Participation of Latin America in MOOCs: Exploring Trends Across Providers

Jorge Maldonado-Mahauad
Universidad de Cuenca
Pontificia Universidad Católica de Chile
jorge.maldonado@ucuenca.edu.ec

José A. Ruipérez-Valiente
University of Murcia
Massachusetts Institute of Technology
jruiperez@um.es

Mar Pérez-Sanagustín
Université Paul Sabatier Toulouse III
Pontificia Universidad Católica de Chile
mar.perez-sanagustin@irit.fr

Matt Jenner
FutureLearn
matt.jenner@futurelearn.com

Ignacio Despujol[1], Carlos Turró[2]
Universitat Politècnica de València
[1]ndespujol@asic.upv.es
[2]turro@cc.upv.es

Thomas Staubitz[3], Tobias Rohloff[4]
Hasso Plattner Institute
[3]thomas.staubitz@hpi.de
[4]Tobias.Rohloff@hpi.de

Germán Montoro
Universidad Autónoma de Madrid
german.montoro@uam.es

Justin Reich
Massachusetts Institute of Technology
jreich@mit.edu

Abstract—**Massive Online Open Courses (MOOCs) have become popular in various regions of the world through the years. Since 2008, this phenomenon has received plenty of attention from higher education and universities across countries began to produce these courses. The countries of Europe and the United States are the world's leading producers of MOOCs and research studies reporting on this topic. This previous research has focused on (1) analysing data from global providers such as edX, Coursera or FutureLearn; (2) describing learners' characteristics from a small sample of courses in these regions; and (3) offering overviews of courses and platforms. However, research in other regions such as Latin America or Africa are very scarce. As a consequence, little is known about local initiatives in Latin America region, and about the needs and characteristics of its learners. Moreover, this has generated an unequal and biased perspective of what we know today about MOOC learners. To close this inequality gap, this work, presents a cross-platform exploratory study in Latin America, using data from more than three million learners and seven different MOOC providers to generate a joint comparable analysis about students' characteristics in this region with others regions in the world. Preliminary results report on the differences and similarities of trends based on level of education, age, gender of students, their level of activity and performance of learners in Latin America through the different providers of MOOCs. These results help us understand the MOOC ecosystem in Latin America and report results to the entire community, while at the same time calling for more large-scale studies between researchers and institutions.**

Keywords—MOOCs, Latin America, global learning

I. INTRODUCTION

Massive Open Online Courses (MOOCs) are a global phenomenon that is transforming teaching and learning while facilitating researchers to reflect on new forms of learning in higher education institutions (HEI) worldwide. The increase in the interest and importance of MOOCs is reflected in the number of courses published and students that registered to them. As of December 2019 and according to Class Central [1], nearly 110 million learners reported to have registered in one of the more than 13.5 thousand MOOCs. This is almost twice as many learners and courses compared to the 6,850 courses published and nearly 58 million students registered through 2016 [2] This accelerated growth and its rapid adoption in different countries, has caught the attention of researchers, who seek to understand what its propagation has been like, its characteristics, the impact on students, as well as the technological, social, cultural and economic barriers that underlie the creation of these courses [3].

A challenge in studying a global phenomenon such as MOOCs lies in the diversity of education systems and the heterogeneity of its learners around the world [4]. Most studies have focused on the global MOOCs providers of Anglo-American universities (such as edX, FutureLearn or Coursera), where the predominant language of these courses is English. Some of these studies have discussed the impact of language and culture on learning [5]. This has generated a systematic problem of inequality in online MOOC learning given the low levels of English proficiency of learners in regions such as Latin America [6]. Thus, MOOC providers failed in their mission of "democratizing higher education" by creating an open, global and free online access to courses from the best universities[5].

In Latin America, the great take-off of MOOCs began only in 2015 thanks to the development of regional initiatives such as the Erasmus+ MOOC-Maker project, where a growing number of MOOCs were produced in languages other than English (i.e., Portuguese and Spanish) [7] by local MOOCs providers. Even more considering that the Spanish language is the second most spoken in the world after the Chinese language [8]. This increase in the number of courses is mainly due to three reasons: (1) the association of Latin American universities to platforms such as Coursera and edX (these platforms began to worry about meeting the demand for courses in other languages); (2) the dissemination and development of MOOC platforms in Latin America such as MiriadaX, Telescopio (Guatemala) or Veduca (Brazil); and (3) that universities with great recognition in the region such as the University of Sao Paulo, the National Autonomous University of México or the Pontifical Catholic University of Chile joined the MOOC initiative in a timely manner, beginning to experiment with new educational models based on this type of courses [9]. However, to the best of our knowledge, studies developed to deeply understand this variation in different regions in the world (with the exception of the United States and Europe) and which account for the kinds of course registration and course-taking patterns influenced by cultural factors, are scarce [10]. Moreover, there are no studies in Latin America that examine trends at a macro level of several MOOC providers that allow us to understand which trends in MOOCs are universal and which of these are context dependent.

In this paper, using data from seven MOOC provider partners, we conduct a cross-platform exploratory study with data from more than three million of enrolled learners from the Latin American region. By researching how learners in Latin America differ from other students globally, it is possible to better understand the variation in the adoption and use of MOOCs regionally. For this, the following research questions have been raised in this paper:

- *RQ1: Which are the demographics of Latin American learners across the MOOC providers?*

- *RQ2: How are Latin American learners distributed across the MOOC providers?*

- *RQ3: What is the level of activity and performance of Latin American learners across providers?*

II. METHODOLOGY

A. Multiplatform MOOC Analysis

In order to conduct this research, we follow a four step process: 1) make an initial call looking for representative partners with access to large MOOC datasets from different platforms; 2) partners have to shape their data into the desired same common format; 3) a Jupyter notebook was provided to the MOOC partners to run it on their datasets, this script expects exactly the common data format specified in 2); lastly in 4) each partner shares the aggregate data output that comes out of the script and we conducted the joint data analysis with data from all different institutions. This methodology is helpful to manage the logistical and privacy concerns of sharing different levels of learners' information, where we performed comparisons with datasets that have the same variables and the analysis is conducted using the same script.

B. Variables and Measures

In order to conduct this research, we have considered the following variables for the analysis: level of education, age, and gender of students. Also, we explored some course activity and performance results based on two percentage metrics and defined these as follows:

a) viewed (they accessed the course)
   % viewed = 100* (# viewed / # registered)

b) completed (they achieved a passing grade)
   % completed = 100 * (# completed / # viewed)

We want to remark that the % of viewed is computed based on the total number of registrations but the % of completed is based on the total number of learners that viewed the course (defined as at least accessing the course contents once). The rationale is that since many MOOC learners enroll to a course but do not even access the contents once, using the total number of learners that viewed the course as a denominator provides a more realistic and less noisy metric for the % of completed.

C. Context and MOOC Data

We provide the description of the context and the size of data of the seven providers that have joined this research. The partners provided a data sample from all of their learners, but we filtered and took into consideration for this study only those learners that come from Latin American countries (i.e., México, Colombia, Argentina, Peru, Venezuela, Chile, Ecuador, Guatemala, Cuba, Bolivia, Dominican Republic, Honduras, Paraguay, Salvador, Nicaragua, Costa Rica, Puerto Rico, Panama, Uruguay, and Brazil). These providers considered in the study are:

- **MITx and HarvardX** (abbreviated as MITxHx) where the nature and target of the courses are diverse. Data collected of above 1,071,450 enrolled learners across 552 MOOCs.

- **FutureLearn** founded by the UK Open University and partners with over 170 organizations. Most courses are in English. Data collected of around 55,647 enrolled learners across 1,545 MOOCs.

- **openSAP** founded by German software company SAP. Most courses are in English. Data collected of around 84,496 enrolled learners across 166 MOOCs.

- **PUC (Coursera)** from the Pontificia Universidad Católica of Chile, all the courses are in Spanish. Data collected of around 697,666 enrolled learners across 25 MOOCs.

- **UAMx** from the Universidad Autónoma of Madrid. All courses are in Spanish and hosted in edX. Data collected of around 160,000 enrolled learners across 66 MOOCs.

- **UPValenciaX** from the Universitat Politécnica de Valencia. All courses are in Spanish and hosted in edX. Data collected of around 940,458 of enrolled learners across 230 MOOCs.

- **UPVx** is another site from Universitat Politécnica de Valencia. Most courses are in Spanish and Catalan hosted in Open edX. Data collected of around 17,354 enrolled learners across 132 MOOCs

In the case of partners MITx, HarvardX, openSAP, PUC, UAMx, UPValenciaX and UPVx, they used data from their entire portfolio of MOOCs. In the case of FutureLearn, they used a data sample from several institutions with a large number of courses. Therefore, we consider that this overall sample is representative of the whole worldwide population of Latin American MOOC learners.

III. RESULTS

A. RQ1: Which are the demographics of Latin American learners through the MOOC providers?

In the next Fig.1 we show the distribution of the age, gender and level of education across the seven MOOC providers for LATAM countries. Regarding age, we grouped learners in different age segments in a blue divergent palette of colors (darker means older) allowing a comparison across providers. The most common age segment registering learners for most providers is between [26, 35) except for openSAP whose learners are between [36, 45). The trend shows that openSAP MOOC provider has the oldest population of learners, while the rest of providers have mainly young professionals. However, FutureLearn shows the most heterogeneous distribution of learners in terms of age, and UAMx MOOC provider has the highest percentage of young learners under 25 years old.

Regarding the distribution of gender by provider, we used a 100% stacked bar chart using two colors. The openSAP MOOC provider has the most percentage of male learners across the LATAM countries excluding Bolivia and Paraguay. Also, the most percentage of learners taking MOOCs in the
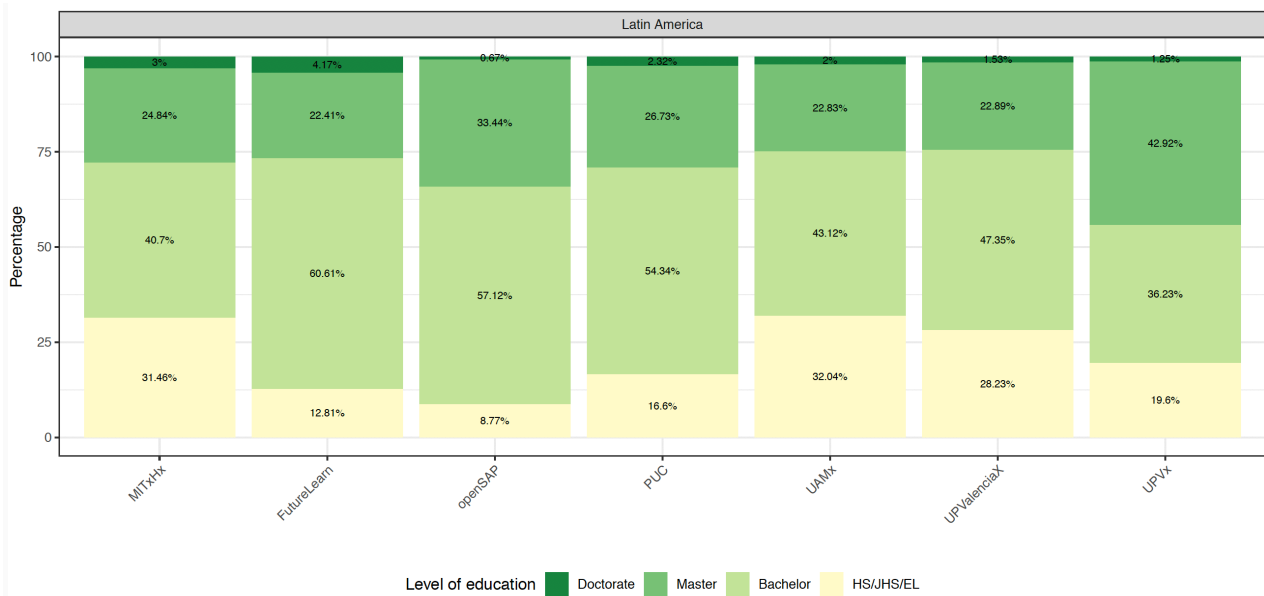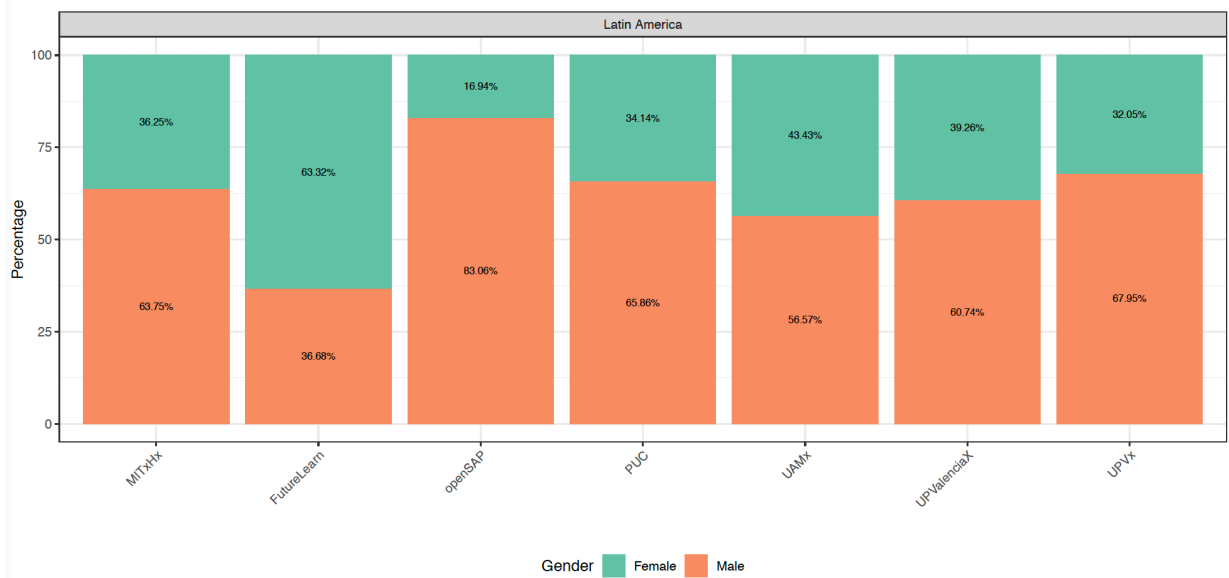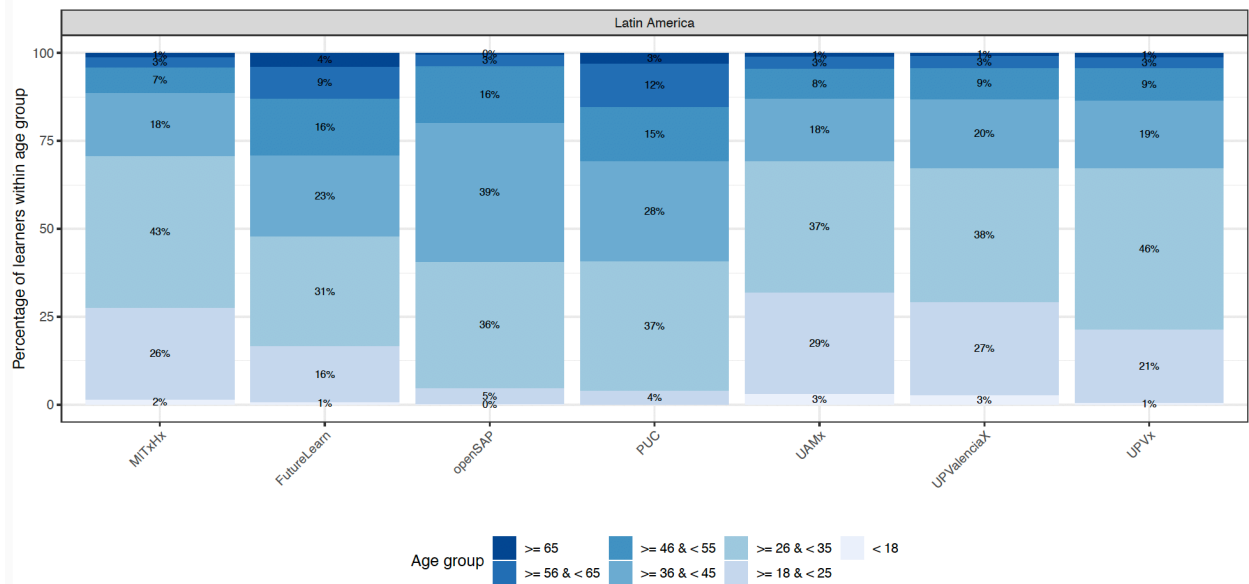
Fig. 1. Distribution by age, gender and level of education across seven MOOC providers in LATAM
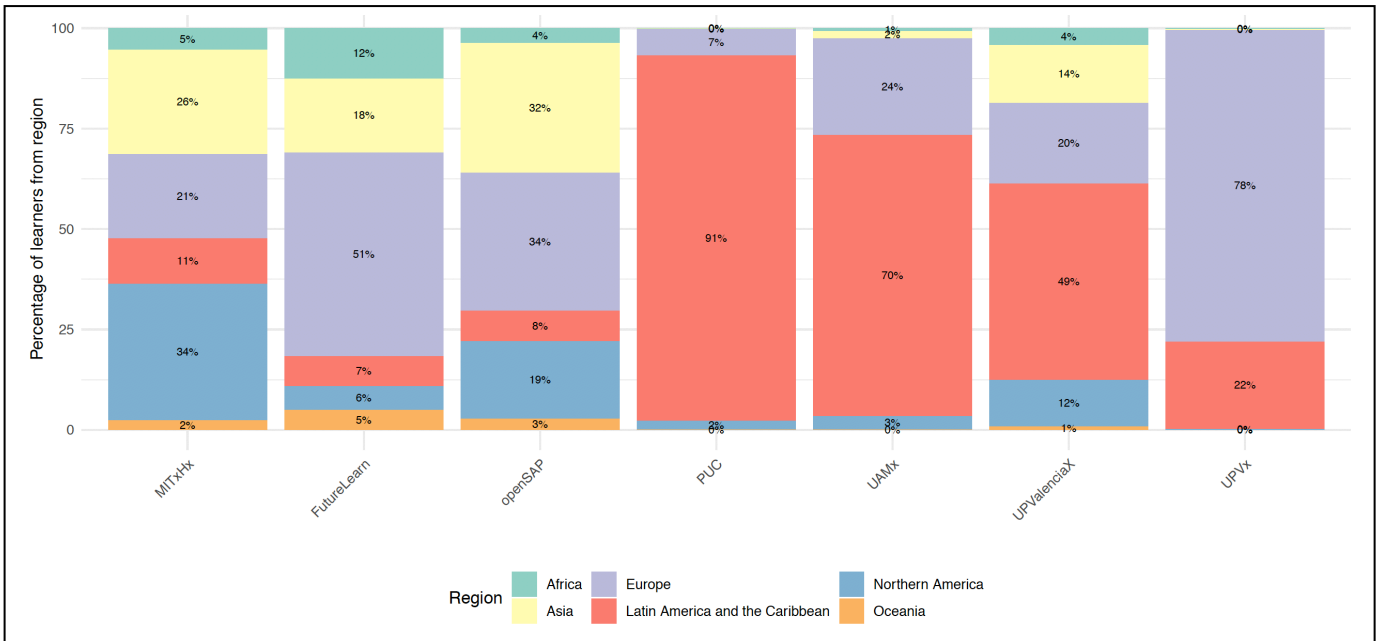
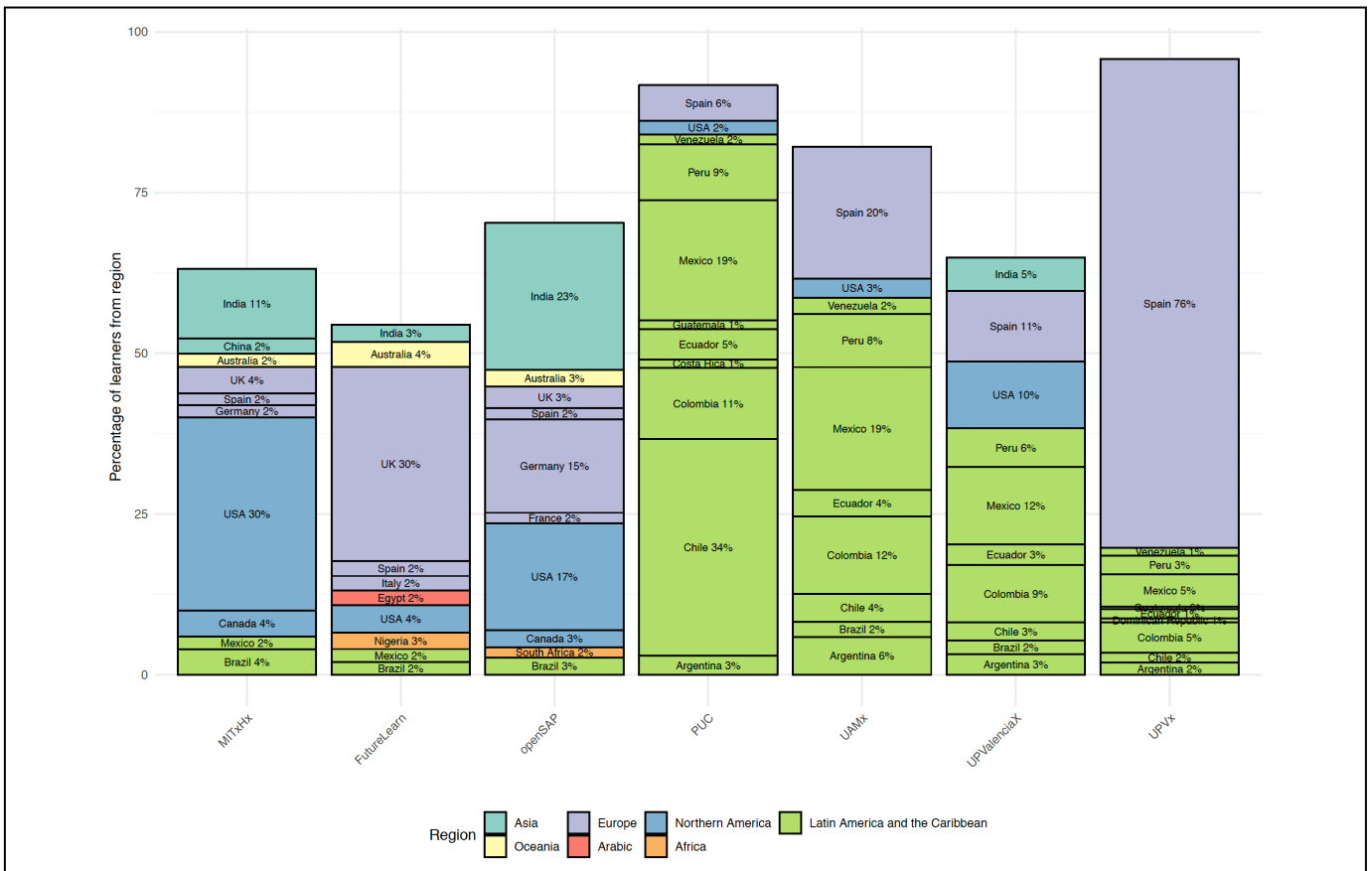Fig. 2. Distribution of learners across the seven MOOC providers for region

Fig. 3 Distribution of the top-ten countries representation region in terms of percentage of learners across the seven MOOC providers in LATA

platforms are mainly male (excepting FutureLearn MOOC provider). In FutureLearn MOOC platform, most of the countries have female learners (excluding Salvador and Guatemala where this trend is different and the most percentage of learners are male). This gender gap may be influenced by the nature of openSAP courses, which are mainly technical, despite the fact that female participation in higher education enrollment in Latin America grew by 3% between 2010 and 2016 according to a report from the Iberoamerican network of Higher Education Indicators [11].

TABLE I.   LATAM MOOC PROVIDERS OVERVIEW

MOOC Provider	MOOCs offered	Unique Learners	Number of Enrollments	Number of Viewed	Number of Completed	Number of Certified	% Viewed	% Completed
MITxHx	552	490,358	1,071,450	619,058	29,117	24,904	57.58	4.7
FutureLearn	1,545	15,146	55,647	17,450	9,458	12,657	29.84	54.2
openSAP	166	23,359	84,496	66,985	9,444	0	80.5	14.1
PUC	25	514,264	697,666	162,374	31,421	20,823	23.38	19.35
UAMx	66	128,714	159,712	143,752	43,215	43,215	89.02	30.06
UPValenciaX	230	592,607	940,458	563,897	52,132	43,561	60.27	9.24
UPVX	132	14,935	17,354	9,661	1,709	0	54.77	17.69

Regarding the level of education by provider, we show the distribution using a 100% stacked bar chart. We present four different educational levels such as 'Doctorate', 'Master', 'Bachelor', and 'High School, Junior High School or Elementary School (HS/JHS/EL)' presented in a palette of colors. Darker shades represent higher level of education. There are interesting distinctions when comparing MOOC providers.

MITxHx and UAMx MOOC providers shows similar proportions of learners with HS/JHS/EL and bachelor's degree, while FutureLearn and openSAP have more proportion of learners with a bachelor's degree. However, UPVx has more percentage of learners with a master's degree and FutureLearn has more percentage of learners with a Doctorate degree. This means that UPVx and FutureLearn attract more educated learners from LATAM. The openSAP MOOC provider have the most percentage of learners with lower level of education with 12.86% of their learners with a HS/JHS/EL education

B. RQ2: How are Latin American learners distributed through MOOC providers?

In Fig. 2 we show the distribution of learners across the seven MOOC providers for region (LATAM countries included) using a 100% stacker bar chart, where the color represents the region in each MOOC provider (LATAM learners are in red).

We explore the distribution of learners across MOOC providers, where PUC MOOC provider attracts more percentage of learners (91%) from LATAM followed by UAMx (70%) while FutureLearn and openSAP attract less percentage of learners from LATAM (7% and 8% respectively). This can be explained given that PUC and UAMx offer their courses mainly in Spanish language and despite UPValenciaX and UPVx that also offer their courses mainly in Spanish, PUC and UAMx can be considered as benchmark for students in the region.

If we look for the other regions, we can see that MITxHx attracts more percentage of learners from North America, UPVx and FutureLearn attract more percentage of learners from Europe, openSAP attracts more percentage of learners from Asia, and FutureLearn is the unique MOOC provider attracting more percentage of learners from Oceania and Africa.

For specific details, Fig. 3 shows a stacked bar chart with the top-ten most representative countries in terms of percentage of learners for each platform, where the color represents the region of the country, which helps to identify the regional course focus on each provider. The MITxHx and FutureLearn attract more percentage of LATAM learners from Brazil and México, openSAP attracts more percentage of learners from Brazil, while PUC, UAMx and UPValenciaX attract more percentage of LATAM learners (85%, 57% and 38% of learners respectively).

C. RQ3: What is the level of activity and performance of Latin American learners?

We explored some course activity and performance results based on two percentage metrics: viewed (they accessed the course), completed (they achieved a passing grade) (see Table I). The total number of unique enrolled learners in Latin America across the seven MOOC providers is around 1.77 million, where 176,496 of these learners completed the course and 145,160 of these learners certified the course (see Table I).

From Table I, we can observe that learners in FutureLearn platform have the highest percentage of completer learners being around the 54,2%, in contrast with MITxHx that reach only 4,7% of completer learners, being the lowest percentage of completion from LATAM learners. However, MITxHx have the highest number of enrolled learners in their platform (above 1.07 millions of enrolled learners), becoming the platform with the worst completion rate of courses. However, it is important to note that the FutureLearn platform is designed to stimulate horizontal communication between tutors and students, while the MITxHx platform is designed for students to take their courses independently.

We can also observe that UPValenciaX has the highest number of unique learners and the highest number of MOOC completions (592,607 learners and 52,132 learners respectively). In contrast, UPVX has the smallest number of unique learners and the smallest of MOOC completions (14,935 learners and 1,709 learners respectively). In the other hand, FutureLearn has the highest percentage of learners that certified the MOOC followed by UAMx (71.38% and 30.64% respectively). However, PUC in Latin America is a benchmark for students in the region, since it has only 25 courses developed on Coursera platform (if compared with MITxHx and UPValenciaX that offer 552 and 230 courses respectively). Additionally, PUC is the second platform in terms of registrations (in this study), with more than half a million unique learners and the third in terms of enrolment with nearly 700,000 enrollments in their courses.

IV. Discussion and Conclusions

This cross-platform study represents an important step in being able to globally understand and explain the MOOC ecosystem by looking at it form a large-scale perspective and by using data from students across seven MOOC providers. Our main findings suggest that age, gender, level of education, activity and performance can provide valuable information about the types of students taking MOOCs and the added value that each of the MOOC providers brings locally in the Latin American region.

We found that openSAP MOOC provider has the oldest population of learners taking their 'technical' courses (36-45 years old), and also has the highest percentage of learners with lower level of education (12.86% of their learners have only HS/JHS/EL degree), where the largest number of LATAM learners on this platform come from Brazil (only 3% of learners) and is the second MOOC provider that registers the highest percentage of learners that viewed at least half of the chapters (79,28% of learners explored the course). They also have the highest percentage of male learners.

Regarding FutureLearn MOOC provider, the most percentage of learners taking their courses are females, also has the highest percentage of learners with higher level of education (4.1% of their learners have a doctorate degree) and the largest number of LATAM learners on this platform come from Brazil and México (2% of learners respectively). Furthermore, this MOOC provider registered the highest percentage of completer learners from LATAM (54,2%).

Regarding UAMx MOOC provider, together with MITxHx and UPValenciaX registered the most percentage of young learners from LATAM (under 25 years old). Also, UAMx and PUC concentrate the most percentage of LATAM learners taking their courses (91% of the LATAM learners in PUC, 70% of LATAM learners in UAMx). However, PUC in Latin America is a benchmark for students in the region, since it has only 25 courses developed and registered more than half a million unique learners and is the third MOOC provider registering nearly 700,000 enrollments.

Previous work [12] conducted a similar regional study focused on the Arab world countries. They also found certain similarities in the distributions influenced by the course topic and so on. One key finding was that they observed a smaller gender, level of education and completion gaps in Edraak, an Arabic MOOC provider than in MITxHx. Therefore, future work should perform better comparisons on the impact of regional and global platforms in Latin American MOOC learners in the existent gaps, as well as on learners' perceptions and preferences.

This multiplatform analysis comparing MOOC providers lead us to see Latin America region as a distinct place, with a distinctive language and culture. This exploratory study shows differences between learners across Latin America MOOC providers, but also some similarities. A number of factors might be affecting these demographic similarities and differences across MOOC providers, such as the topics in the MOOC catalog, language of instruction or geographical location. These factors will be explored in future work. Despite these results, we have seen the potential of conducting analyses at a macro scale, while encouraging the community to perform more large-scale studies through partnerships between researchers and institutions to advance the field forward.

References

[1] D. Shah, "By the Numbers: MOOCs in 2019," Cl. Cent., 2019, [Online]. Available: https://www.class-central.com/report/mooc-stats-2016/.

[2] D. Shah, "By the Numbers: MOOCs in 2016," Cl. Cent., 2016, [Online]. Available: https://www.class-central.com/report/mooc-stats-2016/.

[3] N. M. Castillo, J. Lee, F. T. Zahra, and D. A. Wagner, "MOOCS for development: Trends, challenges, and opportunities," Int. Technol. Int. Dev., vol. 11, no. 2, p. 35, 2015.

[4] T. R. Liyanagunawardena, S. Williams, and A. A. Adams, "The impact and reach of MOOCs: a developing countries' perspective," eLearning Pap., pp. 38–46, 2014.

[5] T. Dillahunt, Z. Wang, and S. D. Teasley, "Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education," Int. Rev. Res. Open Distrib. Learn., vol. 15, no. 5, pp. 177–196, 2014.

[6] M. J. Kieffer, "Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States.," J. Educ. Psychol., vol. 100, no. 4, p. 851, 2008.

[7] M. Sanagustin, J. Maldonado, and N. Morales, "Estado del arte de adopción de MOOCs en la Educación Superior en América Latina y Europa," MOOC-Maker Constr. Manag. Capacit. MOOCs High. Educ. MOOC-Maker. Recuper. http//www. mooc-maker. org/wp-content/files/WPD1. 1_ESPAOL. pdf, 2016.

[8] K. Katzner and K. Miller, The languages of the world. Routledge, 2002.

[9] C. Alario-Hoyos et al., "MOOC-Maker: Tres Años Construyendo Capacidades de Gestión de MOOCs en Latinoamérica," in Proceedings of the II International Conference MOOC-Maker (MOOC-Maker 2018), 2018, pp. 4–14.

[10] J. A. Ruipérez-Valiente and J. Reich, "Participation of the Arab World in MOOCs," in 2018 Learning With MOOCS (LWMOOCS), 2018, pp. 47–50.

[11] A. García de Fanelli, "Panorama de la Educación Superior en Iberoamérica según datos de la Red Índices." Buenos Aires: OEI. http://www. redindices. org, 2018.

[12] J. A. Ruipérez-Valiente, S. Halawa, R. Slama, and J. Reich, "Using multi-platform learning analytics to compare regional and global MOOC learning in the Arab world," Comput. Educ., vol. 146, p. 103776, Jan. 2020, doi: 10.1016/j.compedu.2019.103776.