Práctica de visualización de información

Curso académico 2024-25

Grupo # \_\_\_\_\_\_\_\_\_\_\_\_\_\_

# 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 | Categórico (69) |
| DATOS DEL PACIENTE | Sexo Paciente | Categórico (2) |
| HISTORIA CARDIOVASCULAR | Paciente | categórico (69) |
| HISTORIA CARDIOVASCULAR | Hipertensión arterial | categórico (2) |
| HISTORIA CARDIOVASCULAR | Max valor P. Sistólica | cuantitativo [130 - 200] |
| HISTORIA CARDIOVASCULAR | Hipertensión controlada | categórico (3) |
| HISTORIA CARDIOVASCULAR | Medicamento Antihipertensivo | categórico (2) |
| HISTORIA CARDIOVASCULAR | Valvulopatía | categórico (2) |
| HISTORIA CARDIOVASCULAR | Tipo de valvulopatía | categórico (4) |
| HISTORIA CARDIOVASCULAR | Enfermedad coronaria | categórico (2) |
| HISTORIA CARDIOVASCULAR | Tipo | categórico (3) |
| HISTORIA CARDIOVASCULAR | Cuantos? | cuantitativo [0 - 3] |
| HISTORIA CARDIOVASCULAR | Tipo de intervención | categórico (4) |
| HISTORIA CARDIOVASCULAR | Enfermedad aórtica | categórico (2) |
| HISTORIA CARDIOVASCULAR | Cirugía | categórico (3) |
| HISTORIA CARDIOVASCULAR | Enfermedad carótidea | categórico (2) |
| HISTORIA CARDIOVASCULAR | Cirugía | categórico (3) |
| HISTORIA CARDIOVASCULAR | Enfermedad vascular periférica | categórico (2) |
| HISTORIA CARDIOVASCULAR | Arritmias | categórico (2) |
| HISTORIA CARDIOVASCULAR | Arritmia auricular | categórico (3) |
| HISTORIA CARDIOVASCULAR | Taquicardia ventricular | categórico (3) |
| HISTORIA CARDIOVASCULAR | Taquicardia paroxística supraventicular | categórico (2) |
| HISTORIA CARDIOVASCULAR | Muerte súbita cardíaca | categórico (2) |
| HISTORIA CARDIOVASCULAR | Ablación | categórico (3) |
| HISTORIA CARDIOVASCULAR | Insuficiencia cardíaca | categórico (2) |
| HISTORIA CARDIOVASCULAR | Etiología | categórico (5) |
| HISTORIA CARDIOVASCULAR | NYHA | Ordinal secuencial (clase) |
| HISTORIA CARDIOVASCULAR | Fracción de Eyección del Ventrículo Izquierdo | cuantitativo [30 - 67] |
| HISTORIA CARDIOVASCULAR | Hipertensión pulmonar | categórico (3) |
| HISTORIA CARDIOVASCULAR | Presencia de crepitantes | categórico (3) |
| HISTORIA CARDIOVASCULAR | Ingurgitación de vena yugular | categórico (3) |
| HISTORIA CARDIOVASCULAR | Transplante cardíaco | categórico (2) |
| HISTORIA CARDIOVASCULAR | Fecha del transplante | ordinal secuencia (fecha) |
| HISTORIA CARDIOVASCULAR | Resincronización | categórico (2) |
| HISTORIA CARDIOVASCULAR | Fecha de la resincronización | ordinal secuencia (fecha) |
| HISTORIA CARDIOVASCULAR | Miocardiopatía | categórico (2) |
| HISTORIA CARDIOVASCULAR | Tipo de cardiomiopatía | categórico (3) |
| ANTITROMBOTICOS(ALTA) | Paciente | categórico (70) |
| ANTITROMBOTICOS(ALTA) | Se ha aplicado alguna guía clínica para la definición del tratamiento antitrombótico? | categórico (3) |
| ANTITROMBOTICOS(ALTA) | Se le ha prescrito anticoagulante? | categórico (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 | categórico (4) |
| ANTITROMBOTICOS(ALTA) | Otro tipo | categórico (2) |
| ANTITROMBOTICOS(ALTA) | Tipo de ACOD | categórico (5) |
| ANTITROMBOTICOS(ALTA) | Otro tipo de ACOD | categórico (2) |
| ANTITROMBOTICOS(ALTA) | Tipo de AVK | categórico (2) |
| ANTITROMBOTICOS(ALTA) | Otro tipo de AVK | categórico (1) |
| ANTITROMBOTICOS(ALTA) | Motivo | categórico (6) |
| ANTITROMBOTICOS(ALTA) | Otro motivo | categórico (4) |
| ANTITROMBOTICOS(ALTA) | Se le ha prescrito algún antiagregante? | categórico (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 | categórico (3) |
| ANTITROMBOTICOS(ALTA) | Otro antiagregante 1 | categórico (1) |
| ANTITROMBOTICOS(ALTA) | Plan terapéutico antiagregante 2 | Ordinal secuencial (tiempo en medicación) |
| ANTITROMBOTICOS(ALTA) | Número de meses | cuantitativo [0 - 0] |
| ANTITROMBOTICOS(ALTA) | Antiagregante 2 | categórico (3) |
| ANTITROMBOTICOS(ALTA) | Otro antiagregante 2 | categórico (1) |
| OTROS DIAGNOSTICOS | Paciente | Categórico (67) |
| OTROS DIAGNOSTICOS | Transtorno de la coagulación | Categórico (2) |
| OTROS DIAGNOSTICOS | Discrasia sanguínea | Categórico (2) |
| OTROS DIAGNOSTICOS | Diabetes | Categórico (2) |
| OTROS DIAGNOSTICOS | Dislipdemia | Categórico (2) |
| OTROS DIAGNOSTICOS | Enfermedad hepática | Categórico (2) |
| OTROS DIAGNOSTICOS | Cáncer | Categórico (2) |
| OTROS DIAGNOSTICOS | Anemia | Categórico (2) |
| OTROS DIAGNOSTICOS | Enfermedad renal | Ordinal secuencial (gravedad de la enfermedad renal) |
| FACTORES DE RIESGOS | Codígo | Categórico (69) |
| FACTORES DE RIESGOS | Antecedentes familiares de cardiopatía isquémica precoz (menor de 50 en varones y menor de 55 en mujeres) | Categórico (4) |
| FACTORES DE RIESGOS | Riesgo de caída | Categórico (3) |
| FACTORES DE RIESGOS | Caídas previas conocidas | Categórico (4) |
| FACTORES DE RIESGOS | Toma algún medicamento presente en esta lista | Categórico (15) |
| FACTORES DE RIESGOS | Déficit sensorial | Categórico (7) |
| FACTORES DE RIESGOS | Estado mental | Categórico (4) |
| FACTORES DE RIESGOS | Deambulación alterada | Ordinal secuencia (capacidad de andar) |
| FACTORES DE RIESGOS | Tabaquismo | Categórico (5) |
| FACTORES DE RIESGOS | Consumo diario de alcohol | Categórico (4) |
| FACTORES DE RIESGOS | Tipo de consumo | Categórico (4) |
| HABITOS | Codígo | Categórico (68) |
| HABITOS | Actualizado en este seguimiento | Categórico (3) |
| HABITOS | Realiza control regular de la presión? | Categórico (4) |
| HABITOS | Minutos semanales de actividad aeróbica | Cuantitativo [0-240] |
| HABITOS | Dieta Mediterránea | Categórico (4) |

Eventos.xls

|  |  |  |
| --- | --- | --- |
| Hoja de Excel | Campos (poner el nombre) | Tipo de atributo y cardinalidad |
| SANGRADO | Paciente | Categórico (116) |
| SANGRADO | Número anticoagulantes | Cuantitativo [0 - 1] |
| SANGRADO | Número antiagregantes | Cuantitativo [0 - 2] |
| SANGRADO | Otro medicamentos | Categórico (49) |
| SANGRADO | ANTICOAGULANT\_STRING | Categórico (4) |
| SANGRADO | ANTICOAGULANT\_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | ANTIPLATELET\_STRING | Categórico (6) |
| SANGRADO | ANTIPLATELET\_COUNT | Cuantitativo [0 - 2] |
| SANGRADO | ANALGESIC\_STRING | Categórico (2) |
| SANGRADO | ANALGESIC\_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | NON\_STEROIDAL\_INFLAMMATORY\_STRING | Categórico (2) |
| SANGRADO | NON\_STEROIDAL\_INFLAMMATORY\_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | DIURETIC\_STRING | Categórico (2) |
| SANGRADO | DIURETIC\_COUNT | Cuantitativo [0 - 1] |
| SANGRADO | OTHER\_STRING | Categórico (49) |
| SANGRADO | OTHER\_COUNT | Cuantitativo [0 - 17] |
| SANGRADO | Caracterización de la hemorragia | Categórico (3) |
| SANGRADO | Procedimientos terapéuticos | Categórico (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 | Categórico (6) |
| SANGRADO | ¿El paciente ha subido una transfusión? | Categórico (9) |
| TROMBOTICO | Paciente | Categórico (104) |
| TROMBOTICO | Número anticoagulantes | Cuantitativo [0 - 1] |
| TROMBOTICO | Número antiagregantes | Cuantitativo [0 - 2] |
| TROMBOTICO | Otro medicamentos | Categórico (21) |
| TROMBOTICO | ANTICOAGULANT\_STRING | Categórico (3) |
| TROMBOTICO | ANTICOAGULANT\_COUNT | Cuantitativo [0 - 1] |
| TROMBOTICO | ANTIPLATELET\_STRING | Categórico (6) |
| TROMBOTICO | ANTIPLATELET\_COUNT | Cuantitativo [0 - 2] |
| TROMBOTICO | ANALGESIC\_STRING | Categórico (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 invervención | Categórico (4) |
| TROMBOTICO | option3\_THROMBOTIC\_PRE | Categórico (2) |

# Abstracción de tareas

1. ¿Qué tipo de tares 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.

1. 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 propria 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. 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