



Contributions of the Analysis of Academic Trajectories to Educational Management

Charles H. M. Junior¹, Thiago M. R. Dias¹, Patrícia M. Dias²

¹*Federal Center for Technological Education of Minas Gerais
Rua Álvares de Azevedo 400 - Bela Vista – Zip-Code: 35503-822 - Divinópolis - MG - Brazil
acad.charlesjr@gmail.com, thiagomagela@cefetmg.br*

²*State University of Minas Gerais
Av. Paraná, 3001 - Jardim Belvedere – Zip-Code: 35501-170 - Divinópolis – MG - Brazil.
patriciamdias@gmail.com*

Abstract. Educational management is present and necessary in contemporary times, given the ability of the techniques and dynamics proposed to guide education systems and maximize the investments made. Given this fact, conclusions about social phenomena that can be guaranteed by analyzing academic trajectories are of significant relevance, since they reveal information regarding the effectiveness of education centers, the means of re-establishing training paths that have been broken and the variables in the context of individuals that influence the teaching-learning route. In view of this, this work aims to carry out a case study with an educational institution, so that results can be extracted from the investigation of educational histories that collaborate with the institution's strategies. Based on CVs from the CNPq's Lattes Platform, the research was structured around the stages of Data Selection, Filtering, Summarization and Stylization, with interpretations of educational profiles, geographical mobility for academic and professional purposes, migration between institutions and other related data, from an understanding that ranges from high school to doctoral level. In this way, conclusions were reached, in the first phase of the work, regarding the scope of the institution's pedagogical projects and the spatial retention of qualified individuals by the campuses.

Keywords: Academic Trajectories, Lattes Platform, Data Analysis.

1 Introduction

Given the countless areas of society that are impacted by education, the definition of “quality” is ambiguous and relative to the different views and interests involved [1]. Understanding the meaning of “excellence” presupposes a socioeconomic and cultural analysis of a country, so that the interaction and representativeness of educational policies in the government’s portfolio are observed in conjunction with the local and global economic contexts [2]. With the justification of obtaining the best returns for society, during the last two decades, states and multilateral organizations have been involved in the establishment of regulatory agencies and verification processes for higher education courses on a global scale [1]. This unprecedented categorization was also accompanied by social transitions that still demand transparency, democracy and excellence arising from the process of globalization and massification [2].

With regard to Brazilian higher education, the 1990s were a landmark decade for the development of epistemological systems that measured the effectiveness of educational institutions. It is fitting to consider as a symbol of this new national movement the implementation at the time, by the National Institute of Studies and Educational Research Anísio Teixeira (INEP), of the *provão*, a test intended for students who were completing their respective courses, and of the Assessment of Offer Conditions (ACO), the in-person inspections of

infrastructure, pedagogical projects and teachers that aimed at regulation and restructuring [1]. In order to improve these already established metrics, the National Higher Education Assessment System (SINAES), from 2004, emerged with the objective of contemplating a global, incorporated and comprehensive analysis regarding the diversity and particularities of the influences, structures, relationships, activities, purposes and social responsibilities of Higher Education Institutions (IES) and the courses offered. The SINAES includes institutional assessment, based on self-assessment carried out by the Permanent Assessment Committee (CPA), course assessment, which in addition to peer assessment on site, considers responses from various structural positions of the HEI, and the assessment of the performance of new and graduating students through the National Student Performance Exam (ENADE), a test with questions on general education (25%) and specific content (75%) that aims to assess basic, academic and professional skills, in addition to interdisciplinarity [3]. Since then, several government education plans have followed one another, supported by different political leaders, which culminated in the concept of “quality” sometimes supported by humanist principles, in order to privilege the autonomy and citizenship of the subjects, sometimes aligned with interests that seek an education that integrates individuals into the dynamics of the market, sometimes guided by other sets of ideas that guided and guide, directly and indirectly, republican decisions [4].

However, despite the disagreement about which qualitative factors guide educational primacy, there is a need for management instruments whose application is viable and capable of implementing and operating necessary transformations in education systems in the organization of Brazilian education, since educational management establishes directions and mobilizations that structure and stimulate the methodology of student entities and tend to avoid time-consuming and harmful expenditures for the whole [5]. Therefore, educational management can be seen as the process of managing the dynamics of the education system [6] in line with educational guidelines and policies. The logic of management is guided by democratic precepts, since its conception assumes that large-scale obstacles require collective action and, therefore, is characterized by the dialogue of agents in decisions about guidance, organization, planning and articulation of the pillars that determine actions. In this sense, the differentiation between management and administration is relevant due to the confusion between the terms and the distinct natures of which they are part. Management is based on a rational, linear, fragmented and hierarchical process, whose distant and objective command and control maintains the focus on the use of instruments and resources with utilitarian and mechanistic biases. The action of management establishes assumptions such as the predictability of human behavior and the environment, crises as dysfunctions to be avoided and success as a goal that does not require subsequent maintenance [5]. Thus, the manager must be the one who overcomes dimensional restrictions and decontextualization and builds, systematized by the criteria of interactivity and broad vision, the image of the group that develops articulated and more consistent actions.

In the contemporary world, Information and Communication Technologies (ICT) have an indispensable presence in the fields of human life and especially in education, since new forms of knowledge acquisition take advantage of technological instruments for their perpetuation [7]. In line with this fact, for educational management to be carried out fully, the presence of ICT is essential in the development of better administrative practices with expanded and dynamic procedures [8]. Especially with regard to the focus, positioning and control stages of strategic planning [9], the relevant role that technologies play in data collection and analysis is identified [10]. In this context, academic trajectories are used as a parameter of social phenomena, including for insertion in the job market and for positioning in the social structure. There are efforts to understand what educational trajectories reveal and how their understanding enables the transformation of certain realities. However, it is known that student trajectories are capable of providing information on the effectiveness of the educational system, on demographic and socioeconomic variables, and on necessary interventions that enable the resumption of the educational path [9].

Therefore, faced with an object of analysis that allows the understanding of educational formation as a social phenomenon, the present production aims, from the perspective of a case study, to analyze information regarding the academic trajectories of individuals who passed through the Federal Center for Technological Education of Minas Gerais (CEFET-MG) and the institution's relationships with other agents and environments.

2 METHODOLOGY

To achieve the objective of this production, the team used a database containing resumes from the Lattes Platform in XML format. The application in question emerged as an initiative of the National Council for Scientific and Technological Development (CNPq) and aims to unite academic resume databases from public and private institutions in a single location. Seen as a Brazilian standard for evaluation, Lattes resumes guarantee the history of academic and professional activities of registered researchers [10].

In total, 1,041,206 resumes were explored, which corresponds to practically all masters and doctorates in Brazil. The processing of the cluster was subdivided into four stages, namely: Selection, Filtering, Summarization and Stylization, as shown in the flowchart in Figure 1.

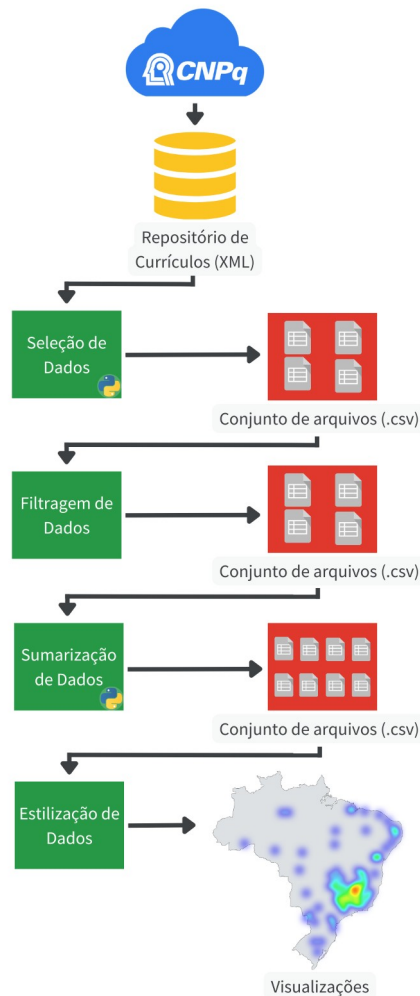


Figure 1: Overview of Project Operation

Data Selection:

Initially, the Data Selection sought to establish the delimitation of the theme by defining a subgroup of those who built part of their academic career at CEFET-MG. Secondary and technical-vocational education and undergraduate, master's, doctoral and specialization degrees were considered. The other levels of education were disregarded due to technical obstacles in obtaining the content of these categories and the focus initially proposed by the work. With the segmentation performed, data whose relevance is notorious for understanding the scope and influence of the federal institution in society were extracted, such as the complete academic

history, place of birth and the most recent institution of professional, academic or scientific activity of those involved. The extraction results were stored in Comma-Separated Values (CSV) format files. Given the considerable presence of a scientific processing and data analysis community that is adept at technology, Python is one of the most important programming languages in data science in professional and university environments [11]. Therefore, to execute a significant part of what is proposed in this work, this resource was used and, especially in the step in question, XPath (XML Path Language) was inserted, since it is a tool capable of accessing different points of an XML document in a suggestive and structured manner.

Data Filtering:

It is expected that, since this is characterized as a database for free data insertion [10], certain spelling, typing, grammatical and similar errors will be present in the records acquired from the resumes. To this end, the solution to these problems, an essential factor in the quality of the desired analyses, is tangible with the Clustering process, which seeks to group similar objects through a pre-established similarity measure [12]. In view of the above, integration with an IDT (Interactive Data Transformation Tool) allows optimized data cleaning operations under a unified interface. When considering accessibility and practicality in implementation, Openrefine and its matching parameters, such as Levenshtein Distance, were used to filter the data.

Data Summarization:

The module entitled Data Summarization performs, after organizing and reducing tabular data, the quantification of certain occurrences and events, with a view to simplifying the processes of interpreting results through more compact, direct, objective and pragmatic visualizations. In this context, more than being on the path to becoming the most powerful and flexible open-source tool for data manipulation and analysis, the Pandas library fulfills the mission of being a fundamental element for real data analysis in Python [13]. Its functionalities cover the manipulation of different types of data, such as tables and observational and statistical data sets, with the treatment of missing data, mutability, cell alignment, indexing, combination of sets and numerous tasks that precede analytical conclusions. The implementation of Pandas in the Data Summarization stage aims to optimize and measure results obtained in previous stages. For this production, new records were obtained that contained information on the number of graduates per course and per year, the number of graduates per course and per year in the vicinity of CEFET-MG, the number and recurrence of cities of birth in common, as well as the number of individuals who worked in the same place.

Data Styling:

However, it is known how important it is to create data visualizations for the dissemination of university production to society, as well as for the extension of activities [14]. Therefore, in order to take advantage of the benefits of visual and scientific resources, three main visualizations of the data processed during this production will be created. The tree map will be used to understand the course conclusions by level of education and its construction will be done using Flourish, a data visualization tool. Gephi, a network visualization software, is suitable for observing the relationship between CEFET-MG and other national and international educational institutions. Furthermore, based on the strengths of the Python ecosystem, the Folium library allows the creation of geographic heat maps based on the number of students born by region and the locations of those who have a history at CEFET-MG.

3 Results

From the initial database, 1,393 CVs were selected from students who had some level of education at CEFET-MG. Of these individuals, approximately 54% were from the city of Belo Horizonte/MG, followed by Contagem/MG and Divinópolis/MG, with 5% and 2.4%, respectively. Regarding the most recent location of work, the capital of Minas Gerais maintained its predominance, guaranteeing the employment of 55% of those

who at some point studied at CEFET-MG, together with Divinópolis/MG and Betim/MG, with two shares of 2.4%. Figures 2 and 3 elucidate this set of percentages.

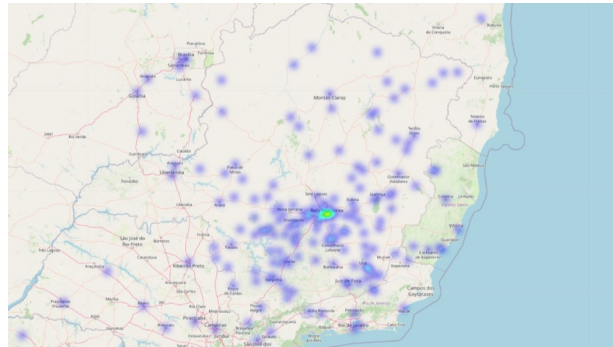


Figure 2: Geographic heat map based on the incidence of municipalities as places of birth.

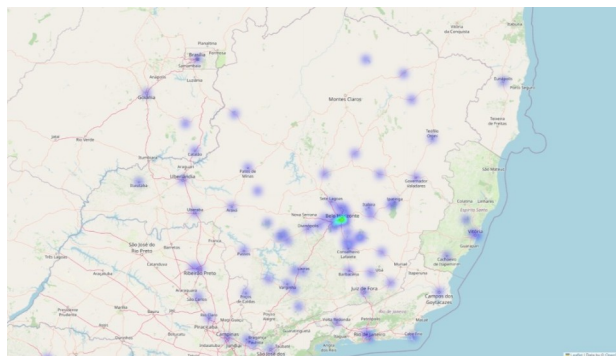


Figure 3: Geographic heat map based on the incidence of municipalities as locations of action.

From the data analysis, it is possible to verify the recurrence of course completion according to the level of education. The areas that had the most graduates registered in the CEFET-MG undergraduate courses were electrical, civil and mechanical engineering and the Special Program for Pedagogical Training of Teachers (Figure 4); in the case of the doctorate degree at the same institution, Mathematical and Computational Modeling and Language Studies gained notoriety (Figure 5). Data on professional technical courses and master's degrees were also collected and are represented in Figure 6.



Figure 4: Number of individuals who completed a degree at CEFET-MG according to the course.



Figure 5: Number of individuals who completed a doctorate at CEFET-MG according to the area.

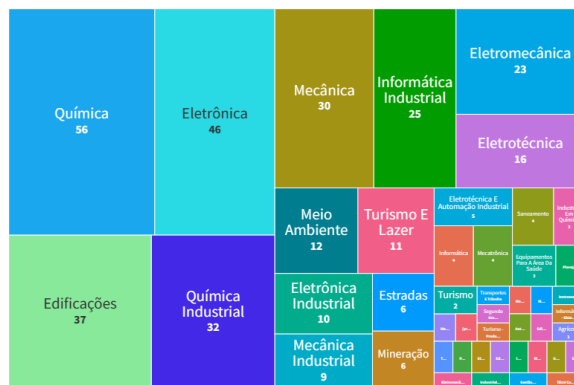


Figure 6: Number of individuals who completed a professional technical course at CEFET-MG.

Regarding relations between universities, Figure 7 shows a significant exchange of students from CEFET-MG to the Federal University of Minas Gerais (UFMG). In addition, there is a significant flow of students from UFMG and the Pontifical Catholic University of Minas Gerais (PUC Minas) to the institution that is the protagonist of this work.

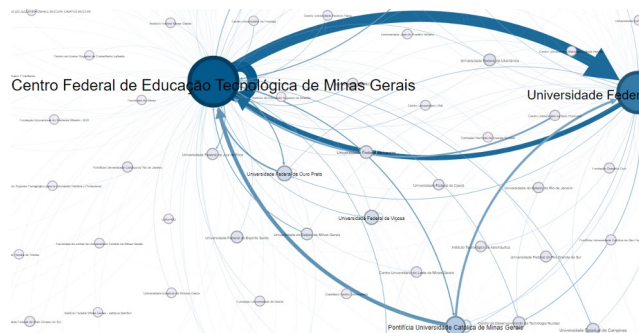


Figure 7: Network of edges directed according to the origin and destination of student exchange.

4 Conclusion

Regarding the relationship between place of birth and place of work, Belo Horizonte (BH) stood out for being the origin and final location of a significant portion of the new students at CEFET-MG. A result that would take this form was expected, given the fact that the city encompasses the central governance of the state and the history of the federal institution, which began in 1909, dating back to the inauguration of the first unit based in the capital of Minas Gerais. As for the other municipalities that have CEFET-MG units, despite the different orders of magnitude compared to BH, Divinópolis and Contagem stand out with regard to the permanence of students in the geographic region. In the first, in particular, the institution of predominant activity

is the Federal Center for Technological Education of Minas Gerais. However, even though the unit has suppressed intermunicipal migrations of qualified citizens, the 2022 Management Report of the Directorate of Extension and Community Development (DEDC) [15] indicates that only 7.6% of the extension actions of the entire network, that is, the execution of courses, events, programs and projects, are aimed at Campus V.

There is a disparity between the number of graduates of the technical course in the area of chemistry and the number of graduates of the undergraduate course in the area of chemistry within the CEFET-MG units. In line with this, it is attested that the higher education course in question, with a much smaller number of graduates, is only available at the Nova Gameleira Campus. According to the official platform, the modality was implemented in 2006 under the justification of meeting regional demands and, therefore, was shaped by the environmental, food, energy and mineral pillars. To solve this problem, it is necessary to consider expanding the Technological Chemistry course as a way of serving all citizens who aim for training and subsequent specialization. Furthermore, the number of graduates, at higher levels of education, of courses that address Language Studies suggests a need to implement, also in professional technical courses, education that guarantees linguistic knowledge and its adjacent aspects.

Of the possible relationships between educational institutions, CEFET-MG, UFMG and PUC Minas establish a connection in which the Catholic university, for the most part, sends students to the other two educational centers. When also considering the large number of students from CEFET-MG to UFMG, the demand for ICT-related courses observed in the tree maps obtained in this work and the international recognition, the continuous investment and dedication to the basic areas of Computer Science, Engineering and related fields corroborate the dynamics of the next generations and agree with the needs addressed.

References

- [1] Júlio C. G Bertolin and Telmo Marcon.. 2015. O (des)entendimento de qualidade na educação superior brasileira – Das quimeras do provão e do ENADE à realidade do capital cultural dos estudantes. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 20, 1: 105–122.. <https://doi.org/10.590/S1414-40772015000100008>.
- [2] Mário César Barreto Moraes and Guilherme Felipe Kalnin.. 2018. Qualidade na educação superior: uma revisão teórica da evolução conceitual no campo da educação superior. *Ensaio: Avaliação e Políticas Públicas em Educação* 26, 100: 530–551.. <https://doi.org/10.1590/S0104-40362018002601114>.
- [3] Márcia Regina F. de Brito.. 2008. O SINAES e o ENADE: da concepção à implantação. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 13, 3: 841–850.. <https://doi.org/10.1590/S1414-40772008000300014>.
- [4] Marília Fonseca.. 2009. Políticas públicas para a qualidade da educação brasileira: entre o utilitarismo econômico e a responsabilidade social. *Cadernos CEDES* 29, 78: 153–177.. <https://doi.org/10.1590/S0101-3262200900020000>.
- [5] Heloísa Lück. 2017. *Gestão Educacional: Uma Questão Paradigmática*. Editora Vozes.
- [6] Rejane Oliveira and Sebastião de Souza Lemes. 2022. A gestão educacional e os referenciais cognitivos e normativos em política pública. *Revista on line de Política e Gestão Educacional*. 26, 00 (mar. 2022), e022002. <https://doi.org/10.22633/rpge.v26i00.16741>.
- [7] Angela Maria Gonçalves de Oliveira and Glaucilene Sebastiana Nogueira Lima. A gestão educacional e a efetivação de políticas públicas para utilização das TIC na educação. 2016. *Revista Exitus* 5, 2: 125–137.
- [8] Geovana Rafael Theisen, Sílvia Maria de Oliveira Pavão and Ana Cláudia Oliveria Pavão. 2021. TIC na gestão educacional: efetividade na academia. *Revista on line de Política e Gestão Educacional*. 25, 2 (ago. 2021), 1264–1281. <https://doi.org/10.22633/rpge.v25i2.14576>.
- [9] Sônia Simões Colombo. 2007. *Gestão Educacional: Uma nova visão*. Artmed Editora.
- [10] Sueli Liberatti Javaroni, Silvana Claudia Dos Santos, and Marcelo de Carvalho Borba. 2011. TECNOLOGIAS DIGITAIS NA PRODUÇÃO E ANÁLISE DE DADOS QUALITATIVOS. *Educação Matemática Pesquisa* 13, 1 (2011), 197-218.
- [11] Alice Plakoudi Souto Maior, Camila Akemi Karino and Joaquim José Soares Neto. 2021. Trajetórias acadêmicas de estudantes na educação superior: uma revisão da literatura. *Perspectivas em Diálogo: Revista de Educação e Sociedade*. 8, 17 (jun. 2021), 464-489. <https://doi.org/10.55028/pdres.v8i17.12226>.
- [12] Aline Grasielle Cardoso de BRITO, Luc QUONIAM, and Jesús Pascual MENA-CHALCO.. 2016. Exploração da Plataforma Lattes por assunto: proposta de metodologia. *Transinformação* 28, 1: 77–86.
- [13] Wes McKinney. 2018. *Python para análise de dados: Tratamento de dados com Pandas, NumPy e Ipython*. Novatec Editora.
- [14] Eduardo Vieira Queiroga. 2017. *Abordagens Meta-heurísticas para Clusterização de Dados e Segmentação de Imagens*.
- [15] Pandas Development Team. 2012. *Pandas: powerful data analysis toolkit*, Release 0.7.3.
- [16] Rafael Ferreira dos Santos and Taiane Ritta Coelho. 2021. Proposta de um modelo de visualização de dados: Impacto social da extensão do ensino superior sob a ótica de Campus Inteligente. *Revista Informação na Sociedade Contemporânea*. 5, (abr. 2021), e23416. DOI:<https://doi.org/10.21680/2447-0198.2021v5n1ID23416>.
- [15] Diretoria de Extensão e Desenvolvimento Comunitário DEDC-CEFET-MG. 2022. *Relatório de Gestão 2022*.