



ANÁLISIS DE COBERTURA HOSPITALARIA EN LATINOAMÉRICA

MASTER EN BUSINESS INTELLIGENCE Y BIG DATA
2020

María Fernanda Calderón
Carolina González Martínez
Hugo González Riera
Alejandro Sánchez

TABLA DE CONTENIDO

1	DEFINICIÓN DEL PROBLEMA	5
2	VALIDACIÓN DEL PROBLEMA	11
2.1	Definición de Hipótesis	11
2.2	Proceso de validación de Hipótesis	11
2.3	Análisis y conclusiones de las entrevistas	11
3	ANÁLISIS PRELIMINAR DE DATOS	13
4	ANÁLISIS DIAGNOSTICO	16
4.1	Análisis de Entorno	16
4.2	Análisis competitivo	16
4.3	Análisis DAFO	17
4.3.1	Debilidades	17
4.3.2	Amenazas	18
4.3.3	Fortalezas	18
4.3.4	Oportunidades	18
5	PLAN ESTRATEGICO	20
5.1	El modelo de Negocio	20
5.1.1	Propuesta de Valor	21
5.1.2	Socios Clave	21
5.1.3	Recursos Claves	22
5.1.4	Relación con Clientes	22
5.1.5	Segmento de Clientes	23
5.1.6	Canales de Distribución	23
5.1.7	Actividades Claves	23
5.1.8	Estructura de Costes	24
5.1.9	Fuente de Ingresos	26
6	PLAN DE ACCIÓN	27
6.1	Definición Alcance y Objetivo Del Proyecto	27
6.2	Definición de Actividades	28
6.2.1	Análisis	28
6.2.2	Diseño	30
6.2.3	Implementación	31
6.2.4	Validación	38
6.2.5	Retroalimentación	39
6.3	Definición de arquitectura técnica	39
6.4	Análisis de recursos	40
6.4.1	Talento Humano	40

6.4.2	Recursos físicos.....	41
6.4.3	Recursos financieros	41
6.5	Gestión del tiempo: Cronograma	42
7	RENTABILIDAD DEL PROYECTO	43
8	BIBLIOGRAFIA.....	46
9	ANEXOS	46

TABLA DE FIGURA

Figura 1 . Tendencias mundiales y regionales de los pagos catastróficos: indicador ODS 3.8.2. Fuente Seguimiento de la cobertura sanitaria universal: Informe de monitoreo global 2017	5
Figura 2 Porcentaje de Población con Instalaciones Mejoradas América. Fuente: Indicadores Básicas Salud de las Américas 2017.....	6
Figura 3 Estadísticas de Aprobaciones de los últimos cinco años, modalidad de Garantía Soberana. Fuente: https://www.iadb.org/es/sectores/salud/perspectiva-general	7
Figura 4 Estadísticas de Aprobaciones de los últimos cinco años, modalidad de Sin Garantía Soberana. Fuente: https://www.iadb.org/es/sectores/salud/perspectiva-general	7
Figura 5 Premisas del Modelo Analíticos Servicio Hospitalario	9
Figura 6 Tipos de Fuentes de Datos	13
Figura 7 Análisis DAFO	17
Figura 8 Modelo de Negocio Canvas	20
Figura 9 Recursos Claves Técnicos	22
Figura 10 Recursos Claves Aprovisionamiento	22
Figura 11 Actividades Claves	23
Figura 12 Fases del Proyecto	28
Figura 13 Mapa de Municipios de Colombia	29
Figura 14 Definición Arquitectura Técnica	30
Figura 15 Fragmento de Código Obtención de Datos Espaciales	31
Figura 16 Fragmento de Código Obtención de Datos Espaciales	32
Figura 17 Fragmento de Código de Obtención de Datos Censo.....	32
Figura 18 Fragmento Código obtención Datos Hospitalares	33
Figura 19 Fragmento Código Cálculo Métricas Municipio.....	34
Figura 20 Fragmento Código Cálculo Métricas Hospitalarias.....	34
Figura 21 Fragmento de Código Modelos Predictivos	35
Figura 22 Algoritmos de Machine Learning	36
Figura 23 Matriz de Confusión	36
Figura 24 Ajuste Modelo con Diferentes Escenarios.....	37
Figura 25 Fragmento Código Extracción Resultados Tableau.....	37
Figura 26 Ejemplificación Cuadros de Mando	38
Figura 27 Arquitectura Técnica	39
Figura 28 Organigrama	40
Figura 29 Tipos de Beneficios	44

LISTADO DE TABLAS

Tabla 1 Costos Tecnológicos	25
Tabla 2 Costos Administrativos Lugar de Trabajo	25
Tabla 3 Asignación porcentual de Costos.....	25
Tabla 4 Costos Administrativos Salario Equipo de Trabajo	26
Tabla 5 Estructura de Costos.....	26
Tabla 6 Fichero de Datos	35
Tabla 7 Cronograma	42

1 DEFINICIÓN DEL PROBLEMA

Los altos costos para el acceso a los servicios hospitalarios a nivel mundial hacen que miles de personas no cuentan con acceso a la salud, tras dicha premisa, el 12 de diciembre del 2012, la Asamblea General de las Naciones Unidas aprobó una resolución en la cual se busca que los países comiencen a asumir compromisos e inversiones en sus sistemas de salud.

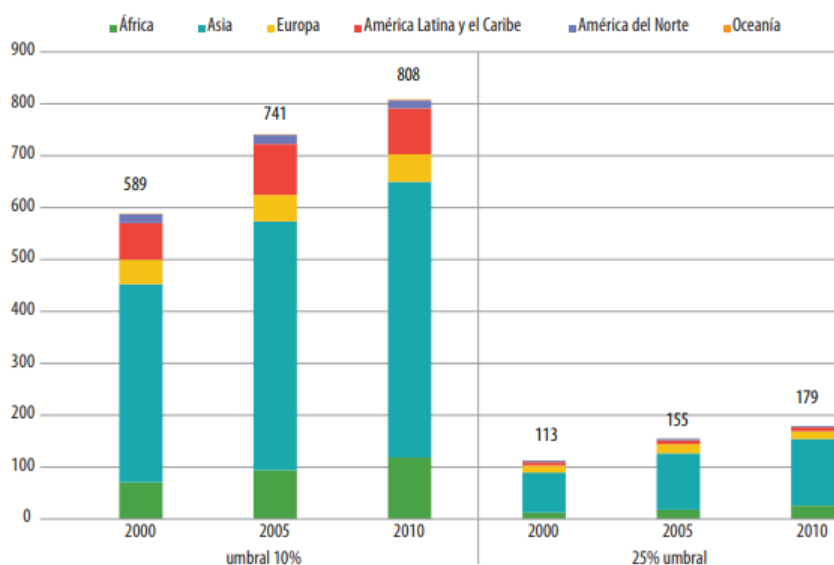


Figura 1 . Tendencias mundiales y regionales de los pagos catastróficos: indicador ODS 3.8.2. Fuente Seguimiento de la cobertura sanitaria universal: Informe de monitoreo global 2017

El secretario General de la Organización de las Naciones Unidas ONU, Antonio Guterres, expuso en esa misma fecha en su cuenta de Twitter.

"Es inaceptable que millones de personas no tengan acceso a una atención médica por los altos costos. Hago un llamado a los líderes mundiales para garantizar que la salud sea una realidad para todos en el Día Internacional de la Cobertura Sanitaria Universal".

Sin embargo, hoy en día Latinoamérica, una región compuesta por 20 países del continente americano, con una población general de 630 millones de habitantes, con índices pobreza de un 30,1%, sigue siendo una de las regiones con mayor deficiencia en los servicios de cobertura hospitalaria a nivel mundial, ya que aproximadamente el 30% de la población no tiene acceso a los servicios de salud pública gratuita, según las estadísticas de la Organización Panamericana de la Salud OPS.

La deficiencia de los servicios de cobertura hospitalaria, se deben en gran medida a las condiciones geográfica, la poca inversión en materia de vacunación y prevención de enfermedades, a esto se le suman los impedimentos demográficos para la movilización a los centros de salud en zonas rurales o de difícil acceso, así como la ausencia en sí de centros médicos adecuados en dichas zonas.

"No es suficiente tener hospitales y centros de salud, estas instituciones deben tener la combinación correcta de recursos humanos, infraestructura y equipamiento, medicinas y otras tecnologías sanitarias, para evitar largos tiempos de espera y ofrecer atención de calidad", Carrisa Etienne directora de la OPS, 12 de diciembre 2019.

La deficiencia de los servicios de cobertura hospitalaria en Latinoamérica varía entre los países, existen países con importantes inversiones en materia de salud como lo demuestran: Cuba, con una

inversión del 28% de sus ingresos, Puerto rico, Panamá y Uruguay. Sin embargo, esto solo evidencia aún más los fuertes desequilibrios de servicio entre distintas zonas de la propia Latinoamérica empezando por los países del Caribe y terminando por el cono sur o los países de centro América.

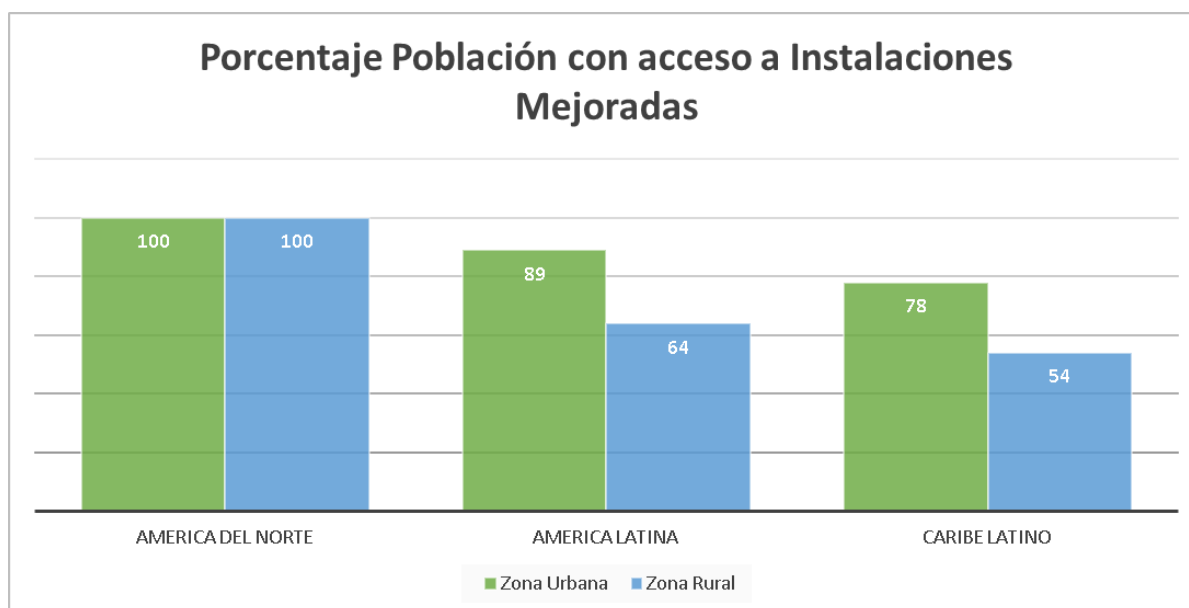


Figura 2 Porcentaje de Población con Instalaciones Mejoradas América. Fuente: Indicadores Básicas Salud de las Américas 2017

Dado lo anteriormente descrito, diferentes organismos financieros han establecido una imperiosa necesidad de establecer programas de ayuda inversión para la generación de nuevos hospitales y servicios hospitalarios en aquellas zonas más necesitadas. Algunos de estos organismos son el Banco Mundial, la Organización Internacional de la Salud (dependiendo de la ONU) o el Banco Interamericano de Desarrollo.

El Banco Mundial, por su parte ha expresado en su página web su compromiso de ayuda a mejorar la cobertura sanitaria.

“El Grupo Banco Mundial (GBM) ha asumido el compromiso de ayudar a los Gobiernos a lograr la cobertura sanitaria universal (CSU) antes de 2030, un avance que tiene el potencial de transformar la salud y el bienestar de los individuos y las sociedades. La cobertura sanitaria universal significa que todas las personas pueden obtener los servicios médicos de calidad que necesitan sin tener que pasar por dificultades económicas. Además, permite a los países aprovechar al máximo el capital humano, su principal activo. La salud es una inversión fundamental en capital humano y en crecimiento económico; sin una buena salud, los niños no pueden asistir a la escuela y los adultos no pueden ir a trabajar....” Banco Mundial, 10 de abril 2018,
<https://www.bancomundial.org/es/topic/health/overview#1>

Un ejemplo de su compromiso ha sido el financiamiento del Proyecto de prevención de Enfermedades no Transmisibles en Uruguay, con el cual se rediseño el modelo de atención médico de dichas enfermedades, con el cual se logró la disminución de la mortalidad vinculada con enfermedades del sistema circulatorio entre la población mayor de 70 años disminuyó de 75,18 % a 60,3 % entre 2006 y 2014.

Otro ejemplo de las acciones que están tomando estos organismos mencionados los podemos ver en las cifras del Banco Interamericano de Desarrollo BID, el cual durante los últimos cinco años ha realizado préstamos a los diferentes países de Latinoamérica, en dos modalidades:

- **Préstamos de Garantía Soberana:** financiamiento a los gobiernos y a las instituciones controladas por el Gobierno para apoyar el desarrollo y proyectos sociales.



Figura 3 Estadísticas de Aprobaciones de los últimos cinco años, modalidad de Garantía Soberana. Fuente: <https://www.iadb.org/es/sectores/salud/perspectiva-general>

- **Préstamos sin Garantía Soberana:** financiamiento para proyectos del sector privado que contribuyan al desarrollo.

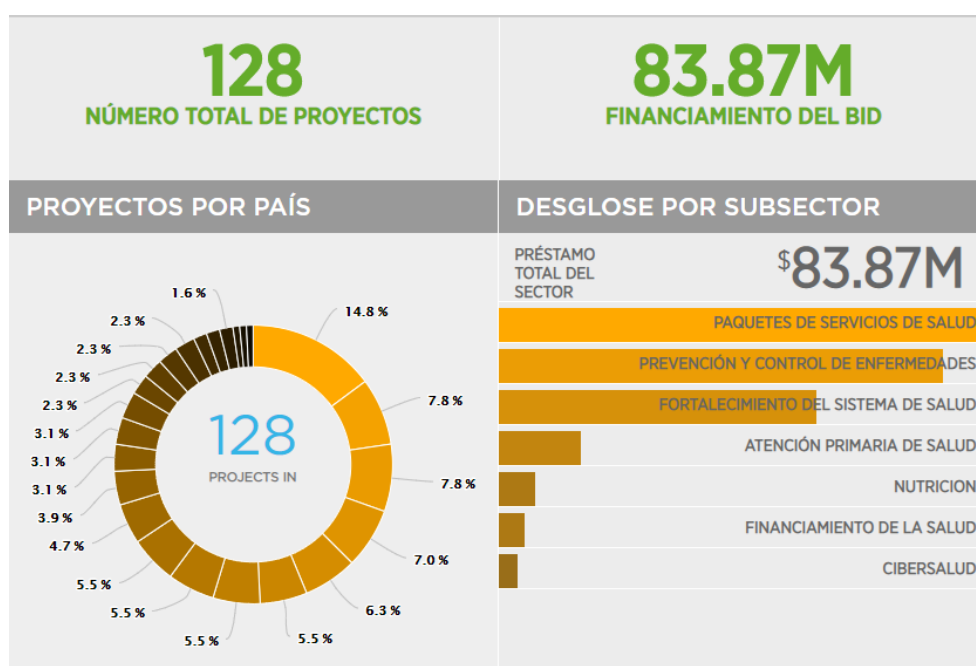


Figura 4 Estadísticas de Aprobaciones de los últimos cinco años, modalidad de Sin Garantía Soberana. Fuente: <https://www.iadb.org/es/sectores/salud/perspectiva-general>

Como observamos en los ejemplos anteriores, cada uno de los distintos organismos públicos o privados son los encargados de definir las políticas y planes de inversión que desean implementar en materia de salud.

Sin embargo, se es consciente de que algunos gobiernos pueden realizar grandes inversiones sanitarias sin que éstas generen la mejora esperadas, ya sea por falta de investigación, mala implementación de los recursos, error en el análisis de su situaciones y población objetivo o cuestiones políticas, derivadas de campañas electorales o favores políticos.



Con el fin de evitar que estas inversiones no obtengan los resultados esperados y teniendo en cuenta los avances tecnológicos en áreas de investigación como Big Data y la importancia del análisis de datos para la toma de decisiones, proponemos la implementación de métodos científicos para el análisis de métricas sanitarias, con el fin de realizar proyección y medición de impactos futuros de cada uno de los proyectos y los objetivos de estos.

Los impactos para evaluar dependerán las situaciones propias de cada país, para ello se deberán definir métricas, entre los que identificamos las siguientes:

- Mejora atención hospitalaria en región específica
- Acceso a los servicios hospitalarios
- Tiempos de Espera en los hospitales
- Tasas de Mortalidad infantil
- Tasas de Mortalidad
- Tasas de Morbilidad
- Tasas de cobertura de especialidades médicas
- Tasas de incidencia de enfermedades, tales como paludismo, Zika, gripe, malaria, etc.

Una vez definidas la métrica, se debe identificar la población objetivo y sus particularidades tanto demográficas, geográficas y comportamiento social y económico, solo con estas podremos realizar una verdadera proyección del impacto.

Para establecer este modelo científico, será necesario generar un modelo de análisis del servicio hospitalario, en la cual se obtenga la información de datos demográficos, geográficos, económicos y sanitarios que se encuentren en la diferentes fuentes públicas y privadas, las cuales servirán como base de análisis y proyección.

Para ello también es importante identificar ciertas premisas que se deben tener en cuenta a la hora de la extracción, transformación y análisis de los datos, entre las principales identificamos las siguientes:

- **Datos actuales**, con los cuales se pueden extrapolar las necesidades a futuro a cubrir cuando los hospitales con servicios hospitalarios estén disponibles
- **Datos fiables**, fuentes públicas y privadas certificadas
- **Implementación de métricas robustas**, relacionadas con el impacto que se quiere conseguir
- **Fácilmente replicable** entre los diferentes países y regiones, evaluando métricas generales de servicios sanitarios

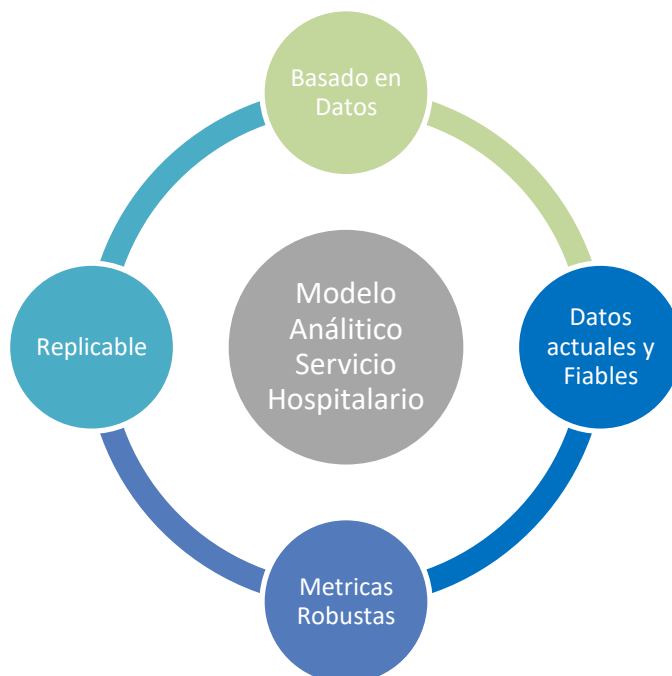
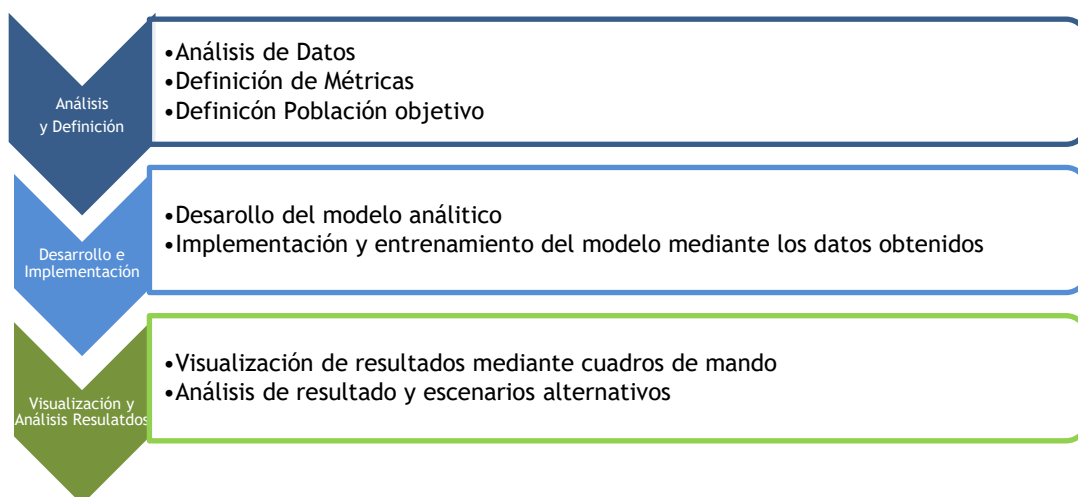


Figura 5 Premisas del Modelo Analítico Servicio Hospitalario

Mediante el uso de métodos científicos, no solo se podrán medir los impactos de una propuesta específica, si no que permitirán evaluar el impacto de alternativas de proyectos y comparar el impacto entre los mismos con el fin de realizar la mejor inversión.

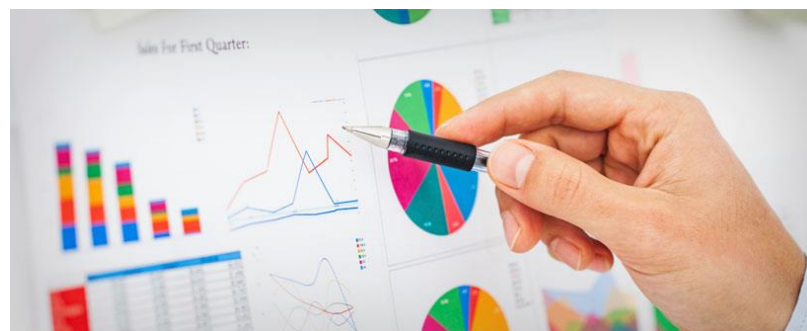
Basado en la anterior, nace como iniciativa del área de I+D de una Entidad Financiera Multilateral de Desarrollo (en adelante Entidad Financiera), generar un modelo para el análisis de la cobertura hospitalaria en Latinoamérica, cuyo objetivo es brindar una evaluación y estudio del impacto que generaría la colocación de un determinado servicio hospitalario sobre una zona o población determinada, basándose en un diagnóstico previo de métricas especialmente definidas sobre la población o región a cubrir.

El desarrollo de este modelo de análisis se compondrá por tres fases conceptuales:



A continuación, se realiza una descripción de cada una de las fases ilustradas en la figura anterior:

1. **Análisis y Definición**, será la primera fase conceptual se compone por las siguientes tres sub-fases:
 - **Análisis de los datos obtenidos de las distintas fuentes de datos:** dependiendo del grado de gran realidad que se tenga los datos se podrá llegar a un alcance mayor o menor tanto del diagnóstico como de impacto a medir. Por ejemplo, si se tuviese a nivel de no solo hospitales si no de servicios médicos dentro de esos hospitales (especialidades, número de médicos, enfermeros, camas disponibles, número y tipo de intervenciones, etc.) se podría mejorar las métricas a usar en la medición y diagnóstico.
 - **Definición de Métricas:** estarán fundamentalmente basadas en información externa pública, aunque en fases posteriores se puedan complementar y mejorar con información privada disponible por los gobiernos y entidades. Parte de la información está referida a temas sanitarios como las tasas de mortalidad o las incidencias determinadas enfermedades y otras serán informaciones menos específicas y también disponibles de forma pública, los censos poblacionales o las ubicaciones de los hospitales y servicios hospitalarios.
 - **Definición de la población objetivo.**
2. **Desarrollo e Implementación**, será la segunda fase conceptual la cual estará compuesta por dos sub-fases:
 - **Desarrollo del modelo de análisis** mediante la implementación de algoritmos de Machine Learning
 - **Implementación del modelo y entrenamiento** de este a partir de diferentes escenarios y métricas.
3. **La Visualización y Análisis de Resultados**, será la tercera fase conceptual, la cual consistirá en dos sub-fase:
 - **La creación de un cuadro de mando** donde se visualizarán los diferentes resultados obtenidos del proceso de modelado
 - **Análisis de Resultado**, de los impactos y posibles escenarios



2 VALIDACIÓN DEL PROBLEMA

Hemos identificado que existe una deficiencia en los servicios de cobertura hospitalario de Latinoamérica, para continuar con el proceso de análisis de la situación se realizará un proceso de validación de hipótesis, análisis preliminar de datos, entorno y competencia.

2.1 Definición de Hipótesis

Para la validación del problema identificado se planteará la siguiente Hipótesis:

El uso de un modelo de Análisis de cobertura hospitalaria para Latinoamérica para identificar donde es adecuado construir un centro hospitalario, será de gran interés y aporte para los gobiernos y a sus servicios de salud a la hora de definir los proyectos de inversión.

2.2 PROCESO DE VALIDACIÓN DE HIPÓTESIS



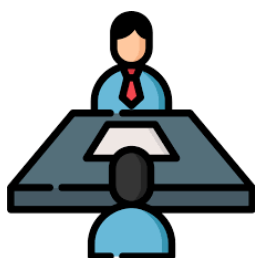
Inicialmente, el proyecto de crear un modelo de análisis de cobertura hospitalaria en Latinoamérica surge de una necesidad planteada por una Entidad Financiera con lo que entendemos que es de gran interés para una institución de esta naturaleza. Aun así, se realizó un proceso de validación de la hipótesis planteada a partir de un proceso de entrevista a personas especialista en s temas de inversión y de salud en Latinoamérica.

Para ello se realizaron dos entrevistas a dos tipos de perfiles de experto:

- Directivo del departamento de tecnología e investigación de la Entidad Financiera
- CEO de una entidad de seguros en Latino América

Ambos entrevistados han solicitado aparecer de forma anónima. Adjuntamos el detalle de las entrevistas en el [Anexo 1: Validación de Hipótesis - Entrevistas](#)

2.3 ANÁLISIS Y CONCLUSIONES DE LAS ENTREVISTAS



Una vez realizadas la entrevista, previamente mencionadas en el apartado anterior, se procede al análisis y entendimiento de cada una de las entrevistas, para posteriormente obtener una conclusión que nos llevara a la validación inicial de la hipótesis planteada.

A continuación, se encontrará el resumen realizado a partir del análisis y entendimiento de cada una de las entrevistas:

- **Entrevista1: Directivo del departamento de tecnología e investigación de la Entidad Financiera**

El directivo del departamento de tecnología, en adelante “el directivo”, confirmó la importancia y usabilidad de un modelo de este tipo. También, confirmó que actualmente se encuentran en transcurso diferentes tipos de proyectos de esta índole, y que el problema de la salud es uno de los principales temas a tratar y prioritario para este tipo de instituciones que guían las inversiones de la región.

Por otro lado, resalto que un modelo de este tipo requiere de mucha experiencia en diferentes especialidades, no solo a nivel técnico de gestión de datos, sino que requiere mucho conocimiento específico del negocio, por ejemplo, conocimiento médico, del sector salud, de hospitales, de transporte, de infraestructura, y una lista inmensa, dependiente de la exactitud con la que queramos definir este modelo y las variables que queramos incluirle. Teniendo en cuenta que en esta instancia la propuesta es académica, nos recomendó que tuviéramos especial cuidado en como acotaríamos la solución.

- **Entrevista 2: CEO de una entidad de seguros en Latino América**

La entrevista con el CEO de la entidad de seguros (en adelante “el CEO”) permitió tener una visión general de la hipótesis, dado que confirmó el interés en un modelo de análisis como este e incito a comenzar un análisis para crear un modelo fácilmente adaptable, teniendo en cuenta la disponibilidad de diferentes variables, y modificando el modelo para que también sea aplicable a otro tipo de negocio. Es decir, extender y potenciar la idea.

Adicional, recomendó la importancia de entender la heterogeneidad de los datos que se irán obteniendo, así como la importancia de la necesidad de “expertise” en datos y en materia de sanidad para la evaluación del modelo.

Finalmente, luego de realizar un completo análisis y entendimiento de las dos entrevistas anteriormente resumidas, se ha llegado a las siguientes conclusiones:

- Confirmación de que el proyecto es considerado de suma utilidad e interés teniendo en cuenta principalmente el impacto social, económico y hasta político que representaría sobre cada una de las inversiones que se realicen en los países de Latinoamérica o en cualquier país en vía de desarrollo.
- Es de vital importancia para la correcta definición, ejecución e implantación de una solución como la propuesta, será necesario contar diferentes áreas de experiencia tanto a nivel tecnológico, de datos, social y en materia de salud.
- Para realizar un modelo robusto y eficiente, será de vital importancia la inclusión de una gran cantidad de métricas no solo sanitarias si no también sociales y económicas que permitan ver en un plano general todos los procesos que se llevan a cabo para poder realizar la construcción de un centro hospitalario.
- Para un primer ejercicio y desarrollo de un **MVP**, Minimum Viable Product por sus siglas en inglés, se debe realizar una primera definición con un alcance acotado pero que permita una escalabilidad en la complejidad.
- El modelo de análisis de cobertura hospitalaria en Latinoamérica puede ser fácilmente aplicable a otras regiones o ámbitos a parte de los sistemas de salud.

3 ANÁLISIS PRELIMINAR DE DATOS



Para realizar el análisis preliminar de los datos, se realizó una búsqueda intensificada de las diferentes fuentes de datos existentes en los últimos 5 años para las diferentes regiones de Latinoamérica.

Para lograr un mejor análisis preliminar de los datos y acotamiento de nuestro alcance, nuestra búsqueda se centró en conseguir información solo de Colombia, siendo este el país en el que se basara nuestro MVP. La búsqueda de datos se centró en tres tipos de fuentes como lo indica la siguiente figura:



Figura 6 Tipos de Fuentes de Datos

A continuación, describimos algunas de las fuentes de datos abiertas identificadas y analizadas para el desarrollo del modelo, basándonos en tres tipos de fuentes de datos:

DATOS SANITARIOS:

- Registro Especial de Prestadores de Servicios de Salud, por clase de prestador IPS- Instituciones Prestadoras de Servicios de Salud.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Registro-Especial-de-Prestadores-de-Servicios-de-S/c36g-9fc2>
- Listado de EPS, IPS y Centros de Salud del Municipio de Yopal.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Eps-Ips-y-Centros-de-salud/cmb8-yyw6>
- Información de contacto de los actores del sistema general de seguridad social en salud, del municipio de Bucaramanga, referente a :
 - **EPS:** Entidad promotora de salud. Son las encargadas de hacer la afiliación, el registro de los afiliados al Sistema General de Seguridad Social en Salud y el recaudo de los aportes que trabajadores y empleadores deben hacer por ley para acceder al servicio.
 - **IPS:** Instituto prestador de salud. Una IPS es contratada por las entidades promotoras de salud (EPS) para que cumpla con los planes y servicios que estas ofrecen (promueven) a sus usuarios.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/IPS-y-EPS-s-Municipio-de-Bucaramanga/kjnk-8pnX>

- Conjunto de Datos de Hospitales y Centros de Salud del Municipio de Palmira, Valle del Cauca
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Hospitales-Y-Centros-De-Salud-Del-Municipio-De-Pa/46em-xah3>
- IPS y EPS Municipio de Sogamoso
<https://www.datos.gov.co/Salud-y-Proteccion-Social/IPS-y-EPS-Municipio-de-Sogamoso/air2-szdy>
- Número de afiliados por departamento, municipio, régimen y administradora para el año 2017
<https://www.datos.gov.co/Salud-y-Proteccion-Social/N-mero-de-afiliados-por-departamento-municipio-r-g/ptvs-rhi5>
- Información básica de las Instituciones Prestadoras de Servicios de Salud de la red pública, que se encuentran habilitados ante el REPS - Registro Especial de Prestadores de Servicios de Salud del departamento de Boyacá.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Red-p-blica-de-Prestadores-de-Servicios-de-Salud-D/8byh-6agx>
- IPS habilitadas en el municipio de Sincelajo a 2017.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/INSTITUCIONES-PRESTADORAS-DE-SERVICIOS-DE-SALUD-HA/3yz8-kdzf>
- Calidad de la atención en salud que reciben los usuarios del Sistema General de Seguridad Social en Salud (SGSSS).
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Calidad-en-Salud-EPS/riak-tm9d>
- Relación de IPS públicas y privadas según el nivel de atención y capacidad instalada Fuente: Base de datos Registros Especial de Prestadores de Servicios de Salud - REPS. Junio 15 2017.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Relacion-de-IPS-p-blicas-y-privadas-seg-n-el-nivel/s2ru-bqt6>
- Estimación de los Indicadores para el seguimiento del Talento Humano en Salud THS
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Saludatos-Estimacion-de-los-Indicadores-para-el-se/5atd-7cac>
- Información del ministerio de Salud Colombianos
<https://www.minsalud.gov.co/Paginas/datos-abiertos.aspx>

DATOS DEMOGRÁFICOS:

- Tasa de natalidad y mortalidad - Causas de morbilidad 2018.
<https://www.datos.gov.co/Salud-y-Proteccion-Social/Tasa-de-natalidad-y-mortalidad-Causas-de-morbilidad/85hy-fkdp>
- Organización de Datos Abiertos de Colombia.
<https://www.datos.gov.co/>
- Organización de Censos poblacionales de Colombia DANE.
<https://www.dane.gov.co/>

DATOS GEOGRÁFICOS:

- Departamentos y municipios de Colombia, Código de la división Político Administrativa del país
<https://www.datos.gov.co/Mapas-Nacionales/Departamentos-y-municipios-de-Colombia/xdk5-pm3f>
- Datos Abiertos de ESRI Colombia
<http://datosabiertos.esri.co/>
- Mapa de figuras geométricas de Departamentos y Municipios de Colombia.
<https://sites.google.com/site/seriescol/shapes>
- Servicios geográficos del SIAC de Colombia
<http://www.siac.gov.co/geovisorconsultas>

Luego de realizar un análisis exhaustivo de las fuentes de datos anteriormente mencionadas, se logra determinar que los datos se encuentran con un fuerte nivel de heterogeneidad, y que gran parte de las estadísticas dadas a nivel de los servicios prestados están determinados por las prestadoras de servicios ya sean IPS y EPS.

Además de esto, se ha contado con la suerte de la existencia de una gran información geográfica, de vías, localizaciones de hospitales y shapefile, con una granularidad permite la división geopolítica desde un nivel estatal, a uno departamental y llegando al municipal, siendo esta la granularidad mínima y el objetivo de este estudio

Sin embargo, se considera que para lograr trabajar con estos conjuntos de datos es necesario realizar un trabajo intensificado de normalización de estructuras, dado que existen numerosos campos de tipo texto libre y sin identificadores claros para poder relacionarlos., se van a homogenizar e ir relacionando cada uno de los valores por los diferentes municipios existentes en Colombia.

Con estos datos, tras aplicarle transformaciones, permitía obtener información relevante en cuanto a los grados de calidad de atención a nivel municipal, y poder usar los datos geoespaciales para poder hacer los primeros gráficos, generando una base de datos robusta.

4 ANÁLISIS DIAGNOSTICO

El análisis diagnóstico del proyecto se comprende por tres tipos de análisis: Entorno y Competitivo, los cuales brindaran un Revisión general de la propuesta, sus fortalezas y debilidades.

4.1 Análisis de Entorno

Latinoamérica, actualmente es una de las regiones del mundo donde se presenta una mayor deficiencia en los sistemas hospitalarios, esto basado en los porcentajes de acceso a servicios hospitalarios, mortalidad infantil y mortalidad en general, como lo indican los estudios realizados en el 2017 por la Organización Panamericana de la Salud (OPS).

Se entiende por acceso a la salud pública, que toda la población sin importar su carácter socioeconómico o étnico pueda ser atendido o hacer uso de los siguientes servicios:

- Sistemas de salud públicos (hospitales y consultorios rurales).
- Vacunación.
- Cobertura en maternidad.
- Sistema de prevención control y tratamiento de enfermedades (como el VIH y cáncer).
- Dotación de equipos médicos de alta tecnología.
- Desarrollar la investigación científica sobre los avances en medicina.

Sin embargo, en el informe de la OPS, se evidencia que más del 30% de la población latinoamericana no tiene acceso a la salud pública, y como solo un 64% de la población rural puede llegar a tener acceso a instalaciones mejoradas.

Actualmente la desigualdad y los altos costos de servicios médicos, representa uno de los mayores dolores de cabeza para la población Latinoamericana, donde la inversión gubernamental en promedio no llega a superar el 6% de los ingresos de las naciones, según datos de la Organización Mundial de la Salud.

La cobertura hospitalaria es esta región también está en gran medida diferenciada por aspectos socio-económicos tales como: Los diferentes estratos sociales, a mayor capacidad económica la población puede acceder a servicios hospitalarios de pago, las condiciones geográficas, las cuales dificultan el acceso e equipamiento y la economías propia de cada país; esto conlleva a que las poblaciones como menores capacidades económicas o en pobreza reciban una menor cobertura hospitalaria en servicios como vacunación, atención primaria y prevención de enfermedades.

*“#SaludUniversal significa participación social, acción gubernamental, trabajo intersectorial y multisectorial, defensa y activismo, para dar voz a aquellos que no son escuchados.
#SaludParaTodos #AlmaAta40 #DiaMundialDeLaSalud”, Carissa Etierre, 7 de abril del 2018, Cuenta oficial de Twitter.*

4.2 ANÁLISIS COMPETITIVO

Por la naturaleza del proyecto y entendiendo que nace como una iniciativa interna del área de I+D de la entidad financiera cuyo único fin es brindarles a los gobiernos o instituciones herramientas para lograr que sus inversiones tengan un mayor impacto, sin esto estar directamente ligado a que sea la Entidad Financiera aquella que otorgue el préstamo al brindar el servicio de análisis.

No consideramos relevante la identificación de competencia en esta área, antes consideramos que en el caso de otras compañías comenzaran con iniciativas como estas, se debería generar un programa de apoyo y mejora continua entre las diferentes entidades con el fin de mejorar y optimizar los modelos de análisis en materia de salud o poder llegar a exportarlos a otras áreas de necesidades de los países.

5 ANÁLISIS DAFO

Se ha realizado un análisis DAFO, mediante el cual hemos identificado a nivel interno las fortalezas y debilidades, y a nivel externo las oportunidades y amenazas, que pueden llegar a incidir en el proyecto:



Figura 7 Análisis DAFO

5.1 DEBILIDADES

A continuación, se realiza una descripción de las debilidades identificadas:

- **Desconocimiento del mercado y del negocio:** existe la posibilidad de que el diagnóstico no llegue a tener un profundo conocimiento del sistema sanitario de la zona a analizar. Esto debido a que cada país tiene diferentes tipos de sistema sanitario, teniendo que ser de vital importancia conocer las diferencias entre sistemas públicos y privados, así como los distintos niveles de cobertura de los individuos y la capacidad de acceder a los servicios ofrecidos. Adicional para obtener mejor resultados es importante conocer la población objetivo, para entender los distintos tipos de hábitos más comunes de la región/población.
- **Datos heterogéneos que requieren fuerte trabajo adicional:** la recolección de datos para el modelo puede ser complejo dependiendo del nivel de variabilidad de los datos y de sus fuentes. Estos datos generalmente variarán debido a distintos factores como la temporalidad (datos en momentos de tiempo distintos) por la granularidad (en algunas zonas se tendrán datos a nivel más detallado y en otra a nivel agregado) o el propio tipo de datos (no es lo mismo analizar una tasa de morbilidad, que un porcentaje de incidencia de una enfermedad, o un ratio de población enferma). Adicionalmente el trabajo de transformación de los datos

provenientes de distintas fuentes hace que se requiera un esfuerzo importante en la fase de ETL.

5.2 AMENAZAS

A continuación, se realiza una descripción de las Amenazas identificadas:

- **Incidencia de factores externos:** existe la posibilidad de que diferentes factores no permitan la edificación del centro médico sugerido en la localización indicada, debido a situaciones no consideradas en el análisis, pero relevantes a la hora de la ejecución real. Unos de los factores ambientales que podrían llegar a afectar es la posibilidad de la existencia de: fallas o actividad tectónicas inestable, zona de alto riesgo volcánico o zonas de riesgo de inundación; o la existencia de factores sociales como: zonas de conflicto, narcotráfico entre otras.
- **Datos no accesibles:** por parte de las entidades sanitarias, debido a que dicha información pertenece a la confidencialidad de los hospitales o información que no estén dispuestos a proporcionar por temas internos de la entidad.

5.3 FORTALEZAS

A continuación, se realiza una descripción de las Fortalezas identificadas:

- **Modelo fácilmente adaptable y escalable:** gracias a esta orientación en el diseño, esto permite la implementación de nuevas variables con el fin de ser continuamente mejorado y especializado. Basado en esto, se ha iniciado con un mínimo MVP que permita una prueba de concepto como punto de partida, la cual luego se ira escalando y mejorando el análisis. La adaptabilidad y escalabilidad también permite poder adaptar el modelo no solo a la sugerencia de un nuevo centro médico sino también a otro tipo de servicios relacionados como farmacias o servicio de ambulancias.
- **Uso de herramientas analíticas y ML:** este proyecto se basa en técnicas de Machine Learning las cuales mejoran sensiblemente el tipo de análisis y la forma de asignar recursos, a como se está realizando actualmente con este tipo de servicios.

5.4 OPORTUNIDADES

A continuación, se realiza una descripción de las Oportunidades identificadas:

- **Proyecto con gran impacto social:** dado que la sanidad es un servicio que afecta a toda la población sin importar su condición social, de raza o género, el hecho de que se esté buscando mejorar la cobertura del servicio mediante un estudio, genera un gran nivel de aceptación y de valor para la región en la que se plantee realizar el proyecto.
- **Tipo de proveedor y fuente de financiamiento:** aunque sea un proyecto interno de la Entidad Financiera, nuestra principal motivación no se basa en que sea quien brinde el préstamo para dicha inversión. Nuestro fin único es que se realice la inversión, ya sea con otras entidades financiera u organizaciones internacionales que al igual que nosotros este dispuestas a apostar por una mejora en la calidad de vida de Latinoamérica.
- **Expectativas del mercado con Big Data:** las fuertes expectativas de la población con respecto a las herramientas de Big Data hace que los organismos estén muy abiertos a la incorporación de este tipo de tecnologías en sus procesos.

6 PLAN ESTRATEGICO

6.1 EL MODELO DE NEGOCIO

A continuación, se describe la definición del modelo de negocio en el cual se representarán relaciones entre las actividades y recursos, intervinientes del negocio (colaboradores, socios y clientes), la propuesta de valor, canales de comunicación y estructura de coste e ingresos.

Socios Claves	Actividades Claves	Propuesta de Valor	Relación con Clientes	Segmento de Clientes
<p>Relaciones Clientes- Proveedores: Empresas o entidades públicas proveedoras de datos de consumo hospitalario y demográfico.</p> <p>Alianzas Estratégicas: Entidades sanitarias públicas y privadas, quienes facilitaran sus propios datos estadísticos internos de consumos hospitalario y su criterio a nivel experto para identificar factores de mejora de los servicios.</p>	<ul style="list-style-type: none"> •Identificación de métricas y KPIs •Definición modelo. •Identificación de región de estudios (ciudad, provincia, país). •Análisis y estructuración de datos hospitalarios, geográficos y demográficos. •Implementación y evaluación del modelo predictivo. 	<p>Sistema de análisis diagnóstico que permita, mediante el uso de métricas sanitarias , determinar la ubicación e impacto de un nuevo centro medico en una región determinada</p>	<p>Relacion de apoyo y asesoramiento en los planes de mejora de servicios sanitarios</p>	<p>Los clientes serán los gobiernos de los países latinoamericanos que estén interesados en invertir en la mejoras del servicio sanitarios de sus países.</p>
	<p>Recursos Claves</p> <p>Técnico:</p> <ul style="list-style-type: none"> •Estructura de base de datos, •Software especializado •Herramienta de visualización <p>Aprovisionamiento:</p> <ul style="list-style-type: none"> •Información demográfica, geográfica y hospitalaria. 		<p>Canales</p> <p>Marketing de le entidad financiera, para la promoción del servicio gratuito de análisis diagnóstico del sistema hospitalario.</p>	
Estructura de Costos		Fuente de Ingresos		
<ul style="list-style-type: none"> •Licencias software •Recursos hardware necesarios •Costes administrativos (Sueldos, equipos) 		<p>Proyecto interno de la entidad financiera, cuyo capital deriva del presupuesto asignado a proyectos de I+D de la entidad.</p>		

Figura 8 Modelo de Negocio Canvas

6.1.1 Propuesta de Valor

La propuesta de valor consiste en un sistema de análisis diagnóstico que permita, mediante el uso de métricas sanitarias, sociales, económicas y políticas, determinar la ubicación e impacto de un nuevo centro médico en una región determinada.



Este sistema de análisis nace como una iniciativa interna de la Entidad Financiera, mediante la cual se quiere incentivar la inversión en la mejora de los sistemas hospitalarios de los países latinoamericanos, brindándoles un sistema basado en métodos científicos mediante el cual puedan a partir de sus datos sanitarios, geográficos y demográficos identificar en que zonas deberían ubicar nuevos centros médicos y predecir el impacto que tendría sobre la población y las estadísticas de atención hospitalaria.

Esta propuesta de valor tiene como único fin darle a conocer a los países el análisis y conclusiones de un estudio personalizado para sus regiones, sin tener que ser la Entidad Financiera aquella con la que contrate el préstamo para lograr esta inversión, puesto que nuestra principal visión como Entidad es ayudar a mejorar la calidad de vida de Latinoamérica.

6.1.2 Socios Clave

Los socios claves para este proyecto serán aquellos que brinden sus datos y criterios de experto con los cuales nuestro sistema diagnóstico se nutre y evoluciona. Para ello hemos identificado dos tipos de socios claves:



- **Clientes-Proveedores:**

Los identificamos como aquellas entidades públicas y empresas encargadas de realizar los estudios estadísticos a nivel sanitario y demográfico. Están serán las encargadas de brindar todos aquellos datos con los cuales se aprovisionará el sistema diagnóstico.

- **Alianzas estratégicas:**

Se establecerán alianzas estratégicas con entidades prestadoras de servicios sanitarios públicas y privadas, quienes brindarán su aporte de dos formas:

- Facilitando sus datos estadísticos internos, tales como niveles de satisfacción, afluencia, principales procesos realizados en salas de urgencias, tiempos de atención, unidades médicas y personal disponible por horario, entre otras.
- Criterio de experto, participaran en la definición de métricas, debido a su experiencia directa con la situación sanitaria, la cual permitirá evolucionar el modelo.

6.1.3 Recursos Claves



Este proyecto cuenta con dos conjuntos de recursos claves importantes: Técnicos y Aprovisionamiento, siendo estos los que nos permitan desarrollar un Sistema de Análisis Diagnóstico robusto y eficientes. A continuación, describimos cada uno de este conjunto de recursos:

- **Técnicos:**
Dentro de esta categoría se identificó:
 - **Hive**, como estructura de bases de datos, donde se almacenará la información hospitalaria, geográfica y demográfica proporcionada por la distintas entidades y empresas identificadas como socios claves
 - **R**, como un entorno y lenguaje de programación con un enfoque al análisis estadístico.
 - **Tableau**, como herramienta de visualización



Figura 9 Recursos Claves Técnicos

- **Aprovisionamiento:**
Los datos hospitalarios, geográficos y demográficos, son nuestra materia prima para el entendimiento, análisis y estructuración de nuestro sistema diagnóstico y las KPIs que se definan. Estos serán brindados por nuestros socios claves y servicios geográficos como Google Maps Api.



Figura 10 Recursos Claves Aprovisionamiento

6.1.4 Relación con Clientes



La relación con nuestros clientes consiste en una relación de apoyo y asesoramiento en la toma de decisiones para gestionar sus planes de mejoramiento del Sistema hospitalario de su región, basándonos en sus propias estadísticas y buscando la mejora de las métricas sanitarias estimadas para tener una mejor cobertura hospitalaria.

Como hemos mencionado nuestro principal fin es que nuestros clientes inviertan en mejorar la calidad de vida en Latinoamérica, en este caso específico en el sector hospitalario por ello estamos

continuamente gestionando un entorno colaborativo con nuestros clientes, para que nos vean como su asesor principal para identificar las soluciones de mejora.

6.1.5 Segmento de Clientes

Nuestro segmento de clientes serán los gobiernos de los países latinoamericanos que estén interesados en invertir en planes de mejoramiento del servicio hospitalarios de sus países. Nuestra solución está orientada a brindar el soporte a este tipo de clientes para que realicen inversiones basados en planes estratégicos definidos a partir de métodos científicos de predicción y que se ajuste a sus situaciones propias como país.



6.1.6 Canales de Distribución



El principal canal de distribución del sistema de análisis diagnóstico de cobertura hospitalaria será mediante herramientas de marketing, las cuales clasificamos en dos tipos:

- **Canales de Comunicación de la Entidad:**
Se realizarán promoción del sistema digital en la página oficial, especificando algunas de las métricas implementadas y pequeños casos de uso realizados para algunas regiones de Latinoamérica y recalando que es un servicio gratuito, con el fin de evidenciar a los clientes el alcance de nuestro estudio y las oportunidades que podría brindarles.
- **Comunicación Directa:**
Se realizará la promoción del servicio gratuito de análisis diagnóstico, en el momento en que los representantes de los gobiernos se encuentren interesados en estudiar las posibilidades de obtener un préstamo para inversión en el sector sanitario en sus países.

6.1.7 Actividades Claves

Las actividades clave necesarias para el desarrollo con éxito del proyecto se centran en cinco fases: Análisis, Diseño, Implementación, Validación y Retroalimentación.

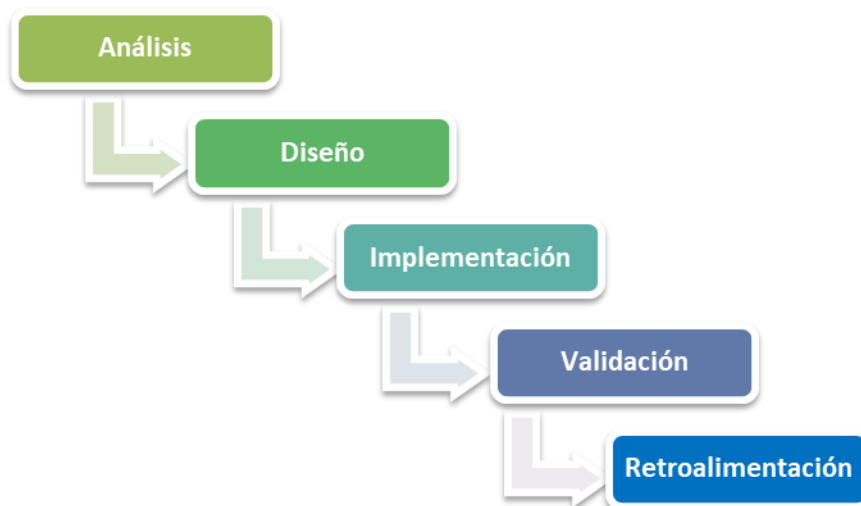


Figura 11 Actividades Claves

Para estas diferentes fases hemos identificados las siguientes actividades claves:

- **Análisis:**
El principal enfoque de esta fase será el análisis de cada uno de los datos con el fin de identificar sus características y relevancia en el estudio.
- **Diseño:**
Esta fase se compone por dos actividades claves:
 - Definición de los KPIs y métricas a evaluar por el modelo
 - Selección de modelo predictivo a implementar
- **Implementación:**
Esta fase se compone por las siguientes actividades claves, cada una es codependiente con la anterior:
 - Definición región a estudiar
 - Obtención de los datos demográficos y sanitarios
 - Obtención de datos geográficos
 - Establecimiento de métricas
 - Desarrollo modelos predictivos
 - Medición de impactos
 - Optimización de resultados
 - Creación de visualizaciones
- **Validación:**
Esta fase está compuesta por la validación de los datos ejecutados para la región definida para a implementación, donde se revisará la coherencia de los resultados con los diferentes KPIs y métricas propuestas.
- **Retroalimentación:**
La principal actividad será recibir retroalimentación del proceso de análisis por parte de nuestros socios claves, haciendo uso de su criterio experto para agregar nuevo KPIs y métricas, logrado robustecer el modelo.

6.1.8 Estructura de Costes



A continuación, se realiza indican los diferentes ítems que representarían un coste para la Entidad Financiera, no se especifican valores de precios debido a temas de confidencialidad de la entidad con respecto a sus costos internos. Los costos los hemos dividido en:

- **Costos Tecnológicos:** se incluyen los siguientes costos:
 - Equipos: se estimaron 6 equipos rentados con un costo de USD 50 por día. No se estimaron computadores para los consultores expertos, ya que no se cree necesario.
 - **Licencias de Cloud (Hive), R y Tableau:** Se estimaron USD 100 mensuales, que incluye cualquier requerimiento de licencia. Cabe destacar que esta es una estimación de alto nivel, ya que el factor precio en este caso depende de muchos factores que se deberán detallar al momento del diseño de la solución real a detalle.

- **Espacio de almacenamiento en los servidores de la entidad:** igual al punto anterior, se estimaron USD 200 mensuales, pero también se deberá revisar al momento de confirmar o estimar una aproximación mejor para el tamaño del conjunto de datos.

De esta forma los costos tecnológicos quedarían de la siguiente manera:

Tecnológicos	Cantidad	Valor mes	Meses	Total
Equipos	6	\$ 50.00	4	\$ 1,200.00
Licencias	1	\$ 100.00	4	\$ 400.00
Almacenamiento	1	\$ 200.00	4	\$ 800.00
				\$ 2,400.00

Tabla 1 Costos Tecnológicos

Hay que tener en cuenta que los costos de Licencias de Tableau y de almacenamiento, según el requerimiento específico del proyecto, podrían volverse en costos recurrentes, y se podría optar por costos mensuales o anuales.

- **Costos Administrativos:** se incluye los siguientes costos:
 - **Costos del lugar de trabajo:** esta estimación se realizó también rentando un espacio para 6 personas, asumiendo que los consultores expertos no requieren espacios dedicados, y se estimó un costo de USD 150 mensual por cada integrante. De esta forma los costos administrativos serían los siguientes:

Administrativo				
Lugar de trabajo	6	\$ 150.00	4	\$ 3,600.00
Equipo de trabajo				\$ 147,200.00
				\$ 150,800.00

Tabla 2 Costos Administrativos Lugar de Trabajo

- **Costo de salarios del equipo de trabajo:** se identificó el equipo de trabajo, el cual tendrá la siguiente composición para ser tomada como referencia para el análisis de los costos, donde se indica el porcentaje de participación tendrán en el proyecto, teniendo en cuenta que el proyecto tendrá una estimación de 16 semanas:

Costo del equipo de trabajo	Asignación
Project Manager	35%
Business Analyst	60%
Technical Team	
Data Extraction Team	50%
Data Transformation Team	50%
Development Team	50%
Reporting Team	50%
Consultores expertos	
Consultores Medicos	10%
Consultores Hospitales	10%
Consultores Economistas	10%

Tabla 3 Asignación porcentual de Costos

A continuación, se detallan estos costos.

Equipo de trabajo	Cantidad	Hora	Día	Asignación	Días	Total
Project Manager	1	\$ 70.00	\$ 560.00	35%	28	\$ 15,680.00
Business Analyst	1	\$ 55.00	\$ 440.00	60%	48	\$ 21,120.00
Technical Team			\$ -		0	\$ -
Data Extraction Team	1	\$ 55.00	\$ 440.00	50%	40	\$ 17,600.00
Data Transformation	1	\$ 55.00	\$ 440.00	50%	40	\$ 17,600.00
Development Team	1	\$ 60.00	\$ 480.00	50%	40	\$ 19,200.00
Reporting Team	1	\$ 55.00	\$ 440.00	50%	40	\$ 17,600.00
Consultores expertos			\$ -		0	\$ -
Consultores Medicos	1	\$ 200.00	\$ 1,600.00	10%	8	\$ 12,800.00
Consultores Hospital	1	\$ 200.00	\$ 1,600.00	10%	8	\$ 12,800.00
Consultores Econom	1	\$ 200.00	\$ 1,600.00	10%	8	\$ 12,800.00
						\$ 147,200.00

Tabla 4 Costos Administrativos Salario Equipo de Trabajo

Como mencionamos anteriormente, estamos suponiendo un equipo con 6 integrantes del equipo base y 3 consultores adicionales expertos. En la columna hora se podrá ver un estimado de costo hora según el rol, en la columna asignación se puede detallar el porcentaje de asignación en la totalidad del proyecto, y en la columna total se puede ver el costo por rol. Como vemos, el equipo de trabajo tiene un costo de USD 147.200.

Como conclusión los costos para la implementación del proyecto serían los siguientes:

Tecnológicos	Cantidad	Valor mes	Meses	Total
Equipos	6	\$ 50.00	4	\$ 1,200.00
Licencias	1	\$ 100.00	4	\$ 400.00
Almacenamiento	1	\$ 200.00	4	\$ 800.00
				\$ 2,400.00
Administrativo				
Lugar de trabajo	6	\$ 150.00	4	\$ 3,600.00
Equipo de trabajo				\$ 147,200.00
				\$ 150,800.00
TOTAL PROYECTO				\$ 153,200.00

Tabla 5 Estructura de Costos

6.1.9 Fuente de Ingresos



Como hemos indicado este proyecto nace como iniciativa interna de la entidad para brindar un apoyo y asesoramiento a sus clientes de forma gratuita, por tanto, su único ingreso será el presupuesto asignado al área de I+D de la entidad para que continúe con este tipo de iniciativas.

7 PLAN DE ACCIÓN

7.1 Definición Alcance y Objetivo Del Proyecto



El sistema de análisis diagnóstico de cobertura hospitalaria en Latinoamérica tiene como objetivo brindar un apoyo a los diferentes gobiernos latinoamericanos que estén interesados en realizar una inversión en materia sanitaria, brindadores un estudio que les permita identificar donde crear nuevos centros médicos y cuál sería el impacto de estos en sus regiones.

Para medir el impacto que se espera generar al ubicar un nuevo centro médico, se han definido un conjunto de métricas, que permiten estandarizar los resultados y medirlos desde diferentes frentes, las cuales se agruparon de la siguiente manera:



- **SOCIO - SANITARIAS:**

Dichas métricas están directamente relacionadas con la situación sanitaria de la zona analizada. Entre ellas podemos identificar las siguientes:

- Índices de satisfacción con la atención sanitaria
- Número de días de espera de para atención
- Tasa de mortalidad infantil, lactantes, embarazada, tercera edad y en general.
- Tasa de morbilidad.
- Índices de incidencias de enfermedades, tales como malaria, cólera, dengue, gripe, entre otras.

- **SOCIO - ECONÓMICAS:**

Estas métricas no se relacionan directamente relacionadas con los indicadores de los servicios de salud, pero si influyen económicamente sobre su calidad y aceptación social. Entre ellas hemos identificado:

- Salario mínimo
- Aportes a la sanidad
- Sueldo medio de la población
- Porcentaje del PIB asignado a la salud.

- **POLÍTICAS:**

Estas métricas, analizaran aquellos factores políticos que puedan llegar a influir en la ejecución del proyecto, entre estas hemos identificado las siguientes:

- Planes de ordenamiento territorial de la zona
- Periodo electoral
- Temporada de elecciones

Con la definición de cada una de estas métricas, se puede adaptar el modelo a cada una de las situaciones sociales de las regiones.

7.2 Definición de Actividades

Este proyecto se ha compuesto por cinco fases: Análisis, Diseño, Implementación, Validación y Retroalimentación, cada una de ellas es complementaria a la anterior, para ello cada una están compuesta por diferentes actividades para obtener un output para la siguiente fase.

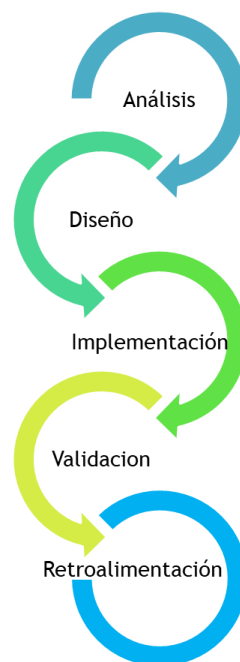


Figura 12 Fases del Proyecto

A continuación, se identifican cada una de las fases del proyecto y sus actividades correspondientes:

7.2.1 Análisis

Esta fase consiste el análisis de los principales recursos claves del proyecto con los cuales daremos inicio al diseño, dentro los cuales identificamos:

- **Estudio y definición requerimientos técnicos:**

Se revisan las diferentes necesidades técnicas del proyecto con las que deberá cumplir la arquitectura que se diseñe en la siguiente fase. Entre necesidades se han identificados los siguientes aspectos:

- Almacenamiento
- Proceso de ETL (Extraction , Transformation and Load, por sus siglas en ingles)
- Proceso de análisis y modelado
- Publicación de resultados.

- Selección de la zona de análisis:

Se deberá seleccionar la zona de análisis, la cual deberá cumplir con unos requisitos mínimos de tamaño, el cuál oscila entre los 500 a 1000 km², dentro de los cuales se pueda medir objetivamente el impacto de una población, dado que la cantidad de información será importante para nutrir el modelo y entrenarlo. Sin embargo, se tiene en cuenta que la zona no debe superar unos límites establecidos puesto que realmente, aunque se evalúe mucha información el impacto no tendrá el mismo nivel puesto que la cantidad de población beneficiada será mayor a la debida.

Para el MVP desarrollado, se ha seleccionado como región objetivo los 1112 municipios del país de Colombia. Se analizará sus áreas de influencia de cada uno de ellos.

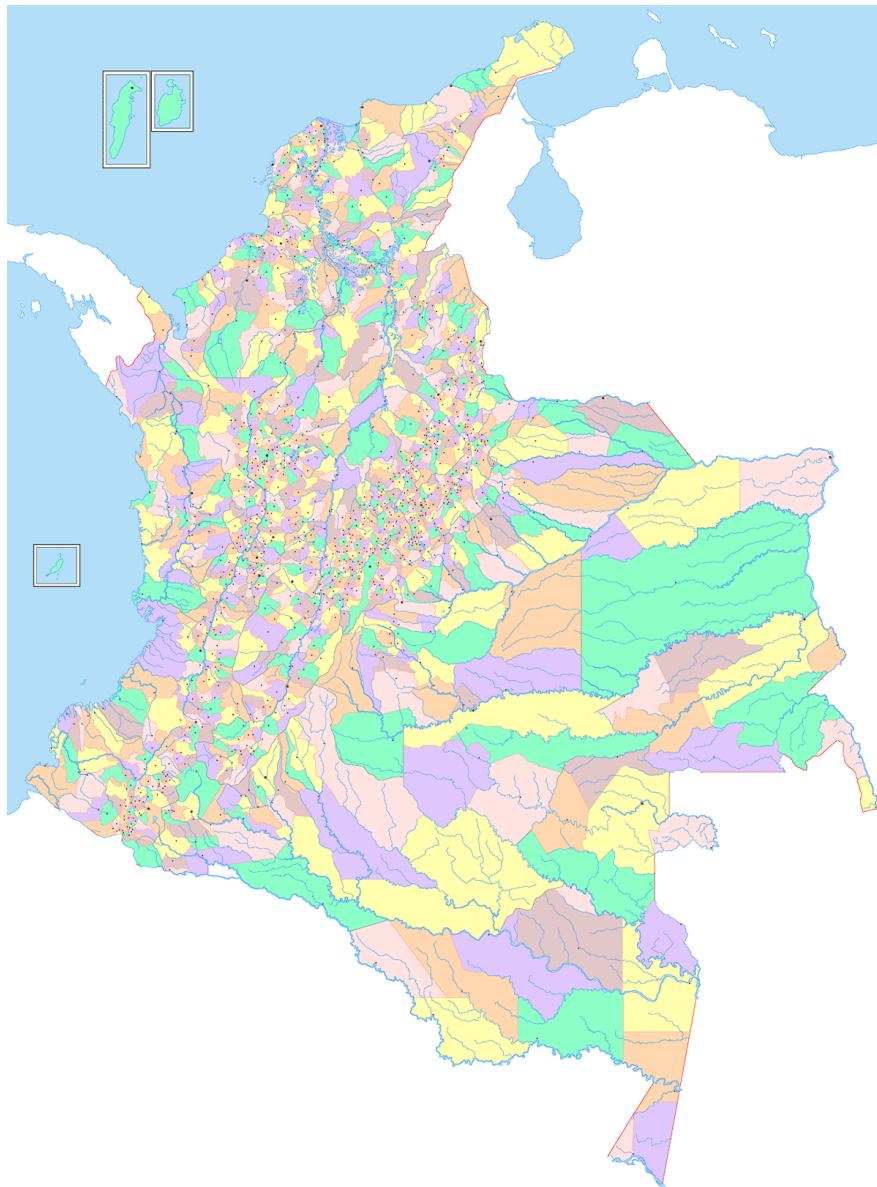


Figura 13 Mapa de Municipios de Colombia

- **Análisis de fuentes de datos:**

Se definirán las fuentes de información, las cuales será aprovisionadas por nuestros socios claves siendo estas de carácter público y privado, además se deberán analizar el tipo de transformaciones iniciales que se deberán realizar para que los datos se puedan almacenar e implementar en nuestro sistema. Además, podremos analizar y definir como se tratarán aspectos como: la utilidad de las variables aprovisionada, el peso que tendrán sobre cada una de las métricas definidas, niveles de granularidad, entre otras, cuando se obtengan los datos de la región de estudio.



Para el MVP desarrollado, se analizaron las fuentes de datos y definieron aquellos valores que serían implementados para el desarrollo del modelo. Las fuentes de datos analizadas fueron las siguientes:

- **Datos sociales: (Fuente DANE)** Datos de encuesta a censo relativos a calidad del servicio sanitario.
- **Datos geográfico: (Fuente GoogleMaps)** Archivos shapefile con granularidad municipal del territorio, vías de acceso, ubicación de Hospitales
- **Datos Demográficos: (Fuente Datos abiertos)** población de cada uno de los municipios .

7.2.2 Diseño

Esta fase consiste en el diseño de funcional y técnico teniendo en cuenta los recursos claves analizados en la fase anterior, con los cuales iniciaremos la fase de implementación, dentro de las cuales identificamos:

- **Definición de arquitectura técnica:**

A partir del análisis de recursos técnicos, se plantea el diseño arquitectónico a implementar en el cual una arquitectura big data pueda suplir las necesidades expuestas.

Para el MVP desarrollado se definió una arquitectura con una base de datos en HIVE, R, como software de modelo y Tableau como herramienta de visualización. Para más información ver el apartado “3.3 Definición Arquitectura Técnica”.



Figura 14 Definición Arquitectura Técnica

- **Definición de métricas y KPIs:**



Teniendo en cuenta las fuentes de información definidas, se identificarán las métricas con las que se establecerán los KPIs teniendo en cuenta sus valores. Además, se definirá el peso, valor y rango en el modelo.

Para el MVP desarrollado, se han definido las siguientes métricas :

- Ratio de satisfacción del servicio sanitario: cuyos valores son de 1 a 4 dependiendo, siendo uno el mayor valor posible, en esta métrica se eliminaron valores de vacíos.
- Número de personas enfermas atendidas en el centro médico.
- Número de personas que asisten a un centro médico.
- Población
- Número de hospitales
- Ubicación de los centros médicos
- Ubicación de los centros de los municipios del departamento

Teniendo en cuenta las métricas anteriores, se definieron los siguientes KPIs:

- Porcentaje de enfermos sobre población total
- Porcentaje de enfermos que se atendieron en un centro de médicos específico.
- Porcentaje de enfermos atendidos por todos los centros médicos.
- Radio de influencia de los centros médicos dentro del municipio (punto céntrico).
- Ratio de personas entre la cantidad de hospitales.
- Distancia media entre centro departamento y hospital más cercano al mismo.
- Ratio de atención a personas en el radio de influencia del hospital más cercano
- Distancia media de hospitales en zona de influencia
- Distancia media de todos los hospitales del departamento respecto al municipio

7.2.3 Implementación

Esta fase consiste en la implementación del proceso de extracción de datos, transformación de datos y desarrollo del modelo predictivo con las métricas y KPIs definidos, con los cuales se harán pruebas en fase producción para entrenar el modelo. Dentro de esta fase se identifican las siguientes actividades.

- **Obtención de los datos:**

Se ha creado un proceso de extracción de los datos sanitarios, geográficos y demográficos, los cuales serán cargados en la base de datos Hive, previamente definida en la fase de diseño.

Para el MVP desarrollado, se generaron distintos procesos de obtención de datos. Estos procesos fueron desarrollados en R para luego generar .csv que fueron almacenados en la base de datos de HIVE, para luego ser analizado.

Para cada uno de los tipos de datos se generaron distintas funciones, encontraremos tres procesos de obtención de datos:

- Datos Espaciales, mediante el cual se obtuvo la información geográfica de los 33 departamentos de Colombia y sus 1112 municipios.
- Datos Censo, esto contenían la información de las encuestas de calidad y los datos demográficos de la población para cada uno de los municipios
- Datos Hospitales, con el cual se obtuvieron los 20 primeros centros hospitalarios de cada uno de los municipios.

A continuación, se ilustra un fragmento de código para cada uno de los procesos de extracción previamente mencionados.

Obtención de Datos Espaciales:

```
#obtiene los datos espaciales y los transforma a un CRS comun
obtenerDatosEspaciales <- funcion() {
  #Obtención de datos .shp de los municipios de colombia
  i_data_gs <- readOGR(dsn = "datos/DatosEspaciales", layer = "mpio")

  #pasamos a numeric algunos de los datos

  i_data_gs$num_DPTO<-as.numeric(i_data_gs$DPTO)
  i_data_gs$num_MPIO<-as.numeric(i_data_gs$MPIO)
  i_data_gs$num_MPIOS<-as.numeric(i_data_gs$MPIOS)

  i_data_gs = spTransform(i_data_gs, CRS("+init=epsg:3116"))
  return (i_data_gs)
}
```

Figura 16 Fragmento de Código Obtención de Datos Espaciales

Obtención de Datos Censo:

```
loadCenso = funcion(idDepartamento, porcentaje_influencia) {
  # función privada para procesar el resumen de censo por porcentaje_influencia
  procesaResumen = funcion(dataCenso){
    numElems = nrow(dataCenso) * porcentaje_influencia
    #cojo porcentaje de casos aleatorios
    muestra = dataCenso[sample(nrow(dataCenso), numElems),]

    resumen_censo <- muestra %>%
      group_by(U_MPIO) %>%
      summarise(numero=sum(contador) ,
                enfermo=sum(P_ENFERMO_T) ,
                atencion=sum(PA_LO_ATENDIERON_T) ,
                calidad=sum(PA1_CALIDAD_SERV_T))

    resumen_censo$porcentaje_influencia <- porcentaje_influencia
    resumen_censo$numero <- resumen_censo$numero * porcentaje_influencia
    resumen_censo$enfermo <- resumen_censo$enfermo * porcentaje_influencia
    resumen_censo$atencion <- resumen_censo$atencion * porcentaje_influencia
    resumen_censo$calidad <- resumen_censo$calidad * porcentaje_influencia
    resumen_censo$MPIO <- as.factor(resumen_censo$U_MPIO)
    return (resumen_censo)
  }
  # obtiene el censo del hash. Para evitar recargar el csv
  censo = censos[[as.character(idDepartamento)]]
  #
  # #si ya lo ha cargado, procesa resumen
  if (is.null(censo))
  {
    return (procesaResumen(censo))
  }
  # si no esta cargado, lo carga

  pathFile = paste(rutaCenso, idDepartamento, ".csv", sep = "")
  censo<-read_csv(pathFile)
  #head(censo[c(1:10),], 10)
  #dim_censo<-dim(censo)
  censo$contador<-1

  #Transformamos las variables para poder luego hacer metricas con ellas
  censo$P_ENFERMO_T<-censo$P_ENFERMO_T
  censo$P_ENFERMO_T[censo$P_ENFERMO_T==2]<-0
  censo$P_ENFERMO_T[censo$P_ENFERMO_T==9]<-0
  censo$P_ENFERMO_T[is.na(censo$P_ENFERMO_T)]<-0

  censo$PA_LO_ATENDIERON_T<-censo$PA_LO_ATENDIERON_T
  censo$PA_LO_ATENDIERON_T[censo$PA_LO_ATENDIERON_T==2]<-0
  censo$PA_LO_ATENDIERON_T[censo$PA_LO_ATENDIERON_T==9]<-0
  censo$PA_LO_ATENDIERON_T[is.na(censo$PA_LO_ATENDIERON_T)]<-0

  censo$PA1_CALIDAD_SERV_T<-censo$PA1_CALIDAD_SERV_T
  censo$PA1_CALIDAD_SERV_T[censo$PA1_CALIDAD_SERV_T==9]<-0
  censo$PA1_CALIDAD_SERV_T[is.na(censo$PA1_CALIDAD_SERV_T)]<-0

  #guardo el censo en memoria ya cargado
  censos[[as.character(idDepartamento)]] = censo
  #proceso resumen
  return (procesaResumen(censo))
}
```

Figura 17 Fragmento de Código de Obtención de Datos Censo

Obtención de Datos Hospitales:

```

loadHospitales = function(mpios) {
  mpio_gs = getMpio(mpios)
  nombreDep = as.character(mpio_gs$NOMBRE_DPT[1])
  nombreMpio = as.character(mpio_gs$NOMBRE_MPI[1])

  pathFichero = paste(rutaHospitales, nombreDep, sep="")
  dir.create(pathFichero)
  rutaFichero = paste(pathFichero, "/", nombreMpio, ".csv", sep = "")

  if (file.exists(rutaFichero)){
    hospitales = (read.csv2(rutaFichero))
    colnames(hospitales) = c("#", "nombre", "direccion", "latitud", "longitud")
    hospitales$latitud = as.numeric(as.character(hospitales$latitud))
    hospitales$longitud = as.numeric(as.character(hospitales$longitud))
    return (hospitales)
  }

  # Clave API googlemaps AIzaSyCTYPNpzF4coInsaYPVqCNIh_gZms_M3o
  API_KEY<- "AIzaSyCTYPNpzF4coInsaYPVqCNIh_gZms_M3o"
  #obtencion de hospitales
  searchString = paste("hospital,", nombreMpio, nombreDep, ", Colombia, America")
  res <- google_places(search_string = searchString, key = API_KEY)
  name_hospitales<-res$results$name
  dir_hospitales<-res$results$formatted_address

  token<- res$page_token

  hospitales =
  data.frame(name_hospitales, dir_hospitales, res$results$geometry$location$lat, res$results$geometry$location$lng)
  #hospitales<-unique(hospitales)
  colnames(hospitales) = c("nombre", "direccion", "latitud", "longitud")
  hospitales$latitud = as.numeric(as.character(hospitales$latitud))
  hospitales$longitud = as.numeric(as.character(hospitales$longitud))

  write.csv2(hospitales, file= rutaFichero)
  return (hospitales)
}

```

Figura 18 Fragmento Código obtención Datos Hospitales

- **Transformación de los datos:**

Se ha creado un proceso de transformación de los datos para normalizarlos, homogenizarlos, con el fin de que sean tratados de forma numérica y por rango para la evaluación del modelo.



En el MVP desarrollado , una vez obtenidos cada uno de los datos anteriormente descritos, se desarrolló un proceso de transformación y calculo de nuevos datos que serán implementado en el modelo.

Inicialmente se relacionan cada uno de los municipios, a partir de su nombre, con los 20 hospitales encontrados para cada uno de ellos y las métricas de los censos de cada uno.

A continuación, se ilustra un fragmento del código implementado en cálculo de métricas para cada municipio.

```
#input: nombre Municipio
calculaMetricasMpio = function(MPIOS){
  #data frame de respuesta
  ret = crearDataFrameVacio()
  #información de los hospitales recuperada de google
  hospitales = loadHospitales(MPIOS)

  for (row in 1:nrow(hospitales)){

    # para cada hospital, hayamos su area de influencia
    hospital = hospitales[row,]
    nombreHospital = as.character(hospital$nombre)
    print(paste('hospital',row,'de',nrow(hospitales),'Nombre:',nombreHospital))
    #primero transformamos las coordenadas
    metricasHospi = calculaMetricasHospital(hospital$longitud,hospital$latitud,
    nombreHospital)

    ret = rbind(ret,metricasHospi)
  }

  return (ret)
}
```

Figura 19 Fragmento Código Cálculo Métricas Municipio

Una vez realizado el cálculo de estas métricas, se procede a calcular las métricas hospitalarias, siguiendo estos procedimientos:

1. Obtiene el área de influencia del hospital
2. Recupera las zonas influenciadas
3. Por cada municipio influenciado obtiene el área de influencia (intersección del municipio con el área de influencia del hospital)
4. Descarta los municipios con una influencia inferior al 10%
5. Obtiene la distancia del hospital al centroide del municipio
6. Obtiene los datos de censo del municipio según el porcentaje de influencia
7. Genera los datos de respuesta

```
calculaMetricasHospital = function (latitudHospital, longitudHospital,
nombreHospital){

  ret = crearDataFrameVacio()

  hospital_gs = transformarPuntos(data.frame(latitudHospital,longitudHospital))
  # coordenadas en epsg4326
  coordenadasHospital = coordinates(transformarPuntosA4326(hospital_gs))
  # creamos el area de influencia
  areaInfluencia_hospital = gBuffer(spgeom=hospital_gs, width =
RADIO_INFLUENCIA_HOSPITAL)
  # recuperamos zonas influenciadas
  zonasInfluenciadas_gs = i_data[areaInfluencia_hospital,]
  arrayZonasInfluenciadas = as.character(zonasInfluenciadas_gs$MPIOS)
  # por cada mpio influenciado hayamos los datos
  n_mpios = length(arrayZonasInfluenciadas)
  for(indexZonaInfluenciada in 1:n_mpios){
    # datos espaciales del mpio
    longMpio = arrayZonasInfluenciadas[indexZonaInfluenciada]
    mpio_gs = getMpio(longMpio)
    #calculo el % area influenciada
    mpioInfluenciado_polygon = as(mpio_gs, 'SpatialPolygons')
    intersection_data = raster::intersect(areaInfluencia_hospital,
mpioInfluenciado_polygon)
    porcentaje_Influencia = raster::area(intersection_data)/raster::area(mpio_gs)
    if (porcentaje_Influencia < 0.10){
      next
    }
  }
  # obtengo informacion a partir de los datos espaciales
  idMpio = as.character(mpio_gs$MPIO[1])
  idDepartamento = as.character(mpio_gs$DPTO[1])
  nombreDep = as.character(mpio_gs$NOMBRE_DPT[1])
  nombreMpio = as.character(mpio_gs$NOMBRE_MPI[1])
  # calculo centroide mpio
  centroMpio = gCentroid(mpio_gs)
  coordenadasMpio = coordinates(transformarPuntosA4326(centroMpio))
  # calculo distancia del hospital al centroide del mpio
  distancia = gDistance(centroMpio, hospital_gs)
}
```

Figura 20 Fragmento Código Cálculo Métricas Hospitalarias

Finalmente se obtiene un fichero de datos con la información relevante para el proceso de modelado.

CAMPO	DESCRIPCIÓN
IDDEP	Identificador del departamento según el código DANE
NOMBREDEP	Nombre del departamento
IDMPIO	Identificador del municipio según el código DANE
NOMBREMPIO	Nombre del municipio
NOMBREHOSPI	Nombre del hospital sobre el que se realiza el análisis
LATITUDHOSPI	Coordenadas de latitud del hospital proyectados en sistema epsg4326
LONGITUDHOSPI	Coordenadas de longitud del hospital proyectados en sistema epsg4326
LATITUDCENTROIDE	Coordenadas de latitud del centroide del municipio proyectados en sistema epsg4326
LONGITUDCENTROID	Coordenadas de longitud del centroide del municipio proyectados en sistema epsg4326
DISTANCIA	Distancia calculada entre el hospital y el centroide del municipio
PERSONAS	Número de personas totales influenciadas en proporción por el hospital
ENFERMAS	Número de personas que han declarado estar enfermas
ATENDIDAS	Número de personas que han declarado estar atendidas
N_MPIOS	Número de municipios que influencia el hospital
P_INFLUENCIA	Porcentaje de influencia del hospital en el municipio
P_ENFERMAS	Porcentaje de personas enfermas
P_ATENDIDAS	Porcentaje de personas atendidas
R_CALIDADATENDIDAS	Ratio de calidad medio de las personas atendidas
R_CALIDADENFERMAS	Ratio de calidad medio de las personas enfermas

Tabla 6 Fichero de Datos

- **Desarrollo de modelos predictivos:**

Se desarrolló el modelo predictivo teniendo en cuenta las características propias del lenguaje R, mediante los cuales se generaron dataframes de prueba y validación para entrenar al modelo basado en un algoritmo de Árboles de Decisión, evaluando cada uno de los KPIs. definidos.

Para el MVP desarrollado, se desarrollaron una serie de modelos predictivos, con los cuales se desea determinar los ratios de calidad del servicio hospitalario.

Se han desarrollado los siguientes modelos predictivos:

- Linear Regression
- Generalized Linear Models
- XGBoost
- Random Forest,

```
#Modelo GLM
model1<-train(x=train_p[,i_indep], y=train_p$r_calidadAtendidas, method = "glm")
model1
summary(model1)
predict1<-predict(model1, test_p)

#Modelo CART
model2<-train(x=train_p[,i_indep], y=train_p$r_calidadAtendidas, method = "rpart")
model2
summary(model2)
predict2<-predict(model2, test_p)

#Modelo Random Forest
model3<-train(x=train_p[,i_indep], y=train_p$r_calidadAtendidas, method = "rf")
model3
summary(model3)
predict3<-predict(model3, test_p)
```

Figura 21 Fragmento de Código Modelos Predictivos

- Análisis y selección de los modelos:**
 Se analiza y se selecciona el modelo de Machine Learning más adecuado a desarrollar para establecer la relación entre la variable objetivo y el resto de las variables. El método dependerá de la variable objetivo y de los predictores a incluir.

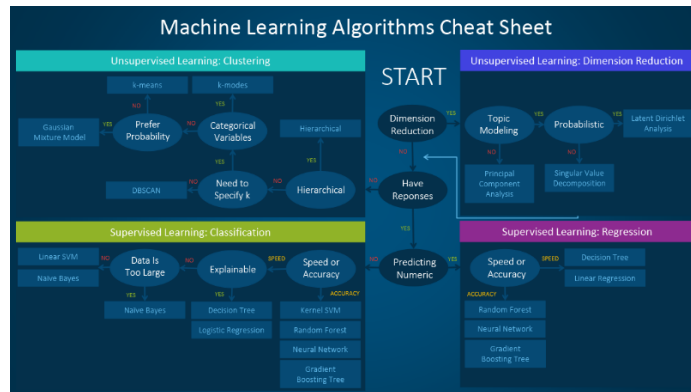


Figura 22 Algoritmos de Machine Learning

Para el MVP desarrollado, se testearon cada uno de los modelos desarrollados mediante una matriz de confusión, mediante la cual se analizaron cada uno de los resultados obtenidos basándose en aquel modelo que diera mayores métricas, algunas de las métricas implementadas fueron AIC, BIC, RSquare. Finalmente el modelo seleccionado ha sido Random Forest.

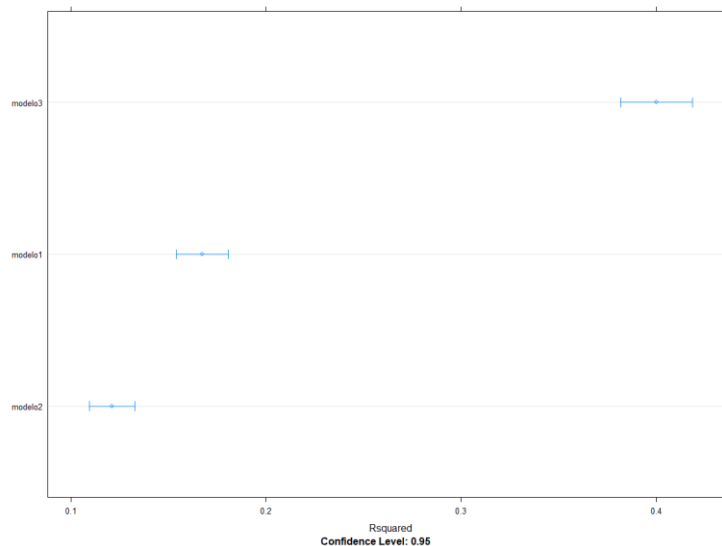


Figura 23 Matriz de Confusión

- Prueba de concepto:**
 Se ejecuta el proceso de generación del análisis diagnóstico, desde la fase de extracción, transformación de datos y ejecución del modelo para la zona seleccionada. Con estos procesos se obtendrá un último dataset con el que se aprovisionarán las diferentes visualizaciones en Tableau, que se crearán en la siguiente actividad.

Una vez realizados los procesos de extracción de datos y transformación de los mismos, se procede a la ejecución del modelo previamente seleccionado Random Forest. Dentro de la ejecución de este modelo, se ajusta para que no solo evalúe un escenario si no contemple la posibilidad de analizar sus resultados basados en la ubicación de distintos hospitales en nuevas zonas con el fin de validar como afecta esto a las métricas y cálculos de calidad.

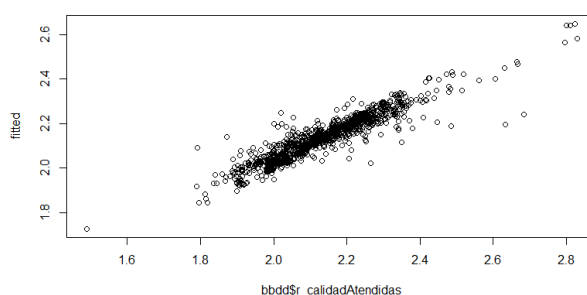


Figura 24 Ajuste Modelo con Diferentes Escenarios

Una vez se definen los escenarios (las latitudes y longitudes en las que ubicar un nuevo hospital) se ajusta el modelo, y se vuelven a agrupar los municipios y preparar la exportación a un fichero csv que Tableau leerá y presentarán

A continuación, se ilustra un fragmento de código del proceso de extracción de resultados para Tableau.

```
#-----
#Añadimos hospital
#-----
#Incluimos las métricas del hospital

#esc1 <- latitudHospital = -76.6866, longitudHospital = 5.14867
#esc2 <- latitudHospital = -75.204390, longitudHospital = 10.069994
#esc1 <- latitudHospital = -76.6866, longitudHospital = 5.14867

metricasHospital = calculaMetricasHospital(latitudHospital = -75.204390
, longitudHospital = 10.069994, 'hospitalX', data_gs = i_data)

bbdd_new_hospital <- rbind(bbdd, metricasHospital)

#Predecimos para el hospital
fitted_esc2 <- predict(model3, bdd_new_hospital)

#Añadimos el nuevo hospital a la base de datos de hspoitka
bbdd_new_hospital$esc2 <- fitted_esc2
bbdd_new_hospital$esc2[1:nrow(bbdd)] <-
bbdd_new_hospital$r_calidadAtendidas[1:nrow(bbdd)]

#Resumimos de nuevo los municipios
bbdd_mpios_new <- bdd_new_hospital %>%
  group_by(nombreMpio) %>%
  dplyr::summarize(distancia=mean(distancia),
    Personas=sum(Personas),
    Enfermas=sum(Enfermas),
    Atendidas=sum(Atendidas),
    n_mpios=mean(n_mpios),
    P_influenzia=mean(P_influenzia),
    S_influenzia=sum(P_influenzia),
    p_enfermas=Enfermas/Personas,
    p_atendidas=Atendidas/Enfermas,
    r_calidadAtendidas=mean(r_calidadAtendidas),
    n_hosp=sum(unidad),
    esc2=mean(esc2)
  )

bbdd_escenarios$esc2 <- bdd_mpios_new$esc2
write.csv2(bbdd_escenarios, "bbdd_escenarios.csv")
```

Figura 25 Fragmento Código Extracción Resultados Tableau

- **Creación de visualizaciones:**

Se crearon las diferentes visualizaciones con la herramienta de Tableau, en la cual se graficaron mapas de calor, identificación de puntos y graficas lineales para visualizar el impacto y la evaluación de los KPIs.

Para el MVP desarrollado en Tableau, inicialmente se insertó un objeto geográficos que contenía la información de cada uno los municipios como polígonos, ya que a partir de esto se podrían llegar ubicar los hospitales y demarcar las zonas de afluencia para cada municipio.

Una vez obtenido esto se relaciono el fichero, obtenido en el proceso de extracción anterior, con los nombres de cada uno de los municipios. Una vez terminada la fuente de datos se procede a generar el cuadro de mando, el cual consistirá en un mapa de calor donde se visualiza como se observa el comportamiento de cada uno de los escenarios con referentes a las métricas de calidad de atención.

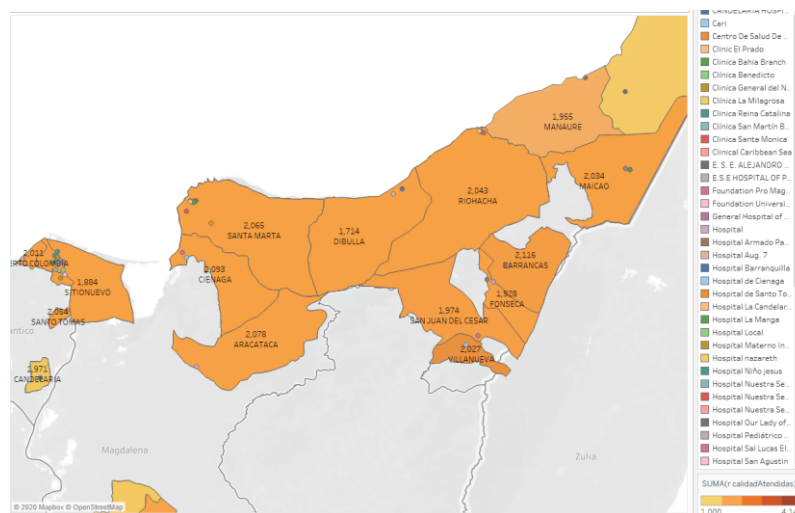


Figura 26 Ejemplificación Cuadros de Mando

7.2.4 Validación

Esta fase consiste en la validación de los resultados mediante la medición de impactos y optimización de los resultados. A continuación, se detalla los procesos de cada una de las actividades.

- **Medición de impactos**

Se evaluaron los resultados luego de la ejecución del modelo, haciendo varias pruebas con los diferentes datos, y evaluando el impacto de los resultados obtenidos, esto nos permite identificar la mejor conjugación de variables que nos darán un mejor impacto en la ubicación de un nuevo centro médico.

- **Optimización de resultados**

Se ha definido una prueba de optimización, creando una matriz con una serie de puntos aleatorios definidos para nuevos centros médicos, con tal de evaluar el impacto de estos y cómo se comporta el modelo frente a La prueba.

7.2.5 Retroalimentación

Esta fase consiste en la implementación del proceso de extracción de datos, transformación de datos y desarrollo del modelo predictivo con las métricas y KPIs definidos, con los cuales se harán pruebas en fase producción para entrenar el modelo. Dentro de esta fase se identifican las siguientes actividades.

- **Mejoras de los modelos:**
Se deben establecer mejoras continuas de la herramienta de cara a hacerla más extensiva, más relevante y fácil de usar. Estas mejoras se relacionarán con las siguientes capacidades:
 - Usabilidad: mejoras relativas al uso y manejo de la herramienta. De forma que se mejore el acceso a los datos pudiendo automatizar el mismo. La transformación de los datos como la visualización
 - Flexibilidad: mejoras destinadas a hacer más flexible el análisis, por ejemplo, dejando la posibilidad de implementar distintas variables objetivas.
 - Predicción: mediante el cual se tomen los datos de fechas anteriores para proyectar los mismo a 5 o 10 años, teniendo en cuenta aspectos como movimientos migratorios, relación poblacional, entre otros. Y con estas proyecciones realizar análisis prospectivos y establecer los impactos y diagnósticos en un futuro hipotético proyectado.

- **Inclusión de nuevas métricas:**
Con la ayuda de nuestras alianzas estratégicas, se recibirá retroalimentación de los procesos realizados con el objetivo de crear métricas más específicas que se alineen cada día más a las situaciones sanitarias actuales.

7.3 DEFINICIÓN DE ARQUITECTURA TÉCNICA

A continuación, se ilustra el modelo de arquitectura diseñado para la ejecución del proyecto, teniendo en cuenta cada uno de los requerimientos definidos en la fase de Análisis.

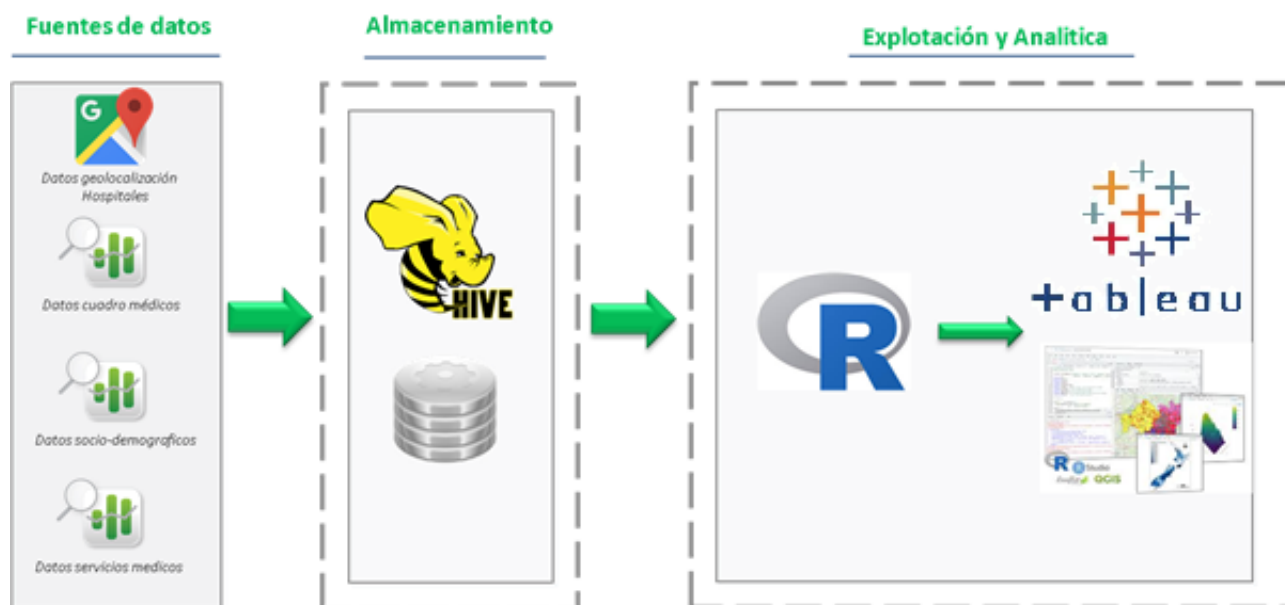


Figura 27 Arquitectura Técnica

7.4 ANÁLISIS DE RECURSOS

A continuación, se describe el proceso de análisis de recursos de Talento Humano, Físicos y Financieros.

7.4.1 Talento Humano

Consideramos que el factor humano representa uno de nuestros recursos principales y la gestión de este supone una actividad de apoyo esencial para el éxito del proyecto. La gestión del personal incluye la selección del talento adecuado para el desempeño de las funciones establecidas.

Para este proyecto hemos definido los diferentes perfiles especializados en: Datos, tanto en su análisis, validación y transformación de estos, desarrollo de aplicaciones en R y modelado y gestión de equipos. Teniendo en cuenta lo anterior hemos definido el siguiente organigrama.

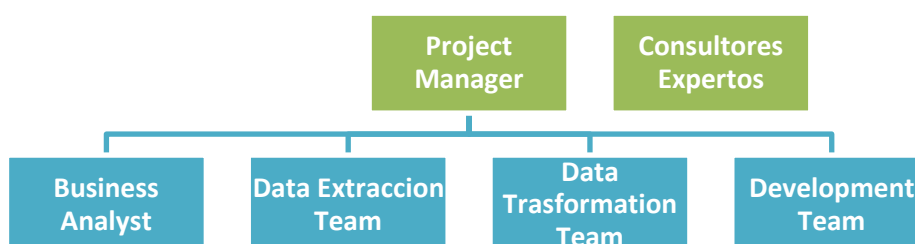


Figura 28 Organigrama

- Project Manager**
 Es conocedor de las necesidades del producto y velará por que las partes estén alineadas, dando solución a los inconvenientes que puedan surgir, definirá las necesidades y tiempos de cada uno de los tres equipos. Debe ser una persona con un perfil proactivo, con inteligencia emocional, pensamiento analítico, con conocimiento y experiencia en aplicación de metodologías ágiles y con liderazgo
- Analista de Negocio**
 Es conocedor técnico y al tiempo entiende y conoce las necesidades del negocio. Es la persona encargada de traducir las necesidades de negocio en requerimientos informáticos. Hay que destacar que en este proyecto los usuarios que ayudarán a definir los requerimientos funcionales podrían estar confirmados por expertos economistas de la Entidad Financiera, o los Consultores Expertos que se contratarán para tal fin.
- Data Extraction Team:**
 Como experto en datos, se requiere de perfiles senior que sea de definir procesos de extracción. Este equipo debe estar conformado con perfiles con experiencia en análisis de datos y conocimientos de lenguajes de programación. Personas proactivas, metódicas, organizadas y de pensamiento analítico.
- Data Transformation Team:**
 Como se ha indicado el origen de los datos son heterogéneos y de diversas fuentes, tanto públicas como privadas, por tanto, se requiere un equipo capaz de llevar a cabo el análisis de

orígenes de datos (fuentes fiables), el modelado de datos (que guardar) y como homogenizarlos en la base de datos.

- **Development Team:**
Se requiere un equipo con conocimiento del lenguaje R y herramientas como Tableau. Este equipo tiene un perfil de gente capaz de estimar las tareas y realizar el modelado funcional de la aplicación, con experiencia en visualización de datos.
- **Consultores Expertos:**
Para la revisión del modelo y selección de variables se requiere mucha experiencia en este tipo de modelos. Por esto se está previendo contratar consultores expertos en sector médico, hospitales y economistas.

7.4.2 Recursos físicos

Como recursos físicos, será necesario contar con un espacio de trabajo adecuado para el desarrollo de un proyecto por parte de un equipo con diferentes tipos de perfiles. Este espacio debe contar con un espacio para reuniones y la tecnología necesaria para las mismas, además con equipos y sitios de trabajo que cumplen con las normas de seguridad internacionales. Todos estos recursos físicos anteriormente mencionados se encuentran como dentro de las instalaciones de la Entidad Financiera, lo cual no representa un esfuerzo para la misma.

7.4.3 Recursos financieros

Como hemos indicado este proyecto nace como iniciativa interna de la entidad para brindar un apoyo y asesoramiento a sus clientes de forma gratuita, por tanto, su único ingreso será el presupuesto asignado al área de I+D de a entidad para que continúe con este tipo de iniciativas.

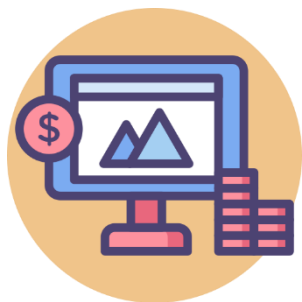
7.5 GESTIÓN DEL TIEMPO: CRONOGRAMA

A continuación, se indica el cronograma establecido para la resolución del proyecto, este se desarrolla a partir de las cinco fases del proyecto.

FASES PROYECTO	RECURSOS	SEMANAS															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
FASE I: ANÁLISIS																	
Estudio requerimientos técnicos	Development Team	█															
Selección de la zona de cobertura	Project Manager		█														
Análisis de fuentes de datos	Data Extracction Team		█	█													
FASE II: DISEÑO																	
Definición de arquitectura técnica	Development Team			█													
Definición de KPIs y métricas a usar	Equipo Completo				█												
FASE III: IMPLEMENTACIÓN																	
Obtención de los datos	Data Extracction Team					█											
Tranfirmación de Datos	Data Transformation Team						█										
Desarrollo modelos predictivos	Development Team						█	█									
Análisis y selección de los modelos	Equipo Completo							█	█								
Prueba de concepto	Development Team								█	█							
Creación de visualizaciones	Development Team									█	█						
FASE IV: VALIDACION																	
Medición de impactos	Data Transformation Team												█	█			
Optimización de resultados	Development Team													█			
FASE V: RETROALIMENTACIÓN																	
Mejoras de los modelos	Equipo Completo																█
Inclusión de nuevas métricas (ad-hoc)	Equipo Completo																█

Tabla 7 Cronograma

8 RENTABILIDAD DEL PROYECTO



Los beneficios que se buscan con la implementación del presente proyecto no incluyen rentabilidad económica, puesto que no está en los planes actuales de la Entidad Financiera la comercialización de la solución.

Como se explicó anteriormente el objetivo de construir un modelo de este tipo es ayudar a los gobiernos de los diferentes países o regiones en la toma de decisiones, apoyándolos con herramientas de Big Data, con modelos robustos y confiables. Estos modelos permiten determinar o estudiar con más exactitud diferentes opciones para que los gobiernos evalúen su implementación y que esto permita mejorar la vida de los habitantes de la región y la población en general. Hay que resaltar que la toma de decisiones apoyada en un modelo analítico tiene un considerable aumento en el factor de éxito, puesto que estas decisiones están fundadas en análisis y datos reales.

Así mismo recordamos que el desarrollo del proyecto será implementado y financiado por una Entidad Financiera de desarrollo multilateral, cuya misión principal es mejorar la calidad de vida de la región de América Latina y el Caribe. Esta idea nace como un proyecto del departamento interno de I+D, junto con el departamento de sectores de sanidad e infraestructura.

Teniendo en cuenta el objetivo y los beneficios explicados anteriormente, podemos confirmar que el proyecto no tiene un beneficio de rentabilidad económica esperado, ya que su principal misión es la obtención de beneficios sociosanitarios y socioeconómicos, por medio del apoyo en la toma de decisiones para los diferentes gobiernos, en la evaluación y construcción de nuevos hospitales en las zonas más necesitadas.

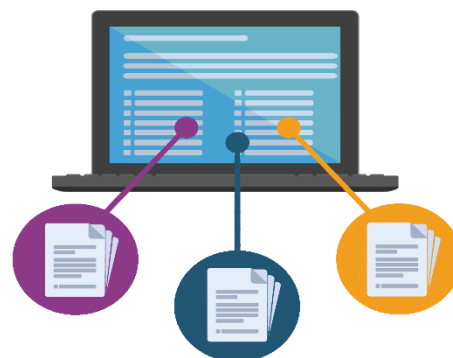
Los diferentes países para poder medir estos beneficios, principalmente beneficios como la rentabilidad social, deberán evaluar el verdadero impacto generado por el nuevo centro médico u hospital propuesto a partir del modelo de análisis de cobertura médica entregado en este proyecto.

Consideramos que esta rentabilidad social, se podrá medir a partir de 5 años después de que se encuentre activo el centro médico, ya que se contara con una gran cantidad de datos para evaluar los criterios inicialmente planteados: sociosanitarios y socioeconómicos y bajo los nuevos criterios:

- Nivel de vida de esperanza de vida de la región
- Nivel de satisfacción de atención médica
- Índice de muertes por incidencias médicas
- Tiempos de espera de atención médica
- Porcentaje de mortalidad en los últimos cinco años

Para lograr este análisis en 5 años, es necesario que se cuente con los datos necesarios y con la mejor calidad posible. Por este motivo, el acompañamiento de la Entidad Financiera no finaliza con la entrega del presente modelo a los gobiernos, sino que es necesario un acompañamiento para mejorar (o implementar de cero en algunos casos) la recolección de datos masivos, que permitan realizar mejores evaluaciones futuras, es decir, gobiernos o centros médicos (para ser más específicos) orientados al dato. Al asegurarnos la existencia de estos datos, no sólo podemos evaluar más eficientemente el impacto en 5 años, sino que también se puede mejorar el modelo para obtener información aún más precisa.

Como conclusión es importante que los gobiernos realicen esta recolección de datos, y el estudio de análisis de impacto del centro médico, luego de los 5 años de implantación, en la que se evalúe los indicadores anteriores junto con los actuales. La Entidad Financiera está evaluando la adición de este análisis o consultoría a los diferentes gobiernos, y de esta forma poder asegurarnos un impacto integral en el sector hospitalario y de salud, por medio de mediciones más certeras. Adicional también se está evaluando la construcción de un cuadro de mando, donde se pueda determinar fácilmente el impacto y los gobiernos puedan evaluar la efectividad de su inversión.



A continuación, describiremos los beneficios que un gobierno o región obtendrían, al decidir construir un centro médico u hospital en una zona óptima propuesta por el modelo de Análisis de Cobertura Hospitalaria:

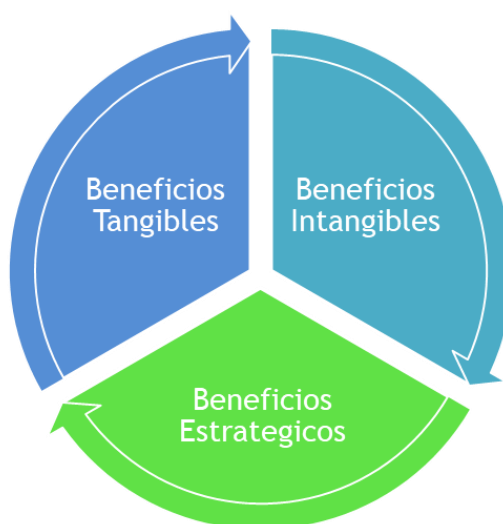


Figura 29 Tipos de Beneficios

- **Beneficios Tangibles:**
 - Generación de ingresos:
 - Para la correcta gestión de un hospital, así sea un hospital público, es necesario que este tenga ingresos. Estos se componen por ingresos públicos, y por ingresos que se generan de las visitas y los seguros de las personas que allí se atienden. Si la construcción de un hospital se realiza en áreas donde realmente se necesiten, nos aseguramos de que tendrán las mínimas visitas para que el sistema funcione.
 - Incrementa la satisfacción de los clientes.
 - Incrementa la confianza en el sistema de salud.
 - Incrementa la economía de la zona (personal para el hospital, restaurantes, transporte, etc.)
 - Reducción de costes:
 - Eliminar ineficiencias. Este proyecto minimiza el error al construir un hospital donde realmente no se necesite o donde el impacto no sea el esperado. Teniendo en cuenta que esta construcción es muy costosa, este es un beneficio muy importante para nuestro proyecto, y que indirectamente ayuda en la buena asignación de los recursos del estado.
 - Al brindarle atención a personas que hoy no tenían una buena cobertura, incentivamos que los pacientes para que acudan al hospital de forma preventiva, y

luego evitar enfermedades más avanzadas, que, aparte de poner en peligro al paciente, son mucho más costosos los tratamientos si lo comparamos con tratamientos preventivos.

- Disminuimos las reclamaciones al aumentar la cobertura.
- **Beneficios Intangibles:**
 - Definición de un modelo analítico con gran impacto social, que puede ir evolucionando o ser ejemplo para desarrollar este tipo de proyectos en otros sectores que también prioricen el impacto social y el desarrollo de países en las regiones más necesitadas.
 - Incrementar la credibilidad del gobierno.
 - Fomentar un gobierno digital.
 - Fomentar la utilización de la tecnología para el desarrollo de proyectos con impacto social.
- **Beneficios Estratégicos:**
 - Mayora habilidad para definir la estrategia del gobierno en el sector salud y de hospitales.
 - Mayora habilidad para la estrategia del gobierno en la utilización de recursos públicos.
 - La información soporta el plan estratégico para los años futuros en el desarrollo de una región.

9 BIBLIOGRAFIA

- Indicadores Básicos Situación de Salud en las Américas 2017.
https://iris.paho.org/bitstream/handle/10665.2/34330/IndBrasicos2017_spa.pdf?sequence=1&isAllowed=y
- Banco Interamericano de Desarrollo: Sector Salud, perspectiva general .
<https://www.iadb.org/es/sectores/salud/perspectiva-general>
- Los sistemas de salud en Latinoamérica y el papel del seguro privado.
<https://app.mapfre.com/ccm/content/documentos/fundacion/cs-seguro/libros/los-sistemas-de-salud-en-latinoamerica-y-el-papel-del-seguro-privado.pdf>
- Seguimiento de la cobertura sanitaria universal: Informe de monitoreo global 2017
<https://apps.who.int/iris/bitstream/handle/10665/310924/9789243513553-spa.pdf>
- Banco Mundial, entendiendo a la pobreza, Temas Salud
<https://www.bancomundial.org/es/topic/health/overview#1>
- Asamblea General de las Naciones Unidas, Resoluciones
<https://www.un.org/es/ga/67/resolutions.shtml>
- América Latina
https://es.wikipedia.org/wiki/Am%C3%A9rica_Latina
- La pobreza aumentaría en 2019 en una América Latina que lucha contra la desigualdad
<https://www.efe.com/efe/america/economia/la-pobreza-aumentaria-en-2019-una-america-latina-que-lucha-contr-desigualdad/20000011-4121519>
- ¿Cómo avanza América Latina en cobertura sanitaria?
<https://www.telesurty.net/news/avances-america-latina-cobertura-sanitaria-salud-publica-20180406-0057.html>

10 ANEXOS

- Anexo A: Entrevistas



Análisis Cobertura
Hospitalaria en Latir

- Anexo B : Código de R



Análisis Cobertura
Hospitalaria en Latir