Visual Analytics

Camilo Andrés Rojas Bulla, Oscar José Marriaga, Monica Katherine Duran Vaca



Fig. 1. Misbirth in Colombia

**Abstract**— This article explains the design of the visualizations for the analysis of legal misbirths in Colombia between 2009 and 2017, based on a dataset of the Ministry of Health, provided by Datasketch, client assigned for this project. Datasketch is a digital platform for investigative and data journalism. The portal allows journalists, data scientists, social scientists and citizens in general to learn and consult on data visualizations, tools, software and in-depth research on various short-term issues.[2]

**Index Terms**—Visualizations,insights...

Introduction

Since 2006, the Constitutional Court of Colombia gave way to the IVE (Voluntary Interruption of Pregnancy or *Interrupción Voluntaria del Embarazo* in spanish), allowing this procedure to be carried out in three circumstances; when the pregnancy endangers the health or life of the woman, when the pregnancy is the result of rape or incest, and when there are malformations of the fetus that are incompatible with life outside the uterus. This interruption is a fundamental law that consists of a procedure to consciously end an ongoing pregnancy, in many countries this procedure is illegal, so women resort to illegally doing it in places that do not meet optimal health conditions, putting your health and your life in danger.

The purpose of the visualizations presented here, is to be able to, from the data used, make an analysis, make comparisons, find trends, that allow to establish strategies that manage to control the causes of legal abortion in Colombia and that mitigate non-legal procedures avoiding deaths in Colombian women.

# Related work

The main objective of the visual analytics project is to present visualizations that allow the client to interact with the data, finding insights that help in the process of making decisions

For this purpose, we propose the design of 5 visualizations that try to solve the following tasks: summary the distribution of the amount of abortions in each department, explore the amount of abortions occurred in each department given a selected year, find the department with the largest amount of abortions, summarize the distribution of the number of abortions by age range over time and finally, finding the year in which more abortions were practiced given a rank of age.

The use of visualization tools used in this project, such as the Tamara framework (what - why - how), the d3 library, and vega lite, allowed us to analyse the data in detail by summarizing, selecting, identifying and comparing trends, achieving get insights, according to the initial problem found in the data used.

## Visualization One

The first visualizations use an idiom bar chart, to present the information of the departments [figure 1], types of abortion [figure 2]and age range [figure 3] with the highest number of legal abortions practiced in Colombia between 2009 and 2017. It is possible to filter the visualization in year, department or age range. Analysing the visual encoding, bar charts use a line mark and encode a quantitative value attribute with one spatial position channel. [2]

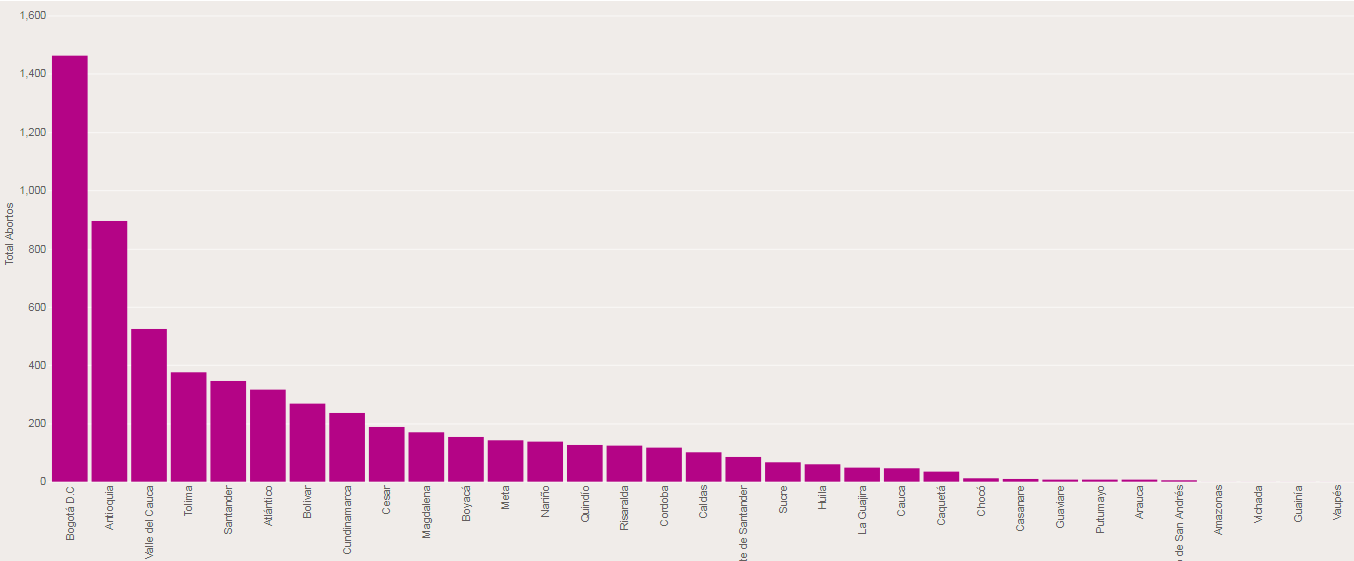


Fig. 1. Number of abortions per department given one year.

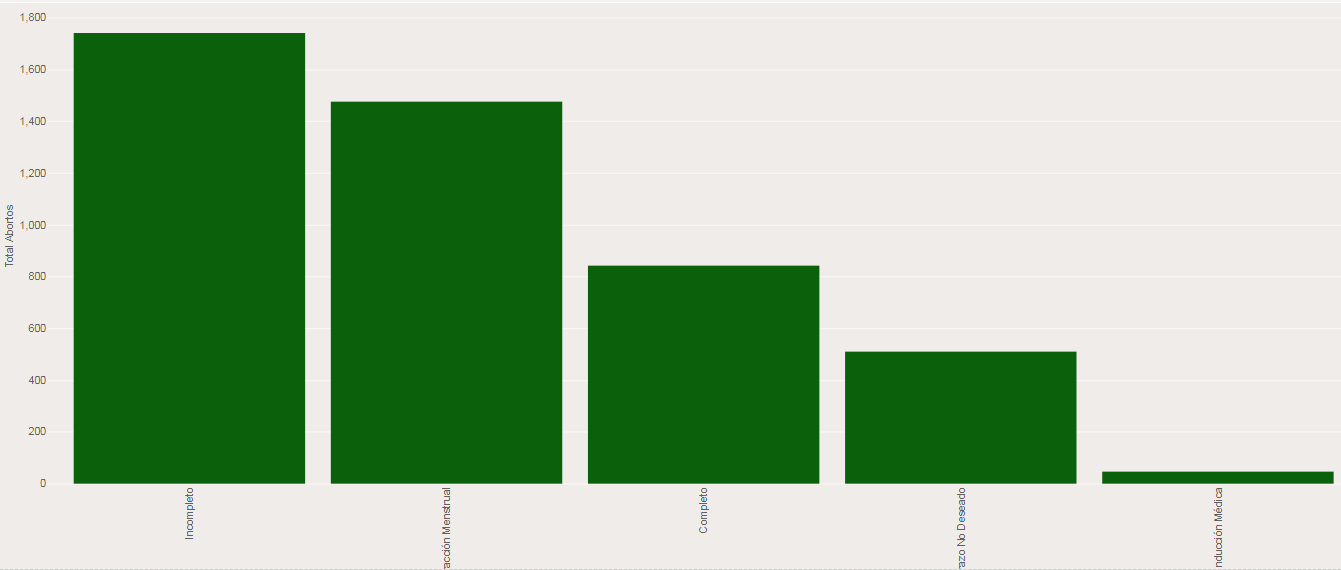


Fig. 2. Type or abortion practiced per year.

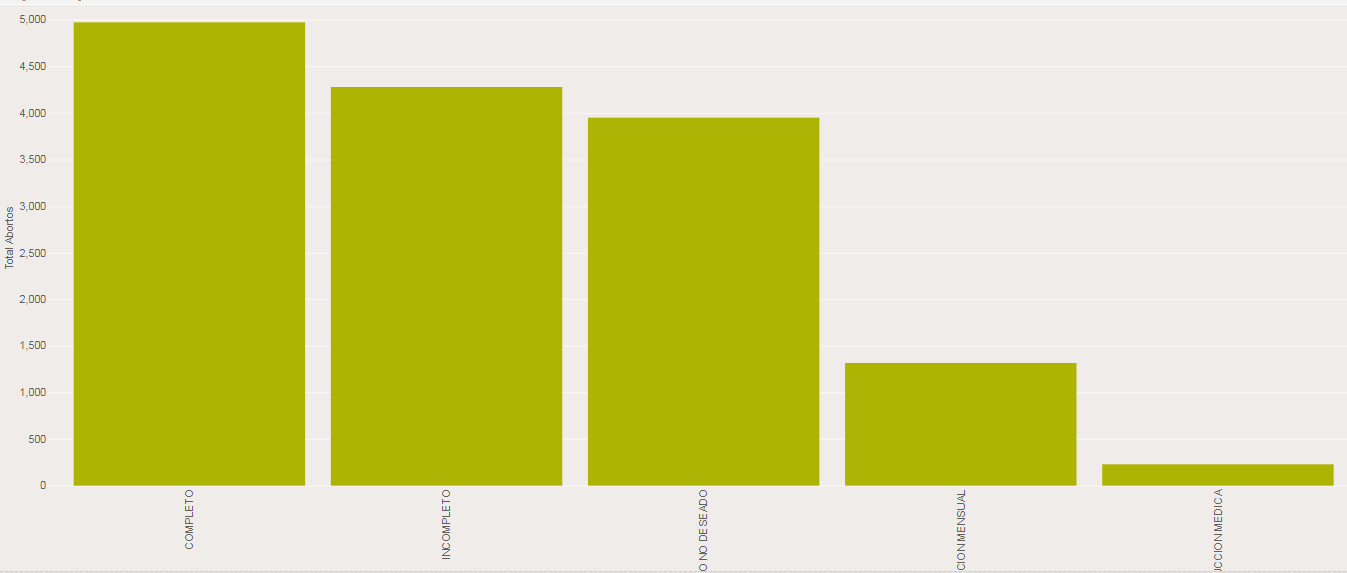


Fig.3. Type or abortion practiced per year range.

## Visualization Two

The second visualization uses an idiom line chart [4], to present the information of the distribution of the number of abortions by age range over time [figure 4].

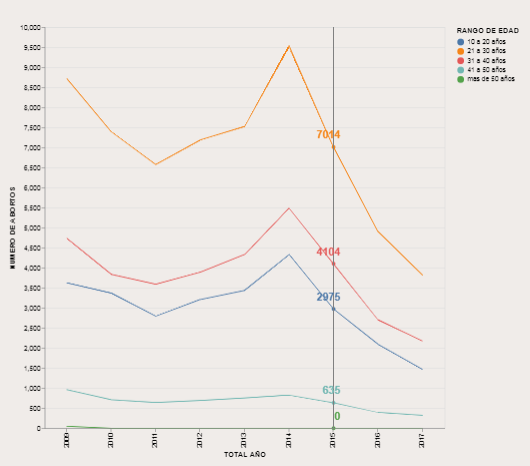


Fig.4. Number of abortions between 2009 and 2017 by age range.

1.3 Visualization Three

The third visualization uses an idiom stacked bar chart to present the distribution of the amount of a type of medical abortion per year [figure 5]. Each component of the bar is separately stacked, so that the full bar height shows the value for the combination of all items in the stack. [5].

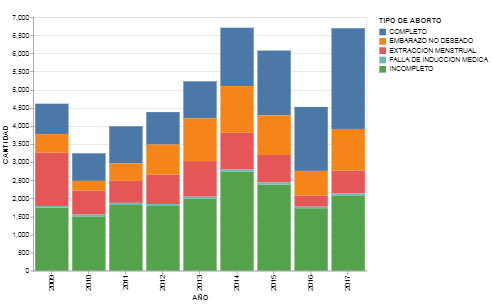


Fig.5. Type of abortion between 2009 to 2017.

The next visualization uses an idiom multiple bar chart to present a type of medical abortion per diagnosis. [figure 6]

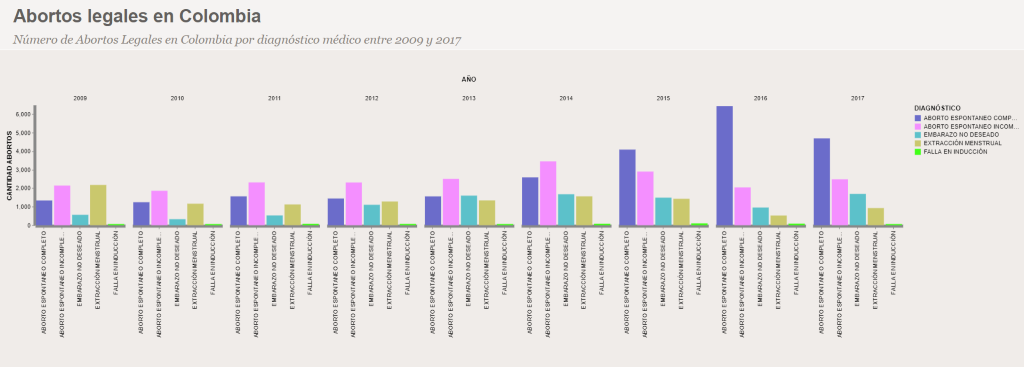


Fig.6. Number of abortions by diagnosis between 2009 and 2017.

# solution

The visualizations presented in the section above show worrisome figures on abortion practices in very young women, procedures that have increasingly increased, causing widespread social problems; the visual analysis, through these visualizations, allow us to take decisions about how to achieve abortions, but above all, create strategies to stop pregnancies at an early age, among others.

Following these figures, different campaigns have been launched in order to prevent these situations, an example of this is the #YoCuidoMiFuturo campaign carried out by the ICBF and Profamilia launched since 2015, this type of campaign seems to have an effect on young people because in last years there has been a decrease in the number of abortions performed.

# Evaluation

The usability proofs focused on three main tasks to know:

* Summarize the number of abortions by age range.
* Find the age range with more and fewer abortions practiced.
* Find an age range.
* Find the year of an age range in which more and fewer abortions were performed.
* Find the number of abortions for a year and an age range.

For the visualization of the number of abortions by age ranges [figure 4], the result indicates incorrect for the task of finding the year where there were more abortions for the age range of 41 to 50 years. This was solved, modifying the scale and the colors for the age range.

# Conclusion

The use of visual analytics is important at the moment of showing figures clearly, but even more important when techniques are involved, as is the case of the Tamara framework (what – why - how), where the interpretation of the data is understood and understood, recognizing its importance in the insights found in the initial problem.

The most important insights found in each visualization are:

* Miscarriage has been the most common diagnosis in authorized health centers to legally terminate a pregnancy. With regard to induced abortions, it can be observed that menstrual extraction has been the most used means to perform the interruption of pregnancy.
* In induced abortions, menstrual extraction has been the most used means to perform the termination of pregnancy.
* The women who have come to perform an abortion legally in health centers are in the age range of 21 to 30 years.
* In 2014 there is a high peak in the number of abortions, the most worrisome figure was in women aged 21 to 30 years with a total of 9,537 abortions.
* In most departments, the year in which there were more abortions was in 2014, in departments like Arauca a fairly large peak is evidenced, tripling the amount with respect to the other years.

References

1. Aborto en Colombia – Servicios para la Mujer, available at https://profamilia.org.co/inicio/mujer-2/servicios-mujer/aborto-colombia/.
2. TAMARA available at https://www.cs.ubc.ca/~tmm/talks.html#minicourse14