

REPORT OF ACTIVITIES

Statistical Methodology and Computing Service

1 September 2020 to 31 August 2021



2020-2021

Statistical Methodology and Computing Service
Report of Activities
1 September 2020 to 31 August 2021

Contents

Introduction.....	4
People and Organization	5
Management and Administration	5
Management committee.....	5
User committee.....	5
People.....	5
Associated Consultants	6
Professors.....	6
Consultants at other institutes	7
Researchers and Research Assistants.....	7
Interns	7
Trainers and Consultants.....	7
Training	8
Consulting.....	11
Servers and Statistical Software	13
Statistical Computing Server	13
UDS.....	13
Available Software.....	13
Usage Statistics.....	14
Help for students.....	14
Larger Projects	15
Research Project - Erasmus+OLS (2015-2021)	15
Research project - CluePoints (2018-2022)	15
Research project - First spin-off (2018-2021).....	15
Research Project - PDR (2020-2024)	15
Survey Support	15
Support to other platforms	16
Project linked to education - Development of teaching materials.....	16
Project linked to education - Dynamo project	16
University Certificate - Analyse de données en R&D (2020-...)	16
Summer/Winter School - Collaboration with the CENTAL	17
External project - Training for journalists.....	17
External project – Proof-Next (2019-2021)	17
Your projects? SMCS can be your partner!	17
Appendix	18
List of Trainings and Seminars 2020-2021.....	18
Trainings	18

Contents

Publications and Presentations by SMCS members	20
Publications.....	20
Seminars and Presentations	20

Introduction

Introduction

“What a long strange trip it’s been” is the name of a Grateful Dead album, and this is how things have felt since spring 2020. Yet, the SMCS managed to continue and even expand its activity as a diversified service provider in practical statistics. We are now 8 full time equivalents of consulting staff and a network of more than 20 associated specialists; all to assist you, UCLouvain researchers and external clients in data collection, processing, analysis, and interpretation. This report describes our activities over the past three academic years focusing on 2020/21.

Consultancy has grown to about 550 projects per year. *Training* reached its highest volume ever with 46 courses (mostly online) representing 189 training days. 662 participants attended for a total of 2394 person days. The *statistical computing and software-sharing servers* had 69 and 39 registered users respectively.

The SMCS continues improving its visibility towards potential new users (researchers, administrators, staff at the university hospitals). As usual, we distributed our newsletters explaining our goals, expertise, and services via the three sectors. We also participated in the “journée plateformes” organized by ADRE, UCLouvain.

Customer satisfaction remained high. In late 2020, we invited more than 800 clients of our services between 2017 and 2020 to an online satisfaction survey. The large majority (97%) of the 149 (18%) respondents were satisfied and would recommend us. The respondents highlighted the quality of the interaction with the team. Aside from building on our strengths, we are working on improving the user experience of our web platform.

The SMCS has continued assisting the UCLouvain administration. This includes urgent work on coordinating the online and face-to-face exams to respect the evolving COVID safety requirements. Another project concerns the data extraction and management for studying student trajectories in undergraduate and graduate programs.

Our expertise in working with survey tools helped in converting institutional elections at UCLouvain to a fully online format. Moreover, we began work on a new survey about the mobility of UCLouvain staff and students. This

survey gathers data to fulfill official legal requirements and will increase understanding of mobility patterns.

Among the new external collaborations, we would like to mention a project with the CIM (Centre d’Information sur les Médias) on handling missing data in the radio audience survey. Another project concerns transformer load patterns for ORES, the largest Walloon distribution grid operator.

Anticipating continuing partial COVID restrictions, we decided to convert our summer school of 2021 to a winter school in January 2022. Let’s hope that this will work out. We also used the pandemic challenges as an opportunity to develop our expertise in e-learning and e-examination and support others in doing so. This includes a collaboration with EPHEC to assist beginning students in Human and Social Sciences. This is very important now as the new cohort of beginning students finished their secondary school under difficult circumstances and who now must cope with additional challenges due to sanitary restrictions. In collaboration with the Royal Statistical Society of Belgium, the SMCS was involved in the first session of a training aimed at journalists about the correct reporting and interpretation of numbers and graphs (in Dutch and French).

Three new university certificate courses have started, a modular version of the Junior Data Analyst, and two versions, basic and advanced, of a program named Data Analysis in Research and Development. In the context of the Junior Data Analyst, we organized a successful Job Day at the end of 2020.

Finally, I would like to congratulate the team for their professionalism and dedication to keep things running. Our daily virtual coffee break has certainly helped. We shall have to continue like that since remote working will remain common for months to come.

Last but not least, the SMCS has more than 500 followers on LinkedIn! Are you already part of this growing community?

As usual, your input, assistance, support, and curiosity will be greatly appreciated. Enjoy reading, Christian Ritter - President of the management committee.

People and organization

People and Organization

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

Management committee

The management committee (comité de gestion) met twice. The Fall meeting has focused on long term planning whereas the Spring meeting has concentrated on finances, project accounting, and short-term planning. A continuing problem is uniting all members on these two dates.

President: Christian Ritter

Director: Nathalie Lefèvre

SSH representative (LIDAM institute): Fabio Mariani

SSS representative (IRSS institute): Niko Speybroeck

SST representative (ELI institute): Claude Bragard

Representative of research activities (ISBA): Catherine Legrand

Representative of training activities (LSBA): Johan Segers

Associate consultant academic representative:

Bernadette Govaerts

Consultant representative: Céline Bugli

With consultative voice:

Administrative coordinator: Nancy Guillaume

Representative of external activities: Antoine Froidure (St Luc)

Representative for each University entity with which the platform has IT collaborations: Thomas Keutgen for CISM and Laurent Buset for SGSI

User committee

With respect to the user committee, the format of an afternoon with public presentations of case studies organized on the same day as the official meeting restricted to the official representatives has again been very well-received. In 20-21, due to the Covid crisis, we

organized a remote meeting with the user representatives.

As a reminder, the user committee is a precious relay: relaying the needs of the users of their entities to the SMCS and relaying the services offered by the SMCS within their entities.

Bruno Schoumaker (IACCHOS), Magali Paquot (IL&C), Marco Saerens (LOURIM), Françoise Muller (INCAL), Michael Andres (IPSY), Cécile Bonmariage (ISP), Pierre Baudewyns (ISPOLE), Frédéric Vesentini (JUR-I), Michael Andres (IoNS), Annie Robert (IREC), Vincent Lorant (IRSS), Laure Elens (LDRI), Laurent Gatto (DDUV), Nicolas Schtickzelle (ELI), Jean-Charles Delvenne (ICTEAM), Tom Leyssens (IMCN), Vincent Destoop (IMMC), Andrea Giammanco (IRMP), Isabelle Donnay (LIBST), Isabelle Thomas (LIDAM), Frédéric Vrins (LIDAM-Mons), Françoise Docq (ADEF), Patrick Watrin (CENTAL), Damien François (CISM), Laurent Buset (SGSI)

People

Given the workload of the SMCS team, the management committee validated the hiring of an additional consultant (1FTE) and financed by the SMCS funds. Benjamin Colling, PhD in Statistics, started in September 2020. The network of data analysis consultants is generally expanding at UCLouvain in particular with the arrival of Gipsi Lima Mendez, hired in 2019 as an SST research logistician to provide precious support in bioinformatics. Following her departure in 2021, it was decided to recruit for a similar function, strategic for bioinfo support.

Consultants



Aurélie Bertrand (UCLouvain/LIDAM/SMCS – 1 FTE)

People and organization

Main areas of activity: Human Sciences 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)

Vincent Bremhorst (UCLouvain/LIDAM/SMCS – 1 FTE - since September 2018)

Main areas of activity: Human Sciences Sector (SSH), Science and Technology Sector (SST), bayesian support.

Doctorat en Sciences – Statistique (UCL, 2017)
Master en statistique (ULB, 2011)
Bachelier en mathématique (ULB, 2009)

Céline Bugli (UCLouvain/LIDAM/SMCS – 0.8 FTE)

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

Doctorat en Sciences – Statistique (UCL, 2006)
Master (DEA) en statistique (UCL, 2003)
Ingénieur Civil en Mathématiques Appliquées (UCL, 2001)

Benjamin Colling (UCLouvain/LIDAM/SMCS – 1 FTE)

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

Doctorat en Sciences – Statistique (UCL, 2017)
Master en Statistique (UCL, 2011)
Master en Mathématique (UCL, 2010)

Lieven Desmet (UCLouvain/LIDAM/SMCS – 0.95 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 (UCLouvain/LIDAM/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 (UCLouvain/LIDAM/SMCS – 0.8 FTE)

Main areas of activity: Human Sciences 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)

Nathalie Lefèvre (UCLouvain/LIDAM/SMCS – 0.5 FTE and UCLouvain/IPSY – 0.5 FTE)

Main areas of activity: Human Sciences 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).

Catherine Rasse (UCLouvain/LIDAM/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).

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

Michel Denuit (LIDAM/ISBA), Sébastien de Valleriola (ICHEC and prof. Invité UCLouvain), Thomas François ILC/PLIN), Laurent Gatto (DDUV), Bernadette Govaerts (LIDAM/ISBA), Marie-Paule Kestemont (IACCHOS), Vincent Lorant (IRSS), Catherine Legrand (IDAM/ISBA), Bruno Masquelier (IACCHOS), Eugen Pircalabelu (LIDAM/ISBA), Christian Ritter LIDAM/LSBA), Bruno

People and organization

Schoumaker (IACCHOS), Nicolas Schtickzelle (ELI), Isabelle Thomas (LIDAM/CORE), Vincent Vandenberghe (LIDAM/IRES)

Consultants at other institutes

Jonathan Dedonder (SSH/IACCHOS), Guillaume Dallemagne (SSH/LIDAM/IRES), Nathalie Lefèvre (SSH/IPSY), Manon Martin (SSS/DDUV), Jérôme Ambroise (SSS/IRES), Aline van Maanen (Institut Roi Albert II - Cancérologie et Hématologie), Gipsi Lima Mendez (SST)

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

Cyrille Nguipdop Lowe, Metnatpar project – mid-June to mid-September 21.

Trainers and Consultants

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

Aurélien Bertrand (SMCS), Vincent Bremhorst (SMCS), Céline Bugli (SMCS), Benjamin Colling (SMCS), Guillaume Dallemagne (IRES), Jonathan Dedonder (IACCHOS), Lieven Desmet (SMCS), Laurent Gatto (DDUV), Bernadette Govaerts (ISBA), Alain Guillet (SMCS), Séverine Guisset (SMCS), Alexandre Jacquemain (ISBA), Eric Lecoutre (WeLoveDataScience!), Nathalie Lefèvre (SMCS), Adam Levai (IRES), Vincent Lorant (IRSS), Rebecca Marion (ISBA), Manon Martin (DDUV), Sophie Mathieu (ISBA), Gilles Mordant (ISBA), Emmanuel Niyigena (ISBA), Eugen Pircalabelu (ISBA), Catherine Rasse (SMCS), Christian Ritter (LSBA), Christophe Robyns (Agilytic), Xavier Rouby (Pfease), Marco Saerens (LRIM), Antoine Soetewey (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. Six of the training sessions took place on the Woluwe campus. Nine training courses have been given in English. We will continue to propose every year more training courses in English but for very advanced trainings, participants generally prefer that they be given in their mother tongue.

Two training sessions (4.5 days) had to be cancelled because there were not enough participants. A solution of personalized coaching has been proposed to the participants who had registered. Based on the two last years' experience, some of our training courses will only be proposed upon specific request. We have the highest statistics ever for training courses, both in terms of the number of training days and person days. Last year, 8 training courses had to be cancelled and reported because of the Covid-19 pandemic. They have been proposed this year, which already explains this increase. Furthermore, the advanced R&D certificate proposed for the first time this academic year was organized twice, which is exceptional.

Almost all training courses were organized at a distance, with a few exceptions in September 2020.

It has been the opportunity for the team to innovate for example, some video supports were created in this context.

The experience of distance learning has led us to think about keeping some courses in distance learning in the future, obviously not all.

Seven new training programs have been developed including 1 new university certificate.

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: "*Modèles multi-niveaux avancés avec SPSS*" or "*Basic concepts and practical applications of Bayesian statistics*" (organized in collaboration with IPSY but open to everyone), "*Analyse de données Omics sur base de techniques multivariées avec R*" (organized for and funded by GSK but open to UCLouvain researchers).
- **University certificate.** In collaboration with the LSBA, the SMCS is also involved in and/or coordinate university certificates. This year, 1 new certificate has been introduced and was organized remotely: "University Certificate: Applied Data Analysis in R&D - advanced level. It has been exceptionally organized twice during the academic year 20-21. This certificate complements the basic level one and the two Junio Data Analyst certificates.
- **Summer/Winter School:** A one-week in-depth training on an advanced topic given by invited speakers who are specialists in the field. The first edition took place in July 2019. It was about Text mining with the Python language (in French) and was organized in collaboration with the Cental (Centre de traitement automatique du langage, UCLouvain). The specificity of the program was to propose, besides courses, some social activities which have been greatly appreciated by the participants. Unfortunately, the 2020 edition, on a similar theme, had to be postponed to 2022 because of COVID. It will be given in January 22.
- **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 46 courses in the academic year 2020-2021 for 189 teaching days. The

Training

decrease observed last year mainly due to the pandemic has resulted in a very significant increase this year in all 3 sectors, due to postponed training.

Table 1. Global statistics on SMCS trainings.

	18/19	19/20	20/21
Courses	52	39	46
Course Days	172	130	189
Participants UCLouvain	493	402*	524
Person-Days* UCLouvain	1425	1435*	2106.5
Participants Ext	253	120	138
Person-Days* Ext	1000	262	287.5

*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. Table 3, Table 4 and Table 5 show the origin of participants from respectively SSH, SSS and SST sectors.

Table 2. Number of participants per sector

	18/19	19/20	20/21
SSH	177	177*	196
SSS	97	82	134
SST	54	71	81
Students	165	71*	110
Central services		1	3

The observed fluctuations between years and sectors reflect natural variation and the organization of special events.

Compared to previous years, there is an overall increase in the number of UCLouvain participants in all sectors.

The statistics training program reaches the entire SSH sector. The high participation rate of IPSY-PSP and IACS-SPLE-ESPO stems from the fact that the IPSY and ISPOLE Institute finances the organization by the SMCS of specific trainings. The high participation of IACS-SPLE-ESPO are also 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 Human Sciences Sector (SSH)

	18/19	19/20	20/21
SSH	247	203*	232
IACS-SPLE-ESPO	54	60*	60
ILC	18	23*	27
LOURIM-LSM	13	3	23
LIDAM	37	19*	29
INCAL-FIAL	5		8
IPSY-PSP	48	65*	47
JURI-DRT	1	1	
RSCS-TECO	1	6	2
Students	70	26	36

In the health sciences sector courses were also 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 updated the training format to take this particularity into account.

This year, the distance learning format of most of the trainings may have made it easier for SSS participants to participate in the trainings than if they had been organized on campus.

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

	18/19	19/20	20/21
SSS	112	94	141
ASS		1	
DDUV	5	16	3
IONS	13	15	56
IREC-MEDE-FSM	44	29	40
IRSS-FSP	17	9	15
LDRI-FASB	14	10	18
Clinics	4	2	2
Students	15	12	7

In the science and technology sector courses were also followed by participants from a wide range of institutes.

*Figures adjusted since the last report - Attendance at training not counted last year

Training

Nevertheless, more than 60% of the SST staff participating in SMCS training comes from the ELI Institute.

such trainings to make sure there is a positive return (financial and others).

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

	18/19	19/20	20/21
SST	134	104*	148
AGRO	6	2	3
ELI	26	47	49
ICTEAM-EPL	2		3
IMCN		1	
IMMC		9	5
IRMP	1		1
LIBST	17	5	10
LIDAM		3	2
LOCI-LAB		3	8
SC	2	1	
Students	80	33*	67

The SMCS is also regularly invited to sessions organized by other entities and these presentations are not included in the training statistics.

Overall, the training activity of the SMCS occupies about 3.4 full time staff equivalents[‡] (FTE). 3.1 FTE of this effort is directly oriented towards the UCLouvain research and teaching units, the remainder towards external clients, such as the "Mutualités Chrétiennes" 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 and to propose new accredited certificates. 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

**Figures adjusted since the last report - Attendance at training not counted last year*

[‡] The calculation of this time is based on the duration of the training courses multiplied by a coefficient usually used to take into account the preparation time of the training courses. This year, this coefficient

should have been increased because the organization of remote learning courses took a lot of time due to the preparation of specific materials.

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 1.9 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 was still particularly important this year and corresponds to more than 4450 hours (1707 hours for external projects and 2750 for internal one’s).

Overall, the consulting activity of the SMCS occupies about 5.8 full time staff equivalents (FTE). 4.2 FTE of this effort is directly oriented towards the UCLouvain research, teaching and administrative units, the remainder towards external clients.

Table 6. Origin of consulting projects

Client split	18/19	19/20	20/21
Research	425	450	495
Educational project	9	14	5
Admin. services	9	23	21
External clients	10	18	33
Total	453	505	554

From 1 September 2020 to 31 August 2021, the SMCS worked on 554 projects of which 521 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 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. We include in our statistics cross-disciplinary supports to each sector. This consists in statistical computing support for all the sector community or projects that target several researchs or teaching entities, such as a website project to bringing together statistical resources to help future students in Human Sciences Sector, or a project of collection, identification, implementation and sharing of good data protection practices (a GDPR project). 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 Human Sciences Sector usually fluctuating around 130 requests per year, increased the last years to now nearly 190 files (1463 hours of work). As in previous years, the institute and the school of psychology had by large the highest share. Therefore, 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). IACCHOS and IRES/ECON also have an internal specialist for data questions matters (Jonathan Dedonder for IACCHOS, Guillaume Dallemagne for IRES/ECON), who is also associated member of the SMCS. Besides IPSY with particularly high activity, requests came from almost all units of this sector.

Consulting

Table 7. Origin of consulting requests from the Human Sciences Sector (SSH)

	18/19	19/20	20/21
SSH	169	187	182
Cross-disciplinary	2	4	3
IACCHOS-ISPOLE-ESPO	16	26	33
ILC-INCAL-FIAL-ISP	16	20	20
LOURIM-LSM	14	7	11
LIDAM	26	20	15
IPSY-PSP	93	105	96
JURI-DRT	2	5	4

Table 8 (SSS) shows that the number of requests from the Health Sciences Sector usually fluctuating around 150 requests per year, increased the last years to now 200 files (846 hours of work). The areas related to clinical research, Health and Society Institute and the medical school account for about 50% of this, other requests are split among the remaining institutes.

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

	18/19	19/20	20/21
SSS	167	198	206
Cross-disciplinary	1	1	1
DDUV	7	6	2
FASB	3	7	3
IONS	18	24	49
IREC-MEDE-FSM	87	104	104
IRSS-FSP	39	30	28
LDRI	3	13	14
Clinics	9	13	5

Table 9 (SST) shows that the number of requests from the science and technology sector fluctuates around 90-120 requests per year, with an increase this year to 121 files (1038 hours of work). The number of hours worked in SST is much higher than in previous years. ELI and LIBST account for about 65% of this, other requests came from almost all units of this sector.

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

	18/19	19/20	20/21
SST	98	92	121
Cross-disciplinary	2	2	4
AGRO	10	7	8
ELI	28	36	53
ICTEAM-EPL	5	3	1
IMCN	1		
IMMC	7	6	19
LIBST	36	29	25
LIDAM	1	1	3
LOCI-LAB	4	2	3
SC	4	6	5

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, **12 projects** also originated from the central administration (AC, LLL, SGSI, BIUL). The last two years, support to central administration and administrative services in general was particularly high, with a total of 1427 hours of work. This is explained first of all by the establishment of a new collaboration between SASP and SMCS on an institutional project to monitor success in BAC and MASTER. It is also explained by the support provided by SMCS on issues arising from the pandemic (remote elections and management of the exam organization). We can add the involvement of SMCS in the implementation and follow-up of the institutional survey on mobility. The statistics shown above do not present in details large-scale internal and external projects documented in a later section of this report.

Computing

Servers and Statistical Software

The SMCS has two servers dedicated to statistical computing managed in collaboration with the CISM. There also is one with the SGSI 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 in October 2019. Our current two performance servers are based on a dual Intel Xeon Gold 6144 processor with 192 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 65 users. People from LIDAM and IACCHOS 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 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 100% on smcs1 and 2% on smcs2 with spikes (500% for smcs1 and 180% for smcs2) 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 and is a good first step before using bigger HPC infrastructures like Manneback or Lemaitre3.

UDS

UDS is a service provided by the SGSI that allow users who asked for an access to use software without having it on their computer. It works through a web browser.

It is possible to add software to UDS. 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 a.k.a. 8282 or even better via the "Catalogue de service". Many UCLouvain members used this service but there still are too many PhD students who don't know about it and buy a software they could get for free.

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 and PASS.

The access conditions and procedures about a given software for the statistical computing servers and the software-sharing server are available in every software section when they are concerned on our website <https://www.uclouvain.be/smcs> (identification is required) or you can ask at the Service desk

Computing

Usage Statistics

Table 10. Users of the statistical computing server

	18/19	19/20	20/21
Total	65	57	69
SSH	35	35	37
IACS-SPLE-ESPO	8	15	17
ILC	1		1
LIDAM	24	20	19
IPSY-PSP	2		
SSS	12	10	14
IONS			
IREC-MEDE	1	2	3
IRSS-FSP	10	8	11
LDRI	1		
SST	18	12	18
ELI	5	3	6
ICTEAM-EPL	1	2	4
IMCN	1	2	
IMMC	5	4	4
IRMP	1	1	2
LIBST	1		1
SC	4		1

Usage of UDS

The most used software was PASS (16 different users), followed by SPSS Premium (11 users), MPlus (7 users) and HLM (5 users) during the year. Surprisingly, only one user of PASS came from SST whereas they need such a software to justify the sample size of their experiment when they enter a new project that the Ethical committee needs to validate. More generally, PASS should be more known since researchers need more and more to justify the chosen sample size.

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

	17/18	19/20	20/21
Total	27	47	39
SSH	17	28	29
IACCHOS	3	4	1
ILC		2	2
IPSY-PSP	8	12	12
JURI		1	
LIDAM	4	7	11
LOURIM-LSM	2	2	3
SSS	10	16	9
DDUV		1	
IONS-FSM	4	6	2
IREC-MEDE	4	6	4
IRSS	1	1	2
LDRI	1	2	
Clinics			1
SST	0	3	1
AGRO			
ELI		2	1
IMMC			
LIBST			
LOCI-LAB		1	
SC			

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. For the first year, the SGSI decided to distribute software only through courses with a support by the professor and the assistants.

Larger Projects

Larger Projects

This section summarizes the largest projects in which the SMCS was involved in 2020-2021. Some of these projects are research projects coordinated by academic members of UCLouvain or other universities. Other projects consist in providing support: regarding surveys, teaching activities, but also data management and data extraction (for the UCLouvain administration). Finally, some university certificates also belong to this category.

Research Project - Erasmus+OLS (2015-2021)

The European project Erasmus+OLS is a seven-year project to provide online language support to several hundred thousand participants in Erasmus+ programs, via a language assessment and an online course.

Our role in this project is to provide statistical expertise, in particular to monitor the assessment, and design and analyze surveys. We use text mining methods to analyze open survey answers and provide regular reports to the project team.

Research project - CluePoints (2018-2022)

In collaboration with CluePoints' R&D department, the SMCS is involved in activities of research and validation of solutions for Central Statistical Monitoring of Clinical Trials.

Research project - First spin-off (2018-2021)

The SMCS helped a research unit from IREC with the elaboration of a first UCLouvain spin-off funding application (Service public de Wallonie). The SMCS is responsible for programming an algorithm for quantifying a new biomarker.

Research Project - PDR (2020-2024)

The SMCS is involved in research project METNATPAR for the identification of the metabolic and/or the biochemical pathways that are triggered by promising natural antiparasitic derivatives in *Trypanosoma brucei* and *Plasmodium falciparum*.

Collaborative research project - Survey among book authors in Fédération Wallonie-Bruxelles

This survey was implemented within the framework of a cooperation agreement between SMCS and Maison des Auteurs/Bela asbl, with a grant from Fédération Wallonie-Bruxelles. The main objective is to study the socio-economic standing of authors in Wallonia and Brussels, and more specifically the diversity of their professional activities and income (type, structure and amounts). The SMCS took part in the questionnaire design, programming and data analysis. The study will lead to a final report presenting the findings.

UCLouvain Mobility Survey

The university will launch a survey on mobility among staff and students in October 2021. The objectives are to gather data to fulfill federal and regional legal requirements while better understanding mobility behaviours and patterns to feed future mobility strategies. We participated in the questionnaire design and programming and organized two tests of the questionnaire. We will collect and analyze the data, as well as prepare various reporting outputs in the autumn.

Survey Support

Survey support consulting was provided for a wide range of applications in all three sectors in 2019-2020 and included advice on survey and questionnaire design, programming in LimeSurvey, sampling, data management, weighting and analysis.

Support was provided to 3 surveys on the European Researchers' Night 2020. SMCS is also providing ongoing support to the team working on the European Social Survey in French-speaking Belgium.

Two to three training sessions linked to surveys are given as standard each year, with additional training sessions tailored to specific requests or audiences.

Several elections for presidents of institutes, deans and vice-rectors took place in a digital format. The SMCS contributed to the implementation and supervision of secure election processes via LimeSurvey for each of these. This process was implemented 8 times with success, including during the election of the SSH Vice-Rector and the elections of the CORTA representatives in SSH, SSS, SST and in the administrative departments.

Larger Projects

The project of a database gathering information on surveys conducted by UCLouvain, scheduled for 2020, was put on hold and will be on the agenda for 2022.

Data management support - (Online) exam sessions (2020-...)

Since the Covid-19 pandemic, the SMCS is involved in the coordination of the exam sessions. The SMCS merges the information on the exam schedule and exam registrations to evaluate the number of students on site at any time of the session on each geographical site of UCLouvain or using a software as Gradescope or Moodle. Several lists of students and summary tables were provided to ensure a proper organization of the online exams.

Data management support - Student path data (2019-...)

In collaboration with the "Service d'Analyse de Stratégie et de Prospective (SASP)", indicators related to the evolution of the duration of studies and the evolution of the percentage of students graduated on time are generated. Dashboards are developed and maintained to summarize such indicators numerically and to graphically illustrate their evolution. Up to now, the first generation UCLouvain student cohorts and the UCLouvain master's first entrance cohorts are considered.

Support to other platforms

The SMCS assists the CICN (Clinical Investigation Centre for Nutrition) platform in writing protocols for clinical studies in nutrition by reviewing study design, calculating sample sizes and creating randomization lists. The SMCS also assists in setting up statistical analyses or realizes some of them.

Project linked to education - Development of teaching materials

Between April and September 2021, the SMCS was involved in the follow-up of a project related to the preparation of new students for higher education that took place between July 2019 and August 2020 in collaboration with the EPHEC and the Pole Louvain. We improved the website (<https://aidealareussite.uclouvain.be/course/view.php?id=65>) that

brings together existing and new resources to help future students in Human Sciences Sector in the Pole Louvain who start their first course in statistics in the best conditions. We also carried out communication campaigns about the existence of this website.

One member of our team is also responsible for the maintenance of the Shiny applications of the project Shiny@UCLouvain (<https://sites.uclouvain.be/RShiny/main.html>).

Project linked to education - Dynamo project

This project aims to give tools to the teacher to evaluate how well a group of students works in using indices in Moodle. Students have to fill in a short survey from which we will give group performances. This index will then be implemented in Moodle interface.

The SMCS provides its statistical expertise in order to find the best possible index.

University Certificate - Analyse de données en R&D (2020-...)

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, the course is since January 2020 valued in 2 university certificates.

Larger Projects

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 (online due to the Covid-19 pandemic) for the fifth edition from October 2020 to December 2020 and is self-financed since the third edition. The sixth edition will start in October 2021. In January 2020, the first edition of a modular JDA has been launched. This modular program is more compatible with a professional activity because half of the courses are organized on Saturdays. The course of the second edition started (online due to the Covid-19 pandemic) in January 2021. A third edition will start in January 2022. The modular program is self-financed since its first edition.

Summer/Winter School - Collaboration with the CENTAL

In 2018-2019, the SMCS launched a new type of training, a Summer School, focused on an advanced topic and given by invited specialists. The first edition was about Text Mining in Python, in collaboration with the Cental. Following this first successful edition, the second one, about Text mining and Social Network Analysis in Python (again in collaboration with the Cental), is finally (after a cancellation in 2020 due to the COVID pandemic) rescheduled in January 2022.

External project - Training for journalists

In collaboration with the Royal Statistical Society of Belgium, the SMCS created, in 2019-2020, material for two versions (one in Dutch, the other in French) of a training aimed at journalists about the correct reporting and interpretation of numbers and graphs. The SMCS was involved in the first session of the training in each language in 2020-2021.

External project – Proof-Next (2019-2021)

In collaboration with the Belgian Respiratory Society, the SMCS is involved in the analysis of the Proof-Next registry of Idiopathic Pulmonary Fibrosis patients.

External Project - Multivariate analysis of the production of household waste and assimilated in Wallonia

This project consists in the analysis of data about production of household waste in Wallonia. The SMCS was involved in the review of multivariate analyses provided by the contractor to the SPW.

External Project – CIM (2020-2021)

In collaboration with the CIM (Centre d'Information sur les Médias), the SMCS is involved in a project consisting in implementing and validating an algorithm that predicts missing listening notebooks in the radio survey of the CIM.

External Project – ORES (2020-2021)

In collaboration with ORES, the SMCS is involved in implementing and validating an algorithm that estimates the distribution of the total amount of charge supported by a given transformer throughout a year.

Your projects? SMCS can be your partner!

As you can see from this report, we are ready to become partner of various types of projects. 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 2020-2021

Trainings

- Certificat universitaire en Junior Data Analyst - Formule modulaire (formation maintenue, partiellement à distance), du 10 janvier 2020 au 15 décembre 2020 (Aurélié Bertrand, Vincent Bremhorst, Jonathan Dedonder, Alain Guillet, Séverine Guisset, Christian Ritter)
- Introduction to data analysis with R - Online training using Teams, du 4 septembre 2020 au 11 septembre 2020 (Alain Guillet)
- Pratique de la statistique avec SPSS, du 8 septembre 2020 au 11 septembre 2020 (Nathalie Lefèvre)
- Utilisation avancée du logiciel Nvivo, du 10 septembre 2020 au 11 septembre 2020 (Jonathan Dedonder)
- Initiation à la statistique avec applications en JMP (Formation intégrée dans un programme intensif plus large du FOREM), du 17 septembre 2020 au 23 septembre 2020 (Lieven Desmet, Catherine Rasse)
- Cours d'initiation à SAS (Formation organisée pour les Mutualités Chrétiennes), du 24 septembre 2020 au 1 octobre 2020 (Catherine Rasse)
- Certificat universitaire en Junior Data Analyst - Formule intensive, du 5 octobre 2020 au 15 décembre 2020 (Benjamin Colling, Vincent Bremhorst, Jonathan Dedonder, Alain Guillet, Séverine Guisset, Christian Ritter)
- **NEW** - Certificat universitaire en Analyse de données appliquées en R&D - Niveau avancé - Formation à distance via Teams, du 8 octobre 2020 au 26 novembre 2020 (Céline Bugli, Lieven Desmet, Bernadette Govaerts, Alain Guillet, Catherine Rasse)
- Initiation au logiciel NVivo, du 13 octobre 2020 au 14 octobre 2020 (Jonathan Dedonder)
- Data management and analysis using Stata - Online training using Teams, du 19 octobre 2020 au 23 octobre 2020 (Alain Guillet, Adam Levai)
- Introduction aux modèles mixtes avec R - Formation à distance via Teams, du 2 novembre 2020 au 26 novembre 2020 (Alain Guillet)
- Pratique de la statistique avec SPSS - Formation à distance via Teams, du 6 novembre 2020 au 27 novembre 2020 (Nathalie Lefèvre)
- **NEW** - Data visualisation and advanced tools in R - Online training using Teams, du 12 novembre 2020 au 13 novembre 2020 (Vincent Bremhorst, Alain Guillet)
- Analyse de données Omics sur base de techniques multivariées avec R (Formation organisée à distance via Teams pour Chemium sprl), du 17 novembre 2020 au 2 décembre 2020 (Manon Martin)
- Plans d'expérience avec JMP (Formation organisée via Teams pour Novasep), du 20 novembre 2020 au 15 décembre 2020 (Catherine Rasse)
- Préparation et mise en place d'une enquête - Formation à distance via Teams, du 26 novembre 2020 au 27 novembre 2020 (Jonathan Dedonder, Séverine Guisset)
- **NEW** - Notions de base en probabilité et inférence statistique, du 30 novembre 2020 au 11 décembre 2020 (Benjamin Colling, Emmanuel Niyigena)
- **NEW** - Vérification de l'invariance de mesure pour l'analyse de données d'enquête multigroupe - Formation à distance via Teams, du 1 décembre 2020 au 3 décembre 2020 (Aurélié Bertrand)
- Analyse de données pré-cliniques et cliniques avec SPSS - Formation à distance via Teams, du 4 décembre 2020 au 11 décembre 2020 (Nathalie Lefèvre)
- Introduction à la réalisation d'enquêtes en ligne avec LimeSurvey - Formation à distance via Teams, du 14 décembre 2020 au 15 décembre 2020 (Jonathan Dedonder)

Appendix

- Pratique de la statistique avec SPSS - Formation à distance via Teams, du 4 janvier 2021 au 8 janvier 2021 (Nathalie Lefèvre)
- Introduction à l'analyse de données avec R - Formation à distance via Teams, du 8 janvier 2021 au 15 janvier 2021 (Vincent Bremhorst)
- Experimental design with JMP - Formation à distance via Teams, du 11 janvier 2021 au 14 janvier 2021 (Rebecca Marion, Catherine Rasse)
- Introduction aux modèles multi-niveaux avec SPSS - Formation à distance via Teams, du 18 janvier 2021 au 22 janvier 2021 (Aurélie Bertrand, Nathalie Lefèvre)
- Analyses multivariées exploratoires en R - Formation à distance via Teams, du 25 janvier 2021 au 26 janvier 2021 (Séverine Guisset)
- Modèles multi-niveaux avancés avec SPSS - Formation à distance via Teams, du 25 janvier 2021 au 29 janvier 2021 (Aurélie Bertrand)
- Initiation au logiciel NVivo - Formation à distance via Teams, du 28 janvier 2021 au 29 janvier 2021 (Jonathan Dedonder)
- Analyses multivariées avec SPSS - Formation à distance via Teams, du 8 février 2021 au 9 février 2021 (Séverine Guisset)
- Certificat d'université : Analyse de données appliquée en R&D - niveau de base - Formation à distance via Teams, du 8 février 2021 au 29 avril 2021 (Aurélie Bertrand, Céline Bugli, Lieven Desmet, Bernadette Govaerts, Alain Guillet, Catherine Rasse)
- Introduction to data analysis with R - Online training via Teams, du 9 février 2021 au 18 février 2021 (Lieven Desmet)
- Analyser la médiation, la modération et la médiation modérée avec SPSS, du 23 février 2021 au 24 février 2021 (Aurélie Bertrand)
- Concepts de base et pratique de la statistique bayésienne avec R et JASP - Formation à distance via Teams, du 1 mars 2021 au 5 mars 2021 (Vincent Bremhorst)
- **NEW** - Comment bien présenter et interpréter des données (Formation organisée en collaboration avec la Société Royale Belge de Statistique), du 30 mars 2021 au 30 mars 2021 (Aurélie Bertrand, Vincent Bremhorst)
- **NEW** - Een inleiding tot correcte en duidelijke communicatie over data (Formation organisée en collaboration avec la Société Royale Belge de Statistique), du 1 avril 2021 au 1 avril 2021 (Lieven Desmet, Séverine Guisset)
- Utilisation avancée du logiciel Nvivo, du 22 avril 2021 au 23 avril 2021 (Jonathan Dedonder)
- **NEW** - Applied statistics with SPSS, du 3 mai 2021 au 7 mai 2021 (Séverine Guisset)
- Certificat universitaire en Analyse de données appliquées en R&D - Niveau avancé, du 4 mai 2021 au 29 juin 2021 (Céline Bugli, Lieven Desmet, Bernadette Govaerts, Alain Guillet, Catherine Rasse)
- Panel-data Analysis using Stata, du 6 mai 2021 au 7 mai 2021 (Alain Guillet)
- Introduction à Jamovi, manipulation et analyse de données (Formation sur mesure organisée pour l'Etnic), du 27 mai 2021 au 22 juin 2021 (Aurélie Bertrand)
- Introduction à Jamovi, manipulation et analyse de données (Formation sur mesure organisée pour l'Etnic), du 2 juin 2021 au 22 juin 2021 (Aurélie Bertrand)
- Introduction à l'analyse de données avec R, du 4 juin 2021 au 11 juin 2021 (Antoine Soetewey)
- Introduction à l'analyse de Réseaux Sociaux avec Gephi, du 9 juin 2021 au 11 juin 2021 (Jonathan Dedonder)
- Introduction aux modèles mixtes avec R - Formation organisée via Teams, du 14 juin 2021 au 18 juin 2021 (Alain Guillet)
- Data visualisation and advanced tools in R, du 24 juin 2021 au 25 juin 2021 (Vincent Bremhorst, Alain Guillet)
- Introduction to Python, du 1 juillet 2021 au 2 juillet 2021 (Vincent Bremhorst)
- Cours préparatoire d'été en statistiques, du 16 août 2021 au 20 août 2021 (Benjamin Colling)

Appendix

- *Certificat universitaire en Junior Data Analyst - Formule modulaire (formation maintenue, partiellement à distance), du 22 janvier 2021 au 15 décembre 2021 (Benjamin Colling, Jonathan Dedonder, Alain Guillet, Christian Ritter, Séverine Guisset, Vincent Bremhorst, Formateurs externes) - Not counted in the statistics of this report because training is still ongoing in 2021-2022*

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

Garcés-Ruiz, M., Calonne-Salmon, M., Bremhorst, V., & Declerck, S. (2021). Diesel fuel differentially affects hyphal healing in *Gigaspora* sp. and *Rhizoglyphus irregularis*. In: *Mycorrhiza*, 31, 413-421.

Lobet, G., Descamps C., Leveau, L., Guillet, A., Rees, J.-F. (2020). *QuoVidi: An open-source web application for the organization of large-scale biological treasure hunts*. *Ecology and Evolution* Volume 11, Issue 8 p. 3516-3526. <https://doi.org/10.1002/ece3.7130>

Ngo Njembe, M.T., Pachikian, B., Lobysheva, I., Van Overstraeten, N., Dejonghe, L., Verstraelen, E., Buchet, M., Rasse, C., Gardin, C., Mignolet, E., Balligand, J.L., & Larondelle, Y. (2021). *A Three-Month Consumption of Eggs Enriched with ω -3, ω -5 and ω -7 Polyunsaturated Fatty Acids Significantly Decreases the Waist Circumference of Subjects at Risk of Developing Metabolic Syndrome: A Double-Blind Randomized Controlled Trial*. By: *Nutrients*, 13(2), 663; <https://doi.org/10.3390/nu13020663> - 18 Feb 2021.

Remacle, A., & Lefèvre, N. (2021). *Which teachers are most at risk for voice disorders? Individual factors predicting vocal acoustic parameters monitored in situ during a workweek*. In: *International Archives of Occupational and Environmental Health*, 94, 1271-1285; <https://doi.org/10.1007/s00420-021-01681-3>.

Rodriguez-Morelos, V.H., Calonne, M., Bremhorst, V., Garcés-Ruiz, M., & Declerck, S. (2021). *Fungicides with contrasting mode of action differentially affects hyphae healing mechanism in *Gigaspora* sp. and *Rhizoglyphus irregularis**. *Frontiers in Plant Science*, Vol. 12; <https://www.frontiersin.org/article/10.3389/fpls.2021.642094>.

Seminars and Presentations

Extra-muros seminars and video

Bugli C., *Les statistiques dans votre mémoire: il vaut mieux y penser avant pour éviter les ennuis après...* Présentation aux mémorants FSM, Louvain-la-Neuve, les 26 octobre et 9 février 2021.



2020-2021