

A quantitative questionnaire for SAP-based data analytics in education

F. Eckhard, K. Haug & H. Wittges

Technical University of Munich, Munich, Germany

Abstract: Data analytics is becoming increasingly important in the business world. As a business software leader, SAP is also an essential player in the analytical ecosystem. With the rising interest in big data and machine learning, an increasing number of experts are required to develop analytical systems. This paper presents the results of the questionnaire about the current state and outlook for SAP-based data analytics in education. 40 lecturers from universities worldwide participate in an online survey with three main research questions. What are current and future data analytics topics in education? Which opportunities and challenges exist using SAP cloud services? Which data analytics offerings are used or planned to be used in education? In conclusion, the paper proposes topics for future research and tasks for curriculum development.

Keywords: Education, SAP, Data Analytics, Cloud

1. Introduction

Data analytics is becoming increasingly critical in the business world. (Whitelock, 2018) (Chen, Chiang, & Storey, 2012) With the development of big data, machine learning and, lately, generative AI, experts are increasingly required to develop, maintain, and evaluate analytical systems. Since the demand for data analytics expertise on the job market will grow even bigger within the next few years, universities and academic institutions should incorporate data analytics topics into their curricula. (Baird & Parayitam, 2019) SAP, as one of the leaders in business software, is an essential player in the analytical ecosystem, as a lot of analytically relevant data is already available in the ERP system. (Torii, et al., 2023) Therefore, introducing SAP software to analytical lectures can be an advantage for students building competences which are demanded in the business world. In this paper, we present our survey about the current state of SAP software in analytical curricula and give an outlook on upcoming topics in this area. For the purpose of this study, data analytics refers collectively to traditional data analysis, machine learning, and artificial intelligence.

Specifically, the following three research questions are defined as the objective of the survey:

- *RQ1*: Which business and data analytics topics are currently and perspectivevely relevant for teaching?
- *RQ2*: Which opportunities and challenges exist in using SAP cloud systems for teaching?
- *RQ3*: Which data analytics tools (SAP and non-SAP) are currently being used and will be used in the future?

This paper is divided into four chapters. Based on the research questions in the second chapter, the methodology for the survey is introduced, specifically the set-up of the questionnaire, the target population, and the pretests. The analysis and interpretation of the questionnaire are presented in the third chapter. Finally the results are summarised in the last chapter and an outlook for further research is given.

2. Research Methodology

This study is based on an online quantitative questionnaire with a range of 25 to 37 questions. The number of questions shown to a participant can vary based on the answers given to some questions. The questionnaire is divided into four question categories:

- 1) General information about the courses/ lecturers
- 2) Presently used analytical systems in education
- 3) Outlook and ideas for the future of analytical systems in education
- 4) Demographic information about the participants

Five questions are related to the first group, eight to the second group, 20 to the third group and the last group contains four questions.

2.1 Setting

The questionnaire is conducted by the SAP UCC (University Competence Center) Munich. The UCC Munich is an Education-as-a-service provider with customers in more than 21 countries. It is part of the Technical University of Munich and member of the SAP University Alliance. Besides the operation of SAP systems for universities and other academic institutions, creating teaching materials is another important aspect of the SAP UCC's services. (SAP UCC Munich, 2024)

2.2 The questionnaire

The process described in (Groves, 2011, S. 41ff.) as shown in Figure 1 is followed to develop the questionnaire. (Groves, 2011, S. 42) For the measurement part the research questions described in Chapter 1 are used as a construct and a general guidance during the measurement creation process. The survey is based on open and closed questions as proposed by (Brace, 2013, S. 38). Regarding open questions, open-ended and pre-coded questions are used. For the most closed questions, adding custom information is possible in an open-ended follow-up question in case it is helpful to get more information answering the research question. An ongoing discussion in the research community is the usage of “don’t know” responses. (Saris, 2014, S. 106) The author (Saris, 2014, S. 107) highlights challenges such as incomplete datasets and participants opting out to avoid cognitive effort. Other authors have a more optimistic perspective. (Oppenheim, 2004, S. 129) (Ryan & Garland, 1999) (Brace, 2013, S. 46f.) argue “don’t know” could be helpful in identifying topics your participants don’t know much about or it can lead to new patterns that should be analysed. Lastly (Lietz, 2010, S. 258) notes that the existence of an explicit “don’t know” option leads more participants to select it, compared to a questionnaire where the option is not explicitly available. Based on the recommendations, we decided to add an explicit “don’t know” option to all questions where it seems like a suitable answer. (Brace, 2013, S. 47) Additionally, only two questions are mandatory. Any other question can be skipped on request. The questions are displayed in the same order for all participants, because the questions are built on each other logically. To avoid primacy and recency effects in pre-coded open questions, answers are shown in a random order to all participants. This does not generally avoid the problem, but can help spread it more evenly. (Brace, 2013, S. 120)

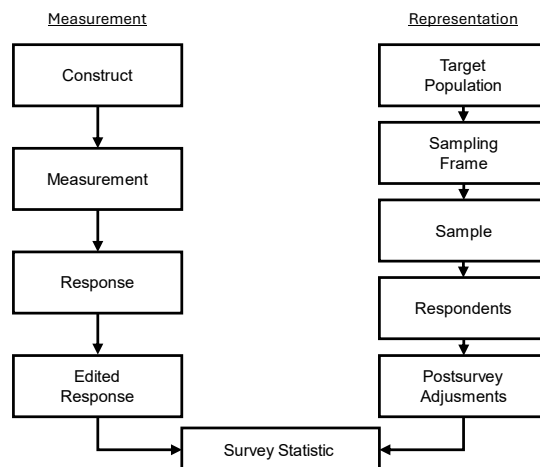


Figure 1: Survey process (Groves, 2011, S. 42)

2.3 Target population

This survey targets all kinds of university lecturers with a focus on data analytics. There is no difference between lecturers using SAP (UCC) products and lecturers not using SAP products. However, because the SAP UCC has the most contact with lecturers using SAP products, there is an imbalance between SAP users and non-SAP users in the frame population. Nevertheless, extending the population to as many non-SAP users as possible is important for a broad overview of used technologies. All lecturers in the final sample are contacted directly via e-mail with information about the questionnaire. We contacted about 2100 lecturers worldwide. Forwarding the invitation to other lectures was explicitly allowed.

2.4 Questionary pretests

Pretests are essential in getting feedback and improving the design of the questions and answers. (Bell, 2019, S. 265) (Baur & Blasius, 2022, S. 443) We conducted six pretests, as shown in Table 1. To make sure the questionnaire does not require SAP domain knowledge, we conducted the

	With domain knowledge	Without domain knowledge
PhD Student	3 persons	-
Postdoctoral Researcher	1 person	1 person
Business Professional	-	1 person

Table 1: Pretest overview

pretests with persons with and without explicit knowledge on SAP data analytics curricula. We used expert reviews and cognitive interviews as methods in our pretests. (Groves, 2011, S. 260) These two approaches were selected because they appear to be the best output based on the existing resources in this study. The expert reviews are used to ensure that the general questionnaire structure and the questions follow good scientific practices and are suitable for receiving scientific results. The expert reviews were conducted with two post-doctoral researchers. For the cognitive interviews, the Think-Aloud technique is used. These interviews aim to determine if all questions are understandable and whether the given answer options are suitable. Cognitive interviews are performed with three PhD students and one business professional. All pretests are performed as an iterative process, where feedback from a pretest is applied to the questionnaire before the next pretest starts. A shortcoming in our pretest approach is that we did not perform a field pretest. (Baur & Blasius, 2022, S. 446f.)

3. Results

The process proposed by (Groves, 2011, S. 47), shown in Figure 1, is extended by the process advised by (Bell, 2019, S. 165) to analyse the results. The combination is shown in Figure 2.

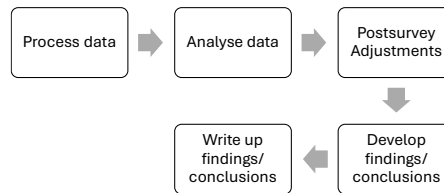


Figure 2: Evaluation process

3.1 Process data

All in all, 64 people started the questionnaire, and 40 finished the questionnaire completely. The dropout rate is 37,5%, which is slightly higher than expected in the literature. (Cape, Lorch, & Piekarski, 2007) (Brace, 2013, S. 13) Considering that only 10 of the 24 dropout cases proceeded beyond the introductory page, the dropout rate is only 20%, which is below the expected rate. Based on these findings, it can be concluded that the overall design of the questions was of adequate quality, and the preliminary tests were successful.

3.2 Analyse data and develop findings

According to the general and demographic information about the final participants, most participants (87,5 %) already use UCC services. Based on the target population (s. Chapter 2.3), this result is expected, but having 12,5% non-UCC service users in the respondents can give important insights for developing the UCC service offering. Most respondents are professors (57,5%), followed by research associates, lecturers, and lecturers for special tasks, with 10% each. Most participants have at least one data analytics course. Another interesting aspect is that while most courses (70 %) use UCC services, there are still 30% of courses using further software. This remains true even when only UCC service users are selected. 31,5% are using additional software in their courses. The other demographic information is shown in Table 2. Regarding the regional distribution, our study focuses on Europe and, more specifically, Germany, with a share of 55%. Only 2,5% of the participants are based outside of Europe.

Country	Share	Institution size	Share	Degree programmes	Share
Germany	55%	< 1000	20%	Computer science	37,5%
Poland	12,5%	1.000 - 5.000	25%	Information Systems	60%
United Kingdom	7,5%	5.000 - 10.000	22,5%	Data Science	30%

Ireland / Portugal	5% / 5%	10.000 - 20.000	25%	Economics	37,5%
Other	15%	> 20.000	7,5%	Other	20%

Table 2: Demographic information

This is related to the fact that most SAP UCC Munich customers are located in Europe, which is a limitation of this study. Regarding the institution size, the study is well distributed with multiple universities of different sizes. Furthermore, participants from several different disciplines are adequately represented.

3.2.1 RQ1: Topics in business and data analytics

The results regarding the first research question are shown in Figure 3. The integration of ERP systems with data analytics is already used by 52,5% of all lecturers, and an additional 40% are interested in introducing such use cases for their lectures. The integration can be seen as a relevant current and future topic. Two other areas of interest are the combination of SAP technologies with other tools and the offering of independent curricula modules (1 - 2 lecture units) that can be used to flexibly extend courses. For both fields, there is a dominant interest in curriculum elements related to machine learning/ artificial intelligence and data analytics. Other relevant topics like supply chain analytics, data visualisation, and process mining are also being considered.

One particular focus of the questionnaire was the relevance of business process intelligence in teaching. While 80% of participants consider this a relevant or highly relevant topic, only 20%

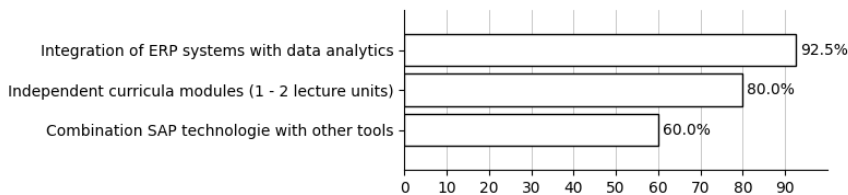


Figure 3: Topics of interest

already use an appropriate software solution. Especially SAP Signavio Process Intelligence is currently not very common for people who already use process intelligence software; only 12,5% of participants use Signavio Process Intelligence. The most relevant competitor is Celonis, with a proportion of 50%. This shows that there lies a huge potential in the process mining education tool market and a need to better integrate SAP Signavio into the current SAP teaching landscape.

3.2.2 RQ2: Opportunities and challenges using SAP cloud systems

The goal of the second research question was to understand the opportunities which SAP cloud systems reveal for teaching and what the challenges are associated with them. In this study, the

focus is on SAP Analytics Cloud and SAP Datasphere as SAP cloud solutions for analytics. Because of the increasing trend at SAP to focus on cloud applications, the integration of these systems in the education sector is necessary. (SAP SE, 2024) Based on the survey results, 40% of the lecturers haven't considered cloud opportunities yet, but 50% acknowledge at least a few opportunities to use the cloud. From a more detailed perspective, the most favoured opportunity is the removal of tool installations (browser only). This is especially true for SAP Analytics Cloud and SAP Datasphere, but only partially for SAP HANA, which has an already integrated web interface into the on-premise version.

Regarding the cloud challenges, nearly 50% of the participants don't see any challenges, and another 25% haven't thought about them yet. For the rest, the challenges mostly concern performance issues, support problems and open questions regarding pricing/ licensing. The SLA-level and scope differences between SAP business support and UCC support is an open issue and reinforces the need for comprehensive UCC support for universities. The issues regarding licensing and pricing remain, but there are already offers via the UCC that can solve this problem. A possible explanation for the low number of challenges seen could be, besides the fact that there may not be many challenges, the low saturation of the product in the education market. Like with all kinds of software, reliability is a characteristic that needs to be measured over time. (Wohlin, Höst, Runeson, & Wesslén, 2003, S. 25) The low saturation is also visible in the questionnaire. Only 2,5% already use an SAP cloud product in their lecture, and many lecturers are unaware of the SAP cloud services (80% SAP Datasphere/ 42,5% SAP Analytics Cloud). For a successful cloud transition, requirements related to the UCC were queried. The results are shown in table Table 3. The requirements A, C and D are partially related to the observed challenges and important aspects that need to be considered by SAP and UCC for a successful

Category	Results
(A) Fixed and predictable costs	60%
(B) Support for problems and challenges in teaching	50%
(C) Support for problems and challenges with the technology	52,5%
(D) Administration of the cloud environment	55%
(E) Providing teaching materials	77,5%
I have no special requirements	5%

Table 3 UCC cloud requirements (multiple choices possible)

cloud transition. B and E are aspects that already play an important role at the UCCs and remain important in cloud environments.

3.2.3 RQ3: Data analytics offerings (SAP and non-SAP) currently being used and are to be used in the future

As mentioned in Chapter 1, SAP is a global leader in business software, but the analytics market seems more fragmented, and the tool usage in education is unclear. Therefore, multiple questions in the questionnaire focus on the tool usage in education. Regarding SAP software, the most used product is SAP HANA (25%), followed by SAP Analytics Cloud (15%), but 37,5% of the lecturers aren't using any SAP analytics software right now. The results for non-SAP software are shown in Figure 4. The majority of lecturers are using Excel 52,5% followed by Python and R

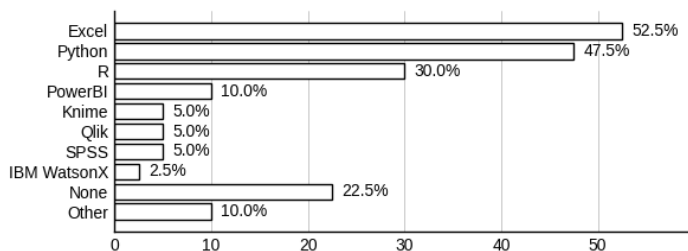


Figure 4 Non-SAP tool usage in education

with 47.5% / 30%. The wide usage of Excel is not surprising because it's one of the standard tools in business analytics and a good starting point for data analytics and visualisation. Also, using Python and R seems plausible because both tools are free of charge and offer a wide number of libraries (e.g. scikit-learn, numpy, matplotlib, keras) which are suitable for beginner and advanced courses. Besides these tools many other smaller products are also used in education, but there is a clear focus on Excel, Python and R. This trend could become even stronger in the future. 47,5% of participants plan to add new non-SAP products to their lectures. Excel, Python and R account for ~73% of this share, with other tools playing a subordinate role. The only exception is PowerBI, which is already used by 10% and has a share of ~11% of new introductions. PowerBI is also free of charge and offers no-code visualisation options, which could be easier to use than Python or R, depending on the course requirements. An open research question is how SAP products can make a difference in this tool environment. Either as completely independent tools or as an addition to the existing landscape.

4. Conclusion

This study conducts a questionnaire about SAP-based analytics in education with three research questions. Concerning the most relevant topics in business and data analytics, one important

finding is the interest in more extensive integration of ERP systems with analytical software. This can be used as a starting point for further curricula development and can be a unique feature offered by SAP in education. Independent curricula modules, which flexibly extend various curricula, are a valuable finding too. An important result about the opportunities and challenges of using cloud services is the low saturation of SAP cloud products in education. Many lecturers are unaware of the cloud offering and have not invested time considering the opportunities and challenges of using a cloud product. It is the task of the SAP UCC to clarify open questions with SAP and provide a suitable offer for cloud products in education. Using free-of-charge software like Python, R, and PowerBI is an educational trend.

In summary, this paper contributes to ongoing research on integrating SAP (analytical) software in education. Based on the information offered by this questionnaire, multiple new tasks and research questions arise. For SAP and SAP UCC, an ongoing question is how to integrate Cloud software successfully in education. Especially the licensing topic is a current issue. From an educational perspective, the insights about current topics and tools can be used as a starting point for further research. For example, developing small independent curricula modules for different knowledge levels and tools in integration with SAP services.

The main limitation of this study is the target population used. While the authors tried including non-SAP UCC customers in the survey, most participants are SAP UCC customers. In future studies, more effort is necessary to get a broader questionnaire sample. The other limitation is the sample size of 40 participants. Although the participants were distributed internationally (with a focus on Europe) from universities of different sizes and with different research fields, a larger overall sample would be more expedient. Nevertheless, we see our questionnaire as a good starting point, which we will continue with further interviews.

5. References

- Baird, A. M., & Parayitam, S. (2019, June). Employers' ratings of importance of skills and competencies college graduates need to get hired: Evidence from the New England region of USA. *Education + Training*, *61*, 622–634. doi:10.1108/et-12-2018-0250
- Baur, N., & Blasius, J. (Eds.). (2022). *Handbuch Methoden der empirischen Sozialforschung* (3., vollständig überarbeitete und erweiterte Auflage ed., Vol. Band 2). Wiesbaden, Germany: Springer VS.

- Bell, E. (2019). *Business research methods* (Fifth edition ed.). (A. Bryman, & B. Harley, Eds.) Oxford: Oxford University Press.
- Brace, I. (2013). *Questionnaire Design* (3rd ed (Online-Ausg.) ed.). London: Kogan Page.
- Cape, P., Lorch, J., & Piekarski, L. (2007). A tale of two questionnaires. *Proceedings of the ESOMAR Panel Research Conference*, (pp. 136–149).
- Chen, Chiang, & Storey. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*, 1165. doi:10.2307/41703503
- Groves, R. M. (2011). *Survey Methodology* (2nd ed (Online-Ausg.) ed.). (F. J. Fowler, M. Couper, J. M. Lepkowski, E. Singer, & R. Tourangeau, Eds.) Somerset: Wiley.
- Lietz, P. (2010, March). Research into Questionnaire Design: A Summary of the Literature. *International Journal of Market Research*, *52*, 249–272. doi:10.2501/s147078530920120x
- Oppenheim, A. N. (2004). *Questionnaire design, interviewing and attitude measurement* (New ed., reprinted ed.). London [u.a.]: Continuum.
- Ryan, C., & Garland, R. (1999, February). The use of a specific non-response option on Likert-type scales. *Tourism Management*, *20*, 107–113. doi:10.1016/s0261-5177(98)00096-x
- SAP SE. (2024). RISE with SAP. *SAP Corporate Website*. Retrieved 05 03, 2024, from <https://www.sap.com/products/erp/rise.html>
- SAP UCC Munich. (2024). Our Mission. *Our Mission*. Retrieved 03 11, 2024, from <https://ucc.tum.de/our-mission/>
- Saris, W. E. (2014). *Design, evaluation, and analysis of questionnaires for survey research* (Second edition ed.). Hoboken: Wiley.
- Torii, D., Grinter, S., Faith, T., Mahendra, N., Ralhan, N., & Anderson, R. (2023). *Magic Quadrant for Cloud ERP for Service-Centric Enterprises*. Tech. rep., Gartner, Inc.
- Whitelock, V. (2018, July). Business analytics and firm performance: role of structured financial statement data. *Journal of Business Analytics*, *1*, 81–92. doi:10.1080/2573234x.2018.1557020
- Wohlin, C., Höst, M., Runeson, P., & Wesslén, A. (2003). Software Reliability. In *Encyclopedia of Physical Science and Technology* (pp. 25–39). Elsevier. doi:10.1016/b0-12-227410-5/00858-9