



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Informática

Estudio del índice de anemia en Perú en niños y niñas de 6 a 36 meses, mediante la evaluación de factores demográficos, económicos, formativos y de salud

Trabajo Fin de Máster

Máster Universitario en Gestión de la Información

AUTOR/A: Mendez Cortez, Katia

Tutor/a: Calduch Losa, Maria de los Angeles

Cotutor/a: Serrano Cobos, Jorge Ignacio

CURSO ACADÉMICO: 2021/2022

Agradecimientos

A mi familia, por apoyar mis decisiones y ser mi soporte siempre.

A Ángeles, por las ideas, el seguimiento, la motivación y el apoyo constante.

Al Gobierno del Perú que con su Programa Nacional de Becas y Crédito Educativo (PRONABEC), me otorgó una beca para poder realizar mis estudios de Máster y realizar el desarrollo del presente Trabajo Fin de Máster (TFM).

Resumen

A través de los años, un ítem que se repite en la agenda del estado peruano es el planteamiento de planes de acción para hacer frente al alto índice de anemia. Hasta el año 2020 se registraba un valor igual o superior al 40%, lo cual ponía a Perú como un país con un problema severo de salud pública. Esta situación, aunada a la escasa información estructurada que se encuentra en las fuentes oficiales, no permiten tener buena visibilidad sobre la problemática, ni cómo se están desplegando los planes de acción a lo largo de los años.

El presente trabajo pretende aportar un análisis inicial de los datos recabados de diversas fuentes de información y del mismo modo, poner a disposición del público en general, una plataforma de acceso abierto con un repositorio de datos y visualizaciones obtenidas, que sirvan de input para investigaciones futuras o análisis de los entes interesados en este tema.

La metodología utilizada incluye como procesos principales, la selección de variables que se quieren analizar; la recolección, limpieza, consolidación y validación de datos; la elaboración de visualizaciones; el análisis de los resultados obtenidos y finalmente, la publicación y habilitación en datos abiertos.

Las principales herramientas utilizadas han sido la web de accesos de datos abiertos de las entidades gubernamentales y tecnologías, como Python, para el procesamiento de datos; HTML/CSS, para el diseño de la web y Tableau, para la visualización de la información.

A partir de esta herramienta, se busca realizar ciclos de mejora continua para generar mayores aportaciones y que la utilidad e influencia sea cada vez más fuerte, ya que perspectivas distintas siempre sumarán para lograr mejores resultados.

Palabras clave: Datos abiertos, Gestión de la información, Índice de anemia, Análisis de datos, Procesos ETL, Anemia niños.

Abstract

Over the years, a recurring item on the schedule of the Peruvian Government has been the proposal of action plans to deal with the high rate of anemia. Up to the year 2020, a value equal or greater than 40% was recorded, making Peru a country with a severe public health problem. This situation, along with the scarce structured information found in official sources, does not allow to have a good visibility of the problem, neither how the action plans are being deployed over the years.

This paper aims to provide a first analysis of the data obtained from various sources of information and to allow the public to have a free access platform with a data repository and visualizations obtained, which serve as input for future research or analysis of the interested entities in this topic.

The method used includes as main processes the selection of variables to be analyzed; the collection, cleaning, consolidation, and validation of data; as well as the elaboration of visualizations; the analysis of the results obtained and finally, the publication and open data release.

The main tools used have been the open data access web of government entities and technologies such as Python, for data processing; HTML/CSS, for web design and Tableau, for information visualization.

From this tool, continuous improvement cycles will be conducted to generate greater contributions and make its utility and influence stronger and stronger, since different perspectives will always add to achieve better results.

Key words: Open data, Information management, Anemia rate, Data Analysis, ETL Processes, Anemia in children.

Resum

A través dels anys, un ítem que es repeteix en l'agenda de l'estat peruà és el plantejament de plans d'acció per a fer front a l'elevat índex d'anèmia. Fins a l'any 2020 es registrava un valor igual o superior al 40%, la qual cosa ens posava com un país amb un problema sever de salut pública. Aquesta situació, conjuntament amb l'escassa informació estructurada que es troba en les fonts oficials, no ens permet tenir bona visibilitat sobre la problemàtica, ni com s'estan desplegant els plans d'acció al llarg dels anys. El present treball pretén aportar una anàlisi inicial de les dades obtingudes de diverses fonts d'informació i de la mateixa manera, posar a la disposició del públic en general, una plataforma d'accés obert amb un repositori de dades i visualitzacions obtingudes, que servisquen d'inici per a investigacions futures o anàlisis de les persones interessades en aquest tema.

La metodologia utilitzada inclou com a processos principals, la selecció de variables que es volen analitzar; la recollida, neteja, consolidació i validació de dades; l'elaboració de visualitzacions; l'anàlisi dels resultats obtinguts i finalment, la publicació i habilitació en dades obertes.

Les principals eines utilitzades han sigut, per una banda, la web d'accés de dades obertes de les entitats governamentals i, per altra, tecnologies com Python per al processament de dades, HTML/CSS per al disseny de la web i Tableau, per a la visualització de la informació.

A partir d'aquesta eina, se cerca tenir cicles de millora contínua per a generar majors aportacions i que la utilitat i influència siga cada vegada més forta, ja que perspectives diferents sempre sumaran per a aconseguir millors resultats.

Paraules clau: Dades obertes, Gestió de la Informació, Índex d'anèmia, Anàlisi de dades, Processos ETL, Anèmia xiquets.

Tabla de contenido

Agradecimientos	3
Resumen	4
Abstract	5
Resum	6
Tabla de contenido	7
Lista de ilustraciones	10
Lista de tablas	12
Acrónimos	13
CAPÍTULO I: Introducción	14
1.1 Motivación	14
1.2 Objetivos	15
1.2.1. Objetivo principal.....	15
1.2.2. Objetivos específicos	15
1.3 Impacto esperado	16
1.4 Metodología	16
1.4.1. Selección de variables	17
1.4.2. Identificación de fuentes de datos.....	17
1.4.3. Selección de las fuentes óptimas	17
1.4.4. Recopilación de datos	17
1.4.5. Limpieza y consolidación de datos	18
1.4.6. Validación de datos	18
1.4.7. Carga de fuentes adicionales.....	19
1.4.8. Elaboración de visualizaciones de datos	19
1.4.9. Publicación y habilitación en datos abiertos.....	19
1.5 Estructura del trabajo	20
1.6 Convenciones	22
CAPÍTULO II: Estado del arte	23
2.1 Marco de referencia	23
2.1.1. Anemia.....	23
2.1.2. Instituto Nacional de Estadística e Informática (INEI):.....	24
2.1.3. Evolución de la anemia infantil en Perú	25
2.1.4. Datos abiertos en Perú	26
2.1.5. El derecho de acceso a la información en Perú.....	28
2.2 Estudios relacionados	28

2.3	Tecnología.....	32
2.3.1.	Observatorio de Anemia	32
2.3.2.	Observatorios regionales.....	36
2.3.3.	Plataforma Nacional de Datos Abiertos	37
2.4	Propuesta y contribución de la investigación	37
CAPÍTULO III: Análisis del problema		39
3.1	Análisis del marco legal	41
3.2	Análisis de riesgos	41
3.3	Identificación y análisis de soluciones alternativas	42
3.4	Solución propuesta.....	42
CAPÍTULO IV: Diseño de la solución		43
4.1	Planificación de las actividades:	43
4.2	Arquitectura del sistema.....	44
4.3	Diseño del proceso	44
4.4	Tecnología utilizada.....	45
CAPÍTULO V: Desarrollo de la solución		48
5.1	Selección de variables.....	48
5.2	Identificación de fuentes de datos	49
5.3	Selección de las fuentes óptimas.....	49
5.4	Recopilación de datos.....	52
5.5	Limpieza y consolidación de datos	55
5.5.1.	Quick wins sobre la data total	55
5.5.2.	Limpieza sobre la base seleccionada.....	55
5.5.3.	Consolidación de la base seleccionada	58
5.5.4.	Creación de campos auxiliares o complementarios.....	59
5.6	Validación de datos	60
5.7	Carga de fuentes adicionales.....	61
5.7.1.	Información relacionada a los objetivos	61
5.7.2.	Información de RENAMU	62
5.7.3.	Geolocalización de las provincias y distritos	63
5.7.4.	Información del gasto devengado anual	64
5.8	Elaboración de visualizaciones de datos	65
5.9	Publicación y habilitación en datos abiertos	67
5.9.1.	Versión inicial del código en Git	67
5.9.2.	Carga de archivos en GitHub	68
5.9.3.	Publicación en un entorno Web.....	69
CAPÍTULO VI: Pruebas y resultados		72
6.1	Pruebas de funcionamiento de la Web	72

6.1.1. Pruebas de adaptabilidad.....	72
6.1.2. Pruebas funcionales	73
6.1.3. Auditoría SEO.....	76
6.2 Análisis de indicadores	77
6.2.1. Análisis Nacional.....	77
6.2.2. Análisis de indicadores regionales	87
6.3 Principales dificultades y soluciones planteadas	91
Conclusiones.....	93
Relación del trabajo con los estudios cursados	95
Trabajos futuros	96
Referencias.....	97
ANEXOS.....	100
Anexo I: Relación con ODS	100
Anexo II: Dimensiones y niveles MELODA.....	104
Anexo III: Tablas y campos seleccionados	105
Anexo IV: Documentación sobre el factor de ponderación	111
Anexo V: Normatividad de compromisos en gobiernos regionales	113

Lista de ilustraciones

<i>Ilustración 1. Metodología de desarrollo</i>	20
<i>Ilustración 2. Prevalencia de anemia en niños de 6 a 35 meses (en porcentaje)</i>	25
<i>Ilustración 3. Open Data Barometer 2016</i>	26
<i>Ilustración 4. Open Data Barometer 2020</i>	27
<i>Ilustración 5. Comparativa de países</i>	27
<i>Ilustración 6. Observatorio de Anemia: Inicio</i>	32
<i>Ilustración 7. Observatorio de Anemia: ¿Cómo estamos?</i>	33
<i>Ilustración 8. Observatorio de Anemia: ¿Qué queremos?</i>	34
<i>Ilustración 9. Observatorio de Anemia: Intervenciones Sectoriales</i>	34
<i>Ilustración 10. Observatorio de Anemia: Intervenciones Regionales</i>	35
<i>Ilustración 11. Observatorio de Anemia: Indicadores Sectoriales</i>	35
<i>Ilustración 12. Observatorio de Anemia - Región Callao</i>	36
<i>Ilustración 13. Observatorio de Anemia - Región Áncash: Prevalencia de anemia</i>	36
<i>Ilustración 14. Plataforma Nacional de Datos Abiertos de Perú</i>	37
<i>Ilustración 15. Índice de anemia por meses</i>	38
<i>Ilustración 16. Índice de anemia real vs objetivo</i>	39
<i>Ilustración 17. Cronograma de actividades inicial</i>	43
<i>Ilustración 18. Arquitectura de la solución</i>	44
<i>Ilustración 19. Modelo de flujo de datos y principales herramientas</i>	44
<i>Ilustración 20. Archivos del INEI - ENDES 2021</i>	52
<i>Ilustración 21. Código para conexión y descarga de archivos</i>	53
<i>Ilustración 22. Código para la descomprimir y leer archivos DTAs y DBFs</i>	53
<i>Ilustración 23. Código para validaciones y carga en la base de datos</i>	54
<i>Ilustración 24. Carga inicial de base de datos</i>	55
<i>Ilustración 25. Vista de verificación de datos en tablas y años seleccionadas</i>	57
<i>Ilustración 26. Tabla consolidada para análisis del alcance objetivo (data_indiceanemia)</i>	58
<i>Ilustración 27. Verificación del índice de anemia calculado</i>	60
<i>Ilustración 28. Metas nacionales de anemia anualizadas 2017-2021</i>	61
<i>Ilustración 29. Metas regionales anualizadas 2017-2021</i>	61
<i>Ilustración 30. Conexión y descarga de data RENAMU</i>	62
<i>Ilustración 31. Tabla de departamentos, provincias y distritos</i>	63
<i>Ilustración 32. Código para obtener datos georreferenciados con la API de Google Maps</i>	63
<i>Ilustración 33. Consulta de Ejecución del Gasto</i>	64
<i>Ilustración 34. Uso de herramienta Postman para obtener datos</i>	65
<i>Ilustración 35. Cuadro de mando indicadores a nivel país</i>	66
<i>Ilustración 36. Números globales y comparativos con objetivo y devengado - Regional</i>	67

<i>Ilustración 37. Carga de versión inicial en Git</i>	<i>67</i>
<i>Ilustración 38. Archivos de programación cargados en GitHub</i>	<i>68</i>
<i>Ilustración 39. Página principal de la plataforma SANITO PE</i>	<i>69</i>
<i>Ilustración 40. Servicio de hosting - Pythonanywhere</i>	<i>70</i>
<i>Ilustración 41. SSL para el dominio</i>	<i>71</i>
<i>Ilustración 42. Página SanitoPe en una versión móvil</i>	<i>71</i>
<i>Ilustración 43. Prueba de adaptabilidad SanitoPe: cabecera y pie de página</i>	<i>72</i>
<i>Ilustración 44. Prueba de adaptabilidad SanitoPe: tableros de control</i>	<i>73</i>
<i>Ilustración 45. Prueba funcional de descarga de datos - Vista Datasets</i>	<i>73</i>
<i>Ilustración 46. Prueba funcional de descarga de datos - Verificación de Información</i>	<i>74</i>
<i>Ilustración 47. Prueba funcional de descarga de datos - Vista Instituciones</i>	<i>74</i>
<i>Ilustración 48. Prueba funcional de descarga de datos - Vista móviles</i>	<i>74</i>
<i>Ilustración 49. Prueba funcional de filtros en tableros de mando: país y región</i>	<i>75</i>
<i>Ilustración 50. Auditoría SEO a la Web SanitoPe: resultados iniciales</i>	<i>76</i>
<i>Ilustración 51. Auditoría SEO a la Web SanitoPe: resultados finales</i>	<i>76</i>
<i>Ilustración 52. Cuadro de mando indicadores globales</i>	<i>78</i>
<i>Ilustración 53. Cálculo de la cantidad de muestra para el 2009</i>	<i>79</i>
<i>Ilustración 54. Índice de anemia según aspectos demográficos y económicos (anualizado: 2009)</i>	<i>80</i>
<i>Ilustración 55. Índice de anemia según aspectos demográficos y económicos (anualizado: 2021)</i>	<i>81</i>
<i>Ilustración 56. Índice de anemia según aspectos formativos y de salud</i>	<i>83</i>
<i>Ilustración 57. Comparación con el índice objetivo y gasto devengado PAN</i>	<i>85</i>
<i>Ilustración 58. Consumo de hierro, anemia durante el embarazo y lactancia exclusiva</i>	<i>86</i>
<i>Ilustración 59. Números globales y comparativos con objetivo y devengado - Regional</i>	<i>88</i>
<i>Ilustración 60. Análisis de principales indicadores por región y año</i>	<i>89</i>
<i>Ilustración 61. Vista global por provincias y distritos</i>	<i>90</i>
<i>Ilustración 62. Dimensiones y niveles de evaluación MELODA</i>	<i>104</i>
<i>Ilustración 63. Factor de ponderación año 2015</i>	<i>111</i>
<i>Ilustración 64. Factor de ponderación año 2020</i>	<i>112</i>
<i>Ilustración 65. Factor de ponderación resto de años</i>	<i>112</i>
<i>Ilustración 66. Documentos normativos de compromisos asumidos por gobiernos regionales</i>	<i>113</i>

Lista de tablas

<i>Tabla 1. Análisis MELODA: Data INEI - ENDES.</i>	50
<i>Tabla 2. Análisis MELODA: Data INS - MINSA.</i>	50
<i>Tabla 3. Evaluación de cobertura de variables.</i>	51
<i>Tabla 4. Tabla de estructura de datos del INEI - ENDES.</i>	52
<i>Tabla 5. Ejemplo de registros para descarga de información de la ENDES - INEI</i>	53
<i>Tabla 6. Resumen de carga inicial en base de datos</i>	54
<i>Tabla 7. Quick wins sobre la data total</i>	55
<i>Tabla 8. Tablas y campos seleccionados</i>	56
<i>Tabla 9. Validación de data por tabla y por año</i>	56
<i>Tabla 10. Limpieza y homologación de datos</i>	57
<i>Tabla 11. Resumen de tabla consolidada "data_indiceanemia"</i>	58
<i>Tabla 12. Identificación del factor de ponderación</i>	59
<i>Tabla 13. Campos adicionales creados</i>	59
<i>Tabla 14. Relación del trabajo con los ODS.</i>	100

BD: Base de Datos.

ETL: por sus siglas en inglés Extract, Transform and Load. En español: extraer, transformar y cargar.

TFM: Trabajo Final de Máster.

ODS: Objetivos de Desarrollo Sostenible.

SGBD: Sistemas de Gestión de Bases de Datos.

WEB: abreviatura de World Wide Web.

DBF: por sus siglas: Data Base File. En español: Archivo de base de datos.

ENDES: ENCUESTA DEMOGRÁFICA Y DE SALUD FAMILIAR.

INEI: Instituto Nacional de Estadística e Informática.

MINSA: MINISTERIO DE SALUD.

MEF: Ministerio de Economía y Finanzas.

RENAMU: REGISTRO NACIONAL DE MUNICIPALIDADES.

MIDIS: MINISTERIO DE DESARROLLO E INCLUSIÓN SOCIAL.

PAN: Programa Articulado Nutricional.

MELODA: METRIC FOR THE EVALUATION OF OPEN DATA. En español: Métrica para la evaluación de datos abiertos.

OMS: Organización Mundial de la Salud.

CRED: Control de CRECIMIENTO Y DESARROLLO DE LA NIÑA Y EL NIÑO.

SEO: Search Engine Optimization. En español: optimización en motores de búsqueda.

API: Application Program Interface. En español: interfaz de programación de aplicaciones

CAPÍTULO I: Introducción

El índice de anemia en niños menores de 3 años en Perú es un tema de mucha preocupación. Hasta el año 2020, el valor de este índice fue igual o superior al 40% (*INEI 2021*), por lo que era considerado como un problema severo de salud pública¹. Además, se ha evidenciado cómo a lo largo de los últimos 10 años no se han presentado mejoras significativas en relación con este indicador, por el contrario, y a pesar de las políticas de estado aplicadas, hubo años en los que se presentaron índices crecientes.

En el año 2017, la meta planteada por el gobierno para el 2021 fue del 19% (*MIDIS 2017*); sin embargo, el índice real para ese mismo año fue 38,8%, siendo más del doble del objetivo planteado.

Sumado a esta situación, se tiene la dificultad de acceso e interpretación de datos abiertos que se tiene en Perú, esto a pesar de que periódicamente se realiza recolección de datos y estudios por las administraciones públicas. Esta situación hace que la vigilancia e investigaciones de la población sea escasa, ya que suele tomar como referencia solo los resultados presentados por las mismas instituciones, sin realizar mayor análisis, ni profundidad en las evaluaciones.

Sin duda, la realidad no es alentadora, pero a la vez, representa un desafío que promueve la búsqueda de estrategias que faciliten la identificación de las causas que hacen que esta situación persista y, asimismo, buscar las maneras de habilitar herramientas que permitan a las personas responsables e interesados en el tema tener mayor interacción y facilidad de análisis. Por ello, con el presente trabajo, se quiere aportar un granito de arena en reducir la brecha de acceso, facilitando datos y visualizaciones relacionados con el alto índice de anemia en Perú en un entorno de acceso abierto.

1.1 Motivación

Desde que inicié mi formación profesional, he pensado que mi país, no solo necesita profesionales destacados, sino a la par, necesita ciudadanos que ayuden a construir de manera sostenible las condiciones de equidad para nuestra población. Por ello, la idea de mi trabajo de fin de máster desde un inicio estuvo orientado a los principales problemas sociales identificados en mi país y que, de alguna manera, he vivido muy de cerca.

Inicialmente, los dos temas en los que estaba interesada eran, la violencia de género y la anemia infantil, sin embargo, al realizar un análisis preliminar de los datos y la evolución de la incidencia de ambos temas, pude notar la poca mejoría que se había tenido en relación con el segundo ítem, ya que en más de una década, y que a pesar que se han implementado diversos programas sociales

¹ La Organización Mundial de la Salud (2011) clasifica el nivel de anemia de acuerdo con el porcentaje de incidencia en: severa (mayor o igual a 40%), moderada (entre 20 y 39,9%), leve (entre 5 y 19,9%) y normal (menor o igual al 4,9%). (*Organización Mundial de la Salud 2011, p. 5*)

y proyectos de inversión nacionales y regionales con el objetivo de reducir este índice, se ha evidenciado que no se ha logrado reducirlo de manera consistente, y por ende, no se ha cumplido con el objetivo.

Al ver esta problemática permanente, la idea inicial del presente trabajo fue identificar los factores relevantes que impactan en este indicador, así como también la relación con la inversión en los programas sociales y alimenticios implementados en los últimos 10 años. Sin embargo, tras evidenciar la dificultad del tratamiento de los datos y la falta de visualizaciones dinámicas de los mismos en entornos abiertos, se ideó que los resultados de este trabajo se presenten en un entorno web de acceso abierto y, además, que se acceso tanto a los datos utilizados para dichas visualizaciones como a los datos recopilados en el proceso. Uno de los principales aportes será que esta habilitación facilite el acceso a la información previamente trabajada a otros investigadores o entes relacionados, como son gobiernos regionales o miembros interesados de las distintas comunidades. La idea es que se pueda realizar un trabajo colaborativo y que se sumen esfuerzos, en vez de repetirlos.

Por todo lo mencionado, mi objetivo y aportación con el presente trabajo de investigación, es ser un agente de cambio que en el marco de mi proyecto de vida y de formación profesional, logre mejorar las condiciones y expectativas de muchos niños y niñas de mi país. Como lo menciona M. Gandhi: *"Seamos el cambio que queremos ver en el mundo"*.

1.2 Objetivos

1.2.1. Objetivo principal

- Ofrecer, mediante una plataforma de acceso abierto, visualizaciones dinámicas, análisis y datos de interés; con relación al índice de anemia infantil y los principales factores que podrían impactarlo.

1.2.2. Objetivos específicos

- Evaluar la relación entre el índice de anemia en el Perú y los factores demográficos, económicos, formativos y de salud a nivel nacional y regional.
- Evidenciar el impacto entre los programas sociales implementados y la evolución del índice de anemia.
- Presentar visualizaciones que sean de fácil comprensión y que brinden gran utilidad para entes impactados, investigadores y público en general interesados en el tema.
- Permitir la descarga centralizada de información para facilitar los análisis y utilizar un entorno colaborativo que permita realizar mejoras para el uso y beneficio de toda la comunidad.
- Ofrecer una web con diseño adaptable que permita una mejor visualización en diferentes dispositivos que puedan ser utilizados.

1.3 Impacto esperado

Se espera que esta herramienta sea utilizada por las entidades gubernamentales, gobiernos regionales, otros investigadores y personas interesadas en general. La idea es que las visualizaciones presentadas, sean de utilidad para los entes interesados en mejorar el indicador de índice de anemia infantil en el Perú, así como también que pueda servir para reducir el tiempo de extracción, consolidación y limpieza de datos relacionados.

El contar con visualizaciones a nivel de regiones, permitirá tener mayor granularidad y una mejor toma de decisiones a los gobiernos descentralizados. Así también, la posibilidad de utilizar la base de datos, los scripts y las visualizaciones trabajadas permitirá a nuevos investigadores tener un punto inicial más avanzado, con ello, un menor tiempo de inversión en tareas ya realizadas y, así, mayor avance en las nuevas investigaciones.

A través de este trabajo, se tendrá un impacto positivo en los siguientes ODS:

1. Objetivo 2: Hambre cero. – En la meta 2.1 se indica que se debe asegurar el acceso a una alimentación sana, nutritiva y suficiente. Y en la meta 2.2. se indica que se debe poner fin a todas las formas de malnutrición. Ambos puntos pueden ser fomentados con el presente TFM.
2. Objetivo 3: Salud y Bienestar. – En la meta 3.2 se menciona como meta reducir la mortalidad en niños menores de 5 años. Si bien esta meta no menciona directamente los temas de malnutrición, se tiene claro que la mejora en la alimentación y reducción de los niveles de anemia es también una forma de reducir la mortalidad en niños.
3. Objetivo 10: Reducción de las desigualdades. – La meta 10.3 menciona que se debe garantizar la igualdad de oportunidades. El plantear políticas de salud que permitan brindar un desarrollo óptimo del infante, permitirá mejorar la igualdad de oportunidades.
4. Objetivo 16: Paz, justicia e instituciones sólidas. – La meta 16.6 indica que es necesario crear instituciones eficaces y transparentes que rindan cuentas. El poner a la disposición información relevante y fácil de entender, permitirá al ciudadano tener una mayor aportación en la vigilancia y cuestionamiento sobre las acciones del estado.

Un mayor detalle de estos puntos se encuentra en el *Anexo I: "Relación con ODS"*

1.4 Metodología

En este apartado se encuentra la secuencia de actividades y criterios tomados en cuenta para la identificación de fuentes de datos, selección de las fuentes óptimas, recopilación de datos, proceso de consolidación y limpieza de la base de datos, validación de datos, elaboración de visualizaciones, análisis de los resultados y finalmente, publicación y habilitación en datos abiertos. Así también, se incluyen la selección de técnicas, herramientas y métricas utilizadas para el procesamiento y planteamiento del modelo.

1.4.1. Selección de variables

Se realiza la identificación y selección de variables que representan factores significativos que influyen en el tema de estudio. Esto permitirá evaluar que fuentes pueden ser consideradas como input para el análisis y desarrollo.

1.4.2. Identificación de fuentes de datos

La identificación de fuentes de datos se realizará mediante una búsqueda de datos abiertos de las instituciones que gestionan la base de datos relacionada con los indicadores de anemia. Se incluyen en este punto tanto la búsqueda en los sitios webs, como también la solicitud a los organismos pertinente.

Para poder tener un grado de fiabilidad aceptable y que la información obtenida sea útil para el estudio, se toman en cuenta las siguientes consideraciones:

- Seleccionar la información de fuentes confiables.
- Identificar que la base de datos corresponda al universo que se está evaluando.
- Verificar que existan los datos necesarios para el estudio.

1.4.3. Selección de las fuentes óptimas

Sobre la información recopilada en el punto preliminar, se utilizará la metodología MELODA, para determinar cuáles serían las mejores fuentes de información para los objetivos planteados (Abella, Ortiz-de-Urbina-Criado, De-Pablos-Heredero 2020). Esta metodología utiliza una calificación en función de los siguientes criterios:

- Tipo de licencia para el uso de los datos.
- Facilidad de acceso.
- Estándar técnico.
- Estandarización.
- Información de geolocalización.
- Frecuencia de actualización.
- Reputación de la fuente.
- Tipo de difusión.

A partir de la puntuación obtenida después de realizar la calificación, se seleccionará la fuente o las fuentes de datos óptimas para el presente trabajo. Cabe resaltar que las fuentes de datos seleccionadas para esta fase del estudio deben ser acotadas

1.4.4. Recopilación de datos

En este paso se diseña el flujo de trabajo que se requiere, a fin de poder obtener los datos desde la fuente seleccionada y luego poder transformarlos al formato que se requiere para su procesamiento.

En este punto, para poder tener un grado de fiabilidad aceptable y trazabilidad de la información obtenida, se deberán tomar en cuenta las siguientes acciones:

- Documentar la fecha de la extracción de la información.
- Permitir la descarga y transformación de más de una extensión.
- Documentar los criterios técnicos que han sido considerados.
- Documentar los medios para recopilar la información.

Con este fin se utilizarán técnicas como el uso de APIs (interfaces de programación de aplicaciones) y/o web scraping, para reconocer las estructuras de los datos, extraer y transformar contenidos y finalmente, almacenar los datos en el formato deseado.

1.4.5. Limpieza y consolidación de datos

En este punto se realiza la consolidación de los archivos, a fin de poder tener una única base consolidada, así como también la limpieza de datos. Las casuísticas que se deben considerar en este proceso son:

- Crear índices para mejorar la velocidad de consulta y consolidación.
- Consolidar la información de archivos o tablas con campos distintos.
- Homologar la información en distintos campos.
- Retirar los valores nulos o inconsistentes.
- Retirar espacios en blanco.

Para la ejecución de todas estas actividades se hace uso de herramientas de programación, tales como Visual Studio Code con Python. Toda la información recopilada se almacenará en una base de datos relacional de MySQL.

1.4.6. Validación de datos

Se realizó las validaciones de los archivos a fin de poder identificar datos inconsistentes o incongruencias con los datos presentados por el ente que proporciona los datos. Los puntos para tener en cuenta son:

- Verificación de que haya información de las tablas necesarias para todos los años de alcance de la evaluación.
- Validar que los valores calculados coincidan con los presentados por el organismo de la fuente de información.
- Comprobar que las columnas utilizadas para la consolidación de tablas cumplan con los mismos criterios y así evitar que se omitan datos.

Para este punto también se hizo uso de Visual Studio Code con Python Colab. Cabe mencionar que, al finalizar este punto, es posible regresar al punto 1.4.5., ya que podrían existir tareas a realizar.

1.4.7. Carga de fuentes adicionales

En este paso se realizó la exploración y recopilación de datos de fuentes adicionales que se pudieran requerir para completar la información base. El detalle de las fuentes y la información que será necesaria se tendrá aún durante la implementación de la solución.

En el punto 5.7., “Carga de fuentes adicionales”, se puede encontrar el detalle de lo referido a este punto. Se debe tener en cuenta que las fuentes que se pueden incorporar son diversas y dinámicas y que el objetivo es ir complementándolas poco a poco los datos que se tienen e ir enriqueciendo el inventario.

1.4.8. Elaboración de visualizaciones de datos

Después de realizar el proceso de consolidación y validaciones, se procede con la elaboración de tableros de control que permita tener una mejor visualización de la información y de este modo, servir a una mejor evaluación. Se plantea una presentación inicial global de todo el país y una presentación por regiones que, en Perú, es un nivel representativo y cuyo presupuesto de inversión para temas de desarrollo e implementación de programas sociales, es independiente al gobierno central. En este punto se deben tomar en cuenta los siguientes criterios:

- Las presentaciones deben mostrar de manera clara las evoluciones por años.
- Se deben representar todos o la mayor parte de factores que se consideran relevantes a partir de la información recolectada.
- Los números deben ser presentados en función de índices y no de totales. De este modo, se podrá tener mejor visibilidad de la incidencia real.
- Se debe tener en cuenta la visualización de medias y ver cómo se desarrollan los valores por encima y por debajo de este indicador.

Para desarrollar este punto, la herramienta que se utilizará será Tableau 2022.1, cuya licencia se pudo obtener de forma gratuita por del uso académico, y Tableau Public, la cual es gratuita y puede ser utilizada por cualquier usuario. Esta es la etapa inicial propuesta, más adelante, de requerirse, se propone migrar a alguna otra herramienta.

1.4.9. Publicación y habilitación en datos abiertos

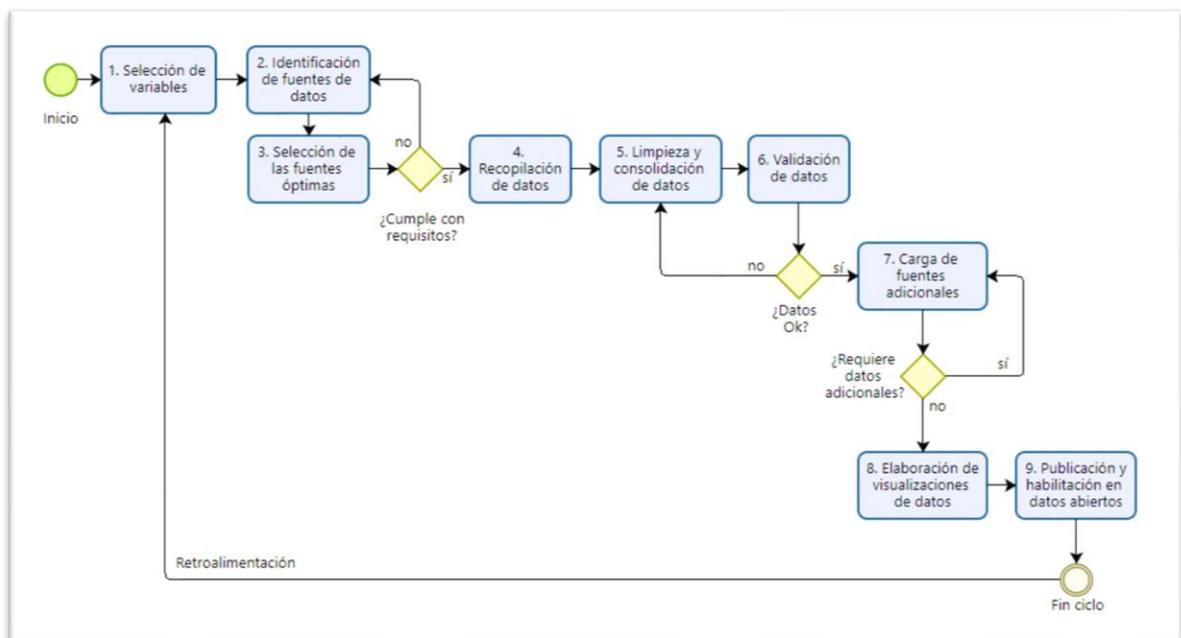
Finalmente, se procede con la publicación de los tableros de control desarrollados, así como también se habilita la base de datos extraída y consolidada. Esto se realiza en un entorno web con acceso abierto tomando en cuenta:

- La licencia que se otorga será Creative Commons CC BY-SA.
- Se permitirá la extracción global y parcial de la base de datos sin tratamiento, así como también las fichas técnicas o documentación que brinde información de esta.

- La visualización de filtros o indicaciones de uso de las herramientas debe ser de fácil uso y brindar una buena experiencia de usuario.
- Se maneja la creación de usuarios para la creación de tableros de control a fin de que figuren los autores de estos.
- La actualización de los datos se realiza de manera centralizada, cada vez que se requiera, por ejemplo, durante la publicación de nuevos datos o por actualizaciones.

Para poder implementar este punto, se hace uso de las siguientes herramientas: WAMP Server, Visual Studio Code, Python, HTML, JavaScript, MySQL y el repositorio de datos de GitHub.

Ilustración 1. Metodología de desarrollo



Fuente: Elaboración propia

El proceso inicial es una versión que puede ajustarse y/o ampliar el alcance, como, por ejemplo, ampliar la cantidad de datos, módulos, tablas o formatos de datos incluidos en los análisis o publicaciones iniciales. Se trabaja en un entorno colaborativo, por ello, el código de todo este trabajo está publicado en repositorios de GitHub.

1.5 Estructura del trabajo

El presente trabajo se ha organizado en las siguientes secciones:

- Acrónimos. - Aquí se han incluido los términos que están compuestos por siglas que son representativas de un nombre más extenso, como puede ser una institución, una encuesta, un método u otros.

- Capítulo I: Introducción. - Esta sección incluye una descripción holística del problema identificado, así como también, la motivación que impulsó a la elección del tema de estudio, los objetivos planteados, el impacto esperado y el detalle de la metodología empleada para el desarrollo del trabajo.
- Capítulo II: Estado del arte. - Se incluye el marco de referencia, en donde se encuentra la descripción de conceptos e instituciones que son importantes conocer. También se incorpora un análisis de las investigaciones afines realizadas y los puntos que se destacan u observan de cada una de ellas. Otro punto importante a considera en este acápite es la tecnología en donde se incluye las herramientas que realizan una función similar a la propuesta en este trabajo. Finalmente, se incluye la propuesta y contribución de esta investigación.
- Capítulo III: Análisis del problema. – En este capítulo se realiza una reflexión y compendio sobre las principales situaciones identificadas que presentan oportunidades de mejora y que se abordan en el desarrollo de este trabajo. También se analizan el marco legal para ver las implicaciones que se podrían presentar y se realiza un pequeño análisis de riesgos. Finalmente, se ponen de manifiesto las posibles soluciones y se selecciona la propuesta más adecuada, explicando el motivo de la elección.
- Capítulo IV: Diseño de la solución. - Se presenta la planificación de las actividades, la arquitectura del sistema propuesto (mediante diagramas), así como también las tecnologías que se han desplegado a lo largo del desarrollo de este trabajo.
- Capítulo V: Desarrollo de la solución. - Aquí se despliega la metodología planteada en el capítulo I: Introducción. Se describe cada paso realizado para la puesta en marcha de la solución propuesta, que incluye desde la selección de variables hasta la publicación en un entorno Web.
- Capítulo VI: Pruebas y resultados. - Se realizan las pruebas de adaptabilidad, funcionalidad y auditoría SEO². También se incluye un análisis de los resultados presentados en las visualizaciones, tanto a nivel nacional como regional. En este capítulo, además, se incluyen las principales dificultades que se presentaron durante la ejecución de este trabajo y cuáles fueron las soluciones planteadas.
- Conclusiones. - Se incluye la interpretación de los resultados y la manera en cómo se han alcanzado los objetivos planteados, enfatizando algunos puntos que se considera, deben tener especial atención. También se incluyen los principales retos y aprendizajes que se han tenido a lo largo de la ejecución del trabajo.

² Auditoría SEO (acrónimo de Search Engine Optimization), es la evaluación que permite identificar brechas que son necesarias superar para la optimización en los motores de búsqueda.

- Relación del trabajo con los estudios cursados. - En este acápite se describe la relación del trabajo realizado en función a los cursos desarrollados en el Máster Universitario de Gestión de la Información (MUGI) y también se han incluido las competencias transversales que fueron importantes adquirir y poner en práctica para el logro de los objetivos.
- Trabajos futuros. - Aquí se incluye la lista de oportunidades de mejoras y desarrollos futuros que permitan fortalecer los objetivos de esta iniciativa.
- Referencias. - Se incluyen los principales libros, estudios, artículos y demás fuentes de información consultadas para desarrollar las diversas secciones de este informe.
- Anexos. - Se documenta la información complementaria que sirve como soporte o fuente de consulta adicional.

1.6 Convenciones

Tomar en cuenta los siguientes puntos:

- Se entrecorillan y se usa cursiva para las citas textuales externas a la obra.
- Las fórmulas se encuentran en letra cursiva y en viñetas.
- Para las referencias bibliográficas se utiliza la norma ISO-690. Las citas se encuentran entre paréntesis y en cursiva, tal como sigue: (*"Primer apellido de autor" "año"*)
- El código fuente se muestra en letra Source Sans Pro, cursiva, con grosor normal.

CAPÍTULO II: Estado del arte

En este capítulo se describen los entes, conceptos, estudios y tecnologías relacionadas que fueron necesarios para el desarrollo del presente trabajo.

Este capítulo se dividirá en las siguientes secciones:

- En la primera, denominada “Marco de referencia”, se aborda la anemia infantil, índice que se ha mantenido en Perú, entes encargados y métodos utilizados para recolectar, procesar y gestionar los datos que ayudan con la medición de este indicador. Así también se incluye un análisis sobre el estado de los datos abiertos en el Perú.
- En la segunda sección, llamada “Estudios relacionados”, se analizan las investigaciones previas que han tenido como foco de estudio, la anemia infantil.
- En la tercera, denominada “Tecnología”, se incluyen herramientas que realizan funciones similares a la propuesta en este estudio.
- En la cuarta y última sección, “Propuesta y contribución de la investigación” se incluirá cuáles han sido las consideraciones tomadas en cuenta y cuál la aportación o diferenciación del presente trabajo de investigación.

2.1 Marco de referencia

2.1.1. Anemia

La anemia es una afección en la que el número de glóbulos rojos o la concentración de hemoglobina dentro de estos es menor de lo normal. La hemoglobina es necesaria para transportar oxígeno y si una persona tiene déficit, se disminuirá la capacidad para trasladar oxígeno a los tejidos del organismo. Las causas más comunes de anemia son las carencias nutricionales, siendo la más frecuente la falta de hierro. La anemia es un grave problema de salud pública en el mundo que afecta particularmente a los niños pequeños y las mujeres embarazadas (*Organización Mundial de la Salud sin fecha*). Se ha comprobado que la anemia por carencia de hierro también trastorna el desarrollo cognoscitivo y físico de los niños y mengua la productividad de los adultos. E incluso, existen estudios que mencionan que, en niños menores de dos años, esta patología podría traer efectos duraderos en la estructura y funcionamiento cerebral (*Beltrán-Navarro, Matute, Vásquez-Garibay 2019*).

La prevalencia de la anemia infantil (niños de 6 a 36 meses) en el Perú fue del 40% para el año 2020 y para el año 2021, aún estaba en un 38,8%, a pesar de los esfuerzos y programas que se pusieron en marcha para que este porcentaje bajara.

De acuerdo con el parámetro de clasificación de la importancia de la anemia para la salud, un valor mayor o igual al 40% representa un problema severo de salud pública y Perú ha estado sobre este valor durante más de 10 años (*Organización Mundial de la Salud 2011*).

2.1.2. Instituto Nacional de Estadística e Informática (INEI):

El INEI es un organismo que pertenece al sector público del Perú, pero que cuenta con autonomía técnica y de gestión. Es el responsable de todas las actividades estadísticas oficiales del país, incluidas todas las encuestas oficiales, tales como la ENDES (Encuesta Demográfica de Salud Familiar), la que se realiza anualmente mediante una entrevista directa e incluye preguntas relacionadas con la salud materna e infantil, además de la fecundidad. Los resultados de esta encuesta son utilizados para la aplicación de programas con el fin de mejorar los indicadores relacionados. (PERÚ Instituto Nacional de Estadística e Informática sin fecha)

La sección en relación con la medición de anemia de la ENDES incluye un sistema de medición para el nivel de hemoglobina, el INEI (Instituto Nacional de Estadística e Informática) utiliza la técnica HemoCue, que es reconocida y recomendada por la OMS como estable y duradera. Las concentraciones de hemoglobina que se utilizan para determinar la anemia y su criticidad en niños son:

- *Anemia severa: < 7,0 g/dl*
- *Anemia moderada: 7,0-9,9 g/dl*
- *Anemia leve: 10,0-10,9 g/dl*

(Organización Mundial de la Salud 2011)

Adicionalmente, se considera el ajuste del nivel de hemoglobina relacionado a la altitud. Para este informe se ha adoptado la fórmula de CDC/PNSS³ y Dirren. Las mediciones iniciales del nivel de hemoglobina han sido llevadas a nivel del mar⁴ con el siguiente procedimiento:

- *Nivel ajustado = nivel observado - ajuste por altura*
- *Ajuste = -0,032*(alt⁵) + 0,022*(alt*alt)*
- *Donde (alt) es: [(altura en metros) /1 000]*3,3*

(INEI 2018)

³ PNSS: Pediatric Nutrition Surveillance System

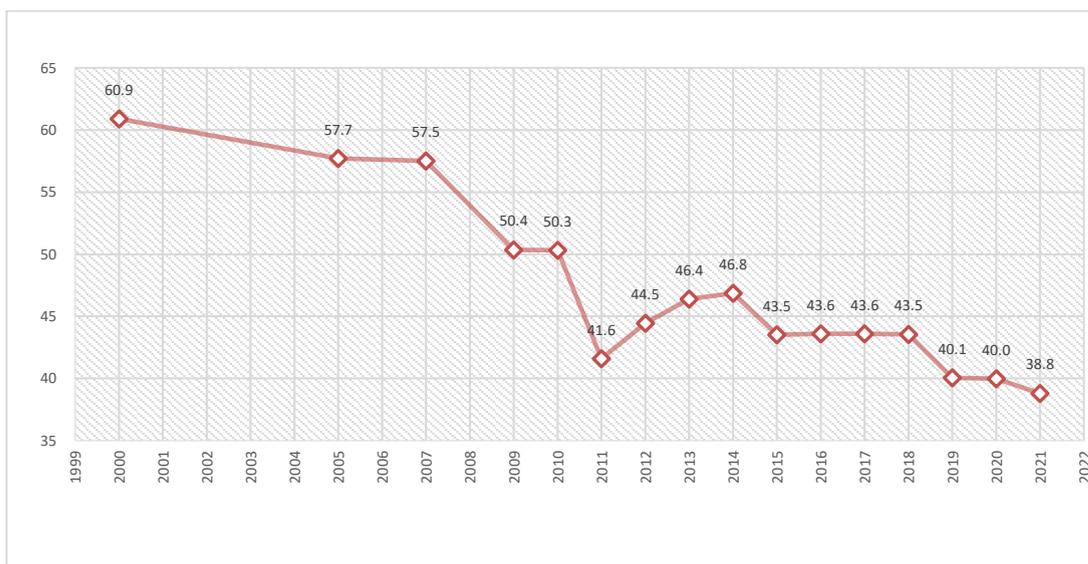
⁴ Al nivel del mar hace referencia a la corrección de la altitud.

⁵ alt: altitud

2.1.3. Evolución de la anemia infantil⁶ en Perú

De acuerdo con los datos recopilados desde el año 2000 por el Instituto Nacional de Estadística e Informática (INEI), a través de la Encuesta Demográfica y de Salud Familiar (ENDES), a través de los años de medición se han presentado variaciones significativas en el índice de anemia en niños de 6 a 35 meses. Como se puede apreciar en la ilustración 2, para el año 2000 se tenía un índice muy elevado de 60,9% y luego tuvo una evolución favorable hasta el año 2011, en el que se alcanzó un índice de 41,6%. Después de dicha fecha se presentaron incrementos hasta el año 2014, en el que el índice fue de 46,8%. Luego, entre el 2015 y 2018 se mantuvo estable en 43,6% y para el año 2019 y 2020 se tuvo una reducción a 40,1% y 40%, respectivamente. Los resultados presentados para el año 2021 tiene un índice de 38,8%, lo cual es favorable, a pesar de las limitaciones que se han tenido con la pandemia. Cabe mencionar que, hasta el año 2005, las evaluaciones del indicador de índice de anemia se realizaban con una periodicidad de 5 años y para los años del 2005 al 2009, se tuvo una periodicidad bianual, por ello, no se cuenta con datos consecutivos para dichos años. Solo a partir del año 2009, las evaluaciones son anuales.

Ilustración 2. Prevalencia de anemia en niños de 6 a 35 meses (en porcentaje)



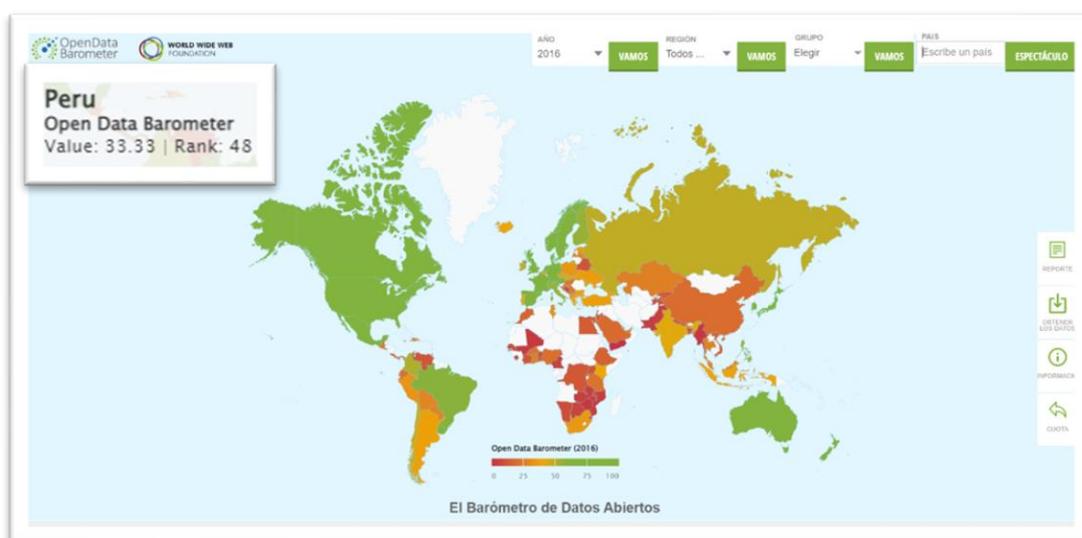
Fuente: INEI - Elaboración propia

⁶ Anemia en niños y niñas de 6 a 35 meses.

2.1.4. Datos abiertos en Perú

En el último reporte de ILDA⁷ realizado el año 2020, Perú obtuvo una puntuación de 45 sobre 100, que es el resultado de las puntuaciones obtenidas en los acápites de preparación para iniciativas, implementación de programas e impacto que se tiene en las empresas, política y sociedad civil. A pesar de que se puede ver un crecimiento de 12 puntos en 7 años, el valor obtenido sigue siendo bajo en relación con el crecimiento esperado. Para hacer una comparativa global, se cuenta con el último reporte general de todos los continentes, realizado en el año 2016 por Open Data Barometer, en el cual Perú ocupaba el número 48 en el ranking con un puntaje de 33, en la Ilustración 3 se puede ver un mapa de calor donde se muestra el barómetro para los países incluidos en el estudio.

Ilustración 3. Open Data Barometer 2016



Fuente: OpenData Barometer

Para el caso de América Latina y el Caribe, Perú pasó del 7º lugar obtenido en el 2016, al puesto al 9º en el 2020. De los parámetros que mide ILDA, que son preparación, implementación e impacto de los datos abiertos, como se aprecia en la Ilustración 4, el indicador más bajo y preocupante es el indicador de impacto, ya que, de la evaluación realizada, solo se obtuvo un valor de 6,67 sobre 100, lo que significa que no se ha encontrado evidencia suficiente de que la publicación de datos abiertos haya generado repercusiones positivas en el país (*Open Data Barometer sin fecha*).

⁷ ILDA, por sus siglas de Iniciativa Latinoamericana por los Datos Abiertos, es el barómetro de datos abiertos para América Latina y el Caribe. El último reporte emitido por esta institución fue en el año 2020

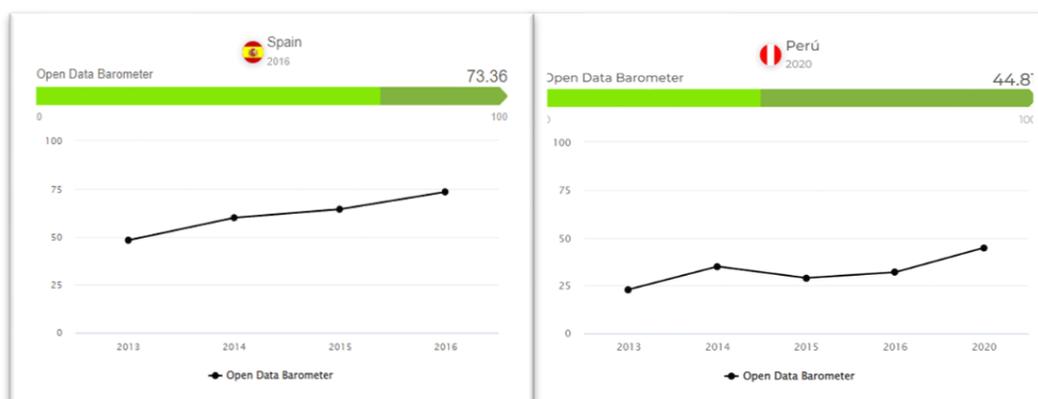
Ilustración 4. Open Data Barometer 2020



Fuente: ILDA

Para poder tener una mejor idea de la situación actual de Perú en relación con el gobierno de datos abiertos, se ha realizado un comparativo con España. En la Ilustración 5 se puede apreciar la evolución de cada uno de estos 2 países en el mismo rango de tiempo y se puede ver que el puntaje obtenido en el año 2020 por Perú (44.8/100), es incluso menor a la puntuación que tenía España el año 2013 (48.1/100).

Ilustración 5. Comparativa de países



Fuente: OpenData Barometer e ILDA

Finalmente, según se menciona en una publicación de ILDA, denominada: *Gobernanza cultural y datos abiertos: Retos y oportunidades para los sectores culturales en el Perú*, las políticas gubernamentales de datos abiertos aún siguen siendo un tema poco desarrollado, caracterizado por una realidad marcada por el abandono, el desorden institucional, la poca relación con la sociedad civil y un liderazgo institucional muy pobre (Delfin 2021).

2.1.5. El derecho de acceso a la información en Perú

En junio del año 2002, se aprobó en Perú la Ley N° 27806, de Transparencia y Acceso a la Información Pública, la cual tiene por finalidad brindar transparencia de los actos del Estado y regular el derecho de acceso a la información consagrado en el numeral 5 del Artículo 2° de la Constitución Política del Perú.

A través de esta normativa, el ciudadano está facultado a solicitar cualquier información (salvo excepciones) que se haya elaborado con recursos del estado. De este modo, poder tener una sociedad más justa y transparente.

2.2 Estudios relacionados

A lo largo de los años, la anemia en niños ha sido un tema de interés nacional en Perú. Diversos entes públicos y privados han realizado estudios sobre la evolución de este indicador, así como también, sobre los factores que pueden ser causantes del alto índice que se mantiene a través de los años. Como parte inicial de este estudio, se realizó la búsqueda y análisis de estudios e investigaciones similares. Se describen a continuación los que se consideran más relevantes:

A. Desnutrición infantil en menores de cinco años en Perú: tendencias y factores determinantes (Sobrino et al. 2014).

En esta investigación se realiza el análisis de los factores que pueden impactar en los indicadores de desnutrición y anemia en niños de 6 meses hasta 5 años para el periodo del 2000 al 2011. El estudio es elaborado por personal sanitario y tiene un respaldo de la Organización Panamericana de la Salud, sede Lima, Perú. Los factores que son incluidos para evaluar su influencia en el índice de anemia son: sexo, edad, zona y región de residencia, educación de la madre, quintil de riqueza, disponibilidad de agua potable, disponibilidad de cloacas, presencia de otros niños en el hogar, orden de los nacimientos, presencia de diarrea y presencia de tos en los 15 días previos.

Una de las conclusiones de este estudio es que para superar la alta tasa de anemia se debe considerar que no es un problema únicamente alimenticio y que también se deben considerar factores determinantes asociados. Este estudio ayuda a identificar, a nivel general, factores que podrían tener correlación directa con los índices de anemia.

Se tiene que considerar que este estudio no incluye el despliegue por regiones y que no se ha podido identificar que se hayan realizado posteriores actualizaciones.

B. Anemia infantil en el Perú: análisis de la situación y políticas públicas para su reducción (Guabloche 2021).

Este estudio aborda la evolución del índice de anemia infantil y su relación con los planes gubernamentales y metas regionales para de los años 2017, 2018 y 2019. La información presentada incluye, entre otros temas: (1) la relación entre el índice de anemia y el gasto en suplementos alimenticios, (2) las brechas identificadas entre los resultados obtenidos según el INEI y las metas regionales planteadas, así también (3) la situación de la anemia infantil en época de pandemia.

Este estudio es importante, porque permite considerar otras variables que podrían generar un impacto en la variación del índice de anemia. En este caso, la principal variable a considerar es la relación con el gasto en suplementos alimenticios que se ha dado por región. En este estudio se muestra esta evaluación, pero no se ha considerado una evaluación evolutiva, sino solo la relación entre dos puntos en el tiempo: variación porcentual del gasto en suplementos e índice de anemia entre los años 2014 y 2019.

C. Anemia infantil, retos y oportunidades (Zavaleta 2021).

El estudio está realizado con la base de niños y niñas de 6 a 35 meses de edad. Aquí se presenta la comparación, entre los años 2012 y 2016, de prevalencias de anemia por grupo etario y por región. Así también se describen las principales causas y consecuencias de la anemia infantil. En esta presentación se incluye un estudio del neurodesarrollo infantil, en donde se visualizan las ventanas de oportunidades por rango de edad, siendo el rango etario de este estudio uno de los más importantes para el desarrollo emocional, motor y cognitivo, que determinará la agilidad mental y motricidad futura del niño. El criterio de selección del rango etario de este estudio ha sido también considerado, ya que el sustento presentado es válido y relevante.

Las causas mencionadas en este estudio, tales como factores culturales, acceso y disponibilidad a los alimentos, bajos ingresos, cuidado prenatal y educación de la madre, también han sido consideradas para ser incluidas.

En este análisis no se ha identificado tener visibilidad sobre la evolución periódica (anual) por región, ni tampoco se ha podido encontrar desplegados los indicadores relevantes que permitan visualizar la influencia de los programas sociales implementados.

D. Análisis del modelo multicausal sobre el nivel de la anemia en niños de 6 a 35 meses en Perú (Ortiz Romaní et al. 2021)(Ortiz Romaní et al. 2021).

Este estudio es de nivel explicativo y toma como año alcance de la evaluación, la información de la encuesta ENDES del 2019. Este estudio incide en el grupo etario de niños entre 6 y 35 meses, ya que refiere que la prevalencia es alta en ellos.

Considero que esta investigación es interesante por la mezcla de profesionales que trabajaron en su elaboración: dos enfermeros, un médico, un nutricionista y un estadístico. Por ello, al realizar el análisis de los datos, no solo se realizó a nivel estadístico, en donde se trabajaron con tablas de contingencia, sino que también, para cada factor identificado como causal, se le brinda una explicación sanitaria o nutricional.

En las conclusiones, este estudio menciona que los factores que pudieran influenciar en la anemia son: diarrea en las últimas dos semanas, edad del niño, suministro de agua, control prenatal, anemia y rango etario de la madre, lactancia materna y quintil de riqueza.

Este análisis presenta una ligera distorsión en relación con las cifras presentadas por ENDES y esto puede deberse a que no se están considerando algunos factores de segmentación que el INEI utiliza en la ENDES para el cálculo de sus informaciones estadísticas.

E. Evaluación del impacto de los multimicronutrientes en polvo sobre la anemia infantil en tres regiones andinas del Perú (Munayco et al. 2014).

En este estudio se incluye la evaluación del impacto de los micronutrientes sobre el índice de anemia en niños de 6 a 35 meses. El estudio está centrado en 3 regiones del Perú, sin embargo, se indica que no se realizó en todos los centros de salud y que no hubo un criterio estadístico para ponderar los centros seleccionados.

Finalmente, el estudio concluye que la suplementación con micronutrientes (MMN) en polvo puede ser una estrategia efectiva en la lucha contra la anemia. De este estudio, se rescata esta conjetura y para el desarrollo de nuestro estudio, se realizará también un análisis para identificar la correlación del índice de anemia con los suplementos alimenticios que se brindan por el estado peruano.

F. Consumo de suplementos con multimicronutrientes Chispitas® y anemia en niños de 6 a 35 meses: estudio transversal en el contexto de una intervención poblacional en Apurímac, Perú (Huamán-Espino et al. 2012).

Similar al estudio previo, en este se incluye la evaluación del impacto de un micronutriente específico sobre el índice de anemia en niños de 6 a 35 meses. El estudio está centrado en la región Apurímac de Perú y se realizó sobre una población de 714 participantes.

Uno de los temas más interesante de este artículo es que se menciona que no se encontró una relación directa entre la cantidad de sobres recibido o consumidos y el índice de anemia, pero sí se pudo apreciar una menor prevalencia de anemia entre los niños que consumieron el suplemento de manera adecuada con respecto a los niños que no lo hicieron. Finalmente, se concluye que no basta con entregar o consumir la cantidad necesaria de micronutrientes, sino que es necesario asegurar el consumo adecuado. Este punto es muy importante tomarlo en cuenta para que, en el análisis, se pueda verificar o cuestionar.

G. Consumo de hierro y prevalencia de anemia en niños y adolescentes en una comunidad a gran altitud en Perú (Ballon-Salcedo et al. 2020)

Estudio realizado en Ayroca, un centro poblado de la región de Arequipa con una altitud de más de 3000 metros sobre el nivel del mar. Según lo mencionado, se realiza en esta población porque representan un grupo vulnerable de padecer anemia. El objetivo fue evaluar el índice de anemia en relación con los conocimientos, actitudes y prácticas sobre el consumo de hierro, teniendo resultados no concluyentes y planteando que puede ser necesario considerar factores de corrección por altitud, distintos a los actualmente utilizados, ya que se debe considerar el factor de adaptación a la hipoxia hipobárica que adquieren las poblaciones que viven a gran altitud.

A pesar de que consideramos que el alcance de este estudio es muy acotado, ya que se basa en una población de solamente 65 niños y adolescentes, una parte muy rescatable es el cuestionamiento en cuanto a los criterios de ajuste de hemoglobina que se realizan en las zonas con elevada altitud. Pensamos, que quizás este ajuste sea incorrecto ya que es probable que la adaptabilidad de la población deje este criterio desfasado y que por ello, se pueda explicar al bajo impacto y mejora que se ha tenido por años en estas zonas.

Finalmente, luego de revisar todos los estudios mencionados, se puede indicar que la principal diferenciación que tiene este trabajo es que el análisis será evolutivo, realizado con información de 12 años, comparativas sencillas y en el alcance geográfico se incluyen todas las regiones del Perú. Además, se mostrará cuadros de mandos que permitan tener visualizaciones fáciles de entender y que con pocos y simples cambios, se puedan adecuar a las necesidades de quien lo requiera. Finalmente, otro punto importante es que el trabajo realizado, así como la data fuente que permitió su realización, serán compartido en datos abiertos para que todo el que necesita, pueda hacer uso de él.

2.3 Tecnología

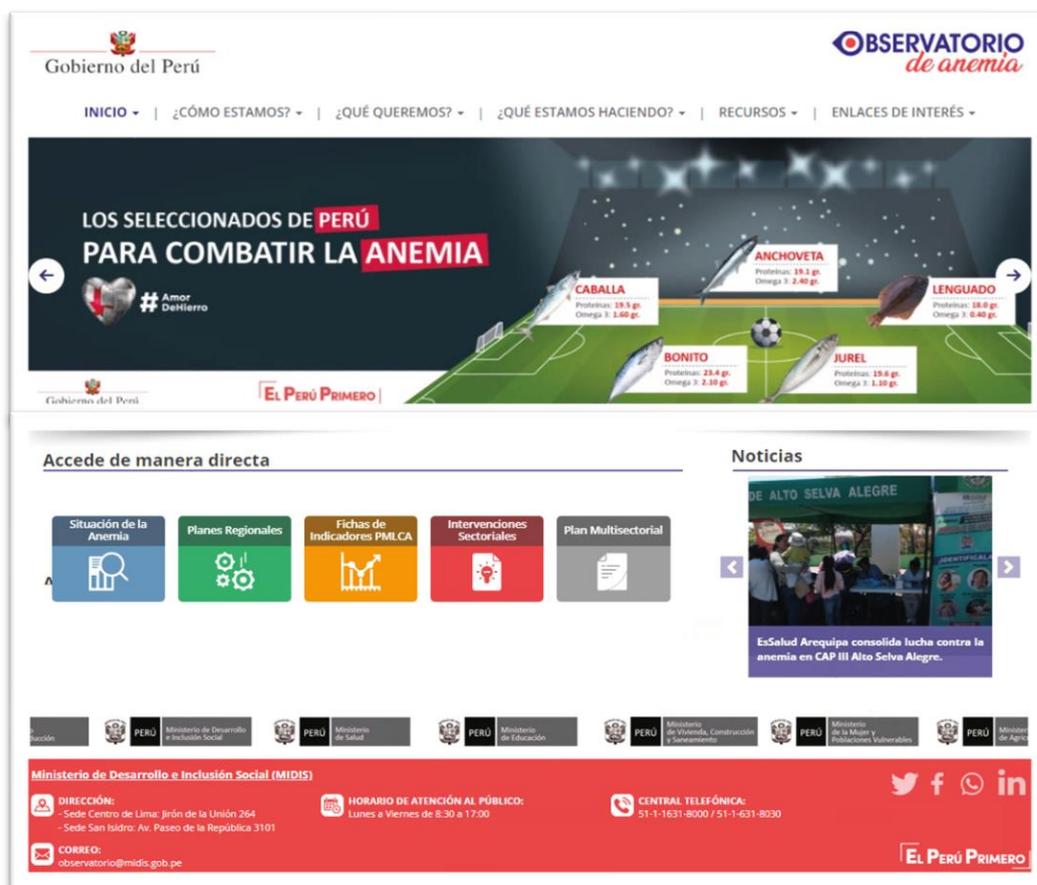
En esta sección se presentarán las herramientas tecnológicas identificadas, que proporcionan un servicio de datos e información con relación a la anemia.

2.3.1. Observatorio de Anemia

De acuerdo con la definición del Gobierno del Perú, este observatorio *es una “herramienta de información implementada en un entorno web, y administrada por el Ministerio de Desarrollo e Inclusión Social (MIDIS)”*.

Esta herramienta incluye datos de los distintos sectores y niveles de gobierno, los avances en torno al objetivo de disminuir la anemia, los planes y programas, así como también los recursos que se tienen. Esta herramienta utiliza la información de la ENDES y la fecha de consulta fue junio 2022 (*Observatorio de Anemia sin fecha*).

Ilustración 6. Observatorio de Anemia: Inicio

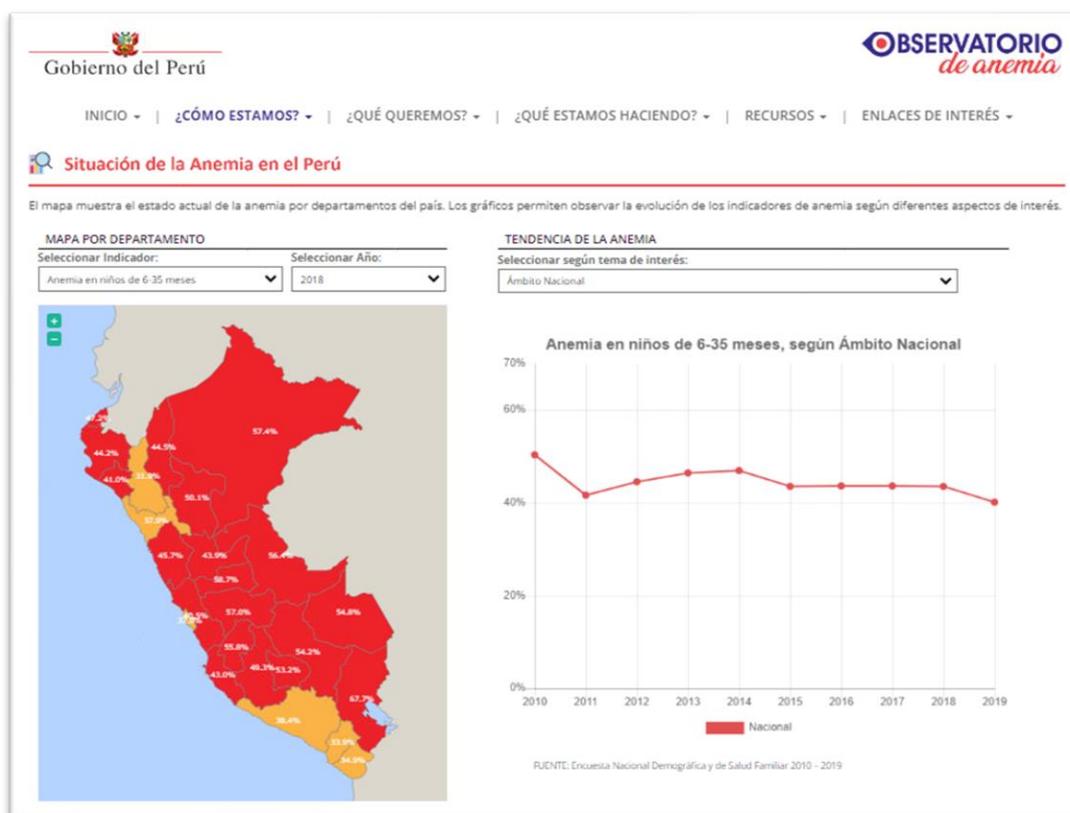


Fuente: Observatorio de anemia inicio⁸

⁸ El portal puede ser consultado en: http://sdv.midis.gob.pe/sis_anemia/home

Al realizar el análisis de la herramienta se apreció que los datos presentados solo están actualizados hasta el 2019 para datos globales y hasta el año 2018 para datos por región. Esto a pesar de que, a la fecha de consulta, la fuente de información que ellos utilizan, ENDES, ya había presentado su informe hasta el año 2021. Otra característica que se puede notar es que los gráficos presentados son preestablecidos, con indicadores globales y que tiene muy poco nivel de detalle. Todo esto se puede ver en la Ilustración 7, presentada a continuación.

Ilustración 7. Observatorio de Anemia: ¿Cómo estamos?



Fuente: Observatorio de anemia ¿cómo estamos?

Con respecto al análisis y comparativa con los objetivos planteados por cada gobierno regional, como se puede ver en la Ilustración 8, sí se notó una clara presentación que permite ver las desviaciones y brechas de cumplimiento. Además, se puede apreciar una comparativa entre el objetivo nacional y los objetivos regionales. Lo que hay que mencionar en este punto es que solo se tienen datos hasta el año 2019, aunque las actualizaciones ya están realizadas hasta el año 2021.

Ilustración 8. Observatorio de Anemia: ¿Qué queremos?



Fuente: Observatorio de Anemia: ¿Qué queremos?

En relación con resto de la página, se puede notar que tiene gran parte de información en formato PDF, algunos enlaces rotos y al igual que las otras secciones de la página, se encuentra actualizada solo hasta el 2019.

Ilustración 9. Observatorio de Anemia: Intervenciones Sectoriales

Intervenciones Sectoriales

MINSA | MINAM | MINAGRI | MINEDU | MINCUL

Presentaciones | Matriz de Seguimiento | Informes | Reuniones

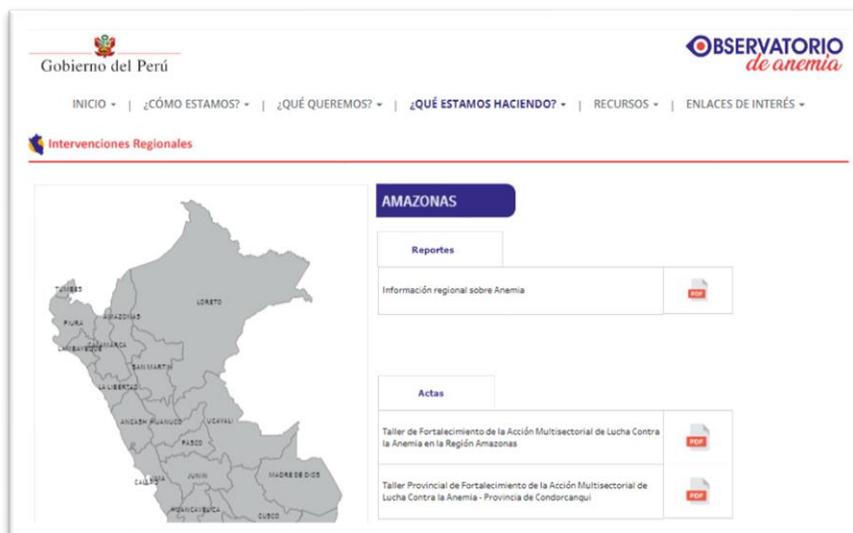
Acciones del MINSA para la lucha contra la anemia
 Ministerio de Salud
 11 enero 2018

El Ministerio de Salud en el marco de la lucha contra la anemia, se encuentra realizando acciones para la prevención y tratamiento de anemia en la población infantil.

Fecha de publicación: enero 2019

Fuente: Observatorio de Anemia: ¿Qué estamos haciendo?

Ilustración 10. Observatorio de Anemia: Intervenciones Regionales



Fuente: Observatorio de Anemia: ¿Qué estamos haciendo?

Ilustración 11. Observatorio de Anemia: Indicadores Sectoriales

The screenshot shows the 'Observatorio de Anemia' website interface for 'Indicadores Sectoriales'. A dropdown menu is set to 'MINSa'. Below, a table displays indicators with columns for 'INDICADOR', 'UNIDAD MEDIDA', 'BASAL', 'META', 'VALOR', 'FECHA CORTE', 'DOC', 'BD', and 'FT'. The table contains four rows of data, all with a 'VALOR' of 35 and a 'FECHA CORTE' of 10/05/2019.

INDICADOR	UNIDAD MEDIDA	BASAL	META	VALOR	FECHA CORTE	DOC	BD	FT
Porcentaje gestantes con anemia.	Número			35	10/05/2019	PDF	PDF	PDF
Porcentaje de mujeres con parto institucional que durante su embarazo tuvieron 4 exámenes auxiliares en el primer trimestre y al menos 6 atenciones prenatales con suplemento de hierro y ácido fólico.	Número							PDF
Porcentaje de Niños de 4 meses de edad que reciben gotas con hierro.	Número							PDF
Porcentaje de Niños de 6 a 11 meses de edad con tamizaje de anemia en los últimos 3 meses.	Número							PDF

Fuente: Observatorio de Anemia: ¿Qué estamos haciendo?

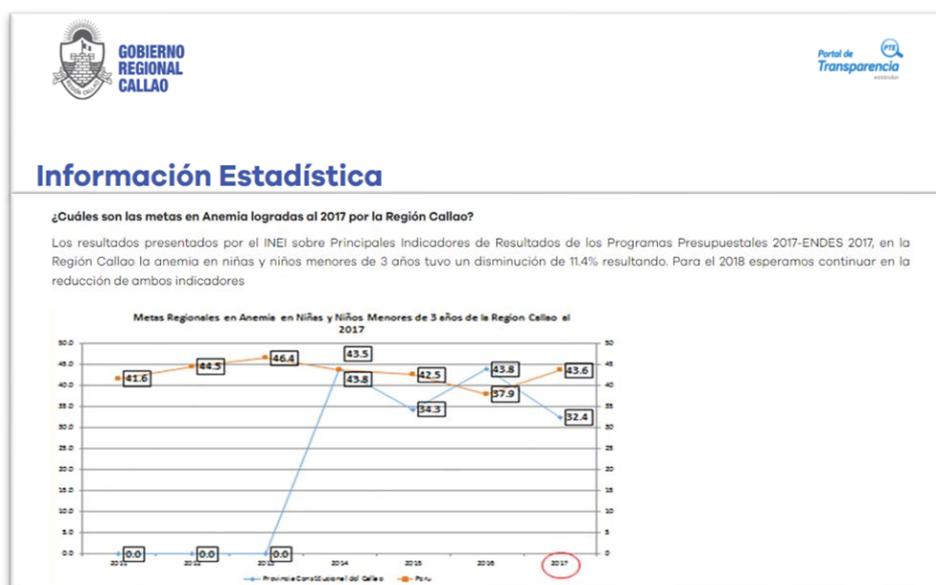
De lo mostrado, se puede mencionar que el Observatorio de Anemia del Gobierno del Perú, en junio de 2022, muestra las siguientes deficiencias:

- Datos desactualizados, ya que solo había datos hasta el 2019 y en algunos casos hasta el 2018.
- El acceso solo se permite desde enlaces de Perú, cuando se intenta ingresar desde otros países se muestra un mensaje de accesos restringidos.
- Si bien presenta una comparativa en relación con las metas regionales, no se puede tener visibilidad sobre los factores que influyen y cómo estos han evolucionado a través de los años.
- La mayor parte de los datos se encuentra en archivos PDF, lo que complica el tratamiento, y que, en la mayoría de los casos, no son reutilizables por su formato técnico, a pesar de ser datos abiertos.

2.3.2. Observatorios regionales

Se analizaron los observatorios regionales que se encontraron hábiles en la fecha de investigación, los cuales fueron de la región Callao y de la región Ancash (2 de las 25 regiones del Perú). Para el caso de la región Callao, la información presentada es estática y solo muestra globales, además que solo se tiene información hasta el año 2017 y en el caso de la región Ancash, a la fecha de redacción del presente informe (junio 2022), los datos relacionados con la prevalencia de anemia no estaban disponibles.

Ilustración 12. Observatorio de Anemia - Región Callao



Fuente: Observatorio de Anemia - Región Callao⁹

Ilustración 13. Observatorio de Anemia - Región Áncash: Prevalencia de anemia



Fuente: Observatorio de Anemia - Región Áncash¹⁰

⁹ El portal puede ser consultado en: <https://web.regioncallao.gob.pe/observatorio-regional-para-la-reduccion-de-la-anemia-y-desnutricion-cronica-infantil/cuales-son-las-metas-en-anemia-propuestas-para-la-region-callao-al-2021/>

¹⁰ El portal puede ser consultado en: http://web.regionancash.gob.pe/gerencia_social/observatorio/prevalencia_anemia

2.3.3. Plataforma Nacional de Datos Abiertos

La plataforma de datos abiertos de Perú, a junio 2022, ha puesto a disposición de la ciudadanía 2.359 conjuntos de datos, que corresponden a diversas categorías e incluyen la información de Ministerios públicos e instituciones del estado.

Ilustración 14. Plataforma Nacional de Datos Abiertos de Perú



Fuente: Plataforma Nacional de datos abiertos de Perú¹¹

Lo que ha podido notar en este portal es que no es amigable y la información no está estructurada, por ejemplo, no se puede tener visibilidad sobre el año o años a los que corresponde la información. Además, se puede notar que muchos de estos datos se encuentran publicados en otros portales y estos últimos tienen incluso actualizaciones posteriores. Como se ha mencionado previamente, hay una falta de gobernabilidad y en ocasiones datos incongruentes.

2.4 Propuesta y contribución de la investigación

Si bien todos los estudios de referencia incluyen la evolución anual del índice de anemia global del país y en algunos casos se incluye la evaluación de la situación en un momento determinado, no se ha podido identificar un análisis específico que permita evaluar la correlación de las variables a través de los años. Tampoco se han encontrado estudios que combinen estas variables con la inclusión de evaluaciones de la ejecución presupuestal.

¹¹ La plataforma puede ser consultada en: <https://www.datosabiertos.gob.pe/>

Adicionalmente, se está incluyendo una base de datos que contiene la información obtenida y tratada de las diversas fuentes en estudio.

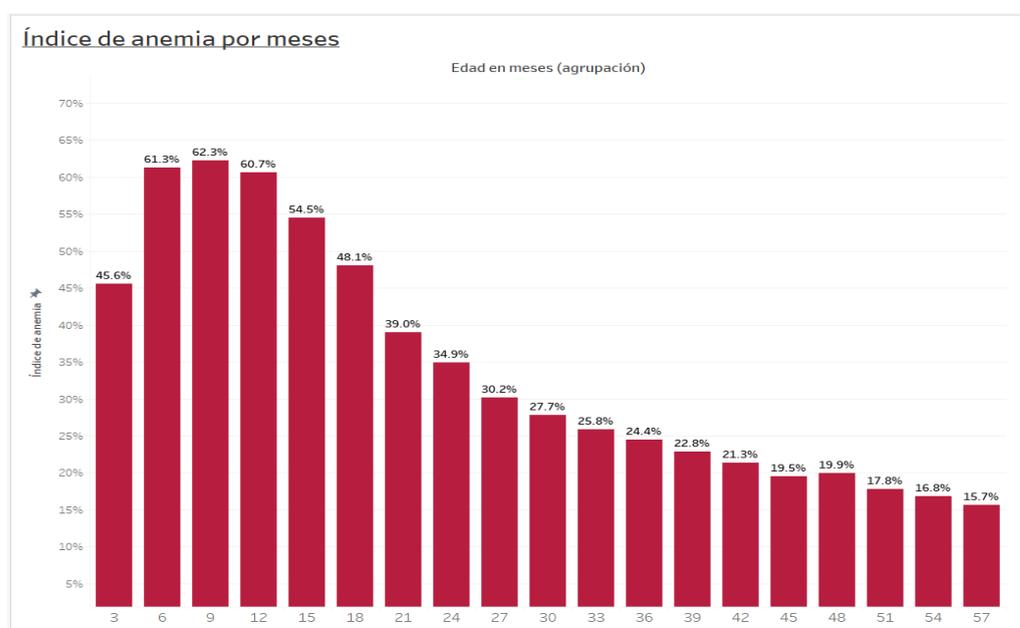
Tanto la información de base de datos, como también los tableros de control, serán compartidos en datos abiertos, de modo que pueda servir para que futuras investigaciones puedan ser más eficientes y así se puedan generar nuevos análisis, identificar otros aspectos relevantes o realizar nuevas visualizaciones de acuerdo con las necesidades presentadas. De esta forma, la base de datos se podrá ir nutriendo poco a poco, mejorando la calidad de esta y por ende, mejorando los análisis relacionados.

¿Porque el rango de edad de 6 a 35 meses?

Se ha seleccionado a esta población por tres motivos principales:

- 1) De acuerdo con estudios que investigan el desarrollo cognitivo, se conoce que desarrollo del cerebro ocurre en los tres primeros años de vida, siendo éste, el eje central del desarrollo integral del niño.
- 2) Como se muestra en la Ilustración 15, hay una mayor incidencia desde los 6 meses y se toma hasta que el índice supere los 25 puntos porcentuales.

Ilustración 15. Índice de anemia por meses



Fuente: INEI - ENDES 2009-2021

Elaboración propia

- 3) La tercera y última razón es que este rango de edad es la población objetivo en la mayor parte de políticas de estado y programas sociales en el Perú.

CAPÍTULO III: Análisis del problema

Como ya se ha mencionado en los capítulos anteriores, el índice de anemia en el Perú es un factor preocupante y, a pesar de los esfuerzos priorizados que el estado está poniendo en este tema, los resultados no son equiparables a los objetivos planteados, como puede verse en la Ilustración 16.

Ilustración 16. Índice de anemia real vs objetivo



Fuente: INEI – ENDES Del 2009 al 2022 y MIDIS – PLAN MULTISECTORIAL DE LUCHA CONTRA LA ANEMIA
Elaboración propia

La situación no es alentadora, ya que, en los últimos 10 años, ha permanecido con un indicador que no se ha podido reducir ni en 5 puntos porcentuales. Sin embargo, esta situación representa una oportunidad para buscar herramientas que permitan tener una mejora sustancial en los años que vienen, además de que hoy en día, con la situación que se vive por la pandemia, hay una gran cantidad de empresas privadas que están sumando esfuerzos para poder contribuir recursos y programas que permitan hacer frente a los problemas nutricionales en la población infantil.

Como un compendio de todo lo identificado en el análisis realizado, a continuación las principales situaciones que presentan oportunidades de mejora:

- No se ha podido evidenciar que se realicen trabajos colaborativos. Cada investigación, previo al análisis, invierte un tiempo relevante en el procesamiento de datos, aun cuando la mayoría de los análisis o investigaciones trabajan sobre los mismos datos.
- La mayoría de las investigaciones realizadas toman como base la información de un año específico, sin ver la evolución de los factores relevantes a lo largo de los años.

- Las investigaciones que tienen presentaciones evolutivas no tienen actualizaciones, según lo referido por algunos autores, esto se debe a la complejidad de procesar los datos actualizados.
- La mayoría de las evaluaciones son a nivel macro (país) y no se encuentran muchos estudios o plataformas que incluyan un despliegue detallado por regiones.
- La información de los distintos entes del estado no se encuentra integrada y los datos representan distintos grupos poblacionales, por lo que es difícil hacerla comparable.
- La información de algunos ministerios públicos aún no está abierta a la población. La transición hacia un gobierno de datos abiertos todavía no está priorizada.
- Aún se tienen pendientes tareas para anonimizar los datos. Esta situación limita la disposición de los recursos en entornos de datos abiertos.
- Hay muchos datos o documentos que solo se puede consultar por una página web o que se presentan como datos no estructurados, con formatos PDF o gráficos.
- Los datos que son publicados se encuentran segmentados en muchos archivos, por módulos, año e incluso tablas específicas. Además, presentan diferentes estructuras y formatos.
- No se publican datos procesados, por ello, los datos presentados no son de fácil interpretación y su procesamiento representa un reto, así como también un tiempo alto de gestión.
- La mayoría de las estadísticas o visualizaciones presentadas son estáticas y no permiten tener visibilidad multivariable, multinivel, ni hacer modificaciones en un rango de años.
- Para los datos que, por la ley de transparencia son públicos, pero no están difundidos, se tarda mucho en crear cuentas de usuarios de acceso o en responder a las solicitudes.
- El indicador de impacto de los datos abiertos tiene un valor muy bajo (6,67/100). Con ello se interpreta que el uso de los datos publicados no ha generado repercusiones positivas significativas en el país.

Como se observa, aún existe un amplio ámbito de mejora en cuanto a la presentación y disposición de los datos recopilados y los conjuntos de los datos procesados. Es importante que los datos sean de fácil acceso y comprensión, para que así sean útiles para fomentar el trabajo colaborativo y su reutilización.

También es importante que las personas interesadas en el análisis tengan la posibilidad de realizar modificaciones y actualizaciones relacionadas al tema que están analizando y al que quieren hacer frente. El no tener visualización adecuada de la información dificulta mucho para realizar un análisis más profundo de la situación y puede ocasionar que se tomen decisiones equivocadas cuando se implementen los planes de acción.

3.1 Análisis del marco legal

En este ítem se incluye un análisis de las restricciones de uso de la data utilizada previo a la publicación. Las fuentes consultadas tienen licencias de uso no comercial y en el caso de la fuente principal de datos que es el INEI, tiene una licencia de uso “Open Data Commons Open Database License (ODbL)”, la cual permite:

- Compartir: Para copiar, distribuir y utilizar la base de datos.
- Crear: Producir obras a partir de la base de datos.
- Adaptar: Modificar, transformar y construir sobre la base de datos.

Siempre que se atribuya y se mantenga el uso de la misma licencia para las bases de datos utilizadas, incluso las que tengan adaptaciones, como es el caso del presente trabajo. Para cumplir con este punto, se incluirá en la página web la aclaración indicando el tipo de licencia por cada base publicada (*OPEN KNOWLEDGE FOUNDATION sin fecha*).

Un segundo punto importante para tomar en cuenta en esta sección sería que cuando se comience a almacenar datos personales de usuarios, tales como nombres y correo, se tenga en cuenta la ley de protección de datos personales que rige en Perú. Este punto se desarrollará en la siguiente etapa de este proyecto, ya que para esta primera, no está en el alcance.

3.2 Análisis de riesgos

Uno de los riesgos principales identificados que podría afectar al logro de los objetivos planteados, es que las instituciones que son fuentes de información restrinjan o dejen de publicar los datos que se requieren para brindar el servicio, por ejemplo, en una de las cláusulas del INEI, indica que *“no se compromete a la actualización inmediata de la información y se reserva el derecho de modificar, aumentar, actualizar, cambiar o corregir cualquier omisión o información contenida en el presente Website en cualquier momento”*. Como plan de acción para esta situación, se propone un plan de contingencia, es decir, si se presenta la situación, se procederá con la solicitud de los datos a los entes pertinentes y bajo la ley de transparencia¹², toda persona tiene derecho de solicitar y recibir información de cualquier entidad de la administración pública.

Otro riesgo importante es que debido al gran volumen de datos que se espera alojar, se tenga procesos de consulta con tiempos prolongados y que ello haga que la experiencia de usuario no sea buena, con lo que se perdería un público objetivo importante. Para este punto, se propone realizar un dimensionamiento de crecimiento y ver cómo abordar esta situación para que se pueda brindar el mejor servicio.

¹² Ley N° 27806: Ley de transparencia y acceso a la información pública

3.3 Identificación y análisis de soluciones alternativas

Para poder hacer frente a los problemas expuestos, se han planteado las siguientes posibles soluciones:

A. Con relación al acceso y disponibilidad de información más legible:

Opción 1: La primera posibilidad sería realizar la solicitud de implantación de esta funcionalidad a las instituciones que manejan los datos relacionados al tema de estudio; sin embargo, creemos que es una opción poco viable, ya que no hay regulación, ni objetivos priorizados que oriente a las instituciones a ejecutar esta tarea adicional. Como prueba de concepto, se ha realizado la sugerencia y solicitud a una de las instituciones relacionadas, pero a la fecha de redacción del presente informe no se ha recibido respuesta a pesar de que ya han transcurrido más de 2 meses.

Opción 2: Replicar la base de datos que contiene los datos que se presentan en el portal e implementar una interfaz de usuario para que realice consultas SQL, de modo que cada usuario pueda realizar las consultas de acuerdo con su necesidad específica. A pesar de que es una solución potente, se considera que no es adecuada, debido a que el público objetivo se limitaría a las personas que tienen conocimientos de lenguaje SQL, lo cual no es viable para el objetivo de trabajo colaborativo planteado.

B. Con relación al análisis y presentación de visualizaciones que aporten valor para la toma de decisiones, se plantea:

Opción 1: Emitir comunicaciones periódicas para todos usuarios que requieran de dicha información y se inscriban para el envío correspondiente. Estos informes serían estáticos y en formato PDF. Con esta opción no se permiten cambios y la reutilización es mínima.

Opción 2: Trabajar con cuadros de mando y publicarlos en el repositorio de la propia herramienta. De este modo, podría ser consultada por cualquier persona interesada, pero el inconveniente sería que habría poca o ninguna retroalimentación con el trabajo que ellos realicen y además el alcance se limitaría a las personas que conozcan y utilicen la herramienta.

3.4 Solución propuesta

Al evaluar las alternativas de solución, se opta por el procesamiento de los datos a través de código propio y para las visualizaciones de datos, se utilizará una herramienta de business intelligence que permita realizar visualizaciones interactivas de los datos recopilados. Estas visualizaciones serán publicadas en una página web que sirva de repositorio y también que permita hacer la descarga de los datos fuente, además de los tableros de mando trabajados.

También es importante que los frentes interesados en el análisis puedan tener posibilidad de realizar modificaciones y actualizaciones y también, puedan retroalimentar el trabajo realizado. Es importante entender que el no tener visualizaciones acordes a la necesidad, limita mucho el análisis y puede generar que los planes o programas implementados no sean los más adecuados.

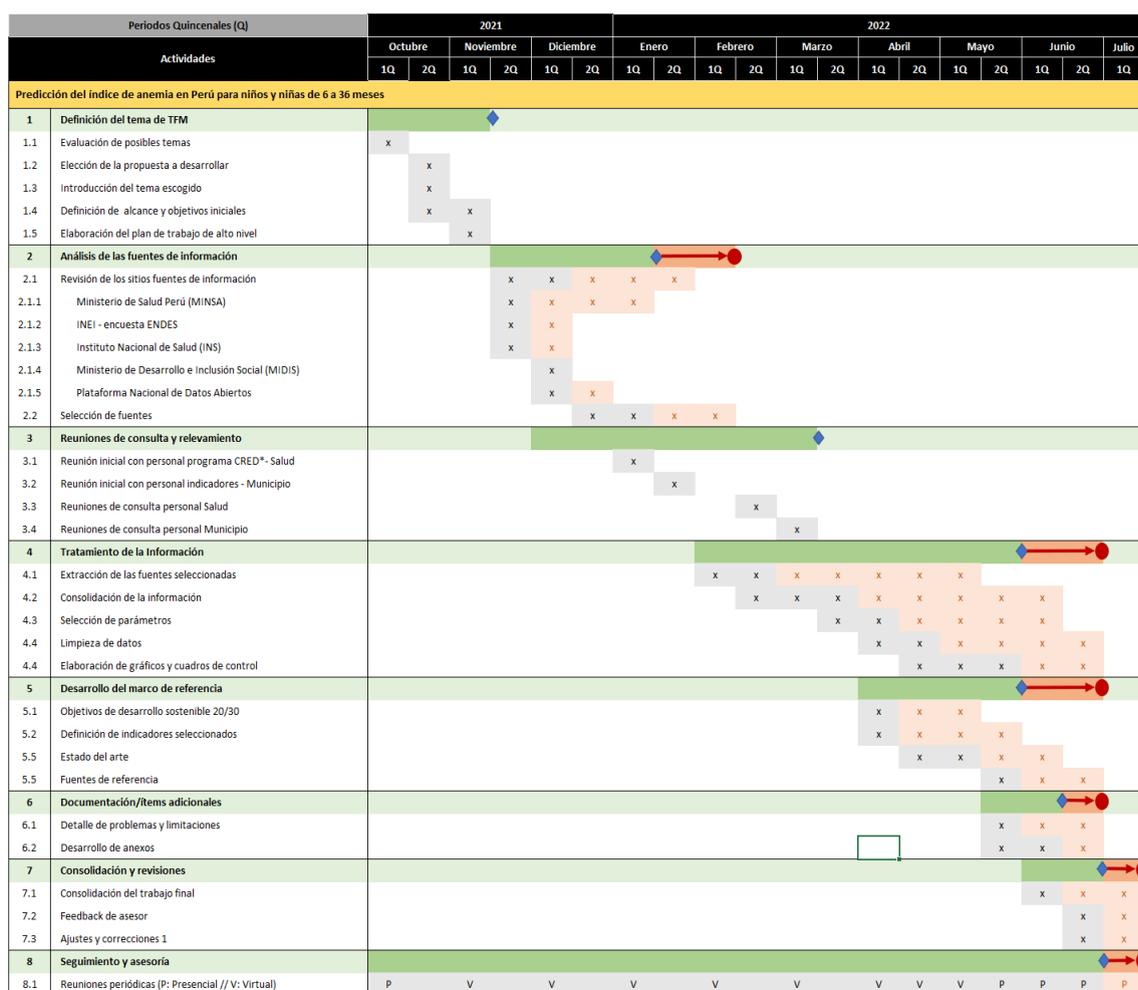
CAPÍTULO IV: Diseño de la solución

A continuación, se describen los pasos necesarios para poner en marcha este proyecto.

4.1 Planificación de las actividades:

En la Ilustración 17 se presenta el plan del proyecto, tomando en cuenta los principales hitos. Los colores verde, azul y plomo representan la planificación inicial, mientras que los colores naranja y rojo, representa las reprogramaciones que fueron necesarias realizar.

Ilustración 17. Cronograma de actividades inicial vs reprogramado



Leyenda	Hito	Actividad	Tarea
Planificado	◆	■	X
Reprogramado	●	■	X

Principales motivos de retrasos:

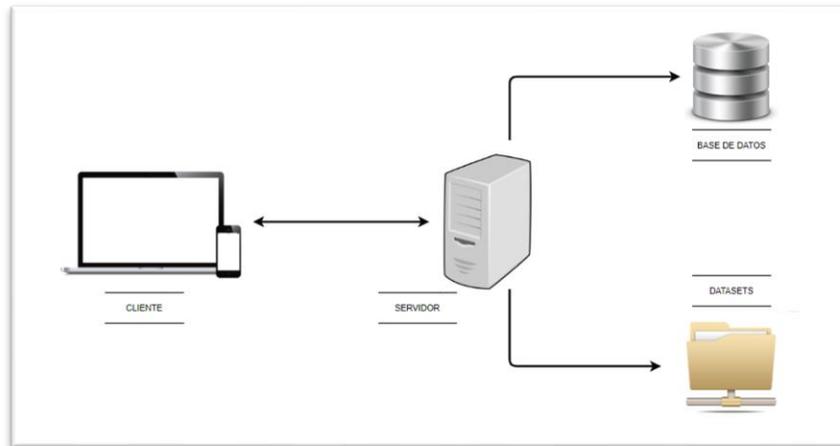
- * Dificultad de acceso a las fuentes de datos.
- * Reprocesos en el tratamiento de los datos por omisión o por corrección.
- * Dificultades en el uso de herramientas nuevas.

Elaboración propia

4.2 Arquitectura del sistema

La aplicación web en donde se comparten los datos, tendrá una arquitectura de cliente-servidor. En el lado del cliente se mostrará al usuario la web con los conjuntos de datos, documentación relacionada (como diccionarios o fichas técnicas) y también las visualizaciones previamente trabajadas; mientras que por el lado del servidor, se realizará la comunicación con la base de datos o con el servidor de archivos donde se almacenará los Datasets trabajados.

Ilustración 18. Arquitectura de la solución

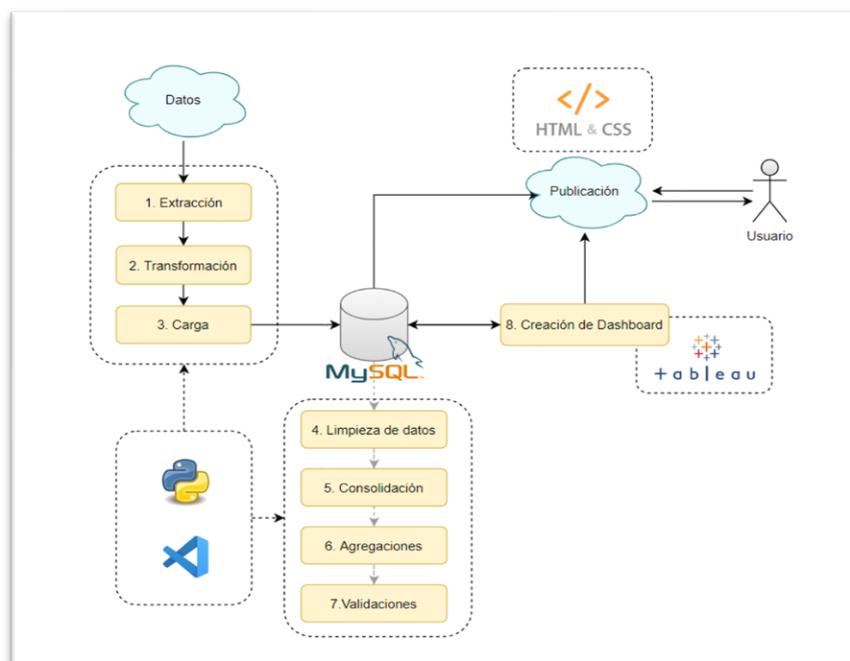


Fuente: elaboración propia

4.3 Diseño del proceso

En la Ilustración 19 se muestra el detalle de los pasos requeridos para la puesta en marcha:

Ilustración 19. Modelo de flujo de datos y principales herramientas



Elaboración propia

Como se puede apreciar, el diseño propuesto es sencillo y parte de la extracción de datos desde distintas fuentes de datos en la web, para transformarlos y cargarlos en una base de datos relacional, en donde se procede, mediante validaciones, a limpiar los datos para luego consolidarlos. Posterior a estos pasos se realiza la creación o agregación de columnas que se consideran importantes y luego se realiza las validaciones finales en contraste con indicadores conocidos (a partir de publicaciones de las instituciones que son fuentes de datos), entonces se procede a la creación y publicación de tableros de mando. Cabe mencionar que también se publican los datos de origen y su documentación correspondiente. El objetivo es que el usuario pueda consumir recursos, pero a la vez también retroalimentar las publicaciones.

4.4 Tecnología utilizada

 <p>Visual Studio Code</p>	<p>Un editor de código fuente, desarrollado por Microsoft, que es libre y multiplataforma: compatible con Windows, Linux, macOS y Web. De gran demanda por sus diferentes bondades en relación con la personalización, multilenguaje de programación, amplio uso de extensiones (plugins) y funciones de sugerencias de código, depuración y control de versiones. Además, tiene una excelente integración con Git y en la actualidad, según una encuesta realizada por Stack Overflow, es el más utilizada cubriendo el 71,6% de la demanda (<i>Stack Overflow 2021</i>).(<i>Stack Overflow Developer Survey 2021 sin fecha</i>)</p>
 <p>Python 3</p>	<p>Es un lenguaje de programación de código abierto que en la actualidad es uno de los más utilizados. Algunas de las ventajas de esta herramienta son las siguientes: tiene un lenguaje de alto nivel, por lo cual la curva de aprendizaje es baja; es polivalente, porque se utiliza para diversos propósitos; de paradigmas múltiples, porque admite programación estructurada, funcional y orientada a objetos; es multiplataforma y portátil, ya que es compatible con todos los sistemas operativos y tiene una comunidad muy sólida y colaborativa que sigue aportando soluciones día a día.</p>
	<p>Es un “micro” Framework escrito en Python que sirve para facilitar el desarrollo de aplicaciones Web. Entre las principales ventajas es que proporciona una estructura de proyecto, además brinda documentación robusta para tener una mayor facilidad de</p>

<p>Flask</p>	<p>aprendizaje y colaboración.</p>
	<p>Sistema gestor de base de datos relacional que tiene una versión de licencia en código abierto. Opera en arquitectura cliente – servidor, es compatible con lenguaje SQL y permite la creación de vistas, procedimientos almacenados y disparadores.</p>
 <p>WampServer</p>	<p>Soluciones para Windows que entre sus funcionalidades permite simular un servidor en tu ordenador personal, para así poder probar, por ejemplo, conexión con base de datos, simulando estar en un servidor, pero desde el localhost.</p>
 <p>HTML & CSS</p>	<p>Son los principales lenguajes de programación para la creación de páginas y aplicaciones web. HTML es el lenguaje de marcado que se utiliza para la creación de la web, mientras que CSS se utiliza para dar el estilo y diseño de la página. También se incluye JS útil para la activación de botones o mensajes emergentes.</p>
 <p>Bookstrap</p>	<p>Framework multiplataforma que brinda funcionalidades para crear un diseño moderno y adaptable en una página web. Funciona mediante una extensa librería de clases. Una de las características principales es su sistema de maquetar por columnas, de modo que permita una adecuada esquematización.</p>
 <p>Google Maps API</p>	<p>API con funcionalidad para la obtención de datos geolocalizados. Es gratuita para la cantidad de consultas equivalente a \$200 al mes, lo cual es equivalente, según menciona Google, a 100.000 mapas estáticos, rango significativamente amplio para trabajos puntuales o consultas no recurrentes. Para este trabajo, es utilizada para la obtención de coordenadas de latitud y longitud de más de 2500 registros.</p>
 <p>Font Awesome</p>	<p>Framework de licencia libre que proporciona una biblioteca de iconos vectoriales y estilos CSS. Se utilizará para dar mejor apariencia a los enlaces de redes sociales y otros iconos similares que se requieran. Es de muy fácil uso y al ser íconos vectoriales permite escalar sin problemas de resolución.</p>

	<p>Potente plataforma de visualización que brinda herramientas para organizar de mejor manera los datos. Entre sus principales ventajas es que es intuitivo, fácil de utilizar, permite múltiples fuentes de datos. Además, su versión Tableau Public es gratuita y como estudiantes, también se tiene licencia gratuita para la versión Desktop.</p>
 Git	<p>Software de código abierto que permite mantener el control de versiones para una buena trazabilidad de los cambios efectuados en los desarrollos de software. Entre sus principales bondades, se pueden mencionar un sólido rendimiento y gran seguridad que permita conservar la integridad del código fuente.</p>
 GitHub	<p>Plataforma de desarrollo colaborativo que permite alojar proyectos utilizando el sistema de control de versiones Git. Actualmente es la más utilizada para proyectos de código abierto. Al igual que Git, permite ver temas de trazabilidad, así como también reforzar el trabajo colaborativo que se quiere fomentar.</p>
 POSTMAN	<p>Herramienta que permite realizar peticiones del tipo REST, para poder tener comunicación con las diversas webs de obtención de datos. Usualmente esta herramienta es utilizada para realizar pruebas y exploraciones iniciales, pero también, permite la automatización de peticiones.</p>
 ILUSTRATOR	<p>Software de edición de imágenes de la suite Adobe, que permite crear o modificar, de manera creativa, imágenes. En este proyecto se utiliza para crear el logo que represente de alguna manera el objetivo de la aplicación, así como también, realizar ediciones a algunas imágenes que se incluyen en la página.</p>
 pythonanywhere	<p>Servicio en la nube para despliegue de aplicaciones web realizadas con código Python. Permite integración sencilla con GitHub y tiene un entorno de desarrollo integrado que permite actualizaciones en línea. En este proyecto, se utilizó este servicio para alojar y publicar en internet, el sitio web de SanitoPe.</p>

CAPÍTULO V: Desarrollo de la solución

En esta sección se detalla la puesta en práctica de lo mencionado en apartados anteriores. Es importante recordar que el proceso de implementación que se ha optado es por una metodología iterativa y que lo presentado este trabajo representa solo un mínimo viable y no es el producto finalizado. Siguiendo los pasos planteados en la metodología de desarrollo descrita en el capítulo 1, acápite 1.6, se realizan los siguientes pasos:

5.1 Selección de variables

Con relación a la bibliografía consultada e investigaciones previas realizadas, así como también de la consulta con especialistas del sector sanitario, tales como enfermeras que tienen a su cargo los programas de control del crecimiento y desarrollo de la niña y el niño (CRED) y con el personal responsable del seguimiento de indicadores en las municipalidades de los gobiernos regionales, se realizó una selección de variables que representan factores significativos que influyen en el índice de anemia y que sería importante que se incluyan en el análisis. Los ítems seleccionados fueron:

A. Generales:

- Sexo: femenino - masculino
- Edad: grupo de edades que seccionan el rango etario del alcance
- Orden de nacimiento: el número de hijo que representa para la familia.
- Tiempo entre nacimientos: el intervalo de tiempo desde el nacimiento previo.

B. Demográficos:

- Zona de residencia: urbana - rural.
- Región natural: Lima metropolitana¹³, resto de costa, sierra y selva
- Región política: división por regiones, provincias y/o distritos.

C. Económicos:

- Índice de riqueza: división por quintiles de acuerdo con el estatus económico.

D. Formativos:

- Grado de educación del responsable del cuidado: que en el Perú en la mayoría de los casos suele ser la madre la persona encargada del bebé.
- Asistencia a consejería: si o no. Hace referencia a las charlas de capacitación sobre alimentación que se brindan en los de salud.

¹³ Lima metropolitana forma parte de la costa peruana, pero debido a que es capital de país y representa el área metropolitana más grande, extensa y poblada del país, se evalúa por separado. Además, presenta una oportunidad de acceso y desarrollo, distinto al resto de la costa.

E. Salud:

- Con seguro de salud: si o no. Es el indicador de cobertura tanto de seguros públicos, como privados.
- Consumo de suplemento de hierro y vitamina A: si o no. Indica la cobertura de uno de los principales planes de acción implementados por el gobierno.
- Peso y tamaño al nacer: como referencia del resultado del proceso neonatal.
- Anemia en la madre: si o no. En conjunto con el indicador de bajo peso, representan indicadores del proceso neonatal y además de la concientización de la madre.
- Lactancia Materna: si o no. Se analiza el proceso de cobertura de lactancia materna en los primeros 6 meses de vida.

Adicionalmente, se considera conveniente incluir la evaluación y comparativas del alcance en función a provincias y distritos, así como también la comparativa con los objetivos planteados, programas sociales y ejecución presupuestal.

5.2 Identificación de fuentes de datos

En Perú, las instituciones oficiales que gestionan la data relacionada con los indicadores de anemia son:

- INEI: Instituto Nacional de Estadística e Informática
- MINSA: Ministerio de Salud
- MIDIS: Ministerio de desarrollo e inclusión social
- INS: Instituto Nacional de Salud.
- Plataforma Nacional de Datos Abiertos.

Se realizó un análisis preliminar de la información que proporcionan cada una de estas instituciones, de modo que se pueda tener información que sea útil y que contenga la población total o una muestra representativa de acuerdo con la distribución geográfica. También se realizó la verificación de que se encuentren datos relacionados a los factores que se han considerado relevantes para realizar el análisis, los cuales se encuentran descritos en el punto previo.

En este punto se identificó que solo el INEI y el INS-MINSA contaban con estos datos y desde fuentes distintas. Las demás instituciones citadas en este acápite hacen uso de los datos de una de estas dos instituciones. Para ambas instituciones se identificó que mantienen un repositorio de datos con acceso abierto.

5.3 Selección de las fuentes óptimas

Sobre el análisis realizado en el punto previo, se realizó el análisis MELODA 5 a las dos fuentes preseleccionadas. Si se quiere tener un entendimiento mejor sobre los niveles por dimensión, se puede consultar el *Anexo 2: "Dimensiones y niveles MELODA"*. A continuación, en las Tablas 1 y 2 se detalla la evaluación realizada.

Tabla 1. Análisis MELODA: Data INEI - ENDES.

DataSet	INEI - ENDES		
Enlace	http://inei.inei.gob.pe/microdatos/		
Dimensiones	Nivel alcanzado	Detalle	Puntaje
<i>Legal licensing</i>	Non-commercial reuse	Open Data Commons Open Database License (ODbL)	3
<i>Access to information</i>	Web access or unique URL parameters to dataset	Una única URL para descarga	1
<i>Technical standard</i>	Open standard reusable	Formatos: .DAT, .DBF, .SAV, .CSV	3
<i>Standardization</i>	Own ad hoc data model standardization published	Tiene anexo la ficha técnica y el diccionario de datos	3
<i>Geolocation content</i>	Simple or complex text field	Tiene data de región, provincia y distrito (descriptivo)	3
<i>Updating frequency</i>	Longer than 1 month	Anual (Desde el 2009)	1
<i>Dissemination</i>	Available resources on updates	Actualizaciones se difunden directamente en redes y otros medios	3
<i>Reputation</i>	Statistics or reports published on users opinions	El estado y la mayoría de las investigaciones se basan en estos datos	3
Puntuación Total	20		
<i>Fecha del Análisis</i>	14/02/2022		
<i>Aspectos positivos</i>	Se tiene información detallada desde el año 1996 y con frecuencia anual desde el 2009		
<i>Aspectos negativos</i>	Se han detectado que datos de año a año difieren.		

Elaboración propia

Tabla 2. Análisis MELODA: Data INS - MINSa.

DataSet	INS - MINSa		
Enlace	https://www.datosabiertos.gob.pe/dataset/sistema-de-informaci%C3%B3n-del-estado-nutricional-de-ni%C3%B1os-y-gestantes-per%C3%BA-inscenen-instituto		
Dimensiones	Nivel alcanzado	Detalle	Puntaje
<i>Legal licensing</i>	Non-commercial reuse	Open Data Commons Attribution License	3
<i>Access to information</i>	Web access or unique URL parameters to dataset	Una única URL para descarga	1
<i>Technical standard</i>	Open standard reusable	Formato: .CSV	3
<i>Standardization</i>	Own data model standardization	La data es entendible, pero no tiene diccionario.	1
<i>Geolocation content</i>	Simple or complex text field	Tiene data solo de la región (descriptivo)	3
<i>Updating frequency</i>	Longer than 1 month	No especifica (publicada 20.08.2021)	1
<i>Dissemination</i>	Available resources on updates	No se tiene evidencia de difusión sistemática	1
<i>Reputation</i>	Statistics or reports published on users opinions	El estado y algunas investigaciones se basan en estos datos	3
Puntuación Total	16		
<i>Fecha del Análisis</i>	14/02/2022		
<i>Aspectos positivos</i>	Se tiene información detallada de los niños evaluados y se encuentra anonimizada.		
<i>Aspectos negativos</i>	Solo se tiene información de los años 2019, 2020 y 2021		

Elaboración propia

Adicional al análisis MELODA, también se realizó la evaluación para ver la cobertura de las variables que son relevantes para el análisis. En la tabla 3 se presenta el detalle:

Tabla 3. Evaluación de cobertura de variables

Criterio de evaluación/variables	INEI - ENDES	INS - MINSA
Generales		
Sexo	Sí	Sí
Edad	Sí	Sí
Orden de nacimiento	Sí	No
Tiempo entre nacimientos	Sí	No
Demográficos		
Zona de residencia	Sí	Sí
Región natural	Sí	Sí
Región política	Sí	Sí
Económicos		
Índice de riqueza	Sí	No
Formativos		
Educación del responsable del cuidado	Sí	No
Asistencia a consejería	No	Sí
Salud		
Con seguro de salud	Sí	Sí
Suplemento de hierro y vitamina A	Sí	Sí
Bajo peso al nacer	Sí	Sí
Anemia en la madre	Sí	Sí
Lactancia Materna	Sí	Sí
Otros		
Programas sociales	Sí	Sí
Datos municipales	No	No
Objetivos estratégicos	No	No
Ejecución presupuestal	No	No

Elaboración propia

A partir de la puntuación obtenida en MELODA, así como también la evaluación de la cobertura de variables, se seleccionó como fuente de datos la ENDES que son publicadas por el INEI. No obstante, la información contenida en la base del INS – MINSA, no es descartada, solo que no será considerada en esta primera etapa. Para los ítems contenidos en la categoría “Otros” que no son proporcionados por esta fuente de datos, pero que son importantes para el análisis, se utilizarán las siguientes fuentes de datos alternas:

- Datos de municipales: Se obtendrá esta data de la encuesta del RENAMU¹⁴ - INEI.
- Objetivos Estratégicos: Se obtendrá de la información publicada por el MIDIS en el documento denominado PLAN MULTISECTORIAL DE LUCHA CONTRA LA ANEMIA.
- Ejecución presupuestal: Se obtendrá del MEF¹⁵, a través de su web CONSULTA AMIGABLE.

¹⁴ RENAMU: Registro Nacional de Municipalidades. Acceso en: <http://iinei.inei.gob.pe/microdatos/>

¹⁵ MEF: Ministerio de Economía y Finanzas. Acceso en: <http://apps5.mineco.gob.pe/transparencia/>

5.4 Recopilación de datos

La fuente principal seleccionada es la Encuesta Demográfica y de Salud Familiar (ENDES), publicada por el Instituto Nacional de Estadística e Informática (INEI) en su portal web, como puede verse en la Ilustración 20:

Ilustración 20. Archivos del INEI - ENDES 2021

The screenshot shows the 'MICRODATOS BASE DE DATOS' portal. It includes a search filter for 'ENCUESTA DEMOGRÁFICA Y DE SALUD FAMILIAR - ENDES' for the year 2021 and period 'Único'. Below the filter is a table listing 13 modules with their respective codes and download options for SPSS, STATA, and CSV formats.

Nro	Año	Periodo	Código Encuesta	Encuesta	Código Módulo	Módulo	Ficha	Descarga
1	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1629	Características del Hogar	SPSS	STATA
2	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1630	Características de la Vivienda	SPSS	STATA
3	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1631	Datos Básicos de MEF	SPSS	STATA
4	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1632	Historia de Nacimiento - Tabla de Conocimiento de Métodos Embarazo, Parto, Puerperio y Lactancia	SPSS	STATA
5	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1633	Inmunización y Salud	SPSS	STATA
6	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1634	Nupcialidad - Fecundidad - Cónyuge y Mujer	SPSS	STATA
7	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1635	Conocimiento de Sida y uso del condón	SPSS	STATA
8	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1636	Mortalidad Materna - Violencia Familiar	SPSS	STATA
9	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1637	Disciplina Infantil	SPSS	STATA
10	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1638	Peso y talla - Anemia	SPSS	STATA
11	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1639	Encuesta de salud	SPSS	STATA
12	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1640	Programas Sociales	SPSS	STATA
13	2021	5	760	Encuesta Demográfica y de Salud Familiar - ENDES	1641	Programas Sociales	SPSS	STATA

Fuente: Portal de microdatos del INEI

Revisando la estructura de publicación de datos, toda la información se encuentra bajo una única URL y se encuentra dividida por años, periodos, módulos y tablas de datos, siendo el alcance el que se muestra en la Tabla 4:

Tabla 4. Tabla de estructura de datos del INEI - ENDES

Estructura de los datos	Cantidad
Años con información	20
Periodos por año (promedio)	1
Módulos por periodo (promedio)	11
Tablas por módulo (promedio)	3
Aproximado de archivos a procesar	660
Formatos de acceso: DBF, STATA, SPSS y CSV	

Elaboración propia

Como se puede apreciar en la Tabla 4, la cantidad de archivos a procesar es un número muy grande y representaría un número alto de horas de procesamiento manual, y considerando además que los archivos se descargaban en carpetas comprimidas, se decidió automatizar la labor de extracción. Para ello se identificó la estructura de los enlaces de descarga, y se obtuvo los datos de número de encuesta y módulo, así como también el tipo de documento/acceso (archivos dbf o dta) y se creó la Tabla 5 con dicha información (DB_Endes):

Tabla 5. Ejemplo de registros para descarga de información de la ENDES - INEI

Nro	Año	Período	Cod_Encuesta	Encuesta	Cod_Modulo	Módulo	Tipo de acceso
1	1996	5	32	Encuesta Demográfica y de Salud Familiar - ENDES	64	Características del Hogar	DBF
2	1996	5	32	Encuesta Demográfica y de Salud Familiar - ENDES	65	Características de la Vivienda	DBF
3	1996	5	32	Encuesta Demográfica y de Salud Familiar - ENDES	66	Datos Básicos de MEF	DBF

Elaboración propia

Con esta información se procedió a realizar la descarga en función a la estructura de la URL:

Ilustración 21. Código para conexión y descarga de archivos

```

conexion = create_engine('mysql+pymysql://root@localhost:3306/indiceanemia')
con = conexion.raw_connection()
cur = con.cursor()

def consulta_enc_mod (registros):
    for i in registros.index:
        Nro_Encuesta = str(registros.loc[i,'Cod_Encuesta'])
        Nro_Modulo = str(registros.loc[i,'Cod_Modulo'])
        año = str(registros.loc[i,'Año'])
        acceso = str(registros.loc[i,'Tipo de acceso'])
        file_url = 'http://inei.inei.gob.pe/inei/srienahto/descarga/'+acceso+'/'+ Nro_Encuesta +'-Modulo'+ Nro_Modulo +'.zip'
        get_zip(file_url,año)
    
```

Fuente: código desde Visual Studio Code

Debido a que los archivos se encuentran comprimidos en carpetas, que a su vez contenían otros tipos de archivos como diccionarios y fichas técnicas, se desplegó un código para descomprimir e identificar los formatos de los archivos válidos para la recopilación de información (en este caso, .dbf y .dta). A continuación, se muestra parte del código utilizado para esta sección:

Ilustración 22. Código para la descomprimir y leer archivos DTAs y DBFs

```

if '.dbf' in Zip_info.filename.lower() or '.dta' in Zip_info.filename.lower():
    myzipfile.extract(Zip_info, '/content/dbfs-dtas')
    ruta = '/content/dbfs-dtas/'+ Zip_info.filename
    if '.dbf' in Zip_info.filename.lower():
        dbf5 = Dbf5(ruta, codec='iso-8859-1')
        df = dbf5.to_dataframe()
    if '.dta' in Zip_info.filename.lower(): # En la web del INEI se tienen algunos .DAT
        df, meta = pyreadstat.read_dta(ruta)
    nombre = Zip_info.filename[0:(len(Zip_info.filename)-4)].lower()
    df.columns = df.columns.str.strip()
    df.columns = df.columns.str.replace('\x00', '')
    columnas_df = df.columns.str.lower()
    df['ano'] = str(año)
    df['fecha_extrae'] = str(datetime.now().date())
    tablas = pd.read_sql_query('SHOW TABLES',con=conexion)[['Tables_in_indiceanemia2']]
    
```

Fuente: código desde Visual Studio Code

Luego, se procedió a realizar verificaciones sobre existencia de la tabla, paridad de columnas y validaciones para evitar la duplicidad de tablas y también evitar duplicidad en los datos que se están ingresando. El código implementado para esta sección fue el siguiente:

Ilustración 23. Código para validaciones y carga en la base de datos

```
df_comprueba = tablas[tablas['Tables_in_indiceanemia'] == nombre]
if df_comprueba.empty:
    print("la tabla no existe, creamos!")
else:
    print("ya está creada, verificamos las columnas y fila")
    columnas_bd = pd.read_sql_query('SHOW COLUMNS FROM `{table_name}`'.format(table_name = nombre), con=conexion)[['Field']]
    info_año = pd.read_sql_query('SELECT ANO FROM `{table_name}` WHERE ANO = {ano}'.format(table_name = nombre, ano = año),
    con=conexion)

    for col in columnas_bd:
        df_prueba = columnas_bd[columnas_bd['Field'].str.lower() == col]
        if df_prueba.empty:
            print('agrega columna a la base de datos')
            print (type(col))
            if type(col) == "int":
                data_formatted = "INT"
            else:
                data_formatted = "VARCHAR(30)"

            base_command = ("ALTER TABLE `{table_name}` ADD column `{column_name}` {data_formatted} NULL")
            sql_command = base_command.format(table_name=nombre, column_name=col, data_formatted = data_formatted)
            cur.execute(sql_command)

        if info_año.empty:
            print("nuevo año a agregar!")
        else:
            print("año existe, eliminando registros e ingresando nueva data")
            base_command = ("DELETE FROM `{table_name}` WHERE ANO = {ano}")
            sql_command = base_command.format(table_name=nombre, ano = año)
            cur.execute(sql_command)

df.to_sql(name=nombre, con = conexion, if_exists = 'append')
```

Fuente: código desde Visual Studio Code

Cabe mencionar que lo presentado aquí solo es un resumen del código y no la funcionalidad completa. El código y las librerías importadas necesarias, se encuentra disponible en GitHub en la siguiente ruta: <https://github.com/KatiaMendez/AnemiaInfantilPeru> (De aquí en adelante, repositorio “AnemiaInfantilPeru” en GitHub), archivos: “02. DescargaDatos.py”.

A la información proveniente de la fuente original, se le adiciona la información del año al que pertenecen los datos y además también se incluye la fecha de extracción, ya que es importante mantener para tener trazabilidad. El resultado de esta etapa fue el siguiente:

Tabla 6. Resumen de carga inicial en base de datos

Ítem	Cantidad
Tablas creadas	60 ¹⁶
Cantidad de filas totales	19,471,801
Tamaño de la Base de Datos	4.8 GB
Tiempo aproximado procesamiento	10 horas

Elaboración propia

¹⁶ En la base de datos se tienen 62, pero 2 de ellas corresponde a las tablas iniciales creadas para poder realizar la extracción y crear los índices de tablas.

Ilustración 24. Carga inicial de base de datos

Nombre de tabla	Número de filas	Tamaño
rech5	473,685	128.6 MB
rech6	235,459	82.4 MB
rech8	135,435	28.1 MB
rech9	128,288	36.7 MB
rech10	73,154	9.7 MB
rech11	64,693	8.5 MB
rech23	561,758	278.6 MB
rechm	9,543	1.0 MB
recho	28,763	6.2 MB
rechobs	34,857	5.0 MB
Total	19,471,801	4.8 GB

Fuente: base de datos inicial

5.5 Limpieza y consolidación de datos

Esta sección se divide en dos partes principales. En la primera parte, se trabajará sobre el total de tablas descargadas y en la segunda parte, a partir del punto 5.5.2, será trabajada sobre los datos seleccionados para el análisis específico.

5.5.1. Quick wins sobre la data total

Si bien el objetivo de este punto no es llevar a una limpieza fina de los datos, es importante considerar que, si se tienen oportunidades de solucionar problemas claramente visibles y que no requieran de gran esfuerzo, se deben realizar, ya que como se mencionó previamente, esta información será compartida. En este caso, se realiza las acciones que se presentan en la Tabla 7:

Tabla 7. Quick wins sobre la data total

Actividad	Objetivo	Pasos	Ejecutada	Cantidad de casos
Crear índices	Mejorar la velocidad de consulta en los procesos de consolidación	Actualización del tipo de dato	✓	95
		Creación de los índices	✓	95
Consolidación inicial	Evitar la redundancia o errores en procesamiento	Unir tablas con nombres distintos	✓	9

Elaboración propia

El código de esta sección también se encuentra en el repositorio “AnemiaInfantilPeru” en Github, archivos: “03. CrearIndex.py” y “04. Correcciones.py”.

5.5.2. Limpieza sobre la base seleccionada

Para poder realizar este punto, se utilizó la selección de tablas y campos específicos que permitieron el análisis de las variables propuestas en el punto 5.1. Para ello se realiza con la lectura de las fichas técnicas y de los correspondientes diccionarios de datos. Del entendimiento obtenido, en la Tabla 8 se presenta el resumen de los datos necesarios para el análisis:

Tabla 8. Tablas y campos seleccionados

Estructura base de datos	Universo	Selección
Tablas promedio por año	33	13
Campos promedio año	2500	235

Elaboración propia

El detalle de las tablas seleccionadas y los campos correspondientes se puede encontrar en el Anexo III “Tablas y campos seleccionados”.

Luego se realiza la validación por tabla para ver que todas contengan la información de todas las columnas de referencia. Para ello también se utilizó código, el cual podrá encontrar en el repositorio “AnemiaInfantilPeru” en Github, archivos: “05.Validaciones.py”. Aquí se identificó los datos faltantes por tabla y por año, lo cual se aprecia en la Tabla 9:

Tabla 9. Validación de data por tabla y por año

AÑO	TABLAS												
	RECH1	RECH4	RECH6	RECH0	RECH23	PS_WA..	REC43	REC95	REC94	REC41	REC21	REC0111	RE22312
1996	✓	✓	✗ ⁵	✓	✗ ⁵	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2000	✓	✗ ⁵	✗ ⁵	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2004	✓	✓	✗ ⁵	✓	✗ ⁵	✗ ¹	✓	✓	✓	✓	✓	✗ ⁴	✗ ²
2005	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✗ ²
2006	✓	✓	✓	✗ ³	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2007	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✗ ²
2008	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2009	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2010	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2011	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2012	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2013	✓	✓	✓	✓	✓	✗ ¹	✓	✓	✓	✓	✓	✓	✓
2014	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2015	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2016	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2017	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2018	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2019	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2020	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2021	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Elaboración propia

Al evaluar estas casuísticas, se pudo identificar los siguientes escenarios y tomar acción sobre cada uno de ellos:

- Para el caso de la tabla PS_WAWAWASI (x^1), que contiene los datos del programa social más relevante para el análisis, se confirma que solo se cuenta con registros desde el año 2014. Se decide mantener, pero tomando en cuenta esa consideración.
- Para el caso de la tabla RE22312 (x^2), se valida que la información se encuentra en otras tablas denominada RE223132 y R22312. Se procede con la consolidación.
- Para el caso de la tabla RECHO (x^3), se valida que la información se encuentra en otras tablas denominada RECHO. Se procede con la consolidación.
- Para el caso de la tabla REC0111 (x^4), se valida que la información se encuentra en otras tablas denominada REC01_11. Se procede con la consolidación.
- Para los demás casos (x^5) se revisaron los diccionarios de datos y se confirmó que tales datos no existían. Por ser la tabla RECH6 tan importante para el análisis (aquí está contenido la medición de anemia en niños), se excluyen de la evaluación los años 1996, 2000 y 2004.

Ilustración 25. Vista de verificación de datos en tablas y años seleccionadas

```

25 def verificar_data (tablas,años):
26     for tabla in tablas:
27         for año in años:
28             info_año = pd.read_sql_query('SELECT ANO FROM `{table_name}` WHERE ANO = {año}'.format(table_name = tabla, año = año),
29             if info_año.empty:
30                 print("No hay información en la tabla: "+tabla+", para el año: "+año)
31
32
33 # Se brindan los valores que se quieren verificar que tengan información
34 tablas = ['RECH1','RECH4','RECH6','RECH8','RECH23','PS_WAWAWASI','REC43','REC95','REC94','REC41','REC21','REC0111','RE223132']
35 años = ['1996','2000','2004','2005','2006','2007','2008','2009','2010','2011','2012','2013','2014','2015','2016','2017','2018','20
36 verificar_data (tablas,años)
    
```

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL

```

No hay información en la tabla: RECH4, para el año: 2000
No hay información en la tabla: RECH6, para el año: 1996
No hay información en la tabla: RECH6, para el año: 2000
No hay información en la tabla: RECH6, para el año: 2004
No hay información en la tabla: RECH23, para el año: 1996
No hay información en la tabla: RECH23, para el año: 2004
No hay información en la tabla: PS_WAWAWASI, para el año: 1996
No hay información en la tabla: PS_WAWAWASI, para el año: 2000
No hay información en la tabla: PS_WAWAWASI, para el año: 2004
No hay información en la tabla: PS_WAWAWASI, para el año: 2005
No hay información en la tabla: PS_WAWAWASI, para el año: 2006
No hay información en la tabla: PS_WAWAWASI, para el año: 2007
No hay información en la tabla: PS_WAWAWASI, para el año: 2008
No hay información en la tabla: PS_WAWAWASI, para el año: 2009
No hay información en la tabla: PS_WAWAWASI, para el año: 2010
No hay información en la tabla: PS_WAWAWASI, para el año: 2011
No hay información en la tabla: PS_WAWAWASI, para el año: 2012
No hay información en la tabla: PS_WAWAWASI, para el año: 2013
PS C:\Users\Katia\Desktop\IndiceAnemia>
    
```

Fuente: código desde Visual Studio Code

Para los datos seleccionados se realizaron validaciones que podrá encontrar en el repositorio “AnemiaInfantilPeru” en Github, archivos: “05. Validaciones.py”.

Tabla 10. Limpieza y homologación de datos

Actividad	Ejecutada	Cantidad de tablas
Unión de tablas que difieren en nombre	✓	4
Homologar campos	✓	13
Retirar espacios en blanco.	✓	13
Reemplazar ceros por valor en blanco	✓	13

Elaboración propia

5.5.3. Consolidación de la base seleccionada

Con este nuevo alcance, se procede a unir las tablas. En este punto es importante mencionar que se hace la creación de tablas intermedias debido a que se requiere información sobre el padre o apoderado por cada niño evaluado. El script de esta sección se encuentra en el repositorio “AnemiaInfantilPeru” en Github, archivo: “Consolidado.sql”

Las consideraciones importantes para tomar en cuenta en esta etapa son:

- Cuando se cree una nueva tabla, se deben crear también sus índices correspondientes para así reducir los tiempos de consulta.
- En concordancia con la población objetivo, la tabla de consolidación toma como base el universo de niños y niñas seleccionados para la evaluación y se asigna a estos registros, la información de la persona responsable de ellos (padre o madre). En la consolidación no se tiene todos los registros que provee la fuente de datos inicial.
- Se realiza el filtro correspondiente al retiro de valores fuera del rango en medición del nivel de hemoglobina, así como también se depuraron los registros con valores nulos. No se realizan filtros de rangos de edad, ni otros factores ya que se espera que estos datos sean de utilidad para aquellos que quieren realizar evaluaciones con otros alcances en su análisis.

Finalmente, la tabla de la base de datos que se obtuvo se muestra en la Ilustración 26:

Ilustración 26. Tabla consolidada para análisis del alcance objetivo (data_indiceanemia)

Fuente: Base de datos MySQL del proyecto

Los datos importantes de este acápite son:

Tabla 11. Resumen de tabla consolidada “data_indiceanemia”

Ítem	Cantidad
Columnas	190
Filas	200,521
Tamaño de la tabla	185.6 MB
Tiempo aproximado consulta	0.0022 seg

Elaboración propia

5.5.4. Creación de campos auxiliares o complementarios

La información procesada es muestral, por ende, requiere que dicha muestra sea representativa a la estructura poblacional. Para poder cumplir con este punto, el INEI utiliza un criterio denominado “Factor de ponderación” (desde ahora FP), el cual se utiliza para reconstruir la muestra poblacional. Todos los años se cuenta con un FP, sin embargo, no siempre se ha almacenado en el mismo atributo, ni en la misma tabla. Para poder realizar una adecuada validación de este campo, se realizó la lectura de las fichas técnicas de todos los años y se obtuvo lo mostrado en la Tabla 12:

Tabla 12. Identificación del factor de ponderación¹⁷

Año	Tabla	Campo
2020	RECH6	HV005A
2015	RECH0	HV005X
Resto de años	RECH6	HV005

Elaboración propia

En función a este cuadro, se creó un campo calculado (FP), en el cual se asignaba el valor que le correspondía de acuerdo con el año. El código se encuentra en el repositorio “AnemiaInfantilPeru” en Github, archivo: “08.1 FPCalculados.py”.

Otros campos que fueron creados y alimentados fueron los correspondientes a equivalencias de acuerdo con el diccionario de datos y los específicos de anemia, grupo de edad, consumo hierro y seguro social:

Tabla 13. Campos adicionales creados

Columna	Descripción	Columna(s) referencia
HV104_name	Detalle del sexo	HV104
HC57_name	Nivel de anemia	HC57
HV024_name	Región política	HV024
HV025_name	Área de residencia	HV025
SHREGION_name	Región natural	SHREGION
HV270_name	Quintil de riqueza	HV270
Anemia	Valor binario “Sí” o “No”	HC57
edad_meses_grupo	Agrupación de la edad	HC1
Consumo_hierro_7d	Consumo de alguna forma de hierro	S465E*
Tiene_seguro_salud	Valor binario “Sí” o “No”	SH11*

Elaboración propia

El código de esta sección también se encuentra en el repositorio “AnemiaInfantilPeru” en Github, archivo: “09. CreaCamposDetalle.py”

Finalmente, como tema adicional identificado en este punto, fue que desde el 2005 al 2008,

¹⁷ Mayor detalle de los documentos relacionados, se podrán encontrar en el Anexo IV: Documentación sobre el factor de ponderación.

los informes fueron bianuales, por lo cual era necesario realizar la actualización:

- Para los años 2005 y 2006 se evidenció que era información repetida, por lo cual se depuraron los datos del año 2006.
- Para los años 2007 y 2008 se evidenció que era complementaria y se procedió con la unión de estas bases.

El código de esta sección también se encuentra en el repositorio “AnemiaInfantilPeru” en Github, archivo: “10. ActualizaAño.py”

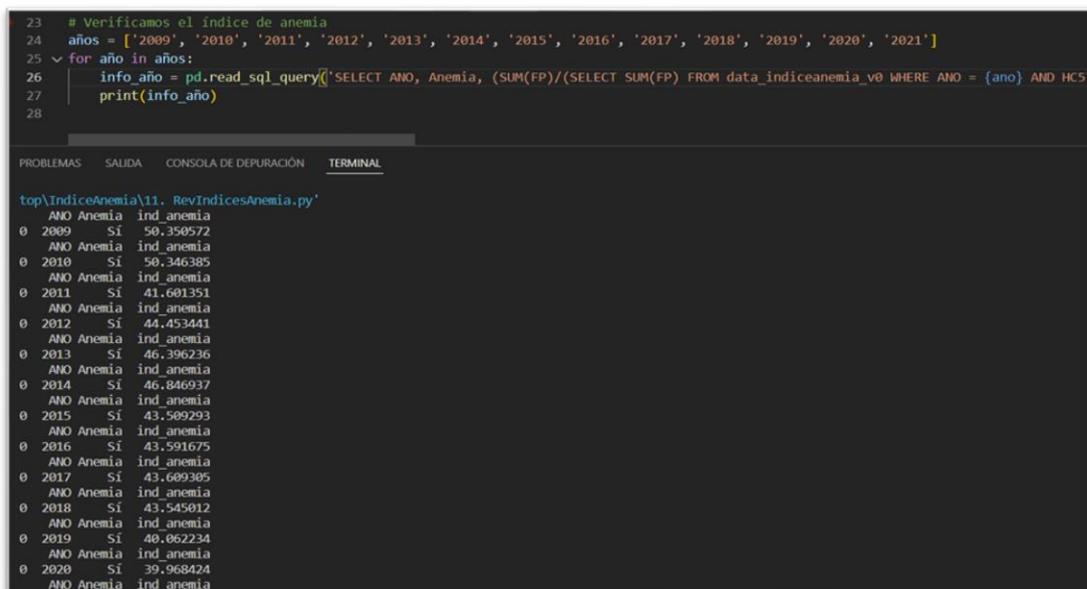
Por no ser representativa para una evaluación anual, se procede a retirar estos años del alcance del análisis¹⁸.

5.6 Validación de datos

Se realizan validaciones a nivel global de la base consolidada creada, para ello se verifica que haya información de todos los años del alcance y que los valores calculados del índice de anemia coincidan con los presentados por el INEI. En este punto, tal como se muestra en la Ilustración 27, la validación resultó satisfactoria.

Ilustración 27. Verificación del índice de anemia calculado

```
23 # Verificamos el índice de anemia
24 años = ['2009', '2010', '2011', '2012', '2013', '2014', '2015', '2016', '2017', '2018', '2019', '2020', '2021']
25 for año in años:
26     info_año = pd.read_sql_query('SELECT ANO, Anemia, (SUM(FP))/(SELECT SUM(FP) FROM data_indiceanemia_v0 WHERE ANO = {ano} AND HC57
27     print(info_año)
28
```



ANO	Anemia	ind_anemia
0 2009	SI	50.350572
0 2010	SI	50.346385
0 2011	SI	41.601351
0 2012	SI	44.453441
0 2013	SI	46.396236
0 2014	SI	46.846937
0 2015	SI	43.509293
0 2016	SI	43.591675
0 2017	SI	43.609305
0 2018	SI	43.545012
0 2019	SI	40.062234
0 2020	SI	39.968424

Fuente: código desde Visual Studio Code

El código de esta sección también se encuentra en el repositorio “AnemiaInfantilPeru” en Github archivo: “11. RevIndicesAnemia.py”. Cabe mencionar que esta fue una validación inicial sobre globales, pero a lo largo de los demás acápites se realizaron validaciones adicionales.

¹⁸ Los datos se excluyen del análisis, pero se mantienen en la base de datos por si otros usuarios los requieren.

5.7 Carga de fuentes adicionales

Si bien este paso no estaba contemplado en la metodología inicial, es importante realizarlo, ya que, como parte de la preparación del modelo de datos para trabajar las visualizaciones, se requirió de datos adicionales que se describen a continuación.

5.7.1. Información relacionada a los objetivos

En el año 2017, el Gobierno del Perú estableció objetivos para la lucha contra la anemia infantil que incluía índices de anemia objetivo para el periodo 2017 – 2021 y que fueron publicados en el “Plan Multisectorial de Lucha contra la Anemia” (*MIDIS 2017*). En este plan se incluyen las metas a nivel de país y a nivel regional. Los valores a nivel de país fueron aprobados por la Comisión Interministerial de Asuntos Sociales en la Tercera Sesión Ordinaria del año 2016, mientras que las metas regionales son asumidas por los gobiernos regionales.

Ilustración 28. Metas nacionales de anemia anualizadas 2017-2021

Indicador	2017	2018	2019	2020	2021
Tasa de anemia en niñas y niños de 6 a 35 meses de edad	43,6	42,0	39,0	30,0	19,0
	(basal)				

Fuente: Metas anuales incluidos en el Plan Multisectorial de Lucha contra la Anemia

Ilustración 29. Metas regionales anualizadas 2017-2021

N.º	Región	2017	2018	2019	2020	2021
1	Amazonas	38,8	34,0	29,2	24,3	19,4
2	Áncash	36,6	32,8	29,0	25,1	21,2
3	Apurímac	46,5	40,7	35,0	29,2	23,3
4	Arequipa	38,7	33,9	29,1	24,3	19,4
5	Ayacucho	46,0	39,2	32,5	25,8	19,0
6	Cajamarca	27,3	24,3	21,3	18,3	15,3
7	Callao	38,1	33,4	28,6	23,9	19,1
8	Cusco	49,1	41,6	34,1	26,6	19,1
9	Huancavelica	50,3	42,4	36,0	30,6	23,3
10	Huánuco	42,5	37,9	33,2	28,4	23,5
11	Ica	35,7	31,3	26,9	22,4	17,9
12	Junín	50,8	46,6	42,4	38,2	33,9
13	La Libertad	33,0	27,5	23,5	20,0	16,0
14	Lambayeque	36,3	34,3	31,3	24,8	19,8
15	Loreto	52,8	46,2	39,7	33,1	26,5
16	Madre de Dios	50,6	45,6	38,4	28,4	20,0
17	Moquegua	36,0	33,0	29,0	24,0	19,0
18	Pasco	52,7	46,1	39,6	33,1	26,4
19	Piura	31,3	27,4	23,5	19,7	15,7
20	Puno	66,0	57,8	49,6	41,4	29,0
21	San Martín	42,0	36,8	31,6	26,4	21,0
22	Tacna	31,2	27,3	23,4	19,5	17,3
23	Tumbes	42,4	37,2	31,9	26,6	21,3
24	Ucayali	50,3	41,4	35,6	29,7	23,7
25	Lima Provincias	35,6	31,1	26,7	22,3	17,8

Fuente: Metas regionales incluidas en el Plan Multisectorial de Lucha contra la Anemia

En el *Anexo V: Normatividad de los compromisos de gobiernos regionales*, se puede ver el detalle de los documentos normativos de los compromisos asumidos por los gobiernos regionales que se encuentran descentralizados.

5.7.2. Información de RENAMU

En el informe ENDES, se tiene información relacionada con la geolocalización hasta el nivel de distrito, sin embargo, la información solo se tiene como un código numérico y no la descripción específica, por ello, se procedió a obtener la información del nombre equivalente desde el RENAMU (Registro Nacional de Municipalidades), que también es una encuesta publicada por el INEI. A continuación, se describe de forma breve los pasos realizados para este punto.

- Se inició revisando las fichas técnicas desde donde se obtuvo la información de los módulos relacionados a la información de códigos y nombres de municipalidades por año. Luego de ello se preparó el archivo base para el paso siguiente relacionado con la descarga de la información.
- El proceso de descarga, descompresión y carga en la base de datos es similar a lo ejecutado en el punto 5.4 Recopilación de datos, solo que cambia la base inicial utilizada para la descarga. En este caso, debido a que la información que se requiere de estas encuestas es puntual, se ha trabajado en otra base de datos que será temporal. El código completo se podrá encontrar en el repositorio “AnemiaInfantilPeru” en Github, archivo: “12. DescargaRENAMU.py”.

Ilustración 30. Conexión y descarga de data RENAMU

```
conexion = create_engine('mysql+pymysql://root@localhost:3306/renamu')
con = conexion.raw_connection()
cur = con.cursor()

def consulta_enc_mod (registros):
    for i in registros.index:
        Nro_Encuesta = str(registros.loc[i, 'Cod_Encuesta'])
        Nro_Modulo = str(registros.loc[i, 'Cod_Modulo'])
        año = str(registros.loc[i, 'Año'])
        acceso = str(registros.loc[i, 'Tipo de Acceso'])
        file_url = 'http://inei.inei.gob.pe/inei/srienaho/descarga/'+acceso+'/'+ Nro_Encuesta +' -Modu
        get_zip(file_url,año)
```

Fuente: código desde Visual Studio Code

- Después de realizar la recopilación y la carga, se procedió con la consolidación y extracción de columnas que son importantes para el análisis. El resultado final fue un total de 1882 registros, que corresponde a los datos a nivel de distritos del Perú.

Ilustración 31. Tabla de departamentos, provincias y distritos



Mostrando filas 0 - 24 (total de 1882, La consulta tardó 0,0040 segundos.) [departamen

```
SELECT * FROM `db_nom_provdiss` ORDER BY `db_nom_provdiss`.`departamen` ASC
```

1 > >> | Número de filas: 25 | Filtrar filas: Buscar en este

+ Opciones

ccdd	departamen	ccpp	provincia	ccdi	distrito
1	AMAZONAS	1	CHACHAPOYAS	1	CHACHAPOYAS
1	AMAZONAS	1	CHACHAPOYAS	2	ASUNCION
1	AMAZONAS	1	CHACHAPOYAS	3	BALSAS

Fuente: Base de datos del proyecto

- d. Se procede a la incorporación de estos datos a la fuente general utilizada para esta evaluación. El código de esta parte se podrá encontrar en el repositorio “AnemiaInfantilPeru” en Github, archivo: “13. NombreProvDistr.py”. Aquí también se hace la división de la región Lima, ya que a nivel de evaluación, resulta conveniente hacer la diferenciación entre Lima provincias y Lima como región.

5.7.3. Geolocalización de las provincias y distritos

Al querer hacer uso de los datos de provincias y distritos en las vistas, se identificó que la base de datos de Tableau no contenía la totalidad de datos de provincia, ni distritos geolocalizados (solo se encontraban los de mayor dimensión poblacional), por ello, se decidió obtener las coordenadas de estos datos, desde Google Maps, haciendo uso de la API que ofrecen para tal funcionalidad.

Se implementó un código sencillo, tal como puede verse en la Ilustración 32:

Ilustración 32. Código para obtener datos georreferenciados con la API de Google Maps

```
def geolocalizar_provincia():
    df_bd_provincias = pd.read_sql_query('SELECT distinct provincia, departamen FROM db_nom_provdiss', con=conexion)

    for i in df_bd_provincias.index:
        provincia = str(df_bd_provincias.loc[i, 'provincia'])
        region = str(df_bd_provincias.loc[i, 'departamen'])
        place_provincia = provincia + ', ' + region + ', PERÚ'
        url = 'https://maps.googleapis.com/maps/api/geocode/json?&address=' + place_provincia + '%C3%BA&key=AQUI VA EL API KEY'
        req = requests.get(url)
        res = req.json()

        try:
            result = res['results']
            geodata_latitud = result[0]['geometry']['location']['lat']
            geodata_longitud = result[0]['geometry']['location']['lng']
            base_command = ('UPDATE db_nom_provdiss SET lat_prov = {geodata_latitud}, lng_prov = {geodata_longitud} WHERE provincia = '
                             + provincia)
            sql_command = base_command.format(geodata_latitud = geodata_latitud, geodata_longitud = geodata_longitud, provincia = provincia)
            cur.execute(sql_command)
        except:
            print('except')
            geodata_latitud = ""
            geodata_longitud = ""
```

Fuente: código desde Visual Studio Code

El código se puede encontrar en el repositorio “AnemiaInfantilPeru” en Github, archivo: “14. ObtenerCoordenadas.py”. La forma de cómo se pudo ocultar la API de Google se describe en el punto 5.9, Publicación y habilitación en datos abiertos.

5.7.4. Información del gasto devengado anual

Para poder hacer una comparativa de las mejoras que se han presentado a lo largo de los años, era importante realizar el comparativo del gasto en los diversos programas y estrategias del estado peruano. De esta forma, es posible evaluar si toda la inversión ejecutada, está viéndose representada en la consecución de los objetivos planteados.

Para obtener esta información se hizo uso de la plataforma denominada “Consulta amigable”, del MEF (Ministerio de Economía y Finanzas), en donde se puede dar seguimiento de la ejecución del gasto.

Ilustración 33. Consulta de Ejecución del Gasto

	PIA	PIM	Certificación	Compromiso Anual	Ejecución			Avance %
					Atención de Compromiso Mensual	Devengado	Girado	
TOTAL	197,002,269.014	219,019,451.356	166,835,365.329	137,539,266.638	82,576,894.813	71,397,313.028	68,995,002.918	32.6

Notas

- Los montos están en Soles.
- La columna Avance % representa la razón del Devengado entre el PIM, expresado en porcentajes.
- A partir del 2007 se comienza a incluir información de los Gobiernos Locales. [Ver más detalles.](#)
- A partir del 2012 el programa cambia de denominación por el de división funcional, y el subprograma por el de grupo funcional.
- La información se actualiza mensualmente. Última actualización: 31 de mayo de 2022.

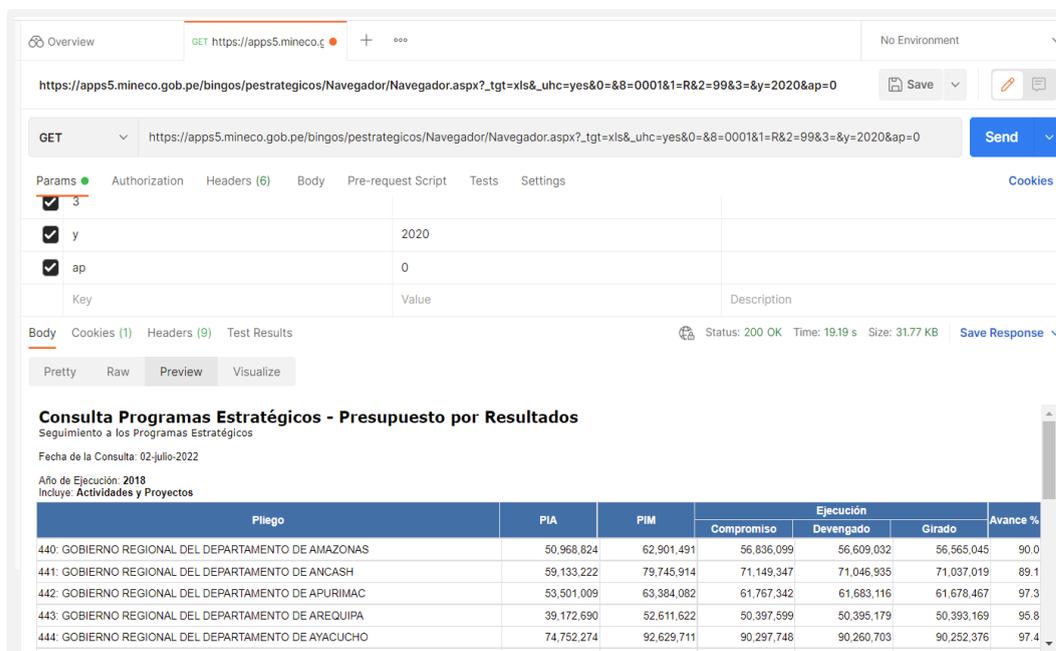
[Sobre la información presentada](#) | [Estadísticas de uso](#)

Fuente: Consulta amigable MEF

Esta plataforma presenta restricciones de acceso geográfico, pero su acceso solo está permitido para Perú y también tiene restricciones a través de código, por ello, en este paso y a la fecha del presente informe, no se pudo realizar la extracción automatizada con Python debido a que el servidor rechazaba la solicitud. Por lo tanto, se utilizó la herramienta Postman para este fin, el uso se puede ver en la ilustración 34. De igual manera, en esta plataforma se puede automatizar la descarga asignando parámetros. Cabe mencionar que la web de Consulta Amigable del MEF realiza una asignación de nombres de archivos, automática y de manera aleatoria y, además, los archivos se descargan en un formato incorrecto.

Finalmente, se trabajaron los archivos para lograr la unificación y carga en la base de datos. El programa seleccionado para la evaluación es el denominado “Programa Articulado Nutricional”, que incluye un indicador denominado “Prevalencia de anemia en menores de 36 meses” y, por otro lado, también se ha incluido la evaluación del producto específico “Niños con suplemento de hierro y vitamina A”, para realizar la comparativa con la cantidad de niños que consumieron hierro y hacer el contraste contra el índice de anemia para ver si este producto está teniendo los resultados esperados.

Ilustración 34. Uso de herramienta Postman para obtener datos



Fuente: Web Postman

5.8 Elaboración de visualizaciones de datos

Con los datos limpios y cargados en la base de datos, se procede con la elaboración de unos tableros de mando en Tableau, de modo que permita tener una mejor visualización de la información e indicadores. Se ha dividido las visualizaciones en un análisis a nivel de país que permita tener un vistazo general del índice de anemia y otro análisis a nivel regional, que permita ver el detalle de la situación por cada región, incluyendo sus provincias y distritos.

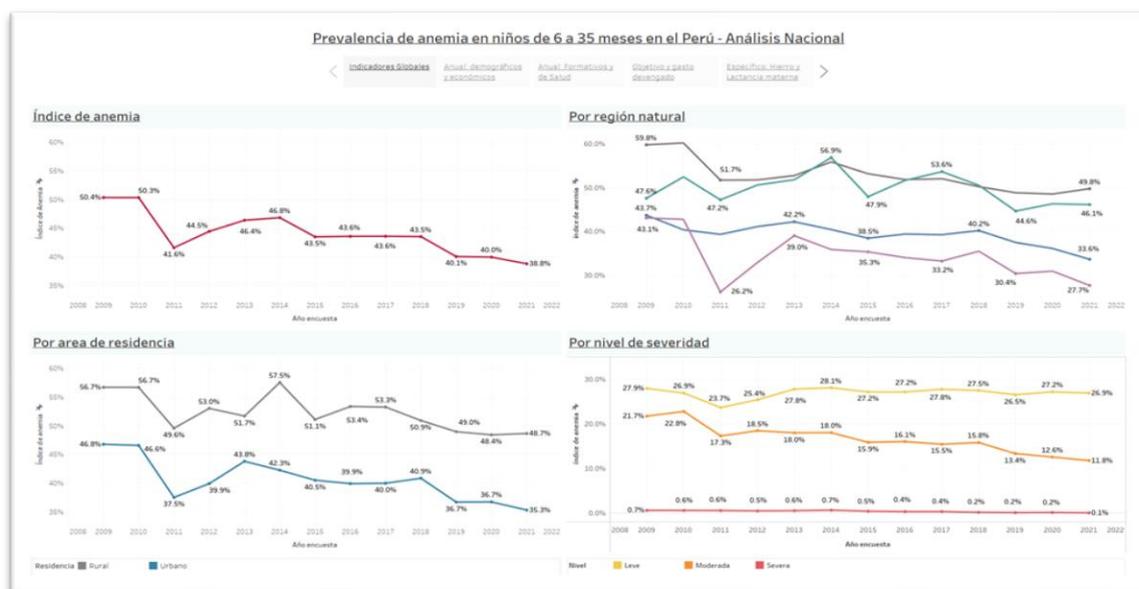
En la Ilustración 35, se puede ver la visualización creada a nivel de país, en donde se ha incluido 5 cuadros de mando para los principales análisis, los cuales son:

- Indicadores globales. - Muestra una comparativa entre regiones naturales, área de residencia y nivel de severidad, además de la evolución del indicador por año.
- Anual: demográficos y económicos. - Como su nombre lo menciona, muestra la relación de los factores demográficos y económicos con el índice de anemia. La periodicidad es anual.
- Anual: formativos y de salud. - Aquí se evalúa los factores más relevantes relacionados a la educación, condición del seguro de salud y condiciones del niño al nacer.
- Objetivo y gasto devengado. - Es una comparativa entre el índice de anemia en contraste con los índices objetivos planteados por el gobierno para los últimos 5 años, así como también, la comparativa con el gasto devengado en el programa PAN¹⁹.

¹⁹ PAN, acrónimo de Programa Articulado Nutricional, que es la principal iniciativa planteada por el gobierno peruano para mejorar la nutrición infantil.

- **Específico: hierro y lactancia materna.** - aquí se evalúa el gasto devengado de un producto específico que es “Niños con suplemento de hierro y Vitamina A” y se contrasta con el % de niños con suplemento de hierro. Luego se compara este último con el índice de anemia. Adicionalmente, se incluye la evaluación respecto a la anemia de la madre en el embarazo y también con respecto a la lactancia exclusiva durante los primeros 6 meses de edad.

Ilustración 35. Cuadro de mando indicadores a nivel país



Fuente: Tableau – Elaboración propia

En la ilustración 36, se presenta la visualización a nivel regional, la cual se han incluido 3 cuadros de mando. A continuación, una pequeña descripción:

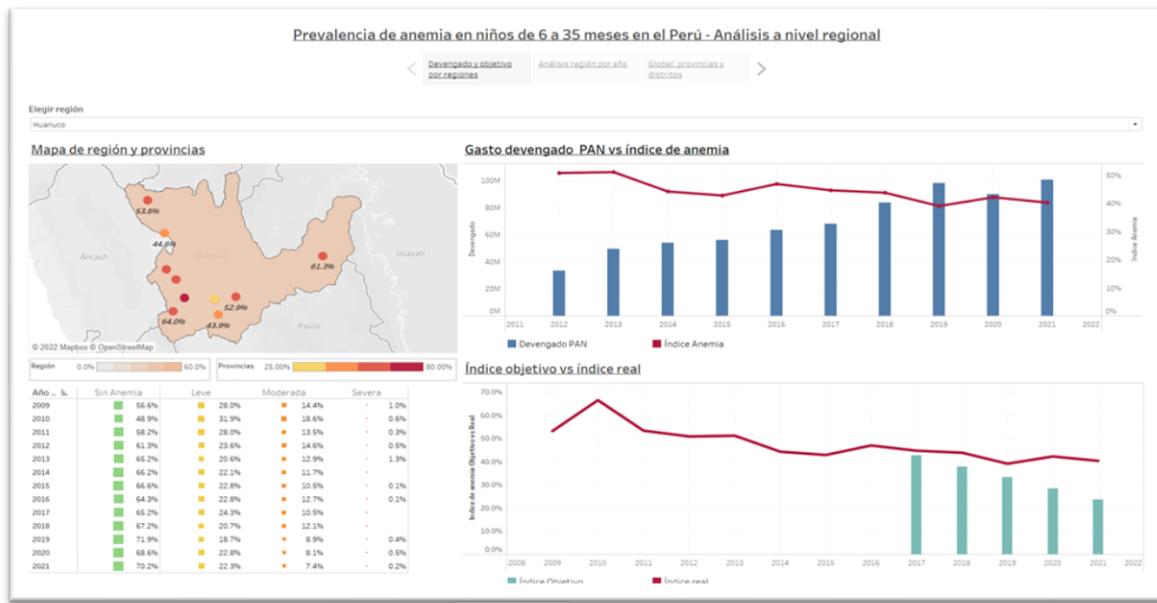
- **Devengado y objetivo por regiones.** - Hace referencia a la comparativa del índice de anemia en contraste con el gasto devengado PAN y de los objetivos específicos en la región.
- **Análisis región por año.** - Se incluye los principales indicadores que son de relevancia y sobre los cuales el gobierno regional, puede tener un mayor ámbito de acción.
- **Global: provincias y distritos.** - Mapa global con el índice de anemia de todo el país. Este debe ser tomado únicamente como referencia, ya que la muestra considerada no asegura que se tenga una estructura poblacional adecuada para provincias y distritos.

Cabe mencionar que, a fin de recibir una retroalimentación de la utilidad de los cuadros presentados, estas visualizaciones fueron mostradas a personas que participan al área de análisis de datos del gobierno regional, personal sanitario encargado de los programas de crecimiento y desarrollo (CRED) y empresas aliadas que realizan campañas para mejora los problemas nutricionales en el Perú. A pesar de tener una respuesta positiva de dichos sectores, se resaltó que a nivel salud, los indicadores que se necesitan deben ser más específicos y llegar a un mayor

nivel de detalle. De igual manera, para las empresas aliadas, se requiere mayor granularidad en relación con la extensión geográfica para poder ver la influencia de sus programas.

En el Capítulo VI: Resultados, ítem 6.1, Análisis de indicadores, se presentará un análisis más detallado de estas representaciones.

Ilustración 36. Números globales y comparativos con objetivo y devengado - Regional



Fuente: Tableau – Elaboración propia

5.9 Publicación y habilitación en datos abiertos

Es este acápite se realiza la inicialización de los archivos en Git, carga de datos en GitHub y la publicación de la página web.

5.9.1. Versión inicial del código en Git

Una vez que se ha culminado la primera versión del código y se ha validado que es funcional, se crea la primera versión en el gestor Git, tal como se muestra en la Ilustración 37:

Ilustración 37. Carga de versión inicial en Git

```

create mode 100644 scripts/Aux-Querys.sql
create mode 100644 scripts/Aux-renamu.sql
create mode 100644 scripts/Consolidado.sql

Katia@LAPTOP-CT8PTAQA MINGW64 ~/desktop/IndAnemiaProcesaDatos (master)
$ git log
commit c18d32eab441ddce75bcb13f5428601385bf2fb6 (HEAD -> master)
Author: Katia <kameco.21@gmail.com>
Date: Sat Jul 2 14:15:04 2022 +0200

    miPrimeraVersionCommit

Katia@LAPTOP-CT8PTAQA MINGW64 ~/desktop/IndAnemiaProcesaDatos (master)
$ git status
On branch master
nothing to commit, working tree clean
    
```

Fuente: Git

Esto servirá para tener una adecuada gestión de versiones, que pueda tener trazabilidad de los cambios que se realicen a partir de este punto. Además, facilitar la comunicación con GitHub, lo que se describe a continuación.

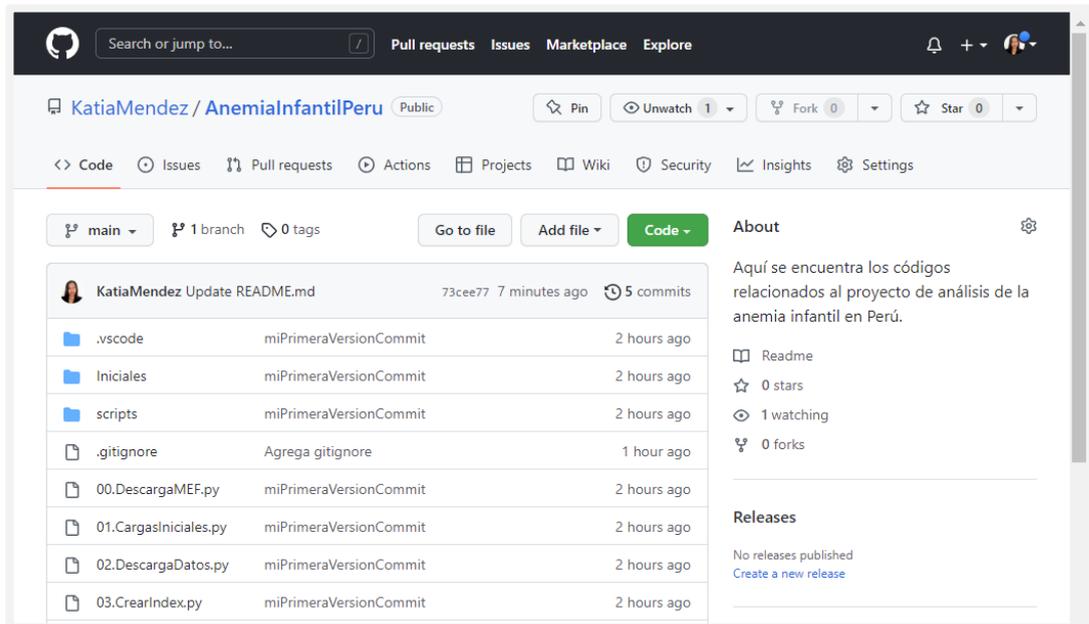
5.9.2. Carga de archivos en GitHub

En GitHub, se realiza la carga de todo el código relacionado al procesamiento de los datos y se despliega de manera pública, para que sea utilizado por quienes lo requieran. El proceso fue sencillo y se realizó desde la consola de Git, siguiendo las indicaciones que se encuentran en la página de GitHub. El repositorio donde se encuentra todos los archivos que han sido necesarios para el procesamiento es:

<https://github.com/KatiaMendez/AnemiaInfantilPeru>

En la Ilustración 38 se tiene el resultado obtenido:

Ilustración 38. Archivos de programación cargados en GitHub



Fuente: Repositorio GitHub

Sobre el punto de la API KEY de Google Maps utilizada en el código para la obtención de coordenadas, inicialmente se pensó en no compartir este archivo, o compartirlo con el ítem del API KEY borrado, pero eso representaría que se tendría que cambiar constantemente el código. Por ello, se analizó e identificó la manera de poder ocultarlo al momento de compartir en Github y esto se realiza mediante el uso del archivo “.gitignore”. En este archivo se agregó el documento que contiene la función que llama a la API KEY. Este archivo sirve para registrar todo aquello que, por algún motivo, no se quiere hacer público.

5.9.3. Publicación en un entorno Web

Finalmente, como último punto de la metodología propuesta, se procede con la publicación de los tableros de control desarrollados, así como también se habilitan los datos extraídos y consolidados. Esto se realiza en una web con acceso abierto, tomando en cuenta que las licencias que son otorgadas por las instituciones permiten, por lo menos, el uso libre - no comercial.

Para poder implementar este punto, se han realizado las siguientes actividades:

- A. Diseño de la Web: la maquetación inicial de la página web, se realizó en local y haciendo uso de las siguientes herramientas: WAMP Server, Visual Studio Code, Python, HTML, CSS, JavaScript, MySQL, Ilustrador para la parte de diseño, Git para la gestión del cambio y GitHub como repositorio del código. El nombre que se ha asignado a la plataforma es: Sanito Pe!, el cual es un acrónimo de SALud para la NIñez de TOvero el PERÚ, pero a la vez tiene una connotación de identidad para todos los peruanos ya que el sufijo “pe”, representa una manera coloquial de lenguaje utilizado mucho en Perú.

Ilustración 39. Página principal de la plataforma SANITO PE



Fuente: Web Sanito Pe!

- B. Adquisición del dominio y hosting: en paralelo al punto anterior, se realiza la búsqueda y adquisición del dominio www.sanitope.com y se realiza la búsqueda de un proveedor de hosting que permita almacenar la página web. Un punto importante que se debió tomar en cuenta para este ítem fue asegurar que el host contratado permita el despliegue de código Python y de base de datos MySQL, ya que la web se encuentra desarrollada con estas herramientas. Para poder realizar una elección adecuada se hizo un comparativo de

proveedores que ofrecen el servicio y además se verificó algunas opiniones de las personas que los utilizan. Por lo expuesto, se eligió la empresa IONOS²⁰, para la obtención de un dominio por el costo de 1€ mensual durante el primer año y para el caso del hosting, se seleccionó Pythonanywhere²¹, que inicialmente se utilizó en su plan gratuito para probar el despliegue correcto y luego se contrató la versión más básica para poder hacer uso de dominio propio adquirido.

Ilustración 40. Servicio de hosting - Pythonanywhere

Hacker	Web dev	Startup	Custom
\$5/month	\$12/month	\$99/month	\$5 to \$500/month
Run your Python code in the cloud from one web app and the console	If you want to host small Python-based websites for you or for your clients	Start a business and don't worry about having to scale to handle traffic spikes	Want a combination that's not on the list? Create your own! All custom plans have:
A Python IDE in your browser with unlimited Python/bash consoles	A Python IDE in your browser with unlimited Python/bash consoles	A Python IDE in your browser with unlimited Python/bash consoles	A Python IDE in your browser with unlimited Python/bash consoles
One web app on a custom domain or your-username.pythonanywhere.com	Up to 2 web apps on custom domains or your-username.pythonanywhere.com	Up to 3 web apps on custom domains or your-username.pythonanywhere.com	Up to 20 web apps, on custom domains or your-username.pythonanywhere.com
Enough power to run a typical 100,000 hit/day website. (more info)	Enough power to run a typical 150,000 hit/day website on each web app. (more info)	Enough power to run a typical 1,000,000 hit/day website on each web app. (more info)	As many web workers as you need to scale your site's capacity. (more info)
2,000 CPU-seconds per day for consoles, scheduled tasks and always-on tasks (more info)	4,000 CPU-seconds per day for consoles, scheduled tasks and always-on tasks (more info)	10,000 CPU-seconds per day for consoles, scheduled tasks and always-on tasks (more info)	Up to 100,000 CPU-seconds per day for consoles, scheduled tasks and always-on tasks (more info)
IPython/Jupyter notebook support	IPython/Jupyter notebook support	IPython/Jupyter notebook support	IPython/Jupyter notebook support
1GB disk space	5GB disk space	50GB disk space	As much disk space as you choose
Please contact support to change your account	Please contact support to change your account	Please contact support to change your account	Please contact support to change your account

Fuente: Web IONOS

C. Migración de local a público: se procede a publicar la primera versión de la web SanitoPe. Para ello se realiza la migración y despliegue en el host contratado. Para poder ejecutar exitosamente esta tarea, se siguieron estos pasos:

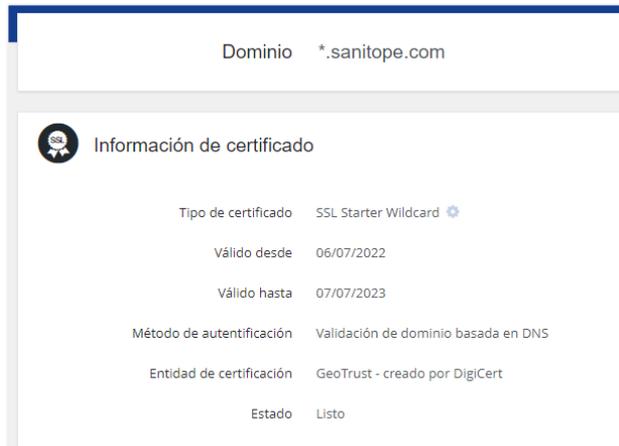
- Exportar la BD en un archivo SQL.
- Activar el control de versiones en Git.
- Migrar los archivos de desarrollo desde el equipo local a Github.
- Realizar configuraciones en el servidor.
- Direcccionar el dominio hacia la página web y activar el SSL (Ilustración 41).

Luego de estas configuraciones se obtuvo la primera versión de la web en productivo. En la Ilustración 42, se muestra la versión en un smartphone.

²⁰ Enlace de IONOS: <https://www.ionos.es/alojamiento/linux-hosting>

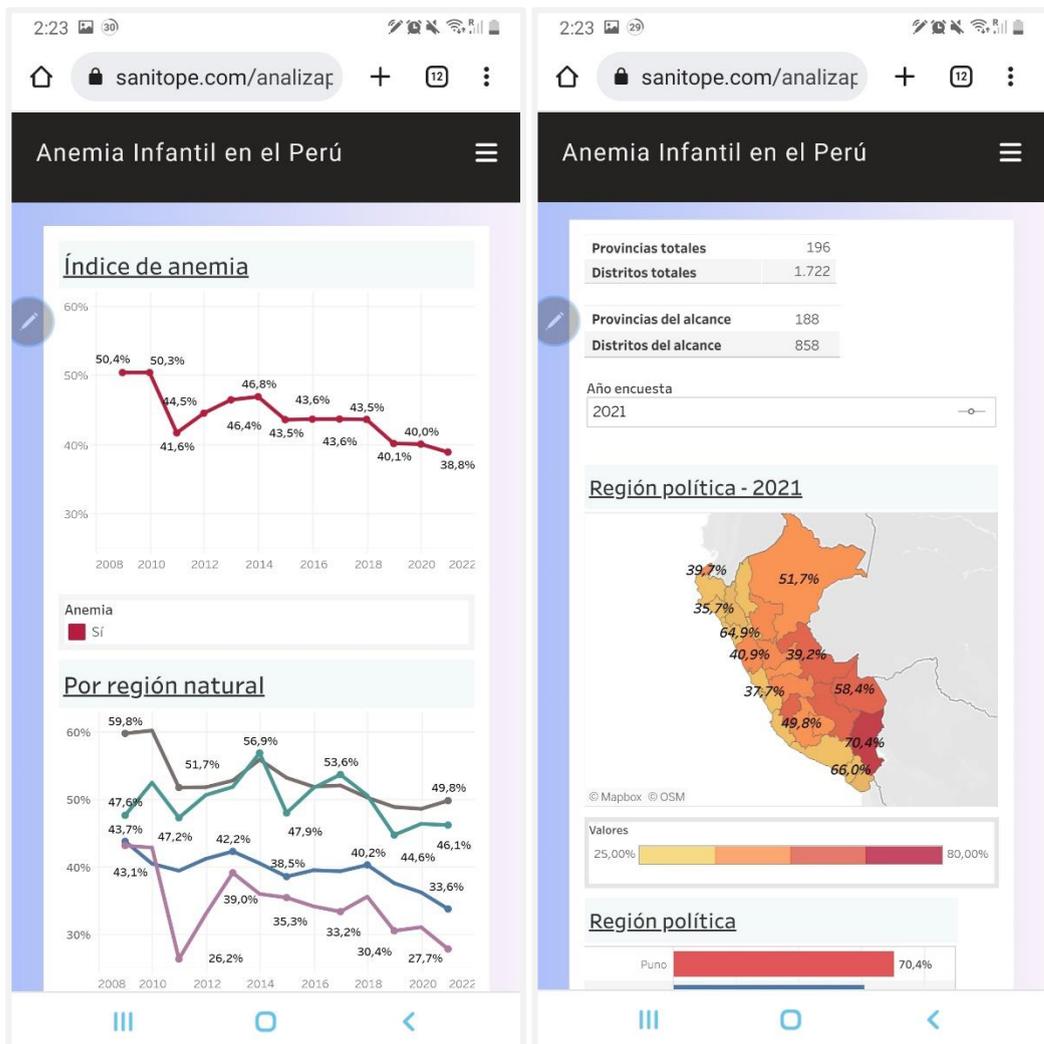
²¹ Enlace de Pythonanywhere: <https://www.pythonanywhere.com/pricing/>

Ilustración 41. SSL para el dominio



Fuente: IONOS, configuración del dominio

Ilustración 42. Página SanitoPe en una versión móvil



Fuente: Web SanitoPe (www.sanitope.com).

CAPÍTULO VI: Pruebas y resultados

En este capítulo se desarrollan las pruebas de funcionamiento de los elementos publicados en el entorno web, y también se incluye el análisis de los principales indicadores presentados en las visualizaciones. Finalmente, se documentarán las principales dificultades identificadas y las estrategias de resolución llevadas a cabo.

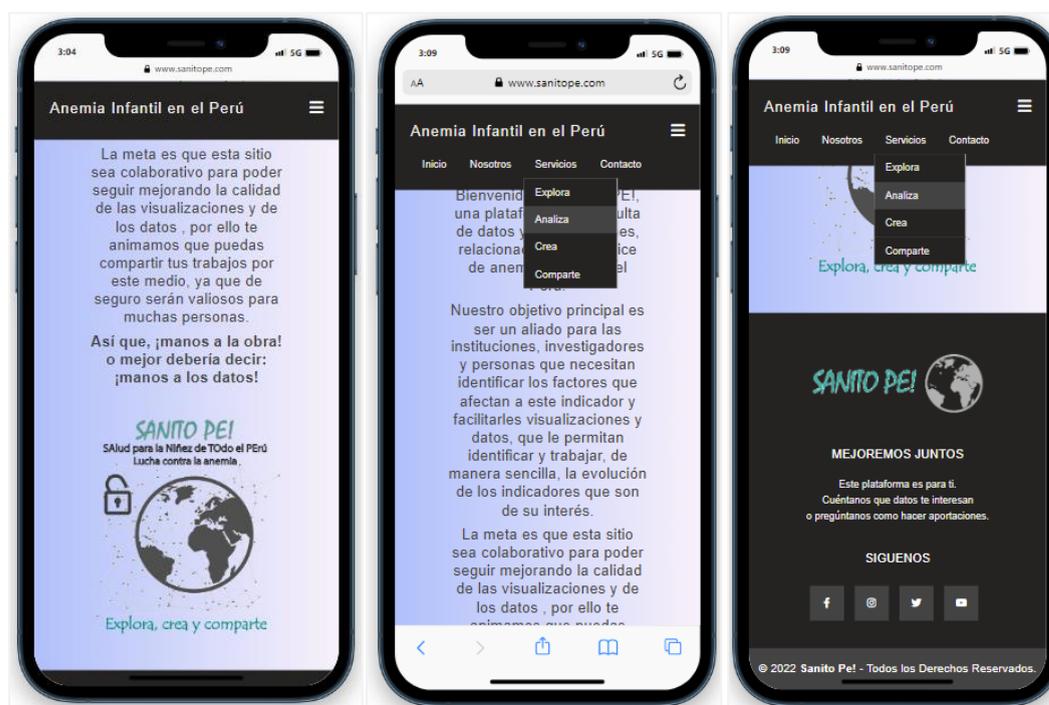
6.1 Pruebas de funcionamiento de la Web

Se realizan las principales pruebas en el dominio www.sanitope.com.

6.1.1. Pruebas de adaptabilidad

En este punto se procede a comprobar que el diseño adaptativo funciona correctamente. Se realizó la prueba sobre la pantalla inicial para la verificación del menú responsive de la cabecera y de la adaptación de los elementos del pie de página. En la Ilustración 43, se puede observar cómo se adecua correctamente de acuerdo con la disponibilidad del espacio del que se disponga:

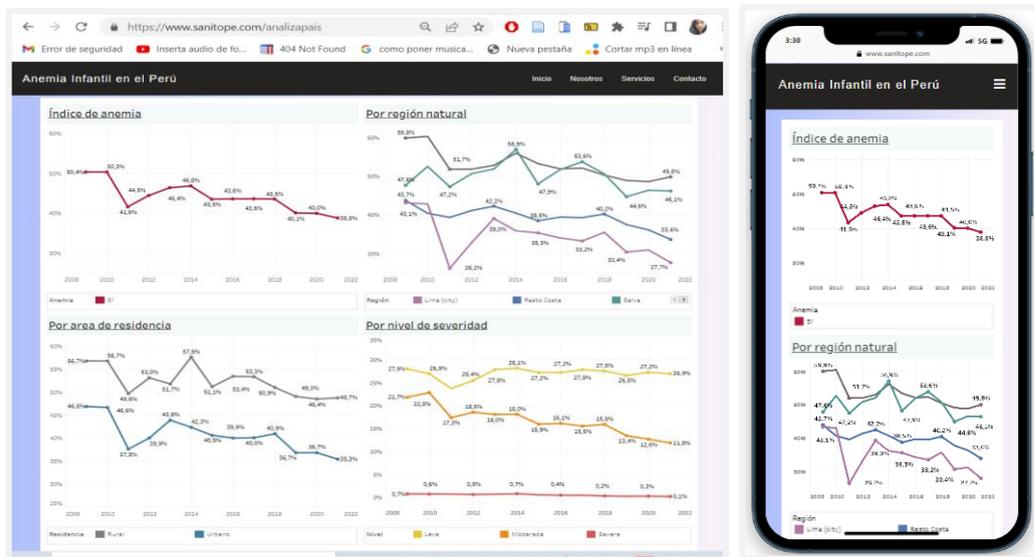
Ilustración 43. Prueba de adaptabilidad SanitoPe: cabecera y pie de página



Fuente: Web SanitoPe

De igual manera se probó la adaptabilidad de las gráficas de los tableros de control. En la Ilustración 44, en la imagen de la izquierda se puede apreciar cómo el dashboard se presenta todo en conjunto de manera horizontal, mientras que en la imagen de la derecha se aprecia que los gráficos están alineados verticalmente.

Ilustración 44. Prueba de adaptabilidad SanitoPe: tableros de control



Fuente: Web SanitoPe

6.1.2. Pruebas funcionales

Se plantean pruebas para ver que las funciones principales del sitio están operando correctamente. En este análisis se incluyen las dos funcionalidades que se consideran principales. Estas son:

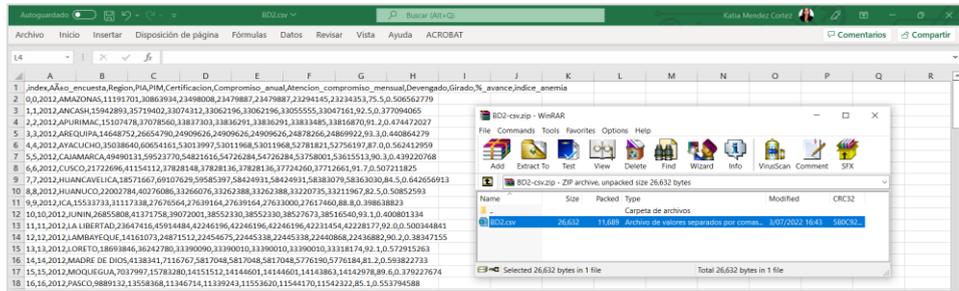
- A. Descarga de datos: se corrobora que la navegación funciona correctamente y que permite la descarga o visualización de los datos que se ofrecen. En la Ilustración 45 y 46 se ve la descarga en la vista Datasets:

Ilustración 45. Prueba funcional de descarga de datos - Vista Datasets

Id	Dataset	Descripción	Años de alcance	Descarga Datos	Descarga Diccionarios
1	Índice de anemia infantil en Perú	Contiene los datos relacionados al índice de anemia en niños de 0 a 5 años, recopilada en la ENDES del INEI. Se incluye la evaluación de factores demográficos, económicos, formativos y de salud.	2009-2021	CSV, JSON, STATA	CSV, PDF, XLSX
2	Gasto devengado PAN vs índice de anemia	A partir de los datos del gasto devengado obtenidos del MEF se ha incorporado la data del índice de anemia para niños de 6 a 35 meses	2012-2021	CSV, JSON, STATA	CSV, PDF, XLSX
3	Gasto devengado suplemento hierro y vitamina A vs consumo de hierro	A partir de los datos del gasto devengado obtenidos del MEF se ha incorporado la data del porcentaje de consumo de hierro y el índice de anemia para niños de 6 a 35 meses	2013-2020	CSV, JSON, STATA	CSV, PDF, XLSX
4	Anemia: índice objetivo vs índice real	Se obtuvieron los datos del índice de anemia objetivo para niños de 6 a 35 meses, obtenidos de las publicaciones del MIDIS, y se contrastaron con los índices reales	2017-2021	CSV, JSON, STATA	CSV, PDF, XLSX

Fuente: Web SanitoPe

Ilustración 46. Prueba funcional de descarga de datos - Verificación de Información



Fuente: Archivo descargado de Web SanitoPe

En la Ilustración 47 tenemos la vista de instituciones: Selecciona el año y se abre la ficha, en una nueva pestaña.

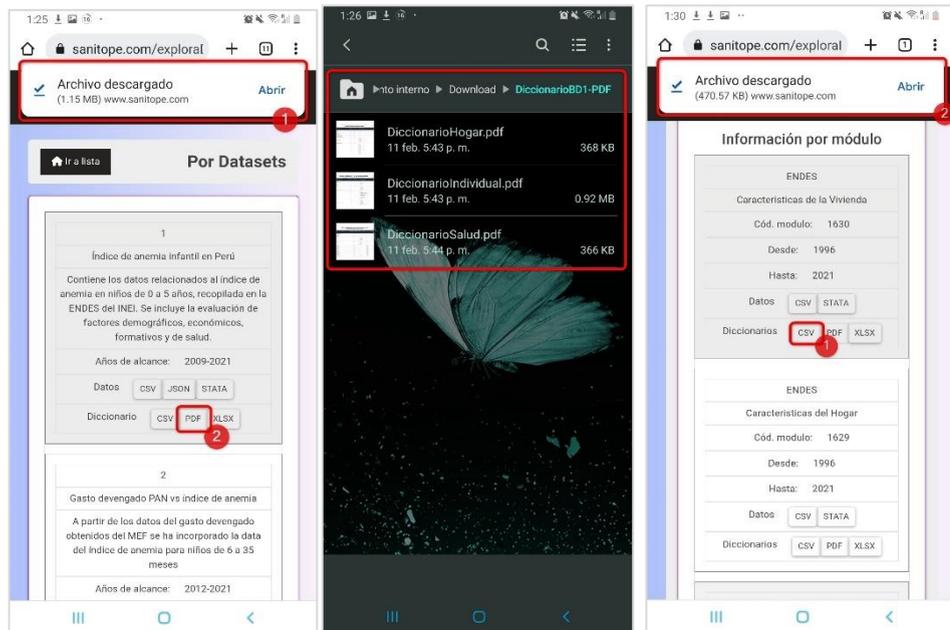
Ilustración 47. Prueba funcional de descarga de datos - Vista Instituciones



Fuente: Web SanitoPe

En la Ilustración 48 se presenta la vista en dispositivos móviles:

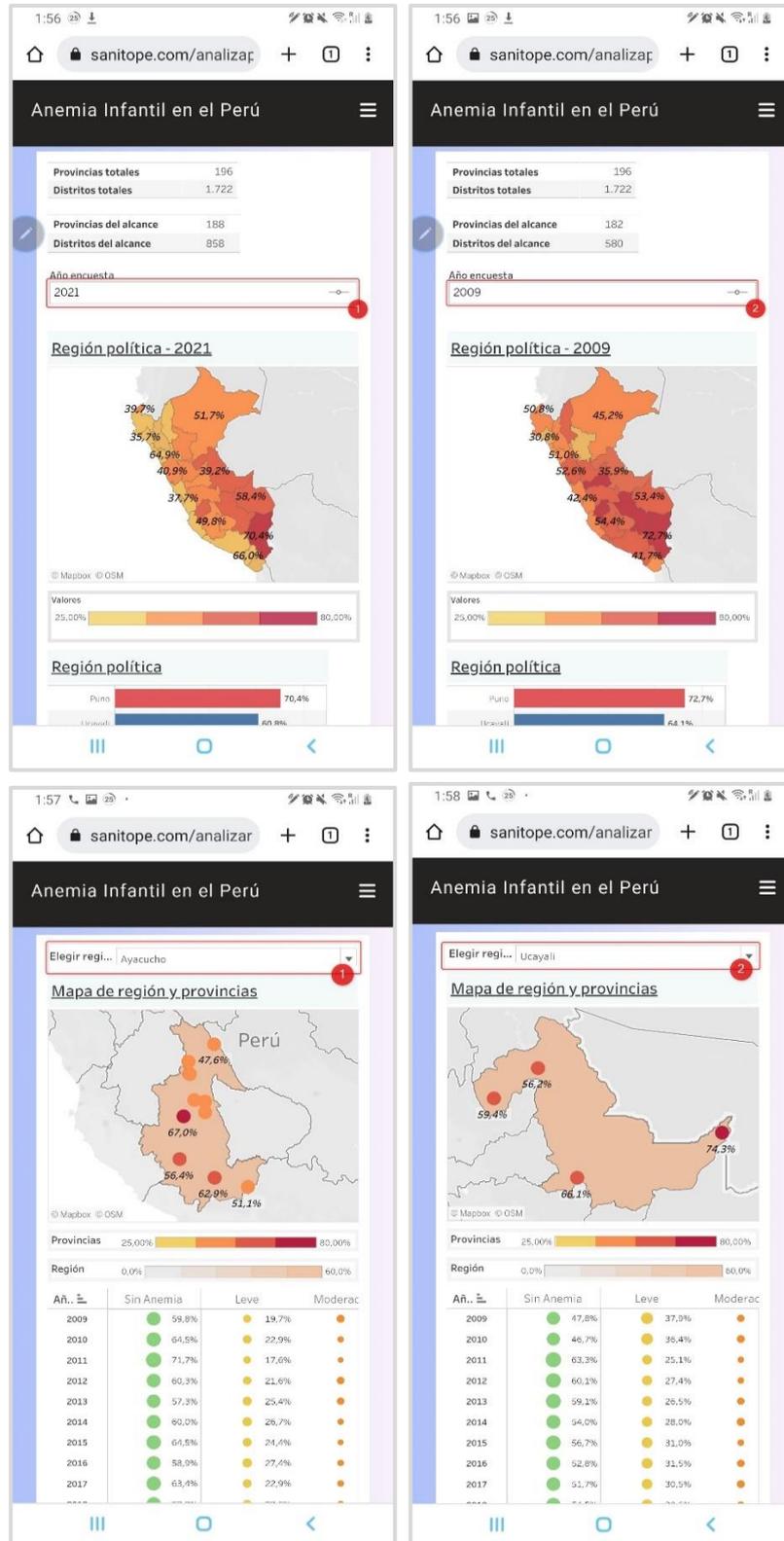
Ilustración 48. Prueba funcional de descarga de datos - Vista móviles



Fuente: Web SanitoPe

B. **Tableros de mando:** se verificó que funcionen los filtros y transiciones en móviles, como puede verse en la Ilustración 49.

Ilustración 49. Prueba funcional de filtros en tableros de mando: país y región

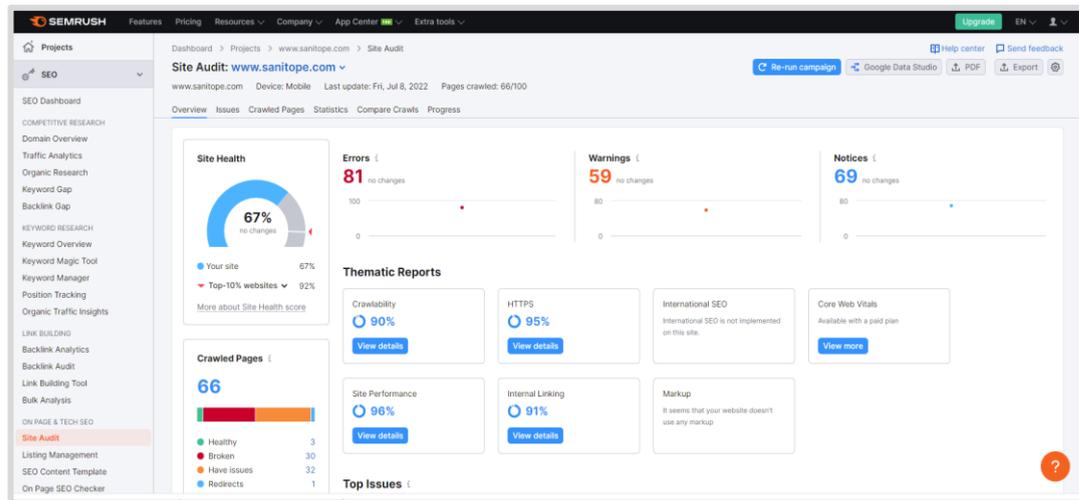


Fuente: Web SanitoPe

6.1.3. Auditoría SEO

Se plantean pruebas de una auditoría SEO general, que permita obtener el estado de salud de la página y ver las principales brechas y oportunidades para poder mejorar. No se pretende superar todos los problemas que se identifiquen, pero sí tener una visión inicial de lo que se tiene, para luego ir mejorando la página.

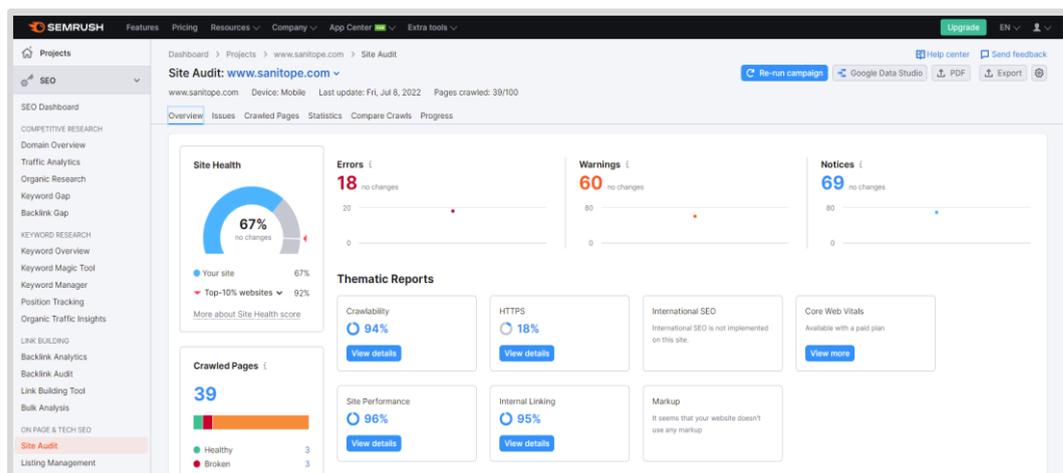
Ilustración 50. Auditoría SEO a la Web SanitoPe: resultados iniciales



Fuente: Semrush – Auditoría SEO

Se revisaron los errores y se corrigieron aquellos que eran de fácil resolución, con ello se redujo de 81 a 18 errores. Se pretende seguir con las validaciones y solucionar las brechas para mejorar el funcionamiento del sitio web y, por ende, mejorar la experiencia del usuario en el uso de la herramienta:

Ilustración 51. Auditoría SEO a la Web SanitoPe: resultados finales



Fuente: Semrush – Auditoría SEO

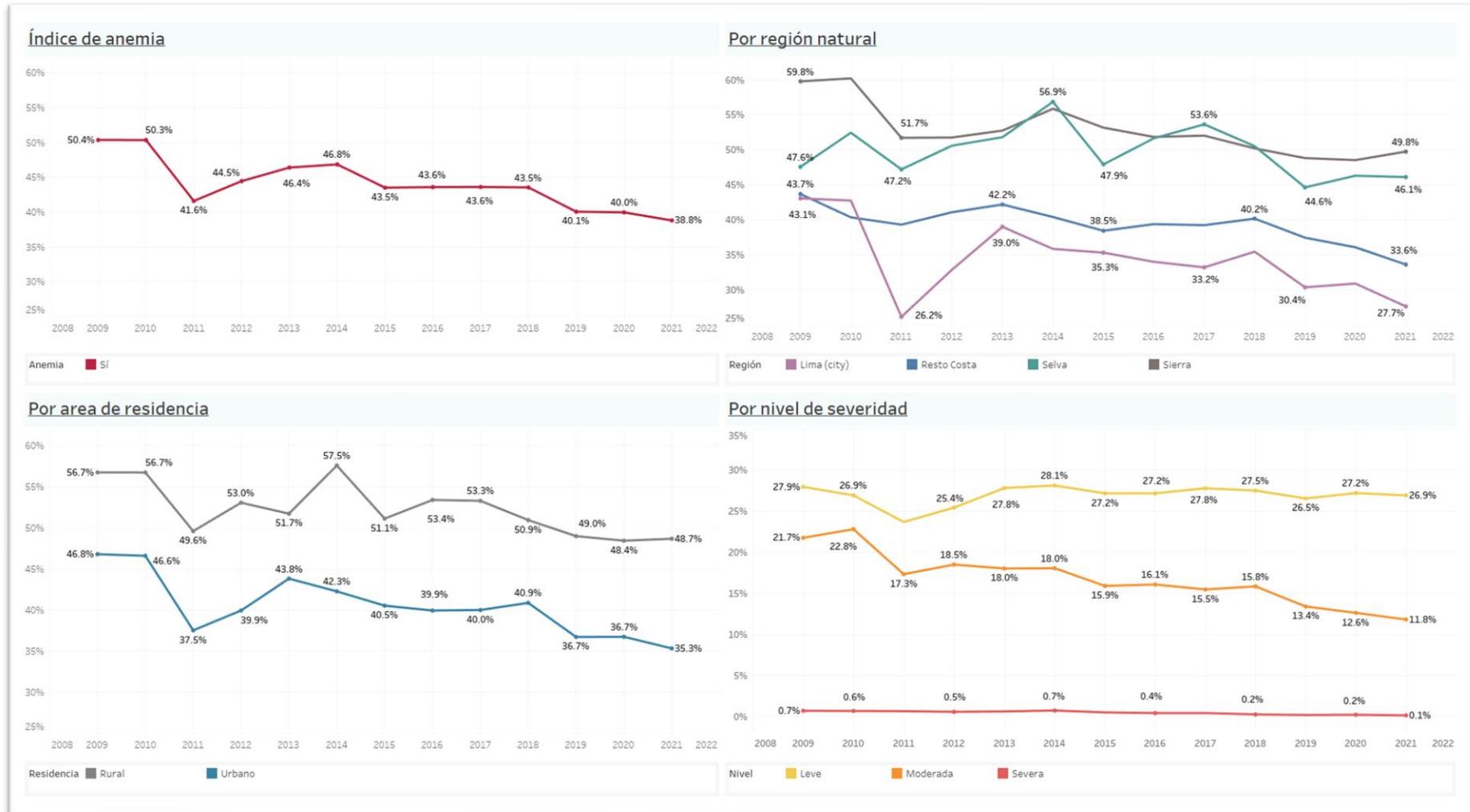
6.2 Análisis de indicadores

6.2.1. Análisis Nacional

Se plantea una presentación inicial global de todo el país, en donde se evalúa el índice de anemia en relación con distintos factores que podrían tener una correlación. A continuación, se tiene un pequeño análisis de los puntos más relevantes.

- A. Indicadores globales: En la Ilustración 52, mostrada líneas abajo, se puede observar claramente lo comentado en todo el trabajo, el índice de anemia no ha presentado mejoría significativa en los últimos años, ya que desde el 2011 que presentó un valor de 41,6%, solo se ha podido reducir 2,8% en los últimos 10 años. Si se analizan las cifras **por región natural**, se podrá notar que la sierra ha sido, casi siempre, la región más afectada y, por el contrario, Lima metropolitana o ciudad, ha sido la que ha mantenido mejor este indicador. Sobre este mismo análisis, sería importante validar el índice de anemia para Lima en el año 2011, ya que como se puede apreciar, representa un valor atípico y puede que esté distorsionando los indicadores para este año. Al tener Lima ciudad una población de 9,6 millones de habitantes, que conforman casi el 30% de la población del Perú, la representatividad sobre el total de la muestra es significativa y puede generar distorsiones relevantes en los indicadores. Es importante esta consideración para tenerla en cuenta en el resto de los análisis realizados. En cuanto al análisis **por área de residencia**, se puede notar que es el área rural el que más padece con este problema, esta situación ha sido evidente desde hace muchos años, por la dificultad de acceso físico que tienen las personas que viven en dichas zonas, por ello los planes de lucha contra la anemia siempre mencionaban un mayor foco en esta población, sin embargo, si se observa la evolución desde el 2018 (resultados desde que se reestructuro el plan de respuesta), se puede apreciar que el área urbana ha tenido una mejoría en 5,8 puntos porcentuales, mientras que en el área rural, en este mismo periodo, solo se ha tenido solo una mejoría de 2,2 puntos porcentuales. Esta situación debería involucrar una revisión de los planes y su aplicación. Finalmente, también se puede apreciar la evaluación **por nivel de severidad**, en donde se percibe que la anemia severa, que es la más preocupante, siempre ha mantenido un valor pequeño y que ha caído un poco más en los últimos años. De igual forma, se ve el mismo reflejo en la anemia moderada; sin embargo, la anemia leve persiste en niveles similares desde el año 2009.

Ilustración 52. Cuadro de mando indicadores globales



Fuente: Tableau – Elaboración propia

- B. Factores demográficos y económicos: En las Ilustraciones 54 y 55, mostrada líneas abajo, se presenta la relación del índice de anemia con factores geográficos, demográficos y económicos. Lo primero que se evaluó en esta presentación es la representatividad a **nivel provincias y distritos**, como se observa, la cantidad total de provincias y distritos es 195 y 1.683, respectivamente. Realizando el cálculo de muestra representativa con la herramienta SurveyMonkey²², se puede ver que la cantidad de registros en el alcance, tanto para el 2009 (182 y 580), como para el 2021 (188 y 858) excedía lo requerido:

Ilustración 53. Cálculo de la cantidad de muestra para el 2009

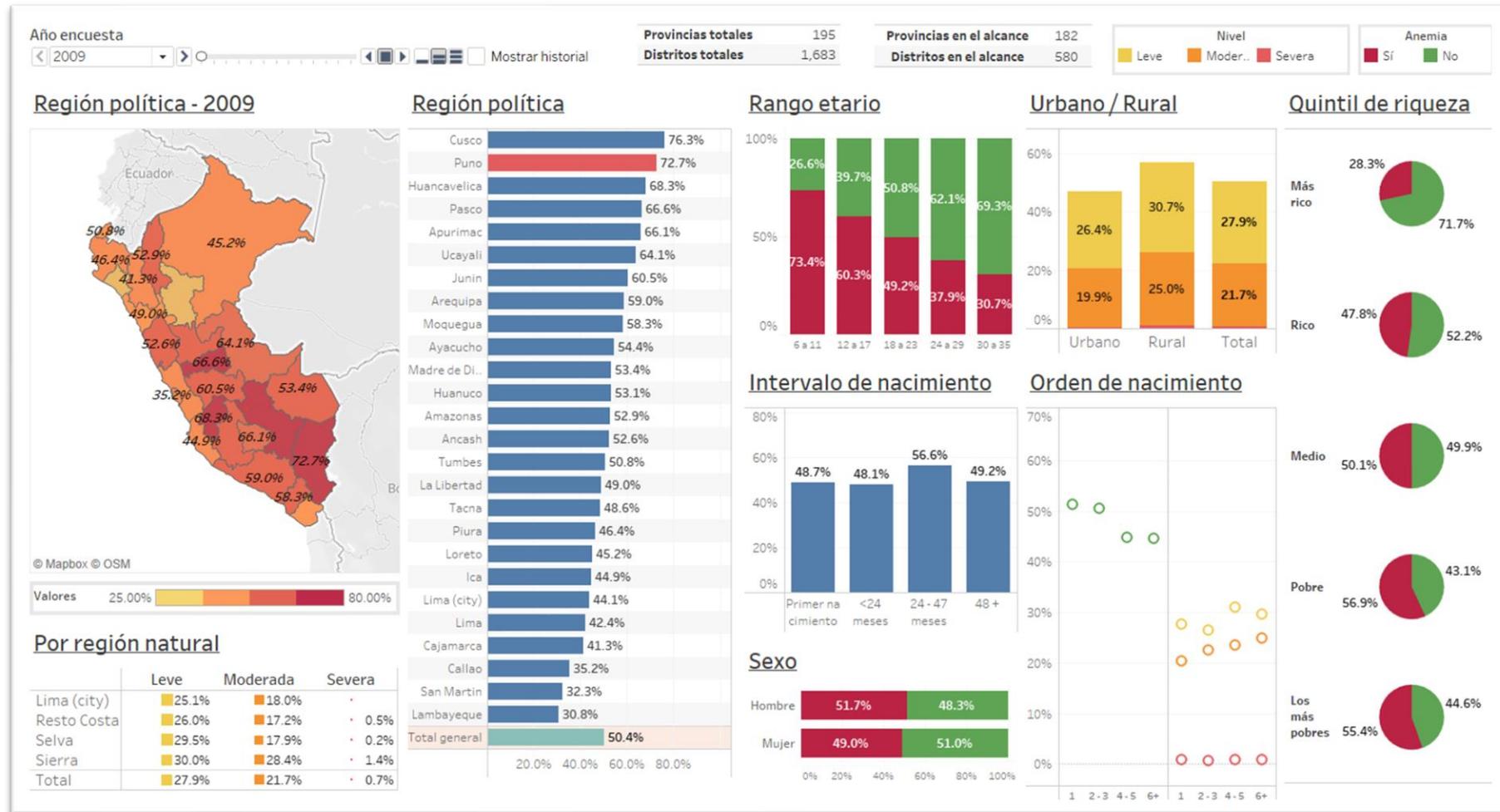
Tamaño de la población	Nivel de confianza (%)	Margen de error (%)	Tamaño de la muestra
195	95	5	130
1683	95	5	313

Fuente: SurveyMonkey

Continuando con el análisis, si bien aquí se tiene la presentación de 2 años, cabe mencionar que el análisis se puede realizar para los 12 años del alcance. En primer lugar, se puede apreciar cómo el mapa de **región política** muestra una buena mejoría en el año 2021 ya que tiene una mayor incidencia del color amarillo, que representa menor porcentaje de anemia. Así mismo, en la gráfica de barras del mismo acápite “región política”, se presenta una preocupante cifra para la región Puno (remarcada de rojo), que luego de 12 años, solo ha disminuido un 2,3%. Esto debe ser un punto de interés y puesta en marcha de programas de que permitan mejorarlo. Con respecto a los demás ítems, se tiene: **(1) rango etario**, se puede apreciar que la anemia disminuye a lo largo del crecimiento; **(2) intervalo de nacimiento**, en la mayoría de los años evaluados, se aprecia menor incidencia en los primeros nacimientos o los nacimientos después de 48 meses, lo que muestra que el cuidado de una mayor cantidad de niños puede representar un factor importante para el índice analizado; **(3) sexo**, aquí se ve un mayor predominio en el sexo masculino; **(4) orden de nacimiento**, se denota que la incidencia es menor en los primeros hijos, esto creemos que se debe a los cuidados que suelen ser mayores al iniciar la maternidad/paternidad; y finalmente, **(5) quintil de riqueza**, sobre este punto es claro que el índice de anemia es mucho mayor para la población más pobre y que a través de los años, la reducción para este sector de la población, es la misma que para los demás sectores.

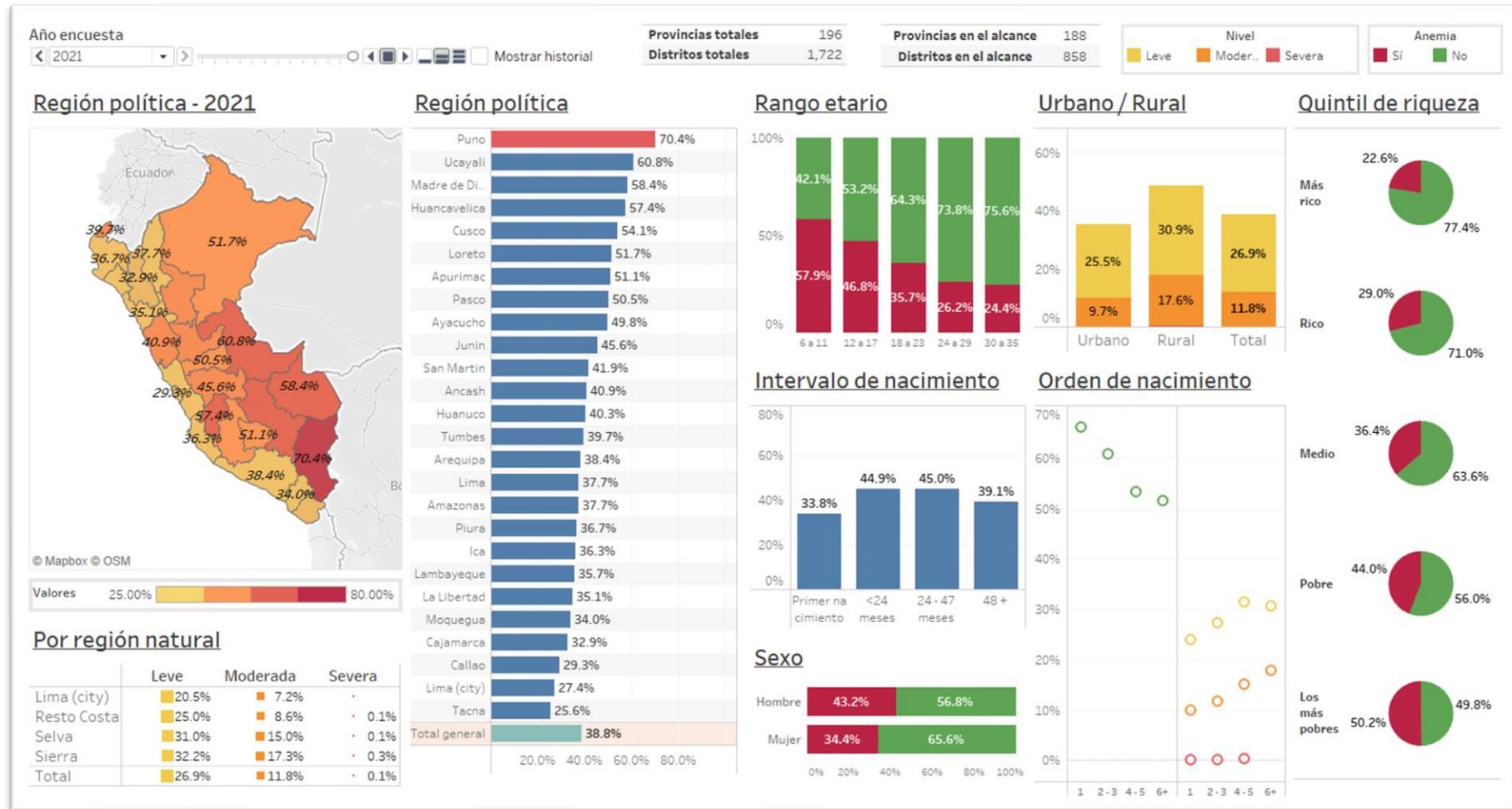
²² Herramienta utilizada SurveyMonkey, accedida en: <https://es.surveymonkey.com/mp/sample-size-calculator/>

Ilustración 54. Índice de anemia según aspectos demográficos y económicos (anualizado: 2009)



Fuente: Tableau – Elaboración propia

Ilustración 55. Índice de anemia según aspectos demográficos y económicos (anualizado: 2021)



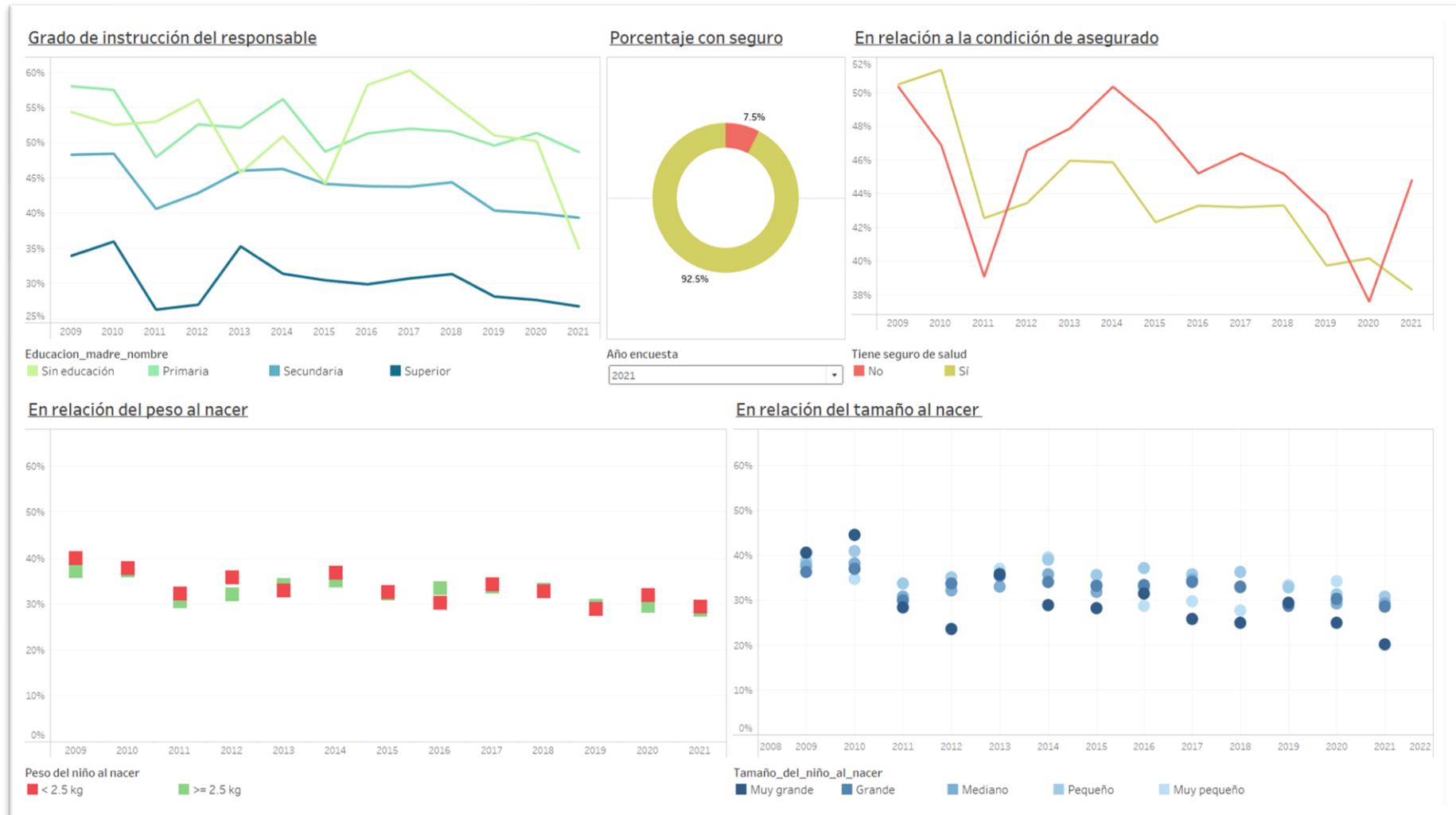
Fuente: Tableau – Elaboración propia

C. **Factores formativos y de salud:** en la ilustración 56, se presenta la evolución del índice de anemia con relación a la educación del responsable del niño, así como también con los indicadores del nacimiento y habilitación del seguro de salud. Esta visualización es un análisis evolutivo, a través de los años del alcance de esta evaluación. Con respecto al **grado de instrucción del responsable**, se puede apreciar que el índice de anemia es menor para niños cuyo responsable tiene un nivel educativo más alto. Por ejemplo, en el caso de responsables con nivel educativo superior (universitario o técnico), se presenta un indicador de anemia que no sobrepasa el 36% en ningún año. Estos resultados orientan a que el factor educativo en los padres influye notablemente en el índice de anemia. Así también se puede mencionar sobre este mismo ítem, que las personas que no tienen estudios formales (“Sin educación”²³), presentan un despliegue poco uniforme que no permite identificar de manera clara la relación.

Por otra parte, se procede a analizar el ítem de seguro de salud, que incluye el **porcentaje de niños con seguro** y la relación del **índice de anemia con la condición de asegurado**, aquí se puede percibir que el nivel de niños con seguro ya sea público o privado, es bastante elevado y para el año 2021 representa el 92,5%. Sobre este punto se verificó la relación del índice de anemia en función de la condición de asegurado y se identificó que, si bien a partir del 2011 el índice de anemia es menor en los niños que cuentan con seguro, se encontró una distorsión para el año 2020, en donde se aprecia que el índice de anemia es menor para los niños que no contaban con seguro. Según las respuestas recibidas por personal salud, esta situación se podría dar debido a que, en el 2020, año de pandemia, el grado de recurrencia de niños a los centros de salud para seguimiento y control disminuyó considerablemente. El siguiente ítem por evaluar es la relación del índice de anemia con respecto al **peso y talla al nacer**, para el caso de peso, se separó en dos grupos, los niños con peso menor de 2,5 kg y los que tenían un peso mayor o igual a 2,5 kg; y con respecto a la talla, se dividió en 5 medidas cualitativas desde muy pequeño, hasta muy grande. Para ambos casos no se pudo identificar una correlación significativa con estas variables, ya que los valores de porcentaje de anemia son similares o no son consistente a través de los años.

²³ Hace referencia a la educación formal impartida por el Ministerio de Educación.

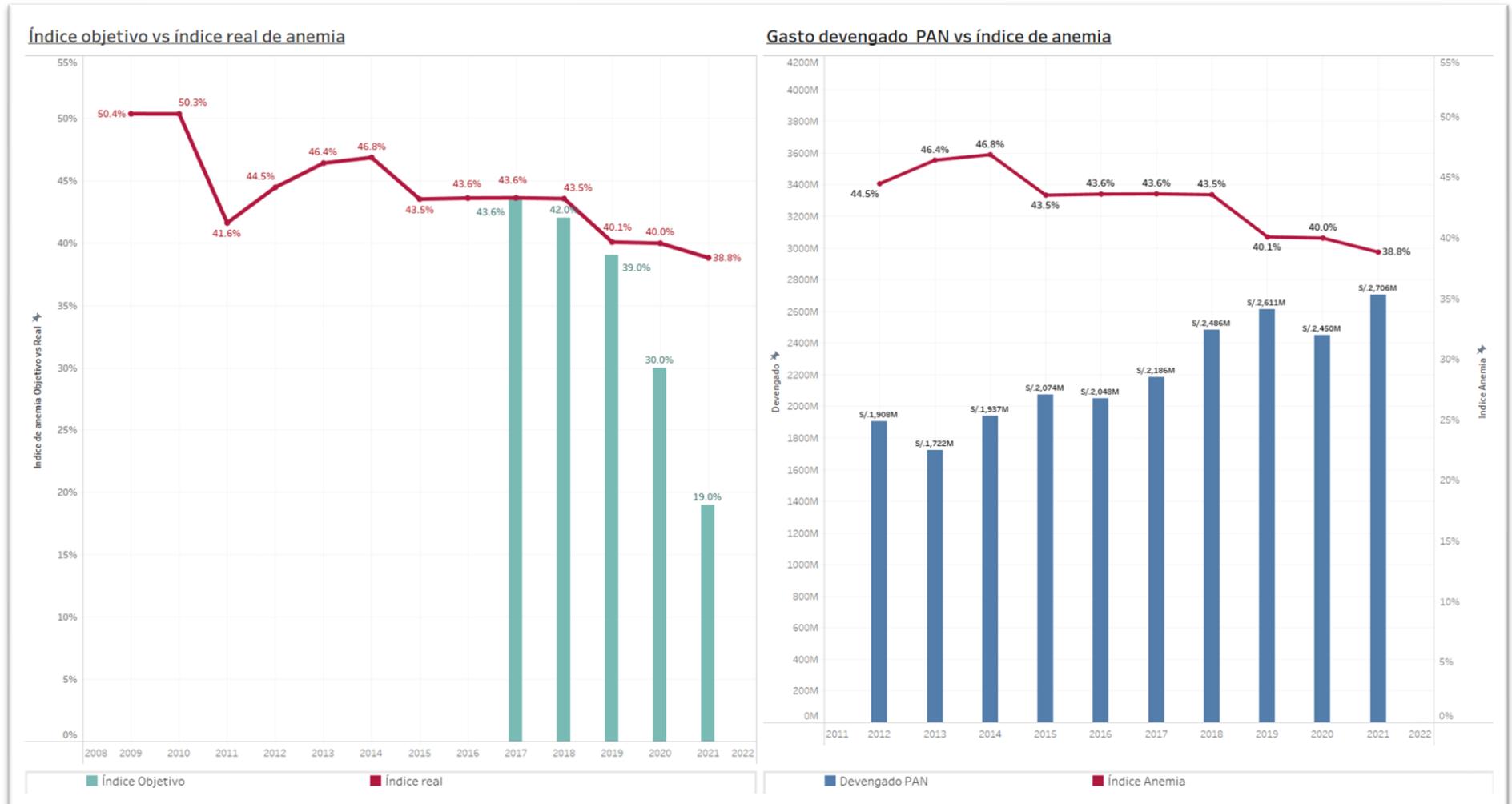
Ilustración 56. Índice de anemia según aspectos formativos y de salud



Fuente: Tableau – Elaboración propia

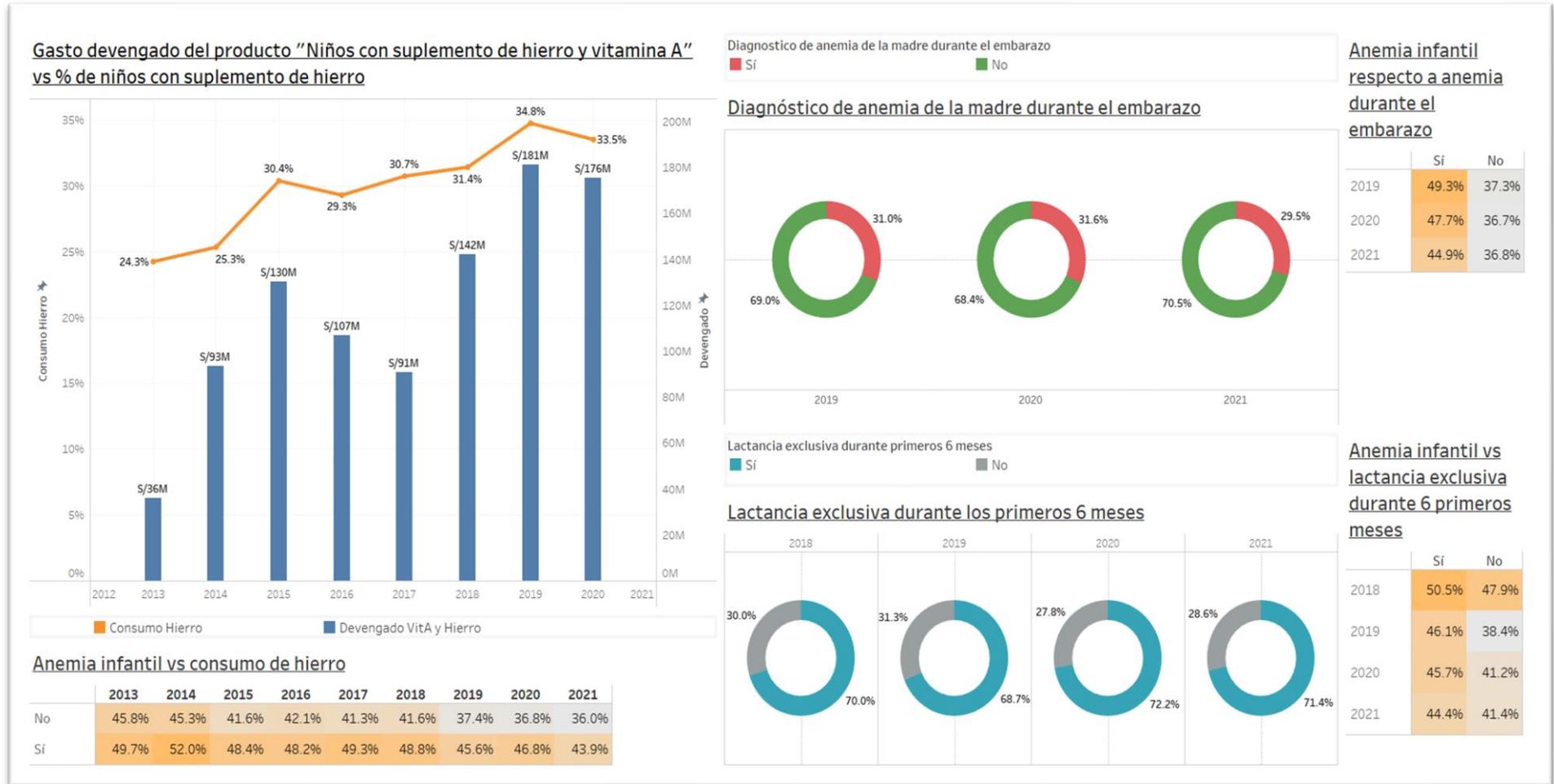
- D. Índice objetivo y gasto devengado PAN: en la Ilustración 57, en la imagen de la izquierda se aprecia un **comparativo del índice real, versus el índice objetivo**, y, como ya se ha mencionado durante el desarrollo del presente trabajo, es preocupante ver la distancia que hay entre estas dos medidas. Como se puede percibir, la brecha alcanzada entre el índice de anemia y el índice objetivo, para el año 2021, representa más del 100% (38,8% de índice real sobre 19% de índice objetivo). Esta situación está siendo evaluada a la fecha por el gobierno central y planeando una nueva estrategia para los años que vienen, por ello, respecto a este punto, es importante mantenernos vigilantes. Sobre el segundo gráfico, el que se encuentra a la derecha, **comparativa entre el gasto devengado PAN y el índice de anemia**, se incluyó al ser el PAN uno de los principales programas contra lucha de las diversas formas de malnutrición, incluyendo la anemia. Tiene entre sus indicadores, la reducción de la prevalencia de anemia en menores de 36 meses. Como se puede apreciar en la gráfica, a excepción del año 2020, el devengado para el PAN se ha venido incrementando año a año, y si bien se aprecia mejoras en el índice de anemia, no se puede ver una mejoría que corresponda significativamente a los incrementos del devengado en el PAN. Habría que analizar más a detalle este punto, incluyendo los otros indicadores que se consideran parte de la medición del logro del PAN, como lo son, la proporción de menores de 5 años con desnutrición crónica y los niños menores de 36 meses con lactancia materna exclusiva en los primeros 6 meses de vida.
- E. Consumo de hierro, anemia en embarazo y lactancia: en la Ilustración 58, en la imagen de la izquierda superior, se ha incluido el comparativo entre el **gasto devengado del producto “Niños con suplemento de hierro y Vitamina A”, en contraste con el porcentaje de niños con suplemento de hierro**, observándose que parecen tener una relación directa (a mayor inversión, mayor cobertura). Por otro lado, al realizar la comparativa entre niños que **consumieron hierro y el índice de anemia** (cuadro de la parte izquierda inferior), se identificó que el índice de anemia es mayor en los niños que consumieron suplemento de hierro, lo cual no resulta coherente, pero podría tener sustento en dos posibles causas: primero, que este indicador no especifica la proporción de suplemento que se brindó, así que en muchos casos podría ser insuficiente y segundo, que el entregar el suplemento, no confirma el consumo y menos, el consumo adecuado. En relación con **la anemia durante el embarazo**, se puede apreciar que el índice de niños con anemia es mayor en el grupo de madres que tuvieron anemia durante el embarazo. Finalmente, con respecto a la **lactancia exclusiva**, se aprecia que el índice de anemia es mayor entre los niños con lactancia exclusiva. Esto es un poco extraño, ya que es una recomendación de la OMS, pero se podría explicar por factores adicionales, como lo es la anemia durante el embarazo.

Ilustración 57. Comparación con el índice objetivo y gasto devengado PAN



Fuente: Tableau – Elaboración propia

Ilustración 58. Consumo de hierro, anemia durante el embarazo y lactancia exclusiva



Fuente: Tableau – Elaboración propia

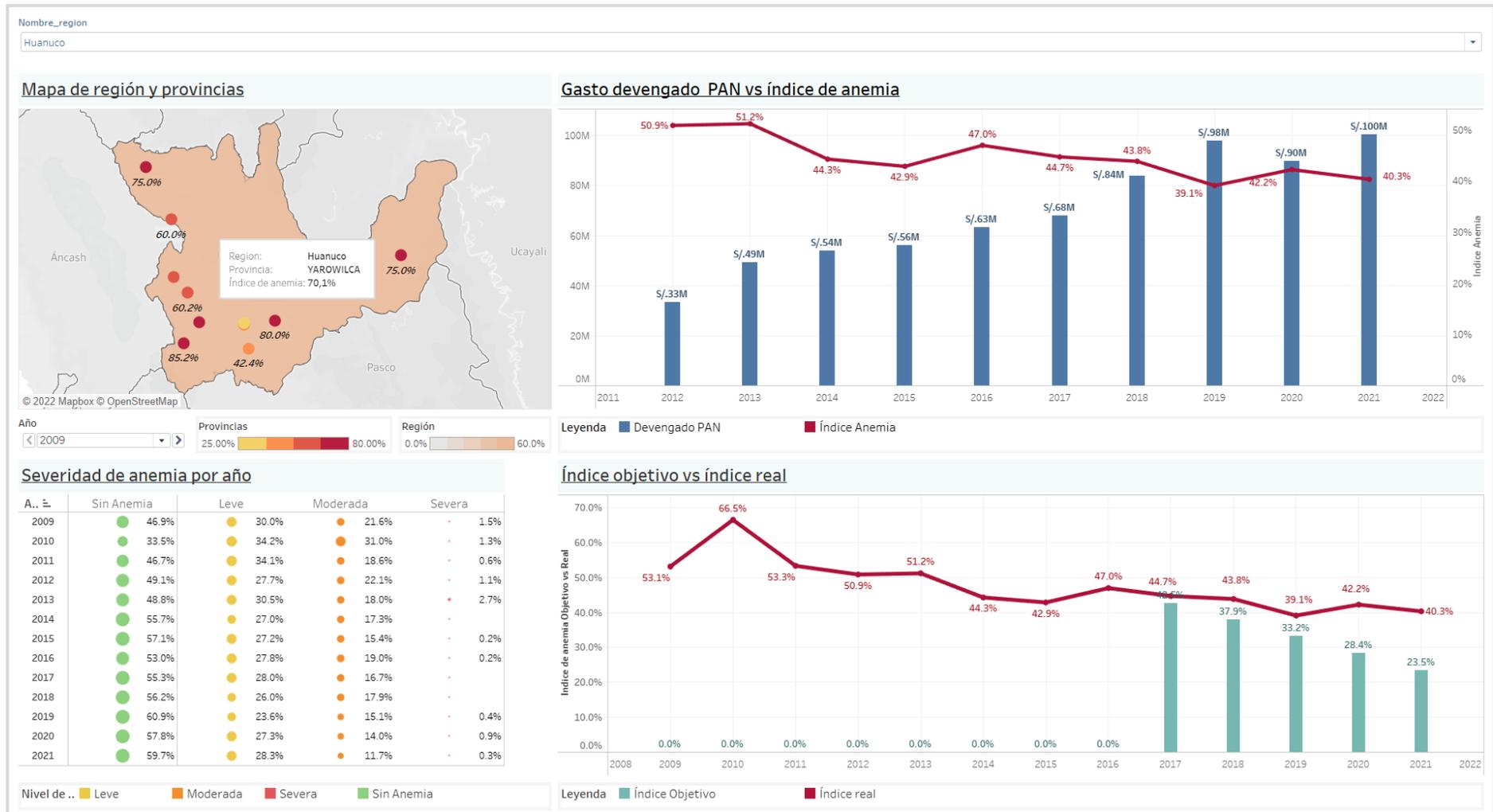
6.2.2. Análisis de indicadores regionales

Aquí se presenta un análisis detallado por regiones, así como también una vista general del índice de anemia en las provincias y distritos que componen la región. Los indicadores son similares a los globales, solo que se han priorizados los que pueden ser de interés para implementar planes de acción por parte de los gobiernos regionales, es decir, se consideran las variables con las cuales se puede trabajar para mejorar.

Como una evaluación general, a continuación, se describen los tableros elaborados y los principales puntos de atención.

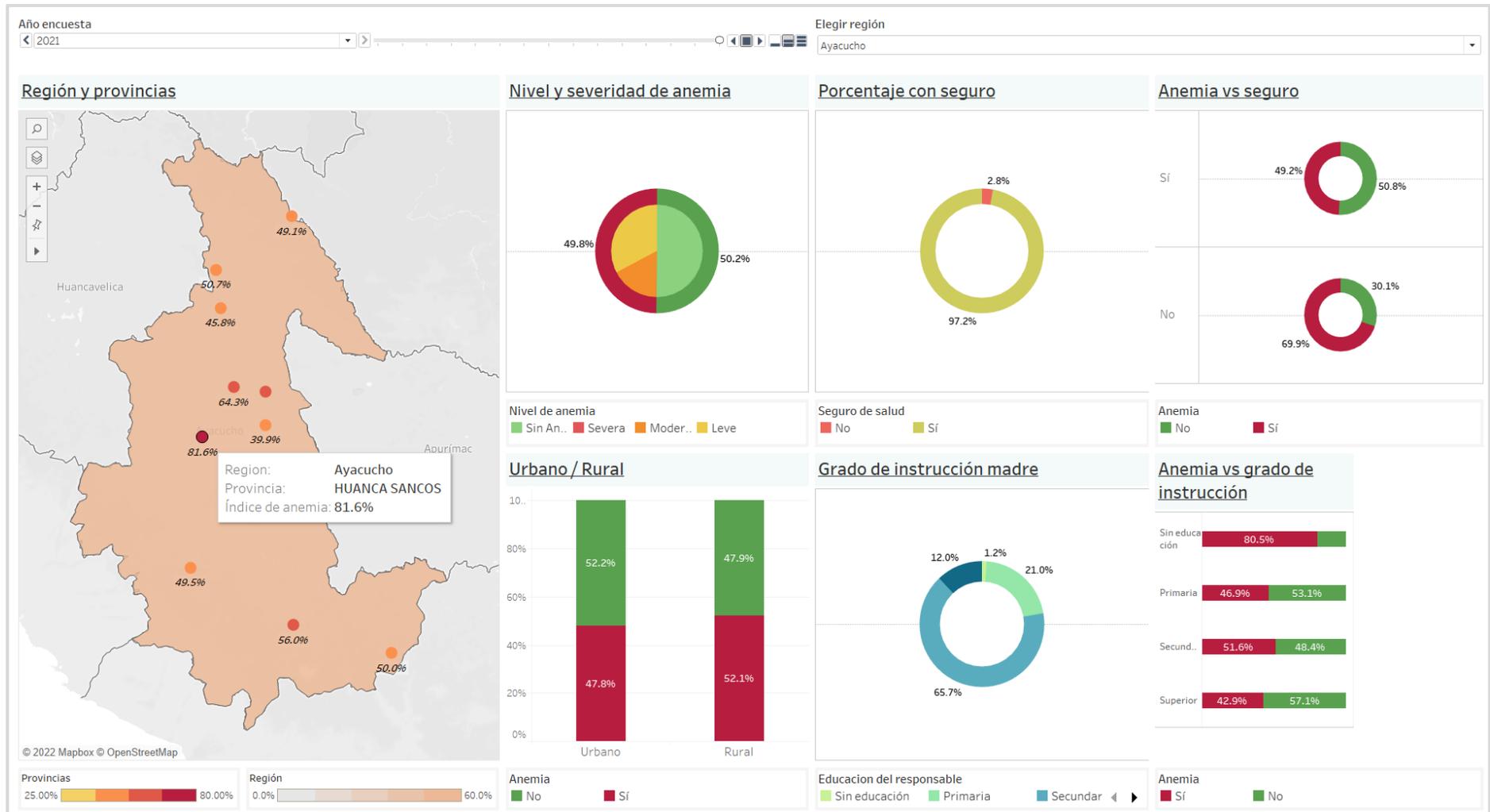
- A. Devengado y objetivo por regiones: en la Ilustración 59 se presenta el despliegue del índice de anemia por regiones y provincias. En la representación denominada **mapa de región y provincias**, se presentan los índices específicos por estos dos niveles geográficos, esto se realiza con el objetivo de brindar mayor visibilidad sobre los lugares en donde se requiere impulsar de mejor manera los programas sociales. En este dashboard de la región Huánuco, como ejemplo, se aprecia que se tiene una caja de texto emergente en el mapa, con lo cual se puede ver que la provincia Yarowilca es una de las más afectadas, con un índice de 70,1%. De igual manera, se realiza el análisis evolutivo con **gasto devengado del PAN y con los objetivos de la región**, en donde se aprecia una inversión creciente del gasto devengado (a excepción del año 2020), pero no se refleja la mejoría en el índice de anemia. Finalmente, también se tiene la evolución del índice de anemia por **nivel de severidad y por año**, el cual se encuentra debajo del mapa.
- B. Análisis por región y año: en la Ilustración 60 se aprecian los principales indicadores en una vista por región y por año. Como el alcance es de 25 regiones y 12 años, el total de vistas posibles es de 300. Es importante tener este nivel detalle porque permite hacer un mejor análisis de acuerdo con lo que requiere cada región. En este caso, a modo de ejemplo se realizará el análisis sobre la región Ayacucho, para el año 2021. A nivel general se aprecia que el indicador de anemia es alto para toda la región, sin embargo, se debe tener especial atención por la provincia Huanca Sancos que presenta un valor de 81,6%. Otro punto relevante, es que a pesar de tener una cobertura de más del 97% de personas con seguro de salud, no presenta mejor indicador en contraste con otras regiones que tienen una menor cobertura de seguro.
- C. Análisis global de provincias y distritos: en la Ilustración 61 se muestran los mapas con el índice de anemia por provincia y distrito y que se puede ir analizando anualmente. La idea es tener una visión general que permita poner foco en las regiones, provincias o distritos que presenten un indicador más desfavorable.

Ilustración 59. Números globales y comparativos con objetivo y devengado - Regional



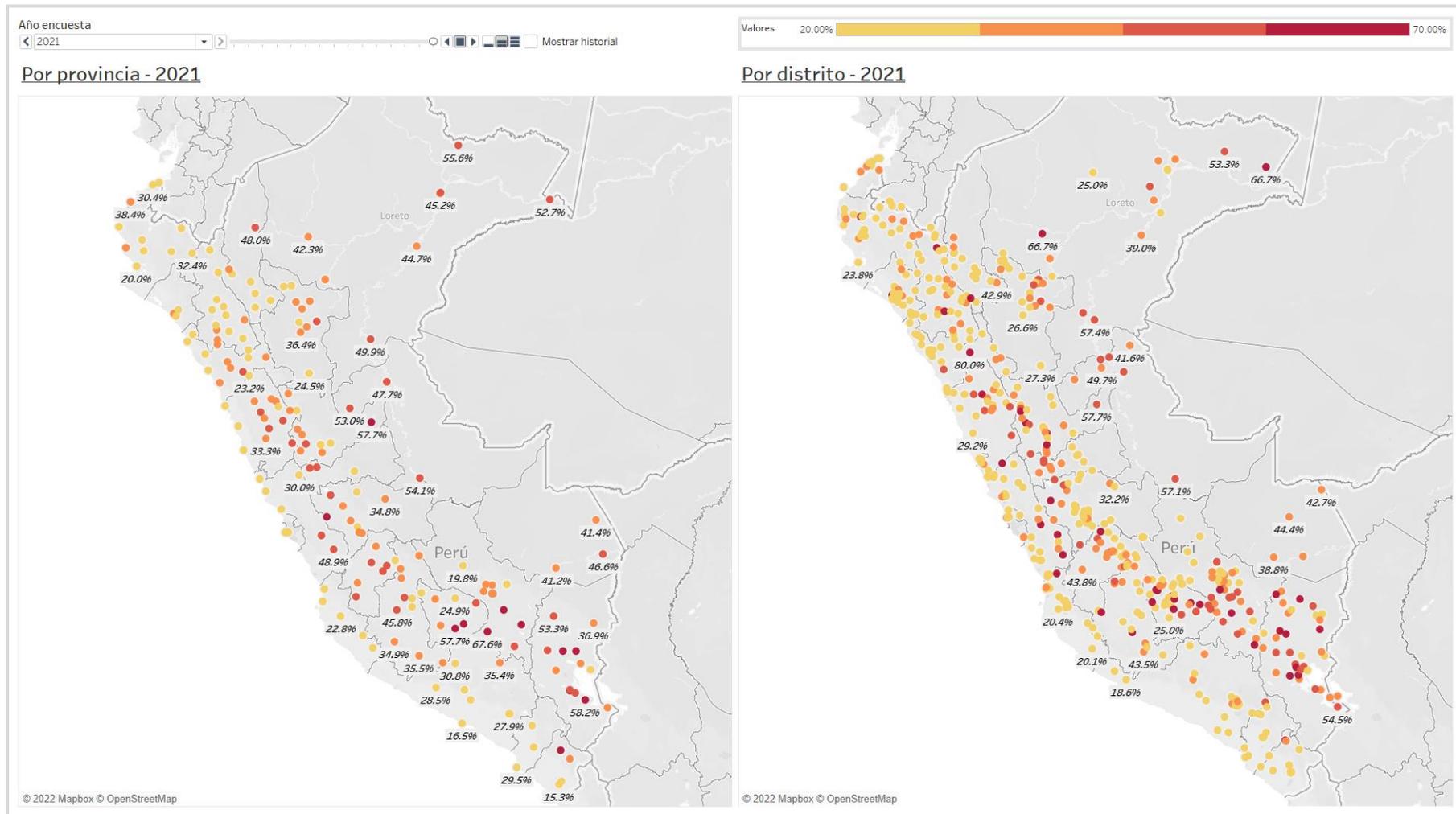
Fuente: Tableau – Elaboración propia

Ilustración 60. Análisis de principales indicadores por región y año



Fuente: Tableau – Elaboración propia

Ilustración 61. Vista global por provincias y distritos



Fuente: Tableau – Elaboración propia

6.3 Principales dificultades y soluciones planteadas

Si bien se logró ejecutar todas las tareas programas según la planificación establecida, cabe mencionar que hubo algunas dificultades que demandaron un gran número de horas y que son importantes incluir para que sirvan de aprendizaje a futuros trabajos.

- Restricciones de geolocalización: Varias páginas webs que han sido consulta para la realización del presente trabajo, tenían restricciones de acceso geolocalizado, es decir, permiten el acceso en el territorio peruano, pero están restringidas para otros países, por lo que realizar el trabajo desde España se dificultó un poco, teniendo que crear una cuenta VPN para poder realizar las consultas iniciales y entendimientos del flujo de datos.
- Reprocesamiento de datos: La consolidación inicial de los datos fue semi automatizada, es decir, se utilizó Python para la descarga de datos, para luego descomprimirlos y cargar solo los que corresponde a la plataforma de trabajo para su consolidación por año. Esta situación demandó gran cantidad de tiempo, sobre todo cuando por omisión no se incluían todas las columnas necesarias y luego se tenía que reprocesar los más de 30 archivos por año y considerando que se debía procesar 12 años, la tarea se volvió insostenible. Por este motivo, se decidió realizar una mayor automatización, de modo que los reprocesos, fueran fáciles de gestionar y consuman menos tiempo de operación.
- Política CORS de restricción de accesos: CORS es acrónimo de Cross Origin Resource Sharing, la cual es una medida de seguridad que evita accesos desde un origen distinto a la web. Esta situación se generó al realizar una solicitud de información a la web denominada “Consulta Amigable”, que administra el MEF, se mostraba un mensaje de acceso denegado, por lo cual las descargas no se podían realizar por este medio. Al inicio fue muy difícil identificar el error, pero luego, se encontró que esta casuística le había sucedido a otro investigador, por ello se tomó como referencia su casuística (*Sebastiá Pérez 2022*). Como solución para la descarga de estos datos, se utilizó la herramienta Postman, ya que dicha herramienta tiene integrada una solución para superar la política CORS.
- Distorsiones en los tipos de archivos de las fuentes de datos: Como se ha mencionado en algunas secciones del informe, las tablas que contenían los datos que se requerían como fuente de información, no estaban homologados, ni a nivel formato, ni a nivel contenido (por ejemplo, en relación con los tipos de datos o la longitud de estos). Esta situación representó un gran reto ya que, por la cantidad de datos procesados, estas situaciones no eran fácilmente identificables y en ocasiones, truncaban la descarga de los datos.
- Errores en los programas utilizados: en la herramienta Tableau, se identificó un problema al hacer uso de la carga de información desde base de datos, ya que en los valores que representaban sumatorias o valores porcentuales, se agregaba de manera arbitraria un número al final del valor, los que generaba distorsiones significativas. Al incluir esta consulta

en foros de ayuda, la respuesta fue que es un error conocido, pero se brindó una forma de hacer frente a esta situación.

- Aprendizaje de nuevas herramientas: Como parte muy provechosa, ha sido el aprendizaje de diversas herramientas que eran necesarias como complemento para poder desplegar la solución propuesta. Este ítem, si bien es productivo en todo sentido, sí ha requerido un tiempo considerable que no estaba estimado inicialmente. La solución fue ampliar ligeramente el cronograma para cumplir con las actividades programadas y poder lograr los objetivos planteados.
- Puesta en marcha del hosting: al realizar el despliegue de la página web, se tuvo varios desafíos; desde solucionar errores de carga, hasta la puesta en marcha de certificados de seguridad. Si bien es cierto que la mayoría de los problemas fueron por falta de conocimientos en el tema, se rescata que, al momento de la elección del hosting, se tomen en cuenta los criterios de costo, facilidad de uso, disponibilidad de soporte, compatibilidad con los lenguajes de programación y bases de datos; pero sobre todo, que se tome como factor muy relevante que exista y se tenga a disposición, suficiente documentación y manuales que ayuden al proceso.

Conclusiones

A modo de cierre, se puede señalar que el estudio realizado ha brindado un punto de partida para poder entender, de mejor manera, la situación actual de la anemia infantil en Perú, y con ello, plantear propuestas para hacer frente a esta realidad.

Como principal aporte, se la logrado la implementación de una versión inicial de la plataforma de acceso abierto llamada “SanitoPe”, que pone a disposición del público en general, visualizaciones dinámicas, análisis realizados y datos consolidados en relación con el índice de anemia infantil, cumpliendo de este modo, el objetivo principal de este proyecto.

Con respecto a las visualizaciones de datos realizadas, han permitido tener mejor visibilidad de la relación entre las variables que fueron seleccionadas, tal es así, que de acuerdo con los datos analizados del índice de anemia y su relación con factores demográficos, económicos, formativos y de salud, se han confirmado algunas situaciones conocidas, como la gran incidencia de anemia que se presenta en los sectores más pobres o en regiones con gran altitud; pero a la vez, se han identificado algunas situaciones que se recomienda analizar con mayor profundidad, como son: el impacto en el índice de anemia, del consumo de hierro y de la lactancia exclusiva en los primeros 6 meses. El primer punto se considera relevante porque, a pesar de que, en los últimos años se ha logrado un mayor alcance en la cantidad de niños en programas de consumo de hierro, no se ha visto una mejora significativa del índice de anemia entre los niños que participaban en el programa y los que no. Como se menciona previamente en este trabajo, esto puede deberse al consumo incorrecto o no consumo del hierro, pero de ser el caso, se deberían aplicar medidas para superar esta situación. Con respecto a la lactancia exclusiva en los primeros 6 meses de vida, se ha notado que el índice de anemia es mayor entre la población de niños que tuvo lactancia exclusiva. Si bien es cierto que la lactancia exclusiva es una iniciativa que se ha fomentado por muchos años y que el impacto positivo tiene respaldo científico, es posible que esta situación se presente por el alto número de mujeres que presentan anemia u otras formas de malnutrición y que al no solucionar estas deficiencias, puede que afecte al perfil nutricional del lactante.

Sumado al párrafo previo, en relación con el impacto de los programas sociales implementados, se ha observado que si bien en los últimos años se viene incrementando los presupuestos, tanto a nivel nacional como regional, esta inversión no se ha visto reflejada de manera proporcional en el indicador que mide el índice de anemia, por lo que se debería realizar un análisis de la estrategia para poder, de ser necesario, reformularla.

A nivel de presentación de las visualizaciones en la web, se ha logrado tener gráficos de fácil entendimiento, adaptativos al tipo de dispositivo y que permiten tener un nivel de detalle por año y por región. Es importante mencionar que se han priorizado la simplicidad de uso y el diseño adaptativo, para que puedan ser utilizadas sin problemas en cualquier dispositivo y sin necesidad de tener

experiencia en tecnologías de desarrollo. Así también, se ha contemplado que la herramienta permita poder reutilizar los conjuntos de datos y visualizaciones trabajadas y que los usuarios puedan compartir los trabajos que realicen, a fin de sumar esfuerzos colaborativos en beneficio de la comunidad.

En relación con el desarrollo del trabajo, es relevante mencionar que los problemas que se presentaron en el camino de ejecución no han sido pocos. Se han encontrado dificultades como restricciones geográficas de acceso, políticas limitantes en los portales web o la ya conocida disparidad en las estructuras y formatos de datos. Sin embargo, estas situaciones han representado una oportunidad de poner a prueba el pensamiento analítico y creativo, y con ello, encontrar una solución a estas casuísticas, mediante el uso de tecnologías alternativas o estructurando soluciones automatizadas.

Compartiendo una de las lecciones aprendidas más significativas, se puede mencionar que uno de los principales errores cometidos en el desarrollo de este trabajo, ha sido el reproceso de descarga y procesamiento de los datos. Esta situación se presentó por la escasa validación inicial y por realizar correcciones puntuales, en vez de buscar las causas raíz. El aprendizaje fue dedicar un poco más de tiempo a la validación, identificación y corrección de las causas que ocasionaban los errores.

Un punto muy favorable que ha sido logrado con la ejecución de este trabajo ha sido poder integrar y poner en práctica los conocimientos adquiridos en el Máster Universitario de Gestión de la Información y sumar nuevos conocimientos y tecnologías que no se habían revisado en profundidad. Estas tecnologías representan una gran utilidad en el presente y entre los principales conocimientos adquiridos se pueden mencionar: Uso de la API Google Maps para la obtención de datos de geolocalización, poner en marcha en un dominio propio una web responsive, uso avanzado de Tableau, mayor aprendizaje de programación web y aplicación de sentencias SQL para consultas en la base de datos. Sobre este mismo punto, se considera que la selección de tecnologías utilizadas ha sido un acierto, ya que, a pesar de presentar dificultades iniciales por la falta de conocimientos, luego de capacitaciones y entrenamiento, han permitido el logro de los objetivos planteados.

Por último, hay que precisar que se conoce que las herramientas presentadas en este trabajo no serán la solución definitiva para disminuir el índice de anemia radicalmente, ni que serán influyentes a corto plazo; sin embargo, se espera que sea una contribución importante para que cada vez más interesados en esta problemática, puedan aportar sus conocimientos y, de esta manera, lograr mejores resultados. Y como menciona Alicorp²⁴ en una conocida frase que reafirma su ideología y la mía: *“Alimentemos un mañana mejor”*, para ti, para mí, para ellos y para ellas. Juntos somos más fuertes.

²⁴ Alicorp es uno de los principales grupos económicos del Perú que apoya la lucha con las distintas formas de malnutrición.

Relación del trabajo con los estudios cursados

Entre las materias impartidas en el Máster Universitario de Gestión de la Información (MUGI), que han facilitado la ejecución del presente trabajo, se pueden mencionar:

- Explotación de Datos Masivos: En esta materia se enseñó a utilizar Python como tecnología para la obtención, análisis y explotación de datos que son de grandes volúmenes, así también, su uso en situaciones que simulaban un entorno real ayudó para un mejor entendimiento y uso adecuado de las herramientas que brinda.
- Datos Abiertos en la Web: A través de este curso se obtuvo el entendimiento de los principales repositorios de datos, así como también los criterios que se deberían tomar en cuenta para poder evaluar la calidad de los datos abiertos. En esta materia se impartió el uso del análisis MELODA, el cual fue utilizado como parte de este trabajo.
- Servicios de Datos y Contenidos: Esta materia se orientó a la identificación de la importancia y el poder que tiene los datos. Además, mediante la aplicación de prácticas con datos reales, permitió identificar las posibles complicaciones y las consideraciones que se deben tener en cuenta cuando se recopilan datos.
- Visualización de la Información: permitió aprender estrategias de presentación de datos de manera gráfica, que permitan comunicar, identificar patrones, entender la información y ayudar a una adecuada toma de decisiones, generando, de esta manera, un impacto positivo en las personas que hacen uso de esta herramienta.

En relación con las competencias transversales, se puede mencionar que el desarrollo del trabajo ha ayudado a desarrollar principalmente las siguientes:

- Aplicación y pensamiento práctico: Por la búsqueda de la mejor alternativa de solución que permita aportar de manera significativa al problema planteado.
- Análisis y resolución de problemas: Porque en diversos puntos de la ejecución del presente trabajo se ha evaluado distintas soluciones a las problemáticas presentadas y se ha podido optar por la mejor opción acorde a la evaluación realizada.
- Responsabilidad ética, medioambiental y profesional: Durante el desarrollo de todo el trabajo se ha tomado en cuenta los tipos de licencia con los que se contaba y cómo utilizar y citar a fin de poder ser consecuentes con los requerimientos de los autores. También en relación con este punto, se ha evaluado cómo el presente trabajo aporta al logro de los ODS.
- Aprendizaje permanente: A pesar de las dificultades y carencias de conocimiento, no se ha escatimado esfuerzos en el estudio y adquisición de nuevas herramientas y se ha buscado seguir la línea de aprendizaje sobre temas de los cuales se tenía un conocimiento limitado.

Trabajos futuros

Nuestra visión es que el presente proyecto sea el inicio de un trabajo colaborativo, que permita nutrir la información contenida en la web de SanitoPe y que ayude a dar valor agregado para hacer frente al problema expuesta.

Algunos trabajos futuros que podrían llevarse a cabo serían:

- Implementar nuevas visualizaciones que sean útiles para sectores como salud, educación o sectores privados que se encuentren interesados en identificar el impacto de los programas nutricionales que fomentan.
- Crear nuevos conjuntos de datos desde información proveniente de otras instituciones que manejan otros datos y que también son importantes analizar. Aquí se incluye los datos más granulares del MINSA y los datos del gasto devengado para otros ítems distintos al analizado en este trabajo.
- Analizar a mayor detalle los puntos comentados en relación con factores de especial interés como los son, el consumo de los suplementos de hierro y la lactancia exclusiva. Así también, incluir en futuras evaluaciones el cuestionamiento a las formas de ajuste por altitud, actualmente utilizadas.
- Configurar perfiles de usuario y brindarles accesos de acuerdo con sus necesidades, con el objetivo de que puedan mejorar su experiencia, sus expectativas sean cubiertas y además poder tener una mejor trazabilidad sobre las actividades desarrolladas.
- Solicitar la difusión de la página en el portal de datos abiertos del Perú, ya que el enfoque del presente trabajo está alineado al objetivo de dicho portal.
- Agregar una funcionalidad para que los usuarios puedan subir nuevas secuencias de comandos que puedan ser reutilizados, modificados y mejorados. Con esto se incrementará el trabajo colaborativo.
- Optimizar el código utilizado a fin de poder lograr consultas más eficientes que permitan dar un mejor servicio a los usuarios de la plataforma. Asimismo, mejorar la apariencia del entorno para que sea más intuitivo y atractivo al usuario de la web SanitoPe.
- Mejorar las funciones de la página a nivel SEO, utilizando auditorías que permitan ver las brechas que se tienen y plantear mejoras continuas.
- Automatizar la descarga y actualización de datos, asignando los periodos que corresponden por ejemplo los datos del ENDES, se deberían actualizar anualmente, pero los datos del devengado ejecutado, se actualizarían mensualmente.
- Evaluar el impacto y cómo el uso de la web promueve el análisis de la problemática expuesta. Se propone realizar encuestas y focus group entre los principales sectores de interés para poder determinar la utilidad de la herramienta.

- ABELLA, Alberto, ORTIZ-DE-URBINA-CRIADO, Marta y DE-PABLOS-HEREDERO, Carmen, 2020. Meloda 5: A metric to assess open data reusability. *El Profesional de la Información* [en línea]. 8 enero 2020. Vol. 28, no. 6, pp. 1-10. [Consulta: 12 mayo 2022]. DOI 10.3145/epi.2019.nov.20. Disponible en: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2019.nov.20>
- BALLON-SALCEDO, Carlos, CCAMI-BERNAL, Fabricio, RAMOS-FLORES, Yury, SIERRA-MORALES, Stefany y MORENO-LOAIZA, Oscar, 2020. Consumo de hierro y prevalencia de anemia en niños y adolescentes en una comunidad a gran altitud en Perú. *Revista Española de Nutrición Comunitaria* [en línea]. 1 octubre 2020. Vol. 26, no. 4, pp. 198-203. [Consulta: 15 junio 2022]. DOI 10.14642/RENC.2020.26.4.5337. Disponible en: https://www.renc.es/imagenes/auxiliar/files/NUTRICION_COMUNITARIA_4-2020_Art3.pdf
- BELTRÁN-NAVARRO, Beatriz, MATUTE, Esmeralda y VÁSQUEZ-GARIBAY, Edgar M., 2019. Efecto de la deficiencia de hierro sobre el desarrollo neuropsicológico en lactantes. *Interdisciplinaria* [en línea]. 2019. Vol. 36, no. 2, pp. 129-150. [Consulta: 8 julio 2022] DOI 10.16888/interd.2019.36.2.9. Disponible en: <https://www.redalyc.org/journal/180/18060566026/html/>
- DELFIN, Mauricio, 2021. Gobernanza cultural y datos abiertos: Retos y oportunidades para los sectores culturales en el Perú. En: *ILDA, documentos de trabajo* [en línea]. 7 abril 2021. DOI 10.5281/zenodo.4670293. Disponible en: <https://zenodo.org/record/4670293>. [Consulta: 13 junio 2022].
- GUABLOCHE, Judith, 2021. Anemia infantil en el Perú: análisis de la situación y políticas públicas para su reducción. *Revista Moneda, Banco Central de Reserva* [en línea]. marzo 2021. Vol. 185, pp. 48-55. [Consulta: 12 mayo 2022]. ISSN: 1991 - 0606. Disponible en: <https://www.bcrp.gob.pe/docs/Publicaciones/Revista-Moneda/moneda-185/moneda-185.pdf>
- HUAMÁN-ESPINO, Lucio, APARCO, Juan Pablo, NUÑEZ-ROBLES, Eloisa, GONZÁLES, Elena, PILLACA, Jenny y MAYTA-TRISTÁN, Percy, 2012. Consumo de suplementos con multimicronutrientes Chispitas® y anemia en niños de 6 a 35 meses: estudio transversal en el contexto de una intervención poblacional en Apurímac, Perú. *Revista Peruana de Medicina Experimental y Salud Pública* [en línea]. 2012. Vol. 29, no. 3. pp. 314-323. [Consulta: 23 junio 2022]. DOI 10.17843/rpmesp.2012.293.363. Disponible en: <https://rpmesp.ins.gob.pe/index.php/rpmesp/article/view/363>
- INEI, 2018. Metodología de la medición de la anemia en línea. En: *INEI (Instituto Nacional de Estadística e Informática)* [en línea]. Disponible en: https://proyectos.inei.gob.pe/endes/2018/documentos_2018/METODOLOGIA_DE_LA_MEDICION_DE_LA_ANEMIA.pdf. [Consulta: 3 mayo 2022].
- INEI, 2020. INEI - Perú: Encuesta Demográfica y de Salud Familiar 2019 - Nacional y Regional. En:

- INEI (*Instituto Nacional de Estadística e Informática*) [en línea]. Disponible en: https://www.inei.gov.pe/media/MenuRecursivo/publicaciones_digitales/Est/Endes2019/. [Consulta: 23 junio 2022].
- INEI, 2021. Informe Principal ENDES 2020. En: *INEI (Instituto Nacional de Estadística e Informática)* [en línea]. Disponible en: https://proyectos.inei.gov.pe/endes/2020/INFORME_PRINCIPAL_2020/INFORME_PRINCIPAL_ENDES_2020.pdf [Consulta: 10 julio 2020].
- INICIATIVA LATINOAMERICANA POR LOS DATOS ABIERTOS, Open Data Barometer, sin fecha. En: *ILDA (Iniciativa Latinoamericana por los Datos Abiertos)* [en línea]. Disponible en: https://barometrolac.org/?_year=2020&indicator=ODB [Consulta: 8 junio 2022].
- MIDIS, 2017. Plan Multisectorial de Lucha contra la Anemia. En: *UNESCO (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura)* [en línea]. Disponible en: https://siteal.iiep.unesco.org/sites/default/files/sit_accion_files/plan_multisectorial_de_lucha_contra_la_anemia_2018.pdf [Consulta: 13 junio 2022].
- MINISTERIO DE SALUD, Observatorio de Anemia, sin fecha En: *MINSA (Ministerio de Salud del Perú)* [en línea]. Disponible en: <https://www.gob.pe/institucion/midis/noticias/19142-observatorio-de-anemia> [Consulta: 26 abril 2022].
- MUNAYCO, César V., ULLOA-REA, María E., MEDINA-OSIS, José, LOZANO-REVOLLAR, Carmen R., TEJADA, Violeta, CASTRO-SALAZAR, Consuelo, MUNARRIZ-VILLAFUERTE, José, DE BUSTOS, Cecilia y ARIAS, Lena, 2014. Evaluación del impacto de los multimicronutrientes en polvo sobre la anemia infantil en tres regiones andinas del Perú. *Revista Peruana de Medicina Experimental y Salud Pública* [en línea]. 10 febrero 2014. Vol. 30, no. 2. [Consulta: 2 mayo 2022]. DOI 10.17843/rpmesp.2013.302.196. Disponible en: <https://rpmesp.ins.gob.pe/index.php/rpmesp/article/view/196>
- OPEN KNOWLEDGE FOUNDATION. Open Data Commons Open Database License (ODbL) — Open Data Commons: legal tools for open data, sin fecha. En: *Open Knowledge foundation* [en línea]. Disponible en: <https://opendatacommons.org/licenses/odbl/> [Consulta: 11 julio 2022].
- ORGANIZACIÓN MUNDIAL DE LA SALUD. Anemia, sin fecha. En: *Organización Mundial de la Salud* [en línea]. Disponible en: <https://www.who.int/es/health-topics/anaemia> [Consulta: 11 junio 2022].
- ORGANIZACIÓN MUNDIAL DE LA SALUD, 2011. Concentraciones de hemoglobina para diagnosticar la anemia y evaluar su gravedad. En: *ORGANIZACIÓN MUNDIAL DE LA SALUD* [en línea]. 31 mayo 2011. Disponible en: <https://www.who.int/es/publications/i/item/WHO-NMH-NHD-MNM-11.1> [Consulta: 28 abril 2022].
- ORTIZ ROMANÍ, Katherine Jenny, ORTIZ MONTALVO, Yonathan Josué, ESCOBEDO ENCARNACIÓN, Josselyne Rocio, DE LA ROSA, Luis Neyra y JAIMES VELÁSQUEZ, Carlos Alberto, 2021. Análisis del modelo multicausal sobre el nivel de la anemia en niños de 6 a 35 meses en Perú. *Enfermería Global* [en línea]. 2021. Vol. 20, no. 64, pp. 426-455. [Consulta: 2 mayo 2022].

DOI 10.6018/eglobal.472871. Disponible en:
https://scielo.isciii.es/scielo.php?script=sci_abstract&pid=S1695-61412021000400426&lng=es&nrm=iso&tlng=es

SEBASTIÁ PÉREZ, Daniel, 2022. *Desarrollo de una aplicación web para acercar datos públicos abiertos a los ciudadanos* [en línea]. Trabajo Fin de Grado. Universidad Politécnica de Valencia. [Consulta: 10 julio 2022]. Disponible en:
<https://riunet.upv.es/bitstream/handle/10251/182108/Sebastia%20-%20Desarrollo%20de%20una%20aplicacion%20web%20para%20acercar%20datos%20publicos%20abiertos%20a%20los%20ciudadanos.pdf?sequence=1&isAllowed=y>

SOBRINO, Manuel, GUTIÉRREZ, César, CUNHA, Antonio J, DÁVILA, Miguel y ALARCÓN, Jorge, 2014. Desnutrición infantil en menores de cinco años en Perú: tendencias y factores determinantes. *Rev Panam Salud Publica* [en línea]. 2014. Vol. 35, no. 2, pp. 104-112 [Consulta: 8 julio 2022]. Disponible en: http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S1020-49892014000200004

STACK OVERFLOW, 2021. Stack Overflow Developer Survey 2021: Explorando los resultados del año. En: *Stack Overflow* [en línea]. 8 agosto 2021. Disponible en:
<https://insights.stackoverflow.com/survey/2021/> [Consulta: 10 julio 2022].

ZAVALETA, Nelly, 2021. ANEMIA INFANTIL RETOS Y OPORTUNIDADES. *Revista Peruana de Medicina Experimental y Salud Pública* [en línea]. 2021. Vol. 34, no. 4, pp. 1-17. [Consulta: 3 mayo 2022]. DOI 10.17843/rpmesp.2017.344.3281. Disponible en:
<https://doi.org/10.17843/rpmesp.2017.344.3281>

Anexo I: Relación con ODS

Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

Tabla 14. Relación del trabajo con los ODS

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No Procede
ODS 1. Fin de la pobreza.				X
ODS 2. Hambre cero.	X			
ODS 3. Salud y bienestar.		X		
ODS 4. Educación de calidad.			X	
ODS 5. Igualdad de género.				X
ODS 6. Agua limpia y saneamiento.				X
ODS 7. Energía asequible y no contaminante.				X
ODS 8. Trabajo decente y crecimiento económico.			X	
ODS 9. Industria, innovación e infraestructuras.				X
ODS 10. Reducción de las desigualdades.		X		
ODS 11. Ciudades y comunidades sostenibles.				X
ODS 12. Producción y consumo responsables.				X
ODS 13. Acción por el clima.				X
ODS 14. Vida submarina.				X
ODS 15. Vida de ecosistemas terrestres.				X
ODS 16. Paz, justicia e instituciones sólidas.	X			
ODS 17. Alianzas para lograr objetivos.				X

Elaboración propia bajo formato establecido por la ETSINF²⁵

²⁵ El formato se puede encontrar en el siguiente repositorio: <https://www.inf.upv.es/www/etsinf/es/plantilla-tfg/>

Reflexión sobre la relación del TFM con los ODS y con el/los ODS más relacionados.

El presente trabajo está orientado en aportar valor para la búsqueda de una solución a una problemática que viene afectando al Perú por muchos años, como es la anemia infantil. Bajo esta premisa, el ODS que está estrechamente relacionado es el ODS 2: Hambre cero.

Por otro lado, el despliegue de la solución representa también una aportación importante y este acápite está relacionado al ODS 16: Paz, justicia e instituciones sólidas. A continuación, se explica la relación del presente trabajo con las metas específicas para cada uno de estos objetivos:

A. ODS 2: Hambre cero

- ✓ *Meta 2.1: Para 2030, poner fin al hambre y asegurar el acceso de todas las personas, en particular los pobres y las personas en situaciones vulnerables, incluidos los lactantes, a una alimentación sana, nutritiva y suficiente durante todo el año.*

Este trabajo pretende aportar una mejor visibilidad sobre los sectores más vulnerables del Perú y que a la fecha, presentan problemas severos de malnutrición, tales como la anemia. Si bien el presente trabajo, solo es una versión inicial, estamos convencidos de que en el camino se seguirá sumando esfuerzos colaborativos para tener herramientas que permitan tener mayor claridad de la realidad y enfocar los esfuerzos en donde realmente es necesario. Así también, mediante el apoyo de empresa y grupos económicos privados enfocados a la mejora del entorno donde se desarrollan, se espera fomentar el acceso de todas las personas a una alimentación sana, nutritiva y suficiente.

- ✓ *Meta 2.2. Para 2030, poner fin a todas las formas de malnutrición, incluso logrando, a más tardar en 2025, las metas convenidas internacionalmente sobre el retraso del crecimiento y la emaciación de los niños menores de 5 años, y abordar las necesidades de nutrición de las adolescentes, las mujeres embarazadas, lactantes y las personas de edad.*

La emaciación²⁶ en infantes es una de las formas más severas de malnutrición, sin embargo, las consecuencias que se pueden dar por una forma severa de anemia podrían ser similares, principalmente, en cuanto a las deficiencias en el desarrollo de la capacidad de aprendizaje cognitivo, por ello, toda herramienta que apoye la lucha contra la anemia infantil, también está apoyando el logro de esta meta; además, como se ha podido ver en el análisis realizado, las condiciones nutricionales de la madre durante el embarazo y la lactancia son temas que también se tienen que abordar para la lucha contra la anemia infantil y con ello se enfatiza

²⁶ Forma de malnutrición potencialmente mortal que provoca una delgadez y debilidad extremas en los niños, y aumenta sus posibilidades de morir o de sufrir deficiencias en su crecimiento, su desarrollo y su capacidad de aprendizaje (Consultado en la página de UNICEF: <https://www.unicef.org/es/comunicados-prensa/otros-67-millones-de-ni%C3%B1os-menores-de-5-a%C3%B1os-podr%C3%ADan-sufrir-de-emaciaci%C3%B3n-este>)

más aún el desarrollo de este ítem.

Por otro lado, también hay que tomar en cuenta que actualmente hay una gran inversión del estado en programas sociales para erradicar los problemas de malnutrición (incluyendo la anemia) y una evaluación adicional de las causas reales del problema, permitirán realizar un mejor reenfoque de los planes de acción; de este modo, estar más convencidos de que las acciones que se están ejecutando, serán mejor enfocadas y permitirán dar un mayor aporte a la solución de la problemática.

B. ODS 16: Paz, justicia e instituciones sólidas

- ✓ *Meta 16.6 Crear a todos los niveles instituciones eficaces y transparentes que rindan cuentas.*
A la fecha actual, las instituciones públicas en Perú se rigen por la Ley N° 27806 de Transparencia y Acceso a la Información Pública y por ello, han puesto a disposición de la ciudadanía una gran cantidad de repositorios de datos abiertos. Sin embargo, muchas veces la población no hace uso de éstos debido a que el entendimiento y procesamiento es complejo e implica un alto esfuerzo. La vigilancia y cuestionamiento de la población es indispensable para lograr instituciones eficientes, por ello, con el presente trabajo se quiere aportar el acceso a los datos de una manera más sencilla, y con ello una mayor facilidad en el entendimiento o cuestionamiento de las decisiones tomadas por el estado.
- ✓ *Meta 16.7 Garantizar la adopción en todos los niveles de decisiones inclusivas, participativas y representativas que respondan a las necesidades.*

Solo podemos opinar apropiadamente sobre lo que conocemos y para poder tener certeza razonable de que las decisiones tomadas por el estado son representativas de la voluntad del pueblo, primero se debe conocer la situación sobre la cual se tomó acción. Si como ciudadanía, no se tiene el conocimiento necesario, será muy complicado brindar una retroalimentación. Mediante la presentación de datos más entendibles y visualizaciones ajustadas a las necesidades, se fomenta que la ciudadanía pueda formarse un criterio a partir del conocimiento y se sienta más involucrada en la vida política del país.

Adicional a estos puntos, en menor medida, pero no en menos importancia, el trabajo también tiene aportaciones sobre los siguientes ODS:

C. ODS 3: Salud y Bienestar

- ✓ *Meta 3.2 Para 2030, poner fin a las muertes evitables de recién nacidos y de niños menores de 5 años, logrando que todos los países intenten reducir la mortalidad neonatal al menos hasta 12 por cada 1.000 nacidos vivos, y la mortalidad de niños menores de 5 años al menos hasta 25 por cada 1.000 nacidos vivos.*

Si bien esta meta no menciona directamente con el objetivo del trabajo presentado, se tiene

claro que la mejora en la alimentación y reducción de los niveles de anemia que se espera lograr a mediano plazo, tanto de la madre como del niño, es también una forma de reducir la mortalidad en recién nacidos y niños en edad preescolar.

D. ODS 10: Reducción de las desigualdades

- ✓ *10.3 Garantizar la igualdad de oportunidades y reducir la desigualdad de resultados, incluso eliminando las leyes, políticas y prácticas discriminatorias y promoviendo legislaciones, políticas y medidas adecuadas a ese respecto.*

El tener visibilidad sobre cuáles son los sectores más vulnerables y cuáles son las causas reales del problema, debería reflejarse en planteamientos de políticas y programas que permitan brindar un desarrollo óptimo del infante, enfatizando las poblaciones más vulnerables. Con ello, se podrá mejorar la igualdad de oportunidades, ya que una malnutrición en las etapas tempranas puede generar una brecha significativa para el desarrollo del niño y por ende complicaciones para el futuro adulto.

Anexo II: Dimensiones y niveles MELODA

Ilustración 62. Dimensiones y niveles de evaluación MELODA

Dimensions (maximum 61 points)	Levels
Legal licensing (max. 6 points)	<ol style="list-style-type: none"> 1. Private use 2. Non-commercial reuse 3. Commercial reuse or no restrictions
Access to information (max. 6 points)	<ol style="list-style-type: none"> 1. Web access or unique URL parameters to dataset 2. Web Access unique with parameters to single data 3. API or query language
Technical standard (max. 6 points)	<ol style="list-style-type: none"> 1. Closed standard reusable and open non reusable 2. Open standard reusable 3. Open standard, individual metadata
Standardization (max. 10 points)	<ol style="list-style-type: none"> 1. Own data model standardization 2. Own ad hoc data model standardization published (harmonization) 3. Local standardization 4. Global standardization
Geolocation content (max. 6 points)	<ol style="list-style-type: none"> 1. No geographic information 2. Simple or complex text field 3. Coordinates or full geographical information
Updating frequency of data (max. 15 points)	<ol style="list-style-type: none"> 1. Longer than 1 month 2. Monthly. Updating period ranges from 1 month to 1 day 3. Daily. Updating period ranges from 1 day to 1 hour 4. Hour. Updating period ranges from 1 hour to 1 minute 5. Seconds. Updating period is lower than 1 minute
Dissemination (max. 6 points)	<ol style="list-style-type: none"> 1. Communication / dissemination not systematic 2. Available resources on updates (i.e., RSS feed) 3. Proactive dissemination / push dissemination (information automatic and timely)
Reputation (max. 6 points)	<ol style="list-style-type: none"> 1. No information about the reputation of the data source 2. Statistics or reports published on user's opinions 3. Indicators or rankings on reputation of the data source

Fuente: Meloda 5 A metric to assess open data reusability (Abella, Ortiz-de-Urbina-Criado, De-Pablos-Heredero 2020)

Anexo III: Tablas y campos seleccionados

Ítem	Descripción	Código
Cuestionario	Cuestionario del hogar	CH
Módulo	Características del hogar	64
Tabla	RECH0 - "Datos generales del hogar"	RECH0

Cód	Descripción	Ejemplo
ID1	Año	2019
HHID	identificación Cuestionario del Hogar	000100201
HV000	Código del país	PE6
HV001	Conglomerado	1
HV002	Vivienda	2
HV002A	Hogar	1
HV003	Número de orden del Informante del cuestionario del hogar	2
HV009	Total de personas en el hogar	4
HV010	Número de mujeres elegibles para entrevista individual	1
HV015	Resultado de la entrevista en el hogar	1
HV023	Dominio	1
HV024	Región	1
HV025	Área de residencia	1
HV026	Lugar de residencia	1
HV042	Selección para medición de hemoglobina	1
HV022	Estrato	3
HV005	Factor de ponderación	142155

Ítem	Descripción	Código
Cuestionario	Cuestionario del hogar	CH
Módulo	Características del hogar	64
Tabla	RECH1 - "Características de los miembros del hogar"	RECH1

Cód	Descripción	Ejemplo
ID1	Año	2019
HHID	identificación Cuestionario del Hogar	000100201
HVIDX	Número de Orden en el Hogar	1
HV102	¿Vive habitualmente aquí?	1
HV103	¿Durmió aquí anoche?	0
HV104	Sexo	1
HV105	Edad	43
HV109	Nivel educativo alcanzado	4
HV112	Número de orden de la madre	
HV114	Número de orden del padre	
HV117	Elegibilidad para entrevista individual de mujeres	0
HV120	Menores de 5 años para medición de peso/talla y hemoglobina	0
QH25A	¿Cuál es su nacionalidad?	PERUANA
QH25B	Vive permanentemente en el Perú	

Estudio del índice de anemia en Perú en niños y niñas de 6 a 36 meses, mediante la evaluación de factores demográficos, económicos, formativos y de salud.

Ítem	Descripción	Código
Cuestionario	Cuestionario del hogar	CH
Módulo	Características del hogar	64
Tabla	RECH4 - "Seguros de Salud"	RECH4

Cód	Descripción	Ejemplo
ID1	Año	2019
HHID	identificación Cuestionario del Hogar	000100201
IDXH4	Número de orden del listado del hogar	1
SH11A	Seguro de salud: ESSALUD/IPSS	1
SH11B	Seguro de salud: Militar	0
SH11C	Seguro de salud: Integral	0
SH11D	Seguro de salud: Compañía de seguros	0
SH11E	Seguro de salud: Seguro privado	0
SH11Y	Seguro de salud: No sabe	0
SH11Z	Seguro de salud: No tienen	0

Ítem	Descripción	Código
Cuestionario	Cuestionario del hogar	CH
Módulo	Características de la vivienda	65
Tabla	RECH23 - "Características básicas de la vivienda" // "Nivel socioeconómico"	RECH23

Cód	Descripción	Ejemplo
ID1	Año	2019
HHID	identificación Cuestionario del Hogar	000100201
HV201	¿Cuál es la fuente principal de abastecimiento de agua para tomar o beber?	11
HV202	¿Cuál es la fuente principal de abastecimiento de agua en su hogar?	
HV204	¿Cuánto tiempo se demora en ir, recoger agua y volver?	996
HV205	¿Qué tipo de servicio higiénico tiene su hogar?	11
HV206	¿Tiene electricidad en su hogar?	1
HV219	Sexo del jefe de hogar	1
SHREGION	Región natural	3
SHSEMES	Semestre	1
SHPROVIN	Provincia	1
SHDISTR	Distrito	1
HV270	Índice de riqueza	4
HV271	Factor de puntuación del índice de riqueza (5 decimales)	1

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Datos básicos de MEF (Ministerio de Economía y Finanzas)	66
Tabla	REC0111 - "Características demográficas y sociales"	REC0111

Estudio del índice de anemia en Perú en niños y niñas de 6 a 36 meses, mediante la evaluación de factores demográficos, económicos, formativos y de salud.

Cód	Descripción	Reg1
ID1	Año	2019
HHID	Identificación Cuestionario del Hogar	100201
CASEID	Identificación del caso	000100201 2
V001	Conglomerado	1
V002	Número de vivienda	2
V003	Número de línea de entrevistada	2
V012	Edad actual - entrevistada	33
V023	Dominio de ejemplo - Departamento	1
V024	Región	1
V025	Tipo de lugar de residencia	1
V040	Altitud del conglomerado en metros	2356
V042	Selección de hogar para hemoglobina	1
V102	Tipo de lugar de residencia	1
V106	Nivel educativo más alto	3
V113	Fuente principal de abastecimiento de agua potable	11
V116	Tipo de instalación sanitaria	11
V119	En su hogar tiene: electricidad	1
V136	Número de miembros del hogar	4
AWFACTT	Factor todas las mujeres - total	100
AWFACTU	Factor todas las mujeres - urbano/rural	100
AWFACTR	Factor todas las mujeres - regional	100
AWFACTE	Factor todas las mujeres - educación	100
AWFACTW	Factor todas las mujeres - índice de riqueza	100
V155	Alfabetización	2
UBIGEO	Código de Ubicación Geográfica	010101
V022	Estratos	3
V005	Factor de ponderación	154803
V190	Índice de riqueza	4
V191	Factor de puntuación del índice de riqueza (5 decimales)	1.23445
MUJERES12A49	Mujeres de 12 a 49 años	2

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Historia de Nacimiento - Tabla de Conocimiento de método	67
Tabla	REC21 - "Datos de los hijos - historia de nacimientos"	REC21

Cód	Descripción	Reg1
ID1	Año	2019
CASEID	Identificación del caso	000100201 2
BORD	Número de orden en la historia de nacimiento	2
BIDX	Número de orden de nacimiento	1
B0	El nacimiento fue parto único o múltiple	0
B4	Sexo del niño	1
B5	El niño está vivo	1
B8	Edad actual del niño	0
Q220A	Duración del embarazo	9

Estudio del índice de anemia en Perú en niños y niñas de 6 a 36 meses, mediante la evaluación de factores demográficos, económicos, formativos y de salud.

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Embarazo, Parto, Puerperio y Lactancia	69
Tabla	REC41 - "Embarazo, Parto, Puerperio y Lactancia por cada hijo"	REC41

Cód	Descripción	Ejemplo
ID1	Año	2019
CASEID	Identificación del caso	000100201 2
MIDX	Orden de historia de nacimiento	1
M4	Duración de la lactancia	95
M5	Meses de amamantamiento	8
M14	Visitas prenatales por embarazo	12
M17	Parto por cesárea	0
M18	Tamaño del niño al nacer	2
M19	Peso del niño al nacer (kilos - 3 dec.)	3460
M19A	Peso del niño al nacer según recuerda	1
M34	Cuando empezó a darle el pecho al niño	0
M35	Cuántas veces le dio pecho durante la noche de ayer	7
M36	Cuántas veces le dio pecho durante el día de ayer	6
M42E	Durante su embarazo en alguno de sus controles: Le hicieron examen de sangre	1
M45	Durante el embarazo, le administraron tabletas, jarabe o inyecciones de hierro	1
M46	Por cuántos días tomó hierro y/o cuántas inyecciones recibió	80
M54	Recibió una dosis de vitamina A en los primeros 2 meses después del parto	0
M70	Durante el primer mes le hicieron al bebe algún control o revisión médica	1
M73	Dónde le hicieron el chequeo o revisión médica al bebé por primera vez	31

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Embarazo, Parto, Puerperio y Lactancia	69
Tabla	REC94 - "Embarazo, Parto, Puerperio y Lactancia por cada hijo"	REC94

Cód	Descripción	Reg1
ID1	Año	2019
CASEID	Identificación del caso	000100201 2
IDX94	Orden de historia de nacimiento	1
S425A	Entrevistada era miembro del seguro de salud materna	
S432	El niño está afiliado al Seguro Integral de Salud	0
S436C	Pusieron al niño en contacto piel a piel con Ud. inmediatamente después del nacimiento	0
QI422A_A	Durante el embarazo algún personal de salud le realizó una prueba o análisis para descartar anemia	1
QI422A_B	Durante el embarazo le diagnosticaron o le dijeron que tenía anemia	2
QI422A_C	Durante el embarazo le indicaron tratamiento con hierro	
QI422A_D	Durante el embarazo consumió hierro tal como le indicó el personal de salud	
QI440B	Durante primeros 6 meses, solo recibió leche materna	1

Estudio del índice de anemia en Perú en niños y niñas de 6 a 36 meses, mediante la evaluación de factores demográficos, económicos, formativos y de salud.

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Inmunización y Salud	70
Tabla	REC43 - "Vacunación y control"	REC43

Cód	Descripción	Reg1
ID1	Año	2019
CASEID	Identificación del caso	000100201 2
HIDX	Nro. orden de historia de nacimiento	1
H11	En los últimos 14 días, ha tenido diarrea la niña(o)	2
H22	Ha tenido fiebre en las últimas dos semanas	1
H31	Ha tenido tos en las últimas dos semanas	0
H34	Vitamina A en los últimos 6 meses	1
H41B	Alguna vez recibió la dosis de vitamina A	1
H42	Tomando píldoras de hierro, jarabe	

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Inmunización y Salud	70
Tabla	REC95 - "Vacunación y CRED"	REC95

Cód	Descripción	Reg1
ID1	Año	2019
CASEID	Identificación del caso	000100201 2
IDX95	Índice de historia de nacimiento	1
S45NM1	Recibió vacuna Neumococo 1ra. dosis	1
S45NM2	Recibió vacuna Neumococo 2da. dosis	1
S45NM3	Recibió vacuna Neumococo 3ra. dosis	0
S45RT1	Recibió vacuna Rotavirus 1ra. dosis	1
S45RT2	Recibió vacuna rotavirus 2da. dosis	1
S466	Le hicieron algún control de crecimiento y desarrollo	1
S466C	Cantidad de controles de crecimiento de control	10

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Peso y talla - Anemia	74
Tabla	RECH6 - "Datos de peso y talla - anemia" (Niño)	RECH6

Cód	Descripción	Ejemplo
ID1	Año	2019
HHID	identificación Cuestionario del Hogar	000100201
HC0	Número de orden en el hogar	4
HC1	Edad en meses	8
HC2	Peso en kilogramos (1 dec.)	107
HC3	Altura en centímetros (1 dec.)	727
HC27	Sexo	1

Estudio del índice de anemia en Perú en niños y niñas de 6 a 36 meses, mediante la evaluación de factores demográficos, económicos, formativos y de salud.

Cód	Descripción	Ejemplo
HC51	Número de orden del padre o responsable	2
HC53	Nivel de hemoglobina (g/dl-1 decimal)	132
HC55	Resultado de medir (hemoglobina)	0
HC56	Nivel de hemoglobina ajustado por altitud (g/dl-1 decimal)	121
HC57	Nivel de anemia	4
HC60	Número de orden de la madre en el hogar	2
HC61	El nivel educativo más alto de la madre	3
HC62	El año más alto de educación de la madre	5
HC63	Intervalo de los nacimientos anteriores al niño	141
HC64	Número de orden de nacimiento	2
HC68	Nivel educativo más alto (CS para informe preliminar y final)	3
HV005A	Factor de ponderación (solo para 2020 Y 2008) - En el 2008 está en RECHO, pero se usa para cálculos regionales	

Ítem	Descripción	Código
Cuestionario	Cuestionario Individual	CI
Módulo	Programas Sociales	569
Tabla	PS_WAWAWASI - Cuna más	PS_WAWAWASI

Cód	Descripción	Reg1
ID1	Año	2019
HHID	identificación Cuestionario del Hogar	001200401
HVIDX	Número de Orden en el Hogar	4
QHCLUSTER	Conglomerado	12
QHNUMBER	Vivienda seleccionada	4
QHHOME	Hogar seleccionado	1
PS107_1A	Hace cuanto tiempo es beneficiario: Años	0
PS107_1M	Hace cuanto tiempo es beneficiario: Meses	3
PS107_1S	Alguien más es beneficiario	2
PS6	Número de Orden	1

Anexo IV: Documentación sobre el factor de ponderación

A continuación, se presentan las tres casuísticas identificadas:

Factor de Niño (HV005X): En el año 2015 el factor de ponderación utilizado fue distinto. Se tuvo una subdivisión para los factores hogar, mujer y niño, tal como se muestra a continuación:

Ilustración 63. Factor de ponderación año 2015

Factor Hogar (HV005)

Es el factor básico de muestreo que ha sido ajustado a la no respuesta de viviendas y se utiliza para el cálculo de los indicadores que están relacionados al hogar, por ejemplo: Saneamiento básico de hogares, agua tratada, entre otros.

En la base de datos de la ENDES-2015 se encuentra en el módulo RECH0.

Factor Mujer (V005)

Este factor permite recomponer la estructura poblacional de las mujeres en edad fértil y está ajustado a la no respuesta, se utiliza para el cálculo de indicadores relacionados con esta población, por ejemplo: Planificación familiar de mujeres en edad fértil, parto institucional, demanda insatisfecha de planificación familiar, control prenatal, entre otros.

En la base de datos ENDES-2015 se encuentra en los módulos REC0111.

Factor Niño (HV005X)

Para el año 2015, se determinó la construcción de un factor de ponderación para niñas y niños menores de 5 años, que permite recomponer la estructura de esta población.

Para ello, se utilizó como información auxiliar el registro de viviendas y está ajustado a la tasa de no respuesta de las mediciones antropométricas, lo cual permitirá que todas las niñas y niños menores de 5 años estén representados en las estimaciones de los indicadores relacionados con esta población, por ejemplo: desnutrición, anemia, control de crecimiento y desarrollo, bajo peso al nacer, inmunizaciones, consumo de suplemento de hierro, entre otros.

En la base de datos ENDES-2015 se encuentra en los módulos RECH0 y REC0111.

También se considera los factores para el módulo de salud:

Fuente: Ficha técnica ENDES - Año 2015

Factor mediciones Niño (HV005A): En el año 2020 el factor de ponderación utilizado tuvo una casuística similar a la del 2015, esto debido a que se tuvo una subdivisión para los factores hogar y mediciones niño, tal como se muestra a continuación:

Ilustración 64. Factor de ponderación año 2020

➤ **Factor Hogar (HV005)**

Es el factor básico de muestreo que ha sido ajustado a la no respuesta de viviendas y se utiliza para el cálculo de los indicadores que están relacionados al hogar

Para la obtención de los indicadores con los softwares estadísticos actuales para determinar el plan del diseño muestral (muestras complejas), se debe considerar las siguientes variables:

HV001: conglomerado, HV022: estrato, peso=HV005/1000000

El factor hogar (HV005) se encuentra en el módulo RECH0 de la base de datos.

➤ **Factor mediciones Niño (HV005A)**

Es el factor básico de muestreo que ha sido ajustado a la no respuesta de viviendas y se utiliza para el cálculo de los indicadores de **desnutrición y anemia en menores de 5 años**. Para la obtención de los indicadores con los softwares estadísticos actuales para determinar el plan del diseño muestral (muestras complejas), se debe considerar las siguientes variables: HV001: conglomerado, HV022: estrato, peso=HV005A/1000000.

El factor mediciones niño (HV005A) se encuentra en el módulo RECH6 de la base de datos.

Fuente: Ficha técnica ENDES - Año 2020

Factor Hogar (HV005): Para todos los demás años, no existió una caracterización especial y se utiliza el factor hogar para los cálculos. En estos años solo se tiene una distinción para el factor mujer. Se obtuvo la captura del informe 2019 (INEI 2020).

Ilustración 65. Factor de ponderación resto de años

➤ **Factor Hogar (HV005)**

Es el factor básico de muestreo que ha sido ajustado a la no respuesta de viviendas y se utiliza para el cálculo de los indicadores que están relacionados al hogar, desnutrición y anemia en menores de 5 años.

Para la obtención de los indicadores con los softwares estadísticos actuales para determinar el plan del diseño muestral (muestras complejas), se debe considerar las siguientes variables:

HV001: conglomerado, HV022: estrato, peso=HV005/1000000

El factor hogar (HV005) se encuentra en el módulo RECH0 de la base de datos.

➤ **Factor Mujer (V005)**

Este factor permite recomponer la estructura poblacional de las mujeres en edad fértil y está ajustado a la no respuesta, se utiliza para el cálculo de indicadores relacionados con esta población, por ejemplo: Planificación familiar de mujeres, parto institucional, demanda insatisfecha de planificación familiar, control prenatal, entre otros.

Para la obtención de los indicadores con los softwares estadísticos actuales para determinar el plan del diseño muestral (muestras complejas), se debe considerar las siguientes variables:

V001: conglomerado, V022: estrato, peso=V005/1000000

El factor mujer (V005) se encuentra en el módulo REC0111 de la base de datos.

Fuente: Ficha técnica ENDES - Año 2019

Anexo V: Normatividad de compromisos en gobiernos regionales

Ilustración 66. Documentos normativos de compromisos asumidos por gobiernos regionales

Región	Actas de consenso de metas anualizadas de anemia	Documento normativo emitido	Pacto regional para reducir anemia	Documento normativo
Amazonas	X		X	
Áncash	X		X	
Apurímac	X	X		OR N° 005-2017-GR-APURIMAC/CR
Arequipa	X	X	X	OR N° 377-AREQUIPA
Ayacucho	X	X	X	OR N° 006-2017-GRA/CR
Cajamarca	X		X	
Callao	X	X	X	OR N° 0008
Cusco	X	X	X	RD N°0418-2017
Huancavelica	X	X	X	OR N° 370-GOB:REG_HVCA/CR
Huánuco	X		X	
Ica	X	X		OR N° 016-2017-GORE-ICA
Junín	X		X	
La Libertad	X	X	X	OR 013-2017 GRLL/CR
Lambayeque	X	X	X	OR N° 019-2017-GR.LAMB/CR
Loreto	X	X	X	OR N° 012-2017-GLR-CR
Madre de Dios	X		X	
Moquegua	X	X	X	OR N° 01-2017-CR/GRM
Pasco	X	X	X	OR N° 409-2017-G.R.P./CR
Piura	X	X	X	OR N° 394-2017/GRP-CR
Puno	X	X	X	OR N° 05-2017-GR PUNO-CRP
San Martín	X	X		OR N° 017-2017-GRSM/CR
Tacna	X	X	X	OR N° 003-2017-CR/GOB.REG.TACNA
Tumbes	X	X	X	OR N° 011-2017-GOB.REG.TUMBES-CR-CD
Ucayali	X	X	X	RER N° 0282-2017-GRU-GR
Lima Provincias	X	X		OR N° 014-2017-CR-GRL

Fuente: Plan Multisectorial de Lucha contra la Anemia