

**Curso: Introducción a Data Science con R**

**Trabajo final: análisis de dataset “Educación Sexual”**

**Alumno: Christian Farnast Contardo**

**Profesora: Mg. Layla Scheli**

**Marzo 2021**

1. **INTRODUCCION**

El trabajo final consiste en el análisis de un dataset provisto por la relatora del curso, Prof. Layla Scheli, llamado “Educación Sexual”, el que está en formato .xlsx

Las variables del dataset son:

1. Condición socioeconómica
2. Años de escolaridad
3. Relación de pareja
4. Número de hijos
5. Edad

Para el análisis se usará el software estadístico R y sus librerías:

* readxl, el que permite leer archivos .xlsx
* ggplot2, librería que permite la visualización de datos por medio de gráficos
* plotly, librería que complementa ggplot2 y otorga al grafico interactividad mostrando los datos en los gráficos obtenidos.

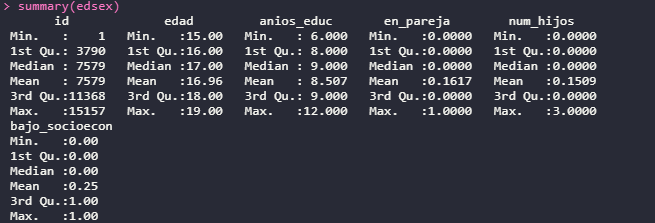
Con todo lo anterior, se deberán concluir o inferir las posibles interacciones entre variables con la finalidad de dilucidar el comienzo de una educación sexual que permita bajar la cantidad de embarazos en etapa escolar.

1. **DESARROLLO: ANÁLISIS GRÁFICO DEL DATASET**

Como se nombró previamente, para el análisis de los datos se ocuparan los paquetes o librerías del software estadístico R:

* readxl.
* ggplot2.
* plotly.
  1. **ESTUDIO DE LOS DATOS**

Aplicando la función summary, se muestran los siguientes resultados:



De la imagen se desprende que:

1. El número de encuestados (“id”) son 15157 personas
2. La edad de los encuestados (“edad”) varía entre 15 y 19 años
3. Los años de educación (“anios\_educ”) van desde los 6 hasta los 12 años
4. La variable en pareja (“en\_pareja”) muestra 0, para quienes están solteros y 1 para una relación de noviazgo.
5. El número de hijos (“num\_hijos”) varia de 0 a 3 hijos y,
6. La variable socio económica (“bajo\_socioecon”) es de 0 para quienes no tienen condición baja y 1 para quienes si tienen una condición baja.
   1. **ANÁLISIS GRÁFICO**

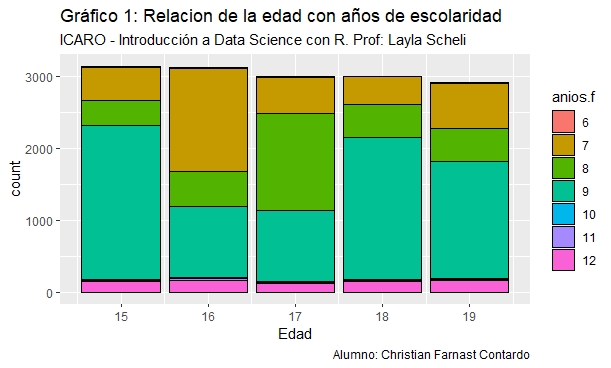
Las variables más significativas para el estudio del dataset son:

* edad y,
* num\_hijos

Estas variables inciden finalmente en los años de escolaridad (“anios\_educ”) de los encuestados si pueden o no mantener la responsabilidad de ser padres con terminar en tiempo y forma su enseñanza secundaria.

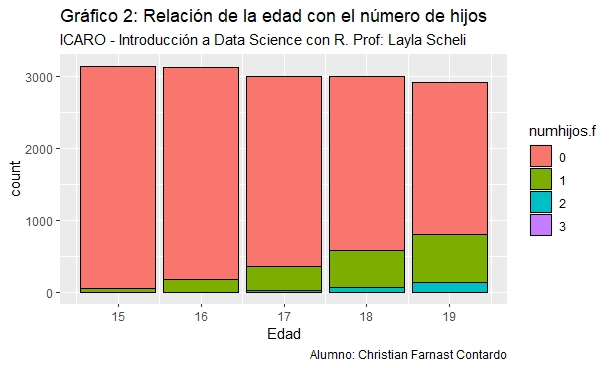
Secundariamente, es también importante entender su condición socioeconómica, su relación en pareja y su relación con las demás variables de estudio.

2.2.1. Edad en relación a los años de escolaridad.



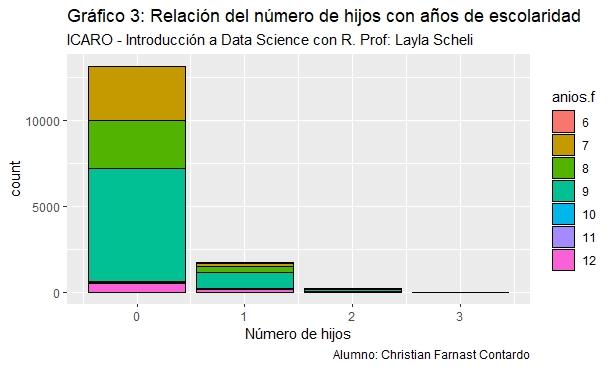
Se muestra una relativa uniformidad de los encuestados en su número versus su edad. Se puede observar que pocos encuestados de la edad de 18 años cursan su doceavo año de escolaridad. En el resto de las edades hay distribución entre los 6 a 12 años de escuela, posiblemente por su maternidad y/o paternidad prematura.

2.2.2 Relación de la edad con la cantidad de hijos



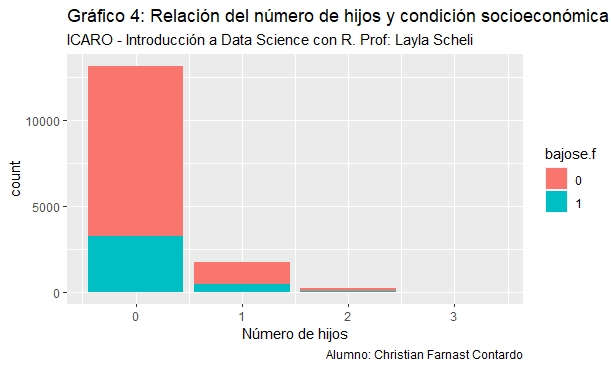
El gráfico muestra que todas las edades tienen al menos un hijo y la cantidad va aumentando a medida que crecen. Sin embargo, son mayoría aquellos que no tienen hijos por lo que este grupo puede tener menos o más años de escolaridad no presumiblemente a la maternidad y/o paternidad prematura.

2.2.3. Relación del número de hijos con los años de escolaridad.



El gráfico 3 muestra que el grupo sin hijos muestra una diversidad en sus años de colegiatura, esta tendencia disminuye en cantidad de encuestados a medida que el número de hijos aumenta. Se observa que el grupo con 9 años de escolaridad ya tiene un hijo, con 15 años de la madre y/o padre.

2.2.4. Relación del número de hijos con la condición socioeconómica.



Se observa en el gráfico 4 que quienes tienen como también aquellos que no tienen hijos, pertenecen en su mayoría a una condición socioeconómica no baja.

**2.3. DISCUSION**

La mayor parte de los encuestados corresponde a la edad de quince años, los que en su mayoría van acorde a sus años de escolaridad (9 años). Un número importantemente bajo son aquellos que teniendo 17 y 18 años van con sus 12 años de escolaridad, sin embargo este parámetro no se relaciona con la cantidad de hijos, ni tampoco con la condición socioeconómica donde este grupo es mayoritario (grafico 6 dispuesto en el código)

En cuanto al número de hijos, esto se puede ver en correlación de la edad y su vida en pareja (grafico 7, dispuesto en el código), no obstante la cantidad de encuestados varia a la baja conforme aumenta los años de escolaridad junto con la condición socioeconómica, siendo este último punto predominante en aquellos padres que no tienen una baja condición socioeconómica.

Complementando el análisis de la condición socioeconómica, el número de hijos varia notablemente entre los que tienen y aquellos que no son padres observándose mayormente aquellos encuestados no tienen hijos, junto con no tener una baja condición socioeconómica. Se hace ver que tampoco hay una relación entre esta variable con los años de escolaridad: para todas las edades hay un predominio de quienes no tienen una condición baja.

De mayor valor en el dataset, hubiera sido tener la información por sexo de los encuestados para poder visualizar este punto con el resto de las variables.

* 1. **CONCLUSION PERSONAL DEL PROYECTO**

De lo anteriormente descrito en los resultados y en la discusión, no se concluye una relación indirecta del número de hijos y los años de escolaridad. De acuerdo con las nuevas responsabilidades que adquieren los jóvenes al asumir una paternidad temprana la relación tendería a ser directa, sin embargo al no ser tal, otros factores pueden influir como deserción escolar por consumo de drogas, falta de trabajo estable en el hogar que obliga a los estudiantes a trabajar en oficios tal vez poco remunerados, desarrollo de un oficio que persiguen los jóvenes como ideales de obtener mayor redito económico, etc.

Se puede observar que los nacimientos aumentan con la vida en pareja siendo directa dicha relación y no importando la condición socioeconómica.

A la edad de 15 años se ven pocos encuestados con al menos un hijo, por lo que es importante una educación sexual previa a esta edad, prevalecer el rol de los padres como educadores primordiales en la sexualidad de los hijos y recalcar a la escuela como reforzador de lo anterior. Además de la prevención de embarazos, se debe hacer hincapié en el contagio y transmisión de enfermedades por contacto sexual como SIDA, sífilis, gonorrea y hepatitis A, B y C.

**2.5. ANEXOS**

Gráficos

