

Women in STEM, Potential in the 4th Industrial Revolution

*Some questions that could be answered with
sources integration*

Juan Daniel Oviedo

New York, February 2019



**El futuro
es de todos**

**Gobierno
de Colombia**

The potential of Women in STEM fields in Colombia

Some examples of what we know from Household Surveys

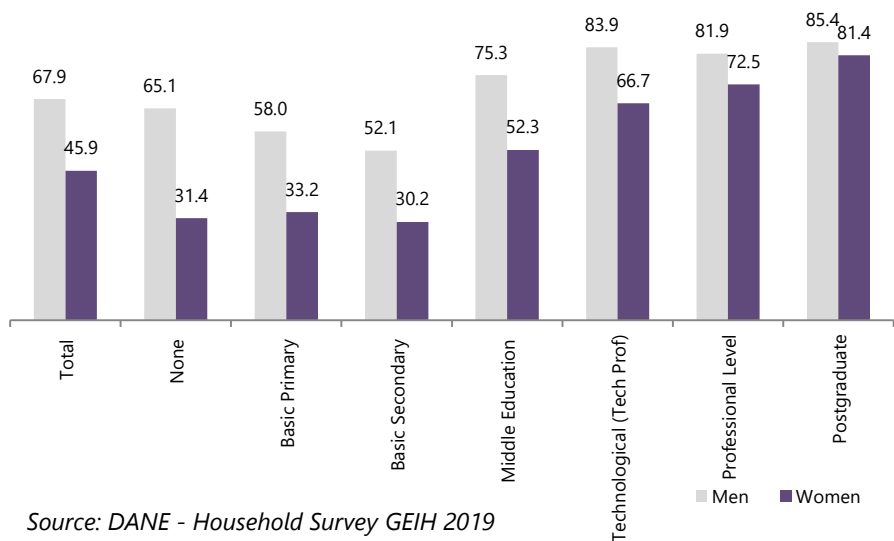


El futuro
es de todos

Gobierno
de Colombia

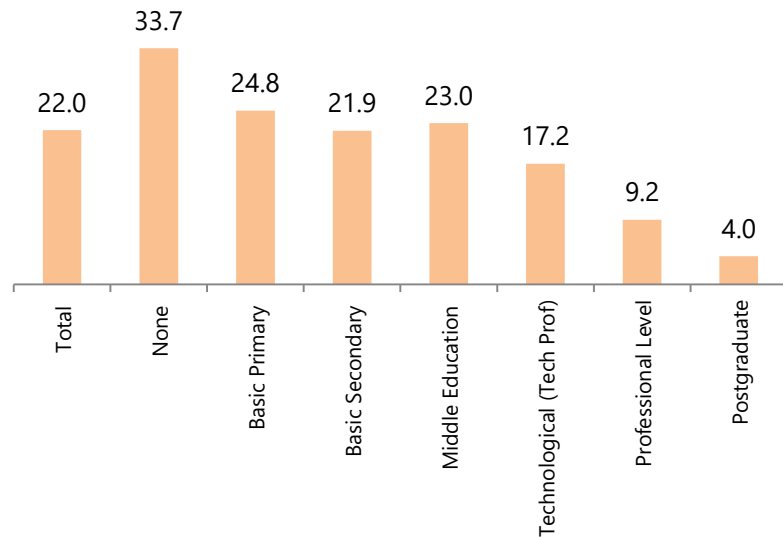
Employment Rate by education level and sex – Household Survey, GEIH

Employment Rate by education level and sex, 2019



Source: DANE - Household Survey GEIH 2019

Employment Rate gap (W-M), 2019



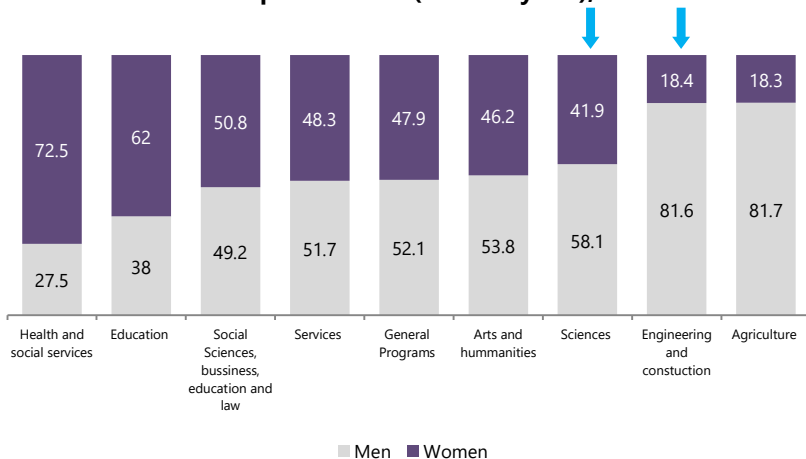
The higher the level of education, the smaller the gender gap in employment rate.

The highest gap is registered when the educational level is none, with **33,7 p.p.**

But in GEIH we **do not have** information about the **area of knowledge** or **type of career**

Career aspirations gender gap in employed and unemployed young people – ETET*

Distribution by knowledge areas in which employed young people study or would like to study according to gender - 13 cities and metropolitan areas (14 to 28 years), 2017

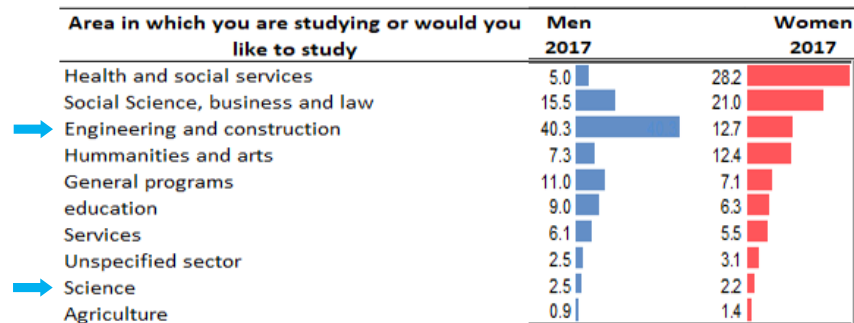


Source: DANE, School-to-Work Transition Survey ETET, 2017.

Note: General programmes include school education, literacy and numeracy, and professional development.

18,4% of employee people who study or would like to study **Engineering, Industry and Construction** are **women**.

Distribution by thematic areas in which unemployed young people study or would like to study according to gender - 13 cities and metropolitan areas (14 to 28 years), 2017



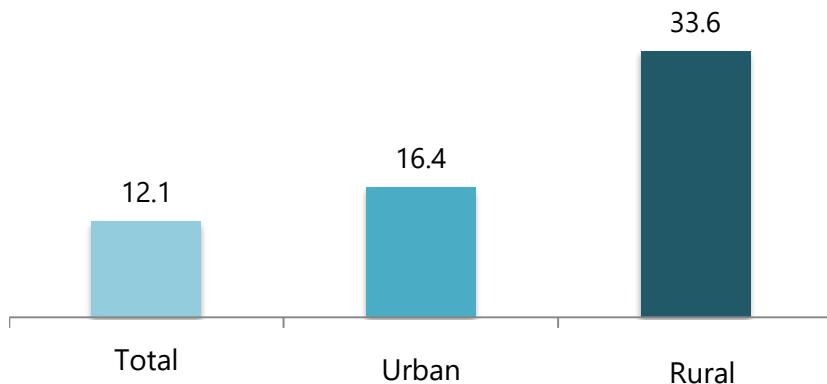
Source: DANE, School-to-Work Transition Survey ETET, 2017.

- **12,7% of unemployed women** are studying or would like to study **engineering, industry and construction**
- **40,3% of unemployed men**, are studying or would like to study **engineering, industry and construction**
- This gender gap has been getting shorter in Science careers.

*ETET was focused on Young People (14-29 years old). It was conducted on 2013, 2015 and 2017. Some of its questions are being included in the new GEIH questionnaire.

Gender pay gap – Household Survey, GEIH

Labor Income gap by geographic area (2018)



Source: DANE – Household Survey GEIH. .

In 2018 the gender pay gap was:

- The **total gender pay gap was 12.2%**.
- In rural areas the gender pay gap was 17.2 p.p higher than in urban areas.

But...

- In GEIH we have general economic sectors and job position, but not detailed.
- In GEIH we **do not have** information about the **area of knowledge** or **type of career**.
- This is a **self reported** income.

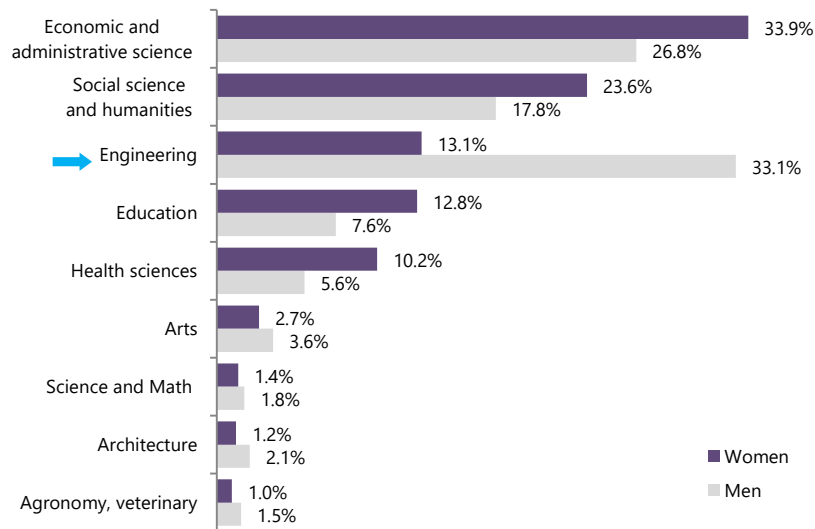


The potential of Women in STEM fields in Colombia

What we know from Administrative Records

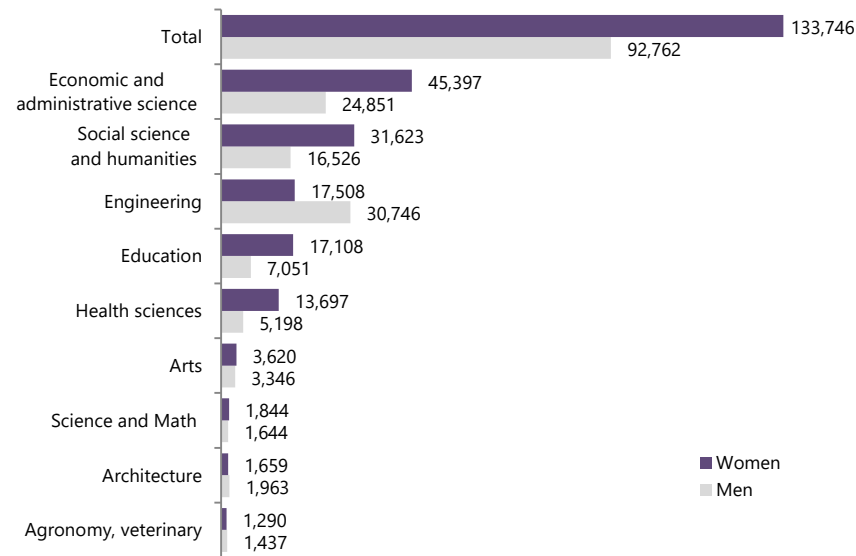
Graduation of professionals in Colombia, by areas – SNIES, MEN

People graduated in 2018, by sex and knowledge area (percentage by each sex)



Source: SNIES, MEN, 2020. Percentage calculated on the total number of persons by sex

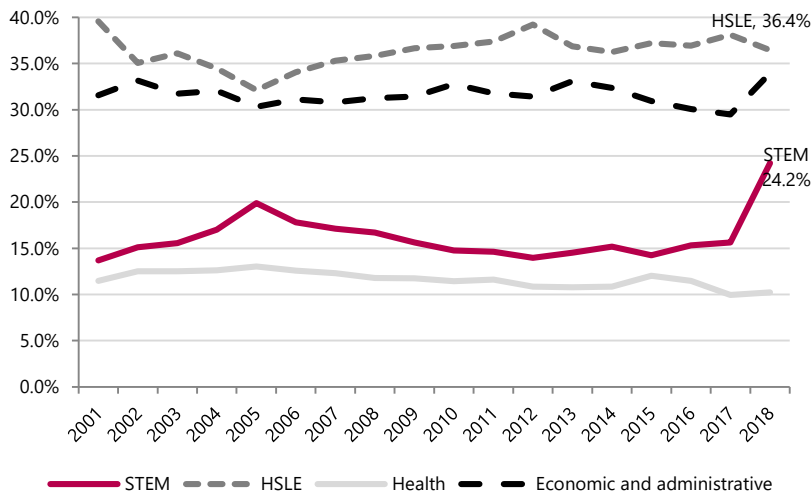
Graduates in 2018, by sex and knowledge area



Source: SNIES, MEN, 2020. Percentage calculated on the total number of persons by sex

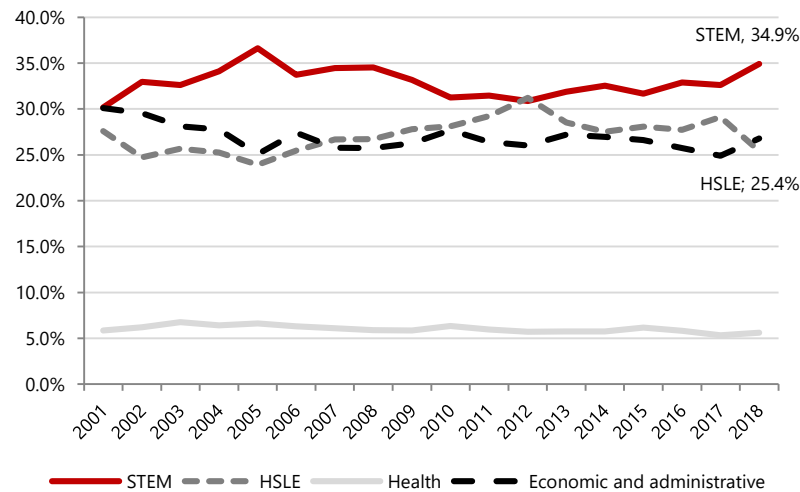
Higher Education - Evolution of Colombian professional degrees by sex

Female higher education graduates by study area, 2001-2018



Source: MEN, SNIES, 2020. Note: Architecture and Arts, Agricultural Sector and unclassified are excluded, with representations less than 5%. Own classification.

Men higher education graduates by study area, 2001-2019



Source: MEN, SNIES, 2020. Note: Architecture and Arts, Agricultural Sector and unclassified are excluded, with representations less than 5%. Own classification.

Questions to be answered

3 forthcoming projects



El futuro
es de todos

Gobierno
de Colombia



1. How are the labor indicators according to the **types of careers or knowledge areas?**



Integration of GEIH and Ministry of Education's Administrative Records (SNIES)

2. What happens **within the labor market? For examples, are there differences in times to get a promotion?**



Big Data - LinkedIn

3. How much is the **gender pay gap, by economic sector, if it is not self-reported?**



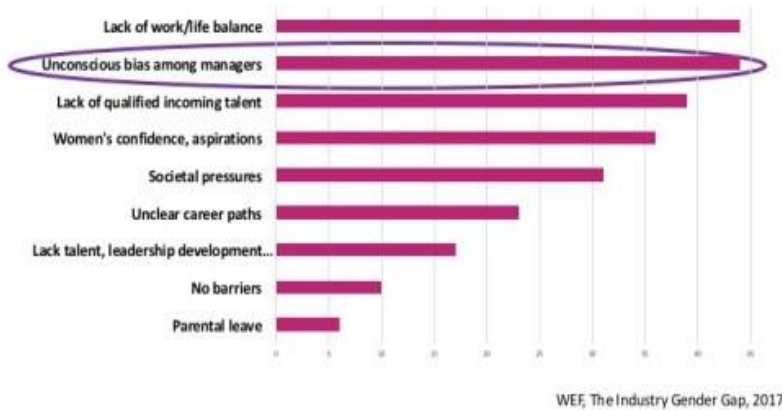
Integration of GEIH and Ministry of Health's Administrative Records on social security (PILA)

Let's talk about the last two:

2. What happens withing the labor market?

Are there barriers that differentially affect women, like in the USA?

Barriers to Gender Parity, All Industries

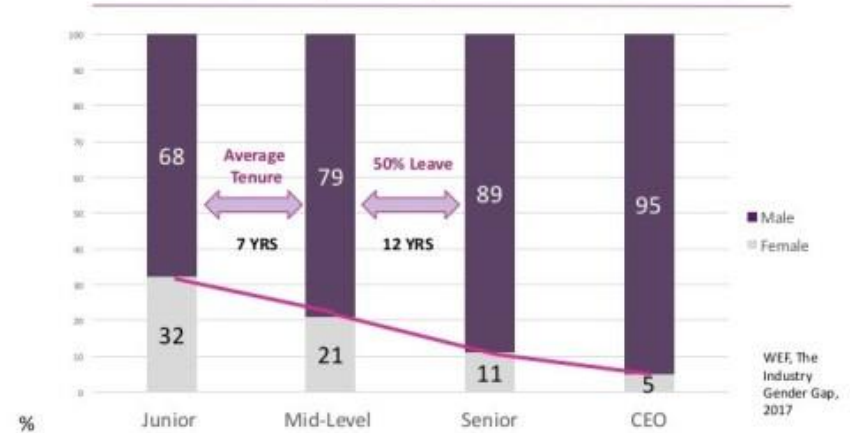


Source: Linda Scott (2018). From Santiago: Gender Equality and the Fourth Industrial Revolution.

*Linda Scott is Emeritus DP World Professor of Entrepreneurship and Innovation, University of Oxford. Founder and Senior Advisor to, the Global Business Coalition for Women's Economic Empowerment

In Colombia we don't have official measurements of this

ICT Career Progression by Gender 2016 (% of Total)



Source: Linda Scott (2018). From Santiago: Gender Equality and the Fourth Industrial Revolution.

"This graph is the key to why women are so few in tech. It is not really because of the intake into the pipeline (which has now risen to 40%), but is more about **what happens in the first decade of their careers in that industry**" Linda Scott

¿BIG DATA? ¿ADMINISTRATIVE RECORDS?

DANE Project: People in STEM areas from LinkedIn

Objective To evaluate women's participation in the labor market in STEM areas through non-traditional Sources. Regarding:

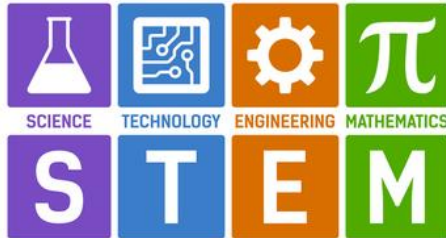
- **Time to achieve promotions.**
- **Time to get a job.**
- Construction of professional profiles: how do people offer their profile to attract employers?
- Geographical disaggregation: differences in gap sizes.
- Does new technological employment platforms break down barriers to labor inclusion?

Some advantages

- Development of more efficient algorithms for processing social networks information.
- The number of Colombian users registered in the LinkedIn social network is approximately **29.5% of the economically active population in size.**
- Preliminary results show the viability of this Source: to generate **indicators with a gender perspective.** .
- **LinkedIn is positioned as one of the main social media platforms focused on professionals.** This is used by employers to search for potential candidates, and by their users as a means of presenting their curriculum.



Challenges of using this Source:



- Database debugging regarding the duplication of information.
- Classification of records by **sex** and **STEM areas**; also, data will be generated based on user values such as names and the current work, to determine the needed variables.
- Python libraries were used for debugging the required variables; therefore, classification methods were applied using word-based dictionaries.
- **Data access** due to information security policies, which are related to the user that makes the request and its contact network. Ethical and legal aspects that may limit access to such information.
- **We have not achieved access to the complete base because it costs**, but just to partial user downloads.

3. Dane Project: Estimation of the Gender Wage Gap by integrating Social Security Administrative Records (PILA) and Household Survey - GEIH

Some advantages

- PILA allows us to **overcome some of the sampling survey errors: does people underestimate or overestimate their wages differentially according to sex?**
- The **imputation process of the personal income in GEIH (*hot deck*) could be improved using PILA data** which is currently used to correct errors due to non-response and outliers.
- To get **greater disaggregation**, improving the design of public policies for closing gender gaps: positive discrimination, elimination of stereotypes since childhood.

Next steps

- It is required **to use other administrative records and the Census** to complement the PILA information and get more intersectional disaggregations.
- To increase the matching rate GEIH-PILA by **exploring record linkage methods where no perfect integration key exists**, using variables such as names, surnames and birth dates.

Some challenges

- **Linkage:** The personal identification number (*cédula*) is not mandatory for respondents in the GEIH. Thus, it is troublesome to match GEIH-PILA.
- **Integration of multiple data sources:** It is necessary to harmonize the information since the administrative records do not have statistical purposes.
- **Statistical Unit:** PILA excludes especial regimes of social security such (as district schools teachers and informal employees without social security payments). This could result in a bias as **there may be differences in gender participation between included and excluded groups.**
- **Income definition:** There is no information of non-labor income in PILA. While in GEIH, there is no information available on social security payments, although there are measures of in-kind income. PILA and GEIH do not include tax information.

Women in STEM, Potential in the Fourth Industrial Revolution

Some questions that could be answered with sources integration

Juan Daniel Oviedo
jdoviedoa@dane.gov.co

New York, February 2019



**El futuro
es de todos**

**Gobierno
de Colombia**



@DANEColombia



@DANE_Colombia



/DANEColombia



/DANEColombia

www.dane.gov.co