www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 18(1), 2027, 418-433. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2027.v18i1.1327

Application of Hierarchical Agglomerative Clustering for the Segmentation of Startups in Latin America

Laura Angélica Décaro Santiago[1], Asdrúbal López Chau*[1], María Guadalupe Soriano Hernández[1], Maria del Carmen Messina Scolaro[2]

[1] Universidad Autónoma del Estado de México, CU UAEM Zumpango

[2] Universidad de la República, Uruguay

Emails: ladecaros@uaemex.mx , alchau@uaemex.mx, mgsorianoh@uaemex.mx, maria.messina@fcea.edu.uy

Abstract. The number of startups in Latin America has been growing quickly in the last few years, generating valuable data that has not been analyzed before in detail. In this paper, we collected data from 1,274 Latin American startups and tested three clustering methods to determine the best among them to analyze this data.
This approach, combined with a statistical analysis, allowed for the identification of common patterns and significant differences, which facilitated the establishment of profiles that help understand business growth and consolidation processes. The findings allow for the modeling of the transitions that occur in startups: initiation, growth, and expansion.
.

Keywords: Unsupervised learning, Hierarchical Clustering, Latin American Startups, Data Mining.

Article Info
Received: Feb 3, 2026
Accepted: Aug 11, 2026

1 Introduction

The increasing rise of startups in Latin America (LATAM) has generated large volumes of heterogeneous data, including text, temporal data, numerical and categorical information. The automatic extraction of patterns from this data is a strategic component for understanding market dynamics in a globalized world. Due to the startup's DNA, its life cycle differs from traditional companies because of the close parallelism with financing (Angel Ventures, 2022; Escartín et al., 2020; Mageste et al., 2024; Peixoto et al., 2023; Solís Tepexpa & Torres Castañeda, 2022), under the understanding that the survival and development of the startup cannot dispense with the injection of capital (Chakraborty et al., 2025).

Various investigations have provided evidence regarding the development process of startups through case studies, such as in Kuusinen et al. (2018), García et al. (2020), Caro et al. (2020), among others. Indisputably, this type of research is difficult for 1,274 LATAM startups, which constitute the sample of the present study. Therefore, it has been chosen to link variables related to age, financing, and digital visibility to detect clusters according to the life stage. Although age and financing are some indicators of success (Chakraborty et al., 2025; Díaz-Santamaría & Bulchand-Gidumal, 2021), including the number of followers responds to the demand for rapid growth, where visibility and audience increase become crucial, and social networks a valuable tool to achieve it.

Much has been written in scientific articles about startups originating from developed countries and regions (to name a few: United States (Banerji & Reimer, 2019; Ewens & Malenko, 2020; Wei et al., 2025), European Union (Bethlendi et al., 2025; Díaz-Santamaría & Bulchand-Gidumal, 2021; Greco, 2025; Kézai & Skala, 2024; Linina et al., 2021; Maarouf et al., 2025; Martielli et al., 2025), and Asia (Daradkeh & Mansoor, 2023; Ressin, 2022; Shabtai et al., 2025; Usman & Sun, 2023; Vongvisitsin & Sun Tung, 2025)), contrary to emerging regions like LATAM, especially using a sample with a significant number of startups from the region.

LATAM finds itself in a context of significant challenges in terms of entrepreneurship; from the persistent difficulty in accessing financing and a limited adoption of digital technologies; to weak institutionality —in contrast to more developed countries—. A situation that demands a more transparent and efficient institutional environment that supports the efforts of the ecosystem actors (Álvarez et al., 2014; Saavedra Leyva & Texis Flores, 2019), as well as security measures and effective actions that incentivize entrepreneurial initiative (Martínez Sidón et al., 2025). Despite these challenges, the region, which houses approximately 8.7% of the world's population and maintains important cultural connections among its member countries, has consolidated itself as an attractive territory for Venture Capital (VC) investment (Angel Ventures, 2022), generating a significant number of startups (BID, 2024).

Likewise, the analysis process identifies correlations and underlying structures that enrich the knowledge framework regarding startup dynamics in LATAM.

The research question is the following: Do age, financing, and digital visibility possess the necessary discriminant capacity to model representative clusters of LATAM startups according to their life cycle? Based on this question, the analysis also addresses two additional interrogatives: What is the degree of statistical correlation between these dimensions? and What multidimensional differentiation patterns emerge among the identified segments in the region? A double purpose is proposed: to validate the efficacy of these variables for the configuration of profiles through unsupervised learning and to identify the differentiating vectors within the sample.

The novelty of this research materializes in the approach to the LATAM startup from a large-scale quantitative and comparative perspective, employing statistical techniques and Machine Learning for pattern identification. While most studies concentrate on developed contexts —particularly in North America, Europe, and East Asia— this work tends to cover part of the scientific vacuum surrounding startups in LATAM, whose ecosystems present outstanding challenges in terms of financing, institutionality, and digital services. Additionally, as this research advances in the characterization of the life cycle, it also proposes a replicable methodological framework for the comparative analysis of emerging business ecosystems.

The rest of the article is organized as follows. The first section consists of the theoretical framework regarding startup attributes, the link between financing and the life cycle, the role of visibility on social networks, as well as the contextual framework of LATAM. The second section corresponds to the methodology that describes data collection, variables, and analysis procedures. Thirdly, the results of the statistical and cluster analysis are presented. Finally, the discussion and conclusions regarding the identified groups and the main significant differences and correlations are provided.

2 Theoretical and contextual framework

2.1. The Startup DNA

The history of the startup dates back to the nineties of the last century, and time has transformed its conception. Over the decades, the definition of the startup has matured (marked by global events). Currently, the conceptualization of the startup concentrates on a group of attributes: young, fast-growing, disruptive in its business model, avant-garde in the use and development of technology and/or knowledge, and aggressive in assuming high risk.

Under the attribute of youth, there is no homogeneous agreement; however, in Greco (2025) the startup is referred to as a temporary organization, that is, the company has a successful exit stage to stop being a startup, through an Initial Public Offering (IPO), merger or acquisition (Angel Ventures, 2022; Escartín et al., 2020).

Regarding rapid growth, Chakraborty et al. (2025) refer to the Organization for Economic Cooperation and Development (OECD) to typify it after a growth of 20% annually for three years with 10 or more employees at the start. While rapid growth goes hand in hand with scalability, it is influenced by the level of innovation in the business model, where significantly, the collaborative economy makes the difference for the creation and delivery of the value proposition.

The startup is "…the vehicle through which the Industrial Revolution 4.0 is being structured." (Escartín et al., 2020, p. 14). In the era of technology and knowledge, these are notable ingredients for product or process innovation. Innovation differentiated between: a) Exploitative innovation that improves, adapts, and extends existing technologies to systematize or

make products or processes more efficient, and, b) Exploratory innovation that discovers new ideas, technologies, and resources in a radical way (Clauss et al., 2021; Daradkeh & Mansoor, 2023).

Finally, the characteristic of non-risk aversion. This position is assumed from the break with conventionalism, undertaking with disruption, opaque information, and a scenario of uncertainty that the entrepreneur accepts because of the competitive advantage that the authenticity of innovation grants. “Investing in innovation and new technologies is a high-risk activity…”, mention Mageste et al. (2024, p. 16).

2.2. Financing and Life Cycle

The organizational life cycle theory originates in the fifties of the last century through its first conceptualization and later by the appearance of various models that describe the cycle of a typical organization, divided into stages (Ionescu & Negrusa, 2007). In contrast, the startup phases are navigated differently. The report “Empresas emergentes (start-ups) en América Latina y el Caribe” from CEPAL points out how financing sources align with each phase of the life cycle (Mageste et al., 2024).

In each stage, certain financing instruments and actors prevail. For their part, in Solís Tepexpa & Torres Castañeda (2022) and Escartín et al. (2020) a final successful exit stage is included. Meanwhile, in Malekmohammadi et al. (2024), the last stage is designated as decline or renewal.

It should be noted that the close relationship between life stages, intrinsic objectives, actors, instruments, and financial amounts is due to their interdependence for survival and consolidation (Peixoto et al., 2023; Pons & Boar, 2020; Shabtai et al., 2025; Wei et al., 2025). In this same line of thought, Maarouf et al. (2025) assume that securing financing is indicative of success in the startup, which is logical, since the crystallized fundraising represents a considerable victory. First, because investors risk their capital and back the proposal that was previously evaluated and validated in terms of market, technological domain, startup management, personality, experience, and founders' capability (Bethlendi et al., 2025; Endeavor & Glisco Partners, 2024). And second, because it requires time and dedication (the process can last from six months to a year (Pons & Boar, 2020)).

Generally, the risk level starts high and then declines, especially because for the launch, the startup has a limited track record and assets, relying only on a minimum viable product, data on market conditions, and information about the founding team (Martielli et al., 2025). Conversely, in the future, the elements for evaluation are more complete (Berre & Le Pendeven, 2023). Additionally, in the transition between stages, the control of the company also changes (the founding entrepreneur shares it incrementally with the VC investors as investment rounds progress, Ewens & Malenko, 2020)).

2.3. Visibility through social networks

Eloquently, much has been written about the importance of social connections and their virtue in providing access to tangible and intangible resources. Banerji and Reimer (2019) evidenced the strong connection between the average number of followers of the founders on the LinkedIn network and fundraising. Gloor et al. (2020) show the relationship between financing traction and the active presence on the social network of members of the board of directors. In the same way, Vongvisitsin and Sun Tung (2025) point out how the development of the founders' social network capital, in addition to facilitating the mobilization of financing and resources, also promotes team formation, alliances, and builds credibility for the project.

While the cited research uses the number of followers as an indicator to quantify the relational capital of the founders, the perspective of this study analyzes it as a factor that allows visibility for exponential growth; since communication through digital media is a potential and economical tool for young companies. In a case study, in Kuusinen et al. (2018), the use of Google AdWords and Facebook Ads provided customers to the startup in its early stages.

Likewise, Peixoto et al. (2023) point out the role of social networks in the growth of the startup, whose content changes as it grows. The authors highlight how a small investment in social networks facilitates connection with prospects and customers, data collection for decision-making, as well as obtaining funds. Additionally, Usman and Sun (2023) find significant evidence of the network effect (positive effect of the value of the proposal as users increase) on the performance of the startup.

2.4. LATAM and the startup ecosystem

Despite great economic and social phenomena facing LATAM, such as high inflation in Argentina, restrictive monetary policy in Brazil, social and political uncertainty in Peru, or the closure of institutions in Mexico (OECD et al., 2024; Startupblink, 2024), LATAM is an attractive region for transferring invention with disruptive business models; therefore, it is imperative to investigate what levels of development Latin startups maintain.

With a population exceeding 657 million, a labor force of 48.29%, and a GDP per capita of 10,796.9 current dollars (World Bank, n.d.), Brazil, Chile, Mexico, Colombia, Argentina, Uruguay, Peru, Costa Rica, Ecuador, Panama, and Jamaica are part of the top 100 best ecosystems (Startupslatam, 2025). Furthermore, (Startupblink, 2024) reports that Brazil, Colombia, Chile, Mexico, and Argentina are positioned in the top 50 places.

Brazil, besides being the country with the largest number of inhabitants in LATAM, is a leader and provider of ten top cities (Startupblink, 2024). The OECD reveals it as the only one in the region with a specific ministry for SMEs (OECD et al., 2024). Meanwhile, the Global Startup Ecosystem Index places Colombia in the second position in LATAM with its nine top cities. Among its main strategies: simplified financial licenses for institutions in the Fintech sector and chambers of commerce for non-traditional industries (Startupblink, 2024). In the third position as the best startup ecosystem and with two top cities, is Chile. A country that has employed interesting mechanisms to attract and retain foreign entrepreneurial talent through Startup-Chile; additionally, it holds the first place in the Insuretech sector of the region (Startupblink, 2024). Mexico is the second most populated country in LATAM; with 17 top cities, it enjoys a strategic location between North America and South America. Mexico has the Law of Financial Technology Institutions that strengthens the Fintech ecosystem (Startupblink, 2024). On the other hand, the worldwide leadership in the export of agricultural products that offers opportunities in the Agtech sector and the large supply of human capital with STEM careers are strengths of Argentina.

Additionally, some general data of the region are: a) Fintech is the leading sector (Startupslatam, 2025), although the AI development sector has begun to gain ground (Endeavor & Glisco Partners, 2024); b) the average ticket of capital raised is 1M USD (Latam Republic, 2024). LATAM VC investors contributed an average of 1.1M, 11.5M, 18.3M, 42.5M, and 242.7M in seed and angel, Series A, Series B, Series C, Series D (onwards), respectively (Secretaría General Iberoamericana [SGIA], 2020), while for the year 2024 in early stage, average capital of 4.1M is raised, and in growth and maturity stages, 36.5M (Endeavor & Glisco Partners, 2024); c) 57% of Latin startups have two founders and only 34% have female founders (OECD et al., 2024); d) the main contextual challenges for the region are the operating environment and the difficulty of processes, as well as access to financing (OECD et al., 2024); e) knowledge about the mechanisms of the Venture Capital sector in LATAM is no more than ten years old (Rockstart, 2023); and, g) The region has significant challenges in access to public and financial services (Angel Ventures, 2022).

3. Methodology

3.1. Data collection

Data collection was carried out through queries on various web platforms specialized in startup information. Among these, the Startupeable platform (Startupeable, 2025) proved to be the source with the largest amount of available and free-access data. This platform provides information on 1,807 startups founded from 1972 to 2023, including country of origin (PAIS), foundation year (AÑO_FUNDACIÓN), financing level (ULTIMA_RONDA_INVERSIÓN), invested amounts (CAPITAL_LEVANTADO), investors (INVERSIONISTAS), social networks (FACEBOOK/ X (TWITTER)/ INSTAGRAM/ YOUTUBE/ LINKEDINN/ OTRA), among other data not used in this study. The collection process took place between November 2024 and March 2025. It was decided to use only the data from those startups that were in operation at the time of collection. To determine their validity, it was verified that the official website of the startup was active during the analysis period.

3.2. Variables

The data for the variables PAÍS and CAPITAL_LEVANTADO were recorded in the original format of the Startupeable platform (Startupeable, 2025). The data for the variable ULTIMA_RONDA_INVERSIÓN were categorized into four groups:

(0) No round, (1) Before series A, (2) Series A, and (3) After series A. Meanwhile, the data for the three remaining variables were organized as shown in Table 1. Some specifications of the variables are shown below:

With the exception of CAPITAL_LEVANTADO and NUM_SEGUIDORES, the variables are categorical.

The variables CAPITAL_LEVANTADO, ULTIMA_RONDA_INVERSION, EDAD, NUM_INV, and NUM_SEGUIDORES are dynamic variables, whose value may change over time.

For the cluster analysis, the variables CAPITAL_LEVANTADO were grouped as: (0) No funding raised, (1) Before series A, (2) Series A, and (3) After series A, considering the amounts from SGIA (2020); likewise, NUM_SEGUIDORES is classified into four quartiles.

Table 1. Description of the transformed variables.

Variable	Transformed variable	Description
AÑO_FUNDACION	EDAD	The year of foundation is used to calculate the age of the startup, and it is classified into three age blocks: (1) $x \leq 6$, (2) $7 \geq x \geq 14$ and (3) $x \geq 15$.
INVERSIONISTAS	NUM_INV	Count of the number of investors mentioned and grouped into one of four categories: (0) No investors, (1) One, (2) Two, (3) More than two.
FACEBOOK, X (TWITTER), INSTAGRAM, YOUTUBE, LINKEDINN, OTRA	NUM_SEGUIDORES	

4.3. Data Processing

Since the data collection process was carried out manually, some records presented inconsistencies or errors, making it necessary to apply a cleaning procedure.

- Unknown values were coded as NN, while the category NO was used to indicate the actual absence of an attribute in a startup.
- The variable AÑO_FUNDACIÓN, which becomes EDAD (age) —originally numerical— was transformed into a variable with three categories (based on the historical chronology of the theoretical framework: founded up to the year 2010, between 2011 and 2018, and from the year 2019 onwards), which originated EDAD in three blocks: $x \leq 6$, $7 \leq x \leq 14$, and $x \geq 15$.

- For 24 startups, the foundation year was not found on the Startupeable platform (Startupeable, 2025), nor on their website or social networks. To avoid the exclusion of these startups, a value was imputed based on the following conditions.
    - If the words ['platform', 'application', 'intelligence', 'artificial', 'algorithm', 'algorithms', 'robot', 'data', 'economy', 'collaborative'] appear in their description, a pseudo-random value greater than 2011 and less than 2025 was assigned. This is based on the premise that LATAM startups began to venture more frequently into these types of technologies and applications starting in 2011.
    - For the remaining startups, pseudo-random values were imputed, considering the following proportions: 10% for $x \leq 6$, 50% for $7 \leq x \leq 14$, and 40% for x $x \geq 15$.

After data processing, a descriptive analysis of each variable was performed.

- For categorical variables, absolute and relative frequencies are highlighted. Concurrently, scale variables report Maximum (Max) and Minimum (Min), as well as Median (Mdn) and Range (Rgo); central and dispersion measures reported after the application of the Kolmogorov-Smirnov normality test.

To identify relationships between variables while respecting their nature, three tests are applied.

- First, the Chi-square statistic to identify the significance of the relationship between categorical variables (Cramer's V test is added).
- Second, the Kruskal-Wallis test to determine significant differences in CAPITAL_LEVANTADO and NUM_SEGUIDORES among groups of categorical variables. Additionally, a post hoc test was integrated (significance values are adjusted using the Bonferroni correction for multiple tests).
- Finally, Spearman's correlation (Rho) for scale variables, also including ordinal categorical variables. Up to this point, the Statistical Package for the Social Sciences software SPSS Version 28.0 is use.
- In addition, to identifying relationships between variables through statistical tests, a cluster analysis was performed with the purpose of discovering groups of startups with similar characteristics.

For this purpose, we apply and test the following clustering algorithms: K-Means, Agglomerative hierarchical clustering algorithm and Gaussian mixture models. We chose the algorithm that gave the best results with the Silhouette index. Other algorithms that create complex shapes were not used (such as DBSCAN), because the interpretation of the identified groups is more difficult for this study.

In addition, we performed an exploration of the optimal number of clusters through visual analysis of the dendrograms generated by the method. From visual observation, the optimal number of partitions (clusters) is determined by identifying the largest vertical jump not crossed by horizontal lines, which indicates a natural separation between groups. Since this algorithm is based on distance calculation, it was necessary to transform categorical variables into numerical format using one-hot encoding, thus ensuring the appropriate representation of the data in a metric space. Once the dendrogram was generated, the number of clusters was confirmed by visual inspection, identifying the most representative cutoff point.

The identified groups were visualized applying the dimensionality reduction technique of Principal Component Analysis (PCA). PCA allows the data to be projected into two dimensions while preserving as much variance as possible. This facilitates the visualization of internal structures, including possible groupings. Upon achieving the identification of the clusters through hierarchical clustering analysis, a qualitative and quantitative interpretation of each group was carried out, with the aim of characterizing the predominant profiles in each one. The interpretation was based on the analysis of the distribution of variables within each cluster. For this, the absolute frequency of the variables was calculated for each group, which allowed for the identification of common patterns and differentiators. Subsequently, an interpretive summary of each cluster was prepared, highlighting the distinctive attributes that define it.

4. Results

4.1. Descriptive analysis

The study excluded 533 startups that are not in operation, finalizing with 1,274 startups (detailed in Table 2). The descriptives for the categorical variables, except for PAIS (country), are located in Table 3.

The data come from all the countries covered by the Startupeable platform (Startupeable, 2025). Individually, the following are differentiated: Mexico (Méx), Colombia (Col), Argentina (Arg), Peru, Chile, and Brazil (Bras). Meanwhile, Costa Rica, Guatemala, Honduras, Nicaragua, El Salvador, and Panama are grouped under the name Central America (CA). The "Others" category aggregates Ecuador, Bolivia, Paraguay, Uruguay, Venezuela, and the Dominican Republic. Table 4 shows the absolute and relative representation by country.

Table 2. Startups by country, total and active.

Country/ Startups	Mex	Chile	Col	Bras	Arg	Perú	Otros	CA	Total
Totals	662	264	268	91	238	136	96	52	**1.807**
Active	454	209	193	89	100	100	77	52	**1.274**
Relative frequency	35.6	16.4	15.1	7.0	7.8	7.8	6.0	4.1	**100**

The variable EDAD, numerical in nature, presented the following descriptive statistics: mean (M) = 8.64, standard deviation (SD) = 5.518, median (Mdn) = 7, and range (Rgo) = 51. Subsequently, this variable was transformed into a categorical one as described in the methodology section. The absolute and relative frequency corresponding to the variables EDAD, last investment round (ULTIMA_RONDA_INVERSION), and number of investors (NUM_INV) is presented in Table 3. Given that the Kolmogorov-Smirnov normality tests indicated a non-normal distribution ($p < 0.001$) in the quantitative variables, non-parametric measures and tests were chosen for subsequent analyses.

Table 3. Distribution of the variables EDAD, ULTIMA_RONDA_INVERSION and NUM_INV

Variable	Category / Absolute frequency (Relative frequency)			
EDAD (n=1274)	x≤6	14≥x≥7	x≥15	
	121 (9.5)	609 (47.8)	544 (42.7)	
ULTIMA_RONDA_INVERSION (n=1072)	No Round	Before Series A	Series A	After Series A
	194 (18.1)	662 (61.8)	153 (14.3)	63 (5.9)
NUM_INV (n=432)	None	One	Two	More than Two
	194 (44.9)	151 (35.0)	39 (9.0)	48 (11.1)

Regarding the CAPITAL_LEVANTADO variable, the descriptive statistics were: n = 728, median = 500,000.00, range = 2,000,000,000.00, with a minimum value of 0 and a maximum of 2,000,000,000.00. On the other hand, the NUM_SEGUIDORES variable presented the following values: n = 1,274, median = 2,840, range = 23,005,000, minimum = 0 and maximum = 23,005,000. These data reflect high dispersion in both variables, which justifies the use of robust measures of central tendency.

4.2. Relationship analysis of categorical variables

The variable PAIS reports a significant relationship with all categorical variables. The results are presented in Table 4.

Table 4. Cluster differences for variables CAPITAL_LEVANTADO and NUM_SEGUIDORES

Variable	CAPITAL_LEVANTADO (n=728)		NUM_SEGUIDORES (n=1274)	
	H (gl)	p	H (gl)	p
PAIS	33.42 (7)	<.001	29,407 (7)	<.001
EDAD	8.989 (2)	<.011	36,599 (2)	<.001
ULTIMA_RONDA_INVERSION	496,438 (3)	<.001	98,440 (3)	<.001
NUM_INV	270,149 (3)	<.001	57,469 (3)	<.001

CAPITAL_LEVANTADO Post hoc analysis:

- **Brazil** (Mdn= 3,900,000.00; Rgo= 1,500,000,000.00) is the only country with significant differences compared to **Central America** (Mdn= 150,000.00; Rgo= 2,800,000.00, p= .043), **Chile** (Mdn= 500,000.00; Rgo= 316,000,000.00, p= .010), **Mexico** (Mdn= 199,000.00; Rgo= 2,000,000,000.00, p < .001), and **Peru** (Mdn= 199,000.00; Rgo= 70,000,000.00, p= .001).
- Regarding **EDAD**, x ≤ 6 Mdn= 200,000.00; Rgo= 445,000,000.00) shows a difference with 7 ≤ x ≤ 14 (Mdn= 500,000.00; Rgo= 2,000,000,000.00; p= .026) and x ≥ 15 (Mdn= 1,000,000.00; Rgo= 2,000,000,000.00; p= .026).
- In **ULTIMA_RONDA_INVERSION** with p < .001, a difference exists between all groups: **No round** (Mdn= 0; Rgo= 0), **Before Series A** (Mdn= 450,000.00; Rgo= 12,999,640.00), **Series A** (Mdn= 7,000,000.00; Rgo= 315,997,000.00), and **After Series A** (Mdn= 26,000,000.00; Rgo= 1,991,000,000.00). The difference regarding **NUM_INV** lies between "no investors" and all other categories.

NUM_SEGUIDORES Post hoc analysis:

- **Peru** (Mdn= 6,377; Rgo= 1,653,700) represents the only country significantly different from **Colombia** (Mdn= 2,000; Rgo= 2,172,002, p= .003) and **Mexico** (Mdn= 2,000; Rgo= 23,005,000, p < .001).
- For **EDAD**, 7 ≤ x ≤ 14 (Mdn= 5,000; Rgo= 9,753,400) has a higher median than x ≤ 6 (Mdn= 2,000; Rgo= 1,241,104, p < .001) and x x≥ 15 (Mdn= 2,001; Rgo= 23,005,000, p < .001).
- In **ULTIMA_RONDA_INVERSION**, significant differences occur between all group pairs, except between **Series A** and **After Series A** (No Round–Series A, p= .008; all other pairs, p < .001). The medians and ranges are: **No Round** (Mdn= 176,000; Rgo= 23,005,000), **Before Series A** (Mdn= 2,093.50; Rgo= 2,126,900), **Series A** (Mdn= 10,000; Rgo= 2,105,455), and **After Series A** (Mdn= 62,497; Rgo= 2,105,455).
- Regarding **NUM_INV**, it shows p < .001 for the "No investors" group compared to the remaining three. "**One investor**" (Mdn= 4,000; Rgo= 900,255) and "**More than Two investors**" (Mdn= 17,695.5; Rgo= 1,797,010).

4.3. Correlation Analysis

Using the Spearman's R statistic, a moderate correlation is detected between the variables **CAPITAL_LEVANTADO** and **NUM_SEGUIDORES** (r= .302, p < .001). Table 5 summarizes other correlation results, including ordinal categorical variables, highlighting the strongest relationship between **CAPITAL_LEVANTADO** and **NUM_INV** (r= .974, p < .001); as well as ULTIMA_RONDA_INVERSION and NUM_INV (r=0.932, p<0.01), and CAPITAL_LEVANTADO and

ULTIMA_RONDA_INVERSION (r=0.931, p<.001). Since all three relationships are positive and very close to one, it can be stated that there is a strong correlation between these variables

Table 5. Variables correlation matrix.

VARIABLES	CAPITAL_LEVANTADO	TOTAL_SEGUIDORES	EDAD	NUM_SECTORES	ULTIMA_RONDA_INVERSION
TOTAL_SEGUIDORES	.302 (<.001)				
EDAD	.107 (.004)	.050 (.077)			
ULTIMA_RONDA_INVERSION	.931 (<.001)	.287 (<.001)	.038 (.210)	.148 (<.001)	
NUM_INV	.974 (<.001)	.360 (<.001)	-.380 (<.001)	.274 (<.001)	.932 (<.001)

4.4. Hierarchical Cluster Analysis

The results of calculating the silhouette score for the clusters identified by each algorithm are shown in Table 6.

Table 6. Evaluation of clustering methos for the data collected.

Algorithm	Number of clusters	Silhouette index
K-Means	2	0.2060
Agglomerative hierarchical clustering	3	0.1170
Gaussian mixture models	7	0.2465

The agglomerative hierarchical clustering algorithm was chosen, as it obtained the best Silhouette index. For this technique, the number of clusters that produces the lowest silhouette index is three.

The dendrogram obtained by applying the hierarchical clustering algorithm is shown in Figure 1. A large vertical separation can be observed at the top of the dendrogram (around a distance of 25–30), just before the tree clearly splits into three groups: a group on the left, a group in the center, and a group on the right. According to the dendrogram, the optimal number of clusters is 3, as cutting at this level produces large distances between the groups, indicating good separation. More than three partitions significantly reduces the distance between branches, suggesting that additional groups are not clearly differentiated. This confirms the results shown in Table 6.

The two-dimensional representation of the data is presented in Figure 2. It shows that the group identified as **Cluster 1** contains the fewest elements but exhibits a clear and defined separation from the other groups, indicating a distinctive structure within the dataset.

On the other hand, **Clusters 2 and 3** present a certain degree of overlap, suggesting that a subset of startups exists whose characteristics do not allow for a sharp differentiation between both groups. Nonetheless, these cases represent a minority of the total; therefore, the general segmentation can be considered appropriate and coherent.

Table 7 shows the distribution of startups by country in each cluster , while Tables 8 shows a summary of the distribution of variables with respect to the cluster to which they belong. Below, we present a description and interpretation of the identified clusters is presented, based exclusively on these variables.

Cluster 1: Mid-aged startups without access to investment

This group, comprised of 194 startups, has a distribution dominated by intermediate-aged companies: 60.8% are between 7 and 14 years old, while only 19.6% are younger than 6 years. They completely lack access to external investment in financing rounds; therefore, 100% of the companies have not reported investors. Nevertheless, they show diversity in their digital presence, although with a higher concentration in the first quartile (Q1, 48.5%), which suggests limited visibility on social platforms.

Cluster 2: Young startups with early access to investment

With 444 startups, the second cluster is composed almost entirely of newly founded startups ($x \leq 6$ years) that, nonetheless, have already accessed investment rounds (n = 433 in round 1) and have at least one investor; additionally, 93.7% have raised pre-seed or seed capital, and others have initiated more advanced capital-raising processes (n=26 Series A and n=2 beyond Series A). Their digital presence is moderate: 42.5% are located in the second quartile (Q2) and 18.2% in the first (Q1). This conglomerate has a significant representation of total Mexican companies (43.9%), as well as startups from Central America (72.9%), Peru (52.6%), and other countries (62.7%).

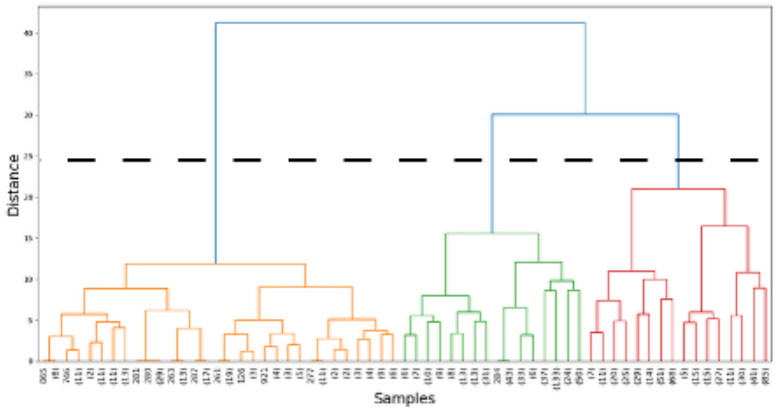

Fig. 1. *Visualization of the cluster hierarchy. The dashed line indicates the optimal cut-off point to define a specific number of groups.*

Table 7. Number of startups per country in each cluster.

	PAIS							
Cluster	Arg.	Bras.	CA	Chile	Col.	Méx.	Perú	Otros
1	18	9	6	30	12	103	13	3
2	32	27	35	64	61	127	51	47
3	49	40	7	112	109	59	33	25

Table 8. Distribution of variables by cluster: EDAD, ULTIMA_RONDA_INVERSIÓN, NUM_INV, CAPITAL_LEVANTADO, NUM_SEGUIDORES.

EDAD			
Cluster	1	2	3
1	38	118	38
2	432	12	0
3	27	352	55

ULTIMA_RONDA_INVERSION				
Cluster	0	1	2	3
1	194	0	0	0
2	0	433	10	1
3	0	229	143	62

NUM_INV				
Cluster	0	1	2	3
1	194	0	0	0
2	0	81	13	20
3	0	35	18	16

CAPITAL_LEVANTADO				
Cluster	0	1	2	3
1	194	0	0	0
2	0	416	26	2
3	0	308	61	65

NUM_SEGUIDORES				
Cluster	Q1	Q2	Q3	Q4
1	94	28	33	39
2	80	189	119	56
3	97	61	116	160

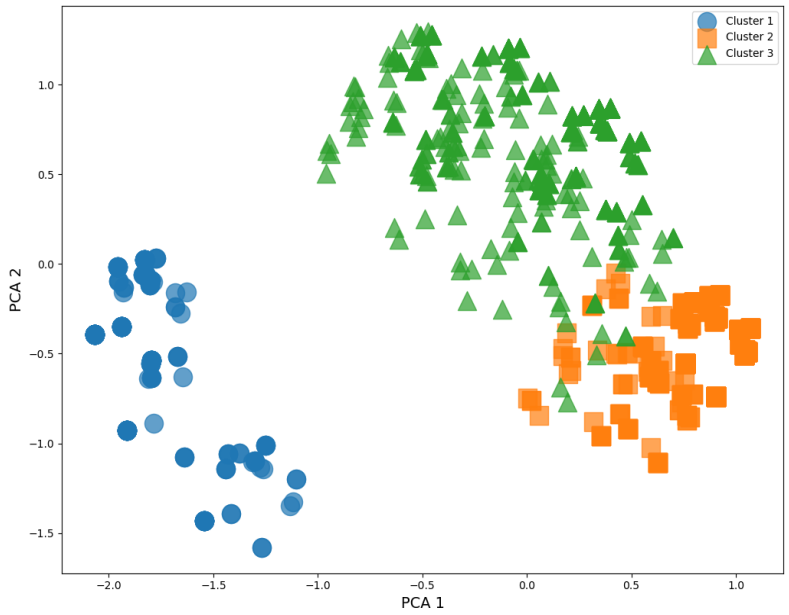

***Fig. 2**. Groupings of Latin American startups projected with PCA for visual exploration.*

Cluster 3: More advanced startups with higher capital and visibility

This cluster groups 434 startups, representing 40.4%, dominated by mid-aged (81.1%) and advanced-aged (12.7%) startups. They exhibit advanced characteristics: presence in multiple investment rounds (143 in round 2 and 62 in round 3), participation of various investors, and the highest proportion of capital raised in Series A (n = 61) and beyond Series A (n = 65) stages. Of the three groups, they are leaders in terms of public visibility: 160 startups are located in the top quartile of followers, indicating a robust digital positioning strategy. The countries with the highest concentration are Colombia (59.9%), Brazil (56.2%), and Chile (54.4%).

4.5. Discussion

The following paragraphs discuss the composition of the clusters, considering not only the quantitative aspect obtained with the agglomerative hierarchical clustering algorithm, but also external aspects or variables that help explain the formation of companies and their evolution.

The sample used in the cluster analysis provides a general sketch of the startup sector in LATAM. The typical startup in the sample has an average age of 8.6 years. The youngest are two years old, while three are the oldest startups in the sample, at 41, 42, and 53 years old. Slightly more than 18% have not raised investment; the rest have accessed some round, especially before Series A, obtaining a central measure of 500,000.00 dollars. This result is consistent with the region's resource attraction, where seed capital investment (43%) by LATAM VC investors prevails (SGIA, 2020).

On the other hand, the conglomerates allow the identification of three large groups of startups: a) Cluster 1. Mid-aged startups that have not achieved progress in capital attraction, with diversity in the number of social media followers; b) Cluster 2. Young startups starting and capturing capital with a medium number of followers on their social networks; and c) Cluster 3. Mid-aged and advanced longevity startups that have gone through various investment rounds and possess larger social media communities.

When contemplating the classification by Mageste et al. (2024), Cluster 2 reveals the initial phase of the life cycle. This stage is characterized by the emergence and validation of the idea, the formation of the work team, and the acquisition of the first customers, using self-financing, investment from Angel investors, and mechanisms such as crowdfunding (Escartín et al., 2020; Mageste et al., 2024; Solís Tepexpa & Torres Castañeda, 2022). Thus, 41.4% of the startups in this sample do not focus

on profitability, but rather on the development and exposure of a promising product, which is the objective of this stage (Angel Ventures, 2022; Kuusinen et al., 2018).

This profile reveals a group of nascent but highly dynamic startups, possibly linked to incubators or innovation programs that bring them closer to seed capital funds. The geographical distribution is more diverse, with an emphasis on Mexico, Peru, Central America, and other countries. The concentration in intermediate quartiles of followers suggests a communication strategy under construction. Despite their youth, these startups already show effective insertion into the regional financing system.

Under the current context, for this group of startups, moving to the next stage will represent a greater challenge than for the group of more mature startups; because by the year 2020, 26% of startups achieved the conversion from seed capital to Series A in less than two years, a percentage that plummeted to 1% in the year 2023 (Endeavor & Glisco Partners, 2024). Additionally, it should be noted that although the founders of the ventures in Cluster 2 have been able to identify a business opportunity and raise their first rounds of capital, many will fall by the wayside if they fail to implement the business idea and overcome the difficulties of establishing the company; characteristics that Peña-Ramírez et al. (2020) point out when referring to "zombie entrepreneurs," whom the authors also identify as common entrepreneurs in LATAM innovation ecosystems.

Continuing with the classification of Mageste et al. (2024), Cluster 3 represents growth and expansion. Blocks of mid-aged and advanced-aged startups dominate, suggesting a broader history of operation, organizational learning, and business model consolidation (Escartín et al., 2020; Malekmohammadi et al., 2024; Solís Tepexpa & Torres Castañeda, 2022). The clustering procedure does not manage to make a distinction to delimit the growth stage from the expansion stage, requiring additional information such as sales volume, customers, financial results, and internationalization data; especially since the growth stage is reached upon exceeding the economic break-even point and the expansion stage is achieved by consolidating financial muscle (Escartín et al., 2020).

In this way, a significant number of startups in the study are under the banner of scaling. In this regard, Kuusinen et al. (2018) points out that scaling starts after five years of life with a product validated in the market and with stable but relatively sustainable income. Importantly, the startups grouped in Cluster 3 are older than 6 years. Although this group of startups has advanced the most in the steps of capital attraction, it is still incipient because knowledge about VC mechanisms in the region is relatively young (Rockstart, 2023). On the other hand, the countries with the greatest presence in this cluster are Brazil, Colombia, and Chile—countries that respond to a more robust ecosystem (Startupblink, 2024).

Although conglomerate number 1 represents the smallest number of startups in the sample, its characteristics trigger questions about its stage of development. These startups have exceeded six years in the market but report no capital raising. Returning to the premises that a capital crisis limits startup growth (Chakraborty et al., 2025), and that capital raising is an indicator of success (Maarouf et al., 2025), one might think these are startups facing stagnation.

On the other hand, longevity is also an indicator of success (Díaz-Santamaría & Bulchand-Gidumal, 2021). In that sense, their permanence without external capture could be justified by more sustainable practices that reduce the Burn Rate (as suggested by Endeavor & Glisco Partners (2024) to avoid quickly consuming capital). A second possibility is the Spinoff figure, since it has support from the company from which it was born. In the study by Acosta et al. (2024) on university spinoffs, it is analyzed how the resources and capabilities of the original entity become decisive factors in the creation of the new company, providing tangible and intangible resources, and even visibility before the industry. These advantages reinforce what Furlan (2016) detects: a Spinoff emanating from a consolidated organization lasts longer than a startup.

Before the search for conglomerates, statistical tests of relationship and differentiation between the research variables were analyzed. These statistical tests revealed a large number of significances; notably, the supremacy of Brazil in capturing capital and positioning itself visibly in Cluster 3 stands out. This result is not surprising; Brazil ranks above other LATAM countries in the 2024 Innovation Index (The Global Economy, 2024), the 2025 Kearney Foreign Direct Investment Confidence Index (Kearney, 2025), and broadband connectivity as part of the National Digitalization Strategy (OECD et al., 2024). It is also the Latin country with the highest level of maturity in the startup ecosystem (Startupblink, 2025).

Continuing with the relationships, the correlation between the value of capital raised and the number of investors coincides with the results obtained in Martielli et al. (2025) and Shabtai et al. (2025). The authors point out that achieving financing is a signal that captures the attention of other investors. However, in practice, entrepreneurs must sustainably distribute equity or use mechanisms that allow the startup to remain attractive to future investors.

Finally, the digital window provides the startup with data to establish strategic decisions as part of social media intelligence and supports the generation of large-scale communities of prospects and customers (Peixoto et al., 2023). As observed, startups with greater participation in investment rounds showed larger audiences, a result backed by significant correlation. This evidence is consistent with the network effect, since as users increase, the cost of acquiring customers decreases, consequently contributing to accelerated growth and allowing for early internationalization (Usman & Sun, 2023).

Thus, starting from the theoretical review of the startup and its life cycle, the analysis of a significant sample of LATAM startups under the lens of statistical and cluster analysis, the study provides a comprehensive view of the development and characteristics of startups in the region. Segmentation through cluster analysis using age, capital attraction, and digital visibility allowed for the clear identification of the initial stage of the life cycle, while the growth and expansion stages appeared merged. For future research, it will be necessary to incorporate financial and market variables.

Regarding the small cluster formed by mid-to-advanced-aged startups without capital raising, it leaves open the question about their sustainability, which is presumed to involve the adoption of a low burn rate strategy or a spinoff figure. This fact justifies a second investigation into this small group of startups with a qualitative approach.

5. Conclusions

In this paper, data collected from more than 1,000 LATAM startups was analyzed. Since the data was unlabeled, it was analyzed using statistical techniques combined with the agglomerative hierarchical clustering algorithm, which was chosen for showing the best performance compared to two other methods.

The results yield the following conclusions:

a) The clustering model managed to clearly isolate the initial stage of the startup life cycle (Cluster 2). This stage is characterized by young companies (6 years old or less) with dynamic access to pre-seed and seed capital. The growth and expansion stages appear merged (Cluster 3), grouping together older companies that have scaled to more advanced financing rounds (Series A or higher).

b) A correlation between capital raising and the size of digital communities (number of followers) was corroborated through statistical analysis. Startups with greater participation in investment rounds boast larger audiences, confirming that social media traction is a key driver for expansion. Furthermore, it was proven that countries with more mature ecosystems (Brazil, Colombia, and Chile) concentrate the majority of companies positioned in the most advanced growth phases.

c) A small group of mid-aged startups that have survived without raising external capital was identified (Cluster 1). This finding suggests that, given the difficulties of obtaining financing in LATAM, some companies opt for high-sustainability strategies or manage to survive by operating as a spinoff backed by the resources of an established organization.

As future work, it is necessary to incorporate market variables and financial metrics (such as sales volume, number of customers, or revenue) that provide the model with the discriminant capacity needed to quantitatively separate the growth phase from the expansion phase.

6 References

Acosta Medina, J. K., Contreras Pacheco, O. E., & Pedraza Avella, A. C. (2024). Factores que inciden en la creación de spin-offs universitarias: Un análisis multinivel desde la literatura. *Estudios Gerenciales, 40*(170), 31–44. https://doi.org/10.18046/j.estger.2024.170.6268

Álvarez, C., Amorós, J. E., & Urbano, D. (2014). Regulations and entrepreneurship: Evidence from developed and developing countries. *Innovar, 24*(1Spe), 81–89. https://doi.org/10.15446/innovar.v24n1spe.47548

Angel Ventures. (2022). *Venture capital fund toolkit: A guide to set up a VC fund in Latin America and the Caribbean*. Inter-American Development Bank. https://doi.org/10.18235/0004618

Banco Interamericano de Desarrollo. (2024, June 20). *Estudio: Ecosistema fintech en América Latina y el Caribe supera las 3.000 startups*. BID source

Banerji, D., & Reimer, T. (2019). Startup founders and their LinkedIn connections: Are well-connected entrepreneurs more successful? *Computers in Human Behavior, 90*, 46–52. https://doi.org/10.1016/j.chb.2018.08.033

Berre, M., & Le Pendeven, B. (2023). What do we know about startup-valuation drivers? A systematic literature review. *Venture Capital, 25*(4), 385–429. https://doi.org/10.1080/13691066.2022.2086502

Bethlendi, A., Hegedűs, S., & Szőcs, Á. (2025). What could we learn from startup failures? *Journal of Innovation and Entrepreneurship, 14*, Article 35. https://doi.org/10.1186/s13731-025-00493-w

Caro, O., Montes, C., Paronian, E., Gerad, C., Carpeta, D., & Bagur Femenías, L. (2020). Aire fresco. *Revista de Contabilidad y Dirección, 30*, 115–133.

Chakraborty, I., Ilavarasan, P. V., & Edirippulige, S. (2025). What helps age-tech startups scale-up? Some insights. *International Entrepreneurship and Management Journal, 21*, Article 70. https://doi.org/10.1007/s11365-025-01081-w

Clauss, T., Kraus, S., Kallinger, F. L., Bican, P. M., Brem, A., & Kailer, N. (2021). Organizational ambidexterity and competitive advantage: The role of strategic agility in the exploration-exploitation paradox. *Journal of Innovation & Knowledge, 6*(4), 203–213. https://doi.org/10.1016/j.jik.2020.07.003

Daradkeh, M., & Mansoor, W. (2023). The impact of network orientation and entrepreneurial orientation on startup innovation and performance in emerging economies: The moderating role of strategic flexibility. *Journal of Open Innovation: Technology, Market, and Complexity, 9*(1), Article 100004. https://doi.org/10.1016/j.joitmc.2023.02.001

Díaz-Santamaría, C., & Bulchand-Gidumal, J. (2021). Econometric estimation of the factors that influence startup success. *Sustainability, 13*(4), Article 2242. https://doi.org/10.3390/su13042242

Endeavor, & Glisco Partners. (2024). *Ecosistema de Venture Capital y Growth Equity en LATAM 2024*. Report page

Escartín, D., Marimon, À., Rius, A., Vilaseca, X., & Vives, À. (2020). Startup: Concepto y ciclo de vida. *Revista de Contabilidad y Dirección, 30*, 13–21.

Ewens, M., & Malenko, N. (2020). *Board dynamics over the startup life cycle* (NBER Working Paper No. 27769; revised May 2024). National Bureau of Economic Research. https://doi.org/10.3386/w27769

Furlan, A. (2016). Who lives longer? Startups vs spinoffs founded as proprietorships. *International Journal of Entrepreneurial Behavior & Research, 22*(3), 416–435. https://doi.org/10.1108/IJEBR-08-2015-0179

García, A., Martín, A., Iglesias, M., Martín, A., & Quintana, L. (2020). Spotify. *Revista de Contabilidad y Dirección, 30*, 81–99.

Gloor, P. A., Fronzetti Colladon, A., Grippa, F., Hadley, B. M., & Woerner, S. (2020). The impact of social media presence and board member composition on new venture success: Evidences from VC-backed U.S. startups. *Technological Forecasting and Social Change, 157*, Article 120098. https://doi.org/10.1016/j.techfore.2020.120098

Greco, F. (2025). Overseas benchmarking for startup ecosystems in Central and Southern Italy. *International Entrepreneurship and Management Journal, 21*, Article 22. https://doi.org/10.1007/s11365-024-01028-7

Hollander Shabtai, R., Tzofi, O., & Blank, T. H. (2025). Unlocking growth potential: Investigating the significance of firm location for music startups. *International Entrepreneurship and Management Journal, 21*, Article 10. https://doi.org/10.1007/s11365-024-01027-8

ICEX España Exportación e Inversiones, & Secretaría General Iberoamericana. (2021). *Global LATAM 2020: Series inversión extranjera*.

Ionescu, G. G., & Negruşa, A. L. (2007). The study about organizational life cycle models. *Review of International Comparative Management, 8*(4), 5–17.

Kearney. (2025). *The 2025 Kearney Foreign Direct Investment Confidence Index*. Kearney report

Kézai, P. K., & Skala, A. (2024). Remarks on the location theories of startups: A case study on the Visegrad countries. *Regional Science Policy & Practice, 16*(9), Article 100063. https://doi.org/10.1016/j.rspp.2024.100063

Kuusinen, K., Sørensen, M. K., Frederiksen, N. M., Laugesen, N. K., & Juul, S. H. (2019). From startup to scaleup: An interview study of the development of user experience work in a data-intensive company. In M. K. Lárusdóttir, M. Winckler, K. Kuusinen, P. Palanque, & C. Bogdan (Eds.), *Human-centered software engineering* (pp. 3–14). Springer. https://doi.org/10.1007/978-3-030-05909-5_1

Latam Republic. (2024). *Venture capital in Latin America: The essential insights from Cuantico's latest report*. Latam Republic article

Linina, I., Arbidans, D., & Vevere, V. (2021). Identification of business management improvement factors for start-up companies in Latvia in the conditions of globalization. *SHS Web of Conferences, 129*, Article 08011. https://doi.org/10.1051/shsconf/202112908011

Maarouf, A., Feuerriegel, S., & Pröllochs, N. (2025). A fused large language model for predicting startup success. *European Journal of Operational Research, 322*(1), 198–214. https://doi.org/10.1016/j.ejor.2024.09.011

Mageste, S., Plottier, C., Rocha, C., & Saporito, N. (2024). *Empresas emergentes (start-ups) en América Latina y el Caribe: Una primera aproximación a su identificación y características*. Comisión Económica para América Latina y el Caribe.

Malekmohammadi, M., Hosseini, S. H. K., & Hanzaee, K. H. (2025). A behavior-based model of the innovativeness lifecycle: A meta-synthesis approach. *Transformation Management Journal, 17*(1), 1. doi:10.22067/tmj.2024.90344.1622

Martielli, F., Laus, V., Salvi, A., & Kotaskova, A. (2025). Unlocking funding success: Intercultural entrepreneurship dynamics in European startups. *International Entrepreneurship and Management Journal, 21*, Article 19. https://doi.org/10.1007/s11365-024-01038-5

Martínez Sidón, G., Osorio Novela, G., & Saavedra Leyva, R. E. (2025). El desempeño institucional y la seguridad pública en el dinamismo emprendedor de México. *Innovar, 35*(97), e107057. https://doi.org/10.15446/innovar.v35n97.107057

Organisation for Economic Co-operation and Development, CAF—Development Bank of Latin America and the Caribbean, & Latin American and Caribbean Economic System. (2024). *Índice de políticas para PyMEs: América Latina y el Caribe 2024: Hacia una recuperación inclusiva, resiliente y sostenible*. OECD Publishing. https://doi.org/10.1787/807e9eaf-es

Peixoto, A. R., de Almeida, A., António, N., Batista, F., & Ribeiro, R. (2023). Diachronic profile of startup companies through social media. *Social Network Analysis and Mining, 13*, Article 52. https://doi.org/10.1007/s13278-023-01055-2

Peña-Ramírez, C., Garcés-Rodríguez, G., & da Silva, S. S. (2020). Los emprendedores zombis: Emprendimiento en etapa inicial que no crea empresa en el contexto latinoamericano. *Estudios Gerenciales, 36*(156), 352–363. https://doi.org/10.18046/j.estger.2020.156.3610

Pons, F., & Boar, A. (2020). Las fases y documentos necesarios en una ronda de financiación. *Revista de Contabilidad y Dirección, 30*, 47–63.

Ressin, M. (2022). Start-ups as drivers of economic growth. *Research in Economics, 76*(4), 345–354. https://doi.org/10.1016/j.rie.2022.08.003

Rockstart. (2023). *¿Por qué fracasan las startups en Latinoamérica? Estudio del fracaso*.

Saavedra Leyva, R. E., & Texis Flores, M. (2019). El factor institucional en el emprendimiento por oportunidad de América Latina y el Caribe. *Innovar, 29*(73), 99–112. https://doi.org/10.15446/innovar.v29n73.78025

Solís Tepexpa, S., & Torres Castañeda, K. (2022). Importancia de las startup en la economía del siglo XXI. *Revista Gestión y Estrategia*, (61), 45–62. DOI: 10.24275/uam/azc/dcsh/gye/2022n61/Solis.

StartupBlink. (2024). *Global Startup Ecosystem Index 2024*. StartupBlink reports

StartupBlink. (2025). *Corporate Startup Activity Index 2025*. StartupBlink report page

Startupeable. (n.d.). *Directorio de Startups y Venture Capital*. Retrieved September 14, 2026, from Startupeable directory

Startups Latam. (2024). *Informe Startups Latam 2024*. Report page

The Global Economy. (n.d.). *Índice de innovación - Clasificaciones*. Retrieved September 14, 2026, from The Global Economy ranking

Usman, M. A., & Sun, X. (2023). The impact of digital platforms on new startup performance: Strategy as moderator. *Heliyon, 9*(12), Article e22159. https://doi.org/10.1016/j.heliyon.2023.e22159

Vongvisitsin, T. B., & Tung, V. W. S. (2025). Technology start-ups in tourism and hospitality: A networked social capital theory perspective from early-stage start-up founders. *Tourism Management, 106*, Article 104996. https://doi.org/10.1016/j.tourman.2024.104996

Wei, C.-P., Fang, E. S.-H., Yang, C.-S., & Liu, P.-J. (2025). To shine or not to shine: Startup success prediction by exploiting technological and venture-capital-related features. *Information & Management, 62*(6), Article 104152. https://doi.org/10.1016/j.im.2025.104152

World Bank. (n.d.). *Population, total—Latin America & Caribbean* [Data set]. World Bank Open Data. Retrieved September 14, 2026, from World Bank data series