SMCs

STATISTICAL METHODOLOGY AND COMPUTING SERVICE

Report of Activities
1 September 2016 to 31 August 2017

2016-2017

Statistical Methodology and Computing Service
Report of Activities
1 September 2016 to 31 August 2017

Contents

Contents

Introduction

Introduction

The SMCS continues its activity as a diversified service provider in the area of practical statistics. Its 4.5 to 5 full time equivalents of consulting staff and a network of more than 20 other specialists assist UCL 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 2016/17.

The consulting activity has been stable at around 400 projects per year. The training activity continued to grow to 40 courses, representing 135 days of training. These courses, focusing on practical statistics and the use of statistical software, were attended by about 650 participants for a total of about 1850 person days. The statistical computing server had about 54 registered users.

Since the current workload started to exceed the staff capacity, the management committee supported the appointment of an extra 1.5 EFT starting in September 2017. 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. We wrote a brochure presenting our services and distributed it widely inside and outside the university and included it on our website. Moreover, we made an effort to communicate more via our contact persons in the different research institutes. We already see improvements and will continue in this direction.

We also gained visibility at the university level. In the framework of the merger UCL – USL-B, the access to the technological platforms such as the SMCS has been highlighted in the UCL Express of April 19, 2017 as a UCL contribution to USL-B to support the research.

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 first edition of the new certificate "Junior Data Analyst" which took place in fall 2016. You will find more details in this report.

A new project in Metabolomics started in 2016 about Macular degeneration. Finally, 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. Another research collaboration (Metabiose) ended in 2017.

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 2017/18 we are considering several new initiatives. Ideas include a summer school for researchers and a stronger methodological and practical support for surveys. 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 comité de gestion 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 2017. Further ideas were gathered to make this event even more attractive to a wider audience.

Personnel

The SMCS hired two consultants, Séverine Guisset and Aurélie Bertrand, who will start working for the team in September 2017. S. Guisset will work more specifically on surveys while A. Bertrand will help on general projects and trainings.

Corine Lhôte, who worked on the Metabiose project, finished her contract in September 2016.

Marie-Pascale Cosse joined the SMCS team in September 2016 to work on promotional and quality management.

Statistical Software

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

Overall, the initiative seems to work well, even though we sometimes receive reports of unusually long waiting times. Most of these cases can be traced back to misunderstandings or peculiarities of the university organization. In particular, the special status of St Luc as a hospital with a separate management but members who may also belong to the UCL research institutes of the health sciences sector still too often creates problems with respect to access to appropriate statistical software. We will have to continue working on this issue in collaboration with the central computing services.

The SGSI launched a new service, UDS, which will probably replace Inuvika/Ucilla in future. 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. In particular the students following the FOREM training were encouraged to use the service, even from home. This helped adjust the system set up and identify some issues. Its use is going to be extended next year to test how this tool reacts when there are more simultaneous users.

The New University Certificate "Junior Data Analyst"

In collaboration with the LSBA, the SMCS has launched in October 2016 a new certificate called "Junior Data Analyst" (JDA), an intensive course for humanities graduates in practical data management, analysis, and interpretation. The courses took place from October 2016 to December 2016. The project is sponsored by the European Social Fund (ESF). In the first year, we received about 50 applications, accepted about half of them, and in the end, 16 people enrolled. The second edition will start in October 2017. 52 candidates have applied and 20 of them enrolled.

UNION EUROPÉENNE
Fonds social européen

News

Promotion, administration, miscellaneous

Marie-Pascale Cosse is in charge of the promotion of the SMCS activities and quality management. In particular, she has worked on the promotion of the JDA certificate, promotional materials such as posters, a catalogue of all the SMCS services, our new website, mailings and newsletters. She will continue working on promotional activities and will start working on quality issues.

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

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

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 EFT)

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 EFT)

Main areas of activity: Medical Sector (SSS), and statistical computing.
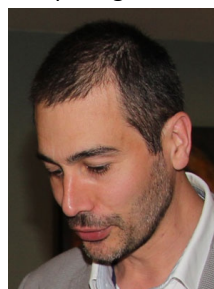
Master professionnel Sciences, Santé, Technologie, mention Mathématiques et Applications, spécialité Ingénierie Mathématique (Université de Nantes, 2005)
DEA en informatique (Université de Nantes, 2004)

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

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).

People

Arnaud Pollaris (UCL/IMMAQ/SMCS – 1 EFT)

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 EFT)

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

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

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

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)

Nathalie Lefèvre (UCL/IPSY)

Jérôme Ambroise (IREC),

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.

Interns

David Stern, August 2017

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.

Nicolas Asin (ISBA), Aurélie Bertrand (ISBA), Céline Bugli (SMCS), Vincent Bremhorst (ISBA), Joris Chau (ISBA), Benjamin Colling (ISBA), Mickaël De Backer (ISBA), Jonathan Dedonder (IACCHOS), Sébastien Fontenay (IRES), Bernadette Govaerts (ISBA), Oswaldo Gressani (ISBA), Alain Guillet (SMCS), Jonathan Jones (IMMAQ), Eric Lecoutre (WeLoveDataScience!), Nathalie Lefèvre (SMCS), Rebecca Marion (ISBA), Manon Martin (ISBA), Gaëtan Montero Redondo (IMMAQ), Arnaud Pollaris (SMCS), Catherine Rasse (SMCS), Christian Ritter (ISBA), Xavier Rouby (Pfease), Marco Saerens (LRIM), Cédric Taverne (Devoteam), Isabelle Thomas (IMMAQ), Nathan Uyttendaele (ISBA), 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. Four of the training sessions took place on the Woluwe campus.

Nine 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 UCL (SPSS, R, SAS, SAS Enterprise Guide, JMP, Stata, NVivo, Excel, Limesurvey).

  Some trainings are organized to respond to users' requests. For example, the R training, "*Pratique de la statistique avec R*", January 2017. The SMCS offers many trainings about R, some as an initiation to the R language and others more advanced. Some trainings organized to respond to user requests thereafter become Regular SMCS trainings.

- **Customized trainings for specific audiences** (UCL and external). Examples: "*Introduction aux modèles multi-niveaux avec SPSS*" (organized in collaboration with IPSY but open to everyone), "*Initiation à la statistique avec application en JMP*" (in collaboration with the FOREM).

- **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.

- **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 first 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 40 courses in the academic year 2016-2017 for a total of 135 teaching days. This represents an increase of 26 days of training due to, among others, the organization of the new Junior Data Analyst certificate (21 days).

Table 1. Global statistics on SMCS trainings.

	2014/15	2015/16	2016/17
Courses	37	38	40
Course Days	112.5	108.5	135
Participants UCL	450	448	537
Person-Days* UCL	837.5	826	1199
Participants Ext	164	117	128
Person-Days* Ext	619	660.5	680

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.

Table 2. Number of participants per sector.

	2014/15	2015/16	2016/17
SSH	161	147	173
SSS	115	77	122
SST	60	65	122
Students	112	159	120
Central services	2		

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 Master's students. The fact that they cannot register in priority can explain why they could not participate in some of our trainings and therefore why there is a decrease in their participation rate.

Training

Table 3 (SSH) shows that 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. However, these trainings remain open to all.

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

	2014/15	2015/16	2016/17
SSH	**206**	**239**	**214**
IACS-SPLE-ESPO	54	53	61
ILC	34	13	7
LOURIM-LSM	9	14	16
IMMAQ	8	16	8
INCAL-FIAL	6	7	5
IPSY-PSP	48	42	71
JURI-DRT	2	1	5
RSCS-TECO		1	
Students	45	92	41

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.

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

	2014/15	2015/16	2016/17
SSS	**135**	**82**	**133**
ASS	1	1	1
DDUV-Ludwig Inst.	30	11	9
IONS	17	14	21
IREC-MEDE-FSM	30	22	36
IRSS-FSP	14	11	28
LDRI-FASB	14	15	20
Clinics	9	3	7
Students	20	5	11

Table 5 (SST) shows an increase in all entities with a larger increase for IMMC. As explained earlier, the spike in attendance corresponds to a special session more specifically dedicated to engineers.

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

	2014/15	2015/16	2016/17
SST	**107**	**127**	**190**
AGRO	6	1	7
ELI	31	39	39
ICTEAM-EPL	2	3	9
IMCN	1		2
IMMC	1	4	32
IRMP		4	9
ISV	7	8	16
LOCI	9	5	6
SC	3	1	2
Students	47	62	68

The SMCS is also regularly invited to sessions organized by other entities, for example, a seminar requested by IREC, "*Why and How Randomise a Clinical Trial*", May 29 (C. Bugli and L. Desmet). These presentations are not included in the training statistics.

Overall, the training activity of the SMCS occupies about 2.5 full time staff equivalents (FTE). 1.5 FTE of this effort is directly oriented towards the UCL 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 UCL. 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 and requires about 1 EFT.

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.

Table 6. Origin of consulting projects

Client split	2014/15	2015/16	2016/17
Research*	318	294	384
Master's projects*	87	73	
Admin. services	12	10	5
External clients	7	10	11
Total	**424**	**387**	**400**

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 UCL research at large.

From 1 September 2016 to 31 August 2017, the SMCS worked on 400 projects of which 389 came from UCL. 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 UCL. This is the first year we include in our statistics a cross-disciplinary support to each sector even though this has been done by the SMCS since the academic year 2014/2015. 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.

Table 7 (SSH) shows that the number of requests from the humanities sector fluctuates around 130-160 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. Only the institute for philosophy has not yet requested assistance from the SMCS.

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

	2014/15	2015/16	2016/17
SSH	**159**	**131**	**133**
Cross-disciplinary			1
IACCHOS-ISPOLE-ESPO	30	20	12
ILC-INCAL-FIAL	10	11	6
LOURIM-LSM	14	7	9
IMMAQ	7	7	10
IPSY-PSP	95	84	94
JURI-DRT		1	1
RSCS-TECO	3	1	

Table 8 (SSS) shows that on average there are almost 150 requests per year from the health sciences sector. The areas related to clinical research and the medical school account for about 60% of this, other requests are split among the remaining institutes.

Consulting

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

	2014/15	2015/16	2016/17
SSS	**139**	**120**	**150**
Cross-disciplinary			1
ASS	1	1	
DDUV-Ludwig Inst.	12	15	10
FASB	0	1	6
IONS	28	16	25
IREC-MEDE-FSM	72	69	92
IRSS-FSP	10	7	8
LDRI	9	11	5
Clinics - Albert II	7		3

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).

	2014/15	2015/16	2016/17
SST	**113**	**124**	**101**
Cross-disciplinary			1
AGRO	13	12	9
ELI	44	66	37
ICTEAM-EPL	2	2	3
IMCN		1	
IMMC	2	8	8
ISV	43	27	35
LOCI	5	5	4
SC	4	2	3
Commission doctorale		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, five projects also originated from the central administration (AC, LLL, SGSI).

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

Computing

Servers and Statistical Software

The SMCS now has two servers dedicated to statistical computing and no software sharing server anymore since Inuvika and UDS can replace it. 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

The system was renewed 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. This server system and its storage are hosted by the CISM (UCL's high performance computing platform) who 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 54 users. This year we counted users in a different way: 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. The server was mainly used by people from IMMAQ and IRSS (the table on the following page provides more details about the users of the two servers). Users may be divided in two categories: people who want to use SAS for 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 12% on smcs3 and 4% on smcs4 with spikes (100% for smcs3 and 90% for smcs4) occurring when users execute paralleled programs in Stata or R.

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

Software Sharing Server

This server was scheduled to stop in the 2015-16 academic year and has been replaced by Ucilla and UDS (see below).

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.

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 UCL 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" (cf. https://www.uclouvain.be/477533.html). Many UCL members used this service – 893 installation requests made in the academic year 2016-2017, (414 demands for SPSS, 212 for JMP, 137 for Stata, 63 for R, 43 for SAS Enterprise Guide and 24 for SAS).

In case of hardware incompatibility, a solution is available via a web portal (UDS) or a client (recommended).

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

Finally, UDS already offers the Big5-STAT softwares and Ucilla gives an access to HLM, MPlus, nQuery, PASS and Simca. All but Simca are paid for by the SGSI; Simca is paid for by the professor B. Govaerts.

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.

	2014/15	2015/16	2016/17
Total	**96**	**125**	**54**
SSH	**46**	**59**	**30**
IACCHOS-ISPOLE-ESPO	7	11	4
ILC		1	1
LOURIM-LSM	6	7	
IMMAQ	27	36	23
IPSY-PSP	6	4	2
SSS	**23**	**29**	**17**
ASS		1	
IONS	4	6	2
IREC-MEDE	3	4	1
IRSS-FSP	15	16	13
LDRI	1	2	1
SST	**27**	**36**	**7**
ELI	13	19	4
ICTEAM-EPL	2	3	
IMCN		2	
IMMC		1	
IRMP			
ISV	5	4	
LOCI			
AGRO	4	5	
EPL	1		1
SC	2	2	2
Central services		**1**	
SGSI		1	

Table 11. Users of the software sharing server.

	2014/15	2015/16	2016/17
Total	**33**	**26**	**25**
SSH	**20**	**20**	**19**
IACCHOS	2	2	1
ILC	1		
LOURIM-LSM	2	2	
IMMAQ	8	2	9
IPSY-PSP	7	14	9
SSS	**9**	**5**	**2**
IONS	4	2	1
IREC	2	2	1
IRSS	2		
LDRI	1	1	
SST	**4**	**1**	**4**
AGTO			1
ELI	1		
ICTEAM	1	1	
ISV	1		
SC	1		3

Usage of the web portal - Ucilla.
The most used softwares were SPSS (278 connections with 45 for SPSS Premium) and JMP (147) because two courses require the former and students who cannot install the later on their computers have access to SPSS on Ucilla. SAS Enterprise Guide is the third most used (108 connections); this can be explained by the fact that it allows Mac users who cannot execute SAS natively on their OS to run it. Following these softwares, we find PASS, nQuery, MPlus Stata and Minitab.

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 2016-2017, the platform helped 158 students of which 149 came for assistance with SPSS. This suggests that the available information for installation may not yet be sufficient.

Larger Projects

Larger Projects

Survey Support

The Survey Project is an initiative of the humanities sector SSH to facilitate the design and implementation of surveys. It has two components: methodology and infrastructure.

With respect to methodology, the SMCS collaborated with other UCL experts to identify and prepare training materials. The resulting training sessions were very well attended and had to be offered twice. They will be further adapted to the audience's needs.

Set up with SSH support, the Survey Project now helps SSS and SST sectors as well. It is an example of how a service such as the SMCS can help identify cross-disciplinary needs and assist in creating joint approaches for dealing with them.

The arrival of our new consultant, Séverine Guisset, will boost the « Survey Support » of our platform via a reflection on new trainings and an increased watch for external projects and/or key-projects for the university.

Research Project - Work and Organizational Psychology

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. modelization in Structural Equation Modeling), saves time and can also get some hints and ideas for future research projects.

Research Project - Erasmus+OLS

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 - Metabiose

In collaboration with ISBA, the SMCS participates in a research consortium on Metabolomics (Metabiose). In this consortium, a research group at the University of Liège is responsible for the medical and analytical aspects and the ISBA/SMCS works on the data analysis and interpretation methodology. Through this project, the SMCS consultants acquired "know-how" with respect to analyzing OMICS data. The project ended in June 2017.

Research Project – Macular degeneration

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 also a metabolomics project.

Research project – First spin-off

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

Course Module - FOREM

Since 2011/12, the SMCS provides a course in bioinformatics/biostatistics for the FOREM. This course is

Larger Projects

aimed at people who have bio-sciences degrees and difficulties finding a suitable job.

In the beginning, the course volume was 19 days. It has since then increased to 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 UCL 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.

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

Course Modules - BioWin

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. This will probably mean that our trainings will no longer be viable as exclusive sessions for government-sponsored job seekers. We will therefore try to open them up to a larger public while keeping a number of places for participants sent by BioWin.

University Certificate - Junior Data Analyst

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

The first edition already drew a lot of attention and we had fifty applications. Half of them satisfied the acceptance criteria and in the end all sixteen available places were filled with motivated participants. In fact, all continued until the end of the course.

In the meantime, many of them found work in the data analytics/science area or were able to improve their positions. We are keeping in contact and their feedback helps us in further improving the program.

For the 2017/18 edition, we again received about fifty applications and this time we made twenty places available. They were all filled.

Collaboration with the LSBA

The LSBA has asked the SMCS to gather documentation as a support for its courses. The information consists in videos, databases, and literature. This collaboration will continue during the coming years.

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 2016-2017

Trainings

- Introduction to R language, 2.5 jours, du 07/09/2016 au 09/09/2016, donnée par Alain Guillet
- Pratique de la statistique avec SPSS, 2.5 jours, du 12/09/2016 au 16/09/2016, donnée par Nathalie Lefèvre
- **NEW** - Formation en statistiques dans le domaine de la kinésithérapie : analyse de données de rééducation et réadaptation (formation organisée en partenariat avec le Centre National de rééducation fonctionnelle et de réadatation (Luxembourg)), 2 jours, du 21/09/2016 au 22/09/2016, donnée par Céline Bugli
- **NEW** - Rappels de statistique et applications avec Stata (formation organisée en partenariat avec l'Agence wallonne pour la sécurité routière), 1 jour, du 21/09/2016 au 21/09/2016, donnée par Alain Guillet
- **NEW** - Introduction au calcul de taille d'échantillon (formation organisée en partenariat avec le comité d'éthique en matière d'expérimentation animale de LLN), 1 jour, du 29/09/2016 au 29/09/2016, donnée par Catherine Rasse, Céline Bugli
- **NEW** - Certificat universitaire en Junior Data Analyst, 21 jours, du 03/10/2016 au 16/12/2016, donnée par Aurélie Bertrand, Jonathan Dedonder, Alain Guillet, Nathalie Lefèvre, Arnaud Pollaris, Catherine Rasse, Christian Ritter, Vincent Bremhorst, Marco Saerens, Formateurs externes JDA
- **NEW** - Initiation à la statistique avec applications en JMP (formation organisée en partenariat avec le FOREM), 4 jours, du 05/10/2016 au 20/10/2016, donnée par Catherine Rasse, Céline Bugli
- Initiation au logiciel NVivo, 1.5 jours, du 10/10/2016 au 12/10/2016, donnée par Jonathan Dedonder, Nathalie Lefèvre
- Formation SAS : programmation de base (formation organisée en partenariat avec SAS), 2 jours, du 11/10/2016 au 25/10/2016, donnée par Formateur SAS, Céline Bugli
- Introduction au logiciel JMP, 1 jour, du 19/10/2016 au 26/10/2016, donnée par Céline Bugli, Lieven Desmet
- Data management and analysis using Stata, 2.5 jours, du 24/10/2016 au 28/10/2016, donnée par Sébastien Fontenay, Alain Guillet
- Initiation à la statistique avec applications en JMP (formation organisée en partenariat avec le FOREM), 4 jours, du 23/11/2016 au 01/12/2016, donnée par Catherine Rasse, Céline Bugli
- Prémisses d'une enquête et création de questionnaires, 2 jours, du 28/11/2016 au 29/11/2016, donnée par Jonathan Dedonder
- Plans expérimentaux pour la recherche, le développement et la production industrielle - Applications à la pharmacie et aux sciences du vivant (formation organisée en partenariat avec BioPhare pour UCB), 3 jours, du 09/12/2016 au 13/12/2016, donnée par Catherine Rasse
- Introduction à la réalisation d'enquêtes en ligne avec LimeSurvey, 1 jour, du 12/12/2016 au 12/12/2016, donnée par Jonathan Dedonder, Nathalie Lefèvre
- Exploration de données et base de modélisation statistique avec JMP (formation organisée en partenariat avec BioPhare), 3 jours, du 15/12/2016 au 20/12/2016, donnée par Catherine Rasse
- **NEW** - Pratique de la statistique avec R, 2.5 jours, du 02/01/2017 au 06/01/2017, donnée par Alain Guillet
- Pratique de la statistique avec SPSS, 2.5 jours, du 09/01/2017 au 13/01/2017, donnée par Nathalie Lefèvre, Jonathan Dedonder
- Plans d'expérience avec JMP (formation organisée en partenariat avec BioPhare), 3 jours, du 12/01/2017 au 16/01/2017, donnée par Catherine Rasse, Bernadette Govaerts

Appendix

- Formation en analyse de données et statistiques pour la recherche et le développement dans la bioindustrie (formation organisée en partenariat avec le Forem), 34 jours, du 16/01/2017 au 16/05/2017, donnée par Aurélie Bertrand, Céline Bugli, Bernadette Govaerts, Alain Guillet, Manon Martin, Catherine Rasse, Lieven Desmet
- Analyses multivariées en RCommander (FactoMineR), 2 jours, du 23/01/2017 au 24/01/2017, donnée par Arnaud Pollaris, Benjamin Colling
- Introduction aux modèles multi-niveaux avec SPSS (formation organisée en partenariat avec l'institut IPSY), 2.5 jours, du 30/01/2017 au 03/02/2017, donnée par Aurélie Bertrand, Céline Bugli, Nathalie Lefèvre
- Initiation au logiciel NVivo, 1.5 jours, du 30/01/2017 au 01/02/2017, donnée par Jonathan Dedonder, Nathalie Lefèvre
- Introduction to R language (formation organisée en partenariat avec la Faculté de Santé Publique), 3 jours, du 07/02/2017 au 14/02/2017, donnée par Joris Chau, Nathan Uyttendaele
- Formation SAS : SAS advanced programming (formation organisée en partenariat avec SAS), 5 jours, du 14/02/2017 au 28/03/2017, donnée par Formateur SAS
- Analyser la médiation, la modération et la médiation modérée avec SPSS, 2 jours, du 16/02/2017 au 17/02/2017, donnée par Arnaud Pollaris
- Initiation aux modèles d'équations structurales (SEM) avec Stata, 2.5 jours, du 13/03/2017 au 17/03/2017, donnée par Arnaud Pollaris, Jonathan Dedonder
- Inférence et modélisation statistique avec JMP, 1 jour, du 15/03/2017 au 22/03/2017, donnée par Céline Bugli, Lieven Desmet
- Initiation au logiciel NVivo, 1.5 jours, du 27/03/2017 au 29/03/2017, donnée par Jonathan Dedonder, Nathalie Lefèvre
- Gestion et analyse de données avec Stata (formation organisée en partenariat avec ESPO), 2.5 jours, du 03/04/2017 au 07/04/2017, donnée par Sébastien Fontenay, Alain Guillet
- Utilisation avancée du logiciel Nvivo, 1 jour, du 20/04/2017 au 21/04/2017, donnée par Jonathan Dedonder
- **NEW** - La statistique, outil du chercheur et de l'ingénieur, 0.5 jour, du 28/04/2017 au 28/04/2017, donnée par Catherine Rasse
- Analyse de données pré-cliniques et cliniques avec JMP, 1 jour, du 10/05/2017 au 24/05/2017, donnée par Céline Bugli, Lieven Desmet
- Cours d'initiation à SAS, 2.5 jours, du 15/05/2017 au 17/05/2017, donnée par Lieven Desmet
- Pratique de la statistique avec SPSS (formation organisée en partenariat avec l'Université de Mons), 2.5 jours, du 17/05/2017 au 19/05/2017, donnée par Arnaud Pollaris
- Panel-data Analysis using Stata, 1.5 jours, du 18/05/2017 au 19/05/2017, donnée par Sébastien Fontenay
- Cours d'initiation au langage R, 2.5 jours, du 12/06/2017 au 16/06/2017, donnée par Alain Guillet
- **NEW** - Cartographie statistique avec R, 1.5 jours, du 19/06/2017 au 21/06/2017, donnée par Isabelle Thomas, Gaëtan Montero Redondo, Jonathan Jones
- **NEW** - Introduction aux analyses factorielles exploratoires avec SPSS (formation organisée en partenariat avec l'institut IPSY), 1 jour, du 30/06/2017 au 30/06/2017, donnée par Arnaud Pollaris, Nathalie Lefèvre
- Introduction à l'analyse de Réseaux Sociaux avec Gephi, 1.5 jours, du 03/07/2017 au 05/07/2017, donnée par Jonathan Dedonder, Arnaud Pollaris

Appendix

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

Daems, Frédéric; Romnee, Jean-Michel; Rasse, Catherine; Froidmont, Eric; Heuskin, Stephanie; Lognay, Georges. *Quantification of Four Isoflavones in Forages with UPLC®-MS/MS, Using the Box–Behnken Experimental Design to Optimize Sample Preparation*, *Chromatographia* 79(11-12), April 2016.

Daems, Frédéric; Romnee, Jean-Michel; Rasse, Catherine; Froidmont, Eric; Lognay, Georges. *Optimized Quantitative Method for Determining Isoflavones and Equol in Bovine Digestive Fluids and Feces*. In: *Analytical Letters*, June 2017.

Dormal, Valérie ; Larigaldi, Nathanaël ; Lefèvre, Nathalie ; Pesenti, Mauro ; Michaël, Andres. *Effect of Perceived Length on Numerosity Estimation: Evidence from The Müller-Lyer Illusion.* Accepted in Quarterly Journal of Experimental Psychology.

Féraud, Baptiste ; Munaut, Carine ; Martin, Manon ; Verleysen, Michel ; Govaerts, Bernadette. *Combining strong sparsity and competitive predictive power with the L-sOPLS approach for biomarker discovery in metabolomics*. In: *Metabolomics*, 13: 130 (2017). https://doi.org/10.1007/s11306-017-1275-y

Guillet, Alain ; Leblet, Jimmy ; Rampon, Jean-Xavier. *Faithful extension on finite order classes*. In: *The Australasian Journal of Combinatorics*, Vol. 69 Part 1, no.1, p. 1 (2017). http://hdl.handle.net/2078.1/187367

Seminars and Presentations

Extra-muros seminars

Bugli, Céline & Desmet, Lieven. *Design and Analysis of Clinical Trials.* Séminaire dans le cours interuniversitaire de la SBMI (Soc. Belge de Médecine Interne), Saint Luc, le 18 novembre 2016.

Bugli, Céline & Lefèvre, Nathalie. *Le défi d'un support statistique à la recherche, à l'enseignement et à la société.* Présentation dans le cadre du Groupe Interdisciplinaire de Recherche, CHU Godinne, le 24 mars 2017.

Bugli, Céline & Desmet, Lieven. *Why and How Randomise a Clinical Trial.* IREC monthly seminar series, le 29 mai 2017.

Bugli, Céline & Desmet, Lieven. *Statistical Methodology and Computing Service dedicated to research: SMCS.* Présentation lors du Phd Day de l'IREC, Woluwé, le 10 mars 2017.

Cosse, Marie-Pascale & Lefèvre, Nathalie. *De la réforme des BAC PSP à un enseignement innovant et collaboratif de la statistique.* Intervention lors du Forum FDP du LLL, Louvain-la-Neuve, 2 juin 2017.

Presentations and applied statistics workshops

Pollaris, Arnaud. *Création de questionnaires. Quelques pièges à éviter…* Intervention dans le cadre de l'Atelier de statistique appliquée, Louvain-la-Neuve, le 17 mars 2017.

Guillet, Alain. *Atelier d'introduction aux forêts aléatoires.* Midi du SMCS, Louvain-la-Neuve, le 21 février 2017.

Lefèvre, Nathalie. *Atelier d'initiation à l'outil Wooclap.* Midi du SMCS, Louvain-la-Neuve, le 18 avril 2017.

Appendix

Collaborations and Supervisions

Collaborations

Bugli, Céline. Aide à la rédaction d'offres de la plateforme Centre d'Investigation Clinique en Nutrition, année académique 2016-2017.

Bugli, Céline. Préparation d'un dossier Région Wallonne pour la société OneLife, année académique 2016-2017.

Bugli, Céline. Supervision d'analyse de données RMN de patients atteins de DMLA, année académique 2016-2017.

Bugli, Céline. Membre depuis 2013 du bureau exécutif de la Royal Statistical Society of Belgium.

Cosse, Marie-Pascale. Organisation des Midis pédagogiques de la LSBA, Louvain-la-Neuve, année académique 2016-2017.

Cosse, Marie-Pascale & Lefèvre, Nathalie. Participation au comité techno-pédagogique de la LSBA, Louvain-la-Neuve, année académique 2016-2017.

Rasse, Catherine. Participation à la commission d'éthique animale de Louvain-La-Neuve, année académique 2016-2017.

Supervision of Master and PhD thesis

Bugli, Céline. Participation au comité d'encadrement de thèse de Florian Pechon, année académique 2016-2017.

2016-2017