International Journal of Engineering Education Vol. 37, No. 5, pp. 1215–1230, 2021
Printed in Great Britain
0949-149X/91 $3.00+0.00

Analyzing Enrollment in Information & Communication Technology Programs and Use of Social Networks Based on Gender*

MARY LUZ MOURONTE-LÓPEZ**
Higher Polytechnic School of Universidad Francisco de Vitoria. Engineering and Women in ICT from Humanism Group. Ctra. Pozuelo a Majadahonda Km 1.800, Pozuelo de Alarcón, 28223 Madrid. Spain. E-mail: maryluz.mouronte@ufv.es

JUANA SAVALL CERES
Faculty of Education and Psychology of Universidad Francisco de Vitoria. Engineering and Women in ICT from Humanism Group. Ctra. Pozuelo a Majadahonda Km 1.800, Pozuelo de Alarcón, 28223 Madrid. Spain.

This paper aims to analyze the gender dependence in the enrollment in the bachelor's or equivalent levels, with a strong focus on the Information and Communications Technology (ICT) field. The study is carried out with 12 nations included in the Organisation for Economic Co-operation and Development (OECD). The relationships with certain socio-economic factors (such as unemployment rate, government budget allocations for Research and Development (R&D), better life index, annual wages per full-time, score in the OECD's Programme for International Student Assessment (PISA report), educational attainment, student skills, Gender Inequality Index, etc.) and other indicators related to the use of ICT are studied. Information was retrieved from the Organization for Economic Co-Operation and Development (OECD) website and from the United Nations Development Program Human Development Reports, which was elaborated in the context of the United Nations Development Programme. A statistical analysis of the enrollment and its relationships with other socio-economic and ICT variables is carried out. In order to perform the analysis, various software programs, in R programming language, were implemented. Gender dependence on the use of the Twitter and Instagram networks is also examined, in order to check whether the interactions relating to particular topics present a similar pattern to those observed in the enrollment, per field of study. For the realization of this study, several software programs were developed in Python. The implementation of these programs followed a waterfall life cycle, including requirements definition, physical implementation and testing activities. Some of the conclusions point out that there is no relationship between the number of students enrolled in ICT programs and the average score obtained by the 15-year-old students in the PISA report 2018. This research shows that countries with strong investment in research and development (Ireland, Poland, New Zealand) have a higher number of women studying ICT programs. The frequency indicators related to the use of technological and digital resources (the use of Internet daily or almost every day (%), the use of a computer (%), downloading and installing of software from the Internet (%), the creation of a web page (%), the installation or replacement of an operating system (%)) do not show significant differences between the sexes. Twitter and Instagram show a significant gender dependence, according to the topic addressed, but the pattern observed is not the same as that detected in enrollment by field of study. Research findings can help to gain a deeper understanding of the situation of women's enrolment in ICT programs, while some suggestions can also be made for the development of educational and social policies, with the intention of increasing women's participation in these programs.

Keywords: ICT; gender; equity; inclusivity; comparative study; tertiary education; social networks; technological resources

1. Introduction

In recent decades, although there has been a relevant improvement in women's access and participation at the different levels of education, differences between sexes with regards to their academic-professional choice exist [1, 2]. With respect to university studies, research shows that the engineering, information and communication technology (ICT) sector does not seem to have been as sensitive to the incorporation of women, and therefore, in most cases, continues to be a subject chosen by men. This situation, together with the growing demand for technological profiles from the labor market [3, 4], has led different supranational bodies to take a special interest in the issue. This has encouraged them to take into account indicators in their reports related to the percentage of women enrolled in the field of science, technology, engineering and mathematics (STEM). In 2017, the Organisation for Economic Co-operation and Development (OECD), which is an intergovernmental economic organisation made up of 37 nations in Europe, the Americas, and the Pacific, had an average access rate of 27.2%, including new students, undergraduates or those taking equivalent studies in the ICT area. Of this percentage only 30% were women [5]. Along these same lines, the United Nations Educa-

** Corresponding author.

* Accepted 2 April 2021.

tional, Scientific and Cultural Organization [6] points out that, in 2016, the representation of women in higher education was especially low in ICT (3%), natural sciences, mathematics and statistics, (5%) as well as engineering, manufacturing and construction (8%), displaying higher numbers in health and well-being (15%). These gender disparities are highly alarming since the careers related to these fields are considered to be the jobs of the future, the driving force for innovation and sustainable development.

If the scientific literature is reviewed, a difference in the interest levels of girls and boys in relation to engineering and technical subjects has already been found in the first courses of secondary education [7]. Lower female presence in the ICT field seems not to be based on cognitive capacity, but on the socialization processes in which the identity of the girls, their beliefs and their life choices are being configured [6]. The research describes reasons why women do not opt for these studies, identifying factors such as the absence of female role models [8], the existence of certain stereotypes [9, 10] or the absence of adequate orientation at the end of the mandatory studies can be cited as possible causes [11–13].

Gender, socioeconomic and institutional factors affect the academic performance of boys and girls [14]. OECD's Programme for International Student Assessment (PISA) 2018 [15] indicates that girls have significantly higher performance than boys in reading, but that boys have better results than girls in mathematics, also showing that the distribution of performance in science is more dispersed than in girls. Performance differences may be partially related to how girls and boys socialize, both at home and at school [16]. PISA 2018 [15] notes that socioeconomic status is a "strong predictor" of performance in mathematics and science in all participating countries. Comparable international data is required to ensure evidence-based planning and policy formulation, as well as future documentation on the effectiveness and impact of the interventions which are being carried out.

The purpose of this research is to study more deeply the gender dependency in the enrollment within bachelor's or equivalent levels, with relevant focus on ICT programs. It will also detect which may be the most influential factors. Based on the findings obtained, some suggestions could be made in order to implement educational and social actions. Twelve member countries of the OECD were selected for the study: Spain, Canada, Estonia, Finland, Germany, Ireland, Korea, New Zealand, Poland, Sweden, the U.K. and the U.S.A. The reason for the selection of these countries was that all of them, except Spain, are located in the top 11 positions of the PISA 2018 report [15]. More specifically, the goals of this research are:

1. To analyse the total enrollment in undergraduate studies (total and determined by sex), relating it to other data (scores in PISA 2018 [15], social and economic information). Examination of the gender dependency per field of study.
2. To study, in more detail, the enrollment in bachelor's or equivalent levels for ICT field, correlating data in technological careers with socio-economic variables and with gender indicators.
3. In order to check whether what is observed in the interactions on certain topics, matches the gender pattern noted in the fields of study, to analyze the gender influence on the use of Twitter and Instagram.

2. Material and Methods

2.1 Overview of Used Resources for the Study

2.1.1 Data Repositories

The data available on the website of the Organization for Economic CoOperation and Development (OECD) [17] are used. Specifically, the datasets:

- Enrolment by field.
- Better life index.
- Access and Usage by Households and Individuals.
- Annual labour force statistics (ALFS).
- National accounts at a glance. Gross Domestic Product (GDP).
- Educational finance indicators: total expenditure on educational institutions per full-time equivalent student relative to GDP per capita.

Data from the United Nations Development Program Human Development Reports [18] are also used.

2.1.2 Software Program

To carry out the statistical study, various software programs in programming language R [19] were implemented. These programs allowed the calculation of correlations, the execution of statistical tests, as well as the clustering studies detailed in section 2.2. This is in addition to the graphical representations that illustrate the results of this research. For the connection to social networks Twitter and Instagram, several software programs were developed in the programming language Python [20]. These programs applied web scraping to obtain data about the interactions that used certain "hashtags", as well as to record the tweets and profiles of the users who interacted. The implementation of these programs followed the typical waterfall life cycle of any software component, including requirements definition, physical implementation, and testing activities. The i.genderize [21] Application Programming Interfaces (API) was also used in order to obtain the sex of users involved in the interactions.

2.2 Overview of Used Method for the Study

2.2.1 Analysis of Volume of Students in Bachelor's or Equivalent Levels

The International Standard Classification of Education (ISCED) 2011 typifies the tertiary education at four levels: short cycle tertiary education (level 5), tertiary education degree or equivalent (level 6), master's level, specialization or equivalent (level 7) and doctorate level or equivalent (level 8). This section studies the volume of students enrolled in level 6. The nations chosen for the analysis are, as mentioned above, Spain, Canada, Estonia, Finland, Germany, Ireland, Korea, New Zealand, Poland, Sweden, the U.K and the U.S.A. All of them, except Spain, are in the top 11 positions of the PISA 2018 report [15]. The position refers to the average scores achieved, with a 95% level of confidence, in science, mathematics and reading comprehension. The study is carried out in the years 2013, 2014, 2015, 2016 and 2017. In the aforementioned countries, to allow identification of possible trends, the average percentage of students enrolled by field of study is analyzed each year. The fields of study are:

- Agriculture, forestry, fisheries and veterinary (AFFV).
- Arts and humanities (AH).
- Health and Welfare (HW).
- Social Sciences, Journalism and Information (SSJI).
- Services (SER).
- Business, Administration and Law (BAL).
- Information and Communication Technologies (ITC).
- Natural Sciences, Mathematics and Statistics (NSMS).
- Engineering, Manufacturing and Construction (EMC).
- Generic Programmes and Qualifications (GPQ). – Education (EDU).

For all OECD countries, in the indicated years, relationships are also identified, and linear correlation coefficients are calculated [22] between the total of students enrolled in ISCED2011 level 6 and the following variables [17]:

- Government budget allocations for R&D in PPP dollars current-prices (GPRD).
- Employees working very long hours. Percentage of company workers whose usual hours of work per week are 50 hours or more (EWLH). Both sexes.
- Unemployment in thousands of persons (UNEM). Both sexes.
- Population (POPU). Both sexes.
- Enrolment in number of persons (ENRO). Both sexes.
- Life satisfaction (LISA) on a scale from 0 to 10, using the Cantril Ladder Scale [23]. Both sexes.
- The average annual wages per full-time equivalent dependent employee. It is presented in US dollars at 2011 prices (PEEA). Both sexes.
- Educational attainment. It considers the number of adults aged 25 to 64 holding at least an upper secondary degree over the population of the same age, as defined by the OECD-ISCED classification (EATT). Both sexes.
- Percentage of total expenditure on educational institutions per full-time equivalent student relative to GDP per capita (GDPE).
- Student skills. Students 'average score in reading, mathematics and science as assessed by PISA (STSK). Both sexes.

The Pearson correlation or the Spearman correlation [24] will be used, depending on whether the variables have a normal distribution or not. The normality of the distribution will be studied using the AndersonDarling test [25].

The identification of other types of relationships between variables will be made through hierarchical clustering algorithms which, using multivariate techniques, will group the variables with maximum homogeneity within clusters and will identify, based on a distance d, the proximity relationships [26]. Specifically, the following methods are used [27]:

- Single Linkage: the distance between two clusters is the distance between the closest points.

$$dmin(A,B) = x_i \in A, x_j \in B\ d(x_i, x_j) \quad (1)$$

  $dmin(A,B)$ Euclidean distance between two points i, j.

- Complete Linkage: the distance between two clusters is the distance between the furthest points.

$$dmin\ (A,B) = x_i \in A, x_j \in B\ d(x_i, x_j) \quad (2)$$

  $dmin(A,B)$ Euclidean distance between two points i, j.

- Average Linkage: The distance between two clusters is the average of distances between all pairs of members.

$$daverage\ (A,B) = \frac{1}{|A||B|} \sum_{x_i \in A, x_j \in B} \qquad (3)$$

$$d(x_i, x_j) = d(\mu_A, \mu_B)$$

μ_A: centroide of A; μ_B: centroide of B; $dmin(x_i, x_j)$: Euclidean distance between two points i, j.

- Ward [28]: the distance between clusters A and B is the increase in sum of squared errors (SSE) when the two clusters are merged.

$$d\ (A,B) = \frac{|A||B|}{|A|+|B|} |\mu_A - \mu_B| \qquad (4)$$

μ_A: centroide of A; μ_B: centroide of B.

To evaluate the clustering algorithms used, firstly, the Silhouette Coefficient (SC) will be calculated, followed by the Dunn Index (DI) [29]. The number of clusters for which the optimal values are achieved will be taken into account. Other considerations:

- Before calculating correlations and applying the hierarchical clustering algorithm, a normalization is carried out on each variable x, MinMax:

$$(x - \min(x))/(\max(x) - \min(x)) \qquad (5)$$

min(x): smallest value of the variable x. max(x): highest value of variable x.

- When in a country, a variable x has missing information for a certain year, its average in that year is taken as the value of the variable.

2.2.2 Study of Dependence on Gender per Field of Study in Education Bachelor's or Equivalent Levels

This section studies whether the number of students enrolled by field of study is dependent on their sex. The steps followed are:

1. Checking for the existence of normality using the Anderson-Darling test [25]. The hypotheses are:
   - Null hypothesis (H_0): "Samples came from a normal distribution".
   - Alternative hypothesis (Ha): "Samples did not come from a normal distribution".
2. Verification of homoscedasticity. If there is normality, the homogeneity of variances will be verified through the Fligner-Killeen test [30], considering the hypotheses:
   - Null hypothesis (H_0): "The samples have a distribution with constant variance".
   - Alternative hypothesis (Ha): "The samples have a non-constant variance distribution".
3. If assumptions 1 and 2 are met, the Analysis Of VAriance (ANOVA) [22] will be used as a hypothesis test to verify sex dependency. Otherwise, the non parametric test of Kruskal-Wallis will be employed [30]. The following hypotheses were considered:
   - Null hypothesis (H_0): "The gender of the students does not impact relevantly on the variable".
   - Alternative hypothesis (H_a): "The gender of the students impacts on the variable".

All the tests were applied with a confidence level of 95%.

2.2.3 Analysis of Volume of Students in Bachelor's or Equivalent Levels in ICT Field

The mean in each of the analyzed years of the total number of women who are enrolled in ISCED2011 level 6, ICT field, is estimated. The linear correlation coefficients between this variable and the variables detailed in section 2.2.1 are calculated. Clusters are also estimated according to the algorithms presented in that section. Additionally, the following variables are studied:

- Gender Inequality Index (GII).
- Total unemployment rate (female to male ratio) (UR).

2.2.4 Gender influence in the values of some social indicators related to ICT

The following indicators were studied [13]:

- Individuals using the Internet daily or almost every day – last 3 m (%) (C5B1).
- Individuals using a computer – last 3 m (%) (C2B).
- Individuals who have found, downloaded and installed software from the Internet – last 12 m (%) (H1H).
- Individuals who have written computer code – last 12 m (%) (H1K).
- Individuals who have created a web page – last 12 m (%) (H1L).
- Individuals who have installed or replaced an operating system – last 12 m (%) (H1M).

For the study, the same procedure indicated in section 2.2.2 is used.

2.2.5 Gender Influence in the use of Social Networks

This research analyzes whether the use of Twitter and Instagram on the subjects of engineering, technology, mathematics, science, architecture, fine arts, humanities and health sciences is dependent on the sex of the users. We study, in these networks, those interactions that have as a "hashtag" any of the indicated subjects. The procedure consists of carrying out, for each "hashtag", 100 connections for each social network. 20,000 interactions are extracted in each connection and each connection constitutes an experiment. In total, 2,000,000 interactions per "hashtag" are studied and processed in each network. Once the profiles of the users who interact on each "hashtag" are obtained, the users are identified, and their sex is deduced. The procedure explained in section 2.2.2 is carried out.

3. Results

The results are shown in the Tables 1–6, as well as in the Figs. 1–5.

For the years 2013–2017, Table 1 shows the Spearman correlation of the variable ENRO with

Table 1. Considering 2013, 2014, 2015, 2016 and 2017 year, Spearman correlation between Enrolment in number of persons (ENRO) and Total expenditure on educational institutions as a percentage of GDP (GDPE), Life satisfaction (LISA), Population (POPU), Unemployment (UNEM), Educational attainment (EATT), Student skills (STSK), Personal earnings (PEEA), Employees working very long hours (EWLH), Government budget allocations for R&D (GBRD).

Year	GDPE	LISA	POPU	UNEM	EATT	STSK	PEEA	EWLH	GBRD
2013	0.21276	–0.20169	0.95514	0.85758	–0.12817	–0.0577	–0.06513	0.31883	0.82875
2014	0.20822	–0.15594	0.95695	0.87946	–0.13589	0.01146	–0.11262	0.31821	0.86029
2015	0.21340	–0.16599	0.95489	0.88879	–0.14446	0.01610	–0.15562	0.33569	0.86196
2016	0.22296	–0.17576	0.95661	0.92541	–0.21310	0.01510	–0.14467	0.39619	0.84411
2017	0.24429	–0.12230	0.95283	0.93352	–0.17459	–0.06839	–0.07182	0.37480	0.88332

Table 2. Considering 2013, 2014, 2015, 2016 and 2017 year, Spearman correlation between Student skills (STSK), Educational attainment (EATT) and Total expenditure on educational institutions as a percentage of GDP (GDPE).

Year	EATT	GDPE
2013	0.36143	0.39392
2014	0.38741	0.34005
2015	0.46277	0.28680
2016	0.42145	0.29622
2017	0.42145	0.35882

Table 3. In ISCED2011 level 6, values for p-value in Kruskal-Wallis test per field of study, Agriculture, Forestry, Fisheries and Veterinary (AFFV), Arts and Humanities (AH), Health and Welfare (HW), Social Sciences, Journalism and Information (SSJI), Services (SER), Business, Administration and Law (BAL), Information and Communication Technologies (ITC), Natural Sciences, Mathematics and Statistics (NSMS), Engineering, Manufacturing and Construction (EMC), Generic Programmes and Qualifications (GPQ), Education (EDU), in 2013, 2014, 2015, 2016 and 2017 years.

Year	2013	2014	2015	2016	2017
AFFV	0.81821	0.79885	0.85724	0.82859	0.91952
AH	0.16329	0.15443	0.14593	0.15721	0.16586
SJI	0.16809	0.16030	0.13780	0.14357	0.15721
SER	0.64931	0.66375	0.67088	0.66499	0.72902
HW	0.00140	0.00259	0.00476	0.00426	0.00440
UNK	0.95898	0.94431	0.81847	0.93102	0.81392
BAL	0.57769	0.55211	0.55211	0.56370	0.58336
ICT	0.00011	0.00011	0.00019	0.00008	0.00009
NSMS	0.95259	0.98804	0.98804	0.97697	100.000
EMC	0.00821	0.00708	0.00708	0.00670	0.01140
GPQ	0.95459	0.66324	0.92102	0.76676	0.92941
EDU	0.00555	0.00647	0.00555	0.00453	0.00511

Table 4 Considering 2013, 2014, 2015, 2016 and 2017 year, Spearman correlation between female enrolment in Information and Communication Technologies (ITC) and Total expenditure on educational institutions as a percentage of GDP (GDPE), Life satisfaction (LISA), Population (POPU), Unemployment (UNEM), Educational attainment (EATT), Student skills (STSK), Personal earnings (PEEA), Employees working very long hours (EWLH), Government budget allocations for R&D (GBRD).

Year	GDPE	LISA	POPU	UNEM	EATT	STSK	PEEA	EWLH	GBRD
2013	0.30331	–0.14891	0.82515	0.70185	–0.00458	–0.05616	–0.12242	0.36465	0.50171
2014	0.30876	–0.05207	0.80378	0.69927	0.11178	0.11793	–0.08637	0.28809	0.75372
2015	0.23628	–0.08339	0.82411	0.75550	0.04561	0.08691	–0.13013	0.29740	0.79280
2016	0.30331	–0.14900	0.82515	0.70185	–0.00460	–0.05616	–0.12242	0.36465	0.50171
2017	0.37007	0.00271	0.55468	0.60398	0.05302	0.11702	–0.27007	0.37201	0.47383

Table 5. Considering 2013, 2014, 2015, 2016 and 2017 year, Spearman correlation between female enrolment, and Gender Inequality Index (GII), Total unemployment rate (female to male ratio) (UR).

Year	GII	UR
2013	0.27372	–0.01727
2014	0.18815	0.04811
2015	0.21144	0.11173
2016	0.27372	–0.04971
2017	0.17185	–0.02795

the variables GDPE, LISA, POPU, UNEM, EATT, STSK, PEEA, EWLH, GBRD. For the same years, Table 2 presents the Spearman correlation of the variable STSK, with the variables EATT and GDPE.

Table 3 shows in ISCED2011 level 6, the p-value obtained in the Kruskal-Wallis test according to the field of study (AFFV, AH, HW, SSJI, SER, UNK, BAL, ICT, NSMS, EMC, GPQ, EDU), in the years 2013–2017. In the same years, Table 4 presents the Spearman correlation between the female enrollment in ICT and the variables GDPE, LISA, POPU, UNEM, EATT, STSK, PEEA, EWLH, GBRD. Table 5 displays, in the same years, the Spearman correlation between the female enrollment variable and the variables GII and UR.

Table 6 presents the p-value obtained in the Kruskal-Wallis test for the variables C5B1; C2B; H1H; H1K; H1L and H1M, in the years 2013–2017. Table 7 demonstrates, in the same years, on Twitter and Instagram networks, for interactions on Engineering, Technology, Mathematics, Science, Architecture, Fine Arts, Humanities, and Health Sciences subjects. the p-value obtained in the Kruskal-Wallis test.

Fig. 1 shows in ISCED2011 level 6, in the years 2013–2017, the average percentage of students per country and field of study (AFFV, AH, HW, SSJI, SER, UNK, BAL, ICT, NSMS, EMC, GPQ, EDU). Fig. 2 presents, in 2013, the Spearman correlation and clustering between the variables ENRO, GDPE, LISA, POBL, UNEM, EATT, STSK, PEEA, EWLH, GBRD, using Average Linkage, Single Linkage, Average and Ward methods.

Fig. 3 lays out, at ISCED2011 level 6, the evolution of the average percentage of women's enrollment in the ICT field of study in the years 2013–2017.

Fig. 4 presents the correlation and clustering between variables ENRO for women in the ICT field, GDPE, LISA, POPU, UNEM, EATT, STSK, PEEA, EWLH, GBRD, GII, UR, in 2016 and 2017. The results obtained applying the Average Linkage and Ward methods are shown.

Finally, Fig. 5 depicts the boxplot diagram for the interactions on technology in Twitter and Instagram networks.

4. Discussion

4.1 Analysis of Volume of Students in Bachelor's or Equivalent Levels

This section studies the number of students in the level of education bachelor's or equivalent

Table 6. In social indicators, p-value in Kruskal-Wallis test for Individuals using the Internet daily or almost every day – last 3 m (%) (C5B1), Individuals using a computer – last 3 m (%) (C2B), Individuals who have found, downloaded and installed software from the Internet – last 12 m (%) (H1H), Individuals who have written computer code – last 12 m (%) (H1K), Individuals who have created a web page – last 12 m (%) (H1L), Individuals who have installed or replaced an operating system – last 12 m (%) (H1M).

Year	5B1	C2B	H1H	H1K	H1L	H1M
2013	0.89177	0.96791	–	–	0.10794	–
2014	0.99198	0.52873	0.98544	0.00617	0.97914	0.01865
2015	0.55209	0.84876	0.48864	0.01452	0.98544	0.98544
2016	0.67944	0.99314	0.30020	0.01224	0.95830	0.98544
2017	0.98799	0.55767	0.33076	0.00078	0.95830	0.98544

Table 7. In Twitter and Instagram networks, p-value in Kruskal-Wallis test.

Matter	Twitter	Instagram	Field of study
Engineering	2.31e-19	0.01267	EMC
Technology	6.37e-06	0.01100	ITC
Mathematics	9.87e-20	0.00858	NSMS
Science	0.00270	0.51811	NSMS
Architecture	0.26590	0.12805	EMC
Fine Arts	4.71e-05	2.36e-07	AH
Humanities	1.88e-07	0.00115	AH
Health Sciences	0.12431	0.01298	HW

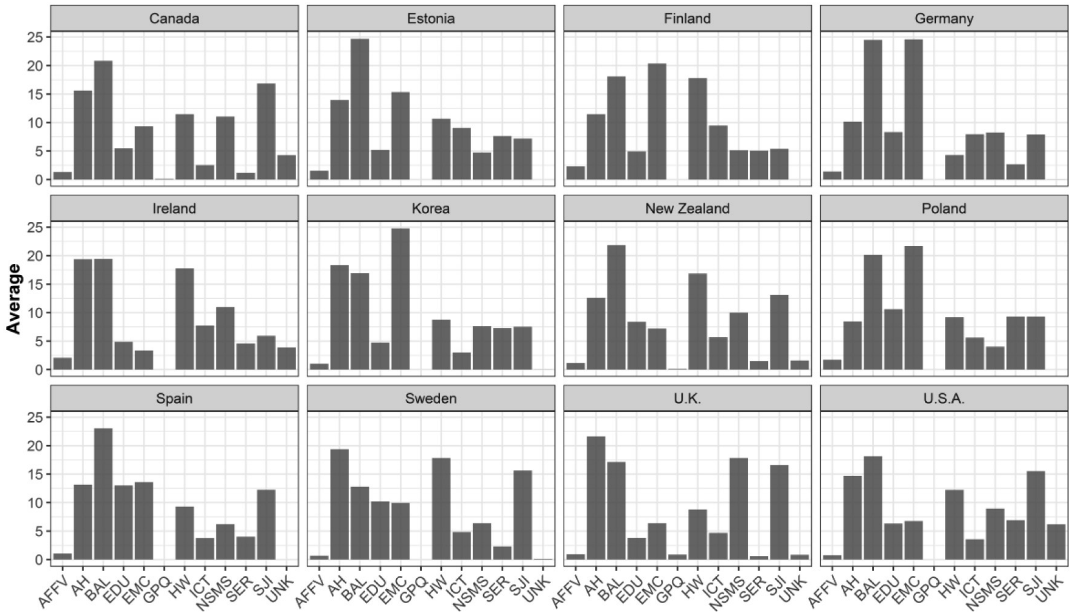


Fig. 1. Considering 2013–2017 year, in ISCED2011 level 6, average of percentage of students per field of study: Agriculture, forestry, fisheries and veterinary (AFFV), Arts and Humanities (AH), Health and Welfare (HW), Social Sciences, Journalism and Information (SSJI), Services (SER), Business, Administration and Law (BAL), Information and Communication Technologies (ITC), Natural Sciences, Mathematics and Statistics (NSMS), Engineering, Manufacturing and Construction (EMC), Generic Programmes and Qualifications (GPQ), Education (EDU).

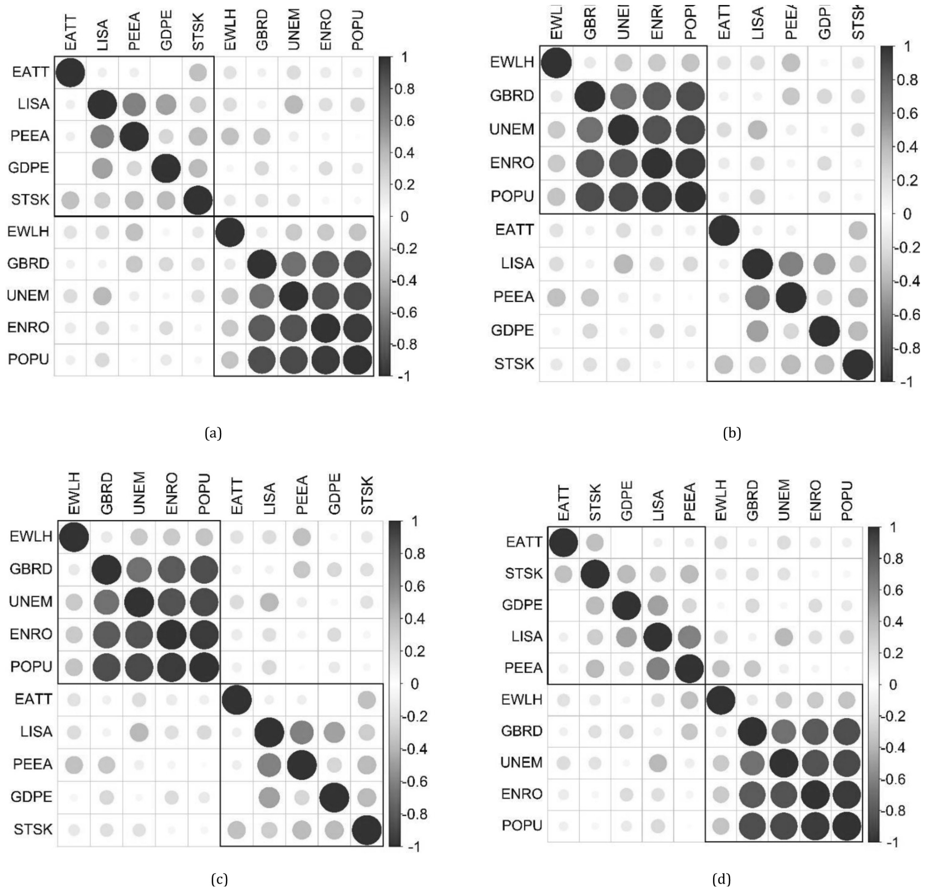


Fig. 2. In 2013, Spearman correlation and clustering between variables: Enrolment in number of persons (ENRO), Total expenditure on educational institutions as a percentage of GDP (GDPE), Life satisfaction (LISA), Population (POPU), Unemployment (UNEM), Educational attainment (EATT), Student skills (STSK), Personal earnings (PEEA), Employees working very long hours (EWLH), (Government budget allocations for R&D (GBRD), using Ward method (a) Complete Linkage method (c) Average Linkage method (d) Single Linkage method.

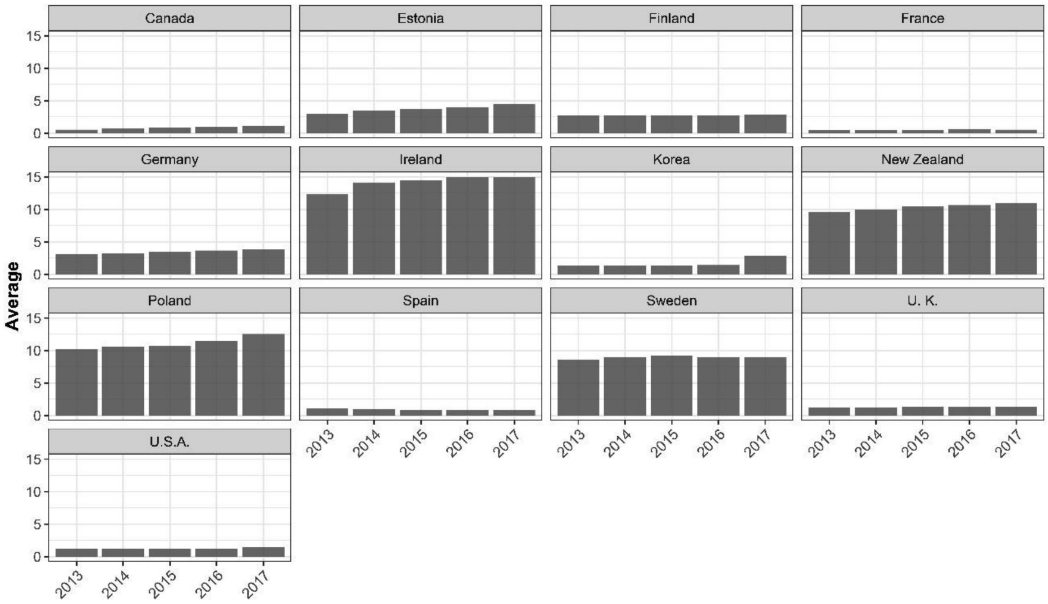


Fig. 3. In level ISCED2011 level 6, in Canada, Estonia, Finland, France, Germany, Ireland, Korea, New Zealand, Poland, Spain, Sweden, U.K. and U.S.A., average of percentage of female students in the field of study Information and Communication Technologies (ICT), per years 2013–2017.

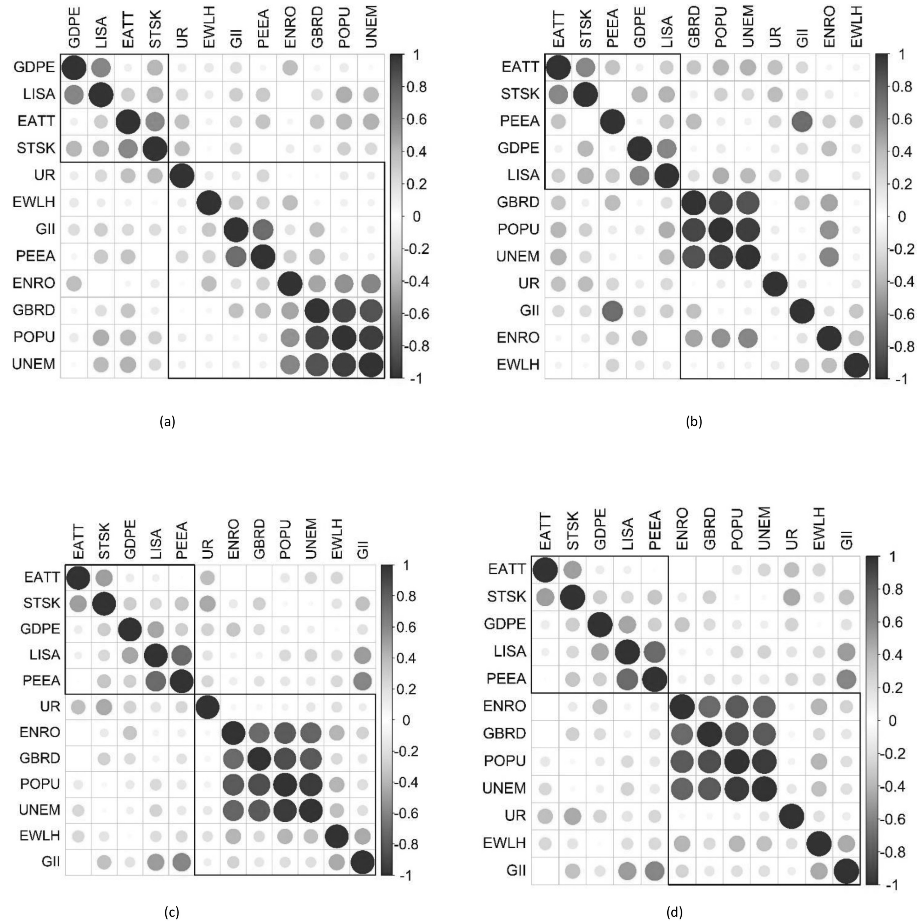
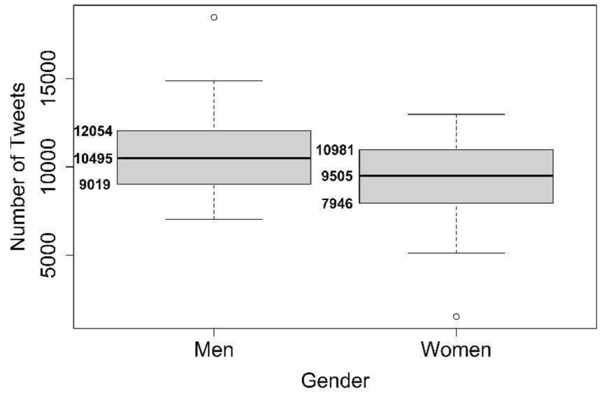
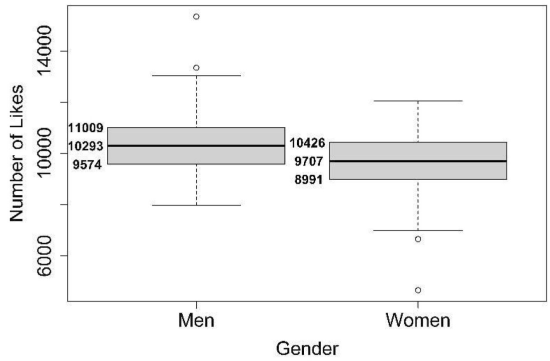


Fig. 4. In 2016 and 2017, Correlation and clustering between variables Enrolment in number of persons (ENRO) for women in ICT field, Total expenditure on educational institutions as a percentage of GDP (GDPE), Life satisfaction (LISA), Population (POPU), Unemployment (UNEM), Educational attainment (EATT), Student skills (STSK), Personal earnings (PEEA), Employees working very long hours (EWLH), Government budget allocations for R&D (GBRD), Gender Inequality Index (GII), Total unemployment rate (female to male ratio (UR). In year 2017, using Average (a) and Ward (b) methods. In year 2016, using Average (c) and Ward (d) methods.



(a) Twitter



(b) Instagram

Fig. 5. In Twitter (a) and Instagram(b), boxplot diagram for technology field. Median of tweets and likes is shown.

(ENRO). The highest mean percentages are found in the fields of study BAL (Canada, Estonia, Ireland, New Zealand, Spain, U.S.A, Germany) and EMC (Finland, Korea, Poland, Germany). In these maximums all countries exceed 18%. The case of Sweden and the U.K. is different, since the maximum average magnitude occurs in the AH field with 19.35% and 21.61%, respectively. In all the countries analyzed, the lowest average magnitudes are in the AFFV field, except in Canada and the U.K., where it is in the SER field. The average percentage calculated over all the years analyzed by field of study and country is shown in Fig. 1. The average percentage calculated over all the years analyzed by field of study and country, presents a standard deviation in the range [0%, 5.91%].

As Table 1 shows, the linear correlation of total student volume is strong with POPU, UNEM, and GBRD variables. There is a moderate correlation with the EWLH variable and a weak correlation with the GDPE, LISA and EATT variables. There is no correlation with the rest of the variables. One of these variables is STSK, which shows that there is no linear dependence between the volume of students in bachelor's or equivalent levels and their abilities. The strong correlation with the GPRD variable indicates that, in the short term, changes in the R&D budget have a high impact. However, modifications in the budget for education show a weak influence. We also note that the LISA variable presents an average value, calculated over all OECD countries and all analyzed years of 6.6 out of 10, with a standard deviation $\pm$ 0.79705, which represents a medium level of life satisfaction. This can explain the low impact of the LISA variable on the enrollment. The high correlation with the variables POPU and UNEM shows a high impact to variations in the population and unemployment. The moderate correlation with jobs that require working 50 or more hours per week is also significant, which suggests that enrollment in these studies can be seen as a chance to achieve other higher level occupations. For example, those that may require less dedication. The level of correlation between the analyzed variables is maintained through the studied years 2013, 2014, 2015, 2016 and 2017.

The Anderson-Darling test shows that there was no normality in the ENRO variable p-value ≤ 0.05, therefore the Spearman correlation is used. Table 2 shows a moderate correlation between STSK and EATT variables. This seems to indicate that the highest level of education held by the population aged between 25 and 64 years has the most significant influence on the best skills of the 15 year old students. The moderate correlation of STSK with GDPE predominates throughout the studied years. In this sense, the PISA results suggest that above a certain threshold of GDP spending on education per capita, there is no impact on the average performance of students in a country [31]. The best-performing countries and economies, among those with the highest incomes, tend to invest more in teachers' qualification, wages and recruitment. In Korea and Hong Kong-China, countries with high performance on the PISA tests, which prioritize the quality of their teachers and have high expectations of their students, the teachers earn more than double the GDP per capita. However, the relationship between teachers' performance and salaries is not maintained among countries with less wealthy economies. Education expenditure only, does not guarantee a good educational system.

As observed in Fig. 2, the hierarchical clustering for all the methods used, points to the existence of a relationship between the variables ENRO, UNEM, GPRD, EWLH, and POPU, since they are all grouped in the same cluster. This occurs for all the years analyzed. The best SC and DI values are obtained for two clusters. For all the methods studied, the SC value in 2013 is 0.5334 and the DI value is 0.9927.

4.2 Study of Dependence on Gender per Field of Study in Bachelor's Education or Equivalent Levels

This section studies whether the volume of students in each field of study depends on their sex. The Anderson-Darling test shows that there was no normality in the variable p-value $\leq$ 0.05. The Fligner-Killeen test returns a (p-value = 1), with homogeneity of variances. Taking the above into account, we used the Kruskal-Wallis non-parametric test. In Table 3 we observe that there is sex dependency (p-value $\leq$ 0.05) in the fields of study HW, ICT, EMC and EDU.

4.3 Analysis of Volume of Students in Bachelor's or Equivalent Levels in ICT Field

This section studies the volume of students enrolled in ICT programs and its relationship with the socioeconomic data showed in the section 2.2.1. Because the qualitative research can delve deeper into the motivations and persons' opinions about the choice of a university degree, rather than the quantitative research [32], the results are related to other existing qualitative research.

The graduated students in the field of ICT in 2015 represented 3.6% of all graduates [33]. According to our research, as can be observed in Fig. 1, considering the average percentage of students per field of study in 2013, 2014, 2015, 2016 and 2017, the countries with the highest volume are Finland with 9.47% and Estonia with 9.06%. The countries with the lowest percentage are Canada with 2.52% and Korea with 3.00%. These percentages may be a consequence of the fact that countries with good results in the PISA report give more importance to teacher training in these subjects, as is the case in Estonia. The University of Tallinn has a specialized center to train teachers who teach in the ICT field. Estonia is considered one of the most digitized societies and one of the leading countries in information technology. Finland includes ICT knowledge among the competences of Primary Education, integrating them in the rest of the subjects. It should be noted that those OECD countries with the highest average PISA score, Estonia (525.33), Korea (519.67), Canada (516.67) and Finland (516.33) are not the nations that show the highest average percentage of students enrolled in the ICT field.

If we look at the percentage of women, which is shown in Fig. 3, the highest value is found in Ireland with 14.19%, Poland with 11.09% and New Zealand with 10.32%. The lowest values correspond to Canada and Spain with 0.85% and 0.90%, respectively. The standard deviation obtained from the average percentage of women enrollment in the ICT field of study during the analyzed years, is in the range [0.04%, 1.08%]. Regarding Ireland, its location, its technological infrastructure, its 25% R&D tax incentive and its low corporate taxes, have encouraged many multinational technological companies to set up in the country. The Irish government recently published collaborative actions between government, education system and industry, the ICT Skills Action plan (2014–2018) [34], in order to enhance and attract talent in the area of computing, software and electronic engineering. This country, in accordance with its digital strategy for educational centers, made very important investments between 2015 and 2020, which enabled the funding for the modernization of relevant infrastructures in schools [35]. In Poland, the students learn mandatory subjects in technology and computer science from the age of 10. New Zealand's ICT companies [36] have an international reputation, with salaries in the ICT area twice the national average. Estonia and Poland [37], in recent years, have also significantly improved their economic level and increased their investments in the ICT sector. The former seems to have favorably affected enrollment in higher education in the ICT field.

As can be seen in Table 4, the volume of women enrolled in the ICT field shows a strong correlation with the variables POPU and UNEM in all the years analyzed. The correlation with GBRD is high in 2013, 2014, 2015 and 2016 and moderate in 2017. With EWLH there is an average correlation in 2013, 2016 and 2017, but is evidently weaker in 2014 and 2015. With GDPE there is an average correlation in 2013, 2014, 2016 and 2017, but weak in the year 2015. The correlations with the GDPE variable are slightly higher than in the case of the total volume of students enrolled for both sexes. The linear correlation with the rest of the variables is weak or non-existent. In this sense, with respect to life satisfaction, some qualitative research, [38] shows that students with disadvantaged socio-economic backgrounds perceive the STEM careers as an opportunity to improve their standing at a social level.

Furthermore, we highlight that the variable of enrolled women in the ICT field that shows low, or no correlation is STSK. This is in line with the conclusions of the PISA 2018 report, [15] which indicates that, on the OECD average, although boys outperform girls in mathematics by 5 points, the girls are ahead by 2 points in science. Meanwhile, 10% of children expect to work in a profession related to ICT, and the percentage of women is only 1%. One in three boys want to be an engineer or scientist at age 30, while only one in five girls wish to carry out these professions in the future. [39] Through using information collected in a survey and recorded written answers, results suggest that, in spite of a similar opportunity of choice existing for both sexes, boys consider themselves to be more successful in science subjects and wish to learn them more than girls. Furthermore, boys adapt to STEM careers more easily than girls.

In accordance with the above, the sex difference between students who hope to develop scientific or technological careers or enroll in ICT programs do not depend on the academic performance. In this sense, the qualitative research [40] shows, collecting what the students expressed in focus group discussions, that there are no differences between male and female ability, and suggests that the messages in the family environment have a significant influence on whether certain careers are associated with one sex or the other. Also, qualitative research [41] identified in the narratives of the high school students that the interest of such students' in choosing certain university studies changes over the time, due to the influence of the close environment. [38] Using a qualitative analysis demonstrates the importance of the family as a driving force towards enrollment in STEM studies. This is particularly noteworthy in the case of women. Therefore, it is not only necessary to promote an increase of female presence in STEM areas, but also to work further to reduce the gender stereotypes [42].

Additionally, we also analyze the linear correlation of the variable of enrolled women in the ICT field with the variables GII and UR. As observed in Table 5, the correlation between female enrolment, with GII is weak and with UR, depending on the year, is weak or non-existent. We note that these countries, on average, have low inequality and a similar unemployment rate between men and women. The GII has an average value, calculated over all the years analyzed, of 0.12621 with a standard deviation of $\pm$ 0.08399, and the UR of 1.02706 with a standard deviation of $\pm$ 0.17071. [43] studies the so-called paradox of equality in science and engineering studies and concludes that the more gender equality there is in a country, according to the World Economic Forum's Gender Equality Index, the lower the percentage is of women studying engineering and technical careers. According to our study, in the OECD countries, the variable of women enrolled in the ICT field has a non-existent or low relationship with the GII variable. Furthermore, the changes in it do not seem to have an impact, and it should be observed that in these countries on average, a low level of gender inequality exists.

Regarding hierarchical clustering, which is depicted in Fig. 4, in 2017, the best results were obtained, taking into account firstly the SC, and then the number of the cluster, with the Ward and Average containing three clusters of an SC equal to 0.3686, and a DI equal to 0.5292. In 2016, something similar occurs with an SC 0.4655 and DI 0.7212. In 2015 and 2014, better results are also obtained for the aforementioned algorithms, and three clusters. In 2013, something identical happens, but with two clusters. In all the optimal cases, the variables ENRO, POPU, UNEM and GPRD appear located in the same cluster, which indicates a sustained relationship between all the analyzed years.

4.4 Gender Influence in the Values of some Social Indicators related to ICT

A study of whether a sex dependency exists in the indicators 5B1, C2B, H1H, H1K, H1L and H1M in the age range 16–24 years, during the years analyzed. The Anderson-Darling test does not show normality in its distributions (p-value ≤ 0.05). The Fligner-Killeen test gives a p-value = 1, for all the indicators, with homogeneity of variances (p-value $\geq 0.05\%$). Considering this, we use the Kruskal-Wallis non-parametric test. Table 6 describes that there is only sex dependency (p-value ≤ 0.05) in the H1K indicator. This indicator is the one that

Table 8. In Twitter and Instagram networks, for women and men, median of tweets and likes.

Matter	Twitter		Instagram		Field of study
	Women	Men	Women	Men	
Engineering	8,294	11,706	10,400	9,600	EMC
Technology	9,505	10,495	9,707	10,293	ITC
Mathematics	8,971	11,029	9,709	10,291	NSMS
Sciences	9,907	10,193	9,709	10,291	NSMS
Architecture	10,129	9,871	10,291	9,709	EMC
Fine Arts	10,471	9,529	9,509	10,491	AH
Humanities	8,971	11,029	10,471	9,529	AH
Health Sciences	10,400	9,600	9,709	10,291	HW

requires the most knowledge about programming techniques and software design methodology.

We can conclude, therefore, that, according to most of the indicators analyzed, girls and boys, at the user level, frequently use ICT in a similar way. This conclusion is in line with the research of [44] and [45], which shows that girls are as frequent users of ICT as boys and that one of the reasons may be that digital technology is the preferred platform for communication between young people today. Certain qualitative research has delved into the analysis of ICT use according to sex [46, 47] explain that the efforts made in the development of academic contents decreased the gap in the ICT knowledge between girls and boys, but it was unable to eliminate the gender differences in the attitudes. [47] shows that no differences are detected in the use of ICT in terms of ideas and knowledge, income and educational background, however, men seem to be more inquisitive, and fast in operating with the ICT.

4.5 Gender Influence in the use of Social Networks

Several authors have studied the use of social networks by men and women, explaining that it shows notable differences. [48] analyzes the differences in terms of general activity, cross-gender interaction, communication style and network structure, [49] investigates gender differences in people's decisions about used information in the context of social networking sites. [50] also explains that females use Facebook for maintaining existing relationships, academic purposes and following agendas more than males. In contrast, males employ it for making new relationships at a rate higher than females. Our research, based on the 2,000,000 interactions for each "hashtag", analyzes whether the volume of dialogue on social networks Twitter and Instagram, on engineering, technology, mathematics, science, architecture, fine arts and humanities, as well as about health sciences shows sex dependence.

According to the Anderson-Darling test, only the variables corresponding to interactions about engineering and mathematics on Instagram showed normality in the distributions (p-value > 0.05). The Fligner-Killeen test showed homogeneity of variances (p-value ≥ 0.05). With these considerations, we use the Kruskal-Wallis non-parametric test, which would indicate, if p-value ≤ 0.05, that there is sex dependency, the results are shown in Table 7.

The results indicate that the use of Twitter and Instagram is strongly dependent on sex according to the topics analyzed, but this difference between the sexes depends on the network and it does not generally coincide with what was observed in the analysis. of the fields of study, as shown in Table 8. On twitter there is no dependency on sex in architecture and health sciences. On Instagram there is no dependency on sex in architecture and science. The SSJI, NSMS study fields displayed no sex dependency. Regarding technology, gender dependency is similar to that observed in the ICT field of study in both networks, with women showing a lower median than men. It is shown in Fig. 5. that the different way of using social networks is in accordance with some research such as [51] and [52]. [53] identifies using both quantitative and qualitative methods results, in a range of different styles and topical interests per sex in Twitter. [54] using qualitative research shows that male as well as female contributors to hashtags reproduce examples in debates on Twitter of the discourses that dictates their gender roles.

In addition to the above, it is necessary to take into account that students who participated in PISA 2018 and who are currently enrolled in higher education, have grown in a context of rapid technological progress and with a dependence on digital devices. PISA 2018 collects the different patterns of behavior of both male and female when faced with these types of devices. Females often use digital devices to carry out online social activities on a daily basis, such as chatting or participating in social networks, whereas males use them in online leisure activities, such as playing online or reading news.

4.6 Limitations

The findings reported in this research come from a limited sample of countries and years. All countries analyzed are OECD members. As mentioned, all of them but Spain are located in the top 11 positions of the PISA 2018 report. The findings and conclusions of this research might have differed in the event that the study was to be carried out for different socio-economic regions in the world.

5. Conclusions and Future Research

The gender dependence in bachelor's or equivalent degrees enrollment was analyzes with a strong focus on the ICT programs. There are many factors that intertwine in order to determine the gender influence on the choice of study fields and professional preferences of men and women. The conclusions obtained for each of the objectives of this research are as follows:

With respect to the objective 1: to analyse the total enrollment in undergraduate studies (total and disaggregated by sex), relating it to other data (scores in PISA 2018, social and economic information). Examination of the gender dependency per field of study.

It has been detected that there is no relationship between the volume of students enrolled and the average score of student skills in the PISA report. Relationships between the enrollment, unemployment, government budget allocations for R&D, employees working very long hours, and population have been found. These relationships have been shown both in the hierarchical clustering and in the correlation studies.

With respect to the objective 2: to study, in more detail, the enrollment in bachelor's or equivalent levels for ICT field, correlating data on technological careers with socio-economic variables and with gender indicators.

A gender dependence is observed in some fields of study (HW, ICT, EMC and EDU). Regarding the volume of women studying ICT, considering the correlation and the clustering studies, there is a strong dependency with the population, unemployment, as well as with the investment made in research and development. The influence of the education budget is greater on the percentage of women who study ICT programs than on the total enrolled students for all fields of study. The above points to the fact that a country with strong research and development, complemented by investments in education, would obtain a higher volume of women studying in the ICT field, and also in the total level of studies.

Since on average, OECD countries present a low level of inequality, a low male-female unemployment ratio, and an average level of life satisfaction, these variables seem to have low or no influence on the volume of female enrollment in the ICT field. Indicators of frequency of ICT use requiring only user-level knowledge do not show gender dependence.

With respect to the objective 3: in order to check whether what is observed in the interactions on certain topics, matches the gender pattern noted in the fields of study, to analyze the influence of gender on the use of Twitter and Instagram networks.

The communications in these social networks show a very high sex dependency according to interaction. However, this is not related to the results obtained by field of study.

In view of the above, the difference in the percentage of women and men who study in the ICT field seems to be related to stereotypes, which cause women to show a lower level of personal efficacy and confidence in their abilities. It is necessary to invest in the training and recognition of the teachers themselves, improve the academic-professional orientation of the students and give recognition to the female leaders in the ICT field. This would result in achieving the required stimulation needed for the student's degrees in this field, particularly among women.

This research can contribute to deepening on the factors which might impact on the lower female presence in ICT studies. Determining which countries and economies have been able to narrow the gender gap in order to help identify conditions and practices that allow boys and girls to break some sexist stereotypes.

The research will be continued studying other social variables that could impact on the volume of women enrolling in ICT studies. These include public or private university, socioeconomic level of the family of the students, and the existence of maternal references in that field of study. Both existing quantitative data in repositories and semi-structured interviews with selected samples of students will be used.

Acknowledgments – This work was carried out as result of the Project: Hopper: Women, Society, Technology and Education which was granted in the internal call for research projects in 2021 at the Universidad Francisco de Vitoria. This research also was partially funded by Telefónica Chair at Universidad Francisco de Vitoria.

References

1. C. Agger and J. Meece, Gender and Academic Motivation, in J. D. Wright, C. Byrne, P. Schmidt, C. McBride-Chang (eds.), *International Encyclopedia of the Social & Behavioral Sciences*, **9**, pp. 677–681, 2015.
2. T. Vaarmets, Gender, academic abilities and postsecondary educational choices, *Journal of Applied Research in Higher Education*, **10**(3), pp. 380–398, 2018.
3. Digitales, 2019, The STEM Vocations Challenge. Why young Spaniards discard science and technology studies. Report. Digitales. https://www.digitales.es/wp-content/uploads/2019/09/InformeEL-DESAFIO-DE-LAS-VOCACIONES-STEM-DIGITAL-AF-1.pdf. Accessed December 24, 2020.
4. S. Fayer, A. Lacey and A. Watson, STEM Occupations: Past, Present, And Future. Report. U.S. Bureau of Labor Statistics. https://www.bls.gov/spotlight/2017/science-technologyengineering-and-mathematics-stem-occupations-past-present-andfuture/pdf. Accessed December 24, 2017.
5. OECD, Education at a Glance 2019: OECD Indicators, OECD, France, 2019.
6. UNESCO, Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM). Technical Report. UNESCO, 2019, https://unesdoc.unesco.org/ark:/48223/pf0000366649. Accessed December 24, 2020.
7. M. L. Mouronte-López, A. García, S. Bautista and C. Cortés, Analyzing the gender influence on the interest in engineering and technical subjects, *International Journal of Technology and Design Education*, 2020.
8. Microsoft. Girls in STEM: the importance of role models. Report. Microsoft. https://news.microsoft.com/europe/features/girls-in-stem-theimportance-of-role-models/, 2018. Accessed December 24, 2020.
9. S-J. Leslie, A. Cimpian, M. Meredith and E. Freeland. Expectations of Brilliance Underlie Gender Distributions Across Academic Disciplines, *Science*, **347**(6219), pp. 262–265, 2015.
10. R. Hermanussen and C. Booy, Equal Opportunity in Higher Technical Education: Past, Present and Future, *International Journal of Engineering Education*, **18**(4), pp. 452–457, 2002.
11. L. R. Ramsey, D. E. Betz and D. Sekaquaptewa, The effects of an academic environment intervention on science identification among women in STEM, *Psychology of Education*, **16**(3), pp. 377–397, 2013.
12. A. Reinking and B. Martin, The gender gap in STEM fields: Theories, movements, and ideas to engage girls in STEM, *Journal of New Approaches in Educational Research*, **7**(2), pp. 148–153, 2018.
13. G. M. Young, L. A. Rudman, H. M. Buettner and M. C. McLean. The influence of female role models on women's implicit science cognitions, *Psychology of Women Quarterly*, **37**(3), pp. 283–292, 2013.
14. L. García and V. Cantillo, Factors influencing the academic performance in standardized tests of computer science/engineering students in Colombia, *International Journal of Engineering Education*, **34**(3), pp. 1073–1084, 2018.
15. OECD, PISA 2018, Technical Report, OECD. https: //www.oecd.org/pisa/PISA â ' resultsENGLISH.png, 2018. Accessed December 24, 2020.
16. A. Hadjar, S. Krolak-Schwerdt, K. Priem and S. Glock, Gender and educational achievement, *Educational Research*, **56**(2), pp. 117–125, 2014.
17. OECD, Development Centre of the Organisation for Economic Co-operation and Development. Technical Report, OECD, 2020. Https://stats.oecd.org/. Accessed December 24, 2020.
18. UNDP, United Nations Development Programme, Human Development Reports, http://hdr.undp.org/en/data, 2019, Accessed December 24, 2020.
19. R, The R Project for Statistical Computing, https://www.r-project.org/, 2020, Accessed December 24, 2020.
20. Python, Python programming language, 2020, https://www.python.org/, Accessed December 24, 2020.
21. GENDERIZE, 2020, Determine the gender of a name, https://genderize.io/, Accessed December 24, 2020.
22. S. Boslaugh, Statistics in a Nutshell, OReilly Media, Estados Unidos, 2012.
23. W. Glatzer and J. Gulyas, Cantril Self-Anchoring Striving Scale, In A. C. Michalos (eds.) Encyclopedia of Quality of Life and Well-Being Research, Springer, Netherland, 2014.
24. K. F. Weaver, V. Morales, S. L. Dunn, K. Godde and P. F. Weaver, *An Introduction to Statistical Analysis in Research: With Applications in the Biological and Life Sciences*, John Wiley and Sons, United States of America, 2017.
25. D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, John Wiley & Sons, Inc, United States of America, 2003.
26. R. S. King 2015. *Cluster Analysis and Data Mining: An Introduction*, Mercury Learning and Information, United States of America, 2015.
27. P. F. Murtagh and P. Legendre, Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm, *Journal of Classification*, **31**(3), pp. 274–295, 2014.
28. F. Murtagh, and P. Contreras, Methods of Hierarchical Clustering, *Cornell University*, pp. 1–21, 2011.
29. O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J. M. Pérez and I. Perona, An extensive comparative study of cluster validity índices, *Pattern Recognition*, **46**, pp. 243–256, 2013.
30. M. Neuhauser, *Nonparametric Statistical Tests: A Computational Approach*, Chapman and Hall/CRC, United States of America, 2017.
31. OECD, Pisa in Focus no 13. Report. OECD, 2012, https: //www.oecd.org/pisa/pisaproducts/pisainfocus/PISA%20in%20 Focusâ 'n%C2%B013%20ESPFinal.pdf, 2012, Accessed December 24, 2020.
32. C. K. Avendaño and D. E. Magaña, Elección de carreras universitarias en áreas de ciencia, tecnología, ingeniería y matemáticas (STEM): revisión de la literatura, *Revista interamericana de educación de adultos*, **40**(2), pp. 155–173, 2018.
33. EC, Women in the Digital Age, Report, European Commission, https://op.europa.eu/es/publication-detail/-/publication/84bd6dea-2351-11e8-ac73-01aa75ed71a1, Accessed December 24, 2020.
34. DESDJEI, ICT Skills Action Plan (2014–2018), Report, Department of Education and Skills and Department of Jobs, Enterprise and Innovation, https://www.education.ie/en/Publications/Policy-Reports/ICT-Skills-Action-Plan-2014-2018.pdf, Accessed December 24, 2020.
35. EC, Digital Education at School in Europe, Report, European Commission, Eurydice, https://eacea.ec.europa.eu/nationalpolicies/eurydice/sites/eurydice/files/en_digital_education_n.pdf, 2019, Accessed December 24, 2020.

36. NZN, Information technology, Report, New Zealand Now, https://www.newzealandnow.govt.nz/work-in-nz/nz-jobs-industries/information-technology-jobs, Accessed December 23, 2020.
37. MEDP, The Polish ICT sector, Report, Ministry of Economic Development of Poland, https: //ict.trade.gov.pl/en/news/238157,el – sector – polaco – de – las – tic.html, 2017, Accessed December 24, 2020.
38. K. Talley and A. Ortiz, Women's Interest Development and Motivations to Persist as College Students in STEM: A mixed methods analysis of views and voices from a hispanic-serving institution, *International Journal of STEM Education*, **4**(5), 2017.
39. L. English, P. Hudson and L. Dawes, Perceived Gender Differences in STEM Learning in the Middle School, *International Journal of Engineering Education*, **27**(2), pp. 389–398, 2011.
40. K. Myers, J. Jahn, B. Gailliard and K. Stoltzfus, Vocational Anticipatory Socialization (VAS): A communicative model of adolescents' interests in STEM, *Management Communication Quarterly*, **1**(25), pp. 87–120, 2011.
41. H. Holmegaard. Performing a Choice-Narrative: A qualitative study of the patterns in STEM students' higher education choices, *International Journal of Science Education*, **9**(37), pp. 1454–1477, 2015.
42. S. Verdugo-Castro, M. C. Sánchez-Gómez, A. García-Holgado and F. J. García-Peñalvo, Revisión y estudio cualitativo sobre la brecha de género en el ámbito educativo STEM por la influencia de los estereotipos de género, *Investigación Cualitativa en Ciencias Sociales*, **3**, pp. 381–386, 2019.
43. G. Stoet and D. Geary, The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education, *Psychological Science*, **29**(4), pp. 581–593, 2018.
44. I. Throndsen and O. E. Hatlevik, Examining Gender Differences in ICT Literacy, Interest, and Use, in E., Eyvind (eds.), Digital Expectations and Experiences in Education, Springer, Germany, pp. 221–240, 2016.
45. I. Vekiri and A. Chronaki, Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school, *Computers & Education*, **51**(3), pp. 1392–1404, 2008.
46. M. Volman and E. van Eck, Gender equity and information technology in education: the second decade, *Review of Educational Research*, **71**(4), pp. 613–634, 2001.
47. W. K. Nantwi and A. Adjei, Gender metamorphoses in the use of ICT Tools: A case study at offinso college of education, *British Journal of Education*, **6**(8), pp. 109–118, 2018.
48. J. Schneider and C. Meske, Gender Differences in Enterprise Social Network Usage and Transformation over Time, 38th International Conference on Information Systems, Korea, 10–13 December, 2017.
49. L. Xiaolin and X. Wang, Examining gender differences in people information-sharing decisions on social networking sites, *International Journal of Information Management*, **50**, pp. 45–56, 2020.
50. G. Mazman and U. Yasemin, Gender differences in using social networks, *Turkish Online Journal of Educational Technology*, **10**, pp. 133–139, 2011.
51. C. Meşe and G. Sancak, The use of social networks among university students, *Educational Research and Reviews*, **14**(6), pp. 190–199, 2019.
52. N. A. Petrocchi, A. Asnaani, A. Piquer, A. Nadkarni and S. G. Hofmann, Differences Between People Who Use Only Facebook and Those Who Use Facebook Plus Twitter, *International Journal of Human Computer Interaction*, **31**(2), pp. 157–165, 2015.
53. D. Bamman, J. Eisenstein and T. Schnoebelen, Gender in Twitter: Styles, Stances, and Social Networks, *Computing Research Repository – CoRR*, 2012, https://arxiv.org/vc/arxiv/papers/1210/1210.4567v1.pdf. Accessed 22 March 2021.
54. L. M. A. Bahammam, Gendered discourses and discursive strategies employed in Twitter-hashtagged debates about Saudi-women's issues, PhD thesis, University of Reading, 2018. http://centaur.reading.ac.uk/82516/. Accessed 23rd March, 2021.

Mary Luz Mouronte-López. PhD in Telecommunications Engineering, Universidad Politécnica de Madrid (UPM). PhD in Law and Society Doctoral Program, Universidad a Distancia de Madrid (UDIMA). Master Business Administration (MBA), UPM. Master in Fundamental Rights, Universidad Nacional de Educación a Distancia (UNED). Professor at the High Polythenic School, Universidad Francisco de Vitoria (UFV). Director of the Telefónica Chair at UFV. Main Researcher of the Research Group Engineering and Women in ICT from Humanism at UFV.

Juana Savall Ceres. PhD in Education, Universidad Autónoma de Madrid. Professor at Francisco de Vitoria University, both in Early Childhood and Primary Education Degrees, as well as in University Master's Degree in Secondary Education and Bachillerato, FP and Language Teaching. Vice-Dean of Integral Formation of the Faculty of Education and Psychology. Member of the Research Group Engineering and Women in ICT from Humanism at UFV.

Appendix

AH	Arts and Humanities
AFFV	Agriculture, Forestry, Fisheries and Veterinary
ALFS	Annual labour force statistics
ANOVA	Analysis Of VAriance
API	Application Programming Interface
BAL	Business, Administration and Law
C2B	Individuals using a computer - last 3 m (%)
C5B1	Individuals using the Internet daily or almost every day - last 3 m (%)
DI	Dunn Index
EATT	Educational attainment.
EDU	Education

EMC	Engineering, Manufacturing and Construction
ENRO	Enrolment in number of persons
EWLH	Employees working very long hours. Percentage of dependent employed whose usual hours of work per week are 50 hours or more
GDP	Gross Domestic Product
GDPE	Percentage of total expenditure on educational institutions per full-time equivalent student relative to GDP per capita
GII	Gender Inequality Index
GPQ	Generic Programmes and Qualifications
GPRD	Government budget allocations for R&D in PPP dollars current-prices
H_0	Null Hyphotesis
Ha	Alternative Hyphotesis
H1H	Individuals who have found, downloaded and installed software from the Internet - last 12 m (%)
H1K	Individuals who have written computer code - last 12 m (%)
H1L	Individuals who have created a web page - last 12 m (%)
H1M	Individuals who have installed or replaced an operating system - last 12 m (%)
HW	Health and Welfare
ICT	Information, Communications and Technology
ISCED 2011	International Standard Classification of Education 2011
LISA	Life satisfaction
MinMax	Minimum Maximum
NSMS	Natural Sciences, Mathematics and Statistics
OECD	Organisation for Economic Co-operation and Development
PEEA	The average annual wages per full-time equivalent dependent employee
PISA	OECD's Programme for International Student Assessment
POPU	Population
PPP	Purchasing power parities
R&D	Research and Development
SC	Silhouette Coefficient
SER	Services
SSE	Sum of squared errors
SSJI	Social Sciences, Journalism and Information
STEM	Science, Technology, Engineering and Mathematics
STSK	Student skills. Students 'average score in reading, mathematics and science as assessed by PISA
U.K.	United Kingdom
UNEM	Unemployment in thousands of persons
UR	Total unemployment rate (female to male ratio)
U.S.A.	United States of America