

# Proyecto Final

**Estudiantes: Emiliano Barone, María Noel Blanco**

## INTRODUCCIÓN

Nuestro trabajo se basa en el análisis de los datos del Estudio Longitudinal de Bienestar del Uruguay llevado a cabo por el Instituto de Economía de la FCEA. Su objetivo general es “el abordaje sistemático del estudio de la evolución del bienestar desde un enfoque multidimensional, a partir de un relevamiento longitudinal representativo de los niños que concurren al sistema de educación primaria público.”

El trabajo tuvo inicio en el año 2004 y continua hasta el día de hoy. La segunda etapa fue en 2006, la tercera en 2011/12, la cuarta 2015/16 y este año comenzó la quinta. Nuestro trabajo se va a focalizar en la tercera ola cuando los que eran niños en 2004 pasaron a tener alrededor de 14 años.

Un aspecto positivo de los estudios longitudinales es que disponemos de información, a lo largo del tiempo, de la misma muestra de niños con la que se comenzó. Esto permite distinguir fenómenos de corto, medio y largo plazo, y su distribución. Además “se espera que una mayor comprensión de estos aspectos permita proporcionar insumos para el mejor diseño de políticas sociales que abarquen desde la temprana infancia hasta la adolescencia.”

Para una primera exploración de los datos utilizaremos el software estadístico R.

## OBJETIVOS

El objetivo de este trabajo es generar un primer análisis de los datos del estudio longitudinal; una mirada general que permita conocer algunas de sus variables e incentivar futuras exploraciones. Como objetivos específicos tenemos:

1. Crear una herramienta que represente una idea de la posición económica de los hogares (ICH).
2. Analizar distintos aspectos de la educación del, ahora, adolescente con respecto al ICH.
3. Analizar aspectos sobre la salud de los adolescentes.
4. Conocer parte de la situación de los padres de estos jóvenes.
5. Creación de una herramienta que permite una visualización interactiva (Shiny App).

## DATOS

Los datos contienen información acerca de los niños, los adultos responsables y las características de sus hogares; abarcando un conjunto amplio de aspectos, como salud, nutrición, desarrollo socioemocional, ingresos, educación, trabajo, participación social y autonomía.

Los datos de este estudio están divididos en dos categorías: “personas” y “hogares”. En nuestra exploración, sólo consideraremos la información proveniente de esta última categoría. Nuestra base de datos contiene 2151 observaciones y 397 variables, de las cuales utilizaremos únicamente algunas de las cualitativas. Como ser: variables relacionadas al confort del hogar utilizadas para la creación de un indicador; asistencia a centro educativo, si es público o privado; asistencia psicológica/psiquiátrica, hospitalización, edad de concepción, expectativas de los padres para sus hijos.

# PREGUNTAS

## Educación

1. ¿Cuál es la proporción de los jóvenes que han repetido algún año escolar o liceal teniendo en cuenta el intervalo de ICH al que pertenecen sus hogares?
2. ¿Cuál es la proporción de adolescentes que asisten a algún instituto educativo según el intervalo del ICH al que pertenecen sus hogares?
3. ¿Cuál es la proporción de jóvenes que asisten a centro educativo privado y público teniendo en cuenta el intervalo del ICH al que pertenecen sus hogares?
4. ¿Cuáles son las expectativas de los padres respecto al futuro educativo de sus hijos en relación al intervalo de ICH al que pertenecen sus hogares?

## Padres

1. Relación entre el ICH de los hogares y la edad en la que la madre concibió al hijo o hija.
2. ¿Cuál es la proporción de hogares en los que padre y madre conviven con el hijo o hija, según el intervalo de ICH que tienen asignado?
3. ¿Cuáles son las expectativas de los padres respecto al futuro de sus hijos en relación al intervalo de ICH al que pertenecen sus hogares?

## Salud

1. ¿Cómo perciben los padres el estado de salud de su hijo según el intervalo de ICH de cada uno de sus hogares?
2. ¿Existe relación entre las dificultades en el aprendizaje de los jóvenes y la asistencia a un profesional psicólogo/psiquiatra?
3. De los jóvenes que fueron hospitalizados en el último año (2011), ¿cuáles fueron los motivos de dicha hospitalización?

# ANÁLISIS EXPLORATORIO

Para llevar a cabo el análisis de los datos decidimos crear un indicador que toma en cuenta las principales características de los hogares y sus comodidades, con el fin de encontrar un sustituto de lo que sería la variable ingreso que perciben estos hogares.

Para ello nos basamos en el INSE (Índice de Nivel Socioeconómico), que toma en cuenta variables relacionadas con la composición del hogar, la educación y atención en salud, elementos del confort del hogar y características de la vivienda. Con esta información se elabora un sistema de puntaje en el que se asigna a cada variable un puntaje mínimo y máximo. Dependiendo de la respuesta proporcionada por el entrevistado, se obtiene una determinada cantidad de puntos por cada variable. La suma de ellos es el puntaje final.

Con la información proporcionada por este estudio longitudinal, no es posible contruir el INSE debido a la falta de muchas de sus variables. Por lo tanto, construimos un indicador utilizando el mismo sistema de puntaje que el INSE pero basado únicamente en las características y elementos de confort del hogar (ya que son los presentes en el estudio) al cual denominamos ICH (Indicador de Confort del Hogar).

Las variables que utilizamos para su creación son: vive en asentamiento, material de construcción de las paredes de la casa, si dispone de electrodomésticos tales como calefón, cocina eléctrica, microondas, equipo de música, heladera, lavarropas, lavavajillas, estufa, plancha, aspiradora, TV cable y computadora; si hay conexión a internet, automóvil y teléfono celular en el hogar. A cada una de ellas le asignamos un puntaje (ver Tablas) y para cada hogar sumamos los puntos obtenidos por cada una de estas variables. Un hogar con ICH perfecto suma un total de 100 puntos.

A continuación mostraremos la tabla de puntajes que elaboramos para construir el ICH.

	Si	No
Vive en asentamiento	0	10
Tiene calefón	7	0
Tiene cocina	7	0
Tiene micro	5	0
Tiene radio	2	0
Tiene heladera	7	0
Tiene lavarropa	6	0
Tiene lavavajilla	3	0
Tiene estufa	5	0
Tiene plancha	2	0
Tiene aspiradora	2	0
Tiene TV cable	4	0
Tiene compu no ceibal	6	0
Tiene internet	6	0
Tiene auto	8	0
Tiene celular	5	0

	Si	No
Mampostería	10	0
Madera	8	0
Metal	5	0
Adobe	3	0
Cartón	0	0
5 a 9 habs	5	0
4 habs	4	0
3 habs	3	0
2 habs	2	0
1 hab	0	0

Cabe destacar que el intervalo donde el ICH está entre 0 y 10, posee únicamente 6 observaciones. Esto implica que a la hora de analizar los gráficos no sería correcto tomar en cuenta este intervalo con la misma relevancia que el resto de los mismos.

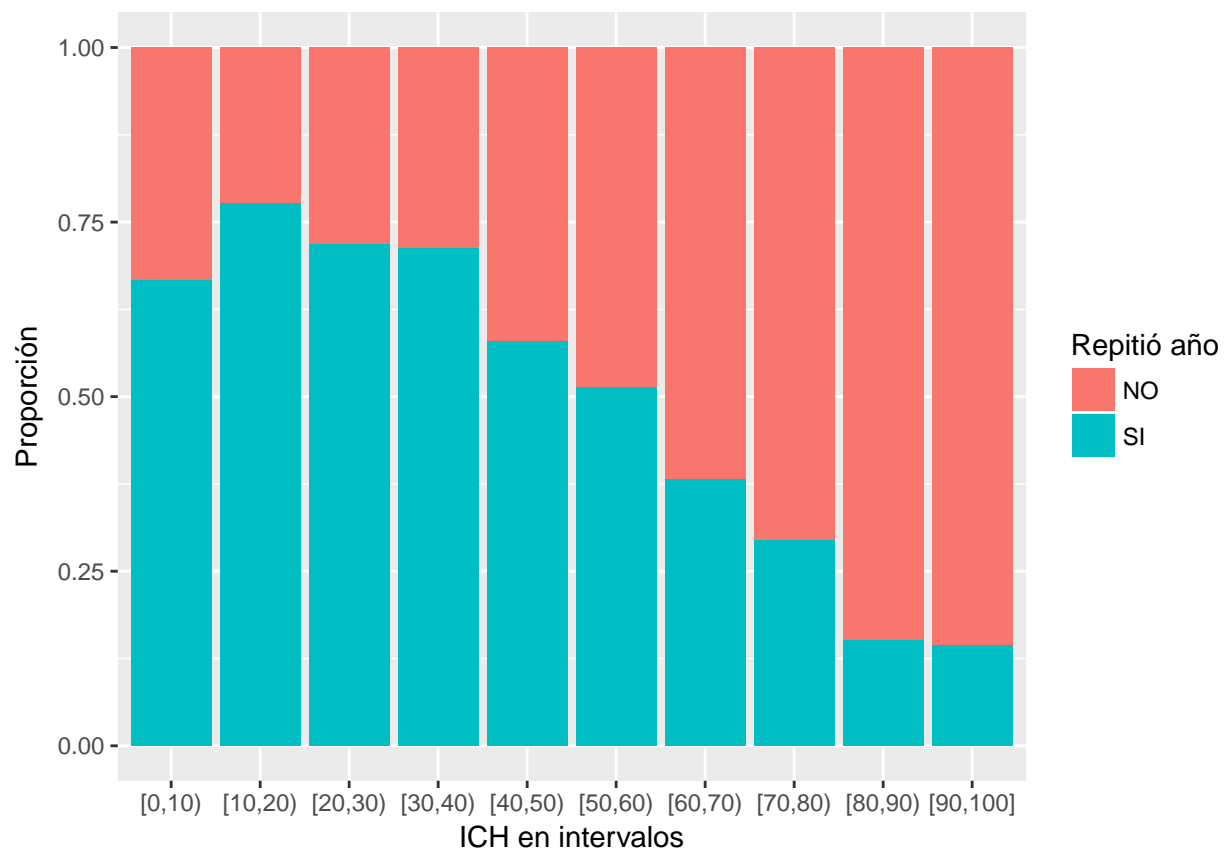


Figure 1: Gráfico de barras apiladas al 100% que muestra en celeste la proporción de jóvenes que repitieron al menos un año educativo y en rojo los que no; según el ICH.

## EDUCACION

En la Figura 1 podemos observar claramente que a medida que el ICH aumenta, disminuye la proporción de los individuos que repitieron al menos un año educativo.

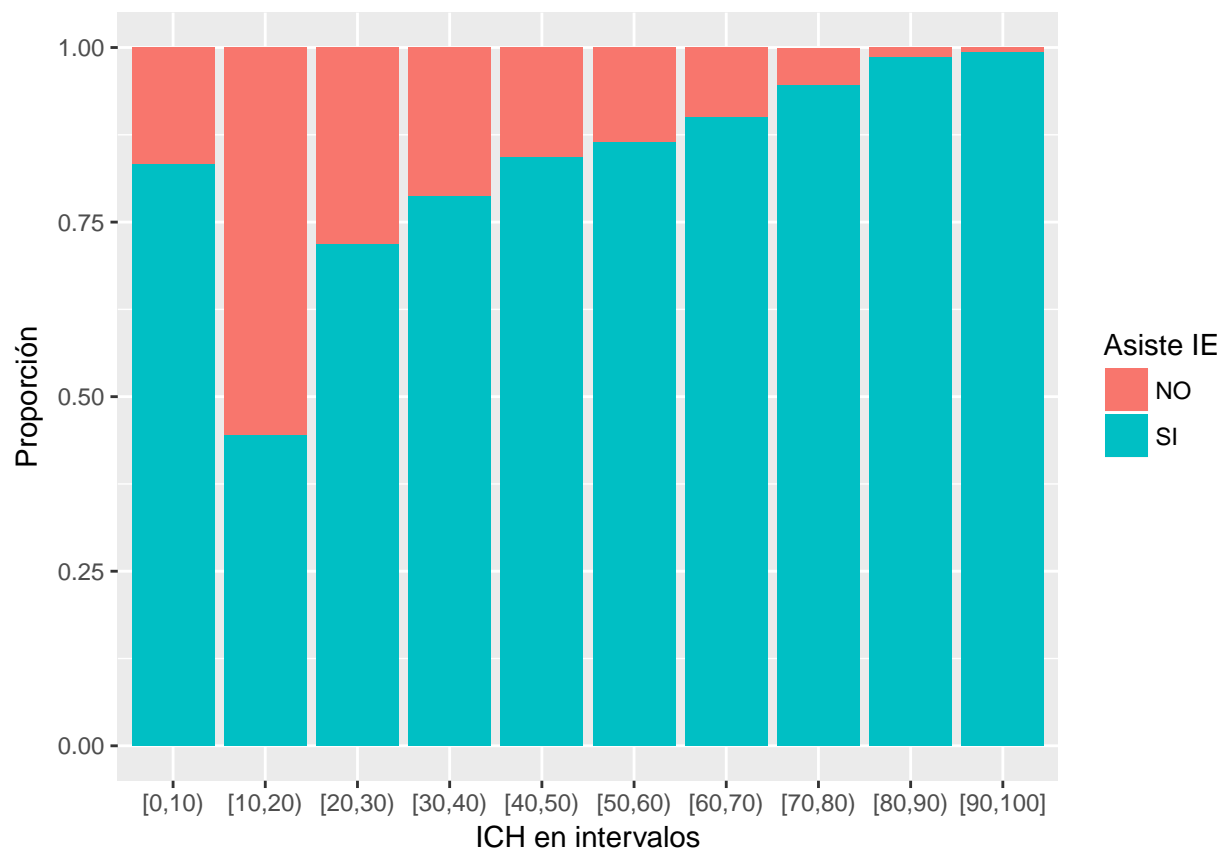


Figure 2: Gráfico de barras apiladas al 100% que muestra la proporción de jóvenes que asisten a algún centro educativo, según el ICH que alcanza cada hogar.

En esta visualización (Figura 2) se puede apreciar como, a mayor ICH hay una proporción más alta de jóvenes que asisten a centros educativos; alcanzando casi el 100% los que tienen un ICH entre 90 y 100.

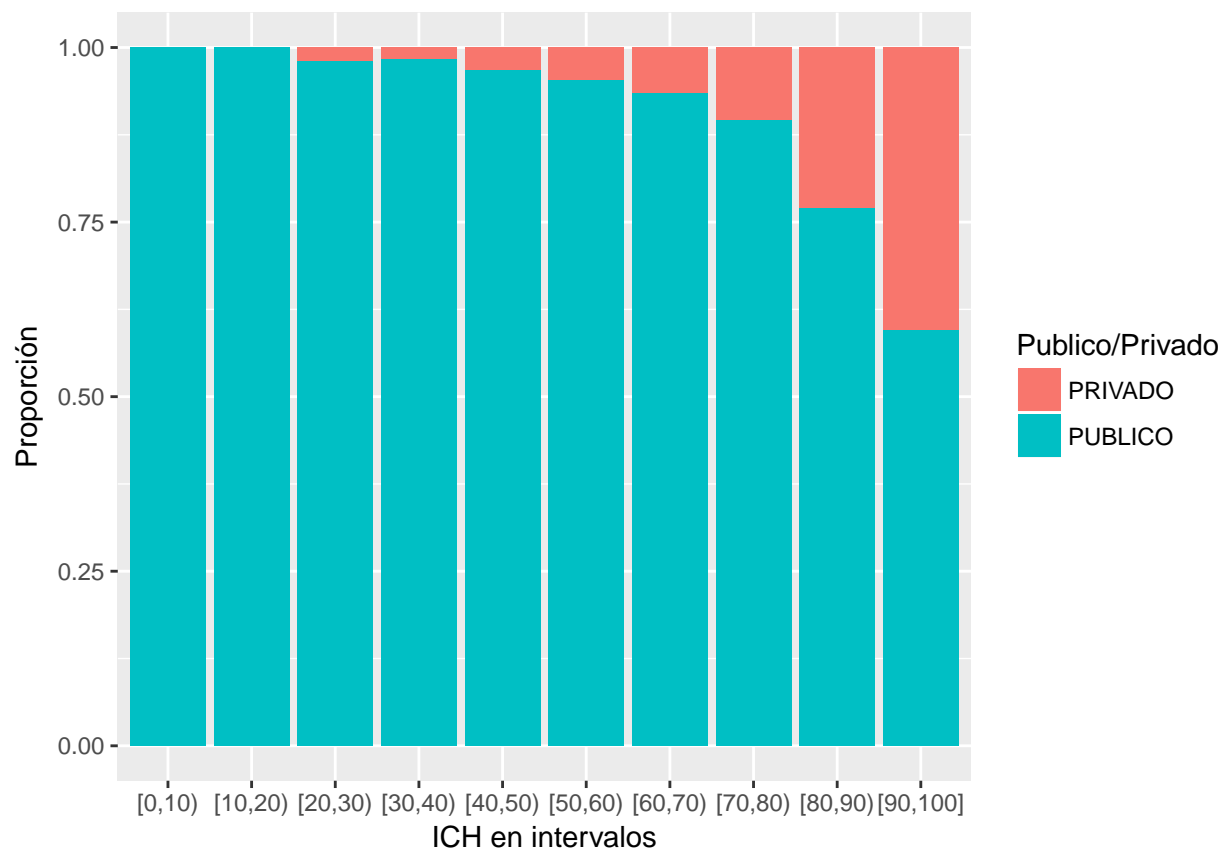


Figure 3: Gráfico de barras apiladas al 100% donde se puede ver la proporción de jóvenes que concurren a centro educativo público o privado, en función del ICH.

En la Figura 3 salta a la vista que la gran mayoría de los jóvenes entrevistados concurren a institutos públicos. Por otro lado, podemos ver que la mayor proporción de los que asisten a instituto privado se encuentra en el intervalo 90-100 del ICH.

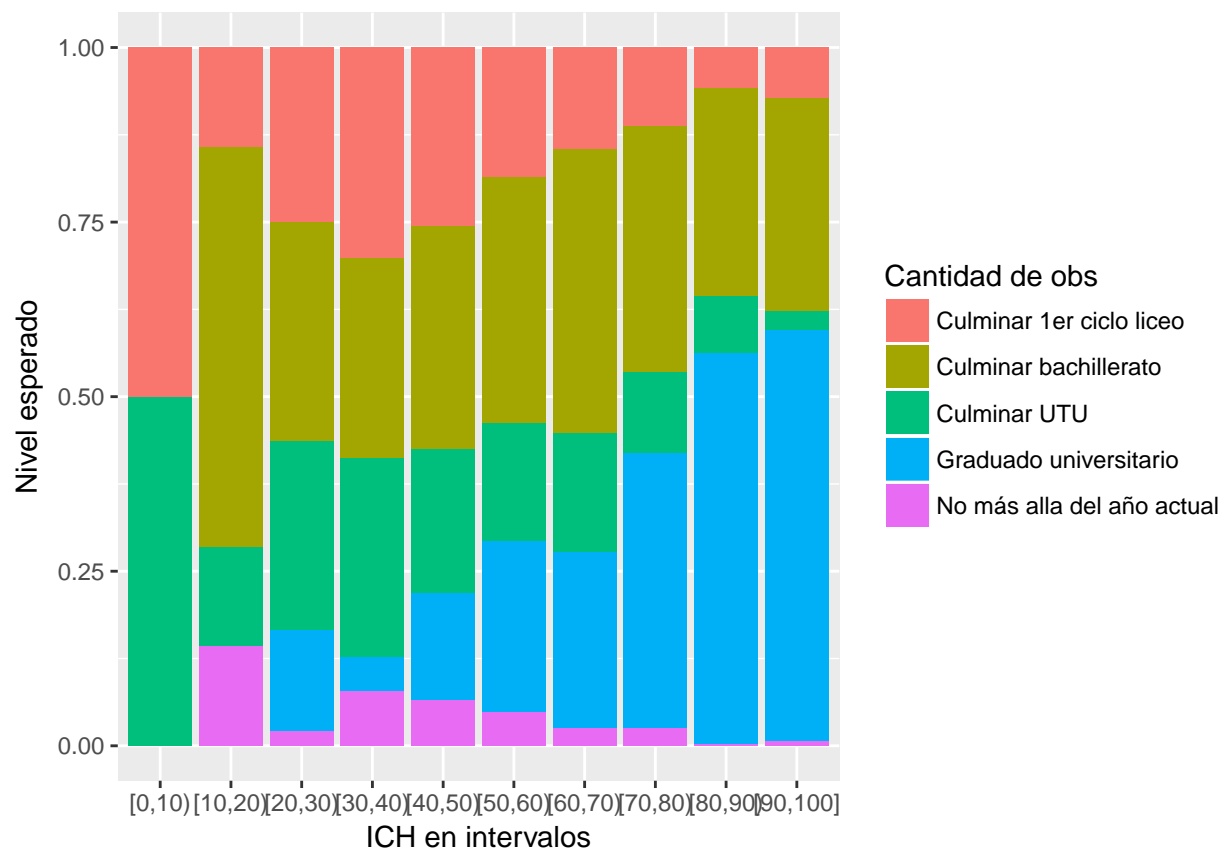


Figure 4: Gráfico de barras apiladas al 100% donde se observa la proporción de cada uno de los niveles (diferenciados por color) que esperan los padres que alcancen sus hijos a nivel educativo.

En la Figura 4 representamos las expectativas que tienen los padres al desempeño educativo de su hijo, según el ICH. Podemos observar que en los hogares con ICH más alto, la proporción de los padres con expectativas más altas (facultad y bachillerato) aumenta. Algo que resulta curioso es que los padres pertenecientes a los hogares con ICH entre 10 y 20 descartan la posibilidad de que su hijo sea graduado universitario, mientras que las expectativas respecto a la finalización del bachillerato o UTU son altas. Este intervalo, además, es el que presenta un porcentaje mayor de expectativas puestas en que el hijo culmine el año que cursa actualmente.



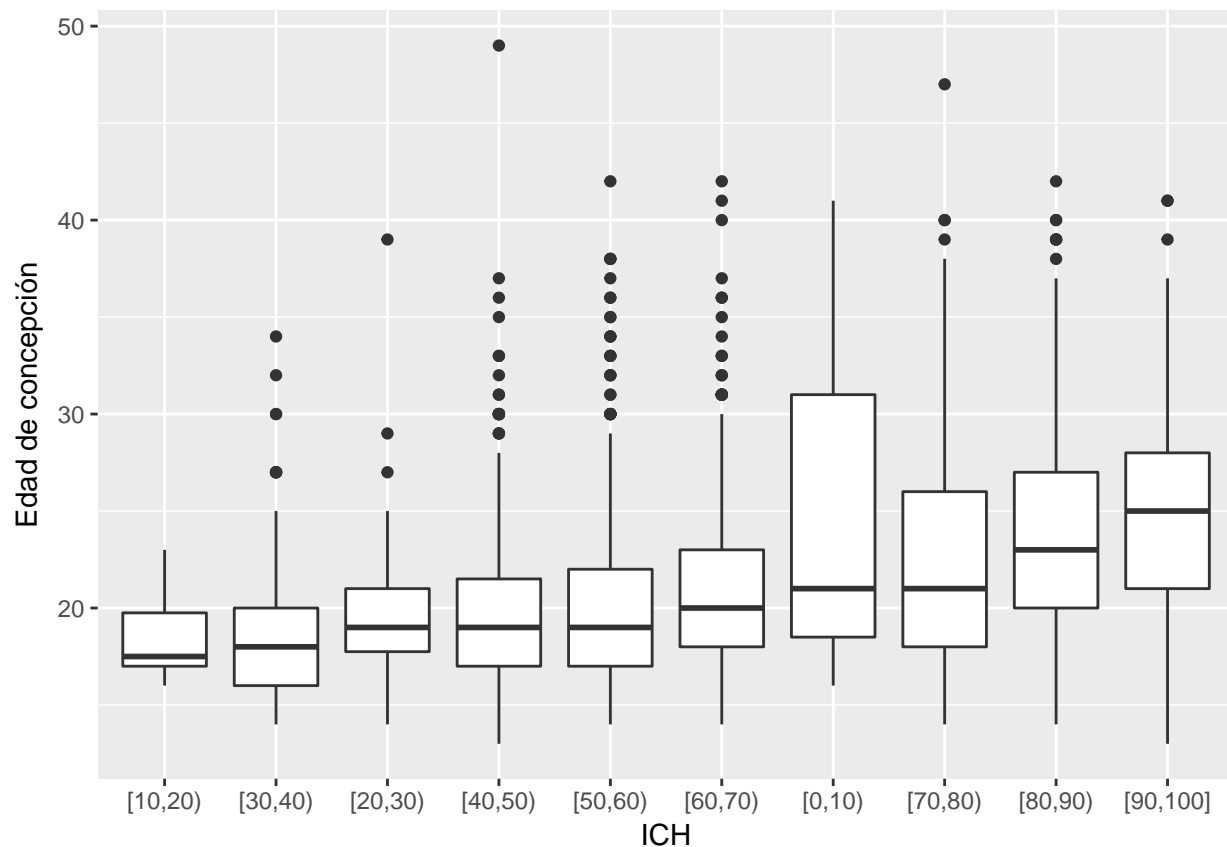


Figure 5: Gráficos de cajas para cada intervalo del ICH que muestran la edad de los padres cuando tuvieron a su hijo (que participa del estudio). Dichos gráficos se encuentran ordenados según la mediana de cada boxplot.

## PADRES

En la Figura 5 podemos notar que para los intervalos más bajos de ICH, se acumula el 50% de las observaciones en las edades menores a 20 años. En los intervalos más altos, la mediana de las edades es mayor a 20 años. Es interesante destacar que el outlier más alto se da a los 49 años de edad; y el valor más pequeño es cercano a los 14 años.

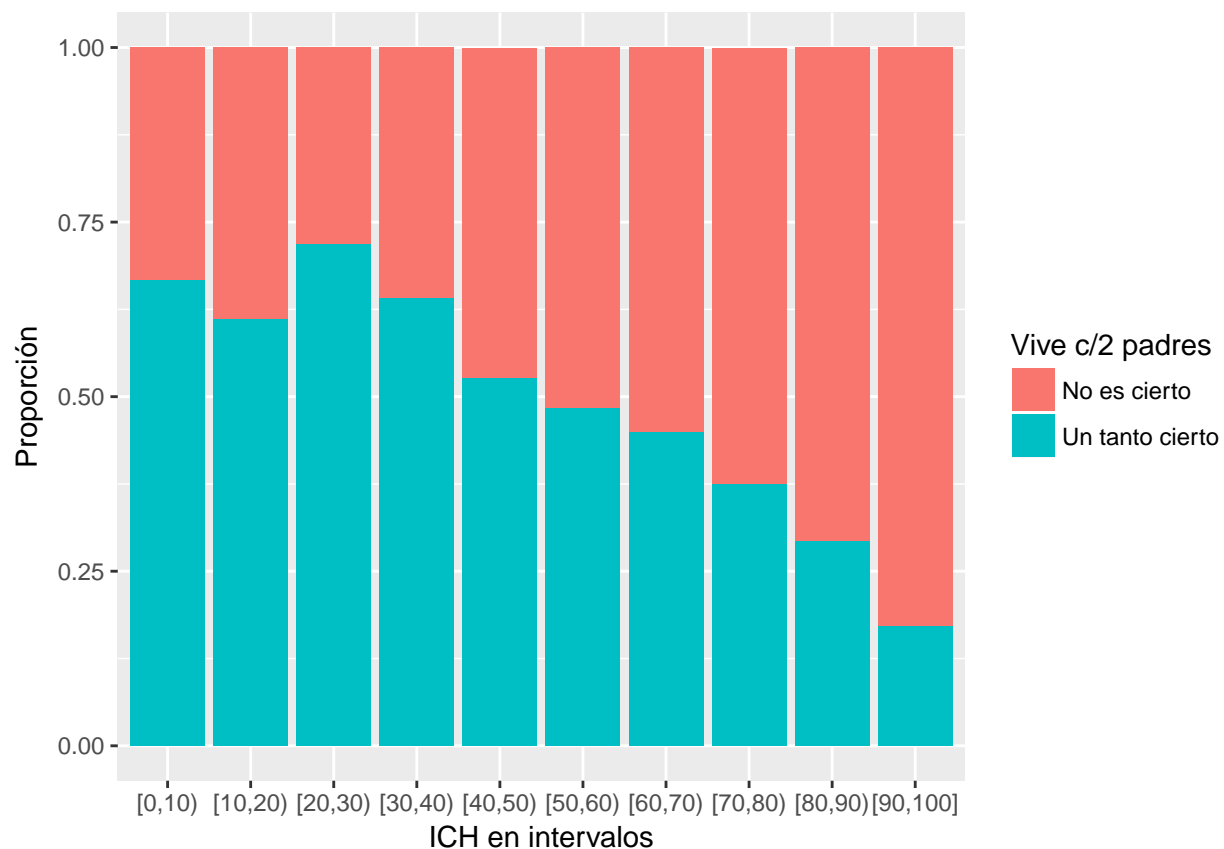


Figure 6: Gráfico de barras apiladas al 100% que permite ver las proporciones de jóvenes que viven con padre y madre, y los que no; contraponiéndolo con el ICH.

La Figura 6 muestra que, los hogares que alcanzan ICH elevados tienen menor porcentaje de padres conviviendo juntos; mientras que los hogares más humildes parecería que tienden a tener parejas más estables.

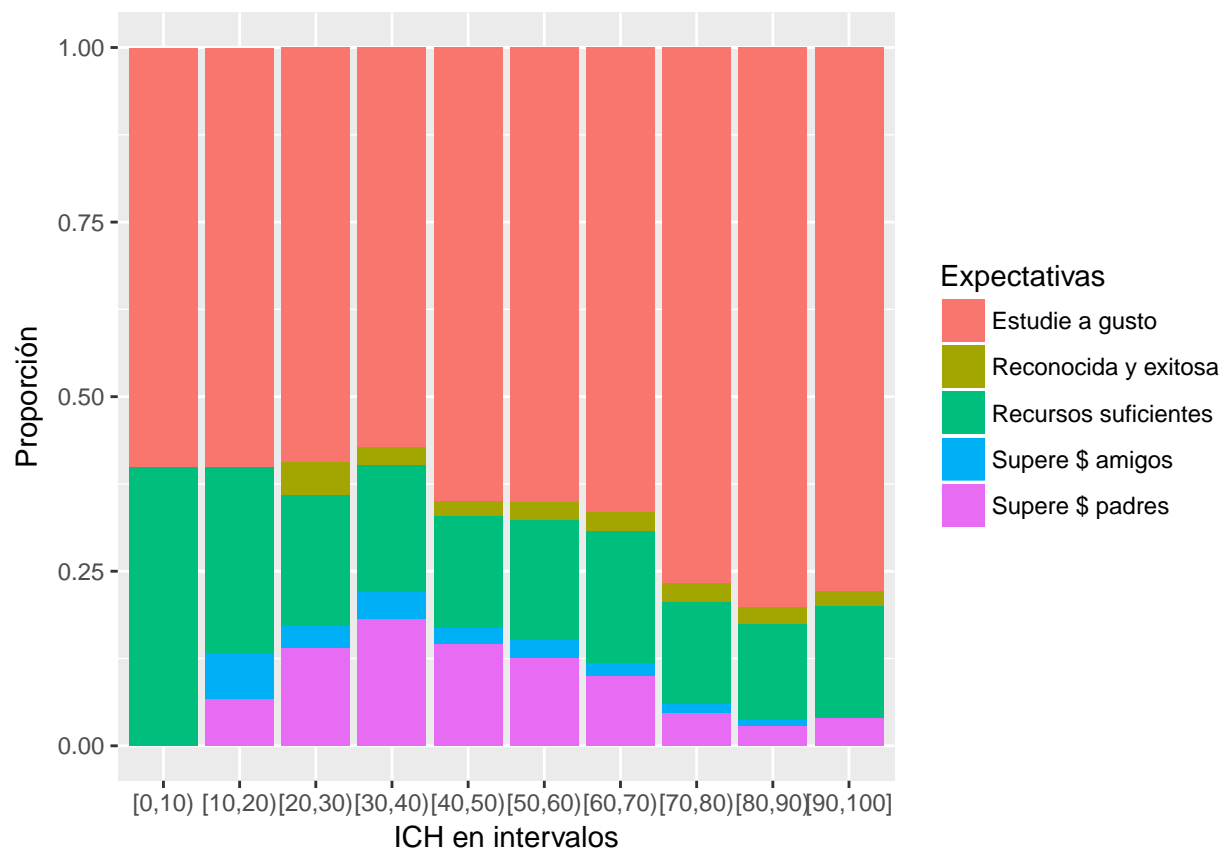


Figure 7: Gráfico de barras apiladas al 100% que muestra las expectativas de los padres respecto al futuro de sus hijos, según el ICH de cada uno de sus hogares.

La Figura 7 denota que la amplia mayoría de los padres aspiran a que sus hijos se dediquen a estudiar lo que les gusta. También se puede ver que la proporción verde (que los hijos tengan recursos suficientes) es bastante similar para todos los niveles de ICH.

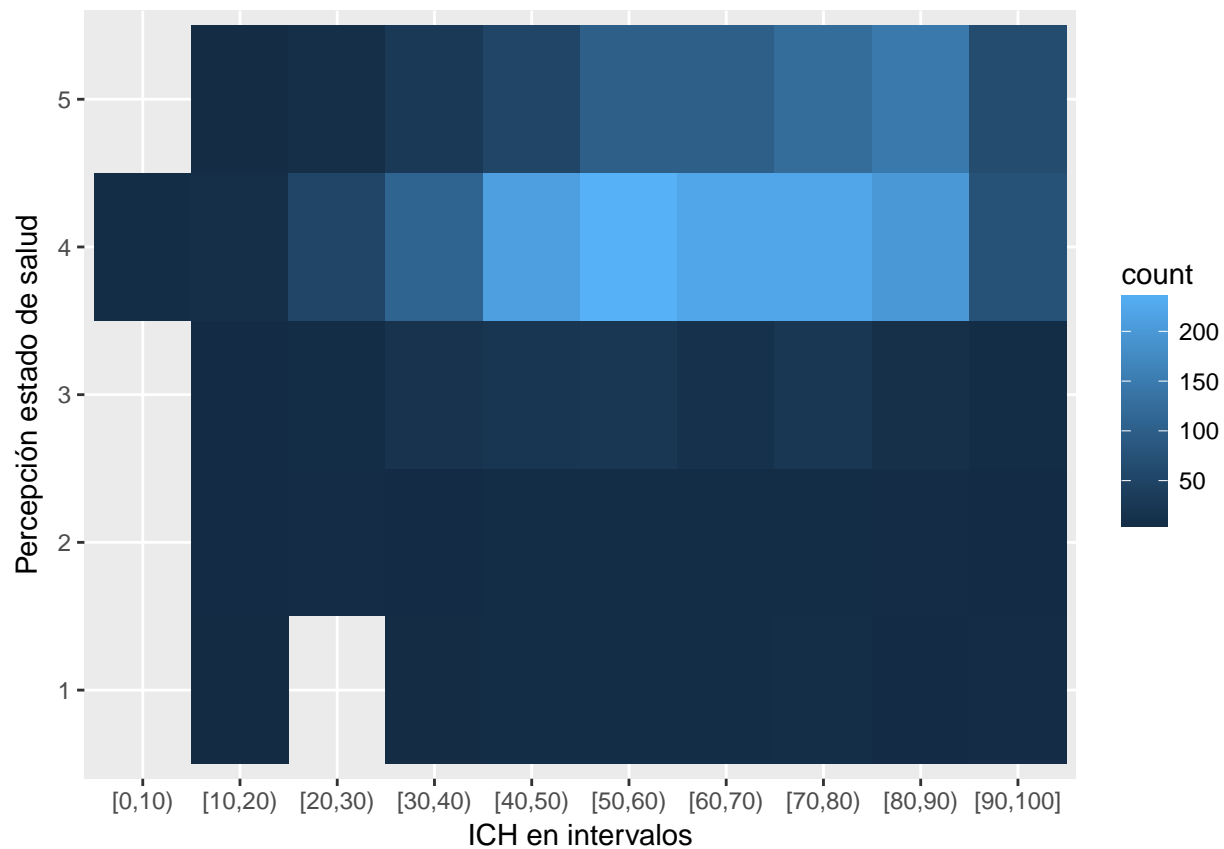


Figure 8: Gráfico geom tile acerca da la percepción de los padres sobre los niveles de salud de los hijos

## SALUD

El gráfico (figura 8) muestra que para una percepción de salud excelente (5) del hijo, la mayor cantidad de observaciones es de padres pertenecientes a hogares de alto ICH (en especial el intervalo 80-90). También podemos ver que el nivel de salud percibido que mas observaciones tiene es el 4 (buena).

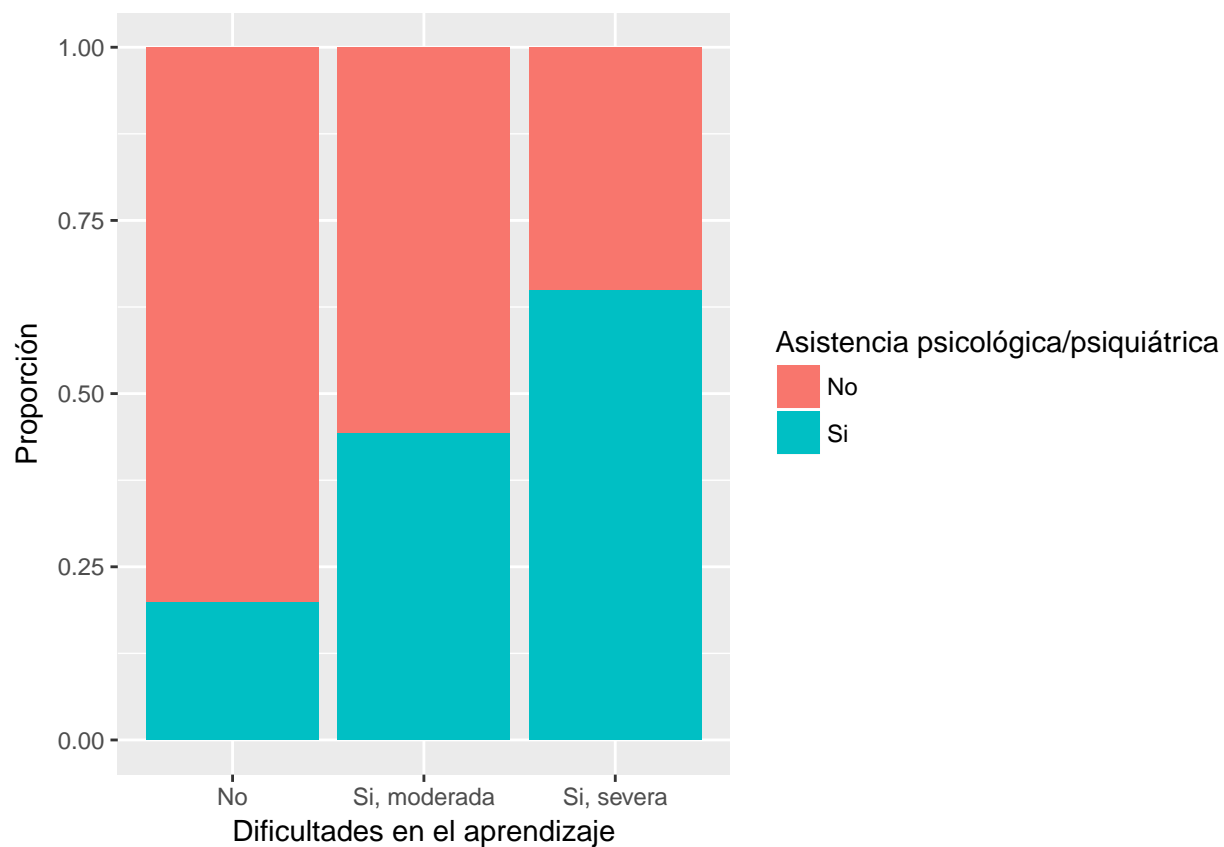


Figure 9: Gráfico de barras apiladas al 100% que contrasta las posibles dificultades en el aprendizaje de los jóvenes con la asistencia psicológica/psiquiátrica.

En la Figura 9 se puede apreciar que, a medida que los adolescentes presentan dificultades en el aprendizaje, tienden a concurrir a un psicólogo/psiquiatra en mayor proporción que los que no las presentan.

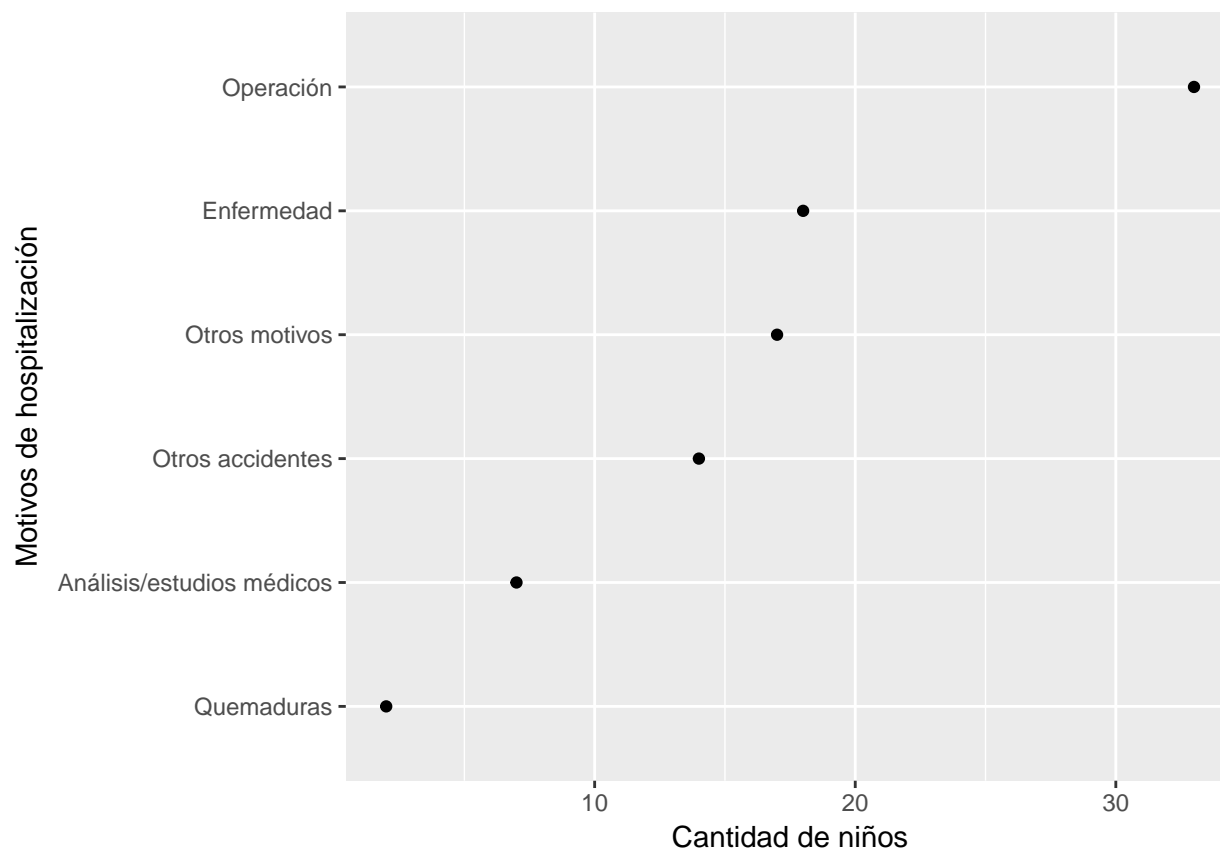


Figure 10: Gráfico de puntos que muestra la cantidad de jóvenes que estuvieron hospitalizados en el último año y los motivos de dicha hospitalización.

De la Figura 10 se puede destacar que la operación es la razón más frecuente por la que los jóvenes permanecen hospitalizados.

## DESCRIPCIÓN DE LA SHINY

Con el fin de presentar nuestro análisis exploratorio nos propusimos utilizar una herramienta que permite visualizar los gráficos y conclusiones de forma interactiva y ordenada: la aplicación Shiny.

Nuestra Shiny está ideada siguiendo el hilo de la exploración y la base de datos. Consta de cuatro pestañas donde cada una describe un aspecto de la vida de los protagonistas del trabajo (los jóvenes que en 2004 estaban en primer año de escuela).

La primera es un histograma interactivo que permite variar el ancho de las barras según la preferencia del lector. Dicho gráfico muestra las observaciones para cada puntaje del ICH que va de 0 a 100.

La segunda pestaña contiene un análisis sobre la educación de los adolescentes. Realizamos 4 gráficos de barras apiladas al 100% donde relacionamos la educación (asistencia, repetición de año, instituto público o privado y expectativas de los padres) con el intervalo del ICH al que pertenece cada hogar.

La tercer pestaña contiene tres gráficas donde se pueden ver algunos datos de los padres (edad a la que concibieron a su hijo o hija, y si ambos padres conviven con él o ella); y también una visualización con las expectativas de los padres respecto al futuro de los jóvenes.

La última pestaña presenta información sobre la salud de estos jóvenes. Podemos encontrar tres gráficos: uno sobre la percepción de los padres sobre el estado de salud de los hijos, otro de los motivos por los que fueron hospitalizados (si es que lo fueron), y uno que muestra una relación entre las dificultades en el aprendizaje y la concurrencia a psicólogo/psiquiatra.

## BIBLIOGRAFÍA

- Hadley Wickham (2017). tidyverse: Easily Install and Load the ‘Tidyverse’. R package version 1.2.1. <https://CRAN.R-project.org/package=tidyverse>
- Michael Hahsler, Christian Buchta, Bettina Gruen and Kurt Hornik (2018). arules: Mining Association Rules and Frequent Itemsets. R package version 1.6-1. <https://CRAN.R-project.org/package=arules>
- Carson Sievert, Chris Parmer, Toby Hocking, Scott Chamberlain, Karthik Ram, Marianne Corvellec and Pedro Despouy (2017). plotly: Create Interactive Web Graphics via ‘plotly.js’. R package version 4.7.1. <https://CRAN.R-project.org/package=plotly>
- Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2018). shiny: Web Application Framework for R. R package version 1.1.0. <https://CRAN.R-project.org/package=shiny>
- <http://www.cinve.org.uy/informesproyectos/indice-de-nivel-socioeconomico-inse/>
- <https://plot.ly/>
- <https://stackoverflow.com/>