

Práctica de visualización de información

Curso académico 2024-25

Grupo # VIS_6

Instrucciones de ejecución de la aplicación interactiva

Es necesario guardar los archivos “*eventos.xlsx*” y “*pacientes.xlsx*” dentro de una carpeta llamada “data”. Previamente a la ejecución de la app de R, hace falta ejecutar todas las celdas del archivo “*limpia_excel.ipynb*” contenido dentro de la carpeta “Final”, con ello se preprocesan los datos como sale explicado en la sección [Conclusiones](#).

Para lanzar la aplicación, ejecutar el archivo “*app.R*” contenido dentro de la carpeta “Final” tras instalar todas las librerías necesarias.

Problema (media página)

¿Quién son los actores involucrados, que quieren obtener con la herramienta de visualización de datos? (media página)

Los principales actores involucrados en el uso de esta herramienta de visualización de datos son profesionales sanitarios, investigadores clínicos y personal técnico o analítico de instituciones de salud.

Esta aplicación interactiva, desarrollada en R, permite analizar datos agregados a nivel de cohorte, aunque también ofrece la posibilidad de consultar tablas detalladas de pacientes dentro de un grupo seleccionado. Esto permite tanto una visión general como un análisis más específico, dependiendo del objetivo del usuario.

El propósito común entre estos actores es identificar asociaciones entre enfermedades cardiovasculares, tratamientos antitrombóticos, factores de riesgo, hábitos de vida y eventos clínicos (como trombosis o sangrados). Al hacerlo, se pretende detectar posibles interacciones clínicas no deseadas, por ejemplo, evitar la combinación de ciertos medicamentos en pacientes con patologías específicas.

Los investigadores pueden utilizar esta herramienta para generar hipótesis o validar observaciones preliminares, mientras que los clínicos podrían apoyarse en ella para adaptar recomendaciones terapéuticas según patrones observados en la cohorte.

Abstracción de datos

Analizar el conjunto de datos

El tipo del conjunto de dato es una tabla, ya que son datos organizado en filas y columnas donde cada fila representa un ítem, cada columna representa un atributo y cada celda contiene el par (fila, columna) con un valor.

En las tablas inferiores vamos a profundizar en los tipos de datos de cada atributo.

Pacientes.xls

| Hoja de Excel | Campos | Tipo de atributo y cardinalidad |
|-------------------------|--|---------------------------------|
| DATOS DEL PACIENTE | Edad | Cuantitativo [49-97] |
| DATOS DEL PACIENTE | Paciente | Categorico (69) |
| DATOS DEL PACIENTE | Sexo Paciente | Categorico (2) |
| HISTORIA CARDIOVASCULAR | Paciente | categorico (69) |
| HISTORIA CARDIOVASCULAR | Hipertensión arterial | categorico (2) |
| HISTORIA CARDIOVASCULAR | Max valor P. Sistólica | cuantitativo [130 - 200] |
| HISTORIA CARDIOVASCULAR | Hipertensión controlada | categorico (3) |
| HISTORIA CARDIOVASCULAR | Medicamento Antihipertensivo | categorico (2) |
| HISTORIA CARDIOVASCULAR | Valvulopatía | categorico (2) |
| HISTORIA CARDIOVASCULAR | Tipo de valvulopatía | categorico (4) |
| HISTORIA CARDIOVASCULAR | Enfermedad coronaria | categorico (2) |
| HISTORIA CARDIOVASCULAR | Tipo | categorico (3) |
| HISTORIA CARDIOVASCULAR | Cuantos? | cuantitativo [0 - 3] |
| HISTORIA CARDIOVASCULAR | Tipo de intervención | categorico (4) |
| HISTORIA CARDIOVASCULAR | Enfermedad aórtica | categorico (2) |
| HISTORIA CARDIOVASCULAR | Cirugía | categorico (3) |
| HISTORIA CARDIOVASCULAR | Enfermedad carótidea | categorico (2) |
| HISTORIA CARDIOVASCULAR | Cirugía | categorico (3) |
| HISTORIA CARDIOVASCULAR | Enfermedad vascular periférica | categorico (2) |
| HISTORIA CARDIOVASCULAR | Arritmias | categorico (2) |
| HISTORIA CARDIOVASCULAR | Arritmia auricular | categorico (3) |
| HISTORIA CARDIOVASCULAR | Taquicardia ventricular | categorico (3) |
| HISTORIA CARDIOVASCULAR | Taquicardia paroxística supraventricular | categorico (2) |

| | | |
|----------------------------|---|---|
| HISTORIA CARDIOVASCULAR | Muerte súbita cardíaca | categorico (2) |
| HISTORIA CARDIOVASCULAR | Ablación | categorico (3) |
| HISTORIA CARDIOVASCULAR | Insuficiencia cardíaca | categorico (2) |
| HISTORIA CARDIOVASCULAR | Etiología | categorico (5) |
| HISTORIA CARDIOVASCULAR | NYHA | Ordinal secuencial (clase) |
| HISTORIA CARDIOVASCULAR | Fracción de Eyección del Ventrículo Izquierdo | cuantitativo [30 - 67] |
| HISTORIA CARDIOVASCULAR | Hipertensión pulmonar | categorico (3) |
| HISTORIA CARDIOVASCULAR | Presencia de crepitantes | categorico (3) |
| HISTORIA CARDIOVASCULAR | Ingurgitación de vena yugular | categorico (3) |
| HISTORIA CARDIOVASCULAR | Transplante cardíaco | categorico (2) |
| HISTORIA CARDIOVASCULAR | Fecha del transplante | ordinal secuencia (fecha) |
| HISTORIA CARDIOVASCULAR | Resincronización | categorico (2) |
| HISTORIA CARDIOVASCULAR | Fecha de la resincronización | ordinal secuencia (fecha) |
| HISTORIA CARDIOVASCULAR | Miocardiopatía | categorico (2) |
| HISTORIA CARDIOVASCULAR | Tipo de cardiomiopatía | categorico (3) |
| ANTITROMBOTICOS(ALTA) | Paciente | categorico (70) |
| ANTITROMBOTICOS(ALTA) | Se ha aplicado alguna guía clínica para la definición del tratamiento antitrombótico? | categorico (3) |
| ANTITROMBOTICOS(ALTA) | Se le ha prescrito anticoagulante? | categorico (3) |
| ANTITROMBOTICOS(ALTA) | Plan terapéutico del anticoagulante | Ordinal secuencial (tiempo en medicación) |
| ANTITROMBOTICOS(ALTA) | Número de meses | cuantitativo [1 - 6] |
| ANTITROMBOTICOS(ALTA) | Tipo de anticoagulante | categorico (4) |
| ANTITROMBOTICOS(ALTA) | Otro tipo | categorico (2) |
| ANTITROMBOTICOS(ALTA) | Tipo de ACOD | categorico (5) |
| ANTITROMBOTICOS(ALTA) | Otro tipo de ACOD | categorico (2) |
| ANTITROMBOTICOS(ALTA) | Tipo de AVK | categorico (2) |
| ANTITROMBOTICOS(ALTA) | Otro tipo de AVK | categorico (1) |
| ANTITROMBOTICOS(ALTA) | Motivo | categorico (6) |

| | | |
|-----------------------|---|--|
| ANTITROMBOTICOS(ALTA) | Otro motivo | categorico (4) |
| ANTITROMBOTICOS(ALTA) | Se le ha prescrito algún antiagregante? | categorico (2) |
| ANTITROMBOTICOS(ALTA) | Número de antiagregantes | cuantitativo [2 - 2] |
| ANTITROMBOTICOS(ALTA) | Plan terapéutico antiagregante 1 | Ordinal secuencial (tiempo en medicación) |
| ANTITROMBOTICOS(ALTA) | Número de meses | cuantitativo [0 - 0] |
| ANTITROMBOTICOS(ALTA) | Antiagregante 1 | categorico (3) |
| ANTITROMBOTICOS(ALTA) | Otro antiagregante 1 | categorico (1) |
| ANTITROMBOTICOS(ALTA) | Plan terapéutico antiagregante 2 | Ordinal secuencial (tiempo en medicación) |
| ANTITROMBOTICOS(ALTA) | Número de meses | cuantitativo [0 - 0] |
| ANTITROMBOTICOS(ALTA) | Antiagregante 2 | categorico (3) |
| ANTITROMBOTICOS(ALTA) | Otro antiagregante 2 | categorico (1) |
| OTROS DIAGNOSTICOS | Paciente | Categorico (67) |
| OTROS DIAGNOSTICOS | Transtorno de la coagulación | Categorico (2) |
| OTROS DIAGNOSTICOS | Discrasia sanguínea | Categorico (2) |
| OTROS DIAGNOSTICOS | Diabetes | Categorico (2) |
| OTROS DIAGNOSTICOS | Dislipdemia | Categorico (2) |
| OTROS DIAGNOSTICOS | Enfermedad hepática | Categorico (2) |
| OTROS DIAGNOSTICOS | Cáncer | Categorico (2) |
| OTROS DIAGNOSTICOS | Anemia | Categorico (2) |
| OTROS DIAGNOSTICOS | Enfermedad renal | Ordinal secuencial (gravedad de la enfermedad renal) |
| FACTORES DE RIESGOS | Código | Categorico (69) |
| FACTORES DE RIESGOS | Antecedentes familiares de cardiopatía isquémica precoz (menor de 50 en varones y menor de 55 en mujeres) | Categorico (4) |
| FACTORES DE RIESGOS | Riesgo de caída | Categorico (3) |
| FACTORES DE RIESGOS | Caídas previas conocidas | Categorico (4) |
| FACTORES DE RIESGOS | Toma algún medicamento presente en esta lista | Categorico (15) |
| FACTORES DE RIESGOS | Déficit sensorial | Categorico (7) |
| FACTORES DE RIESGOS | Estado mental | Categorico (4) |
| FACTORES DE RIESGOS | Deambulación alterada | Ordinal secuencia (capacidad de andar) |
| FACTORES DE RIESGOS | Tabaquismo | Categorico (5) |
| FACTORES DE RIESGOS | Consumo diario de alcohol | Categorico (4) |
| FACTORES DE RIESGOS | Tipo de consumo | Categorico (4) |
| HABITOS | Código | Categorico (68) |
| HABITOS | Actualizado en este seguimiento | Categorico (3) |
| HABITOS | Realiza control regular de la presión? | Categorico (4) |

| | | |
|---------|---|----------------------|
| HABITOS | Minutos semanales de actividad aeróbica | Cuantitativo [0-240] |
| HABITOS | Dieta Mediterránea | Categorico (4) |

Eventos.xls

| Hoja de Excel | Campos (poner el nombre) | Tipo de atributo y cardinalidad |
|---------------|---|--|
| SANGRADO | Paciente | Categorico (116) |
| SANGRADO | Número anticoagulantes | Cuantitativo [0 - 1] |
| SANGRADO | Número antiagregantes | Cuantitativo [0 - 2] |
| SANGRADO | Otro medicamentos | Categorico (49) |
| SANGRADO | ANTICOAGULANT_STRING | Categorico (4) |
| SANGRADO | ANTICOAGULANT_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | ANTIPLATELET_STRING | Categorico (6) |
| SANGRADO | ANTIPLATELET_COUNT | Cuantitativo [0 - 2] |
| SANGRADO | ANALGESIC_STRING | Categorico (2) |
| SANGRADO | ANALGESIC_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | NON_STEROIDAL_INFLAMMATORY_STRING | Categorico (2) |
| SANGRADO | NON_STEROIDAL_INFLAMMATORY_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | DIURETIC_STRING | Categorico (2) |
| SANGRADO | DIURETIC_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | OTHER_STRING | Categorico (49) |
| SANGRADO | OTHER_COUNT | Cuantitativo [0 - 17] |
| SANGRADO | Caracterización de la hemorragia | Categorico (3) |
| SANGRADO | Procedimientos terapéuticos | Categorico (3) |
| SANGRADO | Descenso de hemoglobina | Cuantitativo [0-7] |
| SANGRADO | Gravedad de la hemorragia (TIMI) | Ordinal secuencial (gravedad hemorragia) |
| SANGRADO | Gravedad de la hemorragia (GUSTO) | Ordinal secuencial (gravedad hemorragia) |
| SANGRADO | Gravedad de la hemorragia (BARC) | Ordinal secuencial (gravedad hemorragia) |
| SANGRADO | Tipo de sangrado | Categorico (6) |
| SANGRADO | ¿El paciente ha subido una transfusión? | Categorico (9) |
| TROMBOTICO | Paciente | Categorico (104) |
| TROMBOTICO | Número anticoagulantes | Cuantitativo [0 - 1] |
| TROMBOTICO | Número antiagregantes | Cuantitativo [0 - 2] |
| TROMBOTICO | Otro medicamentos | Categorico (21) |
| TROMBOTICO | ANTICOAGULANT_STRING | Categorico (3) |
| TROMBOTICO | ANTICOAGULANT_COUNT | Cuantitativo [0 - 1] |
| TROMBOTICO | ANTIPLATELET_STRING | Categorico (6) |
| TROMBOTICO | ANTIPLATELET_COUNT | Cuantitativo [0 - 2] |
| TROMBOTICO | ANALGESIC_STRING | Categorico (2) |

| | | |
|------------|---|-----------------------|
| TROMBOTICO | ANALGESIC_COUNT | Cuantitativo [0 - 1] |
| TROMBOTICO | NON_STEROIDAL_INFLAMMATORY_STRING | Categórico (1) |
| TROMBOTICO | NON_STEROIDAL_INFLAMMATORY_COUNT | Cuantitativo [0 - 0] |
| TROMBOTICO | DIURETIC_STRING | Categórico (2) |
| TROMBOTICO | DIURETIC_COUNT | Cuantitativo [0 - 1] |
| TROMBOTICO | OTHER_STRING | Categórico (21) |
| TROMBOTICO | OTHER_COUNT | Cuantitativo [0 - 10] |
| TROMBOTICO | TYPE_THROMBOTIC_PRE | Categórico (2) |
| TROMBOTICO | ¿El paciente ha sufrido un evento trombótico previo a la inclusión? | Categórico (1) |
| TROMBOTICO | Tipo de evento trombótico | Categórico (5) |
| TROMBOTICO | Tipo de intervención | Categórico (4) |
| TROMBOTICO | option3_THROMBOTIC_PRE | Categórico (2) |

Abstracción de tareas

1) ¿Qué tipo de tareas se ha de realizar el usuario?

El usuario debe llevar a cabo principalmente tareas analíticas de tipo *explorar, filtrar, comparar, relacionar e identificar*. La herramienta está diseñada para permitir a los usuarios observar patrones en la cohorte, detectar posibles asociaciones entre variables clínicas y realizar comparaciones entre subgrupos (por ejemplo, por edad, sexo o enfermedades específicas). También se incluyen tareas de *consultar detalles bajo demanda*, como al hacer clic en una barra para acceder a los datos individuales de los pacientes en ese grupo.

2) Identificar los tipos de atributos derivados necesarios para realizar las tareas

Para facilitar las tareas descritas, es necesario derivar ciertos atributos a partir de los datos originales. Algunos ejemplos incluyen:

- Agrupaciones por rango de edad o niveles de actividad física.
- Cálculo de tasas de eventos (frecuencia relativa).
- Correlaciones entre variables clínicas (medidas estadísticas).
- Clasificación de eventos por tipo (sangrado vs trombótico y sus subtipos).
- Media, mediana y rango para variables como presión arterial.

| Usuario | Tarea (usar la terminología propia de la abstracción de tareas) | Justificación |
|---------------------|---|--|
| Profesional clínico | Explorar / Filtrar / Consultar detalles bajo demanda | Necesita examinar la distribución de eventos y acceder a casos concretos que cumplan ciertos criterios (edad, sexo, etc.). |

| | | |
|-----------------------|--|---|
| Investigador | Relacionar / Comparar / Correlacionar | Busca identificar asociaciones entre tratamientos, enfermedades y eventos adversos. |
| Analista de datos | Obtener resumen / Describir distribución | Requiere calcular y visualizar atributos derivados (promedios, frecuencias, correlaciones) para análisis más amplios. |
| Tomador de decisiones | Comparar / Detectar patrones | Le interesa observar diferencias entre subgrupos para definir recomendaciones clínicas o políticas. |

Codificación visual

Describir las opciones de diseño para que el usuario pueda realizar las tareas especificadas anteriormente

| Vista | Opción de diseño (terminología) | Detalles de la tarea (que puede realizar el usuario) |
|----------------------------------|---|--|
| Eventos (gráfico de barras) | Altura / longitud (posición), color, interacción (clic) | Comparar frecuencia de eventos, distinguir tipos de evento (color) y acceder al detalle de los pacientes (clic). |
| Eventos (comparativa) | Altura / longitud, doble panel, filtros interactivos | Filtrar grupos por sexo, edad u otros atributos y comparar visualmente la ocurrencia de eventos entre grupos. |
| Correlaciones (heatmap) | Color (intensidad), posición en matriz, etiquetas | Identificar relaciones entre enfermedades, antitrombóticos y eventos (correlaciones positivas o negativas). |
| Descripción (factores de riesgo) | Altura (barras), color por categoría, filtros múltiples | Explorar cómo uno o varios factores afectan a una enfermedad cardiovascular específica. |
| Descripción (presión arterial) | Altura (boxplot), posición, outliers con forma distinta | Comparar la presión sistólica según factores de riesgo y detectar valores atípicos. |
| Descripción (actividad física) | Tamaño (burbuja), posición (X e Y), color por categoría | Relacionar cantidad de ejercicio físico con enfermedades, comparar entre grupos según la distribución espacial. |

| | | |
|--------------------|---|--|
| Tabla de pacientes | Texto, ordenamiento de columnas, interacción (clic) | Consultar detalles bajo demanda, ordenar por atributos específicos para buscar patrones o agrupaciones relevantes. |
|--------------------|---|--|

Implementación

Detalles de la implementación y funcionalidades usadas. Describir la organización del código, librerías usadas y lógica desarrollada.

Previamente a la visualización de la información se han preprocesado los datos, juntando columnas que contenían información similar, como aquellas que indicaban el número de medicamentos tomados y cuáles han sido, el tratamiento que recibe cada paciente y la duración de este. También se eliminaron las columnas que contenían siempre el mismo valor y no aportaban información relevante, y se mapearon a 0 los valores nulos en determinadas columnas. Todo esto se realizó desde Python usando la librería pandas. Tras este proceso, los datos se exportaron a ficheros CSV.

Desde el módulo *data_processing* se cargaron los datos y se unió la información procedente del Excel de pacientes con los datos de eventos. Además, se mapearon las columnas tabaquismo, consumo de alcohol y minutos de actividad física para poder operar con estas de forma numérica.

El archivo **app.R** carga las librerías usadas, que son las siguientes: shiny, para construir la aplicación web interactiva; shinythemes, que permite aplicar temas de diseño a la interfaz de la app; plotly, para generar gráficos interactivos; ggplot2, para visualizaciones personalizadas; dplyr, que facilita la manipulación y transformación de datos; tidyr, empleada para ordenar y dar formato a los datos; DT, que permite mostrar tablas interactivas dentro de la aplicación; y readxl, utilizada para importar datos desde archivos Excel. A continuación, se cargan los módulos principales, se ejecuta el preproceso de los datos y se lanza la aplicación.

El módulo **server_module.R** define el servidor de la aplicación, encargándose de la navegación entre las distintas pestañas. Se llaman las funciones responsables de generar cada visualización, pasándoles los datos y el contexto de la sesión. Además, se configuran los observadores que permiten cambiar de pestaña desde los botones de la pantalla de inicio.

El módulo **ui_module.R** define la estructura de la interfaz de usuario de la aplicación mediante y la barra de navegación entre las diferentes pestañas cada una asociada a una sección de la aplicación.

El módulo **Descripción de pacientes** busca entender mejor la interacción entre enfermedades cardiovasculares y factores de riesgo. Consta de tres visualizaciones:

Un gráfico que muestra cómo afectan uno o varios factores de riesgo a una enfermedad cardiovascular específica, lo cual puede ayudar a identificar combinaciones particularmente perjudiciales.

Un boxplot que representa la presión sistólica máxima en función de los factores de riesgo seleccionados, permitiendo visualizar la distribución y detectar valores atípicos.

Un gráfico de burbujas que relaciona la actividad física semanal con una enfermedad concreta, con el objetivo de identificar qué patologías están asociadas a un mayor o menor nivel de ejercicio recomendado.

Al hacer clic en cualquiera de los grupos de cualquiera de los tres gráficos, también se despliega una tabla interactiva con todos los pacientes que cumplen esos criterios.

Sobre el apartado **Grafico de eventos**, en la interfaz (*grafico_eventos_ui.R*), se organiza un panel lateral que incluye controles interactivos para seleccionar el tipo de evento, género, rango de edad y según el tipo de evento, filtros adicionales para la gravedad de sangrado o tipo de evento trombótico.

En el servidor (*grafico_eventos.R*) se gestiona la visualización del gráfico interactivo. En él se representan los datos filtrados según las selecciones del panel lateral, diferenciando los distintos datos mediante colores. Además, el gráfico permite la interacción del usuario, que puede seleccionar un rango de edades con el ratón y consultar los eventos correspondientes a ese tramo en una tabla que se muestra en una ventana modal. La visualización se genera con *ggplot2* y consiste en un histograma que muestra la distribución de eventos por edad. El color de las barras varía según el tipo de evento o su gravedad, en función de las selecciones del usuario. El gráfico permite aplicar un zoom sobre los rangos seleccionados.

El módulo de **Comparación de eventos** funciona de forma similar al gráfico de eventos, pero en este caso se muestran dos histogramas independientes, sin interacción directa sobre ellos. Cada gráfico cuenta con su propio panel lateral de filtros para seleccionar el tipo de evento, sexo y características específicas según el evento excepto la edad, que ambos comparten un deslizador de edad común, lo que permite comparar de manera visual distintos grupos de eventos sobre una misma franja de edad, facilitando el análisis paralelo de subgrupos de pacientes.

En el código *comparacion_eventos_ui.R*, se define la estructura con los filtros para cada gráfico y el deslizador de edad. En *comparacion_eventos.R* se filtran los datos según los valores seleccionados en los paneles laterales.

El **módulo de correlaciones** está diseñado para explorar relaciones entre variables categóricas dentro de los eventos registrados. Está dividido en dos pestañas: una para eventos de sangrado y otra para eventos trombóticos. En cada pestaña, el usuario puede elegir dos variables desde un panel lateral y se genera construye una tabla de frecuencias cruzadas con la que construye un mapa de calor que muestra la frecuencia de cada combinación de valores. De este modo, se facilita la identificación de posibles relaciones que puedan asociarse a la aparición de eventos.

El código de la interfaz (*correlaciones_ui.R*) utiliza un *tabsetPanel* para organizar las dos pestañas. Dentro de cada una, aparece un *sidebarPanel* que contiene dos *selectInput* para que el usuario escoja las variables a correlacionar (propias del evento).

El servidor escucha los inputs seleccionados y filtra las columnas necesarias del dataset y usando `table()` genera una tabla que se transforma en `data.frame` para poder graficarla. Luego se construye un heatmap con `ggplot2` donde el color de cada celda representa la cantidad de eventos, y se añade el valor numérico encima de cada celda. El gradiente de color va de blanco (poca frecuencia) a azul intenso (alta frecuencia).

Conclusiones

- Conclusiones sobre la cohorte de pacientes tras el análisis visual implementada
- Identificación de futuras mejoras de la visualización analítica

En las gráficas de comparación por género, se observa una clara diferenciación en la distribución por edad. En el caso de las pacientes mujeres, los eventos se concentran mayoritariamente en edades avanzadas, destacando especialmente a partir de los 78 años. Por el contrario, en los pacientes varones, aunque también predominan los eventos en edades elevadas (en torno a los 85 años), la distribución es algo más homogénea y repartida en diferentes tramos de edad.

Filtros Histograma A

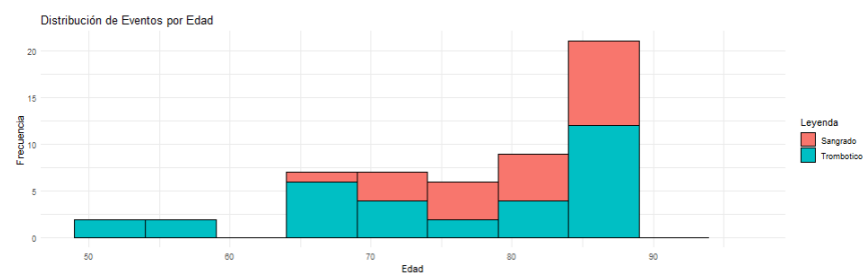
Tipo de Evento:

Todos

Sexo:

MALE

Histograma A



Filtros Histograma B

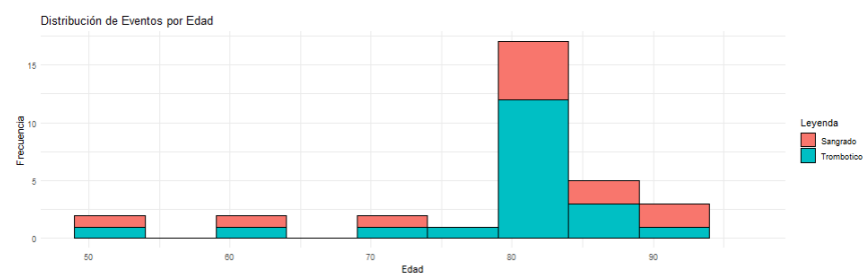
Tipo de Evento:

Todos

Sexo:

FEMALE

Histograma B



Analizando en detalle los eventos de sangrado según género, se identifican patrones diferenciados en hombres y mujeres. En los pacientes varones, se aprecia una tendencia ascendente en el número de eventos de sangrado conforme avanza la edad. Este incremento alcanza su punto máximo alrededor de los 85 años, franja en la que además se concentran los casos de mayor gravedad. Esto permite concluir que, en hombres, tanto la frecuencia como la severidad de los sangrados aumentan progresivamente con la edad.

Por su parte, en las pacientes mujeres, los eventos de sangrado comienzan a registrarse a edades más tempranas que en los hombres. No obstante, su evolución no sigue una progresión lineal tan definida, y se detecta que a partir de los 80 años se produce un incremento notable en la cantidad de casos, situando también este tramo de edad como un punto de especial riesgo.



Al analizar los eventos trombóticos en función del género, se observan varios patrones relevantes. El evento TRANSIENT_ISCHEMIC_ATTACK ocurre principalmente en hombres de aproximadamente 85 años de edad, aunque destaca un caso aislado en una mujer joven que ha sufrido este tipo de evento. Los eventos más frecuentes en la cohorte son MYOCARDIAL_INFARCT y STABLE_ANGINA. Ambos se presentan mayoritariamente en hombres, distribuidos a lo largo de prácticamente todo el rango de edades analizado. Sin embargo, en mujeres, estos eventos se concentran casi exclusivamente en edades avanzadas, repitiendo el patrón ya observado en otros tipos de eventos. En cuanto a STROKE y SYSTEMIC_EMBOLISM, la cantidad de casos registrados en la muestra es demasiado reducida como para poder extraer conclusiones fiables. Sería necesario disponer de una muestra más amplia o más casos para evaluar correctamente su comportamiento en función de la edad y el género.

Filtros Histograma A

Tipo de Evento:

Trombotico

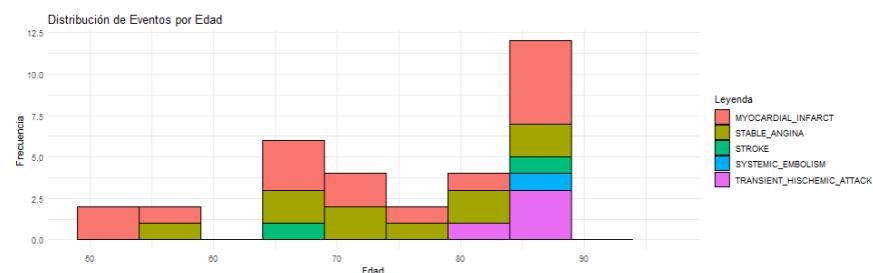
Sexo:

MALE

Tipo Trombotico:

- ☐ STABLE_ANGINA
- ☐ TRANSIENT_ISCHEMIC_ATTACK
- ☐ MYOCARDIAL_INFARCT
- ☐ STROKE
- ☐ SYSTEMIC_EMBOLISM

Histograma A



Filtros Histograma B

Tipo de Evento:

Trombotico

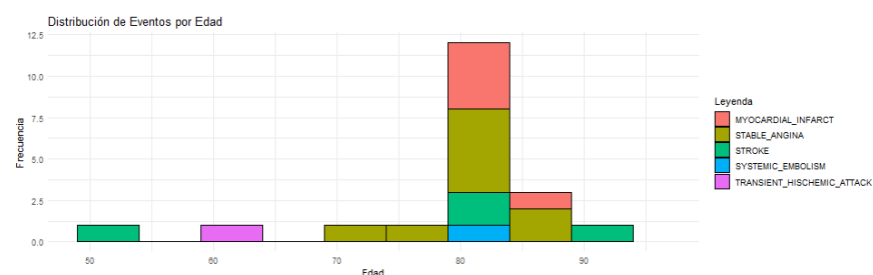
Sexo:

FEMALE

Tipo Trombotico:

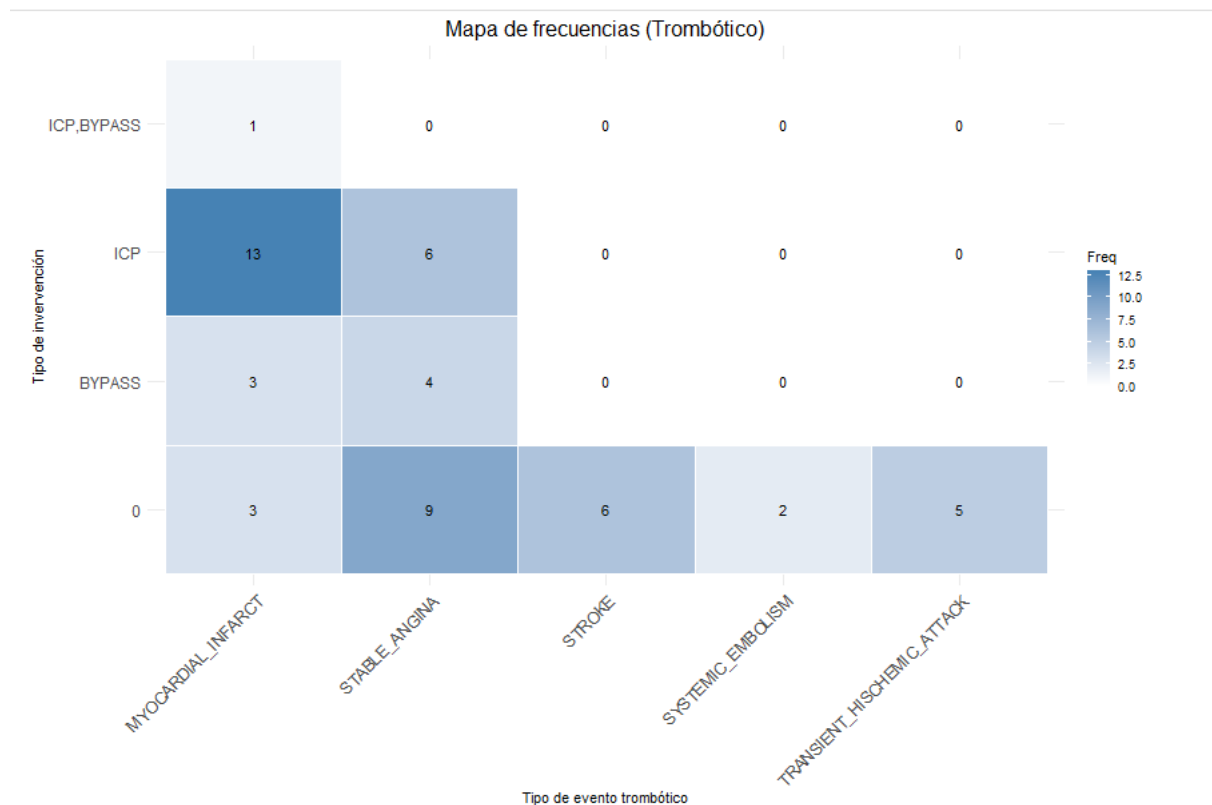
- ☐ STABLE_ANGINA
- ☐ TRANSIENT_ISCHEMIC_ATTACK
- ☐ MYOCARDIAL_INFARCT
- ☐ STROKE
- ☐ SYSTEMIC_EMBOLISM

Histograma B

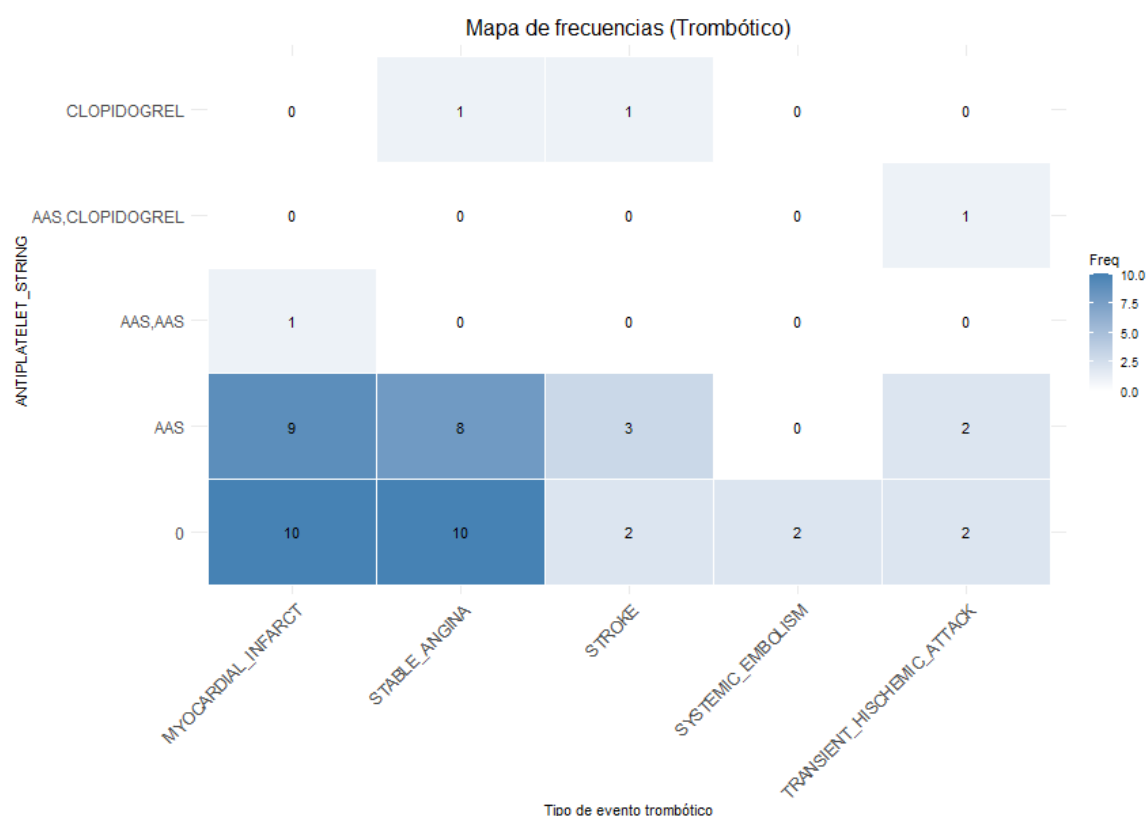


Eventos trombóticos

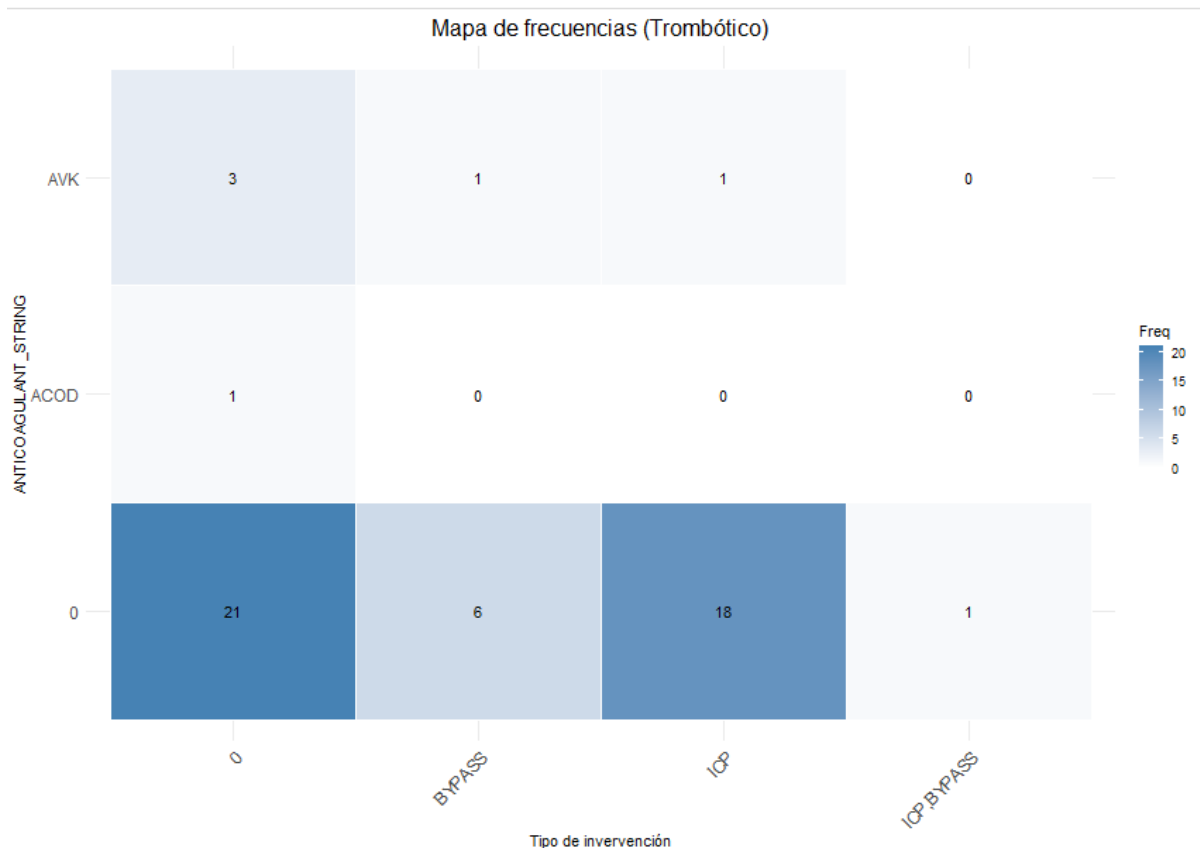
En la visualización de correlaciones para eventos trombóticos, al analizar la relación entre el tipo de evento trombótico y el tipo de intervención realizada, se identifican varios patrones. Los eventos STROKE, TRANSIENT_ISCHEMIC_ATTACK y SYSTEMIC_EMBOLISM no han recibido ningún tipo de intervención terapéutica en la cohorte analizada. Por otro lado, dentro de los eventos de MYOCARDIAL_INFARCT, predominan las intervenciones mediante ICP (Intervención Coronaria Percutánea), destacándose como el abordaje más habitual. En el caso de STABLE_ANGINA, se observa que la mayoría de los pacientes tampoco ha recibido ninguna intervención, aunque en un número relevante de casos también se ha optado por ICP, estableciéndose como la alternativa más frecuente en aquellos que sí recibieron tratamiento invasivo.



Respecto a la relación entre los eventos trombóticos y el uso de antiplaquetarios, se observa que una parte considerable de los eventos registrados en la cohorte no ha recibido ningún tipo de anticoagulación. Sin embargo, entre los pacientes que sí recibieron tratamiento, el AAS (ácido acetilsalicílico) destaca como el fármaco más utilizado. Es relevante señalar que el clopidogrel se ha administrado de forma individual únicamente en los casos de STROKE y STABLE_ANGINA. En los eventos de TRANSIENT_ISCHEMIC_ATTACK, su uso se ha registrado exclusivamente en combinación con AAS, sin casos de administración aislada.

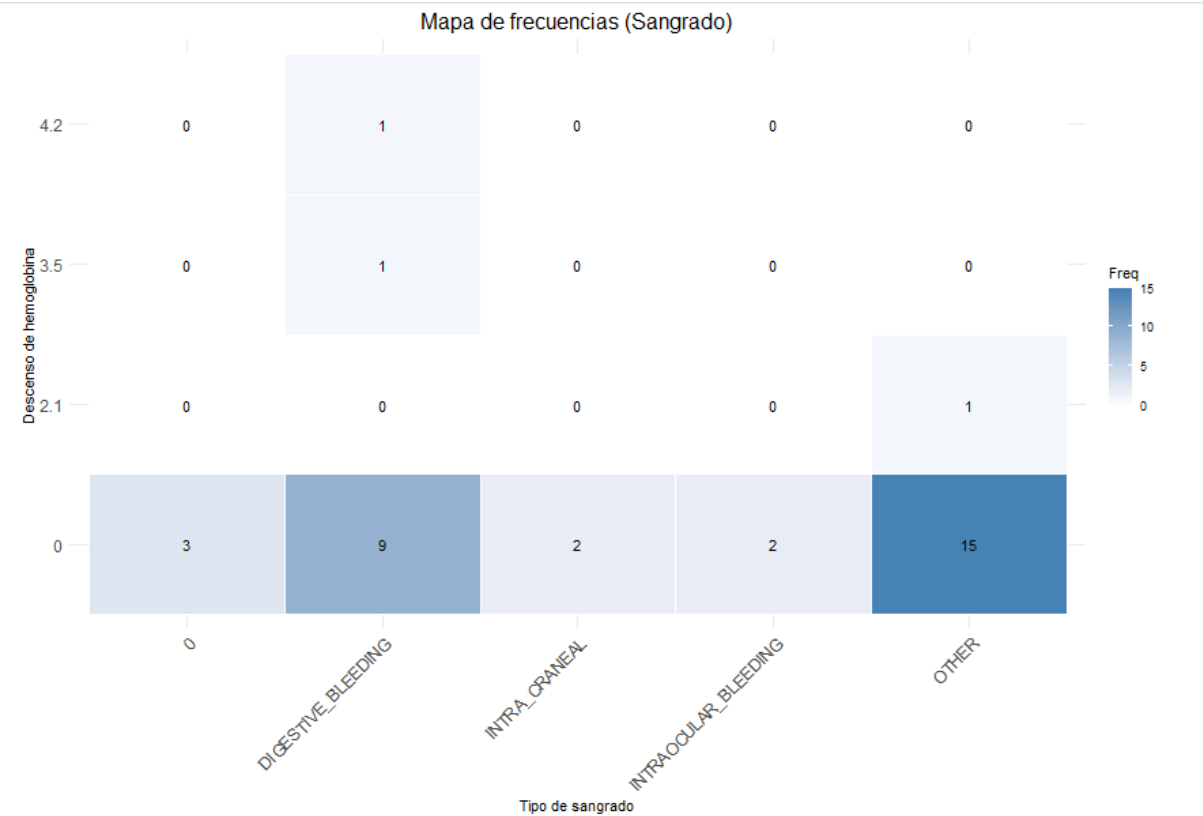


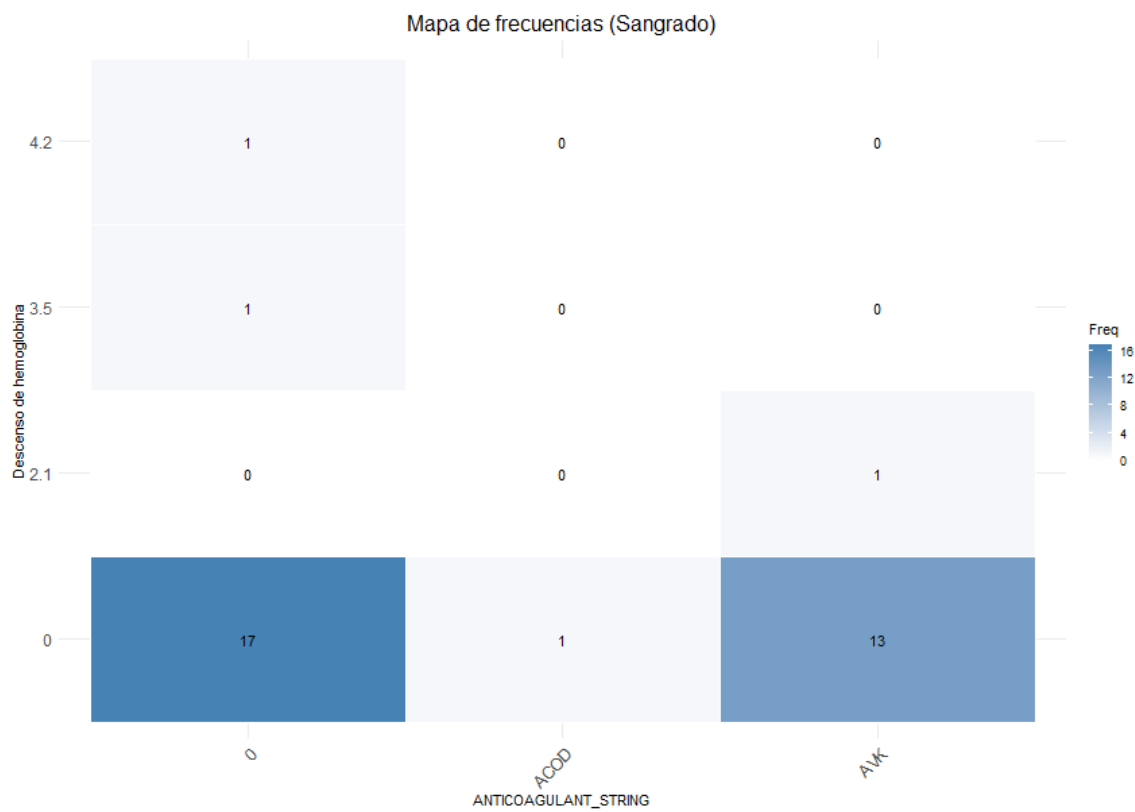
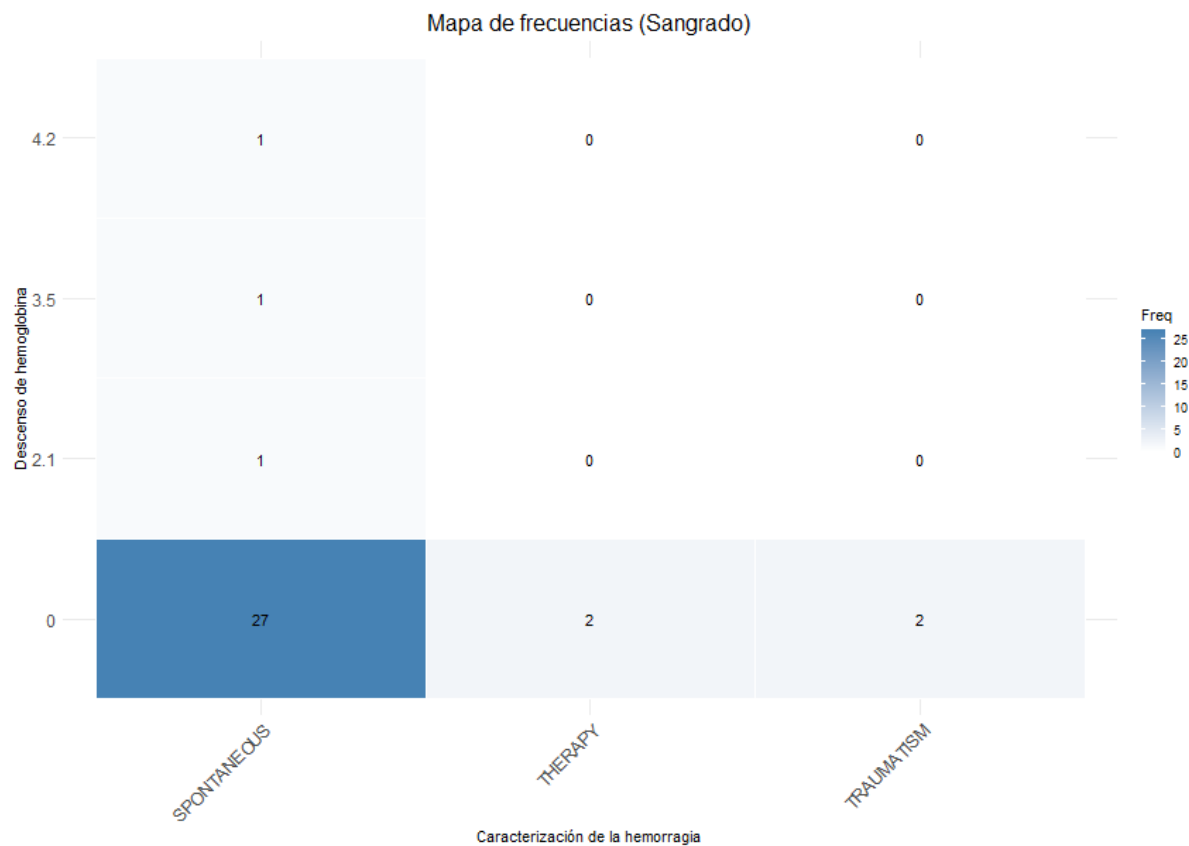
Al analizar la relación entre el uso de anticoagulantes y las intervenciones realizadas en eventos trombóticos, se aprecia que, salvo en un par de casos puntuales, no se ha administrado medicación anticoagulante tras las intervenciones. La mayoría de procedimientos se han llevado a cabo sin tratamiento anticoagulante asociado.

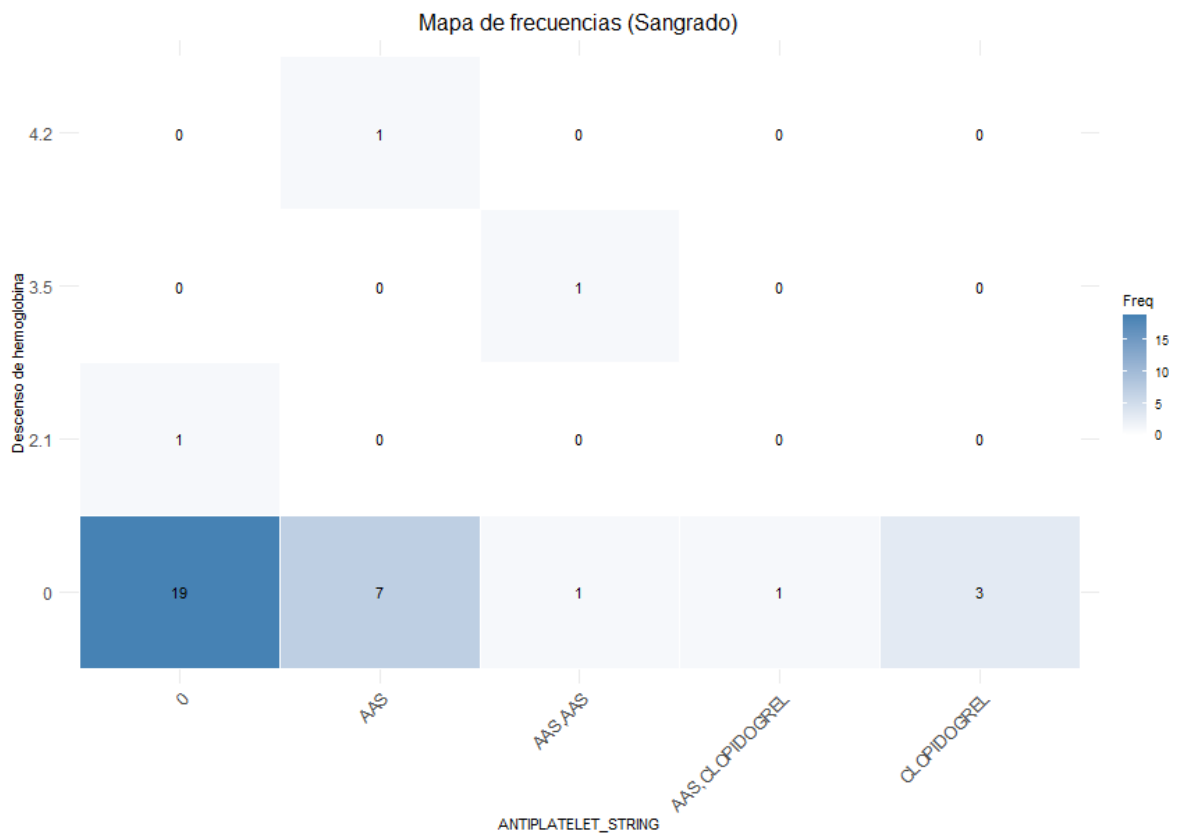


Eventos de sangrado:

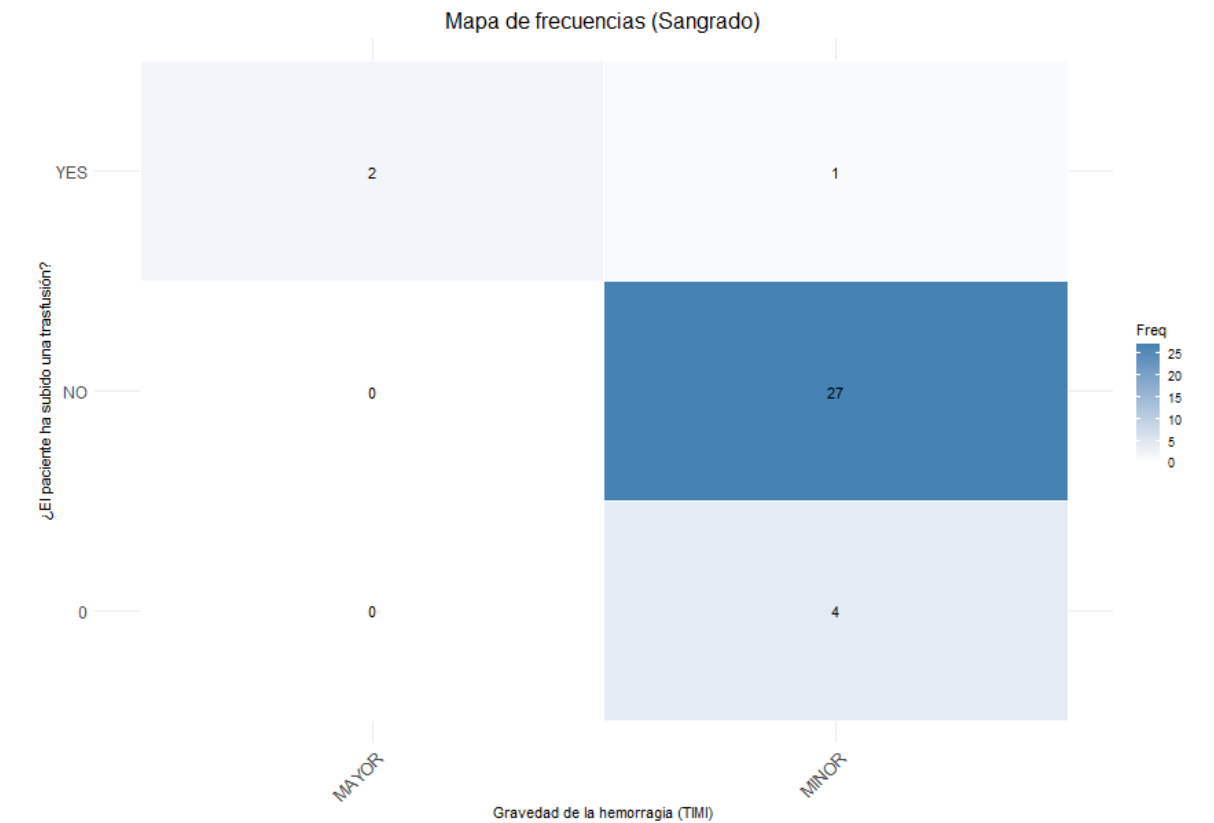
Analizando los distintos tipos de sangrado, se identifica que DIGESTIVE_BLEEDING destaca como el evento con los descensos más pronunciados de hemoglobina. En comparación, otros tipos de sangrado presentan descensos nulos salvo un ejemplo de OTHER, lo que sugiere que los tipos de sangrado son diferenciales entre los tipos de hemorragia. Al visualizar también la caracterización de la hemorragia, se reconoce que todos los descensos de hemoglobina registrados han sucedido en eventos espontáneos, sin relación con intervenciones médicas previas, lo que aporta relevancia al contexto de estos episodios. Además, se detecta que los sangrados digestivos graves ocurren mayoritariamente en ausencia de anticoagulantes, pero bajo tratamiento antiplaquetario, especialmente AAS. Este patrón sugiere que, el tipo de sangrado y su localización deben considerarse a la hora de ajustar tratamientos, pero la cantidad de valores en los datos no es suficiente como para asegurar con mucha seguridad las afirmaciones anteriores, nos haría falta recolectar más información.







Mediante la visualización entre la gravedad de la hemorragia y si el paciente ha subido una transfusión, se observa que en todos los casos clasificados como hemorragia mayor se ha administrado una transfusión. Por otro lado, los casos en los que el valor aparece como 0 corresponden a registros sin información, aunque es razonable inducir que se tratan de hemorragias sin necesidad de transfusión, debido a que todos son casos de gravedad menor.



El resto de atributos de los eventos de sangrado no parecen representar mucha información, ya que la mayoría de los eventos salvo 1 o 2 se concentran en el mismo valor,

Descripción de pacientes

En esta pantalla, dedicada a relacionar enfermedades cardiovasculares con los factores de riesgo, vemos que algunos factores de riesgo, como el de déficit sensorial tiene muy pocos individuos, por lo que, no se puede sacar conclusiones de él.

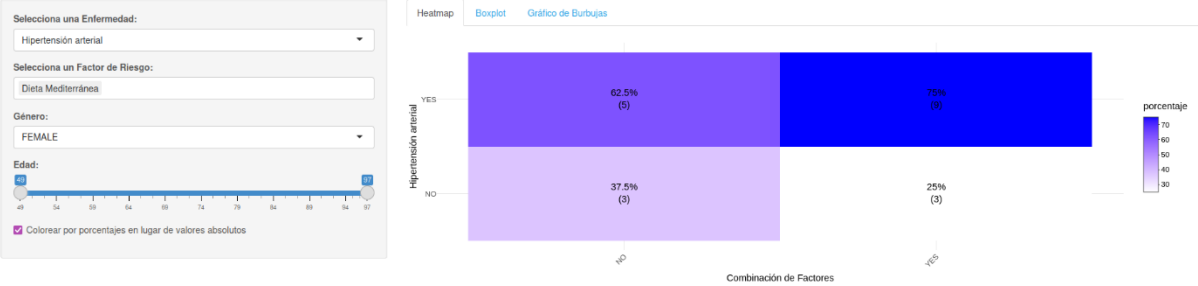
Vamos a empezar sacando información del consumo de drogas (tabaquismo y consumo de alcohol).

Vemos que en nuestro dataset casi nadie bebe, solo beben 3 hombres, en estilo mediterráneo y una mujer en estilo anglosajón.

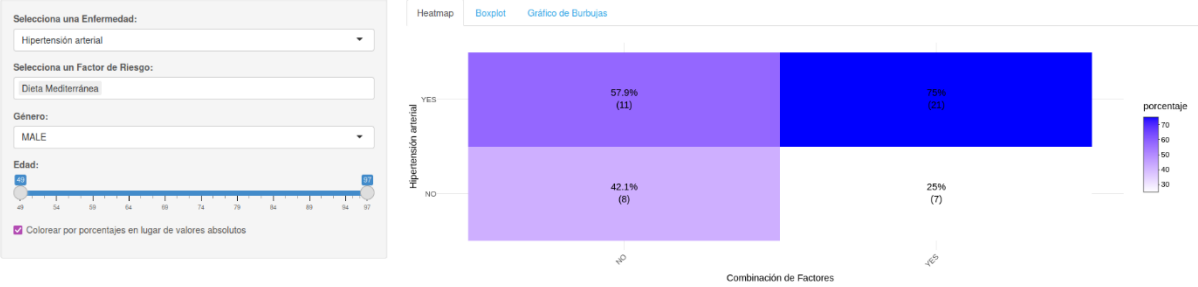
Por parte de las mujeres, tampoco fuma casi ninguna, ya que fuman solo 2 mujeres, aparte de otra que es exfumadora. Por lo tanto, no tiene sentido filtrar por sexo.

Analizamos que la distribución en la dieta entre ambos sexos también es muy similar:

Relación entre Hábitos, Factores de Riesgo y Enfermedades



Relación entre Hábitos, Factores de Riesgo y Enfermedades

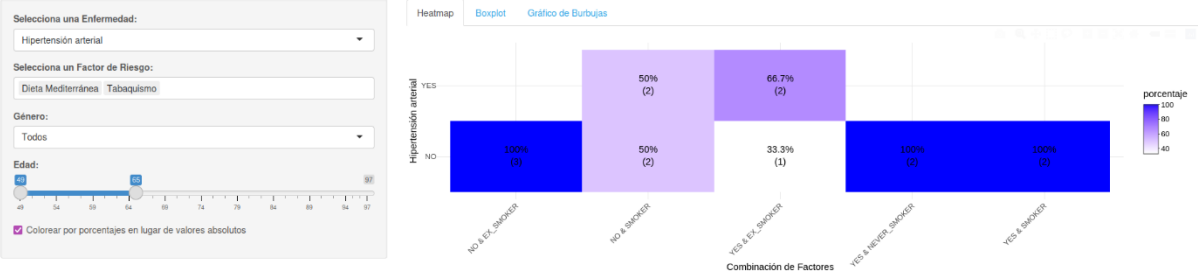


Vamos a analizar cómo afectan el tabaquismo y la dieta mediterránea a las enfermedades cardiovasculares, sin filtrar por sexo, ya que no hay diferencias.

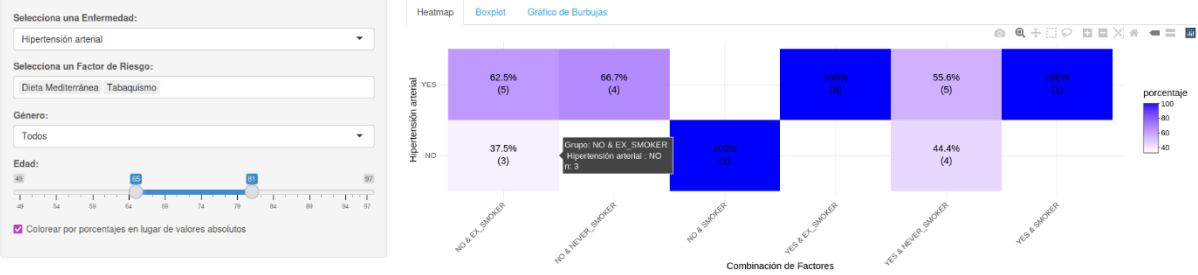
Vamos a dividir la población en 3 grupos de edad: **jóvenes** (de 49 a 65), **mediana edad** (de 65 a 81) y **mayores** (de 81 a 97).

Hipertensión arterial:

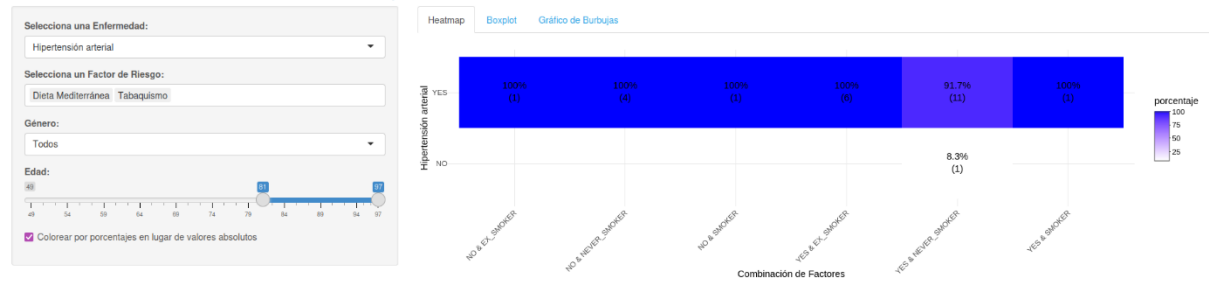
Relación entre Hábitos, Factores de Riesgo y Enfermedades



Relación entre Hábitos, Factores de Riesgo y Enfermedades



Relación entre Hábitos, Factores de Riesgo y Enfermedades



No hay datos concluyentes para decir que los factores de riesgo afecten a tener o no la enfermedad. En cambio, si se puede ver que conforme avanzamos el tramo de edad, es más común tener la enfermedad.

Aun así, no podemos concluir que conforme envejeces es más probable tener hipertensión arterial, porque los datos que tenemos son solo de gente que tiene alguna enfermedad cardiovascular. Entonces, lo único que se puede concluir es que **si tienes una enfermedad arterial, lo más probable es que tengas más de 81 años.**

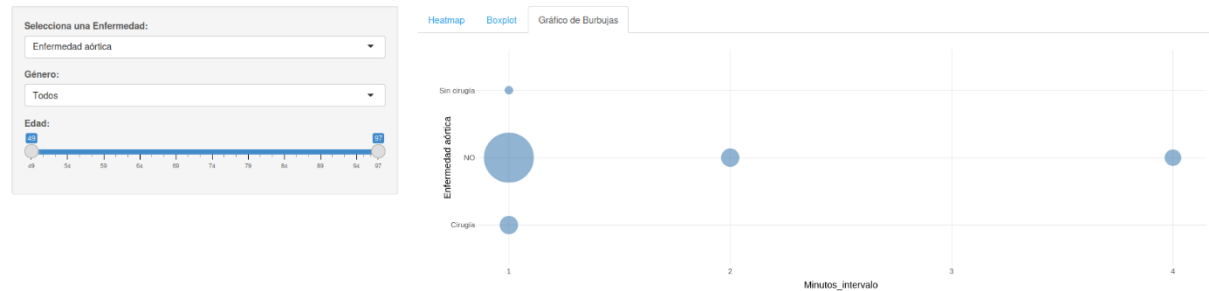
Relación entre Hábitos, Factores de Riesgo y Enfermedades



En esta gráfica podemos ver como varía la máxima presión Sistólica. Se ve que no hay gran diferencia entre las distintas clases.

Por último, analizando el bubble plot de los minutos de ejercicio asociado a una enfermedad cardiovascular, se ve para todas las enfermedades que la gran mayoría de los pacientes no realizan mucha actividad física, con casi todos encontrándose en el primer grupo.

Relación entre Hábitos, Factores de Riesgo y Enfermedades



Aspectos de mejora

Al realizar el análisis, en eventos sí que lo incluimos, pero en la descripción de pacientes, faltaría que, al hacer clic en un grupo, a parte de la tabla, saliese debajo una gráfica con la distribución de la edad y el género. Se puede analizar perfectamente sin este añadido, pero si lo podría hacer más cómodo.

Los datos incluyen muy pocas muestras, por lo que las conclusiones que se pueden obtener son poco fiables. Además, al tener información solo sobre pacientes que tienen alguna enfermedad y no tenemos información de la población sana, no se puede comparar y las asociaciones aparentes no se pueden contrastar, ya que podrían ser casualidad.

Por último, los datos son inconsistentes, hay eventos de pacientes que no están registrados, aparte de muchas columnas vacías, lo que hace que se pierdan muchos datos que aparentemente se tienen.