

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Relación entre embarazo adolescente y desnutrición infantil en
Guatemala**

Trabajo de graduación presentado por Rodrigo Arnoldo Garoz de la
Cruz para optar al grado académico de Licenciado en Ingeniería en
Ciencia de la Computación y Tecnologías de la Información

Guatemala,

2023

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería




**Relación entre embarazo adolescente y desnutrición infantil en
Guatemala**

Trabajo de graduación presentado por Rodrigo Arnoldo Garoz de la
Cruz para optar al grado académico de Licenciado en Ingeniería en
Ciencia de la Computación y Tecnologías de la Información

Guatemala,

2023

Vo.Bo.:


(f) 

Ing. Mario Barrientos

Tribunal Examinador:

(f) 

Ing. Mario Barrientos

(f) 

Msc. Carlos Váldez

(f) 

Msc. Lynette García

Fecha de aprobación: Guatemala, 25 de mayo de 2023.

Este documento es la culminación de 5 años de viaje académico. Es significado y significativo. Lo plasmado aquí sirve como un testamento a lo hecho y el potencial de lo que está por hacerse. Estas páginas marcan el final de una etapa y el inicio de otro camino.

Agradezco a mi familia por apoyarme durante toda mi carrera universitaria. Han sido una parte muy importante de este proceso y el agradecimiento es infinito.

Agradezco a mi asesor por su guía durante el desarrollo de este trabajo.

Agradezco a todas las personas que colaboraron durante el desarrollo de este trabajo, validando todo el proceso realizado.

Agradezco a amigos y amigas con quien compartí esta etapa e hicieron este viaje más agradable.

Por último, agradezco a quien se toma el tiempo de leer este documento, espero su contenido sea enriquecedor.

Prefacio	V
Lista de figuras	X
Lista de cuadros	XI
Resumen	XIII
Abstract	XV
1. Introducción	1
2. Antecedentes	3
3. Justificación	5
4. Objetivos	7
4.1. Objetivo general	7
4.2. Objetivos específicos	7
5. Marco teórico	9
5.1. R	9
5.1.1. R Studio	10
5.2. Computación en la nube	10
5.2.1. Digital Ocean	11
5.2.2. DNS	11
5.2.3. Nginx	11
5.3. Análisis estadístico	12
5.3.1. Análisis secundario	12
5.3.2. CRISP-DM	12
5.3.3. Design Thinking	13
5.3.4. Estadística descriptiva	13
5.3.5. Regresión exponencial y logística	13
5.3.6. Chi cuadrado	16

5.4.	Situación actual en Guatemala	16
5.4.1.	Desnutrición infantil	17
5.4.2.	Embarazos en mujeres adolescentes	19
5.5.	Observatorio de Salud Sexual y Reproductiva (OSAR)	20
5.5.1.	Métodos anticonceptivos	21
5.6.	Mediciones de desnutrición crónica en Guatemala	21
5.6.1.	Factores demográficos	21
5.6.2.	Encuesta Nacional de Salud Materno Infantil (ENSMI) 2014-2015 . . .	22
5.6.3.	Línea de Base de la Gran Cruzada Nacional por la Nutrición (LBGCNN) 23	
6.	Metodología	25
6.1.	Obtención de datos	25
6.2.	Análisis inicial	26
6.3.	Sesiones de retroalimentación	31
6.3.1.	Primera sesión de retroalimentación	31
6.3.2.	Cambios implementados	32
6.3.3.	Segunda sesión de retroalimentación	32
6.3.4.	Cambios implementados	32
6.3.5.	Tercera sesión de retroalimentación	33
6.3.6.	Cambios implementados	34
6.4.	Relaciones adicionales	34
6.5.	Tablero de datos	36
6.5.1.	Despliegue y configuración de tablero	36
7.	Resultados	41
7.1.	Resultados iniciales	41
7.2.	Resultados adicionales	45
7.3.	Tablero creado	48
8.	Discusión	53
8.1.	Gráficas	53
8.2.	Regresión logística	54
8.3.	Pruebas de chi cuadrado	55
8.4.	Tablero	56
9.	Conclusiones	59
10.	Recomendaciones	61
11.	Bibliografía	63
12.	Anexos	67
12.1.	Columnas existentes	67
12.2.	Agenda de reunión	68
12.3.	Infografía creada	69
12.4.	Matrices de confusión	70
12.5.	Código de tablero	71
13.	Glosario	85

Lista de figuras

1.	Fórmula general de regresión exponencial	14
2.	Forma de función exponencial con $b > 1$	14
3.	Forma de función exponencial con $0 < b < 1$	14
4.	Fórmula generalizada de regresión logística	15
5.	Fución logística	15
6.	Fórmula generalizada para chi cuadrado	16
7.	Distribución de chi cuadrado	16
8.	Pirámide poblacional de Guatemala	17
9.	Ejemplo de estándares de talla para edad en niños	18
10.	Ejemplo de estándares de talla para edad en niñas	18
11.	Cambio en embarazo adolescente entre 2000 y 2021	19
12.	Logo de OSAR	20
13.	Portada ENSMI 2014-2015	22
14.	Portada LBGCNN 2022	23
15.	Educación v Nacimientos totales	27
16.	Educación v Nacimientos totales	28
17.	Educación v Nacimientos totales	28
18.	Educación y nacimientos totales	29
19.	Quintil de riqueza y nacimientos totales	29
20.	Edad de las madres y cuántas desviaciones estándar se alejan de la media	30
21.	Edad de las madres por grupo de edad y cuántas desviaciones estándar de la media	30
22.	Edad de las madres y número de hijo, mostrando cuántas desviaciones estándar se encuentran de la media	31
23.	Segunda sesión de retroalimentación	33
24.	Presentación de resultados ante comité	34
25.	Pantalla principal al ingresar al tablero	37
26.	Segunda pestaña en tablero	37
27.	Tercera pestaña en tablero	38
28.	Porcentaje de desnutrición moderada y severa, en dos estudios realizados	41

29.	Porcentaje de desnutrición moderada y severa, en dos encuestas realizadas, según edad	42
30.	Porcentaje de niños que sufren de desnutrición, según grupo de orden de nacimiento	42
31.	Promedio de hijos por mujer, según edad del primer hijo	43
32.	Promedio de hijos por mujer, según si fueron madres adolescentes o no	43
33.	Porcentaje de mujeres, según la cantidad de hijos antes de los 20 años	44
34.	Porcentaje de niños que sufren de desnutrición crónica, según grupo de edad de la madre al primer nacimiento	44
35.	Porcentaje de niños que sufren de desnutrición crónica, según grupo de edad de la madre al primer nacimiento, por área rural y urbana	45
36.	Porcentaje de desnutrición moderada y severa, en dos estudios realizados	47
37.	Porcentaje de desnutrición moderada y severa, en dos estudios realizados	47
38.	Porcentaje de desnutrición moderada y severa, en dos estudios realizados	48
39.	Primera pestaña en tablero	48
40.	Primera pestaña en tablero	49
41.	Segunda pestaña en tablero	49
42.	Segunda pestaña en tablero	50
43.	Segunda pestaña en tablero	50
44.	Tercera pestaña en tablero	51

Lista de cuadros

1.	Dimensiones de conjuntos de datos obtenidos	27
2.	Porcentaje de madres según máximo nivel de educación alcanzado	28
3.	Valores obtenidos de regresión logística	45
4.	5 valores p más bajos en regresión lineal realizada	46
5.	Cambio de porcentaje por unidad cambiada	46
6.	Valores de chi y valores p	46

Este trabajo consistió en un análisis secundario de los resultados de la Encuesta Nacional de Salud Materno Infantil (ENSMI) 2014-2015 y la presentación de estos resultados en un tablero que pueda ser accedido públicamente. Ya que este conjunto de datos es muy extenso, se enfocaron las tareas hacia la desnutrición crónica. Se inició obteniendo los datos de la fuente oficial, DHS. Sobre estos datos se realizó estadística descriptiva, para conocer el comportamiento general de los datos disponibles y empezar a familiarizarse con las columnas existentes.

Luego, se empezó a indagar en los datos obtenidos, obtener información adicional y plasmarlo en gráficas. Se utilizaron múltiples regresiones y clasificaciones con Chi Cuadrado para identificar asociaciones entre diferentes características de la madre y el nacimiento del niño con desnutrición crónica. El proceso se validó constantemente con el Observatorio en Salud Sexual y Reproductiva, quien identificó será el usuario principal del producto realizado. Todos estos resultados se visualizaron y agregaron a un tablero utilizando Shiny. El tablero está públicamente disponible a través del enlace público <https://rgaroz.shinyapps.io/ENSMI/>.

Tras realizar este análisis secundario y generar este tablero, se obtuvieron las siguientes conclusiones acerca de las asociaciones encontradas y del tablero creado. Se encontró una correlación entre un mayor periodo entre nacimientos y un mayor nivel de desnutrición crónica. También, se encontró una correlación inversa entre la edad del primer nacimiento y el número de hijos. Además, la cantidad de hijos antes del nacimiento se asocia con un mayor nivel de desnutrición.

This work consisted of a secondary analysis of the results of the 2014-2015 National Survey of Maternal and Child Health (ENSMI) and the presentation of these results in a publicly accessible dashboard. Since this dataset is very large, we focused the tasks toward chronic malnutrition. We started by obtaining the data from the official source, DHS. Descriptive statistics were performed on these data to understand the general behavior of the available data and to begin to familiarize ourselves with the existing columns.

Then, we began to investigate the data obtained, obtain additional information and translate it into graphs. Multiple regressions and Chi-Square classifications were used to identify associations between different characteristics of the mother and the birth of the chronically malnourished child. The process was constantly validated with the Observatory on Sexual and Reproductive Health, who identified the main user of the product, and all these results were visualized and aggregated into a dashboard using Shiny. The dashboard is publicly available through the public link <https://rgaroz.shinyapps.io/ENSMI/>.

After performing this secondary analysis and generating this dashboard, the following conclusions were obtained about the associations found and the dashboard created. An association was found between a longer period between births and a higher level of chronic malnutrition. Also, an inverse association was found between age at first birth and number of children. In addition, the number of children before birth is associated with a higher level of malnutrition.

Este trabajo se centró en el análisis de desnutrición crónica de los datos de la Encuesta Nacional de Salud Materno Infantil (ENSMI) 2014-2015 y la creación de un tablero que presente estos resultados. El análisis realizado es secundario, ya que se utilizan datos existentes para un análisis diferente del planeado originalmente. La fuente principal son las bases de datos con los resultados de la ENSMI 2014-2015. Se utilizó la librería Shiny del lenguaje de programación R para realizar el análisis secundario y la visualización de datos.

Se eligió esta temática tras identificar una necesidad por parte del Observatorio de Salud Sexual y Reproductiva (OSAR). OSAR nace como parte de un acuerdo entre el Ministerio de Salud Pública y Asistencia Social (MSPAS) y miembros de la sociedad civil, con el objetivo de monitorear y fiscalizar la implementación de las políticas públicas relacionadas a Salud Pública. Esta organización presentó y sugirió las temáticas a utilizarse durante este trabajo.

Las temáticas principales fueron relación entre desnutrición [crónica] infantil y edad de las madres (adolescentes) o número de hijo en la familia. Esta necesidad nace del acuerdo de trabajo que existe entre OSAR y la Secretaría de Seguridad Alimentaria y Nutricional (SESAN), buscando promover la salud y nutrición sostenible de la población guatemalteca. Además, se mencionó la intención de incluir la prevención de embarazo adolescente en la agenda de desnutrición crónica.

Se empezó obteniendo las bases de datos con los resultados de la ENSMI 2014-2015 a través del sitio de las encuestas de demografía y salud (DHS, por sus siglas en inglés). Estas encuestas son realizadas por la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID, por sus siglas en inglés). Se realizan de manera estandarizada en múltiples países en vía de desarrollo, siguiendo un formato estándar. El acceso a estos datos debe ser solicitado directamente a través de la página de DHS, debe ser utilizado para fines académicos y no pueden compartirse los datos con personas ajenas a la investigación.

Estos resultados se encuentran separados en 5 archivos, con información acerca de nacimientos, hogares, personas encuestadas, niños, hombres y otros miembros del hogar. En este trabajo, únicamente se utilizan los datos de nacimientos, individuos y niños, ya que

contienen los datos demográficos de las mujeres encuestadas y los datos nutricionales de sus hijos. Los datos de hombres y otros miembros del hogar no son necesarios para cumplir con el alcance definido.

Estos datos fueron explorados inicialmente utilizando Python y Pandas, para comprender mejor sus características generales y el comportamiento de sus variables. Se identificaron los tipos de columnas, se realizaron agrupaciones por diferentes características demográficas y se observó el rango de datos posibles en cada columna. Se utilizaron estos lenguajes ya que cuenta con herramientas disponibles para este tipo de análisis.

Luego, se empezó a utilizar R para procesar y obtener información adicional sobre estos datos. Se eligió R ya que cuenta con la librería Shiny, la cual permite realizar transformación y análisis de datos y crear tableros de datos web basados en estos resultados. Se realizaron regresiones logísticas y exponenciales para identificar nuevas asociaciones entre diferentes variables. También, se utilizó análisis con Chi Cuadrado para explorar asociaciones entre variables categóricas.

Estos resultados fueron graficados y colocados en una presentación, la cual fue validada durante múltiples sesiones con personas de diferentes organizaciones, como OSAR, SESAN y el Programa Nacional de Salud Reproductiva (PNSR). En estas sesiones se recibió retroalimentación acerca del análisis realizado y sus posibles usos, particularmente en la concientización del embarazo adolescente como factor de riesgo para la desnutrición infantil.

Los resultados validados fueron agregados a un tablero, donde se pueden observar los datos de manera general o agrupar los datos por quintil de riqueza, máximo nivel de educación alcanzado y área de residencia. Se configuró un servidor en una máquina virtual en Digital Ocean para poder servir este tablero de manera pública. A esta máquina virtual se le configuró un registro DNS, para poder ser accedido a través de la dirección <http://osar-desnutricion.com:3838/osar/>.

Tras realizar estas tareas, se concluyó que existe una correlación entre el número de nacimiento, peso y talla de la madre y periodo entre nacimientos y el nivel de desnutrición de un niño. Esto muestra la importancia que tiene reducir el embarazo adolescente para reducir la desnutrición infantil. Atrasar el primer nacimiento significa disminuir la cantidad total de hijos, lo cual disminuye el nivel de desnutrición para cada hijo.

En algunos países, se ha identificado una relación entre la edad de la madre y el riesgo de que un niño sufra de desnutrición crónica. En Ghana se realizó un estudio que estableció que los niños de madres adolescentes eran 8 veces más probables de sufrir de un retraso en el crecimiento que los niños de madres adultas [1].

Este fenómeno también se ha identificado en otros países, como India, donde se encontró que los niños de madres adolescentes tienen menor talla para edad, peso para edad y peso por talla que los hijos de madres adultas [2]. En El Salvador se identificó un patrón similar, donde los hijos de madres adolescentes mostraron un mayor retardo en crecimiento que cuando sus madres eran de edad avanzada [3]. En Bangladesh, se identificó que los niños con madres adolescentes tienen una mayor probabilidad de sufrir de desnutrición crónica [4].

Un estudio realizado con datos de mujeres de toda Latino América entre 1985 y 2003 muestra que las madres adolescentes tenían un mayor riesgo de complicaciones postparto para la madre y para los hijos. En este estudio se observó que todas las madres adolescentes tenían mayor riesgo de hemorragia postparto, partos prematuros y niños con bajo peso al nacimiento. Estos resultados se obtuvieron a través de la regresión logística, para obtener los coeficientes de probabilidades ajustados [5].

En el año 2023, se estima que Guatemala tiene una población de 17.3 millones de personas, de los cuales el 31 % es menor de 15 años, el 63 % tiene entre 15 y 64 años y cerca del 6 % tiene más de 65 años [6].

En Guatemala, la tasa específica de fecundidad en mujeres de 15 a 19 años es de 92 por cada mil mujeres en ese grupo de edad [7]. A nivel mundial, para el año 2015, la tasa específica de fecundidad de mujeres de 15 a 19 años era de 114 por cada mil mujeres en ese grupo de edad. Naciones Unidas estima que para el año 2021 esta tasa es de 61 en Guatemala y de 41 a nivel mundial [8]. El embarazo adolescente y la maternidad temprana tienen consecuencias para la vida de las mujeres y para sus hijos, hay evidencia de que los hijos de madres adolescentes tienen mayor riesgo de bajo peso al nacer y de ser prematuros [1].

Por otro lado, el informe de la Línea de Base de la Gran Cruzada Nacional por la Nutrición reporta que para el año 2022, 46 % de los niños y niñas menores de 5 años tienen algún grado de desnutrición crónica [9]. La desnutrición crónica es un retraso en la talla para la edad, se conoce que sus efectos van más allá de una baja estatura porque tiene “consecuencias irreversibles, como retraso en el desarrollo motor, deterioro de la función cognitiva y bajo rendimiento escolar” en la vida de las personas [10].

En Guatemala, la desnutrición infantil y el embarazo adolescente muchas veces se abordan como dos problemas diferentes. Sin embargo, en otros países se han identificado una relación entre ambos, por ejemplo, un estudio mostró que en Ghana los hijos de madres adolescentes tiene 8 veces más probabilidades de sufrir de desnutrición, en comparación con los hijos de madres adultas [1]. Guatemala tiene a nivel nacional una alta tasa de fecundidad entre las mujeres adolescentes y uno de cada dos niños sufre de desnutrición crónica, pero hay poca información utilizando datos de Guatemala sobre la relación entre estos dos problemas.

OSAR cuenta con un acuerdo de trabajo con SESAN, buscando promover la salud y nutrición sostenible de la población guatemalteca. Este acuerdo de trabajo plantea la necesidad

de fortalecer el tema de atención prenatal y salud reproductiva con salud y nutrición. Además, busca abordar la atención interinstitucional de las adolescentes embarazadas siempre y cuando se encuentra alineadas a la Ley del Sistema Nacional de Seguridad Alimentaria y Nutricional (SINASAN) y GCNN. Dentro de este acuerdo es que nace la necesidad de abordar en conjunto el embarazo adolescente y la desnutrición crónica [11].

El análisis de las relaciones que puedan existir entre ambos problemas y la representación gráfica de los resultados, pueden servir como evidencia a OSAR para sus acciones de monitoreo de políticas públicas. Además, puede mostrar a los tomadores de decisiones la importancia de abordar estos temas de una manera más integral.

4.1. Objetivo general

Identificar qué relación existe entre el embarazo adolescente y la desnutrición infantil en Guatemala.

4.2. Objetivos específicos

- Identificar los indicadores más importantes que se pueden obtener de los datos obtenidos en ENSMI.
- Realizar análisis secundario sobre la relación entre la desnutrición infantil, edad del primer nacimiento, número total de hijos, disponibilidad de servicios básicos y quintil de riqueza.
- Presentar los resultados en un tablero para su fácil acceso y uso.

Dados los objetivos de encontrar asociaciones adicionales en un conjunto de datos existente y construir un tablero con los resultados, es necesario plantear algunas definiciones. Primero, se inicia planteando las herramientas tecnológicas, incluyendo el lenguaje de programación y la librería a utilizar. Luego, los conceptos principales de la computación en la nube, para comprender cómo se publicó la herramienta creada. Por último, se expande y profundiza en la problemática identificada, la desnutrición crónica infantil en Guatemala.

5.1. R

R es un lenguaje de programación utilizado principalmente para computación estadística y para creación de gráficas. Fue creado en 1993 por los estadígrafos de la Universidad de Auckland Ross Ihaka y Robert Gentleman, para ser utilizado para aprender estadísticas introductorias. Actualmente es utilizado principalmente para minería de datos, bioinformática y análisis de datos. A través de su ciclo de vida, la comunidad ha creado múltiples paquetes para expandir las funcionalidades del lenguaje [12].

Según el índice TIOBE, el cual mide la popularidad de lenguajes de programación, en marzo de 2023 R es el decimocuarto lenguaje de programación más utilizado, con una calificación de 0.93 % [13]. Esto muestra una disminución durante el año pasado, cuando se encontraba en decimoprimer posición, con una calificación de 1.37 % [13]. Esta calificación se basa estrictamente en los resultados obtenidos en motores de búsqueda, no implica necesariamente la frecuencia de su uso [13].

Por otro lado, GitHub, el cual mide la cantidad de *pull requests* realizados en este lenguaje de programación, muestra que a finales de 2022, los proyectos escritos en R contuvieron 0.048 % de todos los *pull requests* realizados en GitHub [14]. Esta calificación lo coloca en la posición 42 entre todos los lenguajes de programación. Esta métrica también muestra que el uso de R ha disminuido en los últimos años, en 2021, se encontraba en la posición 33, representando 0.078 % de estos eventos.

5.1.1. R Studio

R Studio es un ambiente de desarrollo integrado (IDE, por sus siglas en inglés) para el lenguaje R. Cuenta con dos versiones: RStudio Desktop, utilizado para escritorio y RStudio Server, utilizado para servidores que sirven instancias de RStudio. Está disponible para múltiples sistemas operativos, Windows macOS y múltiples distribuciones de Linux. La mayoría de RStudio está escrito en C++, con algunas partes escritas directamente en R [15].

Es desarrollado por Posit, una empresa centralizada en la Ciencia de Datos y que también provee herramientas para otros lenguajes de programación, como Python. No tiene ningún tipo de relación con la fundación sin fines de lucro R, encargados de desarrollar el lenguaje de programación R [16]. Además de RStudio, Posit desarrolla y mantiene múltiples paquetes y proyectos adicionales para R, entre ellos *Tidyverse* y *Shiny*. Gracias a esto, RStudio cuenta con funcionalidades específicas para facilitar el uso de estas herramientas [15].

Tidyverse es una colección de paquetes para manejo de datos, que provee capacidades para manejo y modificación de datos, con capacidades adicionales para visualización de datos. *ggplot2* es la herramienta utilizada para visualización de datos. Se basa en la gramática de gráficas, la cual sostiene que todas las gráficas se pueden definir con un único lenguaje, basándose en capas de visualización sobre una capa compartida de estética. Esto permite un mayor control sobre la visualización creada, ya que se tiene control sobre cada componente de las gráficas [17].

Shiny es un paquete que permite construir herramientas interactivas para navegadores web, directamente en R. Esto permite utilizar todas las herramientas existentes en R para análisis y visualización de datos como parte de un tablero web, el cual puede ser compartido y publicado posteriormente, a través de un servidor web. Además, estas visualizaciones pueden ser aumentadas con funcionalidades para hacerlas interactivas, con todas las posibilidades permitidas por las tecnologías web modernas [18].

Las aplicaciones escritas en *Shiny* contienen dos partes, el servidor y la interfaz gráfica. En el servidor se encuentran la definición de las gráficas y el procesamiento de datos necesarios para poder realizar las gráficas. En los archivos de interfaz gráfica, se utilizan las gráficas generadas y se ordenan con un lenguaje parecido a HTML, con conceptos de filas, columnas y diferentes etiquetas para modificar la apariencia de los elementos. Ambas partes forman una aplicación, la cual es armada por *Shiny* y puede ser publicada a través de un servidor local, un servidor público propio o a través del servicio proveído por Posit [19].

5.2. Computación en la nube

La computación en la nube es la disponibilidad de sistemas de computación bajo demanda, sin necesidad de administración directa por parte del usuario. Estos recursos principalmente son almacenamiento de datos y poder de computación. Durante los últimos años ha sido un modelo de consumo de recursos que ha estado apogeo, dada la facilidad y flexibilidad que provee. Además, disminuye los costos capitales, distribuyéndolos a través del tiempo en costos operativos, ya que muchos proveedores utilizan el modelo de pagar únicamente por los recursos que se utilizan [20].

5.2.1. Digital Ocean

Digital Ocean es una empresa multinacional de tecnología y proveedor de infraestructura en la nube. Cuenta con 15 centros de datos a nivel global y desde estos centros provee múltiples servicios digitales. El servicio para máquinas virtuales se conoce como *droplets*, las cuales pueden ser creadas con diferentes tamaños y sistemas operativos, dependiendo de las necesidades identificadas. Además, se ofrecen múltiples configuraciones y funcionalidades adicionales sobre estas máquinas virtuales [21].

Se pueden definir *firewalls* para regular el acceso y las conexiones a estas máquinas virtuales, lo cual se ve reflejado en las configuraciones de *ufw* a nivel sistema operativo. También, se pueden configurar múltiples características de red dentro de las redes de las máquinas virtuales. Esto incluye balanceo de carga, para lidiar con posibles cambios de demanda. También, DigitalOcean cuenta con *nameservers* propios, por lo cual se puede realizar configuraciones de DNS dentro de DigitalOcean. Para esto se debe contar con un dominio, el cual debe registrarse dentro de estos servidores [22].

5.2.2. DNS

El sistema de nombres de dominio (DNS, por sus siglas en inglés), es un sistema jerárquico para nombrar de manera distribuida los recursos a través del Internet. Se utiliza para asociar nombres de dominios humanos en direcciones IP, las cuales son utilizadas por los protocolos de red. Para lograrlo se utilizan *nameservers* autorizados, los cuales cuentan con la información global de esta traducción entre nombres y direcciones IP. Estas configuraciones son propagadas a través de múltiples *nameservers* de manera global, los cuales son accedidos al momento de realizar una consulta [23].

Al momento de configurar un nuevo registro, este debe registrarse en los *nameservers*, para que el tráfico pueda redirigirse correctamente. Debido a la estructura descrita anteriormente, estos cambios no se realizan de manera inmediata. Se debe dar suficiente tiempo para que estos cambios se propaguen a través de toda la red de servidores y todas las consultas puedan utilizar la información más actualizada [23].

5.2.3. Nginx

Nginx es un servidor web que también provee capacidades para balanceo de carga, servidor proxy y un cache HTTP. Puede configurarse fácilmente para servir contenido estático a través de un servidor web. Es una herramienta de código abierto y actualmente es el servidor más utilizados nivel mundial, habiendo superado recientemente al servidor web Apache. La documentación oficial de DigitalOcean utiliza Nginx como el servidor web por defecto para servir aplicaciones [24].

5.3. Análisis estadístico

Existen múltiples técnicas para encontrar asociaciones entre diferentes variables en un conjunto de datos. Muchas veces se busca encontrar la asociación entre una o múltiples variables independientes y el impacto que estas tienen en una única variable dependiente. Muchas técnicas cuentan con un número o estadístico para validar la calidad de la predicción hecha. Esto permite comparar diferentes iteraciones utilizando las mismas técnicas [25].

5.3.1. Análisis secundario

Un análisis secundario se realiza cuando se utilizan datos existentes de manera diferentes al trabajo original. Esta técnica puede tener múltiples ventajas, principalmente en la reducción de costos, ya que no se debe gastar en la recolección de datos. Sin embargo, se sigue necesitando rigor al momento de realizar los procesos de investigación, principalmente al momento de entender la metodología utilizada para la recolección de datos y el impacto que esto pueda tener en las metodologías futuras [26].

5.3.2. CRISP-DM

El proceso estándar trans-industrial para minería de datos (CRISP-DM, por sus siglas en inglés) es el estándar más utilizado para guiar el proceso de minería de datos. Es un estándar abierto creado por un equipo de 200 organizaciones, con fondos de la Unión Europea. Originalmente fue desarrollado para minería de datos, pero se puede utilizar para múltiples estilos analíticos [27]. Se definen 6 fases como parte de este estándar [28]:

1. Entendimiento del negocio: En esta fase se busca entender los objetivos del negocio, los recursos con los que se cuenta y cuál es el objetivo que se tiene. Esta fase sienta las bases para las tareas que se realizarán en las siguientes fases.
2. Entendimiento de datos: Se recolectan, describen y valoran los datos con los que se estarán trabajando. Sirve como la exploración y acercamiento inicial.
3. Preparación de datos: Se empieza a seleccionar, limpiar y construir los datos. Se integran las diferentes fuentes de datos y se les da el formato adecuado.
4. Modelado: Se seleccionan los modelos a utilizar y se construyen los modelos para cumplir con los objetivos establecidos.
5. Evaluación: Se obtienen y evalúan los resultados obtenidos en las fases anteriores. Se revisa el proceso realizado y se determinan los siguientes pasos.
6. Despliegue: Se crea el producto final y se despliega para uso. Además, se monitorea y se le da mantenimiento durante su funcionamiento.

5.3.3. Design Thinking

Design Thinking es un proceso y enfoque para la resolución de problemas e innovación. Es un proceso no lineal e iterativo para entender a usuarios y crear soluciones apropiadas para estas soluciones. Busca centrarse en la solución del problema, más que en el problema mismo. Su carácter iterativo permite recibir retroalimentación constante del cliente e implementarla durante todas las fases de desarrollo de la solución [29]. En la mayoría de la literatura, se conceptualizan las siguientes 5 fases [30]:

1. Empatizar En la fase inicial se busca entender la necesidad del usuario, generalmente a través de empatía.
2. Definir Se busca acumular la información obtenida en la fase anterior. Se concentra la información para obtener los problemas centrales que se han identificados.
3. Idear Se empiezan a generar ideas utilizando el conocimiento adquirido en las primeras dos fases.
4. Prototipar Se busca identificar las mejores soluciones posibles para los problemas encontrados. Se crean pequeños prototipos para ser validados con el usuario.
5. Probar Se validan las posibles soluciones con los usuarios finales. A partir de esta retroalimentación se puede regresar a fases anteriores.

5.3.4. Estadística descriptiva

Las estadísticas descriptivas son un conjunto de técnicas utilizadas para resumir y describir una variable o conjunto de variables dentro de un conjunto de datos. Esto sirve en contraparte a las estadísticas inferidas, las cuales buscan obtener información adicional acerca de la población de la cual fue tomada la muestra analizada. La estadística descriptiva puede utilizarse para resumir únicamente una estadística o para analizar relaciones o asociaciones entre múltiples variables. A pesar de que ambos tipos de estadísticas tienen usos diferentes, muchas veces se utilizan ambas al momento de realizar análisis [25].

La estadística descriptiva permite entender características generales de la población de una manera resumida. En caso de conjuntos de datos extremadamente largos, esto facilita mucho la comprensión de sus características. Estos resultados pueden ser un buen inicio para análisis de datos, ya que dan las características generales de la población con la que se estará trabajando. Este conocimiento puede servir como un punto de partida para análisis posteriores, guiando la elección más apropiada para las técnicas a utilizar [25].

5.3.5. Regresión exponencial y logística

La regresión exponencial se utiliza para encontrar una función de tipo exponencial que mejor se adapte a los datos observados con el tipo:

$$y = ab^x$$

Figura 1: Fórmula general de regresión exponencial

Esta fórmula predice los conjuntos de datos donde el crecimiento relativo es constante a través de todos los datos, mientras que la regresión lineal se utiliza donde el crecimiento absoluto es constante. En datos con este tipo de comportamiento, el crecimiento cercano a 0 es mayor al cambio en los datos posteriores. Si el valor de b es superior a 1, el crecimiento es positivo y si $0 < b < 1$ el crecimiento es negativo [31].

Para calcular la calidad de esta regresión, se puede utilizar el estadístico de prueba R^2 . Este indica qué tanto se acerca la predicción a los valores realmente observados. Se calcula con la suma de todas las diferencias entre todos los valores observados y los valores obtenidos a través de la regresión exponencial [31].

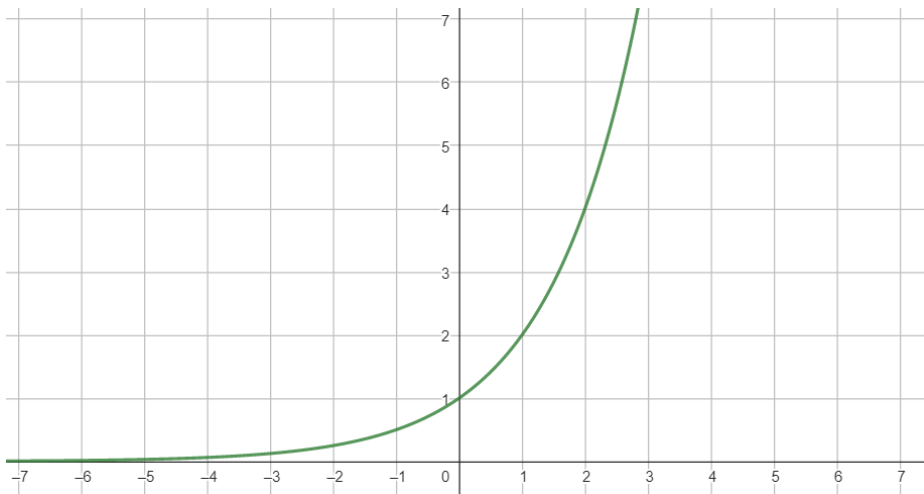


Figura 2: Forma de función exponencial con $b > 1$

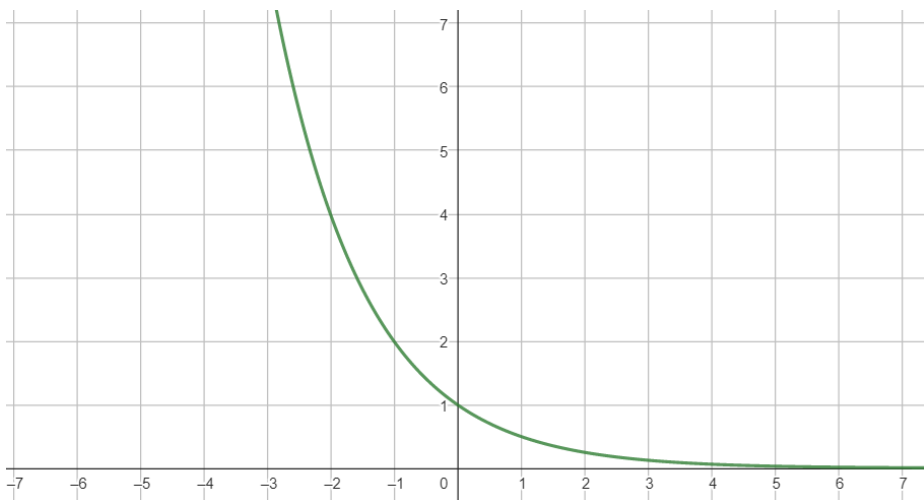


Figura 3: Forma de función exponencial con $0 < b < 1$

La regresión logística es un modelo estadístico que provee la probabilidad de que un evento ocurra a través de una combinación lineal de múltiples variables independientes. También, puede utilizarse para estimar los parámetros en un modelo logístico, obteniendo los factores de la combinación lineal. La combinación lineal puede ser una variable categórica o una variable logística, pero el resultado debe ser binario, contando únicamente con dos posibles resultados. Existen extensiones para este modelo que permite utilizarlo como una herramienta de clasificación [31].

La fórmula general de una regresión logística se puede observar en 5. El valor t son las probabilidades logísticas y β son los parámetros de cada uno de los factores. En caso de ser variables categóricas, se incluye un parámetro por cada una de los valores posibles de cada uno estas variables [31].

$$t = \log_b \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_M x_M$$

Figura 4: Fórmula generalizada de regresión logística

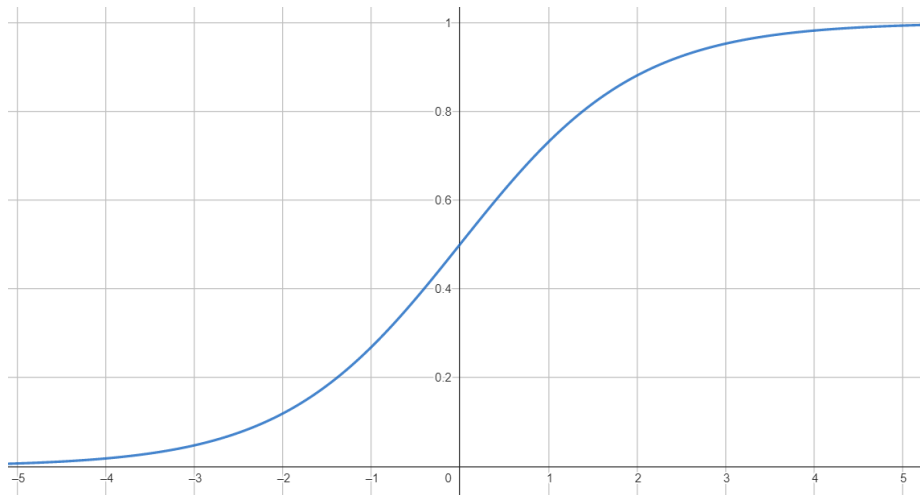


Figura 5: Función logística

La regresión logística puede utilizarse también para obtener la razón de probabilidades, lo cual puede utilizarse para explicar la asociación entre cambios en los factores observados y cambios en la variable dependiente. Esto muestra el cambio de probabilidad que ocurre al cambiar uno de los factores, mientras todos los demás factores se mantienen igual. Los valores obtenidos con este análisis permiten entender cómo diferentes factores influyen en las probabilidades de la variable dependiente [31].

Este modelo es utilizado en múltiples campos, incluyendo en campos médicos. El programa para la mejora de calidad de trauma (TQIP, por sus siglas en inglés) utiliza una regresión logística de 18 variables para estimar el riesgo de mortalidad para los pacientes de trauma. Los resultados proveen los porcentajes de riesgo para cada uno de los factores de este modelo [31].

5.3.6. Chi cuadrado

La prueba de chi cuadrado es una prueba estadística utilizada para determinar si dos variables categóricas influyen el estadístico de prueba. La prueba chi cuadrado de Pearson se utiliza para determinar si existe una diferencia estadísticamente significativa entre la frecuencia esperada y observadas de variables categóricas en una tabla de contingencia [25].

Esto se logra al separar todas las observaciones en clases mutuamente exclusivas. Si las variables no están asociadas, se espera que la frecuencia esté uniformemente distribuida y el estadístico de prueba sigue una distribución de chi cuadrado. Al existir asociación entre ambas variables, el estadístico de prueba aumenta de manera diferente y a partir de esta información es que se obtiene información sobre la relación que existe entre ambas variables [25].

La fórmula para calcular este estadístico de prueba se puede observar en 7. Esta fórmula describe n observaciones distribuidas en k categorías independientes. Por cada categoría se calcula la diferencia entre la cantidad de valores esperados y los valores observados [25].

$$X^2 = \sum_{i=1}^k \frac{(x_i - m_i)^2}{m_i} = \sum_{i=1}^k \frac{x_i^2}{m_i} - n$$

Figura 6: Fórmula generalizada para chi cuadrado

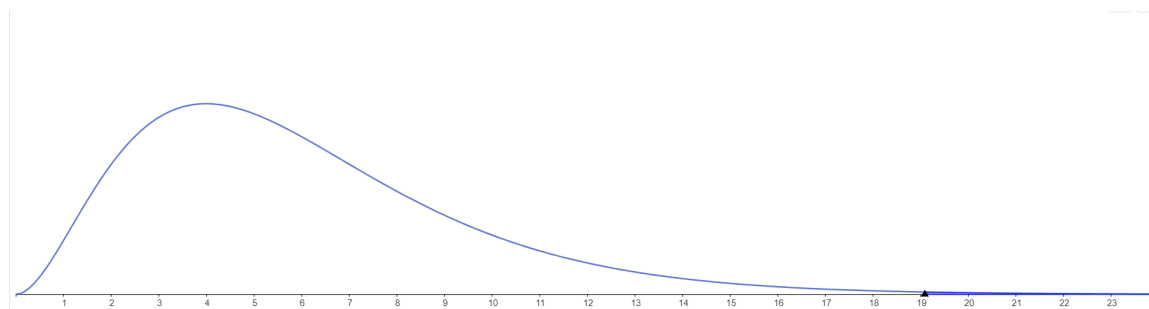


Figura 7: Distribución de chi cuadrado

5.4. Situación actual en Guatemala

Guatemala es un país joven. Las estimaciones realizadas por el Instituto Nacional de Estadística (INE) indican que la población de menores de 5 años era alrededor 1.870 millones en 2022. En 2015 el valor era similar, rondando los 1.868 millones [32]. Además, la población adolescente era de 1.764 millones en 2015 y 1.747 millones en 2022 [32]. Esto muestra la importancia que la salud de este sector tiene para el país.

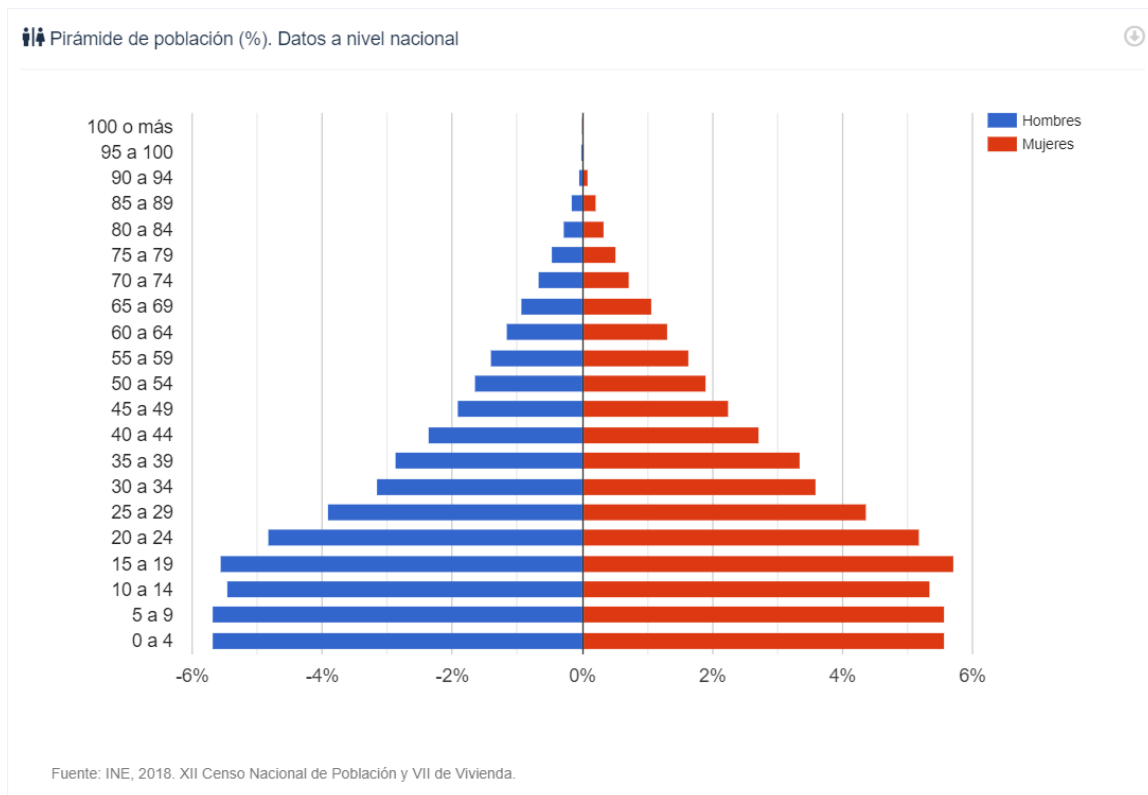


Figura 8: Pirámide poblacional de Guatemala

5.4.1. Desnutrición infantil

Se definen tres tipos de desnutrición, la desnutrición aguda, global y crónica [33]. La desnutrición crónica indica un retraso de la talla esperada con respecto a la edad, que afecta el desarrollo y crecimiento físico y cognitivo de una persona [10]. Este déficit representa el impacto acumulativo de múltiples eventos, principalmente en los primeros 1000 días desde la concepción. [10] Sus causas son multidimensionales, por lo que es complicado comprender el fenómeno y encontrar posibles soluciones. La desnutrición crónica puede tener repercusiones duraderas para las personas a quien afecta o afectó anteriormente [34].

Por otro lado, la desnutrición aguda se identifica cuando el peso de un niño está por debajo de la media para su talla. Esto puede ser causado por no tener suficientes alimentos o por contraer alguna enfermedad o infección. En algunos casos, el peso puede llegar a ser tan bajo que el niño puede estar en riesgo de morir [34].

Por último, la desnutrición global se identifica cuando el peso de un niño está por debajo de la media para edad. Principalmente, se usa para monitorear el crecimiento de los niños. Es un indicador de que se han tenido dificultades en la alimentación en el pasado y en el presente [35].

El estado nutricional muchas veces se mide en términos de antropometría [10]. Esto conlleva comparar el crecimiento de un niño contra estándares internacionales, los cuales han sido creados para mostrar el crecimiento infantil bajo condiciones óptimas, desde el

nacimiento hasta los 5 años. La comparación se realiza encontrando cuántas desviaciones estándar se aleja un niño del estándar utilizado[10].

El estándar que se utiliza en los resultados de la ENSMI y de LBGCNN es el estándar utilizado por la Organización Mundial de Salud, publicado en 2006 y desarrollado utilizando datos de más de 8000 niños en Brasil, Ghana, India, Noruega, Omán y Estados Unidos [36]. Ambos informes definen desnutrición crónica moderada como todos los niños por debajo de -2 desviaciones estándar de la media y desnutrición crónica severa por debajo de -3 desviaciones estándar de la media [10].

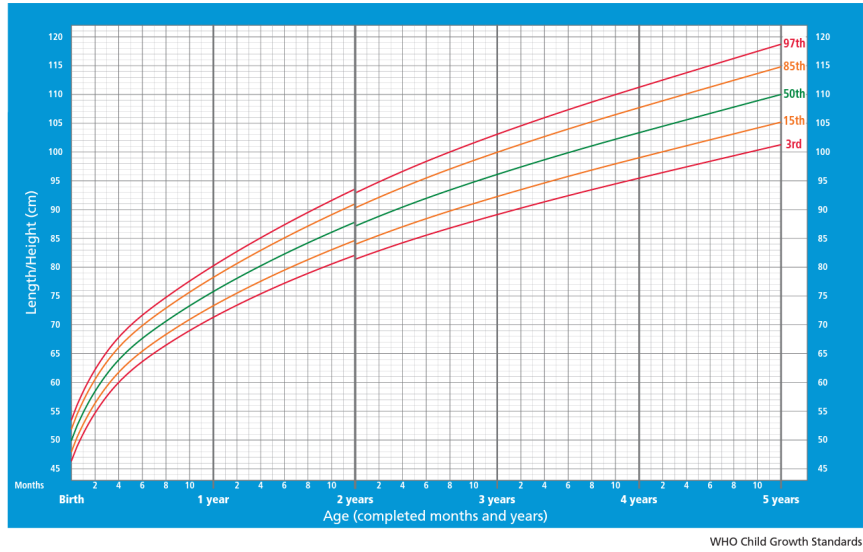


Figura 9: Ejemplo de estándares de talla para edad en niños

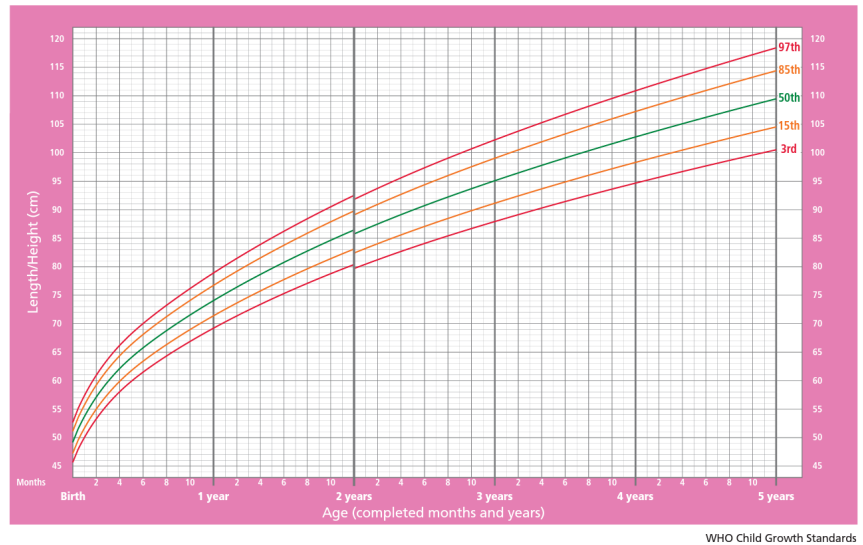


Figura 10: Ejemplo de estándares de talla para edad en niñas

Guatemala es el sexto país del mundo con peores índices de malnutrición infantil y el primero de América Latina [37]. Según la ENSMI 2014-2015 [7], 46.5% de los niños sufren de desnutrición crónica y se estima que la mitad de las enfermedades que causan

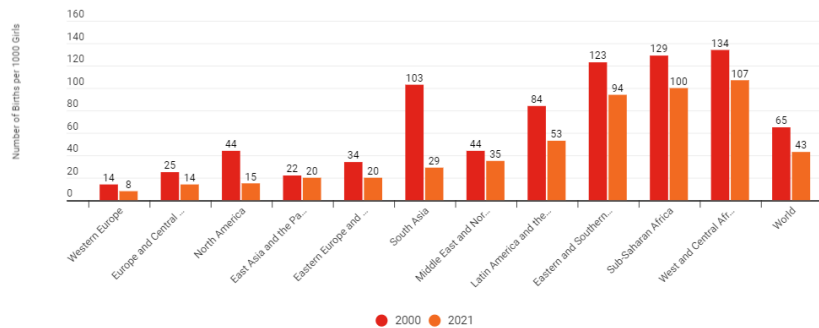
mortalidad infantil están relacionadas a algún tipo de malnutrición. Este fenómeno no se observa de manera homogénea en todo el país [38]. En algunas regiones donde la mayoría de la población es de origen indígena, los índices alcanzan 80 % [37]. Se estima que esto representa una pérdida de 3 mil millones de dólares anuales para el país [38].

5.4.2. Embarazos en mujeres adolescentes

Se estima que cada día 176 niñas y adolescentes son madres en Guatemala [39]. Esto tiene un impacto importante en la vida de las madres y de los niños. Se estima que únicamente 2 % de las mujeres que fueron madres durante la adolescencia logran un título universitario en comparación a 9 % que fueron madres entre los 20 y 29 años [39]. Además, se observa una brecha salarial del 30 % entre las mujeres que fueron madres adolescentes y las mujeres que fueron madres en la edad adulta [39].

Aunque el embarazo adolescente ha disminuido a nivel mundial durante los últimos años, este proceso ha sido más lento de lo esperado. En muchos casos, el descenso ha sido únicamente algunos puntos por década. Además, en algunas regiones del mundo, los patrones observados son muy parecidos a los observados hace más de medio siglo [40]. Estados Unidos tiene la tasa de embarazos adolescentes más alta en el mundo occidental industrializado, 34 nacimientos vivos por cada 1000 mujeres entre 15 y 19 años. Esto es el alrededor del doble que la tasa más baja en Europa occidental y casi 10 veces más que la tasa más baja [10].

El embarazo adolescente puede tener resultados negativos tanto para la madre como para el hijo [41]. Se ha encontrado una asociación entre embarazo adolescente y bajo peso al nacer, embarazo prematuro, muertes neonatales y muerte infantil antes de un año de vida. Estos efectos ocurren con mayor frecuencia en mujeres más jóvenes y con menos posibilidades económicas [10].



Source: United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022

Figura 11: Cambio en embarazo adolescente entre 2000 y 2021

Los factores predictores y que causan el embarazo adolescente son múltiples. El Fondo de Población de las Naciones Unidas (UNFPA, por sus siglas en inglés) creó un marco para comprender los factores inmediatos y motivadores del proceso de embarazo adolescente. Además, explica que no es suficiente conocer si hay o no embarazos adolescentes, debe tenerse también el conocimiento sobre cuándo y cuántos embarazos ocurren [40]. Tomando

esto como base, se identifican los siguientes tres procesos de fertilidad fundamentales e interconectados:

- El tiempo del primer embarazo
- El espacio entre nacimientos adolescentes
- La cantidad o número total de nacimientos para cada madre adolescente

Cuando las mujeres tienen un primer embarazo temprano, posiblemente tengan más embarazos durante su vida reproductiva. Esto, junto a un espacio entre nacimientos más reducido, lleva a una mayor cantidad de embarazos para una madre adolescente. Estos tres procesos están influenciados por factores como la menarquia y el inicio de la vida sexual, la frecuencia de la actividad sexual, los patrones de matrimonio y formación de uniones y el uso de anticonceptivos. Estos factores influyen tanto en el primer nacimiento como en los posteriores. En los países en vía de desarrollo, el inicio de las relaciones sexuales aumenta la probabilidad de formar una unión más temprano, lo cual aumenta la probabilidad de embarazos adolescentes. Una unión temprana aumenta la frecuencia de actividad sexual que aumenta la probabilidad de embarazos posteriores [40].

5.5. Observatorio de Salud Sexual y Reproductiva (OSAR)

El OSAR fue creado por el Congreso de Guatemala, con el fin de monitorear la implementación del marco político y legal. Además, vigila y audita el cumplimiento del Estado en convenios y compromisos internacionales y aboga por cambios políticos para disminuir brechas entre poblaciones. Fue creado en 2008, a través de un convenio entre el Congreso de la República y Organizaciones de la Sociedad. Actualmente, se mueve con 4 líneas de trabajo [42]:

- Uso estratégico de la información
- Incidencia política y auditoría social
- Investigación
- Participación ciudadana



Figura 12: Logo de OSAR

OSAR se involucra en temas de desnutrición infantil a través de un acuerdo firmado con SESAN en 2021, el cual establece un acuerdo de trabajo en conjunto entre ambas organizaciones. Un componente fuerte de la labor en conjunto es el tema de embarazo en adolescentes y su relación con el estado nutricional del niño al nacer. Además, buscan indagar en los beneficios del espaciamiento intergenésico para aumentar la sobrevivencia de niño y la madre. Como parte de este acuerdo, se crearon técnicos de enlace a nivel central, para coordinar y fomentar la comunicación y colaboración entre ambas partes. Bajo esta estipulación, se creó el Comité Técnico de seguimiento a la atención de embarazadas adolescentes, el cual se reúne periódicamente para dar seguimiento a este acuerdo.

5.5.1. Métodos anticonceptivos

Dentro de las encuestas realizadas por DHS, se definen cuatro tipos de métodos anticonceptivos, métodos modernos, tradicionales, folclóricos y ningún método. Los métodos tradicionales incluyen el ritmo y la abstinencia. Los métodos modernos incluyen métodos como inyecciones, esterilización y dispositivo intrauterino. Además, se define métodos folclóricos como “cualquier otro método tradicional”[43].

5.6. Mediciones de desnutrición crónica en Guatemala

En Guatemala, las dos fuentes más recientes para datos de desnutrición crónica en menores de 5 años son la Encuesta Nacional de Salud Materno Infantil (ENSMI) 2014-2015 y el Informe de Línea de Base de la Gran Cruzada Nacional por la Nutrición (LBGCNN) 2021-2022. Ambas son encuestas realizadas con representatividad nacional. Sin embargo, existen algunas diferencias en las maneras en las que se presentan los datos de desnutrición crónica. La diferencia principal entre ambas encuestas son los grupos de edad que se utilizan al momento de presentar resultados agrupados. Esta diferencia puede dificultar la comparación directa de los resultados.

5.6.1. Factores demográficos

La ENSMI 2014-2015 utiliza múltiples características demográficas para presentar y distribuir los resultados obtenidos. Estos componentes se refieren a la composición de la población por sexo y edad, el ambiente en donde viven, nivel de educación y nivel económico. Estas características son utilizadas para apoyar y complementar la información presentada a través de todo el informe [7].

El máximo nivel de educación alcanzado y el área de residencia se recolectaron como parte de la encuesta realizada. Además, se recolecta información sobre alfabetismo y múltiples características del hogar, como el acceso a agua potable, combustible para la cocina y número de bombillas en el hogar. Todas estas características se utilizan para calcular el quintil de riqueza. Este se calcula utilizando un Análisis de Componentes Principales (PCA, por sus siglas en inglés), asignando una calificación de riqueza a cada hogar, valor que se utiliza para separar todos los hogares en 5 categorías con la misma cantidad de hogares [7].

5.6.2. Encuesta Nacional de Salud Materno Infantil (ENSMI) 2014-2015

La Encuesta Nacional de Salud Materno Infantil fue realizada entre octubre de 2014 y julio de 2015 por el Ministerio de Salud Pública y Asistencia Social (MSPAS) a través de la Unidad de Planificación Estratégica (UPE), el Instituto Nacional de Estadística (INE) y la Secretaría de Planificación y Programación de la Presidencia (SEGEPLAN). Los recursos financieros fueron manejados por el Instituto de Nutrición de Centroamérica y Panamá, el Fondo de Naciones Unidas para la Infancia (UNICEF) y del MSPAS. Además, se recibió financiamiento adicional del programa de Encuestas de Demografía y Salud (Programa DHS), quien provee asistencia para encuestas en países en vías de desarrollo [7].

Esta encuesta se realiza alrededor de cada 6 años en Guatemala. La versión más reciente es la sexta edición de la ENSMI, realizada entre 2014 y 2015, con los resultados publicados en 2017. La versión anterior se realizó entre 2008 y 2009, con los resultados publicados en el 2010. La séptima edición de esta encuesta no se ha empezado a realizar, en el año 2022 se publicó el acuerdo 12-2022, que dispone la realización esta encuesta. Sin embargo, al momento de realizar este informe, no se ha iniciado la recolección de datos [44]. Hasta que se realice la séptima edición de esta encuesta y se publiquen los datos, la sexta edición es la más reciente de esta encuesta.

La ENSMI se realiza con el propósito de conocer la situación demográfica de salud infantil, sexual y reproductiva a nivel nacional. Los resultados son presentados en un informe final, el cual es presentado por el Ministerio de Salud Pública y Asistencia Social. [7].

La encuesta incluye preguntas acerca de la fecundidad, inicio de la vida sexual, primera unión, cantidad de hijos, desnutrición, entre otras. Recolectó información general de todos los miembros en cada hogar y se realizaron preguntas adicionales para cada niño menor a 5 años. La encuesta únicamente se realiza a mujeres entre 15 y 49 años. Los hogares donde se realiza las encuestas son elegidos de manera aleatoria y se aplica un formulario con una duración aproximada de 30 minutos [7].

Los conjuntos de datos son publicados en el sitio de las encuestas demográficas y de salud (DHS, por sus siglas en inglés) y únicamente pueden ser accedidos solicitando los permisos correspondientes. Debido a esto, se debe solicitar permiso explícito para utilizar los datos para proyectos investigación. Además, se prohíbe compartir los datos crudos con personas externas [45].

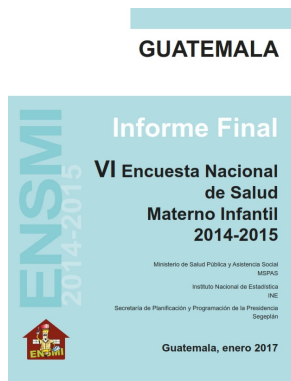


Figura 13: Portada ENSMI 2014-2015

5.6.3. Línea de Base de la Gran Cruzada Nacional por la Nutrición (LBGCNN)

La Línea de Base de la Gran Cruzada Nacional por la Nutrición fue realizada entre los meses de mayo 2021 a marzo 2022, por la Secretaría de Seguridad Alimentaria y Nutricional (SESAN) en colaboración con Ministerio de Salud Pública y Asistencia Social de Guatemala (MSPAS). Esta iniciativa consistió en una encuesta realizada a nivel nacional en temas de salud, alimentación saludable y saneamiento e higiene. Además, se buscó focalizar los esfuerzos en 114 municipios en los 10 departamentos con mayores índices de pobreza extrema, desnutrición crónica e inseguridad alimentaria [9].

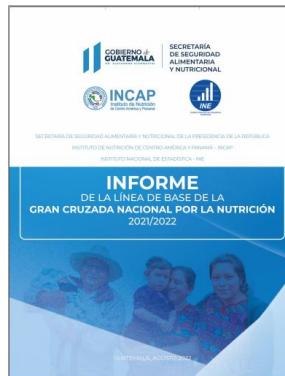


Figura 14: Portada LBGCNN 2022

Se utilizó la metodología CRIPS-DM como marco para el procesamiento de los datos a realizarse. Esta metodología sigue un proceso iterativo para procesar los datos, empezando con obtener y explorar los datos, para luego ser procesados y validados. Estas fases sirvieron como bases para los pasos que se estarían realizando para cumplir con los objetivos establecidos. La fase de entendimiento del negocio se realizó durante los acercamientos iniciales con OSAR y durante la investigación realizada como parte del marco teórico. La fase 2 se realizó con la obtención de los datos de la página de DHS y la exploración de los datos.

La tercera fase se realizó al cambiar los datos al tipo necesario para su procesamiento. Las siguientes dos fases se realizaron al crear y validar los modelos datos creados. Por último, al publicar el despliegue se realizó al publicar el tablero en una página web. Durante todas las fases se realizaron validaciones con OSAR, para asegurar recibir retroalimentación constante.

6.1. Obtención de datos

Los datos se obtuvieron a través de la página de las encuestas de salud y demográficas (DHS, por sus siglas en inglés) de United States Agency for International Development (USAID). Para acceder a ellos, fue necesario crear una cuenta en este sitio y llenar una solicitud sobre el tipo de trabajo a realizar. Una vez se aprobó esta solicitud, se recibió acceso a los resultados originales de la encuesta. Estas bases de datos venían en 4 formatos diferentes, .dta, .dat, .sas7bdat y .sav. Los resultados están divididos en 6 conjuntos de datos, con los siguientes nombres y contenidos:

- Births Recode: Registros de los nacimientos de mujeres encuestadas
- Household Recode: Registros de los hogares

- Individual Recode: Registros de cada mujer que respondió la entrevista
- Children’s Recode: Registros de todos los niños de hasta 5 años
- Men’s Recode: Registros de parejas de mujeres encuestadas
- Household Member Recode: Registro de otros miembros de los hogares de mujeres

Inicialmente, se buscó realizar con el tablero en la plataforma Superset. Explorando esta opción, se realizó una instalación local de la herramienta utilizando contenedores y Windows Subsystem for Linux .

Al intentar cargar los datos a esta herramienta, se observó que los formatos originales no podían ser cargados, por lo que fue necesario cambiar los datos a otro formato. Se decidió cambiar el formato hacia un archivo .csv, el cual puede ser leído directamente por R. La herramienta elegida para realizar esta transformación fue Python con la librería Pandas, ya que tiene funcionalidades implementadas para leer y exportar archivos en múltiples formatos. Todas las bases de datos con el formato .sav se cargaron en Python utilizando Pandas. Esta herramienta crea un *dataframe* el cual se puede utilizar para transformaciones de datos dentro de Python. Estos objetos se exportaron directamente al formato .csv utilizando una función proveída por Pandas, creando un archivo .csv basado en este *dataframe*.

Sin embargo, los datos no funcionaban correctamente, debido a las grandes dimensiones. Se intentó cargar los datos hacia la base de datos PostgreSQL, para acceder a ellos de esta manera. Este manejador de base de datos cuenta con un límite definido de filas, el cual era sobrepasado por la cantidad de resultados proveída. Al notar que esta herramienta no podía ser utilizada para estos fines, se decidió utilizar R con la librería Shiny. Ya que ya se contaban con los datos en .csv, se decidió utilizarlos para cargar los datos a esta herramienta.

Para comprender el significado de cada columna de estos registros, se utilizó un documento proveído por DHS, donde se detalla la información de cada columna registrada y la relación que guardan con otros datos en los diferentes conjuntos de datos. Con este documento se identificaron las columnas a utilizar al inicio del análisis y se consultó múltiples veces durante el análisis realizado para obtener la información necesaria.

6.2. Análisis inicial

Los archivos con el formato .csv fueron cargados en R, con lo que se pudo empezar a trabajar sobre ellos. Se inició cargando las librerías necesarias para su procesamiento, las cuales son las siguientes:

- shiny, para construir el tablero
- ggplot, para visualización de datos
- dplyr, para procesamiento de datos
- tidyr, para el procesamiento de datos

Luego, se empezó a explorar los datos para entender sus características. Para lograrlo, se realizaron tareas de estadística descriptiva, buscando relaciones y características superficiales entre las diferentes columnas presentes en el conjunto de datos. Se inició conociendo la cantidad de columnas y filas presentes en las bases de datos importadas, con lo cual se obtuvo los siguientes datos:

Conjunto de datos	Filas	Columnas
Births Recode	55398	1236
Household Recode	21383	2598
Individual Recode	25914	4780
Children´s Recode	12440	1236
Men´s Recode	11145	833
Household Member Recode	102510	310

Cuadro 1: Dimensiones de conjuntos de datos obtenidos

Se empezó agrupando los conjuntos de datos por algunas características demográficas, como quintil de riqueza, área de residencia y máximo nivel de educación alcanzado, para comprender su comportamiento. Se eligieron estas características ya que también son utilizadas a través del informe de la ENSMI 2014-2015.

Todo este análisis se realizó utilizando Python y Pandas, dado la facilidad que provee para estos análisis sencillos. Estos resultados fueron utilizados para guiar las tareas a realizarse en el futuro y para comprender el comportamiento general de los datos. Ya que los resultados de este análisis no serían enseñados al cliente, no se modificó el nombre de las columnas utilizadas. En anexos se incluye el significado de cada una de las columnas presentadas.

Se inició identificando el promedio del periodo entre nacimientos por quintil de riqueza. Los resultados son los siguientes:

B11	
V190	
Middle	3.050538
Poorer	2.916096
Poorest	2.714926
Richer	3.493890
Richest	3.864447

Figura 15: Educación v Nacimientos totales

Luego, se identificó el promedio número de hijos por quintil de riqueza, los resultados son los siguientes:

V201	
V190	
Middle	2.181599
Poorer	2.492296
Poorest	2.907743
Richer	1.775430
Richest	1.486653

Figura 16: Educación v Nacimientos totales

También, se identificó el promedio de edad para el primer nacimiento, según el máximo nivel de educación alcanzado:

V212	
V106	
Higher	23.544931
No education	18.989987
Primary	19.265840
Secondary	20.680490

Figura 17: Educación v Nacimientos totales

También, se obtuvo el promedio de hijos que tienen las madres adolescentes, obteniendo un promedio de:

4.26

Este valor es superior al promedio de hijos en madres no adolescentes.

Se identificó la distribución de las madres adolescentes por quintil de riqueza, identificando su distribución por el máximo nivel de educación alcanzado. Los resultados se resumen en la siguiente tabla:

Conjunto de datos	Porcentaje de madres <15	Porcentaje de madres >=15 y <20	Porcentaje de madres >= 20 y <25	Porcentaje de madres >= 25
Sin educación	29.56	21.34	30.99	33.02
Educación primaria	59.69	57.14	46.17	41.01
Educación secundaria	10.17	21.34	15.65	15.48
Educación terciaria	0.6	1.57	7.20	10.49

Cuadro 2: Porcentaje de madres según máximo nivel de educación alcanzado

Con este análisis básico realizado, se empezó a trabajar directamente en R, para generar gráficas iniciales, las cuales permitirían empezar a recibir retroalimentación por parte de OSAR y otras partes interesadas. En este punto se seguían utilizando los nombres de las columnas como se encontraban en el conjunto original de datos, ya que se valoró más el

comportamiento de los datos que la estética de las gráficas realizadas. Para su interpretación, se empezó a crear un diccionario de columnas. Además, este fue utilizado para todo el trabajo de desarrollo de gráficas. Las relaciones exploradas para este punto son las siguientes:

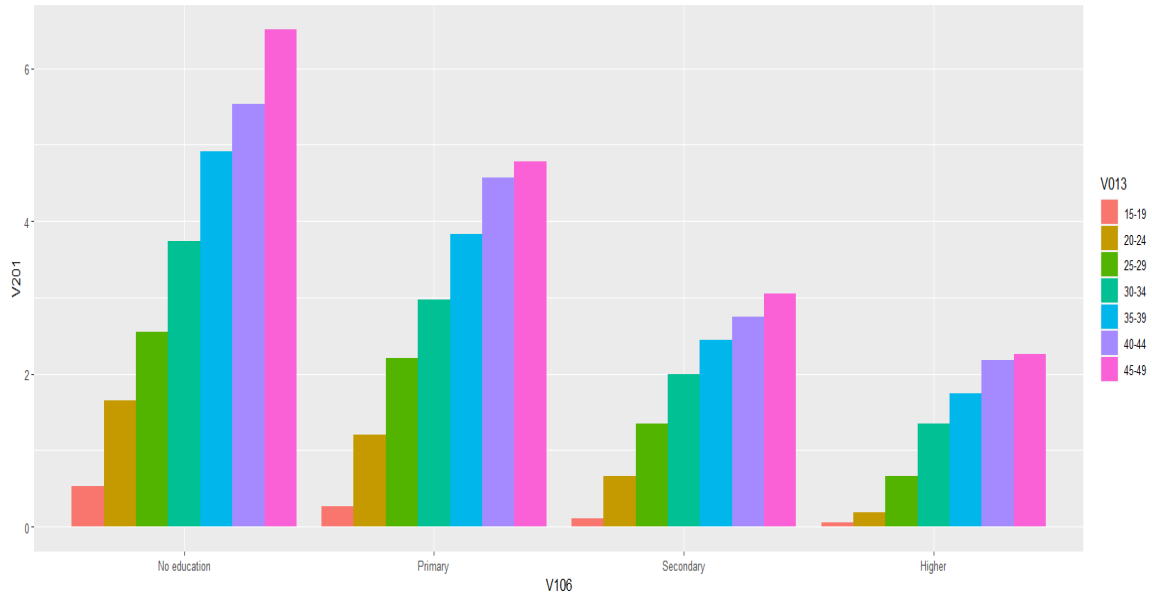


Figura 18: Educación y nacimientos totales

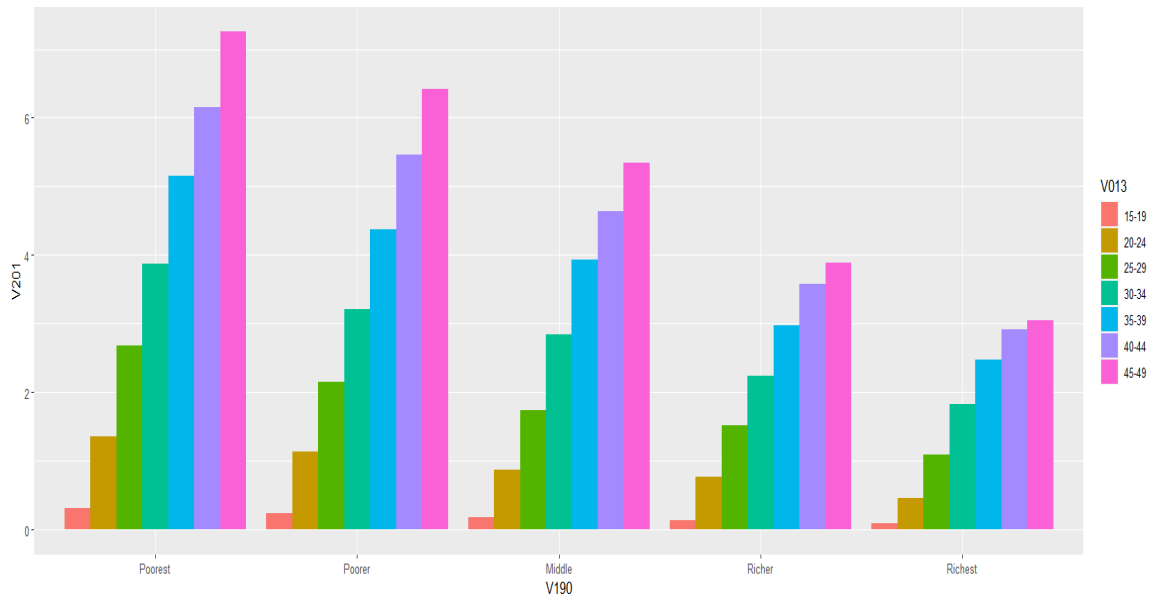


Figura 19: Quintil de riqueza y nacimientos totales

También, se empezó a visualizar la relación que existe entre la talla por edad y la cantidad de hijos que se han tenido:

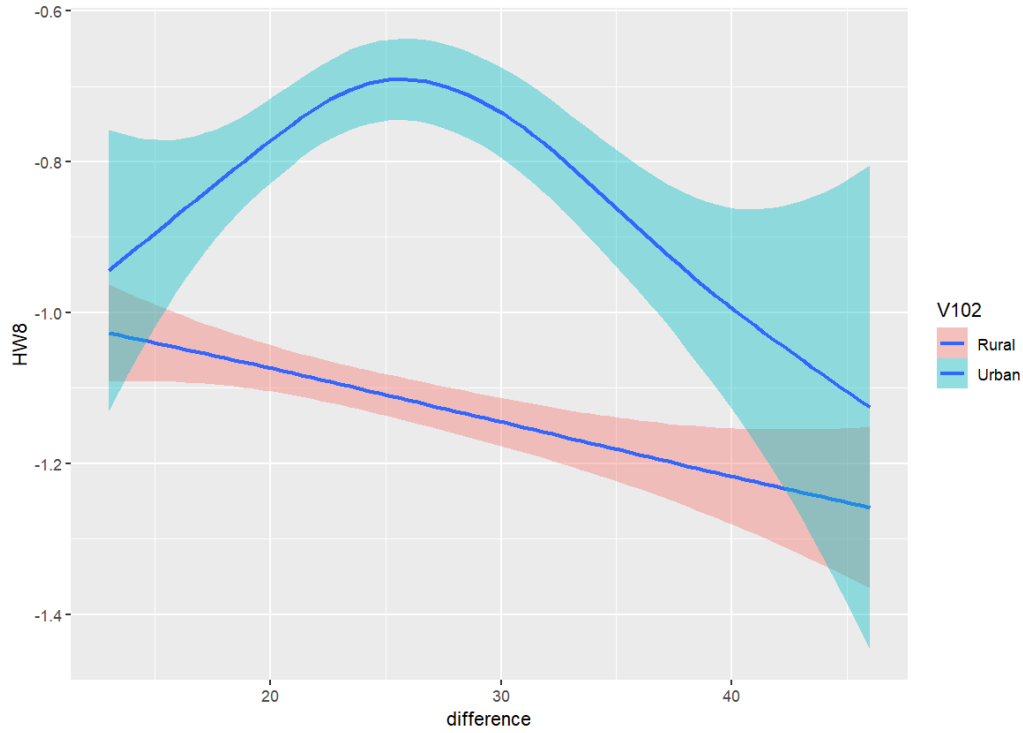


Figura 20: Edad de las madres y cuántas desviaciones estándar se alejan de la media

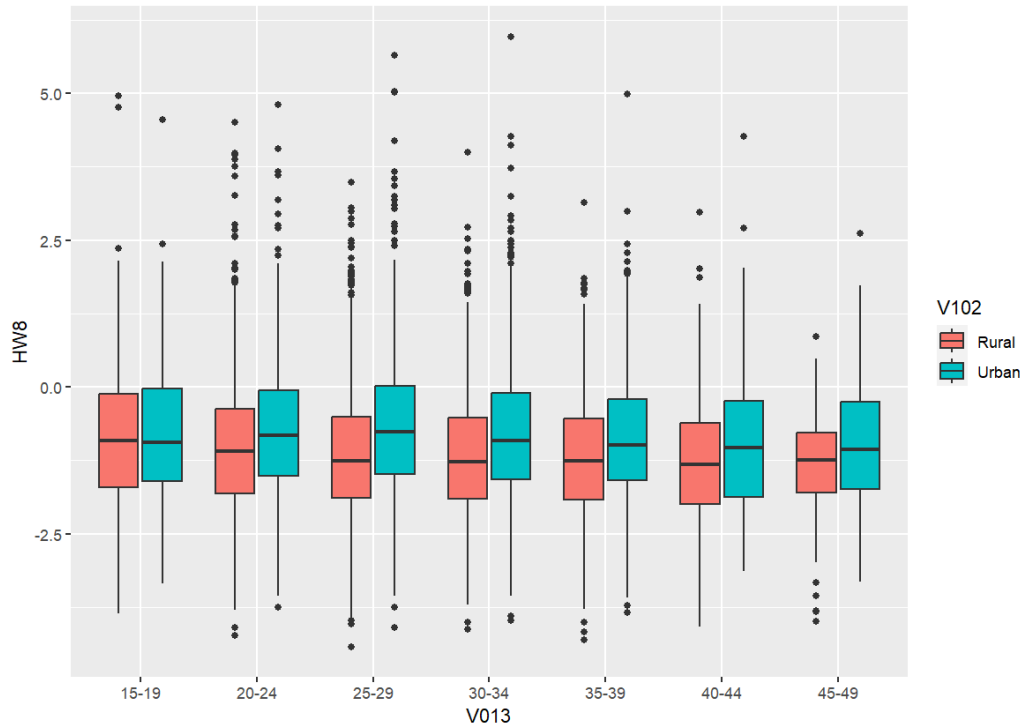


Figura 21: Edad de las madres por grupo de edad y cuántas desviaciones estándar de la media

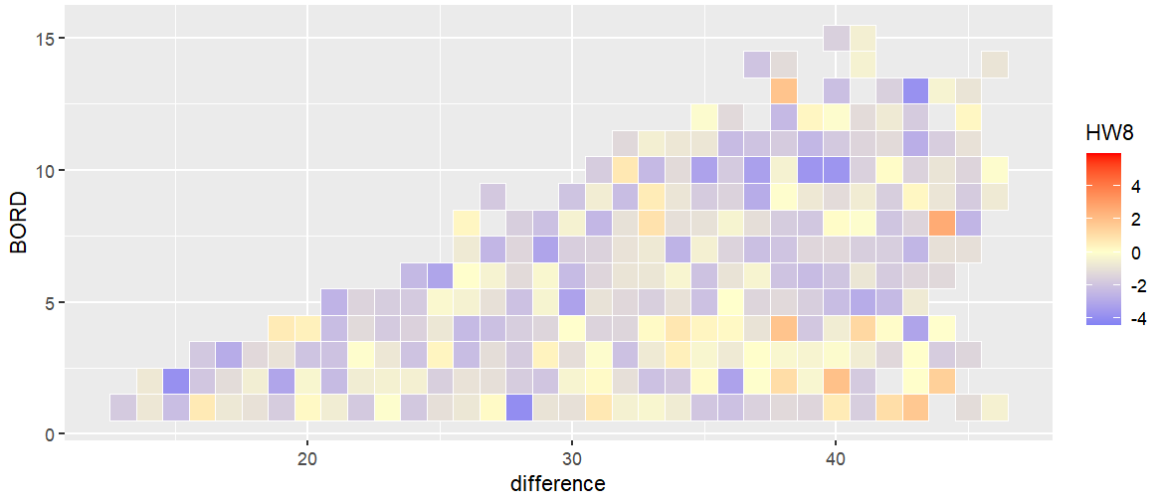


Figura 22: Edad de las madres y número de hijo, mostrando cuántas desviaciones estándar se encuentran de la media

6.3. Sesiones de retroalimentación

Para manejar la interacción con OSAR y desarrollar el cliente final, se decidió utilizar Design Thinking. Esto permitiría trabajar e interactuar de manera iterativa con OSAR. Además, su enfoque en el problema aseguró que la solución creada fuera útil para OSAR y tuviese un uso verdadero después de su desarrollo.

De esta manera, se recibiría retroalimentación constante acerca del avance en el tablero y los cambios se implementan rápidamente, para minimizar la cantidad de cambios implementados. Para estas reuniones, se estaría presentando los resultados a OSAR y en reuniones posteriores se estarían presentando a una audiencia más amplia, conforme el tablero fuese madurando.

Se empezó empatizando con OSAR a través del acercamiento inicial para entender la problemática que se estaba buscando identificar. Con esta información se empezó a definir el alcance de las tareas a realizar y cómo esto impactaría los datos a utilizar. También, se encontró posibles límites de los resultados a generar, dados los datos disponibles.

Conociendo las posibilidades y los límites, se empezó a idear los resultados a generar a través del análisis secundario. Muchas de estas ideas fueron graficadas y validadas con el cliente, para guiar el desarrollo de estas soluciones. Estos prototipos fueron evolucionando mientras se encontraba información adicional y se modificaba el producto final.

6.3.1. Primera sesión de retroalimentación

Con estos primeros resultados, se organizó una primera sesión con la secretaria técnica de OSAR, para presentar los hallazgos iniciales. Durante la reunión se mostraron las gráficas realizadas hasta este momento, se recibió retroalimentación sobre los datos presentados

y se inició la coordinación para sesiones adicionales. Durante la reunión, se recibió retroalimentación positiva sobre las gráficas realizadas, pero se recibieron sugerencias para crear gráficas más simples, para ser utilizadas por las diferentes audiencias de OSAR. También, se recomendó revisar un tablero colocado en la página de OSAR como un ejemplo del tipo de información que se busca recolectar.

También, se identificaron posibles audiencias. Entre estas se mencionó a la Secretaría de Seguridad Alimentaria y Nutricional (SESAN), Ministerio de Salud Pública y Asistencia Social (MSPAS) y miembros departamentales de OSAR para validar la usabilidad y el funcionamiento del tablero creado. Estas sesiones se realizaron en marzo, después de realizar los cambios sugeridos al tablero e incluir resultados adicionales.

6.3.2. Cambios implementados

Después de tener esta sesión de retroalimentación, se implementaron cambios a las gráficas creadas y principalmente a la manera en que se presentaban los resultados. Se empezó limitando la cantidad de variables que se utilizan en cada gráfica, para simplificar el mensaje que se estará enviando. Además, se cambiaron las etiquetas en las gráficas y los nombres de los ejes, para poder ser entendidos de mejor manera y mostrar más claramente la información obtenida. Estos cambios fueron utilizados en una segunda reunión donde se presentaron los resultados a una nueva audiencia.

6.3.3. Segunda sesión de retroalimentación

En esta reunión, realizada de manera virtual, participaron miembros de OSAR, SESAN y el Programa Nacional de Salud Reproductiva (PNSR) del MSPAS. Se presentó el motivo del trabajo, los resultados obtenidos hasta este momento y el tablero que se estaba construyendo. En esta reunión, nuevamente, se recibió la recomendación de utilizar lenguaje más simple al momento de presentar los hallazgos del análisis secundario de la ENSMI 2014-2015 que se estaba realizando. Además, se recibió la sugerencia de utilizar datos de la LBGCNN 2021-2022 y compararlo con los datos obtenidos de la ENSMI 2014-2015.

6.3.4. Cambios implementados

A partir de esta sesión de retroalimentación, se realizaron cambios a las gráficas y se realizó una transformación adicional para compararlos con los resultados de la LBGCNN 2021-2022. Para este análisis fue necesario modificar los datos obtenidos y clasificarlos en los mismos rangos que los resultados presentados en el reporte de la LBGCNN 2021-2022. La diferencia se debe a que ambos reportes utilizan diferentes rangos de edades para presentar los resultados sobre la desnutrición crónica, lo cual imposibilita compararlos únicamente con acceso a los informes finales.

Se empezó leyendo la sección metodológica de este nuevo reporte y tras validar de que la comparación entre ambas encuestas era posible, se obtuvieron los datos sobre la desnutrición crónica con los rangos apropiados. Estos resultados sí pueden ser comparados directamente



Figura 23: Segunda sesión de retroalimentación

con los valores presentados en el informe final de la LBGCNN, por lo cual dan información clave sobre la evolución de la desnutrición en Guatemala durante los últimos 8 años.

Para presentar los datos de una manera más clara, se trabajó en utilizar frases más simples para acompañar a las gráficas relacionadas. También, siguiendo las recomendaciones recibidas, se eliminó el uso de términos particularmente estadísticos, como "Desviación estándar", lo cual se identificó como potencialmente demasiado complejo para el tipo de uso que se le dará a esta información. Se trabajó, en acuerdo con OSAR, para obtener los mensajes clave de cada gráfica y se consolidaron las gráficas y sus mensajes claves en una única presentación, la cual fue validada nuevamente con OSAR. Además, a esta presentación se incluyó los resultados obtenidos de análisis estadísticos que mostraban perspectivas relevantes sobre estos temas. Estas gráficas son las encontradas en la sección Resultados iniciales.

6.3.5. Tercera sesión de retroalimentación

La nueva presentación fue presentada de manera presencial a miembros de SESAN, PNSR y OSAR, como parte de una reunión del comité técnico de seguimiento a la atención de embarazadas adolescentes. La agenda de esta reunión se adjunta en los anexos. Tras presentar los avances hasta este punto, se habló sobre la posibilidad de utilizar estos datos como parte de una iniciativa para incluir la prevención de los embarazos adolescentes en la agenda de la Gran Cruzada Nacional por la Nutrición (GCNN). Además, se planeó presentar los resultados obtenidos en una reunión con la Vicepresidencia de la República, para seguir socializando los resultados y causar incidencia en los tomadores de decisiones a nivel nacional a través de ellos.



Figura 24: Presentación de resultados ante comité

6.3.6. Cambios implementados

Después de realizar esta reunión, se decidió continuar con las gráficas presentadas en esta reunión. Para este punto, el enfoque se centró en cambios estéticos para el tablero y las gráficas creadas. Se modificó el nombre de las categorías a español y se agregó nombre a los ejes, agrupaciones y títulos a las gráficas. Para los títulos se utilizó las perspectivas presentadas en esta reunión.

6.4. Relaciones adicionales

Tras realizar las primeras gráficas y entender las relaciones existentes entre algunas de las características incluidas en este conjunto de datos, se empezó a trabajar en identificar relaciones adicionales entre el conjunto de datos.

Se inició realizando un proceso de reducción de dimensionalidades, con el objetivo de simplificar la gran cantidad de variables existentes y así poder encontrar patrones ocultos. Se utilizó la técnica Factor Analysis of Mixed Data (FAMD), la cual permite utilizar variables cualitativas y cuantitativas. Se utilizaron todas las variables identificadas en la recolección de literatura como punto inicial. Las variables utilizadas son las siguientes:

- Origen de nacimiento
- Edad de primer nacimiento
- Peso de la madre
- Altura de la madre
- Quintil de riqueza
- Edad de la primera unión
- Edad de primer encuentro sexual
- Sexo del niño
- Intervalo al nacimiento anterior

Con estos factores se logró obtener la siguiente varianza con la siguiente cantidad de factores:

Cantidad de dimensiones	Varianza mantenida (%)
1	19.86
2	32.78
3	42.34
4	49.19
5	55.90
6	62.31
7	68.45
8	74.40
9	80.07

También, se realizó una regresión logística para identificar los factores que más influyen en la desnutrición infantil. Ya que la regresión logística sirve para valores categóricos, se creó una nueva variable para mostrar si un niño sufre de desnutrición o no, usando la descripción utilizada por OMS y ENSMI. Esta variable fue utilizada para entrenar el modelo logístico, como la variable dependiente, utilizando variables propias de la madre y el nacimiento del hijo como variables independientes. Con este método se identificaron la influencia que estas categorías tienen sobre si un niño sufre de desnutrición crónica o no. Se inició haciendo una separación de datos, utilizando 80% de los datos para entreno y 20% para probar. Además, se realizó una validación *10-fold*, obteniendo medidas *accuracy*, *recall* y *precision* para validar el funcionamiento del modelo. Para cada uno de estas ejecuciones, se obtuvo la matriz de confusión para evaluar el desempeño del modelo.

Este análisis se complementó utilizando la prueba de chi cuadrado, utilizando únicamente variables categóricas. La elección de estas categorías se realizó basado en los resultados obtenidos en la fase anterior, haciendo únicamente los cambios necesarios para que ser utilizados como variables categóricas. Se decidió utilizar el orden de nacimiento del nacimiento de los niños, si son hijos de mujeres que son o fueron madres adolescentes y su uso de métodos anticonceptivos. Para estas pruebas se obtuvo el estadístico de prueba chi cuadrado para validar el nivel de confianza de los resultados obtenidos. Además, se realizaron gráficas que mostrasen estos resultados.

6.5. Tablero de datos

Todos los resultados descritos en las secciones anteriores se incorporaron y centralizaron en un tablero de datos. De esta manera se puede publicar de manera centralizada, con fácil acceso a los datos y las perspectivas creadas. Ya que se contaron con múltiples perspectivas, se dividieron en diferentes categorías, para facilitar su acceso dentro del tablero. Estas categorías fueron creadas en el tablero a través de diferentes pestañas en el tablero creado. Las pestañas creadas son las siguientes:

- Características generales
- Edad de la madre al nacimiento
- Número de hijos

6.5.1. Despliegue y configuración de tablero

Tras realizar el tablero, se buscó una manera de desplegarlo para que pudiese ser accedido de manera pública. Se intentó hacer el despliegue hacia ShinyApps, que es la opción proveída por Posit, y tiene opciones para hacer el despliegue de manera automática. Provee equipos de hasta 1GB de RAM gratis, con opciones de comprar más recursos, a través de suscripciones mensuales o anuales. La opción por defecto no proveía suficientes recursos para hacer el despliegue, debido a la gran cantidad de procesamiento de datos que se realiza, por lo que se buscó otras opciones para realizar el despliegue.

Se eligió hacer un despliegue manual, a través de una máquina virtual ubicada en DigitalOcean, ya que provee equipos elásticos de cómputo y contiene documentación extensa sobre múltiples temas, incluyendo el despliegue de un servidor Shiny. Se inició configurando una máquina virtual con 4GB de RAM dentro de DigitalOcean. A esta máquina se le configuró un usuario con permisos *sudo*, pero no *root*, el cual se utilizó para las tareas posteriores de configuración.

Con este usuario, se instaló y configuró R, para poder instalar las herramientas necesarias para ejecutar el tablero de Shiny. Luego, se instaló y configuró Nginx para ser el servidor que servirá el tablero. Además, se configuró el *firewall* para permitir todas las conexiones entrantes hacia Nginx. Con estas configuraciones realizadas, se instaló Shiny y el servidor Shiny, a través del manejador de paquetes *apt*. Por defecto, los tableros de Shiny se ubican en el puerto 3838, y todos los proyectos se encuentran en rutas adicionales dentro de esta dirección.

Una vez se verificó que se podían ejecutar correctamente los proyectos de prueba, se copió el código desarrollado hacia la máquina virtual y se ejecutó por primera vez. Al momento de ejecutar el tablero por primera vez, se notó de que no funcionaba correctamente y las conexiones se interrumpían constantemente. Tras indagar entre los *logs* creados por el servidor de *Shiny*, se identificó que esto se debía a la falta de recursos para este proceso. Para solucionarlo, se apagó la máquina virtual y se aumentó la cantidad de recursos disponibles para la máquina virtual. Al colocarle 8GB RAM, se verificó que el tablero funcionaba correctamente y podía ser accedido desde múltiples equipos. Este valor fue elegido tras

aumentar paulatinamente los recursos de esta máquina virtual, hasta asegurar su correcto funcionamiento.

Para este punto, la página únicamente podía ser accedida a través de la dirección IP de la máquina virtual donde se configuró el servidor. Sin embargo, se identificó que de esta manera no se podía compartir fácilmente la dirección y no era fácil de recordar. Para solucionar estos problemas se decidió configurar un dominio dentro de DigitalOcean, para que redirigiese hacia la página configurada. El dominio se compró a través de la página *GoDaddy.com* y se eligió el dominio "http://www.osar-desnutricion.com". Para realizar la configuración necesaria de DNS, se cambió los *nameservers* en *GoDaddy.com* para apuntar hacia DigitalOcean. Dentro de DigitalOcean se crearon dos registros de tipo A para apuntar el dominio hacia la dirección IP. Por último, se esperó a que las configuraciones realizadas se propagaran, lo cual tomó un par de horas. Tras verificar que se podría acceder correctamente a este sitio a través del dominio configurado, el tablero tenía la siguiente forma:



Figura 25: Pantalla principal al ingresar al tablero

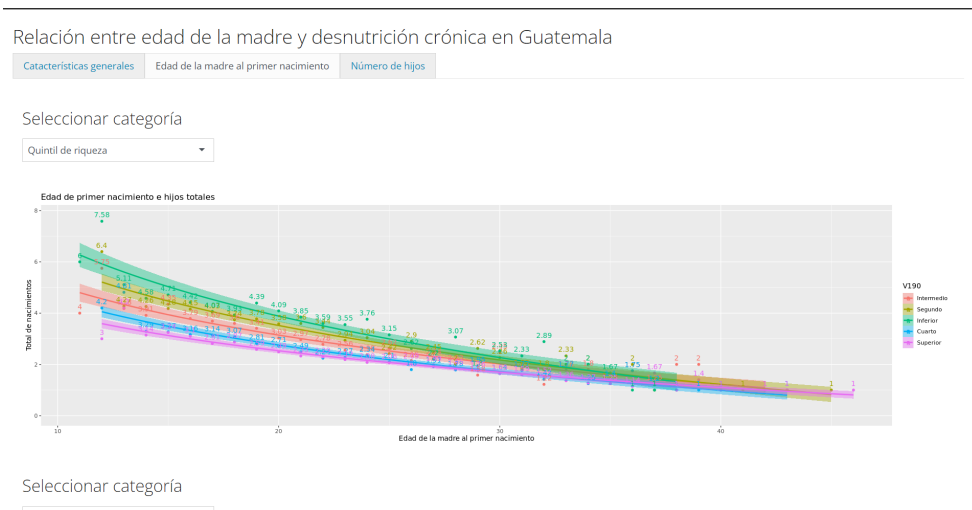


Figura 26: Segunda pestaña en tablero

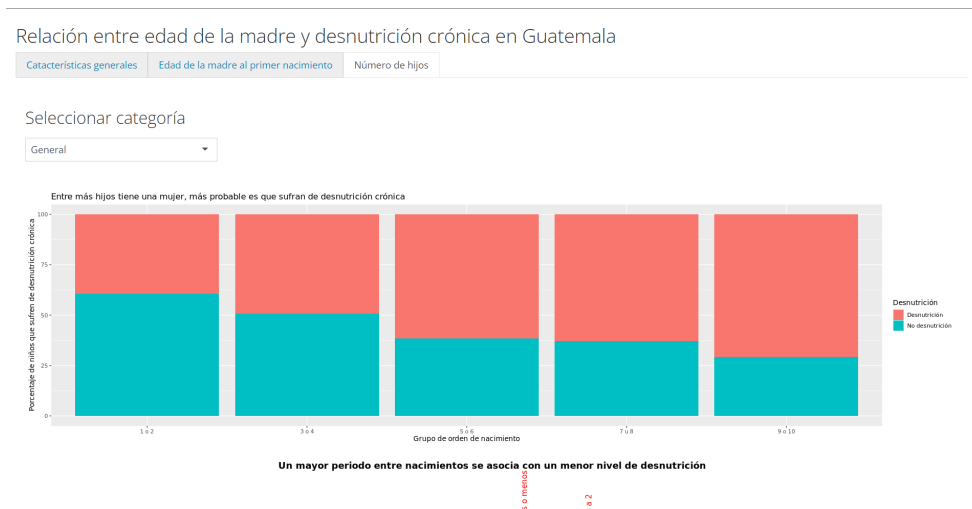


Figura 27: Tercera pestaña en tablero

Tras realizar el despliegue y verificar que pudiese ser accedido correctamente, se trabajó en poder interactuar con el tablero desde el navegador. Con excepción a las gráficas de Chi Cuadrado, se agregó la opción de segmentar todos los resultados por tres categorías, además de la categoría general:

- Quintil de riqueza
- Máximo nivel de educación alcanzado
- Área de residencia

Esto permite indagar más profundamente en el comportamiento de los resultados y poder identificar patrones adicionales. Además, se realizaron cambios estéticos en todas las gráficas. Se cambió el título de las gráficas, para incluir el texto presentado durante las reuniones anteriores. Se cambiaron las etiquetas en muchas de las gráficas, para expresar de mejor manera la información contenida. Por último, se agregó un texto a todas las gráficas para mostrar que habían sido creadas a base de un análisis secundario de la ENSMI 2014-2015.

Todos estos cambios fueron realizados de manera local, y luego se replicaron en el servidor con el tablero de datos. Utilizando RStudio se implementaron los cambios y se verificó que funcionase correctamente de manera local. En el servidor, se realizó un respaldo de los archivos ya configurados, creando una carpeta separada y moviendo los archivos a esta nueva ruta. Luego, se creó nuevamente dos archivos, a los cuales se les copió el contenido actualizado. Al realizar estos cambios, Shiny vuelve a crear la página web, por lo que se puede seguir accediendo a través de la dirección mencionada. En caso el tablero no funcionase correctamente durante este proceso, se pueden regresar los archivos originales a la ruta apropiada para hacer un *rollback* de los cambios realizados.

Debido a los altos precios asociados con la ejecución del servidor dentro de la máquina virtual ejecutada, se decidió reducir los recursos necesarios para su ejecución. Se exportó

los datos que utiliza cada gráfica y se modificó el programa utilizado para cargar únicamente estos datos. Esto permitió reducir la cantidad de RAM necesitada para la ejecución y después de este ajuste, se necesita menos de 512MB de memoria, por lo cual puede colocarse en la versión gratis de ShinyApps. La aplicación puede ser visitada a través de "<https://rgaroz.shinyapps.io/ENSMI/>", donde se encuentra alojado de manera gratuita.

7.1. Resultados iniciales

Las perspectivas obtenidas a través del análisis estadístico inicial son las siguientes:

La desnutrición crónica severa, en menores de 5 años, aumentó entre 2015 y 2022.

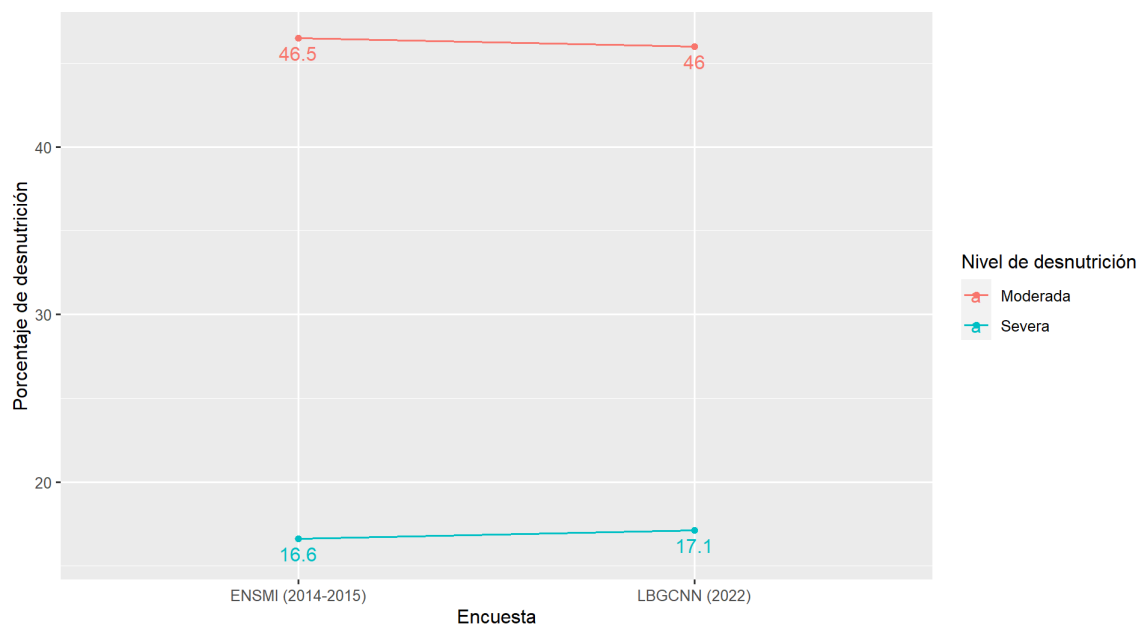


Figura 28: Porcentaje de desnutrición moderada y severa, en dos estudios realizados

La desnutrición crónica en menores de 3 años ha empeorado desde el 2015.

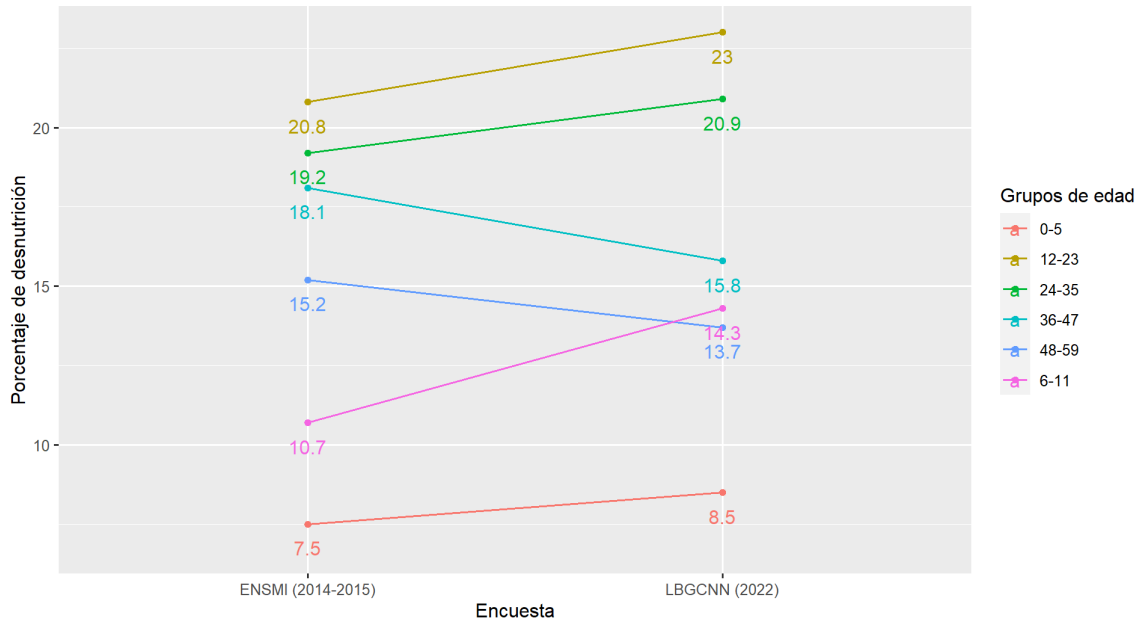


Figura 29: Porcentaje de desnutrición moderada y severa, en dos encuestas realizadas, según edad

Entre más hijos tiene una mujer, es más probable que sufran desnutrición crónica.

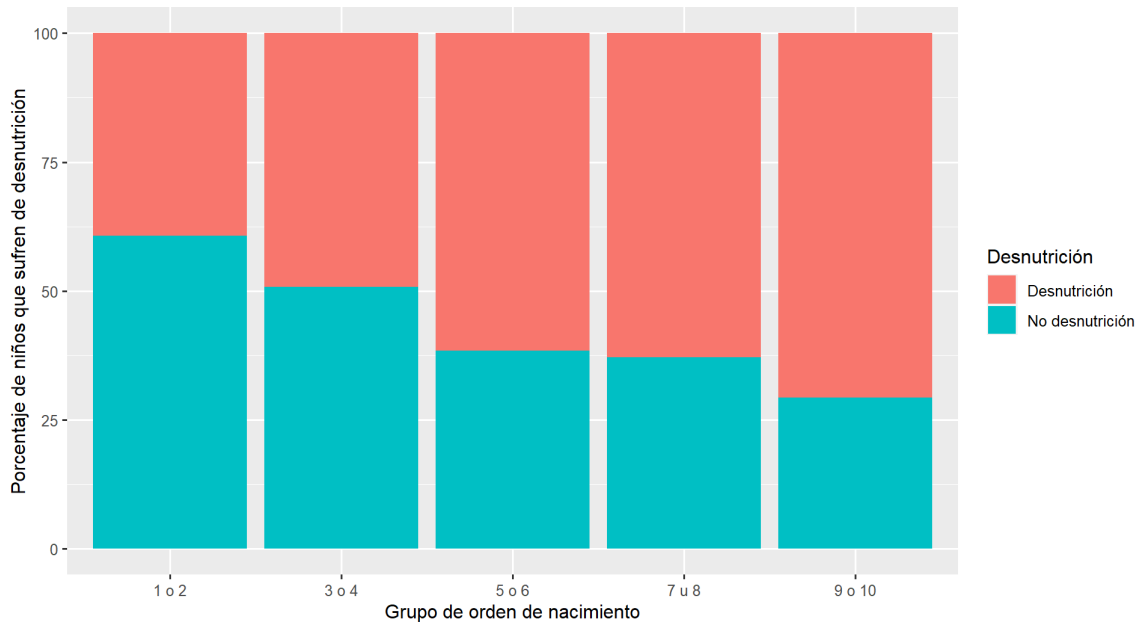


Figura 30: Porcentaje de niños que sufren de desnutrición, según grupo de orden de nacimiento

Entre más temprano suceda el primer nacimiento, más hijos tendrán las mujeres.

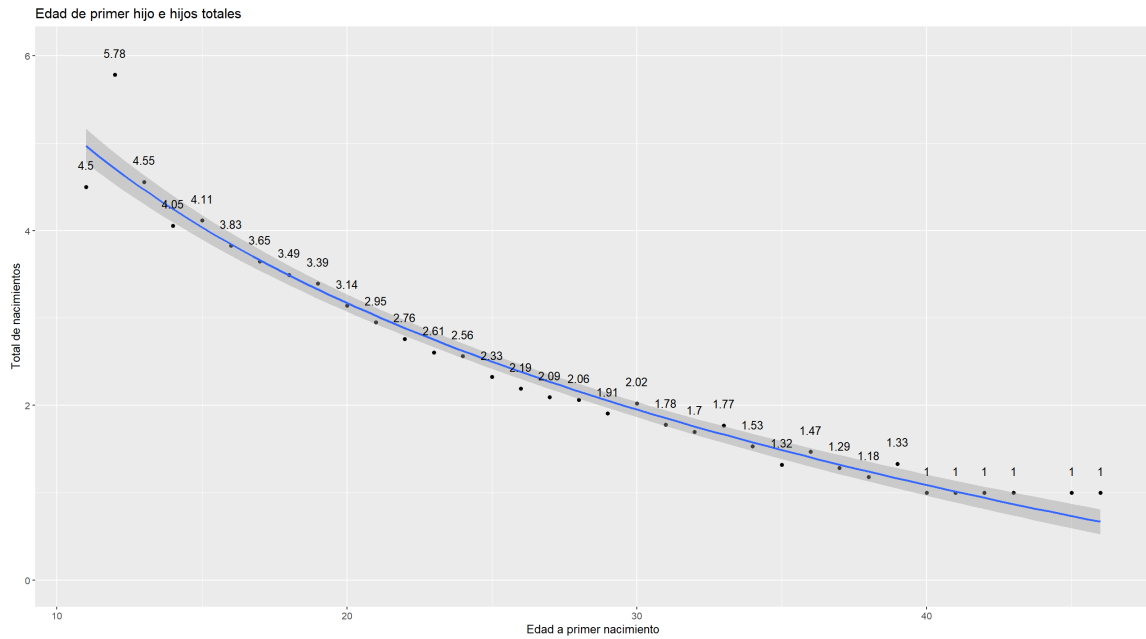


Figura 31: Promedio de hijos por mujer, según edad del primer hijo

Las mujeres que inician la maternidad en la adolescencia tienen en promedio un hijo más.

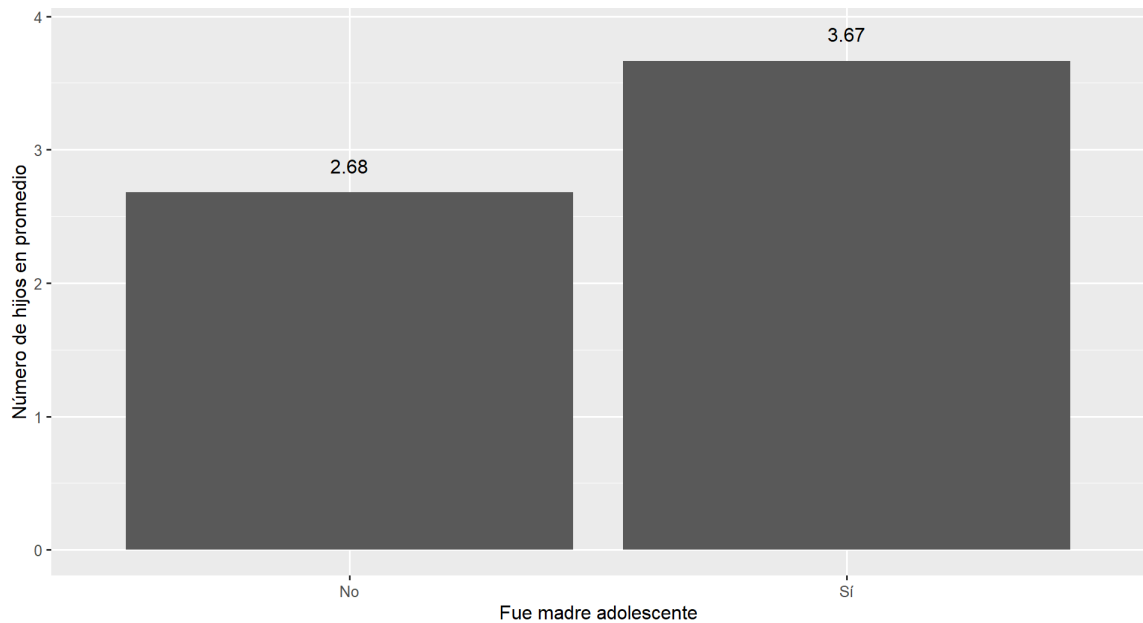


Figura 32: Promedio de hijos por mujer, según si fueron madres adolescentes o no

El 43% de las mujeres tienen por lo menos un hijo antes de cumplir 20 años.

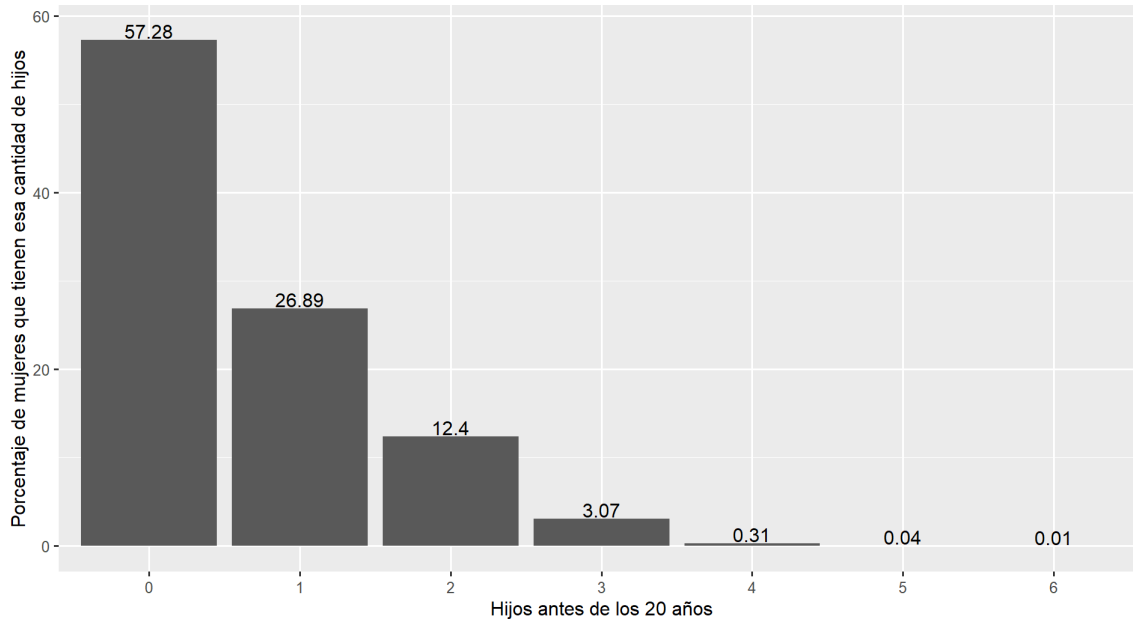


Figura 33: Porcentaje de mujeres, según la cantidad de hijos antes de los 20 años

Casi 60% de los niños con desnutrición crónica son hijos de mujeres que son o fueron madres adolescentes.

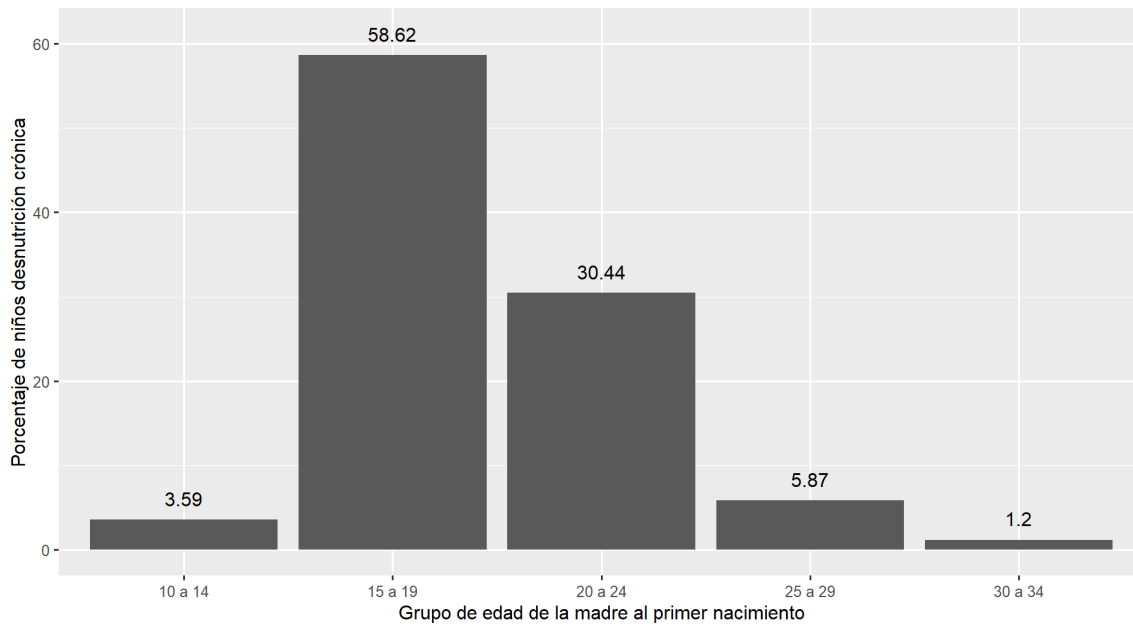


Figura 34: Porcentaje de niños que sufren de desnutrición crónica, según grupo de edad de la madre al primer nacimiento

Los hijos de las mujeres que tienen su primer nacimiento antes de los 20 años tienen más probabilidades de sufrir desnutrición crónica.

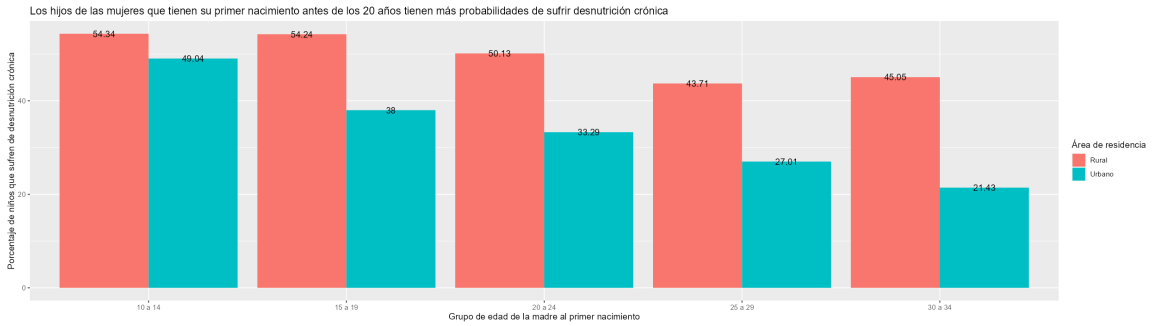


Figura 35: Porcentaje de niños que sufren de desnutrición crónica, según grupo de edad de la madre al primer nacimiento, por área rural y urbana

7.2. Resultados adicionales

Las perspectivas obtenidas tras realizar análisis adicional en los datos existentes son las siguientes:

Los resultados obtenidos tras la validación realizada son los siguientes:

Accuracy	Precision	Recall
0.70	0.74	0.77

Cuadro 3: Valores obtenidos de regresión logística

Ejemplos de las matrices de confusión se presentan a continuación:

\	No	Sí
No	134	49
Sí	37	80

\	No	Sí
No	136	40
Sí	34	89

\	No	Sí
No	120	49
Sí	50	80

El resto de las matrices pueden encontrarse en la sección de anexos.

Después de realizar la regresión logística, se encontraron los factores que más influían en el estado de desnutrición de los niños. Los 5 factores que más influyen en esto y su valor P son los siguientes:

Factor	Valor P
Peso de la madre	<2e-16
Quintil de riqueza inferior	1.12e-10
Periodo entre embarazos	9.72e-8
Altura de la madre	1.51e-6
Quintil de riqueza superior	9.61e-5

Cuadro 4: 5 valores p más bajos en regresión lineal realizada

Los porcentajes en los que estos factores influyen en la probabilidad de que un niño sufra de desnutrición son los siguientes:

Factor	Cambio en porcentaje por unidad
Peso de la madre	3 %
Quintil de riqueza inferior	16 %
Periodo entre embarazos	0.2 %
Altura de la madre	0.34 %
Quintil de riqueza superior	11 %

Cuadro 5: Cambio de porcentaje por unidad cambiada

Los factores con los cuales se identificó una relación significativa al 1 % son los siguientes, con sus valores chi cuadrado y su valor p:

Factor	Valores posibles	Chi cuadrado	Valor p
Periodo entre nacimientos	2 años o menos	71.81	<2.2e-16
	Más de 2 años		
Uso de métodos anticonceptivos	Método moderno	161.41	<2.2e-16
	Ningún método		
	Método tradicional		
Máximo nivel de educación alcanzado	Sin educación	965.81	<2.2e-16
	Primaria		
	Secundaria Superior		

Cuadro 6: Valores de chi y valores p

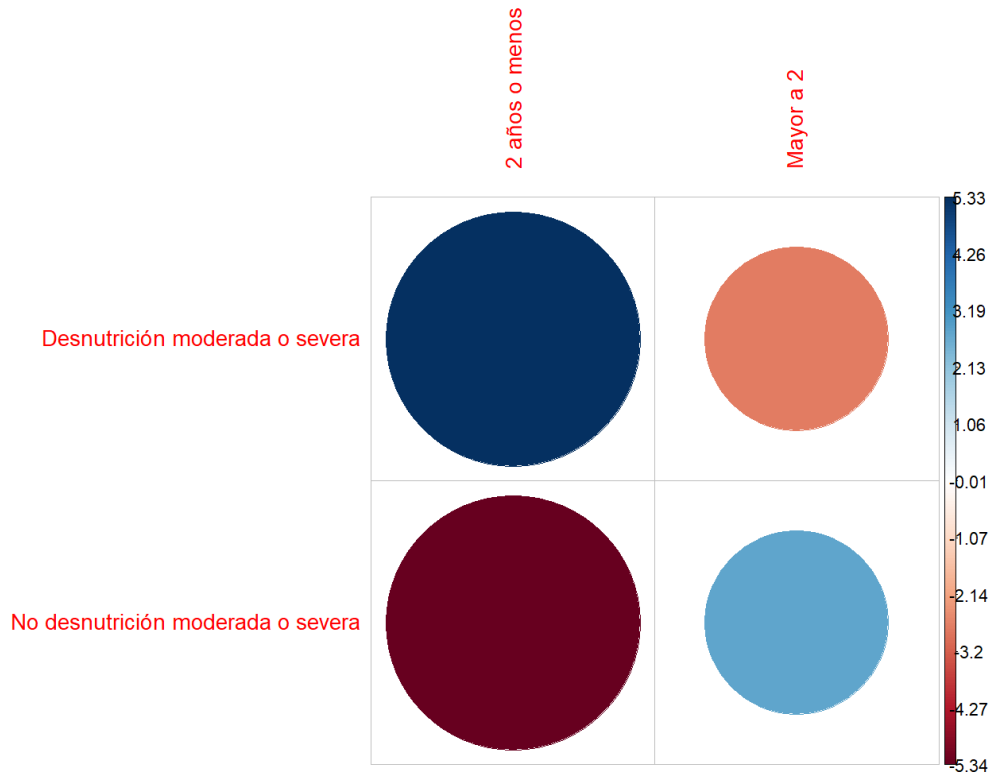


Figura 36: Porcentaje de desnutrición moderada y severa, en dos estudios realizados

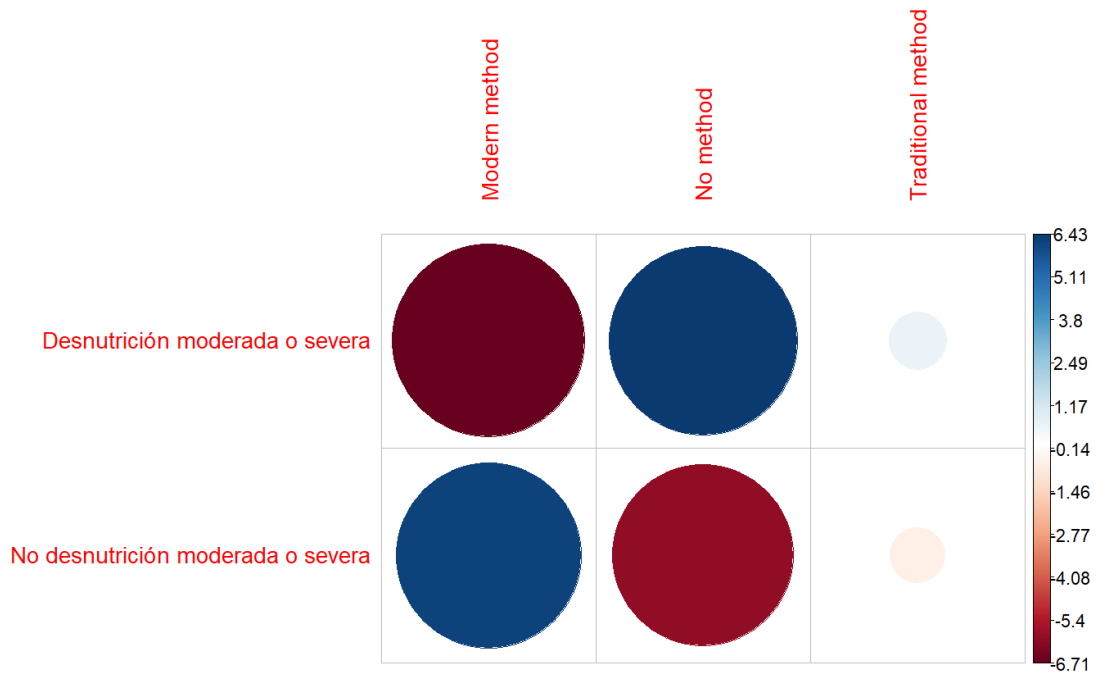


Figura 37: Porcentaje de desnutrición moderada y severa, en dos estudios realizados

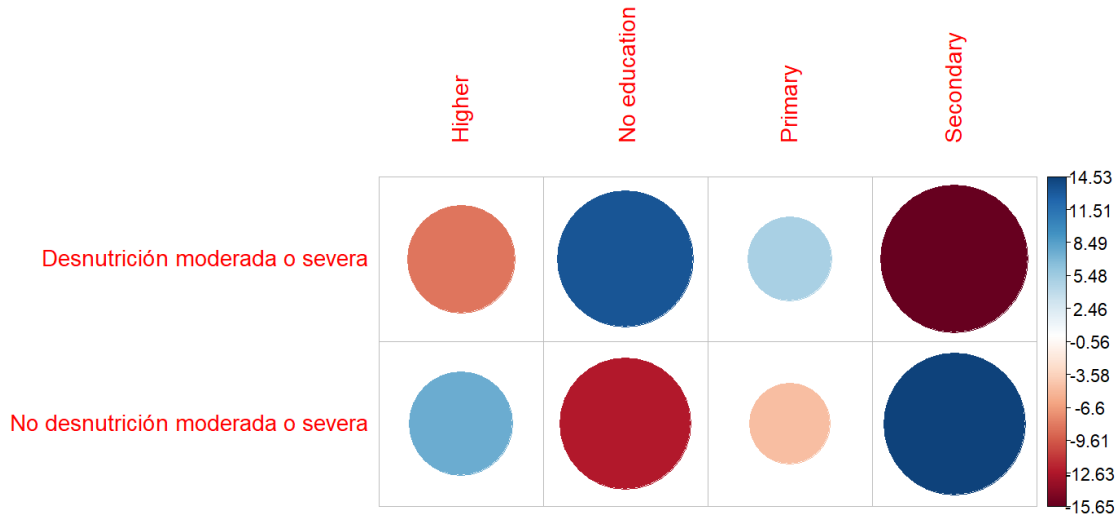


Figura 38: Porcentaje de desnutrición moderada y severa, en dos estudios realizados

7.3. Tablero creado

Después de implementar los cambios en el tablero y colocarlos en servidor apropiado, el tablero final fue el siguiente:

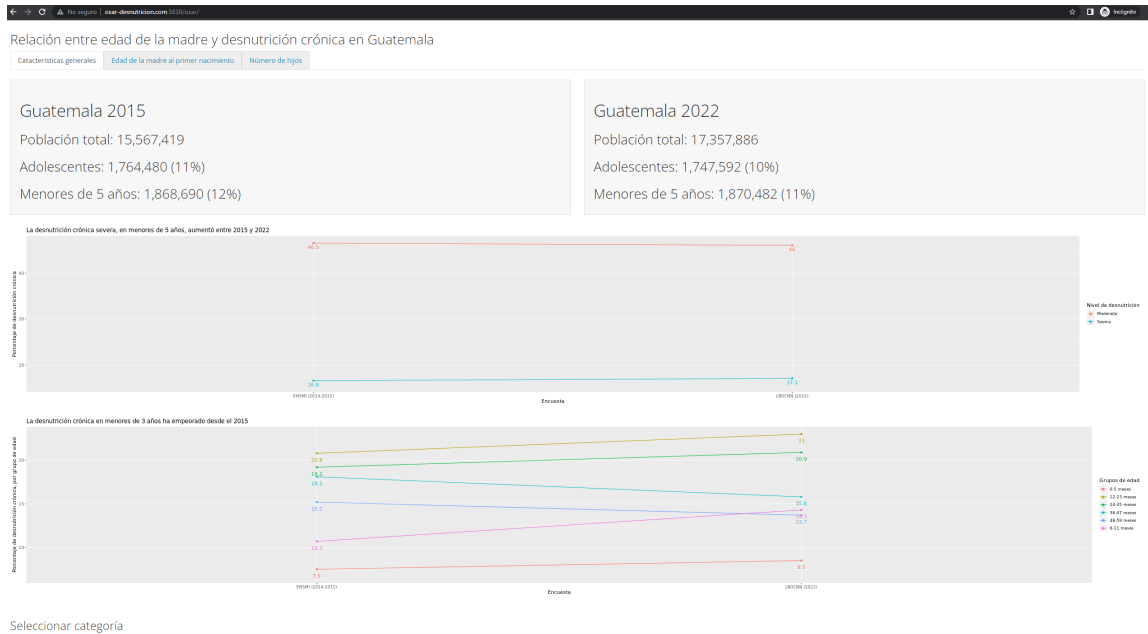


Figura 39: Primera pestaña en tablero

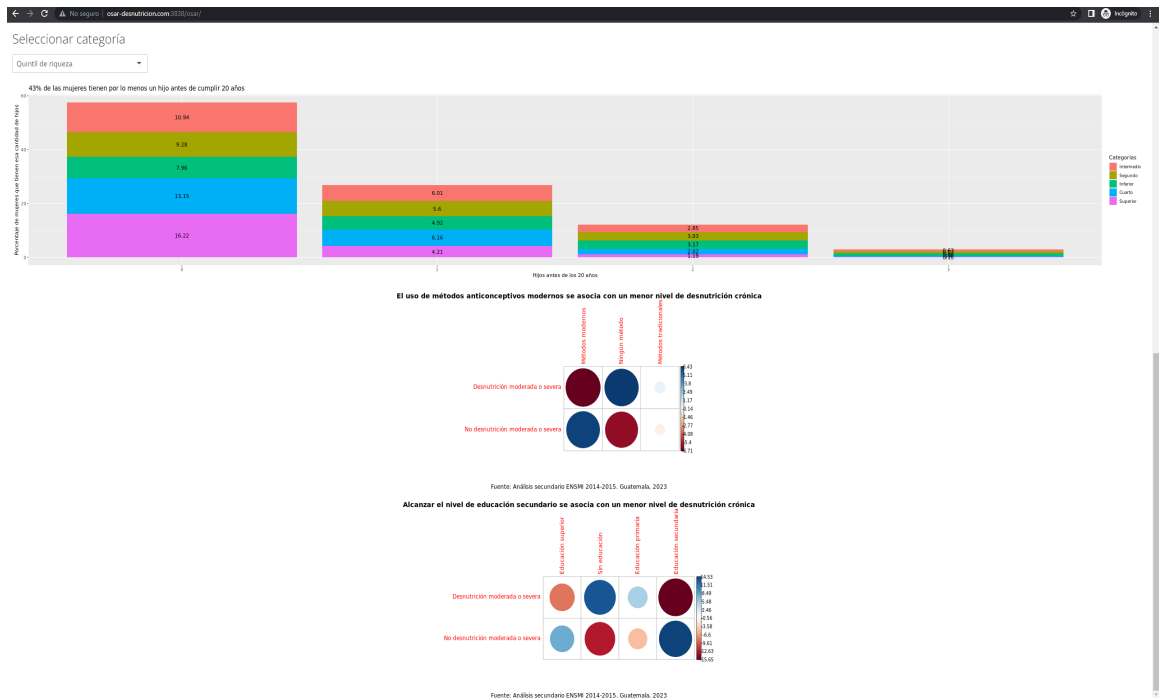


Figura 40: Primera pestaña en tablero

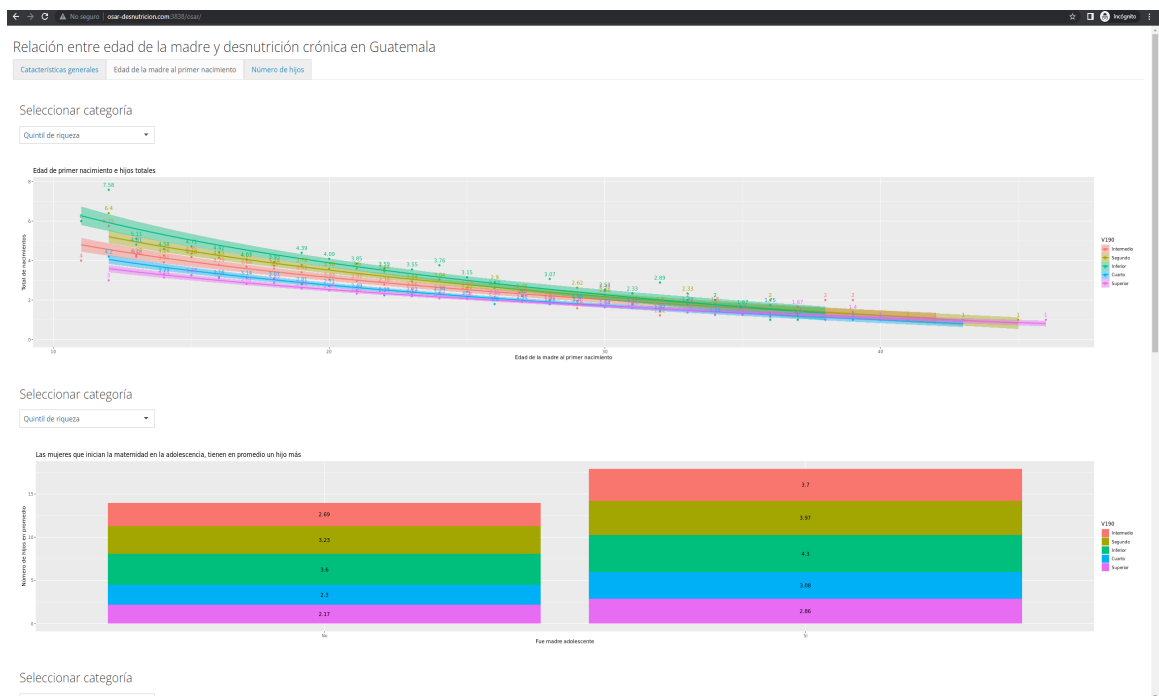


Figura 41: Segunda pestaña en tablero

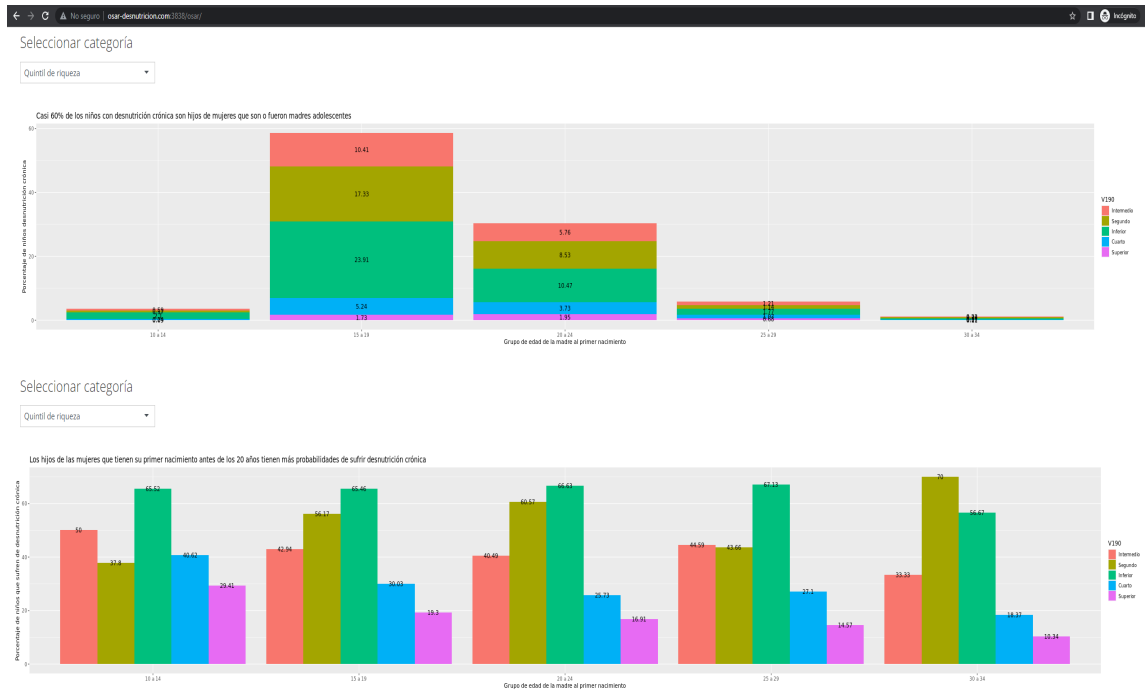


Figura 42: Segunda pestaña en tablero

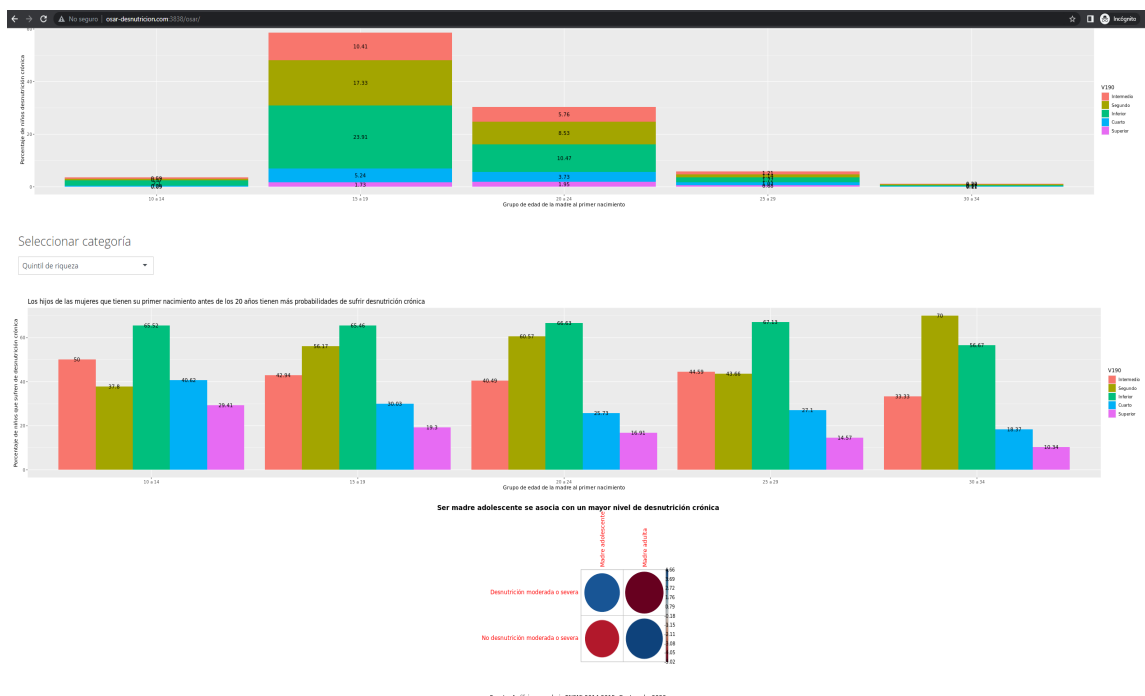


Figura 43: Segunda pestaña en tablero

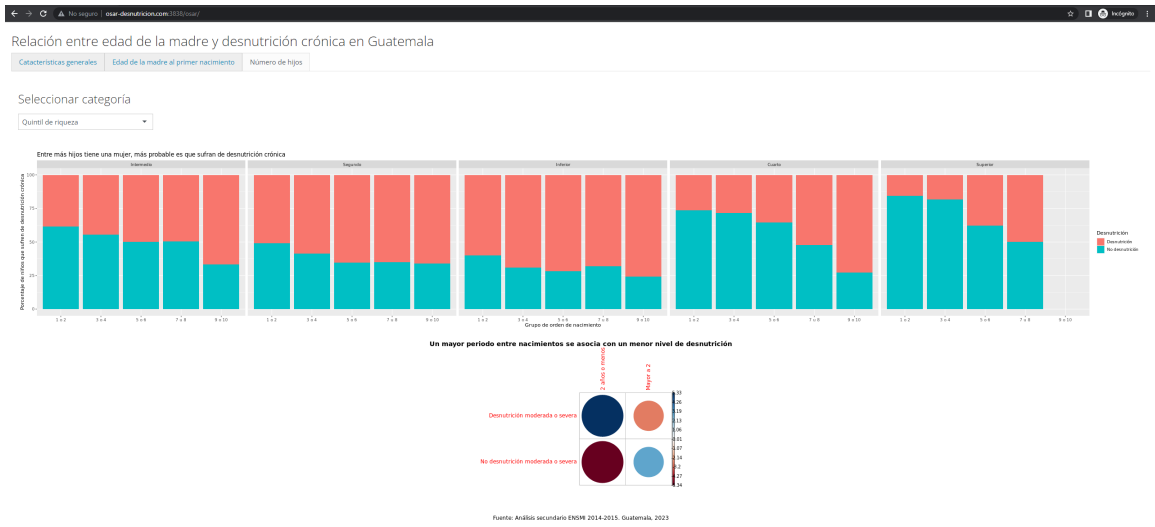


Figura 44: Tercera pestaña en tablero

Los resultados obtenidos muestran la influencia que ciertos factores tienen en el embarazo adolescente, y el efecto que esto tiene en la desnutrición crónica infantil. La relación más importante identificada es la relación entre edad de primer nacimiento, cantidad de hijos y nivel de desnutrición crónica. Entre antes se tenga el primer hijo, más hijos en promedio se tendrán, lo cual implica un mayor riesgo de desnutrición crónica. Además, se identificó que un mayor nivel de desnutrición crónica lleva a mayor desnutrición crónica en niños. Esto indica la existencia de un ciclo completo, donde tener hijos temprano lleva a tener más hijos, lo cual lleva a un mayor nivel de desnutrición crónica, lo cual aumenta el riesgo de desnutrición crónica futura.

8.1. Gráficas

La Figura 28 muestra el cambio de los niveles de desnutrición crónica moderada y severa entre la ENSMI 2014-2015 y LBCGNN 2022. Esta figura muestra que la desnutrición moderada disminuyó en 0.5 puntos porcentuales en este periodo de tiempo, mientras la desnutrición severa aumentó la misma cantidad en el mismo periodo de tiempo.

La Figura 29 muestra cómo ha cambiado la desnutrición moderada o severa entre la ENSMI 2014-2015 y la LBCGNN (2022) según grupos de edad de los niños. La desnutrición crónica moderada o severa ha disminuido en niños de 36 a 47 meses y de 48 a 59 meses. En los otros grupos de edad, la desnutrición crónica ha aumentado en este periodo de tiempo. Particularmente importante es el aumento en el grupo de edad de 6 a 11 meses, donde aumentó en 3.6%, el cambio más importante entre todos los grupos.

La Figura 30 muestra el porcentaje de niños que surgen por desnutrición crónica moderada o severa, según el grupo de orden de nacimiento. Se puede observar que a mayor orden

de nacimiento, mayor porcentaje de niños sufren de desnutrición crónica. alrededor de 40 % de niños con el orden 1 o 2 sufren de desnutrición crónica, y este valor aumenta casi a 70 % en niños con el orden 9 o 10.

La Figura 31 muestra el promedio de hijos por mujer, según la edad del primer nacimiento. Se puede observar que mientras antes se tenga el primer hijo, mayor cantidad de hijos tiene una mujer. Al tener el primer hijo a los 15 años, en promedio se tiene 4.11 hijos. Si el primer nacimiento ocurre 5 años después, a los 20, este número baja a 3.14 y en 5 años más este valor desciende aún más a 2.33 hijos en promedio. Se realizó una regresión exponencial para encontrar la relación entre ambas variables. Esto indica que atrasar el primer nacimiento a una edad joven tiene un mayor impacto en la cantidad de hijos que atrasarlo en años posteriores.

La Figura 32 muestra la cantidad promedio de hijos, según si la madre fue madre adolescente o no. Se observa que las mujeres que fueron madres adolescentes tienen en promedio un hijo más que las mujeres que no fueron madres adolescentes.

La Figura 33 muestra el porcentaje de mujeres, según la cantidad de hijos que se tienen antes de los 20 años. Se puede observar que 43 % de todas las mujeres tienen por lo menos un hijo antes de los 20 años. Se identifican casos de mujeres con 6 hijos antes de los 20 años.

La Figura 34 muestra el porcentaje de niños con desnutrición crónica, según el grupo de edad de la madre al primer nacimiento. Se observa que casi 60 % de los niños que sufren de desnutrición crónica son hijos de madres adolescentes. Además, se identifica que casi 90 % de los niños con desnutrición crónica son hijos de madres que tuvieron su primer hijo antes de los 25 años.

Por último, la Figura 35 muestra que el porcentaje de niños que sufren de desnutrición crónica, según el grupo de edad de la madre al primer nacimiento. Se puede observar que el porcentaje disminuye mientras aumenta el grupo de edad del primer nacimiento. El porcentaje de hijos nacidos de mujeres que tuvieron su primer hijo entre 15 a 19 años es de 54 % en el área rural y 49 % en el área urbana, mientras que este valor disminuye a 50 % y 33 % respectivamente en el grupo de 20 a 24.

8.2. Regresión logística

Además, se realizó una regresión logística para identificar la influencia que ciertos factores tienen sobre el nivel de desnutrición de los niños. Ya que la regresión logística necesita de un valor categórico para predecir, se creó una nueva variable binaria para definir si un niño sufría de desnutrición. Se tomó la definición utilizada en el informe de la ENSMI, la cual indica que un niño sufre de desnutrición crónica moderada o severa si se aleja por más de -2 desviaciones estándar de la media.

Se eligió esta técnica basada en la flexibilidad que provee y los antecedentes que se identificaron con este método. Esta técnica permite procesar diferentes tipos de variables, numéricas y categóricas. La interpretación de estos resultados también resulta muy útil al momento de ser presentado a OSAR, ya que identifican los factores que influyen el resultado y el impacto que tienen. Además, existen múltiples antecedentes en otros países donde se

efectuaron estudios similares utilizando estas técnicas, por lo que se decidió implementar en este contexto.

En algún momento se consideraron otras técnicas, como reducción de dimensionalidades y regresiones lineales múltiples. Sin embargo, estas técnicas no funcionaban correctamente con los tipos de datos existentes, por lo cual se debía reducir mucho las dimensiones del conjunto de datos a utilizar, lo cual afectó los resultados obtenidos.

El modelo entrenado mostró un buen desempeño en las pruebas realizadas notando un buen resultado en el *recall* obtenido. Se notó la importancia de esta métrica para el caso de uso identificado. Si se desea implementar este modelo en un ambiente productivo es importante identificar un alto porcentaje de niños que puedan sufrir de desnutrición crónica, para poder darles seguimiento. No se identificó un desempeño particularmente malo al momento de identificar demasiados falsos positivos o negativos. Por esta situación, se decidió utilizar estas métricas y no realizarse ajustes adicionales a estas métricas.

El Cuadro 1 muestra los 5 valores p más bajos de la regresión logística realizada. El valor p muestra la probabilidad de que el factor de una característica sea 0, lo cual indicaría que no influye en la variable dependiente. Un valor p bajo muestra que los cambios en los valores de esta variable influyen en los valores de la variable dependiente.

El valor más bajo es el peso de la madre, lo cual indica que tiene una mayor influencia en el nivel de desnutrición de un hijo. Además de este valor, las características que más influyen es pertenecer al quintil de riqueza inferior, el periodo entre embarazos, la altura de la madre y pertenecer al quintil de riqueza superior. Esto concuerda con los estudios realizados por Conde Agudelo sobre la relación que el periodo entre nacimientos tiene la salud de la madre y de los hijos.

El peso de la madre también tiene la mayor influencia en la probabilidad de que un niño sufra de desnutrición. Cada cambio en la unidad de peso de la madre implica un cambio del 3% en la probabilidad de que un niño sufriese de desnutrición. La altura de la madre también tiene una influencia importante, con un cambio de 0.34% por cada unidad de cambio. En el caso del periodo entre embarazos es de 0.2% por cada mes adicional. Esto significa que por cada año que pase entre el nacimiento entre hijos, la probabilidad de que un niño sufra de desnutrición disminuye por 2.5%.

Para las variables categóricas, se puede observar los siguientes comportamientos. Pertenecer al quintil de riqueza conlleva un 16% de probabilidad superior de sufrir de desnutrición, dejando todos los demás parámetros iguales. Pertenecer al quintil superior de riqueza conlleva una probabilidad de 11%, bajo las mismas condiciones.

8.3. Pruebas de chi cuadrado

También, se realizaron pruebas de chi cuadrado, para intentar encontrar relaciones entre múltiples variables categóricas, con los siguientes resultados. En todas estas pruebas, se tomaron las variables categóricas y se obtuvo su valor de chi cuadrado con un 95% de confianza. Para todas las variables, se intentó identificar una posible correlación con el nivel de desnutrición, utilizando la misma categoría descrita anteriormente. Con esta prueba, se

pudo identificar que existía una correlación significativa entre el periodo entre nacimientos, uso de métodos anticonceptivos y el máximo nivel de educación alcanzado.

Se eligió esta prueba ya que permite validar los resultados obtenidos con la regresión logística y permiten validar las calificaciones utilizadas por la ENSMI 2014-2015 y LBGCNN. Esto complementa los resultados obtenidos anteriormente y proveen una validación adicional para obtener la misma información. Además, se identificó que esta técnica es utilizada de la misma manera en estudios realizados en otros países centrados en la misma temática.

Para la definición de periodo entre nacimientos, se utilizó la definición encontrada en FUENTE, por lo que se separó el periodo entre embarazos en un periodo mayor y menor a 2 años. Utilizando estas categorías, se volvió a validar la relación que existe entre el periodo entre nacimientos y el nivel de desnutrición crónica.

Además, se validó el impacto que el uso de métodos anticonceptivos modernos tiene en el nivel de desnutrición de los niños. Para esta parte, se tomó la definición de DHS sobre los tipos de métodos anticonceptivos. Se eligió la categorización de métodos anticonceptivos modernos, métodos tradicionales y ningún método. Se identificó que el uso de métodos modernos mostraba una relación con un nivel de desnutrición más bajo. No usar métodos anticonceptivos también se asocia con un menor nivel de desnutrición. Por último, el uso de métodos anticonceptivos tradicionales se asocia con mayores niveles de desnutrición.

Por último, se validó la relación con el máximo nivel de desnutrición alcanzado. Vale la pena mencionar que esto hace referencia al máximo nivel empezado, no necesariamente el máximo nivel terminado. Un mayor nivel de educación se asocia con un menor nivel de desnutrición. Esto se puede explicar ya que existe un mayor quintil de riqueza, lo cual se ha identificado como un factor de mejor acceso a servicios de salud.

8.4. Tablero

El tablero de datos se colocó en una máquina virtual en DigitalOcean. Se eligió este proveedor ya que cuenta con una interfaz amigable y cuenta con documentación extensa sobre múltiples tareas de instalación y de administración. Durante el proceso de instalación, se probaron diferentes cantidades de recursos para la máquina virtual que utilizaría este tablero. Fue necesario utilizar una máquina virtual con 8GB RAM y un procesador compartido, lo cual fue el mínimo de recursos con lo cual el tablero funcionó correctamente.

Esta máquina virtual tiene un costo de 48 dólares al mes. Por otro lado, el dominio comprado tiene un costo de 12 dólares al año. En total, esto significaría un costo cercano a 588 dólares para OSAR. Este costo podría disminuirse al realizar cambios al tablero final, para que necesite menos recursos, y así poder utilizar una máquina que únicamente tenga 4GB RAM, lo cual disminuiría el costo a únicamente 24 dólares al mes, lo cual resultaría en un costo de 300 dólares al año.

La opción proveída por Shiny Apps tiene un costo similar, con 50 dólares al mes. Esta opción también da acceso a despliegues con 8 GB RAM, lo cual podría cumplir con las necesidades identificadas. En esta opción se debería utilizar un dominio que incluya *shinyapps* y el manejo de la aplicación se realiza directamente desde Shiny Apps.

Este tablero podría utilizarse también para encuestas de DHS realizadas en otros países o en Guatemala en años anteriores o posteriores. Es importante mencionar que, durante un periodo de tiempo, los datos fueron recolectados por el MSPAS, no por DHS, por lo cual no pueden utilizarse directamente en estas herramientas. Esto podría resultar en comparaciones y tableros adicionales sin necesidad de modificaciones adicionales de los datos. Esto es posible ya que los nombres de las columnas y lo que representan es estándar en todas las encuestas de este tipo.

Para poder realizar este cambio, debe realizarse la transformación inicial de los datos como se hizo con los resultados actuales. Esto implicaría convertir los archivos oficiales obtenidos en archivos .csv, ya que son los archivos que se utilizan para alimentar el tablero. Estos se importan al inicio del programa, y en este punto se puede hacer la sustitución para poder ser cargados en el tablero.

De igual manera, estos datos podrían ser utilizados en otros tipos de herramientas que ya utilizaron encuestas DHS. Para ambos casos es necesario validar el funcionamiento del tablero antes de entrar en un ambiente de producción. Además, es necesario validar que los datos utilizados puedan utilizarse en este tipo de proyectos, dadas las restricciones existentes al momento de obtener estos datos.

Los resultados de la ENSMI cuentan con múltiples indicadores importantes para identificar embarazos adolescentes. Principalmente, contiene información sobre el primer nacimiento y edad acerca de la primera unión, lo cual es un buen predictor para el embarazo adolescente. Además, contiene información demográfica importante de las madres, lo cual es un factor importante adicional para el embarazo adolescente y los niveles de desnutrición.

Es evidente que existe una relación directa entre el embarazo adolescente y la desnutrición crónica. Tener un embarazo temprano es un buen predictor para la cantidad de hijos que se tendrá y la probabilidad de sufrir desnutrición aumenta con el orden de hijos que se tiene.

La información recolectada y el tablero creado ya está siendo utilizado por OSAR para sus tareas. Esto se ve reflejado en los materiales que han sido creados a partir de esto y el seguimiento que se le ha dado con aliados de OSAR.

Existen múltiples opciones para hacer un despliegue de un tablero de Shiny. La opción oficial es una manera fácil de realizar el despliegue, pero conlleva gastos adicionales para poder contar con recursos suficientes. También, puede realizarse el despliegue hacia una máquina virtual, lo cual provee mayor flexibilidad, pero requiere más trabajo para la instalación y mantenimiento.

Tras realizar el trabajo descrito en este documento, se identificaron algunas recomendaciones para investigaciones posteriores. Estas recomendaciones tienen el objetivo de mejorar el producto final y ampliar su usabilidad en el futuro. Estas podrían aplicarse a la misma herramienta presentada o en productos futuros.

La primera recomendación es modificar el programa en R y el despliegue realizado para automatizar el uso de los resultados de otros años. Esto implicaría configurar la carga de nuevos archivos que utilicen el estándar DHS e identificar posibles errores en el procesamiento, para no mostrarlos. Esto permitiría crear múltiples tableros y realizar comparaciones con otros años y otros países.

También, se recomienda realizar este análisis en una mayor cantidad de temáticas que la desnutrición crónica. Esto podría ser para los tres tipos de desnutrición, ya que la ENSMI contiene suficiente información para realizar este análisis. Además, podrían incluirse otros factores relacionados a las características de vivienda de los niños, información que también se encuentra en los resultados de la ENSMI 2014-2015.

Por último, se recomienda utilizar el tablero creado en este trabajo con conjuntos de datos de DHS de otros países que participen en esta encuesta, los cuales son países en vías de desarrollo. Ya que estos datos se encuentran estandarizados, este tablero también podría ser utilizado con estos conjuntos de datos. A través de todos los países encuestados, las mismas columnas representan los mismos datos, por lo que no es necesario cambiar la definición de las gráficas en el programa. Esto permitiría realizar comparaciones entre diferentes países o entre los mismos países en diferentes momentos.

-
- [1] A. Wemakor, H. Garti, T. Azongo, H. Garti y A. Atosona, «Young maternal age is a risk factor for child undernutrition in Tamale Metropolis, Ghana,» *BMC Research Notes*, vol. 11, n.º 1, 2018. DOI: 10.1186/s13104-018-3980-7.
- [2] P. H. Nguyen, S. Scott, S. Neupane, L. M. Tran y P. Menon, «Social, biological, and programmatic factors linking adolescent pregnancy and early childhood undernutrition: A path analysis of India's 2016 National Family and Health Survey,» *The Lancet Child Adolescent Health*, vol. 3, n.º 7, págs. 463-473, 2019. DOI: 10.1016/s2352-4642(19)30110-5.
- [3] R. E. Henríquez, M. M. Velásquez y R. Sibrián, *Estado nutricional de menores de cinco años de madres adolescentes de las comunidades La Joya y Rancho Quemado del Municipio de Perquín y El Matazano del Municipio de Arambala, Departamento de Morazán, El Salvador*.
- [4] P. H. Nguyen, S. Scott, L. Q. Khuong et al., «Adolescent birth and child undernutrition: An analysis of demographic and health surveys in Bangladesh, 1996–2017,» *Annals of the New York Academy of Sciences*, vol. 1500, n.º 1, págs. 69-81, 2021. DOI: 10.1111/nyas.14608.
- [5] A. Conde-Agudelo, J. M. Belizán y C. Lammers, «Maternal-perinatal morbidity and mortality associated with adolescent pregnancy in Latin America: Cross-sectional study,» *American Journal of Obstetrics and Gynecology*, vol. 192, n.º 2, págs. 342-349, 2005. DOI: 10.1016/j.ajog.2004.10.593.
- [6] I. N. de Estadística, *Censo poblacional*, 2019. dirección: <https://www.censopoblacion.gt/cuantosomos>.
- [7] M. de Salud Pública y Asistencia Social, I. N. de Estadística y S. de Planificación y Programación de la Presidencia. Ministerio de Salud y Asistencia Infantil, 2017.
- [8] G. Institute, *Maternidad Temprana en Guatemala: Un Desafío Constante*, 2006. dirección: https://www.guttmacher.org/sites/default/files/report_pdf/rib-guatemala.pdf.

- [9] S. de Seguridad Alimentaria y Nutricional, I. de Nutrición de Centro América y Panamá e I. N. de Estadística. Secretaría de Seguridad Alimentaria y Nutricional, 2022.
- [10] R. M. Kliegman, B. F. Stanton, J. W. St Geme III y N. F. Schor, *Textbook of pediatrics*. Elsevier Saunders, 2016.
- [11] S. de Seguridad Alimentaria y Nutricional de la Presidencia de la República, *Acuerdo de trabajo conjunto entre la Secretaría de Seguridad Alimentaria y Nutricional de la Presidencia de la República -SESAN- y el Observatorio en Salud Sexual y Reproductiva -OSAR-*, 2021.
- [12] T. R. Foundation, *The R project for statistical computing*, 2000. dirección: <https://www.r-project.org/>.
- [13] T. organization, *Tiobe index*, jun. de 2022. dirección: <https://www.tiobe.com/tiobe-index/>.
- [14] F. Beuke, *GitHut 2.0*, 2016. dirección: https://madnight.github.io/githut/#/pull_requests/2022/4.
- [15] Posit, *RStudio desktop*, mar. de 2023. dirección: <https://posit.co/download/rstudio-desktop/>.
- [16] T. R. F. for Statistical Computing, *Statutes of "The R Foundation for Statistical Computing"*, 2002.
- [17] H. Wickham y G. Grolemund, *R for Data Science: Import, Tidy, transform, visualize, and model data*. O'REILLY MEDIA, 2023.
- [18] Posit, 2014. dirección: <https://shiny.rstudio.com/>.
- [19] Posit, *Welcome to shiny*, 2014. dirección: <https://shiny.rstudio.com/tutorial/written-tutorial/lesson1/>.
- [20] IBM, *What is cloud computing?* Dirección: <https://www.ibm.com/topics/cloud-computing>.
- [21] Bloomberg, *Digitalocean LLC - company profile and news*. dirección: <https://www.bloomberg.com/profile/company/0852730D:US?leadSource=uverify+wall>.
- [22] DigitalOcean, *Droplets: Digitalocean's Scalable Virtual Machines*. dirección: <https://www.digitalocean.com/products/droplets>.
- [23] *What is DNS? | how DNS works | cloudflare*. dirección: <https://www.cloudflare.com/learning/dns/what-is-dns/>.
- [24] Netcraft, *March 2022 web server survey*, mar. de 2022. dirección: <https://news.netcraft.com/archives/2022/03/29/march-2022-web-server-survey.html>.
- [25] D. R. Anderson, D. J. Sweeney, T. A. Williams et al., *Statistics for Business and Economics*. Cengage, 2020.
- [26] J. E. Castle, «Maximizing Research Opportunities: Secondary data analysis,» *Journal of Neuroscience Nursing*, vol. 35, n.º 5, págs. 287-290, 2003. DOI: 10.1097/01376517-200310000-00008.
- [27] M. S. Brown, *What it needs to know about the data mining process*, jul. de 2015. dirección: <https://www.forbes.com/sites/metabrown/2015/07/29/what-it-needs-to-know-about-the-data-mining-process/?sh=631208ae515f>.
- [28] P. Chapman, *The CRISP-DM User Guide*.

- [29] J. Liedtka, *Why design thinking works*, feb. de 2023. dirección: <https://hbr.org/2018/09/why-design-thinking-works>.
- [30] I. D. Foundation, *What is design thinking?* Jul. de 2022. dirección: <https://www.interaction-design.org/literature/topics/design-thinking>.
- [31] I. Pardoe, *Lesson 12: Logistic, Poisson; nonlinear regression*. dirección: <https://online.stat.psu.edu/stat462/node/90/>.
- [32] I. N. de Estadística, *Proyecciones*, 2018. dirección: <https://www.ine.gob.gt/proyecciones/>.
- [33] C. E. para América Latina y el Caribe. «Malnutrición en niños y niñas en América Latina y el Caribe.» (jul. de 2018), dirección: <https://www.cepal.org/es/enfoques/malnutricion-ninos-ninas-america-latina-caribe>.
- [34] A. C. el Hambre. (2022), dirección: <https://www.accioncontraelhambre.org/es/desnutricion-cronica-que-es>.
- [35] INCAP. (2020), dirección: <http://www.incap.int/index.php/es/desnutricion2>.
- [36] W. H. Organization, *World Health Organization releases New Child Growth Standards*, 2006. dirección: <https://www.who.int/news/item/27-04-2006-world-health-organization-releases-new-child-growth-standards>.
- [37] UNICEF, 2019. dirección: <https://www.unicef.es/noticia/desnutricion-en-guatemala>.
- [38] UNICEF, *En Guatemala el 49,8% de los Niños Sufre Desnutrición Crónica*, María Claudia Santizo, *Oficial de Nutrición en unicef Guatemala*, 2019. dirección: <https://www.unicef.es/noticia/en-guatemala-el-498-de-los-ninos-sufre-desnutricion-cronica-maria-claudia-santizo-oficial>.
- [39] P. Salazar, *Prevenir el embarazo en Adolescentes es un compromiso impostergable*, sep. de 2022. dirección: <https://www.prensalibre.com/opinion/columnasdiarias/prevenir-el-embarazo-en-adolescentes-es-un-compromiso-impostergable/>.
- [40] U. N. P. Fund. 2022.
- [41] A. J. WHITE, *Washington Manual of Pediatrics*. WOLTERS KLUWER MEDICAL, 2022.
- [42] O. de Salud Reproductiva, *¿Quiénes Somos?* 2018. dirección: <https://osarguatemala.org/quienes-somos/>.
- [43] Demographic y H. S. Program, *STANDARD RECODE MANUAL*, 2018.
- [44] M. Barrientos, *Salud llevará a Cabo Encuesta de Salud materno infantil*, ene. de 2022. dirección: <https://www.prensalibre.com/ahora/guatemala/politica/salud-llevara-a-cabo-encuesta-de-salud-materno-infantil/>.
- [45] Demographic y H. Surveys, *Using Datasets for Analysis*, 2014. dirección: <https://www.dhsprogram.com/data/Using-Datasets-for-Analysis.cfm>.

12.1. Columnas existentes

Nombre de columna	Valor que contiene
V012	Edad en años completados
V106	Nivel de educación máximo alcanzado
V190	Quintil de riqueza
V201	Nacimientos totales
V212	Edad al primer nacimiento
V509	Edad de primera unión
V525	Edad de debut sexual

12.2. Agenda de reunión



ALIMENTARIA Y NUTRICIONAL
DE LA PRESIDENCIA
DE LA REPÚBLICA

Agenda

Reunión Comité Técnico de seguimiento a la atención de embarazadas adolescentes

Lugar: Hotel Las Américas, pendiente de confirmar el salón.

Fecha: 15 de marzo del 2023

Hora: 7:30 - 10:00 horas

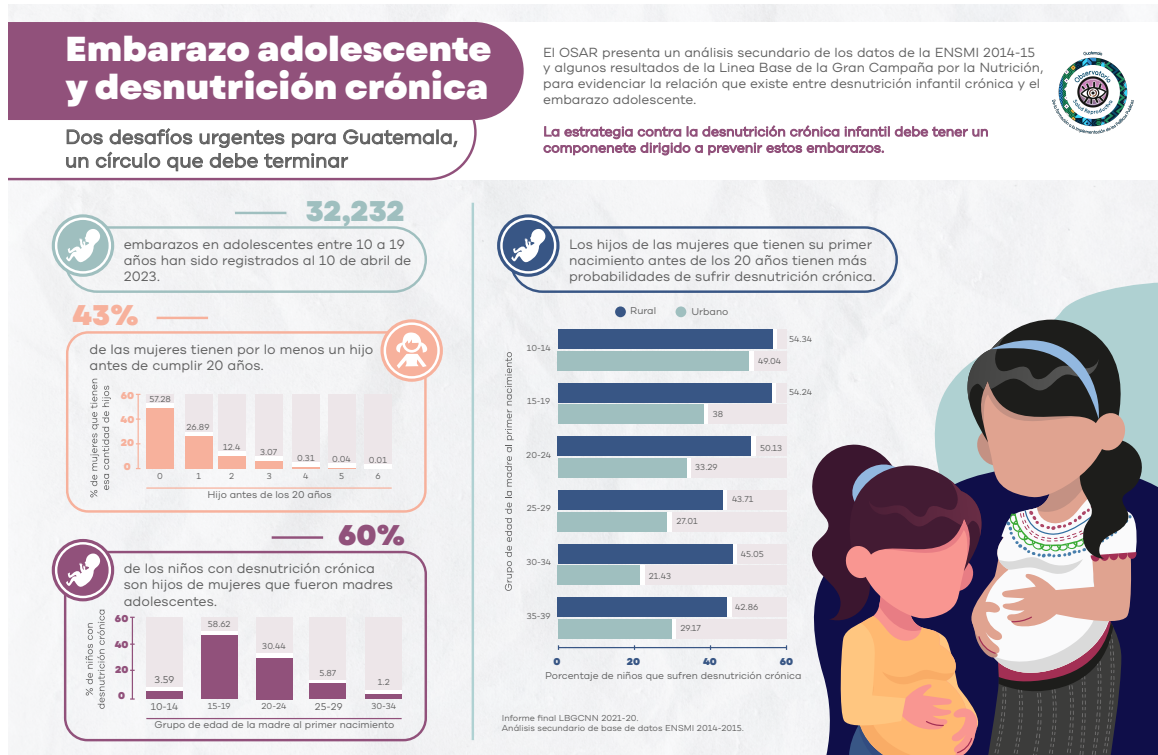
Objetivo: Definir acciones para el seguimiento 2023 de la atención prenatal en embarazadas adolescentes por parte del comité técnico OSAR-SESAN-MSPAS

Hora	Actividad	Responsable
7:30 – 7:40	Bienvenida y objetivo de la reunión	Rosa Elvira Pacheco, técnico de Aprovechamiento Biológico
7:40 – 8:00	Análisis secundario de desnutrición y embarazos según ENSMI 2014 - 2015.	Rodrigo Garzo de la Cruz/OSAR. Universidad Del Valle de Guatemala
8:00 – 8:20	Propuesta de dos Iniciativas para el seguimiento de la atención prenatal de embarazadas adolescentes en municipios priorizados	Mirna Montenegro/OSAR
8:20 – 8:40	Brigadas integrales institucionales	Marlon Figueroa/SIAS MSPAS
8:40 – 9:00	Datos de producción MSPAS relacionados a la atención nutricional de las embarazadas adolescentes	SIGSA
9:00 – 9:30	Plenaria: Plan piloto 2023	Todos
9:30 – 9:50	Acciones de incidencia 2023	Mirna Montenegro /OSAR
9:50 – 10:00	Cierre de la reunión	Karin Medrano/SESAN

Secretaría de Seguridad Alimentaria y Nutricional de la Presidencia de la República - SESAN -
8a. Avenida 13-06 zona 1, Guatemala, Guatemala
Teléfono: (+502) 2411-1900

Escaneado con CamScanner

12.3. Infografía creada



12.4. Matrices de confusión

\	No	Sí
No	138	44
Sí	32	85

\	No	Sí
No	127	58
Sí	43	70

\	No	Sí
No	136	44
Sí	35	85

\	No	Sí
No	134	52
Sí	36	77

\	No	Sí
No	134	46
Sí	36	83

\	No	Sí
No	123	38
Sí	47	91

\	No	Sí
No	130	46
Sí	40	83

12.5. Código de tablero

```
library(shiny)
library(ggplot2)
library(dplyr)
library(tidyr)
library(cluster)
library(factoextra)
library(bslib)
library(ggiraph)
library("FactoMineR")
library("factoextra")
library(pscl)
library(caret)
library(corrplot)
library(shinycssloaders)
library(data.table)

general <- fread(file="general.csv")
df3 <- fread(file="df3.csv")
line_test_no_na_numeric <- fread(file="line_test_no_na_numeric.csv
")
#str(line_mal_test_no_na_numeric)
ir_by_age_V190 <- fread(file="ir_by_age_V190.csv")
ir_by_age_V106 <- fread(file="ir_by_age_V106.csv")
ir_by_age_V025 <- fread(file="ir_by_age_V025.csv")
ir_by_age_general <- fread(file="ir_by_age_general.csv")

teen_children_V190 <- fread(file="teen_children_V190.csv")
teen_children_V106 <- fread(file="teen_children_V106.csv")
teen_children_V025 <- fread(file="teen_children_V025.csv")
teen_children <- fread(file="teen_children.csv")

percentage_V190 <- fread(file="percentage_V190.csv")
percentage_V106 <- fread(file="percentage_V106.csv")
percentage_V025 <- fread(file="percentage_V025.csv")
percentage <- fread(file="percentage.csv")

perc_of_moderate_V190 <- fread(file="perc_of_moderate_V190.csv")
perc_of_moderate_V106 <- fread(file="perc_of_moderate_V106.csv")
perc_of_moderate_V025 <- fread(file="perc_of_moderate_V025.csv")
perc_of_moderate <- fread(file="perc_of_moderate.csv")

moderate_percentage_first_mal_V190 <- fread(file="
  moderate_percentage_first_mal_V190.csv")
moderate_percentage_first_mal_V106 <- fread(file="
  moderate_percentage_first_mal_V106.csv")
```

```

moderate_percentage_first_mal_V025 <- fread(file="
  moderate_percentage_first_mal_V025.csv")
moderate_percentage_first_mal <- fread(file="
  moderate_percentage_first_mal.csv")

chi_df_v313 <- fread(file="chi_df_v313.csv")
test_v313 <- chisq.test(table(chi_df_v313$cat, chi_df_v313$V313))

chi_df_v106 <- fread(file="chi_df_v106.csv")
test_v106 <- chisq.test(table(chi_df_v106$cat, chi_df_v106$V106))

chi_df <- fread(file="chi_df.csv")
test3 <- chisq.test(table(chi_df$cat, chi_df$mother_cat))

chi_df_b11 <- fread(file="chi_df_b11.csv")
test_b11 <- chisq.test(table(chi_df_b11$cat, chi_df_b11$B11_cat))

ui <- fluidPage(
  theme = bs_theme(version = 4, bootswatch = "yeti"),
  titlePanel("Relaci n entre edad de la madre y desnutrici n
    cr nica en Guatemala"),
  tabsetPanel(
    type="tabs",
    tabPanel(
      "Catacter sticas generales",
      br(),
      fluidRow(
        column(6,
          wellPanel(
            h1("Guatemala 2015"),
            h2("Poblaci n total: 15,567,419"),
            h2("Adolescentes: 1,764,480 (11%)"),
            h2("Menores de 5 a os: 1,868,690 (12%)")
          )
        ),
        column(6,
          wellPanel(
            h1("Guatemala 2022"),
            h2("Poblaci n total: 17,357,886"),
            h2("Adolescentes: 1,747,592 (10%)"),
            h2("Menores de 5 a os: 1,870,482 (11%)")
          )
        )
      ),
      br(),
      fluidRow(
        column(12,
          plotOutput("chronic_desnutrition_comparison") %>%
            withSpinner()
        )
      ),

```

```

br(),
fluidRow(
  column(12,
    plotOutput("chronic_desnutrition_comparison_by_age")
  ),
),
br(),
fluidRow(
  column(12,
    selectInput("select_kids_before_age", h3("
      Seleccionar categoria"),
      choices = c("Quintil de riqueza"="V190",
        "Mximo nivel de educacion alcanzado"="V106",
        "Lugar de residencia"="V025",
        "General"="General"),
      selected = "GGeneral")),
    column(12,
      plotOutput("kids_before_age")
    ),
),
br(),
fluidRow(
  column(12,
    plotOutput("chi_pf")
  ),
),
br(),
fluidRow(
  column(12,
    plotOutput("chi_education")
  )
),
),
tabPanel(
  "Edad de la madre al primer nacimiento",
  br(),
  column(12,
    selectInput("select_total_children_first", h3("
      Seleccionar categoria"),
      choices = c("Quintil de riqueza"="V190",
        "Mximo nivel de educacion alcanzado"="V106",
        "Lugar de residencia"="V025",
        "General"="General"),
      selected = "V190")),
  br(),
  column(12,

```

```

        plotOutput("total_children_first")
    ),
    br(),
    column(12,
        selectInput("select_percentage_of_malnutrition", h3("
            Seleccionar categoria"),
            choices = c("Quintil de riqueza"="V190",
                "Mximo nivel de educacion
                alcanzado"="V106",
                "Lugar de residencia"="V025",
                "General"="General"),
            selected = "V190")),
    br(),
    column(12,
        plotOutput("percentage_of_malnutrition")
    ),
    br(),
    column(12,
        selectInput("select_malnutrition_percentage_only", h3(
            "Seleccionar categoria"),
            choices = c("Quintil de riqueza"="V190",
                "Mximo nivel de educacion
                alcanzado"="V106",
                "Lugar de residencia"="V025",
                "General"="General"),
            selected = "V190")),
    br(),
    column(12,
        plotOutput("malnutrition_percentage_only")
    ),
    br(),
    column(12,
        selectInput("select_malnutrition_percentage_all", h3(
            "Seleccionar categoria"),
            choices = c("Quintil de riqueza"="V190",
                "Mximo nivel de educacion
                alcanzado"="V106",
                "Lugar de residencia"="V025",
                "General"="General"),
            selected = "V190")),
    br(),
    fluidRow(
        column(12,
            plotOutput("malnutrition_percentage_all")
        ),
    ),
    br(),
    column(12,
        plotOutput("chi_teen_mother")
    )

```

```

    ),
  ),
  tabPanel(
    "Número de hijos",
    br(),
    column(12,
      selectInput("select_malnutrition_by_birth_order", h3
        ("Seleccionar categoría"),
        choices = c("Quintil de riqueza"="V190",
          "Lugar de residencia"="V025",
          "General"="General"),
        selected = "V190")),
    br(),
    column(12,
      plotOutput("malnutrition_by_birth_order")
    ),
    br(),
    column(12,
      plotOutput("chi_spacing"))
  )
)
)

#write.csv(general, "general.csv")
#write.csv(df3, "df3.csv")
#write.csv(line_test_no_na_numeric, "line_test_no_na_numeric.csv")
#write.csv(ir_by_age_V190, "ir_by_age_V190.csv")
#write.csv(ir_by_age_V106, "ir_by_age_V106.csv")
#write.csv(ir_by_age_V025, "ir_by_age_V025.csv")
#write.csv(ir_by_age_general, "ir_by_age_general.csv")

#write.csv(teen_children_V190, "teen_children_V190.csv")
#write.csv(teen_children_V106, "teen_children_V106.csv")
#write.csv(teen_children_V025, "teen_children_V025.csv")
#write.csv(teen_children, "teen_children.csv")

#write.csv(percentage_V190, "percentage_V190.csv")
#write.csv(percentage_V106, "percentage_V106.csv")
#write.csv(percentage_V025, "percentage_V025.csv")
#write.csv(percentage, "percentage.csv")

#write.csv(perc_of_moderate_V190, "perc_of_moderate_V190.csv")
#write.csv(perc_of_moderate_V106, "perc_of_moderate_V106.csv")
#write.csv(perc_of_moderate_V025, "perc_of_moderate_V025.csv")
#write.csv(perc_of_moderate, "perc_of_moderate.csv")

#write.csv(moderate_percentage_first_mal_V190, "
  moderate_percentage_first_mal_V190.csv")

```

```

#write.csv(moderate_percentage_first_mal_V106, "
  moderate_percentage_first_mal_V106.csv")
#write.csv(moderate_percentage_first_mal_V025, "
  moderate_percentage_first_mal_V025.csv")
#write.csv(moderate_percentage_first_mal, "
  moderate_percentage_first_mal.csv")

#write.csv(chi_df_v313, "chi_df_v313.csv")
#write.csv(chi_df_v106, "chi_df_v106.csv")
#write.csv(chi_df, "chi_df.csv")
#write.csv(chi_df_b11, "chi_df_b11.csv")

server <- function(input, output) {
  # No aplica
  output$chronic_desnutrition_comparison <- renderPlot({
    ggplot(general, aes(x=malnutrition_source, y=
      malnutrition_ensmi, group=malnutrition_type, color=
      malnutrition_type)) +
    geom_point() +
    geom_line() +
    geom_text(label=general$malnutrition_ensmi, nudge_y=-0.9,
      check_overlap = T) +
    labs(title="La desnutrici n cr nica severa, en menores de
      5 a os, aument entre 2015 y 2022", x="Encuesta", y="
      Porcentaje de desnutrici n cr nica", color="Nivel de
      desnutrici n")
  })
  # No aplica
  output$chronic_desnutrition_comparison_by_age <- renderPlot({
    ggplot(df3, aes(x=type, y=value, group=age_groups, color=
      age_groups)) +
    geom_point() +
    geom_line() +
    geom_text(label=df3$value, nudge_y=-0.75, check_overlap = T)
    +
    labs(title="La desnutrici n cr nica en menores de 3 a os
      ha empeorado desde el 2015", x="Encuesta", y="Porcentaje
      de desnutrici n cr nica, por grupo de edad", color="
      Grupos de edad")
  })
  # Listo
  output$malnutrition_by_birth_order <- renderPlot({
    ggplot(line_test_no_na_numeric[line_test_no_na_numeric$BORD <
      11,], aes(x=factor(number_cat), fill=factor(cat))) +
    geom_bar(position="fill") +
    facet_grid(
      switch(input$select_malnutrition_by_birth_order,
        "V190"=~V190, "V106"=~V106, "V025"=~V025)) +
    scale_x_discrete(labels=c("1 o 2", "3 o 4", "5 o 6", "7 u

```



```

      8", "9 o 10", "11 o 12", "13 o 14", "15 o 16")) +
labs(title="Entre m s hijos tiene una mujer, m s probable
      es que sufran de desnutrici n cr nica", x="Grupo de
      orden de nacimiento", y="Porcentaje de ni os que sufren
      de desnutrici n cr nica", fill="Desnutrici n") +
scale_y_continuous(labels=function(x) x*100)
})
# Revisar
output$total_children_first <- renderPlot({
  switch (input$select_total_children_first,
    "V190" = ggplot(ir_by_age_V190, aes(x=V212, y=
      total_births, color=V190)) +
      geom_point() +
      stat_smooth(method=lm, formula=y ~ log(x)) +
      geom_text(label=round(ir_by_age_V190$total_births,
        2), nudge_y = 0.25) +
      labs(title="Edad de primer nacimiento e hijos
        totales", x="Edad de la madre al primer
        nacimiento", y="Total de nacimientos", color="
        Quintil de riqueza"),
    "V106" = ggplot(ir_by_age_V106, aes(x=V212, y=
      total_births, fill=V106, color=V106)) +
      geom_point() +
      stat_smooth(method=lm, formula=y ~ log(x)) +
      geom_text(label=round(ir_by_age_V106$total_births,
        2), nudge_y = 0.25) +
      labs(title="Edad de primer nacimiento e hijos
        totales", x="Edad a primer nacimiento", y="Total
        de nacimientos", color="Nivel de educaci n
        m ximo"),
    "V025" = ggplot(ir_by_age_V025, aes(x=V212, y=
      total_births, fill=V025, color=V025)) +
      geom_point() +
      stat_smooth(method=lm, formula=y ~ log(x)) +
      geom_text(label=round(ir_by_age_V025$total_births,
        2), nudge_y = 0.25) +
      labs(title="Edad de primer nacimiento e hijos
        totales", x="Edad a primer nacimiento", y="Total
        de nacimientos", color=" rea de residencia"),
    "General" = ggplot(ir_by_age_general, aes(x=V212, y=
      total_births)) +
      geom_point() +
      stat_smooth(method=lm, formula=y ~ log(x)) +
      geom_text(label=round(ir_by_age_general$total_births
        , 2), nudge_y = 0.25) +

```

```

        labs(title="Edad de primer nacimiento e hijos
              totales", x="Edad a primer nacimiento", y="Total
              de nacimientos")
      )
    })
# Revisar
output$percentage_of_malnutrition <- renderPlot({
  switch (input$select_percentage_of_malnutrition,
    "V190" = ggplot(teen_children_V190, aes(x=
      was_teen_mother, y=percentage, fill=V190)) +
      geom_col(position="dodge") +
      labs(title="Las mujeres que inician la maternidad en
              la adolescencia, tienen en promedio un hijo m s
              ", x="Fue madre adolescente", y="N mero de hijos
              en promedio", fill="Quintil de riqueza") +
      geom_text(label=round(teen_children_V190$percentage,
        2), position = position_dodge(width=.9)),

    "V106" = ggplot(teen_children_V106, aes(x=
      was_teen_mother, y=percentage, fill=V106)) +
      geom_col(position="dodge") +
      labs(title="Las mujeres que inician la maternidad en
              la adolescencia, tienen en promedio un hijo m s
              ", x="Fue madre adolescente", y="N mero de hijos
              en promedio", fill="Nivel de educaci n m ximo
              ") +
      geom_text(label=round(teen_children_V106$percentage,
        2), position = position_dodge(width=.9)),

    "V025" = ggplot(teen_children_V025, aes(x=
      was_teen_mother, y=percentage, fill=V025)) +
      geom_col(position="dodge") +
      labs(title="Las mujeres que inician la maternidad en
              la adolescencia, tienen en promedio un hijo m s
              ", x="Fue madre adolescente", y="N mero de hijos
              en promedio", fill=" rea de residencia") +
      geom_text(label=round(teen_children_V025$percentage,
        2), position = position_dodge(width=.9)),

    "General" = ggplot(teen_children, aes(x=
      was_teen_mother, y=percentage)) +
      geom_col(position="dodge") +
      labs(title="Las mujeres que inician la maternidad en
              la adolescencia, tienen en promedio un hijo m s
              ", x="Fue madre adolescente", y="N mero de hijos
              en promedio") +
      geom_text(label=round(teen_children$percentage, 2),
        position = position_dodge(width=.9)),
  )
})

```

```

    )
  })
# Listo
output$kids_before_age <- renderPlot({
  switch (input$select_kids_before_age ,
    "V190" = ggplot(head(percentage_V190 , n=20), aes(x=
      factor(total), y=percentage , fill=V190)) +
      geom_col(position="dodge") +
      #geom_col() +
      labs(title="43% de las mujeres tienen por lo menos
        un hijo antes de cumplir 20 a os ", x="Hijos
        antes de los 20 a os ", y="Porcentaje de mujeres
        que tienen esa cantidad de hijos ", fill="Quintil
        de riqueza ") +
      geom_text(label=round(head(percentage_V190 , n=20)
        $percentage , 2), position = position_dodge(width
        =.9)) ,

    "V106" = ggplot(head(percentage_V106 , n=16), aes(x=
      factor(total), y=percentage , fill=V106)) +
      geom_col(position="dodge") +
      #geom_col() +
      labs(title="43% de las mujeres tienen por lo menos
        un hijo antes de cumplir 20 a os ", x="Hijos
        antes de los 20 a os ", y="Porcentaje de mujeres
        que tienen esa cantidad de hijos ", fill="Nivel de
        educaci n m ximo ") +
      geom_text(label=round(head(percentage_V106 , n=16)
        $percentage , 2), position = position_dodge(width
        =.9)) ,

    "V025" = ggplot(head(percentage_V025 , n=8), aes(x=
      factor(total), y=percentage , fill=V025)) +
      geom_col(position="dodge") +
      #geom_col() +
      labs(title="43% de las mujeres tienen por lo menos
        un hijo antes de cumplir 20 a os ", x="Hijos
        antes de los 20 a os ", y="Porcentaje de mujeres
        que tienen esa cantidad de hijos ", fill=" rea de
        residencia ") +
      geom_text(label=round(head(percentage_V025 , n=8)
        $percentage , 2), position = position_dodge(width
        =.9)) ,

    "General" = ggplot(head(percentage , n=4), aes(x=factor
      (total), y=percentage)) +
      geom_col(position="dodge") +
      #geom_col() +

```

```

labs(title="43% de las mujeres tienen por lo menos
un hijo antes de cumplir 20 años", x="Hijos
antes de los 20 años", y="Porcentaje de mujeres
que tienen esa cantidad de hijos", fill="
Categorías") +
geom_text(label=round(head(percentage, n=4)
$percentage, 2), position = position_dodge(width
=.9)),
)

#fill=input$select_kids_before_age

})
# Revisar
output$malnutrition_percentage_only <- renderPlot({
switch (input$select_malnutrition_percentage_only,
"V190" = ggplot(head(perc_of_moderate_V190, 25), aes(
first_child_cat, percentage, fill = V190)) +
geom_col(position="dodge") +
scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
44", "45 a 49")) +
labs(title="Casi 60% de los niños con desnutrición
crónica son hijos de mujeres que son o fueron
madres adolescentes", x="Grupo de edad de la
madre al primer nacimiento", y="Porcentaje de
niños desnutrición crónica", fill="Quintil de
riqueza") +
geom_text(label=round(head(perc_of_moderate_V190,
25)$percentage*100, 2), position = position_dodge
(width=.9)) +
scale_y_continuous(labels=function(x) x*100),

"V106" = ggplot(head(perc_of_moderate_V106, 20), aes(
first_child_cat, percentage, fill = V106)) +
geom_col(position="dodge") +
scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
44", "45 a 49")) +
labs(title="Casi 60% de los niños con desnutrición
crónica son hijos de mujeres que son o fueron
madres adolescentes", x="Grupo de edad de la
madre al primer nacimiento", y="Porcentaje de
niños desnutrición crónica", fill="Nivel de
educación máximo") +
geom_text(label=round(head(perc_of_moderate_V106,
20)$percentage*100, 2), position = position_dodge
(width=.9)) +

```

```

    scale_y_continuous(labels=function(x) x*100),

"V025" = ggplot(head(perc_of_moderate_V025, 8), aes(
  first_child_cat, percentage, fill = V025)) +
  geom_col(position="dodge") +
  scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
    a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
    44", "45 a 49")) +
  labs(title="Casi 60% de los ni os con desnutrici n
    cr nica son hijos de mujeres que son o fueron
    madres adolescentes", x="Grupo de edad de la
    madre al primer nacimiento", y="Porcentaje de
    ni os desnutrici n cr nica", fill=" rea de
    residencia") +
  geom_text(label=round(head(perc_of_moderate_V025, 8)
    $percentage*100, 2), position = position_dodge(
    width=.9)) +
  scale_y_continuous(labels=function(x) x*100),

"General" = ggplot(perc_of_moderate[c(1, 2, 3, 4, 5)
  ,], aes(first_child_cat, percentage)) +
  geom_col(position="dodge") +
  scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
    a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
    44", "45 a 49")) +
  labs(title="Casi 60% de los ni os con desnutrici n
    cr nica son hijos de mujeres que son o fueron
    madres adolescentes", x="Grupo de edad de la
    madre al primer nacimiento", y="Porcentaje de
    ni os desnutrici n cr nica") +
  geom_text(label=round(perc_of_moderate[c(1, 2, 3, 4,
    5),]$percentage*100, 2), position =
    position_dodge(width=.9)) +
  scale_y_continuous(labels=function(x) x*100)
)
})
# Listo
output$malnutrition_percentage_all <- renderPlot({
  switch (input$select_malnutrition_percentage_all,
    "V190" = ggplot(head(
      moderate_percentage_first_mal_V190, 25), aes(x=
        first_child_cat, moderate, fill=V190)) +
      geom_col(position="dodge") +
      scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
        a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
        44", "45 a 49")) +
      labs(title="Los hijos de las mujeres que tienen su
        primer nacimiento antes de los 20 a os tienen

```

```

    m s probabilidades de sufrir desnutrici n
    cr nica", x="Grupo de edad de la madre al primer
    nacimiento", y="Porcentaje de ni os que sufren
    de desnutrici n cr nica", fill="Quintil de
    riqueza") +
    geom_text(label=round(head(
    moderate_percentage_first_mal_V190, 25)$moderate
    *100, 2), position = position_dodge(width=.9)) +
    scale_y_continuous(labels=function(x) x*100),

"V106" = ggplot(head(
    moderate_percentage_first_mal_V106, 20), aes(x=
    first_child_cat, moderate, fill=V106)) +
    geom_col(position="dodge") +
    scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
    a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
    44", "45 a 49")) +
    labs(title="Los hijos de las mujeres que tienen su
    primer nacimiento antes de los 20 a os tienen
    m s probabilidades de sufrir desnutrici n
    cr nica", x="Grupo de edad de la madre al primer
    nacimiento", y="Porcentaje de ni os que sufren
    de desnutrici n cr nica", fill="Nivel de
    educaci n m ximo") +
    geom_text(label=round(head(
    moderate_percentage_first_mal_V106, 20)$moderate
    *100, 2), position = position_dodge(width=.9)) +
    scale_y_continuous(labels=function(x) x*100),

"V025" = ggplot(head(
    moderate_percentage_first_mal_V025, 10), aes(x=
    first_child_cat, moderate, fill=V025)) +
    geom_col(position="dodge") +
    scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
    a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
    44", "45 a 49")) +
    labs(title="Los hijos de las mujeres que tienen su
    primer nacimiento antes de los 20 a os tienen
    m s probabilidades de sufrir desnutrici n
    cr nica", x="Grupo de edad de la madre al primer
    nacimiento", y="Porcentaje de ni os que sufren
    de desnutrici n cr nica", fill=" rea de
    residencia") +
    geom_text(label=round(head(
    moderate_percentage_first_mal_V025, 10)$moderate
    *100, 2), position = position_dodge(width=.9)) +
    scale_y_continuous(labels=function(x) x*100),

```

```

"General" = ggplot(head(moderate_percentage_first_mal,
  5), aes(x=first_child_cat, moderate)) +
  geom_col() +
  scale_x_discrete(labels=c("10 a 14", "15 a 19", "20
  a 24", "25 a 29", "30 a 34", "35 a 39", "40 a
  44", "45 a 49")) +
  labs(title="Los hijos de las mujeres que tienen su
  primer nacimiento antes de los 20 años tienen
  más probabilidades de sufrir desnutrición
  crónica", x="Grupo de edad de la madre al primer
  nacimiento", y="Porcentaje de niños que sufren
  de desnutrición crónica") +
  geom_text(label=round(head(
  moderate_percentage_first_mal, 5)$moderate*100,
  2), position = position_dodge(width=.9)) +
  scale_y_continuous(labels=function(x) x*100)
)
})
output$chi_pf <- renderPlot({
  corrplot(test_v313$residuals, is.corr = FALSE, main="El uso de
  todos los anticonceptivos modernos se asocia con un menor
  nivel de desnutrición crónica", sub="Fuente: Análisis
  secundario ENSMI 2014–2015. Guatemala, 2023", mar=c
  (5,1,1,1))
})
output$chi_education <- renderPlot({
  corrplot(test_v106$residuals, is.corr = FALSE, main="Alcanzar
  el nivel de educación secundaria se asocia con un menor
  nivel de desnutrición crónica", sub="Fuente: Análisis
  secundario ENSMI 2014–2015. Guatemala, 2023", mar=c
  (5,1,1,1))
})
output$chi_teen_mother <- renderPlot({
  corrplot(test3$residuals, is.corr = FALSE, main="Ser madre
  adolescente se asocia con un mayor nivel de desnutrición
  crónica", sub="Fuente: Análisis secundario ENSMI
  2014–2015. Guatemala, 2023", mar=c(5,1,1,1))
})
output$chi_spacing <- renderPlot({
  corrplot(test_b11$residuals, is.corr = FALSE, main="Un mayor
  periodo entre nacimientos se asocia con un menor nivel de
  desnutrición", sub="Fuente: Análisis secundario ENSMI
  2014–2015. Guatemala, 2023", mar=c(5,1,1,1))
})
}

shinyApp(ui = ui, server = server)

```


Dataframe: Estructura de datos en el paquete de Python Pandas que contiene ejes nombrados. Se puede pensar como un contenedor parecido a un diccionario para una serie. 26

Factor Analysis of Mixed Data: Método para reducir la dimensión de variables cuantitativas y cualitativas. Utiliza PCA para las variables cuantitativas y MCA para variables cualitativas. 34

Logs: Registros creados de los eventos ocurridos en un programa de computación. Puede contener información sobre errores. 36

Rollback: Proceso en el que se regresa un archivo a una versión anterior. Generalmente usado para corregir errores inducidos a un archivo previamente funcional. 38

Root: En sistemas UNIX, usuario que tiene todos los permisos en un sistema de archivos. Puede modificar todos los archivos y utilizar cualquier ejecutable. 36

Sudo: Comando en UNIX que permite ejecutar comandos como cualquier usuario. Generalmente se utiliza para ejecutar comandos como el usuario root. 36

Valor p: La probabilidad de que un valor estadístico calculado sea posible una hipótesis nula. Ayuda a diferenciar resultados producto del azar de efectos del muestro. 46

Windows Subsystem for Linux: Funcionalidad en sistemas Windows que permite ejecutar ambientes Linux sin necesidad de la creación de una máquina virtual. 26