

Administración de Medicamentos a Recién Nacidos

Renato Jair Rojas Dominguez, Crhistian Mark Montenegro Valderrama

2024-09-01

Tabla de Contenidos

1. Introducción:	2
2. Resumen:	3
3. Objetivo:	3
4. Metodología:	3
4.1. Obtención de los datos	3
4.2. Limpieza y preparación de datos	3
4.3. Formato de dataset	3
4.4. Limpieza de columnas	4
4.5. Exportar dataset en CSV codificado UTF-8	4
4.6. Importar dataset de trabajo	4
4.7. Conteo de valores NA	4
4.8. Manejo de NA en variable PESO	5
4.9. Manejo de NA en variables FECINGHOSP, FECALTAHOSP Y DIAS_HOSP	6
4.10. Manejo de NA en variable PRESENTACION	7
4.11. Imputación de outliers de Peso por Género	7
4.12. Imputación de outliers CANTIDAD, VALOR_BRUTO, VALOR_NETO y NIVEL_EESS	8
4.13. Exportar el dataset limpio cómo archivo CSV con la codificación UTF-8	9
5. Análisis Exploratorio de Datos (EDA)	9
5.1. FECHA_CORTE	9
5.2. CODIGO_ANONIZADO	9
5.3. EDAD_DIAS	11
5.4. FECATENCION	17
5.5. SEXO	19
5.7. DEPARTAMENTO	23
5.8. PROVINCIA	25
5.9. DISTRITO	27

5.10. UBIGEO.....	29
5.11. NIVEL_EESS	31
5.12. RENIPRESS	32
5.13. NOMBRE_EESS	33
5.14. COD_SERVICIO.....	34
5.15. SERVICIO	36
5.16. FECINGHOSP	37
5.17. FECALTAHOSP	38
5.18. DIAS_HOSP.....	39
5.19. CODDIA	44
5.20. C10_NOMBRE.....	45
5.21. COD_MEDICAMENTO.....	47
5.22. NOMBRE_MEDICAMENTO.....	48
5.23. PRESENTACION.....	50
5.24. FORMA_FARMACEUTICA.....	51
5.25. CANTIDAD.....	52
5.26. VALOR_BRUTO	53
5.27. VALOR_NETO.....	56
6. Análisis.....	58
6.1. Análisis Geográfico.....	58
7. Principales hallazgos	60
8. Limitaciones	61
9. Conclusiones	61
10. Recomendaciones.....	62

1. Introducción:

Análisis de la Administración de Medicamentos en Recién Nacidos: Un Estudio de Datos de Instituciones de Salud Pública en Perú.

Este proyecto de análisis de datos se centra en el estudio de un extenso dataset relacionado con la administración de medicamentos a recién nacidos en diversas instituciones de salud pública en Perú. El propósito del análisis es explorar y comprender los patrones de administración de medicamentos, los diagnósticos asociados y otros factores relevantes en el cuidado neonatal.

El dataset, denominado “Medicamentos”, contiene un total de 1,240,605 registros y 27 columnas.

2. Resumen:

Hacer resumen

3. Objetivo:

Explica brevemente el propósito del análisis (por ejemplo, entender los patrones de administración de medicamentos y su relación con los diagnósticos neonatales).

4. Metodología:

Menciona las técnicas de análisis y las herramientas utilizadas.

4.1. Obtención de los datos

Los datos fueron obtenidos de la “Plataforma Nacional de Datos Abiertos” <https://www.datosabiertos.gob.pe/> en la categoría Salud con el nombre “**Consumo de medicamentos en recién nacidos asegurados al Seguro Integral de Salud - [SIS]**” <https://www.datosabiertos.gob.pe/dataset/consumo-de-medicamentos-en-reci%C3%A9n-nacidos-asegurados-al-seguro-integral-de-salud-sis>

4.2. Limpieza y preparación de datos

```
Medicamentos_raw <- read_csv("Medicamentos_Recien_Nacidos_SIS.csv", locale = locale(encoding = "UTF-8"))
```

4.3. Formato de dataset

El código convierte las columnas de texto del dataset `Medicamentos_raw` de la codificación **Latin-1** a **UTF-8**. Esta conversión es crucial para corregir problemas con tildes y otros caracteres especiales, garantizando que toda la información textual (como nombres de medicamentos y localidades) se visualice correctamente. Esto asegura la integridad de los datos antes de llevar a cabo análisis más avanzados, evitando errores en la interpretación y presentación de los resultados.

```
Medicamentos_raw <- read_csv("Medicamentos_Recien_Nacidos_SIS.csv", locale = locale(encoding = "UTF-8"))
```

Conversión de tipos de datos Una vez corregida la codificación de los caracteres, se procedió a convertir los tipos de variables en el dataset para asegurar que cada columna tuviera el formato más adecuado para el análisis.

```

Medicamentos_raw <- Medicamentos_raw %>% mutate( # Convertir las fechas a tipo
Date FECHA_CORTE = as.Date(as.character(FECHA_CORTE), format = "%Y%m%d"),
FECATENCION = as.Date(as.character(FECATENCION), format = "%Y%m%d"),
FECINGHOSP = as.Date(as.character(FECINGHOSP), format = "%Y%m%d"),
FECALTAHOSP = as.Date(as.character(FECALTAHOSP), format = "%Y%m%d"), #
Convertir variables de texto a factor SEXO = as.factor(SEXO), DEPARTAMENTO =
as.factor(DEPARTAMENTO), PROVINCIA = as.factor(PROVINCIA), DISTRITO =
as.factor(DISTRITO), UBIGEO = as.factor(UBIGEO), NIVEL_EESS =
as.factor(NIVEL_EESS), RENIPRESS = as.factor(RENIPRESS), NOMBRE_EESS =
as.factor(NOMBRE_EESS), COD_SERVICIO = as.factor(COD_SERVICIO), SERVICIO =
as.factor(SERVICIO), CODDIA = as.factor(CODDIA), C10_NOMBRE =
as.factor(C10_NOMBRE), COD_MEDICAMENTO = as.factor(COD_MEDICAMENTO),
NOMBRE_MEDICAMENTO = as.factor(NOMBRE_MEDICAMENTO), PRESENTACION =
as.factor(PRESENTACION), FORMA_FARMACEUTICA =
as.factor(FORMA_FARMACEUTICA), # Variables numéricas EDAD_DIAS =
as.integer(EDAD_DIAS), PESO = as.numeric(PESO), DIAS_HOSP =
as.integer(DIAS_HOSP), CANTIDAD = as.integer(CANTIDAD), VALOR_BRUTO =
as.numeric(VALOR_BRUTO), VALOR_NETO = as.numeric(VALOR_NETO), # Variables
carácter CODIGO_ANONIMIZADO = as.character(CODIGO_ANONIMIZADO), )

```

4.4. Limpieza de columnas

Limpieza del código ubigeo dentro de los datos en las columnas DEPARTAMENTO, PROVINCIA y DISTRITO

```

Medicamentos_raw <- Medicamentos_raw %>% mutate( DEPARTAMENTO =
factor(sub("^\\d+\\s+", "", as.factor(DEPARTAMENTO))), PROVINCIA =
factor(sub("^\\d+\\s+", "", as.factor(PROVINCIA))), DISTRITO = factor(sub("^\\d+\\s+", "",
as.factor(DISTRITO))) )

```

4.5. Exportar dataset en CSV codificado UTF-8

Con el fin de tener el csv limpio y con el formato correcto para ser limpiado y procesado.

```
fwrite(Medicamentos_raw, "Medicamentos.csv", bom = TRUE)
```

4.6. Importar dataset de trabajo

Éste dataset es el que será limpiado y utilizado para el análisis

```

Medicamentos <- read_csv("Medicamentos.csv", # Mantiene UBIGEO locale =
locale(encoding = "UTF-8")) # Falta dar formato de tildes

```

4.7. Conteo de valores NA

Se cuenta la cantidad de valores NA por cada variable para comenzar a procesarlos.

Contar los valores NA por columna

```
na_count <- sapply(Medicamentos, function(x) sum(is.na(x)))
```

Calcular el porcentaje de valores NA por columna

```
na_percentage <- (na_count / nrow(Medicamentos)) * 100
```

Filtrar solo las columnas con valores faltantes

```
na_columns <- na_count[na_count > 0] na_percentage_filtered <-  
na_percentage[na_percentage > 0]
```

Mostrar el número de NA por variable

```
for (col_name in names(na_columns)) { cat("Columna:", col_name, "- Número de  
valores faltantes:", na_columns[col_name], "") }
```

Mostrar el porcentaje de NA por variable

```
for (col_name in names(na_percentage_filtered)) { cat("Columna:", col_name, "-  
Porcentaje de valores faltantes:", round(na_percentage_filtered[col_name], 2), "%") }
```

4.8. Manejo de NA en variable PESO

Se determina que los valores de CODDIA, SEXO, COD_MEDICAMENTO y COD_SERVICIO cómo valores importantes para determinar el peso de los Recién Nacidos.

Se imputa los valores NA de la variable PESO mediante mediana por CODDIA, COD_MEDICAMENTO y COD_SERVICIO, PESO y COD_SERVICIO, SEXO y COD_SERVICIO

```
# Calcular la mediana de PESO por CODDIA, COD_MEDICAMENTO y COD_SERVICIO
```

```
# Calcular la mediana de PESO por CODDIA, COD_MEDICAMENTO y COD_SERVICIO
```

```
mediana_peso_por_grupo <- Medicamentos[, .(PESO_MEDIANA = median(PESO, na.rm  
= TRUE)), by = .(CODDIA, COD_MEDICAMENTO, COD_SERVICIO)]
```

```
# Imputar los valores NA en PESO con la mediana correspondiente
```

```
Medicamentos[is.na(PESO), PESO := mediana_peso_por_grupo[.SD, PESO_MEDIANA,  
on = .(CODDIA, COD_MEDICAMENTO, COD_SERVICIO)]]
```

```
# Imputar los valores NA restantes con la mediana dentro de grupos más amplios
```

```
mediana_peso_por_grupo_amplio <- Medicamentos[, .(PESO_MEDIANA =  
median(PESO, na.rm = TRUE)), by = CODDIA] Medicamentos[is.na(PESO), PESO :=  
mediana_peso_por_grupo_amplio[.SD, PESO_MEDIANA, on = .(CODDIA)]]
```

```
# Calcular la mediana de PESO por SEXO y COD_SERVICIO
```

```
mediana_peso_porsexo_servicio <- Medicamentos[, .(PESO_MEDIANA =  
median(PESO, na.rm = TRUE)), by = .(SEXO, COD_SERVICIO)]
```

```
# Imputar los valores NA restantes en PESO con la mediana correspondiente a la
combinación de SEXO y COD_SERVICIO
```

```
Medicamentos[is.na(PESO), PESO := mediana_peso_porsexo_servicio[.SD,
PESO_MEDIANA, on = .(SEXO, COD_SERVICIO)]]
```

```
# Verificar si quedan valores NA en PESO
```

```
na_peso_despues <- sum(is.na(Medicamentos$PESO)) cat("Cantidad de valores NA en
PESO después de la imputación:", na_peso_despues, "")
```

4.9. Manejo de NA en variables FECINGHOSP, FECALTAHOSP Y DIAS_HOSP

Se calcula los días hospitalizados mediante la diferencia de la fecha de alta y fecha de ingreso de hospitalización, para las siguientes valores NA se imputó mediante CODDIA, SEXO y EDAD_DIAS

```
# Calcular Dias Hospitalizados
```

```
Medicamentos <- Medicamentos %>% mutate( DIAS_HOSP = ifelse( is.na(DIAS_HOSP),
as.integer(as.Date(FECALTAHOSP) - as.Date(FECINGHOSP)), DIAS_HOSP) )
```

```
# Imputación de Dias Hospitalizados por CODDIA, SEXO y EDAD_DIAS
```

```
Medicamentos <- Medicamentos %>% group_by(CODDIA, SEXO, EDAD_DIAS) %>%
mutate(DIAS_HOSP = ifelse(is.na(DIAS_HOSP), as.integer(median(DIAS_HOSP, na.rm =
TRUE))), DIAS_HOSP)) %>% ungroup()
```

```
# Calcular la mediana de DIAS_HOSP por CODDIA y SEXO
```

```
Medicamentos <- Medicamentos %>% group_by(CODDIA, SEXO) %>%
mutate(DIAS_HOSP = ifelse(is.na(DIAS_HOSP), as.integer(median(DIAS_HOSP, na.rm =
TRUE))), DIAS_HOSP)) %>% ungroup()
```

```
# Calcular la mediana de DIAS_HOSP por SEXO y EDAD_DIAS Medicamentos <-
Medicamentos %>% group_by(SEXO, EDAD_DIAS) %>% mutate(DIAS_HOSP =
ifelse(is.na(DIAS_HOSP), as.integer(median(DIAS_HOSP, na.rm = TRUE))),
DIAS_HOSP)) %>% ungroup()
```

```
# Calcular Fecha de ingreso de hospitalización
```

```
Medicamentos <- Medicamentos %>% mutate( FECINGHOSP =
as.Date(ifelse(is.na(FECINGHOSP) & !is.na(FECALTAHOSP) & !is.na(DIAS_HOSP),
as.Date(FECALTAHOSP) - DIAS_HOSP, FECINGHOSP), origin = "1970-01-01"))
```

```
# Imputar los valores NA en FECALTAHOSP con FECATENCION
```

```
Medicamentos$FECALTAHOSP[is.na(Medicamentos$FECALTAHOSP)] <-
Medicamentos$FECATENCION[is.na(Medicamentos$FECALTAHOSP)]
```

```
# Calcular Dias Hospitalizados
```

```

Medicamentos <- Medicamentos %>% mutate( DIAS_HOSP = ifelse( is.na(DIAS_HOSP),
as.integer(as.Date(FECALTAHOSP) - as.Date(FECINGHOSP)), DIAS_HOSP) )

# Verificar la cantidad de valores NA en DIAS_HOSP, FECINGHOSP y FECALTAHOSP
na_dias_hosp_despues <- sum(is.na(Medicamentos$DIAS_HOSP))
na_fecinghosp_despues <- sum(is.na(Medicamentos$FECINGHOSP))
na_fecaltahosp_despues <- sum(is.na(Medicamentos$FECALTAHOSP))

# Mostrar los resultados

```

4.10. Manejo de NA en variable PRESENTACION

Se imputo mediante la moda por medio del código de medicamento y forma farmaceutica.

```

# Convertir PRESENTACION a caracter temporalmente

Medicamentos <- Medicamentos %>% mutate(PRESENTACION =
as.character(PRESENTACION))

# Imputar por la moda dentro de cada grupo definido por COD_MEDICAMENTO y
FORMA_FARMACEUTICA

Medicamentos <- Medicamentos %>% group_by(COD_MEDICAMENTO,
FORMA_FARMACEUTICA) %>% mutate(PRESENTACION =
ifelse(is.na(PRESENTACION), ifelse(sum(!is.na(PRESENTACION)) > 0,
names(sort(table(PRESENTACION), decreasing = TRUE))[1], NA), PRESENTACION))
%>% ungroup()

# Para los valores que aún sean NA, reemplazarlos con la moda global

moda_global <- names(sort(table(Medicamentos$PRESENTACION), decreasing =
TRUE))[1] Medicamentos$PRESENTACION[is.na(Medicamentos$PRESENTACION)] <-
moda_global

# Convertir PRESENTACION de vuelta a factor

Medicamentos$PRESENTACION <- factor(Medicamentos$PRESENTACION)

```

4.11. Imputación de outliers de Peso por Género

Se imputaron valores outliers ya que habían datos de peso registrados en 0kg y 10 kg mediante medianas.

```

# Calcular los cuartiles y el IQR por género

cuartiles_masculino <- quantile(Medicamentos$PESO[Medicamentos$SEXO ==
“MASCULINO”], probs = c(0.25, 0.75), na.rm = TRUE) IQR_masculino <-
cuartiles_masculino[2] - cuartiles_masculino[1] limite_inferior_masculino <-

```

```

cuartiles_masculino[1] - 1.5 * IQR_masculino limite_superior_masculino <-
cuartiles_masculino[2] + 1.5 * IQR_masculino cuartiles_femenino <-
quantile(MedicamentosPESO[MedicamentosSEXO == "FEMENINO"], probs = c(0.25,
0.75), na.rm = TRUE) IQR_femenino <- cuartiles_femenino[2] - cuartiles_femenino[1]
limite_inferior_femenino <- cuartiles_femenino[1] - 1.5 * IQR_femenino
limite_superior_femenino <- cuartiles_femenino[2] + 1.5 * IQR_femenino

```

```
# Calcular la mediana de PESO por género
```

```

mediana_peso_masculino <- median(MedicamentosPESO[MedicamentosSEXO ==
"MASCULINO"], na.rm = TRUE) mediana_peso_femenino <-
median(MedicamentosPESO[MedicamentosSEXO == "FEMENINO"], na.rm = TRUE) #
Imputar los outliers con la mediana correspondiente al género
MedicamentosPESO[MedicamentosSEXO == "MASCULINO" &
(MedicamentosPESO < limite_inferior_masculino | MedicamentosPESO >
limite_superior_masculino)] <- mediana_peso_masculino
MedicamentosPESO[MedicamentosSEXO == "FEMENINO" & (MedicamentosPESO <
limite_inferior_femenino | MedicamentosPESO > limite_superior_femenino)] <-
mediana_peso_femenino

```

```
# Verificación de la imputación
```

```

outliers_masculino <- sum(Medicamentos$SEXO == "MASCULINO" &
(Medicamentos$PESO < limite_inferior_masculino | Medicamentos
PESO > limite_superior_masculino)) outliers_femenino <- sum(MedicamentosSEXO
== "FEMENINO" & (MedicamentosPESO <
limite_inferior_femenino | MedicamentosPESO > limite_superior_femenino))

cat("Cantidad de outliers restantes en MASCULINO después de la imputación:",
outliers_masculino, "")

cat("Cantidad de outliers restantes en FEMENINO después de la imputación:",
outliers_femenino, "")

```

4.12 Imputación de outliers CANTIDAD, VALOR_BRUTO, VALOR_NETO y NIVEL_EESS

Se imputaron valores outliers ya que habían datos con números que superaban los miles mediante medianas.

NIVEL_EESS presentó un valor 0 por lo que se imputó como nivel 2.

```
# Imputar outliers en CANTIDAD con la mediana
```

```

MedicamentosCANTIDAD[MedicamentosCANTIDAD <
quantile(MedicamentosCANTIDAD, 0.05) | MedicamentosCANTIDAD >
quantile(MedicamentosCANTIDAD, 0.95)] <- median(MedicamentosCANTIDAD,
na.rm = TRUE)

```

```
# Imputar outliers en VALOR_BRUTO con la mediana
```

```
MedicamentosVALOR_BRUTO[MedicamentosVALOR_BRUTO <  
quantile(MedicamentosVALOR_BRUTO, 0.05)|MedicamentosVALOR_BRUTO >  
quantile(MedicamentosVALOR_BRUTO, 0.95)] <  
-median(MedicamentosVALOR_BRUTO, na.rm = TRUE)
```

```
# Imputar outliers en VALOR_NETO con la mediana
```

```
MedicamentosVALOR_NETO[MedicamentosVALOR_NETO <  
quantile(MedicamentosVALOR_NETO, 0.05)|MedicamentosVALOR_NETO >  
quantile(MedicamentosVALOR_NETO, 0.95)] <  
-median(MedicamentosVALOR_NETO, na.rm = TRUE)
```

```
# Reemplazar Valores 0 por 1 en la Variable NIVEL_EESS
```

```
Medicamentos <- Medicamentos %>% mutate( NIVEL_EESS = ifelse(NIVEL_EESS == 0,  
2, NIVEL_EESS) )
```

4.13. Exportar el dataset limpio cómo archivo CSV con la codificación UTF-8

Dataset limpio es exportado para su análisis.

```
fwrite(Medicamentos, "Medicamentos.csv", bom = TRUE)
```

5. Análisis Exploratorio de Datos (EDA)

Se realizó un análisis descriptivo de cada variable conocer el comportamiento de los datos en cada uno de ellas.

5.1. FECHA_CORTE

La generación del dataset fué el 26 abril del 2024

5.2. CODIGO_ANONIZADO

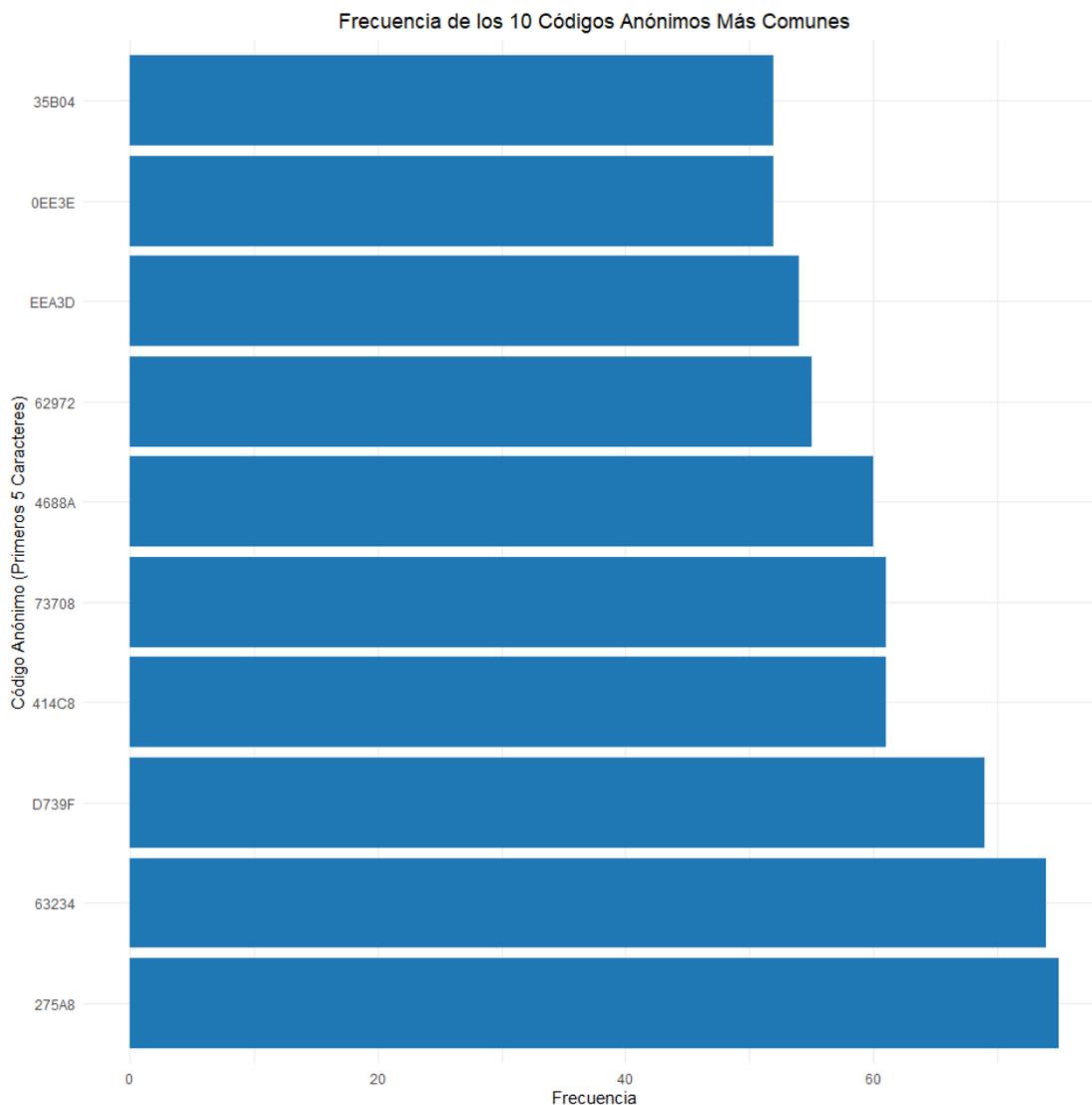
Identificador Anonimizado del Paciente

Se realizó un análisis de la frecuencia de los códigos anonimizados en el dataset para identificar los más comunes. Los resultados muestran que los 10 códigos más frecuentes se repiten entre 52 y 75 veces. Estos códigos representan diferentes pacientes en el conjunto de datos y su frecuencia puede estar relacionada con múltiples visitas o tratamientos registrados para los mismos individuos. La visualización presentada muestra un gráfico de barras que destaca la frecuencia de aparición de estos códigos, utilizando solo los primeros 5 caracteres para facilitar la lectura y comparación.

```
## # A tibble: 10 × 3  
##   CODIGO_ANONIMIZADO
```

Frecuencia

Codigo_Corto	
## <chr>	<int>
<chr>	
## 1 275A83246FF17D86F67D0414410676A97E6B0863BEB4FDAA8F26... 275A8	75
## 2 6323449532FEIBC01CA7D01CC18352A9613BC017E62F11ABDFD1... 63234	74
## 3 D739F4B76ED6D99D94826E413CF0BC4A07524F532C9BA8DCECBD... D739F	69
## 4 414C8318C6CC0F975C7643E69FF6135FC980D21DE91F84804081... 414C8	61
## 5 737084D974E8582A8CE7852C85C0D386AB8D6E4D70E3E000F94F... 73708	61
## 6 4688A2E238F87D26BF09B1BF92120F3CA0D0E60DC37690B45801... 4688A	60
## 7 62972F3C3BA33DF0A1F22E40C8DDE01799C20C1AFF5F8A76CD7D... 62972	55
## 8 EEA3DAD1526E1D220B0DDDD8C70075C56431C35402EFC8EE2A91... EEA3D	54
## 9 0EE3ED9F916D4E3A954875427B28E072CA88166F657742C0D691... 0EE3E	52
## 10 35B04891B698894D34D42A60B089FAE744266AD846A96526BE78... 35B04	52



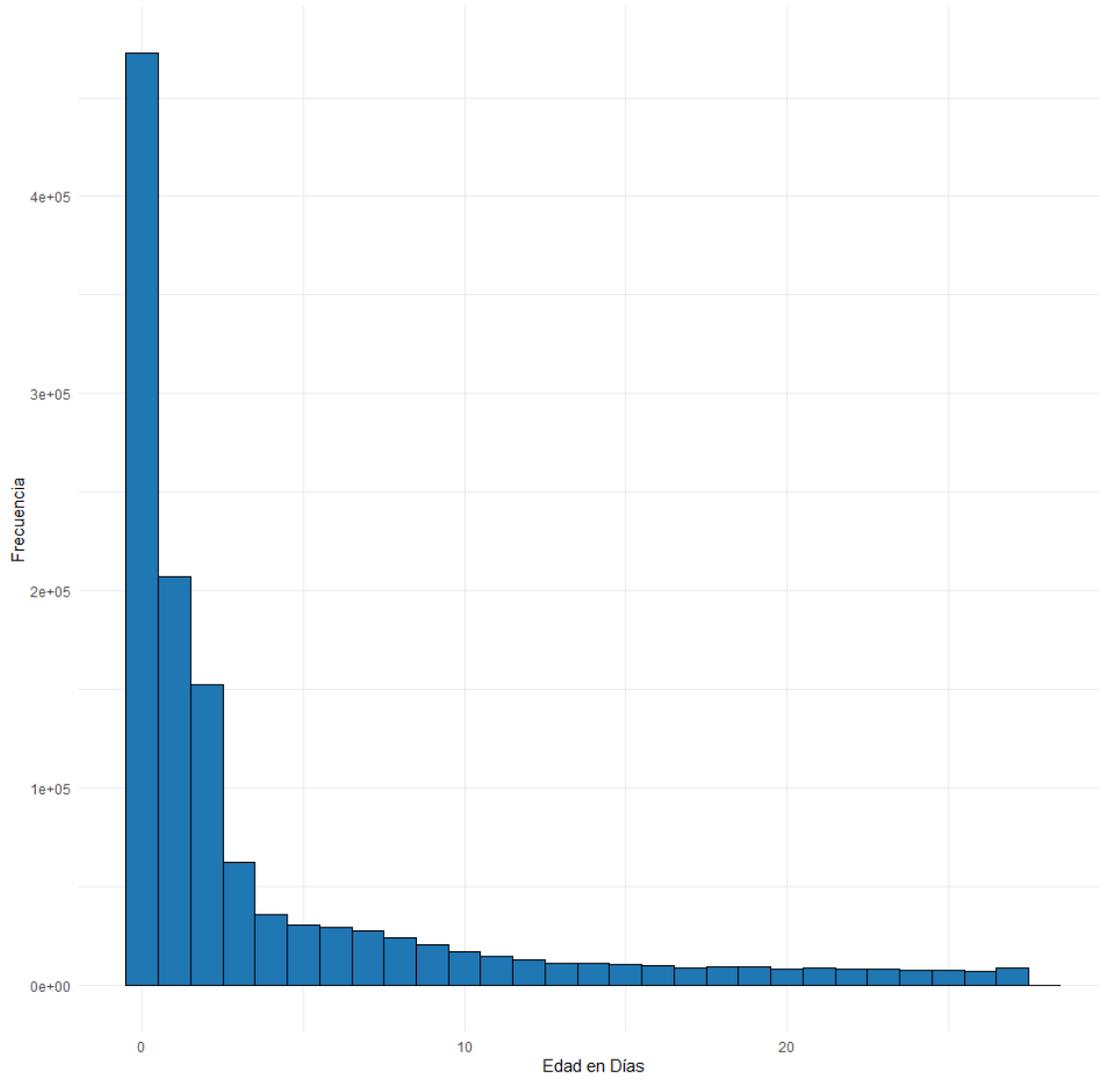
5.3. EDAD_DIAS

Edad del paciente a la fecha de atención. La unidad de medida es días.

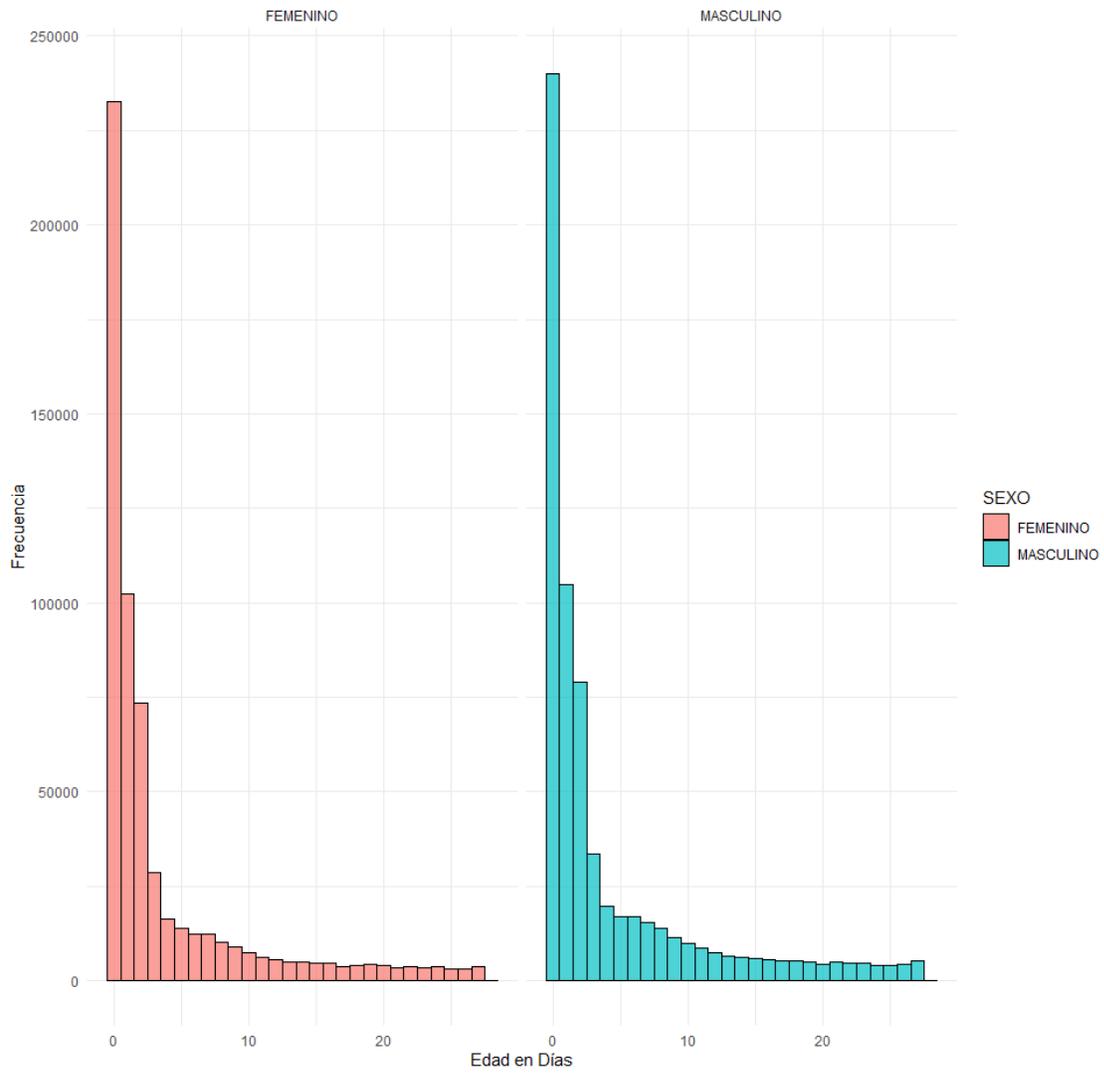
La variable EDAD_DIAS proporciona información sobre la edad de los pacientes en días, desde el momento del nacimiento hasta el final del registro. El análisis estadístico revela que la edad mínima es 0 días (recién nacidos), y la máxima es 28 días, con una media de 3.9 días y una desviación estándar de 6.23 días. La mayoría de los recién nacidos en el dataset tienen entre 0 y 5 días de edad, lo que refleja un enfoque significativo en los cuidados perinatales. Para visualizar esta distribución, se presentan varios histogramas, incluyendo separaciones por sexo, departamento y nivel de centro de salud, además de diagramas de caja (boxplots) para ilustrar la dispersión y los valores extremos en los datos.

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
##	0.000000	0.000000	1.000000	3.896931	5.000000	28.000000	6.233279

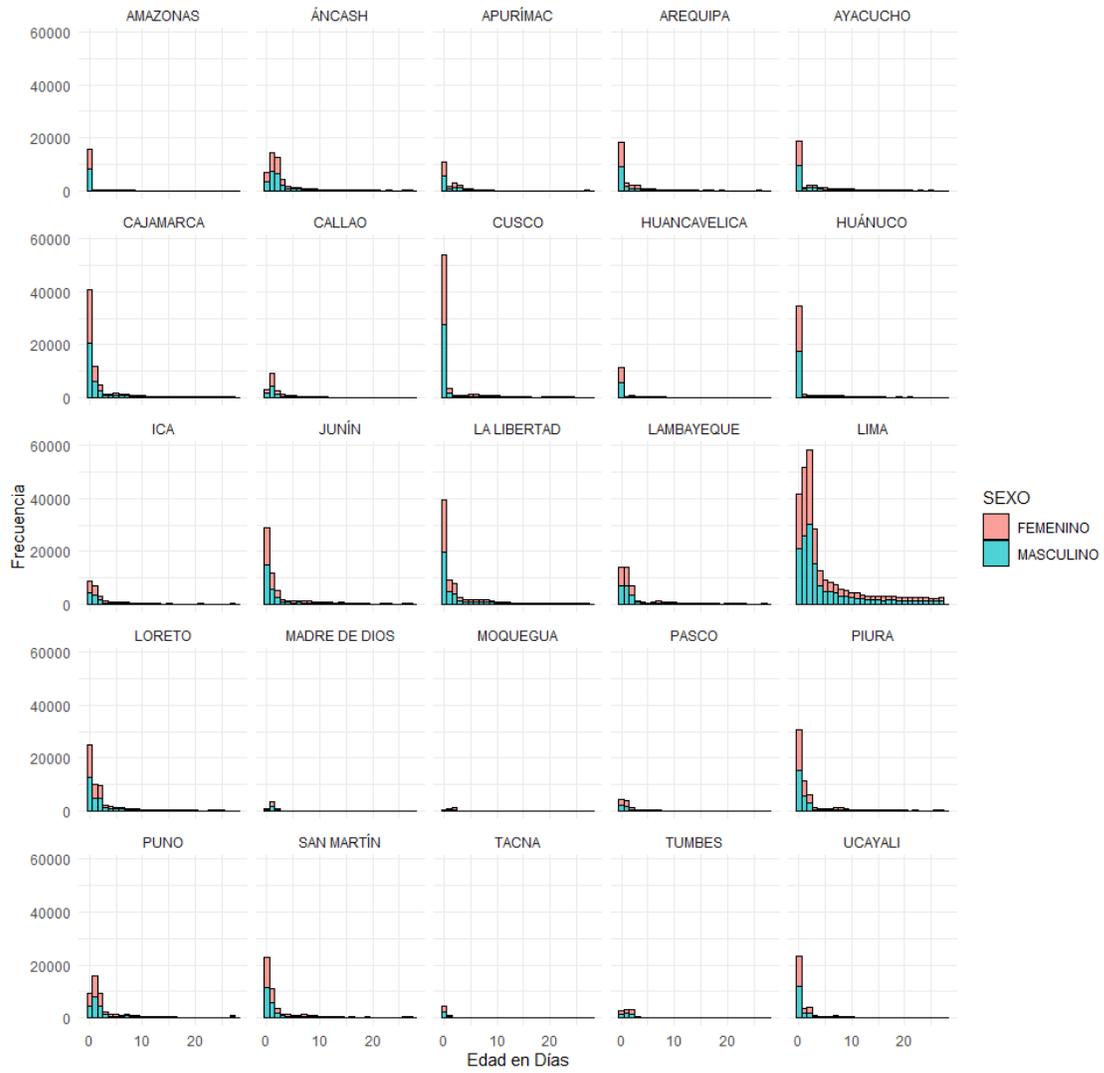
Distribución de la Edad en Días



Distribución de la Edad en Días Separada por Sexo



Distribución de la Edad en Días Separada por Departamento



Distribución de la Edad en Días

Nivel de Centro de Salud

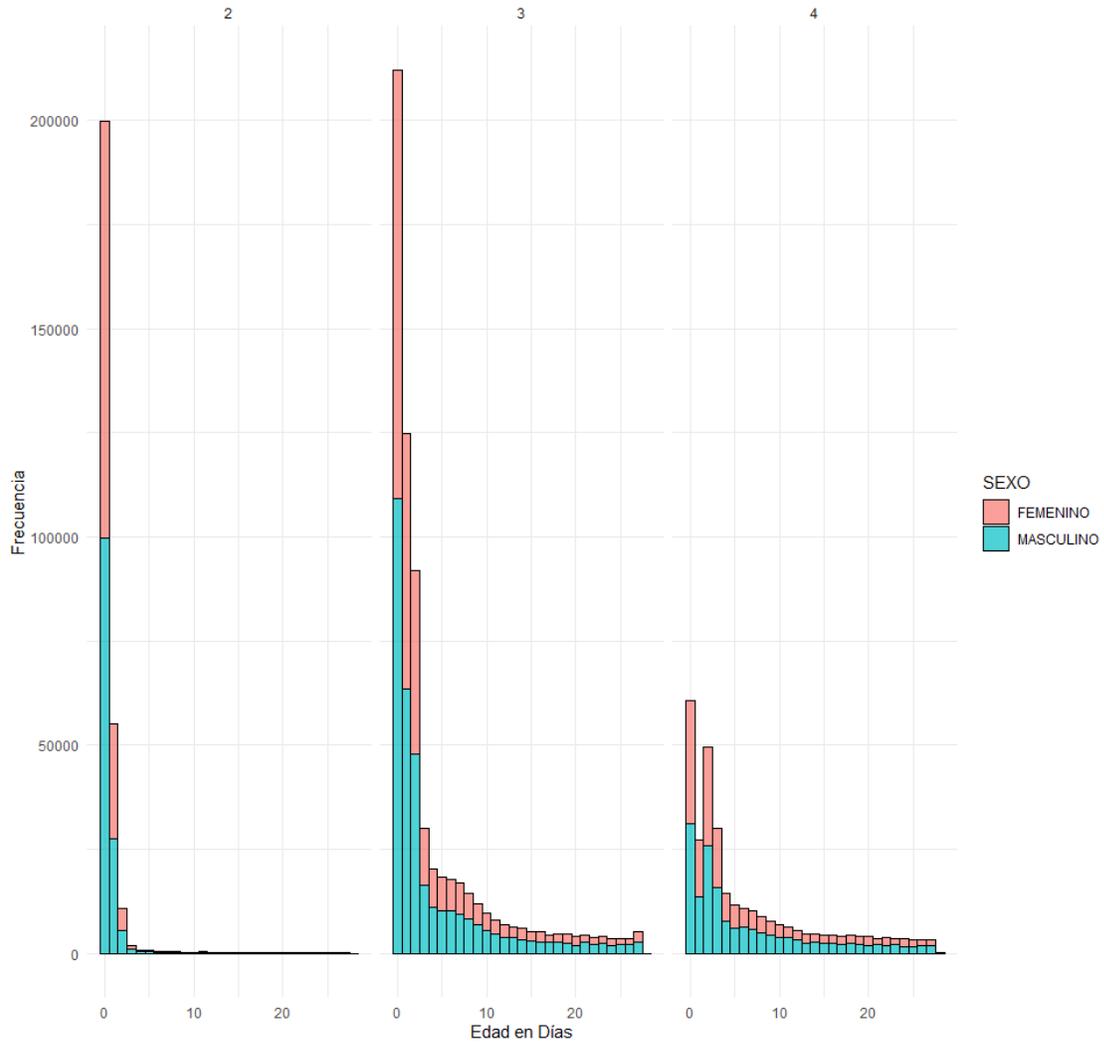
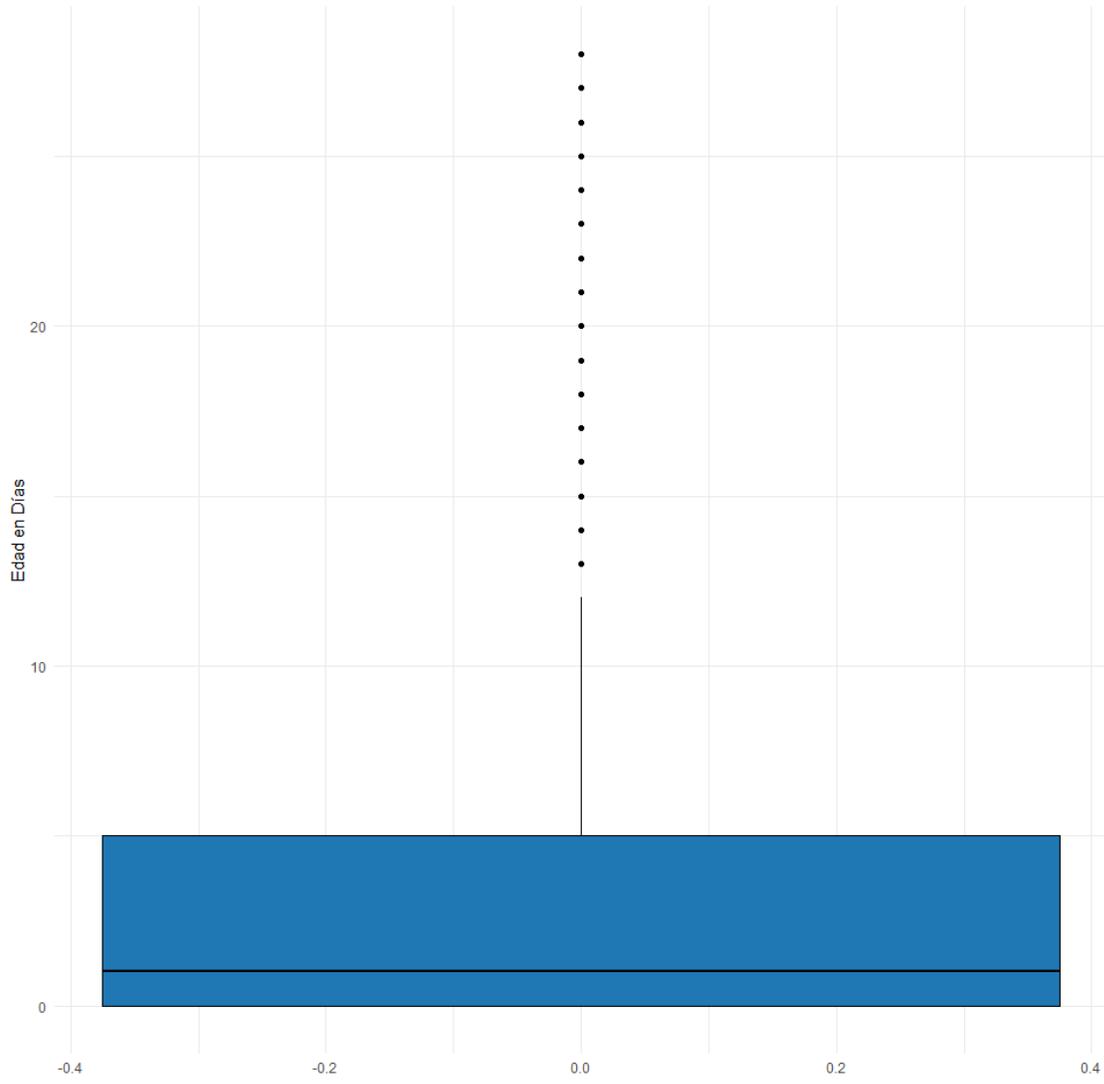
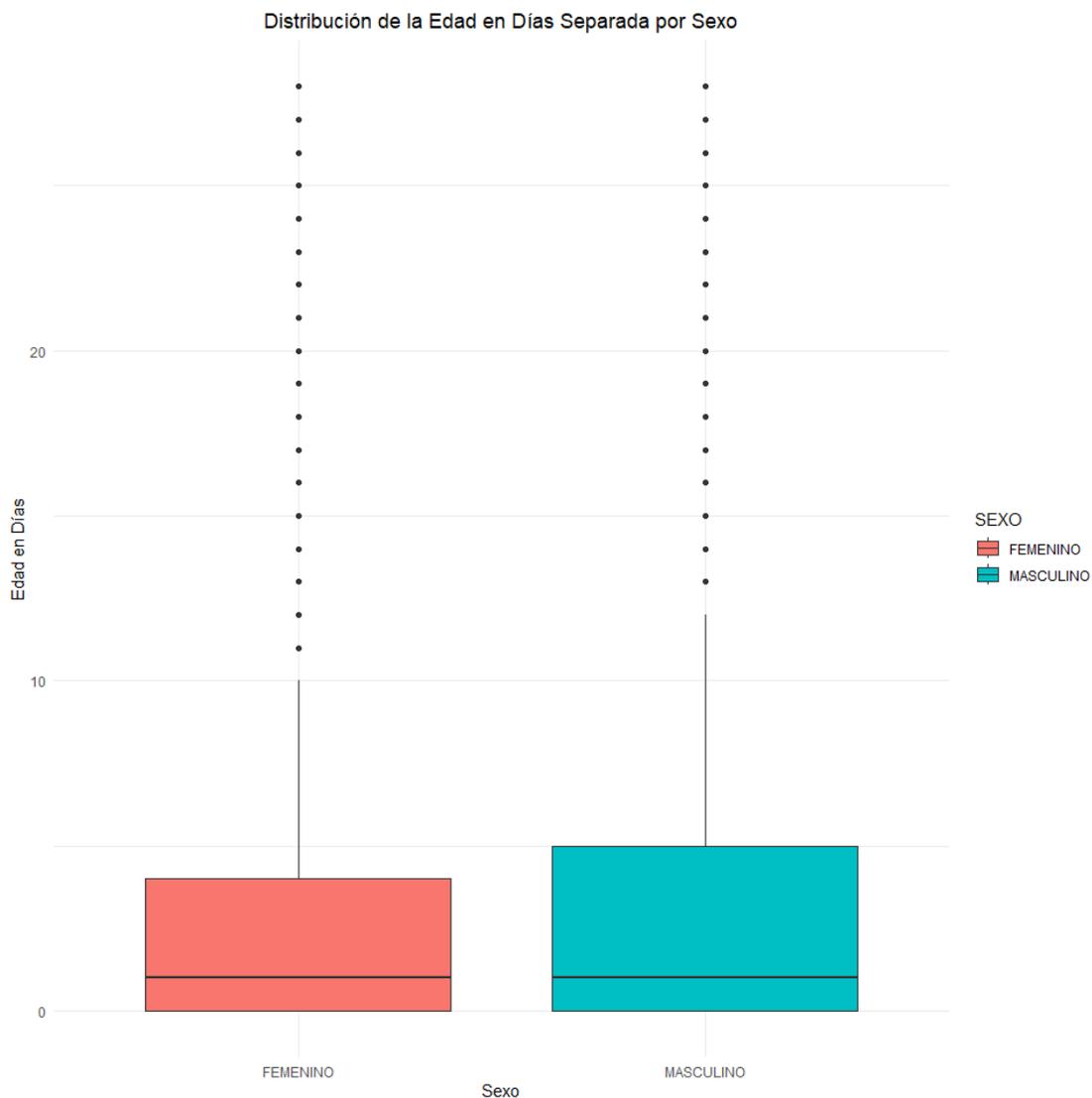


Diagrama de Caja de la Edad en Días



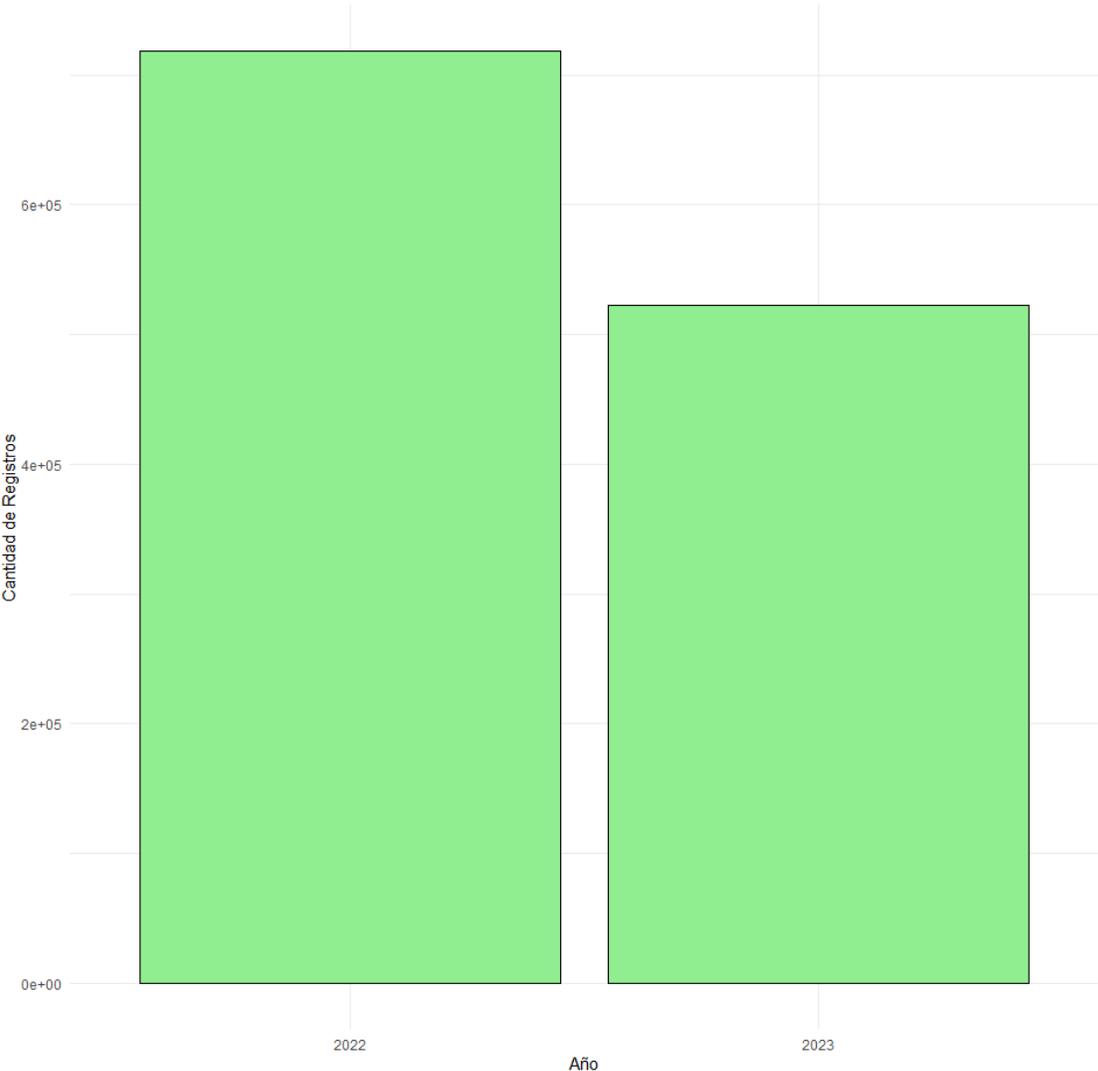


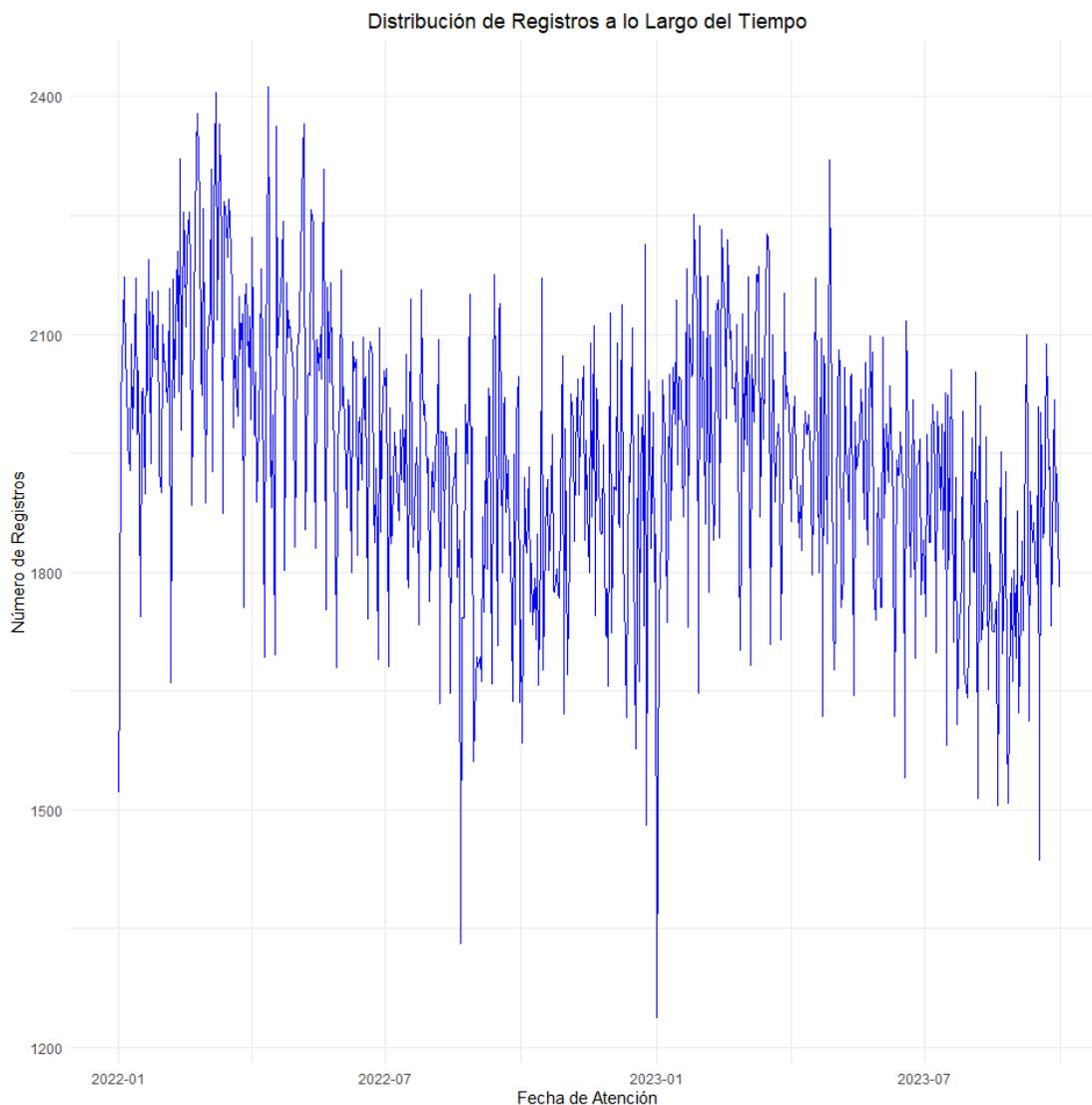
5.4. FECATENCION

La variable FECATENCION se refiere a la fecha en la que se brindó atención médica a los recién nacidos. El rango de fechas abarca desde el 1 de enero de 2022 hasta el 30 de septiembre de 2023, con la mayor concentración de atenciones ocurriendo entre mayo de 2022 y abril de 2023. El análisis muestra una distribución uniforme de registros a lo largo del tiempo, con una ligera concentración en ciertos períodos. Se incluyen visualizaciones que muestran la cantidad de registros por año, así como una serie de tiempo que ilustra la evolución de los registros diarios, destacando las fechas de corte con líneas punteadas para un análisis temporal más profundo.

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	"2022-01-01"	"2022-05-30"	"2022-11-10"	"2022-11-09"	"2023-04-17"	"2023-09-30"

Cantidad de Registros por Año





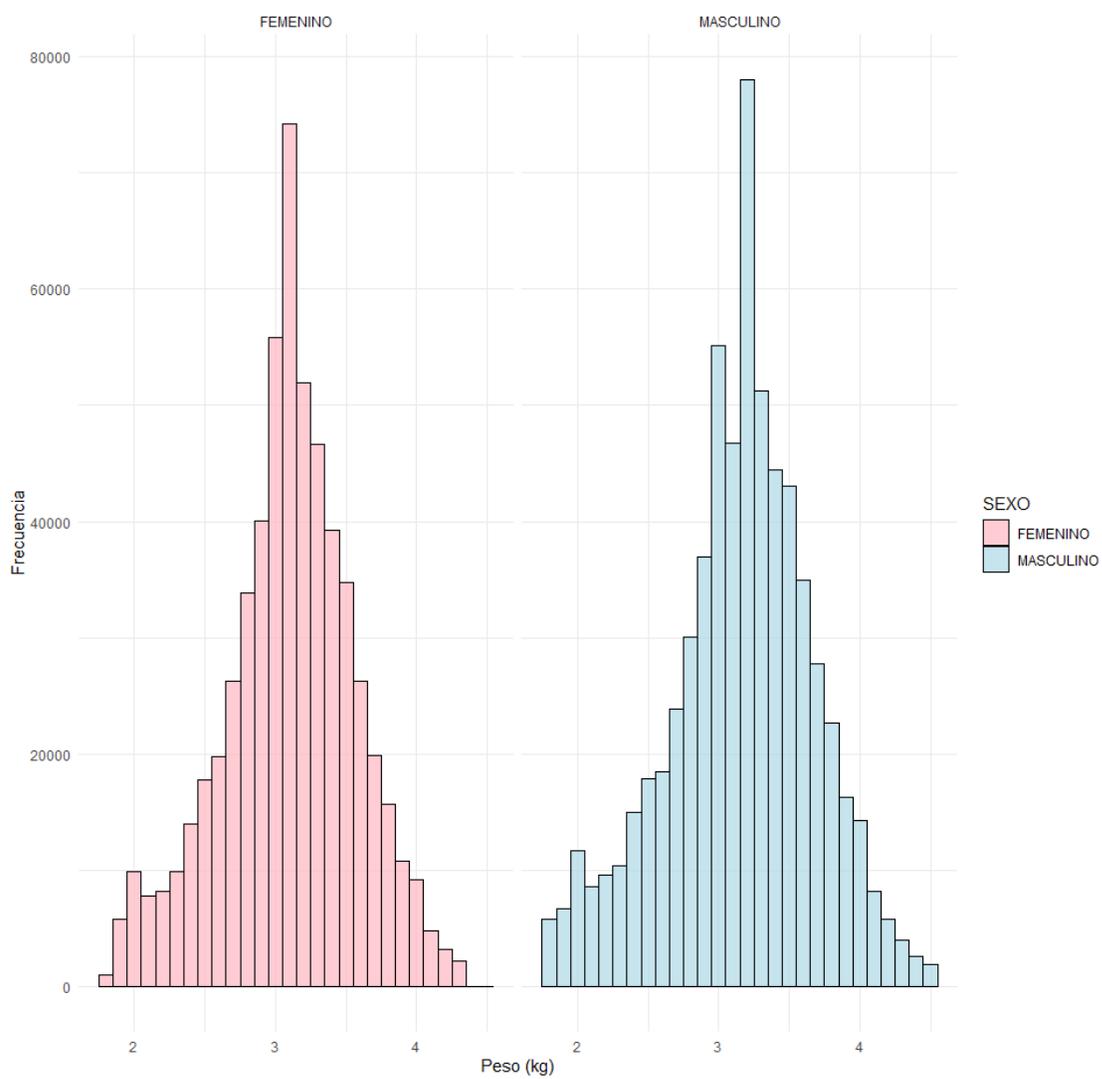
5.5. SEXO

La variable SEXO categoriza a los pacientes en femenino y masculino. El análisis de esta variable revela que hay una ligera predominancia de pacientes masculinos (651,827 registros) en comparación con los femeninos (588,778 registros). Este patrón puede ser relevante para entender las dinámicas de género en la atención médica neonatal en el contexto estudiado. El gráfico de barras asociado muestra la distribución de los registros por sexo, utilizando colores diferenciados (rosado para femenino y azul para masculino) para facilitar la comparación visual. La variable PESO refleja el peso de los recién nacidos al nacer, con un rango que va desde 1.75 kg hasta 4.55 kg. El peso promedio de los recién nacidos es de aproximadamente 3.13 kg, con una desviación estándar de 0.496 kg, lo que indica una dispersión relativamente estrecha en torno a la media. Esta información es crucial para evaluar el estado de salud general de los recién nacidos, ya que el peso al nacer es un indicador clave de bienestar neonatal. Se presentan histogramas para visualizar la distribución del peso

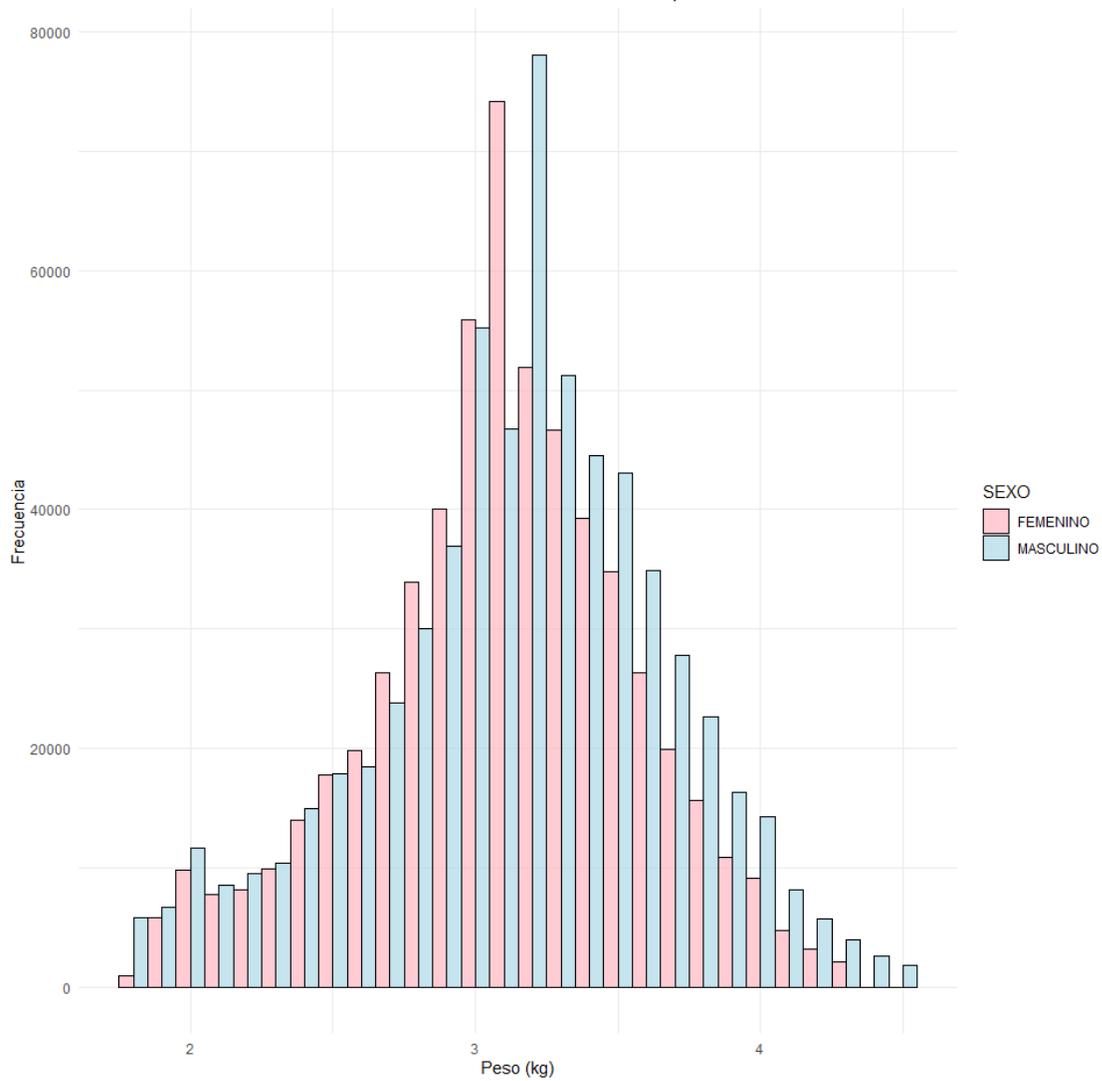
por género, así como diagramas de caja (boxplots) que muestran la variabilidad y posibles valores atípicos en los datos.

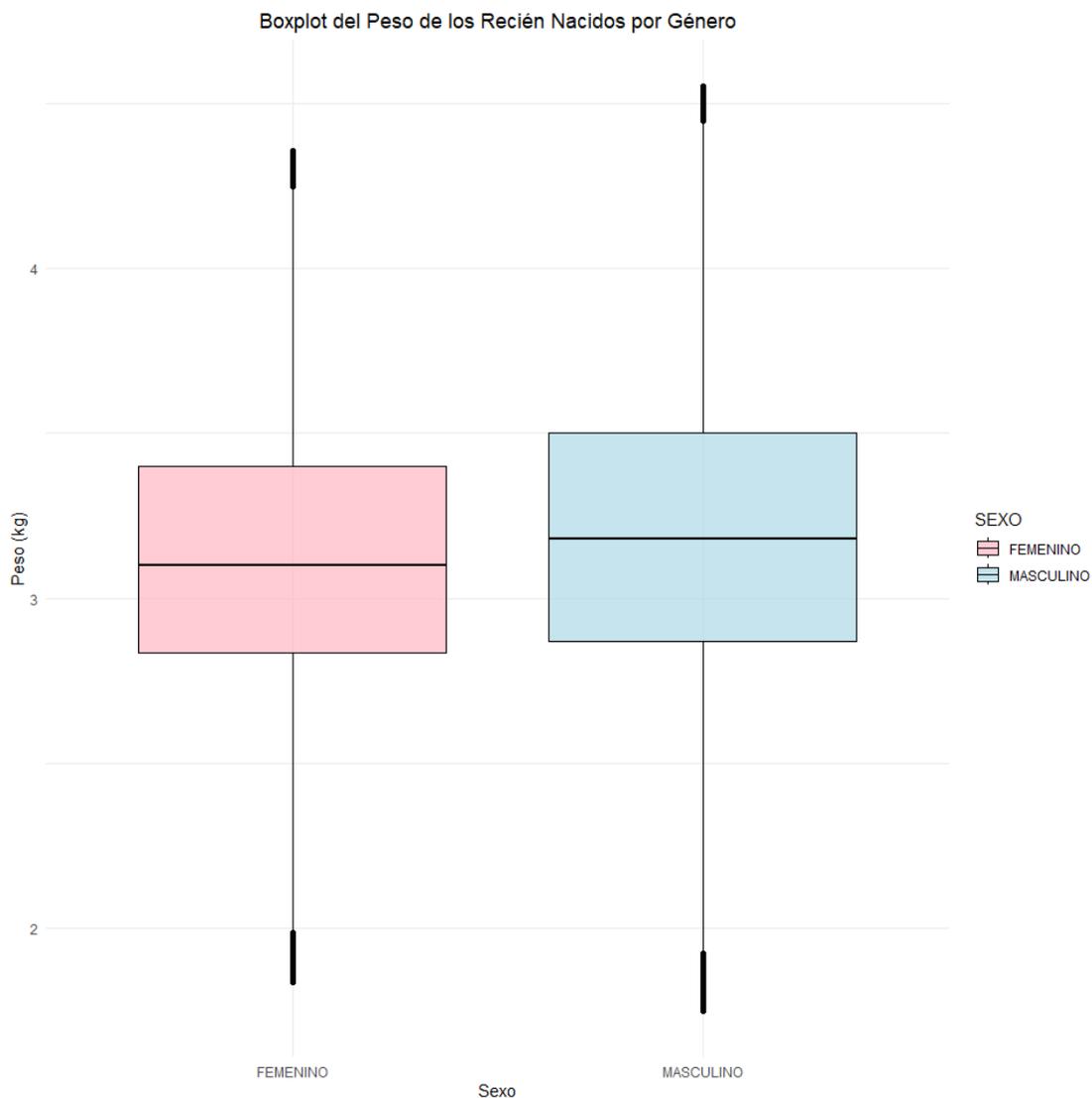
```
## # A tibble: 1 × 7
##   Min `1er Q` Median Mean   SD `3er Q`  Max
##   <dbl>  <dbl>  <dbl> <dbl> <dbl>  <dbl> <dbl>
## 1  1.75    2.85   3.15  3.13  0.496   3.45  4.55
```

Distribución del Peso de los Recién Nacidos por Género



Distribución del Peso de los Recién Nacidos por Género





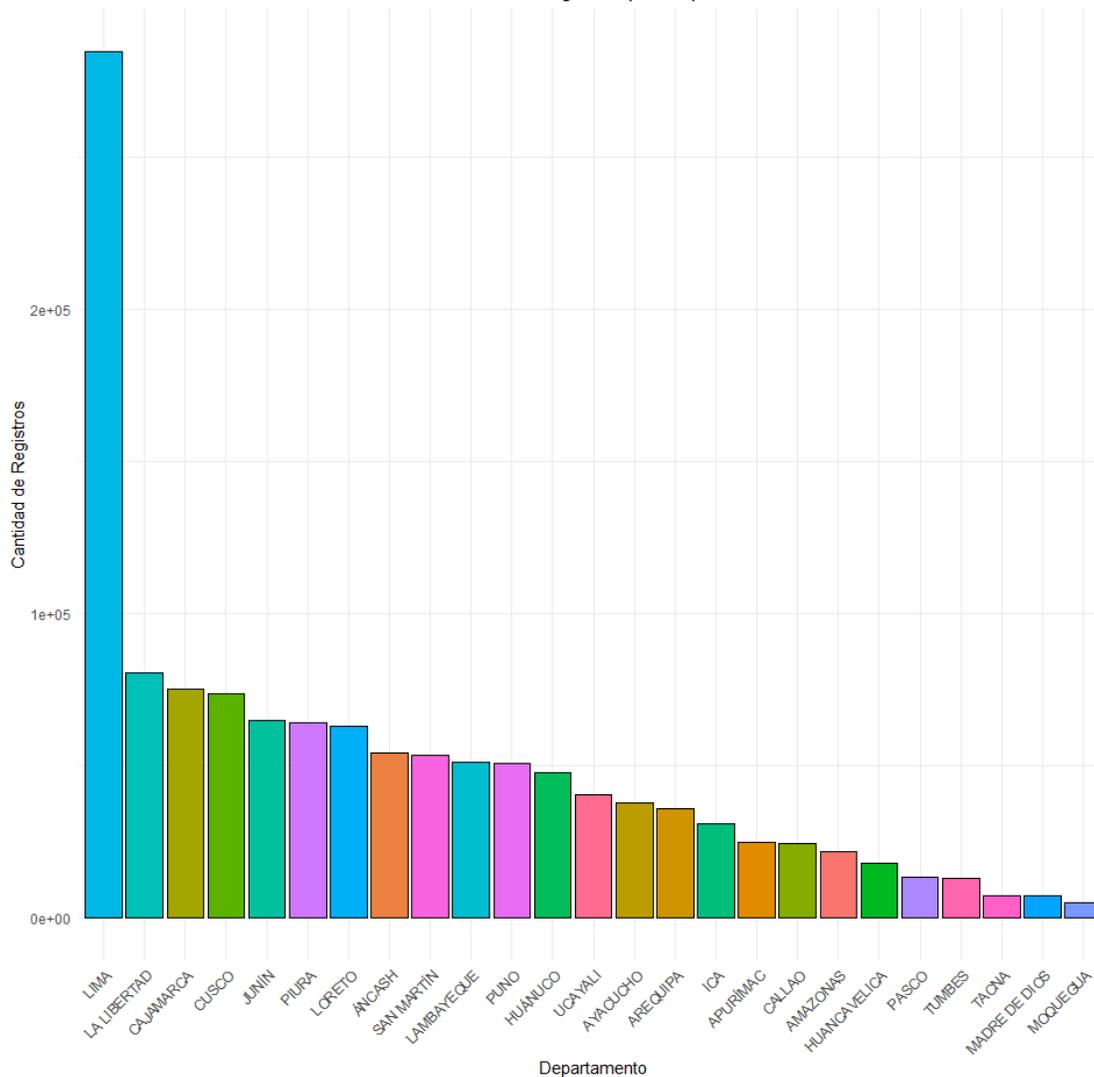
5.7. DEPARTAMENTO

El análisis de la distribución de registros por departamento revela que Lima concentra la mayor cantidad de registros con 284,662, seguido por La Libertad, Cajamarca, y Cusco, que tienen entre 73,592 y 80,447 registros. Esta concentración en Lima es esperada dada su alta población y la centralización de los servicios de salud en la capital. Por otro lado, departamentos como Moquegua, Madre de Dios, y Tacna presentan los menores números de registros, con menos de 8,000 casos cada uno, lo que podría reflejar una menor población o acceso limitado a servicios de salud en estas regiones. Estos datos se visualizan en gráficos de barras, que destacan tanto los departamentos con más como con menos registros.

```
## # A tibble: 10 × 2
##   DEPARTAMENTO Count
##   <fct>         <int>
## 1 LIMA           284662
```

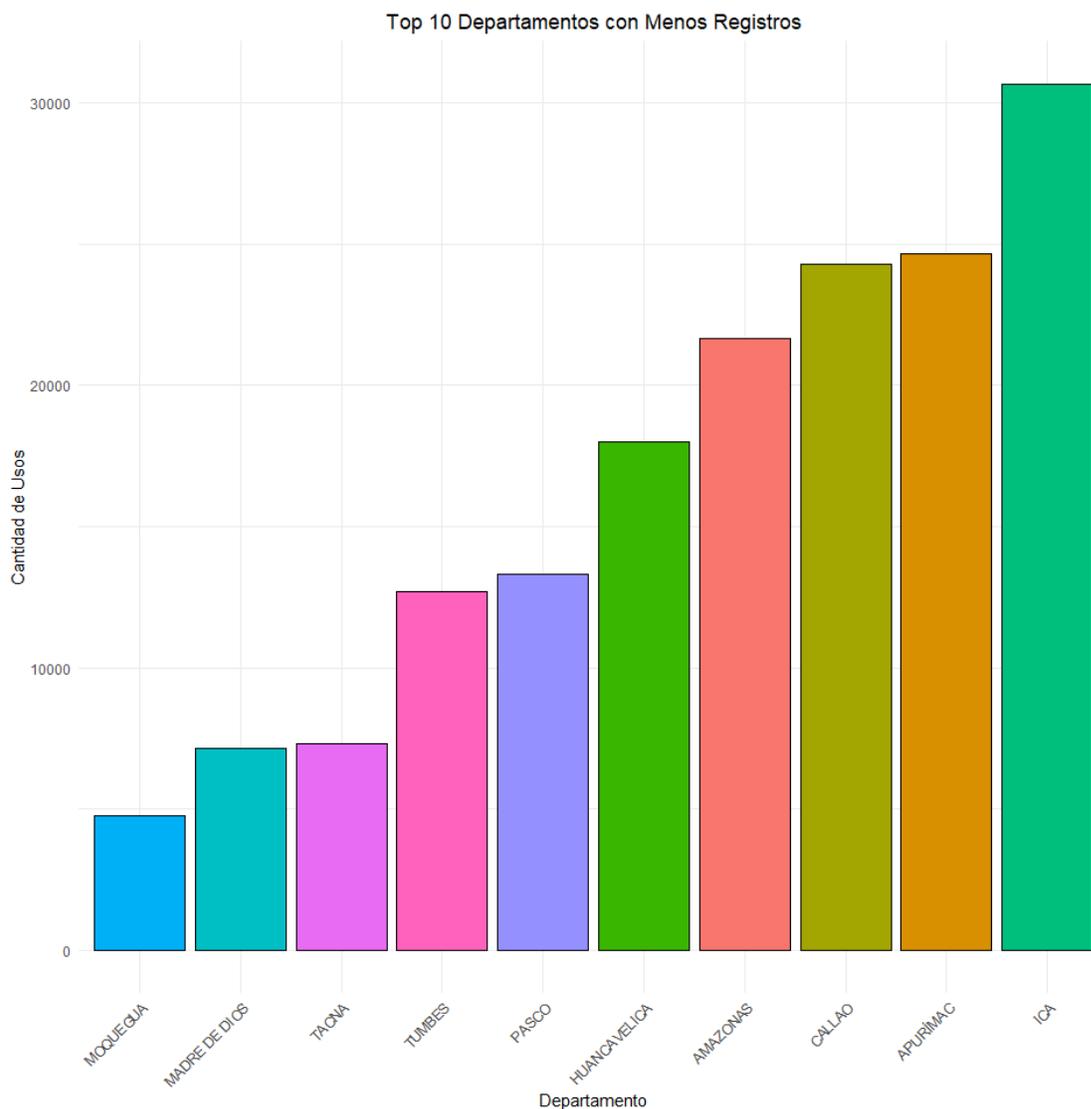
##	2	LA LIBERTAD	80447
##	3	CAJAMARCA	74949
##	4	CUSCO	73592
##	5	JUNÍN	64865
##	6	PIURA	64081
##	7	LORETO	62779
##	8	ÁNCASH	54082
##	9	SAN MARTÍN	53347
##	10	LAMBAYEQUE	51071

Distribución de Registros por Departamento



```
## # A tibble: 10 × 2
##   DEPARTAMENTO Cantidad_Usos
##   <fct>          <int>
## 1 MOQUEGUA      4765
## 2 MADRE DE DIOS 7161
## 3 TACNA         7322
## 4 TUMBES        12700
```

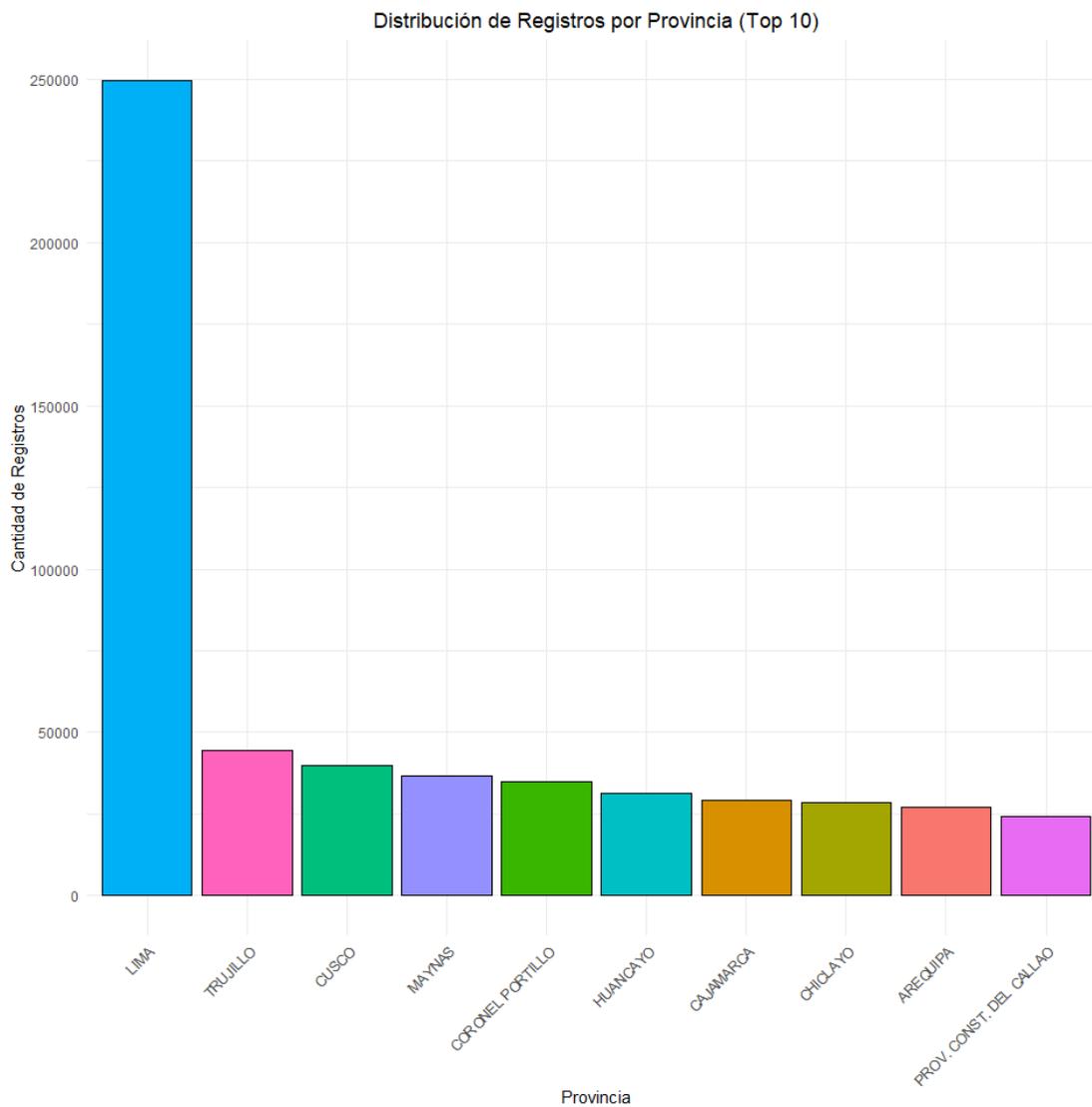
##	5	PASCO	13316
##	6	HUANCAVELICA	17997
##	7	AMAZONAS	21652
##	8	CALLAO	24279
##	9	APURÍMAC	24648
##	10	ICA	30649



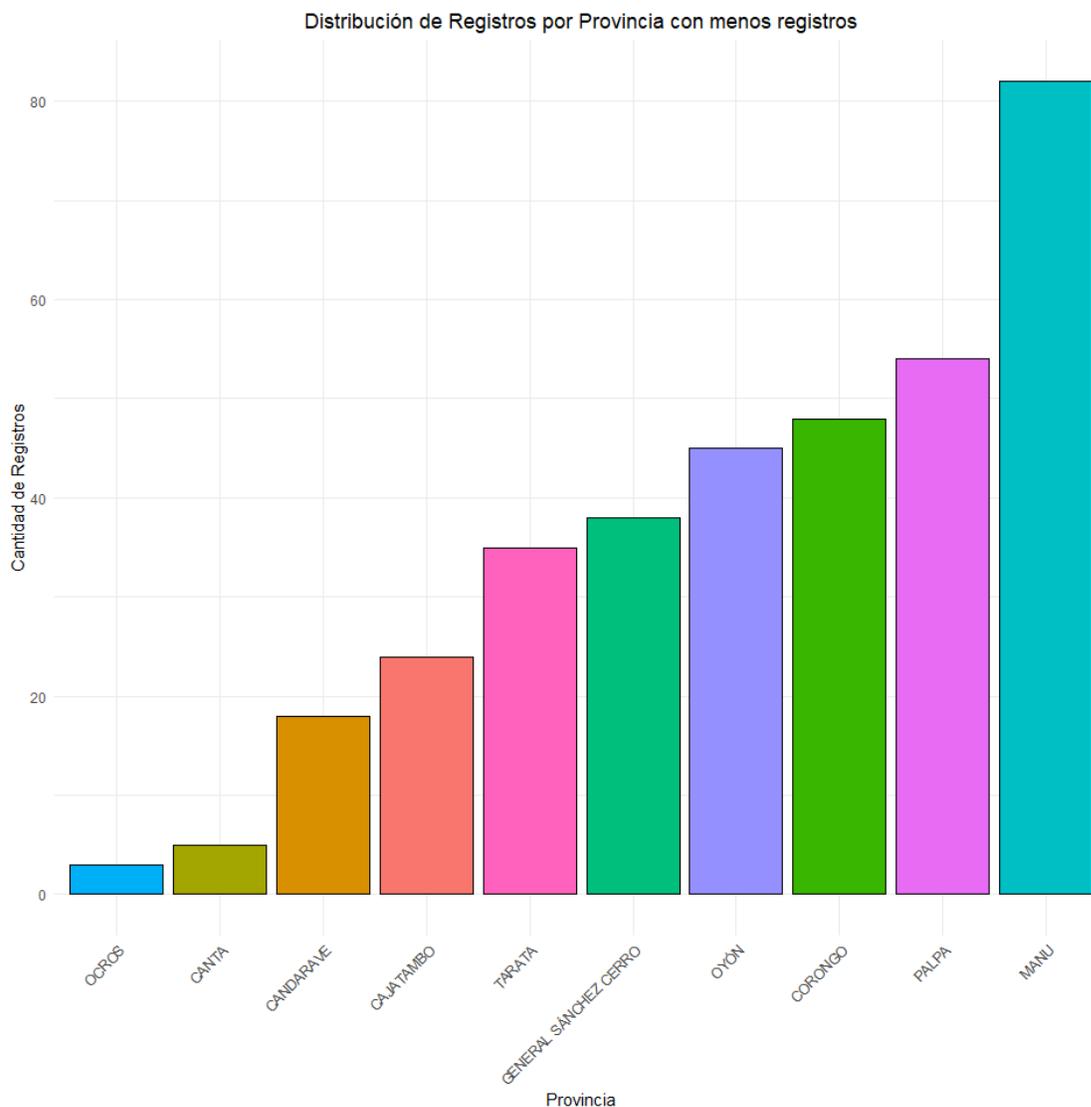
5.8. PROVINCIA

En cuanto a la distribución por provincia, la ciudad de Lima lidera con 83,743 registros, seguida por Trujillo, San Juan de Miraflores, y Cajamarca, que registran entre 26,721 y 32,926 casos. Estas cifras reflejan la alta demanda de atención en áreas urbanas y densamente pobladas. En contraste, provincias como Ocos y Canta, con tan solo 3 y 5 registros respectivamente, se encuentran en el extremo opuesto, lo que podría indicar áreas con menor cobertura de servicios de salud o baja densidad

poblacional. La distribución de registros por provincia se presenta en gráficos de barras que ilustran tanto los valores más altos como los más bajos.



```
## # A tibble: 10 × 2
##   PROVINCIA      Cantidad_Usos
##   <fct>          <int>
## 1 OCROS           3
## 2 CANTA           5
## 3 CANDARAVE      18
## 4 CAJATAMBO      24
## 5 TARATA         35
## 6 GENERAL SÁNCHEZ CERRO 38
## 7 OYÓN           45
## 8 CORONGO        48
## 9 PALPA          54
## 10 MANU           82
```



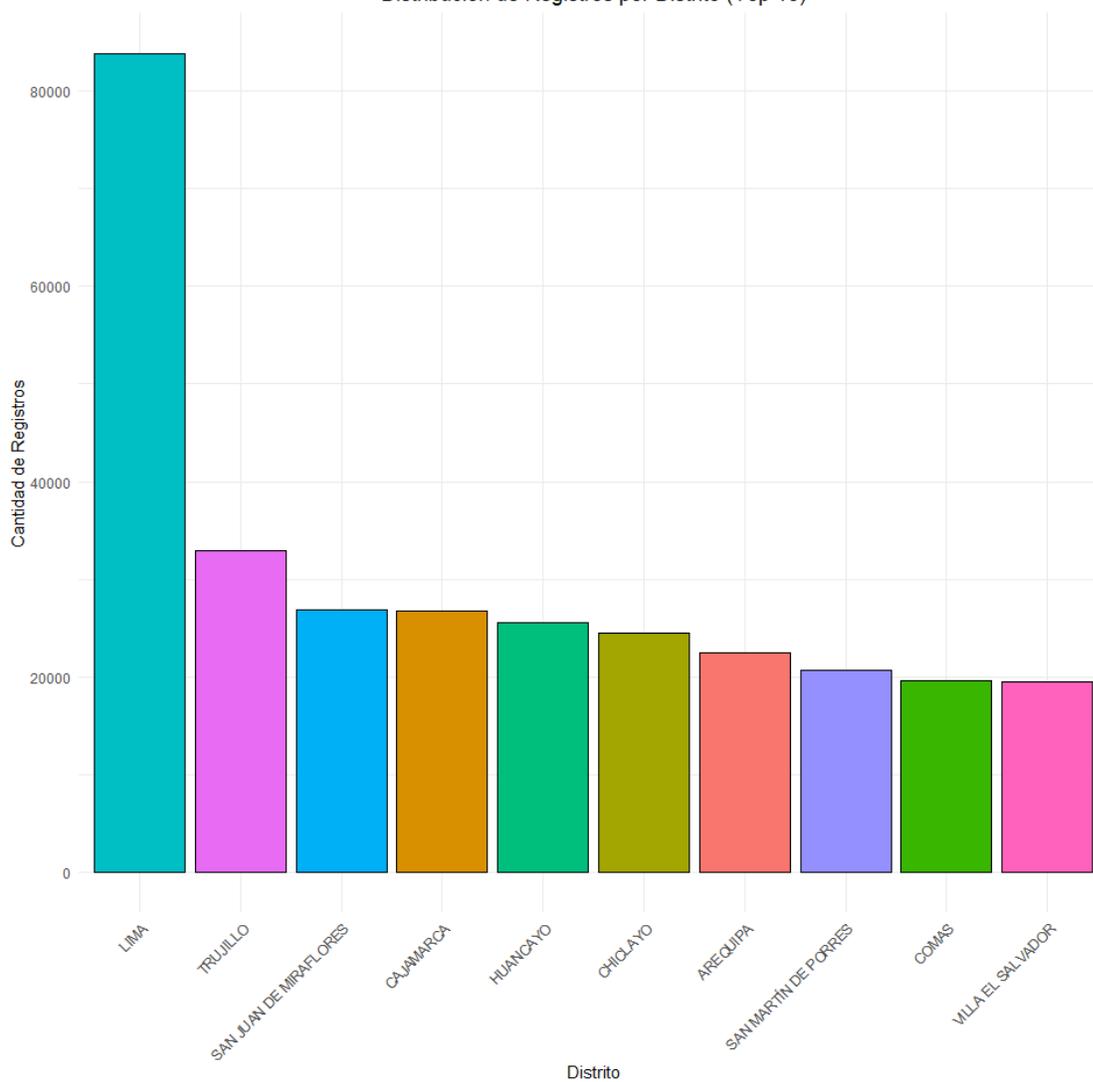
5.9. DISTRITO

El análisis por distrito muestra que Lima es nuevamente el distrito con el mayor número de registros, con 83,743 casos, seguido por Trujillo y San Juan de Miraflores. Estos resultados destacan la importancia de estos distritos como centros de atención médica. Por otro lado, varios distritos, como Acos Vinchos y Capachica, tienen solo un registro, lo que podría reflejar la limitada cobertura de datos o menor acceso a servicios de salud en estas áreas. Para visualizar esta información, se han creado gráficos de barras que destacan tanto los distritos con más como con menos registros.

```
## # A tibble: 10 × 2
##   DISTRITO      Count
##   <fct>        <int>
## 1 LIMA          83743
## 2 TRUJILLO      32926
## 3 SAN JUAN DE MIRAFLORES 26915
```

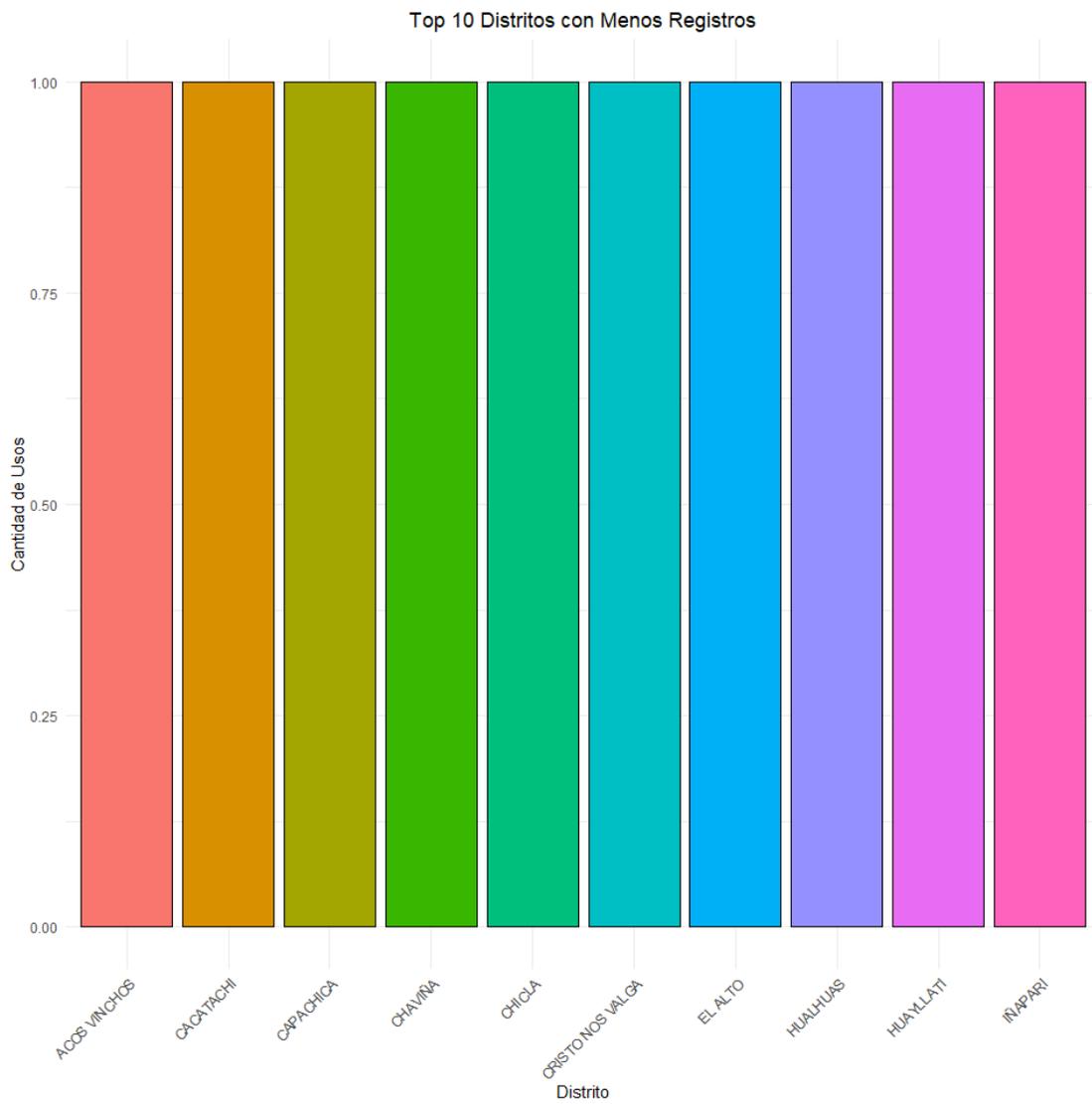
##	4	CAJAMARCA	26721
##	5	HUANCAYO	25555
##	6	CHICLAYO	24458
##	7	AREQUIPA	22478
##	8	SAN MARTÍN DE PORRES	20650
##	9	COMAS	19603
##	10	VILLA EL SALVADOR	19481

Distribución de Registros por Distrito (Top 10)



```
## # A tibble: 10 × 2
##   DISTRITO      Cantidad_Usos
##   <fct>         <int>
## 1 ACOS VINCHOS         1
## 2 CACATACHI            1
## 3 CAPACHICA            1
## 4 CHAVIÑA              1
## 5 CHICLA               1
## 6 CRISTO NOS VALGA     1
```

##	7 EL ALTO	1
##	8 HUALHUAS	1
##	9 HUAYLLATI	1
##	10 IÑAPARI	1

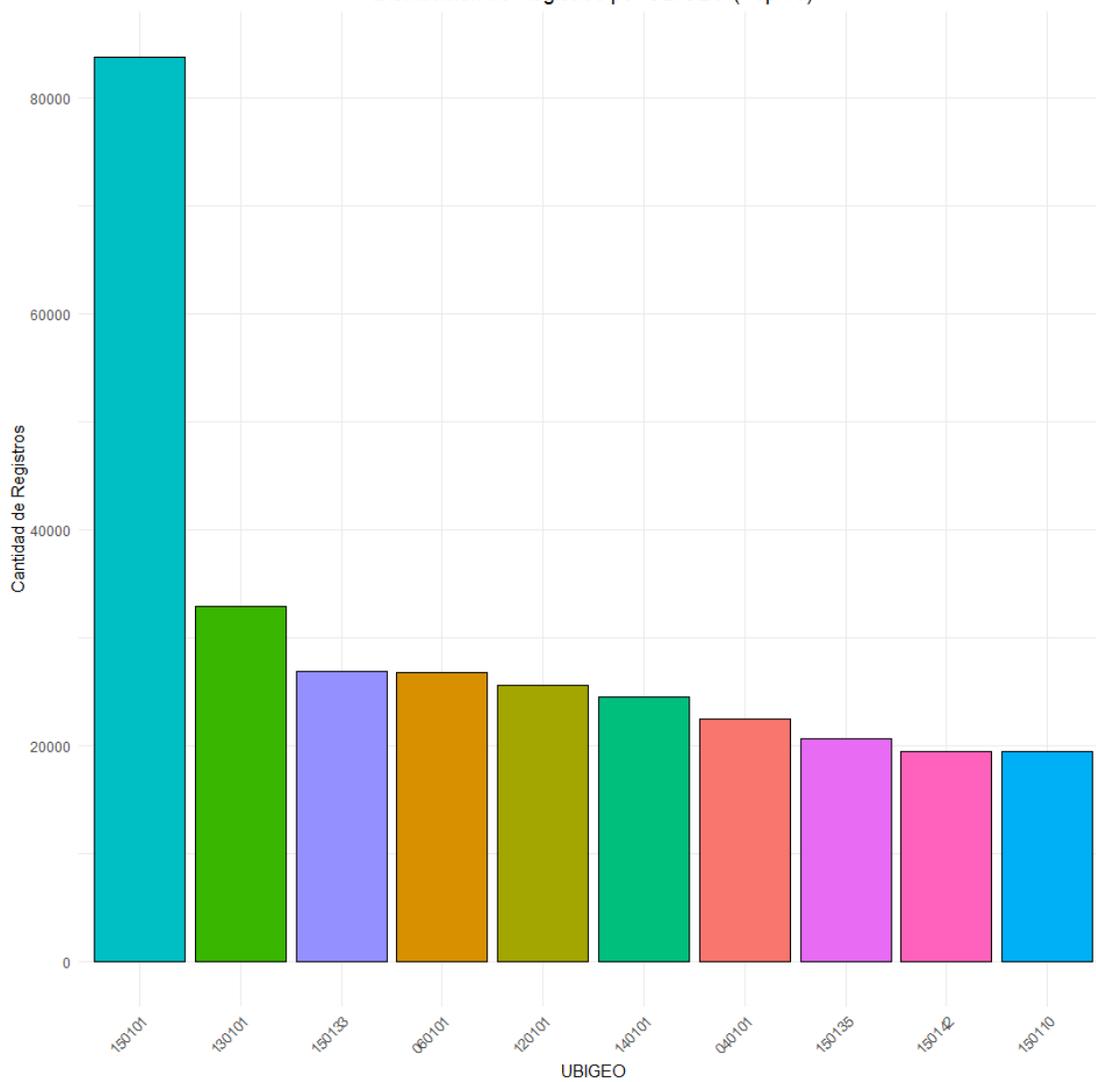


5.10. UBIGEO

El código UBIGEO, que identifica de manera única a las localidades geográficas en Perú, muestra que el distrito de Lima (código 150101) tiene el mayor número de registros, seguido por Trujillo (130101) y San Juan de Miraflores (150133). Estos datos refuerzan los hallazgos anteriores sobre la concentración de atención médica en estas áreas. El gráfico de barras correspondiente visualiza los 10 códigos UBIGEO más frecuentes, permitiendo una rápida identificación de las áreas con mayor actividad en el sistema de salud.

```
## # A tibble: 10 × 2
##   UBIGEO Count
##   <fct> <int>
## 1 150101 83743
## 2 130101 32926
## 3 150133 26915
## 4 060101 26721
## 5 120101 25555
## 6 140101 24458
## 7 040101 22478
## 8 150135 20650
## 9 150142 19481
## 10 150110 19467
```

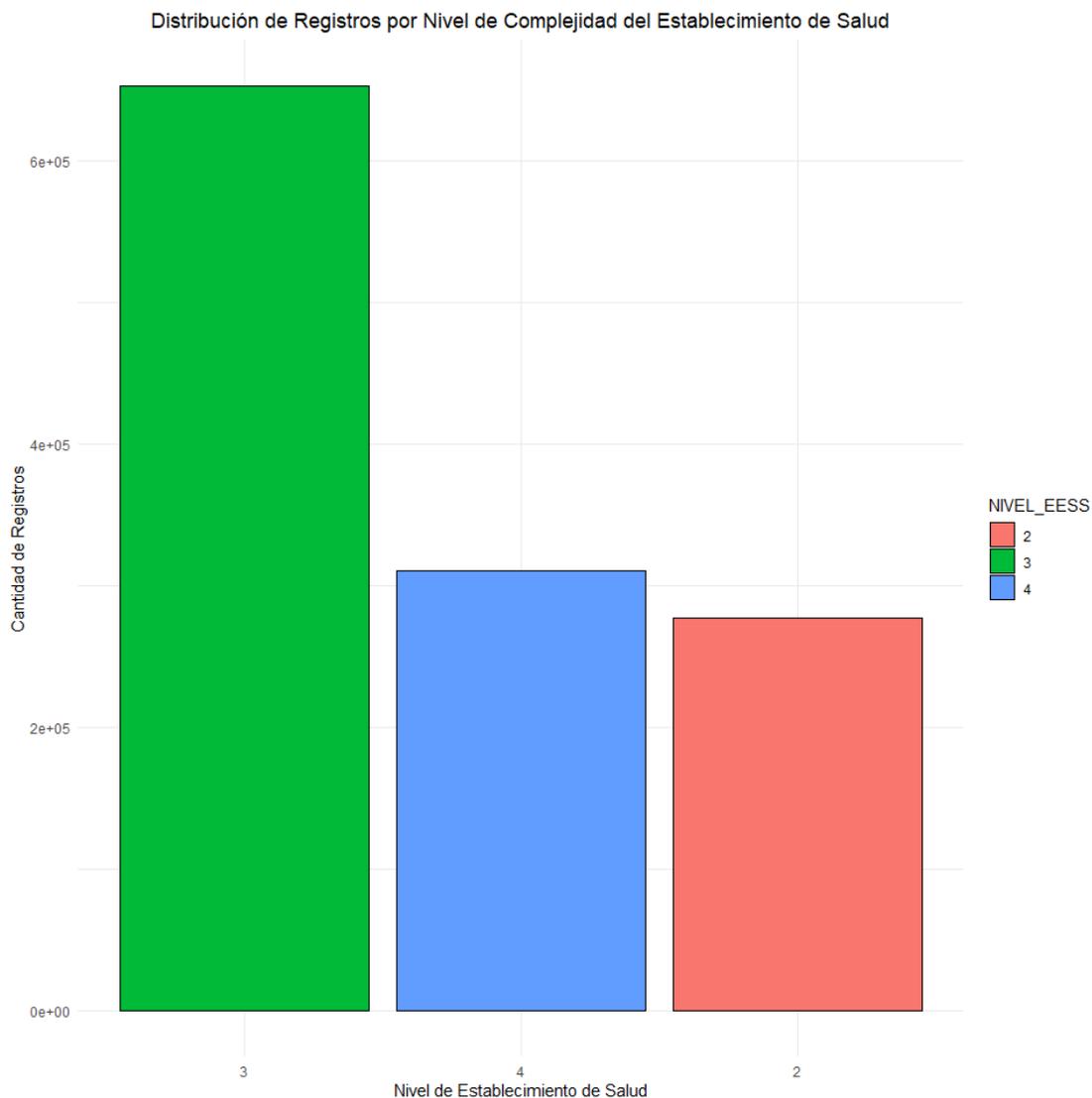
Distribución de Registros por UBIGEO (Top 10)



5.11. NIVEL_EESS

La clasificación de los registros por nivel de complejidad del establecimiento de salud revela que los establecimientos de nivel 3 son los más representados, con 652,948 registros, seguidos por los de nivel 4 y nivel 2, con 310,272 y 277,385 registros respectivamente. Esto sugiere que una gran proporción de la atención se lleva a cabo en establecimientos con una alta capacidad de resolución, lo cual es crucial para la atención neonatal. El gráfico de barras correspondiente ilustra esta distribución, subrayando la predominancia de los centros de salud de nivel 3.

```
## # A tibble: 3 × 2
##   NIVEL_EESS Count
##   <fct>      <int>
## 1 3          652948
## 2 4          310272
## 3 2          277385
```

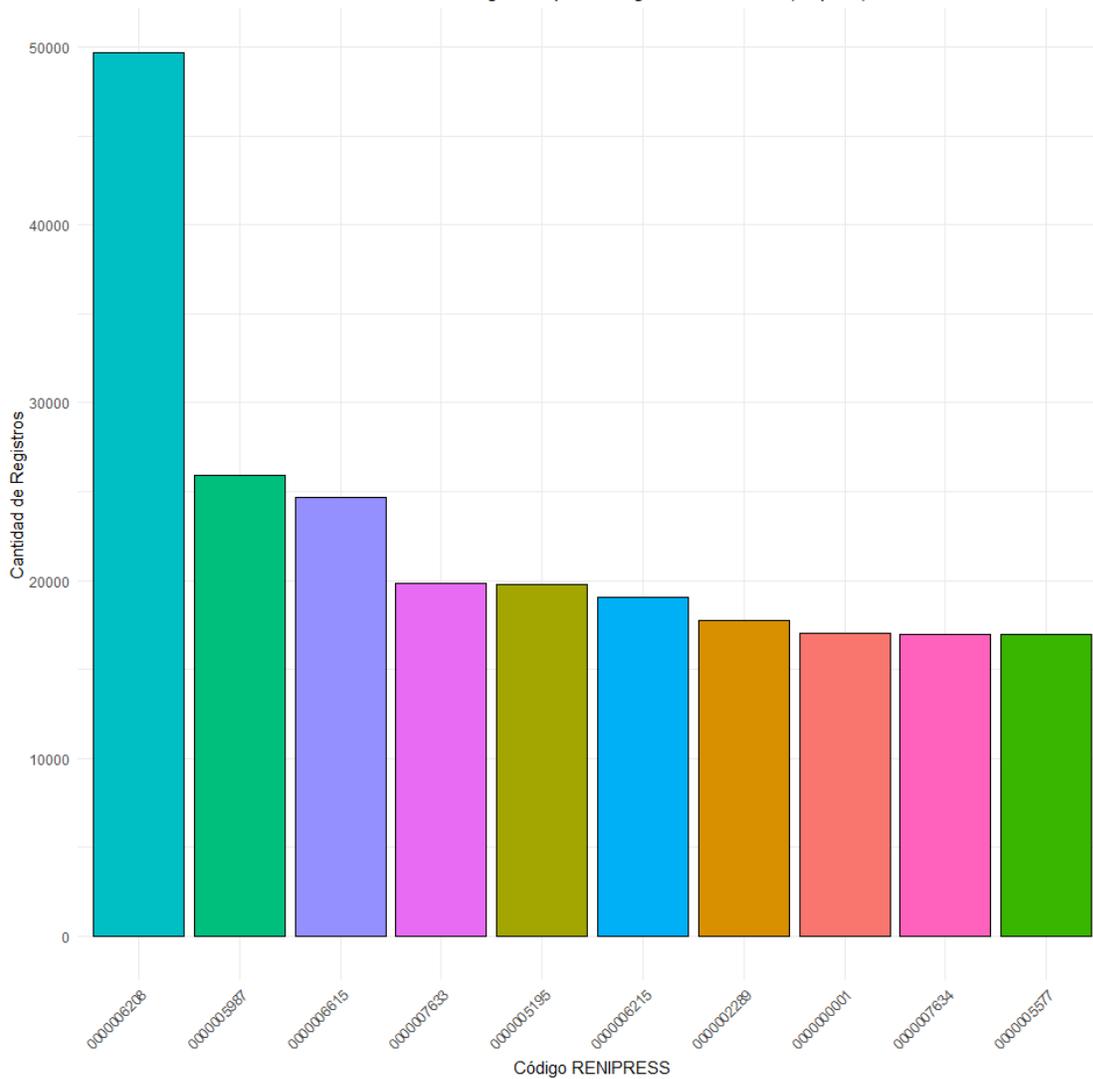


5.12. RENIPRESS

El código RENIPRESS, que identifica de manera única a cada establecimiento de salud en Perú, muestra que el “Instituto Nacional Materno Perinatal” tiene la mayor cantidad de registros, con 49,665 casos. Otros hospitales importantes, como el “Hospital María Auxiliadora” y el “Hospital Regional Docente Materno Infantil El Carmen”, también están entre los más representados. Esto indica que estas instituciones son puntos clave para la atención neonatal en el país. El gráfico de barras correspondiente visualiza los 10 códigos RENIPRESS más frecuentes, destacando a los principales establecimientos de salud en el manejo de estos casos.

```
## # A tibble: 10 × 2
##   RENIPRESS Count
##   <fct>      <int>
## 1 0000006208 49665
## 2 0000005987 25907
## 3 0000006615 24698
## 4 0000007633 19863
## 5 0000005195 19765
## 6 0000006215 19085
## 7 0000002289 17739
## 8 0000000001 17009
## 9 0000007634 16993
## 10 0000005577 16950
```

Distribución de Registros por Código RENIPRESS (Top 10)

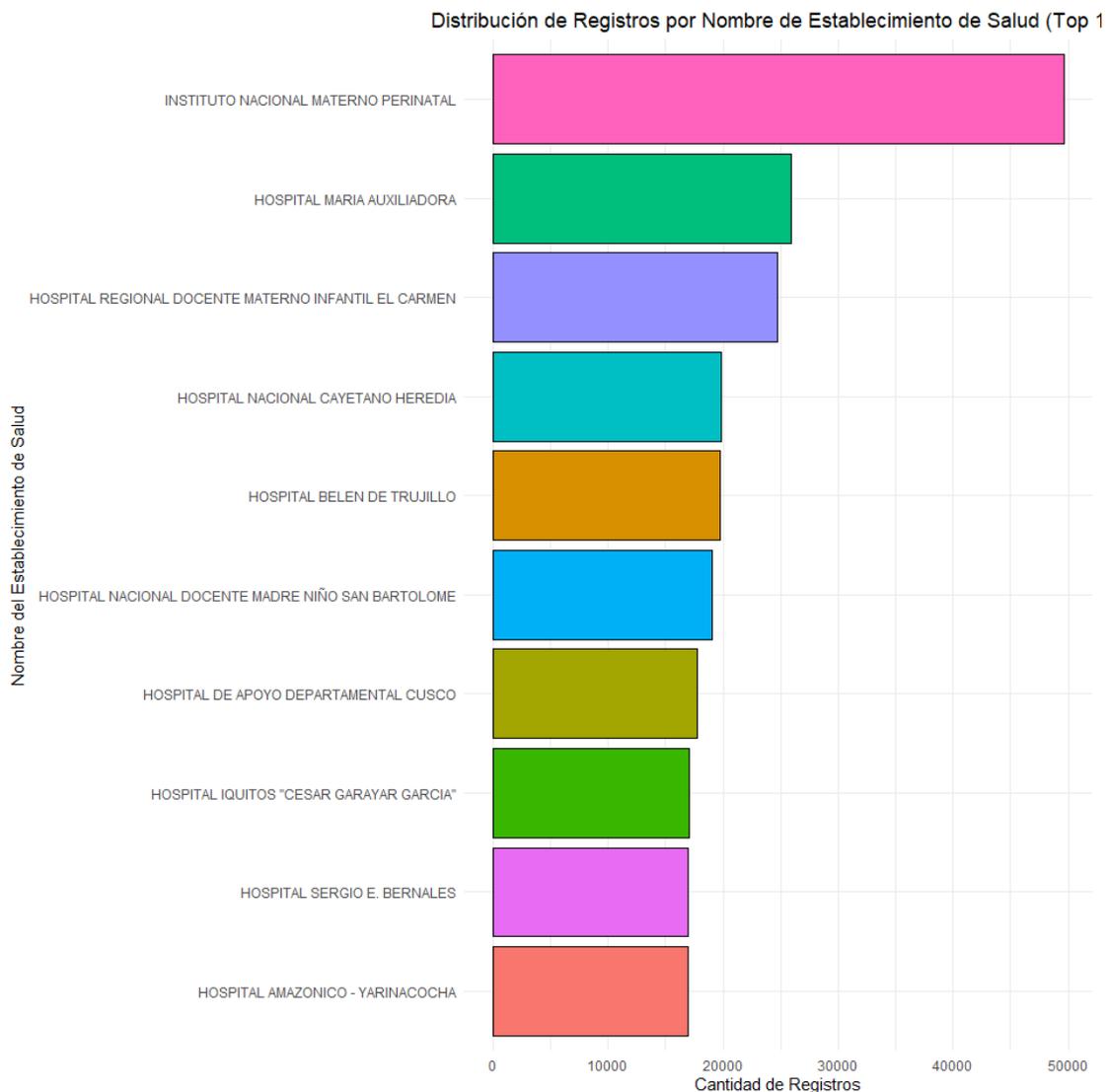


5.13. NOMBRE_EESS

Al analizar la distribución de registros por nombre del establecimiento de salud, se observa que el “Instituto Nacional Materno Perinatal” encabeza la lista, lo que refuerza su rol como líder en la atención materno-infantil en Perú. Otros hospitales importantes, como el “Hospital María Auxiliadora” y el “Hospital Nacional Cayetano Heredia”, también muestran un alto volumen de atención. Este análisis se presenta en un gráfico de barras horizontal, que facilita la comparación directa entre los 10 establecimientos más frecuentemente registrados.

```
## # A tibble: 10 × 2
##   NOMBRE_EESS          Count
##   <fct>              <int>
## 1 "INSTITUTO NACIONAL MATERNO PERINATAL" 49665
## 2 "HOSPITAL MARIA AUXILIADORA"          25907
## 3 "HOSPITAL REGIONAL DOCENTE MATERNO INFANTIL EL CARMEN" 24698
```

##	4	"HOSPITAL NACIONAL CAYETANO HEREDIA"	19863
##	5	"HOSPITAL BELEN DE TRUJILLO"	19765
##	6	"HOSPITAL NACIONAL DOCENTE MADRE NIÑO SAN BARTOLOME"	19085
##	7	"HOSPITAL DE APOYO DEPARTAMENTAL CUSCO"	17739
##	8	"HOSPITAL IQUITOS \"CESAR GARAYAR GARCIA\""	17009
##	9	"HOSPITAL SERGIO E. BERNALES"	16993
##	10	"HOSPITAL AMAZONICO - YARINACOCHA"	16950



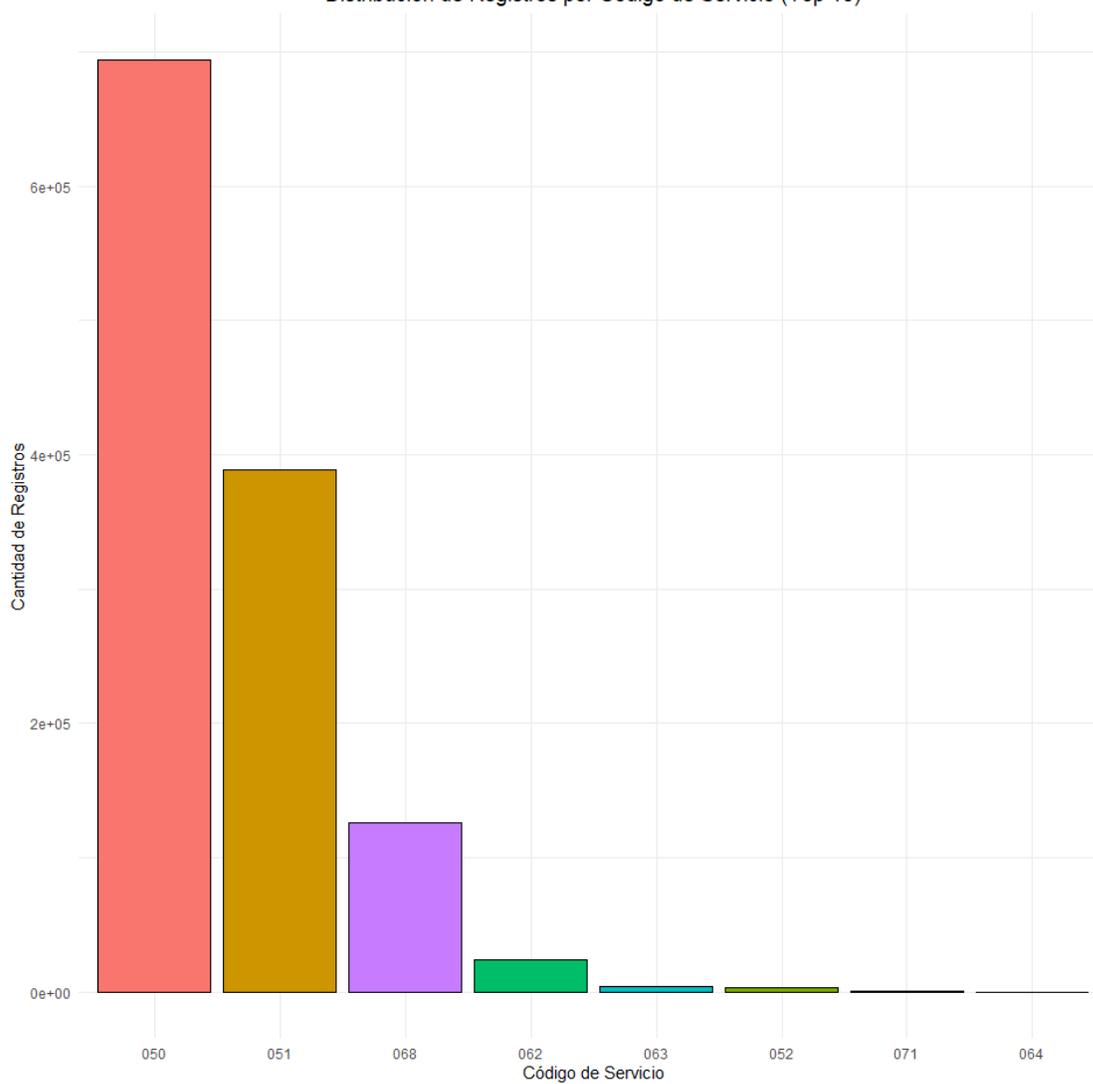
5.14. COD_SERVICIO

El análisis de los códigos de servicio revela que el código 050, correspondiente a la "Atención Inmediata del Recién Nacido Normal", es el más común, representando 693,737 registros. Esto indica que la mayoría de los registros están asociados con nacimientos sin complicaciones que requieren atención estándar. Otros códigos significativos incluyen 051, que corresponde al "Internamiento del Recién Nacido con Patología No Quirúrgica" y el 068, relacionado con "Internamiento en la Unidad de

Cuidados Intensivos”, mostrando la variedad de servicios brindados a recién nacidos. Estos hallazgos se presentan en un gráfico de barras que muestra los códigos de servicio más frecuentes.

```
## # A tibble: 8 × 2
##   COD_SERVICIO Count
##   <fct>         <int>
## 1 050           693737
## 2 051           388895
## 3 068           126197
## 4 062            24100
## 5 063             4211
## 6 052             2978
## 7 071              363
## 8 064              124
```

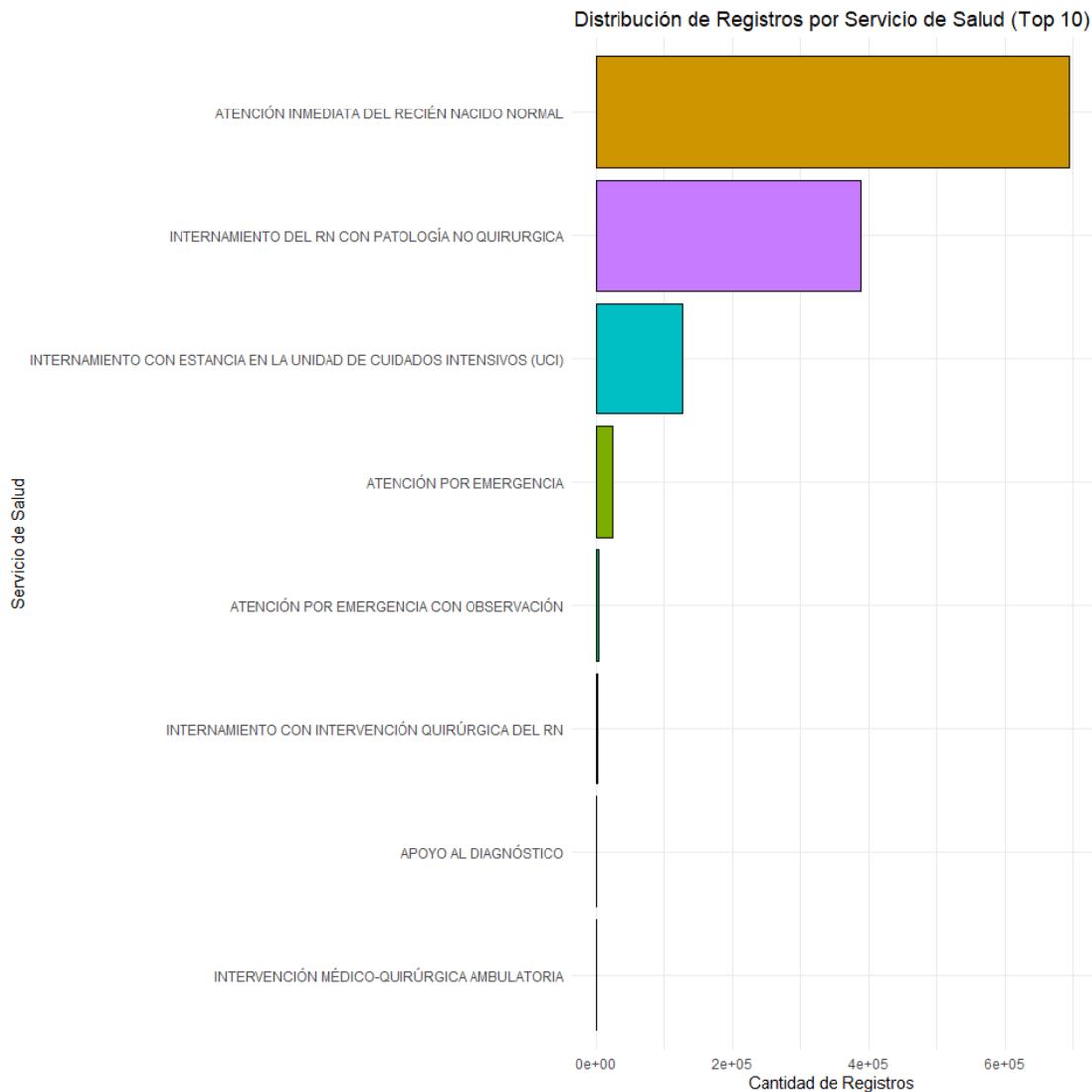
Distribución de Registros por Código de Servicio (Top 10)



5.15. SERVICIO

En cuanto a los servicios específicos brindados, la “Atención Inmediata del Recién Nacido Normal” es nuevamente el más representado, seguido por “Internamiento del Recién Nacido con Patología No Quirúrgica”. Estos servicios reflejan los tipos de cuidados más comunes en las unidades neonatales, desde la atención de rutina hasta el manejo de patologías no quirúrgicas. Servicios más especializados como “Internamiento con Estancia en la Unidad de Cuidados Intensivos” y “Atención por Emergencia” también están presentes, pero en menor número. El gráfico de barras horizontal ilustra la distribución de los registros por tipo de servicio de salud.

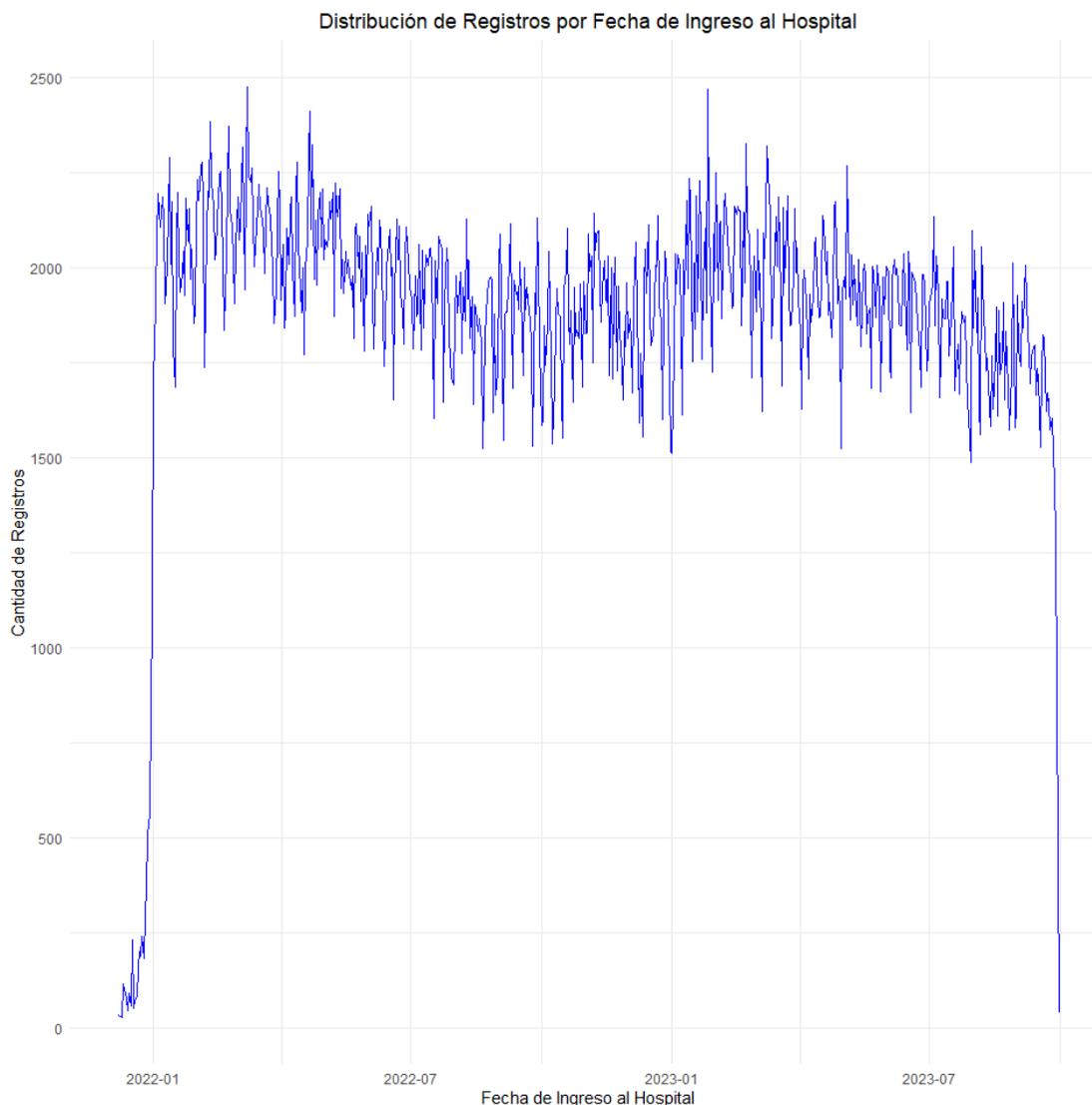
```
## # A tibble: 8 × 2
##   SERVICIO
##   <fct>
##   <int>
## 1 ATENCIÓN INMEDIATA DEL RECIÉN NACIDO NORMAL
693737
## 2 INTERNAMIENTO DEL RN CON PATOLOGÍA NO QUIRURGICA
388895
## 3 INTERNAMIENTO CON ESTANCIA EN LA UNIDAD DE CUIDADOS INTENSIVOS (UCI)
126197
## 4 ATENCIÓN POR EMERGENCIA
24100
## 5 ATENCIÓN POR EMERGENCIA CON OBSERVACIÓN
4211
## 6 INTERNAMIENTO CON INTERVENCIÓN QUIRÚRGICA DEL RN
2978
## 7 APOYO AL DIAGNÓSTICO
363
## 8 INTERVENCIÓN MÉDICO-QUIRÚRGICA AMBULATORIA
124
```



5.16. FECINGHOSP

El análisis de la variable FECINGHOSP, que registra la fecha de ingreso al hospital, muestra que los ingresos hospitalarios se distribuyen a lo largo del tiempo, con un aumento notable en ciertos períodos. El rango de fechas va desde el 7 de diciembre de 2021 hasta el 30 de septiembre de 2023, con una mediana de ingresos el 6 de noviembre de 2022. Estos datos sugieren una constante demanda de hospitalización para recién nacidos en este período. La distribución de los ingresos a lo largo del tiempo se presenta en un gráfico de líneas, destacando los momentos con mayor afluencia hospitalaria.

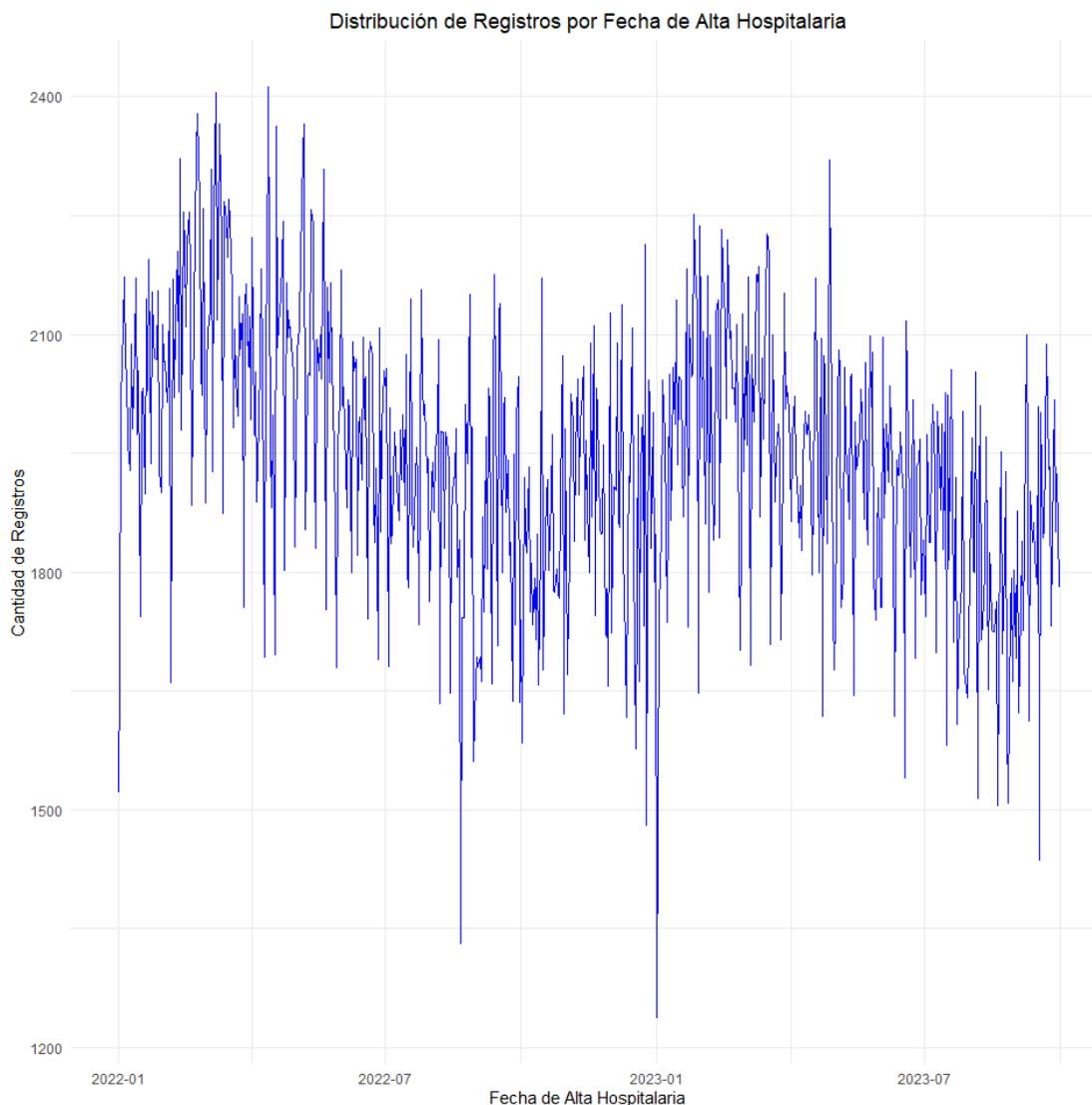
##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	"2021-12-07"	"2022-05-27"	"2022-11-06"	"2022-11-06"	"2023-04-14"	"2023-09-30"



5.17. FECALTAHOSP

La variable FECALTAHOSP, que registra la fecha de alta hospitalaria, tiene un patrón similar al de la fecha de ingreso, con registros que van desde el 1 de enero de 2022 hasta el 30 de septiembre de 2023. La mediana de las altas hospitalarias es el 10 de noviembre de 2022, lo que indica que las hospitalizaciones y altas siguen un curso paralelo a lo largo del tiempo. Este análisis sugiere que la duración de la estancia hospitalaria es generalmente corta, reflejando un manejo eficiente de los casos. El gráfico de líneas muestra la distribución de las altas a lo largo del tiempo.

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	"2022-01-01"	"2022-05-30"	"2022-11-10"	"2022-11-09"	"2023-04-17"	"2023-09-30"



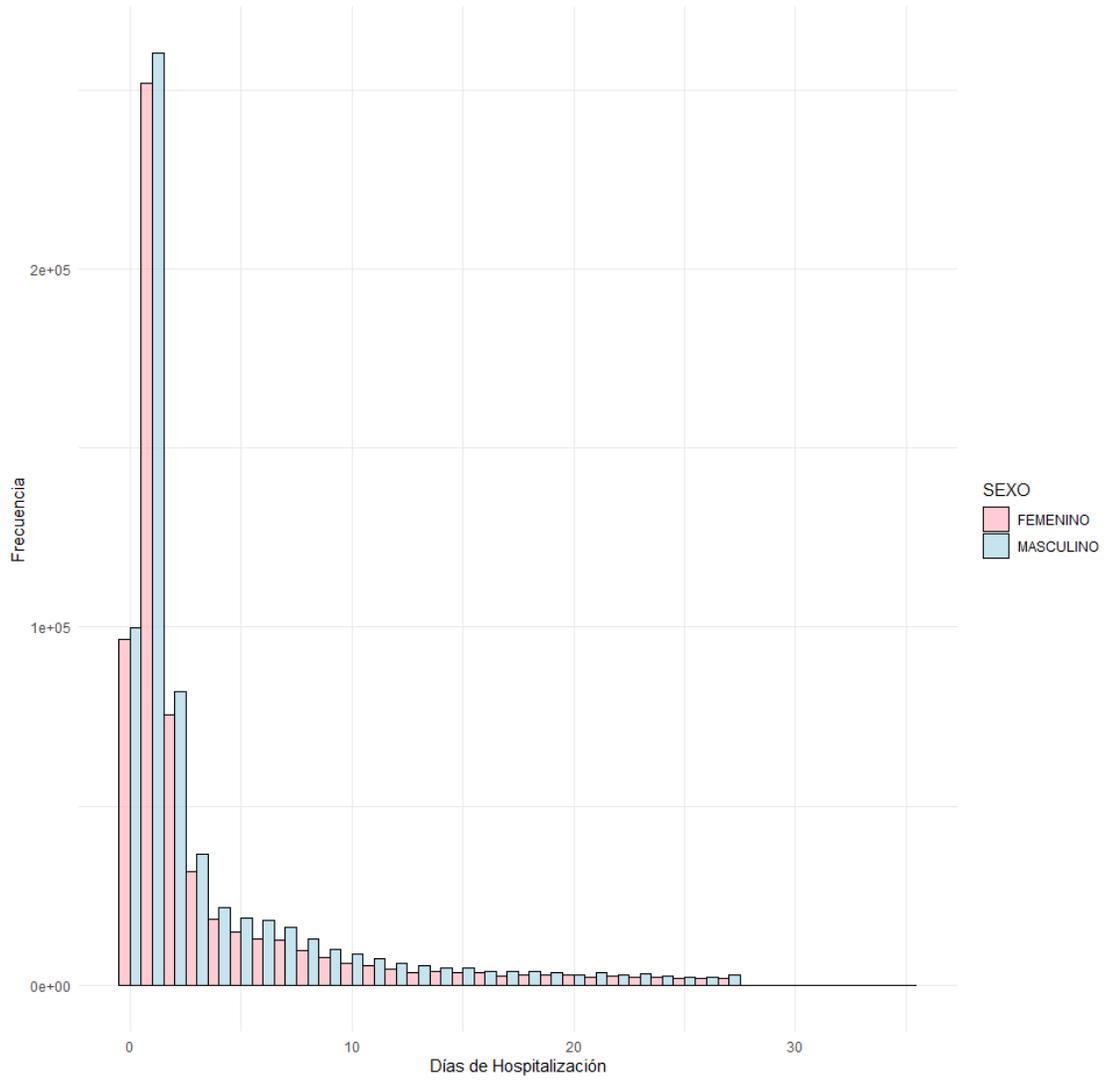
5.18. DIAS_HOSP

El análisis de la duración de la hospitalización muestra que la mayoría de los recién nacidos permanecen hospitalizados por un período corto, con una mediana de 1 día y una media de aproximadamente 3.46 días. La desviación estándar es de 5.25 días, lo que indica variabilidad en las estancias hospitalarias, dependiendo de la complejidad de los casos. Las hospitalizaciones más prolongadas alcanzan hasta 35 días. La distribución de los días de hospitalización se visualiza a través de histogramas y diagramas de caja (boxplots), tanto por sexo como por nivel de establecimiento de salud, destacando cómo varían estos tiempos según diferentes factores.

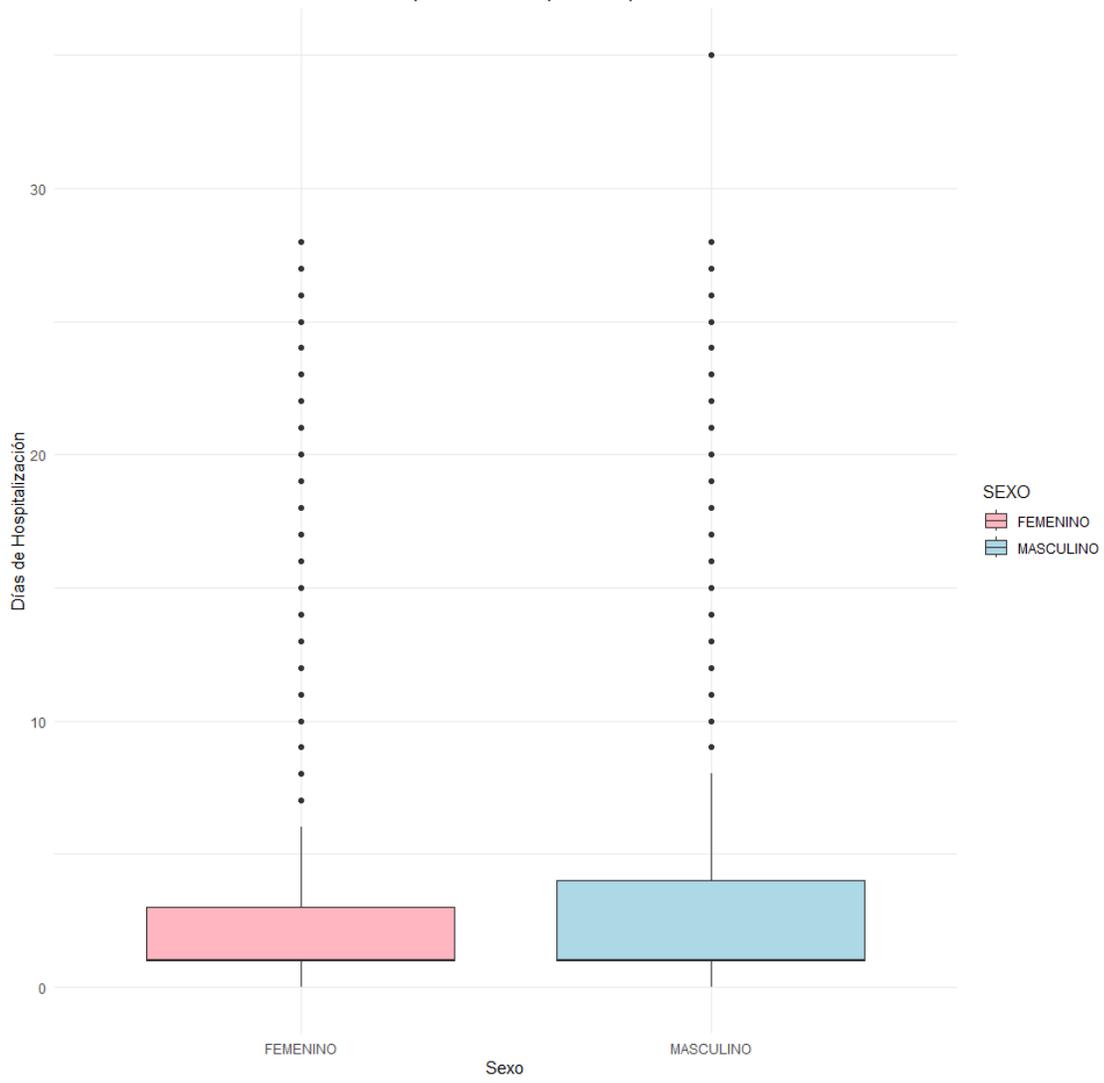
```
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000  1.000   1.000   3.464  3.000   35.000
```

```
## [1] "Desviación Estándar de DIAS_HOSP: 5.24586577857741"
```

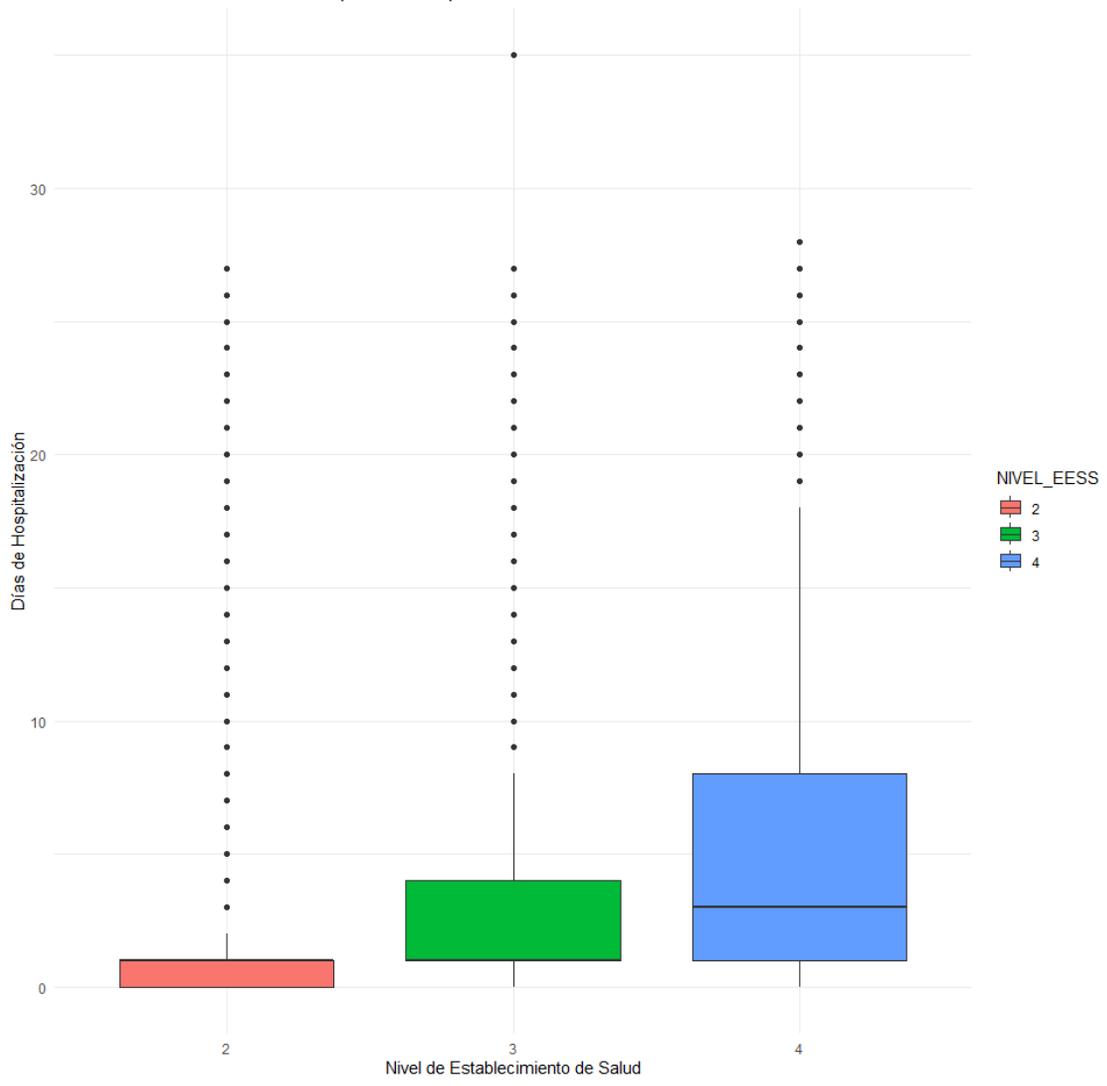
Distribución de los Días de Hospitalización por Sexo



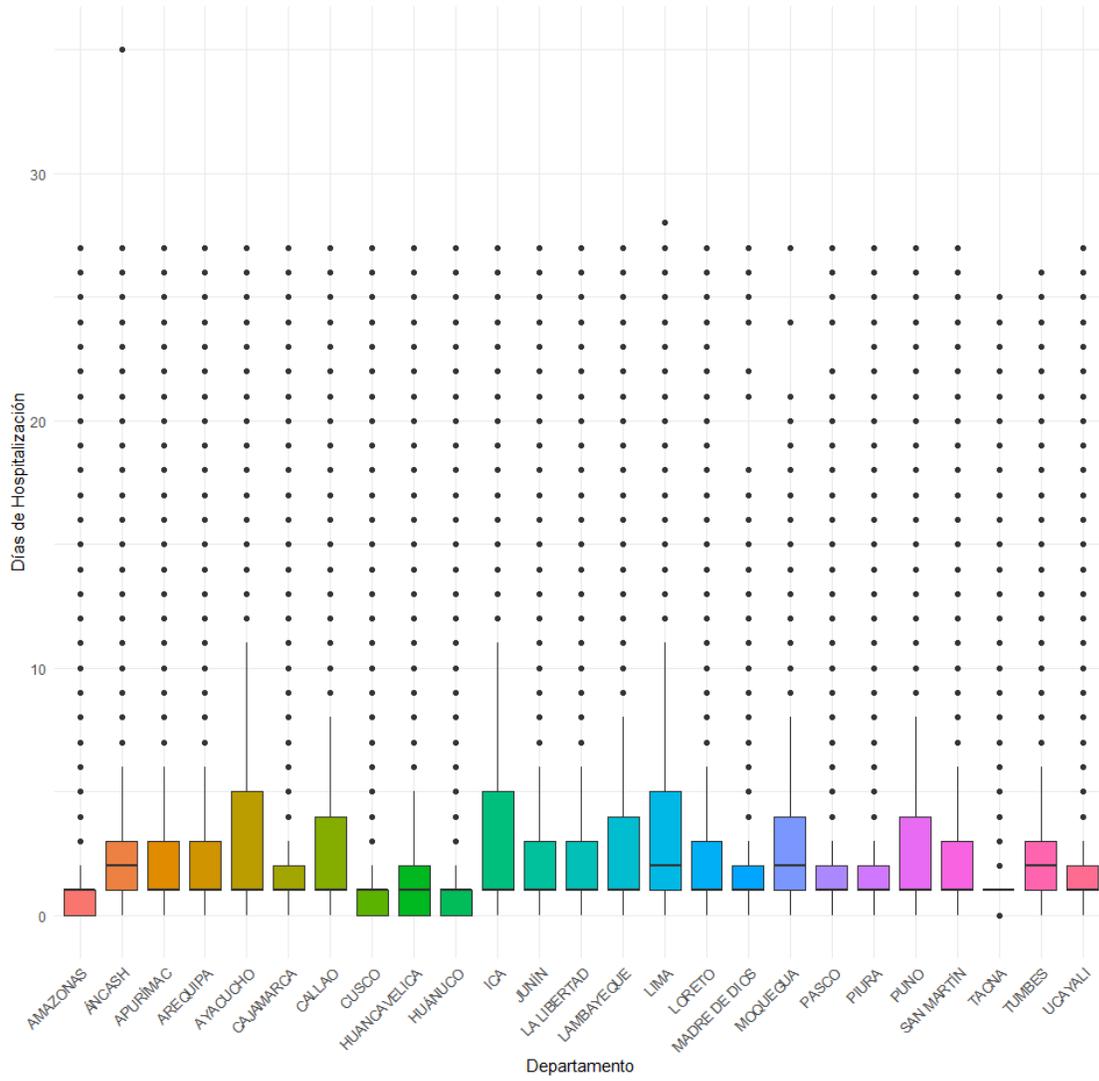
Días de Hospitalización Separados por Sexo

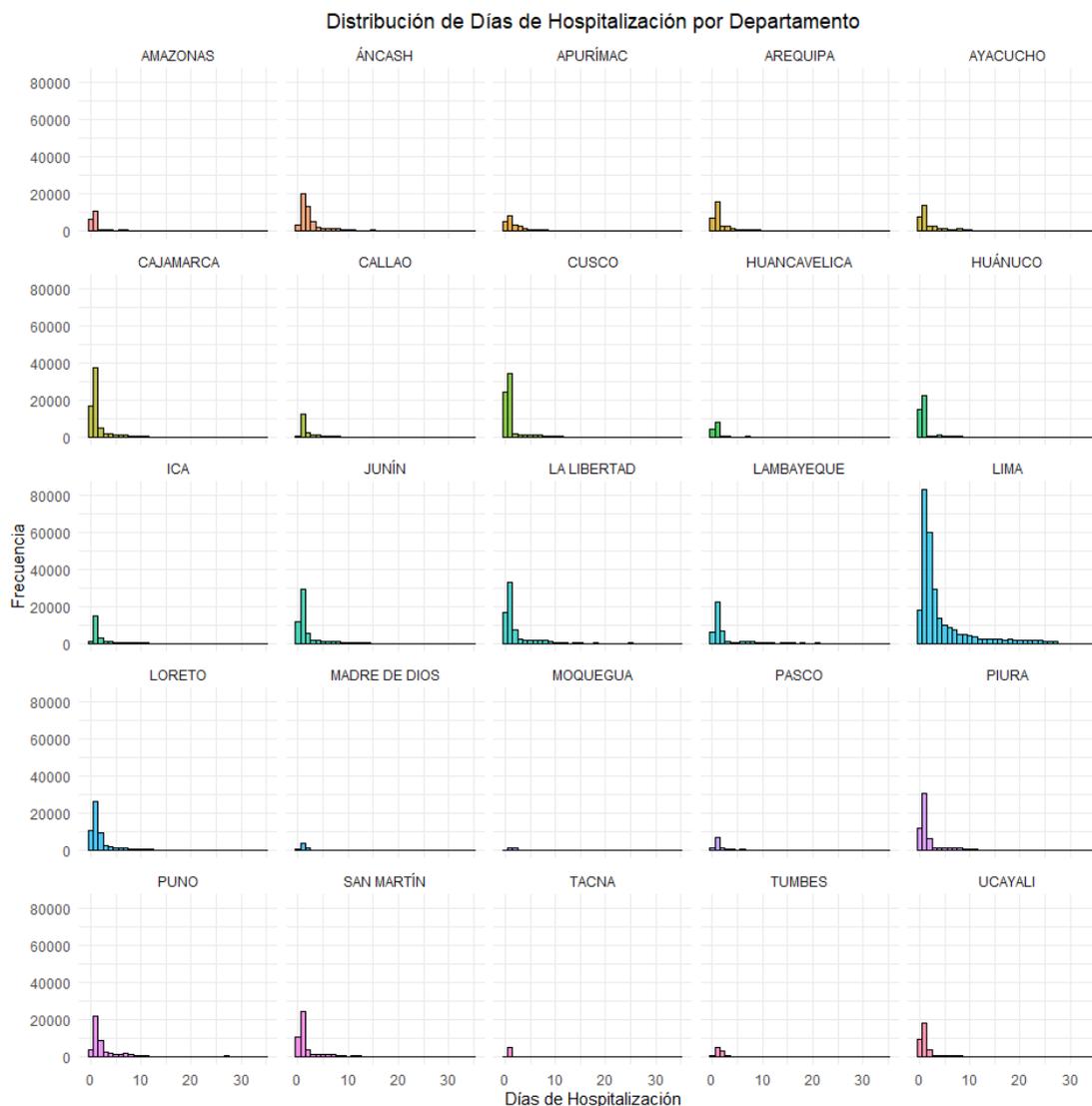


Días de Hospitalización por Nivel de Establecimiento de Salud



Días de Hospitalización por Departamento





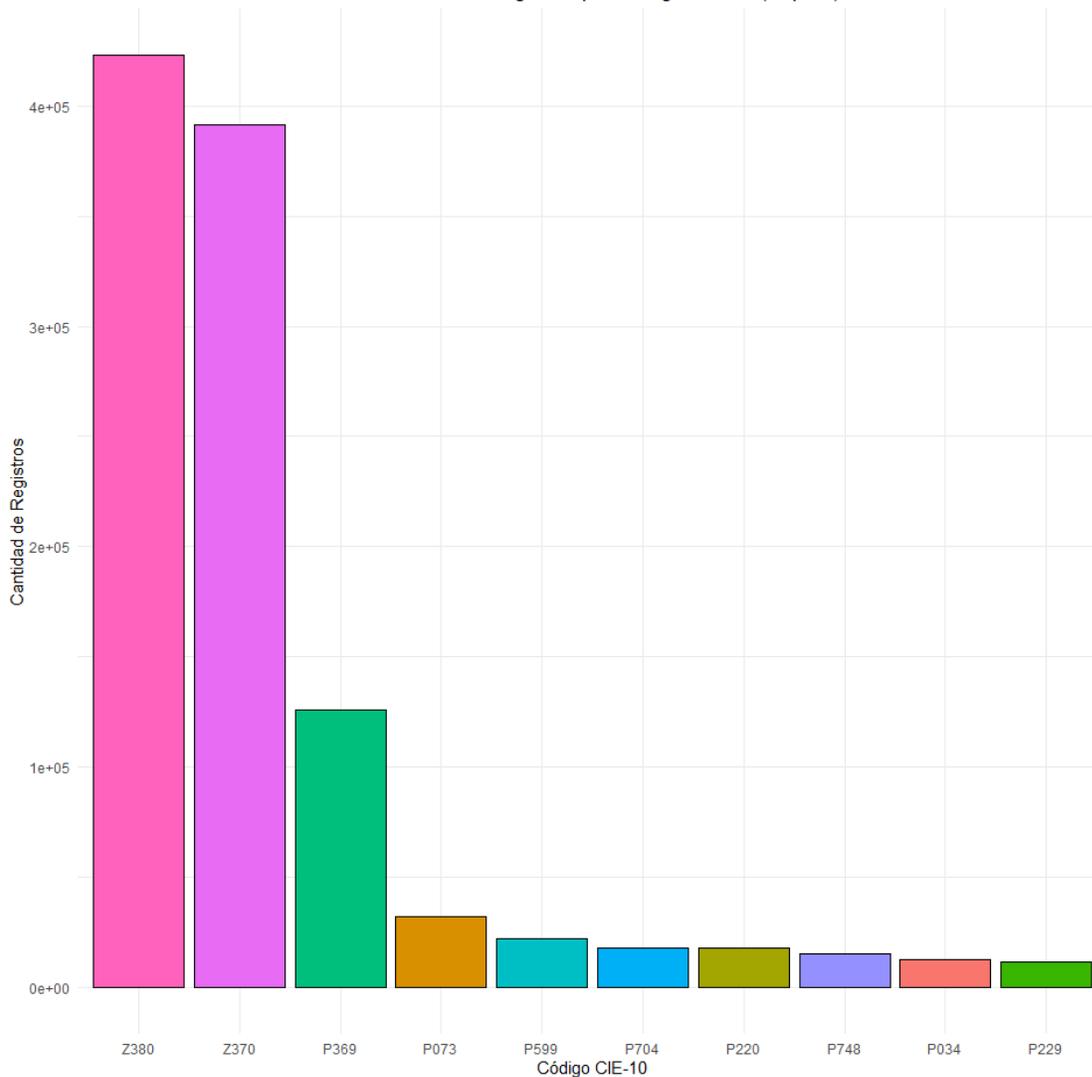
5.19. CODDIA

La variable CODDIA, que corresponde al código CIE-10 del diagnóstico principal, revela que los códigos más comunes son Z380 (“Producto Único Nacido en el Hospital”) y Z370 (“Nacido Vivo Único”), con 423,001 y 391,560 registros respectivamente. Estos códigos reflejan diagnósticos rutinarios asociados con nacimientos sin complicaciones. Otros códigos, como P369 (“Sepsis Bacteriana del Recién Nacido No Especificada”) y P073 (“Otros Recién Nacidos Pretermino”), indican casos más complejos que requieren atención médica especializada. La distribución de los diagnósticos se presenta en un gráfico de barras que muestra los 10 códigos CIE-10 más frecuentes.

```
## # A tibble: 10 × 2
##   CODDIA Count
##   <fct>   <int>
## 1 Z380    423001
```

##	2	Z370	391560
##	3	P369	125998
##	4	P073	31855
##	5	P599	21861
##	6	P704	17928
##	7	P220	17810
##	8	P748	15078
##	9	P034	12386
##	10	P229	11718

Distribución de Registros por Código CIE-10 (Top 10)

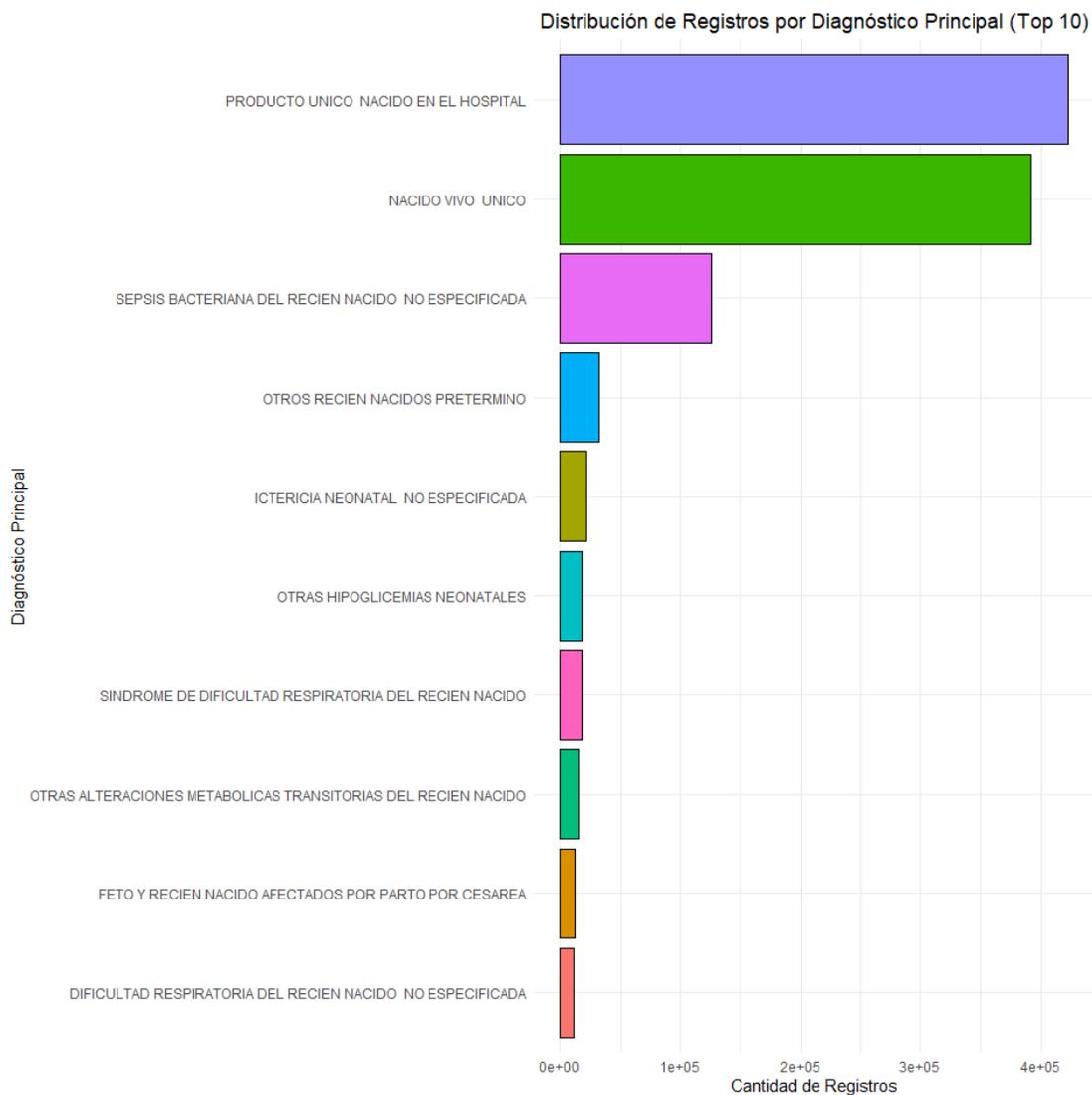


5.20. C10_NOMBRE

El análisis de la variable C10_NOMBRE, que describe el diagnóstico principal, reafirma que “Producto Único Nacido en el Hospital” es el diagnóstico más común, seguido por “Nacido Vivo Único” y “Sepsis Bacteriana del Recién Nacido No Especificada”. Estos datos reflejan los principales motivos de atención en las unidades neonatales, con un

énfasis en la atención de nacimientos normales, pero también destacando la presencia de infecciones neonatales. El gráfico de barras horizontal permite visualizar claramente la frecuencia de los principales diagnósticos en la atención neonatal.

```
## # A tibble: 10 × 2
##   C10_NOMBRE
Count
##   <fct>
<int>
## 1 PRODUCTO UNICO  NACIDO EN EL HOSPITAL
423001
## 2 NACIDO VIVO  UNICO
391560
## 3 SEPSIS BACTERIANA DEL RECIEN NACIDO  NO ESPECIFICADA
125998
## 4 OTROS RECIEN NACIDOS PRETERMINO
31855
## 5 ICTERICIA NEONATAL  NO ESPECIFICADA
21861
## 6 OTRAS HIPOGLICEMIAS NEONATALES
17928
## 7 SINDROME DE DIFICULTAD RESPIRATORIA DEL RECIEN NACIDO
17810
## 8 OTRAS ALTERACIONES METABOLICAS TRANSITORIAS DEL RECIEN NACIDO
15078
## 9 FETO Y RECIEN NACIDO AFECTADOS POR PARTO POR CESAREA
12386
## 10 DIFICULTAD RESPIRATORIA DEL RECIEN NACIDO  NO ESPECIFICADA
11718
```

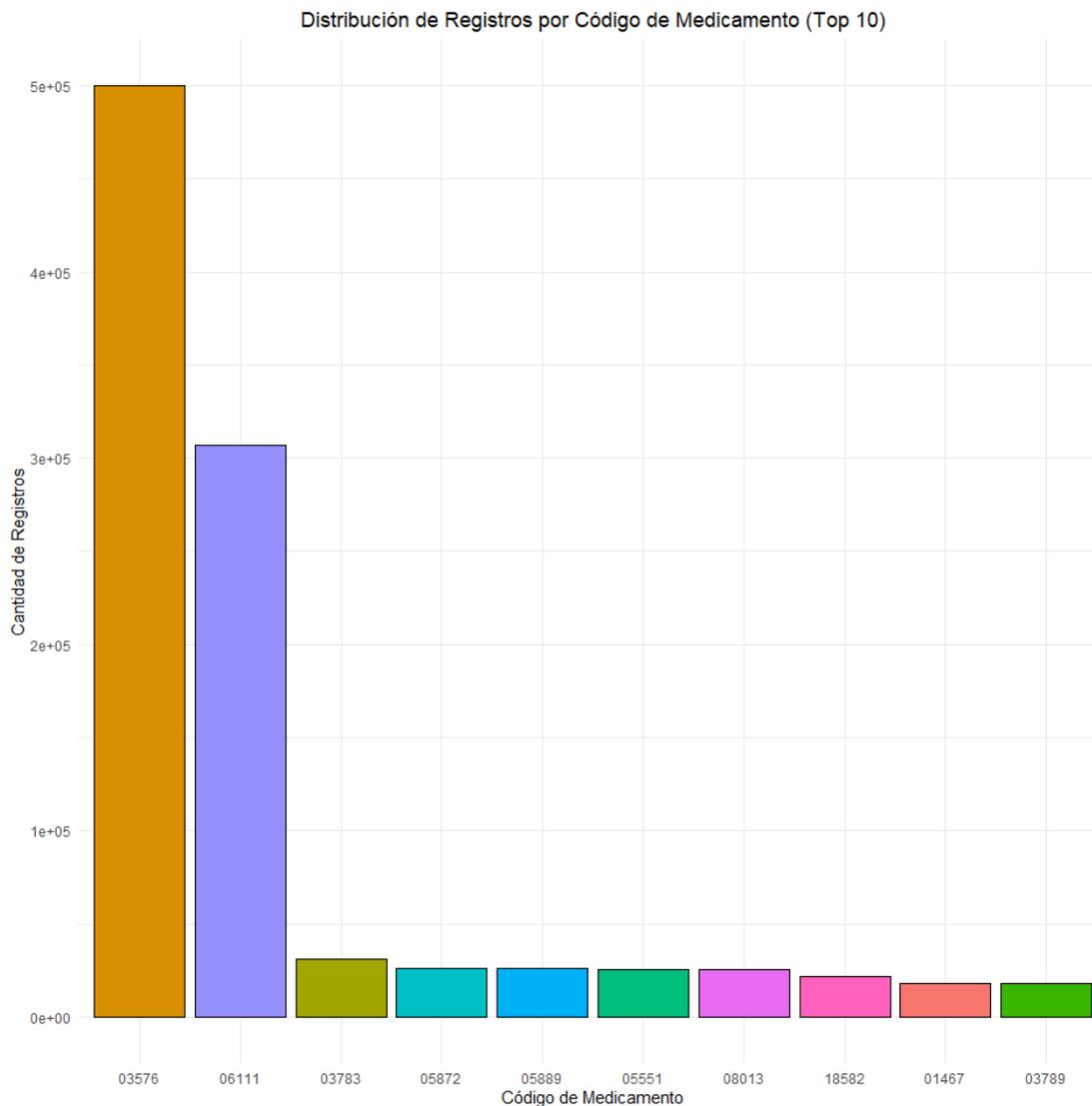


5.21. COD_MEDICAMENTO

El análisis de los códigos de medicamentos muestra que el código 03576 es el más frecuente, con 499,804 registros, seguido por el código 06111, con 307,052 registros. Estos códigos representan los medicamentos más comúnmente administrados en los establecimientos de salud. La lista incluye otros medicamentos que, aunque menos frecuentes, siguen siendo significativos en términos de su uso. El gráfico de barras muestra la distribución de los 10 códigos de medicamentos más comunes.

```
## # A tibble: 10 × 2
##   COD_MEDICAMENTO Count
##   <fct>           <int>
## 1 03576           499804
## 2 06111           307052
## 3 03783           30835
## 4 05872           26209
```

##	5	05889	26052
##	6	05551	25677
##	7	08013	25273
##	8	18582	21565
##	9	01467	18158
##	10	03789	17712



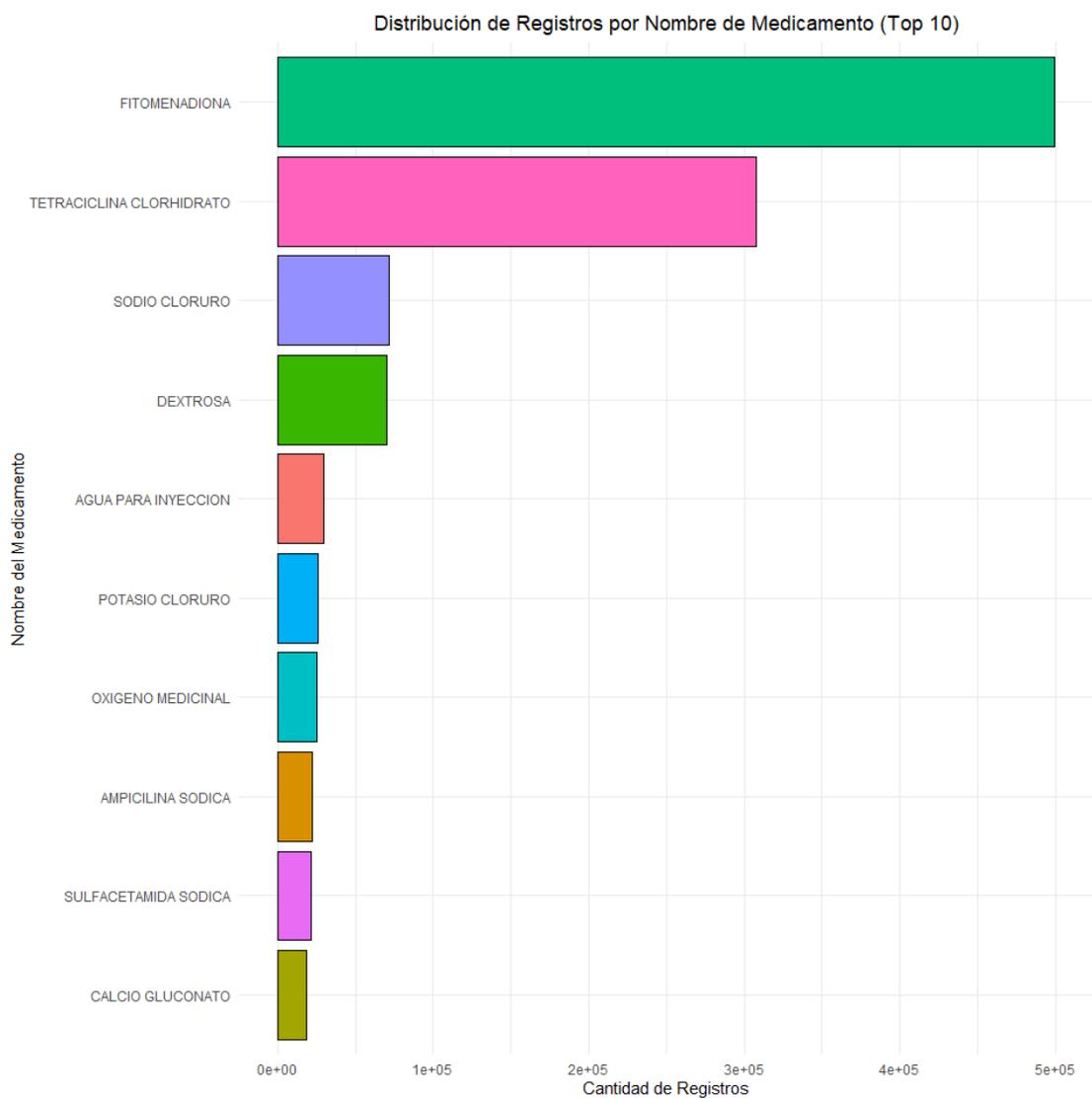
5.22. NOMBRE_MEDICAMENTO

En términos de nombres de medicamentos, la **Fitomenadiona** es el medicamento más registrado, con 499,811 usos. Esto sugiere su alta demanda y uso en tratamientos neonatales, probablemente debido a su uso en la prevención y tratamiento de la enfermedad hemorrágica del recién nacido. Otros medicamentos frecuentes incluyen **Tetraciclina Clorhidrato** y **Sodio Cloruro**, que son esenciales en el manejo de infecciones y equilibrio de electrolitos. El gráfico de barras horizontal ilustra la distribución de los registros por los nombres de medicamentos más comunes.

```

## # A tibble: 10 × 2
##   NOMBRE_MEDICAMENTO      Count
##   <fct>                 <int>
## 1 FITOMENADIONA         499811
## 2 TETRACICLINA CLORHIDRATO 307819
## 3 SODIO CLORURO         71600
## 4 DEXTROSA              69867
## 5 AGUA PARA INYECCION   29706
## 6 POTASIO CLORURO       25700
## 7 OXIGENO MEDICINAL     25173
## 8 AMPICILINA SODICA     22296
## 9 SULFACETAMIDA SODICA  21565
## 10 CALCIO GLUCONATO     18158

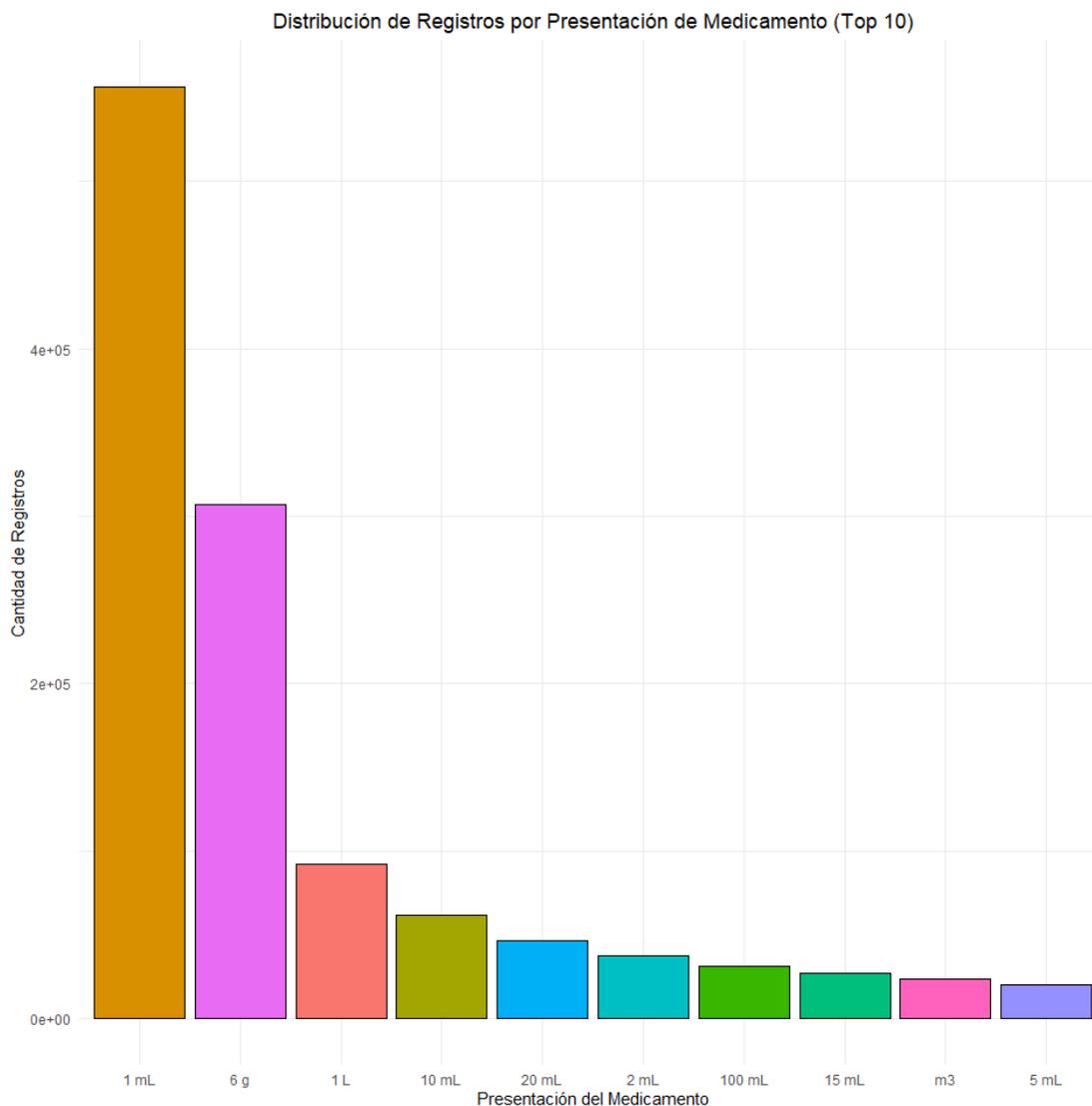
```



5.23. PRESENTACION

La presentación más frecuente de medicamentos es la de **1 mL**, con 556,279 registros, seguida de **6 g**, con 307,053 registros. Estas presentaciones reflejan las dosis comúnmente administradas a los recién nacidos, indicando una preferencia por volúmenes pequeños y precisos en la administración de medicamentos. El gráfico de barras vertical permite visualizar las 10 presentaciones de medicamentos más comunes en el conjunto de datos.

```
## # A tibble: 10 × 2
##   PRESENTACION Count
##   <fct>          <int>
## 1 1 mL           556279
## 2 6 g           307053
## 3 1 L           92073
## 4 10 mL         61370
## 5 20 mL         46432
## 6 2 mL          37232
## 7 100 mL        31037
## 8 15 mL         26924
## 9 m3            23677
## 10 5 mL         20014
```

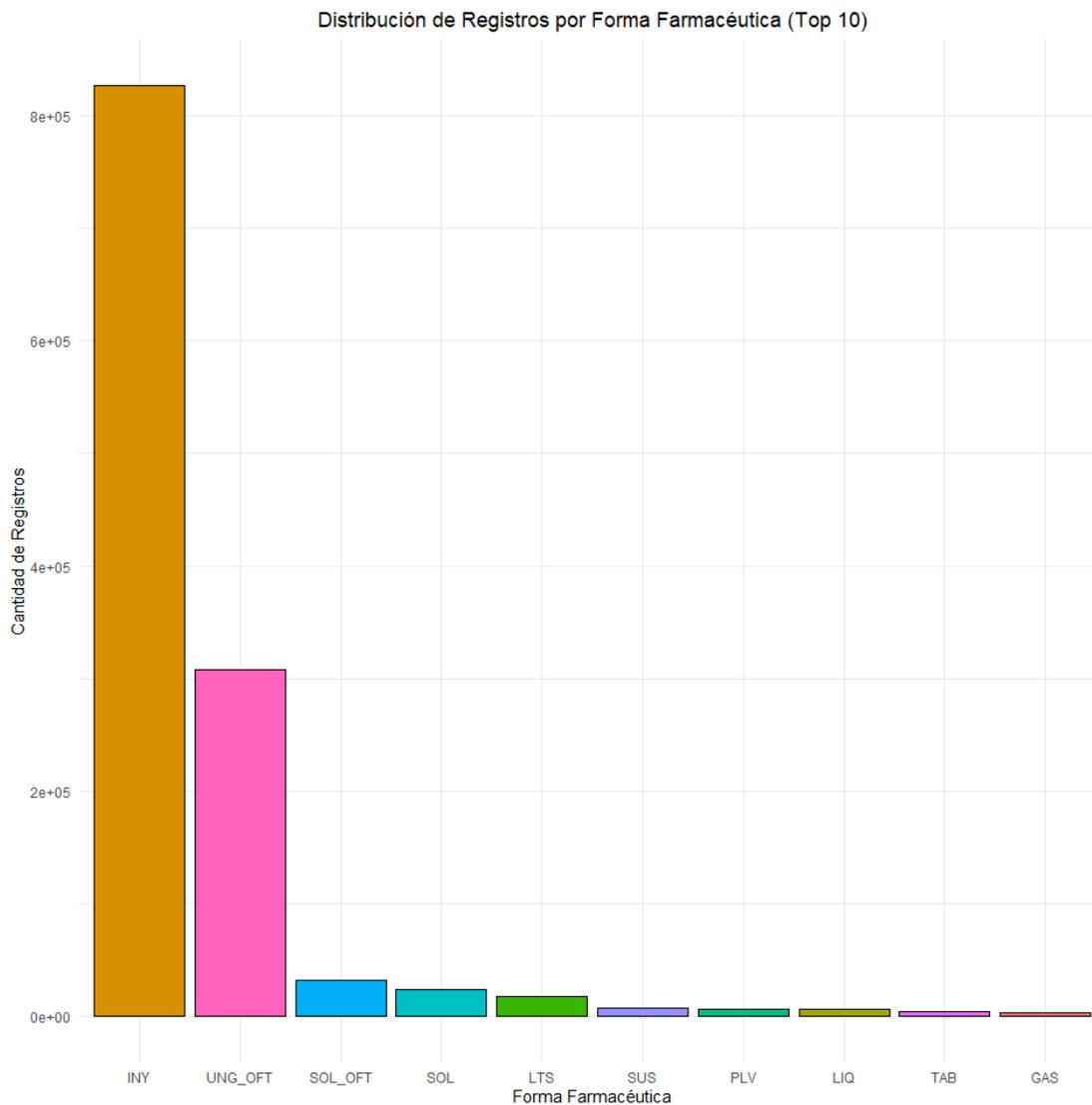


5.24. FORMA_FARMACEUTICA

La **forma farmacéutica** más frecuente es la **inyectable (INY)**, con 826,740 registros, lo que subraya la prevalencia de los medicamentos administrados por vía inyectable en el tratamiento neonatal. Otras formas, como **Ungüento Oftálmico (UNG_OFT)** y **Solución Oftálmica (SOL_OFT)**, también son comunes, reflejando las necesidades de cuidado ocular en recién nacidos. El gráfico de barras vertical muestra la distribución de las 10 formas farmacéuticas más utilizadas.

```
## # A tibble: 10 × 2
##   FORMA_FARMACEUTICA Count
##   <fct>                <int>
## 1 INY                  826740
## 2 UNG_OFT              307822
## 3 SOL_OFT              321730
## 4 SOL                  233140
```

##	5	LTS	16859
##	6	SUS	7204
##	7	PLV	6265
##	8	LIQ	5922
##	9	TAB	4299
##	10	GAS	3294

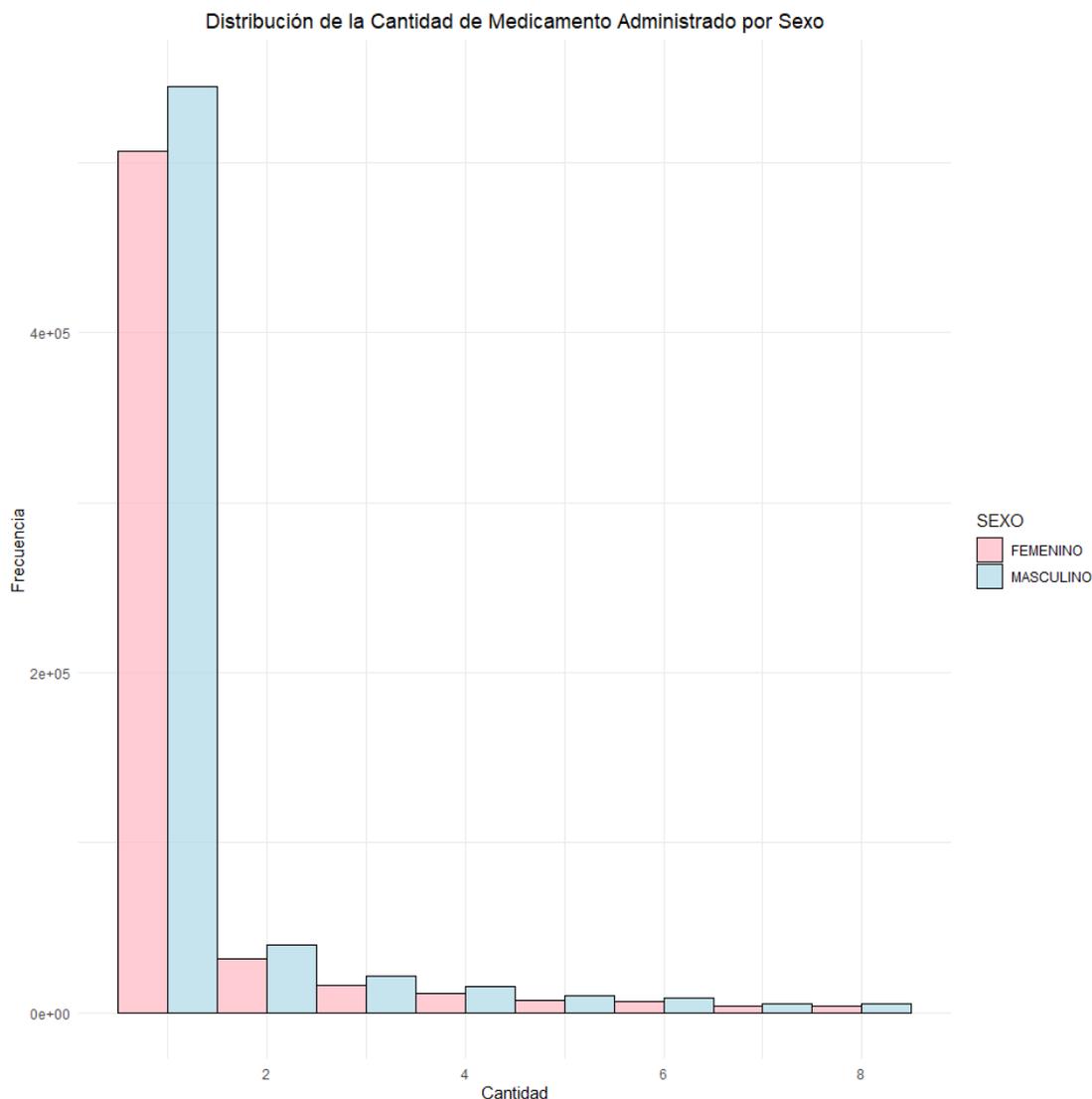


5.25. CANTIDAD

El análisis de la **cantidad de medicamento administrado** muestra que la mayoría de los registros corresponden a una única unidad de medicamento, con una mediana de 1 y una media ligeramente superior a 1 (1.403). La distribución tiene una desviación estándar de 1.18, lo que indica una variabilidad limitada en las dosis administradas. Este comportamiento se muestra en un histograma, que también presenta las diferencias de distribución por sexo.

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.000  1.000   1.000   1.403  1.000   8.000
```

```
## [1] "Desviación Estándar de CANTIDAD: 1.18361390953672"
```



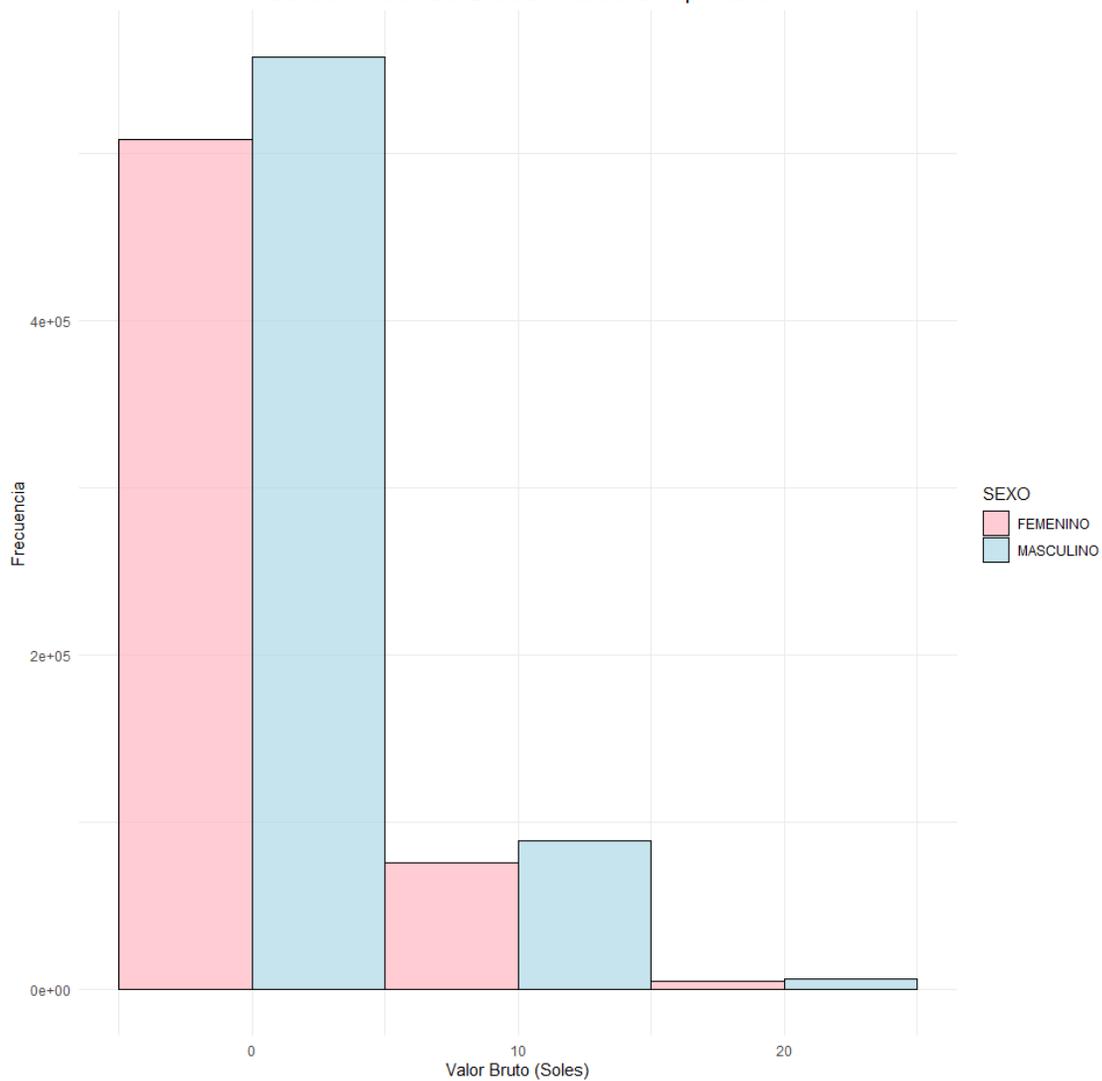
5.26. VALOR_BRUTO

El **valor bruto** de los medicamentos administrados tiene una mediana de 0.69 soles y una media de 2.14 soles, con una desviación estándar de 3.78 soles. Esto indica que la mayoría de los medicamentos tienen un valor relativamente bajo, aunque hay algunos casos con valores más altos, hasta un máximo de 17.213 soles. Los histogramas muestran cómo se distribuye este valor por sexo y por departamento, revelando variaciones significativas en los costos asociados con la administración de medicamentos.

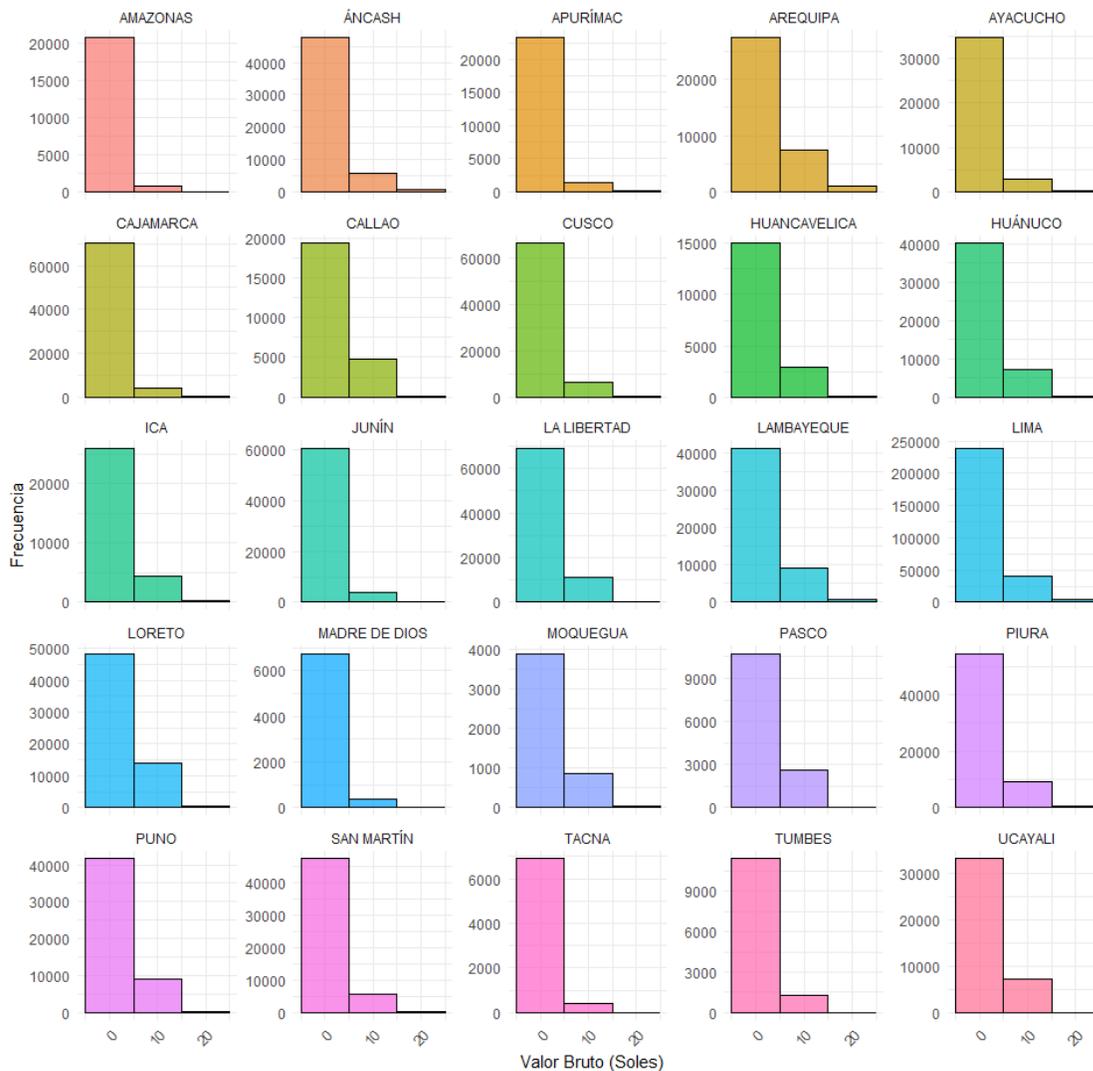
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.000  0.000   0.690   2.143  1.840  17.213
```

```
## [1] "Desviación Estándar de VALOR_BRUTO: 3.77795350756339"
```

Distribución del Valor Bruto de Medicamento por Sexo



Distribución del Valor Bruto de Medicamento por Departamento



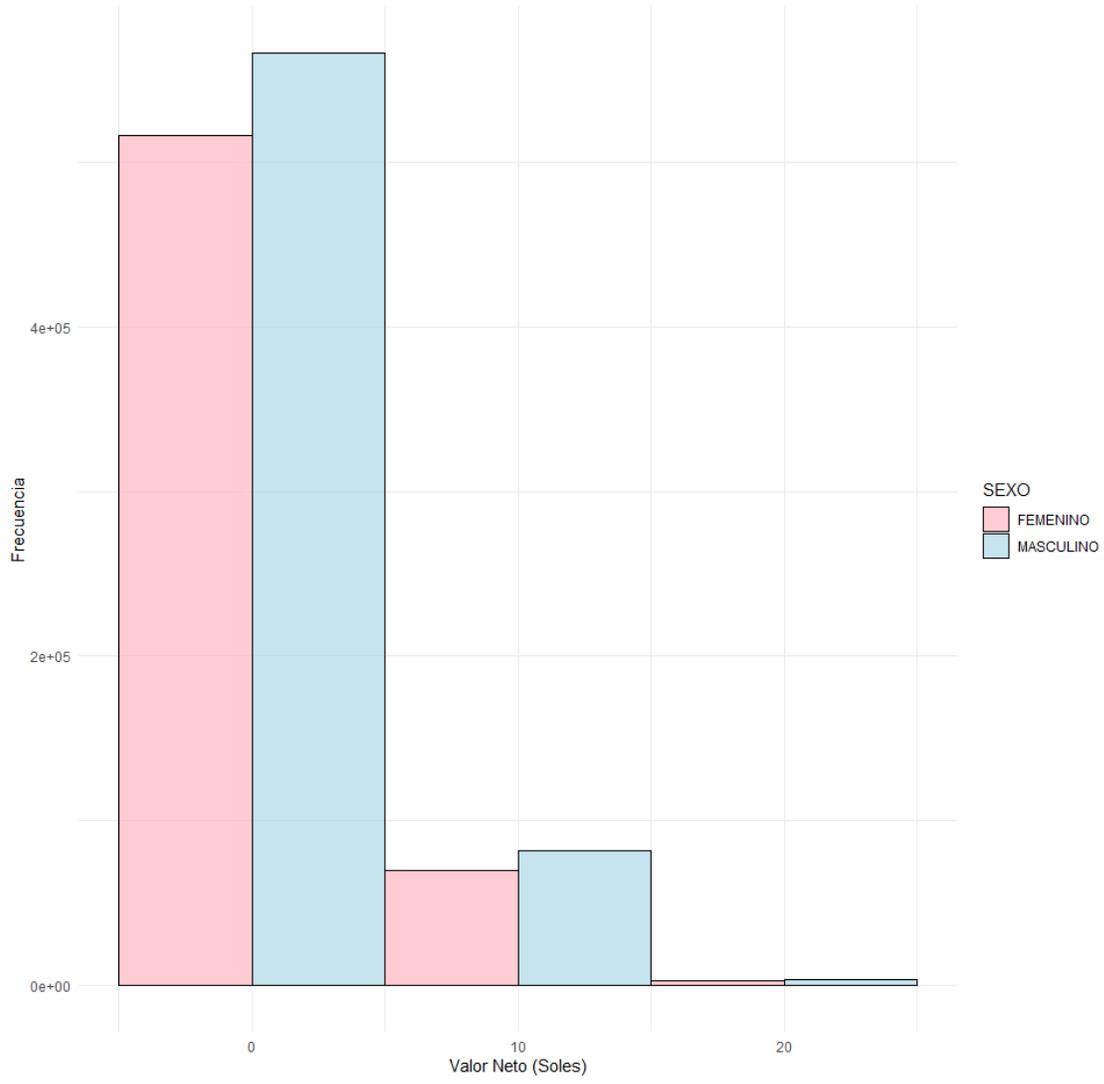
5.27. VALOR_NETO

El **valor neto** de los medicamentos, tras aplicar posibles descuentos u otras consideraciones, presenta una mediana de 0.4 soles y una media de 1.93 soles, con una desviación estándar de 3.59 soles. Al igual que con el valor bruto, la mayoría de los medicamentos tienen un costo bajo, pero existen valores más altos que llegan hasta 16.5 soles. Los gráficos muestran la distribución del valor neto tanto por sexo como por departamento, proporcionando una visión detallada de las variaciones en los costos de los medicamentos.

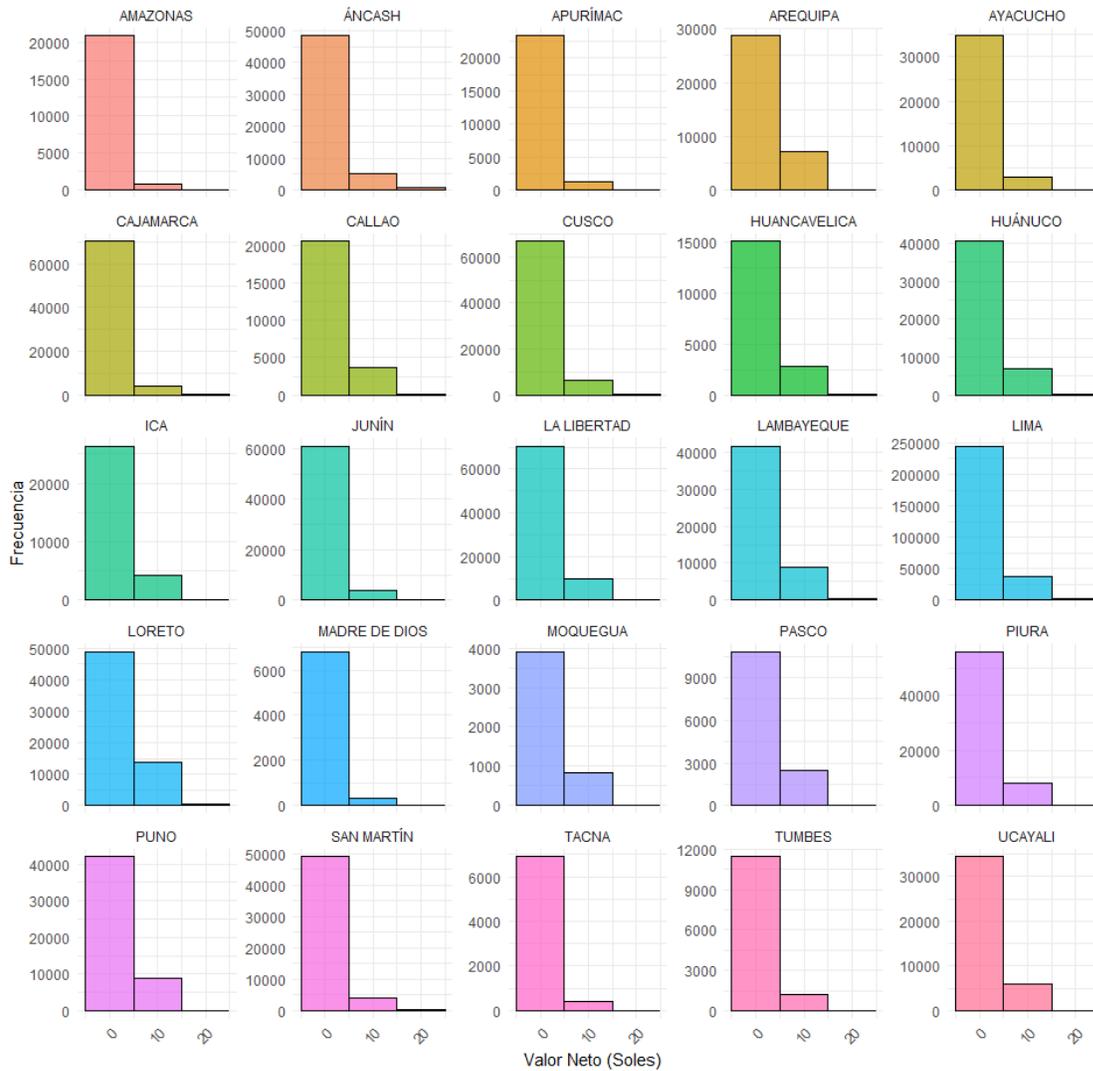
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.000  0.000   0.400   1.931  1.625  16.500

## [1] "Desviación Estándar de VALOR_NETO: 3.59323068831439"
```

Distribución del Valor Neto de Medicamento por Sexo



Distribución del Valor Neto de Medicamento por Departamento



6. Análisis

6.1. Análisis Geográfico

Este mapa muestra la distribución de los registros de atención de salud en los diferentes departamentos del Perú. Cada departamento está coloreado según la cantidad de registros, con un gradiente de color que va desde los departamentos con menor cantidad de registros (colores más claros) hasta aquellos con mayor cantidad de registros (colores más oscuros). Esto permite identificar visualmente las áreas con mayor actividad médica, destacando regiones como Lima, que suele concentrar una gran parte de los servicios de salud.

Mapa de los Departamentos de Perú con Nombres



7. Principales hallazgos

- **Distribución Geográfica:** La mayor parte de los registros de atención de salud se concentra en el departamento de Lima, seguido de otros departamentos como La Libertad, Cajamarca y Cusco. Esto refleja una alta demanda de servicios de salud en las áreas más pobladas y urbanizadas del país.
- **Servicios de Salud Predominantes:** Los servicios de atención inmediata del recién nacido normal y el internamiento del recién nacido con patología no quirúrgica son los más frecuentes. Esto sugiere una alta incidencia de nacimientos y atención neonatal en los establecimientos de salud.
- **Diagnósticos Frecuentes:** Los diagnósticos más comunes están relacionados con el nacimiento en el hospital y sepsis bacteriana del recién nacido. Estos

resultados indican que la atención perinatal y neonatal son áreas críticas de los servicios de salud en el Perú.

- **Medicamentos Más Usados:** La Fitomenadiona y la Tetraciclina Clorhidrato son los medicamentos más frecuentemente administrados, lo que apunta a su uso generalizado en los tratamientos perinatales y neonatales.
- **Variación en la Hospitalización:** La duración de la hospitalización muestra variabilidad considerable, con una media de 3.5 días y una desviación estándar alta. Esto puede indicar diferencias en la gravedad de las condiciones tratadas o en los protocolos de atención.

8. Limitaciones

- **Cobertura de Datos:** La base de datos podría no representar completamente todas las regiones del Perú, especialmente las áreas rurales donde la infraestructura de salud es más limitada. Esto podría sesgar los resultados hacia las zonas más urbanizadas y mejor atendidas.
- **Calidad de los Datos:** La calidad de los datos depende de la precisión y consistencia en la recolección de información en los establecimientos de salud. Errores en la codificación de diagnósticos, servicios o medicamentos podrían afectar la exactitud de los hallazgos.
- **Datos Anónimos:** La anonimización de datos es esencial para la privacidad, pero puede limitar la capacidad de realizar análisis detallados a nivel individual o de identificar patrones específicos en subpoblaciones.
- **Temporalidad de los Datos:** Los datos analizados cubren un período de tiempo limitado. Cambios recientes en políticas de salud o prácticas médicas podrían no estar reflejados en los hallazgos.

9. Conclusiones

- **Concentración en Lima:** La concentración de registros en Lima resalta la centralización de servicios de salud en la capital, lo que puede indicar desigualdades en la accesibilidad a servicios de salud en otras regiones.
- **Foco en la Atención Neonatal:** La predominancia de diagnósticos y tratamientos relacionados con el recién nacido sugiere la importancia de continuar mejorando la atención perinatal y neonatal en el país.
- **Variabilidad en la Hospitalización:** Las diferencias significativas en la duración de la hospitalización indican que es necesario estandarizar los protocolos de atención y mejorar el manejo de casos complejos para reducir la variabilidad en los tiempos de hospitalización.

10. Recomendaciones

- **Mejorar la Cobertura y Calidad de Datos:** Se recomienda ampliar la cobertura de la recolección de datos en áreas rurales y garantizar la calidad y consistencia de los datos recolectados en todos los establecimientos de salud.
- **Descentralización de Servicios:** Fomentar la descentralización de los servicios de salud para mejorar el acceso en regiones fuera de Lima, asegurando que todas las regiones tengan acceso a un nivel adecuado de atención médica.
- **Fortalecimiento de la Atención Neonatal:** Continuar fortaleciendo los programas de atención perinatal y neonatal, con énfasis en la prevención y tratamiento de condiciones críticas como la sepsis neonatal.
- **Estandarización de Protocolos:** Desarrollar y estandarizar protocolos de hospitalización para reducir la variabilidad en la duración de las estancias hospitalarias, optimizando los recursos y mejorando los resultados clínicos.
- **Investigación Continua:** Realizar estudios continuos para monitorear las tendencias de atención de salud, con un enfoque en la evaluación de intervenciones implementadas y la identificación de nuevas áreas de mejora.