Even if this type of support counts for one unit in the table, this can sometimes represent up to 50 hours of work. This support consisted this year in the participation in the shooting of two videos for the MOOC "CRITIX" addressed to the BAC1 students (**two educational projects**).

Table 7 (SSH) shows that the number of requests from the humanities sector fluctuates around 133 requests per year. As in previous years, the institute and the school of psychology had by large the highest share. This is why these units have a 0.5 FTE statistical advisor (Nathalie Lefèvre, who in the remainder of her time acts as the director of the SMCS). Also IACCHOS has an internal specialist for data questions matters (Jonathan Dedonder) who is an associated member of the SMCS. Besides IPSY with particularly high activity, requests came from almost all units of this sector.

Table 7. Origin of consulting requests from the humanities sector (SSH)

	2015/16	2016/17	2017/18
SSH	131	133	133
Cross-disciplinary		1	1
IACCHOS-ISPOLE-ESPO	20	12	18
ILC-INCAL-FIAL	11	6	20
LOURIM-LSM	7	9	7
IMMAQ	7	10	12
IPSY-PSP	84	94	73
JURI-DRT	1	1	2
RSCS-TECO	1		

Table 8 (SSS) shows that on average there are 150 requests per year from the health sciences sector. The areas related to clinical research, Health and Society Institute and the medical school account for about 60%

Consulting

of this, other requests are split among the remaining institutes. The number of requests increase in IRSS probably thanks to the presence of two of our consultants regularly in the institute.

The statistics shown above do not include large-scale external projects documented in a later section of this report.

Table 8. Origin of consulting requests from the health sciences sector (SSS)

SSS	2015/16	2016/17	2017/18
Cross-disciplinary	120	150	151
ASS		1	1
DDUV-Ludwig Inst.	15	10	8
FASB	1	6	3
IONS	16	25	26
IREC-MEDE-FSM	69	92	63
IRSS-FSP	7	8	30
LDRI	11	5	8
Clinics - Albert II		3	12

Table 9 (SST) shows that the number of request from the science and technology sector fluctuates around 100-130 requests per year.

Table 9. Origin of consulting requests from the science and technology sector (SST)

SST	2015/16	2016/17	2017/18
Cross-disciplinary	124	101	117
AGRO		1	2
ELI	12	9	7
ICTEAM-EPL	66	37	36
IMCN	2	3	3
IMMC	1	8	14
LIBST	8	35	45
LOCI	27	4	4
SC	5	3	6
Commission doctorale	2	1	1

Our main consultant for the science and technology sector (SST), Catherine Rasse, also continues to be a member of an ethics committee related to experimentation on animals.

Finally, **eight projects** also originated from the central administration (AC, LLL, SGSI).

Computing

Servers and Statistical Software

The SMCS now has two servers dedicated to statistical computing and no software-sharing server since UDS replaces it (see below). Moreover, the SMCS works together with the SGSI and the CISM on several activities to reduce barriers to using appropriate statistical software for university members. The computing activities require about 0.5 FTE.

Statistical Computing Server

We renewed the system three years ago. Our current two performance servers are based on a dual Intel Xeon E5-2667 processor with 128 GB of RAM and an SSD disk for the SAS temporary folder. The CISM (UCLouvain's high performance computing platform) hosts this server system and its storage and is in charge of the operating system and the user management. The SMCS takes care of the statistical software on these machines. Currently, SAS, R and Stata-MP are available.

The statistical computing server has about 52 users. This year we counted users in the same way as last year (and so in a different way from two years ago): before it was enough to renew one's login to be a server user but from now on you have to renew it and use it. People from IMMAQ and IRSS are the main users of this service (the table on the following page provides more details about the users of the two servers). We can divide users in two categories: people who want to use SAS free in the Big5-STAT project and people in need of a high performance computer mainly for the three installed software packages (large RAM and multi-processors version). The average CPU charge was 15 % on smcs3 and 21 % on smcs4 with spikes (100% for smcs3 and 100% for smcs4) occurring when users execute paralleled programs in Stata or R.

More data analysts at UCLouvain should consider using this server system. It offers more performance, storage, and data security than local PC-based systems.

Ucilla and UDS

Ucilla and UDS are services provided by the SGSI that allow users who asked for an access to use software without having it on their computer. It is recommended

to do so via an installed client on the computer but it also works through a web browser. From September 2018, Ucilla disappears with UDS replacing its functionalities.

It is possible to add software to these platforms. Who pays for it will depend on many criteria so you should send an email to 8282@uclouvain.be to have more details.

Available Software

Via the Big5-STAT initiative, any UCLouvain member can request the installation of or the access to one or several of the five most used statistical software packages, R, JMP, SAS (via Enterprise Guide), SPSS, and Stata. The request should be made via the central computing service desk 8282 or even better via the "Catalogue de service". Many UCLouvain members used this service – 931 installation requests made in the academic year 2017-2018, (444 demands for SPSS, 181 for JMP, 181 for Stata, 61 for R, 34 for SAS and 30 for SAS Enterprise Guide).

High performance solutions in SAS, R and Stata (multi-processor) are available on the two statistical computing servers.

Finally, UDS already offers the Big5-STAT software and gives an access to HLM, MPlus, nQuery, PASS and Simca. The SGSI paid all the software but Simca that is paid by the professor B. Govaerts. Note that nQuery is no more available since it was not renewed (there only was one user during the last year and PASS can replace it).

The access conditions and procedures for the statistical computing server and the software-sharing server are available under the Servers tab of our [official web page](#).

Computing

Usage Statistics

Table 10. Users of the statistical computing server

	2015/16	2016/17	2017/18
Total	125	54	52
SSH	59	30	30
IACS-SPLE-ESPO	11	4	3
ILC	1	1	2
IMMAQ	36	23	24
IPSY-PSP	4	2	1
LOURIM-LSM	7		
SSS	29	17	14
ASS	1		
IONS	6	2	
IREC-MEDE	4	1	
IRSS-FSP	16	13	14
LDRI	2	1	
SST	36	7	8
AGRO	5		
ELI	19	4	6
ICTEAM-EPL	3	1	1
IMCN	2		
IMMC	1		
IRMP			
LIBST	4		
LOCI			
SC	2	2	1
Central services	1		
SGSI	1		

Table 11. Users of the software-sharing server (UDS)

	2015/16	2016/17	2017/18
Total	26	25	50
SSH	20	19	25
IACCHOS	2	1	
ILC			1
IMMAQ	2	9	9
IPSY-PSP	14	9	12
LOURIM-LSM	2		3
SSS	5	2	17
DDUV			4
IONS	2	1	2
IREC	2	1	5
IRSS			2
LDRI	1		4
SST	1	4	8
AGRO		1	2
ICTEAM		1	
IMMC			1
LIBST			2
SC		3	3

Usage of the web portal - Ucilla

The most used software was PASS (111 connections) followed by HLM (48), Mplus (38), SPSS Premium (22) and the other software had at most 10 connections during the year. The increase of users observed in SSS comes mainly from the use of PASS (sample size computation)

Help for students

Students sometimes need help installing a statistical software package on their own computers to prepare exercises or to work on a project. The SMCS helps them to do so since the course assistants or professors do not always know what went wrong during the installation.

In the academic year 2017-2018, the platform helped 100 students of which 68 came for assistance with SPSS. This suggests that the available information for installation may not yet be sufficient even though it decreased from 30 % compared with the previous year.

Larger Projects

Larger Projects

Survey Support

The Survey Project was an initiative of the humanities sector SSH to facilitate the design and implementation of surveys and now helps SSS and SST sectors as well.

10 free survey support consulting projects were implemented in 2017-2018, covering advice on survey and questionnaire design, sampling and survey weights. Furthermore, two paid internal projects aimed at analysing and reporting survey data.

In addition to the review of existing survey-related training courses, an external training course on sampling for auditors was developed.

The SMCS also made use of these skills to implement a review and update of the satisfaction survey sent to training participants.

General interest Research Projects

Since late 2017, a constant monitoring of public tenders linked to statistics and statistical training has been implemented. Thanks to this more proactive approach, three proposals were submitted in response to public tenders. We submitted a tender for a 70-day long Data Analyst training for Bruxelles Formation, but were not selected. We submitted and obtained two other projects, which will be implemented in 2018-2019: the design of a theoretical framework and questionnaire regarding the impact of vocational training on job-seekers for Bruxelles Formation and the identification and gathering of gendered statistics for the Service Public Régional de Bruxelles. These projects have in common to be general interest research projects in cooperation with various experts within UCLouvain (Pr. Isabel Raemdonck for the first one and a team of gender studies experts for the second one). This type of project helps strengthen relationships with other entities, brings forward internal expertise and allows SMCS consultants to develop new skills while providing additional budget to cover staff costs.

Research Project - Work and Organizational Psychology (2016-2018)

In 2016-2017, we have also started other kinds of projects. Following the departure of a PhD student, part of the remaining budget was given to the SMCS to implement the statistical analysis of existing data. In collaboration with D. Desmette from the institute of psychology (IPSY), the SMCS took a large part in the analysis of data collected from ONE's employees (Office of Birth and Childhood). With this type of project, the client benefits from the statistical expertise of the SMCS (e.g. modelisation in Structural Equation Modeling), saves time and can also get some hints and ideas for future research projects.

Research Project - Erasmus+OLS (2015-2021)

In spring 2015, the SMCS became partner in the European project Erasmus+OLS. It is a seven-year project to provide online language support to several hundred thousand participants in Erasmus+ programs. The language support consists of a testing platform to determine the level in the work language of the host institution at the beginning and the end of the stay and of an online learning platform.

Our role in this project is to monitor the performance of a progressive online test (a test in which the difficulty of later questions depends on the answers to earlier ones), to assist in the design and realization of user experience surveys, to facilitate communication of performance statistics, and to provide expertise to the scientific committee.

As a new initiative, we are currently exploring the potential of text mining methods to analyze the answers to open questions of the surveys.

Research Project - Macular degeneration (2017-2019)

In collaboration with ISBA, the SMCS participates in a research consortium on Macular degeneration involving a research group at the University of Liège. This project is a metabolomics project.

Larger Projects

Research project - First spin-off (2018-...)

The SMCS helped a research unit from IREC with the elaboration of a first UCLouvain spin-off funding application (Service public de Wallonie).

Research project - Micro-array analysis for GSK (2018)

The SMCS helped in the analysis of micro-array data a team of UCLouvain (LIBST) involved within a research project in collaboration with GSK. GSK financed the implication of the SMCS and the SMCS was also involved in a master thesis related to this topic.

Course Module - FOREM (2011-...)

Since 2011/12, the SMCS provides a course in bioinformatics/biostatistics for the FOREM. This course is aimed at people who have bio-sciences degrees and difficulties finding a suitable job. The course volume is 33 days (corresponding to more than 200 hours of lectures and supervised exercise sessions).

The results are quite encouraging as most of the former participants quickly found work suited to their abilities and training. Since the volume of this course is substantial, we submitted a proposal to associate it with an accredited certificate. Moreover, the course modules developed for this training program will become available to UCLouvain researchers.

In 2017, SAS certification became mandatory for the participants. The aim is to gradually prepare the training to become as well a university certificate. This year, 100% of the participants were SAS certified!

Moreover, a module of the biostatistics program is now offered twice a year in the bioinformatics program of the FOREM.

Course Modules – BioWin (2010-2018)

The SMCS is also providing short course modules in the context of the BioWin initiative of the Walloon government. This type of collaboration is currently under revision due to budget constraints. Courses offered within this framework are advertised by BioWin and course fees are partially subsidized until end of 2018.

University Certificate - Junior Data Analyst (2016-...)

In collaboration with the LSBA, the SMCS has launched a new certificate called “Junior Data Analyst” in 2016, an intensive course for graduates in humanities in practical data management, analysis, and interpretation. The courses were given for the third edition from October 2018 to December 2018. The project was financed by the European Social Fund (ESF) until the end of 2017 and is now self-financing. Almost the whole SMCS team has been actively involved in the preparation of the course program and the different modules. The fourth edition will start in October 2019.

Collaboration with the LSBA

In 2019, the SMCS will organize in collaboration with the LSBA and CENTAL, a summer school on Text Mining.

Your Project?

As you can see from this report, we are ready to become partners of research projects and grant applications. Do not hesitate to contact us.

We can provide an early estimation of the amount of statistical work needed in your project, plan important stages of this work, and give an outline of the statistical methodology which would be employed.

The advantages of involving us in the early stages of a project include a more thorough study design, efficient allocation of resources and methodologies, and data that are easier to analyze and interpret.

The benefits are higher research efficiency and more credibility towards the granting authorities and scientific journals.

Appendix

Appendix

List of Trainings and Seminars 2017-2018

Trainings

- Analyse de données pré-cliniques et cliniques avec JMP, du 8 juin 2017 au 7 septembre 2017 (Céline Bugli, Catherine Rasse)
- NEW - Analyse de données d'enquête - Formation sur mesure, du 12 juin 2017 au 21 septembre 2017 (Séverine Guisset, Arnaud Pollaris)
- Programme de formation à l'utilisation du logiciel NVivo (formation réservée aux employés de l'entreprise ADE), du 15 juin 2017 au 31 octobre 2017 (Jonathan Dedonder)
- Introduction to R language, du 4 septembre 2017 au 6 septembre 2017 (Alain Guillet)
- Manipulation et exploration de données avec SPSS et réalisation de rapports efficaces, du 11 septembre 2017 au 12 septembre 2017 (Nathalie Lefèvre)
- NEW - Collecter des données d'enquête sur tablette et smartphone, du 12 septembre 2017 au 12 septembre 2017 (Bruno Schoumaker)
- Initiation à la statistique avec applications en JMP (Formation intégrée dans un programme intensif plus large du FOREM), du 15 septembre 2017 au 21 septembre 2017 (Catherine Rasse, Céline Bugli)
- Pratique de la statistique avec SPSS, du 18 septembre 2017 au 22 septembre 2017 (Nathalie Lefèvre)
- Introduction au calcul de taille d'échantillon, du 29 septembre 2017 au 29 septembre 2017 (Catherine Rasse, Céline Bugli)
- Formation de base au logiciel NVivo et coaching associé, du 29 septembre 2017 au 12 décembre 2017 (Jonathan Dedonder, Séverine Guisset)
- Analyse factorielle: approche théorique et mise en application avec SPSS, du 2 octobre 2017 au 3 octobre 2017 (Séverine Guisset)
- Certificat universitaire en Junior Data Analyst, du 2 octobre 2017 au 16 décembre 2017 (Aurélie Bertrand, Jonathan Dedonder, Alain Guillet, Nathalie Lefèvre, Arnaud Pollaris, Catherine Rasse, Christian Ritter, Séverine Guisset)
- NEW - Créez votre liste de randomisation avec le logiciel JMP, du 4 octobre 2017 au 4 octobre 2017 (Céline Bugli)
- Introduction au logiciel JMP, du 4 octobre 2017 au 11 octobre 2017 (Céline Bugli, Lieven Desmet)
- Initiation au logiciel NVivo, du 9 octobre 2017 au 11 octobre 2017 (Jonathan Dedonder, Séverine Guisset)
- Data management and analysis using Stata, du 23 octobre 2017 au 27 octobre 2017 (Sébastien Fontenay, Alain Guillet)
- Initiation à la statistique avec applications en JMP (Formation intégrée dans un programme intensif plus large du FOREM), du 24 octobre 2017 au 27 octobre 2017 (Catherine Rasse, Céline Bugli)
- Formation SAS : programmation de base, du 7 novembre 2017 au 14 novembre 2017 (Formateur SAS, Céline Bugli)
- Pratique de la statistique avec SPSS, du 10 novembre 2017 au 24 novembre 2017 (Nathalie Lefèvre, Aurélie Bertrand)
- Exploration de données et base de modélisation statistique avec JMP, du 16 novembre 2017 au 20 novembre 2017 (Catherine Rasse)
- Prémisses d'une enquête et création de questionnaires, du 27 novembre 2017 au 28 novembre 2017 (Jonathan Dedonder)

Appendix

- **NEW** - Publier votre travail dans un package R, du 28 novembre 2017 au 28 novembre 2017 (Manon Martin, Joris Chau)
- **NEW** - Analyse de données pré-cliniques et cliniques avec SPSS, du 1 décembre 2017 au 1 décembre 2017 (Aurélie Bertrand, Nathalie Lefèvre)
- Analyser la médiation, la modération et la médiation modérée avec SPSS, du 7 décembre 2017 au 8 décembre 2017 (Arnaud Pollaris, Nathalie Lefèvre)
- Introduction à la réalisation d'enquêtes en ligne avec LimeSurvey, du 11 décembre 2017 au 11 décembre 2017 (Jonathan Dedonder, Nathalie Lefèvre)
- Pratique de la statistique avec SPSS, du 8 janvier 2018 au 12 janvier 2018 (Nathalie Lefèvre)
- Plans d'expérience avec JMP, du 11 janvier 2018 au 15 janvier 2018 (Catherine Rasse, Bernadette Govaerts)
- Pratique de la statistique avec R, du 15 janvier 2018 au 19 janvier 2018 (Vincent Bremhorst)
- Initiation aux modèles d'équations structurales (SEM) avec Stata, du 15 janvier 2018 au 19 janvier 2018 (Arnaud Pollaris)
- Introduction to R language, du 23 janvier 2018 au 30 janvier 2018 (Lieven Desmet, Nathan Uyttendaele)
- Analyses multivariées en RCommander (FactoMineR), du 25 janvier 2018 au 26 janvier 2018 (Arnaud Pollaris)
- Initiation au logiciel NVivo, du 29 janvier 2018 au 31 janvier 2018 (Jonathan Dedonder, Nathalie Lefèvre)
- Introduction aux modèles multi-niveaux avec SPSS, du 29 janvier 2018 au 2 février 2018 (Aurélie Bertrand, Nathalie Lefèvre)
- Formation SAS : SAS advanced programming, du 16 février 2018 au 30 mars 2018 (Formateur SAS)
- **NEW** - Pratique de la statistique avec SAS, du 2 mars 2018 au 9 mars 2018 (Catherine Rasse, Céline Bugli)
- Formation en analyse de données et statistiques pour la recherche et le développement dans la bioindustrie, du 5 mars 2018 au 29 juin 2018 (Aurélie Bertrand, Céline Bugli, Bernadette Govaerts, Alain Guillet, Manon Martin, Catherine Rasse)
- Introduction aux outils statistiques pour le suivi de la qualité de processus de production, du 8 mars 2018 au 8 mars 2018 (Alain Guillet)
- Programme de formation en statistique pour la recherche clinique avec exemples en SPSS, du 8 mars 2018 au 10 mars 2018 (Nathalie Lefèvre, Aurélie Bertrand)
- **NEW** - Introduction aux concepts de base de la statistique bayésienne, du 23 mars 2018 au 23 mars 2018 (Vincent Bremhorst)
- Modèles linéaires avec SAS, du 23 mars 2018 au 30 mars 2018 (Catherine Rasse, Céline Bugli)
- Gestion et analyse de données avec Stata, du 9 avril 2018 au 13 avril 2018 (Sébastien Fontenay, Alain Guillet)
- Utilisation avancée du logiciel Nvivo, du 16 avril 2018 au 19 avril 2018 (Jonathan Dedonder)
- **NEW** - Introduction to the basic concepts of Bayesian statistics, du 27 avril 2018 au 27 avril 2018 (Vincent Bremhorst)
- Initiation au logiciel NVivo, du 2 mai 2018 au 4 mai 2018 (Séverine Guisset, Jonathan Dedonder)
- Panel-data Analysis using Stata, du 17 mai 2018 au 18 mai 2018 (Sébastien Fontenay, Alain Guillet)
- Initiation au logiciel NVivo, du 17 mai 2018 au 18 mai 2018 (Séverine Guisset)
- Introduction aux modèles mixtes avec SAS, du 28 mai 2018 au 29 mai 2018 (Catherine Rasse, Céline Bugli)
- Cours d'initiation au langage R, du 28 mai 2018 au 4 juin 2018 (Lieven Desmet, Vincent Bremhorst)
- **NEW** - Introduction au Text Mining, du 29 mai 2018 au 29 mai 2018 (Sébastien De valeriola, Thomas François, Séverine Guisset, Patrick Watrin)
- Cours d'initiation au langage R, du 4 juin 2018 au 8 juin 2018 (Nathan Uyttendaele, Joris Chau)
- Pratique de la statistique avec R, du 7 juin 2018 au 14 juin 2018 (Vincent Bremhorst)
- Introduction à l'analyse de Réseaux Sociaux avec Gephi, du 11 juin 2018 au 13 juin 2018 (Jonathan Dedonder)
- Gestion et analyse de données avec Stata, du 12 juin 2018 au 26 juin 2018 (Alain Guillet)
- Analyses multivariées avec SPSS, du 18 juin 2018 au 21 juin 2018 (Séverine Guisset, Nathalie Lefèvre)

Appendix

- NEW - Introduction aux modèles mixtes avec R, du 18 juin 2018 au 22 juin 2018 (Alain Guillet)
- NEW - Build interactive web applications with the R Shiny Package, du 25 juin 2018 au 26 juin 2018 (Eric Lecoutre)
- NEW - Introduction à l'analyse de parcours avec R (package TraMineR), du 3 juillet 2018 au 3 juillet 2018 (Lieven Desmet)

Publications and Presentations by SMCS members

Although the primary mission of SMCS members is to support other research projects, its members are encouraged to participate in scientific meetings, to present their work, and to publish.

Publications

Journal Articles

Dallemagne, Matthew ; Ghys, Emmanuelle ; De Schrevel, Catalina ; Mwema, Ariane ; De Troy, Delphine ; Rasse, Catherine ; Donnay, Isabelle. *Oxidative stress differentially impacts male and female bovine embryos depending on the culture medium and the stress condition.* In: *Theriogenology*, Vol. 117, p. 49-56 (2018). <http://hdl.handle.net/2078.1/198588>

Desmet, Lieven ; Venet, David ; Doffagne, Erik ; Timmermans, Catherine ; Legrand, Catherine ; et. al. *Use of the beta-binomial model for central statistical monitoring of multicenter clinical trials.* In: *Statistics in Biopharmaceutical Research*, Vol. 9, no. 1, p. 1-11 (2017). <http://hdl.handle.net/2078.1/173822>

Dormal V., Bremhorst V., Lannoy S., Lorant V., Luquiens A., Maurage P. *Binge drinking is associated with reduced quality of life in young students: A pan-European study.* In: *Drug & Alcohol Dependence*. In press.

Dormal, Valérie ; Larigaldie, Nathanaël ; Lefèvre, Nathalie ; Pesenti, Mauro ; Andres, Michael. *Effect of perceived length on numerosity estimation: Evidence from the Müller-Lyer illusion.* In: *Quarterly Journal of Experimental Psychology*, Vol. 71, no. 10, p. 2142-2151 (2018). <http://hdl.handle.net/2078.1/194300>

Seminars and Presentations

Extra-muros seminars and video

Bertrand, Aurélie & Guisset, Séverine. Participation au MOOC Critix coordonné par le Louvain Learning Lab - Vidéos sur l'échantillonnage et les sondages - Eté 2018.

Lefèvre, Nathalie, Cosse Marie-Pascale, Hafner, Christian & Govaerts Bernadette. *Enseignement Collaboratif et Innovant de la Statistique au sein de l'UCL - Auto-Apprentissage en Ligne et Télèveotours.* Colloque Francophone International sur l'Enseignement de la Statistique (CFIES), Grenoble, du 6 au 9 septembre 2017.

Presentations and applied statistics workshops

Bugli, Céline. *Les statistiques dans votre mémoire: il vaut mieux y penser avant pour éviter les ennuis après...* Présentation aux mémorants FSM, le 2 octobre 2017.

Sebban, Ilham, Guisset, Séverine et Ritter, Christian. *Analyser les réponses ouvertes d'une enquête avec des outils automatiques - premiers pas.* Intervention dans le cadre de l'Atelier de statistique appliquée, Louvain-la-Neuve, le 10 novembre 2017.

Appendix

Collaborations, Response to Tenders and Supervisions

Collaborations and Response to Tenders

Bertrand, Aurélie. Organisation des Midis pédagogiques de la LSBA, Louvain-la-Neuve, année académique 2017-2018.

Bugli, Céline. Aide à la rédaction d'offres de la plateforme Centre d'Investigation Clinique en Nutrition, année académique 2017-2018.

Bugli, Céline. Aide à la rédaction d'offre pour un appel de la région Wallonne d'analyse multivariée de la production de déchets ménagers et assimilés en Wallonie pour l'entreprise RDC Environnement, année académique 2017-2018. Offre obtenue.

Bugli, Céline. Aide à la rédaction d'offre pour un appel de la région Wallonne d'élaboration d'un référentiel de calcul des pertes et gaspillages alimentaires en Wallonie pour l'entreprise Comase, année académique 2017-2018. Offre non obtenue.

Bugli, Céline et Guisset, Séverine. Rédaction d'une offre pour un appel de Bruxelles Formation dans le cadre d'un marché de services de formation relatif au métier de « Data Analyst », année académique 2017-2018. Offre non obtenue.

Guisset Séverine. Rédaction d'une offre soumise avec Isabel Raemdonck auprès de Bruxelles Formation concernant le marché de services d'enquête sur les processus d'insertion sur le marché du travail à l'issue d'une formation professionnelle, année académique 2017-2018. Confirmation d'obtention du projet le 26/06/2018.

Guisset Séverine. Rédaction d'une offre soumise en partenariat avec des expertes en Genre de l'UCLouvain sur l'identification et la collecte d'indicateurs et statistiques de genre pour le Service Public Régional de Bruxelles, année académique 2017-2018. Confirmation d'obtention du projet le 04/07/2018.

Guisset, Séverine. Rédaction d'une offre de 2 jours de formation aux « Statistiques et techniques d'échantillonnage » destinée à des auditeurs internes de la Direction de l'Audit interne du Secrétariat général du SPW. Offre retenue.

Lefèvre, Nathalie et Guillet, Alain. Rédaction d'une offre de 10 jours de formation intra-entreprise aux méthodes statistiques avec le logiciel R pour le personnel de l'IWEPS. Offre retenue.

Lefèvre, Nathalie et Guillet, Alain. Rédaction d'une offre de 11.5 jours de formation intra-entreprise aux méthodes statistiques avec le logiciel Stata pour le personnel de l'IWEPS. Offre retenue.

Lefèvre, Nathalie. Aide à la rédaction d'offres diverses en collaboration avec Ritme Informatique, France et Ritme Sprl Benelux, année académique 2017-2018.

Lefèvre, Nathalie. Aide à la rédaction d'un projet de développement d'outils et procédures pour des évaluations dématérialisées asynchrones dans le cadre d'une demande de financement via le Fonds de développement pédagogique (Promoteur : Kestemont, M.P. et co-promoteurs : Lefèvre, N., Masquelier, B., Schoumaker, B.), année académique 2017-2018. Confirmation d'obtention du projet en juin 2018.

Lefèvre, Nathalie et Bertrand, Aurélie. Rédaction d'une offre de 2.5 jours de formation en statistique pour la recherche clinique utilisant des exemples appliqués en SPSS. Offre retenue.

Rasse, Catherine. Participation à la commission d'éthique animale de Louvain-La-Neuve, année académique 2017-2018.

Supervision of Master and PhD thesis

Bugli, Céline. Participation au comité d'encadrement de thèse de Florian Pechon, année académique 2017-2018.



2017-2018