REPORT OF ACTIVITIES

Statistical Methodology and Computing Service

1 September 2021 to 31 August 2022











Statistical Methodology and Computing Service Report of Activities 1 September 2021 to 31 August 2022

Contents

Introduction	4
People and Organization	5
Management and Administration	
Management committee	5
User committee	5
People	5
Associated Consultants	7
Professors	7
Consultants at other institutes	7
Researchers and Research Assistants	7
Interns and jobists	7
Trainers and Consultants	7
Sponsoring	8
Training	9
Consulting	12
Servers and Statistical Software	14
Statistical Computing Server	
UDS	14
Available Software	14
Usage Statistics	15
Larger Projects	16
Research Projects	16
Erasmus+OLS (2015-2022)	16
CluePoints (2018-2022)	16
First spin-off (2018-2021)	16
PDR (2020-2024)	16
Survey among book authors in Fédération Wallonie-Bruxelles (2018-2022)	16
Gender and housing in Brussels (2022-2023)	16
Survey Support	16
Election processes (2020-2022)	
UCLouvain Mobility Survey (2020-2021)	
Working as a trustee survey (2021-2022)	
Consulting	
Data management and statistical support - Strategic decisions of the general services (2019)	
Data management support - Web scraping (2021-2022)	1/

Contents

Support to the CICN platform (2016)	17
Support to the MiiL platform (2021-2022)	17
Development of teaching materials	17
Evaluation of medical internship sites (2021-2022)	17
Training	18
University Certificate - Analyse de données en R&D (2020)	18
University Certificate - Junior Data Analyst (2016)	18
Summer School - Collaboration with the CENTAL (2019)	18
Collaboration with the CIM (2020)	18
Collaboration with ORES (2021-2022)	18
Your projects? SMCS can be your partner!	18
Appendix	19
List of Trainings and Seminars 2021-2022	
Trainings	19
Publications and Presentations by SMCS members	21
Publications	21
Seminars and Presentations	21

Introduction

Introduction

A year on the way back to normality after the pandemic turbulences. We learned a lot about collaborating remotely and will use this in the future. Will we ever go back to look at some of the video recordings of our online and hybrid courses? It would certainly be interesting to do so in a few years, but would it be useful? It might at least be worth to find a (hobby) professional who could look at some of them and tell us if there is value or not.

The SMCS continued 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 2021/22.

Consultancy has grown to about 600 projects per year. Training reached its highest volume ever with 51 courses representing 193 training days. 655 participants attended for a total of 2349 person days. The statistical computing server had 62 registered users.

We further improved our visibility towards potential new internal users (researchers, administrators, staff at the university hospitals). Our external visibility benefits from the new website and newsletter dedicated to technical and service platforms at UCLouvain. Moreover, we currently have more than 700 followers on LinkedIn. Are you one already? If not, check it out.

The SMCS has continued assisting the UCLouvain administration. This includes work on coordinating the online and face-to-face exams to respect the evolving COVID safety requirement for the sessions in September 2021 and January 2022. We also support the data extraction, management, and visualization 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. Furthermore, we assisted the central services with their latest survey about the mobility of UCLouvain staff and students. This survey gathers data to fulfill official legal requirements, and its analysis will increase understanding of mobility patterns.

Four university certificate courses were organized during this academic year, a modular and an intensive version of the Junior Data Analyst, and two versions, basic and advanced, of a program named Data Analysis in Research and Development. The open text comments from the participant surveys associated with these programs emphasize the excellent quality and professionalism of the trainings. We will continue maintaining close contacts with the alumni of these programs to follow relevant market trends in practical statistics and data analytics.

We also organized a summer school on Text Mining and Social Network Analysis in collaboration with the Center for Natural Language Processing (CENTAL).

One of the challenges for the next few years is to find and maintain a good equilibrium between practical work in consultancy, delivering training in statistical practice, and keeping close contact with research in statistics and neighboring disciplines.

I would like to congratulate the team for their professionalism and dedication.

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

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



People and Organization

MANAGEMENT AND **A**DMINISTRATION

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): Catherine Legrand SSS representative (IRSS institute): Niko Speybroeck SST representative (ELI institute): Marnik Vanclooster Representative of research activities (ISBA): Michel Denuit

Representative of training activities (LSBA): Anouar El Ghouch

Associate consultant academic representative:

Bernadette Govaerts

Consultant representative: Aurélie Bertrand

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

In 21-22, 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 Masquelier (IACCHOS), Magali Paquot (IL&C), Gordy Pleyers (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), Sergio Altomonte (LAB), Isabelle Donnay (LIBST), Vincent Vandenberghe (LIDAM), Nathan Lassance (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. As the workload was still very important and the work of B. Colling of high quality, it has been decided to propose him a CDI.

Manon Martin has been hired as an SST research logistician in replacement of Gipsi Lima Mendez in order

to provide precious support in bioinformatics to the SST community (specifically for the LIBST and ELI needs) by helping the researchers to process and analyse their high throughput data and organise trainings that best matches their needs. She is specifically dedicated to the SMCS up to 20% of her time.

We could also count on the support of Sophie Mathieu, former PhD student. She built the materials for a Python training and worked on the realization of Dashboard.

Finally, Guillaume Dallemagne decided to leave the university. He was of a great help in the field of econometrics. His position is vacant.

Consultants



Aurélie Bertrand (UCLouvain/LIDAM/SMCS – 0.975 FTE)

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).

Manon Martin (UCLouvain/LIDAM/SMCS – 0.2 FTE, UCLouvain/LIBST – 0.5 FTE and UCLouvain/ELI – 0.3 FTE)

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

Doctorat en Sciences – Statistique (UCLouvain, 2020) Master en Statistique (UCL, 2014)

Master en Bioingénieur, orientation économie agricole (UCL, 2012)

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), 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 Schoumaker (IACCHOS), Nicolas Schtickzelle (ELI), Sébastien Van Bellegem (LIDAM), Vincent Vandenberghe (LIDAM/IRES).

Consultants at other institutes

Jonathan Dedonder (SSH/IACCHOS), Guillaume Dallemagne (SSH/LIDAM/IRES), Nathalie Lefèvre (SSH/IPSY), Manon Martin (formerly SSS/DDUV – now SST/LIBST, ELI), Jérôme Ambroise (SSS/IREC), Aline van Maanen (Institut Roi Albert II - Cancérologie et Hématologie), Axelle Loriot (SSS/DDUV).

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 AND JOBISTS

During her internship (mid-September to end of August 2022), Laura Wautier created a pedagogical Shiny application to perform logistic regressions; this app will be used in the SMCS training.

Two jobists, Armelle Liénard and Margaux Latour, were hired for 3 months to conduct face-to-face interviews and carry out the qualitative analysis of the interviews and open-ended comments on the surveys conducted as part of the study on the syndic's profession.

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.

Aurélie Bertrand (SMCS), Vincent Bremhorst (SMCS), Céline Bugli (SMCS), Benjamin Colling (SMCS), Guillaume Dallemagne (IRES), Damien De Meyere (CENTAL), Jonathan Dedonder (IACCHOS), Benjamin Deketelaere (ISBA), Morine Delhelle (ISBA), Lieven Desmet (SMCS), Hortense Doms (ISBA), Thomas François (ILC), Laurent Gatto (DDUV), Bernadette Govaerts (ISBA), Alain Guillet (SMCS), Séverine Guisset (SMCS), Fanny Hoogstoel (ISBA), Alexandre Jacquemain (ISBA), Chikeola Ladekpo (ISBA), Eric Lecoutre (WeLoveDataScience!), Nathalie Lefèvre (SMCS), Adam Levai (IRES), Stéphane Lhaut (ISBA), Vincent Lorant (IRSS), Axelle Loriot (DDUV), Manon Martin (formerly DDUV - now LIBST, ELI, SMCS), Sophie Mathieu (ISBA), Hubert Naets (CENTAL), Charles Guy Njike Leunga (ISBA), Eugen Pircalabelu (ISBA), Catherine Rasse (SMCS), Christian Ritter (LSBA), Christophe Robyns Antoine Soetewey (ISBA), Patrick Watrin (Agilytic), (CENTAL), external experts.

SPONSORING

We are grateful to Agilytic, Brainstorming, Cream Consulting and Micropole for their support in the organization of our Junior Data Analyst certificate; and to Business and Decision for their financial support in the organization of our Text mining summer school in Python.

Training

Training

Training is one of our main missions. In general, we focus on hands-on training in practical statistics and the use of statistical software in ways that are not covered by regular university courses. Five of the training sessions took place on the Woluwe campus. Sixteen training courses have been given in English. It is more than the previous year during which 9 training courses were proposed 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.

Three training sessions (5.5 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. Last year was an exceptional year in terms of the number of courses organized due to the postponement of courses cancelled at the beginning of the COVID. Despite this, we have even higher statistics this year in terms of number of training courses and days.

Thirteen of the fifth training courses were still organized at a distance. The experience of distance learning has led us to think about keeping some courses in distance learning in the future, even if the vast majority will remain presential.

Nine new training programs have been developed including 1 new subject for our summer school organized in collaboration with the Cental (Centre de traitement automatique du language, UCLouvain).

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, "Generalized Linear Mixed Models using R".

- Customized trainings for specific audiences (UCLouvain and external). Examples: "Modèles multiniveaux 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, 4 certificates were organized: "Certificat universitaire en Analyse de données appliquées en R&D Niveau de base" + "Niveau avancé" and "Certificat universitaire en Junior Data Analyst Formule intensive" + "Formule modulaire".
- 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 2020 edition, on a slightly different theme, was finally organized in July 2022 because of COVID. The specificity of the program was to propose, besides courses, some social activities which have been greatly appreciated by the participants.
- 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 51 courses in the academic year 2021-2022 for 193 teaching days. The increase observed last year mainly due to a post-covid rebound (with postponed trainings) wasn't followed by a decrease, participation remained very high this year in all 3 sectors.

Training

Table 1. Global statistics on SMCS trainings

	19/20	20/21	21/22
Courses	39	46	51
Course Days	130	189	193
Participants UCLouvain	402*	524	583
Person-Days* UCLouvain	1435	2106.5	1983
Participants Ext	120	138	72
Person-Days* Ext	262	287.5	366

^{*}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

	19/20	20/21	21/22
SSH	177*	196	173
SSS	82	134	160
SST	71	81	124
Students	71*	110	118
Central services	1	3	8

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 SSS and SST sectors.

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

Table 3. Origin of participants from the Human Sciences Sector (SSH)

	19/20	20/21	21/22
SSH	203*	232	226
IACS-SPLE-ESPO	60*	60	42
ILC	23*	27	32
LOURIM-LSM	3	23	18
LIDAM	19*	29	28
INCAL-FIAL		8	3
IPSY-PSP	65*	47	45
JURI-DRT	1		3
RSCS-TECO	6	2	2
Students	26	36	53

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)

	19/20	20/21	21/22
SSS	94	141	180
ASS	1		
DDUV	16	3	23
IONS	15	56	45
IREC-MEDE-FSM	29	40	48
IRSS-FSP	9	15	19
LDRI-FASB	10	18	23
Clinics	2	2	2
Students	12	7	20

In the science and technology sector courses were also followed by participants from a wide range of institutes. Nevertheless, more than 50% of the SST staff participating in SMCS training comes from the ELI Institute.

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

Training

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

	19/20	20/21	21/22
SST	104*	148	169
AGRO	2	3	2
ELI	47	49	64
ICTEAM-EPL		3	3
IMCN	1		2
IMMC	9	5	15
IRMP		1	
LIBST	5	10	9
LIDAM	3	2	1
LOCI-LAB	3	8	28
SC	1		
Students	33*	67	45

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). 3.3 FTE of this effort is directly oriented towards the UCLouvain research and teaching units, the remainder towards external clients, such as the "FOREM" and some private companies and public administrations.

The trainings given to external clients fulfill three objectives. First, they are a part of our mission of service to society. Second, they allow us to develop new course modules for internal use, such as the FOREM training which allowed us to bring together an entire course portfolio in the area of bioinformatics 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 such trainings to make sure there is a positive return (financial and others).

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

^{*}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

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 2 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 coauthor of publications based on the work. This type of support was still particularly important this year and corresponds to more than 5341 hours; this represents 20% more than the previous year and nearly 90% of these hours were spent on internal projects (604 hours for external projects and 2750 for internal one's). Far fewer hours were spent this year on external projects (1707 hours last year vs. 604 this year), which is explained by the fact that internal projects always have priority over external projects. SMCS consultants do not seek to develop external collaborations when they are involved in internal projects. However, some internal projects are still funded by external partners, such as a research project with CluePoint (see the section on larger projects).

Overall, the consulting activity of the SMCS occupies about 4.9 full time staff equivalents (FTE). 4.3 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	19/20	20/21	21/22
Research	450	495	538
Educational project	14	5	8
Admin. services	23	21	22
External clients	18	33	28
Total	505	554	596

From 1 September 2021 to 31 August 2022, the SMCS worked on 596 projects of which 568 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 "Larger projects".

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. Even if each request counts for one unit in the table, some of them can represent up to 450 hours of work.

Table 7 (SSH) shows that the number of requests from the Human Sciences Sector fluctuates around 190 requests per year for the past 3 years (1463 hours of work last year and now 1893 hours). 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 until June 30, 2022), 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)

	19/20	20/21	21/22
SSH	187	182	193
Cross-disciplinary	4	3	1
IACCHOS-ISPOLE-ESPO	26	33	22
ILC-INCAL-FIAL-ISP	20	20	37
IPSY-PSP	105	96	100
JURI-DRT	5	4	2
LOURIM-LSM	7	11	10
LIDAM	20	15	20
RSCS-TECO			1

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)

	19/20	20/21	21/22
SSS	198	206	239
ASS-Cross-disciplinar	1	1	1
DDUV	6	2	6
FASB	7	3	4
FSM	20	13	15
IONS	24	49	56
IREC	52	54	52
IRSS-FSP	30	28	40
LDRI	13	14	14
MEDE	31	37	42
Clinics	14	5	9

Table 9 (SST) shows that the number of requests from the science and technology sector fluctuates around 90-120 requests per year (627 hours of work). We can notice those figures are stable. The number of hours worked and the number of requests in ELI and LIBST account for about 65% of the SST support, other requests came from almost all units of this sector.

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

	19/20	20/21	21/22
SST	92	121	121
Cross-disciplinary	2	4	1
AGRO	7	8	14
ELI	36	53	51
ICTEAM-EPL	3	1	2
IMCN			
IMMC	6	19	14
LIBST	29	25	26
LIDAM	1	3	1
LOCI-LAB	2	3	4
SC	6	5	8

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, **15 projects** also originated from the central administration (AC, AS, SGSI, BIUL). The last three years, support to central administration and administrative services in general was particularly high, with a total of 1017 hours of work this year. 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). 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 62 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 multiprocessors version). The average CPU charge was 15% on smcs1 and 10% on smcs2 with spikes (600% 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 Lemaitre 3.

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

service-desk@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. The most asked software was SPSS far before Stata which also is far from JMP and R (but everybody can install R for free from internet even without administrator rights). The least demanded software is SAS even counting SAS and SAS Enterprise Guide together.

High performance solutions in SAS, R and Stata (multiprocessor) 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 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

	19/20	20/21	21/22
Total	57	69	62
SSH	35	37	42
IACS-SPLE-ESPO	15	17	22
ILC		1	1
LIDAM	20	19	19
SSS	10	14	11
IREC-MEDE	2	3	4
IRSS-FSP	8	11	6
LDRI			1
SST	12	18	9
ELI	3	6	5
ICTEAM-EPL	2	4	2
IMCN	2		1
IMMC	4	4	
IRMP	1	2	
LIBST		1	
SC		1	1

Usage of UDS

The following table lists the number of researchers who requested access to UDS to use software without having

it on their computer. Unfortunately, the SGSI was unable to provide usage data for the 2021-2022 academic year.

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

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

Larger Projects

Larger Projects

This section summarizes the largest projects in which the SMCS was involved in 2021-2022. 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 PROJECTS

Erasmus+OLS (2015-2022)

The European project Erasmus+OLS was 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 was to provide statistical expertise, in particular to monitor the assessment, and design and analyze surveys. We used text mining methods to analyze open survey answers and provided regular reports to the project team. This project ended in its current form in the summer of 2022.

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.

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 was responsible for programming an algorithm for quantifying a new biomarker.

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.

Survey among book authors in Fédération Wallonie-Bruxelles (2018-2022)

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.

Gender and housing in Brussels (2022-2023)

This project consists in the mapping, gathering and processing of data on gender and housing in Brussels, on behalf of the SPRB. The final report will assess 6 dimensions of the topic. SMCS is in charge of the coordination of the project. Prs Salembier (LAB), Wagener (IACCHOS) and Bernard (USaint-Louis) bring their expertise to the project, while post-doc Benjamin Leclercq (LAB) coordinates the qualitative part of the project.

SURVEY SUPPORT

Survey support consulting was provided for a wide range of applications in all three sectors in 2020-2021 and included advice on survey and questionnaire design, programming in LimeSurvey, sampling, data management, weighting and analysis.

Support is currently provided to two surveys about sport and cultural activities offered at UCLouvain, as well as urbanistic developments in Louvain-La-Neuve. Methodological support was also provided to an external project on mobility in Wallonia.

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

Election processes (2020-2022)

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

Larger Projects

these. This process was implemented 6 times with success, including during the election of the SSS Vice-Rector.

UCLouvain Mobility Survey (2020-2021)

The university launched a survey on mobility among staff and students in October 2021. The objectives were 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, programming, pilots, data collection, data treatment, analysis and produced various reporting outputs for the survey.

Working as a trustee survey (2021-2022)

SMCS coordinated a survey on behalf of the Professional Institute of Estate Agents (IPI-BIV) in collaboration with Professors Casini (IPSY) and Raemdonck (IACCHOS). The aim is to better understand the difficulties faced by trustees (co-ownership managers) in their everyday work and identify solutions for the future of their profession. SMCS took care of the methodological and technical side of the project (questionnaire design, programming, data collection and processing), as well as statistical analyses and report writing. Prs Casini and Raemdonck brought their subject-matter expertise to the questionnaire design, analysis and report, with a particular focus on recommendations for the future.

CONSULTING

Data management and statistical support - Strategic decisions of the general services (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. Since January 2022, the collaboration with SASP is extended to a statistical support for the strategic decisions of the general services.

Data management support - Web scraping (2021-2022)

The SMCS was involved in a Proof-Of-Concept (POC) project founded by Innoviris. The SMCS developed two Python programs. The first one aims to automate the web scraping from different URL sources and to create a unique database with all the scraped information. The purpose of the second is, from a Scopus extraction, to prepare a database to achieve a citation networking analysis (CNA).

Support to the CICN platform (2016- ...)

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.

Support to the MiiL platform (2021-2022)

In collaboration with the MiiL platform, the SMCS is involved in the implementation of a weighted algorithm to make the Opinio panel of the RTBF as representative as possible of the Walloon population on the basis of age, gender and region.

Development of teaching materials

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

Evaluation of medical internship sites (2021-2022)

The internship administration of the medical faculty of UCLouvain organizes an evaluation of the different internship sites by the medical students, looking not only at pedagogical aspects but also at practical aspects of the different internship sites in order to assess their continued quality. The SMCS assists the faculty in producing automated reports per internship supervisor.

Larger Projects

TRAINING

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 course is since January 2020 valued in 2 university certificates. The third edition of the 2 certificates took place in 2021-2022 with as much success as ever.

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

In collaboration with the LSBA, the SMCS launched in 2016 a new certificate called "Junior Data Analyst", an intensive course for graduates in humanities in practical data management, analysis, and interpretation. The course was given for the sixth edition from October 2021 to December 2021 and is self-financing since the third edition. In January 2020, the first edition of a modular JDA was launched. This modular program is more compatible with a professional activity because half of the courses are organized on Saturdays. Its third edition started in January 2022. A fourth edition will start in January 2023. The modular program is self-financed since its first edition.

Summer School - Collaboration with the CENTAL (2019-...)

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 (with a focus on the analysis of financial documents), 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), finally took place (after a cancellation in 2020 and postponements in 2021 and early 2022 due to the COVID pandemia) in July 2022.

Collaboration with the CIM (2020-...)

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.

Collaboration with ORES (2021-2022)

In collaboration with ORES, the SMCS is involved in implementing two simulators providing the photovoltaic production for a selected place and period in Belgium, the first one giving the production in an optimal uncovered sky case, the second one taking into account the typical meteorological variations in cloud cover and temperature during a day.

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 2021-2022

Trainings

- Certificat universitaire en Junior Data Analyst Formule modulaire, du 22 janvier 2021 au 17 décembre 2021 (Vincent Bremhorst et Benjamin Colling et Jonathan Dedonder et Formateurs externes et Alain Guillet et Séverine Guisset et Nathalie Lefèvre et Sophie Mathieu et Christian Ritter et Antoine Soetewey)
- Introduction to data analysis with R, du 2 septembre 2021 au 10 septembre 2021 (Lieven Desmet)
- Pratique de la statistique avec SPSS, du 6 septembre 2021 au 10 septembre 2021 (Nathalie Lefèvre)
- Introduction au calcul de taille d'échantillon, le 24 septembre 2021 (Benjamin Colling et Catherine Rasse)
- NEW Data manipulation and visualisation with R, du 28 septembre 2021 au 30 septembre 2021 (Laurent Gatto et Axelle Loriot et Manon Martin)
- Initiation à la statistique avec applications en JMP (Formation intégrée dans un programme intensif plus large du FOREM), du 30 septembre 2021 au 5 octobre 2021 (Lieven Desmet et Catherine Rasse)
- Certificat universitaire en Junior Data Analyst Formule intensive, du 4 octobre 2021 au 17 décembre 2021
 (Vincent Bremhorst et Benjamin Colling et Jonathan Dedonder et Formateurs externes et Alain Guillet et Séverine Guisset et Nathalie Lefèvre et Sophie Mathieu et Christian Ritter et Antoine Soetewey)
- Initiation au logiciel NVivo, du 11 octobre 2021 au 14 octobre 2021 (Jonathan Dedonder)
- Introduction to GraphPad, du 15 octobre 2021 au 22 octobre 2021 (Lieven Desmet)
- Data management and analysis using Stata, du 18 octobre 2021 au 22 octobre 2021 (Guillaume Dallemagne et Alain Guillet)
- Introduction to JMP software, du 25 octobre 2021 au 26 octobre 2021 (Catherine Rasse)
- NEW Visualisation des données et outils avancés de programmation en R, du 8 novembre 2021 au 9 novembre 2021 (Benjamin Colling)
- Pratique de la statistique avec SPSS, du 12 novembre 2021 au 3 décembre 2021 (Nathalie Lefèvre)
- Applied statistics with JMP, du 15 novembre 2021 au 22 novembre 2021 (Catherine Rasse)
- NEW Introduction à la visualisation et à l'analyse de données en Python, du 18 novembre 2021 au 26 novembre 2021 (Vincent Bremhorst et Sophie Mathieu)
- Write and publish your work in an R-package, le 26 novembre 2021 (Vincent Bremhorst et Alexandre Jacquemin)
- Analyser la médiation, la modération et la médiation modérée avec SPSS, du 30 novembre 2021 au 3 décembre 2021 (Aurélie Bertrand)
- Notions de base en probabilité et inférence statistique, du 30 novembre 2021 au 10 décembre 2021 (Vincent Bremhorst et Morine Delhelle)
- NEW Introduction to mixed-effects models using R, du 6 décembre 2021 au 10 décembre 2021 (Lieven Desmet)
- Cours d'initiation à SAS, du 9 décembre 2021 au 14 décembre 2021 (Catherine Rasse)
- Introduction to data analysis with R, du 9 décembre 2021 au 17 décembre 2021 (Alain Guillet)
- Analyse de données pré-cliniques et cliniques avec SPSS, du 10 décembre 2021 au 17 décembre 2021 (Nathalie Lefèvre)
- Pratique de la statistique avec SPSS, du 10 janvier 2022 au 14 janvier 2022 (Nathalie Lefèvre)
- Introduction à l'analyse de données avec R, du 13 janvier 2022 au 21 janvier 2022 (Antoine Soetewey)
- Introduction to experimental design with JMP, du 17 janvier 2022 au 25 janvier 2022 (Catherine Rasse)

Appendix

- Initiation au langage Python A distance via Teams, du 20 janvier 2022 au 21 janvier 2022 (Vincent Bremhorst et Sophie Mathieu)
- Préparation et mise en place d'une enquête, du 24 janvier 2022 au 25 janvier 2022 (Jonathan Dedonder et Séverine Guisset)
- Certificat d'université : Analyse de données appliquée en R&D niveau de base, du 31 janvier 2022 au 28 avril 2022 (Aurélie Bertrand et Céline Bugli et Lieven Desmet et Alain Guillet et Catherine Rasse)
- Introduction à la réalisation d'enquêtes en ligne avec LimeSurvey, du 2 février 2022 au 3 février 2022 (Jonathan Dedonder)
- Analyses multivariées avec SPSS, du 7 février 2022 au 8 février 2022 (Séverine Guisset)
- Introduction to data analysis with R, du 9 février 2022 au 18 février 2022 (Lieven Desmet)
- NEW Advanced Stata: Data Handling and Programming Tools to Make Your Research Life Easier Online via Teams, du 16 février 2022 au 18 février 2022 (Vincent Vandenberghe)
- Initiation aux modèles d'équations structurelles (SEM) avec Stata, du 14 mars 2022 au 17 mars 2022 (Aurélie Bertrand)
- Initiation au logiciel NVivo, du 23 mars 2022 au 25 mars 2022 (Jonathan Dedonder)
- Analyses multivariées exploratoires en R, du 28 mars 2022 au 29 mars 2022 (Benjamin Colling)
- Introduction aux modèles multi-niveaux avec SPSS, du 19 avril 2022 au 22 avril 2022 (Aurélie Bertrand et Nathalie Lefèvre)
- Utilisation avancée du logiciel Nvivo, du 19 avril 2022 au 22 avril 2022 (Jonathan Dedonder)
- Applied statistics with SPSS, du 2 mai 2022 au 6 mai 2022 (Séverine Guisset)
- Certificat universitaire en Analyse de données appliquées en R&D Niveau avancé, du 2 mai 2022 au 28 juin 2022 (Céline Bugli et Lieven Desmet et Bernadette Govaerts et Alain Guillet et Catherine Rasse)
- Panel-data Analysis using Stata, du 5 mai 2022 au 6 mai 2022 (Guillaume Dallemagne et Alain Guillet)
- NEW Modèles d'équations structurelles (SEM) avancés avec Stata, du 5 mai 2022 au 24 mai 2022 (Aurélie Bertrand)
- Modèles multi-niveaux avancés avec SPSS, du 9 mai 2022 au 13 mai 2022 (Aurélie Bertrand et Nathalie Lefèvre)
- NEW Gestion et visualisation de données en Excel, du 30 mai 2022 au 31 mai 2022 (Séverine Guisset et Nathalie Lefèvre)
- Introduction à l'analyse de Réseaux Sociaux avec Gephi, du 7 juin 2022 au 10 juin 2022 (Jonathan Dedonder)
- Introduction à l'analyse de données avec R, du 9 juin 2022 au 17 juin 2022 (Antoine Soetewey)
- Introduction aux modèles mixtes avec R, du 13 juin 2022 au 17 juin 2022 (Alain Guillet)
- NEW Generalized Linear Mixed Models using R, le 21 juin 2022 (Alain Guillet)
- Data visualisation and advanced tools in R, du 23 juin 2022 au 24 juin 2022 (Alain Guillet)
- Initiation au langage Python, du 7 juillet 2022 au 8 juillet 2022 (Vincent Bremhorst)
- Ecole d'été 2022 sur le Text Mining et l'analyse des réseaux sociaux en Python, du 11 juillet 2022 au 15 juillet 2022 (Damien De meyere et Thomas François et Hubert Naets et Patrick Watrin)
- Cours préparatoire d'été en statistiques, du 22 août 2022 au 26 août 2022 (Benjamin Colling et Charles Guy Njike leunga)
- 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

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

Amiodarone-induced thyroid dysfunction in children: insights from the THYRAMIO study - Montenez, Sarah[UCL]; Moniotte, Stéphane[UCL]; Robert, Annie[UCL]; Desmet, Lieven[UCL]; Lysy, Philippe[UCL](2021) Therapeutic Advances in Endocrinology and Metabolism — Vol. 12, p. 1-9 (2021)

Long-term follow-up of boys who have undergone a testicular biopsy for fertility preservation - Kanbar, Marc[UCL]; De Michele, Francesca[UCL]; Giudice, Maria Grazia[UCL]; Desmet, Lieven[UCL]; Poels, Jonathan[UCL]; Wyns, Christine[UCL](2021) Human Reproduction — Vol. 36, no. 1, p. 26-39 (2021)

Tumor apelin and obesity are associated with reduced neoadjuvant chemotherapy response in a cohort of breast cancer patients - Gourgue, Florian[UCL]; Derouane, Françoise; van Marcke, Cedric; Villar, Elodie; Dano, Helene; **Desmet, Lieven**; Bouzin, Caroline; Duhoux, Francois P.; Cani, Patrice D.[UCL]; Jordan, Bénédicte[UCL] (2021) Scientific Reports — Vol. 11, no.1, p. 9922 (2021)

KATP channel blockers control glucagon secretion by distinct mechanisms: A direct stimulation of α -cells involving a [Ca2+]c rise and an indirect inhibition mediated by somatostatin - Singh, Bilal[UCL]; Khattab, Firas[UCL]; Chae, Heeyoung[UCL]; **Desmet, Lieven**[UCL]; Herrera, Pedro L; Gilon, Patrick[UCL](2021) Molecular metabolism — Vol. 53, p. 101268 (2021)

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 25 octobre 2021.



