www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 18(1), 2027, 189-199. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2027.v18i1.1509

Data Pipeline Architecture for the Integration of Technology and Artificial Intelligence in the Competencies of the Future Manager

Morayma G. Orán-Muñoz[1], Mayra L. González-Mosqueda[1], Karla Martínez-Tapia[1], Alfonso Ávila-Pérez Tagle[1], Enrique J. Mohedano-Torres[2], Francisco Chiapa-Téllez[2], Domingo N. Marrón Ramos[2]

[1] Departamento de Ciencias Económicas y Administrativas, Instituto Tecnológico Nacional de México-Instituto Tecnológico de Pachuca,

[2] Departamento de Ingeniería Industrial, Instituto Tecnológico Nacional de México - Instituto Tecnológico de Pachuca.

E-mails: mayra.gm@pachuca.tecnm.mx, karla.mt@pachuca.tecnm.mx, alfonso.ap@pachuca.tecnm.mx, enrique.mt@pachuca.tecnm.mx, francisco.ct@pachuca.tecnm.mx, dmarron22@pachuca.tecnm.mx

✉Corresponding author: morayma.mo@pachuca.tecnm.mx

Abstract: This study proposes an automated Extract, Transform and Load (ETL)-based methodological framework that relates employer-demand data to institutional enrolment records through a sequential evidence-integration procedure to support evidence-informed curriculum planning. The framework was implemented in Python using two complementary data sources: 52 completed questionnaires obtained from an accessible population of 60 active organisations with institutional collaboration agreements, and a historical database of 3,149 student enrolment records. The ETL workflow automated data acquisition, preprocessing, validation and source-specific descriptive statistical analysis to identify priority technological competencies. The findings identified high employer demand for digital transformation, business intelligence, artificial intelligence and process automation competencies, while institutional enrolment records provided complementary evidence to support curriculum planning from an institutional perspective. The proposed framework contributes a structured, reproducible and transparent methodology for relating heterogeneous evidence to support academic specialisation design in higher education.

Keywords: Technological Competencies, Curriculum Design, Digital Management, Business Intelligence.

Article Information
Received on: Jul 01, 2026
Accepted: Aug 01, 2026

1 Introduction

The adoption of Artificial Intelligence (AI), Machine Learning and Big Data architectures has modified the operational structure and traditional decision-making processes in contemporary organisations (Brynjolfsson & McAfee, 2014; World Economic Forum [WEF], 2025). Business management is shifting towards a model based on quantitative evidence, where strategic planning and operational control depend on analytical information flows rather than traditional empirical schemes (Davenport & Mittal, 2023; McAfee & Brynjolfsson, 2012). This scenario poses a fundamental problem for higher education institutions, which must resolve the existing asymmetry between curricula designed under linear contexts and the computational skills that the labour market demands today (Organisation for Economic Co-operation and Development [OECD], 2023). At the regional level of the State of Hidalgo, this dynamic demands professional profiles suitable for integrating sectors in the process of technological modernisation and responding to institutional policies for technological incorporation (Consejo de Ciencia, Tecnología e Innovación de Hidalgo [CITNOVA], 2024). In bachelor's degrees in the economic-administrative area, this gap manifests itself in a purely instrumental teaching of computer science, limiting the ability of graduates to coordinate complex automated systems (Sousa & Rocha, 2019).

Previous studies have addressed data-driven curriculum design and learning analytics in higher education (Chu & Ashraf, 2025; Viberg et al., 2018); however, employer requirements and institutional enrolment evidence are frequently treated through separate analytical processes. Consequently, curriculum design frequently relies either on employer surveys or on institutional indicators, limiting the integration of external labour-market demand with institutional enrolment evidence for curriculum decision-making. This methodological gap highlights the need for analytical frameworks capable of relating heterogeneous educational and organisational evidence within a transparent and reproducible computational workflow. Traditional curriculum-updating processes frequently rely on manual document review, descriptive analysis and expert consensus. Although these procedures remain valuable, they may limit the traceability and systematic integration of evidence obtained from different sources. To support a more timely and structured alignment with the socio-productive context, this study implemented a software-supported workflow for ingesting, cleaning and analysing employer and institutional datasets, including the proportional normalisation of ordinal employer responses, in a reproducible manner. In this context, this study presents and justifies the design of a specialisation for the Bachelor's Degree in Administration at the Technological Institute of Pachuca (ITP). The curricular structuring of the proposed specialisation followed the institutional guidelines established for the integration of specialisations (Tecnológico Nacional de México [TecNM], 2015).

The methodological contribution does not lie in the use of ETL processes, Python or descriptive statistics in isolation. Instead, it lies in the implementation of a traceable and sequential evidence-integration architecture that connects the separate preparation of employer and institutional datasets, the proportional normalisation of ordinal responses, the automated prioritisation of competency dimensions through a predefined threshold and the curricular mapping of the resulting priorities. Unlike conventional survey processing, the proposed framework maintains an explicit relationship between data preparation, competency scoring, institutional context and curricular decision-making. It differs from learning-analytics approaches because it does not predict individual student performance or learning outcomes, and it is not presented as a standalone decision-support software system. Therefore, the novelty of the proposal lies in the transparent linkage and reproducibility of the complete process, rather than in the development of a new ETL algorithm, predictive model or advanced statistical technique.

The organisation of this article is as follows. Section 2 describes the proposed methodological framework, including the sampling strategy for the accessible employer population, the ETL workflow, the reliability assessment of the research instrument and the computational implementation in Python. Section 3 presents the descriptive statistical analyses and discusses the correspondence between employer competency requirements and the proposed curriculum structure. Finally, Section 4 presents the conclusions, limitations and future research directions.

2 Methodology

To support the design of the new Specialisation in Digital Management and Business Intelligence within the Bachelor's Degree in Administration at the Technological Institute of Pachuca, the study followed a quantitative and descriptive approach and employed an ETL-based methodological framework to relate employer-demand data to historical institutional enrolment records through a sequential evidence-integration procedure (Hernández Sampieri & Mendoza Torres, 2018). The two datasets were processed separately because they represented different units of analysis and were not merged at the individual-record level. The methodology was organised into four components: data sources and sampling procedure, ETL pipeline architecture, instrument structure and reliability assessment, and statistical analysis and computational environment.

2.1 Data Sources and Sampling Procedure

Employer data source and accessible population. The employer sampling frame was obtained from the institutional directory of organisations linked to the Technological Institute of Pachuca through collaboration agreements and professional residency activities involving students from the Bachelor's Degree in Administration programme. The original directory comprised 61 organisations located in the state of Hidalgo. One commercial microenterprise was excluded because it had ceased operations at the time of data collection. Therefore, the final accessible population consisted of 60 active organisations.

The accessible population included 26 organisations from the commerce sector (43.3%), 24 from the service sector (40.0%) and 10 from the manufacturing sector (16.7%). Regarding organisational size, it comprised seven microenterprises (11.7%), 18 small enterprises (30.0%), 13 medium-sized enterprises (21.7%) and 22 large enterprises (36.7%).

Sample-size estimation and participation procedure. The finite-population formula was used as a reference to estimate an adequate number of completed questionnaires for the employer component of the study:

$$\boldsymbol{n} = \frac{NZ^2pq}{e^2(N-1)+Z^2pq} \tag{1}$$

Where:

- ***n:*** Sample size required for the employer survey.
- ***N:*** Accessible population of active organisations ($N = 60$).
- **Z**: Confidence-level coefficient ($Z = 1.96$, corresponding to 95% confidence)

- ***P:*** Expected positive variability ($p = 0.5$)
- ***q:*** expected negative variability ($q = 0.5$) and
- ***e*** tolerated margin of error ($e = 0.05$).

Substitution of these parameters yielded an estimated sample size of 52.01 organisations, corresponding to approximately 52 completed questionnaires. The formula was used to establish a reference response target rather than to select organisations randomly, because the electronic questionnaire was distributed to all 60 eligible organisations. A total of 52 complete questionnaires were obtained, representing a response rate of 86.7% of the accessible population. The remaining eight eligible organisations did not provide a completed questionnaire and were therefore not included in the analytical dataset. Characteristics of the participating organisations. The final analytical sample comprised 52 organisations from the commerce, service and manufacturing sectors and included micro, small, medium-sized and large enterprises. Their distribution by sector and organisational size is presented in Table 1.

Table 1. Characteristics of the participating organisations

Variable	Category	n	%
Economic sector	Commerce	22	42.3
	Services	21	40.4
	Manufacturing	9	17.3
Organisational size	Microenterprise	7	13.5
	Small enterprise	17	32.7
	Medium-sized enterprise	11	21.2
	Large enterprise	17	32.7
	Total	**52**	**100.0**

Note. Percentages were calculated using the 52 organisations included in the final analytical sample.

Source: Own elaboration based on the employer database used in the study.

The designated respondents included directors, managers, administrators, unit supervisors, business owners and other organisational representatives with administrative or managerial responsibilities. Organisations were considered eligible when they were active at the time of data collection, were located in the state of Hidalgo and formed part of the Institute's collaboration or professional residency network. For inclusion in the analytical dataset, a complete questionnaire submitted by an eligible organisational representative was required. The organisation that had ceased operations was excluded from the accessible population, whereas the eight organisations that did not submit a completed questionnaire were treated as non-respondents.

The sectoral and organisational-size distributions of the participating organisations were broadly comparable to those of the accessible population. Nevertheless, the sampling frame represented the Technological Institute of Pachuca's employer network rather than the complete business population of the state of Hidalgo. Consequently, the findings are interpreted within this institutional and regional context and are not generalised to all organisations operating in the state. The possibility of non-response bias associated with the eight non-participating organisations is also acknowledged. Institutional data source. The second source consisted of a historical institutional database comprising 3,149 enrolment records from the Bachelor's Degree in Administration programme at the Technological Institute of Pachuca. These records represented a different unit of analysis and were independent of the employer sampling frame. Therefore, the value of 3,149 was not used as N in the employer sample-size calculation. The institutional records were used to describe observed enrolment patterns and provide contextual evidence for curriculum planning, rather than to demonstrate the availability of infrastructure, teaching staff or operational capacity.

Analytical relationship between the data sources. Because employer questionnaires and institutional enrolment records represented different units of analysis, the two datasets were not merged at the individual-record level. Instead, a sequential two-stage evidence-integration procedure was followed. In the first stage, employer responses were used to identify and rank the technological competencies required in the regional work environment. In the second stage, institutional enrolment records were used to describe the size and aggregate enrolment pattern of the target student population as contextual information for planning the proposed specialisation. No numerical weighting was applied between the two data sources. Employer results determined the prioritisation of technological competencies, while enrolment information did not modify the competency scores or their ranking. The outputs from both stages were subsequently considered during the curricular mapping process:

employer evidence informed the competencies to be addressed, whereas institutional records provided the academic context in which the specialisation was planned.

2.2 ETL Pipeline Architecture

The processing of the collected information was carried out through a parameterised Extract, Transform and Load (ETL) architecture implemented in Python, following the general logic of conventional ETL processes (Vassiliadis, 2009). This computational workflow was designed to automate data ingestion, preprocessing, validation and statistical analysis, supporting the reproducibility and consistency of the analytical process. The architecture consisted of three sequential phases: Extraction, Transformation and Loading, each of which performed a specific function within the data pipeline. Because the employer questionnaires and institutional enrolment records represented different units of analysis, the two datasets were processed separately and their outputs were subsequently related through the sequential integration procedure described in Section 2.1. The architecture was therefore designed as a traceable sequence from source-specific data preparation to dimension-level competency prioritisation and subsequent curricular mapping; it did not involve record-level merging, predictive modelling or the development of a standalone decision-support application.

1.Extraction Phase. The primary information collected from the 52 participating organisations was automatically captured through a cloud-based digital questionnaire, reducing the risk of manual transcription errors during the data acquisition process. Each response was stored in a structured dataset, where every record represented one organisation and each column corresponded to a scored competency indicator derived from an ordinal question or matrix sub-item of the research instrument. This automated extraction stage supported data integrity before computational processing.

2.Transformation Phase. During the transformation stage, the extracted data were organised into a structured data matrix, where each row represented a participating organisation and each column corresponded to one scored competency indicator derived from the research instrument. This matrix constituted the computational input for the automated preprocessing routines implemented in Python. Before the calculation of the indicator-level and dimension-level descriptive statistics, ordinal responses obtained from three-, four- and five-category items were proportionally normalised to a common 1-to-5 scale. The proportional transformation was performed using the following equation:

$$\boldsymbol{x_{ij}} = \mathbf{1} + \mathbf{4}\left(\frac{\boldsymbol{r_{ij}} - \mathbf{1}}{\boldsymbol{k_j} - \mathbf{1}}\right) \quad \textbf{(2)}$$

Where:
$\boldsymbol{r_{ij}}$ represents the original ordinal response provided by organisation i for competency indicator j.
$\boldsymbol{k_j}$ represents the number of response categories associated with indicator j, corresponding to three, four or five categories.
$\boldsymbol{x_{ij}}$ represents the resulting normalised score on the common 1-to-5 scale.

The structured dataset can be mathematically represented as:

$$\mathbf{X} = \left[x_{ij}\right]_{n\times m} \quad \textbf{(3)}$$

Where:

- **X** represents the complete normalised analytical data matrix.
- $\boldsymbol{n}$ is the number of participating organisations ($n = 52$).
- **m** is the total number of scored competency indicators (m=43).
- $\boldsymbol{x_{ij}}$ denotes the normalised score obtained by organisation i for competency indicator j.

No missing scored responses were identified in the 52 questionnaires included in the final analytical dataset; therefore, no imputation procedure was required.

Following data validation, the normalised competency indicators were grouped into the four analytical dimensions defined in the research instrument. For each participating organisation, a dimension-level score was calculated as the arithmetic mean of the indicators assigned to that dimension. The resulting dimension-level score was then compared with the predefined high-demand threshold. The automated classification criterion implemented in Python was defined as follows:

$$H_{id} = I(\bar{x}_{id} \geq 4) \quad \textbf{(4)}$$

- Where:
  $\boldsymbol{H_{id}}$ represents the binary high-demand classification assigned to organisation $i$for analytical dimension d.
- $\boldsymbol{I}(\cdot)$denotes the indicator function**.**
- $\bar{\boldsymbol{x}}_{\boldsymbol{id}}$ represents the arithmetic mean of the normalised competency indicators included in dimension $d$for organisation
- $\boldsymbol{i}$**.** A value of 1 indicates High Demand, whereas a value of 0 indicates Medium or Low Priority.

For each analytical dimension, the high-demand frequency corresponded to the number of organisations with a dimension-level score equal to or greater than 4. The high-demand percentage was calculated by dividing this frequency by the 52 participating organisations and multiplying the result by 100. The overall mean and standard deviation of each dimension were also calculated across the 52 organisations.

This automated decision rule enabled the computational prioritisation of the four competency dimensions used in the subsequent curricular-mapping process.

3.Loading Phase. Following the validation of the processed dataset, the normalised employer data were transferred to the loading stage of the ETL pipeline. During this phase, the cleaned employer data were loaded into source-specific descriptive matrices and summary tables to calculate the dimension-level means, standard deviations, high-demand frequencies and percentages presented in the results section. In parallel, the historical institutional enrolment records were organised by academic term to generate descriptive summaries and graphical representations of the observed enrolment patterns. These records were used as contextual evidence for curriculum planning and not to estimate dropout rates or demonstrate infrastructure, teaching-staff availability or operational capacity. The outputs generated during this stage constituted the analytical basis for the statistical results and data visualisations presented in the following section.

2.3 Research Instrument and Reliability Assessment

The questionnaire was collaboratively developed by members of the Academic Committee of the Department of Economic and Administrative Sciences at the Technological Institute of Pachuca. Its content was reviewed and approved by the Academic Committee before institutional application. The instrument was developed specifically for this institutional needs assessment and was not adapted from a previously validated scale. The Academic Committee review constituted an institutional content review rather than a formal psychometric validation. The instrument comprised 32 questions designed to collect information on organisational characteristics, administrative needs, preferred areas of professional specialisation and employers' perceptions of the competencies required by graduates of the Bachelor's Degree in Administration programme.

Of the 32 questions, 25 questions (Questions 6–30) were associated with the four analytical competency dimensions and generated 43 scored indicators because several matrix questions contained multiple sub-items. Business Intelligence and Advanced Analytics comprised seven questions and 12 scored indicators; Process Automation with Artificial Intelligence comprised six questions and six indicators; Innovation Models and Digital Environments comprised six questions and eight indicators; and Strategic Management and Digital Leadership comprised six questions and 17 indicators.

The remaining seven questions (Questions 1–5 and 31–32) addressed organisational characterisation, the organisational areas in which administrators worked, institutional linkage and hiring intention, and open-ended prospective recommendations. These questions were treated separately as descriptive, institutional-linkage, hiring-intention or qualitative variables, as appropriate, and were not included in the dimension-level scores or the internal-consistency assessment.

The questionnaire combined free-text questions, categorical multiple-choice questions and ordinal competency-assessment formats with three, four or five response categories, depending on the content and purpose of each question. Within the instrument, eight questions (Questions 11, 12, 17, 18, 23, 24, 29 and 30) were specifically developed by the Academic Committee to assess emerging competencies related to heterogeneous data architectures, interactive dashboards, prompt engineering, process automation, agile methodologies, digital commerce, information governance, and strategic project and risk management. Before calculating the dimension-level scores, the 43 scored indicators were proportionally transformed to the common 1-to-5 scale described in Equation (2). Higher numerical values represented greater perceived importance, relevance or organisational demand. All 52 questionnaires included in the final analytical dataset were complete; therefore, the reliability and dimension-level analyses were conducted using complete responses.

To assess the overall internal consistency of the 43 scored competency indicators, Cronbach's alpha coefficient was calculated after normalisation and before dimension-level aggregation. The seven questions that did not form part of the four analytical dimensions were excluded from this calculation. The coefficient was computed using the following expression:

$$\boldsymbol{\alpha} = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_T^2}\right) \quad \textbf{(5)}$$

Where:

- α is Cronbach's alpha coefficient.
- k denotes the number of scored competency indicators included in the reliability analysis (k=43).
- σ_i^2 represents the variance of competency indicator i.
- σ_T^2 represents the variance of the total score obtained from the 43 scored competency indicators.

The automated computation yielded a Cronbach's alpha coefficient of 0.89, indicating high overall internal consistency among the scored indicators included in the analytical matrix. Because the instrument comprised four analytically distinct competency dimensions, this coefficient was interpreted as evidence of consistency within the combined scored matrix and not as proof of unidimensionality, construct validity or validity of the questionnaire as a whole.

2.4 Computational Environment and Software Tools.

The data-processing pipeline was implemented in Python 3.10 within the Jupyter Notebook computational environment. Data obtained from the digital questionnaire platform were exported in CSV format and programmatically ingested into the computational environment. Structured data processing, validation, construction of frequency matrices and descriptive summaries were performed using the Pandas library (v1.5). The internal-consistency assessment using Cronbach's alpha and the exploratory homogeneity-of-variance tests were conducted using the Pingouin library (v0.5). Institutional enrolment records were organised and processed using NumPy arrays (v1.23), while the final data visualisations were generated using Seaborn (v0.12). Figure 1 summarises the ETL-based data pipeline implemented in Python for the source-specific processing, validation and analysis of the employer and institutional datasets used in curriculum planning.

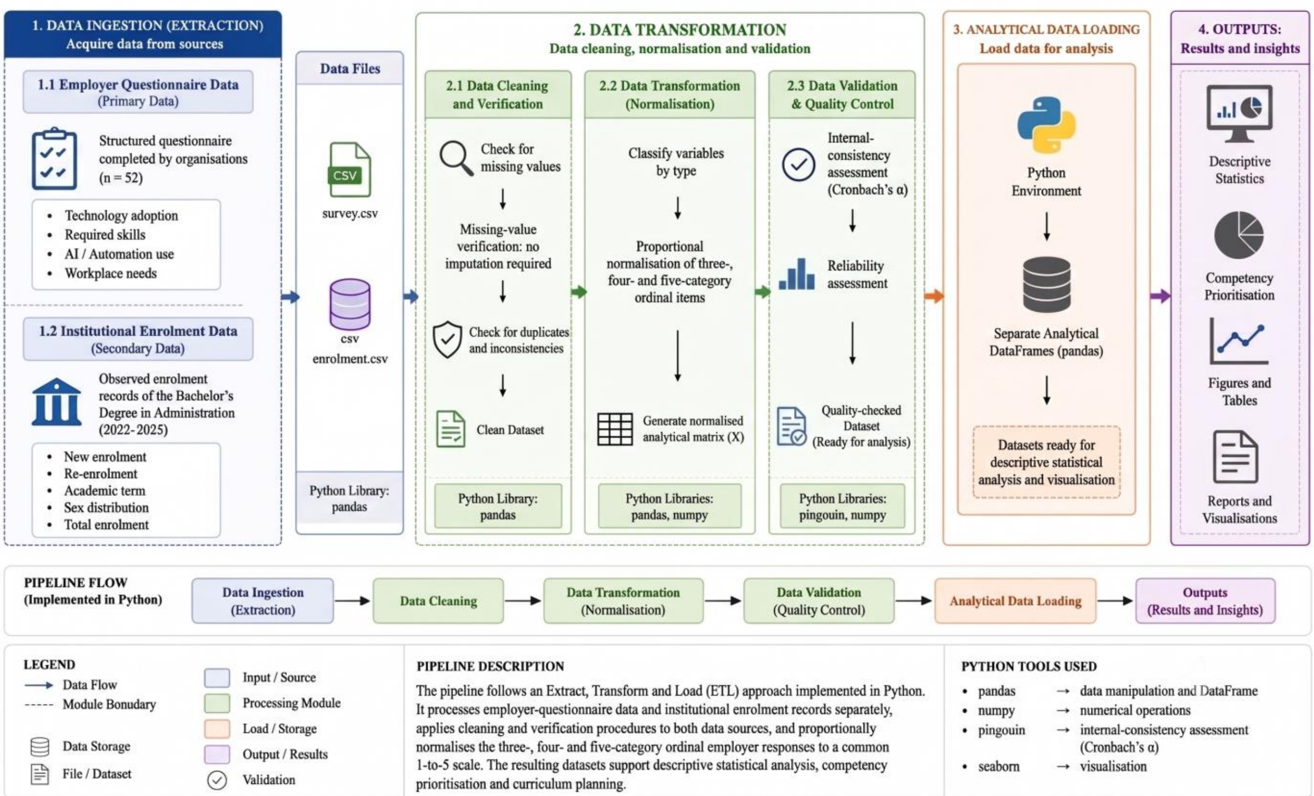

Fig. 1. ETL-based data pipeline implemented in Python for evidence-informed curriculum planning

3 Results and Discussion

The results obtained through the data pipeline are presented in two complementary components: the technological competency priorities identified among the participating employers and the institutional enrolment context of the Technological Institute of Pachuca. These components were analysed separately and subsequently considered during the curricular-mapping process.

3.1 Analysis of the Needs Detected by Employers through the Data Pipeline

The employer-questionnaire dataset was processed using the ETL pipeline implemented in Python. The 43 normalised competency indicators were aggregated into the four analytical dimensions defined in the research instrument. For each dimension, the mean, standard deviation, high-demand frequency and high-demand percentage were calculated across the 52 participating organisations. The resulting dimension-level priorities are presented in Table 2.

Table 2. Dimension-Level Employer Demand for Technological Competencies

Dimension Analysed	Mean (μ)	Standard Deviation (σ)	High-Demand Frequency (Dimension-Level Score ≥ 4)	High-Demand Percentage
Business Intelligence and Advanced Analytics	4.62	0.48	46	88.4%
Process Automation with AI	4.48	0.52	44	84.6%
Innovation Models and Digital Environments	4.26	0.61	42	80.7%
Strategic Management and Digital Leadership	4.12	0.73	37	71.1%

Source: Authors' elaboration based on the dimension-level output generated in Jupyter Notebook using the Pandas library.

As shown in Table 2, Business Intelligence and Advanced Analytics recorded the highest mean score ($\mu = 4.62$) and the highest high-demand percentage, with 46 of the 52 participating organisations (88.4%) obtaining a dimension-level score equal to or greater than 4. Process Automation with AI ranked second, with a mean of 4.48 and 44 organisations (84.6%) classified within the high-demand category. Innovation Models and Digital Environments and Strategic Management and Digital Leadership also recorded mean scores above the predefined threshold, although their high-demand percentages were comparatively lower. Figure 2 presents the high-demand percentages obtained for the four analytical dimensions.

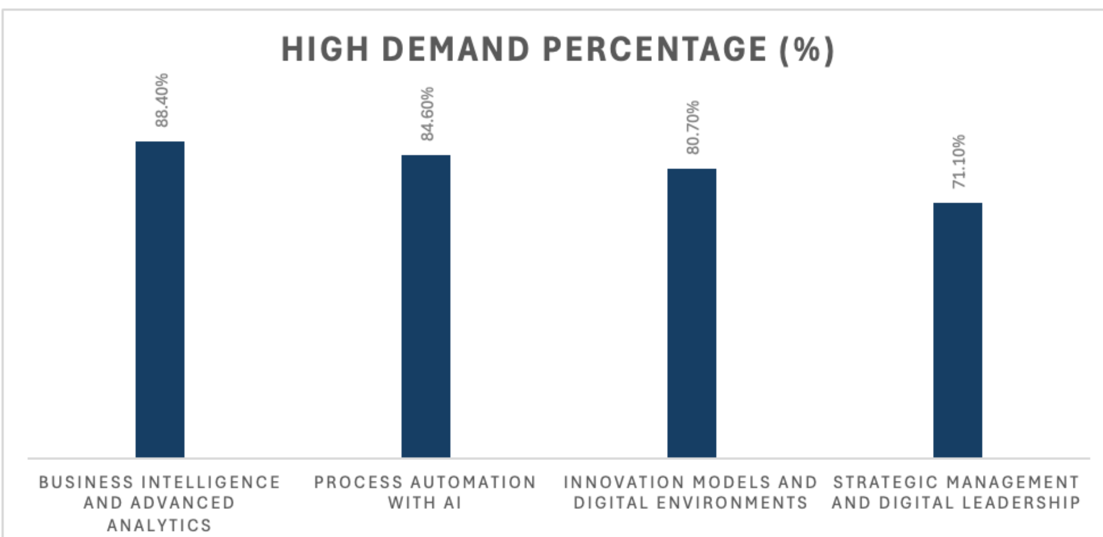

Fig. 2. High-demand percentages for the four technological competency dimensions among the participating organisations

The exploratory analysis of the dimension-level scores by organisational size yielded $p > 0.05$ in the homogeneity-of-variance tests conducted using the Pingouin library. These results indicate insufficient evidence to reject the null hypothesis of equal variances across the organisational-size groups; however, they do not demonstrate that competency requirements were identical or statistically equivalent among micro, small, medium-sized and large organisations. The overall percentages indicate that the four competency dimensions were relevant among the participating organisations, although equivalence between organisational-size groups cannot be concluded from the non-significant results.

3.2 Descriptive Analysis of Institutional Enrolment

The computational analysis of the historical enrolment records for the Bachelor's Degree in Administration at the Technological Institute of Pachuca was performed using NumPy vectors to describe the observed enrolment pattern during the 2022–2025 period. These data were used as contextual evidence regarding the size and aggregate enrolment pattern of the student population considered in planning the proposed specialisation. The analysis was descriptive and was not intended to evaluate infrastructure availability, teaching-staff sufficiency, student retention or institutional operational capacity. The distribution of the active student population by academic term is presented in Table 3.

Table 3. Observed student enrolment by academic term and sex in the Bachelor's Degree in Administration at the Technological Institute of Pachuca (2022–2025)

Academic Term	New Entry (H)	New Entry (M)	Re-entry (H)	Re-entry (M)	Total Enrolment (N)
Jan–Jun 2022	9	19	136	231	395
Aug-Dec 2022	12	38	124	215	389
Jan–Jun 2023	4	21	125	230	380
Aug-Dec 2023	20	45	118	220	403
Jan–Jun 2024	3	11	143	238	395
Aug-Dec 2024	14	42	126	223	405
Jan–Jun 2025	6	13	143	237	399

Source: Authors' elaboration based on institutional records provided by the Department of Economic and Administrative Sciences (Tecnológico Nacional de México, Instituto Tecnológico de Pachuca, 2026).

Across the seven observed academic terms, total enrolment ranged from 380 to 405 students, with an arithmetic mean of 395.1 students per term. These values indicate relatively limited variation in the observed enrolment totals during the analysed period. The academic-term distribution is presented graphically in Fig. 3.

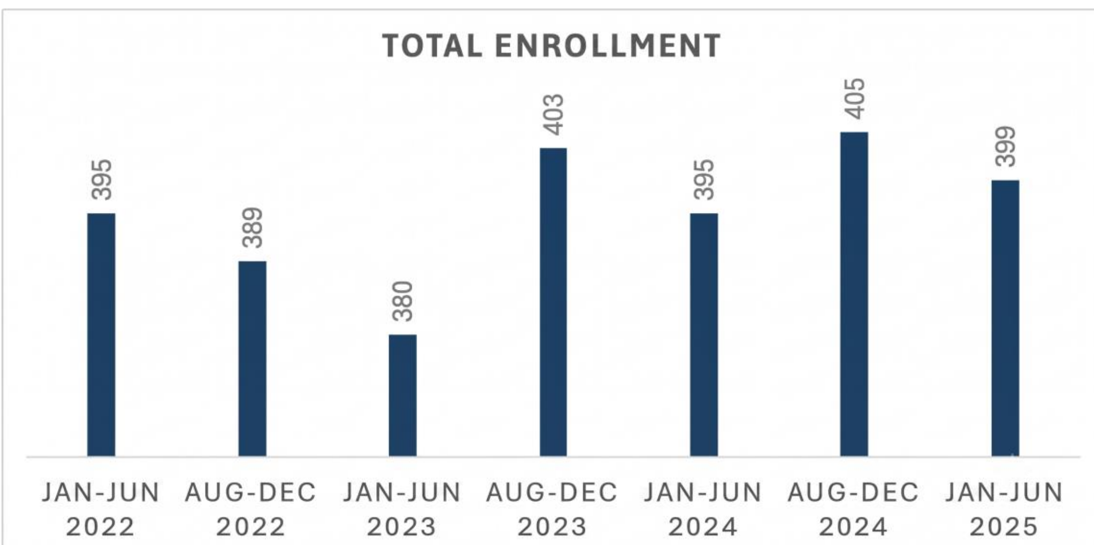

Fig. 3. Observed total enrolment in the Bachelor's Degree in Administration at the Technological Institute of Pachuca, 2022–2025

The observed enrolment series provides contextual information for academic planning by showing relatively limited variation in aggregate enrolment during the analysed period. However, aggregate enrolment totals alone do not demonstrate infrastructure availability, teaching-staff sufficiency or the operational capacity required to implement the proposed specialisation. No dropout rate was estimated from these records because the available academic-term totals do not constitute a cohort-based attrition analysis. Consequently, the results are interpreted exclusively as descriptive institutional context for curriculum planning.

3.3 Evidence-Informed Curricular Structuring and Scientific Discussion

Based on the sequential consideration of the employer competency requirements presented in Table 2 and the institutional enrolment context presented in Table 3, the Academic Committee structured the Specialisation in Digital Management and Business Intelligence (key LADE-GDI-2026-01), incorporated into the LADM-2010-234 curriculum. Employer evidence determined the prioritised competency areas, whereas institutional enrolment records provided contextual information for academic planning. The resulting five-course structure was organised according to the institutional curricular guidelines of the Tecnológico Nacional de México and was subsequently officially registered for implementation beginning in August 2026 (TecNM, 2026). This registration is reported as evidence of institutional adoption of the curricular proposal and not as scientific validation of the analytical framework. The curricular structure is presented in Table 4.

Table 4. Curricular Structure of the Specialisation and Alignment with the Competency Dimensions

Subject	Key	Primary Competency Linkage	Content and Focus of the Subject
Professional Skills Workshop	GDD-26-01	Strategic Management and Digital Leadership; cross-cutting ethical and governance competencies	Information governance models, ethical guidelines in the use of technology and organisational data protection.
Fundamentals of AI Applied to Business	GDD-26-02	Process Automation with Artificial Intelligence	Strategic use of virtual assistants, development of efficient instructions (*Prompt Engineering*) and automation of administrative workflows.
Agile Methodologies for Project Management	GDD-26-03	Innovation Models and Digital Environments	Implementation of collaborative frameworks (*Scrum* and *Kanban*) to coordinate technological projects in organisations.
Business Strategy and Market Analytics	GDD-26-04	Business Intelligence and Advanced Analytics	Application of explanatory statistical analysis, market segmentation and evaluation of trends in the commercial environment.
Business Strategy Management Tools	GDD-26-05	Business Intelligence and Advanced Analytics; Strategic Management and Digital Leadership	Structuring of institutional data, design of interactive dashboards and monitoring of critical indicators (KPIs).

Source: Authors' elaboration in accordance with the curricular guidelines of the Tecnológico Nacional de México.

The four analytical competency dimensions were not translated into courses through a one-to-one correspondence. The Academic Committee used conceptual affinity, progression of learning, complementarity and avoidance of content duplication as curricular-mapping criteria. Business Intelligence and Advanced Analytics was represented across two courses because market analysis and the design of managerial dashboards involve different learning outcomes and applications. Process Automation with Artificial Intelligence and Innovation Models and Digital Environments were each associated with a directly corresponding course. Strategic Management and Digital Leadership was primarily addressed through Professional Skills Workshop and reinforced through Business Strategy Management Tools. Professional Skills Workshop also incorporated cross-cutting content related to information governance, ethics and organisational data protection. This curricular-mapping process explains why the four employer-demand dimensions resulted in a five-course specialisation.

The employer priorities identified in this study are consistent with previous research indicating that the adoption of Artificial Intelligence increases the importance of management, business, digital and data-interpretation skills in occupations exposed to technological change (Green, 2024; McAfee & Brynjolfsson, 2012; OECD, 2023). They are also aligned with international competency frameworks that emphasise the need to integrate applied AI knowledge with human-centred decision-making, ethics and responsible technological use (UNESCO, 2024). Within the proposed specialisation, these global priorities are reflected in the combination of analytical, automation, leadership, governance and project-management competencies.

Unlike large-scale studies that use online job-posting mining to identify changes in labour-market demand, the present study used direct responses from regional organisations linked to the Technological Institute of Pachuca (Manca, 2023). This approach provides context-specific evidence for curriculum planning, although its findings cannot be generalised to the entire business population of Hidalgo or to other regions. The scientific contribution therefore lies not in proposing a universally applicable course structure, but in demonstrating a sequential procedure through which employer evidence, institutional context and academic judgement can be systematically related to support curricular decision-making. Future applications could complement employer questionnaires with graduate tracking or job-posting analysis to examine whether the identified competency priorities remain stable over time.

4 Conclusions

This study indicates that the integration of Artificial Intelligence and data analysis into management education is consistent with the competency priorities identified among the participating employers. The computational processing of employer responses identified the need to strengthen competencies related to Business Intelligence, process automation, data analysis and digital decision-making, providing context-specific regional evidence for curriculum planning. The ETL-based data pipeline implemented in Python provided a structured, traceable and reproducible methodological approach for relating employer-demand evidence to institutional enrolment context through a sequential evidence-integration procedure. Through source-specific data ingestion, cleaning, validation, proportional normalisation, dimension-level aggregation and descriptive analysis, the proposed architecture systematised procedures that are frequently conducted manually in curriculum-planning processes and supported an evidence-informed approach to curricular structuring.

Based on the sequential consideration of the processed evidence and the curricular mapping conducted by the Academic Committee, a five-course specialisation entitled Digital Management and Business Intelligence was designed. Historical institutional enrolment records provided descriptive contextual information about the size and aggregate enrolment pattern of the target student population; however, these records did not by themselves demonstrate infrastructure availability, teaching-staff sufficiency or institutional operational capacity. Following the institutional technical evaluation, the specialisation was officially registered by the Tecnológico Nacional de México for implementation in the Bachelor's Degree in Administration at the Technological Institute of Pachuca beginning in August 2026. This registration is reported as evidence of the practical institutional adoption of the curricular proposal and not as scientific validation of the analytical framework.

The principal contribution of this research lies in presenting a transparent and reproducible methodological architecture through which employer competency evidence, institutional context and academic judgement can be systematically related to support curricular decision-making. The contribution does not reside in the development of a new ETL algorithm, predictive model or standalone decision-support application, but in the traceable linkage between data preparation, competency prioritisation and curricular mapping. Employer evidence determined the prioritised competency areas, whereas institutional enrolment records provided contextual information and did not modify the competency scores or their ranking. Because the employer sample was limited to organisations belonging to the Institute's collaboration and professional-residency network in Hidalgo, the findings should not be generalised to the entire business population of the state or to other regions. The possibility of non-response bias associated with the eight eligible organisations that did not submit a completed questionnaire should also be considered when interpreting the results.

Although the study was conducted within a regional and institutional context, the proposed architecture may be adapted by other higher education institutions seeking to modernise academic programmes, provided that their local sampling frames, institutional conditions and curricular regulations are considered. Future research should evaluate the outcomes of the specialisation after its implementation, particularly student competency development, graduate employability and employer satisfaction. Subsequent studies could also compare employer-survey findings with graduate tracking, curriculum evaluation and job-posting analysis to determine whether the identified competency priorities remain stable over time.

References

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Chu, T. S., & Ashraf, M. (2025). Artificial intelligence in curriculum design: A data-driven approach to higher education innovation. *Knowledge, 5*(3), Article 14. https://doi.org/10.3390/knowledge5030014

Consejo de Ciencia, Tecnología e Innovación de Hidalgo. (2024). *Innovación empresarial y habilidades emergentes para la transformación digital en Hidalgo* [Informe institucional interno]. Gobierno del Estado de Hidalgo.

Davenport, T. H., & Mittal, N. (2023). *All-in on AI: How smart companies win big with artificial intelligence*. Harvard Business Review Press.

Green, A. (2024). *Artificial intelligence and the changing demand for skills in the labour market* (OECD Artificial Intelligence Papers No. 14). OECD Publishing. https://doi.org/10.1787/88684e36-en

Hernández Sampieri, R., & Mendoza Torres, C. P. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill Education.

Manca, F. (2023). *Six questions about the demand for artificial intelligence skills in labour markets* (OECD Social, Employment and Migration Working Papers No. 286). OECD Publishing. https://doi.org/10.1787/ac1bebf0-en

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68. Harvard Business Review article

Organisation for Economic Co-operation and Development. (2023). *OECD employment outlook 2023: Artificial intelligence and the labour market*. OECD Publishing. https://doi.org/10.1787/08785bba-en

Sousa, M. J., & Rocha, Á. (2019). Skills for disruptive digital business. *Journal of Business Research, 94*, 257–263. https://doi.org/10.1016/j.jbusres.2017.12.051

Tecnológico Nacional de México. (2015). *Manual de lineamientos académico-administrativos del Tecnológico Nacional de México*. https://www.tecnm.mx/normateca/Direcci%C3%B3n%20de%20Docencia%20e%20Innovaci%C3%B3n%20Educativa/Manual%20Lineamientos%20TecNM%202015/Manual_de_Lineamientos_TecNM.pdf

Tecnológico Nacional de México. (2026, 14 de julio). *Registro de la especialidad Gestión Digital e Inteligencia de Negocios, clave LADE-GDI-2026-01* (Oficio No. M00.2/1386/2026) [Documento oficial de registro].

Instituto Tecnológico de Pachuca. (2026). *Índices académicos del Departamento de Ciencias Económico Administrativas* [Informe institucional interno].

UNESCO. (2024). *AI competency framework for students*. https://doi.org/10.54675/JKJB9835

Because UNESCO is both author and publisher, APA 7 does **not** require repeating "UNESCO" as publisher.

Vassiliadis, P. (2009). A survey of extract–transform–load technology. *International Journal of Data Warehousing and Mining, 5*(3), 1–27. https://doi.org/10.4018/jdwm.2009070101

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior, 89*, 98–110. https://doi.org/10.1016/j.chb.2018.07.027

World Economic Forum. (2025). *The future of jobs report 2025*. World Economic Forum — The Future of Jobs Report 2025