

**ANÁLISIS SOBRE LA PRESTACIÓN DE SERVICIOS HOSPITALARIOS EN MEDELLÍN
ENTRE LOS AÑOS 2019 -2022**

INTEGRANTES

ILEANA VALENCIA ZULETA

FRANK ALEXANDER LOPERA

ANÁLISIS DE DATOS - EXPLORADOR
BOOTCAMP TALENTO TECH - MINTIC
MEDELLIN
2025

Introducción

El presente análisis se basa en los datos obtenidos de la plataforma MEDATA, correspondiente al proceso de hospitalización en el Municipio de Medellín durante el período comprendido entre los años 2019 y 2022. En este contexto, se entiende por hospitalización el ingreso de un paciente a un centro médico (hospital o clínica) con el propósito de recibir atención médica y cuidados que requieren una estancia prolongada, cama asignada y/o vigilancia en observación.

El objetivo principal de este análisis es comprender el comportamiento del uso de la planta hospitalaria por parte de los habitantes de Antioquia y de algunos pacientes provenientes de otros departamentos. Esto con el fin de encontrar patrones relacionales que permitan diseñar esquemas de atención preventiva dirigidos especialmente a las poblaciones estudiadas en la base de datos, para detectar alta demanda en el sistema de salud o de segmentos vulnerables, lo que facilitaría una atención más oportuna y efectiva.

Para ello, se utilizó una base de datos abierta proporcionada por MEDATA, que contiene información correspondiente a la vigencia 2019–2022. Esta base cuenta con un total de 839,738 registros y 24 columnas, ofreciendo un panorama amplio del sistema hospitalario de la ciudad.

La metodología empleada consistió en el uso de herramientas de análisis de datos mediante la plataforma Google Colab, utilizando el lenguaje de programación Python. Se realizaron procesos de limpieza, transformación y análisis estadístico de los datos, apoyados en librerías especializadas para garantizar resultados confiables y pertinentes.

Objetivos

1. Identificar las relaciones que pueden existir entre los datos demográficos, tipo de consulta y posible segmentación de pacientes por medio de la base de datos.
2. Evaluar los rangos de edad de los pacientes que presentaron mayor demanda de atención hospitalaria durante el periodo de análisis.
3. Identificar y diferenciar las 10 principales causas de hospitalización que concentran el mayor número de atenciones y consumo de recursos
4. Comparar las enfermedades registradas al ingreso hospitalario con los desenlaces al egreso, diferenciando entre pacientes dados de alta y fallecidos.
5. Analizar las tendencias temporales asociadas de hospitalización, con el fin de identificar las edades con mayor hospitalización y posibles complicaciones.
6. Encontrando posibles relaciones de ingresos y salidas, así como las diferentes causales de atenciones hospitalarias.

Desarrollo

Se busca tener un análisis exploratorio de los servicios de salud a nivel hospitalario en la dirección de salud de Medellín, que permita llevar a cabo procesos de mitigación o control de los servicios que se presentan a población vulnerable o que requieran mayor atención.

Para este fin se cuenta con una base de datos recolectados entre las vigencias 2019 a 2022, disponible en Medata (Corporation, 2025).

Se inicia con la carga de las librerías que nos permitirán desarrollar el análisis de la base de datos, tales como, Pandas, Numpy, Seaborn y Matplotlib.

Al cargar la base de datos directamente desde la fuente (url), se procede con la evaluación, detalle y estudio de la información extraída. Donde se identifican valores nulos, tipos de datos no numéricos sino tipo objeto, códigos de diagnóstico sin contexto o estructura, etc.

Se procede a estudiar la estructura de la base de datos para conocer un poco más a detalle la información (anexo 1. Estructura), y se carga una base de datos adicional con la clasificación internacional de enfermedades (Ministerio, 2025), con el fin de homologar los códigos cargados en la base de datos para las causas de ingreso y egreso de las hospitalizaciones.

Continuando así con el análisis exploratorio:

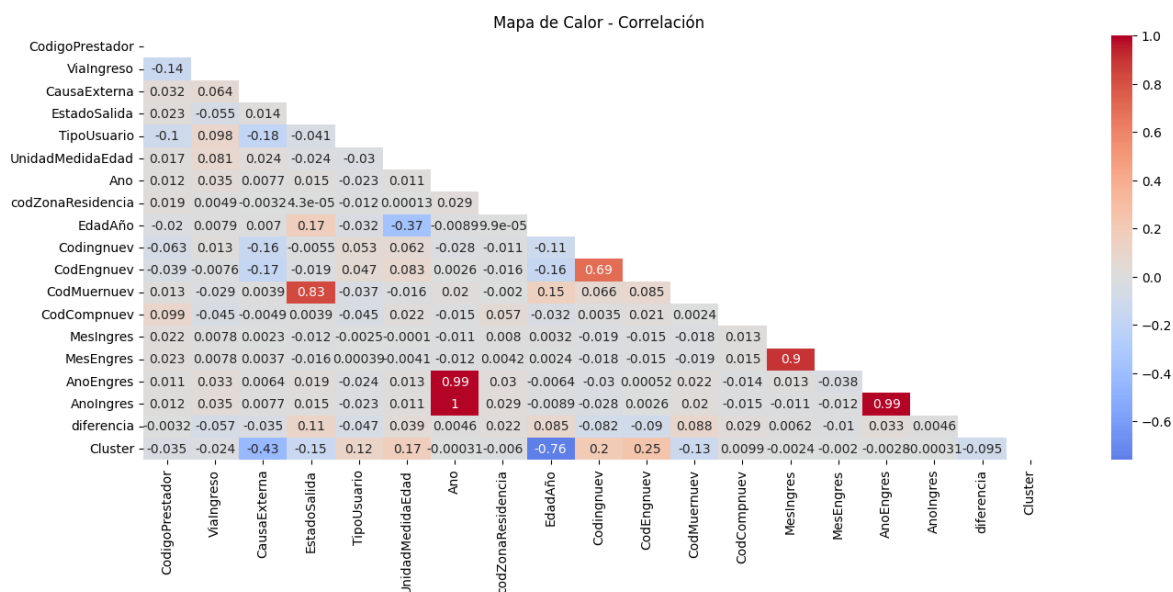
1. Copia de seguridad del DataFrame.
2. Información de tipo de dato de las columnas.
3. Identificar valores únicos de las columnas para conocer posibles errores dentro de las columnas.
4. Limpieza de datos nulos.
5. Cambio de tipos de datos en las columnas.
En este punto fue importante tener la mayor cantidad de datos números, para más adelante proceder con análisis de correlación entre los datos.
6. Agregación de columnas tipo string para una vez tener datos correlacionados, identificas a que se refería cada uno de los datos numéricos.

Análisis de la base de datos

Por medio de la librería Pandas, logramos realizar un análisis descriptivo de la base de datos, encontrando que:

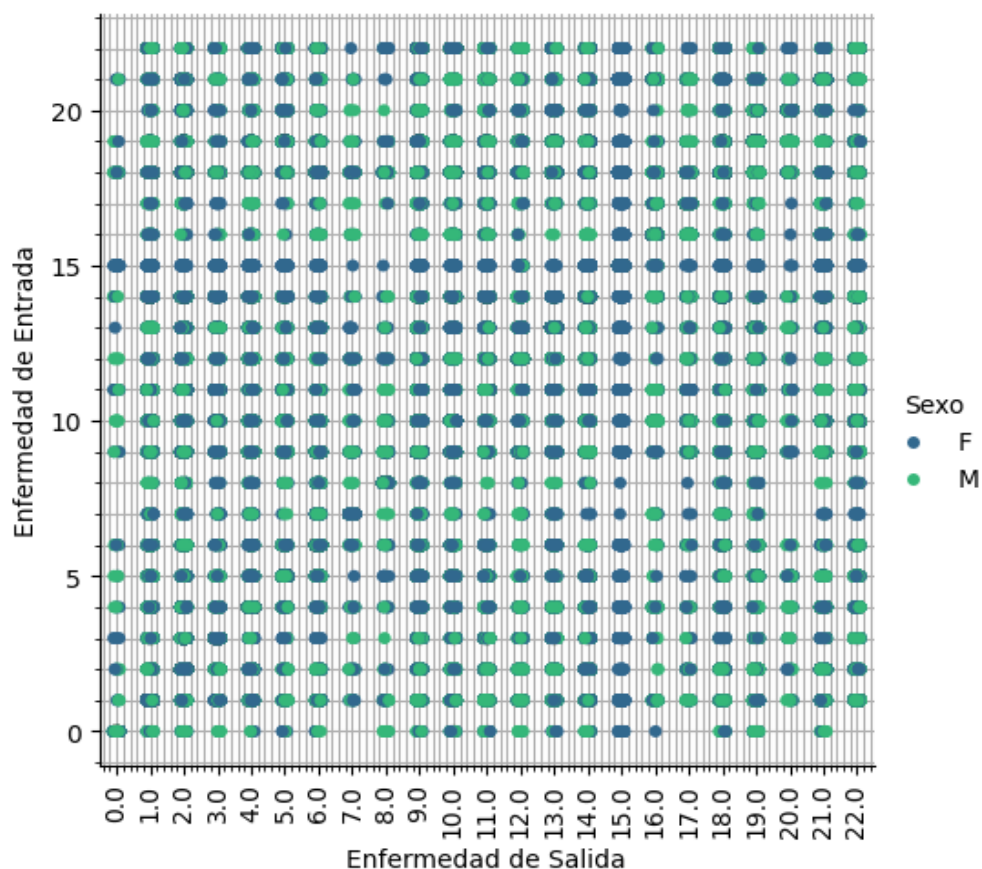
1. La edad promedio de los pacientes hospitalizados es de 46 años, con una desviación estándar de 25 años. Lo que indica que la mayor cantidad de pacientes hospitalizados se encuentra entre las edades de 20 y 70 años de edad.
2. Outlier de un paciente hospitalizado de 119 años (valor máximo).
3. Datos mínimos de 0 años de edad, que corresponden a pacientes neonatos.

Análisis de Correlación



Gráfica 1

Aunque se aplicó una conversión de variables categóricas (strings) a numéricas para calcular correlaciones, el mapa de calor evidencia que **no se encontraron relaciones fuertes entre la mayoría de las variables** de la base de datos. Las pocas correlaciones altas (valores cercanos a 1) corresponden principalmente a variables estrechamente relacionadas por su naturaleza, como Año y Año de Egreso o Mes de Ingreso y Mes de Egreso. Esto indica que no existe una correlación significativa entre variables clínicas o sociodemográficas con los códigos de diagnóstico o tipo de usuario, lo que sugiere que se requieren análisis más avanzados, como técnicas de agrupamiento (clustering), análisis multivariado o modelos de predicción, para extraer patrones ocultos en los datos hospitalarios.



Gráfica 2

En el proceso de verificación se procedió a identificar si los pacientes que ingresan a hospitalización vs su salida son idénticos o existe alguna regla que muestre afinidad con el proceso, encontrando que:

1. Los pacientes en los que no se identificó el código de ingreso y, por lo tanto, quedó en 0, no salen en ninguna causa por los conceptos de: 7. Enfermedades del ojo y sus anexos (H00-H59), 16 Ciertas afecciones originadas en el periodo perinatal (P00-P96), 17 Malformaciones congénitas, deformidades y anomalías cromosómicas (Q00-Q99), 20 Causas externas de morbilidad y de mortalidad (V01-Y98), 22 Códigos para propósitos especiales (U00-U99)
2. Los pacientes con ingresos por valoración del código 7 enfermedades del ojo y sus anexos (H00-H59) no van a salir con valoración de egreso por: 0 sin ingreso de código o 20 ciertas afecciones originadas en el periodo perinatal (P00-P96)
3. los pacientes con ingreso por valoración del código 8 enfermedades del oído y de la apófisis mastoides (H60-H95) no van a salir con valoración 16 ciertas afecciones originadas en el periodo perinatal (P00-P96) o 20 causas externas de morbilidad y de mortalidad (V01-Y98)
4. Los pacientes ingresados con el código 16, ciertas afecciones originadas en el periodo perinatal (P00-P96), no van a salir con valoración bajo los códigos 0 no ingresado u 8, Enfermedades del oído y de la apófisis mastoides (H60-H95).

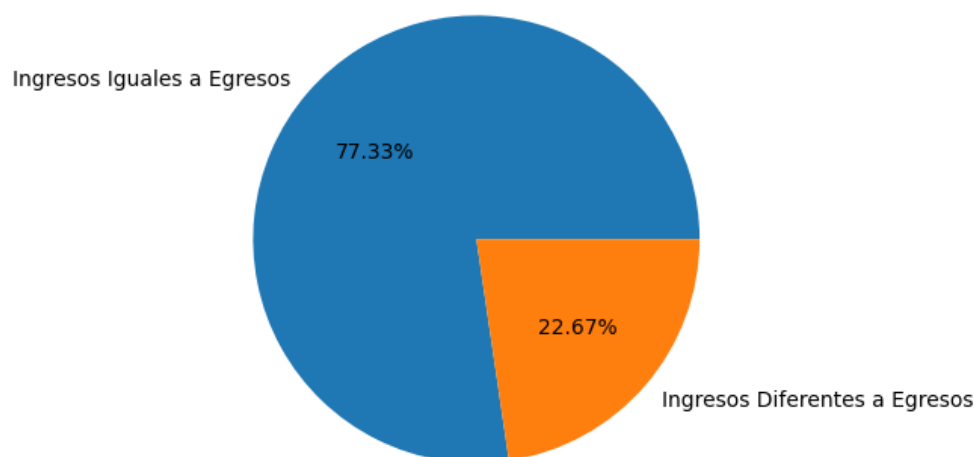
5. Los pacientes con ingresos por 17 malformaciones congénitas, deformidades y anomalías cromosómicas (Q00-Q99), 20 causas externas de morbilidad y de mortalidad (V01-Y98) y 22 códigos para propósitos especiales (U00-U99), no tendrán salida por código 0 los cuales no ingresaron código.

El análisis realizado permite evidenciar la existencia de signos clínicos y causas de ingreso que, por vacíos en el registro codificado, no presentan una asociación directa ni corresponden a un tipo específico de egreso. Esto se puede observar dentro de la asociación realizada ingreso - egreso.

Es de aclarar que existen conceptos registrados de ingresos donde en ningún momento se genera una relación con el evento de salida mostrando que estos no presentan asociación según la información.

También se puede observar que el evento 15 embarazo, parto y puerperio (O00-O99) es característico del sexo femenino y la entremezcla de ambos sexos en los otros tipos de ingreso - salida

Comparativo de cantidad de ingresos de hospitalización a diagnósticos de egreso

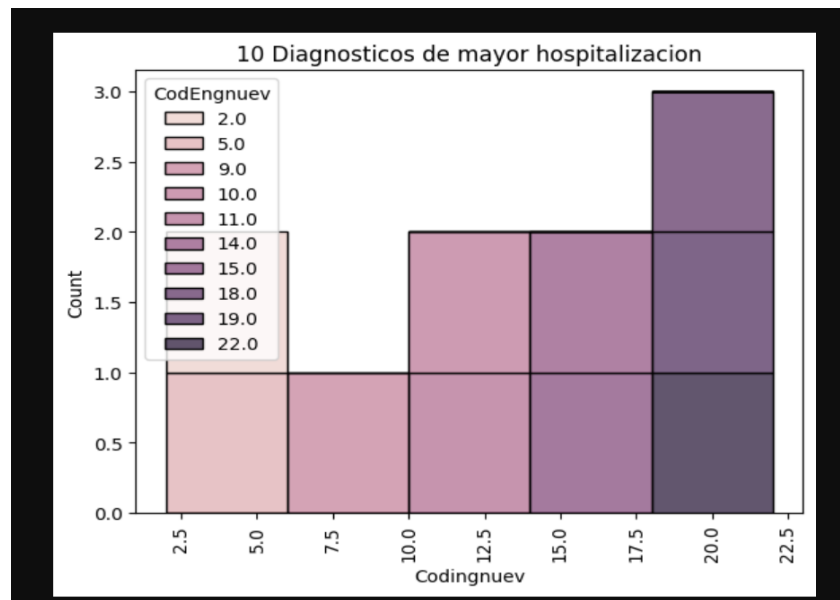


Gráfica 3

Es importante identificar si el ingreso de los pacientes tiene relación con el diagnóstico de egreso, por lo que nos damos cuenta de que un 22,6% de los pacientes son hospitalizados con un diagnóstico de ingreso muy diferente con el diagnóstico con el que se le da de alta.

Codingnuev	CodEngnuev	
15.0	15.0	126733
10.0	10.0	65953
19.0	19.0	57566
9.0	9.0	56359
14.0	14.0	49197
2.0	2.0	43584
11.0	11.0	43387
18.0	18.0	32663
22.0	22.0	30696
5.0	5.0	27731

Tabla 1



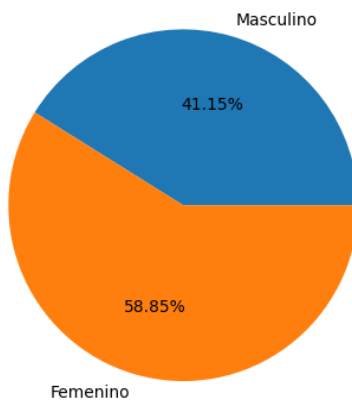
Gráfica 4

Se puede observar los 10 códigos de enfermedades de mayor hospitalización, así como su relación con egresos generados donde podemos observar que en estas está ampliamente relacionados con la misma ingreso-egreso siendo:

1. 15- Embarazo, parto y puerperio (O00-O99).
2. 10- Enfermedades del sistema respiratorio (J00-J99)
3. 19-Traumatismos, envenenamiento y otras consecuencias de causas externas (S00-T98).)
4. 9-Enfermedades del sistema circulatorio (I00-I99).
5. 14- Enfermedades del sistema genitourinario (N00-N99)
6. 2-Tumores [NEOPLASIAS] (C00-D48)
7. 11-Enfermedades del sistema digestivo (K00-K93)
8. 18- Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte (R00-R99).

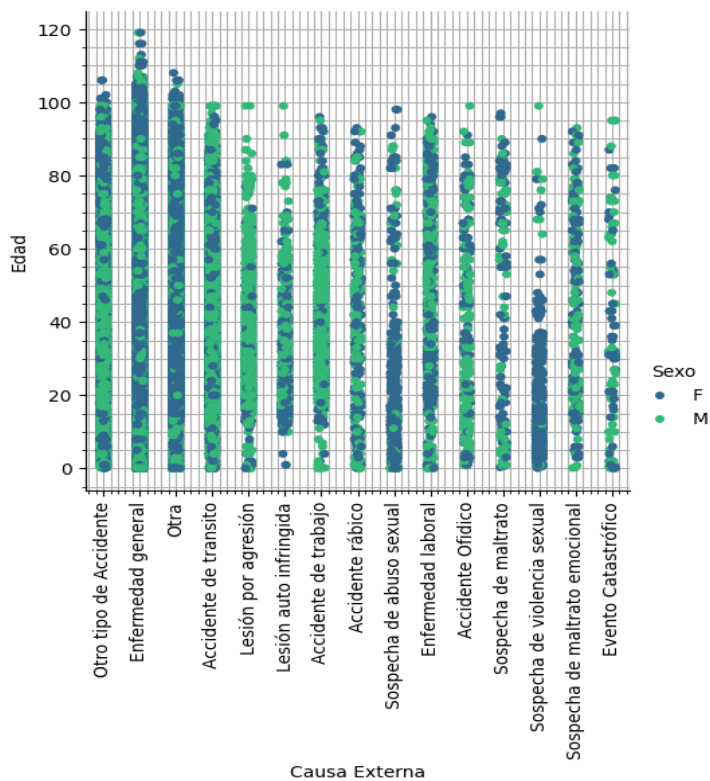
9. 22- Códigos para propósitos especiales (U00-U99
10. 5-Trastornos mentales y del comportamiento (F00-F99).

Comparativo de hospitalizaciones de acuerdo al genero



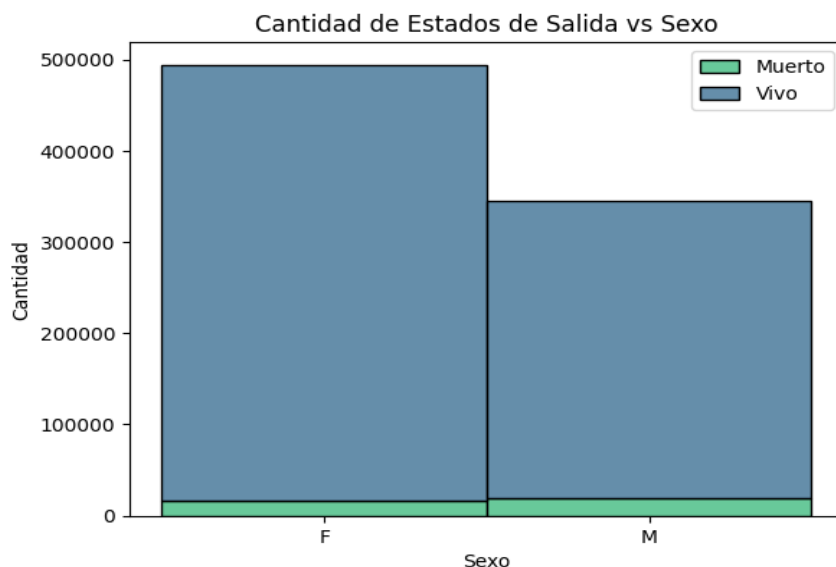
Gráfica 5

Así mismo, las mujeres son quienes mayor proporción de hospitalización tienen, pero esto se entiende más adelante, debido a que son quienes más consultan con relación a los hombres (Gráfica 7).



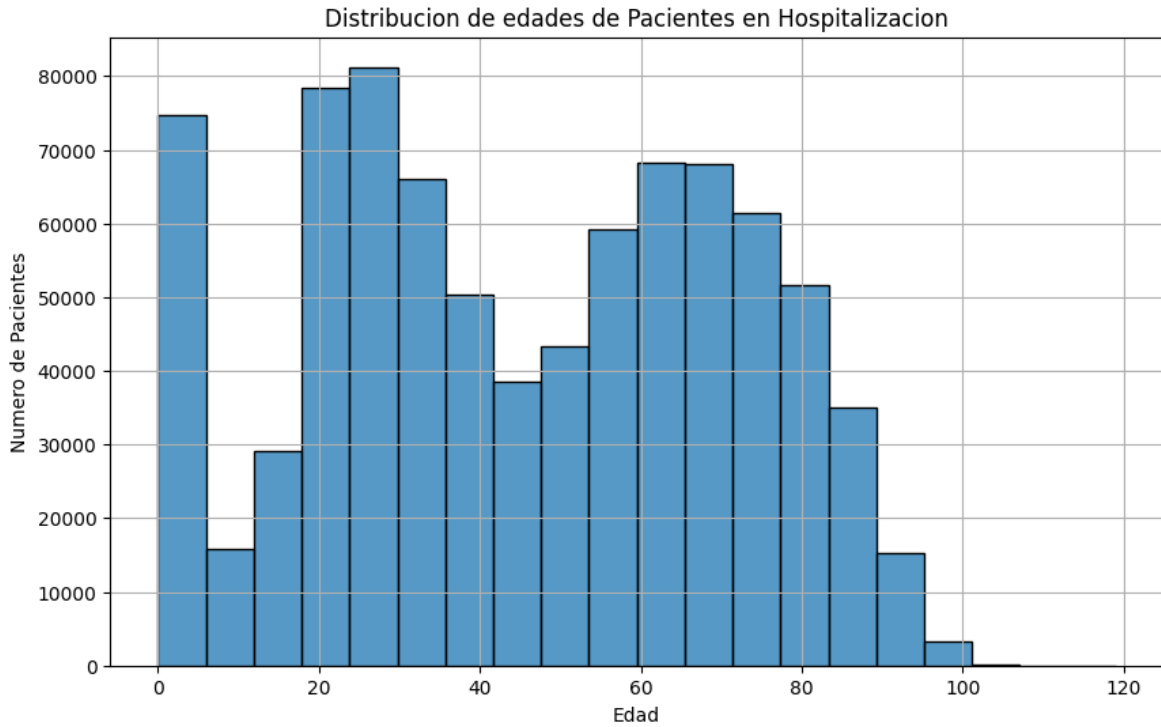
Gráfica 6

La causa externa es aquella por la que los pacientes llegan al hospital y son hospitalizados inicialmente, en esta gráfica podemos ver una comparativa de la causa externa de acuerdo con el género del paciente. Se identifica que los hombres en mayor medida son hospitalizados por lesiones por agresión, accidentes de trabajo y accidentes de tránsito. Mientras las mujeres entre los 0 y 40 años son atendidas por sospecha de abuso sexual y por sospecha de violencia sexual.



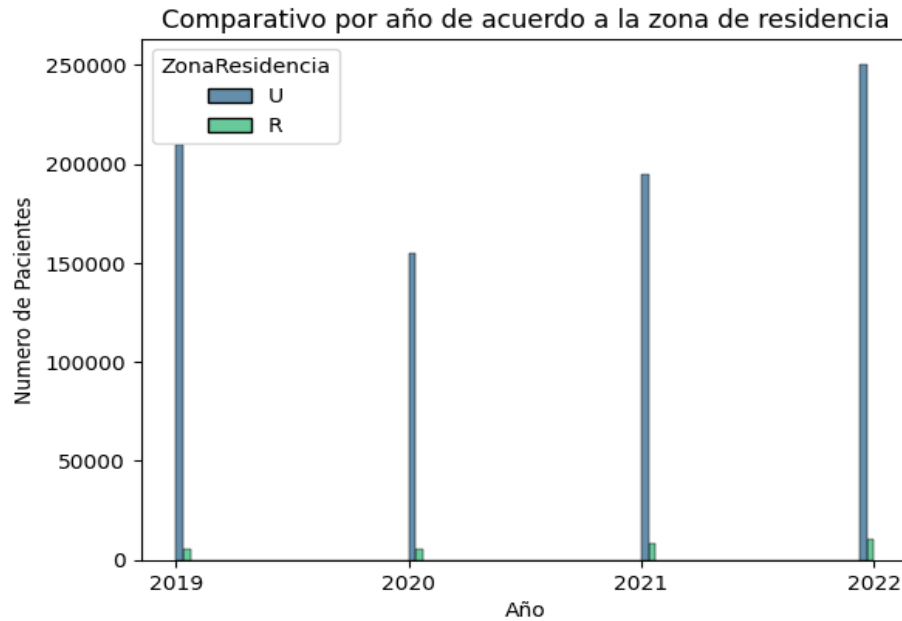
Gráfica 7

En esta gráfica podemos encontrar el estado de salida de los pacientes de acuerdo con el género, donde se nota claramente que las mujeres son quienes más consultan y son hospitalizadas, pero hay algo que notar y es que, los hombres aun siendo en menor medida hospitalizados, son quienes más salidas de muerte después de hospitalización presentan, por lo que se puede inferir que, los hombres hospitalizados llegan en estados críticos que pueden generar complicaciones y posterior a ello la muerte.



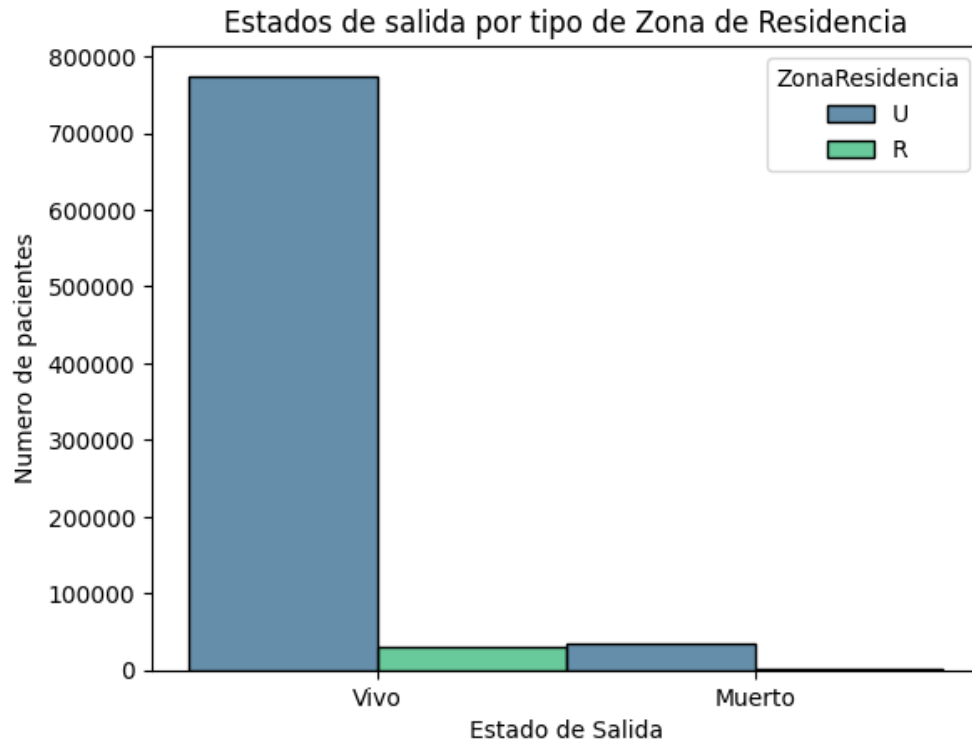
Gráfica 8

El histograma de distribución por edad muestra que los grupos con mayor número de pacientes hospitalizados se concentran en recién nacidos (neonatos) y en el rango de 20 a 30 años, seguidos por un repunte sostenido a partir de los 60 años en adelante, alcanzando un nuevo pico entre los 60 y 70 años. Esto indica que los extremos de la vida, recién nacidos y adultos mayores concentran una alta proporción de hospitalizaciones, probablemente por causas perinatales y enfermedades crónicas, respectivamente. La alta frecuencia entre los 20 y 30 años podría estar relacionada con causas externas como accidentes o partos. En contraste, el grupo de edad entre los 35 y 50 años muestra una menor frecuencia de hospitalización, lo cual podría reflejar un periodo de relativa estabilidad en salud.



Gráfica 9

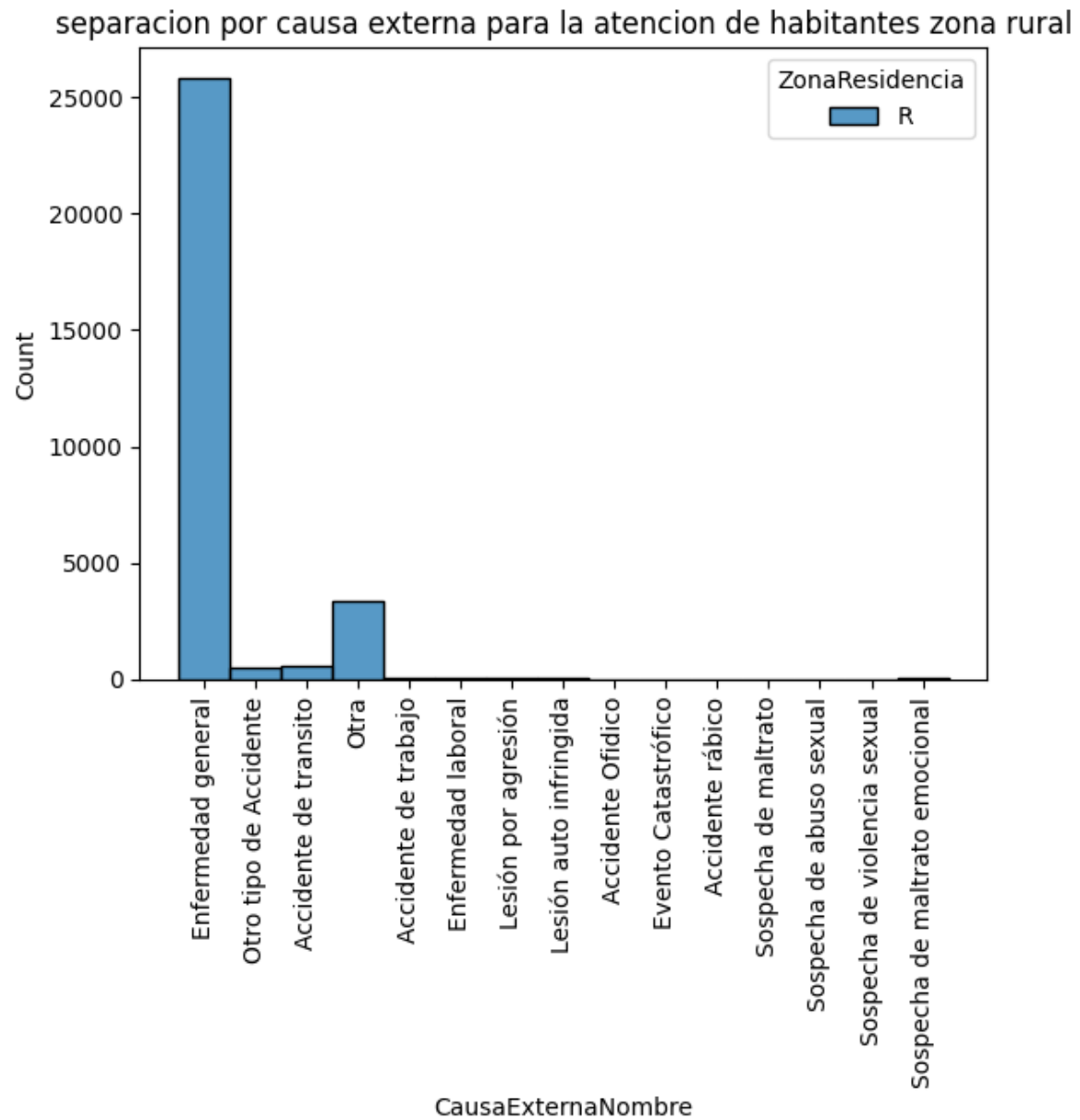
El gráfico comparativo por año y zona de residencia evidencia que en el año 2020 se presentó una notable disminución en el número de pacientes hospitalizados, tanto en zonas urbanas (U) como rurales (R), en comparación con 2019. Esta caída coincide con el periodo más crítico de la pandemia por COVID-19, durante el cual se restringió el acceso a los servicios de salud no urgentes y se prioriza la atención ambulatoria y domiciliaria. A partir del 2021, se observa una tendencia al alza en la hospitalización, superando incluso los niveles de 2019 en el 2022, lo que podría reflejar tanto la reactivación del sistema de salud como la acumulación de casos no atendidos durante la emergencia sanitaria. De forma consistente, se observa que la mayor parte de los pacientes hospitalizados provienen de zonas urbanas, lo cual resalta una brecha significativa en el acceso o la necesidad de servicios hospitalarios entre zonas urbanas y rurales.



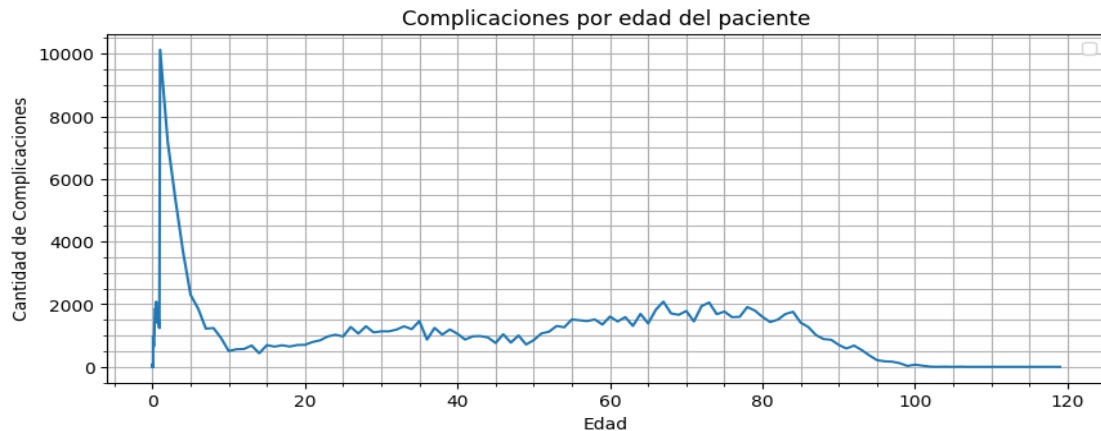
Gráfica 10

El gráfico muestra que la mayoría de los pacientes hospitalizados salieron con vida, tanto en zonas urbanas como rurales. La mayoría de los egresos “vivos” vienen de áreas urbanas (U), algo que tiene sentido pensando que hay mayor cantidad de gente viviendo en ciudades y el mejor acceso a servicios hospitalarios. En cuanto a las zonas rurales, la baja proporción de egresos puede deberse a las dificultades para acceder a servicios hospitalarios, lo que afecta tanto la atención médica oportuna como el registro completo de los casos.

Se podría interpretar que, los casos de hospitalización de los pacientes que vienen de la zona rural, aparentemente pueden ser casos graves, pero no es así, como se nota (Gráfica 11), los pacientes de zonas rurales en hospitalización, son pacientes que consultan por una causa de tipo general como pueden ser: infecciones graves como neumonía, infecciones urinarias complicadas, infecciones de piel y tejidos blandos como celulitis o fascitis, e infecciones del sistema nervioso central como meningitis o encefalitis. Además, enfermedades crónicas como la EPOC reagudizada, insuficiencia cardíaca, o complicaciones de la diabetes.



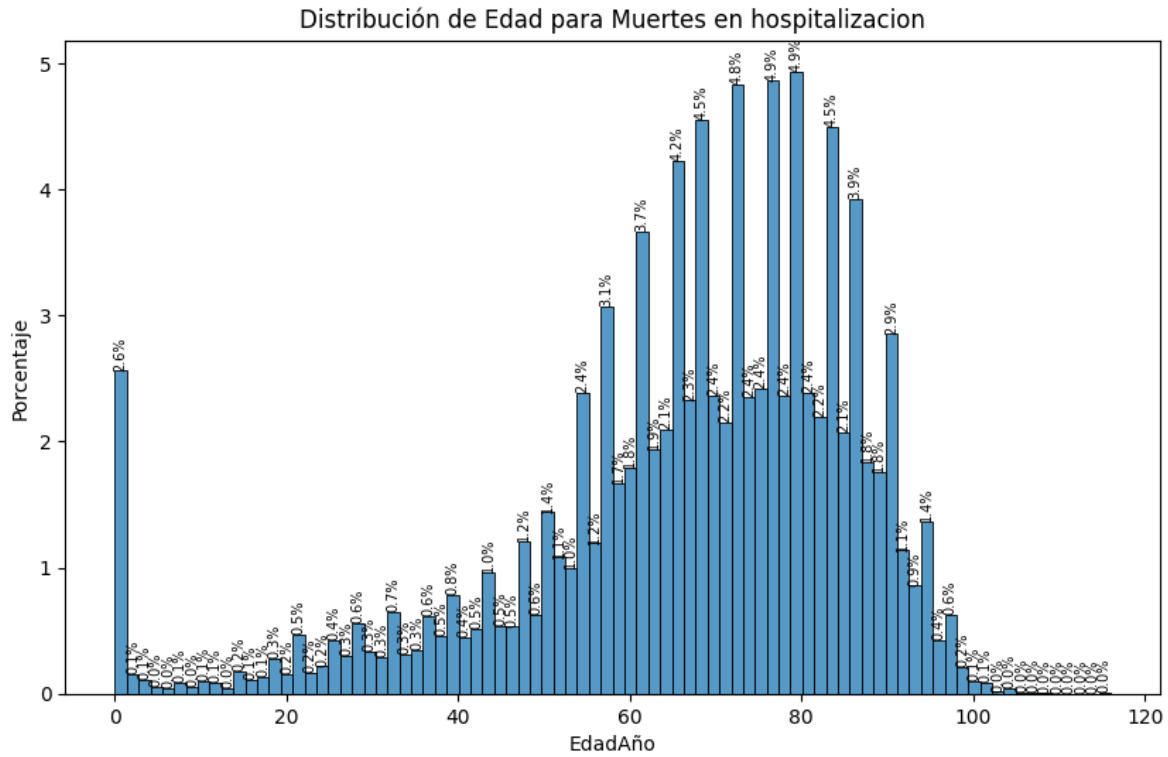
Gráfica 11



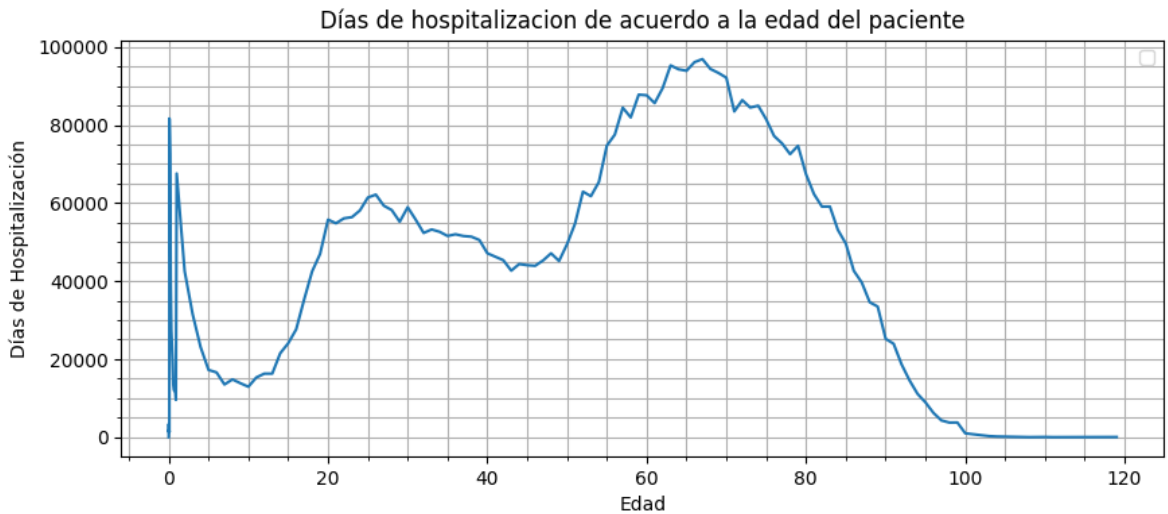
Gráfica 12

La gráfica 12 muestra que la mayor cantidad de complicaciones durante la hospitalización se presenta en pacientes neonatos y menores de 5 años, alcanzando su punto máximo en los primeros días de vida. Este comportamiento puede atribuirse a factores como partos prematuros, malformaciones congénitas, infecciones neonatales o condiciones perinatales críticas, que requieren intervenciones hospitalarias complejas. A partir de los 5 años, la frecuencia de complicaciones disminuye considerablemente y se mantiene relativamente estable en niveles bajos durante la vida adulta, con un leve incremento entre los 60 y 80 años, posiblemente relacionado con el aumento de enfermedades crónicas y degenerativas propias del envejecimiento. Finalmente, la frecuencia cae en mayores de 90 años, lo cual podría deberse a un menor volumen poblacional en ese grupo o de eventos clínicos.

Cabe resaltar que, la distribución de pacientes hospitalizados que son egresados en un tipo de estado “muerto” tiene una alta relación a las edades en las que las hospitalizaciones se complican, sin embargo, las edades tempranas siendo, las edades con mayores complicaciones son egresados en tipo de estado “vivo”, las edades adultas entre 60 y 90 años son las más propensas a pasar a tipo de estado “muerto” en su egreso cuando la hospitalización se complica (Gráfica 13).

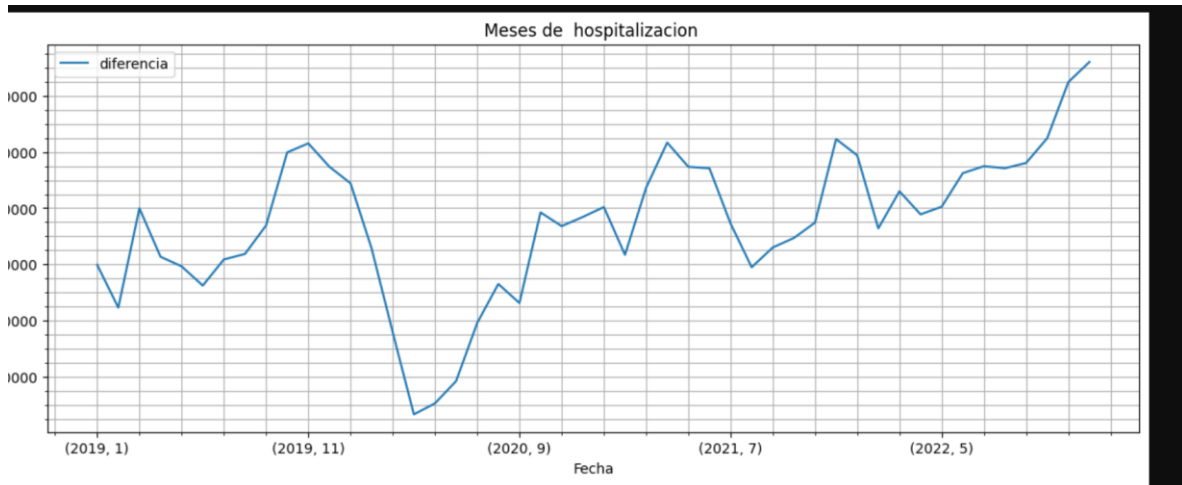


Gráfica 13



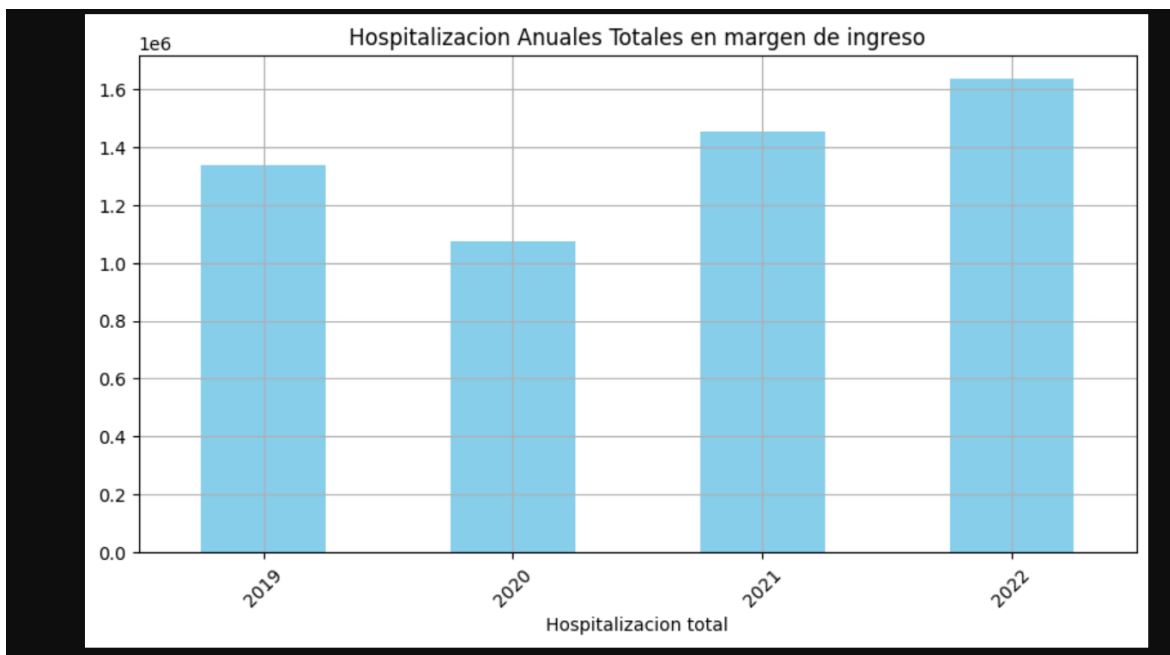
Gráfica 14

Esta gráfica nos permite conocer qué rango de edades permanecen más tiempo en hospitalización, entendiendo que los que se encuentran entre los 60 y 70 años presentan un mayor tiempo en los centros hospitalarios. Probablemente está relacionado con enfermedades crónicas, degenerativas o eventos cardiovasculares.



Gráfica 15

La gráfica de hospitalizaciones por meses muestra una tendencia clara, los días de hospitalización vienen aumentando de manera constante desde 2019 hasta 2022 (Gráfica 16), con varios picos. Los momentos más críticos se ven en noviembre de 2019, julio de 2020, junio de 2021 y finales de 2022, siendo diciembre de ese último año el más alto registrado. Estos repuntes podrían estar ligados a temporadas con más enfermedades respiratorias, los efectos de la pandemia y cómo esta fue afectando cada vez más la red hospitalaria. También resalta una caída fuerte en el primer semestre de 2020, probablemente por las restricciones de movilidad y la baja en consultas no urgentes al comienzo de la pandemia. En general, la gráfica deja ver una presión creciente sobre los servicios de salud con el paso del tiempo.



Gráfica 16

Conclusiones

1. Los recién nacidos y niños menores de 5 años son los que más complicaciones presentan durante la hospitalización, incluso más que los adultos mayores. Esto marca un patrón crítico en la atención de la primera infancia, posiblemente relacionado con condiciones congénitas, partos prematuros o infecciones tempranas.
2. Por su parte, los adultos jóvenes y de mediana edad tienen una baja frecuencia de complicaciones, lo que indica que este grupo genera menos presión sobre los recursos hospitalarios en casos graves. Este comportamiento podría ayudar a segmentar a los pacientes según su perfil de riesgo clínico.
3. Entre los 60 y 80 años se nota un pequeño aumento en las complicaciones, que parece coincidir con hospitalizaciones más prolongadas vistas en otros gráficos. Es probable que esto se relacione con enfermedades crónicas y degenerativas. En general, el gráfico sugiere una conexión entre la edad, la gravedad clínica y el desenlace hospitalario (como egreso vivo o fallecido).

Bibliografía

- Alcaldía de Medellín. (s. f.). *Datos sobre hospitalización en Medellín (2019-2022)*. Medata. Recuperado el 13 de julio de 2025, de <https://medata.gov.co/node/16878>
- Ministerio de Salud y Protección Social. (s. f.). *CIE-10 y CIE-11* [PDF]. Recuperado el 13 de julio de 2025, de <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/IA/SSA/cie10-cie11.pdf>
- Organización Mundial de la Salud. (2022). *Informe mundial sobre la salud: Mejorar la prestación de servicios de salud para todos*. Organización Mundial de la Salud. <https://www.who.int/publications/i/item/9789240061734>
- Ministerio de Salud y Protección Social. (2021). *Análisis de situación de salud (ASIS) nacional 2020-2021*. <https://www.minsalud.gov.co/salud/publica/epidemiologia/Paginas/asis.aspx>